My Beloved IP Is Getting Trashed

Troy Dixler

We're asking a lot of IP, whose initial claim to fame was its simplicity. Are we asking it to do too much?

We're turning the Internet and IP into a virtual sewer. Now before you start throwing stuff, let me tell you where I'm coming from: I'm an IP bigot, and a devoted, long-standing member of North American Network Operators Group (NANOG) and the IETF.

But over the last year, I've become frustrated watching the IP "gurus" introduce a horde of new standards (many still "drafts"), technologies and services, all of which appear to serve a noble purpose: To make IP suitable for any and all types of applications.

We've seen this movie before; it was called ATM. You'll recall, ATM was going to bring about the utopia of convergence—the integration of voice, video and data. It didn't. While it delivered the service guarantees needed for different traffic types, those capabilities came at a high cost, high level of complexity and high overhead (i.e. cell tax). I've seen better movies.

Now the ATM missionaries, having conceded that the world is becoming all-IP, have taken the connection vs. connectionless debate to the IETF. Dozens of working groups, with hundreds of vendors, are working on new standards and technologies. There's work going on to:

- Create ATM-like connections across IP networks.
- Deliver multiple Layer 2 services (i.e., ATM, frame relay, Ethernet) over a single IP network.
- Guarantee service quality end-to-end.
- Construct VPNs across a shared IP backbone *(a la* frame relay or ATM).

The industry is spinning under a "draft-of-the-week" regime. But all these various Layer 2 and Layer 3 proposals, each designed to solve a specific point problem, fail one critical test: They don't make for simpler IP networks.

Instead, they are creating confusion (read: sewage) at the network edge. It's becoming next to impossible for carriers to filter out the winners and focus in on a single architectural approach.

Let me be clear: Many of the new protocols and technologies, such as MPLS and QOS, are useful for creating more predictable traffic flows across a shared IP infrastructure (i.e., the Internet).

But that's also the problem. Resource sharing always involves compromise, and it lacks true predictability and guarantees. This is not only true for logical circuits (e.g., MPLS label switched paths) but also for physical assets like routers, where schemes like virtual routing have been proposed. Because these schemes are based on a fundamentally flawed approach—sharing critical routing resources—they don't result in stability for carriers; instead, they increase complexity.

The real problem is that IP, a connectionless protocol, was never developed to be the *universal* protocol. ATM was developed to serve that purpose and failed.

Myths vs. Reality

MPLS is being touted as the way to deliver many applications where a connection-oriented strategy is required. For applications like traffic engineering, VPN services and maintaining service-level guarantees, for example, the going-forward assumption is that MPLS will be the transport mechanism for future IP networks and services.

But there's a huge gap between what MPLS is supposed to enable, and what the carriers that have actually implemented it are doing. Originally, *online* traffic engineering was intended to allow providers to dynamically re-optimize their networks to take advantage of underutilized paths. But the carriers using MPLS today still do *offline* traffic engineering—they make periodic changes to shift traffic from one link to another when appropriate.

This is largely because the network operators don't trust the algorithms, which creates a Catch-22: Operators lose control of their networks when online traffic engineering is enabled, but doing it offline undermines the efficiency gains the online approach was created to deliver.

Meanwhile, every virtual service application is being promoted as if the carrier networks are MPLS-enabled edge-to-edge. The reality on the ground is quite different. It can take up to two years to schedule and deploy a network redesign based on MPLS's basic building blocks. Moreover, these applications could be deployed on a pure IP backbone without the complexity of MPLS, but doing so requires rethinking how routing technology is deployed, bought and sold (see "Do Routers Need Fixing?")

Troy Dixler (troy@allegronetworks.com) is a founding engineer at Allegro Networks in San Jose.

We also hear about collapsing multiple Layer 2 networks (ATM, frame relay and Ethernet) onto a single IP multiservice edge router, and the IETF's Martini draft (draft-martini-l2circuit-trans-mpls-08.txt), is being developed in the IETF Pseudo Wire Emulation (PWE3) as part of this effort (see *BCR*, February 2002, pp. 29–35). But many of the largest carriers operate discrete ATM, frame relay and IP networks, each of which can have hundreds of Layer 2 and 3 devices. Collapsing all three networks into one will create a single point of failure for everyone on the system. Other alternatives, such as L2TPv3/UTI tunneling, don't require end-to-end MPLS, and enable carriers to offer Layer 2 services over a non-MPLS-enabled IP backbone.

A New Layer In The OSI Model

Another area of IP sewage is quality of service (QOS) and the politics that surround its implementation. As Sandy Borthick wrote in a *BCR* article last fall, "QOS must function end-to-end to function at all, and the need for inter-vendor cooperation is particularly acute. Outside the enterprise network, it isn't just vendors who have to interoperate, it's the carriers and service providers, too." (See *BCR*, October 2001, p. 37).

But so far, few carriers have redesigned their networks to offer QOS within a single domain, let alone across domain boundaries. The harsh reality is that it's next to impossible to deliver predictable, end-to-end services over the Internet and across domain boundaries, whether a carrier uses DiffServ code points (DSCP) with normal IP packets, or MPLS using the experimental (EXP) bits to represent the different DSCP.

This is because inter-domain routing shifts the control point of traffic from one autonomous system to another. This has given way to a new layer of the OSI model, which I call the Political Layer. To understand how the Political Layer works, consider Border Gateway Protocol (BGP).

