Industry Background
The Demand for High-End Networking

Data Center Consolidation and Virtualization

Data centers are the nerve center of the enterprise and hold the majority of corporate strategic information technology assets. Data center performance, reliability, and security directly affects enterprise business and financial performance because these centers house a company's customer, engineering, and back-office applications and data, as well as its computational power, website, and storage. As such, the data center is a critical part of the enterprise switching and routing market. According to industry analyst group IDC, the worldwide data center network market will grow to $10.2 billion by 2011, up from $7.9 billion in 2007.

Many IT organizations are working to solve the problems created by the explosive expansion of IT infrastructure beginning in the late 1990's. In many instances this rapid expansion to meet business unit demand for IT resulted in highly complex infrastructures that offered sub-optimal reliability and security while being difficult to manage and expensive to operate. Driven to efficiently manage IT budgets, corporations are consolidating small and medium-sized data centers into larger data centers. A single consolidated data center is easier and less costly to manage, operate, backup, and secure than multiple small data centers. In addition, capital expenditures for consolidated data centers are lower due to centralized resources, increased asset utilization, and reduced cost per user.

Along with consolidation, we believe the evolution of data centers is currently focused on three technology areas: **infrastructure virtualization**, **cluster computing**, and **grid computing**. These concepts all share the notion that enterprise applications should be able to tap a common pool of computing resources (servers, storage, and networking) as if they were a single large virtual system. Depending on the application, the resources might be located in a single data center or distributed across a number of data centers – or spread throughout the entire network. For all of these technologies, the single system view is accomplished by interposing a middleware layer between applications and the resource pool.

Infrastructure Virtualization software is intended to de-couple existing enterprise applications and software services from underlying infrastructure by virtualizing server processing, storage, and networking. This de-coupling allows hardware management to be completely separated from software management, with the hardware allocated and re-allocated to various software services on the fly via a management console that provides the required resource allocation, load balancing, and monitoring of resource utilization.

Cluster Computing. There are essentially two types of computer clusters: load-balancing clusters and computational clusters. Both types of cluster are based on software that makes an array of homogeneous servers appear as a single virtual system.

Load balancing clusters provide high availability and scalable performance by distributing client connections among the application servers in the cluster. Various load balancing algorithms may be used to distribute the client sessions to achieve relatively equal distribution of the workload. Because each client session is independent of the others, the servers can operate in parallel, but there is no requirement for inter-process communications (IPC) among the servers or modification of existing applications to run on the cluster.

By linking hundreds, thousands – or even tens of thousands – of inexpensive servers together, enterprises can build computational capabilities that were previously only available in the largest and most expensive mainframes, yet at a fraction of the cost. These computational clusters offer an entirely different dimension of virtualization in which a tightly coupled cluster of computers provides supercomputer-like power in executing applications specifically programmed to run on parallel computers.

Grid Computing is taking shape as a general purpose distributed computing model where heterogeneous systems on an intranet or the Internet can publish services (computational services, data services, or other types of services) that may be accessed and utilized by other systems participating in the Grid. With Grid computing, computer systems and other resources aren't constrained to be dedicated to individual users or applications, but can be made available for dynamic pooling or sharing in a highly granular fashion to meet the changing needs of the organization. The vision of the Global Grid Forum (GGF) is that standards-based Grid middleware will allow Internet-wide resource sharing and collaborative problem solving by multi-institutional "virtual organizations."

As a result of these trends, data centers are placing new demands on network infrastructure, including **enhanced resiliency and reliability**, **higher density**, and **greater bandwidth**.

Enhanced Resiliency and Reliability. Because of the data center's critical role in the enterprise, network failure affects many facets of the business. According to Dataquest, network downtime can cost between $50,000 and $1,000,000 per hour for larger enterprises. Since avoiding network failure is of paramount importance, the resiliency and reliability of the network is critical. Resiliency and reliability refer to the ability of network equipment to avoid and recover from attack and failure.

Higher Density. In a high performance data center, every one of the hundreds, thousands (or tens of thousands) of servers must be able to communicate with every other server and with the external network. To accomplish this, each server must connect to an Ethernet switch port. In order to accommodate the increasing number of servers in a cost-effective manner, Ethernet switches must have many switch ports in a single chassis. Density, in the networking context, is a measure of the number of switch ports that are contained in a single system.

Greater Bandwidth. Since each server has the computational power to process and send Gigabits of data per second onto the network, hundreds or thousands of servers can push hundreds or thousands of Gigabits of data onto the network simultaneously creating a tremendous need for bandwidth. Managing this traffic flow requires switches and routers with Terabit capacity and the ability to support hundreds or thousands of Gigabit ports and tens or hundreds of 10 Gigabit ports.

