

SWITCH Builds Agile, Simple, Reliable Data Center with Cumulus Linux and Routing on the Host

INDUSTRY: Higher Education Cloud + Hosting

BUSINESS OBJECTIVE: Agile, Simple and Reliable Solution

PARTNERS: Puppet, OpenStack, Ceph

Using Cumulus Networks-enhanced routing on the host allowed us to eliminate MLAG and spanning tree in our environment while still providing redundancy to the host. Cumulus Quagga's OSPF unnumbered gave us network agility, making it a core functionality for us. Deploying routing on the host with Cumulus Quagga improved our overall system availability while allowing simpler operation and troubleshooting.

— Saverio Proto: Cloud Operator, SWITCH

Company Overview

[SWITCH](#) was founded in 1987 as a non-profit organization and is the largest provider of IT services for universities and higher education in Switzerland. The SWITCH service portfolio includes:

- **SWITCHlan** — connects Swiss universities to other universities worldwide
- **SWITCHaai** — provides online access to learning
- **SWITCHconnect** — provides internet access to university campuses and elsewhere
- **SWITCHinteract** — provides a tool for video and web conferences
- **SWITCHcast** — provides a video management system for recording and streaming lectures
- **SWITCH-CERT** — protects education and research networks against cyber criminals and failures
- **SWITCHengines** — provides cloud IaaS to allow universities to extend their IT infrastructure

SWITCH is a thought leader and is motivated to try new, innovative technologies since they support research institutions. They also share ideas with the community.

Building SWITCHengines

SWITCHengines, an emerging part of SWITCH's portfolio, provides compute and storage resources for the higher education and research community. SWITCH needed to build a simple, reliable, cost effective and agile data center to support the SWITCHengines initiative and enable growth.

SWITCH required a solution that worked well with native OpenStack Neutron and Ceph deployment, allowed both IPv4 and IPv6 addressing, and had easy device configuration while deploying innovative open technologies without locking into a specific hardware platform.

SWITCH needed redundant connectivity to the VXLAN VTEPs on the hosts to support their multi-tenant OpenStack deployment. To meet their uptime Service Level Agreements (SLAs), they wanted to remove network complexity and cumbersome troubleshooting associated with Spanning Tree Protocol (STP) and Multi-chassis Link Aggregation protocol (MLAG), all while reducing the costs associated with deploying two leaf switches per rack.

Solution

After thorough research on architecture options, SWITCH decided on a Clos layer 3 fabric with fixed form factor switches for agility, which allowed for easy scalability. They evaluated many Open Networking software solutions and decided on Cumulus Linux, as it was the most mature and proven operating system. Routing directly on their Ubuntu OpenStack and Ceph hosts with the Clos architecture was considered their preferred option and was tested alongside other architectures to the host. Multiple designs and architectures including MLAG were examined and tested in their lab, but based on test results, SWITCH ultimately decided to run an open source routing protocol on the hosts, the Cumulus Networks-enhanced version of the Quagga package.

The criteria for choosing Cumulus Quagga on the host included:

- **Network Simplification:** Running Cumulus Quagga OSPF unnumbered directly on the host allows for the complete removal of MLAG and STP. Layer 3 throughout the data center is easier to troubleshoot using tools like ping and traceroute. Since both the switches and hosts are Linux, SWITCH uses Puppet to configure the hosts as well as the switches.
- **Network Flexibility and Agility:** Cumulus Quagga OSPF unnumbered provides the flexibility of deploying hosts in the data center without relation to link IP addresses, which enables easy migration and flexibility if needed. Upgrading to a higher bandwidth switch is easier with Routing on the Host since MLAG is not deployed and the uplinks are not bonded to each other — just let the routing protocol converge. Since SWITCH deploys only one leaf switch per rack, they need a solution that provides flexibility of cabling and rack layout since hosts may be connected to a leaf switch on a different rack for redundancy.
- **Network Reliability:** SWITCH deployed VXLAN with OpenStack for multi-tenancy, and the VXLAN Tunnel EndPoints (VTEPs) on the host hypervisors are advertised as a /32 address into the layer 3 data center. This architecture provides redundancy to the hosts via Equal Cost Multipath Routing (ECMP) without deploying MLAG or STP. The cables and hosts are moved at will and connectivity can be restored due to OSPF convergence.

Routing on the Host with Cumulus Quagga deployed at SWITCH enables an affordable, resilient, agile, and scalable network that is very easy to manage and troubleshoot. See Figure 1.

