

Best Practices for Architecting Storage in Virtualized Environments

Leverage Advances in Storage Technology to Accelerate Performance, Simplify Management, and Save Money in Your Virtual Server Environment

Overview

Virtualization enables dramatic increases in utilization and efficiency. However, without proper planning, virtualization can present major challenges, which include greater complexity, higher management costs, and time-consuming backups. One key consideration in launching a virtualization project is choosing the correct storage platform. Unfortunately, today's legacy storage platforms were designed before virtualization, high-density disk drives, and flash memory were introduced.

Understanding the latest developments in storage technology is crucial when planning a virtual server storage architecture to reduce your long-term total cost of ownership (TCO) for primary and backup storage, as well as network-related expenses for backup and disaster recovery. Choosing the wrong storage architecture can dramatically impair your ability to provide application resources and recoverability.

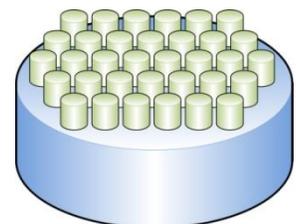
This documentation will help you understand how recent storage advancements can help you gain the maximum benefits from your virtual server environment.

Virtual Storage Layout

Planning a virtual server environment should begin with the selection of the storage layer, which would directly impact the functionality of the entire virtual infrastructure. Several options are available, each with pros and cons, and thus the best choice is not always clear. The storage platform decision requires careful consideration, because the shared nature of the virtual infrastructure can increase implementation times.

All-in-One Storage Layout

Many virtual server storage implementations begin with the traditional method of creating and then formatting a large storage volume with a native VMFS/NTFS file system, or using the NFS shared storage and placing all virtual disks on that volume. This approach eliminates the need to re-configure the storage infrastructure when provisioning new virtual servers, by simply informing virtual administrators to store new VMs and their data virtual

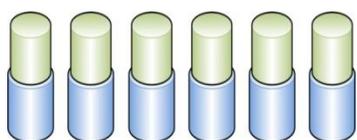


disks on a particular array. However, this approach is susceptible to corruption that can impact your whole environment.

Protection of the shared file system model is all-or-nothing, which dramatically increases the storage requirements for snapshots and the WAN bandwidth for off-site disaster recovery. Thus, though this approach could decrease management time, you may be forced to pay for it many times over in storage waste and monthly network bandwidth costs.

Storage Layout Matched to Application Needs

For applications with divergent performance and recoverability needs, the best solution is to separate virtual disks into their own storage volumes. This approach provides greater

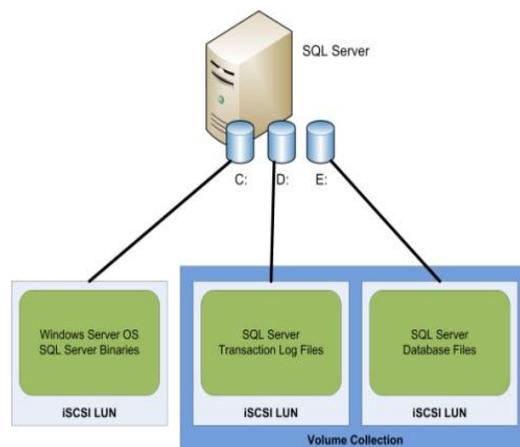


management flexibility. A storage system that offers built-in systems could dramatically simplify the provisioning of the initial volumes, and provide the highest control over the performance and data protection levels of different applications. Separating

volumes also protects against file system corruption and allows you to create protection schedules based on each volume's unique service level agreements.

For example, you might wish to frequently snapshot and replicate a critical database that constantly changes, while using daily snapshots for another server's operating system virtual disk that rarely changes. Individual storage management also provides the ability to monitor performance and usage on a per-application server basis. This approach is advantageous because every server has its own unique requirements that may change over time.

