Video Routing and Formats

Introduction

Life in the DTV world would be simpler if there were only a single video format allowed in ATSC Table 3. That isn't the situation in the real world. The multiplicity of contribution and emission formats complicates the design of a practical broadcast facility.

Some networks have already indicated that they will provide programming to their affiliate stations in more than one format. A number of their affiliate stations, however, are planning to implement a single-format DTV facility. As might be expected, many stations will continue operating the legacy NTSC portion of their plant in its current form and apply the single-format philosophy only in a new island of DTV operations. Perhaps surprisingly though, some stations are embracing this single-format philosophy even more aggressively, with plans to up-convert even NTSC signals and derive their NTSC output from a single-format HDTV master control switcher! A single format used throughout various areas of a DTV facility is referred to as the "plant native format", and will be discussed in detail later.

Network and Affiliate Relationships

The networks typically have affiliation agreements with stations that stipulate the relationship of the network and station on programming. No precedent has yet been set to apply these agreements in the context of DTV, so it is possible that an affiliate station might take a standard definition feed and up-convert the program for broadcast. Alternatively the station might receive a program in a high-definition format and elect to transmit the program in a standard-definition format.

Typical Television Station Operations

DTV program capabilities must include all current NTSC program operations at a minimum. The process of passing through a network DTV signal is simpler than what occurs during current and conventional station operations as well as during emergency conditions. However there are a few non-trivial operations that the station must perform even in a pass-through scenario including addressing the mandatory actions within PSIP (ATSC Standard A/65). Current and future station operations and capabilities are focused on master control and production control.

The DTV operations must be accomplished without adding operational personnel for separate DTV master control and production facilities. Daily television station operations encompass many of the activities listed below.

Master Control

- Network Feed
- Network News Service
- Independent News Service
- Local live feeds—studio, newsroom, ENG truck, satellite truck
- News Programming/production studio
- Spots—National/Local
- Station Promos, PSA's, etc.
- Syndicated Programming-- Legacy and Digital

- Station Branding and ID
- Pushback (DVE for weather, school closing, live news events)
- Video crawls
- Audio Under/over
- Weather Alert Systems
- Automation control
- Multiple outgoing streams
- Audio issues mono, LoRo, LtRt, 5.1
- Multiple channel broadcasts
- Encoder switching SD/HD
- Encoder switching data rates
- Traffic interface with reconciliation
- Live external closed caption

