Leveraging EOS and sFlow for Advanced Network Visibility

As data center architectures consolidate to common infrastructure, shared services and cloud-like two-tier networks, system and network utilization and total capacity continue to increase at a pace that exceeds the current generation of monitoring tools and applications.

Historically it was feasible to leverage a few dedicated hardware platforms at strategic network choke points to provide application level flow visibility. Changing application behaviors, a high degree of infrastructure sharing and widely deployed multi-path networks mean the existing tools cannot meet the goals of providing visibility, or scale and cost expectations of next-generation environments.

Full packet capture is possible through integrated or overlaid telemetry networks, for example leveraging Arista's Data Analysis toolset, however there is often simply too much data to process effectively in real-time creating a data mining and storage challenge.

Overcoming these challenges requires a two tier approach to monitoring - first using a coarse view of network wide activity to identify first-order anomalous behavior, followed by focusing down on particular hotspots or flows for detailed capture and analysis.

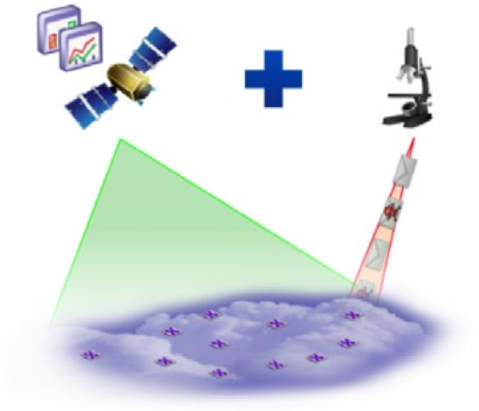

Figure 1: Combining coarse and fine-grained analysis for maximum visbility

Flow Analysis

Flow analysis is not a new concept, with many high-end routers and Ethernet switches offering this capability. However, dramatic increases in total system capacity (10-40Tbps) and interface speed migration from 1GbE all the way to 100GbE means that the network throughput has grown dramatically faster than the overall capacity of the general purpose CPU embedded into the management plane of the network devices. As a consequence considerable thought needs to be given to scaling any kind of real-time or sampled flow monitoring.

Traditional flow-visibility has been very device centric, with each network device leveraging dedicated processing and memory resources to track IP flows traversing the device. Each device's local flow table is processed locally before being sent to a central collector for aggregation and analysis.

This solution worked effectively when traffic patterns were primarily North-South and where the interface count (and overall throughput) was relatively low, however it became disproportionately computationally expensive as port density and interface rates scaled ahead of CPU performance leading to a narrowing of support in most networking portfolios to a very small fraction of the total system capacity.

The ability to solely monitor traffic at choke points in a north-bound direction is at odds with the trend in application flow and overall scaling where East-West traffic patterns and very large numbers of flows make it desirable to monitor accurately at more network touch points rather than fewer.

One solution that has been proposed by some technology companies is to offload flow table generation to dedicated external devices that collect a subset of traffic through network taps or from mirror ports. However, this approach may even compound the loss of visibility and increase costs for a number of reasons:

- Consolidation of flow monitoring, that would have been performed in parallel across multiple devices, into a single central device means that the central device must scale better than the pool of processers it replaces
- Discrete modular devices will have upper scaling limits in terms of density and throughput, which caps the overall ability to monitor an infrastructure to the number of physical interfaces or aggregate performance. The only path to expand these systems is through upgrading and/or replacing hardware.
- The technologies are highly likely to be proprietary in architecture and as a result costly to purchase, license and support, which are further limits to scaling.

At Arista, we believe that open distributed architectures are the key to delivering economically viable, scalable infrastructures and to provide wide-ranging network visibility. The same paradigm applies to monitoring.

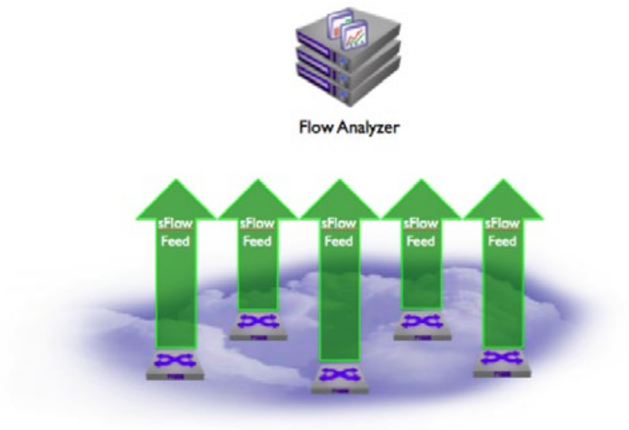

Figure 2: Parallel monitoring out-scales centralized devices

At Arista we have adopted and embedded sFlow (RFC 3176) into all of our products whether they are at the Leaf, Spine, Spline or Telemetry layers of the network, delivering a complete, consistent and distributed stream of flow information unilaterally across the network infrastructure.

sFlow (RFC 3176) In Brief

sFlow is a widely adopted, open standard designed specifically to address the scaling challenges discussed earlier. As a universally embedded technology, it avoids the visibility cliff between device classes often associated with flow monitoring technologies. Most importantly, the ability to monitor your network scales linearly with every new device installed.

