Abstract

Businesses are deploying multiple different disk drive technologies in an attempt to meet DB2 for z/OS service levels as well as to reduce cost. To manage these complex environments, it is necessary to utilize an automated tiering product. This white paper describes how to implement Fully Automated Storage Tiering with Virtual Pools (FAST™ VP) using EMC® Symmetrix® VMAX® with DB2 for z/OS.

September 2012
# Table of Contents

**Executive summary** .......................................................... 4  
  Audience ............................................................................. 4  

**Introduction** ................................................................. 4  
  Virtual Provisioning .......................................................... 5  
  Fully Automated Storage Tiering for Virtual Pools (FAST VP) ................................................. 6  

**DB2 testing** ...................................................................... 8  
  Overview .......................................................................... 8  
  Skew ................................................................................ 9  
  VMAX configuration ......................................................... 9  
  DB2 configuration ............................................................. 10  
  Workload ......................................................................... 10  
  FAST VP policies ............................................................ 10  
  Testing phases .................................................................. 10  
  Testing results .................................................................... 11  
    Run times ...................................................................... 12  
    Response times ............................................................ 12  
    Average IOPS .............................................................. 14  
    Storage distribution across the tiers .............................. 14  
  Summary ........................................................................... 15  

**Best practices for DB2 and FAST VP** ................................. 15  
  Unisphere for VMAX .......................................................... 16  
  Storage groups ................................................................. 16  
  FAST VP policies ............................................................ 16  
  Time windows for data collection ................................... 17  
  Time windows for data movement ................................. 17  
  DB2 active logs ............................................................... 17  
  DB2 REORGs ................................................................. 17  
  z/OS utilities ..................................................................... 18  
  DB2 and SMS storage groups ......................................... 18  
  DB2 and HSM ................................................................. 18  

**Conclusion** ...................................................................... 19  

**References** ....................................................................... 19
Executive summary

The latest release of the Enginuity™ operating environment for Symmetrix® is Enginuity 5876, which supports the Symmetrix VMAX® Family arrays, VMAX® 10K, VMAX® 20K, and VMAX® 40K. The capabilities of Enginuity 5876 to network, share, and tier storage resources allow data centers to consolidate applications and deliver new levels of efficiency with increased utilization rates, improved mobility, reduced power and footprint requirements, and simplified storage management.

Enginuity 5876 includes significant enhancements for mainframe users of the Symmetrix VMAX Family arrays that rival in importance to the original introduction of the first Symmetrix Integrated Cached Disk Array in the early 1990s. After several years of successful deployment in open systems (FBA) environments, mainframe VMAX Family users now have the opportunity to deploy Virtual Provisioning™ and Fully Automated Storage Tiering for Virtual Pools (FAST™ VP) for count key data (CKD) volumes.

This white paper discusses DB2 for z/OS and FAST VP deployments and measures the performance impact of using a DB2 for z/OS subsystem with FAST VP. It also includes some best practices regarding implementation of DB2 for z/OS with FAST VP configurations.

Audience

This white paper is intended for EMC technical consultants, DB2 for z/OS database administrators, mainframe system programmers, storage administrators, operations personnel, performance and capacity analysts, technical consultants, and other technology professionals who need to understand the features and functionality capabilities of the FAST VP implementations with DB2 for z/OS.

While this paper deals with the new features as stated, a comprehensive understanding of all of the mainframe features offered in Enginuity prior to this release can be gained by reviewing the EMC Mainframe TechBook, *EMC Mainframe Technology Overview*.

Introduction

FAST VP is a dynamic storage tiering solution for the VMAX Family of storage controllers that manages the movement of data between tiers of storage to maximize performance and reduce cost. Volumes that are managed by FAST VP must be thin devices.

In order to understand the implications of deploying a DB2 subsystem with VP and FAST VP, it is necessary to have a basic understanding of the underlying technologies. This introduction provides an overview of these technologies for readers unfamiliar with them.
Virtual Provisioning

Virtual Provisioning is a new method of provisioning CKD volumes within Symmetrix VMAX Family arrays. It is supported for 3390 device emulation and is described in detail in the white paper titled *z/OS and Virtual Provisioning Best Practices*.

Standard provisioning, also known as thick provisioning, provides host-addressable volumes that are built on two or more physical devices using some form of RAID protection. The fact that these volumes are protected by some form of RAID, and are spread across multiple disks, is not exposed to the host operating system. This configuration is depicted in Figure 1.

