Deploying 10 Gigabit Ethernet with Cisco Nexus 5000 Series Switches

Introduction

Today's data centers are being transformed as the most recent server technologies are deployed in them. Multisocket servers with multicore processors deliver immense amounts of computing power in rack-mount servers and blade systems. This vast computing power has stimulated deployment of virtualization software that can more effectively put today's server capacity to use. The raw computing power of today's servers, combined with the higher utilization enabled by virtualization technology, places greater demands on IP networks than ever before.

The alignment of these two trends is encouraging the rapid adoption of 10 Gigabit Ethernet. Today, IT departments deploy 10 Gigabit Ethernet through the use of third-party adapter cards installed in server expansion slots. As server manufacturers respond to the demand for this level of bandwidth, an increasing number of rack-mount and blade systems will have multiple 10 Gigabit Ethernet ports integrated on the motherboard. Some blade systems now accommodate 10 Gigabit Ethernet switches integrated into the chassis. Other blade systems are likely to offer 10 Gigabit pass-through connections that make each blade's 10 Gigabit Ethernet ports available to access-layer switches, eliminating oversubscription within the blade chassis.

Regardless of the mechanism by which 10 Gigabit Ethernet enters the data center, IT departments must consider a new set of trade-offs when deploying it in data center racks. 10 Gigabit Ethernet is more than just a means to Gigabit Ethernet; planning is needed to use it as effectively and efficiently as possible. Factors to consider include the following:

- As server density increases, so does I/O bandwidth. 10 Gigabit Ethernet provides the bandwidth needed by servers with powerful multicore processors.
- Gigabit Ethernet is often supported by access-layer switches in intermediate distribution frames (IDFs) integrated into rows of server racks. With Gigabit Ethernet, the low cost of copper cabling made end-of-row switch configurations practical. With 10 Gigabit Ethernet, the cost of fiber transceivers and cabling is substantial. The alternative, 10GBASE-T, uses more expensive cabling, imposes higher latencies, and uses considerably more power. These constraints suggest use of a top-of-rack configuration with a better cabling solution.
- 10 Gigabit Ethernet may be accompanied with an increased demand for uplink capacity from the access to the aggregation layer. Indeed, as virtualization software simplifies the movement of virtual machines from server to server, IP networking requirements become more fluid and less predictable, requiring top-of-rack switching technology that can scale.

The Cisco Nexus™ 5000 Series is a family of line-rate, low-latency, low-cost-per-port 10 Gigabit Ethernet/Fibre Channel over Ethernet (FCoE) switches for data center access-layer applications. The Cisco® Nexus 5000 Series supports a cost-effective top-of-rack deployment model that delivers both the port density and the bandwidth needed by servers in both top-of-rack and pod-style environments. This document presents several scenarios that illustrate the effectiveness of these switches in supporting both rack-mount and blade systems in high-density configurations.

Port Density for Racks of 1RU Servers

The Cisco Nexus 5020 Switch, the first in the Cisco Nexus 5000 Series, has the port density and the low-cost cabling option to handle the densest rack-mount server configurations: racks full of one-rack-unit (1RU) servers, each with dual 10 Gigabit Ethernet interfaces. With a pair of 40-port Cisco Nexus 5020 Switches occupying a total of 4RU, a standard 42RU rack filled to capacity could contain up to 38 servers.

The configuration shown in Figure 1 connects one port from each server to one switch, and the second port from each server to the second switch. This configuration occupies 38 of the 40 fixed switch ports. The Cisco Nexus 5020 supports up to two 6-port expansion modules in each switch, allowing this configuration's number of uplink ports to range from 2 to 14. This range allows the rack-level oversubscription ratio to reach as low as 2.7:1 with a maximum of 14 uplinks configured, enabling IT departments to start with the oversubscription ratio of their choice and scale to 2.7:1 as traffic demands.

