A Practical Guide to Storage Tiering for DB2 for z/OS

Paul Pendle
EMC Corporation

08/02/2010
1:30 PM-2:30 PM

Agenda

- Why should you care about disk?
  - … and how we got here
- Disk technologies
  - A brief introduction to Flash Drives
- Defining Storage Tiers
- Static tiering
- Integration with SMS/HSM
- Dynamic tiering
- Thin provisioning
- Sub-Volume tiering
- Futures …

Why Should You Care About Disk?

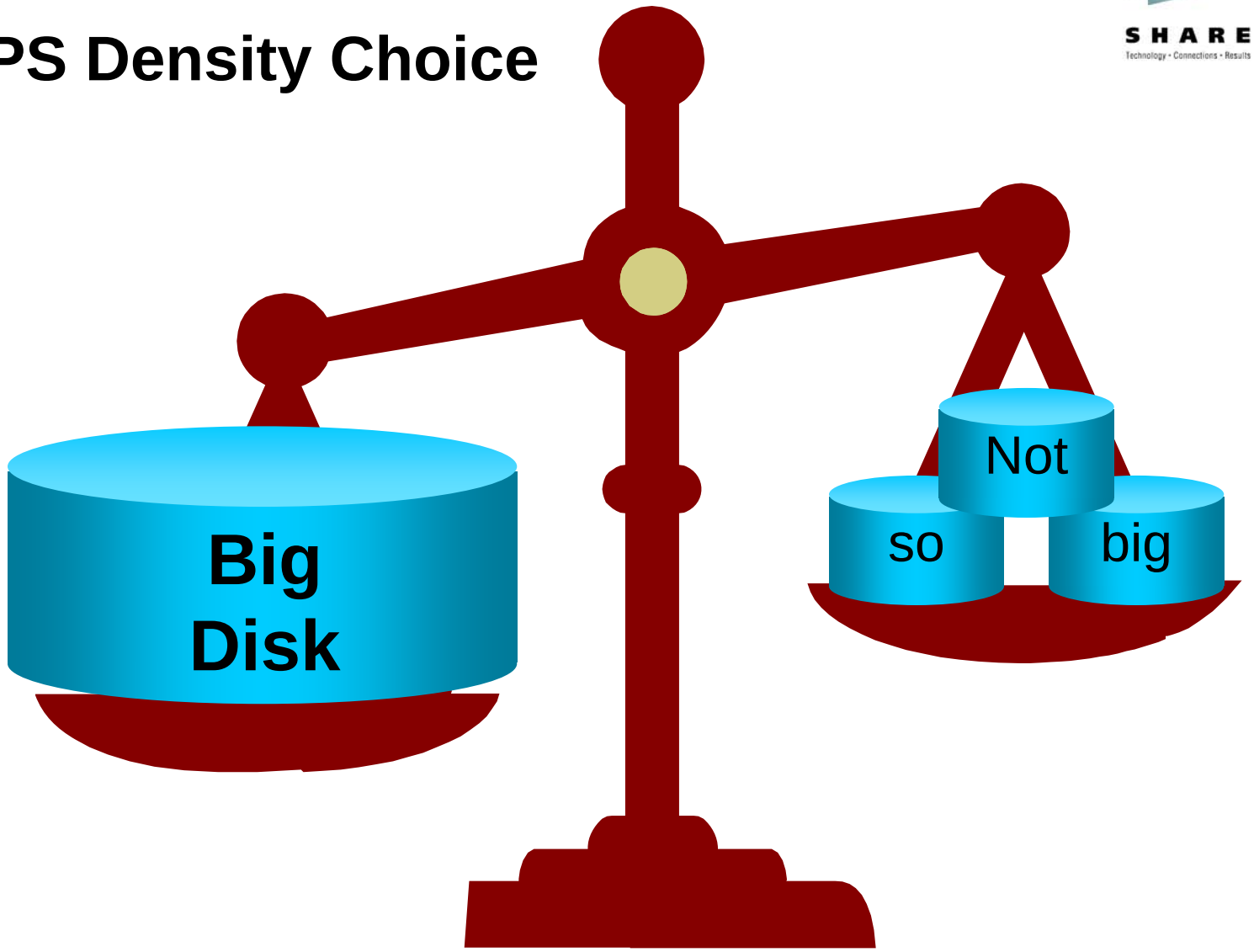
- Processors are getting more and more powerful
  - Faster, more efficient and multi-way
  - More memory (up to 1.5TB)
- Channels are getting faster
  - Bus & tag
  - ESCON
  - FICON
  - zHPF, 8Gbit FICON
- Hard Drives
  - Bigger and bigger
  - But …. not much faster

Why Should You Care (contd.)

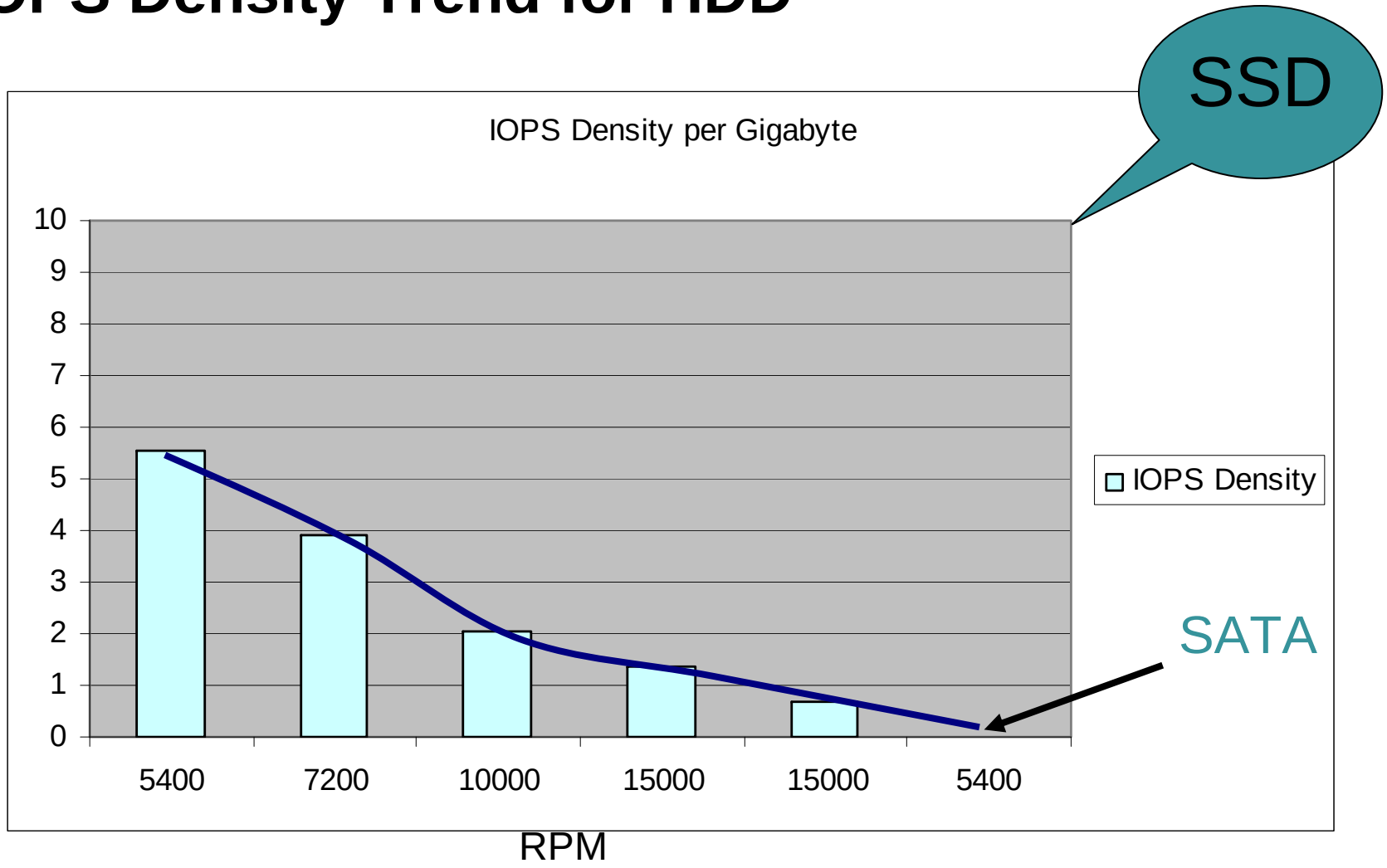
The spectrum of choices is broader than ever before:

