European Broadcasting Union

Society of Motion Picture
and
Television Engineers

Task Force for Harmonized Standards for the Exchange of Program Material as Bit Streams

First Report: User Requirements

April 1997

Contents

CONTENTS i

PREFACE vii

EXECUTIVE SUMMARY ix

0 INTRODUCTION 1

0.1 Open Standards 2

0.2 Scope of Activities 2

1 COMPRESSION ISSUES 5

1.1 Image Quality 6
1.1.1 Coding Method 6
1.1.2 Sampling Structure 6
1.1.3 Compression Pre-processing 7
1.1.4 Data Rate 7
1.1.5 Group of Pictures 7
1.1.6 Constant Quality vs. Constant Data Rate 8
1.1.7 Editing 8
1.1.8 Concatenated Compression 9

1.2 Quality Levels 10

1.3 Operational Considerations 10

1.4 Storage 11
1.4.1 Interfaces 11
1.4.2 Data Rate Requirements 11
1.4.3 Resource Management 11
1.4.4 Audio, Video and Metadata Synchronization 11
1.4.5 VTR Emulation 12

1.5 Interoperability 12

1.6 Compliance Testing 12

1.7 Compression Issues – Recommendations 12

2 WRAPPERS AND METADATA 15

2.1 Purpose of Wrappers 15

2.2 Terminology – What's in a Wrapper 15
2.2.1 Content Structure 16
2.2.2 Essence 17
2.2.3 Metadata 17
2.2.4 Metadata Characteristics 18
2.2.5 Overhead 18

2.3 General Requirements 18

2.4 Breadth of Application and Wrapper Profiles 18

2.5 Metadata Requirements and Metadata Sets 19

2.6 Wrapper Size 19

2.7 Platform Neutrality 20

2.8 Interleaving 20

2.9 Unique Identifiers 20

2.10 Immutability and Generation Numbering 20

2.11 References 20

2.12 Indexing 21

2.13 History 21

2.14 Access Control 21

2.15 Support of Transactions 21

2.16 Property Rights 22

2.17 Asset Management 22

2.18 Application Programming Interface (API) 22

2.19 Compatibility and Conversion 22

2.20 Extensibility 22

2.21 Wrappers and Metadata – Recommendations 23

3 FILE MANAGEMENT, TRANSFER PROTOCOLS, AND PHYSICAL CONNECTIONS 25

3.1 File Management and File Systems **25**

3.2 File System Attribute Requirements **25**

3.3 File System Commands **25**

3.4 Digital Data Transfers **25**

3.5 Transport mechanisms **26**

3.6 Physical and Link Layer Considerations **27**

3.7 High Level Management Functions **27**

3.8 Further work **28**

ANNEX A – ABBREVIATIONS AND SPECIALIZED TERMS 29

ANNEX B – RECOMMENDATIONS 41

B.1 Compression Issues – Recommendations **41**

B.2 Wrappers and Metadata – Recommendations **42**

B.3 File Management, Transfer Protocols, and Physical Connections – Recommendations **42**

ANNEX C – COMPRESSION APPLICATIONS 43

C.1 The Application Model **43**

C.2 Acquisition **43**

C.3 Contribution **43**

C.4 Compression in Live/Real Time Production. **44**

C.5 Post Production/Exploitation **44**
C.5.1 Editing in real time 44
C.5.2 Post production editing 44
C.5.3 Off-line post production. 45
C.5.4 Presentation / Master Control. 45

C.6 Distribution **45**
C.6.1 Satellite links 45
C.6.2 Public carriers (telcos). 45
C.6.3 Emission 45
C.6.4 Packaged media distribution 45

C.7 Archiving **46**
C.7.1 On-line archiving 46
C.7.2 Near-line archiving 46
C.7.3 Deep / long-term archiving 46

ANNEX D – WRAPPERS AND METADATA **47**

D1 – PLATFORM NEUTRALITY **47**

D1.1 Bit ordering **47**

D1.2 Multi-byte data **47**

D1.3 Endian-ness flag issues **48**

D1.4 10-bit data and ITU-R BT.601 C_b Y C_r [Y] sample structure **48**

D2 – WRAPPER REFERENCING **49**

D2.1 Wrapper Varieties **49**

D2.2 Referencing Content **49**

D2.3 Reference Types **50**

D3 – ACCESS CONTROL AND COPYRIGHT **51**

D3.1 Access Control **51**

D3.2 Intellectual Property Rights **52**

D4 – NOTES **53**

D4.1 Applications **53**

D4.2 Usage of Essence and Metadata **54**

D4.3 References and Labels **55**

D4.4 Security **55**

D4.5 API **55**

D4.6 Essence Extraction **55**

D4.7 Efficiency and Completeness **55**

D4.8 Extensibility **56**

D4.9 Immutability and Generation Numbering **56**

D4.10 Endian-ness of 10-bit Sample Structures **56**
D4.10.1 “Straightforward” Packing 57
D4.10.2 “LSBs separate” Packing 57
D4.10.3 “Sparse” Packing 57
D4.10.4 “Tight” Packing 57
D4.10.5 Permutation during the transfer 58
D4.10.6 Conclusion 58

D4.11 Ideas for Wrapper Referencing - SMPTE 258M and HTML **58**

D5 – DAVIC AND DAVIC TERMINOLOGY **59**

D5.1 Monomedia Component Types: **59**

D5.2 Content Structure Definitions **59**

D5.3 Metadata **61**

D5.4 Overall Structure **61**
D5.4.1 Metadata Data Structure Example 62

ANNEX E – FILE MANAGEMENT, TRANSFER PROTOCOLS, AND PHYSICAL CONNECTIONS **63**

E.1 Definitions: **63**

E.2 The OSI Reference Model **64**

E.3 File Management and File Systems **65**
E.3.1 File System Attribute Requirements 65
E.3.2 File System Commands 65
E.3.3 File System Layers 66

E.4 Digital Asset Transfers **66**
E.4.1 “File Transfer” Requirements for point-to-point Connections 66
E.4.2 "File Transfer” Requirements for point-to-multipoint Connections 67
E.4.3 “Asset Play” Requirements for point-to-multipoint Connections 67

E.5 Transport mechanisms **68**
E.5.1 Synchronization of associated data (e.g. Audio-Video-Meta-Data) 68
E.5.2 Transfer modes and QoS parameters 68
E.5.3 Transfer initiation parameters 69
E.5.4 Transfer phase 69
E.5.5 Example of the specification of a class of transfer 70
E.5.6 Transfer interaction management, exception reporting 70
E.5.7 Basic transport protocols 71

E.6 Physical Interfaces **71**
E.6.1 Physical layer 71
E.6.2 User Requirements for Physical Interfaces 71

E.7 High Level Management Functions **73**

E.8 Application example **74**

E.9 Wide Area Network Gateway **74**

E.10 User Requirements for standards **75**

ANNEX F – TASK FORCE PARTICIPANTS **77**

Preface

Radical changes have begun in the methods used in the production, post production, and distribution of television programs. The changes leverage the technologies of digital signal processing, computers, and data communications. They offer the possibilities of significantly enhanced creativity, improved efficiency, and economies of scale in the creation of television programming. Driven by the growing demand for programming to fill the multiplicity of competitive distribution channels to consumers that are being put into place around the world, they are likely to become pervasive in teleproduction over the next five to ten years.

To harness the full potential these changes present, certain prerequisites must exist. These include well-understood methods for dealing with a variety of compression schemes, standard interfaces for the interconnection of equipment, standardized file formats, uniform means for collecting and storing the many elements that contribute to programs, and common mechanisms for describing program elements and for accumulating related information. For maximum benefit to be derived from the new techniques, the underlying supporting structures must be implemented identically world wide.

With this background, 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 on the Exchange of Television Programs as Bit Streams. As its first assignment, the task force was charged with producing a blueprint for the implementation of the new technologies looking forward a decade or more.

This rather remarkable document is the result of that effort. It was produced by a group of some 80 experts from Europe, Japan, and North America, meeting formally five times over a period of less than six months. It represents the first attempt by the several industries involved to look into the future together and to set a direction for all to follow. It takes as its premise the need to identify requirements that users will have as they adopt the new methods. It includes the views of users and manufacturers, both of which are needed in order to get a picture of 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 achieveable. Because of the many trade-offs that exist between compression methods, their parameters, and the performance achieved in specific applications, different techniques will be required in particular situations to meet explicit requirements. Thus the greatest benefit to all concerned will come from providing mechanisms that will permit systems to easily handle various compression schemes while maintaining the maximum quality of the program 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 is one of the most important findings of this effort, a new class of program-related data called “metadata,” which is descriptive and supporting data connected to programs or program elements. Metadata is intended both to aid directly in the use of program content and to support the retrieval of content as needed during the post production process.

The work of the task force is not done. This report is therefore something of a snapshot of where the work stands currently. Over the next six months or so, the task force will build on the results reported herein and devise the development strategies that will permit development to be done, tests to be conducted, and standards to be written that will begin implementing the future envisioned herein. The work products required will be described and assigned to follow-on activities for their completion. A further report is likely upon completion of that step.

This report is divided into an Executive Summary, an Introduction, and three Chapters – on Compression, Physical Links and File Formats, and Wrappers and Metadata, respectively. The chapters are followed by a series of Annexes. The chapters contain the major findings of and are the work of five 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 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 to take a longer term view into the future with the hope of directing its path. As co-chairmen, we are honored to have been given the responsibility to lead the exercise. We wish to thank all those involved, both through direct participation and through financial and travel support of the many participants, for their efforts. There have been many long days spent by a large number of people to produce this initial output. We believe the result has been worth the labor.

Horst Schachlbauer
Co-chairman for the
European Broadcasting Union (EBU)

S. Merrill Weiss
Co-chairman for the
Society of Motion Picture and Television Engineers (SMPTE

Executive Summary

The expected convergence of television, computers, and communications is occurring. It is impacting television production, post production, and distribution through the appearance in television systems of video and audio compression methods, server technology, and digital networking. These technological changes offer the potential of huge benefits in reduced cost, improved operating efficiencies and creativity, and increased marketability of material. They also pose the threats of confusion, complexity, reduced technical performance, and increased costs if not properly managed. They clearly will radically change the way in which television is produced in the future.

In this context, the Society of Motion Picture and Television Engineers (SMPTE) and the European Broadcasting Union (EBU), joined by members of the Association of Radio Industries and Businesses of Japan (ARIB), formed the Task Force for Harmonized Standards for the Exchange of Television Program Material as Bit Streams. The Task Force drew the participation of about 80 experts from around the world who, meeting on five occasions in Europe and the United States, produced this report to guide the industry in its decisions regarding specific implementations of the technology. The report is also meant to steer the future development of standards intended to maximize the benefits and minimize the detriments of implementing such systems.

The Task Force was charged with developing a set of user requirements against which proposed solutions can be measured. In the few meetings available to deal with a subject having myriad implications and applications, the Task Force achieved a remarkable level of understanding of the problems to be solved and in setting the future direction in a number of areas that will facilitate the implementation of digital television systems based on bit streams. Although it was not able to resolve all of the issues that came before it, it nonetheless succeeded in providing significant mechanisms for rationalizing the differences likely to occur in practice.

In carrying out its work, the Task Force divided the effort into five fundamental areas of investigation: compression, physical link and transport layers for networks, wrappers and file formats, metadata, and file transfer protocols. Each was assigned to a sub-group that produced a portion of the report. In some cases, the sub-groups found that their areas of responsibility were so linked to one another that they did their work together and produced a unified output. This has implications for the associated technologies and standards that will derive from this work.

The compression technology used for processing video and audio is at the core of the changes that are occurring. It permits the digital transmission of program material through bandwidths that hitherto could not support digital signals. It permits the storage of that material in servers that make access to the material virtually instant and that allow simultaneous use by multiple users. These all have the effect of improving the efficiency and reducing the cost of producing and post-producing the material. Looked at another way, the compression technology has the effect of dramatically increasing the quality that can be achieved in transmission or storage confined to a particular bandwidth or data space.

The choice of appropriate compression schemes has significant impact on the overall cost / performance balance within production, post production, and broadcast operations as it will affect quality, storage / transmission efficiency, latency, editing / switching of the compressed stream as well as error resiliency.

The participants attempted to offer guidance for the long-term integration of compression into program production. The discussions revealed contrasting views between members on how best to achieve this. The issue is moot in a sense because a number of manufacturers have already introduced compression schemes to the market which target specific areas of applications, using different data rates as well as incompatible compression algorithms. The functional interoperability and exchange of video data between equipment employing different and incompatible compression algorithms is therefore currently achieved by transcoding to the high bit rate level defined in ITU-R BT.601.

The participants acknowledged that the rate of development of technology is moving at a fast rate and that, within the wide gamut of individual operational requirements, only a common methodology in handling different compression schemes can lead to the minimization of difficulty when interchanging program material and will retain predictability of the technical quality of the released output product. A number of key user requirements to achieve these goals were identified. A degree of optimism is necessary that these will be implemented. Users and manufacturers were invited to maintain an ongoing dialogue to reconcile conflicting requirements and to achieve harmonization in the area of compression.

The physical link and transport layers provide mechanisms that allow moving compressed video and audio streams and files from one place to another. Wrappers and file formats are used to contain program elements and other data in ways that it can be stored, retrieved, and transmitted as identifiable objects. Metadata ("bits about the bits") is descriptive and supplementary data that can be used either to interpret other data that carries program elements or in conjunction with those program elements. File transfer protocols are methods for transferring program elements and metadata, contained in wrappers and/or files, through physical and link layers by means of specific transport mechanisms to other locations.

With regard to the physical, link, and transport layers, the requirements were identified in terms of a number of quality of service (QoS) classes that were examined in relationship to specific applications. From this study, there is general agreement that standards are required to provide means for carrying program material as bit streams over two types of physical interfaces – direct connections between individual equipment units and networks where common interconnections are shared between multiple equipment units. With regard to direct connections, the standard needed is one that permits continued use of the existing serial digital interface (SDI) as currently embodied in international standards. For networks, decisions must be made on a limited number of physical means of interconnection together with the protocols that enable their use. This will permit the development of interoperable equipment, which does not require translations and gateways to communicate bit streams or files from one unit to another. While the activities to date have been prioritized on local area networks, it is acknowledged that further work is required on wide area networks.

The purposes of a Wrapper (sometimes called a "container") are to gather together program material (including audio, video, graphics, etc.), called **Essence,** and related information, called **Metadata;** 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. Essence and Metadata together form the **Content** of the Wrapper.

The various kinds of information that go into a Wrapper have been defined. The task force considered a range of applications where a Wrapper may be useful, ranging from capture to editing to distribution. Although users prefer a single solution, it was realized that a single Wrapper format will not satisfy all applications, although the number of distinct formats might be limited to just two. It is especially important for interchange that the different Wrapper formats are defined so that they are compatible. Detailed analysis of the usage profiles for Wrappers must be completed during the development of the formats.

Metadata was identified as a major new class of enablers of systems using bit streams for program material exchange. Metadata is a generic term for all sorts of data captured that relates in one way or another to program material. It ranges from time code and details of technical conditions when material was created to the scripts used, the publicity materials created, and descriptions of shooting locations. It can include standardized descriptive data to help in locating the material through various database entries. This can aid in repurposing of material, thereby significantly increasing its value. A core set of **Mandatory** Metadata items are required to guarantee basic interchange; different core sets will be defined for different uses (for example news acquisition, drama editing, commercial playout). Other Metadata values are optional, and default values may be defined. Metadata must be extensible and be made compatible by registering with a single registration authority.

Specific requirements of Wrappers are defined in the Wrappers section, with additional discussion contained in Annexes D1 through D5

A great deal has been achieved in the initial efforts to unify technology surrounding the exchange of program material as bit streams. Users and manufacturers alike agree, however, that this is just the beginning, and a similar endeavor should be undertaken for the next step if the organizations that handle standards on a routine basis are to do so in a coordinated way.

0 Introduction

The vast increase in content throughput caused by the advent of multi-channel program release together with the ever increasing pace of television program production mandate the exclusive use of digital production technology. These trends require the development of improved production tool functionality and the application of optimized workflow-concepts for the total production chain.

Manufacturers of professional television equipment as well as the computer industry are currently developing radically new system concepts to meet these requirements. These pivotal concepts include:

1. program data transport in the form of compressed bit streams,
2. non-real-time data transfer,
3. simultaneous, multi-user access to random program segments stored on servers,
4. inter-networking on open platforms of all production tools within the post-processing chain, and
5. hierarchical storage concepts based on tape, disk, and solid state media.

In this context, questions of end-to-end interoperability as well as technical quality have to be considered.

Functional interoperability and exchange of video material between equipment of different manufacturers, employing different and incompatible compression algorithms, currently can be achieved by decoding and re-encoding. The accumulation of compression algorithm artifacts within the long chain of production, post production, and distribution can introduce unpredictable, detrimental effects to the quality of the final output product. Concern about this topic has been expressed at all levels up to the International Telecommunications Union (ITU) and various government agencies.

It has been proposed that these detrimental effects be mitigated by limiting the number of generations of decoding and re-encoding and by employing compression algorithms that provide adequate post-production headroom for the end-to-end production chain. With few exceptions, the processing and editing of compressed program material will require decoding to the high bit rates of uncompressed rasters and baseband audio. Editing material based on different compression methods and data rates clearly demands the development of “agile” decoders, with special attention to codec symmetry and delay (or latency). The use of compression schemes which allow users to retain the potential for high quality post production throughout the production process can alleviate the quality problem but may lead to penalties in the area of storage efficiency and network bandwidth sharing.

Metadata is a major new class of enablers of systems using bit streams for program material exchange. Metadata is a generic term for all sorts of data captured that relates in one way or another to program material. It ranges from time code and details of technical conditions when material was created, to the scripts used, the publicity materials created, and descriptions of shooting locations. It can 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 program content and its associated metadata in ways that it can most easily be transferred and most beneficially be used. This means that the metadata may need to be accessible from the outside of containers so that the contents 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.

