#SHAREorg

SHARE
Technology · Connections · Results

Getting Even More (and a lot less) From Storage Virtual Provisioning and Automated Storage Tiering

Tony Negro
EMC Corporation

Tuesday, Februrary 05, 2013 1:30 PM
Session Number 12174

SHARE in San Francisco 2013

Objectives

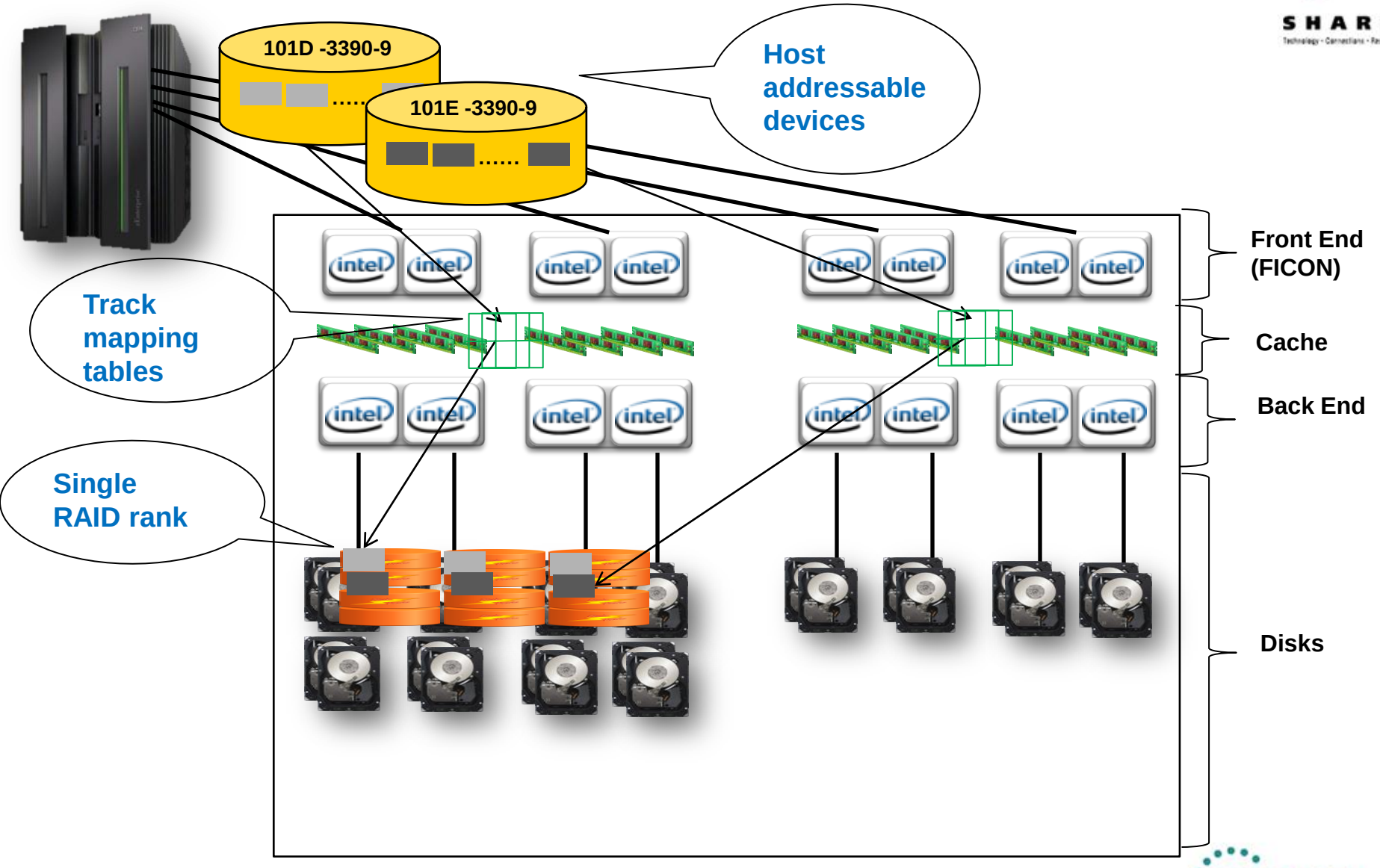
At the end of this session, you will better understand

- Virtual Provisioning and Virtual Pools for Count Key Data (CKD) devices
- Advanced utilities for Virtual Pools and array based compression featues for CKD devices
- Fully Automated Storage Tiering for Virtual Pool (FAST VP) theory of operation
- Potential benefits that can be achieved when deploying these technologies

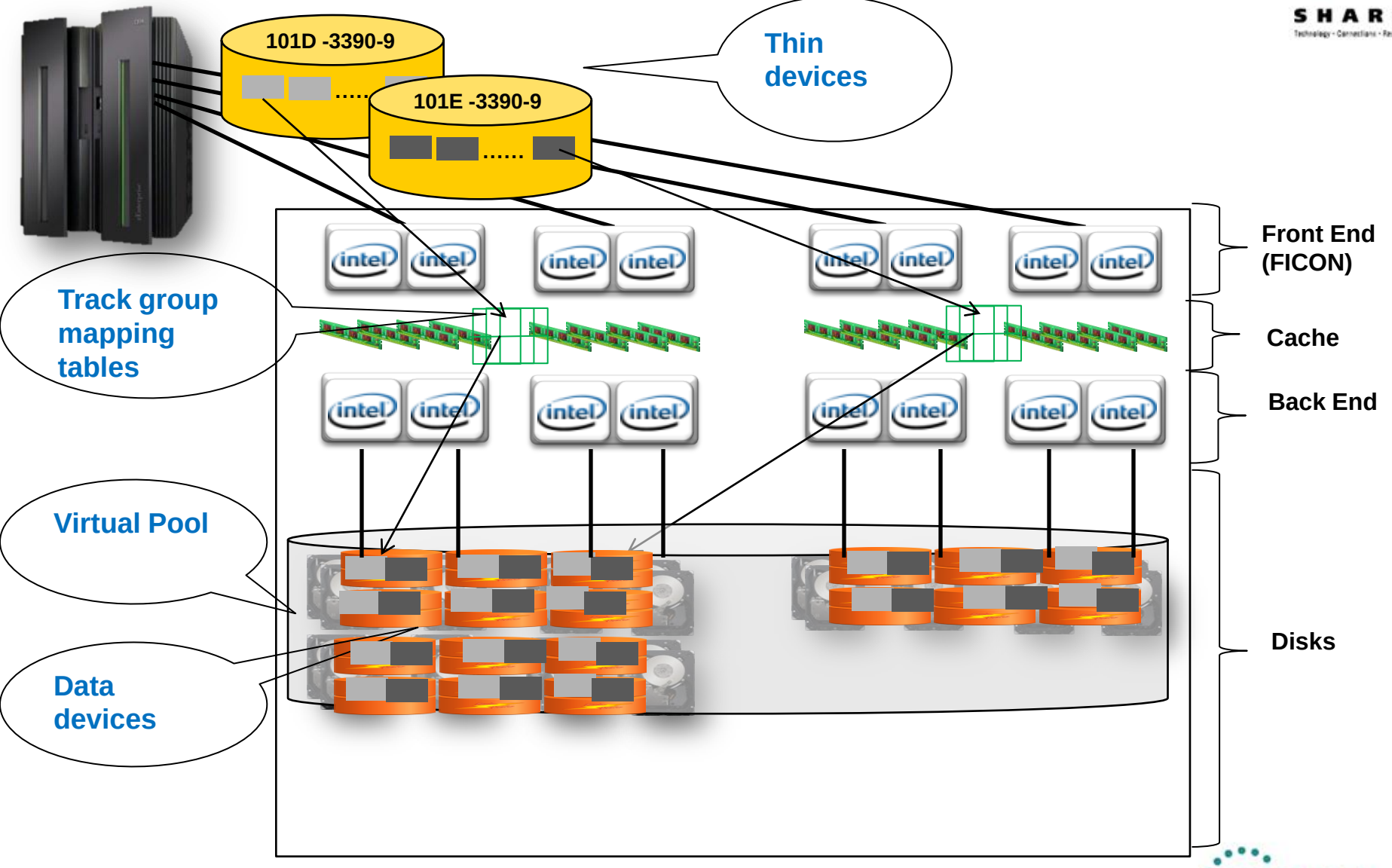
Agenda

- Storage Provisioning Overview
  - Provisioning terminology
  - Thin Provisioning mapping
  - Provisioning challenges
- Storage Tiering Consideration
  - Why use different tiers
  - Data mobility between storage tiers
- Automated Storage Tiering Concepts
  - Fully Automated Storage Tiering for Virtual Pool (FAST VP) components
  - FAST VP Implementation

