

April 20, 2010

Arista Scales Up Ethernet Portfolio with Big Switch

by Michael Feldman, HPCwire Editor

The axis of datacenter networks is shifting and for companies like Arista Networks and other high performance 10 Gigabit Ethernet (GbE) switch vendors, this is opportunity knocking. The North-South network traffic pattern of client-server computing that dominated the industry for so long is giving way to high performance East-West networks where server-to-server communication is paramount. And it's the rise of the HPC and cloud computing markets that are helping to drive this realignment.

On Monday at the **High Performance Computing Linux Financial Markets conference** in New York, Ethernet switch vendor Arista Networks announced its first modular switch for the datacenter. The Arista 7500 provides 384 wirespeed L2/L3 10GbE ports in a compact 11RU chassis that the company is touting as "the highest throughput 10 Gigabit Ethernet switch in the industry." With an aggregate switching capacity of 10 Tbps, a packet memory of 18 GB, a port-to-port latency of 4 to 10 microseconds, and the ability to sling 5.7 billion packets per second, the 7500 is certainly a switch to be reckoned with.

According to Anshul Sadana, Arista's VP of Systems Engineering and a Cisco alum who joined the company in 2007, they've leapfrogged their Ethernet competition. "Even their roadmap slide decks haven't caught up with what we're delivering to the market," he says. In fact, Arista is offering 40 and 100 GbE upgradeability on today's 7500 platform. So when the rest of the ecosystem catches up, a fully tricked out 7500 switch could be transformed into a 48-port 100GbE box.

Until the 7500 was launched this week, the company portfolio was restricted to its 24- and 48-port rack switches -- the 7048, the 7100 T and 7100 S -- that sit at the leaf layer of the network. The new 7500 is aimed at the spine layer -- the super-highway that aggregates the network traffic from the compute and storage leaf switches. Like all of Arista's gear, the 7500 is designed for datacenter applications that require a low latency, non-blocking network fabric. In the past, this type of setup was confined mostly to HPC, but today, with the build-out of cloud computing infrastructure, it's on its way to becoming a much more widespread networking model.

Arista is one of those companies that is riding the GbE to 10GbE industry transition. After less than three years from its launch, Arista has already managed to accumulate over 300 customers spread across more than 25 countries. Sadana says the company is not revealing revenue figures at this point, but claims they are "growing exponentially."

Currently, about 30 percent of their customers are financial institutions within the high frequency trading space, where Arista's ultra-low-latency 7100 S switches are a nice fit for applications that need to execute split-second trades. Although none of these customers are admitting to using Arista gear, Sadana says you'll see plenty of their switches in just about any co-location facility next to the major stock exchanges. According to him, some of these

deployments are in the 100-switch range, deployed across multiple sites. "I would say we're the number one vendor in that space today," says Sadana.

Arista's other customers are spread across the oil and gas industry, universities, US DOE labs, Web 2.0 datacenters, and cloud/hosting providers. Specific customers -- at least ones who will publicly admit to owing Arista gear -- include the National Center for Atmospheric Research (NCAR), SARA in Europe, the University of Colorado, the BBC, Lawrence Livermore National Laboratories, and the San Diego Supercomputing Center, among others

Some of the first customers of the 7500 switch include the Howard Hughes Medical Institute, the San Diego Super Computer Center (SDSC) and a yet-to-be named DOE lab. Early testing on a 504-node (4,032-core) cluster employing an Arista 7508 switch and 7148SX top-of-rack switches yielded a Linpack performance mark of 35.8 teraflops. That represents 84 percent of the machine's peak performance, which is on par with the Linpack efficiency of InfiniBand-connected clusters. Only proprietary interconnects, such the one in Japan's Earth Simulator supercomputer can do much better -- 93 percent efficiency, in this case.

In fact, Arista switching pretty much matches InfiniBand performance and utility in most areas except perhaps raw bandwidth. But compared to traditional Ethernet aggregation switches, the 7500 boasts much better performance, energy efficiency, and price-performance. For example, Cisco's Nexus 7010 switch can only support 64 10GbE ports at wirespeed and just under 1 billion packets per second. And at 21RU, the Cisco switch is nearly twice as large as the Arista chassis.

A fully configured 7500 is listed at \$460,000, which works out to around \$1200 per port. That's 1/30 the cost of its Cisco switch counterpart. Furthermore, since the Nexus box consumes 139 watts per wirespeed 10GbE port, compared to 13.2 watts for the Arista gear, energy costs for the 7500 are going to be about 1/10 that of the 7010.

To be fair, the Nexus 7010 incorporates a lot more routing capability and other features that makes it a more capable end-to-end switch for the datacenter. But that was part of Arista's calculation. By focusing on the spine layer of the network, the company could maximize aggregate bandwidth and minimize latency, while beating the competition significantly in cost and power consumption.

In the short term, more formidable competition may come from InfiniBand vendors -- and not just from InfiniBand switches. Last year, Voltaire **launched** the Vantage 8500, an Ethernet core switch with a capacity of 11.5 Tbps that provide up to 288 10GbE wirespeed ports in a 15U chassis. The 8500 is an L2 core switch, so presumably it lacks some of the L3 routing smarts of the Arista 7500. But since Voltaire employs cut-through rather than the store-and-forward switching of the Arista 7500, the 8500 is able to keep latencies in the 1 microsecond range. (On the other hand, store-and-forward has some performance advantages in the spine layer, especially for storage clusters and TCP traffic.) Power consumption and pricing per port for the 8500 are comparable to the Arista solution. On Monday, BLADE Network Technologies and Voltaire **announced** they are combining forces to offer a lossless converged Ethernet fabric solution based on the 8500 and BLADE's G8124 24-port top-of-rack switch. The BLADE-Voltaire solution is aimed applications that rely on the same type of low latency, non-blocking networks that Arista is targeting, namely HPC and cloud-scale apps. The switch combo is packaged with Voltaire's Unified Fabric Manager to manage switch resources, and is designed to scale up to 3,400 10GbE ports. The

solution also comes with BLADE's VMready (for virtualized environments) and Voltaire Messaging Accelerator software (for high frequency trading applications).

Although InfiniBand is still the interconnect of choice where the lowest latency and highest bandwidth are needed, solutions like those from BLADE, Voltaire, and Arista will make it easier for HPC, high frequency trading, and ultra-scale datacenter applications to transition from GbE to 10GbE networks. And with 40 GbE and 100 GbE just over the horizon, the protocol will be with us for the foreseeable future.

Links referenced within this article

High Performance Computing Linux Financial Markets conference

<http://www.flaggmt.com/linux>

launched

<http://www.hpcwire.com/offthewire/Voltaire-Announces-Vantage-8500-Ethernet-Switch-64998847.html>

announced

<http://www.hpcwire.com/offthewire/BLADE-Voltaire-Partner-on-Highest-Density-10GbE-Fabric-Solution-91527154.html>

Find this article at:

<http://www.hpcwire.com/features/Arista-Scales-Up-Ethernet-Portfolio-with-Big-Switch.html?viewAll=y>

<http://www.aristanetworks.com/>

<http://www.futurea.fi/>