IDC TECHNOLOGY SPOTLIGHT

Sponsored by: Cisco

This paper examines the growing importance of pervasive and real-time data plane streaming telemetry to network performance monitoring in the context of hybrid IT and multicloud. The paper also considers how the Cisco Tetration platform provides diagnostic capabilities and actionable network insights across a wide range of scenarios and use cases.

Addressing Network Performance Monitoring Requirements in Hybrid Cloud Environments

Written by:

Brad Casemore

Research Vice President, Datacenter Networks

AT A GLANCE

KEY TAKEAWAY

Network performance monitoring solutions should provide pervasive and real-time data plane streaming telemetry information, including flow visualization.

Introduction

In the context of digital transformation, applications have gained unprecedented importance, serving as the digital lifeblood for organizations worldwide. Applications increasingly represent the face of the business, generating revenue, facilitating engagement with customers and other stakeholders, driving business outcomes, and ultimately differentiating organizations from their competitors.

Given the paramount importance of applications, their creators — developers — have become the heart of business transformation. By extension, developers also have become invaluable customers of enterprise IT. Applications, designed by developers for customer engagement, directly impact the top and bottom lines and are highly prized for their business value. IT operators, including networking professionals, increasingly provide business value by supporting and delivering applications with greater agility, efficiency, and productivity.

With developers leading the charge, organizations are deploying applications in multiple public and private clouds while continuing to support legacy business applications in on-premises datacenters. This trend will be accentuated and amplified by the rise of microservices, yielding a growing number of business-critical applications developed with innate portability. Indeed, as a result of microservices, developers are beginning to construct highly distributed application environments in which application tiers and data services are spread across datacenters and public clouds.

The network provides essential bandwidth, connectivity, and network services to these all-important applications, regardless of whether they reside in an on-premises datacenter, a private cloud, a public cloud, or multiple clouds. The network also should be capable of providing

invaluable visibility and insight into application behavior, dependencies, and performance. Unfortunately, outdated protocols, technologies, and tools have not kept pace with evolving application environments, and the network has fallen short of providing this capability.

Today's applications are characterized by unprecedented dynamism, leveraging virtualization, containerization, microservices, and workload mobility to create ever-changing communication patterns between distributed application components. Applications are also generating greater traffic volume within datacenters.

Given the limitations of network protocols and many network performance monitoring tools, enterprises lack adequate data plane visibility for performance tracking, troubleshooting, and diagnostics, all of which are critical to ensure and maintain application availability, performance, and responsiveness. Indeed, current network monitoring offerings provide limited visibility into data plane traffic within the network fabric, resulting in higher mean time to repair/resolution (MTTR) and other operational inefficiencies.

With no viable options at hand, organizations are compelled to use manual, inexact, and time-consuming processes to ascertain the network paths traversed by application traffic flows. They also must determine where the greatest latency exists on the network and where packets are being dropped, as well as where latency and packet drops are affecting flows. This task is complicated by the challenge of correlating flows between clients and servers when traffic goes through a Layer 4–7 "middle box," such as a load balancer of a NAT gateway.

To properly meet these challenges, IT operators require systems that can provide pervasive and real-time data plane streaming telemetry information, including flow visualization. They also need the ability to analyze such information at scale using machine earning technologies and to arrive at actionable insights that result in highly efficient application-oriented network operations.

Benefits

With such a system in place, IT operators will possess a comprehensive and modernized approach to network performance monitoring, capable of supporting application environments that are on-premises, in private clouds, or in public clouds. Several key benefits can be realized, including the following:

- Enhanced real-time network visibility and flow visualization result in faster troubleshooting and remediation of performance issues.
- Faster remediation of network issues allows for continuous optimization of network performance.
- Correlation enables operators to determine whether a problem is on the network or on a server.
- By being able to identify exactly where a problem resides, network operators will also achieve faster MTTR. Similarly, network operations can achieve faster "mean time to innocence" through the expedited ability to demonstrate conclusively that a problem exists somewhere other than the network.
- Network forensics are improved through rich network analytics and network visibility.
- Operators can obtain a time series view of topology to understand how it has changed over a period of time.

- Application dependency mapping and application insights enable the prompt identification of performance bottlenecks that, for example, determine whether an application process can handle packet flows at speed.
- Visibility into how hop-by-hop paths traverse server clusters is now available.
- Operators can identify points of network congestion that cause TCP window-size issues and gain additional TCP performance insights into application and network metrics.
- Issues of application latency, session initiation/establishment latency, and flow latency can be seen and addressed.
- Operators gain the capacity to identify whether latency issues are on the fabric or across the WAN, with network insights facilitating faster identification of where problems reside and how to quickly remediate them.
- The ability to apply fabric alerts based on thresholds relating to latency, packet drops, or packet throughput is invaluable. Alerts can be integrated into an operations dashboard, allowing prompt action and remediation.

Key Trends

There is a strong relationship between digital transformation and cloud computing and between the rise of cloud and cloud-native application environments with the growing need for pervasive and real-time network visibility to inform effective datacenter network operations and management.

