

Economics Of QOS On WAN Access Lines

John Bartlett, Peter Sevcik, Sean Moore

Packet loss rates dictate the amount of traffic an access link can carry.

John Bartlett is a VP at NetForecast, where he focuses on real-time traffic, Internet performance and quality of service (QOS) techniques. John can be reached at john@netforecast.com. Peter Sevcik is president of Net Forecast and is a leading authority on Internet traffic, performance and technology. Peter can be reached at peter@netforecast.com. Sean Moore is the chief scientist in the converged infrastructure group at Avaya. He can be reached at smoore@avaya.com.

Supporting critical data, voice and video applications requires an IP network that will transport information with minimal packet loss and delay. Conventional network wisdom tells us that to achieve low loss and delay, we need to keep utilization low, often below 35 percent. Conventional financial wisdom, however, tells us that to keep expenses low, we need to fully utilize our network assets. This tradeoff is a discussion every CIO is having today, as the business demand for clear uninterrupted voice and video collides with the perennial pressure to control and justify IT expenses.

So what does this tradeoff really look like, and how can we pick an operating point—that is, a percentage of link utilization that simultaneously meets the performance and the financial goals of the business? In the following analysis, we answer this question, finding that packet loss rates dictate the maximum link utilization that can be achieved.

For links supporting a mix of high- and low-priority traffic, the highest utilizations come when well-behaved high-priority traffic, such as voice, is mixed with lower priority data traffic. No link utilization gains are observed when both high pri-

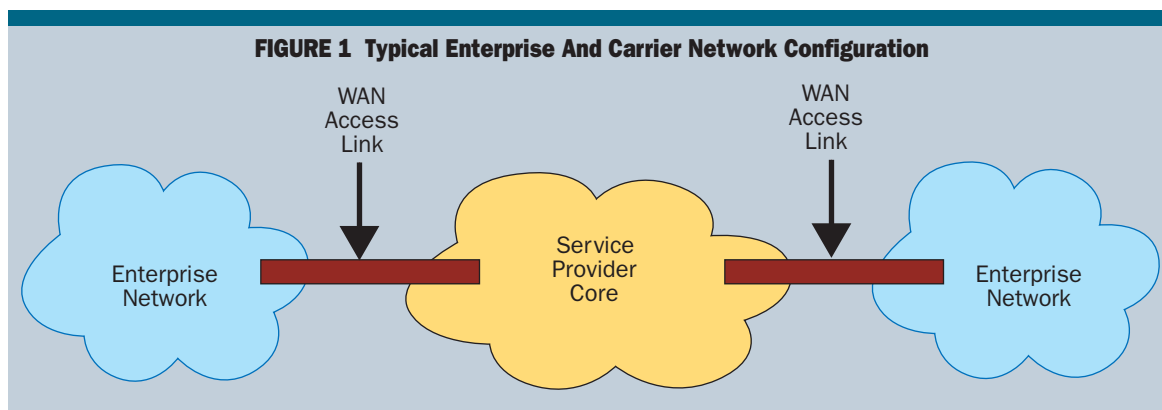
ority and low priority traffic are bursty data, only degraded performance of the low-priority data traffic.

To make proper use of these findings will take the enterprise network planner several steps beyond merely selecting the appropriate packet loss rate(s) for the service level agreement (SLA) with the carrier. That's because typical carrier SLA specifications are an average of the SLA value over the billing period, such as a calendar month. Customers should not be misled into thinking this average SLA will provide the necessary transport quality on the days and at the times when it is needed (see "How To Work Out An End-To-End SLA").

Also, most carrier SLAs cover only the high-speed core of the network, not the much lower-speed access links (Figure 1). Contention, delay, and packet loss are much more likely to occur on the access links, affecting not only the traffic leaving the enterprise and entering the service provider core, but also traffic leaving the service provider core and entering the enterprise. Our analysis focuses on these critical links.

The Factors That Matter

But first, let's review why we care. To support our applications on the network at performance levels that will satisfy business requirements, we need not only enough bandwidth, but also small delays,



and few lost packets. But what exactly does this mean? It means different things for different types of applications.

