EBU Technical Review

Special Supplement 1998

European Broadcasting Union
Ancienne Route 17A
CH-1218 Grand-Saconnex
Geneva
Switzerland

Tel: +41 22 717 21 11
Fax: +41 22 717 22 00
Telex: 415 700 ebu ch
E-mail: techreview@ebu.ch
Web: www.ebu.ch

EBU / SMPTE Task Force for Harmonized Standards for the Exchange of Programme Material as Bitstreams

Editeur Responsable:
P.A. Laven

Editor:
M.R. Meyer

Final Report: Analyses and Results

August 1998

European Broadcasting Union
Union Européenne de Radio-Télévision

Preface

The long-heralded convergence of television, computer and communications technologies is happening. The pace of change in the television industry, as a result of this convergence, is quickening dramatically as countries around the world make commitments to digital television broadcasting in one form or another. The result is the creation of many new, competitive, distribution channels that are driving a constantly-growing consumer demand for programming. Meeting that demand, in turn, requires a leveraging of the technologies used in digital signal processing, computers and data networking in order to yield significantly-enhanced creativity, improved efficiency, and economies of scale in the origination and dissemination of Content. The result of these changes is likely to lead to the wholesale replacement, or new construction, of practically all the television production and distribution facilities world-wide over the next decade or so. A major migration and a huge investment in technology is highly probable, and thus it is critical that the right decisions are made about the technological choices and the management of the transition to the new forms.

The opportunities presented by such a concurrent, world-wide, renovation of facilities are unique in the history of the industries involved. The coming changes will result in the industries literally remaking themselves, with consequent possibilities for new workflows, system designs and cost structures. Indeed, with the proliferation of distribution channels to audiences will come a fragmentation of those audiences which will mean that smaller budgets will be available for specific productions. The only ways to counteract this effect will be to find more uses for the Content produced, or to find more efficient ways in which to produce it.

With these necessities firmly in mind, the European Broadcasting Union (EBU) and the Society of Motion Picture and Television Engineers (SMPTE) formed a joint Task Force for the Harmonization of Standards for the Exchange of Programme Material as Bitstreams. The Task Force was charged with two assignments: (i) to produce a blueprint for the implementation of the new technologies, looking forward a decade or more, and (ii) to make a series of fundamental decisions that will lead to standards which will support the vision of future systems embodied in the blueprint. The first of these assignments was completed in the Task Force's *First Report — User Requirements*, published in April 1997.

The extraordinary document that you hold in your hands – or see on the screen in front of you – is the Task Force's response to the second assignment. It was produced by a group of over 200 experts from Europe, Japan, Australia and North America, meeting some seventeen times over a period of 1½ years. It is intended as a guide, developed co-operatively by the several industries involved, and will set the direction for all to follow. It represents the culmination of a unique effort by those industries – recognizing that they stand at a crossroads in their collective histories – to look into the future jointly and to choose their course together. It takes as its premise the need to identify requirements for the development of standards that will enable the exchange of programme material in the new forms and which will support the construction of complete systems based upon the new techniques. It includes the views of users and manufacturers of all types, both of which are needed in order to get an idea of what should and what will be implemented, and how it can be done.

At the start of this activity, some saw it as an opportunity to select a single video compression scheme to be used at a moderate bit-rate for a wide range of production and post-production applications. After thorough discussion, however, it was recognized that such a goal was not realistically achievable. Because of the many trade-offs that exist between compression methods, their parameters, and the performance achieved in specific situations, different techniques will be required in particular situations to meet explicit requirements. Thus the greatest benefit to all concerned will come from providing the mechanisms that will permit systems to handle easily the various compression schemes, while maintaining the maximum quality of the programme elements.

To this end, significant progress has been made in identifying the structures that will be necessary to support television production using compression, and in initially defining their characteristics. Among these are a new class of programme-related data called *Metadata* and, for the control and management of systems, the use of *object modelling* techniques. Metadata may be defined as the descriptive and supporting data that is connected to the programme or the programme elements. It is intended both to aid directly the use of programme Content and to support the retrieval of Content as needed during the post-production process. Object modelling techniques treat the devices and the Content items as "objects," the properties and parameters of which can be manipulated. Object models are intended to enable the easy integration of new devices into control networks, and the control of those devices from unified control and management systems.

Although the work of the Task Force is over, it is not done. Rather it has just begun. The purpose of the Task Force all along has been to point the way for successor activities to develop the standards, conduct tests, and co-ordinate the implementation strategies that will enable the realization of the future that is envisioned herein. Included as an annex to this document is an initial listing of major standards efforts that are required, all of which are anticipated to be undertaken by the SMPTE (and some of which have already begun). As with any roll-out of a major technological advance, the development of the standards for full implementation is expected to occur over a period of time, as more and more sophisticated applications of the technology evolve.

The subject matter of the Task Force's work is highly complex. This has led to more than a little confusion from time to time among the experts who have laboured to understand and make decisions about the future uses of the new technologies. In writing this document, every effort has been made to clarify difficult concepts, and to make them accessible to those who did not have the opportunity to participate in the discussions. There may be places in the text, however, where complete clarity has not been achieved. This is indicative of the complex nature of the subject and of the fact that this is still uncharted territory where the Task Force is clearly breaking new ground.

This report is divided into an *Executive Summary*, an *Introduction* and four sections which, respectively, cover *Systems*, *Compression issues*, *Wrappers and Metadata*, and *Networks and Transfer Protocols*. These sections are followed by a series of annexes. The sections contain the major findings of, and are the work of, six separate Sub-Groups that were assigned the tasks of investigating each of the subject areas. The annexes contain supplementary and tutorial information developed by the Sub-Groups, as well as information from the various sections brought together in one place. As they were written separately by different authors, the sections do not necessarily have the cohesiveness of style that might come from common authorship. Nevertheless, an attempt has been made to reconcile differences in terminology, so that individual terms have a single meaning throughout the document.

The work of the Task Force and the preparation of this report have provided a unique opportunity to put aside the short-term business of technology development and standards preparation and, instead, to take a longer-term view into the future with the hope of directing its path. As co-Chairmen, we are honoured to have been given the responsibility to lead the exercise. We wish to thank all those who were involved directly in the work, those who provided financial and travel support, as well as those who provided the meeting venues. We especially thank those who have served as Chairmen of Sub-Groups and, in particular, Roger Miles of the EBU who has served as Secretary throughout. There have been many long days spent by a large number of people to produce this output. We believe the result has been worth the labour.

Horst Schachlbauer, co-Chairman for:

The European Broadcasting Union (EBU)

Merrill Weiss, co-Chairman for:

The Society of Motion Picture and Television Engineers

Executive Summary

The convergence of the television, computer and communications industries is well under way, having been anticipated for quite some time. Video and audio compression methods, server technology and digital networking are all making a big impact on television production, post-production and distribution. Accompanying these technological changes are potential benefits in reduced cost, improved operating efficiencies and creativity, and increased marketability of material. Countering the potential benefits are threats of confusion, complexity, variable technical performance, and increased costs if not properly managed. The technological changes will dramatically alter the way in which television is produced and distributed in the future.

In this context, the Society of Motion Picture and Television Engineers (SMPTE) and the European Broadcasting Union (EBU) jointly formed the *Task Force for Harmonized Standards for the Exchange of Programme Material as Bitstreams*. The Task Force has had the benefit of participation by approximately 200 experts from around the world, meeting some 17 times in Europe and the United States over a period of a little less than two years. The Task Force has now produced two reports. The first one, published in April 1997, was called *User Requirements* for the systems and techniques that will implement the new technology. This second report provides *Analyses and Results* from the deliberations of the Task Force. Taken together, these two reports are meant to guide the converging industries in their decisions regarding specific implementations of the technologies, and to steer the future development of standards which are intended to maximize the benefits and minimize the detriments of implementing such systems.

The goals of the Task Force have been to look into the future a decade or more, to determine the requirements for systems in that time frame, and to identify the technologies which can be implemented in the next few years in order to meet these requirements over the time period. This approach recognizes that it takes many years for new technologies to propagate throughout the industries implicated in such sweeping changes. An example of this is the roll-out of straight-forward component digital video technology, which began with the adoption of the first standards in 1981 and has not yet, in 1998, been completed. Nevertheless, many of the techniques developed to support the implementation of component digital video now form the foundation of the move to compressed digital video, together with disk-based server and data networking methods which were developed first in other industry segments. Thus, because of the large and complex infrastructures involved, choices must be made of the methods that can be installed in the relatively near future, but which will still be viable over the time period contemplated by the Task Force's efforts.

To attain its objectives, the Task Force partitioned its work among six separate Sub-Groups, each of which was responsible for a portion of the investigation. These Sub-Groups were responsible for work on *Systems, Compression, Wrappers and File Formats, Metadata, File Transfer Protocols,* and *Physical Link and Transport Layers for Networks,* respectively. Some of the Sub-Groups found that their areas of interest were inextricably linked with one another and, consequently, they did their work jointly and produced a common report. Thus, there are four major sections to this report – with the efforts on Wrappers, File Formats and Metadata, and those on Networks and Transfer Protocols, having been combined into just two chapters. This combination of effort, that proved so useful for the work of the Task Force, will have ramifications for the related technologies and for the standards that will derive from this enterprise.

The Task Force has gone a long way towards identifying the technologies and standards that will be required to carry the converging industries with an interest in television to the next plane of co-operation and interoperation. The vision that results from that effort is expressed in this report. To turn that vision into reality will require even greater efforts by those who follow in the Task Force's footsteps. The Task Force has provided a guide to, or a map of, the directions to be taken. It will now be up to the industry as a whole, and the standards bodies in particular, to put into place a regime that will make the vision one that can be implemented practically. The Task Force members will be participating in those continuing efforts to turn this setting of direction into a pathway well travelled.

Systems summary

The work on Systems is new in this second report; it did not appear in the first report. It was only recognized through the work that went into the several specialized areas for the first report, that an overarching view was required which tied the various technologies together. It was also recognized from that initial work that the Systems which will be built, based on the new technologies, will be significantly more complex than in the past. Thus, it became important to consider implementation from a Systems perspective and to provide mechanisms for the management and control of all the facilities that will use the new techniques.

A high-level view of the overall scheme being considered by the Task Force is portrayed by a System Model which includes criss-crossing Activities, Planes and Layers – all of which interconnnect with one another. The model is intended to bring a visual representation of the relationships between the many complex workflows and technologies that comprise television systems based on the new methods. It also recognizes that the various subsystems covered in the remaining portions of this report must be integrated into a workable total system, if advantage is to be taken of the potential offered by the new methods.

The Systems section also considers two categories of implementation issues: the operations that will be required in systems to integrate the various technologies, and the choices that must be made among options at the several layer levels in order to construct optimized systems. Operations include: control; monitoring, diagnostics and fault tolerance; Data Essence and Metadata management; Content multiplexing; multiplexing of Metadata into containers; and timing, synchronization and spatial alignment. Choices among the interconnection options can be optimized through the use of Templates applied to different combinations of defined transfer types that will be used for specific activities and applications. In addition to study of the implementation issues themselves, consideration is given to the requirements for migration from current systems to those contemplated by the Task Force.

Among the most significant findings of the Task Force, with respect to the control and management aspects of operations, is the solution offered by distributed object-modelling techniques. Object modelling, in general, offers a means to abstract both control functions and representations of Content in a way that allows large systems to be built through the sharing of object definitions, and without the need to provide individual software drivers at every controlling device, for every device to be controlled. This has important consequences for the ability to establish and expand complex resource management methods, both quickly and economically.

Migration of systems to object-modelling techniques will require that suppliers of current equipment make public their control protocols. This will allow translation mechanisms to be developed between new object-based systems and the currently-installed base of equipment. As the migration progresses, new equipment is expected to work directly in the object domain, and standards will evolve from the functions identified in current equipment to meet the requirements of new system elements as they are developed. This will require unprecedented co-operation between equipment developers but, as a result of the Task Force's efforts, it is now generally recognized that little added value comes from unique solutions for control and management, while significantly greater costs can result. Thus, it is in everyone's interest that standards-based solutions take hold.

Compression summary

Compression is the process of reducing the number of bits required to represent information, by removing any redundancy that exists in the bitstream. In the case of information Content such as Video and Audio, it is usually necessary to extend this process by removing information that is not redundant but is considered less important. Audio and video compression schemes are therefore not normally lossless. Consequently, reconstruction from the compressed bitstream leads to some distortions or "artefacts."

The decision to use compression has a significant impact on the overall cost / performance balance within television production and post-production operations, as it affects the quality, storage / transmission efficiency, latency, editing / switching as well as the error resiliency.

Compression of Video and Audio allows functionality with a bandwidth / storage efficiency that is not viable with uncompressed processing. Through a reduction of the number of bits required to represent given programme Content, it makes economical the support of applications such as the storage of material, transmission, faster-than-real-time data transfer, and simultaneous access to the same Content by a number of users for editing and other purposes.

Choices made with regard to compression techniques and parameters have significant impact on the performance that can be achieved in specific applications. It is most important that those choices be made with a clear understanding of the requirements of the associated application. In particular, it is important to make decisions about compression in the studio that take into account the production processes and compression generations across the total production chain. The decisions taken there will be different from those that would be made if the compression were optimized only for presentation to a human observer.

The first part of *Section 3* provides users with general information about Audio and Video compression characteristics to assist in making judgements about appropriate solutions. It further contains recommendations on approaches to be taken to facilitate interoperation to the greatest extent possible between systems within a single family of compression techniques and between families of compression methods.

After the issuing of the First Report, members of the European Broadcasting Union and major manufacturers of broadcast equipment have held in-depth discussions on present and future compression schemes.

Annex C of this report reveals that future technology for networked television production must maintain a close focus on the compression types and on the balances obtained in terms of:

- ➩ ultimate technical programme quality versus data-rate;
- ➩ interoperability of compression schemes using different encoding parameters;
- ➩ editing granularity versus complexity of networked editing control.

Based on an analysis of the market situation and with reference to the list of user requirements established in the first part of the report, the Task Force promotes two different compression families as candidates for future networked television production, to be used for core applications in production and post-production for Standard Definition Television (SDTV):

- ➩ DV / DV-based 25 Mbit/s with a sampling structure of 4:1:1 and DV-based 50 Mbit/s with a sampling structure of 4:2:2, using fixed bit-rates and intra-frame coding techniques exclusively. DV-based 25 Mbit/s with a sampling structure of 4:2:0 should be confined to special applications.
- ➩ MPEG-2 4:2:2P@ML using both intra-frame encoding (I) and GoP structures, and data-rates up to 50 Mbit/s [1,2]. MPEG-2 MP@ML with a sampling structure of 4:2:0 should be confined to special applications.

Each compression family offers individual trade-offs in terms of coding flexibility, product implementation and system complexity, as well as adequate headroom to allow migration from SDTV into HDTV operations.

Standardization of encoding parameters, the mapping of compressed data into various transport streams as well as the interfaces required within different areas of applications, is in progress.

Chip-sets for both compression families will be available on an equitable and non-discriminatory basis.

The coexistence and interoperation of the above compression families within a networked television facility will pose a number of operational problems and will therefore be the exception and not the rule.

Manufacturers are committed to produce silicon-based *agile decoders* which will enable the coexistence and interoperation of members within a single compression family.

The Task Force believes that digital audio in production and post-production will remain uncompressed, although it cannot be totally excluded that external contributions may require the occasional handling of audio in compressed form.

Wrappers and Metadata summary

Starting from the First Report on User Requirements, the Sub-Group started to search for a single comprehensive solution. Examining the work which was already underway elsewhere within the computer industry and within the SMPTE, it was apparent that the Metadata requirements could be addressed through the creation of a *Metadata Dictionary* and a number of formatting standards, all maintained through a registry mechanism.

1. For recording on a VTR, a fixed bit-rate must be agreed for each family mamber.
2. For specific applications, this also includes MPEG-2 MP@ML if decodable with a single agile decoder.

With regard to the Wrappers requirements, the Sub-Group issued a *Request for Technology* (RFT). Several responses were received, covering aspects of the required technology, from established companies in both the computer and the television industries. The responses ranged from discussions on specific items such as Unique Material Identifiers and frame index tables for use inside Wrappers, to complete solutions for specific applications such as multimedia presentation delivery, and also included the specifications for data models and container formats in use today in the industry, within multiple products. These responses were analyzed during repeated meetings, along with comparisons of existing practices in the industry and discussions on the standards development efforts which have been continuing simultaneously.

No single response to the RFT covered all the requirements. In general, however, the sum of the responses on stream formats covered most of the stream requirements, and similarly the sum of those on rich Wrapper formats covered most of the complex Content package requirements.

Not surprisingly, the various proprietary technologies submitted were not immediately fully interoperable to the degree requested in the First Report. However, in their use of established practices – such as the use of globally unique identifiers – some of the responses were more amenable to limited modification than others in order to achieve interoperation.

The final phase of the Sub-Group's work was to issue a second RFT in search of one missing item from the first response – a low-level special-purpose storage mechanism.

During the concluding meetings of the Sub-Group, it became clear that the technology to employ in comprehensively addressing this requirement does exist. However, it was not possible to complete the documentation of this technology within the scope of the Sub-Group. Instead, this should be taken up by the SMPTE, following the plan given in *Section 4.9*.

Networks and Transfer Protocols summary

The Sub-Group on Networks and Transfer Protocols has investigated interfaces, networks and the relevant transfer protocols for the transmission of Content. In the course of these investigations, the Sub-Group defined a Reference Architecture for both Content file and streaming transfers, to meet the demand for interoperability. This Reference Architecture includes interfaces and networks as well as file transfer protocols, protocols for real-time streaming, and methods for file system access. Existing standards are recommended where available, and areas requiring further development and standardization are identified.

The interfaces, networks and transport mechanism recommended include:

- Serial Data Transport Interface (SDTI);
- Fibre Channel according to NCITS T11 FC-AV;
- ATM.

While it was recognized that any of these technologies could be used for streaming within a broadcast studio, recommendations were made for the best operational use of these technologies:

- SDTI was identified as the solution for streaming at the current time;
- Fibre Channel was chosen for file transfers because of its fast transfer capabilities;
- ATM is particularly suited to the wide-area network and allows both streaming and file transfer.

Where necessary, recommendations for new work to be carried out by standardization organizations, and improvements of the transport mechanism and protocols, have been made. Mapping rules are needed to enable the transport mechanisms to be used for the different applications (e.g. transmission of DV or MPEG) and to meet the operational requirement for real-time streaming of Content. A number of these mappings have been established (e.g. DV25 / 50 into SDTI), while others (e.g. DV25 / 50 into ATM) remain to be defined and standardized.

For the file transfer of Content, FTP was chosen as the universal transfer protocol. User requirements for fast and point-to-multipoint file transfers have encouraged the development of FTP+ as an enhanced version of FTP, and the adoption of the eXpress Transfer Protocol (XTP). The Sub-Group has carried out some fundamental work on the definition of such an enhanced file transfer protocol. Standardization has already started in the SMPTE and in the EBU.

Section 1

Introduction

The television industry currently faces both the tremendous challenge and the tremendous opportunity of remaking itself over the next decade or so. The expected change is of historic proportions. It will be driven by (i) the proliferation of new delivery channels to consumers, (ii) the new capability of those channels to carry data of many types in addition to, or in place of, video and audio and (iii) the need to fill those channels with Content. At the same time that more Content is needed, the cost to produce it will have to decline significantly because, on average, fewer consumers will watch or use each programme. Balancing this trend, new business possibilities will open for those who can leverage the ability to transmit new forms of information through channels that formerly carried only television. The net result of all this will be that Content will have to be produced far more efficiently or will have to serve multiple uses, or a combination of these, if the quality of the end product is not to suffer substantially.

The transformation will be aided by its confluence with the dramatic changes occurring in computer and networking technologies, leading to faster processing, larger memories, greater storage capacity and wider bandwidths – all at lower capital and operating costs. These improved capabilities in hardware platforms will permit systems to be based on radically new concepts, aimed at achieving the required improvements in efficiency and utilization. The pivotal new concepts include:

- ⇨ programme data transport in the form of compressed bitstreams;
- ⇨ non-real-time data transfer;
- ⇨ simultaneous, multi-user access to random programme segments stored on servers;
- ⇨ inter-networking on open platforms of all production tools within the post-processing chain;
- ⇨ hierarchical storage concepts based on tape, disk and solid-state media;
- ⇨ the treatment of data on an opportunistic basis.

In this context, end-to-end interoperability as well as optimized technical quality must be considered as prerequisites to successful system implementations. Interoperability comprises not only the capacity to exchange Video, Audio and Data Content between equipment of different manufacturers, but also the capability to fit within common control systems and resource management schemes. Exchange of Content between systems which are based upon different compression mechanisms can be achieved currently only by using decoding and re-encoding. Such concatenation of compression methods can be made less burdensome to systems through the use of “agile decoders”, which are enabled by the combination of advancing technology, automatic identification of signal types and parameters, and publication of the techniques used in the various compression schemes. Similarly, methods that preserve the decisions and parameters used in the original compression process, and then use them again when the content is re-encoded, can have beneficial results with respect to the quality of the resulting product.

Choices relating to the compression system characteristics – even within single families of compression devices – can have an enormous impact on the quality achieved. Applications may place constraints on exercising some of the potential choices; for example, there will be a requirement for limited delay (latency) between points which support live conversations among participants in a programme. The constraints thus placed on system choices can have significant implications elsewhere, such as in the amount of storage required or the bandwidth necessary for particular programme elements.

Future systems are likely to be noted for their complexity relative to current systems. At the same time, it will be desirable for operations to be conducted by personnel less technically skilled than those currently involved in the handling of Content. This can lead to efficiencies through the closer involvement of those skilled in management of the Content itself, rather than in management of the technology. Such a change in skill sets, however, will require that control systems manage the technology transparently from the point of view of operations personnel, making decisions automatically that hitherto would have required technically-skilled personnel. This has manifest implications for the control and resource management systems of the future.

One of the major impacts of the adoption of techniques from the domain of computer technology will be the use of "layered" systems, in which functions necessary to implement a connection between devices are considered to exist in a "stack". Such layering permits the use of different techniques at each layer to suit the particular application, without requiring the replacement of all elements of the interconnection functionality. The concept of layering appears throughout this report, as indeed it will appear throughout future systems. The pervasive nature of layering will be reflected in the need for future standards also to be based upon this concept.

The sections of this report can be thought of as representing layers in a stack. They more or less go from top to bottom of the stack, starting with Systems, continuing with Compression, followed by Wrappers and Metadata, and finishing with Networks and Transfer Protocols. Each of these topics represents a layer in this report and it is likely that they will appear as layers in future systems and in future standards. They can be considered analogous to several of the layers of the classical seven-layer Open System Interconnect (OSI) model used to describe networked systems. It is expected that future systems, and the standards that define them, will use this particular combination of layers as their foundation.

Considering each layer in turn, the Systems layer deals with all of the functionality necessary to integrate a multiplicity of devices and techniques. It provides the means to interrelate the operation of the many elements that can comprise a major function or activity. Systems are extensible in the sense that they can treat smaller groupings of elements as subsystems of larger systems, and can cascade such relationships to ever larger systems. Among the concerns at the Systems layer are such matters as the control and management of the system and its functionality, the necessary interrelationships between the methods chosen at each of the lower layers, and the data structures necessary to sustain the applications to be supported on specific systems. A principal outcome of the work on Systems control is the need to use object-oriented techniques to provide an efficient, extensible control architecture. Object-oriented technology provides an intuitive way of modelling complex real-world problems, by splitting their solutions into manageable blocks which themselves can be subdivided into further blocks. The result is a powerful notion of objects containing their own data and the operations that can be performed on that data.

Metadata is a major new class of enablers of systems which use bitstreams for programme material exchange. Metadata is a generic term for all sorts of captured data that relates in one way or another to programme material. It ranges from timecode and details of technical conditions when material was created, to the scripts used, the publicity materials created and descriptions of the shooting locations. It can also include standardized descriptive data to help in locating the material through various database entries. This can aid in the re-use of material, thereby significantly increasing its value.

Wrappers and file formats are inextricably linked with Metadata in that they contain programme Content and its associated Metadata in ways that it can most easily be transferred and most beneficially used. This means that the Metadata may need to be accessible from the outside of containers, so that the Content of the containers can be properly identified and processed. The need to both contain and provide access to the Metadata requires that the form of the Metadata and of the containers be considered together.

Networks and Transfer Protocols provide the ability to exchange Content (Audio, Video and associated Data and Metadata) easily and reliably between different devices and systems. Agreed methods to move Content within production chains are essential, providing a stable basis for user choice, and encouraging a variety of solutions without putting limits on innovation in product design. For the first time, networks will encompass not only exchanges inside a production complex but will also deal with the complexities and characteristics of digital public carrier distribution systems and of networking through them. This will enable production processes that are physically distributed between sites to be integrated, from the perspective of the operator, as though they were at a single location.

Despite all the new technology that will be applied to the television system, there are certain factors that differentiate television from almost any other activity and which must be acknowledged and included in the design of many future systems. In particular, the requirements of live television place constraints on the speed and bandwidth necessary for transfers, the delay or latency that is acceptable between the place of acquisition and the point of programme integration, and the error rates that are essential in networks since the retransmission of damaged packets may not be possible.

In this report, the Task Force seeks to provide users, system designers and manufacturers alike with a document that addresses these issues in a complete and informative manner. Its purpose is to begin the process of standards development to support implementation of the techniques that are described. The work must be completed by standards development organizations such as the SMPTE – the intended target for many of the tasks that arise from this report – which have the permanence to see it through the transition and to recognize all of the additional work that will be necessary but that has not been identified herein because of the myopia that comes from working in the present.

Readers are advised that, while this document is the result of the efforts of over 200 experts from four continents, meeting over a 1½-year-plus period, it is not a work of literary art. Most of the time has been spent exploring the requirements of users and the technological solutions that will address them. The writing of the findings of that effort has been a parallel effort of a number of teams working independently, in a co-ordinated way. Consequently, different parts of this document may read as though they were written by different authors, which they were. Nevertheless, an effort has been made to consolidate the report so that the same terminology is used throughout and that reasonably consistent conclusions are drawn by the different sections. To the extent that this is not the case, it results from the intense efforts effectively to write a book in a very short period. Any confusion that remains is likely to be reflective of the complexity of the subject at hand.

An open issue is how the work required to fulfil the vision contained herein will be guided and monitored. Clearly the SMPTE intends to take on the work fully. It has already begun to change the organizational structure of its standards development activities to be reflective of the Task Force's output. The EBU and the SMPTE must consult in the future to determine whether any follow-on activities are necessary. A corollary matter is whether any updates to this report should be prepared as the transition to the new modalities progresses. Both these questions require the benefit of a future viewpoint for appropriate responses to be obtained.

Section 2

Systems

Following the Introduction below, this section examines a model that has been developed within the Task Force to provide a platform for consideration of the many issues that will bear on future digital television systems. It is also intended that this section will serve as a vehicle for communicating the concepts derived from the Task Force's efforts to other groups that will continue the work.

One subsection considers the various kinds of operations that must be included in an overall implementation of bitstream-based television programme exchange. Another investigates some of the many issues that must be deliberated before a complete complement of standards can be brought forward. Still another subsection provides a guide to the preferred implementation combinations at the various system layers for specific applications. Finally, a listing is provided of systems-level standards that must be developed as a result of the Task Force's efforts, some of which have already been passed along to the SMPTE in advance of publication of this Final Report.

2.1. Introduction

The brave new world of television, based on the exchange of programme material as bitstreams, brings with it many new and changed considerations when the requirements for systems are examined. Future systems will not only provide new operational functions and features, they will also perform even traditional operations in a new and fundamentally different manner. These systems will be quite different in the elements and techniques that comprise them; they will lead to new workflows in the facilities in which they are installed, and they will lead to new approaches to system design, implementation and support. In this section, the many aspects of the systems that will support bitstream-based programme exchanges are examined from a systems-level perspective.

Systems, by their nature, integrate a multiplicity of devices and techniques. They provide the means to interrelate the operation of the many elements that can comprise the entirety of a function or operation. Systems are also extensible in the sense that what is seen as a system in one view of an operation can be viewed as components (or subsystems) of even larger systems, when seen from a higher level of the operation, i.e. systems may act as subsystems in the larger context. This section thus treats as subsystems the system elements which are described in the other parts of this report. Additionally, it looks at two aspects that appear uniquely in a systems perspective of the future world of television, namely, the integration of the other aspects of systems, and the control, monitoring and management of the overall facilities of which they all become part.

2.1.1. *Systems of the future will do different things*

Future television systems will be called upon to accomplish many functions that are quite different from what they have had to do in the past. For example, many facilities which in the past would have delivered a single service at their outputs will be expected to deliver a potentially large number of services. There will be requirements for "re-purposing" of material so that it can be used across many distribution media and to provide many differing versions on a single medium. Similarly, there will be a need to access Content for multiple simultaneous uses. This will permit, for instance, the efficient creation and release of several versions of the Content. It will be necessary to handle new forms of programme-Content creation and delivery. These forms will range from the scheduled delivery of data in its most basic form to the streaming of video and audio elements that are to be rendered into programme Content in the receiver or monitor. New types of businesses are likely to spring up which will require support from the systems of the future. However, the nature of those future businesses may well be totally unknown at the moment in time when the technology for those systems is being created.

2.1.2. Systems of the future will be quite different

The future systems that are contemplated will be quite different from those built in the past. They will be largely based upon the use of computing techniques and data networking. Conceptually, they will be built upon layered structures that will permit the use of interchangeable parts at each of the layers. They will make widespread use of servers, which will enable the use of non-real-time transfers that, in turn, will allow optimization of the trade-off between the speed of delivery and the bandwidth utilized. The use of data networking and non-real-time transfers will lead to the need to manage end-to-end Quality of Service (QoS) and to make bandwidth reservation services available to avoid unpredicted conflicts. The inclusion of Metadata in the Content will require schemes for its capture, storage, modification and transfer, and for its re-association with the Essence (see *Section 2.5.3.*) to which it is related.

Many aspects of future systems will depend upon the registration of various types of data so that mechanisms for identification of Essence and the control of equipment and systems can be continually extended. This will allow the growth of systems, the addition and modification of equipment elements for improved functionality, and the relatively easy dissemination of new standard parameters for objects of all kinds. These extensions will enable the implementation of increasingly complex systems. The system complexity will naturally result from the intermixture of different equipment families and from the resulting requirement for translation between signal types. Since it can be expected that the choices between equipment families will only broaden with time, it can also be anticipated that increasingly complex systems will result.

2.1.3. New workflows will result from integrated control schemes

Currently there can be enormous differences in the ways that different companies carry out the same operations and the flows of work through their operations. Similarly, there are often vast differences in the way that the same functions are carried out in different countries and regions of the world. This places a strong demand on systems to provide adaptability so that they can work in the many environments in which they may be situated. At the same time, there will be opportunities for changes in workflow that will lead to more efficient operations and / or better output products. Among the potential benefits that can accrue from the use of computer technology and data networking will be integrated control schemes that permit access to a wide range of equipment and processes from a single user terminal and interface. This can be enhanced through inter-facility networking that permits resources at distant locations to be integrated into systems as though they were local, or that enables a single operator at a central location to monitor and control a number of distant facilities.

2.1.4. New types of system design are required

The changes in systems that will come along with the new technologies for programme exchange will bring with them requirements for wholly new types of system designs. The great complexity with which some systems will have to be implemented will cry out for the use of automation in places where it has not been considered before. Automation subsystems will permit the simplification of systems operations from the point of view of the user, making the complexity involved in performing those operations disappear transparently into the background while nevertheless providing the benefits of the complex systems. This, in turn, will lead to the control of operating costs through a requirement for less technically-skilled operators and / or the ability to engage more highly-skilled operators to perform the Content-related functions. Automation will also permit the application of resource management techniques that can lead to maximized efficiency of facility utilization.

Besides the benefits of automation, there will be potential savings in maintenance staffing costs. These savings may be realized through the outsourcing of maintenance work, for which there will be a much larger universe of possible suppliers because of the commonality of equipment with that of the computer industry. Additionally, this same commonality will provide for easier recruitment of “on-staff” maintenance personnel. When coupled with fault tolerant systems, it may be possible to have maintenance personnel on-call rather than on-site. Furthermore, capital costs can be controlled through the use of standardized interfaces, protocols, and command and response structures that are shared with other industries rather than being for the exclusive use of the television industry. Operating costs can be controlled through the increase in throughput that will result from the parallel processing of Content in order to develop multiple versions and to support multiple uses all at the same time. Countering some of these potential savings may be the need to have very highly competent systems administrators and operations support ("help desk") staff who are available around the clock, either on the staff payroll or through outsourcing. Annual software maintenance costs are an additional recurring expense that must be considered.

New methods of system design can also be applied to the systems of the future. For example, in addition to the traditional functions of laying out physical equipment and deciding how it should be interconnected, system design of the future may well entail the selection of appropriate software drivers to be loaded from CD-ROM or DVD-ROM, designing the user interfaces for the displays, and establishing the signal processing rules for loading into the system's resource manager. Part of this task will be the need to establish a configuration documentation and control system for both the hardware and the software that will have to be maintained over time.

2.2. Request for Technology

During its work, the Task Force discovered that the issues pertaining to the harmonized standards for the exchange of programme material as bitstreams were much more complex than the traditional processes in place in the industry. In order to deal with these complex problems, a level of "System Management Services" is required, and a Request for Technology was issued in order to find possible solutions. The RFT document described these problems and is an important document for the understanding of the challenges it poses. Therefore, it is attached to this report in *Annex B*.

Three responses were received in response to this RFT and, from a partial analysis of these responses, the Task Force determined that there is a need to develop a Reference Object Model for System Management in order to insure interoperability in the longer term. These responses are the base material that is being used by a newly-formed group within the SMPTE to work on that reference model.

Since this process will likely be an incremental process that will require some time, intermediate steps have been identified. *Fig. 2.1* shows that there will be multiple interacting control systems, with each system containing or communicating with various devices.

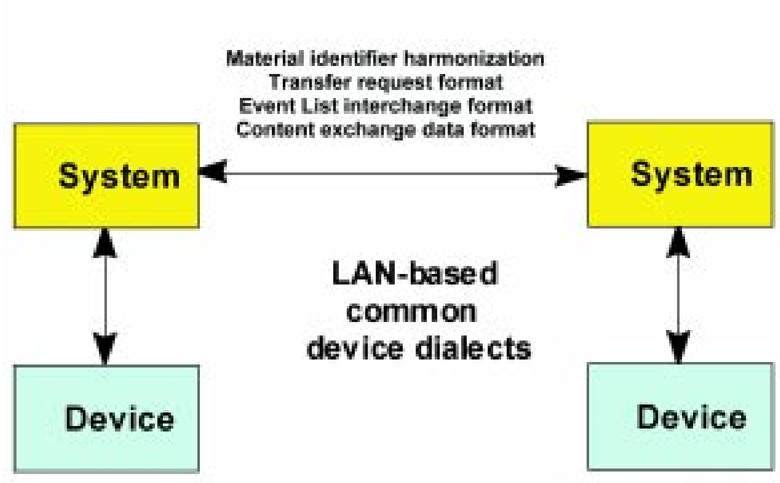


Figure 2.1: Management System relationships.

With respect to *Fig. 2.1*, these intermediate steps are:

⇨ Development of standard, LAN-based, common device dialects for system-to-device communication:

- A standard LAN-based control interface for broadcast devices is being developed to allow these to be connected to control systems that use an IP-based transport.

⇨ Harmonization of a Material Identifier:

- For system-to-system or system-to-device communication, a common unambiguous means of uniquely identifying the Content should be developed. Systems and devices could continue to use their own proprietary notation internally but, for external communication, this should be translated to the standard form.

- See *Section 4* (Wrappers and Metadata) for a discussion on the format specification of Unique Material Identifiers (UMIDs) to be used for unfinished programme material.
- For the smaller subset of completed programmes, a more compact, human-readable, still globally-unique identifier will probably be used for the entire programme Content. Many different specifications for these Unique Programme Identifiers (UPIDs) already exist or are being developed.
- Management systems must be able to deal with the full range of Content identifiers, including both UMIDs and the multiplicity of UPIDs that will exist. They must also accommodate the case in which portions of completed programmes are used as source material for other programmes. To enable management systems to recognize this variety of identifiers, each type of identifier should be preceded by a registered SMPTE Universal Label.

There is also a requirement for Common Interfaces for system-to-system communication:

⇨ Transfer Request format:

- A standard format for messages requesting the transfer of material from one location to another should be developed. This format must allow for varying file system and transfer capabilities in the underlying systems, and should allow different qualities of service to be requested. The possibility of automatic file translation from one format to another as a transparent part of the transfer should also be considered.

⇨ Event List Interchange format:

- A standardized Event List Interchange format should be developed. It should allow vendors to develop a standardized interface between business scheduling (traffic) systems and broadcast / news automation systems.

⇨ Content Database Interchange format:

- To facilitate inter-system communication about media assets, a reference data model for the Content description should be developed. Systems can internally use their own data model, but this must be translated for external communication.

2.3. Object model

Studio operation of today is already more complex and sophisticated than of even a few years ago, and is becoming increasingly more complex as new technologies, new studio automation equipment and new applications (services) are introduced. As a new generation of digital television broadcasting equipment is considered, with its increasing use of embedded computers in studio equipment, new levels of management and control become possible.

The Task Force recognizes that the object-oriented approach is the best choice for a coherent, orderly and extensible architecture for studio control, which not only accommodates today's studio but also is capable of handling future evolution.

2.3.1. *Why object-oriented technology?*

Object-oriented technology provides an intuitive way of modelling complex real-world problems by splitting their solutions into manageable blocks, which themselves can be subdivided into further blocks. This approach mirrors good engineering practice in all engineering disciplines. For example, a system block schematic can be expanded into device block schematics, which can themselves be expanded into engineering drawings / circuit diagrams.

Traditional software design has used the procedural approach, which separates the storage of information from the processes performed on that information. When a process defined by one part of the software is changed, the remaining software must be rewritten in many places.

In the object-based approach, if one object is changed, usually little or no other code needs to be changed, because the changed object's interfaces remain the same. In addition, the object approach lends itself to re-use of existing software by allowing a new object to inherit all the capabilities of the previous object, and enhance these. This will happen in the studio as new devices with enhanced or additional functions are developed and brought on line, using object-oriented technology. Such upgrades can be made without replacing other studio software that interacts with these devices.

Object-oriented technology enables entrants to see rapid economic benefits and a faster turn around of new developments, implementations and benefits from a broader range of supply. Object-oriented technology facilitates the interfacing of production systems with business systems. For example, a programme-scheduling object will be able to communicate with a rights-management database, derive the cost of re-running a programme, and indicate any rights issues.

Using object-oriented technology, existing technologies such as computer networking mechanisms can be leveraged to provide communication between objects.

The EBU / SMPTE Task Force recommends the development of an object-oriented reference model for use in the development of future media Content creation, production and distribution systems. A listing of technologies which are candidates for study in the development of the object model is given in *Section 2.10.1*.

In addition, the Task Force recommends that, for this process, UML (Unified Modelling Language) be used for drawing, and IDL (Interface Definition Language) be used for the definition of APIs.

The Task Force believes that a "Central Registry" for Object Classes will be required, and that a "Central Repository" is desirable.

2.4. System model

In order to understand better the requirements of system design, the Task Force has developed a model based on orthogonal parameters and intersected by an underlying control and monitoring layer (see *Fig. 2.2*). This model is used to explore the relationships between *Signals*, *Processes* and *Control Systems*. It will be applied extensively throughout this section of the report.

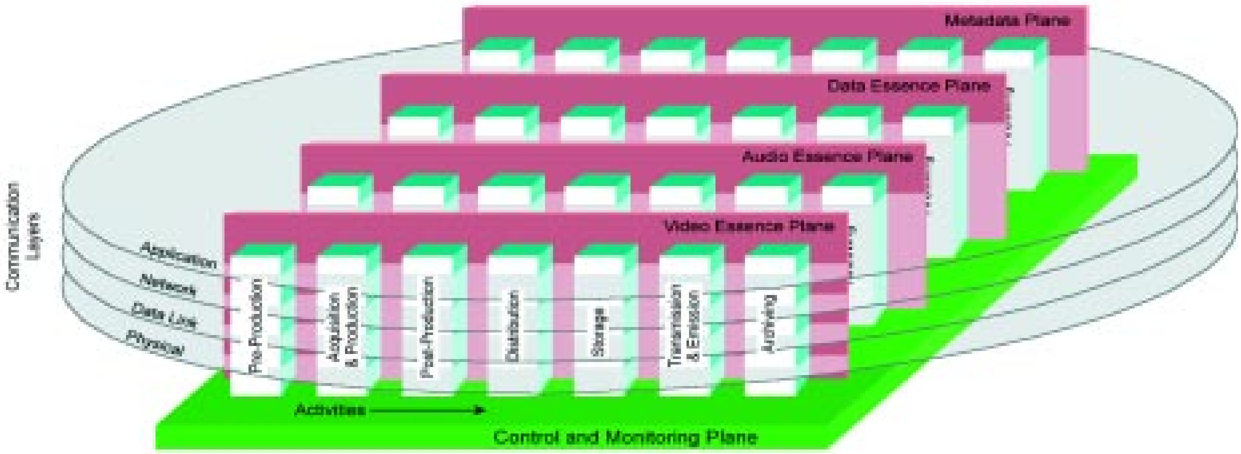


Figure 2.2: System model.

2.4.1. Structure of the model

The three orthogonal axes of the model are:

- ➩ **Activities** – this axis describes the major areas of activity within television production and distribution processes, from acquisition through to delivery and archiving.
- ➩ **Planes** – this axis describes the different types of data encountered throughout the television production chain. Although many different variants of digital information are regularly encountered, the Task Force

has identified the base types of *Video Essence*, *Audio Essence*, *Data Essence* and *Metadata*. All types of programme Content can be placed into one of these information types.

⇨ **Layers** – this axis describes the operating layers which cut through the Activities and Planes. Four layers are defined which have considerable similarities with the ISO / OSI 7-layer model.

These intersecting properties are shown in *Fig. 2.2* with Activities on the horizontal axis, Planes on the depth axis and Layers on the vertical axis.

Underlying the Activities and Planes axes, is a Control and Monitoring plane. It spans the whole model because of the strategic interest of Control and Monitoring functions in the complete television operation and in all digital Content types.

The Task Force model can be used to describe or analyze any type of programme or activity. The description of part of any system can be made in terms of the model by describing the technologies used to carry each of the Planes for any given Layer. It can also describe the Control and Monitoring functions across the Activities and Planes.

A television system can be considered as a number of signal-carrying planes, controlled by an intersecting control plane. Each production task requires the manipulation of signals in all or some of the planes.

In traditional television systems, the planes have been distinct physical systems, i.e. video, audio and data were carried on different cables; Metadata was often simply written on accompanying documentation. Future systems will not necessarily have these distinct physical systems; rather, they are likely to be based on networks or multiplexed signals. It is useful, however, to consider the system in terms of a logical model in which the signal types are distinct. These logical systems are like views of the real physical system.

2.4.2. Activities

Different activities within the television production and distribution process require varying degrees of control over the signal planes. This leads to much commonality in the signals used, with most of the task-specific aspects of the system being embodied in the control plane. As described below, seven generic types of activity have been identified that describe all stages of the television production, storage and dissemination processes.

2.4.2.1. Pre-production

The pre-production process involves the original requests for material to be acquired, together with any rights appropriations. Advance scheduling is also classed as pre-production. The computer systems will generally be standard information-processing systems. Dynamic linkage is required between pre-production and subsequent activities. Access to archive and other library material *will* be required, at least at a browse quality level, together with the ability to mark and retrieve selected material at high quality for subsequent process stages.

2.4.2.2. Acquisition & Production

The acquisition and production process places many demands on the control system; since live production is involved, manual switching occurs and must be executed with little or no delay. The production process uses a wide variety of equipment, usually manually controlled in real-time, with varying degrees of automated assistance. The material being acquired may be unrepeatable and therefore absolute control-system reliability is essential. The control system may also be required to manage the simultaneous capture and / or generation of related Metadata during the acquisition process.

2.4.2.3. Post-production

Post-production activity may involve complex manipulation of signals, requiring the simultaneous, deterministic control of equipment. Modification, iteration, and repetition of sequences is common, thus requiring reliable repetition of sequences. In order to process material, multiple passes may be required which can lead to generation loss. The use of more processing equipment may allow several processes to take place in a single pass, reducing the number of passes required: the control system should support this.

2.4.2.4. Distribution

Distribution requires absolute reliability of control. Distribution may require control of conventional routing switchers, as well as local and wide-area networks.

2.4.2.5. Storage

Storage also requires absolute reliability of control. Storage control requires the use of interfaces to asset-management databases and hierarchical storage-management systems. The storage devices controlled range from disk servers to robotics datatape libraries, as well as systems which use conventional digital videotape.

2.4.2.6. Transmission & Emission

Transmission and emission control systems must manage the integration of Video Essence, Audio Essence, Data Essence and Metadata for transmission. Control of transmission-processing devices, such as multiplexers and conditional-access systems, must be supported. The introduction of complex Data Essence and Metadata into the production, transmission and emission systems may require new and innovative approaches to statistical multiplexing (see *Section 2.5.4*). Transmission environments often have low staffing levels, so automatic monitoring and backup systems must be included.

2.4.2.7. Archiving

Archiving control systems must ensure the long-term integrity of the archived material. It requires effective searching of archived assets in order to allow efficient retrieval. This is often achieved by placing the Data Essence and Metadata associated with the material in a database that is separate from the Audio and Video Essence. The control system must manage the relationship between the user, the Data Essence, the Metadata and the archived Audio / Video Essence.

2.4.3. Planes

The planes described here may be physical or they may only be logical. An example of a logical plane is the Audio Plane when the audio is embedded with the video data. A controlled signal falls into one or more of a number of categories described below. The fact that signals may be multiplexed and routed together in the physical system is embodied in the model as a series of constraints; for example, one constraint of using multiplexed audio and video signals might be that they cannot be separately routed without demultiplexing. These constraints must be consistent with the requirements of the activity.

2.4.3.1. Video Essence

This category includes all Video Essence, whether compressed or not. The video transport may be a network, in which case there is no separate physical transport for the Video Essence. There will be, however, a logical video structure which describes the possible paths for Video Essence and the maximum bandwidth availability between any two points.

It is a point of discussion as to whether Video Essence includes both Video Streams and Video Files. Video Streams are clearly Video Essence. However, Video Files may be low-resolution browse pictures and a convention needs to be developed to classify such video information as either Video Essence or Data Essence. For this report, it has been assumed that Video Files are classified as Data Essence. Thus Video Essence is classified as Video Streams only.

2.4.3.2. Audio Essence

This category covers Audio Essence, including audio description, whether compressed or not. As with the Video Essence, there may be no separate physical audio transport: the Audio Essence may often be carried as a multiplexed signal with the Video Essence but, for the purposes of modelling the control of the systems, it should be considered as a separate logical system.

In common with the above discussion on Video Essence, for the purpose of this report, Audio Files have been classified as Data Essence, thus Audio Essence is classified as Audio Streams only.

2.4.3.3. Data Essence

Data Essence is information other than Video Essence or Audio Essence, which has inherent stand-alone value (unlike Metadata which is contextual and has no meaning outside of its relationship to Essence).

Examples of Data Essence include Closed Captioning text, HTML, programme guides, scripts, .WAV files, still images, Web page associations, video clips (as a file), etc. Again, this may be multiplexed with the other signals but should be considered separately for control purposes.

2.4.3.4. Metadata

Metadata is information other than Essence, that has no inherent stand-alone value but is related to Essence (i.e. it is contextual and has no meaning outside its relationship to the associated Essence). Examples of Metadata include: URL, URI, timecode, MPEG-2 PCR, filename, programme labels, copyright information, version control, watermarking, conditional-access keys, etc. Metadata will either travel through the system, multiplexed or embedded with the Essence from one or more of the other planes, or it will be stored in a known location for subsequent reference.

2.4.4. Layers

Every application domain in the reference model is described by four planes: Video Essence, Audio Essence, Data Essence and Metadata as described in *Section 2.4.1*. For each of these planes (defined by the scope of an application domain), communication between peer entities is accomplished over a four-layered model that is consistent with the International Organization for Standardization (ISO) Open Systems Interconnection (OSI) model. This model describes how communication between peer entities is greatly reduced by layering the communication between entities. The layers discussed here are the *Application layer*, the *Network layer*, the *Data Link layer* and the *Physical layer*, which agree in scope with the equivalently-named layers in the ISO / OSI model.

Unlike many application areas, the diverse nature of the broadcast studio demands that a multiplicity of mechanisms be employed at each layer for differing application domains. For example, within the post-production environment, components may stream the video over a Fibre Channel or ATM infrastructure (the network layer) whereas, within the production environment, SDI may be used. Similarly, control of a video server within the post-production environment may be exercised over an Ethernet infrastructure (the Data Link layer), but over RS-422 within the production environment. This characteristic is derived from the extremely distinct requirements that are placed on applications in the broadcast community, including hard real-time considerations.

2.4.4.1. Application layer

The Application layer defines the specific application entities that are used in the system. The specialized nature of the broadcast community implies that Application layers have a video-centric flavour to them. In particular, many Application-layer entities within the context of this Final Report would be considered as presentation-layer activities in the ISO / OSI model. For example, Application-layer entities in the broadcast community would include MPEG-2 ES, DIF, MPEG-2 TS and FTP process capabilities.

2.4.4.2. Network layer

The Network layer defines the communication protocols between a set of co-operating protocol entities (i.e. processors) that permit the transport of Application-layer services between components in different locations that are not directly attached to one another. Network-layer protocols in the context of the broadcast community include: IP, XTP, ATM, SDTI, Fibre Channel, AES-3 audio and a variety of open and proprietary standards. As with the Application layer, the specialized nature of the broadcast industry requires that consideration be given in the Network layer to protocols that would be considered within the Data Link layer in the ISO / OSI model, e.g. SDI.

2.4.4.3. Data Link layer

The Data Link layer defines the communication protocols between a set of co-operating protocol entities (i.e. processors) which permit the transport of Network layer services between components in different locations that are directly attached to one another. The Data Link layer is responsible for the framing of bits, and for error correction / detection between directly-connected components. Protocols for the Data Link layer, in the context of the broadcast community, include: Fibre Channel, ATM, SDI, Ethernet, RS-422 and a variety of open and proprietary standards.

2.4.4.4. Physical layer

The Physical layer defines the electrical and mechanical characteristics that permit the transport of Data Link layer services between components in different locations that are directly attached to one another. Physical-layer *specifications* in the context of the broadcast community include:

- 75 ohm coaxial cable, terminated using BNC connectors at component video-signalling levels;
- twisted-pair category 5 links, using RJ-45 connectors at Ethernet signalling levels;
- twisted-pair category 3 links, using DB-9 connectors at RS-422 signalling levels.

The Physical layer itself, in the context of the broadcast community, includes the above examples as well as twisted-pair, transmitting at XLR signalling levels, and a variety of open and proprietary standards.

2.4.5. Control and Monitoring plane

The functions of the Control and Monitoring plane are to co-ordinate the transfer, storage, manipulation, monitoring, diagnostics and fault management of signals through the other planes. This plane provides overall management of Content across all activities, planes and layers.

The Control layer can be seen as the co-ordinator of transactions in the other planes. It will allocate and co-ordinate the resources of the system to provide the services for each transaction. Human operators form a critical part of the Control plane in almost all television systems where issues of providing a consistent and reliable Man Machine Interface (MMI) are crucial.

Monitoring, diagnostics and fault management are essential to the smooth running of television facilities, in particular where continuity of output is of the highest importance. Human operators still form a critical part of the monitoring operation in almost all television systems where issues of providing a consistent and reliable MMI are crucial. It is desirable to automate the monitoring functions, wherever possible, in order to assist the human operators of the system.

2.5. Operations

Operations includes the strategic and tactical control of a facility’s work, including scheduling and resource management. In current TV facilities, much of this work is done by people with the assistance of some level of automation or database management. In a fully-networked digital facility, it is anticipated that most of this administrative load will be handled by the system, with human intervention required only for high-level direction and prioritization.

2.5.1. Control

Control encompasses the strategic and tactical control of a facility's work, including scheduling and resource management. Control is exercised at many levels, from highly-abstracted advanced planning to real-time machine control. It is convenient to separate these into strategic and tactical control.

2.5.1.1. Strategic

Strategic control concerns the overall task; it does not concern the details of how the task will be achieved. In a broadcasting facility, the strategic functions include advance programme planning, programme acquisition and marketing planning.

2.5.1.2. Tactical

Tactical control concerns the execution of the strategic plan; it involves the allocation of the physical resources needed to realize the plan, and the management of the required media.

2.5.1.3. Peer-to-peer

Separate control systems may be involved in the production of different programme types. These are complete vertical systems, covering an area of production e.g. News, Commercials. It is frequently required to transfer data from one control system to another where the two systems act as peers. This type of data exchange requires standardization of the transfer format, and is usually concerned with database linkage.

2.5.1.4. Resource management

The allocation of resources to tasks may either be accomplished by having sufficient resources to meet the worst case in each production area, or by dynamically allocating the resources to tasks. In practice, some mixture of the two approaches is likely with low-cost or critical devices being owned by a particular operation and high-value or seldom-used devices being shared.

Technologies to be implemented in the near-term future will pose significant control and system management-related issues. These technologies include: compressed video and audio streams, file servers with file transfer capability, shared networks, associated Data Essence and Metadata, management of output digital multiplexers, delivery of network digital programming to affiliates / service providers, etc.

Automatic backup and / or archiving of critical programme material may be achieved using duplicated video servers at the output point, with suitable networking or other facilities to ensure synchronization of the media stored within them.

There are various signal interconnection schemes that can be used to move Content (both Essence and Metadata) from location to location within (or between) facilities. These can be described in terms of the interfaces and protocols that connect equipment. In the future, these interfaces and protocols may be based on computer networking concepts that can be used in various combinations. This requires that a structure be devised, both for describing and communicating what is involved in a particular transfer. The structure used is a layered model that provides the necessary context, with both file transfer and streaming exchanges supported.

2.5.1.5. Physical control network issues

In general, control will be provided over the Physical and Data Link layers as described above. Control can be implemented using Physical-layer technologies ranging from RS-422 to network technologies such as Ethernet, Fibre Channel and ATM. Consideration must be given to system-design requirements such as latency. It is recognized that several control-related issues are introduced when considering the needs of the broadcast community:

- Distributed Workgroups require scalability, possibly using “virtual LANs” which can be added and dropped without upsetting the overall system. This will permit system architectures to grow without adversely interfering with real-time system performance.
- QoS and bandwidth need to be controllable by the requirements of resources and applications. For example, video streaming may require particular bandwidth to operate effectively and this must be satisfied by the control interface.
- The command latency (the time between command and action) needs to be deterministic for many hard real-time applications. Again, this must be satisfied by the control interface.
- There is a need to provide prioritization of bandwidth, i.e. support for an “Emergency Stop”. Again, the command latency (the time between command and action) needs to be deterministic.
- The control systems may provide redundancy, i.e. there should be “no single point of failure”.
- Reliability may be designed in with the management and monitoring system by using, for example, SNMP2 (Simple Network Management Protocol version 2) with redundant distributed management servers and nodes to provide a hierarchical network management system.
- Redundancy and reliability of the control system may be provided by utilizing redundant distributed servers and nodes.

2.5.1.6. Multiple forms of implementation

The Control plane may be realized as a distributed system, where the control functions of each block in the signal planes are present in each signal-processing, storage or routing element, or as a separate central control system. In practice, most systems will have a combination of centralized control, to co-ordinate activities across planes, and localized control to act on particular devices.

There are two fundamentally different structures used to produce control systems. These can be characterized as *hierarchical* and *peer-to-peer.*

2.5.1.6.1. Hierarchical

In hierarchical systems, the control system is arranged as a series of master-slave relationships. The transaction is usually arranged at the master system, which allocates the resources and passes the commands down to slave systems which control the hardware devices. This type of system has the advantages of simplicity and consistency, it is deterministic, but it can be difficult to expand or adapt. Backup arrangements can usually only be provided by duplicating a large part of the equipment complement.

2.5.1.6.2. Peer-to-peer

In peer-to-peer systems, devices provide services to transaction managers. A transaction manager will hunt for the services it requires and use those to accomplish its goal. The advantages of this type of system are in adding or extending its capabilities to meet unforeseen requirements. The difficulties of this structure are mainly related to ensuring that sufficient resources will be available to meet the requirements at all times.

2.5.1.7. Essential characteristics

The control of large integrated systems presents a new set of challenges. The following is a list of some of the characteristics that should be applied to any new control system under consideration:

- **Extensibility** – new devices will need to be added and removed without upsetting the remainder of the networked devices. Likewise, new features will need to be smoothly integrated. “The necessity to reboot the system” is not allowed.
- **Scalability** – the ability to add more existing devices, without major re-configuration, must be supported.
- The integration of new devices and services, with minimum disruption to existing services, is required.
- The system should support the dynamic allocation of resources.
- The system should provide a consistent interface for common services.

- ➩ The finding and retrieving of assets and resources, with minimum human interaction, is required
- ➩ The system should provide fault tolerance and failure recovery to ensure service continuity (see *Section B.3.18.* in *Annex B*).
- ➩ If distributed resources are used by multiple users, security mechanisms are required (see *Section B.3.11* in *Annex B*).
- ➩ If a distributed system is used, then suitable resource allocation systems must be provided (see *Sections B.3.11* and *B.3.12* in *Annex B*).
- ➩ The system should allow the use and exchange of devices from different manufacturers in a common system (see *Section B.3.12* in *Annex B*).
- ➩ The widest possible manufacturer and user support for any new control system method is required.

2.5.1.8. Logical control layers

The control of devices is generally on one of three layers. Each layer abstracts the one before it, thus providing easier integration of a device into a control system.

2.5.1.8.1. Transaction-based control protocols

These are low-level protocols used to communicate with a device or service using datagram messages. Examples of these protocols are the familiar RS-422 protocols used to control devices such as VTRs. The definition of this type of protocol is usually provided in the form of a document that is used by a software engineer in creating a driver program for each control system.

2.5.1.8.2. Functional APIs

Application Programming Interfaces (APIs) define the control of a device in terms of a set of procedure (subroutine) calls. Each procedure call will be translated into the datagram messages of the transaction-based control protocols. The API may be hosted on the device itself, if it uses a standard operating system, or as a communications driver running on a computer that hosts the application, if the device itself does not.

An API is more abstract than direct control of a device, because it can encapsulate a whole sequence of messages into a single call. Also, if there are multiple variants of the device in use, each of which requires a different command sequence to perform the same operation, this variance can be hidden from the calling program which need only know how to make a single call. This is achieved by installing onto the hosting computer, the version of the API that is appropriate for the model in question (i.e. automatic model-determination by a common shared version of the API). This technique is also a form of "encapsulation of complexity".

2.5.1.8.3. Transition to object-oriented studio control by means of proxies

As illustrated in *Fig. 2.3*, an Object Wrapper encapsulates the API for a device or service, and acts as a proxy for the device. Functions of the device can be controlled or monitored by invoking methods of the object.

Since each layer builds on the one below, it is possible for a device manufacturer to document and supply all three interfaces. This allows new devices to be integrated into traditional control systems, by using transaction-based protocol over a serial link, or into new object-based systems by using a supplied interface object. It is recommended that manufacturers supply both transaction-based protocol documents and higher-level control structures to allow integration of new devices with traditional control systems.

The diagram also shows how the object-oriented approach can accommodate or evolve from the existing transaction-based approach. An Object Wrapper or proxy can be written for the existing device so that it can interact with other objects without the need to re-implement the APIs.

The diagram shows an object representation of a studio which consists of objects representing various devices such as VTRs, video servers and library servers, as well as objects performing system functions that are necessary to support distributed object systems such as an Object Registry and Resource Management.

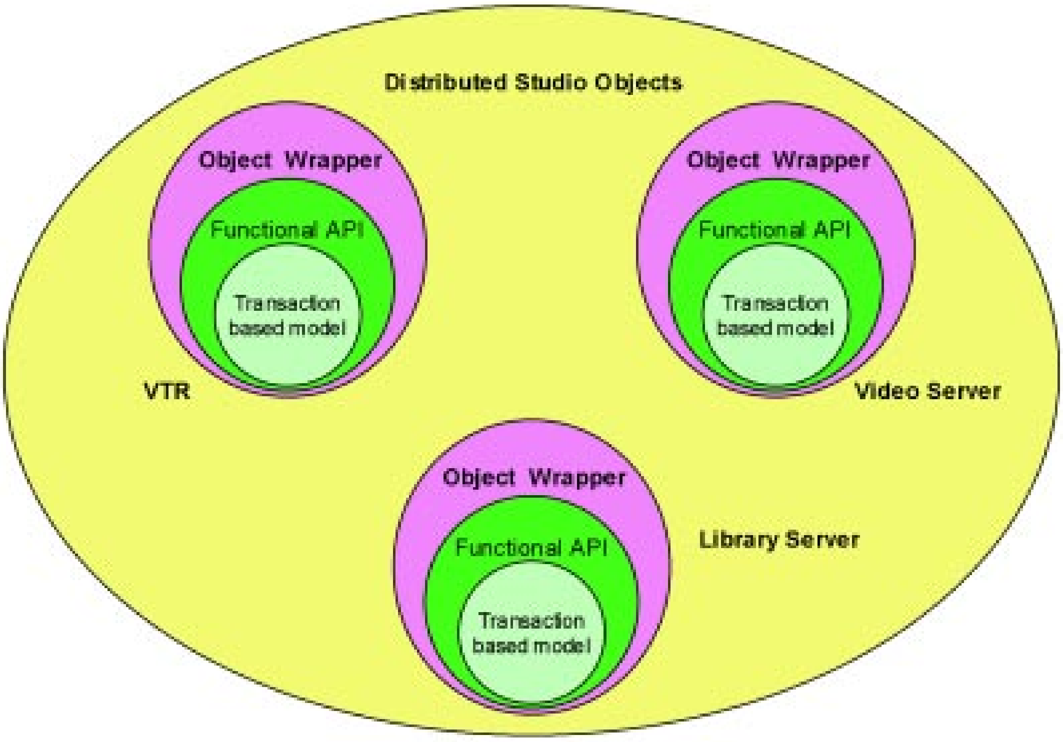


Figure 2.3: Distributed studio objects.

2.5.2. Monitoring, diagnostics & fault tolerance

Monitoring, diagnostics and fault tolerance are essential to the smooth running of television facilities, where continuity of output is of the highest importance. In many cases, automating these functions will make them much more effective.

2.5.2.1. Feedback

Broadcast control systems have traditionally employed closed-loop signalling where any indication is derived from the controlled device, not from the control message. This practice becomes more important as the complexity of the control system increases, thus providing the operators or automation systems with assurance that an action has been carried out.

2.5.2.2. Failure prediction

Degradation of signals can be measured in some digital systems without interrupting the signal. This information can be used by the control system to predict failure and to re-route or otherwise re-allocate the equipment, and also to alert maintenance staff.

2.5.2.3. On-line / Off-line diagnostics

Diagnostic capabilities fall into two classes: on-line, where the function can be performed while the device is in use, i.e. bit-error rate figures; and off-line, where the device must be taken out of service for evaluation. Where on-line evaluation is practical, it is important to use it so that faults may be detected at the earliest opportunity. This is key to the implementation of fault tolerance.

2.5.2.4. Fault tolerance

It must be accepted that no system, however well designed, will operate forever without faults. It is possible to design systems which can tolerate faults, providing that certain accommodations have been made in their component equipment to detect these and work around them. In traditional television facilities, fault tolerance was achieved through the use of redundancy equipment and by human detection of faults followed by instigation of corrective action. By contrast, data-processing systems have long been designed with automated fault detection and correction; thus the migration of digital technology into the television environment provides the opportunity to take advantage of these techniques. Such faults are not restricted to bit errors; of far greater importance is ensuring that the intended bitstream arrives at the correct destination at the correct time.

2.5.2.4.1. Redundancy

The commonest technique for implementing fault tolerance is redundancy; indeed, the two terms are often erroneously assumed to be synonymous. Clearly, if a device fails, making good its loss requires having a functional equivalent available, carrying the same programme stream. However, redundancy in itself is insufficient. It is necessary to have monitoring systems to detect that a failure has occurred, and control systems to initiate the switch-over. Provisions for these must be made in the control architecture.

There are limits to which redundancy can be taken. Network theory teaches us that in a system with a single output, it is impossible to avoid having a single point of failure. The best we can do is to move it as far downstream as is practical. The single point usually ends up being a switch, which in many cases is designed as a bi-stable mechanical device, so that loss of power does not cause it to fail completely.

2.5.2.4.2. Segmentation

Another useful technique is segmentation. It is essential in data network design, where multiple devices must communicate with each other. Through the use of data routing and switching, segments of the network can be isolated from each other and redundant paths can be created. Thus the loss of a link or the flooding of a link with spurious data from a failed node can be isolated and contained. Hardware and software to create and manage such networks are readily available, and it is essential that they be used.

2.5.2.4.3. Monitoring and verification

Redundancy systems cannot be implemented without monitoring. As noted above, in traditional TV facilities, the monitoring was actually done by humans with the assistance of various interpretative devices, including picture monitors, waveform monitors, meters and loudspeakers. More complex equipment, such as VTRs, were capable of monitoring some aspects of their own operation, but could communicate only by displaying that information to an operator who, it was hoped, would be present to note it and take appropriate action.

There are two shortcomings to this type of monitoring. The first is that it relies on the presence and attentiveness of the operator. The second, which is more serious, is that the operator does not become aware of the problem until it is affecting the product, at which point damage avoidance is impossible and damage control is the only remedy.

Automated systems can be constructed, utilizing the error-detecting features possible with digital interfaces, to take corrective action before impairments become apparent to the viewer. They can monitor multiple streams simultaneously and can react far faster than a human can. The SMPTE has standardized a number of tools that greatly facilitate the design of such systems, including SMPTE 269M, 273M and RP-165, and many existing computer industry standards such as SNMP are applicable as well.

To verify that the intended material is arriving at the correct location at the correct time, it is necessary that a *Unique Material Identifier* (see *Section 4.6.*) is transmitted in the bitstream along with the video and audio, and that this particular type of Metadata is monitored throughout the system. This verification has until now been done entirely by humans, as it involves evaluation of Content. By automating it, verification can take place economically at many more places in the system.

2.5.2.4.4. *Implementation*

Fault tolerance can be implemented within a device or within a group of devices. Either approach is valid, as long as the desired results can be attained. All facilities will not require the same amount of fault tolerance; even within a single facility, the level of redundancy required will vary with the time of day and the value of the programming passing through. The reference model must be flexible enough to allow for a range of implementations, without foreclosing options at either end of the spectrum. It must be emphasized that the degree of fault tolerance desired is a trade-off between capital cost on the one hand and downside risk on the other, and that these trade-offs will be different for each system.

2.5.2.5. Content integrity

In the system design, it is necessary that provision be made to ensure not only the unique identification of programme material, but also that it has not been altered. A process must exist to certify the author, the integrity of the file or stream, and its version.

2.5.3. *Data Essence and Metadata management*

This section deals only with the issues of transportation and storage of Data Essence and Metadata (which are defined in *Sections 4.3.4.* and *4.3.5.*). *Section 5.* covers the topics of Video, Audio and Data Essence transfers.

It is important that the formats of Data Essence and Metadata employed be consistent throughout the production, distribution and emission chain. It is not desirable to track identical Data Essence and Metadata with a number of different identification schemes. In the case of Metadata, a mechanism has been established through the Metadata Registration Authority. All suppliers and users are recommended to use this facility. It is further recommended that all Essence formats (Data, Video streams and Audio streams) be registered through the same authority.

Users should be able to share data electronically between databases. These databases may contain either Data Essence (e.g., subtitle text) or Metadata information. Users should not be required to manually re-enter this data at any stage beyond first entry. Indeed in many instances, Metadata and some types of Data Essence ought to be automatically created. There is also a requirement for database connectivity between broadcast and business systems for the exchange of data such as scheduling information and operational results. This connectivity should be handled in a standardized way as far as possible.

2.5.3.1. Layered model

Section 5. describes transfers in a hierarchical or “layered” manner according to the ISO / OSI 7-layer model. This model enables the transport of Data Essence and Metadata structures in a consistent manner with minimal loss of information.

A layered model can also enable different classes and priorities of Data Essence and Metadata, determined by how closely they are coupled with the Video or Audio Essence.

Layered structures can also be used effectively in the formatting of Data Essence and Metadata to allow easier transfers between different systems.

It is recommended that appropriate layering and hierarchical structures should be considered for all formatting, transfers and storage systems wherever practical.

2.5.3.2. File transfer

This is the transfer of Data Essence and Metadata from a source to one or more destinations, with guaranteed delivery achieved by means of the retransmission of corrupted data packets. The transmission rate may not have a fixed value and the transmission, in fact, may be discontinuous. An example is the transfer of a data file between disk servers.

Transfer of time-critical Data Essence and Metadata by file transfer may be inappropriate unless the files have embedded timing information. See *Section 5.* for further information.

2.5.3.3. Streaming

Conventional TV environments involve relatively small delays which can easily be calibrated out by proper plant engineering or by adjustments provided by the manufacturer. Packet-switched networks will involve considerably more delay, sometimes well beyond a single frame period. Plant timing adjustment, timing verification and monitoring will be needed to ensure all Essence and Metadata can be delivered both in time and in sync with other data (lip sync). For systems using queries, round-trip timing may be important even in a streaming environment. Hybrid equipment environments may present especially difficult problems during the transition period.

