Routing Architecture Transformations

Leveraging the principles of "scale out, simplify and software driven control", cloud networks have reaped the advantages of efficiency and cost. These cloud principles have challenged the status quo of various architectures including compute, storage and Ethernet switching. Ethernet has been the ubiquitous Local Area Network (LAN) technology for many years. With its simplicity, interoperability and cost efficiency, coupled with standardization, Ethernet is now the preferred physical layer for Wide Area Networks (WANs) and Metro Area Networks (MANs) replacing SONET/SDH, Frame Relay and other legacy technologies. Large-scale cloud data centers (DCs) are designed as resilient Layer 3 ECMP networks, and deployed with high performance Layer 3 Ethernet leaf and spine switches running rich Layer 3 routing stack and open IP routing protocols. With merchant silicon faithfully following Moore's Law, every new chip generation brings increased forwarding capacity, expanded table sizes, more features and capabilities, and low power.

Given these cloud-scale dynamics, the line between Layer 3 Ethernet switches and traditional Layer 3 routers is indeed blurring. The industry is at an inflection point, where the adoption of cloud principles has intersected the expanding capabilities of the merchant silicon feature set and scale, creating a disruption of legacy routing architectures.

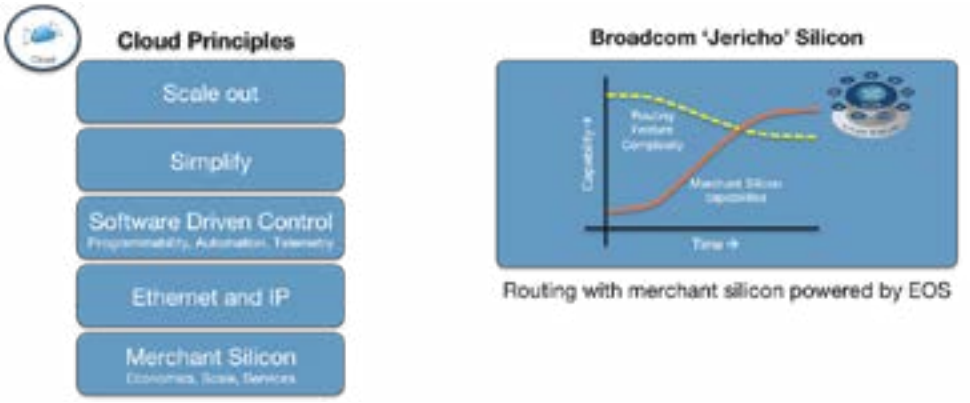

Figure 1: Cloud principles applied to Routing

This paper examines some of the use cases, including DC core, IP Peering, Cloud WAN and Telco NFV Cloud where traditional routing architectures are actively undergoing transformation. It also examines how the Arista 7500R and 7280R series, powered by EOS, applies to these use cases.

Use Case #1: The DC Core Needs A Spine

The first phase of cloud-scale transformation within the DC, focused on eliminating tiers and adopting a resilient Leaf Spine Layer 3 ECMP Universal Cloud network design. However, in most DCs, there still exists a tier above the Spine called the DC Core, which comprises of a pair of purpose built and oversubscribed routers facing the Internet or Data Center Interconnect (DCI). These platforms tend to be limited in system capacity and require wholesale forklifts to upgrade, unable to easily adapt to a rapid growth in application traffic demands or changing traffic patterns. The rationale for this design was as follows:

- Most resources were contained within the data center, and east to west traffic within the data center was far higher than north to south traffic traversing the DC core.
- Purpose built platforms were needed to support Internet scale routing and advanced IP routing protocols.

