Data ONTAP™ 7G—The Ideal Platform for Database Applications

FlexVol™ and FlexClone™ for Database Workloads

Network Appliance Inc. | Blaine R. McFadden | December 2004 | TR 3373

Abstract

Flexible volumes, introduced in Data ONTAP 7G, can provide numerous benefits to an organization through their ability to distribute I/O operations across currently deployed disks in a more efficient manner. This allows the organization the flexibility to provision storage on an as-needed basis. In addition, flexible volumes allow the organization's IT infrastructure to grow and more effectively control data and utilize our Snapshot™ technology in a more granular fashion without having to worry about the size of storage needed.

Network Appliance Inc.

Table of Contents

1. Executive Summary

As both the demand for storage and the complexity of storage solution implementations increase, IT organizations must manage ever-larger and more flexible systems, while simultaneously reducing the management costs of storage and increasing the flexibility of storage deployment. One way to architect a storage solution to allow for increased manageability and flexibility is to use Network Appliance™ Data ONTAP 7G. New features in Data ONTAP 7G, including FlexVol and FlexClone, greatly increase storage utilization, manageability, and flexibility in production, development, and mixed workload database environments. This paper shows the benefits of Data ONTAP 7G in these environments, provides examples of how to move database workloads to Data ONTAP 7G while maintaining current workload throughput, and shows areas that provide the largest throughput gains.

2. Introduction

The purpose of this paper is to illustrate how flexible volumes can be implemented in a traditional mixed database workload environment. This environment can include OLTP (Online Transaction Processing) database transactions, DSS (decision support) database workloads, and both Windows® file sharing (CIFS) and UNIX® file sharing (NFS) type transactions. Data ONTAP 7G will allow the IT organization to provision, grow, and clone storage on an as-needed basis while maintaining separate Snapshot copies.

Customers need the ability to move data sets around in real time without affecting their existing production environment. Customers want the ability to resize storage allocation depending on available resources and needs. In addition, it is important to be able to balance the load in these environments without depending on adding additional physical storage.

Storage allocation and allotment to different applications can be complex, and many customers have a challenging time doing these things effectively. Accurately forecasting application storage requirements is becoming more difficult. Hence, inevitably the initial sizing of the application needs will likely not be a perfect fit, and you'll need to adjust the application storage needs as the business evolves. Without Data ONTAP 7G, these tasks are not simple and easy to do in today's environments. Examples of some of the common questions a storage administrator may have for storage allocation are:

> Do we need one large volume or four smaller volumes?
> Will a smaller four-disk volume provide the type of IOPS necessary to handle my load?
> Will we need to increase the number of disk spindles, although we only need 30GB of storage?
> How will we use the remaining 150GB of storage with negligible effect on performance?
> How can we better maximize our available spindles for all our workloads?
> How can we have the benefit of volume-level Snapshot while using qtrees?

A common practice in enterprise environments is to create extremely large volumes and use qtrees as the basic units of storage provisioning. For example, in a database environment, different qtrees would be created not only for different databases but also for the various data files, log files, and control files. This approach provides some manageability benefits, especially in database environments that use single-file SnapRestore® (SFSR) points for backup and recovery. Although this approach to data management has some strong advantages, careful planning is required before it can be effective. Some common questions while architecting a database solution using the qtree storage provisioning approach are:

> How do I utilize all my available storage in a volume and take advantage of NetApp Snapshot and SnapRestore technology?
> How can I make Snapshot copies of my existing data set writable?
> How can we create Snapshot copies of my qtrees and restore directly at a qtree level?
> How can we copy our database and use it in another environment with minimal impact on the load of the filer?
> How can we copy our data set multiple times without using the I/O-intensive native OS copy or dd commands?
> How can we more efficiently manage our data through the use of Snapshot on a qtree type level?

In this paper, FlexVol and FlexClone are demonstrated to efficiently answer all of the above questions. In addition, three case studies will be presented with performance data comparing the use of both traditional volumes and flexible volumes.

3. Technical Overview
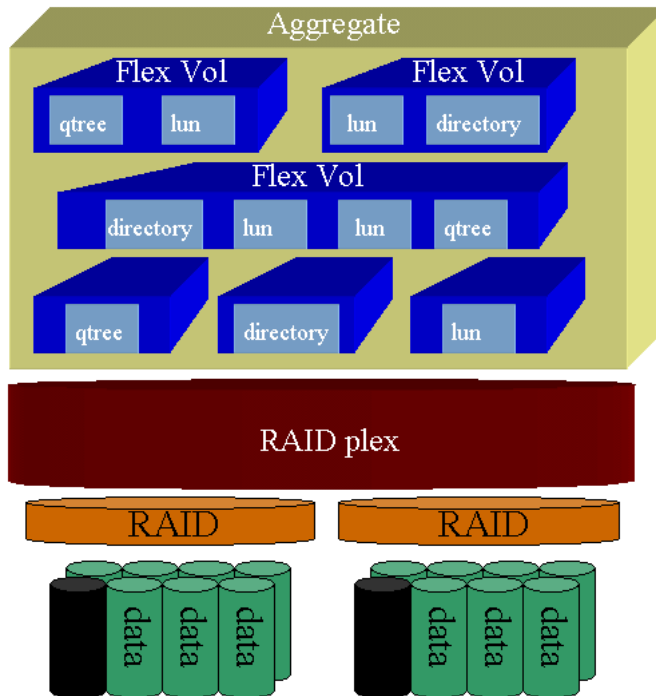
3.1. What Are Flexible Volumes?

Figure 1) Conceptual layout of an aggregate/FlexVol volume.

Prior to Data ONTAP 7G, volumes and aggregates were fused into a single unit where each aggregate contained exactly one volume. In Data ONTAP 7G, Network Appliance Inc. introduces a new kind of volume called a flexible volume, or FlexVol volume. This will allow volumes to be treated as logical entities separate from the physical storage that contains the associated data. The logical containers for pools of disks are now called aggregates. An aggregate can be described as having one or more flexible volumes (Figure 1). The new conceptual storage model of flexible volumes vs. traditional volumes is kept consistent to ease the user's transition to Data ONTAP 7G. Traditional volumes are supported in Data ONTAP 7G for backward compatibility.

A new feature of flexible volumes is the ability to create a space-efficient clone of a flexible volume called a FlexClone volume. This will allow users to make a writable point-in-time copy of the file system based on a Snapshot copy of the existing flexible volume or flexible clone. This is very beneficial in large database environments, allowing DBAs to create a DB template and then create writable clones off the main flexible volume template for use in their testing, development, or production environments. Flexible clones are made possible by the unique NetApp Snapshot architecture—a FlexClone volume only occupies new blocks of storage as needed by changes to the data. Additionally, a clone can be "split" off from the clone tree, thus allowing the administrator to create new permanent, independent flexible volumes for data storage.

