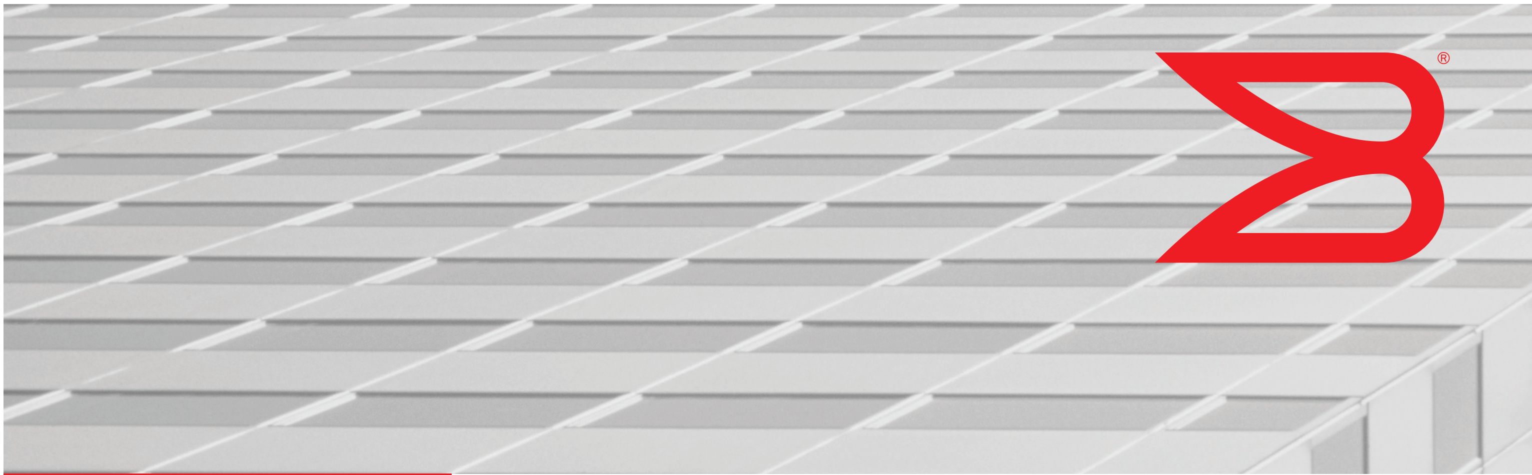


DATA CENTER

# What Is an Ethernet Fabric?

Compared to classic hierarchical Ethernet architectures, Ethernet fabrics provide the higher levels of performance, utilization, availability and simplicity required to meet the business needs of data centers today and into the future.

BROCADE

**Data center networks rely on Ethernet. Over the decades, Ethernet has evolved as new types of application architectures emerged. Today, data center networks carry traffic for a diverse set of applications including client/server, Web services, unified communications, virtual machines, and storage—each with different traffic patterns and network service requirements. Applications are increasingly deployed within virtual machines hosted on server clusters. And Ethernet can be used to build shared storage pools, which place stringent demands on the network including lossless packet delivery, deterministic latency and high bandwidth. Combined, these changes drive the next evolution in Ethernet networks: the Ethernet fabric.**

## EXECUTIVE SUMMARY


Data centers continue to grow as digital assets increase and more applications are deployed. Businesses expect agile application deployment—in minutes, not months—to keep their competitive edge as markets and competitors become global in scale. And data center resources such as rack space, power, and cooling are growing more scarce and costly. When deployed together effectively, high-density, multi-core servers, network, server, and storage virtualization, and IT orchestration tools can be used to pool IT resources and implement cloud architectures. Moving towards cloud computing can reduce capital and operational expenditures by driving the consolidation of applications and improving resource utilization, and at the same time create an infrastructure that rapidly scales and responds to business needs.

