Envivio IP-based Statistical Rate Control

December 2006
Authors: Boris Felts, Jean Kypreos, Thomas Guionnet

Powering IP Video Convergence From Mobile to HD

www.envivio.com
Table of Contents

1. INTRODUCTION 2
2. VIDEO QUALITY ASSESSMENT 2
   2.1 Subjective Evaluation 2
   2.2 Objective Evaluation 3
3. ENVIVIO STATISTICAL RATE CONTROL 6
   3.1 Principle 6
   3.2 Test Scenario 6
   3.3 Results 6
4. OPEN ARCHITECTURE 11
   4.1 Architecture and Principle 11
   4.2 Open vs. Closed Systems 12
5. CONCLUSION 14
Envivio IP-based Statistical Rate Control

1 Introduction
For many years the broadcast industry has used expensive and proprietary hardware to provide statistical multiplexing features in MPEG-2 transport stream-based networks. With the advent of IP-based networks, the notion of multiplexing and de-multiplexing sources that are bundled together is becoming less important, while the requirement to fit multiple channels in a bandwidth-constrained network remains essential.

Drawing from its expertise in MPEG-4 and IP-based video networks Envivio® has developed a next generation, IP-based, statistical rate control module for broadcast mobile TV networks that will provide the highest quality video services in a bandwidth constrained environment. This technology is not limited to DVB-H. It can also be applied to other broadcast mobile networks such as DVB-SH, DMB, 3GPP MBMS and MediaFlo as well as broadband IPTV, broadcast contribution and distribution applications. By adding statistical rate control capabilities within the core of the 4Caster™ M2 Mobile Series encoder, Envivio enables service providers to both optimize the use of their bandwidth, and also avoid the cost of purchasing proprietary statistical multiplexer or stat mux technology.

This paper describes how to measure the video quality of multiple compressed sources and the quality and bandwidth gained by using Envivio statistical rate control technology. This paper also describes the Envivio open approach to system architecture and how this approach provides significant benefits and advantages for mobile broadcast deployments.

2 Video Quality Assessment
This section covers several methods for measuring video quality and how it is possible to measure the cumulative quality of multiple streams simultaneously.

2.1 Subjective Evaluation
Subjective evaluation relies on the comparison of the original and compressed sources by many
observers in a very well-defined context. Test procedures have been defined by the Video Quality Expert Group (www.VQEG.org) to measure subjective quality of television or multimedia signals. While subjective evaluation leads to the most accurate results, it is complex to implement. Consequently, for the purposes of this paper, while checking for accuracy visually using subjective methods was employed, objective quality measurements were the primary methods used for evaluation.

2.2 Objective Evaluation
Among the various objective criterion to evaluate the video quality, PSNR (Peak Signal to Noise Ratio), PQR (Picture Quality Ratio) and SSIM (Structural SIMilarity), Envivio has selected the SSIM method for this evaluation. Although not perfect, this measurement gives results closer to human observations than the two others and also can be adapted to mobile resolutions.

SSIM is a method for measuring the similarity between two images. It compares the original and the compressed sequence and gives a quality score ranging from 0 to 100. A value of 100 indicates that there is absolutely no difference, while a value of 0 indicates extreme differences.

• Single Source
  The figure below shows the evolution of the SSIM score on a particular video sequence for each frame. The average SSIM gives an overall indication of the quality.

Figure 1 SSIM score for a CIF resolution video encoded at 380 kbps in Constant Bit Rate (CBR) mode.
• Multiple Sources
In order to study the efficiency of the compression of various sources simultaneously, the minimum score for all sources combined is computed (see Figure 2 and Figure 3).

This way of assessing the quality on multiple sources corresponds quite well to human observation. Indeed, observers are particularly sensitive to the worst artifacts and less sensitive to smooth video degradations.

The average Minimum SSIM score gives then a fairly good indication of the quality of a group of compressed sources. This score will be used to assess the performances of various compression methods.

Figure 2 SSIM scores for 10 CIF resolution video streams encoded at 380 kbps in CBR. (Low peaks reveal frame dropping on difficult sequences.)
Figure 3 Minimum SSIM score for 10 CIF resolution video streams encoded at 380 kbps in CBR (average = 64).

A histogram of the SSIM scores (number of samples for a given SSIM score) is also an interesting analysis as its results reflect the video quality homogeneity.

Figure 4 Histogram of the SSIM values for 10 CIF resolution video streams in CBR at 240 kbps. (Frequencies give an indication of the quality homogeneity. Cumulated frequencies indicate the proportion of videos under a given quality threshold.)
3. **Envioio Statistical Rate Control**

3.1 **Principle**

In the case of DVB-H, multiple IP services are transmitted in bursts on a fixed bandwidth channel. In order to respect the overall bandwidth constraint, there are generally two ways of transmitting the IP traffic:

- **Fixed bandwidth allocation:** each service utilizes a constant bandwidth and the streams must be constant bit rate (CBR). The burst size and period are fixed.
- **Dynamic, variable allocation:** each service utilizes a variable bandwidth. The burst size and period are variable.

