

Data Center Class

Scaling Data Center Networks

INSIDE

SCALE

The ability for network architectures to accept increased traffic or number of devices without impacting the cost per port

WHY

Today demanding more ports than any switch can support without oversubscription takes six-times the number of network elements

WHO CARES

IT professionals who need to scale performance while maximizing efficiency. IT leaders who are passionate about technology, yet realistic about the business.

WHAT IS NEXT

Network technologies are being disrupted by end-user and business leader demands for lower-cost, consistent uptime, better SCALE.

Virtualization, Cloud Computing, and 10Gb/40Gb/100Gb technologies are all disruptive elements that cause architectural change. A new design is required.

What is Arista's definition of scalability in the context of a data center network? What are the critical requirements and what issues must be solved to address data center scalability?

Scalability of the data center network is the ability to construct and expand a network with simple, repeatable designs that can accommodate increased traffic or new devices without impacting applications, workflows or the cost per port.

Scaling needs be linear for both performance and for cost – i.e., as a system supports more nodes, the price per host and the performance per host stays very close to the same.

Achieving this challenging goal in the data center network has several components:

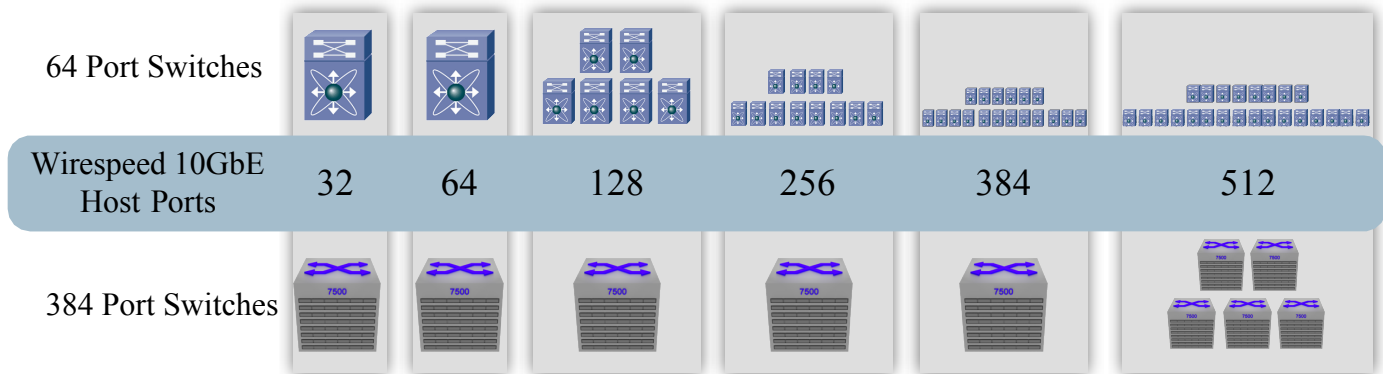
- Scaling up the performance of single systems, as this allows more price/performance linearity
- Extending or stretching the addressing of a network using open standards to embrace workload mobility across increased geographic distances and larger data center facilities
- Scaling the manageability and operational capabilities of the network infrastructure in the data center network so that administrators can effectively provision and manage a larger number of systems. Additionally provide increased visibility to the network, and all of the devices that are required to deliver IT services. This includes integration with automation systems via open and extensible APIs.
- Topologies must continue to self-organize and self-heal while failing over quickly in the event of link or nodal failure. Technologies that make all bandwidth available, reduce the dependency on the artificial insertion of loops into the network, and enable integration of topology change with planned operational procedures are part of Arista's Extensible Operating System - EOS.

SUB-LINEAR SCALABILITY

64-PORT 10GBE SWITCH VERSUS 384-PORT 10GBE SWITCH EXAMPLE

A wire-rate 64-port 10Gb Ethernet switch can linearly scale cost up to 64 ports. After 64 ports the cost per end-device port is severely impacted as it takes a second stage of switches and four access switches to deliver 128 wire-rate interfaces with constant bi-sectional bandwidth, and at 512 ports the wire-rate network needs 16 Access Layer Switches and 8 Aggregation Switches to meet the performance requirements. The below example contrasts the number of devices needed to scale out an Ethernet network contrasting 64 and 384-port devices.