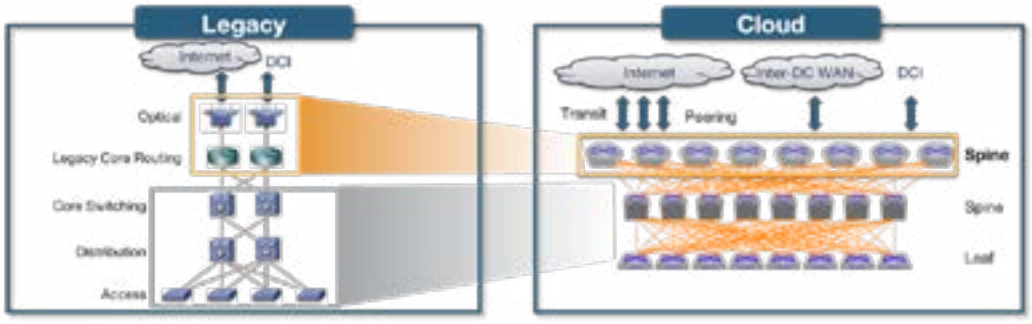

Figure 2: DC Core Transformation

The inter-DC traffic traversing the DC core is growing several-fold per year, due to increased levels of bulk data transfers between application clusters spanning multiple DCs. Whether you are a cloud or a service provider (SP), the requirement to operate several distributed data centers as one logical DC requires the ability to keep pace with the increasing inter-DC application bandwidth demands in a scale-out manner. Even in enterprise networks, critical resources commonly span multiple data centers or a private data center and the public cloud.

These shifts in traffic patterns drive the need for a flexible, cost-effective, scale-out design at the DC Core that is non-blocking, supports large scale ECMP for multi-pathing, enables IP routing with Internet route scale and delivers best in class routing convergence. The DC core needs a platform that can connect to the Internet, DCI or inter-DC WAN and even collapse routing and optical transport tiers into a single tier. The DC core needs a Spine architecture. The Arista 7500R Universal Spine platforms meet these requirements with up to 115 Tbps of fully non-blocking throughput, support for over 2M IPv4 and IPv6 routes using the Arista FlexRoute™ Engine and EOS NetDB with best-in-class routing convergence. With support for a wide range of Ethernet interfaces and integrated Coherent DWDM optics with wirespeed 256-bit MACsec support, the 7500R enables the removal of additional optical transport tiers, lowering cost and complexity, to provide a high bandwidth and flexible interconnect solution. The Arista 7500R series is an ideal choice to transform the legacy DC core to a Universal Spine architecture.

Use Case #2: Internet Peering - Evolution to Content Driven Internet

Looking back over the last two decades at the most impactful transformations in routing architectures, the Internet itself has undergone one of the biggest changes. The primary driver for this has been the ever-accelerating growth of real-time content on the Internet. From the 1990s to the early 2000s, Internet content was primarily email, text and static web pages, so the network interconnects were symmetric or strictly hierarchical. Large Tier 1 Service providers provided transit services to everyone else through their large global backbones, and smaller networks "peered" with each other due to the symmetric traffic flow. Fast forward to today, when content from YouTube, Netflix, Facebook, Spotify and online gaming dominates the Internet traffic. This increase in high-bandwidth content has set in motion a series of changes to the interconnect requirements, altering network economics, interconnection and business relationships between the content providers, the Internet Service Provider (ISP) and the end consumers. New services are driving increased bandwidth, connectivity, uptime and latency requirements, forcing interconnection and network infrastructures to adapt.

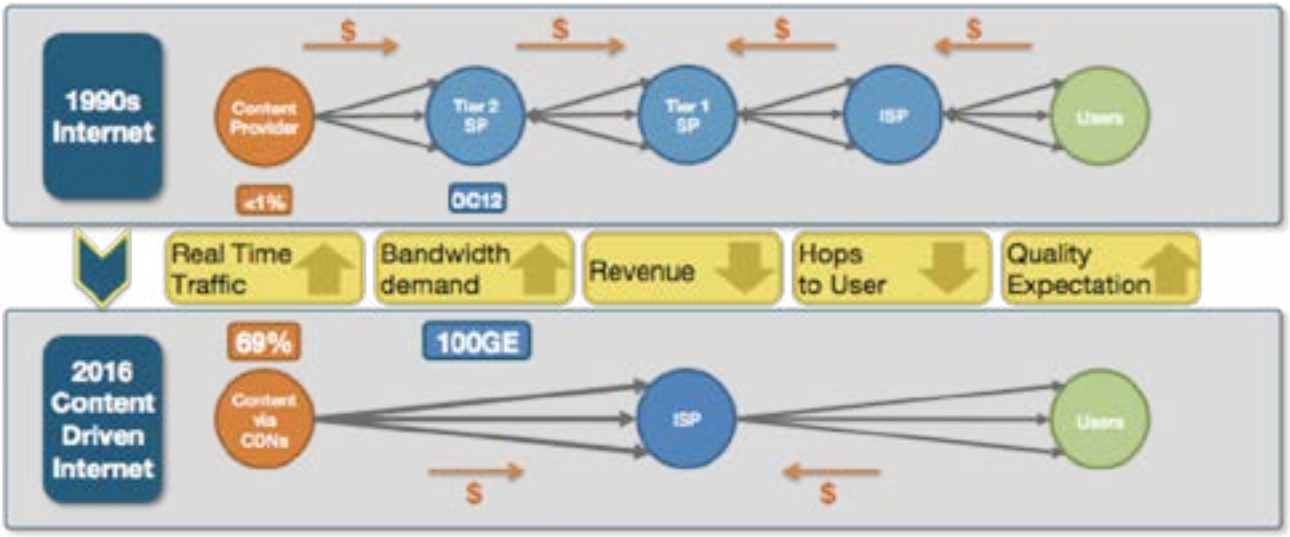

Figure 3: Evolution to Content Driven Internet

In the race to capture eyeballs and provide high quality user experience, content providers are aggressively rolling out Content Distribution Networks (CDNs) that place content closer to the edges of the network to avoid transit hops through the Internet, and are peering directly over high bandwidth peering interconnects with the ISPs. A typical content provider now has the choice of steering content via transit, private or public peering, or private WAN from the CDN pop. On the other hand, amidst growing bandwidth demand from "over the top" (OTT) traffic and declining revenue per user, the ISPs have been forced to adopt cloud efficiencies and principles as they increase capacity in their edge networks and on their IP peering platforms.

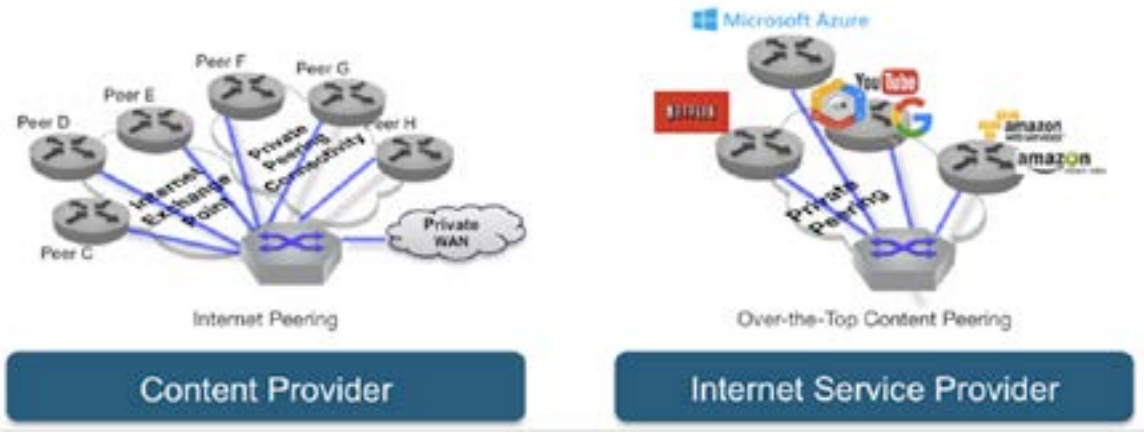

Figure 4: IP Peering with Arista 7500R Series Routing Platforms

Traditional "big iron" router platforms tend to be expensive, inflexible and over-engineered for these roles. By combining both large routing tables and high throughput with low power per port and high interface density, the Arista 7500R and 7280R offer a higher density, greener, programmable, cost-effective 100G IP peering platform for both the content provider and the ISP. The R series platforms can support over 2M routes in hardware thereby providing future-proof Internet Edge routing for customers. In addition to route scale, security at the edge is equally important and Arista AlgoMatch™ delivers industry leading scale for security access control policies on 7500R and 7280R series platforms with uncompromised scale for IPv6 ACLs. Additionally, these platforms support tunneling technologies including MPLS, VXLAN, GRE and MPLSoGRE, along with programmatic traffic steering options that content providers can leverage to optimally route the content.

Use Case #3: Cloud WAN and Segment Routing

To handle the increases in inter-DC traffic due to distributed clusters and changing application traffic patterns, most cloud customers have built a private Cloud WAN to interconnect their DCs. Traditionally these WAN networks would be built using MPLS routers running Traffic Engineering (TE) and bandwidth reservation protocols such a RSVP-TE. In order to keep utilization optimal, networks would use features such as auto-bandwidth, which would resize the bandwidth on the MPLS tunnels, based on monitoring traffic counters on the tunnels, i.e. actual utilization. But these TE path computation and auto-bandwidth algorithms tend to be vendor specific and come packaged in the software stack running on the box. Also, these algorithms don't account for global conditions or changing traffic source-destination patterns. Finally, these protocols tend to be signaled hop-by-hop, are state heavy and introduce significant complexity to network operations.

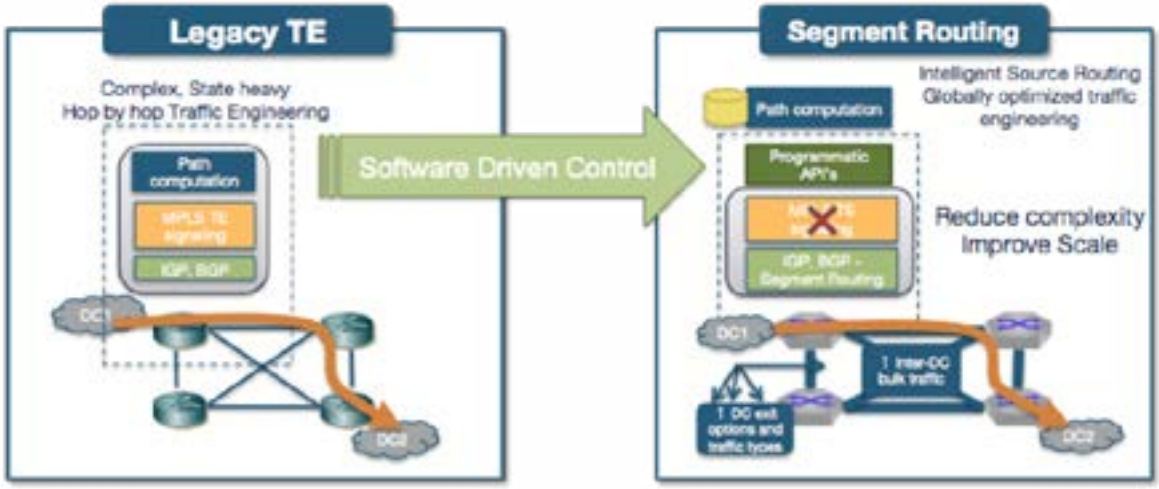

Figure 5: Semi-Collapsed Storage Architecture

Cloud networks need fine-grained control for steering traffic across a wide variety of network interconnects based on a holistic view of their end-to-end network. They need the ability to compute an optimal traffic engineered path based on the global topology, distance, bandwidth availability, congestion conditions, traffic type, latency sensitivity, and business logic. Moreover, the ability to dynamically adapt these computations to changing conditions and criteria is critical. Reliance on the vendor-specific path computation and traffic engineering solutions for their business critical applications doesn't meet that need. Also, simplifying network design, eliminating complexity and removing unnecessary features and functionality is an overarching cloud principle that drives cloud efficiency.

The Cloud WAN requires a software driven approach to traffic engineering that eliminates complexity and enables fine grained control. Segment Routing (SR) provides the perfect paradigm for intelligent software-driven source routing. With this approach, path computation and traffic engineering is centralized and can be customized by the customers to meet their individual deployment needs. Extracting this functionality out of the routing platform eliminates their reliance on vendor specific path computation algorithms. It also helps simplify the network architecture itself by removing the need to run hop-by-hop signaling protocols like RSVP-TE. The physical network can now focus on high performance routing and switching, running IP routing protocols, programmability and rich network telemetry in order to support this new software driven traffic engineering approach.

Segment Routing is being standardized in the IETF in the SPRING working group. MPLS based Segment Routing still leverages MPLS forwarding in the data path. By extending the IGP (ISIS and OSPF) for label distribution, using external path computation engine to compute optimal paths based on varied constraints, then encapsulating the explicit path in the data plane with an MPLS label stack, MPLS Segment Routing solution for TE applications delivers the ideal solution for the Cloud WAN. Segment Routing also offers better control plane and data plane scaling by removing the need for per flow state at every network hop and better ECMP characteristics compared to traditional TE solutions.

Arista 7500R and 7280R series support ISIS-SR extensions, multi-label stack support for BGP-LU, MPLS ECMP hashing, and ability to look deep into the MPLS label stack for hashing decisions. In addition, EOS is built on the foundation of high degrees of programmability, and EOS NetDB provides real-time state streaming and analytics for live monitoring and historic forensic troubleshooting, making the R series the ideal Segment Routing platform for the Cloud WAN.

Use Case #4: Telco Transformation - Service Provider NFV

Telco networks offer various "network services" to their customers - from Ethernet access, transparent LAN or wire service, secure L3VPN to subscriber management. These services are offered from Central Offices (COs), which are located in close proximity to the customer base. COs are geographically distributed, come in various sizes and usually contain a few racks of servers and storage systems with several "big iron" multi service edge routers that terminate the customer circuits and instantiate the network services while providing connectivity to the backbone network. In legacy networks, these edge routers were required to terminate circuits from Frame Relay, TDM or SONET/SDH. The COs are prime real estate in local metros and are designed to survive natural disasters. Telco equipment in COs has strict environmental and physical packaging requirements, with a focus on very high uptime.

Many Telcos face business challenges from increased high bandwidth OTT traffic and competition from the cloud providers. Instead of competing directly with public cloud offerings, the Telcos are instead adopting cloud principles to deliver their network services in a more efficient manner. Some are providing cloud connection services to their existing L3VPN customers to provide secure VPN access to the public cloud. Others are leveraging the user proximity of the COs to offer CDN services or to house 3rd party CDN caches from Google, Netflix, etc. These trends have resulted in a dramatic overhaul of the legacy routing architectures and a transformation to cloud network designs.

The most significant inefficiency in the legacy Telco networks is the reliance on large and expensive hardware service edge platforms. These platforms are rigid, lack programmability, and need wholesale replacement for meaningful bandwidth upgrades. They present a real challenge to the operator in bringing new services to market or adapting to business growth in a cost-effective and agile manner.

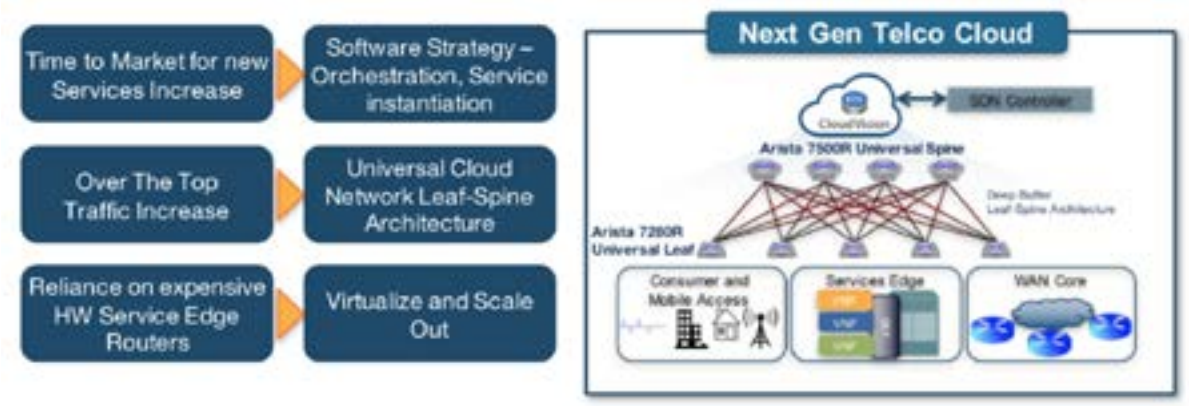

Figure 6: Service Provider NFV Cloud

As Service Providers think through the benefits of cloud principles, the main transformation is a holistic software-centric approach to offering services and minimizing the reliance on hardware edge platforms. Virtualizing the service edge into Virtual Network Functions (VNFs), SP Network Functions Virtualization (NFV) aims to adopt an efficient, scale-out, cost-effective approach to delivering cloud based network services. The service provider can now meet growing demands by instantiating more software VNFs as needed on any server. This requires the network infrastructure to support some Network Virtualization technology to interconnect these VNFs over an overlay network. Additionally, provisioning, orchestration and telemetry functions are envisioned to be software driven in the new Telco NFV Cloud.

With the virtualization of the Service Edge now eliminating much of the complexity from the hardware edge platforms, the physical network can focus on high performance switching and routing and network virtualization. Arista's Universal Cloud Network (UCN) leaf-spine architecture with the 7280R Series Universal Leaf and 7500R Series Universal Spine delivers the best scale-out network architecture for the SP NFV networks. EOS supports multiple overlay options including VXLAN, MPLS, MPLSoGRE and powerful capabilities for network automation with EOS CloudVision®. Arista AlgoMatch technology provides rich visibility with accelerated sFlow at dense 100G that can be used for various applications from optimal traffic steering to DDOS monitoring to trending analysis. AlgoMatch in conjunction with EOS advanced visibility and Telemetry capabilities including real-time state streaming with NetDB, DANZ, Tracers and CloudVision Telemetry provide the needed high visibility for customers building NFV clouds. EOS CloudVision integrates with various SDN controllers and orchestration systems enabling best-of-breed ecosystem choices for the SP. With support for DC power and NEBS compliance, these platforms are an ideal fit for central offices, which are transforming into the next generation Telco NFV Cloud.

Arista 7500R Universal Spine Platforms

The Universal Spine is a non-blocking, resilient and programmable platform, which can be flexibly used in various roles for Layer 2 or Layer 3 switching and routing, with support for Internet scale routing, large scale ECMP with best-in-class convergence and rich network telemetry, visibility and automation.

The Arista 7500R Series delivers the industry's highest performance and system throughput to meet the needs of the largest scaled networks. They combine scalable L2 and L3 resources and high port density with advanced features for network monitoring, precision timing and network virtualization to deliver scalable and deterministic network performance while simplifying designs and reducing OpEx.

Figure 7: Arista 7500R Universal Spine platforms

Deep packet buffers and large routing tables provide complete deployment flexibility and allow the 7500R to be deployed in a wide range of open networking solutions, including large scale layer 2 and layer 3 cloud designs and routing applications.

