



WHITE PAPER



Protecting and Recovering Virtual Server Environments

Introduction

Server virtualization technology provides compelling advantages to enterprises looking to cut administrative costs and improve overall server availability. Gartner Inc. predicts that by 2012, over 85% of enterprises will be using it extensively in production environments. Recent enhancements to popular server virtualization platforms are making them more and more viable for mission-critical applications. As enterprises consider moving their mission-critical applications into virtual server environments, recovery capabilities become an important concern.

Recovery has several facets which must be explored. When it comes to protecting data, there are backup and restore capabilities designed to meet daily operational requirements and data distribution capabilities for disaster recovery. Common concerns in data recovery include the production impacts of data protection operations (e.g. backup windows, overhead, etc.), recovery point objectives (data loss on recovery), recovery time objectives (how fast can recoveries occur), and recovery reliability (am I always able to recover when I need to).

But recovery capabilities must go beyond just data – they must also include application recovery designed to restore the application services that meet end users' business requirements. Generally, enterprises are using one set of tools to protect and restore data, and another set of processes and/or tools to perform application recoveries. If mission-critical applications are running in virtual server environments, then recovery tools need to be able to meet very stringent recovery requirements.

There is one key factor which makes deploying recovery tools in virtual server environments very different from deploying them in physical server environments. Physical servers, because of the difficulty associated with provisioning and reconfiguring them, were generally set up with significant headroom available across CPU, memory, storage and network resources to accommodate growth. Virtual servers, on the other hand, are very easy to provision and reconfigure, and most enterprises run them at very high utilization rates in an effort to squeeze the most out of existing computing resources. This higher utilization is in fact a key component in the cost savings that can be achieved with server virtualization technology. Server utilization rates are a key factor in determining how protection and recovery is handled.

Effectively Scoping Recovery Requirements

The way virtual servers are primarily deployed is by adding them to pre-existing environments. Today, those environments already include heterogeneous servers and storage, different storage architectures (storage area networks, network attached storage, direct attached storage) and protocols (Fibre Channel, SCSI, iSCSI), a variety of different applications, and physical servers. And as virtual servers are added, they may be based on technology from VMware, Citrix, Microsoft, or some combination thereof.

Today, these environments already also have processes in place for protecting and recovering data and applications, both locally and remotely. Most enterprises have a mix of products,

such as backup and restore software, replication products, and/or high availability failover clustering, which are being used, in many cases along with very manual processes, to meet recovery requirements. And if these processes are based on tape infrastructure, array-based replication, and/or shared disk clustering, there's a good chance those enterprises are looking to minimize the impacts on production environments and administrative resources of data protection operations, reduce data loss on recovery, improve recovery times and reliability, simplify the environment, and lower overall costs.

If you've implemented virtual servers and just tried moving your existing physical server recovery processes over, then you've already discovered that this doesn't work very well.

Host-based backup agents and replication software generate too much overhead to be used in virtual servers that are running at or near resource capacity levels. This is why server virtualization vendors like VMware have introduced storage APIs at the hypervisor level that promise to enable “backups” with very little impact on virtual server resources.

But hypervisor-level APIs suffer from a significant problem: they are specific to a particular virtual server platform. If you’re interested in improving efficiency, the last thing you want to do is move towards a model that uses different recovery tools and processes for each platform. This increases training costs, complicates recovery, and introduces needless recovery risk. When recoveries are required, will you always have your most highly skilled administrative resource on that platform available? Possibly, but do you want to bet on that to be the case in every recovery scenario? Do you really want to bet your business on it if a disaster occurs?

