

Arista R3-Series - Multiple Generations of Innovations

The growing demand for higher bandwidth in accelerated time frames is driven by hyper-scale cloud networks and Service Providers. The imminent availability of 5G and Edge computing with higher bandwidth devices, IoT data, new models for video consumption and ongoing Cyber-security challenges raises the expectation of intelligent network services. Customer requirements rarely call for an entirely new network architecture. New technology, however innovative, must preserve existing investments, provide backward and forward compatibility, and enable migration. Many system vendors see this pace of change as a fundamental challenge that should be addressed with new and radically different forwarding architectures, pipelines or operating models that disrupt customers, rather than facilitating the evolutionary approach. Arista takes an evolutionary approach when it comes to facing these challenges with the new R3 generation. The Arista R3-Series comprises a set of products in the latest generation of the highly successful Arista 7500 and 7280 Series platforms. The Arista 7500 Series was first released in 2010 as a high density 10G system and over the last 10 years has evolved with each new generation offering higher performance and more capabilities for routing, telemetry and network security, with investment protection.

The latest generation of these platforms are the high capacity 7800R3, 7500R3, and the 7280R3, offering a breadth of form factor, performance and density that scales from 4Tbps to over 460Tbps, with all systems sharing common and consistent capabilities.

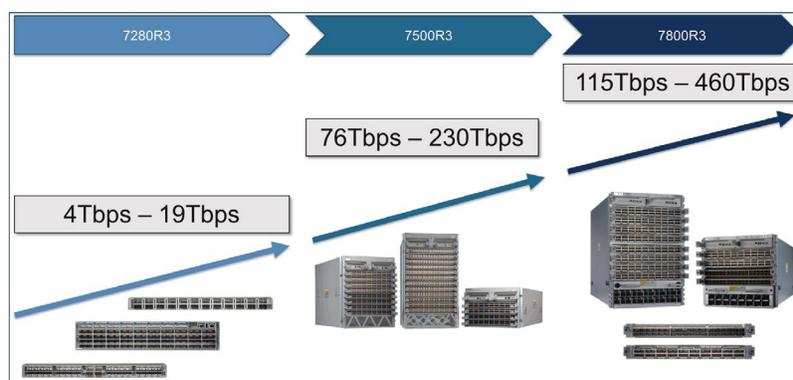


Figure 1: Arista R3-Series from 4Tbps to 460Tbps

This paper introduces the latest generation of the R-Series, outlines the capabilities and innovations that address the constantly evolving network challenges, and the introduction of 400G networks.

Common EOS

Arista EOS® Innovations and 10 Years of Experience

Since 2010, when the very first 10G based 7500 Series and 1G based 7048 Series were launched, each new generation of fixed and modular systems has introduced a set of innovations driven by Arista EOS. The first generation of deep buffer modular and fixed systems drove advanced two-tier designs for Oil and Gas research, High Performance Compute, Online retailers, big data and search that set the standard for the next 10 years. The 7500 Series was so well received that it was awarded Network World Best of Interop in 2010.

The 2013 launch of the 7500E and 7280SE platforms expanded Spine and Leaf networks to 40G and 100G, and introduced a series of innovations including Tap Aggregation, latency monitoring, Coherent 100G DWDM with integrated MACsec encryption for DCI and support for VXLAN for overlay networks. In 2013 the 7500E Series was awarded a second best of Interop, recognizing the investment protection, coupled with these innovations and 3X performance increase.

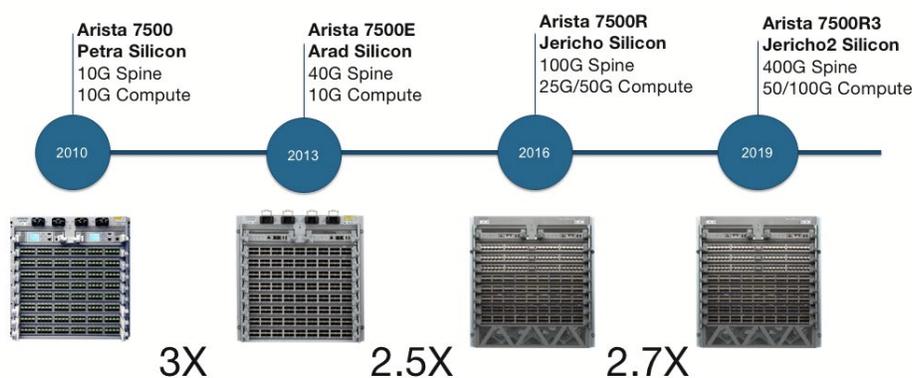


Figure 2: Arista 7500 Series 10 years of continuous innovations and experience

The 2016 release of the 7500R and 7280R Universal Spine brought support for Cloud-grade routing for Internet eXchange (IX) providers, Cloud WAN, next generation cable operators and Over The Top (OTT) services and Content Delivery Networks (CDN). The R-Series of scalable platforms support internet scale routing with [FlexRoute™](#), EOS NetDB innovations, modern protocols and open automation, which allowed for the [transformation of routing architectures](#) and replaced legacy routers. On the 7500R2 and 7280R2 Series, AlgoMatch and Accelerated sFlow provide unprecedented levels of secure access control and deep visibility into high capacity 100G networks and MACsec which was expanded for Data Center Interconnect (DCI) solutions.

The next generation of the R-Series is available in the 7800R3, 7500R3 and the 7280R3 Series with support for more EOS driven innovations.

Proven Architecture

Investment Protection

Customers rarely get a clean sheet of paper with which to design networks. New systems have to coexist with previous networks designs, and offer in place upgrades. The R3-Series provide continued investment protection, from the original 7500 Series through to the 7500R3 Series.

- **Interoperability of the 7500R3 Series** with the previous generation of line cards and fabrics allows mixing generations to preserve existing infrastructure.
- **Reuse of common architecture** across the R3-Series systems provides a common set of building blocks, leveraged to accelerate time to service of new systems.

- **Single OS** across the entire portfolio addresses the challenge of qualifying new products, where existing tools, processes and features work as expected. This is in comparison to the challenge of introducing an all new product with a different operating system and dissimilar features.
- **Feature Rich** building on the code base of EOS and leveraging 1000's of software development years with common platforms allows for faster qualification without having to wait for features to be released.

Arista EOS is a modern operating system and continues to evolve, with ongoing refactoring of major aspects of the code, adding scale, new features and support for new silicon, all without sacrificing support for existing features and code quality.

Table 1: Real Investment Protection

	Characteristic	Why it Matters
H/W Investment Protection	<ul style="list-style-type: none"> • Complete interoperability 3rd/4th generation • Existing line cards, fabric modules, chassis & interfaces work as-is 	<ul style="list-style-type: none"> • True Hardware Investment Protection • Significant CapEx savings • No forklift upgrade
Common Architecture	<ul style="list-style-type: none"> • Deep buffer • Ingress VoQ • Cell based Fabric 	<ul style="list-style-type: none"> • 100% efficient • Handle unpredictable workloads • No hotspots
One OS	<ul style="list-style-type: none"> • Single OS across entire portfolio 	<ul style="list-style-type: none"> • Reduces OpEx • Existing tools & processes work as-is
Proven Platform	<ul style="list-style-type: none"> • Feature-rich day one • High quality on day one 	<ul style="list-style-type: none"> • Leverage 1000s of SW development years • Faster qualification

Consistent VoQ non-blocking fair delivery

The Arista R3-Series leverages a deep buffer virtual output queue (VOQ) architecture with a cell based fabric that is 100% efficient and provides consistently low latency and jitter, eliminates head-of-line (HOL) blocking and virtually eliminates packet drops even in the most congested network scenarios. An advanced traffic scheduler fairly allocates bandwidth between all virtual output queues while accurately following queue disciplines including weighted fair queuing, fixed priority or hybrid schemes. As a result, the Arista 7280R3, 7500R3 and 7800R3 Series can handle the most demanding data center requirements with ease, including mixed traffic loads of real-time, multicast, and storage traffic while still delivering low latency.

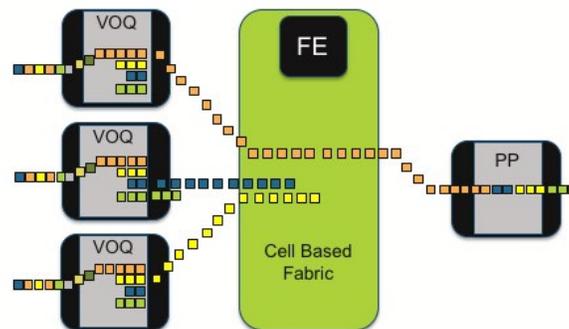


Figure 3: Cell Based Fabric and Ingress VoQ Architecture

The advantages of the VoQ architecture are that traffic is not forwarded to a destination without a credit, and that the cell based fabric uses all paths to the egress packet processor to avoid creating hot-spots. The cell based fabric leverages a variable cell size and handles mixed interface speeds with no impact on system performance.

Enhancements for the R3-Series include multicast expansion in the ingress, fabric or egress, to reduce drops or queueing. In the event of a drop, it can be mirrored to a destination for analysis.

Deep packet buffers with HBM2

The R3-Series is consistent with previous generations of the R-Series with flexible packet buffers combining dedicated on-chip resources with high bandwidth access to 8GB of HBM2 memory per packet processor.

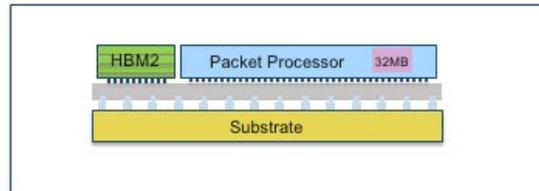


Figure 4: HBM2 and Packet Processor Packaging

The buffer access speed has increased at a factor of five over the previous generation, using advanced memory technology commonly found in graphics and AI systems, aligned to a 5X increase in the packet processor performance. High Bandwidth Memory-2, or HBM2, when integrated to the packet processor package, enables both high bandwidth and greater capacity, requiring 43% less power than the equivalent GDDR memory.

Integrated HBM2 memory allows the R3-Series to deliver a high number of interfaces, increases system reliability, and requires substantially less power than comparable high speed memory options such as HMC, which lack the package integration benefits. HBM2 is available from multiple sources, is backed by the JEDEC consortium (1), in common use on graphics processors and has a technology roadmap to higher bandwidth and performance.

The 32MB of on-chip buffers are used for non-congested forwarding with seamless expansion to the HBM2 based on instantaneous or sustained congestion with no packet drops. Buffers are assigned per VoQ and class of service, with no requirement to tune for optimum behavior.

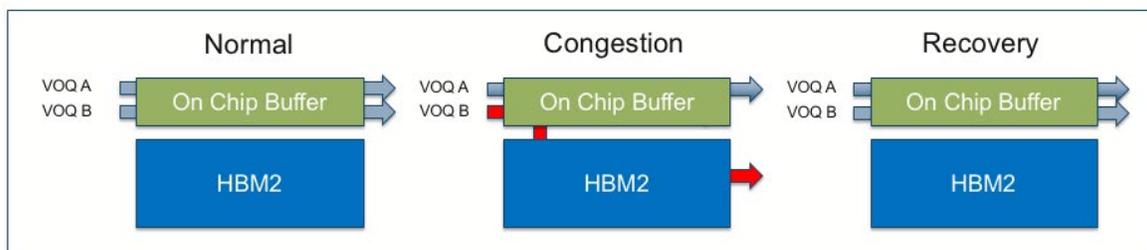


Figure 5: Seamless packet buffer optimization with OCB and HBM2

Memory management is as important as capacity and in the R3-Series the packet buffer memory is optimized for modern networks, that are naturally bursty. Avoiding waste from small frames, memory is allocated in 4K byte segments to accommodate multiple frames, which are queued directly from memory to the destination packet processor, and are not returned via the OCB.

Power Efficient 100G and 400G Ethernet

The R3-Series introduces support for high density 100G and 400G Ethernet in a variety of platforms from 24 ports of 400G in 1U to a modular system with 576 ports, and increase performance by 3X over the previous generation.

A single port of 400G can support 1x400G, 2x200G or 4x100G modes with up to 36 ports in a single line card, equivalent to 144 ports of 100G. While demand for 400G is driven by the largest scale networks the R3-Series 100G systems provide higher bandwidth at 35% lower power per gigabit, enabling the rapid adoption of next generation systems. To bridge the migration from 100G to 400G the R3-Series 400G and 100G systems provide flexible port speed options and media types with breakout cables and parallel optics.

