The Road to Cloud Computing

How to Evolve Your Data Center LAN to Support Virtualization and Cloud

Introduction

Cloud computing is one of the most important topics in IT. The reason for that importance was summarized in a report entitled A Guide for Understanding Cloud Computing[1]. According to that report, the goal of cloud computing is to make a dramatic increase in the cost effective, elastic provisioning of IT services.

One of the companies that has validated that the goal of cloud computing is achievable is Bechtel. Geir Ramleth, the CIO of Bechtel, benchmarked his IT organization vs. leading Internet companies such as Google and Amazon.com. One result of this benchmarking was that Bechtel estimated that Google uses 12 system administrators for every 200,000 servers, or roughly 17,000 servers per system administrator. Bechtel, on the other hand, was operating with 1,000 servers per system administrator. Ramleth was quoted in an article in Network World[2] as saying that "What we learned is that you have to standardize like crazy and simplify the environment. Google basically builds their own servers by the thousands or gets them built in a similar fashion, and they run the same software on it. So we had to get more simplified and standardized."

Another part of Bechtel's benchmarking project analyzed Amazon.com and determined that Amazon.com must have a notably better storage strategy than Bechtel does because Amazon.com is offering disk space for a fraction of Bechtel's internal costs. In particular, the project team determined that while Amazon.com was offering storage for $0.10 per gigabyte per month, Bechtel's internal rates in the United States was $3.75 per gigabyte per month.

The results of Bechtel's benchmarking project didn't convince Ramleth to turn over large portions of his IT infrastructure to suppliers such as Google or Amazon. Rather, he determined that Bechtel would adopt that same approach to computing internally as was used by these vendors. For example, one of the conclusions reached by Ramleth was that the key to reducing his storage costs was to simplify and virtualize his storage environment and to also to drive up utilization. According to Ramleth "Our average utilization was 2.3%," With virtualization, "we now expect to have utilization in the 70% to 75% range."

Because of the constant pressure that all IT organizations are under to become more agile and more cost effective, the majority of IT organizations have begun to face the question of how to best implement virtualization and cloud computing. One of the challenges that could prevent IT organizations from realizing all of the possible cost savings and added agility is that the vast majority of data center LANs were built using a decade old approach to LAN design. As will be explained in this white paper, in order to realize the benefits of virtualization and cloud computing, IT organizations must evolve their data center LANs.

[1] http://www.webtorials.com/content/2009/11/a-guide-for-understanding-cloud-computing.html
[2] http://www.networkworld.com/news/2008/102908-bechtel.html

Cloud Computing

There are three primary classes of cloud computing: public, private and hybrid. Cloud computing service providers (CCSPs) such as Google and Amazon that provide services over a WAN are considered to be providing public cloud computing solutions. There are three categories of public cloud computing solutions. They are:

- Software as a Service (SaaS)

  In one form of SaaS, an independent software vendor (ISV) such as Salesforce.com hosts an application in one or more of their own data centers. In another form of SaaS, an ISV such as Virtual Bridges hosts an application in one or more data centers provided by an Infrastructure as a Service vendor such as Amazon. In either case, the application is provided to users on a usage basis.

- Infrastructure as a Service (IaaS)

  The two primary forms of IaaS are compute and storage. Providers of IaaS solutions typically implement their solutions on a virtualized infrastructure and charge for them on a usage basis.

- Platform as a Service (PaaS).

  PaaS is the delivery of a computing platform and solution stack as a service. PaaS offerings from providers such as Force.com include workflow facilities for application design, application development, testing, deployment and hosting as well as application services such as web service integration, database integration and security.

There is significant interest in public cloud computing solutions, most notably SaaS and IaaS. However, due in part to concerns about security and data confidentiality, most IT organizations will, like Bechtel, decide to adopt the same techniques internally as is used by CCSPs such as Google and Amazon. That approach is referred to as private cloud computing.

A hybrid cloud computing solution involves a combination of services provided by the IT organization itself as well as by one or more CCSPs. For example, an IT organization may either already have, or be in the process of acquiring a 4-tier application. The IT organization may decide that for security reasons that it wants to host the application and database servers itself. However, in order to improve the interaction with the users of the application, the IT organization may also decide to let a CCSP host the web tier in numerous data centers around the globe.

