WHITE PAPER

High-performance AI Infrastructure, Simplified

Boost AI training efficiency and reliability through Arista and Pure Storage networking and storage integration.

Contents

Introduction

Arista Networks and Pure Storage are collaborating to deliver a seamless and high-performance experience for networking and storage in the ever-demanding world of AI training and AI inference deployments. This paper is a high-level walkthrough of the key features, advantages, and business benefits that can be achieved by integrating Arista's industry-leading networking solutions with the innovative Pure Storage® platform and other services. The results are enhanced business outcomes and improved operational efficiency.

Engineering AI-ready Networks and Storage

Artificial Intelligence (AI) is advancing and changing all industries. Whether your core business is self-driving cars, groceries, biosciences, direct patient healthcare, or retail banking, elements of recent AI developments are likely to infiltrate it.

These AI applications introduce new workloads to your infrastructure. They impact every technology pillar within your data centers, including power, cooling, racking, cabling, networking, storage, and security.

Arista Networks and Pure Storage engaged in a project to jointly test and deploy networking and storage solutions that address the demanding needs of AI accelerator clustering.

Arista Networks software and hardware platforms are the industry’s best solution for these demanding workloads. Arista has been tested, selected and deployed in some of the largest accelerator clusters in the world, ranging in size from 2,000 accelerators to 200,000, with even larger clusters under development. Customers choose Arista Networks for these projects for several reasons, including Arista’s unrelenting focus on software quality, hardware resiliency, and the litany of features that aid in the operation of these complex clustered systems.

The current best practices applied to these large XPU clusters call for building several discrete networks. These include a frontend (FE) network, a backend (BE) network, and an out-of-band network (OOB) for management. Some use cases call for an additional dedicated network specifically used for connectivity to storage resources.

The FE network is primarily concerned with providing connectivity for resources such as the CPU, storage, In-Band management and providing the conduit for users to interact with the cluster assets. Contrast this with the BE network, which is solely dedicated to XPU to XPU connectivity. Because of the performance requirements of accelerator-to-accelerator interaction and the intent to avoid any delay in job completion time or the inducement of wait time that could lead to idle XPUs, the current best practice is to eliminate all other data flow in the BE to strictly just XPU to XPU data streams.

Load balancing is arguably the most difficult problem faced by the networking industry today, as is providing optimal load balancing that ensures full utilization of all available bandwidth and fast failure detection and rerouting. Arista employs several key load-balancing features, including Dynamic Load Balancing (DLB) and Collective Load Balancing (CLB).

Pure Storage FlashBlade®, built on a “distribute everything” architecture, ensures that data is evenly distributed across all blades in the array. When combined with Arista’s advanced load balancing capabilities, it creates a highly efficient, distributed I/O pipeline that delivers a competitive edge for AI workloads. This architecture excels in scenarios involving concurrent read/write operations across varied data sizes and handling metadata, with tens of millions of file operations per second. Furthermore, FlashBlade is five times more space- and power-efficient than competing solutions, providing substantial energy and cost savings that can be reinvested into GPUs to accelerate performance further.

Benefits for AI clusters built on Arista and the Pure Storage platforms include:

- **Insight**: Network traffic indicators and contextual data enable faster resolution time, preventing potential bottlenecks and hot spots
- **Stability**: A distributed storage architecture ensures consistent performance, even during unplanned disruptions, safeguarding against downtime
- **Improved training**: Reducing unforeseen congestion and minimizing maintenance time accelerates time to results for AI training workloads
- **Scalability**: Seamless, zero-impact scalability from FlashBlade allows the storage array to grow alongside AI demands, ensuring transparent expansion without disrupting the AI cluster
- **Value**: Predictable, consistent performance as capacity expands provides a more reliable path to achieving faster, more efficient AI training results

The Pure Storage and Arista Hardware Integration

Quality of Service

Priority flow control (PFC) is a hop-by-hop mechanism that temporarily pauses data transmission to prevent packet loss during network congestion. As defined in the IEEE 802.1Qbb standard, PFC enables link-level flow control, allowing each device to independently manage traffic flow based on the IEEE 802.1p class of service or priority level.

This quality of service (QoS) feature enables differentiated traffic treatment, prioritizing critical I/O and preventing disruptions while allowing non-critical, loss-tolerant traffic to be dropped. PFC achieves this by sending per-priority pause frames to the upstream link partner, signaling that the local buffer is nearly full and requesting the sender to pause transmission for a specified period.

In the event of sustained congestion, each device will issue PFC pause frames to its upstream link partner, creating back pressure that eventually prompts hosts to throttle traffic until network conditions improve.

Explicit congestion notification (ECN) is an IP and TCP extension that provides end-to-end network congestion notification without dropping packets. It recognizes early congestion and sets flags that signal affected hosts. These flags are then read by and passed along through the Arista switches, ultimately reaching the furthest endpoints, which are both the sources and destinations of this traffic. ECN usage requires that it be supported and enabled by all endpoints.

To summarize, Arista's QoS features optimize AI workloads by:

- Prioritizing AI traffic for low-latency, high-bandwidth performance
- Implementing advanced congestion control (PFC, ECN) for lossless networks
- Providing real-time analytics for AI-specific traffic patterns
- Enabling intelligent load balancing for distributed AI computations
- Offering RDMA-aware QoS for efficient GPU-to-GPU communication

These capabilities ensure AI applications receive optimal network resources, resulting in faster processing, reduced job completion times, and improved overall performance for end users leveraging AI-driven services.

Visibility

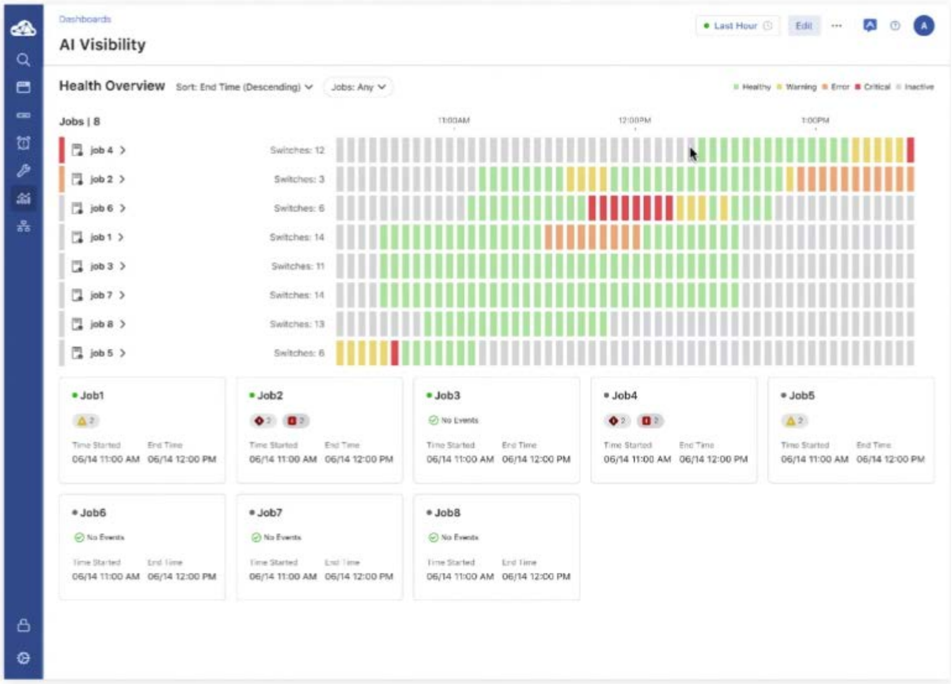
A network operator can’t find and resolve performance issues if they can’t see them. Arista provides several unique, industry-leading visibility, observability, and reporting solutions that aid in an operator’s ability to quickly identify if the network's performance is inducing delay to the job completion time. This visibility also reports on another key metric in these clusters, the TSN (Time Spent in Networking)

Arista Latency Analyzer (LANZ) is a unique tool built directly into Arista’s network operating system, EOS, providing microburst detection and monitoring of interface and queue contention. It allows an administrator an unprecedented ability to visualize events typically occurring at granular rates of the nanosecond and microsecond levels.

LANZ allows the user to customize thresholds for triggering hardware alerts. It works at the byte level to capture reporting data pertinent to the system's overall performance, and can be configured to alert on the smallest levels of interface contention, instrumenting the granularity of just one packet queued behind another. This alerting is made available in real-time via CLI and streaming telemetry to Arista’s award-winning management and configuration tool, CloudVision Portal (CVP).

Building upon the incredibly granular visibility of LANZ is the collection and rendering of tens of thousands of performance-specific data points in our CloudVision Portal product. CloudVision can be installed on-premises or hosted and managed by Arista via CVaaS, CloudVision as a Service. CloudVision consists of two distinct job aids. CVP provides configuration management that includes managing all the way down to the device interface configuration level and continuous vulnerability checks against CVE and EOS software bug databases. These built-in tools save operators thousands of hours by avoiding manual hunting for problems and the avoidance of having to perform manual bug scrubs.

Along with the configuration management side of CVP is the Network Operations aspect, which provides real-time information about the performance of every process, interface, MAC address table, routing protocol neighbor states, and IP routing table on each switch. This incredibly granular collection and rendering of state, counters, and performance shortens the time it takes an operator to determine if and where there are bottlenecks or service degradation in the networking infrastructure.



AI Visibility

Health Overview Sort: End Time (Descending) Jobs: Any

Jobs | 8

job 4 Switches: 12
job 2 Switches: 3
job 6 Switches: 6
job 1 Switches: 14
job 3 Switches: 11
job 7 Switches: 14
job 8 Switches: 13
job 5 Switches: 6

Job1
⚠ 2
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

Job2
◆ 2 ■ 2
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

Job3
✓ No Events
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

Job4
◆ 2 ■ 2
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

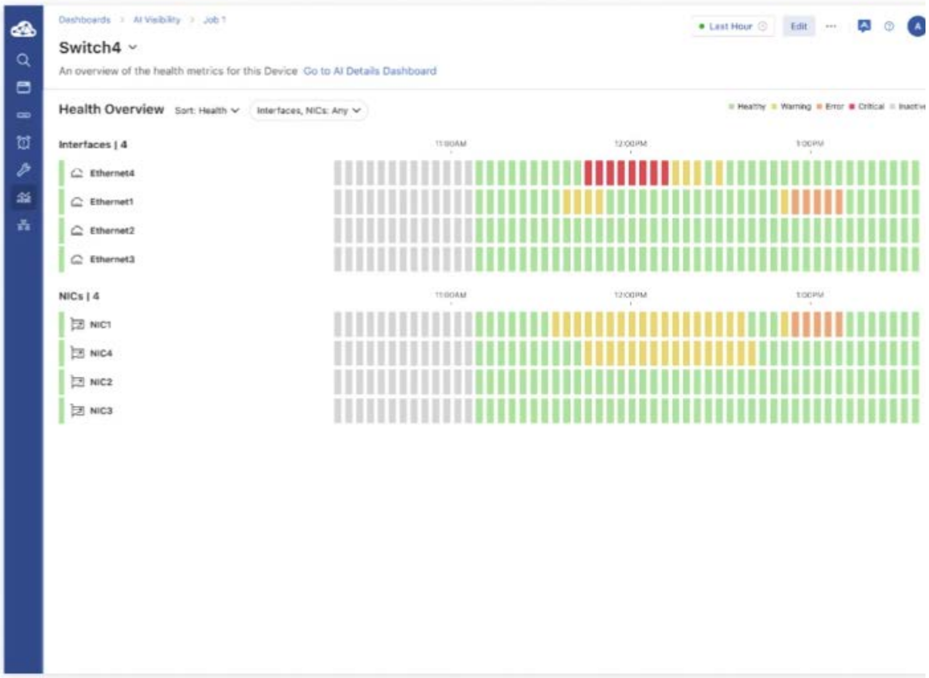
Job5
⚠ 2
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

Job6
✓ No Events
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

Job7
✓ No Events
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

Job8
✓ No Events
Time Started 06/14 11:00 AM
End Time 06/14 12:00 PM

FIGURE 1 AI visibility main dashboard



An overview of the health metrics for this Device Go to AI Details Dashboard

FIGURE 2 AI device interface page

AI Hardware Integration Testing Overview and Goal

The overall testing scenario was created to run NFSoRDMA data traffic using Arista switches and a Pure Storage FlashBlade array. Tests were designed to run various workload types against the FlashBlade array that would vary the block sizes used, iodepth, and job type.

The goal was to perform standard IO tests on the FlashBlade system and Arista switches and monitor the data traffic from them. Arista's LANZ traffic flow view of RoCEv2 allows customers to better understand how data flows on the network are performing and identify potential congestion issues early.

Test Tools and Configuration

FIO

FIO jobs are run in parallel from the sixteen CPU nodes to a single FlashBlade file system export. PDSH was used to distribute the FIO jobs to each CPU node in the cluster.

The test parameters for the FIO ranges were set in advance in the Python script that ran the FIO jobs on each compute node via PDSH:

```
# Test parameter ranges

job_type = ['read', 'randread', 'write', 'randwrite', 'randrw']

io_size = ['4k', '12k', '64k', '1m']

num_jobs = ['12']

io_depth = ['1', '2', '4', '8', '12', '16', '20', '24', '32', '40', '48', '56', '64']
```

Test phases:

Phase	Target	Description
Fill the FlashBlade	90%	FIO Fill job
Baseline	FIO	Run Baseline
Reduce Consumption	73%	
Age data	FIO	Run Baseline test against existing data from fill job twice.
Remove DFM	FIO	Run randrw job while powering off DFM
Remove Blade	FIO	Run randrw job while powering off Blade

TABLE 1 Test phases

RoCE Congestion Policy Configuration

All initiator compute nodes were using the same RoCE policy configuration.

Configuration Type	Settings
PFC Setting	mlnx_qos -i eth4 --pfc 0,0,0,1,0,0,0,0
DSCP Setting	mlnx_qos -i eth4 --trust dscp
TOS	cma_roce_tos -d mlx5_0 -t 106

TABLE 2 Configuration settings

Cluster Details

The test cluster is managed using BaseCommand Manager from Nvidia. All images are delivered to the CPU host hardware via the PXE boot process.

NVIDIA BaseCommand Manager:

Component	Version
Cluster Manager	10.0
CMDaemon	3.0
CMDaemon Build Index	157769

TABLE 3 Component versions

Cluster Size:

Component	Size
CPU Nodes	16
FlashBlade Array	1

TABLE 4 Component sizing

Arista Switch Hardware

Two DCS-7060DX5-64S-F switches are installed in the cluster environment, providing all connectivity between the cluster hosts and FlashBlade//S array.

Component	Version/Value
Hardware version	11.03
Software image version	4.31.1F
Image format version	3.0
Image optimization	Default

TABLE 5 Hardware

Arista Switch Configuration

The Arista switches were configured as follows:

```
!
queue-monitor length
!
queue-monitor length log 5
!
platform trident mmu queue profile RoCELosslessProfile
   ingress threshold 1/16
   egress unicast queue 3 threshold 8
!
!
qos map dscp 0 to traffic-class 0
qos map traffic-class 0 to dscp 0
qos map traffic-class 3 to dscp 26
!
platform trident mmu queue profile RoCELosslessProfile apply
!
!
sflow sample 10000
sflow polling-interval 5
sflow destination 127.0.0.1
sflow source 10.0.6.2
sflow run
!
!
queue-monitor streaming
   vrf mgmt
   no shutdown
!
!
interface Ethernet1/1
   switchport access vlan 110
   qos trust dscp
   priority-flow-control on
   priority-flow-control priority 3 no-drop
   !
   uc-tx-queue 3
      bandwidth guaranteed 20000000
      random-detect ecn minimum-threshold 150 kbytes maximum-threshold 1500 kbytes max-mark-probability 100
weight 0
!
Alpha 1 ECN 256K
uc-tx-queue 3
      no priority
      bandwidth percent 100
      random-detect ecn minimum-threshold 256 kbytes maximum-threshold 512 kbytes max-mark-probability 100
weight 0
Alpha 1 ECN 50K
qos profile ECN-Only
   uc-tx-queue 3
      no priority
      bandwidth percent 100
      random-detect ecn minimum-threshold 50000 segments maximum-threshold 50000 segments max-mark-probability
100 weight 0
```

FlashBlade//S Array Hardware

The FlashBlade system was configured as follows:

Component	Version/Value
FlashBlade//S500	1
Purity Version	4.5.2
XFM8400	2
Chassis	1
Blades	9 (expandable to 50 blades in 5 chassis)
48TB DFM's	40
400Gb/E	16

TABLE 6 FlashBlade configuration.

Hosts

The hosts were configured as follows:

Component	Version/Value
R6515	16
CPU	64 core
RAM	512GB
CX-7 NIC (200Gb/E)	1

TABLE 7 Host configuration.

Storage and Network Architecture

The storage and network architecture is illustrated below.

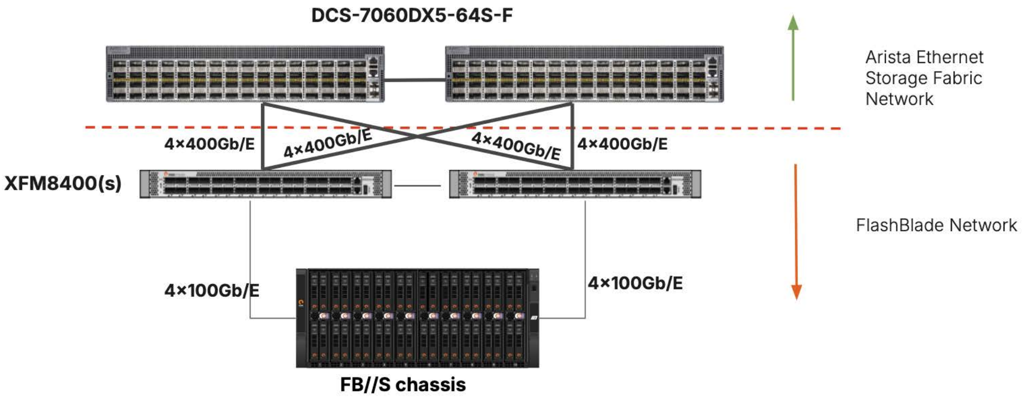


FIGURE 3 storage and network architecture

Test Cluster Architecture

The test cluster architecture is illustrated below.

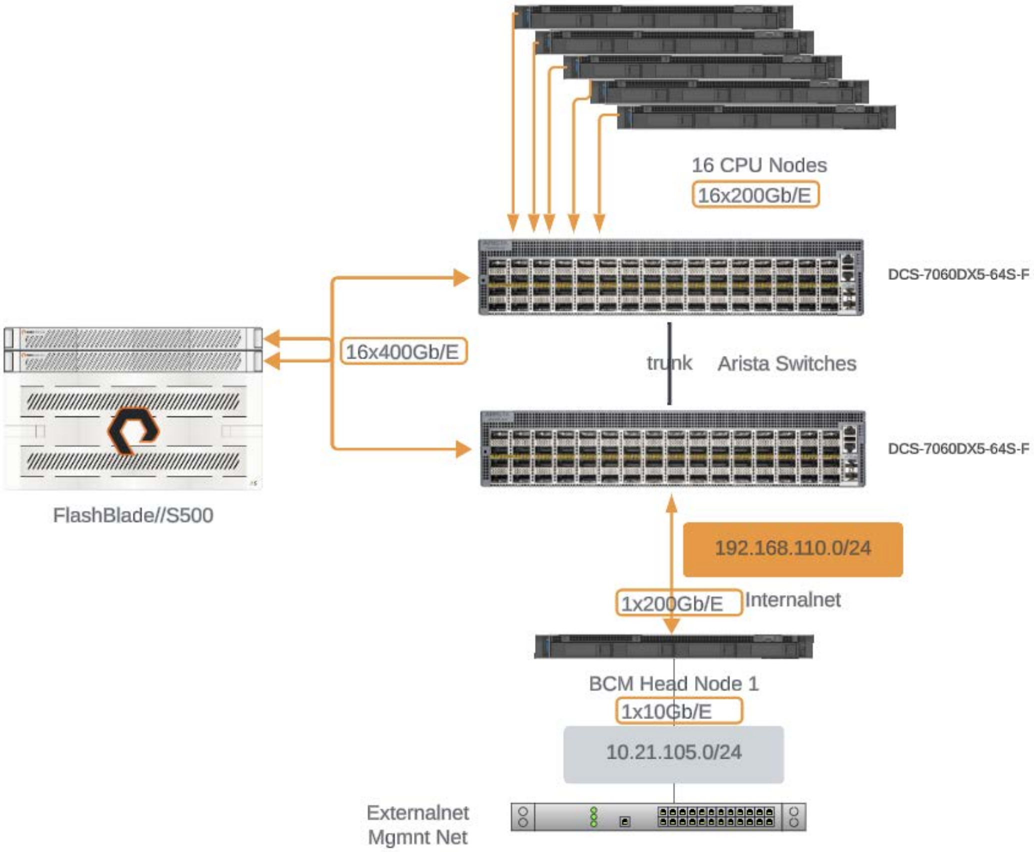


FIGURE 4 Test cluster architecture

Insights and Observations

Key takeaways on the Pure Storage and Arista AI hardware integration include:

- Consistent performance on variable workloads
- Zero performance degradation even at 90% capacity, resulting in cost savings
- Performance was maintained, even after reducing the capacity (blades/drives/dfms) of the system
- No loss in data (visible through the Arista Network dashboard) when the capacity is brought back up or during the reduction stage

Working alongside Pure Storage, Arista deployed switches from the 7060X5 series and an instance of CVaaS to measure the performance of several RDMA (Remote Direct Memory Access) data transfers. In a typical AI cluster, RDMA is the primary method of data transfer. These data packets are encapsulated for transport on an Ethernet-based network into the ROCEv2 (RDMA over Converged Ethernet) format using UDP 4791.

Baseline Data Job

Based on the FIO job examples in Figure 1, FIO data aging workloads were generated on the FlashBlade after the system reached 90% capacity. Running variable workloads on a storage array that is 90% full can introduce significant challenges for traditional, network-based storage platforms, such as performance degradation or I/O bottlenecks. In contrast, the “distribute everything” architecture in FlashBlade is designed to maintain consistent throughput and I/O performance, even at high levels of capacity saturation. This ability to deliver linear and predictable performance at full utilization means training jobs remain uninterrupted, ensuring operational efficiency. By fully leveraging the array’s capabilities, organizations can maximize their hardware investment, avoiding the need for costly over-provisioning or additional storage purchases, resulting in both cost savings and enhanced ROI.

In the example below, you can observe the correlation between IOPS and data throughput. As smaller I/O size jobs were executed, IOPS values were higher, while throughput increased as the I/O size grew. The breakdown of read versus write operations is also shown, highlighting the impact of different job types in the workload loop.

The below graphs show FlashBlade baseline array at 90% full capacity.

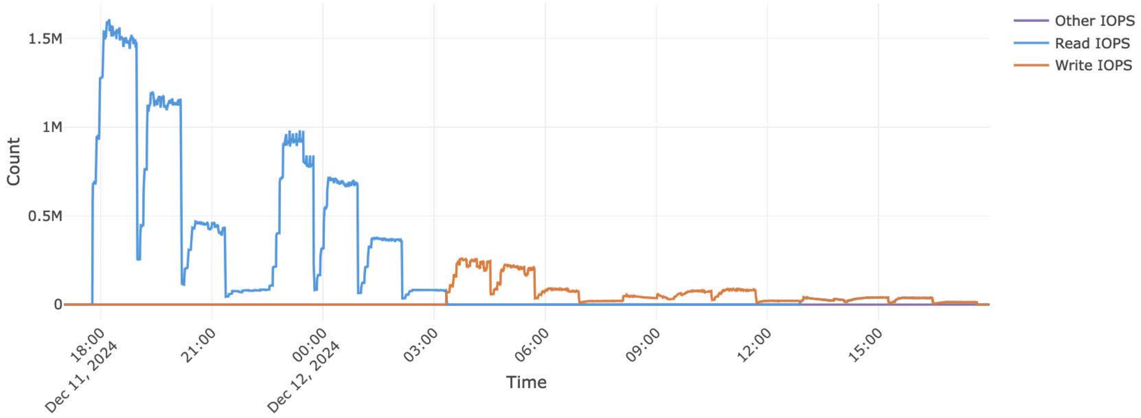


FIGURE 5 Count

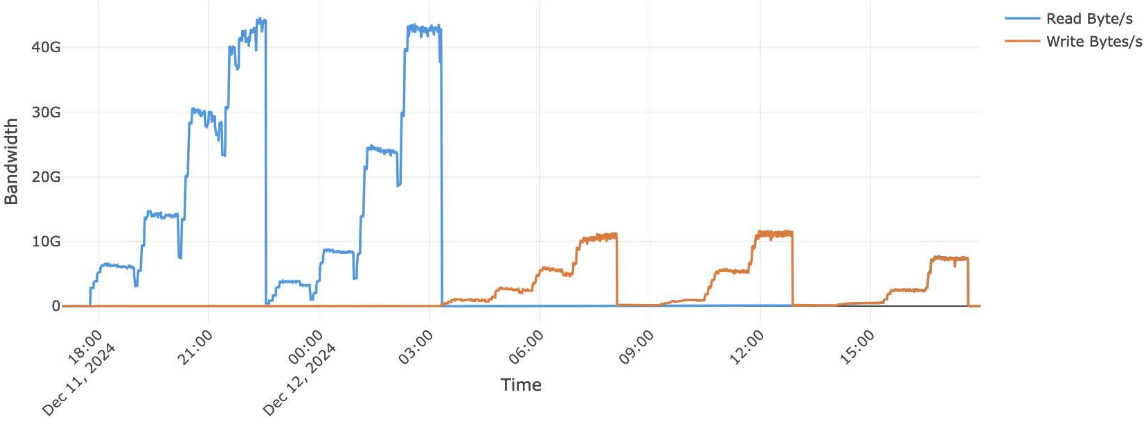


FIGURE 6 Bandwidth

Age Data Job

Reducing the array's capacity and re-running the same FIO jobs to age the data produced the same linear and predictable performance observed at 90% capacity. The jobs were executed twice to further validate, ensuring the FlashBlade filesystem files were modified multiple times, including both ingest loads and reads from the same file. Traditional storage systems, especially those reliant on caching or centralized architectures, often experience performance degradation as requests overload the cache front end. In contrast, FlashBlade was designed to avoid such pitfalls from the outset. By leveraging a distributed key-value store accessible by all blades and distributing traffic across the array at the blade level while maintaining a single IP for data traffic, FlashBlade delivers consistent performance, operational efficiency, and unmatched simplicity.

The below graph shows the FlashBlade aging data run 1 at 73% array capacity

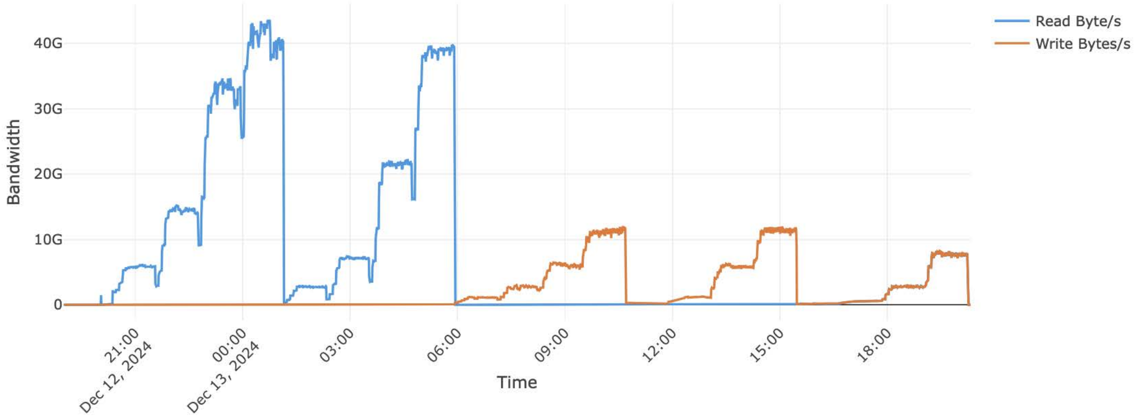


FIGURE 7 Bandwidth

The below graphs show the FlashBlade aging data run 2 at 73% array capacity.

