The 2010 Cloud Networking Report

Dr. Jim Metzler, Ashton Metzler & Associates

The Emerging Data Center LAN

The Emerging Data Center LAN

First and Second Generation Data Center LANs

As recently as the mid 1990s Local Area Networks (LANs) were based on shared media. Throughout this report these shared media LANs will be referred to as First Generation LANs. In the mid 1990s, companies such as Grand Junction introduced Ethernet LAN switches to the marketplace. The two primary factors that drove the deployment of Second Generation LANs based on switched Ethernet were performance and cost. For example, performance drove the deployment of switched Ethernet LANs in data centers because FDDI was limited to 100 Mbps whereas there was a clear path for Ethernet to evolve to continually higher speeds. Cost was also a factor that drove the deployment of Ethernet LANs in data centers because FDDI, in addition to being the only viable, high-speed First Generation LAN technology for interconnecting servers, was fundamentally a very expensive technology.

A key characteristic of Second Generation data center LANs is that they are usually designed around a three-tier switched architecture comprised of access, distribution and core switches. The deployment of Second Generation LANs is also characterized by:

- The use of the spanning tree protocol at the link layer to ensure a loop-free topology.
- Relatively unintelligent access switches that did not support tight centralized control.
- The use of Ethernet on a best-effort basis by which packets may be dropped when the network is busy.
- The need to support applications that are neither bandwidth intensive nor sensitive to latency.
- Switches with relatively low port densities.
- The over-subscription of uplinks.
- The separation of the data network from the storage network.
- VLANs to control broadcast domains and to implement policy.
- Redundant links and fast failover protocols to increase availability.
- Access Control Lists (ACLs) for rudimentary security.
- The application of policy (QoS settings, ACLs) based on physical ports.

Drivers of Change

The Webtorials Respondents were asked “Has your IT organization already redesigned, or within the next year will it redesign, its data center LAN in order to support cloud computing in general, and virtualized servers in particular?” Their responses are shown in Table 4.1.

	Already Have	Will Within the Next Year	No Plans
Cloud Computing in General	28.6%	42.9%	28.6%
Virtualized Servers in Particular	50.5%	30.7%	18.8%

Table 4.1: Redesign of the Data Center LAN

The data in Table 4.1 indicates that one of the key factors driving IT organizations to redesign their data center LANs is the deployment of virtual servers. The Webtorials Respondents were asked to indicate the percentage of their company's data center servers that have either already been virtualized or that they expected would be virtualized within the next year. The responses of The Webtorials Respondents are shown in Table 4.2.

	None	1% to 25%	26% to 50%	51% to 75%	76% to 100%
Have already been virtualized	21.6%	33.0%	18.9%	15.1%	11.3%
Expect to be virtualized within a year	12.4%	25.6%	21.9%	21.9%	18.2%

Table 4.2: Deployment of Virtualized Servers

As pointed out in ***Virtualization: Benefits, Challenges and Solutions***, server virtualization creates a number of challenges for the data center LAN. As previously discussed, one of these challenges is the requirement to manually configure parameters such as QoS settings and ACLs in order to support the dynamic movement of VMs. Other challenges include:

- **Contentious Management of the vSwitch**
  Each virtualized server includes at least one software-based virtual switch (vSwitch). This adds yet another layer to the existing data center LAN architecture. It also creates organizational stress and leads to inconsistent policy implementation.

- **Inconsistent Network Policy Enforcement**
  Traditional vSwitches can lack some of the advanced features that are required to provide the degree of traffic control and isolation required in the data center. This includes features such as private VLANs, quality of service (QoS), and sophisticated ACLs.

- **Layer 2 Network Support for VM Migration**
  When VMs are migrated, the network has to accommodate the constraints imposed by the VM migration utility; e.g., VMotion. Typically the source and destination servers have to be on the same VM migration VLAN, the same VM management VLAN, and the same data VLAN.

Server virtualization, however, is not the only factor causing IT organizations to redesign their data center LANs. Table 4.3 contains a list of the factors and the percentage of The Interop Respondents who indicated that it was the primary factor that is driving their organization to redesign their data center LAN.

Factor	Percentage
To reduce the overall cost	22.4%
To support more scalability	11.6%
To create a more dynamic data center	11.6%
To support server virtualization	11.2%
To reduce complexity	9.9%
To make it easier to manage and orchestrate the data center	9.2%
To support our storage strategy	7.5%
To reduce the energy requirements	6.5%
Other (please specify)	6.1%
To make the data center more secure	4.1%

Table 4.3: Factors Driving Data Center LAN Redesign

The data in Table 4.3 indicates that a broad range of factors are driving IT organizations to re-design their data center LANs. However, as is so often the case: The primary factor driving IT organizations to re-design their data center LAN is the desire to reduce cost.

Third Generation Data Center LAN Architecture and Technology Options

During the transition from First Generation LANs to Second Generation LANs there was considerable debate over the underlying physical and data link technologies. Alternatives included Ethernet, Token Ring, FDDI/CDDI and ATM. One of the few aspects of Third Generation Data Center LANs that is not up for debate is that they will be based on Ethernet. In fact, the Third Generation LAN will provide the possibility of exploiting Ethernet as the single data center switching fabric, eventually displacing special purpose fabrics, such as Fibre Channel for storage networking and InfiniBand for ultra low latency HPC cluster interconnect.

Below is a listing of some of the primary objectives of a Third Generation Data Center LAN and an analysis of the various alternatives that IT organizations have relative to achieving those objectives.

Improved Topologies for Server-to-Server Communications

Many of the on-going IT initiatives that are aimed at improving the cost-efficiency of the enterprise data center (e.g., server virtualization, SOA, Web 2.0, access to shared network storage, and the exploitation of HPC and cluster computing) are placing a premium on IT organizations being able to provide a highly reliable, low latency, high bandwidth communications among both physical and virtual servers. Whereas the hub and spoke topology of the traditional three tier Second Generation LAN was optimized for client-to-server communications (sometimes referred to as 'north-south traffic'), it is decidedly sub-optimal for server-to-server communications, which is sometimes referred to as "east-west" traffic.

One approach for improving server-to-server communications is to flatten the network from three tiers to two tiers consisting of access layer and aggregation/core layer switches.

A two-tier network reduces the number of hops between servers, reducing latency and potentially improving reliability. The typical two-tier network is also better aligned with server virtualization topologies where VLANs may be extended throughout the data center in order to support dynamic VM migration at Layer 2.

Two tier networks require switches that have very high densities of high-speed ports and a higher level of reliability to protect the soaring volumes of traffic flowing through each switch. Reliability and availability also require redundant switch configurations in both tiers of the network. Some of the common characteristics of two tier networks are described in the following subsections.

High Port Density and Port Speed

The network I/O requirements of multi-core physical servers that have been virtualized are beginning to transcend the capacity of GbE and multi-GbE aggregated links. As the number of cores per server increases, the number of VMs per physical server can increase well beyond the 5-10 VMs/server that are typical today. With more VMs per server, I/O requirements increase proportionally. Thankfully the traditional economics of Ethernet performance improvement[1] is falling into place for 10 Gigabit Ethernet (10 GbE). As a result, Third Generation data center LAN switches will need to support high densities of 10 GbE ports to provide connectivity for high performance virtualized servers, as well as an adequate number of 10 GbE ports (and 40 GBE and 100 GbE ports when these are available) for uplinks connecting the access switches to the core tier.

As noted, second generation LAN switches had fairly low port density. In contrast:

[1] Ethernet typically provides a 10x higher performance for a 3-4x increase in cost. This is an example of how Moore's Law impacts the LAN.

The current generation of switches has exploited advances in switch fabric technology and merchant silicon switch-on-a-chip integrated circuits (ICs) to dramatically increase port densities.

Modular data center switches are currently available with up to 256 non-blocking 10 GbE ports, while the typical maximum port density for stackable switches (generally based on merchant silicon) is 48 10 GbE ports. Today, high-speed uplinks are comprised of multiple 10 GbE links that leverage Link Aggregation (LAG)[2]. However, a 40 GbE uplink typically offers superior performance compared to a 4 link 10 GbE LAG. This is because the hashing algorithms that load balance traffic across the LAG links can easily yield sub-optimal load distribution where a majority of traffic is concentrated in a small number of flows. Most high performance modular switches already have a switch fabric that provide 100 Gbps of bandwidth to each line card, which means that when 40 GbE and 100 GbE line cards are available, these can be installed on existing modular switches, preserving the investment in these devices. Most vendors of modular switches expect to be shipping 40 GbE line cards by the middle of 2011, while 100 GbE line cards will take until 2012 or 2013.

In the case of stackable Top of Rack (ToR) switches, adding 40 or 100 GbE uplinks will generally require new switch silicon, which means that existing ToR switches will probably need to be swapped out in order to support the next generation of uplink speeds. The next generation of ToR switches with 40-48 10 GbE ports and 2-4 40 GbE ports are expected to be available in the second half of 2011.

High Availability

As previously noted, IT organizations will be implementing a growing number of VMs on high performance multi-core servers.

The combination of server consolidation and virtualization creates an "all in one basket" phenomenon that drives the need for highly available server configurations and highly available data center LANs.

One approach to increasing the availability of a data center LAN is to use a combination of redundant subsystems within network devices such as LAN switches in conjunction with redundant network designs. A high availability modular switch can provide redundancy in the switching fabric modules, the route processor modules, as well as the cooling fans and power supplies. In contrast, ToR switches are generally limited to redundant power supplies and fans. Extensive hardware redundancy is complemented by a variety of switch software features, such as non-stop forwarding, that ensure minimal disruption of traffic flow during failovers among redundant elements or during software upgrades. Modular switch operating systems also improve availability by preventing faults in one software module from affecting the operation of other modules.

[2] www.ieee802.org/3/hssg/public/apr07/frazier_01_0407.pdf

The bandwidth efficiency of Layer 2 networks 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. This can be accomplished by eliminating loops in the logical topology without resorting to the Spanning Tree Protocol (STP). In the current state of evolution toward a Third Generation data center LAN, loops can be eliminated using switch virtualization and multi-chassis LAG (MC LAG) technologies, which are described below.

Switch Virtualization

With switch virtualization, two or more physical switches are made to appear to other network elements as a single logical switch or virtual switch, with a single control plane.

In order for multiple physical switches to form a virtual switch, they need a virtual switch link (VSL) or interconnect (VSI) that supports a common control plane and data flows between the members of the virtual switch. In redundant configurations, connections between end systems and virtual access switches and between virtual access switches and virtual aggregation switches are based on multi-chassis (MC) link aggregation group (LAG) technology[3], as shown in Figure 4.1. MC LAG allows the links of the LAG to span the multiple physical switches that comprise a virtual switch. From the server perspective, links to each of the physical members of a virtual access switch appear as a conventional LAG or teamed links, which means that switches can be virtualized without requiring any changes in the server domain.

The combination of switch virtualization and multi-chassis LAG can be used to create a logically loop-free topology.

