White Paper

DEPLOYING ORACLE DATABASE ON EMC VNX UNIFIED STORAGE

Best practices for provisioning storage and leveraging storage efficiency features

Abstract

This white paper introduces how the EMC® VNX™ unified storage platform can be effectively used for deploying enterprise Oracle Database applications. This paper also captures most of the Oracle performance testing done by EMC performance engineering. Some best practices for deploying database applications are also covered.

May 2011

ORACLE® Platinum Partner

EMC² where information lives®

Part Number h8242.1

Table of Contents

Executive summary

The EMC® VNX™ family of storage systems represents EMC's next generation of unified storage optimized for virtualized environments. The massive virtualization and consolidation trend of servers demands a new storage technology that is dynamic and scalable as shown in Figure 1. The EMC VNX series offers several software and hardware features for optimally deploying mission-critical enterprise applications. These platforms can also be used to consolidate several block, file and object-based applications in any enterprise data center.

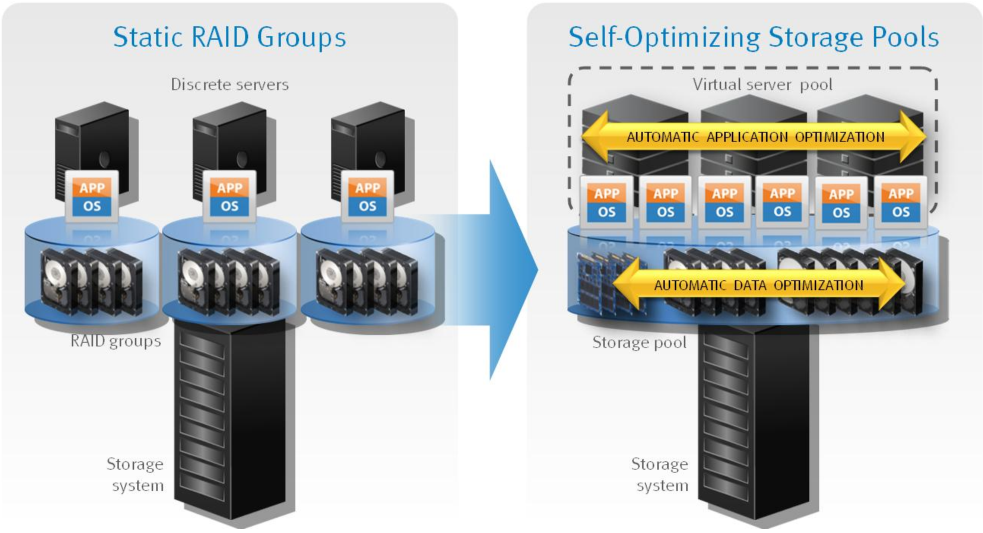

Figure 1. Storage system optimized for virtualized environment

EMC was the first company to introduce Flash drives, also known as solid state drives (SSDs), to the midrange storage platform, as well as various software technologies to fully leverage the performance potential of SSDs. The VNX series with its latest hardware and software advancements improves the story further. For example, FAST Suite technologies eliminate the need for manually identifying hot data by automating identification and migrating data to the right tier. By focusing the use of Flash drives on the most frequently accessed data, Flash drives deliver measurable application performance improvements, and significantly reduce total cost of ownership (TCO) of the solution.

Oracle Database deployed on VNX systems can achieve sizable performance improvements by leveraging some or all of the storage efficiency features. This document discusses several tests conducted at EMC Application Performance Labs. All of the tests consistently show a performance improvement of two to six times when using these storage efficiency features with Oracle Database compared to not using them.

Audience

This white paper is intended for database administrators, storage administrators, and enterprise database application architects.

Introduction

VNX platform

The introduction of the new VNX family of unified storage platforms not only continues the EMC tradition of providing one of the highest industry standards in data reliability and availability, but has factored into the design a boost in performance and bandwidth to address the sustained data access bandwidth rates of modern database applications. The new system design has also placed heavy emphasis on storage efficiencies and density, as well as crucial "green" storage factors, such as data center footprint, lower power consumption, and improvements in power reporting. Figure 2 shows the various platform choices available in the VNX series.

Figure 2. VNX platform models

The VNX platform has been completely redesigned to meet the demands of a modern data center using the latest processors, drive types, and connectivity options. A complete introduction to the VNX series is beyond the scope of this document; however, the References section has links to more information.

Even though most of the concepts and use cases discussed describe VNX block implementations, they also apply to VNX file implementations.

FAST Suite

The FAST Suite is a software bundle offered on the VNX series that contains four products, two of which focus on advanced data efficiencies:

- FAST VP (Fully Automated Storage Tiering for Virtual Pools)
- FAST Cache

Both technologies reduce the number of drives required to meet a given performance requirement, thereby significantly reducing the TCO of a storage system while achieving the desired level of performance.

FAST Suite applicability

Like any other technology, it is important to understand the system challenge thoroughly to maximize the benefits of the FAST Suite software. It is a fact that the Flash drive performance is an order of magnitude greater than rotating drives (HDDs), and they deliver the higher performance for the applications when deployed. At the same time the main goal of the FAST Suite is to optimize use of Flash drives while delivering the required performance. Figure 3 illustrates the expected benefit of using the FAST Suite.

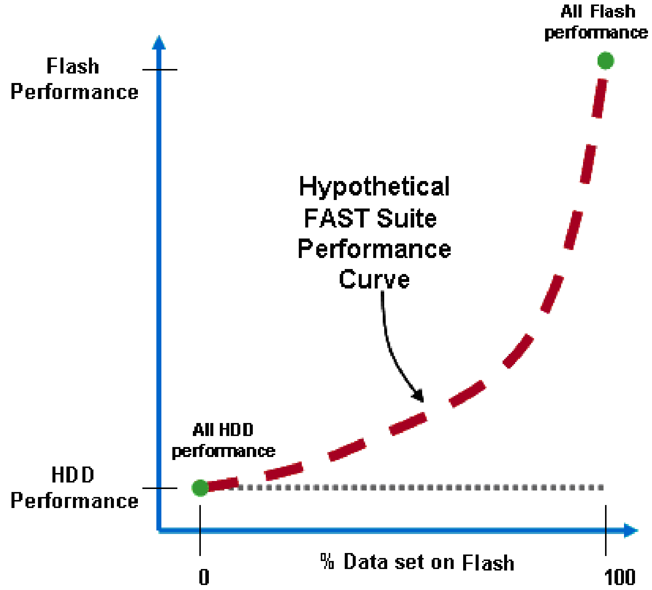

Figure 3. Performance benefit from the FAST Suite

The realized benefit of using FAST Suite is determined by several factors, including:

- The relative performance requirement differences of different applications
- The relative importance of data within an application dataset
- The location of data, determined by partitioning, temporal nature of data, and so on

It is very common to find these conditions in any typical data center. Users generally consolidate several databases onto a single storage system and different databases have different performance requirements. Most Oracle Online Transaction Processing (OLTP) databases tend to be highly random with small amounts of active data compared to the total database size. Any given database also tends to have varied data temperatures because of the following reasons:

- The OLTP databases generally tend to be temporal where the most recent data is more important than older data. This most recent data is generally referred to as a working set of a database.
- The relative importance of data changes from object to object. Some tables may be accessed more often than others.
- The number of I/O operations per second (IOPS) per gigabyte of an object, also known as object access density, changes quite significantly between objects. The best example is an index compared with a table in a database. The relative IOPS received by database blocks occupied by an index object are very high compared to the IOPS received by database blocks occupied by a table object.

Figure 4 shows the total read hits received by a 250 GB LUN at a 1 GB granularity in a 15-minute interval from an Oracle benchmark environment.