Figure 1.   The Cisco Nexus 5020 Can Handle the Densest 1RU Server Configurations with Favorable Oversubscription Ratios
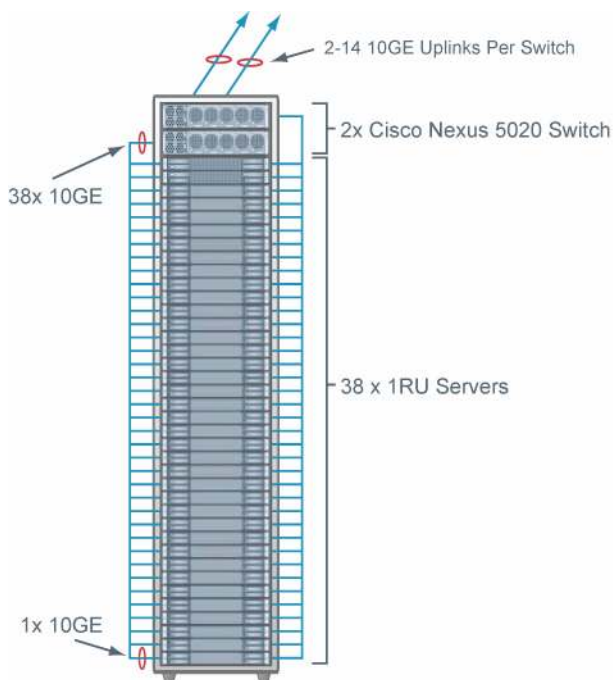
What makes this scenario so attractive is not just the high port density available in the Cisco Nexus 5000 Series; it is the cost effectiveness of the overall solution. The economics begin with the relatively low per-port cost of the switch itself. These economics are enhanced by an innovative solution that integrates SFP+ compatible connectors with Twinax cables into an energy-efficient, low-cost, and low-latency solution (Figure 2). Small Form-Factor Pluggable Plus (SFP+) direct-attach 10 Gigabit copper cables use only 0.1 watt (W) of power per connection and introduce only approximately 0.25 microsecond of latency per link. This solution, available in lengths up to 10 meters, enables both lower capital and operating expenses for in-rack cabling.

Figure 2. The Economics of Top-of-Rack Access Switch Deployment Scenarios Is Enhanced by SFP+ Direct-Attach 10 Gigabit Copper Cabling, Bringing Low Capital and Operating Costs to Rack-Level Deployments

For links to aggregation-layer switches such as the Cisco Nexus 7000 Series Switches, the Cisco Nexus 5000 Series supports multimode, short-reach optical SFP+ transceivers. These optical transceivers use only approximately 1W per transceiver, have a latency of less than 0.1 microsecond, and can reach up to 300 meters.

Both of these options provide lower latency and higher energy efficiency than 10GBASE-T. This standard uses transceivers that consume 4 to 8W per transceiver and contribute a latency of up to 2.5 microseconds per link, making the 10GBASE-T standard a significant contributor to network-level power consumption.

Pod Configuration with Racks of 2RU Servers

Because of the cost of 10 Gigabit Ethernet network interface cards (NICs) relative to the cost of 1RU servers, organizations are more likely to deploy 10 Gigabit Ethernet in racks of more powerful 2RU servers. In this scenario, a pair of Cisco Nexus 5020 Switches can provide redundant connectivity for a pod of two racks, each with 20 2RU servers, each with two 10 Gigabit Ethernet NICs configured for bandwidth and redundancy.

As Figure 3 illustrates, two identical racks could be configured with 20 2RU servers and a single Cisco Nexus 5020 in a top-of-rack configuration. One NIC on each server would connect to its own top-of-rack access-layer switch, and the other NIC would connect to the neighboring rack's switch. Cable length is well within the 10-meter limit for the SFP+ 10 Gigabit direct-attach copper cabling solution. The resulting configuration would use all 40 fixed ports on each of the switches. Uplinks are supported with up to two 6-port 10 Gigabit Ethernet expansion modules. For example, an oversubscription ratio of 10:1 can be established with a single 6-port module per switch populated with 4 uplinks. With two expansion modules and 12 uplinks to the aggregation layer, the oversubscription ratio can be scaled as low as 3.3 as bandwidth requirements dictate.

