The Mandate to Re-Engineer Enterprise Routing to Support Today's Economy

By Jim Metzler, Cofounder, Webtorials Editorial/Analyst Division

Introduction

As economic conditions and the business environment continue to undergo changes, enterprises are adjusting their business strategies to maximize competitiveness, productivity, and profitability. Some of the most commonly cited business imperatives in the current business environment are:

- Containing cost
- Extending competitive advantages
- Maximizing enterprise application availability across the enterprise
- Assuring security, while providing adequate documentation of compliance with government and industry regulations
- Improving business agility
- Accommodating organizational changes, such as mergers and acquisitions

There are two overlapping ways to look at the relationship between a company's business imperatives and its technical strategies. In one of these views, as business imperatives are identified, IT departments must develop technical strategies and initiatives to support the business requirements. In the other view, an IT organization's ability to get the support it needs to upgrade the IT infrastructure is tightly linked to its ability to relate those upgrades to their business benefits. Both of these views are valid and both are supported by a Gartner Executive Programs survey of 1,500 organizations that was conducted in 2008. The results of the survey indicate that the vast majority of senior business managers expect their IT organization to be a key difference-maker in terms of their ability to execute their business strategies.

The arrows in **Figure 1** illustrate the fact that in most enterprises, business imperatives drive technical strategies. Figure 1 also depicts how the business imperatives listed drive a number of technical strategies. One of the goals of this paper is to provide an overview of the technical strategies shown in Figure 1.

Most IT organizations are continually refining their technical strategies in order to ensure that these strategies adequately support the company's business imperatives. While refining their strategies, the IT organization also needs to assess their existing network infrastructure, including their routers, to determine if it has the features and functionality that are necessary to support the emerging business requirements and the corresponding IT strategies. Any gaps in functionality that are identified can be addressed either at the time of the next technology refresh cycle or, in some more timely cases, while performing a re-design of the network infrastructure.

Figure 1: Relationships Between Business Imperatives, Technology Initiatives, and Router Functionality

Business Imperatives	Technical Strategies	Router Functionality
Maximize Application Availability	Centralization and Consolidation	Performance/Security High Availability
Competitive Advantage	Virtualization: Servers, Storage, Desktops, Networks	Power Efficiency
Cost Containment	Optimized Power/Cooling	Single O/S 2-Tier Architectures
Improve Business Agility	Reduced Network Complexity Improved Operational Efficiency	Future Proof Application Functionality
Ensure Security/ Compliance	New Applications/Archs (SOA, Web 2.0)	MPLS
Accomodate Org Changes (Mergers, Acquisitions)	Improved Network Flexibility	Multicast

Figure 1 also depicts the fact that technical strategies drive the need for functionality that routers must have in order to support those strategies. Another goal of this paper is to describe the router functionality that is listed in Figure 1 as well as the linkages between the technical strategies and the router functionality.

In order to support the goals of this paper, three IT professionals were interviewed and were encouraged to be candid with their answers. The names and organizations have been withheld as a result. Those professionals were the manager of network architecture for a major university, the VP and chief architect for a major financial institution, and a senior scientist in the US government's office of advanced scientific computing research. Throughout this white paper the interviewees will be referred to as The Architecture Manager, The Architecture VP, and The Senior Scientist. As will be shown in this white paper, each of the three interviewees is interested in enhanced routing functionality both to support some of the business imperatives that were previously discussed, as well as to support some challenges that are unique to their particular industry.

Current and Emerging Technology Initiatives

As shown in Figure 1, some of highest priority technology initiatives include: centralization and consolidation; virtualization of servers, storage, desktops and networks; optimized power and cooling; reduced network complexity and improved operational efficiency; deployment of new applications and application architectures; and improved network flexibility. The Architecture Manager pointed out the relevance of Figure 1 when he said that within his organization, "It is critical to link technical strategies with business strategies." The Architecture VP stated that from the perspective of the IT organization, that it is a top priority to ensure that technical strategies support business imperatives. He added, however, that this is typically not a high priority for the company's business unit managers. The fact that this is a high priority for IT organizations, but is not a high priority for business unit mangers, means that in most cases that IT organizations must anticipate changing business requirements and build the IT infrastructure to support them.

