



Big Cloud Fabric 3.0

LEAF-SPINE CLOS FABRIC FOR DATA CENTERS

Big Switch's Big Cloud Fabric is a leaf-spine Clos fabric providing physical and virtual workload connectivity in data center pods. Embracing the hyperscale data center design principles, Big Cloud Fabric solution manages physical and virtual switching infrastructure to enable rapid innovation, ease of provisioning and management, while reducing overall costs.

BIG SWITCH NETWORKS

Our mission is to bring hyperscale networking to a broader audience—ultimately removing the network as the biggest obstacle to rapid deployment of new applications.

We do this by delivering all the design philosophies of hyperscale networking in a single solution.

The Big Cloud Fabric Features:

- Open Networking (Britebox or Whitebox) Hardware to Reduce Cost
- SDN Controller Technology to Reduce Complexity
- Core and Pod Designs to Innovate Faster

BIG CLOUD FABRIC OVERVIEW

Big Switch's Big Cloud Fabric™ (BCF) is the industry's first open networking SDN data center fabric bringing hyperscale data center design principles to cloud environments making it ideal for current and next generations data centers. Applications can now take advantage of high east-west bisectional bandwidth, secure multi-tenancy, and workload elasticity natively provided by Big Cloud Fabric. Customers benefit from unprecedented application agility due to automation, massive operational simplification due to SDN and, dramatic cost reduction due to HW/SW disaggregation.

Big Cloud Fabric supports both physical and virtual (multi-hypervisor) workloads and choice of orchestration software¹. It provides L2 switching, L3 routing and, L4-7 service insertion and chaining while ensuring high bisectional bandwidth. The scalable fabric is fully resilient with no single point of failure and supports head-less mode operations.

Big Cloud Fabric is available in two editions:

- **P Fabric** — Leaf-spine physical Clos fabric controlled via SDN Controller
- **P+V Fabric** — Leaf-spine plus virtual switches controlled by SDN Controller

ARCHITECTURE:

SDN SOFTWARE MEETS OPEN NETWORKING HARDWARE

Software Defined Networking (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 control plane are externalized from the hardware device itself, using an SDN controller, with few on-device off-load 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.

Get hands-on experience with our offering, register for a free online trial at:
labs.bigswitch.com

Contact our sales team at
sales@bigswitch.com

¹. See specific hypervisor and orchestration support in a later section of this datasheet

