Multicast Orchestration with Arista Media Control Service

High Performance Cognitive Orchestration for all Multicast Workflows

Introduction

Arista's Media Control Service (MCS) provides cognitive orchestration for multicast flows, that is essential to equipping high-bandwidth critical broadcast workflows. With EOS as the foundational architecture, MCS can infuse EOS's highly available, ultra-low latency, fault-tolerant, and high-performance capabilities into next-gen media workflows. Designed for use as a Network Control Solution, MCS complements the Broadcast industry's adaptation to IP by adding a layer of security, performance, and determinism to real-time on-prem and remote production workflows.

Overview

The requirements of broadcast real-time workflows have advanced during the last decade, from HD to UHD, and beyond. Production quality compressed and uncompressed media streams demand multi-gigabit data rates. On top of that, there are a large number of varied bandwidth multicast flows. To meet these ever growing requirements, the broadcast industry has arrived at an inflection point, where the migration from SDI to IP is the next logical step. Increased capacity, flexibility, and agility in managing broadcast workloads and workflows are all advantages of IP infrastructure. But, migration to IP has its own set of difficulties. When using SDI, a signal is compressed into a single flow, however when using IP, it is separated into at-least three signals (audio, video and ancillary). As a result, systems may quickly scale up to at-least 10,000 flows. Dynamic multicast routing protocols like PIM/IGMP alone would not be sufficient to support this magnitude of multicast and to provide broadcast operators with the same or better performance as SDI routing. This necessitates a cost efficient and high performance solution that aids the broadcast industry to successfully migrate and scale with IP.

Arista switches, which are built on the cloud networking principles of high performance, streaming telemetry, and analytics, will provide broadcasters with reliable, cost-effective, and highly scalable platforms to build and expand their highly dynamic workflows. On top of Arista's best-of-breed hardware and highly reliable and efficient self-healing EOS software, Multicast Flow Orchestration with MCS ensures that on-air broadcast operations are routed and secured as intended.

Most Common ST-2110 Topologies

For ST-2110 multicast distribution workflows, broadcast facilities commonly use Red/Amber and Blue networks. Spine/Leaf networks with Red/Blue networks or an exclusively Purple network with a common fabric are popular choices for on-prem broadcast facilities with distributed architectural needs.

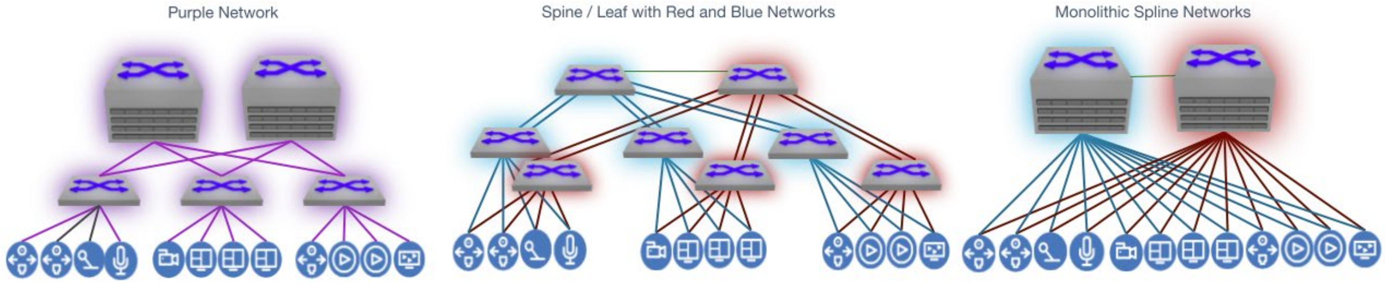

Figure 1: Most Common ST-2110 Topologies

Spline/Monolithic architecture is used in a space-constrained environment, such as OB Trucks, where there is a low necessity for a distributed network but a high requirement for multicast scale.

Hundreds, or even thousands of media endpoints connected to the IP network with significant bandwidth switching requirements exist in both of these situations. In the Spine/Leafs or Shared Fabric environment, there are also multiple links between the leafs and the spines. The signal bandwidth through these networks can range from a few Mbps and Kbps for audio and ancillary data, respectively, to many Gbps for video.

With the requirement of dynamic switching of senders to receivers, there is a need for a high performance bandwidth aware solution that can reliably program multicast flows in a cost and resource efficient manner.

