WHITE PAPER

OPPORTUNITIES AND CHALLENGES WITH THE CONVERGENCE OF DATA CENTER NETWORKS

Juniper Networks is Well Positioned to Deliver a Superior Converged Network Fabric for Next Generation Data Centers

Table of Contents

List of Figures

Executive Summary

Legacy architectures constrain today's data centers as they attempt to cope with the exponential increase in applications, servers, storage, and network traffic. By rethinking the network from the ground up, Juniper Networks provides customers with a long-term, future proof strategy to develop a single, converged data center fabric with the flexibility and performance to scale to super data centers, while continuing to drive down the cost and complexity of managing the data center information infrastructure. Juniper's vision for cloud computing paves the way for super data centers that harness the scale to dynamically allocate any resource—including routing, switching, security services, storage systems, appliances, and servers—without compromising performance.

With standardization efforts well under way, the industry wide adoption of a fully converged network fabric for enterprise data centers will likely gain traction by 2011. Juniper remains actively involved in this standards effort and today is committed to delivering a converged data center fabric that is upwardly compatible with current standards proposals. Consequently, Juniper is well positioned to deliver the next-generation data center fabric that enables converged storage area network (SAN), local area network (LAN), and high-performance computing (HPC) clusters to be built and managed as a single logical infrastructure.

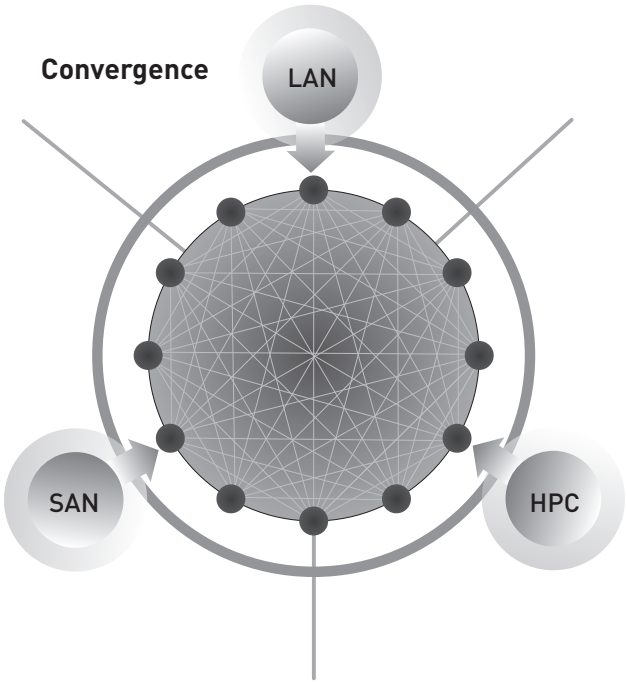

Figure 1: The next-generation data center network fabric enables converged SAN, LAN, and HPC infrastructures to be built and managed as one logical entity.

Introduction

The rising cost and complexity of building and operating modern data centers have led businesses to continually seek new ways to make their data center infrastructures simpler and more efficient. While the cost of data center networking gear is relatively small in relation to the cost of server hardware and software applications, the underlying network fabric is the linchpin that connects all mission critical resources. Therefore, a simpler, more streamlined data center fabric can result in greater efficiency and productivity, as well as lower operating costs.

The legacy approach of deploying multiple networks, switching domains, and tiers to handle a dizzying array of technologies has increased complexity, compromised performance, and prohibited scalability in data center networks. Disparate network infrastructures currently exist for HPC clusters, LANs, and SANs. Multiple oversubscribed switching layers are now in place to meet density and performance requirements as well as to contain costs. Adding to this complexity, each layer often uses a different operating system.