Standard Provisioning Concept



Virtual (thin) Provisioning Concept



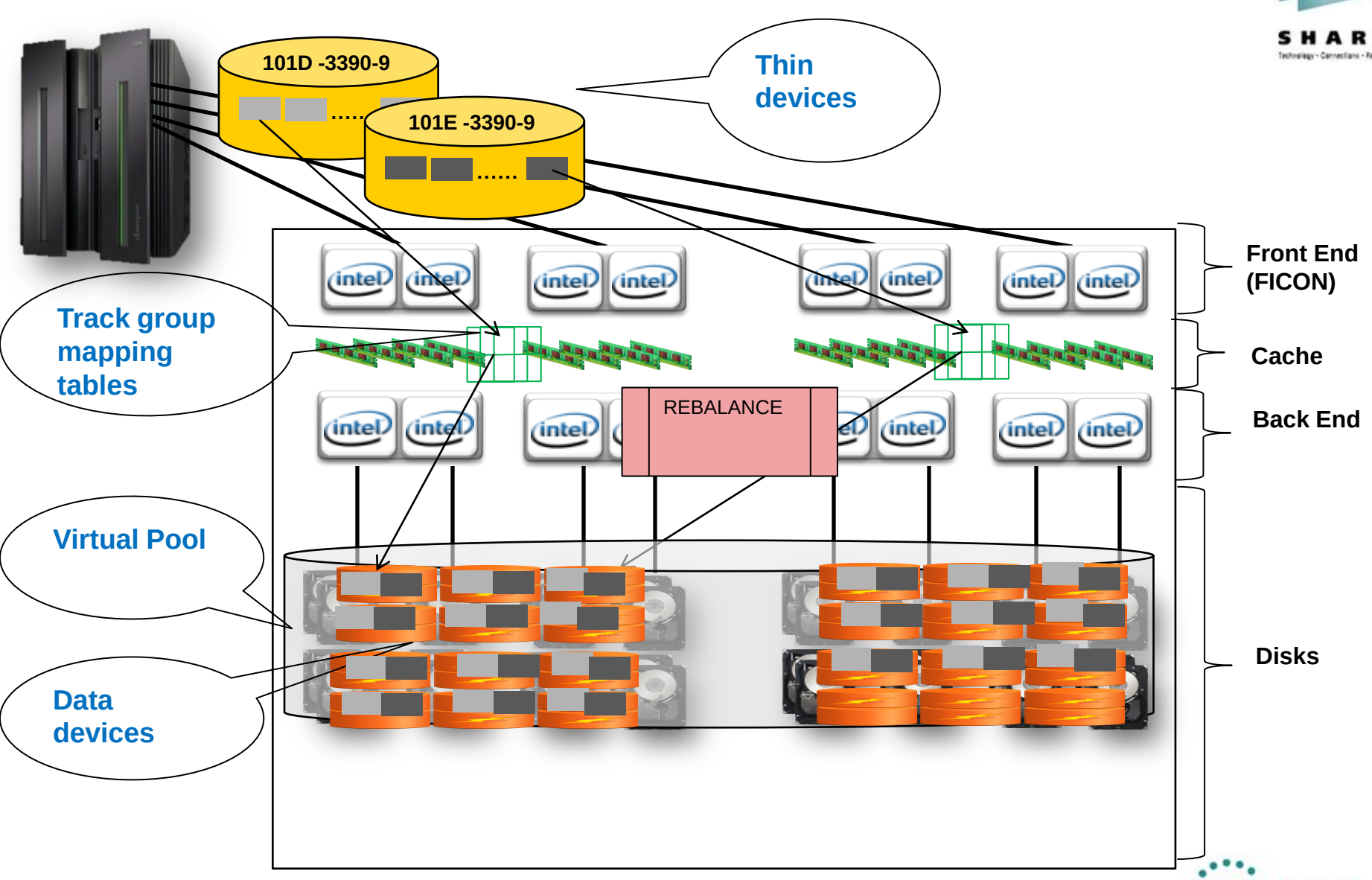
Virtual Provisioning Terminology

TERM	DESCRIPTION
Host Accessible Device	A ‘device’ that is presented on a FICON channel for host use
Thin Device (Virtual Device)	A ‘Host accessible device’ that has no storage directly associated with it
Internal Device	A ‘device’ used for internal function of the array
Data Device	An ‘Internal Device’ that provides storage capacity to be used by a ‘Thin Device’
Thin Pool (Virtual Pool)	A collection of ‘Data Devices’ that provide storage capacity for ‘Thin Devices’
Track Group	The size of the smallest contiguous region of a ‘device’ for which an extent mapping can occur
Extent Mapping	Specifies the relationship between the ‘Thin Device’ and ‘Data Device’ extents.
Bind	The process by which one or more ‘Thin devices’ are associated to a ‘Thin Pool’
Thin Device Allocated Capacity	The capacity that has been allocated from the thin pool capacity for the exclusive use of a thin device.
Thin Device Written Capacity	The capacity on a 'Thin Device' that was written to by a host. In most implementations this is a subset of the 'Thin Device Allocated Capacity'.

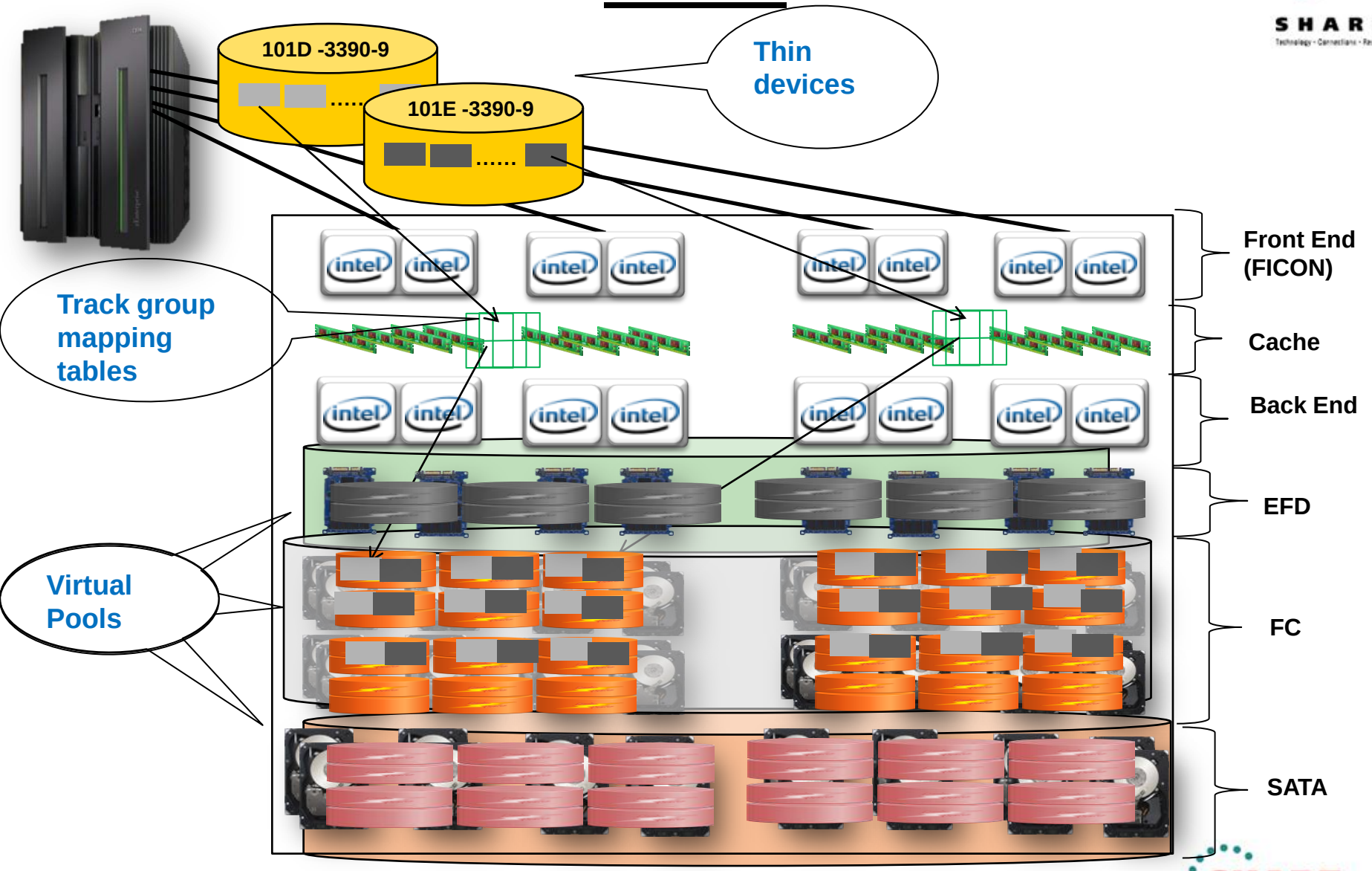
Standard Provisioning issues

- Slow process
- Complex physical layout
- Physical disk contention (“hot spots”)
- Expanding physical capacity can be a challenge