- Different Capacities
  - From 73GB to 2TB
- Different RAID protections
  - RAID 1
  - RAID 1/0
  - RAID 5
  - RAID 6
- Different technologies
  - FC
  - SATA
  - SSD

The IOPS Density Choice



IOPS Density Trend for HDD



Solid State Disk

- They are good for every kind of workload
- They are expensive
- They are not just a large thumb drive …

Solid State Disk (contd.)

- Consumer Flash Technology
  - Typically used for write once, read many applications
  - Optimized for read performance, generally poor write performance
  - Lower cell endurance
  - Examples: memory sticks, mp3 players

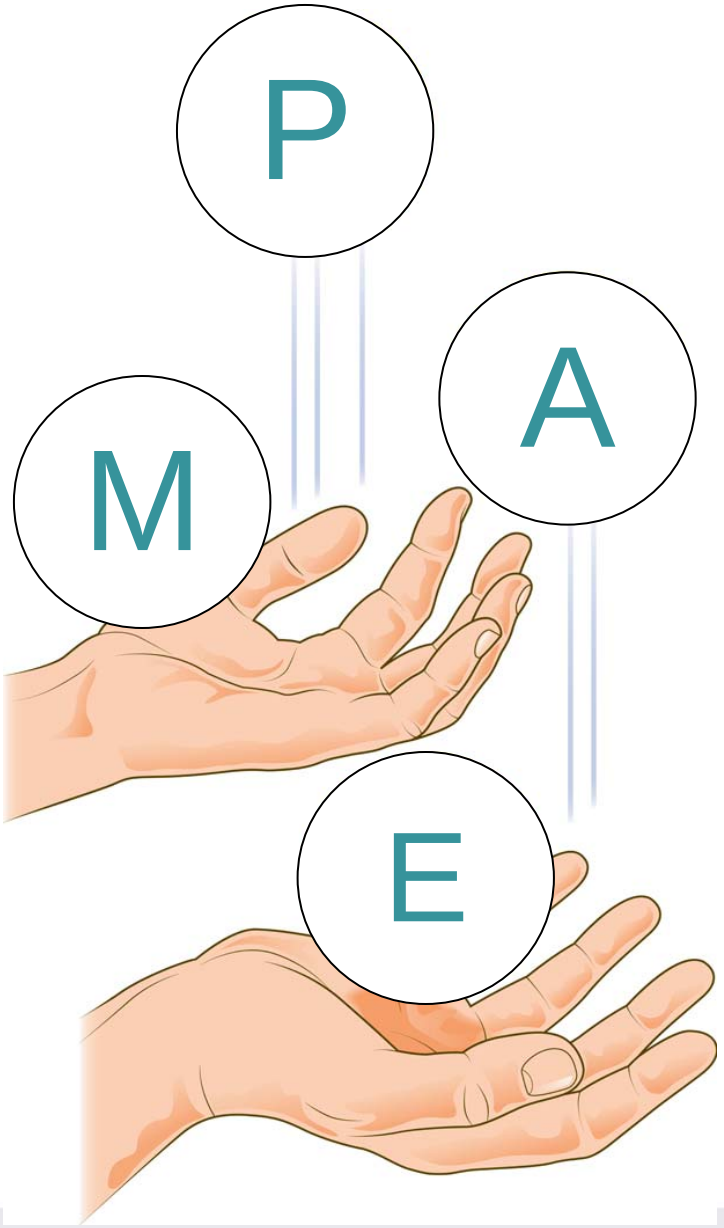
- Solid State Drives (Flash Drives)
  - Dual ported drive interface
  - Higher transfer speeds
  - Higher cell endurance
    - Wear leveling, spare cells
  - Current life expectancy greater than 8 years
  - RAID protection

Solid State Disk – The Details

- Low power usage
- Best for small page I/O (4K, 8K, 16K, 32K)
- Can do 5000 IOPS (small page)
- Best for random read miss
  - Workloads with low storage cache hit rate
- Not fantastic for writes
- Not fantastic for sequential
  - Better than HDD though
- You can't put all your data on SSD