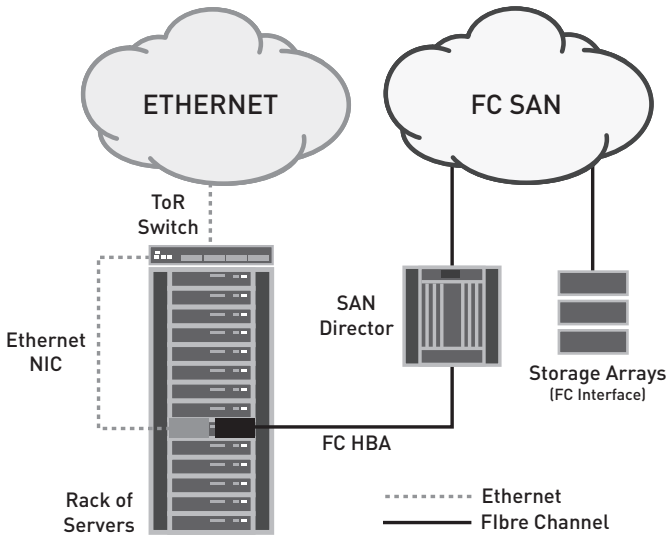

Figure 2: The legacy approach of deploying and managing different network infrastructures has created complexity, compromised performance, and prohibited scalability in the data center.

The deployment of dissimilar hardware and versions of software introduces unpredictable operating behavior and dramatically increases the burden of administration and management. Businesses have tried to curb this complexity by deploying applications in dedicated infrastructure silos, but the results have produced poor overall resource utilization and higher operating costs.

To achieve the lowest total cost of ownership (TCO) for the data center infrastructure, businesses must simplify the network, reduce the number of devices, collapse multiple switching tiers, use fewer operating systems across all network devices, and unfurl one unified, converged network fabric across all data center resources.

Requirements for Converged Data Center Networks

In order to protect the investment of a network upgrade, IT organizations should consider three key factors before embarking upon any effort to converge data center networks into a single, cohesive fabric.

1. Deploy an industry standard convergence solution that creates a single lossless, low latency network fabric.

Multiple networking technologies currently used in data centers, such as Ethernet and Fiber Channel (FC), evolved over time to address a unique set of requirements. The convergence of these data center networks requires identifying the single most promising standards-based infrastructure technology—one that can replace all of the others in a graceful and cost-effective manner. To ensure interoperability between multivendor products, the best data center fabric to support converged environments must support a standards-based approach that guarantees packet delivery and eliminates complex tiers that contribute to latency.

2. Solve the lack of scalability inherent in existing network infrastructures.

A single converged network must connect and unite all virtualized end nodes throughout the data center infrastructure. This increases the total number of devices as well as the diversity of devices that are interconnected by a single data center fabric. Unfortunately, today's data center networks don't scale well due to complex subdivisions of purpose-built domains and tiers. Large networks contain scores of subnetworks and silos that have been deployed to mitigate a variety of legacy Ethernet characteristics, including:

- The unstable and error prone nature of Spanning Tree Protocol (STP)
- The poor utilization of network links when using STP
- The unpredictable onset of congestion and problems associated with eliminating that congestion
- Unpredictable throughput and latency between servers and storage end nodes
- Broadcast storms related to Ethernet's flooding behavior, which impact performance and cause packet loss
- Insufficient security isolation between services that share the same infrastructure without introducing excessive complexity
- The inability to decouple a server's address from its location in the data center

As a result of these factors, a converged data center network with its underlying fabric must have the ability to scale without a significant increase in complexity, capital expenditures, and operating costs.

3. Demand a clear return on investment.

In times of economic downturn and lean budgets, few businesses are willing to invest capital and disrupt data center operations unless there are compelling financial benefits attached. Clear, quantifiable economic metrics must be in place before data center networks can converge into a single, cohesive fabric. The technologies, products, and operational practices that are employed to create this converged infrastructure must yield an extraordinary and unprecedented return on investment (ROI).

Now let us examine each of these factors in greater detail.

Industry Standards for a Lossless, Low Latency Infrastructure

The three primary networking technologies currently in use in today's data centers—Ethernet, Fiber Channel, and Infiniband—were created to address a distinct set of problems. But by the time these technologies were widely adopted as mainstream technologies, they had grown well beyond their original roles, and the fullness of their unique feature diversity and sophistication were no longer relevant to modern data centers.

Additionally, as these technologies continued to evolve independently, they essentially provided multiple ways of performing the same networking tasks such as learning, addressing, forwarding, security, congestion avoidance, quality of service (QoS), and overall management.

Of these technologies, *Ethernet* is the most prominent in global deployment. Initially designed to interconnect a small number of computers over coaxial cable, it has evolved to address the most diverse networking requirements in local, metro, and wide area networks. Ethernet is deployed in data centers in much the same way as it has been deployed in campus networks for years. Today, Ethernet's primary challenge in data centers is to provide greater scale, improved performance, and lossless transport.