Virtual Provisioned Environment



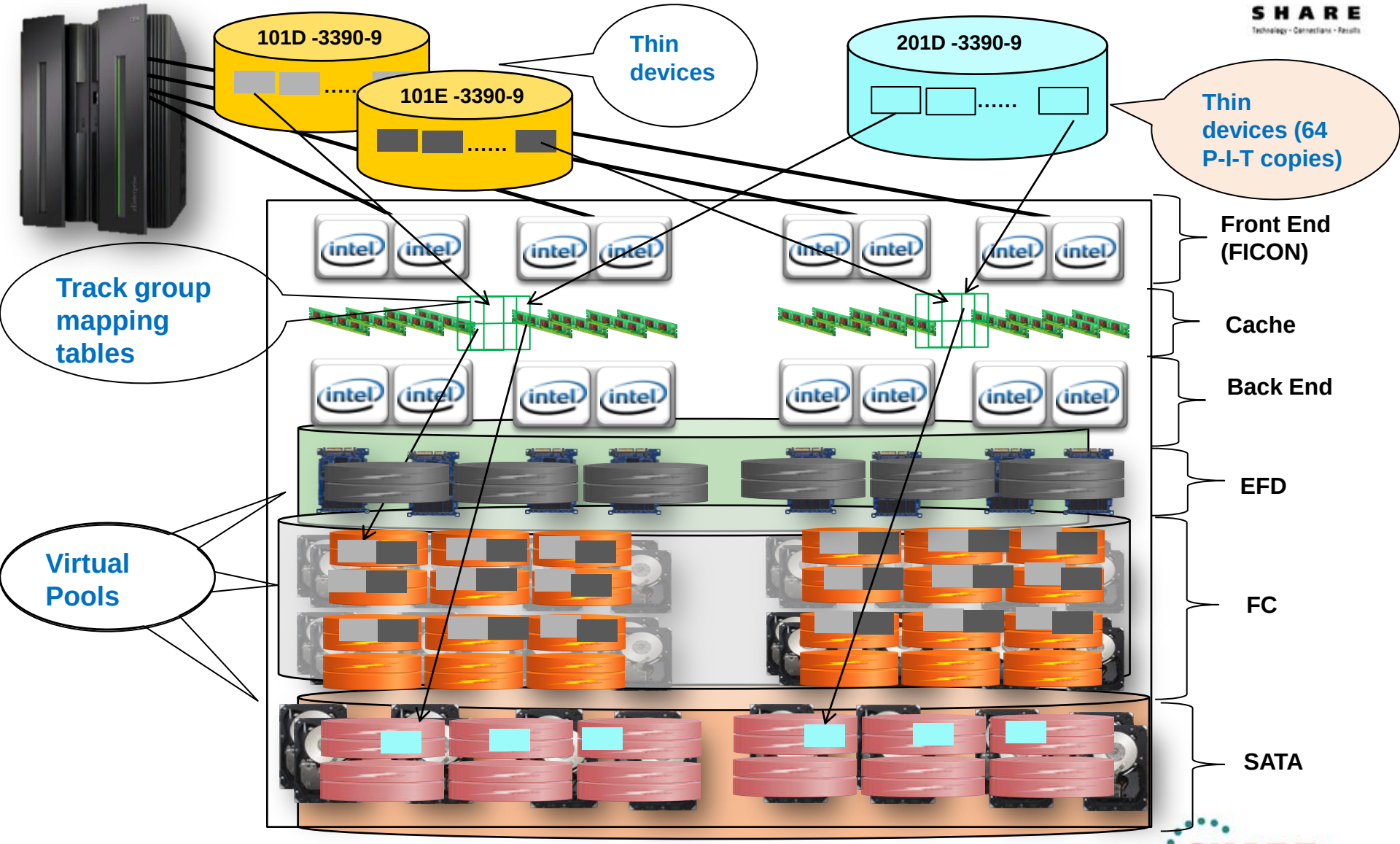
Virtual Provisioned Tiered Environment



2008 – Fundamental Storage Media Shift

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

Virtual Provisioned Tiered Environment



SATA Pool thin device allocation and DSS Backup throughput

EMC Unisphere for VMAX V1.1.0.5

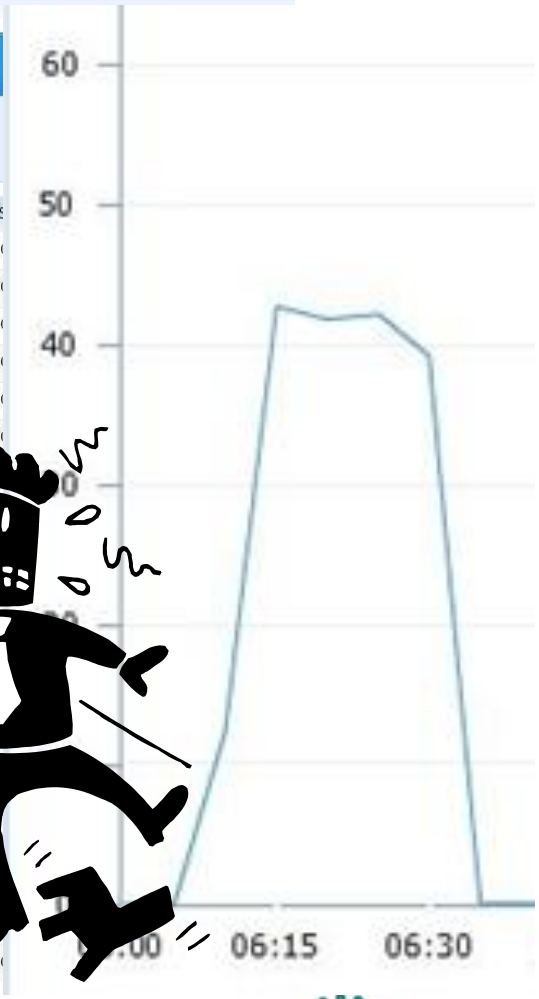
Home | System | Storage | Hosts | Data Protection | Performance | Support

000195700455 > Storage > Thin Pool

Bound Volumes For Thin Pool

Name	Configuration					Total Subscribed %	Total Written (GB)	
0D8B	TDEV					0 %	1.98	N
0D8C	TDEV					0 %	2	N
0D8D	TDEV					0 %	4.06	N
0D8E	TDEV					0 %	4.38	N
0D8F	TDEV					0 %	3.92	N
0D90	TDEV					0 %	1.29	N
0D91	TDEV							
0D92	TDEV						5.7	
0D93	TDEV					0 %	1.3	
0D94	TDEV					0 %	0.9	
0D95	TDEV				4.42	0 %		
0D96	TDEV	CKD-3390			3.75	0 %		
0D97	TDEV	CKD-3390			1.84	0 %		
0D98	TDEV	CKD-3390	2 %	7.93	0.14	0 %		
0D99	TDEV	CKD-3390	2 %	7.93	0.12	0 %		
0D9A	TDEV	CKD-3390	4 %	7.93	0.28	0 %		
0D9B	TDEV	CKD-3390	62 %	7.93	4.88			
0D9C	TDEV	CKD-3390	2 %	7.93	0.18			
0D9D	TDEV	CKD-3390	3 %	7.93	0.23			
0D9E	TDEV	CKD-3390	55 %	7.93	4.37	0 %	4.37	N

Running a “special” DB query for the CIO (against the daily P-I-T copy of PROD DB), and it’s REALLY running S_L_O_W….Help!

