Business white paper

# Advanced data protection for your virtualized environments

## The top 5 challenges and solutions

# Table of contents

# Advanced data protection for your virtualized environments

"As the move toward virtualizing mission-critical applications accelerates, end users should reevaluate their data-protection products, architectures and vendors to ensure adequate backup and recovery in order to meet stringent RPO, RTO and SLA requirements."[1]

—451 Research, Backup in a Virtual World

In today's business environment, CIOs and IT executives are charged with fully ensuring the success, integrity, and security of end-to-end business processes. Part of this responsibility involves protecting application data in virtual environments with the appropriate protection level for the business need, while having minimal impact on the virtual infrastructure.

Virtualization is the leading data center IT transformation that is helping customers deploy IT infrastructure/services in a more agile manner. Virtualization enables organizations to reduce CAPEX/OPEX costs, improve the server deployment process and lower the time to provision IT infrastructure. While it may sound straightforward enough, data protection in virtualized environments is actually one of the biggest information management challenges for organizations to overcome. But with the right systems in place, IT professionals can move forward, leveraging the benefits of virtual environments while protecting virtual data with confidence.

When you consider the compelling benefits of virtualized environments (such as consolidation, improved management and greater levels of reliability/standardization), it's no surprise organizations are choosing server virtualization. There are gains at every turn—from the cost savings of server consolidation to improved business flexibility, agility, and scalability. There is no more efficient way to ramp up infrastructure capabilities in the event of a merger/acquisition or impending litigation than with a virtualized infrastructure. And that's why virtualization technologies are fast becoming a must-have in the modern data center.

1 451 Research, Backup in a Virtual World, 2012, Date accessed 4/15/13: https://451research.com/report-long?icid=2259

But while it's hard to resist the lure of virtual server deployments, it's important to be aware of where you could go wrong. Here are the top five challenges that you should know about when evaluating data protection for your virtual environments:

1. **Resource constraints.** As you probably already know, backup is one of the most taxing processes a server will run. Because virtualization is designed to get more utilization from physical servers, CPUs, memory, and networks are consumed meeting other compute loads, leaving fewer resources available for the backup process. When choices are limited, what do you do? Do you avoid backing up some virtual machines (VMs)? Do you limit the number of VMs to a server?
2. **VM sprawl: adding and removing VMs.** One of the benefits of virtualization is that it makes it easy to create a new virtual machine, but this can also pose challenges. Historically, the process for deploying a physical server was slower. Among other things, you would request a purchase order, get rack space, route power and network connectivity, and then let the rest of the organization (including the backup admin staff) know a new server was coming online. The process to install a new physical server took days or weeks. Now, with virtualization, a new VM can be created in a few minutes. If there isn't a solid process in place to accompany this fast on-ramp, relying on a manual process to apply protection policies isn't reasonable and could create inconsistencies and add risk.
3. **The dynamic nature of VM environments (vMotion, etc.).** Given the dynamic nature of virtualized environments, you could find yourself saying, "Where's my VM now?" Virtualization makes it very easy to move a VM from one physical server to another. This movement can be manual or automatic based on some event such as a failure or for load balancing. For large organizations it's impossible for humans to keep track of what VMs are running on which physical servers. Manually updating the backup application with the location of the VM is inefficient, time consuming and isn't a reliable process.

"IT managers are unwilling to virtualize tier-one applications until they are certain that those applications are well protected and that they can recover VMs, files, applications or parts of applications in a timely-enough fashion to meet the stringent service-level agreements (SLAs) required by mission-critical applications."[3]

—451 Research, Backup in a Virtual World

4. **Distributed big data.** When you consider that each VM has its own copy of an operating system, the reality becomes that there is common operating system (OS) data everywhere. Because the OS copy is often a "golden image" blessed by IT, it's common across all VMs. This has caused an explosion of primary storage needs across the organization because while keeping copies of all this data may be necessary, it also causes an explosion in secondary storage needs as well.
5. **Application-aware protection.** Many organizations are deploying applications in virtual environments to reap the benefits of virtualization as summarized above. But applications running in a VM add another layer of complexity to the backup process. Backing up an application without transactional integration (in the event that the guest operating system can't quiece the application itself), leads to crash consistent backups and can significantly increase the recovery time. Additionally, some applications can consume a significant amount of storage.

