DB2 for z/OS With EMC Storage Tiering: FAST VP

Paul Pendle
EMC Corporation

February 6, 2013
Session Number: 12945
Agenda

• The drivers for tiered storage
  • Technology changes
  • Workload skew

• FAST VP
  • Storage elements
  • Operating parameters
  • Lab testing and results

• Operational/host considerations

• Summary
Drivers Towards Storage Tiering

- Massive data growth
- Faster and faster processors
- Faster and faster channels
- Budgets are flat or decreasing
- Decline of the hard drive
- Arrival of SSD
  - Extremely high price and performance!
- Arrival of SATA
  - Extremely low price and performance!
The RPM Story

Revolutions per minute

1999 2002 2013
IOPS Density Trend for Disk Drives

IOPS Density = IOPS capability / Size GB

SSD

SATA

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
IOPS Comparisons of Drive Technology

I/O operations per second

SSD  FC  SATA

IOPS
Data Life Cycle

Time

Frequency of access

Creation

Disposition
Types of Workload Skew

- **Persistent**
  - To a large extent, historical activity is a good predictor of future activity
  - Good candidates for static tiering

- **Non-persistent**
  - Activity is mostly randomly skewed
  - Hot data today may not be hot tomorrow

- **No skew (at full-volume level)**
  - TPF, DB2 LUW DPF, Teradata
Workload Skew by Volume

Device Activity Report

IOPS

SSD

FC/SAS

SATA
The Static Tiering Conundrum

- How do you know what database objects to place on each tier?????
  - Largely, access patterns to an object change over time
  - The most frequently accessed objects are in the DB2 buffer pool or in the storage controller cache.
  - The biggest objects are not good choices
- What about DB2 logs?
- High write table spaces?
- Sequentially accessed table spaces?
- The whole table space? Partition? Part of a table space?
Sub-Volume Skewing

Typically, only parts of a volume are consistently “hot”
FAST VP: Automated Storage Tiering

- FAST VP—Fully Automated Storage Tiering Virtual Pools
- FAST VP is a policy-based system that automatically promotes and demotes data across storage tiers to achieve performance objectives and cost targets
- Gets the right data, to the right place, at the right time
Evolution of FAST

Traditional

FAST

FAST VP
Tiered storage use case

- 240 x 300GB 15K FC Disks
- 120 x 400GB 10K FC Disks
- 8 x 200GB EFD
- 24 x 1TB SATA

Acquisition $ (2012)

17% Lower Storage Costs
+ reduce maintenance & SW costs

45% More Disk IOPS
+ more aligned with workloads

32% Less Power & Cooling
+ more efficient use of space

37% Fewer Disk Drives
152 EFD+FC+SATA vs. 240 FC

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
FAST VP Elements

Symmetrix Tiers

- R53_EFD_200GB
  200 GB EFD
  RAID 5 (3+1)

- R1_FC_450GB
  450 GB 15K FC
  RAID 1

- R614_SATA_1TB
  1TB SATA
  RAID 6 (14+2)

FAST Policies

- Optimization
  100%
  100%
  100%

- Custom
  10%
  20%
  70%

Storage Groups

- DB2 PROD1
- DB2 PROD2

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
Symmetrix Tiers

- Tiers combine a drive technology and a RAID protection type
- Virtual Pool tiers
  - Contain between 1 and 4 thin pools
    - Thin pools must consist of same drive technology and RAID protection type
    - Pools cannot be shared across multiple tiers
FAST Policies

- FAST Policies manage data placement and movement across Symmetrix Tiers for one or more Storage Groups
- Each Policy can contain up to three Symmetrix Tiers
  - Policies define the upper usage limit of each tier
- Each tier usage rule defines the maximum capacity of a storage group that can be moved to that tier
  - Each tier usage rule may be between 1% and 100%
  - Combined tier usage rules must total at least 100%, but may be greater than 100%
- Symmetrix Tiers may be shared amongst multiple FAST Policies
FAST Storage Groups

