

Achieving Maximum Accuracy With Real-Time Video Quality Measurement

The economics of the media and entertainment industry require that broadcasters and delivery networks provide a quality product every time to every screen to ensure that viewers get the best possible experience. The aim, of course, is to retain — and gain — viewers, subscribers, and advertisers. In order to ensure they are delivering high quality consistently, program originators and distributors alike must be able to measure their audiovisual content and determine if that content meets standards — whether those standards are self-imposed, set by the customer or compliance regulators, or some combination of those. The more quickly and easily the measurements can happen, the better. If program originators and delivery networks simply had to manage output for a few television channels, as they have in the past, then the quality-control process would be relatively simple. But now that they have to accommodate multiple screens well beyond the TV set, and with higher and higher definition, the number of channels, formats, and bit rate configurations is impossible to manage with traditional testing and quality control (QC) solutions. When originators and distributors can find a way to measure all those content streams quickly, efficiently, and automatically, they will not only be able to save time and effort on measurement — resources that can be put to use doing other important tasks — but also ensure they’re meeting SLAs and compliance regulations. Ultimately, automated measurement and QC can help ensure a quality viewing experience, with greater retention rates of both viewers and advertisers.

Why Pay Attention to Video Quality?

In examining the industry, the trend is toward more programming and more devices with varying screen sizes and varying levels of quality. At the same time, new and more efficient encoding standards will allow for additional channels over the same or smaller data rates. This trend drives the need for additional tests by both manufacturers and service providers, which in turn drives the need for tools to make QC accurate and manageable. Content providers must be able to deliver quality content to multiple screens in an environment that is constantly changing. The way to ensure it happens is by testing and QC, not only after the content goes live, but also before.

The idea of “Entertainment Everywhere” is taking hold. Streamed or downloaded programs to a TV, laptop, tablet, or any size mobile device — either through over-the-top (OTT) means from services such as Amazon, Netflix, and Hulu, or through live television channels and download services via a mobile LTE network — must be provisioned for any device, anytime. As in the example in Figure 1, streamed programming must be delivered in a variety of bit rates so that any device

can adapt its playback for any change in network condition on the Internet.

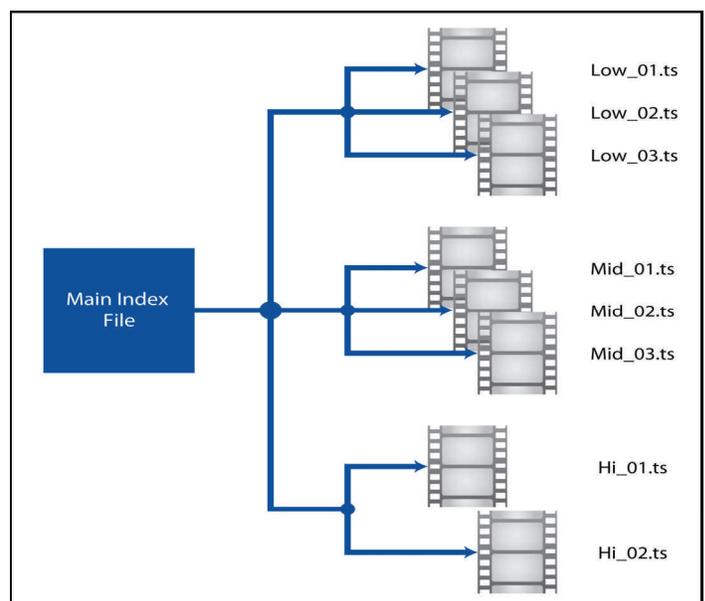


Figure 1: Example of OTT Streams Per Program

Quality Control vs. Quality Measurement

Measuring quality in any type of entertainment delivery network, TV or otherwise, means checking a number of criteria, many of which are IT-based (such as whether the network path is working). The testing process has traditionally existed to check network performance and delivered programming after it has gone to air — the basic definition of QC. But this process is evolving to include quality measurement as well. That is, measuring the true quality of any given content throughout the entire processing chain. IT testing for network performance and testing for A/V quality are very different types of testing that require different types of measurement equipment and different test procedures. Whereas there is a fairly standard, objective QC test to make sure the infrastructure is functioning properly, measuring for A/V quality is far more subjective because it is based on human perception.

