WHITE PAPER

The Mandate for a Strategic Approach to Management Architectures

Ashton, Metzler
& Associates

Leverage Technology & Talent
for Success

Introduction

There is broad agreement that the rate of business change is faster than it has ever been and that the pace will only quicken. As described below, broad based business changes result in rapidly escalating IT complexity, which makes the IT environment increasingly more difficult to manage and defend.

One of the drivers of the escalating rate of business change is the steps that companies are taking to become a digital business. According to Forbes, “digital business is the creation of new business designs by blurring the digital and physical worlds. It promises to usher in an unprecedented convergence of people, business and things that disrupts existing business models - even those born of the Internet and e-business eras.”

This white paper has three goals. One goal is to explain how the rapidly evolving business and technology environments are significantly increasing the complexity of IT. The second goal is to describe how it is difficult, if not impossible, to manage and secure a complex IT environment with a siloed approach to management data. The third goal is to describe the primary sources of management data with a focus on the ability of that data to enable IT organizations to manage the end-to-end IT environment, no matter how complex it becomes.

The Increasing Complexity of IT

According to a recent report, 89% of organizations have already adopted, or have plans to adopt, a digital first strategy. The report also identified nine technologies that IT organizations must implement to become a digital business, including cloud computing and the Internet of Things (IoT).

The introduction of numerous new technologies is an indication of the growing complexity that IT organizations are facing. Several other factors also contribute to the growing complexity of the IT environment, including:

- When IT organizations deploy new technologies, they often don’t eliminate any legacy technologies. As a result, IT organizations need to manage and secure both legacy and emerging technologies.
- Most enterprise applications are comprised of numerous sub-applications which interact with each other and with subtending services, such as DNS and DHCP. All these interactions need to be managed and defended.
- The adoption of cloud computing has led to new cloud native applications based on microservices. This new architecture greatly increases the number of interactions that must be managed and defended.
- Applications and the data they use can be in several locations, including in a public cloud facility, in corporate data centers, at edge locations, or in branch offices.
- When companies adopt cloud computing, they adopt multiple forms of cloud computing and multiple clouds. On average, companies run applications in almost five clouds.

- On average, companies currently use 16 SaaS-based applications. That number will grow dramatically as roughly three quarters of organizations have indicated that by 2020 they will acquire 80% or more of their applications from a SaaS provider.

Three Key Challenges

The emerging business and technology environments present organizations with three key challenges relative to managing and securing the IT environment. One challenge comes from the fact that as businesses continually adopt a digital first strategy, business processes are increasingly reliant on the IT infrastructure and so if the IT infrastructure is not working well, neither are those business processes. As a result, IT organizations are under growing pressure to reduce the amount of time it takes them to identify and resolve the causes of degraded performance to ensure the highest possible customer experience.

The second challenge is that IT organizations must be able to defend against cyberattacks that are growing in intensity and sophistication. The extent of the security challenges that IT organizations must respond to was described in a recent IBM report. According to that report, “in 2018, many organizations across all industries faced unmanageable levels of cyberthreats brought on by the changing threat landscape, the risk of exposure, and an ever-growing attack surface.”

The third challenge is that, as described above, the IT environment is becoming significantly more complex, and hence that environment is becoming increasingly more difficult to manage and defend. Adding to the difficulty of this challenge is the fact that virtually all IT organizations are suffering from a lack of skilled employees.

The Tactical Approach to Management Data

A key contributor to the difficulty that IT organizations are facing relative to managing and securing the evolving IT environment is that the traditional approach to management involves tactically collecting, storing, accessing and analyzing management data within individual technology domains. As noted, the number of technology domains is increasing and many of those domains are becoming highly segmented as an increasing number of technology variants must be managed within each domain; e.g., both 4G and 5G services will need to be managed.

Another contributor to the segmentation of the management proeces is that there is always an outburst in the development and deployment of management tools associated with each new wave of technology. The dramatic increase in the number of tools is highlighted in the titles of the following three reports:

- 26 Best Network Monitoring Tools
- 45 Best APM Tools
- Top 125 Network Security Tools

In most cases, each tool relies on a small subset of management data. The proliferation in the development and deployment of management tools tends to further accentuate the segmentation of management data.

The ongoing micro-segmentation of the end-to-end management process significantly complicates the process of defending against cyberattacks and managing the infrastructure. It also reduces the probability that a performance or security issue can be quickly identified and eliminated. As a result, a segmented approach greatly increases the probability that IT organizations will have to use a time-consuming approach to management based around the concept of getting all the relevant sub-groups into a war room.

A Strategic Approach to Management Data

To overcome the pitfalls associated with a tactical, domain-by-domain, tool-by-tool approach to accumulating management data, IT organizations need to carefully evaluate the sources of management data they will use in order to create an effective visibility architecture. That architecture must enable IT organizations to quickly and effectively identify both the source of degraded application performance and the likelihood of a security intrusion. As illustrated in Figure 1, the management data must provide end-to-end insight across myriad types of end points, sites, last mile networks, WANs, interconnect providers and data centers – both public and private.

