AVERE

Dynamic Tiering – Taking Advantage of the Best of HDDs and Solid State

AVERE SYSTEMS, INC
averesystems.com

April 12, 2010

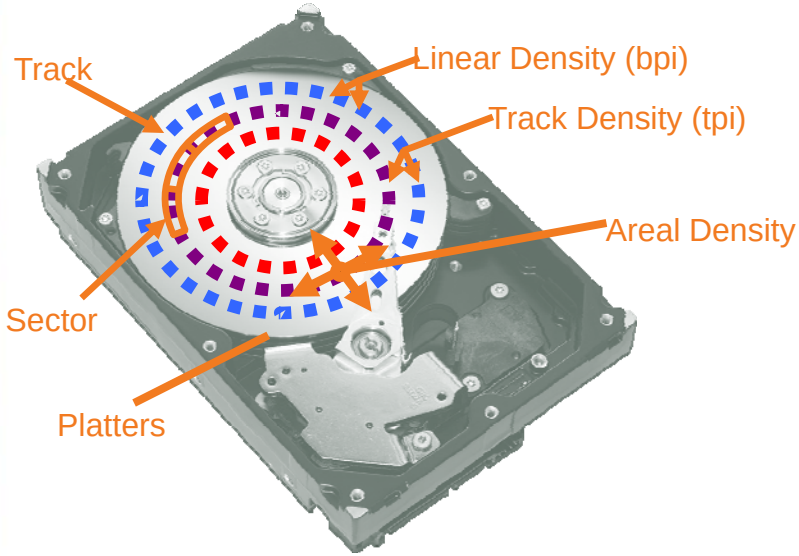
Ron Bianchini, CEO and President

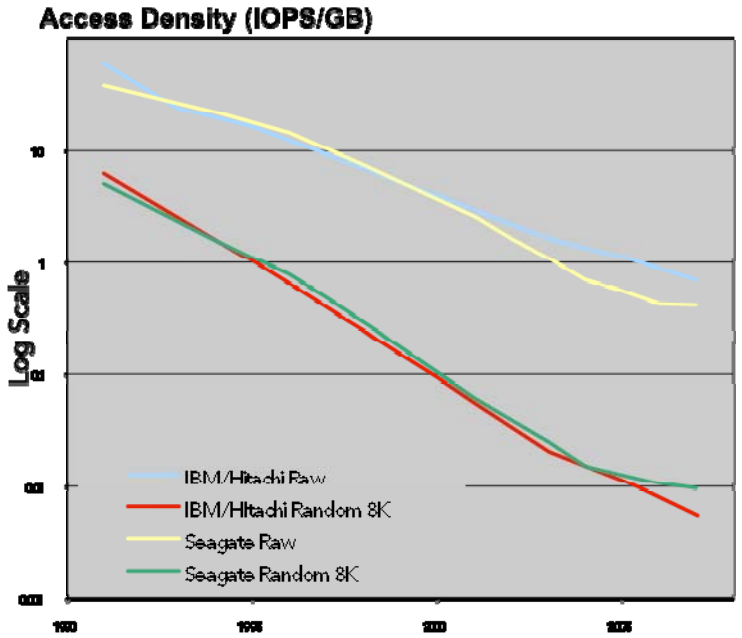
Agenda

- Customer Challenge
- Storage Media Comparison
- Tiered NAS Architecture
- Avere Two Stage Implementation
- Tiered vs. Homogeneous NAS Comparison
- Tiered NAS Roadmap
- Summary

Customer Challenge

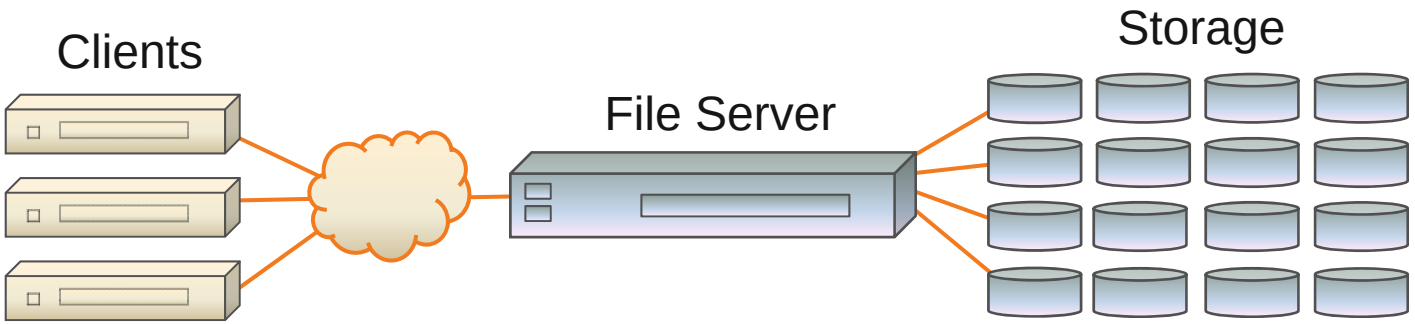
- Hard disk drives (HDDs) are getting bigger not faster
- Many costly 15k RPM drives required to achieve performance
- Challenging due to budget, power, cooling, floor space constraints
- Leverage SSD, but current solutions are expensive and incomplete





Application Example

- Application
  - Working Set: 20% accessed within 20 minutes
- Storage Server
  - 16 attached disk drives
- Tape archive



Application Example (data growth)

	1998	2008 WS %	2008 WS time
Disk Capacity	4.5 GB	450 GB	450 GB
Disk #	16 x	16 x	16 x
Sys. Capacity	72 GB	7.2 TB	7.2 TB
WS %	20% x	20% x	0.35% x
WS Size	14.4 GB	1.44 TB	24.87 GB
Disk Latency	9.50 ms	5.50 ms	5.50 ms
Disk 8K Rate	0.86 MB/s	1.49 MB/s	1.49 MB/s
Sys. 8K Rate	13.80 MB/s	23.83 MB/s	23.83 MB/s
WS Time	17.40 mins	16.78 hrs	17.40 mins

Application Example (no growth!)

