ISSUE PAPER

New Suit of Armor: Securing the New Data Center

John Burke, Principal Research Analyst

Executive Summary

In the dispersing enterprise, the spread of service-oriented architectures, rich internet applications, and virtualization dictate that data center security reshape itself. To secure a dynamic infrastructure with fluid boundaries, security needs to become dynamic and virtualized. To enable access rights tailored to specific relationships with staff, customers, partners, and suppliers, it must be identity-centered. To protect transactions using standardized, upper-level formats such as SIP and XML, security needs to become content-sensitive.

The Issue: Moving Targets

Major tectonic shifts in the way enterprises work with and provision their core applications are forcing changes in the way the enterprise has to think about securing them.

One shift is the continuing opening of the enterprise, with the gradual federation and interpenetration of IT systems between an enterprise and its partners, customers, and suppliers. The figurative walls of the data center are being filled with doors, windows, and access ducts, and now serve more as a framework for structuring the flow of information than as a barrier to it.

Another shift is the rise of service-oriented architectures (SOAs). Enterprises are looking to SOA to provide an integration method for their applications, a development methodology and framework, and an overall architecture and philosophy for deploying new functionality. As enterprise applications gain services interfaces, and sometimes are actually atomized and turned into constellations of loosely-coupled services, each service creates on the network a new set of access points; perhaps tens or hundreds of times as many as

there were before. Things that used to happen within an application, on a single server, become network traffic among servers and even among data centers. Some formerly internal functions even become invocations across the Internet of software-as-a-service (SaaS) packages, or services in partner or supplier data centers. Moreover, components in a SOA can scale independently of each other: new instances of an application running on a Java application server might be created to handle peak loads, and then destroyed as the load subsides.

A third shift involves virtualization, which, like SOA, adds dynamism to the data center. Servers can be provisioned and deprovisioned on the fly, "frozen" and "thawed," and moved from place to place. Problems created by rapid (re)provisioning of physical servers are exacerbated and amplified by virtualization. Combine virtualization with SOA and the security environment becomes, potentially, even more wildly variable.

Last, the security threat landscape is continuing to shift, and formerly solid defenses at the perimeter are falling into new rift valleys as the perimeter erodes. Computer crime continues to move more solidly into the for-profit space, and marketplaces for attacks, attack tools, and the spoils acquired with them make the business easier to get into and easier from which to profit. Attacks are climbing the network stack to evade enterprise defenses at the lowest level and target weaknesses at the higher levels.

The new data center, dynamic, distributed, and under attack, requires a commensurate shift in enterprise thinking about security.

No More Business as Usual

In Nemertes' *Security and Information Protection* benchmark, the majority of participants say they secure virtual servers the same way they secure physical ones. Unfortunately, this means significant reliance on segmentation within the data center network, with security appliances such as firewalls and intrusion prevention systems situated between segments to monitor traffic among them. This puts the burden of securing an increasingly dynamic infrastructure on an essentially static, architectural set of systems.

The biggest drawback to network segmentation for security in the emerging data center is operational: it introduces rigidity to the architecture by drawing artificial lines through the company's infrastructure. Both SOA and virtualization intrinsically undercut the idea of rigidly segmenting which physical systems can talk to each other. SOA undercuts it by breaking open the silos that have been built around applications. When one service may serve the needs of six different orchestrated applications, and another may serve a different but overlapping set of six, and so on, it becomes infeasible to segment and segregate the traffic. Likewise, if virtual servers replace physical ones, segmenting traffic requires that either only servers that are allowed to talk to each other be within the same physical resource pool, or that traffic bound from v-server to v-server all be routed out of the physical pool, through security systems, and then back in. Neither solution is optimal, since both limit the flexibility of the infrastructure.

Outward facing communications are also becoming less segmentable, as the number of external entities with which a large enterprise has a unique

security relationship continues to grow. More outsiders need access to more services year by year. If a separate DMZ has to be spawned for each partner, customer, or supplier to which the enterprise needs to grant specialized access to specific internal services, the number of DMZs can quickly grow from the manageable to the ridiculous. The static associations such structures imply make sense in a world where resources are stationary and separable, but is incompatible with the flexible data center and its increasingly interdependent resources.

Organizations want to ensure that they can respond to changes in their markets and deploy new applications rapidly. With such IT agility as a strategic focus, reliance on static barriers for security will make the infrastructure less flexible, and therefore, make the business less agile. Agility is a critical competitive advantage which should not be sacrificed because of security. Consequently, companies implementing SOA and virtualization for agility must likewise adopt agile security to match.

Higher Level Consciousness

The changeable nature of the new data center is only part of the problem for security, of course. Another major part of the problem is the shift in application and attack focus up the network stack.

Although attacks at layers two through four are still active and dangerous, security at those levels is also relatively strong and increasingly ubiquitous. However, security on layers five through seven lags. At the same time, enterprise applications are rapidly changing in both back-end architecture and front-end implementations. The move to SOA drives applications to swap internal or binary communications for externalized XML interchanges. The move to unified communications puts SIP into the center of converged and integrated voice and data systems.

Enterprise operations are now being driven by XML documents, SIP sessions, SOAP objects, and the like, and this exposes the enterprise to attacks based on those formats and on the content conveyed within them. Criminals now aim attacks at compromising parsers for any or all of these formats, hoping to break into the system hosting an application by feeding it poisoned content in the same way they once sought to crash or compromise routers by feeding them carefully malformed packets. They might also seek to compromise an enterprise not by breaking into or taking control of a system but instead by using systems for their defined purposes but towards bogus ends: in a SOA based inventory system, for example, well-formed but bogus purchase orders created by criminals could empty a warehouse and send millions of dollars worth of goods to random or non-existent addresses. And the back end is only part of the problem: Web 2.0 front ends carry some of the same problems into the client side of the picture, as XML traffic drives the content of client interfaces. Corrupting or hijacking those streams could effect anything from a simple denial of service to password theft, data theft, or operational sabotage.

In order to address such challenges, organizations must adopt security technologies that are able to protect the upper reaches of the network stack in the same way that firewalls protect the lower levels.

Custom Tailored Fit

The security architecture of the emerging dynamic data center has to address both the mutability of the infrastructure and the fact that so much function will be channeled through standards-based, upper-layer formats such as XML and SIP. It must itself be dynamic and virtual, identity aware, and both format- and content-sensitive.

Dynamic security will match servers and applications in its mobility, tracking them through their production lifecycles: encrypted VPNs, access control lists, or even complete virtualized security appliances will come into existence when the servers and services they need to protect do, and will disappear again when those servers or services disappear. Some of this dynamic security will be based within the virtual servers or dynamically-instantiated applications themselves. Some will exist within virtual environments, some at the hypervisor level. The parts outside these virtual environments will be aware of what has happened inside them and adapt as components come and go. (Please see Figure 1: Security in Dynamic, Virtualized Environments, below.)

Identity becomes central to security in this new data center, the base for defining security around data and among systems. Identity management will encompass, not just users, but also these dynamic, transient components and