The capital cost savings of server virtualization have generally met the business mandate to "do more with less." However, the underlying limitations in existing virtualization and system management tools as well as current network architectures have often prevented organizations from meeting the performance, availability, security, and mobility requirements of cloud computing. System management and orchestration tools tend to be extremely complex and require a high degree of customization, making them expensive to deploy and often difficult to use effectively. In order to simplify the management layer and bring the promise of virtualization to fruition, the underlying infrastructure—especially the network—must evolve. It must move from management of physical ports to flows (virtual server to virtual server or virtual server to virtual storage communication). It must be simpler to set up, operate and scale, more flexible, highly resilient, and much more VM-aware. As Gartner explains, "these changes will evolve the network from the traditional tiered-tree topology to a flat-meshed Layer 2 network topology architecture.[1]" In other words, an Ethernet fabric.

This paper reviews classic Ethernet architecture in light of new data center requirements, provides an overview of the differences between traditional Ethernet and Ethernet fabrics, and explains how Ethernet fabrics can be used to address emerging data center challenges.


1 "Eight Key Impacts on Your Data Center LAN Network" (Munch, 4/27/11, ID G00211994)

## THE CLASSIC ETHERNET NETWORK

In order to better understand the Ethernet fabric, consider first a classic Ethernet network. Most data centers need more ports than are available in a single Ethernet switch, so multiple switches are connected to form a network with increased connectivity. For example, server racks often include a switch at the Top of the Rack (ToR) or servers in several racks connect to a Middle of Row (MoR) or End of Row (EoR) switch. All these Ethernet switches are connected, forming a hierarchical, or "Ethernet tree" topology, as shown in Figure 1.

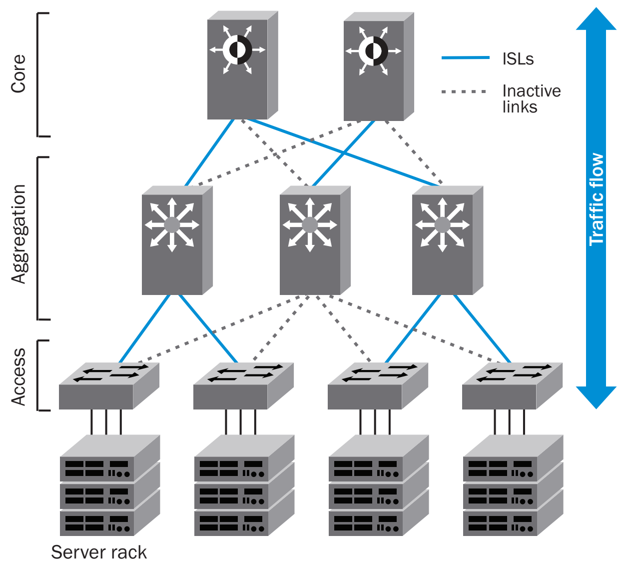



**Figure 1.**
Classic Ethernet network.

## LIMITATIONS OF TRADITIONAL ETHERNET

In a classic Ethernet network, the connections between switches, or Inter-Switch Links (ISLs, shown as solid blue lines in Figure 1), are not allowed to form a loop or frames aren't delivered. Spanning Tree Protocol (STP) prevents loops by creating a tree topology with only one active path between any two switches. (In Figure 1, inactive paths are shown as dotted lines.) This means that ISL bandwidth is limited to a single logical connection, since multiple connections between switches are prohibited. Enhancements to Ethernet tried to overcome this limitation. Link Aggregation Groups (LAGs) were defined so that multiple links between switches were treated as a single connection without forming loops. But, a LAG must be manually configured on each port in the LAG and is not very flexible.

The tree topology requires traffic to move up and down the tree, or "north-south," to get to an adjacent rack. When most of the access traffic is between servers in a rack, this is not a problem. But server clusters, such as those required for clustered applications and server virtualization, have traffic between servers in multiple racks, travelling "east-west," so the tree topology increases latency with multiple hops and restricts bandwidth with single links between switches.

