Bare Metal Switches — Is There a Cost Benefit?

Overview

Network architects face many challenges today with the constant evolution and growing needs of data centers. They need to adapt while maintaining control of their budgets. Despite the fact that most parts of the data center have been automated, the network remains largely a manual operation, and traditional network architectures have been difficult to evolve.

IT staff have been demanding an automated network with resilient and open architectures. In this quest for the perfect solution, many new approaches have been proposed in the industry. But the one very simple approach that can solve these challenges today, and which has been proven in the server industry and in massively scalable data center networks, is to deploy bare metal switches with a Linux operating system on top of them.

Think about servers. Bare metal servers come with a boot loader so you can load the operating system of your choice on top of it, say Linux. You can now do the same on the network: bare metal switches can come with an install environment, ONIE (Open Network Install Environment), so you can load the network operating system of your choice on them, say a Linux OS, Cumulus Linux. You can now leverage Linux's widely deployed automation/orchestration tools for the network, its derived lower OpEx, and the cost savings of an open hardware and application ecosystem.

This paper focuses on the network infrastructure CapEx budget and cost savings when deploying bare metal switches, and looks at one scenario, a green-field deployment for new mid-size data center network PODs.

While the network infrastructure's CapEx considerations are very important as a driver for initial budgetary functions, CapEx represents only part of the savings realized with a bare metal switches solution. The major driver for the solution resides in the automation's OpEx savings.

Key findings:

1. Medium-size environment solutions for 10 racks and 200 nodes dual-attached (400 nodes single-attached) start under $100K
2. Medium-size environments can realize 3 times CapEx savings over incumbent solutions

Modern Data Center Architecture

While traditional architectures based on three-tier access/aggregation/core designs with fixed Layer 2 boxes at the access and modular aggregation/core boxes have been the de-facto standard for many years, the new traffic patterns, virtualization requirements, and growing traffic demands in the data center have prompted organizations to re-evaluate these architectures. Network architects have been demanding simple, resilient, agile, and scalable solutions. As a result, more and more data centers are turning to the scalable and distributed leaf/spine network design based on Layer 3 fabrics. This design itself is simple to troubleshoot, automate, scale, and evolve, and will be the design used as a reference for the purpose of this paper.
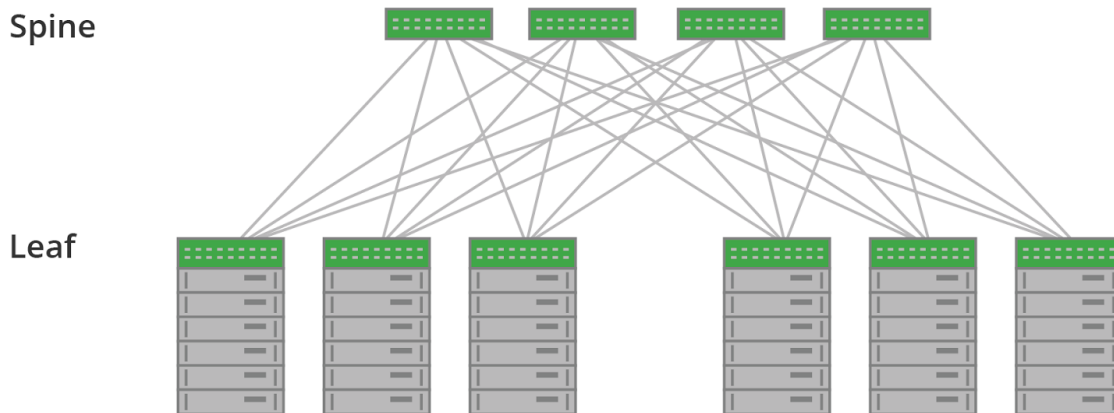
Although the figure represents a two-tier leaf/spine architecture, it is very easy to scale the model to three-tier leaf/spine/spine architectures as the number of servers increases; this architecture can scale to hundreds of thousands of servers.

New Data Center POD Cost Analysis

Customer environments vary from 16-40 servers per rack, either dual attached or single attached and they also vary in size. The design we chose is an example of a medium-size environment design with 10 racks and 200 servers, assuming 20 servers per rack, dual attached at 10G.

One can easily infer the price for similar solutions:

- 40 servers per rack solution, single attached: same price for 400 servers

- 40 servers per rack solution, dual attached: same price for a 200 servers, with half the number of racks

- 24 servers per rack solution, dual attached: same price for networking gear and interconnects for 240 servers (480 switch ports). Add a small delta for 40 additional cables at the ToR

The solution examines the following:

- Cost of hardware at leaves (industry-standard switches, 48x 1/10G + 6x 40G leaves)

- Cost of hardware at spines (industry-standard switches, 32x 40G spines)

- Cost of twinax at server access (industry-standard 3M twinax)

- Cost of optics interconnect between leaf/spine (industry-standard QSFP-SR4)