In the past, the electronic transport of Data Essence and Metadata has been limited to a restricted set such as timecode, subtitle text, etc. It is expected that future studios will depend more heavily on the more extensive streaming of Data Essence and Metadata as described in *Section 4.7.5*. This transport can be accomplished by one of four broad schemes:

- ⇨ No automated connection between Data Essence, Metadata (e.g. stored on discs) and A/V streams (e.g. stored on tapes).
- ⇨ Minimum identifier (e.g. Universal Material ID) embedded with or in the Essence that permits access to associated Essence and Metadata over a separate transport.
- ⇨ Minimum identifier together with partial Metadata embedded with or in Essence that permits access to associated Essence and the remaining Metadata over a separate transport.
- ⇨ Full Metadata embedded with or in all Essence types including Data Essence.

2.5.3.4. Data Essence

Data Essence is information other than Video Essence or Audio Essence, which has inherent stand-alone value (unlike Metadata, which is contextual, and has no meaning outside of its relationship to Essence).

Examples of Data Essence include subtitle text (subtitles carried by Teletext and Closed Captioning), scripts, HTML, Web page associations, still images, video clips (as a file), .WAV files, etc. Again this may be multiplexed with the other signals but should be considered separately for control purposes.

2.5.3.4.1. Repetition rate of Data Essence

During a streaming transfer, there may be certain types of Data Essence that require periodic updating. This allows the user to asynchronously enter the stream and within a certain time period, recover the Data Essence as determined by the repetition rate. For example Teletext is repeated periodically as are programme guides.

On the other hand, it is not necessary to repeat Data Essence within a file during a file transfer, since it is not practical to enter the stream asynchronously and to be able to decode the file successfully.

If a stream consists of a sequence of related files, there may be certain Data Essence types that warrant periodic repetition (e.g. every n^{th} file).

2.5.3.4.2. Registration of Data Essence types

It is recommended that Data Essence types be registered with the SMPTE Registration Authority (see *Section 4.5.3.*).

2.5.3.5. Metadata

The transfer of Metadata within the studio is influenced by a number of studio-specific characteristics. These include limitations on the ability of existing studio infrastructure to effectively store and transfer Metadata at the required repetition rate for essential Metadata delivery. For analogue systems this will require an additional digital infrastructure to store, transfer and track the Metadata.

2.5.3.5.1. Layered structure

There is a need to layer the Metadata into a structure which supports a multiplicity of priorities that indicate how closely the Metadata is to be bound to the Essence.

High priority implies that the Metadata is very closely coupled with Essence, while low priority indicates that the Metadata is less closely coupled. Examples of very closely-coupled Metadata in an SDI infrastructure are: sync, eav, sav, timecode (field by field). Examples in an MPEG-2 TS are: PCR, PTS, DTS.

It is clear that Metadata above some priority threshold should be stored with the primary Essence, while Metadata below this threshold should be stored separately. It is expected that the location of Metadata will depend on this priority. For instance, tape-based systems may be required to store all high-priority Metadata with the Video (either in data space or in unused Video lines), while lower-priority Metadata is located in a database elsewhere. Conversely, server-based systems may be capable of storing all priorities of Metadata "together" although these capabilities are dependent on the manufacturer. The SMPTE should recommend a mechanism for representing this priority structure and include appropriate fields into Metadata structures to reflect this.

It is also recognized that there will be a number of predefined types of essential Metadata. Predefined Metadata includes sync, eav, sav and timecode. These will be established and maintained by the SMPTE Registration Authority (see *Section 4.5.3.*).

2.5.3.5.2. Minimum repetition rate of Metadata types

Considerations for the repetition rate of various Metadata types are similar to those discussed above for the repetition of Data Essence. For example, in MPEG-2 video syntax, the repetition of the picture header can provide Metadata for each picture, whereas the sequence header need only occur once per sequence. The repetition rate will be application-specific.

2.5.3.5.3. Registration of Metadata types

It is recommended that Metadata types be registered with the SMPTE Registration Authority (see *Section 4.5.3.*).

2.5.3.6. Permanence of Data Essence and Metadata

It might normally be assumed that all Data Essence and Metadata should be preserved as far as possible throughout the broadcast chain. However, some Data Essence and Metadata types may intentionally be destroyed after they have served their useful purpose. Examples of Data Essence and Metadata with "short" permanence include:

- Machine control information that must be destroyed after use in order to ensure that it does not inadvertently affect additional, downstream equipment.
- Transfer bandwidth information which is necessary only to execute a specific transfer.
- Error-detection and handling flags related to the active video picture information, that must be replaced by new flags in equipment that modifies the picture Content.
- "Helper" signals that are used to carry encoding-decision information between concatenated compression processes.
- Interim Closed Captioning text or script information that must be altered or rewritten to match properly the final edited audio information.

On the other hand, some Metadata types such as UMIDs (see *Section 4.6.*) and timecode must have "long" permanence. System implementations will need to be sensitive to the varying degrees of "permanence" required for Data Essence and Metadata.

2.5.3.7. Data Essence / Metadata capabilities

The capability of storage systems, transport systems and container formats to support Data Essence and Metadata is limited.

A server-centric view of the studio, as opposed to a tape-centric view, also permits great flexibility in terms of access and cataloguing of stored Content. Direct-access storage devices permit Metadata storage and indexing of all Content within the studio. This allows sophisticated video Content and context-based queries to be followed by immediate data retrieval. Server technology also permits the same Content to be distributed efficiently to multiple recipients simultaneously. Finally, a server-centric view of the studio will map conveniently to a networked transport infrastructure.

For archival storage, both tape and film-based systems with supporting Metadata will continue to be used. It would be desirable if archival storage systems supported a transport infrastructure together with the capacity and formatting to allow recording of Data Essence and Metadata.

2.5.3.7.1. Capability in storage systems

The storage capacity and transfer rates of any storage system may limit the capacity for Data Essence and Metadata. Because of this limitation, applications may need to store data in a hierarchical fashion – with higher-priority Data Essence types coexisting with A/V Essence on the same storage medium, but other lower-priority Data Essence types being stored elsewhere with an appropriate pointer. Tape-based video storage devices will typically be limited in their capacity to store high-priority Metadata. This, coupled with a possible repetition-rate requirement for Metadata, implies that it is necessary to evaluate each tape-based storage system to see if the product of repetition rate and high priority Metadata can be supported.

There are three popular approaches to the storage of video: videotape-based, datatape-based and disk-server-based. From a historical perspective, videotape storage has been the medium of choice. However, disk performance and storage improvements, along with developments in the delivery of compressed video and audio, enables the realization of a server-based model. Servers directly support the manipulation and flexible delivery of compressed bitstreams over multiple interfaces. Additionally, compressed bitstreams can be transferred between servers much faster than real-time over transport infrastructures such as Fibre Channel, ATM and SDTI.

2.5.3.7.2. Capability in transport systems

As above, the transfer rates of any transport system may limit the capacity for data storage. Applications may need to transfer the data types in a hierarchical fashion, with higher-priority data types coexisting with A/V Essence on the same transport medium. Other lower-priority Data Essence types may either not be transported at all, or may be transported by some other means. The application will need to ensure that the pointer to the remaining Data Essence is properly managed / updated.

2.5.3.7.3. Videotape storage

Table 2.1 provides examples of the capabilities of various tape-based storage formats. The preferred location for high-priority Metadata is “Additional Data Space”. Although the SDI input to a VTR or server may contain HANC and VANC data, users should be aware that the recording media may not preserve this data. In most instances, the “audio” AES-3 HANC packets are preserved within the data depth of the recording device (e.g. a 16-bit VTR may truncate a 20-bit packet). The AES-3 stream could contain either audio or data. Compression-based VTRs may also distort some of the data present in the vertical blanking interval. Analogue tape formats are omitted from this table.

2.5.3.7.4. Datatape storage

There are two varieties of data storage on tape-based systems: video recorders that have been modified for general data storage, e.g. DD1, DD2, DD3 and DTF; and data storage tapes that have been improved to support video streaming requirements, e.g. DLT and Magstar. These tape formats have a wide variety of storage capacity and transfer rates.

Table 2.1: Video storage capabilities (tape-based).

Recorder Type	Video Type	Recorded Lines	Audio Type	Audio Tracks	Additional Data Space	Comments
Digital Betacam	Digital 10-bit (compress)	505/525 596/625	Digital 20-bit	4 Track	2,880 Byte/frame	
D1	Digital 8-bit component	500/525 600/625	Digital 20-bit	4 Track	None	
D2	Digital 8-bit composite	512/525 608/625	Digital 20-bit	4 Track	None	
D3	Digital 8-bit composite	505/525 596/625	Digital 20-bit	4 Track	None	
D5	Digital 10-bit component	505/525 596/625	Digital 20-bit	4 Track	None	
DV25 (DVCAM, DVCPRO)	Digital 8 bit compressed	480/525 576/625	Digital 16-bit	2 Track	738 kbit/s	
DVCPRO50 (DVCPRO50, Digital-S)	Digital 8-bit compressed	480/525 576/625	Digital 16-bit	4 Track	1.47 Mbit/s	
Betacam SX	Digital 8-bit compressed	512/525 608/625	Digital 16-bit	4 Track	2,880 Byte/frame	
HDCAM	Digital 10-bit compressed	1440x1080i	Digital 20-bit	4? Track	None	
HDD-1000	Digital 8-bit uncompressed	1920x1035i	Digital 20-bit	8 Track	None	
HDD-2700 HDD-2000	Digital 10-bit compressed	1280x720p 1920x1080i	Digital 20-bit	4 Track	2,880 Byte/frame	

Data storage manufacturers will typically provide a data port. Data storage implementers will augment this with particular capabilities for the storage of video, audio and Metadata. It is recommended that implementers who are using data-storage devices for video-storage applications should include specifications providing the following parameters:

- Storage capacity;
- Sustained data transfer rate;
- Video Essence types supported;
- Audio Essence types supported;
- Data Essence types supported;
- SNR or Corrected BER;
- Interfaces supported;
- Metadata and Data Essence recording limitations.

2.5.3.7.5. *Disk storage*

Disk-based video storage is generally recognized as being more flexible than tape, particularly when considering the transfer of Data Essence and Metadata. Disks permit flexible associations of Data Essence and Metadata with the video and audio. A video server has the characteristic that it views Content from both a network / file system perspective and a video / audio stream perspective and it provides connectivity to both domains. Constraints are derived from internal server performance and from the transport / interface technology that is used when conveying the data to / from the video storage system. Due to the currently-used interfaces, there are limitations on the way that manufacturers have implemented video servers but many of these limitations will be removed as interfaces develop. The manufacturers of disk-based video storage will typically support a wide variety of formats, interfaces and Metadata capabilities. It is recommended that the manufacturer's video storage specifications include the following parameters:

- Storage capacity;
- Sustained data transfer rate;

- Video Essence types supported;
- Audio Essence types supported;
- Data Essence types supported;
- Corrected BER;
- Interfaces supported;
- Metadata and Data Essence recording limitations.

Within the studio there are likely to be several different classes of server, corresponding to the principal work activities within the studio – in particular, play-out servers, network servers and production servers. These will differ in storage capacity, streaming capability and video-processing capability.

2.5.4. Content multiplexing

Content multiplexing is the combining of Essence (Video, Audio and Data) and Metadata elements to preserve exact or approximate timing relationships between these elements.

Multiplexing can be carried out in a single step or in several cascaded stages to achieve different aims:

- the formation of a single multichannel component of one Essence type from the components of individual channels –for example, stereo audio pairs or multi-track audio, or sets of Metadata;
- the formation of a single SDI stream, comprising component video with audio, Data Essence and Metadata in ancillary space;
- the formation of a single Content package from several Content items –for example, a typical video plus stereo audio plus Closed Caption programme in a single FC-AV Container;
- the formation of a multi-programme package from the Content items of several packages –for example, a broadcast of multiple programmes in a single MPEG Transport Stream;
- the formation of a multi-programme package from several single Content packages – for example, multiplexed independent SDTI-CP packages on a single SDTI link;
- the formation of a single-programme package from several single Content packages – for example, multiplexed independent DIF packages on a single SDTI link;
- the multiplexing of a package to achieve faster-than-real-time transfer –for example, 4X transfer over SDTI by mapping four sequential DIF frames or SDTI-CP packages into a single SDTI frame.

Content streaming applications may need to achieve several of these aims at once. The multiplexer may achieve this in a single step or in cascaded stages.

In addition, the multiplexing function may be provided by the Wrapper or, in some cases, by the capabilities of the underlying interconnect technology.

2.5.5. Multiplexing of Essence into containers

When multiplexing Video, Audio, Data Essence and Metadata, there is a need to preserve the timing relationship between these components within some tolerance. This tolerance will depend on the circumstances. For example, with MPEG-2, the transmission tolerance can vary by a margin greater than the presentation tolerance. The maximum delay in timing tolerance will affect the latency that a component exhibits in the studio.

It is recognized that some Metadata is synchronous whereas other Metadata is not. SMPTE timecode is an example of synchronous Metadata. It must arrive within frame timing constraints of the Video Essence. A production script is an example of asynchronous Metadata. It must be delivered sometime during the transmission of the Essence but need not arrive within tight timing constraints.

It would be desirable to define the way in which multiplexers are used in the studio:

- multiplexing may occur immediately before transmission, or;

- ⇨ multiplexing may occur within the studio (for example, streaming using the Real-Time Streaming Protocol and the Real-Time Protocol as defined by the IETF).

It is desirable to have commonality of containers as far as possible; at least, the individual components must be identifiable in each different container design so that gateways between interconnects can easily be built.

Multiplexers should allow easy combination and separation of the multiplex elements.

There can be potential problems of timing alignments / data alignments within containers (for example, 480p 4:2:0P over 360 Mbit/s, by using only 64 bytes HANC instead of 260 bytes), and also in the timing between containers / programmes within a multiplex.

2.5.5.1. Transferring Essence between multiplexes

Formats for containers should be described by recognized standards. This will permit, as far as possible, easy conversion between different systems. For example, the MPEG-2 TS and FC-AV are both recognized standards. The delivery of Essence from one multiplex container to another will be determined by the constraints of both formats. A standardized mapping for transcoding the Video, Audio, Data Essence and Metadata between popular multiplex containers may be needed for some combinations, e.g. SDTI CP to TS. This effort may be minimized by the harmonization of different containers.

There is the potential for problems when maintaining the timing relationship between Data Essence and Metadata elements in a container. For example, the VBV buffer model used within MPEG-2 may require that an Encoder processes video data, even though this may be in conflict with the transmission requirements for Data Essence and Metadata. In such an instance, it is the responsibility of the multiplex control process to accommodate these problems, either by managing the Encoder rate control (for streams that are being encoded in real-time) or by appropriately distributing the Data Essence / Metadata within the bitstream by prior analysis of the stream.

2.5.5.2. How to deal with opportunistic data over transports

The delivery of opportunistic data over either a telecommunications network or a satellite feed is being discussed in ATSC S13. The group has identified two scenarios: *reserved bandwidth* and *opportunistic*. In the reserved bandwidth scenario, a customer pays for a fixed bandwidth channel. Channels will be defined in quanta, e.g. 384 kbit/s, 1.544 Mbit/s, etc. The channel will be available in a continuous fashion (guaranteed bandwidth over short intervals of time). These channels will be associated with the customer by means of entries in the Programme and Systems Information Protocol (PSIP) tables. In the opportunistic scenario, a customer pays for some bandwidth over a period of time (e.g. an average of 2 Mbit/s over a 24-hour period) but without control as to when the data are transferred during this period. Hence, the multiplex system is at liberty to transport the data continuously or in bursts (e.g. a single burst of 2 Mbit/s for 1 hour followed by 23 hours without data transport). Through the use of systems information tables, these channels are not directly associated with a customer.

2.5.5.3. Multiplexing of different systems / formats

There are many instances when a multiplex may contain several different container formats or where several different multiplexes are to be processed concurrently, e.g. when a multiplexer has as input MPEG-2 TS over ATM and STDI-ES over SDI. In these instances, a timing relationship between the container formats is required. For example, it may be necessary to map PCRs recovered from the MPEG-2 TS to timecodes in the SDTI-CP. There are issues associated with the stability and resolution of timing references in these circumstances.

2.5.5.4. Statistical multiplexing considerations

Situations do occur where all inputs to a multiplex have a common, or similar, picture. In these instances, statistical multiplexing may fail when complex scenes are input. This issue will require consideration when planning a studio system. Differing risk assessment will be appropriate for contribution, distribution and emission. In the case of contribution and distribution, multiplexing failures will be unacceptable. In the case of emission (e.g. a cable head-end), the large number of non-live channels should ensure that multiplexing will work. Recommended guidelines include: providing a good mix of live vs. non-live shows on a single multiplex channel, ensuring that many news channels are not placed on the same multiplex, and using pointers within the transport multiplex to a single channel as opposed to transmitting duplicate channels.

2.5.6. Timing, synchronization and spatial alignment

2.5.6.1. Reference signals

The reference signal recommended for both analogue and digital areas will be the analogue colour, black. The original recommended practice, SMPTE RP154, has been enhanced for use in future systems and is being re-issued as a standard. This standard expresses the option of including:

- Vertical Interval Timecode (VITC);
- *(for use with 59.94 Hz and 60.00 Hz related systems),* a sequence of ten field-identification pulses to assist in the locking of 24 Hz related signals and the five-frame sequence that is implicit in the relationship with 48 kHz.

It is recommended that both these options be exploited.

2.5.6.2. Absolute time reference

With the availability of satellite-based highly-accurate frequency references, it is recognized that it is now possible to provide synchronization lock to a common global clock. This offers considerable advantage in the management of delay and latency. It is recommended that, where possible, studio synchronization lock be provided using the Global Positioning System (GPS) or equivalent, with redundancy receivers, as a source of data which offers absolute time. The timecode standard, SMPTE 12M, has been revised to include absolute time, the day and the date. Because these techniques rely on good reception, it is recognized that there will be "impossible" situations (e.g. trucks in tunnels) for which an alternative time reference will be required. However, for fixed locations it should be practical to implement GPS as the time reference.

2.5.6.3. Temporal alignment

When cascading compression processes, significant loss of quality can be expected if the GoP structure of the initial compression is not maintained throughout the processing path. Because of the differing requirements of the different parts of the signal path (inter-facility vs. intra-facility, for example), multiple GoP structures will be the norm, even within a single facility. It is recommended that the multiple GoP structures have a simple integral relationship with each other, and that the I-frames be aligned. However, it is possible that attempting such alignment between GoPs of differing length may induce a cyclic variation in picture quality.

2.5.6.4. Timing constraints on compressed signals

The MPEG toolkit allows the system designer a great deal of flexibility. However, operational requirements for some procedures, such as editing, impose constraints on how those tools may be used. In general, as the constraints increase in severity, the amount of bit-rate reduction that is achievable decreases. For this reason, it is unlikely that any one set of constraints will prove satisfactory for every requirement.

For editing in compressed systems, the Task Force recommends that every GoP of a compressed bitstream sequence contain a constant number of bytes and a constant number of frames. This is necessary for tape-based systems and can be useful in some disk-based systems. There are two ways this requirement can be achieved. The first is to have the coder utilize all of the available bytes in each GoP. The second is achieved by bit-stuffing a stream which has been coded with a constant GoP structure and a constrained maximum number of bytes per GoP. When streaming data over networked environments, bit-stuffing is undesirable because it reduces the transmission efficiency, so the stuffed bits should be removed for this type of transmission. For editing flexibility, a short GoP is recommended.

DV-based systems do not employ motion-compensated compression, so only a single-frame GoP is used. Moreover, as DV was optimized for tape systems, all frames are the same number of bits in length.

When streaming data over networked environments, bit-stuffing is undesirable because it reduces transmission efficiency.

2.5.6.5. Latency and delay

Compression inevitably introduces delay. The amount of delay is roughly proportional to the amount of bit-rate reduction and is a consequence of the techniques used to achieve it. For a given compression scheme, encoder and decoder delays may vary between different manufacturers' implementations. In MPEG systems, changes to the GoP structure can have significant impact on latencies.

In playout systems working entirely with pre-recorded material, these latencies can be compensated out. However, when live material must be dealt with, account must be taken of these latencies. They should be predictable and / or of constant duration, and it will be incumbent on the designer and operator to factor in decoder pre-charge time, analogous to the pre-roll time of earlier VTRs. Moreover, it will be necessary to provide audio clean feeds ("mix-minus" programme feeds) to presenters when round-trip delays exceed the comfort threshold.

2.5.6.6. Spatial alignment

In order to cascade compression processes with minimum quality loss, it is necessary to ensure that macroblocks are aligned throughout the production, distribution and emission process. This is particularly true with MPEG, as the quantization matrices and motion vectors are bound to the macroblock structure. Unfortunately, due to (i) the different MPEG encoder implementations, (ii) the different specifications for production and emission using MPEG compression and (iii) the lines coded by MPEG and DV, it is difficult to maintain this alignment. Of particular concern is the line length of the 480-line formats as coded by the ATSC system.

2.5.6.7. Hybrid analogue / digital facilities

In hybrid analogue and digital equipment, the signal timing necessary for the analogue system is typically much tighter than that required for the digital equipment. The necessity for conversion between the two complicates this further, especially as the analogue equipment is typically composite while the digital equipment is component. One approach which has proven effective is to set a zero-time point for the analogue equipment, but to allow the digital equipment a timing tolerance (a so-called "timing band") which is several microseconds wide. A/D and D/A conversions between component analogue and component digital have latencies that will fit within this band. D/A conversions with composite encoding will often fit as well. A/D conversions combined with composite-to-component decoding are typically over one line long. For these, it is necessary to delay the output signals – and any accompanying audio and time-dependent data – for the remainder of the frame until they are again in sync.

2.6. Interconnection options

The EBU / SMPTE Task Force recognizes the complexities of defining the various interfaces for Control, Video, Audio, Metadata, etc., based on various applications.

When fully defining issues such as file transfer interfaces, communication protocols, physical interfaces, link layers, real-time transfers, streaming in real- and non-real-time, linking of Metadata, transporting of Metadata, etc., the result is a multi-layered matrix that requires further definition and study.

An objective of the follow-up activities by the SMPTE should be to define preferred implementation interfaces and protocols, along with templates or matrix charts that will guide the industry. The output from the SMPTE committee should consist of drawings / charts / text that will provide industry users and manufacturers with templates that can be implemented.

The text that follows in this section is some general guidance on this complex subject, prior to the availability of the templates.

2.6.1. Development of templates

The SMPTE systems committee should be instructed to study the EBU / SMPTE Task Force Final Report as the basis for their deliberations in defining:

- interfaces;
- communication protocols;
- packetizing;
- control layers;
- bit-rates;
- Metadata forms;
- . . . etc.

The end objective should be to develop possible templates that could be implemented. It is recognized that different applications may well result in different implementations.

It is understood by the EBU / SMPTE Task Force that other bodies in the industry such as ATSC / DVB etc. have committees working in similar area: work undertaken by these groups should be used in conjunction with the SMPTE effort.

The SMPTE committee should consider all aspects of the Digital Broadcast model, which will also include Data Broadcasting and Metadata.

2.6.2. Transfer definitions

There are three types of transfer operation:

- **Hard real-time**;
- **Soft real-time**;
- **Non real-time**.

2.6.3. Hard real-time

This is an event or operation that must happen *at* a certain time with no opportunity to repeat, and where there may not be an opportunity to re-do (e.g. a live event). This operation has the highest priority.

2.6.3.1. Play-to-air

2.6.3.1.1. Video

Current practice in a studio facility for digital systems is uncompressed video over SDI.

Future practice will include compressed MPEG and / or DV in addition to the uncompressed video. The transport of compressed video data will be via SDTI and / or networking infrastructures.

Primary distribution over public network and / or satellite links will be via an MPEG Transport Stream.

2.6.3.1.2. Audio

Current and future practice within a studio facility is / will be to use AES-3, either on separate circuits or embedded within the compressed video stream. In acquisition applications, uncompressed 48 kHz PCM audio may be used.

Primary distribution over the public network and / or satellite links will be via an MPEG Transport Stream.

2.6.3.1.3. Data Essence and Metadata

Within a studio facility, current practice is to use RS-422 or Ethernet; future practice will also transport the data using payload space within SDTI and SDI formats and / or other networking infrastructures.

The primary emission format for Data Essence and Metadata is the Teletext transport, embedded in video.

2.6.3.1.4. Control

Current practice is to use GPI, RS-422 and Ethernet, with a range of protocols.

In the future, control should use the IP protocol.

2.6.3.2. Live-to-air

Live-to-air operation will follow the above interconnections. Contribution circuits currently use MPEG or ETSI compression over a variety of bearers. Future contribution circuits will migrate to MPEG over a variety of transport networks and bearers.

2.6.3.3. Live recording

Live recording follows the same practice as Live-to-air.

2.6.4. Soft real-time

This is an event or operation that must happen *by* a certain time and where there may be an opportunity to re-do. At the end of the available time window, this operation becomes hard real-time.

2.6.4.1. Video and Audio (uncompressed)

In the transport of uncompressed video data, SDI is and will be the transport method for both SD and HD signals.

The associated audio signals will be AES-3 audio over separate circuits although, in acquisition applications, uncompressed 48 kHz PCM audio may be used.

2.6.4.2. Video and Audio (compressed)

Transport of compressed video will be via SDTI and / or networking infrastructures carrying MPEG or DV compressed data in both SD and HD systems.

The associated audio signals will be AES-3 audio over separate circuits or embedded in the compressed video data stream.

2.6.4.3. Data Essence and Metadata

Within a broadcast facility, current practice is to use RS-422 or Ethernet; future practice will transport the data using the ancillary data space within the SDTI and SDI formats and / or other networking infrastructures.

2.6.4.4. Control

Current practice is to use GPI, RS-422 and Ethernet, with a range of protocols.

In the future, control will use IP protocol.

2.6.5. Non real-time

Non real-time encompasses operations that need not be completed within time boundaries (e.g. file transfer that is faster or slower than real-time).

Non real-time transfers will either be by file transfer or streaming methods. Both of these methods can be applied to compressed or uncompressed data, although the compressed domain will be the most common.

2.6.5.1. File transfers

File transfers will occur between storage devices (e.g. edit stations and file servers) over network-based infrastructures such as Fibre Channel, Ethernet and ATM.

The file transfer protocol will be FTP or FTP+ as described in *Section 5.* and *Section E.2.3.* File transfer is as fast as possible within the constraints of the channel that has been requested. There may be operational needs to interleave the Video, Audio, Data Essence and Metadata; such requirements are application-dependent as described in *Section 4.3.*

Audio may be separately handled as EBU Broadcast Wave Files (BWF).

2.6.5.2. Control

Control of file transfers will typically be accomplished from user applications over IP protocol.

For file transfers over a unidirectional transport such as SDTI, a back channel such as Ethernet is required to confirm reception of the file.

Other forms of file transfers can be achieved between platforms using tools such as NFS.

2.6.5.3. Streaming

Non-real-time streaming will occur between storage devices over SDTI and / or networking infrastructures.

For SDTI, a control plane using a network infrastructure is required. Non-real-time streaming requires that containers possess explicit or implicit timing structures. It is possible that normal file transfer methods can be used for streaming but, in these instances, there are potential problems with respect to buffer management at the receiving end.

Special processors that can display non-real-time pictures are desirable. Examples include downstream monitoring which can monitor high-speed streaming transfers.

Audio may be separately handled in the form of EBU Broadcast Wave Files (BWF).

2.7. Migration

2.7.1. Placement of existing protocols into the public domain

Protocols need to be available through a single point of contact. This can be achieved by manufacturers keeping protocol documents accessible to the public on their own Web sites, linked to an index of available protocols on the SMPTE Web site. For equipment whose manufacturer support has been discontinued, the SMPTE can act as a central repository for protocol documentation.

This will encourage the wholesale migration to a common, open control environment and will demonstrate an industry commitment towards this process. It will also allow easier integration of individual products into complex systems. Participation in the publication of existing protocols will allow attributes of those protocols to be considered when developing future standards. The Task Force views the establishment and maintenance of the index and repository as services that are essential to the industry.

Issues arising from this migration include provision of high-quality protocol documentation, active support for interface developers using the protocols, and the availability of development and test tools such as simulators.

2.7.2. Essential Common Protocols

Information from the Object Reference Model and documented existing protocols will be used in the development of an Essential Common Protocol set. This set defines a standardized way to perform a given operation on any device that supports that operation.

2.7.3. Interoperability between old and new systems

It must be recognized that, during this migration, users will operate devices and / or control systems which employ a combination of the new and existing protocols. Possible techniques to achieve this are control systems that support old and new protocols, devices that support old and new protocols, or proxies that translate protocols.

2.8. Economic model

The value and cost of digital equipment are now primarily in the software rather than in the hardware. However, the marketing model for broadcast equipment is still based on loading all the development and support cost into the hardware, while the software, including upgrades and maintenance, is provided at low or no cost. This model is outdated and serves neither the manufacturer nor the purchaser well.

In the computer industry, software and hardware are sold and supported separately. Software maintenance costs are an accepted fact of life, and ongoing support is factored in as part of the ongoing cost of ownership. Since the two are separate, it can be relatively painless to upgrade one without upgrading the other. This has evolved to its limit in the PC industry, where software and hardware have been specialized to such an extent that companies manufacture either one or the other, but rarely both.

Currently, and for the foreseeable future, products will be differentiated primarily by their software. As such, it is expected that software will absorb the lion’s share of the development costs. By bundling the cost of software development and support into the purchase price of equipment, manufacturers force users to capitalize not only the cost of the hardware but the initial and ongoing support costs of the software over the useful life of the product. This serves neither party well, as it drives up initial costs to extremely high levels, while encouraging frequent hardware “churn” (turnover) for the sake of supporting the software development.

For future products, manufacturers and users alike should consider shifting to a computer-industry cost model, where hardware and software are un-bundled and ongoing support charges are factored into the pricing model. This will encourage manufacturers to produce (and users to expect) ongoing software upgrades and maintenance, as well as stability in manufacturers’ support efforts. It may also encourage manufacturers to introduce moderately-priced hardware upgrades that preserve and extend their customers’ investments in software.

It is possible that such a pricing model might promote a view of some hardware as undifferentiated commodity items; however, if the revenue can come from the software, which is the real source of utility and differentiation, this should not be a problem. Allocating costs where they truly belong will in the long run benefit all parties.

2.9. Standards

The EBU / SMPTE Task Force has identified areas of further work for due-process standards organizations. This work has already (July 1998) started in the SMPTE which has re-organized its committee structures to speed up the implementation of the required standards.

All of the standards efforts described below must take into account the very real need to provide documented and secure methods for the extension of any protocols or formats. Where possible, this should be by the use of a Registration Authority to register new data types and messages.

2.9.1. Work under way in the SMPTE

- Development of standard, LAN-based common device dialects for system-to-device communication.
  - A standard LAN-based control interface for broadcast devices is being developed to allow these to be connected to control systems that use IP-based transport.
- Harmonization of Material Identifiers.
  - For system-to-system or system-to-device communication, a common unambiguous means of uniquely identifying media should be developed: the UMID. Systems and devices could continue to use their own proprietary notation internally but, for external communication, this should be translated to the standard UMID. See *Section 4.6.* for a discussion on the format specification of UMIDs.
  - For the smaller subset of completed programmes, a more compact, human readable, still globally unique identifier will probably be used for the entire programme Content.
  - Management systems must be able to deal both with UMIDs and with the multiplicity of programme identifiers that will exist and must accommodate the case in which portions of completed programmes are used as source material for other programmes. To enable management systems to recognize this variety of identifiers, each type of identifier should be preceded by a registered SMPTE Universal Label.

A Common Interface for system-to-system communication is:

- **Transfer Request** format:
  - A standard format for messages requesting the transfer of material from one location to another should be developed. This format must allow for varying file system and transfer capabilities in the underlying systems and should allow different qualities of service to be requested. The possibility of automatic file translation from one format to another as a transparent part of the transfer should also be considered.

2.9.2. Standardization efforts yet to be undertaken

- **Event List Interchange** format:
  - A standardized event list interchange format should be developed. The format should allow vendors to develop a standardized interface between business scheduling (traffic) systems and broadcast / news automation systems.
- **Transfer Request** format:
  - A standard format for messages requesting the transfer of material from one location to another is required. This format must allow for varying file system and transfer capabilities in the underlying systems and should allow different qualities of service to be requested. The possibility of automatic file translation from one format to another as a transparent part of the transfer should also be considered (see *Annex B*).
- **Content Database Interchange** format:
  - To facilitate inter-system communication about media assets, a reference data model for the Content description should be developed. Systems can internally use their own data model, but this must be translated for external communication.

The Task Force recommends the development of an object-oriented reference model for use in developing future Content creation, production and distribution systems. A listing of technologies which are candidates for study in development of the object model is given in *Section 2.10.1.*

In addition, the Task Force recommends that for this process UML (Unified Modelling Language) be used for drawing, and IDL (Interface Definition Language) be used for the definition of APIs.

The Task Force believes that a "Central Registry" for Object Classes will be required and that a "Central Repository" is desirable.

2.9.3. Summary of standards required

The SMPTE should provide recommendations for encapsulating Data Essence and Metadata in existing storage devices and transport systems. It should also provide standards guidance for the insertion and extraction of Data Essence and Metadata into containers.

2.10. References

2.10.1. Object-oriented technologies

Table 2.2 provides a list of Websites and printed publications where further information on object-oriented technologies and Java may be obtained.

Table 2.2: Further information on object-oriented technologies and Java.

Object-oriented technologies	
"Object-oriented Analysis and Design with Applications" by Grady Booch, 2nd edition, February 1994:	**Addison Wesley Object Technology Series ISBN: 0805353402.**
"Architecture of the Virtual Broadcast Studio" by Ken Guzik:	**SMPTE Journal Vol. 106, December, 1997** http://www.smpte.org/publ/abs9712.html
"Discovering OPENSTEP: A Developer Tutorial - Appendix A, Object Oriented Programming":	http://developer.apple.com/techpubs/rhapsody/DeveloperTutorial_NT/Apdx_OOP.pdf
OPENSTEP White papers:	http://enterprise.apple.com/openstep/whitepapers.html
"CORBA Overview":	http://www.infosys.tuwien.ac.at/Research/Corba/OMG/arch2.htm#446864
"The Common Object Request Broker (CORBA): Architecture and Specification", Revision 2.2, Feb. 1998:	**Object Management Group, Framingham, MA 01701, USA** http://www.omg.org/corba/corbiiop.htm
CORBA Tutorial:	http://www.omg.org/news/begin.htm
What is CORBA?:	http://www.omg.org/about/wicorba.htm
CORBA / IIOP specification, version 2.2:	http://www.infosys.tuwien.ac.at/Research/Corba/OMG/arch2.htm#446864
Object Management Group home page:	http://www.omg.org
"OMG White Paper on Security", OMG Security Working Group, Issue 1.0, April, 1994:	ftp://ftp.omg.org/pub/docs/1994/94-04-16.pdf
DCOM (Microsoft Distributed Common Object Model):	http://www.microsoft.com
DSOM (IBM Distributed System Object Model):	http://www.rs6000.ibm.com/resource/aix_resource/Pubs/redbooks/htmlbooks/gg244357.00/somwdsom.html
Java information	
"Java Programming Language", 2nd. Edition, by Ken Arnold and James Gosling:	**Addison Wesley, ISBN: 0201310066 5**
Java home page:	http://www.java.sun.com
Java 8-page overview:	http://java.sun.com/docs/overviews/java/java-overview-1.html
Index to Java documents:	http://www.java.sun.com/docs/index.html
Index to Java white papers:	http://java.sun.com/docs/white/index.html
"Java Remote Method Invocation – Distributed Computing For Java":	http://www.java.sun.com/marketing/collateral/javarmi.html

Section 3

Compression issues

3.1. Introduction

This section of the report details the findings of the Task Force's Sub-Group on Compression.

The first Task Force report, issued in April 1997, presented an overview of video compression, covering a broad range of issues related to the use of video compression in television applications. Significant parts of that first report are repeated here, so that the user may better understand the progress and recommendations that appear in later sections of this Final Report.

Since the printing of the First Report, the Sub-Group on Compression has entertained in-depth discussions on the compression schemes available today and in the foreseeable future, and on the balances obtained in terms of:

- ⇨ ultimate technical programme quality versus data-rate;
- ⇨ editing granularity versus complexity of networked editing control;
- ⇨ interoperability of compression schemes using different encoding parameters.

The decision to use compression has a significant impact on the overall cost / performance balance within television production and post-production operations, as it will affect the quality, storage / transmission efficiency, latency, editing / switching of the compressed stream as well as error resiliency.

Compression is the process of reducing the number of bits required to represent information by removing redundancy. In the case of information Content such as video and audio, it is usually necessary to extend this process by removing information that is not redundant but is considered less important. Reconstruction from the compressed bitstream thus leads to the addition of distortions or artefacts. Compression for video and audio is therefore not normally lossless.

Thus it is important to make decisions about compression at the source, taking into account the additional production processes and additional compression generations that will follow. These decisions are quite likely to be different from the choices that would be made if the compression were simply done only for presentation to a human observer.

This section considers a wide range of compression characteristics. Compression of video and audio allows functionality that is not viable with uncompressed processing. Through the reduction of the number of bits required to represent given programme Content, it makes economical the support of applications such as the storage of material, the transmission of a multiplicity of programme elements simultaneously through a common data network, and simultaneous access to the same Content by a number of users for editing and other processes.

Choices made with regard to compression techniques and parameters have significant impacts on the performance that can be achieved in specific applications. Consequently, it is most important that those choices be made with the attributes of the associated application clearly understood. The application includes not only the processing that immediately follows the compression process but also any subsequent downstream processing operations. This section provides users with information about compression characteristics to assist in making judgements about appropriate solutions. It further recommends approaches to be taken to facilitate interoperation to the greatest extent possible between systems within a single family of compression techniques and between families of compression methods.

This section includes examples of 525- and 625-line implementations of interlaced television systems that exchange programme Content as bitstreams, but it is clearly expected that the techniques will be extensible to systems having higher numbers of lines, progressive scanning, and other advanced features.

See *Section 3.10.* for the Task Force's comments on HDTV.

3.2. Image quality

The majority of broadcast production and post-production processes still cannot be performed today by direct manipulation of the compressed data stream, even within a single compression family. Techniques for minimizing the quality loss in production and post-production processes – by direct manipulation of the compressed bitstream or by using special "helper data" – have been proposed and submitted to the SMPTE for standardization. The achievable balance between the gain in picture quality and the increased system complexity remains to be assessed. Furthermore, some proposals for the use of this "helper data" have negative impact when operating with signals which have not been subject to compression. The consequent cascading of decoding and re-encoding processes within the production chain, and the quality losses incurred, therefore require the adoption of compression schemes and data-rates which support the picture-quality requirements of the ultimate output product.

Selection of compression system parameters has a significant impact on the overall image quality. These compression parameter choices must be optimized to preserve the image quality while at the same time fitting the image data into the available bandwidth or storage space. Different combinations of compression parameters may be best for different specific applications.

Compression system parameters which should be considered include: the underlying coding methods, the coding sampling structure, pre-processing, data-rates and the Group of Pictures (GoP) structure used. In choosing the compression system parameters, interaction between the parameter choices must also be considered. Finally, special operational issues such as editing the bitstream or splicing new Content into an incoming bitstream should be considered.

Annex C contains EBU equipment evaluations which are provided as reference information for specific implementations of both MPEG-2 4:2:2P@ML and DV-based compression systems. No inference should be drawn from the inclusion or omission of any implementations, nor are these evaluations included for the purpose of side-by-side comparison. The subjective tests on compression systems used different test conditions for lower data-rates and higher data-rates. These conditions were adapted (and agreed by the manufacturers concerned) to the individual fields of application envisaged. No M-JPEG systems have been offered for evaluation and therefore subjective quality evaluations are unavailable at the time of writing (July 1998).

3.2.1. Coding method

The coding method is the most fundamental of compression choices. There are three compression families used in the television production and distribution chain: MPEG-2, Motion JPEG (M-JPEG) and DV. All of these coding methods are based on the Discrete Cosine Transform (DCT). They use normalization and quantization of the transform coefficients, followed by variable length coding.

In its tool kit of techniques, MPEG includes motion estimation and compensation which may be optionally applied. This allows improved coding efficiency, with some cost penalty in memory and processing latency. M-JPEG and DV are both frame-bound, thereby minimizing the coding cost, but these frame-bound coding methods do not take advantage of the coding efficiency of inter-frame motion estimation and compensation. MPEG-2 and DV both allow motion adaptive processing in conjunction with intra-frame processing.

3.2.2. Sampling structure – SDTV

MPEG-2, M-JPEG and DV can all be used with the 4:2:2 pixel matrix of ITU-R BT.601. MPEG-2 and M-JPEG can both be used with other pixel matrices, multiple frame-rates, and either interlace or progressive scan. Note that the 4:2:2 matrix is sub-sampled from the original full-bandwidth (4:4:4) signal. The pixel matrix can be further sub-sampled to reduce the signal data, with 4:2:2 sampling normally being used for interchange between systems. The following sampling structures are in common use:

- **4:2:2 systems** – such as the MPEG-2 4:2:2 Profile, 4:2:2 M-JPEG and the DV 4:2:2 50 Mbit/s system – which all use half the number of colour-difference samples per line, compared with the number used in the luminance channel. 4:2:2 provides half the horizontal bandwidth in the colour-difference channels compared to the luminance bandwidth, while maintaining the full vertical bandwidth.
- **4:1:1 systems** – such as DV 525 – which use one quarter the number of colour-difference samples per line, compared with the number used in the luminance channel. 4:1:1 reduces the colour-difference horizontal

bandwidth to one quarter that of the luminance channel, while maintaining the full vertical bandwidth. The filters used to achieve the 4:1:1 sub-sampled horizontal bandwidths, like other horizontal filters, generally have a flat frequency response within their pass-bands, thereby enabling translation to and from 4:2:2 with no further degradation beyond that of 4:1:1 subsampling.

- **4:2:0 systems** – such as DV 625 [3],and MPEG-2 Main Profile – which use half the number of colour-difference samples horizontally and half the number of colour-difference samples vertically, compared to the number used in the luminance channel. 4:2:0 therefore retains the same colour-difference horizontal bandwidth as 4:2:2 (i.e. half that of the luminance channel) but reduces the colour-difference vertical bandwidth to half that of the luminance channel. 4:2:0 coding, however, generally does not provide flat frequency response within its vertical pass-band, thereby precluding a transparent translation to the other coding forms. Consequently, systems that use 4:2:0 sampling with intermediate processing will not, generally, retain the full 4:2:0 bandwidth of the prior coding.

Care must be exercised in selecting compression sampling structures where different compression coding techniques will be concatenated. In general, the intermixing of different sub-sampled structures affects the picture quality, so cascading of these structures should be minimized. For example, while 4:1:1 or 4:2:0 signals will have their original quality maintained through subsequent 4:2:2 processing (analogous to “bumping up” of tape formats), the cascading of 4:1:1 and 4:2:0 may generally yield less than 4:1:0 performance.

3.2.3. Compression pre-processing

Video compression systems have inherent limitations in their ability to compress images into finite bandwidth or storage space. Compression systems rely on the removal of redundancy in the images, so when the images are very complex (having very little redundancy), the ability to fit into the available data space may be exceeded, leading to compression artefacts in the picture. In these cases, it may be preferable to reduce the complexity of the image by other methods, before the compression processing. These methods are called pre-processing and they include filtering and noise reduction.

When noise is present in the input signal, the compression system must expend some bits while encoding the noise, thus leaving fewer bits for encoding the desired image. When either motion detection or motion estimation and compensation is used, noise can reduce the accuracy of the motion processing, which in turn reduces the coding efficiency. Even in compression systems which do not use motion estimation and compensation, noise adds substantial high-frequency DCT energy components which might otherwise be zero. This not only wastes bits on extraneous DCT components, but degrades the run-length-coding efficiency as well.

Compression system specifications generally define only the compression functions within equipment, but do not specify the pre-processing before the compression function. An exception is the shuffling which is an inherent part of the DV family, and is not to be confused with the shuffling used for error management in digital recorders.

Since most pre-processing, such as filtering or noise reduction, is not always required, the pre-processing parameters may be selected depending on the nature of the images and the capabilities of the compression system. These choices can be pre-set or can be adaptive.

3.2.4. Video data-rate – SDTV

- The MPEG-2 4:2:2 Profile at Main Level (MPEG-2 4:2:2P@ML) defines data-rates up to 50 Mbit/s;
- M-JPEG 4:2:2 equipment typically operates at data-rates up to 50 Mbit/s;
- DV / DV-based 4:1:1 and DV 4:2:0 operate at 25 Mbit/s;
- DV-based 4:2:2 operating at 50 Mbit/s is currently undergoing standardization within the SMPTE;
- MPEG-2 Main Profile at Main Level is defined at data-rates up to 15 Mbit/s.

3. Note that the proposed SMPTE D-7 format, although based on DV coding, will use 4:1:1 sampling for both the 525- and 625-line systems.

Selection of the data-rate for MPEG-2 4:2:2P@ML is interrelated with the Group of Pictures (GoP) structure used. Lower bit-rates will typically be used with longer more-efficient GoP structures, while higher bit-rates will be used with simpler shorter GoP structures.

Intra-coded images (MPEG-2 4:2:2 Profile [I pictures only], M-JPEG and DV) at data-rates of 50 Mbit / s can yield comparable image quality.

MPEG-2 4:2:2P@ML with longer GoP structures and lower data-rates can provide comparable quality to shorter GoP structures at higher data-rates – albeit at the expense of latency (see MPEG Group of Pictures below).

3.2.5. MPEG Group of Pictures

There are three fundamental ways in which to code or compress an image:

1. The most basic method is to code a field or frame with reference only to elements contained within that field or frame. This is called intra-coding (I-only coding for short).
2. The second method uses motion-compensated prediction of a picture (called a P picture) from a preceding I-coded picture. Coding of the prediction error information allows the decoder to reconstruct the proper output image.
3. The third method also uses motion-compensated prediction, but allows the prediction reference (called an anchor frame) to precede and / or follow the image being coded (bi-directional or B-picture coding). The selection of the reference for each picture or portion of a picture is made to minimize the number of bits required to code the image.

Sequences of images using combinations of the three coding types, as defined by MPEG, are called Groups of Pictures (GoPs). Both Motion JPEG and DV use only intra-frame coding and therefore are not described in terms of GoPs.

MPEG-2 allows many choices of GoP structures, some more commonly used than others. In general, a GoP is described in terms of its total length and the repetition sequence of the picture coding types (e.g. 15 frames of IBBP). The optimal choice of GoP structure is dependent on the specific application, the data-rate used, and latency considerations.

Since I-only pictures are least efficient and B pictures are most efficient, longer GoPs with more B and P pictures will provide higher image quality for a given data-rate. This effect is pronounced at lower data-rates and diminished at higher data-rates. At 20 Mbit / s, the use of long GoPs (e.g. IBBP) may prove useful while, at 50 Mbit / s, shorter GoPs can provide the required quality.

Besides affecting the image quality, the choice of GoP structure also affects the latency. Since a B picture cannot be coded until the subsequent anchor picture is available, delay is introduced in the coding process. Note, however, that this delay is dependent on the distance between anchor frames, not the total length of the GoP structure. This means that a blend of the coding efficiency of long GoP structures together with the lower latency of short GoP structures can be obtained by judicious use of P-picture anchors.

3.2.6. Constant quality vs. constant data-rate

Compression systems are sometimes referred to as Variable Bit-Rate (VBR) or Constant Bit-Rate (CBR). MPEG-2 and Motion JPEG can operate in either VBR or CBR modes; DV operates only with constant bit-rates. In practice, even those systems commonly believed to be constant bit-rate have bit-rate variations, but over shorter periods of time. Another way to characterize compression systems is to compare constant quality with constant bit-rate systems.

3.2.6.1. Constant quality (VBR) systems

Constant quality systems attempt to maintain a uniform picture quality by adjusting the coded data-rate, typically within the constraint of a maximum data-rate. Since simpler images are easier to code, they are coded at lower data-rates. This results in more efficient compression of simpler images and can be a significant advantage in storage systems and in the non-real-time transfer of images. Constant-quality operation is useful for disk recording and some tape recording systems such as tape streamers.

3.2.6.2. Constant bit-rate (CBR) systems

Constant bit-rate (data-rate) systems attempt to maintain a constant average data-rate at the output of the compression encoder. This will result in higher quality with simpler images and lower quality with more complex images. In addition to maintaining a constant average data-rate, some constant data-rate systems also maintain the data-rate constant over a GoP. Constant data-rate compression is useful for videotape recording and for fixed data-rate transmission paths, such as common carrier services.

Constant data-rate processing will, of course, be characterized by a target data-rate. Variable data-rate processing can be constrained to have a maximum data-rate. By ensuring that this maximum data-rate is less than the target rate of the constant data-rate device, constant quality coding can operate into a constant data-rate environment.

3.2.6.3. Interfacing VBR and CBR environments

The interface between a constant quality (VBR) environment and a constant bit-rate (CBR) environment could be accommodated by bit-stuffing a constant quality stream to meet the requirements of the constant bit-rate environment. Further, the stuffing might be removed when moving from a CBR environment to a VBR environment. Additional work on standards and recommended practices is required to clarify whether this function is part of the VBR environment, part of the CBR environment, or part of the interface between the two.

3.2.7. Editing

Consideration of the compression parameters that relate to editing fall into two general applications categories: *complex editing* and simple *cuts-only editing* (seamless splicing). In the case of complex editing, involving effects or sophisticated image processing and analysis, many of the processes will require decoding back to the ITU-R BT. 601 domain. In these cases, the coding efficiency advantage of complex GoP structures may merit consideration. In the case of cuts-only editing, however, it may be desirable to perform the edits entirely in the compressed domain using bitstream splicing. Bitstream splicing can be done between two bitstreams which both use the same compression method. Data-rates and other parameters of the compression scheme may need to be bounded in order to facilitate splicing. Some existing compressed streams can be seamlessly spliced (to provide cuts-only edits) in the compressed domain with signals of different data-rates.

Techniques for operating directly in the compressed domain are still being developed. Issues relating to editing in the compressed domain are being addressed. It has even been suggested that carrying out more complex operations in the compressed domain may be possible. It should be noted, however, that much of the image degradation encountered in decompressing and re-compressing for special effects will similarly be encountered if those effects operations are performed directly in the compressed domain, since the relationships of the DCT coefficients will still be altered by the effects.

If all the compression coding methods used in an editing environment are well defined in open standards, systems could include multi-format decoding. Multi-format decoding would allow receiving devices to process compressed streams based on a limited number of separate compression standards, thereby mitigating the existence of more than one compression standard.

3.2.8. Concatenated compression

To the greatest extent possible, television systems using video compression should maintain the video in compressed form, rather than employing islands of compression which must be interconnected in uncompressed form. Since several compression and decompression steps are likely, the ability to withstand concatenated compression and decompression is a key consideration in the choice of a compression system. The results of concatenated compression systems will be influenced by whether the systems are identical or involve differing compression techniques and parameters.

There are a number of factors which influence the quality of concatenated compression systems. All the systems considered here rely on the DCT technique. Anything which changes the input to the respective DCT operations between concatenated compression systems can result in the transformed data being quantized differently, which in turn could result in additional image information loss. Furthermore, any changes which result in different buffer management will result in different quantization for a transitional period.

In the case of MPEG coding, any change in the alignment of the GoP structure between cascaded compression steps will result in different quantization, since the P- and B-picture transforms operate on motion-compensated image predictions, while the I-picture transforms operate on the full image.

For MPEG, M-JPEG and DV, any change in the spatial alignment of the image between cascaded compression steps will result in different quantization, since the input to any particular DCT block will have changed. Any effects or other processing between cascaded compression steps will similarly change the quantization.

Concatenated compression processes, interconnected through ITU-R BT.601, will have minimal image degradation through successive generations if the compression coding method and compression parameters, including spatial alignment and temporal alignment, are identical in each compression stage.

It is not always possible to avoid mixing compression methods and / or parameters. In some applications, the total image degradation due to cascaded compression and decompression will be minimized by attempting to maintain the highest quality compression level throughout, and only utilizing lower-quality compression levels where occasionally necessary, such as in acquisition or when using common carrier services. For other applications, however, which must make greater use of lower-quality compression levels, the best overall image quality may be maintained by returning to the higher compression quality level only where dictated by image-processing requirements.

Beyond the quality issues just discussed, there are operational advantages to be realized by staying in the compressed domain. Faster-than-real-time transfers, as well as slower-than-real-time transfers, can be facilitated in the compressed domain. Furthermore, some users would welcome image processing in the compressed domain as a potential means of achieving faster-than-real-time image processing.

3.3. Quality levels

While different compression performance levels will be used in different application categories, users will attempt to minimize the total number of performance levels within their operation. Performance differences will be accompanied by differences in the cost of equipment and the operational costs that are appropriate to the application category. For example, a typical broadcast operation might have three levels of compression quality.

- The **highest** compression quality level, generally requiring the highest data-rate, would be used in applications which require the highest picture quality and in applications which involve extensive post-production manipulation. A key attribute of this quality level is the ability to support multiple-generation processing with little image degradation. The highest compression quality level might therefore be used in some higher-quality production applications, but production applications which require the very highest quality will continue to use uncompressed storage and processing. The highest compression quality would also be used for critical imagery and to archive programme Content which is likely to be re-used in conjunction with subsequent further production processing.
- A **middle** compression quality level would be used in applications which require good picture quality and in applications which involve some limited post-production manipulation. This quality level would support a limited number of processing generations, and might be used for news acquisition, news editing, network programme distribution and local programme production. The quality level would also be used to archive programme Content which may be re-used but is not likely to involve significant additional production processing.
- A **lower** compression quality level would be used in applications which are more sensitive to cost than quality. This quality level would not normally support subsequent processing but might be used for programme presentation or mass storage for rapid-access browsing. The lower compression quality would not generally be used to archive programme Content which might be re-used.

These examples of highest, middle and lower compression quality levels do not necessarily correspond to any particular absolute performance categories, but rather should be taken as relative quality levels to be interpreted according to the specific requirements of a particular user's criteria. Further details on particular applications and their use of compression can be found in the First Report of the Task Force, issued in April 1997.

The EBU has acknowledged different levels of compression within the confines of professional television production and post-production (see *Annex C*). Further adaptations will be defined to overcome bottlenecks created by constraints, e.g. bandwidth, tariffs and media cost.

The notion of different quality levels naturally leads to compression families. A compression family can then be defined by the ease of intra-family bitstream transcoding and the availability of an *agile decoder* in integrated form.

The coexistence of different compression families in their **native form** within both local and remote networked production environments requires the implementation of hardware-based **agile decoders**. In many instances, such decoders must allow "glitchless switching" and can therefore realistically be implemented within **one compression family** only. Software-based agile decoding is currently not considered to be a practical option. It is currently still undefined how an agile decoder will output the Audio and Metadata part of the bitstream.

The Sub-Group on Compression concluded that, within the foreseeable future, the coexistence and interoperation of **different compression families** within a networked television facility will pose a number of operational problems and will therefore be the exception and not the rule.

The appropriate selection of a single compression scheme – or a limited number of compression schemes within one compression family, together with the publicly-available specifications of the relevant transport streams and interfaces – will be of overriding importance if efficient exploitation of the potential offered by networked operating environments is to be achieved in the future.

For core applications in production and post-production for Standard Definition Television, two different compression families on the market are currently advocated as candidates for future networked television production:

- ⇨ **DV / DV-based 25 Mbit/s with a sampling structure of 4:1:1**, and **DV-based 50 Mbit/s with a sampling structure of 4:2:2 using fixed bit-rates and intra-frame coding techniques exclusively**.
  DV-based 25 Mbit / s with a sampling structure of 4:2:0 should be confined to special applications.
- ⇨ **MPEG-2 4:2:2P@ML using intra-frame encoding (I) and GoP structures, and data-rates up to 50 Mbit/s [4,5].**
  MPEG-2 MP@ML with a sampling structure of 4:2:0 should be confined to special applications.

3.4. Operational considerations

Systems of all compression performance levels must be fully functional in their intended applications. Equipment that employs compression should function and operate in the same manner as (or a better manner than) similar analogue and non-compressed digital equipment. The use of compression in any system should not impede the operation of that system.

If it is possible to select and alter the compression characteristics as part of the regular operation of a compressed system, such selection and alteration should be made easy by deliberate design of the manufacturer. Variable compression characteristic systems should possess user interfaces that are easy to learn and intuitive to operate. In addition, selections and alterations made to a compressed system must not promote confusion or compromise the function and performance of the systems connected to it.

More than a single compression method or parameter set might be employed in a television production facility. Where this is the case, these should be made interoperable. Compression characteristics used in the post-production process must concatenate and interoperate with MPEG-2 MP@ML for emission.

It is well recognized that integration of compressed video systems into complex systems must be via standardized interfaces. Even with standardized interfaces, however, signal input / output delays due to compression processing (encoding / decoding) occur. System designers are advised that this compression latency, as well as stream synchronization and the synchronization of Audio, Video and Metadata, must be considered. Efficient video coding comes at the expense of codec delays, so a balance must be achieved between

4. For specific applications, this also includes MPEG-2 MP@ML if decodable with a single agile decoder.
5. For recording on a VTR, a fixed bit-rate must be agreed for each family member.

the minimum codec delay and the required picture quality. This may be particularly important for live interview feeds, especially where the available bandwidth is low and the real-time requirement is high. Compressed systems must be designed to prevent the loss of synchronization or disruption of time relationships between programme-related information.

Compressed signal bitstreams should be designed so that they can be formatted and packaged to permit transport over as many communications circuits and networks as possible. Note that compressed bitstreams are very sensitive to errors and therefore appropriate channel-coding methods and error protection must be employed where necessary.

Provision should be made for selected analogue VBI information to be carried through the compression system, although not necessarily compressed with the video. Additionally, selected parts of the ancillary data space of digital signals may carry data (e.g. Metadata) and provision should be made to carry selected parts of this data through a transparent path, synchronously with the Video and Audio data.

3.4.1. Working with existing compression families

The EBU / SMPTE Task Force has completed its studies on digital television compression systems. Emphasis was placed on systems that may be adopted for the exchange of programme material in the form of compressed bitstreams within a networking environment. The principal applications are those of television programme contribution, production and distribution.

In considering the recommendations of the first Task Force report, it was desired to arrive at a single compression family for these applications. There were, however, already a number of compression families in use, including Motion JPEG (M-JPEG), MPEG and DV. With the diversity of compression families already in place, it was not possible to choose one compression family over all the others. The Task Force Sub-Group on Compression recommends that the MPEG and DV compression families be applied to meet those requirements.

M-JPEG has already been used in a variety of professional television applications. Current products take advantage of the simplicity and maturity of JPEG compression components. Due to large investments in both M-JPEG programme material and equipment, some users will continue to work with M-JPEG. As the MPEG and DV technologies mature, it is anticipated that they will displace M-JPEG in many of its current roles. The functions which have been provided by M-JPEG-based compression will, in the future, be served by using intra-frame-coded MPEG or DV compression families.

In order to provide a bridge between M-JPEG, MPEG and DV, it is important to consider the coexistence of multiple compression families during the transition period. A pathway towards interoperability between M-JPEG and future compression formats is essential.

The EBU / SMPTE Task Force recognizes the existence of a diversity of compression formats. It also recognizes the need for diversity to address effectively the different applications and to provide a pathway for the future implementation of new demonstrably-better formats as compression technology evolves.

3.4.2. Agile decoders

Two types of SDTV agile decoder have been discussed:

- ⇨ **common agile decoders** which allow the decoding of multiple compression families;
- ⇨ **intra-family agile decoders** which are confined to the decoding of bitstreams of different parameters within a single compression family.

The Task Force has requested and received written commitments from major proponents of DV and MPEG to provide an *intra-family agile decoder* in integrated form.

It will address the following requirements:

- ⇨ the decoding of different bitstreams with identical decoding delay at the output;
- ⇨ intra-family switching between different bitstreams at the input;
- ⇨ intra-family decoding between different bitstream packets within a single bitstream.

3.4.3. Native decoders

Native decoders designed to operate on non-standard bitstreams – e.g. for optimized stunt-mode performance (shuttle, slow-motion) or for special functions – are acceptable. The decoder chip-set should be available on a non-discriminatory basis on fair and equitable conditions. Details of possible deviations from the standardized input data stream should be in the public domain.

3.5. Family relations

3.5.1. Tools available for intra-family transcoding

For reasons of restricted network bandwidth or storage space, a higher data-rate family member may have to be converted into a lower data-rate member. In the simplest case, this can be performed by simple decoding and re-encoding. Under certain conditions, the quality losses incurred in this process can be mitigated by re-using the original encoding decisions. This can be performed within a special chip or by retaining the relevant information through standardized procedures. The table in *Annex C.4.10.2.* indicates the options available for each family.

3.5.2. Compatible intra-family record / replay

Operational flexibility of networked production will be influenced by the availability of recording devices which can directly record and replay all intra-family bitstreams or which allow the replay of different bitstreams recorded on cassettes. The tables in *Annex C.4.12.* indicate the options available for each family.

3.5.3. MPEG at 24 frames-per-second rates

In some cases, during the transfer of material from film to another storage media, the MPEG coding process may remove what is known as the “3 / 2 pull-down sequence” (60 Hz countries only): in other cases, the material may be transferred in its native 24-frame mode. 24-frame material received by an end user could therefore be integrated into 60-field or 30 / 60-frame material. It is the feeling of the Task Force that 24-frame material should be converted to the native frame-rate of the facility performing the processing. In some cases, it will be appropriate to tunnel the material through a 60 field / s converter so that it might again be used at 24 frames per second, such as for the re-purposing of 24-frame material to 25 / 50 frames per second

3.6. Interfaces

ITU-R BT.601 is the default method of interfacing. However, as network interfaces become available with the required, guaranteed, bandwidth access and functionality, they will allow methods of digital copying between storage devices. Because storage devices can both accept and deliver data representing video in non-real-time, the network should also allow the transfer of files at both faster- and slower-than-real-time for greater flexibility. The network interface should allow the options of Variable Bit-Rate (VBR or constant quality) and Constant Bit-Rate (CBR) at different transfer bit-rates and, optionally, the transfer of specialized bitstreams that are optimized for stunt modes. This will allow a downstream device to copy a file directly from a primary device for stunt-mode replay on the secondary device.

3.6.1. Status of interfaces for MPEG-2 and for DV / DV-based compression

As of July 1998, the only published standard is for the mapping of 25 Mbit / s DV onto IEEE 1394. There is, however, much work in progress on the mapping of MPEG-2 and DV-based compression schemes onto SDTI, Fibre Channel, ATM, and satellite and telco interconnects. *Section 5.7.3.* contains an additional table with annotations which describes the projects in progress, and highlights the areas needing urgent attention.

3.7. Storage

Where a compressed video bitstream is stored and accessed on a storage medium, there may be storage and compression attributes required of the storage medium, depending on the intended application.

3.7.1. Data-rate requirements

Where possible, users would prefer to record incoming data directly as files on a data-storage device, rather than decoding and re-encoding for storage. As there will be different compressed video bit-rates depending on the application, any network connection to the device should be capable of a wide variety of input and output data-rates.

Both a tape streaming device and a disc-based video server will need to be able to store VBR-compressed video streams. This will require an interface that can accommodate the requirements of a VBR data stream.

Furthermore, compressed video streams may be stored on a tape streamer or disc server with each stream recorded at a different average bit-rate.

3.7.2. Resource management

A tape streamer needs to be able to accept and present compressed video files over a range of values. An integrated system will need to know how to control the streaming device for an I/O channel which may have a programmable data-rate rather than a constant data-rate.

The storage devices should specify the range of data-rates which can be recorded and played back. A disk-based video server additionally has the capability of accepting multiple I/O channels. Further signalling may be necessary to ensure that both the channel bandwidth and the number of channels can be adequately signalled to the system.

3.7.3. Audio, Video and Metadata synchronization

Many storage devices may record Video data, Audio data and Metadata on different parts of the media or on separate media for various reasons. Synchronization information should be included to facilitate proper timing of the reconstructed data at normal playback speed.

3.7.4. VTR emulation

Where a storage device which uses compressed video is intended to be, or to mimic, a VTR, it may implement VTR stunt modes. Such stunt modes may include: viewing in shuttle mode for the purpose of identifying the Content; pictures in jog mode and slow-motion for the purpose of identifying the editing points, as well as broadcast-quality slow-motion. However, the removal of redundancy from the video signal by compression will naturally reduce the possibilities for high-quality stunt-mode reproduction.

Compression methods and parameters must allow stunt-mode capability where required in the user's application. If the recording device is required to reconfigure the data onto the recording media to provide better stunt-mode functionality, such conversion should be transparent and should not impose any conversion loss.

3.8. Interoperability

Interoperability can be a confusing term because it has different meanings in different fields of work. Compression systems further confuse the meaning of interoperability because of the issues of programme transfers, concatenation, cascading, encoding and decoding quality, and compliance testing. Programme exchange requires interoperability at three levels: the physical level, the protocols used and the compression characteristics. This section considers only compression while other sections address the physical layer and protocols.

Considering programme transfers, the Task Force has identified that there are several types of interoperability. The first example identified is interoperation through ITU-R BT.601 by decoding the compressed signals to a raster and re-encoding them. This is the current default method and is well understood. Additional methods of interoperation are expected to be identified in the future. Further work is required:

- to categorize the methods of interoperation;
- to explore their characteristics and relate them to various applications;
- to minimize the possible constraints on device and system characteristics;
- to ensure predictable levels of performance sought by users for specific applications.

3.9. Compliance testing

Interoperability between compressed video products is essential to the successful implementation of systems using compression. Although interoperation is possible via ITU-R BT.601, it is desirable to have interoperation at the compressed level to minimize concatenation losses. Compressed interoperation can involve encoders and decoders using the same compression method and parameters, the same compression method with different parameters, or even different compression methods. Compliance testing is a fundamental step towards ensuring proper interoperability.

Compliance testing can be employed by manufacturers and users of compression systems in a variety of ways. Encoders can be tested to verify that they produce valid bitstreams. Decoders can be tested to verify that a range of compliant bitstreams can be properly decoded. Applications can be tested to verify that the characteristics of a given bitstream meet the application requirements; for example, whether the amount of data used to code a picture is within specified limits. In practice, defining and generating the compliance tests is more involved than applying those tests, so the tests employed by manufacturers might be identical to those employed by the users.

In the case of MPEG-2, compliance testing focuses on the bitstream attributes without physical compliance testing, since MPEG-2 does not assume a particular physical layer. A number of standardized tests are described in ISO / IEC 13818-4. The concepts for tests specified in the MPEG-2 documents may be extended to other compression methods, including Motion JPEG and DV. These compliance tests include elementary streams, transport streams, programme streams, timing accuracy tests, video bitstream tests, and audio bitstream tests. The MPEG-2 video bitstream tests include a number of tests specific to the 4:2:2 Profile at Main Level. The development of additional test methods is necessary.

3.9.1. Test equipment

Test equipment is becoming available on the market which allows conformance testing in accordance with the relevant standard specifications of all the system modules.

3.10. HDTV issues

Within the complicated framework of the technical requirements for reliable interoperation of all components involved in television programme production, compression is the cornerstone. It creates the greatest impact on:

- the technical quality of the broadcaster's assets;
- the cost of ownership of the equipment used;

- the operational efficiency that will be achieved in future fully-automated, networked, television production facilities.

Harmonizing all the technical components involved, in order to achieve the above aims, has already proven to be a complex and difficult enterprise for Standard Television applications.

The future coexistence of Standard Television and High Definition Television – operating at very high data-rates within a range of different pixel rasters and frame-rates – will add yet another layer of complexity.

The Task Force therefore strongly recommends that:

- The compression algorithm and transport schemes adopted for HDTV should be based on Open Standards. This implies availability of the Intellectual Property Rights (IPRs) necessary to implement those standards to all interested parties on a fair and equitable basis.
- The number of compression methods and parameters should be minimized for each uniquely-defined application in order to maximize the compatibility and interoperability.
- A single compression scheme used with different compression parameters throughout the chain should be decodable by a single decoder.
- The compression strategy chosen for High Definition Television should allow interoperation with Standard Television applications.

3.11. Audio compression

The Task Force, in its consideration of a technical infrastructure for digital television, has not specifically considered audio issues such as compression schemes. It is understood that digital audio signals, compliant with existing and emerging standards, will be interfaced with, and packaged within, the communication and storage structures discussed elsewhere in this document.

3.11.1. Studio operations

In general, it is expected that studio production and post-production processes, as opposed to contribution and distribution processes, will use linear PCM audio coding to benefit from its simplicity and signal integrity over multiple generations. This will follow existing basic standards:

- the sampling rate will normally be 48 kHz (AES5-1984, reaffirmed 1992), locked to the video frame-rate (AES11-1991), with 16, 20 or 24 bits per sample;
- real-time digital audio signals may be carried point-to-point, in pairs, on conventional cables using the AES / EBU scheme (AES-3-1992);
- packetized formats for streaming, such as the proposed SMPTE 302M, would carry similar audio data within a network system.

As one example, the EBU Broadcast Wave Format (BWF) provides a means of storing this data as individual computer files, e.g. for random access (see EBU Technical Standard N22-1997). Other schemes are also possible.

The use of audio data compression will vary in different regions of the world. Local practice will dictate the level and types of compression used. In some cases, a pre-mixed version of a multi-track surround track may be supplied to the origination studio; in other cases, the origination studio will be expected to create the multichannel mix from the original source tracks. It is anticipated in the near-term future that AES-3 streams will be able to be stored on devices either as audio or data. This will enable compressed audio data embedded in the AES-3 stream to be edited as a normal audio digital signal, including read / modify / write, provided the compressed audio data has defined frame boundaries that are the same as the video frame boundaries. This same constraint will ensure that routing switchers can also switch a compressed audio data stream.

