WHITE PAPER

big switch networks

Big Cloud Fabric™

A NEXT-GENERATION DATA CENTER SWITCHING PLATFORM

TABLE OF CONTENTS

SUMMARY

Big Cloud Fabric (BCF) is the next-generation data center switching fabric delivering operational velocity, network automation and visibility for cloud-native applications and software-defined data centers. Inspired by hyperscale ideas, and designed to work in enterprise data centers of any size, BCF is a scale-out SDN fabric with native ability to support multi-tenancy. BCF eliminates the box-by-box operational paradigm that legacy networks are bound by, driven by a tightly coupled control and data plane in every networking box. BCF leverages a centralized SDN controller on open networking switches to deliver intelligence, agility and deployment flexibility, within flat IT budgets.

WHITE PAPER

INTRODUCTION

Enterprise data centers are challenged today to support cloud-native applications, drive business velocity and work within flat budgets.

The network layer is often cited as the least agile part of data center infrastructure, especially when compared to compute infrastructure. The advent of virtualization changed the server landscape and delivered operational efficiencies across management workflows via automation. Emerging cloud-native applications are expected to demand even greater agility from the underlying infrastructure. Click here for more information on the challenges and the need for a next generation data center network.

Most data centers are built using old network architecture, a box-by-box operational paradigm that inhibits the pace of IT operations to meet the demand of modern applications and software-defined data centers.

With virtualization going mainstream, networks are required to provide visibility into virtual machines, east-west traffic across VMs, and deliver network service connectivity easily. Networks are expected to not adversely impact software-defined data center agility by mandating manual box-by-box network configuration and upgrades.

BCF addresses these problems.

WHAT IS BIG CLOUD FABRIC?

BCF is the next-generation data center switching fabric delivering operational velocity, network automation and visibility for cloud-native applications and software-defined data centers, while staying within flat IT budgets.

BCF leverages software-defined networking (SDN) to make networks intelligent, agile and flexible.

A. **Intelligence** stems from its software-defined fabric acting as “one logical switch” that simplifies operations and provides full visibility & telemetry.

B. **Agility** is delivered via network automation for rapid deployment of applications and services, one-click troubleshooting for faster resolution, and zero touch upgrades.

C. **Flexibility** is provided by BCF’s scale-out fabric to accommodate future growth in east-west traffic without breaking the bank. Leveraging its ability to run on open hardware, BCF provides network hardware vendor choice and lowers costs.

Built using open networking (white-box or brite-box) switches and SDN controller technology, BCF is designed to deliver innovation velocity under flat budgets.

BIG CLOUD FABRIC TECHNOLOGY & ARCHITECTURE

BCF is based on an SDN technology approach. SDN fabric architecture refers to a separation of the network's data and control plane, followed by a centralization of the control plane functionality. In practice, it implies that the network's policy plane, management plane and much of the control plane are externalized from the hardware device itself, using an SDN controller, with few on-device functions, for scale and resiliency. The network state is centralized but hierarchically implemented, instead of being fully distributed on a box-by-box basis across access and aggregation switches. Controller-based designs not only bring agility via centralized programmability and automation, but they also streamline fabric designs (e.g. leaf-spine L2/L3 Clos) that are otherwise cumbersome to implement and fragile to operate in a box-by-box design.

The BCF architecture consists of a physical switching fabric, which is based on a leaf-spine Clos architecture. Optionally, the fabric architecture can be extended to virtual switches residing in the hypervisor. Leaf and spine switches running Switch Light™ Operating System form the individual nodes of this physical fabric. Switch Light Virtual, running within the hypervisor, extends the fabric to the virtual switches. Intelligence in the fabric is hierarchically placed: most of it in the BCF Controller (where configuration, automation and troubleshooting occur), and some of it off-loaded to Switch Light Operating System for resiliency and scale-out.

