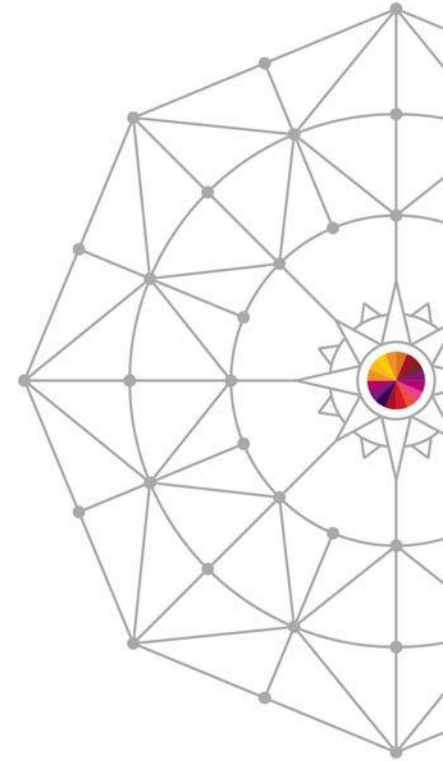


# Rethinking HSM Policies with Automated Storage Tiering

*Session ID 15997*

*Tuesday, August 5, 2014: 03:00 PM - 04:00 PM,  
DLLCC, Room 305*



**SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.**

# AGENDA

- Overview
- Reducing Storage Cost
- Hierarchical Storage Management and Automated Storage Tiering
  - History, Overview of Operations and Design Objectives
  - Values and Benefits
  - Resource Consumption
- Key Resource Measurements
- Numbers for Cost Analysis
- Summary

# Abstract

Today's storage arrays offer Automated Storage Tiering features that have the potential of reducing resources involved with HSM migrate and recall activities. This session will look at a methodology to identify resources used by traditional space management activities and how Automated Storage Tiering could be exploited as an alternative for reducing storage costs.



# Read the Book or Watch the Movie

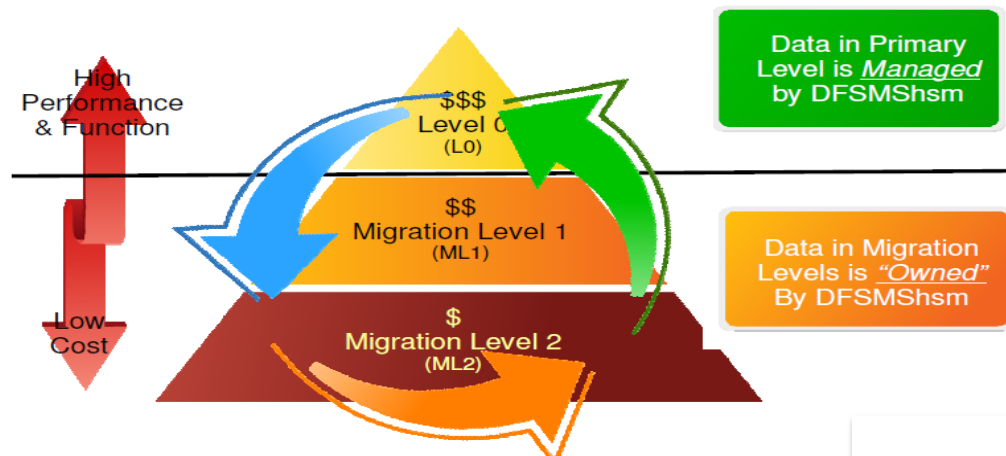


# Sucking in the Seventies

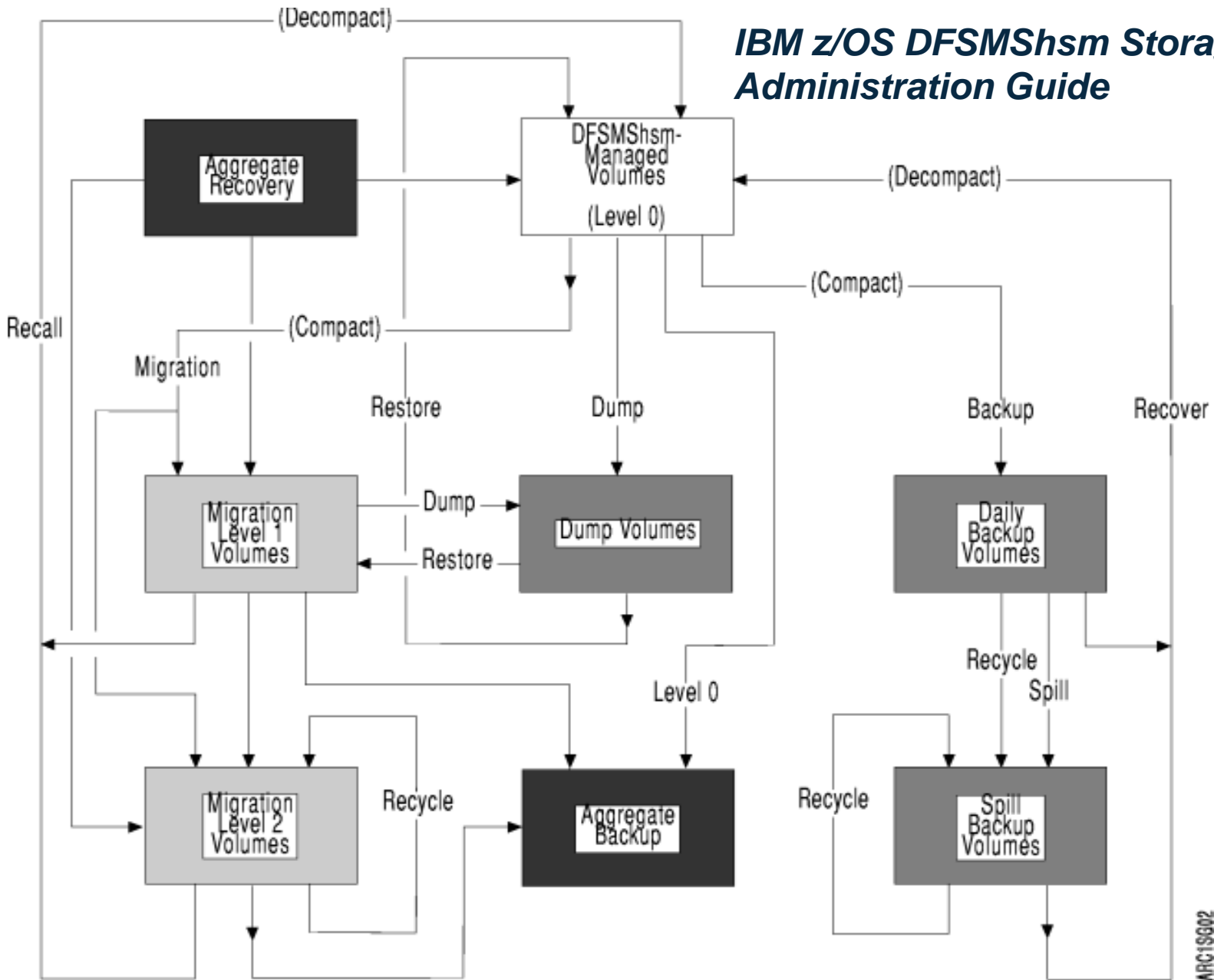
DFSMSHsm – A long history of reducing storage cost