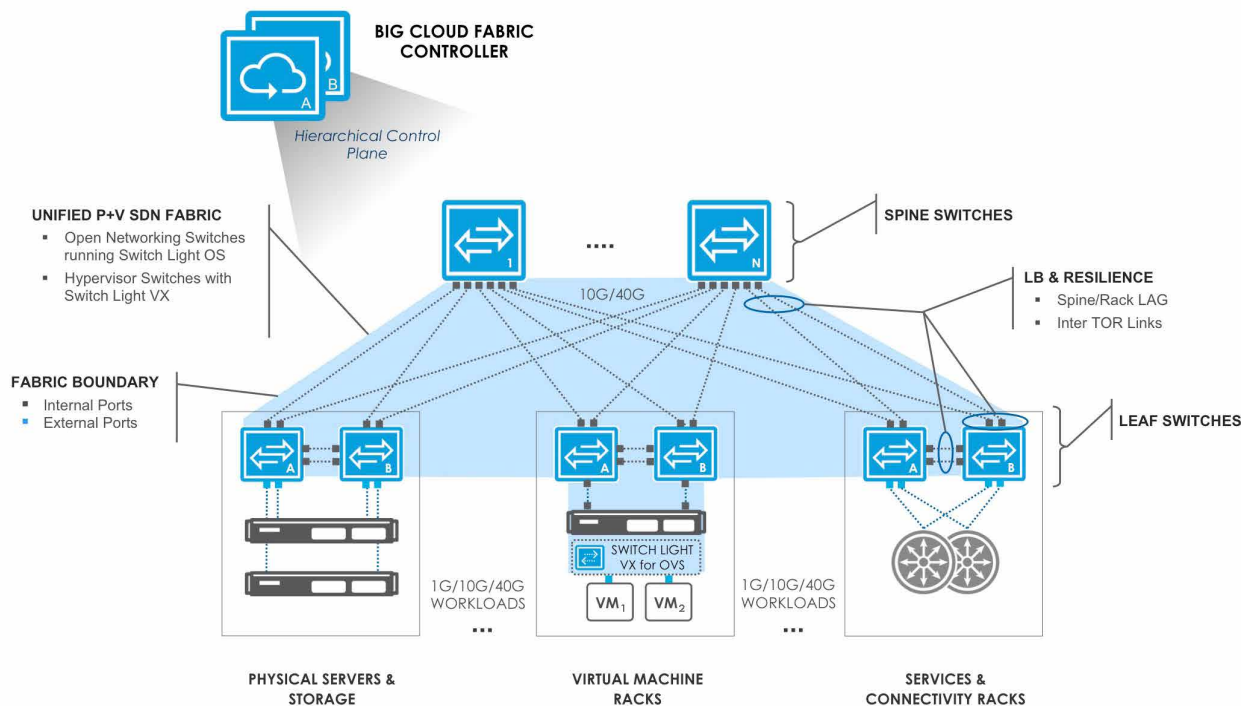


Figure 1: Big Cloud Fabric (Leaf-Spine Clos Architecture)

The Big Cloud Fabric 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 Big Cloud Fabric Controller (where configuration, automation and troubleshooting occur), and some of it off-loaded to Switch Light for resiliency and scale-out.

BIG CLOUD FABRIC SYSTEM COMPONENTS

- **Big Cloud Fabric Controller Cluster** — a centralized and hierarchically implemented SDN controller available as a cluster of hardware appliances for high availability (HA)
- **Open Networking Leaf and Spine Switch Hardware** — the term ‘open networking’ (whitebox or britebox) 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 datacenter 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/40G) and vendors are available on the Big Switch hardware compatibility list.

- **Switch Light™ Operating System** — a light-weight open networking switch OS purpose built for SDN
- **Switch Light VX** — optional high-performance user space software agent for KVM-based Open vSwitch (OVS).
- **OpenStack Plug-In** — optional BSN Neutron plug-in or ML2 Driver Mechanism for integration with various distributions of OpenStack
- **VMware vCenter Extension** — built-in network automation and orchestration for vSphere server virtualization and NSX network virtualization
- **CloudStack Plug-In** — optional BSN Networking plug-in for integration with CloudStack

DEPLOYMENT SOLUTIONS

Big Cloud Fabric is designed from the ground up to satisfy the requirements of physical, virtual or combination of physical and virtual workloads. Some of the typical Pod deployment scenarios include:

- Unified P+V SDN for OpenStack Clouds
- VMware Data Centers — vCenter, NSX or VIO
- High Performance Computing / Big Data / Software Defined Storage Pods
- Virtual Desktop Infrastructure (VDI) Pods
- Specialized NFV Pods

EXAMPLE SCENARIO	SUPPORTED WORKLOADS	LEAF SWITCH CONFIGURATION	SPINE SWITCH CONFIGURATION
Private/Public Cloud—Typical data center pod deployments	1G, 10G	48X10G + 6x40G	32x40G
Cost Optimized Fabric (leverage existing cable infrastructure)	1G, 10G	48X10G + 6x40G	48X10G + 6x40G
High performance 40G storage array and 10G workloads (using splitter cables)	10G, 40G	48X10G + 6x40G or 32x40G	32x40G
High Performance Computing / Software Defined Storage / Big Data Analytics	40G	32x40G	32x40G
Dense 10G Compute (using splitter cables)	10G	32x40G	32x40G

Figure 2: Example BCF Deployment Scenarios

The BCF fabric can be designed to support the above listed deployment scenarios using a combination of Open Networking Ethernet switch options. A few examples are listed in the table shown in Figure 2.

BIG CLOUD FABRIC BENEFITS

Centralized Controller Reduces Management Consoles By Over 30:1

With configuration, automation and most troubleshooting done via the Big Cloud Fabric controller, the number of management consoles involved in provisioning new physical capacity or new logical apps goes down dramatically. For example, in a 16 rack pod with dual leaf switches and two spine switches, a traditional network design would have 34 management consoles. The Big Cloud Fabric 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.

Streamlined Configuration, Enabling Rapid Innovation

In the Big Cloud Fabric 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 Big Cloud Fabric 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.

