nemertes RESEARCH INDEPENDENCE INTEGRITY INSIGHT

issue paper

The Next Stage of WAN Optimization

Days are Numbered for Standalone Optimizers

By John E. Burke
Principle Research Analyst, Nemertes Research

Executive Summary

WAN optimization appliances constitute a growing market now, but the writing is on the wall: Seven years from now, "WAN Optimization Controller" may barely exist as a separate category of solution. Migration of many WAN optimization and application acceleration functions into carrier/cloud services (optimization as a service), virtualization of optimization appliances, and integration of key optimization features such as traffic shaping and compression within routing, switching, and security appliances and even in operating systems make the prolonged existence of standalone WAN optimization appliances far from a sure thing.

The Issue

With adoption by 61% of organizations, nearly all of it in the form of symmetrically deployed appliances, WAN optimization is one of the most broadly deployed application delivery optimization (ADO) technologies. However, penetration of optimization is quite low, especially in companies with large numbers of branch locations. For example, in those with more than 250 locations, the average reach of optimization is only 15% of sites. (Please see figure 1.) This highlights some intrinsic challenges with the appliance-based optimization:

- **You have to put appliances on each site.** This means per-site capital expenditures and increased inertia, as each branch is "heavier" by one or more boxes in the wiring closet. (More than one will be required to provide redundancy, and depending on the solution deployed and the size and needs of the site, additional boxes may be required for management, for visibility and reporting, and for support of mobile users with soft clients.) Heavier branch infrastructure makes the branch that much harder to "turn up," shut down, or repurpose later. This runs counter to the virtualization of the enterprise, with its concomitant spread across space and need to

support agile change in where and how long a branch exists. As the director of networks at a large university put it, "Added boxes are bad."

- **You have to manage them all.** IT takes on both asset and operational management of another layer of appliances, and often (in more than two-thirds of organizations) a new vendor relationship. This adds complexity to the environment, one more thing for IT to monitor and worry about, one more skill set to maintain. The management challenge is heightened for branch locations lacking on-site IT support resources.
- **You have problems with scale.** Given the problems of capital and operating expense, and of complexity, scaling up an appliance-based strategy creates significant challenges for IT. However, the spread of unified communications and other new tools on the one hand and the rise of cheaper but less consistent WAN connectivity on the other are driving increasing interest in deploying optimization more broadly. Increasingly, IT wants to scale up deployments.