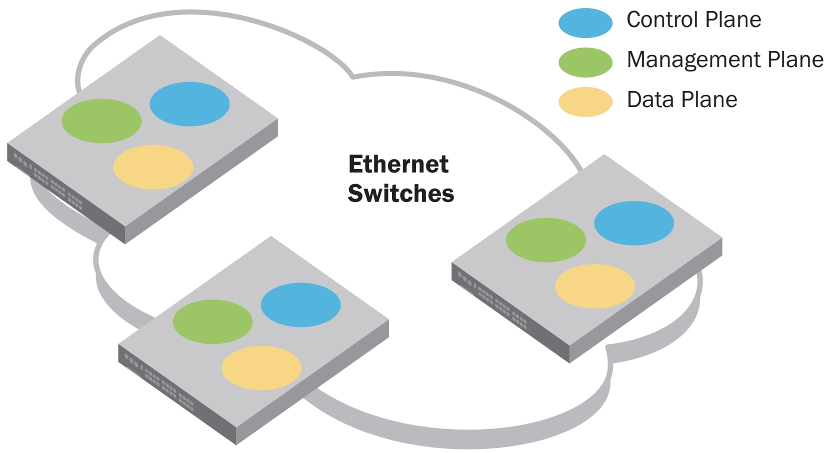
STP automatically recovers when a link is lost. However, it halts all traffic through the network and must reconfigure the single path between all switches in the network before allowing traffic to flow again. Halting all traffic for tens of seconds up to minutes on all links limits scalability and constrains traffic to applications that can tolerate data path blocking to achieve link resiliency. In the past, traffic relied on TCP to handle this interruption in service, but today, with almost all data center applications running in a 24 x 7 high availability mode and storage traffic growing on the Ethernet network, loss of connectivity in the data path for even a few seconds is unacceptable.

Finally, the classic Ethernet switch architecture presents other limitations. Each switch has its own control and management planes. Each switch has to discover and process the protocol of each frame as it arrives on an ingress port. As more switches are added, protocol processing time increases adding latency. Each switch and each port in the switch has to be configured individually, since there is no sharing of common configuration and policy information between switches. Complexity increases, configuration mistakes increase, and operations and management resources do not scale well.

## THE ETHERNET FABRIC ARCHITECTURE

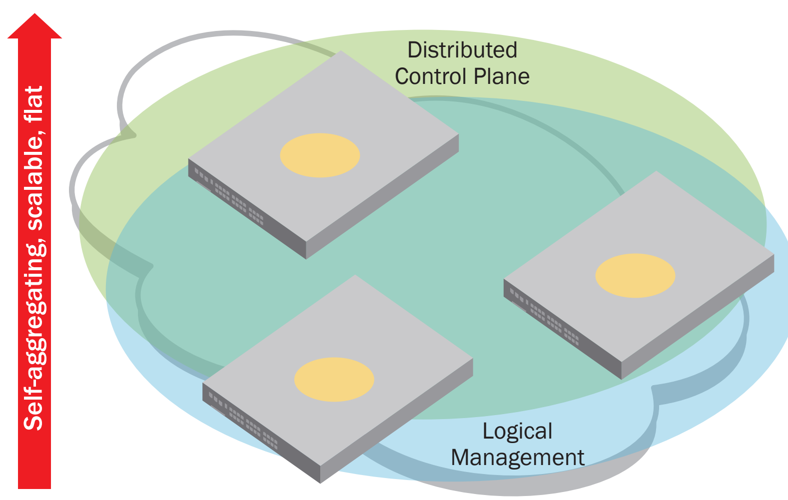
Figure 2 shows the architecture of a classic Ethernet switch. The control, data, and management planes are logically connected to every port via a back plane. Control and management planes operate at the switch level not a network level.

**Figure 2.**
Ethernet switch architecture.




Ethernet fabrics can be thought of as extending the control and management planes beyond the physical switch into the fabric. As shown in Figure 3, they now operate at a fabric level rather than at a switch level.