Centralization and Consolidation

The phrase *centralization and consolidation* refers to two related trends. One trend is to consolidate IT resources such as servers, applications and storage out of branch offices and to place these resources in centralized data centers. The second trend is to consolidate data centers. Centralization and consolidation are driven by the need to:

- Improve control over key information technology resources.
- Contain cost by eliminating the duplication of both the data resources and the IT staff that often reside at numerous sites.
- Improve application availability by placing resources in a more tightly managed environment.

Due to the economy of scale that results from centralizing IT resources, server consolidation tends to result in the need for fewer servers. In addition, the desire on the part of IT organizations to minimize the number of servers means that server consolidation initiatives are often done in conjunction with server virtualization initiatives.

As is discussed in the next section, IT organizations are showing a lot of interest in server, storage and desktop virtualization. Virtualization, however, is not a new topic. IT organizations have been implementing varying forms of network virtualization such as virtual private networks (VPNs) for over twenty years. For example, recent market research[1] indicates that almost two thirds of IT organizations have already implemented Multiprotocol Label Switching (MPLS). As will be expanded upon in a subsequent section of this white paper, MPLS was designed specifically to support virtual network overlays. As will also be discussed, some of the benefits of MPLS are that it is highly resilient and it provides high levels of security. MPLS also supports traffic engineering and fast reroute. Traffic engineering is important for several reasons. For example, once servers and storage have been consolidated into a centralized data center, branch office users access these resources over a wide area network (WAN). In order to ensure acceptable performance, the WAN must consist of services, such as MPLS, that can be engineered to provide low predictable WAN delay.

[1] The 2007 – 2008 MPLS Total Customer Experience, http://www.webtorials.com/abstracts/KubernanSOTM08-1.htm

Virtualization

In a survey that was completed in August 2008, 205 IT professionals were asked to indicate by the end of 2009 how much deployment their organization will have made of a number of forms of virtualization. Their answers are shown in **Table 1**.

Table 1: Anticipated Deployment of Virtualization

	None	Some	Moderate Amount	Significant Amount	Very Significant Amount
Server Virtualization	7.0%	27.6%	24.4%	22.4%	18.6%
Storage Virtualization	16.9%	19.7%	26.1%	22.5%	14.8%
Desktop Virtualization	34.7%	27.2%	21.8%	12.2%	4.1%

As previously noted, the vast majority of IT organizations already use network virtualization. The data in Table 1 indicates that IT organizations have a significant interest in deploying server and storage virtualization. While not as significant, IT organizations also have a strong interest in deploying desktop virtualization.

One of the basic motivations for server virtualization initiatives is to reduce costs by making more efficient usage of servers. For example, a virtualized server running multiple applications may operate at 60 - 70% CPU utilization instead of the typical 10 - 15% CPU utilization for servers dedicated to a single application. Server virtualization can therefore greatly reduce the number of physical servers required in the data center.

Storage virtualization provides a level of abstraction that allows a server to draw on a pool of storage resources as if it were a single storage device. Virtualization of storage can lead to improvements in storage efficiency comparable to those of server virtualization while also improving the manageability of storage devices.

Desktop virtualization refers to a situation whereby a user's entire desktop runs in a virtual machine that is resident on a virtual server in the data center. Desktop virtualization supports cost containment directives by streamlining the management and support of desktops and simplifying operations such as backup and restore. Desktop virtualization can also support higher availability via the deployment of redundant data center servers with automated fail-over amongst the virtual and physical servers.

Many of the initial deployments of desktop virtualization have been at sites that supported both a data center and end users. This approach has allowed IT organizations to leverage the existence of the high speed LAN at those sites in order to provide the high-capacity, low-latency connectivity between the end users and the virtual servers that is required to support desktop virtualization. However, as long as the WAN can be traffic engineered to have low predictable delay (e.g., MPLS), and the appropriate WAN optimization functionality has been deployed, desktop virtualization can be extended to users at remote sites.

Optimized Power and Cooling

Even if IT organizations have implemented centralization, consolidation and server virtualization initiatives, there is typically an ever-increasing demand for more data center capacity. Expansion of the data center is often limited by the capacity of the power and cooling systems. This means that centralization, consolidation and server virtualization initiatives must be accompanied by initiatives the goal of which is to minimize the consumption of power and cooling within the data center. One of the primary ways of achieving this goal is by deploying energy-efficient computing resources and network devices.

