DATA CENTER

Deploying Brocade VDX 6720 Data Center Switches with Brocade VCS in Enterprise Data Centers

At the heart of Brocade VDX 6720 switches is Brocade Virtual Cluster Switching (VCS), a new Ethernet fabric technology that addresses the unique requirements of next-generation data center environments.

BROCADE

Brocade® VDX 6720 Data Center Switches, the first in the new Brocade VDX family, are 10 Gigabit Ethernet (GbE) line-rate, low-latency, lossless switches in 24-port (1U) and 60-port (2U) models. The Brocade VDX is based on multiple generations of Brocade fabric technology and runs the new Brocade Network Operating System.

At the heart of Brocade VDX 6720 switches is Brocade Virtual Cluster Switching (VCS™), a new Ethernet fabric technology that addresses the unique requirements of evolving data center environments. Brocade VCS will fundamentally change how data center networks are architected, deployed, and managed. Customers will be able to seamlessly transition their existing multi-tier hierarchical networks to a much simpler, virtualized, and converged data center network—moving at their own pace.

Brocade VCS capabilities allows customers to start with a small two-switch Top-of-Rack (ToR) deployment and expand to a large virtualization-optimized Ethernet fabrics and to manage these networks as a single logic entity. Furthermore, Brocade VCS-based Ethernet fabrics are convergence ready, allowing customers to deploy one network for storage—Fibre Channel over Ethernet (FCoE), iSCSI, NAS—and IP data traffic.

INTRODUCTION

The Brocade VDX 6720 supports these three deployment models, each of which is described in this paper:

- **Classic 10 GbE active-active access/aggregation.** Represents the classic 10 GbE server access and 1 GbE ToR aggregation deployments enhanced with a broad set of Layer 2 features.
- **Scale-out fabrics for virtualized data centers.** Enables dynamic, large-scale, server virtualization deployments in private (IT customers within an organization) and public (external customers of managed service provider) clouds with proven fabric capabilities such as Distributed Intelligence and Layer 2 Equal Cost Multipath Protocol (ECMP).
- **LAN/SAN convergence.** Converges storage and IP data traffic on a single data center network and supports end-to-end FCoE with fundamental enhancements for iSCSI and NAS.

10 GbE ACCESS/AGGREGATION ARCHITECTURE

In data centers today, a three-tiered architecture has access, aggregation, and core layers, as described below and shown in Figure 1.

Access

The access layer is the first tier or edge of the data center. This is where end devices like, servers and storage, attach to the wired or wireless portion of the network. The servers are front-end to the access and back-end to the Ethernet and Fibre Channel-based storage.

Aggregation

The aggregation layer has a unique role; it acts as a services and control boundary between the access and core. Both access and core are essentially dedicated special-purpose layers. The access layer is dedicated to meeting the functions of end-device connectivity and the core layer is dedicated to providing non-stop connectivity across the entire network. The aggregation layer, on the other hand, serves several functions. It is an aggregation point for all of the access switches; it can be scaled out and acts as an integral member of the access-distribution block providing connectivity and policy services for traffic flows within the access-distribution block. As an element in the core of the network, it participates in the core routing design.

Core

The core layer provides a very limited set of services and is designed to be highly available and operate in an always-on mode. The core layer is a gateway with high-speed connectivity to external entities, such as the WAN, intranet, and Internet. The data center core is a Layer 3 domain in which efficient and expeditious forwarding of packets is the fundamental objective. To this end, the data center core is built with high-bandwidth links (10 GbE) and employs routing best practices to optimize traffic flows. Note that the core should not implement any complex policy services, nor should it have directly-attached user or server connections. The core should have a minimal control plane configuration combined with highly available devices configured with the physical redundancy to provide for non-stop service.

Background on Three-Tiered Network Design

Early LANs were essentially large flat networks that enabled peer-to-peer communication at Layer 2 using Media Access Control (MAC) addressing and protocols. A flat network space, however, is vulnerable to broadcast storms that can disrupt all attached devices and this vulnerability increases as the population of devices on a network segment grows.