The following terms are used throughout the paper and are important for understanding of Data ONTAP 7G:

Aggregate. A physical container of disks from which logical volumes and RAID groups can be created.

Traditional aggregate. An aggregate that contains only one traditional volume. Synonymous with "traditional volume." A traditional aggregate can only be created by the "vol create" command.

Flexible aggregate. A container for one or more flexible volumes, RAID groups, and disks. A flexible aggregate supports the ability to contain multiple, independent user-level file systems named flexible volumes. An aggregate consists of two or more disks and is available starting in Data ONTAP 7G. An aggregate can either be mirrored, using local SyncMirror®, or unmirrored, and aggregates can be copied using the "aggr" command. A flexible aggregate is free of any independent logical file system.

Traditional volume. A single physical volume that is associated with a collection of disks and managed via a single administrative unit.

Flexible volume. A logical volume available in Data ONTAP 7G that resides within an aggregate and can grow or decrease in size. It is only constrained by the soft limits set when created and the hard limits of the aggregate. It is the enabling technology for all other flexible features.

Flexible clone. A volume that is a writable copy of a flexible volume. Flexible clones use the unique Data ONTAP WAFL® files system and Snapshot architecture to only store changed blocks. Initially, the clone and its parent share the same storage. More storage space is consumed only as either one volume or the other changes. FlexClone writes use free blocks in the aggregate.

TPM. Transactions per minute.

LAT. Latency, measured in milliseconds (ms).

IOPS. I/O operations per second.

3.2. Benefits of Data ONTAP 7G for Databases

Data ONTAP 7G allows easier management of your storage environment and an even greater ROI in database environments by providing FlexVol and FlexClone functionality. Easier management of your storage is made available with flexible volumes, and a greater ROI is realized by using flexible clones of flexible volumes.

3.2.1. Benefits of FlexVol in a Database Environment

Flexible volumes deliver many benefits over traditional volumes in a database environment by providing the capabilities of fast backup and restore times as well as allowing quicker deployments of your DB environment.

With the new addition of flexible volumes, NetApp introduces the benefits of volume-based SnapRestore to both a LUN level and a more granular volume level. Flexible volumes solve issues that in the past involved doing many coordinated single-file SnapRestore (SFSR) operations to restore a database. Databases can now be separated into their own FlexVol volume, and one command will now bring the entire database back to a point-in-time Snapshot copy. In a typical production database environment, a DBA may deploy hundreds of data files, log files, and archive files. When configured in a single volume that has 10 to 100 databases, the method for creating Snapshot copies and restoring the database files could be cumbersome. Although Snapshot technology enables extremely quick database backups and restores, flexible volumes further empower DBAs to manage their databases with substantial management and possible performance advantages.

FlexVol provides the ease of management that was previously attained through the use of multiple qtrees. LUN management configuration can be better managed through the use of a FlexVol volume, as opposed to a qtree. Databases can be separated into individual LUNs that reside in a specific FlexVol volume for that specific group and/or department. Flexible volumes simplify configuration scripts and allow DBAs to manage their databases more efficiently. In addition, using SnapRestore with the FlexVol volume improves database restore times tenfold over having to utilize multiple SFSR. Complete details can be found in the technical white paper *TR3356: Introduction to Data ONTAP 7G*.

Flexible volumes allow the DBA or storage administrator to provision storage based on usage needs. Volumes can be configured to allow optimum disk throughput based on the current needs of the application environment. This allows storage administrators to create single or fewer large aggregates and assign flexible volumes based on storage demands. This pool of available storage allows for greater flexibility and performance capacity in meeting the growing demands of the database environment. This combined with the ability to set separate Snapshot

schedules increases the ability to recover individual databases from a FlexVol Snapshot copy and create clone copies of data when necessary.

3.2.2. Benefits of FlexClone in a Database Environment

As seen with FlexVol, a FlexClone volume provides even greater functionality in a database environment. A FlexClone volume is a writable copy of a flexible volume or clone and can be created nearly instantaneously. The FlexClone volume shares unmodified blocks with the parent flexible volume and requires space only for the differences between the two. When a FlexClone volume is initiated, no additional load is imposed on the production flexible volume except the I/O driven to the clone copy.

In a database environment, a FlexClone volume allows the DBA to create an exact copy of a database within seconds when the data resides on a NetApp flexible volume. The DBA can then use that writable mirror copy for development purposes as well as testing and reporting purposes. FlexClone also provides great advantages in a business application environment such as SAP or Oracle® Applications by allowing patches to be applied and tested to a FlexClone volume before deploying to the production FlexVol environment. In addition, a production database can be cloned using FlexClone to allow an IT manager to quickly deploy a test copy of the production environment for problem and fault isolation, leading to quicker problem analysis and resolution.

FlexClone also provides great advantages in a SnapMirror® environment by allowing the SnapMirror destination flexible volume copies to be cloned using FlexClone. This will allow the DBA to create a FlexClone volume of the destination SnapMirror and start the database without having to quiesce and break the SnapMirror copy. This allows the SnapMirror synchronizations to continue while minimizing Snapshot copies. In addition, this allows the remote SnapMirror database copy to be tested on the remote disaster recovery site. Complete details can be found in the white paper *TR3347: A Thorough Introduction to FlexClone Volumes*.

4. Testing Methodology
4.1. Testing and Architecture

The criteria used in the performance comparison were gathered based on studies taken of existing customers using Data ONTAP 6.5.1 of the NetApp filer. These studies were the result of current customer usage trends and usage profiles gathered from the NetApp AutoSupport tool. Further information was gathered from analyzing disk usage patterns while running OLTP-like workload tests and examining NetApp IT production filer usage patterns. Tools used for the testing included SIO (Simple I/O Load Generator V5.48) and a Windows file-serving workload generator.

The goal of the test was not necessarily to see the maximum load the NetApp filer can sustain, but to compare similar workloads on the same filer utilizing traditional volumes and flexible volumes.

The systems under test consisted of 2 Sun™ V4xx class servers connected via an Emulex LP 9002 FCP card through a Brocade Silkworm Entry 3800 fabric switch supporting 2 Gb per second Fibre Channel connections. In addition to the Sun servers, four PCs based on Intel® architecture accessed the same filers connected via Intel Gigabit Ethernet adapters through an Extreme gigabit switch. Two of the Intel PCs were emulating CIFS workloads using a Windows file-serving test suite, and the other two Intel PCs were running Redhat 9.0 using the SIO tool to emulate OLTP-type workloads. The Sun servers were configured to utilize an OLTP-type workload over a raw FCP partition connection.

The NetApp FAS920c was configured and tested in a single-node configuration. The FAS920c was connected to eight DS14mk2-FC shelves totaling 112 disks. The FAS920c was configured with 4GB of RAM and 512MB of NVRAM. The disks utilized in the testing were the 72GB FC disks.

