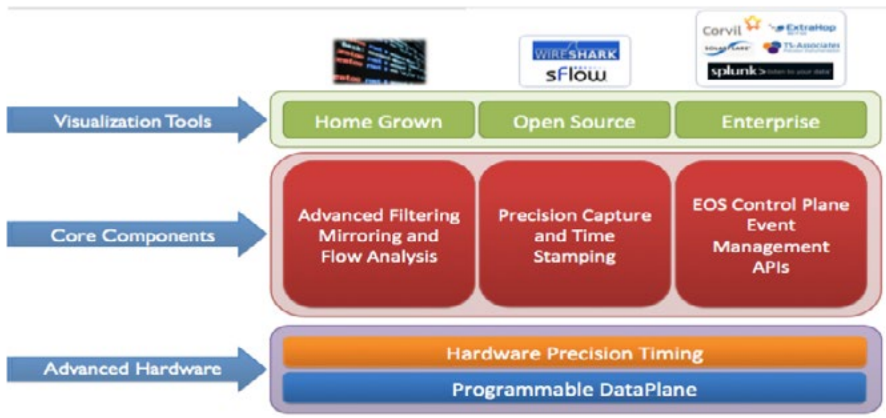
Enabling Arista Advanced Monitoring

Overview

Data center networks continue to grow in scale, capacity, and bit rate and precision analysis has become necessary to effectively plan and operate a modern data center, regardless of whether the priorities are virtualization, performance, capacity, or security. The requirement for robust "always on" monitoring of data flows forces many to build independent observation networks with dedicated appliances, probes, and packet brokers. These dedicated analysis networks are costly and typically have limited bandwidth, which does not scale with the data center network capacity. This leads to asymmetrical monitoring requiring a compromise that ultimately results in less observed data and an inconsistent view of all flows. These observation networks are typically oversubscribed resulting in packet loss and jitter, further rendering application analysis incomplete or inaccurate.

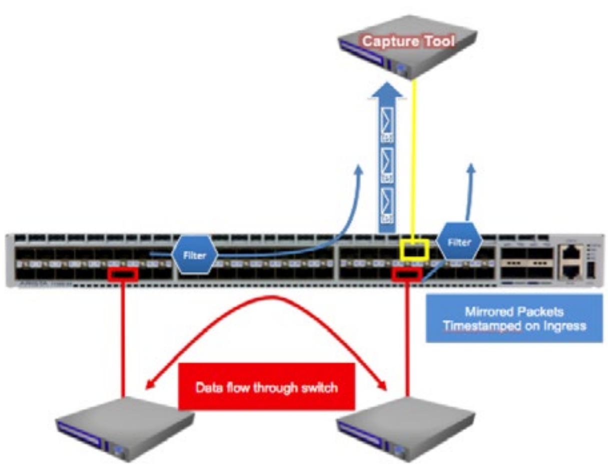
Arista DANZ on the 7150 platform enables scalable end-to-end monitoring directly integrated into the data path, or as a more traditional dedicated observation network. At the core of these capabilities are advanced monitoring, TAP aggregation, and precision time stamping. A rich management API complements these core capabilities, as well as precision time synchronization enabling precision large scale captures and direct integration with open and enterprise application analysis tools. This tech note will expound upon core capabilities of advanced monitoring, aggregation, and time stamping, demonstrating how they can be utilized to remove some of the traditional analysis bottlenecks and design constraints.

Current data analysis and monitoring solutions require dedicated hardware and software, as well as extra physical interconnect, power, and rack space. With the advanced DANZ capabilities, the requirement for additional hardware and data center footprint is reduced or entirely removed.



