QUALITY OF EXPERIENCE IN CABLE NETWORKS: CHALLENGES, TRENDS, AND SOLUTIONS

Alon Bernstein
Cisco Systems

Abstract

Managing quality of experience (QoE) in a multiservice network is one of the most challenging aspects of network design; even more so, in point to multipoint (P2MP) cable networks.

This paper explains how QoE relates to quality of service (QoS), outlines what challenges cable operators see in managing QoE today, identifies the challenges waiting in the near term, and relays how to address them.

QUALITY OF EXPERIENCE VERSUS QUALITY OF SERVICE

A service provider typically sells a service level agreement (SLA) in terms of QoS. Typical parameters are:

- Peak rate – the maximal rate a service can reach in bits per second.
- Committed rate – the minimal rate guaranteed to the service. Even if the network gets congested the user will get at least "committed rate" bits per second.

Parameters that are internally set, but not typically communicated to the user as part of an SLA are:

- Burst size – how many bytes can be sent at line rate. Note that for the duration of the burst, the peak rate is violated, but over a longer period of time, it's maintained.
- Priority – what level of preferential treatment the service gets over other services.

The SLAs above could provide an easy way to compare Internet Service Provider (ISP) offerings: The more megabits per second, the more desirable the service would be. But QoS is only part of the whole user experience.

The data rate of a service is a not a good enough metric of the "experience" for the user. For example, loading a 500 K web page at 100 Mbps (which takes about 40ms) and loading the same page at 1 Gbps (about 4ms) provides the same experience for a user (Note that a full rate video at 30 frames per second switches a image every 33ms – no point in loading pages faster than that).

Quality of experience is, in many cases, a subjective measure which makes it hard to compare one service offering to another. However, several guidelines can be defined per application. For example, voice has Perceptual Speech Quality Measurement (PSQM) and Mean Opinion Score (MOS) as tools that are used to measure quality of experience in an automated way instead of polling a group of listeners each time a system test is run.

Similar attempts have been made to quantify the experience with video as well; for example, Video Quality Metric (VQM) and the "voice quality group" effort.

Quality of experience may also impact quality of service. As an example, a higher codec compression rate might require lower QoS, but may also result in lowered user experience because of the degraded image quality.

Another aspect in which QoE is different than QoS is that it includes all aspects of providing a service. While QoS focuses on the data plane (how many bits per second a service can provide), QoE captures the user experience in its totality, including functions that are associated with the control plane such as the time it takes to change a channel when channel surfing, or the time it takes to establish a voice call.

For control plane QoS, there are "magic numbers" that define the service as acceptable. For example, a channel change time of less than half a second is considered sufficient.

CURRENT BANDWIDTH AND RESOURCE MANAGEMENT CHALLENGES

The following is a list that tracks current issues with bandwidth and resource management in cable networks:

(1) Oversubscription and Modeling: Cable operators count on oversubscription in order to reduce expenses. In other words, the number of users per downstream or upstream is such that if everyone became active at once, their QoE would be unacceptable. A typical number of subscribers per downstream today is 800. With a downstream rate of 38 Mbps, that would be 47.5 Kbps per user – less than an analog modem! However, there is a statistical assumption that not everyone is going to be active at once. Unfortunately, one size does not fit all: traffic patterns are unpredictable. They depend on such things as demographics, penetration levels and other factors that are not within the cable operator's control. All the above makes modeling and capacity planning a challenge.

(2) Admission control of video flows: Video flows can be long lived – 90 minutes or more. If service is blocked because all the bandwidth dedicated for video is used, then this blockage might persist for a while. This means that oversubscription estimates for video need to be more conservative than for voice since the admission patterns are less dynamic.

(3) End-to-End QoS: QoS is typically guaranteed only within the cable operator's network. No uniform end-to-end QoS infrastructure has been standardized. An end-to-end QoS is needed if MSO A routes a phone call (or other service) directly to MSO B without going through the PSTN first.