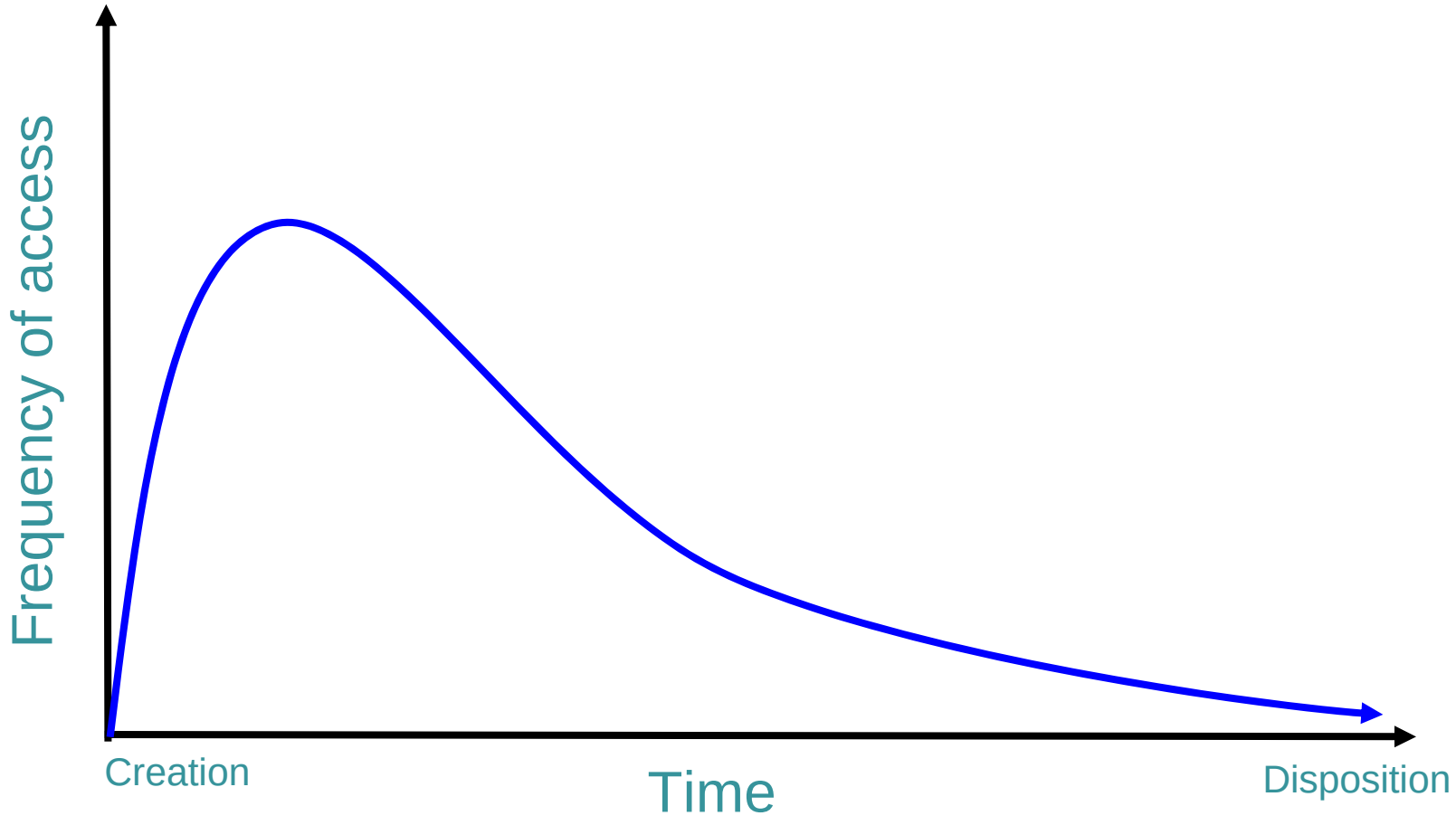
The Juggling Struggle

- Performance
- Availability
- Ecomonics
- Manageability

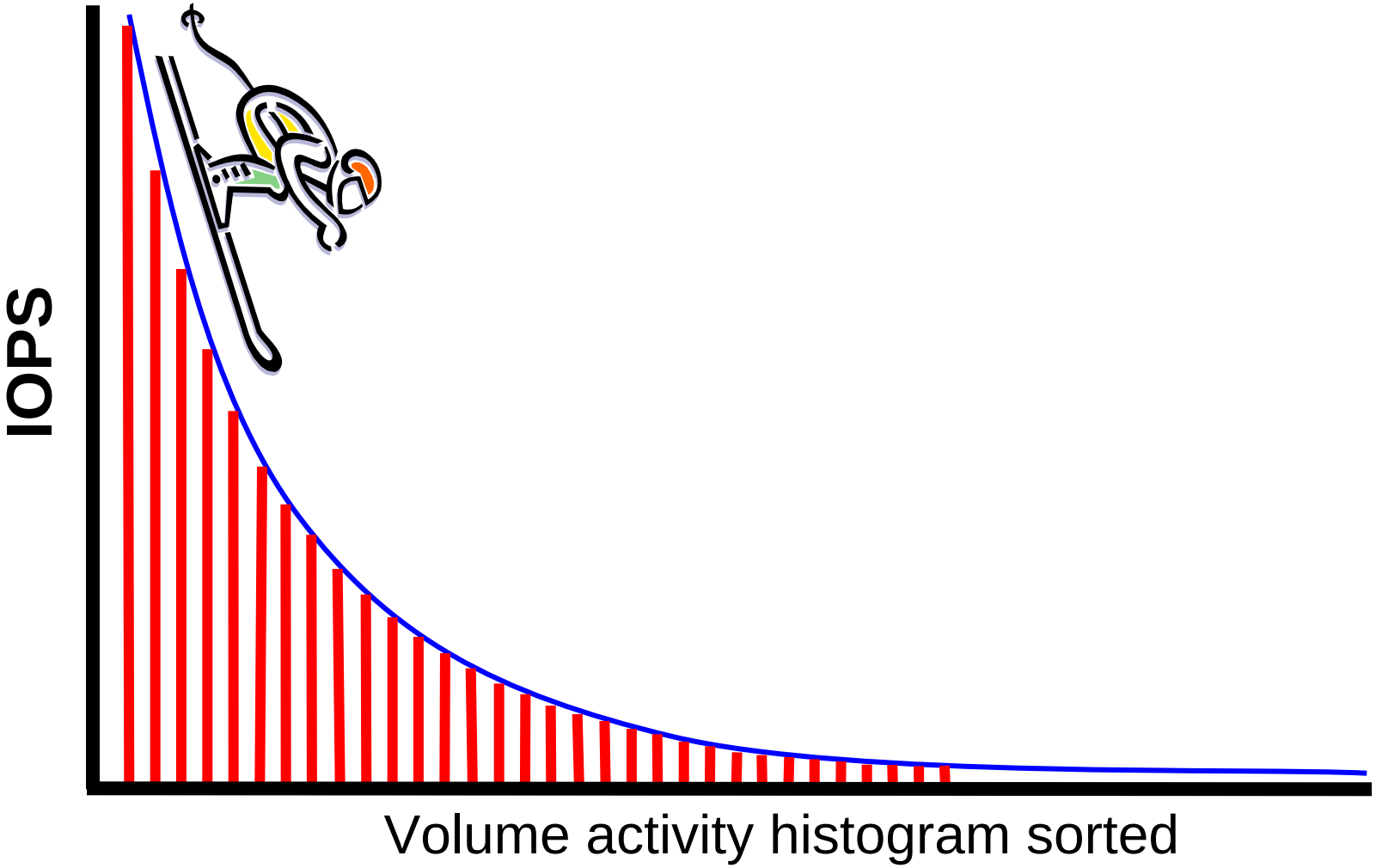


"The right data in the right place at the right time"

