MPEG-2 and SONET

Michel Fortier, Ph. D., eng.
Northern Telecom,
9300 TransCanada Hwy, St. Laurent, PQ, H4S 1K5 (Canada)
Tel.: (514) 956-3473 ♦ E-mail: michel.fortier@nt.com

Abstract

MPEG-2 can be used in many services such as video trunking, pay-per-view (PPV) and Near Video on Demand (NVOD). As MPEG-2 technology becomes commonplace in the cable headend, SONET technology plays a key role in backbone transport of digital content.

With SONET, MPEG-2 can be transported using ATM or non-ATM modes. Both in the core and access network, SONET equipment will be standard. End-to-end connectivity is possible on SONET networks that support both modes. Thus SONET can be used as the transport technology of choice to carry MPEG-2 to interconnect service providers to users with exciting new services.

1 - Introduction

Over recent years, the means for handling and delivering video signals has changed dramatically. Digital video signal processing is now common practice, and is part and parcel of all aspects of producing video programs. Until recently, programs to be used in distribution have been kept in uncompressed format, analog or digital. However, it is now possible to store video programs in digital compressed form. The degree of compression is usually based on the end use - either for long term storage and use in future productions, in which case the resulting picture quality must be of high quality and the compression ratio lower, or for use in server-based services where storage and eventual bandwidth delivery requirements will put constraints on the compression level, and hence, the resulting picture quality.[1]

In today's digital world, a gamut of services exist:

- Cable distribution — In upgrading networks, the cable operators are looking into digital access and distribution technologies to enhance the quality and robustness of the signals to be delivered over long distances, as well as investigating video compression technologies to enable new delivery mechanisms.
- Telco video distribution — With the new regulation in place allowing the telcos to provide service, operators are looking at ways to use broadband channels to carry digitized video signals in their networks.
- Video trunking — Whether for planned services, or for "just in time" services, video networks are being deployed to provide points of presence in many venues (studios, sports centers, convention centers, etc.) which allow the network operator to haul the video signals either to the processing centers, i.e., the broadcast studios, or to other end-points.
- Conversational video services — These services have been addressed until now by H.261 codecs and the higher bit-rate DS3 codecs. Cost and picture quality are of prime concern. MPEG-2 offers opportunities of providing better quality pictures with reasonable bandwidth for premium video services.

Accessibility to the network and the transport mechanisms are the key factors for the operators to establish networks which achieve end-to-end connectivity for video signals in their many formats. This paper presents a view of this connectivity from the perspective of SONET because of its emerging pervasiness as a standard in digital telecommunications networks. The discussion will also show the interworking with digital video, with particular emphasis on MPEG-2 video.

1 Note that the decoded picture quality can be quite good; degradations can even be undetectable. This will depend on the amount of compression used.

