IDC
Analyze the Future

IDC TECHNOLOGY SPOTLIGHT

The Increasing Need for Real-Time Datacenter Analytics with Pervasive Visibility

June 2016

Adapted from *SDN Market to Gain Enterprise Headway, Driven by 3rd Platform and Cloud,* by Brad Casemore and Rohit Mehra, IDC #US40628315

Sponsored by Cisco Systems

In a world characterized by increased virtualization, cloud computing, software-defined networking (SDN), and DevOps — and now containers and microservices — having comprehensive visibility into every flow and packet metadata is essential. This paper examines the need for and business value of pervasive, real-time visibility across the datacenter network. It also looks at the role of Cisco in the market for real-time datacenter analytics solutions.

Introduction

A strategic imperative for enterprises worldwide, digital transformation shifts an ever-increasing proportion of business online and requires flexible and agile datacenter infrastructure that can dynamically adapt to meet business needs.

The datacenter network is the linchpin that underlies 3rd Platform initiatives such as digital transformation. Cloud computing, in particular, is having a profound effect on the datacenter network in ways that are both obvious and less well understood. Clearly, cloud and increased enterprise mobility have generated an unprecedented amount of network traffic. That traffic is not only more voluminous; it is also significantly different from the relatively predictable north-south flows associated with client-server computing on the 2nd Platform. More east-west in nature, cloud traffic moves from server to server and rack to rack and demands increased visibility in addition to greater bandwidth.

Enterprises pursuing digital transformation often are aggressively moving toward harnessing the full power of hybrid cloud, leveraging both private and public clouds to optimize the organization's resources and speed time to market. To improve their business agility and operational flexibility, enterprises want the ability to migrate applications between clouds, attaining cost savings and tapping unique platform capabilities in the process.

As they pursue their multicloud strategies, enterprises also are cognizant of the need for disaster recovery and business continuity plans. It has always made sense to plan for potential disasters and to have validated contingency plans in place, and the cloud era further raises the stakes in that regard. Similarly, forward-looking enterprises also attempt to anticipate scenarios that involve mergers and acquisitions, whereby potentially dissimilar IT environments and infrastructure will have to be reconciled.

In conjunction with their focus on cloud migration, these enterprises often are transitioning to SDN, especially in instances — such as private cloud — where the network needs to possess the agility that derives from increased policy-based automation, programmability, and orchestration.

As enterprises make those journeys, they're also attempting to be policy compliant and to shift toward a zero-trust security model that emphasizes whitelisting rather than blacklisting.

In all of the above scenarios — cloud migration, disaster recovery, mergers and acquisitions, SDN migration, policy compliance, and the transition to a zero-trust security model — pervasive visibility is essential to success.

To be sure, virtualization, cloud, and SDN — and now containers and microservices — are redefining the parameters of effective network optimization and security. An increasingly distributed application environment demands greater network visibility and a decentralized approach to security enforcement closer to the edge of the network.

In this context, network automation and networkwide visibility become priorities. Given the increasing complexity that accompanies the highly distributed nature of modern application environments, real-time application visibility across the network represents both a daunting challenge and an acute need.

Application dependencies are a particularly vexing issue. Many enterprises struggle to identify application dependencies, and some move applications repeatedly because of conflicts that directly result from lack of visibility into such dependencies.

Meanwhile, growth in east-west traffic, including encapsulations such as Virtual Extensible LAN (VXLAN), poses increasing barriers to network visibility, both for ongoing operational efficiency and for network forensics.

Indeed, enhanced visibility is required in the context of managing workloads across multiple infrastructures, including public cloud. In these circumstances, whitelist policy implementation and policy compliance emerge as critical considerations.

In this context, three primary challenges confront datacenter operators:

- Lack of pervasive visibility into their datacenter infrastructure, which results in operational problems and security challenges
- An inability to speedily migrate applications between datacenters and to the cloud, including scenarios that involve disaster recovery and business continuity
- Lack of insight and resources to implement a zero-trust security model and to move from complex and relatively ineffective blacklisting to an approach based on whitelisting

For a number of reasons, existing approaches — including tools and systems — to address these issues have been inadequate. The first failing involves an inability to collect consistent telemetry to support datacenter-scale operations. Existing tools and systems were not designed to collect datacenter telemetry at scale or at line-rate speeds. What's more, these tools and systems (such as syslog, NetFlow, and sFlow) have limitations when collecting telemetry from different sources. NetFlow, for example, aggregates statistics, but it doesn't provide the necessary granularity to deliver pervasive visibility.

