White Paper

Achieve Fabric Resiliency with Cisco Unified Data Center Architecture

What You Will Learn

This document addresses technical professionals who want to increase the resiliency, efficiency, and flexibility of their data center fabric to support virtualization and cloud capabilities. Transformation of the data center is a process with predictable IT challenges along the way. One of the first steps is to create a highly resilient, scalable and secure network fabric that can support virtualized and cloud-based applications. This document focuses on the capabilities and technology options of Cisco® Unified Fabric for building massively scalable and highly resilient fabrics. After reading this document, you will have an understanding of how Cisco Unified Data Center solutions can increase IT simplicity, financial efficiency, and the flexibility of your data center.

Evolving Your Data Center Platform

The data center has undergone numerous evolutionary changes over the past several decades, and each change has been defined by major shifts in computing, including virtual machine mobility, IT resource pooling, and cloud computing. These shifts have resulted in a transition to tighter integration and coupling of the major data center tiers: applications, storage, servers, and the network. Today's virtualized data centers continue to move away from numerous vertically integrated silos and toward a fluid, dynamic fabric capable of moving IT resources to wherever they are needed as the business dictates.

The network is at the core of the data center fabric and is vital to IT's capability to deliver service and value back to the business. Enterprises transitioning to a single unified fabric understand the operational and financial benefits of being able to manage just a single network. New management technologies have made network deployment, configuration, and ongoing maintenance extremely easy and user friendly, which both lowers operating costs and significantly reduces the potential for human error.

Unified fabric is at the center of the fabric resiliency solution. It is tightly integrated with both the computing and storage layers and uses a range of technologies that provide massive scalability and fabric resiliency. With the Cisco Unified Data Center architecture, Cisco delivers exceptional levels of fabric resiliency, high availability, and scalability while preserving customer choice, allowing customers to deploy the technologies they prefer when they design their fabrics. Network technology options for network-wide fabric resilience and high availability include spanning tree, virtual PortChannel (vPC), and Cisco FabricPath.

These capabilities have been tested and validated in Cisco labs on a standardized data center reference architecture called the Cisco Virtualized Multiservice Data Center (VMDC) architecture. The Cisco VMDC reference architecture provides a framework for a building fabric-based infrastructure using the Cisco Unified Data Center platform. Cisco VMDC provides design guidelines that demonstrate how customers can integrate Cisco and partner technologies such as networking, computing, integrated computing stacks, security, load balancing, and system management into a data center architecture that supports critical IT initiatives such as consolidation and virtualization, including desktop virtualization; application migration and rollout; public, private, and hybrid cloud deployments; business continuance and disaster recovery; and the build out of new data centers.

Cisco Virtualized Multiservice Data Center Overview

The Cisco VMDC reference architecture for cloud deployments has been widely adopted by enterprises worldwide. Cisco VMDC has provided design guidance for scalable, secure, resilient, data center infrastructure. This infrastructure is based on hierarchichal data center network designs using leading Cisco platforms and technologies, with necessary network-based services as well as orchestration and automation capabilities to accommodate the needs of various cloud providers and consumers.

Cisco Unified Fabric

Cisco Unified Fabric is a single, flexible, and highly scalable network infrastructure for your data center. Cisco Unified Fabric provides the secure LAN and SAN switching infrastructure with a single network OS that can deliver separate or consolidated I/O with the intelligence to be virtual machine aware and to support high levels of virtual machine mobility and security. Cisco Unified Fabric provides a single point of connectivity and management across physical, virtual, and cloud resources, dramatically reducing management and operating costs.

Cisco Unified Fabric provides:

- Architectural flexibility and scalability
- Consolidated I/O
- Workload mobility
- Simplified management
- Virtual machine-aware networking

Intelligent network services differentiate Cisco's fabric-based platform from commodity infrastructure by helping ensure that virtualized applications remain associated with the required availability, security, acceleration, workload balancing, and performance monitoring services. Unified network services help ensure compliance and provide consistent service delivery using flexible, policy-based provisioning.

Data Center Fabric Resiliency

Today's business environments require nonstop operations. Downtime means lost customers, lost opportunity, and lost revenue. The data center fabric needs to be built with resiliency and self-healing as priorities to help ensure the most uptime with few disruptions.