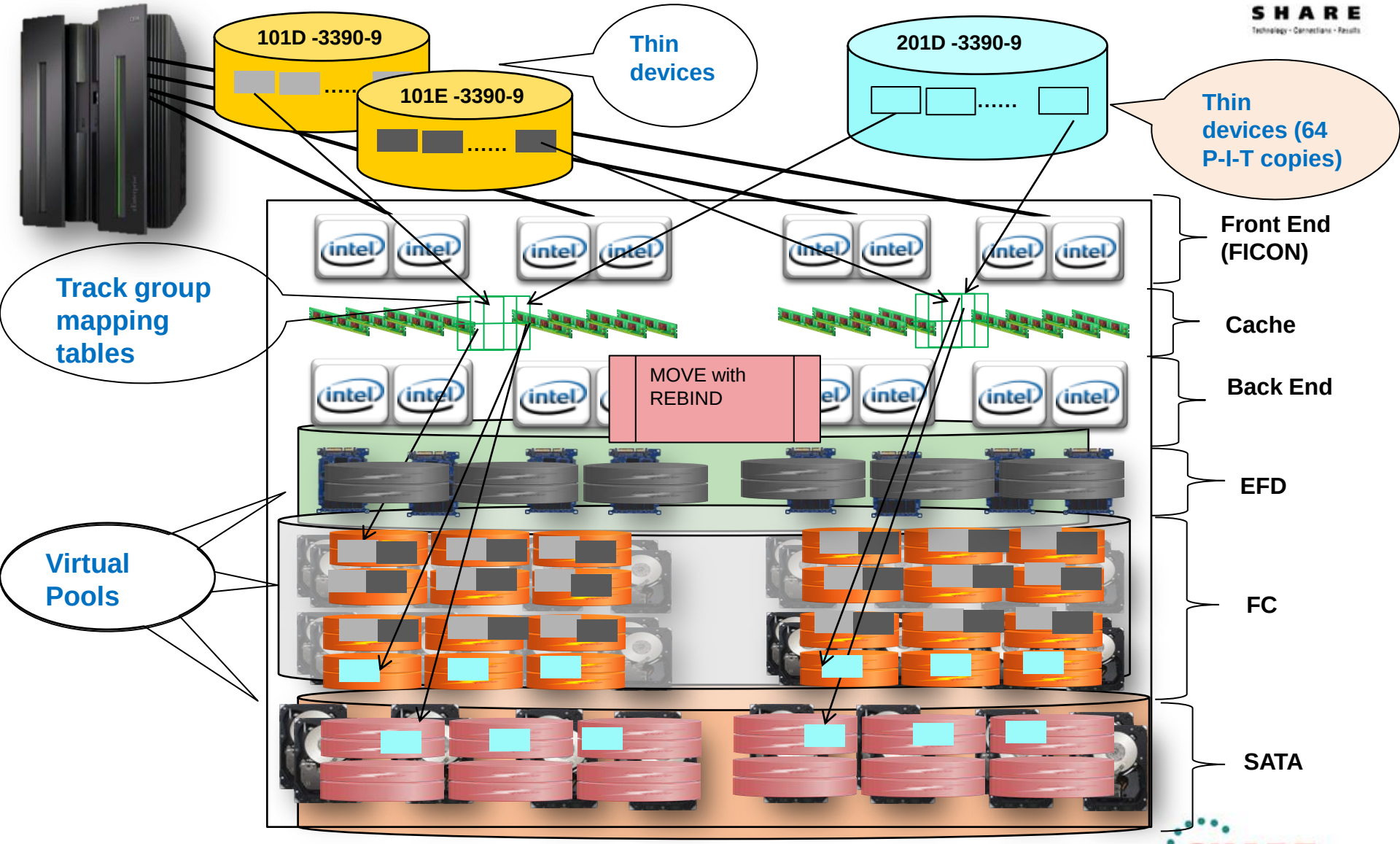


Let's have a closer look….

