

# Using the Multi-Temperature Data Feature with Fully Automated Storage Tiering

**Roger E. Sanders**  
*EMC Corporation*

Session Code: C06-1281

Wed, May 14, 2014 (09:15 AM - 10:15 AM) | Platform: DB2 for LUW



## Topics to be discussed

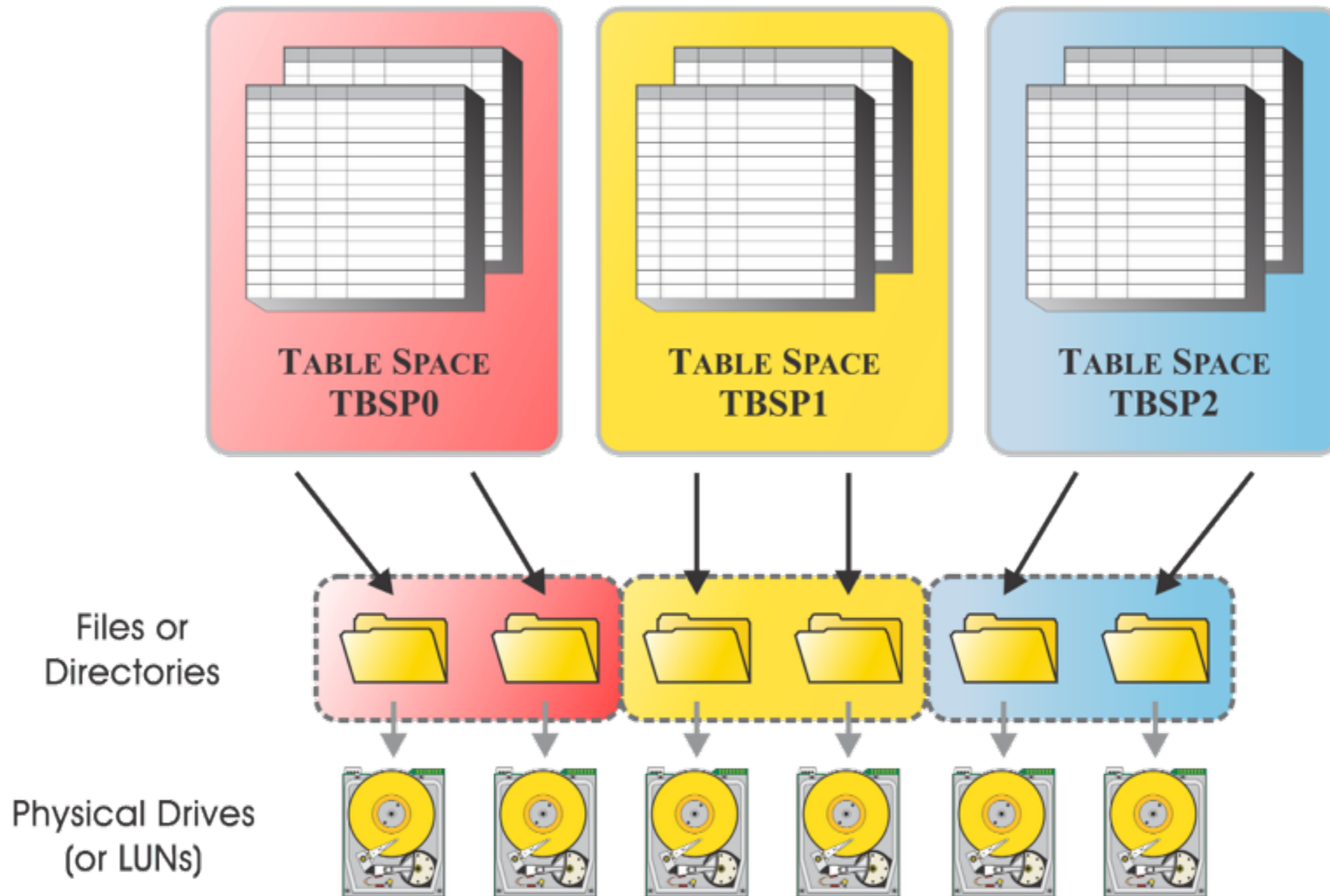
- The Multi-Temperature Data Feature
- EMC Symmetrix VMAX Virtual Provisioning
- EMC Fully Automated Storage Provisioning for Virtual Provisioning
- Using The Multi-Temperature Data Feature with FAST VP
- Conclusion