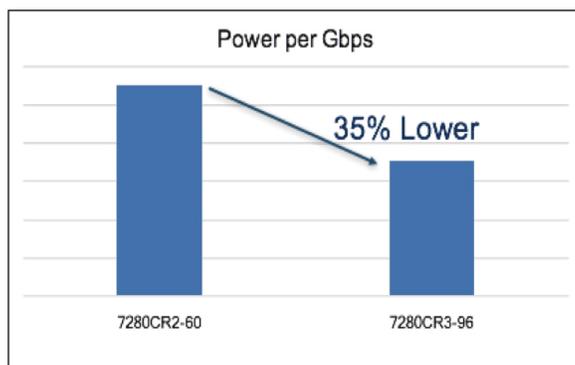


Figure 6: Improved power per bandwidth

FlexRoute Improvements

Internet Route Scale

All R3-Series products support full internet route scale with a capacity of 1.3M routes leveraging a common set of forwarding table resources for the L2 and L3 tables. The R3K-Series with expanded size tables can support up to 2.5M IPv4 and IPv6 routes, allowing for multiple overlapping full tables or for many years of growth of the Internet. Routes are held on the packet processor, without the need for external TCAMs or additional memory, affecting performance, power efficiency or overall system density.

Geoff Huston, Chief Scientist at APNIC, the Asia Pacific Regional Internet Registry has been providing research, analysis and commentary on the global internet routing table for over a decade. The published projection for the next 5 years is shown below. As IPv6 addresses take more space than a single IPv4 it's likely that capacity for 1.3M routes will be required to hold a full table, in the near future, and that 2.5M is required for projected growth beyond 2024.

Table 2: Historic growth of IPv4 and IPv6 announcements

Metric	Jan 2016 (Actual)	Jan 2017 (Actual)	Jan 2018 (Actual)	Jan 2019 (Actual)	Jan 2020 (Prediction)	Jan 2021 (Prediction)	Jan 2022 (Prediction)	Jan 2023 (Prediction)	Jan 2024 (Prediction)
IPv4 Prefixes	587,000	646,000	699,000	755,000	810,000	864,000	919,000	974,000	1,028,000
(% Increase)		9%	8%	7%	7%	6%	6%	6%	5%
IPv6 Prefixes (L)	27,000	37,000	45,000	62,000	75,000	89,000	102,000	116,000	130,000
(% Increase)		27%	18%	27%	17%	16%	13%	12%	11%
IPv6 Prefixes (E)	27,000	37,000	45,000	62,000	83,000	109,000	145,000	192,000	255,000
(% Increase)		27%	18%	27%	25%	24%	25%	24%	25%
Total (linear IPv6)	614,000	683,000	744,000	817,000	885,000	953,000	1,021,000	1,090,000	1,158,000
(% Increase)		6%	6%	6%	6%	6%	6%	6%	6%
Total (exponential IPv6)	614,000	683,000	744,000	817,000	893,000	973,000	1,064,000	1,166,000	1,283,000
(% Increase)		10%	8%	9%	9%	8%	9%	9%	9%

Source: <https://labs.apnic.net/?p=1195>

Next Generation Edge with EVPN and Segment Routing

The R3-Series expands on Segment Routing and EVPN support with more flexible options for underlay and overlay topologies. They support expanded label depth, new tunnel types including v6GRE and VXLAN-v6 and improved load balancing on deep tunneled packets.

As multiple technology transitions evolve at the network edge e.g. multi cloud, workload movement between private and public cloud, content caches near customer proximity and distributed NFV services at the edge, connectivity and segmentation are key to agile service delivery. The R3-Series provides the flexibility to deliver cost optimized and high density interconnect solutions, to accelerate NFV edge services with EVPN, and Segment Routing solutions for the WAN(2).