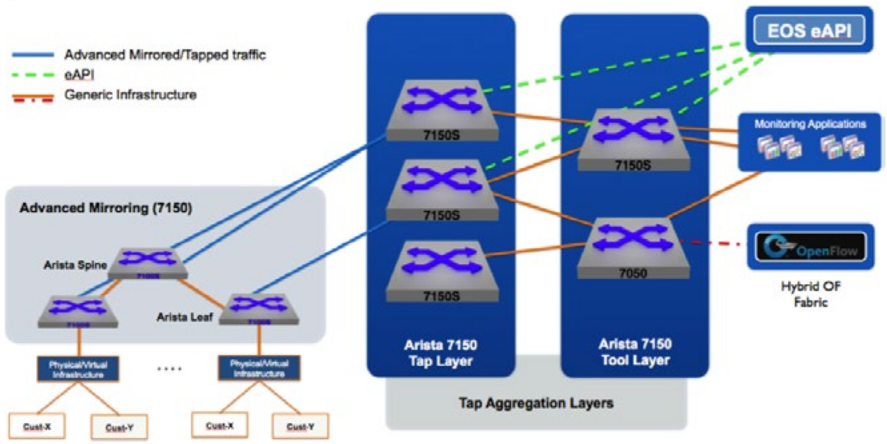
Increasing Scale and Reducing Footprint

Current data analysis and monitoring solutions require dedicated hardware and software, as well as extra physical interconnect, power, and rack space. With the advanced DANZ capabilities, the requirement for additional hardware and data center footprint is reduced or entirely removed. This is because the DANZ features can be enabled on the data-path, avoiding the requirement for a supplement device to filter, truncate, replicate, or time stamp the data. In smaller scale or application specific environments or colocation facilities the integrated functionality on the data path infrastructure provides a common OS and tool set that allows direct integration with open and enterprise third party application analysis software, eliminating a dedicated analysis aggregation network entirely.



In larger scale implementations, a multi-tier approach may be utilized. This design leverages both the integrated advanced monitoring on the data path, as well as the dedicated aggregation mode. At this scale a monitor/tap layer and tool layer allows the analysis applications or appliances to be centralized or virtualized. This design also allows direct integration with OpenFlow or other flow based traffic steering approaches. Arista EOS eAPI allows programmatic control of the multi-tier environment and provides a mechanism for 3rd party applications to configure and control the analysis network provisioning or flow handling.

Arista DANZ integrates a full set of advanced monitoring and aggregation features directly into the data center network.



Both of the above approaches leverage the onboard advanced monitoring feature set enabling the capture and analysis endpoints to receive the specific data of interest in the format expected.

Leveraging DANZ to deliver the right packets to the right tools

One of the issues faced by application analysis or packet capture appliances or software is limited capacity or throughput. There is not enough bandwidth with traditional TAP/SPAN infrastructure to deliver all packets from all points to the tools. Even if there was enough bandwidth, the tools themselves have limited capacity due to application or hardware architecture. This has resulted in the tools being physically distributed in the DC near the islands of aggregation, requiring more tools, physical infrastructure, and centralized management. These tools they require dedicated devices to filter, truncated, and time stamp the data received to effectively process and analyze.

Arista DANZ integrates a full set of advanced monitoring and aggregation features directly into the data center network.

Advanced monitoring provides the following capabilities on the Arista 7150:

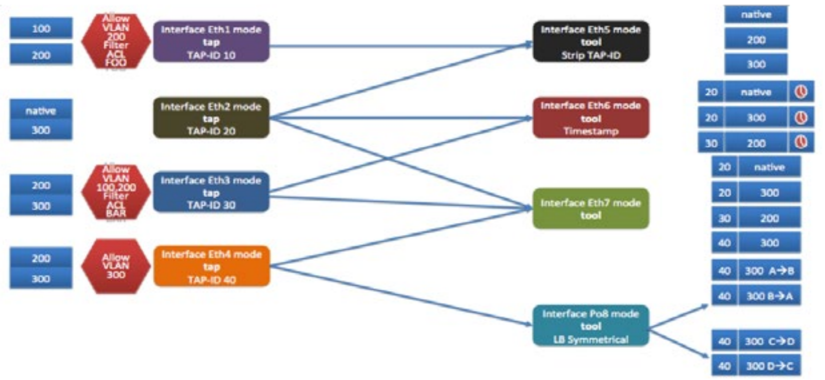
- N:N bi-directional packet mirroring and replication, up to 64 ports
- Monitor to LAG, up to 16 ports
- Monitor to CPU and local storage
- L2/3/4 hardware filtering applied per destination session
- Packet truncation
- Packet time stamping
- Flexible traffic distribution with multiple load sharing scheme
- Hashed, symmetric and dynamic

These two core features are absolutely critical to reduce the amount of data that the analysis network needs to forward to tools and greatly reduces the load on the tools themselves.

The dedicated aggregation mode provides the following capabilities on the Arista 7150:

- N:N packet replication, up to 64 ports
- L2/3/4 hardware filtering
- Packet truncation
- Packet time stamping
- VLAN tag filtering
- VLAN tunneling (4095 inner/outer ID's)
- Flexible traffic distribution with multiple load sharing scheme
- Hashed, symmetric and dynamic

You will note that both modes of operation provide filtering and truncation of data. These two core features are absolutely critical to reduce the amount of data that the analysis network needs to forward to tools and greatly reduces the load on the tools themselves. The aggregation mode provides flexible port and flow identification via VLAN tagging and tunneling to ensure that the data or flows are correctly identified on ingress from a monitor or tap output. Once filtered and identified the flow can be VLAN tunneled through the analysis network for forwarding to the specific tools or output ports that require it. Further processing can be done at the aggregation output, including removing VLAN tags and further filtering or truncation. This concept allows a distributed analysis "matrix" of unique, intelligent inputs and content specific outputs.