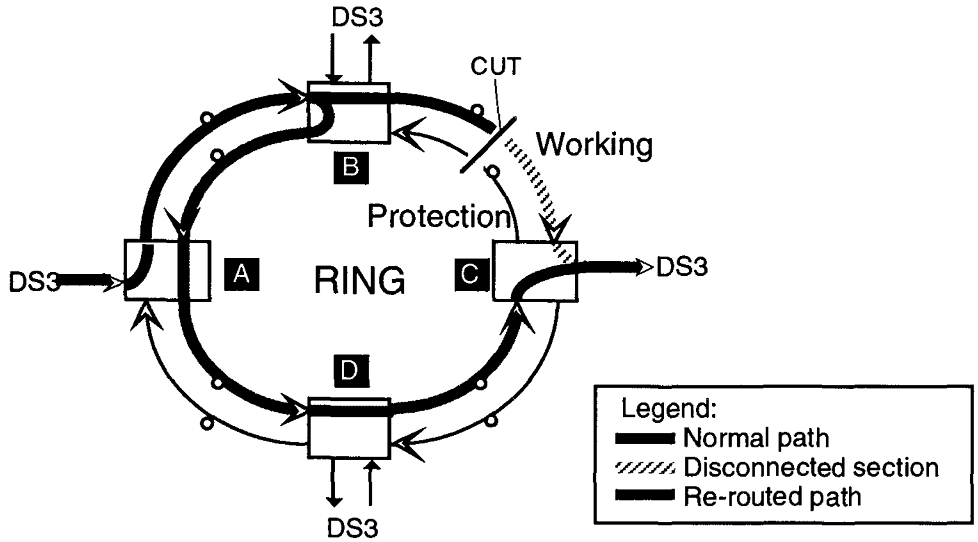

Figure 1 – Example of Bi-directional Line Switched Ring

2 - End-to-end Connectivity
What does SONET offer?

SONET (Synchronous Optical NETwork) is a digital optical transmission infrastructure which has a hierarchy of transmission speeds, or bit-rates – i.e., OC1, OC3, OC12, OC48, OC192. It supports many kinds of channel protocols at the data level. At the physical layer, ATM uses SONET transport at various bit-rates, e.g., OC3c, OC12c, etc. Furthermore, DS3, DS1, and DS0 signals can be mapped into the SONET infrastructure.

Within the SONET infrastructure, there are built-in mechanisms to perform network management as well as carry the related commands using standardized protocols. This network management covers all aspects of the OAM & P process.

SONET being an optical and digital network, the transmission is inherently robust and it can be carried over much longer distances when compared to analog circuits. Moreover, extra fibers can be laid in for redundancy and the reliability of the network can be made very high. The SONET equipment provides for switching time slots on the terminal points, and automatic reconfiguration on failures of ring architectures. For instance, using techniques like bi-directional line switched rings, the network can be made tolerant to single fiber cuts. When activated, this technique re-routes traffic from normal working links to protection links which carries data around the ring in the opposite direction to reach the otherwise disconnected network element. This is shown in Figure 1 where traffic was going from node A to node C. A cut between nodes B and C results in traffic then going from node B to node A via the protection path, then to node D, and finally to node C.

This type of network is becoming all widely accepted in the world of telecommunications. The transport hierarchy equivalent to SONET in countries other than Canada, the United States, Japan, Korea and Taiwan is called SDH (Synchronous Digital Hierarchy). SONET and SDH are compatible with the exception of minor channel maintenance differences. Hence, video signals can then be distributed across continents, notwithstanding the issue of transcoding the video signals from one standard to another, whether compressed or not.

3 - MPEG-2 technology – Technical issues

Most transport of digitized video signals requires compression at some point. This stems from the need to optimize investment in back-bone infrastructures. MPEG-2 is a compression technology which can be applied in many parts of a distribution system. Its applicability will be governed by the user's quality of service criteria. Key technical issues are:

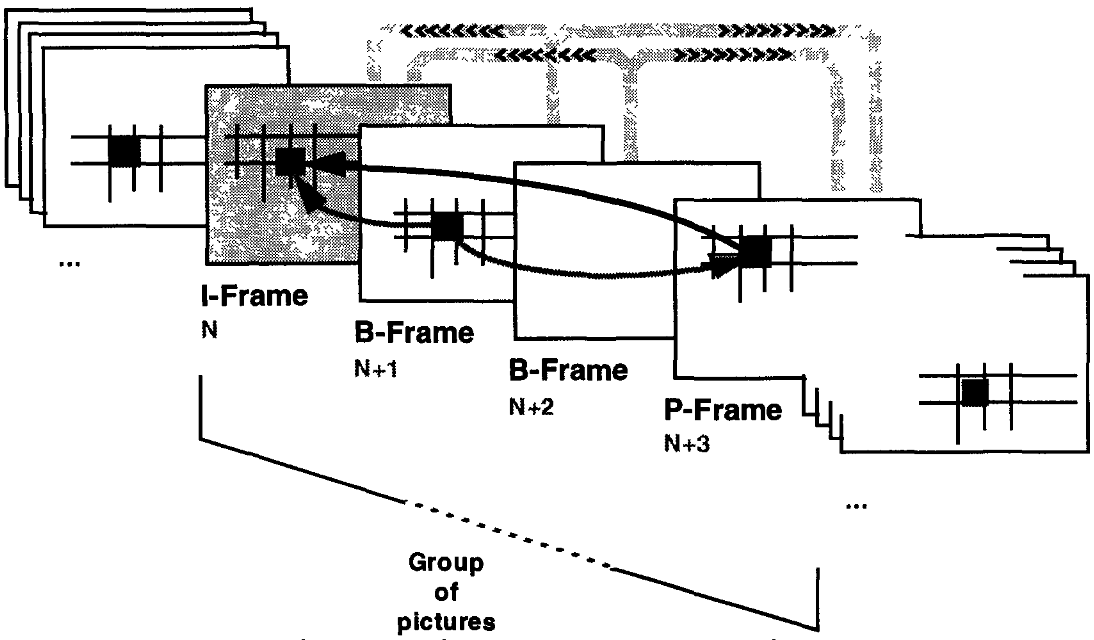

Figure 2 – I, B, P frame relationships

- Bandwidth versus Picture Quality
  MPEG-2 spans a range of bit-rates between 3 and 15 Mb/s. Picture quality in this range can vary between average to very good, and this will depend to some extent on the types of pictures being compressed (e.g., sports scenes, "regular" movies, training material, etc.). Typically, many MPEG-2 streams use the 4-9 Mb/s range, where 4-6 Mb/s may be used for regular and training material, and the higher end, 6-9 Mb/s, will be used for sports, as well as advertising material for insertion into programs.

- End-to-end Encoding-Decoding Latency
  In one-way video services, this aspect is not usually a factor. This latency appears as a combination of the codec's (coder-decoder) use of an IBP coding mode along with the so-called video buffer found in the codecs. The IBP mode refers to the way coded pictures are referenced. This process can be easily visualized with Figure 2.

  I-frames, or intracoded frames, refer to frames that use only information within the current frame being encoded. P-frames, or predicted frames, are differentially coded frames using prediction from previous frames in the coding process which may be I- or P-based frames. This prediction uses motion estimation to optimize the encoding. The B-frames, or bi-directional frames, operate on bi-directional interpolation. This process uses reference information in past and "future" frames to form the best interpolation possible. The overall technique is such that I-frames when coded produce more bits than P-frames which, in turn, produce more bits than B-frames. As a rule of thumb, the I:P:B ratio can be seen as 20:5-to-10:1 but it is far from static. Factors affecting the bits-per-frame production include scene changes, highly textured areas, moving object, etc.

The combined use of all frame types produces the best compression (e.g., the lowest bit-rates) for a given decoded picture quality. However, the IBP mode produces the greatest delay. In general, there are two B-frames between reference I-P, P-P, or P-I frames which accounts for an intrinsic delay of two frames; from the point of reference of a B-frame, the future I- or P-frame must be delayed appropriately before it can be used. This is then combined with the video buffer which must be filled to an operating level to avoid overflows/underflows. The size of the buffer translates into an equivalent delay at a given bit-rate. This can be of the order of ~200 ms. This is further compounded by the video processing delays which can then produce overall delays that are much longer than 200 ms. — in some cases several hundred milliseconds.

As indicated before, one-way video services would not have problems with this type of delay other than synchronization issues where multiple streams need to be integrated from real-time sources. Again, this would be taken care of via buffering mechanisms. For two-way conversational video services, I-frame only coding as well as IP-frame coding provide the best delays because they do not need information from "future" frames and the video buffers can be made to introduce as little delay as possible.