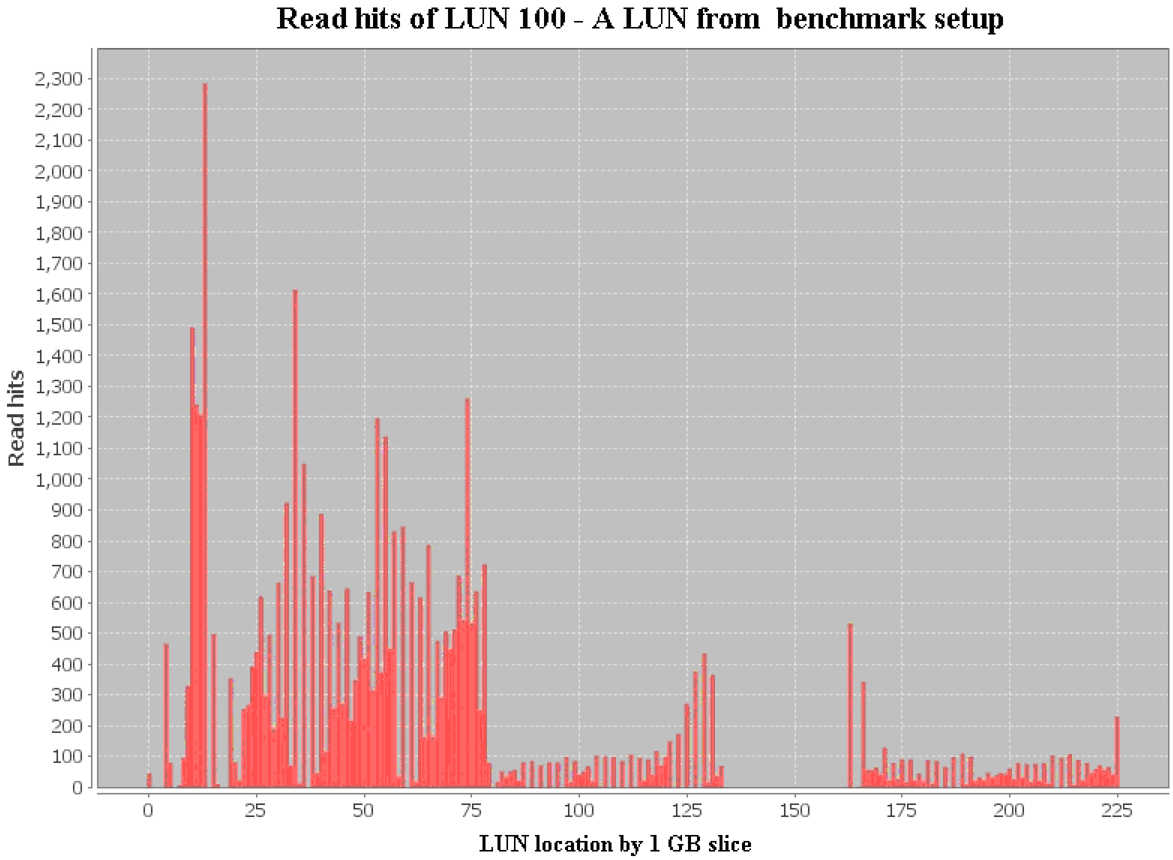

Figure 4. Slice hit chart from benchmark environment

Similarly, Figure 5 shows the total read hits received by a 600 GB LUN at a 1 GB granularity in 15 minutes from an Oracle customer environment.

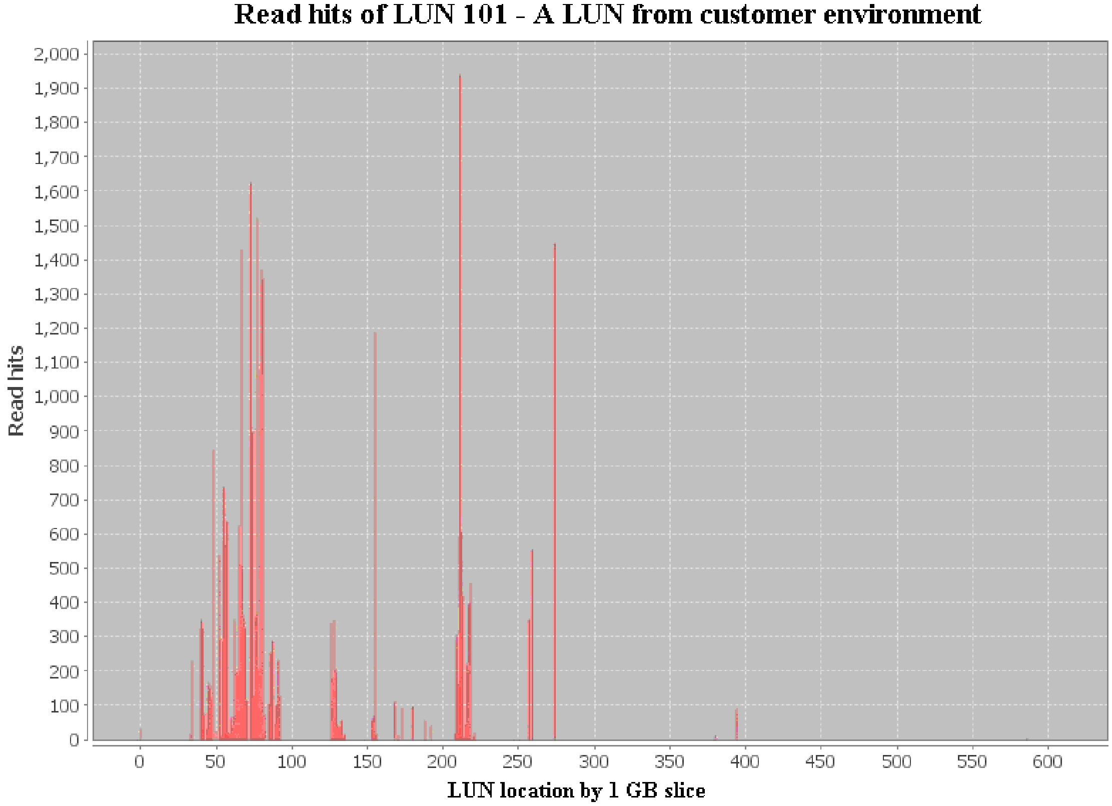

Figure 5. Slice hit chart from a customer environment

Figure 4 and Figure 5 illustrate that only some parts of the LUN receive most of the I/O. You can also see that the LUN from a customer environment has relatively less active data than a LUN from a controlled benchmark environment. This clustered nature of I/O at certain parts of a LUN is called *locality of reference*. The application realizes a sizable performance benefit as long as this small amount of hot data is migrated to a faster storage medium like Flash drives.

Fully Automated Storage Tiering for Virtual Pools (FAST VP)

Traditional tiering approaches have known limitations. The next few paragraphs discuss some of the limitations and how FAST VP solves them with virtual pools of storage and sub-LUN tiering technology.

Due to the mechanics of rotating disk drives, users have limited possibilities for sharing spindles across several workloads, thereby achieving poor capacity utilization, which also reduces storage consolidation ratios. Although using Flash drives has many advantages, users faced challenges in implementing these drives as a tier in the past:

- Users were required to move the entire dataset to the Flash-based tier to take full advantage of these drives.

- Users were required to identify hot data to move to Flash drives. This process was not only complex, but also repetitive because of the following factors and trends:
  - Data temperature constantly changes, so what is “hot” changes with time.
  - Massive data consolidation is driven by virtualization.
  - The size of customer datasets is growing by 60 percent every year.

As long as databases were small enough to fit on available Flash drives, these drives provided the best possible performance for database applications. Once the database outgrew the space available in those drives, database and storage administrators were faced with the daunting task of identifying hot and cold data in order to redistribute data to the right tiers. The Oracle Database provides advanced partitioning features to logically partition the data for easy management and to leverage various tiers of storage. In some cases, however, the database partitioning feature alone cannot manage the entire application data for these reasons:

- Database applications are also associated with a large amount of unstructured data that may be stored outside of the database, and may be accessed by the application through a separate location link.
- The data within a database partition can also have varied temperatures.