- Storage Groups logically group together devices for common management
- A Storage Group can have at most one policy associated with it
- Storage Groups may contain multiple device types
  - Associated FAST Policy will only operate on the devices that match the FAST policy type
- Devices may be “pinned” to prevent FAST movement
  - Performance statistics will continue to be collected for pinned devices
  - Statistics included when generating a new performance movement policy
Time Windows

- **Performance Window**
  - Defines the times of the day, and the days of the week during which performance data is collected
  - Allows for “quiet” periods, or irregular workloads to be excluded from analysis

- **Data Movement Window**
  - Defines the times of the day, and the days of the week during which data movements will automatically be performed
Create FAST Policy
# FAST Policies with New Policy Listed

## EMC Unisphere for VMAX V1.5.0.6

### Storage Policies

<table>
<thead>
<tr>
<th>Policy Name</th>
<th>Tier 1</th>
<th>Tier 1 %</th>
<th>Tier 2</th>
<th>Tier 2 %</th>
<th>Tier 3</th>
<th>Tier 3 %</th>
<th>Tier 4</th>
<th>Tier 4 %</th>
<th># Associated Storage Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB2 FASTVP</td>
<td>xOS_SD_R3</td>
<td>10</td>
<td>xOS_FC_2M</td>
<td>20</td>
<td>xOS_AT_R6</td>
<td>70</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FBA_Initial</td>
<td>Jim_EFDR53</td>
<td>3</td>
<td>Jim_FCR1</td>
<td>100</td>
<td>Jim_SATAR614</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>FBA_LOG</td>
<td>Jim_EFDR53</td>
<td>3</td>
<td>multipool</td>
<td>75</td>
<td>Jim_SATAR614</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Gold</td>
<td>Jim_EFDR53</td>
<td>3</td>
<td>Jim_FCR1</td>
<td>20</td>
<td>Jim_SATAR614</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>paul_test</td>
<td>xOS_SD_R3</td>
<td>50</td>
<td>xOS_FC_2M</td>
<td>50</td>
<td>xOS_AT_R3</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SAP</td>
<td>Jim_EFDR53</td>
<td>2</td>
<td>Jim_FCR1</td>
<td>27</td>
<td>Jim_SATAR614</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>test</td>
<td>multipool</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>zos_bronze</td>
<td>xOS_FC_2M</td>
<td>65</td>
<td>xOS_AT_R3</td>
<td>35</td>
<td>N/A</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>zos_gold</td>
<td>xOS_SD_R3</td>
<td>25</td>
<td>xOS_FC_2M</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>zos_OPT</td>
<td>xOS_SD_R3</td>
<td>100</td>
<td>xOS_FC_2M</td>
<td>100</td>
<td>xOS_AT_R6</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>zos_silver</td>
<td>xOS_SD_R3</td>
<td>1</td>
<td>xOS_FC_2M</td>
<td>79</td>
<td>xOS_AT_R3</td>
<td>20</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Create New Time Window
FAST VP Implementation

**Enginuity**
- Performance Data Collection
- Intelligent Tiering Algorithm
- Execute Data Movement

**FAST Controller**
- Analyze Performance Data
- Allocation Compliance Algorithm
Data Movement

- Data chunks are moved using 6.8MB chunks
- Performance-based
  - Promotions due to high I/O rates
  - Demotions to free up space for promotions
- Compliance-based
Sub-volume Tiering

<table>
<thead>
<tr>
<th>Chunk1</th>
<th>Chunk2</th>
<th>Chunk3</th>
<th>Chunk#</th>
<th>Chunk4</th>
<th>Chunk#</th>
<th>Chunk5</th>
<th>Chunk#</th>
<th>Chunk6</th>
<th>Chunk7</th>
<th>Chunk#</th>
<th>Chunk8</th>
<th>Chunk9</th>
<th>Chunk10</th>
<th>Chunk11</th>
<th>Chunk#</th>
<th>Chunk12</th>
<th>Chunk#</th>
<th>Chunk13</th>
</tr>
</thead>
</table>

