Virtualization and Infrastructure Optimization
REFERENCE GUIDE

IN THIS GUIDE:

Gain the most value from your enterprise infrastructure via virtualization and other optimization solutions.

CDW·G

800.808.4239

CDWG.com/virtualizationguide

Virtualization and Infrastructure Optimization REFERENCE GUIDE

TABLE OF CONTENTS

What is a CDW·G Reference Guide?

At CDW·G, we're committed to getting you everything you need to make the right purchasing decisions — from products and services to information about the latest technology. Our Reference Guides are designed to provide you with an in-depth look at topics that relate directly to the IT challenges you face. Consider them an extension of your account manager's knowledge and expertise. We hope you find this guide to be a useful resource.

Seeking the Infrastructure Ideal

Virtualization and other infrastructure optimization solutions are key to improving operations

IT leaders are on a continuous quest for a more efficient, flexible and better-performing infrastructure that supports the mission requirements of their organization. To meet these goals, they're turning to virtualization and infrastructure optimization tools that help make the most of existing hardware and software investments and reduce the administrative demands placed on overextended IT staffers.

Infrastructure Optimization

Optimizing the infrastructure can be accomplished a number of ways. These options are all aimed at speeding applications across the network and ensuring quicker access to data, while minimizing storage costs.

Optimization can be a complex and multifaceted undertaking. It's often viewed as a never-ending task, as each change can affect network and storage performance in unexpected ways.

Regardless, infrastructure optimization can deliver some strong benefits including:

- **Rightsizing:** Organizations can reap savings on operating expenses (OPEX) and capital expenditures (CAPEX) because of rightsizing and timely provisioning of resources.
- **Reduced staffing costs:** Organizations can gain savings from staff realignment or mitigation thanks to reductions in management requirements.
- **Stable production capacity:** An optimized infrastructure reduces the risk of depleted capacity and the possibility of production outages.
- **Improved visibility:** Organizations gain greater visibility of resource utilization and performance.
- **Better performance:** Enhanced bandwidth efficiency leads to increased application speeds and a resulting reduction in operating costs.
- **Smart storage utilization:** An optimized environment examines real storage needs and allocates dollars and capacity among storage tiers.

Virtualization

Virtualization is a method of decoupling an application and the resources required to run it — processor, memory, operating system, storage and network access — from the underlying hardware host. A layer of software called the hypervisor provides access to the host and communicates directly with the hardware.

In practice, this means a given host can run several virtual machines — up to 320 virtual machines, depending on the solution. By virtualizing servers and applications, an IT shop can operate with fewer physical servers. In theory, the higher the ratio of virtual machines to physical ones, the greater the efficiency.

The concept was first demonstrated in the early 1960s, and since then the word *virtual* has appeared in the branded operating systems of vendors in several successive generations of computing architectures. But virtualization itself isn't only something you buy. It's a process, a state and a set of products. Together they can achieve several benefits for organizations.

WHERE DO YOU STAND?

Forrester Research has developed what it calls the Virtualization Maturity Model for organizations virtualizing their IT.

- **Stage One:** Evaluate and become familiar with the technology.
- **Stage Two:** Virtualize servers: Change operations to backing up virtual machines and begin to move applications to VMs.
- **Stage Three:** Automate the environment and dynamically balance loads; achieve higher ratios of VMs to physical hosts.
- **Stage Four:** Virtualize clients and client applications and thus improve virtual environment disaster recovery.

What Virtualization Offers

Without question, cost reduction is the primary reason that IT organizations launch virtualization projects. Savings occur in the following ways:

- **Reducing the number of servers:** Whether in a stand-alone, rack-mount or blade form factor, a key benefit of virtualization is fewer servers, which in turn results in lower capital expenditures.
- **Reducing energy costs:** Fewer servers require less energy and cooling capacity to operate them.
- **Diminishing real estate costs:** Virtualization often enables data center consolidation, leading to the closure of facilities or a reduced footprint.

There are also productivity savings and efficiencies that can result from virtualization:

- **Faster deployment:** With today's virtualization packages, IT staff can create new virtual machines almost at the touch of a button, or increase capacity on the fly as needed.
- **Greater f exibility:** When requirements are modified, workgroups change or products move into development, virtualization allows supportive IT assets to keep pace more readily than a straight physical environment is able to.
- **Easier administration:** With fewer servers and storage area network (SAN) devices to manage, IT staff, data center operators and network administrators are free to focus on more strategic tasks.

The gains are real enough, but realizing them is a complex undertaking that requires careful planning. Virtualization results in more efficiency and speed, but it's not a simple plug-and-play process. This guide will cover many of the questions and considerations you'll face as you explore this technology.

Virtualization Trends

Although virtualization has been around for a while, its adoption continues to accelerate. Very few IT organizations have not considered it.

Commercially, vendor and product offerings have developed around the main hypervisor vendors. This development takes many forms.

For example, the leading computer hardware manufacturers ship machines loaded only with a chosen hypervisor on their "bare metal" — virtualization-ready machines, including traditional servers and blades. Tools are becoming available that give better visibility into virtualized servers. And cybersecurity vendors are honing hypervisor security tools.

A growing number of CIOs and system administrators are looking for all-in-one monitoring and management tools for their networks, tools that encompass the virtualized segments of the infrastructure. The proliferation of tools has caused a host of management problems, and in the coming months and years the industry will undergo consolidation.

Support for handheld clients is emerging from several virtualization developers, giving secure access to powerful devices that started out in the consumer realm but are migrating into government agencies and educational institutions. Overall costs for virtualization solutions are falling.

Other Infrastructure Optimization Solutions

In networking, a key trend is accelerated convergence of physical transport media and of protocols, which makes WAN optimization all the more important. Ten gigabit is rapidly becoming the standard for Ethernet and iSCSI storage interfaces as prices fall and switch manufacturers build it natively into their ports.

The newer Fibre Channel over Ethernet (FCoE) standard and the decade-old iSCSI protocols pose a tough choice for SANs. IT shops will find applications for both, depending on the distances involved and the tolerance for data loss and latency.

The trend in storage is toward the use of unified fabric to interconnect resources in a data center, and both technologies work in this application. InfiniBand is found in large data centers and also in high-performance or scientific computing.

Regardless of the storage protocol deployed, manufacturers regularly introduce performance enhancements to their wares. Perhaps more important, products surrounding or peripheral to virtual machine farms are becoming hypervisor-aware, giving CIOs the potential to build end-to-end virtual environments that include storage. ■

Server Virtualization

Increasing server efficiency and workload capacity — at the same time

Virtualization enables an organization to more efficiently use its servers and space, processing the same amount of work with fewer machines, less electricity and lower capital costs. Virtualization also facilitates the growth of applications or workloads without a one-for-one corresponding increase in the number of servers.

The old practice of having one application per server certainly worked in its time, but many organizations have found that server proliferation has resulted in poor utilization rates and added costs for underused hardware, not to mention costs incurred by the infrastructure and staff to support it.

How Server Virtualization Works

Abstracting the elements of a physical machine, including the processor, memory, storage and network connection, results in a virtual machine (VM). This is accomplished through server virtualization software, which is available from a variety of manufacturers.

A single physical server can run multiple VMs, each capable of running an operating system plus one or more applications. A control layer, known as the hypervisor, manages the VMs and allocates physical resources according to the demands on them. The hypervisor can move VMs among physical machines on the fly if necessary, unnoticed by users.

Virtualization and virtual instances have existed since mainframe computers first entered the mainstream. Today, the virtualization industry focuses on x86 server architectures.

Server Virtualization Benefits

Much of the benefit of virtualizing servers derives from simply having fewer servers perform a given amount of work, which yields many residual gains.

Smaller physical footprint for servers: Consuming less space can reduce an organization's real estate costs or free up that space to be used for other purposes. And the less floor space a data center consumes, the fewer utilities it requires — heating, air-conditioning, electricity. Many IT leaders raise the question, "If we are reducing office space by deploying mobile solutions to growing numbers of staff, why should we keep expanding the space devoted to data centers?"