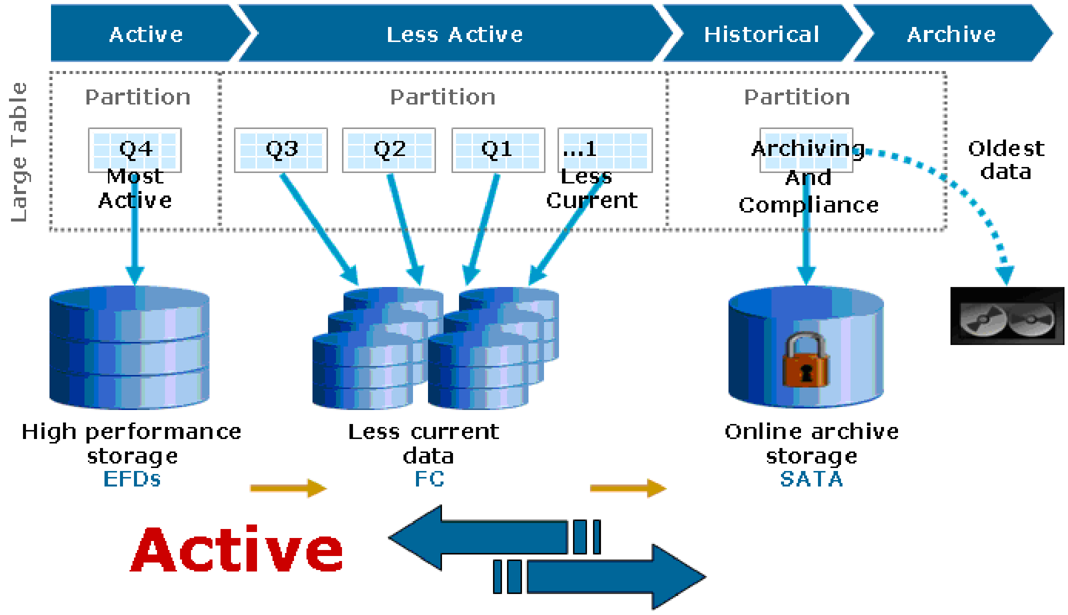

Figure 6. Traditional tiering approach

The traditional data tuning exercise shown in Figure 6 is a repetitive process, often using a scheduled maintenance window, depending on the technology used.

The dynamic nature of today’s enterprise applications makes manual tuning and balancing ineffective over time. Figure 7 shows an example of how enterprise applications’ data usage patterns can change over time after a tuning exercise.

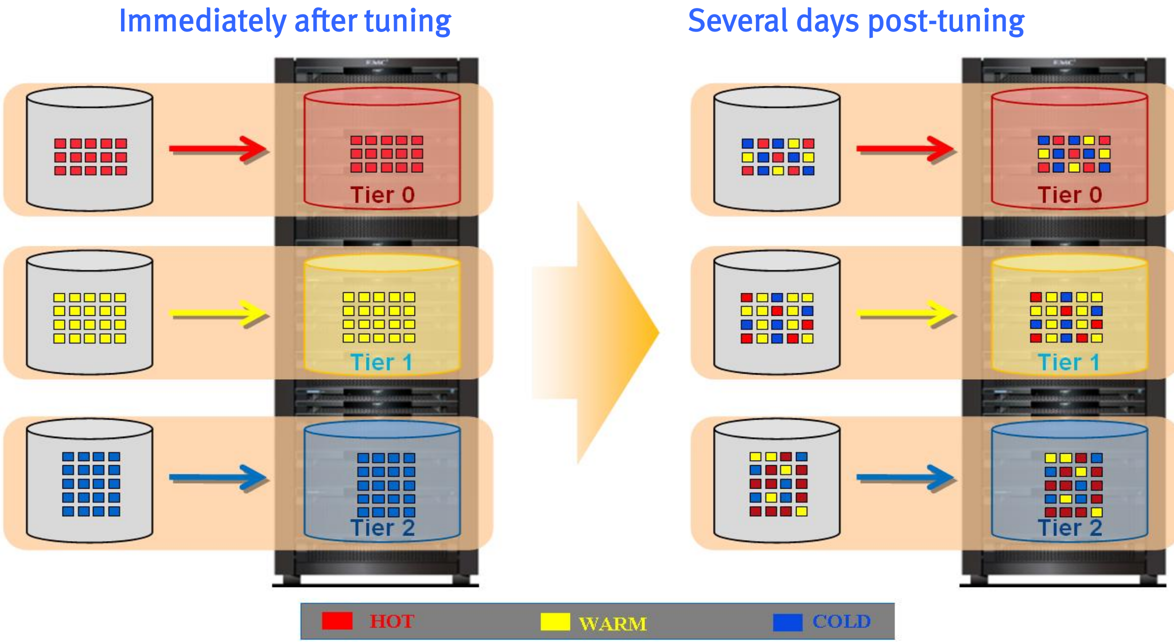

Figure 7. Data temperature changes over time

FAST VP overcomes these problems with the following important features:

- Users can now create blended storage pools composed of various disk types (Flash, SAS, and near-line SAS or NL-SAS). In a highly consolidated virtualized environment this leads to the highest storage efficiency, both from a performance and capacity perspective, since all the drive spindles can be shared efficiently between applications.
- The storage pools can grow non-disruptively as user demands for capacity or performance increase, simply by adding the required storage tier to the pool.
- The storage pools already implement the known best practices for achieving the highest performance from a given set of drives. This eases the burden on storage and database administrators to decide on the right storage layout.
- FAST VP tracks the sub-LUN temperatures at a 1 GB granularity. In turn, these data slices are automatically migrated to the appropriate tiers within a pool depending on temperature. The data migration is completely application-independent, non-disruptive, and can be controlled by simple user-defined policies.
- The tiering policy is very flexible with configuration options for data migration schedules and LUN level tiering preferences. Users have options to set the LUN tiering policy to:
  - Highest Available Tier – Use highest storage tier available in pool
  - Lowest Available Tier – Use lowest storage tier available in pool
  - Auto–Tier – Optimize the data slice location based on their temperature
  - No Data Movement – Do not move the slices of this LUN

Figure 8 compares a traditional repetitive manual tiering process with EMC's FAST VP technology.

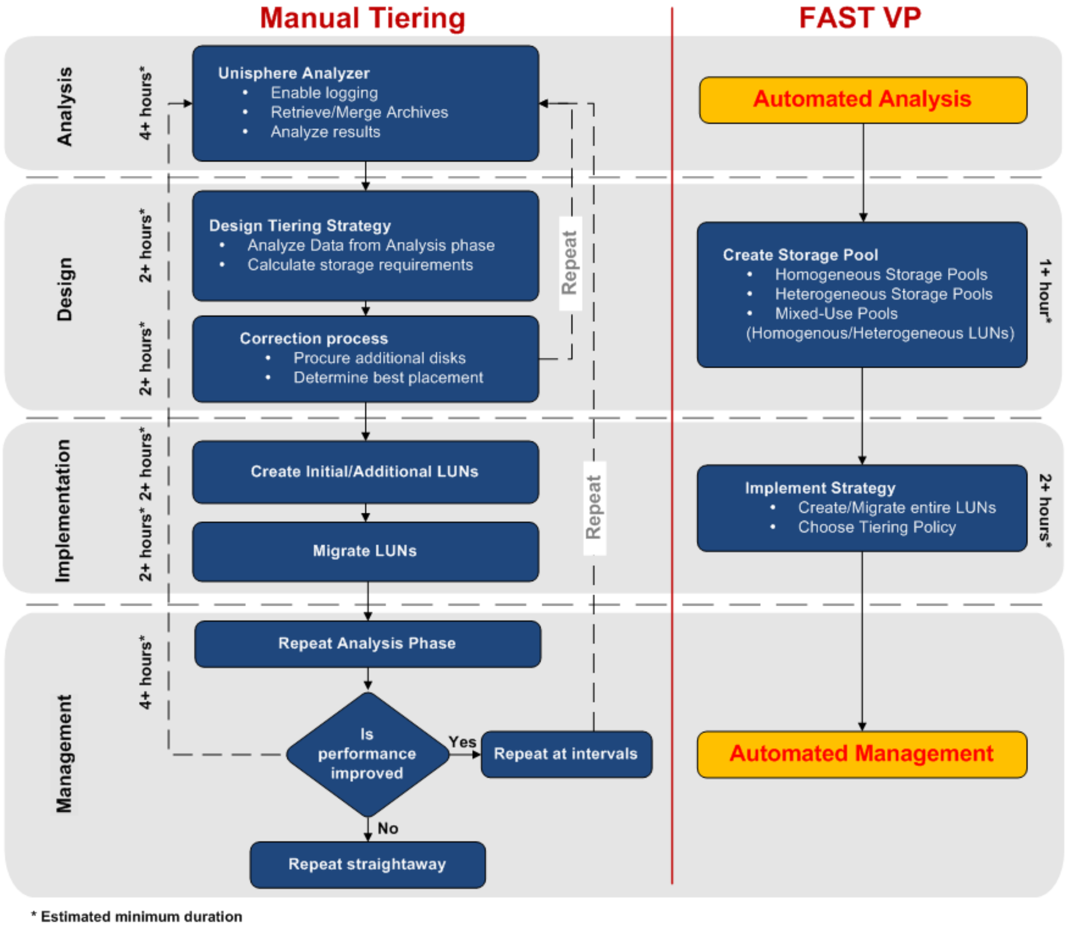

Figure 8. Comparing traditional tiering with FAST VP

A complete introduction to the FAST VP technology is beyond the scope of this document. Please refer to the References section for a list of more detailed documents.

