Carrier-Grade Ethernet Challenges for IPTV Deployment

Sundar Vedantham, Seong-Hwan Kim, and Deepak Kataria, Agere Systems Inc.

ABSTRACT

Carrier-grade Ethernet standardization and deployment is gaining momentum due to the ease of deployment, lower cost, and compatibility with existing networks on the access end. When Internet Protocol Television (IPTV) is deployed using Ethernet as the underlying interconnect fabric infrastructure, meeting fine-grained traffic management (TM) requirements on the service provider side to meet quality of service (QoS), billing, and security features implementation on the user side poses several challenges. Such challenges could be met using the TM features built into network processors (NPs).

INTRODUCTION

Present Ethernet deployments on the LAN side are mature, very well understood, fairly inexpensive, and serve the requirements well. In the last decade, Ethernet technology has evolved from the 10 Mb/s shared wire model to switched operation over unshielded twisted pair (UTP) that support 1000 Mb/s and then on to fiber optic transmission from 100 Mb/s to 10 Gb/s rates with transmission distances spanning from 2 to 2000 km using long-haul dense wavelength-division multiplexing (DWDM) systems [1]. Newer Ethernet versions also support up to eight classes of service and unicast/multicast/broadcast modes via the VLAN technique. This evolution makes it a good candidate for both LAN and WAN interconnection space. Recent projections estimate that close to US$29 billion will be spent worldwide on Ethernet in metropolitan networks between 2004 and 2008 [2]. But Ethernet technology lacks carrier-grade features such as QoS, provisioning, fault tolerance, TDM compatibility, OAM, and self-healing, making it unsuitable for backbone and carrier space. Recently there has been a lot of standards activity at the Metro Ethernet Forum (MEF), the Internet Engineering Task Force (IETF), and the International Telecommunication Union (ITU) with regard to addressing these requirements, thus paving the way for wider deployment of Ethernet as a carrier-class interconnection fabric [3–5]. Early deployments are on in the Asia-Pacific region.

Internet Protocol television (IPTV) is being touted as the next "killer application" that will consume available bandwidth on the Internet and enable future network growth [6]. While the hype has been quite high, real-world use remains minimal and is only in its early stages. If the deployment is to reach its full potential, we need to look at end-to-end scenarios and ensure that there are no stumbling blocks. This article analyzes IPTV requirements and how they could be met when delivered over carrier grade Ethernet.

To put this subject in context, some basic background is instructive. Internet growth is forcing the technology to migrate from simple data services to triple-play services. A standard definition of triple-play services includes voice, data, and video services provided over a single transport infrastructure. In addition, services like audio/video conferencing, collaborative editing, video on demand (VoD), PPV (pay-per-view movies and content stored in the network and streamed on demand), Internet telephony, Internet radio, premium interactive content, and interactive gaming may form an emerging set of services supported. The delivery of such services opens new revenue streams for telecom service providers [7]. Customers will be willing to pay for such a rich suite of services, especially when multiple services are bundled and made affordable.

This migration requires changes in the physical infrastructure that bring service to homes from current cable or satellite TV models. Analog cable TV will not be able to piggyback digital triple-play services on the same network due to bandwidth and interoperability limitations. Satellite TV, being unidirectional, does not lend itself well to triple-play services that require bidirectional communication. Present satellite TV deployments use regular phone lines to receive information from subscribers even for pay per view orders. This dependence on regular phone lines and the technology's inability to provide any direct means of upstream communication (from subscriber to network) makes it an impractical contender in the triple-play arena. As of now, technologies like fiber to the premises (FTTP), xDSL, digital cable, and high-speed wireless seems better positioned to handle multimedia-based rich bidirectional communication

0163-6804/06/$20.00 © 2006 IEEE

services. Noticing the trend, MEF has identified residential triple play with IPTV as one of the carrier-grade Ethernet drivers.

Best-effort service has been traditionally used to provide Internet access services such as Web browsing. Ethernet works well in this context at the subscriber end. But the new set of multimedia services require end-to-end quality-of-service (QoS) guarantees. Since customers are used to viewing television programs and using their telephones without noticing any jitter or delay, QoS guarantees are a must in triple-play service deployment over Ethernet. This becomes critical because, as the available bandwidth per customer increases, emerging suite of services will demand even more bandwidth, generating bottlenecks that can only be handled well via traffic management (TM) features. Thus, for the ubiquitous Ethernet idea to succeed from the end-user to the carrier levels, carrier-grade Ethernet will require good TM capabilities on the network that can be provided by network processors (NPs).

