



NETAPP WHITE PAPER

Helping DBAs Become More Efficient NetApp Efficiency and Manageability Advantages

Steven Schuettinger, Generosa Litton, NetApp
July 2009 | WP-7079-0709

EXECUTIVE SUMMARY

Database administrators face significant and growing data management hurdles in a variety of areas, including provisioning, resource optimization, change assurance, dev/test, backup/recovery, business continuance, disaster recovery, and server virtualization. To succeed, you need to decrease this management complexity while at the same time increasing efficiency and reducing costs. NetApp offers a full range of manageability and data efficiency technologies to help you address database management problems, including technologies for thin provisioning, thin cloning, deduplication, replication, and transparent data movement. This white paper describes how you can apply these solutions to boost efficiency and help reduce the cost associated with your database environment.

TABLE OF CONTENTS

1	INTRODUCTION	3
2	PROVISIONING AND RESOURCE OPTIMIZATION	3
	IMPROVING UTILIZATION AND PERFORMANCE WITH THIN PROVISIONING	4
	MATCHING DATA TO STORAGE	4
3	SIMPLIFIED CHANGE ASSURANCE	6
4	ACCELERATING APPLICATION DEVELOPMENT AND TEST	6
	CREATE MULTIPLE DATABASE COPIES USING MINIMAL SPACE	7
	PROTECTING SENSITIVE USER DATA: DATA MASKING	7
5	BACKUP AND RECOVERY	8
	BACKING UP WITH SNAPMANAGER	8
	RESTORING WITH SNAPMANAGER	9
	CLONING WITH SNAPMANAGER	9
	DEDUPLICATION	9
6	DISASTER RECOVERY AND BUSINESS CONTINUANCE	10
	NETAPP METROCLUSTER AND SNAPMIRROR	10
7	SERVER VIRTUALIZATION	10
	DATA MOTION WITH MULTISTORE	11
8	CONCLUSION	12

1 INTRODUCTION

The job of the database administrator (DBA) never gets easier. Rapid data growth, increasing requirements for data availability, and new demands for database resources seem to overshadow any efficiency gains created by new management tools.

A number of areas continue to act as pain points for many DBAs:

- **Provisioning and resource optimization.** Active databases require almost daily management.
- **Change assurance.** Database patches and upgrades might introduce unacceptable risks.
- **Application development and test.** Test data sets have become large and difficult to manage.
- **Backup and recovery.** These operations are often slow and unreliable and require significant attention.
- **Disaster recovery and business continuance.** Available options are complex, expensive, and slow.
- **Server virtualization.** This new paradigm changes the way databases must be managed.

NetApp understands these problems and offers a full range of manageability and data efficiency technologies that can help you address database management problems. In particular, technologies that help you manage and use storage more efficiently have become an important focus for NetApp. Table 1 shows the key NetApp® storage efficiency technologies that can positively impact database management.

Table 1) NetApp storage efficiency technologies.

NetApp Efficiency Technology	Benefit
Thin provisioning	Storage no longer has to be allocated to a volume up front. Storage is treated as a shared resource, and capacity is consumed only as it is needed.
Thin cloning	Clone complete data sets in seconds without full data copies using NetApp FlexClone®. Storage space is only consumed incrementally as changes occur.
Snapshot™	NetApp Snapshot technology delivers space savings over competing products. ¹ Create up to 255 Snapshot copies without excessive space penalties and virtually no performance penalty.
Deduplication	NetApp deduplication identifies and eliminates redundancy at the block level in both primary and secondary storage.
Thin replication	SnapMirror® and SnapVault® software perform only incremental block transfers—thin transfers—accelerating replication and conserving bandwidth.

This white paper explores how you can apply these and other NetApp technologies to alleviate a variety of database pain points, streamline data-related activities, improve operating efficiency, and help DBAs succeed.

2 PROVISIONING AND RESOURCE OPTIMIZATION

An active database might require daily—if not subdaily—management and maintenance to accommodate growth and optimize operation. Space consumption and performance must be monitored closely, and frequent adjustments are common.