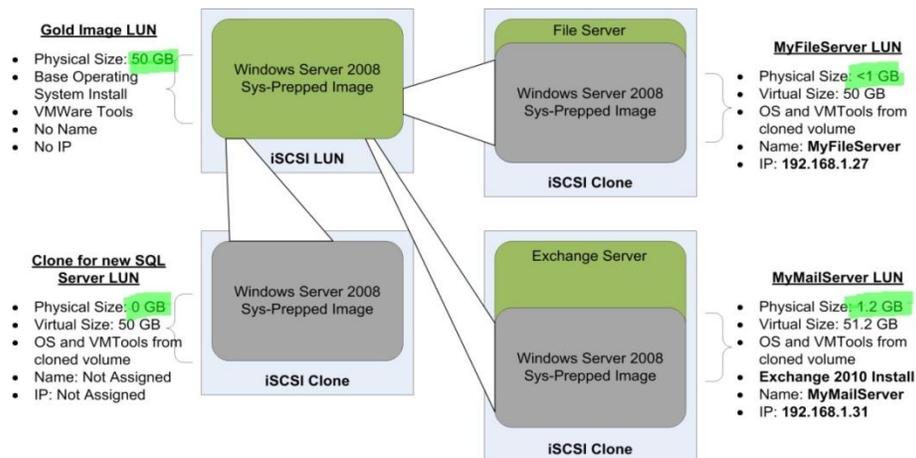
Another key consideration is to separate the operating system and data virtual disks. While this is a legacy best practice from the physical server for proper data protection, highly different I/O requirements exist for OS, database, and transaction log volumes. Matching the application sizing and performance characteristics can gain performance savings in Input/Output Operations Per Second (IOPS). This results in greater virtual machine consolidation and extension of the storage lifespan.



Zero-Copy Cloning

Advanced storage systems can greatly reduce the footprint of commonly used volumes (such as operating systems) by creating zero-copy clones of storage with only the differences between base images. For example, zero-copy cloning allows you to thin-provision a base image volume of 50 GB, and then share those provisioned bytes with clones that only consume space for configuration

changes, such as those on virtual server names as well as unique IP addresses and applications. This system dramatically decreases the time and storage required to provision new systems, as well as maximizes the network bandwidth while copying OS templates.

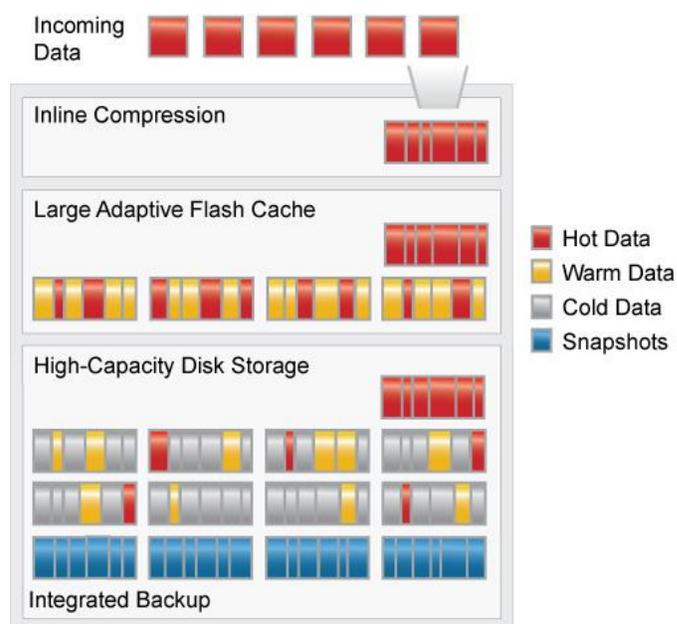


Compression and Block Size

Compression has long been a feature of OS file systems, but is rarely implemented because of the application server that CPU cycles require. Many legacy storage vendors include compression technology but suffer penalties because their internal file systems were not designed to hold blocks that compress at variable sizes. Thus, by using fixed-width blocks, legacy storage vendors cannot efficiently store compressed blocks without the risk of over-running boundaries and forcing additional IOPS to satisfy storage reads and writes. Unused block spaces are thereby wasted. Moreover, file system fragmentation occurs with a higher penalty on storage performance in seeking open holes where a compressed block could fit. This dramatically affects the useful lifespan of the storage investment.