## DB2's Multi-Temperature Data Feature

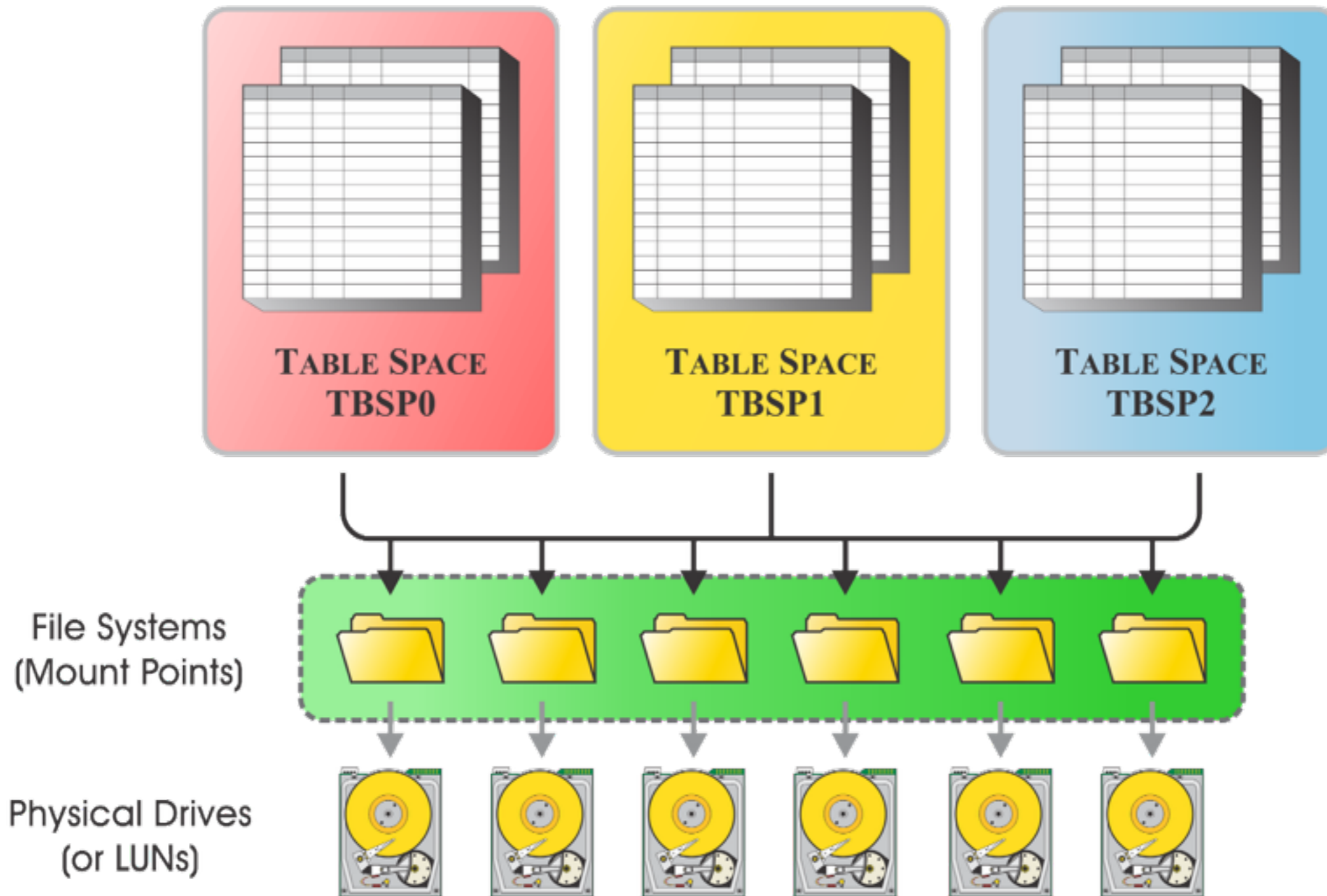
---

What is the Multi-Temperature Data Feature and why would you use it?

# Traditional Approach to Data Storage



# Automatic Storage



## The Multi-Temperature Data Feature

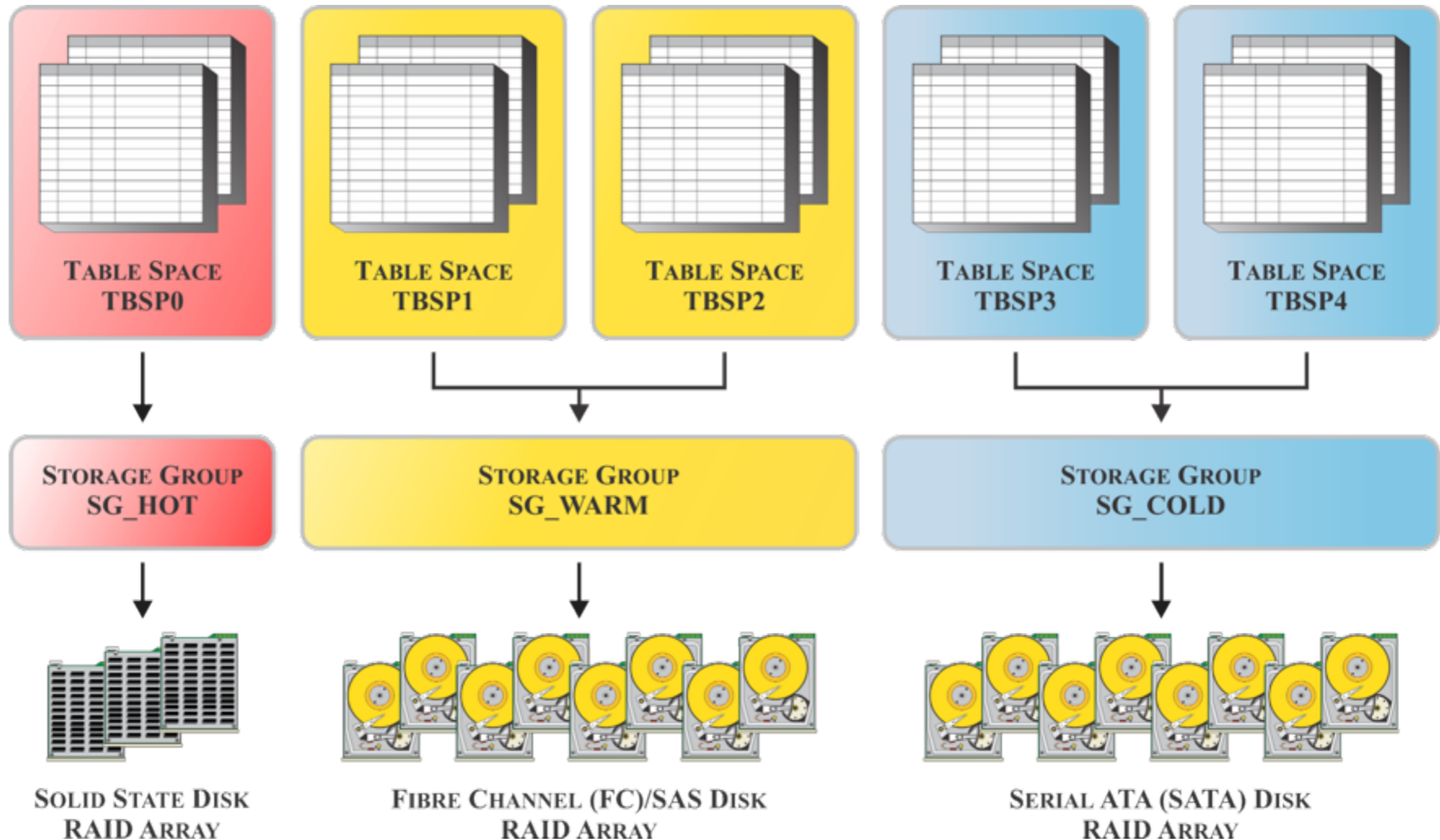
With DB2 10.1, IBM introduced a storage tiering component known as the *multi-temperature data feature*.

