Best Practices for Building Scalable Visibility Architectures

An ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) White Paper
February 2014

IT & DATA MANAGEMENT RESEARCH,
INDUSTRY ANALYSIS & CONSULTING

Best Practices for Building Scalable Visibility Architectures

Table of Contents

Best Practices for Building Scalable Visibility Architectures

Executive Summary

As stakes rise for IT teams, who increasingly find themselves under scrutiny and pressure to deliver high performing, highly reliable, and highly secure application and infrastructure services, the search for strategic advantage continues. Over the past several years, one operations requirement that has been gaining growing attention is the demand for establishing and maintaining deep and definitive packet-based visibility into applications and the infrastructure that hosts and delivers them. Such visibility is essential so that the security team can be aware of normal and abnormal activity and the operations team can shift its alignment from reactive to proactive assurance.

But achieving such visibility is not as simple as we all would wish. Challenges exist with scalability, access, virtualization, and flexibility, all of which render traditional, point-by-point tactical visibility approaches insufficient. Instead, what is needed is a more architectural approach to sustained and reliable visibility – the *visibility architecture*. This paper examines the rise and role of visibility architectures, including essential approaches, attributes and capabilities, including industry best practices as revealed by ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) research.

The Rise of the Visibility Architecture

It's been said that you cannot manage what you cannot measure, and in the context of today's highly flexible, highly dynamic, and high-speed infrastructures, measurement is no small challenge. For network and security operations, measurement means gaining direct insights into application activity from various points around the infrastructure. Such insights can be gained in a number of ways, but most will agree that the most definitive source of truth, when seeking visibility from the network perspective, is gained by inspecting the packets that traverse the networks themselves.

Over the years there have been many techniques employed to gather and analyze streams of packets from the network for security monitoring, network monitoring, application monitoring, compliance monitoring, and other purposes. The number and type of uses for packet-based analysis continues to grow, and yet the number of places where such packet streams can be accessed remains relatively constant. As a result, new strategies are required to build in a more systemic means for harvesting packet streams and leveraging them out for multiple purposes.

A *visibility architecture* represents a strategic investment in packet access and packet stream management technology that will help security, network, and integrated, service–oriented operations teams to establish and maintain a continuous awareness of activity, health, and performance of applications and infrastructure. Further, a visibility architecture represents a new construct and an evolutionary approach versus prior practice of attaching every packet analysis solution to its own dedicated source of packet streams. Conceptually, a visibility architecture is comprised of a number of discrete components and techniques, as represented in Figure 1.

A *visibility architecture* represents a strategic investment in packet access and packet stream management technology

