The Mandate for High Availability Video Programming

Ashton, Metzler & Associates

Leverage Technology & Talent for Success

Introduction

As digital broadband technologies continue to evolve, the market for delivery of standard definition and high definition digital television programming to the residential subscriber is becoming increasingly competitive. Depending on their geographical location, a digital TV customer might be able to choose among a number of service providers, including cable companies, telecom companies with DSL or Passive Optical Networks, satellite companies, and companies delivering services over wireless broadband, such as WiMax or WiFi.

In order to remain competitive, providers of digital TV services will need to achieve high levels of customer satisfaction to help them to gain new customers and to minimize subscriber churn. The high level of competition also makes it imperative to control costs, including the costs of customer support and the costs of providing remedial services to deal with impaired signal delivery.

The key to maximizing customer satisfaction and controlling support costs is to build a highly reliable end-to-end infrastructure that has redundancy and resiliency comparable to that of the traditional POTS telephone network, which is generally regarded as being 99.999% reliable (a.k.a., five 9's of availability). This type of infrastructure can help service providers reduce the number of preventable issues that drive customer calls, which in turn helps improve satisfaction. As will be explained in a subsequent section of this white paper, in order to achieve end-to-end availability of 99.999%, each type of element on the end-to-end path between the signal source and the subscriber (Figure 1) needs to be available more than 99.999% of the time.

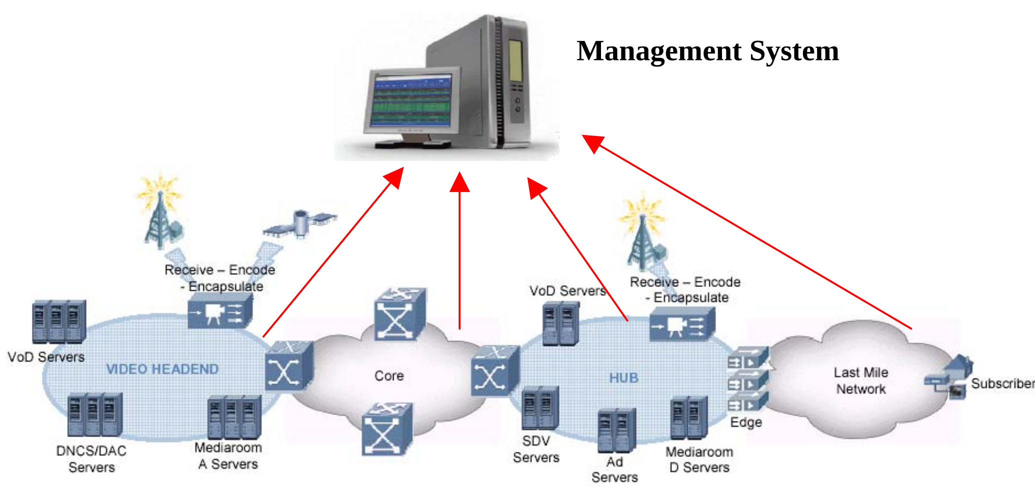


Figure 1: A Digital TV Network with a Routed IP Core

Another characteristic of a highly reliable network is the deployment of effective fault management instrumentation that allows rapid root cause diagnosis and timely repair of the conditions that impair signal quality. The fault monitoring instrumentation needs to be distributed among the various tiers of the network in order to effectively isolate the root cause of errors.

The two primary goals of this paper are to examine how availability should be defined in the special context of digital video programming and to explore the implications for fault monitoring equipment and network equipment testing procedures.

To achieve those goals, six executives at companies who provide video services were interviewed. Two of the executives cannot be referenced by name or company in this white paper. Those executives were a vice president of advanced technology at a cable company and a director of quality assurance at a telecommunications vendor. They will be referred to in this white paper as The Cable Vice President and The Telecommunications Director. Other interviewees were Hung Nguyen, the HMS Committee Lead at SCTE[1]; Stuart Elby, VP of Advanced Technology Networks, Verizon; Jerry Murphy, Emerging Technologies, Telus; and Charlotte Field, Senior Vice President, NETO Infrastructure and Operations, Comcast.

Today's Video Services

There was unanimous agreement amongst the interviewees about the importance of video services. Field, for example stated that, "Video services are critical to Comcast and becoming even more so as additional high definition channels become available." The Telecommunications Director added that video services were driving a transformation within his company and that his company "is betting our future on video services."

