

THE JUNIPER NETWORKS QFABRIC ARCHITECTURE: A REVOLUTION IN DATA CENTER NETWORK DESIGN

Flattening the Data Center Architecture



Table of Contents

Executive Summary	3
Introduction	3
Data Center Network Challenges	3
The Need for an Exponential Data Center	4
Drawbacks of the Legacy Data Center Network	4
Too Complex	4
Too Location Dependent	5
Not All Data Centers are Created Equal	6
The Ideal Data Center Network	6
Simplicity	6
High Performance	6
The Juniper Solution—Introducing QFabric	7
The QFabric Architecture—Delivering Operational Efficiency	7
How QFabric Benefits the Enterprise	8
Conclusion—Juniper Delivers the Next-Generation QFabric Data Center Network	9
About Juniper Networks	10

Executive Summary

In today's data center, the network is the bottleneck preventing businesses from realizing the full benefits of their information infrastructure. The traditional tree structure that has long dominated network design is outdated, and it introduces a host of problems that impede application performance, add operational overhead, and negatively impact the user experience. A change is needed to ensure that businesses can take full advantage of their investments in virtualization, service-oriented architecture (SOA)-based applications, and compute resources. Juniper Networks® revolutionary Quantum Fabric (QFabric) architecture delivers that revolutionary data center design by creating a single tier network that operates like a single Ethernet switch. The QFabric architecture provides a quantum leap in scale, performance, and simplicity, while delivering the ability to support the fully converged and virtualized environments so crucial for today's data centers.

Introduction

Data Center Network Challenges

Today's data center is the foundation of information-driven businesses, delivering the application services and access to data and compute resources required to meet organizational obligations. Over the past decade, businesses have invested heavily in—and benefited greatly from—advances in data center technology such as SOA, server virtualization, and cloud computing. These have provided the operational efficiency, business agility, and network resiliency enabled by a more elastic infrastructure.

What has not evolved, however, is the data center network itself. The network plays a vital role, interconnecting applications and logically extending servers within the data center to form shared pools of resources. Yet the traditional multitiered “tree” architecture that still dominates data center network design imposes a series of performance penalties that hinder enterprises from realizing the full potential of their investments in virtualization and other technologies.

This is because the legacy tree architecture is simply too complex to manage and too topologically diverse to optimally leverage these innovations. In other words, the data center network has too many devices deployed in a very stratified, hierarchical manner—an architecture that prevents businesses from taking advantage of the operational efficiencies and savings afforded by virtualization.