As noted, private cloud computing involves IT organizations implementing the same techniques themselves as are typically associated with public cloud computing solutions. There is not, however, a litmus test to determine the precise set of techniques that determine whether or not a given solution qualifies as being a cloud computing solution. That said, the following techniques are typically associated with cloud computing.

- **Centralization** of applications, servers and storage resources.

- Extensive **virtualization** of every component of IT, including servers, desktops, applications, storage, networks and appliances such as WAN optimization controllers, application delivery controllers and firewalls.

- **Automation** of as many tasks as possible; e.g., provisioning, troubleshooting, change and configuration management.

- The **dynamic movement of resources** such as virtual machines and the associated storage.

- Heavy **reliance on the network** to support high performance connectivity between resources that are dynamic and scalable.

- **Standardization** of the IT infrastructure in terms of the technologies and products that are utilized.

- The **convergence of technologies** such as LAN and SAN as well as the convergence of networks, servers and storage.

- The **consolidation of the IT subgroups** that provide the converged technologies.

Drivers of Change

A recent market research report[3] presented the results of a survey that was given to attendees of the Interop conferences. The survey respondents were asked to indicate the primary factor that is driving their organization to redesign their data center LAN. Their responses are summarized in Table 1.

Number	Factor	Percentage
1	To reduce the overall cost	22.4%
2	To support more scalability	11.6%
3	To create a more dynamic data center	11.6%
4	To support server virtualization	11.2%
5	To reduce complexity	9.9%
6	To make it easier to manage and orchestrate the data center	9.2%
7	To support our storage strategy	7.5%
8	To reduce the energy requirements	6.5%

[3] http://www.webtorials.com/content/2010/11/2010-cloud-networking-report.html

9	Other (please specify)	6.1%
10	To make the data center more secure	4.1%

Table 1: Factors Driving Data Center LAN Redesign

One conclusion that could be drawn from the data in Table 1 is that a broad range of factors is driving IT organizations to redesign their data center LANs. The factors, however, fall into a few major categories. Table 2 depicts the three major categories, the factors that comprise each category and the percentage of respondents that indicated that that category is the primary factor that is driving their organization to redesign their data center LAN.

Category	Factors	Percentage
Support the emerging dynamic, virtualized data center	2, 3, 4	34.4%
Reduce cost	1, 8	28.9%
Reduce complexity	5, 6	19.1%

Table 2: Primary Factors Driving Data Center Redesign

The data in Table 2 presents a simpler view of the factors that are driving IT organizations to redesign their data center LANs. As shown in Table 2, the single most important factor is the need to support a dynamic, virtualized data center. However, reducing cost is almost as important a factor and reducing cost is only somewhat more important than reducing complexity.

The Traditional Data Center LAN

The first generation data center LAN of the early to mid 1990s was based on shared media. In the mid 1990s, driven both by the need to provide improved performance and reduced cost, IT organizations began to implement a second generation data center LAN that was based on switched Ethernet. The second generation data center LAN was usually designed around a three-tier switched architecture comprised of access, distribution and core switches. This design was dictated by the fact that second generation LAN switches had relatively low port densities and hence, in most cases it wasn't possible to provide all of the necessary connectivity with a two-tier design.

The deployment of second generation data center LANs was also characterized by:

- The use of the spanning tree protocol at the link layer to ensure a loop-free topology.
- Switches with relatively low speed ports.
- The separation of the data network from the storage network.

- The use of Ethernet on a best-effort basis by which packets may be dropped when the network is busy.

- The need to primarily support client-to-server communications, which is sometimes referred to as 'north-south traffic'.

- The over-subscription of uplinks.

- Redundant links and fast failover protocols to increase availability.

- The application of policy (QoS settings, ACLs) based on physical ports.

As is described below, many of the characteristics of the second generation data center LANs are being gradually replaced as IT organizations evolve to a third generation of data center LANs.

The Emerging Data Center LAN

This section of the white paper will describe a four-stage evolutionary path that IT organizations can take to evolve their data center LANs. A critical characteristic of the evolutionary path is that it is flexible. In particular, IT organizations are able to traverse the path at their own pace, driven by the demands and constraints of their environment.