(4) Symmetry: Voice, file sharing and video conferencing, drive bandwidth allocation to symmetry. The way the cable plant is wired, however, is inherently asymmetrical with much more downstream than upstream bandwidth.

(5) Shortage of bandwidth: There is not enough bandwidth in general and, in particular, for the next killer application—video. Certain applications, such as peer-to-peer (P2P) file sharing constantly drive bandwidth utilization up.

(6) Statistical multiplexing: The relative size of an upstream (US) or downstream (DS) channel is small compared to the bandwidth consumed by the services offered; for example, 6 Mbps service on a 38 Mbps downstream, reducing statistical multiplexing gain.

(7) Simple billing: Service is typically a flat rate best effort. Best effort might not be the optimal scheduling mode for all applications.

(8) Viruses and DoS (denial of service) attacks are a constant risk since they can reduce usable network bandwidth. QoE can be impacted by viruses and DoS attacks that specifically target the control plane. A DoS attack that floods the network with bogus IGMP JOINs can slow channel change times.

(9) Common language for creating QoS enabled flows: Many applications that may benefit from a QoS service flow do not have the proper hooks in them to trigger a service flow creation automatically. For example, most gaming consoles do not have a standardized way of requesting "delay sensitive" service for their flows.

(10) Benchmarking: Users have tools such as DSL reports, to grade an MSO service. These are primarily bandwidth driven. These are QoS tools and do not reflect the QoE that an MSO has to offer.

(11) Stressing the control plane: QoS is a data plane issue. However, highly dynamic services, such as video channel change or large number of voice calls, strain the control planes as well and impact QoE.

(12) Stressing the data plane QoS: With the DOCSIS® per-flow queuing, a very large number of queues have to be managed for every serving group. This stretches the limits of proprietary and off-the-shelf network processors to their limit.

(13) Commercial services: Cable networks have traditionally targeted consumers. There is a strong drive for providing commercial services over the cable network, including T1/E1 emulation, which is not a natural fit for a packet network that operates in an RF environment.

TRENDS IN QOS AND QOE

The following list is an attempt to predict where the QoE/QoS trends are:

(1) Increased number of services: With PCMM, the range of applications supported by the cable plant will increase, and so will the unpredictability. Tiered services are likely to become more common.

(2) The networked home: Each household is going to have more and more end users and PCs. Each one may be using a different set of applications (for example: data, gaming, video) which means tiered services in the home.

(3) The DOCSIS upstream: Some services on the DOCSIS upstream (UGS/UGS-AD/RTPS) have specific jitter requirements. When the number of different services increases—each with its own

periodicity and jitter requirements—the algorithmic complexity of scheduling these services, and performing admission control on all these varied services, becomes unmanageable. It is mathematically known to be an "NP-complete" problem, meaning that its impossible to find an optimal solution in a finite time).

(4) Increased bandwidth per-user: Because of competition with DSL, and the drive for bandwidth hungry Video On Demand (VoD) applications, the bandwidth demand per-user is increasing. A single high-definition (HD) stream is at about 8 Mbps. Since a cable downstream is about 38 Mbps at the most, there is not much room to grow. Some service providers are already singing up subscribers for 20 Mbps and above. Clearly at these rates, there can not be much statistical multiplexing gains.

(5) Delay sensitive applications: As more delay-sensitive applications, such as voice, video conferencing and gaming, are widely deployed, the need for end-to-end QoS will increase. The need to properly prioritize and schedule all these different services that are all delay-sensitive will also increase.

(6) Billing: As more services are added, a flat rate best effort service would become only a baseline service.

(7) Positive trends in DoS: There seems to be a decline in successful DoS attacks and viruses—most likely as a result of users becoming smarter about protecting their machines, and Microsoft actively working to reduce the number of vulnerabilities that Windows OS has. Router companies, as well, have come up with automated systems to detect and defuse DoS attacked. DoS attacks, however, are still QoS/QoE risks.