The Cisco Unified Fabric architecture is designed to optimize service uptime by enabling availability and fault tolerance at all layers of the data center through the use of a combination of physical redundancy best practices as well as virtualized failover features at the network, computing, and storage layers. At the network layer, physical redundancy starts with hardware redundancy within a node, redundant nodes, and redundant links, with optimized failover convergence across the entire system. At the physical level, redundancy includes dual-homing servers at the access layer and redundant switches in the distribution layer and core. Enterprises also have a choice of redundancy protocols to help ensure link or switch failover should the need occur. Available protocols include link redundancy technologies such as spanning tree and technologies such as vPC and Cisco FabricPath that provide link redundancy and active-active links for increased bandwidth. The Hot-Standby Router Protocol (HSRP) and Virtual Router Redundancy Protocol (VRRP) provide default gateway functions. Routing protocols such as Border Gateway Protocol (BGP) with nonstop forwarding (NSF), Open Shortest Path First (OSPF), and Enhanced Interior Gateway Routing Protocol (EIGRP) are critical to helping ensure the availability and recovery of the routed backbone to which the data center connects. Table 1 summarizes the protocol functions.

Table 1. Some Cisco Network Resiliency Features and Their Functions

Feature	Function
Spanning Tree Protocol	Spanning Tree Protocol is a network protocol that helps ensure a loop-free topology for any bridged Ethernet LAN. The basic function of Spanning Tree Protocol is to prevent bridge loops and the broadcast radiation that results from them. Spanning Tree Protocol also allows a network design to include spare (redundant) links to provide automatic backup paths in the event that an active link fails, without the danger of bridge loops or the need to manually disable and enable these backup links.
Virtual PortChannel (vPC)	A vPC is a PortChannel that can operate between more than two devices. Although multiple devices are used to create the vPC, the terminating device sees the vPC as one logical connection. The main advantages of vPC links are enhanced system availability and rapid recovery in the event of a link failure. This capability is primarily the result of the use of the IEEE PortChannel specification instead of spanning tree for loop management and forwarding.
Cisco FabricPath	Cisco FabricPath is an innovation in Cisco NX-OS Software that brings the stability and scalability of routing to Layer 2. The switched domain does not have to be segmented anymore, providing data center-wide workload mobility. Because traffic is no longer forwarded along a spanning tree, the bisectional bandwidth of the network is not limited, and massive scalability is now possible.
Layer 2 Multipathing	A Layer 2 Multipathing network allows you to remove unused switches in the your network and increase the utilization of all your assets. You no longer need to manually configure Spanning Tree Protocol for redundant paths because they are built into the protocol, and you need less equipment to do the same job and get a lot more bandwidth for your network.
Hot-Standby Router Protocol (HSRP)	HSRP is a Cisco protocol used to provide gateway redundancy for users on the LAN. The protocol establishes a relationship between network routers to achieve default gateway failover in the event that the primary gateway becomes inaccessible, in close association with a rapid-converging routing protocol such as EIGRP or OSPF.
Virtual Router Redundancy Protocol (VRRP)	VRRP is a Cisco protocol that automatically assigns available IP routers to participating hosts, increasing the availability and reliability of routing paths through automatic default gateway selection. The default gateway of a participating host is assigned to the virtual router instead of a physical router. If the physical router that is routing packets on behalf of the virtual router fails, another physical router is selected to automatically replace it. The physical router that is forwarding packets at any given time is called the master router.
Bidirectional Forwarding Detection (BFD)	BFD is a detection protocol designed to provide fast forwarding-path-failure detection times for all media types, encapsulations, topologies, and routing protocols. In addition to fast forwarding-path-failure detection, BFD provides a consistent failure detection method for network administrators.

Fabric Resiliency Design Considerations

Challenges in Current Network Design

Although Layer 2 switching may provide the flexibility critical to the operation of a large data center, it also has some shortcomings compared to a routed solution. The Layer 2 data plane is susceptible to frame proliferation. The forwarding topology, typically but not necessarily computed by the Spanning Tree Protocol, must be loop free at any cost; otherwise, frames could be replicated at wire speed and affect the entire bridged domain.

This restriction prevents Layer 2 from taking full advantage of the available bandwidth in the network, and it often creates suboptimal paths between hosts over the network. Also, because a failure could affect the entire bridged domain, Layer 2 is confined to small islands to contain risk.