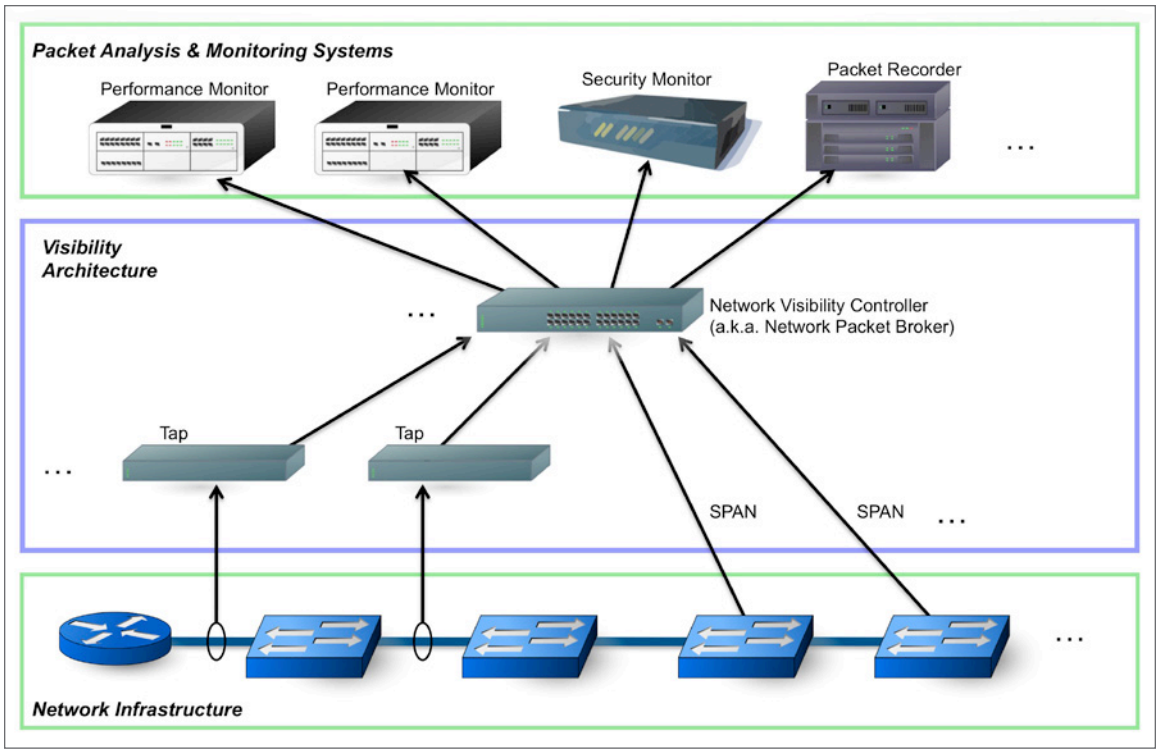

Figure 1: Elements of a visibility architecture

At the first layer, network packets are collected, primarily via one of two techniques – passive taps, or port mirroring sessions (a.k.a. SPAN, RSPAN, ERSPAN). In simple scenarios, such elements are connected directly to the packet analysis tools themselves. But the purpose of a visibility architecture is to support multiple consumers of those packet streams, via an important control element – the Network Visibility Controller (NVC). NVCs represent the latest architectural advancement in packet visibility technology, offering advanced features for collecting, managing, filtering, and distributing packet streams from multiple sources to multiple consuming analysis systems.

EMA chose the term NVC to represent this class of advanced packet visibility products in order to reflect the fact that these elements directly control the destinations of packet streams to be analyzed while also applying filters and manipulations as desired based on the discrete needs of each analysis objective. In the broader industry, this class of solutions has also been called network monitoring switches or Network Packet Brokers (NPBs).

In order to meet its objectives, the visibility architecture must possess a number of important attributes:

- **Scalability**: First, the solution must be scalable so that it can accommodate rapidly rising rates of network link connectivity speeds and, consequently, rapidly rising packet volumes. It must also accommodate growth in the overall managed environment, where a growing number of geographical locations and infrastructure elements must be accommodated and brought into the visibility fold.
- **Sustainability**: Because rising network speeds are a certainty, the visibility architecture must be constructed using technologies that can be readily upgraded to accommodate inevitable higher speeds.

- **Flexibility**: Needs change, new packet-based management technologies emerge and are adopted, and network infrastructures grow and evolve. Solutions must accommodate this change via configuration and adaptation wherever possible, so that actual replacement and upgrade cycles are kept to a minimum.

Drivers for the Visibility Architecture

We already have application architectures, network architectures, data center architectures, and others. Do we really need another "architecture" just for visibility? EMA research[1] indicates the answer to that question is affirmative, due to a number of changes in the nature of the managed environment as well as changes in what is considered best practices for monitoring and management. Following are the key drivers that EMA has revealed in its ongoing research and analysis.

Growth in the Number of Packet-Based Monitoring and Security Tools

Since packets are broadly acknowledged as the most complete record regarding what exactly is running across the network, a growing number of tools and technologies have been enjoying adoption for analyzing and securing the network based on the intelligence available via packet analysis. EMA has documented a steady rise in the number of packet-focused tools that have been deployed within the average enterprise setting, from 2–3 in 2009 to 3–4 in 2013. While this number may seem small, it is significant because only two SPAN sessions can be supported by an individual network switch, meaning that if that same stream of packets is to be processed by more than two analysis tools, some intervening replication technology must be put in place.

