Large SAN Design Best Practices

Introduction

As storage area networks (SANs) continue to grow in size, many factors need to be considered to help scale and manage them. This paper focuses on large SAN deployments within a data center, and provides best practices and design considerations when designing a large physical fabric. It does not address networks implementing Inter-VSAN Routing (IVR), SAN extension, or any intelligent fabric applications (for example, Data Mobility Manager, Storage Media Encryption, SANTap, EMC Invista, or Incipient iNSP).

Design Parameters

In SAN environments, many servers access a shared storage frame. The fan-out ratio varies between environments, but designing the network to maintain this ratio is vital to providing a predictable and scalable network. This document focuses on the following design parameters:

- 1000 or more end devices (servers, storage, and tape devices)
- Majority of end devices with connection speeds of 1G, 2G, or 4G
- Identical dual physical fabrics (Fabric-A, Fabric-B)

Cisco MDS 9000 Fibre Channel Modules

The Cisco MDS 9000 Family provides several Fibre Channel modules to meet various application requirements. Understanding the line cards' capabilities will help in designing the network. The line cards discussed here are presented in Table 1.

Table 1. Cisco MDS 9000 Fibre Channel Switching Modules

Type of Line Card		Cisco Part Number	Purpose
Generation 2 Speeds @ 1/2/4G	FC 12-port	DS-X9112	Recommended for ISLs and high-performance disk and tape array connectivity
	FC 24-port	DS-X9124	Recommended for virtual server or high-performance server connectivity
	FC 48-port	DS-X9148	Recommended for standard server connectivity
	MSM 18/4-port	DS-X9304-18K9	Recommended for SAN extension, iSCSI, and Fibre Channel connectivity
Generation 3 Speeds @ 1/2/4/8G	FC 24-port	DS-X9224-96K9	Recommended for ISLs and high-performance disk and tape array connectivity
	FC 48-port	DS-X9248-96K9	Recommended for virtual server or high-performance server connectivity
	FC 4/44-port	DS-X9248-48K9	Recommended for standard server connectivity

Cisco MDS 9000 Fibre Channel Switching Modules have the capability to provide Persistent Bandwidth Reservation (PBR). This feature allows the SAN administrator to reserve specific bandwidth for a specific port. When the bandwidth is needed for a specific port, it is immediately available. When not needed, the bandwidth is available to other ports in the same port group. This provides greater flexibility in designs involving higher port density modules.

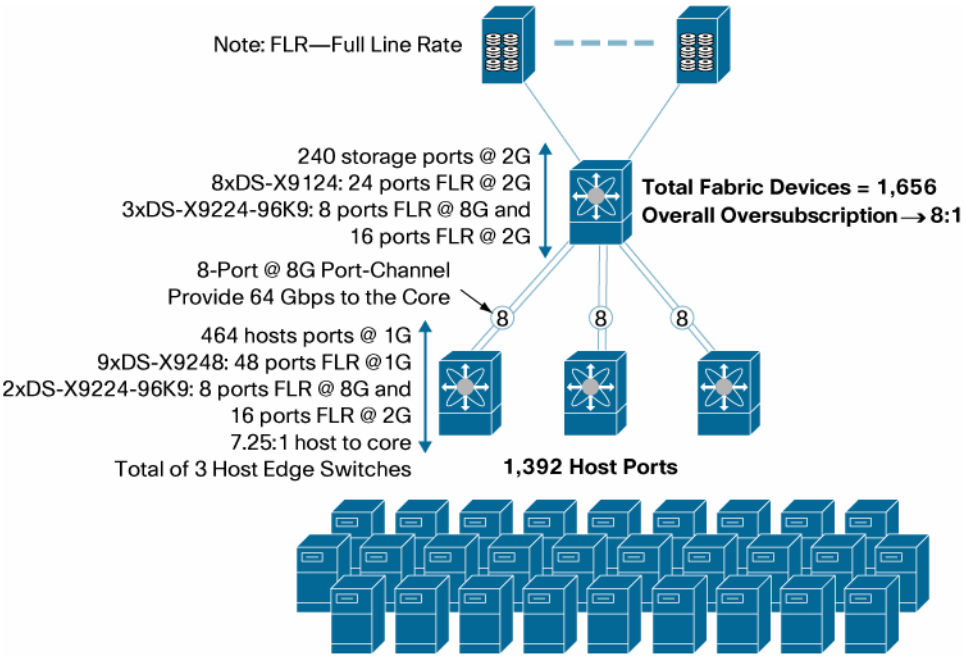
SAN Topology Considerations

It is common practice in SAN environments to build two separate, redundant physical fabrics (Fabric A and Fabric B) in case a single physical fabric fails. When designing for large networks, most environments will fall into two types of topologies within a physical fabric:

1. Two-tier: Core-edge design
2. Three-tier: Edge-core-edge design

Within the two-tier design, servers connect to the edge switches, and storage devices connect to one or more core switches (Figure 1). This allows the core switch to provide storage services to one or more edge switches, thus servicing more servers in the fabric. The interswitch links (ISLs) will have to be designed so that the overall fabric maintains both the fan-out ratio of servers to storage and the overall end-to-end oversubscription ratio.

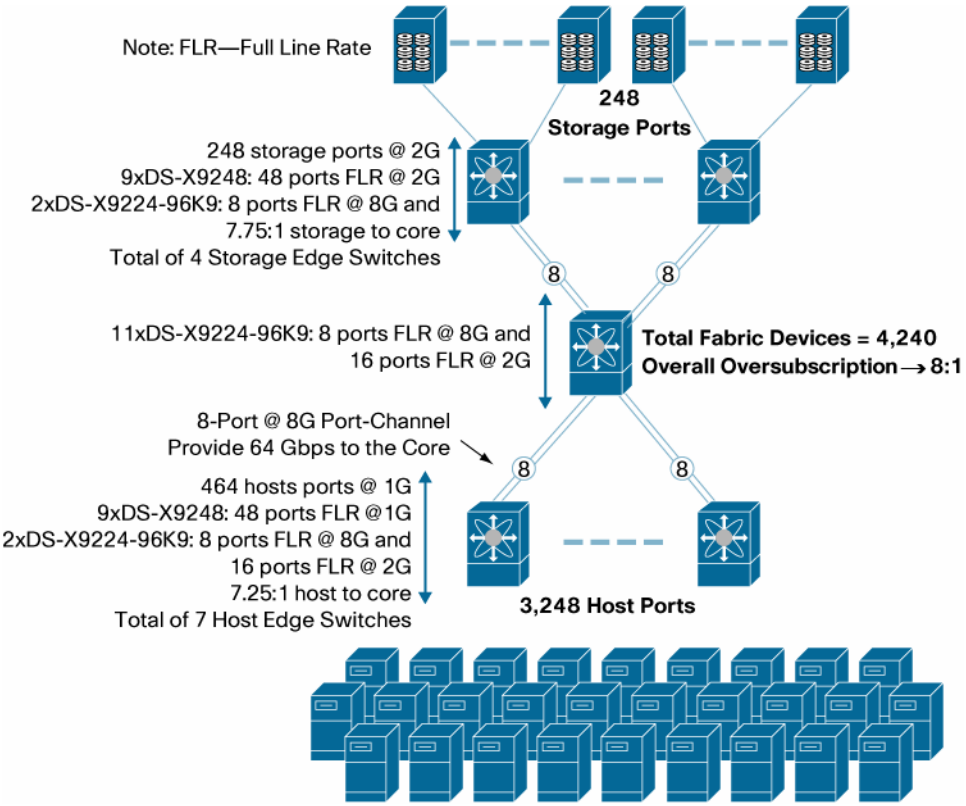
Figure 1. Sample Core-Edge Design



Note: Cisco product IDs are shown in the topology diagram. Please use Table 1 for reference.

In environments where future growth of the network has the number of storage devices exceeding the number of ports available at the core switch, a three-tier design may be ideal (Figure 2). This type of topology still uses a set of edge switches for server connectivity, but adds another set of edge switches for storage devices. Both sets of edge switches connect into a core switch via ISLs.

Figure 2. Sample Edge-Core-Edge Design



Note: Cisco product IDs are shown in the topology diagram. Please use Table 1 for reference.