Current Layer 2 Domain Designs

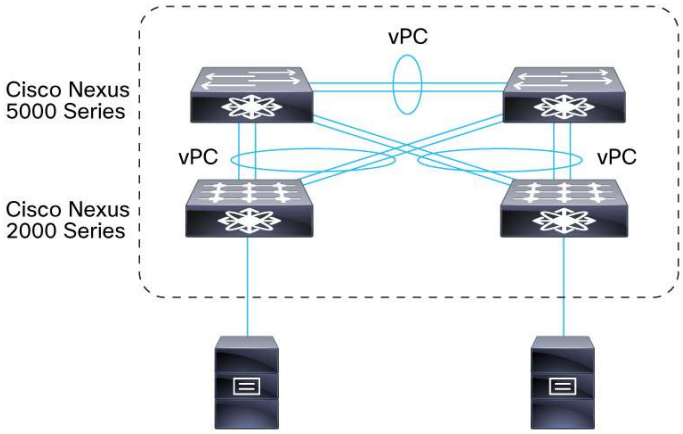
Most existing Layer 2 domains use the traditional Spanning Tree Protocol. However, spanning tree is inefficient in its use of the bandwidth of redundant links. Spanning Tree Protocol deployments usually have at least half the links blocked and do not participate in traffic forwarding. Another disadvantage is convergence; each time a network changes, the spanning tree has to be recalculated. This disadvantage especially applies to Layer 2 domains.

Layer 2 domains designed with vPC technology make better use of the redundant links, but still use Spanning Tree Protocol as a backup mechanism. As a result, a user still must use spanning-tree best practices.

Introducing vPC

vPC technology is a critical adjunct to spanning tree in bringing redundancy and resiliency to data center networks. vPC is easy to set up and configure. The basic function of vPC is to join two switches together and make them act as if they were a single logical switch for all connected devices (Figure 1).

Figure 1. vPC Concept



The vast majority of Layer 2 deployments on Cisco Nexus® platforms are based on the vPC feature. vPC extends standard PortChannel technology across a pair of switches and provides a redundancy mechanism that protects against both link and device failures. Because PortChannels operate at a layer below the Spanning Tree Protocol, the solution is immune to the main problems associated with Spanning Tree Protocol:

- Unlike Spanning Tree Protocol, vPC allows several parallel links to be active. This form of Layer 2 Multipathing provides additional bandwidth between networking tiers.
- No links are blocked by the spanning tree in a vPC environment. The risks associated with Spanning Tree Protocol are thus greatly reduced because loops typically result from the failure of Spanning Tree Protocol to block a port.
- vPC convergence is a local event and generally occurs more quickly than a networkwide Spanning Tree Protocol recomputation.
- vPC convergence is invisible to Spanning Tree Protocol and does not cause ports to be synchronized or MAC address tables to be flushed.

Best-Practice Design Objectives with vPC

A vPC allows links that are physically connected to two different Cisco Nexus 5000 or 7000 Series Switches to appear as a single PortChannel to a third device. The third device can be a Cisco Nexus 2000 Series Fabric Extender or a switch, server, or any other networking device.

vPC Network Benefits

The vPC domain includes both vPC peer devices, the vPC peer keepalive link, the vPC peer link, and all the PortChannels in the vPC domain connected to the downstream device. You can have only one vPC domain ID on each device.

A vPC provides the following benefits:

- Allows a single device to use a PortChannel across two upstream devices
- Eliminates Spanning Tree Protocol blocked ports
- Provides a loop-free topology
- Uses all available uplink bandwidth

- Provides fast convergence if either the link or a device fails
- Provides link-level resiliency
- Helps ensure high availability

The value of vPC is that it allows the implementation of a redundant network that looks like a nonredundant network from the perspective of the Spanning Tree Protocol. As a result, most of the management complexity associated with the Spanning Tree Protocol is eliminated. Otherwise, the vPC is managed like a traditional Layer 2 environment.

Layer 2 Designs with Cisco FabricPath

Until recently, data centers have been designed with high availability as the main priority. Just like the organizations they serve, modern networks must now be agile and accommodate changes in a flexible way. The simple response to this additional requirement would be to increase the size of the Layer 2 domain, because switching allows you to move devices and modify the infrastructure in a way that is transparent to servers. However, existing switching technologies have inefficient forwarding schemes based on spanning trees and cannot be extended to the network as a whole. Therefore, current designs are a compromise between the flexibility provided by Layer 2 and the scalability offered by Layer 3.