Stage One: Server Consolidation

As part of stage one most IT organizations begin the process of consolidating servers out of branch offices and placing them into centralized data centers. This consolidation typically reduces cost and enables IT organizations to have better control over the company's data. In addition to consolidating servers, during this stage many companies also reduce the number of data centers they support worldwide and most begin to virtualize at least some of their data center servers. Driven in part by the need to support the increased I/O that is associated with virtualized servers, during this stage some IT organizations begin to implement LAN switches with 10 Gbps interfaces.

Other than implementing some higher performing LAN switches, the data center LAN at stage one closely reflects the characteristics of the previously described traditional data center LAN. For example, in this stage the data center LAN relies on the spanning tree protocol to eliminate loops. Unfortunately, the spanning tree protocol limits the bandwidth efficiency and scalability of data center LANs. For example, in those situations in which IT organizations have absorbed the significant extra cost that is associated with having multiple links between LAN switches, only one of the links can actively send traffic.

During stage one the various subgroups within IT (e.g., networking, storage, servers) which provide data center functionality operate in silos. By *operating in silos* is meant that these organizations don't have common goals, tools and processes. While this is sometimes problematic, it is not a significant issue in stage one. In addition to having an organization that functions in silos, the overall approach to data center design at this stage is characterized by

having functionality such as servers, storage, LAN switches, firewalls and load balancers be both manually configured and dedicated to a single service or application. This approach to data center design and management results in an increase in the overall cost of the data center. It also increases the time it takes to deploy a new service or application since new infrastructure must be designed, procured, installed, configured and tested before a new service or application can go into production.

Stage Two: Data Center Virtualization

As noted, in stage one, IT organizations consolidated servers out of branch offices and began to virtualize at least some of their data center servers. One of the key characteristics of stage two is that during this stage IT organizations build on these achievements and implement server virtualization more broadly. The increased deployment of server virtualization at this stage leads to greater flexibility and the increased movement of VMs.

In stage two, the previously described traditional data center LAN begins to break down. For example, during stage two there is an increase in server-to-server communications, which is sometimes referred to as "east-west " traffic. Another change that occurs during stage two is that as IT organizations deploy servers with an increased number of cores, the number of virtual machines (VMs) per physical server typically increases proportionally. One impact of the increase in the number of VMs per physical server is that the network I/O requirements of the multi-core physical servers that have been virtualized in stage two begin to exceed the capacity of the existing GbE and multi-GbE aggregated links. Thankfully the traditional economics of Ethernet performance improvement[4] is falling into place for 10 Gigabit Ethernet (10 GbE) and during stage two, the implementation of 10 GbE ports increases significantly. Another impact of the increasing number of VMs per physical server is that the previously adhered to design philosophy of oversubscribing LAN links now leads to performance bottlenecks. As a result, in stage two most IT organizations eliminate over-subscription from their data center LANs.

However, the biggest change to the traditional IT model that occurs during stage two is that driven both by the need to save money and to reduce complexity, many IT organizations implement a two-tier data center LAN, consisting of access and core switches, in at least one of their data centers. As previously discussed, IT organizations deployed three tier data center LANs largely because second generation LAN switches had relatively low port densities and hence, in most cases it wasn't possible to provide all of the necessary connectivity with a two-tier design. During stage two, the ability of IT organizations to implement a two-tier data center LAN is enabled by the deployment of a third generation LAN switch with relatively high port densities. For example, in today's environment third generation stackable LAN switches can support tens of 10 GbE ports and modular LAN switches can support hundreds of 10 GbE ports.