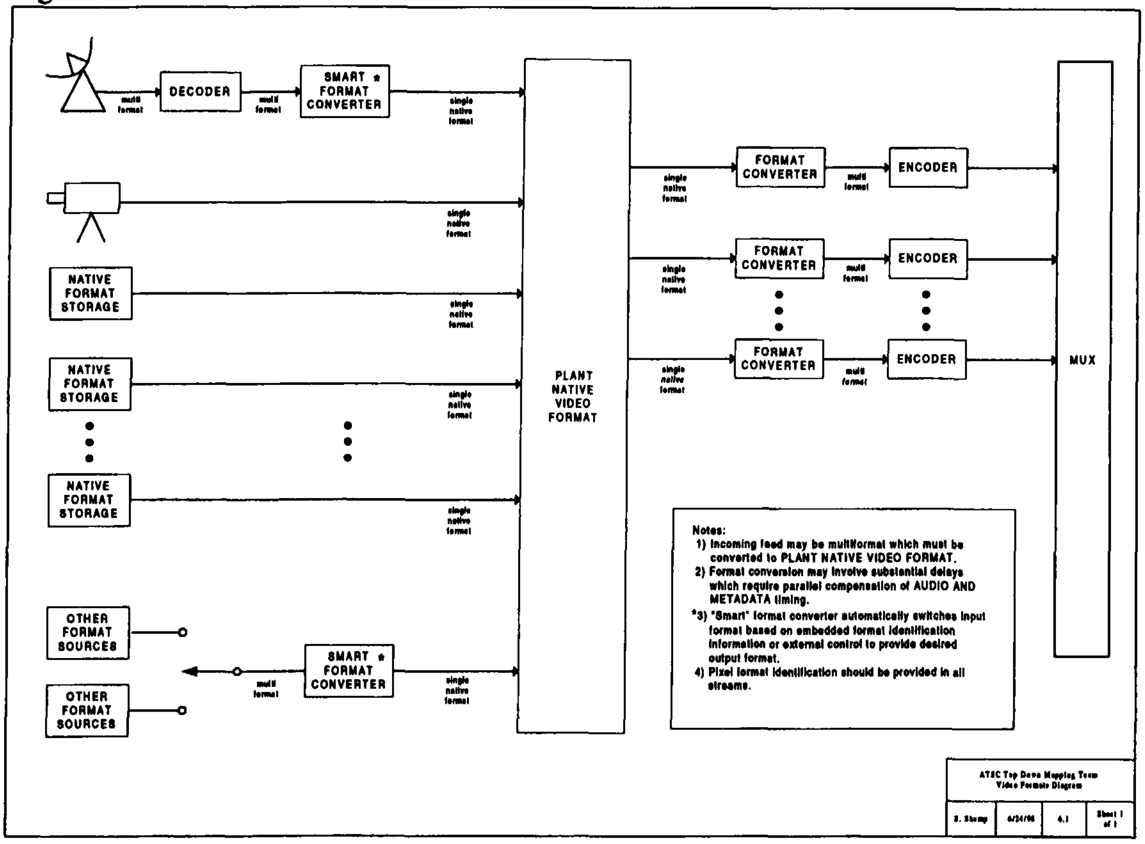
Production Control
- News
- Remote—Parades, elections, disasters, pool feeds
- Field Acquisition
- Legacy formats—U-matic, Betacam, MII, Beta SP, Digital Betacam, DV, DVC PRO, Beta SX, Digital -S, D-1, D-2, D-3, D-5
- Live microwave/satellite
- Computers—control presentations, CG, system playbacks, etc.
- Captioning system/prompter
- Live—Latency problems
- Transitions other then VI (DVE, CG, video mix/effects, chroma-key, virtual sets
- DVE , CGs, SS
- Weather graphics, weather radar
- Cameras
- Audio-- mono, LoRo, LtRt, 5.1
- Live fold-back to the field (mix minus)

Single Format Philosophy: "Plant Native Format"

Several DTV plant architectures were considered. The most basic function, pass through of a pre-compressed stream, with or without local bit-stream splicing, is initially economical but lacks operational flexibility. To provide the operational flexibility required, digital video levels could be added to an existing analog NTSC facility, but managing multiple video formats could get complex and expensive. An interesting alternative is to extend the existing facility utilizing a single-format plant core infrastructure surrounded by appropriate format conversion equipment. This single-format philosophy of facility design is based on application of one format, called the "plant native format", to as much of the facility as practical. The plant native format was chosen as the strawman for video routing and format discussions in the ATSC-IS Top Down meetings. The plant native format approach is shown in Figure 1 on the next page.

(Note that this simplified block diagram differs substantially from the Top Down Map described in the earlier meeting methodology section. The Top Down Map was a generic system block diagram used in video routing discussions and other discussions to identify specific interface standards and issues. The simplified plant native format block diagram was used in discussing implementation specifics.)

Figure 1.



Local station economic realities require minimal capital investment and minimal impact upon existing operations, either practical or financial. The single plant native format philosophy provides the on-air flexibility required and supports the necessary business objectives.

Considerations in Choosing the Plant Native Format

Stations will adopt a native format that makes sense for them. A native format will be based on numerous factors including network's preferred formats, economics, existing legacy equipment and migration strategy, equipment availability, station group format, or marketing reasons.