Broadband Service Requirements

It is clear that each type of service will come with a slightly different set of requirements. Let us consider them case by case as follows:

- VoIP: Minimum bandwidth but strict limits on delay and jitter
- Video conferencing: Higher bandwidth, very strict limits on delay and jitter
- Live video broadcast: Higher bandwidth and limits on delay and jitter; limited packet loss is acceptable
- Video on demand (VoD): Higher-bandwidth statistical guarantees; less stringent delay and jitter requirements, as video streams can be buffered
- Interactive gaming: Low bandwidth but strict bounds on delay
- High-speed data services: Just bandwidth guarantees and less stringent delay requirements; packet-loss limits
- Web Browsing: Medium bandwidth and delay, high reliability, best-effort service
- Email: Low bandwidth and delay, high reliability

Besides QoS, support for additional control functions is needed for various on-demand services. Quick channel-changing ability is one such example. The challenge it presents and possible solutions are discussed in the next section.

Billing and security features will require fine-grained metering of traffic. In addition to meeting these requirements, service providers need to ensure that the available bandwidth is being used to full extent so as to generate maximum revenue. Thus, the deployment will rely heavily on the traffic management functions available on network processor chips.

IPTV Requirements

Customers will be attracted to bundled triple-play services when the bundle costs less than the sum of the service fees for individual services. But to be successful, the bundled services should provide service quality that is at least equal to, if not better than, the service customers are used to. Customers will not migrate to bundled triple-play services if the television picture quality tends to be more grainy with intermittent jitters and the channel switch controls are less responsive than what they are presently used to. Hence, key IPTV features needed for successful deployment include [8]:

- Selection: End users should be able to select their program of choice from a plethora of TV content available. This also requires faster channel selection and channel changing time.
- Storage: TV content should be storable in a local storage device, so that it can be made available to customers whenever they want to watch it (time shifting). Vendors are planning to store about 100 hours of programming (new TV programs or movies) into the set-top boxes (STBs).
- QoS: The service should guarantee bandwidth for video streaming and QoS differentiation for supported classes of traffic.
- Low cost: Per-line/per-customer cost of bundled services must be low.
- Upgrades: Service provider should be able to upgrade codecs and user authentication software without disrupting service.
- Miscellaneous: Service providers should be able to provide high-quality resolution for programs at no substantial additional charge to end users.

Future IPTV service will provide two different types of TV services: standard definition TV (SDTV) and high-definition TV (HDTV). SDTV bandwidth ranges from 1 to 4 Mb/s. HDTV bandwidth ranges from 4 to 13 Mb/s. The typical number of TV channels available from a provider hovers between 250 and 300 channels of SD, and an additional 10 to 20 more channels for HDTV. If each home has approximately four TVs, two to three SDTVs and one to two HDTVs can be supported with about 20 Mb/s bandwidth. At this point, bandwidth management among different traffic classes to homes becomes a critical issue, meaning voice, video, and data services must be handled differently. The following sections discuss how TM features available in network processors address these listed requirements.

Traffic Management and QoS Control

From a QoS control perspective, traffic management in Ethernet-based IPTV networks can be particularly challenging. This is because traffic management solutions have to be implemented at different levels of control granularity. Those levels for xDSL service include:

- Individual services actively used by a given subscriber
- Individual xDSL link-load for the given subscriber
- Aggregate subscribers supported on a given line card
- Aggregate line cards supported on a given uplink card

If each home has approximately four TVs, two to three SDTVs and one to two HDTVs can be supported with about 20 Mb/s bandwidth. At this point, bandwidth management among different traffic classes to homes becomes a critical issue, meaning voice, video, and data services must be handled differently.

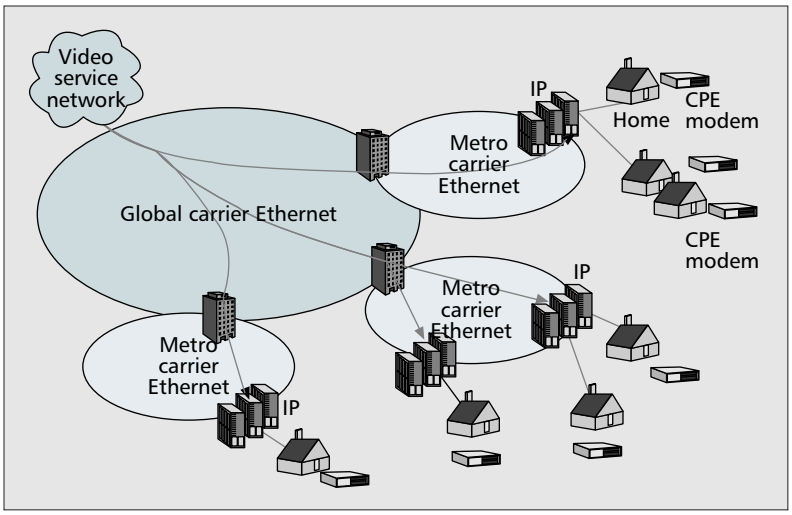

Figure 1. *Carrier Ethernet-based IPTV network.*

There may be other intervening levels of control granularity that have to do with the virtual partitioning of the links at the various levels of the hierarchy for better QoS controls. For example, macro-level controls will be needed to differentiate residential customers from business customers. Figure 1 shows a network topology designed to provide IPTV services based on Ethernet on xDSL networks. Services deployed via wireless or digital cable technology requires very similar service support levels.