Server Virtualization and COOP

Virtualization can be an enabler of an organization's continuity of operations (COOP) plan. But the "plan" is the most important step. A sound COOP plan should include a prioritizing component for backup and recovery for each application, or at least each class of application.

Questions to ask might be: Is the application used throughout the organization or just by one group? If only by one group, how critical is it to the overall organization? Not all applications and data need to have the recovery point objective (RPO) and recovery time objective (RTO) of the most critical ones.

While COOP planning is basic to any IT environment, virtualization brings considerations of its own when it comes to backup and recovery.

The advantage of virtualized servers is that the entire backup process can be much faster than is possible with physical, one-to-one environments. But that requires careful planning, because without the right product mix you can end up with unacceptably long backup times relative to a physical environment.

The key is to update the backup management software to a virtualization-specific version with agents than can "see" into the internal complexity of a multi-VM physical server.

VM-tailored backup software should be able to perform backups (and subsequent rebuilds) from backup images, and not just off the physical server itself. And in the production environment, the latest packages can see into a VM above the operating system level to detect when a particular application has failed. It's also important for the failover mechanism to be configurable such that the cause of failure itself isn't part of a newly provisioned VM.

When setting up the technical strategy to carry out a COOP plan for virtualized systems, there are several choices: Back up incrementally? Back up only image snapshots? Equip each VM with a backup agent and then follow traditional routes?

Each approach brings cost and speed trade-offs, as well as differing degrees of granularity in what is backed up. Regardless of which one is chosen, it is wise to test it thoroughly before relying on it for the production situation.

Lower hardware costs: Virtualization improves machine utilization rates, meaning IT shops need fewer servers. When fewer servers are operating, cost savings begin to appear in other areas, such as operating expenses (OPEX), staffing costs and consulting service costs.

And because utilization rates for virtual server hosts are higher (up to 85 percent versus an average of less than 25 percent for regular application servers), older hardware that's no longer needed in the data center can still be put to use. For example, after a primary data center refresh, older machines can be repurposed for the backup site, or to host virtual clients in a remote office, further reducing hardware expenditures.

More efficient IT administration: Once the virtualization environment is set up, administrators can quickly create and provision virtual machines through a web browser from almost anywhere on the network, which speeds up application deployment.

An example to consider: Suppose a small organization reduces 14 servers and direct-attached storage down to just two physical machines via iSCSI interface. A quick estimate suggests that power consumption could drop by as much as 30 percent and maintenance time by 50 percent, offering substantial savings in these two cost areas.

Reduced software licensing costs: Initially, licensing costs may rise if a software maker charges on a per-server basis. Fortunately, software asset management (SAM) tools can monitor the number of copies running at a given time — a task that must be automated in an environment where VMs are moved and turned off and on, dynamically and constantly.

Moreover, software vendors are updating their licensing policies in acknowledgement of the virtualization boom. You may find fewer licenses are in actual use in a virtual setup, compared with what you might have had in an application-per-server physical environment.

Improved power and cooling efficiency: A single server running close to 40 virtual machines will run at close to full capacity, using more power than the same machine running at 15 or 20 percent capacity. But it won't exceed the power consumption of 40 or so separate servers, each running a single application. Server consolidation means power use reduction, and that's a necessity these days.

Keep in mind that server racks and blade cabinets concentrate the power being used, creating hot spots that most existing data center cooling systems can't handle. But that doesn't translate into a jump in the cooling bill. By reconfiguring cooling and concentrating it where it's most needed, less total energy is required for removing hot air.

Also, the use of three-phase power is another way to improve power efficiency. This power distribution model uses three wires to carry the same current and has a constantly balanced power load, making it very efficient in its delivery of electricity.

High Availability & Disaster Recovery

Another huge advantage of server virtualization is that this computing model facilitates high availability and disaster recovery. Because applications and operating systems are decoupled from physical machines, they can easily be switched to new hardware as needed.

Whether servers are virtualized or not, an organization's backup and recovery strategy for IT interruptions must support operations objectives. For mission-critical applications, managers want a minimum recovery point objective (RPO) — the time/data distance between the last backup and the failure point. How long it takes to actually get the system back to the RPO is the recovery time. The RPO is, of course, supposed to be as short as possible.

Virtual environments can be set up to continuously replicate themselves to a similar-capacity backup recovery site, or perform backups at intervals depending on the RPO and the application. A full-scale mirrored site can be used to load balance under normal conditions. Or the secondary site can be configured as a development and test environment, but be ready to call upon its processing capacity in a disaster situation.

Keep in mind that continuous backup can potentially slow application performance, and that fast backup doesn't necessarily translate to fast recovery. Also, while hypervisors are storage layer-aware, they don't configure storage backup.

There are a number of backup and data recovery setups for virtual server environments.

Full backup: Each VM is backed up as if it were a physical machine. This is input/output (I/O) intensive and might require more backup software licenses.

Dedupe backup: Data deduplication is performed on the VMs and only the changes are backed up. This method lessens LAN traffic, but at recovery time the deduplicated data must be added back in.

Snapshot backup: Periodic VM snapshots are transferred to a proxy server for offsite backup. This practice requires operator attention to manage the proxy server host and works primarily for Windows VMs.

Careful storage assessment must be a part of disaster recovery planning. High availability in virtual environments requires shared storage. This typically calls for iSCSI, Fibre Channel (FC), Fibre Channel over Ethernet (FCoE) and network file systems (NFS) subsystems.

Best Practices for Managing Virtualized Servers

Once an operational case for virtualization has been developed, attention needs to be given to managing it in a practical sense. Here are some suggestions to manage the virtual environment.

Build in limits for VM sprawl. One driver of virtualization is too many physical servers. However, too many VMs can hinder efficiency as well. Virtualization tools make it easy to create virtual servers at almost no cost. But each VM uses system resources such as network bandwidth, processor cycles and storage.

Plus, each VM constitutes a copy of an operating system. Administrators need to have policies in place for what constitutes a legitimate request for a new VM, so uncontrolled growth doesn't endanger service level agreements (SLAs) and reduce the cost benefits of the virtual environment.

Be consistent in technology purchases. While the server virtualization hypervisor and related products are vendor agnostic, it's a smart move to be consistent in hardware brands and components.

For example, if all of the physical servers have the same amount of RAM, the hypervisor can more quickly move VMs and allocate resources. Consistency in host bus adapter configuration and presentation in storage logical unit numbers (LUNs) can prevent VM restart failures and improve performance.

Allocate VMs running a variety of applications to a given host. This approach will more likely result in even patterns of activity across the host resources, rather than having multiple similar applications all calling on the same resources and creating bottlenecks that degrade performance. This will amplify the network and application optimization measures in place.

Avoid single points of failure. A common mistake is putting multiple production hosts on the same storage subsystem. But storage area networks (SANs) should be linked to back one another up. Redundant network interface cards will keep a host connected should one card fail. Use an N+1 approach — one card per machine, plus a fallback — to the cluster of physical machines running under each hypervisor instance to allow headroom should one host crash.

Keep abreast of security developments for the virtualized world. More research is being focused on security at the hypervisor level of virtualization and thereby the improved protection of virtual machines. Most of the threats to the hypervisor have been largely theoretical. Still, hypervisor code sets are growing as more and more VMs per physical machine go into operation.

Virtualization is a means to an end, not an end in itself. It provides more efficient and flexible computing. But some applications will always need dedicated resources, such as large database systems. Some applications can't be virtualized, including kernel-based device drivers, such as Acrobat.