A native format assumes that all material will be processed in the plant in a single common standard to allow production transitions such as keying, bug insertions, etc. The format converters at the input of the plant will convert the contribution format to the native plant format. A format converter would not be necessary if the native format is the same as the contribution format. Legacy streams will also likely require format conversion. The native plant infrastructure may be traditional analog or digital routers, production switchers, master control switchers, or new concepts such as servers and digital networks.

Choosing a native plant format has a different set of criteria than choosing a broadcast emission format. Input signals may arrive in a variety of formats necessitating input signal format conversion. Similarly, changes in broadcast emission format will necessitate output format conversion. Therefore the native

6

plant format should be chosen to help facilitate low-cost, low-latency, high-quality format conversions. It is well understood that format conversions involving interlaced scan formats require careful attention to de-interlacing, while progressive scan formats offer fewer conversion challenges.

The native plant format could be chosen from a variety of candidates, ranging from standard definition 480-line interlaced format to 720 or 1080-line HDTV formats. The specific choice depends on a variety of factors, particularly differing station business objectives.

Choices:

1. Analog NTSC 170M
2. The format of 480i 4:3 or 16:9 (125M) carried by SMPTE 259 Rec. 601
3. 1080i or 720P using SMPTE 292M
4. 480p using SMPTE 293 or 4:3
5. Intermediate compressed formats

Although the full benefits of adopting a plant native format are most apparent with an implementation in which "plant native format" means a single plant native format, many stations will not be able to afford to adopt a single plant native format immediately. For them "plant native format" will, at least on an interim basis, mean a practical mix of perhaps two formats. One of the formats might be the legacy format which is being phased out while the other is the preferred format for future operations. So while the ideal implementation of the plant native format philosophy would truly be a single format, the following implementation scenarios recognize practical constraints of existing facilities.

Implementation Scenarios for Plant Native Format

Individual broadcasters might each start at different points and follow different steps in their implementation of a DTV plant native format. They can migrate toward higher quality over time as new equipment becomes available, depending on their individual economic situation. There could even be islands of different formats or technologies at the same time. Signal distribution could use traditional routing switchers or networks such as FibreChannel or ATM. Storage could use a mix of tape and tapeless technology.

Plant Native Format Scenario 1: NTSC. For some early adopters, the plant will be analog NTSC, with only limited DTV operations in master control.

Plant Native Format Scenario 2: Hybrid. In the first phase, a standard-definition multi-channel component digital master control with format converters would drive both NTSC and DTV outputs. Production and news studio operations would continue in analog NTSC, with outputs converted to the component digital plant native format. News acquisition and editing would continue to use analog NTSC. In the second phase, production control would change to component digital plant native format operation, with news and acquisition still analog NTSC. Finally in the third phase, news acquisition and editing would be upgraded to component digital equipment.

Plant Native Format Scenario 3: 1.5 Gbps. A 1.5 Gbps plant native format based on SMPTE 292M could support both 720P or 1080I as well as all lower resolution formats.[1] The phased

[1] Standardization of HD SDTI is currently under development by standards bodies, and could represent an important plant distribution mechanism in the future.

implementation as described for scenario 2 above could be applied. An individual broadcaster could start with scenario 1 or 2 before migrating to scenario 3.

Plant Native Format Scenario 4: Compressed Formats. A plant native format could be based on compressed formats such as MPEG or DV. Any points in the plant which require full-bandwidth signals, such as master control or production (news) control, would have appropriate converters. Storage devices and, in some cases remote feeds, could operate directly with the compressed plant native format.

Smart Format Converter

There are some applications in which the format converters operate with a single input format and a single output format. In other applications, such as signals from networks or outside sources, the format may change from time to time. In these applications, it would be desirable for the format converter to be able to adapt to a variety of input signal formats.

Commonly available format converters (both up-converters and down-converters) require selection of input and output formats. In the "Native Plant Format" implementation, operation could be simplified if enhanced "smart" format converters could identify the input signal format and automatically convert it to the desired output format.

