![commvault solving forward]

IDENTIFYING AND SOLVING VM SPRAWL
USING AUTOMATED VM ARCHIVING

Today's VMware environments are plagued by out-of-control growth of virtual machines, increasing costs and wasting valuable resources. **LEARN HOW ARCHIVING VIRTUAL MACHINES CAN MITIGATE VM SPRAWL WITH A COMPREHENSIVE APPROACH TO VM LIFECYCLE MANAGEMENT.**

Contents

Introduction

Think back… way back… there was a day in the world of Information Technology in which physical server sprawl was a significant problem in the data center. In this context, the term "sprawl" refers to a situation in which new servers were deployed each and every time a new service or new business workload was needed. Sprawl came about as IT shops adopted a "one to one" mentality with regard to the number of primary workloads to run per server. In order to reduce the potential for two applications to conflict with one another on the same hardware, IT simply deployed two physical servers to support those two workloads. This allowed the applications to each run in their own space without worry that they would negatively impact the operation of other workloads.

However, over time, the constant addition of new servers took a major toll on the data center. Every new server required both rack space and electricity and each new server generated heat that had to be removed from the room. Over time, some of the purposes of these machines was forgotten or the need eliminated, but the systems remained in place out of fear that they were still serving some unknown purpose.

In fact, this situation was what led to the initial push for virtualization. Some of the first major virtualization initiatives were born of CIOs' frustration with this growing state of affairs or from Sys Admins' annoyance with deploying new physical servers. With virtualization came freedom from having to deploy a physical server every single time a new need was identified. Organizations could finally begin to reduce the sheer number of physical server present in their data centers, not to mention ever-increasing power and cooling and even new data center space required as rack and floor space were consumed in existing facilities. After all, most physical servers were overbuilt and had overall resource utilization averaging less than 10% on an ongoing basis.

With virtualization, many of the workloads running on these smaller, heat-spewing, electricity guzzling hosts were converted to virtual machines and migrated to fewer hosts. Instead of a lot of resource inefficiency, organizations suddenly had data centers with hosts that were running upwards of 50% utilized and even up to 80% and 90%. With fewer such hosts, data center electricity costs began to subside and air conditioning units across the planet breathed a collective (and cold!) sigh of relief.

In addition to reining in physical server sprawl and enabling organizations to make full use of the resources on host systems, virtualization also made it possible for organizations to deploy new applications much faster than was possible in the past. Whereas a new physical server deployment required major lead time for hardware orders to ship and for IT staff to install servers in racks and run appropriate cabling, virtual servers could be deployed with just a few clicks of the mouse.

It was a new age in computing.

It was also the beginning of the age of virtual server sprawl. Although organizations rejoiced as server consolidation projects started yielding returns, it turns out that we simply used a new technology – virtualization – to mask a people and process problem.

How Does Server Sprawl Happen?

Server sprawl isn't generally the result of some nefarious scheme by a disgruntled IT staff person. In fact, in almost every case, virtual server sprawl is the result of a bunch of smaller, perfectly appropriate actions that create an aggregated result that is negative to the organization. Sprawl is one of those problems that doesn't get much attention until someone takes a hard look at the environment as a whole and realizes that there is an issue that needs to be resolved – often after it is already a serious issue.

Lack of Policy or Lack of Adherence to Policy
Some organizations have policies that define how and when a virtual machine can be created. In some cases, these policies will define when a virtual machine can be removed. As is the case with all policies, effectiveness is dependent on people actually adhering to the policy.

Combined Test/Dev and Production Environments Without Oversight
Today, more and more organizations are moving towards a DevOps model in the Information Technology area. Under this model, developers and systems administrators work together to provide full systems development life cycle support for applications. This model effectively tears down latency-inducing and error-prone development processes that used to call for a cold handoff from development to systems operations.

In such environments, there may be more virtual machines created and not potentially looked after once the groups are combined. In the old model, there were test/dev machines handled by developers and production machines handled by operations. With the walls coming down, departments now have even less issue deploying or requesting new virtual machines, but there is usually little to no focus on cleaning up stale virtual machines after the project is complete.

As a Normal Part of Operations
In many organizations, dozens, hundreds, and even thousands of virtual machines are created and destroyed every day. In fact, in 2010, Salesforce.com indicated[1] that the company spins up more than 8,000 virtual machines each and every day. In these cases, sprawl is just a part of normal operations, but there are generally well-defined processes to manage the issue. However, there will still be cases of "one off" virtual machines being created that may fall outside these normal operations. In organizations that have adopted self-service provisioning as part of their private cloud infrastructure, VMs are often provisioned through the portal, and then orphaned when they're no longer needed – and without the right policies and processes in place, IT may never be able to identify these orphaned VMs. So, because these administrators are not the owners of these VMs, they have no way of determining which workloads are active and can be safely removed.

> "Private cloud requires significant focus on process and organizational transformation challenges - technology is not a key inhibitor."

PRIVATE CLOUD MATURES, HYBRID CLOUD IS NEXT, THOMAS J. BITTMAN, GARTNER, SEPT 6, 2013.

---

[1] www.computerweekly.com/news/1280097230/The-datacentre-designs-powering-Salesforcecom

Fear plays a big part here, too. When things are working well, good advice indicates that it shouldn't be "fixed." However, at the same time, there is active waste going on, and it's not cost effective to simply keep adding hardware to address the problems. Additionally, there's fear from the lack of knowledge about a particular virtual machine. For example, a virtualization administrator is usually hesitant to delete a virtual machine if they don't know much about it fearing that someone may actually need it. Also, deletion would be immediate and permanent. Even if you have backup copy, it is usually only valid for a few weeks after which the machine is unrecoverable. If at some point the VM owner requires the VM to be brought back, the administrator can't recover it because it could lead to a "Resume Generating Event."

The Impact of Virtual Server Sprawl

Today, virtual server sprawl is a continuing and growing issue. While it may not have the visibility that characterized physical server sprawl, virtual server sprawl still costs real money and has real impact on both IT and on a company's bottom line in numerous ways:

- IT staff spend time managing or supporting servers that may no longer have value to the business.

- Storage consumption; VMs are very large files on tier 1 storage. Regardless of the workload these continue to occupy Tier 1 storage leading to wasted space.

- Each virtual server – whether in use or not – consumes resources, whether that comes in the form of the storage space or network ports used by the virtual machine or it comes from a powered on virtual machine that is actively consuming CPU and RAM. Unnecessary resource consumption leads directly to increased cost as organizations add capacity to deal with resource constraints.

- Backup and disaster recovery operations can be impacted by unchecked server sprawl as well. VMs are not idle instance OS updates are installed, AV checks run all the time, causing changes in the VMs that will cause these VMs to be backed up and replicated unnecessarily leading to longer backup windows.

- Bear in mind that backups require disk space and disaster recovery is often achieved by replicating virtual machines between data centers. If these processes are also supporting servers that may no longer have ongoing use, the organization risks losing focus over what are mission-critical workloads and may end up putting in place backup and recovery processes that are the most efficient. This could be at the cost of neglecting the SLA requirements of the individual applications.

Resolving Virtual Server Sprawl

Unfortunately, solving the problem of virtual server sprawl isn't as easy as just deleting the offending virtual machines. Even though the term "server sprawl" is intended to convey a negative situation, the reality is that each deployed virtual machines was important to someone at some point in time. Moreover, that virtual machine may still be important to someone somewhere in the organization.

Over the years, a number of monitoring products have been developed that attempt to help administrators identify virtual machines that may be contributing to sprawl. These efforts have been met with mixed success, though, since monitoring tools, by their very nature, don't generally manipulate the environment. They are designed to simply keep watch of the environment. Even though administrators may have been notified that a virtual machine was potentially unnecessary, time constraints or uncertainty about the importance of the virtual machine have often held them back from taking further action. The administrator could take manual action, but doing so still requires a decision to be made regarding the eventual disposition of the virtual machine…and, remember, deletion is permanent, thus a risk few admins are willing to take.

A better approach is to implement a system that keeps a constant watch over the environment and takes an automated approach to the issue of sprawl. CommVault Simpana software has much deeper hooks into the virtual environment than traditional monitoring tools and provides a policy-based framework through which virtual machine sprawl can be addressed. CommVault provides a unique lifecycle-based solution to the problem of virtual server sprawl and integrates into the vCenter console everything necessary to provide organizations with the best of both worlds:

- The ability to retain virtual machines for the long term without adversely affecting the overall performance of the environment or consuming production resources.
- Maintaining ongoing access to all virtual machines so that they can be placed back into production at a moment's notice should the need arise.

By moving the solution to virtual server sprawl into the organization's normal backup and recovery cycle, there is focus on the issue and sprawl control becomes a part of the overall data protection lifecycle. In addition, it becomes possible to define different rules for different classes of virtual machines.