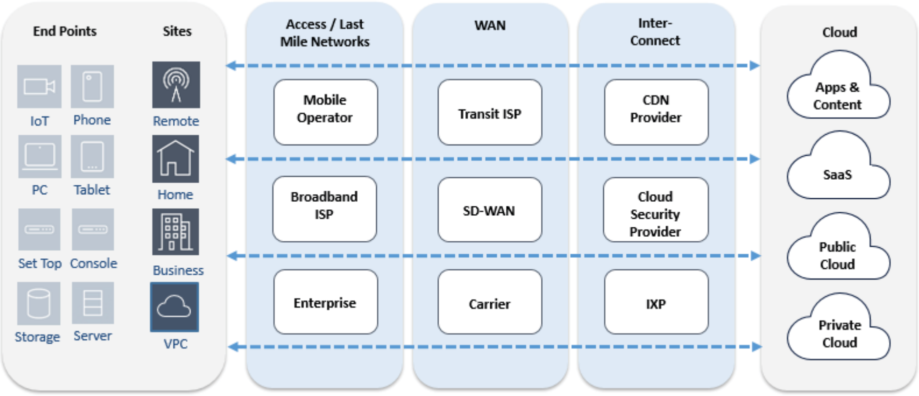

Figure 1: Today's Multi-Domain IT Environment Source: ONUG

Sources of Management Data

There are four primary sources of management data that could be used as part of a visibility architecture. These sources need to be evaluated based on their ability to enable a visibility architecture that it is extensible, scalable, cost effective and which reduces the strain on an IT organization's skill base.

The primary sources of management data are: Agents, Logs, Flows, and Packet Data.

Agents

A management agent is a piece of software that runs on a device and provides an interface to manage that device. While management agents have been in use for decades, this section will focus on a use case for management agents that has grown dramatically in the last several years. That use case is leveraging agents to monitor the performance of applications as part of an Application Performance Management (APM) solution.

APM agents aim to identify bottleneck in application code. Because they are software that runs on a server and generates high volumes of data, APM agents create a performance overhead, as well as large amounts of data that must be uploaded to the APM manager. Increasing resource requirements in this way is particularly burdensome to the adoption of microservices, which, as previously noted, is an increasingly common application architecture.

In addition, managing application agents introduces agility-limiting complexity. One source of this agility-limiting complexity is that agents must be included in an IT organization's configuration and release management processes to ensure that they don't break or interfere with any other software package on that server. This time-consuming process must be performed during the initial installation, as well as any time the agent, the application, or any other software package on the particular server changes.

The biggest limitation of agents stems from the fact that they were developed for well-known application containers line J2EE and .Net at a time with an application was written few programming languages like JAVA and C++. Today's applications use a polyglot of languages like Javascript, Ruby, R, and so on, which APM agents do not support.

As a result, this approach is not only universally not applicable to enterprise management but also creates further segmentation in tool strategy and requires a familiarity with programming and specific languages in use.

Log Files

Logs are an asynchronous and unstructured data source. A log file is a continuous record of events and messages that is automatically generated and contains a record of what happened, when, and by whom. Virtually all components of the IT environment create logs. This includes servers, operating systems, network devices and applications. But in order to be useful, they must be transported to a centralized server for cleanup, indexing, storage, and analysis. For these reasons, logs are considered a noisy or low fidelity source of data.

Logs are popular in part because they are almost ubiquitous. However, there are significant challenges associated with relying too heavily on logs for managing and securing a large IT environment. First is the cost. For example, large companies can generate log files on a daily basis that are hundreds of terabytes or even petabytes in size. That presents challenges relative to collecting, transporting, transforming, and mining all that data, not to mention significant software license fees and hardware costs.

Next is the fact that logs are an asynchronous data source, which in practice are indexed across multiple indexes to better manage the performance issues that result from the volume of logs. Given the different frequency of log categories, the limited storage associated with each index ages out at different intervals. This means there is no guarantee that all the logs will be available when they are needed. They require a continuous engineering and administration effort.

Finally, not everything generates logs. An important example of this is network communications. A fact evident by an early pivot among log management vendors to support virtually all data types, structured and unstructured, including packet data. To successfully use log files, IT organizations must identify upfront the specific use cases for which logs are needed and then turn logs on judiciously and gradually.

Flows

Flows are volumetric information. A flow is a unidirectional sequence of packets between a given source and destination. NetFlow is a Cisco IOS software feature and it is also the name of a Cisco protocol for collecting IP traffic information based on flows. A switch or router outputs a record after it determines that the flow is finished. There are also other vendor's flow technologies, such as JFlow, SFlow, and NetStream, each with less capabilities.

Ten versions of NetFlow and numerous NetFlow variants have been developed. NetFlow and its variants are generally perceived as a low-cost source of management data. That perception is incorrect as there can be significant costs associated with the use of NetFlow. For example, the management data that is generated is sent to a NetFlow manager, which costs roughly $35K to $50K each. Since NetFlow generates a significant volume of traffic, it is common in large environments to have several NetFlow managers. This typically results in a situation where the total cost of the NetFlow managers is hundreds of thousands of dollars.