Figure 2 shows a Multiservice Access Network (MSAN) architecture that can deliver IPTV services using carrier Ethernet devices. The xDSL links from the subscribers terminate in digital subscriber-line access multiplexers (DSLAMs), which is part of the broadband access network. The aggregating system is known as the subtending system; each of the systems connected to it are known as subtended systems.

This many-to-one relationship at different levels is characteristic of DSLAM architectures. This one-way direction of aggregation from the subscribers toward the provider is also known as the upstream direction. The opposite direction from the provider to the subscribers is known as the downstream direction.

In the downstream direction, hierarchical traffic management can provide differentiated services for different subscribers, because each user gets different types of services allotted by different schedulers. In addition, class-dependent traffic isolation can be provided to the end user between different traffic streams. Per-user scheduler configuration can be mirrored to other users who request the same bandwidth and set of services. Hierarchical traffic management, as shown in Fig. 3, would be ideal to handle such a scenario.

In the upstream direction, individual user traffic gets monitored at the lowest level of hierarchical scheduler. And each class of traffic from different users can merge into a separate class scheduler (e.g., voice, data, and video) at the next level of hierarchical scheduler.

Having a separate scheduler for each class will provide class isolation. Another benefit is that bandwidth starvation caused by lower/other class congestion will not be an issue. Having per-user scheduling can also provide end-user isolation for billing and bandwidth control purposes. Without the presence of a network processor, scheduling at this level of granularity using carrier-grade Ethernet gear alone is not possible.

Channel Selection

Contemporary TV technology sends all TV channels to all end users and channel switching is performed by simple filtering of the undesired frequencies. This mechanism is used in existing cable TVs. Specifically, channel changing simply selects one channel functioning like a band pass filter and ignores the rest. IPTV cannot use this approach because of the limited bandwidth reaching each home. Therefore, the telecom service provider only sends the selected TV channel or channels to the customer premises. In case of Fig. 2, this model means that multicast trees should first be established with QoS guarantees for video transmission. When the subscriber sends in a request for a video channel using the STB connected to the CPE, carrier Ethernet filters the request and replicates and forwards only the requested channel to the subscriber.

This creates channel changing delay, since the channel change information has to travel upstream through the network to the service provider. But, if all content/channels are brought to the DSLAM box using high-capacity backplane, the distance IGMP messages need to travel to effect channel switching will be greatly reduced, thus providing dramatic reductions in channel changing delay since IGMP messages need not travel all the way up to a broadband remote access server (BRAS).

The IPTV delivery architecture may also want to stream adjacent TV channels to the one that is being watched to the STB so that moving one channel up/down at a time will be instantaneous. But this approach triples the bandwidth required for each TV set.

Bandwidth Utilization

When the TV sets in a household are all off, customers paying for 20 Mb/s bandwidth would want to get the entire bandwidth diverted to other applications that are currently running, such as browsers or games. While this bandwidth adjustment does not have to be guaranteed (for oversubscription), the service provider should still guarantee the minimum bandwidth for the services that are running (such as browsers and games). To achieve this, the service provider system should monitor the status of TVs at home. If one or more TVs are turned off, the service provider may allocate more bandwidth to the end user's data service with best-effort scheduling technology. When TVs are turned on, the service provider needs to reduce the scheduler rate of other services. Thus, the controller must have a powerful scheduler to dynamically control bandwidth in this fashion.

As stated above, the bundled services should target the right inflection point from cost perspective. That means a reasonably priced network processor would be required to provide proper QoS management, as well as all the benefits of traffic management technology. To be

able to exploit the synergies between telecommunication service providers and cable TV content providers, the NPs handling the traffic flows should be able to balance the demands of the services, while maximizing the bandwidth usage on the network.

Traffic Management for Affordable Broadband

The overall objective of traffic management is to support a wide variety of services with diverse QoS requirements while ensuring efficient sharing of network resources. The former allows providers to offer new revenue-generating services; the latter promotes service affordability. Traffic management includes functions such as policing, buffer management, scheduling, shaping, backpressure flow control, CAC admission control, route selection, and node/network monitoring. These functions must support full bandwidth flexibility such that upstream and downstream rates can be chosen freely and continuously up to the maximum physical limits. The functions also should enable full-service flexibility such that a random mix of services with various bit rates and various traffic requirements can be supported, within available bit-rate limits. The functions should foster maximum sharing of network resources and help minimize network downtime [9].