The EBU / SMPTE Task Force strongly recommends that the AES-3 data stream be utilized for the carriage of all audio signals, compressed or full bit-rate. In some cases, standards will be required to define the mapping of the data into the AES stream.

3.11.2. Compression issues

3.11.2.1. Multi-stage encoding and decoding

All practical audio data-compression schemes are inherently lossy and depend on psycho-acoustic techniques to identify and remove audibly-redundant information from the transmitted data. While a single encode / decode generation may be subjectively transparent, multiple encoding and decoding processes will tend to degrade the perceived audio quality. The number of encode / decode stages that can be tolerated before quality degrades to an unacceptable extent will depend on the particular coding scheme and its degree of compression.

In general, the greater the degree of compression, the fewer generations are tolerable. It is preferable to imagine a system where a comparatively light compression scheme is used for efficient contribution and distribution (with less risk of coding artefacts), while a higher rate of compression is used for emission. The use of a compression coding history, should be encouraged where the compression scheme for contribution / distribution uses a similar tool kit to the emission format. It is important to note that system end-to-end gain in compressed systems must be maintained at unity in order not to introduce any codec artefacts.

3.11.2.2. Coding frame structure

Audio compression schemes tend to have a frame structure that is similar to that of video schemes, with each frame occupying some tens of milliseconds. It is usually possible to edit or switch compressed audio streams at frame boundaries. In some cases, the audio compression frame-rate will be different to its associated video frame-rate; this should be avoided if possible, or suitable precautions should be taken.

3.11.2.3. Latency issues

The processes of encoding and decoding an audio signal with any data compression scheme takes a finite amount of time, typically many tens of milliseconds. An audio process that requires the signal to be decoded will need to be balanced by a similar delay in the associated video to preserve sound-to-picture synchronization. Recommendations in other areas of this report deal with the overall issue of latency; in many instances, the use of SMPTE timecode is suggested as a studio Time Stamp.

3.11.2.4. Mixing

In order to be mixed directly, encoded signals require sophisticated processing. Adding an announcer to an encoded programme requires that the stream to be decoded is mixed with the voice signal; the composite signal is then re-encoded for onward transmission. This process incurs both time-slippage through coding delay, and potential quality loss depending on the type of compression in use and its corresponding coding delay.

An alternative possibility is that the compressed contribution / distribution data stream allows for centre channel sound to be embedded into the compressed data stream without any degradation of the compressed data.

It may be that programme signals compressed at distribution levels can be decoded and processed before re-encoding to a different scheme for emission, without undue loss of quality and with acceptable management of timing.

3.11.3. Use of Audio compression

Each application of audio data compression will define an appropriate balance between complexity, delay, ruggedness and high quality. The criteria affecting this balance vary for different points in the process and, for this reason, it is likely that different compression schemes will need to coexist within the overall system. In some cases it may be possible to transcode from one scheme to another; in other cases this will not be possible. The studio standard of linear PCM provides a common interface in these cases.

3.11.3.1. Contribution

The definition of a contribution-quality signal will vary from region to region, depending upon local and regulatory requirements.

In cases where vital audio material must be communicated over a link of limited bandwidth, audio data compression may be necessary. Compression decoding and, possibly, sample-rate conversion may be required before the signal can be edited and mixed. In some cases the compressed audio mix of a multichannel sound track will be distributed. The degree of usage within the contribution / distribution networks will be regionally dependent. Clearly, the control of overall programme-to-programme level (loudness) must be possible.

3.11.3.2. Distribution

A finished programme will need to be distributed to transmission facilities or to other broadcasters. The need here is for efficient use of bandwidth while retaining quality, and the flexibility for further packaging operations which require switching, editing and mixing.

3.11.3.3. Emission

Broadcast emission requires the minimum use of bandwidth with maximum ruggedness for a defined signal quality. It is expected that coding at this level will only be decoded once, by the consumer, so issues of multiple coding generations are less significant. End users should be aware of concatenation effects when re-purposing off-air signals.

3.11.3.4. Archiving

Dependent on local operational needs, archives may need to store audio data from any of the cases mentioned above.

3.11.4. Channel requirements

For many applications, sound origination will continue in the form of mono and stereo recorded elements. In movie releases and special events programming, the 5.1 surround sound format will exist.

3.11.4.1. Mono

This is the most elementary audio component and it is fundamental within many audio applications for both television and radio.

3.11.4.2. Stereo

The industry is well accustomed to handling two-channel signals. By default, these will represent conventional Left / Right stereo; in some cases, the composite signals (Left / Right) from a matrix surround sound process will be used as direct equivalents.

Exceptionally, a two-channel signal will contain the Sum / Difference transform, or “MS” stereo signal. This can be a convenient format for origination and is intended to be used in post-production to create a conventional Left / Right product. Although these are common practices in traditional “analogue” television, many of these practices can be expected to continue. A word of caution needs to be expressed. In the new DTV world, there will be Content existing in both stereo and surround sound formats. Operational practices will need to be developed to deal with transitions from one format to another.

3.11.4.3. The 5.1 format

All of the emergent television emission standards will be able to present surround sound with five discrete full-range channels, plus the option of a dedicated low-frequency channel (also known as the 5.1 format). Operational considerations within a practical TV environment suggest that this six-channel bundle should be handled as a single entity, rather than as a set of independent mono or stereo signals. Often these six signals will also be associated with other channels. For example, the extra channels could represent different language commentaries, or a separate mono mix, or a descriptive channel for viewers with visual impairments. A bundle of eight channels is therefore proposed in order to carry the complete set in a pre-arranged order.

It is also the case that existing stereo or mono emission services may need to obtain their audio signals from such distribution feeds. In this case, it is important that linear PCM be preserved as far down the distribution chain as possible, in order to minimize the impairments caused by the concatenation of different matrix arrangements.

3.11.4.4. Recording of the 5.1 format on professional videotape recorders

Professional videotape recorders handle typically up to four linear PCM audio channels. In order to record the 5.1 format on conventional videotape recorders, it is suggested that a compression scheme with a low compression factor be used in order to obtain transparent audio quality and the possibility of multi-stage encoding and decoding. Some redundancy should be applied to the error protection of the compressed audio signal. The input and output to recording devices should comply with the AES-3 stream. Other changes will be required to the characteristics of the VTR-server recording channel. These changes are being defined by the SMPTE.

The cascading of several encoding and decoding stages is unavoidable in a complete audio and television broadcasting chain, including programme contribution and distribution. Depending upon regional choice for the audio emission standard, this may well drive the choice of contribution and distribution compression.

3.12. Compression issues – recommendations and current status

The deliberations that went into the preparation of this section led to the following recommendations for the application of compression within television production, contribution and distribution. The bold text following each recommendation indicates the known current (July 1998) status of the issue:

1. Compression algorithms and transport schemes should be based on Open Standards. This implies the availability of the IPRs necessary to implement those standards to all interested parties on a fair and equitable basis. Availability in the marketplace of chip-sets and / or algorithms for software encoding and decoding may give users confidence in the adoption of particular compression methods.	**The necessary standards are either completed or are in progress. Chip-sets from a variety of suppliers are readily available under the above conditions for both DV and MPEG compression.** **Dolby audio compression is also available under licence.**
2. The number of compression methods and parameters should be minimized for each uniquely-defined application in order to maximize compatibility and interoperability.	**Two compression families have been identified to meet the requirements of next-generation network television production.** **For audio, many compression choices exist. Users should be concerned with concatenation effects.**
3. Compliance-testing methods should be available for those building the equipment to standards for algorithms and transport schemes, and for users purchasing and installing equipment to those standards. Standards bodies should adopt standards for compliance-testing methods to support both the manufacturers' and users' needs.	**Compression compliance test work has been done, but transport compliance testing still needs to be done.**
4. A single compression scheme, used with different compression parameters throughout the chain, should be decodable by a single decoder.	**An integrated (economical) intra-family agile decoder will be available for each compression family.** **For audio, this may not be possible.**

5. To support the use of more than one compression family, the development of a common ("agile") decoder is desirable.	**No manufacturer has officially stated a commitment to providing a common agile decoder that would allow decoding of more than one compression family.**
6. Integration of video compression into more complex systems must be via standardized interfaces. Translation through ITU-R BT.601 (i.e. decoding and re-encoding) is the default method of concatenating video signals that have been compressed using different techniques and / or parameters, although other methods are possible. For audio, it is recommended that the AES-3 stream and its associated interfaces should be used for the transport. Full bit-rate audio should be used when necessary to bridge the compression scheme.	**The necessary interface standards are either completed or in progress. Concatenation methods between different compression families remain to be identified.**
7. The compression scheme chosen should not preclude the use of infrastructures based on the serial digital interface (SDI) as embodied in SMPTE 259M and ITU-R BT.656.	**The transport mechanism for compressed signals defined in SMPTE 305M fully complies with the above requirement.** **AES-3 streams should be used to transport audio data streams.**
8. Issues relating to interoperability must be further explored, and standards developed, to allow predictable levels of performance to be achieved in the implementation of specific applications.	**Additional standards and / or recommended practices need to be developed to achieve the above objectives.**
9. Bitstreams carrying compressed signals should be designed so that they can be formatted and packaged for transport over as many types of communications circuits and networks as possible.	**Standards and / or recommended practices for applications currently identified have been completed or are in progress.**
10. Compressed bitstreams are very sensitive to errors, and therefore it is recommended that appropriate channel-coding methods and error protection be employed where necessary.	**Both tape and hard-disk technologies employ these technologies. Public network carriers should provide the necessary ECC strategy to meet application needs.**
11. Compression systems should be designed so that, in normal operation, signal timing relationships (e.g. audio / video lip sync) and synchronization presented at the encoder inputs are reproduced at the decoder outputs.	**Both compression families provide the necessary tools to maintain audio / video lip sync within the limits prescribed for broadcast applications.** **Compressed audio streams should carry an SMPTE timecode to act as a Time Stamp**
12. Signal delays through compression processing (encoding / decoding) must be limited to durations that are practical for specific applications, e.g. live interview situations.	**This will continue to be an issue for live interview situations. It can be controlled for non-live or storage applications.**
13. Provision should be made for selected analogue VBI information to be carried through the compression system, although not necessarily compressed with the video. Additionally, selected parts of the ancillary data space of digital signals may carry data (e.g. Metadata), and provision should be made to carry selected parts of this data through a transparent path, synchronously with the Video and Audio data.	**Some implementations available on the market today provide a limited space for transparent throughput of data. Standards and / or recommended practices for different bearers need to be developed.**
14. The compression scheme chosen for devices that mimic VTRs should allow for (i) the reproduction of pictures in shuttle mode for identifying the Content and (ii) of pictures in jog and slow-motion modes for selecting the edit points.	**Both compression families allow these two possibilities.**
15. Network interfaces and storage devices should provide for both Variable Bit-Rate (VBR) and Constant Bit-Rate (CBR) options, and must be capable of supporting a wide variety of data-rates as required by particular applications.	**Nothing has been done to preclude this as far as networking and storage devices are concerned.** **However current implementations of storage devices may not allow this flexibility.**
16. Storage devices should allow recording and playback of streams and files as data rather than decoding to baseband for recording and re-encoding upon playback.	**Existing hardware implementations have already taken this into account.**
17. The compression strategy chosen for Standard Definition Television should be extensible to High Definition applications, to allow for commonality in the transitional phase.	**Nothing has been done to preclude this in the two compression families.**

Section 4

Wrappers and Metadata

4.1. Introduction

This section of the report details the findings of the Task Force's Sub-Group on Wrappers and Metadata.

The initial discussions of the Sub-Group from September 1996 led to the preparation of the Wrappers and Metadata section of the Task Force's First Report on User Requirements in April 1997. That section included some tutorial information defining the terminology and structure of Content within a Wrapper. It also contained a list of Requirements for Wrapper formats, with explanations and descriptions, some tutorial annexes, and several specific recommendations for standardization.

The First Report was followed by a section within a Request for Technology (RFT) which was published in June 1997. Several responses were received, covering some aspects of the required technology. These responses were analyzed during succeeding meetings, along with comparisons of existing practices in the industry and discussions on the standards development efforts which have been continuing simultaneously.

A second RFT was issued in January 1998, seeking a low-level persistent storage mechanism for the storage of complex Content, to be placed entirely in the public domain. A single but complete response was received. It is recommended that this response is used as the basis for standardization of this portion of the Requirement.

The Sub-Group also received a number of other documents which provided information not directly related to either of the RFTs.

4.2. Purpose of Wrappers

The fundamental purposes of a Wrapper are (i) to gather together programme material and related information (both by inclusion and by reference to material stored elsewhere) and (ii) to identify the pieces of information and thus facilitate the placing of information into the Wrapper, the retrieval of information from the Wrapper, and the management of transactions involving the information.

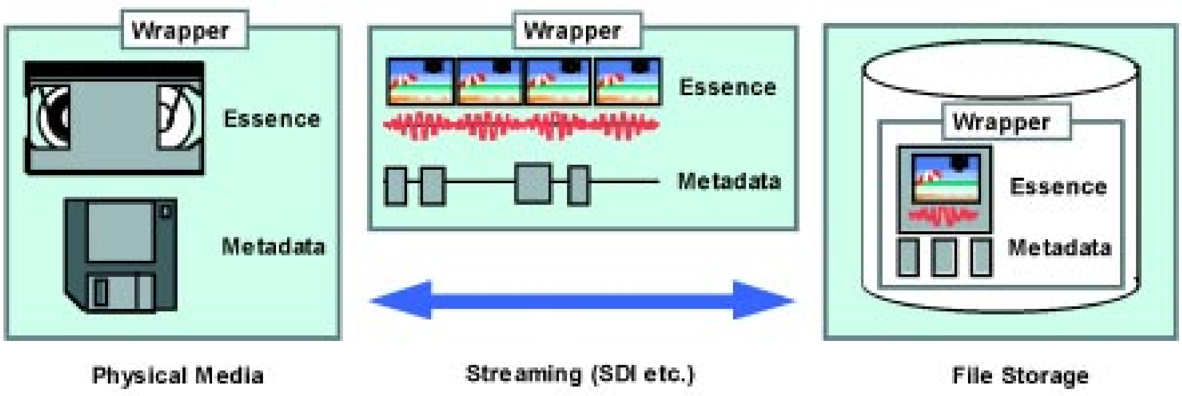


Figure 4.1: Schematic view of Wrappers in use.

4.3. Overall concepts – terminology and structure

4.3.1. General

Wrappers are intended for use in linking physical media together, for streaming of programme material across interconnects, and to store programme material in file systems and on servers. This and other terminology is discussed in this section.

The Sub-Group adopted terminology and structure very close to that defined by the Digital Audio-Video Council (DAVIC) in its Specification V1.2 – the section on Content Packaging and Metadata (baseline document 22) – as follows:

- programme material and related information of any variety is called **Content**;
- the parts of Content which directly represent programme material (such as signal samples) are called **Essence**;
- the parts which describe the Essence and other aspects of the material are called **Metadata**;
- Content is composed of **Content Packages**, which in turn are composed of **Content Items**, which are further composed of **Content Elements**;
- Content is contained in **Wrappers** which must be capable of including Essence, Metadata and other Overhead in differing proportions and amounts, depending upon the exact usage of each Wrapper.

4.3.2. Storage, Streaming, File Transfer and editing of Content

When Content is gathered and placed onto tape or disk for later access, it is kept in a **Storage Wrapper**. The basic form of these Storage Wrappers can be very simple and generic as will be described later in this document.

Applications for Wrappers will include those where Content is primarily to be transferred between origination, display, transmission and storage devices, using signal interconnects such as SDTI or network interconnects such as Ethernet or Fibre Channel. These are referred to as **Streaming** applications, and the Wrappers used are referred to as **Streaming Wrappers**.

These differ from Storage Wrappers in two respects:

- some Overhead may be added to the stream to assist or guide the interconnect layer;
- the specific interleaving and multiplexing structure that are used may be optimized for the interconnect.

Applications will also include those where Content is intended to be randomly accessed, searched, modified and browsed – referred to as **Editing** applications. In these latter applications, the collections of Content are called Content Packages. They will often remain simple, but may also become very complex and contain many inter-relationships.

Editing applications may need to exchange either Content Items or Content Packages with each other; they will also need to exchange Content Items with Streaming applications. For exchanges between Editing applications and Streaming applications (e.g. playout), Streaming Wrappers will be used. For exchanges using File Transfer methods, Storage Wrappers can be used. For exchange between Editing applications resident on the same computer, or sharing the same local or network file system, it is not necessary to use any Wrapper.

While Content is in use by Editing applications, Content Packages will be subject to frequent change – as a result of adding and removing Content Items and Content Elements, and by modification of the Essence and Metadata throughout the Content Package.

When there are constraints on the variety of Content within a Content Package, and there is foreknowledge of the extent of likely modification to the Package, provision can be made within a generic Wrapper to accommodate any growth of the Content as it passes from application to application.

In other cases, to promote efficiency, it may be necessary to store the Content Package in a format which provides for flexible growth and rearrangement of the data. This special-purpose storage format may also provide advanced functions for the management of the Content Package.

Some possible application combinations and Wrapper choices are summarized in the following table (*Table 4.1*):

Table 4.1: Use of Wrappers for Streaming and Editing applications.

Source: → / Destination: ↓	Streaming application	Editing application
Streaming application	Streaming Wrapper	Streaming Wrapper
Remote Editing application via Streaming	Streaming Wrapper	Streaming Wrapper
Remote Editing application via File Transfer	N/A	Storage Wrapper
Local Editing application	N/A	Custom or Storage Wrapper

Note: For most interchange scenarios, a Streaming Wrapper is employed.

4.3.3. Content structure

A Wrapper does more than just contain Content; it also defines and describes the structure of the Content. The microscopic structure of Content is inherent in the Essence itself; the macroscopic structure is built using Metadata and Overhead (see below), and is classified as described here.

Each individual item, either Essence or Metadata, is called a **Content Component** – for example, a block of audio samples, or a timecode word. A Wrapper contains some number of Content Components, built into a logical structure.

A Content Element (CE) consists only of Essence of a single type, plus any Metadata directly related only to that Essence – for example, the blocks of samples of a video signal plus the **Essential Metadata** which describes the sample structure plus the **Descriptive Metadata** which identifies the origin of the signal.

This can be expressed by the simple equation: Content equals Metadata plus Essence.

An exception to this definition is when a Content Element can be generated entirely from Metadata, without the need for Essence – for example, an encoded subtitle. In these cases, the Metadata either refers to external raw Essence or to a device or algorithm which creates Essence.

Types of Essence include **Video**, **Audio** and **Data** of various kinds, including captions, graphics, still images, text, enhancements and other data as needed by each application.

- ➪ A **Content Item** (CI) consists of a collection of one or more Content Elements, plus any Metadata directly related to the Content Item itself or required to associate together the component parts (Content Elements) – for example, a Video clip.
- ➪ A **Content Package** (CP) consists of a collection of one or more Content Items or Content Elements, plus any Metadata directly related to the Content Package itself, or required to associate together the component parts (Content Items and Content Elements) – for example, a programme composed of Video plus Audio plus subtitles plus description.

An example of a such a collection of Content Items and Content Packages contained in a Wrapper is shown in *Fig. 4.2.*

Although these terms describe larger and larger structures of Content, the smaller structures do not have to be fully contained within bigger ones. For example, a single Wrapper could contain Content Elements equal to a full hour of programme source material, and Content Packages describing only a couple of five-minute segments within the material.

Thus, a Wrapper is not restricted to contain any specific quantity or portion of any of these constructs – it may contain only a few Content Components, or as much as several Content Packages. When a Wrapper contains one or more Content Packages in combination with other Content Items, it becomes a **Complex Content Package** as illustrated in *Fig. 4.3*.

Besides using a single Wrapper, two or more Wrappers may be used to transport components of a single Content Item or Content Package where separate transport mechanisms are used. In this case, each of the Wrappers will contain a partial common set of Metadata to allow the Wrappers to be cross referenced. This is the mechanism used where not all of the Metadata can be accommodated in the transport used for the Essence.

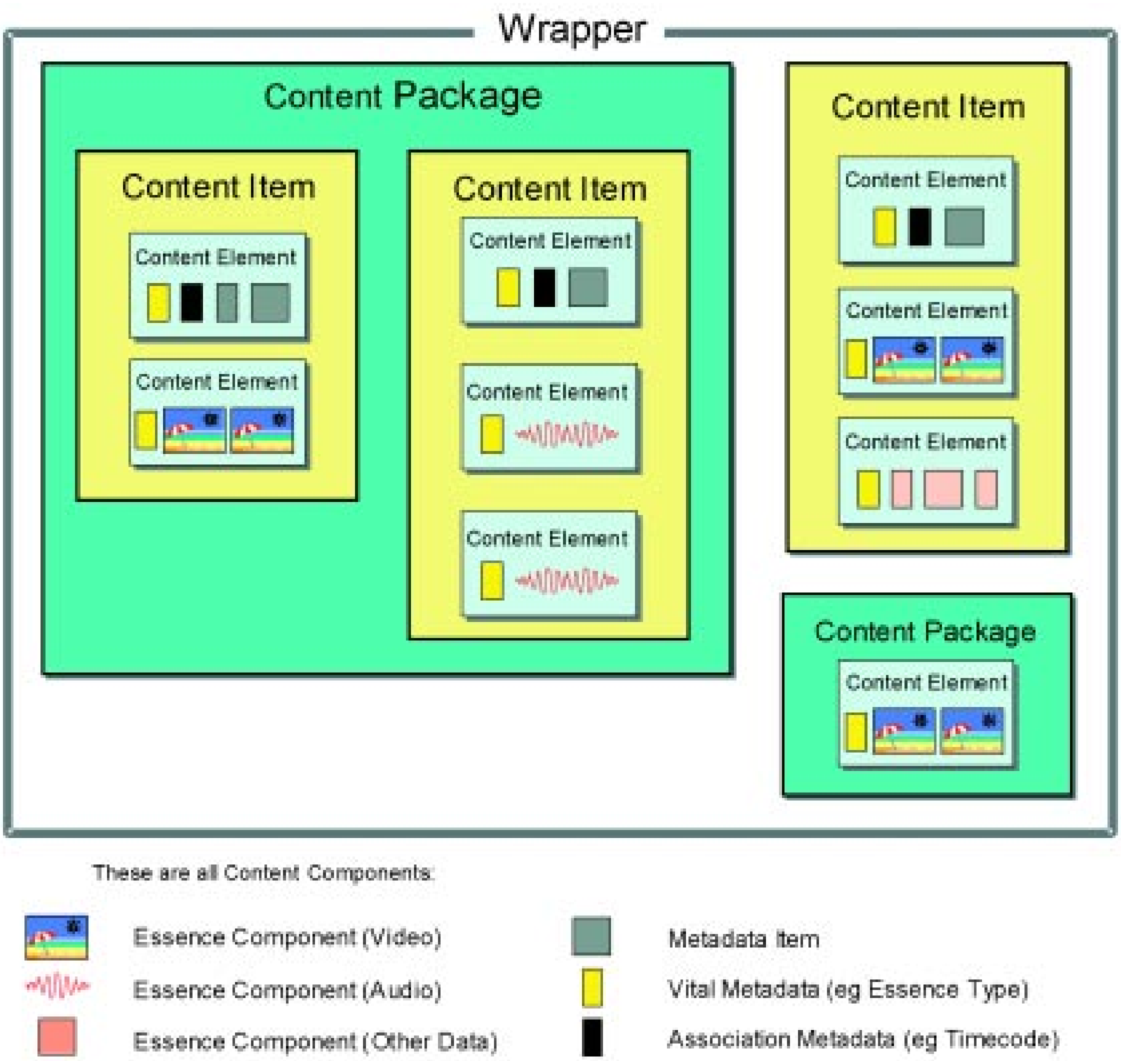


Figure 4.2: Content Package: Packages, Items, Elements and Components.

4.3.4. Essence

Raw programme material itself is referred to as Essence. Essence is the data that represents pictures, sound and text; types of Essence include Video, Audio and Data of various kinds, including captions, graphics, still images, text, enhancements and other data as needed by each application.

Some varieties of Essence may be treated as Metadata in certain circumstances. For example, a sequence of captions should be regarded as a kind of **Data Essence** when it is inserted into a broadcast television signal but, within an Asset Management system, the same captions may be used to index and describe the Content, and should be regarded as Descriptive Metadata. This dual perspective is also presented from the point of view of Metadata in the next section.

The task of interpreting Essence as Metadata may be carried out in an extraction step which results in duplication of the information, or it may be a dynamic process in which the same information is used in either sense upon demand. Additionally, the specifications of some Essence formats include embedded items of Metadata which will need to be extracted and promoted into separate Metadata items for efficient systems operation.

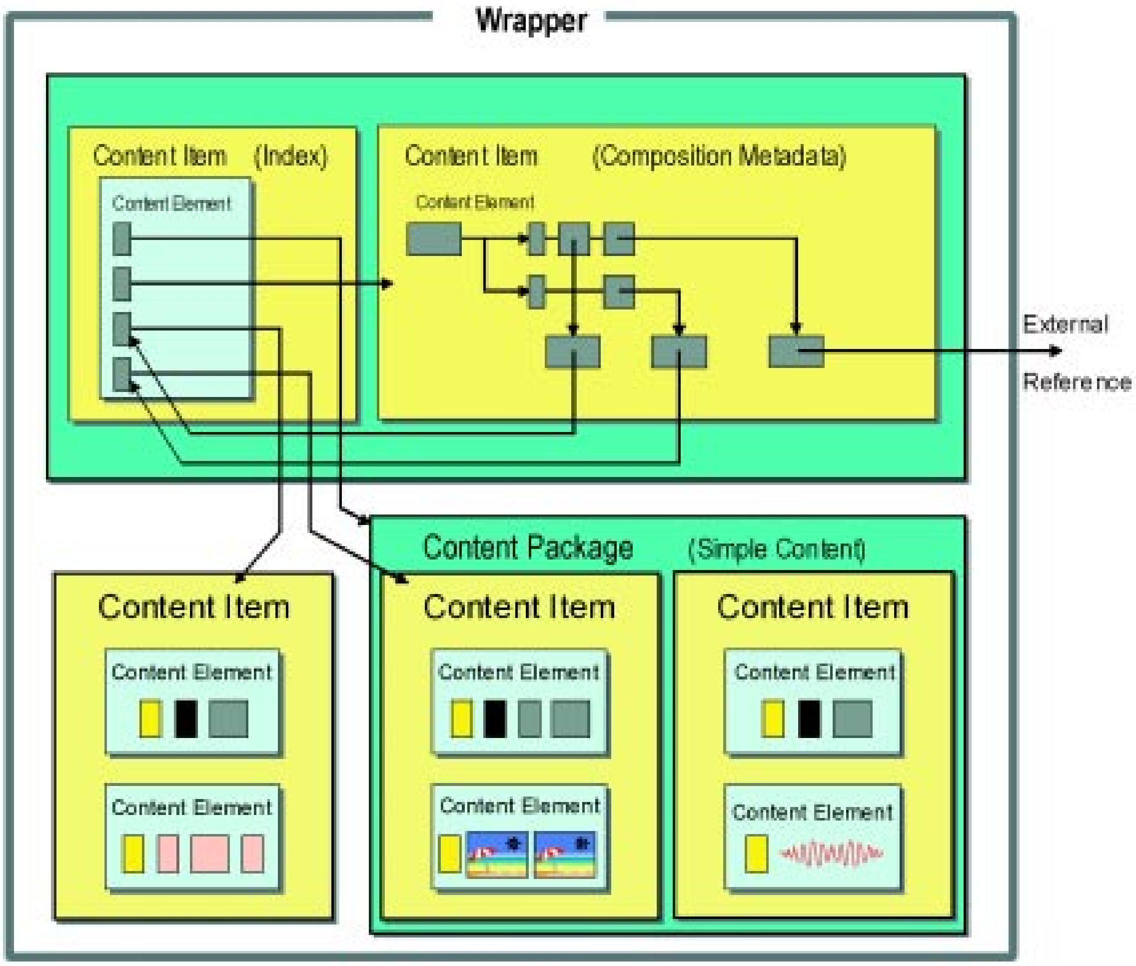


Figure 4.3: Complex Content Package: Packages, Items, Elements, Components and Metadata.

Essence may be encoded or compressed in whatever way is appropriate. It is typically structured in packets, blocks, frames or other groups, which are collectively called **Essence Components**. The microscopic structure of Essence Components depends on the particular encoding scheme used, which in turn is identified by Essential Metadata (see below).

Essence typically has the characteristic of a stream; it provides sequential access whether stored on a file device or a streaming device. Stream data will normally be presented in a sequential time-dependent manner. Essence stored on a file storage device can be randomly accessible. Essence not having the characteristic of a stream (e.g. graphics, captions, text) may still be presented in a sequential time-dependent manner.

4.3.5. Metadata

Other information in the Content is referred to as **Metadata**. Metadata is broadly defined as “data about data”.

The number of distinct varieties of Metadata is potentially limitless. To assist with describing its requirements and behaviour, Metadata is divided into several *categories*, depending upon its purpose, including at least the following:

⇨ **Essential** – any information necessary to decode the Essence. Examples: Unique Material Identifiers (UMIDs), video formats, audio formats, numbers of audio channels.

- **Access** – information used to provide and control access to the Essence. Examples: copyright information, access rights information.
- **Parametric** – information which defines detailed parameters of the Essence. Examples: camera set-up, pan & scan, colorimetry type.
- **Composition** – required information on how to combine a number of other components (e.g. video clips) into a sequence or structure (Content Element, Content Item or Content Package). This may equally be regarded as information recording the heritage or derivation of the Content. Examples: Edit Decision Lists (EDLs), titling information, zoom lens positioning (for virtual studio use), transfer lists, colour correction parameters.
- **Relational** – any information necessary to achieve synchronization between different Content Components, and to achieve appropriate interleaving of the components. Examples: timecode, MPEG SI.
- **Geospatial** – information related to the position of the source.
- **Descriptive** – all information used in the cataloguing, search, retrieval and administration of Content. Examples: labels, author, location, origination date & time, version information, transaction records, etc.
- **Other** – anything not included above. Examples: scripts, definitions of the names and formats of other Metadata, user-defined Metadata.

Within each category, Metadata may be further divided into sub-categories.

As noted in the previous section, some varieties of Metadata may be treated as Essence in certain circumstances. For example, within an Asset Management system, a sequence of key phrases may be used to index and describe the Content, and should be regarded as Descriptive Metadata; but if the same text is converted into captions and inserted into a broadcast television signal, it should be regarded as a kind of Data Essence. The task of interpreting the Captions as Data Essence may be carried out in a rendering step, resulting in duplication of the information, or it may be a dynamic process in which the same information is used in either sense, upon demand.

4.3.6. Metadata characteristics

Metadata which is absolutely necessary for the operation of systems is referred to as **Vital Metadata**. The system must provide values for this kind of Metadata. The specific set of Vital Metadata may be different for each application; but it always includes at least the Essential Metadata. Any item of Metadata may be Vital, independent of its position within the Metadata class hierarchy.

It was realized that the core set of Vital Metadata is relatively small, consisting of the Essential Metadata (UMID and the basic Essence type), and some Relational Metadata. Specific applications may require additional items of Vital Metadata (for example, commercial playout applications require the use of an identifier or ISCI number).

Metadata which is related to the whole of a subsection of the Content (for example a Content Item or Content Package) is referred to as **Static Metadata**.

Metadata which is related to a subsection of the Content (e.g. a single Content Component, a Content Element, or a frame or scene) is referred to as **Variant Metadata**. The variation will frequently be connected to the timing of the Content, but may also be associated with other indexing of the Content. Most categories of Metadata may be Variant.

Further such characteristics of Metadata items have already been identified, while others may well be identified in the future.

During a Streaming transfer, there will be certain types of Metadata that require periodic repetition. This will allow the user to enter the stream asynchronously and, within a given time period, recover the Metadata as determined by the repetition rate. This is called **Repeated Metadata**. It is discussed further below in *Section 4.7.6*.

It might normally be assumed that all Essence and Metadata should be preserved as far as possible throughout the broadcast chain. However, some Metadata types may intentionally be destroyed after they have served their useful purpose. Examples of **Transient Metadata** (with “short” permanence) include Machine Control, QoS Control, Error Management and Encoder Control. Examples of **Permanent Metadata** include Essential Metadata such as Unique Material Identifiers, and timecode.

Metadata may be kept with the associated Essence or kept elsewhere. Factors contributing to the choice include allocation of available bandwidth, ease of access, and concern for systems reliability and error recovery. **Tightly-Coupled Metadata** (above some priority threshold) will be stored and transmitted with the primary Essence without compromise, and will possibly be duplicated elsewhere, while **Loosely-Coupled Metadata** (below the priority threshold) may be stored and transmitted separately.

These characteristics of Metadata must be recorded in the relevant Standards documents and in the SMPTE Registry as part of the definition of the Metadata items (see *Section 4.5.3.*).

4.3.7. Overhead

In addition, the construction of the Wrappers themselves will require some additional items of data. This data is referred to as **Overhead**. Overhead includes such things as flags, headers, separators, byte counts, checksums and so on.

The composition of Wrappers can be expressed by the following simple relationships:

- ⇨ an empty Wrapper is made up of Overhead only;
- ⇨ a full Wrapper is made up of Content plus Overhead;
- ⇨ Content equals Metadata plus Essence;
- ⇨ Hence, a full Wrapper is made up of Overhead plus Metadata plus Essence.

4.3.8. Metadata Sets

Metadata may be grouped into sets which are referenced as a single item. Such groupings are known as **Metadata Sets**. Such sets can be referenced by a single key value rather than by identification through each individual Metadata item; individual items within the sets may be identified with sub-keys, or perhaps implicitly within fixed formatting of the set.

4.4. General requirements

4.4.1. Wrapper requirements

There are a number of issues from the First Report which are assumed in this Final Report. For completeness, these are highlighted below. Please refer to the First Report for full details of each issue.

4.4.1.1. Size of wrapped Content

In some Usage Profiles, the size of some wrapped Content will undoubtedly exceed the capacity of a single storage volume. Wrapper formats must therefore incorporate a mechanism to allow for dividing them into smaller parts if they become too big.

4.4.1.2. Platform neutrality

Wrapper formats must be designed to be “platform neutral”, so that Wrappers may be read by any machine with equal ease (although perhaps with different performance), no matter what machine was used to originally create the Wrapper.

4.4.1.3. Immutability and generation numbering

In most cases, it will not be known how many references have been made to the Content from other Wrappers.

In these cases, it is important to provide identification of the specific generation number (or version number) of the Content, to avoid one user of the Content affecting another user of the same Content.

4.4.1.4. History

Two types of historical information may be included in the Metadata:

- ➯ **Derivation history information**, which may include any Content used to create the current version of the Content (this type of historical information allows the production process to be reversed or reproduced with or without modification). This type of historical information includes any editing history or signal transformation data.
- ➯ **Transaction logging**, which allows the steps taken to produce the current version of the Content from its source material to be traced but not necessarily reversed. This type of historical information includes version and source information.

4.4.1.5. Support of transactions

Wrappers will be subject to many transactions, both for commercial purposes and in the operation of production systems. These transactions will include copying, moving and modification of Content and their surrounding Wrappers. Metadata in support of these transactions may be included within Wrappers.

4.4.1.6. Property rights

Metadata which records the ownership of Content, and the history of ownership and usage, may be stored in the Wrapper in order to facilitate the establishment and preservation of copyright.

As a minimum requirement, it must be possible to tell from the **Property Rights Metadata** the identity of the Content which is contained in the Wrapper. In cases where more than one item of Content is included, multiple instances of the Property Rights Metadata must be permitted.

However, there are many additional requirements for a fully-functioning Property Rights protection system, which are described in “Annex D, Section D3 – Access Control and Copyright” of the First Report. Although not all these functions and capabilities are required in all studio systems, the increasing use of permanently-connected networks implies that protection of Property Rights should be considered in the design of every system.

4.4.1.7. Compatibility and conversion

Wrapper formats must be Compatible with existing formats, including formats for Essence (however stored or transported) and formats for Metadata. In addition, the use of Wrappers must be compatible with established working practices.

It is recognized, however, that when existing Essence and Metadata formats are included within programme material, some of the benefits to be obtained from new Wrapper formats may not be available.

- ➯ A format is **Compatible** with a Wrapper format when Metadata or Essence can be directly placed in a Wrapper from the source format or directly exported from a Wrapper.
- ➯ **Lossless Conversion** is possible when Metadata or Essence cannot be used directly but can be translated into or out of the Wrapper with some processing, and the conversion can be fully reversed.
- ➯ **Lossy Conversion** is possible when Metadata or Essence cannot be used directly but can be translated into or out of the Wrapper with some processing and some loss of meaning or quality, and the conversion cannot be fully reversed.

Users will require Lossless Conversion or better in all cases, except where Content from outside a Wrapper is involved; in this case, users will require Lossy Conversion or better.

4.4.1.8. Extensibility

Any new Wrapper format to be developed is required to be standardized and to have reasonable longevity of decades or more. It is certain that new Metadata types and Essence formats will be required within the life of any standards document. Therefore, every Wrapper format is required to be extensible in the following ways:

- ➩ by the addition of new Essence and Metadata types;
- ➩ by the extension or alteration of data syntax and semantics.

To achieve maximum backwards compatibility, the addition of new Essence and Metadata types must be achieved without change to the underlying Wrapper data syntax; efficient and complete documentation must be provided to ensure that any extensions are equally accessible to all implementations. This will depend upon maintenance of a proper Registry of Data Identifiers.

When unknown identifiers are encountered in the processing of a Wrapper, they (and any attendant data) should be ignored gracefully.

4.4.1.9. Other requirements

Wrappers must be capable of including Essence, Metadata and Overhead in differing proportions and amounts, depending upon the exact Usage Profile of each Wrapper.

For example, a programme replayed from videotape might include video, audio and ancillary data streams, with almost no Metadata; an Edit Decision List might include Descriptive and Composition Metadata, but little or no Essence. Each particular variety of Wrapper will contain a minimum defined level of Essence, Metadata and Overhead.

Wrappers must be capable of including various structures which are combinations of Essence and Metadata, such as the Content Elements, Content Items or Content Packages defined above.

Metadata may be contained in a video or audio data stream (e.g. MPEG or AES-3 streams) but, for ease of access, it could be replicated in a separate Metadata area. Real-time live transfer by streams may require the repeating of Metadata and the interleaving of structures.

As well as directly including Essence and Metadata, Wrappers may contain indirect references to either. This will be discussed later in this document.

4.4.2. APIs

The complexity of management of the Content in the face of all these requirements creates the requirement for an Application Programming Interface (API) to be generally available, at least for the management of Complex Content Packages and probably even for manipulation of Streaming Wrappers. The API should provide functions for locating elements within the file, for reading and writing Essence, Metadata and complete Content elements, and for maintaining the integrity of the data structures and indexes within the file.

While the work of the Sub-Group has been focused at the file level, it has also discussed the need for prototype or reference APIs to speed the adoption of the file format. These APIs should be the basis of future standardization by the SMPTE or other standards bodies.

It is most important for standards bodies to document the formats themselves in the first instance; APIs are inherently more difficult to standardize and such work should be tackled at the same time as implementation of the standard formats.

General guidelines for the reference API are as follows:

- ➩ **Hierarchical API** – it is envisioned that the API and its standardization will occur at multiple levels. At the lowest level, the API will handle core data management. The next level will include facilities that enable access to named objects / parameter-value pairs. Finally, application-level interfaces will provide higher-level functions such as traversing a timeline.
- ➩ **Multi-platform support** – while not mandatory, it is desirable to have implementations of the API that operate on more than one platform.
- ➩ **Full support of the file format** – the API should provide facilities for complete access to all the data included in the file format.
- ➩ **API documentation** – for the API to be useful and be considered for future standardization work, complete and clear documentation is required. Such documentation should be sufficiently complete for anyone to re-implement the API from scratch.

- **Sample code** – a sample application using the provided API will help greatly in showing accepted practice for the use of the API. The sample application does not need to be complex but should implement basic operations on the file format.

4.4.3. Breadth of application and Wrapper profiles

Users would strongly prefer one solution to cover the widest range of applications.

Because of the limitations of technology and the concerns listed below, it is unlikely that a single Wrapper format will fit all applications. However, if multiple formats are to be developed, they must be created with a view to maximum commonality, understanding that programme material may appear in, and be converted between, any or all of the formats during its lifetime.

The range of production processes can be encapsulated into Wrapper Profiles, each calling for one or more of the possible Wrapper formats.

During the study, a wide range of potential activities were listed, which were then grouped into the following categories:

- Pre-production;
- Production and Acquisition;
- Post-production;
- Distribution and Storage;
- Emission and Transmission;
- Archiving.

Every application involves one or more of these processes, and each process makes use of Content in each of several forms:

- Unwrapped (for example, a videotape);
- Streaming Wrapper (for example, on an SDTI channel or as an MPEG stream);
- Storage Wrapper (for file transfer);
- Special Purpose (for example, a database together with signal storage).

These forms reflect the retrieval and editing functionality which is required within each process stage. Furthermore, as well as being used within each process, these forms are all used as interfaces between processes.

There is therefore a requirement for several Wrapper format types (Storage, Streaming and Special Purpose) in addition to the continued use of unwrapped Content.

4.4.4. Framework of a solution

During the Sub-Group's studies, the Recommendations from the First Report were reviewed and extended, and are now as follows:

- The development of an extensible hierarchical classification of Metadata varieties, including the notion of Metadata Sets (templates) appropriate to particular uses.
- The establishment of a single registry of Metadata item names and definitions, including a registry of Essence Formats. Essence Formats are specific values of a particular Metadata item.
- The development of an initial core set of Essence Format specifications.
- The standardization of a single format for a "Unique Material IDentifier (UMID)". It is recognized however that multiple formats are already in use and will continue to be introduced. As a minimum therefore, it should be possible to register existing and new unique identifier formats within the Metadata registry referred to above. We recommend the creation of a new standard to be a target for the industry to converge upon.
- The standardization of a single generic Wrapper for the Streaming of Metadata, which can be mapped onto existing and emerging signal transport layers, and which can be used to stream Content (Metadata

multiplexed with Essence) onto the transport layer. It is understood that there will be many specific solutions for streaming file formats, each one optimized for a given transport layer and application; interoperability can be promoted by allowing the same Metadata to be accessed with equal ease and without translation from each stream.

- The standardization of a single generic Wrapper for the Storage of Content.
- The standardization of a single generic format for representation of Complex Content Packages, for applications requiring arbitrary complexity of Content of all types, including Metadata and Essence. This format must be highly compatible with the Metadata streaming format described above, allowing the same items of Metadata to be accessed when the Content Package is within a stream or within an application. However, it should be noted that Complex Content Packages are usually not expected to be interpreted while they are being moved by a streaming application.

4.5. Metadata requirements

4.5.1. Metadata vision

Many standards development efforts are underway in this area, both within the SMPTE and in other closely-related areas in television and in information systems in general. Metadata is one of the most important emerging technologies of the digital motion imagery (video and audio) age. It is a key enabling technology that will change the very nature of how motion imagery is created, viewed and used.

At its simplest level, Metadata is “data about the data.” It is the data about the motion imagery but it is not the imagery bits themselves. As described above, the imagery bits themselves are called “Essence”.

4.5.2. Requirements for Metadata standards

Metadata standards will be based on the following general approach:

- Specification of a general scheme for the hierarchical naming of items of Metadata, based upon an SMPTE Universal Label (SMPTE 298M). This is explained further in the “Registry” section below.
- Harmonization of diverse Metadata Sets, by defining a common lexicon or dictionary (which defines the place of groups of Metadata items within a hierarchy). This dictionary of defined Metadata items, plus their meaning, purpose and allowed formatting, will be created by reconciling the existing work of bodies such as the “Dublin Core” group, The Library of Congress and others, and by adding Programme-Material-specific items which are known requirements today. Many of the requirements and Metadata types identified in the First Report included: Essential Metadata such as format, basic signal parameters and Unique Material Identifiers; Association Metadata for the synchronizing, referencing and indexing of Content items; Composition or Usage Metadata; Asset Management Metadata, including material generation numbering, and derivation and processing history; Descriptive Metadata including a general description of access and property rights; Access Control Metadata and transactions on Content; and many other types.
- Studies indicate that almost every Metadata type can be grouped easily into one of the following major categories or Metadata Classes previously described:

Class 0		(SMPTE-defined) Transport Stream (or File) Header	
Class 1		Essential Metadata	
Class 2		Access Metadata	
Class 3		Parametric Metadata	
Class 4		Composition Metadata	
	Sub-class 4.x		Heritage Metadata
Class 5		Relational Metadata	
	Sub-class 5.x		Temporal Metadata
Class 6		Geospatial Metadata	
Class 7		Descriptive Metadata	
Class 8		Other Registered Metadata	
	Sub-class 8.x		External Reference Metadata
Class 9		User-defined Metadata	

- ➩ Procedures must be formalized for adding new items into the Metadata dictionary, corresponding to the Class Rules described below.
- ➩ As new items of Metadata are added, particular attention must be paid as to whether they are essential (Class 1), or whether they should be added into the other classes (2~9).
- ➩ A new type of standards document will be taken up by the SMPTE and perhaps similar bodies, called a Dynamic Document. A Dynamic Document is like a regular standards document except that it is used in combination with a public registry of extensions to the standard. The registry is still updated and expanded by due process, but the review and updating cycle can take place much faster than before. This leads to standards which remain more current.

4.5.2.1. Class rules

4.5.2.1.1. Classes 0-7: publicly-defined and due-process standardized (a.k.a. registered)

Consists of due-process standardized specifications. It is generally expected that these types will be known to all systems. In cases where a data type is inappropriate for a particular application, systems should be able to safely identify data types and pass them through the system. It is also expected that any application claiming use of this Registry be up-to-date with all registered data types at the time of release. Ideally, systems should have the ability to import information on new registered data types for the purpose of identification. Registered data types would be reviewed and accepted by the SMPTE and are general-purpose enough that it is not unreasonable for an application to support them all.

4.5.2.1.2. Class 8: registered as private with public description

Class 8 data consists of types that have been well defined but whose detailed description is not publicly available (but can be obtained through proper licensing), or that have been publicly defined but have not yet been accepted as Classes 0-7 types. It is likely that most data types will start as Class 8 and progress to Classes 0-7 over time as they are accepted. Applications may decide to support any number of Class 8 data types. While it is possible for applications to claim full support for all Classes 0-7 data, it is unlikely that any one application will posses the ability to, or have the need to, support all Class 8 data types.

4.5.2.1.3. Class 9: unpublished with no public description (user-defined, proprietary or experimental)

Class 9 data consists of types that are in general private data with a registered identifier. Full specifications are not available. All applications should be able to ignore Class 9 data. Class 9 types may not be exported for interchange.

4.5.3. Registry

The Registry is a central premise of the new paradigm of Wrappers, the Metadata dictionary, and SMPTE "Dynamic Documents". To achieve their full potential, these all require a public repository of registered unique labels and associated descriptions, maintained by due process, with international visibility.

The concept of unique labels was developed by the SMPTE in the years up to 1996, resulting in SMPTE 298M-1997 "Universal Labels for Identification of Digital Data". SMPTE Universal Labels will be applied liberally to Wrappers and Metadata, for all of the following purposes:

- ➩ as the names of Metadata items;
- ➩ as the names of Metadata Sets and templates;
- ➩ as the preamble for Unique Material Identifiers (UMIDs);
- ➩ as a management tool for "dynamic documents",
- ➩ to be assigned as "magic numbers" as required to uniquely identify digital bitstreams, formats and files, in order to promote global interoperability in the digital domain.

The Sub-Group recommends that the SMPTE Registration Authority, Inc. (SRA) be designated as the central Registration Authority for the motion imagery and audio technology domains. The SRA will be operational in early 1998. The Sub-Group notes that the SRA has been designated by the ISO to be the Registration Authority for MPEG-2 "format-identifiers" or RIDs, and has been designated by the ATSC as the Registration Authority for ATSC programme identifiers.

4.5.3.1. Registration Authority activities

The SMPTE Registration Authority will carry out the following activities:

- Manage the Universal Label System.
- Use the Universal Label System to define the "keys" of the Metadata dictionary. The dictionary will include domain-specific vocabularies, detailing the sets of Metadata items appropriate for specific applications.
- Specify and publish an initial baseline universal dictionary of Metadata for use in many applications.
- Establish a due-process procedure for extending the dictionary in discrete revisions (versions 1.0, 1.1, 1.2, etc.) as the basis for compliance, where new versions must be backwards compatible with previous versions. Due process will be conducted by a designated SMPTE Engineering Committee which will delegate daily operations back to SRA staff.
- Establish liaisons with other due-process bodies, as appropriate, to enable interoperable Metadata.
- Establish a Registry which is capable of containing both standardized specifications and also user definitions for installation- or system-specific Metadata items, to serve the requirement for extensibility.

4.5.4. Metadata Templates

The Metadata subset required by and supported by all applications is also known as the Essential Metadata.

Beyond the Essential Metadata Set, Templates will establish conventions for the Vital Metadata Set (i.e. the repertoire, level and complexity of Metadata) necessary for each category of application. A key set of Templates will be defined by due-process standards.

A default behaviour must be defined within each Metadata Set for Templates that do not support that set.

4.5.4.1. Production planning

In this Template, pure Metadata is created, e.g. storyboards, scripts, crew schedules.

4.5.4.2. Acquisition and playback

The acquisition and playback systems only need Essential Metadata support in order to identify and document the Content represented by the Essence. However the system should allow full extensibility. This Template makes use of both Streaming and Storage Wrappers.

4.5.4.3. Editing and mixing

Editing and mixing applications demand a very high level of functionality and complexity in the area of Metadata. The storage format likewise, requires a sophisticated system which allows in-place editing of Metadata along with full random access. Due to the nature of these systems, a hierarchical storage mechanism may be a requirement for management of such a large amount of Metadata and associated relationships.

4.5.4.4. Storage of consolidated programmes

The result of the production process is a consolidated programme. In consolidation, it is usually acceptable to reduce the Metadata to the minimum required for such use. The requirements for in-place editing can also be relaxed, allowing simpler storage formats to be used for the Metadata. Depending on the target distribution format (see next section), both the Essence and Metadata may need to be converted to formats more suitable for such transmission. It should be noted that, while such conversion usually results in a reduction of Metadata, it may be necessary to add additional Metadata which describes specific distribution and transfer attributes (such as licence rights).

4.5.4.5. Emission

Efficient storage and transmission of Content often requires Wrapper formats that are optimized for minimal latency and efficient access. This would generally mean that the amount of Metadata can be reduced to the minimum necessary. Furthermore, the transmission channel may impose additional requirements on the maximum quantity of Metadata that can be handled. On the other hand, additional Metadata may be required to handle transmission-specific requirements.

It is explicitly recognized that different transmission domains (e.g. SDTI versus IP) require a different set of Metadata to accompany the Content, due to the varying needs of each environment. For example, in an IP system, one may want to add additional Metadata in the form of a URL which points to related text and graphics for the Content. Such Metadata, however, would not probably be necessary when Content is being transmitted in a studio environment over SDTI.

4.5.4.6. Archiving

Ideally, an archive system needs a superset of the acquisition Template plus production history. Sufficient Metadata needs to exist to allow fast and efficient identification of the Content. Additional Metadata may be included to describe detailed characteristics of the Content to allow precise searching of the Content. Extensibility is once again a high requirement in order to allow inclusion of customer-specific data such as links to computer databases or billing systems. Applications may use hierarchical storage to contain such Metadata.

4.6. Unique Material Identifiers

To enable the tracing of Content as it passes through a system, it is extremely important to provide a unique identifier of the actual Content contained in the Wrappers. Such a Unique Material Identifier (UMID) is necessary, since multiple identical or near-identical copies of the Content may exist within the system in various locations, and they must be referred to, independent of their location. It is also necessary to refer to Content which has been temporarily removed from the system, for example when transferred to archival storage, or when delivered out of the system, or even before the Content has been created.

Unique Material Identifiers are also used to trace copyright information and ownership of finished programmes. In the publishing industry, these identifiers are called ISBN numbers and such like; a well- known example in use in television today is the ISCI number used for identifying commercials. Work is also under way to define an International Standard Audio-visual Number or ISAN. The Advanced Television Systems Committee (ATSC) has also specified a Programme Identifier within the US Digital Television bitstream. Besides these Unique Programme Identifiers (UPIDs), in television production there is also a need to uniquely identify unfinished material, including both source material and transient or intermediate Content Elements.

File formats and Asset Management systems must be able to continue to work with existing methods of material identification but, at the same time, it would be beneficial to define a standard scheme for constructing UMIDs to enable greater interoperability of new systems as well as to provide the potential to upgrade existing systems.

4.6.1. Standard UMID Core

A standard UMID must be constructed from several parts:

1a. Prefix;

1b. Unique Material Number;

1c. Unique Instance Number;

2. Core Data;

3. Status Metadata;

. . . plus (optionally), Extension Metadata.

The Prefix is an SMPTE-administered Universal Label which identifies that this is an SMPTE UMID. It may additionally include other registered information such as the identification of the country and facility.

The second part of the UMID is a globally-unique material number whose method of generation is yet to be defined.

The third part is a locally-unique “instance” number whose purpose is to indicate the current version of the material. The source material would have a null value while other instances of the same material (copies etc.) would be assigned non-zero numbers through an algorithm (yet to be defined) which is intended to ensure uniqueness of every instance.

These first three parts of a standard UMID could be contained within a short field, perhaps as small as a 16-byte binary number, but certainly no larger than 24 bytes.

All fields in the UMID could potentially be automatically generated for every Content Element. Where automatic generation is not possible, the data could be assigned from a database following rules that guarantee to ensure uniqueness. For automatic generation, various algorithms are possible, employing techniques such as time counters and random number generators. Exact algorithms are expected to be defined during the standardization process.

The second major part is the Core data. The Prefix data, in conjunction with the Core data, forms the **UMID Core** itself. Allocation for the UMID Core should be assigned as early as possible in the production chain, preferably at the source. Once a value has been assigned to any UMID Core component, it cannot be changed and must always remain attached to the associated Content Element.

The third (Status) part of a UMID consists of modifiers and status indicators to identify near-identical copies, versions, generations and representations of the Content Element – the length, format and use of this third part should be specified by each particular system. This might also include element identifiers to allow a single UMID Core to be applied to an entire Content Item, not just a Content Element.

In some cases, encoding the **Status Metadata** as a single qualifier field can simplify or speed up operation of a system. In other cases, it may be more appropriate to keep the information as defined items of Metadata, with individual names and specifications.

All UMID Core data can be automatically generated using, for example, Geospatial Metadata such as the position of the source and the data, and the time of Essence capture.

The Status Metadata allows the Content to be labelled with the current status of the current instance, as it progresses through the production chain. This Status Metadata is editable when the Content is cloned, copied or in any way changed from the source. Ultimately, if the Status Metadata database file were lost, the UMID Core could still provide enough information to assist in manually identifying the Content source.

Note that the UMID Core applies to the whole Content Element as static data. Within a Content Element, additional resolution is provided by supplying the datecode and timecode numbers, or other Temporal Metadata.

The Essential Metadata Set which is needed to access each Content Component is thus comprised of the UMID Core plus some portion of the Status Metadata, dependent upon the precise application.

Some applications may also allow the addition of optional **Extension Metadata**. This is categorized as Metadata which is useful to distinguish a particular aspect of the current representation of the Content but which, if lost, would still allow the Content to be used in a useful way. Extension Metadata is freely editable and often manually, rather than automatically, entered.

The exact formatting and total data requirement for the UMID Core plus Status Metadata will be determined during standardization. Examination of “strawman” proposals indicates that the total size will not exceed 64 bytes per frame.

4.6.2. UMIDs as a linkage between Streams and Databases

Fig. 4.4 gives a general view of the linkage between the UMID Core and the Status and Extension Metadata. The reason for such linkage is that, in most instances, the UMID Core will always be carried with the Content on a per-picture basis, whilst the Status and Extension Metadata will be stored on a server, on a clip basis. Thus, in order to provide a unique link between the UMID Core and the Status and Extension Metadata, the UMID Core value at the clip start point is included on the server. Therefore, even in the event of a broken link, the UMID Core and its associated Status and Extension Metadata file can be re-linked by an automated search process. This may take some time, but it is at least possible where current systems offer no such possibility.

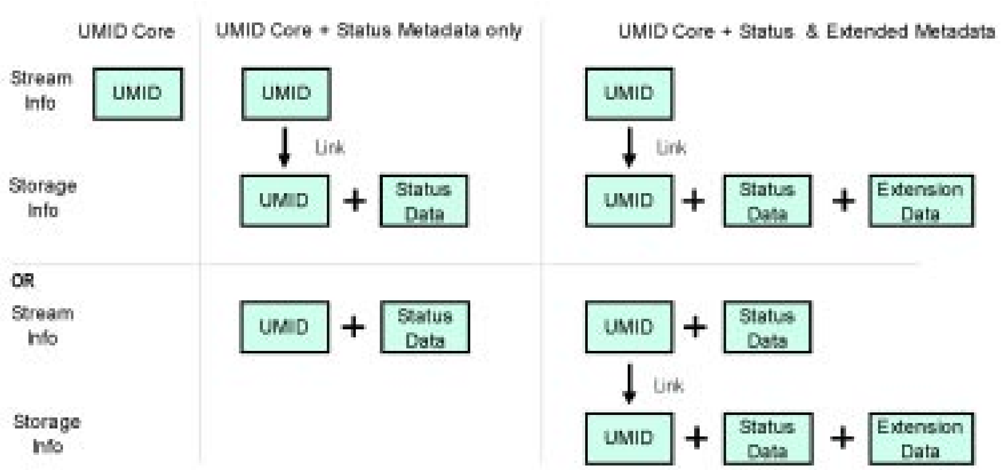


Figure 4.4: Illustrated concept of the layered UMID.

4.6.3. *The UMID Core in Streaming applications*

For video signals, a format will be needed to carry the UMID Core data in the VBI as ancillary data. The coding for the UMID Core data should be able to pass through not just the digital channels but also through compressed VBI channels (as available in MPEG-2 4:2:2P@ML and analogue videotape recorders). This is an important aspect of ensuring that the UMID Core can pass through the A / V production chain. Standards will be required for the formatting of the UMID Core through various parts of the broadcast chain, whether they are carried as Metadata packets alongside of the Essence or embedded within the Essence.

4.7. Wrapper formats

4.7.1. *First Request for Technology*

Early in the study it was decided to treat the streaming and storage of Wrappers as separate processes and this was matched by responses to the First Request for Technology (RFT). Overall, the formats which were targeted directly at just one or the other (streaming or storage) showed greater flexibility, and a simpler implementation model. Another factor considered in review of the RFT responses was the level of commitment to the standardization of various formats. At this time (July 1998), it is clear that a single format will not satisfy all the requirements.

A number of technologies were separately identified in addition to the RFT responses which are already in the process of standardization or available in the public domain.

RFT responses were received for the following Storage format technologies:

- Open Media Framework Interchange;
- Structured Storage;
- QuickTime®.

RFT responses were received for the following Streaming format and base technologies:

- Advanced Streaming Format;
- QuickTime®;
- SDTI Elementary Streams;

- ⇨ SX Native and MPEG-2 Transport Streams Multiplexing formats;
- ⇨ Fibre Channel-AV Simple Container;
- ⇨ Frame Tables;
- ⇨ Unique Material Identifier.

Many of these technologies also contain native structures for the formatting of Metadata and other topics outside of the direct requirements.

As promised in the RFT process, features from all the RFT responses were combined in this report and, where this occurs, the specific sources of each is identified.

4.7.2. *Storage Mechanism and Second Request for Technology*

From the Wrapper structure diagrams given above, it is clear that a common element of all the RFT responses which address Complex Content Packages ("Rich Wrappers") was the use of a special-purpose persistent storage mechanism at a relatively low level of complexity. In all cases, the data model of the overall system was mapped onto such a storage subsystem.

In addition, most of the Streaming Wrapper responses had an embedded underlying Streaming Mechanism at the Network Layer of the OSI model, equivalent to the Storage Mechanism in the storage domain. Just as there are several Streaming Mechanisms optimized for particular stream types, we might expect several Storage Mechanisms optimized for particular storage media and functionality.

Hoping to find a broadly applicable solution for a special-purpose Storage Mechanism, the Sub-Group issued a Second Request for Technology for this specific item. One such proposal was received for a structured storage system, and it is recommended that this be forwarded to the SMPTE for standardization.

By adopting an SMPTE Universal Label as an essential requirement in the special-purpose Storage Mechanism, and as a requirement in the actual file, it becomes possible to identify uniquely an SMPTE Storage Wrapper which is independent of the file name and operating system.

4.7.3. *Interoperability of Wrapper formats*

There are three levels of interoperability:

- ⇨ **level 1** – which ensures the carriage of Essence and Metadata types whatever they may be;
- ⇨ **level 2** – which allows the Essence and Metadata types to be carried and successfully decoded;
- ⇨ **level 3** – which, in addition, allows Complex Content Packages to be interpreted.

Work is already under way to standardize a number of streaming formats through the normal standards processes. In general, this provides level 1 interoperability.

Storage Wrappers have proved to be a more difficult area, but several technologies have been identified which, when applied in simple forms, will help to provide level 2 interoperability.

Achieving level 3 interoperability is still more difficult. However, progress has been made towards this, outside the traditional television equipment domain, and it is now hoped that engagement of new participants in the standards process, in combination with the general techniques studied by the Sub-Group and discussed here, will result in full level 3 interoperation.

Interoperability can be greatly aided by the following mechanisms:

- ⇨ a Central Registry for new and existing data objects;
- ⇨ abstraction of items such as compression type, and timing and control streams;
- ⇨ refraining from developing ambiguities, by having data defined by both a custom data type and a well-known data type;

- modification and adoption of existing Templates whenever possible, rather than developing entirely new models;
- cross-reference mapping of new Metadata types and Templates, when adding them to the Registry;
- Support for software APIs and component models which use a platform-independent referencing system.

The use of all of these mechanisms is a common feature of all the submissions the Sub-Group has studied.

4.7.4. Interchange

For interchange between different systems, simpler specifications generally lead to a greater possibility for the interchange of Essence and Metadata. Likewise, well-prepared public specifications and recommended practices lead to greater interoperability.

In accordance with the discussion above, there is a real prospect of achieving level 2 interoperability by focusing upon the documentation of a simple application of the technologies which have been reviewed. As techniques improve and standards become better defined and understood, more complex interchange formats can be contemplated for full adoption.

4.7.5. Streaming and Storage Wrappers

Wrapper formats are divided into two primary classes: Storage types and Streaming types.

The following schematic diagram helps to appreciate the different components of a Wrapper and how the different elements interact.

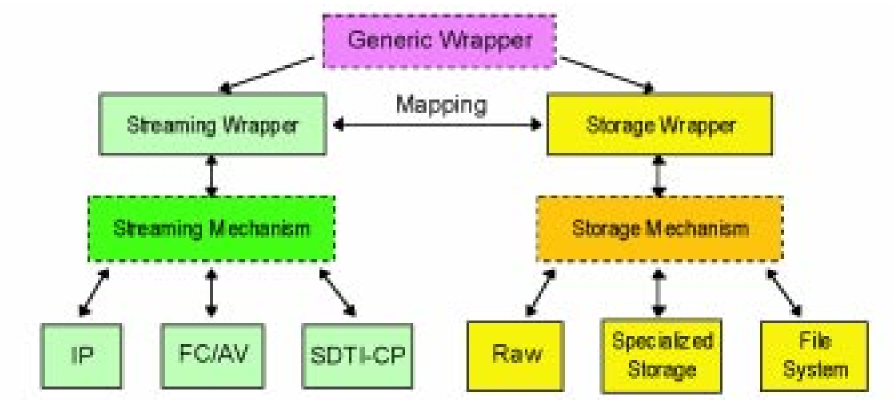


Figure 4.5: Schematic diagram of Storage and Streaming Wrappers.

The highest level is an abstract item called the **Generic Wrapper** which is a general representation of all Wrapper types.

For the purpose of clarifying the implementations of Generic Wrappers, the diagram shows a division into two primary uses: **Streaming Wrappers** and **Storage Wrappers**.

Streaming Wrappers have basic mechanisms to enable Wrappers to be transferred over a connection, whereas Storage Wrappers offer a different but comparable set of basic mechanisms to enable Wrappers to be persistently stored.

Under each of the basic mechanisms are a number of different instances of both Streaming and Storage types.

With appropriate documentation of each instance of a mechanism, conversions can be made between all identified mechanisms whether Streaming or Storage. It should be noted that conversions may be incomplete, depending on the level of compatibility between the instances.

4.7.6. Repetition rates of Vital Metadata

In a multicast or broadcast environment, the various receivers must be able to decode the stream even when they join the transmission in the middle. This implies that the Vital Metadata must be repeatedly inserted into the stream, with a frequency appropriate to the application. This will typically be once per editable unit of the stream, i.e. once per frame in frame-bound systems, or once per GoP in interframe systems. This same repetition of Vital Metadata can be used to provide increased confidence in Vital Metadata, when the signal is transported over links with significant bit error rate or packet loss.

To a large degree, the repetition rate is implementation-specific, as the effects of the implementation upon the render delay are great. For instance, if a service provider encodes all of its channels using the same parameters, once a user has received Vital Metadata for one channel, the delay for other channels will be governed not by repetition rates of Vital Metadata but rather by repetition rates of key frames. Because there is already a potential lag in the ability to render any given stream which is based upon key frame repetition rates, it makes sense to specify that repetition rates do not exceed key frame-rates. Thus we conclude that the maximum repetition rate is the same as the minimum key frame-rate.

When dealing with high-density programme streams, it is likely that any Overhead caused by Vital Metadata repetition will be minimal to the point of irrelevancy. Because of this relationship, it is recommended that Vital Metadata be repeated along with key frames in these systems. Network-based transmissions present a more complex problem. When bandwidth permits, the delay should be kept to a minimum. Many if not most network-based transmission systems have alternatives to repetition of Vital Metadata and will likely find a better alternative where bandwidth is precious.

4.7.7. Streaming Wrapper formats

Streaming Wrapper formats are primarily to be used when Programme Material is transferred between origination, display, transmission and storage devices. In different applications, the underlying transmission medium may be inherently reliable or unreliable, and the transmission mode may be point-to-point, multicast (point-to-multipoint), or broadcast. Moreover, the range of bit-rates may be from a few hundreds of kilobits per second to several hundred megabits per second. Finally, the stream may be composed of a single Essence type or it may be a Complex Content Package. *Fig. 4.6* outlines the basic stream structure.

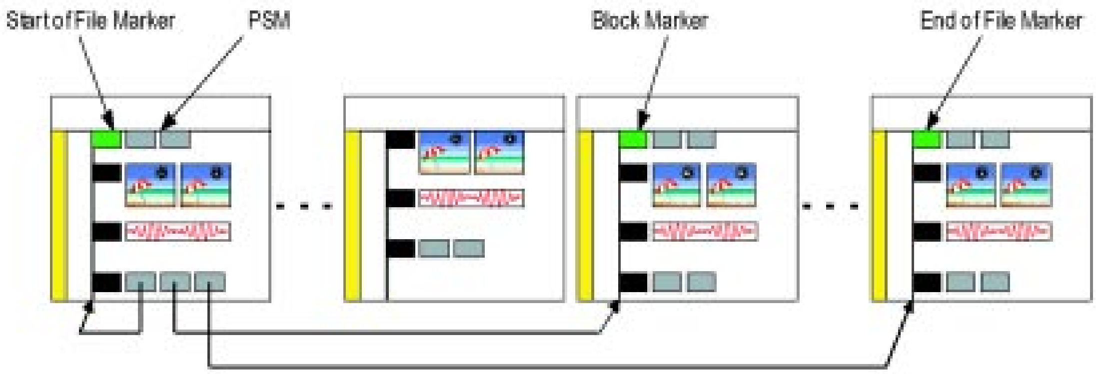


Figure 4.6: Basic stream structure.

It is extremely unlikely that a single format can be stretched to cover all of these possibilities. Anyway, continual innovation will result in new formats being developed which are optimized for each new important combination of variables. In this context, the best contribution of the Sub-Group is to document the basic structures necessary for mapping Content into Stream formats, and to point out the importance of achieving interoperability between formats.

Content Items will pass through several different transmission media between origination and consumption, and it is desirable to avoid transcoding as much as possible. This is equally true for the Essence and for the Metadata.

In addition, during the production and delivery processes, there will be a need to insert items of Metadata to record editing decisions or to guide downstream equipment, and to extract items of Metadata for use in local operations or to discard them as they become unnecessary or irrelevant.

An important criterion for the evaluation of Stream formats is, therefore, how well each format conforms to the standards for Metadata which are already fixed or are under development.

4.7.8. Storage Wrapper formats

The situation for Storage Wrapper formats is a little different. In this case, it is not always important to repeatedly insert the Vital Metadata where it is assumed that a file is always transferred in its entirety. However, the Wrapper format can provide facilities for random access to the Content, and allow modification of the Content in place (i.e. without rewriting the unchanged data).

The format may also accommodate arbitrary complexity of Content, including recursive structures to represent the composition of elements from other elements, the use of pointers and references between elements both inside and outside the file, and the mingling of Essence encoded in many different ways. This is known as the Complex Content Package as embodied in *Fig. 4.3.*

Amid all this complexity, it was realized that a very few basic structures can be used to construct a representation of all the Content in a Storage Wrapper. For example, a multimedia composition can always be represented as a synchronized collection of tracks; each track can be represented as a sequence of clips; and each clip is basically a reference to a portion of a separately-described Essence element. Process steps and effects can be represented as a collection of Metadata items attached to one or more clips, within one or more tracks; and multiple layers of composition can be represented by source clips referring to a sub-master composition, instead of referring directly to the source material.

