MPEG-4 to MPEG-2 HDTV Network Video Transcoding

John Hartung, EGT CTO

Introduction

Much to the surprise of many cable operators, the industry's MPEG-4 era appears to have abruptly begun. Due to HBO's startling June announcement that it will begin beaming all 26 of its HBO and Cinemax multiplex channels in HDTV using the far more bandwidth-efficient video compression format, MSOs suddenly face the daunting challenge of figuring out how to carry all that MPEG-4 programming from cable's biggest network brand. With more than 34 million MPEG-2-only digital set-top boxes already deployed throughout the U.S., cable providers also face the unappetizing prospect that other major cable and broadcast networks, eager to save bandwidth and cut transport costs as well, will likely follow HBO's lead sooner rather than later.

Besides coping with these coming shifts in content distribution, cable operators face the challenge of competing with rival satellite TV and telco TV providers that are already carrying out the upgrade to the advanced MPEG-4 codec. As part of their ambitious HD expansion drives, which call for the carriage of hundreds of national HD channels, both DirecTV and EchoStar are already signing up new high-def subscribers and switching over existing customers to their new MPEG-4 set-top boxes. At the same time, AT&T has started delivering its new U-verse HD programming lineup using the MPEG-4 codec.

As a result, cable operators need a technological solution that will allow them to convert fresh MPEG-4 content to the MPEG-2 format quickly, cheaply, and seamlessly. They need a different system architecture and special tools that will enable them to transcode digital video signals from the new format to the older one without hurting the quality of the pictures. And they need a high-density approach that will permit them to convert many video signals from MPEG-4 to MPEG-2 at the same time.

This EGT white paper addresses the promising market opportunity that MPEG-4 to MPEG-2 transcoding offers to cable operators as HDTV increasingly becomes a mainstream medium in North America. It tackles the myriad technical challenges of carrying out that conversion swiftly and effectively. Finally, the paper examines the techniques and system architectures needed to conduct MPEG-4 to MPEG-2 transcoding at high quality, low cost, and high density.

The MPEG-4 Transcoding Opportunity

Although a switch to the MPEG-4 format offers plenty of technical challenges for the cable industry, it offers a host of enticing market opportunities as well. First off, an upgrade to MPEG-4 would enable MSOs to cut their burgeoning HDTV bandwidth requirements by at least 50%, allowing them to reclaim loads of spectrum for even more HD channels or other advanced digital services. For instance, HBO plans to transmit its MPEG-4 high-def channels over the satellite at a rate of just 8 Megabits per second (Mbps), compared with the standard 19.2 Mbps rate for

linear high-def channels in the MPEG-2 format and about 15 Mbps for cable's HD video-on-demand (VOD) programming.

Second, an upgrade to MPEG-4 distribution would pave the path for the distribution of HDTV content over cost-effective Ethernet networks, rather than more costly satellite frequencies. By enabling both cable operators and network programmers to squeeze much more HD traffic into their existing bandwidth, the codec promotes the use of IP standards for the transmission of video signals.

Third, a cable embrace of MPEG-4 transcoding would insure that the industry maintains its competitive parity with satellite TV and telco TV providers. With DirecTV, EchoStar, and AT&T all moving in the MPEG-4 direction already, cable operators must do so as well to match, and ideally exceed, the HD and other advanced digital service offerings of their prime video rivals. Otherwise, cable operators may find themselves in the same position as several years ago, when they had to play catch-up with DirecTV and EchoStar on the digital programming, video-on-demand (VOD), and digital video recorder (DVR) fronts.

Lastly, the proposed upgrade to MPEG-4 offers an especially intriguing opportunity for mid-sized and smaller cable operators that haven't really introduced HDTV programming packages yet. Without a large base of MPEG-2-only HD set-tops already out there in the field, smaller MSOs can more easily and quickly deploy the new generation of hybrid MPEG-2/MPEG-4 set-tops that can handle both formats. In turn, this will allow them to launch new HD programming services rapidly and efficiently.

The MPEG-4 Transcoding Challenge

The process of transcoding MPEG-4 programming to MPEG-2 presents cable operators with a number of technical challenges. Most importantly, there are some key differences between MPEG-4 and MPEG-2 encoders that must be addressed and overcome.

To start with, both the MPEG-4 and MPEG-2 encoding algorithms use block-based, motion-compensated prediction, quantized transform coding of residuals, and entropy coding. However, MPEG-4 introduces additional coding modes for motion compensation and spatial prediction within frames, along with a new transform and entropy coder. These new coding modes do not directly correspond with the MPEG-2 encoding tools. Thus, they represent the greatest challenge in transcoding video signals from MPEG-4 to MPEG-2.

