VIRTUAL CHASSIS TECHNOLOGY ON EX8200 ETHERNET SWITCH MODULAR PLATFORMS

Table of Contents

Table of Figures

Executive Summary

This paper describes the benefits of deploying Virtual Chassis configurations based on the Juniper Networks® EX8200 line of Ethernet switches deployed in a variety of environments. The EX8200 Virtual Chassis provides the most resilient campus/data center core design with no single point of failure, high scalability, high 10 gigabit Ethernet (GbE) port densities, and network simplification by reducing the number of managed devices.

Introduction

EX8200 switches with Virtual Chassis technology address the fundamental requirements of a core or collapsed core/aggregation switch, delivering a solution for implementing a network fabric in campus and data center environments.

The EX8200 Virtual Chassis solution supports multipathing and eliminates the inefficiencies associated with Spanning Tree Protocol (STP), provides a highly resilient system, and simplifies management and control plane operations at scale. Virtual Chassis technology on the EX8200 modular platforms also reduces the bandwidth inefficiencies associated with STP, accelerating network convergence and simplifying the network architecture.

Juniper Networks Virtual Chassis Technology

Juniper's Virtual Chassis technology enables customers to interconnect multiple individual switches (physical chassis) to create a single logical switch from a management, control and data plane perspective.

True Layer 2 multipath technologies and Layer 2 networks depend on a loop-free network topology for their operation. While traditional Spanning Tree Protocol (STP) technologies allow redundant links to exist in a network by blocking all but one of the connections at any given time, these technologies also have several significant negative side effects. First, they force half of the network's available capacity to sit idle at any time, increasing the cost required to achieve a particular level of performance. Second, if STP is incorrectly configured for any reason, the unblocked redundant connections cause a Layer 2 loop, resulting in traffic storms that are not only extremely difficult to troubleshoot but also bring useful data transfers to a halt. Finally, running STP in a virtualized network with redundant switches requires compute-intensive protocols such as Virtual Router Redundancy Protocol (VRRP) on each switch, limiting the number of simultaneous logical connections that can be supported.

True Layer 2 multipath technologies enable customers to build L2 domains without having to rely on STP to eliminate loops or lose redundant connectivity and achieve full utilization of available link capacity (active/active load sharing of redundant links). Also, true Layer 2 multipath technologies should deliver the resilience of multichassis network designs without imposing the scaling limitations of protocols such as VRRP. In addition, existing network components such as servers and storage devices, other L2 and L3 switches, security appliances, and routers should be able to attach transparently to the L2 multipath-enabled device.

Juniper's Virtual Chassis technology meets all of these requirements. By allowing multiple physical switches to appear as a single "virtual" switch to other attached network devices, Virtual Chassis technology allows multiple, simultaneously active L2 connections to any network device using link aggregation rather than STP. Juniper has been shipping Virtual Chassis technology on the fixed-configuration EX4200 line of Ethernet Switches since 2008, and now the technology is also available on the EX8200 line of Ethernet switches.

Virtual Chassis Technology on EX8200 Line of Ethernet Switches

Virtual Chassis technology on the EX8200 switches provides a scaled solution for core and aggregation layers while eliminating any single point of failure. Virtual Chassis technology enables multiple EX8200 switches to appear as a single "virtual" core/aggregation switch. Currently, two EX8200 chassis can be interconnected to form a single Virtual Chassis configuration. The technology is expected to be extended to four EX8200 chassis in the near future. EX8200 switches with Virtual Chassis technology can be deployed in a collapsed aggregation or core layer configuration, creating a network fabric for interconnecting access switches, routers and service-layer devices such as firewalls and load balancers using standards-based Ethernet LAGs.

EX8200 Virtual Chassis configurations consist of two EX8200 chassis—Juniper Networks EX8208 Ethernet Switches, EX8216 Ethernet Switches, or one of each—and two Juniper Networks XRE200 External Routing Engine devices to provide an active/standby pair of Routing Engines for control plane and management plane redundancy. Each of the EX8200 chassis also includes an internal RE, which is dual-connected to the two XRE200 devices. The XRE200 devices can also be optionally connected to each other via GbE interfaces.