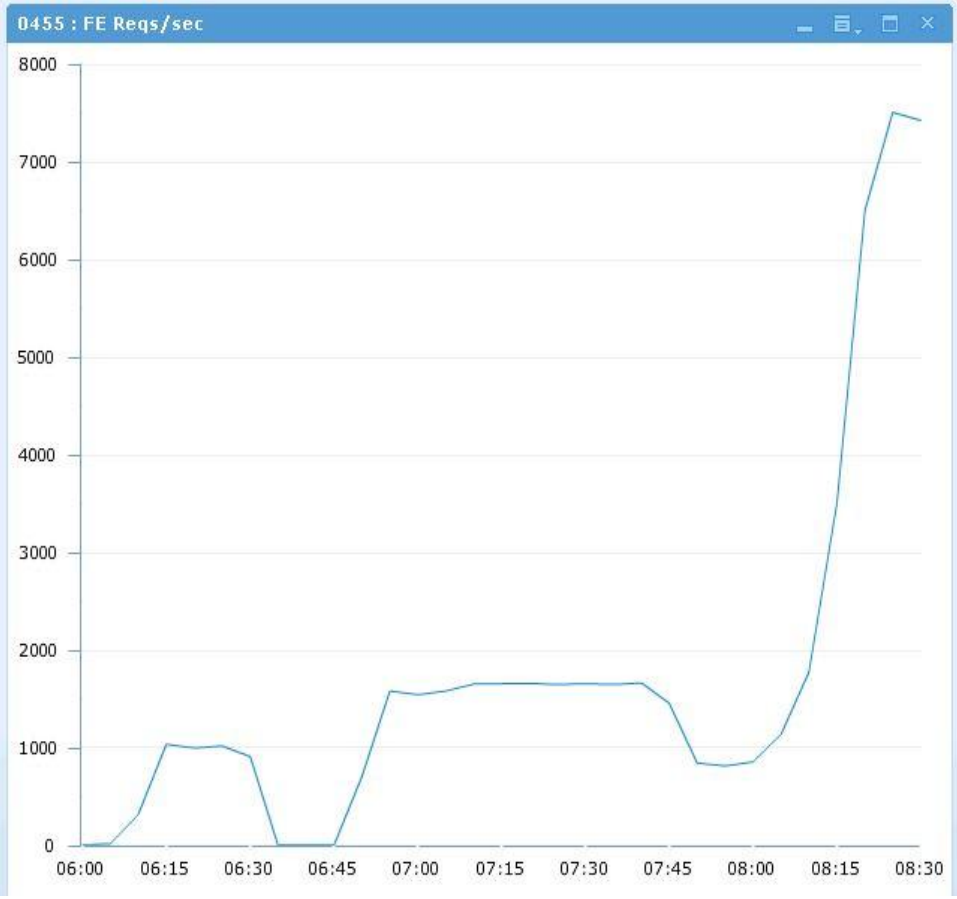
F V11A DEV

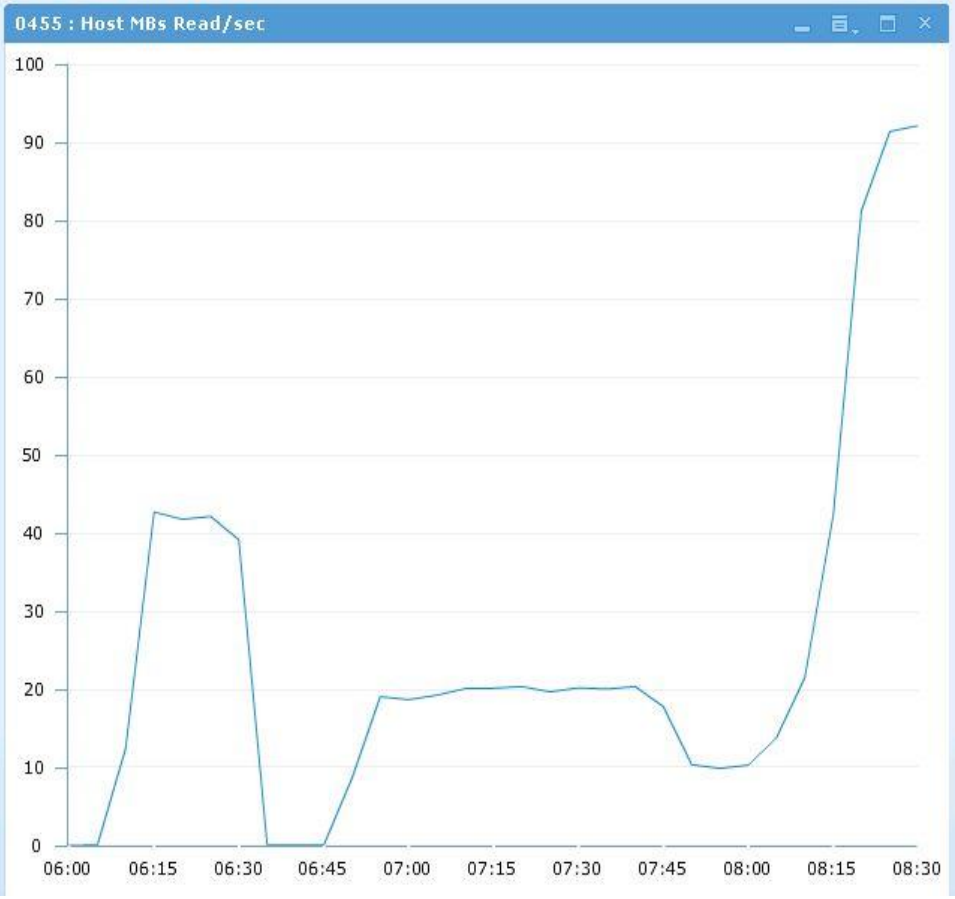
10:52:32	I=53%	DEV			ACTV	RESP	IOSQ	-DELAY-		PEND	DISC	CONN	%D	%D
STG GRP	VOLSER	NUM	PAV	LCU	RATE	TIME	TIME	CMR	DB	TIME	TIME	TIME	UT	RV
	SYMD7A	16C8	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
	SYMD7B	16C9	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
	SYMD7C	16CA	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
	SYMD7D	16CB	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
	SYMD7E	16CC	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
	SYMD7F	16CD	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
	SYMD80	16CE	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
	SYMD81	16CF	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG	DBS001	16D0	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG	DBS002	16D1	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG	DBS003	16D2	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG	DBS004	16D3	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSDSG	DBS005	16D4	1.6H	0013	31.22	32.5	1.33	.00	.00	.104	30.7	.390	62	0
DBSDSG	DBS006	16D5	1.0H	0013	148.1	2.40	.281	.00	.00	.106	1.88	.131	30	0
DBSDSG	DBS007	16D6	1.3H	0013	36.47	27.7	1.14	.00	.00	.106	26.1	.349	75	0
DBSDSG	DBS008	16D7	1.9H	0013	36.07	33.8	0.58	.00	.00	.107	32.8	.397	64	0
DBSDSG	DBS009	16D8	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSDSG	DBS010	16D9	1.0H	0013	4.222	30.1	.000	.00	.00	.107	29.7	.350	13	0
DBSDSG	DBS011	16DA	1.4H	0013	39.91	30.2	2.09	.00	.00	.105	27.6	.356	78	0
DBSDSG	DBS012	16DB	1.7H	0013	31.28	31.8	0.67	.00	.00	.103	30.6	.401	57	0
DBSDSG	DBS013	16DC	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0

Non-disruptive volume mobility (VLUN)



Has Random Workload picked up?





Has Response Time improved?

F CPU= 16/ 15 UIC= 65K PR= 0 V11A DEV

12:06:18 I=29%	DEV			ACTV	RESP	IOSQ	-DELAY-		PEND	DISC	CONN	%D	%D
STG GRP VOLSER	NUM	PAV	LCU	RATE	TIME	TIME	CMR	DB	TIME	TIME	TIME	UT	RV
SYMD7A	16C8	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
SYMD7B	16C9	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
SYMD7C	16CA	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
SYMD7D	16CB	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
SYMD7E	16CC	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
SYMD7F	16CD	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
SYMD80	16CE	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
SYMD81	16CF	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG DBS001	16D0	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG DBS002	16D1	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG DBS003	16D2	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSLSG DBS004	16D3	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSDSG DBS005	16D4	1.5H	0013	142.4	6.27	.000	.01	.00	.123	5.77	.374	58	0
DBSDSG DBS006	16D5	1.3H	0013	715.8	.502	.000	.02	.00	.131	.202	.168	21	0
DBSDSG DBS007	16D6	1.5H	0013	169.5	5.58	.111	.01	.00	.124	5.01	.332	60	0
DBSDSG DBS008	16D7	1.0H	0013	173.9	6.06	.109	.01	.00	.122	5.47	.361	T	0
DBSDSG DBS009	16D8	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0
DBSDSG DBS010	16D9	1.0H	0013	19.15	5.20	.000	.01	.00	.118	4.77	.314	10	0
DBSDSG DBS011	16DA	1.3H	0013	183.9	5.60	.103	.01	.00	.121	5.03	.342	79	0
DBSDSG DBS012	16DB	1.5H	0013	142.6	6.48	.265	.01	.00	.121	5.72	.375	58	0
DBSDSG DBS013	16DC	1.0H	0013	0.000	.000	.000	.00	.00	.000	.000	.000	0	0

Support for VP Compression

- Provide configuration, management, and reporting on the VP compression state for thin devices
- NOTE: Compression is supported for devices NOT bound with PERSIST
- New device level commands:
  - COMPRESS – compress data for thin device(s)
  - DECOMPRESS – decompress data for thin device(s)
- New parameters on the POOLATTR command:
  - COMPRESSION(ENABLE) – enable compression for a thin pool
  - COMPRESSION(DISABLE) – disable compression for a thin pool
- Note: Compression is disabled by default and must be enabled by issuing the POOLATTR command with COMPRESSION(ENABLE) parameter before thin devices bound to that pool can be compressed
- Compression is a background task that precludes other background tasks from running (such as reclaim or allcoation)

Recommendations for Compression

- Compression should be used against very idle data - best not to run medium or greater workloads against it.
  - Can impact the performance
- Auto-compression can only be obtained through use of FAST.
- With the exception of perhaps an occasional backup to tape, it is best to fully decompress data before accessing it.
- Suggested use of manual compression (not FAST) for:
  - archiving old user accounts
  - decompress-use-recompress end of quarter activity
  - use it for low cost/low performance active data

Virtual Provisioning Benefits

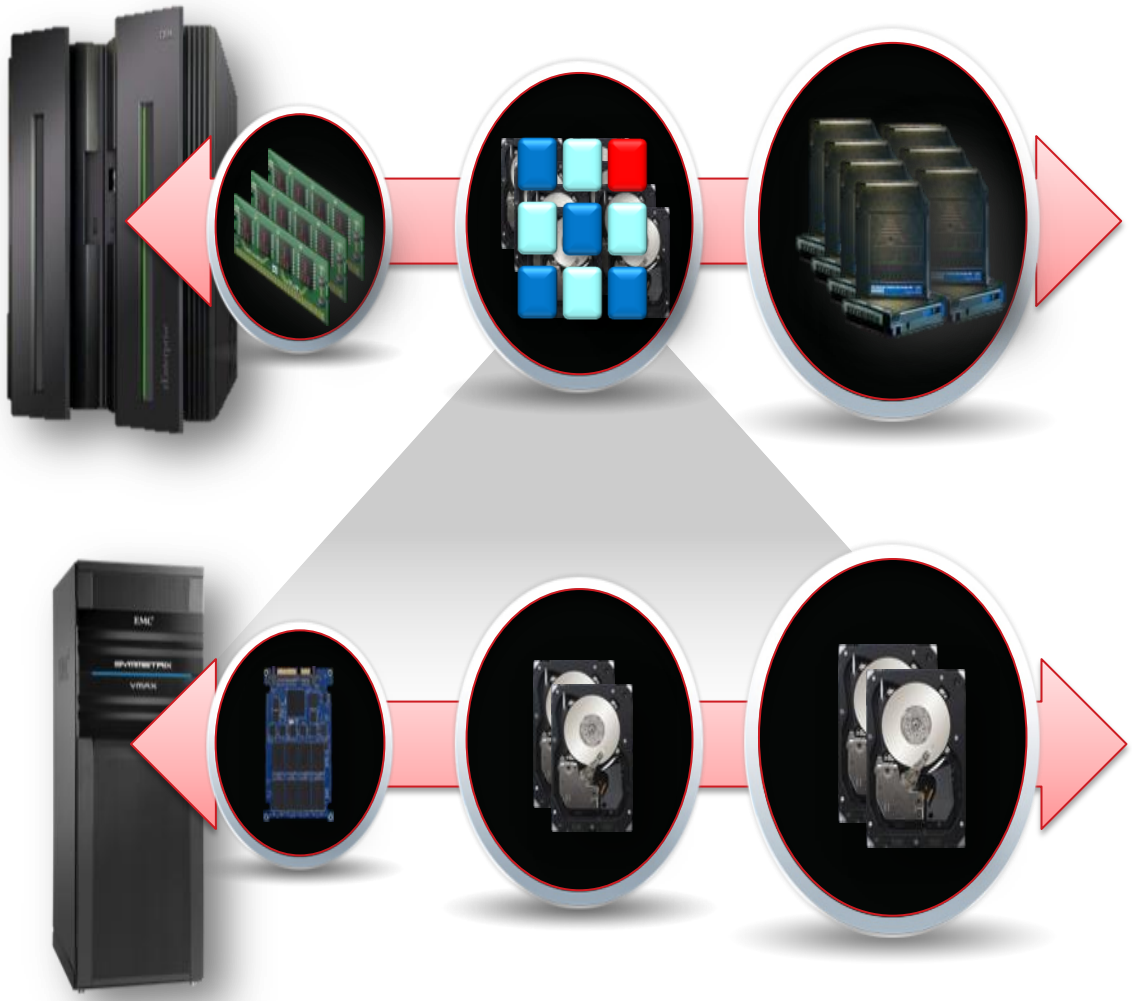
- Enables efficient utilization of available resources
  - Virtual Pools
    - Wide striping distributes I/O across spindles and back end processors
    - Advanced utilities allow greater flexibility
- Provides flexibility when deploying multiple tiers
- Opportunity for ‘Over provisioning’
  - Provision more space than actually exists
  - Consume space as required
- Basis for FAST VP
  - Active performance management at a sub-volume, sub dataset level