The Sub-Group has identified a single data model which is sufficient to address the widest range of applications in television production with reasonable simplicity. This process should continue through standardization.. The proposed data model is extensible to cover new applications which are not yet developed.

Beyond the data model, the Storage Format must also provide a container in which the data structures, the Essence elements, and the necessary indexes can be stored. For interoperation with the Metadata Registry and with the Stream formats, the object identification scheme employed by the container suite should be compatible with the Registry-naming scheme.

Several such container formats and associated APIs have been developed in the computer industry, and have been proposed or solicited for contribution to the Sub-Group. In order to be acceptable for standardization, however, these formats must either be documented to a level which allows their re-implementation from scratch on multiple platforms, or reference implementations of a complete API to create the format must be made available on non-discriminatory terms or placed in the public domain. The response to the second RFT showed that there is a good likelihood that at least one format can pass the standardization process.

4.7.8.1. Wrapper varieties

Programme material will involve at least six different varieties of Wrapper:

A) **Unwrapped Content** – for example, signals from today's traditional recording equipment, or from foreign, non-conforming systems.

B) **Wrappers containing Content Items which are predominantly Essence but with some Metadata including at least a UMID**. These files are the modern equivalent of tape recordings.

C) **Wrappers containing Content Items which have no Essence, and comprise only Metadata**. The Metadata may include Descriptive, Relational and Reference Metadata. If Essence is involved, it will be kept elsewhere (unwrapped or wrapped), and these Wrappers will include references to it. These files are the modern equivalent of Material Logs and similar files. Another name for these files is **Proxy Wrappers** – where the Items in the Wrapper are proxies for the Content which is elsewhere.

D) **Wrappers containing Complex Content Packages which include both Composition Metadata and Essence embedded within the Content Packages**. These Content Packages are in common use today in non-linear editing (NLE) systems.

E) **Wrappers containing Complex Content Packages which are predominantly Composition Metadata and therefore include additional Relational or Reference Metadata which, in turn, refers to Content (unwrapped or wrapped) kept elsewhere**. These Content Packages are the modern equivalent of Edit Decision Lists (EDLs) and similar files.

F) **A variation of type C above, where the Wrapper contains both Content Items and Complex Content Packages**. Reference Metadata in the Complex Content Packages refers to the other Content Items in the same outer Wrapper.

Reference Metadata is one variety of Composition Metadata. A Reference might point to the following:

- ⇨ an entire Content Item or Package;
- ⇨ a region within a Content Element (e.g. a "sub-clip" of a shot);
- ⇨ a point within some Content (e.g. a single frame, or instant);
- ⇨ a specific item of Metadata;
- ⇨ a specific item of Reference Metadata which in turn points to one of the above.

The mechanism employed for specifying References is recommended to be via a UMID. In some simple cases, the References might be given directly as offsets (as in type D above) or as local identifiers (as in type F); but even in these cases, the UMID should be specified in addition to the shortcut reference.

4.7.8.2. Interchange Storage Wrapper format

Starting from the above definitions, the development of specific Wrapper structures becomes remarkably simple. The following example is of an Interchange Wrapper of type C above (a "Proxy Wrapper"):

The file starts with a Universal Label which identifies it as a Proxy Wrapper. This might be combined with an item of Relationship Metadata which counts the number of proxy references contained in the file.

This is followed by a set of UMIDs, each with optional Status and Extension Metadata. So, the Proxy Wrapper file can be specified as follows (see *Fig. 4.7*):

- An SMPTE Universal Label (up to 16 bytes)
  Key = "Proxy Wrapper" (value TBD).
- A Relational Metadata Item
  Key = "Group" (value TBD)
  Value = Number of Components referenced.
- For each Component, add .....
  Key = "Proxy" (value TBD)
  Value = UMID, Status Metadata (including URL), Extension Metadata.
- A termination value.

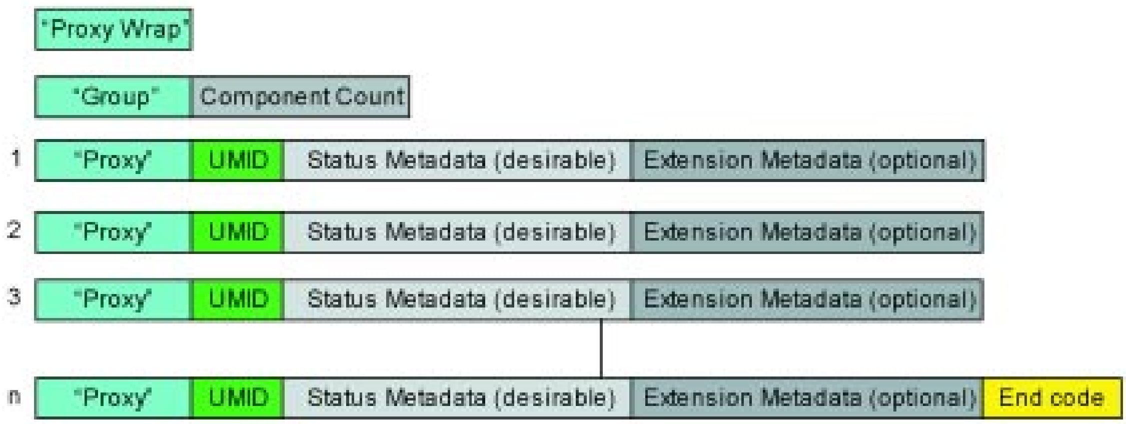


Figure 4.7: Example of the Interchange Proxy Wrapper file structure.

4.7.8.3. Content repertoire

The definition of each new variety of interchange Wrappers must include the repertoire of allowable Content Packages, Items and Elements within the format.

This definition should be contained within the Registry entry for the label which identifies the Wrapper type.

4.7.8.4. Data model

In the analysis of responses to the First Request for Technology, it was realized that several of the responses to the request for a "Rich Wrapper" (now called the Complex Content Package) offered complete systems whose capabilities overlapped. Although reconciliation of these systems into a single system might be achieved, the resulting system would be of such wide scope that the process of standardization, and implementation of the resulting standard, would take a very long time. This does not meet the TFHS goal of selecting a Wrapper format which can be standardized and deployed in the near future.

At the same time, it was clear that all responses used startlingly-similar data models which described "programmes" as a combination of "tracks", and made references to "media data" which is divided into "chunks", described by "hints", incorporating "effects" which are controlled by a set of "tweaks" and so on (this is intentionally a mixture of terminology).

There was renewed agreement to define an overall data model which is adequate for all known existing tasks and is extensible to admit new applications and definitions. It is expected that every proposed system would be mapped onto this model, either directly or through the use of translators.

A large part of this data model is simply the creation of a glossary of terms for Complex Content Packages – which is a subset of the task of creating the Metadata Dictionary with a "vendor-neutral" vocabulary.

In addition to this vocabulary, terminology is needed to describe the logical structure of the data model. The DAVIC terms of Content Package, Content Item, Content Item Element, plus the other defined terms in the First Report form the basis of this class of terminology. In many cases, this terminology is not just a set of words, but also has a real existence as data inside files. It is therefore another class or sub-class of Relational Metadata and should be addressed in the Metadata Dictionary.

4.8. Newly-identified work

The following items of work have been identified to be taken up by the SMPTE or other standards bodies; but these items have not yet been initiated.

4.8.1. Interchange Storage Wrapper

A Proposed Standard on an Interchange Storage Wrapper is required for interoperability between different equipment. The work summarized in this document should be passed to the SMPTE for consideration as a Standard in P18.

4.8.2. Wrapper profiles and Content repertoires

As new examples of Concrete Wrappers are produced, they should be documented within the SMPTE Registry.

4.8.3. Metadata Sets

As specific Metadata Sets are identified, they should be documented within the SMPTE Registry.

4.8.4. Essence format documentation

The SMPTE Registry should also be used as the machinery to document Essence formats precisely (for example, Video Rasters, Audio Packet formats, Graphics formats) and codify them for use within Wrappers.

To some extent, this has already been achieved in the context of data formats for SDTI (notably, for DV-based Content Packages), and is underway for PCM Audio and for MPEG-2 Video Elementary Streams.

However, this approach to documenting Essence formats as stand-alone specifications should be pursued for primary Essence types which will be encountered in the Streaming and Storage Wrappers identified above. Other types of Essence may be documented, but it is to be encouraged to convert these Essence types to primary types in order to maximize interoperability.

4.8.5. Mapping onto transport mechanisms

Standards must be written to map the Wrappers described here onto the transport mechanisms called out in *Section 5*. Several of these projects are under way; however, no proposals have been seen for documenting a Streaming Wrapper format for transport over ATM.

4.9. Conclusions

Starting from the First Report on User Requirements, the Sub-Group started to search for a comprehensive solution, and in the process issued a first Request for Technology. Several responses were received, covering aspects of the required technology, from established companies in both the computer and television industries. The responses ranged from discussions of specific items such as Unique Material Identifiers and Frame Index Tables for use inside Wrappers, to complete solutions for specific applications such as multimedia presentation delivery, to specifications for data models and container formats that are in use today in the industry, within multiple products. These responses were analyzed during repeated meetings, along with comparisons of existing practices in the industry and discussions on standards development efforts which have been continuing simultaneously.

No single response to the RFT covered all Requirements; however, in general, the sum of the responses on Stream formats covered most of the Stream requirements, and similarly the sum of those on Rich Wrapper formats covered most of the Complex Content Package requirements.

Not surprisingly, the various proprietary technologies submitted were not immediately fully interoperable to the degree requested in the First Report. However, in their use of established practices such as the use of globally-unique identifiers, some of the responses were more amenable than others to limited modification to achieve interoperation.

The final phase of the Sub-Group's work was to issue a Second Request for Technology in search of one missing item from the first response – the low-level, special-purpose, Storage Mechanism.

During the concluding meetings of the Sub-Group, it became clear that the technology to employ in comprehensively addressing the requirements does exist. However, it was not possible to complete the documentation of this technology within the scope of the Sub-Group. Instead, this should be taken up by the SMPTE following the plan below.

4.9.1. Work in progress

The following standardization activities are already under way in the SMPTE and in other organizations.

Note: The SMPTE Committees mentioned in this section are active at the time of writing (July 98). However, the SMPTE is shortly going to change its Technology Committee structure to better absorb the work that will arise from the Task Force's efforts, in which case the work will be undertaken by appropriate successor committees to those listed.

SMPTE PT20.07	Working Group on Metadata is processing several Proposed Standards for Metadata.
SMPTE PT20.04	Working Group on Packetized Interconnect Technologies is balloting a proposed Standard on the SDTI-CP Streaming Wrapper.
SMPTE P18	Ad Hoc Group on Unique Material Identifier is creating a Proposed Standard based upon the existing "strawman" proposal.
SMPTE P18.27	Working Group on Editing Procedures is documenting the second RFT as "Requirements for a Complex Storage Mechanism".
SMPTE P18.27	Working Group on Editing Procedures is also creating a Proposed Standard for a format of a "Complex Content Package".
SMPTE P18.27	Working Group on Editing Procedures is also working on a revision of SMPTE 258M Interchange of Edit Decision Lists.
SMPTE M21.21	Working Group on Content Formatting and Authoring has formed an Ad Hoc group on DVD Authoring.
NCITS T11	Standards Committee on Fibre Channel (formerly ANSI X3T11) is documenting the FC-AV Container as part of a Proposed Standard for Fibre Channel Audio/Video.
WG SC-06-01	Networks for Audio Production of the AES Standards Committee is working towards standards for Audio file interchange.
ISO SC29 WG11 MPEG	This group is working on an "Intermedia format" for storage of MPEG-4 data, to be included in version 2 of ISO 14496-1 MPEG-4 Systems, which is planned to be published by the end of 1999.

4.9.2. Which documents to hand over

Note: The SMPTE Committees mentioned in this section are active at the time of writing (July 98). However, the SMPTE is shortly going to change its Technology Committee structure to better absorb the work that will arise from the Task Force's efforts, in which case the work will be undertaken by appropriate successor committees to those listed.

The work summarized here and in the references on the Unique Material Identifier (UMID) will be passed to the SMPTE P18.x Ad Hoc Group on Unique Material Identifiers.

A Proposed Standard on SDTI Content Packages will be studied by the SMPTE P18 Ad Hoc Group on the SDTI-CP Wrapper.

The proponents of an Advanced Streaming Format are encouraged to formulate a proposed Standard for consideration by a new Working Group within SMPTE P18.

The proponents of the Structured Storage mechanism are encouraged to formulate a proposed Standard for consideration by the Working Group on Editing Procedures within SMPTE P18.

4.9.3. Note on proprietary technology

Considering the importance of Wrappers to the industry as a whole, it is clear than there is no place in the output of the Task Force for the selection of proprietary technologies. This is so much so that the SMPTE's usual rules for accepting technology for standardization cannot apply. The TFHS specifically recommends full documentation of the technologies considered here and conformance to the following guidelines.

To be standardized, any technologies must be open in these respects:

- ⇨ existing implementations must be freely licensable;
- ⇨ technologies must be documented sufficiently to permit new implementations from the ground up, without fee;
- ⇨ a due-process standards body must have control of the specification in the future.

The Sub-Group also recommends compliance-testing by an external organization.

Section 5

Networks and Transfer Protocols

5.1. Introduction

There is an increasing use of systems which apply packetization to video, audio and all data types. Interoperability, or simply data exchange in the form of files and streams between different systems, is a strong requirement.

To meet the demand for interoperability, a Reference Architecture (RA) for Content transfers is recommended. The RA recommends interfaces as well as file transfer protocols, protocols for real-time streaming of video, audio and all data types, and methods for file system access. Existing standards are recommended if available, while areas requiring future developments and standardization are identified.

5.1.1. File transfer and streaming

File transfer involves the moving or copying of a file, with the dominant requirement that what is delivered at the destination is an exact bit-for-bit replica of the original; retransmission of parts of the file initially found to suffer errors will be used to achieve this. Although the transfer may often be required to take place at high speed, there will be no demand that it should take place at a steady rate, or be otherwise synchronized to any external event or process.

Streaming is the transfer of television programme material over a transport technology from a transmitter to one or more receivers, such that the mean reception frame-rate is dictated by the sending frame-rate. The transmitter plays out the material without receiving feedback from the receivers; consequently, there is no capability for flow control or for the re-transmission of lost or corrupt data. It is a continuous process in which the transmitter "pushes" programme material to receivers that may join or leave the stream at any time. The transmission frame-rate is not necessarily equal to the material's original presentation frame-rate, thus allowing faster- or slower-than-real-time streaming between suitably-configured devices.

5.1.2. Quality of Service considerations

Ensuring that the needs of file transfers and streaming are met, in particular their respective *dominant* requirements as described above, requires the definition and allocation of "Quality of Service" (QoS) parameters. Some transport systems have implied and fixed QoS values, while others have values obtained by configuration or negotiation between users and the transport system. The latter is particularly the case with transport systems that allow the sharing of common resources (such as bandwidth) between multiple users with disparate QoS needs. This is necessary so that changes in the total loading on available resources, for example when a new user joins existing sharers of the transport mechanism, do not adversely affect the QoS of users already present. In some systems, the time needed for QoS negotiation is itself a QoS parameter, but this is beyond the scope of this document.

Major QoS parameters therefore are:

- bandwidth (which may be expressed as peak and average bit-rates);
- bit error rate;
- jitter and delay (latency);
- access set-up time.

5.1.3. Reference Architecture

The concept of a Reference Architecture (RA) is useful for designers and users of Content in production and distribution facilities. *Fig. 5.1* illustrates the domain in which the RA is applied.

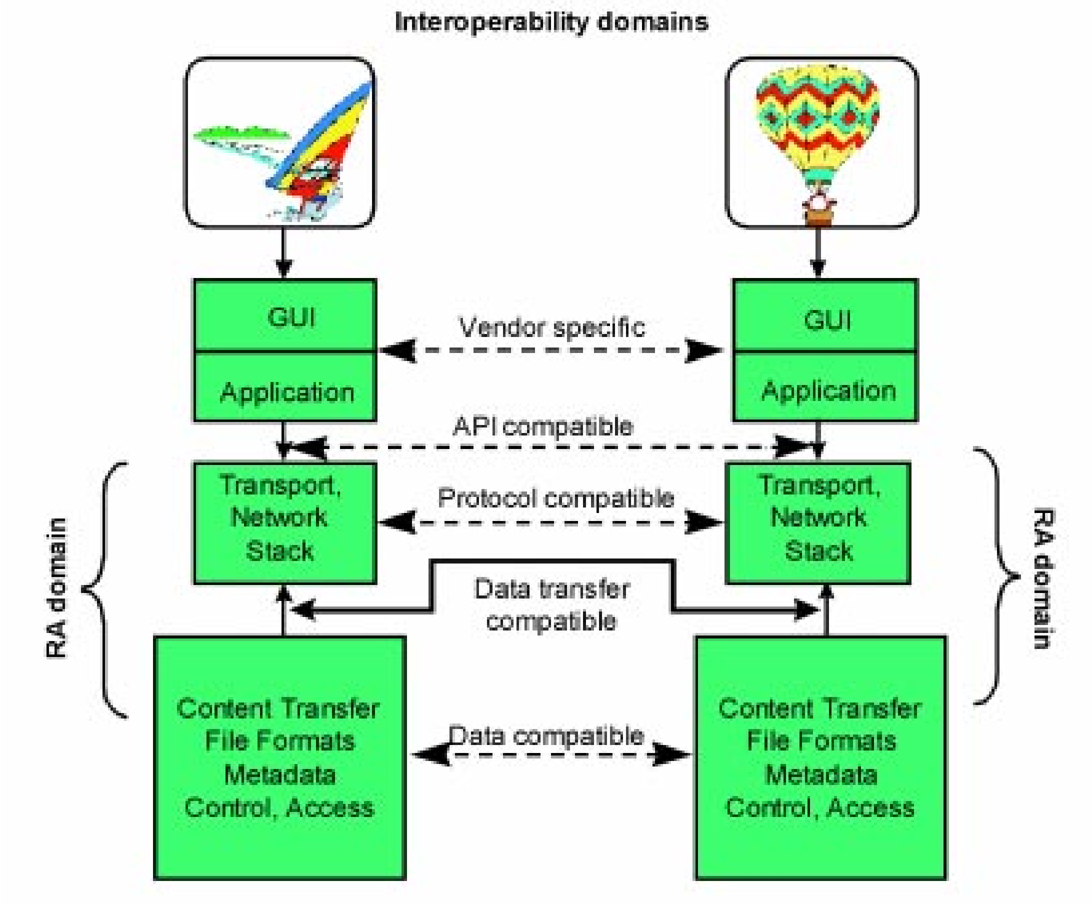


Figure 5.1: Interoperability domains.

5.2. File transfer methods

File transfer entails the asynchronous, error-free transfer of Content, either point-to-point or point-to-multipoint. In some applications, it may be necessary to specify QoS parameters for the transfer (e.g. a specified maximum bit-rate to avoid network congestion).

Four file transfer protocols are considered in this document:

- Universal FTP (based on TCP / IP);
- point-to-multipoint transfers using the eXtended Transfer Protocol (XTP);
- Fast File Transfer (methods which use hardware or lightweight software protocols over Fibre Channel, ATM and other transports);
- an enhanced version of FTP, called FTP+.

Also, a method to initiate the file transfer is specified. NFS – a widespread and standardized file system access method – is recommended, even though it may not provide the high-performance file system access that is required. The definition of an enhanced file system access method may be necessary in the future.

Annex E.2 gives more details about how file transfer is achieved over these and other transports.

Fig. 5.2 illustrates the file transfer methods which depend on application spaces.

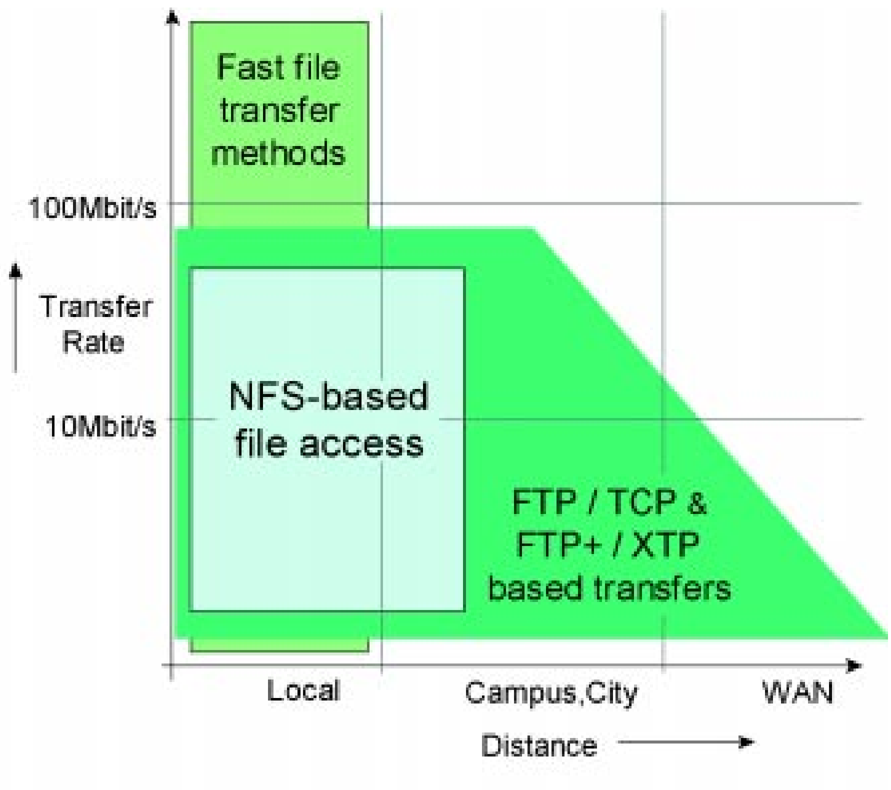


Figure 5.2: File transfer / access application spaces.

5.3. Streaming methods

The required characteristics of streamed Content are:

- the "bounded quality" of the received signal [6];
- the received quality is directly related to the QoS of the link and any Forward Error Correction;
- isochronous / synchronous links or Time Stamps.

Examples of transports / protocols capable of supporting streaming are:

- IP;
- ATM;
- SDI / SDTI;
- Fibre Channel;
- IEEE 1394;
- Dedicated purpose:
  - ANSI 4-40 (AES / EBU);
  - ITU-T T1, T3, E1, E3, ISDN;
  - DVB ASI, SSI.

Annex E.3. gives more details about how streaming is achieved over these and other transports.

6. There is usually no return path to request a retransmission, so the receiver must make the best of received data. Methods which do use a return path for retransmission of packets require buffering and, consequently, they insert a delay.

Fig. 5.3 maps various application spaces to these streaming transports.

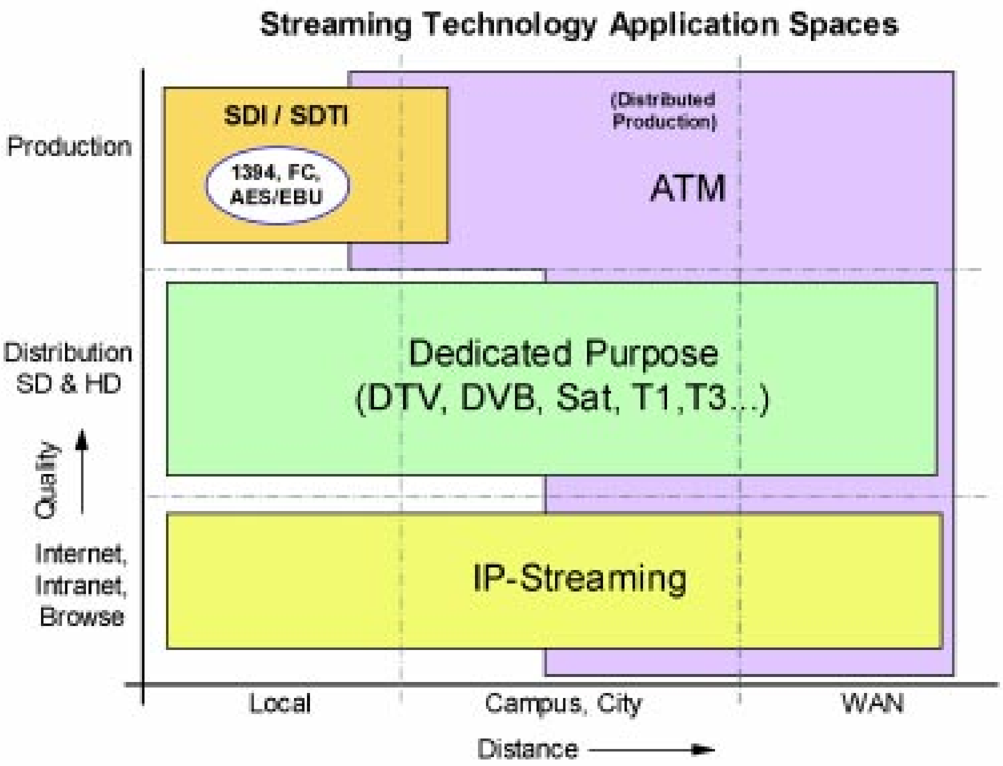


Figure 5.3: Streaming Transports and their application space mapping.

5.4. Recommended transport systems for Content

Recommended interfaces and networks include SDTI, Fibre Channel, Ethernet and ATM. While it is recognized that any of these technologies could be used for streaming within a broadcast studio, **SDTI** is currently (July 1998) recommended for streaming:

- ⇨ **Fibre Channel** is recommended for fast / large file transfers;
- ⇨ **ATM** is recommended for wide-area network file and stream transfers.

Note: Ethernet was not investigated in great detail by the Sub-Group but, due to its wide availability and complete standardization, it may be used as a general-purpose, low-performance, file transfer mechanism.

5.4.1. *Serial Data Transport Interface (SDTI)*

An additional layer above SDI (ITU-R BT.656, SMPTE 259M) has been defined to enable the transport of packetized data. This layer is called SDTI and has been standardized as SMPTE 305M.

SDTI is a universal transport layer which describes in detail the packetization and basic signalling structure for the data to be transported. The use of the underlying SDI interface determines the technical capabilities of SDTI. For example:

- ⇨ SDI is a **unidirectional interface**, making SDTI a unidirectional transport system per connection. To enable the re-transmission of corrupted data, an additional "control" data path from destination to source will be needed to signal the requirement. SDI / SDTI should therefore be used primarily for streaming and not for conventional file transfers [7], as described in *Section 5.2*.

7. There are methods under discussion within the SMPTE and the IETF for the transfer of files over unidirectional links.

- ➩ SDI is a **point-to-point interface** and defines the bit-rate (e.g. 270 / 360 Mbit / s), jitter, delay and loss which constrains the SDTI layer. To overcome the point-to-point restrictions, addressing capabilities are embedded in the SDTI header and may be used in the future for dynamic routing.
- ➩ SDI is a **synchronous interface**; SDTI payloads can therefore be synchronized easily to the TV environment. This capability makes SDTI suitable for use in studio applications for the streaming of Content in real-time and faster-than-real-time, and it also facilitates multiplexing.

All applications using SDTI as a transport require individual documentation.

5.4.2. ATM

ATM is a network technology for use in local and wide-area applications, accommodating a wide range of QoS requirements for the transport of data, voice and video traffic over a common network infrastructure. Payload data is encapsulated in 53-byte containers (ATM cells). Each cell contains a destination address and can be multiplexed asynchronously over a link.

Connections through ATM networks (Virtual Circuits – VCs) can either be pre-configured (Permanent Virtual Circuits – PVCs) or established on demand (Switched Virtual Circuits – SVCs) using standard protocols. A Quality of Service "contract" may be defined for each VC.

5.4.2.1. File transfer over ATM

File transfer over ATM is achieved through its capability to transport IP. This can be achieved through one of the following protocols:

- ➩ Classical IP over ATM (IETF RFC1577);
- ➩ LAN emulation (LANE) (ATM Forum 94-0035);
- ➩ Multi-protocol over ATM (MPOA) (ATM Forum 96-0465);
- ➩ Multi-protocol Label Swapping (MPLS) (IETF Draft).
- ➩ Streaming over ATM

ATM is suitable for Content streaming in both the local and wide area. ATM provides QoS guarantees and error-correction capabilities. In order to meet the jitter and error-rate requirements of professional broadcast studio applications, engineering guidelines for the transportation of streams are necessary. Such guidelines would identify, for example, the ATM network characteristics that must be supported by devices for streaming over ATM.

Synchronization between transmitting and receiving devices is achieved through the transport of timing references within the stream. ATM's scalability, bandwidth efficiency and support for multiple traffic types make it a suitable choice for streaming Content between studios over the wide area. International standards exist which define how the MPEG-2 Transport Stream is streamed over ATM.

5.4.3. Fibre Channel

Fibre Channel is a network that is suited for use in broadcast studio applications. It has been accepted as the high-performance computer peripheral interconnect. It is also being used as a building / campus-wide high-speed network.

In broadcast applications, Fibre Channel is being used by the vendors of video disk recorders as a studio packet network (using IP) and for shared storage attachment (using SCSI protocol). Standards exist to define the transport of IP packets over Fibre Channel.

Commonly-available Fibre Channel links (1 Gbit / s gross) support payload rates of about 800 Mbit / s (a proposed Fast Fibre Channel transfer protocol will enable a net bit transfer rate of up to 760 Mbit / s).

Fibre Channel solutions are limited to local-area transfers and an FC / SCSI-based solution will not be as flexible as one based on IP routing, but has the advantage of higher performance. Fibre Channel offers limited WAN functionality.

5.4.3.1. File transfers over Fibre Channel

Fibre Channel may be used for FTP transfers using IP over its FC-4 layer. The alternative *Fast File Transfer* should follow the rules defined in the Fibre Channel Audio / Video (FC-AV) standard presently under development by the NCITS T11 group. To allow a predictable transfer time, fractional bandwidth capabilities must be available from Fibre Channel component suppliers and implementers. For very fast transfers, therefore, transport protocols implemented in hardware are required.

5.4.3.2. Streaming over Fibre Channel

The Fibre Channel System (FCS) is designed for asynchronous transport but can also support synchronous transport by embedding timing references within the stream and by using FCS Class-4 (fractional bandwidth) – see the section above.

5.4.4. IEEE 1394-1995 high-performance serial bus

The IEEE 1394 bus was designed to support a variety of digital audio / video applications in a desktop environment. The version encountered in this environment relates to a specific cable and connector type, and is limited to cable lengths of about 4.5 m. Some companies have announced cables which work up to 100 m.

The physical topology of IEEE 1394 is a tree or daisy-chain network with up to 63 devices, which need not include a dedicated bus manager. Devices act as signal repeaters; the physical connections between nodes are made with a single cable that carries power and balanced data in each direction.

The base data-rate for the IEEE 1394 cable environment is 98.304 Mbit / s, and signalling rates of 196.608 Mbit / sec (2X) and 393.216 Mbit / s (4X) have been defined.

It is noted that different and incompatible protocols exist (CAL and AVc). Harmonization is strongly encouraged through communication with EIA-R4.

5.4.4.1. File transfers over IEEE 1394

Unlike most other Data Link protocols, IEEE 1394 provides the capability for isochronous as well as asynchronous transmission. This capability has a significant impact on how IP is supported. The IP1394 working group of the IETF is working on an architecture document and appropriate protocol documents for the usage of these link layer properties. Both IPv4 and IPv6 will be addressed.

5.4.4.2. Streaming over IEEE 1394

The isochronous mode of operation provides a streaming capability. In addition to the absence of confirmation, the principal difference from the asynchronous mode is QoS: isochronous datagrams are guaranteed to be delivered with bounded latency. Bandwidth for isochronous data transport (up to about 66% of the total) is reserved on initialization and after any change in the network configuration.

5.5. Other transport systems

Other transport systems are available which may well be suitable in certain instances. These include Gigabit Ethernet, HIPPI, HIPPI-6400 and other IP-based transport systems. However, system implementers are advised that these alternative systems were not considered by the Task Force and their capabilities should be compared to the criteria for data exchange described in this document. The requirements for file transfer are in general more likely to be met than the requirements for streaming.

5.6. Bridging and Tunnelling

Bridging is the transfer of payload between different transport mechanisms such as the moving of MPEG-TS from SDTI or Fibre Channel to ATM (often needed between LANs and WANs).

Tunnelling is a way of transporting a complete interface data structure, including payload, through another interface; for example, IP over SDTI. IP multicast can then be used to transport both Data Essence and Metadata which are associated with the SDTI Audio / Video Essence carried in the main SDTI programme.

Both methods are discussed in *Annex E.*

5.7. Recommendations for future work

5.7.1. File transfers

To meet all requirements for file transfers, the implementation of an enhanced FTP (i.e. FTP+) over the eXpress Transfer Protocol (XTP) is recommended. The FTP+ profiles listed in this document are merely requirements and are neither developed nor standardized. The development and standardization of FTP+ and XTP has started in the SMPTE. Chunking for files also needs consideration by the SMPTE.

The FC-AV committee is working on a standard for a high-performance file transfer protocol. The completion of this work is of great importance and should be actively encouraged by the EBU and the SMPTE. It is important to communicate to manufacturers the requirement for an early implementation of this standard into products, and to encourage further liaison with the FC-AV Group.

A method of setting up and determining the file transfer capabilities between different devices needs to be standardized, probably in the form of a recommended practice.

End-to-end resource control and bandwidth management, when performing file transfer or streaming, is not addressed in the Reference Architecture. This larger type of interoperability should be a long-term goal. However, the standards addressed in the Reference Architecture will provide much utility even without a complete Systems Architecture.

5.7.2. Streaming / mapping

Concerning streaming recommendations, the previously mentioned FC-AV protocol is also suitable for a synchronous real-time / streaming data transport. An implementation of this is presently not available but of great importance. It is important to communicate to manufacturers the requirement for an early implementation of this standard into products, and to encourage further liaison with the FC-AV Group.

The streaming of Content in an interoperable way requires detailed information about the data structure of the Content to be streamed, and a method of how the data structure is mapped into the transport mechanism. In this document, the so-called mapping tables which describe the required Content mappings need to be finalized. It should be stated that, for a given interface (transport mechanism), only one kind of mapping for a given Content should be standardized, in order to avoid multiple standards for the same purpose. Work on several mapping documents is ongoing, proposed or required.

The SMPTE is already working on a draft proposal for tunnelling IP over SDTI links. It is also working on non-conventional file transfer over SDTI links, as part of a Generic Wrapper structure.

5.7.3. Standards for mapping the Content / containers into transport mechanisms

Table 5.1 gives an overview of **C**ompleted, **O**ngoing, **P**roposed and **R**equired mappings between containers and transport technologies. Cells which are not filled out are found to be not applicable or are not required in the short term.

Table 5.1: Mapping standards which are required to be developed and standardized for the Container / Transport technology.

Application/ Transport Technology	ITU-R BT.601 Uncompressed baseband video	Content Package	DV (IEC 61834)	DV-Based 25/50 (Proposed)	M-JPEG ([23])	Component Coded (ITU-T, ETSI)	MPEG-2 TS (ISO13818)	SNF (Sony)	AAF / ASF [22] (SMPTE)
SDI (SMPTE)	C [1]								
SDTI (SMPTE)	C [5]	O [18]	P [12]	P [3]	R	R	P [17]	R [8]	R
FC-AV transport [19] **(ANSI)**	O [2]	R [20]		O [2]	R	R	P [2]		R
1394 (Firewire) (IEEE)	R		P [15]	R	R				
ATM (ITU)	R		P [16]	P [16]	R	C [13]	C [4]		R
IP (IETF)			R	R			C [6]		
PDH (ITU)						C [14]	C [9]		
SDH (ITU)	R						C [10]		
DVB ASI, SSI, SPI (ETSI)							C [11]		

<u>Notes to Table 5.1</u>

1. TU-R BT.656 (SMPTE259M).
2. NCITS T11 FC-AV (Project 1237-D).
3. SMPTE Proposed Standard: Data Structure for DV-based Compressed Systems over Serial Digital Transport Interface (SDTI).
4. ITU-T J.82 MPEG-2-TS in ATM.
5. SMPTE 305M (with code=00).
6. IETF RFC 2038 (MPEG-2 over IP).
7. SMPTE Proposed Standards: MPEG-2 Elementary Streams over the Serial Digital Transport Interface (SDTI).
8. Sony Native Format (SNF) in SDTI (publicly available document).
9. ETSI ETS-300 813 DVB Interfaces to PDH Networks.
10. ETSI ETS-300 814 DVB Interfaces to SDH Networks.
11. ETSI EN 50083-9 Interfaces for CATV / SMATV head-ends and similar professional equipment for DVB / MPEG-2 Transport Streams.
12. Draft SMPTE Specifications for DVCAM (IEC 61834) in SDTI.
13. ATM Forum af-vtoa- 0078.000 Circuit Emulation Service Interoperability Specification 2.0.
14. ITU-T J.81: Transmission of Component-Coded Digital Television Signals for Contribution-Quality Applications at the Third Hierarchical Level of ITU-T Recommendation G.702.
    ETSI ETS 300-174: Digital coding of component television signals for contribution quality applications in the range 34 - 45 Mbit/s.
15. IEC 61833.
16. SMPTE Proposed Standard for DV over ATM.
17. SMPTE Proposed Standard for MPEG-2 TS over the Serial Digital Transport Interface (SDTI).
18. SMPTE Proposed Standard for Content Packages over Serial Digital Transport Interface (SDTI). This should include MPEG-2-ES.
19. FC-AV standards activity includes a simple container, a fast file transfer protocol, and a streaming protocol.
20. MPEG-2-ES need to be mapped into FC-AV simple container.
21. Wrappers are defined in the section on Wrappers and Metadata. These need to be transported over the recommended transport mechanisms.
22. AAF / ASF over IP. For SDTI applications a mapping of IP over SDTI needs to be defined.
23. The M-JPEG container specification is not available yet and needs to be standardized.

Note: The MPEG-4 file format derived from Apple QuickTime needs to be considered in further standardization.

5.7.4. Interfaces / networks

For streaming in a point-to-point and point-to-multipoint studio environment, use of the SDTI interface is recommended. Further enhanced versions of SDTI should be investigated and standardized if necessary (this includes also the mappings of applications into SDTI).

Bridging of specific containers between different transport mechanisms needs to be developed. For example, the transport between studios of DV-encapsulated-in-SDTI requires bridging over ATM.

The implementation of the Fibre Channel FC-AV standard requires the development of new Fibre Channel functionalities in order to supporting functions such as bandwidth reservation. This work should be started as soon as possible and should be supervised by the EBU and SMPTE working groups.

The use of ATM and SDH / SONET in a WAN environment requires the implementation of mechanisms to overcome network jitter and wander, such that the resultant studio signal meets the ITU-R BT.656 and ITU-R BT.470 standards. Care should be taken in selecting equipment which ensures that the specifications mentioned above are met. Guidelines and recommendations need to be developed to help users to determine the wander and jitter performance of the equipment used, as well as the overall functionalities (signalling, AAL, QoS, etc.) that are required for real-time streaming. Moreover, the TFHS has identified two different ATM adaptation layer methods (AAL 1 and AAL 5). A recommendation needs to be developed for the streaming of signals over ATM in a compatible way. Also, the development of a broadcast-specific adaptation layer (AALx) needs to be considered.

Annex A

Abbreviations and specialized terms (Glossary)

Throughout this document, specialized terms are used and defined as and when they occur. In addition to these, many of the words and acronyms used are a jargon familiar to those in the television production, post-production, broadcasting, telecommunications and computer industries. To assist readers who are unfamiliar with these terms, an alphabetic listing of some of the more common terms is given below. The terms given in **bold** text in the right-hand column are separately defined in this glossary.

-A-	
A/D	Analogue-to-digital conversion.
A/V	Audio/Video or Audiovisual. This abbreviation is often used on the socketry of consumer equipment.
AAL	ATM adaptation layer. The AAL translates digital voice, images, video and data signals into the **ATM** cell format and vice versa. Five AALs are defined: • AAL1 supports **connection-oriented** services needing constant bit-rates (**CBR**s) and specific timing and delay requirements (e.g. DS-3 circuit). • AAL2 supports connection-oriented services needing variable bit-rates (**VBR**s), e.g. certain video transmission schemes. • AAL3/4 supports both **connectionless** and connection-oriented variable-rate services. • AAL5 supports connection-oriented variable-rate data services. Also known as Simple and Efficient Adaptation Layer (SEAL).
Access setup time	The amount of time taken to set up a transmission path between a source and a destination from the moment of commencing the connection process.
Adaptive predictor	A predictor whose estimating function is made variable according to the short-term spectral characteristics of the **sampled** signal. For **ADPCM** in particular, an adaptive predictor is a time-varying process that computes an estimate of the input signal from the **quantized** difference signal.
Adaptive quantizing	Quantizing in which some parameters are made variable according to the short-term statistical characteristics of the **quantized** signal.
Address Translation	The process of converting external addresses into standardized **network** addresses and vice versa. It facilitates the interconnection of multiple networks in which each have their own addressing scheme.
ADPCM	Adaptive Differential Pulse Code Modulation. A compression algorithm that achieve bit-rate reduction through the use of **adaptive prediction** and **adaptive quantization**.
AES3-1985	The AES Recommended Practice for Digital Audio Engineering – a Serial Transmission Format for Linearly Represented Digital Audio Data. This is a major digital audio standard for serial interface transfer. It is substantially identical to **EBU** Tech. 3250-E, CCIR Rec. 647, SP/DIF, IEC 958, EIA CP340 and EIA DAT. These standards describe a uni-directional, self- clocking, two-channel standard based on a single serial data signal. The AES format contains audio samples up to 24 bits in length and non-audio data including channel status, user data, parity and sample validity. The differences between these standards lie in electrical levels, connectors, and the use of channel status bits. The AES3 standard is better known as the AES /EBU serial digital audio interface.
Analogue (video) signal	A (video) signal, one of whose characteristic quantities follows continuously the variations of another physical quantity representing information.
Analogue transmission	A type of transmission in which a continuously variable signal encodes an infinite number of values for the information being sent (compare with "digital").
Anisochronous	The essential characteristic of a time-scale or a signal, such that the time intervals between consecutive significant instants do not necessarily have the same duration or durations that are integral multiples of the shortest duration.
ANSI	The American National Standards Institute is a US-based organization that develops standards and defines **interfaces** for telecommunications systems.

API — Application Programming Interface. A set of **interface** definitions (functions, subroutines, data structures or **class** descriptions) which together provide a convenient interface to the functions of a subsystem and which insulate the **application** from the minutiae of the implementation.

Application — A computer program designed to perform a certain type of work. An application can manipulate text, numbers, graphics or a combination of these elements.
An application differs from an **operating system** (which runs a computer), a utility (which performs maintenance or general-purpose chores) and a programming language (with which computer programs are created).

Application layer — The seventh and highest layer in the International Organization for Standardization's Open Systems Interconnection (**OSI**) model. The application layer contains the signals sent during interaction between the user and the **application**, and that perform useful work for the user, such as **file** transfer.

ASCII — American Standard Code for Information Interchange. A coding scheme that assigns numeric values to letters, numbers, punctuation marks and certain other characters. By standardizing the values used for these characters, ASCII enables computers and computer programs to exchange information. Although it lacks accent marks, special characters and non-Roman characters, ASCII is the most universal character-coding system.

Asset — An Asset is any material that can be exploited by a **broadcaster** or service provider. An asset could therefore be a complete **programme file**, or it could be a part of a programme, individual sound, images etc.

Asset transfer — The transfer of an **Asset** from one location to another

Asynchronous transmission — A term used to describe any transmission technique that does not require a common **clock** between the two communicating **devices**, but instead derives **timing signals** from special bits or characters (e.g. start/stop bits, flag characters) in the data stream itself. The essential characteristic of time-scales or signals such that their corresponding significant instants do not necessarily occur at the same average rate.

ATM — Asynchronous Transfer Mode. A form of **digital transmission** based on the transfer of units of information known as **cells**. It is suitable for the transmission of images, voice, video,and data.

ATM Layer — The protocol layer that relays **cells** from one **ATM** node to another. It handles most of the processing and routing activities including: each cell's ATM header, cell **mux**ing/**demux**ing, header validation, payload-type identification, Quality of Service (**QoS**) specification, prioritization and flow control.

ATSC — (US) Advanced Television System Committee.

-B-

Back channel — The segment of a two-way communications system that flows from the consumer back to the **Content provider**, or to a system component, to provide feedback.

Backbone — The top level in a hierarchical **network**.

Bandwidth — The frequency range of an electromagnetic signal, measured in hertz (cycles per second). The term has come to refer more generally to the capacity of a **channel** to carry information, as measured in data transferred per second. Transfer of digital data, for example, is measured in bits per second.

Bandwidth reservation — The process of setting aside **bandwidth** on a specific broadcast **channel** for a specific data transmission. A **Content server application** reserves bandwidth on a Microsoft **Broadcast Router** by calling the msbdnReserveBandwidth function. This function forwards the request to a Microsoft® Bandwidth Reservation Server. The server returns a unique reservation identifier if the bandwidth can be reserved.

Baseband — Describes transmissions using the entire spectrum as one **channel**. Alternatively, baseband describes a communication system in which only one signal is carried at any time. An example of the latter is a composite video signal that is not modulated to a particular television channel.

Baud — Number of bits per second, a measure of data-transmission speed. Baud was originally used to measure the transmission speed of telegraph equipment but now most commonly measures modem speeds. The measurement is named after the French engineer and telegrapher, Jean Maurice-Emile Baudot.

BER — Bit Error Ratio (or Rate).

B-frame — MPEG-2 B-frames use bi-directionally-interpolated motion prediction to allow the decoder to rebuild a **frame** that is located between two reconstructed display frames. Effectively the B-frame uses both past frames and future frames to make its predictions. B-frames are not used as reference frames but for further predictions. However, they require more than two frames of video storage in the decoder, which can be a disadvantage in systems where low cost is of the essence. By using bi-directional prediction, B-frames can be coded more efficiently than **P-frames**, allowing a reduction in video bit-rate whilst maintaining subjective video quality.

Broadband — A service or system requiring transmission channels capable of supporting rates greater than the Integrated Services Digital Network (**ISDN**) primary rate (1.544 Mbit/s (e.g. USA) or 2.048 Mbit/s (e.g. Europe)). Broadband is also sometimes used to describe high-speed **networks** in general.

Broadcast	In general terms, a transmission sent simultaneously to more than one recipient. There is a version of broadcasting used on the Internet known as **multicast**. In multicast, each transmission is assigned its own **Internet** Protocol (**IP**) multicast address, allowing clients to filter incoming data for specific **packets** of interest.
Broadcast (Messages)	Transmissions sent to all stations (or nodes, or devices) attached to the **network**.
Broadcast Router	A component that enables a **Content server** to send a data stream to a multiplexer (**MUX**) or other broadcast output device. A Broadcast Router calls a virtual **interface** to transmit a stream at the appropriate rate and in the appropriate **packet** format.
Broadcaster (Service Provider)	An organization which assembles a sequence of events or **programmes**, based upon a schedule, to be delivered to the viewer.
Buffer	An area of storage that provides an uninterrupted flow of data between two computing **devices**.
BWF	Broadcast Wave File. The **EBU** has defined a file format which contains the minimum information that is considered necessary for all broadcast **applications**. The basic information, together with the audio data, is organized as "Broadcast Wave Format" (BWF) **files**. From these files, using an **object**-oriented approach, a higher level descriptor can be used to reference other files containing more complex sets of information which can be assembled for the different specialized kinds of applications.

-C-

CA	Conditional Access. A system to control subscriber access to **services**, **programmes** and events.
CBO	Continuous Bit-stream Oriented. Services which require an ordered and uninterrupted sequence of data to represent them. **PCM**-coded video is an example of a CBO service.
CBR	Constant bit rate. A type of traffic that requires a continuous, specific amount of **bandwidth** (e.g. digital information such as video and digitized voice).
CCITT	The Consultative Committee on International Telephony and Telegraphy, part of the **ITU**, develops standards and defines **interfaces** for telecommunications systems.
Cell	A transmission unit of fixed length, used in cell relay transmission techniques such as **ATM**. An ATM cell is made up of 53 bytes (octets) including a 5-byte header and a 48-byte data payload.
Cell Relay	Any transmission technique that uses **packets** of a fixed length. **ATM**, for example, is a version of the cell relay technique, using 53-byte cells. Other versions use **cells** of a different length.
CEPT	The Conference on European Post and Telegraph is a European organization that develops standards and defines **interfaces** for telecommunications systems.
Channel	A means of unidirectional transmission of signals between two points.
CHAP	Challenge Handshake Authentication Protocol.
Chip-set	Several integrated circuits (ICs) which work together to perform a dedicated task. Subsequent development of the chip-set usually decreases the number of ICs needed, and often a single IC implementation is achieved.
Chroma / chrominance	The colour portion of the video signal that includes hue and saturation information. Hue refers to a tint or shade of colour. Saturation indicates the degree to which the colour is diluted by **luminance** or illumination.
Chunking	The process of "chunking" converts a large **file** into two or more smaller ones.
Circuit Switching	A switching technique in which a dedicated path is set up between the transmitting **device** and the receiving device, remaining in place for the duration of the connection (e.g. a telephone call is a circuit-switched connection).
Class	In general terms, a category. In programming languages, a class is a means of defining the structure of one or more **objects**.
Class Driver	A standard driver provided with the **operating system** that provides hardware-independent support for a given **class** of **devices**. Such a driver communicates with a corresponding hardware-dependent minidriver, using a set of device control requests defined by the operating system. These requests are specific to the particular device class. A class driver can also define additional device control requests itself. A class driver provides an **interface** between a minidriver and the operating system.
Client	Generally, one of a group of computers that receive shared information sent by a computer called a **server** over a broadcast or point-to-point **network**. The term client can also apply to a software process, such as an Automation client, that similarly requests information from a server process and that appears on the same computer as that server process, or even within the same **application**.
Clock	Equipment that provides a **timing signal**.
Closed Captioning	Real-time, written annotation of the currently displayed audio **Content**. Closed Captioning – mainly used in 525-line countries – usually provides subtitle information to hearing-impaired viewers or to speakers of a language other than that on the audio track.
Codec	A combination of an encoder and a decoder in the same equipment.

COM	Component Object Model. An **object**-oriented programming model for building software **applications** made up of modular components. COM allows different software **modules**, written without information about each other, to work together as a single application. COM enables software components to access software services provided by other components, regardless of whether they involve local function calls, **operating system** calls or **network** communications.
Component (Elementary Stream)	One or more entities which together make up an event, e.g. video, audio, teletext.
Compression	The process of reducing the number of bits required to represent information, by removing redundancy. In the case of information **Content** such as video and audio, it is usually necessary to extend this process by removing, in addition, any information that is not redundant but is considered less important. Compression techniques that are used include: blanking suppression, **DPCM**, sub-Nyquist sampling, transform coding, statistical coding, sub-band coding, vector coding, run length coding, variable length coding, fractal coding and wavelet coding.
Connectionless	A type of communication in which no fixed path exists between a sender and receiver, even during a transmission (e.g. **packet switching**). Shared media **LAN**s are connectionless.
Connection-oriented	A type of communication in which an assigned path must exist between a sender and a receiver before a transmission occurs (e.g. **circuit switching**). **ATM** networks are connection-oriented.
Content	Programme Content can be *Video Essence*, *Audio Essence*, *Data Essence* and *Metadata*. Content can therefore include television programming, data and software **applications**.
Content provider	A person or company delivering broadcast **Content**.
CPU	Central Processing Unit. In a personal computer, the CPU is the microprocessor which *is* the computer.
CRC	Cyclic Redundancy Check. A common technique for detecting errors in data transmission. In CRC error checking, the sending **device** calculates a number based on the data transmitted. The receiving device repeats the same calculation after transmission. If both devices obtain the same result, it is assumed the transmission was error-free. The procedure is known as a redundancy check because each transmission includes not only data but additional, redundant values for error checking.
CVD	Cell Delay Variation. A measurement of the allowable variation in delay between the reception of one **cell** and the next, usually expressed in thousandths of a second, or milliseconds (ms). Important in the transmission of voice and video traffic, CDV measurements determine whether or not cells are arriving at the far end too late to reconstruct a valid **packet**.

-D-

D/A	Digital-to-analogue conversion.
Data Link layer	The second of the seven layers in the International Organization for Standardization's Open Systems Interconnection (OSI) model for standardizing communications. The Data Link layer is one level above the **Physical layer**. It is involved in packaging and addressing information and in controlling the flow of separate transmissions over communications lines. The Data Link layer is the lowest of the three layers (Data Link, Network and Transport) that help to move information from one **device** to another. There is also a Data Link layer in the **EBU** / **SMPTE** four-layer **object** model.
Data service	A mechanism offered by a **broadcster (service provider)** for sending broadcast data to broadcast **clients**. Such data can include **Programme** Guide information, **WWW** pages, software and other digital information. The data service mechanism can be any broadcast process.
Data streaming	The data broadcast specification profile for data streaming supports data broadcast services that require a **streaming**-oriented, end-to-end delivery of data in either an **asynchronous**, **synchronous** or synchronized way through broadcast **networks**. Data which is broadcast according to the data streaming specification is carried in **Programme** Elementary Stream (PES) **packets** which are defined in MPEG-2 Systems. Asynchronous data streaming is defined as the streaming of only data without any timing requirements (e.g. RS-232 data). Synchronous data streaming is defined as the streaming of data with timing requirements in the sense that the data and **clock** can be regenerated at the receiver into a synchronous data stream (e.g. E1, T1). Synchronized data streaming is defined as the streaming of data with timing requirements in the sense that the data within the stream can be played back in synchronization with other kinds of data streams (e.g. audio, video).
Datagram	One **packet** of information and associated delivery information, such as the destination address, that is routed through a **packet-switching network**. In a packet-switching network, data packets are routed independently of each other and may follow different routes and arrive in a different order from which they were sent. An **Internet** Protocol (**IP**) **multicast packet** is an example of a datagram.

DAVIC	Digital Audio VIsual Council. DAVIC has been convened along similar lines to **MPEG** but with no affiliation to a standards body; it therefore has the status of a world-wide industry consortium. Its purpose is to augment MPEG and to collect system specifications for the delivery of a range of audio-visual services which can be applied uniformly on a world-wide basis.
DCT	Discrete Cosine Transform. A DCT process basically involves dividing the picture up into 8 x 8 pixel blocks, then replacing the discrete **luminance** and **chrominance** values of each pixel by the amplitudes of the corresponding frequency components for the horizontal and vertical directions respectively. In this way, the information is transformed from the time domain to the frequency domain. No information is lost in this process, except perhaps by the rounding of the last digit of the frequency coefficient values.
Delivery system	The physical medium by which one or more multiplexes (**MUX**s) are transmitted, e.g. a satellite system, wide-band coaxial cable, fibre optics, terrestrial channel of one emitting point.
DEMUX	Demultiplexer. A **device** that performs the complementary operation to that of a multiplexer (**MUX**).
Descrambler	A **device** that performs the complementary operation to that of a **scrambler**.
Device	A unit of hardware, for example a videotape machine or a **server**.
Device class	A group into which **devices** are placed for the purposes of installing and managing **device drivers**, and for allocating resources.
Device driver	A software component that allows an **operating system** to communicate with one or more specific hardware **devices** attached to a computer.
Device object	A programming **object** used to represent a physical, logical or virtual hardware **device** whose **device driver** has been loaded into the **operating system**.
DIF	Digital InterFace. All the DV-based **compression** schemes share the so-called DIF structure which is defined in the "Blue Book" (IEC 61834).
Digital (transmission) channel	The means of unidirectional **digital transmission** of digital signals between two points.
Digital connection	A concatenation of **digital transmission channels**, switching and other functional units, set up to provide for the transfer of **digital signals** between two or more points in a **network**, in support of a single communication.
Digital demultiplexing	The separation of a (larger) **digital signal** into its constituent **digital channels**.
Digital multiplexing	A form of time-division-multiplexing applied to **digital channels** by which several **digital signals** are combined into a single (larger) digital signal.
Digital signal	A discretely-timed signal in which information is represented by a number of well-defined discrete values that one of its characteristic quantities may take in time.
Digital transmission	The transmission of **digital signals** by means of a **channel** or channels that may assume, in time, any one of a defined set of discrete states.
Digital-S	JVC trademark. This tape system is based on DV technology, and has a data-rate of 50 Mbit/s.
Downstream	One-way data flow from the **head-end** to the **broadcast client**.
DPCM	Differential Pulse Code Modulation. A process in which a signal is **sampled**, and the difference between each sample of this signal and its estimated value is **quantized** and converted by encoding to a **digital signal**.
DSP	Digital signal processor.
DTS	Data Time Stamp.
DV	Digital Video. A digital videotape format originally conceived for consumer applications.
DVB	Digital Video Broadcasting.
DVB-C	DVB **framing** structure, **channel** coding and modulation scheme for cable systems (EN 300 429).
DVB-S	DVB baseline system for digital satellite television (EN 300 421).
DVB-T	DVB baseline system for digital terrestrial television (EN 300 744).
DVC	Digital Video Cassette.
DVCPRO, DVCPRO50	Panasonic trademarks. Based on DV technology and having a data-rate of 25 Mbit/s and 50 Mbit/s respectively.
DVD	Digital Versatile (Video) Disk

-E-

EBU	European Broadcasting Union. Headquartered in Geneva, Switzerland, the EBU is the world's largest professional association of national broadcasters. Following a merger on 1 January 1993 with the International Radio and Television Organization (OIRT) – the former association of Socialist Bloc broadcasters – the expanded EBU has 66 active members in 49 European and Mediterranean countries, and 51 associate members in 30 countries elsewhere in Africa, the Americas,and Asia.
ECM	Entitlement Control Message.
Enhancement	A **multimedia** element, such as a hypertext link to a **WWW** page, a graphic, a text frame, a sound or an animated sequence, added to a broadcast show or other video **programme**. Many such elements are based on Hypertext Markup Language (**HTML**).
EPG	Electronic Programme Guide.
Error ratio [error rate]	The ratio of the number of digital errors received in a specified period to the total number of digits received in the same period.
Error, digital error	An inconsistency between a digit in a transmitted **digital signal** and the corresponding digit in the received digital signal.
ESCR	Elementary Stream Clock Reference.
ETR	ETSI Technical Report.
ETS	European Telecommunication Standard.
ETSI	European Telecommunications Standards Institute.

-F-

FEC	Forward error correction. A system of error correction that incorporates redundancy into data so that transmission errors can, in many cases, be corrected without requiring retransmission.
Field	In broadcast television, one of two sets of alternating lines in an **interlaced** video **frame**. In one field, the odd-numbered lines of video are drawn on the screen; in the other, the even-numbered lines are drawn. When interlaced, the two fields combine to form a single frame of on-screen video.
File	An organized collection of related records, accessible from a storage device via an assigned address. The relationship between the records and the file may be that of common purpose, format or data source, and the records may or may not be sequenced.
Frame	In broadcast television, a single screen-sized image that can be displayed in sequence with other slightly different images to animate drawings. In the case of **NTSC** video, a video frame consists of two **interlaced fields** of 525 lines; NTSC video runs at 30 frames per second. In the case of **PAL** or **SECAM** video, a video frame consists of two interlaced fields of 625 lines; PAL and SECAM video run at 25 frames per second. By way of comparison, film runs at 24 frames per second. A variable-length packet of data is used by traditional **LAN**s such as Ethernet and Token Ring, as well as **WAN** services such as X.25 or Frame Relay. An edge switch will take frames and divide them into fixed-length **cells** using an **AAL** format. A destination edge switch will take the cells and reconstitute them into frames for final delivery.
FTP	File Transfer Protocol. A protocol that supports file transfers to and from remote systems on a **network** using Transmission Control Protocol / Internet Protocol (**TCP/IP**), such as the **Internet**. FTP supports several commands that allow the bi-directional transfer of binary and **ASCII** files between systems.
FTP+	FTP+ is an enhanced version of **FTP**, and uses the same base set of commands. FTP+ includes new commands that enable traditional features and which also provide the ability to embrace **network** protocols other than **IP**.

-G-

Gbit/s	Gigabit per second. A digital transmission speed of billions of (i.e.10^9) bits per second.
Genre	A category of broadcast **programmes**, typically related by style, theme or format, e.g. TV movies or television series.
GoP	Group of Pictures. An MPEG-2 GoP begins with an **I-frame** and extends to the last **frame** before the next I-frame. The GoP sequence is known as an open GoP – the last frame in the GoP uses the first frame of the next GoP as a reference. Another type of GoP is a closed GoP, which has no prediction links to the next GoP and, by definition, always ends in a **P-frame**.
GSM	Global System for Mobile communication.
Guaranteed bandwidth	**Bandwidth** that is reserved only if the requested bandwidth is available for the requested period. Once reserved, such bandwidth can be relied upon to be available.

-H-

HDTV	High Definition TeleVision. Television that is delivered at a higher screen resolution than that of **NTSC**, **PAL** or **SECAM**.

Head-end	The origin of signals in a terrestrial, cable, satellite or network broadcast system. In Broadcast Architecture, the **server** infrastructure that gathers, coordinates and broadcasts the data is generally located at the broadcast head-end.
HEX	Hexadecimal. A numbering system with a base of 16 (binary numbers have a base of 2, and decimal numbers have a base of 10). In HEX notation, the decimal numbers 0 to 9 are extended by the addition of the uppercase letters A to F, i.e. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F (which is equivalent to the numbers 0 to 15 in decimal notation).
Host	A **device** where one or more **modules** can be connected, e.g. a VTR, a PC ...
HTML	Hypertext Mark-up Language. A mark-up language used to create hypertext documents that are portable from one platform to another. HTML files are text files with embedded codes, or mark-up tags, that indicate formatting and hypertext links. HTML is used for formatting documents on the **WWW**.
HTTP	Hypertext Transport Protocol. The underlying, **application**-level protocol by which **WWW clients** and **servers** communicate on the **Internet**.

-I-

ID	Identifier
IDL	Interface Definition Language. Used to describe **interfaces** that **client objects** call, and object implementations provide. It is a purely descriptive language which has mappings provided for several programming languages such as C++, C and Java. It has the same lexical rules as C++.
IEC	International Electrotechnical Commission. Based in Geneva, the IEC is the world organization that prepares and publishes international standards for all electrical, electronic and related technologies.
IEEE	(US) Institute of Electrical and Electronic Engineers. The world's largest technical professional society, with more than 320,000 members. The technical objectives of the IEEE focus on advancing the theory and practice of electrical, electronic and computer engineering, and computer science.
IETF	Internet Engineering Task Force. The IETF is a large open international community of **network** designers, operators, vendors and researchers concerned with the evolution of the **Internet** architecture and the smooth operation of the Internet. It is open to any interested individual.
I-Frame	Intra-coded Frame. I-frame pictures make use only of information already contained within that **frame**. They are not dependent on other frames and can act as the starting point to enable decoders to begin working on a **GoP** containing a sequence of other types of frame. The amount of **compression** achievable is typically less than for the other types of frame.
IIOP	Internet Inter-**ORB** Protocol.
Interactive television	The interactive combination of a video **programme** and **multimedia enhancement** elements such as hypertext links, graphics, text frames, sounds and animations.
Interface	The common boundary point where two elements connect so that they can work with one another. In computing, the connection between an **application** and an **operating system** or between an application and a user (the user interface) are examples of an interface. In C++ programming, an interface is a collection of related methods exposed by a given class of **objects**. These methods are procedures that can be performed on or by those objects.
Interlacing / interlaced	A video display technique, used in current analogue televisions, in which the electron beam refreshes (updates) all odd-numbered scan lines in one **field** and all even-numbered scan lines in the next. Interlacing takes advantage of both the screen phosphor's ability to maintain an image for a short period of time before fading, and the human eye's tendency to average subtle differences in light intensity. By refreshing alternate lines, interlacing halves the number of lines to update in one screen sweep. An alternative video display technique, used in computer monitors, is *progressive scanning*. In progressive scanning, the image is refreshed one line at a time.
Internet	Generically, a collection of **networks** interconnected with **routers**. The Internet is the largest such collection in the world. It has a three-level hierarchy composed of **backbone** networks, mid-level networks and stub networks.
IOR	Interoperable Object Reference.
IP	Internet Protocol. The primary network layer of **Internet** communication, responsible for addressing and routing **packets** over the **network**. IP provides a best-effort, **connectionless** delivery system that does not guarantee that packets arrive at their destination or that they are received in the sequence in which they were sent.
IP Address	An identifier for a **network** node; expressed as four **fields** separated by decimal points (e.g. 136.19.0.5.); IP address is site-dependent and assigned by a network administrator.
IPCP	Internet Protocol Control Protocol.

IP-over-ATM	The adaptation of **TCP/IP** and its address resolution protocol for transmission over an **ATM** network. It is defined by the **IETF** in RFCs 1483 and 1577. It puts **IP packets** and ARP requests directly into protocol data units and converts them to ATM **cells**. This is necessary because IP does not recognize conventional **MAC**-layer protocols, such as those generated on an Ethernet **LAN**.
IS	Interactive Service.
ISDN	Integrated Services Digital Network. A type of dial-up service. Data can be transmitted over ISDN lines at speeds of 64 or 128 kbit/s, whereas standard phone lines generally limit modems to top speeds of 20 to 30 kbit/s.
ISO	International Organization for Standardization, based in Geneva.
Isochronous	A term used to describe signal-timing techniques that require a uniform reference point (usually embedded in the data signal).
ITU	International Telecommunication Union, part of the United Nations, based in Geneva.

-J-

Java	An **object**-oriented, platform-independent computer programming language developed by Sun Microsystems. The Applet subclass of Java can be used to create **Internet applications**.
Jitter	Short-term non-cumulative variations of the significant instants of a **digital signal** from their ideal positions in time.
Jitter, delay, latency	See *Latency*

-K-

kbit/s	kilobits per second. A digital transmission speed expressed in thousand of bits per second.

-L-

LAN	Local Area Network. A **network** dispersed over a relatively limited area and connected by a communications link that enables each **device** on the network to interact with any other.
LAN Emulation	The process of implementing enough of the media access control layer protocol of a **LAN** (e.g. Ethernet or Token Ring) to allow existing higher layer protocols (and **applications**) to be used unchanged over another **network**, such as an **ATM** network.
Latency	The time delay inherent in a manipulative process. In particular, the time that it takes to process an input bitstream through a **compression** and decompression process. Buffering and transmission can be major contributors to processing delays.
Link	Any physical connection on a **network** between two separate **devices**, such as an **ATM** switch and its associated end point or end **station**.
Log on	To provide a user name and password that identifies you to a computer **network**.
LSB	Least Significant Bit. In any related grouping of bits (i.e. a *word*), there will be one which quantifies the zeroth power of 2 (i.e. the value is 0 or 1). This bit is the LSB of the word.
Luminance	A measure of the degree of brightness or illumination radiated by a given source. Alternatively, the perceived brightness component of a given colour, as opposed to its **chroma**.

-M-

MAA	MPEG ATM Adaptation.
MAC	Media Access Control.
MAN	Metropolitan area network.
Master clock	A **clock** that is used to control the frequency of other clocks.
Mbit/s	Megabits per second. A digital transmission speed expressed in millions of bits per second.
MBONE	Multicast backbone. A virtual, **multicast**-enabled **network** that works on top of the **Internet**. The most popular application for the MBONE is video conferencing, including audio, video and whiteboard conferencing. However, the essential technology of the MBONE is simply multicast – there is no special support for continuous media such as audio and video. The MBONE has been set up and maintained on a co-operative, volunteer basis.
Metadata	Data describing other data.
MIB	Management Information Base.
MIME	Multipurpose Internet Mail Extensions.
MJD	Modified Julian Date.
MMDS	Microwave Multi-point Distribution Systems (or Multichannel Multi-point Distribution Systems). Also known as wireless cable.
MMI	Man Machine Interface. The MMI of a door is its doorknob. That of a PC is a combination of keyboard, mouse and monitor.

Module	A small **device**, not working by itself, designed to run specialized tasks in association with a host – for example, a conditional access sub system, or an electronic **programme** guide application module – or to provide resources required by an **application** but not provided directly by the host
MPEG	Motion Picture Experts Group. MPEG-1 is a standard designed for video playback from CD-ROM. It provides video and audio **compression** at rates up to 1.8 Mbit/s. MPEG-2 refers to the ISO/IEC 13818 standard, and it provides higher video resolutions and **interlacing** for broadcast television and high-definition television (**HDTV**). Both standards were created by the Motion Pictures Experts Group, an International Standards Organization / International Telegraph and Telephone Consultative Committee (ISO/CCITT) group set up to develop motion video compression standards. The MPEG system makes use of three different types of compressed video **frames**, (**I, P and B frames**), which are stored so as to enable temporal prediction of missing or incomplete frames as received by the decoder.
MPEG TS	MPEG Transport Stream.
MPI	MPEG Physical Interface.
MPLS	Multi-protocol Label Swapping.
MSB	Most Significant Bit. In any related grouping of bits (i.e. a *word*), there will be one which quantifies the largest power of 2. This bit is the MSB of the word.
MTU	Multiport Transceiver Unit.
Multicast	A point-to-many networking model in which a **packet** is sent to a specific address, and only those computers that are set to receive information from this address receive the packet. On the **Internet**, the possible **IP** multicast addresses range from 224.0.0.0 through 239.255.255.255. Computer **networks** typically use a **unicast** model, in which a different version of the same packet is sent to each address that must receive it. The multicast model greatly reduces traffic and increases efficiency on such networks.
Multicast Messages	A subset of "broadcast" in which a transmission is sent to all members of a pre-defined group of **stations**, nodes or **devices**.
Multimedia	Online material that combines text and graphics with sound, animation or video, or some combination of the three.
Multipoint	Generally encountered in the term "point-to-multipoint" which describes a broadcast topography.
MUX	Multiplex or multiplexer. A **stream** of all the digital data carrying one or more **services** within a single physical **channel**. In general terms, a multiplexer is a device for funnelling several different streams of data over a common communications line. In the case of broadcasting, a multiplexer combines multiple television channels and data streams into a single broadcast.
MVDS	Multipoint Video Distribution System.

