

PHD Virtual Backup and Replication: Replication Overview

Frank Musacchia
www.phdvirtual.com
December 2011

Introduction

Delivering the highest performance and most scalable cross-platform backup solutions on the market, PHD Virtual Technologies has been transforming data protection for virtual IT environments since 2006. Our mission is to preserve the agility, flexibility and cost savings that drive your IT virtualization goals.

PHD Virtual solutions are purpose-built for virtualization and leverage the award winning PHD VBA™ (Virtual Backup Appliance) Architecture to provide virtualized backup and recovery for VMware and Citrix environments. This evolutionary approach allows PHD Virtual Backup to deliver high performance data protection that seamlessly scales for large and distributed deployments. Unlike alternative solutions, PHD Virtual Backup removes the need to deploy and manage separate physical servers, additional software, scripts or agents for backup and recovery of the virtual environment.

PHD Virtual Benefits

PHD provides a completely virtualized solution leveraging a virtual backup appliance VBA™.

Using a virtual appliance removes the need to deploy and manage a separate physical infrastructure to backup your virtual environment. The VBA™ is purpose-built for data protection and can leverage the advanced capabilities of virtualization.

PHD snaps into your virtualization platform architecture and scales seamlessly.

Unlike other alternatives, our solution deploys directly on your virtualization platform and provides integrated single pane of glass management. Scaling up for increased demand is easily accomplished by deploying additional VBA's, without the need to add additional physical servers.

PHD provides the best backup optimization for LAN/WAN environments.

Leveraging our TrueDedupe™ Technology we can provide the most highly efficient LAN/WAN based backup solution with the minimum impact to your network. Deduplicating and compressing backup data on the source side at the host ensures that only the minimum amount of unique data needed for backup is transmitted across the wire, unlike other solutions which can have a dramatic impact on the network or require more bandwidth.

PHD dramatically lowers your backup storage requirements and solution costs.

With our TrueDedupe™ Technology, customers regularly experience a 25:1 dedupe ratio reducing ongoing backup storage needs by 96%. Since our solution is delivered as an integrated and optimized virtual appliance, management overhead is significantly reduced and there is no need to purchase and manage additional hardware, software or agents.

PHD Virtual Architecture

VBA™ Architecture

A VBA™ is a Virtual Backup Appliance - a virtual machine that backs up other virtual machines

- Deploys as a small virtual appliance on either a VMware vSphere or Citrix XenServer host
- Need a minimum of one VBA per VMware Cluster or XenServer Pool.
- Performs image-based block level backup of virtual machines
- Fully integrated with VMware and Citrix API's
- Supports Change Block Tracking (CBT) for VMware
- Leverages the Virtual infrastructure such as vMotion/Live Motion, DRS, HA and Resource Pools
- Requires no additional proxy/media servers or agents
- Purpose built for VM backup and recovery

Replication Overview

PHD Virtual Machine Replication

With PHD Virtual Backup, you have the ability to further protect your virtual environment by replicating production virtual machines to a Disaster Recovery (DR) Data Center. This allows you to populate a DR Data Center with replicas of mission critical VMs, which can be powered on immediately in case of a catastrophic event.

PHD virtual machine Replication takes advantage of existing virtual machine backups that were created by PHD VBAs. PHD leverages these pre-existing backups, to create exact replicas of your virtual machines at your DR Data Center. After PHD creates the replica VMs, you will have the option to either “Test or Failover” each replica VM, based on your needs.

After the virtual machine replication is finished, it is recommended to back up those virtual machine Replicas. This provides data consistency across the backups at the Production Data Center, Replicas at the DR Data Center, as well as Backups at the DR Data Center.

PHD Replication Advantages

The following is a list of Replication attributes and features which will be explained in further detail:

- Replication from Backups
- Intelligent Incremental Replication
- Test and Failover Features

Replication from Backups

As mentioned earlier, PHD leverages Backups instead of other mechanisms such as virtual machine snapshots to replicate virtual machine data. There are three main reasons for doing this:

1. No impact to the Production Virtualization infrastructure.
2. No impact to the Production Network.
3. No need to introduce additional snapshots, after the backups are completed.

Competition

Other virtualization replication techniques usually rely on virtual machine snapshots, in order to copy production virtual machine data from one data center to another. This means that two sets of VM snapshots are being executed by the backup vendor, one for backup and one for replication.