The latter case is better for video transmission. Given the heterogeneous nature of video signals, it is more appropriate to use more bandwidth on parts of the video that are hard to encode and less on the easier parts. Statistically, there is a high chance that different sources do not need high bandwidth at the same time, so there is an interest in dynamically assigning the bit rate for the transmission of each source. This method is called Statistical Variable Bit Rate mode, or Statistical VBR, and the allocation mechanism Statistical Rate Control.

Therefore, testing will compare the efficiency of the first approach based on CBR, and the Statistical VBR approach.

3.2 **Test Scenario**

The goal of these tests is to measure the efficiency of Statistical VBR against the approach most commonly used today in DVB-H, CBR.

**• Bit Rate Model**

In the discussion below, instantaneous bit rates are indicated \( r \) while average bit rates are indicated \( R \).

The simulations have been done with \( N \) video sources on a channel with a total available bandwidth \( R_{total} \).

In the CBR case, the \( N \) video sources are separately encoded at a constant bit rate, each encoded at \( R_{service} = R_{total} / N \).

In the statistical VBR, the \( N \) video sources are encoded in VBR mode, and at any time their bit rates satisfy:

\[ \sum R_{service} = R_{total} \]

To be close to DVB-H conditions, the following codec parameters were used:

- **Format:** CIF (352x288) at 25 Hz
- **H.264 Baseline with 1s I-frame interval**
- **Measure made on 375 frames (15s)**
• No audio

• Sources
The benefits of statistical rate control are particularly visible with heterogeneous sources. Therefore, a set of video sources were selected with various coding complexities, representing what is usually available on standard TV channels: sports, concerts, movies, news and animated movies.

3.3 Results
For the simulations we used $N = 3, 6$ or $10$ and $R_{\text{service}}$ varying from 200 to 400 kbps. The curves in Figure 5 through Figure 7 show the average Minimum SSIM of all sequences according to $R_{\text{service}}$ in CBR and Statistical VBR modes.

![Figure 5 Average Minimum SSIM of CBR vs Statistical VBR with 3 video sources.](image-url)
Figure 6 Average Minimum SSIM of CBR vs Statistical VBR with 6 video sources

Figure 7 Average Minimum of CBR vs Statistical VBR with 10 video sources
• Observations
As expected, the results demonstrate that statistical VBR offers far superior video quality than CBR, and the positive effect of this mode becomes more and more important as the number of video sources increases. The bit rate allocation algorithm has more options to efficiently distribute the bits where they are needed most.

Since the statistical rate control algorithm focuses on improving the "worst" video, the results becomes more homogenous, and the number of samples with poor quality is reduced. As shown on the histogram of Figure 8, the number of samples with poorer quality is significantly smaller in the statistical VBR case compared to CBR. Visually, this has extremely good consequences – the heavy compression artifacts disappear and all services have the same homogenous quality.

• Statistical VBR Gains
All the curves show some significant video quality improvements at any given bit rate. More interestingly, the potential bit rate gain for a given quality level can be estimated. Looking at Figures 5 to 7 (measured here at Minimum SSIM = 65) this gain is also confirmed by the

<table>
<thead>
<tr>
<th>VIDEO SOURCES</th>
<th>CBR</th>
<th>STATISTICAL VBR BIT RATE</th>
<th>BANDWIDTH SAVINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>310</td>
<td>220</td>
<td>29%</td>
</tr>
<tr>
<td>6</td>
<td>340</td>
<td>220</td>
<td>35%</td>
</tr>
<tr>
<td>10</td>
<td>380</td>
<td>240</td>
<td>37%</td>
</tr>
</tbody>
</table>

Table 1 Comparative bit rates for an equivalent quality and potential bandwidth savings.
histogram on Figure 8. For both CBR at 380 kbps and Statistical VBR at 240 kbps, 95% of the video is above SSIM=65. Furthermore, the video quality is more homogeneous as seen on Figure 9. On this figure, the blue portion indicates areas where samples in Statistical VBR 240 kbps dominate, while the red indicates areas where samples in CBR 380 kbps dominate. The blue portion is more compact than the red indicating the video quality is more homogeneous in Statistical VBR than CBR mode, even at low bit rates. The quality scores are more widespread in the case of CBR at 380 kbps compared to the ones in Statistical VBR at 240 kbps. These results are also confirmed by visual observations. The variation in quality is much less noticeable in Statistical VBR modes and thus leads to a better overall subjective quality.

In conclusion, Statistical VBR mode brings significant improvements over CBR. Even with a limited set of sources, the benefits of this method are quite dramatic in terms of artifact reduction. The efficiency increases with the number of heterogeneous sources, and it can be estimated that 30% to 40% performance improvement can be achieved in average over CBR, which was already greatly optimized.
4. Open Architecture

4.1 Architecture and Principle

The Envivio statistical rate control implementation is not dependent on an external module. This functionality is available as an option directly from the encoder and does not require any extra components.

For an Envivio implementation, an encoder working in Statistical VBR mode is assigned to a group. This group represents all the encoder profiles which are acting under the same statistical rate control and will later be sharing the same bandwidth.