The Architecture VP pointed out that one of the factors that drives the ever-increasing demand for additional data center capacity is that most organizations are not able to decommission any applications. He added that another factor is that most servers support interfaces for a wide range of functions; i.e., storage, networking, management, etc. The fact that there are so many interfaces drives the need for additional networking links and additional ports on switches and routers. From a power and cooling perspective, “this becomes a pyramid that collapses on top of itself.”

Even if data center capacity is not an issue, the cost of power typically is. In particular, due to the escalating cost of energy, the power and cooling that is associated with data centers can account for a major portion of the IT budget. Controlling these expenses and maximizing the lifetime of the existing data center facilities requires a comprehensive approach to reviewing the thermal design of data centers and maximizing the power efficiency of every device located in the data center. Power efficiency metrics are being developed for servers, storage devices, switches and routers and other devices that are deployed both in data centers and in other parts of the IT infrastructure. Power efficiency metrics typically take the form:

Power Efficiency = units of useful work/watt

The power efficiency of servers is optimized by increasing utilization efficiency via server virtualization and by using hardware that delivers good performance per watt (e.g., peak and average SPECmarks per watt). Similar efficiency metrics are being developed for switches and routers (e.g., Gbps of throughput per watt) and storage devices (e.g., Gbytes per watt).

The Architecture Manager highlighted both of the drivers of optimizing power and cooling when he expressed his interest in implementing green initiatives in part to reduce the cost of power and in part because “the physical plant is near capacity”. He added that he was also concerned by the fact that the power requirements of many of the devices in his infrastructure are increasing and because “The concept of lights out management[2] is more prevalent in the server space than it is in the networking space”.

[2] Lights out management refers to the ability to either remotely manage devices such as servers or to manage them automatically. In either case, no people are required onsite to manage the devices.

Reduced Network Complexity and Improved Operational Efficiency

As new technologies are added to the network infrastructure, the complexity of the network grows due to proliferation of device types, device operating systems, management interfaces and applications. The added complexity directly translates to increased operational expenses due to the resulting additional workload and requisite technical training. Hence, another strategy in support of cost containment is to counteract the trend toward added complexity through programs that aim to consolidate and simplify the network infrastructure wherever this is practical.

The Architecture Manager stated that reducing complexity has been a focus of his for the last ten years. To reduce complexity, he has implemented a campus network based on a strict adherence to an architecture predicated on a layered architecture. To further reduce complexity, his organization strictly limits the number of disparate devices that are deployed within a given layer of the campus architecture.

Deployment of New Applications and New Application Architectures

The requirement for increased business agility is motivating IT organizations to adopt new application architectures that reduce the time it takes to implement a new application or a new business process. The most popular new application architectures include Service Oriented Architectures (SOAs), Software as a Service (SaaS), Cloud Computing, as well as Web 2.0 inspired techniques such as mashups. Each of these distributed application styles has the potential to significantly change the traditional network traffic patterns.

The requirement for increased business agility is also driving the deployment of applications designed to increase collaboration; e.g., traditional video conferencing and telepresence. The combination of the deployment of new application architectures and highly delay sensitive applications such as video conferencing and telepresence increases the requirement for the network to be flexible enough to accommodate new traffic patterns as well as additional classes of application traffic that require stringent levels of QoS.

The Architecture Manager stated that the faculty and students at his university are “eating up Web 2.0 applications” and that his organization is struggling to keep up with the impact of those applications. He added that the use of Web 2.0 applications combined with the university’s evolution to become more of a 7 x 24 operation has increased the university’s reliance on the network and has “caused significant pain when we fail to anticipate new requirements.”

The Senior Scientist stated that the experiments that they conduct produce huge volumes of data. He added that as the experiments become larger and most costly, that they perform fewer of them. The reduction in the number of experiments drives the need for scientists around the world to collaborate more closely on the analysis of the results of the experiments that are conducted. However, in addition to being voluminous, the data is also highly distributed and the analysis of this data occurs on hundreds of computers. The Senior Scientist emphasized that these factors drive a highly sophisticated workflow in terms of where the data is generated and stored, and in terms of the varying levels of analyses that are performed at research institutions and universities.