But otherwise, the decision on whether to virtualize comes down to the resources demanded by an application. Behind a cluster of VMs there is still a server with finite power, and it might be wise to leave highly processor- and bandwidth-intensive applications for virtualizing at some point in the future, when the IT team has gained some experience with its virtualized environment. ■

Client & Application Virtualization

Centralizing computing resources

For many organizations, virtualizing desktops presents the final frontier in gaining better control over costs and management of IT resources. Client and application virtualization is a simple concept: decoupling the operating system or application from the end-user computing device. Essentially, the computing takes place on a server in the data center and presents a desktop or application to the user.

Client Virtualization Options

Virtualization can take different forms, depending on the needs of the organization (see *Client Virtualization Options* diagram, Page 9).

Local (or client-hosted) virtualization: This approach enables a single-user machine to run several virtual machines by separating the operations from the hardware. In effect, this setup is a single-user server.

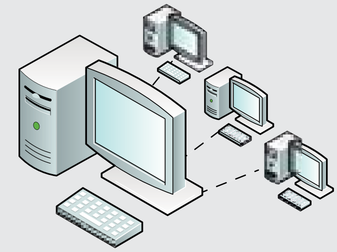
Intelligence, engineering and other technical staff might benefit from having two or more environments running locally on a virtualized client. This methodology doesn't produce the cost reductions that more comprehensive virtualization yields, but it does have some advantages. For example, host-independent virtual machines can be stored on USB drives and thus made portable.

Hosted applications: With this setup, software and user interfaces are shared via terminal protocols. The server typically runs the operating system configured for multiple users, with applications running locally. An alternative setup would be to move the applications to the server, with only presentation occurring at the desktop. This requires applications that can run in a multiuser, server mode.

Virtual desktop infrastructure (VDI): This strategy is full-bore virtualization that closely resembles server virtualization. Each desktop image — consisting of an operating system, applications, data, settings and other particulars — is encapsulated as a virtual machine.

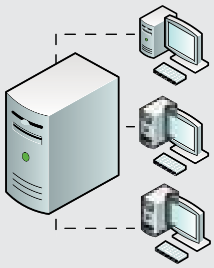
VMs are loaded on servers to the limit permitted by the hypervisor. The VDI can be static, in which the user's VM is complete and simply accessed session by session. Or it can be dynamic, with resources pulled together from other sources when the user logs on. The advantage is that a single image

Client Virtualization Options



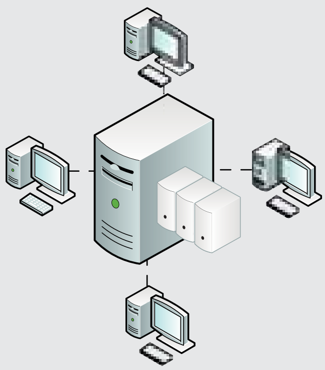
Local (or client-hosted) Virtualization

A local client runs multiple virtual environments.



Hosted Applications

A server hosts an operating system shared by multiple users, and applications run on the local clients.



Virtual Desktop Infrastructure (VDI)

A server hosts multiple desktop images, each a separate virtual machine (VM) on the server.

can serve as many users as needed, thus enabling a smaller footprint of VMs.

With VDI, communications take place over the LAN to the virtual desktops, using vendor-specific protocols layered on top of standard protocols. Manufacturers thereby tune their communications to the hypervisor they support.

PC-over-IP (PCoIP) is a proprietary standard used by several vendors. It compresses and encrypts traffic between VMs and desktop displays. It is useful in call centers, and in CAD, financial and healthcare environments where users need only a display and applications that are common to the workgroup.

Remote users can access applications over the organization's WAN. For branch facilities large enough to have local IT support, the virtual desktop infrastructure can be moved to a local data center or closet, with local storage that is backed up over the WAN. Mobile or small groups of remote workers connect either to a branch facility or to the organization's data center using a virtual private network (VPN).

A current trend is using stateless PCs via OS streaming, in which everything is stored or managed centrally, but sent to a traditional client on login. In some cases, large organizations are moving beyond first-generation, static virtual desktop infrastructures to more flexible, dynamic ones.

Often, such organizations have deployed initial virtualization for workgroups with essentially the same requirements, such as transactional staff in call centers or help desks. Dynamic virtualization is better suited to managerial and knowledge staff who typically use a greater variety of applications and devices.

Benefits of Client & Application Virtualization

Client virtualization benefits are so numerous that it's a wonder that this technology wave didn't occur before server virtualization. Now that many executive, managerial, knowledge and field service or end-user service staff have at least two computing devices, the wave has arrived — and with it a panoply of benefits.

Better security: The threat environment, populated by sophisticated, financially motivated cybercriminals, has worsened appreciably, and there is no end in sight. Equally dramatic is the increase in sensitivity about protecting personally identifiable information, whether held by corporations or government agencies.

No matter how robust the cybersecurity measures an IT shop puts in place, people are the weakest link. Because data is stored on servers rather than end-user computing devices and configurations are locked down, virtual clients can greatly reduce data loss stemming from lost or stolen devices. Virtualization can also enhance security management because it enables faster mass deployment of patches and updates than is possible in traditional PC deployments.

Fast provisioning: New users, users who switch departments or take on new roles, and mobile workers all can be provisioned much more efficiently and quickly when images and computing resources are maintained centrally and rendered device-independent.

With the right authentication credentials, which can be issued on a USB device, users in a virtual client environment are forever freed from a single desktop. Their state, access to organizational resources based on their role, and their work data files are streamed to them wherever they are, on whatever device.

Application availability: Maintaining applications on a virtual machine improves uptime and performance. In the event of file corruption, a VM can be instantly replaced with a "clean" image from backup, rather than having a manual reinstall requiring a cubicle visit by tech support.

Lower hardware and software costs: Because most of the computing takes place on the server, virtualized clients don't need to be replaced as often as full-function PCs. As a result, PCs used to access virtual clients have a longer useful life.

In addition to aiding systems lifecycle management, client or application virtualization can trim software licensing costs. Organizations need to purchase only enough licenses for the average number of simultaneous users, which is generally a lower number than the total number of people who use the program.

Easier administration: Even for 10,000 users or more, an OS upgrade or application rollout can be reduced from a four-month process to a few days. Virtualization eliminates the costs and procedures associated with users moving, because they can log in from anywhere, and their image streams to them. Patches can be installed more quickly as well.

Centralized control: Virtualization can bring a steady proliferation of software versions and desktop images under control. For example, suppose a large organization experiences "image explosion" and ongoing compatibility and upgrade obstacles, with nearly every one of its more than 10,000 machines having a different image with different OS versions and applications.

After virtualizing, the organization is down to 50 images maintained in a virtual desktop center. Upgrades and patches are accomplished monthly for all images instead of quarterly, and with fewer staff hours logged.

Utilizing Thin Clients

Once users' individual inventories of resources are no longer stored and maintained at their desks, organizations can address the issue of computing devices. For many, the answer is a thin client, which is essentially a display for the pixels streamed from their virtual machine in the data center, and a keyboard.

Although stateless and (mostly) lacking storage, a thin client is more than a terminal. Properly specified and configured, a thin client is capable of giving the user a full PC experience in terms of multimedia and response. Thin clients support high-resolution monitors and have sufficient memory to load and display multiple applications.

In other words, there is still a device that resides on the desktop, but it is typically much smaller than a PC (some products incorporate the thin client into the monitor) and requires no visits from the tech staff. Moreover, thin clients require the use of server-based software running atop the virtualization package that is designed specifically for configuring and managing them. It ships free with most thin clients.

Thin client manufacturers use a variety of modular approaches. Some vendors push the local memory to the display, while others keep it within the thin client device. Some incorporate speakers. Most avoid the use of fans, which reduces energy consumption and the attraction of dirt that requires periodic vacuuming by tech support.

Although they simplify user desktop management, rolling out thin clients requires planning because they are available in a variety of configurations and form factors.

Client Virtualization and COOP

If best practices are used for maintenance of service in the event of a data center disaster, it stands to reason that similar availability will exist for users whose virtual clients are hosted in the same data center.