![Figure 1. Standard thick provisioning in Symmetrix VMAX Family arrays](image)

A virtual provisioned volume, that is a thin volume, disperses a 3390 volume image across many physical RAID-protected devices using small (12-track) units called track groups. These devices are protected by the same RAID protection as provided for normal thick devices and are organized into virtual pools (thin pools) that support a given disk geometry (CKD3390 or FBA), drive technology, drive speed, and RAID protection type.

Thin devices are associated with virtual pools at creation time through a process called binding, and can either be fully pre-allocated in the pool, or allocated only on demand when a write occurs to the volume. This configuration is depicted in Figure 2.
The dispersion of track groups across the disks in a pool is somewhat analogous to wide striping, as the volume is not bound to a single RAID rank but exists on many RAID ranks in the virtual pool.

The mapping of a device image to a virtual pool through the track group abstraction layer enables a concept called thin provisioning, which allows a user, who chooses not to pre-allocate the entire volume image, the option to present more storage capacity by way of the thin volumes than is actually present in the thin pool. Presenting more capacity on the channel than is actually in the pool is called over subscription, and the ratio the storage presented on the channel to the actual storage in the pool is called the over-subscription ratio.

Virtual Provisioning also provides these important benefits:
1. The data is effectively wide-striped across all the disks in the pool, thereby eliminating hot spots and improving overall performance of the array.
2. The array is positioned for active performance management at both the sub-volume and sub-dataset level using FAST VP.

**Fully Automated Storage Tiering for Virtual Pools (FAST VP)**

Fully Automated Storage Tiering for Virtual Pools is a VMAX feature that dynamically moves data between tiers to maximize performance and reduce cost. It non-disruptively moves sets of 10 track groups (6.8 MB) between storage tiers automatically at the sub-volume level in response to changing workloads. It is based on, and requires, virtually provisioned volumes in the VMAX array.

EMC determined the ideal chunk size (6.8 MB) from analysis of 50 billion I/Os provided to EMC by customers. A smaller size increases the management overhead to an unacceptable level. A larger size increases the waste of valuable and expensive Enterprise Flash drive (EFD) space by moving data to EFD that is not active. Tiering
solutions using larger chunk sizes require a larger capacity of solid-state drives which increases the overall cost.

FAST VP fills a long-standing need in z/OS storage management: Active performance management of data at the array level. It does this very effectively by moving data in small units, making it both responsive to the workload and efficient in its use of control-unit resources.

Such sub-volume, and more importantly, sub-dataset, performance management has never been available before and represents a revolutionary step forward by providing truly autonomic storage management.

As a result of this innovative approach, compared to an all-Fibre Channel (FC) disk drive configuration, FAST VP can offer better performance at the same cost, or the same performance at a lower cost.

FAST VP also helps users reduce DASD costs by enabling exploitation of very high capacity SATA technology for low-access data, without requiring intensive performance management by storage administrators.

Most impressively, FAST VP delivers all these benefits without using any host resources whatsoever.

FAST VP uses three constructs to achieve this:

- **FAST storage group**
  A collection of thin volumes that represent an application or workload. These can be based on SMS storage group definitions in a z/OS environment.

- **FAST policy**
  The FAST VP policy contains rules that govern how much capacity of a storage group (in percentage terms) is allowed to be moved into each tier. The percentages in a policy must total at least 100 percent, but may exceed 100 percent. This may seem counter-intuitive but is easily explained. Supposing you have an application that you want FAST VP to determine exactly where the data needs to be without constraints, you would create a policy that permits 100 percent of the storage group to be on EFD, 100 percent on FC, and 100 percent on SATA. This policy totals 300 percent. This kind of policy is the least restrictive that you can make. Mostly likely you will constrain how much EFD and FC a particular application is able to use but leave SATA at 100 percent for inactive data.

  Each FAST storage group is associated with a single FAST policy definition.

- **FAST tier**
  A collection of up to four virtual pools with common drive technology and RAID protection. At the time of writing, the VMAX array supports four FAST tiers.

  Figure 3 depicts the relationship between VP and FAST VP in the VMAX Family arrays. Thin devices are grouped together into storage groups. Each storage group is usually mapped to one or more applications or DB2 subsystems that have common performance characteristics. A policy is assigned to the storage group that denotes
how much of each storage tier the application is permitted to use. The figure shows two DB2 subsystems, DB2A and DB2B with a different policy.