It used to be that both types of tests were done manually, that is, with one or more people viewing the material and deciding the quality with their own eyes — a subjective, laborious, imprecise process. That method has evolved to the point where certain products can accurately and automatically execute QC to determine whether something is performing to a certain level over time — such as lip sync (audio alignment) and loudness in audio or caption corruption in video — in real time using scientifically proven algorithms. They can also accurately measure video performance automatically in real time using similar advanced, perceptual methods of determining quality without subjectivity.

Perceptual Quality Analysis: The Key to Accurate Quality Measurement

In-depth A/V quality analysis is highly subjective. The most precise way to measure it is to gather human observers and ask them to judge the quality, which is an expensive, time-consuming and possibly inconsistent approach. What makes accurate A/V analysis possible in media operations today are test methods that use a numeric quality score that is correlated to standardized, subjective databases. The databases are derived from human perceptual studies based on Recommendation ITU-R BT.500-13 (01.12) — Methodology for the Subjective Assessment of the Quality of Television Pictures. The studies yielded two databases of mean opinion scores (MOS). A number of algorithms have been developed to estimate perceived quality in a precise way. The results of these algorithms are then correlated against correctly produced subjective data under ITU-R BT.500-13 (01.12). The result is a measurement of subjective quality very closely approximating a human's perception of the picture quality.

The algorithms are divided into three general types, the most accurate of which are full-reference algorithms, which compare the processed and reference sequences. Full-reference

algorithms yield the most accurate results because the equipment performing the test is supplied with two copies of the content: a source (or reference) version of the video content, and a version that has been processed through some type of network or equipment. The powerful full-reference technique allows measurements to be made in service, so that normal video delivery can continue uninterrupted while tests are being performed.

Multistream Testing: The Next Step in Quality Measurement

Thanks to the advanced analysis methods mentioned above, we are able to measure both performance and perceptual quality accurately in real time. The next step in the evolution of quality measurement is to develop tools that can do it with maximum accuracy on multiple content streams.

When it comes to quality measurement, the requirements are moving targets. The challenge for delivery networks is to tailor and test for in-network quality and customer quality of experience on a significant number of screens and devices at different data rates, and to accommodate a range of processing power, resolutions, available bandwidth, sizes, and more. There will continually be new variables to address, and the variables are constantly changing and evolving. Then there are service requirements to meet, a dynamic and expanding slate of content, and a rapidly growing number of channels. All of these factors make multichannel quality measurement incredibly complex.

The process will always require experienced engineers to “eyeball” the content, but having highly accurate quality measurement methods to match their perceptions is critical.

Ideal Candidates for Multistream Quality Measurement

1. Content originators — Organizations that are producing content and delivering it over broadcast networks (including terrestrial, satellite, cable, and IPTV) and then also doing OTT services with those same assets. Content originators have an ever-increasing number of delivery requirements and must ensure their content is delivered at both constant and adaptive bit rates within set parameters for video quality, audio quality, loudness, and lip sync. Needs have evolved such that more people in the organization must not only understand and adhere to quality levels, but also be able to demonstrate it to management — so that management can prove it to advertisers and compliance bodies. There's also a basic need for broadcasting and satellite TV providers to budget their bandwidth, so testing is critical to reducing bit rates and helping those downstream providers make the best use of their bandwidth. With the number of channels growing exponentially, and everyone throughout the delivery chain having different implementa-

