The New Edge Battle

Net Forecasts – Peter J. Sevcik
BCR Volume 29, Number 11
November 1999

As the Internet grows in size and function, its complexity will grow as well. Service providers, equipment vendors and end users will face decisions about how to deal with this complexity, and the major issue will be which parts of the network should be "smart" and which "dumb." There are three major approaches to where the intelligence should reside:

End station (desktops, servers)--This is the extreme case of the "dumb network" philosophy, and it's actually the way that much of the Internet operates today to deliver Web, email and other services. The goal is to make the best use of otherwise limited-function access and core networks; toward that end, the transport network supplies best-effort service from IP-address to IP-address.

Network core--This entails simplified end stations and access networks, and is the traditional "telephone" model that presumes users and access technicians cannot cope with the complexity of data networking. It is expected that personal digital assistants (PDAs), network appliances and other very simple end stations will require such a network model.

Access edge--In this model--which many network service providers (NSPs) and vendors are working toward--both the end stations and core are kept simple, and functional complexity resides somewhere between the two. The problem is exactly where in this space is the best place for the intelligence to reside: On the CPE, at the first service provider central office (CO), both? The network edge can be quite broad.

I believe that all three models will succeed as substantial markets over the next decade. Traffic patterns user preferences and economic drivers will not be homogenous, and so there will be opportunities for approaches in each of the three areas to find a market.

However, a gold rush is under way among vendors favoring the third model above. Communications Network Architects predicts that the market for IP access devices will reach $8 billion by 2003. This is turning into a battle over alternative ways to connect small businesses and branch offices to the Internet. The IP and broadband revolutions are opening up this market to new challenges and new entrants. IP-based office connectivity will likely impact the frame relay business much as frame relay affected the private line business. The stakes are high, but the rewards will be tremendous.

More Edge Services

The edge has traditionally been where local area network (LAN) technology must be adapted to wide area network (WAN) technology. This adaptation must account for the service differences between the LAN and WAN communications media at Layer 2, plus a multitude of functions supported at Layer 3, and maybe even higher up the protocol stack. As the number of alternatives at each layer increases, the adaptation problem is getting more complex.

There have always been legacy protocols or non-data applications that need to be integrated onto the limited-bandwidth access line, and the usual approach has been to use multiprotocol access devices that only try to match up services at Layer 2, leaving Layer 3 alone. The best current examples are frame relay access devices (FRADs); these devices typically keep things simple and transparent (they do not try to examine the traffic at higher levels, just pass it on in PVCs --sometimes even mixing voice and data).

However, even in the supposedly settled area of Layer 2, the number of technologies keeps climbing on both the LAN and the WAN. Even a standard as old as Ethernet has changed twice in the past few years, to accommodate 802.1Q and Gigabit Ethernet. More significant growth is coming in the WAN with PPP, PPP over Ethernet, L2TP, frame relay, SONET, ATM, Packet over SONET, various point-to-point SONET alternatives and the emerging lambda services (see BCR, August 1999, pp. 10-11).

The story is not much better at Layer 3, where the world has standardized a single protocol, IP. But it's not that simple: For IP to cope with the vastly different requirements of data, voice and video, and for it to scale as will be required, new services will

have to be added. There is a renaissance in defining value-added IP services based on technologies including IP Security (IPsec), MPLS tagging, RSVP, TOS, DiffServ, various forms of traffic shaping and QOS policy functions such as COPS (Common Open Policy Service). Additional services may need to be performed, such as firewalling, intrusion detection, bandwidth management, and H.323 gateway functions. The list appears to be growing by the month!

The LAN-WAN meeting place--the public-private edge--is traditionally defined as the last-mile connection to the customer site. However, debate has begun over where each of the above-listed IP value-added functions should best be performed: At the customer premises (CPE) or the central office (CO).

CPE Access

Some functions clearly are better performed at the CPE side of the access line, including rate shaping, bandwidth management and QOS policy enforcement. These functions need to be performed before the traffic leaves the customer site, so that the network can handle it appropriately at all times. Waiting to apply these functions until the traffic hits the CO leaves open the possibility of problems caused by unmanaged congestion on site or in the "first mile," between the site and the CO. In addition to QOS matters, customers would probably rather perform functions like encryption and other security measures on their premises.