Consequently, Layer 3 routing and the IP address scheme were introduced to subdivide the network into manageable groups and to provide isolation against Layer 2 calamities. Multiple Layer 2 groups interconnected by Layer 3 routers facilitates optimal communication within and between workgroups and streamlines network management and traffic flows. Over the years, this basic layered architecture has been further codified into a commonly deployed three-tier network design.

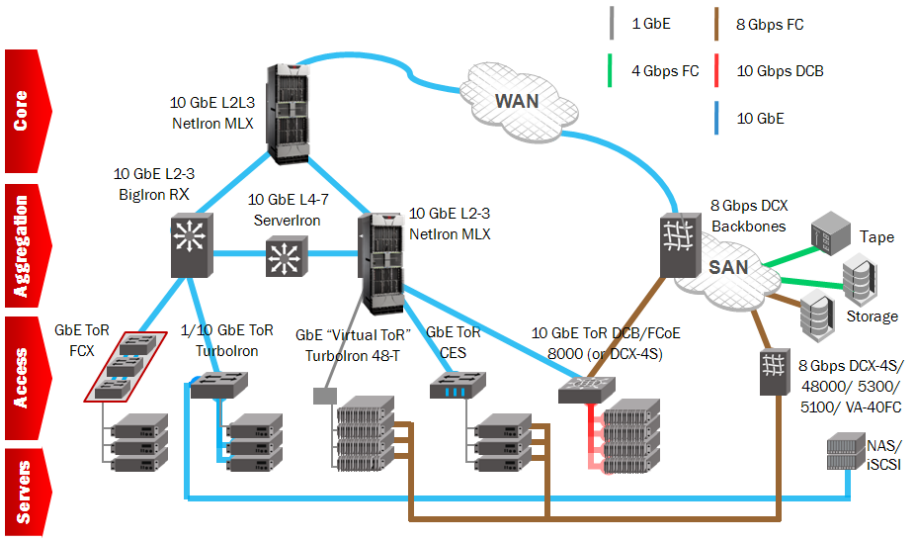

Figure 1.
A Brocade data center network topology.

DEPLOYING THE BROCADE VDX 6720 INTO CLASSIC 10 GbE ARCHITECTURES

For classic Ethernet architectures, Brocade VCS enables organizations to preserve existing network designs and cabling and to achieve active-active server connections without using Spanning Tree Protocol (STP).

In the classic 10 GbE access/aggregation architecture, ToR switches are usually connected to rack servers via a GbE or 10 GbE connection. ToR switches are deployed in pairs in an active-passive configuration that connects to Network Interface Card (NIC)-teamed rack servers. ToR switches are connected to the aggregation layer via Link Aggregation Groups (LAGs) for upstream traffic. But although this design works in many environments, it has some design limitations. The design uses STP and is inefficient, because 50 percent of the ToR ports remain idle until a failure occurs. In a traditional data center, the access layer is dependent on STP to forward and block ports to avoid loops. When a failure occurs, spanning tree needs to recalculate to allow the passive connection to begin forwarding traffic, which creates a time delay (about 30 seconds for standard STP and 2 seconds for Rapid STP).

Integrating Brocade VDX into the classic 10 GbE access/aggregation architecture requires no change to the existing enterprise data center network architecture. In fact, the Brocade VDX design utilizes a single standard LAG, consisting of multiple 10 GbE connections from a pair of Brocade VDX switches for upstream traffic. The Brocade VDX pair uses a vLAG to allow the two switches to look like a single switch to the core routers. The Brocade VDX ToR design locates two Brocade VDX switches per rack, which supports a number of rack servers (dependent on the subscription ratio used). Each rack server has two 10 GbE connections configured as a LAG and operating in an active-active multi-homed mode. A new IETF protocol, Transparent Interconnection of Lots of Links (TRILL), provides active multipath support, so that the rack server sees only a single ToR switch, which provides the rack server with 20 Gigabits per second (Gbps) or more of throughput.