Improved Network Flexibility

The possibility of business mergers and acquisitions provides another impetus for improving network flexibility. For example, in the long run it is usually advantageous to rationalize the merged networks by migrating to a single network architecture. However, in most cases there is an immediate requirement to provide network connectivity to enable the appropriate flow of information between the merged organizations and to also provide at least some cost savings due to the consolidation of WAN services. If one of the merged networks has the flexibility to efficiently support some form of network overlays, it can serve as the foundation upon which to build both the required connectivity in the short term, and the necessary architectural rationalization in the longer term.

The Architecture VP stated that he believes that most IT organizations lack a framework for how to support mergers and acquisitions. He also stated that inside of highly regulated industries such as finance and health care that many IT organizations struggle with how to best segment traffic to ensure compliance with both government and industry regulations. Both the need to support mergers and acquisitions, as well as the need to segment traffic, has made The Architecture VP "a huge believer in enterprise MPLS."

Requisite Routing Functionality

As IT organizations plan and develop technology strategies such as those described above, they may need to re-assess the design and implementation of the current network infrastructure to ensure that it has the needed functionality and performance to support these strategies. To implement the technology strategies discussed in this white paper, data center class WAN/LAN routers should have the functionality depicted in **Table 2**.

Table 2: Required Routing Functionality

Requirement	Key Functionality
MPLS	Traffic Separation, Traffic Engineering, Fast Reroute
Carrier-class reliability	Redundant Hardware, In Service Software Upgrades, Non-stop Operations, Fast Reroute, Hitless Fail-Over, Hitless Process and Protocol Restarts
Security	Integrated Firewalls, Network Address Translation, IPSec
Reduced Complexity	Single Operating System, Support for a Two Tier Architecture
Power Efficiency	Validated Power Efficiency Ratings, Reduced Number of Interfaces
High Performance	Maximum Throughput, Throughput Unaffected by Turning on Features
Multicast	Support for Low Latency Multicast Applications

The rest of this section will provide additional detail on the requisite routing functionality.

Performance, Security, and High Availability

After implementing centralization and consolidation initiatives, the corporate data center becomes even more critical to business operations and hence must deliver the highest levels of availability and security, while accommodating much higher levels of traffic aggregation. The resulting requirements for data center routers include:

- Highly scalable performance that is not adversely affected when a wide array of features is enabled. For example, within the consolidated data center it may be necessary or desirable to enable a number of security features, such as access control lists, firewall feature sets, and traffic event logging. Additional router features and services that are frequently required include application recognition and QoS policy enforcement.

- High availability features including redundant hardware subsystems that can support non-stop operations in spite of component failures, and high availability network operating system software that supports features such as in-service software upgrades, hitless fail-over among redundant route processors, and hitless process and protocol restarts.

The Architecture VP stated that his organization has implemented a number of forms of virtualization and that the fundamental driver of these initiatives was cost optimization. He added that as a result of virtualization that their previous goal of five 9's of availability for their data centers is no longer acceptable and that their new goal is one hundred percent availability.

Power Efficiency

The power efficiency of a router depends on the system architecture and the power characteristics of the various components employed, which tend to differ significantly among different router models. Therefore, power efficiency can be a valid differentiator within a given class of routers and should be considered along with other product attributes and features that are normally factored into router selection. All other things being equal, preference should be given to the routers with the highest validated power efficiency ratings. In any case, router power efficiency should be carefully considered in the light of data center power budgets and total cost of ownership (TCO) calculations.