Except in a certain number of very limited cases, all initial data recoveries, regardless of whether they are file-level or system-level, should be performed from application-consistent recovery points. That is because recovering from application-consistent points in time is the best way to ensure a rapid, reliable recovery. Crash-consistent recovery points can result in reliable recoveries, but even if they do they take significantly longer. Backup agents are used to create application-consistent recovery points in physical server environments, and because these servers are generally overprovisioned, backup agent overhead is not a critical concern. But in virtual server environments where virtual machines (VMs) tend to be run at utilization rates approaching 85% or more, backup agent overhead is a huge concern. The use of conventional backup agents is just not viable in most virtual server environments due to overhead issues.

The recovery requirements for virtual server environments should be as follows:

- Centralized management of a recovery platform that covers both physical and virtual environments, is not tied to any single virtual environment, and supports heterogeneous servers, storage, and applications
- Comprehensive coverage that handles both data and application recovery, locally and/or remotely, by leveraging disk-based recovery and recovery automation to reduce risk and improve recovery reliability – in other words, a single integrated solution that covers the entire recovery continuum

- Low overhead design that provides the needed recovery capabilities (however stringent they may be) to meet critical application requirements yet minimizes resource utilization on production VMs and networks
- Application-consistent recovery options

It’s important to understand the limitations inherent in a recovery solution that is specific to a single virtual server technology. VMware may be the most widely deployed server virtualization technology in production environments today, but it’s a safe bet that over time, both Citrix and Microsoft will gain ground against them. Any recovery products that leverage platform-specific APIs will lock you into that platform or force you to buy redundant recovery products. Heterogeneous support is key not only for making things simpler, but also for keeping costs down.

The consolidation that occurs when server virtualization technology is deployed helps create a simpler, more cost-effective environment. A centralized, recovery management platform that can be used to comprehensively address all recovery requirements across all existing environments will offer the same advantages.

InMage: One Solution For Your Recovery Needs

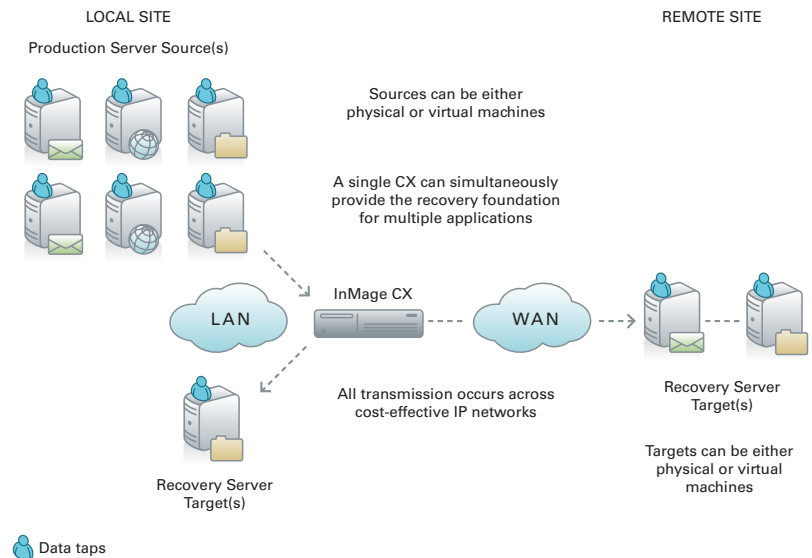


Figure 1. InMage uses filter drivers, called data taps, to capture changes to production servers on a real time, write-by-write basis, and send them across a LAN to the InMage CX, an appliance which can then replicate the data to one or more targets either locally or remotely.

InMage is a software-based solution that supports transparent backup, disaster recovery (DR) from catastrophic failures within minutes, and automated application failover/failback for faster, more reliable restoration of application services. It is a single recovery platform, managed from a centralized management GUI, that covers physical and virtual machine environments (including VMware, Citrix, and Microsoft), heterogeneous servers (Windows, Linux, Unix) and storage (DAS, SAN or NAS), and any applications.