An example of this use case is shown in Figure 2, in which the pair of Brocade VDX ToR switches support a 4:1 subscription ratio for 10 Gbps to the aggregation for the rack server design. This means that 72 ports (36 ports per switch) are available to the rack servers and 20 ports (10 ports per switch) to the aggregation via a single LAG.

vLAGs

A vLAG is a fabric service that allows LAGs to originate from multiple Brocade VDX switches. It acts the same way as a standard LAG using the Link Aggregation Control Protocol (LACP), a method to control the bundling of several physical ports together to form a single logical channel.

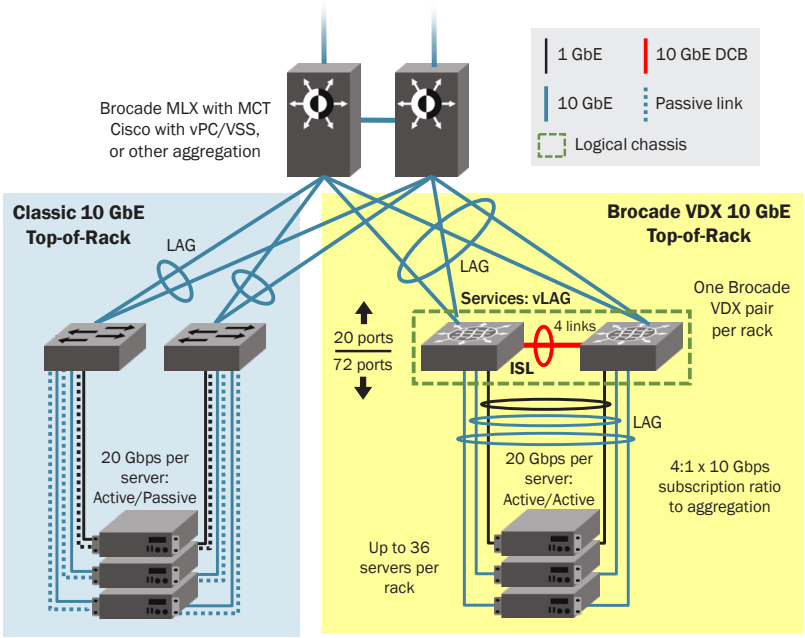

Figure 2.
Classic 10 GbE access/aggregation rack server architecture with Brocade VDX 6720 switches.

When Brocade VDX ToR switches in VCS mode are physically linked together:

- If present, multiple links auto-configure to form an Inter-Switch Link (ISL).
- The Brocade VDX then uses the VCS control plane protocol to create the network topology database, which in turn computes equal-cost shortest paths.
- Since the Brocade VDX switches understand each other's control plane, an Ethernet fabric forms.

In a Blade Server Environment

This design can also be implemented with blade servers with many of the same capacities described in the previous example. In comparison, the following differences apply to blade servers:

- 4 blade server chassis supported per rack for a total of 64 servers:
- 8 links required to support the ISL
- 8 links per blade switch
- 2:1 subscription ratio through the VCS switches

This means that 64 ports (32 ports per switch) are available to the blade servers and 32 ports (16 ports per switch) to the aggregation via a single LAG, as shown in Figure 3.