In the MPEG-2 format, motion-compensated prediction is performed using motion vectors for 16 x 16 pixel macroblocks or, in the case of interlaced frames, 16 x 8 pixel macroblocks. The reference frames used for prediction come from a single preceding and/or following decoded frame. The MPEG-4 motion estimation tools have been expanded to include additional block shapes and sizes, and multiple reference frames can be used to predict a macroblock.

Both MPEG-2 and MPEG-4 provide macroblock coding modes for interlaced fields. But the organization of the fields into macroblocks is different for the two standards. In MPEG-2, field

coding is performed by partitioning a 16 x 16 macroblock into two 16 x 8 fields. In MPEG-4, on the other hand, the two fields are represented by vertically displaced 16 x 16 macroblocks (MBAFF). The MPEG-4 format also introduces a new intra-frame spatial prediction mode that has no correspondence to the MPEG-2 coding modes.

In addition to these new coding modes, MPEG-4 also introduces changes in the encoding loop. The new format replaces the DCT transform used in MPEG-2 with a smaller integer transform, and augments the older VLC entropy coding with an optional adaptive binary arithmetic entropy coder (CABAC). The MPEG-4 standard also calls for a filter in the encoding loop that helps mitigate encoding artifacts, such as blocking at low encoding rates. Although this does not improve the Peak Signal to Noise Ratio (PSNR) performance, it results in more acceptable subjective artifacts.

The MPEG-4 Transcoding Solution

One approach to transcoding video from MPEG-4 to MPEG-2 is to fully decode the MPEG-4 frames and then re-encode them with an MPEG-2 encoder. This decoding and re-encoding process can be carried out with an entirely separate decoder and encoder.

Unfortunately, this approach is computationally expensive because it does not reuse the MPEG-4 encoding parameters and does not produce the highest quality MPEG-2 encoding. The method compromises quality by using quantized frames to re-calculate coding modes and motion vectors, unlike the original MPEG-4 encoded stream, which could provide more accurate decisions based on the initial sequence. It also fails to match frame coding types so that high-quality reference frames, such as I and P frames, are re-encoded with the same types in MPEG-2 for use in subsequent prediction. Failure to do this leads to lower quality reference frames and a resultant propagation of their coding distortion.

An alternative approach offers far more promise. This method calls for decoding the MPEG-4 input to video frames while simultaneously passing the MPEG-4 encoding parameters to the MPEG-2 encoding stage. The encoding step then follows five stages: mapping of MPEG-4 encoding modes to MPEG-2 modes where possible; generation of candidate MPEG-2 motion vectors from MPEG-4 vectors; refinement of motion vectors; bit allocation; and MPEG-2 reference frame generation. Figure 1 shows a block diagram of this approach.