These functions can be implemented in a centralized manner where all the control functionality resides in an uplink card that handles traffic from all the connected line cards, as well as all the subscriber traffic from each line card. In the distributed approach, the control functions are distributed between the uplink card and the line card. The distributed approach provides for a more scalable and flexible architecture.

Despite the aggregation in the upstream direction, the subscribers' traffic needs to remain segregated. This can be achieved by providing an equivalent mirrored structure of controls in the downstream direction.

In the upstream direction, policing ensures traffic conforming to the service level agreement (SLA) is allowed into the network. Marking of traffic may be employed to signal the relative treatment of traffic in the network. Differentiated traffic scheduling is needed to treat the delay-sensitive real-time services (telephony or traffic with real-time video content), preferentially over the data services.

Class-specific backpressure flow control is needed to handle the periods of temporal overload, when the incident traffic load exceeds the available bandwidth in the upstream direction. Due to the fan-out topology, no traffic aggregation exists in the downstream direction. Rate shaping is employed so that downstream traffic flows smoothly with no likelihood of buffer overflows. Temporary rate excesses are accommodated by limited downstream buffering. Traffic multicasting is needed at the edge node such that the WAN links upstream can be efficiently used. Priority-based traffic handling is needed so low latency can be provided for high-quality telephony, audio and video conferencing and other types of video services with real-time content.

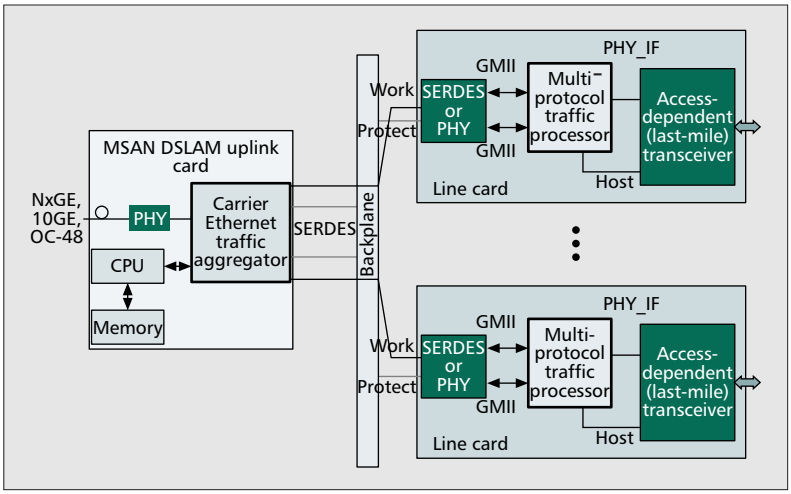

Figure 2. *MSAN DSLAM architecture.*

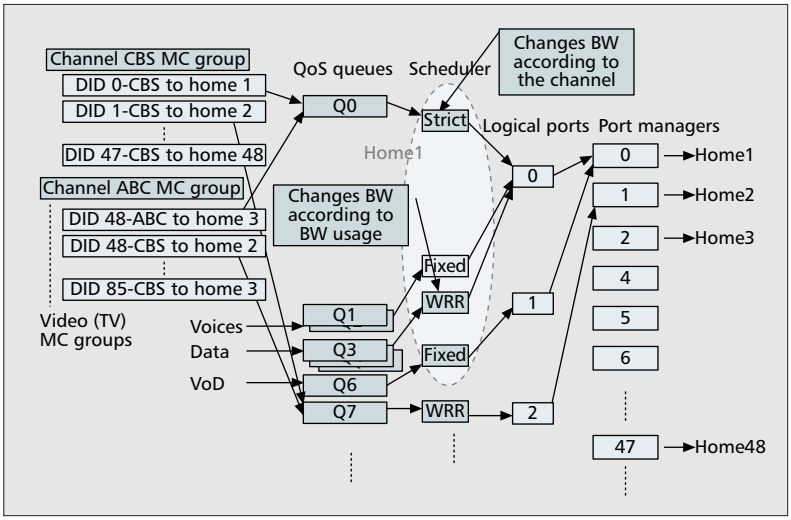

Figure 3. *Downstream traffic management in a line card.*

The call admission control (CAC) function keeps track of upstream and downstream bandwidth along all the nodes and links in the path. Any new connection requesting a certain QoS level is only admitted if the resulting bandwidth use is within a predefined admissible region. It is possible for the flow control flag from the upstream node to be turned off for class 2 traffic, while the flow control flag for class 1 is turned on.