Fiber Channel was initially designed as a transport mechanism for the Small Computer System Interface (SCSI) protocol to connect and share a set of remote disks with a group of servers. The technology was progressively enhanced over the years, and today it includes sophisticated services and management tools that make it the preferred technology for moving mission critical storage traffic within data centers. Scale, performance, and cost pose the greatest challenges to the future viability of this technology in the modern data center.

Originally designed to replace the Peripheral Component Interconnect (PCI) that links I/O devices to servers, *Infiniband* has emerged as the de facto standard for low latency server-to-server interconnections in HPC grids and other cluster topologies. Scale and security issues continue to plague Infiniband in data centers today, and advances in Ethernet technology are quickly closing the gap in terms of throughput, latency, and cost.

Selecting One Standards-Based Converged Network Technology

As a technology, Ethernet has several advantages over the alternatives. First, it is the dominant networking technology in deployments for network areas beyond data centers, and most customers and vendors are already familiar and comfortable with it—certainly more so than any other networking technology.

Ethernet also has the richest ecosystem in terms of a technology vendor community and the most credible roadmap in terms of high-speed, standards-based transmission. Repeatedly delivering on its promise as the lowest cost solution, Ethernet has prevailed over other technologies like Token Ring, Fiber Distributed Data Interface (FDDI), and ATM, which have faded into obscurity.

Based on its rich history of successfully building high-performance networking solutions with a singular focus on IP and Ethernet, Juniper Networks advocates Ethernet as the optimal converged networking technology for data center infrastructures. With the increased affordability and rapid adoption of 10-Gigabit Ethernet (10GbE) in the data center, Ethernet is poised to take on the connectivity tasks formerly relegated to Infiniband and Fiber Channel to become the dominant infrastructure technology for data centers.

While standards-based solutions for iSCSI and network-attached storage (NAS) are readily available, well supported, and widely deployed across an extensive range of Ethernet networks today, Fiber Channel remains the dominant protocol for mission critical storage in data centers.

The challenge of transporting Fiber Channel-based SAN traffic over Ethernet is currently being addressed by standards developments in Fiber Channel over Ethernet (FCoE) and Converged Enhanced Ethernet (CEE), described later in this paper.

For a considerable time, Infiniband had clear advantages over Gigabit Ethernet with respect to latency and performance. However, those advantages have largely eroded with the arrival of 10GbE and the availability of affordable, special-purpose 10GbE network interface cards (NICs).

Enhancements to Ethernet for Converged Data Center Networks: FCoE and CEE

Ethernet was originally developed to handle traffic using a best-effort approach. Conversely, storage networks were designed to guarantee delivery. Because Ethernet is the ideal infrastructure to accommodate converged data and storage networks, the next logical step is to enhance the Ethernet standard so that it can ensure zero packet loss, support in-order packet delivery, and better manage congestion avoidance. Enhanced industry standards that support these key capabilities will propel Ethernet into the forefront as the preeminent infrastructure for LANs, SANs, and HPC clusters.

An enhanced Ethernet transport will benefit existing storage networking solutions like NAS and iSCSI while enabling native Fiber Channel frames to be encapsulated within Ethernet packets, thereby providing a cost-effective migration path for existing Fiber Channel SAN investments.

These enhancements are being addressed by the two standardization efforts mentioned earlier—the FCoE protocol for Fiber Channel packet transport over Ethernet, and CEE to ensure lossless and reliable transport over Ethernet.

FCoE: The Protocol Developed within T11. The proposed FCoE protocol is being developed by the T11 Technical Committee—a subgroup of the InterNational Committee for Information Technology Standards (INCITS)—as part of the Fiber Channel Backbone 5 (FC-BB-5) project.

While the development of FCoE as an industry standard will bring the deployment of unified data center infrastructures closer to reality, FCoE by itself is not enough to complete the necessary convergence. Many additional enhancements to Ethernet are required to make it a viable, useful, and pragmatic implementation. These additional enhancements have been proposed by the CEE committee, which involves a broad range of industry representatives.