	1998	2008 Capacity	2008 Performance Short Stroking!
Disk Capacity	4.5 GB	450 GB	450 GB
Disk #	16 x	1 x	10 x
Sys. Capacity	72 GB	450 GB	4.5 TB
WS %	20% x	3% x	0.3% X
WS Size	14.4 GB	14.4 GB	14.4 GB
Disk Latency	9.50 ms	5.50 ms	5.50 ms
Disk 8K Rate	0.86 MB/s	1.49 MB/s	1.49 MB/s
Sys. 8K Rate	13.80 MB/s	1.49 MB/s	14.90 MB/s
WS Time	17.40 mins	2.7 hrs	16.1 mins

Typical Approaches to Challenge

Type	Limitation
NAS Server	• Expensive over provision & short stroke • Forced to select expensive drive types • Mgmt overhead to scale system
Caching Appliance	• Read only work loads (non-persistent) • One protocol (NFS) limitation typical
SSD Adapter	• Inability to scale separately from server • Proprietary
SSD Array	• Limited/none Tier0 management • High media cost
Switch	• Inability to scale outside switch • Non-transparent

Historical Approach: Evolutionary

- Industry following evolutionary path
  - Denser disk drives, reduced access density
  - Increase data management needed
  - Increase performance via over provisioning & caching
- Start from scratch
  - Start with specification of NAS Server
  - Consider media to satisfy performance requirements
  - Identify allocation algorithms to map data to media

Storage Media Comparison

		Cap	Price
SATA HDD (1x2TB)		2,000G	$150
SAS HDD (1x300GB)		300G	$270
SLC Flash (1x64G SATA)		64G	$700
DRAM (8x4GB DIMMs)		32G	$1280

- Costs shown for media-only, unit quantity, 4/10

Storage Media Comparison

clear winner for archive

		Cap	$/GB
SATA HDD (1x2TB)		2,000G	$0.08
SAS HDD (1x300GB)		300G	$0.90
SLC Flash (1x64G SATA)		64G	$11.00
DRAM (8x4GB DIMMs)		32G	$40.00

Storage Media Comparison

clear winner for archive

	Cap	$/GB	Reads (8KB/sec)	$/Read	Writes (8KB/sec)	$/Write
SATA HDD (1x2TB)	2,000G	$0.08	130	$1.15	130	$1.15
SAS HDD (1x300GB)	300G	$0.90	400	$0.70	360	$0.75
SLC Flash (1x64G SATA)	64G	$11.00	24,500	$0.03	1,000	$0.70
DRAM (8x4GB DIMMs)	32G	$40.00	325,000	$0.004	325,000	$0.004

good for random reads

winner for small R/W

- Can we leverage multiple tiers?

Storage Media Comparison

clear winner for archive

	Cap	$/GB	Reads (8KB/sec)	$/Read	Writes (8KB/sec)	$/Write
SATA HDD (1x2TB)	2,000G	$0.08	130	$1.15	130	$1.15
SAS HDD (1x300GB)	300G	$0.90	400	$0.70	360	$0.75
SLC Flash (1x64G SATA)	64G	$11.00	24,500	$0.03	1,000	$0.70
DRAM (8x4GB DIMMs)	32G	$40.00	325,000	$0.004	325,000	$0.004

can we hide erase time?

winner for small R/W

Storage Media Comparison

clear winner for archive

can we hide access latency?

	Cap	$/GB	Reads (8KB/sec)	$/Read	Writes (8KB/sec)	$/Write
SATA HDD (1x2TB)	2,000G	$0.08	130	$1.15	130	$1.15
SAS HDD (1x300GB)	300G	$0.90	400	$0.70	360	$0.75
SLC 2M write buf (1x64G SATA)	64G	$11.00	24,500	$0.03	4,200	$0.16
DRAM (8x4GB DIMMs)	32G	$40.00	325,000	$0.004	325,000	$0.004

good for random access

winner for small R/W

Storage Media Comparison

clear winner for archive

good for big sequential

	Cap	$/GB	Reads (8KB/sec)	$/Read	Writes (8KB/sec)	$/Write
SATA HDD (1x2TB)	2,000G	$0.08	130	$1.15	130	$1.15
SAS 2M buf (1x300GB)	300G	$0.90	17,500	$0.02	17,500	$0.02
SLC 2M write buf (1x64G SATA)	64G	$11.00	24,500	$0.03	4,200	$0.16
DRAM (8x4GB DIMMs)	32G	$40.00	325,000	$0.004	325,000	$0.004

good for random access

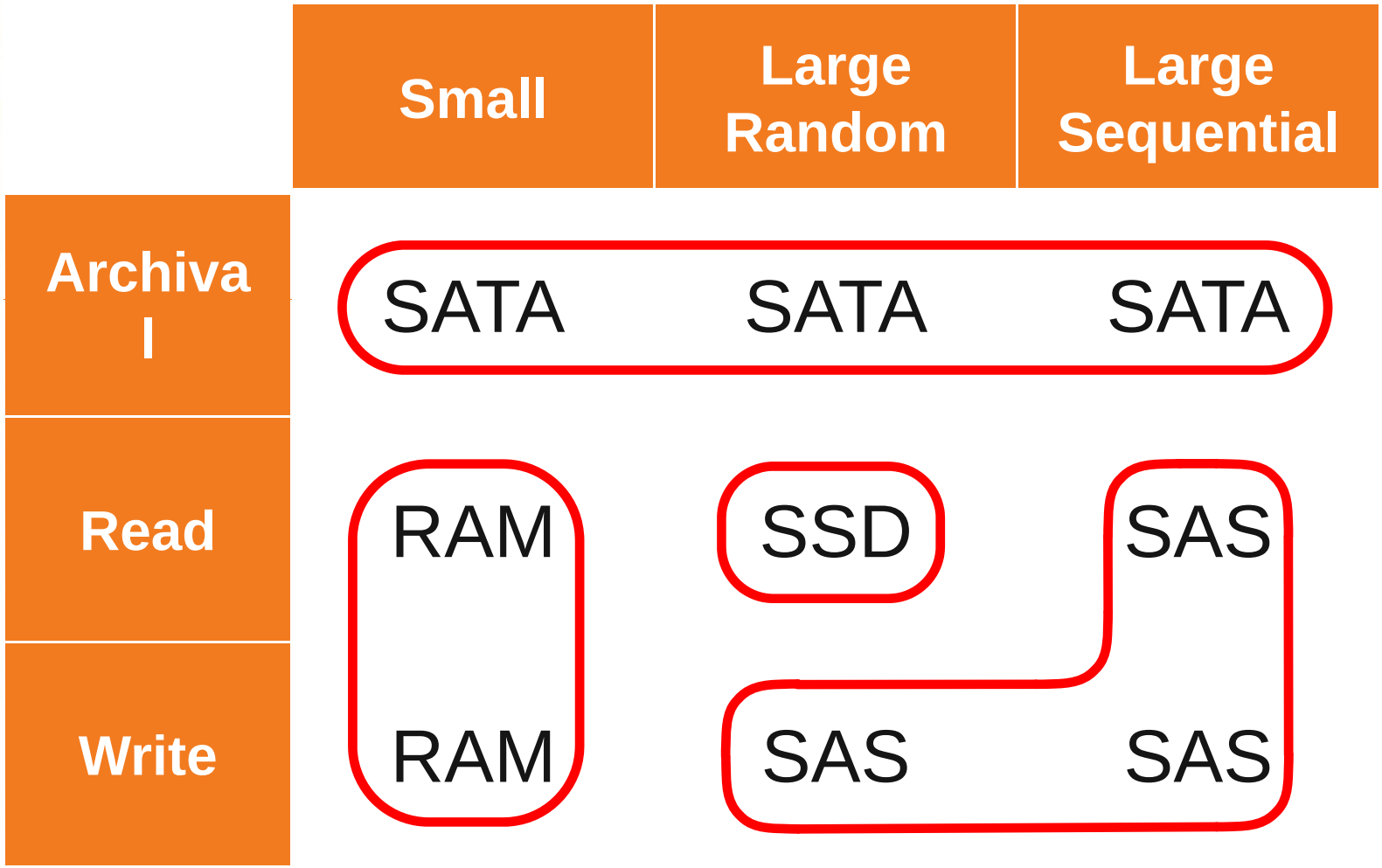
winner for small R/W

- No single technology is best for all workloads

Building a New Solution - Which Media?

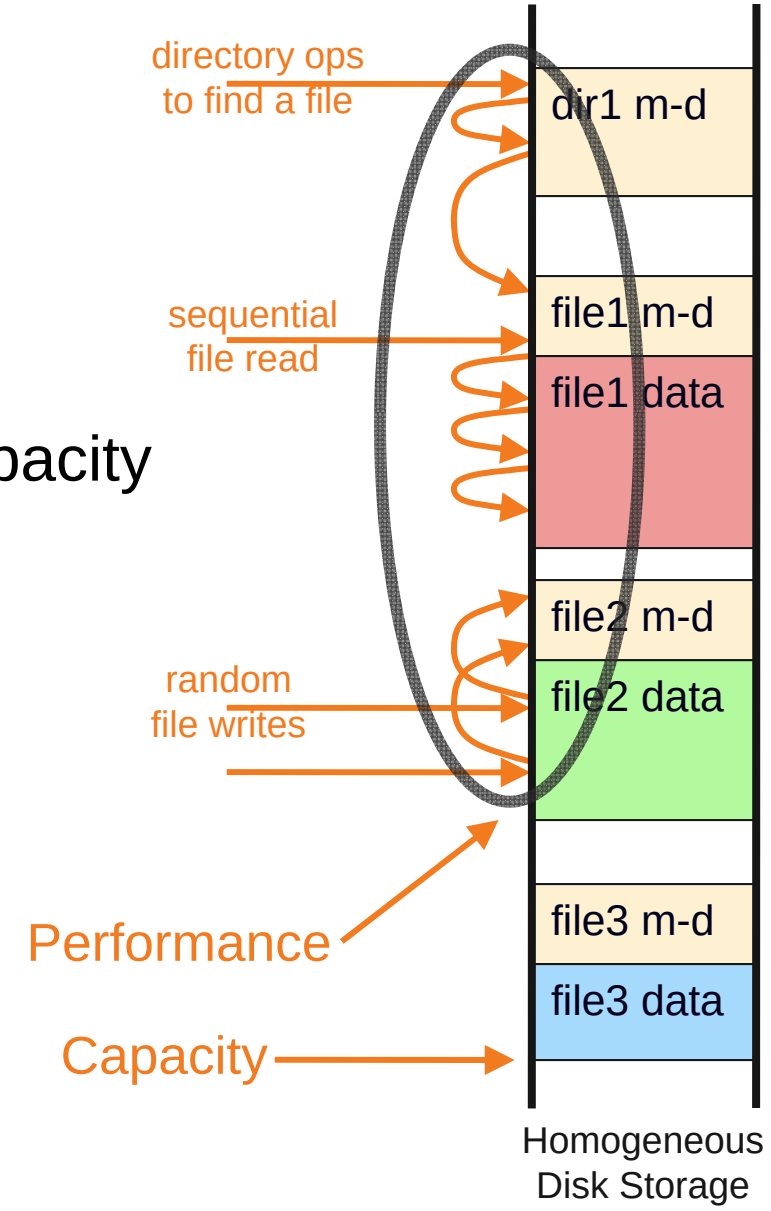
- No single technology is best for all workloads
- A tiered solution leverages best of multiple media
  - SATA for long term archive storage
  - RAM/NVRAM for small file random access
  - Flash is best for random read
  - SAS/RAM is best for sequential read
    - RAM hides access latency for sequential reads
  - SAS/NVRAM is best for random and sequential write
    - NVRAM hides access latency for sequential writes
    - Log-based file systems minimize access latency for random writes

Storage Media Comparison Summary

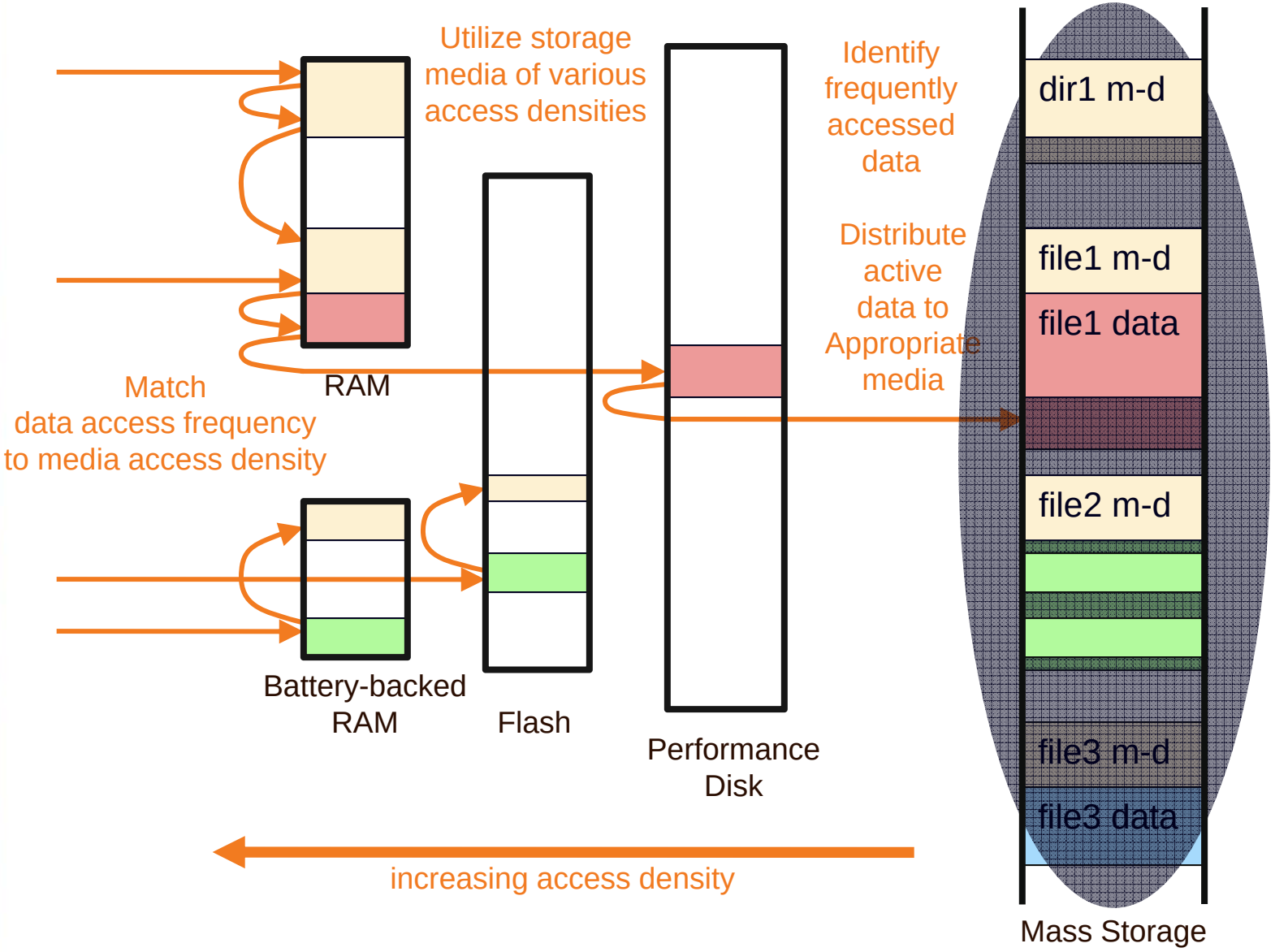


Current NAS Systems

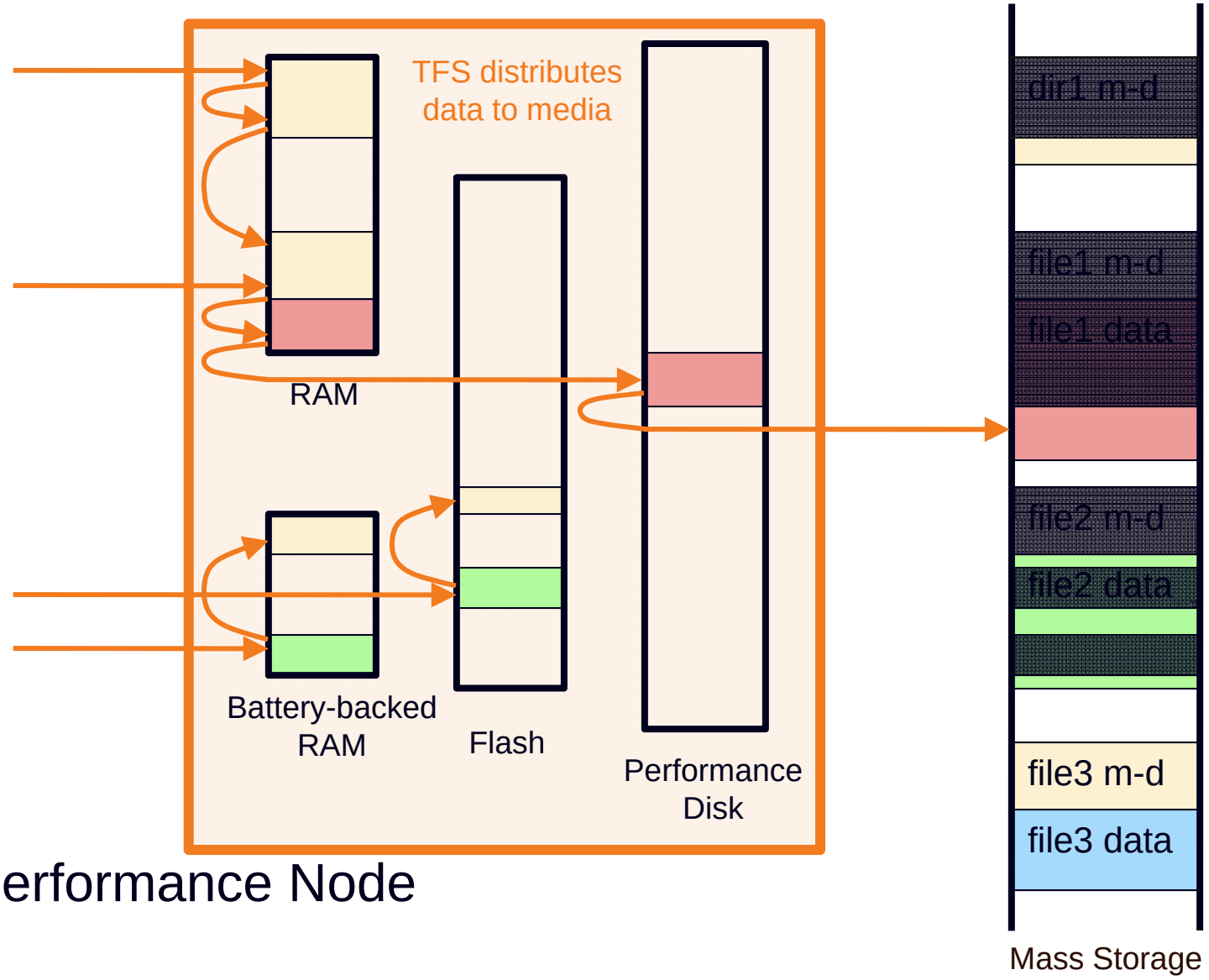
- All NAS operations on homogeneous storage
  - Data/meta-data
- Drive array defines capacity and performance
- Capacity (platters)
  - add drives
- Performance (arms)
  - choose FC vs SATA
  - add drives
    - short stroke