-N-

NE	Network Element.
Network	In computing, a data communications system that interconnects a group of computers and associated **devices** at the same or different sites. In broadcasting, a collection of MPEG-2 Transport Stream multiplexes that are transmitted on a single delivery system, e.g. all the **digital channels** on a specific satellite or cable system.
NFS	Network File System. This is defined in RFC 1813. **File** system access is different from file transfer, in that **Network** File Systems generally employ a **client**-**server** model in which the server computer actually has the file system as local data. The client-host is allowed to "mount" the network file system to get access to the directories and files as if they were locally available. Multiple clients are permitted to simultaneously "mount" the server's file system and get access to its **Content**.
NCITS	National Committee for Information Technology Standards. NCITS T11 is responsible for standards development in the areas of Intelligent Peripheral Interface (IPI), High-Performance Parallel Interface (HIPPI) and Fibre Channel (FC).
NNI	Network-to-Network Interface. In an ATM network, the interface between one ATM switch and another, or an ATM switch and a public ATM switching system.
NTSC	National Television System Committee – which originated the NTSC standard for analogue television signals in North America, and which has also been adopted in Japan and parts of South America. The NTSC system is based on a power supply frequency of 60 Hertz (Hz) and can display 525 scan lines at approximately 30 **frames** per second. However, non-picture lines and **interlaced** scanning methods make for an effective resolution limit of about 340 lines. The bandwidth of the system is 4.2 Megahertz (MHz).

-O-

Object	A computer programming term describing a software component that contains data or functions accessed through one or more defined **interfaces**. In Java and C++, an object is an instance of an object class.
Octet	A group of eight binary digits or eight signal elements capable of representing 256 different values operated upon as an entity (also known as a "word").
Operating system	Software responsible for controlling the allocation and usage of computer hardware resources such as memory, **CPU** time, disk space and peripheral **devices**.
Opportunistic bandwidth	**Bandwidth** granted whenever possible during the requested period, as opposed to **guaranteed bandwidth** which is actually reserved for a given transmission.
OSI	Open Systems Interconnection. This refers to the ISO / OSI seven layer model for standardizing communications.

-P-

Packet	A unit of information transmitted as a whole from one **device** to another on a **network**. In **packet-switching** networks, a packet is defined more specifically as a transmission unit of fixed maximum size that consists of binary digits (bits) representing both data and a header containing an identification number, source and destination addresses, and sometimes error-control data.
Packet Switching	A switching technique in which no dedicated path exists between the transmitting **device** and the receiving device. Information is formatted into individual **packets**, each with its own address. The packets are sent across the **network** and reassembled at the receiving **station**.
PAL	Phase Alternation by Line standard. The analogue television standard for much of Europe – except France, Russia and most of Eastern Europe, which use **SECAM**. As with SECAM, PAL is based on a 50 Hertz (Hz) power supply frequency, but it uses a different encoding process. It displays 625 scan lines and 25 **frames** per second, and offers slightly better resolution than the **NTSC** standard used mainly in North America and Japan. The PAL **bandwidth** is 5.5 Megahertz (MHz).
Partial Transport Stream	Bitstream derived from an **MPEG-2 TS** by removing those TS **packets** that are not relevant to one particular selected **programme**, or a number of selected programmes.
PCM	Pulse Code Modulation. A process in which a signal is **sampled**, and each sample is **quantized** independently of other samples and converted by encoding to a **digital signal**.
PCR	Programme Clock Reference.
PDH	Plesiochronous Digital Hierarchy.
PDU	Protocol Data Unit. A unit of information (e.g. a **packet** or **frame**) exchanged between peer layers in a **network**.
PES	Packetized Elementary Stream.
P-frame	MPEG-2 P-frames use a single previously-reconstructed **frame** as the basis for temporal prediction calculations; they need more than one video frame of storage. Effectively the P-frame uses the nearest previous frame (I or P) on which to base its predictions, and this is called forward prediction. P-frames serve as the reference frame for future P- or **B-frames**, but if errors exist in a particular P-frame, they may be carried forward to the future frames derived from them. P-frames can provide a greater degree of compression than **I-frames**.
Physical Layer	The first of the seven layers in the International Organization for Standardization's Open Systems Interconnection (OSI) model for standardizing communications. It specifies the physical **interface** (e.g. connectors, voltage levels, cable types) between a user **device** and the **network**.
PID	**Packet** Identifier.
Plesiochronous	The essential characteristic of time-scales or signals such that their corresponding significant instants occur at nominally the same rate, any variation in rate being constrained within specified limits. Two signals having the same nominal digit rate, but not stemming from the same **clock**, are usually plesiochronous.
PLL	Phase Locked Loop.
Plug and Play	A design philosophy and set of specifications that describe changes to hardware and software for the personal computer and its peripherals. These changes make it possible to automatically identify and arbitrate resource requirements among all **devices** and buses on a computer. Plug and Play specifies a set of application programming interface (**API**) elements that are used in addition to existing driver architectures.
Point-to-point	A term used by **network** designers to describe network links that have only one possible destination for a transmission.

Port	Generally, the address at which a **device** such as a **network interface** card (NIC), serial adapter or parallel adapter communicates with a computer. Data passes in and out of such a port. In Internet Protocol (**IP**), however, a port signifies an arbitrary value used by the Transmission Control Protocol / Internet Protocol (**TCP/IP**) and User Datagram Protocol / Internet Protocol (**UDP/IP**) to supplement an IP address so as to distinguish between different **applications** or protocols residing at that address. Taken together, an IP address and a port uniquely identify a sending or receiving application or process.
PRBS	Pseudo Random Binary Sequence.
Predictor	A **device** that provides an estimated value of a **sampled** signal, derived from previous samples of the same signal or from a **quantized** version of those samples.
Printf	A symbol in the C programming language.
Programme	A concatenation of one or more events under the control of a broadcaster, e.g. a news broadcast, entertainment show.
PSI	MPEG-2 Programme Specific Information (as defined in ISO/IEC 13818-1).
PSK	Phase Shift Keying.
PSTN	Public Switched Telephone Network.
PTS	Presentation Time Stamp.
Push model	A broadcast model in which a **server** sends information to one or more **clients** on its own schedule, without waiting for requests. The clients scan the incoming information, save the parts they have been instructed to save, and discard the rest. Because the push model eliminates the need for requests, it eliminates the need for a **back channel** from the client to the server. The push model contrasts with the *pull model*, in which each client requests information from a server. The pull model is more efficient for interactively selecting specific data to receive, but uses excessive **bandwidth** when many clients request the same information.
PVC	Permanent Virtual Circuit. A generic term for any permanent, provisioned, communications medium. Note that PVC does not stand for permanent virtual channel. In **ATM**, there are two kinds of PVCs: permanent virtual path connections (PVPCs) and permanent virtual channel connections (PVCCs).
-Q-	
QAM	Quadrature Amplitude Modulation.
QoS	Quality of Service. The **ATM** Forum, for example, has outlined five categories of performance (Classes 1 to 5) and recommends that ATM's QoS should be comparable to that of standard **digital connections**.
QPSK	Quadrature Phase Shift Keying.
Quantizing / quantized	A process in which a continuous range of values is divided into a number of adjacent intervals, and any value within a given interval is represented by a single predetermined value within the interval.
Query	A request that specific data be retrieved, modified or deleted.
-R-	
RAID	Redundant Array of Independent Disks. A means of constructing a **server** by interconnecting several hard disk units such that the data is distributed across all of them. If an individual hard disks fails, the remainder can continue working and the defective unit can be replaced, usually without taking the server out of service.
RAM	Random access memory. RAM is semiconductor-based memory within a personal computer or other hardware **device** that can be rapidly read from and written to by a computer's microprocessor or other devices. It does not generally retain information when the computer is turned off.
Reference clock	A **clock** of very high stability and accuracy that may be completely autonomous and whose frequency serves as a basis of comparison for the frequency of other clocks.
Regeneration	The process of receiving and reconstructing a **digital signal** so that the amplitudes, waveforms and timing of its signal elements are constrained within specified limits.
Registry	A hierarchical database that provides a repository for information about a system's hardware and software configuration.
Resource	A unit of functionality provided by the host for use by a **module**. A resource defines a set of **objects** exchanged between the module and the host by which the module uses the resource. An example of a resource is a piece of static data, such as a dialog box, that can be used by more than one **application** or in more than one place within an application. Alternatively, it is any part of a computer or **network**, such as a disk drive, printer or memory, that can be used by a program or process.
RFC	Request For Comment.
RFT	Request for Technology.

RMS / rms	Root Mean Square.
Router	A **device** that helps local-area networks (**LAN**s) and wide-area networks (**WAN**s) to connect and interoperate. A router can connect LANs that have different **network** topologies, such as Ethernet and Token Ring. Routers choose the best path for a **packet**, optimizing the network performance.
RS-422	A serial data **interface** standard. RS-232 has been around as a standard for decades as an electrical interface between Data Terminal Equipment (DTE) and Data Circuit-Terminating Equipment (DCE) such as modems, and is commonly the serial interface found on PCs. The RS-422 interface is a balanced version of the interface, and it is much less prone to interference from adjacent signals.
RSVP	Resource reSerVation Protocol. RSVP is a **QoS** signalling protocol for **application**-level streams. It provides **network**-level signalling to obtain QoS guarantees.
RTP	Real-time Transport Protocol. RTP permits real-time **Content** transport by the inclusion of media-dependent Time Stamps that allow Content synchronization to be achieved by recovering the sending **clock**.

-S-

S/N (SNR)	Signal-to-Noise Ratio. The amount of power by which a signal exceeds the amount of **channel** noise at the same point in transmission. This amount is measured in decibels and indicates the clarity or accuracy with which communication can occur.
Sample	A representative value of a signal at a chosen instant, derived from a portion of that signal.
Sampling / sampled	The process of taking samples of a signal, usually at equal time intervals.
Sampling rate	The number of samples taken of a signal per unit of time.
Satellite uplink	The system that transports a signal up to a satellite for broadcasting. Signals usually come to the uplink through multiplexers (**MUX**s).
SCPC	Single Channel Per Carrier transmission.
Scrambler	A **device** that converts a **digital signal** into a pseudo-random digital signal having the same meaning and the same digit rate.
SDH	Synchronous Digital Hierarchy. International version of **SONET** that is based on 155 Mbit/s increments rather than SONET's 51 Mbit/s increments.
SDTV	Standard Definition TeleVision. Television service providing a subjective picture quality roughly equivalent to current 525-line or 625-line broadcasts.
SECAM	Sequential Couleur á Memoire, or Sequential Colour with Memory. The television standard for France, Russia and most of Eastern Europe. As with **PAL**, SECAM is based on a 50 Hertz (Hz) power supply frequency, but it uses a different encoding process. Devised earlier than PAL, its specifications reflect earlier technical limitations.
Server	A computer or other **device** connected to a **network** to provide a particular service (e.g. print server, fax server, playout server) to **client** devices connected to the network.
Service	A set of elementary streams offered to the user as a **programme**. They are related by a common synchronization. They are made of different data, e.g. video, audio, subtitles and other data. Alternatively, it is a sequence of programmes under the control of a **broadcaster** which can be broadcast as part of a schedule.
Service_id	A unique identifier of a service within a TS.
SI	Service Information. Digital data describing the delivery system, **Content** and scheduling / timing of broadcast data streams etc. It includes MPEG-2 **PSI** together with independently-defined extensions (ETS 300 468).
Signalling (ATM)	The procedures used to establish connections on an **ATM network**. Signalling standards are based on the ITU's Q.93B recommendation.
Slip	The loss or gain of a digit position or a set of consecutive digit positions in a **digital signal**, resulting from an aberration of the timing processes associated with transmission or switching of a digital signal.
SMPTE	(US) Society of Motion Picture and Television Engineers. The Society was founded in 1916, as the Society of Motion Picture Engineers. The T was added in 1950 to embrace the emerging television industry. The SMPTE is recognized around the globe as a leader in the development of standards and authoritative, consensus-based, recommended practices (RPs) and engineering guidelines (EGs). The Society serves all branches of motion imaging including film, video and **multimedia**.
SNMP	Simple Network Management Protocol.
SNMP2	Simple Network Management Protocol version 2. An enhancement of the simple gateway monitoring protocol, and which was designed as a **connectionless application**-level protocol within **TCP/IP** that uses **UDP** as a **Transport layer**.
SONET	Synchronous Optical NETwork. A set of standards for the **digital transmission** of information over fibre optics. Based on increments of 51 Mbit/s.

Station	An establishment equipped for radio or television transmission.
STM	Synchronous Transfer Mode / Synchronous Transport Module. In **ATM**, a method of communications that transmits data streams synchronized to a common **clock** signal (reference clock). In **SDH**, it is "Synchronous Transport Module" and is the basic unit (STM-1 = 155 Mbit/s, STM-4 = 622 Mbit/s, STM-16 = 2.5 Gbit/s) of the Synchronous Digital Hierarchy.
Streaming	A collection of data sent over a data **channel** in a sequential fashion. The bytes are typically sent in small **packets**, which are reassembled into a contiguous stream of data. Alternatively, it is the process of sending such small packets of data.
Streaming architecture	A model for the interconnection of **stream**-processing components, in which **applications** dynamically load data as they output it. Dynamic loading means data can be broadcast continuously.
String	Data composed of a sequence of characters, usually representing human-readable text.
SVC	Switched Virtual Circuit. A generic term for any switched communications medium. Note that SVC does not stand for switched virtual channel. In **ATM**, there are two kinds of SVCs: switched virtual path connections (SVPCs) and switched virtual channel connections (SVCCs).
Switch	**Device** used to route **cells** through an **ATM network**.
Symbol rate	The number of signal elements of the signal transmitted per unit of time. The **baud** is usually used to quantify this, one baud being equal to one single element per second.
Synchronization	The process of adjusting the corresponding significant instants of signals to make them **synchronous**.
Synchronous	A term used to describe a transmission technique that requires a common **clock** signal (or timing reference) between two communicating devices to co-ordinate their transmissions.
Synchronous network	A **network** in which the corresponding significant instants of nominated signals are adjusted to make them **synchronous**.

-T-

Task Scheduler	A scheduling service and user **interface** that is available as a common **resource** within an **operating system**. A Task Scheduler manages all aspects of job scheduling: starting jobs, enumerating currently running jobs, tracking job status, and so on.
TCP	Transmission Control Protocol.
TCP/IP	Transmission Control Protocol / Internet Protocol. A networking protocol that provides reliable communications across interconnected **networks** made up of computers with diverse hardware architectures and **operating systems**. The TCP portion of the protocol, a layer above **IP**, is used to send a reliable, continuous stream of data and includes standards for automatically requesting missing data, reordering IP **packets** that might have arrived out of order, converting IP **datagrams** to a **streaming** protocol, and routing data within a computer to make sure the data gets to the correct **application**. The IP portion of the protocol includes standards for how computers communicate and conventions for connecting networks and routing traffic.
TDM	Time-division multiplexing. Multiplexing in which several signals are interleaved in time for transmission over a common **channel**.
Telecommunication	Any transmission and/or emission and reception of signals representing signs, writing, images and sounds or intelligence of any nature by wire, radio, optical or other electromagnetic systems.
TFHS	The Joint **EBU** / **SMPTE** Task Force for Harmonized Standards for the Exchange of Programme Material as Bitstreams.
Theme	A category to which individual television **programmes** are assigned within the Guide database. A theme allows a programme episode to be associated with multiple **genre** / subgenre pairs.
Timing recovery [timing extraction]	The derivation of a timing signal from a received signal.
Timing signal	A cyclic signal used to control the timing of operations.
Traffic Policing	A mechanism used to detect and discard or modify **ATM cells** (traffic) that do not conform to the Quality of Service (**QoS**) parameters specified in the call setup procedure.
Traffic Shaping	A mechanism used to control traffic flow so that a specified **QoS** is maintained.
Transmission	The action of conveying signals from one point to one or more other points.
Transparency, digital transparency	The property of a **digital transmission channel**, telecommunication circuit or connection, that permits any **digital signal** to be conveyed over it without change to the value or order of any signal elements.

Transport layer	The fourth of the seven layers in the International Organization for Standardization's Open Systems Interconnection (OSI) model for standardizing communications. The Transport layer is one level above the Network layer and is responsible for error detection and correction, among other tasks. Error correction ensures that the bits delivered to the receiver are the same as the bits transmitted by the sender, in the same order and without modification, loss or duplication. The Transport layer is the highest of the three layers (**Data Link**, Network and Transport) that help to move information from one **device** to another.
Transport_stream_id	A unique identifier of a **TS** within an original **network**.
TS	Transport Stream. A TS is a data structure defined in ISO/IEC 13818-1 for the MPEG-2 Transport Stream. It is the basis of the **ATSC** and **DVB** standards.
TV	Television.
Twisted-pair cable	A communications medium consisting of two thin insulated wires, generally made of copper, that are twisted together. Standard telephone connections are often referred to as "twisted pair."

-U-

UDP	User Datagram Protocol. UDP, as defined in RFC 768, can be used as an option to enable bounded-quality transfers on top of the **IP** layer. It allows broadcast transmissions and is a **datagram**-oriented protocol.
UDP/IP	User **Datagram** Protocol / Internet Protocol. A networking protocol used to send large unidirectional **packets** across interconnected **networks** made up of computers with diverse hardware architectures and **operating systems**. The UDP portion of the protocol, a networking layer above **IP**, is used to send unidirectional packets of up to 64 kilobytes in size and includes standards for routing data within a single computer so it reaches the correct **client application**. The IP portion of the protocol includes standards for how computers communicate and conventions for connecting networks and for routing traffic.
UML	Unified Modelling Language. The UML is a language for specifying, visualizing, constructing and documenting the artefacts of software systems. It assists the complex process of software design, making a "blueprint" for construction.
UNI	User-to-Network Interface. A connection that directly links a user's **device** to a **network** (usually through a **switch**). Also, the physical and electrical demarcation point between the user device and the switch.
Unicast	A point-to-point **networking** model in which a **packet** is duplicated for each address that needs to receive it.
UNO-CDR	Universal Networked Object – Common Data Representation.
Upstream	One-way data flow from the **broadcast client** to the **head-end**.
URI	Uniform Resource Identifier. Also known as a URL.
URL	Uniform Resource Locator. URLs are short **strings** that identify resources on the **WWW**: documents, images, downloadable **files**, services, electronic mailboxes and other **resources** ,etc. They may be thought of as a networked extension of the standard filename concept, in that not only can you point to a file in a directory, but that file and that directory can exist on any machine on the **network**, can be served via any of several different methods, and might not even be something as simple as a file.
User mode	Software processing that occurs at the **application layer**.
UTC	Universal Time Co-ordinated.
U-U	User-User

-V-

VBI	Vertical Blanking Interval. The time period in which a television signal is not visible on the screen because of the vertical retrace (that is, the repositioning of the trace to the top of the screen to start a new scan). **Data services** can be transmitted using a portion of this signal. In a standard **NTSC** signal, perhaps 10 scan lines are potentially available per channel during the VBI. Each scan line represents a data transmission capacity of about 9600 **baud**. In 625-line systems, about 20 scan lines are available in the VBI.
VBR	Variable Bit-Rate. A type of traffic that, when sent over a **network**, is tolerant of delays and changes in the amount of **bandwidth** it is allocated (e.g. data **applications**).
VBV	Video Buffer Verifier. The **MPEG** concept defined in ISO/IEC 13818-2 (MPEG-2, Annex C) employs a fixed-size buffer to handle the transition of the **channel** bit-rate to the rapidly fluctuating coded bit-rate of individual MPEG pictures. The scope of the VBV is only within a sequence. The VBV is built upon a framework of several axioms of decoder behaviour which are unfortunately not very well described in the specification.
VC	Virtual Circuit. A generic term for any logical communications medium.
VCC	Virtual Channel Connection. A logical communications medium identified by a **VCI** and carried within a **VPC**.

VCI	Virtual Channel Identifier. The field in the **ATM cell** header that labels (identifies) a particular virtual **channel**.
VCR	Video Cassette Recorder.
VHF	Very High Frequency.
VHS	Video Home System.
Virtual LAN	A logical association of users sharing a common broadcast domain.
VPC	Virtual Path Connection. A logical communications medium in **ATM**, identified by a Virtual Path Identifier (VPI) and carried within a link. VPCs may be permanent virtual path connections (PVPCs), switched virtual path connections (SVPCs), or smart permanent virtual path connections (SPVPCs). VPCs are uni-directional.

-W-

WAN	Wide Area Network. A communications **network** that connects geographically-separated areas.
Wander	Long-term non-cumulative variations of the significant instants of a **digital signal** from their ideal positions in time.
Wrapper	A function that provides an **interface** to another function.
WWW	World Wide Web / the Web. A hypertext-based, distributed information system created in Switzerland and used for exploring the **Internet**. Users may create, edit or browse hypertext documents on the Web.

-X-

XTP	eXtended Transport Protocol. A **network**-level **interface** appropriate for **file** transfer. XTP can operate in a "raw" mode in which it encompasses both the Network and **Physical** layers, or it can operate on top of **IP**. XTP in raw mode achieves some efficiency and has the possibility of using features of the underlying physical media (such as the **QoS** for **ATM**) that is not possible when XTP is used on top of IP.

-Y-

YUV	True-colour encoding that uses one **luminance** value (Y) and two **chroma** values (UV).

Annex B

Systems

B.1. Object model tutorial

The following paragraphs provides a brief introduction to the terminology of current object model technology. A summary is provided in *Section B.1.4.*

B.1.1. Object systems

In an **object-based system**, software services are provided in terms of individual building blocks that have the ability to contain their own data and to perform all of their logical operations on themselves. The general term for describing this type of software package is a **class** and, within it, are defined all of the operations it can perform and the type of data on which it is allowed to operate. A simple example of this is a VCR class which contains basic information about a tape deck, such as the format of the tape that it accepts (see *Fig. B.1*) as well as operations that will control the basic functions of the tape deck. An **instance** of this class, known as an **object**, contains its own specific values for the data, and the operations defined by the class will operate on this data. In *Fig. B.1*, each instance represents a different physical device, each with its own specific format of tape. For example, if there are two tape decks that accept tapes of different formats (say D1 and Betacam SP), two instances of class VCR can be created: one that represents and controls the D1 tape deck, and the other that controls the Betacam SP deck. This notion of objects containing their own data and the operations that can be performed on that data is known as **encapsulation**. When looking at how object systems can link together to perform complex tasks, this idea becomes critically important.

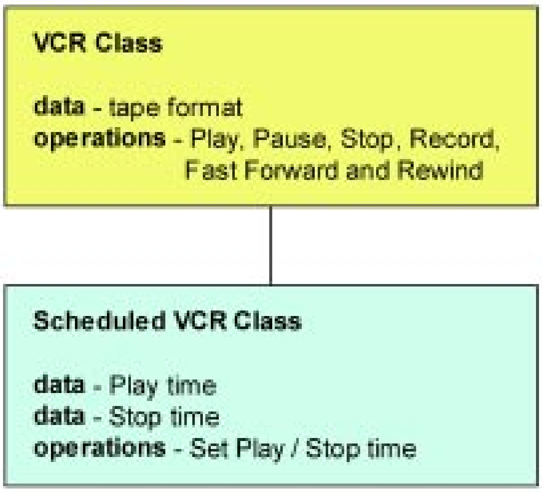


Figure B.1: Class and Sub-class relationships.

Aside from the benefits of data encapsulation, another important characteristic of classes is their ability to **inherit** functionality from other classes. This idea is known as **sub-classing** and can be illustrated very simply by looking at the example VCR class in *Fig. B.1.* As previously described, instances of this class store information about the specific tape device and provide control of the tape deck's basic functions. Therefore, the tape format information and the control operations encapsulate all of the **attributes** of the class. As can be seen in *Fig. B.1*, a second class is created, the Scheduled VCR class, which inherits all of the attributes of the VCR class, but adds an additional bit of functionality for starting and stopping the VCR at specific times. In this case, the sub-class does not replicate the attributes of its **ancestor** class, but inherits them at no extra cost and only needs to provide the new functionality.

This idea may be extended to create sub-classes of the Scheduled VCR Class that add yet more functionality. For each successive sub-class, only the new attributes and operations need to be specified. Everything else is inherited from the ancestor classes. This ability to add and change functionality without replication of effort is what makes the object-oriented methodology so compelling. The functionality of existing classes can be augmented without duplicating it, by simply adding in the desired operations. In systems implemented using this methodology, the consequence is that new functionality can be added to existing systems with very little effort and without risk of breaking the current functionality.

In addition to passive data, objects can contain other objects as part of their attributes. With the ability to both inherit functionality from existing classes, and to use the functionality of other objects, very large and complex systems can be built that are easily maintained, easily enhanced and flexible to change. Furthermore, the use of object systems permit software developers to divide up large systems into “natural” pieces that model the intuitive view of the solution, and provide much more easily-understood interfaces. As an example, the VCR class closely models the real-world view of a tape deck.

Another useful feature of many object systems is a mechanism that allows objects to **observe** the state changes of other objects. An example of this might be a Monitor object that watches the VCR class so that when the tape starts to play, it automatically turns itself on to display the output (see *Fig. B.2*).

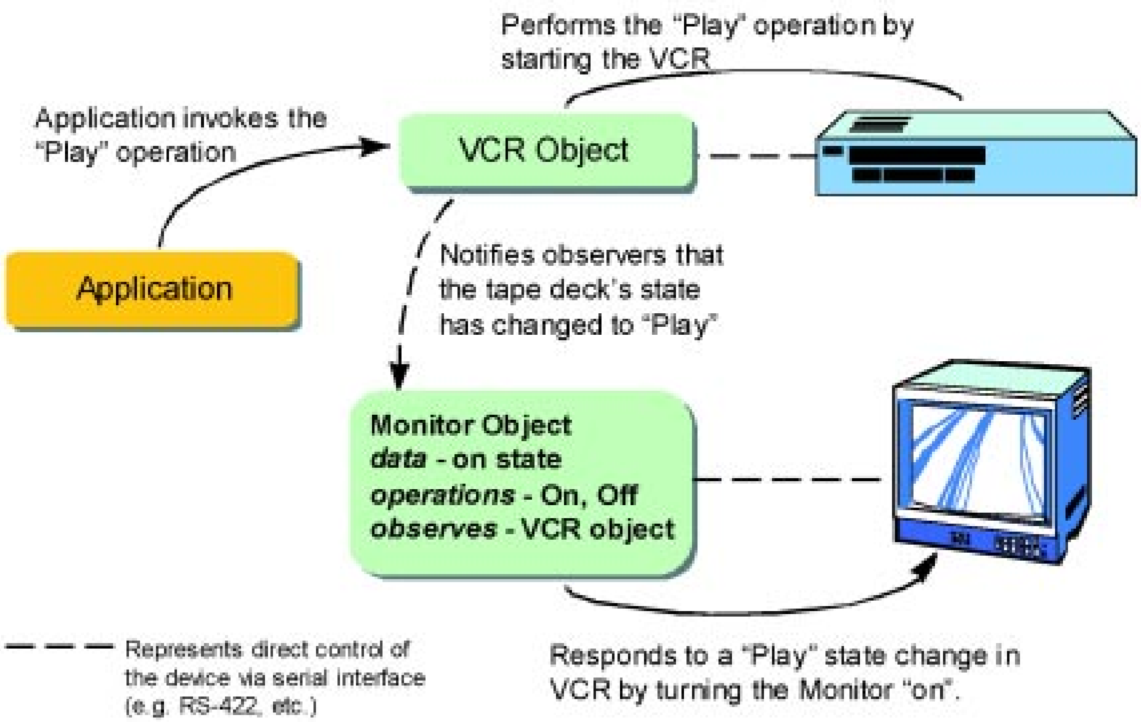


Figure B.2: Example of the observation mechanism.

In this case, the monitor will request notification of change-of-state events in the VCR, which will in turn broadcast these events to its observers when the changes take place. The key here is that the VCR does not need to know anything about the objects that are observing it or even the number of objects interested in its state changes. It simply has to broadcast a notification using a standard mechanism that is part of the object system, and all of the observers will receive the appropriate events. This is a powerful concept in that it allows us to add new objects to the system that can interact with existing ones without their knowledge and without disruption of the current system activities.

By extending these concepts to a **distributed object system,** the picture can be completed. Here, objects reside on devices anywhere on a network and are still accessible to all other objects and software clients. This can be accomplished by creating a **registry** of objects on the network that knows the location and capabilities of all objects in the system. When a client wants to sub-class an existing class, or wants to use an existing object, the registry will provide the information necessary to access their attributes.

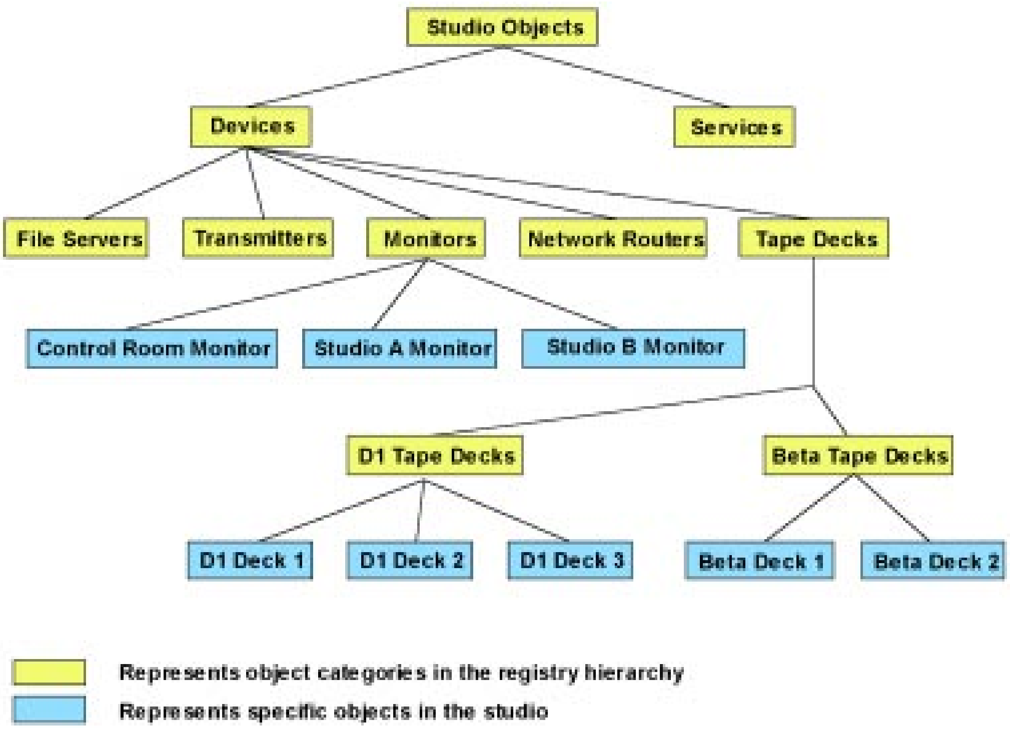


Figure B.3: Network Object registry.

Fig. B.3 shows an example of a simple registry for studio devices and services that organizes objects in the studio into a well-defined hierarchy depending on their capabilities. If the registry allows objects and classes to be registered with information about their nature, i.e. a “Monitor” type, or a “Tape Deck” type, then a system is created where the presence of an object or class on the system can be queried when needed, and the client need not have any prior knowledge of the specific makeup of the network. An example of where this is useful is an application that wants to direct the output of a tape deck to all the monitors in the studio. The client retrieves the object controlling the specific VCR it wants to play, requests the registry for all of the currently-registered objects of type “Monitor” and instructs the tape deck to output to those devices.

Note that with distributed object systems, one or more objects can live in the same device, or a single object can be spread over several devices (e.g. for objects that contain other objects as part of their attributes). This allows an object system to attain high levels of extensibility as well as a high degree of reliability through distributed implementations.

The ideas of network-distributed objects can be applied directly to the broadcast studio if all of the studio devices are viewed as network objects. That is, devices such as tape decks, monitors, video file servers, cameras, network routers, and even transmitters will all be implemented as network objects that employ well-known operations to control their functions. The above example of sending output from a VCR to a studio monitor shows this as a simple case of the general idea.

B.1.2. The networked studio

In a studio designed around a network-distributed object system (hereafter simply called a *networked studio*), all physical devices are attached to the network either by containing the necessary software within themselves, or by attaching themselves to a device “proxy” that is on the network and is able to control their functions. In each of these cases, a software implementation of the object that controls the device is present on the network and registered with appropriate information about its nature. An example of this is a tape deck that is attached to a computer through an RS-422 connection, which in turn is attached to a network and implements all of the functions of the “Tape Deck” object. Such a proxy may control many attached devices. If a client wants to start the tape deck playing, he queries for the specific tape deck object and invokes the Play operation.

Fig. B.4 shows a diagram of the network with a variety of devices connected to it. Notice that objects are viewed the same way on the network whether they represent real physical devices, or are purely software objects such as a playlist editing application.

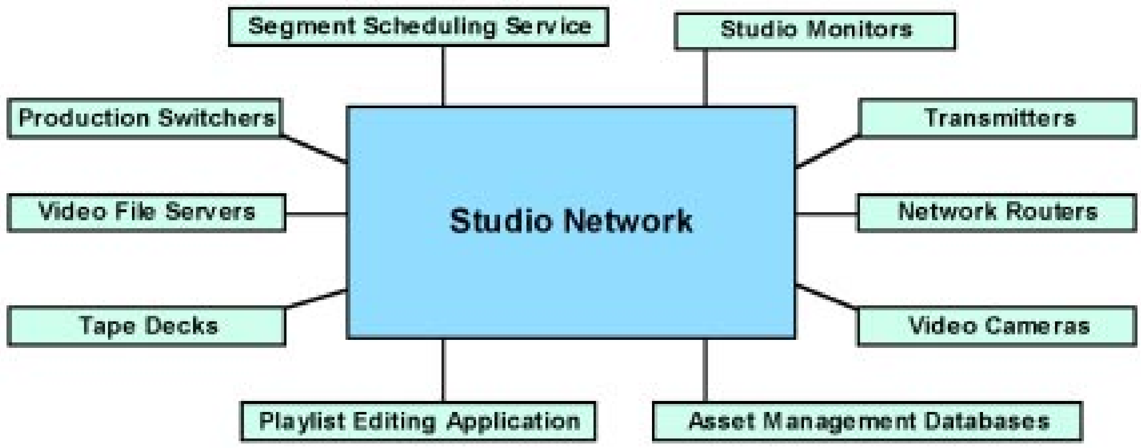


Figure B.4: The networked studio.

B.1.3. Security

The preceding sections demonstrated how the networked studio is readily extensible in its functionality and is easily accessible by the wide variety of network connection options. One of the potential problems with both of these attributes is that of network security. In one scenario, unauthorized people need to be prevented from accessing the network in order to protect the studio's proprietary information. In another scenario, it must be ensured that malicious services that can disrupt studio operations are not allowed to operate within the system. Both of these scenarios have been the focus of much study in the computer community for many years, and methodologies and protocols have been developed that provide for extremely secure access to network resources. These services are currently available in most modern commercial network systems and discussion of their specifics is beyond the scope of this document.

However, in addition to these well-known network security services, it is also possible to implement additional security models that augment this functionality by providing secure access at the object level. In these cases, security services take into account more issues than simply a user name and password. In some systems, objects are stamped to indicate their origin or author, and security models have been developed that require a client to be authenticated as "trusted" before it can invoke the operations of another object [8]. The notion of an object being trusted can take on a variety of meanings, but in general it refers to an object of known authorship or origin that is assured of not being malevolent. Beyond this, an object can also allow for varying levels of access to its operations and data, depending on how well-trusted the calling client is. For example, an object may provide some functionality to all objects, while other operations can only be accessed by well-known clients. As can be seen, a great deal of technology is available to ensure that access to the network and to the network objects occurs only by authorized personnel, and that malicious services will not be allowed to operate within the system that can disrupt studio operations.

B.1.4. Summary of Terminology and Structure

A **class** is a kind of software package, a building block, which defines a unit of data storage and a set of associated **operations**, sometimes called **methods**, which typically operate on this data. Every **object** is an **instance** of one of the classes, and has its own private instance of the data unit. This data can be thought of as the **attribute**s of the object. One class may be defined as a **sub-class** of another class, in which case, it **inherits** the data and operations of its **super-class** or **ancestor** class. **Encapsulation** is the notion that the private data

8. CORBA and Sun Microsystems's Java™ are examples of systems that use this security model.

instance, and the details of the implementation of the object's operations, are generally not directly accessible by other objects; they are only permitted to invoke (call) the object's operations. If objects in a system are located on two or more different computers (each having its own address space), the object system is described as **distributed**. In this case, a **registry** is useful to resolve an object's names into its host computer's name and its local name on that computer. If one object monitors the actions and state changes of another object, the first object is said to **observe** the second.

B.2. The RFT - New systems management services

After the phase of work culminating in the release of its First Report, the Task Force moved on to examine fully-developed technologies that can be applied to the installation of television production systems in the immediate future. As part of this effort, it sought input from experts and organizations in related industries, regarding technologies for the management of systems that may be applicable for use in this environment. The vehicle used for obtaining that input was the following Systems Request for Technology (RFT), which had December 1997 as a deadline for responses.

The transition from current methods of television production to new methods that will be enabled by the incorporation of computer and networking technologies will lead to the potential for significantly increased complexity, both in the systems to be used and in the operation of those systems. This is likely to lead to the need for partially- or fully-automated methods of managing and controlling the new systems, at least in facilities of moderate and larger sizes.

To enable management and control of the various types of equipment and networks that will be designed into future systems, it will be necessary to have standardized interfaces and protocols that can be used for the exchange of information and control instructions between system devices and controllers. Such standardized techniques will permit the design of systems using elements from various suppliers, while relieving the vendors (of both the controllers and the controlled devices) from the burden of having to create specialized interfaces for each type of equipment to which their devices might be connected. This will benefit both the suppliers and the owners of the equipment.

The purpose of the technology sought in this RFT is to enable the design and implementation of the resource managers, schedulers and system controllers envisioned as solutions to the potential problems of increased complexity, without at the same time defining those devices. The goal is to establish the infrastructure of communication networks, protocols on those networks, and control ports on system equipment, as well as to define the types of messages that will flow from controlled and scheduled equipment to the controllers and schedulers.

The ultimate aim of this work is to make it possible for the operators of television production and distribution systems, based on bitstreams, to take advantage of the benefits that derive from the use of techniques such as video compression, while making transparent to the user the additional system complexity that often comes with such methods: the user's job is to manipulate the Content, rather than deal specifically with the technology. Thus, for example, when an operator needs to use certain Content as input to an editing process and that Content has been captured and stored on servers in two formats – one at full resolution, the other with "thumbnail" resolution – then, depending upon the type of editing system the operator uses, the correct form of the Content must be delivered without the operator specifying, or even knowing, which one is required and on which server it is stored. If the editing was done on a thumbnail editor, then when the edited material is required for use in a programme or for distribution to viewers, a conformed version using the full-resolution format must be created automatically by using the Edit Decision List created on the thumbnail editor stored in the system and related to all of the material involved in the editing process. This simple example shows the sort of functionality that this RFT seeks to enable and that will be required on a much more complex level in future real systems.

B.2.1. Description of the problems to be solved

The following sections discuss many of the considerations that will influence the requirements for the management and control of systems. Understanding these aspects of system implementation should help in devising and / or explaining the capabilities and functionality of the technology to address the needs expressed in this RFT. This collection of issues, techniques and practices is meant to illuminate the areas of concern and utilization, while not providing an exhaustive description of a particular application which would inherently limit the range of future potential uses of the technology.

Full understanding of many of these topics requires detailed knowledge of television operations. It is not the purpose of this document to serve as a tutorial on these matters. If information in more depth is required, potential respondents are invited to contact Roger Miles at the EBU address. He will be able to refer the caller to a knowledgeable individual on the subject in question.

B.2.1.1. Push and Pull models

In the development of system management specifications, one must identify the difference between the "Push" and "Pull" models of Content distribution, and understand the industry's movement toward the Pull model. The Push model refers to the process of broadcasting information from a source to passive receivers without an explicit request, while the Pull model refers to the process of a receiver requesting the information from a passive source. An example of the Push model is traditional television broadcasting, while the World Wide Web is an example of a Pull model.

The major difference for the Content provider concerns the very different demands on timeliness, storage and distribution of the two models. The Push model gives a provider total control over the system resources. A successfully implemented system can allow for relatively limited demands on Content storage and access, since distribution is straightforward and not dependent on audience size. Pull models, to be successful, must understand the scope of the demands that the audience will place on the Content delivery mechanism.

It is understood that a broadcast studio will require both Push and Pull styles of Content delivery. The system must be flexible enough to accommodate both models simultaneously, as well as being adaptable enough to allow roles to change as the studio evolves. The demands and structure of these models are significantly different and, on the surface, require different types of Content, facilities and system management. In the desired solution, however, the demands placed on system storage and access facilities for both models will be accommodated by the same system.

B.2.1.2. Equipment

The typical television facility has a variety of equipment, most of it quite specialized. This is due to the way the technology evolved. In recent years, this has been supplemented with a considerable amount of generally standard computer hardware, both for control and for audio / video generation and storage.

Purpose-built television equipment uses signals, interfaces and control systems developed and optimized specifically for the industry. Some of these, like SMPTE Timecode, have proven so useful that they have been adopted almost universally. Others are quite narrow in their application. Control protocols currently are in most cases manufacturer-specific although, in some cases, a particular device has been so widely adopted that other manufacturers have found it makes sense to emulate it.

Equipment replacement cycles in the broadcast industry are much longer than in the data-processing business. It is not uncommon for users in the broadcast industry to expect a 15-year service life, and to justify capital expenditures accordingly. Therefore, any technology offered for use in this industry must be flexible enough to accommodate legacy devices in addition to making use of all the capabilities of new equipment built in recognition of the bitstream approach: developers must be prepared to provide support for older systems while introducing features to take advantage of newer ones.

B.2.1.3. Incompatibilities between data and devices in a system

Due to the variety of sources found in a system, from acquisition to archive, incompatibilities may occur between data and devices.

The response should explain how the proposed management system manages the different formats, how it resolves (if it can) these incompatibilities, and what information will be available when it does. For example, a server with an SDTI or network connection could store either SX or DV files. How does an edit suite or an On-air suite, dedicated to one format, use the other? How does the system provide circuits with necessary conversion? In the same way, how does the operator, or the device itself, know what restrictions there are in terms of delay or latency due to the circuits and treatments?

B.2.1.4. General device control

The responses should describe the general device control architecture, including the kinds of communications supported (i.e. whether via connection, connection-less, point-to-point, network, bi-directional or piggy-backed, etc.). The functional principles of the system should be outlined in detail, including the approach taken for integration of legacy equipment and the compatibility of the proposed control system with current standard and industry practices. The expandability of the device control system should be described, along with the areas of application it covers, e.g. whether local or distant.

B.2.1.5. Data timing

Within a TV facility, the timing and synchronizing of related data, such as video and audio, is a critical part of system design. In future TV environments, where separate compression systems will cause differing latency of both the video and audio, the task of maintaining “lip sync” will become an essential part of the system design. To add to this complexity, there may well be associated data and Metadata which will have to be synchronized to the V / A stream. In addition, many complex timing relationships between a multiplicity of video, audio, and data streams will have to be maintained.

Metadata within the production centre and that arriving via contribution and distribution connections will need to be managed and tagged for future use. Ideally, the stored Metadata and the playout list will be linked to provide automatic retrieval and playback synchronization. Furthermore, as Content is moved within a production centre, the links will have to be maintained, even though storage of the Essence and its associated Metadata may be on separate devices.

B.2.1.6. Time line management and real-time control

Within the production facility, there may exist many types of control systems. One of the most difficult control parameters is Hard Real Time. In most simplistic terms, Hard Real Time means that, with the push of a button, a function is executed with minimal predicted delay, (e.g. insertion of a commercial during a live event where there is no prior knowledge of the insertion time).

Television systems that employ compression suffer from varying degrees of latency that may change on a dynamic basis. End-to-end timing of a typical network feed for a sporting event could be as much as 10-15 seconds, due to various points along the chain having different latencies. Between a network origination point and the local station or head-end, there will exist a significant time offset, so the overall system management has to be able to deal with or anticipate this time offset.

In some cases, a local event is triggered from a remote source. In such a case, the latency of the control system and the pre-roll time, if any, should all be accounted for.

Television, as it exists today, is a frame-accurate system, and this accuracy cannot be relinquished. The system manager must ensure, in the case of compressed signals, that any preloading (pre roll) of buffers be taken into account to guarantee frame-accurate switching. Topics which must be considered are:

- remote / local control;
- deferred commands;
- Hard Real Time commands;
- device interfaces / pre-roll;
- network / local timing.

B.2.1.7. Network services and Quality of Service (QoS)

The TV production facility of the future is likely to be interconnected using a mix of technologies ranging from unidirectional dedicated circuits to bi-directional shared networks. When shared networks are used, various degrees of interconnection reliability, bandwidth and latency – termed Quality of Service (QoS) – are possible and must be managed (see the first Task Force report for details of QoS considerations).

A network service is a level of performance and / or function that is provided by the network to applications and users. When the level of performance is unpredictable and unreliable, the service is termed “best-effort”. Predictable, reliable and guaranteed services can also be made available on the network. Services can be applied to support specific high-priority applications and users, or can be applied at protocol or organizational levels for bandwidth management to make the network more efficient.

Service guarantees are based on defining the network characteristics (QoS characteristics) that can be configured in the network elements. For network services to be effective in supporting high-priority applications and users, they must follow three rules:

- **Rule 1**: Services are applied end-to-end, between source and destination, at all network elements in the path of the application flow. This includes the systems’ device drivers, operating systems and application interfaces.
- **Rule 2**: Services are configurable using QoS characteristics (described below) at each network element in the path of the application flow.
- **Rule 3**: Services are verifiable within the applicable network.

These rules are necessary conditions for services to be meaningful within the network, and to their high-priority applications. For example, if a network service cannot be applied to all network elements in the path of the application flow, then the service cannot meet its guarantees to the application / user.

B.2.1.7.1. Quality of Service

Each service can be described by its QoS characteristics. For network performance, QoS characteristics are measured in terms of bandwidth, delay and reliability. Examples of QoS characteristics for performance include:

- **Bandwidth**: Peak Data-Rate (PDR), Sustained Data-Rate (SDR), Minimum Data-Rate (MDR).
- **Delay**: End-to-End or Round-Trip Delay, Delay Variation (Jitter) and Latency (delay to first receipt of requested data).
- **Reliability**: Availability (as % Up-time), Mean Time Between Failures / Mean Time To Repair (MTBF / MTTR), Errors and Packet Loss.

A response to these and other QoS characteristics will be needed to define services in the network, as well as to develop metrics that will be used to verify services within the network.

B.2.1.8. Multichannel operation management

Multichannel management ranges from a 4- to 6-channel system, up to a 500-channel system. It may be that control systems exist that are, in fact, extensible. On the other hand, systems that are optimized for small systems may only be suitable for those systems. Listed below are some of the functions that are necessary in a multichannel playout facility and where contributions are welcome:

- Contribution / Distribution Multiplex Control;
- Emission Multiplexer Control;
- Automatic Channel Monitoring (of Stream Attributes and Errors);
- Multiple Play List Handling;
- Network Conflict Resolution (Network in this context is a shared data network);
- Network Configuration / Router Set-up (Network – see above);
- EPG Creation / Update;
- Interface to Billing Services;
- Resource Assignment;
- Multiple Audio Channel control.

B.2.1.9. Multicasting

Video data within the studio must often be distributed to multiple destinations for simultaneous viewing, recording, redistribution or archiving. To accomplish this, a system must be able to direct a single video / audio stream to multiple destinations without a reduction in the expected stream quality. For some systems, this requires the delivery of the stream to all recipients with identical stream characteristics as the source, while in others the QoS may vary in response to the recipient's requirements. In both cases, however, stream delivery is initiated simultaneously to all recipients from a single source. The stream should proceed reliably and with very low error rates to all destinations within the limits of the system resources. Addition or removal of stream recipients can occur at any point in the stream, and the system must handle these situations transparently for all active recipients without reduction in the stream quality or reliability to all other destinations. Control of the multicast stream is commonly restricted to the source of the stream, but systems may provide levels of control to individual recipients if desired. In all cases, changes to the stream source by any client will propagate to all stream recipients simultaneously.

B.2.1.10. Connections, streams and transfers

The movement of data through a broadcast studio involves three distinct functions: connection of devices within the studio, transferring data from one device to another, and the streaming of media data between devices. Generally speaking, the requirements for each of these elements are necessarily related. However, for the purposes of defining the specific operations that occur when moving data about the studio, it is useful to define them separately and to identify the specific requirements of each.

B.2.1.10.1. Connections

A connection is defined as the path between two devices in the studio for the purpose of controlling, transferring, or streaming data from one device to the other. Connections can be accomplished either through the use of serial cabling for high-speed dedicated point-to-point connections, or by using an addressable digital network channel for point-to-point or point-to-multipoint connections. In most studios, both types of connections are supported and a mix of operations requiring combinations of each is not uncommon.

In the case of a digital network, connections refer to the specific path that a signal must take, including all router and switch links, to provide a contiguous data path from the source to the destination devices. For serial links, connections refer to the source and destination device ends, the switches in between, and the physical cabling that makes up the entire path.

B.2.1.10.2. Streams

A stream is the controllable, continuous flow of data (video, audio, etc.) between devices in either a synchronous or an asynchronous manner. For critical studio operations (playout, output monitoring, video editing, etc.), streams are assumed to be continuous, with extremely low error rates. For less-critical operations (video browsing, etc.), the dataflow must still be continuous, but may use lower bandwidth and exhibit higher error rates depending on the quality of the stream required.

The reliability of a stream is dependent on the connection it flows across. If the connection is interrupted or broken, the stream is stopped and cannot proceed unless a new connection is established, or the interrupted connection is resumed.

B.2.1.10.3. Transfers

Data transfer is the controllable flow of data, either synchronously or asynchronously, between devices. As opposed to a stream, the flow of data in a transfer operation is not always required to be timely or continuous, but is assumed to be reliable and error-free. Operations such as the movement of video data between servers, or to off-line storage and communications operations are examples of data transfer.

As was the case for streams, the reliability of a data transfer is dependent on the connection it flows across. If the connection is interrupted or broken, the transfer is stopped and cannot proceed unless a new connection is established, or the interrupted connection is resumed.

B.2.1.11. System resource management

Since connections require resource allocation of both source and destination devices and ports, as well as the network devices and physical switches prescribed by the connection path, the management of resource scheduling must be carefully done to ensure that over-subscription of resources does not impact on the execution of studio operations. Specifically, this requires that the system provides a device and network resource management facility for computing and allocating an optimal network path between devices. This system should allow the clients to reserve system resources in advance of their actual use, it should prescribe priorities for resource usage between clients, and should provide real-time interruption and resource reallocation to accommodate critical studio tasks. In addition, the system must provide facilities for monitoring the system resource usage and should provide adequate feedback to interested clients. It should also deploy a security model that is capable of preventing access to studio resources by unauthorized users.

Since clients will use system resources sparsely, an important task of the resource manager is to provide scheduling facilities that will allow clients to request resources and priorities for a given set of connections ahead of time. The nature of this request is determined by the type of connections being reserved. For network connections, information about the required stream bandwidth, QoS parameters, stream or data transfer type, data format, and the start time and duration of the connection are required. For serial connections, only the start time and the duration of the connection are necessary.

For network connections, the specification of routers and switches that make up a connection are generally not of interest to the client requesting the connection. As such, the system resource manager has a great deal of leverage in defining the exact path to use between devices; it can choose a path that either optimizes the overall network usage, or optimizes the timeflow of data between the devices. It is the responsibility of the system resource manager to provide heuristics for computing the exact path information, based on the interests of both the client and the overall system.

Within the studio, there are always a large number of connections in use simultaneously. While the system resource manager should optimize the connections so as to accommodate the largest number of clients possible, it must be able to handle situations where critical network or device resources saturate, and resource reallocation is required to accommodate the critical studio tasks. To facilitate this, the resource scheduling service must define priorities for all resource reservations and should provide a mechanism for interrupting lower-priority connections when a critical need arises. For example, live programme Content will generally have the highest priority, scheduled programme Content the next highest, etc., with non-real-time tasks such as video library browsing or editing having relatively low priorities. Where a given system resource is oversubscribed, the system resource manager will make a best effort to accommodate all tasks if possible. If not, it is responsible for breaking the connection of a lower-priority client to allow higher priority tasks to proceed. In all cases, the system resource manager must provide proper notification to clients, informing them of a critical status change.

As a studio expands to accommodate more devices, the ability of the system resource manager to accommodate a larger number of connection requests becomes stretched. To ensure that the infrastructure of the studio is capable of providing increased demand on its resources, the system resource manager must provide qualitative feedback about the system resource usage to interested clients. When critical resources become overbooked frequently, the resource manager must be capable of conveying this information to interested clients, with enough information that studio management will be able to make efficient decisions about deployment of new resources to avoid problems. In addition, the system resource manager should include a mechanism for providing qualitative information about the overall system-resource impact of requested studio operations prior to their execution. This is necessary to facilitate decision-making processes and to allow for the consideration of alternative actions based on resource constraints.

Since digital networks provide a wider range of connection points than do serial connections, the system resource manager must employ a security model that prevents unauthorized users from gaining access to the system resources. Generally speaking, this involves security at two levels:

- at the network access point to ensure that users have authorized access to the network in general;
- at the resource itself, to ensure that resource APIs cannot be invoked without proper authorization.

Both levels of security are necessary to ensure that access to the studio is properly protected.

B.2.1.12. Device control and resource usage

For most studios, a great variety of devices and services must exist to handle daily operations. It is generally useful for the studio to provide a method of controlling these devices and of managing the system services in a standard and well documented manner, so as to accommodate the widest number of studio device vendors and to provide an open platform for new studio service development. While this is not an absolute requirement for all studios, maximal compatibility between vendor devices is strongly encouraged, as is the highest quality and greatest reliability possible for the overall studio. In the case of serial device connection and control, this generally involves a standard command protocol while, for networked devices and services, this involves a standard network transport protocol for communication and a set of software APIs for device and service control.

In addition, it is necessary for the studio to provide a system of feedback to interested studio clients about overall system resource usage, specific resource audit information, and real-time information about device and service status changes. The requirements of each of these types of information will vary depending on the clients, but this information is generally necessary to inform studio managers about the overall state of studio resources, as well as for service clients that rely on up-to-date device status information to perform their tasks. These mechanisms should be functional for both serial and network devices and should involve well-defined protocols and / or software programming interfaces to communicate with clients.

B.2.1.13. Wrapper management

The various data packets to be managed within the system will in the long-term be associated with and attached to a Wrapper. The Wrapper will contain information about the programme (see below) which the system manager will need to know about. The Task Force is interested in technology or systems that can track and manage this data in the areas of:

- ⇨ Access Control;
- ⇨ Identifiers & Labels;
- ⇨ Version Control;
- ⇨ IPR Management;
- ⇨ Data Access;
- ⇨ Essence Tracking;
- ⇨ Contribution / Distribution Information;
- ⇨ Data Base Management;
- ⇨ Play-list-Essence Matching.

B.2.1.14. Content / Essence management

Essence and associated Metadata (Content) needs to be tracked, catalogued and accessed throughout its life from creation / acquisition to post-production, and on to consumption and archiving. Individual Essence elements (images, audio, data) need to be identified, located from a variety of locations in a distributed environment, and merged to create the final programme. Furthermore, the final programme in its various versions needs to be tracked and distributed to the proper destinations at the appropriate time for emission / viewing and ultimately for archiving. Locating the Essence by using intelligent agents to “query” the database servers (containing the Metadata of the Essence) is a method of acquiring Essence. Access rights, versioning, copyright and billing are a few of the processes that can be “served” through databases and agents.

B.2.1.15. Multiplexer control

Multiplexer control may take on two distinct requirements, one dealing with contribution / distribution, the other dealing with emission. While there may be similarities in most areas, there will be different demands placed upon the multiplexer.

In the case of contribution / distribution, the bandwidth between the sending site and the receiving site(s) can be considered a very flexible, configurable pipeline. In particular, the types of Metadata connected with the video and audio signals may consist of public and private data. The pipeline must be reconfigurable in hard real-time. The pipeline is likely to have bandwidths up to 150 Mbit/s, although the more common bandwidth will be in the 40 -60 Mbit/s range. The types of control will vary depending on the multiplexer manufacturer. However, control over the compressor bit-rate, the latency of the compression engine, the allocation of bandwidth for Metadata and the insertion of system information are examples of some of the many commands that will be necessary.

In the case of the emission multiplexer, there are certain mandated constraints imposed by the ATSC or DVB standards. For example, even though in both cases the packetizing of the data must comply with the MPEG Transport Stream definitions, differences exist between the ATSC and DVB standards in some instances; e.g. in the handling of private data packets, Closed Captioning and the like. As was the case for the contribution / distribution multiplexer, there will be the need for both deferred commands and hard real-time functions. In addition to the Video / Audio information control, the emission format requires the insertion of complex system information, such as a unique station ID, virtual channel table, electronic programme guide, etc. It is anticipated that somewhere in the overall system design, the input sources to the multiplexer – video, audio and data of all types – will need to be time-linked in some way. Although not yet mandated by any of the authorities, it is likely that a record of the daily events that take place will be required, or at least desired.

B.2.1.16. Systems Information management

Systems Information (SI) management is a new function that the broadcast industry will have to control. The transmitted bitstream contains some information necessary for the reception of a digital channel. Depending upon the local authorities, the required data may vary from region to region.

The SI has segments which range from mandated to optional, so the system manager has to ensure that mandated information is available and is transmitted. Optional information must also be tracked and inserted as required. The SI must be linked in some way to the playout schedule.

Listed below are some of the required interfaces and information that it will be required to be inserted.

- Interface to Contribution / Distribution Multiplexer;
- Emission Multiplexer Control;
- Insertion of MPEG, or other System Information;
- Remapping of Contribution / Distribution SI into the Emission Format;
- Insertion of Navigational Package (Virtual Mapping Table);
- Ratings Information (Note: In some systems the ratings table is required to be transmitted, even though it may contain nothing).

B.2.1.17. Data services

With the introduction of digital television transmission (DTV), there is the opportunity where regulations permit for a separate data transmission service to be established. In addition, there may be programme-associated data transmitted along with the video and audio information.

B.2.1.17.1. Contribution / distribution data Content

In some instances where the programme emission site is separated from the point of origination (network headquarters to local site), it may be that the data contained in the contribution / distribution bitstream differs from the data finally to be transmitted. In other instances, programme- / Content-related data may be sent from the network HQ in advance of the associated programme.

B.2.1.17.2. Emission Data Content

The Emission Data Content may range from programme-related data at a relatively low bit-rate to a situation where the entire bandwidth of the emission bitstream is considered to be data. Rules which relate to the embedding of data are defined by the DVB or ATSC standards.

B.2.1.17.3. Management of data service sales

In some emission standards, there exists the possibility to "sell" data space on the transmitted bitstream. A means should exist in the management system to manage the multiplexer output and to create billing output to the facility's financial services function. There exists a need to create a data transmission "play-list." As this function is new to the broadcast industry, it is unlikely that any legacy system considerations need to be taken into account.

B.2.1.17.4. Interfaces to transactional services

It is anticipated that, over a period of time, broadcast data services will be linked to external transactional services. Consideration on how system interfaces should be implemented, including possible Application Programming Interfaces (APIs), are of interest to the Task Force.

B.2.1.17.5. Formation of data carousels

Formation of data into carousels may be considered part of the data management process. Complexities and conflicts may occur when centrally-generated data formats need to be integrated into locally-generated carousels. Clearly these conflicts have to be managed or resolved. It is not clear where and how these issues are to be addressed.

B.2.1.17.6. Programme-related data insertion

Within the compression engine or within the multiplexer, certain programme-related data will be inserted into the transport stream. This programme-related data may be in the form of mandated Closed Captioning, or in the transmission of programme IDs. There is also the option to transmit time-sensitive programme-associated data, such as statistics about sporting events, recipes during cooking programmes, or educational information during schools programmes. Co-ordination and insertion of this data is likely to be under the general control of the resource manger, the automation system or the data management system. The combination of this programme-related data and the video, audio and data streams cannot exceed the bandwidth of the emission system.

B.2.1.17.7. Conditional Access activation

Conditional Access falls into two categories: that dealing with contribution / distribution systems and that dealing with emission.

Contribution / distribution CA and its characteristics will mostly be defined by the service provider. There must be provision for control of the individual packets that are being transmitted, with control in hard real-time made possible.

Emission CA must have similar characteristics to that of contribution / distribution, with the CA provided individually for the video, audio and the various data packets, commonly called Private Data Packets. CA standards and interfaces are currently being discussed with the industry standards bodies. Although interfaces may be necessary to the CA system, the Task Force at this time is only looking for general ideas of how the system control and management system would deal with the requirement to implement a CA module. In some cases there may also be a requirement for a "return" channel from the receiving device which could be a TV set, a set-top box or a computer.

B.2.1.17.8. Interfaces to third-party providers

The resource manager or systems manager will have a requirement to interface to third party vendors such as automation companies, billing services, audience research / ratings providers, etc. Interfaces may also exist for Electronic Programme Guide vendors, data carousel suppliers, and the like. Information to integrate these services into the transmitted bitstream and to allocate bandwidth, billing information, etc., goes way beyond what is done today, including dynamic two-way interfaces.

B.2.1.17.9. Private data services (network to local station or head-end)

Within the contribution and distribution system there is a need to include in the bitstream both public and private data. In this context, private data can be characterized as that information which is transmitted by the service provider to the affiliate, but which is not part of the emitted signal. The system management system must be capable of coding this information in such a way that it can be detected and routed to the final destination without human intervention.

B.2.1.18. Fault tolerance

Various components of a broadcast studio system may have different requirements for error and fault tolerance, depending upon the application and budget. An On-Air server, for example, is a mission-critical function. A failure to complete a broadcast of a commercial would result in loss of revenue and inventory, precluding the opportunity to sell the particular time period in which the failure occurred. A system must provide a means to ensure the success of such functions. The system must also support the operation of less-critical functions, and provide error and fault recovery appropriate to each task. For example, a non-linear editor may more readily accept data transfer errors than allow an interruption to dataflow, while an archive operation will require data integrity, with less concern for sustaining data throughput.

Storage networks (such as RAID) can provide a means of recovering from storage media faults, through the use of redundant data. Redundant, independent, paths should be provided by storage and communications networks to allow access to data in the event of network faults. Redundant servers should have access to storage and communication networks, to provide back-up of critical system functions. For example, a media or file server function may be transferred to a back-up server, provided the back-up has access to the media required.

A management server may also represent a critical function, as other servers and clients will require access to the management tasks to continue operation. The system should allow for the continued operation of critical tasks following a fault in the management system. Manual or automated means could be used to bypass the management server for critical tasks until the management server operation is restored.

The control system should provide for a timely and effective health check for inter-dependent systems, to allow rapid recovery from a server fault. The recovery may be to switch over to a secondary method of operation, bypassing the failed server, or by switching over to a back-up server.

B.2.2. Major system-level applications

There are three major enterprises within a “typical” television operation: Production (including Post-Production), Operations and News. While these are conceptually distinct and have different objectives, they share many requirements and are not entirely separate in their execution. Production includes the creation of Content (programmes, spots, logos, promos and commercials) for use either within the creating facility or for distribution beyond the facility. Operations (or “Operations and Engineering”) encompasses all activities required to broadcast the facility’s daily transmissions. News is the creation, production and transmission of news programming. All networks engage in all three enterprises; all local stations and many local cable operations run Operations and do at least some Production; many local stations create a significant amount of Content (local programming) and also do local News. Thus, there is a need for scalability and modularity in the solutions offered to service Production, Operations and News applications. Since Sport, at both the network and the local levels, combines all three endeavours, it is not specifically detailed here but depends on all three types of activities.

B.2.2.1. Operations & Engineering

The goal of Operations is to get the facility’s programming, live or pre-recorded, to air on schedule and without dead air-time. The sources of the Content to be transmitted are varied (tapes, disk-based servers, live cameras) and must be scheduled and carefully managed to ensure Content availability with frame accuracy. Today, the major functional components of Operations are Media Library Services (including Archives), external Input and Output, Master Control, Control Automation, Graphics & Titling, Studio Control and Studio Operations. In the future, with the advent of all-digital operations, these functions may not be categorized in this way and, indeed, some may no longer exist, while new functions may be introduced into the Operations process.

The function of Media Library Services is to be the repository of recorded Content and to make Content available when needed, including the playout of Content on command (playback). As the industry evolves, Media Library Services will support a mixed-media combination of tapes (usually in cassettes) and disk-based servers. Increasingly, Media Library Services will be called upon to support Content re-use and low-resolution browsing as well as high-resolution playback for Production and News activities. It may also provide media management services for cataloguing, searching and retrieving media Metadata. Thus, Media Library Services is an important functional component of all three applications that are the focus of this RFT although, from a systems management perspective, the demands of Operations will take precedence over conflicting playout requests from the other two application areas. Interoperability standards that will impact on the Media Library Services function include the requirement to link sources and destinations together, the need to manage infrastructure capacity, and the need to deal with Content in multiple formats.

External Input and Output (a.k.a. Feeds) is the function that records live feeds, network programmes, time-shifted material, commercials and other satellite-delivered material, and also handles incoming and outgoing taped materials, and transmits certain outgoing material to satellite. Its primary interfaces are to the Media Library Services (moving Content in and out and providing Metadata about that Content), to Master Control for direct provision of Content, and to Control Automation for determining what incoming material to capture.

Control of what is actually transmitted resides in the Master Control function, which switches between a variety of available inputs (usually according to a schedule) and also co-ordinates a Control Automation function which manages slave-mode devices such as graphics and titling equipment. These functions will extend to encompass the management of computer-based Content sources (such as video servers), arrayed on protocol-based networks. The primary interfaces to Master Control are Streaming Content sources (tapes, servers, live transmissions) and the promulgation of commands to control these sources.

Facilities that create Content have functions for Studio Control (which makes Content decisions for locally-produced material, i.e. what to shoot) and Studio Operations (the actual management and operation of the sound stages, cameras, microphones, cueing devices, etc). The studio management function (including mobile operations) is the source of live Content that needs to be managed within the system, along with tape and computer-based Content.

B.2.2.2. Production

The goal of Production is to create Content of various types, both for local use within the facility and for distribution to external locations. This process encompasses all the steps generally performed in the creation of media Content intended for distribution. These include Production Planning, Scripting, Shooting and Editing. Planning and Scripting will become standard workstation-based activities that will benefit from workflow and data management capabilities, as well as data-sharing services. Shooting and Editing will generate transactions (both real-time and non-real-time) against the Media Library Services function. These will include browsing, depositing and retrieving low-resolution and high-resolution media as well as composition information in various states of readiness. It is probable that multiple servers will exist within a facility, and that different servers will provide low-resolution and high-resolution media. It is also possible that some work locations (or even work groups) will generate great demand on the Media Library Services for high-resolution media (editing), while others will more often be dealing in low-resolution requests (browsers). These behaviour patterns within a sufficiently large production environment can be used to design and efficiently manage the networks and infrastructure within the facility to maximize the availability at appropriate cost levels.

There are additional activities in Programming and Promotion that provide programme schedules and plans for promotional material. These activities provide to the Operations function information and material essential to maintain the style and rhythm of the facility’s programming. They do not complicate the system considerations. For example, the information can be incorporated into the lists used to manage the master control operation along with other schedules of network operations and commercials to be inserted. It may also tie into the inventory of time availability in the output of a particular channel, which is composed from the information supplied about the programming and used in the selling of commercials.

B.2.2.3. News

The goal of News is to create the news programming and to get it distributed on schedule. Although News is a mixture of Production, Post-Production and Operations, the need to deliver stories in a timely fashion creates additional pressures on the process in comparison to the production of other Content.

The News Room Computer System (NRCS) is a computer-based program which is used to manage the production and distribution of a News broadcast. The system consists of a database that tracks the needed information (stories, scripts, assignments, rundowns) and application interfaces to the data to support various client applications that create, use and modify the information. Clients are usually desktop systems connected to the NRCS via standard networks (e.g. 10-baseT).

Typical clients include Journalist Workstations, Technical Director Workstations and News Editing Workstations. Journalists, of course, write the scripts for stories, identifying approximate matches between available Video clips and script Content. In the future, the browsing of low-resolution media and the making of edit decisions will become standard activities for journalists at their workstations. In some cases, news editors will perform the final editing of high-resolution media for stories, following the specifications of the journalists, and will post the Content for airing. In other cases, the system will automatically conform the high-resolution media to the journalists" edit decisions. Both the journalists and the news editors put heavy demands on the Media Library Services and will both need interfaces for querying, browsing, acquiring, and creating media at differing resolutions, as well as for performing similar functions on media Metadata. The Technical Director controls the overall operation of a News broadcast, including the functions of Master Control during airtime. There is also a Machine Control function within a News operation that controls the output of various devices used during the broadcast (character generators, vision mixers, audio mixers, effects devices, cueing systems). This function is similar to that used in Master Control, but is often separate from it and controlled by the Technical Director.