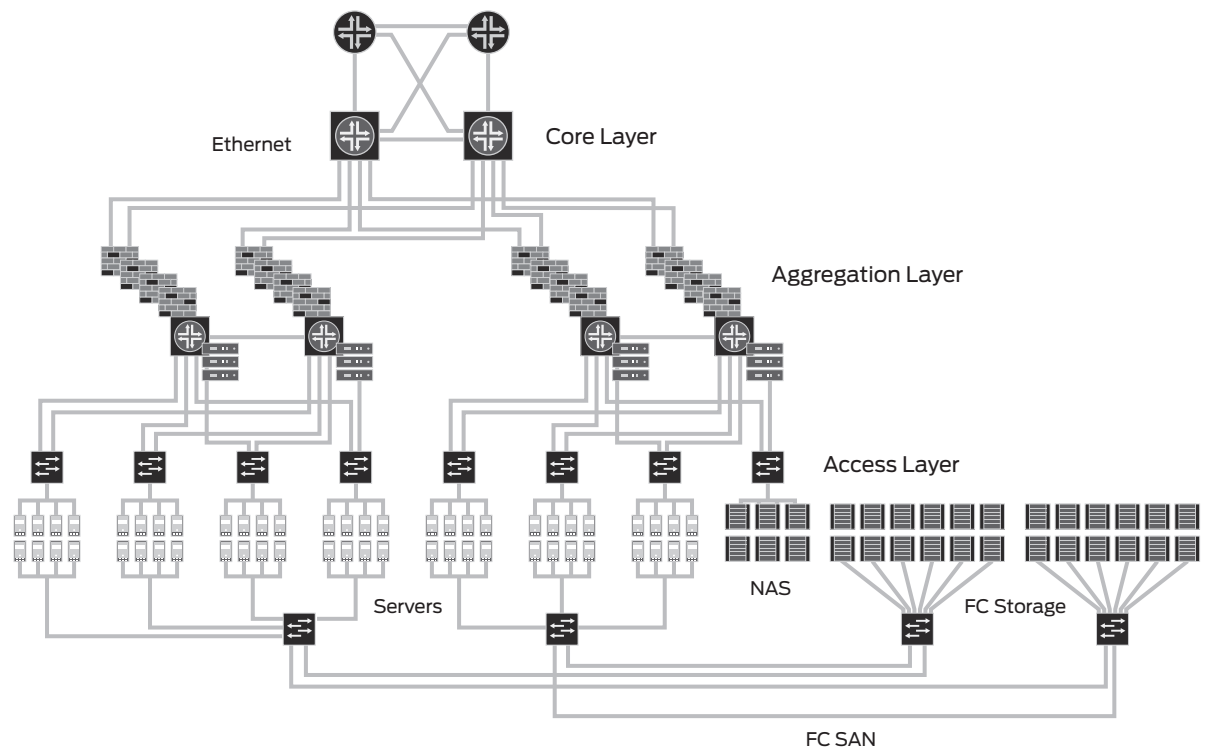


Figure 1: Today's data centers feature a complex, three-tiered tree architecture that adds latency and impedes application performance.

Here's why. With the adoption of SOA-based applications and technologies such as server virtualization, most of today's data center traffic is server-to-server or "east-to-west"—i.e., among peer devices at the same level within the hierarchy. The legacy tiered architecture, however, forces every packet passing from one server to another to first transit multiple switches up and down the tree in a "north-south" flow. This imposes significant latency on the transaction, which severely impacts application performance, virtualization features such as workload mobility, and real-time cloud computing interactions.

The current data center architecture is also costly from an operational standpoint. As VMs migrate between virtual servers, the network paths between them change, adding tremendous management overhead as IT struggles to keep track of and maintain connectivity between VMs through the tree structure in a highly dynamic environment. Ultimately, managers are forced to make trade-offs between network performance, scalability, and management complexity—deciding between user experience and economics to meet their particular data center network requirements.

The Need for an Exponential Data Center

A change in data center network design is needed to ensure that organizations can take full advantage of their investments in SOA-based applications, virtualization, and storage and compute resources. Since the most efficient way for resources to interact is for them to be no more than a single hop away from each other, the ideal next-generation network architecture would directly connect all processing and storage elements in a flat, any-to-any connectivity-based network fabric. Optimized for performance and simplicity, this next-generation architecture would address the latency requirements of today's applications, eliminate the complexity of legacy hierarchical architectures, scale elegantly, and support virtualization, cloud computing, convergence, and other requirements for the next-generation data center.

Juniper Networks QFabric data center architecture delivers this revolutionary design by creating a single tier network that operates—and is managed—like a single, logical, distributed switch. Consisting of edge, interconnect, and control devices all running the Juniper Networks Junos® operating system, QFabric technology combines the low latency, flat architecture, and operational simplicity of a single switch with the scalability and resiliency of a fabric. The result is a revolutionary approach that unleashes the exponential power of the data center, delivering any-to-any connectivity that reduces complexity and lowers capital, management, and operational expenses.

Drawbacks of the Legacy Data Center Network

Legacy data center architectures have two fundamental problems: They are too complex and too location dependent, which impacts performance and the user experience.

Too Complex

As a data center grows and the number of physical servers increases, IT must add more switches to interconnect them.