All IP data transaction applications and Web pages run over TCP/IP, and TCP automatically adjusts to the available bandwidth using its congestion avoidance algorithm. As a TCP connection gets congested, some packets will be lost, but TCP recovers them through retransmission and slows its transfer rate to keep all applications running. This means that response time to the user slows down as well. For some applications that is OK, but for mission critical applications, this may lead to lost productivity or lost sales.

Table 1 (p.18) shows the effect of packet loss on two public websites: as packet loss increases,

so do the page load times. The general rule of thumb with Web-based applications is that the bandwidth must be sufficient to not throttle an application for a single user, and then sufficient to keep packet loss below 1 percent for the aggregate user population during the busy period.

Voice and video applications behave quite differently. Unlike transaction and Web apps, voice and video run over UDP, which does not employ the congestion algorithm. Furthermore, these applications fail quickly when they experience congestion, rather than just slowing down. They need a network connection that will not become congested in the first place. The rule of thumb is that voice and video need less than 0.1 percent packet loss.

The rule of thumb is that voice and video need packet loss rates lower than 0.1 percent

How To Work Out An End-To-End SLA

When a carrier commits to an average SLA of 0.1 percent packet loss in a month, actual performance could be as poor as 0.6 percent packet loss when you discount nights and weekends from the averaging. This could be as bad as a 5-minute period, occurring 10 times per workday, when packet loss is 5 percent. If those 5-minute periods occur when enterprise communication quality is most important, the network could fail to provide adequate support.

So how can an enterprise specify an SLA that will actually meet its needs for application traffic? Here is an approach to consider. First, determine the required packet loss rate for high and low priority traffic, based on the applications you must support and the performance requirements of those applications. Diagram your network as shown in Figure 1, p.16, and determine the number of router hops in the access network(s) as packets move from one enterprise location to another.

Next, take the carrier's packet loss rate SLA for the core network, and degrade it by an order of magnitude to account for the extensive averaging used. This number is the best possible SLA value for the busy hour periods of your business, assuming your busy hour periods coincide with the majority of users of the WAN core. If this number is not sufficient for your SLA, search for another carrier.

Assuming the core network number arrived at above is sufficient, allocate the remaining packet loss across the number of hops in your access network(s), and size the links to meet that specification. Let's work through an example for clarity:

1. Suppose we have a 2-hop access network, one hop into the carrier cloud and one hop out, and we want to meet an end-to-end packet loss rate of 1 percent for data applications.

2. The carrier specifies a monthly average

packet loss SLA of 0.05 percent. We degrade the carrier SLA to 0.5 percent to represent the reality of the workday or the busy hour.

3. This leaves us with 0.5 percent (1 percent - 0.5 percent = 0.5 percent) loss rate to allocate to our access networks, giving us a spec of 0.25 percent for each link.

Referring to Figure 3, p. 19, we can determine how much traffic the link will support and still meet our SLA. We find 0.25 percent on the vertical axis and follow it to the right to the bursty traffic line, finding that a utilization of about 38 percent (on the horizontal axis) is the maximum we can allow if we wish to stay within our specification.

If we were hoping to use this same wide area network for voice traffic, and want an SLA of 0.1 percent packet loss to support that, we may be in trouble, given the 0.5 percent discounted SLA on the core network. Testing ought to be done at this point, during multiple busy hour periods, to determine the core network behavior during enterprise busy hour periods. Prioritizing voice traffic on the access links will not help if the core will not support the required SLA.

Your carrier might tell you that there is no need for QOS (preferential treatment for high-priority traffic) in the network core, because they specify such a high-quality SLA (low packet loss rate) for that core. While this may be true at face value, remember that the carrier is not measuring the loss on the access links as a part of their SLA, and that is where you really need to prioritize.