CEE: A Set of Proposals to the IEEE. In order to guarantee lossless transport, Ethernet must undergo a multitude of enhancements. To drive that change, a consortium of network, storage, CPU, ASIC, server, and NIC vendors have joined forces to create the CEE proposal. The original version of this proposal, which was submitted to the IEEE in March 2008, includes recommendations on four vital components of the IEEE 802 standard that governs Ethernet's physical and logical properties:

1. Priority Flow Control (PFC) – IEEE 802.1Qbb
2. Enhanced Transmission Selection (ETS) – IEEE 802.1Qaz
3. Ethernet Congestion Management (ECM) – IEEE 802.1Qau
4. Data Center Bridging Exchange Protocol (DCBX) – IEEE 802.1 DCBX

There is some skepticism in the technical community as to whether the currently proposed ECM mechanisms will work well in an end-to-end fashion across a data center. While the current CEE proposal could achieve limited success, an even better approach would be to avoid congestion altogether, rather than just reacting to it after it has already set in and spread across the network. Reducing the number of network hops eliminates the cause of the problem more directly and offers a cleaner approach to managing the complexity of congestion across multiple hops.

According to Pat Thaler, chair of the IEEE DCB Task Group, any standard resulting from the initial CEE proposal will not be ratified as a full-fledged standard until March 2010 at the earliest [1].

Meeting the Scalability Requirements

The term "a single converged network fabric" involves connecting the full diversity of virtualized end nodes to one unified data center network. Unfortunately, today's data center networks are incapable of accommodating this model by design. Existing equipment and practices continue to divide large network domains into smaller subnetworks in order to deal with the limitations of legacy Ethernet architectures.

This approach runs counter to the notion of a single large virtualized pool of resources—long considered the *Holy Grail* of data center productivity and efficiency. As a result, it is imperative for next-generation data center networking technologies and architectures to support unprecedented scalability.

Next-generation data centers must be able to accommodate large numbers of servers without having to divide the network into many smaller subnetworks. The use of Spanning Tree Protocol should also be minimized whenever possible. In addition, next-generation data center architectures must decouple server addresses from their locations in data centers. And there should be no compromise between the secure isolation of services and scalable performance.

[1]DCB standards will be final in March 2010, four months later than initially planned due to outstanding, but not insurmountable issues, according to Pat Thaler, chair of the DCB Task Group in the IEEE. **www.networkworld.com/news/2008/102008-cee.html?page=2**

Modern applications and programming models pose additional challenges to data center network scale. Distributed applications, such as map-reduce and memory-cache, change the profile of communication patterns within data centers. They significantly increase server-to-server communication, require no assumptions of network locality to keep software development simple and efficient, and thrive on predictable and uniformly low latencies.

Another demanding application is server virtualization. Server virtualization software uses IP subnets to define the network range for virtual machine (VM) live migration. Today, limitations in legacy Ethernet's architecture restrict IP subnets to just a few hundred servers. So there is a conflict between creating large, scalable networks for VM migration and managing large, scalable networks hampered by legacy Ethernet architectures. This conflict must be resolved in order to reap the full benefits of a large scale virtualized server environment that runs on a converged Ethernet fabric, thus making any-server/any-service computing a reality[2].

A Clear Return on Investment

Even when converged network technologies become viable and deployable, data centers will not likely undergo rapid and sweeping infrastructure changes. As servers are phased out, or as new racks are deployed in new or existing facilities, data center architects will take the opportunity to thoroughly evaluate their options and carefully consider the economic benefits of upgrading to a converged network fabric.

Naturally, there must be compelling proof of significant ROI before any new data center infrastructure technology is adopted. In addition to being less expensive to purchase and operate, the accrued savings from new technology must be substantial enough to offset the risk of further change. Keeping mindful of the potential rewards and risks, customers who migrate to a converged network fabric in the data center will likely do so in two phases.

Phase 1: I/O Consolidation within a Rack

With pre-standard versions available in 2009, FCoE will consolidate Ethernet and Fiber Channel interfaces on servers using converged network adapters (CNAs) within a rack. A top-of-rack Ethernet switch would then separate the server traffic into distinct LAN and SAN clouds.

Businesses will have two deployment options when it comes to Phase 1, which involves I/O consolidation within a rack. The first option, which is also the simplest and most cost effective, places the Fiber Channel-to-FCoE gateway function in the SAN director so that Fiber Channel traffic is received in Ethernet frames by the top-of-rack switch (see Figure 3). A more complex and costly second option places the Fiber Channel-to-FCoE gateway function in the top-of-rack Ethernet switch with converged network adapter (see Figure 4).