There are two predominant points of management in the network—the switches and the interactions between the switches. Since each switch is an autonomous device, IT must manage each one individually, which creates significant overhead in large data centers. These switches also communicate with each other by using shared network protocols such as routing, link aggregation groups (LAGs), quality of service (QoS), and security—all examples of protocols that must be applied across the network and configured on every device to support the overall design. This causes the number of interactions between switches (often referred to as "control traffic") to grow geometrically, even if they are not directly connected. This explosion of interactions follows a simple formula of $n(n-1)/2$ where "n" represents the number of switches. Based on this formula, 10 switches can generate 45 such interactions, 100 switches can generate nearly 5,000 potential interactions, and 1,000 switches can create 5 million interactions (see Figure 1).

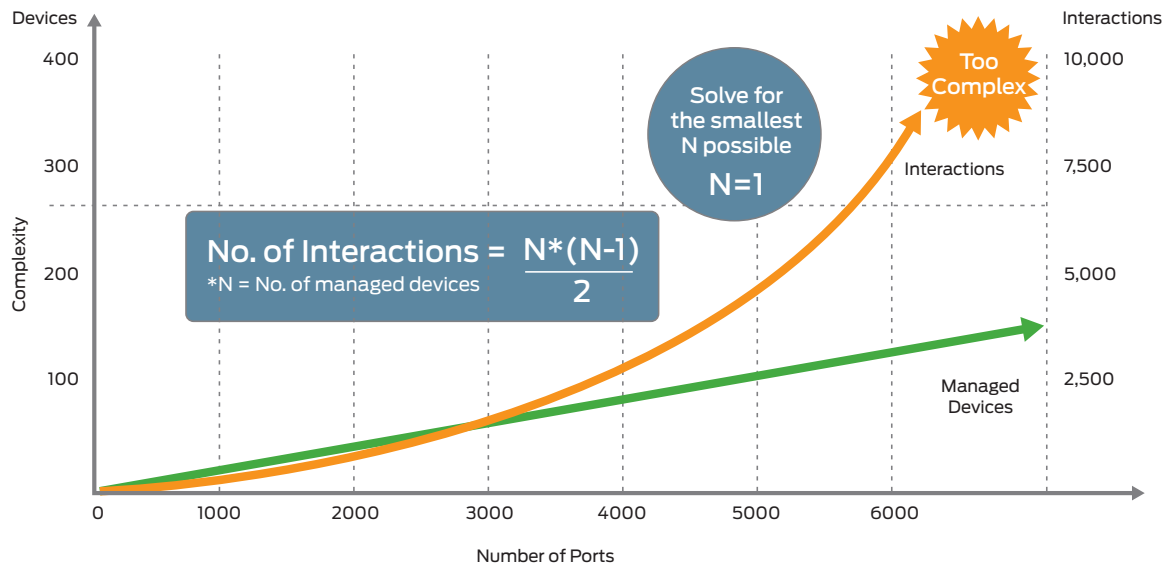


Figure 2: The number of interactions between switches grows geometrically as new devices are added to the network.

Since it is nearly impossible to deal with this level of increasing complexity, network operations will typically segment the network into smaller subnetworks. Unfortunately, this runs counter to the desire to consolidate all servers and storage in the data center into a single large, more efficient, and more elastic pool of resources—the vision behind cloud computing.

The tree-based network imposes a huge economic penalty as well. Up to 50 percent of a network's ports—the most expensive ports—are used to interconnect switches, not to link servers, storage, or other devices. At the same time, as much as 50 percent of a network's bandwidth is unavailable if Spanning Tree Protocol (STP) is being run, because STP disables up to half the available links to prevent the creation of loops in a tree-based network. As a result, organizations must purchase, deploy, and manage a higher number of switches relative to the number of servers, storage, and other data center devices that need to be connected.

Too Location Dependent

In today's data center network, application performance is often dependent on the physical location of servers that are trying to communicate—that is, where each of them resides in the tree hierarchy relative to one another. The more hops required to complete a transaction, the more latency that transaction incurs, and this contributes to unpredictable application performance.

As mentioned earlier, IT is forced to expand data center networks by interconnecting multiple layers of switches in complex hierarchical tree topologies—an architecture that forces each packet to transit multiple switches up and down the tree. This forces packets to take divergent, poorly defined paths through the network, depending on the source of the request and the destination. The risk of performance degradation multiplies with each switch transit and every layer of the switch hierarchy. In such an environment, it becomes very difficult to ensure consistently low application latency, which in turn makes it virtually impossible to guarantee service levels or to meet service level agreements (SLAs).