- Leverages several technologies to reach objective
  - non-Cached and Cached DASD
  - Native tape and Virtual tape
  - Mass Storage Systems and optical media
  - Host and/or Control Unit based Data Compression

**System z processor and channel resources are used to move the data across the storage hierarchy**



# IBM z/OS DFSMSHsm Storage Administration Guide



ARC1SG02



# HSM - Benefits and added Value

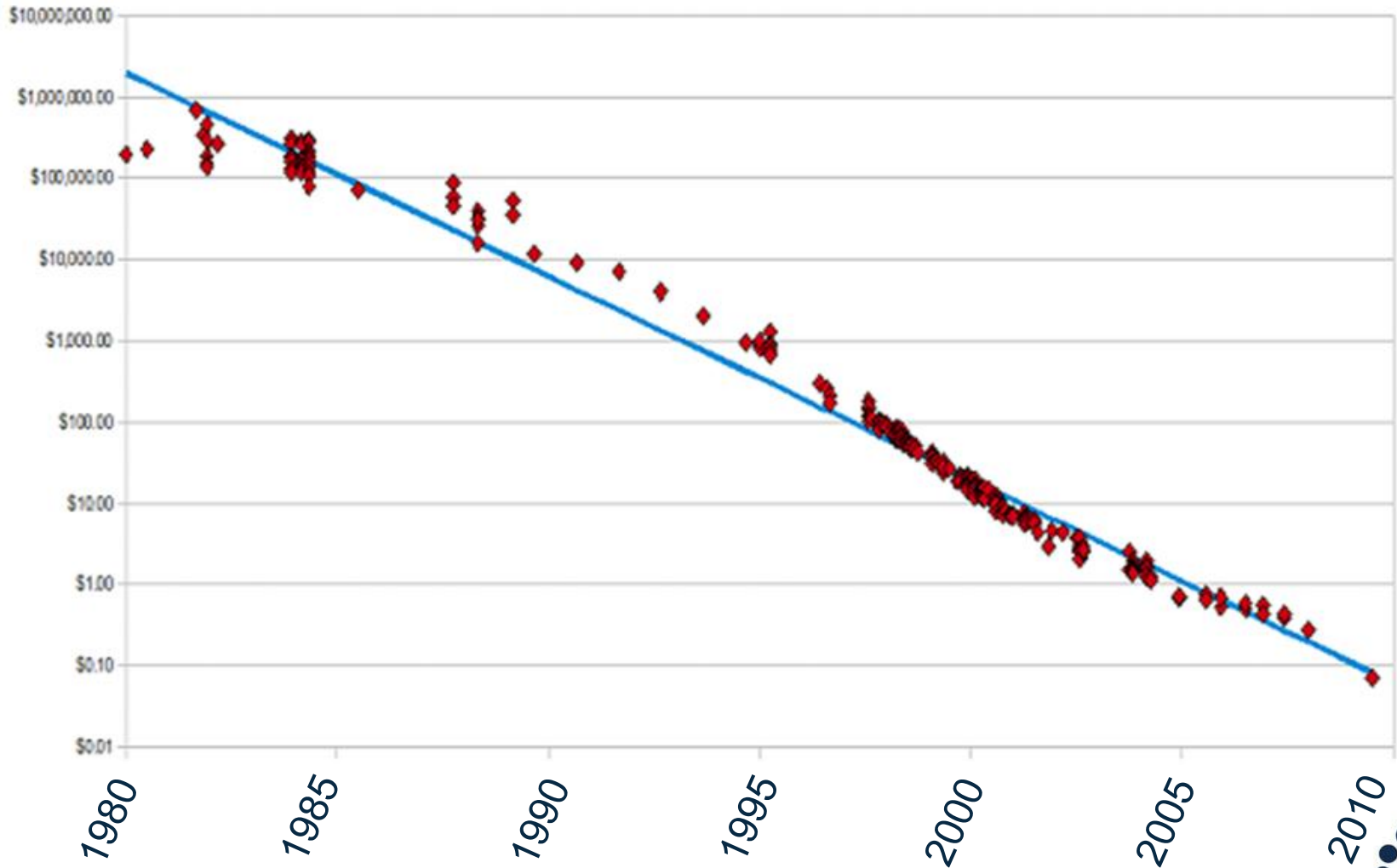
- Automatically places data on storage media with the appropriate cost profile at all points during the data's life cycle
- Ensures data availability through a robust backup and restore capability at the dataset and volume level
- Space management automation to avoid out of space conditions
  - Migration/Recall processing
  - Partial release
  - Extent reduction
  - Small Data Set Packing (SDSP)
- “New” Storage Tiers support - Class Transitions within the Primary Hierarchy