The following attributes make sFlow ideal for modern network designs:

- **Computationally cost effective** – sFlow can be widely implemented in hardware in products designed to sit at all cost and scale points in the network meaning no visibility gap between network spine/core and leaf/edge.
- **Performance scales with interface speeds** – sFlow supports 10Gb, 40Gb and 100Gb Ethernet and beyond.
- **Integrated hardware based statistical sampling** – packets do not need to be processed in real time so the data load on devices and monitoring tools is reduced and data integrity is maintained, resulting in accurate and unbiased telemetry.
- **Naturally distributed** – every time a new device is added to the network topology, it adds sFlow processing scale. There is no central bottleneck and both horizontal and vertical visibility is available by default.
- **Collector applications are hardware independent** – sFlow analysis tools run on standard servers with open operating systems.
- **Real-time, raw export** – sFlow does not require pre-correlation or processing on each device and does not summarize flow data. The collector gets a complete view of traffic profiles, including payload data.
- **Lack of local correlation** – local device CPU performance does not affect the ability of the device to provide accurate flow data.

sFlow is Tremendously Versatile in Usage

sFlow data is routinely used for multiple parallel applications including:

- Application performance management
- Troubleshooting
- Capacity planning

- Congestion management
- Accounting, billing and charge-back
- Anomaly and attach detection and audit
- Path profiling

sFlow in Telemetry Overlays

With support of sFlow in all Arista products, flow information can be generated in a distributed fashion across the entire data center. However, most network infrastructures contain at least a few legacy devices or transit links that would benefit from additional monitoring.

The deployment of a DANZ equipped tap aggregator (packet broker) enables the generation of sFlow datagrams from sources such as network taps or traffic derived from port mirrors.

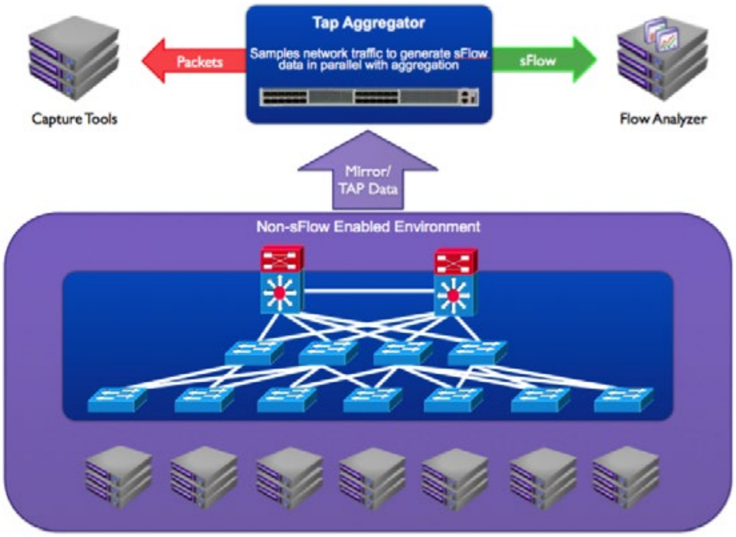
Tap Aggregator
Samples network traffic to generate sFlow data in parallel with aggregation
Packets
sFlow
Capture Tools
Flow Analyzer
Mirror/ TAP Data
Non-sFlow Enabled Environment

Figure 3: Generating sFlow from the TAP Aggregation layer

The advantages of a consistent monitoring paradigm across all devices to allow a unified toolset is clear, however a secondary benefit also exists. As many devices implement traditional flow monitoring in the CPU or in data plane hardware unloading this processing requirement may also unlock network performance and scale improvements.

Thanks to the lightweight, scale optimized nature of sFlow generation this solution provides for up to 1152 ports of 10GbE, 288 ports of 40GbE or 96 ports of 100GbE in the Arista 7500E Series modular platform guaranteeing coverage for a wide range of use cases, including high bandwidth links and wide area links.

Leveraging the Power of EOS and sFlow to Overcome Information Overload

So far we've seen how selection of an appropriate lightweight technology can provide a high degree of visibility across the infrastructure without unduly loading devices or analysis tools. This provides only part of the story.

To close the loop between anomalies detected through flow analysis and the requirement for full packet capture to perform deeper forensic investigation, collector applications require a channel to trigger captures from either infrastructure or telemetry hardware.

With the universal JSON/RPC API native within Arista EOS any partner application can drive the creation and modification of monitoring and packet capture policy within the network and tap aggregation policies in the telemetry layer.

Automating the process of detection through to capture allows creation of a system that quickly identifies, captures and alerts operators to with minimal costs in infrastructure or tooling.