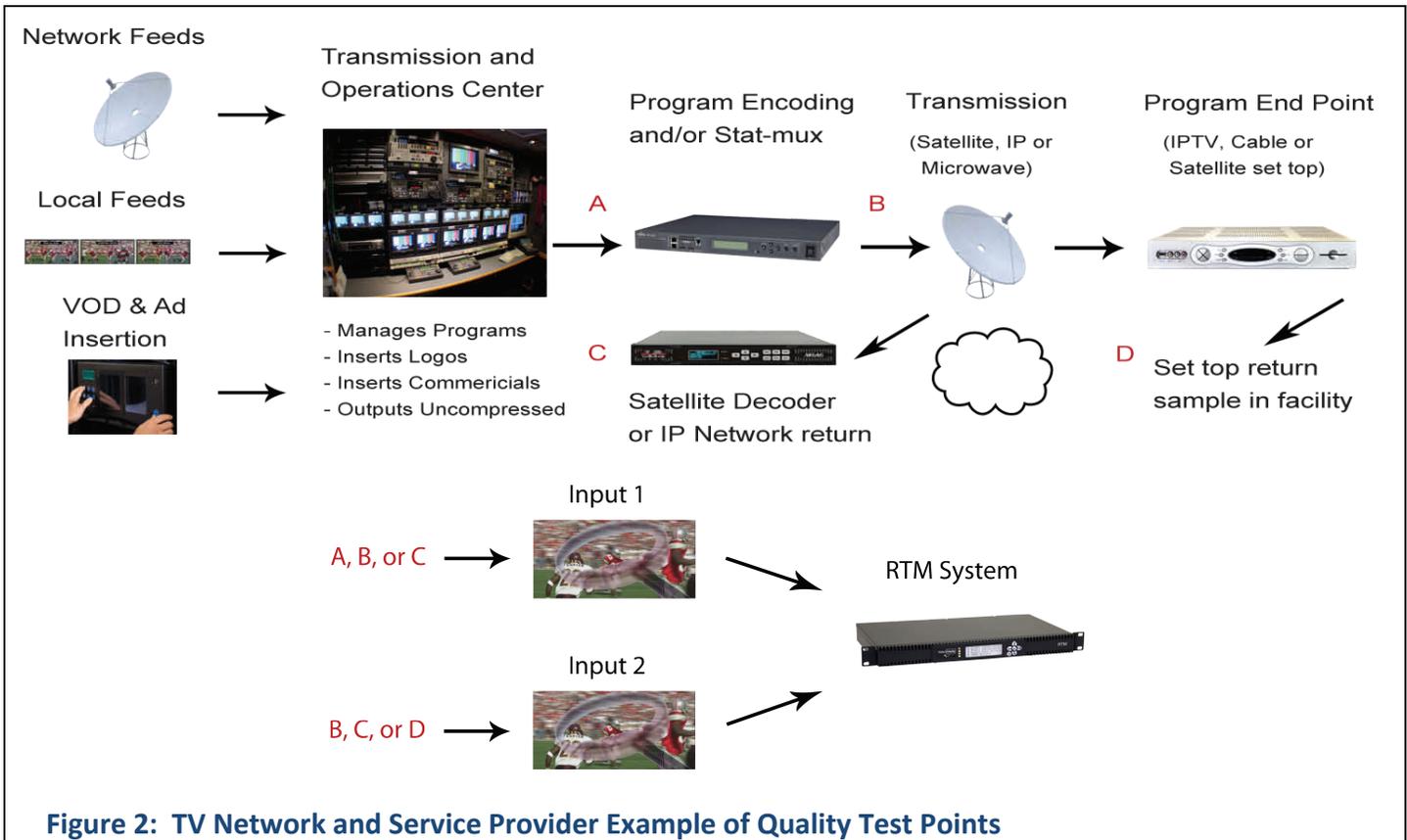


Figure 2: TV Network and Service Provider Example of Quality Test Points

tions, it is more important than ever for content originators to be able to provide objective, numbers-based guidance up and down the chain.

2. Secondary delivery service providers — Organizations that take programs (channels or assets) from content originators and deliver them to the end customer. This category includes cable, IPTV, direct broadcast, and satellite providers, as well as ISPs, CDNs, and hybrid organizations that operate as both originator and delivery service. Audio loudness is measured using the ITU standard for LKFS measurement. With this capability, any video segments where the loudness threshold is exceeded can be captured for later analysis. This capability allows a delivery network to monitor compliance with the CALM act and other regulations

Why Existing Options Are Not Enough

Some existing solutions do a fairly good job of identifying syntax and/or data and IP issues in both the file and the IP domains. Those well-deployed solutions can tell customers whether the data packets/streams are arriving at their destinations intact, but they can't identify with a high degree of accuracy the quality that's being delivered. They can't do enough to monitor what comes out of the endpoint decoder or define the true, reference-based quality consistently throughout the chain.

At this point, "eyeballing it" is the norm, but as more and more revenue becomes associated with OTT content, eyeball testing won't be good enough. Users must be able to measure accurately the quality of the source to the quality of the endpoint, correlated back to the measurements taken on their own out-bound signal. The only way to do that is with a solution that relies on full-reference test methods for perceptual quality, but there has been no such viable solution.

The Benefits of a Full-Reference Solution

The accuracy of a full-reference solution, especially for a testing environment, cannot be overstated. Such a solution would yield benefits for content originators and second-level delivery providers.

First, it would eliminate the need to rely solely on an engineer to make quality decisions. While they would still do visual assessments, engineers would have an objective, highly accurate, numbers-based method and scoring system for quality that correlates to the subjective score they would assign during visual evaluation. The reliable scoring system could be used throughout the delivery chain and applied consistently to all types of content no matter what the delivery channel.