# Change Happens – mostly Good things

- New application workloads
- New Technologies
- Requirements (RTO/RPO)
- Laws, Regulations
- Disk costs
- System z CPU costs



# Hard drive cost per GB 1980-2010



# 2008-2009 – The Storage Media Shift

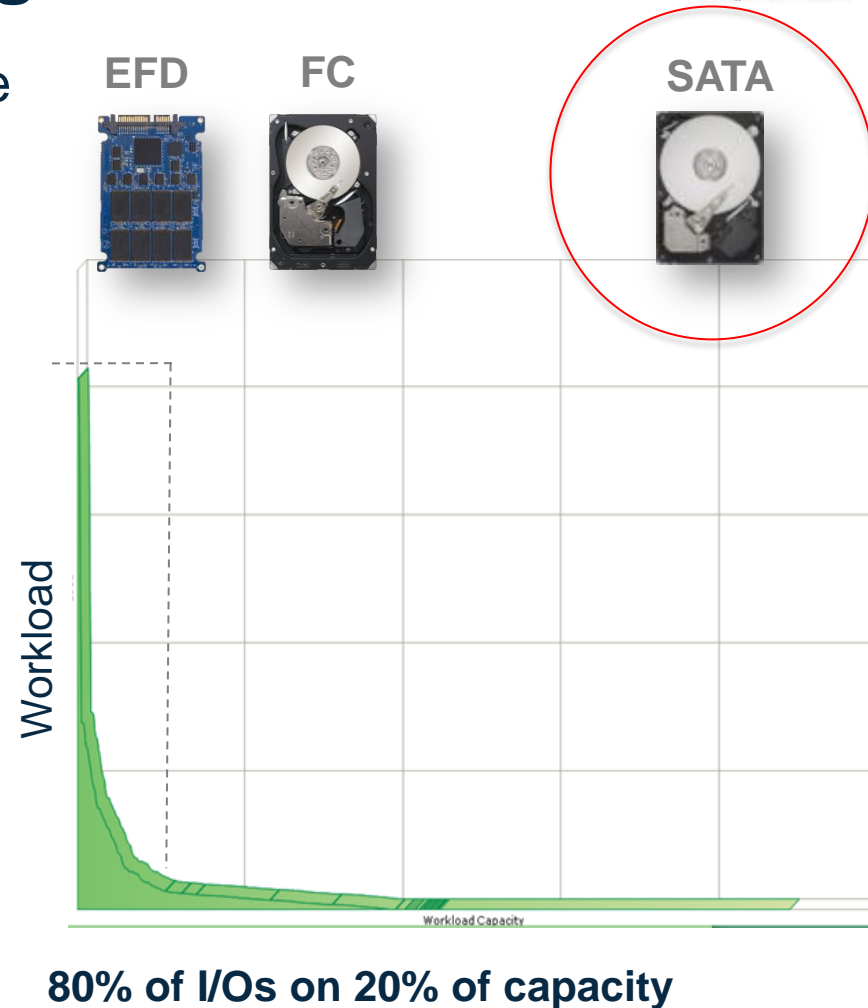
| Disk                            | Cost/GB | IOPS/GB | Response Time |
|---------------------------------|---------|---------|---------------|
| 15 K RPM                        | 1       | 1       | 6 ms          |
| Serial ATA (SATA)<br>7,200 RPM  | 1/3     | 1/6     | 12 ms         |
| Enterprise Flash<br>Drive (EFD) | 8       | 30      | < 1 ms        |

# Automated Storage Tiering

- With information growth trends, all Fibre Channel (FC) configurations will:
  - Cost too much
  - Consume too much energy
  - Take up too much space
- Automated Storage Tiering *leverages* disk drive technologies

**Skew:** At any given time, only a small address range is active – the smaller the range, the better

**Persistence:** If an address range is active (or inactive), it remains so for a while – the longer the duration, the better