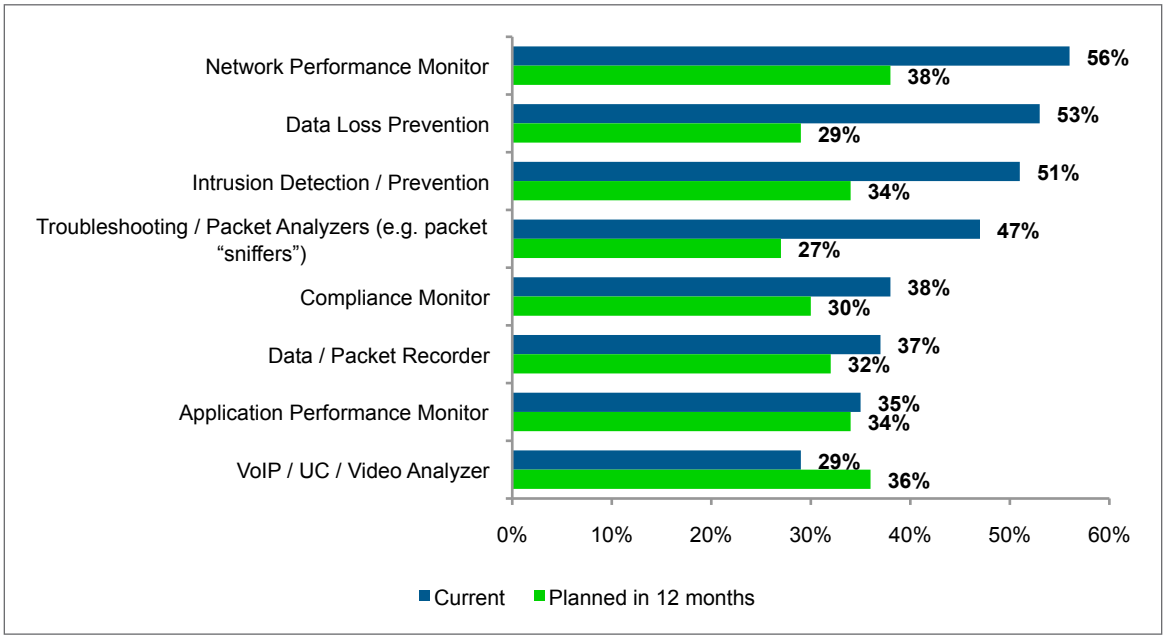

Figure 2: Most common monitoring and analysis tools attached via NVC (September 2013)

Shown in Figure 2 are recent EMA research results reflecting the various types of tools most commonly attached to NVCs. Clearly, there is a mix of security and network tools being deployed and sharing the visibility architecture in most organizations. Based on these findings, it is also clear that future plans are balanced as well, very evenly distributed between monitoring and analysis in both the operations and security realms.

[1] *Network Visibility Controllers: Best Practices for Mainstreaming Monitoring Fabrics*, EMA, October 2013

Growth in Network Speeds

No one can refute the simple fact that network speeds continue to rise in a steady, continuous manner. Regular network refreshes have pushed the mainstream to 10Gbps in the core network and 1Gbps in much of the rest of the network – including growing portions of the access layer. New data center and network build-outs are increasingly based on 40G or even 100G technologies. EMA research validates the expectation of significantly greater deployments of these ultra-high-speed networks in the months to come (see Figure 3). A visibility architecture can act as a resilience layer between network link speeds that rise faster than the ability of packet analysis tools to support higher line rates, by offering the ability to load balance across multiple analysis tools and to adapt/translate between dissimilar network interface line rates.

