



ETHERNET FABRIC SOLUTIONS FOR HPC

Micah Staggs - Sr. Systems Engineer

November 09, 2011

Legal Disclaimer

All or some of the products detailed in this presentation may still be under development and certain specifications, including but not limited to, release dates, prices, and product features, may change. The products may not function as intended and a production version of the products may never be released. Even if a production version is released, it may be materially different from the pre-release version discussed in this presentation.

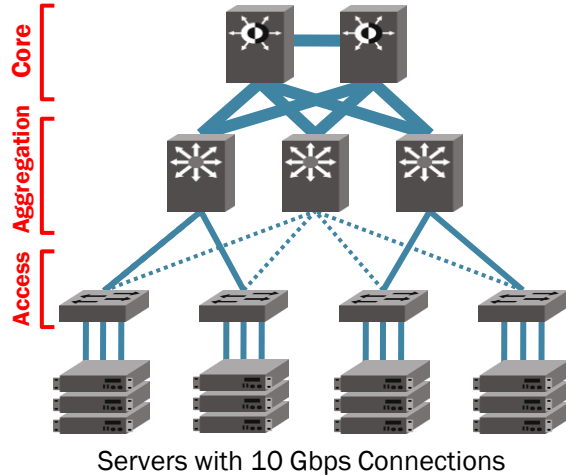
NOTHING IN THIS PRESENTATION SHALL BE DEEMED TO CREATE A WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, STATUTORY OR OTHERWISE, INCLUDING BUT NOT LIMITED TO, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NONINFRINGEMENT OF THIRD-PARTY RIGHTS WITH RESPECT TO ANY PRODUCTS AND SERVICES REFERENCED HEREIN.

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.

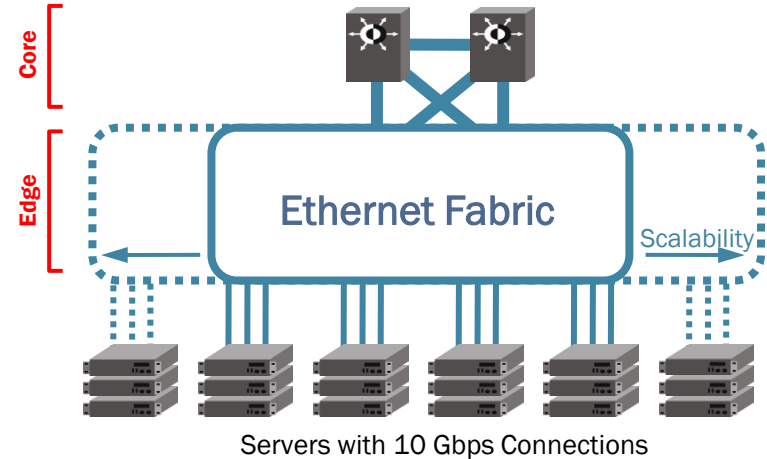


Ethernet Fabrics

Classic Hierarchical Ethernet Architecture



Ethernet Fabric Architecture



Ethernet Fabric Properties

- Lossless, deterministic
- All paths are available, traffic travels along the shortest path
- Self-aggregating, with distributed intelligence
- Managed and functions as a single logical chassis
- LAN-SAN convergence ready

Technologies

- DCB
- TRILL
- Brocade VCS
- Brocade VCS
- FCoE and DCB



Virtual Cluster Switching (VCS)