In addition, there are other significant costs associated with running NetFlow. One cost stem from the fact that performing functionality such as Deep Packet Inspection (DPI) on devices such as routers and switches consumes a significant portion of their processing power even if it is only basic accounting. As a result, these devices prematurely reach capacity and need to be upgraded. IT organizations also pay a price for their use of NetFlow when the switches and routers experience a bottleneck. When this happens, the devices stop sending NetFlow data. This is a high price to pay because one of the times that IT organizations need management data the most is when the network infrastructure is experiencing bottlenecks.

NetFlow has several other major weaknesses. For example, NetFlow does not enable network organizations to understand the user experience; it doesn't provide sub-second granularity; nor does it track UCC call set up and health metrics to name a few. In addition, the number of NetFlow variants further adds to the segmentation of the end-to-end management process.

The latest to join the set of flow-based solutions is flow-cache, the meta data engine behind Cisco's Tetration. Fundamentally, it is another flow scheme with the same limitations. Flow-cache is designed to hold telemetry generated by ASIC sensors in the latest Nexus switch architectures, which are not pervasive yet. In turn this has led to the product offering software agents on workloads to generate the flow information. Flow-cache is largely a solution in support of Cisco's ACI micro-segmentation policy verification and enforcement, as opposed to a full-fledged monitoring tool. While some day the network device may also become a full DPI sensor, that day has not yet arrived. And, to achieve the full potential of such technologies across multi-vendor architectures, that task is best left to purpose-built vendor-agnostic solutions.

For the reasons listed above, most IT organizations that heavily depend on NetFlow almost always also deploy some form of DPI capability. While the motivation to use NetFlow tends to be cost savings, the result is an incomplete and reactive tools strategy, compared to that of modern DPI solutions.

Packet Data

Packet data, a.k.a. wire data or DPI, has been around almost as long as networking itself. It usually requires some passive means of packet acquisition (e.g., taps) and inspection of the packet header and payload to glean insights. In the early days, this was a manual exercise of comparing the hexadecimal content of packet to protocol documentation. Sniffers came next to automate that process, and they were followed by RMON probes that replaced the ad-hoc investigations into continuous meta data.

Today's enterprise class packet data solution is a combination of sophisticated meta data engines, diagnostic and reporting applications, and the means to continuously record the packets to storage for later forensic inspection. This represents a great deal of workflow automation for network and security operation teams. Given that today's solution engines leverage the well-defined structure of communication protocols to generate meta data, they tend to be efficient and highly scalable. For the same reason, they are also highly extensible: when a new protocol is developed, its packet decodes could easily be added to the solution.

Packet data is a powerful management data source, not only because it contains health metrics about the network, itself, but it also contains the unadulterated interactions of the users, their applications, and the infrastructure they depend on. It provides visibility into the pulse of all users, all applications and all the infrastructure from a single normalized vantage point. The ability to retrieve the actual communication, the packets, after the fact without having to recreate the issue is also a game changer, which again provides significant leverage for operations and security teams.

Nevertheless, packet data is not without limitations. One has been the cost associated with deploying taps and appliances throughout the network to achieve pervasive visibility and other the ability to observe inter-process communication on a server. Both limitations have been addressed by industry innovations in recent years. First through transition to software and COTS solutions, some vendors have dramatically reduced the cost of pervasive deployments of packet data solutions, which now include support for virtualization, cloud, containers, and Kubernetes. Next, the transition to microservice architectures has resulted in the inter-process communication being played out across container bridge, Kubernetes pod-to-pod communication, a hypervisor switch, or an API gateway, all of which are easily monitored by the modern packet data solution.

In addition, cloud service providers, such as AWS and Azure, are also responding to market demands by introducing their own Virtual Tap products. These factors combined position packet data as fundamental and powerful building block of the tooling strategy.

Summary and Call to Action

Driven in part by the movement on the part of companies of all types and sizes to become a digital business, the rate of business change is faster than it has ever been and that the rate of change will only increase. This rate of change, combined with the growth in the number of technologies, has drastically increased the difficulty of managing and defending IT environments.

Adding to this difficulty is the continued segmentation of management data that is a result of several factors. One of those factors is that IT organizations typically choose their sources of management data in a tactical manner, in which they use one source of data for a given technology, such as containers, or for a given component of the environment, such as end points. The increasing micro-segmentation of management data significantly increases the amount of time it takes to identify and eliminate performance and security issues.

To be able to effectively and efficiently respond to the emerging business and technology challenges, IT organizations must adopt a strategic approach to monitoring by reducing tool glut and streamlining monitoring architectures. The goal must be an extensible, scalable, cost-effective architecture that reduces the strain on an IT organization's skill base. Advancements in packet data make it the ideal choice as a foundational building block to provide pervasive visibility, which can then be augmented by either logs or agents for specific use cases. Packet data acts as a universal line of visibility into every application, service, server, and user, while its reach into virtualized, cloud, and container environments address the skill gap by reducing the need for niche tools in those environments.