For downstream traffic, the CAC function needs only ensure that aggregate bandwidth use for guaranteed traffic remains within the available capacity. For upstream traffic, the admissible region is the set of bandwidth use values for which all guaranteed flows receive their contracted QoS. Network/node monitoring functions allow the detection of resource problems proactively such that remedial actions can be taken for maximum availability. For example, switch-level redundancy at the line card, and switch fabric, common control, power supply levels, and network-level redundancy (such as automatic routing, failure recovery, and congestion con-

For cell-based traffic, well-known GCRAs are used to check for conformance to limits on how fast the cells may arrive and limits on the number of cells that may arrive back-to-back during a specified interval. Actions on detection of nonconformance include dropping or tagging cells.

trols) are important functions that help minimize network downtime. While carrier-grade Ethernet standards efforts are on to address these requirements, it is clear that traffic management plays a critical role to support new revenue generating services for service providers.

Efficient traffic management also plays a major role in maximizing use of network resources. Traffic management helps increase provider revenue according to the following list of factors:

- Using the minimum amount of buffer space for the maximum good throughput
- Satisfying the QoS requirements
- Using the least amount of bandwidth
- Supporting the maximum number of connections over a given network link
- Minimizing network downtime

This in turn makes broadband services more affordable and facilitates logarithmic increase in cost of service with bandwidth demand. This occurs by deriving maximum possible use from existing and newly deployed network and system resources. In the absence of this, existing and new capital investments are underused, leading to the cost of service growing linearly with bandwidth demand.

Efficient traffic management also offers several hooks for increasing provider revenue. By dynamically adjusting the scheduling weights for both improving QoS for higher-priority users — those that suffer from a high population of users with lower priority — revenue earned by the service provider can be increased. However, this strategy should not completely block the users who may have subscribed for cheaper service.

Users can also be charged additional fees depending on the throughput improvement, for example, when real-time services are not being used. Traffic management functions in this case must detect the change and provide the new rate to the willing subscriber.

Thus, efficient traffic management plays a major role in delivering QoS and increasing provider revenue, both of which are key elements in making affordable Ethernet based personal broadband a reality.

Management Factors

In this section we look at four significant TM features available in NPs that affect IPTV implementation in carrier-grade Ethernet.

Policing

Policing is the function of monitoring traffic generated by the user. Its purpose is to ensure traffic conformance to the contracted temporal behavior and take appropriate action upon detection of nonconformance.

For cell-based traffic, well-known generic cell rate algorithms (GCRAs) are used to check for conformance to limits on how fast the cells may arrive and limits on the number of cells that may arrive back-to-back during a specified interval. Actions on detection of nonconformance include dropping or tagging cells. During the upstream queuing and buffer management functions, tagged cells are treated with lower priority compared to untagged cells [10].

For frame-based traffic, the service eligibility test, or frame-based generic cell rate algorithm (F-GCRA), is defined in addition to conformance tests. A frame is deemed conforming if all its cells are conforming and nonconforming if one or more of its cells are nonconforming (GCRA test).

A frame is eligible for service guarantees if it is conforming and passes the F-GCRA test for the contracted minimum rate and the associated tolerance. This tolerance limits the number of frames that can arrive back-to-back during a specified interval. These notions of conformance and service eligibility make possible the specification of guarantees of service at the frame-level for the variable-sized frames when using the underlying Ethernet infrastructure for transport. For frame-based traffic, ineligibility is handled by dropping complete frames or tagging all the cells of the frame to indicate lower priority during the upstream queuing and buffer management functions.

The ultimate objective of the policing function is to protect network resources from malicious or unintentional source misbehavior that can affect QoS commitments to other users. This policing is done by detecting violations of contracted parameters and taking appropriate discarding or tagging actions.

Buffer Management

Buffer management follows the policing function and has the dual objectives of satisfying the heterogeneous QoS requirements of connections or flows. This is done by first providing the necessary protection and isolation, and then simultaneously extracting multiplexing gains from sharing the different buffer allocations across the hierarchy of classes/port, and across all ports.

Without distinct buffer allocations, it is not possible for the scheduling function to differentiate between the various classes of traffic competing for a single output port. Even within a single class, there may be QoS requirements of different connections or different flows, requiring separate queues or buffers. Providing a hierarchy of buffer allocations with static thresholds achieves the needed differentiation, but does not allow efficient sharing across the connections or flows within a class, between different classes per port, or even between ports. As indicated above, buffer management should also promote maximal buffer use, because it has a direct effect on the cost of the service provided.