**DB2A**
This has a policy labeled Optimization, which allows DB2A to have its storage occupy up to 100 percent of the three assigned tiers. In other words, there is no restriction on where the storage for DB2A can reside.

**DB2B**
This has a policy labeled Custom, which forces an exact amount of storage for each tier. This is the most restrictive kind of policy that can be used and is effected by making the total of the allocations equal one hundred percent.

More details on FAST VP can be found in the white paper *Implementing Fully Automated Storage Tiering for Virtual Pools (FAST VP) for EMC Symmetrix VMAX Series Arrays.*

![Diagram of DB2A and DB2B with Optimization and Custom policies](image)

---

**Storage Groups**

**Policies**

**Tiers**

*Figure 3. FAST VP storage groups, policies, and tiers*

**DB2 testing**

**Overview**
This section provides an overview of the DB2 for z/OS workload testing that was performed in the EMC labs to show the benefit of running DB2 transactions on z/OS volumes managed by FAST VP. The testing was performed using DB2 V10 and a batch transaction workload generator that generated high levels of random reads to the VMAX array.

DB2 workloads on a Symmetrix subsystem are characterized by a very high cache hit percentage, 90 percent and sometimes higher. Cache hits do not drive FAST VP algorithms since they cannot be improved by placing the associated data on
Enterprise Flash drives. So the testing simulated a 5 TB DB2 subsystem with a 90 percent cache hit rate. The 500 GB of cache miss data was created using a randomizing unique primary key algorithm that created 100 percent cache misses on the 500 GB. This simulated a DB2 subsystem with these characteristics, but without having to demonstrate 90 percent cache hits that would be irrelevant to this test.

**Skew**

Workload skew is found among nearly all applications. Skew happens when some volumes are working much harder than other volumes. Also, at a sub-volume level, parts of the volume are in demand much more than other parts. Based on analysis of many customer mainframe workloads, skew at the volume level is around 20/80: 20 percent of the volumes are doing 80 percent of the work. This proportion also applies at the sub-volume level. If you do the math, you can calculate that around four percent of the disk space is accounting for 96 percent of the workload. FAST VP exploits this skew factor by determining where this four percent is (or whatever the actual percentage is) and noninvasively relocating it to Enterprise Flash drives.

One important factor of skew is that if the systems that are being managed only have a small amount of capacity, the skew causes the data to be held in either the VMAX cache or in the DB2 buffer pool. For example, if the database is a small two terabytes database, four percent skew would result in an 80 GB working set. This can easily be held in VMAX cache, even on a small controller, and thus is not an appropriate application for FAST VP, unless there are many more applications running on the VMAX array.

It should be noted here that the one aspect of the testing did create skew is the unique index that was used to process the individual singleton SELECTs. This index was 32,000 tracks and was spread across two volumes. This resulted in those two volumes being heavily hit, but also resulted in the index being cached by the VMAX array, resulting in very fast, memory speed response times.

**VMAX configuration**

The DB2 workload was run against a VMAX 40K with the following configuration:

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enginuity</td>
<td>5876.82.57</td>
</tr>
<tr>
<td>Enterprise Flash drives</td>
<td>200 GB RAID 5 (3+1)</td>
</tr>
<tr>
<td>Fibre Channel drives</td>
<td>600 GB 15K RAID 5 (3+1)</td>
</tr>
<tr>
<td>SATA</td>
<td>16 2TB 7.2K RAID 6 (6+2)</td>
</tr>
<tr>
<td>Cache</td>
<td>86GB (usable)</td>
</tr>
<tr>
<td>Connections</td>
<td>2x 4Gb FICON Express</td>
</tr>
<tr>
<td>Engines</td>
<td>1</td>
</tr>
</tbody>
</table>

Although EMC was not explicitly performing tests to move data to SATA drives, these drives are still an important component of FAST VP configurations. SATA drives can
augment primary drives in a HSM environment. Inactive data on the primary drives migrates to SATA over time and is still available without HRECALL operations. Note that FAST VP does this data movement without using expensive host CPU or I/O. Exactly how much SATA space in the FAST VP configuration is appropriate is site-dependent. Systems with larger capacities of SATA in use are more economical.