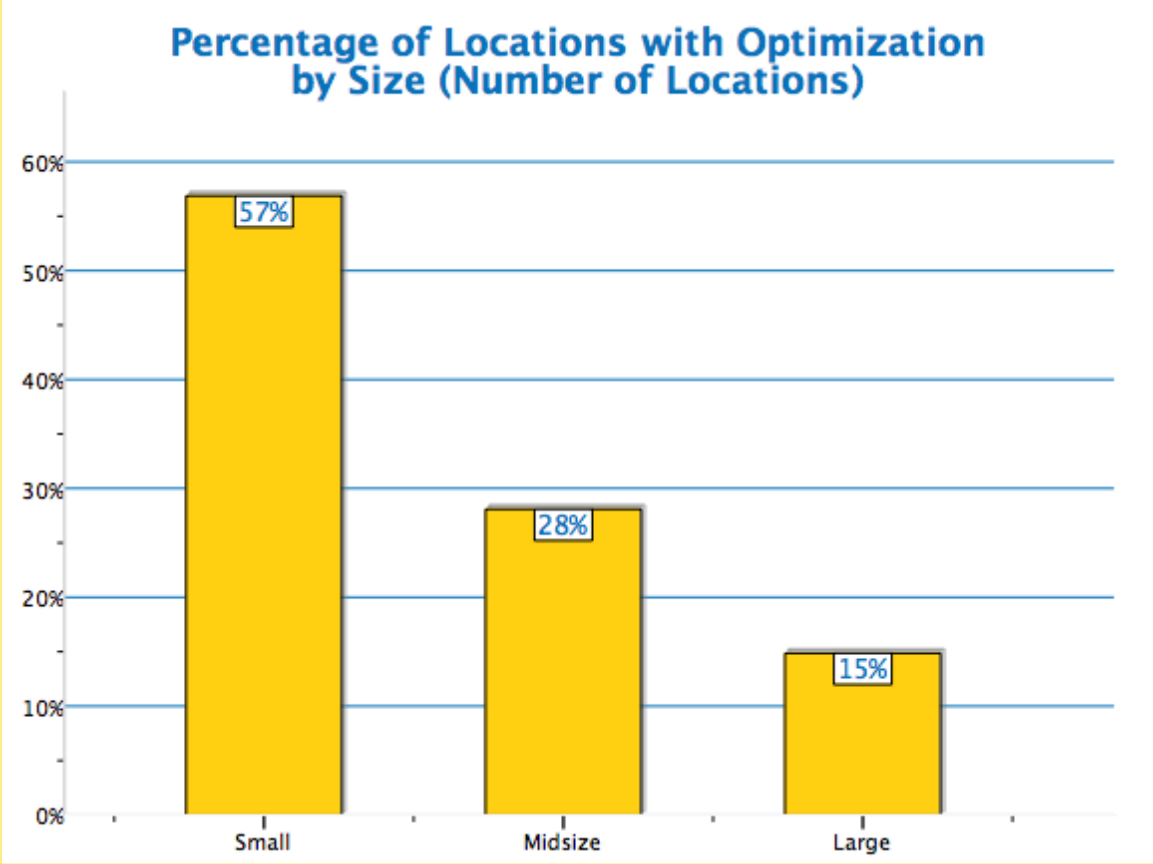

Figure 1: Depth of Optimization, by Number of Sites

As a result of these challenges in the appliance-per-site model, vendors and service providers are delivering alternative technology strategies and deployment. Solutions fall into three categories:

- Optimization as a workload,
- Optimization as a feature,
- Optimization as a service,

Together these new options increase the opportunity for IT architects to realize the promise of optimization.

Optimization as a Workload: Virtualized Appliances and Point Solutions

In recent years, WAN optimization controller (WOC) vendors have delivered virtualized versions of their wares. The idea is that for smaller branches, where IT is already looking at using virtual servers to consolidate essential branch services (DNS, DHCP, print services, etc.) onto just one or two host servers, the virtual WOC can share hardware with these other services.

This eliminates the need for a separate box in each location, *if* there is already adequate hardware in a typical branch, and thereby reduces administrative costs and reduces the weight of the deployment. Virtual appliances are easier and faster to deploy, and to move.

If there is not adequate hardware on site already to handle the optimization load (e.g. no server in the branch, or only servers with insufficient processing power or memory, or inadequate network interfaces), then IT will have to deploy new or beefier hardware. This means increased capital outlay to make sure the virtual WOC can do its job. (And of course, if there is no server there yet, then IT is just adding another box to the site, just as if it were adding an appliance; although on the plus side, it is not a "unitasker" and IT can use it to deliver other services, too.)

The virtual appliance approach also allows some shopping around for best-cost platforms, or, conversely, the ability to use the organization's preferred server hardware vendor. This last scenario is an option IT administrators welcome—Nemertes benchmarking finds 50% would be just as happy running a virtual optimizer as a physical one, if performance was equal.

Of course, running virtual appliances even on existing hardware doesn't eliminate capital costs, as pricing for virtual appliances is carefully gauged to forestall cannibalizing physical appliance revenues. And although it can reduce administrative costs, there remains the need to manage a box (although virtual) in every location, meaning additional management expense.

Running virtual appliances even on existing hardware doesn't eliminate capital costs, and vendors price virtual appliances to limit competition with their own physical appliances.

Although it can reduce administrative costs compared to physical appliances, there remains the need to manage a (virtual) box in every location.

As an alternative to deploying optimization as an encapsulated workload in the form of a virtual appliance, IT can run optimization as a workload alongside the services being optimized. For example, a Web traffic optimizer can run alongside the Web server on the same virtual or physical server, as a system service rather than a virtual appliance. This approach is often enormously cheaper than deploying appliances (physical or virtual) to branches, but it only optimizes one kind of traffic, such as Web traffic, and can only control traffic to the one server. Proliferating point solutions for different services would lead to complexity and management issues, as well as the likelihood of exceeding the cost of a network-based approach.

Optimization as a Feature: Integrated Optimization

Another way to get out of the "box per branch" trap is to get optimization integrated into the feature set of other branch office devices already on-site. For most sites, this would mean a router or a security appliance.

As with virtual appliances and host servers, the immediate issue is: does the branch have a beefy enough box in place? It now has to take on optimization duties as well as what it was initially deployed to do, e.g. route or filter traffic, and optimization is often IO, CPU, and memory intensive. Where the existing box is not capable of the additional work, IT needs to beef it up to add the new features, a capital cost on the same order as putting a separate appliance in place. An IT manager in a midsized professional-services firm, commenting on the idea of integrating optimization into his routers, says, "We would like to have fewer boxes, but we would have to replace *everything*." He, and many like him, have their eyes on integrated optimization as they count down to the next refresh of their routing infrastructure, but in the meantime are looking elsewhere for optimization solutions.

