White Paper

Managing the WAN's Migration from Plumbing to Platform

Jim Metzler, Vice President
Ashton, Metzler & Associates
February 2007

Table of Contents

Introduction

During the dot com era, many enterprises[1] believed that Information Technology (IT) would allow them to fundamentally transform their business model to become more efficient, enter new markets and fend off competition. As a result, the typical IT organization was highly valued and IT budgets increased accordingly.

Shortly after the dot com era ended there was the belief that the IT infrastructure was like plumbing. By that is meant that it was recognized that an enterprise could not function without an IT infrastructure, the same as it could not operate an office building without the basic utilities: heat, electricity, or water. However, the belief at this time was that there was no linkage between an enterprise's IT infrastructure and their ability to achieve their business goals, the same as there was no linkage between how an enterprise used the basic utilities and their ability to achieve their business goals. As a result of this belief, the typical IT organization was not perceived as providing much business value and IT budgets reflected that perception.

Currently there are some fundamental shifts underway in the business environment. For example, enterprises are under constant pressure to become more agile. Enterprises are responding to this pressure by taking steps such as re-engineering their business processes, distributing key business functions and personnel, and placing more emphasis on customer service. To enable these shifts in the business environment the application environment is also changing. For example, many enterprises are consolidating servers out of branch offices, reducing the number of data centers, deploying applications such as Voice over IP (VoIP), and implementing more distributed applications.

Just as the application environment is changing in order to support the changing business environment, the typical enterprise network is changing to support the changing application environment. For example, at the end of the dot com era WAN design was not focused on adding additional functionality, but was focused on insuring that the WAN exhibited utility-like characteristics — such as being low cost and highly available. The point being that at that time the WAN, like the rest of the IT infrastructure, was regarded as plumbing.

That approach to WAN design is changing. Today WAN design has two distinct goals. One goal is to ensure that the WAN continues to exhibit utility-like characteristics. The other goal is to ensure that the WAN can also support the key business and application trends by adding sophisticated WAN functionality such as:

- QoS (Quality of Service)
- Network optimization techniques such as caching, compression and protocol optimization
- Security functionality such as firewalls as well as IDS (Intrusion Detection Systems) and IPS (Intrusion Protection Systems)
- Network Attached Storage (NAS) and Storage Area Networks (SAN)
- Server Load Balancers (SLB) and Application Front Ends (AFEs)

This white paper will describe some of the key business, application and network trends in general, and will also discuss how these trends impact certain key industries. The goal of this description is to create a framework that will be used to both identify the impact that these trends have on network management and to identify selection criteria that IT organizations can use to choose an appropriate network management solution.

The Changing Business Environment

Kubernan[2] recently conducted a survey that will be referred to in this white paper as The Kubernan Survey. As part of The Kubernan Survey, IT professionals were asked to indicate their enterprise's two most important business goals. The top three responses and the percentage of times that they were referenced are contained in Figure 1.

Business Goal	Percentage of Respondents
Improve the Customer's Experience	37%
Reduce Cost	26%
Streamline Business Processes	24%

Figure 1. Enterprise's Top Business Goals

One of the conclusions that can be drawn from Figure 1 is that unlike the situation immediately following 9/11, reducing cost is no longer the top business goal. By a wide margin, that distinction belongs to a goal that is often listed among the most important business goals — improving the customer's experience.

[1] Throughout this white paper the term enterprise will be used to refer to any type of organization, independent of size and industry.

[2] Kubernan is a consulting and analyst joint venture of Jim Metzler and Steven Taylor

Improving the customer's experience is important to all industries around the globe. For example, Gabriel Haerring of Helsinn Chemicals in Biasca, Switzerland pointed out the importance of customer service when he stated[3], "Soft factors, such as customer services are key to winning and keeping customers." He went on to say, "In chemistry terms, there is now no difference between firms from the U.S., Europe, China, or India, [As a result] quality and customer service are our key differentiators."

Customer service is particularly important in the financial services industry. As was recently reported[4], major bank chains in the United States are taking cues from retailers and emphasizing customer service. "Our customers are our most important asset and we strive to meet all their financial needs and provide exceptional customer service, "a spokesman for New York-based Citibank told Bank Marketing International in the fall of 2006. "To ensure that customer service quality remains high and that we quickly identify any issues, Citibank recently implemented a pilot program in 30 branches called 'How are we doing?' As part of this effort, customer feedback is solicited and responded to daily. "We have completed the pilot and plan to deploy the program nationwide."

The reason that streamlining business processes made the list contained in Figure 1 is because enterprises of all types are under constant competitive pressure to become increasingly agile, and one of the ways that they are becoming more agile is by streamlining their business processes. In this context, an enterprise is considered to be agile if it can respond in real time enough fashion to situations that impact the health and well being of the enterprise and its stakeholders. For example, the U.S. Department of Homeland Security has to respond to threats for which they receive little if any notice. Hence, for the U.S. Department of Homeland Security real time enough fashion means having to respond in a matter or minutes or hours. Large retailers need the ability to modify their supply chain in order to respond to unexpected demands for products at any of its stores. In this situation, real time enough fashion means having to respond in a matter of two or three days.