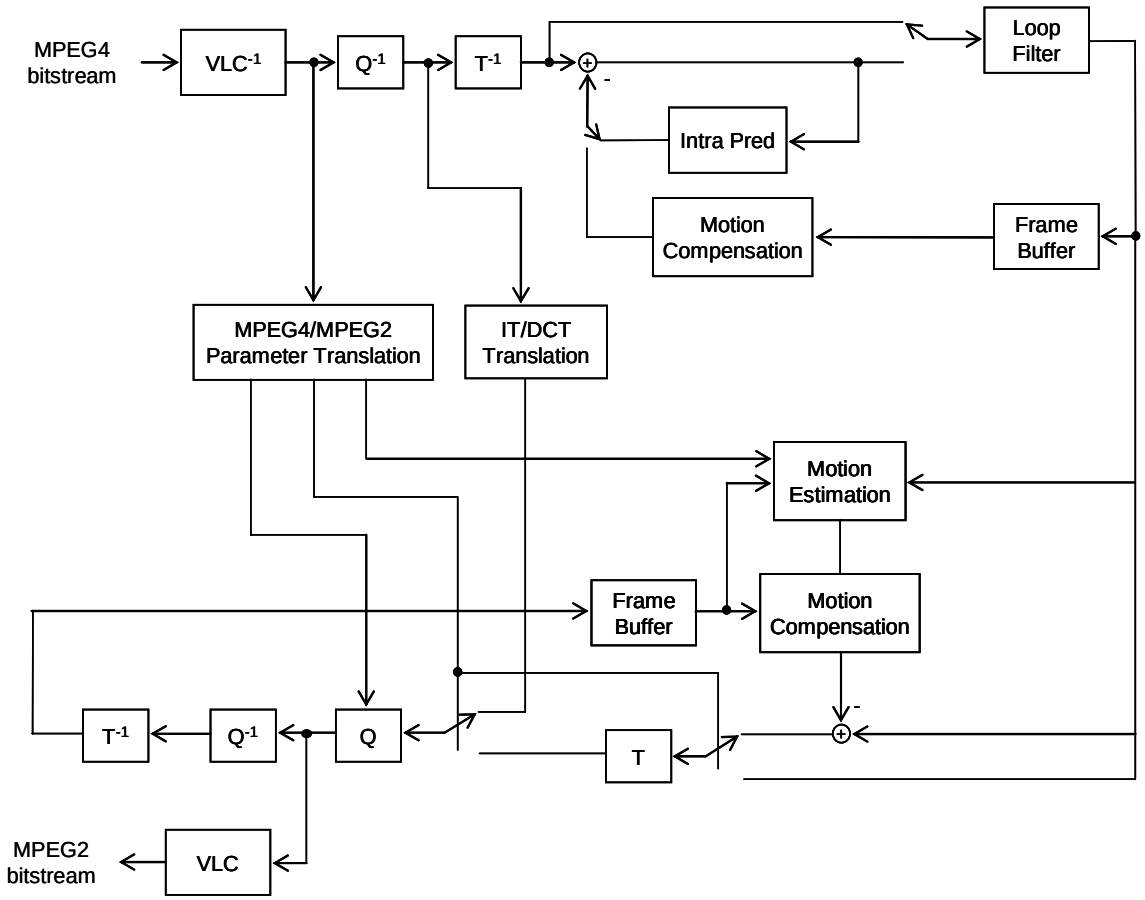

Figure 1
Optimal MPEG-4 to MPEG-2 Transcoding Architecture

In some instances, the encoding parameters in MPEG-4 can be directly mapped to MPEG-2 parameters, significantly reducing the complexity. This mapping consists of converting the MPEG-4 integer transform coefficients to MPEG-2 DCT coefficients, and MPEG-4 motion vectors to MPEG-2 motion vectors. Efficient techniques for transforming coefficient conversion are described in the literature [1].

When 16 x 16 macroblocks are involved, the MPEG-4 motion vectors can sometimes be directly mapped to MPEG-2 motion vectors [2]. However, parameter mapping is more complex in MPEG-4 prediction modes with interlace, prediction using partitioned macroblocks, and multiple reference frames. Field prediction modes in MPEG-4 can be translated to MPEG-2 parameters in the following ways. For MPEG-4 frame pictures not using MBAFF, frame mode coding can be used in MPEG-2. In cases where MPEG-4 uses MBAFF, the corresponding macroblocks in MPEG-2 can use field prediction. Field pictures in MPEG-4 can be handled by using frame mode in MPEG-2 with field mode prediction for all macroblocks.

Once the field/frame mode decision has been made, the remaining macroblocks can be encoded. These macroblocks are more complex to map because they use prediction modes not available in

MPEG-2. The lineup includes blocks having reference frames displaced further than the immediately preceding decoded images, as required in MPEG-2, or using sub 16 x 16 or 16 x 8 prediction block shapes. In these cases, a single mode decision and feasible candidate motion vectors must be determined for each 16 x 16 MPEG-2 macroblock.

We have designed good solutions for this mode-mapping by formulating the problem as a statistical classification issue. This classification returns both the MPEG-2 parameters and an indication of the decision confidence. For those blocks without a clear mapping, a full mode and motion vector search is performed. Candidate motion vectors can be generated by assuming linear motion and then scaling the vectors as a function of the distance to the MPEG-4 reference frame.

After determining the macroblock coding modes and candidate motion vectors, the transcode process can generate the prediction error and use it to identify blocks requiring motion vector refinement. Because candidate motion vectors have been determined from the original MPEG-4 frames, only a small search is required to generate the optimal motion vectors. At this point, the DCT transform of all macroblocks, inter-frame- or intra-frame-coded, can be calculated and quantized to meet bit-rate requirements. Using the MPEG-4 bit allocation to estimate frame and macroblock complexity enables the performance of an effectively “two-pass” MPEG-2 encoding. The succeeding MPEG-2 reference frame is then generated in preparation for the next MPEG-4 frame slated for input.

The described technique dramatically reduces the complexity of transcoding because it avoids the search required for determining optimal MPEG-2 encoding modes and motion vectors, as well as the need for multiple-pass encoding. At the same time, this approach leads to more efficient encoding mode decisions because they rely on MPEG-4 information derived from the original un-coded frames.

