WHITE PAPER · ENTERPRISE NETWORK ARCHITECTURE

QoS Doesn’t Create Bandwidth

Where congestion management actually belongs — and what multi-class MPLS is really buying

Queuing only does something when a link is congested. Across a 500-site, 60-country WAN, the congested link was almost never the carrier backbone — it was the local access circuit. That single observation changes what you should be willing to pay a provider to arbitrate.

BY **JACK NASH** 500 SITES · 60 COUNTRIES ABOUT 10 MIN READ

THE ARGUMENT IN BRIEF

- Queuing only does something when a link is congested. **QoS doesn't create bandwidth** — it decides who loses when there isn't enough.
- In our environment the congested link was almost never the carrier backbone. It was the **local access circuit**, commonly the narrowest hop in the path.
- Congestion management only helps **before** the constrained link. That produces a real problem for inbound traffic — and the best case for buying carrier CoS.
- Our answer was to **remove the constraint** where broadband made that cheap, and to place application-aware policy at both ends where we controlled it.
- Underneath it: **if the backbone is congested, don't buy the backbone. If it isn't, don't buy multi-class MPLS.**

The objection

Between approximately 2007 and 2009, while at Cadbury, I proposed moving a global corporate network of roughly 500 sites across 60 countries off multi-class MPLS and onto locally purchased Internet access.

Application performance was the first serious objection, and it was a fair one.

The conventional enterprise answer at the time was multi-class MPLS. Voice received one Class of Service. Video received another. Business-critical applications received another. Everything else received best effort. You bought the classes, you mapped your applications into them, and performance became something the carrier managed on your behalf.

I didn't set out to argue against that. I set out to understand what it was actually doing, because we were paying a significant premium for it and I couldn't explain to myself what the premium bought.

What I found changed how I thought about buying network performance — and it started with a fairly basic question about queuing.

What queuing actually does

Take Cisco's Class-Based Weighted Fair Queuing as a worked example. Under CBWFQ, configured bandwidth guarantees become relevant during congestion. When an egress interface has sufficient capacity to transmit the traffic presented to it, there is no constrained bandwidth for the queuing system to arbitrate among competing classes. Packets are simply transmitted.

When congestion occurs, the situation changes. Now the router needs rules determining how constrained capacity should be allocated, and the class configuration decides which traffic is protected and which is not.

That distinction sounds obvious stated plainly, but its implications are easy to miss when Class of Service is presented as a product line rather than a queuing behavior:

> **QoS doesn't create bandwidth. It determines how constrained bandwidth is allocated.**

Which raises the question that ought to precede any CoS purchase. If queuing only matters where a link is congested, where in the path is the congestion?

Where the constraint actually was

Consider the shape of a traditional MPLS path.

A TRADITIONAL MPLS END-TO-END PATH

The shaded links are the customer access circuits — commonly the narrowest hops in the entire path.

The provider backbone generally operated at substantially greater capacity than any individual customer's access circuit. The local CE-to-PE access link was commonly one of the smallest links in the end-to-end path.

That wasn't universally true, and I wouldn't claim it as a law. But in our environment it was frequently the case. A provider might operate an enormous backbone while one of our sites reached it through a relatively small T1/E1, DS3/E3, or similar circuit. The mismatch was often an order of magnitude or more.

Our own operational data pointed the same direction. Approximately 75 to 80 percent of our total WAN cost was associated with local access circuits, and approximately 98 percent of our circuit outages involved local access rather than the provider's core. The access layer was where the money went, where the failures happened, and where the capacity was tightest.

So if congestion management was going to deliver meaningful value anywhere in that path, the constrained access link was the obvious candidate — not the backbone.

But knowing where the constraint is only gets you halfway. Congestion management also has to happen in the right place relative to it.

Before the bottleneck, not after

Congestion management is useful before traffic traverses the constrained resource. After is too late.

For traffic leaving one of our sites, this was straightforward. If our CE was about to transmit more traffic than the local access circuit could carry, we could classify, queue, and prioritize that traffic before placing it onto the circuit. The decision happened at the right point. Outbound was a solved problem, and it was solved on equipment we owned.

Inbound traffic was a different matter. Once the provider had transmitted traffic across the constrained access circuit toward our CE, the scarce bandwidth had already been consumed. Dropping a low-priority packet after it arrived at our router did not give that bandwidth back. The packet had already occupied the circuit. The congestion-management decision had been made too late to matter.

> **The best place to manage congestion is before the constrained link.**

This is worth sitting with, because it leads somewhere uncomfortable for the argument I was making.

The best argument for carrier CoS

If the constrained link is the access circuit, and congestion management has to occur before the constrained link, then for inbound traffic the only party positioned to act in time is the provider — queuing at the PE, on egress toward the customer.

That is precisely what multi-class MPLS sells. It is the strongest argument for buying carrier Class of Service, and it deserves to be stated at full strength rather than glossed over. A customer cannot queue traffic that has already crossed the link.

We had two answers, and I want to be honest that they are answers rather than refutations.

The first was to remove the constraint. Queuing is an arbitration mechanism for scarcity, so scarcity is the thing worth attacking. Local ISPs were increasingly delivering DSL and other broadband technologies built for a mass market, frequently offering considerably more bandwidth, faster installation, and dramatically lower cost than the enterprise access circuits we were buying. Where that was true — and by 2007 it was true in a growing number of our markets — we could often purchase several times the capacity for less than we were paying to carefully arbitrate a scarce T1 or E1.

> **The best congestion-management strategy is to eliminate the constraint.**

The second answer was architectural. We deployed WAN optimization at every site, which meant that for any given flow there was equipment we controlled at both ends. The appliance at the sending site could shape and prioritize before transmission — which is inbound congestion management for the receiving site, performed at the correct point in the path, by us.