Integration of optimization as a tool set should dramatically decrease management loads. There is no separate asset to track, no separate vendor to work with, and no separate "golden" configuration file to manage: everything is integrated into the existing solution. However, that integration has a less-positive side, too: it increases vendor lock-in. Combining the layers of function makes it harder to seek alternate solutions for optimization or for the underlying tool set (routing or security). Separate layers can be replaced independently, as changing requirements necessitate or changes in features and pricing allow. Where integrated, the organization is more likely to have to make do with a feature set that doesn't meet all its optimization or other needs; the old "best of breed" versus "single vendor" dilemma again.

In Nemertes' 2010-2011 benchmark of IT professionals, in more than 200 organizations, 57.1% of participants said they would like to have optimization integrated into their core packet-delivery gear (e.g. routers), and 23.8% said they would not like that. Rather fewer were interested in integrating security and optimization, with only 35% preferring that and 40% saying they were not interested at all. About 21% are uninterested in either kind of integration.

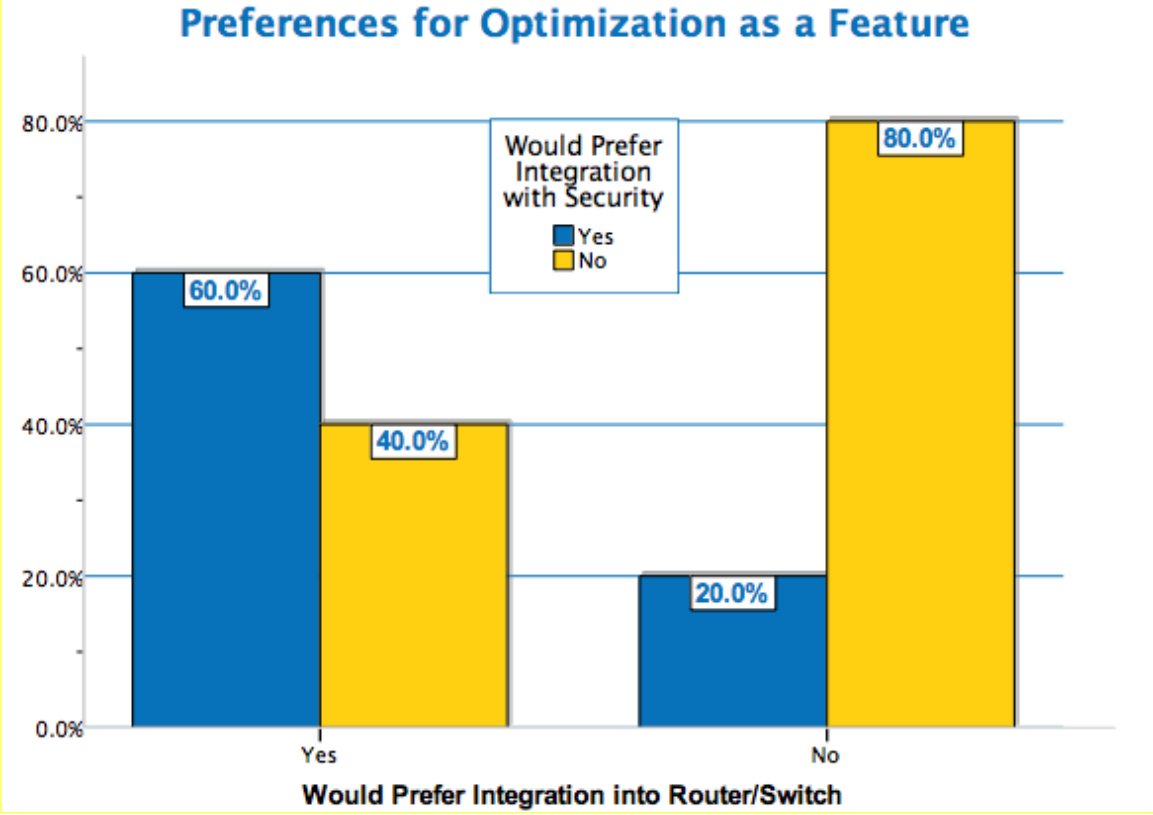

Figure 2: Relationship Between Integration Preferences

Optimization as a Service: Managed Optimization

The third major alternative to an in-house deployed and managed optimization appliance layer is to purchase optimization as a service. This can happen in one of two ways: via a managed on-premise solution, or via a pure-play service-provider cloud solution.