Digital transformation is an imperative worldwide for organizations of all sizes and across all vertical markets. Worldwide spending on digital transformation technologies (hardware, software, and services) is expected to reach nearly $1.3 trillion this year, an increase of 16.8% over the $1.1 trillion spent in 2017. What's more, IDC forecasts that spending on digital transformation will maintain robust growth through 2021, when it will reach nearly $2.1 trillion.

Cloud is a primary means through which organizations seek to realize their digital transformation objectives, and multicloud is fast becoming the dominant enterprise posture. Indeed, by 2021, more than 90% of enterprises will have intensively multicloud environments with on-premises, off-premises, public, and private cloud — and multiples and combinations of all of them — as default environments.

Containers and microservices, which can facilitate agile application development and deployment as well as application portability, also will proliferate and become ubiquitous in datacenters. IDC expects that nearly 50 million container hosts (physical and virtual) and about 3 billion container instances will be installed by 2021.

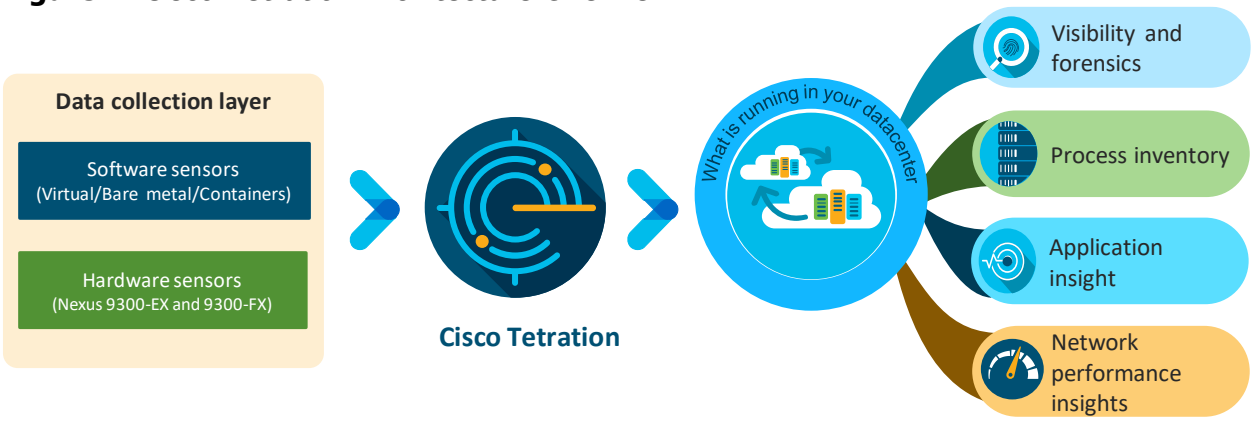
As a result of these trends, enhanced network visibility will be essential for digital business and cloud infrastructures. Enterprises will prioritize network visibility investments to enable faster and improved digital outcomes for internal and external application users. IDC anticipates that during the 2018–2019 time period, average enterprise spending on network visibility will increase sharply, rising from approximately $100,000 for a new corporate investment in network monitoring to almost $400,000 to deliver network visibility into mission-critical cloud infrastructures supporting the digital enterprise.

Considering Cisco

The Cisco Tetration platform extends machine learning capability to provide actionable insights into network performance yield across a wide array of diagnostic capabilities. The platform provides network insights capability that enables datacenter operations teams to better visualize network communication characteristics for mission-critical applications running in on-premises datacenters and in the public cloud. It supports these capabilities with comprehensive traffic telemetry information collected from servers as well as from Cisco Nexus 9000 Series datacenter switches.

The platform performs advanced analytics processing and monitors various network performance statistics. It also discovers and tracks the physical network topology in time series for the Cisco ACI fabric and external devices such as servers connected to it. Advanced features enhance flow-search capabilities, allowing operations teams to search all flows through a link or a queue. They can then pivot around flows that are affected by network or server bottlenecks, all of which can be correlated with discoverable application components (see Figure 1).

Figure 1: *Cisco Tetration Architecture Overview*



Source: Cisco, 2018

The Cisco Tetration platform uses telemetry data collected from hardware sensors on Cisco Nexus 9000 Series switches to provide detailed time-series views of traffic flows, packet-drop indications, burst detection, and hop-by-hop latency insights. With such extensive network visibility, operations teams can gain insights into the following:

- Where congestion occurs in the network
- Which applications are affected by link congestion
- Which applications traverse which links (including return trip)
- What the fabric latency is on a per-hop basis
- Which flows are experiencing and affected by packet drops
- Which links are experiencing traffic bursts, potentially causing packet drops
- How the hop-by-hop paths appear for each application flow
- Where flow end-to-end latency is occurring

» What class of service (CoS) applies per flow

» How thresholds can be set on any combination of these parameters to trigger actionable alerts

The Cisco Nexus 9000 Series switches that support the Cisco Tetration Network Insights capability include top-of-rack (ToR) switches such as the Cisco Nexus 93108TC-FX, the Cisco Nexus 93180YC-FX, and the Cisco Nexus 9348GC-FXP. Spine switches that support the Cisco Tetration Network Insights functionality include Cisco Nexus 9500 Series switches with the following line cards: N9K-X9736C-FX and Fabric Modules for the N9K-C9504-FM-E, N9K-C9508-FM-E, and N9K-C9516-FM-E.