This means that after completing VM backups, additional unnecessary snapshots need to be

created and left open for the duration of the virtual machine replication. This practically means that snapshots could be opened for a substantial part of a day, for performing both backup and replication of virtual machines.

While it is very common for a backup vendor to leverage snapshots to perform backups, it is not very common or necessary to use snapshots for virtual machine replication. Leveraging snapshots for replication complicates the management process and puts additional strain on the production virtual infrastructure.

PHD Virtual Benefits

PHD chose to replicate virtual machine data from backups, believing this is a much better alternative than using snapshots. By using the backup data for replication PHD eliminates the need to leverage any production virtual infrastructure.

When PHD Replication is being performed, the DR Data Center VBA is reading backup data from the Production Data Center backup datastore, which never traverses the production network. This actually means virtual machine replication could potentially be configured to run during the course of a business day, allowing the replication to take place between the backup storage and the DR Data Center.

This is a breakthrough in how organizations can obtain Disaster Recovery/High Availability, with no impact to their production virtual environments.

Intelligent Incremental Replication

PHD Replication is intelligent in respect to how it actually transfers the replicated virtual machine data. As similar to the backup process, the first pass of any replication job is considered a “Full” replication. Every subsequent replication pass will always be an “Incremental” replication, never having to perform a full replication pass ever again.

Each time a replication jobs runs, **only the changed data** is copied to the replicated VMs thus optimizing:

- The speed at which VMs can be replicated.
- Performance across the non-production network.
- WAN Bandwidth utilization is kept to a minimum.

After each replication job is completed, PHD Virtual creates a snapshot of the replicated, powered-down virtual machines, in order to preserve the replicated VMs state. Since these VMs are powered down, the snapshots consume very little storage space and do not grow over time. The snapshots are in place to provide a few key benefits:

- Testing
- Failover
- VM integrity

Testing, Failover and VM Integrity

PHD Virtual has added several advanced features into its replication software, to ensure flexibility as well as virtual machine integrity. Below is a summary of those advanced features:

Test Mode

A built-in “Test Mode” feature enables the user to validate replicated VM’s at the DR data Center. This feature allows the end user to test the VMs in any way they see fit, to ensure their integrity. This is an extremely valuable tool given that in a DR event of the Production virtual infrastructure, these replicated VMs will most likely be powered on and become the Production VMs.

When virtual machine replicas are in Test Mode, the PHD Replication will be suspended for those replicated virtual machines. This will ensure virtual machine testing will go uninterrupted due to a replication job being started. All other virtual machines not in Test Mode will be replicated as normal.

Failover Mode

A built in “Failover Mode” allows the end user to power on replicated VMs, which can then be used for Production purposes. PHD will acknowledge the VMs are live in a Production environment, and ensure those selected VMs are automatically removed from the next replication job.