The smart format converter could be facilitated by inclusion of format identification in the serial data stream, as described earlier. This is an issue for standardization.

Choosing an Emission Format

The affiliate TV station may broadcast their network's HD format, but when local needs develop, they may change their network feeds to SD and add a second SD for such events as local election returns, emergency weather, and live news. This requires emission format converters for each of the station's program feeds.

Therefore affiliate TV stations must operate one functional legacy NTSC and/or DTV plant, using one basic native digital bit-stream format instead of switching multiple formats. That in turn requires inputs to have format converters from legacy and other HD formats to the station's native format and from the native format to the emission format(s).

A format converter(s) on the output will be necessary if you plan to emit a different format other than the native format. This architecture supports single and multiple channel output streams.

A station might choose an HDTV native format and convert all differently formatted material to this format and pass it to an encoder without conversion. However, there may be a time when the station desires to emit a lower resolution format in order to accomplish a special purpose. The output format converter accommodates this requirement.

A station might choose an SDTV native format and convert all material to this format and pass it to the encoder with or without conversion. Even the case where the native format is required by economics to be composite NTSC, the output will pass to a format converter and on to the ATSC encoder.

SDTV Production vs. Emission Format Discrepancy

Some DTV production formats exactly match the image format definitions of the corresponding DTV emission formats. For example in the case of 720P or 1080I HDTV formats, the structures of pixels per line and lines per frame are exactly the same in production and emission formats. A troublesome counter example is the well established ITU-R BT.601 and SMPTE 125M standard definition production format, for which there is no corresponding perfectly matched emission format. In particular, digital production facilities around the world are based on 720 pixels per line, while the ATSC Table 3 emission format has only 704 pixels per line.

Various processes in the production process introduce artifacts at the edges of the image. This was recognized some time ago with the introduction of the concepts of "Clean Aperture" and "Production Aperture" in SMPTE RP187. Transient filtering effects have been dealt with using a number of techniques.

Beyond these requirements for filtering in the production process, it is recognized that both transmission and display processes include functions that may introduce transient filtering effects. The introduction of edge effects can occur not only in production, but in transmission and display as well. It is therefore advantageous to pass the entire production aperture as defined in SMPTE RP187, through the digital production/emission/display chain in order that any filtering and edge effects be dealt with in a consistent manner throughout the production, transmission, and display processes. The 1920x1080 and 1280x720 formats already enjoy such a consistent approach, while the 720x480 format has different production and emission parameters that preclude a consistent approach to filtering and edge effects. This suggests that the 704x480 entry in ATSC Table 3 may not be adequate.

If a 720 pixel/line production format must be used in conjunction with a 704 pixel/line emission format, a number of undesirable alternatives are available. The simplest way to get from 720 pixels/line is truncation. Truncation may, unfortunately, result in undesirable ringing artifacts due to excessively fast risetimes. Another alternative is use of format converters at numerous points throughout the plant to convert between 720 pixels and 704 pixels and back. These format converters introduce additional expense and an inconsistent approach between different DTV signal formats.

The preferred solution to this discrepancy between 720 pixel/line production format and 704 pixel/line emission format is to match the production and emission format, thus applying an approach which is consistent with the approach used in the case of the 1920x1080 and 1280x720 formats. Failing in this rationalization of formats, format conversion between production and emission formats is recommended.

Further work is required to identify whether, in fact, consumer receivers might be fully capable of decoding 720 pixels/line. In this case, broadcasters would have freedom to transmit 720 pixels/line corresponding to Rec. 601.

Interconnect Signal Formats for Different Image Formats

As discussed earlier, different individual broadcasters may choose different migration paths to their plant native format. Recent proposals for carrying a wide variety of different image formats on either SDTV or HDTV serial digital interfaces bring an interesting degree of freedom to the facility design. A number of considerations related to the plant native format make this flexibility attractive. Facilities will have to contend with multiple image formats, possibly at multiple frame rates, but related to a single clock. Facilities will not want to deal with different layers of routing for each format. Material will be converted to a common native format to permit intercutting, mixing, etc. in hard real time.