Cisco FabricPath is a routed alternative to Spanning Tree Protocol and Cisco vPC technology. Cisco FabricPath provides routed protection against spanning-tree loops, it keeps all links forwarding (unlike Spanning Tree Protocol), and it is a little easier to configure than vPC, especially as your data center becomes larger. However, Cisco FabricPath is not as mature as these other technologies, and it is proprietary to Cisco, whereas TRILL provides the standard for behavior similar to that of Cisco FabricPath.

By introducing a new control protocol (based on the Intermediate System-to-Intermediate System [IS-IS] Protocol) and a new data plane, Cisco FabricPath can work around most of the limitations that affect a traditional Ethernet network. In fact, Cisco FabricPath offers the benefits of both Layer 2 and Layer 3 technologies. Equal Cost Multipath (ECMP) allows Cisco FabricPath to use the total bandwidth of multiple parallel links.

Cisco FabricPath frames, including time-to-live (TTL) and reverse-path forwarding check (RPFC) frames, are applied to multidestination traffic. Also, unlike vPC, Cisco FabricPath can handle an arbitrary network topology.

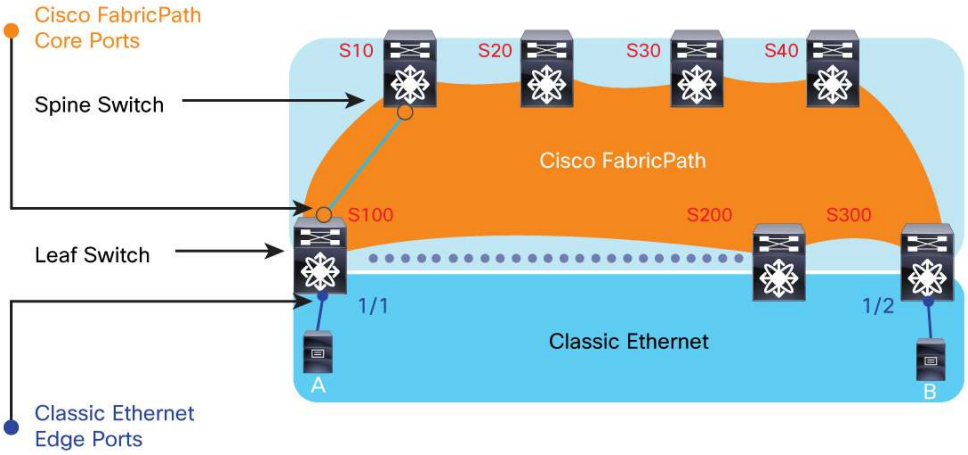
All the elements borrowed from Layer 3 technologies make Cisco FabricPath safe for extension to an entire data center without the risk of looping and allow easy reconfiguration of a network with little disruption.

Cisco FabricPath Design Considerations

A typical Cisco FabricPath network topology is the Clos fabric, shown in Figure 2. A Clos fabric consists of two kinds of node: leaf switches and spine switches. A particular leaf switch is connected to all the spine switches, and a particular spine switch is connected to all the leaf switches.

The goal for the network is to provide optimal connectivity between the leaf switches, with the hosts attached using Classic Ethernet (edge) ports.