During stage two, the vast majority of IT organizations still operate in silos. However, unlike the situation in stage one, during stage two the fact that the IT organization functions in silos begins to have a significant negative impact on the IT organization. One source of that negative impact stems from the fact that once a server is virtualized, that physical server is equipped with a hypervisor-based virtual switching (vSwitch) capability. The primary role of the vSwitch is to

[4] Ethernet typically provides a 10x higher performance for a 3-4x increase in cost.

facilitate connectivity among VMs on the same physical platform as well as to switch traffic that is destined to external destinations. However, the existence of the vSwitch raises a fundamental question: Is a switch inside of a server the responsibility of the network organization or the server organization? Unfortunately, in many IT organizations the deployment of virtual servers has caused organizational stress as both the network and the server organizations claimed or disclaimed responsibility for the vSwitch. One common manifestation of this organizational stress is that changes to the data center LAN occur more slowly and are more error prone than they are in a traditional data center. Recognizing that organizational issues can limit their ability to fully leverage the advantages of emerging LAN technologies, during stage two some IT organizations begin to implement initiatives such as cross training their employees on multiple IT disciplines.

Stage Three: Cloud Optimized Data Center LAN

One of the incremental changes to the traditional data center LAN environment that occurs in this stage is that IT organizations build on the success that they had in implementing two-tier data center LANs in stage two, and implement two-tier data center LANs in additional data centers. During stage three, IT organizations also further increase their deployment of server virtualization.

As previously noted, in the traditional data center LAN environment, the data network is kept separate from the storage network. In stage three, some IT organizations begin to experiment with deploying a unified data center switching fabric. However, as previously noted, traditional Ethernet only provides a best effort service that allows buffers to overflow during periods of congestion and which relies on upper level protocols (e.g., TCP) to manage congestion and recover lost packets through re-transmissions. The best effort nature of Ethernet presents a challenge relative to supporting a unified fabric as well as for supporting other traffic classes such as video and telepresence. In stage three, many IT organizations respond to this challenge by beginning to deploy a new generation of Lossless Ethernet technologies that are based on a collection of standards that are commonly referred to as IEEE Data Center Bridging (DCB).

A major change in the data center environment that has been building over the previous two stages and which peaks in stage three is that the combination of the ongoing movement both to consolidate and to virtualize servers has created an “all your eggs in one basket” phenomenon in terms of the amount of critical IT resources that most IT organizations have placed in a small number of data centers. In this stage that phenomenon drives the need for highly available data center LANs with greater fault tolerance.

During stage three the bandwidth efficiency and availability of Layer 2 data center LANs with redundant links can be greatly improved by assuring that the parallel links from the servers to the access layer and from the access layer to the core layer are always in an active-active forwarding state. As previously noted, the spanning tree protocol prevents this from happening. In stage three, IT organizations implement two different approaches to eliminating loops without using the spanning tree protocol. One of these approaches involves switch virtualization and multi-chassis (MC) link aggregation group (LAG) technology. In this context *switch virtualization* refers to two or more physical switches being made to appear to other network elements as a

single logical switch or virtual switch, with a single control plane. In this approach when there are redundant configurations the connections between the end systems and the virtual access switches and between the virtual access switches and the virtual core switches are based on MC LAG that allows the links of the LAG to span the multiple physical switches that comprise a virtual switch. The second approach is based on TRILL (Transparent Interconnection of Lots of Links). TRILL is an Internet Engineering Task Force (IETF) project to develop a Layer 2 shortest-path first (SPF) routing protocol for Ethernet. The associated set of standards is sufficiently developed that in stage three some IT organizations begin to deploy TRILL.

Another major change that occurs in stage three is that many IT organizations make progress with reducing the negative impact caused by the previously discussed organizational silos. They do this in part by continuing to cross train key employees and by creating goals and a reward system that encourages employees to take a more holistic view of IT. They also do this in part by beginning to implement common tools such as service orchestration. By automatically coordinating the provisioning and reuse of resource across servers, storage, and networks, service orchestration enables IT organizations to streamline operational workloads and overcome some of the impact of the organizational silos that were previously discussed. For example, an orchestration engine can automatically initiate the creation of virtual machines while simultaneously deploying the network access and security models across all of the required infrastructure components. This includes routers, switches, security devices, and core infrastructure services. The entire process can allow setup and deployment of network routes, VPNs, VLANs, ACLs, security certificates, firewall rules and DNS entries without any time consuming manual entries via device-specific management systems or CLIs.