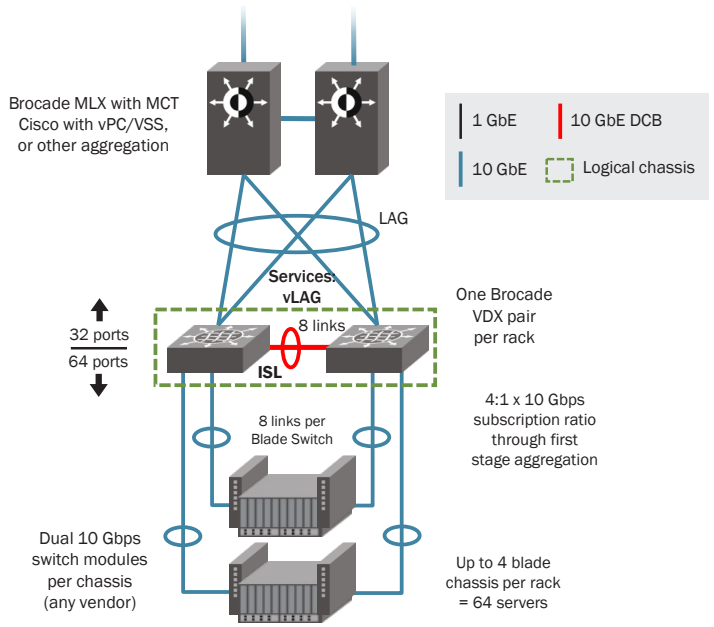

Figure 3.
Classic 10 GbE access/aggregation blade server architecture with VCS.

These ToR (access) solutions use Brocade VCS, which can also be used as a scalable aggregation layer. The Brocade VCS aggregation layer has the ability to deploy additional Brocade VDX switches without network disruption, while providing redundancy throughout the network. It uses a single LAG (multiple active 10 GbE connections) to provide connectivity to the core. Note that Multi-Chassis Trunking (MCT), vPC/VSS, or another type of core virtualization design are also used.

MULTI-CHASSIS TRUNKING (MCT)

Multi-Chassis Trunking, currently available in Brocade MLX routers, is a Brocade technology that allows multiple switches to act as single logical switch connected to another switch using a standard LAG. MCT at the core with vLAG in the edge fabric enables a single LAG between the two logical elements—which results in a fully active-active network.

In a Stacked Configuration

The next example (see Figure 4) shows a three-switch stack configuration at the Top of Rack and two 10 GbE links per switch for a total of 6 links to the Brocade VCS aggregation layer via a single LAG. This configuration provides 144 ports for server connectivity. The VCS aggregation layer can be designed to provide 1:1 wire-speed through the logical chassis resulting in 192 usable ports. A 4:1 subscription ratio through the aggregation layer provides 154 ports to the access layer and 38 ports to the core. Based on the three-switch stack configuration, the 4:1 subscription supports a total of 900 servers, or 25 racks of servers.

This building block design enables you to "pay as you grow." To increase port count, simply add a Brocade VDX switch non-disruptively into the fabric. Since the Brocade VCS fabric looks and acts like a single logical entity, minimal management is required moving forward.

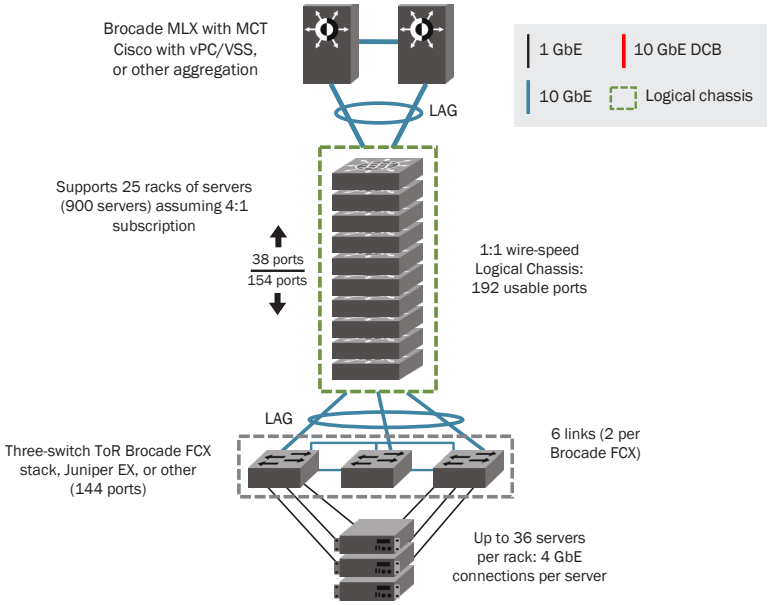

Figure 4.
10 GbE Brocade VCS aggregation layer architecture.