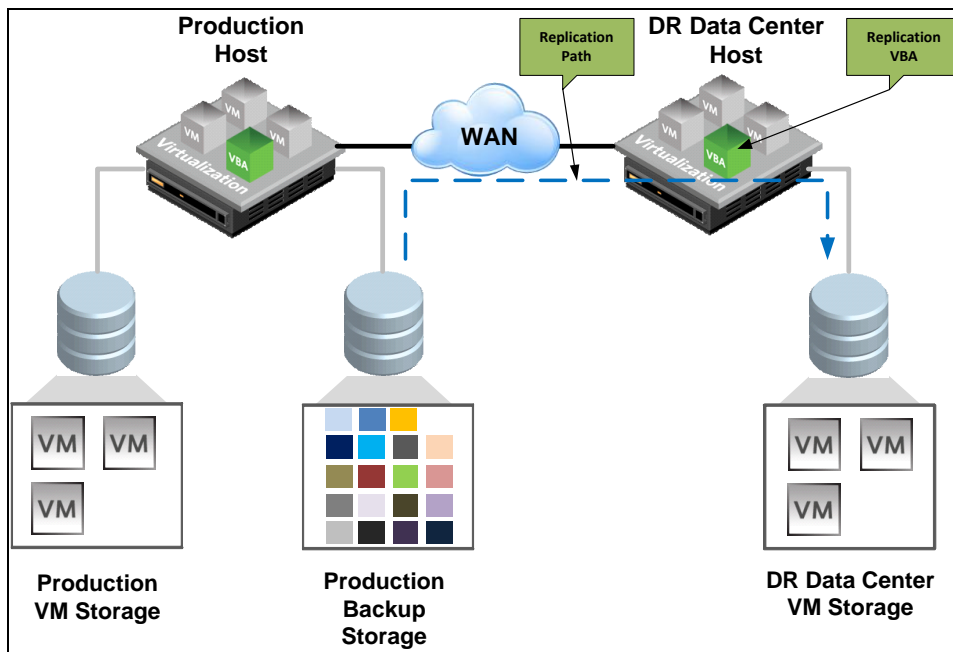
VM Integrity

By allowing a snapshot to be maintained on PHD Virtual replicated virtual machines, any inadvertent change to the data in those virtual machines will be discarded at the start of the next replication pass. This will ensure the intelligent incremental replication will be preserved and there will never be a need to perform another full replication again.

Replication Options

Replication to a DR Data Center

It is possible with PHD Virtual to replicate virtual machines between data centers. This feature allows Virtual Machines which are running in the Production Data Center to be replicated to a DR Data Center.

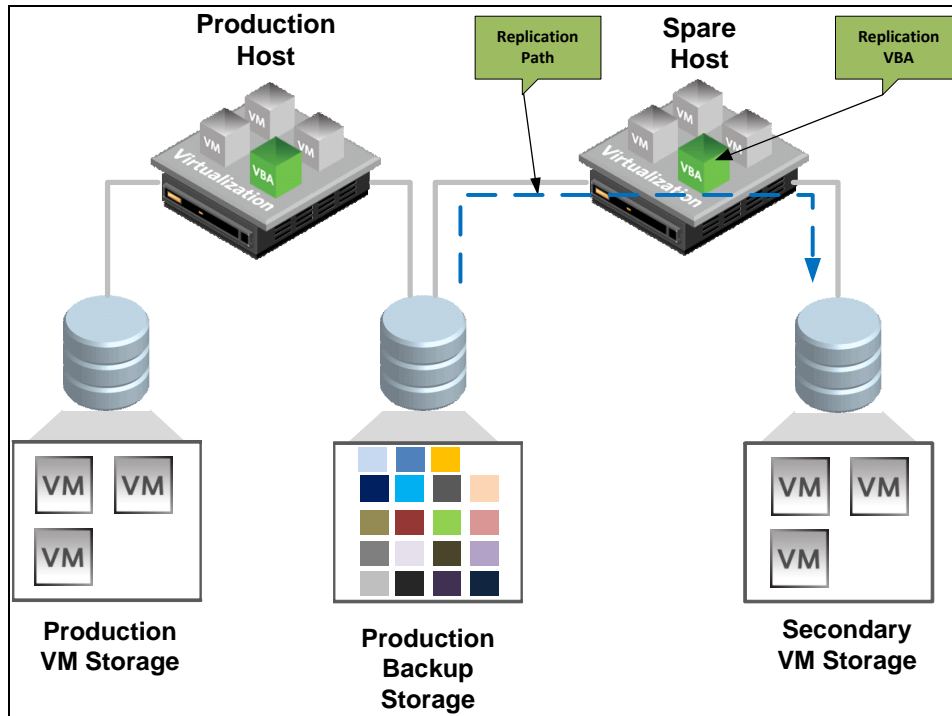


Setup and Configuration

When replicating to a DR Data Center, it is recommended to install a Replication VBA at the DR Data Center. This Replication VBA will be configured to read from the Production Backup Storage and Replicate those VMs to the DR Data Center VM Storage.

Replication within the same Data Center

It is possible with PHD Virtual to replicate virtual machines within a data center. This feature allows Virtual Machines which are running in the Production Data Center to be replicated to spare infrastructure within the same data center.