With regard to **networks and program file transfers**, users require the ability to exchange audio, video, and associated data easily and reliably between different systems. Agreed methods to move content within the production chain are therefore essential, will provide a stable basis for user choice, and will encourage a variety of solutions without putting limits on innovation in product design.

In general, users in the analog world have not had to deal with the complexities and characteristics of digital public carrier distribution systems and of networking through them. In an all-digital environment, the selection of one or more compression schemes and their associated bit rates should take into account the desired level of technical quality to be retained in the final output product of a given application. This will minimize cascading errors, fully exploit the bandwidth of the interconnection, and place economic utilization of any interconnecting networks and storage devices under user control.

In this report, the Task Force seeks to provide users and manufacturers alike with a document that addresses these issues in a complete and informative manner.

0.1 Open Standards

Users are asking that the future of digital video be based on the research and adoption of standards, such as those of the International Telecommunications Union (ITU), the Society of Motion Picture and Television Engineers (SMPTE), and the European Broadcasting Union (EBU). It may be useful to consider standards as providing “90% solutions.” Users and user groups may then need to specify “profiles” that adjust basic standards to meet the last 10% of functionality that broadly based commercial systems were never designed to meet. Lastly, users may have to specify "recommended practices" and/or “engineering guidelines” to facilitate final and specific technical interoperability. An example of the need for clearly defined recommended practices is the case where standards may broadly define a capability that maximizes flexibility but does not guarantee interoperability. By carefully selecting “nominal” values from the ranges of choices within a standard, users can better shape interoperability for their classes of applications.

As a general design philosophy, user organizations should attempt to choose Open Standards to which all qualified vendors can design and field systems. The EBU has recently published a Statement on Open Standards which emphasizes this requirement. By selection of international standards wherever possible, global competition can be maintained, providing all international players with opportunities to contribute their technologies to common systems and data exchange.

0.2 Scope of Activities

Users are well aware that the technology to treat television programs as bit streams is developing at a rapid pace. They also understand that within the wide gamut of system proposals advocated today, only the careful matching of technical options offered for system components involved in the total production chain will deliver the desired results. The provision of a single option, or at least a transparent gateway between different options within each subsystem is therefore of vital importance to users.

The Task Force has identified the following elements as critical for both interoperability between systems of different vendors and the overall performance of program transfers based on bit streams:

The digital encoding (compression) format.

Wrappers and file formats that are used to contain program elements and other data in ways so that they can be stored, retrieved, and transmitted as identifiable objects.

Metadata – a major new class of enablers of systems using bit streams for program material exchange. Metadata is a generic term for all sorts of data that relates to program material but is not directly content. It can be used to locate and identify content, help in its use and interpretation, and support its exchange between systems of differing characteristics.

Interfaces suitable for program data transfers.

The physical link and transport layers that provide mechanisms that allow the movement of compressed video and audio streams and files from one place to another.

These components have been carefully studied, and conclusions about them in this report are contained within individual chapters on the subjects of:

Compression Issues (the digital encoding format),

Wrappers and Metadata (methods for collecting and describing program content and other elements)**,**

File Management, Transfer Protocols, and Physical Connections (the combination of data transport mechanisms and interfaces).

Each of these chapters is largely independent of the others, although the topic of compression underlies the whole document.

Note that it is not the intention of the Task Force to include all possible compression formats but to limit the range initially to:

- 525 line and 625 line interlaced systems,
- Video Compression schemes using bit rates of 50 Mbit/s or less (excluding audio), and
- Coding resolution of one frame or more.

The considerations discussed, however, are broadly applicable to advanced television systems, such as those being implemented in North America.

1 Compression Issues

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 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 in addition information that is not redundant but is considered less important. Reconstruction from the compressed bitstream thus leads to the addition of distortions or "artifacts." Compression for video and audio is therefore not normally lossless.

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

This chapter considers a wide range of compression characteristics. Compression of video and audio allows functionality not viable with uncompressed processing. Through reduction of the number of bits required to represent given program content, it makes economical the support of applications such as storage of material, transmission of a multiplicity of program 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 chapter provides users with information about compression characteristics to assist in making judgments about appropriate solutions. It further recommends approaches to be taken to facilitate interoperation to the extent possible between systems within a single family of compression techniques and between families of compression methods.

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

The chapter is subdivided into six parts and one associated Annex:

Section 1.2 introduces the topic of compression through an examination of the relationship between its various aspects and image quality. This topic is very large and is developing as new techniques are devised.

Section 1.3 considers the need for several quality levels within a single facility.

Section 1.4 considers some of the operational aspects of compression.

Section 1.5 discusses the key attributes of storage media when used in conjunction with compression.

Section 1.6 deals briefly with the complex topic of interoperability.

Section 1.7 looks at the techniques of compliance testing and the benefits they produce.

Section 1.8 lists the recommendations for the application of compression within the television production and post production chain.

Annex C treats specific applications and their uses of compression techniques.

1.1 Image Quality

Selection of compression system parameters has a significant impact on overall image quality. These compression parameter choices must be optimized to preserve 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 compression system parameters, interaction between parameter choices must also be considered. Finally, special operational issues such as editing the bit-stream or splicing new content into an incoming bit-stream should be considered.

1.1.1 Coding Method

The coding method is the most fundamental of compression choices. There are four main compression methods used in the television production and distribution chain: MPEG-2 Main Profile at Main Level (MP@ML), MPEG-2 4:2:2 Profile at Main Level (4:2:2@ML), Motion JPEG, and DV. All of these coding methods are based on the Discrete Cosine Transform (DCT). All of these coding methods use normalization and quantization of the transform coefficients, followed by variable length coding.

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

1.1.2 Sampling Structure

MPEG, Motion JPEG, and DV can all be used with the 4:2:2 pixel matrix of ITU-R BT.601. MPEG and Motion 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 subsampled from the original full-bandwidth (4:4:4) signal. The pixel matrix can be further subsampled to reduce the signal data, with 4:2:2 sampling normally used for interchange between systems.

4:2:2 systems, such as the MPEG-2 4:2:2 Profile, 4:2:2 Motion JPEG, and the DV 4:2:2 system (which operates at 50 Mbits/s), all use half the number of color difference samples per line compared to the number used in the luminance channel. 4:2:2 provides half the horizontal bandwidth in the color difference channels compared to the luminance bandwidth and full vertical bandwidth.

4:1:1 systems such as DV 525 use one quarter the number of color difference samples per line compared to the number used in the luminance channel. 4:1:1 reduces the color difference horizontal bandwidth to one quarter that of the luminance channel while maintaining 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 passbands 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[1] and MPEG-2 Main Profile, use half the number of color difference samples horizontally and half the number of color difference samples vertically compared to the numberused in the luminance channel. 4:2:0 therefore retains the same color difference horizontal bandwidth as 4:2:2 (i.e., half that of the luminance channel) but reduces the color 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 passband, thereby precluding transparent translation to the other

---

[1] 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.

coding forms. Consequently, systems that use 4:2:0 sampling with intermediate processing generally will not 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, intermixing different subsampled structures impacts 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), cascading 4:1:1 and 4:2:0 will generally yield less than 4:1:0 performance.

1.1.3 Compression Pre-processing

Video compression systems have inherent limitations in their ability to compress images into finite bandwidth or storage space. The compression systems rely on 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 artifacts in the picture. In these cases, it may be preferable to reduce the complexity of the image through other methods before compression processing. These methods are called pre-processing, and include filtering and noise reduction.

When noise is present in the input signal, the compression system must expend some bits encoding the noise, 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 coding efficiency. Even in compression systems which do not use motion estimation and compensation, noise adds substantial energy in high frequency components of the DCT which might otherwise be zero. This not only wastes bits on extraneous DCT components, but degrades 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 system, 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.

1.1.4 Data Rate

The MPEG-2 4:2:2 Profile at Main Level defines data rates up to 50 Mbits/s. Motion JPEG 4:2:2 equipment typically operates at data rates up to 50 Mbits/s. DV 4:1:1 and DV 4:2:0 operate at 25 Mbits/s. DV 4:2:2 operating at 50 Mbits/s is undergoing standardization. MPEG-2 Main Profile at Main Level is defined at data rates up to 15 Mbits/s.

Selection of the data rate for MPEG-2 4:2:2@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 Mbits/s, can yield comparable image quality. MPEG-2 4:2:2@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 Group of Pictures below.)

1.1.5 Group of Pictures

There are three fundamental ways in which to code or compress an image. The most basic is coding 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). The second way in which to code an image 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. 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 Mbits/s, the use of long GOPs (e.g. IBBP) may prove useful, while at 50 Mbits/s shorter GOPs can provide the required quality.

Besides affecting image quality, the choice of GOP structure also impacts 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.

1.1.6 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 variable bit rate or constant bit rate modes; DV operates only with constant bit rate. In practice, even those systems commonly believed to be constant data rate have data rate variation, but over shorter periods of time. Another way to characterize the compression systems is to compare constant quality with constant data rate systems.

Constant quality systems attempt to maintain a uniform picture quality by adjusting 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 non-real-time transfer of images. Constant quality operation is useful for disk recording and some tape recording systems such as tape streamers.

Constant 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.

1.1.7 Editing

Consideration of compression parameters relating to editing falls 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 bit-stream splicing. Bit-stream splicing can be done between two bit-streams 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.

1.1.8 Concatenated Compression

To the 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 means that additional image information is lost. Further, any changes which result in different buffer management will result in different quantization.

In the case of MPEG coding, any change in the alignment of 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 in 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 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. Further, some users would welcome image processing in the compressed domain as a potential means to achieving faster-than-real-time image processing.

1.2 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 cost of equipment and of operation 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 program 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 program distribution, and local program production. The quality level would also be used to archive program 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 to quality. This quality level would not normally support subsequent processing but might be used for program presentation or mass storage for rapid-access browsing. The lower compression quality would not generally be used to archive program 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 Annex C: Applications.

1.3 Operational Considerations

Systems of all compression performance levels must be fully functional in their intended applications. Equipment employing compression should function and operate in the same or better manner as similar analog 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 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 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 inter-operate 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 synchronization of audio, video, and metadata must be considered. Efficient video coding comes at the cost of codec delay, so balance must be achieved between 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.

Compressed signal bit streams 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 bit streams 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 analog vertical interval 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.

1.4 Storage

Where a compressed video bit stream 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.

1.4.1 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 at non-real time, the network should also allow the transfer of files at both faster and slower than true 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 transfer specialized bit streams 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.

1.4.2 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. Since there will be different compressed video bit rates depending on the application, any network connection to the device must 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 variable bit rate compressed video streams. This will require an interface which 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.

1.4.3 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 disc based video server additionally has the capability of accepting multiple I/O channels. Further signaling may be necessary to ensure that both channel bandwidth and the number of channels can be adequately signaled to the system.

1.4.4 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.

1.4.5 VTR Emulation

Where a storage device using 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 and slow motion for the purpose of identifying editing points, and 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 impose no conversion loss.

1.5 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 program transfers, concatenation, cascading, encoding and decoding quality, and compliance testing. Program exchange requires interoperability at three levels: the physical level, the protocols used, and the compression characteristics. This chapter considers only compression while other chapters address physical links and protocols.

Considering program 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 for this Task Force is to categorize methods of interoperation, explore their characteristics and relate them to various applications, and develop possible constraints on device and system characteristics necessary to ensure predictable levels of performance sought by users for specific applications.

1.6 Compliance Testing

Interoperability between compressed video products is essential to successful implementation of systems using compression. Although inter-operation is possible via ITU-R BT.601, it is desirable to have inter-operation at the compressed level to minimize concatenation losses. Compressed inter-operation 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 toward 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 bit-streams. Decoders can be tested to verify that a range of compliant bit-streams can be properly decoded. Applications can be tested to verify that the characteristics of a given bit-stream 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 compliance tests is more involved than applying those tests, so tests employed by manufacturers might be identical to tests employed by users.

In the case of MPEG-2, compliance testing focuses on the bit-stream 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 transport stream tests, program stream tests, timing accuracy tests, video bit-stream tests, and audio bit-stream tests. The MPEG-2 video bit-stream tests include a number of tests specific to the 4:2:2 Profile at Main Level. Other tests will need to be developed.

1.7 Compression Issues – Recommendations

The deliberations that went into preparation of this chapter led to the following recommendations for the application of compression within television production, contribution and distribution:

1. Compression algorithms and transport schemes should be based on Open Standards. This implies availability of the intellectual property 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.
2. The number of compression methods and parameters should be minimized for each uniquely defined application in order to maximize compatibility and interoperability.
3. Compliance testing methods should be available for those building 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 manufacturer and user needs.
4. A single compression scheme used with different compression parameters throughout the chain should be decodable by a single decoder.
5. To support use of more than one compression family, the development of a common (“agile”) decoder is desirable.
6. Integration of video compression into more complex systems must be via standardized interfaces. Translating through ITU-R BT.601, i.e., decoding and re-encoding, is the default method of concatenating video signals compressed using different techniques and/or parameters, although other methods are possible.
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.
8. Issues related to interoperability must be further explored and standards developed to allow predictable levels of performance to be achieved in the implementation of specific applications.
9. Bit streams 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.
10. Compressed bit streams are very sensitive to errors, and therefore it is recommended that appropriate channel coding methods and error protection be employed where necessary.
11. Compression systems should be designed so that, in normal operation, signal timing relationships (e.g., audio/video lip sync) and synchronization presented at encoder inputs are reproduced at decoder outputs.
12. Signal delays through compression processing (encoding/decoding) must be limited to durations that are practical for specific applications, e.g., live interview situations.
13. Provision should be made for selected analog vertical interval 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.
14. The compression scheme chosen for devices that mimic VTRs should allow for the reproduction of pictures in shuttle mode for identifying content and of pictures in jog and slow motion modes for selecting edit points.
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.
16. Storage devices should allow recording and playing back of streams and files as data rather than decoding to baseband for recording and re-encoding upon playback.

17. The compression strategy chosen for standard television should be extensible to high definition applications to allow for commonality in the transitional phase.

2 Wrappers and Metadata

The Joint EBU / SMPTE Task Force has been considering the formatting of collections of audio-visual program material and related information for exchange within and between studios and other centers which process or store that information. The goal was to establish a maximum degree of interoperability independent of the encoding format for the audio-visual signal.

Collections of information include both streams of program material (to be transported by means under discussion elsewhere within the TFHS), and files of program material and related information, to be held in storage systems and manipulated by computer-based equipment.

The particular role of the TFHS Sub-Group on Wrappers is to determine the user requirements for the access to and manipulation of these collections of information. These requirements may influence the characteristics of the Wrappers used to group and label the information.

The particular role of the TFHS Sub-Group on Metadata is to determine the user requirements for the related information within these collections, including the types of information appropriate for each application, its formatting, its relation to the program material, and the relative importance of each type of information. These requirements may further influence the characteristics of the Wrappers used to group and label the information.

The results of these discussions are contained in this Chapter, and consist of:

- some Terminology
- a set of Requirements, grouped into topics

Background information and additional discussion of topics is contained in Annexes D1 through D5.

2.1 Purpose of Wrappers

The fundamental purposes of a Wrapper are to gather program material and related information together (both by inclusion and by reference to material stored elsewhere), 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.

2.2 Terminology – What's in a Wrapper

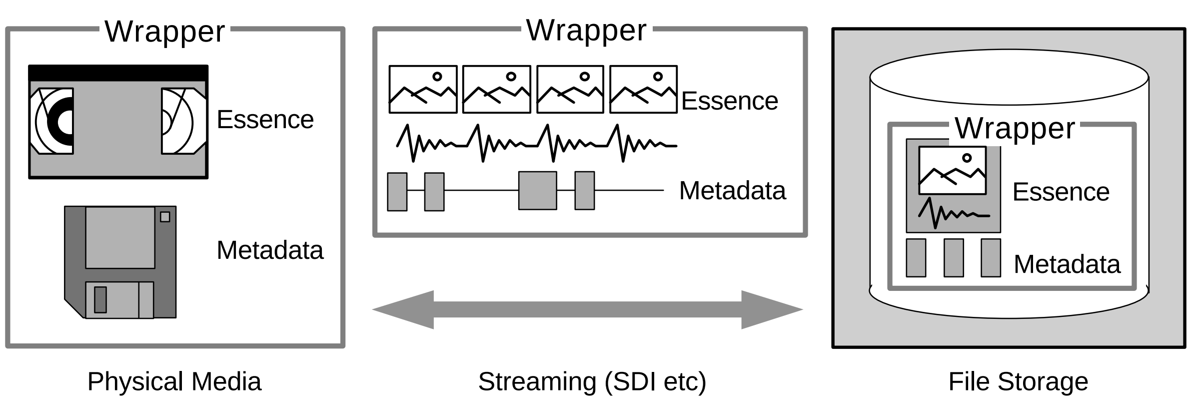

Figure 1: Schematic View of Wrappers in Use

Program material and related information of any variety is called **Content**. The parts of Content which directly represent program material (such as signal samples) are called **Essence** (see section 0 below); the parts which describe the Essence and other aspects of the material are called **Metadata** (see section 0 below).

Wrappers are intended to be used to link physical media together, for streaming of Content across interconnects, and to store Content in file systems and on servers.

This and other terminology is discussed in this section.