**DB2 configuration**

The entire DB2 subsystem was deployed on 64 MOD-9s. The partitioned table space containing the data that was accessed was deployed on 26 4 GB partitions that were spread across 26 volumes in the SMS storage pool. Two additional volumes contained the index that was used for the random access.

**Workload**

The workload was generated using 32 batch jobs running simultaneously. Each batch job ran 200,000 transactions, each of which generated 40 random reads to a partitioned table space spread across 26 MOD-9s. This resulted in almost 100 percent cache miss for the individual row fetches. The index used to locate the individual rows in the partitions resided on two volumes and was 32,000 tracks in total. These two volumes were the most heavily hit during the runs of the workload and also had a high cache hit rate.

**FAST VP policies**

For this particular test, EMC was more interested in having FAST VP move data to EFD than have it archive data to the SATA tier. This is because FAST VP can proactively move data between tiers by driving a workload on those tiers. For SATA, there is a need to wait for the aging algorithms to determine the appropriate time to move the data to the SATA tier. This process just requires inactivity, and thus it was decided it would be best to just use two tiers for the testing.

Each of the policy settings below describe how much of the DB2 subsystem was permitted to reside on a given tier. For the tests, FAST VP policies were established to have the following allowances:

- **TEST2**: EFD 5%, FC 100%
- **TEST3**: EFD 10% FC 100%
- **TEST4**: EFD 15% FC 100%

A quick recap on what these percentages mean: Each percentage determines how much of the 500 GB DB2 subsystem was able to reside on the designated tier. For instance, in the case of TEST2, up to five percent (5%) of the subsystem (approximately 25 GB) can reside on EFD, however the entire subsystem can reside on the Fibre Channel drive tier.

**Testing phases**

The testing methodology consisted of a number of steps that involved testing on thick devices first, and then three tests with FAST VP using varying policy settings. The thick
testing phase provided a baseline for comparison to the FAST VP. In between the FAST VP runs and policy changes, the same workload was also run to give FAST VP performance statistical data to make decisions regarding data movements. (These runs were not measured.)

The following is a list of the steps that were performed for the four measured phases of the testing:

1. The workload was first run on a baseline of thickly provisioned devices. The purpose was to provide a baseline for comparison for the following tests. In the following charts, the data for this phase is labeled TEST1.

2. The complete DB2 subsystem was copied from the source thick volumes to 64 thin MOD-9s. The original source volumes were varied off and the thin volumes were varied on and the DB2 subsystem was started on the thin devices.

3. The next step was to assign the TEST2 FAST VP policy to the DB2 storage group. The amount of time to run the workload was set before data was to be moved to the minimum two hours (the default is 168 hours) and made the movement time window unrestricted. To get the data to move from the FC tier to the EFD tier was simply a matter of running the workload again. After the workload was finished, FAST VP completed its data movement in a short period of time.

4. The same workload as in step 1 was run again, and the performance data was collected. The charts designate this data as TEST2.

5. The next step was to assign the TEST3 FAST VP policy to the DB2 storage group. This was an attempt to measure the impact of increasing the capacity on the EFD by five percent. To get the next five percent of the data to move from the FC tier to the EFD tier was simply a matter of running the workload again. After the workload was finished, FAST VP completed its data movement in a short period of time.

6. The same workload as in step 1 was run again, and the performance data was collected. The charts designate this data as TEST3.

7. The next step was to assign the TEST4 FAST VP policy to the DB2 storage group. This was an attempt to measure the impact of adding another five percent EFD capacity, totaling 15 percent. To get the next five percent of the data to move from the FC tier to the EFD tier was simply a matter of running the workload again. After the workload was finished, FAST VP completed its data movement in a short period of time.

8. The same workload as in step 1 was run again, and the performance data was collected. The charts designate this data as TEST4.

**Testing results**

The four workloads were measured using RMF data and STP (Symmetrix Trends and Performance) data that was retrieved from the VMAX service processor. The STP data was input into SYMMERGE to analyze the data.
**Run times**

Figure 4 shows the various run times for the aggregate work of the 32 batch jobs. Multiple runs were executed to ensure the validity of the measurements.