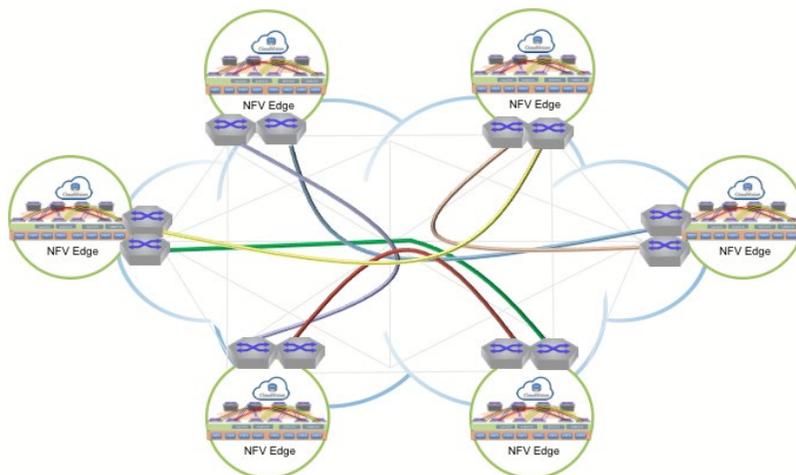


Figure 7: Next Generation Telco Edge with EVPN and Segment Routing

Programmability

Programmable Processing with Elastic Pipeline

The R3-Series delivers advanced programmable packet processing with a new elastic pipeline. Complex features requiring additional lookup cycles are implemented through a pool of stages in the pipeline. Existing advanced features, including Accelerated sFlow, are implemented in the programmable pipeline avoiding external offload engines, and increasing flexibility.

Flexible Packet Processor Tables

In large scale networks the total scalability is bound by one or more resources: Routes, Hosts, MAC addresses or MPLS Labels. In these environments the ability to increase the capacity by flexibly reallocating tables allows for more efficient designs.

The R3-Series leverage a common central database of forwarding and lookup resources, available to all ingress and egress stages for the common forwarding entries. These resources are allocated using forwarding profiles that ensure the optimal allocation to different tables, for a wide range of use-cases. The L3 optimized profile expands the routing and next-hop tables to address networks where larger route table capacity is required, while the Balanced profile allocates the resources across all tables and is suited for leaf and spine data center applications.

The R3K-Series systems and linecards include support for double the capacity of forwarding tables for environments that require resources to be expanded further. The R3K packet processors have 2x larger integrated forwarding tables, and the profiles optimize for Peering, DCI and IP/MPLS applications with up to 2.5M routes or 2.8M MPLS Labels.

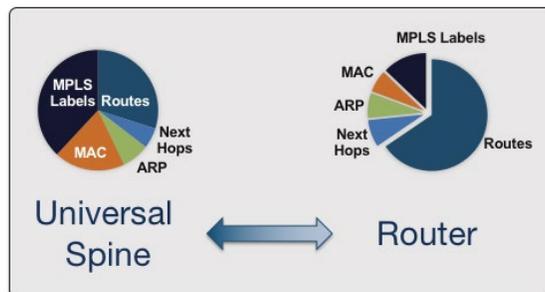


Figure 8: R3-Series Packet Processor Flexible Tables

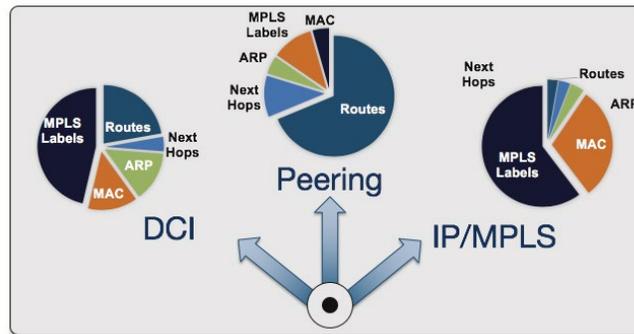


Figure 9: R3K-Series Packet Processor Flexible Tables

Advanced Network Telemetry

Accelerated sFlow with 1:100 at 400G

The R-Series support for Accelerated sFlow enabled support for high rate sampling across a system with dedicated sFlow engines. The R3-Series expands Accelerated sFlow to be supported on all systems, with each packet processor leveraging the programmable pipeline for high sampling rates at 400G and 100G on all ports, with the additional metrics of AS-Path and Next-Hop. Telemetry information on traffic flows is valuable when determining the sources of rogue and misbehaving applications within the network, together with flow analysis and traffic engineering of paths.