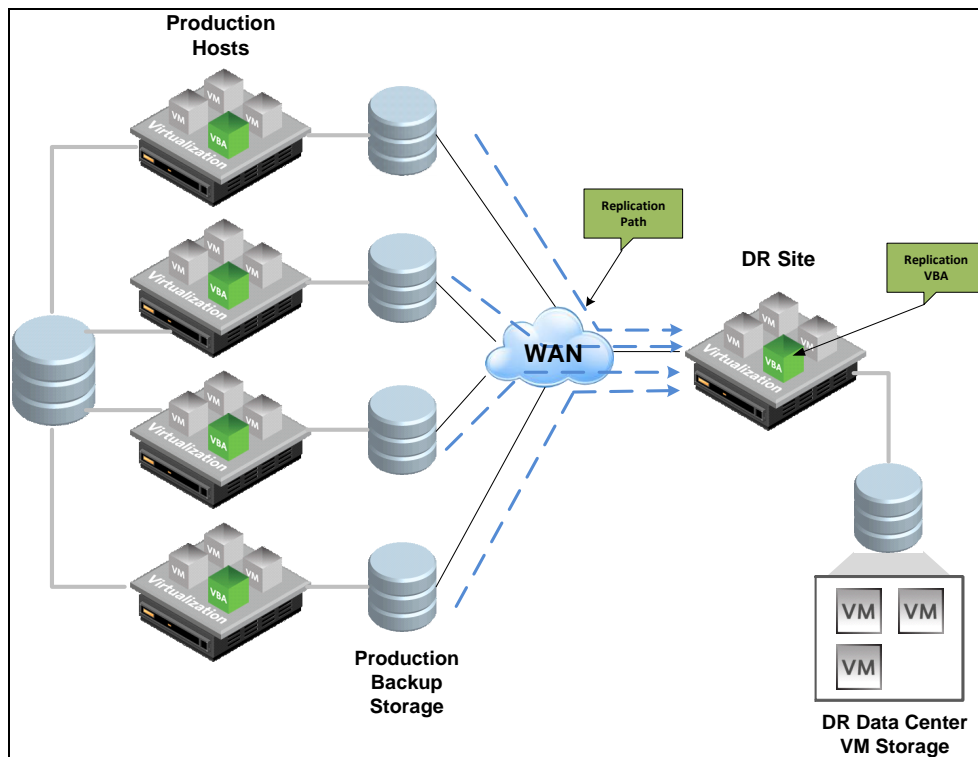
Setup and Configuration

In this scenario it is recommended to install a Replication VBA on a spare host in the Production Data Center. This Replication VBA will be configured to read from the Production Backup Storage and replicate those VMs to the Secondary VM Storage.

Many to Many Replication

A VBA that is configured for Replication at the DR Data Center has the capability of replicating from multiple Production Backup Storage devices at the Production Data Center to the DR Data Center.

This allows PHD to limit the amount of resources necessary to be deployed at the DR Data Center for Replication purposes. With PHD Virtual the DR Data Center could contain as little as one host, in order to optimize resources and costs at the DR Data Center for Replication purposes.



Replication Seeding

In order to speed up the process of replication and limit the amount of WAN bandwidth used between two data centers, it is recommended to use a Seeding process. Seeding is the process of pre-populating a storage device with virtual machine data, prior to beginning regular scheduled replication.

Seeding can be accomplished by taking the data from a production site and moving it to a DR Data Center, using portable media or other methods. Doing this avoids having to send a larger amount of data over the WAN, which can take a very long time depending on the speed and configuration of the network. There are typically two types of seeding methods:

1. Seeding from Backup Data
2. Seeding from Replication Data

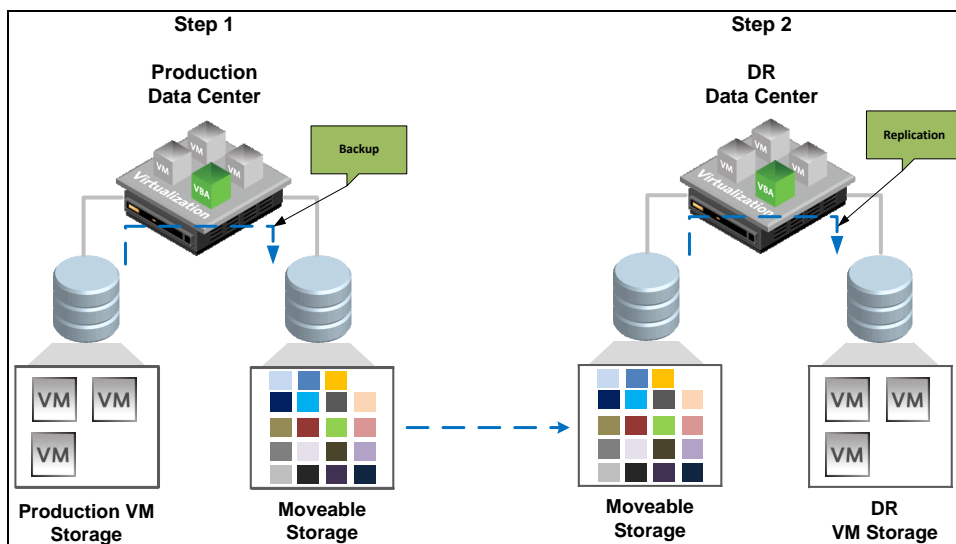
Seeding from Backup Data

Step 1

At the Production Data Center, backup your VMs to some form of “Moveable” Storage. This Moveable Storage could be a portable NAS device, or some form of a USB device, configured for either CIFS or NFS.

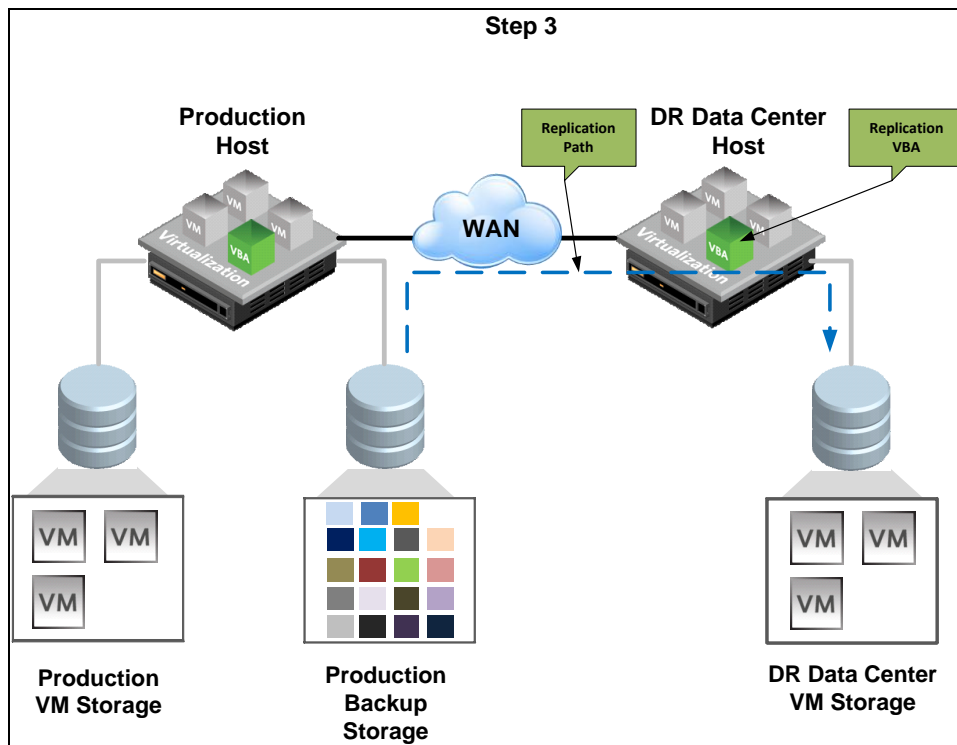
Step 2

Relocate the Moveable Storage to the DR Data Center. At the DR Data Center, proceed to use PHD Replication to replicate the VMs from the Moveable storage to the DR VM Storage. The first pass of the replication will be a “Full” replication of all of the VMs.