Given the current resource constraints that affect most IT operations, there are two imperatives with regard to database provisioning:

- Simplify provisioning and minimize the amount of manual allocation adjustments that must be made.
- Reduce the overall cost of storing large databases.

These problems can be addressed through the use of thin provisioning and information lifecycle management.

¹ <http://media.netapp.com/documents/ar1038.pdf>

IMPROVING UTILIZATION AND PERFORMANCE WITH THIN PROVISIONING

In the traditional approach to provisioning, you allocate the space you think you'll need for a particular application up front. Unfortunately, because it is often impossible to estimate space requirements with any accuracy, you inevitably end up with some storage volumes that are overprovisioned, while others are chronically short of space, as illustrated in Figure 1.

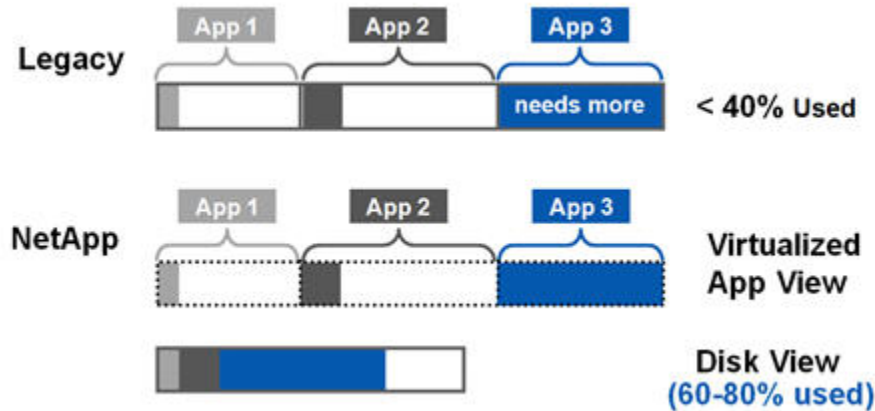


Figure 1) Legacy approach to provisioning versus thin provisioning.

With NetApp thin provisioning, you still estimate how much space you think you'll need up front, but that storage space is drawn from a single pool of shared free space and only consumed as it is actually needed. The results are easier capacity management and much higher utilization rates for disk resources. NetApp users typically report disk utilization rates of 60% to 80% versus 40% or less using traditional disk volumes with traditional provisioning². With NetApp, you are also free to grow or shrink a volume or LUN as necessary without disruption and without having to worry about the possible performance impact.

Performance is also addressed by the NetApp approach to provisioning. The performance of a storage volume is directly related to the number of disks that make up the volume: the more disk spindles, the greater the aggregate I/O. With traditional provisioning, the number of underlying disk spindles is tied to the size of the volume: small volumes typically have few underlying spindles; large volumes have many. As a result, performance is often a problem, and volume size must be increased to raise performance, even in situations where the extra capacity is not needed.

NetApp flexible volume (FlexVol®) technology virtualizes space allocation so that you no longer have to worry about the underlying disks. All volumes regardless of size are immediately spread across the maximum number of disk spindles for optimum performance with no tuning required. More flexible provisioning and higher space utilization are the result.

The net result of the NetApp approach is that you have to tune your storage configuration less often; you achieve higher utilization and better performance with much less effort.

MATCHING DATA TO STORAGE

With most databases, data is accessed less often as it ages. You can take advantage of this fact to save money by migrating data to less expensive storage over time, as illustrated in Figure 2.

