Big Cloud Fabric P+V Solution Validation with OpenStack

2015

Introduction to CloudLabs

CloudLabs is an initiative within Ciii chartered with streamlining the efforts to service its customers' cloud related needs. With the emergence of scale out computing workloads for e-commerce, big data analytics and private cloud, Ciii customers are demanding integrated solutions encompassing rack scale hardware with integrated software stack. In many cases, the complete stack is a converged infrastructure that includes optimized hardware with software architectures that solve specific workload requirements around speed, performance, power and cost.

The trend of disaggregation and open source software proliferation has led to the emergence of new scale-out blueprints over the last decade that has dramatically changed the infrastructure landscape.

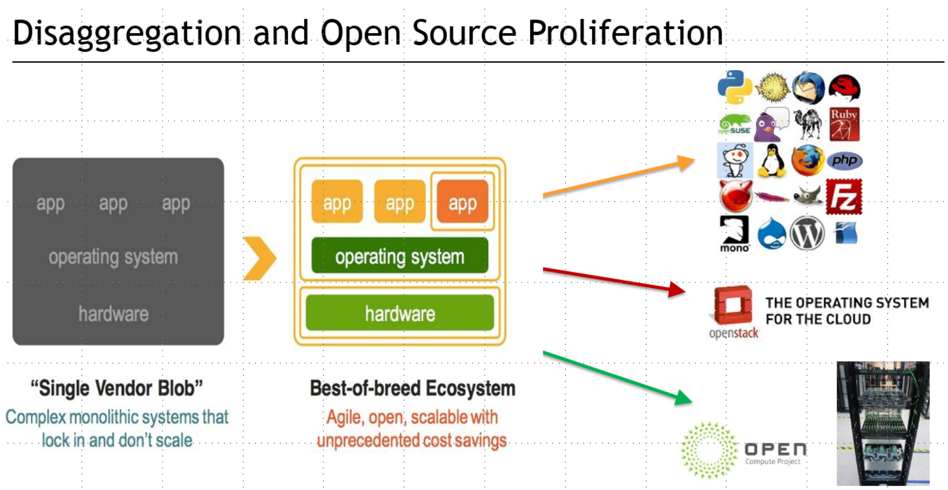

Figure 1: Disaggregation and Open Source Proliferation

CloudLabs provides engineering and design services to optimize such rack level solutions with emphasis on multivendor equipment integration. CloudLabs enables customers to accelerate a spectrum of cloud, converged infrastructure, and datacenter strategies. For this purpose, we have created a customer showcase staging area running a live multi-rack cloud instance on hardware developed within Flex. This staging lab facilitates our customers and partners to deploy latest technologies to allow for faster diffusion of technologies in a controlled environment for identifying value proposition for the end customer. It also provides our customers an access to a sand box environment that can be leveraged to test performance of their workloads and evaluate the solution capabilities.

SDN Trends and Solutions

Networking technology along with compute and storage constitute the core of the information technology. Traditionally the network control has been an integral part of the networking hardware such as switches and routers. However the constant push for disaggregation of software and hardware has given rise to Software Defined Networking (SDN) technology that centralizes network control, allowing the networking hardware to be simple and focused on data forwarding. As a result of this disaggregation, the networks become highly programmable with flexible and dynamic topologies. With the advent of cloud computing and scale out architectures, networking is clearly undergoing a paradigm shift with SDN.

Emerging SDN Deployment Models

Overlay/Underlay

- The physical networking devices are static, and all the automation is moved to the virtual switches
- Exemplified by Nicira/Vmware's Network Virtualization Platform (NVP), Juniper Contrail, and others
- Combination of overlay, underlay and gateways can make the overall solution complex and expensive

Unified Physical + Virtual model

- The physical networking devices are automated along with the virtual switches
- Big Switch Networks' (BSN) approach with Physical and virtual networks fully integrated and managed by the SDN controller.
- Expected to simplify deployment and operation while reducing costs

Hybrid (Overlay, OpenFlow, Other Protocols and APIs)

- Put forward by established networking players such as Cisco, and upstarts like Pica8
- Support for distributed control plane for network programmability

There are also layers of SDN as shown below each served by multiple vendors.

Layer	Examples	Benefits
Orchestration, Cloud Mgt.	OpenStack, cloud/REST APIs, CloudStack, vCloud Director	Automation, provisioning and orchestration
SDN apps, Network services	Layers 4–7 services, traffic engineering, network analytics, Firewall/Security Apps	Personalization, monetization, application intelligence and responsiveness.
Virtualization/ Control	OpenFlow/SDN controllers, vSwitches, Overlay Networks, tunneling protocols (VXLAN, NVGRE, STT), etc.	Advancements in routing controls/resource utilization, network flexibility and agility, and network management
Network Infrastructure	Vendor-specific OS for network devices (e.g., Cisco IOS), can be deployed as a Fabric, OpenFlow Tables reside	Scalability, operational simplicity and low latency

Figure 2: Layers of SDN

SDN technology curve in terms of functionality deployed is not uniform as expected. The current path towards realization of true SDN networks requires several intermediate steps as shown below, each with its own timeline of its technology maturity.