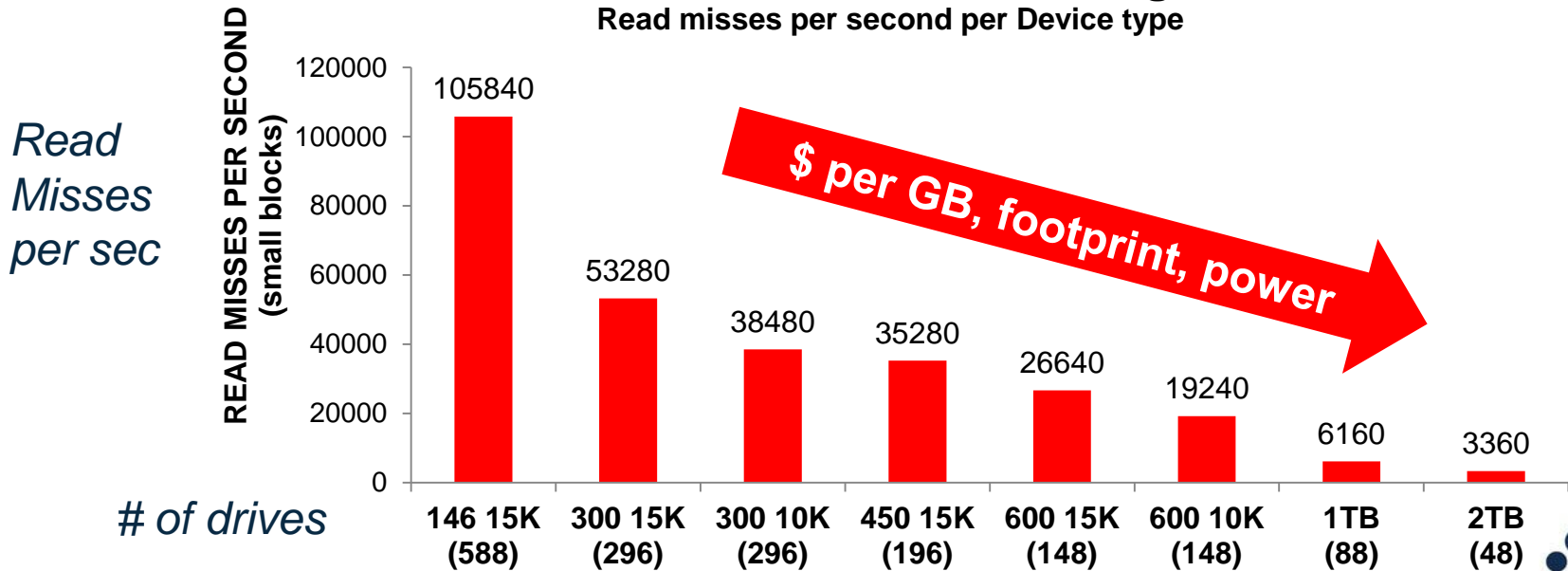


# Wide striping and short stroking are common practice

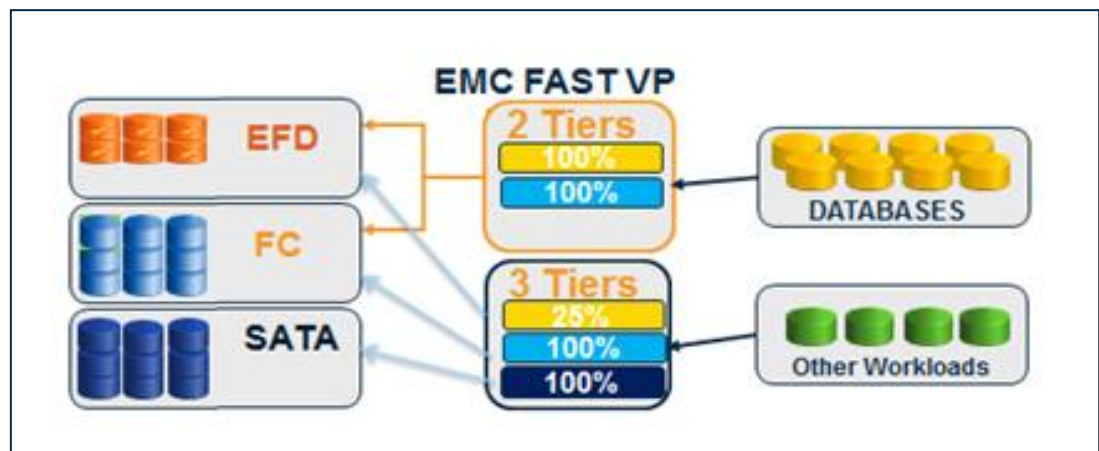
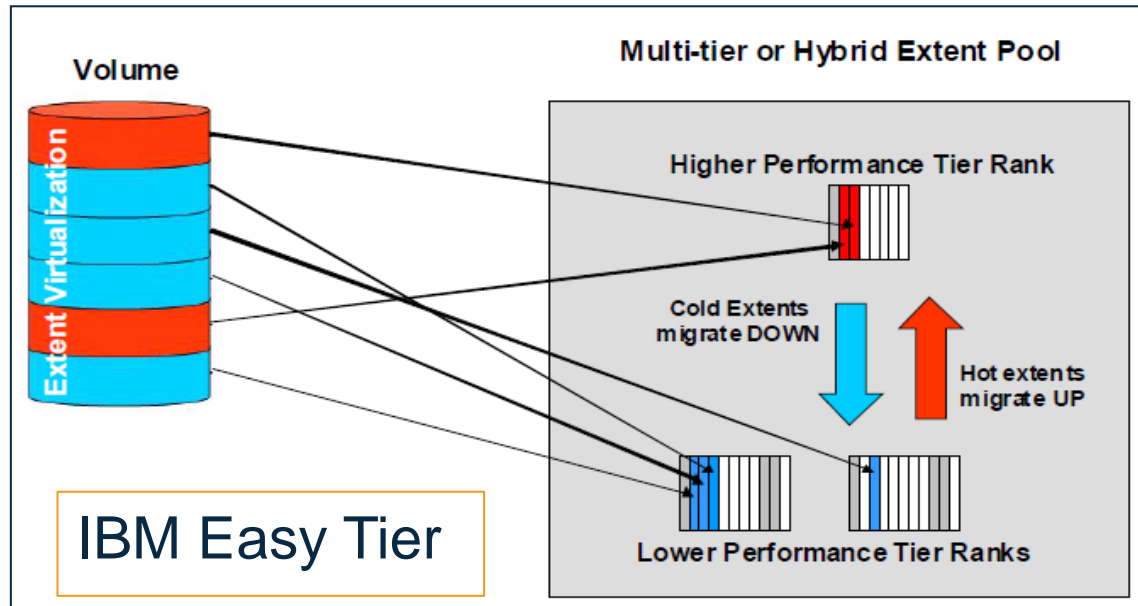
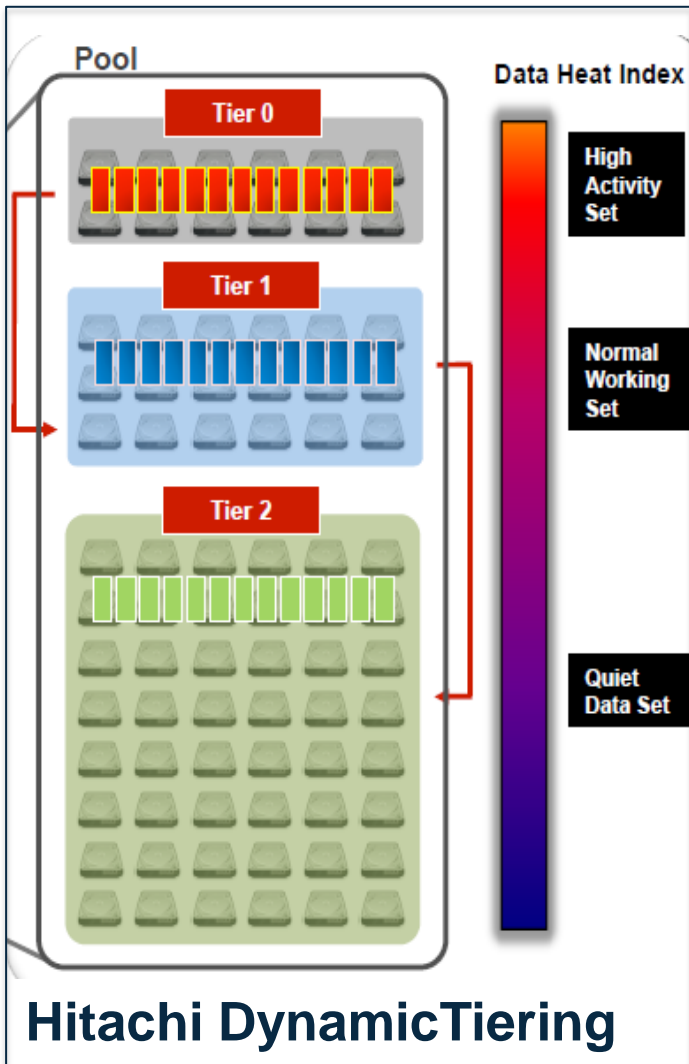
The vast majority of online workloads enjoy high cache-hit percentages, but **service levels are dictated by read-misses during transitional periods**