Automation at all Layers

SMS &
HSM

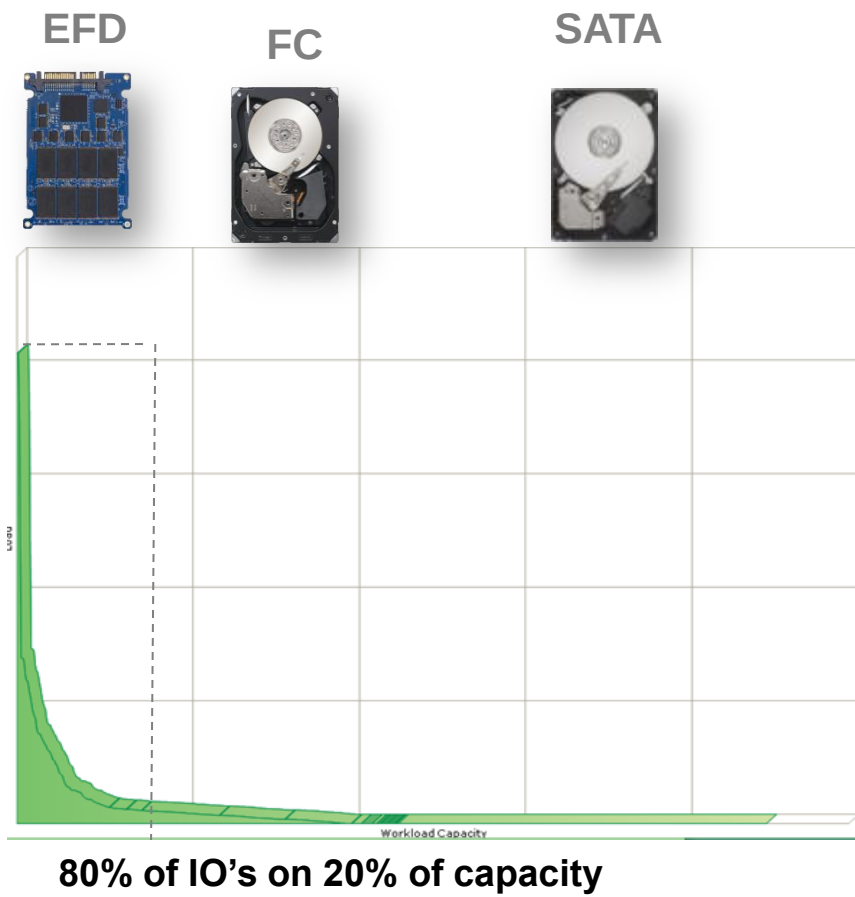
+

FAST VP



Basis for FAST

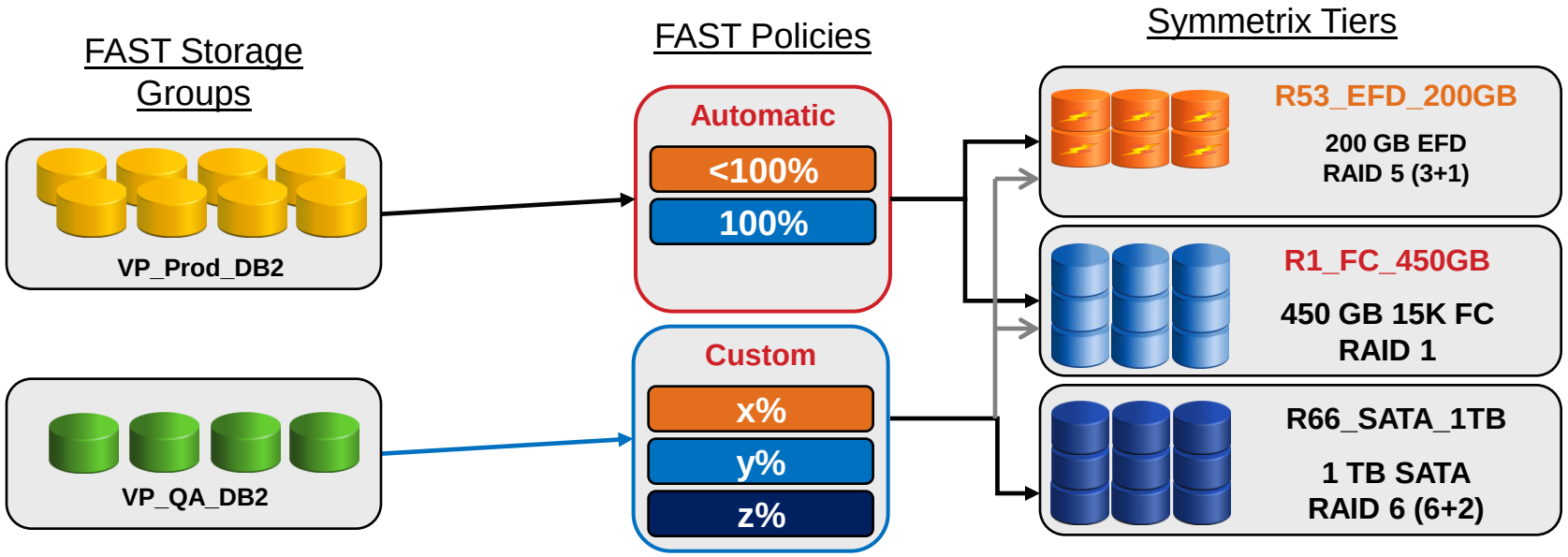
- With information growth trends, all Fibre Channel (FC) configurations will:
  - Cost too much
  - Consume too much energy
  - Take up too much space
- FAST helps by leveraging disk drive technologies
- What makes FAST work in real-world environments?
  - **Skew**: At any given time, only a small address range is active – the smaller the range, the better
  - **Persistence**: If an address range active (or inactive), it remains so for a while – the longer the duration, the better



80% of IO's on 20% of capacity

Storage Elements

- Symmetrix Tier – a shared storage resource with common technologies (Virtual Pools)
- FAST Policy – manage Symmetrix Tiers to achieve service levels for one or more Storage Groups
- FAST Storage Group – logical grouping of thin devices for common management