This means that data center LANs can be built without using the spanning tree protocol (STP) and first hop router redundancy protocols (e.g., VRRP). This is important because these protocols prevent all available forwarding resources in a redundant network design from being simultaneously utilized.

In Figure 4.1, loops are eliminated because, from a logical perspective, there are only two switches with a single LAG from the server to the access switch and a single LAG from the access switch to the aggregation switch. The traffic load to and from each server is load balanced across the two links participating in the multi-chassis LAG connecting each server to the virtual access switch. Therefore, both server connections are actively carrying traffic in both directions rather than being in an active state for some VLANs and in an inactive state for others. In the same fashion, traffic between the access virtual switch and the aggregation virtual switch is load balanced across all four physical links connecting these devices. Both physical switches participating in the aggregation layer virtual switch are actively forwarding traffic to the network core (not shown in the figure). The traffic is load balanced via the LAG hashing algorithms rather than being based on VLAN membership, as is the case with more traditional redundant

[3] http://en.wikipedia.org/wiki/Link_aggregation

LAN designs. The virtual switch not only improves resource utilization but also enhances availability because the relatively long convergence times of STP topology calculations are circumvented. Virtual switch technology also simplifies management because multiple physical switches can be managed as a single entity.

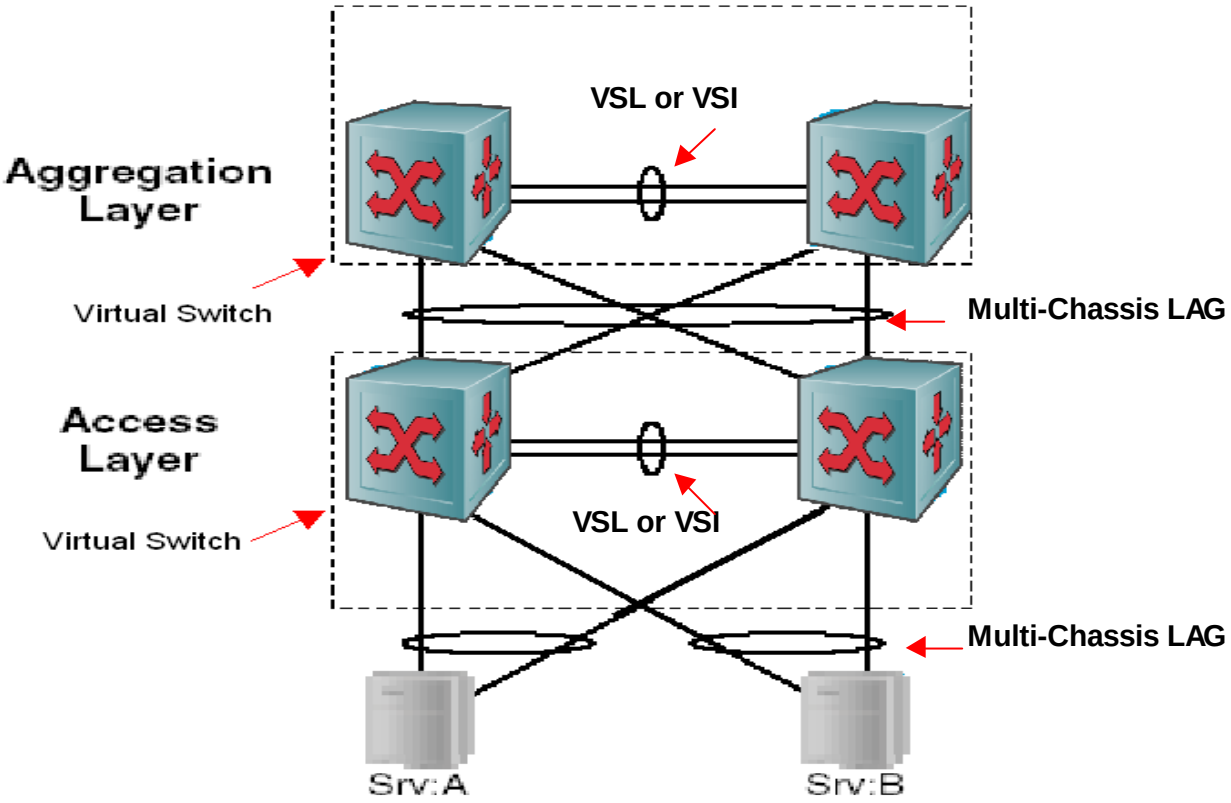


Figure 4.1: Switch Virtualization and Multi-Chassis LAG

Most vendors of data center switches support switch virtualization and MC LAG in their ToR and modular switches, and these technologies are fully utilized in the two-tier LAN designs that they are currently recommending to enterprise customers. As a result, most two tier LAN designs being proposed will not be based on STP for loop control. There are some differences among vendors in the VSL/VSI technology and in the LAG hashing algorithms. For example, some vendors of stackable ToR switches take advantage of the stacking interconnect as the VSL/VSI link, while other vendors will use 10 GbE or 40 GbE ports (when available) for VSL/VSI. Most LAG implementations conform to the IEEE 802.3ad standard. However, LAG hashing algorithms are outside the 802.3ad standard and more sophisticated hashing algorithms can provide differentiation by improving load balancing across LAG links. In addition, there are some differences in the number of ports or link that can participate in a LAG. Some vendors support up to 32 links per LAG, while 8 links per LAG is the most common implementation.

It must be noted that two-tier LANs and switch virtualization are far from the final word in the design of data center networks. Standards bodies have been working on

technologies that will allow active-active traffic flow and load balancing of Layer 2 traffic in networks of arbitrary switch topologies. TRILL (Transparent Interconnection of Lots of Links) is an Internet Engineering Task Force (IETF) project to develop a Layer 2 shortest-path first (SPF) routing protocol for Ethernet. A similar effort is being pursued by the IEEE 802.1aq working group which is defining a standard for shortest path bridging of unicast and multicast frames (based on the IS-IS protocol) and which supports multiple active topologies. With TRILL or 802.1aq, it would be possible to achieve load-balanced, active-active link redundancy without having to resort entirely to switch virtualization and VSL/VSI interconnects.

SPF bridging should have major implications for data center LAN designs and most of the larger switch vendors are well along in developing switches that can support SPF bridging and network designs based on these technologies. While the TRILL standard is possibly still a year or more away, some vendors are preparing pre-standard and proprietary enhancements for introduction well before the standard is ratified. It may well turn out that two-tier networks based on switch virtualization and MC LAG are just a mid-way point in the evolution of the Third Generation LAN.

With technologies like TRILL, the difference between access switches and core switches may shrink significantly.

As a result, the switch topology may shift from a two-tier hub and spoke, such as the one in Figure 4.1, to a highly meshed or even fully meshed array of switches that appears to the attached devices as a single switch. SPF bridging can support a variety of other topologies, including the fat tree switch topologies[4] that are popular in cluster computing approaches to HPC. Fat trees are also used by Ethernet switch vendors to build high density, non-blocking 10 GbE switches using merchant silicon switch chips. Figure 4.2a shows how a 48 port 10 GbE TOR switch can be constructed using six 24-port 10 GbE switch chips. By increasing the number of leaf and spine switches, larger switches can be constructed, The maximum density switch that can be built with a two-tier fat tree architecture based on 24 port switch chips has 288 ports. A number of high density 10 GbE switches currently on the market use this design approach.

[4] www.mellanox.com/pdf/.../IB_vs_Ethernet_Clustering_WP_100.pdf

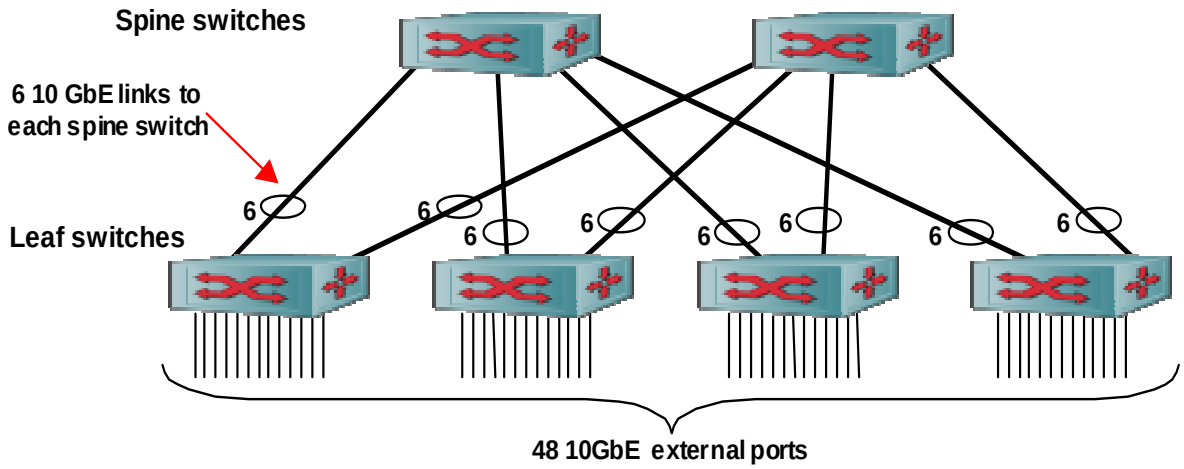


Figure 4.2a: TOR Switch Fat Tree Internal Architecture

With the emergence of TRILL/SPF bridging, a number of Ethernet switch vendors are adopting the fat tree as the topology for the 10 GbE data center LAN. With core/aggregation switches playing the role of spine switch and access switches playing the role of leaf switches. TRILL/SPF bridging allows traffic between two access switches to be load balanced across the numerous parallel paths through the core/aggregation layer switches (Equal Cost Multi-Path bridging). Using 72 48-port 10 GbE TOR switches as the common building blocks, a two- tier fat tree data center network can be built with 1,152 10 GbE ports. With 48 256 port 10 GbE switches, a two-tier data center fat tree network with 8,192 10 GbE ports would be possible, as shown in Figure 4.2b, The upper limit for two-tier fat trees based on 256 port switches is 32,768 ports using 384 switches.