Besides the basic transcoding function, there are system aspects that must be considered in deploying MPEG-4 to MPEG-2 transcoders. These aspects include encryption, conditional access, and grooming of the resultant MPEG-2 streams. Because the MPEG-4 signals are carried in an MPEG-2 transport stream, cable operators can use their existing satellite receivers (IRDs) to deliver a transport stream to the transcoding equipment, thereby making better use of these receivers.

For broadcast channels, cable providers can gain additional compression efficiency by incorporating closed loop statistical multiplexing in the transcoding process, instead of using existing open loop statistical re-multiplexers. Feeding the output of the existing IRDs to transcoders that have integrated statistical multiplexing enables the delivery of three HD MPEG-2 channels per QAM at low cost and high quality.

For operators deploying these new HD channels as switched services, operators can better control the per-service bit rate and quality; while at the same time enable MPEG-4 efficiency over the distribution network, transcoding to MPEG-2 closer to the edge.

Proposed System Architectures for MPEG-4 to MPEG-2 Transcoding

Receiver-Based Transcoding

To date, two basic types of system architectures have been proposed for transcoding video signals from the MPEG-4 format to MPEG-2. The first type of system architecture, known as receiver-based transcoding, relies upon the satellite receivers to carry out the actual transcoding from MPEG-4 and then relay the freshly minted MPEG-2 signals to the Switched Digital Video infrastructure or to the variable bit-rate (VBR) broadcast re-multiplexer. See Figures 2 and 3 below for diagrams of these approaches.

Receiver-based transcoding may have certain strengths. But it also has several major drawbacks. For one thing, under this approach, video signal quality in a non-switched application is hampered by the two generations of recoding that must be performed, first at the satellite receiver level and then at the VBR re-multiplexer level. Plus, the open-loop nature of the VBR re-multiplexer keeps it from benefiting from the original MPEG-4 coding decisions, which were discarded at the output of the receiver-transcoder. As a result, the open-loop design can further compromise signal quality.

Another issue is low density. With just one channel of receiving/transcoding per rack-unit, a receiver-based transcoding system requires one full rack of equipment for each set of 24 channels. So it takes numerous racks to support dozens, never mind hundreds, of channels.

Lack of simplicity is a third concern. Under this approach, a cable operator would need to install extensive cabling, creating many more points of failure and increasing the overall potential for system outages.

Further, a receiver-based transcoding system is relatively lacking in flexibility. Once the cable network or individual set-top box becomes MPEG-4 capable down the line, the satellite receivers lose their transcoding value. So there's no longer any payback for the hefty upfront investment in transcoding receivers.

Lastly, and perhaps most importantly, by transcoding MPEG-4 HD to MPEG-2 HD at the reception site, cable operator are simply unable to avail themselves of the bandwidth efficiencies enjoyed by the programmer. Significant, growing numbers of cable operators are distributing their content to multiple headends and/or hubs over an IP backbone. By combining the content reception with transcode operation, MSOs have no choice but to transport the HD content at an MPEG-2 rate that is roughly twice that of its MPEG-4 counterpart.