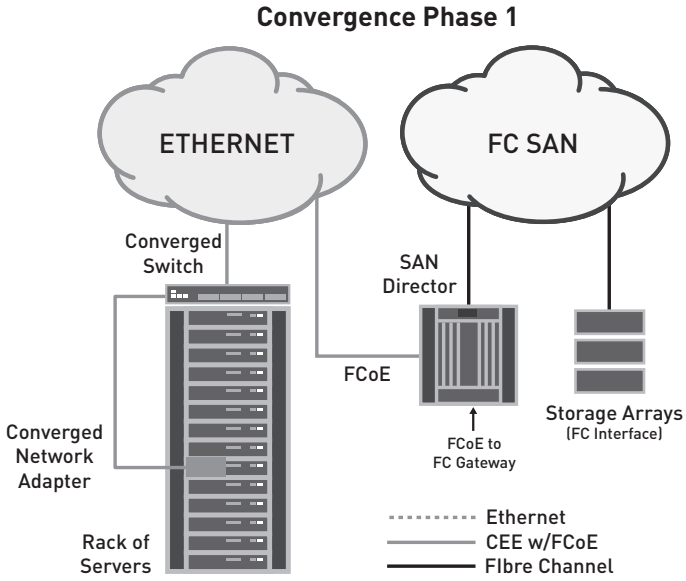

Figure 3: Today, the simplest and most cost-effective path to I/O consolidation within a rack involves placing the Fiber Channel-to-FCoE gateway function in the SAN director so that Fiber Channel traffic is received in Ethernet frames by the top-of-rack switch.

[2]The Cost of a Cloud: Research Problems in Data Center Networks by Albert Greenberg, James Hamilton, and David A. Maltz. **http://ccr.sigcomm.org/online/files/p68-v39n1o-greenberg.pdf** See section 3.2

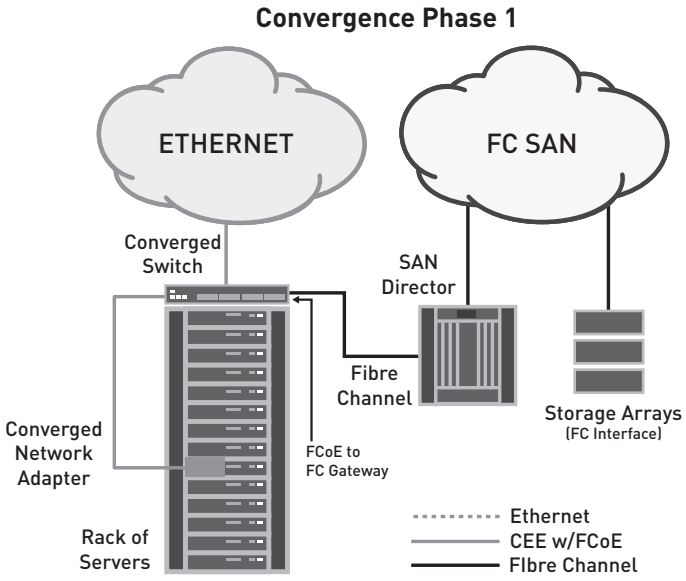

Figure 4: Alternatively, I/O consolidation within a rack may be performed by employing the Fiber Channel-to-FCoE gateway function in the top-of-rack Ethernet switch using a pre-standard converged network adapter (CNA).

If reducing TCO is an overriding requirement for a converged data center infrastructure, several issues might hinder a successful FCoE deployment today:

- **Cost.** The cost of CNAs today is significantly more than the cost of the host bus adapters (HBAs) and NICs they are trying to replace. At the time of this writing, here is a sampling of the current average street price for NICs and HBAs: CNA ($1,900), 4 Gbps FC HBA ($500), 10 GbE NIC ($750), GbE NIC ($90). Including also the cost of the top-of-rack switch ports and the cabling, we find that in all cases (over a range of data bandwidth from 2 to 14 Gbps, with Fiber Channel traffic assumed to be half of the data traffic), it is always cheaper today to use NICs and HBAs instead of CNAs and their respective switch types. This analysis includes all infrastructure including access to the LAN and SAN fabric, and is not burdened by back end topology that is common across both options.
- **Feature Completeness.** First-generation CNAs implement the PFC feature only. ETS, an essential part of the CEE proposal that is not included on current CNAs, will enable intelligent provisioning of priority and bandwidth guarantees across multiple hops of a network. The absence of ETS on first-generation CNAs means users must rely on randomly available bandwidth to resolve different requirements of LAN and SAN traffic. Also, given that first-generation, hardware-based FCoE CNAs can only go up to 4 Gbps for SAN traffic due to ASIC limitations, the price paid for a 10GbE interface is wasted in the absence of a good bandwidth partitioning scheme.
- **Incomplete Standards.** Because most CEE features are baked into a CNA's ASICs, changes to the current proposal will most certainly require a firmware upgrade, if not a complete hardware replacement. As noted earlier, the last CEE feature regarding congestion notification is not slated for standardization until March 2010.
- **Power.** Several CNA vendors have announced plans to consolidate multiple fixed-function ASICs into fewer multifunction ASICs. This will allow a universal CNA to be used for regular Ethernet, CEE, or FCoE ports. However, until such CNAs are available, the multiple ASICs deployed today will consume all power available for these adapters. This compares unfavorably with the much lower power consumption of separate Fiber Channel HBAs and Ethernet NICs.
- **Interoperability.** There are no multivendor interoperability certifications available today for new CEE and FCoE-enabled top-of-rack switches and the dominant director class SAN switches. This lack of interoperability, especially given the pre-standard nature of current data center convergence equipment, locks customers into a single vendor's products.

Phase 2: Fabric Convergence Throughout the Entire Data Center

The goals of productivity and efficiency can only be achieved when network convergence is applied throughout the entire data center, not just within the narrow confines of a single rack. This is the second and final phase in the adoption of a converged data center network fabric.

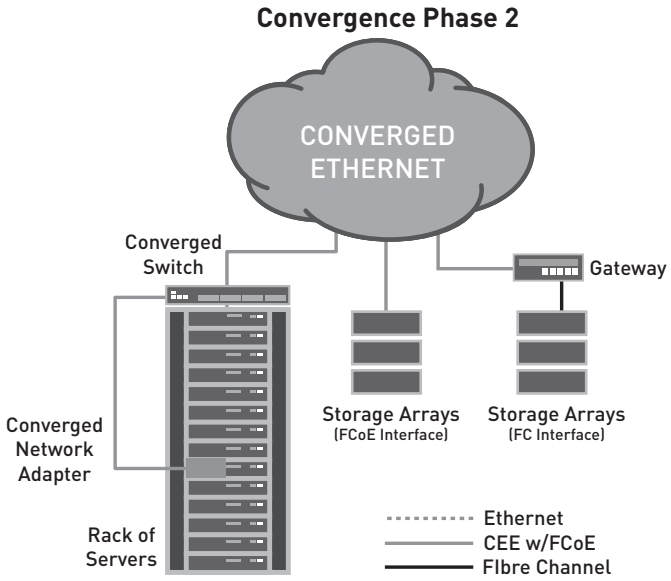

Figure 5: The convergence of LAN, SAN, and HPC clusters represents the second and final phase in the adoption of a unifying next-generation network fabric in the data center.

The CEE proposal will need to be standardized and vetted in a lab environment to prove that it is fully operational above and beyond what was required in Phase 1. This kind of vetting is required because the congestion management mechanisms of the CEE proposal kick in after the network is already congested, instead of the preferred approach of avoiding congestion from setting in.

Additionally, specifications in the CEE proposal must apply to the full breadth of LAN and SAN aggregation switching equipment to deliver on the end vision of a truly seamless, converged data center fabric. Unfortunately, no hardware that would enable this is available today, and this will delay the deployment of a fully converged data center network fabric until about 2011.

Conclusion

The Stratus Project[3], announced by Juniper Networks in early 2009, enables the creation of a single data center fabric that delivers a quantum jump in scale, performance, security, and simplicity, with the flexibility to support fully converged and virtualized data center environments.