The old model of recovery involved restoring physical machines from backup tapes, a process that could take hours or days. Virtualization tools now enable mirrored sites to synchronize continuously. This provides two benefits. First, the sites act as load-balancing resources for one another. And second, they provide near instant failover should one center become unavailable.

Virtualization also works in favor of continuity in situations where the data center might be perfectly fine but users can't make it to the office, such as during extreme weather. Traditional telework setups don't always give full access to resources as if they were in the office. However, virtualizing client and client applications changes that.

Thin clients are designed to work optimally with specific client virtualization software brands (Citrix, Microsoft or VMware) and with specific operating systems (Windows 7, Linux). They ship with many combinations of software for display, communications, management and virtualization package support.

Several manufacturers offer portable thin clients, incorporating most features of notebook computers except for hard drives. These devices have nonvolatile flash storage in addition to RAM, but they eliminate the recurring security problem of lost notebooks containing hundreds of gigabytes of stored information.

The virtualized images are housed in the data center on servers. Blade servers offer an efficient host environment for virtualized images because they have a smaller footprint than rack-mount servers and can share power supplies and local storage.

Best Practices

Virtualizing clients and applications represents a commitment to a new approach to computing. Deployments require upfront investment in software, training for IT staff and, often, reconfiguration of facilities. All this anticipates a payback through lower capital expenditures, power consumption, maintenance and support. Following a few best practices can make for a smooth deployment.

Plan out user management. Client and application virtualization change the nature of managing users, in such areas as configuration and provisioning, security, and backup and recovery. Consider how the organization will handle these issues before proceeding with a deployment.

Start deployments slowly. With an initial foray into client virtualization, it's best to target users who perform tasks with limited rosters of applications — staff in call centers for end-user or technical support, for example. A single software image can apply to everyone. But don't strip away all aspects of the user's identity. Some degree of user customization, such as desktop wallpaper and application display settings, should also be enabled.

Think through security. Thin clients have USB ports, which raises the threat of data loss via thumb drives. Look for data loss prevention solutions that enable USB read/write privileges to be controlled via the management console. Beyond data I/O, USB ports are useful for wireless peripherals and as receptacles for encryption and access tokens when the system requires two-factor authentication.

Allocate proper storage to virtualized applications. Storage access is a large factor in application responsiveness that users will experience. Consider the number of VMs per disk resource, frequency with which applications read and write, the type of caching used and the mix of solid-state storage in the SAN or local disk subsystem.

Don't forget about mobile computing devices, such as tablets or smartphones. These endpoints are architected from the outset around web applications and should be considered as components in your client virtualization plan.

Review your software licensing arrangements with software makers. License tracking is not a strong aspect of virtualization software suites, but add-on packages can provide visibility into license statistics. ■

Data Center Optimization

Improvements that reverberate throughout the organization

Virtualization elevates the importance of the data center to the organization. Besides being the repository of applications, the data center hosts the computing power and desktop images for all users once virtualization is complete. Full client and server virtualization results in greatly increased control, security, manageability and energy efficiency.

All of this brings fresh challenges to optimizing the data center so that it provides the best possible service levels to applications and users.

Migrating to Blade Servers

The data center has changed a lot in the past decade. A given cubic meter of space can pack far more computing capacity than the same amount of space could 10 years ago. Moore's Law, predicting the continued exponential growth of computing power, is still very much on display in the data center.

Today, the preferred packaging for x86 architecture servers is the blade format. Blade servers typically mount vertically in enclosures of varying sizes, which in turn fit into standard 19-inch racks. With blades, it's possible to double or triple the computing power in a standard rack versus rack-mount servers. Blades reduce the data center footprint and use less power.

Blade servers come in a wide array of prices, capabilities and capacities. At the high end, they can support up to four processors as well as 256 gigabytes of RAM. Onboard storage can be solid state, flash memory drives or spinning disk drives.

More typically, blades incorporate storage controllers to connect through the backplane communications fabric to external storage, usually storage area networks (SANs). Vendors also offer blade-format cards that comprise terabyte-size disk arrays.

Blade servers lack power supplies, so they're powered from either a blade format power supply in the enclosure or from an external power unit serving one or more enclosures.

One of the top challenges in migrating from rack-mounted servers to blade servers is contending with changed patterns of heat generation. Because blades are packed so densely, with dozens of microprocessors

Data Center Optimization and COOP

The data center is a critical element of an organization's continuity of operations (COOP) plan.

Organizations want to ensure that their primary sites are as reliable as possible and have a backup site ready to go. As a bonus in normal circumstances, the backup site can function as a load balancer for both applications and storage.

Data centers rely on communications and power. Power backup operates in tiers. First comes an alternate utility, if it is available. When main power is interrupted, failover goes to battery backup, and then to power from generators. Test power backup periodically, including generator starts.

Where possible, run data lines from two separate carriers to the data center, with physically separated connection points. A nearby hot site can be mirrored continuously with microwave communications.

With one-application-per-server setups, there is a need for one-to-one duplication at the remote backup site, and uncertainty that the backup environment is always identical to the production environment. A mismatched or out-of-phase backup can be only marginally more helpful than having no backup, and will probably leave the organization in a difficult position with respect to legal compliance.

Virtualization encapsulates the OS, the application and the data as a unit functioning through the hypervisor. The entire runtime package can be replicated or mirrored to the hot backup site and simply restarted on an available server.

Although they must be logically similar, with virtualization, backup sites don't have to be physically identical thanks to the intercession of the hypervisor layer. Organizations gain the flexibility to back up offices via two-way active connections, or between the operational data center and a dedicated backup site.

Snapshots of servers and other system components can be stored on the storage area network (SAN) for replication or full restoration. Snapshots and machine images can also be stored remotely to tape.

The COOP is incomplete without assigned roles and responsibilities and a playbook for prioritizing which applications are brought back in sequence. Above all, no matter what the environment, best practice dictates regular testing.

operating within an enclosure encompassing less than a cubic meter, blade setups produce hot spots that existing cooling systems cannot adequately address.

The brute-force approach to this problem, essentially doubling the cooling capacity for the data center, might keep the blades cool, but it will drive up costs and waste electricity.

A better strategy is to keep track of the accumulated power consumption of the infrastructure. It is simple to calculate the wattage used by a given enclosure and therefore the amount of heat that will have to be removed. Also, be sure to factor in the maximum ambient temperature that the technology can tolerate; it tends to be higher with newer processors.

Alterations in distance between racked enclosures can reduce the power needed for cooling, but that might work against the larger strategy of using the least amount of real estate. In any case, it may prove helpful to retain a heating, ventilation and air-conditioning (HVAC) specialist to design a cold aisle/hot aisle setup that ensures adequate airflow to the blade enclosure while preventing hot exhaust from getting pulled back in.

Power is the other consideration in migrating to a virtualized environment housed in blade servers. Although the resulting blade infrastructure might be physically smaller than what it replaces and use less total power, it might also be more power-dense, and therefore might require more power per square foot.

There is no universally accepted way of calculating power per square foot. Some methodologies take into account the total square footage of the data center, while others calculate based on the floor area taken up by the equipment.

However power needs are figured, when planning a change in infrastructure, the design should include an overlay of power and cooling distribution. That way power distribution and cooling apparatuses, together with all associated cabling, are laid out just once and in a way so as not to degrade cooling efficiency.

MAXIMIZING CAPACITY

Implement these basic application acceleration and WAN optimization steps for best use of storage capacity.

File compression: Implement file compression for non-streaming applications such as database updates and file transfers. Header compression can reduce bandwidth use when TCP/IP overhead matches or exceeds the information in the packets.

Object and byte caching: Perform object and byte caching at user sites to reduce network traffic. Object caching stores units of static information that are needed frequently, while byte caching creates local tokens that identify repetitive streams of compressed traffic. Both techniques can speed response times across the WAN. Wide area file services (WAFS) appliances at end locations act as proxy servers for frequently requested files.