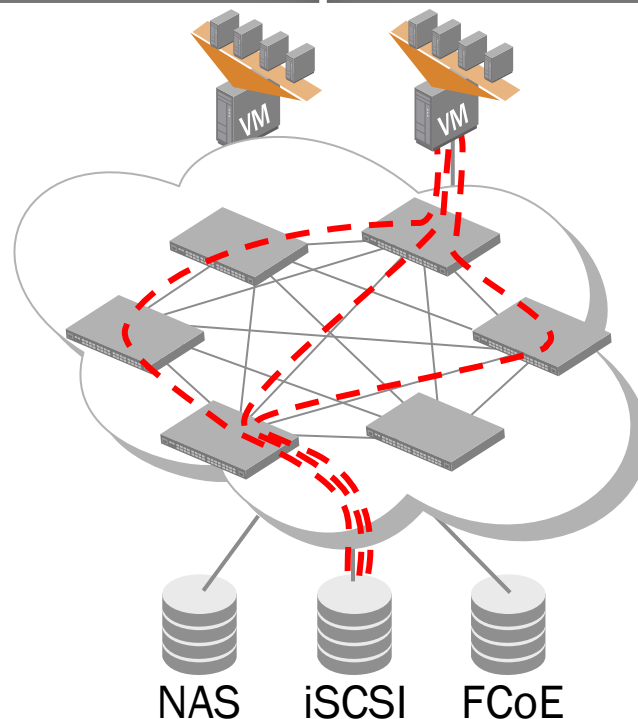
ETHERNET
FABRIC

DISTRIBUTED
INTELLIGENCE

LOGICAL
CHASSIS

DYNAMIC SERVICE
INSERTION

- First data center Ethernet fabric
- No Spanning Tree Protocol
- Multi-path, deterministic
- Auto-healing, non-disruptive
- Lossless, low latency
- Built for convergence



CORE TECHNOLOGY

Virtual Cluster Switching (VCS)

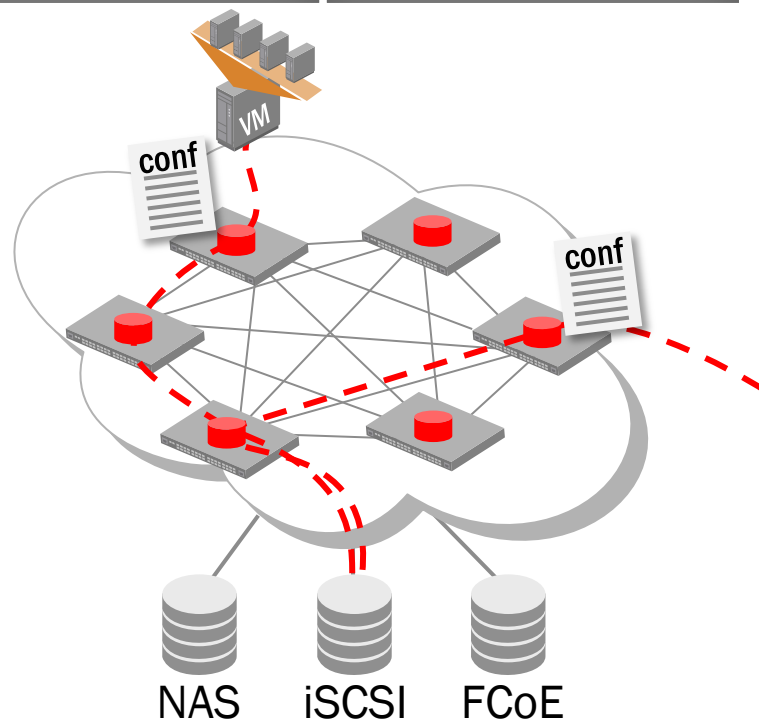
ETHERNET
FABRIC

DISTRIBUTED
INTELLIGENCE

LOGICAL
CHASSIS

DYNAMIC SERVICE
INSERTION

- Fully distributed control plane
- Arbitrary topology, self-forming
- Network-wide knowledge of all members, devices, VMs
- Automatic Migration of Port Profiles (AMPP)



CORE TECHNOLOGY

Virtual Cluster Switching (VCS)