Advantage of VCS in the Access/Aggregation Layers

A pair of interconnected Brocade VDX switches is deployed into the 10 GbE classic access/aggregation at the top of each new server rack, alongside the existing ToR infrastructure. This provides significant investment protection as the existing architecture, using a three-tier architecture, can be combined with the new capabilities of the Brocade VDX ToR solution. In the aggregation layer, the Brocade VDX switches serve as cost-effective 10 GbE End-of-Rack (EoR) or Middle of Rack (MoR) aggregation for the existing GbE server ToR switches. You can also use existing cabling and ToR switches. Additional Brocade VDX switches can be added to support ongoing growth without disruption of network traffic.

The Brocade VDX access/aggregation model provides the following benefits over the classic 10 GbE access/aggregation model:

- High density
- 10 GbE ToR server access with active/active multi-homed server connectivity (for the VCS access design) redundancy
- Simplified management at the rack level
- Investment protection, no "rip and replace," integrate into a classic three-tier data center architecture
- Elimination of STP in the VCS design

In this deployment, Brocade VCS provides the clustering capability of small, discrete numbers of switch elements to scale much more cost effectively, of logically larger aggregation switch. Note that these use cases do not take advantage of other capabilities of Brocade VCS, such as virtualization and convergence, shown in upcoming use cases.

COLLAPSING THE NETWORK WITH A SCALED FABRIC DESIGN

For scale-out fabric architectures, Brocade VCS allows organizations to flatten the network design, provide Virtual Machine (VM) mobility without network reconfiguration, and manage the entire fabric as a single logical chassis.

In the previous example, the Brocade VDX was added into the access or aggregation layer of an existing network. In this second use case, the access and aggregation layers are collapsed onto a single layer, essentially flatting the network. How is this done? Fabrics are self-aggregating, eliminating the need for a separate aggregation layer.

Brocade VCS protocols remove the need for STP and allow for all equal-cost paths to be active, which results in no single point of failure. This architecture is designed to support virtual environments and functions such as VM migration. It also allows for growth in the access/aggregation layer accomplished by adding a switch to the VCS fabric. The Automatic Fabric Configuration capability allows the Brocade VDX to automatically obtain its configuration and share information about the fabric. Although composed of many switches, the Brocade VCS fabric looks and acts like a single logical entity to the network.

Since VCS is a flexible topology, it's ideal for virtual environments that support VM migration, during which port profiles must be moved with the VM to ensure policies and configurations are maintained. Brocade Automatic Migration of Port Profiles (AMPP) provides this capability in a VCS design. In the VCS topology, AMPP enables a seamless migration: using the Command-Line Interface (CLI), port profiles and MAC address mapping are created on any switch in the fabric. The mapping provides the logical flow for traffic from the source port to the destination port. As a VM migration finishes, the destination port in the fabric learns of the MAC move and automatically activates the port profile configuration.

The fabric can designed using a Clos fabric (two-stage architecture), shown in Figure 5, which allows for growth as the network grows. The first stage does not connect to the edge devices but connects to the second stage, which provides the connectivity to the edge devices. This design allows for multiple VCS fabrics to connect to the same first stage.

CLOS FABRIC ARCHITECTURE

Clos is one of the most common network topologies used to build multistage switching systems that can be scaled to virtually any level.

Clos networks have three stages: the ingress stage, middle stage, and egress stage. Each stage is made up of a number of switches. Each call entering an ingress switch can be routed through any of the available middle-stage switches to the relevant egress switch. A middle-stage switch is available for a particular new call if both the link connecting the ingress switch to the middle-stage switch and the link connecting the middle-stage switch to the egress switch are free.