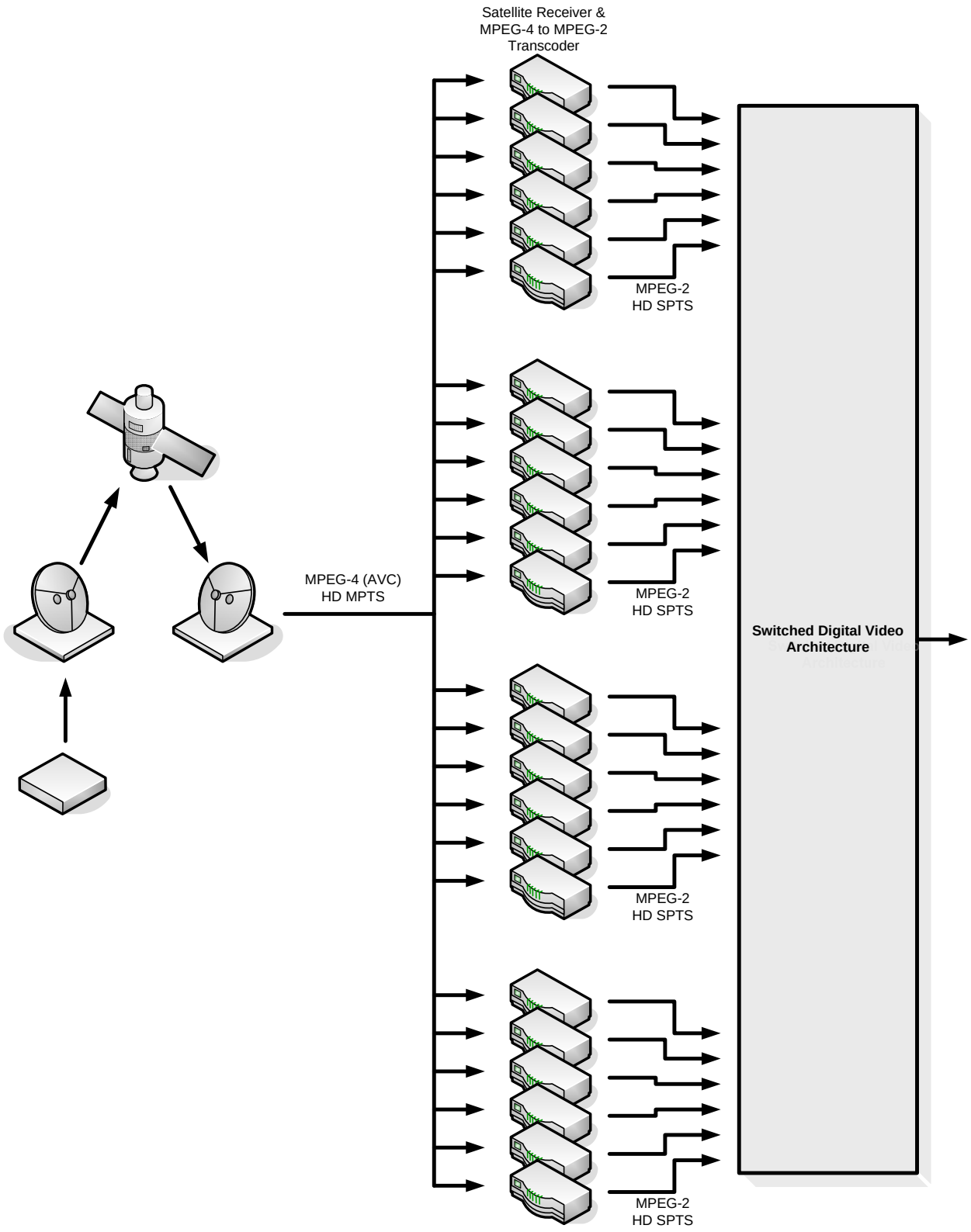

Figure 2
Receiver-based Transcoding Architecture in a Switched CBR Application

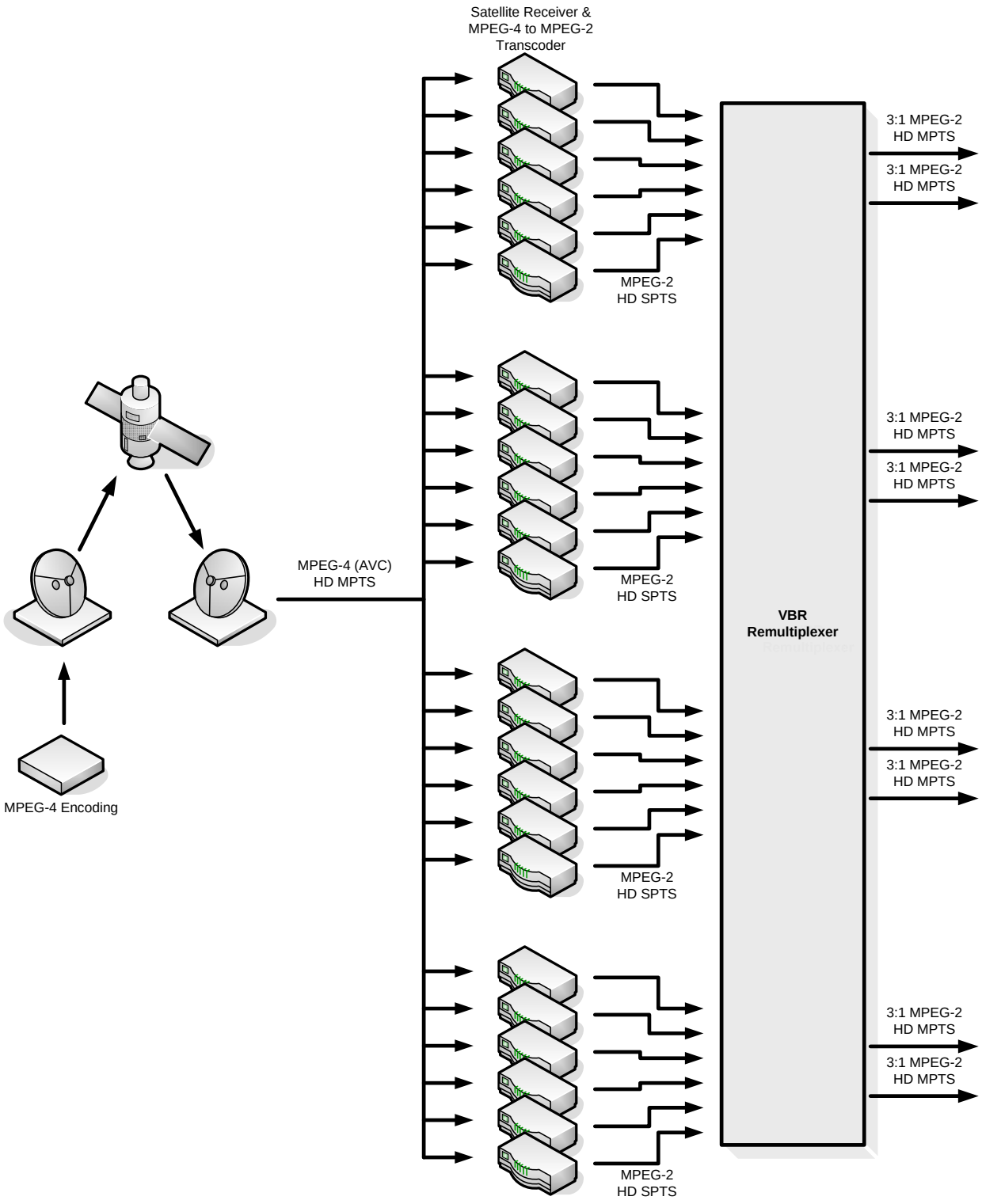

Figure 3
Receiver-based Transcoding Architecture in a Non-Switched Broadcast 3:1 Application

EGT's Suggested Approach: Dense GigE Transcoding

We believe that a second type of system architecture offers far more promise. This system, known as Dense GigE Transcoding, depends upon the MPEG-4 to MPEG-2 conversion occurring at a later stage in the signal transmission process. Instead of the various satellite receivers performing the transcoding, the conversion takes place later at the network level. See Figure 3 for a diagram spelling out this approach.

The relative benefits of the Dense GigE Transcoding method are clear. First, greater signal quality can be achieved because this system uses just one generation of recoding, not two. The Dense GigE system also takes advantage of closed-loop efficiencies, rather than the less efficient open-loop approach of receiver-based transcoding.

Second, Dense GigE Transcoding provides much higher density than receiver-based transcoding. In fact, this system requires just one-half of the rack space required by receiver-based transcoding. So one rack of equipment can support twice as many digital video channels.

In addition, the Dense GigE Transcoding architecture offers significantly greater simplicity than receiver-based transcoding. Under our recommended approach, cable operators will have fewer devices on their networks to manage, monitor, and maintain.

Further, Dense GigE Transcoding promises far greater flexibility than the receiver-based transcoding method. Rather than invest in special transcoding receivers, cable operators can reuse their standard satellite receivers in the Dense GigE architecture. Plus, MSOs can repurpose the MPEG-4 to MPEG-2 transcoders for MPEG-2 to MPEG-4 conversions as needed. And they can easily use this system to transcode video signals for other distribution platforms, such as mobile video.

Lastly, and most importantly, by separating the reception from the transcode, cable operators can choose where to turn their MPEG-4 transmissions back into MPEG-2 signals. They are free to deploy the transcoding at the edge of the network and enjoy the efficiencies of using the MPEG-4 format on their transport backbone.