Built-in Orchestration Support Streamlines DC Operations

Big Cloud Fabric Controller natively supports integration with various Cloud Management Platforms (CMPs)—VMware (vSphere, NSX Manager & VIO), OpenStack and CloudStack—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

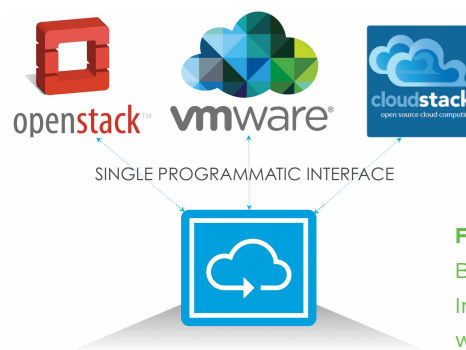


Figure 3:
BCF Supports
Integration
with CMPs

CMPs. Data center admins benefit from streamlined application deployment workflows, enhanced analytics and simplified troubleshooting across physical and virtual environments.

Unified P+V SDN Fabric Enables Deep Visibility and Resilience for OpenStack Networks

The BCF OpenStack Neutron plugin for L2/L3 networking provides resiliency necessary for production-grade OpenStack deployments – including support for distributed L3 routing and distributes NAT (Floating IP). The Big Cloud Fabric Controller acts as the single pane for provisioning, troubleshooting, visibility and analytics of the entire physical and virtual network environment. This enables data center operators to deploy applications rapidly, simplifies operational workflows and provides immediate root-cause analysis when application performance issues arise.



Figure 4: BCF Graphical User Interface (GUI)

Network/Security/Audit Workflow Integration

The Big Cloud Fabric controller exposes a series of REST APIs used to integrate with application template and audit systems, starting with OpenStack. By integrating network L2/L3/policy provisioning with OpenStack HEAT templates in Horizon GUI, the time to deploy new applications is reduced dramatically as security reviews are done once (on a template) rather than many times (on every application). Connectivity edit and audit functions allow for self-service modifications and rapid audit-friendly reporting, ensuring efficient reviews for complex applications that go beyond the basic templates.

Scale-out (Elastic) Fabric

The Big Cloud Fabric'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 Resiliency

The Big Cloud Fabric 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 entire controller cluster is unavailable (headless mode). Swapping a switch (in case of HW failure or switch repurpose) is similar to changing a line card in modular chassis. After re-cabling and power up, 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.

USING BCF: A 3-TIER APPLICATION EXAMPLE

The Big Cloud Fabric supports a multi-tenant model, which is easily customizable for the specific requirements of different organizations and applications. This model increases the speed of application provisioning, simplifies configuration, and helps with analytics and troubleshooting. Some of the important terminology used to describe the functionality include:

Tenant — A logical grouping of L2 and/or L3 networks and services.

- Logical Segment — An L2 network consisting of logical ports and end-points. This defines the default broadcast domain boundary.
- Logical Router — A tenant router providing routing and policy enforcement services for inter-segment, inter-tenant, and external networks.
- External Core Router — A physical router that provides connectivity between Pods within a data center and to the Internet.
- Tenant Services — Services available to tenants and deployed as dedicated or shared services (individually or as part of a service chain).

Tenant Workflow

In the most common scenario, end consumers or tenants of the data center infrastructure deal with a logical network topology that defines the connectivity and policy requirements of applications. As an illustrative example, the canonical 3-tier application in Figure 5, shows various workload nodes of a tenant named "BLUE". Typically, a tenant provisions these workloads using orchestration software such as OpenStack, VMware vSphere, or BCF Controller GUI/CLI directly. As part of that provisioning workflow, the Big Cloud Fabric Controller seamlessly handles enabling the logical topology onto the physical and virtual switches.

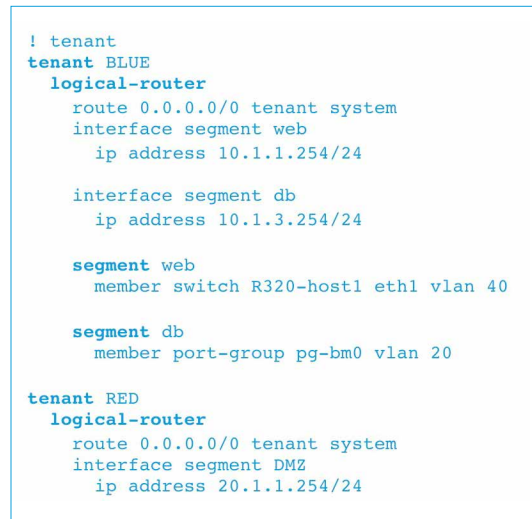


Figure 6: Application Centric Configuration

An illustrative sample set of entries in various forwarding tables highlight some of the salient features of the Big Cloud Fabric described in earlier sections.