The Kubernan Survey determined that roughly two thirds of enterprises are currently re-engineering one or more business processes and that 45% of companies are re-engineering two or more business processes. Given the business importance of improving the customer experience, it is not surprising that the primary aspect of business that is being impacted by business process re-engineering is customer service.

A technique that enterprises are using to increase business agility is to distribute key business functions (i.e., R&D, Marketing, Manufacturing and Customer Service) around the U.S., and in some cases around the world. Another factor driving the distribution of business functions is that many enterprises want to have a presence close to their customers, suppliers and distributors. There are several reasons why having a local presence is advantageous for an enterprise. For example, enterprises doing business internationally typically find that having a local presence is driven by a combination of logistical and economic necessity, as well as legal requirements.

One of the primary reasons that enterprises are distributing the customer service function is that enterprises have begun to value their branch offices very differently than they did only a few short years ago. At that time, the conventional wisdom was that branch offices offered little business value and as a result were heading towards extinction. Today many companies have come to realize that branch offices provide them with a critical touch point with their customers. The vast majority of companies are looking to fully leverage this customer touch point by deploying business critical applications such as VoIP out to branch offices.

One of the results of the fact that enterprises are distributing their key business functions is that currently relatively few employees spend the majority of their time in a headquarters facility. In particular, the 80/20 rule that was in place up until a few years ago stated that 80% of an enterprise's employees were in a headquarters facility and accessed an application over a high-speed, low latency LAN. The new 80/20 rule states that 80% of an enterprise's employees access applications over a relatively low-speed, high latency WAN.

One of the business trends that have been underway for the last decade is a growing reliance on information to support critical business functions. For example, information enables:

- Retailers to know which items are selling quickly in which stores. This allows the retailer to adjust its supply chain to ensure delivery of those hot items to the stores that need them in a timely manner; i.e., just-in-time inventory.
- Enterprises to respond to myriad governmental regulations, such as the Sarbanes-Oxley Act, HIPAA, and the Gramm-Leach-Bliley (GLB) Act.

[3] Chemical Week April 5, 2006.

[4] Bank Marketing International. September 26, 2006

- Enterprises to develop and implement plans for business continuance and disaster recovery (BC/DR). While many enterprises implement BC/DR plans on a voluntary basis, some industry segments are required to implement these plans. For example, implementing a disaster recovery plan is a U.S. government mandate for the financial services sector[5].

Information also allows enterprises to achieve their goal of focusing on the customer. One way this happens is through customization — or the trend towards consumer-made products. As part of this trend, enterprises are designing products for the mass market that consumers can personalize to meet their needs or fulfill their desires. For example[6], consumers can design their own stamps through a website set up by the U.S. Postal Service, or they can customize their own soda labels with Jones Soda. They can also customize their own Starbucks cup, and select the color of their M&Ms. These capabilities build on the success of TiVo and iPod, which allow consumers to control the media they watch and the music they listen to.

The Oakland Athletics recently announced[7] their vision for a new baseball stadium to be built in Fremont, California. Personalization of the fan's experience is at the heart of this vision. For example, fans will be able buy to tickets on their cell phones and have access to game-day video on their handsets if they arrive late. Kiosks at the stadium will use networking to identify ticket holders and sell them upgrades or other products. In addition, location-based services are also a part of the plans which include wireless personal-area networks which will trigger digital signs to display ads tailored to the preferences of fans who pass near them. In the stands, fans will be able to order food and drink that would be brought to them, thanks to the location-based services.

The Oakland Athletics are not the only enterprise that intends to leverage information to up-sell and cross-sell products and services. Another example of this business trend is the bundling of data, voice and video services is proving to be critical in order for service providers to up-sell and cross-sell products which helps the service providers to both maximize the lifetime value of a customer and to reduce churn.

A recent article[8] that stated that companies that don't bundle services are finding themselves at a competitive disadvantage.The article pointed out that DirecTV of El Segundo, California, not only competes against EchoStar's Dish network for satellite customers, but also against cable companies and will soon be competing with phone companies that are offering television over fiber optic lines. The article concluded that aggressive competition from cable companies that bundle services is one of the reasons why DirecTV's subscriber base growth rate has been slowing.

Key Application Trends

Key application trends are described below:

Mainframe Computing

As will be discussed in section 3.4, many enterprises have begun to adopt an application architecture based on distributed computing. While servers play a major role in supporting distributed computing, so do mainframe computers. In particular, mainframe computers are often used to perform backend processing such as running large databases. Consequently, IBM has redesigned its mainframes to handle the latest technology. Another reason for the continued viability of the mainframe is that the machines are smaller and cheaper than ever before and still offer the most powerful and reliable computing environment available. To put this in perspective, an IBM executive recently stated[9] that 95% of the Fortune Global 500 use mainframes and that 80% of the world's business transactions are processed on them.