Existing tools do not possess the ability to analyze data volumes in real time or to address operational challenges comprehensively. Most tools restrict themselves to a single use, such as application performance. These tools also do not provide long-term data retention capabilities needed for extensive forensics. Consequently, customers end up with discrete, unconnected tools for different purposes, with no correlation between them.

Even when systems and tools possess the technological capabilities to address aspects of these new challenges, they tend to be too complex, requiring advanced skills and resources to implement effectively. As a result, they are expensive to own, cumbersome to deploy, and difficult to maintain.

What's needed is an integrated approach to datacenter operations in which infrastructure telemetry provides real-time insights into application dependencies — an increasingly intractable problem in many datacenters — as well as pervasive visibility into compliance and policies.

Fortunately, the concepts at the heart of big data analytics can be applied to network intelligence. As a result, datacenter network operators can obtain actionable insights into application dependencies, east-west traffic flows, security vulnerabilities, forensics, compliance, and policies.

To address these requirements, enterprise IT departments must leverage unsupervised machine learning, pervasive visibility, and behavior analysis to gain actionable insights into 3rd Platform application traffic. That is especially true as enterprises shift to being proactive and predictive in pursuit of pervasive digital transformation.

Benefits

Comprehensive application behavior–based network visibility has become a primary need for the 3rd Platform, especially as it encompasses increased virtualization, containerization, and microservices. IDC believes that application behavior–based network visibility will gain unprecedented importance within the context of network automation and SDN.

The shift to more efficient cloud-based provisioning brings with it the need to alter the diagnostic and operational characteristics of the datacenter network. During the client-server era, operations teams knew where a server was located and exactly what it was doing. That certainty changed as a result of virtualization, and it will change more as Linux containers proliferate. The ability to fully understand the state of the infrastructure and the state of the applications running on the infrastructure will require deeper and more comprehensive telemetry than has been available until now.

In addition, multisite datacenter interconnect fabrics are increasingly prevalent for use cases such as business continuity and disaster recovery, especially for hybrid and public clouds. Thus there's a need for visibility into popular encapsulation and label-based protocols, such as Virtual Extensible LAN, multiprotocol label switching (MPLS), and Network Service Headers (NSH).

Given these changes, IDC believes the use of real-time analytics will grow as IT organizations move toward adopting proactive measures to preclude service slowdowns and outages. Use of machine learning to establish baseline conditions of utilization, response times, and relationships between key operational variables will set the stage for proactive, real-time trend and forecast analysis; anomaly anticipation and detection; incident prevention; and ongoing performance and capacity optimization. With increasing adoption and acceptance of IT operations analytics, IT organizations will look to automate responses to key conditions and use cases to further improve service levels delivered to end users.

It will become increasingly important to harvest and meaningfully analyze application-based data, its behavior, and its impact on the network. Immediate benefits can accrue from untangling application dependency and adopting a fine-grained network security posture. Prominent among these benefits will be the ability for enterprise IT departments to avoid future security attacks and quickly detect and resolve security issues before they cause incalculable damage.

Real-time analytics are also applicable to long-term forensics and auditing. IDC believes the need for detailed forensics will grow as more organizations are held to strict regulatory requirements and as cyberinsurance becomes more widespread. Such forensics will be able to ascertain exactly what data was compromised, how long it had been compromised, and the method of compromise. Such forensics will be an important capability for enterprise IT departments.

Considering Cisco

As Cisco evaluated 3rd Platform trends and the challenges they create, the company found that several criteria had to be met in real time, including the ability to:

- Rewind what has happened in the past, view what is happening in the present, and model what could happen in the future.
- Provide pervasive visibility into every packet and flow across the datacenter infrastructure.
- Identify baseline communication patterns among all application components through machine learning.
- Detect when/where there are deviations with automatic behavior analysis.
- View what happened in the past with forensics and long-term data retention.

The Cisco Tetration Analytics platform provides application behavior insight, policy simulation and impact analysis, automated whitelist policy generation, long-term data retention for forensics, and analysis. It can also detect policy deviations in minutes.

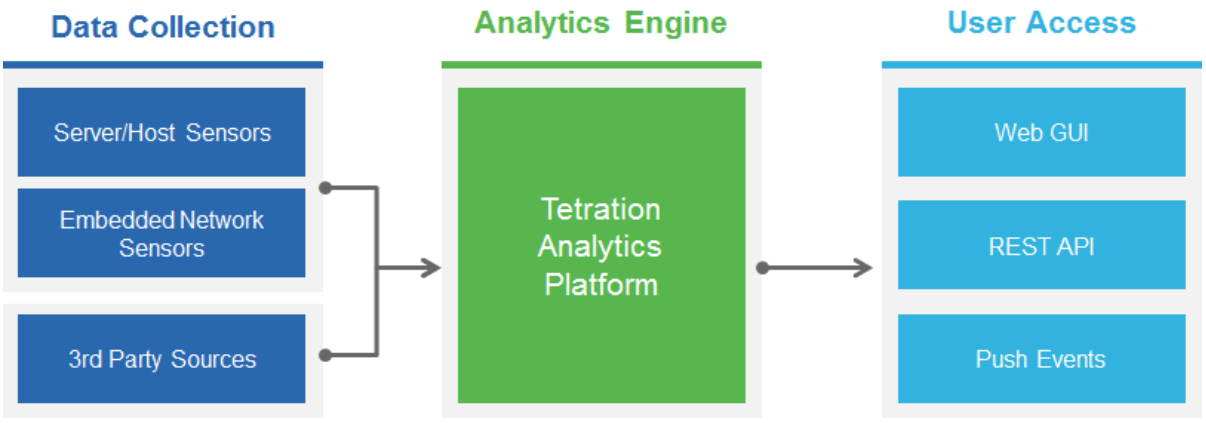
The platform leverages big data technologies as its foundation. Atop this foundation are algorithms that support various use cases, giving customers the flexibility to deploy and operate the platform without the cost-prohibitive requirement for data scientists and in-house big data expertise. It processes millions of flows per second and possesses enough capacity to retain tens of billions of flow records without the need for data aggregation. This feature enables users to perform natural language and granular flow searches and to benefit from long-term data retention.

Cisco Tetration Analytics comprises three essential components (see Figure 1):

- Data Collection, which includes server/host sensors, embedded network sensors, and third-party sources
- Analytics Engine, which is formed by a Tetration Analytics cluster
- User Access, which includes a Web GUI, a REST API, and push events

FIGURE 1

Cisco Tetration Analytics Solution Overview



Source: Cisco

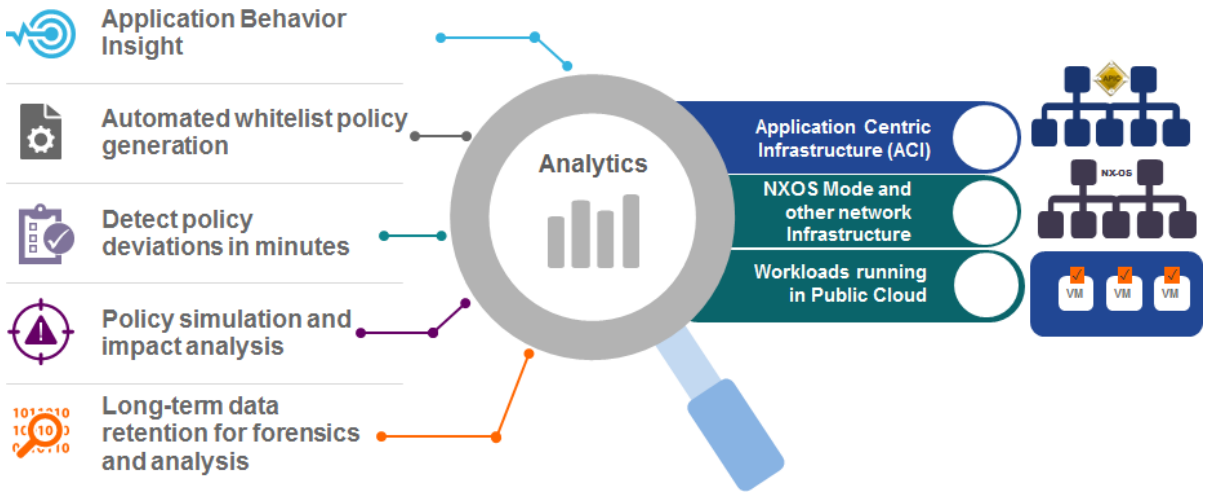
The data collection capabilities are predicated on pervasive sensors — on the host, on the network, and those provided by third parties. The host sensors include those for Linux VMs, for Windows Server VMs, for bare metal, and — in forthcoming releases — for hypervisors and containers. Network sensors are available for Nexus 9200-X and Nexus 9300-EX switches. Meanwhile, third-party sensors will be provided for functions such as geo, who is, IP watchlists, and network and security services including functions such as load balancing.

The host sensors have low CPU overhead and can be enforced by SLA, and the network sensors are built in to the Cisco ASIC. All sensors are secure — with all code signed and authenticated — and while all flows are analyzed, no payloads are exposed.