VM Archiving Guides the Way

CommVault Simpana software provides organizations with a rules-based framework for automatic remediation of virtual server sprawl. This features takes a multiphase, progressive approach to addressing server sprawl with a focus on ensuring that ongoing operations remain stable while addressing the excess resource consumption cause by virtual sprawl.

Identify Target Virtual Machines

The initial step in managing potentially unused virtual machines is to simply shut down those machines and leave them in place for a period of time. Bear in mind that the longer a virtual machine goes unused, the less likely it is to be needed eventually. They key word here is "unused." As such, there are a number of different rules that should be met before an automated systems determines that a running virtual machine is truly unused. These rules will likely differ from across different types of organizations, so tools that help identify idle or stale VMs should be configurable to use one or multiple metrics and thresholds

First, the virtual machine shouldn't be using a whole lot of CPU. If CPU usage is above an administrator-defined threshold, it probably means that the virtual machine is actually doing something and may not be contributing to sprawl. In addition, an unused virtual machine should not consume much in the way of disk I/O and network based resources. Once some or all of these resources (Simpana software allows you to configure which resources to measure and their thresholds) have remained under their individually defined thresholds for a period of time, Simpana software will automatically shut down a virtual machine.

With that action, Simpana software takes the first step in assisting organizations in reclaiming unused compute and RAM resources, which than then be redirected to other, more mission critical needs. Should that virtual machine be required to be placed back into production, it can simply be powered on in vCenter. This identification process does not require any additional processing on the VM or vCenter resources as all of the metrics are collected as part of the CommVaultbackup process. You can even configure auto-discovery subclients that let you automatically assign VMs that are most likely to be archived (e.g.: Test/Dev Resource Pools) to separate archive-enabled backup policies so you don't have to fear inadvertent archiving of production VMs.
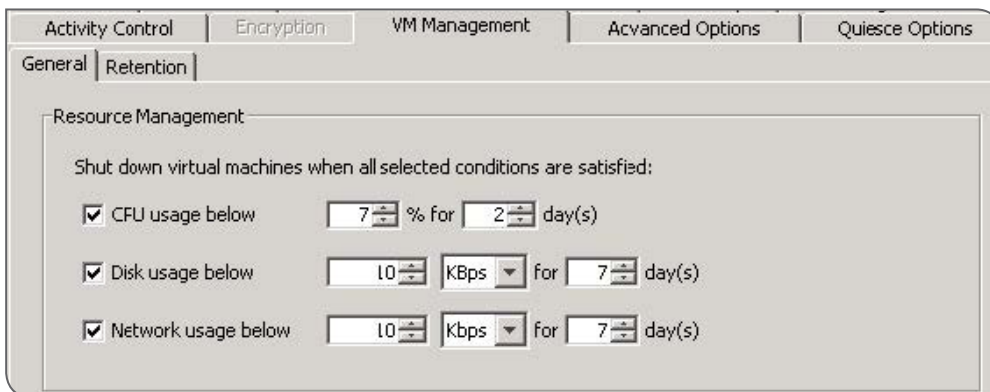

Figure 1: Based on administrator-defined policies, CommVault® Simpana® software will pro-actively shut down virtual machines that meet policy conditions.

Perhaps most importantly is the fact that the administrator can choose whether or not the shutdown action is automatic. In a test/dev scenario, it may be easier to accept a mistake in shutting down a virtual machine, but in certain production environments, doing so could result in business downtime. In order to accommodate such scenarios, VM Archiving can be configured in monitoring-only mode. In this mode, the tool will report to the administrator the list of virtual machines that would have been shut down, but will not take proactive steps to shut them down. It's important that administrators, even when automating certain functions, retain full control over the environment.

Relocate Virtual Machines

As we've already identified, if a powered down virtual machine is still needed, we can simply place it back into production. However, it's more than likely that the virtual machine will simply lie dormant, consuming disk resources while providing no active value to the organization. Bear in mind that powered down virtual machines may have been powered down via the CommVault resource management function or may have been manually powered down by the owners or administrators of the virtual machine.

Often, we want to reclaim the tier 1 storage consumed by a powered down virtual machine, but we usually want to keep it more accessible for a time. After a virtual machine has been shut down for an administrator-defined time period, production storage resources are reclaimed by using Storage vMotion to relocate dormant virtual machines away from primary storage. As a part of this migration, Simpana software allows the administrator to define this interim storage destination and also enables the administrator to choose to store the virtual machine using thin provisioning, which can conserve additional storage resources.