## 60 TB of RAID 5 3+1 Storage

Read misses per second per Device type



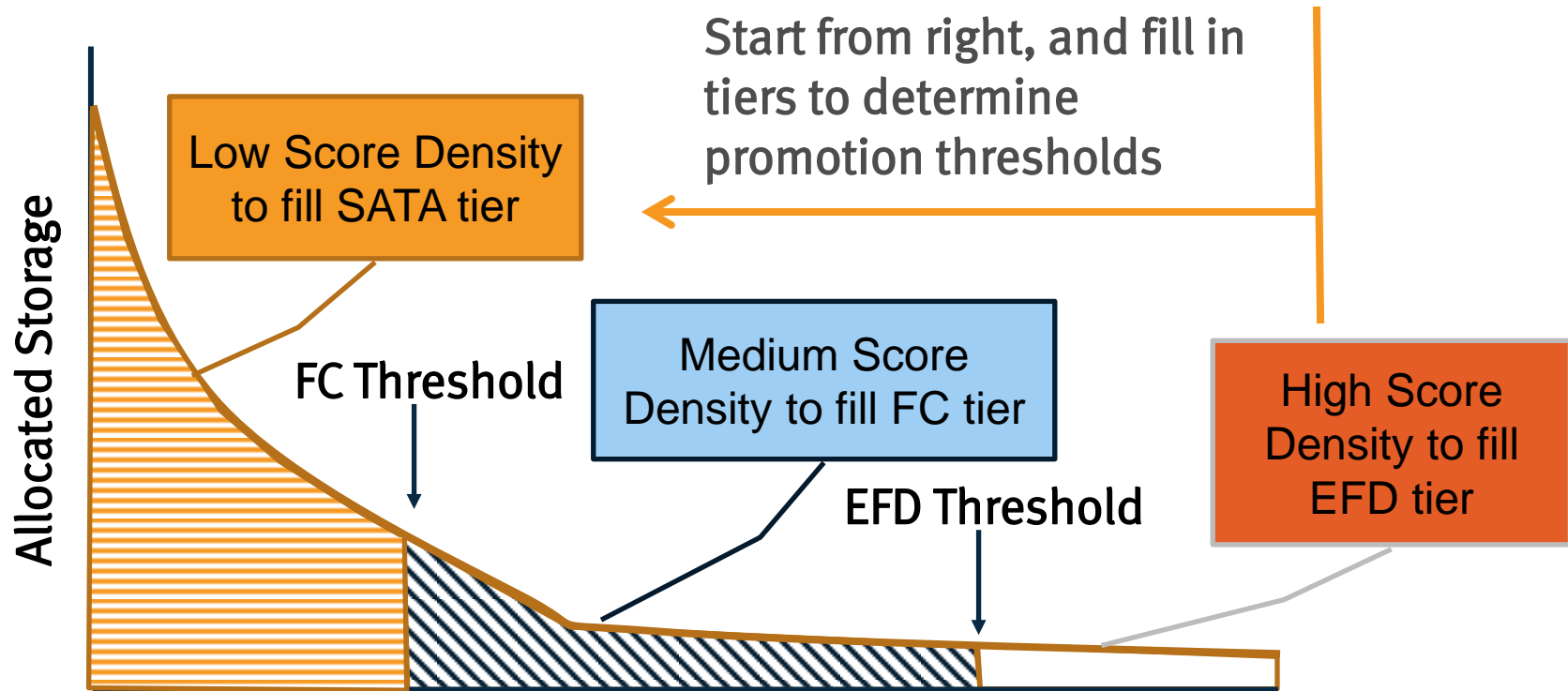
# Automated Storage Tiering Solutions



## Extent/Page Based Movement between Tiers

Complete your session evaluations online at [www.SHARE.org/Pittsburgh-Eval](http://www.SHARE.org/Pittsburgh-Eval)