Single Operating System, Support for Two Tier Architectures

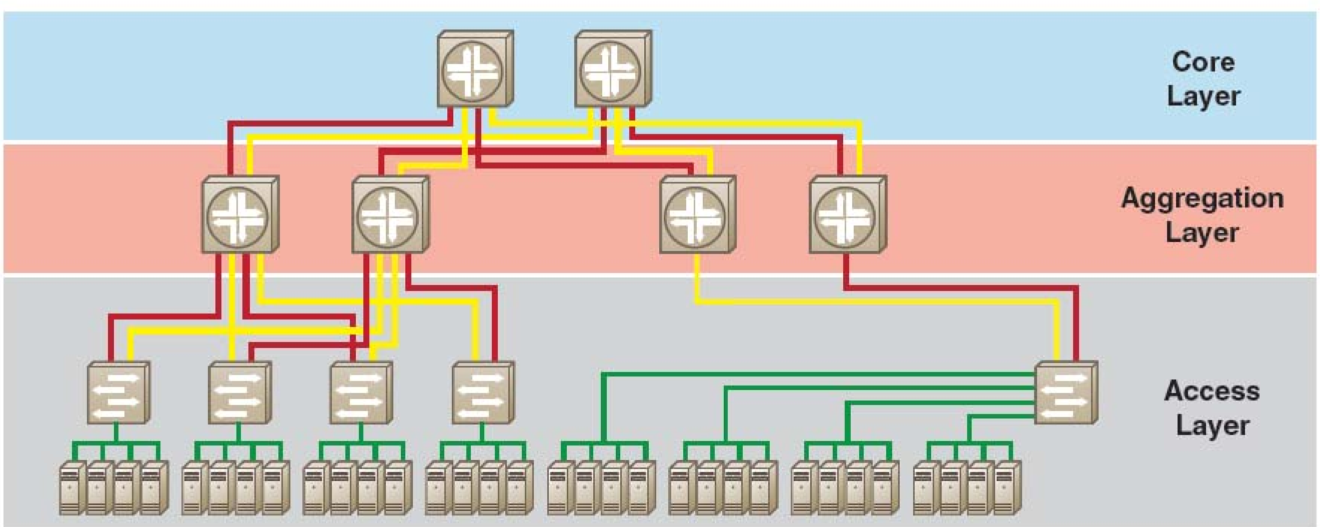
The Architecture Manager highlighted one of the generally acknowledged best practices for reducing complexity and simplifying network management. That being to minimize the number of different models of routers deployed. Another best practice is to minimize the number of different versions of network operating systems that are in use. Excess diversity among operating systems adds considerable complexity to a wide range of management tasks that are listed below.

The ultimate reduction in complexity can be achieved when all routers and switch/routers in the network share a single operating system that has been adapted to, or compiled for different hardware platforms from a common base of source code. Consolidation of the network operating system environment reduces complexity and the workload of management and administration in a number of ways:

- Simplified patch processes and version management
- Improved network reliability and predictability
- Fewer discontinuities in features and functions across network tiers
- Fewer interoperability problems among devices
- Less expertise required in navigating disparate user interfaces (e.,g, CLIs)
- Simplified troubleshooting and fault management
- Simpler recovery from security vulnerabilities and intrusions

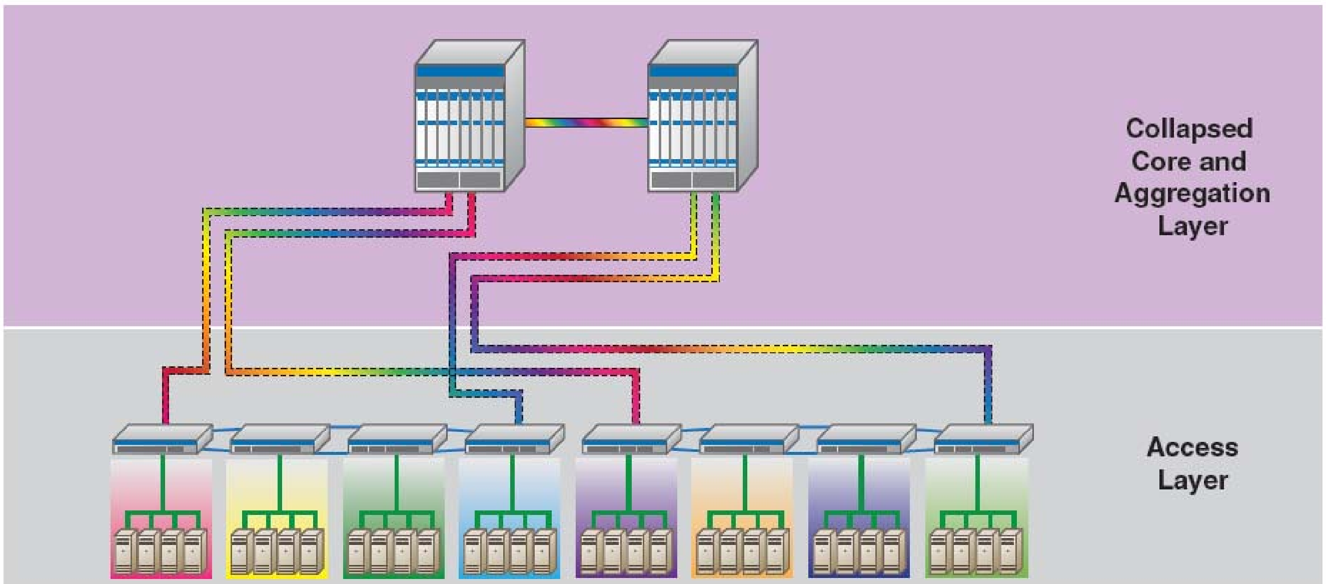
Complexity can be further reduced through router and switch consolidation based on highly scalable devices that have the performance and functionality to support a simpler data center architecture. **Figure 2a** depicts the traditional LAN architecture that is comprised of three tiers: access, distribution and core.