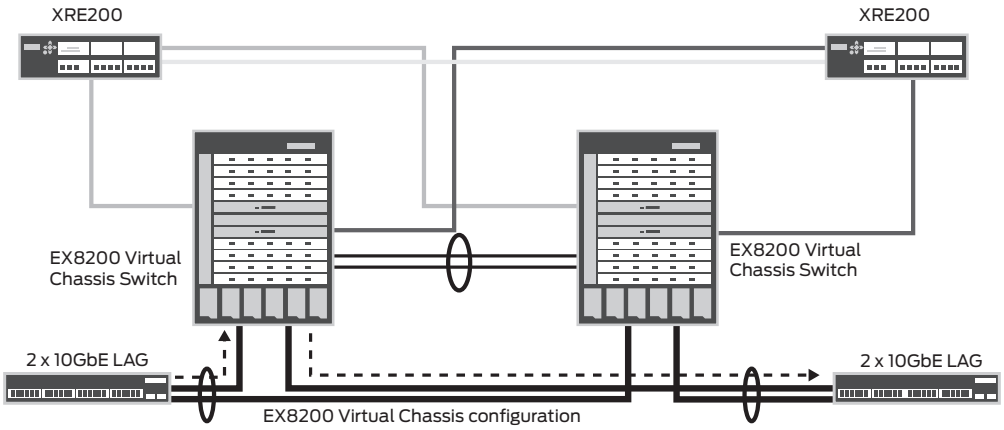

Figure 1: EX8200 Virtual Chassis configuration

An EX8200 Virtual Chassis configuration can include a mix of EX8208 (8-slot) and EX8216 (16-slot) switches, which can be interconnected using standard line-rate 10GbE interfaces as Virtual Chassis intra-connections. The connection between any two chassis in a Virtual Chassis configuration can either be a single line-rate 10GbE link or a LAG with up to 12 10GbE line-rate links. If the Virtual Chassis members are located in the same or adjacent racks, low-cost direct-attach cables (DACs) can be used as the interconnect mechanism. Member switches can be separated by up to 40 kilometers using small form-factor pluggable transceiver (SFP+) interfaces, allowing a single EX8200 Virtual Chassis configuration to span multiple floors or buildings.

Benefits of EX8200 Virtual Chassis Technology

- **Resiliency**—EX8200 Virtual Chassis configurations are highly resilient, with no single point of failure. This means that no single element—whether a chassis, a line card, a Routing Engine, or an interconnection—can render the entire fabric inoperable following a failure.
- **Support for Virtualization**—Virtual Chassis technology also provides server virtualization at scale. This is feasible due to the fabric's ability to provide simple L2 connectivity over a very large pool of compute resources located anywhere within a data center, whether those resources are across racks or across pods.
- **Scaled deployment**—EX8200 Virtual Chassis configurations provide high port densities and high performance by combining two switches (with support for four switches coming soon) as a single virtual device. An EX8200 Virtual Chassis configuration can be used to extend VLANs across rows of pods within and between data centers, interconnected via dark fiber. This is accomplished by placing an equal number of Virtual Chassis switch members in both data centers, or by interconnecting two separate Virtual Chassis configurations using a simple L2 trunk.
- **Remove STP inefficiencies**—The network fabric created by an EX8200 Virtual Chassis configuration does not present network loops, eliminating the need for protocols such as STP or for forklift upgrades to support unproven technologies such as TRILL.
- **Network simplification**—A Virtual Chassis fabric simplifies the network by reducing the number of managed devices in the core layer to one and eliminates the need to run protocols such as VRRP. In addition, the Virtual Chassis Control Protocol (VCCP) used to form the EX8200 Virtual Chassis configuration (as in the EX4200 line of switches) does not affect the function of the control plane. Control plane protocols such as 802.3ad, OSPF, Internet Group Management Protocol (IGMP), Physical Interface Module (PIM), and BGP do not require any modification. They function in exactly the same way as those running on a standalone chassis.
- **Investment Protection**—The EX8200 with Virtual Chassis technology is the only solution that supports all existing and future modules in the EX8200. In contrast, the incumbent solutions typically have an embedded implementation where only a small subset and mostly new cards can work in the corresponding virtual chassis, making it an expensive rip-and-replace solution. EX8200 Virtual Chassis provides investment protection for all the shipping and future modules of the EX8200.