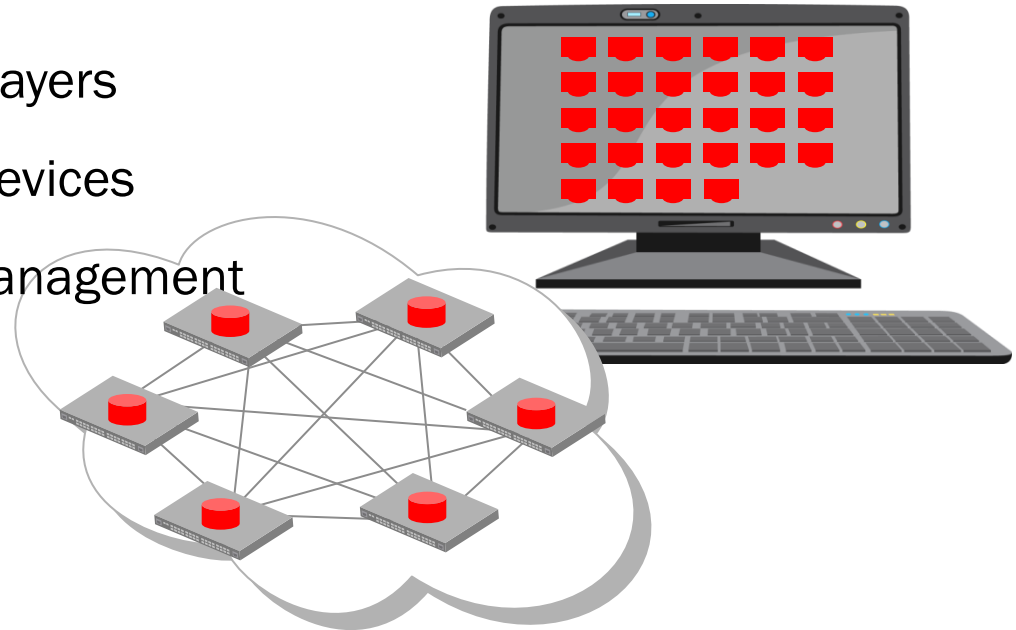
ETHERNET
FABRIC

DISTRIBUTED
INTELLIGENCE

LOGICAL
CHASSIS

DYNAMIC SERVICE
INSERTION

- Managed as a single switch
- Logically collapses network layers
- Auto-configuration for new devices
- Centralized or distributed management
- Radically reduces managed elements



CORE TECHNOLOGY

Virtual Cluster Switching (VCS)

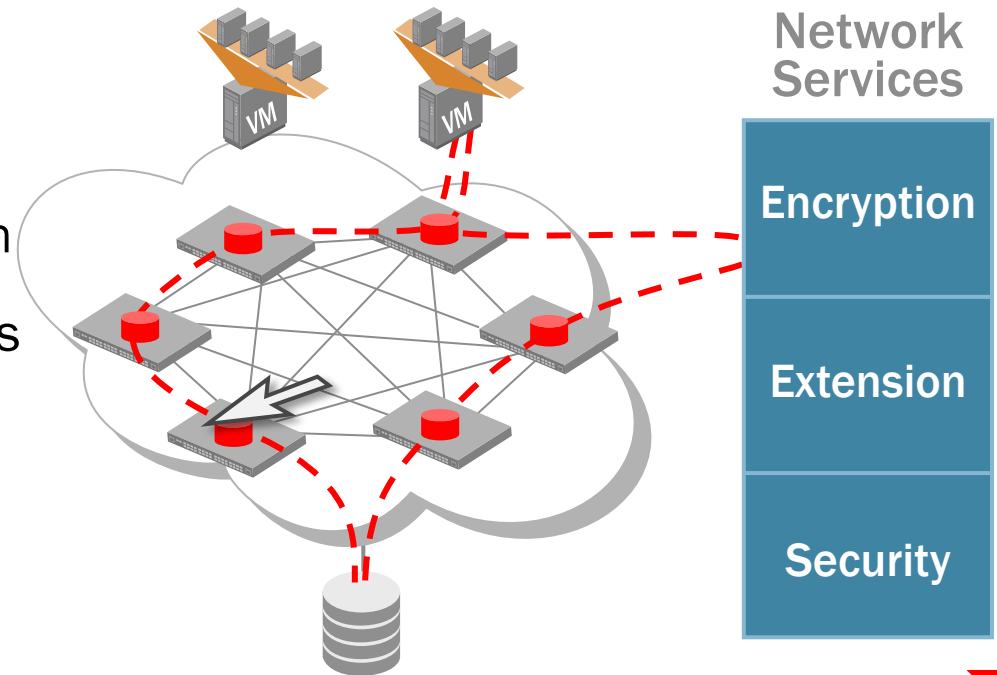
ETHERNET
FABRIC

DISTRIBUTED
INTELLIGENCE

LOGICAL
CHASSIS

DYNAMIC SERVICE
INSERTION

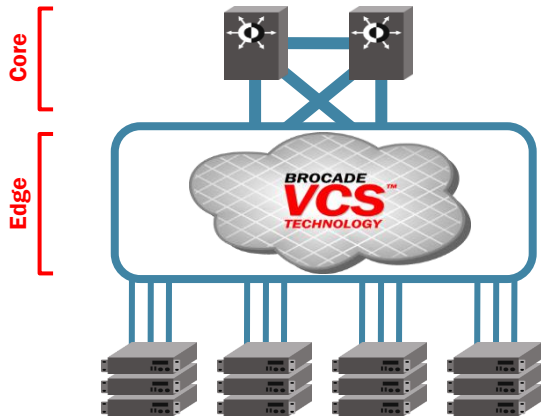
- Reconfigure network via software
- Hardware-based flow redirection
- Incorporation of partner services
- Non-stop service insertion
- Minimizes cost and physical moves