- L3 routing decision is made at the first hop leaf or vleaf switch (distributed virtual routing — no hair-pinning)
- L2 forwarding across the pod without special fabric encapsulation (no tunneling)
- Full load-balancing across the various LAG links (leaf and spines)
- Leaf/Spine mesh connectivity within the physical fabric for resilience



DATASHEET

BIG CLOUD FABRIC FEATURES.

FEATURE	DESCRIPTION / BENEFIT
Zero Touch Fabric (ZTF)	ZTF enables complete control and management of physical switches within BCF without manually interacting with the switches. It tremendously simplifies day-to-day network operations: • auto-configuration and auto-upgrade of Switch Light OS, • automatic topology updates and event notifications based on fabric link state changes. • auto-scaling of the fabric—adding or removing nodes and/or links within the fabric requires no additional configuration changes on the controller.
Fabric LAG	Fabric LAG combines the underlying LAG functionality in switching ASICs with the centralized visibility of the SDN controller to create a highly resilient and efficiently balanced fabric. Compared to spanning tree protocols or even traditional MLAG/ECMP based approaches to multi-path fabric formation, Fabric LAG technology enables significantly reduced convergence time on topology changes and dramatically reduced configuration complexity.
Fabric Sync	Fabric Sync intelligently synchronizes Controller Information Base (CIB) with fabric node's Forwarding Information Base (FIB) using the OpenFlow protocol. During a topology change, only delta updates are synced across impacted switches. Fabric Sync ensures strong CIB-FIB consistency, as it is the single point of control for maintaining all forwarding and associated policy tables.
Resilient Headless Mode	In situations when both controllers are unreachable, fabric nodes are considered to be running in Headless mode. In this mode, all provisioned services continue to function as programmed prior to entering the Headless mode. Additionally, multiple levels of redundancy enable a highly resilient and self-healing fabric even during headless mode.
Centrally-managed Fabric (GUI, CLI & REST APIs)	Big Cloud Fabric Controller provides single pane of glass for entire fabric. • Administrators can configure, manage, debug or troubleshoot, and upgrade the fabric nodes using CLI, GUI, or REST API. • REST APIs, CLI and GUI have application and tenant awareness. Single Pane of Glass fabric management enhances operational simplicity by providing a centralized dashboard for fabric management as well as quick and easy access to troubleshooting, analytics and telemetry information. Additionally, it provides simplified workflow for network operators and administrators.
Fabric Analytics	Fabric Analytics is the set of features that provides Advanced Multi-node Troubleshooting, Analytics & Telemetry in the Big Cloud Fabric solution.
API-first Fabric	Big Cloud Fabric Controller is highly programmable due to its “API-first” design principle and can be implemented as a closed loop feedback system. For example, security applications can dynamically detect threats and program the BCF controller for mitigation. The BCF GUI and CLI utilize the underlying REST APIs—hence are by definition consistent and hardened.
Tenant-aware Fabric	Big Cloud Fabric provides built-in multi-tenancy via tenant-aware configurations, tenant separation and fine-grain inter-tenant access control. Configuration in the CLI, GUI or REST API is based on the concept of logical tenants.
Service-aware Fabric	Big Cloud Fabric supports L3 virtual and physical service insertion and service chaining. Services can be shared across tenants or dedicated to a specific tenant. L2 services will be supported in subsequent releases.