2.2.1 Content Structure

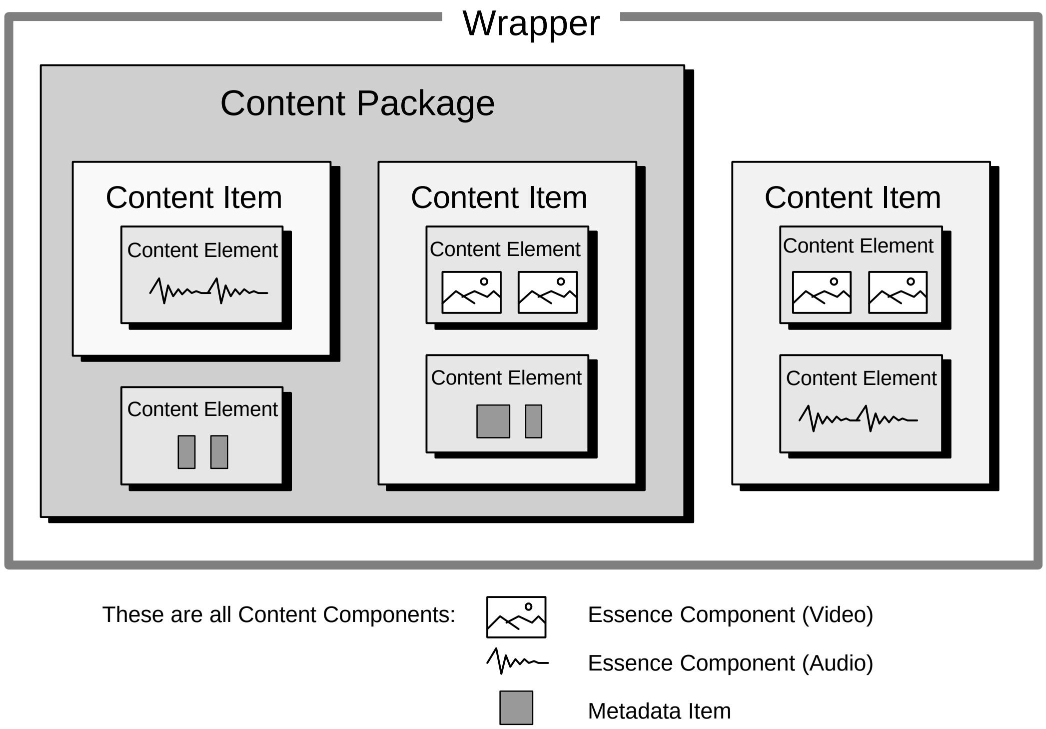

Figure 2: Content Structure and Content Components

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 Format Metadata describing the sample structure plus the Descriptive Metadata identifying the origin of the signal.

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.

Types of Essence include Video, Audio, Graphics, Still Images, Text, and other sensor 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 the component parts (Content Elements) together – 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 the component parts (Content Items and Content Elements) together – for example, a program composed of video plus audio plus subtitles plus description.

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 program 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.

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.

2.2.2 Essence

Program material itself is referred to as **Essence**. Essence includes all data that represents pictures, sound and text; types of Essence include Video, Audio, Graphics, Still Images, Text, and other sensor data as needed by each application. Essence may be encoded or compressed in whatever way is appropriate. and 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 Format Metadata (see below).

Essence typically has the characteristic of a stream, with sequential access whether stored on a file device or 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.

2.2.3 Metadata

Other information 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 requirements and behavior, Metadata is divided into several **categories**, depending upon its purpose, including at least the following:

- **Format -** any information necessary to decode the Essence.
  Examples: video formats, audio formats, numbers of audio channels, aspect ratio, pan & scan etc.
- **Descriptive** – all information used in the cataloguing, search & retrieval, and administration of Content.
  Examples: unique material identifiers (UIDs), labels, author, location, date & time, geospatial (information related to the position of the source), copyright information, access rights information, modification time stamps, version information, transaction records, etc.
- **Association –** any information necessary to achieve synchronization between different Content Components, and to achieve appropriate interleaving of the components.
- **Composition** – information required 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 derivation of the Content.
  Examples: edit decision list, titling information, zoom lens positioning (for virtual studio use), color correction parameters, etc.
- **Other** – anything not included above.

Examples: Scripts. Definitions of the names and formats of other Metadata.

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

2.2.4 Metadata Characteristics

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.

Other such **characteristics** of items of Metadata may be identified.

2.2.5 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.

2.3 General 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 program 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 Content Elements, Content Items, or Content Packages defined above.

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

As well as directly including Essence and Metadata, Wrappers may contain indirect references to either. This is discussed further below and also in D2 – Wrapper Referencing, which also includes a list of some different varieties of Wrapper.

2.4 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 program material may appear in and be converted between any or all of the formats during its lifetime.

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

A wide range of potential activities were listed (see D4 – Notes), which were then grouped into the following categories:

- Pre Production
- Production and Acquisition
- Post Production
- Distribution and Storage

- Emission and Transmission
- Archive

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

- Unwrapped (for example, a Videotape)
- Streaming (for example, on an SDI channel or as an MPEG stream)
- Rich (for example, a database together with signal storage, or an Edit Decision List)

As well as being used within each process, these three forms are all used as interfaces between processes.

There is therefore a requirement for at least two Wrapper formats (Streaming and Rich) in addition to the continued use of unwrapped Content.

2.5 Metadata Requirements and Metadata Sets

Recognizing the breadth of varieties of Metadata, the naming scheme used for varieties of Metadata should be hierarchical. The hierarchy of varieties (categories, sub-categories and so on), the actual names and the definitions should be registered by a single independent registration authority such as SMPTE. The names must have a plain text representation.

Each application will employ different combinations and varieties of Metadata. There is therefore a requirement for Metadata Sets to provide a guideline as to what combination to employ in a particular application. The Metadata Sets should be developed as part of the recommended Metadata standardization process. A preliminary list of the required Metadata Sets is contained in D4 – Notes.

To maximize compatibility, there is a strong preference for Metadata to have a defined representation in plain text form, using an international character set such as ISO 646.

It is recognized that some varieties of Metadata are inherently not representable as plain text. Other varieties carry information in local language, and must be represented using a regionalized character set. Metadata which names, defines and describes other Metadata must be represented in the international character set.

Within each Metadata Usage Profile, a core set of **Mandatory** Metadata items must be provided with each Content structure (Content Component, Content Element, Content Item, or Content Package). This small core set provides for the basic management of the Content structure.

A further set of **Essential** Metadata items must either be provided, or else a sensible default value can be automatically inferred. These items are typically within the class of Descriptive Metadata.

Within certain Profiles, an additional arbitrary assortment of **Optional** Metadata items may be required to be carried.

2.6 Wrapper Size

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

This may require that some Metadata is repeated in each of the parts; alternatively, Metadata needed by each part may be held in a Wrapper of its own, whence the data will be obtained by Referencing

In this situation, Composition Metadata will be required, to describe the relationship between the parts.

2.7 Platform Neutrality

Wrapper formats must be designed to be "Platform Neutral", so that Wrappers may be read by any machine with equal facility (although perhaps with different performance), no matter what machine was used to originally create the wrapper.

Typical considerations are byte ordering, and the organization of sample structures in customized word formats. The need for platform neutrality does not preclude creating a Wrapper in an optimal format for a particular machine.

Annex D1 **-** Platform Neutrality gives more information on this topic.

2.8 Interleaving

It is likely that information in Wrappers, particularly Essence Components, must be interleaved in various ways for optimization of storage, retrieval, presentation, and transmission.

Existing transport layers such as SDI or MPEG-2 Transport Stream may dictate the interleaving scheme. Wrapper formats must therefore permit conversion between interleaving schemes.

The Wrapper formats should insulate the user from the specific interleaving scheme which is used, so that both Essence and Metadata may be manipulated with equal simplicity (although perhaps with different performance) no matter how they are interleaved.

2.9 Unique Identifiers

Content must be identified by some species of Unique Identifier. Unique Identifiers are classified as Mandatory Descriptive Metadata.

Unique Identifiers serve to identify the Content, irrespective of the physical location of the content, and independent of whether the Content is the original or a copy. This is different from the function of the Filename (discussed by the Sub-Group on File Transfer Methods in Chapter 3).

When Content is duplicated, it must retain the same Unique Identifier; however, whenever processing is performed on the copy, or when a copy is made of only a subsection of the Content, a new Unique Identifier must be assigned. In some Usage Profiles, traceability to the original Unique Identifier is required (see also "History" below).

In some other cases, identification of each specific instance or copy of the Content is required in addition to the Unique Identifier. This will probably involve linkage between Unique Identifiers and Filenames.

2.10 Immutability and Generation Numbering

In most cases, it is not known how many References have been made to 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.

2.11 References

The Wrapper formats must allow Metadata to refer to points and regions within other wrappers, or within the same wrapper, or within external material by means of indexing. Indexing is discussed in the next section.

This basic capability is required for many purposes, including:

- Creation of associations between Essence and Metadata
- Inclusion of external material (for example, from videotape or camera) within programmes
- Description of editing operations within Composition Metadata

Annex D2 – Wrapper Referencing gives more information on this topic.

Although the use of References can improve the efficiency of systems by decreasing the use of copy operations, this may be offset by increases in the complexity of systems to manage the proliferation of separate Wrappers.

It is expected that both techniques will be required to accommodate different operational requirements; but Wrapper Profiles may indicate a preference for one or the other method.

2.12 Indexing

For the purpose of implementing References, Wrapper formats must allow indexing of points and regions within a Wrapper in either of the following ways:

- Systematic indexing (for example, timecode, sub-frame or sample index)
- Specific indexing (for example, named cue points, key frames)

Note that there are many issues of consistency of indexing when dealing with the diverse sample rates and synchronization methods in present televisions systems – such as the relationship between audio samples and video frames, and the relationship between film frames and video frames. These relationships must be accommodated by the Indexing method, in combination with Association Metadata.

Annex D2 – Wrapper Referencing gives more information on this topic.

2.13 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 includes any editing history or signal transformation data.
- Transaction logging, allowing the steps taken to produce the current version of the Content from it's source material to be traced but not necessarily reversed. This includes version and source information.

2.14 Access Control

Features for operational security may be included in a Wrapper format to prevent unauthorized access to Content.

Operational security requires the use of a log-in procedure (or decryption key) and supports the user as an individual or a member of a group. Files may be protected by their location and by time limits. Access rights may be provided at several levels. Encryption is the only feasible mechanism of protecting components within a Wrapper.

Annex D3 - Access Control and Copyright gives more information on this topic. This subject is also discussed in the section on File Transfer Methods.

2.15 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 Wrappers.

Metadata in support of these transactions may be included within Wrappers.

Other aspects of the operation of transaction systems are outside the scope of this section, and are discussed in the section on File Transfer Methods.

2.16 Property Rights

Metadata recording the ownership of Content and the history of ownership may be stored in the wrapper in order to facilitate the establishment and preservation of copyright.

Annex D3 - Access Control and Copyright gives more information on this topic. This subject is also discussed in the section on File Transfer Methods.

2.17 Asset Management

Wrapper formats must support indirect references to content – that is, references to objects which are themselves references to Content. This is a basic requirement used to support all manner of different material management systems.

Effective asset management is required by the users. This may be provided by either manual or automatic methods as appropriate. Wrapper referencing of Content can work most effectively where automation tools are provided for storage administration tasks and to ensure cohesive referencing when files are moved or copied.

Specification of material management systems is outside the scope of this section, and this is discussed in the section on File Transfer Methods.

2.18 Application Programming Interface (API)

Specific Wrapper Profiles, particularly those emphasizing richness of data description may require a standard application programming interface (API) to simplify the process of reading and writing the Wrapper format.

2.19 Compatibility and Conversion

Wrappers 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 program 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 directly used but can be translated to or from the Wrapper with some processing, and the conversion can be fully reversed.
- Lossy Conversion is possible when Metadata or Essence cannot be directly used but can be translated to or from the Wrapper with some processing, and some loss of meaning or quality, and the conversion cannot be fully reversed.

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

2.20 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 addition of new Essence and Metadata types,
- By 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 with an efficient but complete documentation process, 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.

Other issues related to Version management are discussed in D4 – Notes.

2.21 Wrappers and Metadata – Recommendations

The specific recommendations on Wrappers and Metadata are as follows:

- The development of an extensible hierarchical classification of Metadata varieties, including the notion of Metadata Sets appropriate to particular uses.
- The establishment of a single registry of Metadata identifiers and definitions.
- The standardization of a single generic Wrapper format for streaming of Metadata, which can be mapped onto existing and emerging signal transport layers.
- The standardization of a single generic Wrapper format for applications requiring arbitrary richness of Content of all types, including Metadata and Essence. This must be highly compatible with the streaming format described above.
- The standardization of a single format for a “unique identifier” would also be of assistance; however it is recognized that multiple formats are already in use. As a minimum therefore, it should be possible to register existing and new unique identifier formats within the Metadata registry referred to above.

3 File Management, Transfer Protocols, and Physical Connections

The transfer of program content can take place either as a continuous stream (as with existing analog playout) or in the form of a discontinuous file transfer in packetized form.

The transfer requires a number of operations, including the identification of the file or stream to be transferred, the destination for the transfer, establishment of a connection and the control of the transfer. In addition, file transfers require a logical connection, the identification of the file at its destination and notification to the user of the completion of the transfer.

These operations involve file management, transfer protocols and physical interfaces, and apply to both local-area and wide-area networks.

3.1 File Management and File Systems

It is necessary to specify components on which content management or other applications may be built. File systems for audio/video servers are similar in concept to normal computer file systems but they do have some unusual requirements in terms of file size, real-time access and reliability. Although some of these issues have been addressed in large-scale and real-time systems, the facilities are not available on a wide range of broadcast products. The primary requirement is to have a consistent feature set between broadcast file systems.

3.2 File System Attribute Requirements

The file system should support a hierarchical name space and have consistent file name syntax and semantics. The path and file naming must permit at least 256 characters.

The file system should support a read/write permission access mechanism, based on usernames, passwords and user groups. The chosen mechanism should also support default parameters for users that do not require access control. See Annex D section 3 for further details.

3.3 File System Commands

Commands are required in order to manipulate files and navigate the directory system. The local machine (the one the user is connected to) must respond to these commands. If remote machine access is provided then these same commands must be implemented on the remote machine.

The set of commands must at least permit:

listing of the contents of the current directory name space and sub-tree name space;

- display of the files and directory properties;
- copying, renaming and deletion of files;
- creation, rename and removal of directories.

3.4 Digital Data Transfers

At least two distinguishable types of data transfer are required:

- File Transfer. The File is any data wrapped with a known cover. See Annex D on Wrappers and Metadata.
- Streaming of raw data from a serving device to one or more receivers.

The two main methods to be supported are:

- Guaranteed-delivery “File Transfer”.

- Bounded-quality streaming transfer. Data may be streamed from a serving device at user-selectable payload rates for real-time or slower/faster than real-time.

Note: The term "Bounded quality" is used for a transfer method that is designed to move the payload from source to destination(s) but without the absolute certainty of true guaranteed delivery. Traditional analogue video is moved with bounded quality in a " content playback" method today. Also, payload data that is moved over the ITU-R Rec. BT.656 serial digital interface is moved in a bounded-quality way. Usually, bounded-quality links are used to transport streamed, real-time content. On the other hand, "guaranteed delivery" indicates that the entire payload will reach the destination without bit errors, barring a failure of the physical link.

3.5 Transport mechanisms

The transport mechanism is used to build the link between the data to be transported (i.e. the file format) and the physical layer. Specifications are required on the following transport mechanism functions (for details see Annex E):

- **Transfer modes:**
  In a mainstream television production environment three types of time relationship between the source data-clock and the received data-clock of a signal are necessary: synchronous, isochronous and asynchronous transmission in both bounded and guaranteed quality transmission.
- **Synchronization of associated data:**
  The bonding of audio and video to certain types of metadata has been identified as a vital issue and requires further investigation.
- **Transfer initiation and transfer phase parameters:**
  During the initiation or pre-roll phase of a transfer (independent from the transfer modes) certain parameters such as set-up-time of a process and the response time for a process need to be defined.
- **Quality of service (QoS) for the transfer:**
  Once the a transfer has commenced, a defined transmission quality must be guaranteed. QoS-parameters for controlling such a transmission are: bit-rate, delay, Bit-Error-Rate (BER) and jitter.
- **Transfer interaction management and error reporting:**
  The transfer interaction management such as establishment, maintenance and release of the data link, error and flow control usually runs in the background of the system and is not visible to the user. However all of these functions need to be specified to achieve interoperability and the system must report to the user non-recoverable exceptions and distortions if specified limits are exceeded.
- **Basic interworking protocols:**
  Two families of transport mechanisms have been identified for the intra-studio transfer mechanism: transport mechanisms based on the serial digital interface (ITU-R BT.656) and network-based transport mechanisms. These need to fully specified.
    - For data transfer via the serial digital interface (ITU-R BT.656) in packetized form, users require a single open transport mechanism. A standard for such a transport mechanism is currently under development in SMPTE and EBU.
    - For transport mechanisms based on networks, users require at least one open standardized baseline protocol such as, for example, FTP/ IP. However these protocols have not been developed to meet the mainstream television production requirements and optimized protocols which better meet the user requirements defined above need to be developed and standardized.

Interworking between the serial digital and network based transport mechanisms is a strong user requirement and open, standardized gateways need to be defined.

3.6 Physical and Link Layer Considerations

The choice of physical and link layer implementations are restricted by the choice of network, and this choice is itself dependent on system and application requirements such as the required QoS and bit rate.

Network interfaces must be chosen from the existing industry-standard interfaces. At the present time these are considered to be:

- IEEE 802 Ethernet (see Note 1)
- Fibre Channel ANSI X3230-1994 and related standards(see Note 1)
- ATM (see Note 1)
- IEEE 1394 (see Note 1)
- SMPTE [XXX] (the SMPTE Serial Data Transport compatible with SMPTE 259M)

Note 1: Not all classes of service, particularly real-time full bit-rate video transfers, are available at the present time

It is desirable that equipment network interfaces be available in a variety of forms to suit the industry-standard networks listed above and that an equipment interface be capable of upgrading to take advantage of higher network bandwidths as these become available.

Gateways between the different interface implementations are required (e.g. between Fibre Channel and ATM) to enable networks to be interconnected and for local area networks to be connected to those serving the wider area. This report is currently limited in scope to operation within a local area network environment; the wide area network context requires further study.