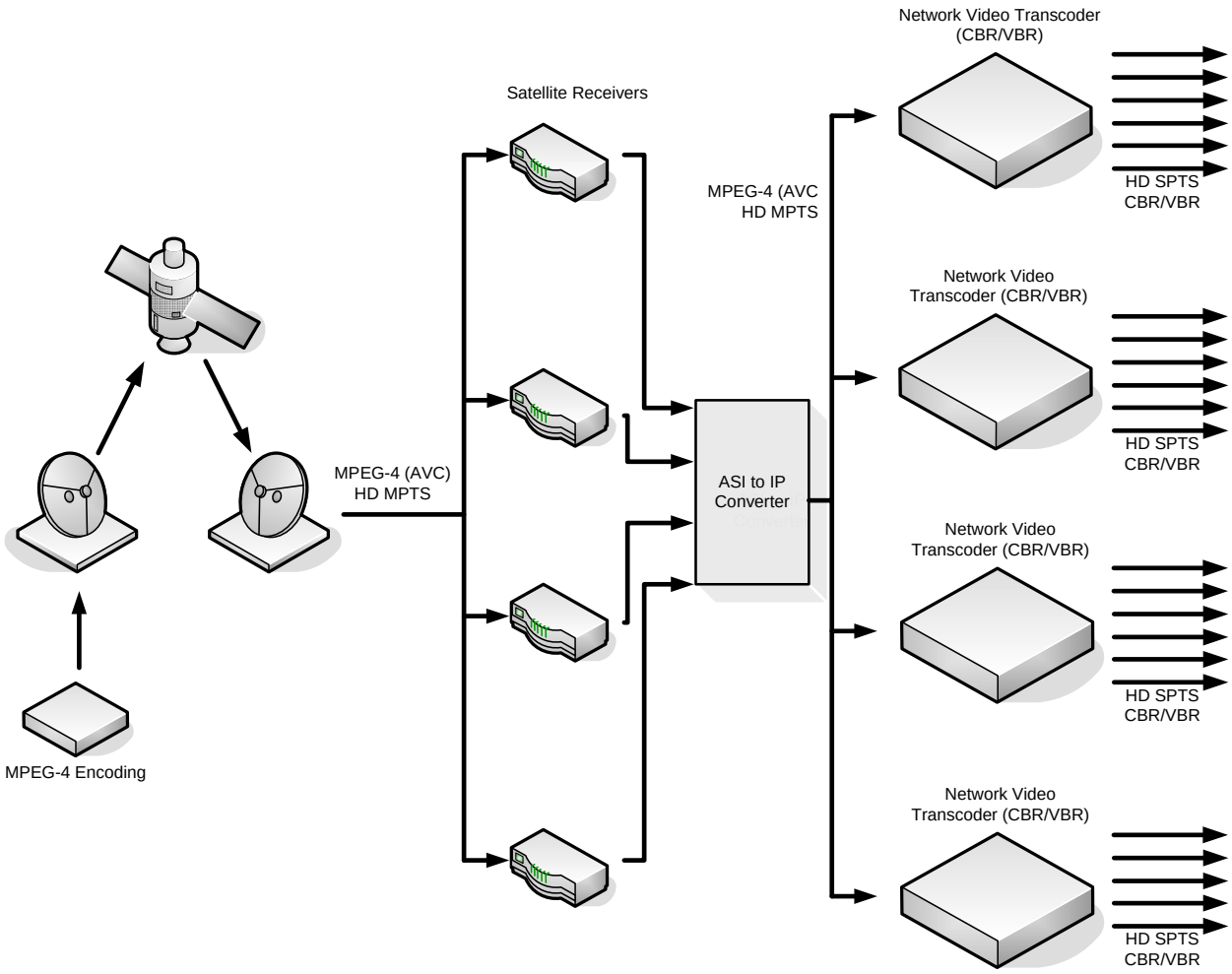

Figure 3
Dense GigE Transcoding –Switched Digital Video Architecture

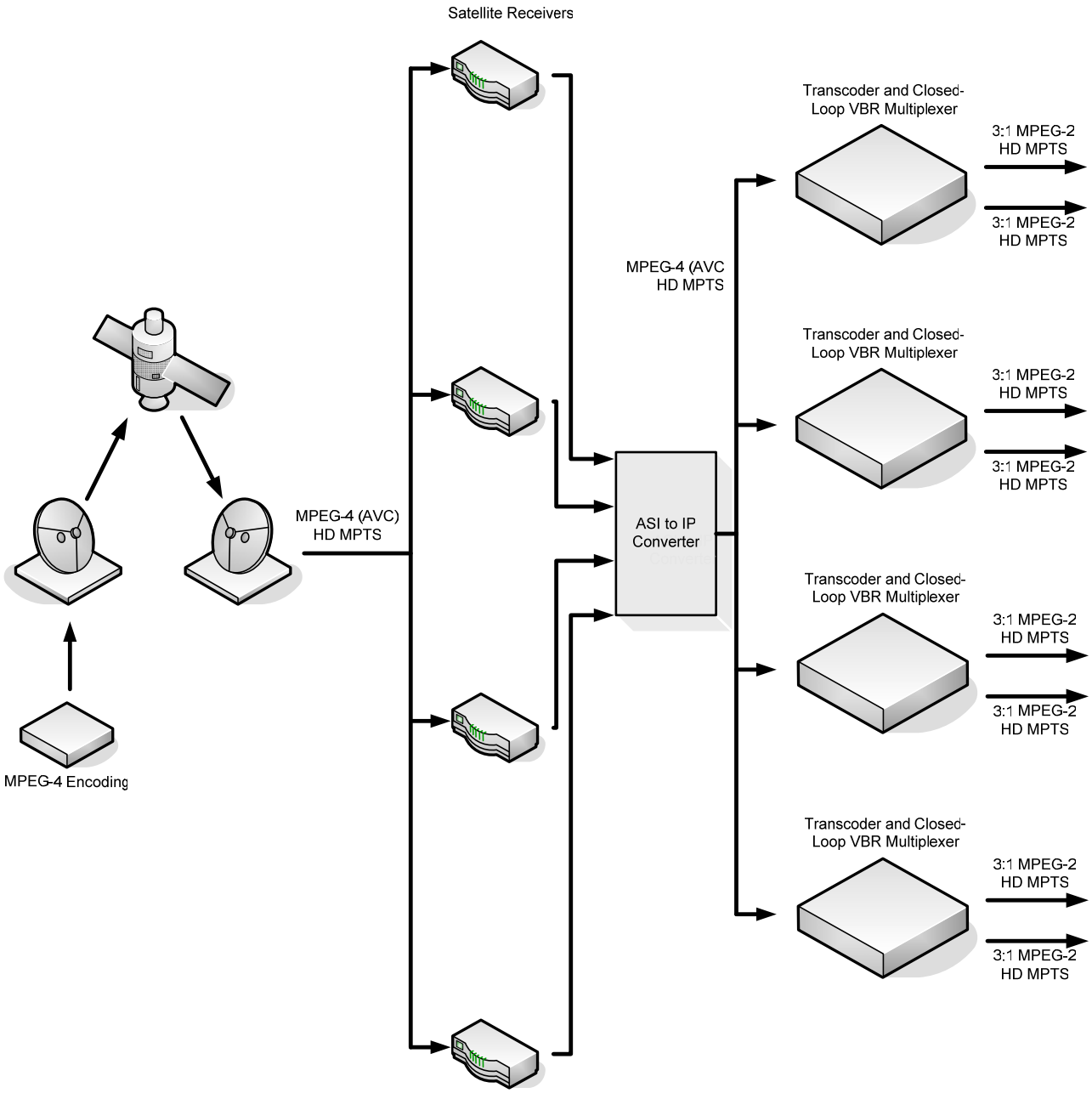

Figure 4
Dense GigE Transcoding – 3:1 Closed-Loop Broadcast Architecture

Conclusion

With HDTV sets now in more than 30 percent of U.S. homes and that number rising sharply, cable and broadcast networks are scrambling to pump out more HD programming to TV viewers. Over the next year, for instance, such major cable programmers as HBO, Discovery Communications and Starz Encore plan to launch a total of more than new 50 high-def channels for distribution.

Under increasing market pressure to carve out enough bandwidth for all this high-def programming, both cable networks and operators will need to upgrade to the more efficient MPEG-4 video compression format. Cable providers will also need to carry MPEG-4 signals to keep up with their satellite and telco TV rivals. DirecTV alone plans to launch more than 100 national HD channels later this year, most, if not all, of them in the advanced MPEG-4 format.

The big challenge for cable operators is that, with more than 34 million MPEG-2-only digital set-tops already deployed in the field, they need a technological solution that will enable them to switch MPEG-4 programming to MPEG-2 swiftly, efficiently, and cheaply. They need a system architecture that will allow them to transcode