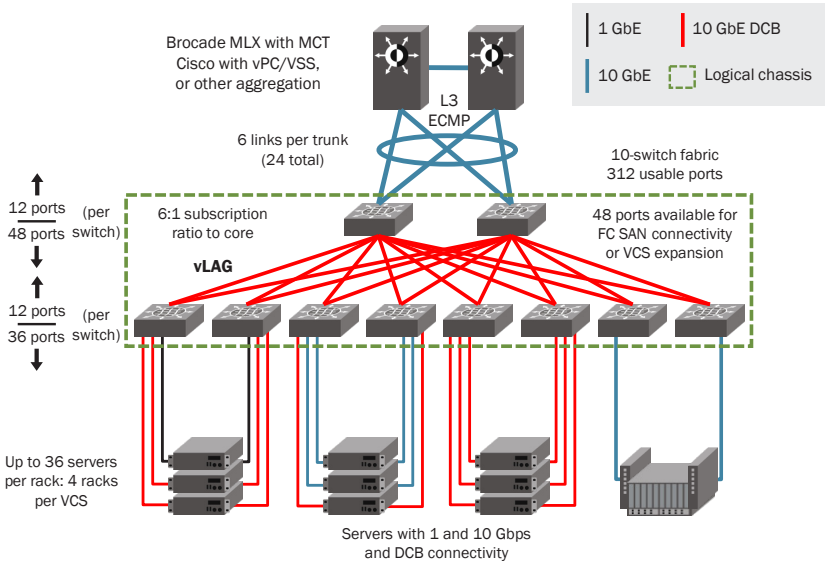

Figure 5.
Brocade VCS collapsed network using Clos fabric architecture.

Advantages of Brocade VCS in a Collapsed Network Architecture

These environments are characterized by scaled-out, dense server installations running tens of virtual workloads per physical machine. This design transforms the traditional three-tier approach, in which each switch and tier must be managed individually and differently; the entire fabric can be managed as a single logical switch.

Another benefit is the scalability of the Brocade VCS design. Growing the traditional data center was constrained to the physical limitations of the equipment itself in supporting more ports. For additional ports, the network had to be redesigned to support another switch or to replace a current switch with a denser switch. In the flat Layer 2 Brocade VCS design, increasing the number of ports simply means adding another switch to scale out the fabric without reconfiguring or removing any other equipment. Unlike switch stacking technologies, this Ethernet fabric is masterless, which means that no single switch stores configuration information or controls fabric operations. Any switch can be added, removed, or fail without causing disruptive fabric downtime or delayed traffic while a new master switch is selected.

Brocade VDX platforms in this deployment form the true, flat, scaled-out Layer 2 Ethernet fabric using one of several physical topologies (such as Clos, mesh, star, or ring) with AMPP—provided by the fully distributed intelligence capability of Brocade VCS. Distributed intelligence also supports a more virtualized access layer. Instead of distributed software switch functionality residing in the virtualization hypervisor, access layer switching is performed in switch hardware. This approach improves performance, helps ensure consistent and correct security policies, and simplifies network operations and management. AMPP supports VM migration to another physical server, ensuring that the source and destination network ports have the same configuration for VMs.

CONVERGING THE LAN AND SAN

For data centers that can take advantage of the converged LAN/SAN environment, Brocade VCS provides end-to-end Data Center Bridging (DCB) capabilities, enabling traditional IP and storage traffic to exist on the same network. Organizations can use Brocade VCS to prioritize FCoE and iSCSI traffic to make sure it receives sufficient bandwidth and remains lossless