One enterprise[10] that has stuck by mainframes is Winnebago Industries Inc., the recreational vehicle maker. Winnebago wanted to cut costs by using packaged software on cheaper hardware. The enterprise's IT staff concluded, however, they were messing with a system that worked well. Winnebago kept its manufacturing, inventory, production scheduling and email applications on the mainframe.

[5] "Sound Practices to Strengthen Resilience of the US Financial System", Board of Governors of the Federal Reserve System Docket No. R-1128, Office of the Comptroller of the Currency Docket No. 02-13 and Securities and Exchange Commission Release No. 34-46432; File No. S7-32-02.

[6] Trend Spotting with Author Robyn Waters http://www.businessweek.com/magazine/content/06_48/b4011410.htm Nov 27, 2006

[7] Cisco, Oakland A's Sketch Out Field of Dreams http://www.eetimes.com/news/latest/showArticle.jhtml?articleID=194400271, Nov 15, 2006

[8] Associated Press Financial Wire, August 7, 2006, 4:09 PM GMT

[9] The Straits Times (Singapore) Nov. 7, 2006

[10] InformationWeek Oct 21, 2002 12:00 AM

Server Consolidation

In order to reduce cost, comply with regulations, and reduce the challenges associated with security, many companies are consolidating servers out of branch offices and are also reducing the number of data centers. In particular, consolidating servers into a centralized data center results in the need for fewer servers, due to the economy of scale that results from centralization. The reduction in the number of servers results in a lessening in the cost of the servers as well as the cost of the relevant software licenses, support and real estate. In addition, an IT organization can demonstrate that it is managing risk since one of the other primary factors driving server consolidation is responding to the ongoing enactment of government and industry regulations, such as the Sarbanes-Oxley Act (SOX). One common characteristic of the recent spate of regulations such as SOX, is the requirement for enterprises to put a greater emphasis on assuring the accuracy, security and confidentiality of data. It is very difficult to do this when there are multiple copies of the enterprise's data on servers in branch offices. Consolidating the servers into centralized data centers makes these tasks notably easier to accomplish.

Delay Sensitive Applications

Many companies have begun to deploy delay sensitive, business critical applications, such as SAP out to branch offices. Some of the SAP modules are notably delay sensitive; i.e., the Sales and Distribution (SD) module of SAP that is used for sales order entry. If the SD component is running slowly, an enterprise can compute the lost productivity of the enterprise's sales organization as they waste time waiting for the SD module to respond. In addition, if the SD module times out, this can irritate the customer to the point where they hang up, taking their business elsewhere.

Another example of a latency-sensitive, business-critical application that has been deployed in many branch offices is VoIP[11]. Some network organizations view VoIP deployment as just adding one more application to an IP network. While there certainly is some validity to that argument, there are also a number of very unique aspects of VoIP that distinguish it from more traditional data applications. One such aspect is that VoIP deployment places availability and performance requirements on the network that are far more demanding than those placed by the typical data application.

For example, a traditional data application, such as email, can accept a small amount of downtime. That is not true with voice. The vast majority of people expect that their voice call will go through every time they pick up the phone and dial a number. Another requirement that distinguishes VoIP from a more typical data application is the rigorous demands that voice places on the underlying IP network. These rigorous demands include the requirement to have exceedingly low levels of delay, jitter, and packet loss.

For many data applications, end-to-end delay is not a critical issue. That is not the case with voice. The International Telecommunication Union (ITU) recommends that the end-to-end delay associated with a voice call not exceed 150 milliseconds (ms). Experience has shown that it is possible to exceed that goal by a small amount. However, if the delay becomes too large, the quality of the voice call degrades noticeably.

Jitter is a measure of how packet delay changes over time. The vast majority of data applications are not impacted by jitter. Again, voice is different. Using the RFC 1889 definition of jitter, jitter should not exceed 30 ms or else the voice quality degrades noticeably.

Voice, as well as virtually all data applications, is sensitive to packet loss. However, in the case of data applications it is typical to re-transmit any packets that were lost. In a VoIP deployment, lost voice packets should not be re-transmitted. This follows in part because modern voice coding algorithms (CODECs) are explicitly designed to deal with the occasional loss of a packet. As such, it is better to let the CODEC do what it was designed for, than to re-transmit a voice packet and introduce significant delay and jitter.

Distributed Computing

In many instances, the phrase *distributed computing* is used interchangeably with the phrase *n-tier applications*. N-tier applications are applications in which the functionality that had at one time been provided by a monolithic application is now segmented into multiple tiers. Since these tiers are implemented on separate systems, n-tier applications typically make more demands on the network than do monolithic applications.

The first generation of n-tier applications was referred to as client-server applications. These 2-tier applications were structured around the use of a PC that accessed an off-the-shelf database product using a language such as Structured Query Language (SQL). The information flow in these applications went from the PC to the database and back again over an enterprise's private network.