Available as a choice of 16, 12, 8 and 4 slot, the Arista 7500R is the next generation of the 7500 Series and sets a new standard for performance, density, reliability, and power efficiency. The 7500R can support up to 576 ports of wire speed 100GbE and 40GbE and offers over 150 Tbps of total capacity with a broad choice of Ethernet line cards. Every 100GbE interface supports a choice of five speeds including 25GbE and 50GbE providing unparalleled flexibility and the ability to seamlessly transition data centers to the next generation of Ethernet performance.

Table 1: Arista 7500R Series Ethernet Port Combinations and System Performance

	7516R	7512R	7508R	7504R
Maximum 10GbE Density	2,304 Ports	1,728 Ports	1,152 Ports	576 Ports
Maximum 25GbE Density	2,304 Ports	1,728 Ports	1,152 Ports	576 Ports
Maximum 40GbE Density	576 Ports	432 Ports	288 Ports	144 Ports
Maximum 50GbE Density	1,152 Ports	864 Ports	576 Ports	288 Ports
Maximum 100GbE Density	576 Ports	432 Ports	288 Ports	144 Ports
Maximum Throughput / PPS	150Tbps / 69Bpps	115Tbps / 51Bpps	75Tbps / 34.5Bpps	38Tbps / 17.3Bpps

Designed for high availability, all components are hot swappable, with redundant supervisor, power, fabric and cooling modules with efficient front-to-rear airflow. The system is purpose built for co-location and data centers, with a choice of AC or DC power, and are designed for NEBS compliance. These attributes make the Arista 7500R an ideal platform for building reliable and highly scalable data center networks.

The Arista 7500 Series uses a deep buffer virtual output queue (VOQ) architecture that eliminates head-of-line (HOL) blocking and virtually eliminates packet drops even in the most challenging network scenarios. An advanced traffic scheduler fairly allocates bandwidth between all virtual output queues, while accurately following queue disciplines, including weighted fair queueing, fixed priority, or hybrid schemes. As a result, the Arista 7500 can handle the most demanding traffic requirements with ease, including mixed loads of real-time, multicast, and storage traffic, while still delivering low latency.

Highest 100G Density with Power Efficiency

The 7500R Series delivers up to 576 ports of 100GbE and 150 Tbps of system throughput in a range of compact systems, at breakthrough price/performance compared to traditional routers. Designed with titanium rated power supplies that are over 94% power efficient and combined with power efficient merchant silicon that needs less than half the power/100G compared to similar routing platforms this power advantage provides savings sufficient to power as much as two additional rack of servers in a typical installation.

Arista FlexRoute Engine and EOS NetDB

The Arista innovation of FlexRoute Engine provides support for the full Internet routing table, in hardware, with IP forwarding at Layer 3 and with sufficient headroom for future growth in both IPv4 and IPv6 route scale to more than 2 million routes. Coupled with Arista EOS NetDB evolution, the Arista 7500R supports hundreds of BGP peers and millions of routes, with large scale ECMP and proven best-in-class routing convergence.

Figure 8: Arista FlexRoute Engine and EOS NetDB

Arista AlgoMatch

AlgoMatch is a unique Arista innovation for modern cloud networks, combining both software and hardware to enable more flexible and scalable solutions for access control, policy based forwarding and network telemetry. By combining general purpose memory with advanced software algorithms AlgoMatch delivers higher scale, performance and efficiency with lower power and is more cost effective than traditional TCAM-based solutions. AlgoMatch provides a more efficient packet matching algorithm that enables flow matching for access control, policy and visibility. The net benefits are a high performance policy engine with both increased functionality and scale in a cost and power efficient solution. AlgoMatch is available on the 7500R and 7280R Series of products.