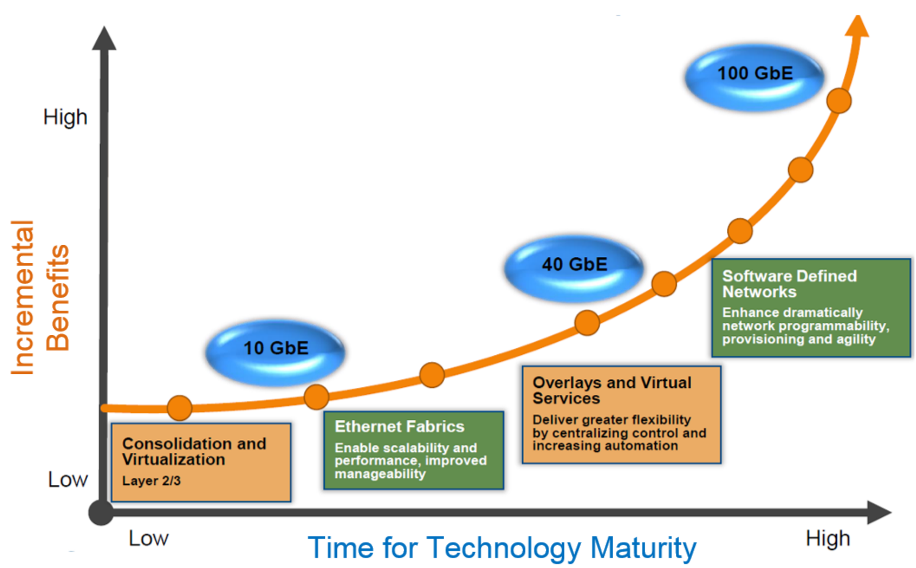

Figure 3: SDN Technology Curve

As explained before, multiple vendors and technology steps are required to on-board SDN technology. As a result of this, a collaborative effort is required across ecosystem of partners who are involved in delivering SDN HW+SW stack to the customer.

This white paper highlights one such effort of deploying and characterizing an SDN solution provided by Big Switch Networks.

Big Switch Networks Background

Big Switch Networks is bringing hyperscale data center networking technologies to a broader audience. Using bare metal and open Ethernet switch hardware, sophisticated SDN control software, and modern data center designs—Big Switch delivers fit-for-purpose data center networking solutions designed for enterprises, cloud providers and service providers.

Big Switch was among the very early startups (alongside Nicira) that started working on SDN technology. The company is widely considered to be one of the original SDN pioneers and has one of the industry's broadest portfolios in enabling SDN technologies.

Today Big Switch Networks is taking hyperscale networking design principles and applying them to its products such as Big Monitoring Fabric, an entry level solution to monitor existing networks and its flagship Big Cloud Fabric (BCF), the most advanced bare metal SDN switching fabric intended for new data center pods such as private cloud, big data and virtual desktop infrastructure (VDI).

Working alongside the technical marketing team of Big Switch Networks provides CloudLabs with a unique opportunity to bring a compelling and robust SDN stack to showcase for our customers.

CloudLabs and BSN Partnership Objective

The main objective of the partnership between Big Switch Networks and CloudLabs was to work together on deploying and verifying several new features of the flagship Big Cloud Fabric product BCF-3.0.0 as well as provide feedback on the overall user experience for the BCF. Big Switch provided the requisite hardware including the BCF controller nodes as well the network switch nodes. CloudLabs provided the compute nodes, the lab facility, and acted as an independent customer validation site by testing the features of BCF-3.0.0 exhaustively.

Following are the key features that were tested against various use cases:-

- Hardware Setup and Fabric Deployment
- BCF P+V and OpenStack Integration using Fuel 6.1
- Deploying application workloads using OpenStack
- Intra-tenant workload connectivity
- Inter-tenant workload connectivity (including tenant isolation)
- Fabric Resiliency

- Floating IP
- Network Throughput tests
- Router Grid Tests
- Upgrading the Cluster
- Network Test Path
- GUI Tests

CloudLabs' goal was to perform an exhaustive test effort over 3 months. Further, all the bugs, issues and user-experience feedback were reported back to Big Switch adding value to the partnership. The following section provides project details around the partnership and key feature validation effort.

Project Overview

Project Definition

The project definition phase included defining the scope, timeline, and hardware requirements for the project. The hardware setup in CloudLabs was based on a logical 2-rack topology recommended by Big Switch Networks using six Accton switches - two AS6700-32X spine switches and four AS5710-54X leaf switches. In addition, the cluster included 2 BCF controller nodes and 6 Flextronics servers for OpenStack cloud deployment.

Project Execution

The BCF-3.0.0 P+V solution validation project was carried out in multiple phases.

Initial deployment with BCF 3.0.0 Beta1 software release

This first phase used BCF Beta1 version of the software using Packstack tool for OpenStack deployment. After integrating BCF P+V Edition with OpenStack, and deploying Switch Light Virtual, various tests were conducted using OpenStack Horizon GUI, BCF GUI and BCF CLI. This phase also involved constant interaction with the BSN team and discussions on product features and some of the limitations encountered during testing.

Upgrade to BCF 3.0.0 Beta2 release

The second phase of the project involved upgrading BCF software to Beta2 release using the same hardware setup. Again, Packstack tool was used for deploying OpenStack. After successfully deploying the P+V solution, many of the test cases were repeated in this phase to verify the functionality and the bug fixes from Beta1. This phase of testing also included many of the disruptive test cases, such as headless operation, switch failure etc.

Upgrade to BCF 3.0.0 GA release and Mirantis Fuel 6.1 Integration

The focus of this phase was to upgrade the cluster to BCF 3.0.0 GA release and to deploy BCF P+V solution using Mirantis Fuel 6.1. Deploying standalone OpenStack cluster using Fuel 6.1, deploying Switch Light Virtual on the compute nodes, testing the functionality of the P+V solution, and testing the BCF GUI features were the key steps during this phase.

Lab Environment

The figure below shows the 2-rack topology deployed in CloudLabs using the BCF 3.0.0 P+V OpenStack integrated solution.

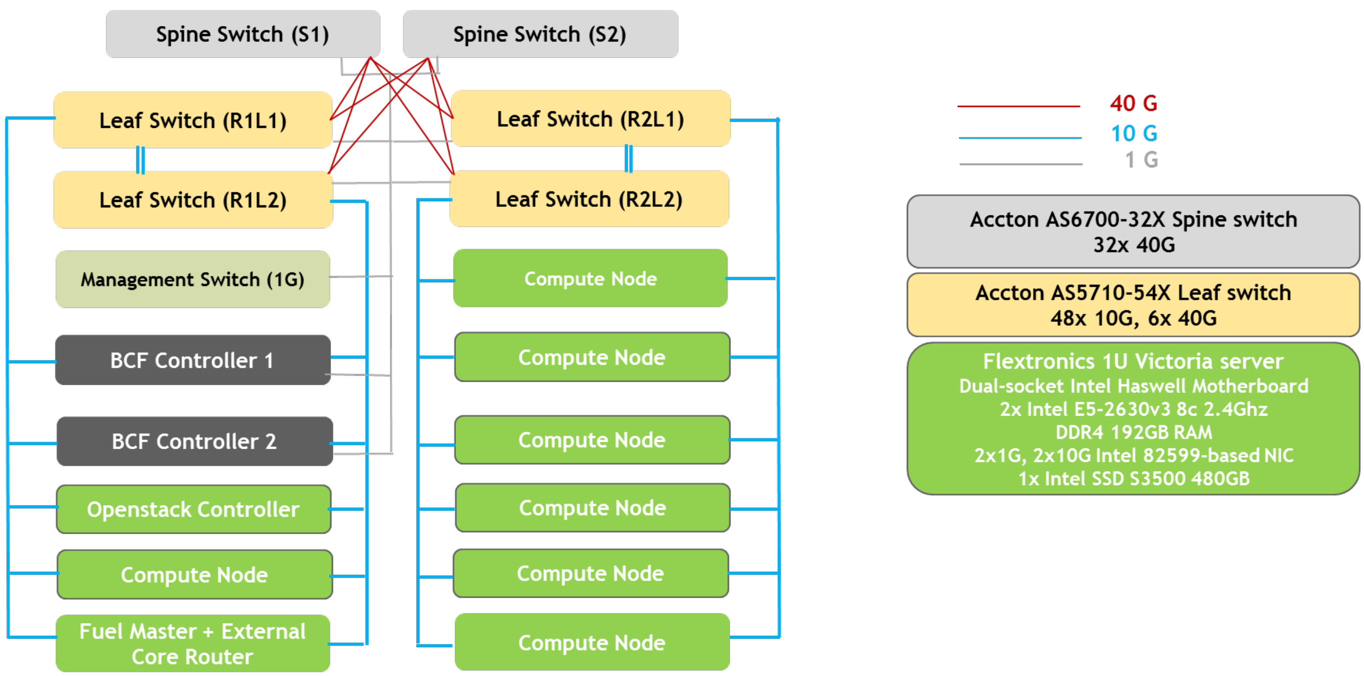

Figure 4: BCF 3.0.0 P+V OpenStack integrated solution in CloudLabs

The first rack has all the key control components for the OpenStack cluster, including OpenStack Controller, BCF SDN controllers (active and standby), and a general purpose gateway sever for hosting the Fuel master VM and providing external core router functionality for the OpenStack VMs. The internal OpenStack networks setup in the cluster follow the guidelines outlined in the Fuel 6.1 deployment guide. The figure below describes the logical network topology of these internal OpenStack networks. The first rack can also host compute nodes. The second rack in the CloudLabs 2-rack topology is a pure compute node. Most of the compute nodes used on this rack are the powerful next generation S5140 1U servers, codenamed Victoria.