EX8200 Line External Routing Engine (XRE200)

The switches employ redundant independent XRE200 external Routing Engines to externalize control plane functionality and separate it from the data plane, delivering the most resilient design with no single point of failure. All control protocols such as OSPF, IGMP, Link Aggregation Control Protocol (LACP), 802.3ah, VCCP, etc.—as well as all management plane functions—run or reside on the XRE200.

With its 2.1 GHz dual-core CPU, 4 GB DRAM, 160 GB RAID hard disk, and dual redundant power supplies, the XRE200 supports control plane processing requirements for large-scale systems.

Juniper Networks Junos® operating system high availability (HA) features such as GRES, NSR, and nonstop bridging (NSB) are enabled on the two XRE200 devices in an EX8200 Virtual Chassis configuration. In the event of an active XRE200 failure, the standby XRE200 takes over and Junos OS HA features ensure that the state of the Virtual Chassis, L2/L3 protocols and forwarding information are not lost.

Figure 2: XRE200 External Routing Engine

Two XRE200 devices are required in an EX8200 Virtual Chassis configuration to provide an active/standby pair for control plane and management plane redundancy.

Each XRE200 features:

- 2.1 GHz Intel® Core 2 Duo processor with 2 MB L2 cache
- 4 GB DRAM
- Internal 4 GB flash
- Redundant 160 GB hard disk drive
- LCD panel
- 10/100/100BASE-T RJ-45 port for out-of-band management
- Console port for out-of-band management
- USB drive for file storage
- One 10/100/100BASE-T RJ-45 for XRE200-to-XRE200 connection or XRE200-to-Virtual Chassis connection
- Two slots for 4x1GbE I/O cards, with customers able to choose between two I/O cards
  - 10/100/100BASE-T RJ-45 (one I/O module included with XRE200)
  - 1GbE SFP (future)
- Two redundant hot-swappable power supplies
- Two redundant hot-swappable fans

The XRE200 ships with the 4xGbE RJ-45 I/O card. Customers can choose to buy a second I/O card from the two available models. The fiber I/O card (future) can be used to connect the active and standby XRE200 devices when the distance between them exceeds the span of the Cat 5 or Cat 6 cable—for instance, when a Virtual Chassis configuration is spread between buildings on a campus. It is important to note that the active and standby XRE200 devices need not be directly connected, although there are advantages to connecting them directly as shown in the next section.

Advantages of Externalizing the Routing Engine

The "out-of–the-box" XRE200 provides resiliency and scale at many levels.

- **Link Resiliency**—First, if the intra-Virtual Chassis connection goes down, traffic flowing from an access switch to any other access switch or to any core/WAN router connected to the same EX8200 Virtual Chassis member switch is not affected, as shown in Figure 3. This is a significant improvement over alternate solutions, where the loss of the intra-system link leads to complete connectivity loss between any nodes (access switches or core routers) interconnected via the aggregation layer. Second, if the link from the XRE200 to the Virtual Chassis goes down, there are redundant links via the other XRE200. This advancement is made possible by externalizing Routing Engine functionality. Third, if one or more switches in the Virtual Chassis configuration lose their connections to adjacent chassis, access switches connected to the EX8200 Virtual Chassis configuration do not lose connectivity with the network. Figure 3 shows the traffic flow in the event of an intra-Virtual Chassis connection failure.

