Enterprise Storage

Leah Schoeb, Member of SNIA Technical Council

SNIA Emerald™ Training

SNIA Emerald Power Efficiency Measurement Specification, for use in EPA ENERGY STAR®

July 14-17, 2014

SNIA™ Green Storage Initiative

SNIA Emerald™

What is an Enterprise Storage System?

SNIA Green Storage Initiative

- Storage is more than a device
  - In general, we think of a box
  - We look at the device attributes
    - Capacity
    - Performance (speeds and feeds)
  - It's really about storing and retrieving information with integrity, availability, data protection, cost
  - Need a strategy and efficiency to address how we deal with information

60% — **Storage**

% of IT Hardware Spend

40% — **Servers**

Source: IBM

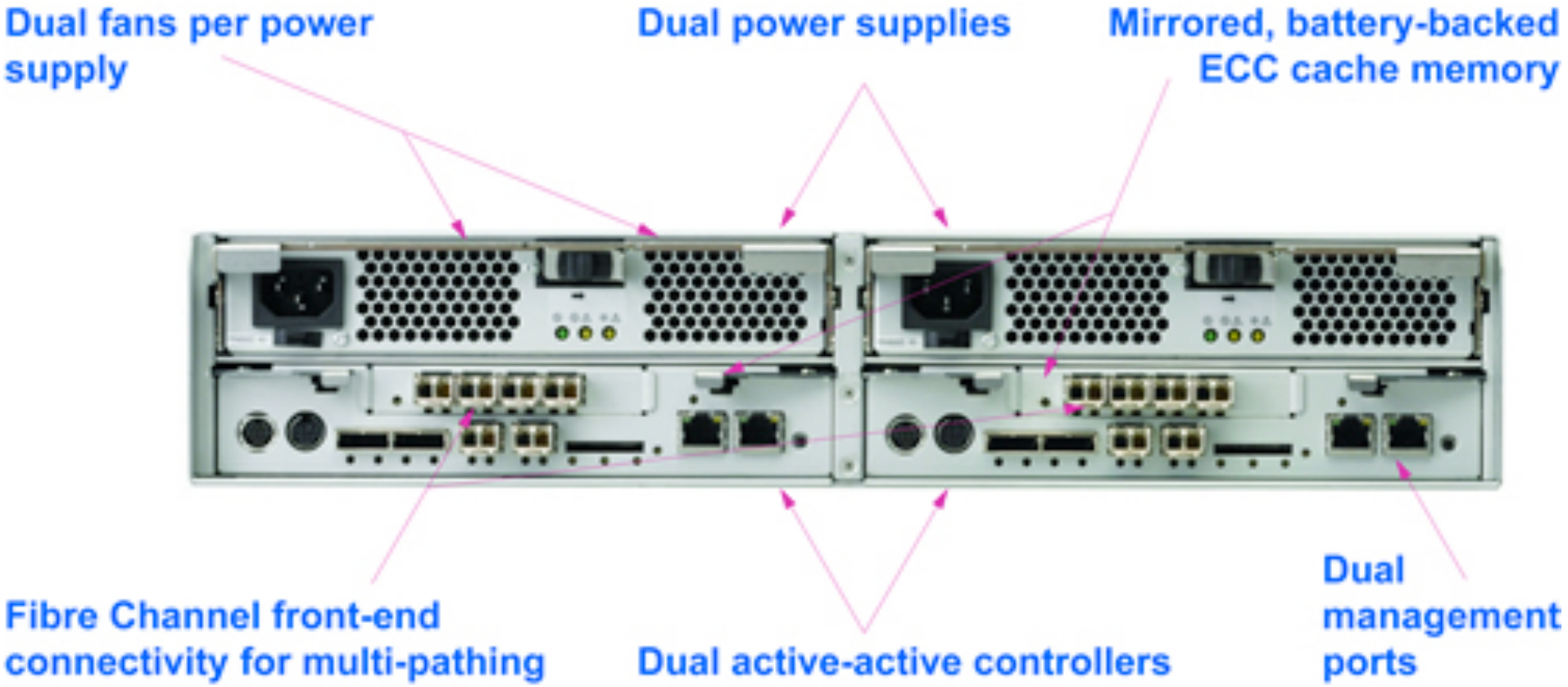
Enterprise Storage System Basics

- Storage System Basic Types
  - Just a Bunch Of Disks (JBOD)
  - Block Storage System
  - File/NAS Storage System
  - Solid State Storage Systems
- Architectural Designs
  - Scale up
  - