Government organizations are working hard to do more with less, consolidating data centers and striving to deliver more services online to help lower operating costs and comply with federal regulatory mandates such as the Office of Management and Budget’s 25-Point Plan to Reform Federal IT.

A big part of the ongoing consolidation efforts includes virtualization, putting more applications on fewer servers to save space and reduce complexity in government data centers. Trouble is, virtualization implementation efforts in many organizations have slowed, partly due to the difficulties that arise in the management and control of rampant storage growth.

Recent analysis from the Gartner Group reports most organizations both in the public and private sectors are still only roughly 30% virtualized, meaning less than half of all X86 server workloads have been migrated to virtual operations. And in the aftermath of Federal CIO Vivek Kundra’s June departure, news reports indicate data center consolidation efforts may be slowing as the pace of closures (previously expected by the OMB to run at about 12 per month) are becoming difficult for agencies to sustain.

Although virtualization has been around for nearly a decade, its penetration in government may be slowing for a number of reasons. First, government organizations generate mind-boggling amounts of data. Simultaneously, numerous regulations require data remain safely stored for years, and sometimes even decades. Due to continuous belt-tightening, storage budgets have faced extreme pressure, impacting costly capital investments in expensive disk arrays. Operating expenses are also under scrutiny, as electricity costs rise and data centers face limits in floor space and power capacity.

All of this leads to a strong requirement to investigate alternatives, including updating agency networks and investing in advanced storage virtualization tools and techniques, such as tiering and deduplication, to keep storage costs down and ensure virtualization efforts continue to move forward. Jon Toigo, CEO of Toigo Partners International and Chairman of the Data Management Institute, maintains that storage efficiency is a ‘Holy Grail,’ challenged by a combination of poor resource management, unwieldy methods for allocating capacity to applications, and a nearly complete absence of intelligent data management. “Budgetary belt-tightening has savaged the ranks of storage administrators, leaving server and application managers with additional work in capacity, performance and data protection management. And most customers didn’t anticipate the enormous storage I/O requirements created by virtual ‘multi-hypervisor’ environments,” he said, adding that “very little is being done to develop real strategies for moving beyond the mess.”

Because server virtualization is considered crucial
to helping federal organizations consolidate physical IT environments, a robust and efficient storage infrastructure designed to support virtual server environments is equally as important. To realize the greatest economies of scale and achieve optimal performance, government organizations must investigate advanced storage management tools to support virtual server environments.

“For 80% of the work being done, storage virtualization solutions can provide an enormous increase in speed and performance in virtual server environments,” Toigo said.

While layers of storage virtualization already exist in most organizations, it’s smart to look closely at the newer tools that can help to reduce costs and gain greater storage efficiency, he added.

Because storage remains one of the most poorly utilized, and expensive components of IT hardware acquisition budgets, Toigo advocates the concept of ‘strategic storage’ to help IT managers and planners better understand the variety of storage options available to them.

Technologies to Address Storage Management Headaches

Industry observers and suppliers maintain there are real advantages to implementing a ‘shared storage’ infrastructure to help improve overall utilization and performance in virtual server environments. A shared storage infrastructure is a multi-workload computing environment that uses the same storage back end to gain efficiencies in storage use, scalability, and overall performance. While this typically involves some form of virtualization at each layer of the stack, including storage, networking and computing, it does not exclude the use of traditional physical servers.

By building a shared storage infrastructure, organizations can achieve a much higher level of utilization, for lower operational costs, observers said. A more tightly integrated virtual server/storage environment is also easier to manage, and can scale to meet ongoing growth requirements as they arise. While there are multiple ways to achieve greater storage efficiency, industry observers and suppliers maintain that deduplication and thin provisioning are technologies that can increase the amount of storage available without having to add physical disk capacity.

In traditional methods of provisioning storage disk capacity is preallocated, long before its needed, while thin provisioning consumes storage on an as-needed basis. Deduplication, meanwhile, refers to the elimination of redundant, duplicate data, as only one unique instance of the data is actually retained on storage media, such as disk or tape. Redundant data is replaced with a pointer to the unique data copy.

Other storage technologies, such as volume-level Snapshot copies, provide subsecond copies that take no additional disk storage space. Snapshot software provides a point-in-time copy of data that can then be backed up to a controller or dumped to an archive medium for long-term storage. These backups have no performance impact on data protection and consume minimal storage space. Snapshot copies also enable backups from single files or an entire volume.

“The primary benefit of incorporating advanced storage virtualization technologies is an overall reduction in physical hardware, which in turn will help lower both operating and capital expenses,” said Jon Benedict, Systems Architect, Infrastructure and Cloud Engineering for NetApp Inc., Sunnyvale, CA.

Storage Virtualization at Salsa Labs

Salsa Labs, Washington, D.C., provides a cloud-based ‘software-as-a-service’ (SaaS) offering to more than 2,000 specialized user groups, supporting 50 million members, donors and activists around the world. In the past 18 months, Salsa Labs has grown from using one small, virtualized server implementation, to a nearly 70% virtualized data center, thanks to the addition of a shared storage infrastructure that has delivered flexibility, scalability and a tight integration with the lab’s virtual server environment based on Red Hat Enterprise Virtualization (RHEV).