It also becomes increasingly difficult to engineer the network so that it performs consistently under varying loads. Given that traffic in the data center is predominantly server-to-server (east-west), today's hierarchical network architecture imposes a significant performance penalty. This is due to the fact that all traffic must first travel north and south before it can go east and west, creating bandwidth congestion and contention on those vertical north-south paths. Therefore, it follows that traffic flows that are forced to make fewer hops—that is, those sourced from and destined to servers that are relatively close to one another—have a performance advantage over flows that are sourced by and destined to opposite ends of the data center (the extremities of the tree). Clearly, application performance and eventually the end user experience will be greatly influenced by the location of servers and the applications that run on those servers.

Not All Data Centers are Created Equal

Unfortunately, businesses often have to make trade-offs between performance, scalability, management complexity, and cost to meet their particular data center network needs. Some organizations such as stock exchanges and investment houses need very high-performance data center networks; if they can't complete a trade within a few microseconds, they can lose out to competitors. To ensure high performance, they trade cost for speed. For other companies, such as cloud providers, scalability is critical to their business success so they may have to trade complexity and cost for scalability. For still other businesses, the data center network is an expense, not a revenue source, so they must optimize for economics and often have to sacrifice speed, scalability, and other performance needs.

Recognizing that different types of organizations have different data center requirements and business goals, networking vendors have designed their products to address specific performance, scalability, or cost requirements. Rather than having different products from different vendors deliver one or two of these capabilities, enterprises should be able to deploy a next-generation data center network that can address all of their needs, without any trade-offs or sacrifices.

The Ideal Data Center Network

Given the problems with legacy data center network designs, a new architectural approach is required, one that is transformational and innovative, not incremental. As Yankee Group researchers put it, "The data center network is now the backplane of the virtual data center."¹

The ideal next-generation network architecture for modern data centers would directly connect all processing and storage elements in a flat, any-to-any network fabric. Optimized for performance and simplicity, this next-generation architecture would address the latency requirements of today's applications. It would support virtualization, cloud computing, convergence, and other data center trends. And, it would scale elegantly; and eliminate much of the operational expense and complexity of today's hierarchical architecture.

Juniper's flat, fabric-based architecture offers several distinct advantages.

Simplicity

To create a simple data center network, Juniper drew inspiration from switch design, determining that the best way to address data center complexity was to make the network itself behave like a single switch. Inside every switch is a fabric—a mesh that is completely flat (not tiered like a tree) and provides any-to-any connectivity between ports. A next-generation data center fabric architecture extends the simplicity of this switch fabric concept and scales it beyond a single device to the entire data center network. When this fabric design is applied, it retains a level of simplicity by allowing multiple physical switches to behave as, and be managed like, a single, logical switch. Management and operational expenses are an order of magnitude lower than with traditional network architectures.

A network fabric design enables any-to-any connectivity, where every device is a single hop away from any other device. A network fabric also greatly simplifies the network architecture by eliminating unnecessary switch layers and the corresponding switch-to-switch interactions. In addition, because the network fabric functions as a single device, there is no need for shared network protocols like STP, as the configuration and management of the network fabric does not need to be exposed to the customer. Consequently, network management and operations are greatly simplified, and performance improves even if the network scales from tens of ports to tens of thousands of ports. Servers, storage, and other data center devices can be connected to the fabric seamlessly because of the fabric design's high degree of scalability.

High Performance

With a flat, fabric-based network architecture, the data path is greatly simplified and extremely fast. Since there is no tree structure, there is no need for multiple hops between sending and receiving devices. Location is no longer an issue—by providing direct any-to-any connectivity, a network fabric minimizes latency by ensuring that all members of the fabric are just one hop away from each other. In addition, each connection benefits from dedicated bandwidth, thus avoiding congestion and ensuring optimal network performance for applications. North-south traffic flows are eliminated, while server-to-server traffic is carried over multiple east-west paths. This results in performance improvements that accelerate business critical, productivity enhancing applications that have a positive impact on the user experience.