FEATURE	DESCRIPTION / BENEFIT
L2 Features	• Layer 2 switch ports and VLAN trunks • IEEE 802.1Q VLAN encapsulation • Support for up to 4K VLANs (i.e. 4K Logical Segments) • Support for up to 3K ingress access control list (ACL) entries • BPDU Guard • MLAG (up to 16 ports per LAG) • Link Aggregation Control Protocol (LACP): IEEE 802.3ad • Jumbo frames on all ports (up to 9216 bytes) • VLAN translation
L3 Features	• Layer 3 interfaces: Routed ports, Switch Virtual and Interface (SVI) • Support for up to 48K host prefixes (i.e. 48K Endpoints) • Support for 1K Virtual Routing and Forwarding (VRF) entries (i.e. 1K Logical Routers) • 1K Tenants • 10K static routes • Up to 16 ways Equal-Cost Multipathing (ECMP) • 1K Equal-Cost Multipathing (ECMP) groups • 3K flexible ACL entries • Policy-Based Routing • ACL: Routed ACL with Layer 3 and 4 options to match ingress ACL • Jumbo frame support (up to 9216 bytes) • DHCP relay • NAT/PAT support
QoS	• Layer 2 IEEE 802.1p (class of service [CoS]) • CoS-based egress queuing • Tenant/Segment based classification • Marking • Control-Plane Policing (CoPP) or Rate-Limiting
High Availability	• Controller HA • Headless mode (fabric forwards traffic in absence of Controller) • Redundant Spine • Redundant Leaf • Redundant Links • Controller cluster with single Virtual IP • Support for redundant out-of-band management switch
Security	• Ingress ACLs • Layer 3 and 4 ACLs: IPv4, Internet Control Message Protocol (ICMP), Transmission Control Protocol (TCP), User Datagram Protocol (UDP), etc. • ACLs on controller management interface • ACL logging (IPv4 only)
OpenStack Integration	• Nova-network & Neutron ML2 Driver mechanism support • Neutron L3 Plugin support (distributed routing, Floating IP, PAT and Security Group visibility) • Auto Host Detection & LAG Formation • Dynamic Provisioning of the Physical Fabric • OpenStack Horizon Enhancements (Heat Templates, Fabric Reachability Test)
CloudStack/CloudPlatform Integration	• Auto Host Detection & LAG Formation • Dynamic Provisioning of the Physical Fabric

DATASHEET

FEATURE	DESCRIPTION / BENEFIT
VMware vCenter Integration	Provides Fabric Automation and Visibility including: • Auto Host Detection & LAG Formation • Auto L2 Network Creation & VM Learning • Network Policy Migration for vMotion/DRS • VM-level Visibility (VMname, vMotion) • VM-to-VM Troubleshooting (Logical & Physical)
VMware NSX-v Support	Close the overlay/underlay gap for visibility and troubleshooting. Features include: • Auto Host Detection & LAG Formation • AAuto Network Creation for VTEP Port-Group & VTEP Discovery • Underlay Troubleshooting – VTEP-to-VTEP connectivity • Underlay Visibility through Fabric Analytics (VM-name, VXLAN ID, Logical Switch)
Fabric Management	• GUI • CLI-based console to provide detailed out-of-band management • REST API Support • Switch management using 10/100/1000-Mbps management through controller • Beacon LED (based on underlying switch) • Configuration synchronization • Configuration rollback • Secure Shell Version 2 (SSHv2) • Telnet • Username and passwords authentication • TACACS+ / RADIUS • Syslog (4 servers) • SNMP v1 and v2c • sFlow support • SPAN with Policy/ACL • Fabric SPAN with Policy/ACL • Ingress and egress packet counters per interface, per segment, and per tenant • Network Time Protocol (NTP) • Test Path — Enhanced Troubleshooting & Visibility with logical and physical fabric views (VM <--> vLeaf <--> Leaf <--> Spine <--> Leaf <--> vLeaf <--> VM) • Fabric Analytics including telemetry and enhanced analysis
MIBs Bare Metal Switch Vendors	Documented in a separate MIB's document
Industry Standards	• IEEE 802.1p: CoS prioritization • IEEE 802.1Q: VLAN tagging • IEEE 802.3: Ethernet • IEEE 802.3ae: 10 Gigabit Ethernet • IEEE 802.3ba: 40 Gigabit Ethernet
Support for Ethernet-Based	Support for 10G and 40G switches from Dell and Accton / EdgeCore. The common supported switch configurations are: • 48x10G + 6x40G (Broadcom Trident-II ASIC) • 32x40G (Broadcom Trident-II ASIC) Support for 40G to 10G breakout cable. For the complete list of supported switch vendors/configurations as well as optics/cables, included in the Big Cloud Fabric Hardware Compatibility List (HCL), please contact the Big Switch Sales Team at sales@bigswitch.com.