Using this feature, a DBA can create what is known as a *storage group* and then assign one or more table spaces to that group.

Storage groups are typically used to represent the different classes of storage that are available in a SAN or NAS storage array (for example, SSD, FC, and SATA); table spaces are assigned to storage groups based on the frequency at which their data is accessed.



# Multi-Temperature Data Feature – Illustrated



## Benefits of the Multi-Temperature Data Feature

- Different storage groups can be defined for “hot” and “cold” data
- Data placement can be controlled based on anticipated data access needs
  - “Hot” data can be assigned to storage groups with fast devices
  - “Warm” or “cold” data (such as historical data) can be assigned to slower, cheaper devices
- Performance for target workloads can be improved
  - Workload Manager can identify and prioritize work based on the temperature of the data the work touches
- Total cost of ownership (TCO) can be reduced
  - InfoSphere Optim Configuration Manager can process data migration aging policies that automatically move table partitions to cheaper storage



## The CREATE STOGROUP Statement

Storage groups are created by executing the **CREATE STOGROUP** statement; basic syntax for this statement is:

```
CREATE STOGROUP [Name] ON [Path,...]  
<OVERHEAD [Milliseconds]>  
<DEVICE READ RATE [MegabytesPerSecond]>  
<SET AS DEFAULT>
```

```
CREATE STOGROUP sg_hot  
ON '/mnt/fsystem1', '/mnt/fsystem2'
```

Creates a storage group named SG\_HOT with the paths “/mnt/fsystem1” and “/mnt/fsystem2”

## Assigning a Table Space to a Storage Group

If the database was created with DB2 Version 10.1 or later, automatic storage table spaces are automatically assigned to the *currently designated default storage group* at the time they are created.

However, an automatic storage table space can be assigned to a different storage group (at creation time by) including the **USING STOGROUP [*StorageGroupName*]** clause with the **CREATE TABLESPACE** statement used.

## Moving a Table Space From One Storage Group to Another

Automatic storage table spaces can be moved from one storage group to another by executing an **ALTER TABLESPACE . . . USING STOGROUP** statement. When this statement is executed:

- New containers are allocated on the target storage group's storage paths.
- All of the table space's original containers are marked "Drop Pending" and new allocation requests are satisfied from the new containers.
- Data is moved, in the background, from the original containers to the new containers.
- All of the original containers are physically dropped.

## Planning for a Multi-Temperature Storage Environment

- To help ensure predictable performance, all the paths that are assigned to a storage group should have the same media characteristics (latency, device read rate, and size).
- For optimum performance, set the storage group media device attributes (**OVERHEAD** and **DEVICE\_READ\_RATE**) to accurately reflect the underlying storage device's capabilities.
- Physically separate hot, warm, cold, and dormant data by storing such data based on its temperature. That is, store hot data on the fastest storage (for example, SSD), warm data on fast storage (FC), and cold or dormant data on slower storage (SATA).

## Planning for a Multi-Temperature Storage Environment (Continued)

- Use table partitioning when possible and arrange partitions in table spaces such that the granularity of the data stored in each table space matches your schedule for moving data from one temperature tier to another.
- Before storing multiple tables in the same table space, ensure that they have the same temperature characteristics. (All the data in a table space is moved when a table space changes storage groups.)
- Be careful to not overload a system by moving too many table spaces at one time. Each **ALTER TABLESPACE ... USING STOGROUP** statement initiates an implicit background rebalance operation at commit time, which can be resource intensive.