Figure 3. The Cisco Nexus 5020 Can Support Two-Rack Pods of 2RU Servers with an Oversubscription Ratio as Low as 3.3:1
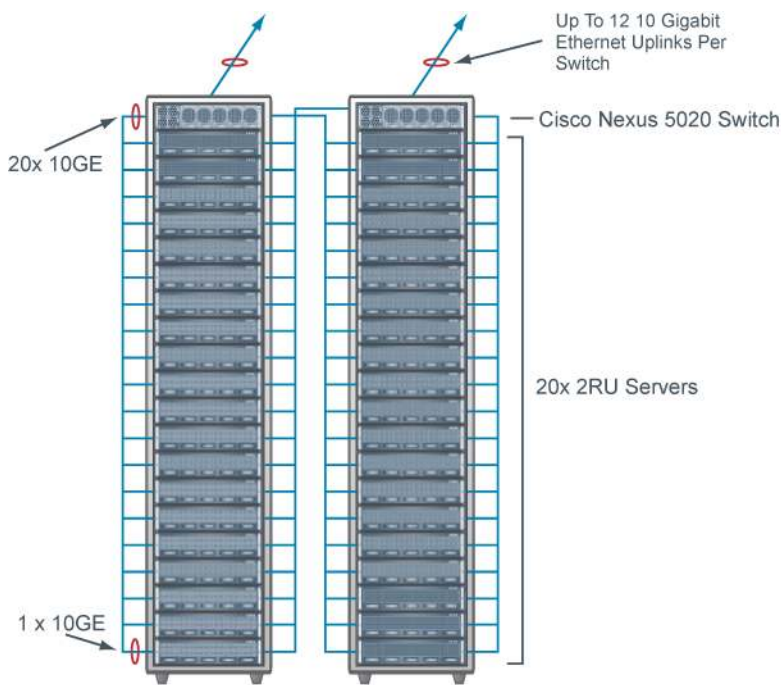
10 Gigabit Ethernet Networks for Blade Servers

Two configurations will be common for 10 Gigabit Ethernet networks on blade servers. Blade servers with internal 10 Gigabit Ethernet switches connect to each blade internally, with a set of oversubscribed 10 Gigabit links extending to the access-layer switch. Pass-through configurations extend each blade's 10 Gigabit Ethernet interfaces to the blade server chassis, where they can be extended to an access switch at the top of the rack. Trade-offs between these two technologies include the following:

- Blade server switches add the complexity of another layer of switching and management. They impose their own oversubscription ratios, requiring careful allocation of applications to blades so that the chassis-level switch does not become a bottleneck. The benefit of blade server switches is that they reduce in-rack cabling and port requirements on access-layer switches.
- Pass-through configurations send all network traffic to the access-layer switch, where it can be managed along with the rest of the data center network infrastructure, and where uplinks can be oversubscribed based on actual traffic requirements, rather than blade server switch limitations. Pass-through configurations demand high port density at the top of the rack, an area in which the Cisco Nexus 5020 excels

The Cisco Nexus 5020 is an ideal companion in both blade server networking scenarios.

Pods of Blade Servers with Chassis Switches

The Cisco Nexus 5020 provides a cost-effective means of aggregating uplinks from multiple blade servers, each with built-in 10 Gigabit Ethernet switches. Assuming that each blade server has four 10 Gigabit Ethernet uplinks, a pair of switches can be deployed to serve a pod of three racks with full redundancy. Figure 4 shows two racks configured with three blade servers each, and a central rack configured with two blade servers and two Cisco Nexus 5020 Switches. This configuration would use a total of 32 switch ports, leaving 8 switch ports available for uplinks. This configuration yields an oversubscription ratio of 4:1. If the two layers of oversubscription ratios, one in the chassis and one at the top of the rack, provide insufficient bandwidth, addition of one or two 6-port 10 Gigabit Ethernet expansion modules can scale the oversubscription ratio to as low as 1.6:1.

This configuration supports scaling to even higher densities: with two 6-port 10 Gigabit Ethernet expansion modules per switch, the pod can support three more blade systems with a total of 56 connections to servers and up to 8 uplinks, for an oversubscription ratio as low as 5.5:1.

The use of blade systems with in-pod switching becomes particularly compelling as SFP+ 10 Gigabit direct-attach copper technology becomes available in blade switches, reducing power consumption and latency compared to 10GBASE-T technology.