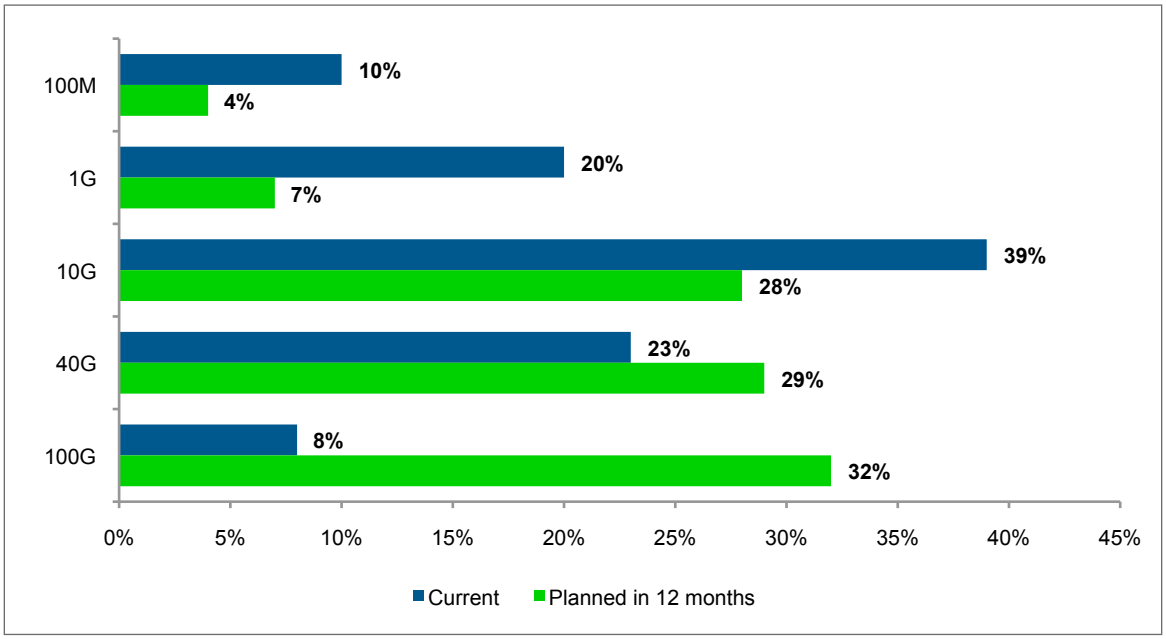

Figure 3: Maximum network link speeds within data center/core network (September 2013)

Resilience/Performance of In-line Security Technologies

While in-line security technologies such as DLP (Data Loss Prevention) and IPS (Intrusion Prevention Systems) have been around for many years, reliability and performance concerns have prevented many organizations from deploying them in a true in-line mode, where they are able to intercept and interrupt traffic. Visibility architecture components are able to provide load balancing across multiple security components, improving throughput and reliability in case any individual element suffers degradation or failure. Interest in this deployment model is growing rapidly, as confirmed in recent EMA research (see Figure 4).

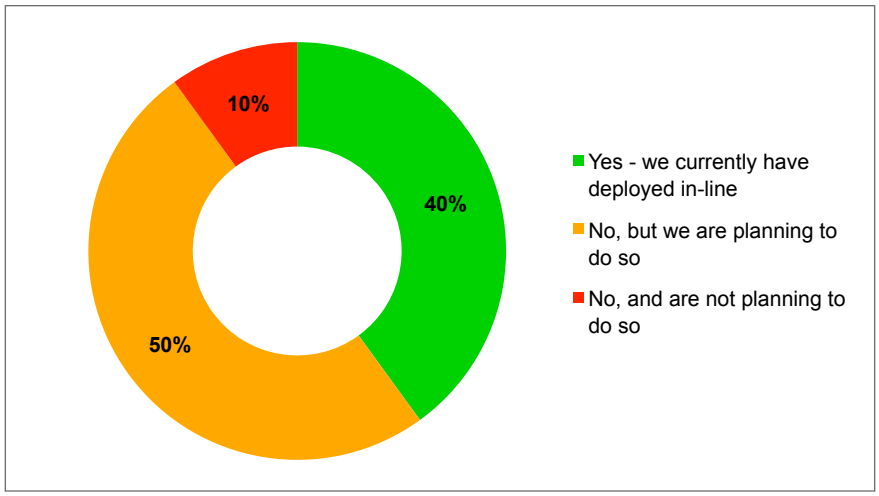

Figure 4: Current and planned in-line use of NVCs (September, 2013)

Alignment with Network Initiatives

The field of network engineering and design is currently facing one of the most disruptive rounds of innovation that has occurred in many years – namely that of automation, virtualization, and Software-Defined Networking (SDN). These innovations will fundamentally change the manner in which networks are deployed and the way in which network behavior is managed, including direct participation in cross-domain, cloud–like orchestration and automation layers. Despite these fundamental shifts the need for visibility remains, and a visibility architecture offers the means to maintain constant vigilance over and into these new dynamic infrastructures.

Building the Visibility Architecture: Best Practices for Enterprise-Wide Packet Monitoring

When considering a visibility architecture, it is important to delve into questions of what, where, and how to proceed. In particular, choices must be made regarding where packets will be sourced, what types of access techniques will be used, whether or not in-line deployments are required, and how virtualized infrastructure will be accommodated. Each of these decision points is discussed in greater detail below.