¹Source: "Data Center Evolution Is Dependent on a Network Fabric," copyright 2010—sponsored by Juniper Networks

Eliminating tiers of switches and enabling all ports has ancillary benefits as well. This fabric-based data center network requires fewer network devices and fewer device interconnections, which reduces the capital outlay required to build or refresh the data center network as well as cutting power, space, and cooling costs.

A fabric-based architecture can address a broad array of data center requirements. By giving organizations high performance, simplicity, and scalability while lowering capital and operational expenses, a fabric-based data center architecture eliminates the need for IT to trade off one capability for another. Businesses benefit by being able to cost-effectively refresh or build the data center network they need to be competitive and to take advantage of emerging business opportunities. Whether a company is looking to re-architect its data center to sell cloud services, streamline service delivery, or bring new products to market more quickly, a fabric-based network architecture can help them get the most from their investments in virtualization, SOA, and other data center technologies, and be well positioned for future growth.

The Juniper Solution—Introducing QFabric

Juniper Networks has developed a comprehensive strategy designed to make this new fabric-based data center network design a reality. At the heart of this strategy is Juniper's 3-2-1 Data Center Network Architecture, which eliminates layers of switching to flatten today's three-tier tree network structure to two layers and, with the Quantum Fabric architecture, to just one. QFabric is the "1" in 3-2-1.

The QFabric architecture complements Juniper's existing data center product portfolio. Juniper Networks EX Series Ethernet Switches with Virtual Chassis technology have been shipping for the past three years. Virtual Chassis technology enables enterprises to build two-tier data center networks and benefit from the high performance, low cost, and operational simplicity of a flatter network. QFabric expands on these benefits by flattening the data center network to a single tier.

The Juniper Networks QFX Series of products consists of three building blocks—the QF/Node, the QF/Interconnect, and the QF/Director devices. These building blocks work in concert to create a data center network that behaves with the simplicity and efficiency of a single switch, with all devices running the Junos operating system. In developing QFabric—the culmination of more than 1 million engineering man-hours and \$100 million in R&D expenses—Juniper is leveraging its expertise in switching, routing, and security, and it is drawing on its innovative engineering and technological advantages in software, hardware, silicon, and network architecture.

The QFabric architecture, which features more than 130 filed patents, includes a distributed data plane that directly connects all ports to one another, and an SOA-based control plane that provides tremendous scalability, resiliency, and a single point of management. New compute nodes can be added in plug-and-play fashion, allowing the fabric to grow seamlessly with minimal administrative overhead.

The QFabric Architecture—Delivering Operational Efficiency

In designing QFabric, Juniper has essentially taken the three basic components of a self-contained switch fabric—line cards, backplane, and Routing Engines—and broken them out into independent, standalone devices—the QF/Node, the QF/Interconnect, and the QF/Director respectively.

- **QF/Node:** In the QFabric architecture, the line cards that typically reside within a chassis switch become a high density, fixed configuration 1 RU edge device called the QF/Node, which provides access into and out of the fabric. The first member of the QFX Series product family, the QFX3500, not only acts as a QF/Node edge device in a QFabric architecture, it can also serve as a high-performance standalone switch in highly demanding data center environments.
- **QF/Interconnect:** Similarly, the backplane of a single switch becomes the QFabric/Interconnect device, which connects all QF/Node edge devices in a full mesh topology.
- **QF/Director:** The Routing Engines embedded within a switch are externalized in the QFabric architecture and called the QF/Director, which provides the control and management services for the fabric.

Each QF/Node edge device supports standards-based L2, L3, and I/O convergence functionality, with throughput of 1.28 Terabits per second and ultra low latency (<5 microseconds through the QF/Node and QF/Interconnect from any-to-any server). In addition to server facing ports, each QF/Node edge device includes four very high-speed interfaces for linking to up to four QF/Interconnect devices.