FAST Cache

FAST VP solves problems with traditional tiering approaches and can deal with data temperature changes at a macro level or over long periods of time. FAST VP technology makes data migration decisions based on *historic data temperature* information. In a highly consolidated data center environment, sudden changes or small bursts in data access patterns can occur. The scheduled reaction time of FAST VP may not always be sufficient to handle these sudden surges in data access patterns over small periods of time. FAST Cache, a logical extension to DRAM cache of the storage system, bridges this gap. FAST Cache tracks the data temperature at a 64 KB granularity and copies hot data to a collection of designated Flash drives. Once FAST Cache caches the data, Flash drives handle subsequent access to the promoted data, thereby providing extremely low latency for that data. As the temperature of the copied data decreases over time, the data is evicted from FAST Cache, making room for new, hot data.

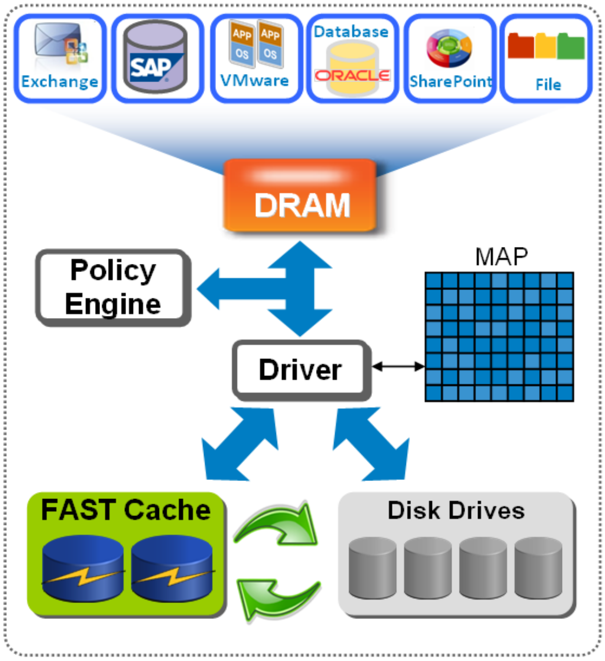

Figure 9. FAST Cache architecture

FAST Cache includes the following features:

- FAST Cache is read/write in nature, so it also provides acceleration to data overwrites.
- FAST Cache tracks data temperature at a 64 KB granularity, which complements the 1 GB slice granularity of FAST VP.
- FAST Cache can be enabled/disabled at a storage pool level.
- Users can create FAST Cache by selecting the set of Flash drives that are going to be part of FAST Cache without requiring any storage system downtime.
- FAST Cache is created on Flash drives, which can be accessed by both storage processors. This approach has several advantages:
  - FAST Cache does not consume any extra PCI-E slots inside the storage processor. This also makes it possible to implement efficient read/write cache.
  - In the event of a failure on one storage processor, the surviving storage processor can continue to access the FAST Cache through a shared SAS back end.
  - Upon power restoration following a complete power outage, both storage processors can reload the FAST Cache metadata from a persistent medium instead of repopulating it from scratch by observing the data access

patterns again. Therefore, no long warm-up time is required after a power outage.

FAST VP and FAST Cache can be used alone or together to achieve the maximum possible benefits. FAST Cache uses a memory bitmap to track the hit rates of data at 64 KB granularity. This limits the size of the FAST Cache that a given storage system can support. By combining both FAST Cache and FAST VP, users can scale their Flash tier to any size their applications require, a unique capability in the midrange storage market.

With FAST Suite, users have the following options:

- Using Flash drives as a separate manual tier of storage by themselves (that is, a homogeneous pool or stand-alone RAID group)
- Using Flash drives as an automated tier within a pool of storage using FAST VP
- Using Flash drives as a logical extension to DRAM cache using FAST Cache

The FAST Suite and FAST VP algorithms also work together to prevent data from being stored on a Flash drive twice. Whenever data is in pools or RAID groups, FAST Cache does not cache that data if it is already residing on Flash drives.

Applying the FAST Suite to Oracle Database

EMC Application Performance Labs conducted several tests to characterize the impact of the FAST Suite on Oracle Database application performance. This section describes in detail the workload used, the benchmark setup and use cases tested.

Workload description

EMC used an OLTP database workload that simulates a typical Order Entry system as a benchmark. Generally, most OLTP databases tend to have a read/write mix of 80/20. To demonstrate that the FAST Suite can handle workloads involving a high write mix, EMC created an aggressive test scenario with a 60/40 read/write mix. The Oracle OLTP database benchmark used in the test has the characteristics shown in Table 1.

Table 1. Workload characteristics

Property	Value / Description
Database size	1.0 TB
Database version	Oracle Database 11g Release 2 single instance
Storage type	Oracle Automatic Storage Management (ASM)
Storage container size	2.0 TB – ASM disk group created on 8 x 250
Read/write ratio	60 / 40
Database metric	Transactions per minute (TPM)
Working set	Around 200 to 300 GB of data receiving most of the I/O
Number of users	150

EMC kept consistent workload characteristics for all the use cases described next so users can analyze the application metrics.

Benchmark setup

To characterize FAST Suite with Oracle OLTP workloads, EMC used the following benchmark setup: The benchmark testing was done on a single-instance Oracle Database 11g Release 2 database deployed on an EMC VNX5300™ storage system and a Dell R910 server. The hardware setup is shown in Figure 10.

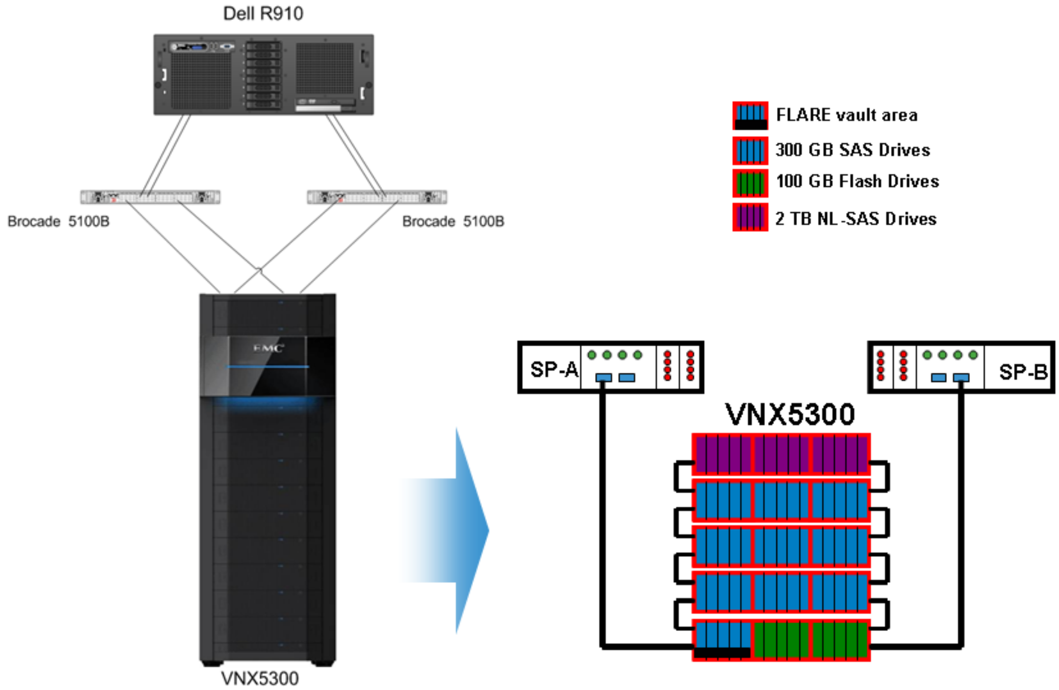

Figure 10. Benchmark setup

Use case description

EMC tested the following use cases to understand the impact of the FAST Suite on Oracle Database applications, especially to understand the impact of FAST Suite on application-level and database-level metrics. The main goal of these use cases was to show that EMC's FAST Suite can significantly reduce the number of IOPS rotating spindles received once the hot data migrates to Flash drives, thereby delivering significant improvement to the application-level metrics. The following use cases were tested:

- **Baseline**: The baseline metric is established on 45 x 15k rpm 300 GB SAS drives
- **FAST Cache use case**: Added FAST Cache to the configuration
- **FAST Suite use case**: Added both FAST Cache and FAST VP to the configuration

Baseline

The baseline is established on an all-SAS rotating drive configuration. This setup uses a total of 45 x 300 GB 15k rpm SAS disks with Oracle data files deployed on a 40-drive virtual storage pool and database online redo logs on a 5-drive storage pool. Figure 11 shows the exact database and Oracle ASM layout used.