Figure 2 shows the architectural layout of the hardware used in the testing.
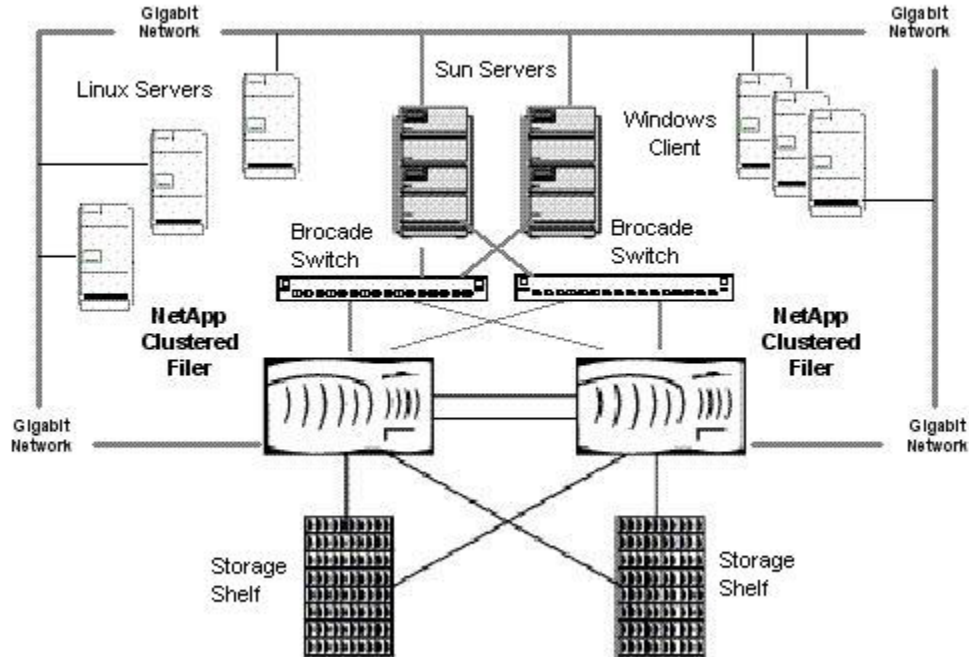
Host/Filer Architecture

Figure 2) Host/filer architecture.

The configuration shown in Figure 2 includes several Window clients as well as Linux® and Solaris™ servers that are connected via FCP as well as NFS to the NetApp 920c filer.

4.2. Database Case Studies

Three case studies will be presented. Two case studies will demonstrate the performance impact between traditional volumes and flexible volumes. The third case study will present the performance impact of running a flexible clone while under a filer performance load.

4.2.1. Case Study A—Single-Database Workload (NFS)

Traditional Volumes

A customer has a small workload and is using a filer to manage a single database. The customer is currently using the filer to host databases in a traditional volume and wants to better understand how to optimize the number of spindles in use. In addition, the customer would like to understand how using flexible volumes will impact performance while still adding greater manageability with FlexVol Snapshot copies and FlexClone cloning. The ultimate goal is to better manage a single Oracle Database and grow to multiple databases using the FlexVol and FlexClone technologies.

The current configuration includes one server and two disk shelves. One 10-disk volume (volume A) will be used to hold the data files and one four-disk volume (volume D) to hold the redo logs and archive logs. Since this is a rather small database, it is a good low-level benchmark to understand the impact of using FlexVol. In addition, the second NetApp DS14 shelf will be utilized once the customer understands the impact of upgrading to Data ONTAP 7G.

Figure 3 shows the current database configuration.

Database Architecture for a 10-Disk Traditional Volume and a 4-Disk Traditional Volume
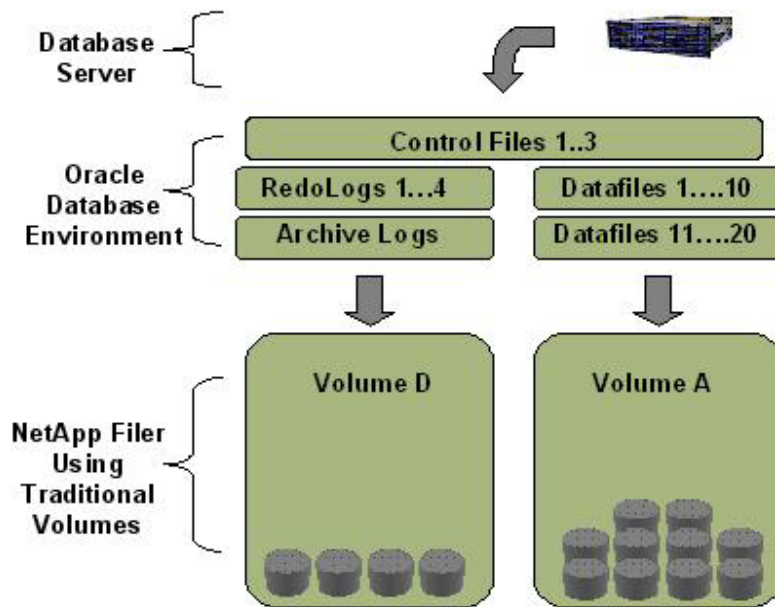

Figure 3) Conceptual view of a 14-disk traditional volume used in a database workload.

As you can see in Figure 3, the database is separated into two separate volumes: one for the Oracle data files and one for the Oracle redo logs and archive logs when enabled. Transaction loads to be used in this configuration are as follows:

Volume layout case study A using traditional volumes:

Volume Name	Number of Disks	RAID Type	Access Method	Volume Capacity
Volume A	10	RAID4	NFS	50%
Volume D	4	RAID4	NFS	50%

In the above table, the workload will consist of two volumes. Volume A will consist of 10 disks, and volume D will consist of four disks. The access method used in this study is NFS.

Volume load case study A using traditional volumes:

	Read %	Write %	Random %	Block Size	Threads/Procs	Protocol
Volume A	80	20	80	4096	2	NFS
Volume D	0	100	0	4096	32	NFS

The workload pattern includes the above scenario. We will emulate 80% random reads to volume A vs. 100% sequential writes to volume D. This workload is representative of a typical OLTP-type workload, and the intent was to provide a worst-case scenario.

Network Appliance Inc.

8

Results of Traditional Volumes

As seen in Figure 4, when traditional volumes are used, the filer is able to maintain 7382 filer operations per second when the OLTP workload test is performed while using traditional volumes.