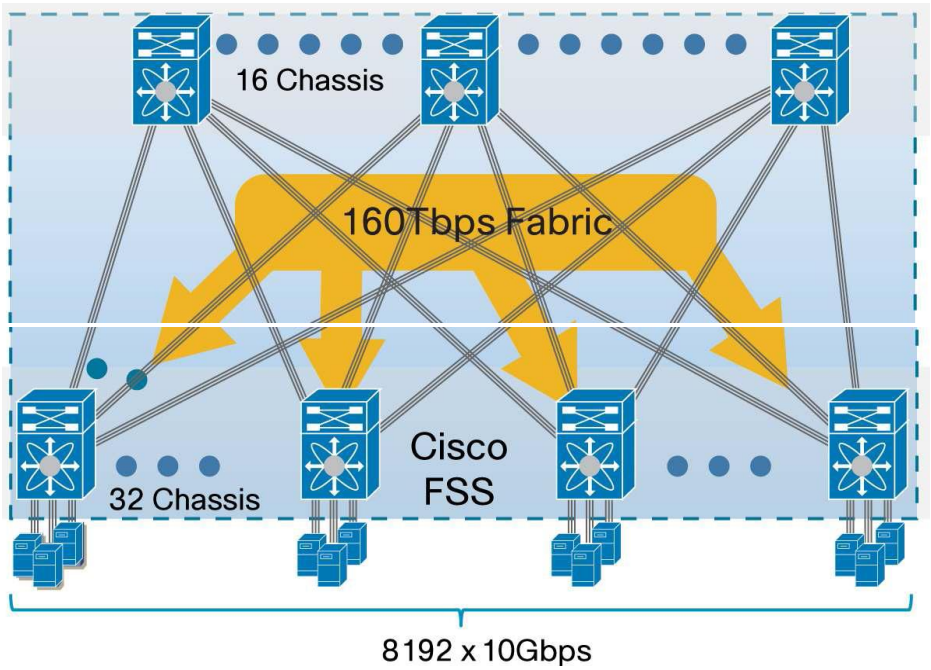


Figure 4.2b Two-Tier Fat Tree Data Center LAN with 8,192 10 GbE Ports (Source: Cisco Systems)

Controlling and Managing Inter-VM Traffic

With server virtualization, each physical server is equipped with a hypervisor-based virtual switching capability that allows connectivity among VMs on the same physical platform. Traffic to external destinations also traverses this software switch. From the network perspective, the hypervisor vSwitch poses a number of potential problems:

1. The vSwitch represents another tier of switching that needs to be configured and managed, possibly requiring an additional management interface. This can partially defeat an effort to flatten the network to two–tiers.

2. The vSwitch adds considerable complexity, because there is an additional vSwitch for every virtualized server.

3. vSwitch control plane functionality is typically quite limited compared to network switches, preventing a consistent level of control over all data center traffic

4. As more VMs per server are deployed, the software switch can place high loads on the CPU, possibly starving VMs for compute cycles and becoming an I/O bottleneck.

5. VM-VM traffic on the same physical server is isolated from the rest of the network, making these flows difficult to monitor and control in the same fashion as external flows.

6. The vSwitch functionality and management capabilities will vary by hypervisor vendor and IT organizations are increasingly deploying hypervisors from multiple vendors.

   The vSwitch presents a number of concerns related to management, security, functionality and organizational responsibilities.

There are two approaches to the problems posed by the early generation vSwitch: Distributed Virtual Switching (DVS) and Edge Virtual Bridging (EVB). With DVS, the control and data planes of the embedded vSwitch are decoupled. This allows the data planes of multiple vSwitches to be controlled by an external centralized management system that implements the control plane functionality. Decoupling the data plane from the control plane makes it easier to tightly integrate the vSwitch control plane with the control planes of physical access and/or aggregation switches and/or the virtual server management system. Therefore, DVS can simplify the task of managing a large number of vSwitches, and improve control plane consistency, but it doesn't address the other issues listed above.

With EVB, all the traffic from VMs is sent to the network access switch. If the traffic is destined for a VM on the same physical server, the access switch returns the packets to the server over the same port on which it was received (e.g., a "hair pin turn"). With Edge Virtual Bridging, the hypervisor is relieved from all switching functions, which are

now performed by the physical access network. With EVB, the vSwitch now performs the simpler function of aggregating hypervisor virtual NICs to a physical NIC. Basic EVB can be supported by most existing access switches via a relatively simple firmware upgrade. The IEEE 802.1bg Working Group is creating an EVB standard based on a technology known as Virtual Ethernet Port Aggregator (VEPA) dealing with hair-pin turns and a definition of a multi-channel service for remote ports to access local VMs. A companion effort, IEEE 802.1bh Port Extension is defining a tagged approach to deal with frame replication issues in the EVB. EVB/VEPA standards supported in switches and hypervisors will address all of the issues listed above.

Essentially all vendors of data center switches support the IEEE's EVB standards efforts. Some vendors are waiting until the standard is finalized and are supporting hypervisor vSwitches in the interim. Other vendors have pre-standard implementations of basic EVB/VEPA already available or under development.

Network Convergence/Fabric Unification

In contrast to Second Generation Data Center LANs:

> ***A key characteristic of Third Generation Data Center LANs will be the convergence of block-level storage and data traffic over a common high-speed Ethernet data center switching fabric.***

This unified fabric offers significant cost savings in multiple areas including converged network adapters on servers and reductions in rack space, power and cooling capacity, cabling, and network management overhead.

Traditional Ethernet, however, 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. In order to emulate the lossless behavior of a Fibre Channel (FC) SAN, Ethernet needs enhanced flow control mechanisms that eliminate buffer overflows for high priority traffic flows, such as storage access flows. Lossless Ethernet will be based on the following standards, which are commonly referred to as IEEE Data Center bridging (DCB):

.

- **IEEE 802.1Qbb Priority-based Flow Control (PFC)** allows the creation of eight distinct virtual link types on a physical link, with each virtual link mapped to an 802.1p traffic class. Each virtual link can be allocated a minimum percentage of the physical link's bandwidth. Flow is controlled on each virtual link via the pause mechanism which can applied on a per priority basis to prevent buffer overflow, eliminating packet loss due to congestion at the link level. In particular, block-level or file-level storage traffic on one of the virtual lanes can be protected from loss by pausing traffic on one or more of the remaining lanes.

- **IEEE 802.1Qau Congestion Notification (CN)** is a traffic management technique that eliminates congestion by applying rate limiting or back pressure at the edge of

the network in order to protect the upper network layers from buffer overflow. CN is intended to provide lossless operation in end-to-end networks that consist of multiple tiers of cascaded Layer 2 switches, such as those typically found in larger data centers for server interconnect, cluster interconnect and extensive SAN fabrics.

- **IEEE 802.1Qaz Enhanced Transmission Selection (ETS)** will specify advanced algorithms for allocation of bandwidth among traffic classes including the priority classes supported by 802.1Qbb and 802.1Qau. While the queue scheduling algorithm for 802.1p is based on strict priority, ETS will extend this by specifying more flexible drop-free scheduling algorithms. ETS will therefore provide uniform management for the sharing of bandwidth between congestion managed classes and traditional classes on a single bridged network. Priorities using ETS will coexist with priorities using 802.1Qav queuing for time-sensitive streams. ***Data Center Bridging Exchange (DCBX)*** protocol is also defined in the 802.1Qaz standard. DCBX protocol is an extension of the Link Layer Discovery Protocol (LLDP) that allows neighboring network elements to exchange request and acknowledgment messages to ensure consistent DCB configurations. DCBX is also used to negotiate capabilities between the access switch and the adapter and to send configuration values to the adapter.

- **IEEE 802.1aq Shortest Path Bridging (SPF)** is a standard for shortest path bridging of unicast and multicast frames (based on the IS-IS protocol) supporting multiple active topologies. SPF is part of the IEEE standards efforts relative to the data center, but is not strictly required for standards-compliant lossless Ethernet.

DCB Lossless Ethernet will play a key role in supporting Fibre Channel over Ethernet (FCoE) technology that will allow the installed base of Fibre Channel storage devices and SANs to be accessed by Ethernet-attached servers (with converged FCoE network adapters) over the unified data center switching fabric. DCB will benefit not only block-level storage, but also all other types of loss and delay sensitive traffic. In the storage arena, DCB will improve NAS performance and will make iSCSI SANs based on 10/40/100 GbE a more competitive alternative to Fibre Channel SANs at 2/4/8 Gbps. In order to take full advantage of 10 GbE and higher Ethernet bandwidth, servers accessing iSCSI storage resources may also need intelligent converged NICs that offload iSCSI and TCP/IP processing from the host.

All data center vendors are planning to support the DCB standards when they are available. In some cases the timing of the availability of that support may differ between access and core switches. In addition, some vendors are offering pre-standard support for DCB capabilities, including PCF, CN, ETS, and DCBX.

Fibre Channel over Ethernet (FCoE) is an industry standard that is being developed by the International Committee for Information Technology Standards (INCITS) T11 committee.

The FCoE protocol specification maps Fibre Channel upper layer protocols directly over a bridged Ethernet network. FCoE provides an evolutionary approach to the migration of

FC SANs to an Ethernet switching fabric while preserving Fibre Channel constructs and providing reliability, latency, security, and traffic management attributes similar to those of native FC. FCoE also preserve investments in FC tools, training, and SAN devices (e.g., FC switches and FC attached storage). Implementing FCoE over a lossless Ethernet fabric requires converged server network adapters (CNAs with support for both FCoE and IP) and some form of FC Forwarding Function (FCF) to provide attachment to native FC devices (FC SAN switches or FC disk arrays). FCF functionality can be provided by a FCoE switch with both Ethernet and FC ports or by a stand alone gateway device attached to a FCoE passthrough switch, as shown in Figure 4.3.

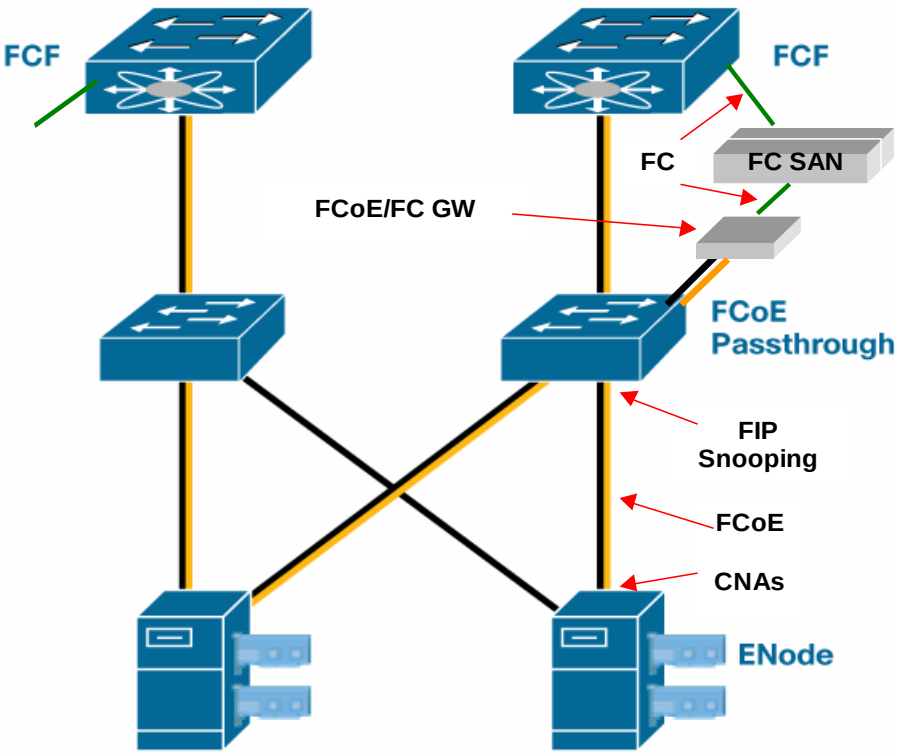


Figure 4.3: FCoE Converged LAN (Source: Cisco Systems)

As shown in the figure, End Nodes (servers) do not need to connect directly to a FCF capable switch. Instead the FCoE traffic can pass through one or more intermediate FCoE passthrough switches. The minimal requirements for a simple FCoE passthrough switch is support for lossless Ethernet or DCB. The FCoE Initialization Protocol (FIP) supports handshaking between a FCoE End Node and an FCF in order to establish and maintain a secure virtual FC link between these devices, even if the end-to-end path traverses FCoE passthrough switches. For DCB passthrough switches that support FIP Snooping, the passthrough switches can inspect the FIP frames and apply policies based on frame content. FIP Snooping can be used to enhance FCoE security by preventing FCoE MAC spoofing and allowing auto-configuration of ACLs.