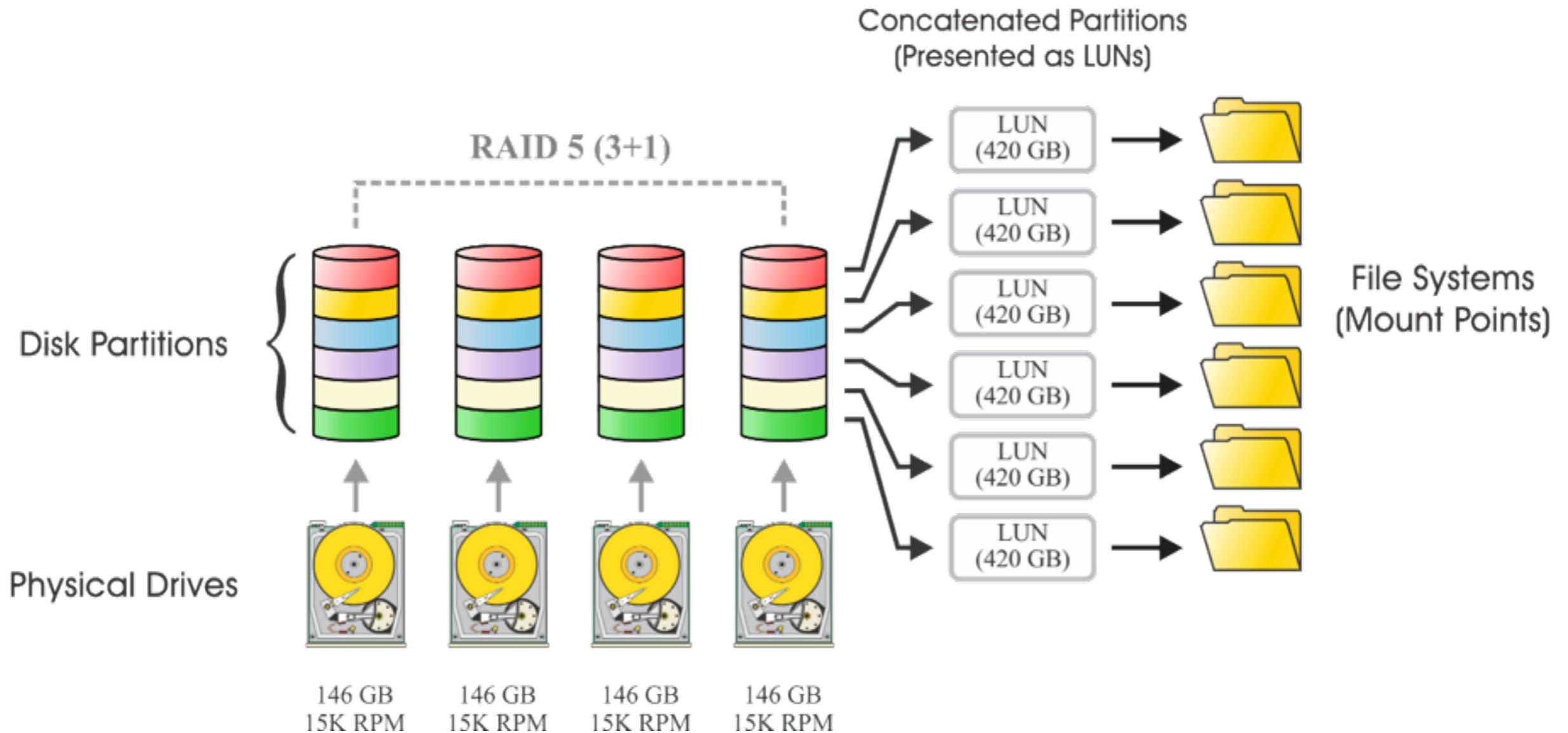


## EMC Virtual Provisioning

---

A flexible way of provisioning storage for database environments

# Traditional Physical Disk Storage Provisioning



# Challenges of Provisioning Storage for Database Environments

- Storage Administrators typically allocate space based on anticipated future growth
  - Done to prevent having to incrementally increase storage allocations
  - Often results in more storage being allocated than necessary and higher initial costs
  - Still may be necessary to provision additional storage in the future
- A second layer of over-provisioning can occur when a Database Administrator over allocates storage for a table space to ensure sufficient free space exists
  - Operating system sees the space as being completely allocated but in reality, only a fraction of the allocated space may be used
  - Over time, the table space's High Water Mark may prevent unallocation and repurposing of any unused space

## EMC Virtual Provisioning

EMC Virtual Provisioning provides the ability to configure and present a large “thin” device to a host while consuming physical storage from a shared pool only as space is needed.

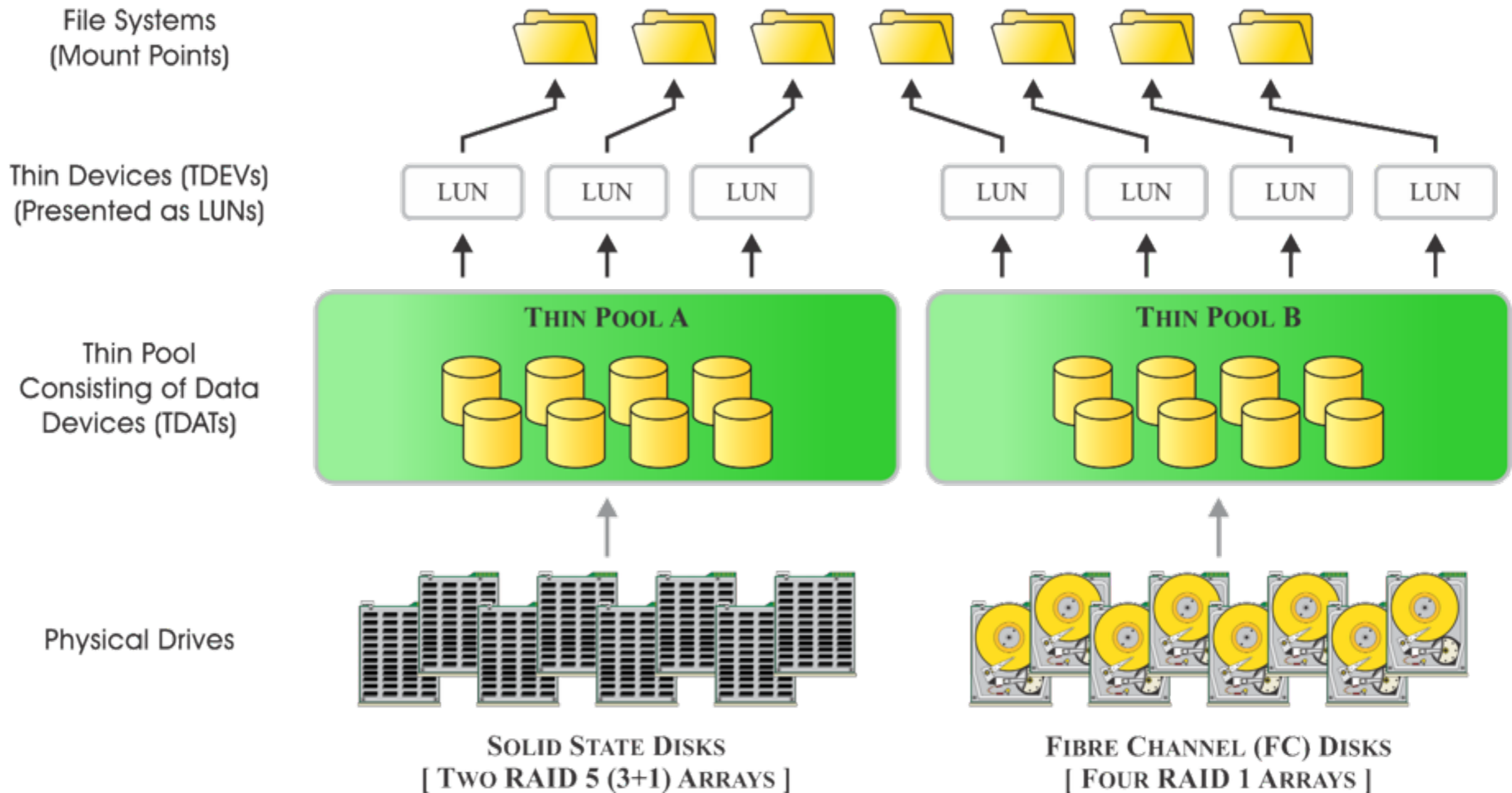
Virtual Provisioning also simplifies data layout – automated wide striping, which can deliver equivalent or potentially better performance than standard provisioning, means less planning and labor is required to configure storage.