MPEG-2 has techniques for timing recovery to help reconstruct the stream at its original bit-rate. This allows for display and synchronization of the far end operating frequency to be linked back to the source. This timing recovery can be done in ATM and non-ATM working environments.

An MPEG-2 Bitstream can contain audio, video, and other ancillary streams. These are initially formed into packetized elementary streams (PES). Then, a multiplexer engine creates a stream from these PES streams in one of two formats: program streams (PS) and transport streams (TS). The PS format is to be used in "error-free" contexts while the TS format is to be used mostly for error prone environments (e.g., SONET transport or ATM). However, there are no specific rules in their use, only their intrinsic adaptation to certain environments.

While it is true that the TS format and AAL5 can be used in ATM environments, other means are possible. Carriage of streams can also be done using DS3/E3-level signals. Unfortunately, the standards bodies have not addressed such transmission channels. Users will want interconnecting networks, i.e., streams coming into, and multiplexed by one vendor's equipment into a DS3 for instance, should be capable of being demultiplexed by another vendor's equipment. Flexible multiplexing and demultiplexing will be required.

Nevertheless, SONET inherently transports either mode of transfer, non-ATM and ATM. The ATM mode offers advantages with respect to giving operators the possibility of transferring streams composed of single or multiple program streams into one transport stream. With the latter, the end-user can switch from one program to another within the same transport stream. The user can also switch from one transport stream to another. Non-ATM transfers can be used in point-to-point connections, and also in drop-and-continue scenarios. Point-to-point refers to users who have a link established with a remote point, whether permanent or temporary. Drop-and-continue refers to the capabilities offered on SONET rings to drop the signal at the network element and transport the same signal on the ring to be dropped at the next network element, and so on. This capability offers a mechanism for broadcasting around the ring. Resources can also be reallocated as needed with switching functions. The maximum number of network elements in the ring is sixteen to preserve switching times in case of failure as set by Bellcore standards.

4 - MPEG-2 and SONET

Integrated video networks – The bandwidths of channels required to transport data of all kinds are becoming larger. Voice, data, and video networks can be merged together and managed as one, whether internally the channels are mixed or segregated. Hence, the network operator need not contend with overlay networks which, in turn, provides efficiencies, e.g., economies of scale. All services are carried on one SONET network therefore reaping the benefits of economies of scale, maintenance, staff training, system integration, and others. MPEG-2 sources can therefore be distributed across networks in a transparent fashion. This applies whether transporting PS or TS formats in non-ATM or TS in ATM environments.

• Hierarchical Accesses to Servers

In setting up networks for services like video on demand (VOD), or near video on demand (NVOD), or access to multi-media servers, a hierarchy will be set up in the network. Let us address one view of a NVOD scenario. In the example shown in Figure 3, a two-level hierarchy is displayed. In the first level, source material can be gathered and stored on global servers, holding program and advertising contents which are to be used in later programming. In the second level, the local servers get data from the global servers and are replenished according to scheduled downloads which provide the contents for daily programming. The output of the servers goes to the home via the RF distribution plant in one of two ways:

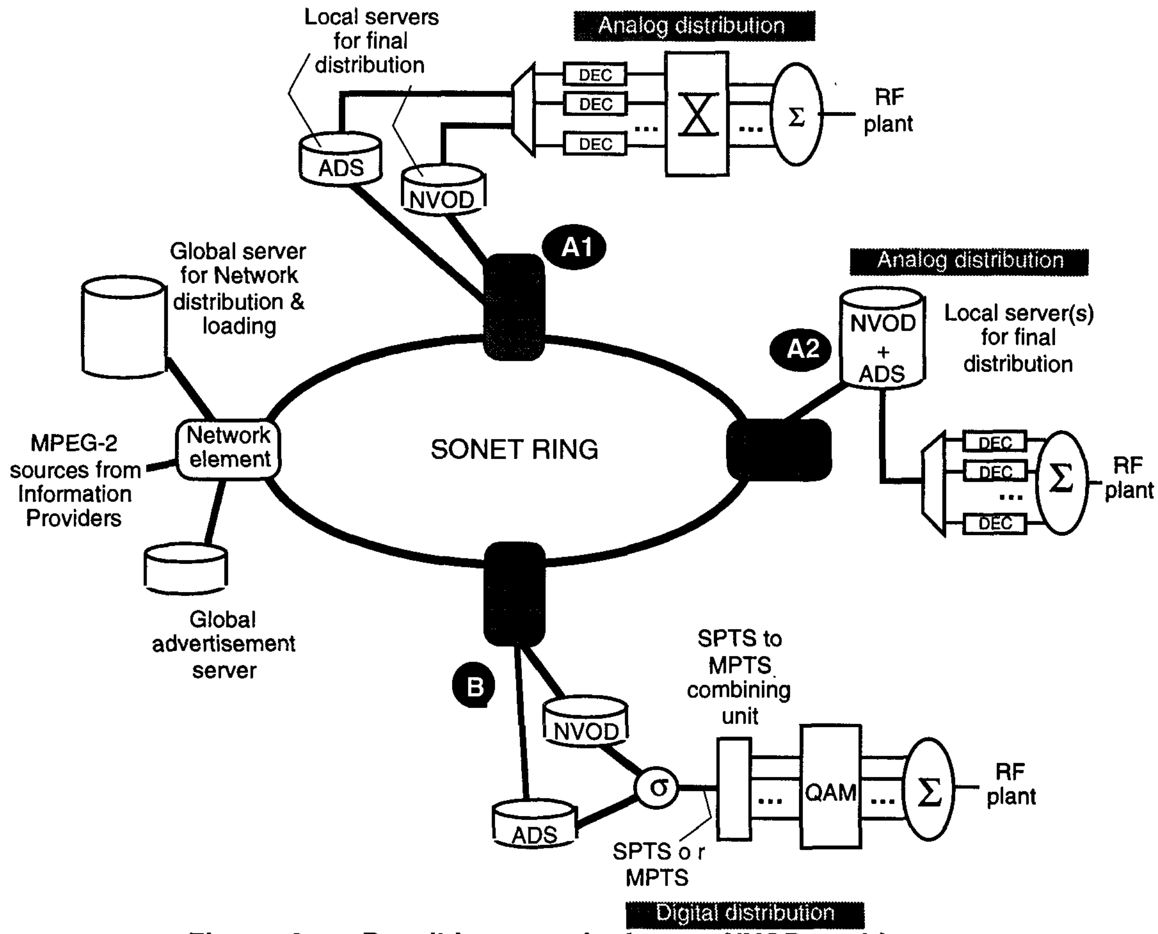

Figure 3 — Possible scenario for an NVOD architecture

A - analog video signal to the set-top box, or

B - a digital stream to the set-top box.

Method A can be done in two ways. One, shown by the reference mark A1, would be to have separate servers for the program and the ads. Both servers send data to a demultiplexer for distribution to a bank of decoders. The decoders' outputs then go through a switch to assign video signals to a modulator before going to the RF plant. In this way, a program runs until a commercial needs to be inserted. Then, the appropriate signal from the commercials' server is routed by the switch to the right modulator.

The second method, marked A2, stores the programs and the ads on the same server, eliminating the need for an analog video switch. In this technique, a program runs until the time for the commercial. Then, the server is instructed to get the appropriate file for the ad.

In Method B, similar scenarios can be used, separate or combined program/ad servers. The difference is that the streams are first put through a stream combining unit which multiplexes multiple MPEG-2 streams to form a multiple program transport stream. The output of these stream multiplexer units are used as input to QAM modulators before entering the RF plant, i.e., the streams are sent in digital format to the end-user. At the user's home, a set-top box receives the QAM-formatted signal, extracts the bitstream, and then, demultiplexes the appropriate program for decoding by the set-top box decoder.