BCF CONTROLLER APPLIANCE SPECIFICATION

The Big Cloud Fabric Controller can be deployed as a Hardware Appliance.

BCF Controller — Hardware Appliance Specification The BCF large controller is available as an enterprise-class, 2-sockets, 1U rack-mount hardware appliance designed to deliver the right combination of performance, redundancy and value in a dense chassis.

FEATURE	TECHNICAL SPECIFICATION	
Controller	HWB (Standard)	HWBL (Large)
Processor	Intel Xeon E5-2620 v3 2.40GHz, 15M Cache, 8GT/s QPI, Turbo, 6 Cores, 2 Sockets, 85W	Intel Xeon E5-2670 v3 2.30GHz, 30M Cache, 9.60GT/s QPI, Turbo, 12 Cores, 2 Sockets, 120W
Form Factor (H x W x D)	1U Rack Server (4.28cm x 43.4cm x 60.7cm)	
Memory	4 x 16GB RDIMM, 2133 MT/s, Dual Rank, x4 Data Width	
Hard Drive	2 x 1TB 7.2K RPM SATA 6Gbps 3.5in Hot-plug Hard Drives; RAID 1 for H330/H730/H730P	
Networking	Embedded NIC: Four 10/100/1000 Mbps Network Adapter: Intel X520 Dual Port 10Gb DA/SFP+ server adapter	
Power	2 x Hot Plug Power Supplies 550W	
Additional Features	Fan fault tolerance; ECC memory, interactive LCD screen; ENERGY STAR® compliant	

ENVIRONMENT	SPECIFICATION
Temperature—Continuous Operation	10°C to 35°C (50°F to 95°F)
Temperature—Storage	-40°C to 65°C (-40°F to 149°F) with a maximum temperature gradation of 20°C per hour
Relative Humidity—Continuous	10% to 80% with 29°C (84.2°F) maximum dew point
Relative Humidity—Storage	5% to 95% at a maximum wet bulb temperature of 33°C (91°F), atmosphere must be non-condensing at all times
Altitude—Continuous	-15.2m to 3048m (-50ft to 10,000ft)
Altitude—Storage	-15.2m to 12,000m (-50ft to 39,370ft)

SUPPORTED WORKLOADS & ORCHESTRATION SYSTEMS

FEATURE	TECHNICAL SPECIFICATION
Physical Workloads	Bare-metal server workloads
Virtual Workloads	see <i>Virtual Workloads</i> table below
Cloud Orchestration	OpenStack Juno (Nova Networking, Neutron ML2 driver) CloudStack / CloudPlatform 4.5

VIRTUAL WORKLOADS FOR P-FABRIC

HYPERVISOR	OPENSTACK—KILO*	OPENSTACK—JUNO*	CLOUDSTACK / CLOUDPLATFORM 4.5	CUSTOM ORCHESTRATION / STAND ALONE
KVM		Ubuntu 14.04 (Mirantis OpenStack)	Ubuntu 12.04	Ubuntu 12.04, 14.04
		CentOS 7.1 (RDO)	CentOS 6.5, 6.6	CentOS 6.5, 7.1
	RHEL 7.0 (RHOS)			RHEL 6.5, 7.0
ESXi	vSphere 5.5 U2 (Mirantis OpenStack)	vSphere 5.5 U2 (Mirantis OpenStack, VIO)	vSphere 5.5 U1, 5.5 U2	vSphere 5.5 U2, 6.0
Hyper-V				Windows Server 2012 R2 UR4 (SCVMM 2012 R2 UR4)
XenServer			XenServer 6.2	

*OpenStack installation options: Manual, Mirantis Fuel, RDO PackStack, RedHat Foreman. For verified combinations please contact the Big Switch Sales Team at sales@bigswitch.com.



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

Copyright 2014 Big Switch Networks, Inc. All rights reserved. Big Switch Networks, Big Cloud Fabric, Big Tap, Switch Light OS, and Switch Light VX are trademarks or registered trademarks of Big Switch Networks, Inc. All other trademarks, service marks, registered marks or registered service marks are the property of their respective owners. Big Switch Networks assumes no responsibility for any inaccuracies in this document. Big Switch Networks reserves the right to change, modify, transfer or otherwise revise this publication without notice. BCF 3.0 Datasheet V1 August 2015.