There was also wide spread agreement that delivering high quality video services is very challenging. Murphy stated that his company has not been able to take their background in managing the traditional telecommunications network and apply that background directly to video services. He added that the lack of effective troubleshooting tools results in a lot of unnecessary truck rolls. Field stated that poor video quality tends to frustrate customers and that frustration is increased if the customer stays home for a half day for a service call that does not resolve the problem. Elby pointed out that customers often call to complain about a problem that occurred sometime during the previous few days and that is the problem is not currently occurring, identifying the cause of the problem is "very, very difficult".

Nguyen highlighted a fundamental challenge associated with delivering high quality video services. He stated that the industry has done a good job of delivering relatively high quality analog video signals. However, he added that delivering a digital video signal is relatively new, and is much more complex because it requires so many components, each of which could either fail or could introduce some form of degradation.

The Telecommunications Director pointed out another fundamental challenge in delivering high quality video services is that the networking equipment that they deploy to support these services lacks the same embedded management capabilities and product resiliency that they are used to having in the traditional telecommunications environment. Murphy and Field agreed with The Telecommunications Director. Murphy stated that, "There is nothing built into routers that will give us meaningful management data about the quality of video services." He went on to say that, "Troubleshooting video quality based just on use the command line interface (CLI) of a router is pretty much impossible." Field added that part of the problem that they are experiencing is that they "are pushing our routers to a different level and that issues that may seem miniscule from a router perspective can be very harmful from a customer perspective."

The Telecommunications Director discussed some of the cultural challenges associated with managing video quality. He said that the part of his company that focuses on traditional telecommunications expect metrics that demonstrate the quality of the service that they are providing. He stated that, "Many of the old time video folks prefer a visual inspection of the

[1] HMS refers to the Hybrid Management Subcommittee. SCTE refers to the Society of Cable Telecommunications Engineers.

video to a reliance on metrics." The Telecommunications Director concluded by saying that the engineering of his company's network has been very good but that now the challenge is how to better manage the services that transit that network.

The Cable Vice President stated that, "The cable industry is heading for a problem that nobody has seen yet." He elaborated by stating that, "VoIP was the first product that cable operators got into that is time sensitive and crosses multiple physical boundaries. Our experience with VoIP highlights a massive hole in terms of what it takes to support services such as VoIP and video."

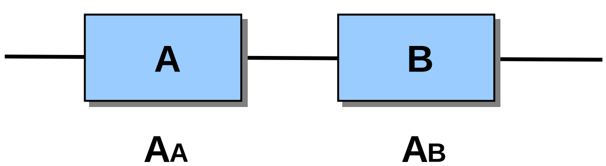
Availability

Availability is traditionally thought of as the probability that a component, system, or service is operable at a given instant in time. Based on this definition, availability is calculated (Equation 1) as the percentage of the time that a component, system, or service is capable of performing its intended function over a measured time interval. Downtime is the measured time interval where the entity is not available.

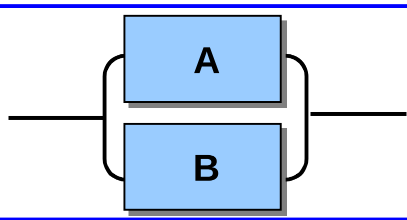
$$\text{Availability} = (\text{time interval} - \text{downtime}) / \text{time interval}$$

Equation 1: Availability Calculation

When a system is composed of numerous components connected in either series or parallel configurations, the system availability can be calculated from the availability of the individual components[2] using the equations shown in Figure 2.



$$\text{Combined Availability of A,B in series} = A_A \times A_B$$



$$\text{Combined Availability of A,B in parallel} = 1-[(1-A_A) \times (1-A_B)]$$

Figure 2: Availability of Series and Parallel Subsystem Configurations

[2] In Figure 2, AA and AB refer to the availability of component A and B respectively.

To exemplify the equations shown in Figure 2, assume that there are two components each of which has 99% availability. If the two components are arranged in series, then the combination has an overall reliability of 98%. However, if the two components are arranged in parallel, the combination has an overall reliability of 99.99%. Continuing with the example, now assume that there are ten components, each of which has 99% availability. If the ten components are arranged in series, then the combination has an overall reliability of 90%. However, if the ten components are arranged in parallel, the combination has an overall reliability of virtually one hundred percent.

Elby stated that from purely a network perspective his company is providing “a little better than four 9’s of availability.” He pointed out that even if the network is available, the customer could experience degradation in the video services that transit the network. The Telecommunications Director said that, “Coming from a telecommunications background, five 9’s of reliability is the Holy Grail. We are currently struggling to get to four 9’s of reliability with the hope of getting to five 9’s.”