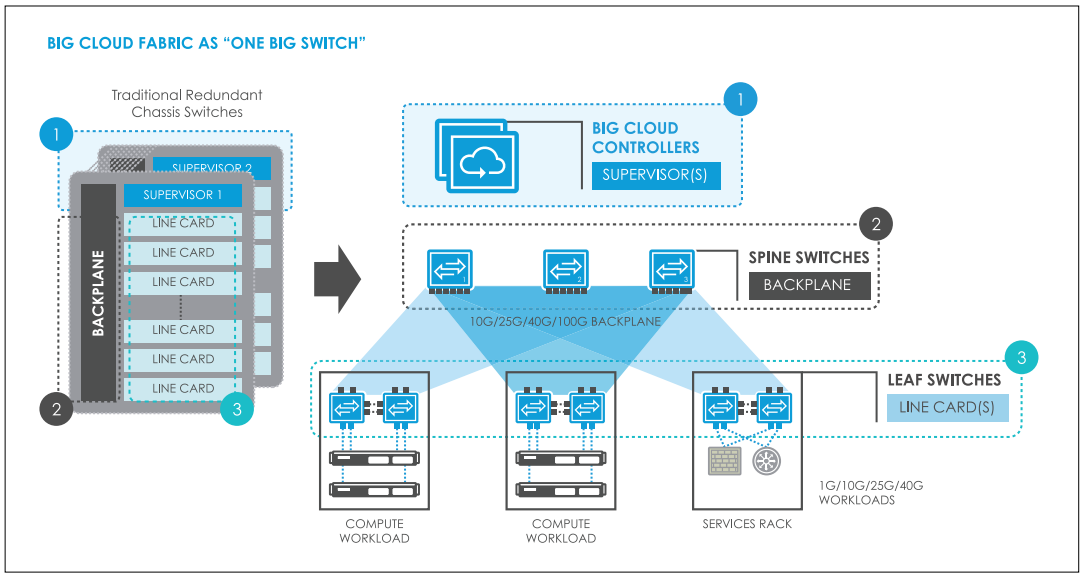


Figure 1:
BCF Leaf-Spine Clos Architecture Overview

The BCF architecture can be viewed as software-driven/software-defined disaggregation of the traditional chassis:

Supervisor Card(s) --> SDN Controller Cluster

Chassis Backplane --> Spine Switches

Line Card(s) --> Leaf Switches

BCF's logical chassis architecture—or a one logical pod switch—enables data center users to construct high-performance, resilient and scale-out (pay-as-you-scale) pod fabrics. SDN controller (a.k.a. chassis supervisor) is a critical and necessary component without which chassis disaggregation is not possible. It enables zero-touch fabric operation, single-point of fabric management, and centralized fabric analytics—thus significantly speeding up application deployment and reducing operational cost. Moreover, with HW/SW disaggregation, BCF supports a variety of high-performance (10G/25G/40G/100G), high-density switch hardware enabling full vendor choice and tremendous CapEx cost optimization.

BIG CLOUD FABRIC PRODUCT COMPONENTS

- **BCF Controller Cluster** — a centralized and hierarchically implemented SDN controller available as a pair of hardware appliances for high availability (HA)
- **Switch Light Operating System** — a light-weight open networking switch OS purpose built for SDN
- **Open Networking Leaf and Spine Switch Hardware** — the term 'open networking' (white-box or brite-box) refers to the fact that the Ethernet switches are shipped without embedded networking OS. The merchant silicon networking ASICs used in these switches are the same as used by most incumbent switch vendors and have been widely deployed in production in hyperscale data center networks. These bare metal switches ship with Open Network Install Environment (ONIE) for automatic and vendor-agnostic installation of third-party network OS. A variety of switch HW configurations (10G/25G/40G/100G) and vendors are available on the Big Switch hardware compatibility list.

Optional components include:

- **VMware vCenter Extension / GUI Plugin** — built-in network automation and VM Admin visibility for vSphere server virtualization and NSX network virtualization
- **OpenStack Plugin** — BCF Neutron plugin or ML2 Driver Mechanism for integration with various distributions of OpenStack
- **Container Plugin** — BCF plug-n for various container orchestrators for container-level network automation and visibility
- **Switch Light VX** — high-performance user space software agent for KVM-based Open vSwitch (OVS) for OpenStack and container use-cases.