3.7 High Level Management Functions

These are functions or principles which must permeate all parts of the system and be extended or contracted in their reach automatically or semi-automatically when the physical extent of the system is expanded or contracted. They must provide::

- Data security: backup/archiving must provide opportunity for restoration when any content element is lost or corrupted
- Data shadowing: redundant copy of data being accessed must allow real-time hot-switching to redundant copy in event of sudden non-availability of main copy
- Automatic Backup: must run in the background and not require human vigilance nor affect functionality or security of foreground operations
- Localized recovery from local failure: recovery must not affect remainder of system (e.g. Ctl-Alt-Del is not an option for the ABC network)
- Robust operation: no concentrations of failure points; extent of system failure caused by failure of a system element must be proportional to the fraction of the system represented by the failed element - there must be no propagation or escalation effect (e.g. no centralized bandwidth allocation manager or domain name server); this applies equally to hardware and to software elements.
- Failure condition notification: must trigger automatic system notification and logging of onset and end of failure (e.g. assist in localization of failed element, even across heterogeneous operating systems)
- Distributed Resource Management: automatic calculation and provisioning of system resources needed for transfers, across physically-separated but logically-bound system entities (e.g. file size notification & comparison with remote free storage capacity)
- Dynamic resource detection (e.g. hot plugging): user must be able to add or remove system resources without rebooting entire system; when added, new resources must be fully functional and system must be fully aware of their presence (vice versa when removed) - e.g. switching a VTR in and out of a Remote

- Multi-operating system support: file names, properties, etc., must be interchangeable between different operating systems

3.8 Further work

To satisfy the requirement for interoperability between systems from different vendors, a number of open standards need to be defined:

- File formats for content transfer (see also Annex D)
- Link protocols for guaranteed and bounded-quality transfer over LAN and WAN applications
- Flow control
- Error control
- Networks and interfaces
- Payload and mapping documents (e.g. the mapping of content blocks into the packets of a transport mechanism)
- Gateways between networks, LAN and WAN and SDI-based transport mechanism
- Machine control interfaces (Note1)
- Platform independent file system management interfaces
- timecode-based scheduling for file system actions such as browsing

Note 1: Machine control commands need to be openly defined: a central, external control system must be able to control all system elements without the need for multiple customized drivers or command translators; (e.g. would support ES-LAN Server dialect)

Annex A – Abbreviations and Specialized terms

Throughout this document specialized terms are used and defined as and when they occur. In addition to these, many of the words and acronymic abbreviations 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 ones is given below.

-A-

A/D	Analog to Digital conversion
AAL (ATM adaptation layer)	The AAL translates digital voice, image, 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 and specific timing and delay requirements. (e.g., DS-3 circuit) AAL2 supports connection-oriented services needing variable bit rates. (e.g., certain video transmission schemes) AAL3/4 supports both connectionless and connection-oriented variable rate services. AAL5 supports connection-oriented variable bit rate data services. AKA: Simple and Efficient Adaptation Layer (SEAL)
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 computer 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)	ADPCM algorithms are compression algorithms that achieve bit rate reduction through the use of adaptive prediction and adaptive quantization.
Analog	A type of transmission in which a continuously variable signal encodes an infinite number of values for the information being sent. (Compare with "digital.")

Analog signal — A signal one of whose characteristic quantities follows continuously the variations of another physical quantity representing information.

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 insulate the application programmer from the minutiae of the implementation.

Asynchronous — The essential characteristic of time-scales or signals such that their corresponding significant instants do not necessarily occur at the same average rate.

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 (i.e., start/stop bits, flag characters) in the data stream itself. (Compare with "synchronous.")

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 image, 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 muxing/demuxing, header validation, payload-type identification, quality-of-service specification, prioritization, and flow control.

-B-

Bandwidth — A measure of capacity, usually, the capacity of a communications line to transmit voice, data, video, or image traffic through a network. Bandwidth is usually expressed in bits per second (bps), thousands of bits per second (kbps), millions of bits per second (Mbps), or billions of bits per second (Gbps).

BER — Bit error ratio (or rate)

Binary digit *(bit)* — A member selected from a binary set. Bit is an abbreviation for binary digit.

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)).
Broadcast (Messages)	Transmissions sent to all stations (or nodes, or devices) attached to the network.
Buffer	An area of storage that provides an uninterrupted flow of data between two computing devices.

-C-

CBO	Continuous bit-stream oriented (services)
CBR	Constant bit rate
CBR (Constant Bit Rate)	A type of traffic that requires a continuous, specific amount of bandwidth over the ATM network (e.g., digital information such as video and digitized voice).
CCITT	The Consultative Committee on International Telephony and Telegraphy, now the International Telecommunications Union (ITU), is an international organization that 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 cell relay 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, transmission channel	A means of unidirectional transmission of signals between two points.
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 plain old telephone call is a circuit-switched connection)
Clock	Equipment that provides a timing signal.
Codec	A combination of an encoder and a decoder operating in opposite directions of transmission in the same equipment.

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 information that is not redundant but is considered less important.
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 LANs 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.
CRC	Cyclic redundancy check
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-

DCT	Discrete cosine transform
DEMUX	Demultiplexer
Descrambler	A device that performs the complementary operation to that of a scrambler.
Digital	A type of transmission that encodes a discrete value (e.g., "0" or "1") for each unit of information being encoded. (Compare with "analog.")
Digital channel, 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, to support 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.
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

-E-

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.

-F-

Frame	Variable-length packet of data used by traditional LANs 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.
FSK	Frequency shift keying

-G-

Gbit/s (Gigabit per second)	A digital transmission speed of billions of bits per second.

-H-

Header Error Control (HEC)	An 8-bit Cyclic Redundancy Code (CRC) computed on all fields in an ATM header; capable of detecting single bit and certain multiple bit errors. HEC is used by the Physical Layer for cell delineation.

-I-

Interface	The common boundary between two associated systems.
Internet Protocol (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.

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.
ISDN	Integrated services digital network
Isochronous	A term used to describe signal timing techniques that require a uniform reference point (usually embedded in the data signal).
Jitter	Short-term non-cumulative variations of the significant instants of a digital signal from their ideal positions in time.
-K-	
kbit/s (kilobit per second)	A digital transmission speed of thousands of bits per second.
-L-	
LAN (Local Area Network)	A system consisting of computer and communications hardware and software connected by a common transmission medium, usually limited to a scope of a few miles.
LAN Emulation	The process of implementing enough of the MAC layer protocol of a LAN (i.e., 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 that it takes to process an input bit stream 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.
LSB	Least significant bit
-M-	
MAN	Metropolitan area network
Master clock	A clock that is used to control the frequency of other clocks.
Mbit/s (Megabit per second)	A digital transmission speed of millions of bits per second.
Megabits Per Second (Mbit/s)	A digital transmission speed of millions of bits per second.

Metadata	Data describing other data
MSB	Most significant bit
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.
Multipoint	A term used by network designers to describe network links that have many possible endpoints.
MUX	Multiplexer

-N-

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.
Octet	A group of eight binary digits or eight signal elements representing binary digits operated upon as an entity.
OSI	Open system interconnection

-P-

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.
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.
PDH	Plesiochronous digital hierarchy
PDU (Protocol Data Unit)	A unit of information (e.g., packet or frame) exchanged between peer layers in a network.
Permanent Virtual Circuit (PVC)	A generic term for any permanent, provisioned communications medium. NOTE: PVC does not stand for permanent virtual channel. No such term has been defined by any standards organization. Neither has the term "permanent virtual path (PVP)." In ATM, there are two kinds of PVCs: permanent virtual path connections (PVPCs) and permanent virtual channel connections (PVCCs).
Physical Layer	The first layer in the OSI Model. It specifies the physical interface (e.g., connectors, voltage levels, cable types) between a user device and the network.

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.

Point-to-point — A term used by network designers to describe network links that have only one possible destination for a transmission.

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.

-Q-

QoS (Quality of Service) — The ATM Forum has outlined five categories of performance (Classes 1 through 5) and recommends that ATM's quality of service should be comparable to that of standard digital connections.

Quantizing — 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.

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.

-S-

Sample — A representative value of a signal at a chosen instant, derived from a portion of that signal.

Sampling — 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 time.

SCR (Sustainable Cell Rate) — A measure of the maximum throughput that can be achieved by bursty traffic over a given virtual connection without the risk of cell loss.

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.

Signal — A physical phenomenon one or more of whose characteristics may vary to represent information.

Signaling (ATM)	The procedures used to establish connections on a ATM network. Signaling 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.
SONET (Synchronous Optical NETwork)	A set of standards for the digital transmission of information over fiber optics. Based on increments of 51 Mbit/s.
SPVC (Soft Permanent Virtual Circuit)	A generic term for any communications medium which is permanently provisioned at the end points, but switched in the middle.
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.
SVC (Switched Virtual Circuit)	A generic term for any switched communications medium. NOTE: SVC does not stand for switched virtual channel. No such term has been defined by any standards organization. Neither has the term "switched virtual path (SVP)." 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 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 coordinate their transmissions. (Compare with "asynchronous.")
Synchronous network	A network in which the corresponding significant instants of nominated signals are adjusted to make them synchronous.

-T-

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.

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 cells (traffic) that do not conform to the Quality of Service parameters specified in the call setup procedure.
Traffic Shaping	A mechanism used to control traffic flow so that a specified Quality of Service 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.

-U-

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.

-V-

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)
VC (Virtual Circuit)	A generic term for any logical communications medium.
VCI (Virtual Channel Identifier)	The field in the ATM cell header that labels (identifies) a particular virtual channel.
Virtual Channel Connection (VCC)	A logical communications medium identified by a VCI and carried within a VPC. VCCs may be permanent virtual channel connections (PVCCs), switched virtual channel connections (SVCCs), or smart permanent virtual channel connections (SPVCC). Further, VCC is an end-to-end logical communications medium. Another acronym, VCL (virtual channel link), is more precise, referring to the single segment object identified by a VCI and carried within a VPC. Similarly, a VPC is an end-to-end object and a Virtual Path Link (VPL) is identified a VPI within a link.
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
Wander	Long-term non-cumulative variations of the significant instants of a digital signal from their ideal positions in time.

Annex B – Recommendations

For convenience of reading, this Annex groups together the recommendations made throughout the document.

B.1 Compression Issues – Recommendations

1. Compression algorithms and transport schemes should be based on Open Standards. This implies availability of the intellectual property 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.
2. The number of compression methods and parameters should be minimized for each uniquely defined application in order to maximize compatibility and interoperability.
3. Compliance testing methods should be available for those building 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 manufacturer and user needs.
4. A single compression scheme used with different compression parameters throughout the chain should be decodable by a single decoder.
5. To support use of more than one compression family, the development of a common (“agile”) decoder is desirable.
6. Integration of video compression into more complex systems must be via standardized interfaces. Translating through ITU-R BT.601, i.e., decoding and re-encoding, is the default method of concatenating video signals compressed using different techniques and/or parameters, although other methods are possible.
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.
8. Issues related to interoperability must be further explored and standards developed to allow predictable levels of performance to be achieved in the implementation of specific applications.
9. Bit streams 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.
10. Compressed bit streams are very sensitive to errors, and therefore it is recommended that appropriate channel coding methods and error protection be employed where necessary.
11. Compression systems should be designed so that, in normal operation, signal timing relationships (e.g., audio/video lip sync) and synchronization presented at encoder inputs are reproduced at decoder outputs.
12. Signal delays through compression processing (encoding/decoding) must be limited to durations that are practical for specific applications, e.g., live interview situations.
13. Provision should be made for selected analog vertical interval 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.

14. The compression scheme chosen for devices that mimic VTRs should allow for the reproduction of pictures in shuttle mode for identifying content and of pictures in jog and slow motion modes for selecting edit points.
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.
16. Storage devices should allow recording and playing back of streams and files as data rather than decoding to baseband for recording and re-encoding upon playback.
17. The compression strategy chosen for standard television should be extensible to high definition applications to allow for commonality in the transitional phase.

B.2 Wrappers and Metadata – Recommendations

1. The development of an extensible hierarchical classification of Metadata varieties, including the notion of Metadata Sets appropriate to particular uses.
2. The establishment of a single registry of Metadata identifiers and definitions.
3. The standardization of a single generic Wrapper format for streaming of Metadata, which can be mapped onto existing and emerging signal transport layers.
4. The standardization of a single generic Wrapper format for applications requiring arbitrary richness of Content of all types, including Metadata and Essence. This must be highly compatible with the streaming format described above.
5. The standardization of a single format for a “unique identifier” would also be of assistance; however it is recognized that multiple formats are already in use. As a minimum therefore, it should be possible to register existing and new unique identifier formats within the Metadata registry described above.

B.3 File Management, Transfer Protocols, and Physical Connections – Recommendations

In order to achieve interoperability between different vendors, which is a basic user requirement, the following standards need to be defined:

1. File Formats for guaranteed and bounded quality transfer
2. Link Protocols: e.g. protocols for different classes of service, transfer mode commands, transfer initiation commands, transfer phase commands, the format of data being interchanged, number of bits per element, type of encoding scheme, synchronization information
3. Flow Control: methods to control the flow of data between two devices (e.g. to avoid overflow of storage at the receiver)
4. Error Control: e.g. type of error detection and control
5. Inter-Layer commands: e.g. error messaging between different layers
6. Machine Control
7. Networks and interfaces
8. Payload and mapping documents
9. User interfaces
10. Platform independent command sets for file management

Annex C – Compression Applications

C.1 The Application Model

The figure below illustrates a generic view of the program chain identifying all the primary application areas. Each of these functions is described in detail below.

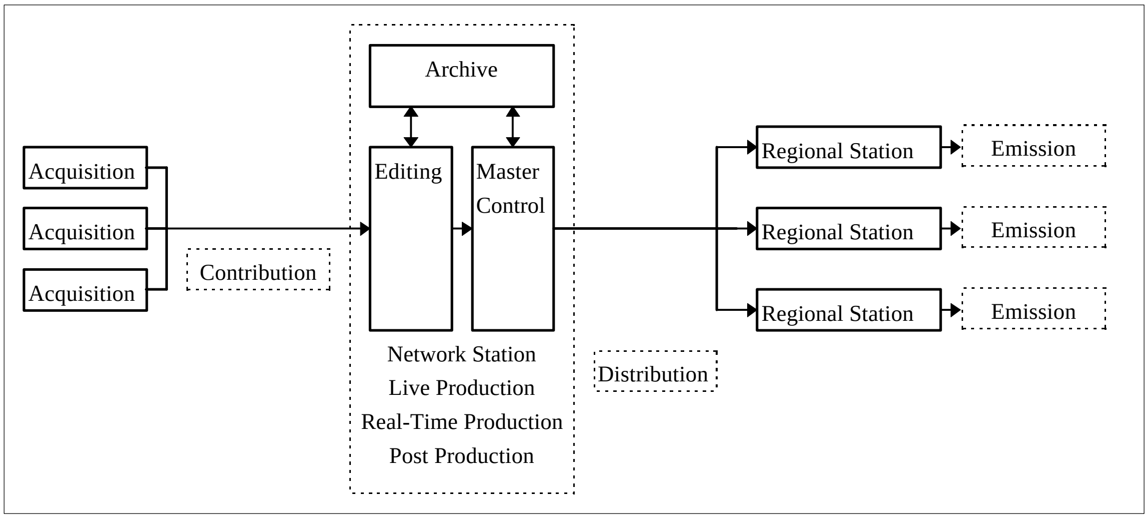

Figure C1 – Generic television program chain.

C.2 Acquisition

Compression offers considerable benefits for field acquisition, e.g., in news, sports, electronic field production (EFP), consumer, and allied uses.

The requirements for equipment for field application are:

high portability,
low weight,
low power consumption (long running time), and
ruggedness.

The application of compression plays a significant role in meeting these requirements because of the reduced bit rate and reduced need for storage capacity that it yields.

The compression scheme used for acquisition should allow field editing of the signal.

For applications such as EFP and Sports applications a 4:2:2 sampling structure is preferable. Usage of a sampling structure of less than 4:2:2 (such as 4:1:1 and 4:2:0) should be acceptable for News Gathering applications provided extensive post processing is not required.

C.3 Contribution

Content acquired in the field and edited on the spot using a portable field editor or fed live must be delivered to the studio for further processing and final transmission.

Communication channels available for this purpose may have a different bandwidth than the recorded compressed signal and possibly a different quality of service. It would be desirable for the compressed signal from an acquisition source to fit perfectly into the constraints of the communication medium. However, due to the wide variety of choices in the marketplace (e.g., DS3, E3, and OC-1 through OC-192, with data rates ranging from tens of megabits to gigabits per second) this cannot be practically achieved without resort to the data rate of the lowest bit bandwidth channel.

Where a compression system has a bit rate greater than that of the communications channel, the options are slower-than-real-time transmission (which may not be practical for some applications) or an additional layer of compression, which will lead to concatenation errors and will result in some degradation of picture quality.

In the case where the compression system used has a data rate lower than the communications channel it will be preferable to use a single compression system rather than use concatenated compression systems. The unused bandwidth may be used for other purposes (e.g. extra error protection or ancillary service information)

Satellite contribution service between remote trucks and studios is not always available due to overcrowding of the spectrum for this type of communication. A choice of low bit rate compression may be helpful in this communication shortfall. However, using this as the only criterion for selection of the bit rate of the compressed signal may have other negative consequences.

Significant input to television news is received in the form of fixed satellite communication. This type of a signal is recorded at the station on a 24 hour basis and selection of the proper compression format is dependent on other criteria which may be different from those of field acquisition equipment.

C.4 Compression in Live/Real Time Production.

Use of compression in field production will ease demand on the communication channel between remote acquisition equipment locations and the studio where final production is completed. The application of compression in Live and Real-Time Production has been covered in the sections on Acquisition and Contribution.