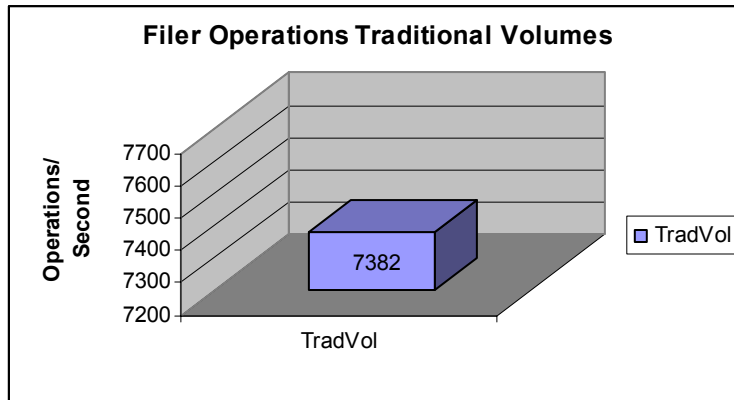**Filer Operations Traditional Volumes**

Figure 4) Traditional volumes operations per second.

The performance metrics in Figure 5 show an average disk utilization of 55% for traditional volumes.
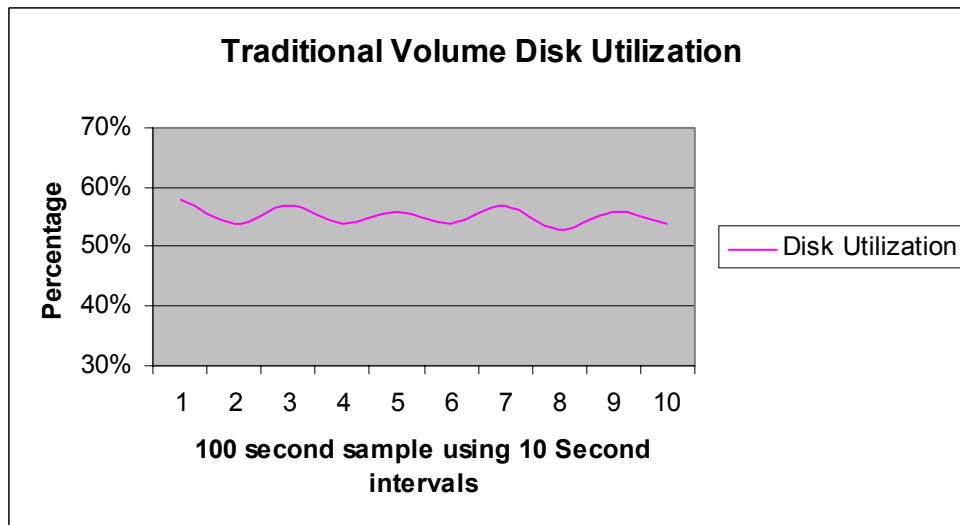
Traditional Volume Disk Utilization

Figure 5) Disk utilization statistics in a traditional volume workload test.

The workload introduces moderate disk utilization as a whole. However, to look at the statistics individually, the volume D (10-disk) workload had more moderate disk operations, while the volume A (four-disk volume) disks were operating at an average of 272 disk operation per second. The disks were working much harder to service the requests of the sequential writes required for log data. Further shown is the randomness of the peak disk throughput (disk utilization percentage) while under load when using traditional volumes for that workload.
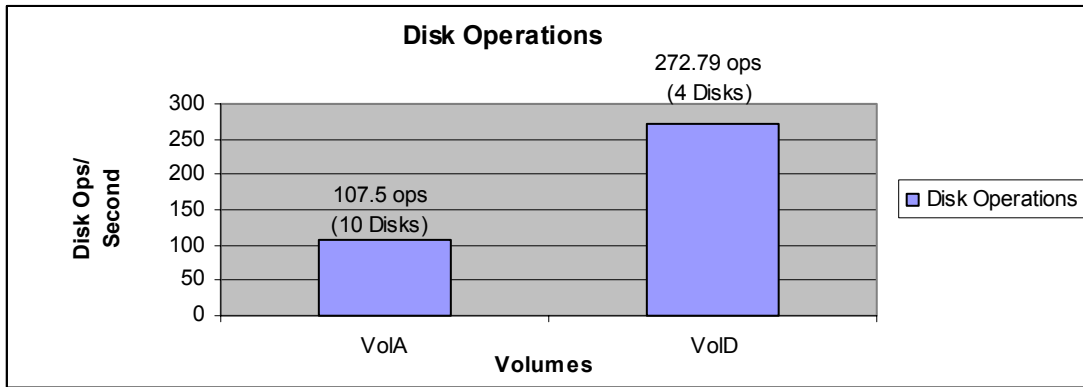
Disk Operations

Figure 6) Disk operations per second on a traditional volume for volume A vs. volume D.

Further illustrated in Figure 6 is volume A (10-disk volume) under a moderate load and volume D (four-disk volume) under an extremely busy load. Volume A is running at 107.5 disk operations per second on the data volume, while volume D is running at 272.79 disk operations per second on the smaller log file volume. The traditional volume shows a constrained random workload, while the four-disk volume (volume D) is showing heavy disk operation usage. The heavy write usage to the database log files (volume D) is common in a high-transaction workload database environment.

Flexible Volumes

In Data ONTAP 7G the database architecture will be conceptually the same as the traditional volume tests. However, the data is striped collectively over the entire set of 14 disk drives instead of two separate volumes of 10 and four disk drives. Figure 7 shows the flexible volume architecture used for testing. The flexible volume architecture allows Snapshot copies to be managed for each volume, while the larger aggregate will spread the I/O across all the hard disk drives.

Network Appliance Inc.

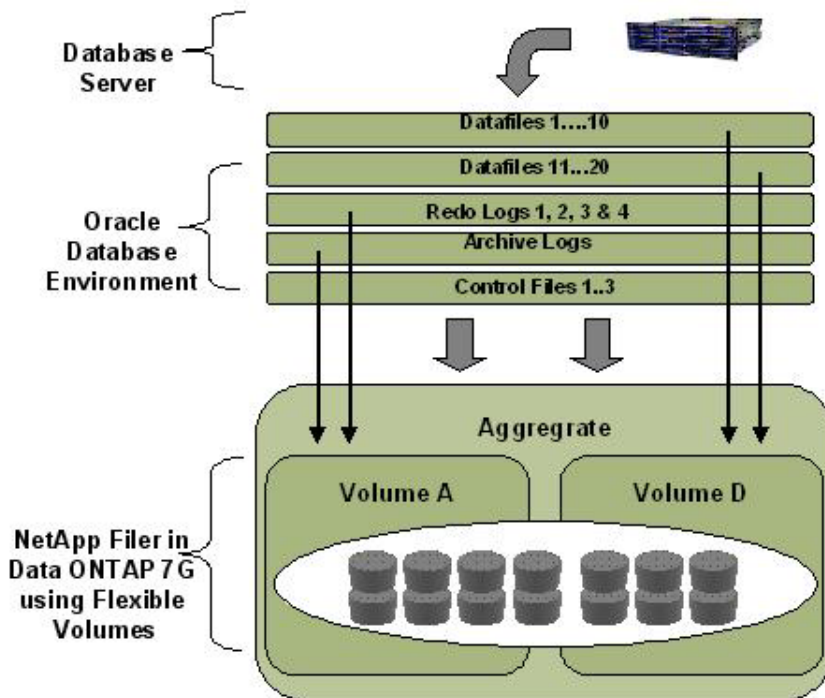## Database Architecture for 14 Disk Volume in Flexible Volumes

Figure 7) Conceptual view of a 14-disk FlexVol volume used in a database workload test.