# Automated Storage Tiering – Real Time Performance Analysis



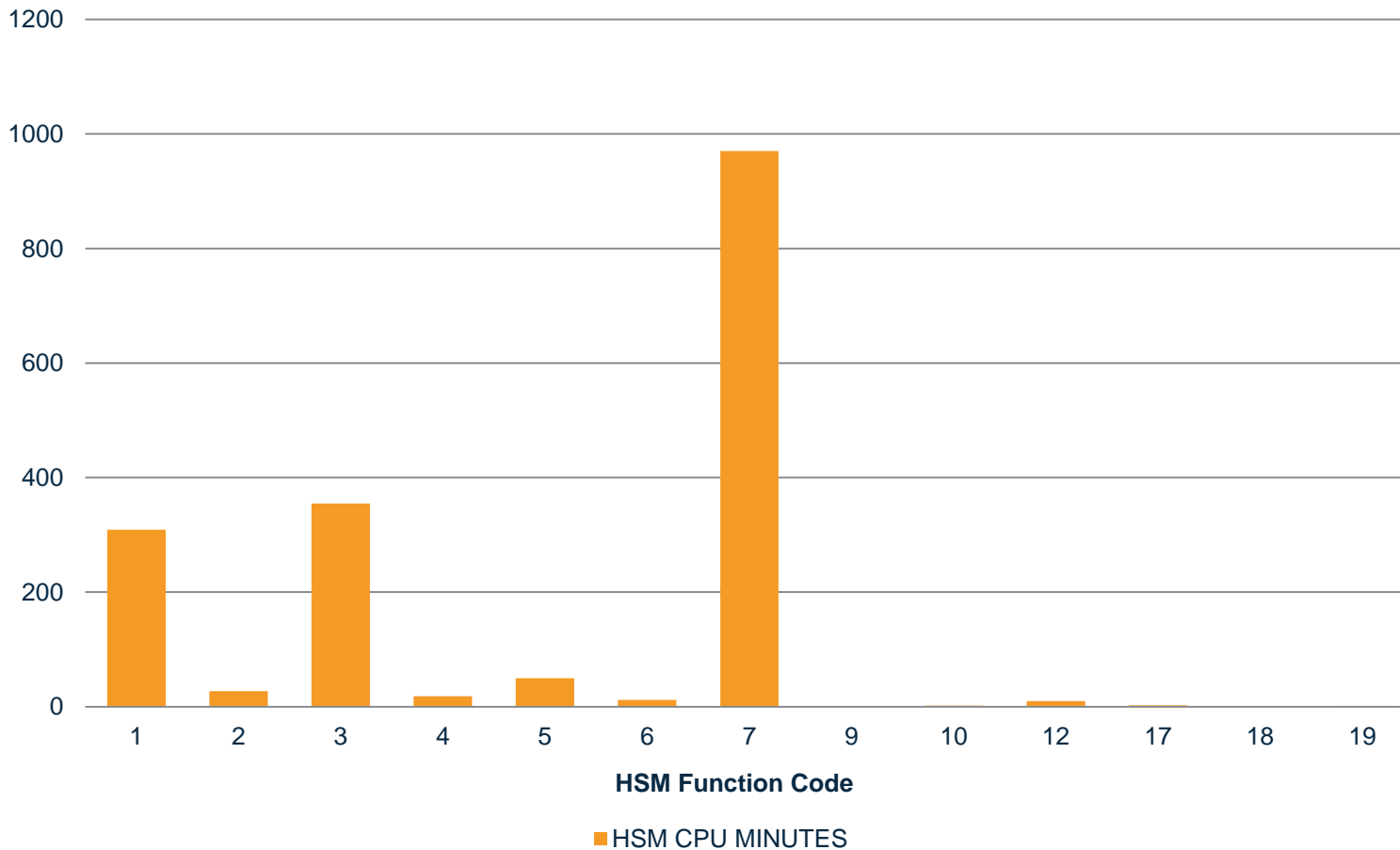
Does not use System z processor or channel resources to move the data across the storage tiers