![Run time (mins)](image)

**Figure 4. Batch jobs runtimes**

Clearly, the more data that FAST VP promoted to the Enterprise Flash drives, the faster the batch workload completed. Since the workload was 100 percent random read with a very low cache hit rate, this is to be expected. This test emulated the 10 percent read miss workload that a 5 TB database might experience. So the other 90 percent of the database activity was at memory speed, that is, cache hits.

**Response times**

The average response times for each of the four tests are depicted in Figure 5. The individual components for response times are shown. As can been seen, the addition of more space on the EFD tier caused an almost linear drop in response time. This is one aspect of having a completely random workload without any skew.

The graph also shows an increase in connect time when the I/O rate is increased due to the use of the Enterprise Flash drives. This is because only two FICON channels were being used in the test, and when the I/O rate started to increase, a little queuing on the FICON port on the VMAX array became evident.
The consistently large DISCONNECT times (shown in green in Figure 5) are due to the fact that the workload that was architected was almost 100 percent read miss. As explained earlier, this was a deliberate setup to try and emulate that component of the subsystem that is not getting cache hits. FAST VP algorithms do not base their calculations on I/Os that are satisfied by cache.

What is seen in Figure 5 is consistent with the job run times seen in Figure 4.
**Average IOPS**

The average IOPS for the four workloads was measured and is shown in Figure 6.

![Average IOPS](Image)

**Figure 6. Average IOPS for each test**

The behavior seen in the graph corresponds directly to the reduced response time and reduced run time depicted in the prior two figures.

**Storage distribution across the tiers**

It is possible to interrogate the VMAX array to determine how much of a thin device is on each tier. This can be accomplished in one of three ways: Running Unisphere®, using SCF modify commands in the GPM environment, or running batch JCL pool commands. The following is a truncated output from the batch command to query allocations on a series of thin devices (400-43F). This command was run after the 15 percent EFD policy was in place.

```
EMCU500I QUERY ALLOC               -
EMCU500I (                    -
EMCU500I LOCAL(UNIT(144C))    -
EMCU500I DEV(400-43F)         -
EMCU500I ALLALLOCS            -
EMCU500I )
EMCU060I Thin Allocations on 0001957-00455                API Ver: 7.40
EMCU014I Device      Alloc                  Pool
EMCU014I 00000400    150396                  ZOS11_FC_2MV
EMCU014I 00000401    150396                  ZOS11_FC_2MV
EMCU014I 00000402    150396                  ZOS11_FC_2MV
EMCU014I 00000403    150396                  ZOS11_FC_2MV
EMCU014I 00000404     91836                  ZOS11_FC_2MV
EMCU014I 00000404     58560                  ZOS11_SD_R5V
EMCU014I 00000405    132384                  ZOS11_FC_2MV
EMCU014I 00000405     18012                  ZOS11_SD_R5V
EMCU014I 00000406    103896                  ZOS11_FC_2MV
```
The output is truncated for brevity. Note that some volumes only have tracks on the Fibre Channel tier and some have tracks on both the FC tier and also the EFD tier.

When totaled, the following track counts are seen for the pool:

- Enterprise Flash tier (ZOS11_SD_R5V): 1,236,620
- Fibre Channel tier (ZOS11_FC_2MV): 8,180,124

Note that Symmetrix volume 405, which is one of the volumes that contained the active index for the application, has no tracks in the solid-state tier. This is because its intense, continuous activity kept it in the DB2 buffer pool and also in Symmetrix cache, resulting in either no I/O or a read hit, respectively. This type of I/O pattern will not cause FAST VP to move the data on the volume to the EFD tier.

Also note that all the volumes in the pool were pre-allocated. This means that all the tracks for the volumes were assigned track groups in the pool, which accounts for many volumes being allocated the maximum number of tracks (150,396). This number exceeds the host-visible number of tracks (150,255) due to the host-invisible cylinders that are allocated out of the pool (CE cylinders, and so on).

**Summary**

FAST VP dynamically determined which active data needed to be on Enterprise Flash drives and automatically moved that data up to the Flash tier based on the policies that were established. The movement to the Flash tier was accomplished using the storage controller resources and was transparent to z/OS, apart from the significant improvement in performance that was observed. It is important to realize how impossible this would be to accomplish this kind of dynamic, automatic, tiering in response to an active, changing workload by using manual methods.