In its initial release, the QFabric architecture will support the interconnection of up to 128 QF/Node edge devices to create a single fabric capable of supporting 6,000 10GbE ports, with the ability to grow to an even larger fabric over time. The QF/Node and QF/Interconnect devices together create the distributed data plane for the QFabric architecture over which all of the data traffic to and from servers and storage is carried.

The control plane services are provided by the QF/Director, QF/Node, and QF/Interconnect devices. The QF/Director has a completely separate out-of-band control plane network that connects the QF/Node, QF/Interconnect and QF/Director devices and is used exclusively for carrying control traffic. Separating the data and control plane has always been a key design principle for Juniper because it ensures very high reliability. This principle is supported in the QFabric architecture wherein data and control traffic is carried over two separate networks. The distributed nature of the control plane is the key to the QFabric architecture's scalability and simplicity. At the same time, it also provides tremendous reliability since it eliminates any single point of failure in the system.

The QF/Director also provides all management services for the QFabric. Since it directly communicates with all QF/Node and QF/Interconnect devices, it builds a global view of the entire QFabric, which could consist of thousands of server facing ports. As a result, network managers get a single point of visibility, control, and management for the entire data center fabric. Rather than having to manage each switch individually, the QF/Director allows the data center network to appear as a single switch, providing a singularity of behavior and management to network operators. This dramatically reduces the operational costs typically associated with managing a data center network.

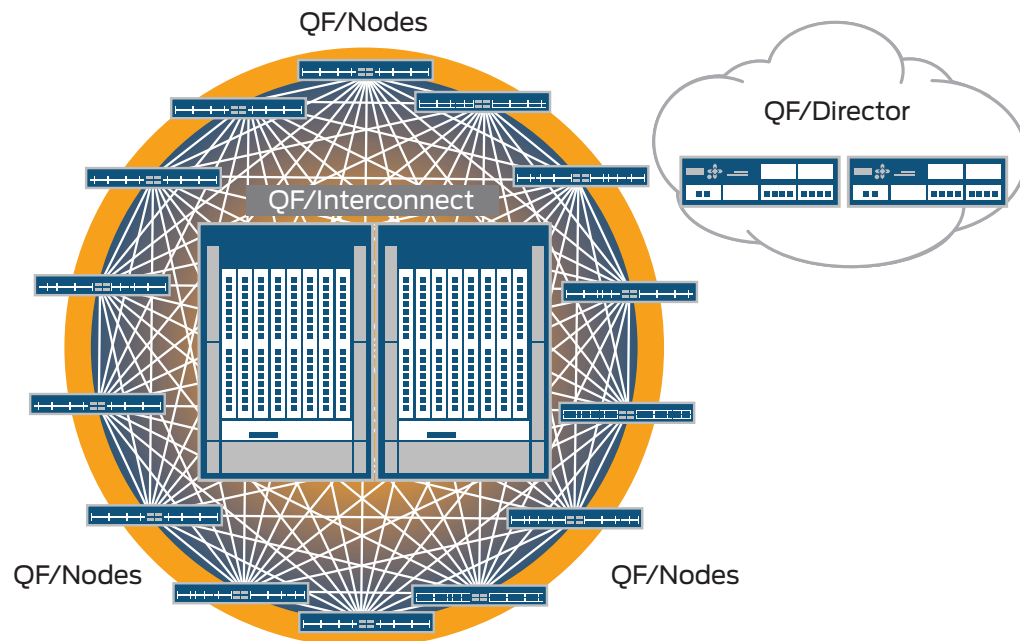


Figure 3: The QFabric architecture dramatically simplified the data center network, enabling any-to-any connectivity.