Tiered NAS Architecture



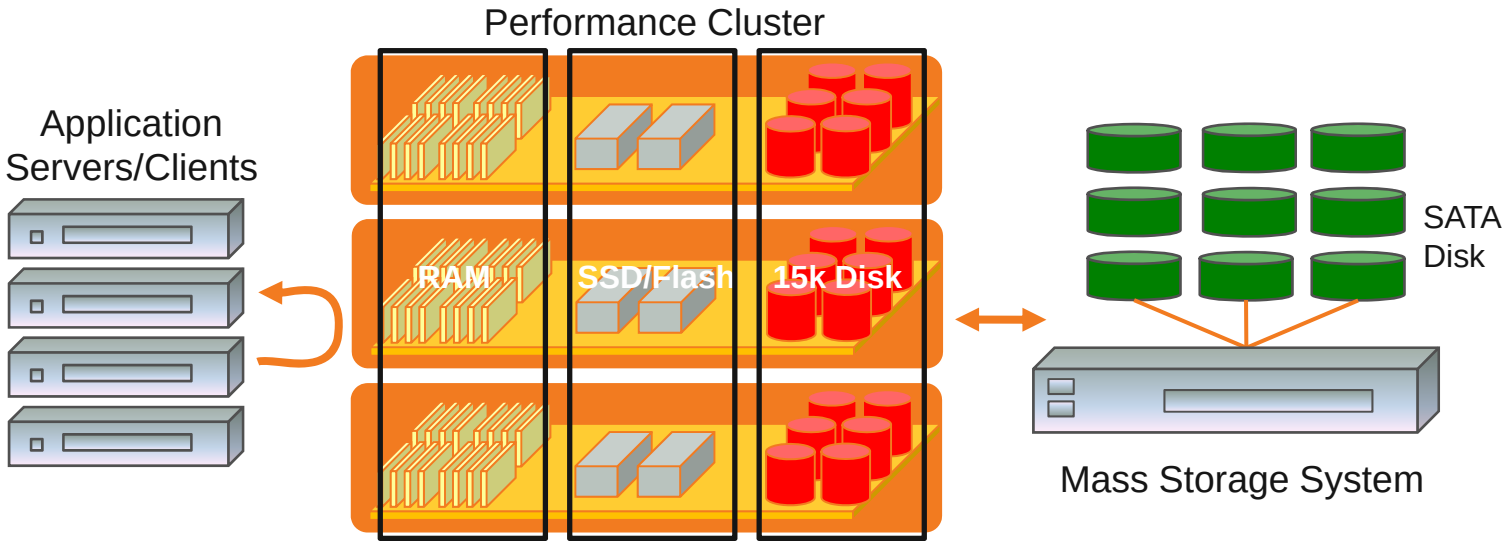
Two Stage Implementation



Performance Node

Two Stage Tiered NAS Architecture

- NAS function implemented in two stages
  - Stage 1: Data Delivery
    - Contains three faster storage media: RAM/NVRAM, SSD, SAS disk
  - Stage 2: Mass Storage
    - Intended for densest, low cost storage media: SATA



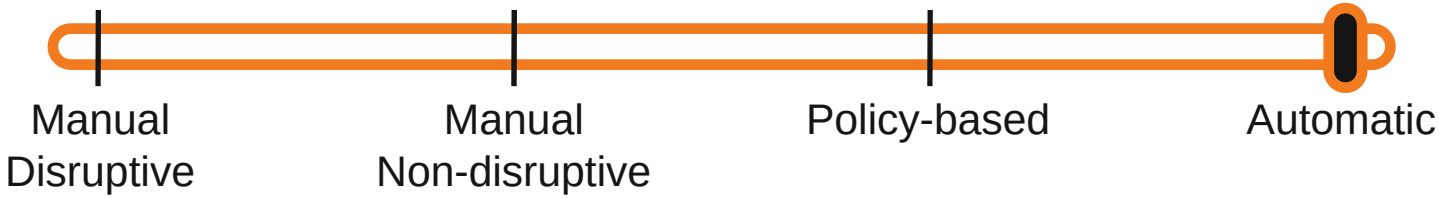
- Need More Capacity? Add SATA disks to Mass
- Need More Performance? Add performance nodes to cluster
  - Media tiered globally among FXT nodes
  - Single names space across all FXT nodes

True Dynamic Tiering

- What? Finest level of granularity

- When? Data is tiered on-the-fly

- How? Automatic movement between tiers



  - Automatic by frequency, access pattern and size

Tiered NAS Comparison

- Spec SFS ’08 Industry Benchmark
  - http://www.spec.org/sfs2008/results/

The Old Way – Homogeneous NAS

- Estimate storage and performance
- Choose server family based on performance
- Compromise on storage technology for both
- If wrong, iterate

The New Way – Tiered NAS

- Buy highest density storage (Low-NAS)
- Buy Performance nodes to achieve performance
- If wrong, linearly scale either independently

- Tiered NAS architecture optimizes technologies
  - Minimizes cost, real estate, power
  - Maximizes density, performance