[11] 2005/2006 VoIP State of the Market Report, Steven Taylor, www.webtorials.com

The success of client-server computing led to the deployment of more sophisticated application architectures, such as 3-tier applications. The typical 3-tier application is comprised of a Web browser, an application server(s) and a database server(s). The information flow in a 3-tier application travels in a linear fashion from the Web browser to the application server(s) and to the database and then back again over the Internet using standard protocols such as HTTP. This form of application has become quite popular despite the addition of significant traffic to the WAN. In particular, applications that use HTTP consume five to ten times as much WAN capacity as applications that use a thin client.

The next step in the development of distributed computing is the deployment of Web services-based applications. Web services refer to reusable software modules that encapsulate business functionality and which can be accessed over an IP-based network. One of the primary reasons why Web services-based applications have become so popular is that they enable business agility. In a Web services-based environment, it is easy to create a service that is a network-enabled component representing business functionality that can be easily reused. As a result, IT organizations that adopt a Web services approach to applications development can modify business processes such as supply chain management and customer relationship management in a fraction of the time and for a fraction of the cost than would be possible in a more traditional application environment.

However, Web services-based applications demand a highly available, highly functional WAN in part because in most cases the servers that run the Web services are housed within:

- Multiple data centers owned by a given organization, or
- Multiple data centers owned by different entities.

These application deployment options place tremendous demands on the WAN infrastructure for both security and performance. Applications that would have previously executed within a single data center now execute in multiple data centers. This results in a very significant amount of additional traffic placed on the WAN. For example, Kobielus[12] estimates that XML requires 2 to 100 times more bits than non-XML binary transfers. However, in spite of this additional traffic and the latency that is inherent in a WAN, these Web services-based applications are expected to perform as if they were running over a LAN in a single data center.

From Plumbing to Service Delivery Platform

As mentioned in the introduction, until a few years ago WANs were designed primarily to provide reliable, low cost connectivity. This is sometimes referred to as the plumbing or utility approach to WAN design. This approach certainly describes the TDM (Time Division Multiplexing) networks of the late 1980s and the Frame Relay and ATM networks of the early to mid 1990s.

Many enterprises are changing their approach to WAN design. In particular, while it is still important that WANs provide reliable, low cost connectivity, WANs are also being designed as service delivery platforms. By service delivery platform is meant that sophisticated functionality is being added to the WAN so that the WAN can support the business and application trends that were discussed in the preceding sections. The rest of this section will highlight some of the functionality that is being added to the enterprise WAN.

QoS (Quality of Service)

QoS refers to the ability of the network to implement policies that ensure that preferential treatment is given to certain classes of traffic. QoS is required in those situations which bandwidth is scarce and there are one or more delay sensitive, business critical applications. Note that in this context, preferential treatment could mean limiting the bandwidth that certain applications (i.e., email or Internet Radio) receive while simultaneously ensuring fairness for all the users of some business critical, delay sensitive application such as VoIP or SAP. In some cases, enterprises implement QoS via queuing functionality in their routers. In other instances, enterprises implement QoS by using an MPLS (Multi-Protocol Label Switching) service from a carrier.

Security

The interest in security is driven by multiple factors. One factor is the desire to reduce the number of security incidents and the corresponding financial losses. Another factor is the requirements set out in government mandates such as HIPAA and the Gramm-Leach-Bliley Act. Given both the great interest in security as well as the complexity of the problem being solved, many enterprises have implemented myriad security technologies. This includes using protocols such as SSL, HTTPS and IPSec that are designed with security in mind.

[12] "Wrestling XML Down to Size", James Kobielus, BCR, December 2004

It also includes implementing some or all of the following technologies:

- Firewalls
- Encryption
- Wired Equivalent Privacy
- Authentication
- Intrusion Detection
- Intrusion Protection
- Network Access Control
- Virus Scanning
- Virtual LANs
- Access Control Lists
- Digital Certificates
- Digital Signatures
- Smart Cards
- Biometrics

Storage

The traditional way to provide storage is called Direct Attached Storage (DAS), and refers to a couple of slightly different approaches. For example, DAS refers to having storage that is internal to a computer, as is the case with most PCs. DAS also refers to having hosts and storage devices attached over point-to-point connections, as shown in Figure 2.

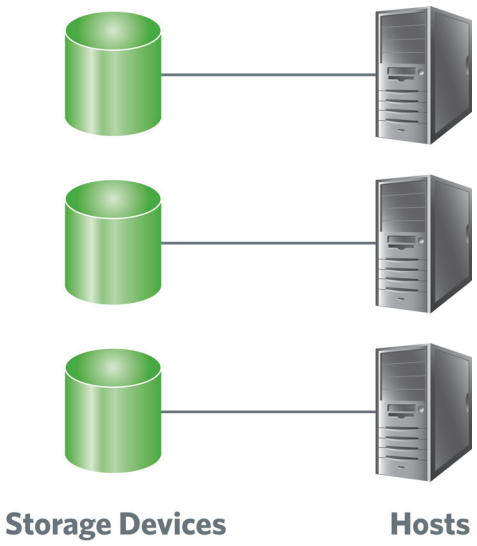

Figure 2. A DAS Environment.