# Reporting and Analytics

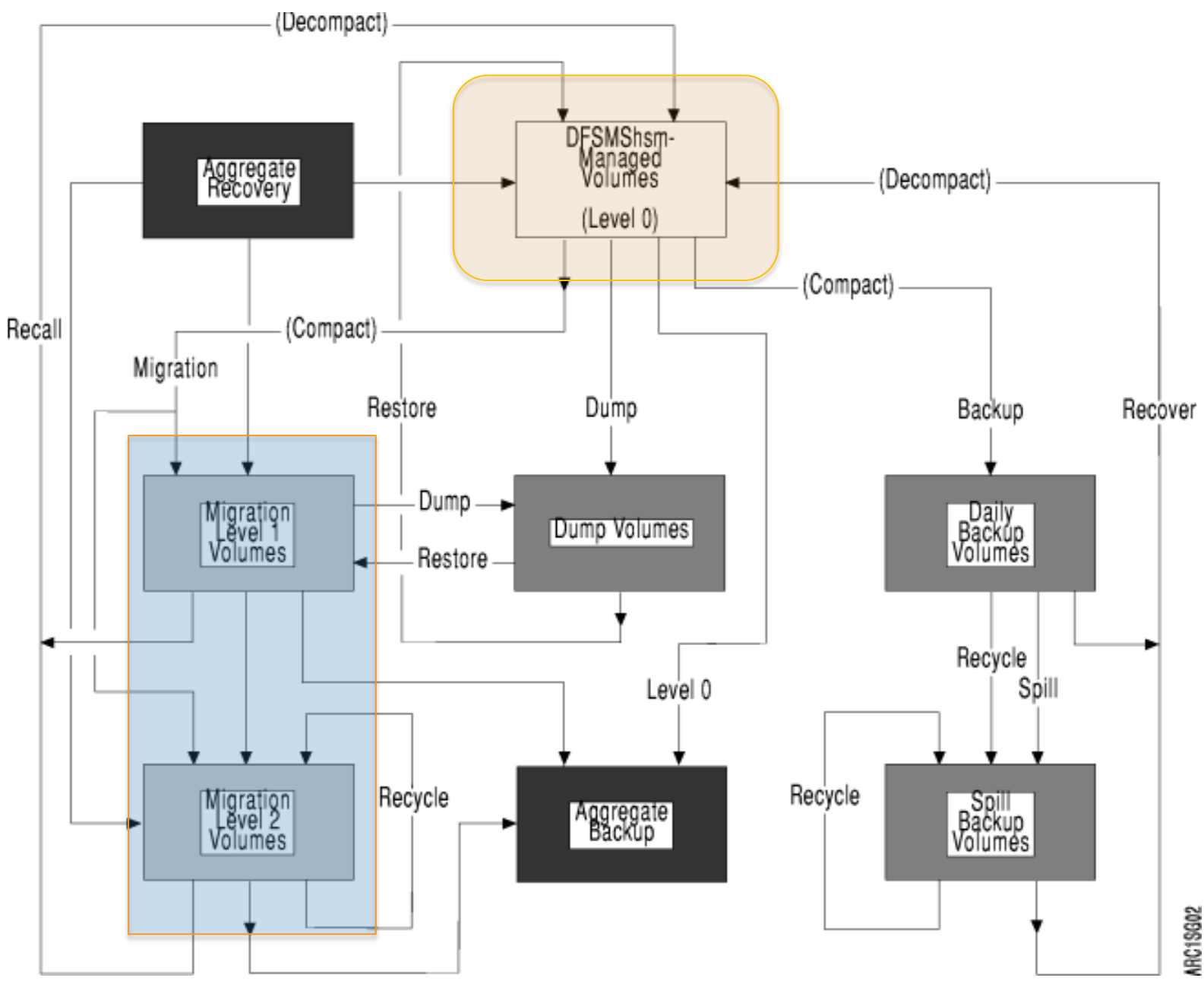
- Peak week of Functional Statistics Records (FSR)
- Corresponding SMF 30-2 Records (Interval Termination)
- IDCAMS DCOLLECT D, M and V Records
- Roll your own or Vendor Management Tool
  - SAS (etc.) or CA, BMC (Software Vendor)
- SHARE Sessions –
  - ICE-PAK 21<sup>st</sup> Century -15981: Customer Maximizes ROI - Applying data intelligence to storage and backup \*SHARE Live\*
  - 16128: DFSMSHsm CDS Deep Dive – Glenn Wilcock IBM
  - 15766: Reinventing How You Maintain DFSMSHsm with Tivoli Advanced Tools - Chris Taylor and Steve W. Clar

# Resource Consumption - CPU

## HSM CPU Minutes by Function







ARC19G02

## HSM FUNCTION CODES

- 1=PRIMARY TO LEVEL 1 MIGRATION
- 2=LEVEL 1 TO LEVEL 2 MIGRATION,
- 3=PRIMARY TO LEVEL 2 MIGRATION
- 4=RECALL FROM LEVEL 1 TO PRIMARY
- 5=RECALL FROM LEVEL 2 TO PRIMARY
- 6=DELETE A MIGRATED DATA SET
- 7=DAILY BACKUP
- 8=SPILL BACKUP
- 9=RECOVERY
- 10=RECYCLE BACKUP VOLUME
- 11=DATA SET DELETION BY AGE
- 12=RECYCLE MIGRATION VOLUME
- 13=FULL VOLUME DUMP
- 14=VOLUME OR DATA SET RESTORE
- 15=ABACKUP FUNCTION
- 16=ARECOVER FUNCTION
- 17=EXPIRE PRIMARY OR MIGRATED DATA SETS
- 18=PARTREL FUNCTION
- 19=EXPIRE OR ROLL OFF INCREMENTAL BACKUP VERSION
- 20=(H)BDELETE INCREMENTAL BACKUP VERSION