As this discussion illustrates:

There are several levels of support that data center switch vendors can provide for FCoE.

For example:

1. The lowest level of support is FCoE passthrough via lossless Ethernet or DCB alone.

2. The next step up is to add FIP Snooping to FCoE passthrough switches

3. A third level of support is to add a standalone FCF bridges/gateways to front end FC SAN switches or disk arrays.

4. The highest level of support is to provide DCB and FIP Snooping for FCoE passthrough switches and also to provide FCoE switches that incorporate FCF ports, creating hybrid switches with both DCB Ethernet and native FC ports.

Most vendors of Ethernet data center switches that do not also have FC SAN switches among their products are planning FCoE support at levels 1, 2, or 3 described above. In fact, most of these Ethernet-omly vendors are considerably more enthusiastic about iSCSI SANs over 10/40/100 GbE than they are about FCoE.

The primary drivers of FCoE are the vendors that offer both Ethernet and FC products.

These are the vendors that are already shipping pre-standard lossless 10 GbE Ethernet switches and hybrid lossless 10 GbE/FCF switches. Even among the vendors providing early support for FCF there are some significant differences, as shown in Figure 4.4.

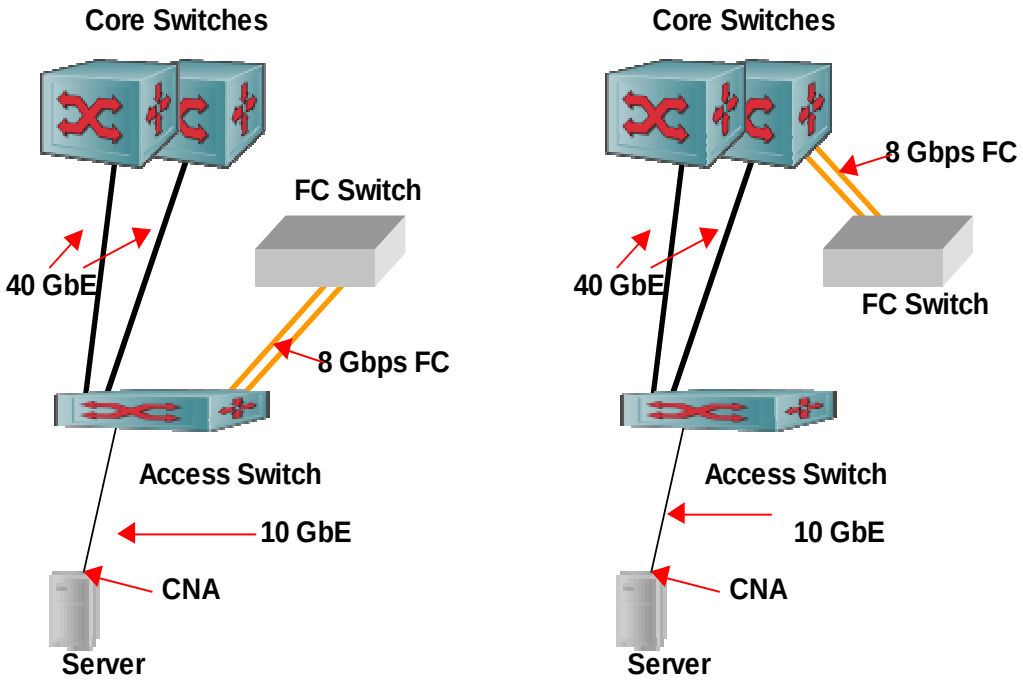


Figure 4.4: FCF Support Options

The left side of the figure shows single hop FCoE with the FCF function integrated into the access switch. It would also be possible to use intervening FCoE/FCF gateways (either standalone or incorporated in the FC switch) which would be connected to the access switch via 10 GbE, making the access switch an FCoE passthrough switch, as shown in the previous figure. The advantage of single hop FCoE is that the storage traffic doesn't compete for bandwidth in the uplinks or the core switches and the core switches aren't required to support DCB or FIP Snooping. The right side of the figure shows multihop FCoE with the FCF function integrated into the core switch, and the access switch in FCoE passthrough mode. Again it would be possible to use FCoE/FCF gateways (either standalone or incorporated in the FC switch) connected to the core switch via 10 GbE. FC SANs and disk arrays connected at the core offer the advantage of a more centralized pool of storage resources that can be shared across the data center LAN.

Network Support for Dynamic Creation and Migration of VMs

As noted earlier, the requirement to support the dynamic creation and movement of VMs is one of the primary factors driving IT organizations to redesign their data center LANs. As noted earlier, the requirements for VM migration with VLAN boundaries has provided a major impetus for flattening the LAN with two-tier designs featuring Layer 2 connectivity end-to-end.

Many of the benefits of cloud computing depend on the ability to dynamically provision VMs and to migrate them at will among physical servers located in the same data center or in geographically separated data centers. The task of creating or moving a VM is a relatively simple function of the virtual server management system. There can, however, be significant challenges in assuring that the VM's network configuration state (including VLAN memberships, QoS settings, and ACLs) is established or transferred in a timely fashion. In many instances today, these network configuration or reconfigurations involves the time-consuming manual process involving multiple devices.

Regulatory compliance requirements can further complicate this task. For example, assume that the VM to be transferred is supporting an application that is subject to PCI compliance. Further assume that because the application is subject to PCI compliance that the IT organization has implemented logging and auditing functionality. In addition to the VM's network configuration state, this logging and auditing capability also has to be transferred to the new physical server.

The most common approach to automating the manual processes involved in VM provisioning and migration is based on communication between the Hypervisor Management system and the switch element management system (EMS) via APIs supported by both vendors.

When a Virtual Machine is created or when a virtual machine move is initiated, the Hypervisor manager signals to the EMS that the event is about to occur and provides a partial VM network profile including a virtual MAC, VLAN memberships, and the target hypervisor. Based on existing policies, the EMS extends the VM network profile to include appropriate QoS and security parameters such as ACLs. The EMS can then determine the target hypervisor's access switch and configure or reconfigure it accordingly. Where VLANs need to be created, the EMS can also create these on the uplinks and neighboring switches, as appropriate. In a similar manner, when a VM is deleted from a hypervisor, the EMS can remove the profile and then prune the VLAN as required. All of these processes can be triggered from the hypervisor.

Most data center switch vendors have already implemented some form of VM network profile software, including linking their switches to a least one brand of hypervisor. Some differences exist between the range of hypervisors supported and the APIs that were used. Distribution of VM network profiles is only one of many management processes that can benefit greatly from automation, so it would benefit IT departments to develop expertise in open APIs and powerful scripting languages that can be exploited to streamline time-consuming manual processes and thereby reduce operational expense while improving the ability of the data center to dynamically reallocate its resources in response to changes in user demand for services.

A somewhat different approach to automating data center configuration, including the provisioning and migration of VMs is based on Orchestration engines, which are discussed in more detail in the management section of this report. Service orchestration is a centralized server function that can automate many of the manual tasks involved in

provisioning and controlling the capacity of dynamic virtualized services. In the case of VM provisioning and migration, the Orchestration engine would function as the point of integration between the network device's EMS and the hypervisor management system. Orchestration solutions are available from a number of network management vendors and hypervisor vendors.

The data center LAN is on the cusp of a number of quite dramatic technology developments, as summarized in Table 4.4. As shown in the table, most the items on this list are still in flux and require additional development, and/or additional work from the standards bodies[5].

[5] Exceptions to this statement are entries number 1, 2, 4 and to some extent 11.

Technology Development	Status
1: Two-tier networks with Layer 2 connectivity extending VLANs across the data center.	On-going deployment
2: Reduced role for blade switches to eliminate switch tier proliferation.	On-going
3: Changing role for the hypervisor vSwitch as a port aggregator (VEPA) for EVB, essentially eliminating the vSwitch tier.	A standard is in progress. Pre-standard implementations are occurring.
4: STP displaced by switch virtualization and multi-chassis LAG technology.	On-going deployment
5: Multi-core servers with notably more VMs per server and 10 GbE connectivity to the LAN.	Early adoption stage.
6: 40 GbE and 100 GbE uplinks and core switches.	A standard is in place: 40 GbE due in 2011 100 GbE due in 2012
7: DCB delivering lossless Ethernet for 10 GbE and higher speed Ethernet	Standards are in progress. Pre-standard implementations are occurring.
8: FCoE approach to fabric unification	FCoE standard is in place. Early implementations of FCoE are occurring over pre-standard DCB.
9: 10 GbE iSCSI approach to fabric unification	Early implementations over pre-standard DCB.
10: TRILL/SPB enabling new data center LAN topologies; e.g., fully meshed, fat tree.	Standards are in progress. Pre-standard implementations are imminent.
11: Management tools that integrate, coordinate, and automate provisioning and configuration of server, storage and network resource pools	These are proprietary and have varying levels of maturity.

Table 4.4 Status of Data Center Technology Evolution

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.

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

The Network Platform for Cloud

Introduction

From Cisco's perspective, cloud is a model in which IT resources and services are abstracted from the underlying infrastructure and provided on demand and at scale in a multi-tenant environment.

ROLE OF THE NETWORK PLATFORM IN CLOUD

Access to Critical data, Services, Resources, and People

- Core fabric connects resources within the data center and data centers to each other
- Pervasive connectivity links users and devices to resources and each other
- Network provides identity- and context-based access to data, services, resources, and people

Granular Control of Risk, Performance, and Cost

- Manages and enforces policies to help ensure security, control, reliability, and compliance
- Manages and enforces SLAs and consistent QoS within and between clouds, enabling hybrid models and workload portability
- Meters resources and utilization to provide transparency for cost and performance

Robustness and Resilience

- Supports self-healing, automatic redirection of workload and transparent rollover
- Provides scalability, enabling on-demand, elastic computing power through dynamic configuration

Innovation in Cloud-Specific Services

- Context-aware services understand identity, location, proximity, presence, and device
- Resource-aware services discover, allocate, and pre-position services and resources
- Comprehensive insight accesses and reports on all data that flows in the cloud

- "On demand" means that resources can be provisioned immediately when needed, released when no longer required, and billed only when used.
- "At scale" means the service provides the experience of infinite resource availability to meet whatever demands are made on it.
- "Multi-tenant environment" means that the resources are provided to many consumers - for example, business units - from a single implementation.