Stage Four: The Near Term Future

While projecting too far into the future is risky, there is no doubt that data center LANs will continue to evolve past what IT organizations achieved in stage three. For example, upon completion of stage three, virtually all IT organizations will continue to expand their implementation of a unified fabric within data centers and to implement continuingly more sophisticated automation. Many IT organizations will also increase their efforts to reduce the impact of organizational silos. It is also highly likely that in stage four that IT organizations will make a notably broader deployment of TRILL than was accomplished in stage three. TRILL allows the difference between access switches and core switches to shrink significantly. As a result, in stage four some IT organizations shift their data center LANs away from a two-tier hub and spoke design to a highly meshed or even fully meshed array of switches that appears to the attached devices as a single switch.

Another possible change to data center LANs in the near term future that will build on top of what was already achieved is that many IT organizations are likely to extend their unified fabric both between their own data centers as well as between their data centers and one or more data centers provided by a cloud computing service provider (CCSP). One obvious advantage of doing this is that it enables IT organizations to efficiently move workloads between data centers and hence enhances the ability of the IT organization to implement disaster recover/business continuity solutions.

Another advantage of extending the unified fabric between data centers is that it will enable two emerging techniques that are intended to maximize the efficiency of cloud computing: cloud bursting and cloud sharing. Cloud bursting refers to taking an application that currently runs in a data center controlled by an IT organization and dynamically deploying that application and the subtending storage in a data center controlled by a CCSP. Cloud balancing refers to routing service requests across multiple data centers based on myriad criteria including proximity to the user, business metrics, response time, capacity, cost and power consumption. Unlike cloud bursting, cloud balancing assumes that the application is already running in each of the data centers. Both cloud bursting and cloud balancing allow IT organizations to efficiently respond to a spike in demand.

Summary

As demonstrated by Bechtel's benchmarking project, the effective implementation of virtualization and cloud computing enable IT organizations to be dramatically more cost effective and agile. However, virtualization and cloud computing make some unique demands on the network and IT organizations will not realize the benefits of these techniques if they continue to design their data center LANs based on traditional design principles.

The good news facing IT organizations is that there is a wide range of new and emerging LAN technologies that will enable them to fully support the demands of virtualization and cloud computing. Given both the range of the technologies and emerging nature of some of the technologies, the vast majority of IT organizations will evolve their data center LANs over a period of several years. While it is not possible to predict exactly how all IT organizations will evolve their data center LANs, it is highly likely to happen in a set of four stages.

At the beginning of stage one, the typical data center LAN closely resembles the traditional data center LAN. As this stage evolves, most IT organizations begin to consolidate servers out of their branch offices into centralized data centers. They also begin a modest deployment of both server virtualization as well as high-speed LAN switches.

During stage two, IT organizations increase their deployment of server virtualization and high-speed LAN switches and also begin to eliminate over-subscription from their data center LANs. The biggest technology change that occurs during stage two is that IT organizations implement a two-tier LAN design in at least one of their data centers. Another significant change that occurs during stage two is that the organizational silos that characterize the typical IT organization begin to have a significant negative affect and as a result, some IT organizations begin to take steps to change the organizational dynamics.

One of the incremental changes to the traditional data center LAN environment that occurs in stage three is that IT organizations increase their deployment of server virtualization as well as their deployment of two-tier data center LANs. In addition, some IT organizations begin to experiment with deploying a unified data center fabric while many IT organizations begin to deploy technologies such as DCB and TRILL. During stage three many IT organizations focus on transforming their organizational culture to be less based on technology domains and more focused on a holistic view of IT. One common way that IT organizations develop more of a

holistic focus during stage three is by implementing service orchestration to automatically coordinate the provisioning and reuse of IT resources.

Many of the changes that occur during stage four are extensions of existing initiatives. This includes expanding the implementation a unified fabric within a data center, deploying more sophisticated automation, increasing the deployment of technologies such as TRILL and expanding the efforts to reduce the impact of organizational silos. One of the most significant changes that occurs during stage four is that many IT organizations extend their unified fabric both between their own data centers as well as between their data centers and one or more data centers provided by a CCSP. This capability enables IT organizations to efficiently move workloads between data centers and enables emerging techniques that are intended to maximize the efficiency of cloud computing.