Network Considerations

When designing a large Cisco MDS fabric, the following should be taken into consideration:

- Number of VSANs
- Number of ISLs
- Effects of fabric logins
- Type of zones

Virtual SANs (VSANs)

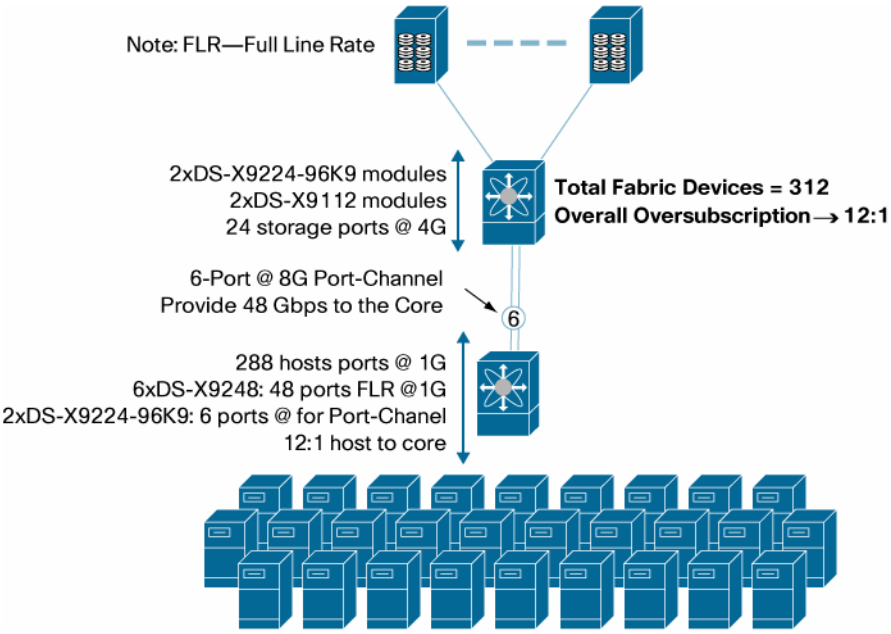
Cisco MDS switches offer VSAN technology, which is a simple and secure way to consolidate many SAN islands into a single physical fabric. Separate fabric services and separate role base management are provided for each VSAN, while providing separation of both the control plane and the data plane.

There are multiple use cases for VSANs, such as creating a VSAN for each type of operating system (e.g., VSAN for Windows or HP-UX), or utilizing them on the basis of business functions (e.g., a VSAN for development, for production, or for lab). VSAN 1 is created on the Cisco MDS switch by default and cannot be deleted. As a best practice, VSAN 1 should be used as a staging area and another VSAN should be created for the production environment. With each VSAN having its own zonesets and zones, Cisco MDS switches enable a scalable, robust network.

ISLs

The number of ISLs between the Cisco MDS switches will depend on the desired end-to-end oversubscription (Figure 3). The fan-out ratio from a single storage port to multiple servers can help determine the number of ISLs needed for each edge-to-core connection.

Figure 3. Number of ISLs Needed to Maintain a 12:1 Oversubscription Ratio



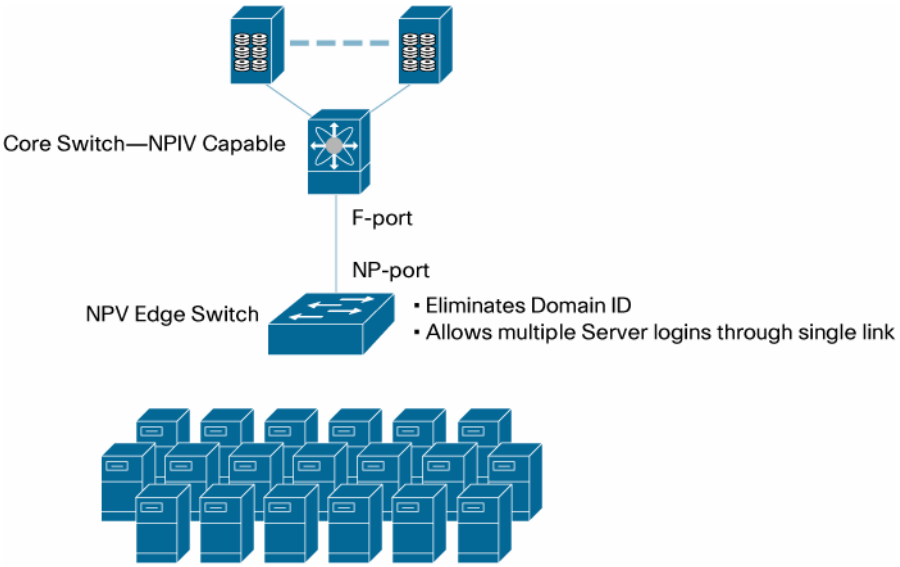
Note: Since there are 24 storage ports @ 4G in this example, a total of 48 Gbps of bandwidth is needed to maintain the oversubscription ratio in this environment. Since the ISLs are 8G-capable, six ISLs are required. With this configuration, the environment still has room to grow at the core and expand with more edge switches to support a growing number of servers.