The Cisco Tetration Analytics platform is a scalable appliance that the customer deploys on-premise. It features a secure design, with role-based access and two-factor authentication. One-touch deployment, self-monitoring, and self-diagnostics make it relatively simple to deploy, use, and manage. It comes with a Web UI, a REST API, and event-based publish and subscribe (pub/sub) notifications (see Figure 2).

FIGURE 2

Cisco Tetration Analytics Platform



Source: Cisco

Tetration's real-time analytics are applicable for long-term forensics and auditing as well as for responding to events as they occur. Tetration's analytics also can automate migration to Cisco's Application Centric Infrastructure (ACI), providing application-level policy enforcement and visibility, self-documentation of the network, and real-time change notification.

Tetration Analytics accommodates agile development processes. The platform can discover and map application dependencies, provide real-time traffic monitoring for policy compliance, deliver real-time policy simulations for impact analysis, and understand how policy modifications affect application performance.

Tetration use cases include:

- **Accelerated application behavior insights:** Understanding application dependencies is critical to successful datacenter operation, application migration, disaster recovery planning, and policy enforcement. With Cisco Tetration AppInsight, datacenter operators can continuously monitor application behavior, establish a baseline for normal behavior, quickly identify any deviation in communication patterns, and plan for disaster recovery. The application behavior insights of Tetration Analytics are derived from real-time communication data between different application components and behavior analysis–based algorithms to identify application groups and their communications patterns and service dependencies.
- **Automated whitelist policy recommendation:** Whitelist policy provides better security, enforces consistent policy across different infrastructures (including workloads running in the cloud), and enables identification of anomalies. This translates into a zero-trust operational model. Once the application dependency mapping is approved, an administrator can download the whitelist policy for the application in three programmatic formats: JSON, XML, or YAML. This policy can be implemented as firewall rules or access lists on Nexus switches. In an ACI context, this can be automatically converted to Endpoint Groups and contracts through import functionality available in the ACI toolkit.
- **Policy simulation for impact analysis:** Cisco Tetration Analytics allows for informed operational decisions through policy simulation that provides in-advance validation of the impact of planned policy changes. Using a Cisco Tetration Analytics cluster, an administrator can simulate a whitelist policy and perform an impact assessment before applying the policy on the production network. An impact assessment can be done using historical data or real-time data, and it does not affect production traffic. This enables an administrator to see in advance how a whitelist policy would impact actual traffic flowing through the network. It also shows which flows will be classified as compliant, noncompliant, or "exceptions." Network administrators can also use this functionality to test policy to see whether an anomaly could have been identified if the policy were in place. Based on simulation and analysis, an administrator has the flexibility to further refine application mapping and regenerate the whitelist policy to accurately reflect application behavior.

Challenges

Cisco faces a number of competitors, including entrenched analytics vendors and several analytics and security start-ups. Some customers will be inclined to look to specialists for datacenter analytics, and they will perceive Cisco as a newcomer that has yet to establish its credentials and qualifications in the market. However, one of Cisco's strengths is its prodigious installed base, including many loyal customers well disposed to new offerings from the network giant.

On its own merits, however, Tetration Analytics has much to recommend it. These include its standing as a relatively novel offering that derives deep telemetry from lightweight software sensors that run on servers and built-in hardware sensors in Nexus 9K platforms, as well as its capacity to deliver real-time analytics to achieve actionable insights from searching billions of records in seconds.

Conclusion

Enterprises confront significant business and technological challenges in pursuit of digital transformation. Yet they often find they lack the systems and tools that can provide pervasive, real-time visibility across their datacenter network. This is true whether organizations are:

- Implementing multicloud strategies that involve cloud migration
- Migrating to an SDN from traditional network architectures and operational models

- Attempting to move to a zero-trust security model
- Detecting policy deviations in real time to be policy compliant
- Planning for disaster recovery and business continuity in the cloud era
- Integrating disparate IT systems following a merger or acquisition

In a world characterized by increased virtualization, cloud computing, SDN, and DevOps — and now containers and microservices — having comprehensive visibility into every flow and packet metadata is essential.

Real-time analytics — based on recent innovations in machine learning, behavior analysis, and visualization — is what facilitates the pervasive visibility required to gain insight into and control over application mobility, security, and datacenterwide infrastructure. For the enterprise, the business benefits that derive from utilization of real-time analytics in the datacenter are compelling. As such, this approach is the missing ingredient needed to enable datacenter operations to respond to the challenges of digital transformation.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com