In the late 1990s, the need to improve the utilization of storage lead to the development of alternative approaches, such as Storage Area Networking. As shown in Figure 3, a Storage Area Network (SAN) consists of hosts and storage devices, connected via one or more Fibre Channel switches.

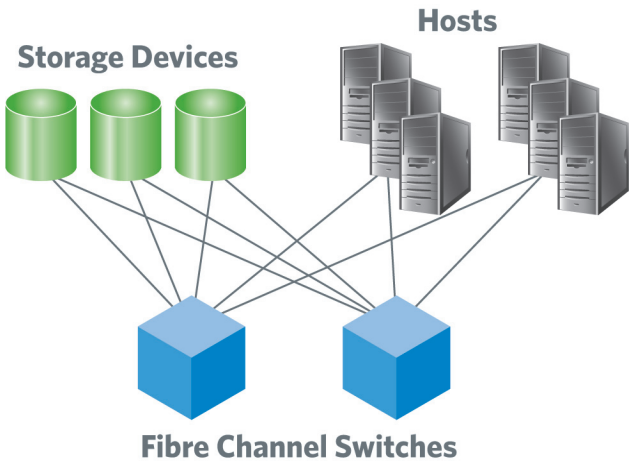

Figure 3. A SAN Environment.

One of the very clear advantages of a SAN is the ability to flexibly scale the amount of storage that is available to an enterprise's key applications.

Branch Office Optimization Solutions

The goal of Branch Office Optimization Solutions is to improve the performance of applications delivered from the data center to the branch office or directly to the end user. These solutions are designed in part to compensate for characteristics of the WAN that negatively impact application performance. These WAN characteristics include insufficient bandwidth, high latency, packet loss and congestion.

Some of the techniques that comprise branch office optimization solutions include:

- Caching
- Compression
- Congestion Control
- Differencing
- Forward Error Correction
- TCP Acceleration
- HTTP Acceleration
- Request Prediction
- Request Spoofing

Application Front Ends (AFEs)

The genesis of this category of product dates back to the IBM mainframe-computing model of the late 1960s and early 1970s. Part of that computing model was to have a Front End Processor (FEP) reside in front of the IBM mainframe. The primary role of the FEP was to free up processing power on the general purpose mainframe computer by performing communications processing tasks, such as terminating the multi-point private lines, in a device that was designed just for these tasks.

From a more contemporary perspective, the current generation of AFEs evolved from the earlier generations of Server Load Balancers (SLBs) that were deployed in front of server farms and, as the name implies, balanced traffic over multiple servers. While an AFE still functions as an SLB, the AFE has assumed, and will most likely continue to assume, additional roles. For example, AFEs serve to both manage application traffic within the data center, as well as accelerate the delivery of applications asymmetrically from the data center to individual remote users. One of the primary new roles played by an AFE is to offload processing-intensive tasks that are network related and that consume considerable CPU cycles. An example of server offload is the SSL processing in the data center. SSL offload allows the intranet Web and Internet eCommerce servers to process more requests for content and handle more transactions.

AFEs implement some of the optimization techniques that were listed in section 4.4. However, AFEs also implement functionality not found in a Branch Office Optimization Solution. An example of this additional functionality is reverse caching. Reverse caching refers to having a cache inside of the AFE the role of which is to accelerate the delivery of Web pages. All Web pages sent from a server back to the user's browser move through and are stored in the reverse cache. If the next request for a Web page has already been stored in the cache, it is retrieved from the cache, minimizing the involvement of the servers in rendering Web pages to users.

The Service Delivery Platform — A Reality Check

There is a wide variability in terms of how the trends discussed previously in this paper impact any given enterprise. Part of this variability is a result of the differing dynamics of each vertical industry. To exemplify both the similarities and differences in terms of how the business and technology trends impact a given enterprise, this section will take the trends that were outlined in preceding sections and discuss them in the context of four specific industries — financial services, health care, government, media and entertainment.

Financial services

The financial services industry faces several challenges. Some of the challenges faced by the banking component of the financial services industry include:

- **Implementing Intelligent Customer Management.** Leading banks are using portals to better connect customers and the bank with the goal of providing a comprehensive, consistent view of the customer to the bank's service employees.
- **Improving Security.** Many banks are working on multiple initiatives the goal of which is to eliminate identity fraud and to control access to customer data.
- **Building an Enterprise Payment Structure.** Banks are deploying a new payment structure in order to leverage the synergies between different payment methods. The goal of this initiative is to reduce costs and to eliminate redundant processes.
- **Training the Workforce.** Many banks are using eLearning techniques to continually train the workforce on new products and services.