The managed appliance model has the potential to solve many of the problems of the in-house appliance model. It shifts management onto a third party. Although IT necessarily is still responsible for setting policy on what kinds of things get preference and what gets stepped on when congestion threatens, someone else gets to translate those policies into settings and configurations. Using a managed-appliance model can relieve IT of the burden of buying/owning the appliances, although this point may require fierce negotiation with a service provider. An OpEx-only cost model makes branches more lightweight from the organization's perspective. Likewise, IT can negotiate (and typically pay extra) for ad-hoc access to application performance and usage data on top of basic canned reporting.

On the flipside, as with all managed services, the organization loses both responsiveness and a measure of control in such scenarios. "We don't know what we need until we need it, but these contractual relationships tend to be inflexible,"

notes a CIO in a small healthcare company. And, while it is possible with some service providers to get an OpEx-only arrangement, it is far from the norm; generally, the organization is still expected to buy and own the appliances, even if the purchase cost is rolled into the monthly service fees. Nemertes sees interest in managed-appliance optimization highest among managed-router users.

Optimization as a Service: Cloud Optimization

Contrast the managed device model with the emerging option of optimization in the cloud. Instead of deploying appliances (or having someone else do so) IT routes WAN traffic through a provider cloud and all optimizations would take place there. Connect the data center to the provider cloud, connect branches to the cloud, and traffic is optimized.

In a truly cloud-modeled optimization service:

- No capital costs!
- Radically reduced management cost. IT worries about policy only. IT does not have to worry about the deployment, management, or space/power impact of optimization appliances in branch network closets, as there aren't any. There are no hardware, firmware, or software updates to worry about, and services can be consistently up to date in all locations. Obsolescence becomes a service provider issue, and IT can focus on service-level agreements and performance metrics instead.
- Pay-as-you-go optimization: apply the service where needed, when needed, as long as needed. True cloud services are distinguished in part by "pay by the drink" pricing and on-demand access. A cloud-modeled service will apply these concepts to optimization, allowing organizations to add and remove optimization to a branch on a short time scale, according to business cycles or needs: month to month, potentially day by day or even hour by hour. Yearlong or multi-year commitments should become a thing of the past, reducing costs and making the organization more agile.

One significant shortcoming of the pure cloud model is that it eliminates the possibility of compression on the branch-to-cloud leg of the WAN. If part of the goal of optimization at a particular site is to reduce bandwidth, the pure-cloud model breaks and the provider must install customer-premise optimizers (whether physical or virtual). This necessarily increases the "weight" of the solution—the pure-cloud service is weightless—and increases provisioning times and overhead (for optimizers to be sent/installed/configured) as well as de-provisioning overhead. If the service is cloud modeled, though, any end point equipment remains property of the provider, and will be entirely managed by the provider. Consumer deployments of DSL and cable modems suggest the "weight" of such boxes can be brought very low.

Another shortcoming of the pure-cloud model is the increased potential for vendor lock-in, as with integrating routing or security with optimization. The

more services you layer onto your WAN, and the more your provider's services determine your ability to manage applications' network performance, the harder it is for you to easily or transparently change providers. A benchmark participant in financial services says that he explicitly keeps performance management as a layer above his MPLS provider so that he can swap out MPLS providers without affecting performance.

A third shortcoming, pointed out by benchmark participants, is that IT often has enough complaints about carriers' ability to carry traffic to make adding performance management on top seem like a bad idea. A CTO in a financial services firm says, "I need the carriers to do less, not more. They need to get back to the business of providing pipes. I'm in a business where milliseconds matter. I can't turn that over to the carriers." However, carriers see the "just pipes" business as a losing one: plumbing is a commodity and has accordingly slim margins for them. Layered services are their revenue future.

Despite these concerns, as cloud optimization offerings multiply and mature, Nemertes expects them to become the platform of choice for a significant percentage of enterprises. Organizations where the benefits (especially in agility) can strongly outweigh the shortcomings include organizations that:

- Need to bring new sites on line frequently and would benefit from being able to do so as quickly as possible.
- Have short-lived sites.
- Have seasonal or cyclical need for optimization.
- Are worried more about application acceleration and strengthening network reliability (through packet loss mitigation, for example) than about bulk data compression.
- Want visibility, but don't need optimization (a carrier cloud can easily provide usage and performance data for all connections from a single access point).
- Prefer managed services to capital investment and staff growth.
- Are scaling up use of cloud infrastructure and SaaS.

IT's inclination to use similar services is a good indicator of interest in optimization-as-a-service models. Those who would prefer a carrier/cloud solution already spend about three times as much of their IT budget on carrier services as those who are not interested. (Please see Figure 3.)