BGP allows each provider or enterprise to control its own policies on how routes are handled through its domain, and to announce changes to other autonomous systems as they occur. Associated BGP policies also let providers influence other peers on the best way to reach their destination networks inside a given peer's domain.

But when a company attempts to use inter-domain policies to influence routes into another domain, the additional policy information is typically stripped off the routes right at the entrance of the second domain. Even with new technologies, such as MPLS, DiffServ or the other QOS schemes, providers won't allow other entities to control traffic flows across their domain.

In a report on Internet backbones, Michael Kende, director of Internet Policy for the FCC's Office of Plans and Policy, asserted that "if backbones are unable to overcome the economic, administrative and technical hurdles to interconnect to exchange traffic flowing from new services requiring QOS, then the Internet faces the risk of balkanization." He's right. (See Kende's report, "The Digital Handshake: Connecting Internet Backbones," September, 2000 at www.fcc.gov/Bureaus/OPP/working_papers/oppwp32.pdf).

A Virtual Mess

VPNs are another area of IP sewage. The IETF's Provider Provisioned VPN (PPVPN) working group is looking at approaches that enable carriers to offer private VPNs to corporations— in other words, to find ways to circumvent the commoditized market for broadband Internet connectivity. IPSec VPNs, though complicated, expensive and riddled with many unresolved interoperability issues, is one approach.

But again, the issue of control becomes paramount. VPNs based on network-based IPSec require customers to give up control to the carrier. It's one thing for an enterprise to outsource a VPN, but quite another for it to outsource its firewall, often the main security point to the entire business. Many IT execs see the Internet as a place where script kiddies with nose rings hack their network and distribute the enterprise's most precious data, just for the thrill of it. Outsource their firewalls? You've got to be kidding.

Moreover, the cost of network-based IPSec services is high compared to CPE-based solutions that enable the enterprise to retain control. Again, carriers are in a Catch-22: Network-based IPSec equipment is expensive to purchase, operate and maintain, so the cost for network-based IPSec services is high. As a result, the service has not achieved the uptake the carriers had hoped.

An alternative is MPLS-based VPNs (RFC 2547bis). This approach is based on monolithic or virtual routers that use virtual routing and forwarding instantiations (VRFs) to separate VPN customers. Proponents claim MPLS-based VPNs can scale and provide service guarantees, they don't have the overhead and expense of encryption, and they use extensions to standards-

Do Routers Need Fixing?

The success of new IP technologies will depend on the foundation upon which they are implemented—IP edge routing. While the performance and functionality of this technology have evolved during the last two decades, the basic architecture hasn't changed.

From the smallest CPE to the largest core routers, today's routers are architected with a centralized, sometimes redundant, route processor. This architecture includes a single CPU and bank of memory, operating system and routing protocols that run on the system—each of which is a single point of failure.

New technology is evolving to address these problems, by allowing the processing of multiple services on separate routing engines. Each service can be isolated from others, so if a problem occurs within one environment or service, the potential for a "domino effect" can be controlled□

Back to the future: Dedicated, separate IP networks

based routing protocols on traditional edge routers.

But there has been debate around just how much scalability RFC 2547bis really can deliver (see *BCR*, December 2001, p. 39). Central to this debate is the fact that all enterprise routing tables must be redistributed into the provider edge (PE) router. This put massive pressure on conventional edge router architecture—as more customers come on to a MPLS-VPN service, the requirement grows for more memory in the MPLS-enabled PE routers. Today, these routers have a finite amount of physical memory in both the control and data plane, so solving this problem will involve a completely new approach to edge-router architecture.

Another alternative is to construct Layer 2 VPNs across a carrier network. This is also being proposed in the IETF under the name Virtual Private LAN Service (VPLS). Instead of a carrier redistributing the routes from each corporation into the PE router, why not offer a transparent Layer 2 LAN service across the carrier network?

While driven by noble intentions, this would return us to the days of bridged services, repackaged with MPLS. Enterprises would confront a full mesh of LSPs, similar to the full-meshed ATM networks of a few years ago. Just when you thought spanning tree had left the building—replaced with full-duplex Ethernet—the old arguments for determining which paths are forwarding versus blocking will be heard once again. The vicious circle continues. Even members of the PPVPN working group are asking questions like: "How do you plan to handle multicast?" and "This looks a lot like LAN Emulation (LANE) from a couple years ago. Didn't we learn our lesson?"

This is not to say that VPLS is a completely bad idea. It has a place in some metro applications. But how many metro Ethernet networks are going to survive in the current market? Today, Layer 2 services can be offered over an IP network without requiring MPLS as the transport. But you won't hear that from the vendors pushing VPLS.

What Not to Do

Given these IP sewage problems, is a "Roto-Rooter" waiting? Perhaps, but it won't be easy. Many new IP technologies have been created to help make the Internet more predictable, but they create at least as many problems as they solve.

So, maybe we're approaching the problem from the wrong perspective. Instead of creating more technologies to ride on top of IP—which turns IP into something it's not—let's go back to basics: Construct dedicated, separate IP networks. This approach would minimize IP sewage—sadly, it wouldn't eliminate it—and restore what we all love about IP: its simplicity.

One thing is clear: Trying to turn IP networks into a pseudo-ATM network *is not* the answer. As we learned the first time ATM came around, that path creates complexity, cost and performance problems that will take years to solve□