Figure 2: Relocate powered off virtual machines to less expensive storage

Archive

Once a dormant virtual machine has been powered down and possibly offloaded to secondary storage, they can be moved to longer-term archival storage. Again, understanding that different organizations have different needs, the options around archiving dormant virtual machines are fully configurable.
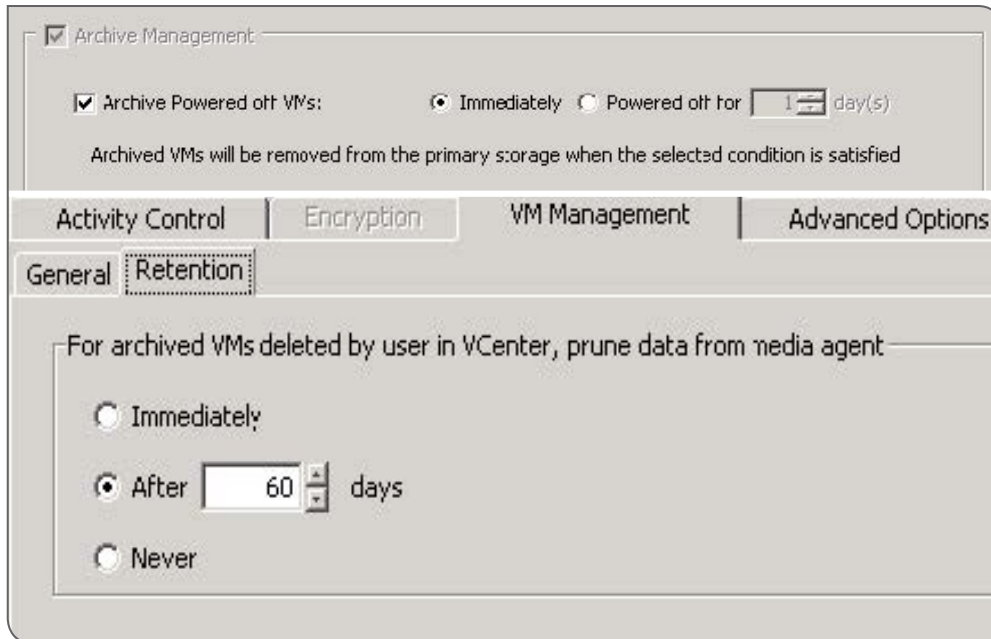

Figure 3: An administrator defines how dormant virtual machines are treated.

In addition, a stub for the virtual machine remains present in vCenter so that the virtual machine, allowing admins to see and report on archived VMs within their vCenter console. Stubbed VMs appear just as any active VM in vCenter, but are identifiable by an icon next to the virtual machine. The stub simply points to the latest version of the VM captured as part of the backup process, so no extra copies are created. So, the VMs remain visible in vCenter, while the VMDKs no longer occupy storage on primary datastores. Eventually, an archived virtual machine may be deleted from vCenter by an administrator. Once this action is taken, Simpana software will retain the virtual machine for an administrator-defined length of time after which it will be permanently deleted.

Recovery

A tool that manages the lifecycle of VMs, but leaves them inaccessible at any point is a non-starter. Access is critical to the successful implementation of any sort of lifecycle management or planning If, at any point, the virtual machine is needed, it must be recovered on demand to the production environment. This process should be able be initiated by an administrator or by the person that needs the virtual machine through the use of a self-service portal. Oftentimes, a full VM is not required, but individual files may be needed from an archived VM. Simpana software enables users to browse these contents and restore single files without bringing a full VM back to a primary datastore.

SUGGESTED VIDEO >>

CommVault VM Archival and Retirement

Learn more about VM archiving with CommVault's Virtual Machine archive tool from this short video. [ii]

WATCH

Understanding that an administrator may not always be available, developers and testers who own archived machines. There is no need to involve the IT helpdesk and wait for the virtual machine to be recovered. Users are provided with on demand real time recovery access to not only their archived virtual machines, but to all currently running active virtual machines they own.