**Figure 3.**
Ethernet fabric architecture.



In an Ethernet fabric, the control path replaces STP with link state routing, while the data path provides equal-cost multipath forwarding at Layer 2 so data always takes the shortest path using multiple ISL connections without loops. Combined with the fabric's control plane, scaling bandwidth is made simple. For example, it becomes possible to automate the formation of a new trunk when a new switch connects to any other switch in the fabric. If a trunk link fails or is removed, traffic is rebalanced on the existing links non-disruptively. Finally, if an ISL is added or removed anywhere in the fabric, traffic on other ISLs continues to flow instead of halting as with STP.

With this architecture in place, a group of switches can be defined as part of a "logical chassis", similar to port cards in a chassis switch. This simplifies management, monitoring, and operations since policy and security configuration parameters can be easily shared across all switches in the logical chassis. In addition, because information about connections to physical and virtual servers and storage is now known to all switches in the fabric, the fabric can ensure all network policies and security settings continue to be applied to any given virtual machine no matter whether it moves or where it resides.

## ADDRESSING DATA CENTER CHALLENGES WITH ETHERNET FABRIC

In modern, virtualized data centers, IT groups need to better scale virtual server environments, provide application mobility, and reduce infrastructure complexity and management overhead.

### Scaling Virtual Server Environments

When organizations attempt to scale virtual server environments, the network often presents challenges due to Spanning Tree Protocol (STP, shown in Figure 2), the growing number of GbE connections per server, low utilization, and link failure recovery.

Enabling virtualization capabilities, such as Virtual Machine (VM) mobility, requires VMs to migrate within a single Layer 2 network, since non-disruptive migration of VMs across Virtual LANs (VLANs) using Layer 3 protocols is not supported by virtualization hypervisors. In traditional Layer 2 Ethernet networks, to create a highly available network, organizations designate paths through the network as active or standby using STP. While this provides an alternate path, only one path can be used at a time, which means that network bandwidth is not well utilized. Since one of the goals of server virtualization is to increase utilization of the physical server, increased utilization of network bandwidth should also be expected.

To increase network utilization, Multiple Spanning Tree Protocol (MSTP) and similar protocols allow for separate spanning trees per VLAN. While this improves bandwidth utilization, the STP limit of one active path between switches remains. And, because traffic paths are manually configured with MSTP, complexity increases.

Another challenge with STP is network behavior when links fail. When failures occur, the spanning tree needs to be redefined. This can take anywhere from five seconds with Rapid Spanning Tree (RSTP) up to several minutes with STP—and this convergence can vary unpredictably even with small topology changes. The demands for non-stop traffic flow increases with server virtualization, and consequently network convergence times have to shrink. STP does not provide an adequate solution for these requirements. Finally, when a spanning tree is reconverging, broadcast storms can occur and result in network slowdown. All of these limitations of STP are why Layer 2 networks are typically kept small in the data center.

### Ethernet Fabrics

Compared to classic hierarchical Ethernet architectures, Ethernet fabrics provide higher levels of performance, utilization, availability, and simplicity. They have the following characteristics at a minimum:

**Flatter.** Ethernet fabrics eliminate the need for Spanning Tree Protocol, while still being completely interoperable with existing Ethernet networks.

**Flexible.** Can be architected in any topology to best meet the needs of any variety of workloads.

**Resilient.** Multiple "least cost" paths are used for high performance and high reliability.

**Elastic.** Easily scales up and down at need.

More advanced Ethernet fabrics borrow further from Fibre Channel fabric constructs:

- They are self-forming and function as a single logical entity, in which all switches automatically know about each other and all connected physical and logical devices.
- Management can then be domain-based rather than device-based, and defined by policy rather than repetitive procedures.
- These features, along with virtualization-specific enhancements, make it easier to explicitly address the challenges of VM automation within the network, thereby facilitating better IT automation.
- Protocol convergence (eg Fibre Channel over Ethernet, or FCOE) may also be a feature, intended as a means of better bridging LAN and SAN traffic.