Ethernet Fabrics for HPC

Very low node-to-node latency, in a flat lossless Layer 2 network, without STP

Ethernet Fabric Architecture



- Significantly lower network latency over classical Ethernet with STP
 - Max 2 or 3 low-latency switches to cross in a well designed L2 fabric
 - No need to go through aggregation/core in larger networks
- The network is much more responsive and resilient in the presence of errors
- Only one device to manage (logical chassis)

Fabric designs can be tailored to maximize application performance

- Ethernet Fabrics are optimal for interconnect networks
- Knowing the application communication pattern – inter-node and I/O – the network architecture can be tailored to maximize performance:

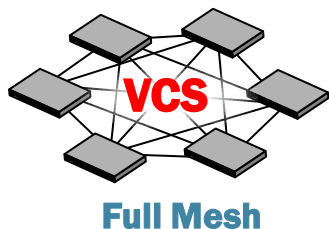
$$\text{network_design} = F (\# \text{devices, connectivity, oversubscription, target_latency})$$

where:

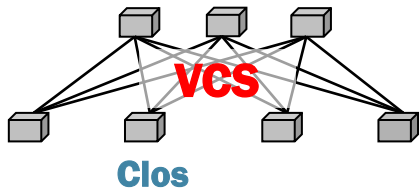
#devices	= total servers and storage devices in the cluster
connectivity	= type of server/storage connectivity, 1 or 10GbE
oversubscription	= desired/tolerated network oversubscription level
target latency	= desired latency level to servers and to storage



Example: large 10GbE non-oversubscribed fabric

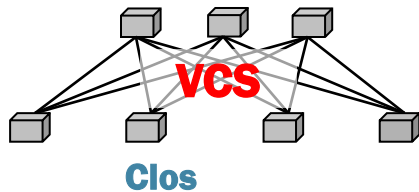


- up to 156 10GbE servers – use full mesh fabric with the 24 ports switch; 13 switches needed; max 2 switches to cross
- up to 930 10GbE servers – use full mesh fabric with the 60 ports switch; 31 switches needed; max 2 switches to cross



- up to 288 10GbE servers – use Clos fabrics with the 24 port switch; 36 switches needed; max 3 switches to cross
- up to 1800 10GbE servers – use Clos fabric with the 60 ports switch; 90 switches needed; max 3 switches to cross

If some oversubscription is acceptable, single fabric sizes can grow significantly



24-port switch in a Clos fabric

- 1:1 – 288 x 10 GbE servers with 36 switches/fabric
- 2:1 – 384 x 10 GbE servers with 32 switches/fabric
- 4:1 – 456 x 10 GbE servers with 29 switches/fabric

60-port switch in a Clos fabric

- 1:1 – 1,800 x 10 GbE servers with 90 switches/fabric
- 2:1 – 2,400 x 10 GbE servers with 80 switches/fabric
- 4:1 – 2,880 x 10 GbE servers with 72 switches/fabric

Brocade VCS Technology and VDX Switches

Low Latency Fabric infrastructure

Scalability and Resiliency without compromising Latency

Proven Technology

Support



Thank You



Brocade NetIron MLXe Series Routers

Product highlights



Leading performance and scalability

- Up to 15.36 Tbps forwarding capacity in a chassis
- Up to 1536 1 GbE and 256 10 GbE ports
- Wire-speed, non-blocking performance

Optimum flexibility

- 4-, 8-, 16-, and 32-slot chassis
- 20-, 48-port copper, 20-port fiber 1 GbE modules
- 4- and 8-port 10 GbE modules
- 4- and 8-port OC48 and 2-port OC192 modules
- 40/100 GbE-ready
- Field-programmable network processors

High availability

- Management, fabric, and M+N power redundancy
- Hitless failover and upgrade; graceful restart

Advanced, scalable software features

- Advanced virtualization with multi-VRF
- Multiservice feature set (IPv4, IPv6, MPLS, QoS)
- sFlow for granular network traffic accounting

Operational efficiency

- Best-in-class power efficiency
- Front-to-back airflow



100 GbE MLX Module

Product Highlights

Industry's first **2 port** 100 GbE module

Massive 100 GbE density of **32 wire-speed** ports

Multiple **full** 100 Gigabit packet processors.