Packet Sources

Network packets can be drawn from a number of sources and locations around the managed environment. There is no need (nor typically sufficient budget) to draw them from every single link and every single location; rather it is important to find and take advantage of locations in the network that provide the most advantaged viewpoints. EMA research has found that NVCs, at least, are deployed in many places within the infrastructure – commonly in three or more different topological locations. Most common locations include data center core network, top of rack, and data center edge (ingress/egress). These all represent locations where there will be aggregation of network activity, and thus gathering packet streams will yield broad visibility results. Other topological locations that have seen some interest for establishing visibility include campus backbone, remote sites, DMZ, end of row, and backhaul links. While remote sites and backhaul links may represent narrower, more specific monitoring needs, the other locations will also enjoy some degree of traffic aggregation, and thus leveraged visibility.

Also interesting are the results seen within the most recent EMA research regarding where additional visibility deployments are planned within the next 12 months. While teams are still planning to deploy more visibility into the data center core, nearly identical levels of investment are planned for data center edge and remote site locations.

Access Techniques – Tap, SPAN, or Both?

Another important choice lies in the technique by which packets themselves will be harvested and fed into the visibility architecture. There are two primary choices here. The first is the use of passive, layer one physical devices known as taps, and the second involves using intrinsic port mirroring functions, known most commonly as SPAN, available as a feature of network switches. There are advantages and disadvantages to each of these techniques, and most organizations will use some combination of both:

- Taps are relatively simple, passive layer 1 devices that are put in line and split off a copy of packets without interrupting normal network function. This affords the opportunity to avoid generating any load on the infrastructure. The disadvantages of taps are that they require manual installation, including a maintenance window during which connectivity will be broken temporarily, and they must be upgraded whenever network speeds change. Further, some organizations tend to shy away from the use of taps because they represent a break in the delivery path and thus introduce both a resilience and security risk that, while small, is still nonzero.
- SPAN, and related functions known as RSPAN and ERSPAN, offers more flexibility and deployment, as it can be configured directly on the network switch without requiring downtime or breaking a network link. The disadvantages of SPAN are significant, however. Network switches can typically only support two SPANs per device, and are limited in their ability to deliver multiple SPAN sessions to any outbound port by the bandwidth limitations of that port. Thinking of this a different way, you cannot SPAN more than ten 1G interfaces running at line rate out a single 10G outbound port. Further, some switch architectures do not guarantee 100% SPAN integrity, particularly when a device is under load, and thus there is a nonzero risk of incomplete packet transfer and subsequent gaps in visibility. Further still, the use of SPAN can introduce timing inaccuracies, making it difficult to use in extremely time/latency-sensitive settings. And lastly, SPAN commonly introduces packet duplication, whenever aggregating streams across more than one switch interface.

Despite what seems a greater list of disadvantages than advantages for either approach, a visibility architecture must be built upon some combination of these two techniques. Figure 5 shows the four-year trends related to choices regarding the use of these access techniques.

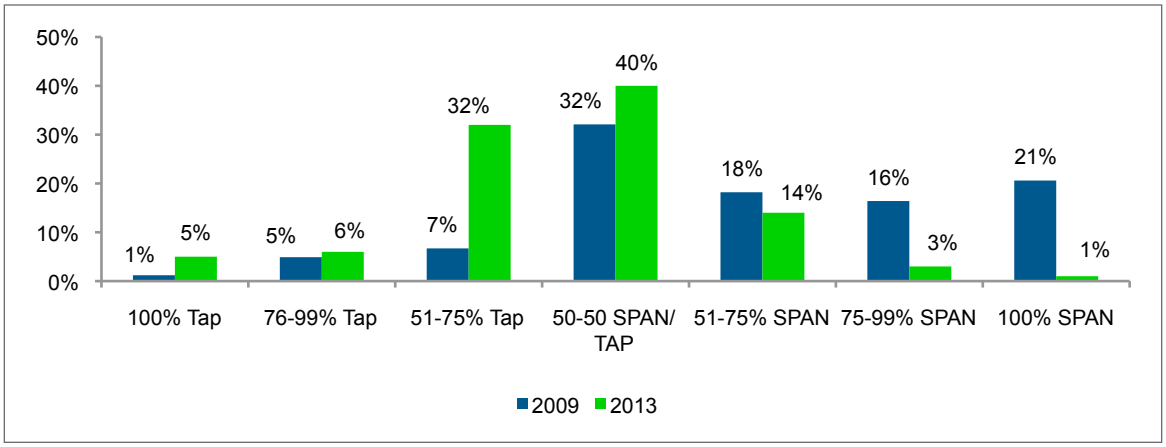

Figure 5: Trends in packet access technique usage

Clearly, very few organizations have chosen to go purely with one approach or the other, and the trend over time has been towards greater use of taps. Fortunately, the two SPANs per switch limitation can be overcome via use of an NVC within the visibility architecture, so those packet streams can be shared with multiple consumers.