We believe as enterprise data centers move through consolidation phases toward next-generation architectures that increasingly leverage virtualization technologies, the importance of high performance Ethernet switch/routers will continue to grow. The high volume of Ethernet products continues to spur rapidly declining prices and a constant stream of enhancements and innovations including, 10 GbE WAN PHY, Ethernet MAN/WAN services, iWARP RDMA/TOE NICs, 10GBase-T for 10 GbE over twisted pair, and the next generation of the Ethernet bandwidth hierarchy at 40 or 100 Gbps.

Enterprise LANs

Wiring closets in the enterprise LAN are undergoing a number of significant changes. Perhaps the most important transition stems from the emergence of the enterprise IP network and IP/Ethernet LAN as the converged infrastructure for both data networking applications and real-time communications applications. This convergence is driven by the cost reductions achievable through network consolidation and the productivity gains that can be achieved through innovative linkages among communication and data applications, resulting in converged applications such as unified messaging. As application convergence gains further momentum, the LAN infrastructure must continue to evolve to support a widening range of real-time applications, including:

- IP telephony
- IP video conferencing (bidirectional or multi-directional)
- Wireless LAN mobility solutions
- IPTV (unidirectional)
- IP multimedia conferencing
- Instant messaging (IM) with presence determination

All of the above applications (except IM) require PoE technology to get power. While some of these applications are just beginning to appear on the LAN, IP telephony has entered a phase of mainstream adoption by enterprises of all sizes. IDC expects the number of VoIP phones to increase by 42 percent in 2008, and the Dell'Oro Group expects IP PBX revenues to exceed $7.5 billion in 2011. Also according to the Dell'Oro Group, more than 80 percent of new wiring closet ports installed in 2008 are estimated to be PoE-capable. These figures show that wiring closets are being prepared for VoIP phones and other devices that will be powered over the LAN.

As the breadth of applications supported by the LAN continues to diversify, the wiring closet will need to accommodate a wider range of attached devices, including:

- VoIP desk phones and video phones
- WiFi access points for laptops, PDAs, and wireless VoIP phones
- Digital surveillance cameras
- IP enabled badge readers

According to IDC, the three primary concerns of IT managers deploying VoIP are: network availability to meet expectations for "always on" communications applications, latency/jitter, and sufficient bandwidth to support the new application mix. As VoIP and other emerging applications and edge devices are deployed, the edge of the network will need to be capable of evolving to provide additional security, manageability, serviceability, and power provisioning functionality.

Another major change for the wiring closet will come when desktop PCs and other edge devices are transitioned to Gigabit Ethernet (GbE) connectivity. While there is no imminent "killer application" that absolutely requires GbE to the desktop, there are a number of potential benefits of increasing desktop bandwidth beyond 100 Mbps. In the near term, the primary benefit of GbE to the desktop will be improved response time for client/server and peer-to-peer data applications that involve transferring bursts of large blocks of data.

The response time advantage of GbE is now achievable due to improvements in 10 GbE intelligent NICs that allow servers and network attached storage (NAS) devices to support numerous simultaneous, wire-rate GbE client/server streams at low levels of CPU utilization. Beyond improved response time, GbE provides the headroom that can accommodate the diverse bandwidth and latency requirements of the widening range of applications and edge devices:

- VoIP: <100 Kbps per call, but tight latency/jitter requirements for small packets
- Video conferencing/telephony: <1 Mbps per endpoint, low latency required for variation for variable length packets
- Streaming video: ~1 Mbps for 1280 x 720 resolution at 30 frames/second, low latency variation for large packets
- IPTV: Up to 19 Mbps per channel for full HDTV with MPEG4 compression
- WiFi: Up to 540 Mbps is required for IEEE 802.11n

Given the significant changes expected in LANs, we endeavor to maximize the lifetime of investments being made in wiring closet infrastructure. Future-proof investments will be based on Ethernet switching platforms that have the functionality and flexibility to accommodate the application-rich environment.

Service providers

Service providers face a stark reality today. The demand for bandwidth has been growing geometrically. Streaming video, pod casts, peer-to-peer traffic – generated and viewed by users with desktops, laptops, and cell phones — are filling network pipes. At the same time, the price that service providers are able to charge for bandwidth has dropped precipitously, while the cost to create additional bandwidth – primarily consisting of traditional core routers – has remained substantially unchanged. Faced with falling prices and stagnant costs, service providers are seeking lower cost alternatives to the expensive Internet core routers that they have come to rely on in the past.

Service providers have long been aware of the fundamental cost advantage of Ethernet. However, two factors have impeded their adoption of Ethernet-based products: Historically, SONET/SDH-based Internet core routers were more resilient and reliable than their Ethernet-based counterparts; and, in each concurrent generation of Ethernet and SONET/SDH technology, SONET/SDH has historically maintained a substantial performance advantage.