Arista's Cloud Network architecture is designed to deliver far more linear scalability to hundreds of 10Gb Ethernet interfaces and thousands of Gb Ethernet interfaces



The critical requirements of data center network scalability are:

- A portfolio of high-performance switching platforms with a consistent operating model across platforms
- Ability to provision and manage multiple devices as a single system without trading off functionality to scale operations
- Ability to construct high-performance multi-system Leaf/Spine topologies using open standards without reducing the capacities of each system, introducing blocking architectures or systems with insufficient buffers to handle large numbers of TCP flows
- Ability to linearly scale to a larger number of high-performance interfaces that map cleanly to a servers ability to generate I/O and a storage systems ability to consume/source data

The technology innovations required to deliver on the promise of linear scalability for data center networks with increased systems densities, high capacity HVAC environments, virtualized and multi-tenant environments fall into several

key areas:

- Cost effective host and network Ethernet interfaces that match the CPU PCIe speeds - i.e. 10Gb and 40Gb Ethernet
- 40/100Gb Ethernet interfaces for uplinks that can be deployed and aggregated at a density that preserves clean network designs
- Efficient, stable and standards-based link aggregation technologies to scale current 10Gb Ethernet uplinks
- Scalable network operating system architecture that has open, well defined, programmable and machine readable interfaces, other than CLI, for provisioning of network services
- Trusted segmentation mechanisms that are auto-negotiated between hosts and networks and then carried across the network fabric
- Scalable solutions for deploying global network policies and the ability for these to be set and coalesced with external management systems from multiple-vendors

Scaling Metrics For Two Tier Networks

Leaf Switch	Spine Switch	Host Ports MLAG	Host Ports ECMP (n-way)	End-to-End Latency (μ s)	Total RU MLAG/ECMP	Total Power (KW) MLAG/ECMP
7048	7050S	1,440	3,072 (4)	10	32/68	5.4/11.6
7048	7508	9,120	18,432 (4)	14	212/428	40.6/82.0
7050S	7050S	336	3,072 (16)	2.5	9/80	1.1/10.0
7050S	7508	2,256	18,432 (16)	6.1	69/560	13.5/108.8

What are the relative economics and costs/benefits of Arista's technologies?

Arista aims to have the best price/performance ratio from both a CapEx and OpEx perspective in the data center switching market - thus the best linear scalability.

Using only 1RU switches, Arista can scale networks up to 64 racks (more than 3,000 servers) in a Leaf/Spine architecture, while saving space, critical in a co-location environment where space consumed can be directly measured as operations expense. For larger networks, the Arista 7500 provides scalability up to 384 racks (more than 18,000 servers).

Open standards play a critical role in building scalable networks that preserve vendor choice. Arista supports standards-based technologies to build scalable Layer 2 and Layer 3 networks, such as Multi-Chassis Link Aggregation (MLAG) and ECMP. Arista has been a leader in delivering VXLAN, which extends Layer 2 domains across Layer 3 boundaries while eliminating the 4K VLAN limitation.

When deploying and expanding data center networks, provisioning and on-going configuration management is a major issue. Arista's Zero Touch Provisioning (ZTP) provides standards-based features to automate network and server provisioning. The programmability of Linux-based EOS makes it possible to leverage tools such as Chef and Puppet to manage the switch in a similar way to other Linux-

based systems.

What are the right metrics to measure my network scalability on?

Key metrics in the data center networking market are:

- Price per Wire-Rate Host 10GbE Port
- Price per Million Packets per Second (for L3 devices)
- Power Draw, Overall
- Power Draw per Wire-Rate 10GbE Port
- Power Conversion Efficiency
- Provisioning time for new switch
- Failover time for nodal failure
- Failover time for link failure
- Failover time for planned maintenance with default commands
- Failover time for software upgrade
- Disruption/loss during system boot/insertion
- Port-to-Port Latency in the data center (single-switch)
- Port-to-Port Latency in the data center (any port to any other, multi-system)

Arista's solutions are compelling on all above metrics. Arista provides full wirespeed switches optimizing price/performance. Wirespeed switches with larger port counts results in fewer switches in the data path, and a more scalable data center network.

Additional information and resources can be found at: www.arista.com/en/633/c