In-line vs. Out of Band

While the prime objective of the visibility architecture is to provide insights into performance, health, and security via a passive, "out-of-band" deployment mode, the underlying technologies are capable of going beyond passive techniques to offer active participation in the managed environment via "in-line" deployments. This is particularly applicable for security use cases, where active intervention is highly desirable as a means of mitigating attacks or data leakage.

NVCs that are deployed as part of the visibility architecture contain technology that allows direct inspection of packet streams, so that specific actions may be taken based on what is evident within those streams. Further, a common functionality/feature of any NVC is its ability to provide load balancing across multiple downstream consumers of packet streams. Lastly, bypass technology is commonly available from visibility product suppliers, either embedded within an NVC or as a modular add-on.

This combination of capabilities opens the door to deploying NVCs using an in-line configuration, whereby traffic flowing through the network is actively routed through the NVC as part of the expected delivery path. The NVC then redirects traffic to one or more security devices that provide protective controls, such as active manipulation or blocking of particular sessions or streams. As mentioned earlier, the primary driver behind the in-line deployment approach is the improved performance and resilience that is possible for in-line security technologies. While in-line technologies such as DLP and IPS are considered to be relatively mature, operational concerns persist regarding levels of performance and risks related to any downtime or degradation. In a very real sense, NVCs can mitigate the risks of deploying risk mitigation technologies.

Accommodating Virtualized Infrastructure

The arrival of virtualization within the data center has brought with it a new set of difficulties with respect to visibility. Infrastructure virtualization, be it computer, network, or storage brings with it new components, relationships, and behaviors that cannot be fully recognized, or in some cases seen at all, via traditional monitoring approaches that focus on the physical infrastructure. Some specific areas of difficulty include:

1. **Virtual Switching**: Hypervisors contain virtual networking components, in the form of virtual switches (vSwitches) that can carry packets between Virtual Machines (VMs) without ever crossing the physical network infrastructure. This blind spot means that application and network performance issues, as well as potential security concerns, remain hidden from the view of traditional physical monitoring approaches.
2. **Cloud Environments**: A similar problem occurs when workloads are placed in external cloud environments, particularly when multiple interrelated workloads are placed in the cloud together. Cloud providers typically guarantee aggregate performance, but offer little or no detailed insights to help understand workload behavior and troubleshoot performance issues. Further, cloud providers almost never provide direct access to packet-level visibility as part of their service offerings.

3. **SDN**: Technologies associated with SDN are designed to make the network more flexible and adaptable to application needs, but along the way introduce new layers of abstraction that can interfere with visibility. While EMA research indicates that as of late 2013, only one in five shops was actively using OpenFlow-based SDN, the count of those using some form of virtual network overlays, as indicated by the presence of encapsulation technologies such as NVGRE and VXLAN, is already approaching one in three.

So what are the best methods to address the visibility gap introduced by virtualization? While organizations recognize the need to fill this gap, EMA research indicates that only about half had deployed new monitoring products and configurations specifically for this purpose, although another 40% indicated such steps are going to be necessary. There are a number of techniques available for re-establishing this visibility, and EMA research reflected adoption of these approaches (see Figure 6.)