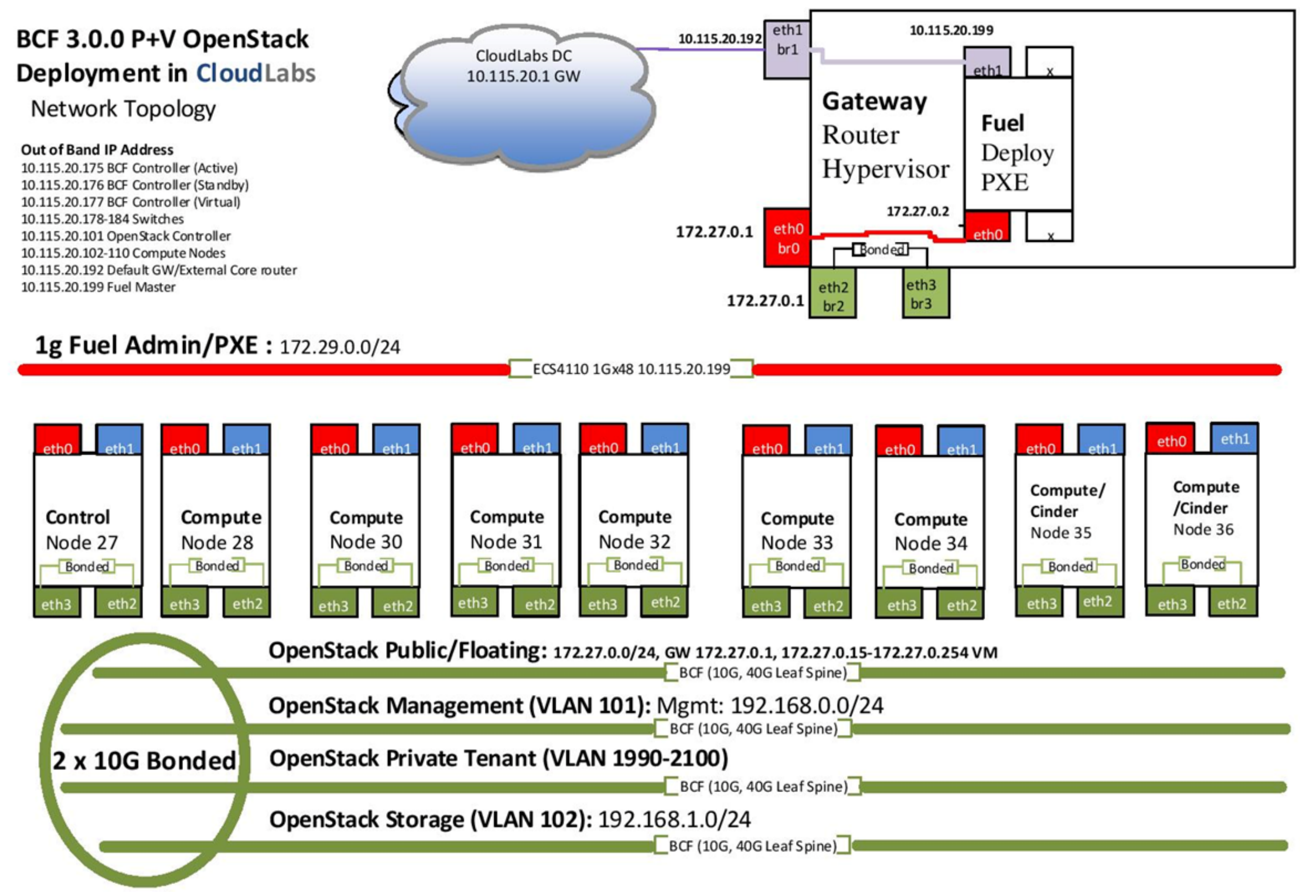

Figure 5: OpenStack Network Topology in CloudLabs

CLOS Network Topology

The network topology deployed at CloudLabs, shown in the figure below, is a leaf-spine topology, where all devices are exactly the same number of segments away from each other, and contain a predictable and consistent amount of latency. Every physical leaf switch is interconnected with every spine switch (full-mesh). Leaf switches in the same rack are connected together with two links (10-GbE ports) to create a leaf group. Dual links increase the available bandwidth and provide link-level redundancy. 10-GbE dual-bonded links from the compute nodes are connected to both ToR leaf switches in the same rack. 10-GbE dual-bonded links from the Big Cloud Fabric Controller nodes are connected to both ToR leaf switches in the same rack for vSwitch management.