- AlgoMatch enables IPv4 and IPv6 access control at the same scale
- L4 rule ranges are programmed efficiently without expansion or reduced capacity
- Multiple actions can be performed on a single packet or flow
- User defined filters allow flexible packet classification based on offsets for custom actions
- Supports rich policy with consistent semantics that would exhaust classical resources

Arista 7280R Universal Leaf Platforms

The Arista 7280R series Universal Leaf platform is a set of purpose built 10/25/40/50/100GbE fixed configuration 1RU and 2RU systems designed for the highest performance environments such as IP Storage, Content Delivery Networks, Data Center Interconnect and IP Peering and NFV clouds. The 7280R platforms are designed to be the Universal Leaf for various workloads - compute, storage, routing, digital media; etc. Furthermore, with up to 10 Tbps of forwarding capacity, and with the same innovative FlexRoute and AlgoMatch technologies as the 7500R series, the 7280R series are extremely versatile for the broad and varied routing use cases discussed.

Figure 9: Arista 7280R Universal Leaf platforms

Arista EOS

EOS is built on the strong foundations of a multi-process state-sharing architecture with modularity, programmability, fault containment and resiliency as the core software building blocks. EOS combines these attributes with key infrastructure innovations like EOS SDK, Go programming language, NetDB for improved route scale and convergence, a real-time state streaming infrastructure, and support for Docker containers, etc. These strengths allow the deployment of Arista platforms in various roles beyond the traditional Cloud and Service Provider data center environment, while ensuring seamless customer experience and high software quality.

Some of the key EOS feature strengths and features include:

- *A rich routing stack for layer 3 IP unicast and multicast protocols* - BGP, ISIS, OSPF, IGMP, PIM-SM, PIM-SSM, PIM Bidir
- *Support for IPv6* - BGPv6, OSPFv3 and ISIS
- *MPLS protocols* - LDP, ISIS-SR, BGP LU (w multi-label stack), LDP based pseudowires and MPLS ECMP, allowing customers to deploy the 7500R Series as MPLS label-switched routers
- *Tunneling options* - Network virtualization and choice of tunneling with VXLAN, MPLS, GRE, MPLSoGRE, IP-in-IP along with programmatic traffic steering allows customers interested in exit point selection for traffic exiting the data center to steer traffic over optimal paths
- *EOS NetDB* - enables large scale routing capability with industry leading route convergence and wide ECMP along with real-time network telemetry and visibility features
- *EOS CloudVision* - provides network wide orchestration and automation and a single point of integration for a wide variety of orchestration systems including OpenStack and OVSDB based SDN controllers

Summary

This paper reviewed changing trends and how cloud networking principles are driving transformation of legacy routing network architectures. All these transformations strive to eliminate the feature complexity from the physical network and instead increase the focus on high performance, high density platforms with programmability and rich telemetry, using an underlying scale-out design. With solid software foundations and innovations in both platform and EOS, Arista's open and extensible solutions expand the capabilities in density, table sizes, programmability, tunneling, traffic steering and software features. These enable the Arista 7280R Universal Leaf and 7500R Universal Spine platforms to be deployed in various roles that require high performance switching and Internet scale routing, combined with high port density, low power, programmability and automation, where traditional router platforms are lacking.

Santa Clara—Corporate Headquarters
5453 Great America Parkway,
Santa Clara, CA 95054

Phone: +1-408-547-5500
Fax: +1-408-538-8920
Email: info@arista.com

Ireland—International Headquarters
3130 Atlantic Avenue
Westpark Business Campus
Shannon, Co. Clare
Ireland

Vancouver—R&D Office
9200 Glenlyon Pkwy, Unit 300
Burnaby, British Columbia
Canada V5J 5J8

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

India—R&D Office
Global Tech Park, Tower A & B, 11th Floor
Marathahalli Outer Ring Road
Devarabeesanahalli Village, Varthur Hobli
Bangalore, India 560103

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

Nashua—R&D Office
10 Tara Boulevard
Nashua, NH 03062