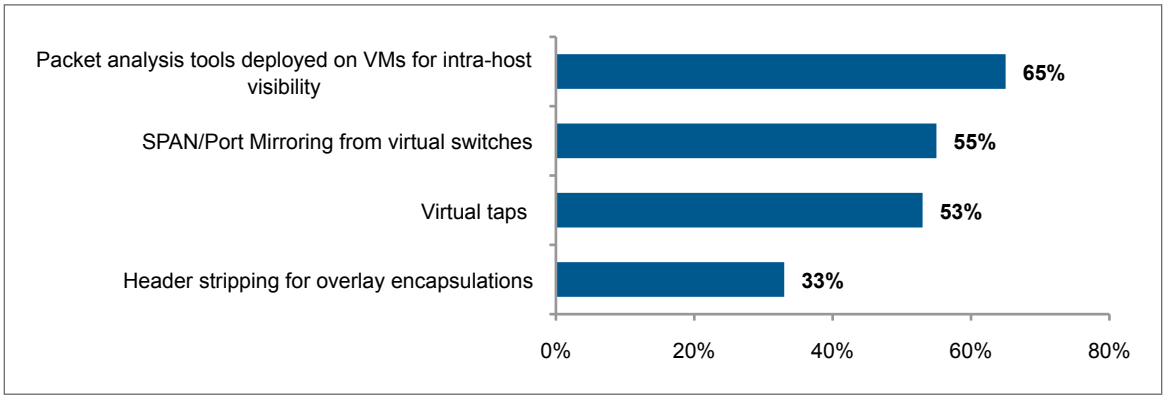

Figure 6: Techniques in use to restore visibility into virtualized environments (September 2013)

Each of these approaches has advantages and disadvantages, and making the best choice depends on specific situations and objectives. With that said, techniques fall into three general categories:

1. **Installing monitoring tools directly into the hypervisor environment.** This was the early, initial approach introduced by monitoring tool vendors to overcome this issue, allowing intra-hypervisor collection and analysis, and typically forwarding metadata results to a central management system for monitoring and reporting. Such an approach can readily restore visibility wherever such tools are deployed. Unfortunately, this approach creates its own problems, as more and more VMs (and thus hypervisor/computer resources) are consumed as the number of monitoring tools increases. Further, re-establishing full visibility means deploying these tools (and more than one) into every hypervisor – an exhaustive approach that has no precedent or equivalent in the traditional, non-virtualized world.
2. **Using new approaches to gain access to the packet streams**: These techniques access packets from inside the hypervisors and route them out to the visibility architecture, where NVCs can be brought to bear for distributing those packets out for monitoring and analysis. Primary options here are the use of virtualized taps (implemented as a VM appliance or a hypervisor kernel agent) or by using the equivalent of traditional SPAN/port mirroring features from the vSwitches. This approach is far more scalable, in that only one virtual tap need be installed per hypervisor, and SPAN functions can be applied to individual intra-hypervisor vSwitches or to distributed vSwitches that tie together multiple hypervisors. The downside of this approach is that it does require use of physical Network Interface Cards (pNICs) on hypervisor host servers, in order to get the packets out of the virtual world and into the visibility architecture.

3. **Header stripping for overlays**: Finally, in order to accommodate the increasing use of virtual network overlays and the associated encapsulations, packet streams that are sent to an NVC can have encapsulation headers stripped before the packet streams are forwarded to monitoring and analysis tools. This reduces processing load at the monitoring tool level. It is worth noting that while current adoption of such features is relatively low, the usage rate is consistent with a number of organizations reporting the presence of encapsulation protocols within their network, and EMA expects this technique to grow in usage concurrent with expected growth in virtual network overlay deployments.

Key NVC Features

Within a visibility architecture, the NVC plays an essential keystone role, offering the ability to enjoy flexibility and resilience for packet based monitoring and analysis product deployments. NVC products offer many specific features intended to facilitate this role. EMA research has found that those deploying NVC solutions have consistently found the greatest value from a specific subset of commonly supported features and capabilities, across all sizes and types of organizations, while other features tend to be used most heavily by certain subgroups.

In general, the following NVC features have found the most immediate broad value within networks across all organizational verticals, types, and sizes:

- **Load Balancing**: One of the core capabilities of an NVC is its ability to act in a load balancing role across multiple packet stream inputs and multiple monitoring and analysis outputs. Load balancing offers the opportunity to stretch tool investments while also improving tool resilience. As mentioned above, it is also an essential part of the in-line deployment model.
- **Filtering**: NVCs offer the ability to apply either inbound or outbound filtering in order to narrow and focus a packet stream on specific activity and traffic of interest. Inbound filters can be applied to restrict visibility into sensitive payload data and/or eliminate traffic streams that are not of interest for monitoring or troubleshooting. Outbound filters allow specific streams to be narrowed and forwarded to specific monitoring tools, for example sending only VoIP traffic to a VoIP monitor/analyzer.
- **Decryption**: As more and more traffic is encrypted as a means of improving security, decryption becomes increasingly important for tools that must be able to peer inside packet streams to understand and troubleshoot performance issues. By having the visibility architecture support decryption, downstream monitoring and analysis tools are relieved of that processing effort.