Dynamic thresholding achieves the dual objectives of QoS differentiation while maximizing buffer use. The thresholds are dynamic because they change, depending on the amount of free buffer space available. The larger the free buffer space, the higher the threshold. These thresholds are usually provided at all hierarchical levels: per connection/flow, per traffic class per port, and per port. The dynamic threshold function permits allocation of buffers to individually overloaded connections when there are large reserves of unoccupied buffers. This buffer

allocation avoids losses that would be incurred by the overloaded connections under a static threshold policy.

Conversely, as the overall traffic-class occupancy approaches a per-traffic-class, per-port threshold, the extra allocation is withdrawn to prevent any one traffic class from occupying a disproportionate share of buffer space at a given port. For example, non-real-time services can use large buffers. By contrast, for real-time services, delay variation tolerance, and transfer delay, constraints dictate smaller guaranteed buffer requirements. Likewise, buffers across multiple ports can be shared up to a guaranteed minimum per port. Thus, an intelligent buffer management function plays a major role in providing needed service differentiation while maximizing buffer use.

Scheduling

The scheduling mechanism selects a queue to service at each cell/packet dispatch time such as to meet the associated QoS requirements. Such requirements are usually specified in terms of acceptable statistical bounds on maximum delay, delay variation, loss rate, or some combination of these, depending on service type.

While satisfying the QoS requirements, the scheduling mechanism must deliver the bandwidth committed as part of the traffic contract. The scheduling mechanism must ensure sufficient isolation between connections such that the traffic characteristics and QoS requirements of one connection do not adversely impact the bandwidth and QoS provided to another connection. It is also expected that the scheduler will provide fair share to connections of excess link bandwidth whenever it is available.

The scheduling mechanism should promote high use of both bandwidth and buffers. It should be scalable with respect to the number of connections and over a wide range of link speeds. While isolating connections requiring guarantees on bandwidth and QoS, the scheduling mechanism must allow maximal sharing among connections requiring no bandwidth guarantee or QoS.

It is also desirable that scheduling functions in NPs work in conjunction with buffer management functions to provide tunable parameters that can maximize coverage of operating points and facilitate design of the connection admission control (CAC), which operates at the connection set-up phase.

Providing QoS Differentiation

Similar to the buffer management function requirements, the need to provide QoS differentiation while enabling the best possible sharing of bandwidth resources and fairness requires a hierarchical scheduler active at the per-connection/flow level, per-service-class level, and per-port level to provide the needed bandwidth guarantees, delay bounds, and fair share of excess bandwidth. The scheduling function should also be augmented to deal with transient backpressure conditions, especially for real-time services like IPTV.

At the first level, each connection/flow is usually provided its own queue for scheduling. The time to serve the queue, called the starting time, is computed based on the system potential function (also called the virtual time). System potential function keeps track of the total work done by the scheduler and is thus a nondecreasing function of time. A queue becomes eligible when the starting time is less than or equal to the system potential function.

Timestamp values (finishing times) are assigned to queues based on the system potential function. It advances each cell/packet departure instant by an amount equal to the reciprocal of connection/flow service rate. At any time, the queue with the minimum eligible timestamp value is selected for service. Thus, eligibility is tested with respect to starting time, and scheduling is done based on finishing time. Because packets can be variable in size, packet lengths should also be considered for packet scheduling.

Work done by the scheduler should be accurately tracked (the actual byte count) in this case, because it affects the ability to deliver bandwidth and delay guarantees. The objective of such a scheduler type is to assure bounded delay independent of the number of active connections, while maintaining fairness of service among competing connections.

For connections/flows that do not require delay guarantees but need bandwidth guarantees, frame-based approaches can be used [11]. In this type of approach, weights are assigned to each connection/flow, and at any time the length of the frame is equal to the sum of all the weights of all the active connections/flows. The service for connection/flow is deemed complete in a given frame if the connection/flow is served a number of times equal to the assigned weight in the frame; such service is repeated in the next frame, and so on. Rather than serve each connection/flow consecutively based on its weight, service is interleaved among all the connections/flows in a round-robin fashion.

For variable-sized packets, packet lengths are considered in scheduling and the frame-based packet scheduler keeps track of the deficit accrued at each frame service. To meet the long-term average commitment of bandwidth guarantees, the scheduler applies it during the next frame service and so on.

The scheduling schemes at the first level are extended at the second level, where the scheduling is among the different service classes. At this level, bandwidth and delay guarantees are met at the aggregate service class level. Any leftover bandwidth from the guaranteed class level scheduler can be made available to different service classes in a fair manner. Similarly, scheduling at the third level can provide bandwidth and delay guarantees at the port level, while providing the capability to distribute excess bandwidth among ports. This hierarchical scheduling provides QoS differentiation and promotes efficient bandwidth use.