Some of the challenges faced by the insurance component of the financial services industry include:

- **Streamlining the Operations.** Most insurance companies have too many systems; i.e., too many claim management systems, too many billing systems. This results in unnecessary cost as well as extra delay in responding to customers.
- **Protecting the Enterprise.** Insurance companies are responding to a widening array of security concerns and regulatory requirements.
- **Optimizing the Workforce.** Leading insurance companies are implementing systems that can automate processes such as insurance underwriting as well as many customer care activities; e.g., allowing customers to pay bills directly from their bank account.

- **Improving Agent Productivity.** Many insurance companies are deploying portals to make it easier for agents to get access to important information. In many cases this also involves implementing E-Learning initiatives.

An example of a bank that is using its network to respond to business challenges is the City National Bank of West Virginia. The bank has been growing rapidly, due in part to the expansion of branches into Wal-Mart stores as well as the acquisition of branches from other banks. In order to support future expansion, City National deployed a new converged voice-data network. This network allowed the bank to integrate its disparate phone systems and to support a wide range of data networking applications such as teller systems, new accounts systems, ATMs, research systems, check imaging, loan documents, email, and accounting and financial systems.

Health Care

The health care industry is comprised of two primary components. One component is the health care providers. These are the people and organizations that administer care. This includes hospitals, clinics, nursing homes and home health care organizations. The second component is the health care payers. Health care payers are entities that receive claims for the cost of medical services and are responsible for payment.

The key issues facing the health care providers include the need to:

- Improve the quality of care
- Respond to ongoing labor shortages
- Control the cost of health care
- Increase operating efficiencies and effectiveness
- Conform to the Health Insurance Portability and Accountability Act (HIPAA)

For example, the Children's Hospital and Medical Center of Seattle needed a way to provide its roving employees with uninterrupted access to important medical applications on the network, as well as to provide patients and their families with Internet access. It also needed a real-time tracking system for high-value mobile hospital equipment, and it needed to find a way to keep costs down.

A hospital-wide LAN installation — wireless coupled with mobile devices and computers on wheels (COWs) — brought applications to the point of care. The hospital deployed wireless (WiFi) networking equipment and uses Radio Frequency Identification Technology to do real-time tracking of its high-value equipment. One of the most important benefits of the new wireless network is that hospital staff can now perform critical tasks without leaving a patient's side.

The key issues facing the health care payers include the:

- Rising cost of health care
- Consumerism movement, in which patients demand more user-friendly medical facilities, as well as more input into both the health care benefits and the treatments that they receive
- Requirement to reduce the costs associated with customer service

The University of Alabama's Health Care Services Foundation needed a cost-effective way to collect on patient accounts with small balances. Their system, which involved multiple patient contacts, was costing the foundation more than patients owed. The foundation deployed a new contact center that enables patients to both pay their bills by credit card and to check their balances over the phone through an automated system.

Entertainment, Communications and Media

The entertainment, communications and media industry faces several challenges. This includes:

- **Reducing Customer Churn.** This industry derives significant revenue from customer subscriptions and customers are faced with a growing number of choices.
- **Increasing Advertising Revenue.** The traditional advertising model, particularly of print media, is changing rapidly. Companies in this industry are attempting to find advertising opportunities in other media.
- **Leveraging Industry Consolidation.** This industry is characterized by a number of mergers and acquisitions. The companies that are the result of these mergers and acquisition are now challenged to leverage the diverse set of assets that they have accumulated.

Jerry Bruckheimer Films, is the California-based entertainment-production company that has produced Black Hawk Down, Pirates of the Caribbean as well as the CSI television series. However, the phone system at Bruckheimer was having trouble adding new lines and accommodating the new voice mailboxes that were needed as each new production brought with it a whole new group of professionals. The system also couldn't support sophisticated applications or the latest media-rich conferencing tools.

To modernize its phone and messaging system, Jerry Bruckheimer Films installed a converged network to support voice, video and data communications. For example, the network provides support for phone conferences with multiple participants. In addition, unified messaging, using pre-existing Microsoft Exchange platform, lets employees check both email and voice mail from a single mailbox.

Government

One of the challenges facing state and federal government agencies is the mandate to implement a variety of E-Government initiatives. The goal of these initiatives is to:

- Make it easy for citizens to obtain services and interact with the Government
- Improve Government efficiency and effectiveness
- Improve Government's responsiveness to citizens

For example, in order to improve customer service the Social Security Agency (SSA) has begun an initiative to implement a paperless office by 2010. As part of this initiative the SSA has expanded online customer service options including retirement claims, Medicare replacement cards, online account access, and the ability for a customer to change components of their profile, such as their address and phone number.