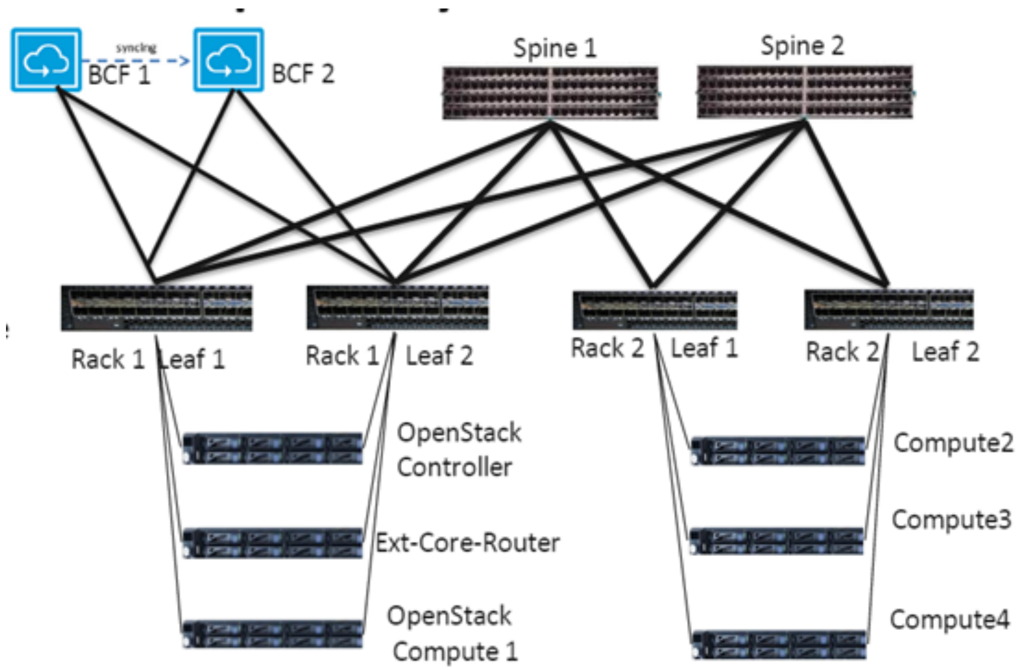

Figure 6: CLOS Network Topology

Flex S5140 1U Server

The S5140 (Victoria) is a 1U rack-mount 4- or 8-drive system designed to be a multi-purpose, high-performance rack server. The server leverages modular-design principles to provide a robust, highly-configurable and energy-efficient platform. The server has a tool-less design to simplify maintenance and servicing, and is equipped with 80Plus Platinum redundant power supplies.

Figure 7: Flex S5140 1U Server

Component	Quantity	Description
CPU	1 or 2	Includes one Intel® Xeon® E5-2600 v3 series CPU, a second CPU is optional
DIMMs	Up to 16	Up to 16 32 GB, 2133-MHz, DDR4 DIMMs
OS Drives	1 or 2	120 GB SSD, Intel DC S3500

Data Drives	Up to 8	Up to 8 2.5” SATA drives (900 GB Seagate ST900MM0006
PCIe disk controller card	1	LSI 8-port controller (9260-8i)
NIC ports	Up to 4	Included: 2-port 1 GbE NIC, or Optional: 2-port 10 GbE NIC
Power supply	2	750 W each; valid input voltage range is 100 to 240 VAC, and line frequency is 50 to 60 Hz
USB ports	2 or 4	2 USB 3.0 (rear), 2 USB 2.0 (front), 1 vertical Type A USB 2.0 (on motherboard), 1 dual port USB 2.0 header (cabled to front USB 2.0 connector)

Table 1:Flex S5140 1U Server Specs

Big Cloud Fabric Deployment and Configuration

The Big Cloud Fabric controller cluster is responsible for the control, management, and policy of the switching fabric. The switching fabric is implemented as a single management domain running Big Switch Networks' Switch Light operating system.

The first step in the deployment process was to install and configure the BCF controller software on the active and standby controllers. Then the switch MAC addresses and the roles (spine or leaf) were defined manually via the BCF controller CLI. When each switch was powered on, the switch's built-in Open Network Install Environment (ONIE) boot loader searched the network for boot servers. The BCF controller, acting as a boot server, responded and downloaded Big Switch Network's Switch Light operating system. The BCF fabric deployment and configuration steps are very simple and automated and significantly reduce the time required for the setup.

OpenStack Installation

The OpenStack installation for the BCF Beta software releases was done using PackStack and for the GA release of BCF, Fuel 6.1 was used.

PackStack based deployment

The procedure detailed in the BCF deployment guide was followed in CloudLabs for Packstack-based deployment. This deployment option first installs OpenStack in the standalone mode. Then, Big Switch OpenStack Installer (BOSI) script was run on the OpenStack controller node to install BSN plugins and Switch Light Virtual (VX) instances in the cluster to achieve full integration of the vSwitches into the Big Cloud Fabric.

Mirantis Fuel 6.1 based deployment

The deployment procedure detailed in the Mirantis Fuel 6.1 deployment guide and BCF 3.0.0 deployment guide were followed in CloudLabs. This involves deploying Big Cloud physical fabric with Mirantis Fuel and running the Big Switch BOSI script to install and configure Switch Light Virtual for OpenStack compute nodes.

BCF GUI Examples

BCF GUI provides an effective way to get a big picture view of the fabric, retrieve fabric analytics and also debug connectivity problems between intra-tenant and inter-tenant VMs in an OpenStack deployment. The following examples show screen-shots of the BCF GUI usage in the CloudLabs' 2-rack cluster.