In Data ONTAP 7G flexible volume sizes are based only on the amount of storage needed, as opposed to being directly related to the number and capacity of disk drives assigned. The next layout identifies the testing to be used for the next suite of tests.

Volume layout case study A using flexible volumes:

Volume Name	Aggregate Name	Number of Disks	RAID Type	Access Method	Volume Size	Volume Capacity
Volume A	Aggr1	14	RAID-DP™	NFS	300GB	50%
Volume D	Aggr1	14	RAID-DP	NFS	120GB	50%

In the above table the workload will consist of one aggregate (aggr1) consisting of 14 disks. Two volumes, volume A and volume D, are created in the aggregate. The access method used in this study is NFS.

Volume load case study A using flexible volumes:

	Read %	Write %	Random %	Block Size	Threads/Procs	Protocol
Volume A	80	20	80	4096	2	NFS
Volume D	0	100	0	4096	32	NFS

The workload pattern used was identical to the workload pattern used for the traditional volume test. We will emulate 80% random reads to volume A vs. 100% sequential writes to volume D. The workload is an OLTP workload and the same workload used in testing FlexVol volumes.

Results of Flexible Volumes

Comparisons between flexible volumes vs. traditional volumes shows a 4.1% overall gain in the number of filer operations that are capable with a similar DB workload when compared to traditional volumes. Figure 8 shows the

Network Appliance Inc.

number of NFS filer operations averaged doing the same database workload as described earlier in "Volume Load Case Study A." The increase in NFS operations is partly due to the number of hard disk spindles that the random read I/O is able to utilize more efficiently when aggregated over the 14 disk spindles.
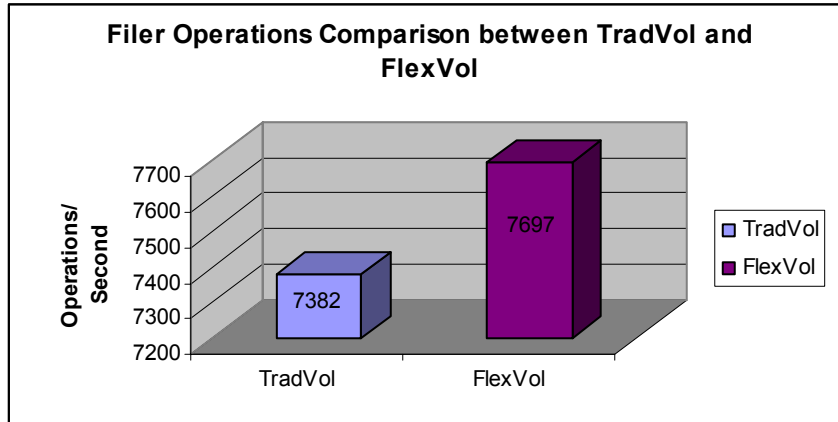
Filer Operations Comparison between TradVol and FlexVol

Operations/Second	
TradVol	7382
FlexVol	7697

Figure 8) Traditional volumes vs. flexible volumes workload test.

The performance metrics in Figure 9 show an average disk utilization of 57% for flexible volumes. The CPU utilization difference between the two workloads is ~2% higher with FlexVol, while the disk utilization workload is ~2% higher than used in the traditional volume test. So the results are relatively the same performance from a resource perspective. However, FlexVol adds a layer of value that traditional volumes do not offer while utilizing the filer's resources more efficiently.
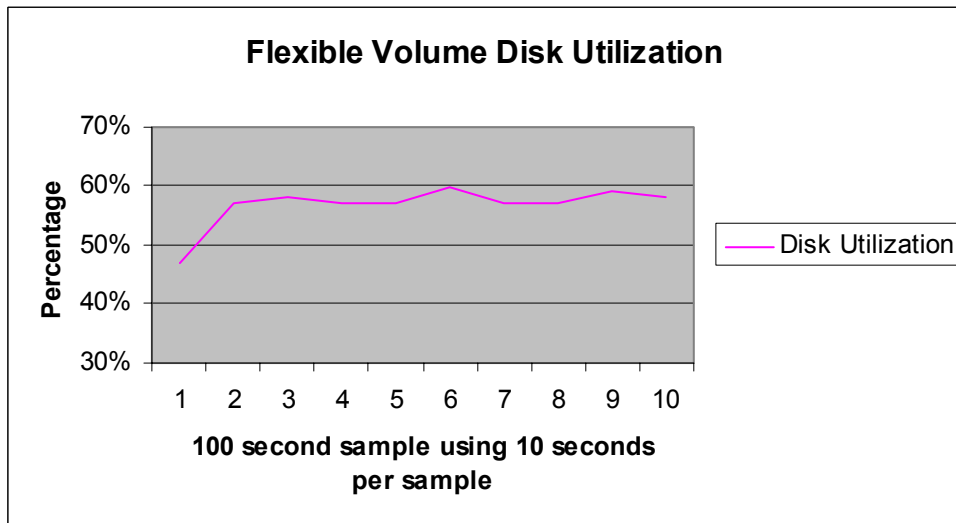
Flexible Volume Disk Utilization

Disk Utilization

100 second sample using 10 seconds per sample

Figure 9) CPU and disk utilization statistics in a FlexVol workload test.

In Figure 9 the disk utilization is more consistent, since the I/O load is able to aggregate over more disk spindles. When utilizing FlexVol in an aggregate, the average utilization of the disk spindles is more consistent, since the data pattern is spread across more disks. In Figure10 the total number of disk operations per second decreases for the 14-disk RAID-DP volume. However, the total throughput increases, and the CPU resources are relatively consistent.
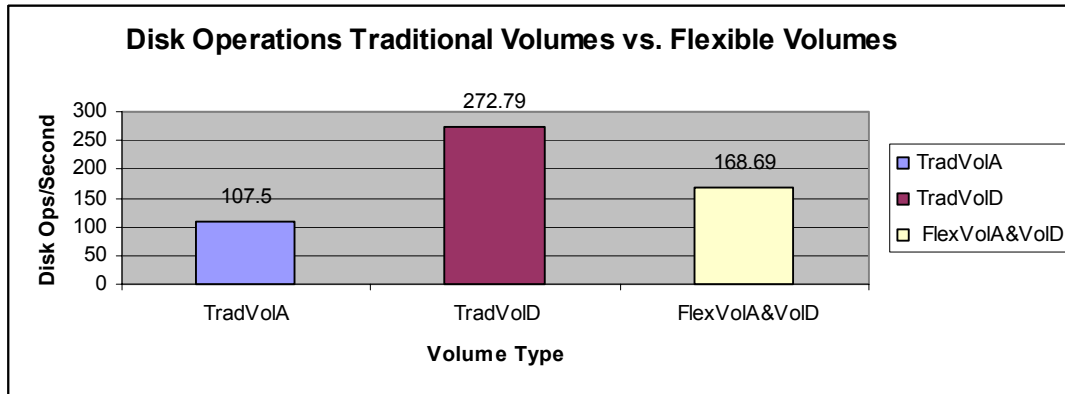
Disk Operations Traditional Volumes vs. Flexible Volumes

Figure 10) Disk operations per second comparison with traditional volume vs. flexible volume for volumes A and D.