Built upon next generation recovery technologies like continuous data protection (CDP), asynchronous replication, application failover/failback and WAN optimization, InMage allows customers to completely dispense with backup as a discrete operation and move to a “continuous backup” paradigm. This approach eliminates backup windows, supports very granular recovery capabilities and optimizes WAN efficiency for DR configurations. **Data is protected and recoverable as soon as it is created, not just once it’s backed up** – a very important consideration for critical application environments. Solutions commonly deployed by InMage customers include local backup, remote disaster recovery, backup consolidation, and remote office backups.

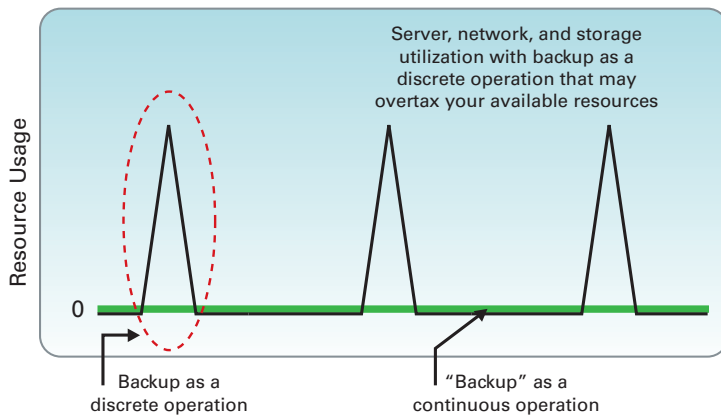


Figure 2. The move to “continuous backup” that CDP enables has a significant impact on resource utilization, significantly decreasing CPU, memory, and bandwidth requirements relative to discrete or periodic backups. Resource usage spikes drive the resource overprovisioning that unnecessarily increases costs.

The use of CDP technology as a core component is important for two reasons. First, by capturing and recording changes in real time, **CDP can meet requirements for near zero RPOs**. Compared to “point in time” backup approaches that can only reduce data loss on recovery as backups are done more frequently, CDP will generally be able to recover back to the last write performed by an application. And second, CDP splits the daily data capture “load” across a 24 hour period instead of trying to bundle all changes up at once and shove them down a

network. This reduces the instantaneous data protection load on servers and networks to almost zero, **effectively making backups transparent**. With InMage, you’ll never have to do another scheduled backup, yet you’ll be able to recover much faster and much more reliably, with less data loss, than you ever could before with tape-based infrastructure.

With this collection of next generation recovery technologies under the hood, InMage provides a 4 in 1 recovery solution that recovers data and applications, both locally and remotely, and includes storage capacity optimization (SCO) technologies to reduce network bandwidth and storage capacity requirements:

- Disk-based backup and recovery that can meet the most stringent local recovery requirements while replacing a large percentage (and possibly all) of your existing backup agents
- Heterogeneous, asynchronous replication that enables recoveries, with current data, from remote sites literally within minutes
- Integrated application failover/failback that speeds the restoration of application services driving the business and effectively replaces products like shared disk clustering, VMware SRM, and others
- Integrated capacity optimization technologies like WAN optimization, change-only driven data capture, delta differencing, etc. that provide a “deduplication effect” without requiring separate WAN accelerators or deduplicating virtual tape libraries and appliances

Finally, it is important that InMage does not require an expensive SAN, yet can clearly support SAN just as it supports storage environments with NAS and DAS. InMage’s unique hybrid recovery architecture leverages the best of host-based and appliance-based models without their disadvantages. Note that with InMage, data is captured at the host (or VM) level, sent across the LAN to the CX appliance (where dedicated resources are used to power the rich functionality InMage supports), and from there replicated to server (not appliance) based targets. This creates a model that puts very low overhead on protected servers, has ample processing power to support rich recovery functionality, supports both file-level and system-level recovery (including bare metal recoveries in VM environments), and can be deployed in a variety of ways to support backup, DR, application recovery, and/or data migration needs across the enterprise.