² <http://media.netapp.com/documents/ar1038.pdf>

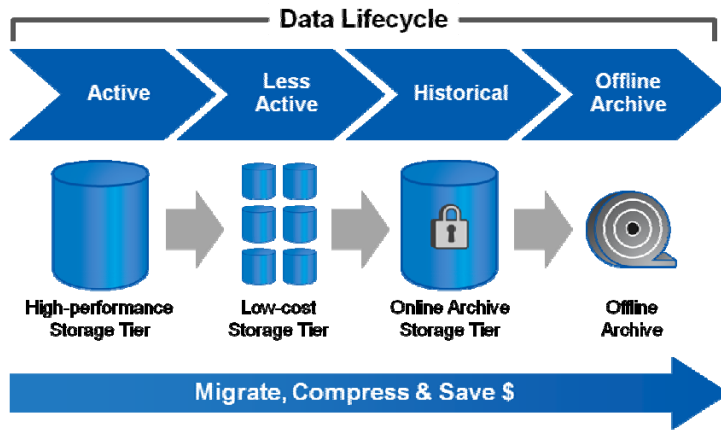


Figure 2) Information lifecycle management can save storage costs.

The only problem with this approach is that with most storage vendors each storage tier is supplied by a separate storage system that is incompatible with those in the other tiers. The results can be more storage systems, multiple management interfaces, incompatible data protection methods, and more coordination required between DBAs and storage administrators. The management costs that result from this increased complexity can outweigh the direct savings on storage in many cases.

NetApp addresses this problem with its unified storage architecture. All online storage tiers—high-performance storage, low-cost storage, and online archives—can reside on the same storage system. Fibre Channel disks provide the high-performance tier, SATA provides low-cost storage, and NetApp SnapLock® software facilitates the permanence, accuracy, integrity, and security of data by enabling database files to be both unalterable and rapidly accessible online for long periods of time. With NetApp ILM for Oracle®, you can use Oracle ILM assistant to create SnapLock volumes on low-cost storage (as illustrated below) or on high-performance storage to accelerate discovery.

Combining all storage tiers on a single storage platform can result in less cost and complexity and dramatically simplified management for both DBAs and storage administrators.

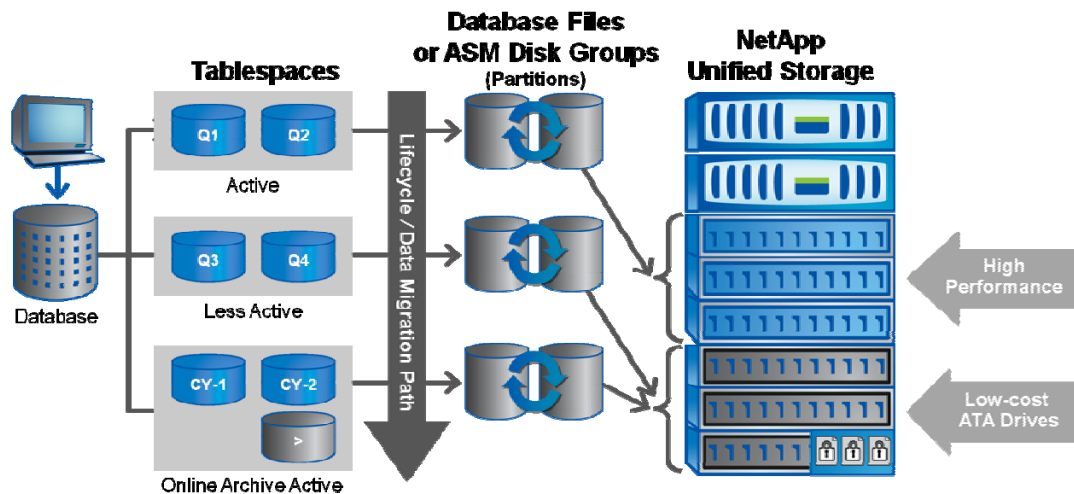


Figure 3) Mapping multiple tiers of storage to a single NetApp unified storage system can reduce storage costs and eliminate the complexity associated with managing tiers of storage across multiple, incompatible arrays.

3 SIMPLIFIED CHANGE ASSURANCE

Another task that DBAs must perform is the installation of database patches and upgrades. Unfortunately, it can be difficult to adequately assess the impact of such changes to your production database before applying them to production systems. The last thing you want is to have to back out a change from a running database in the middle of the day.

NetApp helps you avoid potential problems by simplifying the process of change assurance. With NetApp FlexClone technology you can create a consistent copy of your entire database environment nearly instantaneously, using only incremental storage, because additional storage space is only consumed as changes are made. Cloning tasks can often be performed with little or no support from a storage administrator.

By creating a FlexClone copy of your production database, you can perform both functional and automated regression testing of database patches and upgrades with your full production data set before you apply those changes to production systems.

NetApp Snapshot lets you protect yourself further by creating a point-in-time image of the production database before you apply the changes. Should a problem arise, you can revert to the saved state in a matter of seconds using NetApp SnapRestore®.

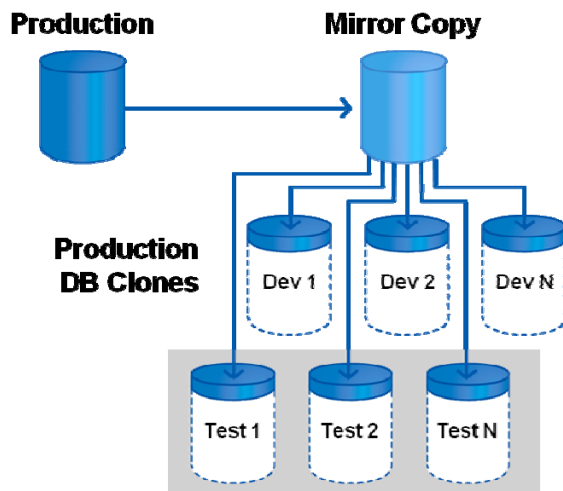


Figure 4) NetApp FlexClone lets you make many copies of a production database for change assurance, dev/test, or other purposes using a fraction of the space of other methods. You can also keep multiple historical database versions for problem triage or to serve as “gold images.”

The space efficiency of FlexClone means that you can also retain a library of historical versions of your database and associated applications to serve as “gold images.” Having such a resource on hand can make problem triage much simpler.

4 ACCELERATING APPLICATION DEVELOPMENT AND TEST

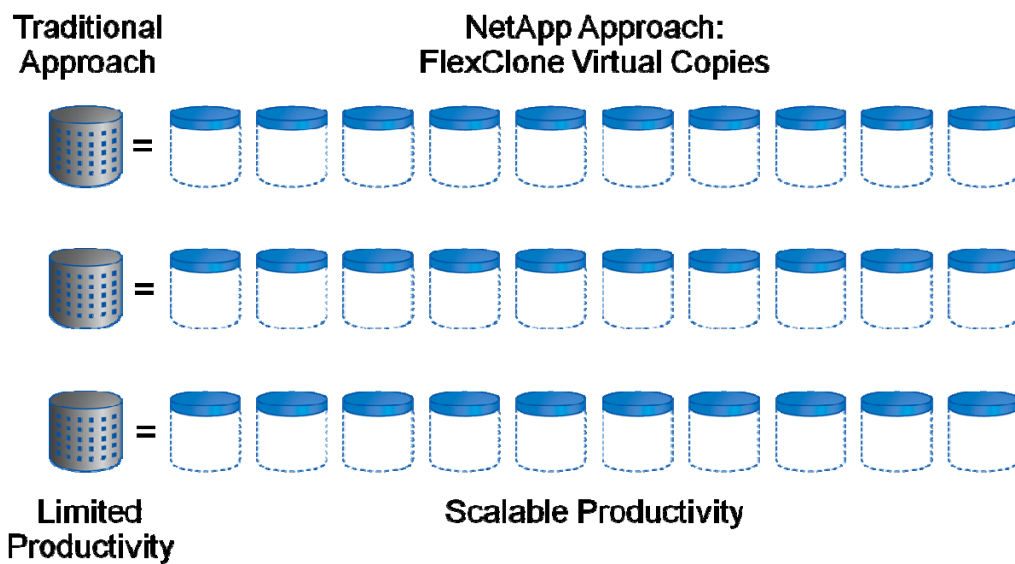
Application development and test can be a huge drain on both storage resources and DBA time. NetApp offers a unique approach that can help accelerate your development activities and reduce time to market while simplifying critical management tasks. Using FlexClone, you can provision virtual database copies for everyone that needs one and simplify the process for that sensitive user data to be protected.