Finally, remember that the enterprise edge router can only prioritize traffic headed into the core. The carrier must perform the same function for traffic coming into the enterprise from the core—even if they are not implementing QOS within the core itself□

Packet loss rates for TDM emulation must be less than 0.001 percent

TABLE 1 How Packet Loss Affects Web Page Load Times

Packet Loss	Page Load Time (Seconds)	
	Motorola	Amazon
0.01 percent	3.4	8.8
0.10 percent	3.5	9
1 percent	4	10.2
2 percent	4.5	11.6
4 percent	5.6	14.4
6 percent	6.8	17.2

Video applications, in particular, require high network quality. Where voice over IP (VOIP) can lose a few packets and still sound quite good, video has a more difficult time. Video compression relies on sending incremental information, the portions of a video image that have changed since the last image was sent. When an update is lost, the subsequent incremental updates are useless because they are building on top of information that is missing.

One of the most stringent new IP data applications is TDM emulation. As is often the case with a new technology, the old technology must continue to exist at the edge of the network until applications for it finally age out, long after the network core has changed. TDM emulation provides traditional TDM interfaces at the edge of a network (T1, E1, T3, E3, etc.), converting these streams into IP packets for transport across the IP core, and then reconstructing the TDM data stream and timing at the far end. Packet loss of no worse than 1 in 10^{-6} (.001 percent) is required to properly emulate legacy TDM services.

So a “level of quality,” specified as a minimum bandwidth, maximum packet loss and maximum delay, is needed that will ensure that our critical applications will work properly and work with sufficient performance to support our business objectives. We also need a way to predict the quality we will obtain from the complete network system so we can make sure our applications are going to work sufficiently well to support our business objectives while also meeting our budget requirements. Let’s first dive into the question of how IP links behave, and then come back to the economic question.

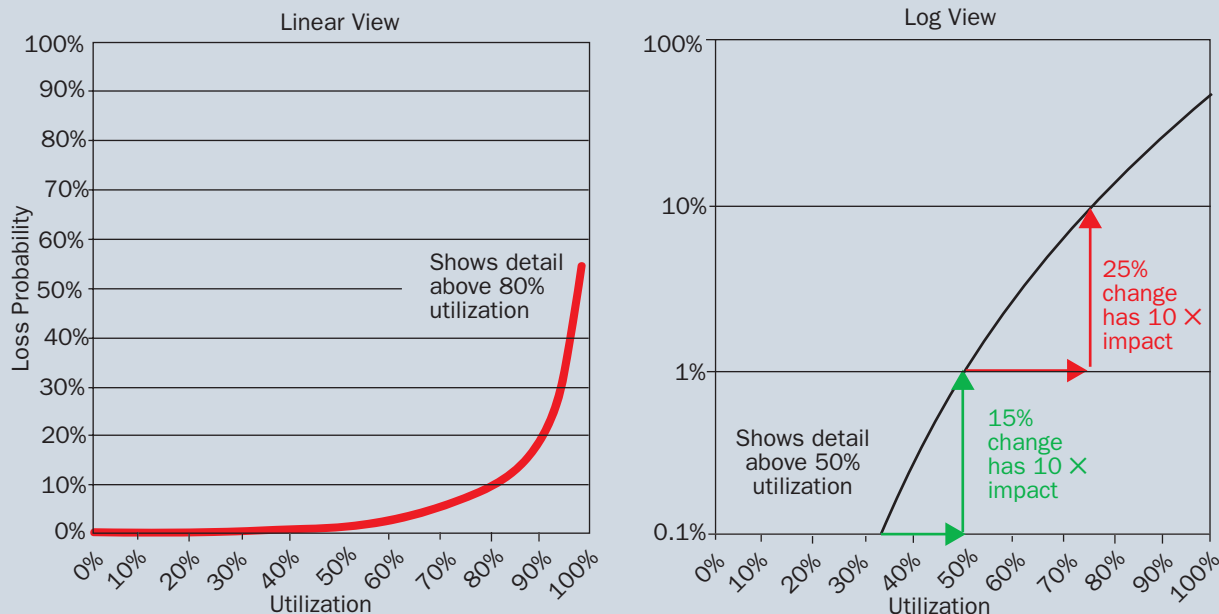
The Relationship Between Utilization And Loss

