Pipeline
Knowledge Is Power

www.pipelinepub.com **Volume 6, Issue 5**

Meeting the Challenges of Quality of Experience with Hierarchical QoS

By Manuel Stopnicki

Service providers know that managing the end-user experience is critical to reducing subscriber churn and the associated costs of capturing new customers. But offering a high quality customer experience today is no easy task. With an ever increasing array of new services, customers have varying degrees of quality expectations.

For example, a residential customer may want a fast Internet connection when downloading a YouTube movie, but five seconds give or take won't make a huge difference. That same customer will not be happy, however, if it takes even two seconds to change the channel when watching IPTV. The expectation is that changing channels should be instantaneous, and there should be no flickering while watching IPTV shows.

The expectations of end-users are influenced not only by the level of service that they've purchased but also by the particular application they are using. Service providers manage these different expectations by using a queues mechanism in the network to treat packets differently. This quality of service (QoS) system must be monitored closely to ensure that the expectations of end-users are being met or exceeded.

There are different types of network traffic, often broken down into the categories of voice, video and data. Network traffic must be treated differently depending upon the application characteristics it supports. This is basic QoS, and it is typically accomplished by having queues within a router, with each queue allocated a certain amount of time for a certain type of traffic. Data traffic may be given, for example, 10 megabits per second (Mbps) to allow for downloading a lot of YouTube, while the voice queue is allocated just 1 Mbps. However, the moment that router detects a voice packet in its queue, it must process it immediately because any delay will translate into echo in the voice conversation. So while the voice queue is allocated less bandwidth, it is still treated like an express line at the airport for frequent fliers. The moment that voice packet shows up, it must be processed no matter how many packets are waiting in line in the other queues.

Obviously, successful service providers are the ones who best optimize the use of their network bandwidth to serve their customers with the right level of quality for the different services they have subscribed. This is an increasingly complex task that calls for increasingly sophisticated mechanisms.

More Service Providers Turning to Hierarchical QoS

To overcome the lack of flexibility of standard QoS mechanisms, one of the major trends in QoS over the last couple years is called hierarchical QoS. To understand this concept, let's compare two customers receiving voice, video, and data traffic over broadband. Each customer is given designated queues based on their subscribed service level. Say one customer purchased the premium broadband service and gets 20 megabits per second, whereas the other bought the basic service that provides 10 megabits per second. Service providers typically design networks using a concept called over-subscription. Just because one customer bought 20 Mbps and the other bought 10 Mbps doesn't mean the network will be designed to handle 30 Mbps. Like an airline company that always overbooks because it knows some passengers are going to cancel at the last minute or not show up, service providers know that not everyone is going to be using the Internet at the same moment. Usually, there is a ratio based on how much has been sold, say an 8-to-1 ratio. So the network is designed to handle 1/8 the capacity that has been sold.

Because of this infrastructure capacity, a further limit, an additional gate, is needed beyond the limits put on customers based upon their subscription. These additional limits are placed on how many resources can be allocated toward each type of network traffic: data, voice and video. This is what is meant by hierarchical QoS. The queues of each customer, in our example the 20 Mbps and the 10 Mbps, are going to be put together into a different queue, called a "data queue", or a "nested queue." There are limits placed on this nested queue. So each customer has their own queue with its own limits defined, but then the parent queue has another limit defined, which is an aggregate of everything. So while a packet may fall within the limit of a customer's own queue, it may nevertheless get dropped by the parent queue because everybody in the neighborhood is watching the same Internet video, or the amount of traffic sent is more than the parent queue can handle for some other reason.

OSS Solutions Face Increasingly Complex Networks

Operating support systems (OSS) have been monitoring queues for a long time. It's important for network operators to know how many packets show up in a queue, how many are processed, how many are forwarded, how many are dropped because the queue was full or could not be processed, and how many were buffered for a long time. But when hierarchical queues are introduced to a network, OSS systems not only have to discover these new queues, they have to discover the relationships between parent and child queues. These performance management systems must analyze and report on both the parent and child queues and allow network operators to navigate around and have the visibility to see these relationships. So, the growing number of hierarchical systems makes network monitoring much more complex. In addition, when a customer is not receiving the expected quality of experience, you have multiple queues to troubleshoot. Was it in a customer's dedicated queue where a packet was dropped, or in an aggregated parent queue?

OSS systems are evolving to meet the challenges of growing network complexity. Having the right OSS system is critical to managing the end-user experience. Today's performance management systems should be capable of working with hierarchical QoS as they track and record - in real time - latency, jitter and packet loss.

Ultimately, for the best management of end-user experience, performance management solutions should be proactive in monitoring network and application efficiency in real time. These solutions should detect potential problems before they happen to reduce, or entirely eliminate, negative impacts on end-users. This proactive approach to service management is essential to meeting and exceeding customer expectations. Customers will be drawn to a truly reliable service, which means less customer churn and faster return on infrastructure investment—and that is good business.