Figure 4. The Cisco Nexus 5020 Can Support a Pod Containing Three Racks of Blade Servers, Each with Four 10 Gigabit Ethernet Uplinks from Chassis-Level Switches
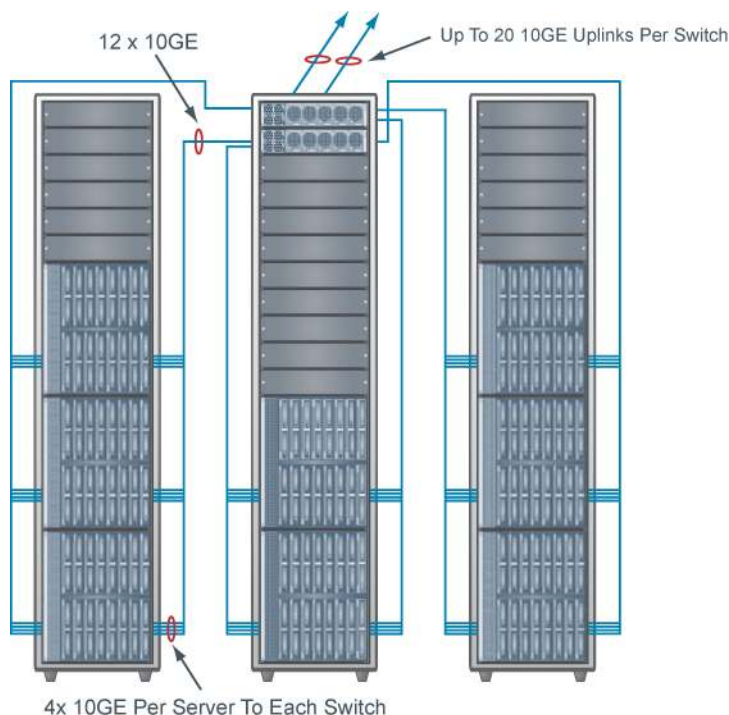
Port Density for Blade Servers with Network Pass-Through Connections

In a rack configuration with two blade servers with 16 blades each and two 10 Gigabit Ethernet NICs on each blade, 32 network links connect to each of a pair of Cisco Nexus 5020 Switches at the top of the rack (Figure 5). This configuration offers full redundancy and a 4:1 oversubscription ratio based on the switch's providing eight uplinks to the aggregation layer. Lower oversubscription ratios can be configured with the addition of one or two 6-port 10 Gigabit Ethernet expansion modules. One or two modules brings the number of uplinks to 14 or 20, respectively, with oversubscription ratios of up to 2.29:1 or 1.6:1, respectively.

Figure 5. In a Configuration with Two Blade Servers with 16 Blades Each, a Pair of Cisco Nexus 5020 Switches Provides Full Redundancy and an Oversubscription Ratio as Low as 1.6:1

Up To 20 10GE Uplinks Per Switch

2x Cisco Nexus 5020 Switch

32x 10GE

2x Blade Server Chassis With 16 Pass-Through 10GE Connections

16x 10GE
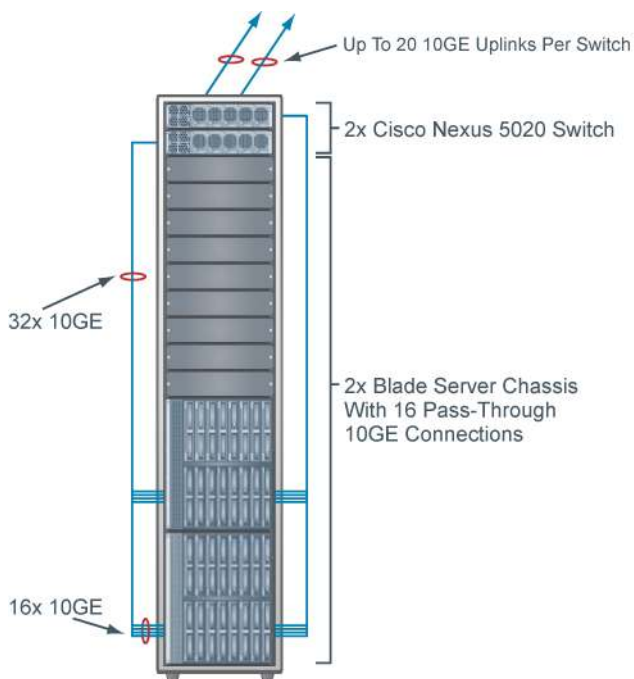
Conclusion

Cisco Nexus 5000 Series Switches pave the way for the adoption of 10 Gigabit Ethernet in data centers today, while providing the port density to support a range of server configurations well into the future. Today, as data centers deploy individual servers and racks of servers with 10 Gigabit Ethernet NICs, the Cisco Nexus 5020 can support high-density 1RU and 2RU server scenarios. As 10 Gigabit Ethernet pass-through connections and chassis-level switches become standard, the Cisco Nexus 5020 delivers the port density to support racks and pods of blade servers.

For More Information

http://www.cisco.com/go/nexus

http://www.cisco.com/go/dc

Printed in USA   C27-489248-01   05/09