Fabric Logins

In a Fibre Channel network, the actual number of physical ports in the fabric is not the most critical concern when designing for large SAN fabrics. Since Fibre Channel ports consist of E/TE ports and F/FL ports, the main consideration is the number of fabric logins in the network. The number of actual physical ports in the fabric is larger than the number of end devices (server, storage, and tape ports) in the physical fabric. The Cisco MDS Family supports up to 10,000 fabric logins in a physical fabric, regardless of the number of VSANs in the network.

Between the increase in blade server deployments and the consolidation of servers due to server virtualization technologies, the design of the network will be affected. With features such as N_Port ID Virtualization (NPIV) and Cisco N_Port Virtualization (NPV), the number of fabric logins has further increased (Figure 4). The proliferation of NPIV-capable end devices such as host bus adaptors (HBAs} and Cisco NPV-mode switches makes the number of fabric logins on a per-port, per-line-card, per-switch, and per-physical-fabric basis a critical consideration. These fabric login limits will determine the design of the current SAN, as well as future growth.

Figure 4. Cisco NPV-Enabled Switches and Fabric Logins



Note: Prior to NPIV and NPV, a single port had a maximum of one fabric login. With NPIV and Cisco NPV-enabled switches, a single port can now support multiple fabric logins.

Zones

Within each VSAN, there is only one active zoneset that contains one or more zones. Each zone consists of one or more members to allow for communication between the members. The Cisco MDS SAN-OS and NX-OS provide multiple ways to identify zone members, but the commonly used ones are listed below:

- device-alias: Allows creation of easy-to-read single name associated with a single device's Port World-Wide Name (pwwn)
- pwwn: Port World-Wide Name of the device (most commonly used)

Depending on the requirements of the environment, choosing the type of zone members is a matter of preference. A recommended best practice is to create a "device-alias" for end devices when managing the network. The device-alias provides an easy-to-read name for a particular end device. For example, a storage pwwn 50:06:04:82:bf:d0:54:52 can be given a device-alias name of Tier1-arrayX-ID542-Port2. In addition, with device-alias, when the actual device moves from one VSAN (VSAN 10) to a new VSAN (VSAN 20) in the same physical fabric, the device-alias name will follow that device. So there is no need to re-enter the device-alias for the new VSAN.

Cisco MDS switches support up to 8000 zones and 20,000 zone members in a physical fabric. There are things to consider for very large environments that may reach this limit. Since the MDS switch implements hardware-based zoning, the number of members in the zone may affect how many entries get used in the hardware.

For example: Server-A needs access to Array-A and Array-B.

Type of Zone		Effects of the Zone
Two zones	Zone1 (Server-A and Array-A)	This will allocate two entries in the hardware: Server-A → Array-A Server-A → Array-B
	Zone2 (Server-A and Array-B)	

Type of Zone		Effects of the Zone
Single zone	Zone1 (Server-A, Array-A, and Array-B)	This will allocate three entries in the hardware: Server-A → Array-A Server-A → Array-B Array-A → Array-B