How QFabric Benefits the Enterprise

Designed from the ground up to meet the needs of today's data center, the QFabric architecture provides a flat, fabric-based network architecture that delivers the following benefits:

- **Simplicity.** By presenting itself as a single switch running Junos OS, the QFabric architecture greatly simplifies data center management. By reducing the number of switches required in the data center network, the QFabric architecture significantly lowers complexity and operational expenses, as well as power, space, and cooling costs. In addition, to maximize network uptime, Juniper has designed the QFabric architecture as a reliable carrier-class infrastructure with no single point of failure or downtime for reconfiguration and maintenance.
- **Performance.** The QFabric architecture's any-to-any connectivity supports the high-speed server-to-server communication so crucial for modern data center networks. Ultra low deterministic latency provides an order of magnitude improvement in performance over traditional network architectures, making it ideal for supporting latency-sensitive applications, east-west traffic flows, virtualization, cloud computing, and other high-performance data center initiatives.

- **Scalability.** The QFabric architecture offers unprecedented scalability, from tens of ports to tens of thousands of 10GbE ports, with the flexibility and performance to scale to super data centers. Because QFabric operates as a single logical device, scaling the network is as simple as adding ports to an existing switch. Likewise, new services, software, and features can be seamlessly integrated into the fabric. With the revolutionary QFabric control plane architecture, IT can easily scale the data center network with minimal additional management and operations overhead.
- **Support for virtualization and convergence.** The QFabric technology's flat, non-blocking, and lossless architecture is optimized for server virtualization, making it easier for enterprises to eliminate "stranded capital" that results from the underutilization of existing server and storage assets. QFabric also supports converged traffic at 10GbE access port speeds. With the QFabric technology, enterprises can easily run Fibre Channel, L2, L3, and Fibre Channel over Ethernet (FCoE) on a single data center network. Juniper's QFabric technology eliminates the trade-offs that three-tiered tree networks must make with respect to simplicity, performance, scalability, and cost while delivering the same scale, performance, and enhanced features as traditional networks.
- **Green.** The QFabric architecture is environmentally conscious, allowing enterprises to optimize every facet of the data center network while consuming less power, requiring less cooling, and producing a fraction of the carbon footprint of legacy multitiered data center networks.
- **Investment Protection.** Significantly, the QFabric architecture is a scalable solution that protects existing investments in the network infrastructure. QFabric components such as the Juniper Networks QFX3500 Switch can be introduced to data centers incrementally, interoperating with the current switch infrastructure while contributing its unique high-performance top-of-rack capabilities. Over time, the rest of the QFabric architecture can be introduced and the QFX3500s can be easily converted to a QF/Node edge device through a simple software configuration, allowing customers to transition to the QFabric solution on their own timetable—without requiring a wholesale replacement of the existing network.

Conclusion—Juniper Delivers the Next-Generation QFabric Data Center Network

Over the past few years, SOA, virtualization, cloud computing, and other technologies have revolutionized the data center. With the QFabric architecture, Juniper Networks is revolutionizing the network with a flat, any-to-any architecture that is the ideal topology for today's data centers. The fabric-based design dramatically reduces the cost and complexity associated with data center networks while improving performance and scalability.

QFabric technology fundamentally simplifies the data center network by reducing the number of network devices and connections and centralizing management. In addition, by optimizing application performance and supporting virtualization, convergence, cloud computing, and other key technologies, QFabric technology makes the entire data center more efficient.

As the leader in high-performance networking, Juniper Networks has the expertise to deliver the next-generation data center fabric. With QFabric technology, Juniper is helping customers build a next-generation network from which they can deliver services at unmatched scale, while protecting their existing investment and minimizing their total cost of ownership.

About Juniper Networks

Juniper Networks is in the business of network innovation. From devices to data centers, from consumers to cloud providers, Juniper Networks delivers the software, silicon and systems that transform the experience and economics of networking. The company serves customers and partners worldwide. 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
www.juniper.net

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
EMEA Sales: 00800.4586.4737
Fax: 35.31.8903.601

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

Copyright 2011 Juniper Networks, Inc. All rights reserved. Juniper Networks, the Juniper Networks logo, Junos, NetScreen, and ScreenOS are registered trademarks of Juniper Networks, Inc. in the United States and other countries. All other trademarks, service marks, registered marks, or registered service marks are the property of their respective owners. Juniper Networks assumes no responsibility for any inaccuracies in this document. Juniper Networks reserves the right to change, modify, transfer, or otherwise revise this publication without notice.