Chatty apps: Block or limit chatty applications that may not be mission critical or even authorized, such as Skype and peer-to-peer sharing programs.

Scheduled batch processing: Schedule traffic such as daily constituent tallies or files to be batch processed outside of regular operation hours.

Data dedupe: Deploy data deduplication to eliminate redundant bits of data traveling over the network.

Common Internet Files System (CIFS) acceleration: Implement CIFS acceleration to speed up activity on the network.

Upgrading Network Capacity?

An organization's WAN needs to have enough capacity to carry virtualized desktop and application traffic from the data center to distant users. Before adding network capacity, however, apply technology to get the most out of the bandwidth the organization is already paying for, such as those mentioned in the "Maximizing Capacity" sidebar above.

A recent study found a slowdown in the rate at which organizations were adopting OC-12 and bonded OC-3 lines. That's because merely adding bandwidth doesn't necessarily solve application performance over a WAN. This trend reiterates the value that organizations stand to gain from WAN and application techniques that help optimize communication between remote locations.

Load Balancing

Load balancing is a technique for keeping server response times within acceptable parameters. It ensures the most efficient use of all the machines in a cluster. Loads vary according to applications and user groups. Within an application, loads can vary throughout the day, the month or the year.

When applications are independent of physical hardware, virtual machines can be subject to load balancers that monitor response times and dynamically allocate transactions to the appropriate hardware that has the least load. Network load balancers receive pings from the servers they are interacting with and assume a machine is at its capacity when it stops signaling its ability to process requests.

Some load balancers include application acceleration features, such as shifting secure sockets layer (SSL) processing to dedicated hardware and issuing redirects among secure and nonsecure servers.

Boosting Power & Cooling Efficiencies

The efficiency of power and cooling strategies can be measured with a metric called power usage effectiveness (PUE), the ratio of total power into the data center to power used by IT. A PUE below 2 is considered highly effective; and a PUE above 3 suggests that there are bigger problems than the IT gear.

In too many data centers, cooling strategies haven't fully kept pace with the microprocessor revolution. The standard approach to cooling — low ambient temperature; raised, perforated flooring; overhead venting — is inefficient and possibly inadequate for high-density blade environments.

By concentrating resources in smaller physical spaces, virtualized servers and storage devices tend to have high utilization rates. The resulting hot spots call for more current cooling techniques.

The arrangement of equipment can enable more efficient cooling by creating hot and cold aisles. The hot aisles are isolated by tenting materials and vented, while fans pull in air from the cold aisles.

Another technique is the installation of chimney-like ducts at the ends of racks where the head is. One organization designed this method into its new data center housing 4,700 blade servers in racks, each consuming 30 kilowatts and weighing one ton. This example shows that there are workable, large-scale alternative cooling options available.

With either architecture, the goal is to get rid of the hot air before it can mix with the cold. Both of these approaches also mean raised flooring can be removed, at least for cooling purposes.

Reducing Power Use

On the power front, virtualization and consolidation are the principal strategies IT shops are using to reduce overall consumption. With recognition that the data center is a high power user, organizations are trying new and novel ways to be more efficient. Some are channeling data center heat to help warm offices elsewhere in the building.

Other organizations are converting excess heat into electricity. A few of these heat recovery strategies are aimed at power generation to run power supplies or networking equipment. Servers equipped with thermoelectric devices, which convert heat into electricity sufficient to operate fans or charge batteries, are in the experimental stage.

Still other organizations are performing air-to-heat exchange. This is done by utilizing devices such as heat wheels, which are large, slow-turning devices connected to a duct system that exchanges cooler outside air with the data center's heated air. This approach to cooling consumes less than a quarter of the electricity needed for traditional cooling systems.

At nearly 20 feet across, one heat wheel can theoretically remove 850 kilowatts of heat energy. But, together with massive ductwork, the wheels require a great deal of space and specialized installation. In addition, they don't work in all weather conditions. This example highlights one of the many alternatives available for reducing power use.

Best Practices for Data Center Optimization

To get the greatest return on investment from the data center, organizations need to look holistically at the equipment it houses. Because virtualization reduces the number of physical machines used, there's an opportunity for a concomitant reduction in real estate and electricity requirements. Consider applying these common best practices.

Virtualize and install blades. Combine virtualization with server consolidation and switch to blade servers. Blades can save up to 25 percent in capital outlays relative to other server form factors.

Redesign the data center. Re-envision the data center before virtualizing. Rather than merely reconfiguring the racks, redesign the whole space. At a minimum, a new server infrastructure likely calls for an overhaul of the cooling and power subsystems. Use the opportunity to reduce square footage and install greener lighting and peripherals.

Evaluate software licensing. Virtualization can cause duplicative copies of applications even though usage is the same. Many software manufacturers are altering their licensing terms to account for virtualization. ■

Storage Virtualization

Better management of data retention and protection

Storage is sometimes the last component of IT that organizations apply virtualization or consolidation to. Perhaps that's because data is often the fuel that powers the organization, and if data is lost, corrupted or unavailable, everything comes to a halt. So there is some justifiable reticence to make changes that could affect an organization's data.

The principal goal of storage virtualization is maintaining fast availability of data to applications, regardless of physical location. This goal is pursued through a number of strategies.

CONSOLIDATED STORAGE STRATEGIES

- Minimize occurrences of identical data.
- Utilize tiering methodologies to apply the least expensive media consistent with service level requirements.
- Ensure that virtual machines don't lose their connectivity to associated data, even when the VMs are moved among physical servers.
- Apply thin storage provisioning to users and applications to strike a balance between wasting storage and crimping performance.
- Manage protocols for speed and interoperability.
- Maintain high availability and ensure consistent, secure backup and recovery.

Data Storage Today

Storage has in one sense been virtualized since the advent of redundant arrays of independent disks (RAID). RAID stripes data across multiple disks, speeding access and providing a measure of backup and fault tolerance, depending on the RAID level.

Today, storage virtualization has advanced shared storage to the point where local storage at the user level isn't needed. Depending on roles and rights, users may have access to data and storage of their own anywhere within the organization. All the while, the organization gains the security benefits of not having confidential information stored in devices that can be lost or stolen.

These gains are predicated by sound network and storage protocol design that alleviates unacceptably slow response times for users. And they are sometimes accompanied by user device configurations that bar organizational data insertion via local disks or USB flash drives.

Virtual storage architectures can result in the use of media separate from their associated protocols. Backup tape libraries have mature protocols, but tape itself is a slow medium. So, many IT shops are switching to virtual tape libraries — disk subsystems made up of relatively inexpensive serial ATA (SATA) disks operating under tape protocols.

At the other end of the data hierarchy, for data that requires high availability, solid-state disks are gaining favor. Based on direct access semiconductor memory, they are relatively expensive per unit of storage. But used judiciously, they eliminate disk-access latency in critical applications.

Deduplication

Just as a given amount of bandwidth can be used efficiently by application and network optimization, organizations can get the most from a given amount of storage by reducing data redundancy through deduplication. Redundancy blossoms for a variety of reasons.

For example, multiple users typically save a copy of a PowerPoint presentation or some other file to multiple drives. Successive backups can also cause redundant copies. If virtualization presents a single logical face from multiple physical storage locations, it makes no sense for duplicative data to drive constant investment in more storage.

Deduplication technology works by sampling data blocks and using predictive algorithms to discover likely multiples. It then substitutes a pointer to a single physical instance of the file or, in the case of a database, the block. (Of course, even after deduplication, good continuity of operations practice dictates one other copy in backup on another physical server.)

When combined with compression, "delta differencing" (saving only changed bytes) yields benefits throughout the IT ecosystem, especially during backup. Storage requirements are met and many other benefits are derived: reduced cooling, electrical and real estate costs; faster backups and recoveries; and narrower recovery point objectives (RPOs).