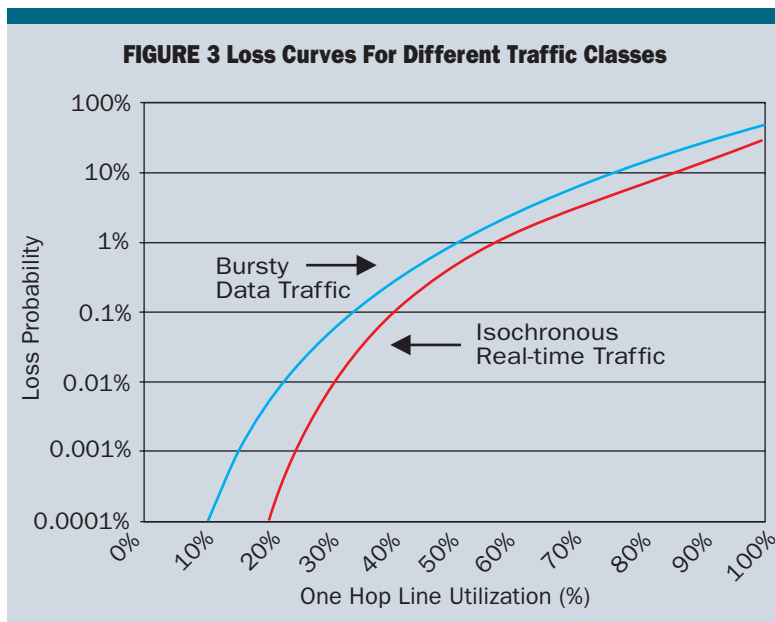
Delay and loss performance in an IP network is a function of router buffers, link capacities and link utilization, i.e., the traffic load carried by the link relative to its capacity. Buffer resources and link capacities are fixed, so determining the number and the size of links to buy and subsequent operational costs boil down to a simple efficiency question: For a pair of average loss and delay requirements, what is the maximum permissible utilization of a link?

Recall that the rule-of-thumb packet loss requirements are 0.1 percent for voice and video and 1 percent for bursty data. Delay requirements also vary by application type, with the most stringent requirements for voice and video. Most people notice round-trip delays in conversation of more than 100 milliseconds, and find the experience of delays over 250 milliseconds (a quarter of a second) objectionable.

Some forms of delay, including propagation delay, cannot be reduced, but others can be controlled. For example, one can control the

FIGURE 2 Bursty Traffic Loss Curve





queue size limit, the link speed and the utilization of the link, thus reducing these sources of queuing delay. Because access lines are small by comparison to the ISP core and the enterprise network, they are typically a large contributor to the overall end-to-end delay. So even though a large queue buffer would provide a good cushion against loss, it is often not prescribed, in order to keep the delay short. Ensuring that an SLA's delay requirement can be met means limiting the size of the queues such that the maximum possible queuing delay, summed across all the routers in the path, plus propagation delay, does not exceed the SLA delay requirement.

Therefore, with a delay boundary fixed by the combination of link capacity and queue size limit, the method for controlling packet loss is to control the average traffic loading of the link, or link utilization. A lower level of utilization or a higher-capacity link will result in fewer episodes during which queues are full, which in turn results in lower packet loss rates.

Bursty Data Traffic

Modeling bursty traffic is not straightforward, and it has generated much interest and debate in the network traffic research community during the past decade. We based our models on those developed by researchers at MCI/WorldCom because their models accurately track empirical data collected from the WorldCom backbone (see L. Yao, "Queuing Analysis And Control Of Long Range Dependent Traffic: Applications to Internet Traffic Engineering," Ph.D. Dissertation, George Washington University; and L. Yao et al., "Long Range Dependence in Internet Backbone Traffic," *IEEE International Conference on Communications (ICC2003)*, May 2003, pp. 1611-1615). The basic loss curve for bursty traffic on a 1.5 Mbps access line is shown in Figure 2. This modeling was done

with a queue depth limit of 10 milliseconds.

The reader will be familiar with the linear-scale view (left chart) in Figure 2. It shows the classic problem of a queue building up and packets getting dropped as utilization increases past 50 percent. At 70 percent utilization, the loss starts to get significant, and the circuit becomes essentially unusable above 90 percent utilization. However, what appears flat in the linear view (i.e., well below the 50 percent utilization point) is

the important section of the curve. As the log scale chart on the right shows, an order of magnitude increase in packet loss occurs not only when utilization grows from 50 percent to 75 percent, but also when utilization increases from 35 percent to 50 percent. The lesson is that a small change in utilization at the low end has a big impact on loss.