Modern storage platforms take advantage of new technologies, such as multi-core CPUs, to provide compression within their platforms and off-load processing cycles from application servers. In addition, modern file systems efficiently use variable-length blocks for compressed storage and write data in block stripes with parity to avoid fragmentation.

Several advanced file systems write in stripes of variable-length blocks to reduce the amount of random physical I/O. However, certain write stripes do not maximize the full length of the disk group/array, creating holes in the file system that degenerate to performing



small random I/O. Write striping in full array, variable-length blocks maximizes the efficiency of hard disks to minimize movement of physical magnetic heads that limit performance.

File systems that seamlessly integrate flash disks as a caching mechanism can greatly improve read performance by altogether avoiding the disk for “hot” data. Consolidation occurs through a sweeping process that runs during idle time to efficiently move variable-length block spaces caused by deleted data, such as those from expired snapshots.

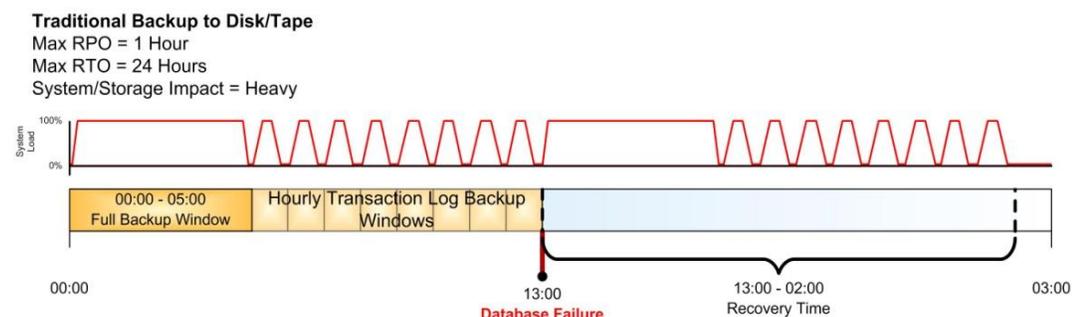
Backup, Restore, and Disaster Recovery Considerations

Storage architecture decisions for the virtual server environment similarly impact backup capabilities. Three methods are used to protect the virtual environment, depending on where the backup process occurs: guest-based, host-based, or storage-based.

Guest-based backup works fundamentally similar to a traditional physical server backup. This method requires installation and an update of one or more backup agents on the server. Agents coordinate with the server and applications to place them into a quiescent state that ensures data integrity. For most data sets, quiesced data are of utmost importance during the backup process to ensure that changes are not made by users that would invalidate data that has not been backed up yet.

For example, while backing up customer order data, a customer address changes. The backup completes, but the corruption on the application data is only found after a system failure when that backup is restored. The customer order record now references the wrong customer address. This is the job of the backup agent: to ensure that data cannot be modified during the backup process.

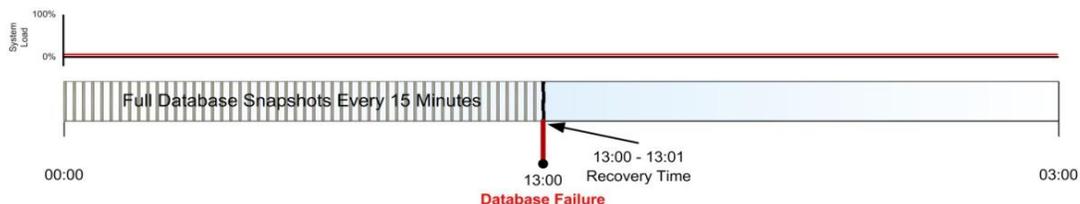
Backup agents have obtained a negative reputation due to the excessive management involved in installation and updates over time. An agent is required to ensure the data integrity of all but the most basic data types. Backup software vendors leave sufficient gaps in their offerings to allow for the rise of host-based backup solutions that integrate with the host hypervisor, rather than each individual virtual machine. Periodically, the backup process requests the hypervisor to place virtual machines into a quiesced state to allow backup of the underlying virtual disk drives.