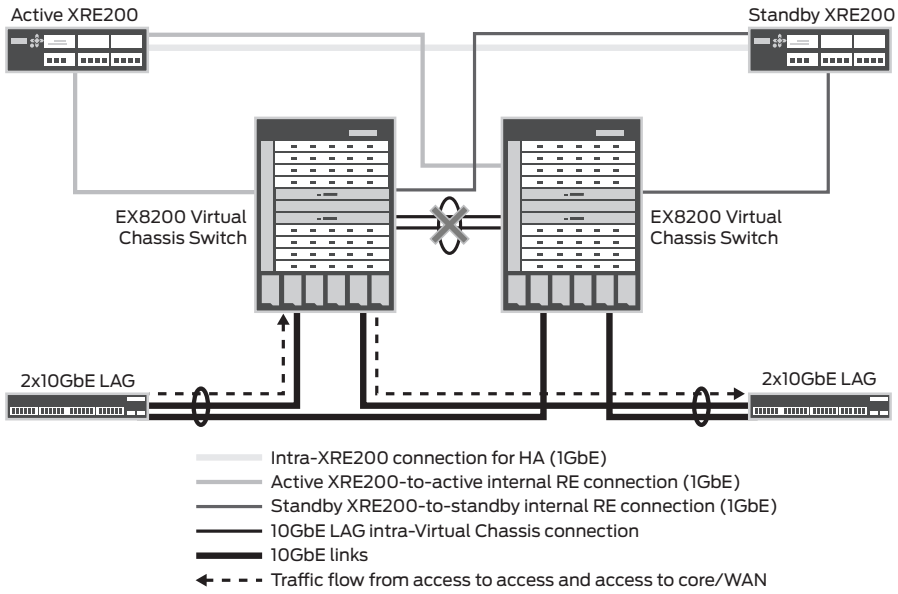

Figure 3: Intra-EX8200 Virtual Chassis configuration connection failure

- **Scale**—An EX8200 Virtual Chassis configuration can scale to include more than two member switches with the same two XRE200 devices, leading to a highly scaled solution. The "out-of–the-box" XRE200 provides significant control plane scalability advantages. The solution's scalability is captured in Table 1.

Table 1: Product Capability Comparison

SOLUTION	STANDALONE LINE RATE 10GBE PORTS	VIRTUAL CHASSIS PORT INTERCONNECT	VIRTUAL CHASSIS LINE RATE 10GBE PORTS	TYPICAL 10GbE SERVER COUNT*	TYPICAL 1GbE SERVER COUNT*
EX8208 Virtual Chassis	64	20-120 Gbps	104-124	600	6,000
EX8216 Virtual Chassis	128	20-120 Gbps	232-252	1,200	12,000

*There is a 5:1 oversubscription from server access to aggregation in the EX8200 line Virtual Chassis configuration.

Connecting an XRE200 to an EX8200 Line Virtual Chassis Configuration

The GbE interfaces on the active XRE200 (up to eight) can be used to connect to the active Routing Engines in each of the EX8200 chassis participating in the Virtual Chassis configuration. Similarly, the GbE interfaces on the standby Routing Engine (again, up to eight) can be used to connect to the standby Routing Engines in each of the EX8200 chassis in the Virtual Chassis configuration. The two XRE200 devices can also be connected to each other directly over any available GbE interface (Figure 4).