In contrast, consider the benefits of a Layer 2 network that:

- Is highly available
- Guarantees high-bandwidth utilization over equal-costs paths
- Does not stall traffic when links are added or removed due to failure or network reconfiguration
- Makes latency deterministic and lossless
- Can transport IP and mission-critical storage traffic over the same wire

These are some of the Ethernet fabric features that enable efficient scaling of virtual server environments without VM mobility constraints and potential network downtime.

### Application Mobility

When an application is running in a VM rather than on a physical server, it is not tied to a specific physical server. This in theory allows a VM to move between physical servers when application demands change, when servers need to be maintained, and when a quick recovery from site disasters is necessary.

VM mobility can occur within a cluster of physical servers that are in the same IP subnet and Ethernet VLAN. This is required for the migration to be non-disruptive to client traffic as changes in the IP subnet are necessarily disruptive. As described in the review of STP limitations, the sphere of VM migration can be further constrained. The solution for flexible VM mobility is a more scalable and available Layer 2 network with higher network bandwidth utilization.

For a VM to migrate from one server to another, many server attributes must be the same on the origination and destination servers. This extends into the network as well, requiring VLAN, Access Control List (ACL), Quality of Service (QoS), and security profiles to be the same on both the source and destination access switch ports. If switch port configurations differ, either the migration pre-flight will fail or network access for the VM will break, as shown in Figure 3. Organizations could map all settings to all network ports, but that would violate most networking and security best practices. The distributed virtual switch in VMware vSphere 4 addresses some of these issues, but at the cost of consuming physical server resources for switching, added complexity in administering network policies at multiple switch tiers, and a lack of consistent security enforcement for VM-to-VM traffic.

With automated VM migration, network administrators will have limited visibility to the location of applications. This makes troubleshooting a challenge, and pinpointing issues to a specific VM will be like finding a needle in a haystack.

Now, again consider a Layer 2 network that:

- Places no physical barriers in the way of VM migration
- Is aware of VM locations and consistently applies network policies
- Does not require manual intervention when a VM moves to a new machine
- Removes the overhead of switching traffic from the hypervisor for maximum efficiency and functionality
- Supports heterogeneous server virtualization in the same network

More advanced, highly VM-aware Ethernet fabrics allow IT organizations to broaden the sphere of application mobility, provide VM awareness, and optimize server resources for applications.

### Network Management

Similar to data center LANs today, multi-tier architectures involve considerable complexity (as shown in Figure 4), paired with the long list of Layer 2 and 3 protocols with which administrators have to be familiar. And the network and its intersection with other domains have gotten much more complicated to manage. The access layer is no longer managed via a single switch, but now includes multiple stages of switching that extend from the software switch in the hypervisor (called a “softswitch”) to the top-of-rack or end-of-row access switch. Each time a new rack of servers is deployed to host VMs, each switching layer has to be configured, driving up cost and complexity.