Step 3

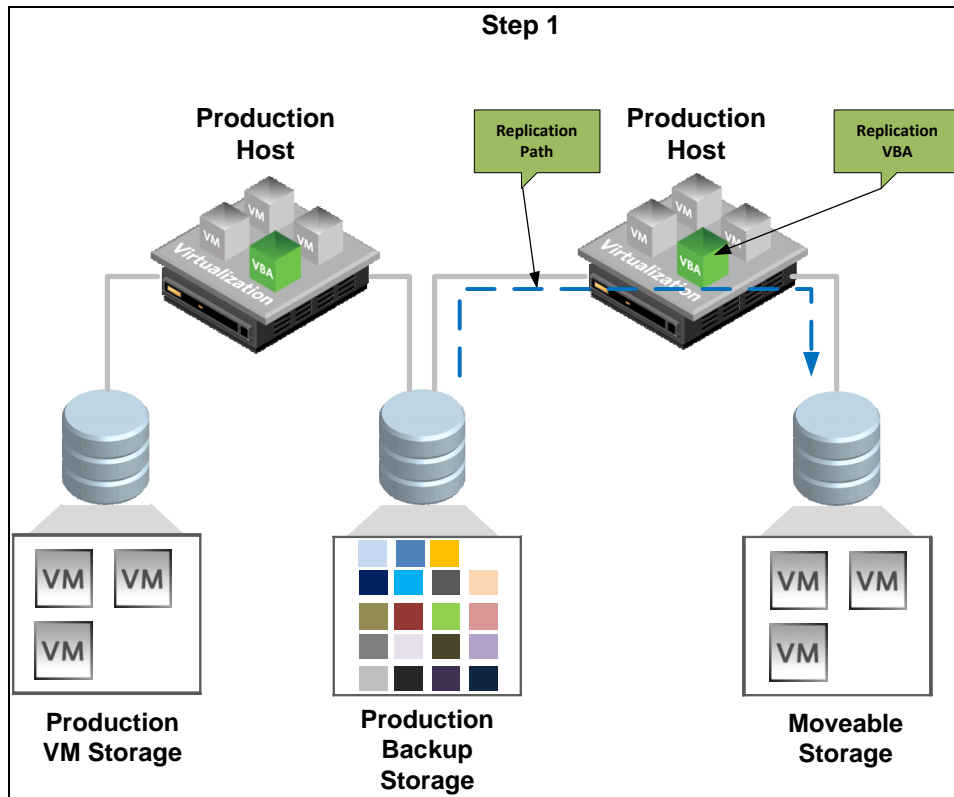
You can now disconnect the Moveable Storage and move it back to the Production Data Center, where it will become the Production Backup Storage. Re-configure the PHD Replication VBA to Replicate the VMs from the Production Backup Storage to the DR Data Center VM Storage. At this point only changed data will be replicated to the DR Data Center.



Seeding from Replication Data

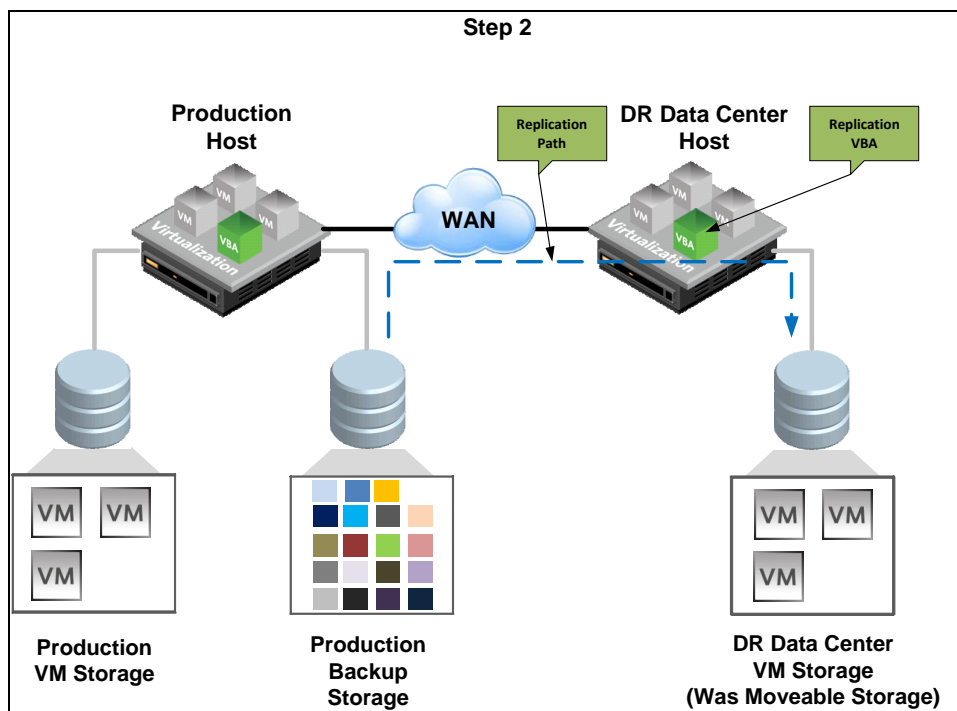
Step 1

At the Production Data Center, replicate your VMs to some form of “Moveable” Storage. This Moveable Storage could be a portable NAS device and will need to be configured as either a VMware Datastore or a XenServer SR.