Host-based backup solutions are not natively application-aware. They only provide crash consistency, not application consistency. Microsoft Hyper-V does provide an integration pass-

through between the host and its virtual guests, passing the request to quiesce all the way to the running application. However, this method is only available on Microsoft Hyper-V; thus, without the help of agents running in the guest, VMware backups can only be considered crash-consistent and not application-consistent. Without application consistency, data cannot be guaranteed to be in a safe state.

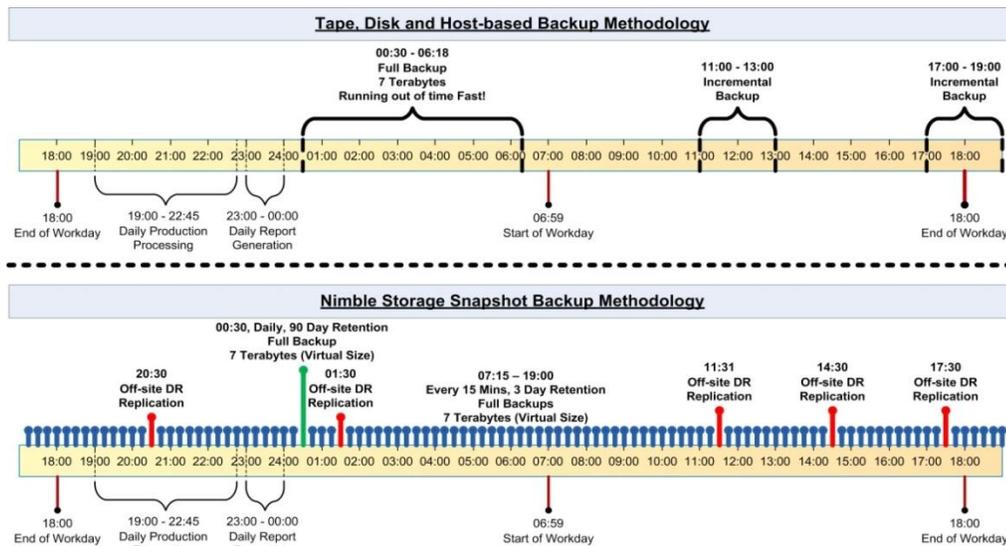
Nimble Storage Snapshot Backup
 Max RPO = 15 Minutes
 Max RTO = 1 Minute
 System/Storage Impact = Minimal



Traditional storage-based backups likewise lack application awareness. Thus, though agent management is a necessary function to guarantee the data integrity of backups – and vendors continue to improve at self-updating their software – the agent requires installation in a base image and automatic deployment through storage cloning.

The ideal backup solution eliminates guest and host overhead by offloading the backup process to storage without sacrificing application awareness. Modern storage platforms provide thin agents that properly coordinate application quiescence with negligible overhead to enable efficient storage-based backup that eliminates the possibility of application data corruption.

Protection Frequency (RPO)



Another consideration is the frequency of point-in-time backups. Ideally, backups should be performed at very short intervals to avoid data loss. One challenge associated with host- and guest-based backups is the application server’s CPU cycles and I/O required to identify which

changed data requires backup. This step creates a huge impact that takes resources away from the application and any other virtual servers running on that host during a backup.

A penalty is also associated with both of these backup solutions, requiring time and network bandwidth to copy the changed data to the backup media. This challenge is overcome by using storage-based snapshot backups. However, as previously mentioned, a backup agent is necessary in the guest to ensure data integrity. Thus, storage-based backups combined with application-aware agents provide the best solution that quiesces I/O without affecting the production application by offloading the backup process to the storage layer. This approach enables more frequent backups compared with those of either host- or guest-based backup solutions.