Thin Provisioning

Thin clients without local storage are often paired with virtualization. Users still need personal storage allocation, but because it is coming from the network, it can now be thin.

In essence, while a client (and a virtual server as well, for that matter) can "see" a large amount of storage available to it, the provisioning software

Storage Optimization and COOP

When virtualizing server environments, organizations have a great opportunity to improve system and application data protection and the performance of backing up data, according to Forrester Research.

Continuity of operations (COOP) planning and disaster recovery (DR) are holistic activities in that they involve the entire network and not merely storage. But there are storage-related steps that can be taken to ensure that this vital component is ready for whatever might happen.

Mirrored data storage is a basic building block for DR and should be part of any COOP strategy. But not all mirroring is equal. Different setups are available depending on the distance to the hot site.

Full, synchronous real-time mirroring can be suitable for geographically close sites with high bandwidth. Where there is less bandwidth, such as between more-distant sites, asynchronous backup is a less costly option in which data changes are transmitted periodically.

A well-designed backup and recovery plan also enables active-active data center mirroring. This allows for multiple uses of a recovery site — such as development and testing, for example — when there's not a disaster.

Disaster recovery becomes more efficient when combining virtualization with storage-based replication of data and images. This gives the IT shop the ability to rapidly restart virtual machines from the recovery site.

As virtualization technologies mature, one flaw surfacing is that fast recovery of a VM in its crashworthy state can lead to another crash. Now manufacturers are introducing technologies to, in effect, back up earlier images of a VM, but at some cost of failover speed.

However, some potential disruptions, such as inclement weather, are often known in advance. So IT can move clusters of VMs to hot sites beforehand.

allocates physical storage stingily, a block at a time, thereby increasing utilization of storage. Under traditional, "thick" allocation systems, a given disk might be only 10 percent utilized, whereas thin provisioning can sometimes result in a far more efficient 80 percent storage utilization.

Thin provisioning can lead to occasional oversubscription. Technology for virtual storage application coming to market soon will either use available backup storage or storage not in use by another VM to address this problem.

When planning a virtualization deployment, keep in mind that storage must keep up with the movement of VMs among servers with regard to availability. This favors shared storage resources as opposed to server-attached resources.

Clustering VMs in fewer physical servers can hamper backup and recovery performance. That's where data deduplication and thin provisioning come in, and why thin provisioning, an older concept, is gaining traction in the virtualization era.

Beyond the importance of access, organizations will want to consider data protection, failover times and disaster recovery when planning out a storage virtualization strategy. Keep in mind that when the virtual machines are relocated, they must interact with data storage, including snapshots and backup. Replication should take place at the storage array level.

Tiered Storage

Another older form of virtualization is tiered storage. This approach moves data in and out of higher- and lower-cost storage media, depending on demand patterns. The tiers are RAM, solid-state disk, disk, tape and optical, although in practice most organizations keep three tiers online in addition to memory accessed directly by CPUs. Data coming out of tape or optical, or a third or second layer of disk, can be held in cache temporarily.

Manual allocation of data among tiers is too labor-intensive for virtual environments, so IT managers should consider automated tools for this task. The trend now is dynamic allocation of Tier 1 storage for mission-critical or high-demand applications.

Manufacturers are developing software that monitors storage performance and demand and that will automatically allocate storage among the tiers and move data sets to their appropriate tier. Benefits include investment cost avoidance, faster seek times, less disk I/O and lower power consumption.

But allocation tools require careful tuning to maintain response requirements while avoiding overinvestment in the highest-performing tier. Keep in mind that often 80 percent of data stored in Tier 1 environments has not been accessed in the past 90 days, according to industry estimates.

Tiering can also take place within a storage array if service level agreements (SLAs) call for it, rather than among physically separate

DATA AVAILABILITY — to the Nines

Data availability is often discussed in terms of nines: four nines, five nines, six nines. The nines refer to the amount of uptime or availability as in 99.999% (five nines). The greater the number of nines, the less allowable downtime, and the greater the data availability.

arrays. These arrays use a mix of disks that may contain some solid-state flash memory, cabinets that integrate flash, Fibre Channel or iSCSI, and SATA drives. Such integration may include virtual tape libraries.

Flash drives are useful not only in scientific and other classic high performance computing, but also in busy database and online transaction processing.

Storage Area Networks

When considering virtualizing storage, it might be time to update the communication protocols used to access storage as well. Emerging trends may favor 10 gigabit iSCSI and Fibre Channel over Ethernet (FCoE) as opposed to traditional iSCSI, Fibre Channel (FC) or Network File System (NFS).

Network attached storage (NAS) and storage area networks (SANs) are both forms of storage attached to the network. The distinction is that NAS systems are oriented toward file sharing, while SANs are data-block based. NAS systems generally run over TCP/IP over Ethernet using a variety of protocols (NFS, CIFS or HTTP), while SANs — the faster systems — use FC and now FCoE and the SCSI and iSCSI protocols for 10 Gig-E environments.

Because SANs don't incorporate the file and operating system overhead required of NAS, they are generally more oriented toward remote access and abstraction of logical storage from physical (an important quality for virtual environments). That's largely the definition of virtualization.

Fibre Channel storage is more expensive than NAS, but that's mitigated partially by the FCoE protocol that lets SANs connect via Ethernet cabling, doing away with the fiber. Just as virtual tape libraries preserve existing protocols over different media, FCoE encapsulates the Fibre Channel protocols over long distance via Ethernet and enables 10 Gig-E links within virtualized environments.

Storage area networks also are well suited to iSCSI, all the more so in virtualized environments thanks to its TCP/IP core. A bonus is that this allows remote booting using a Dynamic Host Configuration Protocol (DHCP) server.

Maintaining High Availability

The fundamental step for high availability is assured and regular backup of all data to a second site. In virtual setups, both sites can be used for daily access provided that mirroring is continuous.

Virtual systems software can create snapshots of applications and storage at given intervals. How far apart depends on the recovery time objective (RTO) and RPO. Both are application-specific. They take into account how granular a recovery needs to be, access patterns, distances involved and how much data changes over time.

PRIMARY THREATS TO HIGH AVAILABILITY

- Deletion, corruption or exfiltration of data
- Hardware failures, including interfaces and SAN switches, not simply storage units themselves
- COOP events including loss of power

Storage Optimization Best Practices

To get the most from the organization's storage investment, consider these best practice tips.

Don't over-provision storage. With an inventory of applications and storage usage patterns, coupled with thin provisioning, organizations can avoid having a lot of storage with low utilization rates.

Plan and allocate storage depending on the application. StorageIO, the tech analyst and consulting firm, points out that for applications relying on data retention, cost per gigabyte is the overriding criterion. But for I/O-intensive, high-count transactional applications, performance is paramount.

Make use of deduplication. Back up virtual servers and physical environments to a consolidated storage system where deduplication can reduce the storage footprint. When time or bandwidth (or both) are constrained, host-based or source deduplication can reduce those requirements, and reduce storage needs as a matter of course. ■

Network Optimization

Achieving efficiencies and improving management

All of the work that IT departments do to make the most efficient data centers and user platforms can fall short if the network is operating subpar. A number of bottleneck issues can degrade network performance, and virtualized environments can make these problems more difficult to manage.

Network complexity is increasing. This is especially true of organizations with multiple locations. That might sound counterintuitive given the trend toward unified communications in which more forms of traffic are moving to a single network. But those unified networks are branching out to even more physical locations, and content is expanded by the *Internet of things*.

The term *Internet of things* refers to the linking of billions (maybe trillions) of devices, with many, if not all, utilizing Internet Protocol version 6 (IPv6) to connect to the Internet. IP audio and video signals from surveillance devices or data generated by sensors in industrial or scientific applications are examples of this growing complexity, which add to the demands on an organization's network.

Designing an Efficient Network

An efficient network is one that supports the requirements of the organization by delivering applications and data. This implies a careful planning effort, taking into account several key factors.