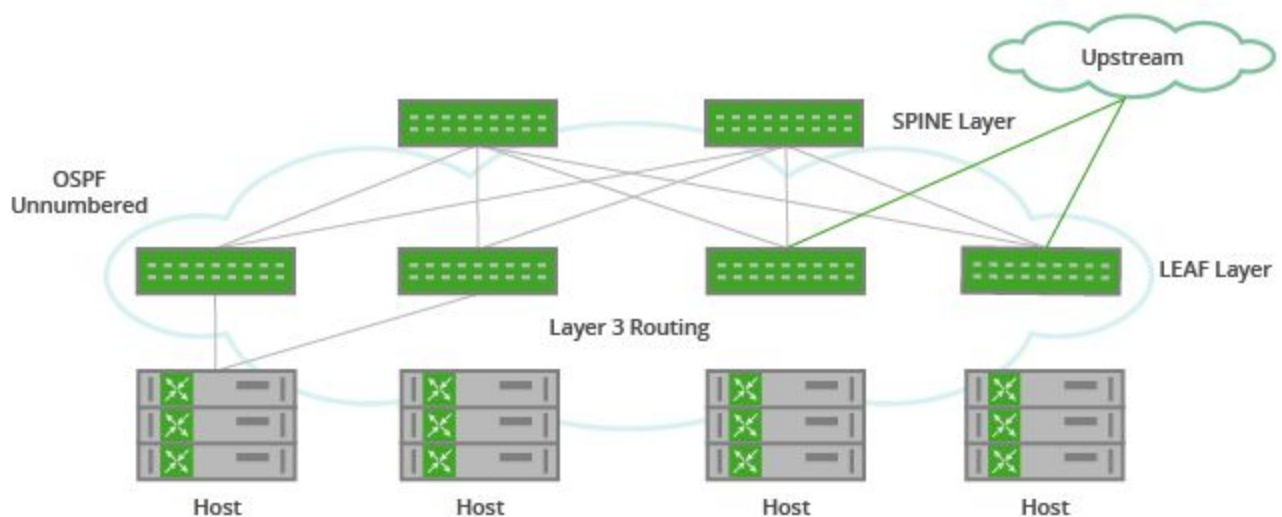


Figure 1 - SWITCH Data Center Architecture

Fancy proprietary multi-chassis L2 redundancy mechanisms often look magical and attractive. But when they fail—and I've seen several of them fail in various ways—and you have to debug them, they quickly stop being fun. With a proven L3 routing protocol like OSPF, we get fast routing around failures without any magic. And thanks to Cumulus' work on Quagga, handy extensions like unnumbered links are also available on servers now.

— Simon Leinen: Cloud Architect, SWITCH

Results and Recap

Deploying Routing on the Host with Cumulus Quagga allows SWITCH to operate a network that is:

Agile

A Clos network with one ToR switch requires agility to move and re-cable hosts for redundancy. Any host needs to be able to connect to any rack. Routing on the Host with Cumulus Quagga OSPF unnumbered provides this capability.

Simple

SWITCH's network is now very simple. A Clos design connecting all leaf switches to spine switches along with OSPF unnumbered means minimal configuration differences among the switches. Neither MLAG nor spanning tree is deployed anywhere in the network. This results in easier troubleshooting with tools such as ping and traceroute, and a smaller failure domain should a link or hardware fail.

Reliable

ECMP routing to the compute and storage hosts allows unlimited redundancy to the VTEPs without needing to deploy spanning tree or MLAG. During a failure scenario, OSPF converges and access to the host is quickly restored while it's connected to any switch in the data center. OSPF unnumbered also simplifies leaf switch hardware upgrades and re-cabling a host to any leaf switch on any rack without requiring configuration changes.

As seen above, Routing on the Host with Cumulus Quagga provides SWITCH with a simple, agile, reliable network that also has the ability to scale. They are now positioned to provide the vast compute power and storage solutions that the Swiss education and research community requires.

About Cumulus Networks

Cumulus Networks demystifies the complexity of networking and enables better, faster, easier networks to support your business. Our network operating system, Cumulus® Linux®, allows you to build and operate your network with the mindset of the web-scale pioneers like Google and Amazon, radically reducing the costs and complexities of modern data center networks. More than 400 organizations, including some of the largest-scale data center operations in the world, run Cumulus Linux. Cumulus Networks has received venture funding from Andreessen Horowitz, Battery Ventures, Sequoia Capital, Peter Wagner and four of the original VMware founders. For more information visit cumulusnetworks.com or follow [@cumulusnetworks](https://twitter.com/cumulusnetworks).

CUMULUS, the Cumulus Logo, CUMULUS NETWORKS, and the Rocket Turtle Logo (the "Marks") are trademarks and service marks of Cumulus Networks, Inc. in the U.S. and other countries. You are not permitted to use the Marks without the prior written consent of Cumulus Networks. The registered trademark Linux® is used pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis. All other marks are used under fair use or license from their respective owners.