- Master/Slave Roles

Within a group, a particular encoder acts as “master,” another one as backup of the master, and the other encoders are “slaves.” The master gets the overall bandwidth assignment from the operator or the management system, listens to the slaves’ requests and allocates the bit rates dynamically. All communication is done via IP protocols.

![Figure 10 DVB-H head-end with master and slave encoders](image)

- Multi-Output and Statistical Rate Control

Most Envivio 4Caster™ encoders, and notably the Envivio 4Caster M2, have the ability to output multiple profiles simultaneously from the same input. Each of these profiles can obey different statistical rate control parameters. For example, 4Caster M2 can output 3 CIF DVB-H profiles, each belonging to a statistical rate control group.
This flexibility can have great cost advantages when it is necessary to compose various statistically rate controlled DVB-H groupings, or “bouquets,” sharing common channels. Instead of using dedicated encoders for each bouquet, the same encoder can be reused to output the same channel to various bouquets under different rate controls. Thus significantly decreasing the head-end cost compared to encoders tightly linked to a DVB-IPE (IP Encapsulator) where equipment needs to be replicated for each bouquet.

Figure 10 National/Regional combined head-ends with multiple stat mux groups, each with various master encoders

• Future Extensions
The “master” functionality has been placed on one encoder, and answers the need for small to medium size deployments which can typically be found with DVB-H. In the future, this functionality will also be available from the central management system, 4Manager, in order to control multiple groups simultaneously and address scalability and redundancy issues in much larger deployments.

4.2 Open vs. Closed Systems

• Closed Systems
Quite often, statistical multiplexer implementations rely on strong ties between the encoders and a multiplexer. Similarly, in the case of DVB-H, most vendors rely on proprietary communication mechanisms between their encoders and their DVB-IPE to insure the statistical rate control function.
In these closed systems, the DVB-IPE is the main "brain" for the service bit rate allocations. It dynamically splits and distributes the bandwidth between all encoders according to the needs of each source. It then performs the proper burst allocation before transmission.

While this kind of system is efficient in performing statistical multiplexing, it is extremely difficult to decouple the encoders from the multiplexer, and there is no standard way of negotiating rate allocation between these two kinds of elements. Practically, this means that it is not possible to mix equipment (encoders or multiplexers) from different vendors without losing the statistical multiplexing functionality. Until now, this has helped broadcast equipment vendors lock customers into a single brand solution.

Open Systems
The Envivio approach is radically different from this closed loop implementation. The Envivio advanced statistical rate control does not rely on any proprietary mechanism between encoders and the IPE. Instead, the rate control allocation is done between encoders, while the encapsulator only takes care of burst allocation and is not obliged to send any control or indication to the encoders.

The statistical rate control feature is provided within a module of the encoder. This module communicates with all the other encoders belonging to the same group and simulates the emission buffer of the DVB-IPE. It can be set to operate at fixed bandwidth, as was done here to perform the comparisons, or set at dynamic bandwidth. In this latter case, an external mechanism can increase or decrease the overall bandwidth attributed to the stat mux group. This can be interesting if other services need to take priority over the video services or if extra bandwidth is available for transmission.

The open architecture provides significant advantages:
- encoders and IPEs can come from various vendors
- encoders and IPEs do not have to be collocated
- no lock-in factor means best-of-breed approach
- higher reliability (elements are more autonomous, so the entire system is less prone to bottleneck failures)

Summary

<table>
<thead>
<tr>
<th>ADVANTAGES</th>
<th>DRAWBACKS</th>
</tr>
</thead>
</table>
| Closed Systems  
Efficient  
Potentially lower delay  
No vendor interoperability required | No possible equipment mix  
More bottlenecks in the system  
Locked into one brand solution  
Equipment needs to be collocated |
| Open Systems  
Efficient  
Best of breed approach  
More architecture solutions available  
Higher reliability  
Vendor interoperability  
Higher number of vendors | |

Envivio IP-based Statistical Rate Control
5. Conclusion

Envivio statistical rate control technology and its instantiation in the 4Caster Mobile M2 encoder bring two significant advantages to the market:

**Picture Quality Improvement or Bandwidth Savings**
- A very efficient Statistical Variable Bit Rate mode, which brings significant advantages over Constant Bit Rate. This leads to either dramatic picture quality improvement for the same overall bit rate, or a reduction in the bandwidth required per service, allowing more services to be carried.

**Open Standards-Based Implementation**
- An open and IP-oriented architecture, allowing advanced services and the creation of best of breed solutions with various component vendors.

Though the benefits of statistical rate control for mobile, and particularly DVB-H were covered in this paper, the same technology and principles can be applied to other kinds of environments, which traditionally use CBR, such as:

**Contribution of Multiple Feeds**
- Statistical rate control delivers some quality or compression improvements even with a small number of encoders, from which satellite and IP contribution could draw benefits.

**IPTV or EnterpriseTV backbone cost reduction**
- In order to save bandwidth on the IP distribution networks (backbone in the case of IPTV, or private networks in the case of Enterprise video distribution) this technology could be applied

Thanks to its efficiency and openness, Envivio statistical rate control enables some significant quality and service improvements to the operator, for mobile, IPTV and contribution networks.