C.5 Post Production/Exploitation

C.5.1 Editing in real time

Editing in real time usually involves the simplest of production effects – the cut, implemented by stream splicing in compressed systems. The splice should be done in the compressed domain without transcoding.

It is highly desirable to be able to perform cuts editing on compressed streams with different bit rates. This area of compression processing is currently in its infancy. However, it is becoming clear that this will be possible between streams of the same compression family.

C.5.2 Post production editing

Post Production editing requires the highest attainable quality. During such production processing, layering is very often used. Any processing effects beside cuts require decoding. Layering, as a process, may involve many subsequent re-encodings; therefore loss of picture quality can be a real possibility.

To avoid this loss of picture quality phenomenon, the producer should use methods similar to those used in the past for layering in analogue environments. One of these methods involved the use of higher quality processing and storage formats to ensure that the original signal is degraded as little as possible. Extensive studies have been performed on concatenation of compressed systems and these studies have shown that concatenated compression will have a negative impact on overall picture quality. The analogue practice of editing and using intermediate storage having the highest quality format available is also applicable in the compressed environment. This means that the processing should be done in the uncompressed domain or in the highest compressed bit rate available to the user.

C.5.3 Off-line post production.

It is assumed that for off-line post production, the user has less demanding picture quality requirements because this process is being used only for creation of an Edit Decision List (EDL). The EDL is later used for actual editing of the program master. In this manner, a post production facility can utilize high quality, high cost post production equipment efficiently. An important feature of the Off Line Editing station is its ability to facilitate proper editing decisions which can be accurately followed in the on-line environment.

C.5.4 Presentation / Master Control.

This type of processing controls the output stream of the station or network. In this environment decisions are always made live and usually under dynamic conditions. The operator must make decisions based on specific situations and therefore response time and repeatability are of the highest importance. At this processing point, the operator has little control over picture quality, only over delivery of picture content.

Reliability of the equipment is also of the utmost importance, and any possibility of failure could have catastrophic consequences for the facility. It is clear that the use of compression will not improve this situation. However, if a facility operates with compressed signals as its primary methodology, it is logical that master control should follow the processing pattern of the facility and switch compressed signals.

Switching in Master Control will take place at the Packetized Elementary Stream (PES) level. This process may be followed by encoding into Transport Streams. Nevertheless, the output picture quality will have been pre-determined by previous encoding processes, and the final picture quality will be decided by the bit rate and structure of the released compressed stream.

C.6 Distribution

C.6.1 Satellite links

Satellite links are frequently available as analog channels within which a suitable digital modulation process can be utilized. This allows considerable flexibility in the choice of video compression system bit rates. Where input bit streams have been previously encoded, satellite encoders should use the same coding process as that of the feeder content, or the pre-encoded content should be directly passed to the satellite modulator to ensure no concatenation losses. Where this is not possible (because the available bit-rate is less than that of the pre-encoded bitstream) concatenation losses will be incurred.

C.6.2 Public carriers (telcos).

The issue of Public Carriers has already been covered in the section on ‘Contribution’ where the same points apply.

Other features of both Satellite Links and Telco Links include the possibility of faster-than-real-time transfers for both cost and operational benefits.

C.6.3 Emission

Transmission to the home is subject to low bit-rate limitations and likely to be MPEG-2 MP@ML, probably limited to less than 8 Mb/s but certainly less than 15 Mb/s.

Insertion of local content may be required in some circumstances through the use of compressed stream “splicing.”

C.6.4 Packaged media distribution

Distribution of content to the home is, like Emission, subject to low bit-rate limitations. A major difference, however, is likely to be the use of the VBR technique and a single channel output.

Insertion of local content will not be required, but provision for interactive branching within the stored content will, in some cases, be necessary.

C.7 Archiving

C.7.1 On-line archiving

On-line archives will generally directly record compressed bit streams to avoid the concatenation effects of another compression system. The archive may also be associated with highly compressed browse mode images and metadata to aid recovery of archive content. The quality level of the browse mode images is only required to support picture recognition, with no expectation that these pictures will be used for any other purpose.

Metadata and browse mode images will normally be located on the same storage device for rapid access to the content. Full editing capability should be possible with on-line archive content.

C.7.2 Near-line archiving

A near-line archive is a mid-way archive containing copies of the content and metadata. Typically, browse mode images and metadata will still be stored in the on-line archive for rapid access whereas content will be stored off-line on a remote server. Full editing capability of near-line archive content is still possible.

C.7.3 Deep / long-term archiving

In a deep archive, content and all associated metadata will be stored in the archive, although only the metadata and browse mode images may be readily accessed for archive browsing. Depending on the deep archive requirements, the content may be subjected to a further compression process such as MPEG-2 MP@ML using a low bit rate.

Annex D – Wrappers and Metadata

D1 – Platform Neutrality

Platform Neutrality normally refers only to the byte-ordering of multi-byte data items in Files. It also refers to the limitations of file lengths through the storage filing system (a common limit is 4Gbyte). A format is Platform-Neutral if the complexity of implementation is roughly equal on any platform. For the business of the Task Force, Platform Neutrality must have a wider scope: it needs to cover efficiency of encoding and decoding simultaneously on different platforms.

There are several topics discussed here:

- bit ordering
- transmission of multi-byte data
- Endian-ness flag issues
- 10-bit data and C_b Y C_r [Y] sample structure

D1.1 Bit ordering

All computers and many other items of broadcast equipment use a minimum symbol size of 8 bits. Bit ordering within each byte therefore presents no problems for media interchange. However, if symbols are serialized for transmission, the bit order must be defined to ensure that the decoded bytes match the bit order from the transmitter. Both 'MSB first' and 'LSB first' are used for different inter-connects but provided a given order is defined by the transmission protocol, byte reconstruction at the receiver will be consistent with the transmitted bytes. There is thus no requirement for a "bit-Endian-ness flag" nor the accompanying complexity.

D1.2 Multi-byte data

Suppose a 32-bit word, 0x76543210 is contained in a Wrapper. If it was written on a Little-Endian (Intel) machine, it will be stored in the sequence: 10 32 54 76. If it was written on a Big-Endian (Motorola and others) machine, it will be stored in the sequence: 76 54 32 10. Clearly this presents problems for interchange between the two Endian forms whether by media transfer or transmission.

Words which are part of Overhead (i.e., not Essence or Metadata) would be expected to always be one way (Little- or Big-Endian) in order to avoid the definition of decoders which first need to refer to an Endian-ness flag before deciding how to decode.

In the case of words which are part of the Essence or Metadata: if the data is Presented or Transported as a File (i.e. it is going from A to B as a unit), and is transported transparently, the decoder will accurately reconstruct the original File contents. This applies whether the File is being transmitted or transferred by media exchange.

However, if the data is being Presented or Transported as a Stream in which interpretation of the data is required (e.g. for a downstream MPEG decoder), then the Byte order of file transfer becomes very important.

In some cases, we are assisted because the data format defines the answer. For example, MPEG bit stream syntax is one-bit serial, and elementary streams are mapped onto a byte stream, MSB first; the byte stream is Endian-free (a byte string is always stored the same way on both Intel and Motorola processors). Also, any Metadata which is character string based is Endian-free.

It would be desirable to have Essence data which is Endian-free but allow encoded Metadata (for example, position measured as a 16 bit number of pixels) to be either way.

D1.3 Endian-ness flag issues

Byte order is simply handled by having an Endian-ness flag (Little- or Big- Endian). Then all data types: short, long and long-long use the Endian-ness flag to change the byte order if necessary.

It is not possible to enforce a particular byte order in all cases, so a byte order flag is a requirement. If a Wrapper contains data with the wrong byte order value, and the platform cannot operate with that byte order, then the data has to be converted. This is necessary and time-consuming, but unavoidable.

For operational simplicity, it would be desirable to have all the data within a wrapper be of the same byte order to ensure the most efficient playout.

For applications which merge data of different byte order, there are two possibilities:

- convert the data at the time it is copied, or use a reference between two Wrappers, and
- allow a change of byte order from one Wrapper to the other.

The byte order flag within a Wrapper can occur in two places:

- as early as possible in the Wrapper Overhead
- as part of the Composition Metadata referring to another wrapper.

Note that only one of these two is required; and if both are present, they must be the same. From one Wrapper to the next, there would be two opportunities to signal the Endian-ness: once in the reference pointing out of the parent Wrapper, and once in the Overhead of the Wrapper being referred to.

D1.4 10-bit data and ITU-R BT.601 C_b Y C_r [Y] sample structure

The bit packing of Essence data must be designed with a view to efficient decoding on all platforms since no perfect Endian-neutral scheme exists. Complex data packing such as required to accommodate, for example, a 10 bit per sample RGB sampling structure occupying 30 bits per Word, must be organized to avoid placing the optimal pattern in favor of one value of Endian-ness.

This is discussed further in D4 – Notes.

D2 – Wrapper Referencing

D2.1 Wrapper Varieties

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

A. Unwrapped Content - for example, signals from today's equipment, or from foreign, non-conforming systems.

B. Wrappers whose Content is predominantly Essence, but which may include some Metadata.

C. Wrappers whose Content includes no Essence, and only includes Metadata. If Essence is involved it will be kept elsewhere (wrapped or unwrapped), and these Wrappers will include references to it.

D. Wrappers which include predominantly Composition Metadata, and presumably therefore include many references to Content of type A or B kept elsewhere.

E. Wrappers which include Composition Metadata and Essence.

F. Wrappers which include Composition Metadata and additional Descriptive Metadata, which in turn refers to Content kept elsewhere.

Type C is particularly intriguing. The Metadata forms an Index or Directory of Content. This is one variety of Association Metadata.

A simple example from current practice is to refer to a segment of video tape from within an EDL. This is achieved today using "reel numbers" and two timecodes, the "source in-point" and "source out-point".

(Note: SMPTE 258M calls these the SOURCE_IDENTIFIER, SOURCE_ENTRY and SOURCE_EXIT fields).

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

- an entire Wrapper
- a point within some Content (e.g. a single frame, or instant)
- a specific item of Metadata
- a region within some Content (e.g. a "sub-clip" of a shot)

D2.2 Referencing Content

There are at least five ways that Content may be referenced from within a Wrapper (typically from within Composition Metadata):

1. Content is contained within the same file.
2. Content is referenced in an external unwrapped file (e.g. media in some "raw" or "native" format). Note that in this case there is no way to guarantee that this Content is the correct material, except that data format, length, and perhaps name if available in that external format can be checked against the Metadata description of the Content.
3. Content is referenced in an external Wrapper with the same UID as the original reference. (There may be indirection here, where a referenced Wrapper does not contain Content but references yet another wrapper that does.)
4. Content is referenced in an external wrapper, but has been replaced by new Content from the same original source by an application that retains the original Metadata (perhaps recreated at a different resolution, or perhaps because it was deleted to conserve storage and then re-created when needed again).

It would be desirable to have two levels of UID – a “handle” and one for the actual Content. All Content references would be to the handle’s UID. Then a Content item could be replaced with a new UID, and its handle would have its reference to the Content updated. All external references to the handle would remain valid.

5. Content is recreated in a separate environment from the same original source. In this case there is no way to have UID references. An application could examine the original source information (e.g. tape and timecode range) and determine that this Content is equivalent to the original Content, and update the reference to the new UID.

D2.3 Reference Types

The cases of references are therefore as follows:

- 1a. from a “user” Wrapper to the whole of a “source” Wrapper
- 1b. from a “user” Wrapper to a point within a “source” Wrapper
- 1c. from a “user” Wrapper to an item within a “source” Wrapper
- 1d. from a “user” Wrapper to a segment of a “source” Wrapper

example: the “user” Wrapper is Composition Metadata, the “source” Wrapper is Essence

- 2a. from a Wrapper to another point within the same Wrapper
- 2b. from a Wrapper to another item within the same Wrapper
- 2c. from a Wrapper to a segment elsewhere in the same Wrapper

example: the user Wrapper is a conglomeration of Composition Metadata and Essence

- 3a. from a Wrapper to the whole of an external file, tape or other storage means
- 3b. from a Wrapper to a point within an external file, etc.
- 3c. from a Wrapper to an item within an external file, etc.
- 3d. from a Wrapper to a segment of an external file, etc.

example: the user Wrapper is Composition Metadata, the external tape is a timecoded videotape

Note that in this example, the external file has a “natural” labeling method (in this case, timecode); so, even though it doesn't have a header or anything to say exactly what range of timecodes is on the tape, the Wrapper can point directly to the portion required.

A more general case would employ a directory (or index) between the user and the source, to locate the actual Content.

- 4a. from a Wrapper via a directory to the whole of an external file, etc.
- 4b. from a Wrapper via a directory to a point within an external file, etc.
- 4c. from a Wrapper via a directory to a segment of an external file, etc.

example: the user Wrapper is Composition Metadata, the directory contains Association Metadata to translate timecode into byte address (e.g. "12:23:34.07 is at offset 0x157A3C in the file"), the external file is a data file on a disk array.

We should also consider:

- 5a. from a Wrapper via a directory to the whole of a source Wrapper
- 5b. from a Wrapper via a directory to a point within a source Wrapper
- 5c. from a Wrapper via a directory to an item within a source Wrapper
- 5d. from a Wrapper via a directory to a segment of a source Wrapper

D3 – Access Control and Copyright

D3.1 Access Control

Operational security should be implemented to prevent unauthorized operational access to the data. There are a number of access control methods available which address the users requirements from which the most appropriate one can be selected. The access control method may involve randomization or encryption of the Metadata in order to prevent unauthorized access.

It is expected that all users will initially access a system via a log-in process on a local machine. Permissions will be granted to a user according to a log-in status map defined below.

Authorized users (i.e. those who have passed the log-in process) will be granted access to Wrappers at three levels of user identification:

- as a named user,
- as a member of a named group,
- as any user.
- Additionally, users may be restricted to access of Wrappers within the limits of the following locations:
- local machine,
- site (e.g. the company building),
- corporation,
- anywhere.

Users may be further restricted to access of Wrappers by time limitations as follows:

- from a defined time,
- to a defined time,
- between two defined time intervals and
- anytime.

(Note: the method of identifying defined times is not specified but would normally be through a script defining the valid times and allowing repeats such as days in a week, hours in a day, holidays etc.)

Authorized users may be granted access to Wrappers according to a profile of user capabilities including:

- the right to access (i.e. make the Wrapper visible to the user),
- the right to play (read),
- the right to create (new write),
- the right to modify (re-write),
- the right to erase (delete)
- the right to administer (change any of the restrictions above).
- Rights such as browse and other rights may be added but need further consideration.

Note that these rights are granted only for the specific file accessed and may change when accessing another file.

The granting of these rights may also be controlled by an API interacting with Metadata describing Ownership, Ratings control and so on.

Operational security measures as above may be implemented in each Wrapper so that access requests and responses may be coordinated through the Application Programming Interface (API). In order to reduce duplicate security measures, equipment may bypass one or more of the above access levels where the underlying operating system provides equivalent security level replacements.

It would be desirable if the access control could be monitored to provide an audit trail for the tracking of faulty operations, unauthorized access, and commercial transactions.

D3.2 Intellectual Property Rights

The Wrapper should contain ownership parameters for Content as follows:

- Content Originator,
- Content Copyright and
- Content Owner.

The Content originator is the name of the creator of that content. It is a permanent name since the creator can never change. This value will be maintained over any Content operation such as copy, move, modify etc.

The Content Copyright is the name of the owner of any copyright contained in the content material. An example is a picture of a work of art, where the copyright of the artist must be associated with the Content. The Content Copyright is permanent except where the copyright owner has given express permission for a change of owner.

The Content Owner is the current owner of the Content and may change.

Each of these values may be associated with any defined frame or segment of the Content. Furthermore, since the Content can be represented as a sequence of Content Components (e.g. a frame of video), all parts of that Content Component is associated with the values of the whole Content Component. For example, if a small still picture is created from a frame of video, that small picture must copy the Intellectual Property Rights (IPR) values from that frame.

If any parameter has no owner, then the value is null.

A charging transaction mechanism is required for the automatic billing of resource usage, which might be provided through an API.

D4 – Notes

Several discussions which contributed the output of the present work on Wrappers and Metadata should be noted for future work. These included an analysis of applications, ideas for future implementations, consideration of possible solutions to the requirements given, or notes on the constraints of present technology.

These topics are noted here.

D4.1 Applications

In considering the categories of Applications for the Wrapper Profiles and Metadata Sets, as many potential activities as possible were enumerated, and these were then grouped into the major categories used in the main Wrappers and Metadata Annex.

This list also forms the preliminary list of Metadata Sets which will be required.