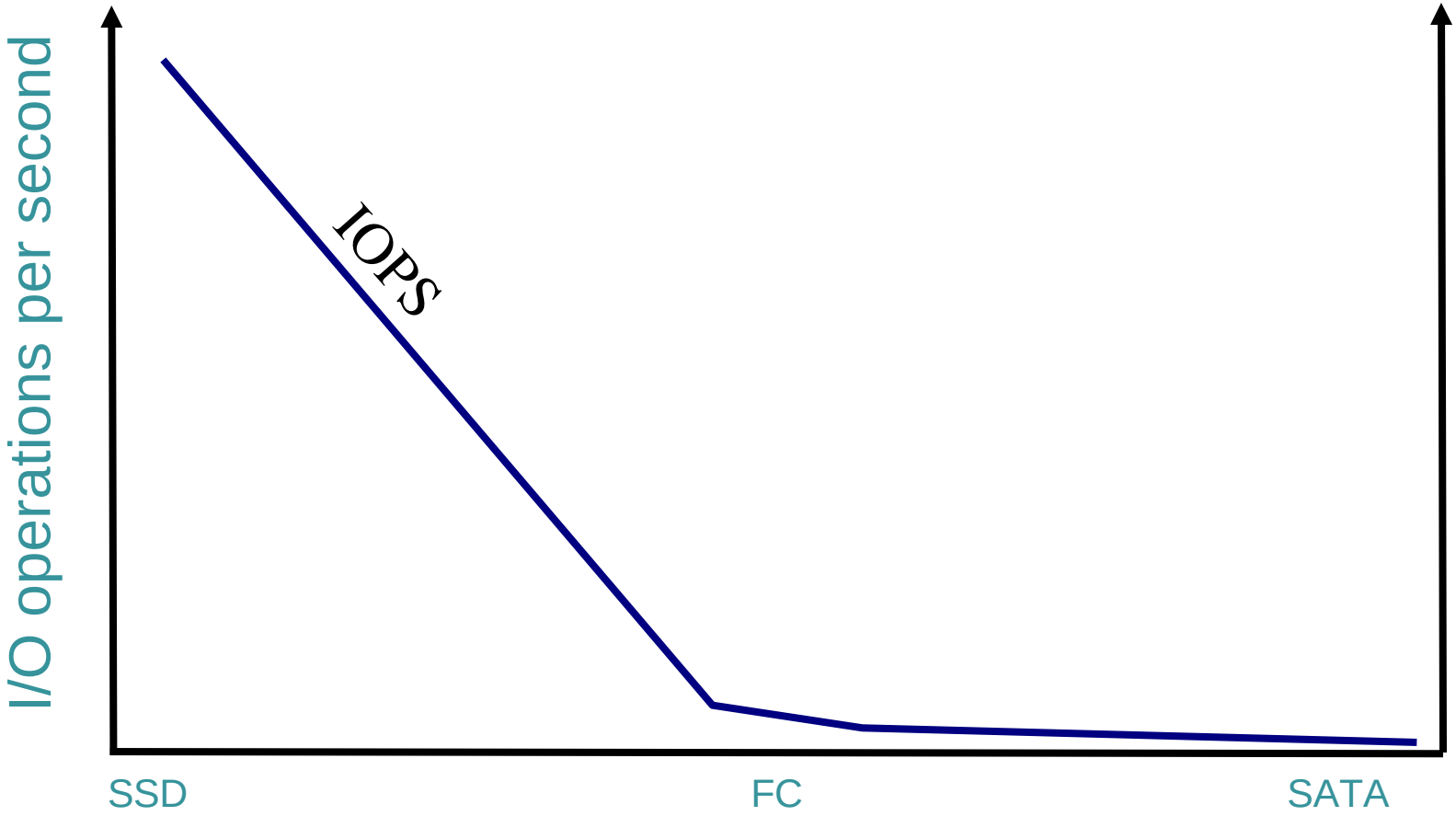
Tiered Storage - ILM



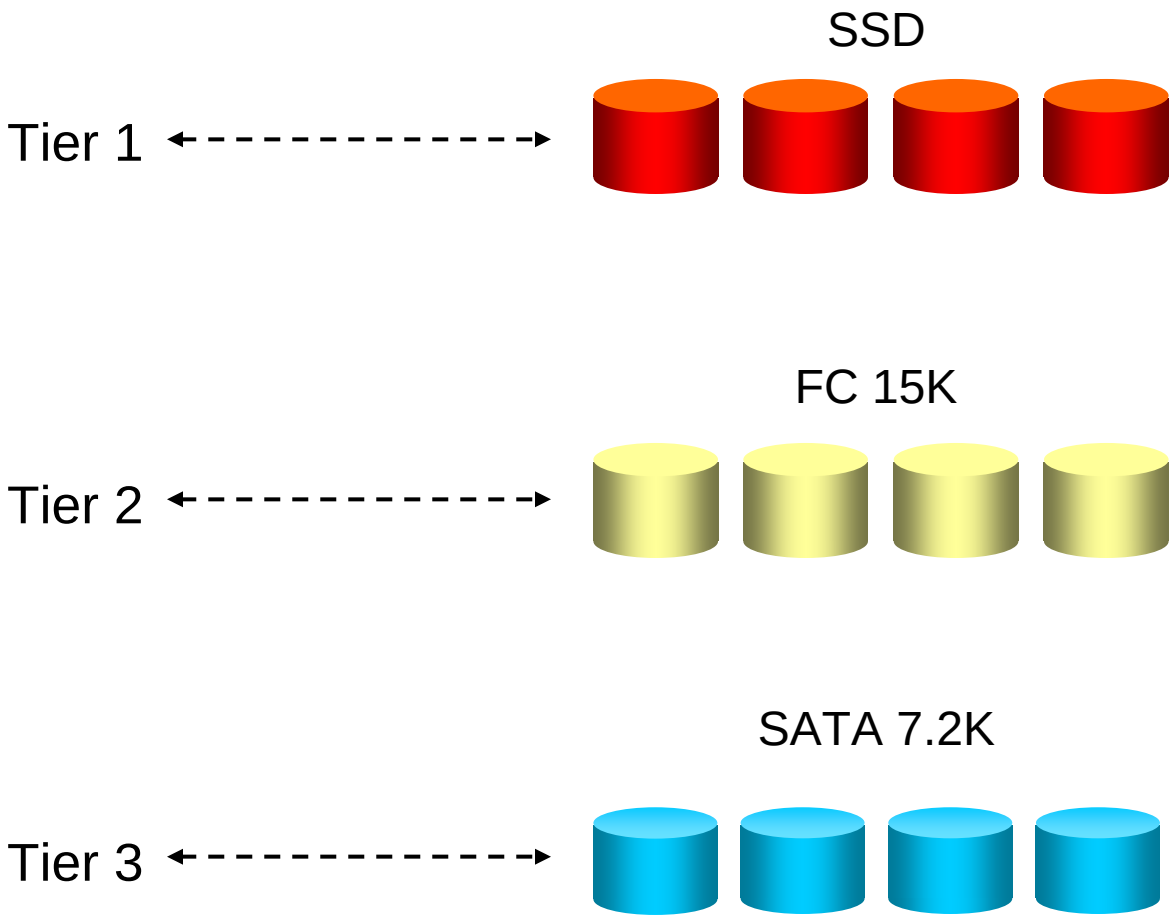
The Disk Performance Ski Slope



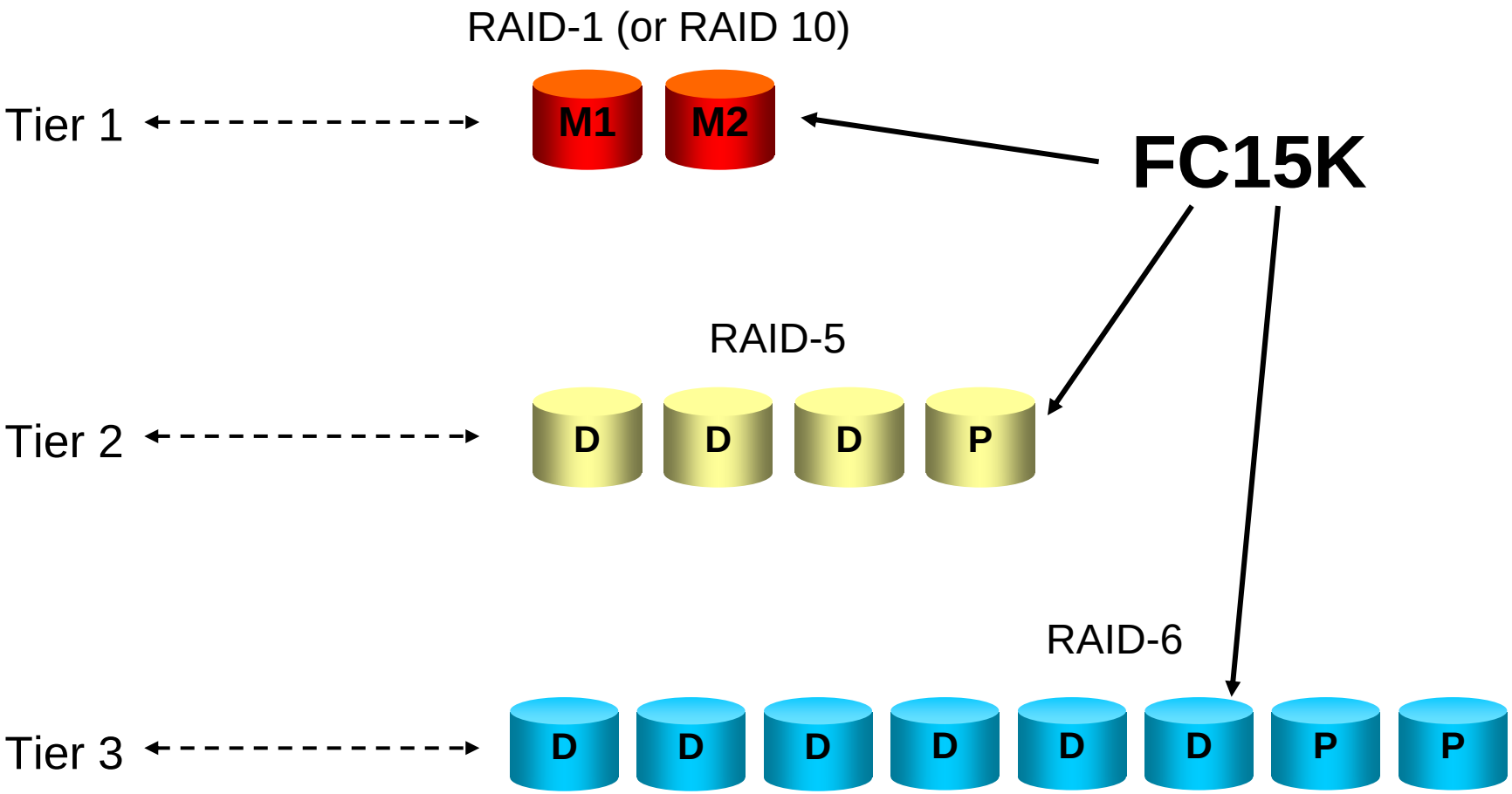
IOPS Comparison



Tiering Using Physical Disk Metrics



Tiering Using Protection Schemes

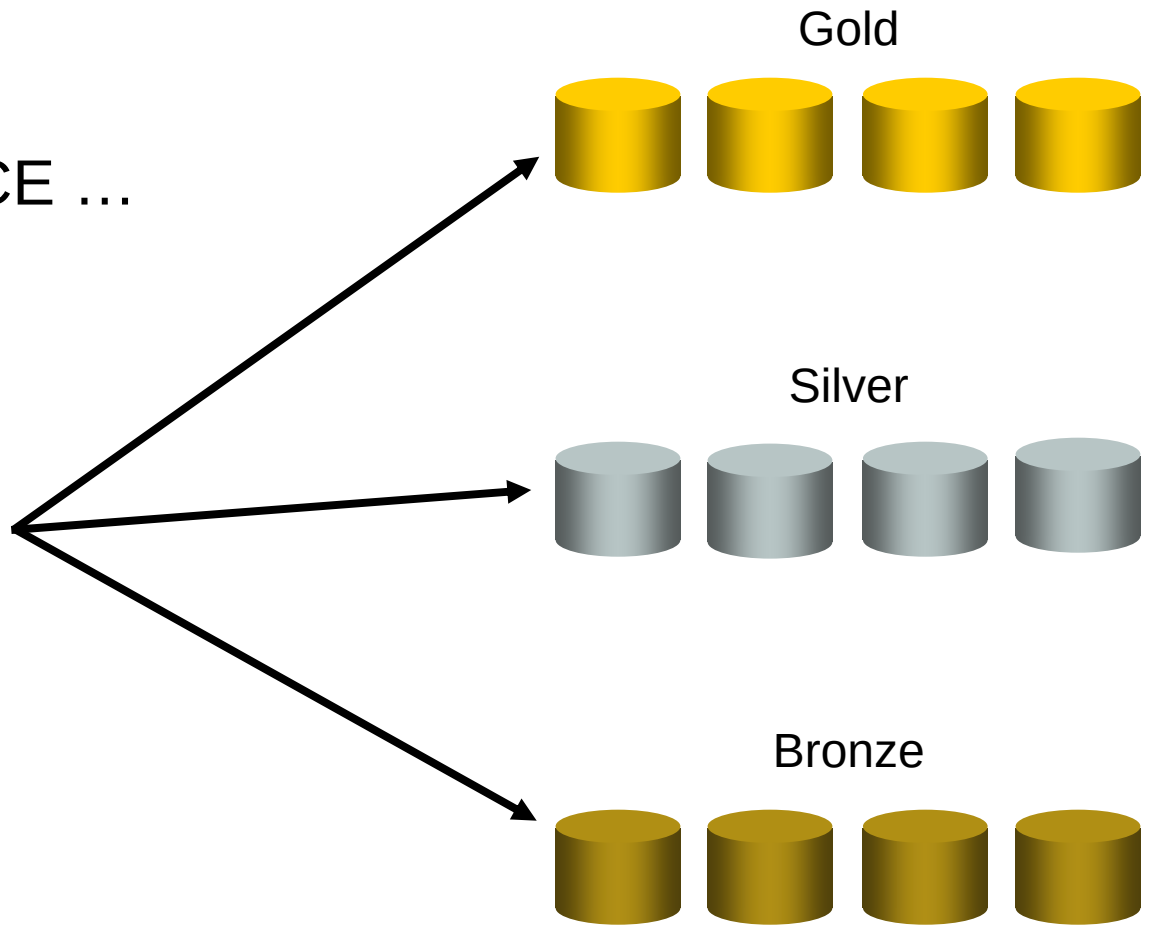


SMS Static Tiering

CREATE TABLESPACE …

HLQ

ACS
+ SC



SMS Static Tiering (Contd.)

- Mixed storage capabilities in a Storage Group?
  - DFSMS maintains a table of device performance metrics
    - DIRECT MSR = 1 (for SSD)
    - DIRECT MSR = 10 (for HDD)
  - z/OS 1.10 needs APAR OA25559 to get SSD settings
  - z/OS 1.11 includes SSD settings
- Deficiencies of the table …*
  - MOD27, MOD54, EAV, …
  - No disk geometry info
  - No understanding of RAID overhead
  - No empirical evaluation

* DFSMSdpf Storage Administration, Table 3 page 82, (z/OS 1.11)

All I/Os Are Not Equal

- Writes go to storage cache
  - Are (mostly) asynchronous from the DB2 buffer pool
  - Active logs – you could use the space for other things
- Sequential reads
  - Cause DB2 prefetch
  - Cause Array prefetch
  - Reads are asynchronous after initial few I/Os
  - Spinning disks stream faster than SSD

The Key Metric – Miss Density

- The best I/O is one serviced from bufferpool/storage cache
- Synchronous reads hurt response time
  - Especially when a “miss” in storage cache
- How to determine the best tablespaces to place on SSD?
  - The data sets with the highest miss rate may not be the best candidates!
- SMF 42 subtype 6 records – look for high DISC
  - Usually high DISC=Storage Cache miss
- How to measure miss density
  - The number of misses per GB of allocated storage

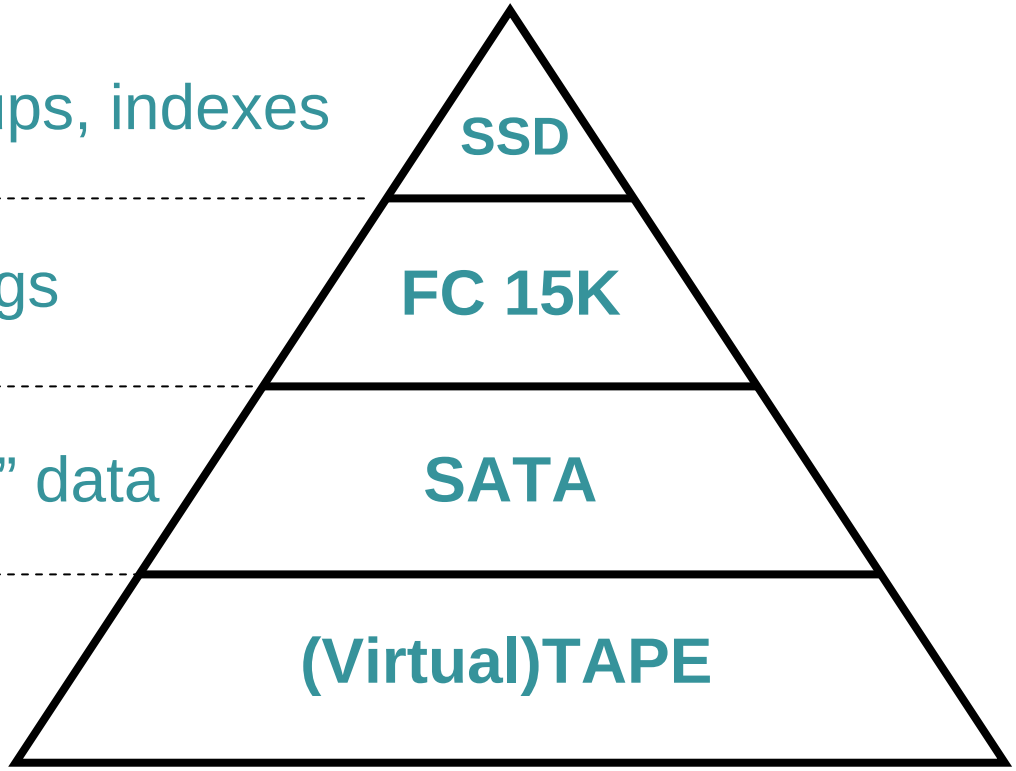
Static Management

Dimension tables, rollups, indexes

Active tables, active logs

Online archives, “quiet” data

Backups, archives



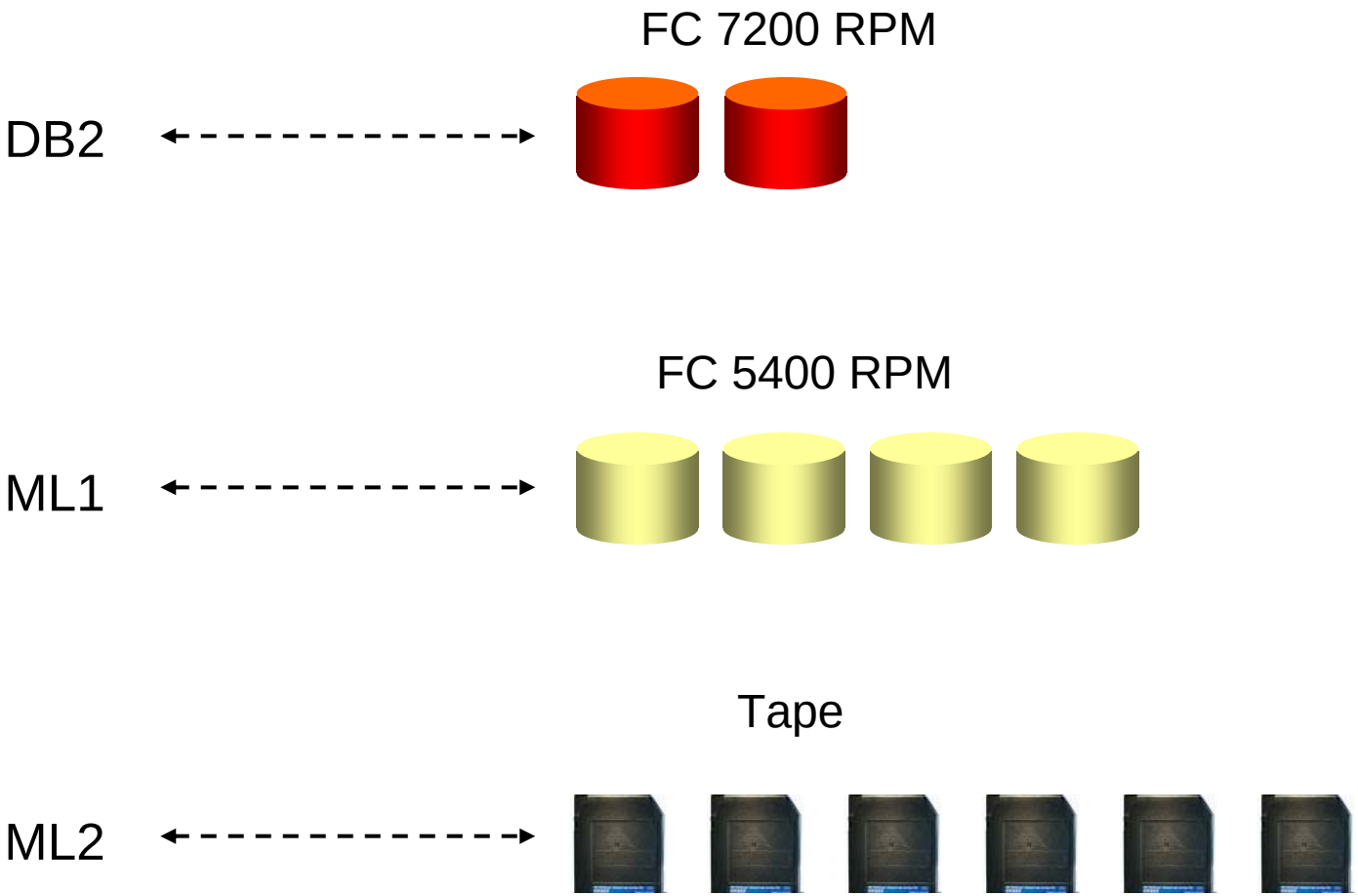
Data Movement Tools - Manual

- DB2 tools
  - Reorg utility
  - Partitioning data – partial solution
- z/OS tools
  - IBM Sotftek zDMF
  - IBM Softek TDMF
  - Innovation FDR PAS
  - EMC z/OS MIGRATOR
  - EMC V-LUN
  - HITACHI AUTOLUN

Data Movement Tools Considerations

- Ease of use / manageability
- Transparency
  - Can you do it while the table space is being used
  - Do you have to close/open it to complete the process
- Volume-based or dataset-based
- Host-based or array-based
- Reliability and risk

HSM (The First Automated Tiering Solution)