Snapshot Methodology Considerations

Historically, storage snapshots were not suitable for point-in-time backups at short intervals or retention beyond a week. Snapshots caused too much interruption and consumed large amounts of disk space. These legacy copies are collectively referred to as copy-on-write (COW) snapshots, because a data block that is about to change is copied to another location on the storage to make room for the new block. This is a very expensive process in terms of performance, because it adds a read-and-write for every application write.

Despite advances in snapshot technology, new techniques cannot be retrofitted onto legacy storage architectures that are tightly intertwined within the underlying storage file system. Newer storage vendors have the benefit of creating their platforms using recent advancements in storage technology, such as flash memory and redirect-on-write (ROW) snapshots. Highly efficient, ROW snapshots simply update the volume metadata pointers to reflect the changes and thus do not require an additional read and write IO for each application write.

New storage platforms can use high-performance persistent memory to cache file system metadata. This provides the ability to take snapshots at much more frequent intervals (measured in minutes) and maintain historical point-in-time snapshot backups for longer retention times (measured in months). Combining compression and block-sharing with snapshot features further reduces the amount of storage footprint necessary to meet recovery point objectives (RPOs).

Restoration and Disaster Recovery

Virtual server backup success is measured by the ability to properly restore data to a certain point in time. Though tape media failure remains a possibility that could affect RPOs, this is largely mitigated by the use of disk-based backup solutions. However, both methods require network bandwidth and time to restore data to the primary storage before applications can restart.

Large datasets routinely take hours or even days, which when compounded by a busy network can extremely stretch recovery time objectives. This problem created the need for a new generation of

storage products that use the paradigm of converged storage — joining primary and backup data within the same architecture, eliminating network copies for both backup and restoration.

Disaster recovery has long been an expensive proposition for IT. Replication bandwidth and its TCO have now surpassed server costs as the chief blocking point. Traditional storage technologies are highly inefficient at performing replication, which is a functionality that was added as an afterthought and therefore a hacked retrofit on primary storage.

One of the challenges hindering the efficiency of legacy storage replication is the lack of application awareness that forces storage products to maintain application writes in block containers much larger than the actual changes in application data. This method forces the entire storage block to be snapshot and replicated, because isolating the changes at a more granular level is impossible. For example, a SQL server writes an 8 KB change to a 1 TB database. Most storage architectures will snapshot and replicate a minimum of a 64 KB block, requiring eight times more bandwidth than the actual data change. For the same amount of database change, some storage platforms even replicate up to a 256 KB block, requiring 64 times more bandwidth.

Conclusion

New storage technologies — such as application awareness, variable block-sizing, and compression — deliver significant improvements over legacy storage architectures. Combining flash memory with low-cost, high-capacity drives eliminates the need for expensive, high-RPM drives for primary storage and separate disk-based backup solutions.

Converged storage provides primary and backup storage, as well as application-aware efficiency for disaster recovery into the same architecture. This solution improves performance and eliminates the complexity of managing separate devices, dramatically reducing the TCO for disaster recovery implementations. Finally, the acquisition costs of a converged storage solution can be offset by the significant cost savings of reduced network bandwidths as well as the elimination of most existing backup software and hardware solutions.



Learn how Nimble Storage's converged iSCSI storage, backup, and disaster recovery solutions can help you save money and simplify data management in your virtual server environment. Refer to these informative resources:

- [The Nimble Storage Website](#)
- [Nimble Storage Three-Minute Overview](#)
- [Nimble Storage Solutions for VMware](#)
- [CRN Test Lab Review of Nimble Storage](#)



Nimble Storage, Inc.

2740 Zanker Road, San Jose, CA 95134

Tel: 408-432-9600; 877-364-6253 | www.nimblestorage.com | info@nimblestorage.com

© 2012 Nimble Storage, Inc. All rights reserved. CASL is a trademark of Nimble Storage Inc. BPG-SVE- 0812