Even though the risks and pitfalls discussed above can seem daunting, there are solutions available that can reduce your risk, eliminate the fear, and enable you to ensure the success, integrity, and security of your end-to-end business processes. Here are five elements of a consistent, efficient, and policy-based approach to data protection for your virtual environments:

1. **Flexible protection options.** Every business has its own unique requirements, which means you can't drop in one-size-fits-all solution and expect it to work every time. For that reason, choose a solution that offers flexible protection options, so you can adjust to fit your needs. A comprehensive, enterprise solution can be tailored to meet specific protection needs, as the environment changes. For instance, having the ability to offload backup processing to other servers and storage arrays is a must.

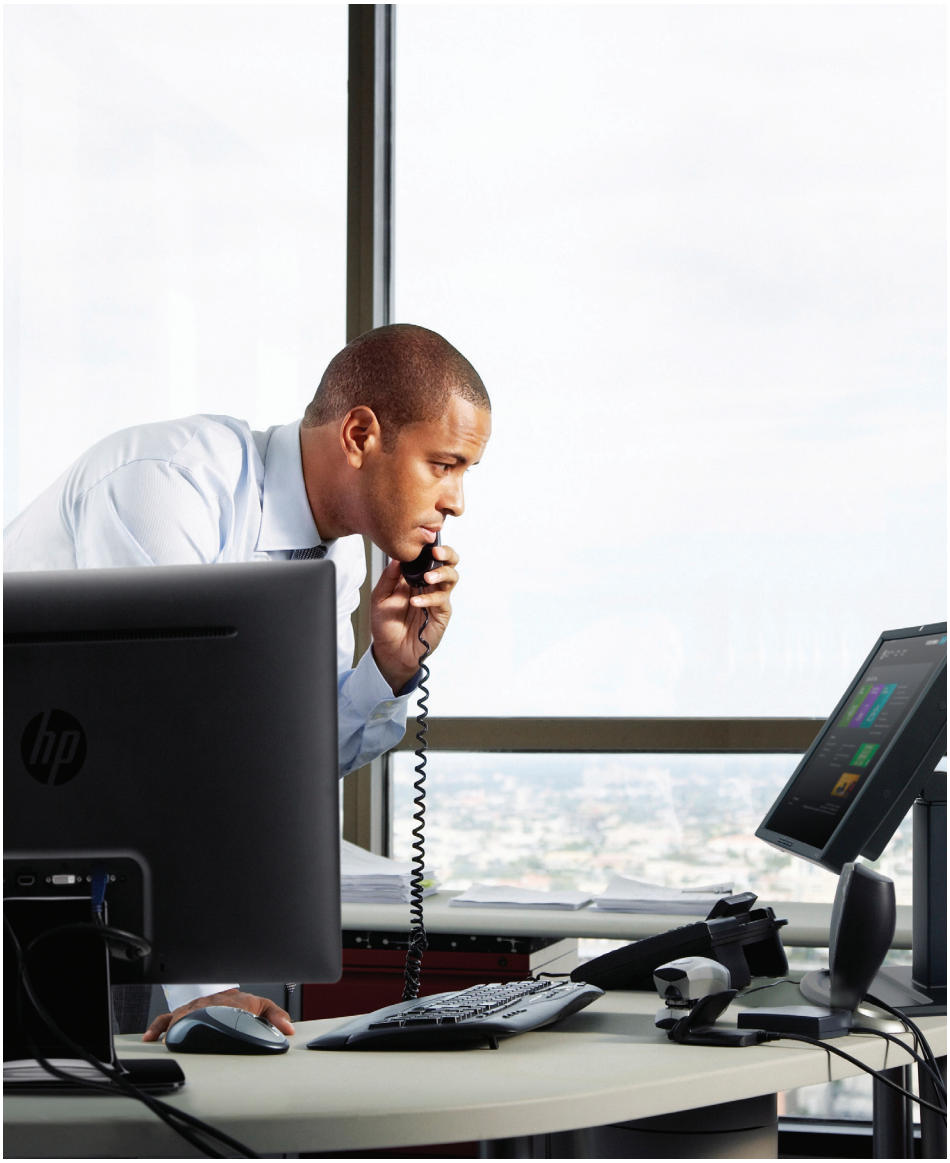
2. **Policy-based protection.** With the ability to rapidly add—and remove—VMs, an enterprise solution needs to be able to automatically provision backup to new VMs. The solution should recognize when a new VM is created and apply an appropriate protection policy automatically. You should be able to specify different protection levels so the right protection can be applied to the new VM. Similarly, when a VM is removed from the environment the protection policy should be removed as well.