1 Pre Production
- 1.1 Scripting
- 1.2 Music Composing
- 1.3 News Assignment (Event)
- 1.4 Planning/Design
  - 1.4.1 Storyboarding
  - 1.4.2 Location Research
  - 1.4.3 Budgeting and Contracting
  - 1.4.4 Sets, Props, Costume
  - 1.4.5 Modeling

2 Production /Acquisition
- 2.1 Live News and Sports
- 2.2 Live Production
- 2.3 Video/Audio Recording,
- 2.4 Film Shooting,
- 2.5 Graphics
- 2.6 Animation
- 2.7 Motion Capture

3 Post Production
- 3.1 Editing
  - 3.1.1 Off-line Editing
  - 3.1.2 Field Editing
  - 3.1.3 A/B roll and On-line Editing
  - 3.1.4 Film Editing
  - 3.1.5 Screening, Workprint and Negative Conforming
- 3.2 Compositing/ Manipulation
  - 3.2.1 DVE, Keying, Paint, Rotoscoping, Colorizing
  - 3.2.2 Real-Time Graphic Workstations,
  - 3.2.3 Multiple M/E Linear On-line Editing
- 3.3 Sound
  - 3.3.1 Dialogue Editing
  - 3.3.2 Foley, ADR
  - 3.3.3 Music and Effects Editing,
  - 3.3.4 Mixing, Audio Sweetening, Re-recording
- 3.4 Multimedia Authoring
  - 3.4.1 Pre-Mastering, Assembling, Linking, Encoding, Bit-budget Allocation
- 3.5 Film Negative Cutting

3.6 Foreign Language Dubbing, Titling, Captioning, Sub-Titling, Internationalization

4 Distribution / Storage

4.1 Routing, Client/Server Access, DDRs VTRs and ATRs

4.2 Receiving Feeds, Internet Download, Archive Retrieval, Inter-facility Transfer, Relay, Backhaul

4.3 Standards Conversion

4.4 Lay-back

4.5 Quality Control

4.6 Asset Management

4.7 Uplinking

5 Emission

5.1 Playlist Preparation, Log Creation

5.2 Wholesale Delivery

5.2.1 Uplinking

5.2.2 Cable Headends

5.2.3 Satellite Headends

5.2.4 Shipping to Theaters,

5.2.5 Shipping Master to Duplicator

5.3 Broadcast

5.4 Commercial Insertion

5.5 Motion Picture Projection

6 Archival

6.1 Near-Line Storage

6.2 Long Term Storage

6.3 Deep Archiving

6.4 Asset Management

D4.2 Usage of Essence and Metadata

Various attributes of the Content were evaluated in each category of activity. The attributes were:

- the absolute amounts of Essence and Metadata
- the absolute bit rates of the Essence
- the relative proportions of each category of Metadata
- the extent to which Content is consolidated into a single Wrapper
- the kinds of access to Content

Approximate ranges were assigned to each attribute as noted in the table headings (except for access kinds), and the values were normalized to the range 0-10.

It was clear that with few exceptions, each activity required formats optimized for both streaming and richness, and for both sequential and random access.

Table D4.1: Essence and Metadata Attributes in each Activity

CATEGORY of ACTIVITY	ESSENCE		METADATA							
	AMOUNT	BITRATE	AMOUNT	DESCRIPTION	COMPOSITION	ASSOCIATION	CONSOLIDATION	RANDOM	MIXED	SEQUENTIAL
Scale 1 Scale 10	Little Lots	1Mbps 200Mbps	Kilobytes Megabytes				Many Wrappers Single Wrapper			
Pre-Production	1	1	3	3	1	0	1		X	

CATEGORY of ACTIVITY	ESSENCE		METADATA							
Production / Acquisition	10	7–10	2–7	7	1	2	5–8			X
Post - Production (Edit, Composite, Sound, Multimedia Authoring)	3	8–10	10	9	10	2	1–5	X	X	X
Distribution / Storage	2	7–10	1–5	5	2	4	1(in) 7–10 (out)	X		X
Emission / Mastering	1	5–10	1–3	1	1	2	10			X
Archive	2–10	1–10	10	10	10	2	5 (Regular) 7–10 (Deep)	X		X

D4.3 References and Labels

The internal method of storing references may change according to the adopted file system and may include methods such as sample offset and sub-files.

D4.4 Security

It might be desirable to include a platform-independent decryptor within the Wrapper (such as a Java applet) unless standard encryption methods are used.

D4.5 API

It is desirable to have a standard API available to lower the barrier for reading and writing the Wrapper format. It might be desirable to include platform-independent executable code such as Java within the Wrapper itself for self-unpacking/packing of both the Wrapper Metadata and Content. An advantage of a built-in API is that it hides the internal organization and specific storage methods of the Wrapper data, so the structure can be changed as long as a correct API is included.

D4.6 Essence Extraction

Means for coding and/or decoding Content might be provided in the API by supporting various methods such as:

- hardware devices accessed by device drivers
- executable code specific to a processor architecture and
- platform independent executable code such as Java.

These methods may be implemented as plug-in modules to provide user extensibility and future upgradability. Software code may be attached to the Content to provide features such as self extraction.

D4.7 Efficiency and Completeness

Different uses of Wrappers place different requirements on the performance of the Wrapper format.

For Content which is to be presented as a stream, there is an emphasis on encoding efficiency and efficiency of information retrieval. For other data, there is an emphasis on richness of data description.

It is recognized that there may be a conflict between efficiency and completeness in some applications. To help resolve this conflict, mechanisms to automatically prune or ignore optional Metadata may be required within Wrapper format converters, importers or application programmer interfaces (APIs). In the interests of efficiency, it is also desired to avoid the copying of data (particularly Essence) when converting between Presentations.

Real-time live transfer by streams may require repeating of Metadata and interleaving of structures. Certain synchronous Metadata falls into a category closely bound to the Essence, for example Timecode. There is little value in separating such Metadata from the Essence when it does not need to be accessed independently of the Content.

D4.8 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 addition of new Essence and Metadata types,
- By 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 with an efficient but complete documentation process, 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.

The extension or alteration of data syntax poses a greater problem in providing for backward compatibility. To facilitate future extensions, every Wrapper format is required to carry a Version Number, to be managed as follows:

- The Wrapper Version Number will be carried in every Wrapper, within the Overhead.
- The Version Number will be assigned by the standards body which documents the Wrapper format, and will be published in a Registry.
- Every device or application which accepts the Wrapper format must decode and check the Version Number.
- Every device or application which accepts the Wrapper format must decode every earlier published version of the format.
- The Version Number will be changed every time that that the Wrapper format is changed in a way which is not absolutely transparent to every implementation of an earlier version.
- The Wrapper format must never be changed in a way that makes decoding of the Version Number impossible or unreliable.

D4.9 Immutability and Generation Numbering

In the general case, the Content in a Wrapper should not be changed because of the possibility of unknown references to the Content in this Wrapper. Changes can be achieved by altering a copy of the Content with a different UID (this does not apply where it is known that there are no references to this Content – but reference counting is possible only within closed systems).

A possible solution is to include a Generation Number as a part of the Reference, so that the Composition Metadata in a downstream Wrapper can indicate the generation which was current at the time of creation. When content is changed, it is not altered in place, but a new copy is made with a new generation number. Then, if the previous generation is deleted for whatever reason, the downstream user can be offered the choice of trying to use the later generation or sending off a request to retrieve the older one.

D4.10 Endian-ness of 10-bit Sample Structures

When 10 bit data is carried in an 8 bit channel, issues of word sync exist. For example, 10 bit Y C_b C_r components might be packed into 32 bit words at 3 components per word, so that we have a packing of 6 samples into 4 words (16 bytes):

C_b	Y	C_r	[Y]	C_b	Y	C_r	[Y]	C_b	Y	C_r	[Y]
w1			w2			w3			w4		

There are many choices for the actual bit-by-bit packing of the components into the words. Four are considered here – “straightforward”, “LSBs separate”, “sparse”, and “tight”. The choice of packing method may strongly affect the efficiency of decoding the components on a different platform from the originator.

D4.10.1 "Straightforward" Packing

"Straightforward" packing is as follows

AAAA AAAA aaBB BBBB BBbb CCCC CCCC ccxx

where `aa bb` and `cc` are the LSBs of the 10bit words, and `A B C` are the MSBs.

After being stored big-Endian, read little-Endian (or vice versa), this bit pattern would become:

CCCC ccxx BBbb CCCC aaBB BBBB AAAA AAAA

All similar alternatives create similar or even more complex permutations. For example:

xxAA AAAA AAaa BBBB BBBB bbCC CCCC CCcc

becomes:

CCCC CCcc BBBB bbCC AAaa BBBB xxAA AAAA

Software algorithms to untwist these are inefficient. Hardware implementations may require an additional stage of multiplexers or shifters plus an Endian-ness flag.

D4.10.2 "LSBs separate" Packing

"LSBs separate" packing is as follows

AAAA AAAA BBBB BBBB CCCC CCCC xxaa bbcc

After being stored big-Endian, read little-Endian (or vice versa), this bit pattern would become:

xxaa bbcc CCCC CCCC BBBB BBBB AAAA AAAA

This still requires untwisting, but the permutation is the same for reading and writing in all cases, whether the Endian-ness of the destination is the same as the source or is the opposite.

Hardware implementations of this scheme still require an Endian-ness flag, but no additional multiplexing or shifting is required.

D4.10.3 "Sparse" Packing

"Sparse" packing expands 10 bit samples to 16 bits each, and is as follows:

AAAA AAAA aaxx xxxx BBBB BBBB bbxx xxxx

After being stored big-Endian, read little-Endian (or vice versa), this bit pattern would become:

bbxx xxxx BBBB BBBB aaxx xxxx AAAA AAAA

This is quite simple for either software or hardware to process, but at the expense of considerable storage overhead.

D4.10.4 "Tight" Packing

"Tight" packing leaves no bits unused, and is as follows:

AAAA AAAA aaBB BBBB BBbb CCCC CCCC cc<u>AA</u> <u>AAAA</u> <u>AAaa</u> <u>BBBB</u> <u>BBBB</u> <u>bbCC</u> <u>CCCC</u> etc.

After being stored big-Endian, read little-Endian (or vice versa), this bit pattern would become:

CCCC cc<u>AA</u> BBbb CCCC aaBB BBBB AAAA AAAA <u>Cccc</u> *AAAA* <u>bbCC</u> <u>CCCC</u> <u>BBBB</u> <u>BBBB</u> etc.

(where **<u>AA</u>** is from the second group and ***AA*** is from the third sample group). As can be seen, this method causes the most involved permutations of all.

D4.10.5 Permutation during the transfer

Certain transport and interconnect technologies (for example, Fibre Channel) provide some facilities for automatic conversion of byte order during the transfer.

These facilities must also be taken into account when defining data packing. In some cases, automatic conversion can improve the efficiency of importing data from a foreign platform. However, without proper identification of the original byte order and adjustment of the flag after conversion, the automatic process might actually worsen compatibility.

D4.10.6 Conclusion

The conclusion to be reached from this discussion is that 10 bit samples add considerable complexity to the choice of a platform-neutral scheme, probably implying that Endian-ness flags must be carried as part of the Format Metadata for each Essence Component.

D4.11 Ideas for Wrapper Referencing - SMPTE 258M and HTML

One approach to References might be based upon extensions of HTML to reflect current practice within EDLs, as described in this section.

Within HTML, the concepts of labeling points within a file and referring to them from within another file are provided (by "<a name=nnnn>" and "<a href=rrrr>" respectively).

In addition, the concept of a directory or index is provided (in HTML, this is achieved by the "<img ismap src=dddd>" construct).

However, some issues are not dealt with cleanly by employing the exact scheme used in HTML. In particular, the need to address a segment of a source file, and the very common use of timecode in television do not presently map to HTML.

It is possible that a new variety of URL could be devised to address these needs, for example:

tfhs://server/mount/path/filename#hh:mm:ss.ff--hh:mm:ss.ff

In this example, the fields are as follows:

- "tfhs:" identifies the service, just like "http:" or "ftp:"
- "server" identifies the location of the material. It may be a physical address, or a local name, or an absolute name; following HTML precedent, it might include user names and passwords for access control.
- "mount" and "path" are the route through the filesystem.
- "filename" is the name of the Wrapper; in the case of a videotape, it might be the tape reel name.
- "#hh:mm:ss.ff--hh:mm:ss.ff" identifies a segment of material. There might also be "#hh:mm:ss.ff" to identify a point, and "#anchor" to identify a specific item of Metadata, and so on.

Further work is required to understand how this notation would be applied in all the cases above. But it seems that there are some intriguing possibilities:

The source reference in today's EDL:

0001 tape66 VA1A2 C 12:12:12.00 12:12:17.00 10:00:00.00 10:00:05.00

Could become:

tfhs:tape66#12:12:12.00--12:12:17.00

(following normal HTML "defaulting" rules, if server, mount, path are omitted, they would refer to obvious defaults).

D5 – DAVIC and DAVIC Terminology

The Digital Audio-Visual Council (DAVIC) is a non-profit Association registered in Geneva. Its purpose is to advance the success of emerging digital audio-visual applications and services, initially of the broadcast and interactive type, by the timely availability of internationally-agreed specifications of open interfaces and protocols that maximize interoperability across countries and applications or services. The DAVIC concept of Digital Audio-Visual Applications and Services includes all applications and services in which there is a significant digital audio video component. The goals of DAVIC are to identify, select, augment, develop and obtain the endorsement by formal standards bodies of specifications of interfaces, protocols and architectures of digital audio-visual applications and services.

The Content structure and terminology used by the Wrappers Sub-Group is intended to be compatible with that employed in DAVIC documents, although there are some slight differences. It is thus possible for DAVIC to be a subset or an application profile of our more general definition. The following discussion reviews the DAVIC concepts contained in DAVIC baseline document #22, to show the similarity of approach.

All programming content is represented in the DAVIC system as **Multimedia Components**. Multimedia Components comprise one or more **Monomedia Components** coupled with the logical relationships between the Monomedia Components. The Multimedia Components will be created by content providers.

D5.1 Monomedia Component Types:

The basic Monomedia Components in DAVIC include the following elements:

- Characters
- Text
- Language Information
- Service Information
- Compressed Audio
- Linear Audio
- Compressed Video
- Still Picture
- Graphics

The term Essence as used by the Wrappers Sub-Group is equivalent to the actual data of a DAVIC Monomedia Component excluding any Metadata.

Multimedia Components comprise one or more Monomedia Components.

DAVIC specifies a number of different profiles. In a specific profile there may be support of a subset of the Monomedia Components.

D5.2 Content Structure Definitions

A **Content Item Element** (CIE) is indivisible, and consists of Essence of a single monomedia type, plus any Metadata directly related only to that Essence.

Wrappers need to be concerned with even smaller units of Essence, reflecting the division of Content Item Elements for interleaving and multiplexing. The Wrappers Sub-Group calls these smaller units Content Components (or Essence Components when specifically referring to Essence).

The Wrappers Sub-Group uses the term Content Element instead of Content Item Element.

A **Content Item** (CI) consists of a collection of one or more Content Item Elements (CIE), plus any Metadata directly related to the CI itself or required to associate the component parts (CIEs) together.

A **Content Package** (CP) consists of a collection of one or more Content Items or Content Item Elements, plus any Metadata directly related to the CP itself or required to associate the component parts (CIs and CIEs) together.

For each Monomedia and Multimedia Component the coding format is specified, as well as applicable constraints for coding of the Components. The coded representation of each Monomedia Component is defined to be packetized in specific ways that permit to include time stamps to support mutual synchronization of multiple Monomedia Components.

Among the important objectives for such a structure are (with no priority in the listing):

- Random access to individual Content Item Elements, e.g. for transfer, updating, deletion and adding.
- Extensibility of elements, i.e. adding new attributes must be possible with backward compatibility.
- Global referencing possibilities in and out of packages.

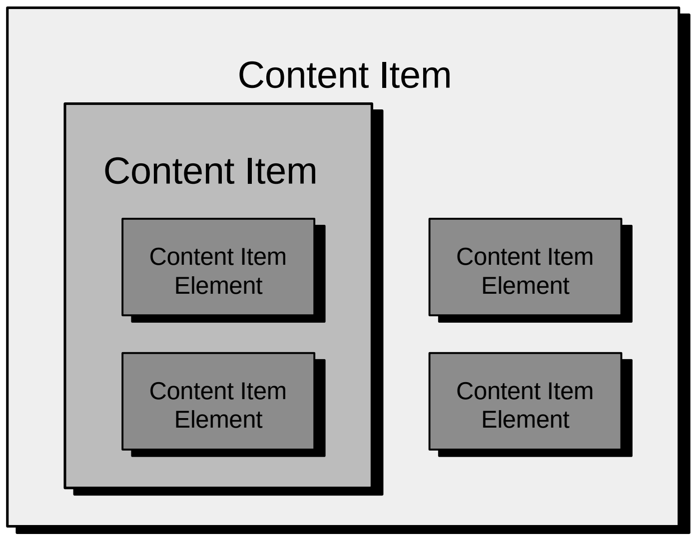

Figure W.5.1 Illustration of Content Items and Content Item Elements

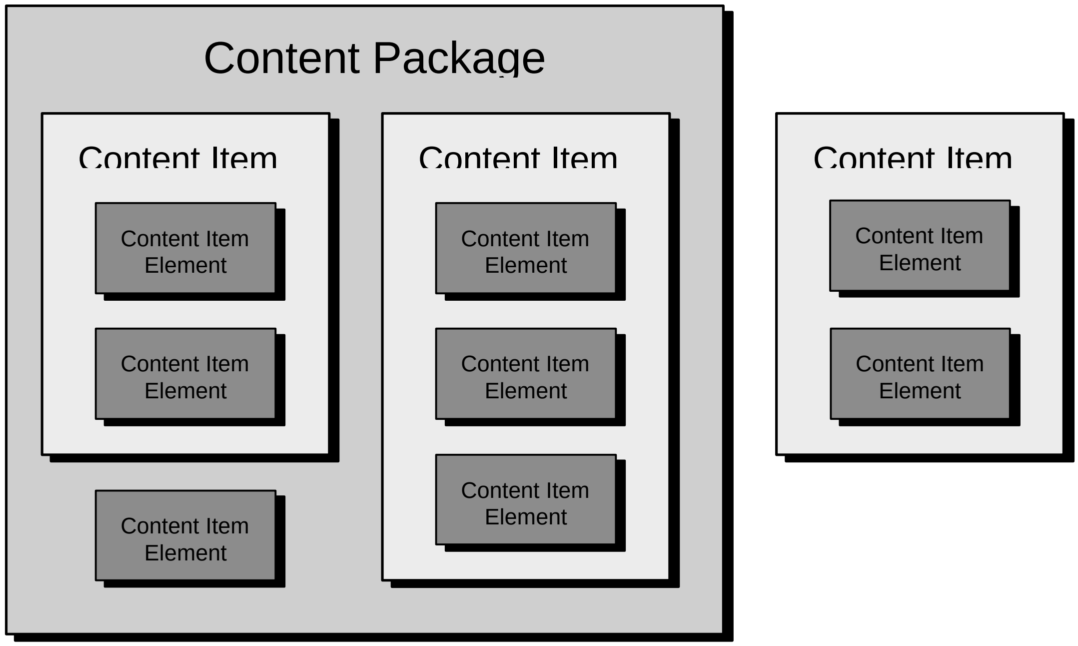

Figure W.5.2 The relationships between Content Packages, Content Items and Content Item Elements

D5.3 Metadata

DAVIC divides Metadata into two categories: Content Management and Navigational. Content Management Metadata refers to Metadata that allows the Content to be produced, processed and transferred. Navigational Metadata is used to enable use of the server content, as controlled by the client and server application executables. This may include navigation, embargo and deletion control, copyright management, and user access control.