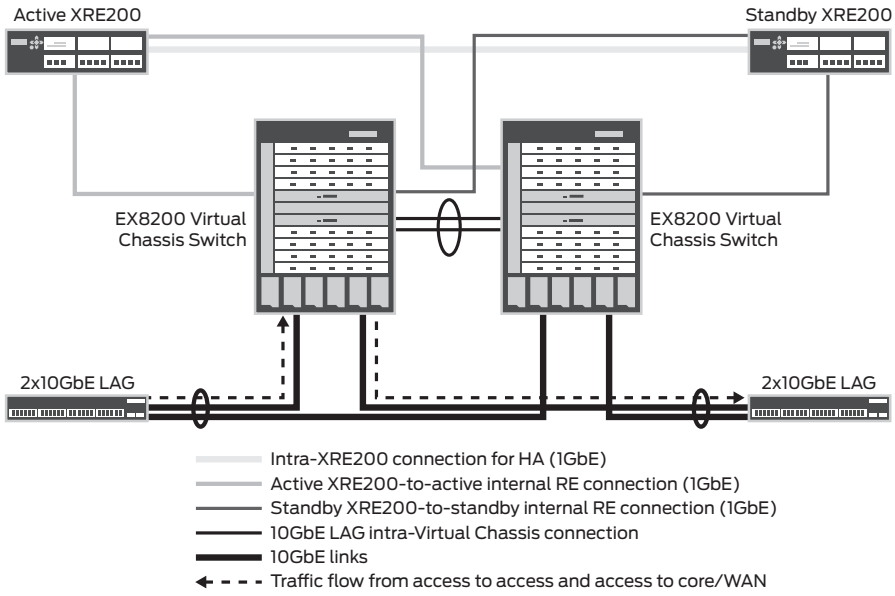

Figure 4: EX8200 Virtual Chassis configuration with XRE200 connecting to every member

Other methods of connecting the XRE200 to the EX8200 Virtual Chassis configuration are also available. In Figure 5, each of the two XRE200 devices is connected to only one internal Routing Engine of the EX8200 Virtual Chassis member switch. In this connection mode, control plane messages received by any Virtual Chassis member switch are relayed over the intra-Virtual Chassis connections to the member switch that is directly connected to the XRE200.

One of the benefits of this connection mode is that the active and standby XRE200 devices can be deployed in two physically different locations.

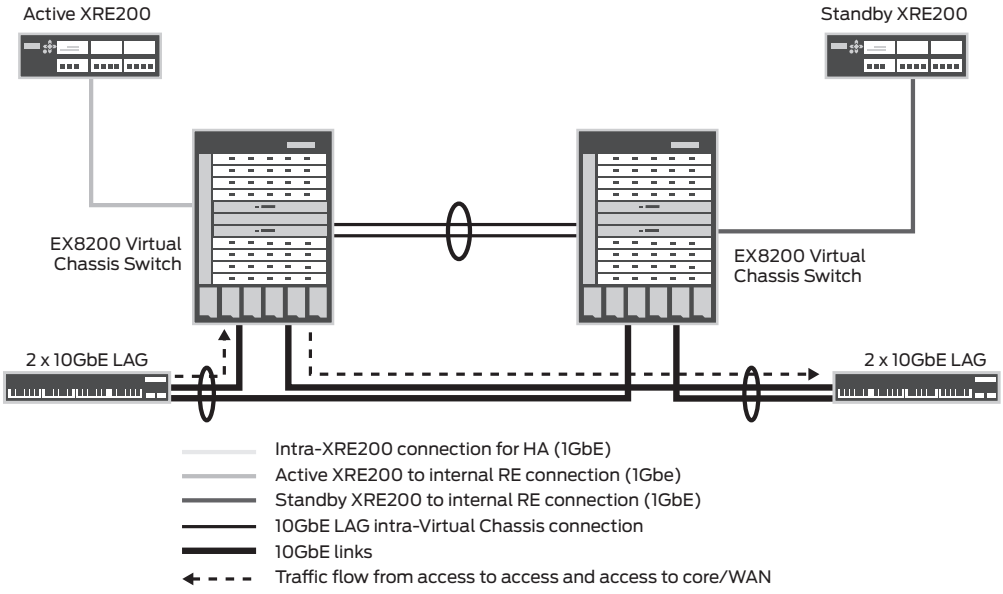

Figure 5: EX8200 Virtual Chassis configuration with XRE200 connecting to one member

Function of Internal Routing Engine in an EX8200 Virtual Chassis Configuration

In addition to providing direct connectivity between the EX8200 chassis and the XRE200, the internal Routing Engines in an EX8200 switches serve the purpose of controlling, monitoring and maintaining the chassis. Chassis-based functions like chassis and line card bring up, environmental monitoring, and power management are some of the typical functions performed by internal Routing Engines when the chassis is a member of a Virtual Chassis configuration. However, the internal Routing Engines do not process any control plane functions. Any L2/L3 control plane protocol packets received on an interface are sent to the XRE200 via the shortest path available. When the XRE200 devices are directly connected to the EX8200 chassis via the GbE port on an internal Routing Engine, all protocol data units (PDUs) are transmitted from the chassis to the XRE200 via the directly connected GbE link.