Many more packet manipulation features are provided within NVC solutions, including time stamping, port labeling, tunneling, masking, de-duplication, packet slicing, header stripping, and more. While EMA research has revealed that all of these are considered valuable, many often become priorities amongst specific organizations or in specific settings. For instance, financial organizations place a relatively higher priority on time stamping features, driven by specific needs for managing ultra-low latency trading environments. As another example, manufacturing organizations see relatively higher value in tunneling capabilities, in part driven by the need to remotely manage sites that are broadly distributed on a geographical basis and that often lack local technical personnel.

From a solutions architecture perspective, great value is placed on administrative efficiency, particularly when the number of visibility devices grows both in quantity and in distribution. The most effective answer to this is the availability of a centralized management platform

Great value is placed on administrative efficiency, particularly when the number of visibility devices grows both in quantity and in distribution.

that can become a single point of access for configuring NVCs and monitoring their status and health. EMA research among those who have deployed NVCs indicates that such centralized management is a truly critical aspect of the solution, particularly among those who have deployed and used the solution for more than six months.

Finally, the visibility architecture is a unique element within the broader IT infrastructure, however it must follow and comply with standard operating tools and procedures as applied to IT as a whole. In the case of NVCs, this translates into two important integration points:

1. **Access and security**: Access to the visibility architecture and the management consoles thereof must be closely controlled and allowed only to those with appropriate skill and clearance, as it is essential to robust security monitoring and the data flowing through the NVCs is often sensitive in nature. As a result, NVC administrative systems must leverage existing user access and security protocols, including systems such as TACACS or RADIUS.
2. **Broader management integration**: The visibility architecture represents an important element within what is typically a broader integrated management and control systems environment. At a minimum, the visibility architecture must include the ability to forward events and log entries to appropriate cross-domain platforms as well as basic health and activity metrics for infrastructure availability and performance monitoring. More advanced capabilities would involve automated closed-loop control actions, whereby external systems (or internal NVC features) support function and configuration change based on observed events or triggers. The most advanced levels would include direct participation in cross-domain orchestration, by which new application loads are deployed and monitoring regimes subsequently defined and deployed within the visibility architecture, all in an automated, programmatic manner.

EMA Perspective

Network and security pros have no shortage of challenges laid at their feet, the latest of which includes dynamic, virtualized, programmable environments as well as an ever widening scope of mobile end points and new application deployments. The dual objectives of ensuring security while also proactively assuring performance require constant, continuous, reliable visibility. Packet-based monitoring and analysis products, both for security and network/application operations purposes, continue to provide the most definitive and complete view into the activity, health, and performance of the infrastructure. But packet-based tools and the efforts required to deploy them fully bring with them significant challenge and risks of their own. The move toward a visibility architecture, put in place to collect, manage, and distribute packet streams for monitoring and analysis purposes, is emerging as the best approach to achieving cost-effective, reliable, and resilient packet-based monitoring and analysis.

While the goal of providing a complete visibility architecture solution is shared by many technology vendors, not all can fully address the needs for today's managed environments. Fully functional NVCs (a.k.a. Network Packet Brokers) are the capstones of a visibility architecture, but by no means the only element. When selecting a visibility technology supplier, EMA recommends that network and infrastructure managers pay particular attention to solution completeness, scalability, and flexibility, with further emphasis on manageability and integration per specific organizational context and needs.

EMA recommends that network and infrastructure managers pay particular attention to solution completeness, scalability, and flexibility.

About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that provides deep insight across the full spectrum of IT and data management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help its clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise line of business users, IT professionals and IT vendors at www.enterprisemanagement.com or blogs.enterprisemanagement.com. You can also follow EMA on Twitter or Facebook.

Corporate Headquarters:
1995 North 57th Court, Suite 120
Boulder, CO 80301
Phone: +1 303.543.9500
Fax: +1 303.543.7687
www.enterprisemanagement.com

2843.021014