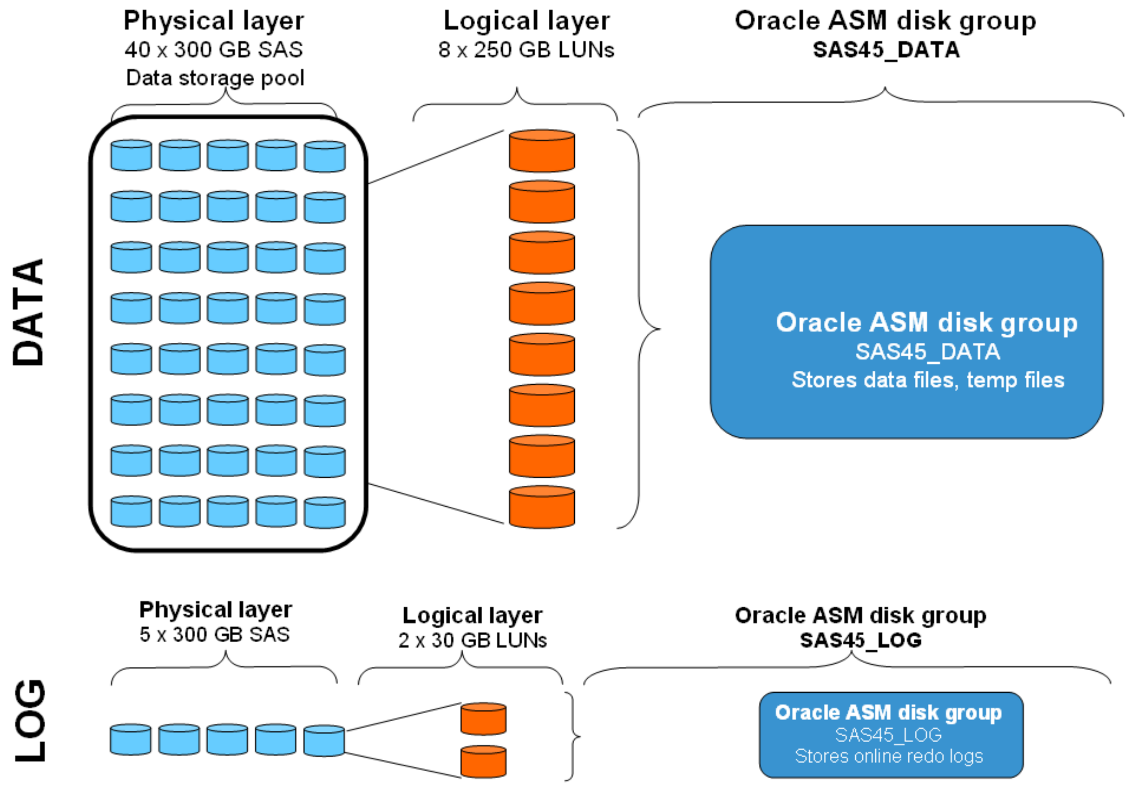

Figure 11. Baseline disk and database layout

FAST Cache use case

The database layout in this use case is identical to the baseline scenario except that FAST Cache is created on 8 x 100 GB Flash drives and is enabled on the data storage pool. FAST Cache is not enabled for the online redo log storage pool since it would not benefit it. In this configuration, once FAST Cache cached the frequently used data, the I/O latencies improved significantly, which resulted in enhanced overall application performance. This particular configuration delivered 5.6 times the number of transactions at one-fifth the latency when compared to the initial baseline configuration without FAST Cache. This is a significant return on a small FAST Cache investment.

A close analysis of the performance data from underlying rotating spindles revealed that once FAST Cache cached the hot data, the 40 rotating drives containing Oracle data files receive less than 15 IOPS each. A lower number of SAS drives or even NL-SAS drives can meet such low per-drive IOPS requirements. Figure 12 shows the reduction in IOPS count as FAST Cache caches more and more hot data.

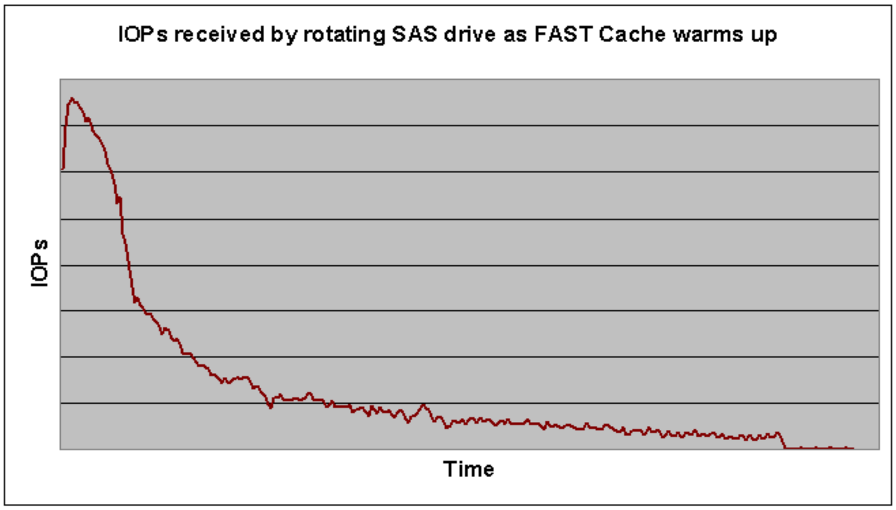

Figure 12. Rotating SAS drive IOPS as data gets promoted to FAST Cache

This test was repeated by creating the data storage pool on just 15 SAS drives instead of 40 drives to determine if FAST Cache still delivers the same improvement when the original database is created on fewer rotating drives. The layout shown in Figure 13 was used for this modified test. Please note that no changes were made to the database log volume.

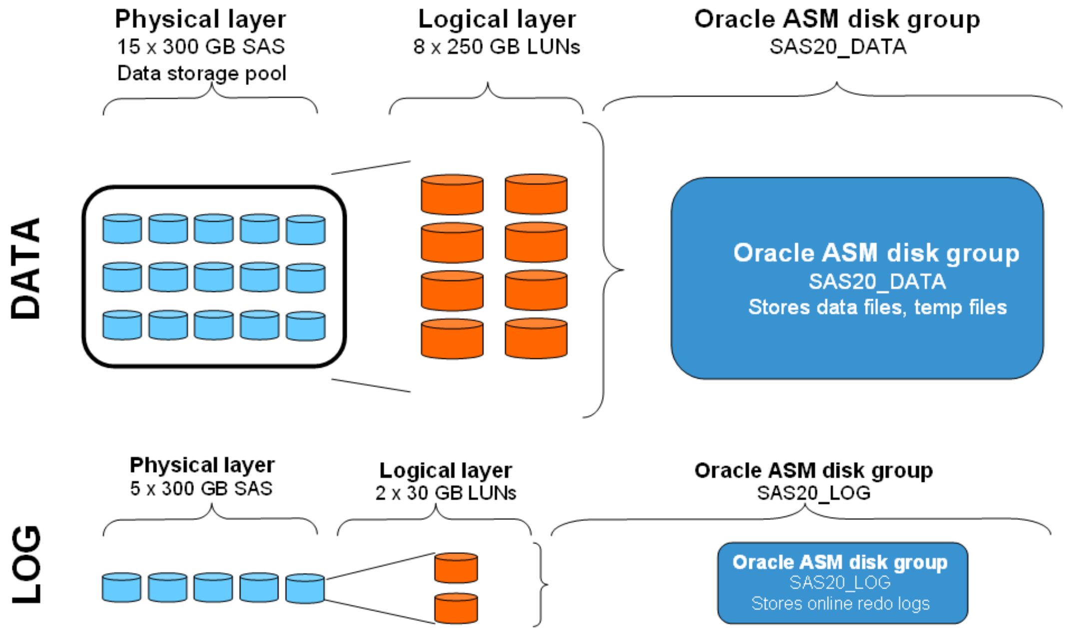

Figure 13. Database layout with reduced number of spindles

FAST Cache was enabled on the newly created 15-drive data storage pool. Once FAST Cache cached the hot data, the new pool with a reduced number of drives yielded almost the same level of performance as that of the 40-drive data pool. Figure 14 shows the improvement in both cases compared to the baseline.

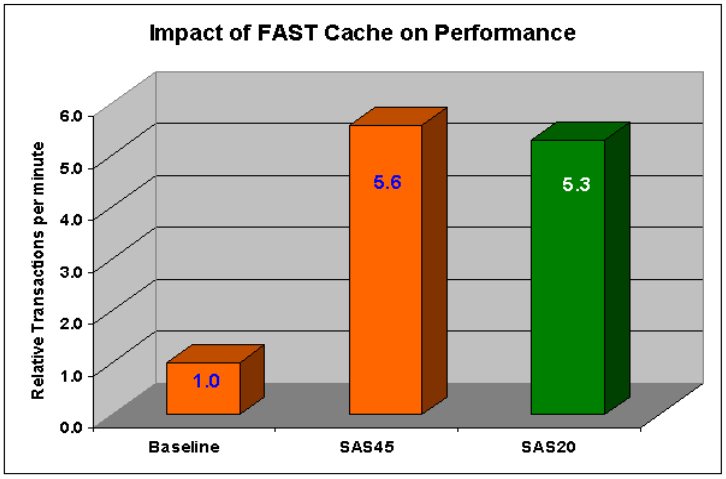

Figure 14. Database-level improvements from FAST Cache

In Figure 14 SAS45 represents the performance of the 45 (40 data, 5 log) SAS drive configuration with FAST Cache on 8 x 100 GB Flash drives, and SAS20 represents the performance of the 20 SAS drives (15 data, 5 log) with the same size FAST Cache. Both configurations yielded almost the same performance even though the database resides in fewer spindles in the SAS20 configuration. The reduction in number of spindles directly translates to a reduced data center footprint, and power and cooling savings. Table 2 shows the corresponding Oracle-level statistics from the Oracle AWR report.

Table 2. Oracle AWR statistics for FAST Cache use case

Baseline - 45 SAS drives

Event	Waits	Time(s)	Avg wait (ms)	% DB Time	Wait Class
db file sequential read	80491412	1994125	25	90.7	User I/O
db file parallel read	2183756	146736	67	6.7	User I/O
log file sync	5019183	19555	4	0.9	Commit
CPU time		12052		0.5	
log file parallel write	3780655	9406	2	0.4	User I/O

SAS45 - 45 SAS + 8 Flash drives

Event	Waits	Time(s)	Avg wait (ms)	% DB Time	Wait Class
db file sequential read	416575329	1887858	5	87.5	User I/O
log file sync	24893532	96776	4	4.5	User I/O
db file parallel read	10945247	85244	8	3.9	Commit
CPU time		61882		2.9	
log file parallel write	5749730	13148	2	0.6	User I/O

SAS20 - 20 SAS + 8 Flash drives

Event	Waits	Time(s)	Avg wait (ms)	% DB Time	Wait Class
db file sequential read	301012478	1878339	6	86.4	User I/O
db file parallel read	7922448	146224	18	6.7	User I/O
log file sync	18032281	65746	4	3	Commit
CPU time		44604		2.1	
log file parallel write	5652509	12026	2	0.6	User I/O

Before you reduce the number of underlying spindles or change the type of the drives is a possibility, you should complete a thorough analysis to determine the right number of spindles. Underestimating the number of rotating drives can impact new data ingest rates since the new data always starts at the rotating drive before it is promoted to FAST Cache.

FAST Suite use case

This use case leverages both FAST VP and FAST Cache. This configuration contains a data storage pool created on 5 Flash drives, 40 SAS drives, and 15 NL-SAS drives. The Oracle online redo log pool is created on 5 SAS drives. Rather than completely tearing down and rebuilding the benchmark setup, the new configuration was created by adding 5 Flash drives and 15 NL-SAS drives to the already existing 40-drive SAS data storage pool. As discussed earlier, the storage pools can be dynamically grown non-disruptively as requirements for capacity and performance increase. When new drives are added to a pool, FAST VP makes data migration recommendations based on historic data temperature information it already has on that storage pool. Users can either wait until the next scheduled data migration window for data movement or can implement the recommendations immediately by starting the data relocation operation manually on that pool. In this use case, a manually initiated FAST migration was started to implement the recommendations immediately. Once all the hot data was migrated to the Flash tier inside the pool, a FAST Cache was also created on 2 x 100 GB Flash drives and was enabled on the data storage pool. FAST Cache was not enabled on the online redo log storage pool since it would not be beneficial. A small FAST Cache configuration was sufficient to handle leftover hot data in this use case. Five Flash drives already existed in the pool and a majority of the hot data should have migrated to these Flash drives during the manually initiated migration step discussed earlier. The modest FAST Cache handled sudden intra-run spikes over small data ranges. Figure 15 shows the exact storage layout for this configuration.

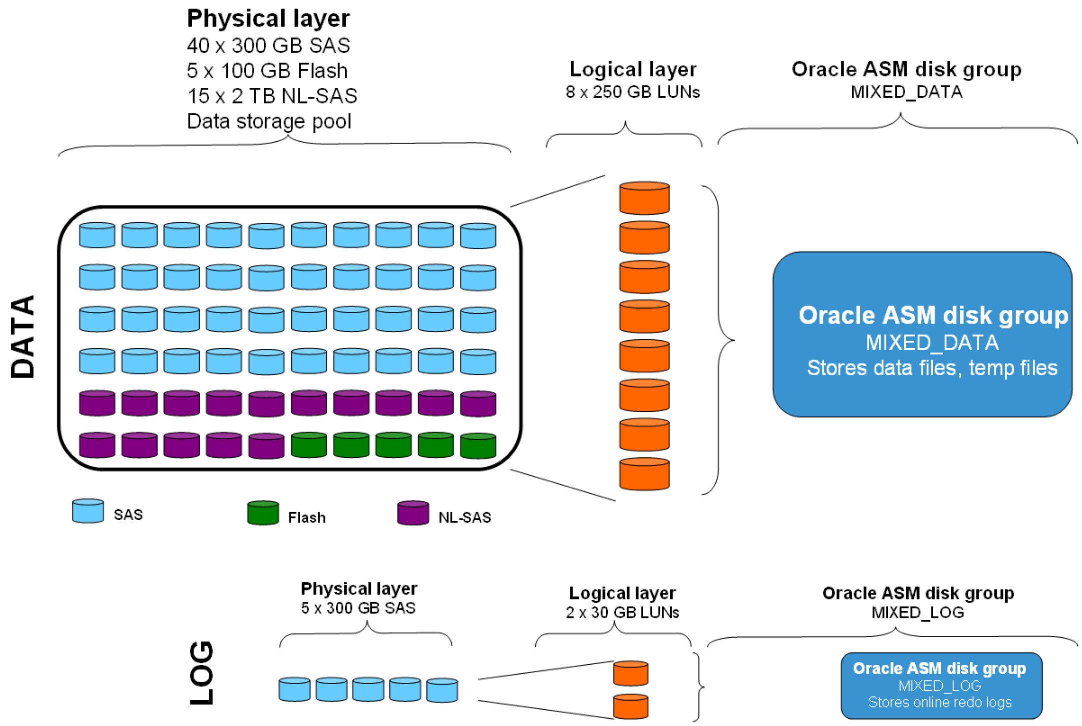

Figure 15. Database layout for the FAST Suite use case

Once FAST Cache cached the remaining hot data, the configuration yielded almost 5.2 times more transactions at roughly one-fifth the latency when compared to the initial baseline configuration without FAST Suite. This is a tremendous improvement given the small number of Flash drives added to the configuration. Figure 16 and Table 3 show the application-level improvement achieved by deploying FAST Suite, and the corresponding Oracle AWR statistics when compared to the baseline, respectively.