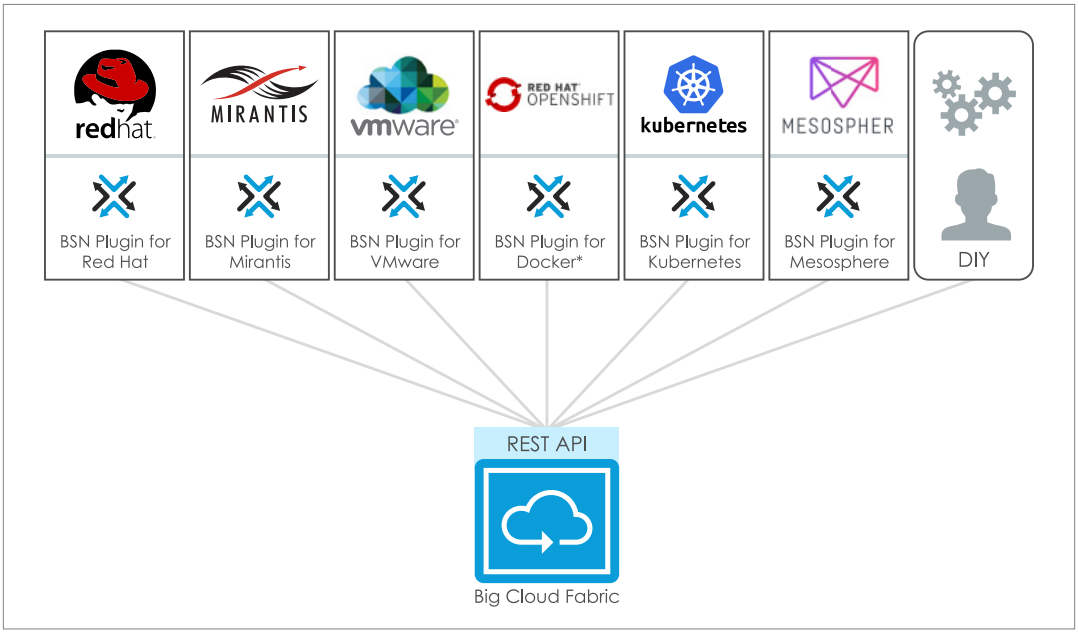


Figure 2: Supported BCF Plugins

BCF BENEFITS

Centralized Controller and Network Automation improves operational agility

With configuration, automation and most troubleshooting done via the BCF controller, a network administrator can work within a single console, regardless of the number of racks the BCF fabric extends over, for tasks such as provisioning new physical capacity, deploying new applications, logical segments, and addition/removal of tenants. For example, in a 32 rack pod with dual leaf switches and four spine switches, a traditional network design would have 68 management consoles. The BCF design has only one—the controller console—that performs the same functions. The result is massive time savings, reduced error rates and simpler automation designs. As a powerful management tool, the controller console exposes a web-based GUI, a traditional networking-style CLI and REST APIs.

Tenant-native construct enables rapid innovation

In the BCF design, configuration in the CLI, GUI or REST API is based on the concept of logical tenants. Each tenant has administrative control over a logical L2/L3/policy design that connects the edge ports under the tenant's control. The BCF controller has the intelligence to translate the logical design into optimized entries in the forwarding tables of the spine, leaf and vleaf.

Open Networking Switch Hardware Reduces CapEx Costs by Over 50%

By adding up hardware, software, maintenance and optics/cables, a complete picture of the hard costs over three years shows that the savings are dramatic. The operational savings delivered via network automation are also significant across various network workflows, as can be seen from the below table.

Workflow	8 Rack OpenStack Pod	16 Rack VMware Pod
	Incumbent vs. Dell and Big Cloud Fabric	Incumbent vs. Dell and Big Cloud Fabric
Application Deployment	16x Faster	16x Faster
Initial Set-up	8x Faster	8x Faster
LAG/Fabric Formation	12x Faster	18x Faster
Software Language	20x Faster	30x Faster
Connectivity Troubleshooting Time	12x Faster	12x Faster
Pod Expansion	12x Faster	n/a

Table 1: BCF delivers operational velocity across various network workflows

Built-in Orchestration Support Streamlines DC Operations

BCF Controller natively supports integration with various Cloud Management Platforms (CMPs)—VMware (vSphere, NSX & VIO), and OpenStack—through a single programmatic interface. This is tremendously simpler and scalable compared to box-by-box networking which demands an exponentially larger number of programmatic interactions with CMPs. Data center admins benefit from streamlined application deployment workflows, enhanced analytics and simplified troubleshooting across physical and virtual environments.