## EMC Virtual Provisioning (Continued)

- Two types of devices are used
  - Thin Device (TDEV) → A host accessible device that has no storage directly associated with it
  - Data Device (TDAT) → An internal device that provides storage capacity to be used by thin devices
- A pool-based approach to storage
  - Thin Devices are not usable until they have been bound to a shared storage pool (known as a Thin Pool)
  - Each Thin Pool can span multiple Data Disks
  - Each pool has one disk drive type (SSD, FC, or SATA) and disk drive speed (10K RPM or 15K RPM)
  - Each pool has one RAID protection scheme (RAID 1, RAID 5, or RAID 6)
  - Multiple (up to 512) Thin Pools can exist



# EMC Virtual Provisioning – Illustrated



## Read Operations

When a read operation is performed on a Thin Device, the data being read is retrieved from the Data Device in the Thin Pool to which the Thin Device is associated.

If for some reason, a read operation is performed against an unallocated portion of a Thin Device, zeros are returned instead of data.

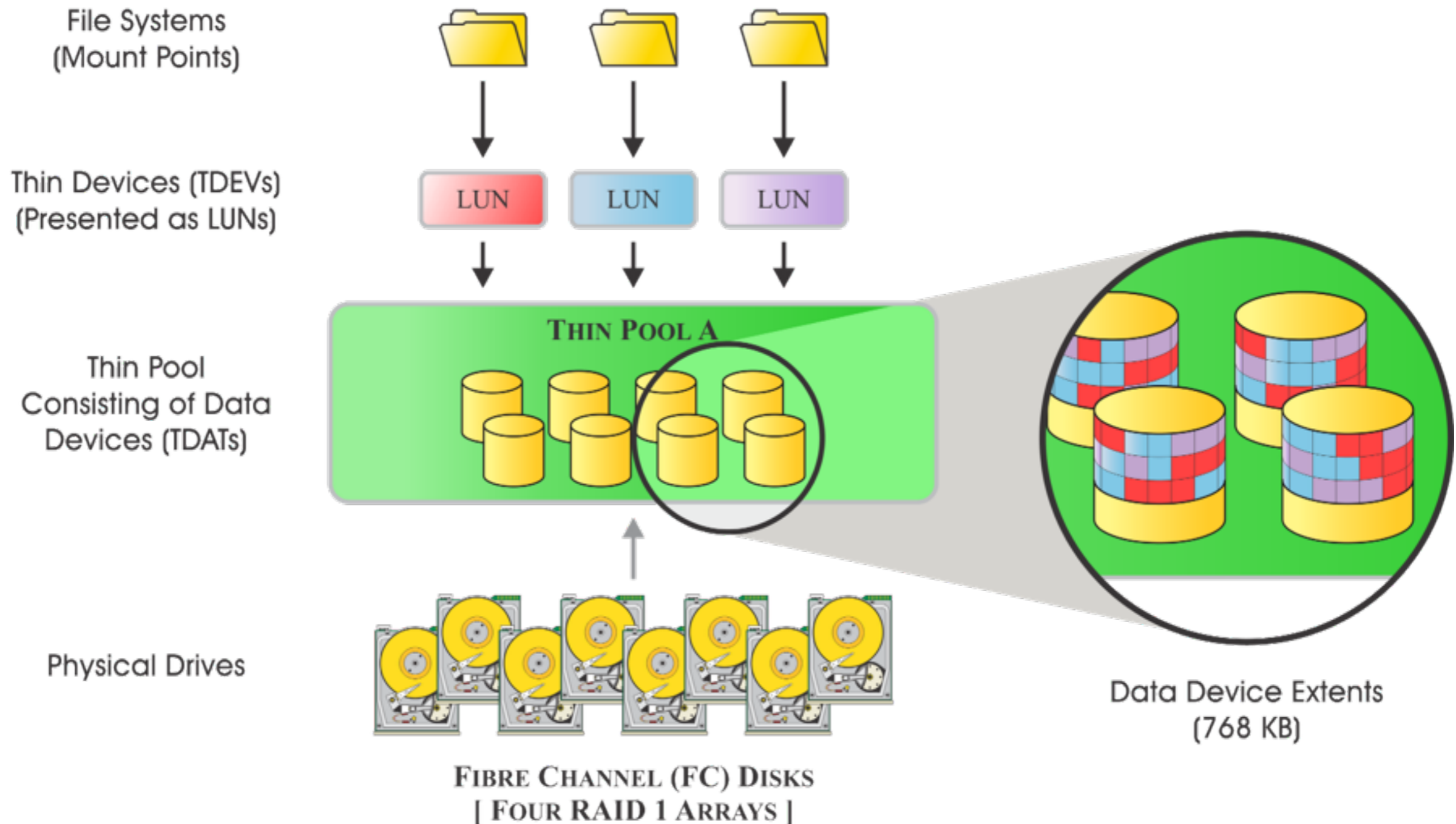
## Write Operations

When a write is performed to a Thin Device for which physical storage has not yet been allocated, a block of storage from the Thin Pool, known as a Thin Device Extent, is supplied.