One key factor is determining the required response times for each application. In transactional or call center environments, often users need LAN-type responses over the WAN. By contrast, more traditional batch data processing can typically be served by setting a lower priority for its traffic.

Other key factors are figuring out what the goals are for data center consolidation, the use of cloud computing services and virtualization and continuity of operations. These factors will all affect network design and layout.

An efficient network is also one that is manageable. Fundamentally, to be efficient, a network has to be running at its maximum practical speed.

Organizationwide networks are really amalgamations of numerous wired and wireless LANs, storage networks and WANs. And servers containing multiple virtual machines pose their own management challenges. Managing a complex network requires visibility into the details so that administrators can pinpoint and eliminate performance vulnerabilities.

If designing an efficient network is a matter of wiring to avoid bottlenecks, than a well-managed network allows administrators to anticipate or eliminate sources of traffic growth, configuration errors and security problems.

Network Management Software

With a good management strategy in place, operators of organizationwide networks have reported 90 percent reductions in staff time to repair problems and a high-nines percentage of uptime.

NETWORK MANAGEMENT TOOL APPLICATIONS

Here are some areas where network management tools can help.

Compliance: Financial compliance, such as the Visa Cardholder Information Security Program (CISP) and Sarbanes-Oxley (SOX), call for audit trails of network configuration changes. Other federal regulations include those related to the Health Insurance Portability and Accountability Act (HIPAA), Gramm-Leach-Bliley Financial Services Modernization Act (GLBA) and still-emerging financial regulatory overhaul provisions.

Fault management: Gain visibility into detection and reporting, physical devices and virtual devices; zero in on problem locations and causes; reduce mean time to recovery; determine if a network or server is slow; put fault and performance information in a single console; and reduce multiple point products and the reports they generate.

Configuration and systems administration: These tasks include Quality of Service (QoS) settings, changes in access control lists and device logins. They also ensure the ability to monitor application performance, including networking routing, to more quickly free up bottlenecks and points of failure. Manual performance of routine tasks is prone to errors and deviations from organizational standards, such as the granting of role-based permissions.

Both private and public sector organizations face growing compliance requirements, thanks in part to several recent pieces of legislation covering everything from financial services to oversight of grantee performance. A visible and reliable network will help reduce compliance costs.

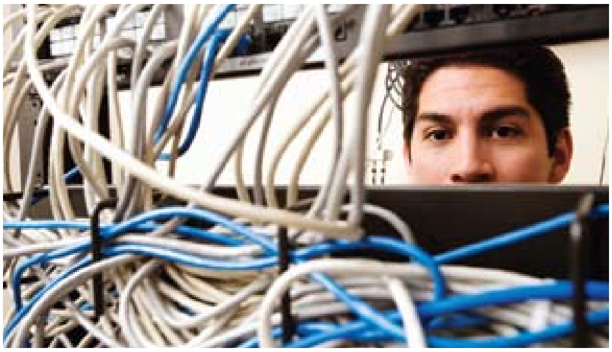
Moreover, networks are by necessity a never-ending work in progress. New technology, applications and operations requirements mean that the IT shop is always tweaking the infrastructure. With effective management tools in place, routine tasks such as provisioning users, installing patches and updates, and adding or subtracting assets (such as wireless LANs or servers) become less time- and labor-intensive.

As noted, networks are only as good as an organization's ability to manage them. The administrator's challenge in managing complex networks is simplifying and unifying the reporting of network events occurring at servers, peripherals and endpoints, as well as at network devices such as switches, routers and bridges.

A huge variety of tools are available for network and application management. Some use Simple Network Management Protocol (SNMP) information from infrastructure devices. Others use proprietary communications with applications to report on performance anomalies or whether versions and patches are current. Still others are tuned to cyber-related issues, searching for patterns in external access.

These tools all produce extensive logs that themselves present a management challenge in teasing out useful information. More recently, an emerging class of tools essentially lumps all network data — cybersecurity and application and network performance data — into a single repository that can be searched to create reports depending on what the user is looking for.

Whether by scripted search-and-report function or traditional log examination, network monitoring and response mechanisms should be integrated and automated as much as possible. Management automation has been shown to substantially reduce the number of network incidents that rise to a so-called Level 2 response — an emergency level — from 75 percent to 1 percent, mainly because of faster identification of problems.



Optimizing the WAN

Several trends are driving increased WAN traffic, and therefore the need for WAN optimization. The first is virtualization itself, which tends to reduce the amount of local storage and the number of application servers.

Other important drivers are organizations reducing their number of data centers; web application deployment both within the organization and outside to partners and constituencies; increased mobility enabled by wider telework acceptance and more robust end-user devices; more sources of WAN traffic including voice and video applications, and IP surveillance.

The purpose of WAN optimization is to enable fast application performance across distances, effectively mimicking LAN performance. Not all traffic behaves in the same way, and optimization takes that into account.

The WAN can become application-aware and automatically apply Quality of Service (QoS) policies and manage port assignments with the right tools. This implies that applications must also be tuned to behave well over the WAN by using compression, data caching off-WAN and assigning application-specific traffic priority.

Overall, it is possible to reduce WAN bandwidth utilization by as much as 60 percent. In one study of organizations using WAN optimization, the tech research firm IDC found an improvement in restore time of 83 percent and a reduction in help desk calls of 17 percent.

Application Networks

In a nonvirtualized approach, each application gets its own server. That approach is very costly — and given today's virtualization capabilities, not necessary. Meanwhile, applications themselves are driving growth in network and traffic. That in turn is driving deployment of application networks, also called application delivery networks.

An application network is the sum of several technologies all aimed at maintaining application performance, meeting SLAs, avoiding downtime, maintaining security and keeping costs in check.

An application delivery controller (ADC) appliance offloads server tasks that are computationally intensive but not strictly components of the application. These include Secure Sockets Layer (SSL) termination and compression. Other ADC capabilities include content caching, TCP optimization and QoS policy execution that is media-aware (wireless, broadband or cellular).

KEY WAN OPTIMIZATION TECHNIQUES

Compression and encryption of data, combined with prepositioning some data: An example of prepositioning is Domain Name System (DNS) caching, or holding query data for a period of time in client machines, provided they are configured for it. Another related technique is server caching.

Data deduplication: This technique can open up WAN bandwidth (see Chapter 5).

TCP flow optimization: This process dedicates links to chatty traffic, such as Voice over Internet Protocol (VoIP) or transactional applications and one-way streaming such as video distribution, so that both links perform more efficiently. Keep in mind that QoS protects some traffic from latency or dropped packets but doesn't provide more bandwidth.

Application delivery networks typically operate at Layers 4 to 7 of the Open Systems Interconnection (OSI) model. This is where policies and granular application intelligence reside, and the levels at which an ADC can identify the application generating a packet and then apply the proper policy, including routing the most available pathway.

An ADC-powered network improves load balancing, enhances network efficiency and ensures application performance.

Optimized Storage/ Network Interaction

Data growth affects network performance and costs. If storage access and retrieval is slow across the WAN, everything from the online user experience to backup and recovery time objectives are affected.

There's never a data or information recession, regardless of economic conditions. Organizations typically experience 25 percent growth in data every year.

For optimal data traffic in and out of storage, the key is minimizing traffic. One way to mitigate the overhead is data deduplication, reducing the dozens or hundred of copies of the same material stored indefinitely. As mentioned in Chapter 5, dedupe is simple in concept — only unique data is stored, along with small pointers in various directories taking users to the single stored copy.

Once data redundancy has been reduced, storage virtualization can extend to the application protocols for communication with storage subsystems.

Many organizations are looking at Fibre Channel over Ethernet (FCoE), which effectively encapsulates the Fibre Channel (FC) protocols so they can share Ethernet cabling in the data center. FCoE enables a single, unified data center fabric, simplifying cabling and other physical requirements. It is not suitable for routing because it operates on the network protocol stack. But FCoE is gaining acceptance as 10 Gigabit Ethernet gains primacy.