Role of the Network

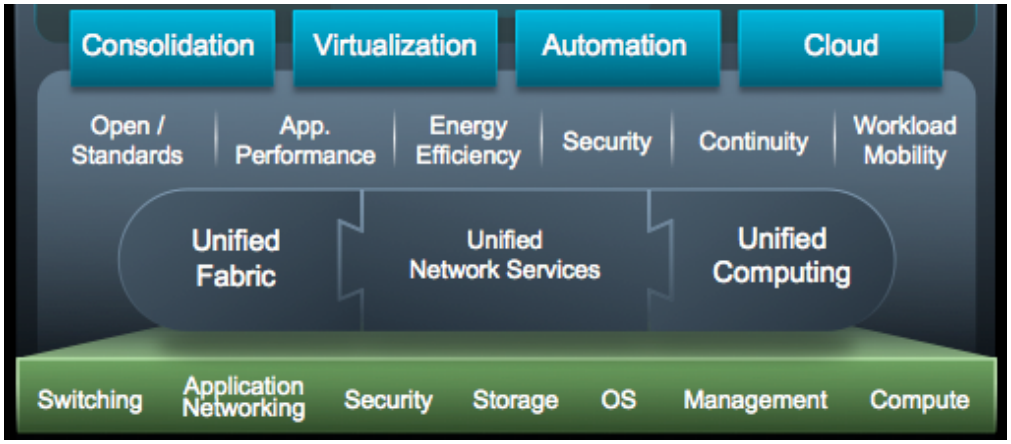
With users, devices and partners accessing virtualized resources and applications within the data center, the network is the essential platform for accessing and delivering cloud computing models. This includes the network in the cloud data center, the network between data centers, and the network connecting users from around the world.

This ubiquity creates a unique opportunity to build and take advantage of capabilities that can be delivered from the network to drive greater value out of cloud infrastructures.

Today's networks are already adopting key innovations for cloud computing: 10Gb Ethernet, WAN and application acceleration, Virtual Machine (VM) level traffic awareness, and enablement of VM mobility within and across data centers.

Cisco's networking capabilities align with three technology pillars: Unified Computing, Unified Fabric, and Unified Network Services. Together, these pillars are woven into Cisco's new Data Center Business Advantage architectural framework enabling enterprises to go from simple system consolidation and virtualization through to enabling infrastructure automation and secure private cloud deployment.

Figure 1. Network Capabilities for Cloud from Cisco's Data Center Business Advantage Architectural Framework



Unified Computing

Cisco's Unified Computing System (UCS) aims to provide scalable, dynamic compute resources for open, physical and virtualized environments. It does this by bringing together compute, network and storage access with virtualization to deliver better resource utilization, operational simplicity and workload mobility. It leverages the network intelligence and scale of Unified Fabric and the service readiness of the Unified Network Services.

UCS brings several innovative capabilities to data center servers, including:

- Extended memory technology allowing very dense VM hosting with up to 384GB of RAM per blade.
- Complete hardware abstraction through server profiles that allow mapping of configurations to the stateless compute blades in minutes.
- Native 10Gb Fiber Channel over Ethernet (FCoE) support.
- High Performance Virtual I/O (Ethernet NIC and FC HBA).
- Open, XML-Based API to provision, orchestrate and manage the UCS system.

Unified Fabric

Unified Fabric provides a simplified and integrated physical network for *all* I/O and communications in the cloud, including data, storage, voice and video. The fabric provides a converged network at scale with embedded intelligent capabilities that enable cloud.

With the widespread deployment of 10Gb Ethernet technology today, a roadmap to 40 and 100 Gb speeds, and the ratification of FCoE standards, Cisco views Ethernet as the fundamental layer for a unified fabric that can support multiple types of storage and data traffic simultaneously.

In addition to traffic within a data center, the unified fabric concept includes the extension of networks across facilities or geographic locations and the capabilities required to enable workload mobility.

Cisco delivers Unified Fabric across the breadth of its data center portfolio, including but not limited to the following:

- Unified Fabric in data center switching, from the hypervisor to the core with the Nexus 1000v, 2000, 5000 and 7000 series, interconnected with storage networks on MDS switches—all leveraging the consistent data center class operating system NX-OS.
- Cisco FabricPath Switching System (FSS) enabling broad Layer 2 data center networks, expanded VM mobility and efficient use of all available network bandwidth.
- Cisco Overlay Transport Virtualization (OTV) allowing Layer 2 continuity between geographically dispersed networks over any transport that supports IP, which in turn enables live migration of VMs between networks, data centers, and clouds.

Unified Network Services

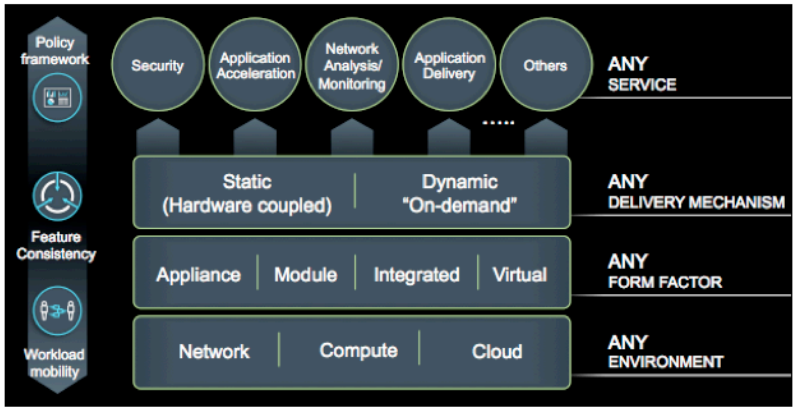
Unified Network Services (UNS) is architected deliver services such as WAN optimization, firewalls, and load balancing in a concerted way across traditional appliances, inside other network devices such as switches and routers, and as virtualized appliances delivered on a hypervisor. This pillar executes a simple vision: to deliver any network service (security, WAN optimization, application delivery and load balancing, etc.), in any form factor (physical, virtual, appliance, integrated), in any environment (network, compute) and with any delivery mechanism (hardware-coupled or dynamic on-demand).

In addition to the industry leading physical appliances and network services that are embedded in different switches and routers either in software or via network modules, Cisco is now tapping into a new inflection point in the data center with the introduction of virtualized network services as part of Unified Network Services.

Cisco VSG works with the Cisco Nexus 1000v virtual switch's vPath capability and the Cisco Virtual Network Management Center (VNMC) to:

- Secure segmentation with zone-based firewall.
- Provide VM-level traffic visibility and granularity with context-aware rules.

Figure 2. Cisco's Unified Network Services Vision



Policy-based centralized management. vWAAS is the industry's first cloud-ready WAN optimization solution. vWAAS works with the Cisco Nexus 1000v virtual switch's vPath capability to:

- Enable on-demand orchestration and policy-based application of rules down to the level of specific VMs.
- Provide separation of compute and storage with cache stored on SAN.
- Support multi-tenancy for cloud providers.
- Designed for optimizing traffic between and to clouds, both within the enterprise and from service providers.

Open Ecosystem and Market Success

Cisco's Data Center Business Advantage architecture is committed to delivering best-of-breed, open-standard networking solutions for cloud. Leveraging technology innovation and new delivery models, Cisco is giving customers greater choice than they've ever had within the Data Center.

- 11x World Record performance – Cisco Unified Computing System.
- 3x "Best of VMworld" winner (Cisco UCS, Nexus 1000v, Cisco OTV).
- Over 1.5 million 10Gb Ethernet ports shipped on Nexus switches.
- Over 40 ISV partners leveraging the UCS-API.
- VCE Coalition (VMware, Cisco, EMC) Vblock Infrastructure Packages.
- IVA Alliance (VMware, Cisco, NetApp) SMT Architecture.
- Cisco, Citrix and NetApp VDI Architecture.
- Application partnerships with Microsoft, Oracle, SAP and many others.
- Management partnerships with BMC, CA and many others.

For More Information

As you begin your own journey to the cloud, we invite you to discuss the right approach for your organization with Cisco. For additional information about:

Cloud: http://www.cisco.com/go/cloud

Data Center Business Advantage: http://www.cisco.com/go/dcba

Unified Computing: http://www.cisco.com/go/unifiedcomputing

Unified Fabric: http://www.cisco.com/go/unifiedfabric

Unified Network Services: http://www.cisco.com/go/unifiednetworkservices

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Open Network Automation is Critical to the Virtual Data Center

Authored by Stephen Garrison,
Vice President Marketing,
Force10 Networks, Inc.

The Evolving Data Center

The data center has undergone several significant transformations since the birth of computing. The data center has evolved from mainframe computing to client server to Internet computing to SOA. Now we sit on the precipice of another major technology shift – the move to a fully virtualized data center (Figure 1). With each transition, the cost of computing was driven down by orders of magnitude and organizations were able to increase the efficiency of data center operations, software development, and most importantly, corporate workers.

Figure 1. Computing through the ages

The shift to a virtual data center will be the most significant IT transformation since the invention of the mainframe as it promises to bring together the network stack, storage and the computing layer to optimize application performance. In a fully virtualized data center, compute resources exist as VMs (virtual machines), storage becomes virtualized "pools" that can exist anywhere, and the network fabric connects these virtual elements to form a flexible, scalable computing environment (Figure 2).

The use of virtualization technology is widespread. A recent enterprise survey revealed that 82% of organizations today are using virtualization technology[1]. The primary driver for almost all companies using virtualization is to consolidate the number of servers. Obviously, this can have a huge impact on TCO since the number of servers can be dramatically reduced, sometimes by a factor of 10.

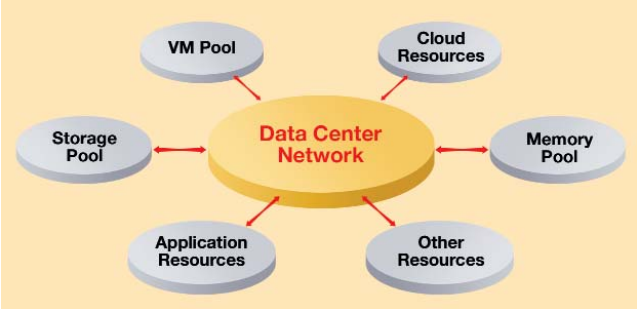


Figure 2. The virtualized data center is connected by the network

However, there are many other reasons for deploying virtualization technology, including:

- It allows software developers or other users to provision their own virtual machines. This will allow developers, engineers or others to have near instantaneous access to compute resources without having to involve several other departments.
- It ensures application performance is maintained when the workload is increased by provisioning additional computing resources.
- It increases the uptime of applications by mobilizing virtual workloads. In the event of an outage, the VM can move across a rack, across the data center or across the network whenever required.
- It acts as the bridge between physical and cloud-based data centers. Resources must be virtualized if they are to easily migrate between private data centers and cloud-based data centers.