Recently, 10 Gigabit Ethernet achieved parity with the currently available SONET/SDH speed, OC-192. In addition, future Ethernet-based technologies are anticipated to meet or exceed the rate of speed available in future generations of SONET/SDH technologies such as OC-768 SONET/SDH technology, which is only recently starting to be deployed. The remaining challenge for Ethernet-based product acceptance in service provider networks is to achieve resiliency and reliability on par with SONET/SDH-based core routers. Over the past three years, our products have demonstrated levels of reliability and resiliency that meet and beat these traditional routers in the most demanding network environments.

Core alternatives. High-bandwidth, best-effort content is driving the need for simple, high-capacity, low-cost networks. In light of these market realities, many carriers are considering the merits of a 10 GbE-based IP routed core.

In the face of falling profit margins and rising traffic volumes, service providers are finding it difficult to justify the expense of their SONET/SDH networks, particularly if MPLS is enabled. With traffic growing exponentially, Ethernet offers a higher density, more scalable, cost-effective alternative to SONET/SDH. In addition, basic IP core switch/routers are significantly less expensive than comparable MPLS-enabled devices. MPLS (Multi protocol label switching) is a means to create a pseudo wire or circuit-like model within an IP network.

Service providers have various architectural options for accommodating best-effort traffic. Building cores that support specific traffic types is one option that allows service providers to leverage their "premium" MPLS-based cores and offload wholesale traffic onto an Ethernet-based IP-routed core. Another alternative is to have a service-rich edge feeding into a fast IP core. Some service providers are rethinking their core strategies entirely.

Best-effort traffic will continue to grow and represent the dominant traffic type for the foreseeable future. To remain profitable, service providers need to trim their network costs to match the economics of moving wholesale traffic. The combination of 10 GbE and IP routing gives service providers a cost-effective, high-bandwidth alternative to an MPLS-based network core – and a way to eliminate further surprises on the way to converged networks.

End-user demand

Web 2.0, user-generated content, and social networking have forever changed the role of the network. While this transition is well under way on the consumer side, it is only now beginning to impact the enterprise. As college students, among the largest users of Web 2.0 technologies, graduate and enter the workforce, they bring with them new communications habits and tools. People who use Facebook as a way to build social networks will want a similar experience in the work force. Before Facebook, the same transition occurred with instant messenger and mobile phones. Today, people increasingly communicate through Web 2.0.

As the new generation of employees begins to create content and applications within the workforce, we believe enterprise IT will be forced to grapple with a host of new network challenges. The more significant consumer technologies that will impact network design include: File sharing, instant messaging (IM), and many voice over IP (VoIP) services and products work via peer-to-peer applications. As organizations orient business processes around collaborative technologies, peer-to-peer applications and traffic will continue to rise. For network managers, this means an increase in unpredictable traffic patterns as more traffic moves between end users rather than from the end user to the data center.

Application Delivery

Enterprises are shifting how new applications are deployed and provisioned to their users using software-as-a-service (SaaS). SaaS is a software application delivery model where enterprises rely on web-based or Internet-based software applications that operate and host the application online and provide services to customers over the network. The concept of hosting critical applications is nothing new; however, software as a service is the first successful implementation of the model. Nearly every software company has a strategy to deliver its software as an Internet-based service. Like peer-to-peer technologies and wikis, SaaS applications make the wiring closet more important and increase the unpredictability of traffic on the network by enabling users to bypass the data center. According to IDC, the key characteristics of SaaS include:

- One-to-many application model (single instance, multi-tenant architecture)
- Network-based access to commercially available software
- Managed from central locations rather than at each user site
- Centralized feature updating, which obviates the need for downloadable patches and upgrades

SaaS architectures have new and significant impact on the enterprise wiring closet. Primarily, workers' productivity with remote, hosted applications is specifically tied to network uptime. In fact, many of these services – such as hosted CRM solutions – have limited to zero value to workers when those workers cannot access and maintain a reliable connection to the Internet. Whereas user productivity could be guaranteed with reliable client applications in the past, this blending of data center core to server edge reliability is considered a necessity in the enterprise for SaaS.

Many companies are going beyond point SaaS solutions delivered as services and are now using virtualization tools like Virtual Desktop Infrastructure (VDI) from Citrix and other virtualization vendors. Virtualization of client services is blurring the lines between the traditional data center compute "center" and the virtualized "campus," wherein all remote and edge clients are consuming data center services for their desktop "platform." Whether a business is using one service, or committing to a more complete virtual client infrastructure, these solutions make similar demands on the campus LAN. Specifically, they require:

- Massive bandwidth that ensures no network congestion
- High availability through full suite of NIC teaming capabilities
- Network simplicity and agility to deploy new services without complex network traffic engineering
- ***Complete reliability, which is paramount***