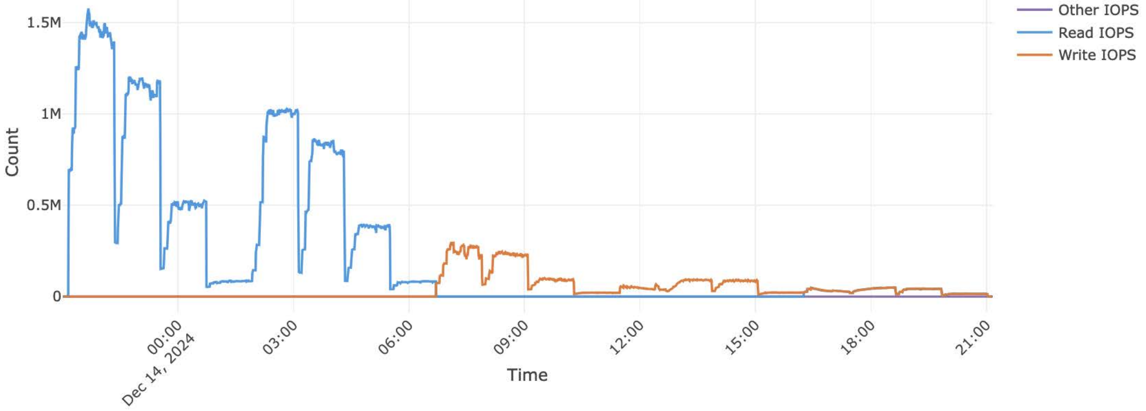


FIGURE 8 Count

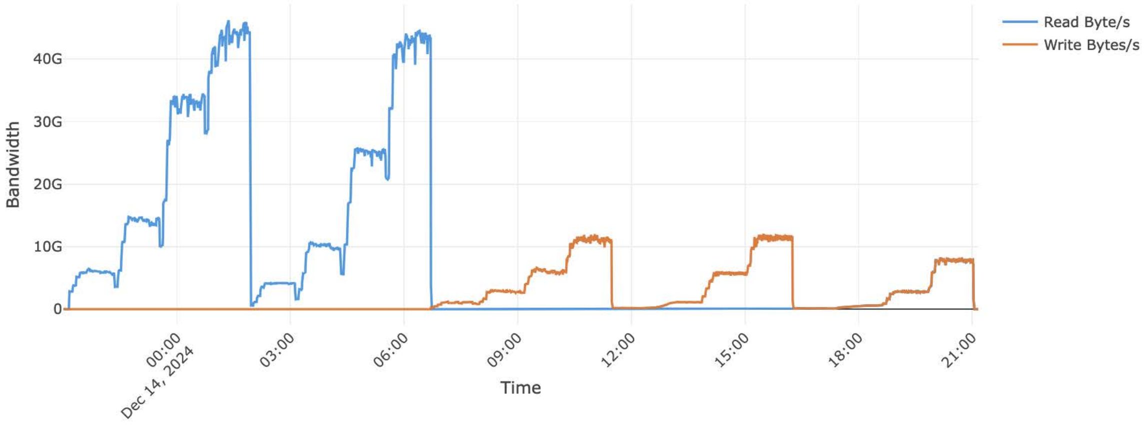


FIGURE 9 Bandwidth

Remove DFM Job

To further stress the array, both a DirectFlash Module® (DFM) and a single blade containing four Direct Fabric Modules were removed from the FlashBlade during a 50:50 random read/write FIO job, which was run against existing files on the array. This test was conducted with the array at 73% storage capacity. The primary goal was to assess the array's ability to maintain performance during these events. In the first example below, you can see the single DFM removed at the 21:08 mark. While there was a slight reduction in total bandwidth, the array quickly adapted and continued serving data without significant disruption.

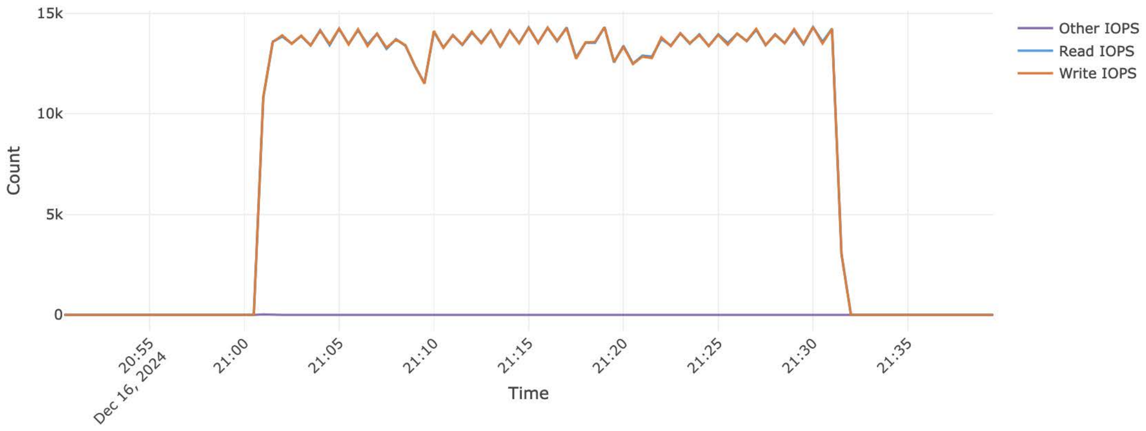


FIGURE 10 Count when DFM was pulled at 21:08.

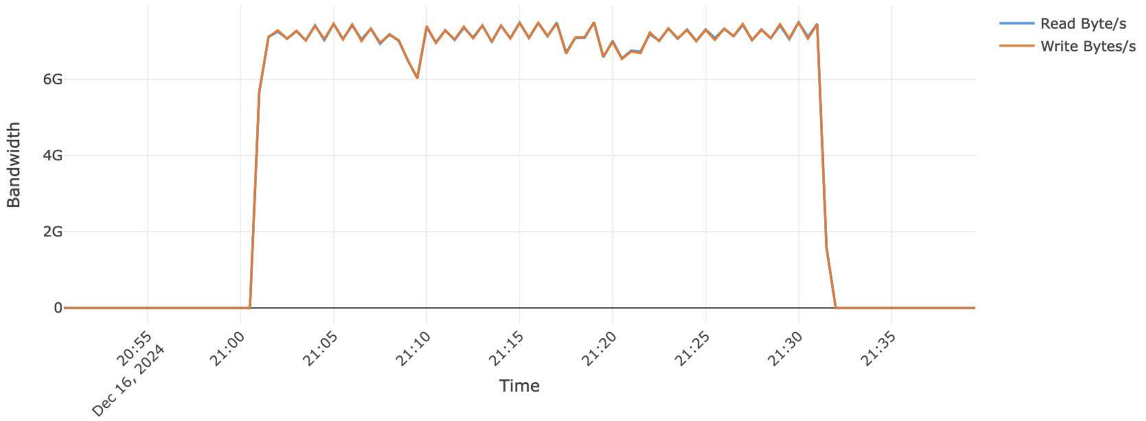


FIGURE 11 Bandwidth when DFM was pulled at 21:08.

Remove Blade Job

In this example, we removed a blade from the FlashBlade, which also removed all associated compute and storage resources (four DFMs) while a 50:50 random read/write job was running. The blade was removed at timestamp 18:30, and the impact was observed shortly thereafter.

Of note is the linear and predictable performance of the FlashBlade, even when an entire blade was removed. Despite a brief reduction in overall throughput, the array maintained consistent performance, continuing to serve data without interruption. This resilience demonstrates the ability of FlashBlade to adapt dynamically to hardware changes, ensuring seamless operation even under stress.

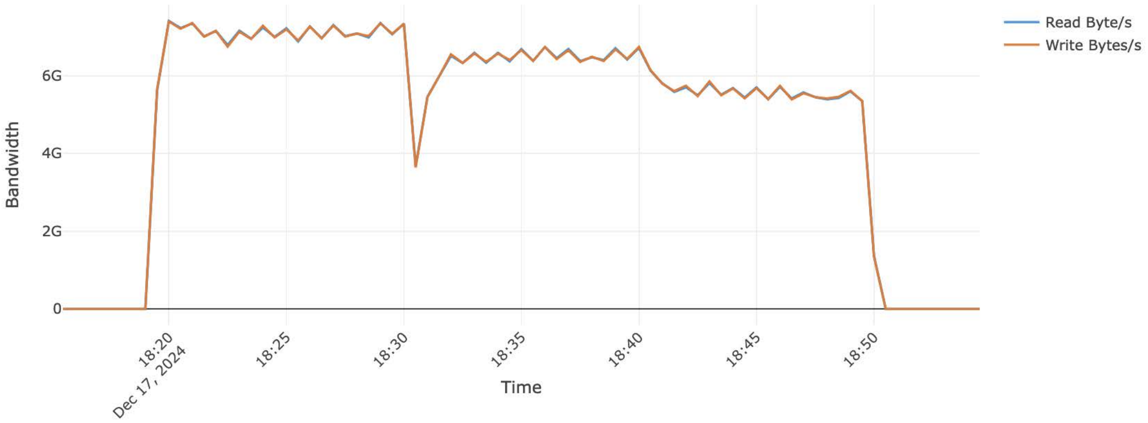


FIGURE 12 Bandwidth with blade removed at 18:30.

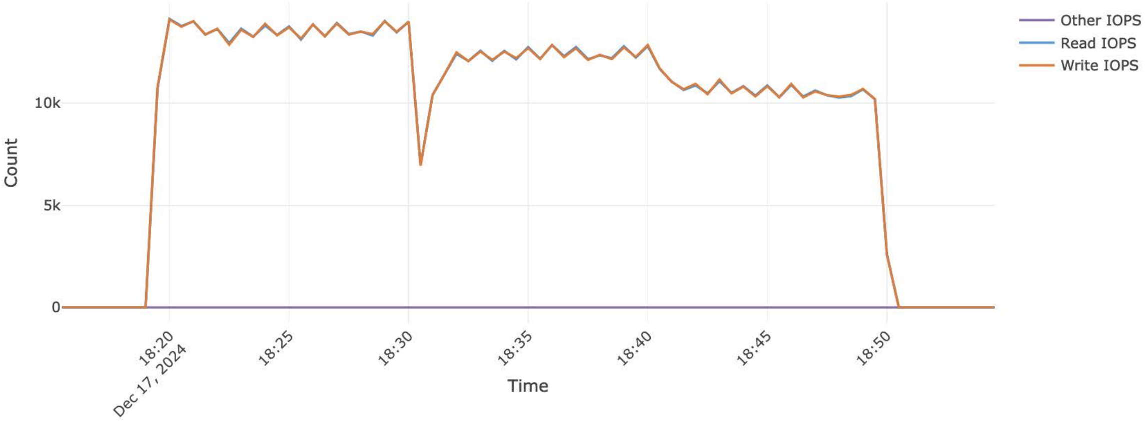


FIGURE 13 Count with blade removed at 18:30.

In both test runs, the single DFM and single blade with four DFMs were reinstalled into the array without downtime or outages. The array was simplified, and the DFM and blade were added back to the cluster. No outage window was required, and there was no downtime.

Looking at the Arista logs via CloudVision for the same time period, we can see the associated network traffic and observe what is happening with the RoCE packets. In this example, we look at the traffic from the initiator side or host.

This graph shows the load coming in and that it generated PFC Frames. However, the big call out here is that there were zero drops. The switch was able to handle the burst:

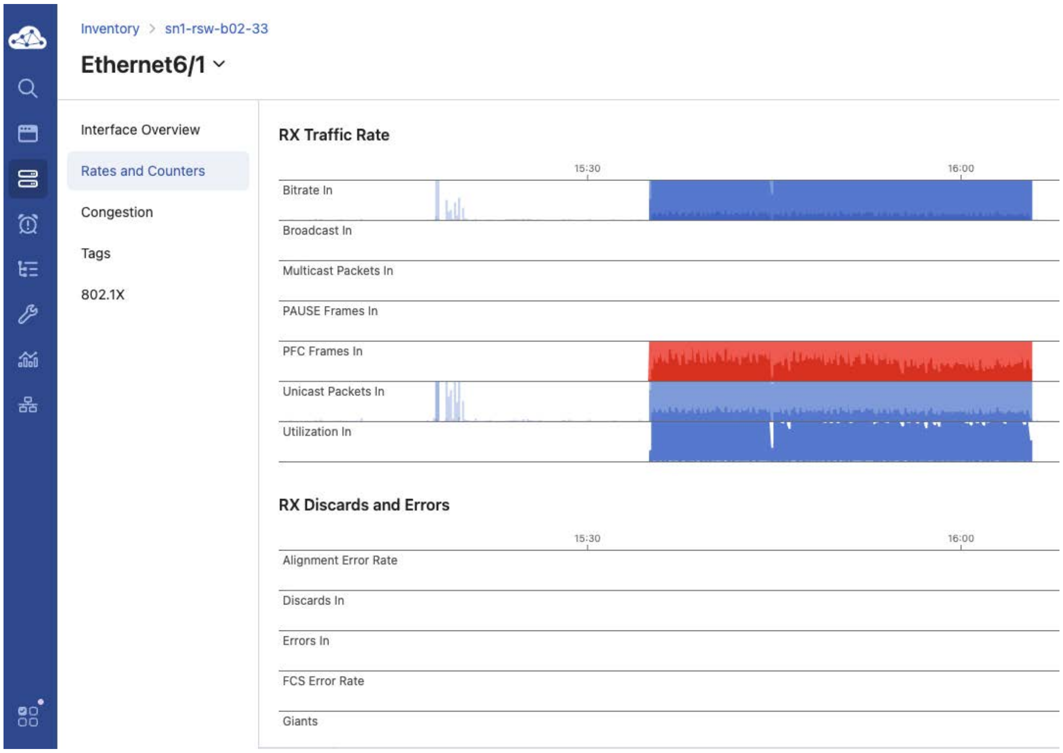


FIGURE 14 Network Traffic

Another graph showing the same time frame but focuses on congestion. Again, zero drops.

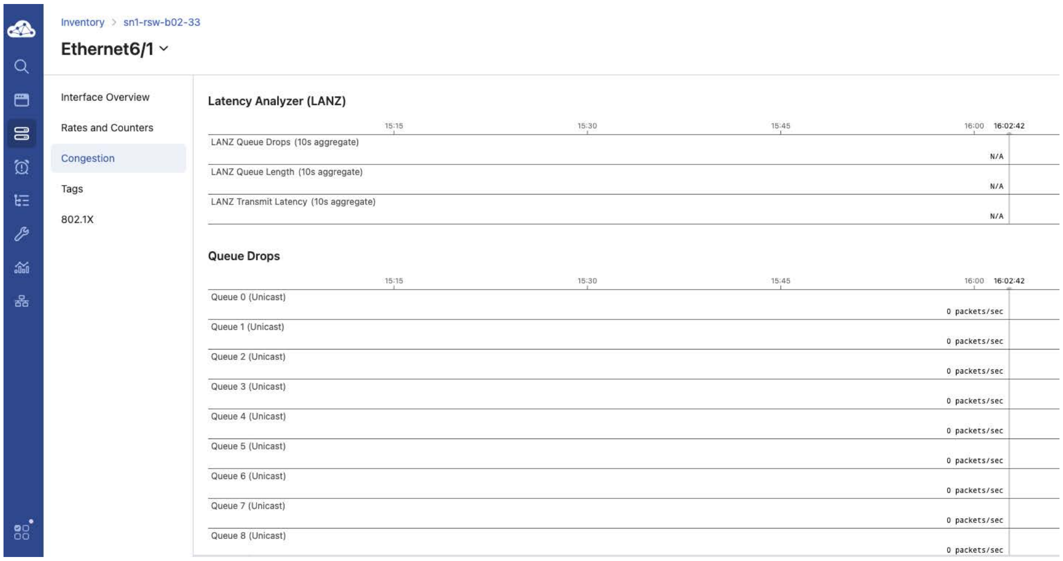


FIGURE 15 Packet drop rates

The next series of graphs are the 400GbE uplink ports at the same time. This shows load sharing and that the 400GbE aren't overloaded.

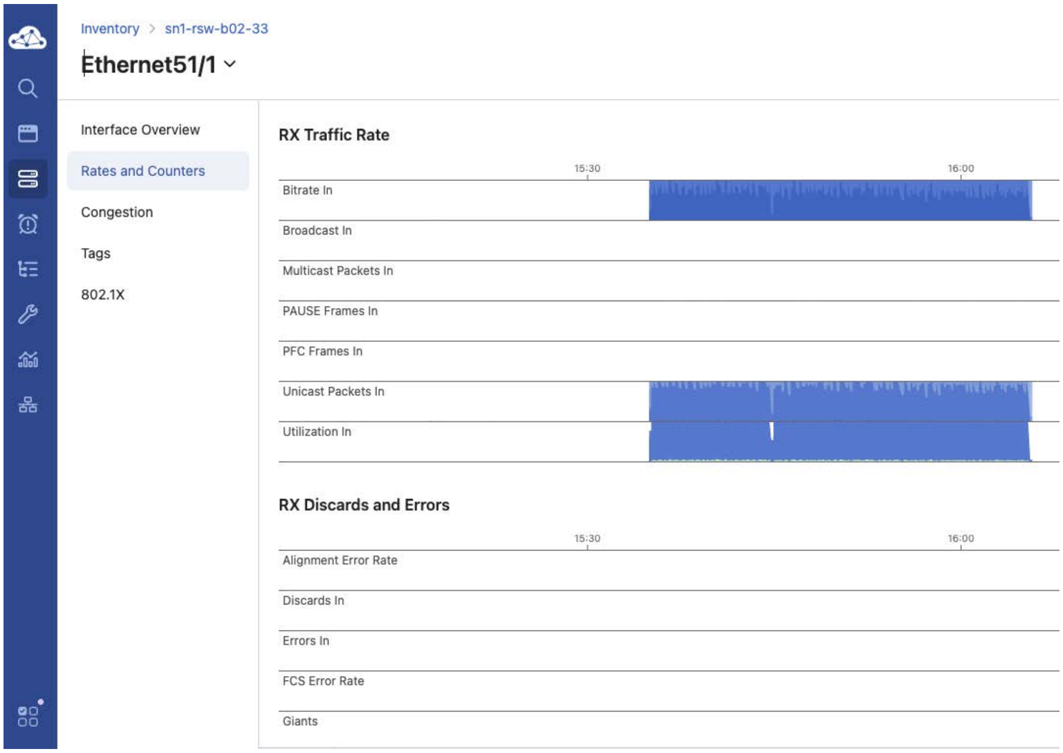


FIGURE 16 Load share 51/1

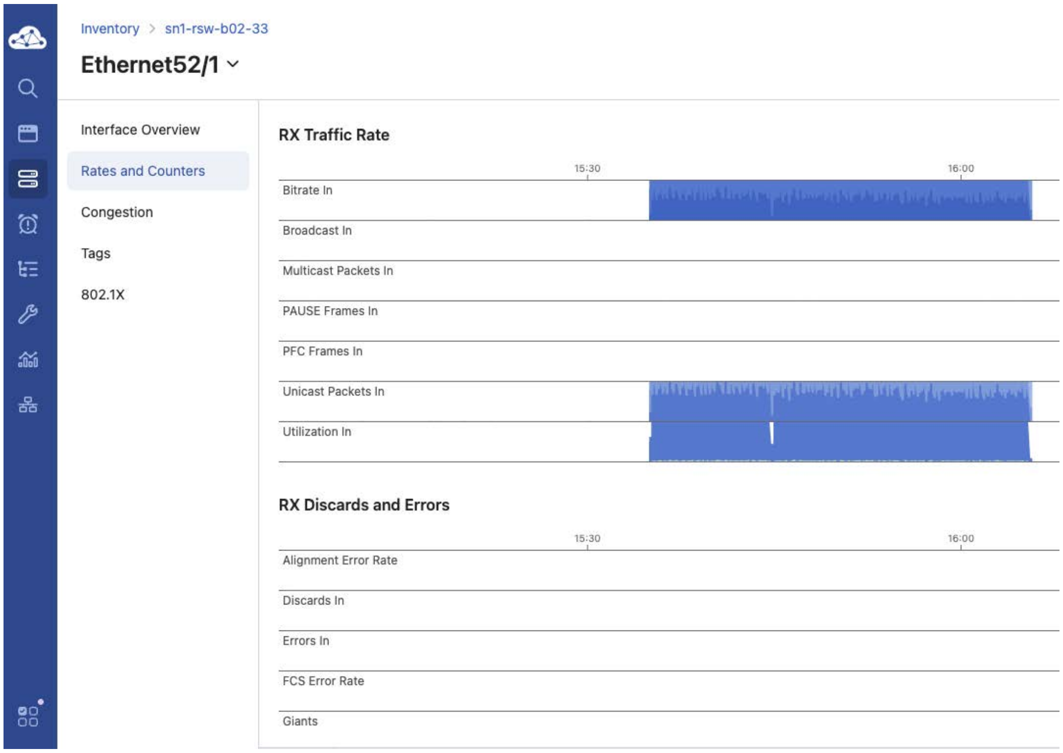


FIGURE 17 Load share 52/1, inputs from Arista

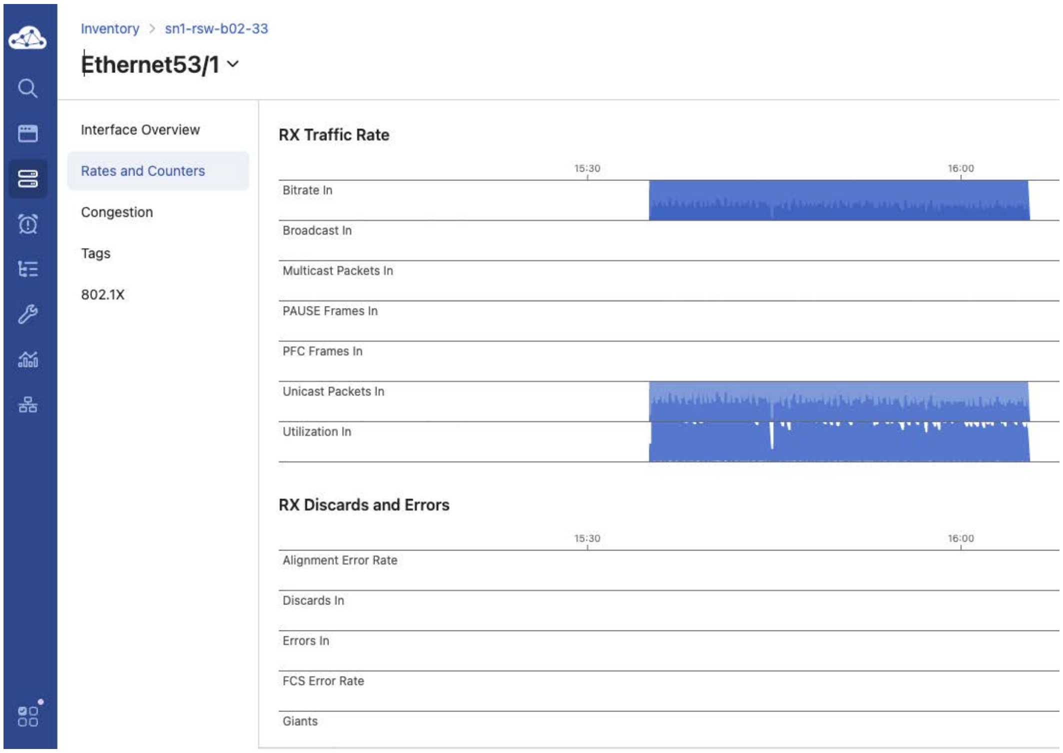


FIGURE 18 Load share 53/1.

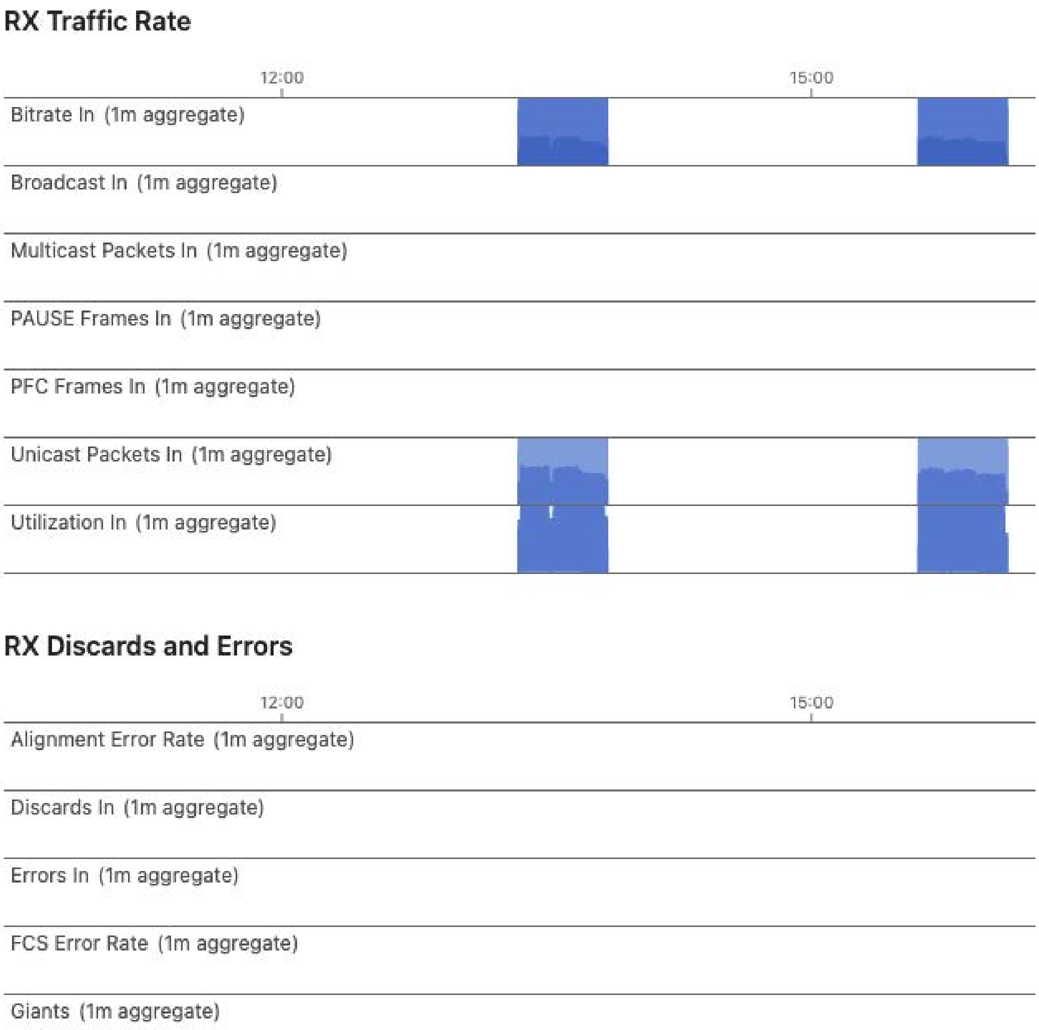

RX Traffic Rate
12:00
15:00
Bitrate In (1m aggregate)
Broadcast In (1m aggregate)
Multicast Packets In (1m aggregate)
PAUSE Frames In (1m aggregate)
PFC Frames In (1m aggregate)
Unicast Packets In (1m aggregate)
Utilization In (1m aggregate)
RX Discards and Errors
12:00
15:00
Alignment Error Rate (1m aggregate)
Discards In (1m aggregate)
Errors In (1m aggregate)
FCS Error Rate (1m aggregate)
Giants (1m aggregate)

FIGURE 19 Traffic rates and errors

During that same time period, CVP provides details about any queuing, delay, or, worst case, discards, that may have occurred when the load was high.

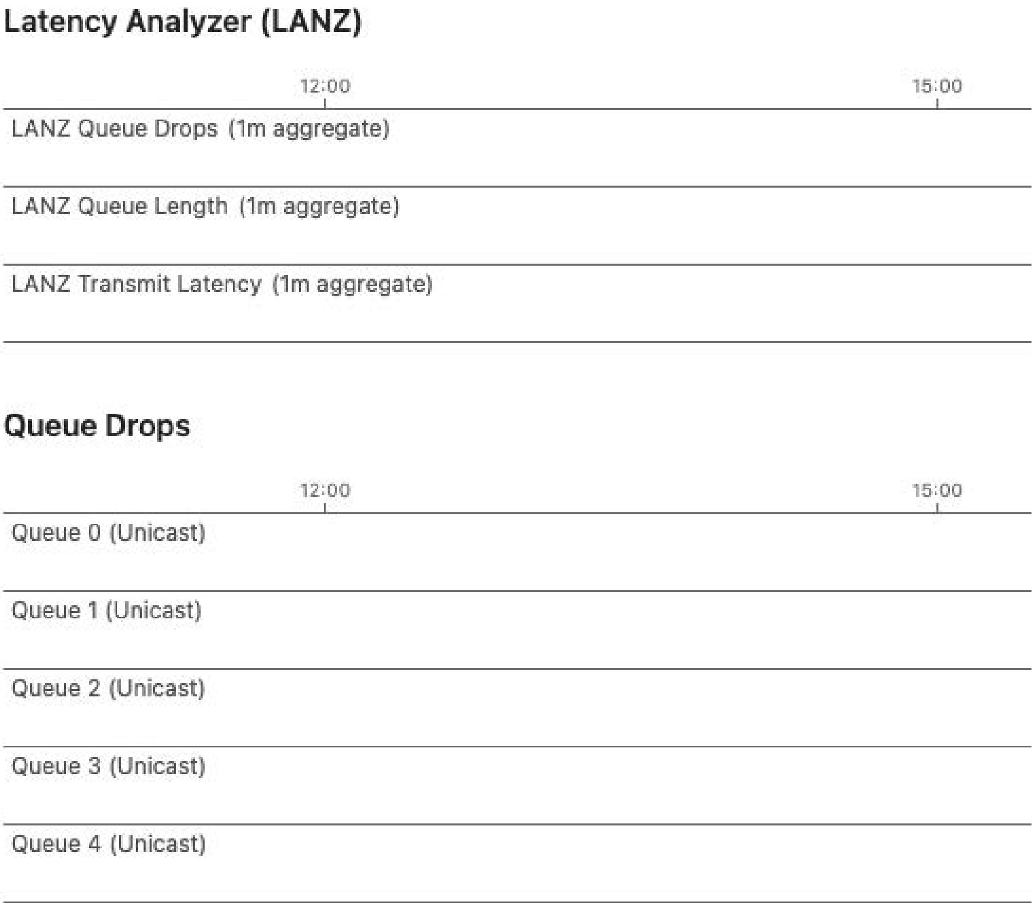


FIGURE 20 Latency and queue drops

Arista CVP Notifications

In a separate test, bandwidth was intentionally restricted to limit the total available bandwidth during job execution. The resulting Quality of Service (QoS) statistics were then measured, with a focus on queueing, induced delays, and packet drops. This test was conducted independently of the previously defined tests and demonstrated how Arista notifications are triggered when QoS thresholds are exceeded.

Examples of Intentional Overload

In this example, the load exceeded the available bandwidth, leading to QoS kicking in. While queueing resources were engaged, no Ddiscards occurred. Also visible in one quick reference is the existence of ECN (Explicit Congestion Notification) messages passing through the switches from end point to end point.

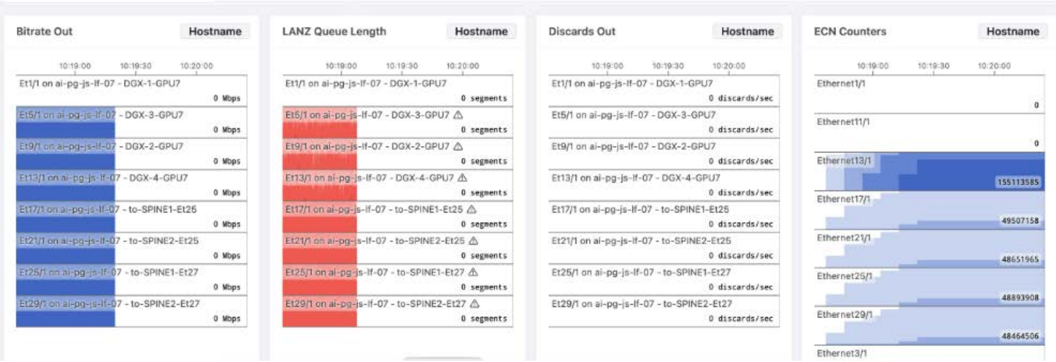

Bitrate Out
Hostname
LANZ Queue Length
Hostname
Discards Out
Hostname
ECN Counters
Hostname
10:19:00
10:19:30
10:20:00
Et1/1 on ai-pg-js-lf-07 - DGX-1-GPU7
Et5/1 on ai-pg-js-lf-07 - DGX-3-GPU7
Et9/1 on ai-pg-js-lf-07 - DGX-2-GPU7
Et13/1 on ai-pg-js-lf-07 - DGX-4-GPU7
Et17/1 on ai-pg-js-lf-07 - to-SPINE1-Et25
Et21/1 on ai-pg-js-lf-07 - to-SPINE2-Et25
Et25/1 on ai-pg-js-lf-07 - to-SPINE1-Et27
Et29/1 on ai-pg-js-lf-07 - to-SPINE2-Et27
0 Mbps
0 segments
0 discards/sec
Ethernet1/1
0
Ethernet11/1
0
Ethernet13/1
155113585
Ethernet17/1
49507158
Ethernet21/1
48651965
Ethernet25/1
48893908
Ethernet29/1
48464506
Ethernet3/1

FIGURE 21 Bit rates and counters before QoS

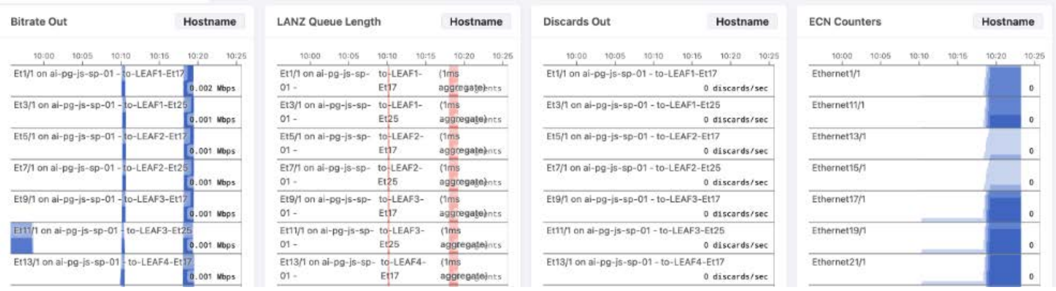

Bitrate Out
Hostname
LANZ Queue Length
Hostname
Discards Out
Hostname
ECN Counters
Hostname
10:00
10:05
10:10
10:15
10:20
10:25
Et1/1 on ai-pg-js-sp-01 - to-LEAF1-Et17
0.002 Mbps
Et3/1 on ai-pg-js-sp-01 - to-LEAF1-Et25
Et5/1 on ai-pg-js-sp-01 - to-LEAF2-Et17
Et7/1 on ai-pg-js-sp-01 - to-LEAF2-Et25
Et9/1 on ai-pg-js-sp-01 - to-LEAF3-Et17
Et11/1 on ai-pg-js-sp-01 - to-LEAF3-Et25
Et13/1 on ai-pg-js-sp-01 - to-LEAF4-Et17
0.001 Mbps
(1ms aggregate)
0 discards/sec
Ethernet1/1
Ethernet11/1
Ethernet13/1
Ethernet15/1
Ethernet17/1
Ethernet19/1
Ethernet21/1
0

FIGURE 22 Bit rates and counters post-QoS

In the example below, not only are we alerted to the high load, but we can see the reception of PFC frames received on one of the interfaces colored in red:

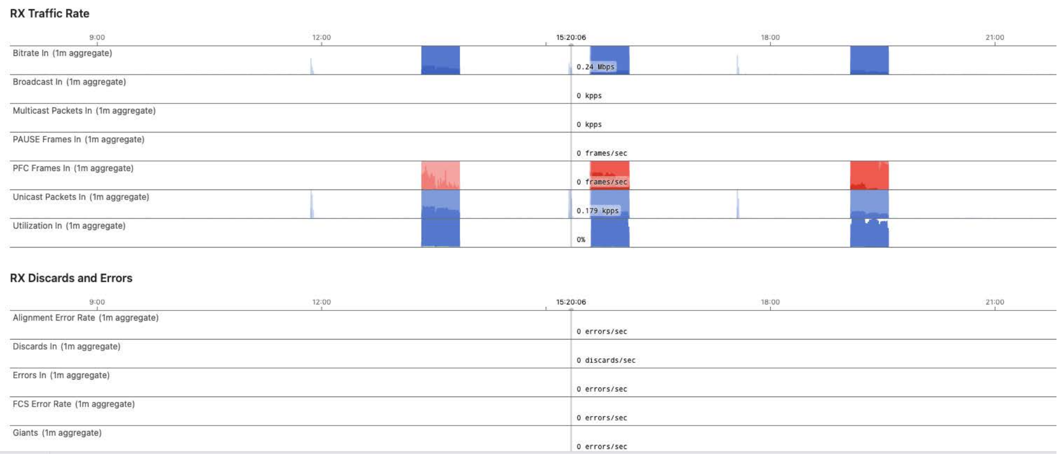
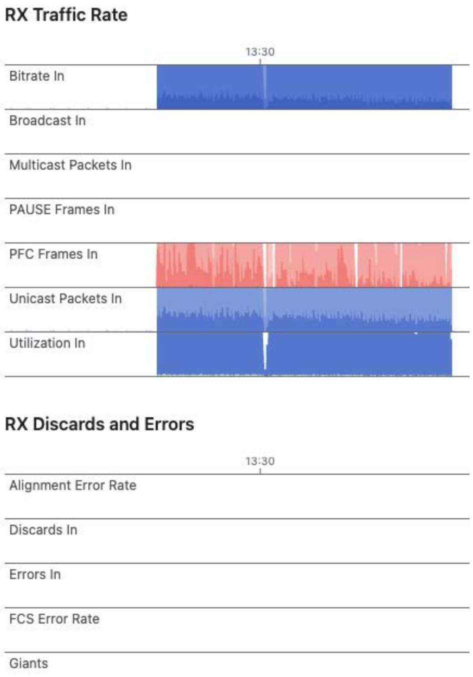

RX Traffic Rate
9:00
12:00
15:20:06
18:00
21:00
Bitrate In (1m aggregate)
0.24 Mbps
Broadcast In (1m aggregate)
0 kpps
Multicast Packets In (1m aggregate)
0 kpps
PAUSE Frames In (1m aggregate)
0 frames/sec
PFC Frames In (1m aggregate)
0 frames/sec
Unicast Packets In (1m aggregate)
0.179 kpps
Utilization In (1m aggregate)
0%
RX Discards and Errors
9:00
12:00
15:20:06
18:00
21:00
Alignment Error Rate (1m aggregate)
0 errors/sec
Discards In (1m aggregate)
0 discards/sec
Errors In (1m aggregate)
0 errors/sec
FCS Error Rate (1m aggregate)
0 errors/sec
Giants (1m aggregate)
0 errors/sec

FIGURE 23 PFC frame alert 1

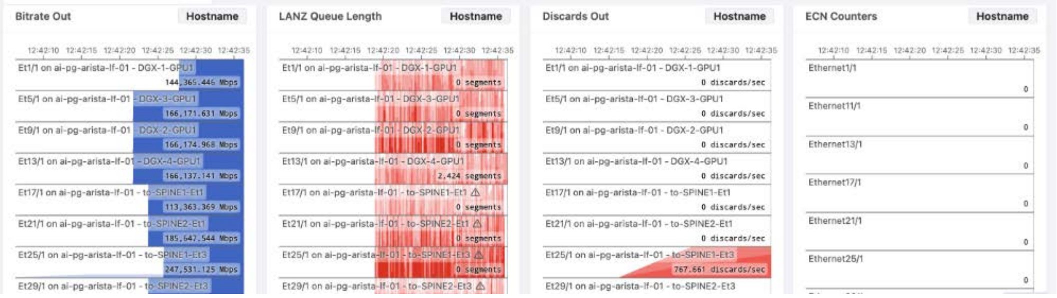


RX Traffic Rate

RX Discards and Errors

FIGURE 24 PFC frame alert 2

By creating a non-production scenario where the provided bandwidth is a fraction of what was required for a specific job, we can see below that ECN kicked in, PFC protects the control plane packets, and discards are induced for the unprotected data plane traffic.



Bitrate Out — Hostname

Et1/1 on ai-pg-arista-lf-01 - DGX-1-GPU1 144,365.446 Mbps
Et5/1 on ai-pg-arista-lf-01 - DGX-3-GPU1 166,171.631 Mbps
Et9/1 on ai-pg-arista-lf-01 - DGX-2-GPU1 166,174.968 Mbps
Et13/1 on ai-pg-arista-lf-01 - DGX-4-GPU1 166,137.141 Mbps
Et17/1 on ai-pg-arista-lf-01 - to-SPINE1-Et1 113,363.369 Mbps
Et21/1 on ai-pg-arista-lf-01 - to-SPINE2-Et1 185,647.544 Mbps
Et25/1 on ai-pg-arista-lf-01 - to-SPINE1-Et3 247,531.125 Mbps
Et29/1 on ai-pg-arista-lf-01 - to-SPINE2-Et3

LANZ Queue Length — Hostname

Et1/1 on ai-pg-arista-lf-01 - DGX-1-GPU1 0 segments
Et5/1 on ai-pg-arista-lf-01 - DGX-3-GPU1 0 segments
Et9/1 on ai-pg-arista-lf-01 - DGX-2-GPU1 0 segments
Et13/1 on ai-pg-arista-lf-01 - DGX-4-GPU1 2,424 segments
Et17/1 on ai-pg-arista-lf-01 - to-SPINE1-Et1 ⚠ 0 segments
Et21/1 on ai-pg-arista-lf-01 - to-SPINE2-Et1 ⚠ 0 segments
Et25/1 on ai-pg-arista-lf-01 - to-SPINE1-Et3 ⚠ 0 segments
Et29/1 on ai-pg-arista-lf-01 - to-SPINE2-Et3 ⚠

Discards Out — Hostname

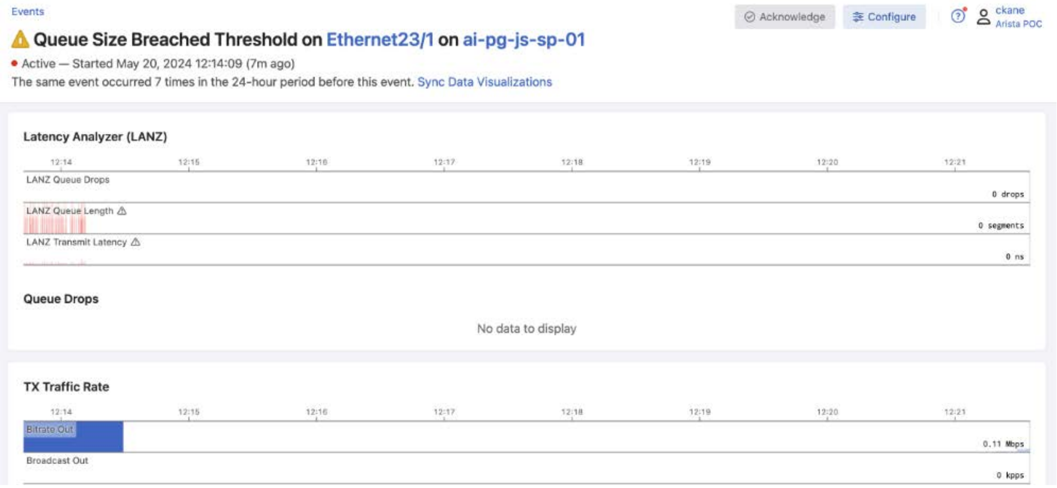
Et1/1 on ai-pg-arista-lf-01 - DGX-1-GPU1 0 discards/sec
Et5/1 on ai-pg-arista-lf-01 - DGX-3-GPU1 0 discards/sec
Et9/1 on ai-pg-arista-lf-01 - DGX-2-GPU1 0 discards/sec
Et13/1 on ai-pg-arista-lf-01 - DGX-4-GPU1 0 discards/sec
Et17/1 on ai-pg-arista-lf-01 - to-SPINE1-Et1 0 discards/sec
Et21/1 on ai-pg-arista-lf-01 - to-SPINE2-Et1 0 discards/sec
Et25/1 on ai-pg-arista-lf-01 - to-SPINE1-Et3 767.661 discards/sec
Et29/1 on ai-pg-arista-lf-01 - to-SPINE2-Et3

ECN Counters — Hostname

Ethernet1/1 0
Ethernet11/1 0
Ethernet13/1 0
Ethernet17/1 0
Ethernet21/1 0
Ethernet25/1 0

FIGURE 25 Induced discards

CVP dashboards were used to drill into the details of the load and congestion a specific interface was experiencing.



⚠ Queue Size Breached Threshold on Ethernet23/1 on ai-pg-js-sp-01

● Active — Started May 20, 2024 12:14:09 (7m ago)

The same event occurred 7 times in the 24-hour period before this event. Sync Data Visualizations

Latency Analyzer (LANZ)

LANZ Queue Drops 0 drops
LANZ Queue Length ⚠ 0 segments
LANZ Transmit Latency ⚠ 0 ns

Queue Drops

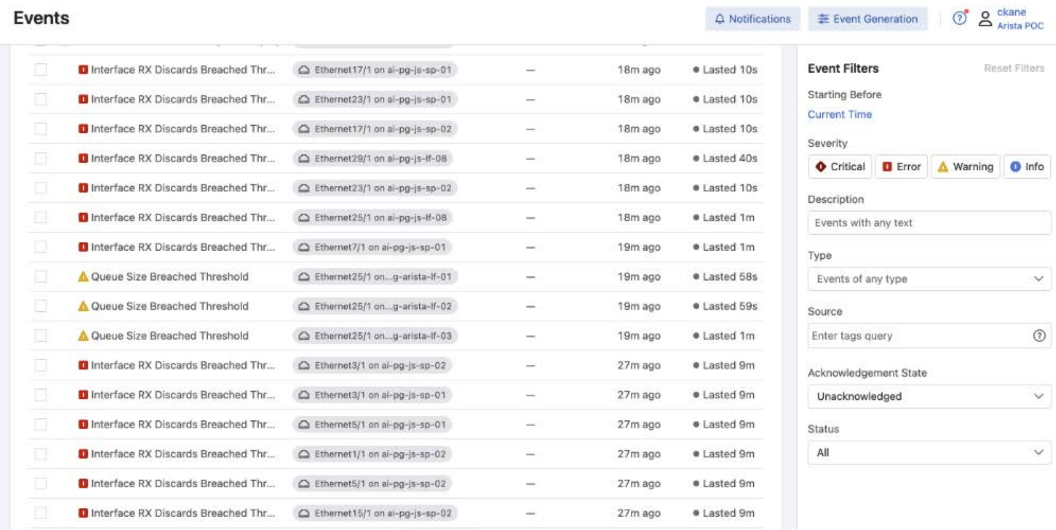
No data to display

TX Traffic Rate

Bitrate Out 0.11 Mbps
Broadcast Out 0 kpps

FIGURE 26 CVP dashboard

During the purposely rate-limited configuration, we can see the events that CVP generated, which were immediately available to the Network Operations Center personnel.



Events

□	Interface RX Discards Breached Thr...	Ethernet17/1 on ai-pg-js-sp-01	—	18m ago	● Lasted 10s
□	Interface RX Discards Breached Thr...	Ethernet23/1 on ai-pg-js-sp-01	—	18m ago	● Lasted 10s
□	Interface RX Discards Breached Thr...	Ethernet17/1 on ai-pg-js-sp-02	—	18m ago	● Lasted 10s
□	Interface RX Discards Breached Thr...	Ethernet29/1 on ai-pg-js-lf-08	—	18m ago	● Lasted 40s
□	Interface RX Discards Breached Thr...	Ethernet23/1 on ai-pg-js-sp-02	—	18m ago	● Lasted 10s
□	Interface RX Discards Breached Thr...	Ethernet25/1 on ai-pg-js-lf-08	—	18m ago	● Lasted 1m
□	Interface RX Discards Breached Thr...	Ethernet7/1 on ai-pg-js-sp-01	—	19m ago	● Lasted 1m
□	⚠ Queue Size Breached Threshold	Ethernet25/1 on...g-arista-lf-01	—	19m ago	● Lasted 58s
□	⚠ Queue Size Breached Threshold	Ethernet25/1 on...g-arista-lf-02	—	19m ago	● Lasted 59s
□	⚠ Queue Size Breached Threshold	Ethernet25/1 on...g-arista-lf-03	—	19m ago	● Lasted 1m
□	Interface RX Discards Breached Thr...	Ethernet3/1 on ai-pg-js-sp-02	—	27m ago	● Lasted 9m
□	Interface RX Discards Breached Thr...	Ethernet3/1 on ai-pg-js-sp-01	—	27m ago	● Lasted 9m
□	Interface RX Discards Breached Thr...	Ethernet5/1 on ai-pg-js-sp-01	—	27m ago	● Lasted 9m
□	Interface RX Discards Breached Thr...	Ethernet1/1 on ai-pg-js-sp-02	—	27m ago	● Lasted 9m
□	Interface RX Discards Breached Thr...	Ethernet5/1 on ai-pg-js-sp-02	—	27m ago	● Lasted 9m
□	Interface RX Discards Breached Thr...	Ethernet15/1 on ai-pg-js-sp-02	—	27m ago	● Lasted 9m

Event Filters

Reset Filters

Starting Before: Current Time

Severity: Critical | Error | Warning | Info

Description: Events with any text

Type: Events of any type

Source: Enter tags query

Acknowledgement State: Unacknowledged

Status: All

FIGURE 27 CVP event generation

Conclusion

Arista and Pure Storage deliver an integrated solution for the high-performance networking and storage required by today's demanding AI workloads. Arista's proven network performance and real-time health insights are essential for teams focused on accelerating time to insight. When combined with a Pure Storage FlashBlade, AI-driven organizations can rely on a stable, high-performance platform that grows with their needs.

Additional Resources

- Explore Pure Storage FlashBlade.
- Learn more about Arista.