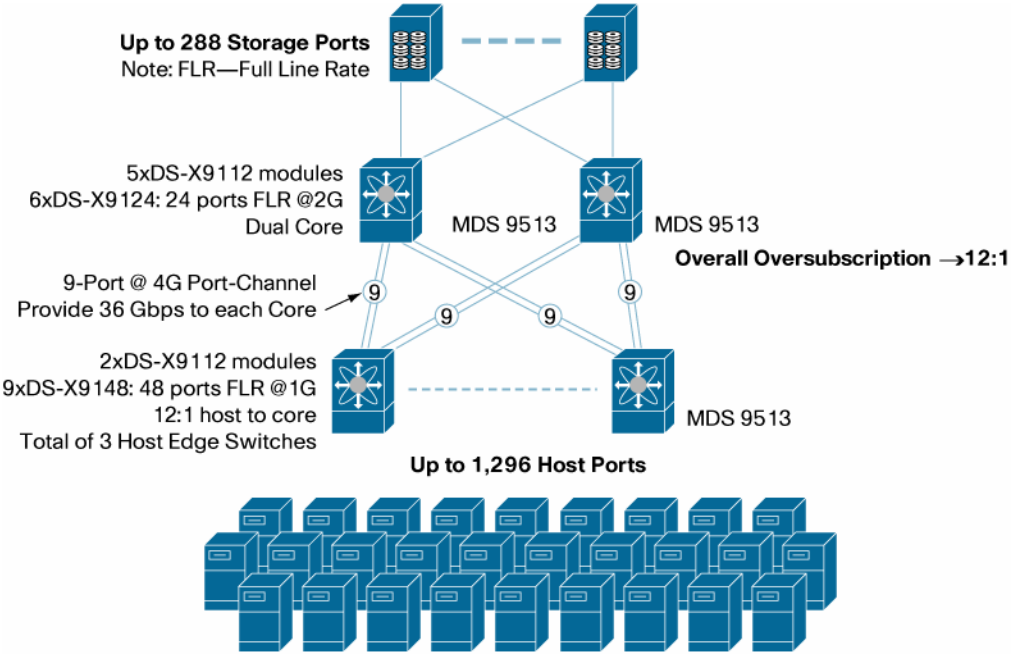
Note: As a best practice for large SAN deployments, it is recommended to have more zones with two-member zones versus a single zone with three or more members. This is not a concern in smaller environments.

Sample Use Case Deployments

Below are two sample deployments of large-scale Cisco MDS fabrics, one with more than 1000 devices in the fabric (Figure 5) and one with more than 4000 (Figure 6).

Sample Deployment 1

Figure 5. Use Case 1 Topology



Note: This deployment allows scaling to nearly 1500 devices in a single fabric. The actual production environment has approximately 190 storage ports and roughly 1050 host ports. The environment required a minimum of 12:1 oversubscription within the network, which required each host edge switch to have a 36-Gbps port channel, using nine physical links. Storage ports will not grow quite as rapidly and the core switch has room to grow to add more host edge switches.

In this environment, the following was used in managing the network:

- Total of four VSANs created
  - VSAN 1 used for staging new devices
  - VSAN based upon operating systems