To summarize, Ethernet networks designed for delivery of triple-play services should be able to perform policing well in the upstream direction, support sophisticated buffer management schemes, and provide hierarchical scheduling with enough queues and schedulers to maintain differentiated QoS for a large number of traffic flows. These requirements can be supported by

Timestamp values (finishing times) are assigned to queues based on the system potential function. It advances each cell/packet departure instant by an amount equal to the reciprocal of connection/flow service rate. At any time, the queue with the minimum eligible timestamp value is selected for service.

NPs monitoring individual user traffic and blocking malicious IGMP messages can prevent overloading of the network resources. The same NP can also identify and block malicious traffic getting into the system using fast pattern matching abilities built into NPs.

designing in network processors on the access end of the network.

Security

Developing an architecture that adequately protects the network resources so that good end-user experience is maintained while malicious users/software are kept out is extremely important in converged networks. Since the services are bundled and delivered through one medium, attacks carried out on one service can end up bring down all the services simultaneously, thus making the system more vulnerable.

A denial of service (DoS) attack aims to prevent legitimate users from obtaining services from desired resources by typically flooding the network with unwanted traffic. This flooding overloads the provider of the service (such as Web servers), thereby preventing services from being delivered. DoS attacks could also be attempts to prevent individual systems from communicating with each other.

In a distributed DoS (DDoS) attack, the attack uses a "many-to-one" approach, typically by taking control of several zombie machines (i.e., systems taken over without the owner's knowledge to perpetuate such attacks), and coordinating the attack from several machines to the targeted service provider system at one synchronized time. In case of IPTV, infected STBs could be forced to send out endless IGMP join/leave messages upstream as if individual TV viewers are switching the channels continuously. A coordinated attack initiated from every STB in a system can overwhelm the provider network if the IGMP messages flood the network all the way up to the service provider's office. The solutions discussed above would address this security concern as well.

Solutions have been proposed to handle such attacks [12, 13]. Typical approaches involve separating legitimate traffic from the flood of DDoS traffic and continuing to provide the service to legitimate traffic while ignoring the pseudo-traffic trying to exhaust resources. One of the techniques used is enforcing a quota system when attackers use legitimate IP addresses to initiate the attack. This ensures no client system consumes an unacceptably high share of the available resource, thereby eliminating starvation (i.e., service denial) for any of the participating clients. Monitoring the extent to which a given user is consuming resources requires fine-grained traffic flow monitoring. Thus, NPs monitoring individual user traffic and blocking malicious IGMP messages can prevent overloading of the network resources. The same NP can also identify and block malicious traffic getting into the system using fast pattern matching abilities built into NPs [14].

Implementing algorithms that measure traffic flow anomalies [15] even though what is being looked for may not be well known ahead of time is another powerful technique in this realm. Fine-grained flow measurement capabilities available in network processors can be put to use to identify suspicious flow patterns so they can be isolated for investigation or blocked from entering the local network.

Reliability

Application service resiliency (ASR) is a means to ensure that end-user access to the network resources is never totally denied, even if cables and/or equipment providing the service become faulty. This end has always been achieved through redundancy in networks. In the standard redundant models, the entire traffic reaching a destination (such as a curbside DSLAM box) would be replicated by simply provisioning double the bandwidth required for the destination using a different physical medium. This results in half of the bandwidth going unused most of the time.

The ASR technique leverages the fine-grained traffic management features available in NPs to individually isolate and track Layer 4 services reaching the same home. If and when the main route providing the service goes down due to equipment/cable malfunction, a selected subset of Layer 4 services is immediately restored to the home via redundant lines, while the rest of the service restoration may be completed as soon as possible. Thus, to give an example, even though a single medium brings in triple-play services into a home, when there is a network failure, the service provider can choose to provide VoIP service alone via the redundant line immediately while letting other services restore on a normal healing schedule. This technique allows service providers to lease a smaller-capacity (say, just 20 percent of the original bandwidth) line and use it as a redundant line, resulting in cost savings. In order to implement this technique, since the Ethernet gear is not designed to work at a fine-grained level, the NPs used in the system should be capable of supporting multiple scheduling queues to serve a single customer, and be able to identify failures and switch the selected services to redundant lines instantaneously while maintaining FIFO order of packets being routed.

Conclusion

Efficient use of network resources while providing differentiated QoS for multiple service classes requires a comprehensive traffic management framework. The carrier-grade Ethernet effort underway does not fully address these requirements, since the focus is oriented more towards scalability, protection, TDM support, minimum QoS control, and OAM issues. An integrated traffic management solution implemented using network processors that employs appropriate controls on sophisticated policing, buffer management, and hierarchical scheduling for differentiated services can provide a hierarchy of economies and revenue benefits valuable to service providers. This maximizes service affordability and fills an important void in making the ubiquitous Ethernet model successful. Thus, leveraging traffic management is key for successful deployment of triple-play services, including IPTV, in the near future.