The minimum amount of physical storage that can be reserved at one time for use by a Thin Device is referred to as a Data Device Extent.

- An entire Data Device Extent (12 tracks or **768** KB) is physically allocated to a Thin Device at the time a storage allocation request is made (in response to a write operation)
- The Data Device Extent is automatically allocated from any one of the Data Devices available in the associated Thin Pool
- Allocations across Data Devices are balanced to ensure that an even distribution of allocations occur from all Data Devices available in the Thin Pool

# Virtual Provisioning Write Operations – Illustrated



## Benefits of EMC Virtual Provisioning

- More storage can be presented to an application than is physically available (if additional storage is needed, it can always be added later)
- Physical storage is allocated only when it is actually written to
- The inherent waste in storage space overallocation is prevented
- The administrative overhead associated with subsequent storage allocation operations is reduced



# EMC Fully Automated Storage Tiering for Virtual Provisioning (FAST VP)

---

What is EMC Fully Automated Storage Tiering for Virtual Provisioning (FAST VP) and how is it used?

## EMC FAST VP

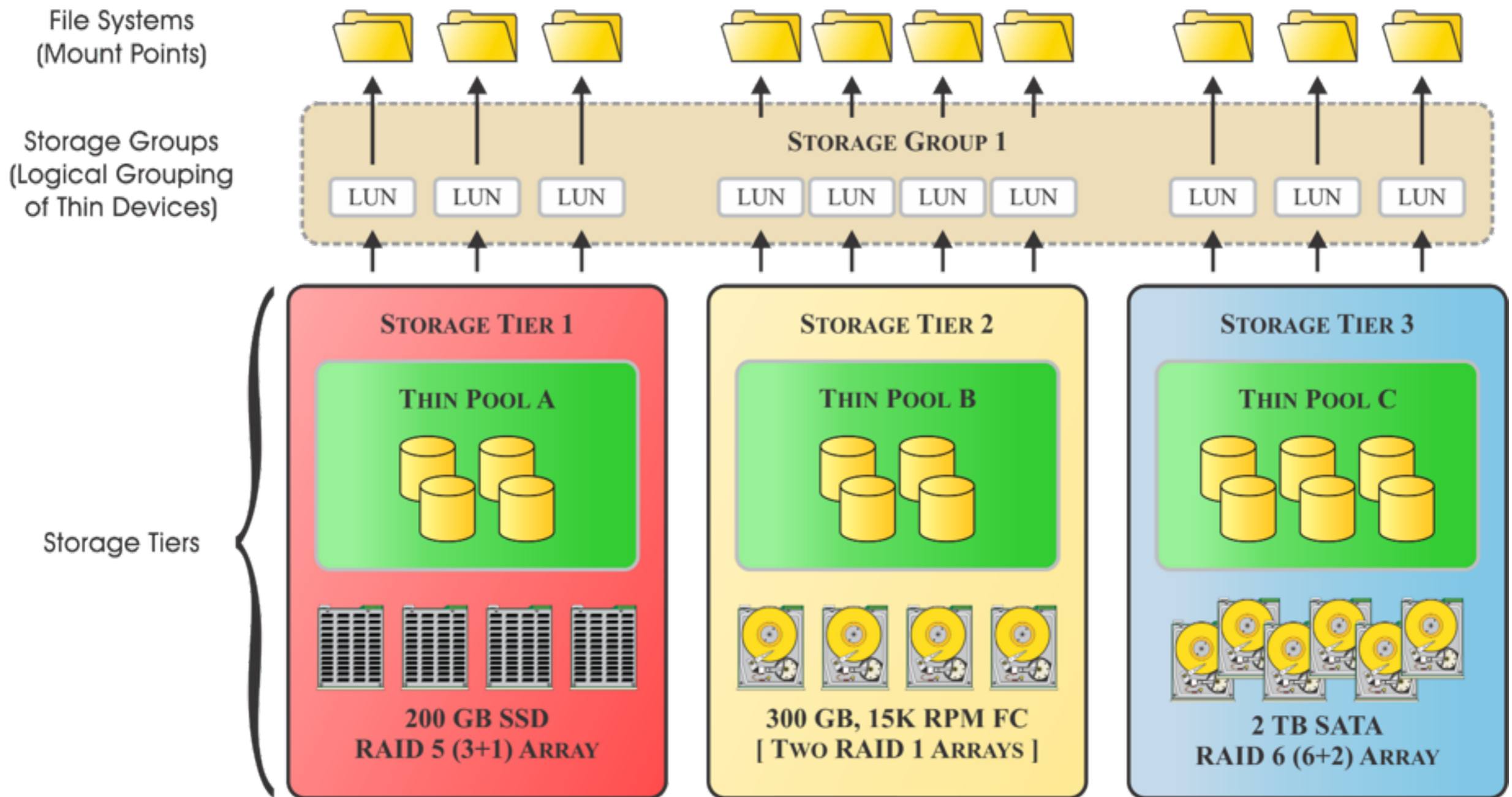
EMC Fully Automated Storage Tiering for Virtual Provisioning (FAST VP) monitors the workload and performance characteristics of a LUN and dynamically moves data between storage tiers, according to a set of tier usage rules that have been defined.