Limits of IGMP and PIM

IGMP and PIM were heavily utilized to deliver multicast when broadcast workflows migrated to IP. The most significant disadvantage of these protocols is that they are not bandwidth or path aware, making them inefficient for programming and maintaining multicast flows in a live production environment. In these instances, the multicast flow distribution is very unpredictable, unequal, and nondeterministic.

The problem of indeterminism is magnified when a multicast receiver requests a signal via IGMP and the multicast group hashes to a link with insufficient bandwidth via the PIM Join. Even if PIM/IGMP distributes the flows randomly with no oversubscription, it will automatically redistribute the multicast flows in the event of a network link or device failure, which can result in oversubscription. This bandwidth-insensitive approach to automatic multicast flow redistribution could result in packets being discarded or dropped, which would be disastrous for important live broadcast workflows.

Furthermore, dynamic protocols do not support broadcast tallies natively. Without tallies, the broadcaster has no way of knowing whether or not the takes were successful without looking at the actual signal. Without accurate tallies, broadcast operators will be forced to operate in the dark for on-air procedures.

During a network link/switch failure, tally comes in handy because it alerts the operator about the network failures and which multicast flow is impacted and gives them a chance to recover important flows. This choice is essential for the operator since it allows them to determine which flow takes priority at any given time during broadcast.

PIM/IGMP does not have the capability to alert the operator of a failure; instead, it re-provisions the multicast flows to recover them. Automatic re-provisioning without alerting the operator not only strips away the operator's capacity to prioritize broadcast operations on demand, but it also increases the risk of losing a live broadcast . A live broadcast can be lost for two main causes in this case. One, because one leg of the network is functional in SMPTE-2022-7 flows, a network failure can go unnoticed without a notification. Second, automatic re-provisioning has the potential to take down another critical multicast stream that is currently in use on-air. Critical broadcast can be protected in either of these instances if there is an intelligent system that can alert the operator about failures and provide them control over recovering multicast flows with dynamic priorities that are only known to them at the time.

Path diversity for multicast flows is not supported by IGMP/PIM by default. The multicast traffic distribution in Purple/Shared Fabric networks may result in both legs of an SMPTE 2022-7 sender routed through the same spine, causing a single point of failure. In these workflows, path diversity is critical since it reduces failure domains and increases flexibility in the event of a failure or maintenance.

Broadcast workflows also demand fast, dependable, and deterministic multicast routing capabilities that mimic SDI router behavior. Multicast behavior that is unpredictable or slow degrades application performance and puts the Live Production environment at risk of becoming unresponsive. This necessitates the use of deterministic flow orchestration strategy to overcome the shortcomings of PIM/IGMP.

Arista MCS Overview

Arista MCS is a multicast flow orchestration solution designed with the many facets of broadcast workflow standards in mind, as well as a satisfying broadcast operator experience. Arista MCS acts as an interface between the control panel of the broadcast operator and the IP network. It facilitates the conversion of broadcast operations to network operations. Arista MCS manages and monitors real-time broadcast workflows over IP networks with a deterministic high-performance service and an easy-to-use API interface. Broadcast controllers integrate with this API interface to provision multicast flows required to support SMPTE ST2110, ST2022, AES67 and many other flow types.

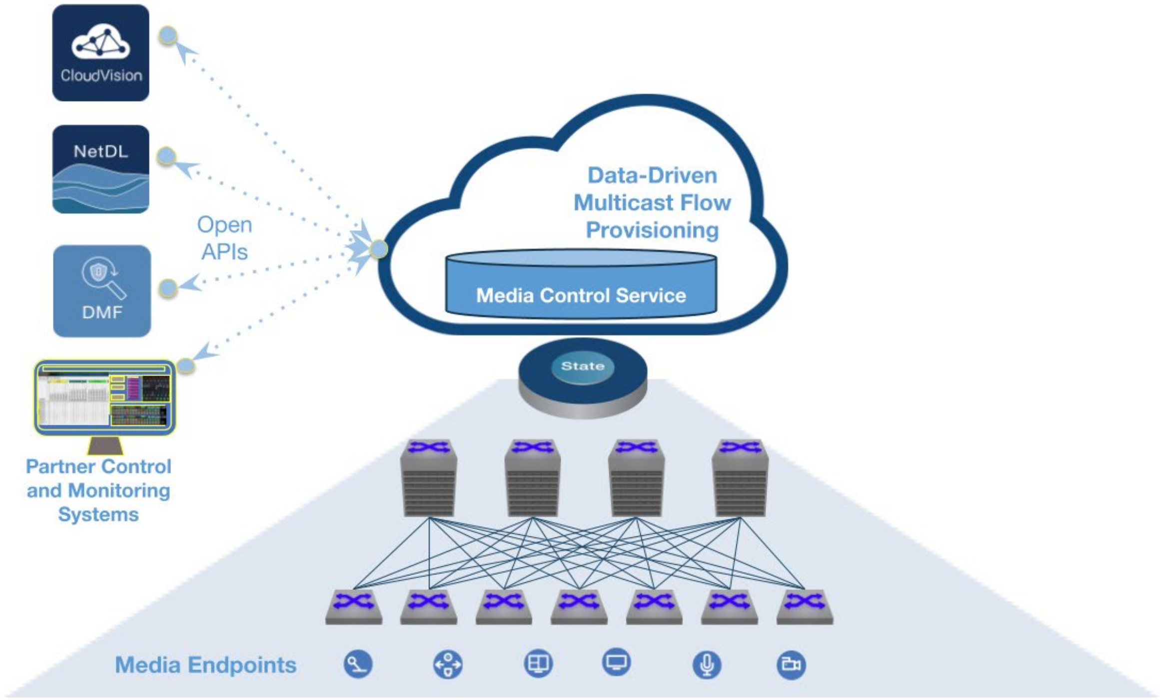

Figure 2: Arista MCS Overview

MCS abstracts topology and bandwidth information and accelerates multicast provisioning, allowing broadcast controllers to remain topology and bandwidth agnostic. MCS is a CloudVision eXchange(CVX) service at the architectural level. MCS takes advantage of CVX's ability to discover and aggregate real-time topological data and programs multicast configuration to the IP Network. MCS exposes two different types of APIs. REST-based APIs and Streaming based notification APIs.

The Broadcast controller integrates with both the API types. For network discovery and multicast provisioning, REST APIs are utilized, and for real-time network and multicast programming status/tally, streaming notification APIs are utilized. The broadcast controller uses REST APIs to POST multicast sender, receiver, and multicast stream bandwidth data to MCS upon integration. When MCS receives this data, it determines the optimum path to set up a multicast route connecting the receivers to the senders depending on bandwidth availability and path diversity. MCS enables near-parallel static multicast route programming and igmp snooping across L2/L3 interfaces. MCS additionally configures a rate policer for every multicast sender on ingress to protect against flows which have an incorrect bandwidth assignment attached to them.

MCS provides graceful and protected recovery modes for multicast flows re-establishing from failures or maintenance windows. It knows exactly where the multicast flows are programmed, what bandwidth is available on the links, and how healthy these multicast flows are because it consists of real-time network-wide streaming telemetry. MCS can consolidate this information in the event of a network link/device failure and provide a real-time notification to the operator regarding the network failure and which multicast receivers are impacted. It is critical to notify the operator about failing multicast receivers in a flow redundant ST-2110 deployment since they can go unreported until the entire feed is lost. Another benefit of notifying users about impacted receivers is that it allows the operator to prioritize which multicast receivers should be re-provisioned at that time. Furthermore, this functionality allows operators to undertake switch or link maintenance without disrupting on-air workflows. MCS can autorecover the multicast flow state once the switch / link health has been restored if the operator does not want to reprogram the flows. MCS can help with gracefully resyncing to the desired broadcast operator state, especially in the case of switch reboots. This is significant because, when a switch reboots, an operator can choose to reroute the affected flows to another switch. User mistakes with multicast configs that are written to the switch's start-up configuration can also occur. MCS restores the desired multicast status on the switch after the switch has restarted in these instances. If a multicast flow was deleted during the reboot, the status is reflected when the switch reconnects with MCS. MCS programmed multicast flows have a higher priority than other multicast flows in the system, therefore even if there are old, overlapping CLI-based static multicast streams, MCS flows will take precedence when the switch is rebooted, and traffic will be unaffected.

CloudVision eXchange (CVX) provides high availability for MCS-controlled broadcast processes. It has at least three controllers, one master, and two standbys deployed as a CVX cluster. Depending on the type of deployment, customers can deploy any number of clusters, each controlling its own subset of switches. In addition to HA, MCS features data redundancy, which is crucial for reducing downtime in broadcast operations. Because a CVX Cluster is set up with virtual IP, the broadcast controller will only have one point of contact with MCS.

If the Master MCS fails, CVX automatically selects the standby instances and selects one of them as the next Master. The network data is re-converged without any traffic loss once the new master is operational.
The health and multicast configuration state of MCS controlled switches are also streamed into CloudVision, which provides near real-time telemetry monitoring and troubleshooting. MCS users may examine and monitor multicast state on user-friendly dashboards with the custom-built and configurable CloudVision dashboards. MCS+CloudVision dashboards show top talker multicast sources, multicast flow paths, unsuccessful multicast flows, and the ability to filter and debug multicast flows using labeled mnemonics, such as Camera1 or SwitcherIn1.

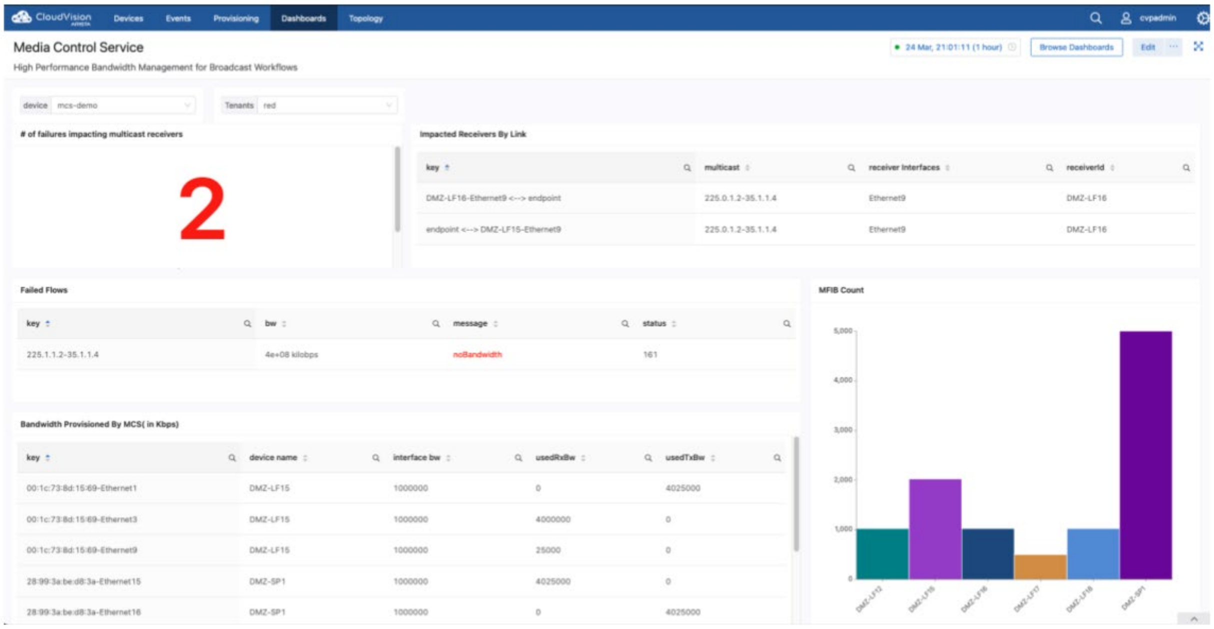

Figure 3: MCS Dashboards

Deployment Examples
Regardless of the deployment environment or workload scale, Arista's hardware and software portfolio offers a wide range of versatile and scalable features that can be used for broadcast and media workflows.

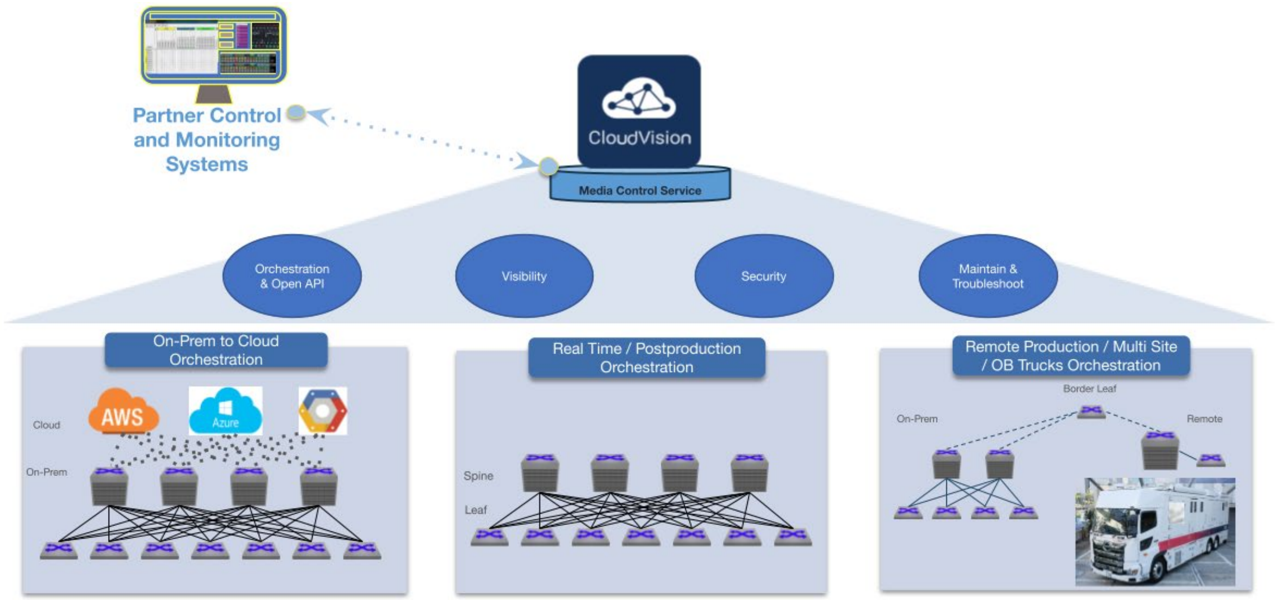

Figure 4: Arista MCS Across Deployments

Because the operators are using a single EOS image across all of these different environments, they can easily enable MCS just like any other service in EOS.

MCS can support orchestrating any of these workflows, whether it is enabling broadcast workflows for real-time/post production in on-prem environments, managing live telecasts across multi-site deployments to mobile OB Trucks, or migrating workloads from cloud to on-prem or on-prem to cloud. MCS can orchestrate multicast flows across a RED/AMBER and BLUE network or a purely purple network in any of these deployments in a bandwidth and path efficient manner.

MCS can coexist with PIM/IGMP flows, NAT-ed workflows, and interface across these environments in all of these installations. Customers can design and implement MCS into any setup that requires managing multicast traffic. It can also work in a hybrid system where both PIM/IGMP zones and MCS zones must be managed to send traffic between one another. In cross-zone setups, a border leaf switch is typically used to connect multicast traffic across these zones.The decision where MCS's boundary ends in relation to the IGMP Zone is dependent on the deployment requirements and is supported by MCS.

Sample Workflows with MCS API

When a broadcast controller is started, it connects to MCS's REST and notification APIs to discover the topology of the media endpoints. The broadcast controller also maintains a connection to the notification APIs in order to receive any further network state changes.The broadcast controller leverages MCS's APIs to program senders and receivers once the endpoint devices have been located, as demonstrated below.

```
curl --location --request POST 'http[s]://<mcs server i.p>/mcs/multicast/programFlows' \
--header 'Content-Type: application/json' \
--data-raw '[
    {
          "data": [
          {
                "destinationIP": "225.0.1.2",
                "sourceIP": "35.1.1.4",
                "bandwidth": 25,
```

```
                    "bwType": "m",
                    "inIntfID": "001c.7356.4f12-Ethernet12"
            }
            ],
            "flow-action": "addSenders",
            "transactionID": "GS#1",
            "trackingID": 2098
        },
        {
            "flow-action": "addReceivers",
            "transactionID": "test",
            "trackingID": 2099,
            "data": [
            {
                    "destinationIP": "225.0.1.2",
                    "sourceIP": "35.1.1.4",
                    "outIntfID": [
                    "001c.7357.6d61-Ethernet8"
                    ]
            }
            ]
        }
    ]'
```

In a high-performance near-parallel method, MCS analyzes the best available path and pushes the required multicast configuration to every switch. MCS configures a QoS policy on the sender interface, "001c.7356.4f12-Ethernet12", as well as static multicast routes on every switch required to provision the multicast flow from sender "001c.7357.6d61-Ethernet8" to receiver "001c.7357.6d61-Ethernet8". If senders/receivers are deployed as L2/SVI interfaces, MCS will also configure the necessary IGMP snooping statements, and multicast routes with the VLAN ID in the IIF or OIF[1]. MCS sends back a notification/tally regarding the success of flow programming after the hardware programming is complete. If a flow fails to be programmed for whatever reason, such as a lack of available bandwidth, MCS sends a notification or tally back to the broadcast controller indicating that the flow programming was unsuccessful.

MCS continuously monitors the flow's routability from senders to receivers after it has been successfully provisioned. If there are any link or switch failures along the flow path that affect the multicast traffic for this flow, MCS notifies the broadcast controller, informing the broadcast operator that the flow is impacted. This is increasingly essential in ST-2110 networks when data path redundancy has a means to disguise the failure on one side of the network.

All MCS-controlled network telemetry, flow programming status, and failures can be abstracted and visualized using custom-built MCS Dashboards on CloudVision, as well as exposed via REST/Notification APIs.

---

[1]IIF = Ingress interface

OIF = Egress interface

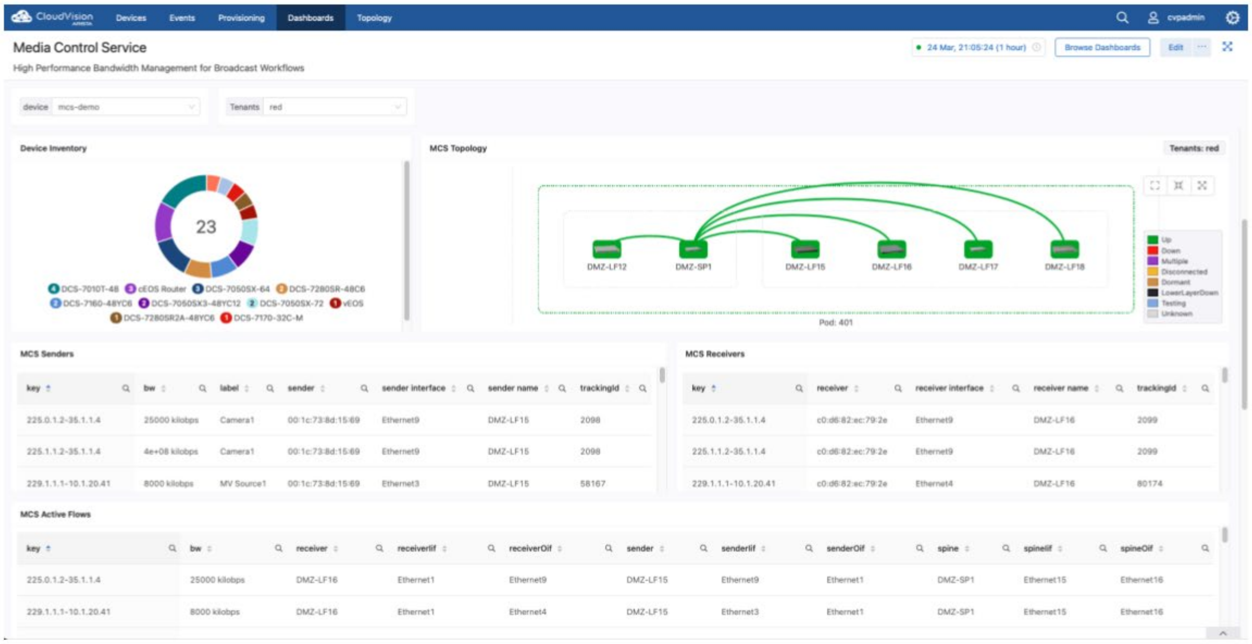

Figure 5: MCS Dashboards

Summary

Arista is enabling the media and entertainment industry in building a resilient, agile, and scalable architecture for transitioning and expanding their broadcast workflows. The requirement for a resilient and adaptable solution that can secure and transfer multicast flows with high determinism and performance . As broadcast workflows' demands and complexities grow, there is a growing requirement for a resilient and adaptable solution that can secure and transfer multicast flows in a highly predictable and performant way. Arista knows and understands broadcast operations and leads the industry with a highly open and programmable modern operating architecture. Arista MCS is built with the highest standards of cloud networking principles and broadcast resiliency in mind. MCS can reduce CapEx and OpEx while increasing production and programming capabilities by delivering consistent and superior performance across flexible leaf/spine, cloud, and remote deployments.

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, 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