References

[1] White Paper, "Understanding Intelligent Carrier Ethernet: Bringing the Advantages of Ethernet to the Service Provider," http://www.cisco.com/warp/public/cc/techno/lnty/etty/ggetty/prodlit/intgn_wp.pdf, 2003.

[2] Infonetics Report, "Carrier Ethernet Booming," http://www.lightreading.com/document.asp?doc_id=71770, Apr. 12, 2005.
[3] Metro Ethernet Forum, "MEF 2: Requirements and Framework for Ethernet Service Protection in Metro Ethernet Networks," http://www.metroethernetforum.org/PDFs/Standards/MEF2.pdf
[4] Metro Ethernet Forum, "MEF 4: Metro Ethernet Network Architecture Framework — Part 1: Generic Framework," http://www.metroethernetforum.org/PDFs/Standards/MEF4.pdf
[5] Metro Ethernet Forum, "MEF 12: Metro Ethernet Network Architecture Framework Part 2: Ethernet Services Layer" http://www.metroethernetforum.org/PDFs/Standards/MEF12.pdf
[6] B. Alfonsi, "I Want My IPTV: Internet Protocol Television Predicted a Winner," IEEE Distributed Systems Online, IEEE Computer Society 1541-4922, vol. 6, no. 2, Feb. 2005.
[7] S. Cherry, "The Battle For Broadband," *IEEE Spectrum*, Jan. 2005, pp. 24–29.
[8] S.-H. Kim and D. Kataria, "Delivering the Next Big Motion Picture: IPTV," http://www.networksystemsdesignline.com/showArticle.jhtml?articleID=170700184, Nov. 2005.
[9] D. Kataria, "Traffic Management for Affordable Broadband," Electronicstalk.com, Mar. 22, 2004
[10] ATM Forum, "Traffic Management Specification Version 4.1," AF-TM-0121.000, Mar. 1999.
[11] N. Giroux and S. Ganti, *Quality of Service in ATM Networks: State of the Art Traffic Management*, Prentice Hall, 1999, p. 10.
[12] J. Xu and W. Lee, "Sustaining Availability of Web Services Under Distributed Denial of Service Attacks," *IEEE Trans. Comp.*, vol. 52, no. 2, Feb. 2003, pp 195-208.
[13] X. F. Wang and M. K. Reiter, "Defending Against Denial-of-Service Attacks with Puzzle Auctions," *Proc. 2003 IEEE Symp. Security & Privacy*.
[14] S. Vedantham, "Network Processors Battle e-Crimes," http://www.eetimes.com/showArticle.jhtml?articleID=173600 898, Nov. 2005.
[15] F. Hao, M. Kodialam, and T. V. Lakshman, "Real-Time Detection of Hidden Traffic Patterns," *Proc. 12th IEEE Int'l. Conf. Network Protocols*, 2004.

BIOGRAPHIES

SUNDAR VEDANTHAM (sundar@agere.com) is a senior systems engineer working in the Telecom, Enterprise & Networking (TEN) division of Agere Systems, Allentown, PA. He received his Ph.D. in computer science from Louisiana State University in 1997. He is working on system integration issues related to DSLAM, RNC/Node-B systems as part of Agere's telecom system integration group. His research interests include network traffic and congestion management, security, high-speed networking, theoretical computer models, and evolutionary computing.

SEONG-HWAN KIM (skim20@agere.com) received a Ph.D. degree in electrical engineering from State University of New York at Stony Brook in 2000. Since 2001 he has been with Agere Systems, where he is a Distinguished Member of Technical Staff in the System Integration group of the TEN division. He is currently working on data networking and wireless system integration, which includes residential gateway, SMB, RNC, and DSLAM application systems for which he has contributed to the architectural designs.

DEEPAK KATARIA (kataria@agere.com) holds a B.S. in electronics and communications engineering, and M.S. and Ph.D. degrees in electrical engineering from Rutgers University. He is systems integration manager in the TEN division of Agere Systems. He manages the development of enabling and value-added software for reference platforms targeting the home gateway, small-medium business gateway, wireline/wireless access networks, and carrier-Ethernet-based multiservice edge systems. His research interests include networking, traffic management, network security, multihop wireless, timing over packet, fixed-mobile convergence, reliability, and storage area networking.

An integrated traffic management solution implemented using network processors can provide a hierarchy of economies and revenue benefits valuable to service providers. This maximizes service affordability and fills an important void in making the ubiquitous Ethernet model successful.