Terabit trunks with 1.6 Tb/s per trunk.

Ports on Demand enabling pay as you grow strategy

Classic XMR and MLX chassis support 1-port 100 GbE

Full featured card with Advanced MPLS and IPv4/IPv6 capabilities

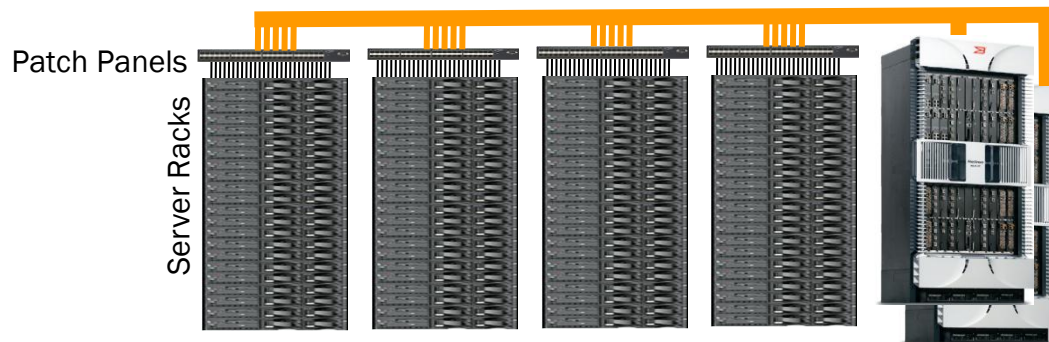
1 million IPv4 and 240 K IPv6 FIB capacity

802.3ba compliant and supports CFP based optics



Brocade NetIron MLXe Virtual ToR Solution

Simplified and Resilient Architecture for Virtual Server Connectivity



- Servers connect to MRJ21 passive patch panels within server racks
 - Up to eight MRJ21 (six 1 GbE links) cables per panel for 48 1 GbE ports
 - No active switches in server racks to manage or configure
- Middle- or end-of-row high-density server connectivity
 - Up to 32 NetIron MLX-48T modules per chassis; 1536 line-rate 1 GbE ports
 - Can be configured with multi-chassis trunking for active-active links
 - Supports up to one million MAC addresses for the largest VM deployments



NetIron MLX-48T 48 1 GbE Module with six MRJ21 Ports



Passive 48-port Patch Panel



MRJ21 Cable

NetIron CES/CER Series

Full-featured, Compact Edge Switch



- **Carrier-class performance**
 - Wire-speed performance: 48 to 136 Gbps
 - 128,000 MAC addresses, 32,000 IPv4 routes
- **Optimum flexibility**
 - 24- and 48-port copper and fiber models
 - All models are field-upgradable to 2-port 10 GbE
- **High availability**
 - NEBS Level 3 certified
 - Hot-swappable, redundant, load-sharing power supplies with AC and DC options
 - N+1 redundant, replaceable cooling system
- **Designed for high-bandwidth, media-rich applications**
 - Deep buffering up to 192 MB
 - 4G Cell backhaul, metro-E business services, IPTV, content networks
- **Advanced software features**
 - MEF 9 and MEF 14 certified
 - Multiservice feature set (IPv4, IPv6, MPLS, QoS)
 - MPLS VPLS and VLLs
 - Comprehensive OAM capabilities
 - E-LINE, E-LAN, and E-TREE services
 - New Ethernet Service Instance (ESI) framework
 - L2PT, MRP, MRP2, PB and PBB