(8) Higher compression: While higher compression reduces the need for bandwidth, it also increases the sensitivity to packet drops because each bit of information becomes critical as compression ratio goes up. This is not an issue for "over the top" services which are likely to use Transmission Control Protocol (TCP), but will be an issue for a more optimized real time protocol (RTP) which currently does not have re-transmission capabilities (an RFC draft only).

(9) Multicast: Multicast greatly improves bandwidth utilization when the same content is viewed by a large number of users at the same time. Multicast services are going to be more common all the way to the Customer Premise Equipment (CPE).

(10) "Over the top" services: Content and services can be provided by outside companies. The same way that voice is delivered by Sktype and Vonage, video distribution would follow the same path. Providers external to the MSO network will provide video services. For these services, end-to-end QoS may also become critical.

(11) P2P file sharing: P2P file sharing networks can be viewed as systems for delivering video over the Internet, since video is (in terms of traffic volume) what is driving the

P2P usage. In other words, even if a service provider manages to clamp down on P2P traffic by some means, it does not mean that the bandwidth demands would decrease– at least on the downstream direction. They will be replaced by legitimate bandwidth hogs in the form of VoD streams.

(12) More subscribers, more bandwidth: For each household passed, each subscriber is going to have more bandwidth. If one assumes that penetration rates are becoming saturated that would mean that the number of users will remain fixed. But this is not the case. DOCSIS Set-top boxes (STBs), as well as other DOCSIS-enabled devices (power meters for example), will increase the number of devices that the network has to support, and with it the constraints on QoS.

SOLUTIONS

Before outlining a set of solutions, a basic question has to be answered: Why does QoE need to be managed ? Can't all QoE problems be solved by providing enough resources to the network? After all, QoS is not an issue if there is enough bandwidth in the network and QoE should not be as well.

There are several answers to this question. First of all, a network can be built more economically if it is not built on worst case assumptions regarding bandwidth utilization. For example, we can assume that each household needs 50 Mbps to allow for 3 HD streams and Internet services. For 500 households passed, that would be 25 Gbps – exceeding the capacity of a single fiber node – and more than what is needed in practice.

Even for a network that is over-provisioned, having QoE/QoS enforcement is needed for mission critical services (such as 911 calls) so that even if the statistical model fails on extreme circumstances, the critical services are not impacted.

Because of the shared media nature of cable networks, a service might degrade as more users are added to the same serving group. In the past, customers have preferred a QoE where bandwidth availably is restricted, but constant, over one where bandwidth availability fluctuates between high and low. Actively managing bandwidth achieves the target of having predicable and stable bandwidth.

The concept of "net neutrality" is hotly debated these days. It's a question on what types of preferences an MSO can give its own traffic versus externally sourced traffic. There seems to be a general agreement that it is not okay to intentionally degrade external traffic, and that it is acceptable to provide a higher level of service to internally sourced traffic.

However, even the latter is debated since an MSO can end up leaving very little bandwidth for external services – enough for web browsing, but no more than that. It's still not clear what type of regulation, if any, would be enforced in these cases. It's likely that QoE tools will be needed to manage it, if indeed it is enforced.

The following list outlines solutions to the challenges and trends presented in earlier sections. Note that not all items can be addressed. This list covers the ones that can:

(1) DOCSIS 3.0 downstream channel bonding: CableLabs® is close to releasing a first version of the

DOCSIS 3.0 specifications. DOCSIS 3.0 increases the bandwidth of a DOCSIS channel by means of channel bonding. This means it does not increase the physical capacity of the channel. Instead, it uses a Multi Link PPP (MLPPP) like technique to spread packet across independent L2 links. The overall effect from an L3 perspective is a faster link.

DOCSIS 3.0 comes with its own set of QoE issues. Because packets are sequenced and sent on independent links, they might be delayed until all of them are received in sequence. This issue is exacerbated when a packet is lost. In certain cases, the re-sequencing engine will have to wait a full re-sequencing window (up to 18 ms) before resuming operations.