In some deployments a single tool or application may not be able to scale to the throughput required, or the desire may be to analyze a higher bit rate interface with lower bit rate interfaces which is currently a bottleneck in 40G/100G monitoring. In this instance Arista EOS enables load-balancing the traffic over multiple interfaces, effectively creating an output port set at a lower bit rate that services the full capacity required. In some scenarios flow symmetry may also be desired to ensure that both RX and TX packets from the same flow are captured at the same output port. This mode of operation is independent of other monitoring or aggregation features and can be configured to symmetrize on L2 or L3/L4 header data.

Arista is working with multiple partners to enable time stamp decoding in their tools and enable the highest levels of resolution in their application visualization tools.

Gain precision insight via hardware time stamping

There are now many applications that require precision time stamping as part of their analysis. In financial services applications latency analysis and event correlation is done at nanosecond scale. In other applications the desire may be to centralize tools while distributing capture over a regional or global scale. In order to have precise event correlation, the time stamp for events must be captured locally before the data is forwarded to a central point which may be 100's of milliseconds away.

Time stamping is typically done at the analysis tools via a hardware time stamping capable network interface card. This is typically not cost effective to scale and requires direct hardware integration for the highest precision. One workaround has been to aggregate flows via an aggregation device before output to a timestamp capable tool, but this results in lost data due to fan in oversubscription, jitter due to congestion, or flow interleaving before the data can reach the tool.

The Arista 7150 provides a line rate hardware time stamp insertion capability on every port. On ingress the time each packet arrives in noted within ~10ns. Once an egress interface is enabled for time stamping, all packets destined for that interface will contain the ingress time stamp of each packet.

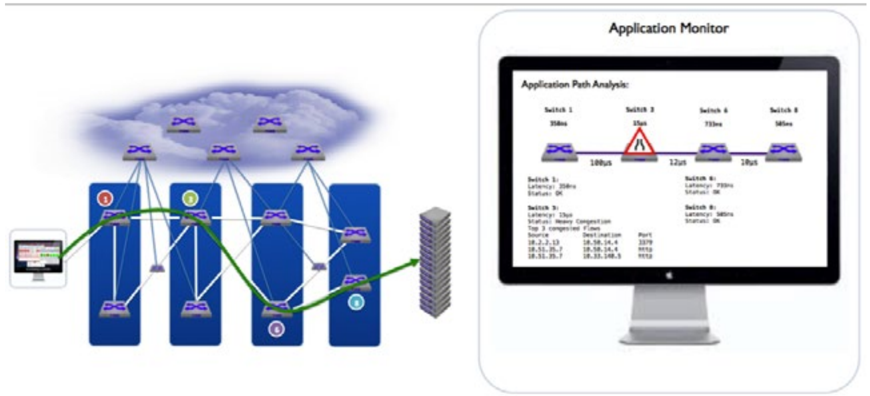
There are two timestamp modes: replacing the Frame Check Sequence (FCS) or appending to the payload. Mode selection is on a per-port basis, so a switch can potentially have different ports configured in each mode. As part of the timestamp implementation, the 7150 also can be configured to generate a key frame in hardware at fixed intervals up to 100 times per second. This key frame provides all the data needed for a software based tool to interpret the Arista timestamp and correlate to UTC. It contains the source port and switch ID, UTC via PTP sync or local clock, the ASIC counter value, and other data to ensure reliable correlation.

The Arista time stamp and key frame format are discussed in detail here: https://eos.aristanetworks.com/?p=1250

This implementation provides home grown or 3rd party applications hardware based nanosecond resolution on captures without requiring dedicated and costly network interface cards to timestamp. Packets arrive with the timestamp appended, with the keyframe in line. Once enabled end to end analysis with hardware time stamps at each hop or observation point is possible. The typical concerns of locality or congestion for observed data are no longer valid as the data receives the timestamp at point closest to ingress. This truly distributed, hardware based time stamp infrastructure enables tool centralization with the hardware precision formerly only possible with direct connections from TAP to tools.

Arista is working with multiple partners to enable time stamp decoding in their tools and enable the highest levels of resolution in their application visualization tools. These partners provide data capture, application analysis, and performance monitoring tools that directly leverage and integrate with the core DANZ features discussed in this document. The list of current partners along with their solutions is maintained here: http://www.aristanetworks.com/solutions/technologypartners

Yhe advanced monitoring, aggregation, and time stamp features of the Arista 7150 provide solutions for many of the problems seen in traditional observation networks or application analysis tools.



Conclusion

In conclusion the advanced monitoring, aggregation, and time stamp features of the Arista 7150 provide solutions for many of the problems seen in traditional observation networks or application analysis tools. By directly integrating these features into the data path, distributing precision instrumentation at the edge while aggregating to a common tool deployment is now feasible, reducing cost while expanding analysis capacity and resolution.

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