Scale-out (Elastic) Fabric

The BCF's flexible, scale-out design allows users to start at the size and scale that satisfies their immediate needs while future proofing their growth needs. By providing a choice of hardware and software solutions across the layers of the networking stack and pay-as-you-grow economics, starting small scale and growing the fabric gradually instead of locking into a fully integrated proprietary solution, provides a path to a modern data center network. Once new switches (physical or virtual) are added, the controller adds those switches to the fabric and extends the current configuration hence reducing any error that may happen otherwise. Customers take advantage of one-time configuration of the fabric.

DC-grade Resilience

The BCF provides DC grade resiliency that allows the fabric to operate in the face of link or node failures as well as in the rare situation when the controller pair is unavailable (headless mode). Swapping a switch (in case of HW failure or switch repurpose) is similar to changing a line card in a modular chassis. After re-cabling and power up, the switch boots up by downloading the right image, configuration and forwarding tables. Additionally, the BCF Controller coordinates and orchestrates entire fabric upgrade ensuring minimum fabric down time. These functionalities further enhance fabric resiliency and simplify operations.

BCF ARCHITECTURAL ADVANTAGES

The benefits listed above are driven primarily by BCF's architectural advantages. Some of the key BCF architectural advantages are:

Tenant-native configuration model

BCF has a tenant-native architecture, that allows L3 network policies to be applied for logically defined tenants. This architectural capability allows a single BCF fabric to share multiple tenants across vCenters. This allows BCF to deliver distributed L3 services such as add/remove logical router, logical segment etc., across the entire physical SDN fabric on a per-tenant basis. Thus BCF makes it very easy to extend tenant-based application isolation in the compute layers to the SDN fabric as well, via tenant-native configuration and policies.

Zero-touch workflows

BCF is a zero-touch fabric, driven by its network automation capabilities. In legacy box-by-box networks, the operational workflows are mostly manual (sometimes scripted, but the scripts must be maintained), resulting in loss of agility. BCF architecture has network automation built-in. See list of automated workflows in figure 4. Rapid network upgrades are up to 20x faster, compared to legacy box-by-box networks. That results in lower network downtime, higher agility and significant OpEx savings.

Zero-Touch Fabric
(REST APIs, GUI, CLI)

Feature	Big Cloud Fabric
Switch OS Install	Automatic
Link Aggregation	Automatic
Fabric Formation	Automatic
Add/Remove Switch	Automatic
ESX host and connectivity to fabric	Automatic
Network Policy Migration with vMotion	Automatic
Hitless Upgrade	Automatic

Table 2

Fabric-Wide Visibility, Troubleshooting and Analytics

The BCF GUI provides (a) health statistics for controller, leaf and spine switches, (b) fabric topology, connectivity graph and traffic heat map, (c) physical and logical resource inventory, (d) top talkers, (e) path tracing and fabric-wide visibility across both *logical paths* (segment, router, policy) and *physical paths* (vSwitch/VTEP, Leaf and Spine switches) to simplify troubleshooting, and (f) fabric analytics to visualize fabric-wide historical events and logs, plus VM/container-centric information from associated orchestrators such as VMware vCenter, and Kubernetes.