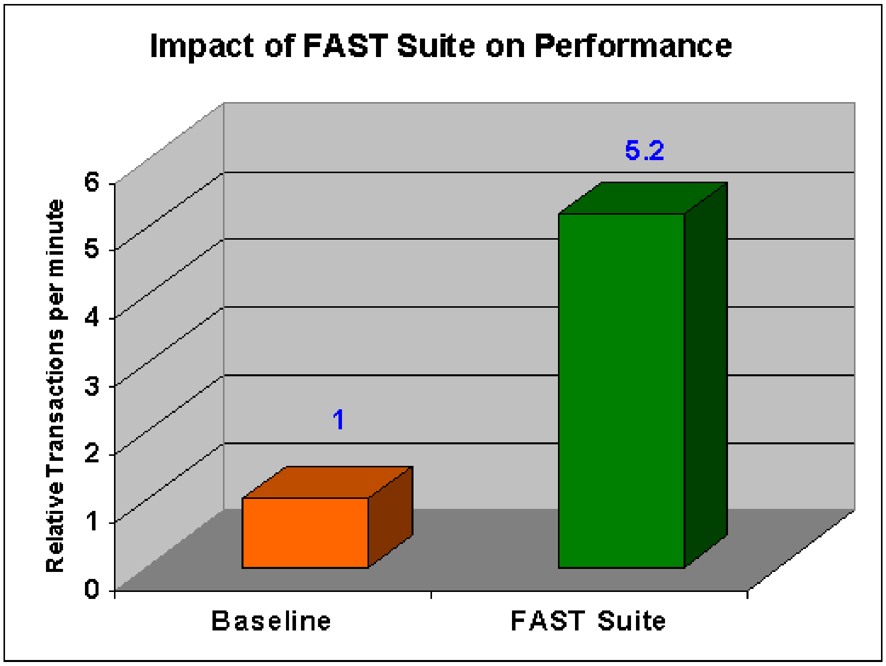

Figure 16. Database-level improvements from FAST Suite

Table 3. Oracle AWR statistics for the FAST Suite use case

Baseline - 45 SAS drives

Event	Waits	Time(s)	Avg wait (ms)	% DB Time	Wait Class
db file sequential read	80491412	1994125	25	90.7	User I/O
db file parallel read	2183756	146736	67	6.7	User I/O
log file sync	5019183	19555	4	0.9	Commit
CPU time		12052		0.5	
log file parallel write	3780655	9406	2	0.4	User I/O

FAST Suite - 45 SAS, 5 Flash, 15 NL-SAS drives in a pool + 2 x 100 GB FAST Cache

Event	Waits	Time(s)	Avg wait (ms)	% DB Time	Wait Class
db file sequential read	405599593	1923059	5	88.4	User I/O
log file sync	23708098	87017	4	4	User I/O
db file parallel read	10460316	81816	8	3.8	Commit
CPU time		59481		2.7	
log file parallel write	6350130	13343	2	0.6	User I/O

Frequently asked questions about FAST Suite and Oracle

Do Oracle ASM and the FAST Suite work together?

Yes, Oracle ASM and the FAST Suite technologies complement each other. Oracle ASM does a very good job of distributing the data within a disk group among all of its associated LUNs including the hot data. FAST Suite takes it one step further by moving the hot data slices from within these equally busy LUNs to a faster storage medium while still maintaining the uniform temperature distribution. It is important to note that all testing and benchmark results discussed here have been performed on an Oracle Database 11g Release 2 single instance deployed on Oracle ASM.

What FAST Cache size is good for my application?

The FAST Suite applicability section touches on this subject to some extent. You would need to determine the hot data set of an application. This is a slightly complex process depending on your current deployment and tools. Here are some of the tools options:

- **Application-level tools:** Database engine-level performance statistics can provide a rough idea about the hot working set of an application. For example, Oracle AWR reports have table space-level and data file-level I/O statistics that can provide information about hot data containers. Oracle V$ tables, like V$SEGMENT_STATISTICS, can provide information about hot database segments.
- **Third-party tools:** Some third-party tools from Precise and ZettaPoint can help determine the hot data set of an Oracle Database application.
- **Storage system level tools:** If the current deployment is on an EMC CLARiiON®, Celerra® unified NS, or VNX series storage system, array-level tools are available to EMC field and support personnel for determining the right size FAST Cache. Contact your EMC sales teams for guidance.

VNX users can also follow some rules of thumb to determine the right FAST Cache size. Generally most of the OLTP databases have 3 percent to 5 percent of active data compared their capacity. It is very rare for database applications to have more than 10 percent active data. So, sizing FAST Cache as 5 percent to 10 percent of the total database size should handle the requirement for more than 90 percent of the existing databases. It is very important to note that sizing for just capacity is not sufficient; also follow the performance guidelines for sizing the number of Flash drives based on parity level and application read/write ratio. The FAST Cache is R/W in nature and uses RAID 1 pairs to protect the writes that have not yet been copied to the HDDs in the event of a disk failure. The following example shows how to size FAST Cache for a typical Oracle OLTP database.

Table 4. Requirements

Requirement	Value / Description
Database size	1.0 TB
IOPS requirement	10,000
Read to write ratio	80/20

By using the simple 10 percent rule, based on the requirements listed in Table 4, the size of the FAST Cache would be just 100 GB (that is, 10 percent of 1.0 TB) or a pair of 100 GB Flash drives; however, this may not be sufficient to meet the IOPS requirement. The back-end storage IOPS requirement depends on RAID type and the read/write ratio. The sizing for the above requirements is shown in Table 5.

Table 5. Sizing example

Property	Value / Description
Front-end IOPS	10,000 IOPS
Read to write ratio	80/20
RAID type for FAST Cache	RAID 1 (RAID 1 is required to allow R/W cache)
Back-end IOPS for RAID 1	= Reads + 2 x Writes = 0.8 * 10000 + 2 * 0.2 * 10000 = 12000 IOPS
Number of Flash drives	= 12000 / 3500 ~= 4

The example indicates how the Flash drive requirement for FAST Cache can be more when considering the performance requirement compared to considering only the capacity requirements. EMC recommends using 3500 IOPS per Flash drive as a rule of thumb for sizing. Although this number appears low compared to the information in most Flash drive information available online, this is still more than 17 times the performance of a single rotating 15k rpm drive. Figure 17 shows the impact of the FAST Cache size on application performance.

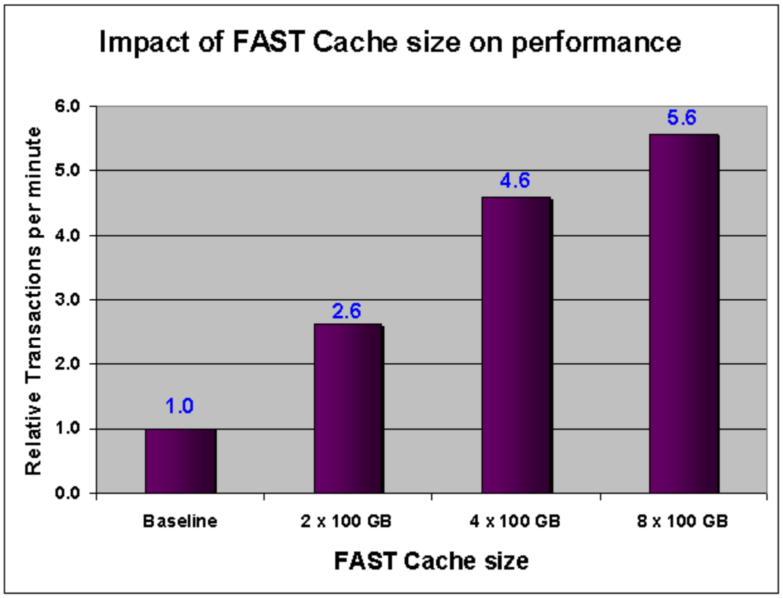

Figure 17. Impact of FAST Cache size

Adding more Flash drives to FAST Cache than is required may not yield any incremental benefit. Once FAST Cache caches the active working set of the database, adding more drives will simply spread the same data/IOPS over the newly added drives, effectively driving fewer IOPS from each Flash drive. The data importance curve in Figure 18 represents a hypothetical curve that shows the percentage of IOPS that a 1.0 TB Oracle Database received. The first 200 GB of data receives 69 percent of the I/O, the second 200 GB receives only 20 percent of the I/O, and the remaining 600 GB receives remaining 11 percent of the I/O. In this example, the first 200 GB of FAST Cache receives the maximum benefit and adding more FAST Cache beyond that yields a relatively lower incremental benefit. In fact, the real-world data importance curves will be even steeper where less than 10 percent of the data receives almost 90 percent of the I/O.