When compared to traditional volumes (volume A = 10 disks and volume D = 4 disks), as shown in Figure 10, the number of disk operations has leveled out to 168 disk operations per second in the flexible volume scenario. The FlexVol volume is able to spread the load among the increased spindles more efficiently, as compared to the traditional volume–only test, as shown in Figure 6 and again in Figure 10. In Figure 10 the FlexVol volume (volume A and volume D) is composed of 14 spindles.

Single-Database Performance Study Results

Flexible volumes provide great advantages in a single-DB workload and even greater advantages in a multiple DB workload environment. In the single-database case study we showed how the filer is able to maintain a heavy workload and maintain a similar data throughput rate as compared with the traditional volumes workload test. This is done while having the additional advantages of volume-level and LUN-level cloning and the ability to manage and provision your data on an as-needed basis.

The throughput demands in this FlexVol scenario are greater than they were was before when using traditional volumes. In our scenario we were still able to show an overall increase in the average throughput numbers as compared to the traditional volume testing of 4%.

Calculation of bandwidth increase when moving to a FlexVol volume for single DB:
7697 – 7382 = 315 / 7382 = 4.3%

The biggest gain is seen in the random I/O load because of the increase in available disks. More gain will be noticed with the increase in number of spindles, depending on the load pattern (random vs. sequential).

4.2.2. Case Study B—Mixed Database Workload (CIFS, NFS, and FCP)

Company ABZ has deployed a FAS920c filer in an environment and uses the filer in a mixed use scenario. The company has split the applications onto separate volumes, and the peak times vary depending on the applications in use:

Oracle DB (peak times: 8:00 a.m. and 3:00 p.m.)
CIFS home directory shares (peak times: 11:30 a.m. and 4:30 p.m.)
Typical NFS traffic (peak times: 8:30 a.m. and 3:30 p.m.)

In the current configuration the filer is used most heavily during the morning time around 8:00 a.m. and then again in the afternoon around 3:00 p.m. Each environment has its own mixture of workload access patterns. The current volume configurations are as follows.

The company would like to utilize the available spindles so that it can add and subtract to volume sizes when necessary. The following table discusses the volume layout and capacities used in the testing.

Volume layout case study B using traditional volumes:

Volume Name	Number of Disks	RAID Group Size	Access Method	Load Type	Volume Capacity
Volume A	10	8	NFS	DB	50%
Volume B	8	8	CIFS	Home Dir	50%
Volume C	6	8	CIFS	Home Dir	50%
Volume D	4	8	NFS	DB	50%
Volume E	10	8	FCP	DB	50%

Case study B required a move from traditional volumes to flexible volumes. The storage space needed was approximated, and one large 42-disk aggregate was created. The flexible volumes were sized based on the actual volume size necessary to complete the tests. Having one big pool of disks allows for growing and shrinking the flexible volume when needed. This allows for even greater flexibility and manageability by carving up the volumes within the aggregate. The following table describes the data layout as composed within a FlexVol volume.

Volume layout case study B using flexible volumes:

Aggregate Name	Number of Disks	Volume Name	Size of Volume (GB)	RAID Group Size	Access Method	Volume Capacity
Aggr1	42		2146			
		Volume A	300	RAID-DP 14	NFS	50%
		Volume B	250	RAID-DP 14	CIFS	50%
		Volume C	200	RAID-DP 14	CIFS	50%
		Volume D	150	RAID-DP 14	NFS	50%
		Volume E	300	RAID-DP 14	FCP	50%

The test workload mixture is described in the following table. The workloads are based on varied database workloads using FCP and NFS protocols as well as a CIFS workload characterized by NetApp internal IT organization. This workload is performed using a SPEC SFS CIFS (System File Server Suite Release 3.0) to drive the CIFS workload, as agreed upon by the SPEC organization.

The purpose of the workload is not to get the best CIFS results but to be a measurement tool for comparing traditional volumes vs. flexible volumes.

Workload mixture case study B:

	Read %	Write %	Random %	Query Path %	Block Size	Threads/Procs	Protocol
Volume A	100	0	100	N/A	8192	8	NFS
Volume B	25	15		54	8192	4	CIFS
Volume C	25	15		54	8192	4	CIFS
Volume D	50	50	50	N/A	4096	4	NFS
Volume E	80	20	50	N/A	8192	10	FCP

In the above scenario we are mixing a heavy read-intensive workload with a CIFS home directory workload and an FCP workload using an OLTP-type transaction load on volume E.

Results of Traditional Volumes Mixed Workload

Performance results using traditional volumes in a mixed configuration are shown in Table 1.

	IOPS	Throughput (kb/sec)
Volume A	1125	8996
Volume B	1002	1002
Volume C	1002	1002

Volume D	2624	20,996
Volume E	2792	22,100

Table 1) Traditional volume throughput results (case study B).

The results in Table 1 show a filer that is disk limited in one of the volumes, while other volume disks are highly utilized. In an ideal scenario the load would be spread evenly over the disks that are not utilizing the available disk operations.

Results of Flexible Volumes Mixed Workload

Case study B for flexible volumes illustrates that all loads ran concurrently and represents a worst-case scenario.

	IOPS	Throughput (kb/sec)
Volume A	1370	10,964
Volume B	1000	1000
Volume C	1000	1000
Volume D	1383	11,060
Volume E	6466	51,726

Table 2) Flexible volume throughput.

If we were to separate out the peak loads, the filer would be able to benefit from the increased spindles, since in a mixed case scenario typically not all peak traffic happens at the same time. While this shows a concurrency to the load when run in a worst-case scenario, the overall throughput is considerably higher than in a traditional volume scenario. In addition, the benefits of FlexVol and FlexClone are even more significant.

Mixed Workload Database Performance Study Results

Overall throughput gains are seen with flexible volumes in a mixed workload scenario. Efficient usage of disk spindles based on I/O load is realized. Table 3 shows the benefits in terms of overall throughput level that FlexVol provides in a total mixed workload environment. This environment may include databases and home directories as well as NFS traffic in a normal day-to-day scenario.