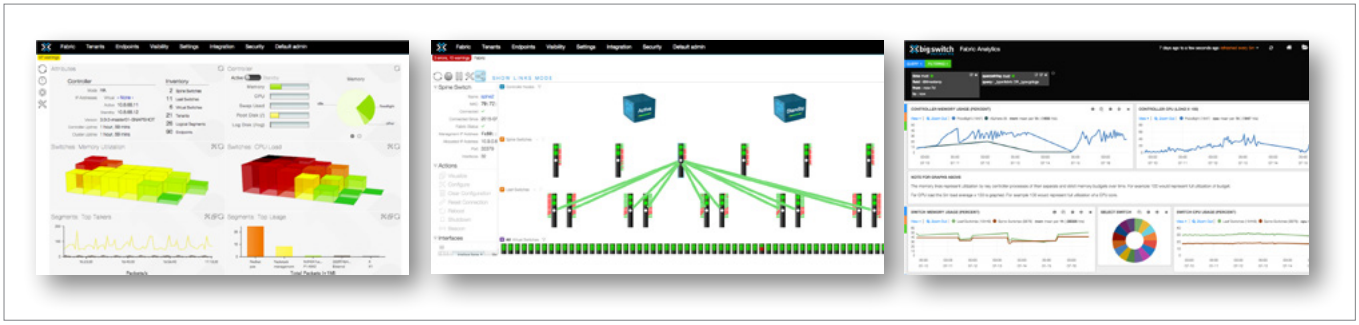


Figure 3:
BCF delivers fabric-wide graphical views and analytics on health statistics, connectivity, inventory, and network events over both logical & physical paths. Significantly simplifies troubleshooting.

BCF vPODs

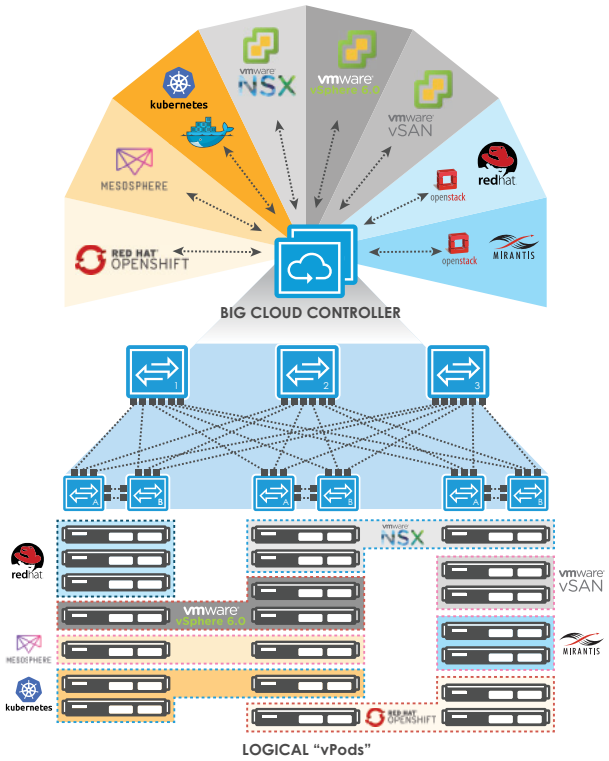
With virtual pod (vPod) technology, BCF can support overlapping IPs and VLANs across *isolated vPods*, in a single SDN fabric. BCF's orchestration integration allows multiple isolated vCenter instances and/or multiple OpenStack instances in parallel. It is ideal for managed private clouds, engineering (dev & test) environments. BCF offers a fully distributed logical router that enables shared services for host elements connected to multiple vCenters. The vPod approach is extensible to container environments as well.

High Performance & Resilience

BCF runs on open network hardware switches to provide up to 100G network. Its centralized controller architecture with leaf spine architecture allows BCF to accommodate multiple failures within seconds without any impact to application performance. BCF delivers best in class high availability at scale.