Figure 2a: Traditional 3-Tier Network Architecture



As shown in **Figure 2b**, the phrase 2-tier architecture refers to collapsing the traditional distribution and core tiers into a single tier supported by routers that provide high performance and high density as well as switching functionality. The complexity associated with a two tier data center architecture is much less than the complexity associated with either a three or four tier data center architecture. In particular, consolidating a number of low-density devices into a smaller number of high density devices reduces management complexity, power consumption, and the number of required interconnects.

Figure 2b: 2-Tier Network Architecture



Future Proof Application Functionality

The number of different Web-based applications traversing the enterprise network is growing rapidly due to the webification of enterprise applications such as ERP and CRM, plus the utilization of emerging Web-based application architectures such as SOA, SaaS, Cloud Computing, and mashups. This trend increases the importance of identifying which Web applications are business critical so that the IT organization can provide preferential treatment to these applications vs. the more mundane or recreational applications that are also Web-based; e.g., Internet radio. Routers that can base QoS scheduling and forwarding behavior on deep packet inspection (DPI) will be able to parse application headers allowing all critical business applications, including VoIP and videoconferencing, to receive preferential treatment and enabling recreational or unwanted application traffic to be either eliminated or rate limited. Implementing router-based QoS at key points of aggregation within the network may offer an attractive alternative to managing an end-to-end QoS scheme involving numerous client end systems.

IT organizations that are looking to future proof their router purchases should also insist that any router they purchase support IP Multicast. IP Multicast provides for the efficient use of WAN bandwidth by enabling the simultaneous delivery of content to large numbers of recipients dispersed throughout the network. Applications leveraging IP multicast include IPTV for corporate communications or distance learning, video conferencing, as well as the distribution of software, stock quotes, and news.

The Architecture VP stated that most major financial services firms are "racing to zero latency" and that these firms are currently measuring latency in microseconds. As a result, The Architecture VP believes that performance based routing will be a necessity on a going forward basis, for at least this segment of the market.

MPLS

As noted, MPLS has been specifically designed as a core network technology that can support VPNs. A VPN refers to overlaying multiple traffic streams on a common infrastructure in a way such that each traffic stream appears to be running over a private network. One way that these overlays can be used is to allow the network of one merger partner to serve as the host network supporting the network of the other merger partner as a client network. Because MPLS supports overlays of IP, ATM, Frame Relay, and Ethernet networks, the core of the merged network can quickly provide the connectivity and network consolidation that can enable IT organizations to reduce WAN service expenses significantly. With MPLS, there is no need for the merging networks to change their IP addressing schemes. This facilitates the rapid integration of the network's physical and data link layers.

Security is a top of mind issue for any IT organization that is concerned with deploying and managing a WAN in part because one of the WAN's central functions is to enable communications with customers, suppliers and distributors. Because it connects with entities outside of the enterprise, the WAN is a source of security vulnerabilities. In addition, today it is widely accepted that the majority of security incidents originate from within the enterprise. As such, in addition to keeping separate the traffic between an enterprise and its customers, suppliers, and distributors, the WAN must also provide separation between the communications of individual departments and work groups. One of the security mechanisms that is inherent in MPLS-based VPNs is traffic separation. In order to separate traffic, each MPLS-enabled VPN is assigned to a unique Virtual Routing and Forwarding (VRF) instance. Traffic destined for each VRF carries its own label value. As such, each VPN is kept logically and physically separate from every other VPN.