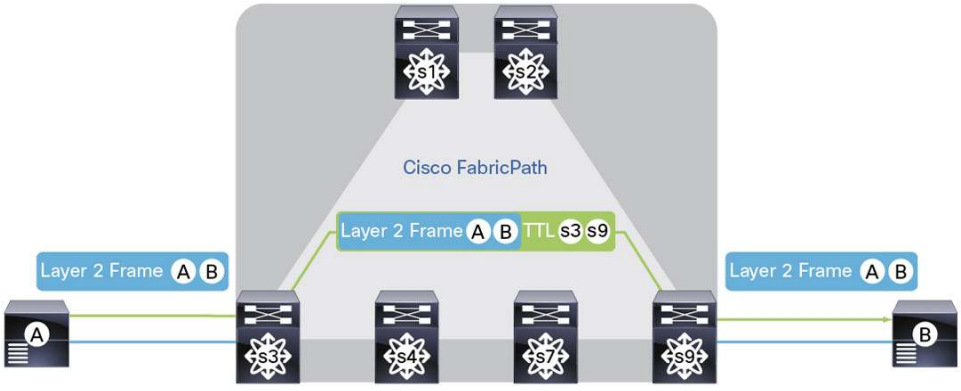
Figure 2. Typical Clos Fabric Design



Cisco FabricPath Traffic Routing within the Fabric

Cisco FabricPath brings the stability and performance of routing to Layer 2 by introducing an entirely new Layer 2 data plane. Cisco FabricPath takes over as soon as an Ethernet frame transitions from an Ethernet network (referred to as Classic Ethernet) to a Cisco FabricPath fabric. Ethernet bridging rules do not dictate the topology and the forwarding principles in a Cisco FabricPath fabric. The frame is encapsulated with a Cisco FabricPath header, which consists of routable source and destination addresses. These addresses are simply the address of the switch on which the frame was received and the address of the destination switch to which the frame is heading. From there on, the frame is routed until it is reaches the remote switch, where it is deencapsulated and delivered in its original Ethernet format, as shown in Figure 3.

Figure 3. Combination of Cisco FabricPath and Classic Ethernet Design



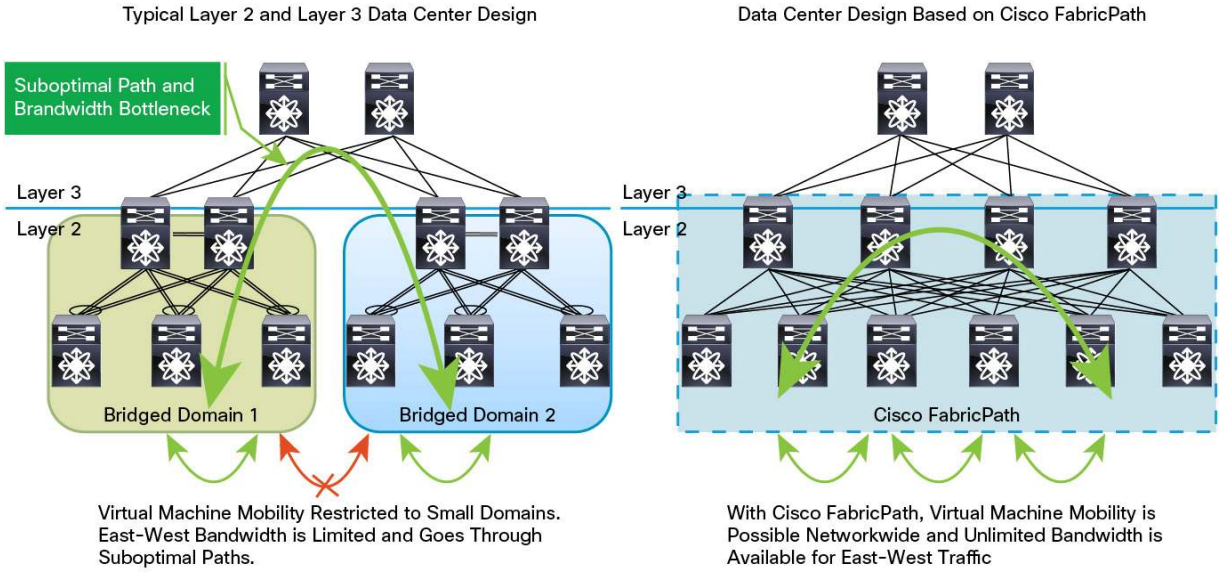
The fundamental difference between Cisco FabricPath and Classic Ethernet is that with Cisco FabricPath, the frame is always forwarded in the core using a known destination address. The addresses of the bridges are automatically assigned, and a routing table is computed for all unicast and multicast destinations. The resulting solution still provides the simple and flexible behavior of Layer 2, while also using the routing mechanisms that make IP reliable and scalable.

Scaling the Typical Data Center Design with Cisco FabricPath

This section demonstrates how Cisco FabricPath can bring additional significant scalability, availability, and flexibility improvement by reorganizing the cabling of an existing data center. Figure 4 shows two data centers using the exact same number of links and switches. In data center A, each access switch is connected through a 4-port PortChannel to two aggregation switches in a vPC domain. In data center B, which supports Cisco FabricPath, each access switch is instead connected through a single uplink to four aggregation switches.