    VSAN XX for Windows/Linux

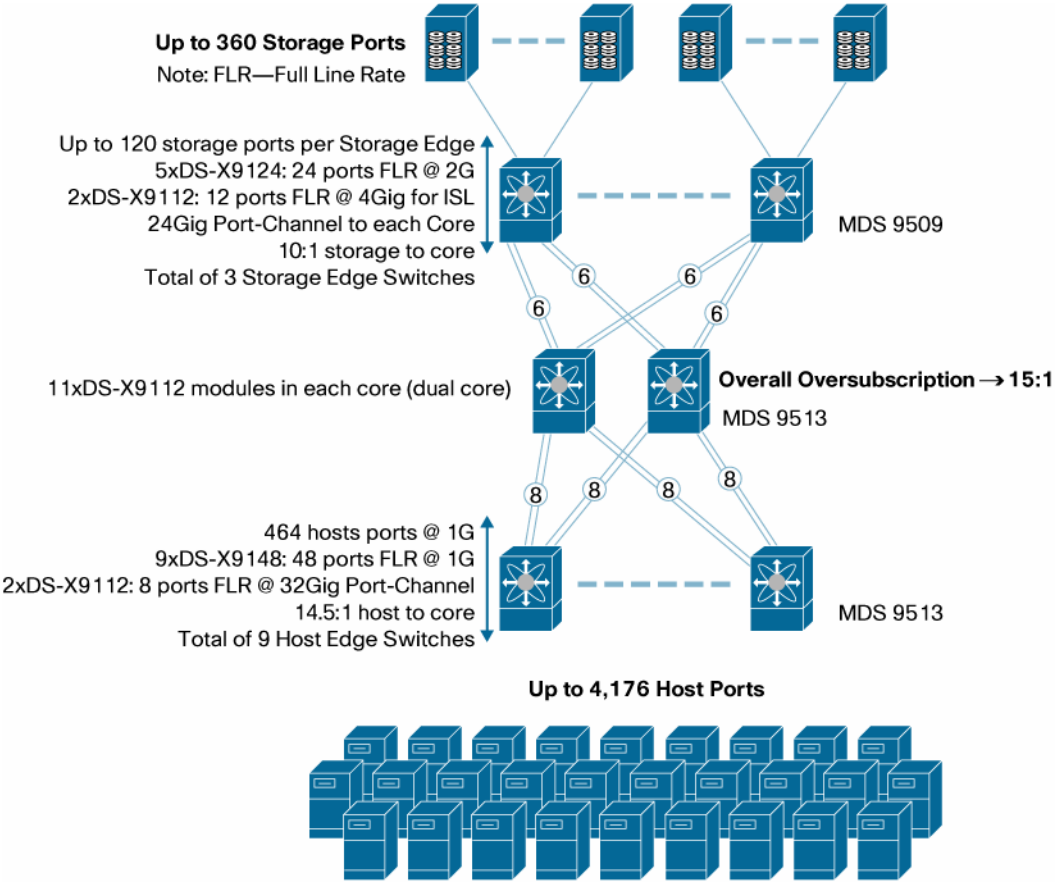
    VSAN YY for HP-UX environment

- VSAN ZZ for tape backup

- Used TACACS+ for authorization and auditing of MDS switches
- Created separate administrative roles for VSANs
- Created device-alias for environment
- Created two member zones of device-alias

Sample Deployment 2

Figure 6. Use Case 2 Topology



Note: This deployment scales to nearly 4500 devices in a single fabric and still has room to grow. The actual production environment has approximately 280 storage ports, 20 tape devices, and roughly 3800 host ports. The environment required a minimum of 15:1 oversubscription within the network, which required each host edge switch to have a 32-Gbps port channel, using eight physical links at the host edge. The storage edge was designed for 10:1 oversubscription, which will give certain applications the proper bandwidth when needed.

In this environment, the following was used in managing the network:

- Total of seven VSANs created
  - VSAN 1 used for staging new devices
  - Four VSANs based upon business operations
  - VSAN for tape backup

    - VSAN for storage replication
- Used TACACS+ for authorization and auditing of Cisco MDS switches
- Created separate administrative roles for VSANs
- Created device-alias for environment
- Enabled Dynamic Port VSAN Membership (DPVM) feature
- Mixture of two and three member zones

Summary

With data centers continually growing, SAN administrators must design networks that meet their current needs and can scale for demanding growth. Cisco MDS switches provide embedded features to help SAN administrators in these tasks. SAN administrators deploying large Cisco SAN fabrics can use the design parameters and best practices discussed in this paper to design optimized and scalable SANs.

Appendix

Fabric Scalability Numbers: Cisco SAN OS 3.2 or Higher

Limits per Physical Fabric	Cisco Validated Numbers	Cisco Maximum Limits
Number of VSANs	80	4000
Number of switches	75	239
Number of fabric logins	10,000	10,000
Number of fabric logins per switch	2000	2000
Number of fabric logins per line card	400	400
Number of fabric logins per port	256	256
Zone members	16,000 (includes all VSANs)	16,000 (includes all VSANs)
Zones	8000 (includes all VSANs)	8000 (includes all VSANs)
Zonesets	500 (includes all VSANs)	1000 (includes all VSANs)

NPV Scalability Numbers: Relevant to Cisco MDS 9124, MDS 9134, and Blade Switches in NPV Mode

Limits per Physical Fabric	Validated Numbers
Number of logins per port	114
Number of logins per port group	114
Number of logins for Cisco MDS 9124 (six port groups)	684
Number of logins for Cisco MDS 9134 (eight port groups)	912
Number of logins for HP MDS 9124e (six port groups)	684
Number of logins for IBM FC Blade Switch (five port groups)	570
Number of NPV switches connected in a physical fabric	105

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

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

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

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