The adoption of new technology always creates new challenges for data center managers, and virtualization is no exception. While server consolidation can dramatically reduce the number of physical servers, an unfortunate side-effect is that it results in an explosion in the number of virtual machines. Managing this so-called "sprawl" of virtual machines is much more difficult than managing physical resources. As organizations move from hundreds of VMs to thousands, questions such as "Where is that VM?", "Who created that VM?", "Who owns that VM",

[1] *Yankee Group Survey 2010*

"Why did it migrate?" and "Where is the data?" become more common. This new complexity results in additional work for server administrators as they shift their workload from managing tens or hundreds of physical servers to managing hundreds to thousands of virtual machines.

But the challenge does not stop there. With virtual machines, data center managers must also provision virtual storage pools and virtual network resources. In earlier times, managing the computing environment, which consisted of a static stack of compute, network and storage resources, was much simpler. But with virtual compute, storage and network resources, complexity has dramatically increased, resulting in more work for system, network and storage administrators.

The Role of Automation

The solution to the additional complexity caused by the extensive use of virtualization in the data center is automation. Automation will play an important role in helping data center engineers better manage virtual resources. Without automation, data center managers need to manually re-provision and optimize server, storage and network resources every time the smallest change in the environment is made. Keeping all of the virtual resources in sync is a near-impossible task for any data center of significant size. In fact, only 17% of respondents polled in Yankee's recent survey[2] feel that the tools to virtualize mission critical applications exist today. This leaves a big gap between the vision of the fully virtualized data center and the current market reality.

The challenge associated with managing a virtual environment is not limited to just deploying new technology, as data center operations and organizational structure are also impacted in a significant way. Today, most large data centers have administrative staff for supporting server, network and storage resources (Figure 3), and each of these groups have expertise in managing their respective technology. Prior to the adoption of virtualization technology, these groups could successfully operate in what were essentially independent groups. But the adoption of virtualization, combined with the need to quickly shift resources as demanded by the business, is now requiring these groups to work closely with each other.

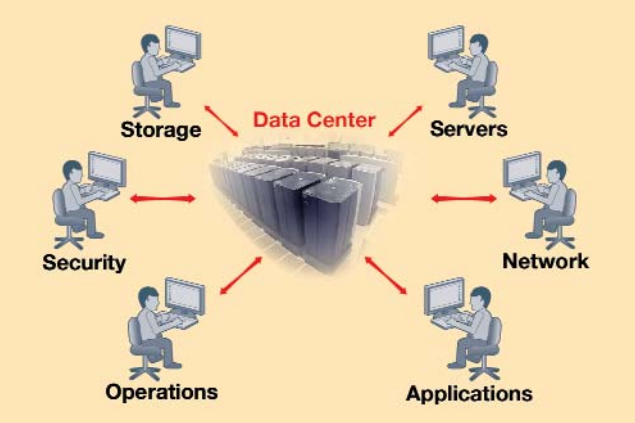


Figure 3. The virtual data center is everyone's responsibility

Automation

The additional complexity caused by the explosion of VMs, the need to tightly coordinate the provisioning of virtual resources, and the organizational challenges of managing this new virtual environment are best solved by automation. Automating the monitoring, management and provisioning of common tasks can greatly reduce the additional workload caused by virtual environments. Automation can also help standardize data center configurations, enforce best practices and increase availability.

For the network, automation can improve data center operations in the following ways:

- Instantly adjusts to changes in data flows, without manual reconfiguration, to optimize application performance. Virtualization, cloud computing, web 2.0 and other trends have given rise to bursty and unpredictable traffic flows. A congestion free network that provides non-blocking switching and routing performance can reduce the end to end latency of the transaction. This will also lead to the flat, layer 2 network that VMotion requires.
- Delivers an "always on" data center fabric. A high capacity, modular, fully redundant network can shift resources almost instantly to withstand any outage. Additionally, the network architecture can be simplified by increasing the density of the ports in the network devices. This means less hardware, a simpler architecture and increased uptime.
- Provides on demand resource allocation through automated network reconfiguration. The network can adhere to any business SLA (service level agreement) to automate tasks such as reallocating resources by moving VLANs, changing priorities through QoS policies, reallocation of bandwidth or reducing power consumption by shutting off underutilized resources.

[2] *Yankee Group Survey 2010*

Because the network is at the heart of the virtual data center, it is unique in its ability to enable organizations to maximize their investments in virtualization and cloud architectures.

The Different Approaches to Network Automation

The goal of network automation is to provide a self-optimized network that is capable of dynamically allocating virtual resources to where they are needed in a timely fashion. Several approaches to network automation have emerged, and data center architects, CIOs and others involved in designing virtual data centers need to be aware of the differences. The network vendors can be seen as falling into one of three high-level approaches:

Approach 1: Integrated Network Automation

This approach involves the vendor adopting a highly integrated, proprietary architecture that requires the customer to source all elements in the stack from a single vendor, or closed system of vendors. The upside of this "vertically integrated" approach is that it delivers a solution that works "out of the box", so there is some short-term benefit. Long-term however, this approach means vendor lock-in, which deprives customers of the power to choose the best technology for their specific environment. To date, Cisco has chosen to adopt this approach.

Approach 2: Network Controlled Automation

In this environment, the monitoring, management and provisioning of virtual environments is controlled from, or by, the network. When a new virtual environment is required, or if an existing virtual environment needs more resources, network management tools provision the network, compute and storage resources. This is a network-centric strategy that requires all of the data center functions to fall under the control of the network rather than working in a cooperative manner. This requires a huge cultural and operational shift by data center managers. This approach has been adopted by Brocade and Extreme Networks.

Approach 3: Open Network Automation

The third approach toward network automation is one that leverages open standards that allow the data center network fabric to be controlled by existing automation or middleware tools. Because this approach is server and application centric, it is consistent with current data center operations, allowing an organization to adopt network automation more seamlessly because current best practices can remain in place. With an open strategy, the network infrastructure aids the operations of the virtual data center but doesn't take on the role of managing the virtual environment. Managing the virtual environment is done by existing virtualization management or system management tools designed for this express purpose. Additionally, standards based protocols are used for exchanging information between the network fabric and hypervisors or virtual switches to manage network configurations. This allows companies to choose best-of-breed technologies and still have the assurance that the solution will work. The open, standards based approach to network automation provides the best long-term benefits for the customer, as it retains the current data center operational structure but still provides a path to the future. Force10 Networks is an example of a vendor that utilizes this approach

What to Look for in a Solutions Provider

As network automation continues to evolve, more and more vendors will claim to have solutions that can help an organization make the transition to a virtual data center. Considering the important role the network will play in the evolution of the data center, it is critical that the following be considered when making a purchase decision:

- An open, standards based approach. There are many solution providers that claim to be open and many that claim to be standards-based. However, it is crucial that the network truly be both. Some vendors that claim to be both will actually be including a number of proprietary features that are "based on standards".
- Hypervisor, virtual switch and server agnostic. If this isn't the case, the organization may lose its choice in compute platforms. Considering the rate of innovation and the reach of virtualization, it's important the network be able to support any of the hypervisor vendors.
- Non-blocking, congestion free architecture. This will minimize the end-to-end latency of traffic flowing across the network. Solutions that are "near non-blocking" or over-subscribed could lead to congestion problems that impair the performance of applications.

- Future proofed technology – high density, 40 GbE and 100 GbE ready. The network infrastructure being purchased today should be thought of as a five year investment. So, the hardware being procured needs to provide sufficient density to allow simplification of the network and upgradability to both 40 and 100 GbE. This will avoid a rip and replace event in the future.
- A vendor with a history of data center innovation. Networking in the data center has many demands that are unique. Choose a vendor that understands the demands placed on the data center network. Vendors who grew in the wiring closet may not have the right culture to meet the challenges of a data center.
- A broad ecosystem of partners. No single vendor can deliver on the vision of the virtual data center. The network solution provider used should have solutions that work with all of the major compute, virtualization, storage and management vendors.
- A solution provider that utilizes common scripting languages. Data center operations today are driven by scripts written in perl, python and UNIX. A network vendor that utilizes the de facto standard scripting tools can help bridge the gap between networking and computing more efficiently and more quickly.

Conclusions

The data center is on the verge of another major transition – the shift to a fully virtualized data center This will lower the cost of computing, improve uptime and application performance and raise corporate productivity to new heights. However, along the way, data center managers will encounter new challenges in managing a data center built on pools of virtual resources instead of physical ones.

Open network automation can help meet many of these challenges by delivering a network that works with the compute infrastructure to automate many of the mission critical, time sensitive tasks needed to run a virtual data center. Open network automation will:

- Enable a virtual infrastructure that can scale to handle unpredictable traffic demands.
- Create an elastic environment where virtual resources can be allocated where and when they are needed based on business policy.
- Improve application uptime by instantly adapting and applying network configuration changes that arise due to changes in the compute environment.
- Provide a bridge to cloud computing by allowing companies to coordinate the movement of resources to the cloud at their own pace.
- Help move customers towards the vision of a virtual data center much faster than solutions that use vertically integrated technology.

Force10 Networks, Inc.
350 Holger Way
San Jose, CA 95134 USA
www.force10networks.com

408-571-3500 PHONE
408-571-3550 FACSIMILE

WP39SG 1010 v1.2

JUNIPER NETWORKS SOLUTION FOR CLOUD COMPUTING

Juniper Networks is dedicated to building simplified, scalable, agile, and secure networks that deliver the best performance and greatest efficiencies for cloud-ready data centers, while simultaneously controlling costs.

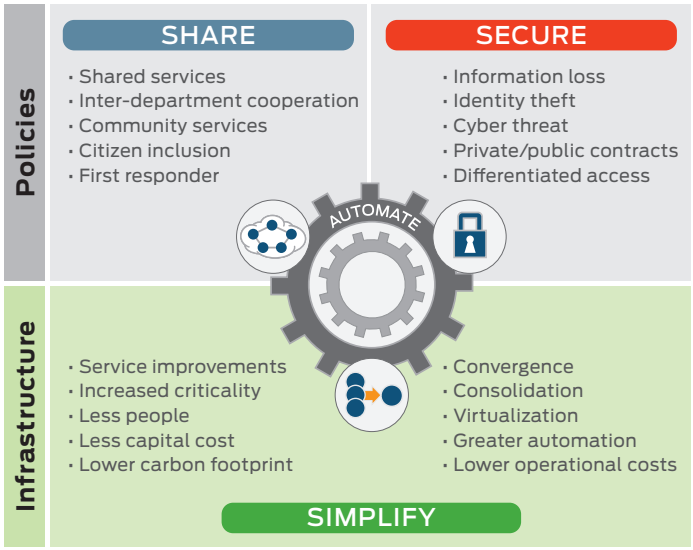
Getting Ready

Success in building a cloud-ready data center network requires three steps: (1) simplify, (2) share, (3) secure, as well as applying automation for smoother operations at each step. Whether you are running your internal IT infrastructure to be cloud-like or plan to connect with public cloud services, designing a cloud-ready data center network gives you significant advantages that can help you lower costs, increase efficiency, and keep your data center agile enough to accommodate any changes in your business or your technology infrastructure.