Industry Benchmark: Spec SFS '08

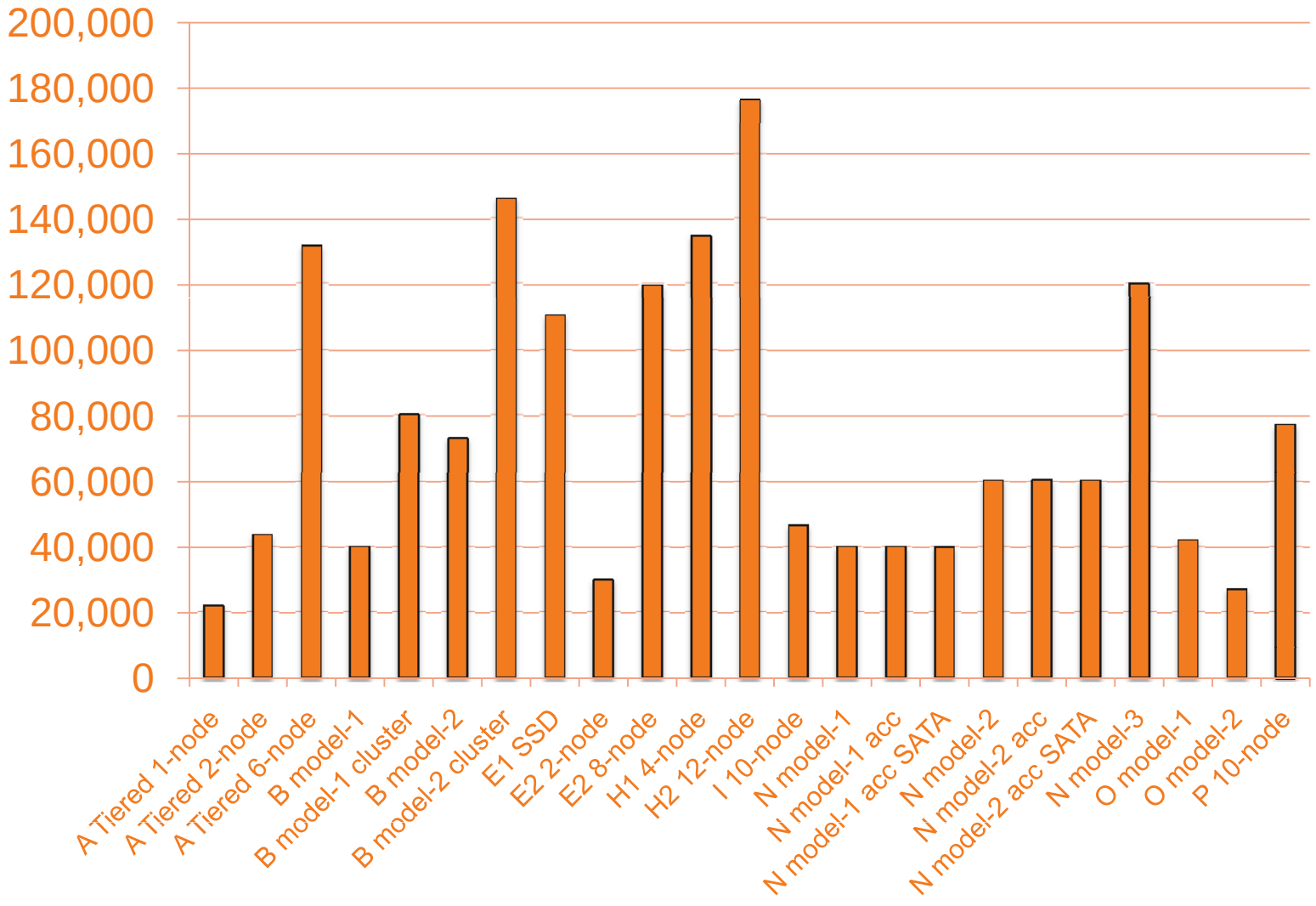
Posting	Op Rate	Latency	#FileSys	# Disks
A Tiered 1-node	22,025	1.30	1	14
A Tiered 2-node	43,796	1.33	1	26
A Tiered 6-node	131,591	1.38	1	79
A2 model-1	8,053	1.37	6	49
A2 model-2	18,511	2.63	16	65
A2 model-3	9,189	2.18	32	65
A2 model-4	18,784	2.67	32	65
B model-1	40,137	3.38	1	74
B model-1 cluster	80,279	3.42	2	148
B model-2	72,921	3.39	1	146
B model-2 cluster	146,076	3.34	2	292
E1 SSD	110,621	2.32	8	100
E2 2-node	29,921	1.96	1	148
E2 8-node	119,550	2.07	1	592
H1 4-node	134,689	2.53	48	584
H2 12-node	176,728	1.67	6	960
I 10-node	46,635	1.91	1	120
N model-1	40,109	2.59	2	224
N model-1 acc	40,107	1.68	2	112
N model-1 acc SATA	40,011	2.75	4	112
N model-2	60,409	2.18	4	224
N model-2 acc	60,507	1.58	2	56
N model-2 acc SATA	60,389	2.18	8	96
N model-3	120,011	1.95	2	324
O model-1	42,111	1.74	32	224
O model-2	27,078	1.99	16	112
P 10-node	77,137	2.29	1	190
SGI	10,305	3.86	1	242

Industry Benchmark: Spec SFS '08 (20K+)

Posting	Op Rate	Latency	#FileSys	# Disks
A Tiered 1-node	22,025	1.30	1	14
A Tiered 2-node	43,796	1.33	1	26
A Tiered 6-node	131,591	1.38	1	79
B model-1	40,137	3.38	1	74
B model-1 cluster	80,279	3.42	2	148
B model-2	72,921	3.39	1	146
B model-2 cluster	146,076	3.34	2	292
E1 SSD	110,621	2.32	8	100
E2 2-node	29,921	1.96	1	148
E2 8-node	119,550	2.07	1	592
H1 4-node	134,689	2.53	48	584
H2 12-node	176,728	1.67	6	960
I 10-node	46,635	1.91	1	120
N model-1	40,109	2.59	2	224
N model-1 acc	40,107	1.68	2	112
N model-1 acc SATA	40,011	2.75	4	112
N model-2	60,409	2.18	4	224
N model-2 acc	60,507	1.58	2	56
N model-2 acc SATA	60,389	2.18	8	96
N model-3	120,011	1.95	2	324
O model-1	42,111	1.74	32	224
O model-2	27,078	1.99	16	112
P 10-node	77,137	2.29	1	190