Deploying an EX8200 Line Virtual Chassis Configuration

Initially, an EX8200 Virtual Chassis configuration consists of a maximum of two member chassis. The EX8200 switches can be interconnected via a single 10GbE port or through a LAG consisting of multiple 10GbE links. In either case, the 10GbE interfaces used for Virtual Chassis connectivity must be on a line-rate 10GbE line card installed in the chassis.

Once the EX8200 switches are linked, the two XRE200 devices—one serving as the primary Routing Engine and one serving as a backup—must be connected. This is most easily accomplished by connecting the XRE200 devices to the out-of-band management ports on each switch's internal Routing Engine to provide the necessary redundancy and resiliency. For additional resiliency, and to reduce the number of hops between devices, a direct link between the two XRE200 devices can also be implemented (see Figure 4).

Virtual Chassis Configuration Options

When adjacent devices such as switches, routers or security devices must be connected to the EX8200 Virtual Chassis configuration, one of two options can be used—fabric mode and non-fabric mode.

a) **Fabric mode**—With the fabric mode option, network devices connect to the EX8200 Virtual Chassis configuration via a standards-based LAG, with members terminated on both switches participating in the Virtual Chassis. Network traffic entering either of the two EX8200 Virtual Chassis switch members has a direct connection to any other network device, and it uses that direct connection to exit the Virtual Chassis (Figure 4).

   For campus and data center environments, fabric mode is the recommended method because it provides the lowest latency for any network traffic and avoids use of the intrachassis VCP link, thereby limiting potential congestion.

   Fabric mode is also highly resilient. Should an interchassis link in the EX8200 Virtual Chassis configuration fail, the system continues to operate since the external control plane running on the two independent XRE200 devices ensures that there is a path between any two members of the Virtual Chassis, whether an EX8200 chassis or an XRE200. When such a failure occurs in alternative technologies that do not implement an external control plane, it typically leads to a "split-brain" scenario—causing a network outage. With the EX8200 Virtual Chassis solution, however, there is no single point of failure.

b) **Non-fabric mode**—In non-fabric mode (see Figure 6), network devices are not directly linked to all members of the Virtual Chassis configuration. In this configuration, traffic between any two network devices might have to traverse the interchassis Virtual Chassis link. While this is not the recommended method for connecting devices to the Virtual Chassis configuration, users might be forced to employ non-fabric mode if the two EX8200 switches are in separate locations and there is not enough fiber to connect every device to both chassis. Users can also use non-fabric mode when deploying an EX8200 Virtual Chassis configuration as a dense 1GbE or 10GbE server access switch.