5 - Interworking of non-ATM and ATM video networks

Presently, there is a need for interworking scenarios between ATM and non-ATM networks. The current networks have not all converted to ATM operations and there will be a transitory phase where streams will either originate in non-ATM format or will cross network boundaries.

Two kinds of migration scenarios can exist to migrate from non-ATM to ATM transport: (i) using ATM to emulate non-ATM circuits, and (ii) taking the non-ATM bitstream into an adaptation interface to convert it into ATM cells. In the first case, the SONET ATM network transports the constant bit-rate data without mapping the MPEG-2 stream into ATM. At the other end of the network, it then reconstitutes the constant bit-rate stream. In the second case, an MPEG-2 bitstream is converted into ATM cells using an AAL5 protocol. At the other end, the cells' payloads are demultiplexed to form a bitstream to be delivered to the end user, or to continue on in the network.

A method is also needed to provide ATM to non-ATM adaptation between networks. Transporting ATM cells via DS3 UNI circuits is one method to adapt between networks. At the receiving end, the cells can then be extracted and put back into an ATM network to continue on to their destination. An alternative is for the cell's payload to be converted with interface units into a constant bit-rate stream. This alternative can only be utilized if the bit-rates selected for transport on the ATM network are well adapted to the bit-rates used in the non-ATM networks. In other words, the non-ATM network may well use interfaces for MPEG streams that handle only specific bit-rates, for instance based on DS2 granularities. Then there will issues in trying to convert an MPEG-2 stream at say 4.5 Mb/s. Either, the interfaces will not accept such bit-rates, or stuffing bits will be needed and a varying degree of network inefficiency will result depending on the proportion of stuffed bits relative to the payload. Flexible interfaces will be required, but more importantly, a standardized way of multiplexing into one or various non-ATM networks will be needed.