Conclusions and Recommendations

WAN optimization appliances constitute a growing market now, but the writing is on the wall: A few years from now "WAN Optimization Controller" may barely exist as a separate category of physical network appliance. As WAN refresh cycles play out over the next seven years or so, Nemertes expects a majority of companies to migrate to one of the other optimization models now available: carrier/cloud services (optimization as a service), point solutions and virtualized optimization appliances (optimization as a workload), or integrated

router-switch/optimizer or security/optimizer appliances (optimization as a feature).

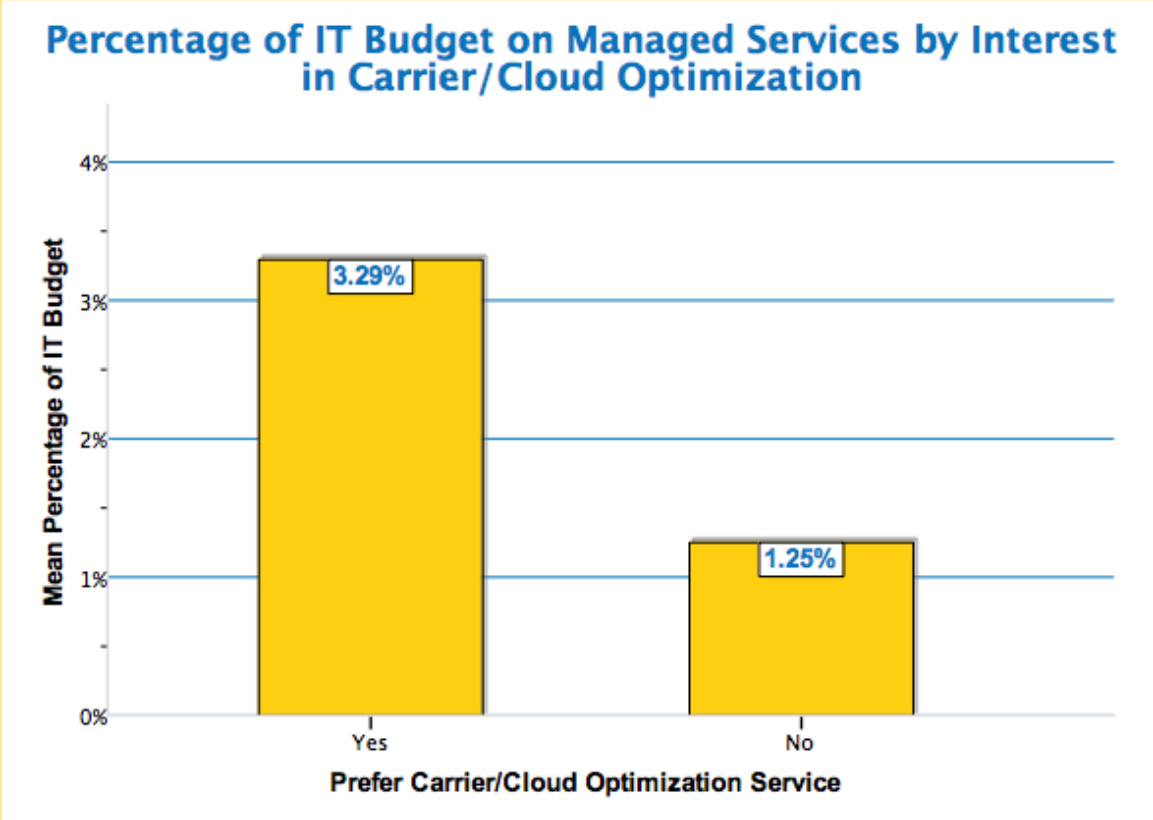

Figure 3: Relative spend on managed services versus interest in carrier optimization

Network architects and WAN managers should make plans accordingly:

- Look at both capital and operating expenses for traditional deployment.
- Look at how management of the optimization layer integrates with performance management and network management overall.
- Look at organizational plans for change in branch numbers and lifecycle.
- Determine your tolerance for the risks associated with various models.
- See whether a managed or cloud service fits your sourcing philosophy.
- Look at optimization as a feature, a workload, and a service, and pick the model—or mix of models—best fitting current and expected needs. It could be a mix, say, of in-house among data centers and cloud for branches; or managed for data centers and virtual for branches. One size does not fit all businesses, nor necessarily all use cases in any one.
- And don't forget to take into account geography and compliance. If an optimization requires caching of data, where will that data live and how will it be protected?

About Nemertes Research: Nemertes Research is a research-advisory firm that specializes in analyzing and quantifying the business value of emerging technologies.

You can learn more about Nemertes Research at our Website, www.nemertes.com, or contact us directly at research@nemertes.com.