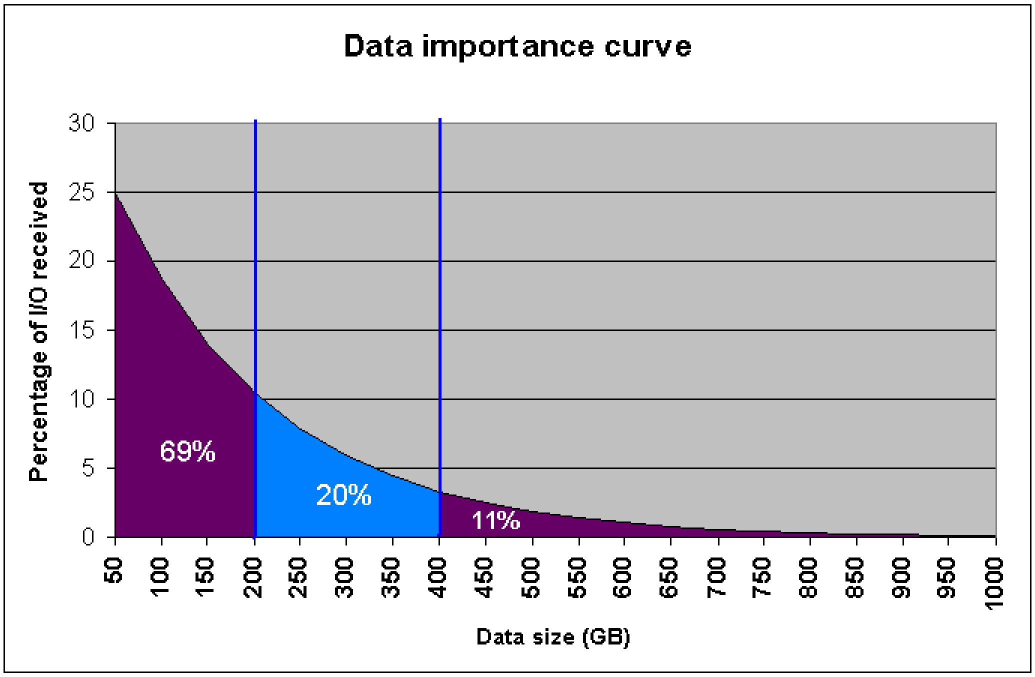

Figure 18. Data importance curve

Does the FAST Suite benefit Oracle log files?

Almost all experiments done by EMC Application Performance Labs indicate that using Flash drives in any form to support Oracle online redo logs and archive logs is not going to be very cost-beneficial. This does not mean that Flash drives cannot be used with database log files at all. However, it is more beneficial to deploy them with other database objects like tables, indices, undo table space, or even temporary table space before using them for online redo logs or archive logs. The same recommendations also apply to FAST Cache and FAST VP. The I/O pattern on Oracle Database logs files, both online and archive, tend to be highly sequential in nature. The archive logs have very large sequential I/O patterns that rotating drives can handle very efficiently. Moreover the archive files are not generally read back unless database or table space recovery is necessary. Similarly, the online redo log average

I/O size may vary depending on the database transactional boundaries, but will still be sequential. The write cache of the storage system can easily coalesce these small writes into bigger back-end I/O stripes. Therefore, a few rotating drives can handle even the online redo logs so that Flash drives can be deployed elsewhere. EMC's general recommendation is *not to use* either FAST Cache or FAST VP for Oracle archive files and online redo log files.

Can the data and logs be deployed inside the same pool?

EMC does not recommend placing Oracle data files and log files in the same pool for the following reasons:

- **Reliability:** The transaction logs play a very pivotal role in Oracle Database recovery. In the event of data file corruption, the database administrator can go back to an older copy of the data file and apply the logs. Similarly if logs are lost, the Oracle Database can guarantee zero or minimal data loss if online redo logs are multiplexed to different sets of spindles. By putting data and logs into the same pool, the fundamental best practices of fault domains is ignored. Unless the user has some other data recovery plan like CDP (continuous data replication), EMC does not recommend putting both data and logs in the same pool.
- **I/O type and size:** The I/O profile of log files tends to be highly sequential. By mixing log files with data files, the sustained write bandwidth of a drive drops as the spindle begins to seek more often.

FAST VP by itself does not impose any restrictions for placing data and logs together. If users still want to deploy both data and logs within the same pool, EMC recommends setting the tier preference of log LUNs to "Lowest Available Tier" in a mixed pool scenario, since the drives in the lowest tier of a mixed pool receive very low traffic and can handle sequential I/O very efficiently.

What is the right database layout to use in a highly consolidated environment?

In a typical data center, it is rare that just one application is deployed on any given storage system. The storage systems themselves are becoming very powerful, like their server counterparts, and can be used to consolidate several workloads. The traditional method of creating isolated sets of application-centric RAID groups or smaller pools increases the complexity of storage deployment. At the same time storage and database administrators want to guarantee service-level agreements (SLAs) to their application owners. FAST VP with its various configuration options can solve this problem and ease storage deployment complexity. Rather than creating several small single-purpose storage containers, users can use FAST VP to create fewer multipurpose pools but still guarantee SLAs, taking into consideration the criticality, fault domains, and I/O characteristics of the data.

Figure 19 shows a possible layout scenario for a highly consolidated Oracle Database deployment with tens of databases.

Oracle VNX Layout			
	Pool 1	Pool 2	Pool 3
Application			
Data Type	Redo Logs	Data	Archive
FAST Suite Policies			
FAST Cache	No	Yes	No
FAST Policies	No DM	Prod. = Auto Tier	No DM
FAST Policies		Dev. = Lowest Tier	
Tier 1			
Drive Type	15K SAS	Flash	
Raid Protection	R 1/0, R5	R5,R6	
Tier 2			
Drive Type		15K SAS	
Raid Protection		R5,R6	
Tier 3			
Drive Type		7.2K NL-SAS	7.2K NL-SAS
Raid Protection		R5,R6	R5,R6

Figure 19. Database layout for a consolidated Oracle environment

In the layout in Figure 19, the mixed/homogeneous storage pools are created based on the I/O type and their fault domain requirements. There are three distinct pools and some simple provisioning policies are defined to ensure some level of SLAs. For example, a production database has the tiering preference set to "Auto-Tier" on the data LUNs compared to "Lowest Available Tier" for the LUNs in the test/dev databases. FAST Cache is also enabled appropriately depending on the data container. Whenever a new database requires storage, the storage administrator simply provisions a few LUNs from Pool1, Pool3, and Pool2 for database online redo logs, archive logs, and everything else, respectively. This guideline can cover many database deployments. The exact layout requirements will change from deployment to deployment based on the exact customer requirements.

Conclusion

EMC VNX series with its massive scalability and flexibility is an ideal midrange storage system for any virtualized data center. The VNX series is a true unified box with block, file, and object storage capabilities that a simple but intuitive EMC Unisphere™ management interface can manage. By leveraging advanced VNX storage functionality such as FAST VP and FAST Cache, users can achieve increased performance and also reduce TCO for their Oracle deployments. All use cases discussed in this paper prove that just by deploying a few Flash drives and leveraging the FAST Suite, users can significantly reduce the total number of drives required for any Oracle Database implementation, while maintaining and even improving performance. With its advanced data features, the VNX series not only reduces the initial cost of the deployment but also significantly reduces the complexity associated with day-to-day data management by automating the complex and time-consuming storage tiering process.

The EMC FAST Suite complements Oracle Database's advanced storage features like ASM and partitioning. These three technologies can be used together to achieve the best benefit and simplify data management of any enterprise Oracle Database deployment.

References

The following can be found on EMC.com and Powerlink:

- VNX Family Data Sheet
- EMC's VNX Family—A Unifying Force in Storage ESG white paper
- Deploying EMC VNX Unified Storage Systems for Data Warehouse Applications EMC white paper
- EMC Unified Storage Solutions: Oracle Database 11gR2 with EMC VNX and Storage Replication Consistency EMC white paper
- Using EMC FAST Suite with SAP on EMC Unified Storage EMC white paper
- EMC Unisphere: Unified Storage Management Solution EMC white paper