A broadcaster who choose to emphasize standard definition operations in the early stages of DTV implementation could design a SMPTE 259M (270 Mbps/s or 360 Mbps/s) serial digital (SDI) facility. As more HDTV programming was provided, the facility could use SMPTE 305M transport to carry a mix of HDTV and SDTV without changing the basic serial digital routing core. This possibility is explored in more detail in the "Compressed Serial Formats" section later.

Alternately, a broadcaster who was anxious to implement full HDTV operations much sooner could install a SMPTE 292M (1.5 Gbps) serial digital core. Assuming development of appropriate mapping and HDTV SDTI standards, that 1.5 Gbps core could some day be used for a mix of HDTV and SDTV, much like the SMPTE 305M example. This possibility is explored in more detail in the "Uncompressed Serial Formats" section below.

Compressed Serial Formats

If an individual broadcaster has either already installed a 601 component digital routing core or expects to emphasize SDTV in the initial phases of DTV implementation, the option of handling a mix of SDTV and HDTV signals on that same basic infrastructure is an important degree of freedom. Particularly in markets where broadcasters expect a relatively slow consumer acceptance of HDTV, a logical migration from SDTV to HDTV is critical.

Proper application of compression in a DTV facility can provide this migration path. If we use the familiar 270 or 360 Mbps/s serial digital routing and distribution within a facility, we can support either lightly compressed HDTV or non-compressed SDTV. A facility could convert to digital now, implement 16:9 wide-screen operation at the appropriate resolution for the individual market, and know that the basic infrastructure is future-proofed to handle either SDTV or HDTV. This lightly compressed video has been shown to be virtually transparent in typical broadcast operations and can offer system design options which are synergistic with compressed disk or tape recording systems.

Uncompressed Serial Formats

As multi-format HDTV equipment for both 720P and 1080I became available, it also became clear that this equipment could also have included 480P or even 480I capability were it not for differences in serial interface specifications. One solution which could provide a common interface for all formats of video, from SDTV through HDTV, would be to map all of the image formats onto a single serial interface at 1.485 Gbps (or 1.485/1.001 Gbps). This interface would be a superset of SMPTE 292M, using the same scrambled NRZI channel coding.

One possible approach to develop this serial digital interface would treat video as packetized data, with each active line delimited by SAV at its start and EAV at its end. Spacing of the EAV and SAV timing reference signals would be determined entirely by the characteristics of the video being conveyed. The remaining space would be available for other purposes, but would be stuffed to bring the interface to a standard rate. The format of the video signal would be explicitly declared by metadata embedded in the signal.

Latency Considerations

Whether discussing a lightly compressed interconnect to put HDTV onto an SDTV interconnect, or mapping SDTV onto an HDTV interconnect, format conversions into and out of the plant native format will incur a delay penalty. In some cases this may be on the order of lines of delay; in other cases it could be a field or a frame. Even where no video format changes are involved, the possibility of compressed audio transport to provide 5.1 channel capability on 2 or 4 channel devices introduces latency issues. All these possibilities point to the need for careful system design and for development of standardized approaches to guarantee correct audio/video timing relationships.

Different Approach for Audio

The "native format" philosophy developed for pictures does not apply to audio because the "up-conversion" process is not applicable to audio. Once multiple channels are downmixed to fewer channels, there is no process for deriving the original individual channels.

First, we have to define the difference between audio "format" and the number of channels a plant can handle. The number of channels in a plant is just that. The audio format can differ from the number of channels available because a four channel format (such as Dolby Surround) can be carried in a two channel plant, when properly encoded. In the same sense, a plant with a single layer of AES3 digital audio can handle a 5.1 channel format audio signal if that signal has been rate reduced by a contribution quality encoder.