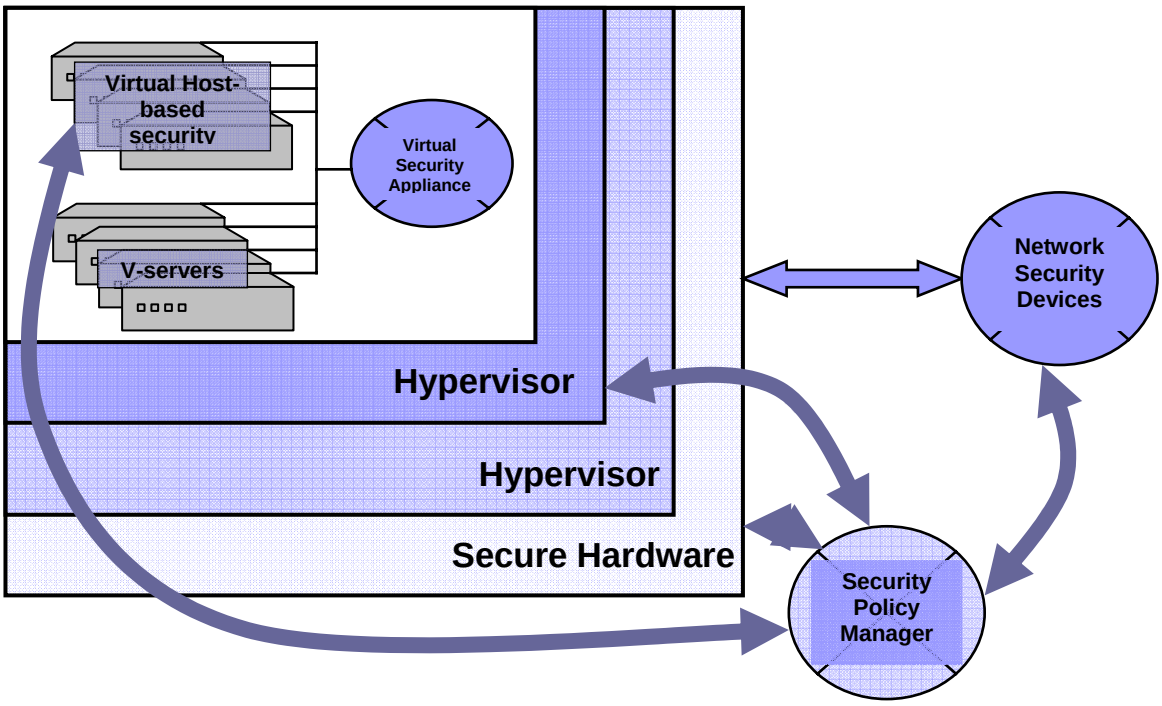

Figure 1: Security in Dynamic, Virtualized Environments

systems. Policy will define which of them are allowed to talk to which, with identity independent from (but not necessarily insensitive to) location. Digital signatures will allow components to be identified and trust assigned, and will form the basis of conversation encryption. Federation will allow the integration of security rules across cooperating organizations. User identity will be propagated among components as well, of course: component identity can be used to make sure that only the right parts of the infrastructure try to access protected resources, and user identity to make sure that they are doing so only on behalf of those users with the appropriate privileges.

Content-aware security will both ensure traffic among components is well-formed, and watch for and alert on unusual traffic that is otherwise well-formed. Traffic may be inbound, and indicate an attempt to compromise systems, or it might be outbound, evidence of an attempt to leak sensitive data. Either way, data center security will have to be aware of the content in order to properly secure it.

In order to achieve all these ends, securing the new data center will ultimately require integration of security across all categories of data center systems: networks, appliances, servers, storage, and applications. This, in turn, places a high premium on strong, standards-driven interoperability. Such integration will have to be both reactive and proactive. Reactive, in that any component should be able to alert the others that something odd is happening. Proactive, in that the configuration management and provisioning tools driving the creation and destruction of virtual servers and services will be able to trigger the necessary changes to security in anticipation of those events.

Conclusion

Security has once again trailed production environments somewhat, an afterthought dealt with once the operational bugs have begun to be shaken out of virtual environments, SOAs, and Web 2.0. Enterprises with functionality out well ahead of matching security will have to play catch up on security again, with the predictable and oft-repeated consequences of confusion and expense. Enterprises just embarking on their own quests for agile and dynamic IT infrastructures will have the chance to build a well-fitted suit of armor as they go.

About Nemertes Research: Founded in 2002, Nemertes Research specializes in analyzing the business value of emerging technologies for IT executives, vendors, and venture capitalists. Recent and upcoming research includes Web services, security, IP telephony, collaboration technologies, and bandwidth optimization. For more information about the analyst, please contact Nemertes at research@nemertes.com.