**TABLE SPACE**

- SSD
- FC/SAS
- SATA
Sequential Tablespace Scan

- Long Seek (costly)
- Rotational delay (costly)
- Read
- Read
- Read ...
FAST VP TESTING WITH DB2
What was tested

- DB2 V10
- Symmetrix VMAX SE (single Engine)
- 2x4Gb Channels
- 4x200GB Enterprise Flash drives
- 32x300GB 15K FC drives
- 10% SSD in FAST VP policy
DB2 Workload

- 26 4GB partitions on 26 MOD 9s
- Highly random OLTP workload driven by 32 batch jobs
- 4x200GB Enterprise Flash drives
- 32x300GB 15K FC drives
- 507GB DB2 subsystem
- FAST VP policy set to 10% EFD use
Policy display before workload

Details: Storage Group: ZOSI_0455_DB2 SG

- Name: ZOSI_0455_DB2 SG
- FAST Policy: zos_gold
- FAST Priority: 2
- Total Capacity (GB): 507.5
- Host Name: N/A
- Volumes: 64
- Pending Views: 0

Related Objects:
- Contains: Volumes - 64
- Associated With: FAST Policy - 1

FAST Compliance Report:

<table>
<thead>
<tr>
<th>Tier</th>
<th>Protection</th>
<th>Technology</th>
<th>Max SG Demand (%)</th>
<th>Limit (GB)</th>
<th>Fast SG Used (GB)</th>
<th>Growth (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>z0S_SD_R3</td>
<td>RAID-5 (3+1)</td>
<td>EFD</td>
<td>10</td>
<td>+50.75</td>
<td>0</td>
<td>+50.75</td>
</tr>
<tr>
<td>z0S_FC_2M</td>
<td>RAID-1</td>
<td>FC</td>
<td>100</td>
<td>+507.48</td>
<td>+507.95</td>
<td>-0.48</td>
</tr>
</tbody>
</table>

Complete your sessions evaluation online at SHARE.org/SanFranciscoEval
IOPS Measured: Unisphere for VMAX Performance View

- Workload starts
- 10% of data on SSD
- FAST moves data
Policy display after workload
HOST CONSIDERATIONS
Automation at all Layers

SMS & HSM

+ FAST VP
HSM and FAST VP Intersect

- FAST VP
  - Performance Management
  - Capacity Management
- HSM
  - Availability Management
SMS Storage Groups

- **Gold**
  - 100%
  - 100%
  - 0%

- **Custom**
  - 10%
  - 100%
  - 100%

- **Test**
  - 0%
  - 20%
  - 100%

- **DB2 PROD1**
- **DB2 PROD2**
- **DB2 PROD3**
- **DB2 TEST1**
- **DB2 TEST2**

**Symmetrix Tiers**

- **R53_EFD_200GB**
  - 200 GB EFD
  - RAID 5 (3+1)

- **R1_FC_450GB**
  - 450 GB 15K FC
  - RAID 1

- **R614_SATA_1TB**
  - 1 TB SATA
  - RAID 6 (14+2)
Operational Considerations

- Storage tiering interactions with z/OS
- SMS
  - Performance-based allocations
  - DIRECT MSR Values (what do they mean now?)
- Thrashing
  - DB2 REORGs
  - HRECALLs
  - Dataset moves
  - Volume restore
FAST VP with Remote Replication Integration

SITE A

SRDF REPLICAITION

SITE B
Operational Considerations (contd)

• How to manage charge back
• How to influence decisions in the performance engine
• How to determine where everything is/was
Benefits of Storage Tiering

- Autonomic/automatic operation
- Optimized performance
- Reduced cost (power and cooling)
- Reduced footprint
- Better capacity utilization
- Ease of management
DB2 for z/OS With EMC Storage Tiering: FAST VP

Paul Pendle
EMC Corporation

February 6, 2013
Session Number: 12945