3 451 Research, Backup in a Virtual World, 2012, Date accessed 4/15/13: https://451research.com/report-long?icid=2259

3. **The ability to move protection with VMs and track with the virtual data center.** As VMs are moved around the organization, across sites, and into cloud infrastructure, it's crucial that you are able to keep track of each VMs location so you can properly apply protection. This requires integration with the virtual infrastructure. To correctly and automatically apply protection, your protection solution must capture and utilize metadata about the virtual data center. This type of integration is also critical to speed recovery processes.

4. **Advanced deduplication.** As VMs move around the organization, you must be able to move data along with them. Being able to deduplicate the OS, application, and data within a VM enables significant cost savings. But deduplicating at the target system isn't enough; you also need to be able to optimize data transferred across the network. A solid solution should enable data deduplication to occur at different points in the infrastructure, depending on available resources and data protection needs.

5. **Integrated application protection.** In addition to meeting all the challenges already outlined, you also need to protect your information within the context of an application. Protection of an application within a VM requires integration with the virtual infrastructure, as well as the application itself. This integration should also maintain transaction logs and automatic restart of the application upon recovery. Since many applications rely on a large amount of data, the solution should also integrate with array based snapshots to offload the data movement from the virtual machine infrastructure.



## HP Autonomy for data protection in your virtual environment

HP Autonomy's data protection solutions for virtual environments offer you the flexibility and scale to meet the your most demanding data protection challenges. By leveraging a robust architecture that enables the offloading of processing to different points in the infrastructure, HP Autonomy data protection solutions are a natural fit for optimizing protection for the often resource-constrained data protection process found in virtual environments.

**Key benefits**

- **Best in class integration** with HP Storage to enable array-based snapshots for even the most demanding SLAs
- **Automatic discovery and provisioning** of protection for VMs to minimize IT intervention
- **Policy-based protection** to meet the various VM protection needs
- **VMware vCloud Director integration** for vApp protection whether the VM is on premise or running in a cloud
- **Federated StoreOnce deduplication** enables data reduction at the VM, backup server, or target system to facilitate the efficient movement of data across the environment
- **Comprehensive application integration** including point in time recovery, transaction log management, and automatic restart of application upon restore

## About HP Autonomy

HP Autonomy is a global leader in software that processes human information, or unstructured data, including social media, email, video, audio, text and web pages, etc. Autonomy's powerful management and analytic tools for structured information together with its ability to extract meaning in real time from all forms of information, regardless of format, is a powerful tool for companies seeking to get the most out of their data. Autonomy's product portfolio helps power companies through enterprise search analytics, business process management and OEM operations. Autonomy also offers information governance solutions in areas such as eDiscovery, content management and compliance, as well as marketing solutions that help companies grow revenue, such as web content management, online marketing optimization and rich media management.

Please visit **autonomy.com** to find out more.

**Sign up for updates**
**hp.com/go/getupdated**

Share with colleagues



autonomy.com

20130418_RL_WP_HP_Advanced_DP_Virtualized_Environment