

End of Row and Middle of Row with Arista 7300 Series

Today's data center is a dynamic and rapidly changing landscape, with the adoption of server virtualization driving 10GbE attachment speeds at the access layer and application architectures powering an exponential growth in server-to-server communication. This increase in bandwidth and change in traffic patterns is placing an overwhelming challenge on traditional End of Row (EoR) or Middle of Row (MoR) designs where the placement of a Cisco Catalyst 6500 or similar switch is now becoming a bottleneck for adding scale, bandwidth and performance.

The Arista 7300 Series is the next generation of data center switch purpose built to allow the design of a high performance, high density EoR solution for the challenges virtualization brings to the data center today. Delivering the performance and scale, with a virtualization feature set for both green-field and brown-field sites, allowing investment protection for existing 1GbE copper cabling plants for high density 10GbE server attachment while seamlessly introducing next generation DC technologies like VXLAN and Network Tracers.

TRADITIONAL END OF ROW DATA CENTER DESIGNS

The term End of Row (EoR) design is derived from the fact that a dedicated networking rack or cabinet is placed at either end of a row of servers for the purpose of providing network connectivity to the servers within that row. Traditionally hosted within the networking racks at either end of the row for redundancy reasons, would be a Cisco Catalyst 6500 or similar chassis based switch.

To address the legacy requirement for 1GbE and even 100Mbps server attachment, each server cabinet in the row would have twisted pair copper cabling (typically Category 5e or 6/6A) routed through overhead cable trays or below a raised floor to the "End of Row" switch. To provide redundant connections these copper cables would typically be run to a pair of Catalyst 6500 switches, installed at either end of the row.

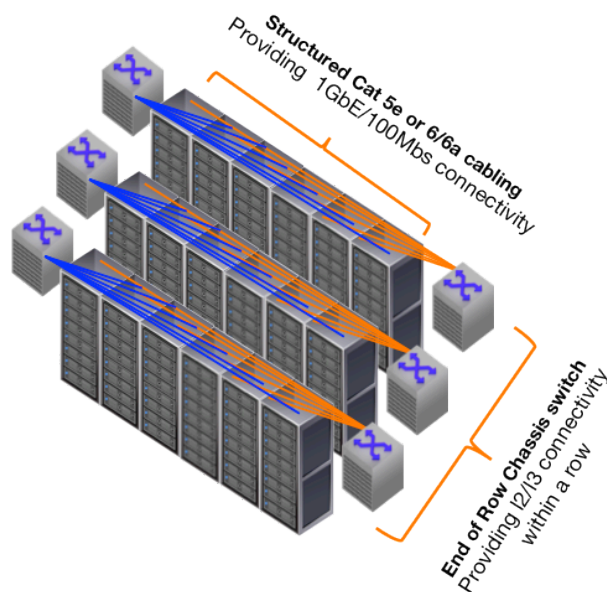


Figure 1: Traditional design with structured cabling for 1GbE

Generally, from a logical perspective the Catalyst 6500 switch, would act as the layer-2/layer-3 (L2/3) boundary for all servers within a row, with applications and individual business units placed in dedicated 802.1Q VLANs. The Catalyst 6500 switch, providing the L3 gateway for the associated subnets. For active-active L2 connectivity to the servers and increased server bandwidth the Cat 6500 switches are often deployed in a pseudo cluster or VSS configuration. The VSS configuration allows the two EoR switches to appear as one logical switch, with servers attaching via 2-port Link Aggregation Group (LAG). The ports of the LAG are split across the physical Catalyst 6500 switches for resiliency reasons providing a dual active and redundant set of connections.

THE CHALLENGES FOR TRADITIONAL EOR DESIGNS IN A VIRTUALIZED DATA CENTER

The EoR solution presents a number of advantages over 1G Top of Rack (ToR) design approaches in smaller environments. It allows collapsing of the access and distribution layers into a single tier of high-density chassis switches. This reduces the number of switches to purchase and manage within the design. In addition it provides improved performance by reducing the level of over-subscription, which would be present in a traditional multi-tier approach.

While offering these advantages the traditional adoption of the Cat 6500 in the architecture restricts its suitability and effectiveness for the changing landscape of today's data center, which now has to facilitate the rapid expansion of higher performance CPUs, server virtualization, flash storage, Big Data and cloud computing.

10GBE SERVER ATTACHMENT

The adoption of server virtualization allows the optimization of compute nodes, by enabling multiple virtual machines (VMs) to reside on a single physical server. A typical physical server can host anywhere from 8 to 16 VMs or even more on a single CPU core. This consolidation dramatically increases the bandwidth egressing a single physical server driving the demand for 10G server attachment. With a typical rack being deployed with anything from as few as 16 to 24 servers, each configured with dual-attached 10GbE ports, support for over 40 10GbE ports within a single rack is not an uncommon requirement. In an EoR design this in turn raises a requirement for high density 10GbE ports on the Catalyst 6500 switch. With the Catalyst 6500 being restricted to just 16 10GbE ports per line-card, which are in turn 4:1 oversubscribed, the scale of 10GbE server attachment and the level of virtualization that can be achieved with a single system is severely restricted.

COST EFFECTIVE 10GBE MIGRATION PATH

The purchasing of 10GbE enabled servers and the adoption of a virtualized strategy will often be an evolutionary process, with the inflection point occurring when the majority of servers are attached at 10G, forecast between 2013 and 2015 (figure 2). So while this migration will eventually drive the need for increased 10G density in the Catalyst 6500, the increase needs to be achieved with minimal effect on the remaining 100Mb and 1G server infrastructure, as both will need to co-exist for some time. In an EoR Catalyst 6500 design originally architected for a 1GbE server estate this need for a phased migration presents a costly and very complex operational challenge.

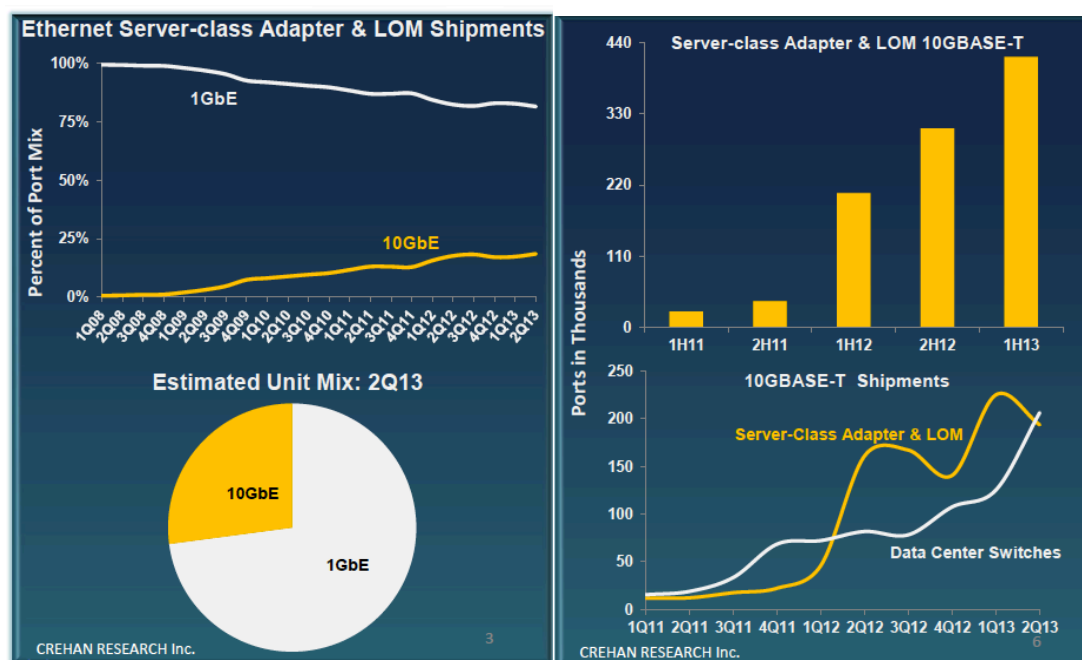


Figure 2: Outlining the projected 10G server adoption rates, and 10Gbase-T market percentage
Source: Crehan Research

To minimize disruption and cost it is preferable that any new 10GbE connections utilize the same structured cabling infrastructure for 1GbE server attachments. In an attempt to achieve this the Catalyst 6500 switch provides a 10G/1G RJ45 copper line-card, but only at a third of the port density (16-ports) of the 1GbE line-card (48-ports), and at a 4:1 oversubscription ratio. This reduction in port density limits the potential for any non-disruptive on-demand 10G migration plan, as the enablement of a single 10G server will reduce the number of 1G servers that can be supported by the infrastructure. A problem, which can only be solved by migrating multiple 1G servers simultaneously to ensure all connectivity, is retained. This is a complex process, which introduces costly and complex network and server engineering cycles to plan and deploy successfully.

BLADE CHASSIS AND NETWORK ATTACHED STORAGE

Standard 1RU/2RU rack servers are not the only server models deployed within the data center. To provide centralized management or to optimize rack space, blade chassis are often deployed. In the blade chassis architecture, a switch is commonly integrated into the chassis providing 1GbE or 10GbE connectivity to the servers and presenting multiple 10G ports upstream to the first hop or EoR switch. A further driver for 10G connectivity is network attached storage (iSCSI and NAS) which is growing at approximately twice the rate of Fiber Channel (FC) storage in the data center.

While these both drive 10G port density requirements within the EoR switch they are typically deployed using SFP+ (with Fiber or Twinax) attachments as opposed to UTP copper cabling. As a result there is a need for mixed copper and SFP+ connectivity within the EoR switches, and ideally without compromising the overall copper 1GbE and 10GbE density.

40GBE CONNECTIVITY

The migration to a 10G server infrastructure in turn drives the bandwidth requirements for upstream connectivity and the need for 40GbE ports. This is also true for any inter-switch links when the Catalyst 6500 are deployed in a VSS configuration. The Catalyst 6500 switch provides just 4 ports of 40GbE per line-card and they are limited in

performance - 2:1 oversubscribed. The introduction of the 40GbE line-card significantly reduces the number of 1GbE/10GbE servers the platform can support.

EAST-WEST TRAFFIC PATTERN GROWTH

In a virtualized data center, applications and services are not tied to a physical location, but rather spread across different racks within a row to ensure optimal usage of all compute capacity within the row. Application architectures are commonly multi-tiered (Web, Application, Database Server), load-balanced and distributed for scale and resilience. These demands result in an exponential increase in server-to-server or east to west traffic flows between servers in different racks and even across rows of the data center.

This dynamic location of servers and their inter-process communication (IPC) places performance and engineering challenges on the server networking infrastructure. The network needs to ensure consistent throughput, minimal over-subscription ratios and consistent latency regardless of where a server is placed and its specific traffic flow. Any perceived changes in performance based on connectivity, will restrict the operational flexibility as placement of a server has the potential to disrupt all compute resources within a row.

This sprawl of resources and increased server-to-server communication is often the driver for the adoption of an EoR design as the IPC communication can then be contained within a single switch. A single switch can potentially provide a non-blocking architecture for the server-to-server communication, rather than a 2:1/3:1 oversubscription ratio which would be common in a two-tier approach. The Catalyst 6500 switch however is a blocking architecture, with all line-cards 40G (2:1), 10G (4:1) and 1G (1.2:1) oversubscribed, thus the placement of servers can become severely restricted.

STRETCHED LAYER 2 DOMAINS

To restrict the scope of the Layer 2 (L2) broadcast domains the logical topology of the EoR design would typically have the Catalyst 6500 acting as the Layer-2/Layer-3 boundary for all 802.1Q VLANs created within the row. The switches can optionally be configured in a VSS pair to provide an active-active L3 gateway for redundancy, and upstream layer 3 connectivity for connecting to a distribution layer for both external communication and between neighboring rack rows.

In a virtualized environment this legacy design approach restricts the scope of mobility and placement of virtual machines, and storage clusters requiring Layer 2 connectivity, to within a row. As any communication between rows, is only possible via layer 3 routing.

ARISTA 7300 SERIES MODULAR SWITCH

The Arista 7300 Series introduces a new architecture for scalable 10GbE EoR deployments, providing up to 768 wire-speed 10GbE ports plus 64 40GbE wire-speed ports for uplink connectivity within a single chassis. The 7300 Series delivers performance, which is over 10X greater than the current Catalyst 6500 solutions and uses existing Cat5e and Cat6/6A structured cabling for seamless on-demand migration from 100Mb to 10GbE

The Arista 7300 Series is a fully redundant modular chassis platform, delivered in a 4-slot (8RU high), 8-slot (12RU high), or 16 slot (21RU) footprint. For consistent operation and reduced sparing the three chassis share common line cards, supervisor modules and power supplies. The chassis are optimized specifically for the data center environment with the PSU and fans delivering both front-to-back and back-to-front airflow options.

The 7300 Series supports three high performance wire-speed layer 2/3 line-cards for high-density 1G/10GbE and 40GbE deployments within the data center. All line-cards support same hardware architecture and as a result have the same consistent Layer2, Layer 3 and virtualization feature sets.

- **7300X-32Q:** A 32 port 40GbE QFSP+ module, providing wire-speed Layer 2/3 and advanced VXLAN functionality. Each 40G QFSP+ port is capable of being software provisioned as four individual 10GbE ports, to provide support for 128 line-rate 10GbE ports in a single line-card.
- **7300X-64T:** A 64 port 10GbE module, providing wire-speed Layer 2/3 and advanced VXLAN functionality. The first 48 ports of the line-card are 10G/1G/100Mb Ethernet copper ports, allowing 10GbE connectivity over Cat5e/6/6a cabling at distances up to 100m. In addition the module also provides four 40GbE QFSP+ ports, which can be individually configured to support a total of sixteen 10GbE ports.
- **7300X-64S:** A 64 port 10GbE module, providing wire-speed Layer 2/3 and advanced VXLAN functionality. The first 48 ports of the line-card are presented as 48 10G/1GbE SFP+ ports for fiber and Twinax connectivity. In addition the module also provides four 40GbE QFSP+ ports, which can be individually configured to support a total of sixteen 10GbE ports.

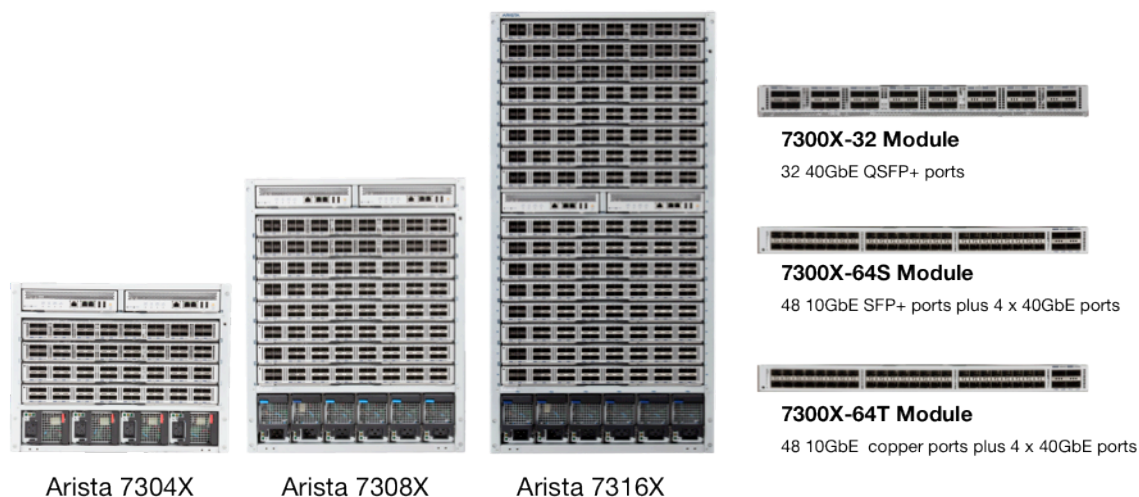


Figure 3: Arista 7300 Series and line cards

ARISTA 7300 SERIES SPLINE™ ARCHITECTURES

The throughput, port density and form factor of the 7300 Series platform are purpose built to allow existing 1GbE EoR designs to seamlessly migrate to next-generation 1/10G/40G virtualized data center solutions, while providing the required performance to facilitate any-to-any server communication. This next generation architectural approach for EoR deployments, Arista terms as a collapsed Spine/Leaf or Spline™.

In the Spline™ architecture the Arista 7300 Series seamlessly replaces a traditional Catalyst 6500 as the EoR switch. With the dense form factor of the 7300 Series the replacement can be achieved while reducing both the rack space and power requirements. The copper port density and triple speed support (100Mbps/1G/10G) of the 7300X-64T line module allowing the re-use of existing structured Cat5e/6/6A cabling to providing seamless on-demand support for 10G server attachment.

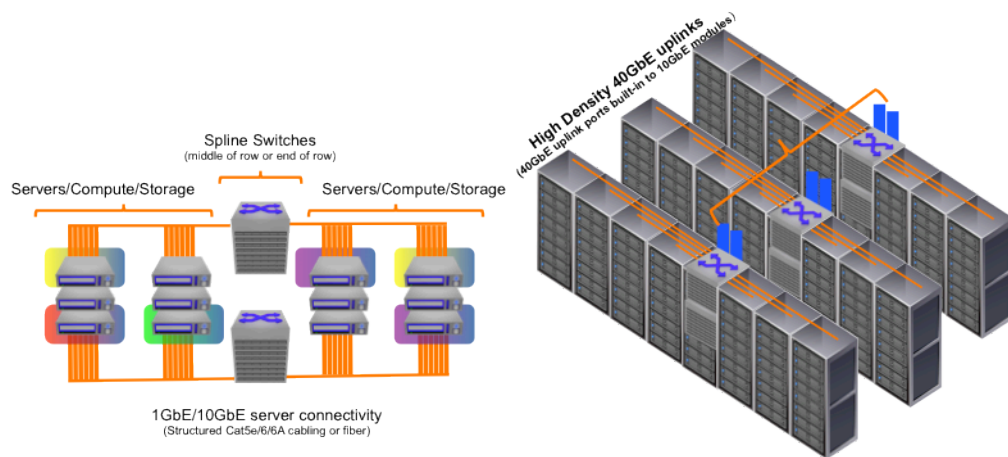


Figure 4: Arista Spline architecture for EoR deployments

While providing a seamless migration for 10GbE server attachment the 7300X-64T/7300X-64S line cards each provide an additional 4 x 40GbE QFSP+ ports. These ports are ideal for fiber connectivity to 10GbE attached storage or blade servers, and each can be software provisioned as four 10GbE ports, providing 16 10GbE ports per line-card for direct connectivity within the server row. By integrating the 40GbE ports into the line-cards they can also be used as 40GbE uplink connectivity, while still maintaining 48 x 10GbE/1GbE ports per linecard for server connectivity and avoiding dedicating a slot or two slots for “uplink” linecards.

MODERN WORKLOADS BENEFIT FROM HIGH PERFORMANCE

The 7300 Series is a high performance system where all modules provide wire-speed throughput. With this performance consistent across all line cards, regardless of whether configured for 1G, 10G or 40G port speeds, the Spline™ architecture is able to provide wire-speed server-to-server communication regardless of the rack or line-card a server is physically located on. As a result physical and virtual servers can be dynamically deployed within the architecture as demand arises, without additional concerns for performance or throughput. This wire-speed performance can be compared to a typical two-tier approach which would only be capable of a 3:1 oversubscription ratio or even worse a 20:1 or greater ratio in traditional EoR Catalyst 6500 designs.

Switch Platform	1GbE Copper Density (Subscription Ratio)	10GbE Copper Density (Subscription Ratio)	40GbE Port Density (Subscription Ratio)	Rack Space
Catalyst 6506(E)	192 (1.2:1)	64 (4:1)	16 (2:1)	11 RU
Arista 7304	192 + 16 x 40GbE (1:1)	192 + 16 x 40GbE (1:1)	128 (1:1)	8 RU
Catalyst 6509(E)	336 (1.2:1)	112 (4:1)	28 (2:1)	15 RU
Arista 7308	384 + 32 x 40GbE (1:1)	384 + 32 x 40GbE (1:1)	256 (1:1)	13 RU
Catalyst 6513(E)	528 (1.2:1 or 2.4:1)	176 (4:1 or 8:1)	44 (2:1 or 4:1)	21 RU
Arista 7316	768 + 64 x 40GbE (1:1)	768 + 64 x 40GbE (1:1)	512 (1:1)	21RU

Table 1: 10GbE/40GE port density scale of the Arista 7300 compared to Catalyst 6500

MULTI-CHASSIS LAG (MLAG) FUNCTIONALITY

To provide server resiliency and active-active traffic forwarding in an EoR design, a pair of 7300 Series installed at each end of a row can be deployed in an Multi-Chassis LAG (MLAG) configuration within the Spline™ architecture. The MLAG technology allows the two physical 7300 systems to appear as a single logical layer 2 switch to all servers within the row. The servers, blade or rack mounted, connect to the MLAG switches using standard LACP technology with the links of the LAG group split across both switches for redundancy. To provide active-active gateway redundancy for the subnets within the row the MLAG pair would be configured as a VARP gateway allowing both 7300 switches to actively route traffic upstream or to any other local subnet.

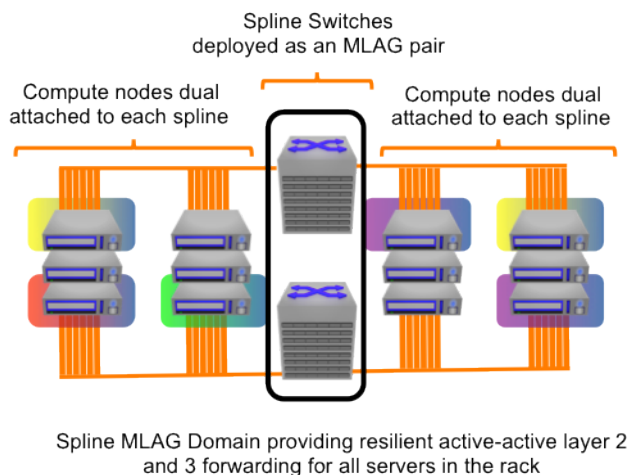


Figure 5: Spline MLAG domain for active-active L2/L3 forwarding

LEVEL OF SCALE THAT CAN BE ACHIEVED WITH THE SPLINE APPROACH

Table 2 provides guidance on the level of scale that can be now achieved with a Spline architecture using the 7300X platform. In each example, a 7300 chassis is fully populated with, either a 7300X-64S or 7300X-64T line-card, to provide the relevant 10GbE and 40GbE port density. The assumption is made that each physical server is deployed with two 10GbE ports, with the ports being split across a pair of 7300X in an MLAG configuration.

10GbE ports are RJ45 Copper 100/1000/10GBASE-T or SFP+ for server connectivity. 40GbE ports are included as part of the design for MLAG peer-links and uplink connectivity. In this example Arista 7300 Series switches are fully populated with 7300X-64S or 7300X-64T linecards:

Switch Platform	10GbE Ports per switch	40GbE Ports per switch	Total Number of Racks (16 physical servers per rack)	Virtual Machines (20 VM per physical server)
Arista 7304 MLAG Pair	192	16	12	3,840
Arista 7308 MLAG Pair	384	32	24	7,680
Arista 7316 MLAG Pair	768	64	48	15,360

Table 2: Arista 7300 Spline design port counts with 10GbE ports

POWER AND COOLING

In any EoR or middle of row design where network equipment will share rack space within a row with server equipment it is critical the airflow of the switch is consistent with the installed servers, whether that be front-to-back or back-to-front airflow. With a Catalyst 6500, which has side-to-side airflow, this can require additional equipment to be installed to redirect the cold air within the rack and prevent the hot exhaust raising the intake temperature of adjacent servers.

The 7300 Series platform is specifically designed for next generation EoR and MoR deployments, and the need to be compliant with industry standard data center hot aisle/cold cooling solutions. For this reason the 7300 Series platform provides front-to-back and back-to-front airflow configurations, allowing the correct airflow regardless of the whether the switch is deployed front-side or back-side to the server's network interface.

VXLAN FUNCTIONALITY

In traditional EoR architectures the layer 2/3 boundary for any server communication (physical or virtual) is defined as the EoR switch, which would act as the L3 default gateway. In a virtualized environment this can severely restrict the placement of virtual machines and their IPC communication, which is often constricted to a single layer 2 domain/subnet. To overcome this restriction and provide optimal use of resources within the data center, the 7300 Series platform provides hardware support for emerging encapsulation technologies such as VXLAN.

VXLAN is a multi-vendor industry-supported network virtualization technology that enables much larger networks to be built at layer 2 without the inherent scale issues that underpin large layer 2 networks. It uses a VLAN-like encapsulation technique to encapsulate layer 2 Ethernet frames within IP packets at layer 3 and as such is categorized as an 'overlay' network. From a virtual machine perspective, VXLAN enables VMs to be deployed on any server in any location, regardless of the IP subnet or VLAN that the physical server resides in. The VXLAN hardware gateway functionality of the 7300 Series allows the VXLAN encapsulation to be performed on the switch and routed transparently across an existing IP core.

By overcoming the need to build a flat, layer 2, network the 7300 Series can be placed within any existing EoR design and transparently provide layer 2 connectivity across rows of the data center. This connectivity is achieved without any uplift or change (physical or logical) to the existing layer 3 routed data center core. Allowing Virtualization administrators to make optimal use of all compute resources within the data center, while maintaining the existing routed layer 3 infrastructures.

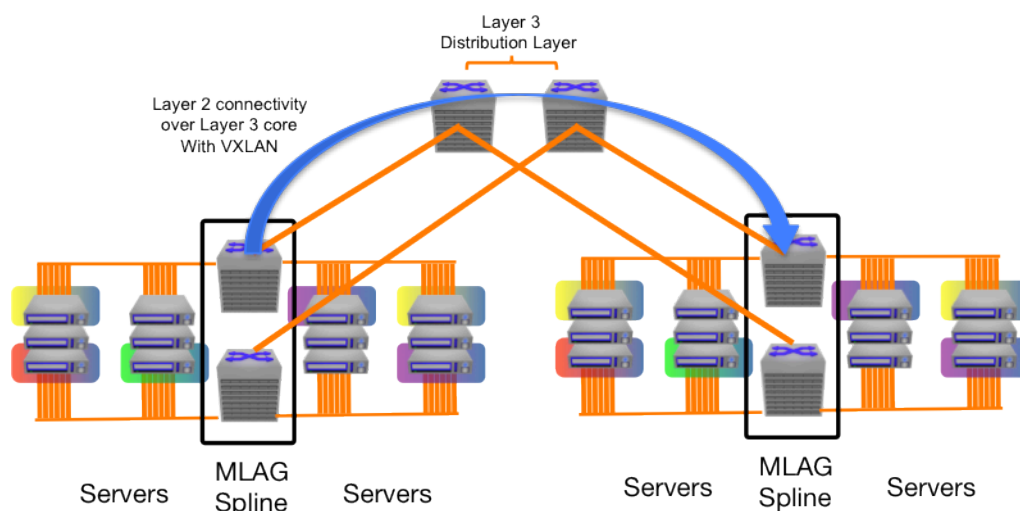


Figure 6: VXLAN providing Layer 2 connectivity over a routed core/distribution layer

VM TRACER

As virtualized data centers have grown in size, the physical and virtual networks that support them have also grown in size and complexity. Virtual machines connect through virtual switches and then to the physical infrastructure, adding a layer of abstraction and complexity. Server side tools have emerged to help VMware administrators manage virtual machines and networks, however equivalent tools to help the network administrator resolve conflicts between physical and virtual networks have not surfaced.