Step 2

Relocate the Moveable Storage to the DR Data Center where this will become your DR VM Storage. Configure the PHD Replication VBA to Replicate the VMs from the Production Backup Storage to the DR Data Center VM Storage. At this point only changed data will be replicated to the DR Data Center.



Failback Procedures

The following is a description of several options for restoring Virtual Machines from the DR site back to the original Production site, after a DR event. There are three options being described for performing this restoration. The physical network topology and set of circumstances for recovery, will most likely govern which scenario best fits your needs.

1. Restoration over the Network
2. Move the DR Backup Storage
3. Storage Based Replication

The following examples will assume the following:

- You were replicating to the DR Data Center with PHD Virtual before the outage occurred at the Production Data Center.
- You have been backing up the Replica VMs with PHD backup software.
- You have been running the Production Virtual Environment out of the DR Data Center for the duration of the outage.
- You are ready to restore the Production Virtual Environment back to the Production Data Center, from the DR Data Center.

Scenario #1 – Restoration over the Network

When it is time to failback to the Production Data Center, the suggested steps are as follows:

1. Shut down the DR Virtual Infrastructure at the DR Data Center, so data will not be changing.
2. Perform one last PHD Virtual backup of the Replica VMs at the DR Data Center.
3. From the Production Data Center and using the PHD Backup Storage at the DR Data Center, perform a PHD Mass Restore of all necessary virtual machines into the Production Virtual Infrastructure.
4. Turn on the VMs in the Production Data Center.

Pros

- Best suited for environments that have either a large link between the Production and DR Data Centers, or a small amount of production data.
- One step Restoration.

Cons

- You need a fairly large link between the Production and DR Data Centers.
- Could take a very long time with a small link.

Scenario #2 – Moving DR Backup Storage

When it is time to failback to the Production Data Center, the suggested steps are as follows:

1. Shut down the DR Virtual Infrastructure at the DR Data Center, so data will not be changing.
2. Perform one last PHD Virtual backup of the Replica VMs at the DR Data Center.
3. Move the PHD backup storage from the DR Data Center, to the Production Data Center.
4. Perform a PHD Mass Restore of all necessary virtual machines into the Production Virtual Infrastructure.
5. Turn on the VMs in the Production Data Center.

Pros

- After moving the storage, will possibly provide the fastest restore option.

Cons

- Have to physically move the DR Backup Storage.

Scenario #3 – Storage Based Replication

When it is time to failback to the Production Data Center, the suggested steps are as follows:

1. Shut down the DR Virtual Infrastructure at the DR Data Center, so data will not be changing.
2. Perform one last PHD Virtual backup of the Replica VMs at the DR Data Center.
3. At the DR Data Center, perform final storage based Replication of the PHD Backup Storage to the Production Data Center.
4. Perform a PHD Mass Restore of all necessary virtual machines into the Production Virtual Infrastructure.
5. Turn on the VMs in the Production Data Center.

Pros

- If using storage based replication this whole process can be automatically scheduled.

Cons

- Need to have storage based replication.
- Can be complex.

About PHD Virtual

As the pioneer of virtual backup appliances (VBAs), PHD Virtual Technologies has been transforming data protection in virtual IT environments since 2006. Its award-winning data protection solution for virtual infrastructures, PHD Virtual Backup (formerly, esXpress), is used today by more than 2000 enterprises worldwide to achieve unlimited dynamic growth, high availability, no single point of failure and scalable performance. PHD Virtual is committed to helping our customers and provides free, easy-to-use virtualization utilities to assist with the administration and management of virtualized environments.

For more information contact a PHD Representative today!

Call: 1.866.710.1882 Email: sales@phdvirtual.com