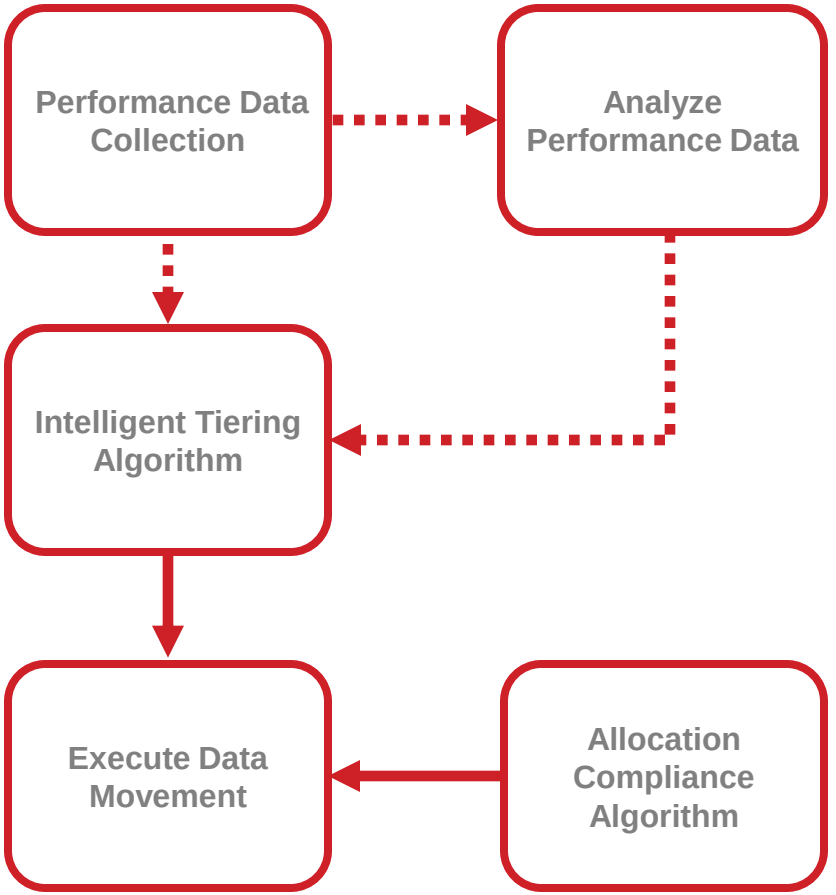


FAST VP Time Windows

- **Performance time window** defines when statistics are collected and decayed
- **Workload Analysis Period**
  - Affects decay rates
  - Affects time to respond to changes
  - Default of 7 days (168 hours)
- **Data movement time window** defines when FAST VP is allowed to move data
  - Windows can be customized, but recommendation for initial implementation is 24x7

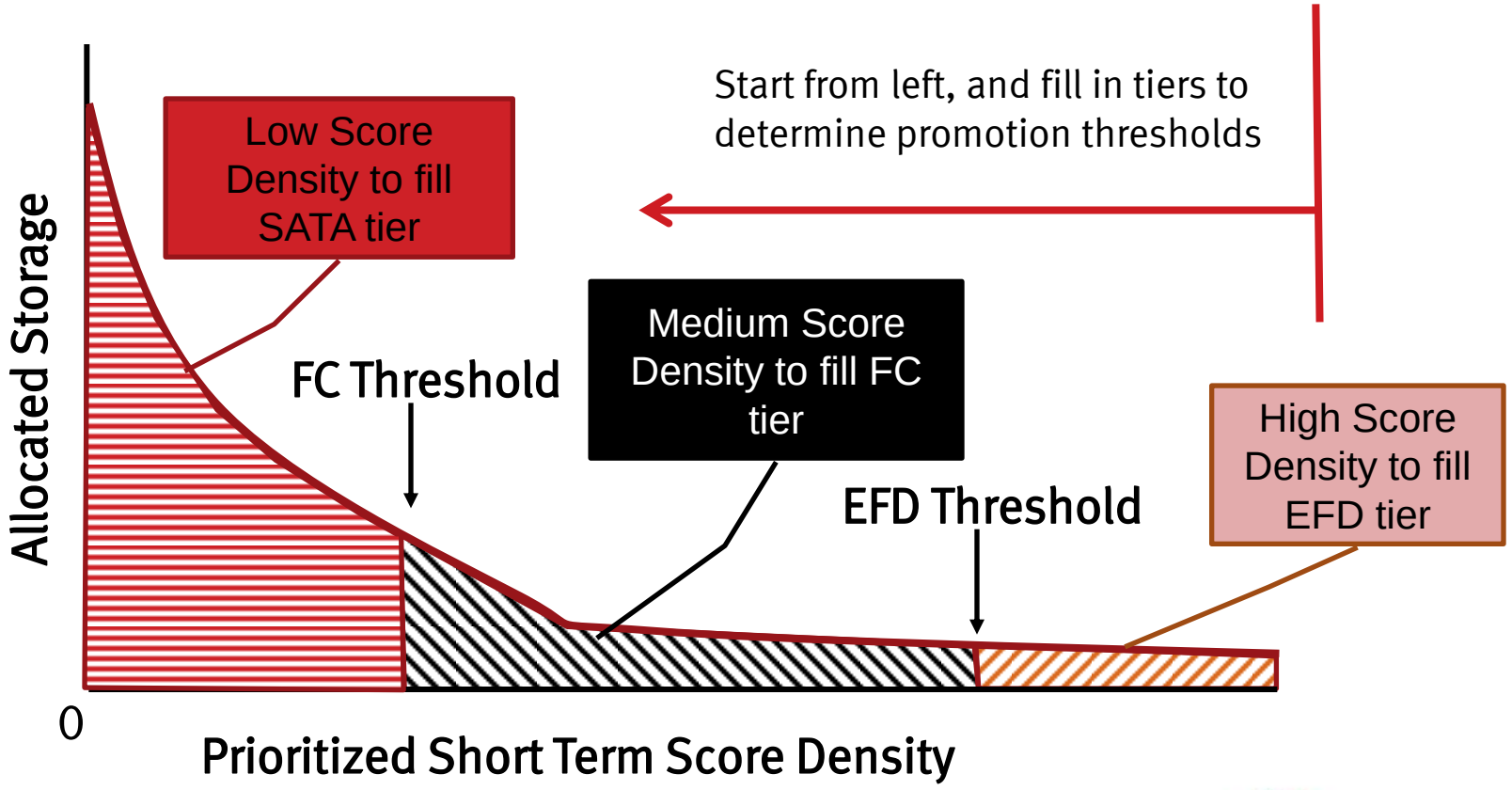
FAST VP Implementation

- Performance data collected by the system
- **Intelligent Tiering** algorithm generates movement requests based on performance data
- **Allocation Compliance** algorithm generates movement requests based on capacity utilization
- Algorithms continuously assess I/O statistics and capacity use, and make decisions for promotion and demotion