Key Components

Juniper recommends three steps to make your network infrastructure cloud-ready, reducing the cost and complexity of networking while improving application and business performance:

- **Simplify the architecture** — Consolidate siloed systems and collapse inefficient tiers using innovative fabric technology and a single network operating system. This results in fewer devices, a smaller operational footprint, reduced complexity, and easier management from a "single pane of glass."
- **Share the resources** — Segment the network into simple, logical, and scalable partitions for your various applications and services with privacy, flexibility, high performance, and quality of service (QoS) as primary goals. This sharing enables agility for multiple users, applications, and services.
- **Secure the data flows** — Integrated and dynamic security services resident in the network can provide benefits to users and applications sharing the infrastructure. Comprehensive protection secures data flows between external, internal, and inter-data center endpoints. Implement centralized orchestration and enforcement of dynamic, application- and identity-aware policies.



SIMPLIFY

The network design that used to work for the business might not be capable of supporting new demands on IT infrastructure and, most importantly, new business requirements. Networks built on fragmented and oversubscribed tree structures have problems with scaling and consistent performance. Design and management complexity and costs increase exponentially as more devices are added.

3-2-1 Data Center Network Architecture

Juniper simplifies the data center network and eliminates layers of cost and complexity with a "3-2-1 Data Center Network Architecture." Using fabric technologies such as Virtual Chassis technology, Juniper helps flatten data center networks, reducing them from three layers to two or even one layer. In the future, Juniper's Project Stratus will manage a 10GbE network at scale, as a single logical switch.

In addition, to help further simplify operations, Juniper consolidates multiple services into single high-performance platforms such as Juniper Networks® SRX Series Services Gateways, and utilizes the Juniper Networks Junos® operating system as the single OS across routing, switching, and security platforms.

SHARE

The cloud-ready data center requires network resources to be elastic, so that they can be allocated on-demand and at scale. Juniper's uniquely architected platforms deliver the agility and scaling required by virtualizing network configurations, segmenting services into logical domains, and using industry-leading hardware designs to scale without complexity. With a large pool of resources to draw from, customers can efficiently partition those resources to meet service requirements, remain flexible, and ensure operational performance, security, and control.

Edge Service Consolidation and Management

Juniper accomplishes this by building an intelligent network where these high-level policies can be enforced at the port level, and even at the date center's edge where connections to other data centers and networks occur over the WAN, the Internet, or a partner's network—effectively creating an even larger pool of resources to share across the organization. The Juniper Networks M Series Multiservice Edge Routers and MX Series 3D Universal Edge Routers are powerful, reliable, and the industry's most scalable solutions for the intelligent edge and inter-data center mobility.

SECURE

Security administrators must protect client-to-server traffic as well as traffic between physical and virtual servers, applications, and systems in other data centers. Security solutions need to be flexible to adapt to the changes in traffic volumes and data flows that occur because of virtualization, Web 2.0 applications, and cloud services. The increasing user access and the rising sophistication of security threats in a cloud-ready data center require expanded protection. Appropriate policies affect availability of business critical applications and operations.

To address these challenges, security services must be consolidated and

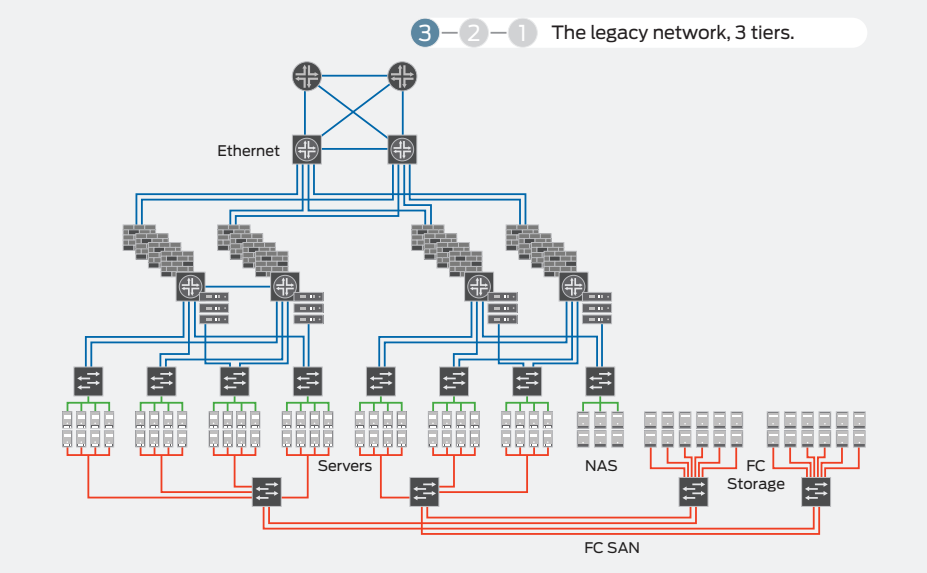


Figure 1: The legacy network

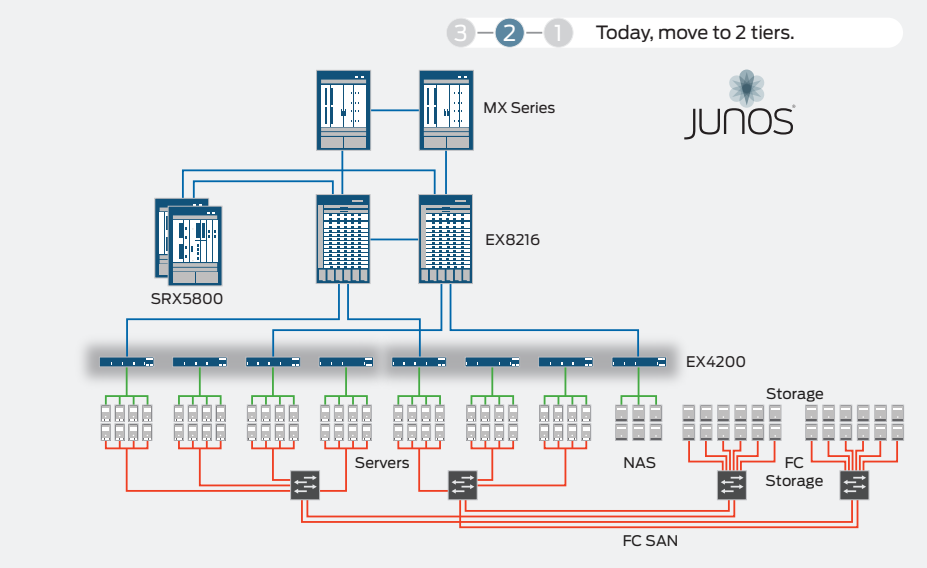


Figure 2: Juniper delivers a simplified two-tier network today with Virtual Chassis fabric technology.

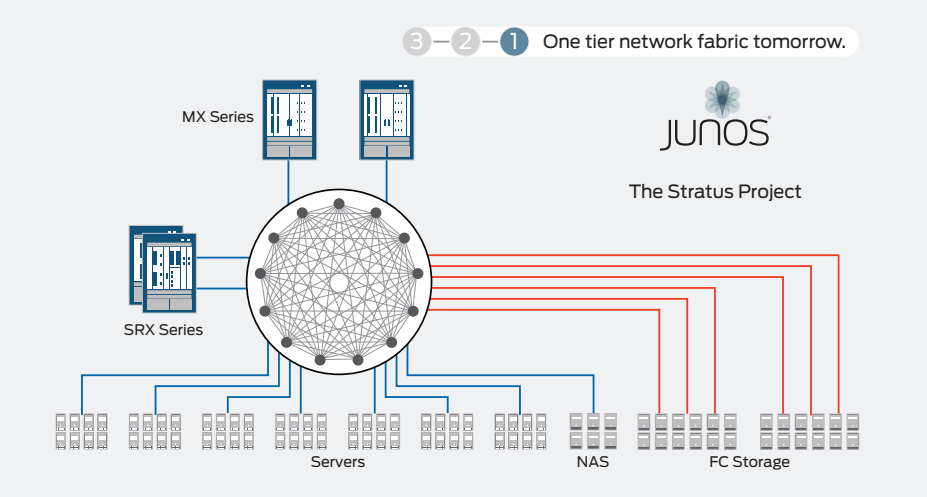


Figure 3: Juniper's vision for the ultimate simplification of the data center is Project Stratus, delivering a single fabric that unites Ethernet, Fibre Channel, and Infiniband networks.

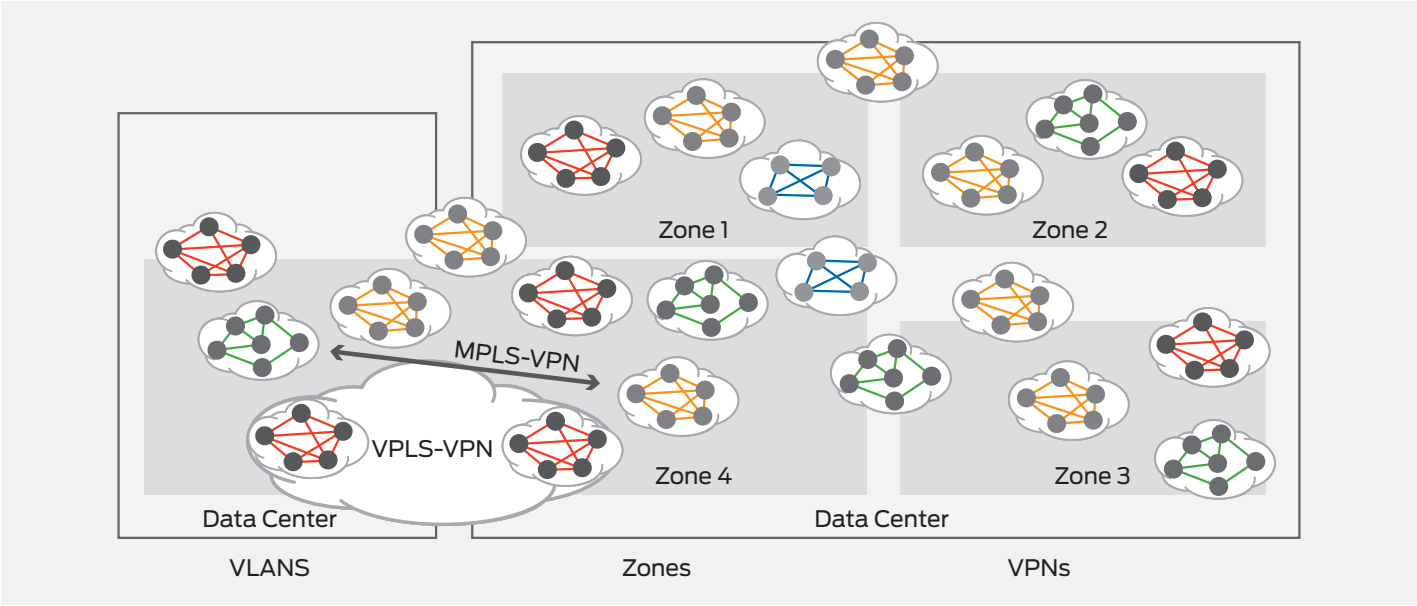


Figure 4: Scalable network virtualization technologies

pooled in a coordinated fashion to complement the simplification and sharing of the network. This approach enhances the flexibility and efficiency of the entire solution.