Video Clarity RTM Solution with RTM Scheduler

RTM Scheduler is a new tool for users of the company's RTM real-time audio and video monitoring solution. RTM Scheduler allows operators to use a single RTM unit to monitor several programs or channels in a series based on a user-defined schedule.

How It Works

RTM Scheduler generates a sequence of RTM server commands according to a schedule defined by a user-generated, tab-delimited input text file usually called "rtmcron.tab." The input file should contain fields for:

- Date and time to launch the command sequence
- The IP address for the target RTM unit
- ";"-delimited sequence of RTM commands

Rows that specify a "daily" schedule indicate that the corresponding commands should be invoked at the specified time each day, while an "hourly" schedule dictates that the corresponding commands should occur at the specified time each hour of each day.

The tool will continue to execute as long as there are RTM commands scheduled sometime in the future. If the input text file contains lines starting with "daily" or "hourly," RTM Scheduler will continue to run until it is manually stopped. Figure 3 shows a sample input text file and schedule.

Figure 3: In this example, the scheduler is telling RTM to switch unicast addresses every 10 minutes. A series of commands — separated by semicolons — directs RTM to check the status, stop RTM, configure the new input, and then start back up again.

To run the scheduler, users select the tool in Windows Explorer, which opens a command window that also captures log status. Alternatively, users can start it directly from a command window. In either case, status for all runs is logged in rtmcron.txt. If users modify the input file while the tool is running, RTM Scheduler will immediately interpret the changes and regenerate the schedule internally while it is executing.

Typical Usage

Although RTM systems monitor a single channel at a time, RTM Scheduler makes it easy for users to program an RTM unit to perform multiple sequential tests, thereby making it possible for a single RTM unit to perform quality measurements and fault-monitoring sessions on several channels in sequence automatically. This capability is especially useful when the RTM unit is monitoring an SDI source against an IP-processed source. As a source for the test reference, RTM can select from an IP stream or up to four SDI interfaces sequentially. The RTM unit can then select any processed IP stream (or series of streams) for quality measurement and monitoring.

Conclusion

Given today's complex and dynamic content delivery environment, where more and more revenue is associated with adaptive-bit-rate OTT content, the ability to measure the quality of multiple streams automatically in a highly accurate way is critically important. It helps content originators and distributors ensure that viewers get the best quality of experience, advertisers are satisfied, and regulations are met. RTM Scheduler, a tool designed for use with Video Clarity's RTM real-time monitoring system, allows operators to use a single RTM unit to monitor several programs or channels in a series automatically — using full-reference methodology that yields highly accurate measurements of perceptual quality. RTM Scheduler makes it possible to test multiple channels for quality and performance.

```

# Change unicast every 10 minutes
# Date      HH:mm:ss      IPaddr      RTM commands
daily      00:00:01      127.0.0.1   version;versiondate
hourly     00:00:01      127.0.0.1   status;stop;configinputip 1 231.1.1.1 1111 udp 1;start
hourly     10:00:01      127.0.0.1   status;stop;configinputip 1 232.2.2.2 2222 udp 1;start
hourly     20:00:01      127.0.0.1   status;stop;configinputip 1 233.3.3.3 3333 udp 1;start
hourly     30:00:01      127.0.0.1   status;stop;configinputip 1 231.1.1.1 1111 udp 1;start
hourly     40:00:01      127.0.0.1   status;stop;configinputip 1 232.2.2.2 2222 udp 1;start
hourly     50:00:01      127.0.0.1   status;stop;configinputip 1 233.3.3.3 3333 udp 1;start
  
```

Figure 3: RTM Scheduler Example

For more information, see the following documents:

- Application Note: RTM Scheduler: <https://videoclarity.com/PDF/RTMScheduler.pdf>
- RTM data sheet: <https://videoclarity.com/PDF/RTMDataSheet.pdf>
- Video Clarity published paper, “Understanding MOS, JND and PSNR”: <https://videoclarity.com/videoqualityanalysis/casestudies/>
- Video Clarity published paper, “Real Time Monitoring Explained”: <https://videoclarity.com/videoqualityanalysis/casestudies/>
- ITU-R BT.500-13 (01.12) “Methodology for the Subjective Assessment of the Quality of Television Pictures”: <http://www.itu.int/rec/R-REC-BT.500/en>