Provisioning Flexibility

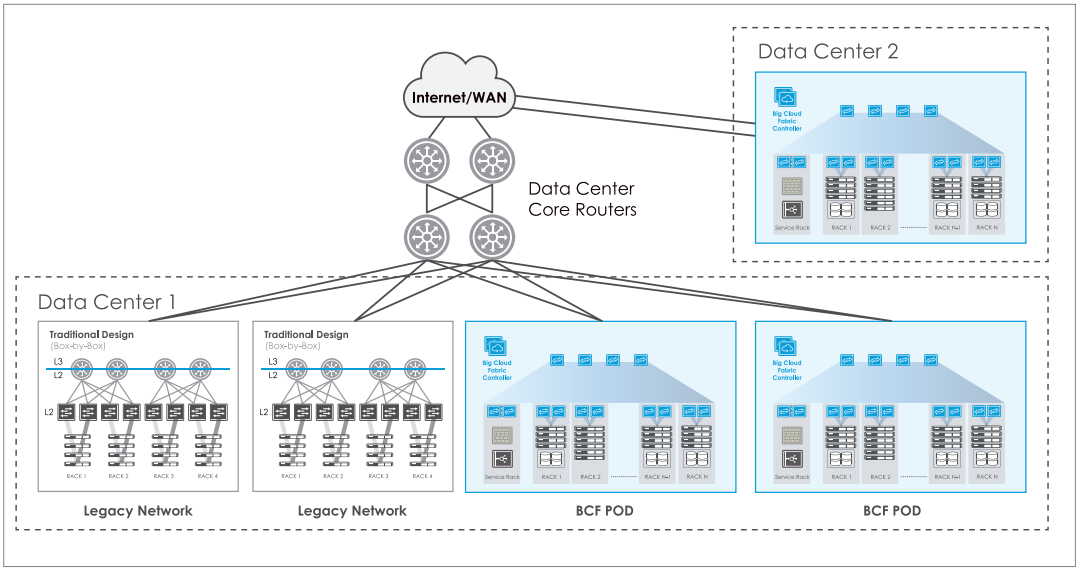
The fabric may be configured via BCF GUI, CLI or REST APIs. The CLI commands leverage the network admins' familiarity with traditional networking CLI commands, while providing a tenant-native construct to apply policies, instead of box-by-box configuration of interfaces. The REST APIs coupled with tenant-native architecture allow multiple orchestrators to program and leverage a single BCF fabric for multiple tenants, while maintaining consistent network operational workflows. This is in contrast to legacy box-by-box networks, where lack of centralized control and logical tenant separation forces network operators to manage each box independently, significantly increasingly operational overhead.



Brownfield Deployments

BCF has a modular architecture that allows enterprises to avoid a rip and replace approach to introduce next generation data center networking technology in their environments. BCF may be introduced in brownfield data centers using core-n-pod network designs, managed by a centralized BCF controller. Inter-pod communication via core routers, allows for various BCF pods situated across data centers co-existing with legacy networks, thus enabling a granular and methodical approach to expand BCF footprint, depending on use-cases.

Figure 4:
BCF interoperates with legacy networks. Pod based BCF deployment allows enterprises to introduce next-generation data center network incrementally.



Foundational for Intent-based Networking

Intent-based networking is in very nascent stages. It attempts to promote an operational approach of specifying "the what" with the underlying networking layer to automatically compute, provision, and deliver "the how". Clearly, network automation is a foundational element to realize intent-based networks' purported benefits. BCF today, supports several network-based intents in an automated manner such as (a) deploying switches based on its role (leaf or spine), (b) verifying VM/host connectivity, and (c) validating Clos fabric design. The intent-based approach is very new and hence, immature at this stage. However, BCF architecture is foundational in nature, so it stands to future-proof your next generation data center network from newer emerging paradigms as well.

Figure 5:
BCF architecture has foundational attributes to support intent-based networking.

TYPICAL USE CASES

Even as software-defined data centers and cloud-native application drive the next leg of innovation in the network layer, a product like BCF is applicable in several environments.

BCF is relevant for use cases that involve:

- VMware SDDC workloads (vSphere, NSX, Virtual SAN)
- OpenStack including NFV.
- Containerized workloads.
- Private clouds
- VDI workloads
- Big Data
- Software Defined Storage (SDS).

BCF can also be easily deployed in existing (brownfield) data centers, as a pod fabric, and supports full communication between modern BCF pod and legacy network via core router.

More details on the various ways to introduce a next-generation data center network in your environment may be found here.

RESOURCES

- The links provide more information on why Big Cloud Fabric is needed, overview of next-generation DC switching architecture, and its drivers and benefits.
- Big Cloud Fabric Community Edition is free and can be downloaded here.
- To try BCF on-line, sign up for our free Big Switch Labs and test drive variety of use cases.

big switch networks

Headquarters
3965 Freedom Circle, Suite 300
Santa Clara, CA 95054

+1.650.322.6510 **TEL**
+1.800.653.0565 **TOLL FREE**

www.bigswitch.com
info@bigswitch.com

Big Cloud Fabric: A Next-Generation Data Center Switching Platform Whitepaper JAN 2017