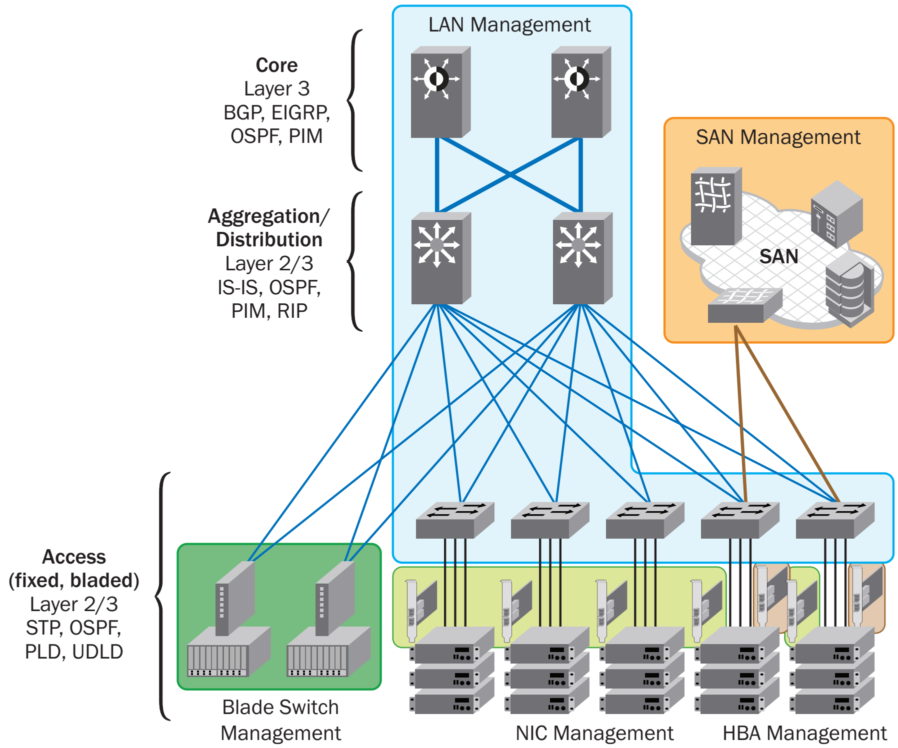



**Figure 4.**
Multi-tier network architectures and many Layer 2 and 3 protocols increase complexity and drive up management costs.

Contributing to management complexities are the separate tools used to manage the LAN, SAN, blade server connectivity, Network Interface Cards (NICs), and Host Bus Adapters (HBAs). Often administrators can see only what is in their direct line of responsibility and do not have the overall view of the entire network environment. Data center management and orchestration tools by themselves will not take care of this any more than virtualization alone does. Infrastructure itself needs to become far simpler and more able to share information automatically and laterally across devices, not just upwards to the management stack.

Ethernet fabrics address this need by:

- Logically eliminating the management of multiple switching layers
- Applying policies and managing traffic across many physical switches as if they were one switch
- Scaling network bandwidth without manual reconfiguration of switch ports and network policies
- Making a unified, customized view of network status available to server, network, and storage administrators

## BROCADE VCS ETHERNET FABRIC TECHNOLOGY

Brocade VCS™ fabric technology allows data center teams to create efficient data center networks that just work. Ethernet fabric architectures built on Brocade VCS technology share information across nodes and may be managed as a single logical chassis, greatly simplifying management and reducing operational overhead. Brocade VCS technology offers unmatched VM awareness and automation versus traditional architectures and competitive fabric solutions and supports storage over a converged fabric when the organization is ready.

Brocade VCS technology, backed by a heritage of proven fabric innovations, delivers IT agility and assures reliability, with a cost-effective point of entry to allow you to transition gracefully to elastic, highly automated, mission-critical networks in the virtualized data center.

VCS technology is embedded in the Brocade VDX™ Data Center Switch portfolio. Brocade VDX Data Center Switches are available today to enable the construction of Ethernet fabrics to support cloud-optimized networking and greater enterprise agility. For more information on Brocade VCS technology, please go to:
http://brocade.com/solutions-technology/technology/vcs-technology

## ABOUT BROCADE

Brocade® networking solutions help the world's leading organizations transition smoothly to a virtualized world where applications and information reside anywhere. This approach is based on the Brocade One™ unified network strategy, which enables a wide range of consolidation, convergence, virtualization, and cloud computing initiatives.

Offering an industry-leading family of Ethernet, storage, and converged networking solutions, Brocade helps organizations achieve their most critical business objectives through unmatched simplicity, non-stop networking, application optimization, and investment protection. To ensure a complete solution, Brocade partners with world-class IT companies and provides a full range of education, support, and professional services offerings. Learn more at 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.com

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



BROCADE®