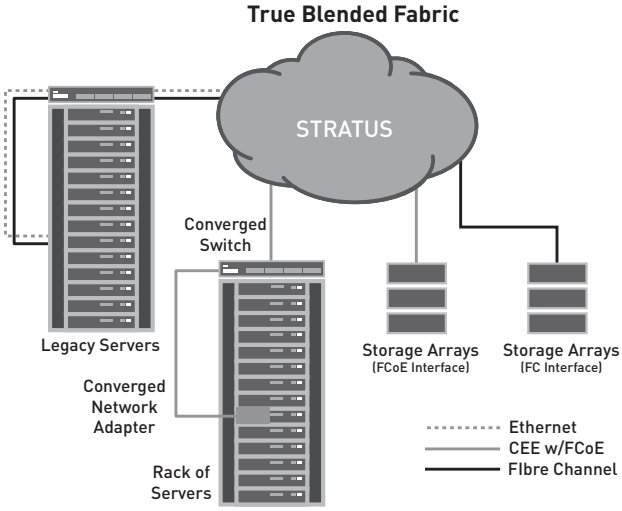

Figure 6: Juniper Networks is poised to deliver a true blended network fabric in the data center that delivers a quantum jump in scale, performance, security, and simplicity, with the flexibility to support fully converged and virtualized environments.

The underlying data center fabric defined by the Stratus Project is easy to manage, flat, lossless and non-blocking—core ingredients for any simple, converged networking solution with a low TCO. In the future, several of Juniper's currently shipping Ethernet switching products will evolve to support the proposed CEE standard as it matures and becomes stable enough to implement.

Given the current state of the CEE standard proposal and the not-so-compelling economics of first-generation CNAs, Juniper does not expect immediate, widespread deployment of I/O consolidation using FCoE. A few early adopters might move forward and pilot some type of FCoE implementation in an effort to become familiar with the technology. However, most businesses will likely wait for the CEE standard to stabilize until the economic benefits of deployment are clear and compelling.

In the wake of today's tough economic climate and tight IT budgets, Juniper recommends that businesses begin to evaluate lower risk options that will enable data centers to migrate gracefully and cost effectively toward a converged networking fabric that offers a low TCO.

Many businesses have already experienced the advantages of consolidating and virtualizing server resources in the data center, but few have yet to consider the benefits of employing similar solutions in the networking and security infrastructure. Juniper advocates the following:

- Collapse multiple layers of legacy switch infrastructures by combining industry-leading, 10 GbE performance Juniper Networks® EX8200 line of Ethernet switches with Juniper Networks EX4200 Ethernet Switches to ensure pay-as-you-grow scalability, simplified network design, and carrier class resiliency.
- Consolidate and virtualize the myriad of security appliances in the data center using Juniper Networks SRX Series Services Gateways.
- Optimize core routing in the data center and across multiple data centers using Juniper Networks M Series Multiservice Edge Routers and Juniper Networks MX Series Ethernet Services Routers.
- For unprecedented simplicity, deploy a single network operating system across all data center switching, routing, and security resources. Juniper Networks JUNOS® Software delivers this using a consistent implementation of all control plane features across the entire infrastructure.
- Adopt a common management platform for data center switching, routing, and security using Juniper Networks Network and Security Manager. NSM manages the life cycle of data center equipment and includes reporting and diagnostic tools for data center-wide visibility.

In combination, Juniper's simplified data center solutions can lead to a demonstrably lower TCO, saving businesses up to 52 percent in capital expenditures alone compared to legacy equipment and architectures.

[3]The Stratus Project at Juniper has been underway for more than one year. The name was coined after the meteorological term "stratus," which is defined as a flat, single-layer cloud. The project currently spans multiple business groups at Juniper and has resulted in more than 30 patent applications filed to date.

About Juniper Networks

Juniper Networks, Inc. is the leader in high-performance networking. Juniper offers a high-performance network infrastructure that creates a responsive and trusted environment for accelerating the deployment of services and applications over a single network. This fuels high-performance businesses. Additional information can be found at www.juniper.net.

Corporate and Sales Headquarters
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

APAC Headquarters
Juniper Networks (Hong Kong)
26/F, Cityplaza One
1111 King's Road
Taikoo Shing, Hong Kong
Phone: 852.2332.3636
Fax: 852.2574.7803

EMEA Headquarters
Juniper Networks Ireland
Airside Business Park
Swords, County Dublin,
Ireland
Phone: 35.31.8903.600
Fax: 35.31.8903.601

To purchase Juniper Networks solutions, please contact your Juniper Networks representative at 1-866-298-6428 or authorized reseller.

2000315-001-EN July 2009


Printed on recycled paper.