That is the part worth understanding. Carrier CoS is not the only way to act before the constrained link. It is the only way to act before the constrained link *if the enterprise owns nothing at the far end.* We chose to own something at the far end.

Neither answer is free, and neither is universal. If your access circuits are genuinely scarce and cannot be economically upgraded, and you have no presence at the remote end of your flows, carrier CoS is doing real work for you and you should buy it. Our environment simply wasn't that environment.

The CoS paradox

Now turn from the access circuit to the provider backbone, where multi-class MPLS is also sold.

Suppose a global service provider routinely operates its backbone with enough congestion that sophisticated queuing is necessary to protect my applications. Then I have a larger problem than queuing: I am buying connectivity from the wrong provider.

A properly engineered global backbone should maintain sufficient capacity and headroom. At the time, our experience and understanding of major carrier engineering practice was that backbone capacity was generally groomed well before sustained utilization approached saturation, often somewhere around the 70 percent range. I did not want to purchase an inadequately provisioned backbone and then pay that same provider an additional premium so my packets would be favored when its infrastructure became congested.

Now take the opposite case. The provider operates a properly engineered backbone. There is sufficient capacity. Its interfaces are not congested. There is no constrained resource requiring CBWFQ to choose between my application classes — which means the class configuration I am paying for is, on that portion of the path, largely inert.

THE PARADOX OF MULTI-CLASS MPLS

If the provider's network is congested, don't buy the network.

If the provider's network isn't congested, don't buy multi-class MPLS.

The paradox applies to the backbone portion of the service. It does not dissolve the inbound access problem above, which is real and which I addressed separately. Keeping those two apart is what makes the argument survive scrutiny.

What the premium was actually buying

There was one more realization that reframed the purchase for me.

The global telecommunications providers we dealt with were not necessarily operating completely separate physical backbones for every service they sold. Their Internet services and their premium enterprise services could share substantial portions of the same underlying infrastructure. We were paying a premium largely for how the service was packaged, managed, prioritized, and supported — not necessarily because every packet traversed an entirely different physical network.

That caused me to separate two things enterprise telecommunications had always bundled together: connectivity and network architecture.

Once separated, the question changes. It stops being *which class structure should we purchase?* and becomes *which party should own the traffic policy?*

Own the policy instead

Eliminating carrier-provided multi-class MPLS did not eliminate our requirement for application performance. It moved responsibility for it onto us.

Each corporate site had a WAN optimization appliance positioned behind its edge firewall, which wrapped the network in an end-to-end, application-aware performance layer.

Instead of purchasing generic traffic classes and mapping our applications into someone else's taxonomy, our own infrastructure could recognize and prioritize actual applications. We made performance decisions at Layer 7 according to our business requirements rather than according to a carrier's product catalog.

And the policy belonged to us. Changing ISPs didn't remove it. Changing circuits didn't remove it. Different countries could use completely different access providers without requiring us to redesign application performance around each carrier's class structure — which, across 60 countries, was not a small consideration. It was one of the things that made the access layer genuinely interchangeable.

How to test this in your own environment

None of the above is a general law. It was a conclusion about a specific network, reached from that network's own data. The useful part is the method, not my answer.

If you are renewing a multi-class MPLS contract, these are the questions I would want answered first.

Where is the narrowest link in your typical path? Compare a representative site's access bandwidth against the backbone capacity your provider advertises. If your access circuit is an order of magnitude smaller, your congestion lives at the edge, and backbone class structure is not where your performance problem is.

How often are your access circuits actually congested? Pull utilization on the CE side. If sustained utilization rarely approaches capacity, you are paying to arbitrate a resource that isn't contended. If it regularly saturates, you have a capacity problem that queuing will manage but not solve.

What would more bandwidth cost? Price the local broadband alternative in your actual markets, not the headline markets. Then compare that against the CoS premium. In a number of our countries, the additional capacity cost less than the arbitration.

Do you own equipment at both ends of your important flows? This determines whether you can manage inbound congestion yourself. If your significant traffic is site-to-site between locations you control, you have options. If it terminates at third parties you don't, the carrier's position in the path is genuinely more valuable.

Which applications actually map cleanly into the classes you're buying? If your traffic doesn't decompose neatly into voice, video, business-critical, and best effort — and by the late 2000s ours increasingly didn't — then Layer 7 classification you control may fit your business better than four generic buckets.

What does your incident data say? Ours said approximately 98 percent of circuit outages were local access. If yours says something similar, spend your attention where the failures are.

Answer those six honestly and the decision usually makes itself. It may well make itself in favor of buying CoS. That's a legitimate outcome — the point is to reach it from your own measurements rather than from a product structure you inherited.

What this is and isn’t an argument for

This was never an argument against QoS. QoS is an excellent engineering tool, and when a constrained resource genuinely needs to be managed, queuing is exactly the right mechanism. I used it extensively on equipment we controlled.

It is an argument for putting congestion management where the constraint actually exists, and for being precise about which party is positioned to act before the bottleneck rather than after it.

And it is an argument against a specific habit: paying a global carrier a premium to provide multiple traffic classes throughout a backbone that should not routinely be congested in the first place, while the link that is actually congested sits outside that arrangement.

Buy connectivity. Own the intelligence.

Understand what the premium is arbitrating before you pay it.

A note on the account. This is a first-person recollection of work carried out between approximately 2007 and 2009. Percentages and cost figures are approximate and drawn from internal analysis of the period. CBWFQ is referenced as a worked example of weighted queuing behavior; product and category names are used descriptively and belong to their respective owners.