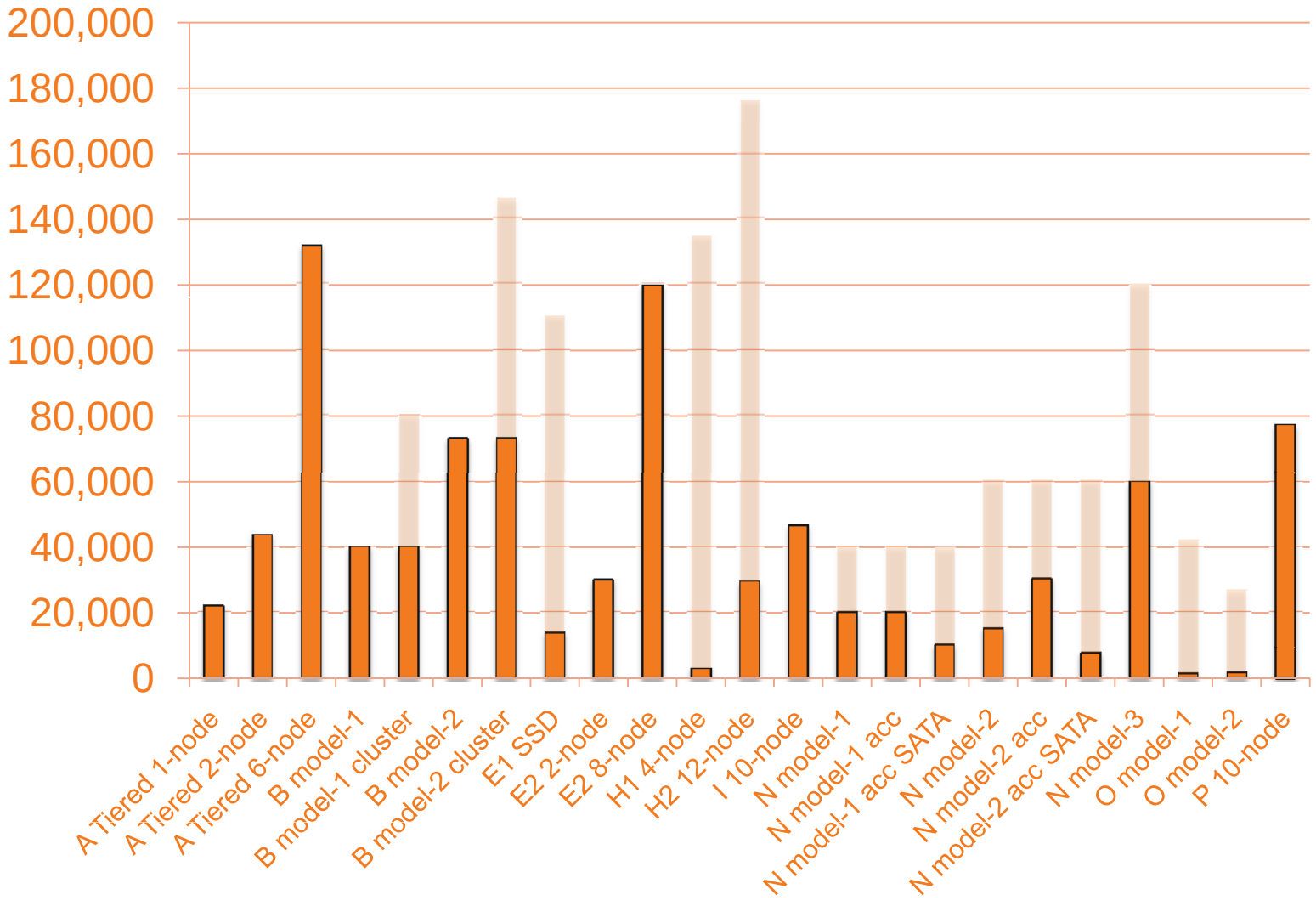
Mass inclusive

Spec SFS '08: Industry Benchmark

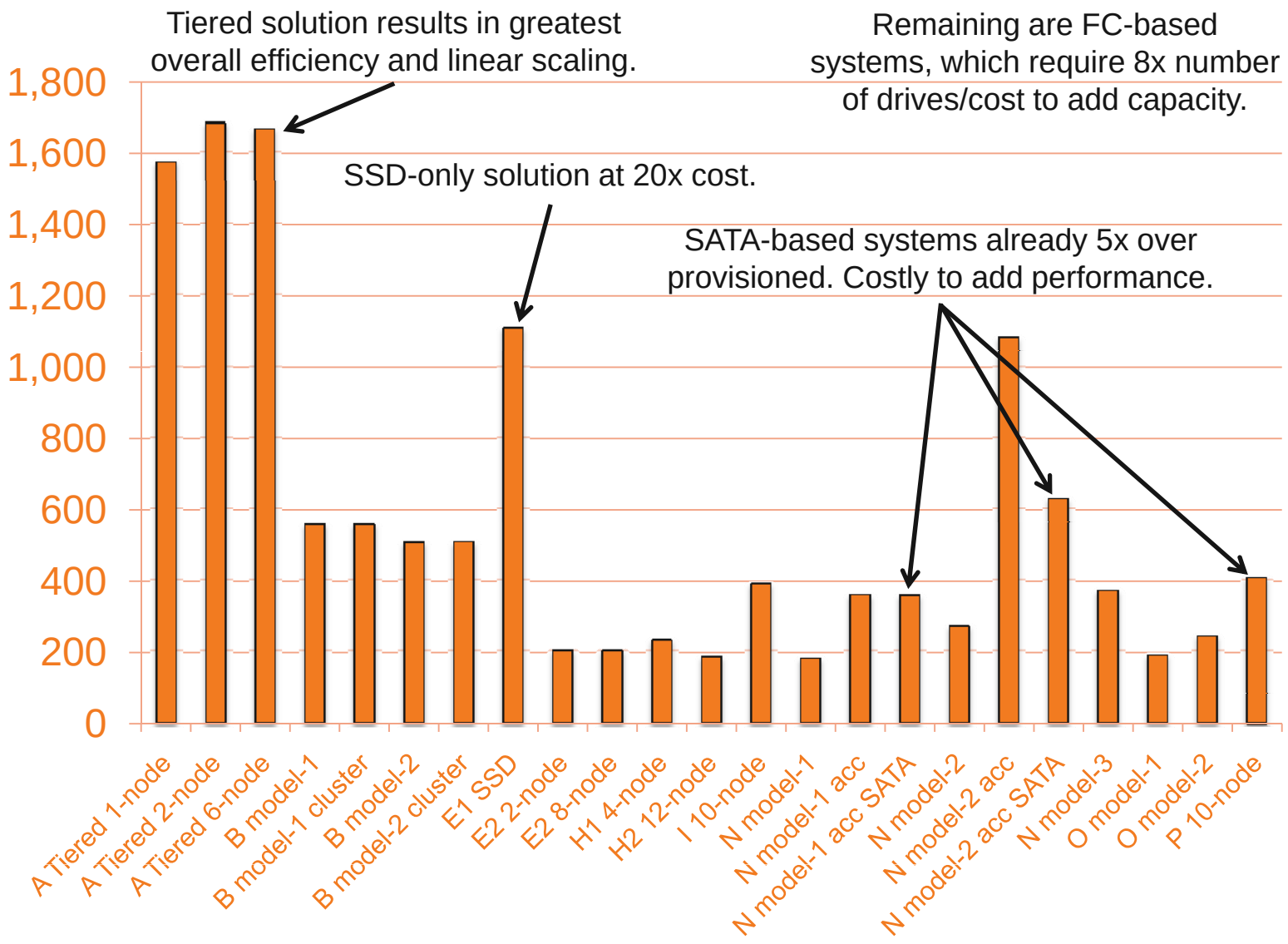


SNIA
SNW
COMPUTERWORLD
April 12-15, 2010
Rosen Shingle Creek Resort
Orlando, Florida

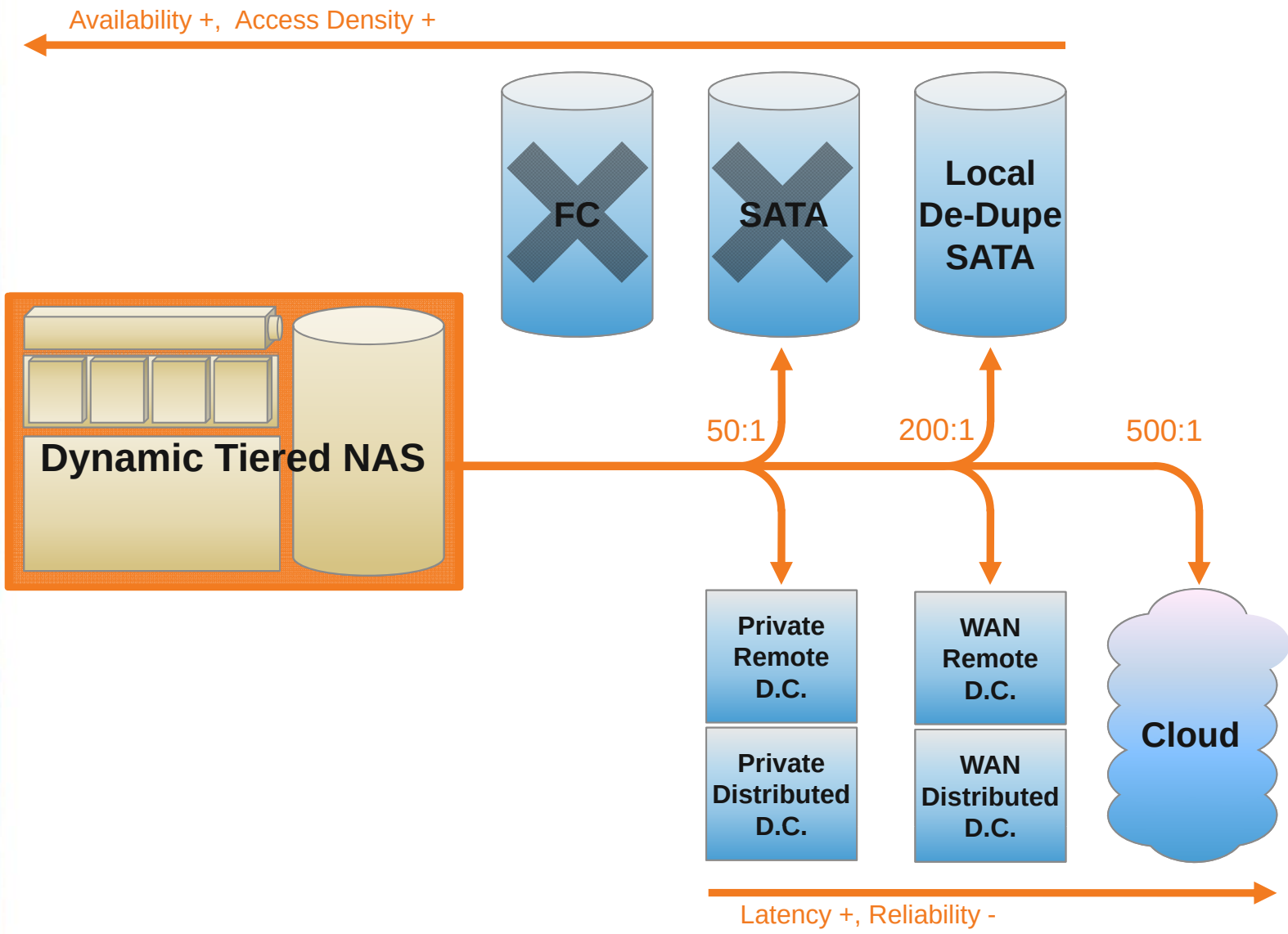
Spec SFS '08: Ops / File Sys



Spec SFS '08: Ops / Disk



Tiered NAS Technology Roadmap



Summary

- No single storage media is best for all applications
- Tiered NAS Architecture
  - Leverage all media types
  - Most efficient
  - Optimize cost/performance
  - Support wide-range of application workloads
- Two Stage Implementation
  - Simple to deploy and manage
  - True separation of performance and capacity scaling
    - Need more capacity? Add SATA disks
    - Need more performance? Add FXT nodes

AVERE

Thank you!

AVERE SYSTEMS, INC
averesystems.com

April 12, 2010

Ron Bianchini, CEO and President