BCF Fabric View

The Fabric view, shown in the figure below, displays the fabric topology. The active/standby controller pair is shown at the top, with green bars indicating healthy status. The health of each switch is displayed by red or green icons. Also shown is the state of each network interface, green for connected, and red for disconnected or error. All of the virtual switches running in the OpenStack environment are auto discovered and displayed.

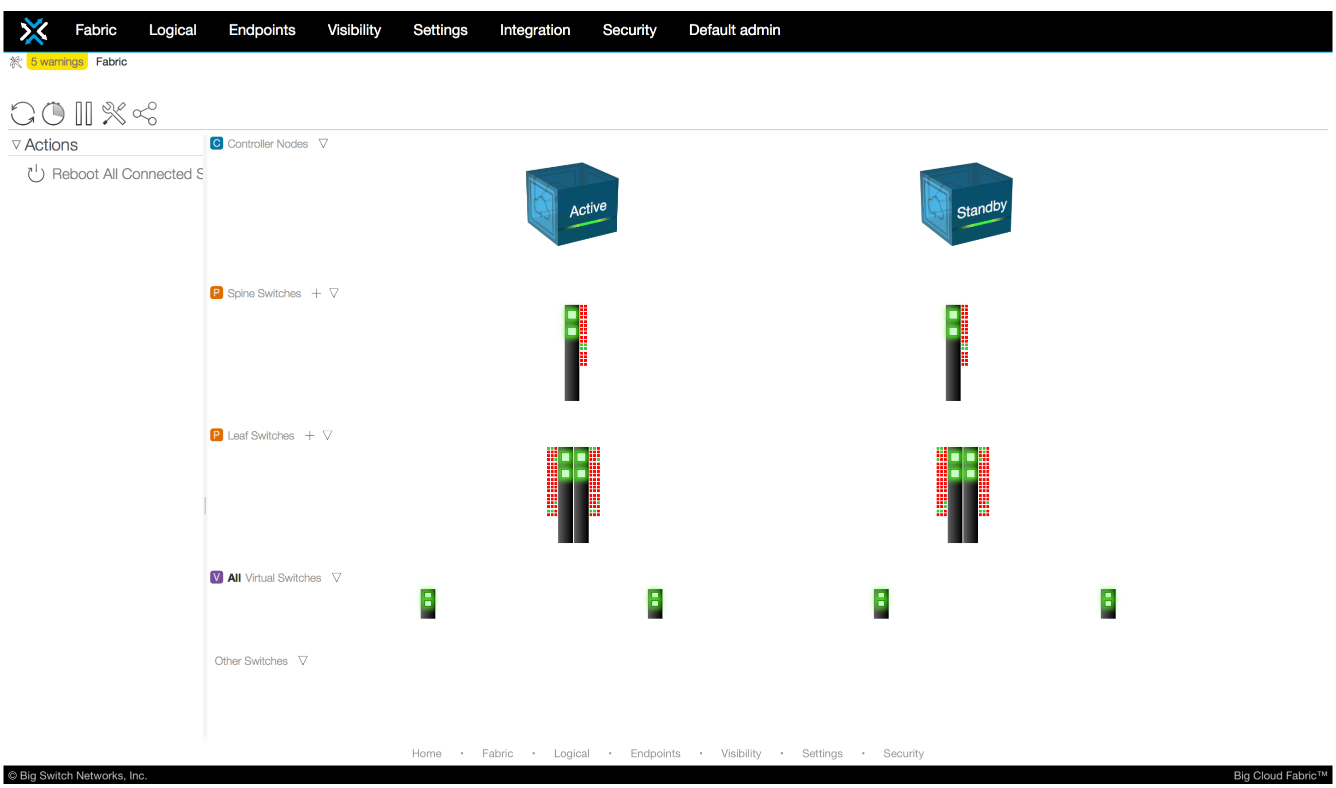

Figure 8: BCF Fabric View

Fabric Analytics

BCF provides full visibility across the physical and virtual fabrics for network statistics, analytics, and end-to-end troubleshooting. The analytics section provides a quick graphical snapshot of the state of the controller, showing memory, CPU, and disk utilization. The dashboard also displays information related to network activity, including the top talkers and top usage for tenants, segments, and interfaces.