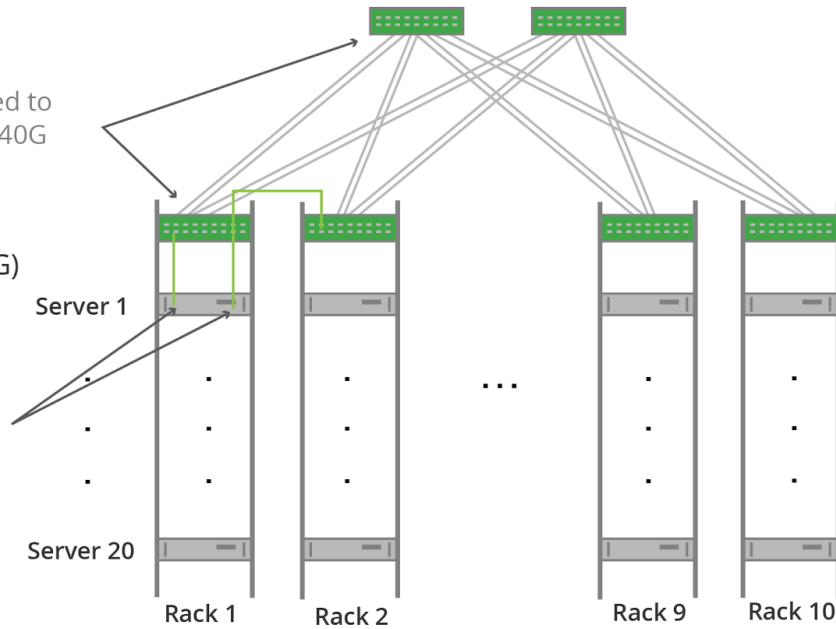- Default hardware partner warranty

Solution: 200 servers, 10 racks

Spines: 2x (32x40G)

Each Leaf interconnected to Spine with 4x40G (QSPF-SR4)

Leaves:
10x (48x10G + 6x40G)

Server 1

Each Server dual attached at 10G (3M Twinax)

Server 20

Rack 1   Rack 2   ...   Rack 9   Rack 10

BARE METAL VS LEGACY SWITCH COMPARISON			
CAPEX for 200 nodes in 10 racks	**Bare Metal Switch List Price***	**Cisco List Price****	**Arista List Price****
Leaves & Spines	$66,980	$432,000	$439,028
Cables & Interconnect	$29,199	$323,600	$202,000
Total	$96,179	$755,600	$641,028

Solution based on Quanta T3048-LY8 at the leaf and Quanta T5032-LY6 at the spines (Source Internet List price[i])
*** Cisco solution is a list price based on Nexus 9396PX at the leaves and Nexus 9332PQ at the spines (Source Internet List price - plug in your own discount for comparison)*
**** Arista solution is a list price based on Arista 7050SX-72 at the leaves and Arista 7050QX-32 (Source Internet List price - plug in your own discount for comparison)*

Summary

Organizations interested in evolving the network side of their business can extract infrastructure savings in a similar manner as the compute side by taking advantage of bare metal switch solutions to get a 3x savings over incumbent solutions. The savings can easily be reinvested in additional hardware and new innovations.

Savings associated with CapEx only scratches the surface when it comes to the benefits associated with a Linux OS. Beyond the savings in actual infrastructure costs, a significant chunk of the benefits lie in OpEx by having an automated network. About 80% of an engineer's time is spent manually configuring the network (Source: ISI Technology Research, Nov 5 2013). Where one needs to count on large teams for manual network configuration and troubleshooting, a simple, resilient network only needs a handful of DevOps staff for data center automation, and network architects can spend their leftover cycles designing efficient networks, providing differentiation, and investigating other innovative software defined network technologies.

Check your cost savings with the TCO calculator at cumulusnetworks/tco!

Appendix

Methodology used in the solution pricing:

- Using end-customer price for bare metal switches
- Using publicly available list prices for Cisco and Arista, use your own discount for comparison
- Solution comprises 10 leaf switches (includes L3 license as CapEx where applicable), 2 spine switches (includes L3 license as CapEx where applicable), 400 3M twinax cables, 4 QSFP-SR4 optics per leaf (80 total)

[i] *Source - Various prices available from neobits.com, digikey.com, fiberstore.com, sfpcables.com, bm-switch.com*

For more information, check out these resources:

- TCO/CapEx savings calculator at cumulusnetworks.com/tco
- Cumulus Workbench remote lab at cumulusnetworks.com/cw
- Product information at cumulusnetworks.com/product/overview/

About Cumulus Networks

Unleash the power of Open Networking with Cumulus Networks. Founded by veteran networking engineers from Cisco and VMware, Cumulus Networks makes the first Linux operating system for networking hardware and fills a critical gap in realizing the true promise of the software-defined data center. Just as Linux completely transformed the economics and innovation on the server side of the data center, Cumulus Linux is doing the same for the network. It is radically reducing the costs and complexities of operating modern data center networks for service providers and businesses of all sizes. 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.