By examining what Salsa has done, public sector organizations can learn more about what it takes to successfully meld server and storage virtualization to help streamline data center operations and lower costs. According to Justin Nemmers, chief operating officer, two NetApp FAS 2050 storage subsystems were deployed...
in the Salsa Labs data center in early 2010 to help him quickly provision services for customers of the company’s SaaS offering. Much like government institutions, Salsa Labs must provide enhanced security and redundancy across multiple locations. In addition, “We must also provide quick and easy provisioning of services,” Nemmers said.

By implementing integrated server and storage virtualization, Salsa Labs has centralized and simplified storage management, increasing the Lab’s end-to-end visibility into storage operations, and improving both utilization and efficiency. The Lab’s RHEV workloads currently in use include web server, load balancing, support, and ancillary email systems that support customer activities such as advocacy messaging. The close integration between Red Hat and NetApp enables Salsa Labs to easily create virtual machines as needed. “We can set up and test a single instance of Red Hat Enterprise Virtualization on our NetApp storage platform in minutes,” he said.

Using the NetApp’s storage management tools, Nemmers can also manage a growing server and application environment on a limited budget, while continuously optimizing storage space, he added. For example, Salsa Labs uses NetApp’s Snapshot technology to create backup copies of internal business data, easily and instantly. And by using NetApp Operations Manager, Nemmers can view, monitor and quickly provision storage resources required to meet customer requirements as they leverage the lab’s SaaS offering.

Now with 70% of its production workloads virtualized on NetApp storage, Nemmers has turned his focus on additional methods for increasing storage efficiency...

“Now with 70% of its production workloads virtualized on NetApp storage, Nemmers has turned his focus on additional methods for increasing storage efficiency...”

Now with 70% of its production workloads virtualized on NetApp storage, Nemmers has turned his focus on additional methods for increasing storage efficiency across the virtual environment. For example, Salsa Labs has implemented deduplication, reducing storage...
often referred to as a ‘boot storm.’ While this scenario may be unavoidable, there are ways to leverage faster storage devices, and invest in automated storage tiering tools that can assist in reducing the load. Also, it’s important to stagger backups to avoid too many systems accessing virtual servers at once.

**LEARN TO CONSERVE** – Investigate modern storage management tools to gain capabilities that can ultimately reduce the agency’s storage footprint as well. Techniques such as thin provisioning reduce the amount of storage allocated for each virtual server. Deduplication can also drastically reduce redundant or duplicate information stored on virtual servers, conserving space and lowering overall storage costs.

**KEEP TABS ON PERFORMANCE** – Just like security measures that require continuous monitoring, storage too, should be monitored to avoid potential problems or bottlenecks before they impact overall system performance. Also, it’s important to keep in mind that monitoring a virtual environment isn’t as simple as monitoring a physical environment. Products designed for virtual environments can help pinpoint exactly which resource or component may be causing a bottleneck. And when multiple storage virtualization tools are used, it’s even more important to automate monitoring rather than performing this task manually, setting alarms to alert personnel when specific data stores may be overcommitted.

**A Combination That Really Works**

Pairing the RHEV platform with NetApp storage provides a powerful, efficient, scalable virtualization platform. Rob Washburn, RHEV Solutions Architect for RedHat explained how customers must understand the storage bottlenecks that emerge when 40+ servers try to access storage on local disks. “In most cases, I/O performance suffers,” he said.

The good news is that virtual machines running on Red Hat Enterprise Virtualization can easily store and maintain data on NetApp’s technology leveraging advanced features such as snapshots, mirroring and replication to alleviate the potential for bottlenecks and
optimize virtual application performance. Ultimately, he adds, “agencies must pay close attention to storage, networks and server memory to fully optimize their virtual IT operations.”

NetApp’s Benedict authored a technical paper on server and storage virtualization that advocated the need for greater planning and integration between servers and storage in virtual environments. “Deploying servers has gone from a complex, convoluted process in traditional computing environments to a quick, easy online process via virtualization,” he explained.

The challenge now is, once users understand they can add virtual machines ‘at will’ that doesn’t necessarily mean the network and/or storage can handle the added workload burden. “Storage virtualization can help agencies to handle multiple workloads, abstracting physical disks to match each organization’s requirements, primarily by customizing the configuration of the controller to manage multiple, different tiers of storage from a single controller,” Benedict explained.

Jeff Boles, senior analyst at the Taneja Group, maintains that while most organizations have until recently been fixated on storage capacity, resolving storage performance issues is a primary focus in the months ahead: “Storage solutions chosen today should provide a significant performance boost and a promise to scale beyond what the organization requires today,” he said.

Also, when it comes to performance, even small improvements can transform how well storage works in a virtual environment. Looking ahead, new tools and storage solutions are likely to leverage cloud computing. Taneja Group has identified storage products associated with the cloud as growing into a $10 billion market by 2014.