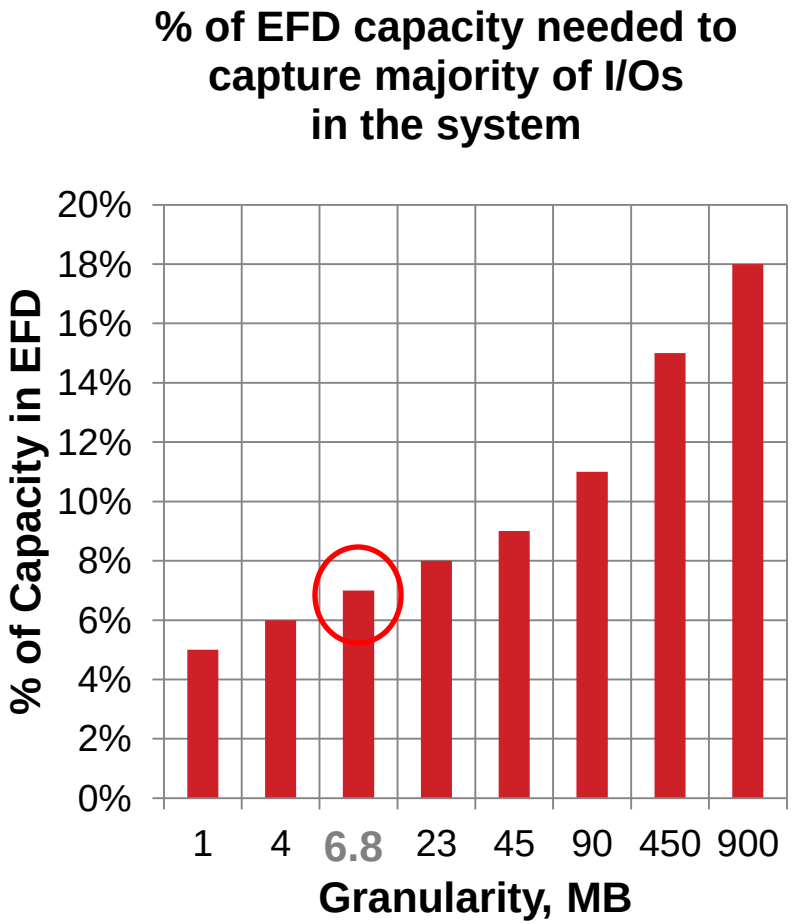
FAST VP – Score Analysis

- Extents Group Sets are grouped in a histogram according to the Score

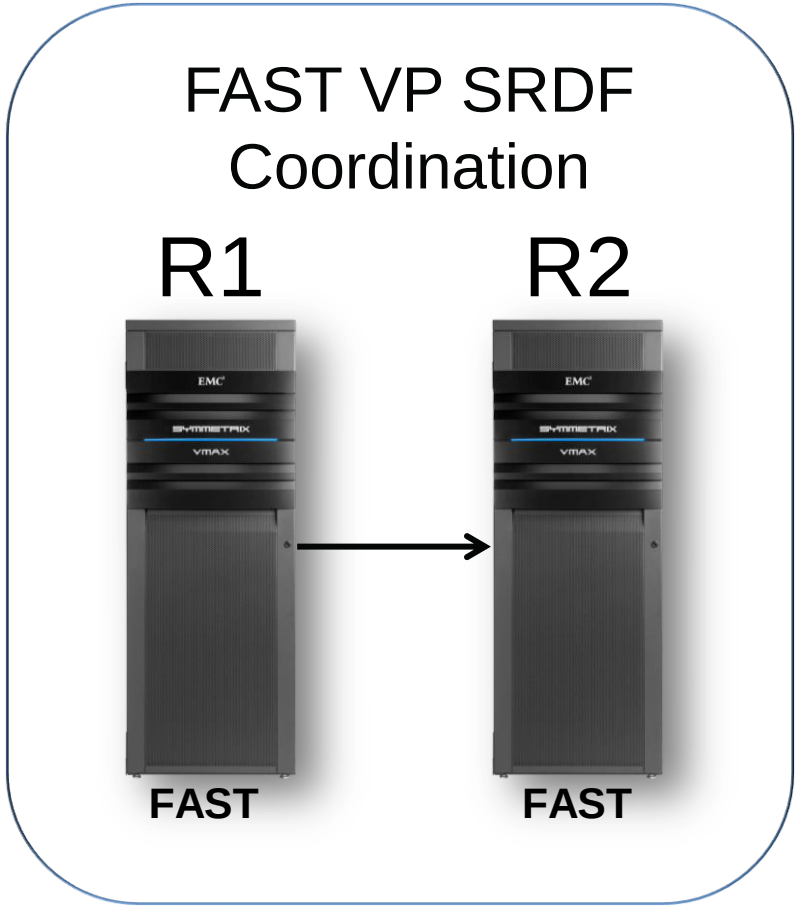


Data Movement Granularity Trade-offs

- Larger granularity
  - Uses EFD ineffectively
- Smaller granularity
  - Uses EFD effectively
  - Requires more system resources to maintain statistics
- There is a sweet spot that maximizes the benefits through better use of EFD and reasonable system resource use



Remote Replication Awareness for FAST VP

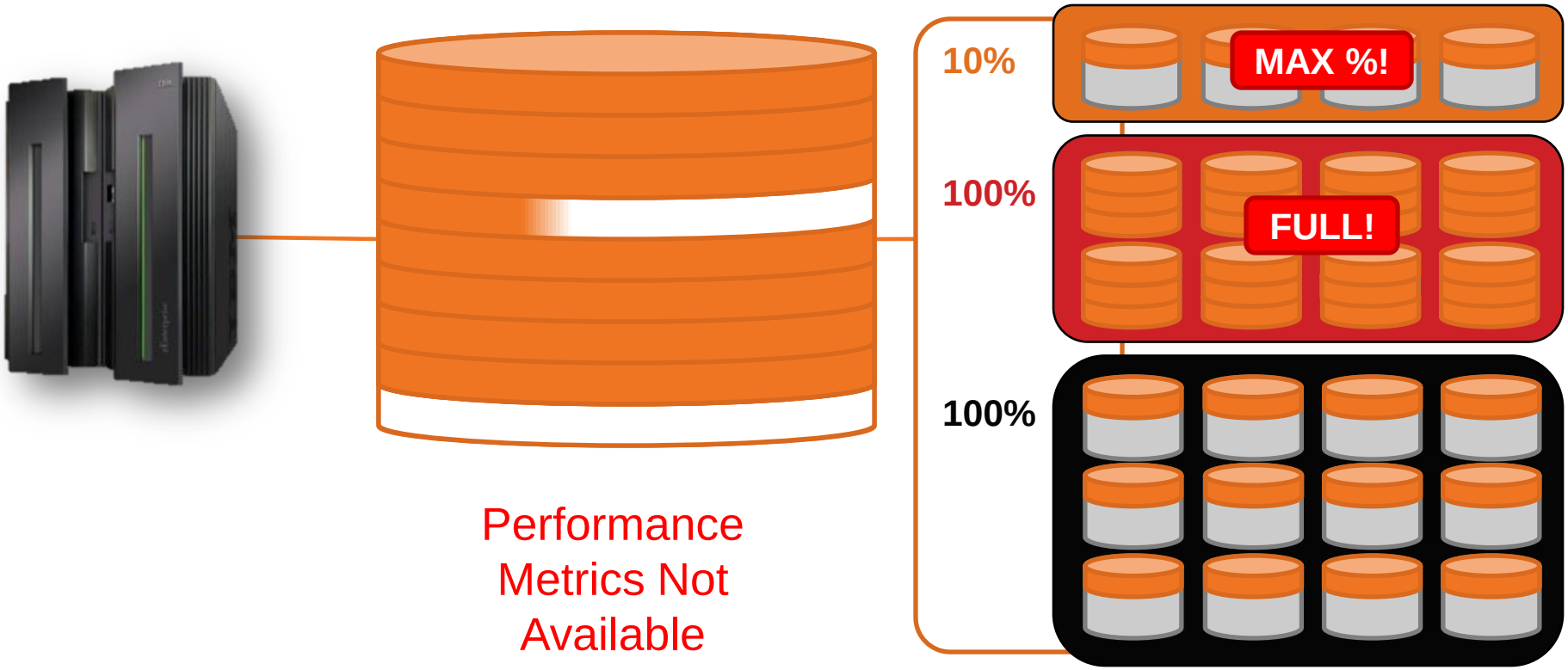


- Align FAST VP performance between R1 and R2 sites
- Performance metrics are shared from the R1 to the R2 site
- In case there is a swap where production is run from the R2 site, the performance on the R2 site will closely match the R1 performance

Policy-based allocation

- Allocations can come from any pool contained within the tiers of the associated FAST VP policy
  - FAST VP will attempt to allocate from most appropriate tier
    - Based on FAST VP performance metrics and policies
  - If pool is full, alternative pool will be chosen for allocation
    - If array capacity is not oversubscribed new writes should not fail
- If performance metrics exist for the region of the thin device, allocation request will attempt to place new data on the appropriate tier
- If no performance metrics exist request will go to the “bound” pool
- If bound pool is full, allocation request will follow the FAST VP policy percentages

Policy-based allocation – Enabled



Summary

- FAST VP is a policy-based system that promotes and demotes data at the sub-volume, and more importantly, *sub-dataset* which makes it responsive to the workload and efficient in its use of control unit resources
- FAST VP introduces active performance management, a revolutionary step forward in storage management
- FAST VP delivers all these benefits without using any host resources

For more Virtual/Tiered storage information:

- Other SHARE sessions
  - 12945: DB2 for z/OS With EMC Storage Tiering: FAST VP – Wed. @ 8AM in Golden Gate 8
  - 13154: EMC Disk Tiering Technology Review – Wed. @ 12:15 in Golden Gate 2
  - 12708: What's New With EMC Symmetrix VMAX and Enginuity? – Wed. @ 4:30PM in Golden Gate 7
  - 12317: Less=More with Thin Provisioning and Linux on System z – Thur. @ 3:00PM in Franciscan D
- EMC.COM Mainframe Page
  - http://www.emc.com/storage/mainframe.htm
- Follow EMC Symmetrix

  @EMCSymmetrix

  www.facebook.com/Symmetrix

#SHAREorg

SHARE
Technology · Connections · Results

Getting Even More (and a lot less) From Storage Virtual Provisioning and Automated Storage Tiering

Tony Negro
EMC Corporation

Tuesday, Februrary 05, 2013 1:30 PM
Session Number 12174

Thank You

SHARE in San Francisco 2013