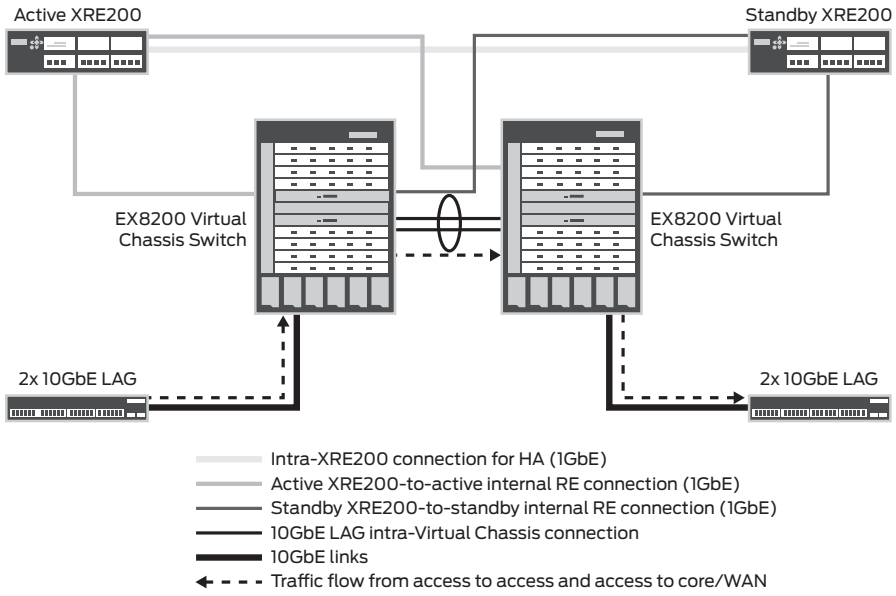

Figure 6: Connecting adjacent devices to the EX8200 switches in a Virtual Chassis configuration using non-fabric mode

Building Virtual Chassis Configurations over Long Distances

In large campus or data center environments where the distance between XRE200 devices and the EX8200 line chassis exceeds the maximum reach of a Cat 5 or Cat 6 cable, dedicated low-end Layer 2 switches such as the Juniper Networks EX2200 Ethernet Switch can be deployed in each location to act as media converters.

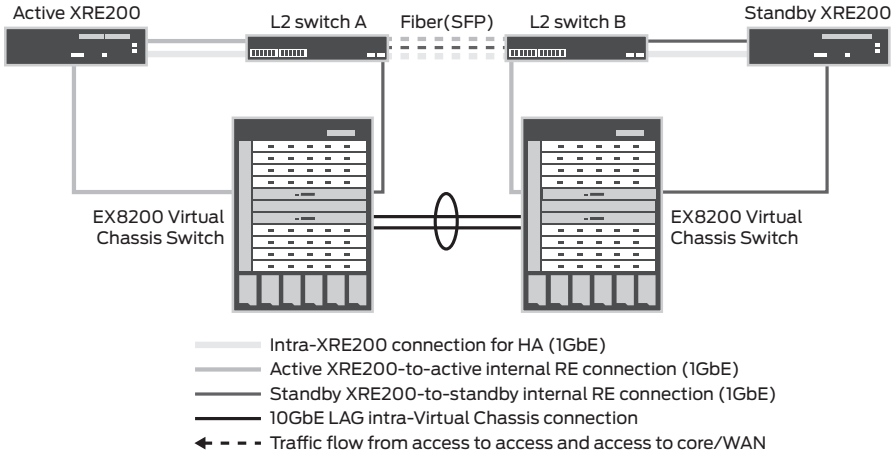

Figure 7: Using L2 switches to extend an EX8200 Virtual Chassis configuration

A port pair consisting of an RJ-45 and an SFP port on each L2 switch is required for every long-distance connection desired. All pair ports dedicated to supporting a long-distance connection are assigned to the same static VLAN. This simple configuration enables users to easily deploy EX8200 Virtual Chassis configurations in a wide variety of environments.

Nonstop Software Upgrade on the EX8200 Virtual Chassis Configuration

Upgrading the Junos OS on an EX8200 Virtual Chassis configuration can be accomplished in a non-disruptive fashion by using the nonstop software upgrade (NSSU) capabilities enabled by Juniper's Virtual Chassis technology. As long as all network-attached devices are dual-attached to more than one line card, network traffic continues to flow during the software upgrade process.

Conclusion

EX8200 switches with Virtual Chassis technology provide a simple, scaled and resilient solution for campus and data center core and collapsed core/aggregation layers, removing loops and eliminating the need for Spanning Tree Protocol by delivering multichassis LAG. EX8200 Virtual Chassis configurations also simplify the network by providing a single point of management and a scalable, resilient control plane.

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. **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.

2000348-004-EN May 2011 Printed on recycled paper