video signals without hurting picture quality or swapping out all their existing digital set-tops. And they need a flexible, high-density solution that will permit them to convert many MPEG-4 signals to MPEG-2 simultaneously.

Fortunately, such a technological solution exists. The Dense GigE Transcoding system architecture proposed by EGT offers both superior video signal quality and high density, enabling cable operators to convert many channels seamlessly at the same time. Unlike receiver-based transcoding, Dense GigE Transcoding also offers great simplicity and tremendous flexibility, especially for MSOs looking to distribute programming on mobile video and other complementary platforms.

Now that the MPEG-4 era for cable has started, the time has come for MSOs to act. With the aid of the Dense GigE Transcoding architecture, MPEG-4 to MPEG-2 transcoding offers the best way for cable operators to embrace HD programming and start carrying out the industry's long-awaited transition to the digital video compression standard of the future.

References

1. Ricardo Marques, Sergio Faria, Pedro Assuncao, Vitor Silva, Antonio Navarro, “Fast Conversion of H.264/AVC Integer Transform Coefficients into DCT Coefficients,” SIGMAP, pp. 5-8, Aug. 2006.
2. Sandro Moiron, Sergio Faria, Pedro Assuncao, Vitor Silva, Antonio Navarro, “H.264/AVC to MPEG2 Video Architecture,” Conftele 2007, May, 2007.