CREATE MULTIPLE DATABASE COPIES USING MINIMAL SPACE

The traditional approach to cloning a database is a heavyweight activity. Since every copy requires 100% of the storage capacity of the original database, it places a major burden on storage resources. Getting necessary approvals, working with the storage admin to locate storage capacity, and copying the data also typically take a long time to accomplish. The only alternative to having a full clone, however, is to do testing with only a subset of your production data, and that can often lead to inconsistent results.

Because full database copies are an expensive resource, they are always in short supply and often not as current as you would like them to be. DBAs, developers, testers and line-of-business specialists all queue up waiting on limited resources.

NetApp FlexClone provides an alternative that can alleviate the resource contention that results from not having enough database copies. With FlexClone you can create multiple clones using minimal incremental storage space. For example, you can create 30 database clones for the same space as three database copies. As a result, everyone who needs a database copy can gain immediate access, so productivity can



increase.

Figure 5) NetApp FlexClone delivers 30 clones using the storage space required by three traditional copies. (Assumes that on average 10% of the data is updated in each copy.)

PROTECTING SENSITIVE USER DATA: DATA MASKING

An important issue with nonproduction database copies is the need to protect sensitive user data (such as social security numbers, phone numbers, addresses, and so on) contained within your production database. Regulations around the world, including, but not limited to, the Sarbanes Oxley Act of 2002 (SOX), and other privacy related regulations in the United States, and the European Union's Data Protection Initiative, now mandate the protection of confidential or sensitive information contained in corporate databases.

Many companies already have third-party data masking tools or homegrown scripts to accomplish the masking process. A difficulty arises in making sure that these tools have been appropriately applied to every database copy.

The use of NetApp Snapshot and FlexClone technologies can simplify this process through the creation of a "master clone". With data masking correctly applied, one can reduce the potential for errors and omissions..

Typically, you'll start with a mirrored copy of your production database that resides on NetApp storage. A NetApp Snapshot copy is made to capture a consistent, point-in-time copy of the database. Data masking is then performed on the Snapshot copy (this can be conveniently performed using the postclone scripting capability of SnapManager® as described in a later section), and the success of the masking process is verified. This then serves as the “master clone.” FlexClone can then be used to create as many clones of the master as required, enabling all copies to be masked correctly in a space efficient manner.

5 BACKUP AND RECOVERY

The prospect of having to perform a full database recovery can give even the most seasoned DBA nightmares. Backup and recovery operations are time consuming and expensive, and reliability is frequently an issue.

NetApp offers a full portfolio of manageability software that can significantly reduce the time you spend doing backups while speeding recovery and greatly increasing reliability. NetApp SnapManager for Oracle and SnapManager for SQL Server® provide database-consistent data management capabilities that automate and simplify backup and recovery activities. These tools integrate fully with NetApp Protection Manager so you can easily manage all your data protection needs. Protection Manager adds off-site vaulting capabilities and policy-based controls that allow storage administrators to delegate storage tasks to DBAs so that DBAs no longer have to wait for a storage admin to accomplish data protection tasks.

BACKING UP WITH SNAPMANAGER

To create consistent backups, the SnapManager® tools are designed to integrate with and coordinate their operations with the underlying database. The database is put into hot backup mode, a Snapshot copy is created, and the database is returned to normal operation. NetApp Snapshot technology creates a space-efficient, point-in-time copy of your data that can serve as an online backup on your primary storage. No additional storage space is consumed at the time a Snapshot copy is created (other than a minimum amount of metadata). Because the NetApp WAFL® file system never overwrites blocks on disk, as changes are made to the database after the Snapshot copy, the original data blocks are simply retained as part of the Snapshot copy instead of being freed. No work is required to retain these blocks, so many Snapshot copies can be kept on primary storage without affecting performance.