Examples of CPE products with IP value added services are Abatis (www.abatis-sys.com), SBE (www.packeteyes.com), Xedia (www.xedia.com), Top Layer (www.toplayer.com), and Packeteer (www.packeteer.com). The PacketEyes product is the most integrated product: it comes with a built-in hub, router and WAN interface (ISDN, T1, DSL, FR). In addition, it performs bandwidth management, security, policy management, and system management tools (SMNP, COPS, LDAP), with a browser interface; all for a list price of less than $3,000!

The TollBridge TB50 (www.tollbridgetech.com) performs most of the same data functions on a single Ethernet port (add your own hub) but in addition it supports 24 analog voice lines (yes basic analog, use any phone you have) with toll quality voice over the single access line for a mere $2,000! Furthermore, the TB50 is designed for a telephone installer, who does not even have to enter an IP address to make the data functions work. This is a good example of CO device self-configuration and NSP-side control that proves you can perform a lot if IP functions at the CPE without needing a network engineer at each site.

CO Aggregation

The alternative view is that value-added IP services are best handled by the NSP at an aggregation point where the cost of the functions can be amortized over many subscribers. Also, since even the largest service providers wouldn't be able to install and manage such complex devices at each customer location, it is better to perform these functions in a centralized location for many customers at a time--think Centrex for data. Furthermore, if the traffic flow is predominantly from the network to the site, as in Web traffic, then it is more important to ensure QOS from the network to the customer over the limited-capacity access line.

A good argument that these vendors make is the need to connect (or multiplex) the site or even individual users to multiple WAN services. For example, a branch office may need some traffic delivered to headquarters via frame relay, while its general Internet traffic needs to be delivered to the client's ISP (e.g., UUnet, GTE). Since there is this need to perform basic service mapping at the CO, why not just do it all and supply the IP value added services as well?

Vendors in this camp range from those trying to provide basic service mapping such as Redback (www.Redback.com), Cisco 6400 (www.cisco.com), and Alcatel-Assured Access (www.ind.alcatel.com), to very complex value add IP switching and processing engines such as Spring Tide (www.springtidenet.com), Nortel-Shasta (www62.nortelnetworks.com) and CoSine (www.cosinecom.com). CoSine has the most ambitious plans to build an "Class 5" IP services engine for the CO. They have an impressive number of compute processors integrated with a switch in order to take on the most number IP value-add features.

Conclusion

Both approaches will have market success. However, I think that the CPE approach has strong advantages.

The two competing camps are indistinguishable from the point of view of the customer site users and local administrator. Both approaches will provide Web-based management and control of the services. In fact, since even the CO-based approaches give sites a great deal of control, they weaken their case that "it can't be done at the site."

Another argument that the CO-centric vendors make is that this is the only way in which service providers can get added revenues for added services. It is a familiar Centrex-like argument. However, the CPE-centric vendors argue just as compellingly that value added services can be provisioned at the customer premises by the NSP. In the end, I think that this argument will also fade into no distinction.

The CPE vendors will be able to create many targeted niche solutions that are out of reach for the CO-based one-size fits all approach. There simply is more flexibility on what you do and how you introduce new functions at the small-device level.

The distributed approach always has an advantage when it comes to localized functions. The CO vendors are walking into a processing nightmare; for example, imagine the processing requirements to encrypt thousands of individual data streams at 1 Mbps each. The site-based box only has to support tens of such streams or less, and if you push the job to the desktop, then it only requires a fraction of the desktop's CPU to perform the job. A thousand processors working at their own pace always win.

The correct answer to this debate is that there is a rational need for devices on both sides of the edge. Spring Tide has the more enlightened view of focusing on multiplexing traffic at the CO while they partner with CPE vendors that will focus on high-touch shaping of the traffic at the site.

There are simply more arguments for performing value-added IP functions closer to the user and at the point of maximum bandwidth discord (LAN to access line). So the CPE vendors have good technical arguments along with good economics in their favor. CO-based vendors should stick to scalable service mapping and multiplexing, with the assumption that the site did all the IP tricks before the packet got to the CO.

Peter Sevcik is President of NetForecast in Waltham, MA, and is a leading authority on Internet traffic, performance and technology. He has contributed to the design of more than 100 networks, and led the project that divided the Arpanet into multiple networks in 1984, which was the beginning of today's Internet. He can be reached at peter@netforecast.com.