Some of the other relevant government initiatives include:

- **GovBenefits.gov.** This is a self-service Web site implemented by the Department of Labor. This system helps citizens gauge their eligibility for 55 different Federal benefit schemes, covering more than one trillion dollars in eligible entitlements.
- **E-Government and Customer Care: Defense Logistics Information Service (DLIS).** When visitors log on to the DLIS Web site, they are greeted by the voice of Phyllis, a virtual representative with a computer-generated voice that provides customized assistance to help users complete online transactions, and advanced information searches.
- **Workforce Optimization: E-Travel Initiative.** The Federal Government is working on a plan to move its travel arrangements online, an initiative that it estimates could save taxpayers up to $2 billion.
- **Workforce Optimization: Air Force E-Forms.** The Air Force is implementing software for operational tests of a four-year, multi-million dollar contract aimed at securing the myriad electronic forms used by service personnel. The Air Force has about 14,000 E-Forms that are used by more than 700,000 members worldwide, and new technology will help transform those static electronic documents into XML-based records.

The Management Challenge

As the business and application trends that were discussed previously continue to unfold, the increased diversity of applications will require the network infrastructure to provide an appropriate level of service for each application, rather than simple, utility-like connectivity. This requires the network to be application-aware and have the functionality (e.g., security, QoS, SLB, acceleration, caching, etc,) to accommodate specific application characteristics. As this level of network functionality is deployed, the network becomes a key component of the end-to-end service delivery platform. From an application perspective, the complete service delivery platform includes the application itself plus the servers, appliances, storage, and network resources that work in concert to determine user response time for information delivery.

In the service delivery model, the computing and network resources which support each application service are not generally dedicated to the service, but may be drawn from a pool of resources shared with a number of other services. As a result, it is desirable to establish and track a set of service delivery metrics for each application and to consider establishing service level agreements (SLAs) based on these metrics.

Managing the IT infrastructure as a service delivery platform requires an approach in which the resources delivering a service are viewed by the management system as components of a complete system, rather than a series of standalone boxes. The end-to-end view at the service/system level helps to ensure that services meet expectations and SLAs are met. The transition to managing the infrastructure as a service delivery system involves a number of significant challenges for the IT management staff including:

- **Integration.** Managing service and application delivery requires close collaboration between system and network specialists who should share a consistent view of the end-to-end delivery system. Therefore, organizational silos and isolated management solutions need to give way to more tightly integrated management of business applications, server systems, databases, and the network infrastructure.
- **Complexity.** The IT environment continues to become more complex, with more distributed applications, different types of applications, and more specialized devices to provide security and performance enhancements. The current trend toward virtualization of applications, servers, storage, as well as network functions has numerous benefits in terms of resource efficiency, but adds even more management complexity.
- **Data Analysis and Reporting.** With the quality of service delivery measured in terms of service metrics and SLAs, there will be an increased emphasis on the generation of real time and historical data to ensure that the platform is performing up to expectations. This performance data will need to be provided to both IT management personnel and various stakeholders and needs to be provided in a manner that is relevant to each constituency.

- **Productivity.** In spite of the growing IT workload that may be implicit in a transition to a new management model and increasing complexity, in most companies the IT staff size and budget are not going to be increased in any significant way. This will require greater emphasis on automation of management functions in order to increase productivity.

Selection Criteria for Management Solutions

Organizations that are adopting a service delivery model for infrastructure management will need to make some changes and adjustments in the tools used to support this function. The evaluation of management tools considered for deployment should be based on a set of criteria that addresses the challenges of managing a service delivery platform at the system level. Some of the criteria that should be considered as part of the evaluation include:

- **Level of Integration.** The suite of tools that are used to manage the infrastructure and the application services needs to be well integrated. Ideally, a single, integrated management platform can provide a span of control that extends across all layers of the infrastructure from the application and servers to the physical networks and network services that provide connectivity. Going forward, a key aspect of the integration will be the capability to deal with both physical and logical views of the entire infrastructure. For example, system management should support server virtualization by integrating management of both real and virtual servers. In addition, the network topology function should be capable of discovering and monitoring both the physical and logical topologies of the network. Integrated service management is an important enabler that provides the foundation for virtualization of the infrastructure, as well as virtualization of applications via implementation of a Service Oriented Architecture (SOA).
- **Support for Heterogeneity.** The typical IT environment includes a range of technologies and application services supported by a number of different vendors. This diversity will tend to increase as more services, such as telephony and videoconferencing, move to the converged IP network. A management platform that can accommodate a wide diversity of vendors and technologies facilitates the introduction of new services and also eases technology transitions, such as moving from TDM voice to VoIP or from Fibre Channel to iSCSI. In addition to supporting a highly diverse distributed computing infrastructure the ideal management system also can provide support for the mainframe environment.
- **Support for Infrastructure Interdependencies.** Tools for service-level management should have the intelligence to represent the interdependencies among the various components of the system that support each application. These capabilities should include a mapping between an application service and the components of the physical and logical topology of the infrastructure. Knowledge of these inter-relationships allows an event occurring within the infrastructure to be expanded into a list of the application services and the SLAs that may be affected.
- **Support for Intelligent Service Assurance.** In order to support SLAs, the fault and performance management functions need advanced levels of intelligence. De-duplication of alarms needs to be supplemented with automated root cause analysis based on a highly detailed model of the physical and logical topologies of the network. Root cause analysis that can identify likely causes at the level of field replaceable units greatly accelerates remedial action. Integrated performance management can also provide proactive notification of degraded performance allowing remedial action to be taken, often preventing a fault occurrence that would interrupt service delivery. When a fault or performance degradation does occur, the mapping between services and the infrastructure topology can be used to determine the severity of the event in terms of which key services are affected.
- **Reporting.** The enhanced intelligence of the management platform should be leveraged to provide a wide range of notifications and reports customized to deliver the right information and detail for individuals at different levels within the IT staff, executive staff, and the business units. For example, a fault could result in a notification for technical personnel of the fault occurrence and its probable root cause together with notifications to the stakeholders of the affected applications that SLAs are being threatened or that certain business operations or deadlines are being compromised.
- **Automation.** The ability to automate infrastructure management functions offers the potential to increase the productivity of IT staff. While management automation has traditionally been restricted to mainframe environments, management of distributed systems is becoming more automated. This trend is driven in part by the adoption of data center virtualization as a step toward utility computing and SOA. Management functions that lend themselves to automation include: problem notification and escalation, systems and storage provisioning, workload management, software and configuration updates, system restart and other simple remedial actions, and backup/restore operations.