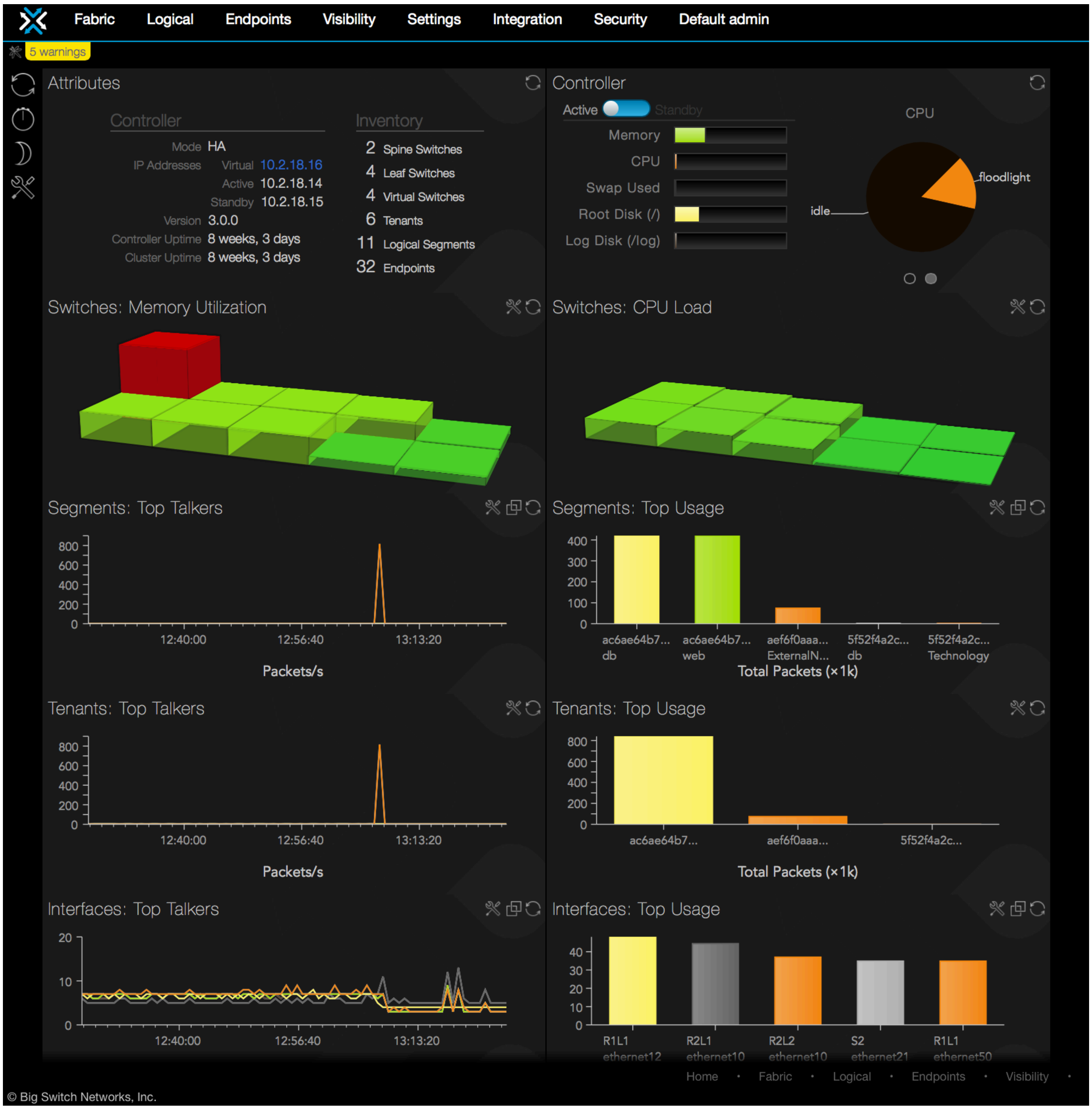

Figure 9: Big Cloud Fabric Dashboard

Trouble Shooting Network Traffic Path

When communication fails, either because of physical component failures or misconfigurations, diagnosing the root cause of the failure can be difficult—significant complexity is hidden behind the integration with orchestration software, and the tenant and segment abstractions. Big Cloud Fabric includes diagnostic tools to aid the system administrator with debugging.

To troubleshoot traffic forwarding problems in the data plane, we used the test path command, which determines if the controller is correctly programmed to forward packets from a given source endpoint to the specified destination. The Controller view is divided into three sections:

- Logical paths: This shows the path using the tenant and segment names as defined in the BCF.
- Physical paths: This shows the physical path using switch names.
- Summary: This states whether the path exists, lists any errors, and states the status of the path in the reverse direction.

Troubleshooting paths in this model extend from the physical fabric to the vSwitch, with end-to-end visibility of every physical and logical hop between any two pairs of endpoints.