Real-Time Voice And Video Traffic

Aggregated packetized voice is well modeled by a (non-bursty) Poisson process, one justification being that voice packet devices such as typical IP telephones generate isochronous packet streams (e.g., uniformly spaced packets), and a large collection of isochronous processes with random phases forms a Poisson process. The models used to generate the non-bursty curves in Figure 3 are found in D. Bertsekas, R. Gallager, *Data Networks*, Prentice Hall, NJ, 1992.

The graph in Figure 3 shows two log-scale curves that describe the packet loss behavior of bursty (the higher, blue curve) and non-bursty (the lower, red curve) traffic. The higher curve represents data applications running TCP, the type of traffic that dominates the Internet today. The lower curve represents isochronous real-time traffic like VOIP or video over IP. These so-called "well-behaved" traffic types have a transmission rate limit imposed by design. This means that individually, they do not exceed some value—for example, a voice call only runs to 80 kbps using G.711, and an H.323 video call using H.261 at 384 kbps only consumes 420 kbps. They never burst to higher values, unlike data applications.

To use the graph in Figure 3, pick a utilization operating point on the X-axis, follow it up until you cross the line representing the type of traffic you support, and then follow that point left to the Y-axis to determine the resulting packet loss probability. Conversely, if your goal is to obtain a

Packet loss can be reduced by lower link utilization or higher link capacity

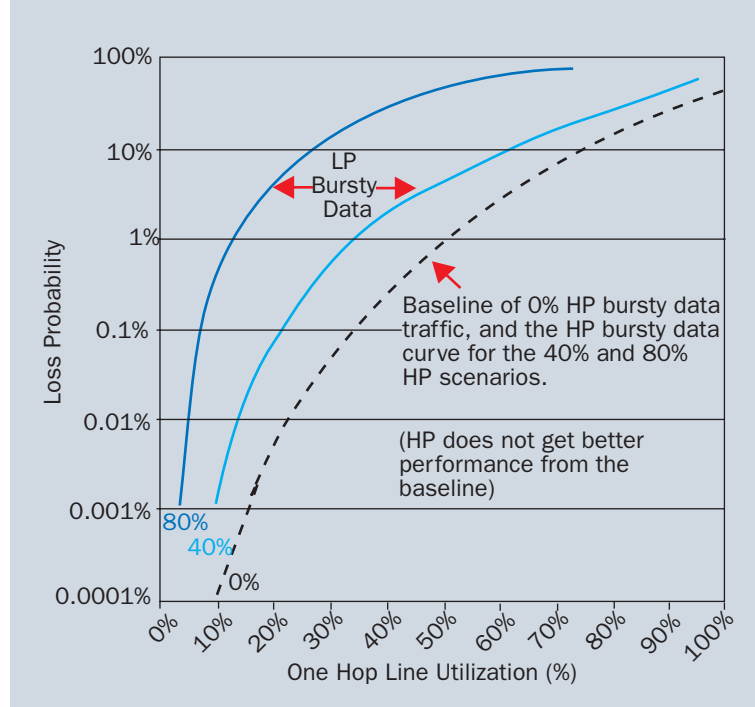
Avoid mixing too much voice with data traffic, or the data applications will suffer

specific packet loss probability, often defined in a service level agreement (SLA), choose the loss probability on the Y-axis, and work the graph in the other direction to find a suitable utilization operating point. As discussed in the sidebar, carrier-supplied average SLA commitments are just a starting point for end-to-end performance engineering on enterprise networks.

You can use the curves in Figure 3 to estimate where you would operate a circuit that supports

both voice and data. For example, in order to support a 0.1 percent loss rate for voice, you need to keep the voice below 40 percent utilization, while data at a 1 percent loss rate must stay below 50 percent utilization. Without a way to prioritize voice, however, you must pick the lower number (40 percent) in order to get the performance you require. Clearly there is some room to use this asset more efficiently—if real-time and best-effort traffic loads are known, if real-time traffic is prioritized, and if the two classes of traffic are accurately modeled for specific mix scenarios. Let's explore a few scenarios to see what mixes of traffic can be carried.