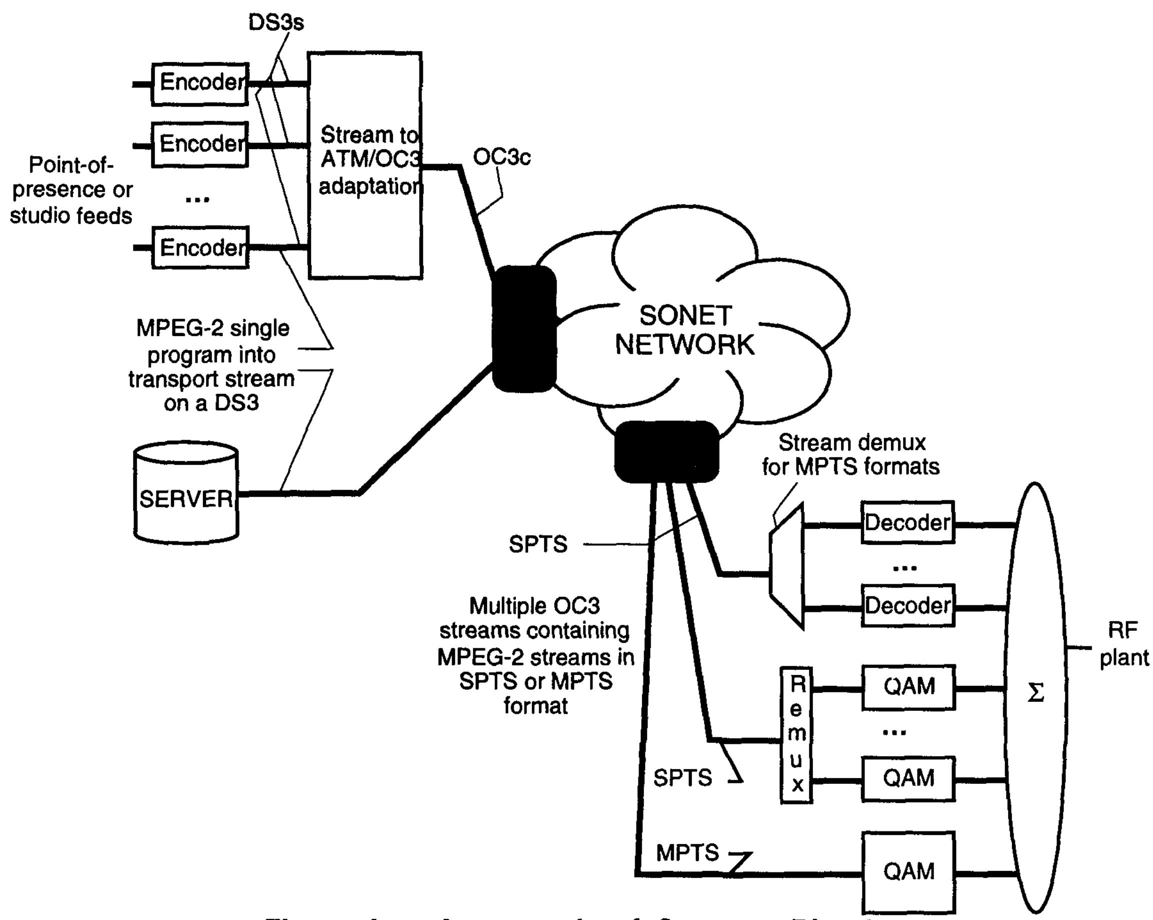

Figure 4 — An example of Constant Bit-rate Streams Adapted to ATM/OC3

Figure 4 shows one of the adaptation scenarios just discussed from non-ATM to ATM transport. Essentially, a number of encoders perform compression on sources coming from venues where major events are taking place, or from studio feeds for instance. The compressed signals are transmitted over non ATM channels such as DS3. The diagram only shows DS3 in seemingly the access portion, however, the access could be followed with a network part that is DS3-based until it gets to the portion that is ATM-based.

At this juncture, the signals can be converted for ATM transport over SONET using AAL5 stream multiplexers. The streams are placed then into OC3c connections and directed to the network element for entry into the SONET network where they get routed to appropriate destinations. At the network exit/destination, the single program transport streams (SPTSs) or multiple program transport streams (MPTSs) are directed to equipment that will either:

- convert the streams back to analog video using decoders, or
- have the SPTS streams remultiplexed into MPTS to be put through QAM modulators, or
- have the MPTS go directly to QAM modulators.

The resulting signals can then be combined as the final output for the RF plant.

■ *6 - Conclusion*

The previous presentation has discussed the issues of MPEG-2 video on non-ATM and ATM networks. Inherently, MPEG-2 has been designed as network independent. It can be transported by ATM as well as non-ATM networks. SONET can be seen as a fabric which can carry both types of networks as well as providing links between the two.

The previous presentation has discussed the issues of MPEG-2 video on non-ATM and ATM networks. Inherently, MPEG-2 has been designed as network independent. The standards have addressed transport of MPEG-2 with ATM cells and the ATM networks can be based on SONET infrastructures. However, MPEG-2 streams can also be transported via non-ATM circuits, like DS3 channels, in a non-ATM format. This has not yet been addressed for standardization. At this moment, SONET is a widely accepted standard in the digital telecommunication field. Networks across the country, across the continent, and between continents can be connected and interwork. Moreover, the networks support voice and data, and will be the backbone of the multimedia highways. SONET is a network fabric for many markets: the public telephone network, the private enterprise networks, the CATV networks, etc. and MPEG-2 can benefit from either mode of transport, ATM or non-ATM.

■ *7 - References*

[1] ISO/IEC, , IS 13818, Coding of Video and Associated Audio in the 3-15 Mb/s Range, 1995

[2] Daniel Minoli, Video Dialtone Technology, McGraw-Hill, 1995