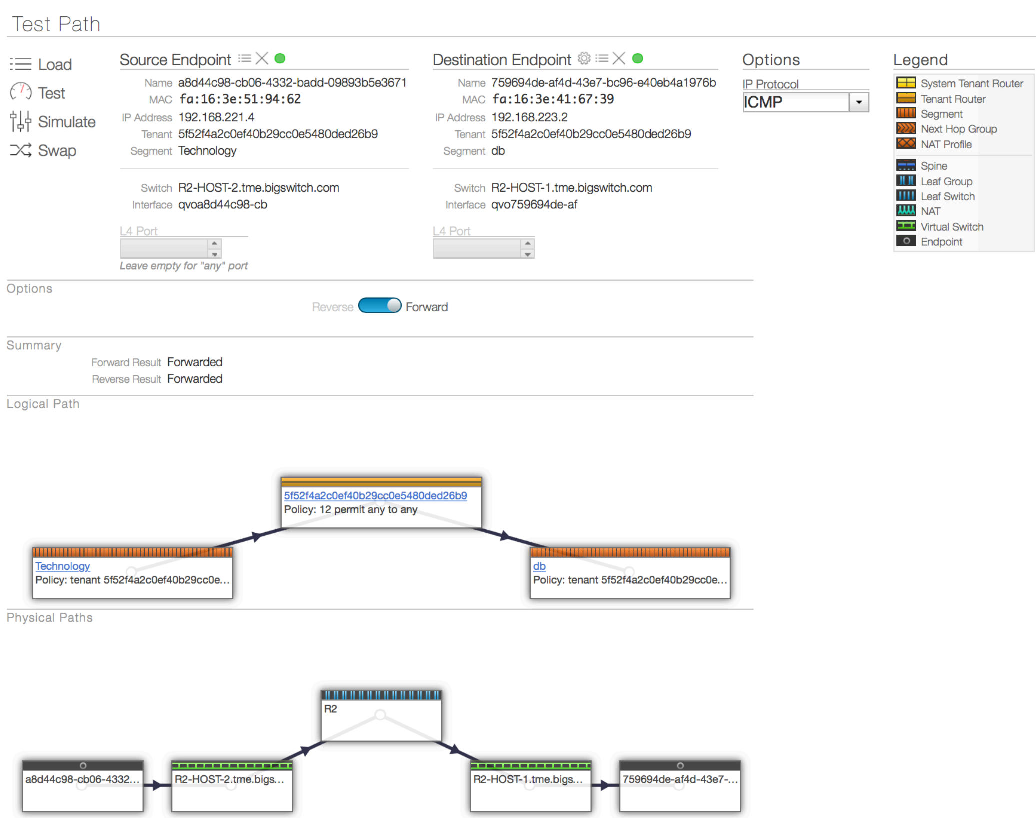

Figure 10: Network Traffic Path

Test Cases

The table below lists the main tests that were executed in CloudLabs after deploying the BCF P+V solution and integrating OpenStack.

SN	Group	Test
1	HW + Big Cloud Fabric deployment	Initial Setup of the servers and Controller Cluster Configuring the Active Controller with the First Boot Script Joining the Standby Controller to an Existing Cluster Verifying Cluster Configuration, Configuring the Cluster Virtual IP Installing and Configuring Fabric Switches with Zero Touch Networking
2	BCF P+V & Openstack Integration	Deploying OpenStack using Packstack [For Beta1 and Beta2] Deploying OpenStack using Mirantis Fuel 6.1 [for GA release]
		Deploying BSN Plugin and Switch Light Virtual (BOSI)
		Configuring External Connectivity (OpenStack ExternalNetwork)
3	Deploying application workloads using OpenStack	Automated L2/L3 provisioning of P+V fabric for multiple tenants via OpenStack 1) Running Heat Orchestration Template to create 3-tiered Web app topology 2) Creating tenants, networks, subnets via horizon
4	Intra-tenant workload connectivity	Intra-tenant workload connectivity configured using Security Groups in OpenStack
5	Inter-tenant workload connectivity (including tenant isolation)	Inter-tenant workload connectivity configured using provider ACLs using BCF Controller GUI/CLI
6	Fabric Resiliency	Application connectivity during negative events such as link flap or switch / controller failures. Also try Headless Mode (both controllers offline)
7	Floating IP	Floating IP - Tenant private network isolation from external network using floating IP
8	Network Throughput tests	1) iPerf Tests 2) IXIA tests
9	Router Grid Tests	1) Tests using the router grid 2) Tests using the router rules
10	Upgrading the Cluster	1) Add more Compute nodes to the cluster 2) Deploy additional compute nodes using Fuel 6.1 3) Deploy BCF plug-in and vSwitchLight on new compute nodes
11	Network Test Path using BCF GUI & Openstack	Test path for intra tenant connectivity on same compute node Test path for intra tenant connectivity on different compute node Test path for inter tenant connectivity on same compute node Test path for inter tenant connectivity on different compute node Test path for inter tenant connectivity using system router Delete an openstack tenant using fabric GUI Delete an openstack segment using fabric GUI New policy configuration using fabric GUI

Table 2: Test Cases

Partnership Impact

This project validated the ease of deployment and full functionality of the BCF P+V solution in a customer lab environment. Based on the deployment and testing done, CloudLabs provided detailed feedback to Big Switch Networks. This included observations, comments, questions, and problems encountered while deploying and using an early version of the solution. These were mainly focused on Fabric deployment, OpenStack integration, BCF CLI and GUI usage, BCF documentation, and overall user experience. This collaboration and the feedback provided helped BSN in making the product more robust and user-friendly. For CloudLabs, it is helpful in our overall objective of validating multi-vendor rack-level cloud solutions with Ciii designed servers. This partnership is also expected to contribute to faster go-to-market for BSN and continued collaboration between CloudLabs and BSN teams.


Ciii