The activities that comprise these three applications, Production, Operations and News, are sufficiently distinct in requirements and separable in execution that it can be acceptable to support different technological solutions within each environment, as long as there are robust and efficient interfaces between the three to achieve interoperability. Of course, the ideal would be one set of technologies (compression, interconnect, file formats, etc.) to support all three areas, but this is probably not possible in the timeframe of this RFT. We encourage the proposal of technologies that would provide system interoperability and management capability even within just one area of activity.

B.2.3. Summary

The EBU / SMPTE Task Force is seeking technology proposals to address the requirements of the management and control of systems These will be built, using as their foundation the exchange of programme material represented as bitstreams. The primary functions to be managed have been briefly described here. The technologies sought are those which must be standardized to enable implementation of the required management and control mechanisms. Those mechanisms themselves will not be standardized as they are likely to vary from system to system ,and from application to application. Rather, the services necessary to support the management and control mechanisms are the focus of this Request for Technology.

B.2.3.1. RFT checklist – System management services

The table on the next page should be completed with marks made to indicate the areas of compliance. This table will be used to categorize the submissions and to check broad compliance with the objectives of the RFT.

Detailed information can be added by using separate pages. Reference to this additional information can be made by indicating “Note 1”, “Note 2” etc. in the table.

The table headings are defined as follows:

- **Offered** indicates whether the response covers this aspect of the requirement.
- **Defined** indicates that the response covers a defined way of dealing with this aspect of the requirement.
- **Universal** indicates that the response covers this aspect for all application areas in its current form.
- **Extensible** indicates that the offering can be extended to cover this aspect of the requirement.
- **Timeframe** indicates the timeframe in which the response will meet this requirement.

RFT checklist: System Management Services.

Ref.	Topic	Offered	Defined	Universal	Extensible	Timeframe
B2.1	Push and Pull Models					
B2.4	General Device Control					
B2.5	Data Timing					
B2.6	Timeline Management (RT Control)					
B2.7	Network Services and QoS					
B2.8	Multichannel Operations Management					
B2.9	Multicasting					
B2.10	Connections, Streams and Transfers					
B2.11	System Resource Management					
B2.12	Device Control and Resource Usage					
B2.13	Wrapper Management					
B2.14	Content / Essence Management					
B2.15	Multiplexer Control					
B2.16	SI Management					
B2.17.1	Data Services: Contr. / Distr. Data Content					
B2.17.2	Data Services: Emission Data Content					
B2.17.3	Data Services: Data Service Sales Management					
B2.17.4	Data Services: Transactional Services Interface					
B2.17.5	Data Services: Data Carousels Formation					
B2.17.6	Data Services: Programme Related Data Insertion					
B2.17.7	Data Services: CA Activation					
B2.17.8	Data Services: Third Party Providers Interfaces					
B2.17.9	Data Services: Private Data Services					
B2.18	Fault Tolerance					
B3.1	Operations and Engineering Applications					
B3.2	(Post) Production Applications					
B3.3	News Applications					

Annex C

Networked Television Production – Compression issues

European Broadcasting Union
Union Européenne de Radio-Télévision

An EBU Status Report

C.1. Introduction

The integration of new digital video data formats based on compression into existing digital production environments is already occurring at a rapid pace, creating a remarkable impact on storage media cost and post-production functionality. The widely-used Digital Betacam recording format [9] is an obvious example for the successful use of compression in digital television production and post-production operations. Compression based on M-JPEG as the key enabling factor for opening Hard Disk technology for broadcast non-linear editing (NLE) applications is yet another. Compression further allows for cost-efficient bandwidth utilization of contribution / distribution links. The routing of programme data in its compressed form through local-area as well as national and international Telco circuits is therefore expected to become the predominant form of distributed programme production in the future.

Although compression can be applied to all data elements relevant to programme production – Video, Audio and Metadata – this report focuses exclusively on the implications of applying compression to the video signal. It is current thinking that digital audio in production and post-production should remain uncompressed although it cannot be totally excluded that external contributions may require the handling of audio in compressed form. In this case, the considerations described in this report will also apply. It is further understood that compression applied to Metadata would have to be totally lossless and reversible.

Interfacing between equipment that uses identical or different compression formats is currently effected through the Serial Digital Interface (SDI) format in base-band exclusively. On this condition, the existence of different and incompatible compression formats within manufacturers' implementations reflects on the achievable picture quality and the storage efficiency exclusively.

This situation is expected to slowly evolve into a state where programme data composed of compressed video, audio and related Metadata will be processed and routed in its native form directly, employing methods and protocols borrowed from the IT community and adapted to meet the QoS requirements of professional television production.

9. Sony advises that Digital Betacam will continue to support digital interfacing at the SDI baseband level and does not recommend interfacing in the native compressed form.

The salient benefits of that approach are: (i) improved operating efficiency by means of multi-user access to identical programme segments and (ii) reduced data transfer times for dubbing and transfer to and from different storage and processing platforms. Although the recording formats used in production and for programme exchange will continue to be subject to constant change, due to the ever-decreasing cycles of storage media development, the significance of guaranteeing future-proof replay of digital compressed television signals from a particular recording support will gradually be replaced by the need for standardized protocols for data transfer across different and changing recording platforms. The compression scheme chosen for that purpose will then no longer be a kernel feature of a particular implementation, but will bear the potential of becoming the core element of a total television production chain, including a hierarchy of tape- and disk-based storage devices offered by different alliances of manufacturers. The integration of compression and network technology into broadcast operations is therefore expected to increase both the operating flexibility and the universal access to television archives.

The majority of broadcast production and post-production operations cannot be performed today by direct manipulation of the compressed data stream, even within a single compression scheme [10]. The consequent cascading of decoding and re-encoding processes within the production chain and the quality losses incurred therefore require the adoption of compression schemes and bit-rates which support the quality requirements of the ultimate output product.

In the framework of the joint EBU / SMPTE Task Force, members of the EBU have entertained in-depth discussions with major manufacturers involved in the development of technology for future networked television production, with a close focus on the compression schemes available today and in the foreseeable future, and on the balances obtained in terms of:

- ultimate technical programme quality versus data-rate;
- interoperability of compression schemes using different encoding parameters;
- editing granularity versus complexity of networked editing control.

In the course of these proceedings, the EBU has acknowledged different quality levels [11] within the confines of professional television production and post-production. There is agreement that further adaptations may be required to overcome bottlenecks created by constraints, e.g. bandwidth, tariffs and media cost. The appropriate selection of a single compression scheme – or a limited number of compression schemes within one compression family, together with the publicly-available specifications of the relevant transport streams and interfaces – will be of overriding importance if efficient exploitation of the potential offered by networked operating environments is to be achieved in the future.

C.2. Compression families for networked television production

For core applications in production and post-production for Standard Definition Television, two different compression families on the market are currently advocated as preferred candidates for future networked television production:

- DV / DV-based 25 Mbit/s with a sampling structure of 4:1:1, and DV-based 50 Mbit/s with a sampling structure of 4:2:2, using fixed bit-rates and intra-frame coding techniques exclusively. DV-based 25 Mbit/s with a sampling structure of 4:2:0 should be confined to special applications.
- MPEG-2 4:2:2P@ML using both intra-frame encoding (I) and GoP structures and data-rates up to 50 Mbit/s [12,13]. MPEG-2 MP@ML with a sampling structure of 4:2:0 should be confined to special applications.

The EBU strongly recommends that future networked television production should focus on compression families based on DV and MPEG-2 4:2:2P@ML which have been identified as being appropriate for television production operations.

---

10. Techniques for minimizing the quality loss in production and post-production operations, by direct manipulation of the compressed bitstream or by using special "helper data", are the subject of research.
11. EBU Test Report on New Recording Formats for News and Sports.
EBU Test Report on Subjective Tests of DV-based 50 Mbit/s Picture Quality.
12. For specific applications, this also includes MPEG-2 MP@ML if decodable with a single agile decoder.
13. For recording on a VTR, a fixed bit-rate must be agreed for each family member.

The EBU has issued Statement D-82: “M-JPEG in Networked Television Production”, to discourage its future use [14].

The discussions also revealed that the co-existence of different compression families [15] in their **native form** within both local and remote networked production environments would require the implementation of hardware-based **common agile decoders** [16]. In many instances, such decoders must allow “glitchless switching” and can therefore realistically be implemented within **one compression family** only. Manufacturers have stated that, within the foreseeable future, the coexistence and interoperation of **different compression families** requiring a “common agile decoder” within a networked television plant will pose a number of operational problems and will therefore be the exception and not the rule.

The positioning of the above compression families within a future networked digital production scenario requires careful analysis and differentiated weighting of the current and future potential influence of various technical constituents on that scenario.

C.3. Requirements for networked operation

In *Sections C.4.* and *C.5* of this annex., there is a discussion on compliance with the criteria introduced immediately below, and also a brief discussion on the results of official EBU tests that are relevant to future Networked Television Production. The details given are based on the current status of development and will be updated and amended as technology progresses and new implementations are introduced into the market place.

Members of the EBU and manufacturers participating in the work of the EBU / SMPTE Task Force have analyzed the following elements which are of particular and immediate importance to broadcasters.

C.3.1. Format stability

- availability of chip-sets;
- format commitment by each manufacturer;
- status of standardization.

C.3.2. Picture-quality ceiling, post-production potential, storage requirements

Information on the subjective and objective quality assessments carried out by the EBU on members of both compression families, within different applications and production scenarios, are briefly outlined in this report.

14. At the time of writing (July 1998), M-JPEG implementations had no defined structure to aid interoperability at the bitstream level. In order for M-JPEG to become acceptable for use in programme exchange, the following requirements have been identified:
    - The specification of the sampling structure and the arrangement of DCT coding blocks into a macroblock structure (containing associated luminance and chrominance blocks) must be defined.
    - The JPEG specification is defined for file storage. This must be extended to define the format for a sequence of JPEG files to create a M-JPEG stream. The stream specification must specify the format at both the sequence and picture layers of the stream and should include all parameters necessary for successful downstream decoding by a third party decoder.
    - Recommendations should be made for preferred modes of operation such as: the type of scanning, the resolution, the type of entropy coding etc.
    - Multi-generation tests should be completed to be able to assess the likely visual effects of the artefacts created. These tests should be carried out at bit-rates appropriate to the application area.
    - Further to that, it was acknowledged that M-JPEG does not provide features that one of the two preferred compression families could not provide as well.
15. A compression family is defined by its ease of intra-family bitstream transcoding and the availability of an "agile decoder" in integrated form.
16. Software-based agile decoding is currently not considered to be a practical option. **It is still undefined how an agile decoder will output the audio and Metadata part of the bitstream**.

As a first step, the EBU has divided the requirements for picture quality and post-production margin of networked broadcast applications into the following categories:

- News and Sports applications;
- Mainstream Broadcasting applications requiring more post-processing overhead.

See the EBU Statement given in *Section C.6.* of this annex.

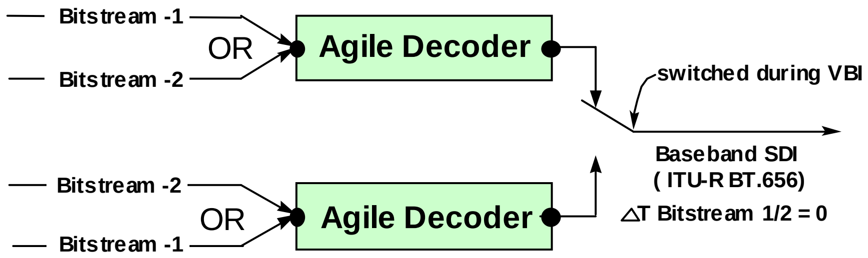
C.3.3. Interfaces

To allow smooth migration towards networked production operations, a stream interface will be required for use within a television production facility – for the flexible transport of packetized video, audio and Metadata over coaxial cable. Further to that, interfaces for different bearers, applications and functionalities will need to be standardized in the near future. (See *Section C.7.*)

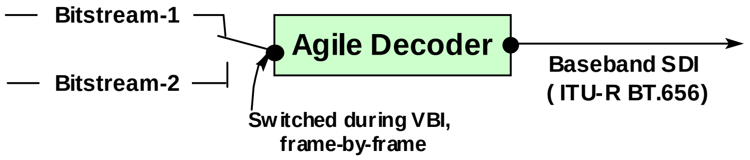
C.3.4. Intra-family agile decoders

Agile decoders for intra-family decoding [17] must be available in integrated form. They are expected to decode streamed real-time packetized video only. Such decoders should comply with the following requirements:

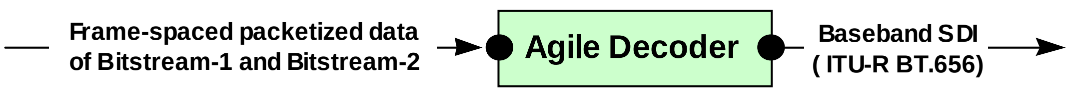
A. Decoding of different bitstreams with identical decoding delay at the output:



B. Intra-family switching between different bitstreams at the input:



C. Intra-family decoding between different bitstream packets within a single bitstream:



17. As an example, bitstream-1/2 in the above block diagrams could be:
Within the DV family – DV-based 25 Mbit/s (4:2.0 or 4:1:1), or DV-based 50 Mbit/s.
Within the MPEG family – MPEG-2-based 4:2:2P@ML, 18 Mbit/s, IB, or MPEG-2-based 4:2:2P@ML, 50 Mbit/s, I.

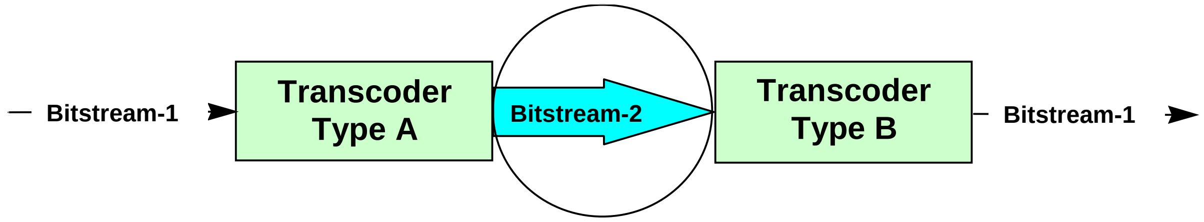
C.3.5. Native decoders

Native decoders, designed to operate on non-standard bitstreams, e.g. for optimized stunt-mode performance (shuttle, slow-motion) or for special functions, are acceptable. The decoder chip-set should be available on a non-discriminatory basis on fair and equitable conditions. Details of possible deviations from the standardized input data stream should be in the public domain.

C.3.6. Family relations

C.3.6.1. Tools available for intra-family transcoding

For reasons of restricted network bandwidth or storage space, a higher data-rate family member may have to be converted into a lower data-rate member. In the simplest case, this can be performed by simple decoding and re-encoding:



Under certain conditions, the quality losses incurred in this process can be mitigated by re-using the original encoding decisions. This can be performed within a special chip or by retaining the relevant information through standardized procedures.

C.3.6.2. Compatible intra-family record / replay

Operational flexibility of networked production will be influenced by the availability of recording devices which can directly record and replay all intra-family bitstreams or which allow the replay of different bitstreams recorded on cassettes.

C.3.7. Editing flexibility and complexity

For compressed data streams employing temporal prediction, the editing granularity of the compressed bitstream on tape – without manipulation of pixels within the active picture – will be restricted. Remote replay sources will require special control data and internal intelligence to allow frame-accurate editing.

C.3.8. Examples of commercial format implementations

- television tape recorders;
- disk storage;
- file servers.

C.3.9. Format development criteria

- A compression family **must** offer the potential for flexible interoperation between family members;
- It would be conceived as a **benefit** if the family allowed expansion to cope with restrictions imposed by special conditions in the areas of storage and Telco interconnection.

C.3.10. Test equipment

➩ Test equipment should be available on the market which allows conformance testing with the respective Standard specifications of all system modules.

C.4. Television production based on DV compression

C.4.1. Format stability

C.4.1.1. DV compression chip-set

The DV chip-set has been developed for consumer applications. It provides a broad application base with resultant economies of scale in commercial production. The chip-set can be configured for either processing a 4:1:1 sampling raster (“525-line countries”) or a 4:2:0 sampling raster (“625-line countries"”). This chip-set is used within camcorders for domestic and industrial use, designed by different manufacturers but increasingly used in professional ENG and studio applications. The 4:2:0 sampling raster requires additional vertical pre-filtering of the colour-difference channels to avoid aliasing. However, considerations of the cost and size of the vertical filter result in sub-optimum performance for professional applications.

25 Mbit/s compression is based on a DV compression chip-set, processing the video with a sampling raster of 4:1:1. The pre-filtering applied to the luminance and colour-difference signals is fixed and does not comply with the figures derived from a “real” 4:1:1 template. Details can be found in the EBU Report: “Tests on Panasonic DVCPRO / EBU Project Group P/DTR.”

50 Mbit/s compression is based on a combination of DV compression chip-sets. The chip-set processes standard 4:2:2 digital Video signals without additional pre-filtering. The chip required for pre-shuffling 4:2:2 DV-based 50 Mbit/s is manufactured by JVC exclusively. Details of compression performance can be found in the EBU Report: “Tests on JVC Digital-S / EBU Project Group P/DTR”.

Chip-set: (Note 1)	Available
Cost:	Consumer Oriented
Application base:	Consumer and Professional, Video and PC Market
Source DV@25 Mbit/s	Matsushita, Sony (Note 2), Toshiba, JVC
Source Shuffling@50 Mbit/s	JVC
Independent source:	Next Wave Technology
Application base:	Consumer, PC
Standards: (Notes 3, 4)	DV: IEC 61834 DV-based 25 Mbit/s, DV-based 50 Mbit/s: Draft SMPTE Standard (PT20.03)

Note 1: Panasonic and JVC have publicly stated their commitment to make the chip-set and appertaining documentation available to all interested parties on an equitable and non-discriminatory basis. This is the reason why DV chip-sets can already be found in a variety of different NLE and PC-based applications.

Note 2: DV consumer and DVCAM only

Note 3: The SMPTE is currently finalizing the details of the Draft Standard for the DVCPRO recording format, (D-7). Details of the mapping of DV macroblocks as well as mapping of digital audio and video data into the SDTI transport stream have recently been submitted to the SMPTE for standardization. The 4:1:1 filtering characteristic is an inextricable part of the Standard which allows broadcasters to retain a degree of predictability of resultant subjective picture quality after cascading. The DV chip-set does allow a degree of fine tuning for motion adaptation as a manufacturers option. In 50 Mbit/s configurations, the shuffling chip further allows a degree of flexibility to handle DCT coefficients.

Note 4: The SMPTE is currently finalizing the details of the Draft Standard for the Digital-S recording format, (D-9). Details of the mapping of DV macroblocks as well as mapping of digital audio and video data into the SDTI transport stream have recently been submitted to the SMPTE for standardization.

C.4.2. 25 Mbit/s intra-frame DV-based compression – basic characteristics for News and Sport

A/V Data-rate-Net Storage capacity / 90 min: (Note 1)	ca. 28 Mbit/s - ca. 19 Gbyte
Sampling raster / Net Video data-rate: (Note 2)	4:1:1 / 25 Mbit/s
Compression scheme: (Note 3)	DCT Transform, VRL with Macroblock pre-shuffling
Editing granularity: (Note 4)	One TV-frame
Quality at 1st Generation	Good, comparable with Betacam SP
Quality at 4th Generation:	Good, comparable with Betacam SP
Quality at 7th Generation:	Still acceptable, better than Betacam SP
Post-processing margin: (Note 5)	Small
Error concealment: (Note 6)	Acceptable

Note 1: The net A/V data-rate and the storage capacity required for a 90 min programme are within the data transfer- and storage volume capabilities of modern tape and hard-disk-based mass data-storage devices. The integration of DV-based compression transport streams into fully networked, robot-driven hierarchical storage-management systems, operating within a broad application base is therefore feasible.

Note 2: The picture quality achievable with the 4:1:1 sampling raster is inferior to the one defined for the 4:2:2 studio and is more closely related to best-possible decoded PAL-I quality. Although this has been obvious to the experts participating in the EBU tests, there was agreement however that, on average, the resultant resolution was still adequate for the applications envisaged.

Note 3: All DV-based compression formats feature special pre-sorting of the macroblocks prior to DCT and VRL encoding. With that exception, DV compression can be considered a member of frame-bound, conventional compression systems. The achievable signal quality of such a system has been tested by the EBU Project Group P/DTR.

Note 4: DV-based compression is frame-bound and allows simple assemble and insert edits of the compressed signal on tape and disk, thus avoiding lossy decompression and re-compression. However, for edits requiring access to individual pixel elements (wipes, re-sizing, amplitude adjustments), the signals have to be decoded.

Note 5: Post-production potential with 4:1:1 DV-based 25 Mbit/s compression is limited, due to the combined effects of reduced chroma-signal bandwidth and the progressive accumulation of compression artefacts.

Note 6: Compressed video signals require elaborate Forward Error Correction schemes to guarantee data integrity if routed through noisy channels. An overload of the Forward Error Correction system results in the loss of complete macroblocks. Concealment is the obvious solution to cope with such situations; completely erroneous macroblocks can be substituted with spatially-adjacent ones although this will achieve only limited results. The DV-based 25 Mbit/s compression format allows for the substitution of erroneous macroblocks by spatially-coinciding macroblocks from the preceding frame with acceptable results. Frame-bound compression prevents error propagation in this case.

C.4.3. 50 Mbit/s, 4:2:2 intra-frame DV-based compression – basic characteristics for mainstream broadcast production

A/V Data-rate-Net Storage capacity / 90 min (Note 1)	ca. 58 Mbit/s - ca. 39 Gbyte
Sampling raster:	4:2:2
Compression scheme: (Note 2)	DCT, VRL with Macroblock pre-shuffling
Editing granularity: (Note 3)	One TV-frame
Quality 1st Generation: (Note 4)	Identical to Digital Betacam
Quality 4th Generation: (Note 4)	Similar to Digital Betacam
Quality 7th Generation: (Note 4)	Comparable, slightly worse than Digital Betacam
Post-processing margin: (Note 5)	Adequate
Error concealment: (Note 6)	Acceptable

Note 1: The net A/V data-rate and a storage capacity required for a 90 min programme are within the data transfer- and storage volume capabilities of modern tape- and hard-disk-based mass data-storage devices. The integration of DV-based 50 Mbit/s compression transport streams into fully networked, robot-driven hierarchical storage-management systems, operating within a broad application base is therefore feasible.

Note 2: All DV compression formats feature special pre-sorting of the macroblocks prior to DCT and VRL encoding. With that exception, DV compression can be considered a member of frame-bound, conventional compression systems.

Note 3: DV-based compression is frame-bound and allows simple assemble and insert edits of the compressed signal on tape and disk, thus avoiding lossy decompression and re-compression. However, for edits requiring access to individual pixel elements (wipes, re-sizing, amplitude adjustments), the compressed signals have to be decoded.

Note 4: At the normal viewing distance, the picture quality of 1st generation DV-based 50 Mbit/s was practically indistinguishable from the 4:2:2 source. At normal viewing distance, experts had difficulty to identify differences between the performance of DV-based 50 Mbit/s through all generations for non-critical sequences. No significant decrease of picture quality was observed up to the 7th generation. In direct comparison with the source, critical sequences processed by DV-based 50 Mbit/s showed a certain softening of sub-areas containing high picture detail. This effect could be observed with a slight increase through each generation. In general, the level of impairment of 7th generation does not compromise picture quality.

Note 5: DV-based 50 Mbit/s compression does not employ pre-filtering. Post-processing margin up to the 7th generation has been rated as adequate for mainstream broadcasting applications.

Note 6: Compressed video signals require elaborate Forward Error Correction schemes to guarantee data integrity if routed through noisy channels. An overload of the Forward Error Correction system results in the loss of complete macroblocks. Concealment is the obvious solution to cope with such situations by substituting complete erroneous macroblocks with other ones. These can be spatially and / or temporally adjacent macroblocks. Concealment is independent of DV-based 50 Mbit/s compression and can be implemented in different ways depending on the application in actual products. DV-based compression processes in segments of five macroblocks, thus, preventing error propagation beyond one video segment of five macroblocks.

C.4.4. Subjective test results when following ITU-R Recommendation BT.500-7

Picture quality of 4:1:1 DV-based 25 Mbit / s and 4:2:2 DV-based 50 Mbit / s compression schemes have been evaluated subjectively and objectively within a variety of different operating scenarios. The sequences below were presented in the test in 4:2:2 quality (**absolute reference**) and in Betacam SP [18] quality (**relative reference)**.

The subjective tests were performed in accordance with the rules given in ITU-R BT 500-7 for the application of the "Double Stimulus Continuous Quality Scale, (DSCQS)" method which entails two different viewing distances: four times picture height (4H) for the critical viewing distance, and six times picture height (6H) for the normal viewing distance. The range of quality ratings extends from bad - poor - fair - good - excellent within a linear scale. The difference between the perceived quality of the reference and the system under test is subsequently evaluated and presented on a scale ranging from 0 to 100%. The 12.5% border is defined as the Quasi Transparent Threshold (QTT) of visibility. The processed subjective quality results do not scale linearly. In pictures rated 30%, degradation is quite visible.

C.4.5. Picture quality of a 4:1:1 DV-based 25 Mbit/s compression scheme

With this compression scheme, the proposed operating scenarios range from acquisition-only to hard news and magazine production. The picture Content of the sequences represents actions that are frequently encountered in both News and Sport.

C.4.5.1. Results obtained for sequences subjected to 1st generation post-processing

C.4.5.1.1. Comments for viewing distance 4H

- The average picture-quality ratings (see *Fig. C.1*) were dependent on picture Content and source quality. Even for the most demanding source picture, "Mobile and Calendar", the rating for picture-quality degradation was still below the "transparency" limit of 12.5%.

18. Different Betacam SP recorders were used for the 4:1:1 DV-based 25 Mbit/s and 4:2:2 DV-based 50 Mbit/s tests. In both cases, the recorders were in current post-production use and were not specially selected or realigned. The histograms for the 7th generation performance in both test series clearly show the variance of test results achieved with analogue equipment.

- In general, the average picture-quality degradation within the range of pictures under test was well below the 12.5% mark.
- In general, picture-quality degradation caused by the compression algorithm were rated as more visible than those of Betacam SP. These differences were within the range of the standard deviation and are therefore statistically insignificant.

C.4.5.1.2. Comments for viewing distance 6H

- The same tendency was found in the voting here as for the 4H case given above, but was less pronounced due to the reduced eye sensitivity at a viewing distance of 6H.
- In general, the average picture-quality degradation for DVCPRO within the range of pictures under test was well below the 12.5% mark.

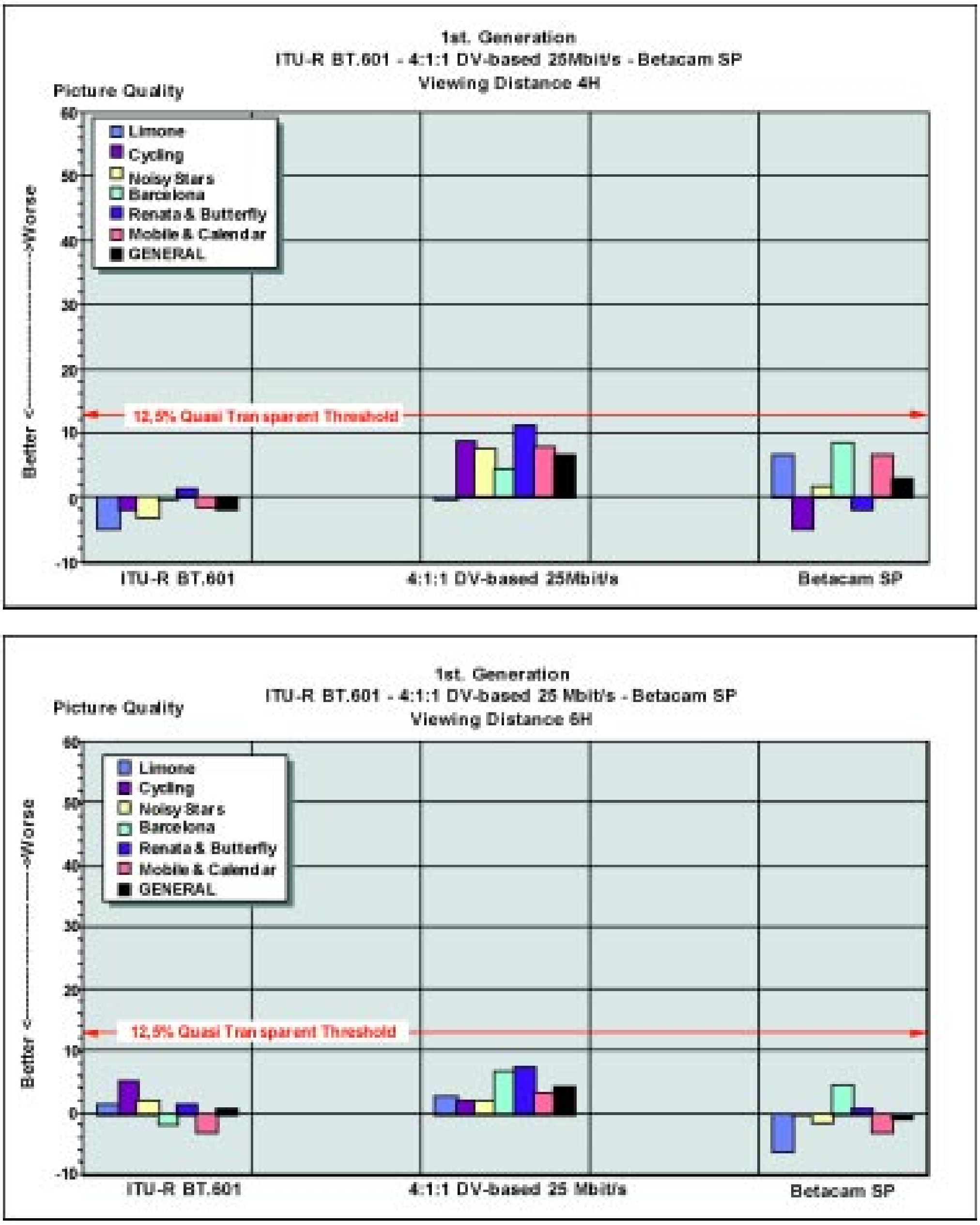


Figure C.1 : 4:1:1 DV-based 25 Mbit/s compression scheme – first-generation picture quality at viewing distances of 4H and 6H.

C.4.5.2. Results obtained for sequences subjected to 4th generation post-processing

The post-production scenario encompassed four generations of 4:1:1 DV-based 25 Mbit/s processing, two of which involved one temporal shift and one spatial shift each.

C.4.5.2.1. *Comments for viewing distance 4H*

- In general, the average picture-quality degradation for the range of pictures under test was still below the 12.5% mark as the defined limit for "transparency" for both DV compression and Betacam SP (see *Fig. C.2*).
- The artefacts produced by DV compression in the natural scenes of the test cycle remained below the threshold of visibility, even at this critical viewing distance.
- For natural scenes, differences in quality shown in the histogram between DV compression and Betacam SP are statistically insignificant.

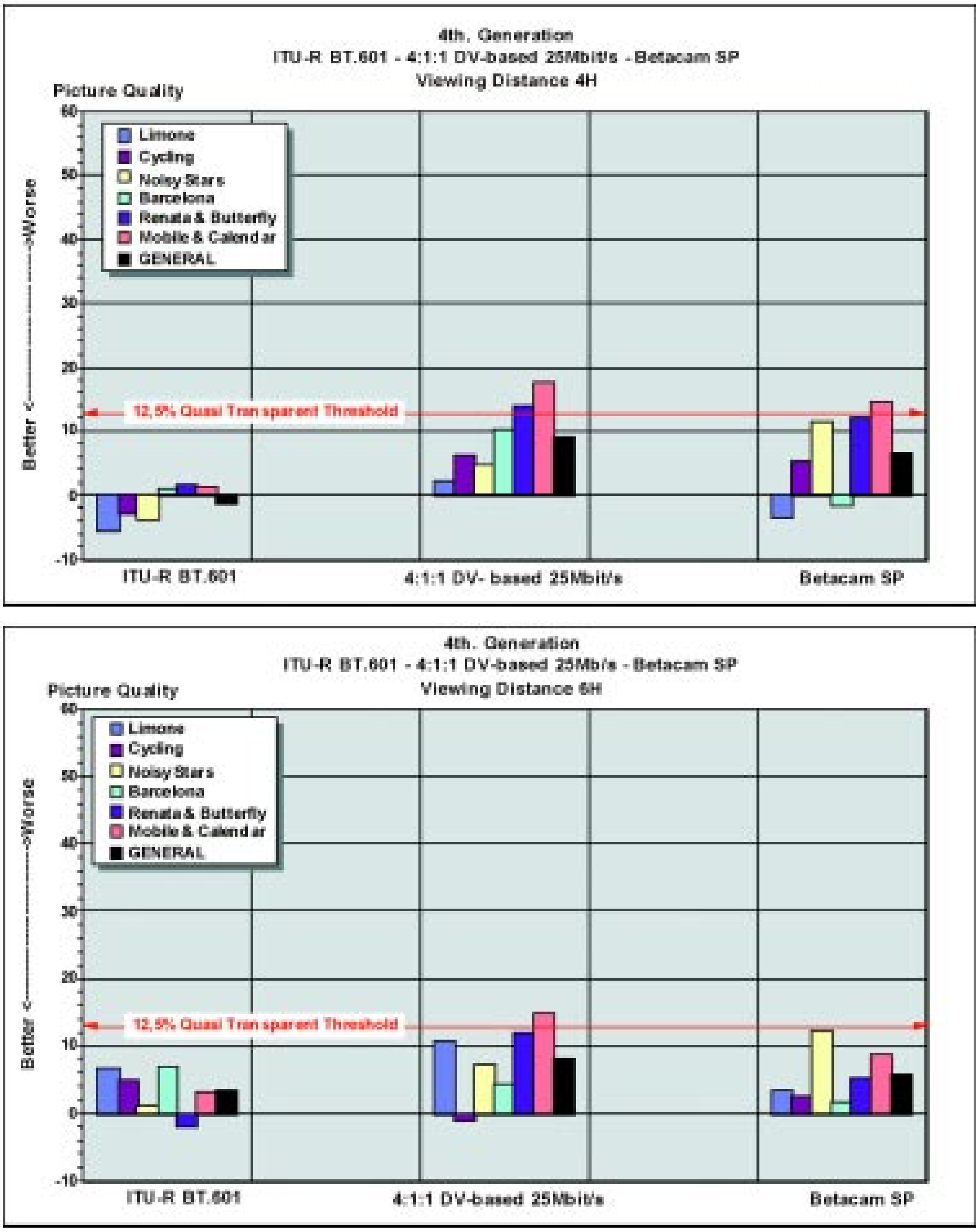


Figure C.2 : 4:1:1 DV-based 25 Mbit/s compression scheme – fourth-generation picture quality at viewing distances of 4H and 6H.

- Only for critical test scenes “Renata-Butterfly” and “Mobile and Calendar”, DV compression exceeded that limit and the picture quality of Betacam SP was judged better than that of DV. The differences are statistically insignificant and the quality of both formats could therefore be rated as comparable.

C.4.5.2.2. Comments for viewing distance 6H

- The absolute ratings for both DVCPRO and Betacam SP are lower than in the 4H case shown above. For natural pictures, differences shown in the histogram (*Fig. C.2*) between DVCPRO and Betacam SP are statistically insignificant.
- Only in the case of “Renata & Butterfly”, DVCPRO slightly exceeded the transparency limit of 12.5%.
- In general, the average picture-quality degradation for the range of pictures under test was still below the 12.5% mark as the defined limit for “transparency” for both DVCPRO and Betacam SP.

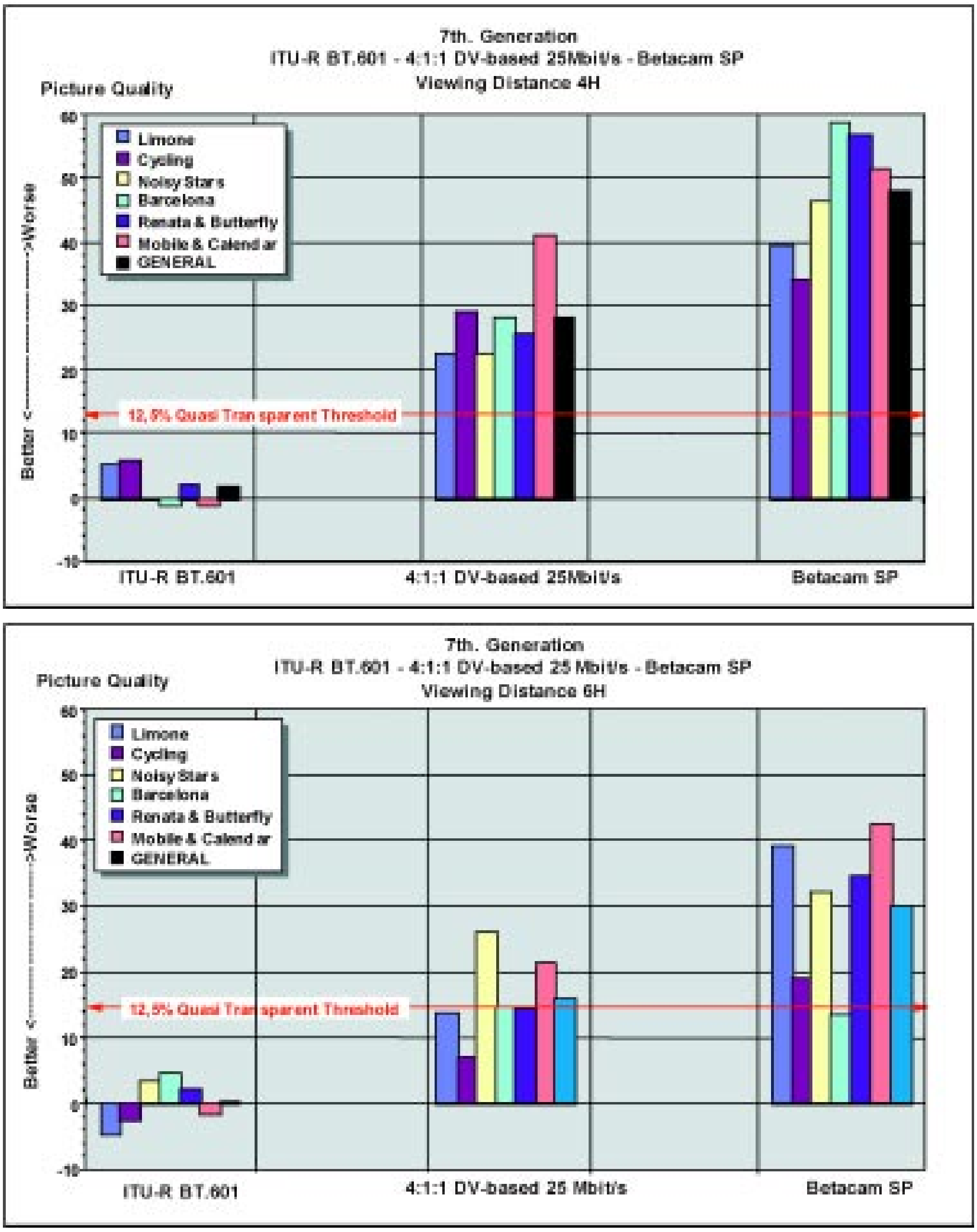


Figure C.3 : 4:1:1 DV-based 25 Mbit/s compression scheme – seventh-generation picture quality at viewing distances of 4H and 6H.

C.4.5.3. Results obtained for sequences subjected to 7th generation post-processing

The post-production scenario encompassed seven generations of 4:1:1 DV-based 25 Mbit / s processing, three of which involved one temporal shift and two spatial shifts each.

C.4.5.3.1. Comments for viewing distance 4H

- The picture degradation produced by DV compression in this operating scenario considerably exceeded the threshold of “transparency” for all test sequences (see *Fig. C.3*).
- For the critical test sequences “Mobile & Calendar”, the limit was exceeded significantly.
- On average, for both normal and critical pictures, the footprints created by DV compression were rated far below the degradation generated by Betacam SP.
- Although the threshold of visibility was exceeded in all cases, the acceptance level of picture quality achieved within this DV post-production scenario will depend on the individual broadcaster’s attitude on the acceptance of Betacam SP picture quality in an identical operating scenario.

C.4.5.3.2. Comments for viewing distance 6H

- The absolute ratings are lower than in the 4H case described above.
- In all but one case, the ratings for DVCPRO exceeded the transparency limit.
- Analogue Betacam was rated markedly worse than DVCPRO in practically all cases.

C.4.6. *Picture quality of a 4:2:2 DV-based 50 Mbit/s compression scheme*

The proposed operating scenario is that of networked mainstream broadcast operations. To assess the picture quality and post-processing ceiling obtainable with 4:2:2 DV-based 50 Mbit / s compression, Digital Betacam was included in the test as an established high-end compression system.

The results given below were obtained for viewing distances at 4H (34 observers) and 6H (26 observers) from a subjective test carried out by the RAI and the IRT on a 4:2:2 DV-based 50 Mbit / s compression scheme.

C.4.6.1. Results obtained for sequences subjected to 7th generation post-processing and pixel shift

The picture sequences were subjected to 7th generation post-processing with the pixel shift characteristics given in the table below.

Processing	Horizontal Shift (Pixel) +1 = 2 Y pixel shift right ll-1 = 2 Y pixel shift left	Vertical Shift (Line) +1 = 1 line shift down ll -1 = 1 line shift up
1st generation →2nd generation.	no shift	+1
2nd generation →3rd generation.	no shift	+1
3rd generation →4th. generation.	no shift	+1
4th generation →5th generation.	+1	no shift
5th generation →6th generation.	no shift	-1
6th generation →7th generation.	-1	-2

Note: The “Diva with Noise” sequence was originally included in the test. This sequence is an extreme test for all compression systems. The “General” result, expressed as numerical values on the histograms above, represents the average over the sequences tested without inclusion of the “Diva with Noise” test sequence.

C.4.6.1.1. Multi-generation performance of Digital-S (DV-based 50 Mbit/s) compression

- At the normal viewing distance, the picture quality of 1st generation Digital-S compression (see *Fig. C.4*) was practically indistinguishable from the 4:2:2 source.
- At the normal viewing distance, experts had difficulty in identifying differences between the performance of Digital-S compression through all generations for non-critical sequences. No remarkable decrease of picture quality was observed up to the 7th generation.
- In direct comparison with the source, the critical sequences processed by Digital-S compression showed a certain softening of sub-areas containing high picture detail. This effect could be observed with a slight increase through each generation. In general, the level of impairment of seventh generation does not compromise the picture quality.

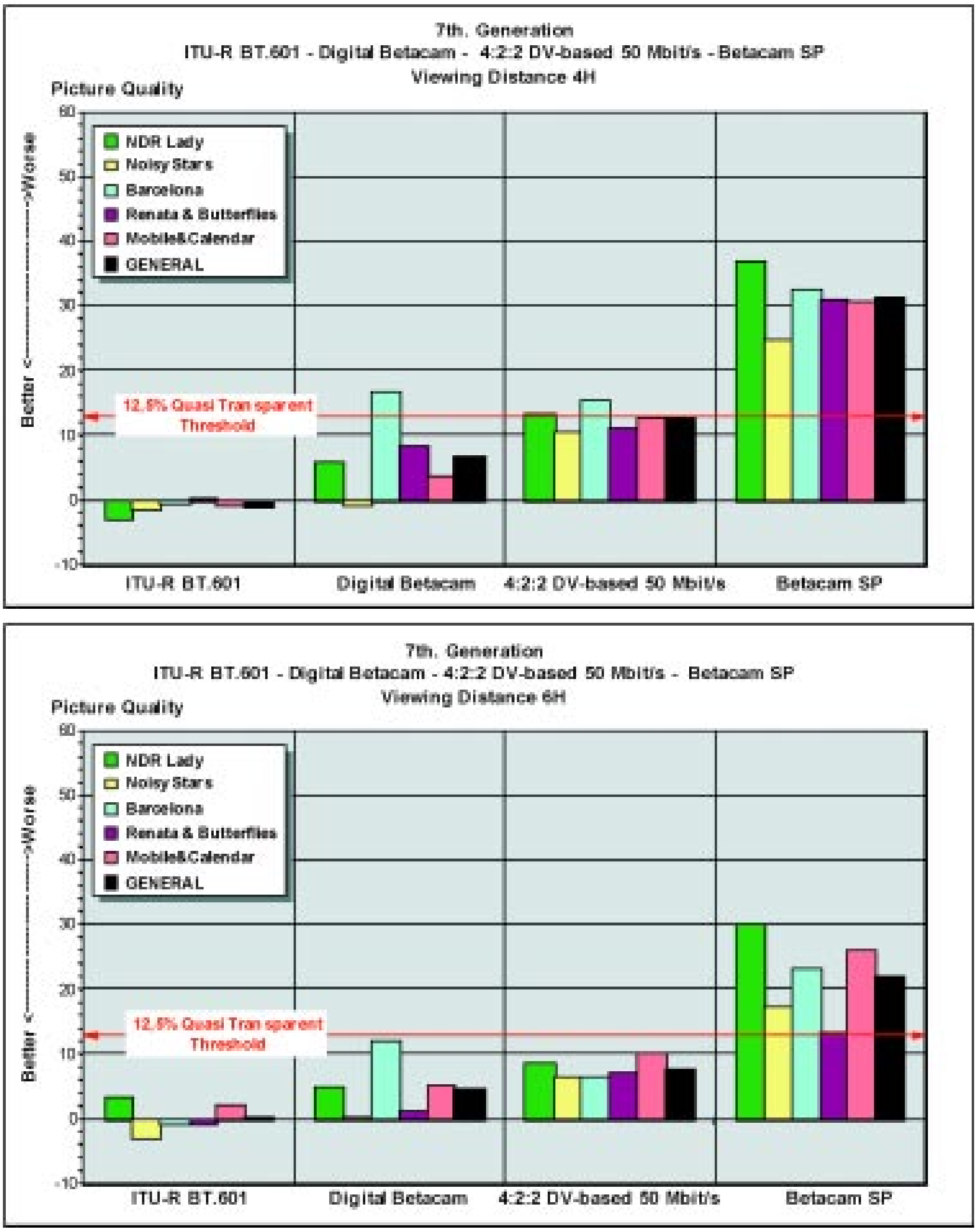


Figure C.4 : 4:2:2 DV-based 50 Mbit/s compression scheme – seventh-generation picture quality at viewing distances of 4H and 6H.

C.4.6.1.2. Multi-generation performance of Digital-S (DV-based 50 Mbit/s) compression and the compression used in Digital Betacam

- ⇨ At the normal viewing distance and for moderately critical sequences, experts had difficulty to identify differences between the performance of the algorithms of the two digital compression systems.
- ⇨ For the first generation, the performance of Digital-S compression and the compression used in Digital Betacam was rated to be identical.
- ⇨ At fourth generation, the performance of Digital-S compression and the compression used in Digital Betacam in the multi-generation scenario is similar. The picture quality of non-critical sequences is practically preserved by both systems. Differences between system performance are detectable on closer scrutiny and can be described as “softness in areas of high picture detail” for Digital-S compression and “increased coding noise” for compression used in Digital Betacam.
- ⇨ At seventh generation, the different behaviour of the two compression algorithms becomes more apparent. The effects described for the fourth generation performance are slightly accentuated. For moderately critical sequences, the level of impairment was very low and did not compromise the overall picture quality. On direct comparison, picture quality provided by Digital Betacam compression was considered to be slightly better than that achieved with Digital-S compression. This is mainly due to the different subjective effects of “softness” and “coding noise” on perceived picture quality.

C.4.7. Digital interfaces

The table below indicates the type of interface required for a DV-based compression family and the respective status of the specification.

Interface:	Status			
	Defined	In progress	Not defined	Standard Document
SDTI	✓			SMPTE 305 M
ATM		✓		
FC	✓			
IEEE-1394	✓			
T-3			✓	
OC-3			✓	
Satellite			✓	

C.4.8. Intra-family agile decoders

C.4.8.1. Market prospects

Panasonic and JVC have stated their commitment to produce an agile decoder chip which performs 4:2:0 / 4:1.1 DV-based 25 Mbit/s and 4:2:2 DV-based 50 Mbit/s decoding. Both companies have further stated that the agile decoder will also decode DVCAM bitstreams.

C.4.8.2. Decoding of different DV bitstreams with identical decoding delay at the output

The general feasibility of seamless switching between DV-based 25 Mbit/s and DV-based 50 Mbit/s bit input streams at SDI output level has been demonstrated.

The agile DV decoder will comply with **requirement A** in *Section C.3.4*.

C.4.8.3. Intra-family switching between different DV bitstreams at the input

The agile DV decoder will comply with **requirement B** in *Section C.3.4*.

C.4.8.4. Intra-family decoding between different DV packets within a single bitstream

The agile DV decoder will comply with **requirement C** in *Section B.3.4*.

C.4.9. Native decoders

DV decoders are native by definition.

C.4.10. Family relations

C.4.10.1. Tools available for intra-family transcoding

No special tools are currently available. All DV-based compression schemes feature the same basic compression and macroblock structure. However, different sampling structure and horizontal / vertical pre-filtering requires transcoding via baseband decoding.

C.4.10.2. Compatible intra-family record / replay

Input	DV		DVCAM		DVCPRO@25		DVCPRO@50		Digital-S	
Tape Cassette	**REC**	**PLAY**	**REC**	**PLAY**	**REC**	**PLAY**	**REC**	**PLAY**	**REC**	**PLAY**
DV Small	Y	Y	Y	Y	N	Y(*)	N	Y(*, **)	N	N
DV Large	Y	Y	Y	Y	N	Y	N	Y(*)	N	N
DVCAM Small	N	Y	Y	Y	N	Y(*)	N	Y(*, **)	N	N
DVCAM Large	N	Y	Y	Y	N	Y	N	Y(**)	N	N
DVCPRO@25 Medium	N	N	N	N	Y	Y	Y	Y	N	N
DVCPRO@25 Large	N	N	N	N	Y	Y	Y	Y	N	N
DVCPRO@50 Medium	N	N	N	N	N	N	Y	Y	N	N
DVCPRO@50 Large	N	N	N	N	N	N	Y	Y	N	N
Digital-S	N	N	N	N	N	N	N	N	Y	Y

Y = yes, N= no

(*): With adapter

(**): Presently-available equipment does not support this functionality. Equipment under development will support play-back of DV and DVCAM formats.

C.4.11. Editing flexibility and complexity

- ➪ No Edit restrictions because of frame-bound prediction window. Editing granularity is 1 frame.
- ➪ No special host / client interactions required. Assemble and insert edit at the bitstream level is possible.
- ➪ Edit control via RS 422 or RS 232 protocol.

C.4.12. Examples of some commercial format implementations

C.4.12.1. DV-based 25 Mbit/s tape recording format ("DVCPRO")

Tape:	6.35 mm Metal Particle, 8.8m
Cassettes:	Medium (up to 63 min.), Large (up to 125 min)
Robustness:	Good, adequate for News & Sports (Note 1)
Slow-Motion and Shuttle:	Limited, requires adaptation to picture build-up (Note 2)
DV replay:	Yes (Note 3)
Audio channels:	2 channels, 48 kHz, 16 bit (Note 4)
Audio 1 / 2 cross-fade:	5 frames (Note 5)
Editing granularity on tape:	1 frame
Transparency of VBI:	Transparent, limited to ca. 553 kbit/s, (Note 6)
SDTI Interface:	Availability pending (Note7)

Note 1: The error rates measured under realistic stress conditions during the EBU test have confirmed that DVCPRO 25 Mbit/s is adequately robust within the application field envisaged for the recording format. The error rates measured off-tape with error correction completely switched off were in the range of 10-5 and 10-6, thus indicating a solid design of the complete head-to-tape area.

Note 2: For Slow-Motion, the replay quality is heavily dependent on the chosen setting. The identification of short audio and video inserts during Shuttle is heavily dependent on the individual setting of the shuttle speed. Identification is further influenced by Content and length of the segment to be located. Details can be found in the EBU Test Report.

Note 3: This requires the use of an adapter cassette. Replay of DV and DVCAM recordings in the original DV format is possible without quality loss.

Note 4: One out of 65536 codes is reserved for error signalling. Code 8000h becomes 8001h.

Note 5: This operational mode seems to be uniquely confined to Europe. The 5-frame delay measured with current implementations is not acceptable and will be remedied, according to Panasonic, by an optional, external RAM delay.

Note 6: Access is currently not implemented. Data located within this area will be recorded transparently. No cross influence between VBI data and active picture.

Note 7: The SDTI interface for the transport of audio, compressed video and Metadata in real-time and non-real-time has recently passed the final ballot with the SMPTE (SMPTE 305 M); practical implementation into digital equipment is planned for 4Q/98.

C.4.12.2. DV-based 50 Mbit/s tape recording format ("DVCPRO50")

Tape:	6.35mm Metal Particle, (61 min,8.8m; 92min, 6,5m)
Cassettes:	Medium (up to 31 min.), Large (up to 92 min)
Robustness:	Good, adequate for Mainstream Broadcast Applications (Note 1)
Slow-Motion and Shuttle:	Significant improvement in shuttle performance compared to DVCPRO (Note 2)
DV replay:	Not tested, will be implemented in a specific model only
Audio channels	4 channels, 48 kHz, 16 bit (Note 3)
Audio 1 / 2 cross fade:	6 frames (Note 4)
Editing granularity on tape:	1 frame
Transparency of VBI:	Transparency limited to 2 uncompressed lines / frame, (Note 5)
SDTI Interface:	Availability pending (Note6)

Note 1: The error rates measured under realistic stress conditions during the EBU test have confirmed that DVCPRO50 Mbit/s is adequately robust within the application field envisaged for the recording format. The error rates measured off-tape with error correction completely switched off were in the range of 10-5 and 10-6 , thus indicating a solid design of the complete head-to-tape area. The EBU tests were carried out with 12.5m tapes only. The behaviour of 6.5m tape remains to be assessed.

Note 2: The identification of sequence transitions with increasing shuttle speeds has been greatly improved. This also applies to DVCPRO recordings in shuttle playback on DVCPRO50 machines. Details can be found in the EBU Test Report.

Note 3: One out of 65536 codes is reserved for error signalling. Code 8000h becomes 8001h.

Note 4: This operational mode seems to be uniquely confined to Europe. The 6-frame delay measured in the EBU test is not acceptable and will be remedied, according to Panasonic, by an optional, external RAM delay.

Note 5: Access currently not implemented. Data located within this area will be recorded transparently. Cross influence between VBI data and active picture by 9 compressed VBI lines possible.

Note 6: The SDTI interface for the transport of audio, compressed video and Metadata in real-time and non-real-time has recently passed the final ballot with the SMPTE (SMPTE 305 M); practical implementation into digital equipment is planned for 4Q/98.

C.4.12.3. 4:2:2 DV-based 50 Mbit/s tape recording format ("Digital-S")

Tape:	12.5 mm Metal Particle, 14,4 µ (104'), 12.4 µ (124')
Cassettes:	One size, up to 124 minutes
Robustness:	Mainstream broadcast applications, News and Sports (Note 1)
Slow-Motion and Shuttle:	Limited, requires adaptation to picture build-up (Note 2)
Audio channels:	2 channels, (Note 3); 48 kHz, 16 bit (Note 4)
Audio 1 / 2 cross fade:	O.K.
Editing granularity on tape:	1 frame
Transparency of VBI:	Transparent, limited to ca. 2,88Kbytes / frame,
SDTI Interface:	planned, but not yet implemented (Note 6)

Note 1: The error rates measured under realistic stress conditions during the EBU test have confirmed that Digital-S is adequately robust within the application field envisaged. The error rate of 10-6 measured off-tape with error correction completely switched off indicates a solid design of the complete head-to-tape area.

Note 2: For Slow-Motion, the replay quality is heavily dependent on the chosen setting. The identification of short audio and video inserts during Shuttle is heavily dependent on the individual setting of the shuttle speed. Identification is further influenced by Content and the length of the segment to be located. Details can be found in the EBU Test Report.

Note 3: Will be upgraded to four channels

Note 4: One out of 65536 codes is reserved for error signalling. Code 8000h becomes 8001h.

Note 5: Data located within this area will be recorded transparently. No cross influence between VBI data and active picture.

Note 6: The SDTI interface for the transport of audio, compressed video and Metadata in real-time and non-real-time has recently passed the final ballot with SMPTE (SMPTE 305 M); practical implementation into Digital-S equipment is therefore imminent.

C.4.13. NLE equipment

- ➪ 4:1:1 DV-based 25 Mbit/s is available;
- ➪ 4:2:2 DV-based 50 Mbit/s is under development.

C.4.14. Format development potential

The following options within the DV-based tape format family *potentially* exist. Implementation will depend on user demand:

- ➪ Replay-compatibility of DV (DVCAM) recordings on DVCPRO;
- ➪ Replay-compatibility of DVCPRO recordings on DVCPRO50;
- ➪ An integrated decoder for three DV-based tape formats (DVCPRO, DVCPRO50 and Digital-S) has been demonstrated;
- ➪ Optional recording of DVCPRO signals on a DVCPRO50 recorder;
- ➪ 1/4 reduction of transfer time from tape / tape and tape / Hard Disk / tape with DVCPRO and DVCPRO50;

- Integration of DV-based tape-format supports into workstations for direct PC-based post-processing;
- Two times real-time digital audio and compressed digital video transfer using SDTI interface (SMPTE 305 M) for Digital-S.

C.4.15. Test equipment

Will be available.

C.5. Television production based on MPEG-2 4:2:2P@ML compression

C.5.1. Format stability

C.5.1.1. MPEG-2 4:2:2P@ML compression chip-sets

Chip-sets:	Available
Source 1:	IBM
Operating range:	Up to about 50 Mbit/s
Cost:	Oriented towards professional market
Application base:	Telecom, computer and professional
Source 2:	C-Cube DV^X
Operating range:	Up to 50 Mbit/s, GoP: I, IB, others
Cost:	Oriented towards consumer & professional market
Application base:	Telecom, computer, consumer, and professional (Note 1)
Source 3:	Sony(Notes 2, 3)
Operating range:	15-50 Mbit/s, GoP: I, IB
Cost:	Oriented towards professional market
Application base:	Telecom, computer and professional
Source 4:	Fast
Operating range:	50 Mbit/s and higher
Cost:	Oriented towards professional market
Application base:	Intra-frame only. Telecom, computer and professional
Standard:	MPEG-2 4:2:2P@ML standardized by MPEG group. Transport protocols submitted to SMPTE for standardization (Note 4)

Note 1: Chip-sets are available for both professional and consumer applications

Note 2: Sony has publicly stated its commitment to make the SX chip-set available together with the appertaining software documentation to all interested parties on an equitable and non-discriminatory basis. For reasons of optimum VTR stunt-mode operation, the arrangement of coefficients within macroblocks within the Betacam SX native data stream differs from that of an MPEG-compliant data stream.

Note 3: Sony has announced its intention to produce “data re-ordering” chips which transparently translate the Betacam SX native data stream to a fully MPEG-compliant data stream (and vice versa). These chips will be made available to all interested parties on an equitable and non-discriminatory basis.

Note 4: MPEG compression allows great flexibility in encoder design. The balance of the great number of encoding parameters to achieve optimum quality is a manufacturer’s choice and need not be documented. Since most pre-processing, such as filtering or noise reduction, is not always required, the pre-processing parameters may be selected depending upon the nature of the images and the capabilities of the compression system. These choices can be pre-set or can be adaptive. The multi-generation performance of individual codec designs with identical data-rate and GoP structure but from different manufacturers and with different pre-settings will therefore be difficult to predict and will require subjective testing in each case.

C.5.2. MPEG-2 4:2:2P@ML – an estimate of first-generation performance

The diagram below (*Fig. C.5*) shows the data-rates required to encode picture sequences of different coding complexity. Software codecs were used and data-rates were adjusted to achieve equal output performance in terms of noise power of the differences between the original and the compressed / decompressed picture sequence (PSNR =40 dB). The prediction window could be adjusted and varied between GoPs in the range of 1 to 15 (Notes 1, 2, 3).

Note 1: The curves shown in the diagram should be taken as an **indication of first-generation performance** within the wide span of MPEG-2 encoding options only. Taking signal differences as a measure of picture quality only allows coarse evaluation of actual quality performance. The variance of encoding parameters allowed in MPEG-2 encoding structures to achieve the desired flexibility will require subjective testing of each individual encoder design to determine the actual quality performance at a given data-rate and GoP structure. The arrows indicate possible members of the MPEG-2 4:2:2P@ML compression family envisaged for Mainstream Broadcasting, News and Sports and for Contribution, as implemented in current industrial designs.

Note 2: Sony has demonstrated an MPEG-2 4:2:2P@ML at 50 Mbit/s (Intra-frame) implementation to the EBU. A formal subjective test of the picture quality obtained with the parameter settings chosen has been carried out by the EBU.

Note 3: The EBU has evaluated the performance of MPEG-2 4:2:2P@ML at 21 Mbit/s operation as selected for the EBU contribution network. Results of a subjective test will be incorporated in the document at the time of IBC 98.

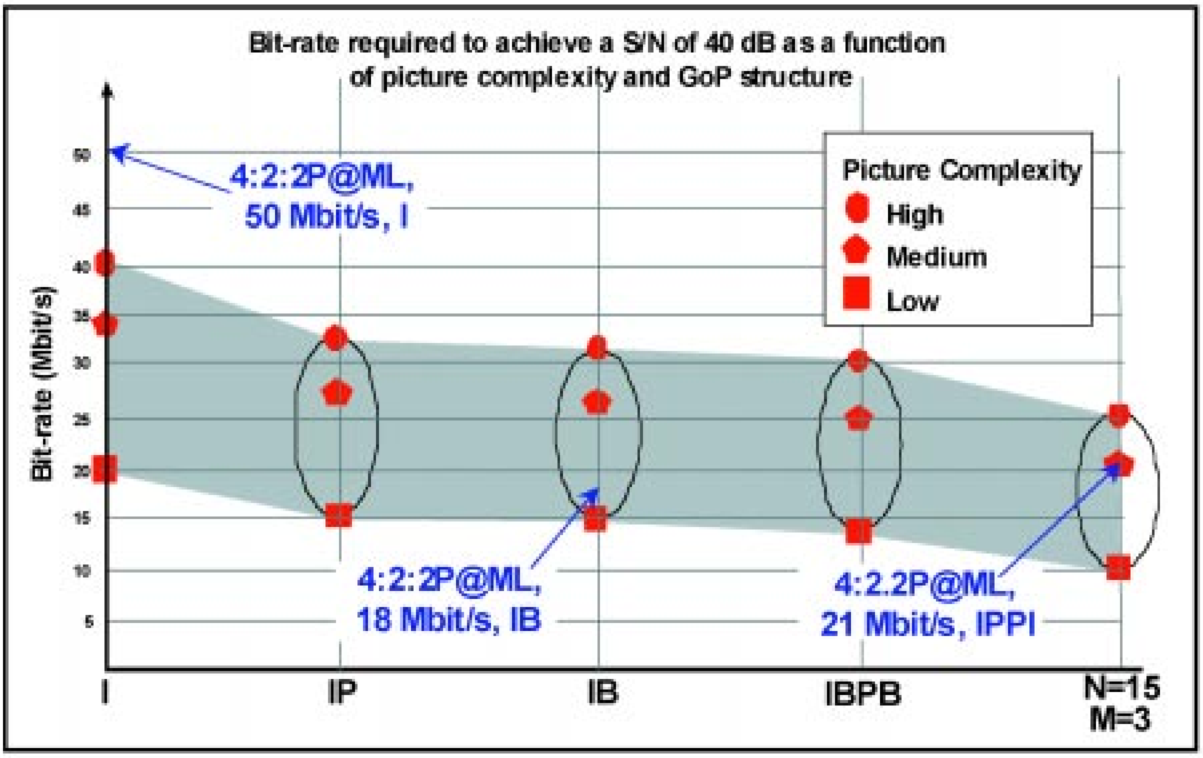


Figure C.5: Basic characteristics of compression for News and Sports – MPEG-2 4:2:2P@ML, 18 Mbit/s, IB (Note 1).

A/V Data-rate-Net Storage capacity / 90 min:	ca. 21 Mbit/s - ca.14 Gbyte (Note 2)
Sampling raster:	4:2:2 (Note 3)
Compression scheme:	DCT, VRL MPEG-2 4:2:2P@ML, GoP=2, IB
Editing granularity:	2 frames without frame modification (Note 4)
Quality at 1st Generation:	Good, comparable with Betacam SP
Quality at 4th Generation:	Good, comparable with Betacam SP (Note 5)
Quality at 7th Generation:	Still acceptable, better than Betacam SP (Note 6)
Post-processing margin:	Small (Note 7)
Error concealment:	Not practicable (Note 8)

Note 1: Betacam SX compression can be described as a subset of MPEG-2 4:2:2P@ML compression with a GoP of 2, based on an IB structure. For reasons of optimum VTR stunt-mode operation, the arrangement of coefficients within macroblocks within the Betacam SX native data stream differs from that of an MPEG-2-compliant data stream.

Note 2: The net A/V data-rate and the storage capacity required for a 90 min programme are within the data transfer and storage volume capabilities of modern tape and hard-disk-based mass data-storage devices. The integration of SX transport streams into fully networked, robot-driven hierarchical storage-management systems operating within a broad application base is therefore possible.

Note 3: The pre-filtering applied to the luminance and colour-difference signals does not comply with the figures derived from a "real" 4:2:2 template. Expert viewing tests have confirmed that, due to the pre-filtering used in the colour-difference channels, the picture quality obtainable with a 4:2:2 sampling raster compliant with the digital studio standard has not been achieved. Resolution obtainable with the current SX implementation is comparable to the one achievable with a 4:1.5:1.5 sampling raster. There was agreement however that, on average, the resultant resolution of picture details was still adequate for the applications envisaged.

Note 4: Simple frame-accurate assemble and insert edits of MPEG-2 compressed signals are strongly dependent on the relative position of the GoP structures within the data streams to be edited. Details can be found in the EBU Test Report: Tests on Sony SX / EBU Project Group P/DTR. This restriction is removed however if the GoP structure around the edit point can be changed. For edits requiring access to individual pixel elements (wipes, re-sizing, amplitude adjustments), the signals have to be decoded.

Note 5: The limits of the SX compression scheme become conspicuous once more-elaborate post-processing of sequences originating from ENG and Sports is required. Average picture quality has been rated to be still acceptable, but no longer good. Picture quality is still better than that obtained from Betacam SP under identical circumstances. In the case of Betacam SP, increasing number of generations beyond about 5 to 6 cause increasingly conspicuous spatial displacement of contiguous luminance and colour-difference samples.

Note 6: For the application envisaged, despite the artefacts of the SX compression scheme accumulated in progressive generations, post-production headroom for processes requiring access to individual pixels is still considered adequate.

Note 7: Post-production potential outside the recommended applications for SX compression is limited, due to the combined effects of reduced chroma-signal bandwidth and the progressive accumulation of compression artefacts.

Note 8: Because of the wide temporal prediction window, error concealment is not practicable and will lead to error propagation. Therefore sufficient margin must be allocated to the error-correction scheme and the format robustness to compensate for this.

C.5.3. MPEG-2 4:2:2P@ML, 50 Mbit/s, I, CBR – basic characteristics for mainstream broadcast production

A/V Data-rate-Net Storage capacity / 90 min	ca. 53 Mbit/s - ca. 36 Gbyte: (Note 1)
Sampling raster:	4:2:2 (Note 2)
Compression scheme:	DCT, VRL MPEG-2 4:2:2P@ML, 50 Mbit/s, GoP=1
Editing granularity:	One TV-frame
Quality 1st Generation:	Identical to Digital Betacam
Quality 4th Generation:	Similar to Digital Betacam
Quality 7th Generation:	Comparable, slightly worse than Digital Betacam
Post-processing margin:	Adequate
Error concealment:	Acceptable

Note 1: The net A/V data-rate and the storage capacity required for a 90 min programme are within the data transfer and storage volume capabilities of modern tape- and hard-disk-based mass data-storage devices. The integration of a 50 Mbit/s transport streams into fully networked, robot-driven hierarchical storage-management systems, operating within a broad application base, is therefore possible.

Note 2: No pre-filtering of luminance and colour-difference inputs was applied.

C.5.4. MPEG-2 4:2:2P@ML – subjective test results when following ITU-R Recommendation BT.500-7

The picture quality of MPEG-2 4:2:2P@ML compression schemes operating at different data-rates and GoP structures were evaluated subjectively and objectively within a variety of different operating scenarios.

The sequences presented in the test were in both 4:2:2 quality (**absolute reference**) and in Betacam SP (Note) quality (**relative reference).**

Note Different Betacam SP recorders were used for the MPEG-2 4:2:2P@ML, 18 Mbit/s, IB and MPEG-2 4:2:2P@ML, 50 Mbit/s, I-only tests. In both cases, the recorders were in current post-production use and were not specially selected or realigned. The histograms for the 7th generation performance in both test series clearly show the variance of test results achieved with analogue equipment.