Another complexity in DOCSIS 3.0 comes from the fact that not all devices and not all services use the same number of channels. This creates a system of multiple, partially overlapping, groups on which a packet can be stripped. Scheduling and load balancing across these groups is a complex task. However, if executed correctly, it will provide both higher bandwidth per modem, as well as statistical multiplexing gains.

Bonding provides a new set of knobs that can be fine tuned to provide different QoE levels: the number of channel a flow is sent across, the timeout for re-sequencing, including/omitting sequence numbers, etc.

DOCSIS 3.0 upstream channel bonding will also address some of the issues with symmetry. Although the physical bandwidth of the plant is still heavily skewed towards downstream bandwidth, upstream bonding gives the ability of sending up to a 100 Mbps on the upstream bandwidth, so selected cable modems can have symmetry.

(2) DOCSIS 3.0 multicast: Another improvement in DOCSIS 3.0 is enhanced multicast support. This will further enhance the bandwidth savings that multicast enables. It’s important to note that even though the trend is to have a stream per subscriber (VoD), multicast still has a role to play for real-time viewing events, and for off-line downloading of popular content to PVRs.

(3) DOCSIS 3.0 commercial services: DOCSIS 3.0 also addresses the issue of transporting T1/E1 services over DOCSIS, which require specific clocking to maintain the appropriate QoE.

(4) Applying PCMM to non-PCMM applications: The same deep-packet-inspection tools that are typically used to filter unwanted traffic can be used to detect certain types of flows: for example, gaming, and request a specific QoE level on their behalf so that even non-PCMM enabled applications can have dedicated QoE services.

(5) Dynamic bandwidth management: Since QoE is not about bits per second, it’s possible to tailor an SLA that fits a user profile instead of selling a limited set of SLAs. A user

that is doing only web browsing does not need a high SLA. One option an MSO has is to sell packages such as "web browsing" and "video" instead of selling a number of mega bytes per second. Another option is to sense what applications a user is running and dynamically assign a profile that fits the users traffic patterns. This will help in modeling and optimizing bandwidth utilization.

(6) End-to-end bandwidth reservations: The issue of end-to-end bandwidth reservation is beginning to be addressed in various forums, but no standard is emerging. There are many issues with customer ownership, responsibly in case of failures and communication of the service levels (not all ISPs mark "high priority" the same way) that are not fully resolved yet. Of special note is the IP sphere effort, which tries to dynamically set business agreements, on a call-by-call basis, between service providers.

(7) Improved modeling: Several tools for advanced modeling are becoming available from CableLabs® and other sources. In addition to that, a good approach to planning network capacity is to start an iterative process: make some assumptions about network utilization and the number of subscribers that can use the network. Raise the right flags as the system begins to cross certain thresholds (for example, aggregate data rate, or CPU utilization or control plane load) so that the MSO has enough time to properly update the network before a QoE disruption occurs.

(8) Flexible jitter bounds for the upstream: A way to address the upstream scheduling problem is to provide a statistical guarantee on jitter/delay, rather than an absolute guarantee. This approach would make the upstream look more like a packet transport (same as the downstream) where the jitter/delay is a function of utilization levels and not guaranteed in any way. Using admission control tools, it is possible to keep the jitter/delay at a certain statistical guarantee.

CONCLUSION

The vision of a converged video/voice/ data network has been around for years. We have reached a point where this vision is becoming a reality. The competition between DSL and cable has driven cable to provide voice services, and DSL to providing video services. While cable operators still manage video as a separate network, they will soon provide video services over their data pipes—either from externally sourced servers, or as an upgrade to their current video offerings (IP video over DOCSIS can provide better bandwidth usage, enhanced interactivity, better network connectivity, etc). Other services will be added: gaming, backup services, video conferencing and others. Managing the QoE for these services will be critical for future successful deployments.