The transmitted audio format is determined by two factors. If a high level format (multi-channel) signal is received, and the plant infrastructure can carry that signal format through the entire process, then it can be transmitted as such, if the plant infrastructure can only transport a lower level format, then the incoming signal must be mixed down to the supported signal format, and transmitted as such. If the incoming signal is at a lower level than the plant can support, the signal must be transmitted at the received level.

The audio layer map development group assumed that there was a digital audio infrastructure. This allows the plant to carry the multi-channel signal (using contribution quality rate reduction if the AES3 signal is used).

Use of compressed audio in the plant introduces latency concerns corresponding to those described under compressed video earlier.

Control, Monitoring & Data

Control of video systems, as well as control of data emissions are additional considerations to be referenced in the documentation of the control group. Likewise, there are significant monitoring and metadata issues that need to be considered in a new DTV plant that are referenced in the Monitoring and Data group output.

Recommendations for Video Formats

1. We recommend standards specifying video (pixel) format metadata in the uncompressed domain, for such purposes as the identification of varying input formats to facilitate automatic switching of format converters.
2. We recommend a common serial (SDI) transport standard that encompasses all production formats being considered for the DTV environment, including not only ATSC formats but also other emerging formats. We recommend efforts to facilitate the use of SMPTE 292M for this purpose.
3. We recommend more study of latency issues, to develop solutions to guarantee synchronization between video, audio, closed-caption information, and program-related data/metadata and to facilitate news-crew cueing and other remote-location interaction.

    - We recommend a new set of standards for metadata for time stamping of video, audio, captioning, and data/metadata.
    - We recommend that every format conversion stage have an associated standard delay of audio, captioning, and data/metadata and/or time stamping. Time stamping will require compensating audio, video, captioning, and data/metadata delays at monitoring, recording, and encoding points.
    - We recommend that every point of separation between audio, video, captioning, and data/metadata streams have time stamping and/or associated delays to ensure time coincidence (after compensating delays) at monitoring, recording, and encoding points.
    - We recommend that the ATSC inform broadcasters that latency issues may render the current practice of radio-network-based FM simulcasting impossible. Multichannel DTV audio may not be a viable substitute for such simulcasts.

4. We recommend that, if the SDTV video format closest to ITU-R Recommendation 601 must remain 704 x 480, there be format conversion between 720-active-pixel production equipment and 704-active-pixel emission encoders or decoders. The group calls the ATSC's attention to cost and latency issues associated with such format conversion.
5. The generic top-down map developed in first meeting served well in identifying interface requirements, but is not suitable for explaining preferred implementations. Therefore the top-down map should be supplemented with a combination of preferred implementations developed in the various top-down sub-groups.