The subjective tests were performed in accordance with the rules given in ITU-R BT 500-7 for the application of the "Double Stimulus Continuous Quality Scale, (DSCQS)" method which entails two different viewing distances: four times picture height (4H) for the critical viewing distance and six times picture height (6H) for the normal viewing distance. The range of quality ratings extends from bad - poor - fair - good - excellent within a linear scale. The difference between perceived quality of the reference and the system under test is subsequently evaluated and presented on a scale ranging from 0 to 100%.

The 12.5% border is defined as the Quasi Transparent Threshold (QTT) of visibility. The processed subjective quality results do not scale linearly. In pictures rated 30%, degradation is quite visible.

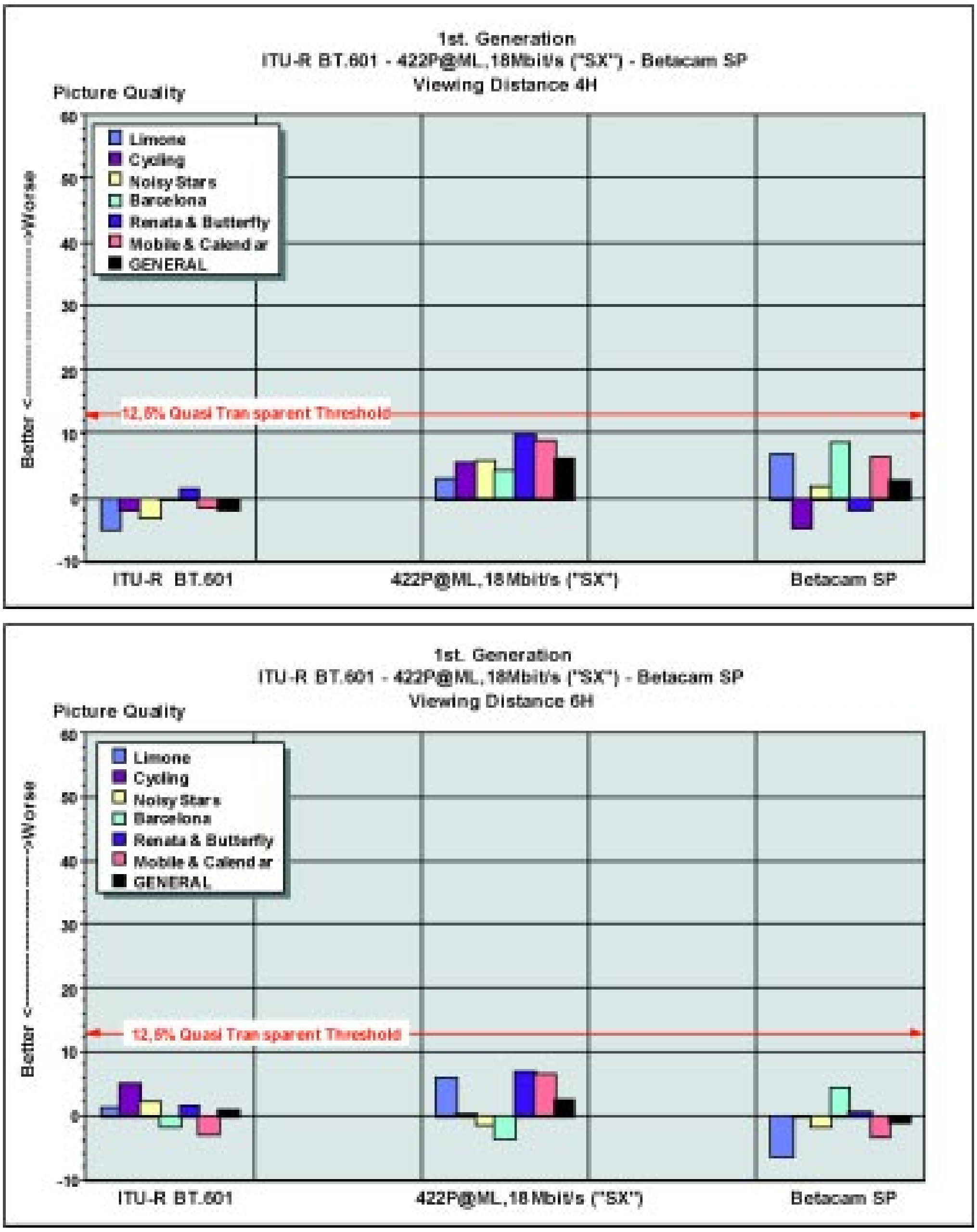


Figure C.6 : 4:2:2P@ML,18 Mbit/s, IB compression scheme – first-generation picture quality at viewing distances of 4H and 6H.

C.5.5. MPEG-2 4:2:2P@ML, 18 Mbit/s, IB – picture-quality issues

The proposed operating scenario ranges from acquisition only, to hard news and magazine production. The picture Content of the sequences represent actions that are frequently encountered in both News and Sport.

C.5.5.1. Results obtained for sequences subjected to 1st generation post-processing

C.5.5.1.1. Comments for viewing distance 4H

- The average picture-quality ratings (see *Fig. C.6*) were dependent on picture Content and source quality. Even for the most demanding source picture, "Mobile and Calendar", the rating for picture-quality degradation was noticeably below the "transparency" limit of 12.5%.
- In general, the average picture-quality degradation for the MPEG-2 4:2:2P@ML, 18 Mbit / s, IB compression within the range of pictures under test was well below the 12.5% mark as the defined limit for visibility.
- In general, the picture-quality degradation caused by the MPEG-2 4:2:2P@ML, 18 Mbit / s, IB compression algorithm were more visible than those of Betacam SP. These differences are within the range of the standard deviation and are therefore statistically insignificant.

C.5.5.1.2. Comments for viewing distance 6H

- The same tendency was found in the voting here as for the 4H case described above, but was less pronounced due to the reduced eye sensitivity at a viewing distance of 6H.
- In general, the average picture-quality degradation for Betacam SX within the range of pictures under test was well below the 12.5% mark as the defined limit for visibility.

C.5.5.2. Results obtained for sequences subjected to 4th generation processing

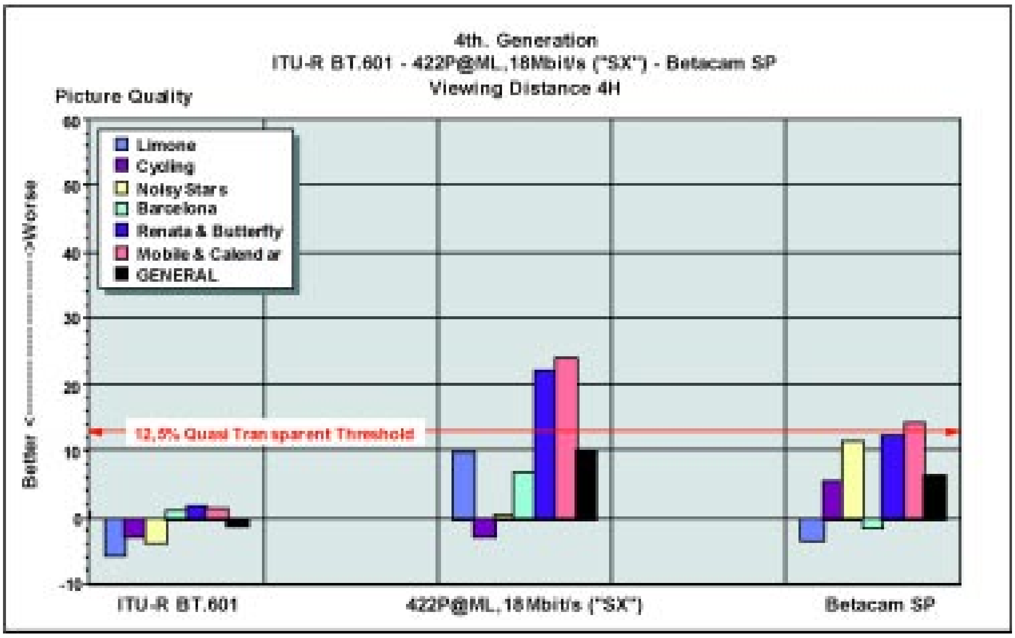
The post-production scenario encompassed four generations of MPEG-2 4:2:2P@ML, 18 Mbit / s, IB processing, two of which involved one temporal shift and one spatial shift each.

C.5.5.2.1. Comments for viewing distance 4H

- In general, the average picture-quality degradation for the range of pictures under test was still below the 12.5% mark as the defined limit for "transparency" for both MPEG-2 4:2:2P@ML, 18 Mbit / s, IB compression and Betacam SP.
- For natural scenes, differences in quality shown in the histogram (*Fig. C.7*) between Betacam SP and MPEG-2 4:2:2P@ML, 18 Mbit / s, IB compression are statistically insignificant.
- Only for the critical test scenes, "Renata-Butterfly" and "Mobile and Calendar", did the MPEG-2 4:2:2P@ML, 18 Mbit / s, IB compression scheme clearly exceed that limit; the picture quality of Betacam SP was judged to be better during these critical test scenes.
- The artefacts produced by MPEG-2 4:2:2P@ML, 18 Mbit / s, IB compression in the natural scenes of the test cycle remained below the threshold of visibility, even at the critical viewing distance.

C.5.5.2.2. Comments for viewing distance 6H

- The absolute ratings for both Betacam SX and Betacam SP are lower than in the 4H case. For natural pictures, differences shown in the histogram between Betacam SX and Betacam SP are statistically insignificant.
- In general, the average picture-quality degradation for Betacam SX within the range of pictures under test was well below the 12.5% mark as the defined limit for visibility.
- Only in the case of "Renata & Butterfly" did Betacam SX exceed the transparency limit of 12.5%.



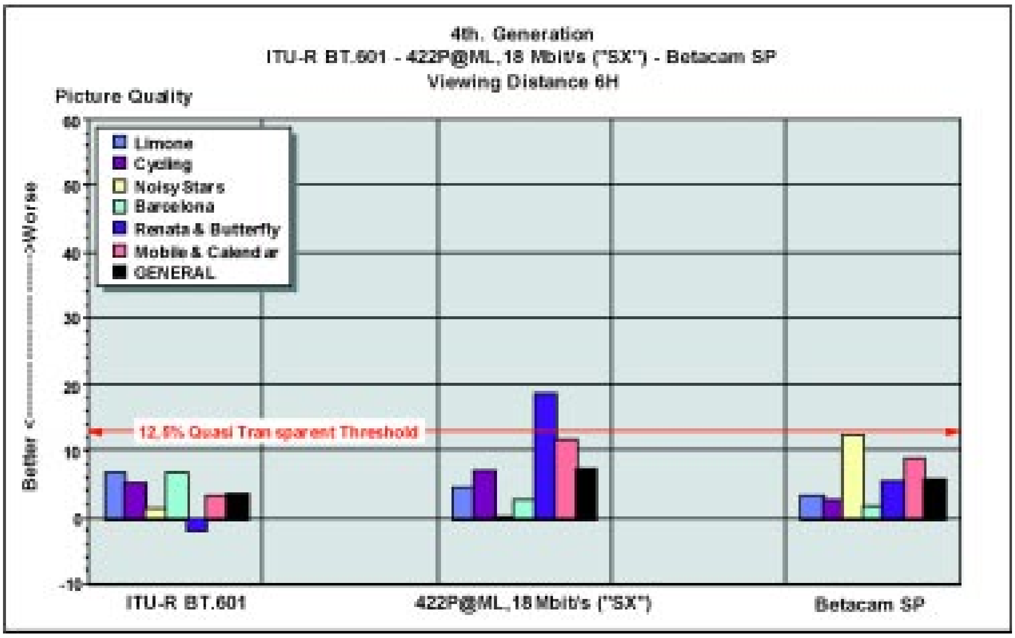


Figure C.7 : 4:2:2P@ML,18 Mbit/s, IB compression scheme – fourth-generation picture quality at viewing distances of 4H and 6H.

C.5.5.3. Results obtained for sequences subjected to 7th generation processing

The post-production scenario encompassed seven generations of MPEG-2 4:2:2P@ML, 18 Mbit/s, IB processing, three of which involved one temporal shift and two spatial shifts each.

C.5.5.3.1. Comments for viewing distance 4H

- The picture degradation produced by MPEG-2 4:2:2P@ML, 18 Mbit/s, IB, in this operating scenario exceeded the threshold of "transparency" for most natural test sequences (see *Fig. C.8*).
- For the critical test sequences "Renata & Butterfly" and "Mobile & Calendar", the limit was exceeded significantly.

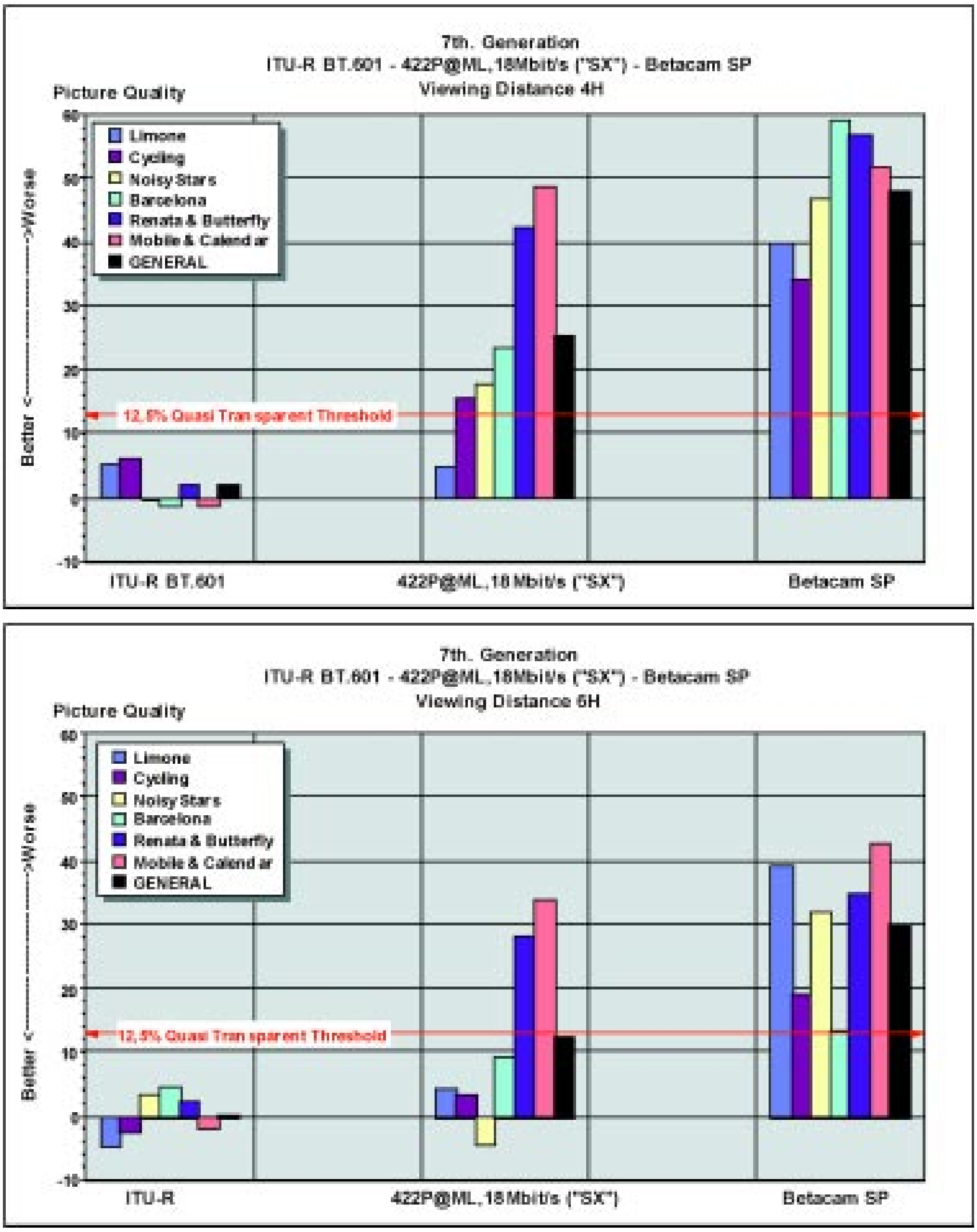


Figure C.8 : 4:2:2P@ML,18 Mbit/s, IB compression scheme – seventh-generation picture quality at viewing distances of 4H and 6H.

- ➪ On average, for both normal and critical pictures, the footprints created by MPEG-2 4:2:2P@ML, 18 Mbit/s, IB compression were rated far below the degradation generated by Betacam SP.
- ➪ Although the threshold of visibility was exceeded in all but one case, the acceptance level of picture quality achieved within this MPEG-2 4:2:2P@ML, 18 Mbit/s, IB post-production scenario will depend on the individual broadcaster's attitude on the acceptance of Betacam SP picture quality in an identical operating scenario.

C.5.5.3.2. Comments for viewing distance 6H

- ➪ The absolute ratings were lower than in the 4H case described above.
- ➪ Analogue Betacam SP was rated markedly worse than Betacam SX for all test pictures.
- ➪ For natural scenes, Betacam SX picture quality was rated to be transparent.

- For critical test pictures, Betacam SX exceeded the limit of transparency considerably.
- Even in the 6H case and for all test sequences, the ratings for Betacam SP exceeded the transparency limit.
- On average, the picture quality of Betacam SX was at the limit of transparency.

C.5.6. MPEG-2 4:2:2P@ML, 50 Mbit/s, intra-frame – picture-quality issues

The histograms (*Fig. C.9*) show the results obtained for viewing distances at 4H (34 observers) and 6H (26 observers) from a subjective test carried out by the RAI and the IRT on an MPEG-2 4:2:2P@ML, 50 Mbit / s, intra-frame compression scheme.

Digital Betacam – as an established high-end compression system – was included in the test in order to assess the picture quality and post-processing ceiling obtainable with non-pre-filtered MPEG-2 4:2:2P@ML, 50 Mbit / s, CBR, intra-frame compression (which is advocated as one option for use in networked mainstream broadcast operations).

C.5.6.1. Results obtained for sequences subjected to 7th Generation post-processing with pixel shift

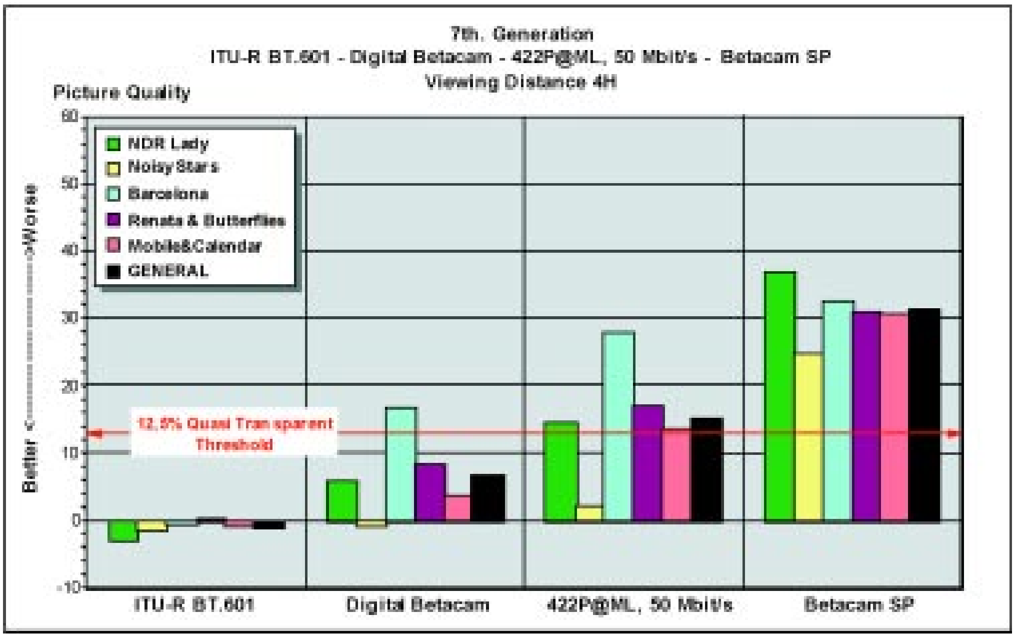
The picture sequences were subjected to 7th generation post-processing with the pixel shift characteristic given in the table below:

Processing	Horizontal Shift (Pixel) +1 = 2 Y pixel shift right ‖-1 = 2 Y pixel shift left	Vertical Shift (Line) +1 = 1 line shift down ‖ -1 = 1 line shift up
1st generation →2nd generation.	no shift	+1
2nd generation →3rd generation.	no shift	+1
3rd generation →4th generation.	no shift	+1
4th generation →5th generation.	+1	no shift
5th generation →6th generation.	no shift	-1
6th generation →7th generation.	-1	-2

Note: The "Diva with Noise" sequence was originally included in the test. This sequence is an extreme test for all compression systems. The "General" result, expressed as numerical values on the histograms (see *Fig. C.9*), represents the average over the sequences tested without inclusion of the "Diva with Noise" test sequence.

C.5.6.1.1. Multi-generation performance of MPEG-2 4:2:2P@ML (50 Mbit/s, I-Frame) compression

- At the normal viewing distance (6H), the picture quality of 1st generation MPEG-2 4:2:2P@ML compression was practically indistinguishable from the 4:2:2 source.
- At the normal viewing distance, experts had difficulty identifying differences between the performance of MPEG-2 4:2:2P@ML compression through all generations for non-critical sequences. No significant decrease of picture quality was observed up to the 7th generation.
- In direct comparison with the source, critical sequences processed by MPEG-2 4:2:2P@ML compression showed some coding noise and a certain loss of resolution in sub-areas containing high picture detail. This effect could be observed with a slight increase through each generation. In general, the level of impairment of the 7th generation does not compromise the picture quality (see *Fig. C.9*).



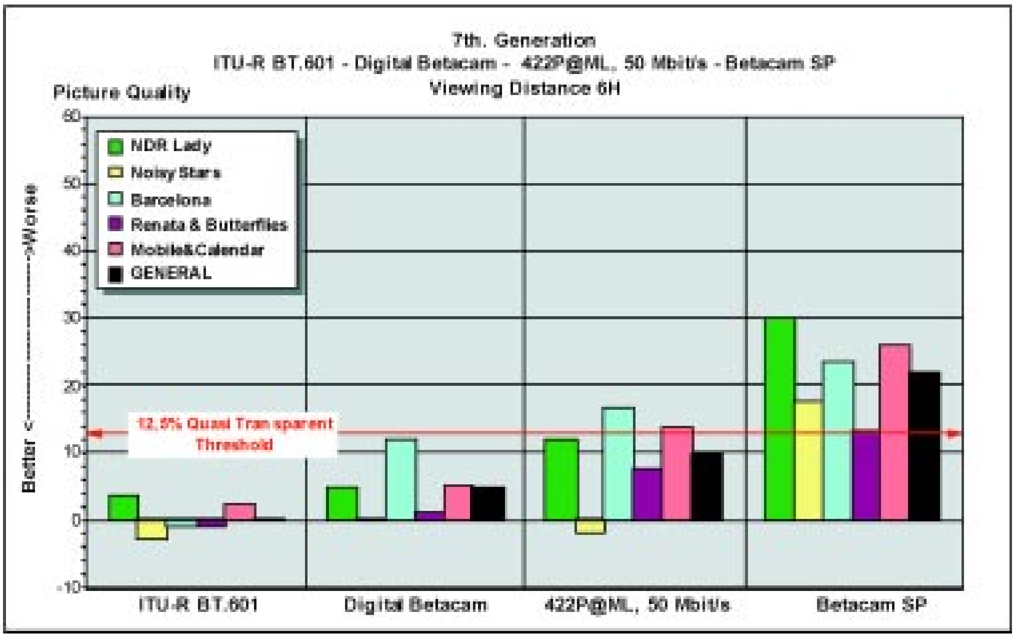


Figure C.9 : 4:2:2P@ML,50 Mbit/s, I-Frame compression scheme – seventh-generation picture quality at viewing distances of 4H and 6H.

C.5.6.1.2. Multi-generation performance of MPEG-2 4:2:2P@ML (50 Mbit/s, I-Frame) compression vs. the compression used in Digital Betacam

- At normal viewing distance and for moderately critical sequences, experts had difficulty in identifying any differences between the performance of the algorithms of the two digital compression systems.
- For the first generation, the performance of MPEG-2 4:2:2P@ML compression and the compression used in Digital Betacam was rated to be identical.
- At fourth generation, the performance of MPEG-2 4:2:2P@ML compression and the compression used in Digital Betacam in the multi-generation scenario is similar. The picture quality of non-critical sequences is practically preserved by both systems. Differences between system performance are only detectable on closer scrutiny. Some loss of resolution and increased coding noise in areas of high picture details were detected in the case of MPEG-2 4:2:2P@ML compression when compared with Digital Betacam.

- At seventh generation, the differences between the systems are marginal. The effects described for the fourth-generation performance are slightly accentuated. For moderately-critical sequences, the level of impairment was very low and did not compromise the overall picture quality. On direct comparison, the picture quality provided by Digital Betacam compression was considered to be slightly better than that achieved with the MPEG-2 4:2:2P@ML compression. This is mainly due to some loss of resolution and a slightly increased coding noise with more critical sequences in the case of the MPEG-2 4:2:2P@ML compression.

C.5.7. Digital interfaces

The table below indicates the type of interface required for a compression family based on MPEG-2 4:2:2P@ML, and shows the status of the various specifications.

Interface:	Status			
	Defined	**In progress**	**Not defined**	**Standard Document**
SDTI	✓			SMPTE 305 M
ATM	✓			AAL 5 / AAL 1
FC		✓		
IEEE-1394			✓	
T-3		✓		
OC-3		✓		
Satellite		✓		

C.5.8. Agile decoders

Throughout the broadcast “programme chain”, a number of different MPEG profiles, levels and bit-rates may be used. These include MPEG-2 4:2:2P@ML for production and contribution, and MPEG-2 MP@ML for distribution and transmission.

MPEG offers several methods for interoperation between bitstreams of different data-rates. Agile decoders working automatically over the 15 - 50 Mbit/s range of MPEG-2 4:2:2P@ML have been implemented and demonstrated by several manufacturers. Furthermore, these agile decoders may also have to operate at MPEG-2 MP@ML. Broadcasters require that source material coded at different rates can be seamlessly edited, mixed and processed without additional intervention.

C.5.8.1. Decoding of different MPEG-2 4:2:2P@ML bitstreams with identical decoding delay at the output

The general feasibility of seamless switching between different MPEG-2 4:2:2P@ML encoded input streams at SDI output level has been publicly demonstrated.

The integrated agile MPEG-2 decoder chip will comply with requirement A in *Section C.3.4.*

C.5.8.2. Intra-family switching between different MPEG-2 4:2:2P@ML bitstreams at the input

The intra-family agile MPEG-2 decoder chip will comply with requirement B in *Section C.3.4.*

C.5.8.3. Intra-family decoding between different MPEG-2 4:2:2P@ML packets within a single bitstream

The agile MPEG-2 decoder chip will comply with requirement C in *Section C.3.4*.

C.5.9. Native Betacam SX bitstream decoders

The native Betacam SX bitstream does not comply with the specification of an MPEG-2 4:2:2P@ML Transport Stream for reasons of optimum stunt-mode operation. A native decoder or a bitstream converter at 4x speed is therefore required (and will be made available) to process the Betacam SX bitstream. A formal specification of the Betacam SX native format has been prepared as a general reference document. This document is not a Standard but a freely-available published specification.

C.5.10. Family relations

C.5.10.1. Tools available for MPEG-2 intra-family transcoding

The following options within both the MPEG-2 4:2:2P@ML and the MPEG-2 MP@ML formats exist:

a) Transcoding between MPEG bit-rates can be achieved by restricting the decoding processes to a minimum and by transferring information extracted by the decoding process to the re-encoder. Information such as the previous quantization levels, motion vector information and GoP information can help the re-encoder more accurately reverse the decoding process at the new bit-rate. By utilizing techniques explored in the ACTS Atlantic Project [19], flexible adjustment of data-rate to different post-production, storage, distribution and transmission requirements by transcoding into different GoP structures can be achieved with reduced quality loss. However, this applies only in the case where the picture Content is not changed between transcoding.

   Complete transparency in the cascading of identical encoding structures can be achieved if the relatively complex step of re-using all the relevant information is taken.

   Transparent cascading in conjunction with intermediate video or audio processing requires the routing of helper information through the processing chain, the maintenance of strict synchronism between macroblocks and helper data, and the provision of dedicated I/O ports for encoders and decoders. Specifications of automatically routed and synchronized helper data (MOLE ™) for both the video and audio have been submitted to the SMPTE for standardization.

b) Sony will provide optimized transcoding between bit-rates to support the flexibility benefits of MPEG-2. The transcoders will not necessarily be working on the same principles as the Atlantic project

C.5.11. Editing flexibility and complexity

At bit-rates in excess of 40 Mbit/s it is likely that I-frame-only structures will be implemented. In the case of lower bit-rates and GoPs of greater than 1, interoperability between equipment of different GoPs may require some additional processing. Assemble-and-insert edits of extended MPEG-2 GoP structures require either decoding of the compressed data steam or a forced P-encoding of an original B picture, depending on the actual positioning of the GoP structure relative to the edit point. This requires a degree of interactivity between the devices involved in editing, e.g. pre-load of data if the edit segment is streamed off an HD server or a confidence replay head, in the case of compressed data replay off and to tape.

19. The work of the “Atlantic” project, led by the BBC and Snell and Wilcox (and described in: Seamless Concatenation of MPEG-2 bitstreams - An introduction to MOLE ™ Technology), is directed towards providing benefits such as:
   - transparent cascading;
   - bit-rate changing;
   - frame-accurate video switching, editing and full post-production;
   - better-than-frame-accurate audio switching and editing.

C.5.12. Examples of some commercial format implementations

C.5.12.1. 422P@ML, 18 Mbit/s, IB tape recording format ("Betacam SX")

Tape:	12.65 mm Metal Particle, 14.5 µ
Cassettes:	Small (up to 60 min); Large (up to 184 min)
Robustness:	very good, adequate for in News & Sports (Notes 1, 2)
Slow-Motion and Shuttle:	good off-tape, very good off HD (Note 3)
Betacam SP replay:	O.K. (Note 4)
Audio channels:	4 channels, 48 kHz, 16 bit
Editing granularity on tape:	Equipment dependent. 1 frame with pre-read heads possible (Note 5)
Cross fade-delay of Audio Ch1 / Ch2:	O.K.
Transparency of VBI:	608 lines recorded , transparent for 1440 samples/field. (Note 6)
SDTI Interface:	Already implemented (Note 7)
Extensibility to a format family:	Inherent in MPEG

Note 1: The replay capability of Betacam SP tapes has been one of the key requirements in the development of the SX format. The use of ½-inch MP tape, together with a moderate data-rate of 44 Mbit/s, allows generous dimensioning of all recording parameters and contributes significantly to the excellent robustness of the whole system. This is confirmed by the error rates measured during the EBU Tests which document the excellent robustness under a variety of different and realistic stress conditions. The actual error-rate performance of the SX Camcorder could not be measured however. The error rate of 10^{-6} measured off-tape with the Hybrid-Recorder provides a testimony of the robust design of the head and tape area.

Note 2: Compressed video signals require elaborate error-correction schemes to guarantee data integrity through noise-infected channels. An overload of the Forward Error Correction system results in the loss of complete macroblocks. Concealment – as the obvious panacea to cope with such situations (by substituting complete erroneous macroblocks by spatially adjacent ones) – will achieve only limited results. SX compression does not allow the substitution of erroneous macroblocks by spatially coincident macroblocks from the preceding frame. This would lead to significant error propagation due to the prediction range embracing several frames. The FEC used in SX, and the use of relaxed recording parameters possible with the available tape area, compensate this however.

Note 3: The audio / video replay quality was good on all settings chosen for Slow-Motion and Shuttle replay. The identification of short audio and video inserts during shuttle is also strongly dependent on the setting of the shuttle speed. Identification is further influenced by scene Content and the length of the inserted segment to be located. Details can be found in the EBU Test Report.

Note 4: When replaying original Betacam SP recordings on SX, a reduction of the video signal / noise ratio of 3 dB was measured, when compared to replay on Betacam SP. Betacam SP FM audio can be replayed from SX.

Note 5: Frame-accurate assemble-and-insert edits of MPEG-2 compressed signals is strongly dependent on the relative position of the GoP structures within the data streams to be edited. Details can be found in the EBU Test Report. For edits requiring access to individual pixel elements (wipes, re-sizing, amplitude adjustments), the signals have to be decoded.

Note 6: With the exception of 1440 samples, input data in this area will not be conveyed transparently through the SX recording and replay channel. The mutual influence of VBI data, and data in the active picture area, can therefore not be excluded. For details, see the EBU Test Report. Cross-talk between data in the active video area and the VBI lines subjected to compression encoding has been measured.

Note 7: The SDTI interface for the transport of the audio, compressed video and Metadata stream in real-time and non-real-time has recently passed the final ballot at the SMPTE; a practical implementation into SX equipment is therefore expected imminently.

C.5.13. NLE equipment

MPEG-2 4:2:2P@ML and MPEG-2 MP@ML non-linear editing (NLE) equipment is available.

C.5.14. Format development potential

MPEG-2 compression operates over a wide range of profiles, levels, GoPs and data-rates. This provides the user with the flexibility to select the combination of picture quality, operational flexibility and economy which is relevant to the specific application. A wide range of applications can be addressed within the toolbox provided by MPEG-2, ranging from distribution at lower bit-rates using MPEG-2 MP@ML, to production and post-production using MPEG-2 4:2:2P@ML at higher data-rates, up to HDTV applications using MPEG-2 MP@HL and MPEG-2 4:2:2P@HL.

Note: Constant quality can be achieved if the application supports variable data-rate.

C.5.15. Test equipment

Available on the market.

C.6. Supplement A: EBU Statement D79-1996 – Open standards for interfaces for compressed television signals

EBU Committee: PMC. First Issued: November 1996.

The proposals to use compressed signals on a number of new television recording formats have raised a number of questions about interfaces.

There are a number of advantages in interconnecting equipment associated with these formats using interfaces which carry the compressed signals.

These advantages include:

- the avoidance of multiple coding and decoding;
- cost effective storage on disk and tape;
- the possibility of non-real-time transfer, particularly at faster than real-time.

However, to best exploit these advantages, broadcasters should be able to interconnect equipment from a variety of manufacturers.

Therefore, **the EBU requires that**:

- a single interface should be defined to carry compressed television signals;
- all elements of the interface should be open and fully specified.

C.7. Supplement B: EBU Statement D80-1996 – Compression in television programme production

EBU Committee: PMC. First Issued: November 1996.

At the present time, broadcasters are faced with a choice between incompatible compression algorithms used on different non-linear editing and acquisition devices. Systems based on the new tape recording formats SX and DVCPRO operate on compression algorithms at 18 and 25 Mbit/s respectively and are intended to be used in the acquisition of Sports and News material. New tape recording formats for mainstream television applications have already been announced. One is based on an extension of the DVCPRO compression algorithm to the 4:2:2 signal format and will operate at about 50 Mbit/s. Other formats based on the 4:2:2profile@ML of MPEG may follow.

It is possible to integrate devices using compression systems into existing digital facilities if they are equipped with the standard serial digital component interfaces in accordance with ITU-R Recommendation BT.656. However, the compressed signals must first be decoded into ITU-R Recommendation BT.601 format.

The following consequences also arise:

- ➩ any further re-encoding and decoding of the previously compressed signal, such as may be required for further non-linear editing, will further increase the loss of signal quality;
- ➩ Even for simple assemble editing, programme segments encoded with different compression algorithms would each need to be decoded into BT.601 format. Subsequently, a decision may have to be made on which format is used for the edited programme material for future storage on a server or in the archive.
- ➩ The cost and operational benefits of an integrated tape and disk strategy using a single algorithm would be nullified by the time required to transfer the programme material between different media. This is because there is little possibility of faster-than-real-time transfer between the acquisition, processing and storage devices when using signals in ITU-R BT.601 form.

The provision of a single interface standard to carry compressed signals would alleviate this situation but the interface signal formats based on existing algorithms would not be compatible with each other or with other MPEG-based standards. Unfortunately, the EBU sees little likelihood of achieving harmonization at bit-rates in the range 18-25 Mbit/s.

The situation is different for compression algorithms operating at higher bit-rates, which may possibly be used in mainstream television studio operations. No significant amount of equipment is installed in this area of activity and hence the possibility still exists for achieving harmonization.

The EBU is encouraged by the continued improvements in performance and cost of disk storage and considers that:

- ➩ there are real economic benefits to be achieved through the use of a single compression algorithm and file format for programme exchange;
- ➩ intermediate storage and long term archival of material in a variety of formats is inefficient and creates problems extending into the future;
- ➩ disk-based editing produces time and cost benefits over tape-based systems;
- ➩ there are technical and system benefits for programme production through an ability to select equipment from different suppliers as appropriate for different applications;
- ➩ that compression algorithms operating in an I-frame-only format at about 50 Mbit/s have been demonstrated and they are likely to offer a picture quality and a headroom for post-processing which are appropriate for all but the most-demanding studio operations.

The EBU firmly believes that:

- ➩ for high-end programme production, uncompressed signals according to ITU-R Recommendation BT.601 or systems using lossless compression or systems using lossy DCT-based compression with a compression factor not exceeding 2 should be used;
- ➩ for mainstream programme production and for programme acquisition using low bit-rate compression formats where the operational advantages of compression are obvious, only a single, open compression algorithm should be applied for storage or file transfer applications.

Furthermore, this system should be operating at 50 Mbit/s and should use an I-frame-only format.

Annex D

Wrappers and Metadata

D.1. First Request for Technology

D.1.1. Introduction

This section of the RFT covers the format of Metadata and the Wrappers used to transport or store this data together with any supported Content items. In general, these formats will be split into stream formats and file formats. A file is defined here as analogous to a computer file and is normally dealt with as a unit, whereas a stream is defined as a continuous flow of data with the capability for receivers to join at any time during a streaming operation (within some published limits). The purpose of these formats is to encourage the maximum practical interoperability between diverse systems.

Background information to this RFT is to be found in Section 2 and Annex D of the EBU / SMPTE document "Joint Task Force for Harmonized Standards for the Transfer of Programme Material as Bitstreams, First Report: User Requirements" (April 1997).

D.1.2. Submission details

D.1.2.1. Description

A description of the offering is required, including available materials, supporting documentation and any commercially-available implementations.

D.1.2.2. RFT table

The table on the next page should be completed with marks made to indicate the areas of compliance. This table will be used to categorize submissions and to check broad compliance with the objectives of the RFT.

Detailed information can be added by using separate pages. Reference to this additional information can be made by indicating "note 1", "note 2" etc. in the table.

The table headings are defined as follows:

- **Offered** indicates whether the response covers this aspect of the requirement;
- **Defined** indicates that the response covers a defined way of dealing with this aspect of the requirement;
- **Universal** indicates that the response covers this aspect for all application areas in its current form;
- **Extensible** indicates that the offering can be extended to cover this aspect of the requirement;
- **Timeframe** indicates the timeframe in which the response will meet this requirement.

D.1.2.3. Footnotes to the table

Please attach additional supporting information as needed (use "note: x" as a reference from the table to the additional information).

D.1.2.4. Supporting materials, APIs and protocol documents

Please attach any additional supporting material as separate documents.

Ref. (note)	RFT Topic (note)	Offered	Defined	Universal	Extensible	Time frame
1	Does the format support streaming (ref. 2.4)?					
2	Does the format support interleaving (ref. 2.8)?					
3	Does the format support file storage (ref. 2.4)?					
4	If 1 and 3are answered with a "Y", are the file and stream formats interoperable?					
5	Does the system support a unique identifier (ref. 2.9)?					
6	Is a registration procedure for Metadata types in place?					
7	Can Essence types be added to the format (ref. 2.19)?					
8	Can Metadata types be added?					
9	Does the system have a flexible class hierarchy?					
10	Are user-defined hierarchies supported?					
11	Are different Metadata Sets (templates) supported?					
12	Does the system support the Content structure described in ref. 2.2.1.?					
13	Does the system support multiple character sets i.e. Unicode, ISO extended language support?					
14	Is a machine-readable plain text format supported?					
15	Can files be partitioned into smaller files (ref. 2.6)?					
16	Are multiple hardware / software operating platforms supported?					
17	Are different byte orderings supported (ref. 2.7)?					
18	Are multiple generations or versions of Metadata supported (ref. 2.10)?					
19	Can Metadata be referred to by pointer (indirection) (ref. 2.11)?					
20	Are methods of systematic indexing supported (ref. 2.12)?					
21	Are methods of specific indexing supported (ref. 2.12)?					
22	Is derivation history supported (ref. 2.13)?					
23	Are security features implemented (ref. 2.14)?					
24	Is transaction logging supported within the format (ref. 2.15)?					
25	Are property rights supported within the format (ref. 2.16)?					
26	Is a material management system available for this format (ref. 2.17)?					
27	Is an API available to the format (ref. 2.18)?					

Note: These references can be found in the First Report of the Task Force on User Requirements, April 1997.

D.2. Second Request for Technology

D.2.1. Introduction

In the analysis of responses to the first Request for Technology, it was realized that several of the responses offered complete systems whose capabilities overlapped. Although reconciliation of these systems into a single system might be achieved, the resulting system would be of such wide scope that the process of standardization and implementation of the resulting standard would take a very long time. This does not meet the TFHS goal of selecting a file format which can be standardized and deployed in the near future.

It was also clear that a common element of all responses was the use of a general-purpose persistent storage mechanism at a relatively low level of complexity. In all cases, the data model of the overall system was mapped onto such a storage subsystem.

Therefore, the Sub-Group determined to proceed with two tasks in parallel:

- to solicit specific solutions to the storage mechanism, and either choose a single solution or create an independent public domain solution;
- to specify an overall data model, to which any commercial or pre-existing system could in principle be mapped either directly or by the use of translation software.

This Second Request for Technology addresses the first of these tasks.

D.2.2. Licensing issues

Considering the importance of this element of technology, the TFHS is specifically requesting full documentation of the storage mechanism and conformance to the following guidelines:

To be standardized, any technologies must be open in three respects:

- existing implementations must be licensable;
- technologies must be documented sufficiently to permit new implementations from the ground up, without fees;
- a due-process standards body must have control of the specification in the future;

The TFHS will also recommend compliance testing by an external organization.

D.2.3. Evaluation of responses

Responses were evaluated during the meeting in Atlanta over the period March 5th~ 7th, 1998.

If no single response met all requirements, the Sub-Group proposed to create a universal format in the public domain under the direction of a standards body, probably the SMPTE.

However, one of the responses addressed all the requirements. The documentation was studied and forwarded to the SMPTE as the proposed universal standard.

D.2.4. Documentation requirements

- A full description of the bitstream generated by the storage mechanism. The description must indicate through diagrams and tables how each data construct appears in the functional requirements, when mapped onto a sequence of bytes.
- An item-by-item response to each functional requirement listed below, as a simple Yes or No plus a single paragraph of explanation.
- A statement on licensing issues, covering the points listed above, and listing any fees charged for the use of existing implementations, and indicating if any patents, trademarks or intellectual property rights exist and the proposal for freeing the technology from them.
- Optionally, additional documentation may be provided in separate documents. This may include reference specifications, or descriptions of additional features of the technology. However, this documentation must not be required to be read in order to understand and evaluate the response.

D.2.5. Functional requirements

- ➩ The storage mechanism shall provide a method to uniquely identify the file format within the file itself. The identification shall be independent of the byte orderings of the file and of the platform. The identification is necessary to distinguish the bitstream from any other bitstream. The identification data shall be an SMPTE Administered Universal Label in accordance with SMPTE 298M-1997.
- ➩ The storage mechanism shall provide a method to define objects. The objects shall consist of a set of properties. Each property shall have a property name, property type name and property value.
- ➩ The storage mechanism shall not constrain the number of objects or properties within the file, and it shall not constrain the number of properties in any object.
- ➩ The storage mechanism shall provide a method to define a property name. The storage mechanism shall allow a property name to be an SMPTE Administered Universal Label in accordance with SMPTE 298M-1997.
- ➩ The storage mechanism shall provide a method to define a property type name. The storage mechanism shall allow a property type name to be an SMPTE Administered Universal Label in accordance with SMPTE 298M-1997.
- ➩ The storage mechanism shall provide a method to read and write a property value. The storage mechanism shall allow a property value to be a sequence of bytes. The storage mechanism shall not restrict the value of any byte in the sequence and shall not restrict the length of property values to a length less than 2^{64}-1 bytes.
- ➩ The storage mechanism shall provide a method to read and write a property value with a byte ordering the same as or different from the default byte ordering in the file, without requiring that the file or the property value be reordered before or after the operation.
- ➩ The storage mechanism shall provide a method to read and write a property value with a byte ordering the same as or different from the native byte ordering of the platform, without requiring that the file or the property value be reordered before or after the operation.
- ➩ The storage mechanism shall provide a method to access a header object, which can also be used as an index to access the objects in the file.
- ➩ The storage mechanism shall provide a method to access each object in a file, in any order or sequence.
- ➩ The storage mechanism shall provide a method to specify that a property value is a reference to an object in the same file and shall provide a method to access an object in the file by using such a property value.
- ➩ The storage mechanism shall provide a method to access all of the properties of an object.
- ➩ The storage mechanism shall provide a method to access a property of an object by specifying the property name and the property type name.
- ➩ The storage mechanism shall provide the following methods to access a property:
  - a method to get the property name of a property;
  - a method to get the property type name of a property;
  - a method to get the length of the property value in bytes;
  - a method to determine if a property exists in an object;
  - a method to add a property to an object by specifying the property name, the property type name and the property value;
  - a method to read or write the entire property value;
  - a method to read or write a portion of the property value.
- ➩ The storage mechanism shall allow an object to have a property with a property name that is the same as the property name of a property on a different object, and to have a property type name that is different from the property on the other object.
- ➩ The storage mechanism may optionally provide a method to control the relative placement of the various data constructs within the file.

D.3. Responses to the Requests for Technology

The documents referenced below were accurate at the time of publication and represent a snapshot of that time. Many items below are subject to further development so the reader is encouraged to locate more recent documents where available. In many instances these documents will eventually be superseded by published standards. These documents are grouped into Storage, Streaming and Metadata responses. The sequence order is not significant.

D.3.1. Open Media Framework Interchange

The Open Media Framework Interchange (OMFI) Specification is a data model for representation of Essence and Metadata in Complex Content Packages. It is intended to be used in conjunction with a low-level storage mechanism for managing the actual storage of information in the file.

It is recommended that the OMFI data model is translated to use standardized SMPTE Labels and to operate together with an SMPTE-standardized low-level storage mechanism, and is then extended as new requirements for Complex Content Packages are documented. The proposers have committed to implement these changes and have produced a new specification called the Advanced Authoring Format (AAF).

A proposed standard to document the format is under consideration by SMPTE P18.27.

D.3.2. Structured Storage

The Microsoft Structured Storage file format has been submitted as a response to the EBU / SMPTE Joint Task Force's Metadata and Wrappers RFT for a low-level storage mechanism, and has been redrafted as a proposed standard and submitted to SMPTE P18.27. It is also included within the specification of the Advanced Authoring Format (AAF).

Structured storage is a portable and scalable interchange file format, designed to store a complex hierarchy of Essence and Metadata, and can be viewed as a file system within a single file.

D.3.3. Advanced Streaming Format

Microsoft Advanced Streaming Format (ASF) is an extensible presentation file format (i.e. a "Wrapper") designed to store synchronized multimedia data. Although ASF can be edited, it is not an editing format per se, but rather is designed to work as an output streaming format, potentially created from the existing editing formats (such as OMFI or AVI / Wave). ASF was explicitly designed to address the media streaming needs of the Internet community. It supports data delivery over a wide variety of networks and protocols while still proving to be well suited for local playback.

A proposed standard will de drafted and submitted to SMPTE P18 for consideration.

D.3.4. QuickTime®

Apple Computer's QuickTime® software includes a file format which serves as a Wrapper for multimedia Content throughout its lifecycle. The file format is supported by a large number of tools, primarily but not exclusively through the cross-platform multimedia layer of the QuickTime® software from Apple Computer Inc. The file format was designed to be a flexible Wrapper for a wide variety of Content formats (e.g. Video, Audio, MIDI etc.), and multiplex formats (e.g. AVI, WAV etc.). The design of the format is adaptable to various streaming protocols, and for various stages in the Content lifecycle.

The design of the "Intermedia format" in the draft MPEG-4 Systems (ISO 14496-1) Version 2 is derived from the QuickTime® file format with additions and changes to accommodate the specific requirements of MPEG-4.

D.3.5. "Strawman" proposal for the structure of a Wrapper

A document was received with proposals for an Interchange Wrapper file which included a number of issues discussed in the main part of this report. The topics discussed included a proposal for a UMID, a proposal for a Proxy Wrapper interchange structure, a proposal for a Content Element and a list of suitable file structures for the basis of interchange.

The sections on a Proxy Wrapper interchange structure and the structure of a Content Item element are available as the basis of an SMPTE Standard in the P18 group. The other topics are already in progress.

D.3.6. SDTI Content package

A proposed format for the transmission of elementary streams over an SDTI (SMPTE 305M) connection has been submitted to the SMPTE for preparation as a standard using Content Packages formatted as a stream. SDTI offers a synchronous stream transfer carrier and this is applied in the proposal by providing specifications for picture-bounded Content Elements.

As of the publication date of this report (September 1998), this proposed standard is being balloted by the SMPTE Packetized Television Technology Committee.

D.3.7. Betacam SX® format over SDTI

A document has been made available detailing the format used by Betacam SX® for the transfer of coded bitstreams as a Content Package over SDTI. This document is not standardized, but is freely available to all parties on request to Sony.

D.3.8. MPEG-2 Transport Stream over SDTI

A document is in the process of standardization through the SMPTE to provide for the transfer of MPEG-2 Transport Stream packets over SDTI. MPEG-2 Transport Streams are a Content Package formatted in a Streaming Wrapper.

D.3.9. Fibre Channel AV Simple Container

The Fibre Channel AV Simple Container provides a mechanism for multiplexing Content Elements of various types onto a Fibre Channel interconnect. It is proposed for standardization by NCITS T11.

It is broadly interoperable with similar proposals targeted on SDTI.

D.3.10. Content Metadata Specification Language

This method of documenting templates for the description of Content Elements, Items and Packages was proposed. It has also been considered by the Digital Audio-Video Council (DAVIC) and is adopted as part of their V1.3 Specification.

D.3.11. Frame Tables

The use of Frame Tables was proposed, to provide indexing into bitstreams where each Content Component is of varying length. This general approach was echoed within several other responses in various forms.

It is expected that Frame Tables will form part of the standardized data model for Complex Content Packages, and will probably become an item of Metadata in their own right.

D.4. Related Essence format documents

D.4.1. DV-based formats over SDTI

A document for the transfer of DV-based formats (e.g. DVCPRO and DVCPRO50) is now available as an SMPTE standard.

D.4.2. DV-based formats over Fibre Channel

A similar document is under consideration within NCITS T11.

D.4.3. Standard DV over SDTI

A document is in the process of standardization through the SMPTE to provide for the transfer of IEC 31864 DV bitstreams over SDTI. It includes extra control data over the IEC standard, to allow finer control of dataflow which may be removed to conform perfectly to the IEC standard.

D.4.4. Audio

Several standards and proposals exist for the formatting of Audio Essence, including:

- EBU Broadcast Wave Format, Tech. 3285;
- SMPTE 302M Formatting of Linear PCM Audio as MPEG-2 Private Data Packets;
- SMPTE A12.40 Audio Access Units for Use in MPEG-2 Systems.

Annex E

Networks and Transfer Protocols – Reference Architecture for Content Transfer and Streaming

E.1. Introduction

The transfer and sharing of files containing Content is a fundamental and important issue. The computer industry has developed several standards which specify methods for performing file transfer and sharing. However, there are problems and requirements unique to the video production industry which are not fully addressed by the existing computer industry standards. To overcome these limitations – and in order to guarantee interoperability in the interface, network and protocol domains for file transfer and streaming – the following rules and guidelines in this Reference Architecture for Content Transfer and Streaming are recommended.

In *Section E2* of this annex, methods for adapting the current computer industry standards, as well as proposals for a new standard, are presented. The use of the ubiquitous File Transfer Protocol (FTP) is described as it relates to transferring large Content files. The use of a Network File System (NFS) is described as it relates to shared file access. Examples of how these existing protocols can be mapped onto various computer network technologies are given in the framework of the “Reference Architecture”.

Some of the important requirements that cannot be met using existing protocols include:

- ⇨ performing transfers simultaneously to multiple destinations (point-to-multipoint transfers);
- ⇨ locating files that are managed by an Asset Management System (i.e. not within a simple hierarchical file system);
- ⇨ the transfer of Data Essence or Metadata separately from the associated Video and Audio Content;
- ⇨ the transfer of parts of a file;
- ⇨ the setting of a maximum transfer rate or QoS for a transfer;
- ⇨ the use of network interfaces which may not support TCP / IP (for example, Fibre Channel).

A new file transfer protocol, FTP+, is defined which provides a number of additional functions. FTP+ builds on FTP and uses the same base set of commands. FTP+ includes new commands that enable these additional features and also provides the ability to embrace new network protocols as they become available. *Section E2.3* of this annex presents an overview of the existing FTP protocols as well as details of the new FTP+ command set. FTP+ is under standardization by the SMPTE so that these enhanced file transfer capabilities will become widely available.

As important as transferring and sharing Content files is the ability for real-time “streaming” of Content across various network interfaces. Streaming refers to the transmission of programme material from a source to one or more destinations such that the material is “played”, either at the original frame-rate or at a faster or slower rate (trick modes). In contrast to file transfer there is no guarantee that all receivers will receive every frame of data. Streaming can be performed on a variety of network interfaces, each of which is appropriate for different bit-rates and Essence formats.

Section E3 of this annex presents methods for streaming which use SDTI, ATM, IP and Fibre Channel networks.

In addition to the underlying network interface, a particular stream will also employ a container format to encapsulate the Content. Examples of container formats include the MPEG Transport Streams and DIF streams. This section also provides detailed charts and information showing how the various combinations of networks, containers and compression types can be utilized to perform streaming of audio and video files over the proposed interfaces and networks.

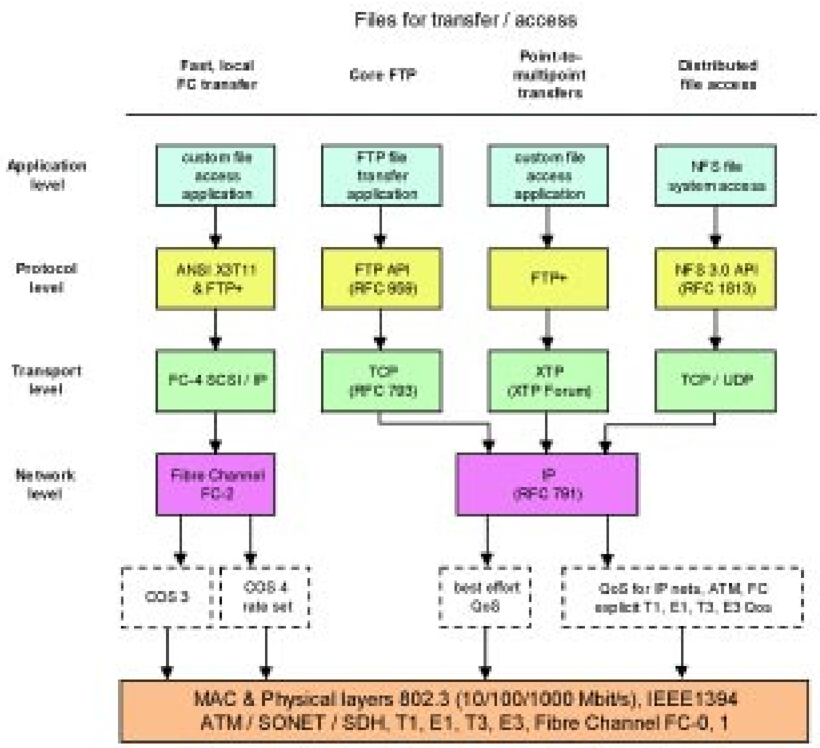


Figure E.1: Protocol Stacks.

E.2. File transfer and access protocols

E.2.1. Scope

This document outlines the general methods of Content transfers by means of file transfer protocols using a Reference Architecture (RA) model. A File System Access method is also considered. Details of the file format are not within the scope of this document.

E.2.2. The protocol stack view of the RA for file transfer and file access

Fig. E.1 shows four protocol stacks. Reading from left to right, the first stack shows the use of Fibre Channel for performing “Fast, local file transfers”. The second stack shows the use of Universal FTP to perform “normal” transfers over IP networks. The third stack shows the use of XTP to perform point-to-multipoint transfers. The fourth stack shows the use of NFS to perform distributed file access.

E.2.2.1. File transfer protocols

It is the intention of the RA model to consolidate the three file transfer methods into a single protocol (at the protocol layer). Why do this? It is preferable to have one common protocol that can work across any of the methods. For example a file may be moved from server to server using the FC links or from server to server across a WAN with the same protocol being usable for either. It should be stated that the protocol will be profiled into sets and each set will be used as needed. For example, some of the actual protocol commands used for an FC transfer are not needed for a standard FTP transfer. Also, some of the enhanced commands for FTP+ are not used by standard FTP and so on. The important aspect of this consolidation (i.e. the commands specified by the FTP and FTP+ protocols) is that most of the protocol commands are common for all three transfer methods.

Note: The stacking shown in *Fig. E.1*, as it relates to file transfer, is a simplified view. File transfer based on an FTP model uses two stacks; one for the actual file data exchange and one for control data ("set up" and "tear down" of the transfer session). For all three methods, the control stack is based on TCP/IP. No actual file data travels over the control connection. For the Fibre Channel method, the data channel will be defined by the NCITS T11 group in consultation with the SMPTE. The SMPTE will define the complete stacking solution for FTP+. All of the new commands needed for FTP+ will be sent over the control path and not the data path. However, for some new features (such as partial file read) and some FC features, the data path R / W protocol is modified. (The IETF's RFC 959 contains a good explanation of the dual stack method used for FTP.) Also, this document does not specify that both the control and the data paths be implemented on the same physical connection.

The network-level interfaces appropriate for file transfer include Fibre Channel, XTP and IP. Fibre Channel specifies both the network layer and the physical layer. XTP can operate in a "raw" mode in which it encompasses both the network and physical layers, or it can operate on top of IP. XTP in raw mode achieves some efficiency and has the possibility of utilizing features of the underlying physical media (such as QoS for ATM) that is not possible when XTP is used on top of IP.

E.2.2.2. Distributed file systems

It is recommended that vendors should provide a facility for sharing files in a network environment. This capability is generally known as a distributed file system. An example of this is the Network File System as defined in RFC 1813. File system access is different from file transfer.

Network file systems generally employ a client-server model in which the server computer actually has the file system as local data. The client host is allowed to "mount" the network file system to get access to the directories and files as if they were locally available. Multiple clients are permitted to simultaneously "mount" the server's file system and get access to its Content. The server may grant different sets of access privileges to a set of users via export control.

Note: Network file systems do not support the streaming of files yet; this capability may be necessary in a professional broadcast environment and should be the subject of future work on distributed file systems.

With the advent of file servers and remote drives, file system access is a commonly-available technology. File system access offers features not available with File Transfer Protocols such as FTP. The following functions need to be provided:

- full file read / write access to all or parts of a shared file;
- the ability to perform any operation on a mounted remote file system that may be performed on a local file system;
- file system navigation and maintenance;
- complete control over file permissions.

E.2.3. File transfer protocols

E.2.3.1. FTP

FTP (file transfer protocol) is chosen as the baseline file transfer protocol which must be supported by all manufacturers to guarantee interoperability. The FTP protocol stack is defined as three levels.

- A GUI (Graphical User Interface), an API and a command line interface offer FTP commands to PUT and GET files [20]. The definition of the GUI, API or the command line interface is a function of specific vendor choice.
- The FTP client-server protocol is standardized by RFC 959 and the Host Requirements by RFC 1123.
- The delivery of the file Content is guaranteed by the TCP protocol for FTP. RFC 1323 (TCP tuning parameters) is also useful when files must be transferred over long distances at high speeds. Implementation of RFC 1323 is optional.

20. A common misconception is that PUT, GET and associated user-level semantics are defined by a standard. They are not. Only the client-server interaction is defined by a standard: RFC 959.

E.2.3.1.1. Recommendations for the chunking of files

The 64-bit FTP is recommended. However, to achieve interoperability, 32-bit files must also be supported. In the 32-bit case, chunking [21] of files must be used if the file size exceeds 2 Gbyte, in accordance with the following rules:

- ➪ If either the local or remote host is limited to 32-bit file lengths, then files exceeding 2 GB must be chunked into two or more smaller files. The smaller files should be named with extensions which indicate their position in the chunk sequence. Files should be sent in the same order as the original file's Content (the first byte of the large file is the first byte of chunk #1 and so on).
- ➪ Chunked files should be named as follows (where "name" is the name of the original large file):
  - <name.c00> is chunk file #1
  - <name.c01> is chunk file #2
  - <name.cxx> are the succeeding chunks with the exception that <name.c99> is always the last chunk.
- ➪ The file <name.c99> is always sent last and its presence indicates that all the chunks have been received.
- ➪ It is the responsibility of the receiving system to de-chunk the files into a larger one if needed.

E.2.3.2. Enhanced FTP – "FTP+"

This section outlines the philosophy and goals for designing enhancements to FTP and allowing for different transport and network layers for data transfer. As described earlier in this document, the requirements for file transfer cannot be met purely by means of standard FTP. Also, standard FTP assumes that the data transport layer is TCP. To meet our needs, we must allow for FTP+ to transfer data across non-TCP data channels as well. In many ways, FTP+ is a super set of FTP.

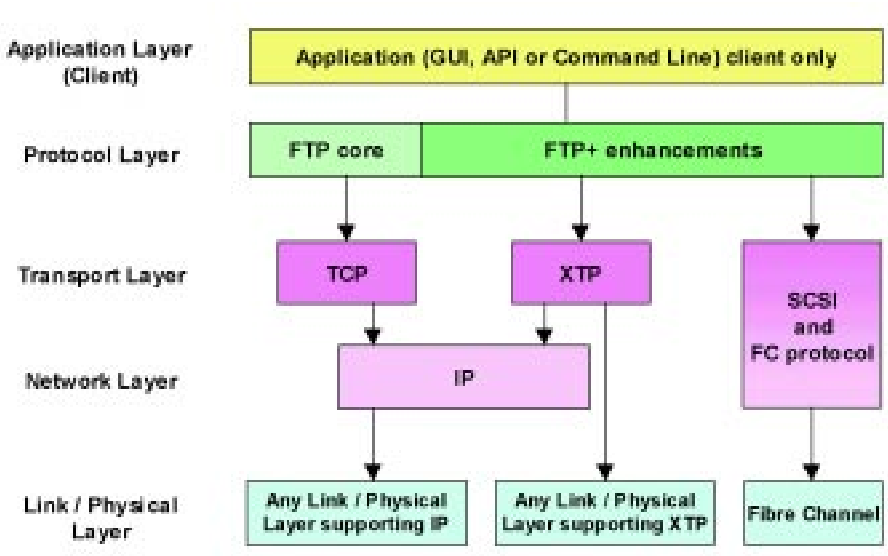


Figure E.2: Protocol, Network and Transport layers for FTP / FTP+ Client or Server (file-data path only shown).

FTP+ is based as much as possible on the FTP RFC 959 specification. Understanding RFC 959 is imperative for a proper grasp of the changes that this document proposes. In its basic form, FTP uses a data path for the actual transfer of data and a control connection for the set up, tear down and status of the data transfers. Both the normal Client / Server model and the less-used Proxy Client model are supported.

To understand the changes that are being made to FTP, it is important to recognize the environment in which FTP+ exists. *Fig. E.2* shows a model of how FTP+ will exist in relation to the layers below it. Also, the control

21. The process of "chunking" converts a large file into smaller ones.

path (which is not shown in *Fig. E.2*) between the local and remote machines (or from the proxy client to each server process) will always be based on TCP regardless of the transport choice for the data path.

It should be noted that an FTP server may be included with the operating system which is installed on most devices. This FTP server may coexist with FTP+, and methods of accessing FTP+ are outlined in this annex.

In *Fig. E.2*, the protocol layer may access either a Fibre Channel, TCP or XTP link for a given file-data transfer session. It should be noted that the FTP+ enhancements are specific to the choice of transport layer used. Ideally, one common set of enhancements would suffice for use with any of the three different transport choices. Unfortunately, this makes the actual use of the transfer protocol ambiguous and difficult to use. (By way of an example, some of the enhancements required to achieve full functionality with XTP are not compatible for use with Fibre Channel, and vice versa.)

To meet our needs, therefore, we propose a family of four basic profile sets of FTP+ enhancements. The sets are designed with particular transport layer types in mind and provide features related to the respective type. These are described below.

E.2.3.2.1. FTP+ profile sets

Table E.1 shows the proposed profile sets for the FTP+ family of protocols. Each set is named, based on its usage profile. The protocol commands defined in standard FTP are called **FTP Core** for this discussion. Standard 100% FTP must be supported on any device that offers file transfer services. Support for FTP+ is optional. The FTP+ members are classed in the following ways. Remember, the FTP and FTP+ application programs may coexist in harmony on the same machine. *Section E.2.6.* outlines the usage of the new commands.

⇨ **FTP (Standard Profile)**: 100% FTP as defined by RFC 959 and RFC 1123. FTP+ includes all the commands and functionality that exist in 100% FTP with the exception that the control and data ports are not defined to be on ports 21 and 20 respectively.

Table E.1: Profiles of the FTP+ family

General Profile	XTP Profile related commands	Fibre Channel profile related commands	Optional Profile commands
STAT (new return values defined)	RATE (new)	"To be defined" (new)	RETR (new use of existing parameter)
XPRT (new command)	MCGM (new)		STOR (new use of existing parameter)
XPAS (new command)	MCPV (new)		STAT (new use of existing parameter)
SITE (new return values defined)			LIST (new use of existing parameter)
RETR (added optional partial file start, length parameters)			NLST (new use of existing parameter)
STOR (added optional partial file start, length parameters)			

Notes to Table E.1:

1. Some commands have newly-defined return values. For example, the STAT command now returns additional values for statistics of the transfer in progress.
2. The STOR, RETR commands have two new features. Firstly, there is an optional set of parameters for partial file transfer read and write. The start byte offset and the length of the desired data are now allowed. When a file is read in partial fashion, the receiving end will create a new file with its Content being the desired partial data. When an existing file is modified by a write, partial file Content is either over-written by the new data or appended to the end of the file, depending on the address of the written data. Secondly, the <pathname> parameter has been expanded to allow for location of Content via means other than an actual pathname.
3. There are new commands to support the multipoint feature of XTP. There is also a new command to support the maximum rate setting available with XTP.
4. There are new commands to support features inherent to Fibre Channel.
5. At this time all of the commands in the Optional profile are backwards compatible, modified, versions of existing commands and are used to support Content-finding methods.
6. XPRT, XPAS, MCPT, MCPV are new commands related to the existing PORT and PASV commands.

- **FTP+ (General profile)**: FTP Core plus general enhancement commands for use over TCP or XTP. When used with XTP the multipoint and maximum rate functions that are available with XTP are not available for *this* profile. XTP offers speed advantages even when used as a direct TCP replacement.
- **FTP+ (XTP profile)**: General profile plus specifically-related XTP commands. This profile supports all XTP features including multipoint transfers and transfer maximum rate setting.
- **FTP+ (FC profile)**: General Profile plus specifically-related Fibre Channel commands. This includes the yet-to-be-standardized NCITS T11 FC-based file-data-related transfer protocol. TCP / IP is not used but rather SCSI / FCP will be used to guarantee reliability.
- **FTP+ (Optional profile)**: These commands are independent of the transport layer choice and may be added to the General, XTP or FC profile sets. These new commands are used to support Content-finding methods.

The exact formats for the new and modified commands are given in *Section E.2.5.* of this annex.

E.2.3.2.2. Use of the profiles

There is no requirement for an FTP+ server process to support all the profiles. The STAT command will return the supported sets of a particular FTP+ server.

When should a particular Profile be used? This depends on the application requirements. As discussed earlier, the choice is a function of the transfer rate desired, the distance between the local and remote devices and the stack support available on a given machine. When deciding to use a particular profile, it must be remembered that switching to a different transport layer will often require a change to the FTP+ profile as well.

There are still some remaining issues with FTP+. Some of these are:

- The finalization of (i) new commands for the FC profile and (ii) the SMPTE work with NCITS T11 to define the requirements for the file-data protocol portion of the session. Additional descriptions needs to be developed for merging this SMPTE work and the efforts of ANSI. It is the intention of the Task Force that the SMPTE should standardize the control portion of the transfer session while ANSI should standardize the actual file-data transfer portion of the session.
- Discussion on the implications of designing a server or client process that supports various profile sets of FTP+. For example, a client process must “know thy server” before sending any commands associated with profile sets.
- The FC Profile does not support point-to-multipoint transfers. This is only possible when using the XTP transport layer.
- Error return values are to be finalized and some state diagrams are needed to describe a complete transfer session.

E.2.3.2.3. FTP+ file transfer using XTP

- FTP+ includes support for the eXpress Transport Protocol (XTP). XTP exists at the same layer as TCP. It is in many ways a superset of TCP but is not compatible with it.
- XTP has maximum-rate and other QoS-setting features, whereas TCP has no equivalent functions and will move data as fast as possible. XTP also supports point-to-multipoint transfers over IP. XTP is defined by the XTP Forum, and the SMPTE will refer to their specifications for its use within FTP+. It is undecided how XTP will achieve formal standardization. It is an open protocol with a history of use and stability.
- XTP is being selected to provide functionality that is lacking in TCP but which is required to meet user needs.
- XTP version 4.0b is recommended.