Unlike most WAN services, MPLS supports traffic engineering. Traffic engineering refers to the process of selecting the paths that the traffic will take as it transits through the network. Traffic engineering can be used to accomplish a number of goals. For example, a network organization could traffic engineer their network to ensure that none of the links or routers in the network are either over or under utilized. Alternatively, a network organization could use traffic engineering to control the path taken by voice packets in order to ensure appropriate levels of delay, jitter, and packet loss; or to ensure that the traffic between users in a branch office and the centralized data centers has low delay; or to enable desktop virtualization to be deployed to branch office employees.

MPLS-based traffic engineering also supports the rerouting of traffic around a failed link or router quickly enough so as to not adversely affect the users of the network. To achieve this fast restoration time, a backup path can be established at each node. The fail-over mechanisms are triggered by physical link or routing events that indicate that the link or node is down. The traffic can be switched immediately to this backup path once the failure has been detected. MPLS with this fast re-route capability can re-route traffic in under 50 ms, which is similar to SDH and SONET networks that carry the public telephony network.

The Senior Scientist stated that the data-intensive, highly distributed workflow that his organization needs to support requires stringent guarantees from the network. Without these guarantees, network congestion occurs which "interrupts the workflow and ripples through the entire system". The Senior Scientist pointed out that they have adopted MPLS in large part because it allows them to do traffic engineering. He added that in addition to bandwidth guarantees, that their network needed the ability to separate diverse traffic types (e.g., best effort, network control, etc.) and also needed an effective reservation system that would only accept a reservation if there was enough capacity to support the request.

Summary

As IT plays an increasingly important role in the execution of enterprise business strategies, IT executives will need to place greater emphasis on developing technology strategies and initiatives that are tightly linked to, and highly supportive of business requirements. However, as emphasized by both The Architecture VP and The Architecture Manager, in many cases the IT organizations will have to anticipate these requirements with little input from the business unit managers.

The agility and the flexibility of the network to respond to new business priorities is highly dependent on the functionality and capabilities of the fundamental network infrastructure. IT executives can solidify the strategic role of their network by ensuring that the infrastructure's most critical components, including data center class routers and switch/routers, are capable of supporting both current and emerging initiatives.

Because the infrastructure cannot be refreshed every time there is an adjustment in the business strategy, network designers need to provide some headroom in terms of both functionality and performance that anticipates possible future technology initiatives to the degree that is possible in today's rapidly changing business and technology environments. While it is not possible to predict with certainty the exact business and technology changes that will impact a given IT organization over the next year or two, it is possible to predict with certainty that change will occur. The router functionality discussed in this white paper is a key enabler to a wide range of technical strategies that will allow IT organizations to respond to these changes.

About the Webtorials® Editorial/Analyst Division

The Webtorials® Editorial/Analyst Division, a joint venture of industry veterans Steven Taylor and Jim Metzler, is devoted to performing in-depth analysis and research in focused areas such as Metro Ethernet and MPLS, as well as in areas that cross the traditional functional boundaries of IT, such as Unified Communications and Application Delivery. The Editorial/Analyst Division's focus is on providing actionable insight through custom research with a forward looking viewpoint. Through reports that examine industry dynamics from both a demand and a supply perspective, the firm educates the marketplace both on emerging trends and the role that IT products, services and processes play in responding to those trends.

For more information and for additional Webtorials® Editorial/Analyst Division products, please contact Jim Metzler at jim@webtorials.com or Steven Taylor at taylor@webtorials.com.

Webtorials Briefs
Vol 3, Number 5

Published by Webtorials Editorial/Analyst Division
www.Webtorials.com

Division Cofounders:
Jim Metzler
jim@webtorials.com
Steven Taylor
taylor@webtorials.com

Professional Opinions Disclaimer
All information presented and opinions expressed in this publication represent the current opinions of the author(s) based on professional judgment and best available information at the time of the presentation. Consequently, the information is subject to change, and no liability for advice presented is assumed. Ultimate responsibility for choice of appropriate solutions remains with the reader.

For editorial and sponsorship information, contact Jim Metzler or Steven Taylor. The Webtorials Editorial/Analyst Division is an analyst and consulting joint venture of Steven Taylor and Jim Metzler.