Summary

There is currently a domino effect impacting IT organizations. The first domino to fall is the fundamental shifts underway in the business environment. To respond to these shifts, most enterprises are making changes to their applications environment. The next domino is that the typical enterprise network is changing in order to support the changing business and application environments. In particular, many IT organizations have shifted from positioning their WAN to be just plumbing and currently position their WAN to also be a platform. Therefore, the typical enterprise WAN is being designed to support the shifts in both the business and application environments by offering sophisticated WAN functionality such as QoS, security, as well as network and application acceleration.

Just as the enterprise WAN is responding to changes in the business and application environments by becoming a platform, so must enterprise management transition in order to support the emerging WAN platform. Some of the challenges IT professionals are faced with include:

- **Integration.** To enable system and network specialists to work collaboratively requires a management solution that tightly integrates the management of business applications, servers, databases and the network infrastructure.
- **Complexity.** The IT environment continues to become more complex, with the movement to deploy more distributed applications, implement virtualization and deploy specialized devices that provide a wide range of added functionality. As the IT environment becomes more complex, the management environment also becomes more complex.
- **Data Analysis and Reporting.** As enterprises place increased emphasis on service metrics and SLAs, there is an increased emphasis on the generation of real time and historical reports that provide value to both IT management personnel and various stakeholders, including business unit managers.
- **Productivity.** To support the increased workload that is associated with the added complexity, requires greater emphasis on automation of management functions in order to increase productivity.

In addition, as IT organizations make this management transition they will need to implement new management tools. Some of the criteria that these organizations should use when evaluating new management tools include the:

- **Level of Integration.** The suite of tools used to manage the infrastructure and the application services needs to be well integrated. Ideally, a single, integrated management platform can provide a span of control that extends across all layers of the infrastructure from the application and servers to the physical networks and network services providing connectivity.
- **Support for Heterogeneity.** The typical IT environment includes a range of technologies and application services supported by a number of different vendors. This diversity will tend to increase as more services, such as telephony and videoconferencing, move to the converged IP network. A management platform that can accommodate a wide diversity of vendors and technologies facilitates the introduction of new services and also eases technology transitions.
- **Support for Infrastructure Interdependencies.** Tools for service-level management should have the intelligence to represent the interdependencies among the various components of the system that support each application. These capabilities should include a mapping between an application service and the components of the physical and logical topology of the infrastructure.
- **Support for Intelligent Service Assurance.** In order to support SLAs, the fault and performance management functions need advanced levels of intelligence. For example, de-duplication of alarms needs to be supplemented with automated root cause analysis based on a highly detailed model of the physical and logical topologies of the network. Root cause analysis that can identify likely causes at the level of field replaceable units greatly accelerates remedial action.
- **Reporting.** The enhanced intelligence of the management platform should be leveraged to provide a wide range of notifications and reports customized to deliver the right information and detail for individuals at different levels within the IT staff, executive staff and the business units.
- **Automation.** The ability to automate infrastructure management functions offers the potential to increase the productivity of IT staff. While management automation has traditionally been restricted to mainframe environments, management of distributed systems is becoming more automated.

In order to be successful, IT organizations must continually place this management transition into business perspective. In particular, network management provides business value because it enables enterprises to implement, manage and secure the type of fully functional enterprise WAN that is necessary to support the emerging business and application environments.

About the Author:

An authority on network technology and its business application, Dr. Jim Metzler has over 28 years of professional experience, assisting vendors to refine product strategies, service providers to deploy technology, and enterprises to evolve network infrastructure. He has a Ph.D. in Numerical Analysis from Boston University, co-authors a weekly online column for Network World on Wide Area Networking and co-authored, “Layer 3 Switching: A Guide for IT Professionals” published by Prentice Hall.

MP312080207