Brocade TurboIron 24X Switch

Product Highlights



- **Data center class performance:**
 - DC modularization with 10 GbE ToR solution
 - Up to 488 Gbps line rate, non-blocking performance
 - Up to 24 10 GbE SFP+ ports in 1RU
 - Cut-through architecture with ultra-low latency
 - Front-to-back cooling
- **Greatest flexibility:**
 - 1GbE/10GbE ports for seamless migration
 - 4 10/100/1000 Mbps copper ports
 - Redundant, removable, and load-sharing AC power supplies
 - Hot-swappable triple-fan tray
- **Advanced software features:**
 - Advanced Layer 2 & Layer 3 features
 - Robust security suite
 - sFlow for detailed network traffic accounting
- **Operational efficiency:**
 - Best in class power efficiency
 - Rack space saving



Brocade FCX 624 and 648

Product highlights



Purpose-built data center design

- 24 or 48 10/100/1000 Mbps ports in a 1U form factor
- Optional 4 ports of 10 GbE for high-speed uplinks
- Optional 4 100/1000 Mbps fiber uplinks
- Efficient SFP+ 10 GbE optics
- 112 W max power draw

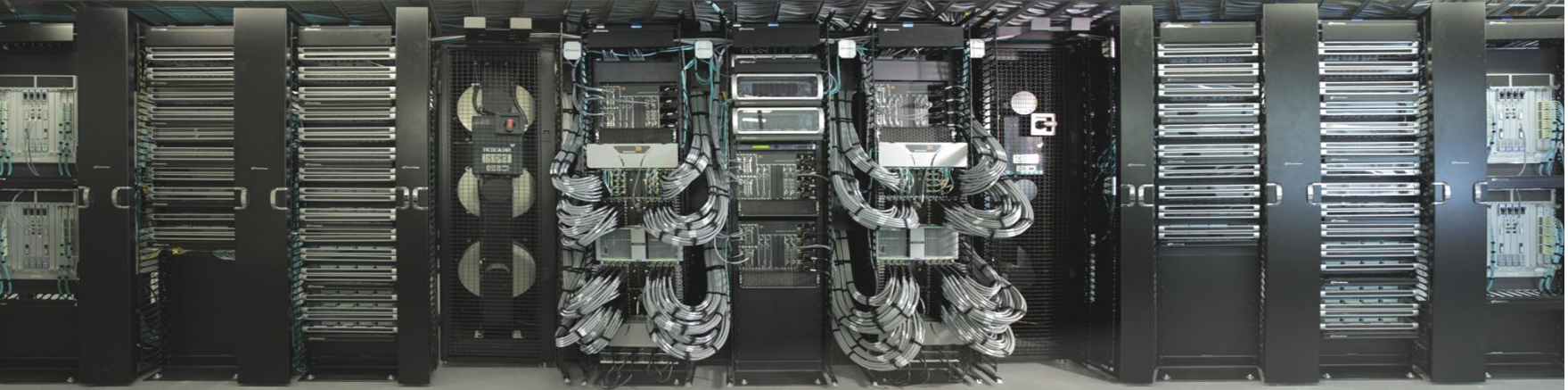
Optimal flexibility and reliability

- Horizontal stacking to create one logical switch per row
- Hot unit insertion/removal from the stack
- Cross-unit link aggregation from the stack
- Single-IP stack management
- Dual removable and load-sharing power supplies
- Reversible front-to-back airflow
- Field replaceable fan unit with 4 redundant blowers

Advanced software features

- Robust and secure Brocade IronWare OS
- sFlow for detailed network traffic accounting
- Enterprise-class Layer 3 routing
- Optional BGP routing

Brocade's Next Generation Product (VCS)



VDX6720 Product Overview

Enhanced Access/Aggregation Layer Switching

- **Best-In-Class Performance and Density**
 - VDX6720-24(24 SFP+ ports, 1RU) and VDX6720-60(60 SFP+ ports, 2RU) DCB ports
 - Non-blocking, cut-through, wire-speed
 - Extremely low latency (~600nsec-1.8usec)
- **Optimized for the Virtualized Data Centers**
 - Front-to-rear and Rear-to-Front airflow
 - Best-in-class energy efficiency (~5–6W per port)
 - Virtual Cluster Switching (VCS) technology
- **Maximum Flexibility**
 - Flexible topologies with VCS enable simplistic design
 - 10 GbE and 1 GbE supported on every port
 - Supports Twinax, Active Optical
- **Enables Network and I/O Convergence**
 - Support for DCB, End to End FCoE
- **Highly Resilient and Efficient Design**
 - Runs new modular Brocade Network OS



VDX 6720