A core set of Metadata items are required to be provided with each Content Item. This small core set provides for the basic management of server items.

Metadata can be related to Monomedia Components (one channel of audio, video, stills, graphics, text) or to Multimedia Components (Content Item Element, Content Item, Content Package).

Metadata which is related to a Content Item Element, (a subsection of the Content, e.g. a frame or scene), is referred to as Time Variant Metadata. A subset of those is formed by Synchronous Metadata that are needed for real time live transfer.

Metadata **Types** define broad classes of Metadata. Metadata **Attributes** define specific parameters within a Metadata Type and take either an explicit value or a reference to a data source.

D5.4 Overall Structure

The structure used should be hierarchical. It is recommended that certain sets of types and attributes are allocated to application dependent Metadata Sets. These would be registered by an independent registration authority like SMPTE. The following example shows one possible arrangement of Content Packages, Content Items and Content Item Elements, Metadata Types and Attributes within a Wrapper.

D5.4.1 Metadata Data Structure Example

```
[Wrapper]
    [content package]
        description.title = “Days of our Lives”
        description subtitle = “Episode 24”
        description.owner = “XYZ-TV”
            [content item]
                description.title = “first program segment”
                temporal.duration = 00:00:15:00
                    [content item element]
                        video.format = “525 4:2:2”
                        video.location = “server1\days24_1.vid”
                    [end content item element]
                    [content item element]
                        audio.format = “48kHz AES/EBU”
                        audio.location = “server2\days24_1.aud”
                    [end content item element]
            [end content item]
            [content item]
                description.title = “second program segment”
                temporal.duration = 00:00:15:00
                    [content item element]
                        video.format = “525 4:2:2”
                        video.location = “server1\days24_2.vid”
                    [end content item element]
                    [content item element]
                        audio.format = “48kHz AES/EBU”
                        audio.location = “server2\days24_2.aud”
                    [end content item element]
            [end content item]
    [end content package]
[end Wrapper]
```

Annex E – File Management, Transfer Protocols, and Physical Connections

The transfer of program content can take place either as a continuous stream (as with existing analog playout) or in the form of a discontinuous file transfer.

All transfers require a number of operations, including:

- identification of the file or stream and its source
- the destination for the transfer
- the establishment of a physical connection
- the control of the transfer

File transfers also require:

- a logical connection
- the identification of the file at its destination
- notification to the user of completion of the transfer (or the reason for a failure)

Consequently, issues of file management, transfer protocols and physical connections must be considered.

This chapter sets out the user requirements in each of these areas.

The main area of application considered in this chapter is the transfer of content within a production facility. This implies a local area network environment. Wide Area Networks, on the other hand offer the possibility of incorporating remote storage devices within a logical cluster ("distributed video production"), but different requirements for content transfer, particularly the protocols, will be necessary. This report is currently limited in scope to operation within a local area network environment (LAN); the wide area network (WAN) context requires further study. Some pointers to such studies are given at the end of this chapter in “Wide Area Network Gateways”.

In addition, attention is drawn to the desirability of common machine control interfaces to enable users to switch easily between devices from different suppliers.

E.1 Definitions:

Synchronous “The essential characteristic of time-scales or signals such that their corresponding significant instants occur at precisely the same average rate.”

Asynchronous “The essential characteristic of time-scales or signals such that their corresponding significant instants do not necessarily occur at the same average rate.”

Isochronous “The essential characteristic of a time-scale or a signal such that the time intervals between consecutive significant instants either have the same duration or durations that are integral multiples of the shortest duration.”

Note: In practice, variations in the time intervals are constrained within specified limits.

Interworking "The practice of enabling disparate entities to achieve some level of cooperative working via the parts they have in common; an example would be filename compatibility between two different computer operating systems."

Latency "The delay in time after a command or initiation before an associated action or flow starts; in the case of a data flow, the latency is not to be confused with the time the data spends traversing the channel between source and destination"

metadata	"data about data, e.g. about its identity, attributes, history or means of combination with other data; examples include timecode, an edit decision list and SMPTE Video Index"
multicast	"a data flow from a single source to multiple destinations; a multicast may be distinguished from a broadcast in that the number of destinations may be limited."
packetized data	"data such as that representing video that is segmented into small pieces that are, for example, wrapped, labeled, numbered, addressed, error protected so as to survive transit through a heterogeneous or hostile environment; the small pieces are reassembled on completion of their journey."
Pathname	"the location of a file in a hierarchical file system consisting of levels of nested directories, normally described in terms of a starting point (the "root" directory) and the names of all the directories that must be traversed until the local or "home" directory in which the file resides is encountered; the directory names making up the path name are separated in a human readable or machine-readable representation by a character not allowed within the directory names themselves—this may be a "/" (front slash), "\" (back slash), ":" (colon), or other character."
point-to-multipoint	"an arrangement, either permanent or temporary, in which the same data flows or is transferred from a single origin to multiple destinations; the arrival of the data at all the destinations is expected to occur at the same time or nominally at the same time."
QoS	**("Quality of Service")** "for a given data flow or transfer, the set of values of a range of parameters that determines the speed, reliability and other attributes of the data flow or transfer though a network, or part of a network; it is common for the QoS to be known or determined prior to the start and maintained throughout the duration of the data flow or transfer, and not to be changed because of the initiation of another flow within the same network; it is therefore common to refer to a "QoS contract" being set up that applies to all the physical network elements involved in the data flow or transfer."
Streaming	"a data flow which is continuous and at a fixed rate of delivery; contrasts with "file transfer"

E.2 The OSI Reference Model

The terms physical and link layers are borrowed from the telecommunications vocabulary and refer to the 7-layer OSI reference model. This is a useful technique for identifying the interface specifications necessary to provide independence between different aspects of a communications system. As an example, Ethernet can be implemented on several different types of physical infrastructure because the physical layer has no impact on the definitions for network or transport layers. The OSI reference model is shown in Table E1 - the shaded areas indicate the scope of this chapter:

Table E1: OSI seven layer reference model

Level	Name	Description	Unit of exchange
7	Application	User application process and management functions	Message
6	Presentation	Data interpretation, format and code transformation	Message
5	Session	Administration and control of sessions between two entities	Message
4	Transport	Transparent data transfer end-to-end control, multiplexing, mapping	Message
3	Network	Routing, switching, segmenting, blocking, error recovery, flow control	Packet
2	Data link	Establish, maintain, and release data links, error, and flow control	Frame
1	Physical	Electrical, mechanical, functional control of data circuits	Bit

The data transport function descriptions which follow respect the above model.

E.3 File Management and File Systems

The intent of this section is to specify components on which asset management or other applications may be built.

File systems for audio/video servers are similar in concept to normal computer file systems but they do have some unusual requirements in terms of file size, real-time access and reliability. Although some of these issues have been addressed in large scale and real-time systems, the facilities are not available on a wide range of broadcast products. The primary objective is to have a consistent feature set between different broadcast file systems.

E.3.1 File System Attribute Requirements

- It should have consistent file name syntax and semantics. The path and file naming must permit at least 256 characters.
- It should support a read/write permission access mechanism.
- It should support access control based on user-names, passwords and user groups. The chosen mechanism should also support default parameters for users that do not require access control.
- It may support file access logging. The system access log may contain data about file creations, deletions, read operations, write, appends, attribute management functions, etc.
- The file system should support a hierarchical name space.
- Files can be specified by providing a full name path or by setting a working directory and supplying partial path and file name. If remote device access is provided, then the machine name of the remote device must be specified (or a default may be defined or referenced) in the full path.
- Support for remote file systems shared over networks is optional. If it is implemented it must provide a consistent feature set.
- Files of at least 2^{48} bytes in length should be supported. This implies that a single file of length 281 Tera Bytes could be referenced. This limit does not place any restriction on the number of files in a file system(s) or their location addresses. Files are referenced by name and not by absolute address.
- The [content package] at the end of a pathname reference may be a wrapped [file] or some other data type.

E.3.2 File System Commands

This section describes the set of commands that must be provided to manipulate files and navigate the directory system. The local machine (the one the user is connected to) must respond to these commands. If remote machine access is provided then these same commands must be implemented on the remote machine.

- A command that lists the contents of the current directory name space and subtree name space.

- The properties of files and directories must be accessible from the list commands. The properties displayed should include: file size, creation date, last used date, protection attributes and possibly some content attribute information.
- A command set to move, copy, rename, and delete a file.
- A command set to create, rename and remove directories and to move (change directory) within the directory space are required.
- Optional ability to mount and dismount a remote file system and servers must be implemented in a consistent way.
- An optional command to remove files from the active name space, but not actually delete the data, and an expunge command to actually remove the hidden (inactive) files. A command to undelete the hidden (inactive) files is also required

E.3.3 File System Layers

The following layers of file system functions must be allowed. The goal is to have each higher layer be a proper superset of the lower layers.

- A flat tape file system (this is an exception to the requirement of a hierarchical file system)
- A hierarchical tape file system (actually a flat file system with extended name syntax).
- A hierarchical disk file system on the local machine
- A hierarchical disk file system on remote machines

E.4 Digital Asset Transfers

Digital assets fall into several different categories. At least two distinguishable types of asset transfer are needed. One type of asset is the File. The File is any data wrapped with a known cover. See the section on data Wrappers contained in Annex D of this document. The second type of asset is raw data that is streamed from a serving device to one or more receivers. Typically this raw data represents video or audio. A method to transfer Files and to Stream real time video/audio is required. The two main methods of sending assets are (see Table E2):

- Guaranteed-delivery “File Transfer” methods.
- Bounded-quality, asset streaming (“Asset or Content Play”) transfer methods. Digital assets may be streamed from a serving device at user-selectable payload rates for real time or slower/faster than real-time.

The term “Bounded quality” is used for a transfer method that is designed to move the payload from source to destination(s) but without the absolute certainly of true guaranteed delivery. Traditional analogue video is moved with bounded quality in a "asset play” method today. Also, payload data that is moved over the ITU-R BT.656 serial digital interface is moved in a bounded quality way. Usually, bounded-quality links are used to transport streamed, real time content. “Guaranteed delivery” indicates that the entire payload will reach the destination without bit errors, barring a failure of the physical link.

E.4.1 “File Transfer” Requirements for point-to-point Connections

A pure "File Transfer" needs to meet the following minimum user requirements:

- Error-free transfer (by means of using a return path requesting the re-send of corrupted packets).
- The connection between sender and receiver must be bi-directional.
- File Transfer “as fast as possible” (AFAP) must be supported. When the link connecting the source to the destination has unknown payload capacity, the AFAP mode is used. An AFAP transfer only

guarantees that the file will eventually be moved to the destination but the time of completion is undefined.

- User-selectable payload transfer rates for real time or < or > real time must be supported when the links that connect the sender and the receiver support a known and definable payload transfer rate. For example, SMPTE 259M supports a known rate so files moved over this link can be sent at user-selectable rates within the limits of the link, sender and receiver.
- File transfer initiation from either the destination or the source must be supported
- Machine and file access and permission control are required.
- A universal default file transfer protocol (see Note 1) must be provided to accomplish the needs of this application.

Note 1: to be defined

E.4.2 "File Transfer" Requirements for point-to-multipoint Connections

This application is the same as for the point-to-point case above with the following exceptions;

- This file transfer mode is optional.
- If implemented, point-to-multipoint must be supported with at least 32 simultaneous receiver nodes.
- The file transfer protocol may be different from that used for the point-to-point case.

E.4.3 "Asset Play" Requirements for point-to-multipoint Connections

An "Asset play" has to meet the special user requirement of "playing" video, audio, data etc. in a television production environment. Because re-sends of video data (and the resulting delay or latency) caused by corrupted data cannot be tolerated, some errors must be accepted and a return path must not be used. The actual payload to be streamed might not be a file but the raw video/audio assets. For example, streaming MPEG2 compressed data over a serial digital interface link (with appropriate data framing to carry compressed data) at real time rates is a typical application.

Other important unique television production requirements are:

- Bounded quality based on the QoS of the link. The Rate, Delay, Jitter and Loss of the link determine the quality of the streamed assets.
- Links that support the streamed assets must be unidirectional in nature.
- There is no AFAP mode as required in the "File Transfer" case above.
- Streaming rates may be at real time or faster or slower than real time with real time being the default mode.
- The machine control methods for the sending device are not specified by this document.
- File access and permission control are required.
- Standards need to be provided for carrying payload data in the connecting link(s); e.g. how to transport MPEG2 data as payload in a serial digital interface link?

Browsing, which is also a "content-play" application, has requirements which differ from the "on-air" content-play application in the following respects:

- error rate, bandwidth and delay of the links may not be well defined.
- Point-to-point is typical. Point-to-Multipoint is not required.
- Access/permission controls are required

E.5 Transport mechanisms

There are many choices and tradeoffs available in the transport of data and the suitability is based on the particular application. The type of transport is in general defined by the quality and cost constraints of the application. The quality aspect is determined primarily by the video type. The higher the quality and performance requirements, the higher the interface and physical layer requirements.

The transport mechanism is used to build the link between the data to be transported (i.e. the file format) and the physical layer. It has to provide attributes such as bandwidth reservation and synchronization as required by the application (e.g. content play requires synchronous delivery).

Two families of transport mechanisms have been identified:

- serial digital interface (ITU-R BT.656) based mechanism
- network-based transport mechanism

A clear user requirement is the interoperability in each of the transport mechanism families by themselves and between the two families (see 1.3.6).

The following sections describe requirements which are necessary to meet the users' expectations on transport mechanism and to achieve interoperability in the transport layer between different systems.

Requirements need to be defined in the areas of:

- synchronization
- transfer modes and quality of service parameters
- transfer initiation and transfer phase parameters
- transfer interaction management and error reporting
- basic interworking protocols

A method is presented for categorizing QoS parameter values according to the application.

E.5.1 Synchronization of associated data (e.g. Audio-Video-Meta-Data)

Maintaining the synchronization of audio and video and associated data (e.g. Time Code) is a critical issue for streaming, isochronous and synchronous transfers. Users require that future implementations of both SDI-based and network-based transmission should meet the specifications described in ITU-R BT.XXX (or EBU Tech. R37).

E.5.2 Transfer modes and QoS parameters

In a mainstream television production environment three types of time relationship (depending on the application as described later) between the source data-clock and the received data-clock of a signal are necessary. These are: synchronous, isochronous and asynchronous transmission in both bounded and guaranteed-quality transmission.

The synchronization aspects which are of importance for professional television transmission differ greatly in some respects from those used in the computer industry.

These time relationship aspects combine with two main methods of transfer described in section 1.2 to produce six combinations applicable to a mainstream television production environment. These are shown in Table E2.

Table E2: Transfer modes

	guaranteed delivery	bounded quality
isochronous	B1	A1
synchronous	B2	A2
asynchronous	B3	A3

The notation A1 through B3 is used as a shorthand for these six modes and it is used in the examples and in Table E8, below.

E.5.3 Transfer initiation parameters

Essential parameters during the initiation or pre-roll phase of a transfer (independent of the transfer modes) are:

- response time for the interaction (e.g. press/click on play, copy etc. and the delay until the system confirms the action)
- set-up-time (e.g. time required for the establishment of a connection-oriented transmission over, for example, a Fibre Channel or ATM link).

Suggested ranges of values for these parameters are shown in Table E3.

Table E3: Transfer initiation parameters (Note 1)

Identifier	0	1	2	3	4
Parameter					
Response time	Unspecified	≤3s	≤1s	≤0.25s	≤10ms
set up time	Unspecified	≤1s	≤100ms	≤50ms	≤5ms

Note 1: in typical applications the response time and the set-up-time will differ from the first initiation of, for example, a connection over a network and any subsequent requests over the same connection. The numbers shown in the table are maximum values for the first initiation of the process.

The parameters are carried forward to Table E5, below, as part of an example.

E.5.4 Transfer phase

Once a transfer has commenced, a defined transmission quality must be guaranteed. Parameters which determine this are: bit-rate, delay, Bit-Error-Rate (BER) and jitter.

These parameters are specified in Table E4 for end-to-end applications and are referred to as Quality of Service (QoS) parameters.

Table E4: Transfer phase parameters and values

Identifier	0	1	2	3	4
Parameter					
Bit-Rate (Note 1)	Available Bit-Rate	> 10 Mbits/s	>50 Mbits/s	≥ 270 Mbits/s	≥ 1.2Gbits/s
Delay	Unspecified	≤ 1 s	≤ 500 ms	≤ 250 ms	≤ 20 ms
BER better than (Note 2)	10^{-6}	10^{-8}	10^{-10}	10^{-12}	0
Jitter (for Wander see Note 3)	Unspecified	≤0.6 UI	≤0.4 UI	≤0.2 UI (see SMPTE 259M)	≤0.135 UI (see SMPTE 297M)

Note 1: Bitrate depends on the application. For example, if a faster than real-time transfer is required interfaces with higher bit rates may be used. For example, a packetized studio interface with a transfer data rate of 50Mbit/s is unlikely to be used. 4 x transfer would lead to data rates in the 200 Mbit/s region which would be more applicable.

Note 2: when BER = 0 is required by the application a resending capability is required

Note 3: Wander is a serious problem for video transmission over wide-area networks and is presently under consideration by SMPTE

The parameters are carried forward to Table E5, below, as part of an example.