Other vendors offer technology that appears superficially similar to NetApp Snapshot, but these solutions either create full copies (space prohibitive) or use a copy-on-write (CoW) algorithm. With CoW, every time a block that is part of a “snapshot” is overwritten, the original block must first be copied. This extra work can result in a 20% to 60% performance penalty on file systems maintaining just a single “snapshot.” By comparison, you can keep up to 255 NetApp Snapshot copies with no performance impact.

Because the NetApp Snapshot backup process occurs in a matter of seconds, you can create many such Snapshot copies throughout the day for increased protection. You can also optionally vault Snapshot copies to a secondary location using the capabilities of NetApp SnapVault and NetApp Protection Manager. Each Snapshot copy can be verified on either the local system or a secondary storage system for integrity.

SnapManager for Oracle can optionally register each Snapshot backup with Oracle Recovery Manager (RMAN) to make sure not only that RMAN can take advantage of Snapshot backups, but also that a DBA can preserve the unique value of RMAN. This can be important for block-level recovery. Restores can be initiated from either RMAN or SnapManager for Oracle. For flexibility, SnapManager for Oracle provides the option of either full or partial backups.

In summary, the advantages of backup using NetApp SnapManager products include:

- Speed and storage efficiency
- Increased backup frequency
- No performance degradation
- Immediate access to backup data

- Automation eliminates manual errors
- Full integration with Oracle RMAN

RESTORING WITH SNAPMANAGER

Restoring data with SnapManager offers many of the same advantages as backup. Restores occur much more rapidly than with other common methods and can be performed either from local Snapshot copies or from vaulted Snapshot backups stored at a secondary site (if you have them).

When you schedule a recovery with SnapManager, it can automatically restore the appropriate Snapshot copy and play back database logs. The potential for operator error can be greatly reduced, and full recovery can be accomplished in a matter of minutes. You can also quickly perform logical recoveries of database objects such as schemas, tables, and so on. However, you must be aware that logical recoveries can result in relational inconsistencies and decide which recovery method is best for a given situation.

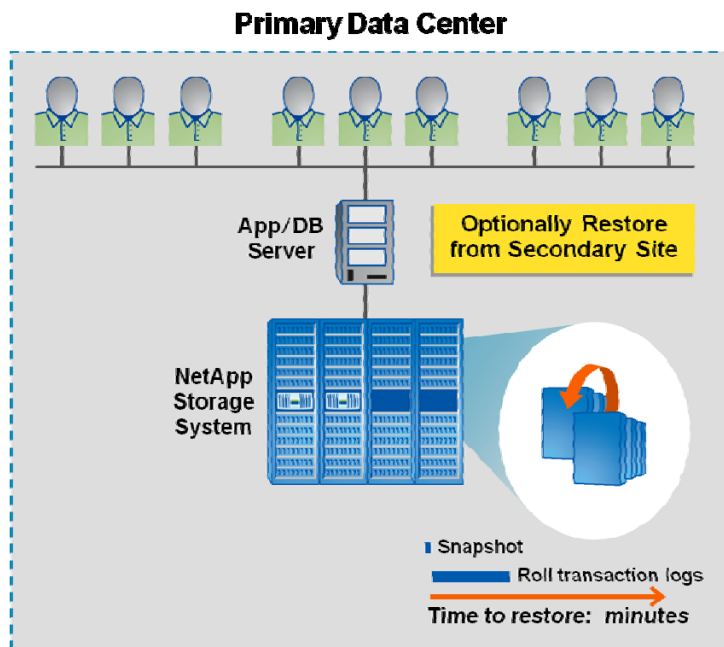


Figure 6) Restoring a database with NetApp SnapManager for Oracle or SnapManager for SQL Server.