E.2.4. File transfer between different operating systems

As always, users must use care when moving files between machines with different operating systems. For example, the Windows NT File System (NTFS) only supports case-insensitive file names while UNIX systems have case-sensitive names. Also, the file access permissions in NTFS differ from those in UNIX.

E.2.5. Overview of FTP (RFC 959)

E.2.5.1. Background

The FTP+ protocol is based on FTP as defined in RFC 959. This section provides an overview of the commands and operation of FTP. The reader is strongly encouraged to read RFC 959 for more complete details on the operation of an FTP server and the specific FTP commands. This background, however, highlights the particular parts of RFC 959 most relevant to the FTP+ protocol.

E.2.5.2. FTP service overview

E.2.5.2.1. FTP control channel

The FTP protocol specifies the use of TCP port 21 as the destination port number for establishing a control channel with a FTP server. This can be done in a simple client / server configuration, in which a client program creates a single control-channel connection to a single FTP server. File transfer is then performed between the client host and the FTP server host. The FTP protocol also supports a "two server" configuration in which a client program establishes two control channel connections to two different FTP servers. In the "two server" configuration, file transfer is performed between the two FTP server hosts on command from the client program, via commands issued on the two control channels.

E.2.5.2.2. FTP data connections

In the simple client / server case, the FTP server will attempt to establish a data connection upon the receipt of a data transfer command (e.g. STOR or RETR). By default, this data connection is made from the FTP server's TCP port 20, back to the TCP port used by the client to establish the control channel. Note that this requires the client to "re-use" the same TCP port (see section 3.2 of RFC 959). Alternatively, the client may choose to request the FTP server to establish a data connection to a different TCP port. To do so, the client uses the PORT command to specify this different port number. The PORT command must be issued prior to any data transfer commands.

In a "two server" configuration, one FTP server must be "passive" and wait for a connection. This is accomplished using the PASV command. Upon receipt of a PASV command, the FTP server opens a new TCP port and begins to "listen" for an incoming data connection. The server sends back to the client, on the control channel, the TCP port number on which it is listening. The client then uses the address information returned by the PASV command to form a PORT command to the second server. The client then issues a data transfer command to this second server, which then initiates a data connection to the "passive" server, using this address information.

E.2.5.2.3. FTP server command processing

All commands and replies sent over the control channel follow the Telnet specification (RFC 854). These commands and replies are sent as separate "lines", each terminated by the Telnet EOL sequence (i.e. Carriage Return followed by Line Feed). Upon receipt of a command, the server will always reply with at least one line of information. These responses are sent over the control channel. They always begin with a 3-digit reply code (see section 4.2 of RFC 959) followed by either a space character or a dash (minus sign) and a text description of the reply terminated by the Telnet EOL sequence. If a minus sign follows the reply code, then there will follow one or more lines of text. These lines of text are then terminated by a line, again beginning with the identical 3-digit reply code, and immediately followed by a blank character. The multi-line response is thus bracketed by these reply code lines.

Here is an example of a one-line response:

```
220 bertha.acme.com FTP server ready
```

Here is an example of a multi-line response:

```
211- status of foobar
-rw-r--r-- 1 guest guest 1234567 Feb 2 11:23 foobar
211 End of status
```

Each digit of the reply code has a particular meaning. RFC 959 defines these codes using the symbol "xyz" where x is the first digit, y the second digit and z the third digit. The x digit is used to describe whether the reply is a positive, negative or incomplete reply. For example, a reply code with a "1" as the first digit indicates a positive preliminary reply. This will then normally be followed by either a reply code with a "2" as the first digit to indicate a positive completion reply, or a reply code with a "5" as the first digit to indicate a negative completion reply. The y digit defines the grouping of the error, for example a "0" as the second digit indicates a syntax error. The z digit is used to define finer gradation of errors. A client program need only check the first digit to verify the success or failure of a command, but the second digit may be of diagnostic help. Note that the reply text associated with a reply code is not fixed, i.e. the same reply code may have different text in different contexts. However, FTP client programs need only interpret the reply code, not the reply text, to determine the correct course of action.

The server does not accept any more commands until it completes (either successfully or with error) the command issued, at which time either a 2yz or 5yz reply code is issued. In some cases it is desirable to issue a command while waiting for a long command to finish (for example, to abort or inquire about the status of a transfer). To do so, the client must "interrupt" the FTP server process. This is accomplished using the Telent "IP" and "Synch" sequences (as described in RFC 854 and in the last paragraphs of section 4.1 of RFC 959). After sending the Telnet "IP" and "Synch" sequences, the client may issue one FTP command, to which the server will send a single line reply.

E.2.5.3. Reference RFC 959 commands

The FTP+ protocol defines new commands beyond those defined in RFC 959. However, it draws strongly from the example of the RFC 959 command set. In particular, the reader is referred to the following 959 commands. Each of them is modified or forms the basis of a new command in the FTP+ protocol.

E.2.5.3.1. TYPE

Defines the way (as text or as binary data) the file data is encoded and packed in the data stream as well as how the receiver should write it.

E.2.5.3.2. PORT

An FTP process which opens a TCP port and listens for an incoming connection will issue this command to another FTP process so that it will know how to connect to the listening FTP process.

E.2.5.3.3. PASV

Requests an FTP process to listen for an incoming connect and returns the TCP address on which it is listening.

E.2.5.3.4. RETR

Issued to start an FTP server process reading from a file and to connect to a client or passive FTP server process. *Section E.2.5.2.2.* describes how the data connection is made.

E.2.5.3.5. STOR

The same as RETR except that the data is now written to a file instead of being read. *Section E.2.5.2.2.* describes how the data connection is made.

E.2.5.3.6. LIST

Returns a listing of files and file information on the data channel. The format of the file information is not defined by RFC 959 and will vary with different operating systems. *Section E.2.5.2.3.* describes how the data connection is made.

E.2.5.3.7. NLST

Returns a list of file names only (i.e. a shortened form of the LIST command). *Section E.2.5.2.2.* describes how the data connection is made.

E.2.5.3.8. STAT

The STAT returns different information, based on the context in which it is issued. This command may be issued with or without an optional "pathname" argument. When issued with the "pathname" argument, this command behaves identically to the LIST command with the exception that the file information is sent back on the control channel, rather than the data channel.

If the STAT command is issued without an argument, the result depends on whether a transfer is in progress or not. If sent during the course of a transfer and with no argument, it will return the status of that transfer (i.e. the number of bytes transferred and the total number of bytes to be transferred). If used with no argument when a transfer is not in progress, it returns information about the state of the server process. See *Section E.2.5.2.3.* for a description of how a command can be issued during the course of a transfer.

E.2.6. FTP+ protocol specification

E.2.6.1. FTP+ components

As with FTP, FTP+ utilizes a control channel and a data channel. FTP+ provides additional capabilities for the data channel as well as for the identification of the Content to be transferred via the control channel.

E.2.6.1.1. FTP+ control channel

FTP+ will **always** employ a TCP connection for the control channel. FTP has been pre-configured (IETF-registered) to "listen" for control channel connections on TCP port 21. FTP+ will need to obtain another registered TCP port number from the Internet Assigned Numbers Authority (E-mail: iana@isi.edu). FTP+ client programs will always first establish a TCP connection to this port.

E.2.6.1.2. FTP+ data channel

FTP+ provides support for XTP transport, TCP and IPv6 networks, in addition to IP version 4. With the addition of XTP transport (and eventually Fibre Channel as specified by NCITS T11) will come the ability for data channels to become multicast channels that involve more than two hosts.

As with FTP, FTP+ supports both simple client / server connections as well as "two server" configurations in which data is transferred between two FTP+ servers using two control channels. These "two server" configurations may be of particular interest to automation systems. For example, an external automation program running on a workstation could use a "two server" configuration to initiate a transfer between two video servers or between a video server and a library management server.

In the case of XTP multicast transfers, the data channel will be "connected" simultaneously to multiple receiving servers. By nature, these multicast transfers are always done as "pushes". That is, each receiving server is instructed to join a multicast group by using a particular multicast address. Once every receiver has successfully joined the multicast group, the sending server can write data to this multicast data channel and it will be received by each receiving server.

E.2.6.2. Required settings for RFC 959 commands in FTP+

E.2.6.2.1. MODE

This command is used to specify the transfer mode as either Stream, Block or Compressed. FTP+ will always use Stream mode. (Stream mode sends the data as a simple stream of bytes). *Section E.2.6.6.2.* discusses how FTP+ will support the restarting of transfers using only the Stream mode.

E.2.6.2.2. STRU

This command is used to specify the File structure as having Page structure, Record structure or no structure (i.e. File structure). FTP+ will always use File structure, in which the file is considered to be a continuous stream of bytes.). *Section E.2.6.6.2.* discusses how FTP+ will support the restarting of transfers using only the Record structure.

E.2.6.2.3. TYPE

The only TYPE values supported are "A" for ASCII and "I" for Image. EBCDIC character and "non-8-bit byte" data representations are not supported.

E.2.6.3. Common features of FTP+ and FTP

This section documents some of the areas in which FTP+ uses features which are identical to those of FTP, and which are specified in IETF document RFC 959.

E.2.6.3.1. Error code syntax and interpretation

The reply code syntax described in *Section E.2.5.* of this annex, and specified in RFC 959 section 4.2, will also be used by FTP+. Some of the new FTP+ commands define additional reply codes which adhere to the FTP reply code syntax. These new reply codes are given in *Section E2.6.4.* of this annex along with the FTP+ command descriptions.

E.2.6.3.2. Multi-line control channel responses

In some cases it is necessary to send multiple lines of response over the control channel (for example, the output resulting from a STAT command which returns information about many files). The mechanism used by FTP+ for sending multi-line textual responses over the control channel is described in *Section E2.5.* of this annex and also specified in RFC 959 section 4.2.

E.2.6.4. FTP+ commands

The FTP+ commands are presented below, grouped into Transfer Protocol commands, Content Manipulation commands and System Status commands. As described in the FTP+ Overview, an FTP+ server will choose to implement the FTP+ commands, based on a "Profile" that is wishes to support. These profiles are the General Profile (GEN) for standard FTP+ operations, the Enhanced XTP Profile (EXTP) for XTP multicast operations, Fibre Channel Profile (FC) for all Fibre Channel operations and the Optional Profile (OPT) for providing general-purpose FTP+ enhancements not related to a specific transport layer.

The command descriptions below will state in which profile a command resides.

E.2.6.4.1. Transfer protocol commands

These commands are used to allow FTP+ to support transport protocols other than TCP on the data channel.

E.2.6.4.1.1. PORT h1,h2,h3,h4,p1,p2

Profile: GEN

This is the normal RFC 959 command, duplicated without modification. It is provided so that existing code, which performs FTP file transfer operations, can still be used with the FTP+ daemon without modification.

E.2.6.4.1.2. XPRT <protocol_specifier> <end_point_address>

Profile: GEN,XTP,FC

The XPRT (eXtendedPoRT) command supplements the RFC 959 command, PORT. XPRT serves a similar functional purpose yet offers enhanced capabilities and a different set of arguments. Namely, it specifies to the receiving FTP+ process, a network end-point to which a data connection shall be established. However, XPRT is generalized to allow the description of non-IPv4 networks and non-TCP transport addresses for the data connection. The ***protocol_specifier*** will describe the transport and network protocols to be used. The ***end_point_address*** will be the actual network address for the given network and transport, and which should be used for the data channel.

The <protocol_specifier> will be given as a " / " (slash) followed by a text string. Optionally, another " / " and another text string may be given, the first string being a ***transport_specifier*** and the second optional string being the ***network_specifier***. For example:

/<transport_specifier>/<network_specifier>

The <network_specifier> values that have been defined are: "IP4" for IPv4 and "IP6" for IPv6. The <transport_specifier> values that have been defined are "TCP", "XTP" and "FCP", which specify the transport layer as either TCP, XTP or Fibre Channel. The <network_specifier> is optional because some transports (XTP and FC) support a native mode in which no additional network protocol is used. However, both XTP and FC transport can be run on top of IP or some other network protocol.

The <end_point_address> specifies the network end-point. This must include both the network and the transport end-points (for example, an IP address and a TCP port number). The syntax of this string is dependent upon the network / transport being used. In the case of "/TCP/IP4", "/XTP/IP4" or "/XTP", it will be a dotted decimal 32-bit IP address followed by a ":" (colon) and the decimal port number. In the case of "/TCP/IP6" or "/XTP/IP6", it will be an address in string representation, as specified in RFC 1884, followed by a ":" (colon) and the decimal port number. The syntax for specifying FCP transport addresses has still to be defined by the Fibre Channel group.

Examples:

1. The following XPRT command specifies the use of TCP transport over IP version 4 networking and the issue of a data connection at the IP version 4 network address, 192.10.20.1, using TCP port 3456:

   XPRT/TCP/IP4 192.10.20.1:3456

2. The following XPRT command specifies the use of raw XTP transport and the issue of a data connection to the address 192.148.20.17 at port 1958 (Note that XTP, when used as both the transport and the network layer, uses the same IP version 4 network address definition and a 2-byte port number, but it does not use IP packets):

   XPRT/XTP 192.148.20.17:1958

3. The following XPRT command specifies the use of TCP transport over IP version 6 networking, and the issue of a data connection at the IP version 6 network address, 10A0:0:0:0:0:0.192.20.148.17, using TCP port 3458:

   XPRT/TCP/IP6 10A0:0:0:0:0:0.192.20.148.17:3458

Upon receipt of a valid XPRT command, the FTP+ process must respond with the standard "200 Command OK". The error code 501 would indicate a syntax error in the parsing of the XPRT arguments (e.g. insufficient arguments or an invalid network or port syntax). The new error code 522 would indicate that the FTP+ process does not support the transport / network combination specified. The text portion of the 522 error must include the list of transport / network protocol combinations that are supported by the FTP+ implementation, enclosed in parentheses and separated by commas, and followed by language- or implementation-specific text. For example, the response to an invalid XPRT might be:

522 (/XTP, /TCP/IP4) Supported protocols

This would indicate that only raw-mode XTP and TCP over IP version 4 are supported by that FTP+ implementation.

E.2.6.4.1.3. PASV

Profile: GEN

This is the normal RFC 959 command, duplicated without modification. It is provided so that existing code, which performs FTP file transfer operations, can still be used with the FTP+ daemon without modification.

E.2.6.4.1.4. XPSV <protocol_specifier> [<interface_name>]

Profile: GEN,XTP,FC

The XPSV (eXtendedPaSsiVe) command replaces the RFC 959 PASV command, serving a similar purpose yet offering enhanced capabilities and a different set of arguments. Namely, it requests that the receiving FTP+ process begin to listen for an incoming data connection and to respond to this command with the network address on which it is listening. In FTP+, the response to an XPSV command will be an address specification corresponding to the transport / network protocol passed as the <transport_specifier> (see the XPRT command in *Section E.2.5.3.* for a description of the syntax of the <transport_specifier>). This address specification shall be returned in the same format as required for use as the arguments for a corresponding XPRT command. The response to a valid XPSV command must be a "229" code followed by an XPRT-format network / transport protocol specification.

The optional ***interface_name*** may be used when connecting to a multi-homed FTP+ server. In some cases, it may be ambiguous as to which network interface the server should begin listening. In such cases, this string may be used to specify the interface. See *Section E2.6.6.3.* for a discussion on multi-homed host issues.

Here is an example of an XPRT command followed by a "229" response:

```
XPRT/XTP/IP4
229 (/XTP/IP4 192.20.10.1:2234) Entering passive mode
```

This would indicate that the FTP+ process on the host 192.20.10.1 is listening for a data connection on the XTP port 2234. The error code 501 would indicate an error in parsing the network or transport arguments (e.g. insufficient arguments). The error code 522 and 523 would be used to indicate that an invalid transport / network combination was specified (as in the XPRT command).

E.2.6.4.1.5. XPRT and XPSV examples

The following is an example of how the XPRT and XPSV commands could be used to set up a data connection between two FTP+ processes by means of a separate client process running on a third host (i.e. a proxy FTP+ session). First the client process would establish TCP control channels to both FTP+ processes and login to both FTP+ servers. Assuming the client wished to perform an XTP transfer over IP version 4, it would then send an XPSV command to the first FTP+ process. It would then use the response to this XPSV command as the arguments of an XPRT command to the second FTP+ process. At this point, a STOR or RETR command could be issued to either FTP+ process and the second FTP+ process would establish the IPv4 / XTP connection to the first FTP+ process. This interaction is shown below. Commands are given in bold text and responses in non-bold italics.

Host 192.20.10.1	Host 192.20.10.2
<**USER/PASS** commands>	
	<**USER/PASS** commands>
XPSV/XTP/IP4	
229 (/XTP/IP4 192.20.10.1:4623) Entering passive mode	
	XPRT(/XTP/IP4 192.20.10.1:4623)
	200 Command Okay
<STOR or **RETR** command**>**	

E.2.6.4.2. Content manipulation commands

This set of FTP+ extensions to the RFC 959 commands allows much more sophisticated control of how Content is located and identified for transfers. Specifically, FTP+ defines additional syntax and arguments to the RFC 959 commands RETR, STOR, STAT, LIST and NLST. Note that the STAT command is also used to obtain system and transfer information (as discussed in *Section E2.6.4.3.1.*). Common to all of these FTP+ commands is the use of a "Content Identifier".

E.2.6.4.2.1. Content Identifiers

Profile: OPT (The use of non-pathname Content Identifiers)

In RFC 959 each of the commands described in this section accepts only one argument, namely <***pathname***>. This is still supported and will still encompass full pathnames, directory names and wildcard pathnames. In addition FTP+ allows a Content to be identified using a <***content_identifier***> which is defined as:

<content_identifier>	=	***<syntax_id>://<content_spec>***
<syntax_id>	=	***<syntax_name>[<content_component>]***
<syntax_name>	=	[UMID | <***vendor_syntax_name***>]
<content_spec>	=	***<content_id>***
<content_component>	=	_META | _ESSENCE
<content_id>	=	<string>
<vendor_syntax_name>	=	<string>

The purpose of these new Content Identifiers is:

- to permit the storage, retrieval and listing of Content by plain text representation of the SMPTE Unique Material Identifier (UMID), or by vendor-specific Content identification schemes;
- to permit STOR and RETR to send and receive a Content's Metadata (_META), independently from its Essence (_ESSENCE).

A vendor who provides a particular query language for searching for Content on his Content server would register a name for this language with the SMPTE. For example, if vendor "XYZ" implemented an "XYZ Query Language" in his Content management system, the registered name might be "XYZ_QL", giving the following as a possible format for a STAT command:

```
STAT XYZ_QL://(version_name such as "J%")
```

. . . where the string ("J%") corresponds to the XYZ query language.

As another example,

```
STOR UMID_META://<universal id>
```

. . . could be used to store the Metadata-only of a Content that is identified by an SMPTE Unique Material Identifier (UMID), and,

```
LIST XYZ_QL://(ClipId like "B%")
```

. . . could be used to list information about all XYZ Content which has an attribute of ClipId whose value starts with a "B". Note that for the LIST and NLST commands, the _META and _ESSENCE options have no meaning.

The files transferred by FTP and FTP+ should be stored in the standardized Wrapper format being developed by the EBU / SMPTE Task Force. When the _META or _ESSENCE options are used to extract Metadata or Essence separately from Content, the server should send the Metadata or Essence wrapped in this standard format. Note that the receiving FTP+ process may choose not to store Content as files at all. The FTP+ specification only requires that the on-wire format be standardized to the EBU / SMPTE Wrapper formats and that Content stored with a given UMID or Content name should be retrievable using that same UMID or Content name.

E.2.6.4.2.2. STOR <content_identifier> [<start_byte_offset> [<length>]]

Profile: GEN, additional options in OPT.

In addition to the <content_identifier>, the STOR command is enhanced by FTP+ to include two new optional arguments, the ***start_byte_offset*** and the ***length*** of bytes to be transferred. These additional options are not applicable when transferring Metadata only. When the <start_byte_offset> is used (with or without <length>), the STOR command will write over or append to the existing Content with either the entire remainder of the Content's data (if no <length> is given) or with the exact <length> (number of bytes) to be transferred.

The reply text for a STOR should be enhanced by FTP+ to include the number of bytes written by the receiving FTP+ server. This will allow a client program to verify that the entire file was received. (See *Section E2.6.6.2.* for a discussion of restarts and failed file transfers). The format of this reply message should be:

```
226 Transfer completed.  12345678 bytes written
```

E.2.6.4.2.3. RETR <content_identifier> [<start_byte_offset> [<length>]]

Profile: GEN, additional options in OPT.

In addition to the <content_identifier>, the RETR command is enhanced by FTP+ to include two new optional arguments, the ***start_byte_offset*** and the ***length*** of bytes to be transferred. These additional options are not applicable when only transferring Metadata. The RETR command will always create a new Content (or a new version of Content) with exactly the number of bytes specified by the combination of <start_byte_offset> and <length>.

E.2.6.4.2.4. LIST <content_identifier>

Profile: GEN

The LIST command is enhanced only by the new <content_identifier>. The LIST command returns one line of information about each matching Content on the data channel. While FTP+ supports a wide range of new types of data channels for carrying out Content transfers, the LIST command need only support sending Content information over standard TCP data channels. The format of the information returned may be different on different implementations of FTP+.

E.2.6.4.2.5. NLST <content_identifier>

Profile: GEN

The NLST command is enhanced only by the new <content_identifier>. The NLST command returns Content names, one per line, for each matching Content on the data channel. While FTP+ supports a wide range of new types of data channels for carrying out Content transfers, the NLST command need only support sending Content information over standard TCP data channels. The format of the information should always be just the Content name terminated by the Telnet EOL sequence.

E.2.6.4.2.6. STAT <content_identifier>

Profile: GEN.

This form of the STAT command (i.e. with an argument) is enhanced only by the new <content_identifier>. The STAT command returns one line of information about each matching Content on the control channel, using the multi-line response syntax. The format of the information returned may be different on different implementations of FTP+. See *Fig. E.3* for a description of the different forms of the STAT command.

E.2.6.4.3. System information commands

These commands are used to provide additional information needed about the capabilities of an FTP+ system.

E.2.6.4.3.1. STAT

Profile: GEN.

When issued without an argument, the STAT command is used to return system information. In this context, the FTP+ version of STAT remains largely unchanged. As in RFC 959, if a transfer is in progress at the time when the STAT command is issued, then the status of the transfer operation is returned on the control channel. However, FTP+ will extend this status information to include an estimate of the time to completion. The format of this response must be:

213 Status: xxx of yyy bytes transferred; estimated time remaining hh:mm:ss

If the STAT command is given without an argument when no transfer is underway, then a 211 response code followed by FTP+ server information is returned. The first part of the response may be implementation-specific, but it is recommended that the current conventions of most FTP servers are followed.

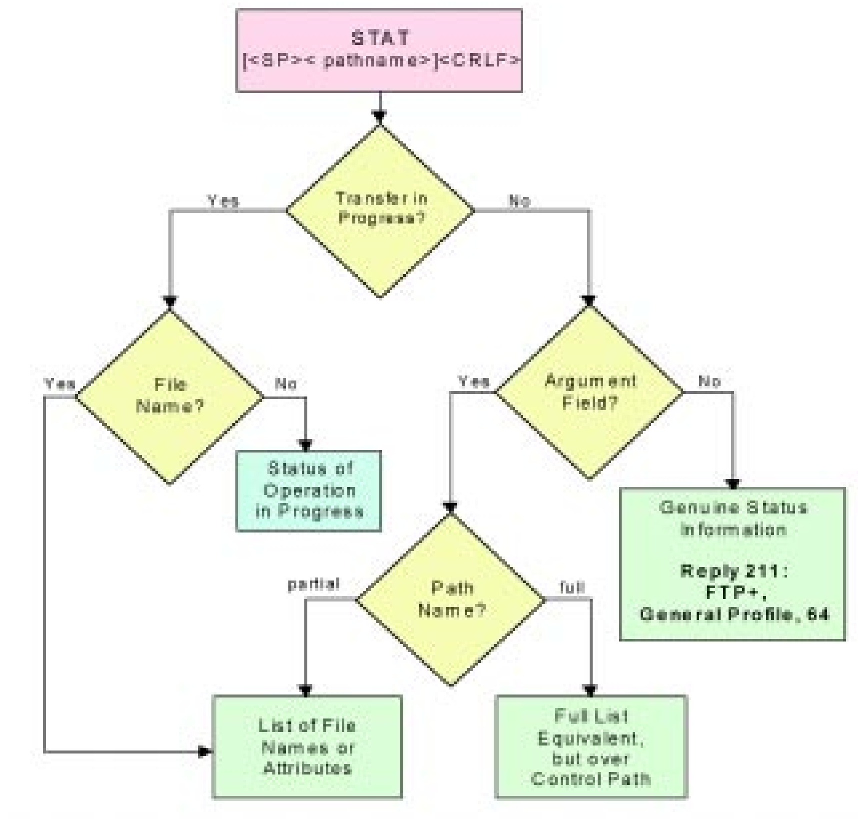


Figure E.3: Different forms of the STAT command.

Example 1:

Version X.X ACME_OS 01/25/98

Connected to foo.bar.com (10.20.1.2)

Logged in as username

TYPE: BIN, FORM: Nonprint; STRUcture: File; transfer MODE: Stream

No data connection

This must be followed by three lines of specifically-required FTP+ information:

Profiles: <ftp_plus_family_list>

Protocols: <protocol_list>

Size: 64 | 32

Where:

ftp_plus_profile_list	=	*ftp_plus_family* [,*ftp_plus_family*…]
protocol_list	=	*protocol_specifier* [,*protocol_specifier*…]
ftp_profile_family	=	GEN | EFTP | OPT | FC

. . . and <protocol_specifier> is as defined in *Section E2.6.4.5.3*.

The intention is that this version of the STAT command will show which set of FTP+ profiles this implementation supports, which transport / network protocol combinations are supported, and whether 64- or 32-bit file sizes are supported. In the following example, the following reply would indicate that the FTP+ implementation supports enhanced XTP capabilities and the optional FTP+ enhancements, and it runs only on IPv4 networks with both TCP and XTP transports of 64-bit files:

Example 2:

211- server.acme.com FTPPLUS server status
Version X.X ACME_OS 01/25/98
Connected to foo.bar.com (10.20.1.2)
Logged in as username
TYPE: BIN, FORM: Nonprint; STRUcture: File; transfer MODE: Stream
No data connection
Profiles: EXTP, OPT
Protocols: /TCP/IP4, /XTP/IP4
Size: 64
211 End of status

E.2.6.4.3.2. SITE RATE | META | CONT

Profile: OPT.

These extended variants of the RFC 959 SITE command will return some additional information regarding the features provided by the Optional profile. This information is returned as text responses to these commands.

The response to the SITE RATE command will be:

200 <maximum_rate>

. . . where ***maximum_rate*** (a decimal number of bytes per second) is the largest value accepted as an argument to the FTP+ RATE command.

The response to the SITE META command will be:

200 [NO | YES] (The _META option is [or is not] supported)

. . . which specifies whether or not the _META option is supported by the STAT, LIST, STOR, RETR and NLST commands.

The response to the SITE CONT command will be:

200 [NO | YES] (The _ESSENCE option is [or is not] supported)

. . . which specifies whether or not the _ESSENCE option is supported by the STAT, LIST, STOR, RETR and NLST commands.

E.2.6.4.4. Overview of multicast transfers

Multicast transfers are supported by XTP, either in "raw" mode or over IP.. However, the FTP+ has been designed to encompass other transports which may support multicast data transfers. To perform a multicast transfer, one server must first create a multicast group which other receivers will then join. The multicast group will be identified by a unique address that the sending server determined when it created the group. When the sending server later sends any data to this address, the underlying transport protocol will transmit this data in packets that will be received by each member of the group.

The XTP protocol was designed with the intention of supporting this type of operation. Specifically, it provides capabilities for multiple receivers to acknowledge the receipt of these multicast packets and to request, if need be, retransmission of lost packets.

E.2.6.4.5. Multicast transfer commands

These commands are used when the enhanced XTP option is being supported to provide multicast transfer capabilities.

E.2.6.4.5.1. RATE <rate_value> [<burst_value>]

Profile: EXTP.

The RATE command is added by FTP+ to allow a client to set the maximum transfer rate that will be used during the data transfer. The ***rate_value*** specifies the maximum average transfer rate and the ***burst_value*** specifies the maximum burst transfer rate. Both of these values are given as decimal bytes per second to be used in a transfer. The SITE RATE command (see *Section E.2.6.4.3.2.*) can be used to determine the maximum value supported by an FTP+ server. This command is only applicable to XTP which inherently supports this type of rate throttling.

E.2.6.4.5.2. MCGM Leaves [Yes | No]

Profile: EXTP

The “MultiCastGroupManagment (MCGM) command” controls (and shows) what the policy should be for FTP+ servers which attempt to join or leave a multicast group while a transfer is in progress. If set to “No”, then the sending server will abort the transfer and will return an error code and message on the control channel. If set to “Yes”, and at least one receiver still remains in the group, then the transfer will continue normally. If given with no arguments, the response to the MCGM command should be the current setting.

E.2.6.4.5.3. MCPV <protocol_specifier>

Profile: EXTP

The “MultiCastPassiVe (MCPV) command” is analogous to the XPSV command (see *Section E.2.6.4.1.4.*). The argument <protocol_specifier> is exactly the same. The purpose of the MCPV command is to request that the receiving FTP+ servers create a multicast group and return the <end_point_address> for that group so that it can be sent to the receiving FTP+ servers in an MCPT command.

The only transport that can currently support multicast transfers is XTP. A <protocol_specifier> is included in this transport so that future transports which support multicast can also be included.

E.2.6.4.5.4. MCPT <protocol_specifier> <end_point_address>

Profile: EXTP

The “MultiCastPorT (MCPT) command” is analogous to the XPRT command (see *Section E.2.6.4.1.2.*). The arguments <protocol_specifier> and <end_point_address> are exactly the same. The purpose of the MCPT command is to make it unambiguous that the specified <end_point_address> is a multicast address. The FTP+ server that receives an MCPT command does not simply establish a connection to the <end_point_address> as in the case of an XPRT command. Instead, it must join the multicast group corresponding to that address. The underlying system calls that perform these two tasks will be different. By having a separate MultiCastPorT command, it is clear (without relying on interpreting the address given) that a multicast connection is desired.

E.2.6.4.6. Examples of multicast transfers

In this first example, we will consider the non-proxy case where there is client code running on the sending server which wishes to send Content to three receiving FTP+ servers. First, this client code must obtain a multicast address (i.e. group) to use. Next, it must establish TCP control connections to each of the receiving FTP+ servers. Then, it will then issue MCPT commands to each of the receiving FTP+ servers. For example:

```
MCPT / XTP / IP 224.0.0.2:1234
200 Command Okay
```

This statement informs each receiving server that the multicast IP address 224.0.0.2 and XTP port 1234 will be used for this transfer. Next, the client code would issue STOR commands to each of the receiving FTP+ servers. Upon receipt of the STOR command, the receiving FTP+ servers would join the multicast group using, the address 224.0.0.2:1234, and would wait for data to arrive. Finally, the client code on the sending host can start writing data to the multicast address which will then be received by each receiving FTP+ server. After all of the data has been written, the client code on the sending server would close the multicast group. Each receiving

FTP+ server would detect this and finish writing data to its Content and would then write completion information back to the client server via the control channel.

In the second example, we will consider the proxy case in which the client code is running on a machine not participating in either the sending or receiving of the data. This client code would first establish TCP control channels to each of the involved FTP+ servers. Next, it would issue an MCPV command to the sending FTP+ server which would then return the multicast address that has been assigned for this transfer. For example:

> MCPV / XTP / IP
>
> 229 (/ XTP 224.0.0.1:4623) Entering passive mode

At this point, the client will send MCPT commands to each of the receiving FTP+ servers, using the multicast address returned by the MCPV command as in the first example. Next, the client will send STOR commands to each of the receiving FTP+ servers, which will cause them to join the multicast group. To start the actual transfer, the client will then send a RETR command to the sending FTP+ server. Once the transfer has finished, the client will receive completion information from all the involved parties via their respective control channels.

E.2.6.5. Writing commands for an FTP+ server

A particular implementation of FTP+ will probably not support all profiles, transports and network protocols. To do so, would entail a very large and complex piece of software. Instead, the writer of commands for an FTP+ server may wish to adopt the following implementation guidelines.

The FTP+ server can logically be divided into (i) a Core Component that receives most of the GEN or OPT profile commands involved in defining the mode of operation of the FTP+ server and (ii) Transport Components which are responsible for reading / writing Content data and sending / receiving this data over a particular transport.

E.2.6.5.1. The Base Component

The Base Component should be responsible for handling the GEN profile commands which are not related to any form of file transfer. These include the SITE commands, MODE, STRU, TYPE, LIST, NLST and STAT (with no arguments and no transfer in progress). The Base Component should also be responsible for looking for the “port commands” (i.e. PORT, XPRT, PASV and XPSV) to invoke the other components. One of the “port commands” is always required before issuing any other commands which will send data over a data connection. By looking at the particular “port command”, the Base Component can determine if the requested transport and network protocols are supported. If they are it can then determine whether an XTP, TCP or FC component should be invoked to handle it.

Once a Transport Component is invoked, the Base Component will continue to process any Content manipulation commands, either involving the OPT profile for extended Content Identifiers or not. Once Content to be transferred has been located, a handle to this Content – along with the command and arguments – will be passed to the Transport Component.

E.2.6.5.2. Transport components

These components implement the actual sending / receiving of data over the data connection of a particular transport.

- ➩ The TCP component will perform all operations which involve sending the data over a TCP connection. As this is the standard case (and is required for sending LIST or NLST data), it may be incorporated into the core component.
- ➩ The XTP component will perform all operations which involve sending the data over an XTP connection. This component will choose to implement either only the standard XTP operations (i.e. no multicast) or the full set.
- ➩ The FC component will perform all operations which involve sending the data over a Fibre Channel connection.

E.2.6.6. Miscellaneous FTP+ issues

E.2.6.6.1. Detecting partial Content transfer

FTP, as specified in RFC 959, has difficulty with Stream mode in detecting the difference between a broken data connection and the end of the data connection, because the receiving FTP server sees both cases as a closed connection. Block mode can provide a solution to this problem, but it is not supported by all FTP implementations. In the case where a sending or receiving FTP+ server crashes while in the process of performing a transfer, the fact that the control channel will also break can be used to detect that the whole Content may not have been transferred.

If the data channel breaks while the sending server is still writing data, the sending server will detect this condition and will send an error message over the control channel. However, if the data channel breaks near the end of a transfer, it is possible that the sending server may already have written all of its bytes into the system buffer, but that not all of these bytes have been emptied from the buffer and acknowledged by the receiving server. To solve this problem, the FTP+ specification proposes to enhance the reply response to the STOR command by including the number of bytes written to permanent storage and to enhance the RETR command by including the number of bytes written to the network. Using this enhancement, the client code on another FTP+ server should be able to verify that this number matches the number of bytes sent / received by it.

E.2.6.6.2. Restarting failed transfers

RFC 959 provides Block mode as a way to “checkpoint” data transfers. This is very important when two machines do not have the same byte size, or do not use Record Structures. However, since 8-bit byte sizes are now universal and Content Essence is universally stored as a byte stream, this problem can largely be solved by knowing how many bytes a server has successfully written.

To this end, STOR and RETR should include in their reply codes the number of bytes written or transmitted, even when the transfer is interrupted by a data channel connection failure. Then using the optional <start_byte_offset> and <length> arguments to STOR and RETR, a client can restart a failed transfer. In the case where an FTP+ server crashes, the STAT command can be used to determine the size of the file that was created on the receiving server to determine how much of it was successfully transferred.

E.2.6.6.3. Multi-homed hosts

When an FTP+ server is connected to more than one type of network interface, it can become problematic to perform certain types of transfers. One solution is to state that the interface on which the control channel has been established should be the interface for all data connections. For FTP+, this does not work well because of the variety of data channels supported and the lack of simple TCP connections on each type of transport protocol.

For example, consider the case of a server that has three Ethernet interfaces; one supports TCP / IP and the other two support raw XTP. A client program then wishes to initiate a two-server transfer between this server and another FTP+ server, using XTP on one of the XTP Ethernet interfaces. To do so, it must state on which interface it wishes this server to listen for a connection (i.e. on which interface it should perform the XPSV command). The additional XPSV command argument ***interface_name*** addresses this problem. It assumes, however, that the client program knows the “name” of the interface on which it wants the server to listen.

E.2.6.6.4. Fibre Channel addressing and FTP+ commands

The syntax for addressing Fibre Channel data connections is still not known at the time of writing (July 1998). More information concerning the use of Fibre Channel will be required before this issue can be resolved. Additionally, the FTP+ command set may need enhancements to support the setting up of a Fibre Channel connection (for example, buffer sizes). As the work of the Fibre Channel group proceeds, the FTP+ protocol must be enhanced as required to support these additional features.

E.2.6.6.5. "Raw-mode" XTP

The use of XTP without IP may be considered as optional. The benefits of using "raw-mode" XTP are limited at the present time. In the future, it is possible that "raw-mode" XTP might take advantage of the underlying physical media, for example using ATM QoS to implement the transfer rate setting. However, at the time of writing, no such capability exists. "Raw-mode" XTP does offer some small increase in efficiency at the present time, but this is not a significant factor.

E.3. Streaming

E.3.1. Scope

This section provides an introduction to the means by which television programme material may be streamed over transport technologies. A description of the processes, technologies and issues associated with streaming is provided, together with an overview of current practice.

E.3.2. Introduction to Content streaming

Streaming is the transfer of television programme material over a transport technology from a transmitter to one or more receivers, such that the mean reception frame-rate is dictated by the transmission frame-rate. The transmitter plays-out the material without receiving feedback from the receivers; consequently there is no capability for flow-control or for the re-transmission of lost or corrupt data. It is a continuous process in which the transmitter "pushes" programme material to receivers that may join or leave the stream at any time. The transmission frame-rate is not necessarily equal to the material's original presentation frame-rate, thus allowing faster- or slower-than-real-time streaming between suitably-configured devices.

E.3.2.1. Stream containers

Digital television programme material consists of Video, Audio and Data Essence, along with Metadata elements, which are segmented, multiplexed together and encapsulated within containers. Containers are transferred by the transport technology and include timing information to enable the receiving devices to synchronize the material's reception rate to its transmission rate; they may also provide mechanisms for the detection and correction of errors. A variety of element formats and container types are used within the television industry, and international standards define the rules for mapping between them. Certain types of containers are capable of encapsulating other containers, thus allowing the creation of hierarchies of programme material.

E.3.2.2. Transport technologies for streaming

The ideal transport technology for the streaming of programme material would provide a lossless and constant-delay channel. In practice, few networks exhibit such qualities and it is therefore important to select a transportation technology that supports the desired streaming requirements. To achieve this, it is first necessary to define the bounds within which the streaming is to be performed. This is achieved through the specification of a set of Quality of Service (QoS) parameters which include bit-rate, error-rate, jitter, wander, and delay and synchronization aspects.

The specification of QoS parameters is based on the requirements of the container and the application. Container requirements might include bit-rate and jitter; application requirements might include error-rate and delay. The selection of a particular transport technology is therefore based upon its capability for providing the required QoS. In some circumstances however, the suitability of the transport technology is also dependent upon the design of the devices being used to map the containers to that transport technology (e.g. the capability of the jitter removal system).

Network transport technology should support the transparent routing of streams over multi-node network topologies between devices identified through a unique address. Additional capabilities, such as mechanisms for multicasting streams in point-to-multipoint applications, are also beneficial.

E.3.2.3. Mapping stream containers to transport technologies

In order to transport streams over networks, it is necessary to use standards-based rules that define how the containers are mapped to and from the network transport technology, using network protocols. The mapping rules define how containers are segmented and reassembled, how synchronization is achieved and how error correction and detection are performed.

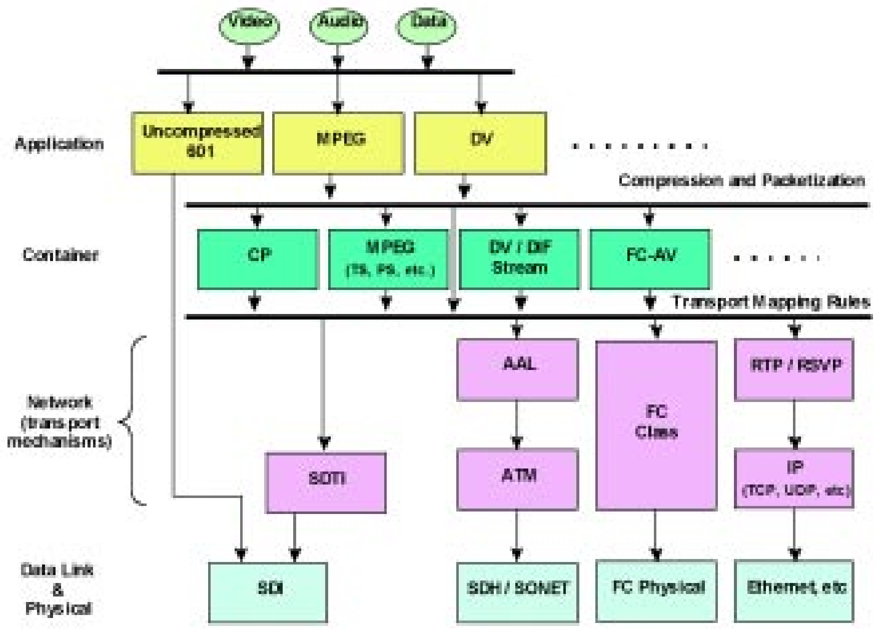


Figure E4: Stream delivery over selected packetized transports.

In a number of cases, standards bodies have defined, or are in the process of defining, such mapping rules. In cases where no mapping exists, the EBU / SMPTE Task Force will decide if a mapping will be of value (whereupon the appropriate standards bodies will be solicited) or will deem the mapping inappropriate. *Fig. E.4* provides a Reference Architecture showing the relationship between the functions.

The function marked "Container Mapping Rules" places the Content as a sequence of bits into a container (e.g. a multiplex). The function marked "Transport Mapping Rules" places the container into the data payload section of a selected underlying transport. This mapping is either performed through the use of a protocol (e.g. MPEG-2 TS payload into ATM, using the ATM AAL1 protocol) or through a direct mapping of the payload into frames (e.g. SDI).

E.3.2.4. System-level issues for streaming

The control of streaming devices requires the implementation of communication protocols and interfaces that provide capabilities for the following:

E.3.2.4.1. Initiation and termination of streaming

- configuration of the sending devices (e.g. bit-rate, MPEG-2 GoP structure etc.);
- selection of material to stream;
- configuration of multicast;
- access control mechanism.

E.3.2.4.2. Joining and leaving streams at the receiver

- configuration of receiving devices;
- selection of stream to join;
- access control mechanism.

E.3.2.4.3. Management of the service

- accessing network performance data (e.g. SNMP);
- defining the management information base (MIB) including monitor points and controls;
- providing remote access to performance data, e.g. Simple Network Management Protocol (SNMP).

No such protocols have currently been standardized and proprietary schemes are likely to prevail for the foreseeable future.

E.3.3. Considerations when streaming

The implementation of a Content streaming solution requires the following issues to be examined:

- Quality of Service requirements;
- Error Management capabilities;
- Timing and Synchronization;
- Addressing and Routing;
- Mapping streaming containers between transport technologies;
- Tunnelling and Bridging.

These issues are discussed in detail below.

E.3.3.1. Quality of Service

The following QoS parameters provide a basis for determining the suitability of a transport technology for streaming.

- **Peak bit-rate** (bit / s) – the bit-rate which the source may never exceed.
- **Minimum bit-rate** (bit / s) – the bit-rate at which the source is always allowed to send.
- **Sustained bit-rate** (bit / s) – the mean bit-rate at which the source transmits.
- **Jitter** (or Delay Variation).
- **End-to-end delay** (seconds) – i.e. propagation delay.
- **Bit Error Rate** (errors / bit) – the average number of errors per bit.
- **Set-up delay** (seconds) – the maximum delay between requesting a connection and receiving it.

It is recognized that some transport technologies permit other QoS parameters to be defined. These may be included in the mapping if necessary.

E.3.3.2. Error management

The implementation of a suitable error management system is essential for the streaming of high-quality Content over network technologies, since none provide completely error-free transport.

Error management is achieved through the implementation of mechanisms that provide error correction or error concealment. In either case, it is first necessary to detect the error in the streamed Content. Re-transmission is usually not possible due to the timing relationship that exists between the transmitter and the receiver. The selection of the error management mechanism is determined by examining the QoS requirements of the streaming application and the QoS characteristics of the transport technology. To realize any level of error management, a transmission overhead is necessary.

E.3.3.2.1. Error detection

Error detection enables the identification of lost or corrupt data. The granularity of error detection (i.e. whether errors are detected at the bit level, container level or frame level) plays a large part in the success of the error concealment.

E.3.3.2.2. Forward Error Correction

Forward Error Correction (FEC) enables the recovery of corrupt or lost data through the addition of redundancy in the stream. This typically introduces an overhead. There are a number of different FEC mechanisms, both proprietary and standards-based (e.g. ITU-T J.82 Reed-Solomon FEC for MPEG-2 over ATM, and ATSC A53 FEC).

E.3.3.2.3. Error concealment

Error concealment enables the effect of lost or corrupt data to be masked through the implementation of techniques such as replication and interpolation. The success of error concealment is dependent upon the sophistication of the error detection mechanism and the redundancy embedded in the Content.

E.3.3.3. Synchronization

Streaming data requires timing synchronization between the transmitters and receiver(s). This timing synchronization may be achieved through either the recovery of timing references embedded within the stream, or through the distribution of a system-wide clock to all participating devices.

In a transport environment there are two aspects to synchronization; the synchronization of the stream and the synchronization of the transport technology. The list below identifies a selection of video service synchronization tolerances (transport technology tolerances are outside the scope of this document):

- Composite video timing requirements are constrained by 0.23 ppm (PAL) and 2.8 ppm [22] (NTSC) as defined in ITU-R 470.
- Component video clock accuracy in hybrid environments is constrained to that of composite video.
- For MPEG Transport Streams, the System Clock accuracy is constrained by 30 ppm as defined in ISO 13818-1.
- SDTI, as defined in SMPTE305M, provides frame and line synchronization facilities through the use of the SDI interface according to ITU-R BT.656.

DV, as a deterministic system, uses a system of clocks directly related to 13.5 MHz. The DV clock is phase-locked to the main clock of the 525- and 625-line TV systems, and so its accuracy is the same as for the basic 525- and 625-line systems.

The proposed standard for carrying MPEG-2 Transport Streams (SDTI-TS) over SDTI defines that incoming MPEG Transport Packets are justified to one line length. The input and output receiver buffers will need to have a buffer capability of 1 line in order to buffer a constant TS bitstream at the input and output.

22. ±10 Hz / 3,579,545 Hz.

It is necessary for streaming receivers to implement mechanisms to remove any jitter introduced by the network technology in order to ensure the accuracy of the recovered stream time references.

E.3.3.4. Addressing and routing

Connections over transport technologies may be pre-defined or created dynamically. In either case, the implementation of an addressing scheme is necessary to identify the attached devices. The addressing scheme enables each attached device (i.e. transmitters and receivers) to be identified uniquely. The means through which this is achieved is dependent upon the transport technology. Typical addressing schemes are IP (v4 and v6), NSAP, E.164 etc.

Methods for routing streams over network technologies is beyond the scope of this document. Nevertheless, it can be stated that the routing function should be transparent when establishing connections between streaming devices.

E.3.3.5. Local- and wide-area streaming

Streaming is performed both within a facility and between facilities. Different considerations are required for each application. Local-area streaming within the facility is performed over a network that is owned and managed by the facility whereas wide-area streaming is generally performed over a network owned by a third-party service provider and which is shared between many subscribers. Consequently, differences in technology, scalability, economics, performance and QoS must be taken into account.

Many network technologies deployed within the local area are not suitable for use in the wide area, due to their restricted transmission distance and limited capability for scaling to large numbers of nodes. Conversely, network technologies for the wide area are often not economic for use in the local area. Local-area network technologies used within a facility include SDI / SDTI, Fibre Channel, ATM, IEEE 1394 and the DVB family. ATM is generally used to provide wide-area connectivity between facilities (over PDH or SDH / SONET). Consequently, gateways are required to provide inter-working between local-area technologies and wide-area technologies.

There are differences in the performance and QoS characteristics of local- and wide-area technologies. These include bandwidth, latency, jitter and synchronization.

E.3.3.6. Mapping containers between transport technologies (bridging)

Containers may be mapped from one transport technology to a different transport technology. This function is provided by gateways that are capable of supporting two or more transport technologies. The mapping is reversible and does not alter the Content or structure of the container.

Mapping will necessarily introduce delay, and re-timing (e.g. updating timing references within the container) may be necessary to meet synchronization requirements.

A bridging function is often required between a local studio network and a wide-area network, due to the different network technologies employed therein. Examples include the mapping of DIF blocks from SDTI to FCS, and MPEG-2 TS from FCS to ATM.

E.3.3.7. Tunnelling

Tunnelling is a way of transporting a complete interface data structure, including payload, through another interface; for example, IP over SDTI. IP multicast can then be used to transport both Data Essence and Metadata which are associated with the SDTI Audio / Video Essence carried in the main SDTI programme.

E.3.4. Supported mappings of containers to transport technologies

A mapping of container-to-transport technologies is provided in *Section 5.7.3*.

E.3.4.1. Overview of containers

E.3.4.1.1. The DV family

The DV compression family consists of a number of different schemes which are summarized as follows:

- DV Consumer at 25 Mbit / s:
  - 4:2:0 sampling in 625 / 50 countries;
  - 4:1:1 sampling in 525 / 60 countries.
- DVCPRO at 25 Mbit / s in 525 and 625 countries (4:1:1 sampling);
- DVCPRO50 at 50 Mbit / s in 525 and 625 countries (4:2:2 sampling);
- Digital S at 50 Mbit / s in 525 and 625 countries (4:2:2 sampling);
- DVCAM at 25 Mbit / s:
  - 4:2:0 sampling in 625 / 50 countries;
  - 4:1:1 sampling in 525 / 60 countries.

All the compression schemes share the so-called DIF (Digital InterFace) structure which is defined in the "Blue Book" (IEC 61834).

- Each **DIF block** is 80 bytes; comprising 3 bytes of header and 77 bytes of payload;
- A **DIF sequence** is composed of 150 DIF blocks;
- In the 25 Mbit / s version, a **DIF frame** consists of either 10 DIF sequences (525-line system) or 12 DIF sequences (625-line system). In the 50 Mbit / s version, the number of DIF sequences per frame is double that of the 25 Mbit / s version.

The Content of DIF blocks and the number of sequences will depend on the compression scheme and on the bit-rate (25 or 50 Mbit / s).

E.3.4.1.2. The MPEG-2 family

The ISO 13818 2 standard defines the following containers:

- **Elementary Streams** (ES). An MPEG-2 ES is a continuous stream which contains no timing information or headers. The transport of ES requires the addition of timing information.
- **Packetized Elementary Streams** (PES). MPEG-2 PES packets include timing and headers and are variable length (up to 64k).
- **Programme Stream Packets** (PS). MPEG-2 PS packets are 188 bytes long and contain only a single programme with embedded timing information.
- **Transport Stream Packets** (TS). MPEG-2 TS packets are 188 bytes long and support a multiplex of programmes with embedded timing information.

E.3.4.2. Transport technologies appropriate to streaming

E.3.4.2.1. SDI- and SDTI-based streams

SDTI, as defined in SMPTE 305M, uses the SDI interface according to ITU-R BT.656 as the physical transport system. Because of its TV frame-synchronous nature, SDTI is well suited for use in real-time critical applications such as streaming Content. The Content to be streamed is inserted in the predefined SDTI packets and therefore frame synchronization is automatically achieved. Because of the use of the SDI / ITU-R BT.656 interface, the data integrity depends on the performance parameters (BER, jitter, delay) determined by the SDI interface.

SDI is a unidirectional interface without re-transmit capabilities for corrupted data. Consequently, SDTI packets can be corrupted as well. In order to overcome this fact, Content needs to be protected by FEC.

SDTI streaming allows real-time streaming and faster-than-real-time streaming. For example, in a 270 Mbit/s SDI environment using SDTI, a payload rate of about 200 Mbit/s is achievable. Subtracting further overheads due to signalling etc., a 50 Mbit/s compressed video stream can easily be transferred at twice real-time speed.

Although SDTI provides addressing capabilities, its use in a streaming environment will be limited to point-to-point or point-to-multipoint applications. Usually the Content is headed by control packets for the receiving device. Examples of such control packets are machine control parameters or transfer speed information.

Each application which takes advantage of SDTI requires (i) full documentation of the data to be transmitted (e.g. DV or MPEG-2), (ii) the data structure of the source stream to be inserted (DIF or MPEG TS including the control information, if used) and (iii) the mapping of the source stream into the structure provided by SDTI.

E.3.4.2.2. ATM-based streams

ATM is suitable for streaming digital television programme material over both the local and wide area, and standards have been defined for the transport of MPEG-2 Transport Streams.

The following issues are addressed by the standards:

- ⇨ Service Class Selection;
- ⇨ mechanisms for MPEG-2 Transport Packet encapsulation;
- ⇨ clock synchronization and de-jitter;
- ⇨ error detection and correction.

MPEG-2 requires an ATM service-class that is connection-oriented and which supports real-time transfer. Class A (CBR) and Class B (VBR-RT) service classes support these requirements. However, the current (July 1998) immaturity of standards for Class B invariably means that Class A is used.

The MPEG-2 Transport Stream that is presented to the ATM adaptation must therefore also be CBR. This is achieved within MPEG-2 encoders by the use of buffering and the implementation of a rate control mechanism which alters the quantization level (and hence the bits per frame).

The AAL defines how the MPEG-2 Transport Stream (TS) mapping, error handling and synchronization are performed. The selection of the appropriate AAL is important since it has a significant impact on the QoS. The ITU-T J.82 standard specifies how either AAL1 (CBR) or AAL5 (VBR) can be used for MPEG-2 streaming applications.

The mapping of MPEG-2 Transport Streams to ATM cells is specified by the AAL Segmentation and Re-assembly function. An MPEG-2 Transport Stream consists of 188-byte packets which must be mapped into the 48-byte payload of the ATM cells. AAL1 uses one of the payload bytes for sequence numbering (which enables the detection of lost cells), thereby allowing an MPEG-2 TS packet to be mapped into exactly four ATM cells. AAL1 allows two MPEG-2 TS packets to be mapped into eight ATM cells (together with the CPCS-PDU trailer).

MPEG-2 uses a 27 MHz system clock to synchronize the operations of the decoder to those of the encoder. The synchronization of the clocks is achieved through the use of MPEG-2 TS Programme Clock References (PCRs). An ideal network maintains a constant delay between each PCR. In practice, the CDV within the ATM cell stream can result in unacceptable PCR jitter at the decoder. It is therefore necessary to implement mechanisms for overcoming the jitter.

AAL 1 is CBR and it provides two mechanisms for timing recovery: Synchronous Residual Time Stamp (SRTS) and Adaptive Clock Recovery. Since AAL 5 is VBR, Adaptive Clock Recovery is used. It should be noted however, that the jitter-removal capability of all ATM devices is design-dependent.

Error-correction and detection mechanisms can significantly improve the quality of the received stream. The AAL 1 specification (ITU-T I.363) includes a FEC and byte interleaving mechanism that is capable of recovering up to four lost cells in a group of 128 cells. In addition, the cell sequence numbering provided by AAL 1 allows errors to be detected on a per-cell basis, thus aiding error concealment. The AAL 5 specification contains no mechanism for error correction and thus the error detection takes place at the SDU level.

Although the ATM Forum recommends the use of AAL 5 for consumer-quality Video-on-Demand (VoD) services, the benefits of low jitter and standards-based error detection and correction have led the DVB Project to recommend the use of ATM with AAL 1 for the transport of professional quality video over PDH and SDH networks (see the ETSI Standards, ETS 300 813 and ETS 300 814). AAL 1 is therefore recommended for the transport of MPEG-2 TS packets over wide-area networks.

E.3.4.2.3. IP-based streaming

IP (Internet Protocol) streaming is not synchronous compared to the video delivery rate. Synchronization of streams over IP therefore requires timing references to be embedded within the stream. IP streaming requires the support of IETF RTP (Real-time Transport Protocol). RTP permits real-time Content transport by the inclusion of media-dependent Time Stamps that allow Content synchronization to be achieved by recovering the sending clock. RSVP (Resource reSerVation Protocol) provides network-level signalling to obtain QoS guarantees.

RSVP is a QoS signalling protocol for application-level streams (called “flows” in the IP community) and is defined by a number of RFCs. The most relevant are: RFC 2205 (Protocol Specification), RFC 2208 (Applicability Statement) and RFC 2209 (Message Processing).

Session-control protocols for streaming (in draft form, July 1998) include RTSP (Real-Time Sessions Protocol), SDP (Session Description protocol) and SAP (Session Announcement Protocol).

UDP, as defined in RFC 768, can be used as an option to enable bounded-quality transfers on top of the IP layer. It allows broadcast transmissions and is a datagram-oriented protocol.

Some of the application scenarios for IP streaming are:

- ➩ Intranet, Internet browsing of Content (e.g. browsing an archive);
- ➩ Internet broadcasting (lectures, conferences, etc.);
- ➩ Web-based off-line editing;
- ➩ Web-based “channels” (IP-TV).

E.3.4.2.4. Fibre Channel-based streaming

Fibre Channel is suitable for streaming Content in local-area and campus-wide networks. Standards have been completed or are under development that can transport uncompressed material and Content streams. These standards cover:

- ➩ Basic network services:
- ➩ Fractional Bandwidth services.

Work is underway on an extended FC standard that covers the following:

- ➩ Encoding and encapsulation of uncompressed Content;
- ➩ Simple container (Wrapper) model;
- ➩ DV compressed streams;
- ➩ MPEG Transport Streams.

The FC Fractional Bandwidth service by itself does not provide the high-quality jitter control required by broadcast systems. This can be accomplished using the Fractional Bandwidth service with another synchronization mechanism.

Timing reference information can be embedded in the data stream (as it is in MPEG Transport Streams) or a common clock can be used.

The FC-AV project has developed an additional bandwidth management scheme. This protocol can be implemented using current FC hardware. This scheme uses larger buffers so that the additional jitter of a higher-level bandwidth management protocol can be supported.

Fibre Channel is defined by the following base standards:

- **FC-PH (X3.230-1994)** covering the basic interface;
- **FC-PH-2 (X3:297-1997)** which extends the basic network with the Fractional Bandwidth service, a Class 3 (datagram service) multicast model and other enhancements;
- **FC-AL (X3:272-1996)** which defines the basic arbitrated loop model;
- **FC-LE (X3:287-1996)** which describes part of IP on FC models;
- **FC-FCP (X3:269-1996)** which describes SCSI-3 encapsulation on FCS.

Other work of interest includes:

- **FC-PH-3 (X3:303-199x)** which covers further enhancements to the basic interface, including higher bandwidth links and a multicast service for Class 1 (a connection based service);
- **FC-AL-2 (Project 1133-D)** which covers enhancements for arbitrated loops.

Work is also underway on protocols and other extensions for A / V applications in FC-AV (project 1237-D).

Annex F

Acknowledgments

Grateful thanks must be extended to all those people, and their generously supportive employers, who have worked long and hard to create the two reports that have been the tangible output of the Joint EBU / SMPTE Task Force for Harmonized Standards for the Exchange of Programme Material as Bitstreams.

Particular mention must be made of the co-Chairmen of the Task Force, Merrill Weiss and Horst Schachlbauer, whose farsightedness and sense of purpose (not to mention their black belts in diplomacy) contributed fundamentally to the work process. Other than these, naming the people who really got to grips with the work throughout the life of the Task Force would perhaps be invidious here, but they (and we) know who they are, and their standing in our industry is correspondingly, and deservedly, very high.

Meetings don't happen for free. Considerable investment in the hosting of meetings has been made in terms of time, conference facilities and catering – in particular by Sony, the EBU, Microsoft, Turner Entertainment Network, Panasonic, Philips, Planon Telexpertise and Pro-Bel. Their invaluable assistance to the Task Force has been much appreciated.

Thanks must also go to the BBC Research and Development Department and David Bradshaw, in particular, who very kindly provided the FTP site and the e-mail reflector which the Task Force has been using to great effect.

The intrepid Secretary to the Task Force throughout has been Roger Miles, of the EBU Technical Department. His patience with the Chairmen, and his persistence in pulling together the many undertakings of the Task Force, has been, in many ways, the glue that has held the operation together.

Finally, mention must be made of Mike Meyer, Editor of the EBU Technical Review, who spent a considerable amount of his valuable time on the raw text and the layout, fashioning this report into an admirably lucid document for publication both by the EBU and the SMPTE.

Annex G

Task Force participants

Given below is a list of the individuals who participated in the deliberations of the EBU / SMPTE Task Force and their company affiliations.

Name		Organization
ABRAMS	Algie	Microsoft
ACKERMANS	Merrick	Turner Entertainment Net.
ACOSTA	Juan	IBM
AGOSTON	György	MTV (Hungary)
ALLBRITTON	Robin	C-Cube
ARDITO	Maurizio	RAI
ATKINSON	Steve	Avalon Consulting
BANCROFT	Dave	Philips DVS
BANTHORPE	Steve	BT
BASTANI	Behzad	Philips Research
BEERS	Andy	MS NBC
BENJAMIN	Shai	ECI Telecom
BERNARD	Frank	SGI
BIRKMAIER	Craig	Pcube Labs
BIRKS	David	Mitsubishi
BOCH	Laurent	RAI
BØHLER	Per	NRK
BOUDIGUES	Pierre	TF1
BRADSHAW	David	BBC
BRATTON	William F.	Turner Broadcasting Syst.
BRENNAN-TURKOWSKI	Rita	Apple Corp.
BROOKS	David	Snell & Wilcox
BRYDON	Neil	
BUKOWSKI	John	Quantel
CAMERON	Don	JVC
CASTLE	Gordon	CNN R&D
CHOU	Philip A.	Microsoft
CINQUINI	Maurice	Philips DVS
CLAY	Nicholas	Microsoft
COCKRAM	Clare	BBC
COLEY	Bryan	Turner Entertainment Net.
COLLIS	Ian	Sony BPE
COTTAM	Seb	Sony BPE
CUDEL	Damian	TF1
DABELL	Adrian	
DALE	Andy	Avid

Name		Organization
DARE	Peter	Sony
DAVIS	Franklin	OMF
DELAIGLE	Jean-Francois	UCL
DIMINO	Giorgio	RAI
DRISCOLL	Sean	Merrill Weiss Group
DUCA	Jim	Consultant
EBNER	Andreas	IRT
EDGE	Bob	Tektronix
ELMER	Peter	BT Labs
EMMETT	John	BPR Ltd.
FANKHAUSER	Eric	Gennum
FARJEON	Marc	Warner Bros.
FELL-BOSENBECK	Frank	Panasonic B'cast
FENTON	John C.	Mentat Inc.
FERNANDO	Gerard	Sun Microsystems
FERRANTE-WALSH	Lisa	OMF / AVID
FIBUSH	David	Tektronix
FLORENCIO	Dinei	Sarnoff Corp.
FUHRER	Jack	Hitachi
GABRIEL	John	The Duck Corp.
GAGGIONI	Hugo	Sony
GARSHA	Chuck	Paramount Pics.
GILMER	Brad	Gilmer Associates
GIROD	Carl	SMPTE
GOLDSTONE	Ira	Tribune
GOLMI	Nada	NIST
GRAY	Bob	Pluto
GUZIK	Ken	Sun Microsystems
HAAS	Norman	IBM
HARRIS	Brooks	B. Harris Film & Tape
HART	Jim	Panasonic B'cast
HAYCOCK	David	Artel Video
HEDTKE	Rolf	FH Wiesbaden
HEITMANN	Juergen	AV Media Tech.
HILKES	Rob	Gennum
HOFFMANN	Hans	IRT
HOPPER	Richard	BBC
HORI	Akihiro	NTV
HULME	Pete	Sagitta
HUNTER	Kurt	Microsoft
IDE	Akifumi	Matsushita
IRVIN	Martin	Sony BPE
IVE	John G.S.	Sony
IWASAKI	Yasuo	Sony
JACOBY	Ronald	SGI

Name		Organization
JOHNSEN	Reidar Otto	NRK
JONES	Chris	Drake Automation
JORGENSEN	Chris	NBC
JULIAN	Marcus	Vibrint Tech.
KAISER	Martin	RBT
KANE	Joseph	ImagingScience
KAPUCIJA	Tom	Gennum
KEANE	James	NBC
KELLEHER	Terry	Pathlight Tech.
KERSHAW	Sam	ECI Telecom
KISOR	Robert	Paramount Pics.
KLEMETS	Anders	Microsoft
KNÖR	Reinhard	IRT
KOBAYASHI	Akira	JVC
KOLLER	Herbert	AMEC
KÖNIG	Peter	NDR
KOVALICK	Al	H.P.
LEE	Henry	Microsoft
LEGAULT	Alain	Matrox
LEWIS	Lisa	Excell Data
LIRON	John	Tektronix
LIU	John C.	SGI
LIVINGSTON	Philip	Panasonic
LONG	Stephen	DOD / NIMA
LUAN	Laura-Zaihua	IBM
MAIER	George	Artel Video Systems
MAJIDIMEHR	Amir	Microsoft
MALTZ	Andrew	Digital Media Technologies
MARDER	Stan	The Duck Corp.
MARSH	Dave	Snell & Wilcox
MAYCOCK	Neil	Pro-bel
MCDERMID	Edward	Avid
McMAHON	Thomas	Microsoft
MIHARA	Kanji	Sony
MILES	Roger	EBU
MILLER	William C.	ABC-TV
MILLER	Daniel	The Duck Corp.
MINZER	Oren	ECI Telecom
MORGAN	Oliver	Avid
MORIOKA	Yoshihiro	Matsushita
MORRIS	Ken	IBM
MURPHY	Robert	Avid
NAGAOKA	Yoshimichi	JVC
NAKAMURA	Shoichi	NHK
NELSEN	Don	Avid

Name		Organization
NELSON	Quentin	Utah Scientific
NEUBERT	Neil	JVC
NEWELL	Chris	BBC R&D
NICHOLSON	Didier	Thomson CSF
NOTANI	Masaaki	Matsushita
NOURI	Semir	JVC
NOVAK	Ben	Microsoft
ODONNELL	Rose	Avid
OSTLUND	Mark	HP
OWEN	Peter	Quantel
OZTEN	Saruhan	Panasonic
PALMER	Douglas	VideoFreedom
PALTER	D.C.	Mentat Inc.
PAULSEN	Karl	Synergistic Tech. Inc
PENNEY	Bruce	Tektronix
PETERS	Jean-Jacques	EBU
PIPPENGER	Den	VideoFreedom
PITTAS	John	Seachange
POE	C.	VideoFreedom
POINT	Jean-Charles	Thomson TBS
POLGLAZE	David	Sony BPE
PUGH	Michael	Digital Acoustics
RADLO	Nick	Journalist
RAHIMZADEH	Auri	Envisioneering Gp.
RAHMAN	Altaf	Philips
RAYNER	Andrew	BT Labs
RODD	Spencer	Pharos
ROE	Graham	Pro-Bel
ROMINE	Jeff	Philips BTS
ROSEN	Andrew	Microsoft
ROSSI	Gerry	The Duck Corp.
RYNDERMAN	Richie	Avid Cinema
SADASHIGE	Koichi	NTA
SAFAR	Johann	Panasonic
SAKAI	Prescott	Cypress S/C
SAVARD	Stéphane	Planon Telexpertise
SCHAAFF	Tim	Apple
SCHACHLBAUER	Horst	IRT
SCHAEFER	Rainer	IRT
SHELTON	Ian	NDS
SHISHIKUI	Yoshiaki	NHK
SIEGEL	Jon	Object Model Group
SLUTSKE	Bob	Nat. TeleConsultants
SMITH	Garrett	Paramount
SMITH	Clyde	Turner Entertainment Net.

Name		Organization
STEPHENS	Roger	Sony BPE
STEVENS	Paul	Silicon Graphics
STOREY	Richard	BBC R&D
STREETER	Dick	CBS
SUTHERLAND	Joe	C-Cube
SUZUKI	Tomoyuki	JVC
SYKES	Peter	Sony BPE
TANAKA	Masatoshi	Matsushita
TATSUZAWA	Kaichi	Sony
THIRLWALL	Clive	Sony BPE
TICHIT	Bernard	Thomson Broadcast
TOBIN	Bill	Sun
TUROW	Dan	Gennum
WALKER	Paul	4 Links
WALLAND	Paul	Snell & Wilcox
WALTERS	David	Snell & Wilcox
WARD	Chris	Sarnoff
WEINSTOCK	Mitchell	QuickTime Co.
WEISS	S. Merrill	Merrill Weiss Group
WELLS	Nicholas	BBC
WENDT	Hans	JVC Germany
WHITE	Jason	Microsoft
WILKINSON	Jim	Sony
WODE	Gary	Wode Designs
WRIGHT	Colin	7 Network Australia
YAGI	Nobuyuki	NHK
YATES	Anthony	Yates & Co.
YONGE	Mark	Solid State Logic
YOSHIDA	Junko	EE Times
ZACCARIAN	Paulo	Consultant