**Best practices for DB2 and FAST VP**

In this section, some best practices are presented for DB2 for z/OS in a FAST VP context. DB2 can automatically take advantage of the advanced dynamic and automatic tiering provided by FAST VP without any changes. However, there are some decisions that need to be made at setup time with respect to the performance and
capacity requirements on each tier. There is also the setup of the storage group, as well as the time windows, and some other additional parameters. All of the settings can be performed using Unisphere for VMAX.

**Unisphere for VMAX**

Unisphere for VMAX can be used to manage all the necessary components to enable FAST VP for DB2 subsystems. While details on the use of Unisphere are beyond the scope of this document, the following parameters need to be understood to make an informed decision about the FAST VP setup.

**Storage groups**

When creating a FAST VP storage group (not to be confused with an SMS storage group), you should select thin volumes that are going to be treated in the same way, with the same performance and capacity characteristics. A single DB2 subsystem and all of its volumes might be an appropriate grouping. It might also be convenient to map a FAST VP storage group to a single SMS storage group, or you could place multiple SMS storage groups into one FAST VP storage group. Whatever is the choice, remember that a FAST VP storage group can only have thin devices in it.

If you have implemented Virtual Provisioning and are later adding FAST VP, when creating the FAST VP storage group with Unisphere, you must use the option Manual Selection and select the thin volumes that are to be in the FAST VP Storage Group.

**FAST VP policies**

For each storage group that you define for DB2, you need to assign a policy for the tiers that the storage is permitted to reside on. If your tiers are EFD, FC, and SATA, as an example you can have a policy that permits up to 5 percent of the storage group to reside on EFD, up to 60 percent to reside on FC, and up to 100 percent to reside on SATA. If you don’t know what proportions are appropriate, you can use an empirical approach and start incrementally. The initial settings for this would be 100% on FC and nothing on the other two tiers. With these settings all the data remains on FC (assuming it lives on there already). At a later time, you can dynamically change the policy to add the other tiers and gradually increase the amount of capacity allowed on EFD and SATA. This can be performed using the Unisphere GUI. Evaluation of performance lets you know how successful the adjustments were, and the percentage thresholds can be modified accordingly.

A policy totaling exactly 100 percent for all tiers is the most restrictive policy and determines what exact capacity is allowed on each tier. The least restrictive policy allows up to 100 percent of the storage group to be allocated on each tier.

DB2 test systems would be good targets for placing large quantities on SATA. This is because the data can remain for long times between development cycles, and the performance requirements can be somewhat looser. In addition, test systems do not normally have a high performance requirement and most likely will not need to reside on the EFD tier. An example of this kind of policy would be 50 percent on FC and 100 percent on SATA.
Even with high I/O rate DB2 subsystems, there is always data that is rarely accessed that could reside on SATA drives without incurring a performance penalty. For this reason, you should consider putting SATA drives in your production policy. FAST VP will not demote any data to SATA that is accessed frequently. An example of a policy for this kind of subsystem would be 5 percent on EFD, 100 percent on FC, and 100 percent on SATA.

**Time windows for data collection**

Make sure that you collect data only during the times that are critical for the DB2 applications. For instance, if you REORG table spaces on a Sunday afternoon, you may want to exclude that time from the FAST VP statistics collection. Note that the performance time windows apply to the entire VMAX controller, so you need to coordinate the collection time windows with your storage administrator.

**Time windows for data movement**

Make sure you create the time windows that define when data can be moved from tier to tier. Data movements can be performance-based or policy-based. In either case, it places additional load on the VMAX array and should be performed at times when the application is less demanding. Note that the movement time windows apply to the entire VMAX controller, so you need to coordinate them with other applications requirements that are under FAST VP control.

**DB2 active logs**

Active log files are formatted by the DBA as a part of the subsystem creation process. Every single page of the log files is written to at this time, meaning that the log files become fully provisioned when they are initialized and will not cause any thin extent allocations after this. The DB2 active logs are thus spread across the pool and incur the benefit of being widely striped.

FAST VP does not use cache hits as a part of the analysis algorithms to determine what data needs to be moved. Since all writes are cache hits, and the DB2 log activity is primarily writes, it is highly unlikely that FAST VP will move parts of the active log to another tier. Think of it this way: Response times are already at memory speed due to the DASD fast write response, so can you make it any faster?

For better DB2 performance, it is recommended to VSAM stripe the DB2 active log files, especially when SRDF® is being used. This recommendation holds true even if the DB2 active logs are deployed on thin devices.

