BROADCOM

Product Brief

In-band Telemetry

Key Use Cases

- Real-time congestion analysis
- End-to-end NFV monitoring
- Link-level and switch-level latency analysis

Introduction

Datacenter networks have become cloud scale. Deployment of hyper-converged networks is increasing. Telecom networks will enable connectivity everywhere with higher bandwidth delivering 5G wireless services. All of these next-generation networks not only require much higher bandwidth, but also require real-time telemetry to deliver services with good Quality of Experience (QoE).

A network with increased and real-time visibility enables better reliability and real-time control.

Figure 1: Cloud-scale Datacenters

In traditional network monitoring, an application polls the host CPU to gather aggregated telemetry every few seconds or minutes which doesn't scale well in next generation networks. In-band telemetry enables packet-level telemetry by having key details related to packet processing added to the dataplane packets without consuming any host CPU resources.

Using In-band telemetry, it becomes easier to analyze when a packet enters and exits the network, at what rate packets arrive at a particular hop, what path a packet takes, how long the packet spends at each hop, and which switches experience congestion.

Gathering such details at the packet level enables real-time telemetry with microsecond granularity.

In-band Telemetry Mechanisms

A number of in-band telemetry mechanisms are being developed.

A few standards are being defined through the Internet Engineering Task Force (IETF)

- "In-band or In-situ OAM" defines one specific mechanism for adding telemetry data to the dataplane packets. It complements out-of-band probe mechanisms such as ping and traceroute.
- Facebook UDP probe is another IETF draft that defines a scheme for gathering telemetry metadata embedded inside the dataplane packets generated by active probes at high frequency.

The P4 project is defining one more framework under the name In-band Network Telemetry

In-band Telemetry and Programmability

Network switch programmability enables in-field upgradability of the pipeline to support new packet processing capabilities. Trident 3 delivers flex-parsing, flex-table lookup, and flex-metadata insertion in the pipeline—all through programmability!

In-band Telemetry Enables Real-time Congestion Analysis

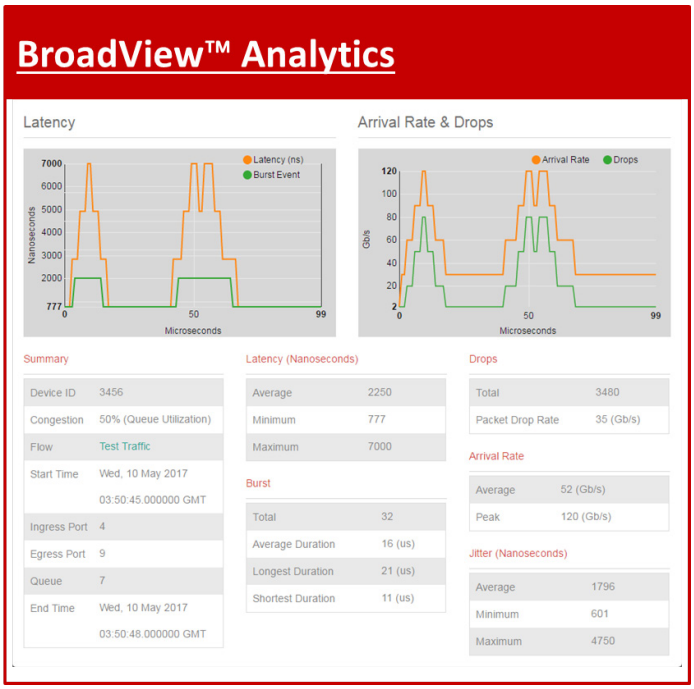
A common scenario is congestion monitoring including detailed microbursts characterization.

Datacenter traffic is bursty by its very nature and congestion often gets created in leaf and/or spine switches resulting in packet loss and consequent degraded application performance.

Congestion characterization at microsecond granularity, such as understanding the duration of the burst and the arrival rate, at a microsecond granularity, will help operators better optimize their network and provide higher service-level guarantees.

Trident 3's in-band telemetry is capable of embedding metadata, such as queue-id, ingress and egress timestamps, transmit rate, and drop count in the forwarding plane packets so that a remote analytics application such as BroadView™ Analytics, can build histograms for latency and burst characterization at microsecond granularity. In the following picture, the BroadView Analytics application shows histograms using the data enabled by the BroadView Agent running on the switch.