Availability of Network Devices

The definition of availability traditionally used by network equipment manufacturers is based on the Mean Time to Repair/Recover (MTTR) after a fault occurs. MTTR may reflect the time it takes for a service technician to diagnose the problem and replace a circuit board, or it may reflect the time it takes for a subsystem within the device to fail-over to a redundant backup subsystem. Typical fail-over times are in the range of 50 milliseconds to 30 seconds. As is shown in Equation 2, another common way to calculate availability is based on the Mean Time Between Failure (MTBF) and the MTTR.

$$\text{Availability} = (\text{MTBF}-\text{MTTR})/\text{MTBF}$$

Equation 2: Traditional Availability Calculation for Network Devices

Assuming a MTTR of one hour, a device that is intended to be 99.999% available would need to have an MTBF of 12 years. Since measuring an MTBF of this duration in a lab is generally not feasible, this leads to the common practice of estimating or predicting the device’s MTBF based on the expected MTBF of the each of the components included in the device.

Device MTBF is initially calculated by categorizing and counting all the components within each subsystem of the product design and by consulting either a handbook (MIL-STD 217 or Telcordia TR-332) or data from the component vendor to determine each component’s expected MTBF. Later in the product life cycle, the calculated MTBF may be adjusted to take into account real data from a large number of devices operating in production environments.

There are a number of shortcomings of typical device level availability calculations:

1. There is no standard that specifies how vendors should compute availability specifications. For example, some calculations may assume all components are connected in series arrangements, while other calculations may take into account redundancy relationships among components or subsystems.

2. The focus is on the catastrophic failure of a hardware component, possibly ignoring software errors altogether. As a result, the primary value of the calculated MTBF is in estimating the number of spares required and the number of service technicians that will be required to maintain the system with an acceptable MTTR.

3. Transient errors that occur in production networks are not considered. Errors in this category include dropped packets due to congestion, out-of-order packets, or excessively delayed packets.

4. Because of the definition of downtime, the MTBF and availability specifications do not reflect in any meaningful way the user's quality of experience, especially for real-time applications such as voice and video services.

Availability of Video Programming

In order to assure an acceptable quality of experience (QoE) for digital video subscribers, service providers must adopt a different definition of availability than has been used traditionally. The new definition of availability must be based on the notion that the video service is not available whenever the quality of the service is below some threshold. After the QoE threshold has been defined, measurements of downtime can be made and the availability can be calculated and compared to the targeted level; e.g., 99.9% or 99.99%. Well-defined video QoE thresholds would also be valuable in communicating reliability requirements to network equipment manufacturers (NEMs). Ideally, the NEMs would be able to modify their testing procedures to support an additional set of QoE availability specifications that would be meaningful for digital video service providers.

While there would be significant benefits from an industry standard metric that specified exactly what constitutes an acceptable QoE for video services, such a standard has not yet been established. However, considerable progress is being made toward defining such a standard. The Society of Cable Telecommunications Engineers (SCTE) defines via standards document SCTE-142 different categories of performance degradation and attempts to prioritize them. In addition, the SCTE HFC Management Subcommittee (HMS) has been working on specifications and best practices for video systems monitoring and video systems test procedures based on input from both service providers and NEMs. The current draft of the HMS-158 Digital System Video Monitoring document bases its definition of service downtime on the *errored second.* With this approach, any second that contains one or more errors (e.g., lost program packets, out of order or duplicate packets, excessive jitter) is considered to be a second of downtime. The errored second can also be used as the basis for other less stringent definitions of downtime (e.g., downtime might be considered to begin after several consecutive errored seconds). The SCTE HMS-158 draft "Recommended Practice for Monitoring" suggests that an HD video program is of acceptable QoE with one errored second per four hours of viewing, while a SD video program is acceptable at a threshold of one errored second per hour of viewing. Per Equation 1, these error thresholds correspond to required system-wide QoE availability of 99.993% for HD and 99.972% for SD.

These HMS recommended errored second thresholds and availability targets are undoubtedly too aggressive for many service providers to implement in the near future due to constraints related to equipment replacement schedules and CAPEX limitations. However, the mandate is clear. The successful delivery of digital TV requires that service providers begin to establish meaningful QoE thresholds that are supported by video monitoring equipment capable of detecting even short duration errors that detract from the viewing experience. Such errors may occur at the IP packet layer or within the MPEG payload and are experienced by the viewer in a number of ways, including the inability of digital TV receivers to tune properly, lack of on-screen program information, noticeable lip sync errors, and compromised or missing audio or video.

Field and Elby agreed on the need to adopt a different definition of availability. Field stated that, "The old way to manage the network was to focus just on availability, and the new way is to focus on both availability and performance." Elby added that his company is working with the rest of the industry to come up with a metric that quantifies video quality that is the equivalent of a mean opinion score (MOS) for voice. He concluded by stating that developing a metric for video quality will be "an order of magnitude more difficult than developing a MOS but that we need to change the video monitoring paradigm."