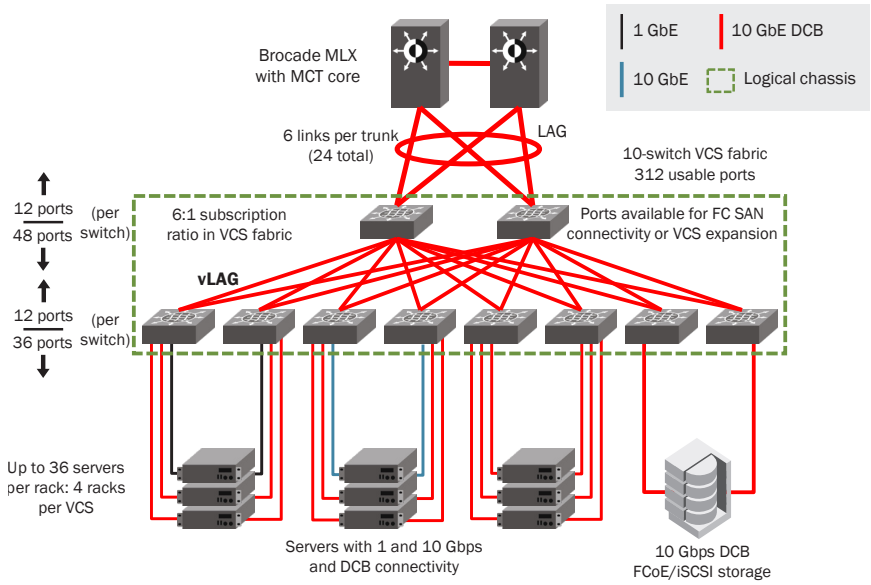

Figure 6.
Brocade VCS converged network.

Previous examples focused on the Ethernet side of the network. In the Brocade VCS fully converged network, use cases include Ethernet and storage in the same network, using the collapsed fabric design (shown in Figure 5) with the addition of storage.

Advantage of VCS Converged Network Architecture

This converged design allows for a truly lossless communication within the fabric and uses the IEEE Data Center Bridging (DCB) protocol to provide Priority-based Flow Control (PFC) and Enhanced Transmission Selection for FCoE traffic. The architecture, shown in Figure 8, provides shared storage access and connectivity for servers so that end-to-end FC and iSCSI applications can be run over a high-performance, multi-pathing, reliable, resilient, and lossless converged fabric. As a result, storage traffic is insulated and 10 Gbps connections maximized.

SUMMARY

As IT organizations look for better ways to support virtualized data centers, they are turning to high-performance networking solutions that increase flexibility through leading-edge technologies. Brocade VDX 6720 Data Center Switches re specifically designed to improve network utilization, maximize application availability, increase scalability, and dramatically simplify network architecture in virtualized data centers. By leveraging Brocade VCS technology, the Brocade VDX builds data center Ethernet fabrics—revolutionizing the design of Layer 2 networks and providing an intelligent foundation for cloud computing.

ABOUT BROCADE

Brocade leads the industry in providing comprehensive network solutions that help the world's leading organizations transition smoothly to a virtualized world where applications and information reside anywhere. As a result, Brocade facilitates strategic business objectives such as consolidation, network convergence, virtualization, and cloud computing. Today, Brocade solutions are used in over 90 percent of Global 1000 data centers as well as in enterprise LANs and the largest service provider networks.

To find out more about Brocade products and solutions, visit www.brocade.com.

Corporate Headquarters
San Jose, CA USA
T: +1-408-333-8000
info@brocade.com

European Headquarters
Geneva, Switzerland
T: +41-22-799-56-40
emea-info@brocade.co

Asia Pacific Headquarters
Singapore
T: +65-6538-4700
apac-info@brocade.com

© 2011 Brocade Communications Systems, Inc. All Rights Reserved. 07/11 GA-WP-1535-01

Brocade, the B-wing symbol, BigIron, DCFM, DCX, Fabric OS, FastIron, IronView, NetIron, SAN Health, ServerIron, TurboIron, and Wingspan are registered trademarks, and Brocade Assurance, Brocade NET Health, Brocade One, Extraordinary Networks, MyBrocade, VCS, and VDX are trademarks of Brocade Communications Systems, Inc., in the United States and/or in other countries. Other brands, products, or service names mentioned are or may be trademarks or service marks of their respective owners.

Notice: This document is for informational purposes only and does not set forth any warranty, expressed or implied, concerning any equipment, equipment feature, or service offered or to be offered by Brocade. Brocade reserves the right to make changes to this document at any time, without notice, and assumes no responsibility for its use. This informational document describes features that may not be currently available. Contact a Brocade sales office for information on feature and product availability. Export of technical data contained in this document may require an export license from the United States government.

BROCADE