The Cisco Tetration platform enhances performance monitoring for applications by collecting telemetry data from software sensors installed on the servers as well as from the hardware sensors on the Nexus 9K switches. The platform provides insights that identify the processes that cannot handle packets fast enough, helping operations teams identify whether the problem is on the client side or the server side (see Figure 2).

Figure 2: *Network Insights Using Software and Hardware Sensors*

Source: Cisco, 2018

Data collected from the software sensors is subjected to analytics processing to deliver the following performance metrics:

- Smoothed Round-Trip-Time (SRTT) latency
- Latency perceived by the application
- TCP retransmissions
- TCP window-size tracking
- Long TCP handshakes
- Location of bottleneck — whether network or application

Moreover, by leveraging Internet Protocol Flow Information Export (IPFIX) data from "middle boxes" including application delivery controllers (ADCs)/load balancers such as Citrix NetScaler and F5's BIG-IP, Tetration can correlate client/server flows that traverse the load balancer (or other middle box). These TCP and network performance statistics are independent of network hardware and can yield end-to-end insights into network performance resulting in reduced MTTRs relating to application and network issues.

"Flow stitching" correlates IPFIX telemetry received from load balancers with the telemetry received from Tetration hardware and software sensors and stiches the flows together to provide an end-to-end view of client/server traffic, which represents a notable advance for network operators. Previously, they were unable to correlate flows between clients and servers when the traffic passed through a Layer 4–7 device such as an ADC/load balancer. As a result of flow stitching, however, they can link these flows into a single related entity, enabling faster identification of performance or network issues anywhere on the continuum. Flow stitching can also help identify misconfigurations in load balancers and assist in the selection of optimal load-balancing algorithms.

When it leverages both hardware and software sensors, the Cisco Tetration platform provides insight into the hop-by-hop path and performance statistics for an application, extending all the way from the server process to the client. Consequently, it can help flag flow bottlenecks with the client or the server and with the network or the application. With visibility into the location network congestion, operations team can make informed decisions about network optimizations and bandwidth upgrades.

Analytics-based insights into these performance statistics enable service providers to gain a unique perspective on the datacenter's operations and serve as a catalyst for increased efficiency.

Challenges and Opportunities

Cisco's primary challenge involves market inertia. Many enterprise IT buyers have a traditional view of network performance monitoring, and they have made investments in technologies and tools in that area. For Cisco to succeed with Tetration in the realm of network performance monitoring, a sufficient number of those customers must perceive that the platform is demonstrably superior to what they use today. For some customers, especially those that have traditional, less dynamic application environments, price might also be a consideration.

Regarding opportunities, Cisco Tetration is well placed to compete as a next-generation network performance management platform, capable of addressing many use cases and delivering a range of benefits that will be considered increasingly valuable to enterprises as they pursue digital transformation through strategic investments in IT. IDC expects enterprises to invest increasingly in the sort of network visibility that Tetration provides.

Conclusion

Regardless of where modern applications reside (in private or public clouds), they are distributed and dynamic, predicated on intensive virtualization and, increasingly, on the proliferation of containerization and microservices. Moreover, applications are generating unprecedented traffic flows within datacenters, with most of that traffic now east-west in nature rather than north-south.

As cloud environments proliferate, however, network performance monitoring solutions are not keeping pace with the needs of the enterprise. The limitations of traditional network protocols and network performance tools mean that enterprises lack the breadth and depth of data plane visibility required for comprehensive and expeditious performance tracking, troubleshooting, and diagnostics. In turn, they find themselves at a disadvantage in ensuring and maintaining application availability, performance, and responsiveness.

Enterprise IT organizations need systems that can provide pervasive and real-time network visibility. They need systems that enable them to take informed action in troubleshooting and remediation with agility and at scale. In addition, they want to better align network operations with agile application and digital business requirements.

Responding to those needs, the Cisco Tetration platform leverages artificial intelligence and machine learning to provide diagnostic capabilities and actionable insights for a variety of network performance scenarios. As a result, datacenter operations teams that adopt Cisco Tetration have the means of visualizing and understanding how the network, at any given time, is supporting business-critical applications running in on-premises datacenters and in public clouds. To the extent that Cisco can address the challenges described in this paper, the company has a significant opportunity for success in this marketplace.

IDC Custom Solutions

IDC Corporate USA

5 Speen Street
Framingham, MA
01701, USA
T 508.872.8200
F 508.935.4015
Twitter @IDC
idc-insights-community.com
www.idc.com

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.