**DB2 REORGs**

Online REORGs for DB2 table spaces can undo a lot of the good work that FAST has accomplished. Consider a table space that has been optimized by FAST VP and has its hot pages on EFD, its warm pages on FC, and its cold pages on SATA. At some point, the DBA decides to do an online REORG. A complete copy of the table space is made in new unoccupied space and potentially unallocated part of the thin storage pool. If the table space can fit, it is completely allocated on the thin pool associated
with the new thin device containing the table space. This new table space on a thin device is (most likely) all on Fibre Channel drives again. In other words, de-optimized. After some operational time, FAST VP begins to promote and demote the table space track groups when it has obtained enough information about the processing characteristics of these new chunks. So, it is a reality, that a DB2 REORG could actually reduce the performance of the tables space/partition.

There is no real good answer to this. But on the bright side, it is entirely possible that the performance gain through using FAST VP could reduce the frequency of REORGs if the reason for doing the REORG is performance based. So when utilizing FAST VP, you should consider revisiting the REORG operational process for DB2.

z/OS utilities

Any utility that moves a dataset/volume (for instance ADRDSSU) changes the performance characteristics of that dataset/volume until FAST VP has gained enough performance statistics to determine which track groups of the new dataset should be moved back to the different tiers they used to reside upon. This could take some time, depending on the settings for the time windows and performance collection windows.

DB2 and SMS storage groups

There is a natural congruence between SMS and FAST VP where storage groups are concerned. Customers group applications and databases together into a single SMS storage group when they have similar operational characteristics. If this storage group were built on thin devices (a requirement for FAST VP), a FAST VP storage group could be created to match the devices in the SMS storage group. While this is not a requirement with FAST VP, it is a simple and logical way to approach the creation of FAST VP storage groups. Built in this fashion, FAST VP can manage the performance characteristics of the underlying applications in much the same way that SMS manages the other aspects of the storage management.

DB2 and HSM

It is unusual to have HSM archive processes apply to production DB2 datasets, but it is fairly common to have them apply to test, development, and QA environments. HMIGRATE operations are fairly frequent in those configurations, releasing valuable storage for other purposes. With FAST VP, you can have the primary volumes augmented with economic SATA capacity and use less aggressive HSM migration policies.

The disadvantages of HSM are:

- When a single row is accessed from a migrated table space/partition, the entire dataset needs to be HRECALLeled.
- When HSM migrates and recalls datasets, it uses costly host CPU and I/O resources.

The advantages of using FAST VP to move data to primary volumes on SATA are:
• If the dataset resides on SATA, it can be accessed directly from there without recalling the entire dataset.

• FAST VP uses the VMAX storage controller to move data between tiers.

An example of a FAST VP policy to use with DB2 test subsystems is 0 percent on EFD, 50 percent on FC, and 100 percent on SATA. Over time, if the subsystems are not used, and there is demand for the FC tier, FAST VP will move the idle data to SATA.

**Conclusion**

As data volumes grow and rotating disks deliver fewer IOPS per GB, organizations need to leverage select amounts of Enterprise Flash drives to be able to meet the demanding SLAs of their business units. The challenge is how to optimize tiering and the use of the Flash drives by ensuring that the most active data is present on them. In addition, it makes good economic sense to place the quiet data on SATA drives, which can reduce the total cost of ownership. The manual management of storage controllers with mixed drive technologies is complex and time consuming.

Fully Automated Storage Tiering for Virtual Pools can be used with DB2 for z/OS to ensure that DB2 data receives the appropriate service levels based on its requirements. It does this transparently and efficiently. It provides the benefits of automated performance management, elimination of bottlenecks, reduced cost through use of SATA, and reduced footprint and power requirements. The granularity of FAST VP makes sure that only the most demanding data is moved to Enterprise Flash drives to maximize their usage. FAST VP and DB2 are a natural fit for those who have demanding I/O environments and want automated management of their storage tiers.

**References**

*DB2 for z/OS Best Practices with Virtual Provisioning*

*z/OS and Virtual Provisioning Best Practices*

*New Features in EMC Enginuity 5876 for Mainframe Environments*

*EMC Mainframe Technology Overview*

*Implementing Fully Automated Storage Tiering for Virtual Pools (FAST VP) for EMC Symmetrix VMAX Series Arrays.*