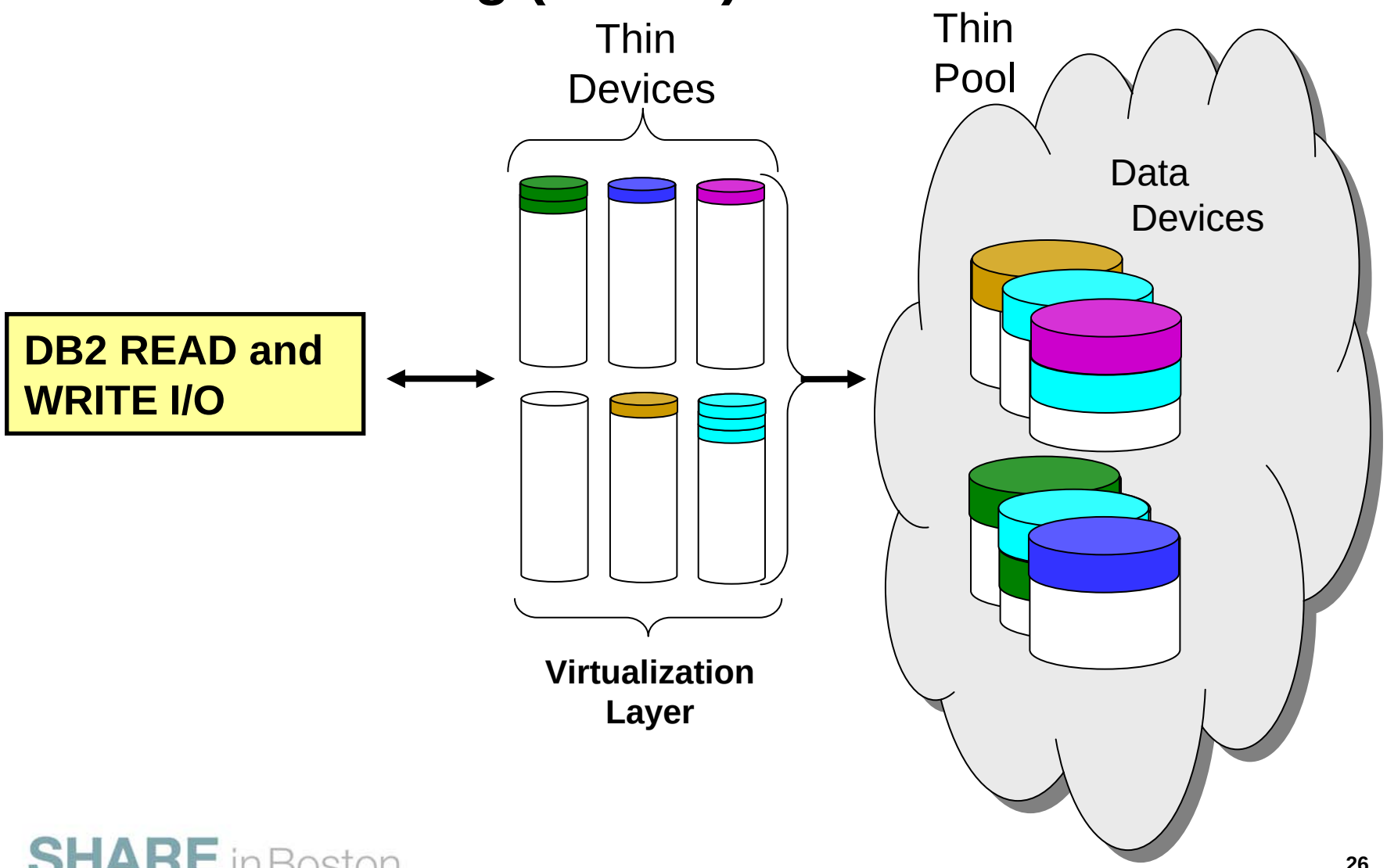
Dynamic Tiering Tools - Automated

- IBM Easy Tier
  - Released April 2010
- EMC FAST
  - Volume only at this time
- Hitachi Tiered Storage Manager

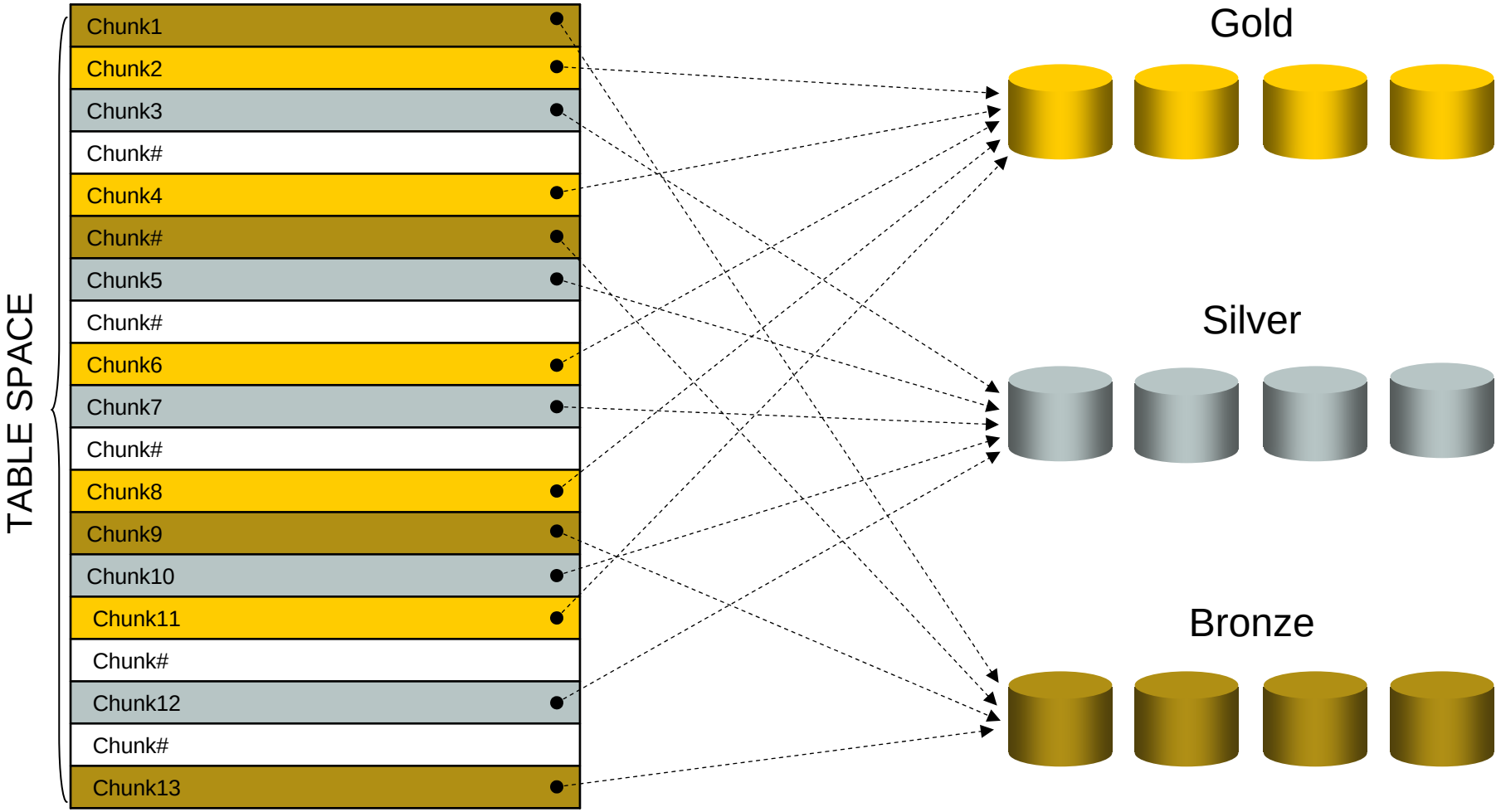
Thin Provisioning

- Virtualized storage layer
  - Host is agnostic to the underlying storage
  - Wide striping across array
  - Wide striping across storage tiers
- Provides economies against over-allocation
- Storage allocated on demand
  - DB2 V8 2 cylinders on table space creation
  - DB2 V9 16 cylinders on tablespace creation

Thin Provisioning (contd.)



Sub-volume tiering – the real requirement



Sub-volume tiering considerations

- Can only be executed inside the storage
- What size chunk?
  - Page? CI? Track? CA? Other?
- Automation in array
  - Based on policy
  - Chunk-based must be automatic!
- How to manage chargeback?
  - Gym membership approach

What is needed? Make demands!

- Policy based movement between tiers
  - Set it and forget it!
- Host hinting
  - Predict instead of react
- Exception processing
  - Never move; always move; other?
- Heuristic mechanisms
- Interface to return once-used storage
- Transparency (to application and operations)
- Billing integration

Takeaways/Takeouts

- Storage tiering is possible with DB2 now
  - Automated tiering is coming
- A virtualization layer is needed
- Storage arrays must control it
- Chunk-based movement is required

A Practical Guide to Storage Tiering for DB2 for z/OS --- THE END

Paul Pendle
EMC Corporation
Pendle_paul@emc.com

08/02/2010