Arista VM Tracer provides this bridge by automatically discovering which physical servers are virtualized (by communicating to VMware vCenter APIs), what VLANs they are meant to be in (based on policies in vCenter) and then automatically applying physical switch port configurations in real time with vMotion events. This results in automated port configuration and VLAN database membership and the dynamic adding/removing of VLANs from trunk ports.

VM Tracer also provides the network engineer with detailed visibility into the VM and physical server on a physical switch port while enabling flexibility and automation between server and network teams.

EOS OPERATING SYSTEM

The Arista 7300 Series runs the same Arista EOS software as all Arista products, simplifying network administration. Arista EOS is a modular switch operating system with a unique state sharing architecture that cleanly separates switch state from protocol processing and application logic. Built on top of a standard Linux kernel, all EOS processes run in their own protected memory space and exchange state through an in-memory database. This multi-process state sharing architecture provides the foundation for in-service-software updates and self-healing resiliency. With Arista EOS, advanced monitoring and automation capabilities such as Zero Touch Provisioning, VM Tracer and Linux based tools can be run natively on the switch with the powerful quad-core x86 CPU subsystem.

CONCLUSION

Arista's single-tier Spline™ and two-tier leaf and spine network designs take the principles that have made cloud computing compelling (automation, self-service provisioning, linear scaling of both performance and economics) and combine them with the principles of Software Defined Networking (network virtualization, custom programmability, simplified architectures, and more realistic price points) in a way that is neither proprietary nor a vendor lock-in.

The Arista 7300 Series provides a unique combination of 100Mb to 40Gb port choices in a range of compact high performance systems that provides a simple, low-risk migration strategy for both brown-field and green-field data centers that accommodates the needs for virtualization, consolidation, network attached storage, low latency, power efficiency and does it without requiring multiple upgrade cycles. Combined with EOS a new set of monitoring, visibility, virtualization and provisioning tools are unleashed providing cloud scale benefits to collapsed Spline™ solutions.



Santa Clara—Corporate Headquarters

5453 Great America Parkway
Santa Clara, CA 95054
Tel: 408-547-5500
www.aristanetworks.com

San Francisco—R&D and Sales Office
1390 Market Street Suite 800
San Francisco, CA 94102

India—R&D Office
Eastland Citadel
102, 2nd Floor, Hosur Road
Madiwala Check Post
Bangalore - 560 095

Vancouver—R&D Office
Suite 350, 3605 Gilmore Way
Burnaby, British Columbia
Canada V5G 4X5

Ireland—International Headquarters
Hartnett Enterprise Acceleration Centre
Moylish Park
Limerick, Ireland

Singapore—APAC Administrative Office
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

ABOUT ARISTA NETWORKS

Arista Networks was founded to deliver software-defined cloud networking solutions for large data center and computing environments. The award-winning Arista 10 Gigabit Ethernet switches redefine scalability, robustness, and price-performance. More than one million cloud networking ports are deployed worldwide. The core of the Arista platform is the Extensible Operating System (EOS®), the world's most advanced network operating system. Arista Networks products are available worldwide through distribution partners, systems integrators, and resellers.

Additional information and resources can be found at www.aristanetworks.com.

Copyright © 2014 Arista Networks, Inc. All rights reserved. CloudVision, Extensible Operating System, and EOS are registered trademarks and Arista Networks is a trademark of Arista Networks, Inc. All other company names are trademarks of their respective holders. Information in this document is subject to change without notice. Certain features may not yet be available. Arista Networks, Inc. assumes no responsibility for any errors that may appear in this document. 03/14