A Process for Improving Video Program Availability

Nguyen said that "monitoring video is a huge effort in part because there are so many platforms and in part because each piece of equipment has its own monitoring." Field stated that within six months to a year that "More and more service providers will look proactively at the customer package across all services." The Cable Vice President stated that, "We are at the cusp of new services that have monitoring requirements that have not been defined yet." He questioned, "How do you tell one bit from another?" He concluded that you cannot tell one bit from another and that every bit needs to be treated "as if there were an SLA assigned to it."

Given the challenges described by Nguyen, and The Cable Vice President, service providers that are looking to improve the QoE associated with their video services can initiate a reliability improvement program similar to the one outlined below:

1. Establish a simple QoE threshold metric that can be used for measuring subscriber-perceived downtime. This threshold could follow SCTE/HMS recommendations or be based on customer research within the provider's subscriber base. The selected metric should be compatible with what is possible given the existing monitoring and test equipment.

2. Participate, at least to some extent, in the standards body efforts to establish industry metrics for QoE thresholds.

3. Deploy QoE monitoring equipment at the edge of the existing network to gather per program downtime statistics as experienced by the customer, and use this data to characterize the current availability of video services. The results gathered at this phase of the process may require some modification of the QoE threshold metric followed by additional monitoring.

4. Deploy additional monitoring equipment as required to allow the availability of different subsystems within the infrastructure to be measured or estimated. These results can be used to identify the major contributors to video downtime.

5. Establish availability test procedures that can be used to aid in the selection of key system components. Typically the test procedures should involve full line rate testing with real video streams that are continuously monitored for QoE errors at both the input and the output of the network device. Stress tests should also include parallel data and voice traffic streams, as well as the normal throughput, latency, and jitter tests for these non-video streams. The video program test setups and results should be shared with NEMs and independent test labs to help encourage the wider adoption of video program availability testing procedures.

6. As higher levels of program availability are achieved, install monitoring equipment within multiple tiers of the network (e.g., headend, core, hub, and edge). More comprehensive monitoring will provide granular information to help guide system upgrades and to help control OPEX by facilitating the rapid diagnosis and repair of faults. Figure 3 provides an example of a possible availability report gathered from a number of monitors distributed across two regional systems, each consisting of a core feeding a pair of egress hubs. This example shows that a fault detected in Core A has reduced Animal Planet program availability below a 99% threshold. This fault effects all the subscribers being fed from both egress hubs, even though as reported by monitors in the hub tiers the availability of the hubs themselves is satisfactory. The example also shows that in Core B, Egress Hub_1 has a fault that has reduced Animal Planet program availability below the 99% threshold. The report shows that the fault is not in Core B itself, but is restricted to Egress Hub_1 and the subscribers it serves, as confirmed by monitors in Core B and in the Egress Hub_2 tier.

	Data Collection Point					
	Core_A	Core_A Egress Hub_1	Core_A Egress Hub_2	Core_B	Core_B Egress Hub_1	Core_B Egress Hub_2
	Measures	Measures	Measures	Measures	Measures	Measures
Program	Availabity	Availabity	Availabity	Availabity	Availabity	Availabity
(A&E (37))	99.442%	99.443%	99.443%	99.443%	99.269%	99.443%
(ABCFMLY (26))	99.453%	99.454%	99.454%	99.454%	99.280%	99.454%
(ACTNMAX (343))	99.443%	99.444%	99.444%	99.444%	99.444%	99.270%
(AMC (59))	99.443%	99.443%	99.444%	99.444%	99.444%	99.444%
(ANIMLPL (63))	98.967%	99.181%	99.181%	99.179%	98.925%	99.179%
(BIOGRPHY (243))	99.442%	99.443%	99.443%	99.443%	99.443%	99.269%

Figure 3: Availability Report Based on Data Gathered from Video Downtime Monitors

Conclusion

Leading service providers are moving to improve the reliability of the digital video networks in order to improve customer satisfaction and control the OPEX associated with responding to customer's trouble reports. The key to improving a video subscriber's QoE is to implement a real time, per-program video monitoring solution that has the flexibility to measure program downtime in accordance with service provider specified QoE metrics. These monitoring systems will allow service providers to achieve significantly higher levels of program availability, as well as allowing further reductions in OPEX due to more granular root cause analysis. Monitoring systems adopted by the service providers can also be leveraged by the NEMs to augment their existing test procedures to include video downtime measurements and video program availability calculations that are more meaningful to the providers of digital video services.