(Ten sequential Data Device Extents are moved, resulting in a transfer of 7.6 MB of data between storage tiers →  $10 \times 768 \text{ KB} = 7.6 \text{ MB.}$ )

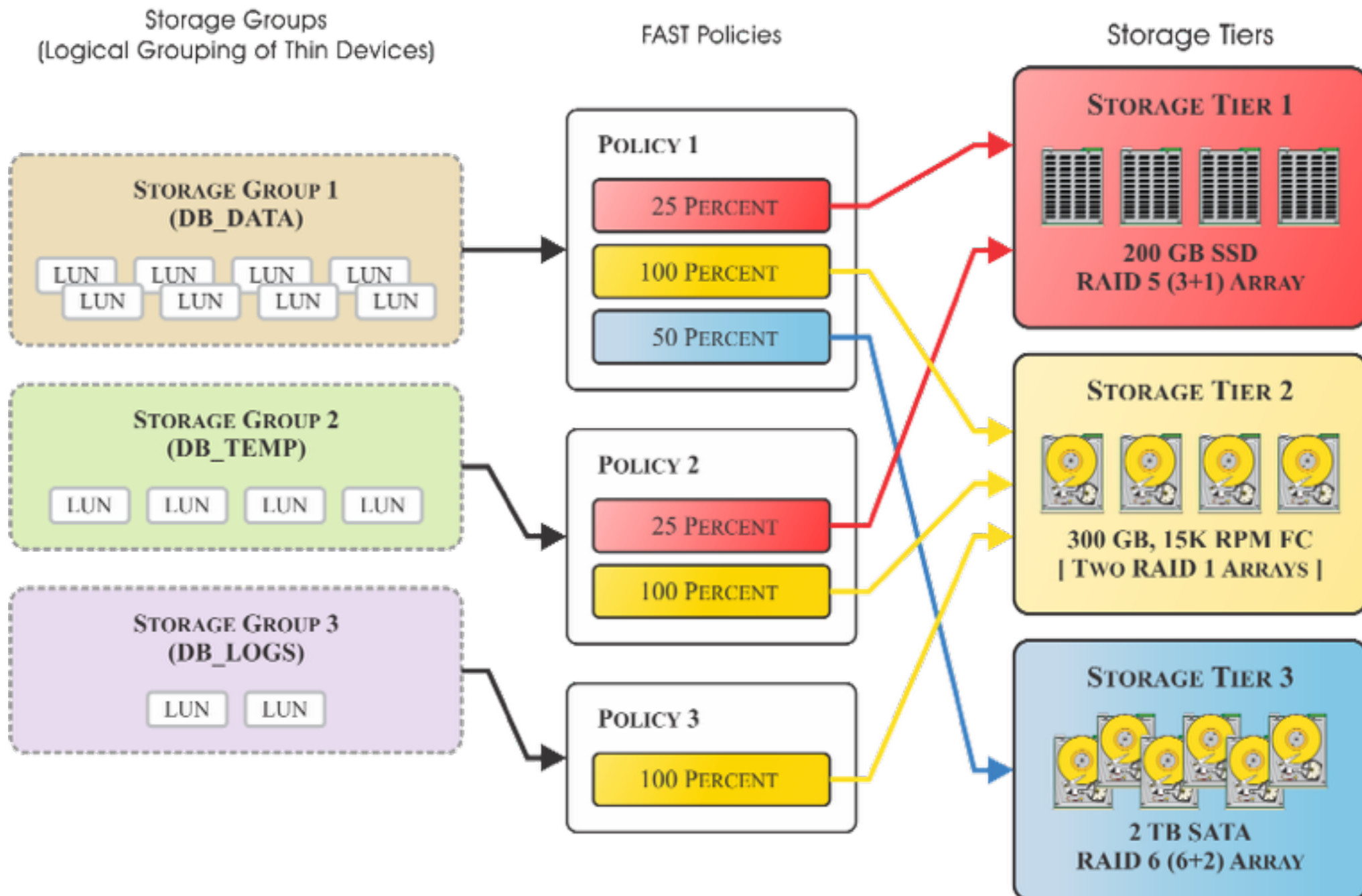
## The Main Elements of FAST VP

- Storage groups
  - A logical grouping of Symmetrix host addressable devices
- Storage tiers (or VP tiers)
  - One or more Thin Pools with common drive technologies and RAID protection types (for example, SSD/RAID 5, FC/RAID 1, and SATA/RAID 6)
- FAST policies
  - A set of tier usage rules that provide guidelines for data placement and movement across storage tiers; FAST policies tie Storage Groups to Storage Tiers and defines the configured capacities, as a percentage, that a storage group is allowed to consume on a particular tier

# Storage Groups and Storage Tiers



# Storage Groups, Storage Tiers, and FAST Policies





## Planning for FAST VP in a DB2 Database Environment

- For optimum I/O performance, use the following RAID protection schemes:
  - EFD/SSD – RAID 5 (3+1)
  - FC – RAID 1 (2-Way-Mir) or RAID 10 (RAID 1 + Striped Meta)
  - SATA – RAID 6 (6+2)
- Distribute the host I/O load by using as many Front-end Adapter (FA) ports as possible; at a minimum, four separate FA ports should be used.
- Create 4 to 8 Thin Device LUNs from a FC RAID 1 Thin Pool for active logs – combine to create a striped-meta with the same number of members as the number of FA available to create a single mount point/file system that is spread across all FA's.

## Planning for FAST VP in a DB2 Database Environment (Continued)

- Thin Device LUNs for data and temporary table spaces can be created from the same single large thin pool that LUNs used for active log are created from; however, LUNs for data and temporary table spaces should be bound to a FC/RAID 1 or RAID 10 storage tier initially.
- Do not assign the Thin Device LUNs for active logs to a FAST VP policy. (Or assign them to a policy that has an 100% FC tier only.)
- Avoid over-subscribing Thin Pools; if you believe a significant amount of data will be moved to other tiers, you can oversubscribe the default (mid-tier) FC pool where most LUNs are bound initially. But even then, you want to avoid exceeding the configured capacity for the array.