The gain by utilizing the available spindles in a FlexVol environment is substantial when compared overall.

Test Type	Average Total Ops	Max Ops	Min Ops	Disk Utilization %	Avg Disk (Reads)	Avg. Disk (Writes)
Traditional Volumes	8241	11,151	6010	53	27,841	25,019
FlexVol	10,766	12,940	2994	24	49,092	15,562

Table 3) Throughput results of case study A when used with traditional volumes.

Figure 11 shows that aggregating the spindles allows as much as 30.6% gain of available resources. This gain is achieved by the increase in the spindle count, which allows the random read workload to increase substantially over what was tested in the traditional volume test.

Calculation of bandwidth increase when moving to a FlexVol volume:
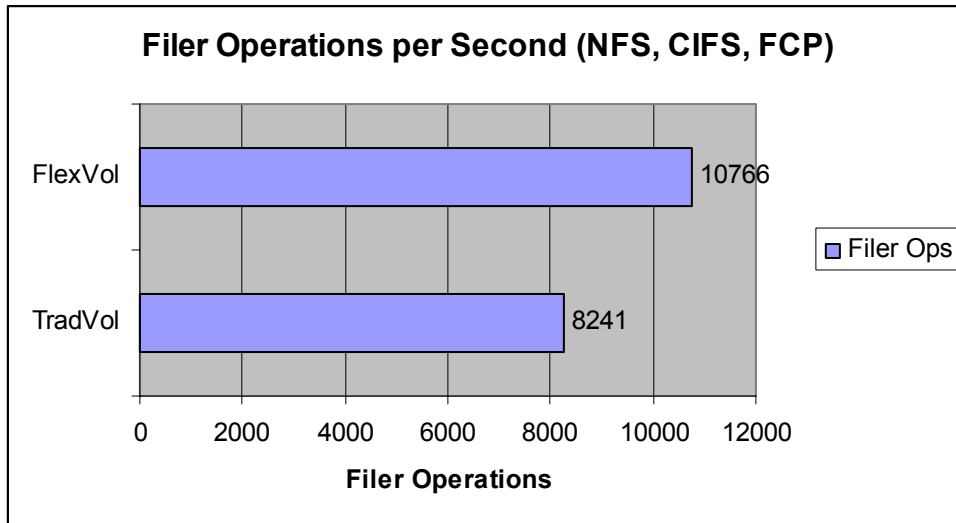10,766 – 8241 = 2525 / 8241 = 30.6%

Figure 11) Comparison of overall filer operations between traditional volumes vs. flexible volumes.

The filer is able to provide increased filer operations while using the same load pattern as in the traditional volume configuration. In an ideal scenario we would like to spread the load evenly across a wide number of disks. FlexVol provides that capability while still allowing the DBA to provision, manage, and schedule Snapshot copies and clones of the data when necessary. Although the goal of this paper is to show equivalent performance with added flexibility, we do see in some workloads (heavy random read workloads) where you will gain performance in addition to the added flexibility. One caveat: when the filer is at maximum CPU utilization, careful planning is necessary before moving from traditional volumes to flexible volumes to avoid further bottlenecks.

4.2.3. Case Study C—FlexClone Performance Impact on Database (NFS and FCP)

A large outsourcer, Company ZBA, has 10TB of NetApp storage to run multiple Oracle DBs. One of the databases is having a problem that Oracle support believes is Oracle related. The Oracle support team needs to debug this problem further without impacting the load on the other environments. The first option is to debug the DB while running in a production environment, which is the least preferred method due to the risk of affecting the production system. The second option would be to put the DB into hotback mode and make a copy of the database having the problem to another system. This method is not optimum because it would require copying all the files to another system for troubleshooting. The copying process would cause performance degradation to the production load and could take hours to complete—hours that could be better spent debugging. The third option would include restoring the DB from tape to another location for testing purposes. This method could take hours to complete and require additional volumes with available space to make it happen. This method may not even be an option if the backup was created prior to the problem's existence.

The criteria we want to use to select the best method are:

Minimize performance impact of the production DB
Make a point-in-time copy of the DB after the problems began occurring
Test in a new environment, emulating the production DB as closely as possible
Be able to apply a patch if necessary to verify whether it solves the problem
Minimize downtime to the production DB while resolving the problem as quickly as possible

Each environment has its own mixture of workload access patterns. The current volume configurations are shown in Table 4.

Volume Name	Number of Disks	RAID Group Size	Access Method	Volume Capacity
Volume A	10	8	NFS	50%
Volume B	8	8	NFS	50%
Volume C	6	8	NFS	50%
Volume D	4	8	NFS	50%
Volume E	10	8	FCP	50%
Volume AA	10	8	NFS	50%
Volume BB	8	8	NFS	50%
Volume CC	6	8	NFS	50%
Volume DD	4	8	NFS	50%
Volume EE	10	8	FCP	50%

Table 4) Volume configuration for volume layout (case study C).

Tables 5 and 6 illustrates the layout to be used in a FlexVol test to be done for this test.

Aggregate Name	Number of Disks	Volume Name	Size of Volume (GB)	RAID Group Size	Access Method	Volume Capacity
Aggr1	42		2146			
		Volume A	300	RAID-DP 14	NFS	50%
		Volume B	250	RAID-DP 14	NFS	50%
		Volume C	200	RAID-DP 14	NFS	50%
		Volume D	150	RAID-DP 14	NFS	50%
		Volume E	300	RAID-DP 14	FCP	50%
		Volume AA	300	RAID-DP 14	NFS	50%
		Volume BB	250	RAID-DP 14	NFS	50%
		Volume CC	200	RAID-DP 14	NFS	50%
		Volume DD	150	RAID-DP 14	NFS	50%
		Volume EE	300	RAID-DP 14	FCP	50%

Table 5) Volume layout using a FlexVol volume—volume size (case study C).

	Read %	Write %	Random %	Query Path %	Block Size	Threads/Procs	Protocol
Volume A	100	0	100	N/A	8192	2	NFS
Volume B	5	95	0	N/A	8192	2	NFS
Volume C	5	95	0	N/A	8192	2	NFS
Volume D	50	50	50	N/A	8192	2	NFS
Volume E	100	0	20	N/A	8192	2	FCP
Volume AA	100	0	100	N/A	8192	2	NFS
Volume BB	5	95	0	N/A	8192	2	NFS
Volume CC	5	95	0	N/A	8192	2	NFS
Volume DD	50	50	50	N/A	8192	2	NFS
Volume EE	100	0	20	N/A	8192	2	FCP

Table 6) Workload mixture (case study C).