FIGURE 4 Applying QOS To High Priority Bursty Data Traffic

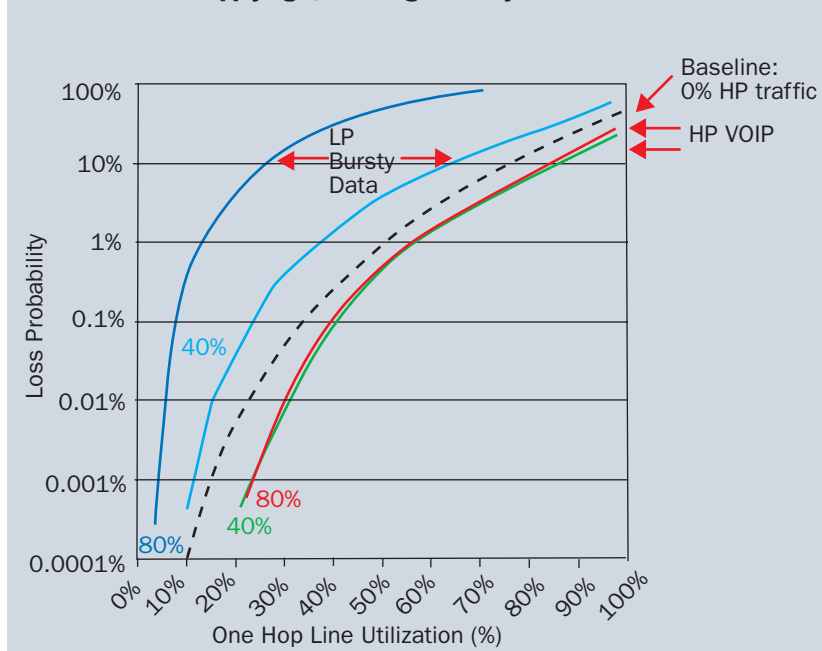


How QOS Improves Performance—But Only For High-Priority Traffic

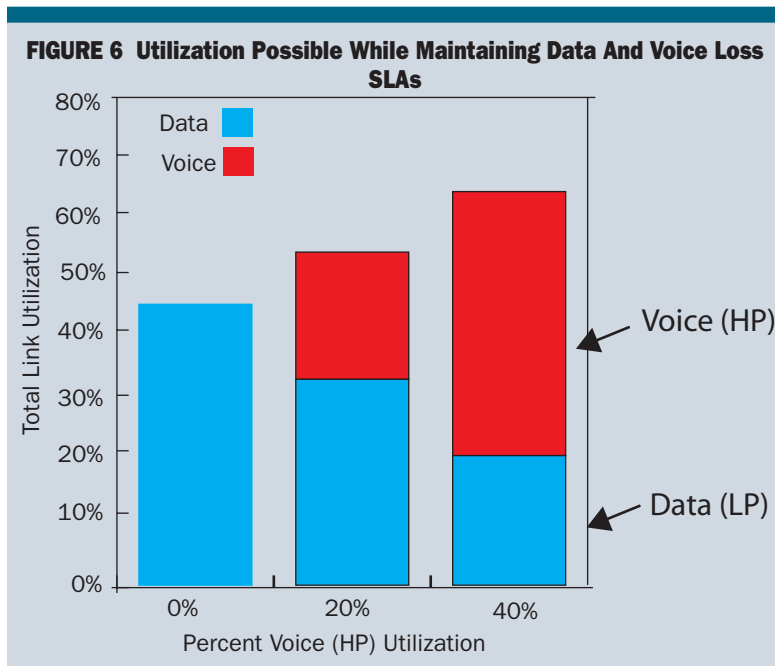
Let's start with a bursty data traffic mix, some of which is mission-critical, hence high-priority, and the rest of which is best effort. Knowing the mix, or the percentage of high-priority traffic being carried, is critical to this analysis.