## Integrating the Multi-Temperature Data Feature with FAST VP

---

How the DB2 Multi-Temperature data feature and EMC FAST VP can be used cohesively to prevent storage “hot spots” and optimize data centric temperature management.

## Multi-temp and FAST VP

The DB2 multi-temperature data feature and the EMC Symmetrix VMAX FAST VP technology complement each other *and should not be considered mutually exclusive*.

Each offering does something entirely different – one moves data at the table space-level via DBA intervention and the other moves data at the block-level, automatically, based on data access patterns.

## Integration Strategies

If data access requirements and the temperature of data is the driving factor for determining how and where to move data, there are basically five different integration strategies to choose from:

1. Data movement via storage groups only
2. Data movement via FAST VP – using hybrid tiers – only
3. Hot/cold data movement via storage groups; warm data movement via FAST VP – using both static and dynamic storage tiers
4. Hot data movement via storage groups; warm/cold data movement via FAST VP – static SSD tier and dynamic FC and SATA tiers
5. Identified hot, warm or cold data movement via storage groups; “lukewarm” data movement via FAST VP – using both static and dynamic storage tiers

## Use Case 1: In-Memory Critical Tables

### **Situation:**

- ✓ DBA has a strong understanding of the data (and how the data is accessed)
- ✓ Available storage consists of FC and SATA drives only
- ✓ Data movement looks something like this:
  - Medium or slow storage → DB2 Buffer Pools

### **Suggested Integration strategy:**

Data movement via the multi-temperature data feature (storage groups) only



## Use Case 2: Critical Mid-Month Report Generation

### Situation:

- ✓ DBA has a strong understanding of the data (and how the data is accessed)
- ✓ Available storage consists of SSD, FC, and SATA drives
- ✓ Data movement looks something like this:
  - Medium storage → Fast storage → Medium storage

### Suggested Integration strategy:

Data movement via FAST VP (using hybrid tiers) *or*

Hot data movement via the multi-temperature data feature (storage groups); warm/cold data movement via FAST VP (static SSD tier and dynamic FC and SATA tiers)

## Use Case 3: Critical Month-End Report Generation

### Situation:

- ✓ DBA has a strong understanding of the data (and how the data is accessed)
- ✓ Available storage consists of SSD, FC, and SATA drives
- ✓ Data movement looks something like this:
  - Medium storage → Fast storage → Slow storage

### Suggested Integration strategy:

Data movement via the multi-temperature data feature (storage groups) only

## Use Case 4: Critical Quarter-End Report Generation

### Situation:

- ✓ DBA has a strong understanding of the data (and how the data is accessed)
- ✓ Available storage consists of SSD, FC, and SATA drives
- ✓ Data movement looks something like this:
  - Slow storage → Medium or Fast storage → Slow storage

### Suggested Integration strategy:

Hot/cold data movement via the multi-temperature data feature (storage groups); warm data movement via FAST VP, using both static and dynamic storage tiers

## Use Case 5: Critical Year-End Report Generation

### Situation:

- ✓ DBA has a strong understanding of the data (and how the data is accessed)
- ✓ Available storage consists of SSD, FC, and SATA drives
- ✓ Data movement looks something like this:
  - Slow storage → Medium or Fast storage → Slow storage

### Suggested Integration strategy:

Hot/cold data movement via the multi-temperature data feature (storage groups); warm data movement via FAST VP, using both static and dynamic storage tiers

## Use Case 6: High Service Level Agreements (SLAs)

### Situation:

- ✓ DBA has a moderate understanding of the data (and how the data is accessed)
- ✓ Available storage consists of SSD and FC drives (80% SSD / 20% FC)
- ✓ Data movement looks something like this:
  - Medium storage → Fast storage → Medium storage

### Suggested Integration strategy:

Hot data movement via the multi-temperature data feature (storage groups); warm/cold data movement via FAST VP – static SSD tier and dynamic FC tiers

## Use Case 7: Medium Service Level Agreements (SLAs)

### Situation:

- ✓ DBA has a moderate understanding of the data (and how the data is accessed)
- ✓ Available storage consists of SSD, FC and SATA drives (33% each)
- ✓ Data movement looks something like this:
  - Medium storage → Fast storage → Medium storage → Slow storage

### Suggested Integration strategy:

Data movement via FAST VP (using hybrid tiers) only *or*

Identified hot, warm or cold data movement via storage groups;  
“lukewarm” data movement via FAST VP – using both static and dynamic storage tiers



## Use Case 7: Low Service Level Agreements (SLAs)

### Situation:

- ✓ DBA has a moderate understanding of the data (and how the data is accessed)
- ✓ Available storage consists of FC and SATA drives (20% FC / 80% SATA)
- ✓ Data movement looks something like this:
  - Medium storage → Slow storage

### Suggested Integration strategy:

Data movement via FAST VP (using hybrid tiers) only *or*

Hot/cold data movement via storage groups; warm data movement via FAST VP – using both static and dynamic storage tiers

## Conclusion

---

A brief summary and a final opportunity to address questions and/or comments.

## Summary ...

- The multi-temperature data feature can be used to assign one or more table spaces to the different classes of storage that are available in a SAN or NAS storage array (for example, SSD, FC, and SATA).
- EMC FAST VP can be used to monitor the workload and dynamically move data between storage tiers, according to a set of pre-defined tier usage rules.
- The DB2 multi-temperature data feature and the EMC Symmetrix VMAX FAST VP technology complement each other:
  - The multi-temperature data feature doesn't take data access patterns into account; and while FAST VP can manage the most frequently accessed data, it does not have knowledge of the importance of the data from a user perspective.

## Questions?