# Migrate to Tape/Disk (Compression) L2 VS Store Data on SATA Tier for Primary Space

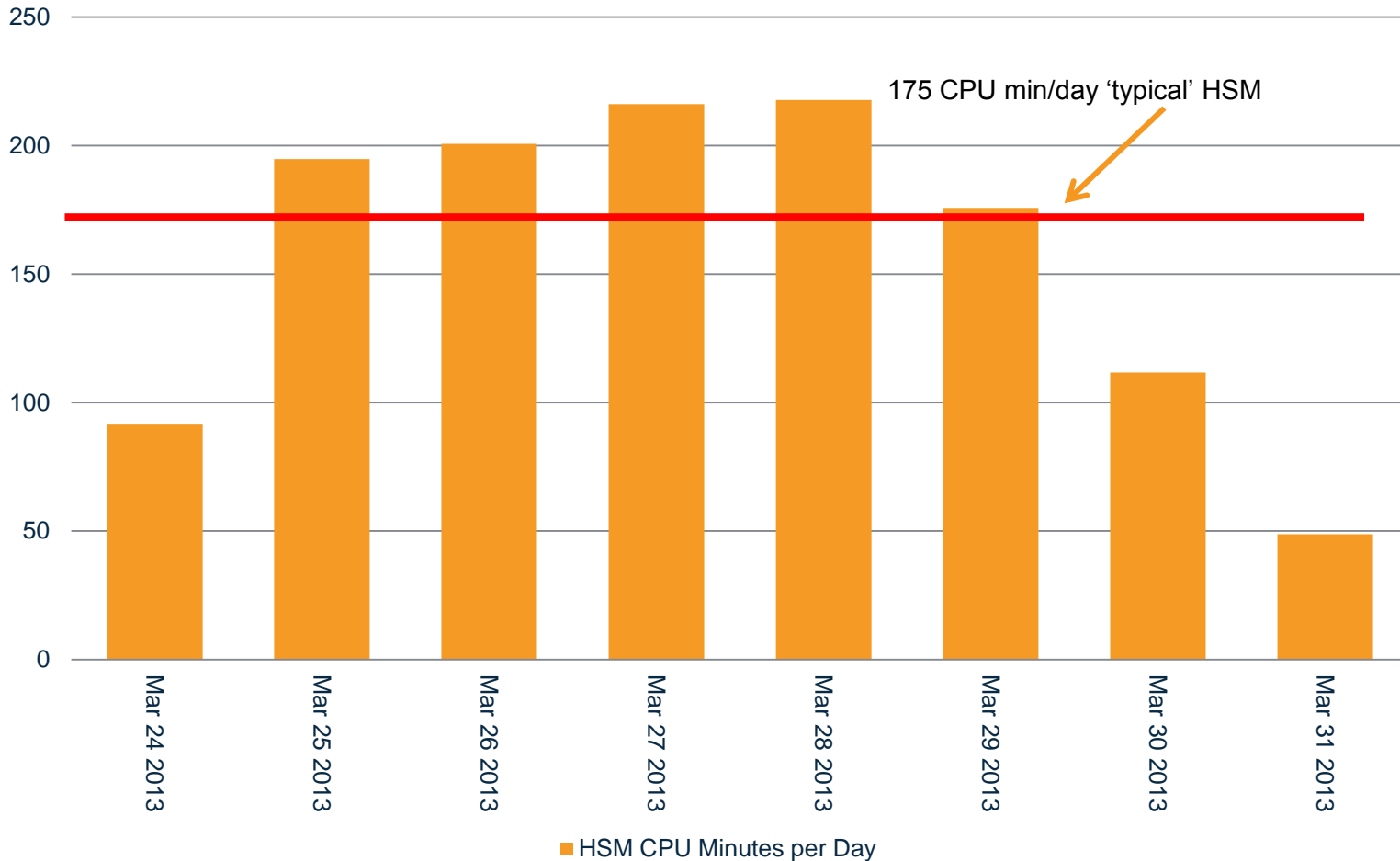


# Measurements and Calculation Needed

- Eliminate Primary to ML1 migration
  - Potential CPU and channel resource savings, particularly if data is also being compressed.
    - Function code 1 for migration activity
    - Function code 4 for recall activity.
- Increase the amount of time data is allowed to stay on primary (ML0) storage before being eligible for migration to ML2
  - Reducing the frequency of ML2 migration to reduce CPU and channel resources consumed migrating data to ML2.
- Migration and recall activity can be reduced by increasing the number of days data is allowed to stay on primary storage.
  - Determine necessary z/OS Storage space increase (VTOC free-space) which will be added in SATA tier (targeted for inactive extent moves by Automated Storage Tiering Solution)

# Potential Reduction in CPU Minutes

## HSM CPU Minutes per Day



# Cost Analysis – The Numbers



## Estimating Financial Benefits of Automated Storage Tiering Solution

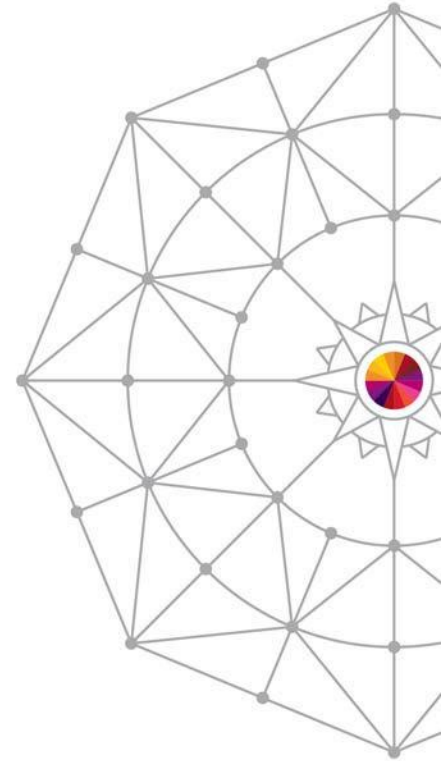
1. Migration + Recall CPU use in Minutes
2. Cost of CPU Minute
  1. \$ / CPU Minute
3. Cost of SATA capacity
  1. Be generous – (see slide 12) – SATA ~70 IOPS RRM
  2. Capacity Based Priced Software
4. Analysis Period



# Summary

- Storage management in the z/OS environment will clearly be undergoing significant change in the coming years with Automated Storage Tiering
- 
- Achieving benefits described in this session requires a thorough understanding of the environment and how HSM and tape is used for space management
- A study of the existing DFSMSHsm environment is an essential first step in gaining this understanding.

# QUESTIONS? & FEEDBACK!



#SHAREorg



SHARE is an independent volunteer-run information technology association that provides **education, professional networking and industry influence.**



# Rethinking HSM Policies with Automated Storage Tiering

*Session ID 15997*

*Tuesday, August 5, 2014: 03:00 PM - 04:00 PM,  
DLLCC, Room 305*



SHARE is an independent volunteer-run information technology association that provides education, professional networking and industry influence.