CLONING WITH SNAPMANAGER

NetApp SnapManager also gives you the ability to quickly create fresh copies or clones of Oracle and Microsoft® SQL Server databases for dev/test, training, reporting, or other purposes. Clones can be created either on primary or on secondary storage (if you have vaulted Snapshot copies on secondary storage).

SnapManager includes the ability to perform post processing on clones using scripting. This allows you to automate the creation of clones that are fully data masked or to perform other modifications after a clone is created.

DEDUPLICATION

NetApp deduplication identifies and removes duplicate blocks on NetApp storage. Deduplication scans a storage volume looking for blocks with duplicate fingerprints. When potential duplicates are found, a complete comparison is performed, and if the blocks match, one of the duplicates is removed and replaced with a pointer to the other. NetApp deduplication works on both primary and backup volumes.

Because backup volumes typically contain large amounts of duplicate data resulting from repeated backup operations, you can typically achieve very high rates of deduplication on backup volumes.

6 DISASTER RECOVERY AND BUSINESS CONTINUANCE

Disaster recovery and business continuance (DR/BC) in database environments can be complicated. Appropriate solutions must be fast, flexible, and reliable, and they must also integrate with native database tools and capabilities.

There are several possible approaches for creating DR/BC infrastructures. Block shipping approaches transfer changed blocks from the primary database to a secondary. This is the approach used by replication tools such as NetApp SnapMirror. With log shipping, database change logs are transferred to the secondary site and replayed against a standby database. This is the approach used by native tools such as Oracle Snapshot Standby as well as many third-party tools. In practice, many NetApp customers choose to combine these approaches. Block shipping can minimize the amount of change logs that have to be replayed to complete a recovery and thus can decrease recovery time.

Important considerations when designing a disaster recovery or business continuance solution include:

- **What are your SLAs?** The final design of your DR infrastructure might depend on your availability, recovery point objective (RPO), recovery time objective (RTO), and other commitments.
- **How will you replicate data?** Synchronous replication can minimize the risk of downtime and data loss, but will only work effectively over limited distances, and it comes at a premium. Asynchronous replication is more economical and works over longer distances, but, because the DR site will not be completely up to date, logs will need to be replayed before restarting operations, slowing recovery.
- **Physical or logical standby?** Oracle provides the option of a physical standby, a block-by-block copy of the primary database, or a logical standby database recreated from shipped logs. Which option you choose will dictate how you architect disaster recovery/business continuance.

NETAPP METROCLUSTER AND SNAPMIRROR

NetApp offers two complementary solutions to help meet business continuance and disaster recovery needs for database environments. NetApp MetroCluster uses synchronous mirroring and automated failover to help achieve continuous data availability, while NetApp SnapMirror® offers asynchronous replication to span longer distances for disaster recovery.

NetApp MetroCluster combines synchronous mirroring with clustering at the storage level to create a simple yet highly resilient storage environment for critical database applications. MetroCluster is a self-contained solution operating at the storage level, it is able to transparently recover from failures with no disruption to business-critical applications. Automatic recovery from any failure within your data center eliminates the disruption that can sometimes result during failback operations to restore normal operations. You can deploy MetroCluster within a single data center, across a campus, or at remote sites up to 100 kilometers away.

For full protection against regional disasters, you might need a solution that can go beyond the metropolitan distances spanned by MetroCluster. NetApp SnapMirror software uses an asynchronous mirroring methodology that can span huge geographic distances. SnapMirror builds on the NetApp Snapshot technology to perform thin replication—efficiently identifying and replicating only changed blocks.

The most critical database applications might require a combination of MetroCluster for continuous availability and SnapMirror for full disaster recovery.

7 SERVER VIRTUALIZATION

Server virtualization technologies such as Microsoft Hyper-V™, VMware® ESX Server, Citrix XenServer, and Oracle VM offer significant features that are attractive in many database environments, including virtual machine migration and disaster recovery. If you use or plan to use these solutions as part of your database infrastructure, they will significantly affect your approach to data management.