Results of Flexible Clones Performance Impact

The final test result involves creating a FlexClone copy of a volume holding a database while still running the OLTP workload. After completing the cloned copy of the FlexVol volume, the original database load is still running on the original FlexVol volume. The FlexClone volume is now ready for additional testing or load if necessary. Figure 12 represents the amount of time taken to create a FlexClone volume while a DB load is running against the FlexVol volume.
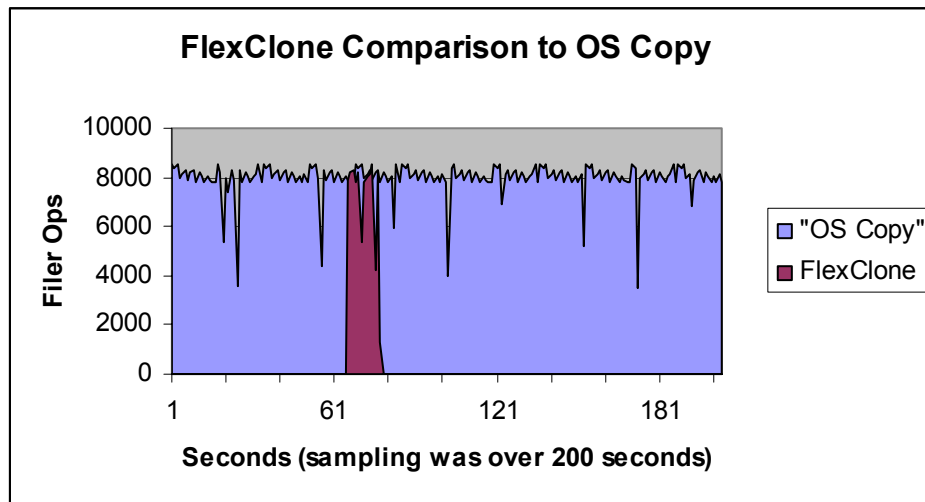
FlexClone Comparison to OS Copy

Figure 12) Comparing the amount of time taken creating a FlexClone volume against a normal OS-level copy of 300GB of data.

Figure 12 illustrates that the amount of time taken to create the FlexClone volume is a magnitude of 100 times less than if the same set of operations were taken using traditional OS copy commands, which on average can take hours to complete. In order to keep Figure 12 readable, only a 200-second sampling is shown, because the OS copy took a significant amount of time longer than the FlexClone copy. The flexible clone was created during a heavy OLTP-type workload and was completed within a 20-second time frame. In addition, this is done within a flexible volume framework, which allows the cloned copy to be writable. Previous versions of Data ONTAP only allowed for clones to be created at a LUN level. In Data ONTAP 7G, writable Snapshot copies (FlexClone volumes) are available at a volume level.

FlexClone Workload Database Performance Study Results

A FlexClone copy provides all the benefits of Snapshot with negligible overhead. The initial FlexClone copy will show a temporary hit in system load and then resume to normal throughput levels. From a performance perspective, less than 10% overhead was seen on the initial copy. This operation typically took less than 10 to 20 seconds under heavy load, if not less.

Part of the test in case study C was to create a clone copy of the data while under load. The goal was to minimize the performance impact in order to minimize the downtime of the production DB. As shown in Figure 12, the performance impact was minimal compared to that of a normal OS-level copy. In addition, a key benefit of making flexible clones of databases for debugging is that the databases will run on a performing system rather than a separate system with few disks.

5. Summary of Results

Flexible volumes are comparable in speed to traditional volumes. In most cases flexible volumes will provide a 10% to 20% performance increase over a traditional volume by utilizing resources that are now currently available in one large disk pool (aggregate). However, this does depend greatly on the type of load pattern you are testing against and the type of clients and protocols you are using. In our testing we used three protocols.

Data ONTAP 7G provides great advantages in the database environment. FlexVol and FlexClone empower the DBA or storage administrator to architect the environment that best suits the situation's needs. There is no need to overprovision storage to utilize the available bandwidth on the spindles. In Data ONTAP 7G these spindles can be consolidated into one or more aggregates, and FlexVol can then be used. This can be done while still delivering

the throughput demands of the current application. It also allows multiple FlexVol volumes to be created and used for other purposes with their own Snapshot schedules.

Instead of setting up multiple smaller volumes to take advantage of NetApp Snapshot technology or using one larger volume and creating your databases inside qtrees, you can now do this at a volume level and benefit with the new features of Data ONTAP 7G, which include FlexVol and FlexClone at the core of the new release. If the overall aggregate size is small, it is important to separate random and sequential traffic. With large enough aggregates, this no longer becomes a concern. If sequential write throughput is a concern in a combined random/sequential write aggregate, then creating multiple aggregates and allowing the sequential write volumes to be consolidated into one aggregate would provide the best solution. You can then use the other aggregate for your volumes that need high random read throughput.

FlexVol allows you to take advantage of NetApp Snapshot technology while still allowing you to benefit from our new FlexClone capabilities.

If you are using SnapMirror technology, SnapMirror copies of the database on the destination copy can be cloned multiple times and tested at a remote DR site. The SnapMirror destination clones can then be recloned and brought online without breaking the SnapMirror relationship.

6. Conclusion

The most challenging tasks of a database administrator include the following:

> Performance troubleshooting
> Patch upgrades
> Planning database deployment
> Backup and recovery issues
> Replication and synchronization
> Resource utilization
> Disaster recovery

Data ONTAP 7G provides enormous benefits in all these areas by allowing you to stripe and mirror all your disks with our WAFL file system and then manage your resources by allowing you to create flexible volumes within one large pool of disks. Patch upgrades are made seamless with both our Snapshot capability and our ability to clone a FlexClone volume specifically for testing purposes.

You can now do backup and recovery with our volume-level Snapshot copies in less time than it takes to enter the command from the command line. Replication and synchronization are made simpler in Data ONTAP 7G because of the ability to utilize SnapMirror or SnapMirror Sync combined with the ability to create a FlexClone copy of a FlexClone volume on the destination mirror site. Database planning and resource utilization are optimized by using large aggregates, and storage is assigned on an as-needed basis, providing greater ROI.

Data ONTAP 7G database planning is simplified as compared to traditional volumes. Organizations that have relied on qtrees to manage separation between work groups and departments can now add another level of manageability that allows even more functionality with flexible volumes. A brief summary of the benefits that flexible volumes bring to a database environment includes:

> Availability of Snapshot at a FlexVol level
> Availability of FlexClone copies (writable Snapshot copies) at a FlexVol level
> Easier management of your data and placement of data with similar access patterns
> The ability to create a FlexClone copy of a FlexClone volume
> The ability to use SnapMirror in conjunction with FlexClone for onlining of your remote database

For more information about flexible volumes and aggregates in Data ONTAP 7G, please refer to *www.netapp.com/tech_library*. Refer to the performance section in the Data ONTAP 7G technical overview for performance guidelines per filer type.

7. Acknowledgements

The author would like to thank the following individuals for their contribution:

Network Appliance, Inc.
Glen Colaco, Andrew McCumiskey, Quinn Summers, Paul Turner