E.5.5 Example of the specification of a class of transfer

According to the previous definitions the transfer of data can be specified as shown in Table E5.

Table E5: Specification example

Application	Transfer modes	Transfer initiation parameters	Transfer phase parameters
Live interview between two performers in the studio	A2 (synchronous, bounded quality)	4.1 (response-time<10ms, set-up-time<1s)	3.4.3.3 (bit-rate≥ 270 Mbits/s, delay ≤ 20 ms, BER≤10^{-12}, jitter=0.2UI)

See also Table E8 for examples of transfer initiation and transfer phase parameter usage

E.5.6 Transfer interaction management, exception reporting

The transfer interaction management usually runs in the background of the system and is not visible to the user.

Some functions required of this interaction management are:

- establishment, maintenance and release of the data link
- the framing, mapping of data to be transmitted
- error control
- flow control

However the system must report non-recoverable exceptions to the user, for example, as shown in Table E6.

Table E6: Exception Reporting

	SDI based transport	Network
Alarm when Link-BER degrades to:	≤10^{-10}	degrades more that 10^{-2} from the QoS agreed on.
Synchronization errors	bit, character, frame, picture	bit, character, frame, picture
Bandwidth degradation	unlikely	due to traffic management error
Priority change	due to traffic management error	due to traffic management error
delay change more than	10% from QoS value	10% from QoS value
Recovery time after a break of 50ms exceeds	500ms	500ms

E.5.7 Basic transport protocols

For content transfer via the serial digital interface (ITU-R BT.656) in packetized form, users require a single open transport mechanism. A standard for such a transport mechanism is currently under development in SMPTE and EBU. This transport mechanism must be implemented in products according to the user requirements and specification made above.

For content transfer via a network, users require at least one open standardized baseline protocol such as IP. However existing protocols have not been developed to meet the mainstream television production requirements; therefore optimized protocols which meet the user requirements defined above need to be developed and standardized.

Interworking between the serial digital and network-based transfer protocols is a strong user requirement and open standardized gateways between them need to be defined.

E.6 Physical Interfaces

E.6.1 Physical layer

This layer deals with the physical medium used to provide the means to transmit data between devices as well as defining the interface between devices.

Examples of physical media include:

- Twisted pair
- Coaxial cable
- Optical fiber
- Radio

For example in the 270 Mbit/s and 360 Mbit/s Serial Digital Interface (SDI) used for serial digital video, the physical medium is coaxial cable with 75-ohm characteristic impedance.

Details of the interface include:

- Mechanical Type of connector
- Electrical (voltage levels, bit rates, rise times, distance requirements)
- Functional (Data, control, timing)
- Data coding

Similarly, for the ANSI/SMPTE Recommendation 259M, the mechanical requirement defines a 75-ohm BNC connector. The electrical requirements define an unbalanced circuit with signal levels of 800 mV p-p across a 75-ohm load with rise times between 0.75 ns and 1.5 ns. Bit rate for 4:2:2 standard video is 270 Mbit/s.

Functionally the data word length is 10 bits with the LSB of any data word transmitted first. Flow control is not required due to the streaming nature of video data.

Examples of existing physical interfaces are given in Table E7.

E.6.2 User Requirements for Physical Interfaces

It is recognized that the choice of physical and link layer implementations is restricted by the choice of network. The choice of network is itself dependent on system and application requirements such as QoS and bit rate.

Given these restrictions, the following user requirements have been identified:

- The operation of the interface must be invisible to the system user.

- Network interfaces for content transfer must be chosen from the available industry-standard interfaces. At the present time these are considered to be:

- IEEE 802 Ethernet (see Note 1)
- ANSI X3T11 Fibre Channel (see Note 1)
- ATM (see Note 1)
- IEEE 1394 (Firewire)
- SMPTE [ ] SDT (the SMPTE [ ] Serial Data Transport compatible to SMPTE 259M)

Note 1: Not all classes of service, particularly real time video transfers, are available at the present time

- Implementations must comply with all relevant aspects of the chosen network interface.
- Connectors must be of appropriate robustness and durability for a professional broadcasting environment. Where, for a particular network interface, a choice of connector exists, the chosen connector should ensure the maximum level of compatibility between equipment
- It is desirable that equipment network interfaces be available in a variety of forms, to suit the industry-standard networks listed above.
- It is desirable that an equipment interface be capable of upgrading to take advantage of higher network bandwidths as these become available.

Table E7: Physical Interfaces

	Electrical Interface		**Optical Interface**	
	point-to point	Network	point-to point	Network
	transmission according to ITU-R BT.656 should be transmitted according to SMPTE 259M		transmission according to ITU-R BT.656 should be transmitted according SMPTE 297M	Network Multi-Mode Fibre Interface according to ISO/IEC 11801 Single Mode Fibre Interface according to ITU-T G957
Bit rate	see Note 1 (>270Mbit/s)	Application-specific, see Note 1	see Note 1 (>270Mbit/s)	Application-specific, see Note 1
Delivered error rate after error correction	better than 10^{-12}	better than 10^{-12}	better than 10^{-12} for LED based MM for SW Laser based MM for Laser based SM	better than 10^{-12} for LED based MM for SW Laser based MM for Laser based SM
Medium	coaxial cable 75Ω nominal impedance	a) coaxial cable 75Ω nominal impedance; b) Cat 5 UTP	Fibre 62,5/125 micron for MM 8-10/125 micron for SM	Fibre 62,5/125 micron for MM 8-10/125 micron for SM
Interconnection length	length up to 250m	application dependent	MM up to 500m SM: short reach up to 2km intermediate reach up to 15km	MM up to 500m SM: short reach up to 2km intermediate reach up to 15km
Transmitter and receiver characteristics	as defined in ITU-R BT656-4	network specific	as defined in SMPTE 297M	as defined in ITU-T G957 for single- and multi-mode
Jitter	defined in SMPTE RP 184 and 192	network specific	defined in SMPTE RP 184 and 192 for SDI based transmission and in ITU-T G958 for	as defined in ITU-T G958 (see Note 2)

			networks	
Connector	standard BNC type (IEC Publication 169-8), and its electrical characteristics should permit it to be used at frequencies up to 850 MHz in 75-ohm circuits.	a) BNC/TNC b) RJ45 c) IEEE-1394 d) DB-9	SC types as defined by IEC874-14 and IEC 11801	SC types as defined by IEC874-14 and IEC 11801

Note 1: bitrate depends on the application. E.g. if a faster than real-time transfer is required, interfaces with higher bit rates must be used.

Note 2: After buffering and jitter removal the residual jitter must meet that specified in SMPTE 259M.

ITU-T G957: optical interfaces for equipment and systems relating to the synchronous digital hierarchy

ITU-T G958: digital line systems based on the synchronous digital hierarchy for use on optical fiber cables

E.7 High Level Management Functions

Due to the complexity of networked systems with multi-user access and the involvement of interconnected server-based systems, further high level operational user requirements must be provided by the system management. These include, for example:

- Mission critical network requirements:

  Data security/shadowing/backup must run automatically in the background

  Localized failure recovery where possible

  No concentrations of failure points

  Automatic failure condition notification and logging

- Distributed Resource Management

  The distributed resource manager has to provide all the functions which are necessary to make the transfer between systems possible and which parameters are necessary for the transfer to be accomplished (and inform the user about exceptions if necessary).

  Examples are:

  - provide information about the slowest/fastest rate any destination can accept, (e.g. whether the intended bit rate is higher than the destination can accept), and what storage is available at a particular destination

  - allow a transfer with a certain speed and a certain file size in, for example, a point-to-multipoint topology

  - perform payload compatibility check (e.g. DV-to-MPEG) and notify transfer initiator

  - determine whether a prior negotiation is needed

  - updating pathname of associated file(s) automatically after move operation

- Plug and Play (hot plugging) and operating system issues

  This is defined as dynamic resource detection without reboot. For example, if a new server is connected to an existing server environment, the system management must detect and install the new system automatically with appropriate messages to the user and with no disruption to ongoing operations during the reconfiguration. This also requires multi-operating system support in order to achieve an easy data exchange and interoperability in the operating system and file association domain.

- Control interfaces

  Machine control commands need to be openly defined in order to support multi-vendor equipment.

E.8 Application example

Table E8 gives examples of different applications and the required QoS parameters.

Table E8: Application examples and QoS parameters (see also Table E5)

	Application example	Transfer mode (Table E2)	Transfer initiation (Table E3)	Transfer phase params. (Table E4)	Example Data Types	Notes
1	**Hard real-time live multicast** (e.g. interaction between performers, live 2 way interview, people and machines)	A1	4.4	3.4.3.3	601, DV422, MPEG422 + Audio +Ancillary +time code +metadata	link is a channel
2	**Isochronous real-time delivery with delay** (e.g. Program distribution channel) [def. Required contribution versus distribution]	A2	3.3	3.3.3.3	601, DV422, MPEG422 + Audio +Ancillary +time code +metadata	link is more a channel
3	**Selectable transfer rate** (e.g. 4 times transfer, 200 Mbit/s moves (50 Mbit/s) at 4x play speed	A3	2.2	3.2.3.2	601, DV422, MPEG422 + Audio +Ancillary +time code +metadata	transfer is As Fast As Possible (AFAP)
4	**File transfer with guaranteed delivery** (e.g. digitized content and/or metadata transfer)	B3,B2	2.2	2.0.4.0	data	transfer is AFAP and uses self routing
5	**File transfer “bounded quality/ multicast version”** (e.g. video email over cable modem)	A3	2.2	2.0.3.0	data	application still evolving; time/transfer is not "an issue".
6	**As 5, but point to multipoint** (e.g. digitized content and/or metadata transfer)	A3	2.2	2.2.3.0	data	time/transfer is "an issue" (related to cost of transmission)
7	**As 5, but at selectable rate** (sending a payload within a specified amount of time)	A3	2.2	2.2.3.0	data	time/transfer is "an issue" (related to cost of transmission)
8	**Browsing** There is a UR for browsing acquired material.	?	?	?	data	requires further study

E.9 Wide Area Network Gateway

Wide Area Networks (WAN) used to interconnect facilities for production and distribution present additional diverse and difficult challenges beyond those of Local Area Networks. The QoS of WANs especially in the areas of latency, rate, jitter, BER require special equipment to interface the facility to the WAN, i.e. through a gateway. The gateway not only has to adapt the QoS of the WAN to match the facility requirements stated earlier but must also perform protocol and signal translation as needed. This includes segmenting and addressing bit streams (digital video and audio) as well as self routing capabilities so they may find their way through a WAN. Protocols on the WAN side of the gateway must adhere to national and international conventions. Gateway and network management services need to be included for set-up, accounting and control including security. A co-ordination intercom may be needed for set-up and maintenance. Discussion of the myriad of protocols is beyond the scope of this document. A concept diagram (Figure E1) is attached showing a possible gateway configuration.

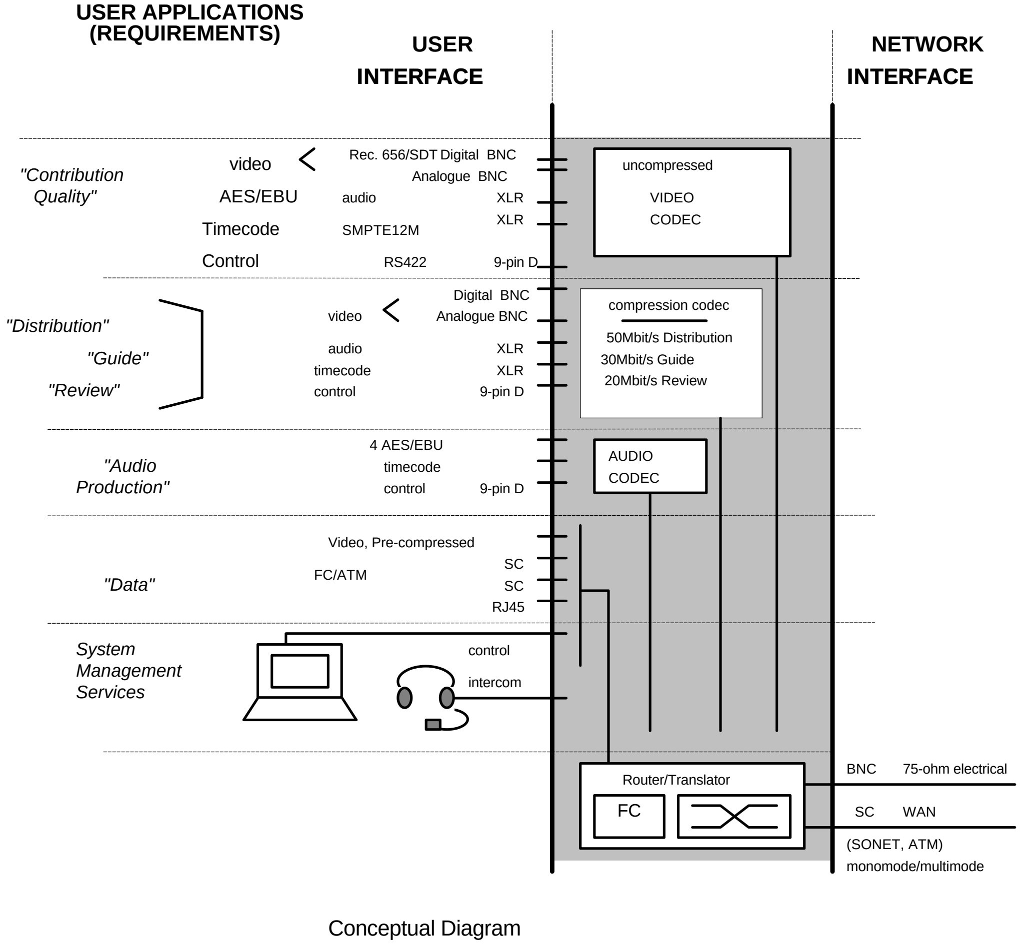

Figure E1. Wide Area Network Gateway

E.10 User Requirements for standards

In order to achieve interoperability between different vendors, which is a basic user requirement, the following standards need to be defined:

- File Formats for guaranteed and bounded quality transfer
- Link Protocols: e.g. protocols for different classes of service, transfer mode commands, transfer initiation commands, transfer phase commands, the format of data being interchanged, number of bits per element, type of encoding scheme, synchronization information
- Flow Control: methods to control the flow of data between two devices (e.g. to avoid overflow of storage at the receiver)
- Error Control: e.g. type of error detection and control
- Inter-Layer commands: e.g. error messaging between different layers
- Machine Control
- Networks and interfaces
- Payload and mapping documents
- User interfaces
- Platform independent command sets for file management

Annex F – Task Force Participants

Following is a list of the individuals who participated in the deliberations of the Task Force and their company affiliations.

Ackermans	Merrick	Turner Entertainment Networks
Agoston	György	MTV (Hungary)
Ardito	Maurizio	RAI
Bancroft	Dave	Philips BTS
Bernard	Frank	Silicon Graphics Inc.
Birkmaier	Craig	PCUBE Labs
Boch	Laurent	RAI
Böhler	Per	NRK
Bradshaw	David	BBC
Bratton	William F.	Turner Broadcasting System
Bryden	Neil	ITV
Coley	Bryan	Turner Entertainment Networks
Dabell	Adrian	Panasonic
Dare	Peter	Sony
Davis	Franklin	OMF / Avid
Dimino	Giorgio	RAI
Duca	Jim	Pluto
Ebner	Andreas	IRT
Edge	Bob	Tektronix
Ferrante-Walsh	Lisa	OMF / Avid
Fibush	David	Tektronix
Gaggioni	Hugo	Sony
Garsha	Chuck	Paramount Pictures
Girod	Carl	SMPTE
Haycock	David	Utah Scientific
Hedke	Rolfe	FH Weisbaden
Heitmann	Jürgen	Philips BTS
Hoffmann	Hans	IRT
Hori	Akihiro	NTV
Ide	Akifumi	Matsushita
Ive	John G.S.	Sony
Iwasaki	Yasuo	Sony
Jorgensen	Chris	NBC
Kaiser	Martin	RBT
Kisor	Robert	Paramount Pictures
Knör	Reinhard	IRT
König	Peter	NDR
Kovalick	Al	Hewlett Packard
Legault	Alain	Matrox
Liron	John	Tektronix
Liu	John C.	Silicon Graphics Inc.
Livingston	Philip	Panasonic
Long	Stephen W.	TASC Inc.*
McMahon	Tom	Del Rey Graphics
Mihara	Kanji	Sony
Miles	Roger	EBU
Miller	William C.	ABC Television
Minzer	Oren	ECI Telecom
Morgan	Oliver	Sony
Morioka	Yoshihiro	Matsushita
Nagaoka	Yoshimichi	JVC
Nakamura	Shoichi	NHK
Neubert	Neil	JVC
Notani	Masaaki	Matsushita
Ostlund	Mark	Hewlett Packard
Owen	Peter	Quantel
Penney	Bruce	Tektronix
Peters	Jean-Jacques	EBU
Pittas	John	Seachange International
Rahman	Altaf	Philips BTS
Rodd	Spencer	Pro-Bel
Roe	Graham	Pro-Bel
Romine	Jeff	Philips BTS
Sadashige	Koichi	NTA
Safar	Johann	Panasonic
Schachlbauer	Horst	IRT / EBU
Schaefer	Rainer	IRT
Shishikui	Yoshiaki	NHK
Slutske	Bob	National TeleConsultants
Smith	Clyde	Turner Entertainment Networks
Streeter	Dick	CBS
Tanaka	Masatoshi	Matsushita
Tobin	Bill	Sun Microsystems
Turow	Dan	Gennum
Weiss	S. Merrill	Consultant / SMPTE
Wendt	Hans	JVC Germany
Wilkinson	Jim	Sony
Wode	Gary	Wode Designs
Yagi	Nobuyuki	NHK

* Representing the National Imagery and Mapping Agency of the U.S. government.