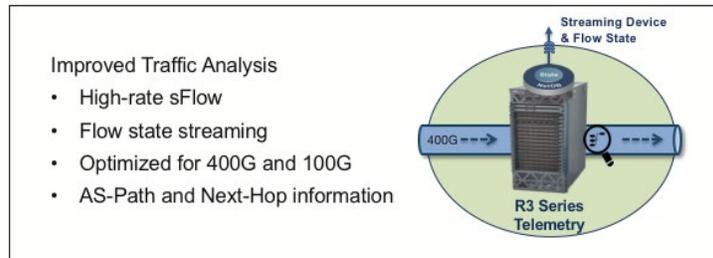


Figure 10: Accelerated sFlow with state streaming

In-band Network Telemetry for flow and packet forensics

In-band network telemetry, or INT, is a standards approach to providing deep visibility into traffic in real-time, with no impact on switch performance. INT is available on the R3-Series and provides per-flow monitoring of drops, latency, congestion and can reconstruct the network topology and paths. INT information can be exported in IPFIX or sFlow formats to a management system or collector such as Arista CloudVision, for predictive analytics and deep forensics.

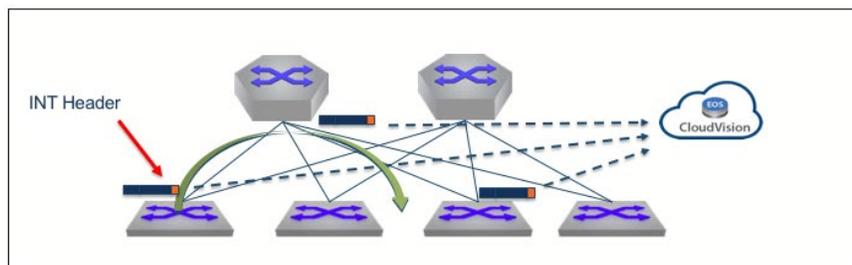


Figure 11: Inband Network Telemetry in Action

Algorithmic ACLs

The R3-series provide support for algorithmic ACLs on all systems that increase both the scale and flexibility leveraging the programmable pipeline. With support for over 100K rules the R3-Series are suited for internet edge and cybersecurity roles with stateful ACLs and rich counters identifying and filtering with no loss in performance.

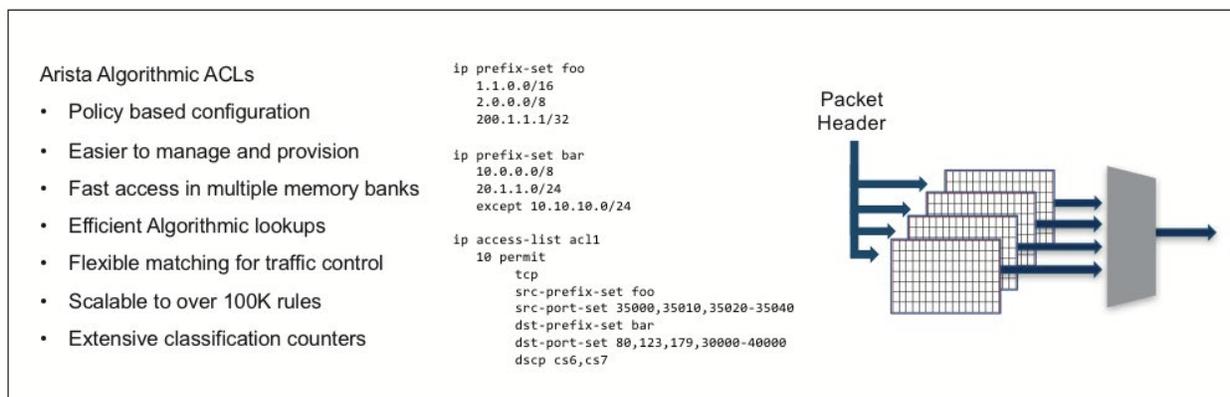


Figure 12: Algorithmic ACLs combine software and hardware innovation

Summary

The Arista R3-Series expands the industry leading spine and leaf systems, preserving investment with a consistent architecture, EOS features and the introduction of support for 400G along with new innovations for programmability, telemetry, routing and security.

- (1) <https://www.jedec.org/standards-documents/docs/jesd235a>
- (2) [EANTC Test reports validating EVPN and Segment Routing features](#)
- (3) APNIC <https://labs.apnic.net/?p=1195>

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



Copyright © 2019 Arista Networks, Inc. All rights reserved. CloudVision, 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. April 6, 2019 02-0082-02