This example shows how Cisco FabricPath brings the stability and scalability of routing to Layer 2. The switched domain does not have to be segmented anymore, providing data center-wide workload mobility. Because traffic is no longer forwarded along a spanning tree, the bisectional bandwidth of the network is not limited, and massive scalability is now possible.

Figure 4. Link Optimization with Cisco FabricPath



vPC+ Environment Migration: Bringing vPC to Cisco FabricPath Design

The vPC+ feature was introduced to allow interoperability between Cisco FabricPath and vPCs. The functions and behavior of vPC+ and vPC are identical. The same rules apply in both technologies: that is, both require peer-link and peer-keepalive messages, the configurations of the vPC peers must match, and consistency checks still take place. In a vPC+ domain, a unique Cisco FabricPath switch ID is configured, and the peer link is configured as a Cisco FabricPath core port. This Cisco FabricPath switch ID in the vPC+ domain is called the emulated-switch ID. The emulated-switch ID must be the same for the two peers and must be unique per vPC+ instance.

The benefits of vPC+ at the edge of a domain are:

- Allows you to attach servers to the device using Link Aggregation Control Protocol (LACP) uplinks
- Allows you to attach other Classic Ethernet devices in vPC mode
- Allows you to attach Cisco Nexus 2000 Series Fabric Extenders in active-active mode
- Prevents orphan ports in a failure scenario; when a peer link fails in the vPC+ domain, the orphan port still has Cisco FabricPath uplinks for communication
- Provides numerous paths

Cisco FabricPath Benefits

Cisco FabricPath provides the following benefits:

- Simplicity that reduces operating expenses.
  - Cisco FabricPath is extremely simple to configure. In fact, the only necessary configuration consists of distinguishing the core ports, which link the switches, from the edge ports, where end devices are attached. You do not need to tune any parameter to get an optimal configuration, and switch addresses are assigned automatically.
  - A single control protocol is used for unicast forwarding, multicast forwarding, and VLAN pruning. The Cisco FabricPath solution requires less combined configuration than an equivalent spanning-tree-based network, further reducing overall management costs.
  - Static network designs make some assumptions about traffic patterns and the locations of servers and services. If those assumptions are incorrect, a situation that often occurs over time, a complex redesign may be necessary. A network based on Cisco FabricPath can be modified as needed in a nondisruptive manner for the end devices.
  - The capabilities of Cisco FabricPath troubleshooting tools surpass those of the tools currently available in the IP world. The ping and traceroute features now offered at Layer 2 can measure latency and test a particular path among the multiple equal-cost paths to a destination within the fabric.
  - Switches that do not support Cisco FabricPath can still be attached to the Cisco FabricPath fabric in a redundant way without resorting to Spanning Tree Protocol.
- Reliability based on proven technology.
  - Although Cisco FabricPath offers a plug-and-play user interface, its control protocol is built on top of the powerful IS-IS routing protocol, an industry standard that provides fast convergence and that has been proven to scale up to the largest service provider environments.
  - Loop prevention and mitigation is available in the data plane, helping ensure safe forwarding that cannot be matched by any transparent bridging technology. Cisco FabricPath frames include a TTL field similar to the one used in IP, and RPFC is also applied.
- Efficiency and high performance.
  - Because ECMP can be used at the data plane, the network can use all the links available between any two devices. The first-generation hardware supporting Cisco FabricPath can perform 16-way ECMP, which, when combined with 16-port 10-Gbps PortChannels, represents bandwidth of 2.56 terabits per second (Tbps) between switches.
  - Frames are forwarded along the shortest path to their destination, reducing the latency of the exchanges between end devices compared to a spanning-tree-based solution.

Conclusion

Cisco Unified Fabric is at the center of the fabric resiliency solutions and contains a range of technologies that provide the massive scalability and fabric resiliency observed in customer data centers. Every Cisco customer has unique fabric design requirements, and because of this, Cisco offers a number of technology options to preserve customer solution choice and flexibility. Fabric resiliency design options include PortChannel technology, which vPC extends to remove Spanning Tree Protocol as a loop management technology in large-scale Layer 2 Ethernet networks. For an alternative to vPC, Cisco FabricPath technology combines the flexibility of Layer 2 with the scalability and performance characteristics of routing and provides a solution that is simple, scalable, and efficient for traditional, virtualized, and cloud environments.

For More Information

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

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

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

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.

Printed in USA C11-721020-00 11/12