Figure 2: Congestion Monitoring



In-band Telemetry Enables End-to-End NFV Monitoring

Overlay technologies such as VxLAN have increased the management complexity and reduced visibility into traffic flows

Operators have been demanding the ability to monitor packet flows across the switch and server in a seamless way, so that packet-flow related issues can be isolated quickly to underlay or overlay for easier network troubleshooting and improved SLA compliance.

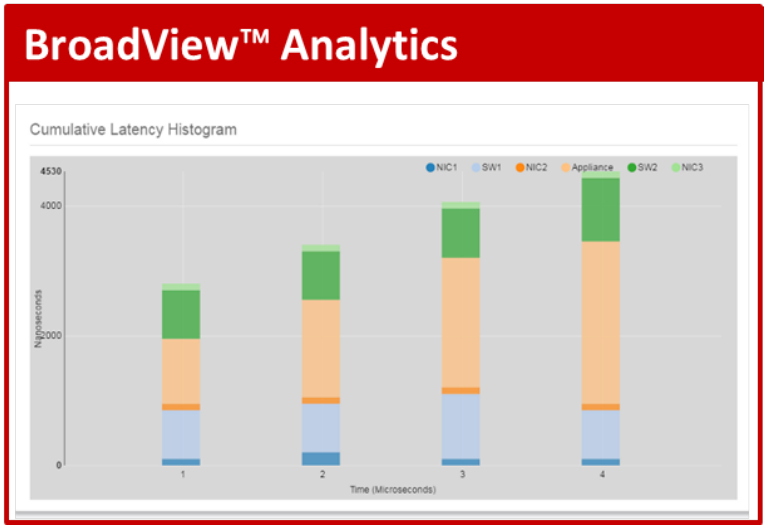
When a packet traverses through the switch and server hops that are in-band OAM compliant, each hop adds timestamps and device-id information to the packet. This enables applications to seamlessly gather end-to-end telemetry.

The instructions specifying the type of telemetry data to collect are inserted into transmitted packets by the originating server. These instructions are a component of the in-band telemetry headers that are added to packets at the dataplane layer. In addition to the collection instructions, servers also insert the same type of telemetry metadata as switches (e.g., time stamps, congestion information, etc.) into the data plane headers.

Trident 3's in-band OAM capability enables real-time network latency analysis and facilitates SLA compliance.

Figure 3 shows a cumulative latency histogram captured using BroadView Analytics application for a packet traversing multiple nodes in a path, and the latency experienced at each node.

Figure 3: Cumulative Latency Histogram



An end-to-end system-level telemetry solution requires in-band telemetry capabilities on both the switches and the servers. To aid implementation of end-to-end telemetry solutions for datacenter networks, Broadcom has successfully demonstrated In-band Telemetry solution with Trident 3 and Broadcom's NetXtreme E-Series NICs. The telemetry capabilities of Broadcom's NetXtreme E-Series NICs include initiation and termination of telemetry sessions, insertion of telemetry data into the data plane packets and reporting the collected telemetry data to a remote analytics application such as BroadView Analytics.

Summary

Trident 3's programmable pipeline enables the addition of completely new features via in-field upgrade without increasing costs or power consumption.

The BroadView Instrumentation suite consists of BroadView Agent which runs on the switch system, the BroadView Analytics application which typically runs in an x86 server outside the switch system, and plug-ins for various SDN/Cloud frameworks. The BroadView Instrumentation suite enables the fastest time-to-market for the latest instrumentation features by leveraging Trident 3 programmability.

Broadcom's market leading 10/25/50/100G NetXtreme E-series NICs are generally available today from most tier 1 server OEMs as well as from Broadcom directly.

Broadcom is the number one merchant silicon switch vendor, and is the industry leader with the broadest Ethernet switch portfolio spanning every layer of the network.

Broadcom offers the largest portfolio of switch chips from 6.4 Tb/s to < 100 Gb/s, spanning from cloud datacenters to enterprise campuses.

Visit www.broadcom.com for more information about the Trident 3 Ethernet switch, the BroadView instrumentation suite, and related Broadcom technology.

BROADCOM®
connecting everything®

Visit the Broadcom website at: www.broadcom.com/products

In-Band Telemetry-PB100 06.14.17