NetApp offers some unique advantages for virtual environments. Although you might have a significant number of virtual servers in your environment, it's likely that you still have physical servers as well and will need to support both as you continue to evolve over time to greater virtualization. NetApp can provide a single pool of storage that meets the needs of both physical and virtual servers. This pool can span metropolitan distances using MetroCluster and longer distances using SnapMirror. Because NetApp's unified storage architecture supports Fibre Channel, iSCSI, and NAS protocols, this single storage pool can simultaneously meet the storage needs of any physical or virtual environment.

From a BC/DR perspective, virtual environments offer features to simplify restarting database services from a secondary site. For instance, VMware HA makes it simple to restart failed virtual machines on alternate servers in the same or a different location. Providing a storage pool that spans location with synchronous mirroring or replication facilitates this process.

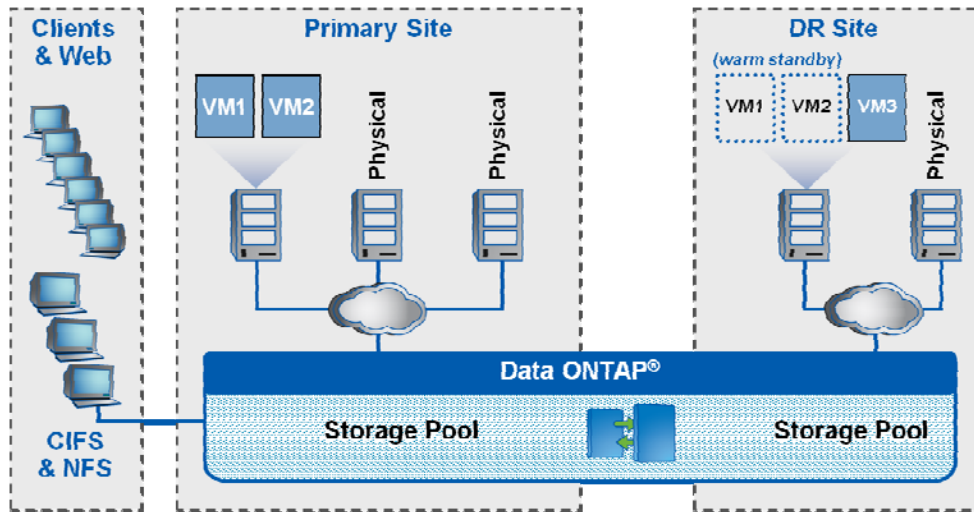


Figure 7) NetApp unified storage supporting an environment with both physical and virtual servers.

DATA MOTION WITH MULTISTORE

NetApp MultiStore® software provides multidomain virtualization that can also be used to match the capabilities of server virtualization solutions. The functional unit of MultiStore is vFiler™, a contained virtual storage system. Multiple vFiler instances can run on a single NetApp storage system. You can assign each database application its own vFiler instance such that the application can only see the storage contained on that vFiler instance. This feature also serves to compartmentalize storage in situations requiring high levels of security.

Any vFiler instance can be transparently migrated from one storage system to another using a Data Motion operation that is analogous to the migration of virtual machines using VMware VMotion®, XenServer XenMotion, or Hyper-V Quick Migration. Combining the ability to migrate both virtual machines and storage at will you can create a highly flexible database environment that requires little or no planned downtime and is highly resilient to failures.

8 CONCLUSION

New technologies, increasing business demands, and rapid data growth all combine to make a DBA's job harder. With a unified storage architecture and some unique storage efficiency technologies, NetApp can help simplify your database environment for easier management, rationalize your approach to storage, and help reduce both capital expenses and management costs. NetApp storage is fast, flexible, reliable, and NetApp tools can integrate closely with your database environment.

Whether your pain point is storage provisioning, making sure patches and upgrades go smoothly, dev/test support, data protection, business continuity, or addressing the unique needs of virtual servers, NetApp solutions are simple and efficient, so they can help solve your problems. As a result, your database environment becomes more productive and simpler to manage.

NetApp provides no representations or warranties regarding the accuracy, reliability or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.