Simpana software provides users with such a self-service portal through which users are able to manage these virtual machine assets. This self-service portal, such in Figure 4, enables a user to simply log in and click a Restore button in order to recover a virtual machine. Self-service portals are increasingly important, particularly as organizations begin to embrace a DevOps-oriented culture in which developers and systems administrators share responsibility for managing and maintaining business critical workloads.
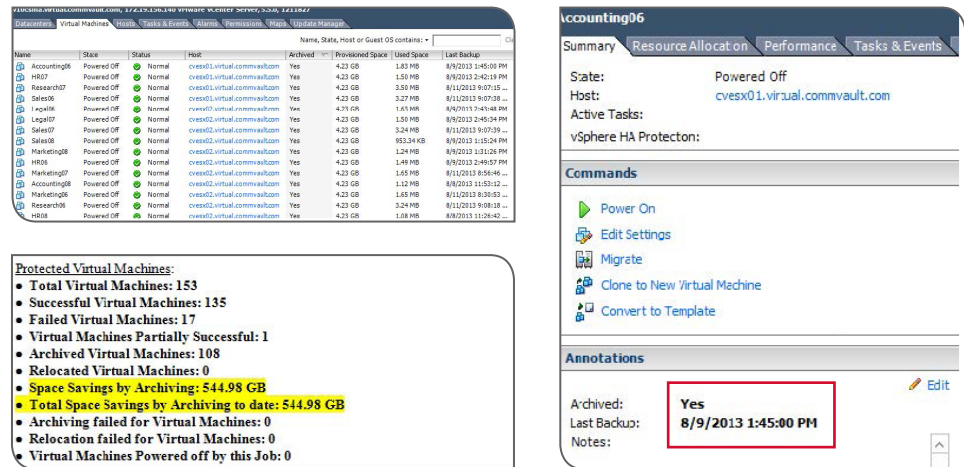

Figure 4: Views of stubbed virtual machines in VMware vCenter.

In addition to the user self-service portal, Simpana software provides a number of additional options so that dormant virtual machines can be placed back into production as quickly as possible. As mentioned, users can use the self-service portal or administrators can make use of the CommVault administrative console. In addition, any IOS or Android powered device can be used to achieve this goal.
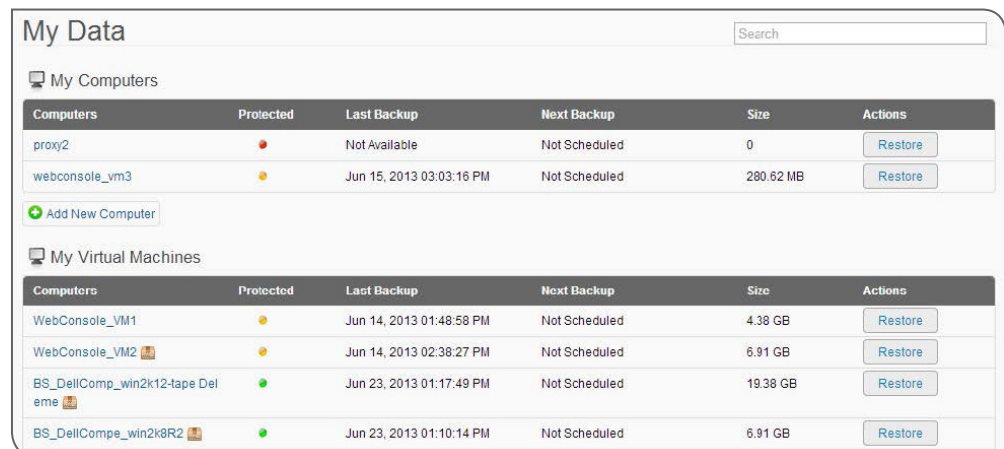

Figure 5: CommVault self-service portal

Summary

Although server sprawl is hardly a new phenomenon, today's interconnected environments provide administrators with the foundation for CommVault Simpana software VM Archiving feature to build upon. Assuming, on average, 30-40% of VMs in a typical enterprise eventually become stale (and, thus, good candidates for archiving), that results in the waste of 30-40% of production resources. That's not just storage. CPU and vRAM, in particular, are also costly resources that would be valuable to reclaim. So, by archiving these VMS, you can increase ESX host utilization increase again (by as much as 30-40%), often deferring the purchase of additional ESX hosts and licensing. By helping administrators take safe, proactive steps to eliminate dormant virtual machines – while still keeping an eye on the need to recover these virtual machines – organizations can reduce costs and fit more mission critical workloads into existing environments.

To learn how Simpana software can help you gain new efficiency out of your VM Ware virtual infrastructure, please visit **commvault.com/virtualization**.

Resources

i   commvault.com/resource-library/1850/reining-in-vm-sprawl-with-vm-archiving.mp4
ii commvault.com/resource-library/2047/commvault-vm-archival-and-retirement.mp4

www.commvault.com • 888.746.3849 • get-info@commvault.com

COMMVAULT REGIONAL OFFICES: UNITED STATES • EUROPE • MIDDLE EAST & AFRICA • ASIA-PACIFIC • LATIN AMERICA & CARIBBEAN • CANADA • INDIA • OCEANIA