Figure 4 shows the effect on the bursty traffic that is being handled by a high priority (HP) and low priority (LP) QOS system. The high priority traffic gets access to all the link bandwidth, so the HP curve (the dashed line in Figure 4) is the same as the baseline no matter what percentage of the traffic is high priority. In other words, tagging some of the bursty packets as high priority and putting them ahead does not improve their performance over the baseline. Note also that the dashed baseline curve is the same as the bursty data traffic curve (the blue curve) in Figure 3.

FIGURE 5 Applying QOS to High Priority Voice Over IP Traffic



As the HP traffic increases, the low priority bursty traffic experiences an increasingly significant negative impact, because LP traffic only has access to its percent-



age of the bandwidth after the high priority traffic has consumed its percentage.

For example, take a mix of 40 percent HP and 60 percent LP traffic on a total link utilization of 60 percent. The HP traffic (60 percent \times 40 percent = 24 percent) will experience 0.02 percent packet loss, while LP traffic will see 3 percent packet loss.

To determine the loss for the low-priority traffic we first calculate that HP traffic is consuming (60 percent \times 40 percent =) 24 percent of the link bandwidth. Thus the available bandwidth for LP traffic is (100 percent - 24 percent =) 76 percent, so low priority traffic is operating at 36 percent/76

percent = 47 percent of its available bandwidth.

Good News/Bad News For Prioritizing VOIP

Now let's look at prioritizing traffic which is not as bursty, specifically voice. Consider the example of voice and data traffic mixed on a single network. If we prioritize the voice packets, they will always be sent ahead of any data packets in the queue. Thus the entire bandwidth of the link is always available for voice traffic, because whenever voice traffic appears, it gets to be forwarded immediately.

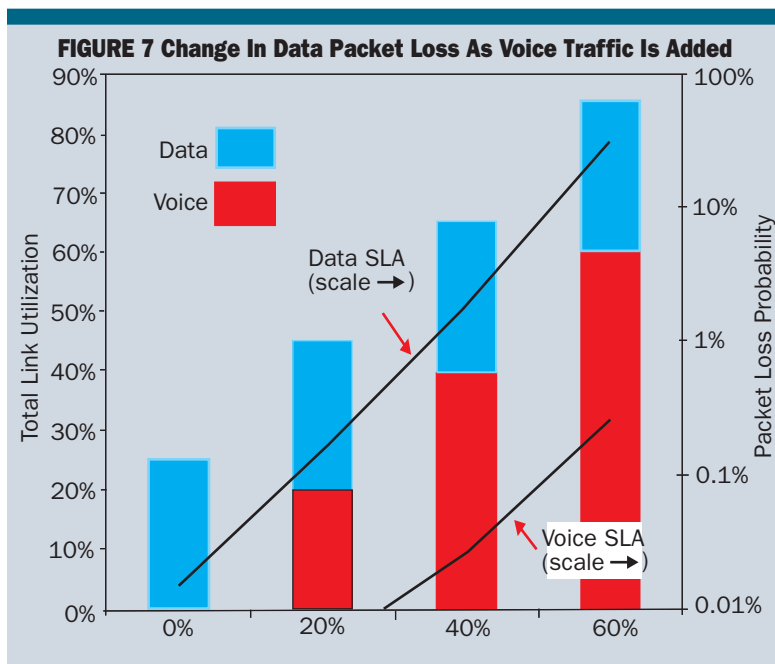
Figure 5 shows the curves for different percentage mixes of high priority voice traffic and low priority data traffic.


There is good news and bad news in this graph.

The effect of giving HP service to the voice traffic is to place it on the same curve it had when it was the only traffic on the link (the red and green curves in Figure 5 are in the same location as the red curve in Figure 3). So QOS does provide the expected effect of giving voice traffic an unencumbered line when it needs it. There is some slight impact to the voice packets from low priority traffic that may still be in the process of being clocked onto the line



A traffic mix of 40 percent voice and 20 percent data will meet the packet loss SLA requirements for both





We have heard conflicting reports on carrier link utilization, delay and loss performance

when the high priority voice packet arrived, but the effect is minimal in this analysis.

The bad news is that the low priority data traffic is just as severely affected by prioritizing VOIP traffic as it is by prioritizing some data traffic. The blue curves in Figures 4 and 5 match up very well.