The other major TCP/IP storage protocol is iSCSI. Because it operates above TCP/IP on the stack, iSCSI can be applied across switched networks and is the standard protocol for storage access across the WAN.

Look for applications unbundled from and operating independently of storage subsystems to maintain operating system independence as well as storage architecture flexibility. ■

NETWORK OPTIMIZATION BEST PRACTICES

Get organized. Inventory and categorize applications by traffic, so the service requirements are fully understood. Carefully plan out future application deployments from a network optimization perspective.

Survey network nodes. Prepare a detailed map of the major network nodes, the initiators of traffic calls. This survey should include the topology of how they connect.

Automate administrative functions. This will minimize faults and unauthorized access. This step should be integrated with performance tools so that performance-related redirects don't run ahead of the ability to detect problems.

Maximize bandwidth usage. Consider application acceleration technologies that operate over the whole network protocol stack to utilize the existing bandwidth as much as possible.

Reduce WAN traffic. Consider implementing data caching and related services via appliances at local nodes such as branch offices, which will help lower the amount of traffic on the WAN.

Schedule VM backups. Rather than trying to back up an entire suite of applications and virtual machines at once, move them in groups to proxy servers and back those up serially, thus minimizing required bandwidth.

Maintaining Storage PACE

To leverage the most value out of a tiered storage arrangement, an organization needs to determine the level of service it wants to maintain. Four important service points to consider are performance, availability, capacity and energy (PACE). Budget is another obvious consideration when planning out an effective storage tiering strategy.

Glossary

This glossary serves as a quick reference to some of the essential terms touched on in this guide. Please note that acronyms are commonly used in the IT field and that variations exist.

Blade

A blade is a server on a circuit board that shares storage and power with other blades in the same mounting cabinet. Blades are typically half the size of a rack-mounted appliance.

Caching

This term refers to the storage of repeatedly used data sets at either end of the LAN or WAN to reduce network traffic. Caches can contain bytes in a standard-size block, by file or by a software object that might comprise more than one file.

Compression

This technique reduces file size by encoding it according to a specific algorithm, using a subset of the file's bits, such that the file can be decoded and restored to its readable form.

Deduplication

This technique eliminates redundant copies of data stored on the network by replacing most copies with pointers to the data. The objective is to reduce storage needs and speed performance.

Encapsulation

This technique renders separate elements of a computer or software program as a single entity, such that it is decoupled from the underlying hardware. This is a process within virtualization.

Encryption

Encryption applies an algorithm to plainly readable information so as to render it unreadable. Encryption is a method for securely transporting data.

Ethernet

Ethernet is the near-universal standard for data movement over a LAN. It operates at the physical, media access control and data link layers of the Open System Interconnection networking model.

Fibre Channel (FC) & Fibre Channel over Ethernet (FCoE)

Fibre Channel is a protocol for moving data to and from storage networks at high speeds. It is used mainly for storage area networks (SANs) that store data in blocks as opposed to files. FCoE encapsulates FC frames so they can travel over Ethernet networks.

Flash memory & flash drive

Flash memory is nonvolatile semiconductor memory that is programmed and erased electronically. A flash drive is a unit of such memory formatted to act as a disk drive.

Hypervisor

A hypervisor is a software program used to enable virtualization by allowing multiple operating systems and applications to run simultaneously on a single host computer.

InfiniBand

This copyrighted term refers to an architecture to interconnect computing resources in high-performance networks.

Latency

Latency refers to the amount of time between when a query is made to a data system and the response is received by the initiator (a user or another computer).

Load balancing
Load balancing ensures efficient utilization of resources by dividing workloads among two or more computers, storage subsystems or network links to avoid bottlenecks.

Network-attached storage (NAS)
NAS is a storage technology that holds data in file formats, as opposed to blocks. NAS communicates with applications via NFS, CIFS or Hypertext Transfer Protocol (HTTP).

Optimization
Optimization covers the steps taken to make sure components in a networked computer system operate at the highest possible level of performance. For example, network optimization might require procedures to minimize traffic, or to get like traffic together on one link.

Power usage effectiveness (PUE)
PUE is a measurement of how efficiently a data center uses electricity. It is the ratio of the power coming in (as measured at the utility meter) to the power available to the data equipment (what remains after cooling, power backup and lighting). A ratio of 2.5 is average; a highly efficient data center would have a PUE of 1.5 or 1.6.

Quality of Service (QoS)
QoS is a set of techniques for managing data traffic on a network according to the application supplying or using the data. Traffic that is highly sensitive to latency, such as VoIP or video, is assigned a higher priority than less sensitive data, such as writes to a database.

Recovery point objective (RPO)
RPO is the maximum acceptable data loss in a system interruption or failure, expressed as time since the most recent replication.

Recovery time objective (RTO)
RTO is the maximum time it takes to restart applications and restore service after a failure. It is part of the organization's operations continuity plan.

Redundant Array of Independent Disks (RAID)
RAID is a method for increasing disk drive fault tolerance and performance by writing data to two or more disks. It allows read and write cycles to occur in an alternating or interleaved fashion, speeding up overall performance.

Service level agreement (SLA)
An SLA is the minimum performance of a system or service acceptable to the providing and using parties. SLAs can apply to system uptime, response, restoration time after failure, security and many other metrics.

Snapshot
A snapshot is a copy of a complete virtual machine made at a given point in time. When the snapshot is recoded, it prevents writes to the VM disk file and becomes the VM to which reads and writes take place.

Storage area network (SAN)
A SAN is an architecture of interconnected disk drives or other data storage devices that record data by blocks rather than files. Typically, host communication is done using the SCSI (or iSCSI) or Fibre Channel (or Fibre Channel over Ethernet) protocols.

Thin client
A thin client is a desktop device that provides the display component of a computer system or network and taps centralized storage and processing. As display devices for virtual machines, they require little maintenance.

Thin provisioning
This term refers to the allocation of storage on a dynamic, as-needed basis to virtual machines. A technique to optimize storage utilization, thin provisioning contrasts with the setting aside of large blocks of storage space that may go unused by a server or client.

Tiered storage
This approach for optimizing data storage costs and performance allocates data to the least expensive medium consistent with application requirements for availability. Tiered storage consists of a fast first layer, a middle layer of disk storage for data likely to be used within a specified time, and longer-term tape or optical media for rarely used or archived data.

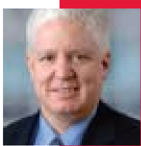
Virtualization
Virtualization is a technique for enabling multiple instances of operation systems and applications to run on a single physical host. In practice, virtualization is an approach to building computing facilities that minimize power consumption and capital expenditures, while ensuring high availability and ease of maintenance and operation.

Disclaimer

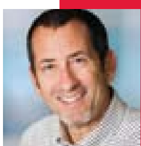
Index

About the CONTRIBUTORS

NATHAN COUTINHO is a Solutions Manager for CDW, with a focus on virtualization. He has more than 11 years of experience in IT, covering various roles in management, technical sales and consulting. His current responsibilities include evaluating and educating clients about trends and directions in the server, client and storage virtualization spaces.



PAUL SCHAAPMAN is a Solutions Architect for CDW. With more than three decades of experience in IT infrastructure, he has a strong background in virtualization (server and client), server and storage engineering, IT architecture, and IT consulting. Paul was awarded VMware's Virtual Vanguard Award in 2007 for his work on a large virtual infrastructure for the Virginia Farm Bureau.



TOM TEMIN is a Potomac, Maryland-based writer, editor and broadcaster with 25 years of experience covering high technology and its application to business.

LOOK INSIDE
for more information on:

- Choosing the right client virtualization solution
- Managing an efficient network
- Optimizing data center performance
- Maintaining storage high availability

CDW·G

800.808.4239
CDWG.com/virtualizationguide

110118
85334AB