InMage AppShots: The Best Application-Consistent Recovery Solution for Virtual Machine Environments

Because of the rapid, reliable recovery that it provides, most administrators prefer to recover from application-consistent points in time. Traditionally, backup software agents interacted

with application-level snapshot APIs, like Windows Volume Shadowcopy Services (VSS) and Oracle RMAN (among many others), to create application-consistent recovery points while applications remained on-line. In virtual machine environments, running backup agents in VMs just creates too much overhead, and few enterprises use this approach.

Many enterprises are, however, using hypervisor-level snapshot APIs to create recovery points with minimal impact on production VMs. But any scripting done to support hypervisor-level recovery in VMware environments is not transferrable to Citrix or Microsoft environments, and clearly can't be used in physical environments. This means that as soon as you introduce VMware into your environment, you will have to have at least two separate backup processes, and possibly more. And if additional virtual server technologies are deployed, you'll add additional backup processes.

InMage definitively resolves all these problems. InMage data taps operate in each physical machine or at the guest operating system level so they can leverage application snapshot API mechanisms like Windows VSS, Oracle RMAN, etc. to reliably produce application-consistent recovery points. Given the way the data tap captures changes, it generates only 1-2% overhead, making it a very viable solution for VMs (as well as physical machines). InMage supports simple file-level restores, so you do not have to restore an entire *vmdk* file (or its equivalent) to do a restore. When you generate an AppShot for recovery purposes, it is already mounted on a running server (the recovery target, which can be physical or virtual) and immediately ready for use - you just copy one or more files from the recovery image directly over to a running VM. This cuts down the number of recovery steps required with other methods. And best of all, InMage deploys exactly the same way across all environments - physical or virtual - so that you have one simple, consistent recovery process used for everything. InMage does not use virtual server platform-specific APIs like VCB or vStorage and therefore is portable across all VMware, Citrix, and Microsoft virtual server environments.

And here's another nuance that many people may not know. The lengthy part of the process when using an application snapshot API is creating the snapshot, during which time the application is effectively quiesced. InMage only tags application-consistent recovery points in the CDP data stream using application snapshot APIs, it does not create them (i.e. move any data) as they are being tagged. Tagging the points takes

literally only a few seconds. Conventional products actually move data to create disk-based snapshots each time a VSS writer (or equivalent) is called - a process that can take minutes - so they impose more impact on production applications, even when they are using application snapshot APIs. Any point that InMage has tagged as an AppShot can be immediately turned into a disk-based recovery point when it is needed **with zero impact on production applications**. And all other data states along that application's recent timeline are also available, something which contributes greatly to root cause analysis in a way that "point in time" snapshots just cannot do. And with InMage, these recovery points are created retroactively when they are needed, which means that zero time is spent managing and pruning snapshots that you may never use.

Conclusion

The increasing penetration of virtual server technology is changing the computing landscape and driving a new set of recovery requirements for the enterprise. What the enterprise needs is a comprehensive recovery solution that covers both data and applications and addresses both local and remote recovery requirements. This centralized recovery management platform must support the heterogeneity that is a fact of life in enterprises today, encompassing heterogeneous servers and storage, physical and virtual machine environments, all types of storage architectures (SAN, NAS, and DAS), and any application environment. And it must do so in a way that meshes well with the needs of all environments: a very low overhead model that provides application-consistent recovery points that can be used to drive data and application recovery both locally and remotely. InMage provides that platform.

If application-consistent recovery with minimal data loss is important to you, InMage is the best solution for all your open systems environments, and with its low overhead model it is particularly well-suited for VM environments from VMware, Citrix or Microsoft. It provides a solution that does not lock you into specific virtual machine products but is deployed the same way, with the same robust recovery functionality, across all virtualized server environments. This centralized recovery management approach leads to the simplest, most scalable and most cost-effective solution to give you the recovery you need to keep your business running with peak performance. And it allows you to replace multiple existing products and tools - backup agents, replication software, failover clustering, etc. - to simplify your environment.