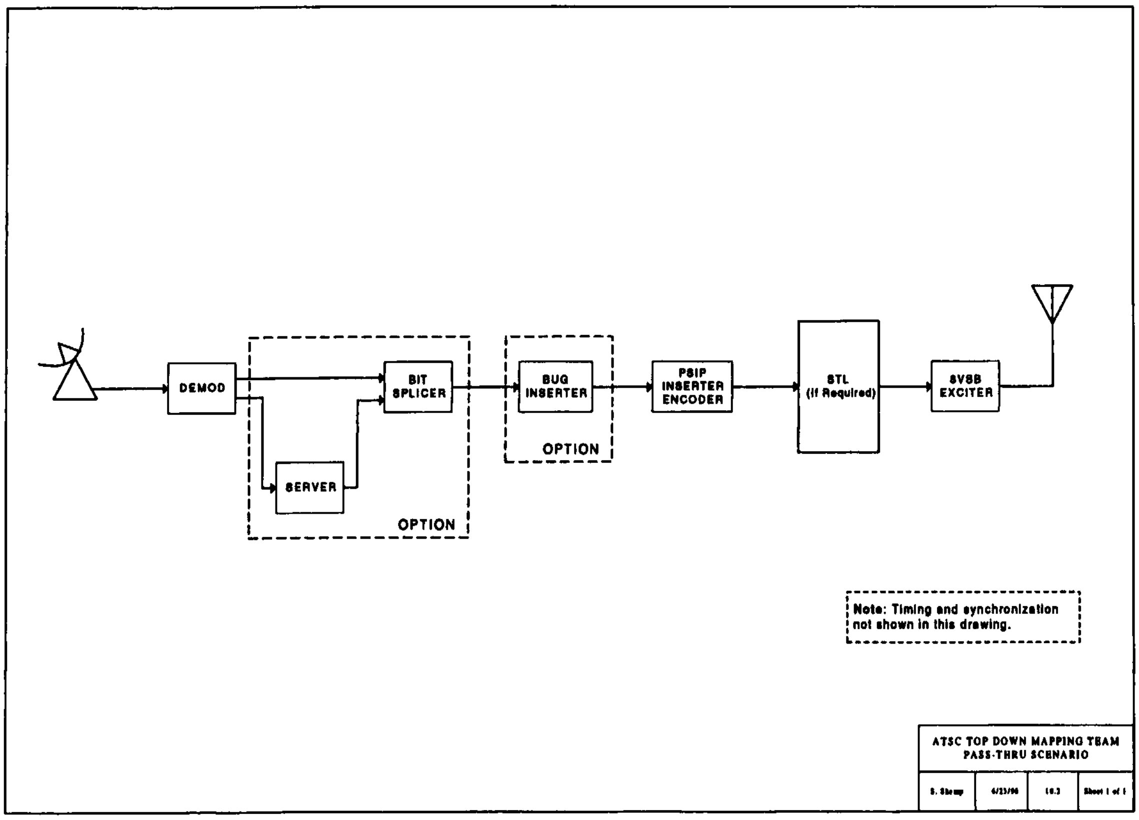
Pass Through of ATSC 19.4 Mbps Transport Stream

Most commercial networks will be providing DTV feeds to their affiliate stations which have adequate headroom to decode signal and operate with conventional non-compressed master control switching. This will allow DTV master control to operate with all the capabilities expected in current NTSC master control environments. Due to economic considerations, some stations may initially opt to pass their network DTV signal, with minimal local manipulation, directly to their transmitters.

Stations do, however, wish to brand their signals. This may involve keying a logo in the corner of the video. Many operations like keying a logo generally involve decoding the compressed network DTV signal to baseband, doing the desired image processing, and re-encoding. Some station operations can, however, be done entirely in the compressed domain. This approach offers an alternative for some stations, since it mitigates the need for a DTV encoder at each local station. A PSIP manipulator is required.

A station could store pre-encoded program or interstitial material on a transport stream server. Bit stream splicers can provide seamless switching between the network and local material, but this functionality is dependent on characteristics of the compressed bit stream.

Figure 2. Pass Through Scenario



Encoding and Multiplexing

Introduction

At the heart of the transition to digital broadcasting is the utilization of compression systems to optimize the program quality through limited bandwidth networks. Additionally, the transition to digital allows multiple streams (audio, video, data and control) to be multiplexed to share a common communication link. In addition to the encoding and multiplexing that must occur to provide an ATSC compliant transmission, an examination of the various paths that program information flows within a station clearly indicates that compression and multiplexing will have extensive use within the plant.

Video compression schemes have been under development for a significant period, and standards activities have resulted in the MPEG-2 standard being well documented and adopted by various manufacturers. The flexibility of this compression standard allows each manufacturer to exercise a number of options to provide the best tradeoff of cost and performance that they feel appropriate for their target applications. While this flexibility is key to ongoing improvements in video quality and programming choice, it can also lead to a multiplicity of design choices that potentially wreak havoc on interface standards. This confusion is multiplied by the application of compression encoders in various

13