Juniper Networks has developed high-performance, cloud-enabled dynamic security services to meet today's security and performance requirements while accommodating future on-demand growth. Services such as application monitoring, stateful firewall, intrusion detection and prevention, and VPNs are consolidated on an expandable platform that flexibly and dynamically assigns resources as needed. Security services must be application- and identity-aware, while providing secure access for the mobile workforce to data center applications. Juniper provides best practices implementation guides to minimize risk and speed time to deployment when configuring security solutions for cloud-ready data centers.

AUTOMATE

Juniper's open, extensible network automation software makes it easier to manage and administer the data center by simplifying repetitive and complex tasks, defining and implementing policies within the network, and orchestrating implementation across multiple systems using network-based software. This greatly lowers operational expenses by reducing configuration errors, measurably improving reliability, and freeing up labor resources to innovate rather than administer.

The Juniper Networks Junos Space network application platform was designed to provide end-to-end visibility and control to enable network resources to be orchestrated in response to business needs. Operators can significantly simplify the network life cycle, including configuration, provisioning, and troubleshooting with an open automation platform.

Improve the Economics and Experience of Information Technology to Deliver Greater Business Value

Many organizations can benefit from cloud-ready data center networks, whether building a cloud-like infrastructure for internal purposes, connecting to public cloud services, or preparing to connect to public cloud services in the future. Juniper Networks, as a partner with wide-ranging experience, can help organizations reduce complexity and overall IT costs while accelerating delivery of IT services to users over a secure, simplified network.

For more information, please visit: **www.juniper.net/us/en/solutions/enterprise/data-center/**

Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, CA 94089 USA
Phone: 888.JUNIPER (888.586.4737)
or 408.745.2000
Fax: 408.745.2100
www.juniper.net

Oct 2010

Axxia™ Communication Processor Accelerates Cloud Networking

APPLICATION INTELLIGENCE COMPONENTS

Application Visibility

- Who is accessing what?
- Top N applications
- Bandwidth consumed per application

Application Profiling and Control

- Network readiness for applications
- Troubleshoot application performance
- Application access control and QoS

Application Acceleration

- Application caching
- Application proxies
- WAN acceleration

Axxia Communication Processor

Cloud computing is all the rage. By 2014, Gartner expects worldwide spending on cloud computing to reach almost $150 billion. The goal of cloud computing is to enable IT organizations to achieve an order of magnitude improvement in the cost effective, elastic provisioning of IT services, including Software as a Service (SaaS), Infrastructure as a Service (IaaS), and Platform as a Service (PaaS).

Significant networking needs challenge this goal: dynamic scalability, lower latency, real-time resource management, self-healing reliability, lock-tight security, and guaranteed application performance. Existing networking infrastructure is already stressed to the breaking point. Faster ports, greater bandwidth, and flatter networking topologies can mitigate some of the challenges, but ultimately we need solutions that will scale with unforeseen demands. We need new cloud networks.

At LSI, we believe application intelligence is the essential ingredient, the key, to new cloud networks. Application intelligence allows every application across the network to obtain its fair share of resources, bandwidth, quality of service (QoS), and Service Level Agreement (SLA) in the presence of all other applications. Key ingredients for delivering application intelligence include:

- Application visibility
- Application profiling and control
- Application acceleration

The next-generation data center demands a new processor. A communication processor with a highly optimized architecture that enables each task to be allocated to the right resources for the job.

Cloud Networking: Lofty challenges, down to earth solution.

Application Visibility

Traditionally, applications were classified for QoS on network port plus IP address source+destination pair. That's no longer good enough. Cloud networks need to peer into data packets for fine-grained application visibility. In addition, applications such as unified communications, IP video, and telepresence require reliable real-time performance. The LSI™ Axxia Communication Processor features hardware-based deep packet inspection (DPI) for fine-grained application visibility with reduced packet latency and increased per-flow performance versus common approaches. DPI also allows the analysis of application signatures to eliminate common security threats like viruses, worms, and denial of service (DOS) attacks.

Application Profiling & Control

The ability to view and gain insight into how applications behave while flowing through network infrastructure can lead to improved design, better user experience and improved business innovation. The LSI Axxia Communication Processor incorporates a high-performance stateful flow processing architecture that targets the right on-chip resource for the job, from classification, to data and control plane processing, to traffic management, all necessary for profiling & control. True scalability, low latency, and deterministic performance result from this unique architecture.

Application Acceleration

Application visibility, profiling, and control enable real application acceleration and WAN optimization. Dramatic improvement in response times can be achieved with compression, application caching, content proxies, and virtualized application hosting.

Axxia has impressive CPU processing power and optimized application–specific resources to allow OEMs to deliver on the promise of cloud networking. In addition to application acceleration with Axxia, LSI offers media acceleration and storage acceleration solutions targeted at cloud networking.

Axxia Communication Processor

Powered by Virtual Pipeline™ Technology

The Axxia Communication Processor (ACP) is designed to meet the increased performance and lower power demands of next-generation communication networks. Using an innovative asymmetric multicore architecture, the ACP delivers fully deterministic performance with up to 20 Gb/s of data throughput, regardless of packet size, system loading, or protocol.

At the heart of each ACP is a high-performance multicore PowerPC® processor made by IBM® capable of reaching 2GHz operating frequency. Function-specific acceleration engines deliver fast path processing without unnecessarily taxing the multicore complex. These acceleration engines are derived from silicon-proven, cores used extensively on the broad product portfolio from LSI, including deep packet inspection, security, packet processing, and traffic management abilities. The ACP architecture uses Virtual Pipeline, a patented message-passing technique, for intra-processor communication between the acceleration engines, multicore complex and system on chip (SOC) subsystem components.

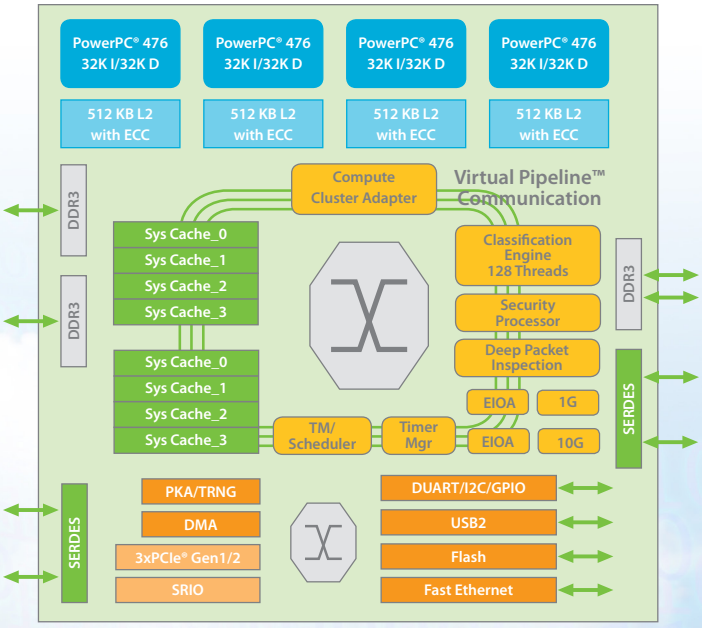
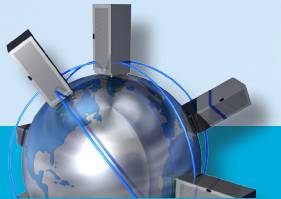


Figure 1 - ACP Block Diagram



Axxia Intelligent Network Interface Card

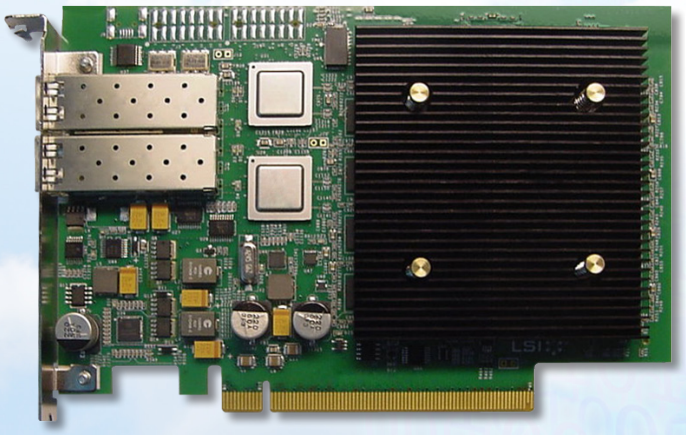
This PCI Express® (PCIe®) NIC delivers an integrated cloud networking solution in a small footprint. Based on the Axxia Communication Processor, this turn-key solution provides application intelligence in cloud servers for security and monitoring applications, as well as server offload capabilities.

Axxia Intelligent NIC Hardware Features

- Based on Axxia Communication Processor
  - Built-in hardware accelerator engines including classification, deep packet inspection, packet integrity check, timer, packet assembly, programmable scheduler/ buffer manager, stream editor engines, and quad-core PowerPC® processor
- Dual 10GbE small form-factor pluggable (SFP) Network Interfaces
- PCIe Gen2 x4
- On board flash for board boot-up
- 10/100 Fast Ethernet port for initial development and debugging
- Optional serial port for management and debugging
- Supports fiber loopback
- Small foot print – PCIe half length card, full height
- On board system and configuration DDR3 SDRAM memory

Axxia Intelligent NIC Software Features

- Throughput up to 20 Gb/s (cut-through mode)
- Pattern recognition and replacement based on powerful classification and DPI engines
- Application recognition with ACP classification and DPI engines, and quad-core PowerPC® processor
- IEEE® 1588 support: flow classification and time-stamping, message type mapping, PTP egress processing
- IPsec: various encryption/ integrity/ authentication algorithms
- TCP proxy server offload: with ACP packet assembly and classification engine
- Packet delivery to host x86 CPU
  - Mechanism for transferring data block over PCIe
  - Large data block transfer; either per flow basis for each transfer or a mix of flows in same transfer



Axxia Intelligent NIC Software Architecture