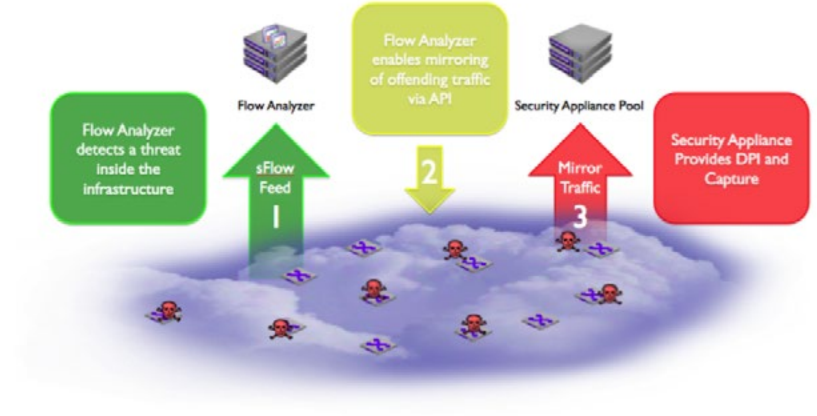
Flow Analyzer
Flow Analyzer enables mirroring of offending traffic via API
Security Appliance Pool
Flow Analyzer detects a threat inside the infrastructure
sFlow Feed
1
2
Mirror Traffic
3
Security Appliance Provides DPI and Capture

Figure 4: Automating detection and focused capture using open APIs

Autonomy for Maximum Scale

Further extending the automated and closed loop architecture Arista's robust and open Linux/x86 architecture opens a host of possibilities for the application of local intelligence to behavioral analysis.

Each Arista device offers the local processing of a small server, coupled with process and memory management of the standard Linux OS. Additionally EOS provides open access to data plane statistics and monitoring. Each device is able to parse critical telemetry data locally and make decisions on analysis and capture for the first time.

Consider a simple case of local application monitoring sFlow data in real-time. Even without the context of wider network sampling, local trends can be an invaluable early warning to impending performance, capacity or security issues. For example, excessive connection rates, traffic types, interface congestion, high latency or packet loss can signal not only an event that may need to be captured for analysis but may be important factors in decisions made by upstream orchestration tools about the placement of new workloads and load balancing distribution.

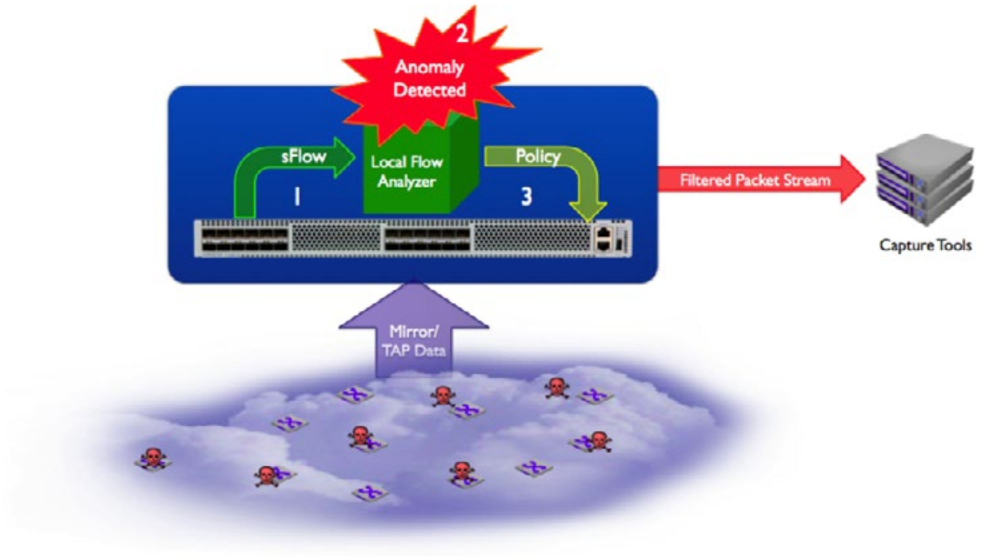

Figure 5: Localized automation enables intelligent response to anomalies

The straightforward pure Linux environment of EOS combined with a plethora of telemetry data derived from sFlow, LANZ and the data plane opens a large number of possibilities for intelligent network control and capture that scales linearly as the infrastructure expands.

Summary

Addressing telemetry at data center scale and cloud economics requires an evaluation of the traditional centrally managed, big-iron approach to monitoring. Linear scalability can only be achieved by distribution of intelligence and consistent visibility across the infrastructure.

The challenge is not to simply replace one means of data generation with another but to dramatically increase the scale and reduce the operational overheads where high degrees of automation are required.

Arista's portfolio addresses these requirements with a simple and open approach that enables operators to first increase visibility and later to build automated workflows that improve response times and application intelligence. By removing the barrier between telemetry and in-path functionality, scalable solutions can be built at minimum cost.

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