This analysis shows that QoS allows voice to perform nearly as well as it would on an empty link. But the impact on data traffic is significant. The mix of voice and data is critical, and it is easy to slip into scenarios where voice gets the service it needs, but negatively affects the data applications.

The Economic Value Of QoS

Let's look at the relationship between utilization and packet loss in a slightly different way. Assume, for example, that we want to maintain a packet loss rate of less than 0.1 percent for the voice (HP) traffic, and less than 1 percent for the data traffic (LP). Packet loss can occur on both of the access links—the hops into and out of the ISP network backbone—so we must add the loss characteristics of the two access links. At low loss rates (1 percent and below), simple addition of the rates from each link is a good approximation of the overall loss rate. For example, to find the utilization operating point in Figure 5 for a two-hop network with an overall loss rate of 1 percent, we look for the intersection of utilization and the 0.5 percent loss point on the graph (0.5 percent + 0.5 percent = 1 percent).

Figure 6 shows the utilization on a link possible for different mixes of voice and data (high priority and low priority) traffic. The first column in Figure 6 shows that 45 percent is the maximum link utilization that can be achieved for data traffic alone while keeping packet loss at 1 percent or less. As voice traffic is added to the mix, the amount of data traffic must be reduced to maintain the 1 percent (data) packet loss rate.

In this example, when voice traffic reaches 40 percent of the link bandwidth (in the right hand column), we can no longer add voice traffic without potentially violating its 0.1 percent loss rate, and data traffic can't be more than about 20 percent while maintaining its 1 percent loss rate. Adding either voice or data traffic beyond these numbers will violate one SLA or the other.

Note that overall utilization climbs as the voice traffic is added. This occurs because the voice traffic is less bursty than the data traffic; it can consume more total bandwidth without violating its packet loss SLA. Gains beyond 60 percent total utilization are not possible without violating the packet loss SLAs set for this example. Thus, under these conditions and SLA requirements, the most efficient use of the total bandwidth is achieved with 2:1 voice:data mix.

Figure 7 shows another way of viewing the same problem. In this example, the link is assumed to have 25 percent data traffic, which will

be maintained as voice traffic is added to the link. The chart shows where the packet loss SLAs are violated for both the voice (high priority) and the data (low priority) traffic. Although voice doesn't violate its packet loss SLA of 0.1 percent until it passes 40 percent utilization, the data SLA of 1 percent is violated below 40 percent voice utilization. This confirms the previous example, in which we had to limit data traffic to 20 percent of the link when voice traffic moved to 40 percent to maintain the SLA.

What Happens In The WAN?

High-speed WAN links that operate above 100 Mbps have different behavior from the 1.5-Mbps model we have been discussing. Recall that queue length is often determined by the latency requirement of an SLA; a shorter queue will help deliver a lower latency, but shorter queues can also lead to higher loss. In the high-speed WAN core, however, latency requirements can be met with much larger queues, since the queues are drained so quickly by the high-speed links.

For example, a 1,500-byte packet (Ethernet's maximum) takes 8 milliseconds out of the end-to-end delay budget at 1.5 Mbps. However, the same packet is off the trunk queue in 12 microseconds on a 1-Gbps line. This means much larger buffers can be used in a WAN switch/router without violating the delay budget.

We have heard conflicting reports on how well carriers minimize delay and loss. There is much confusion and mystery about the planned, practical, and actual utilization operating points of carrier networks. Enterprises have to take the carrier's SLA statements as starting points and do the analysis required to engineer a solution.

Enterprises that operate a private WAN are not likely to have Gigabit trunks, but they should have a better handle on the actual WAN delay and loss rates, since they can both measure the WAN's performance and model it as a complete end-to-end system with fairly accurate results.

Conclusion

The first law of thermodynamics says there is no free lunch, and the second law says you can't even break even. It appears that the same rules apply to stuffing packets into network links. Because our traffic is bursty, overhead is required to ensure a quality connection. Newer applications like voice and video require better quality, and are less tolerant of the packet loss caused in poor connections. SLA design and management will be required as we implement more of these loss-sensitive applications, and as we require our traditional data applications to meet performance goals critical to running the business of the enterprise □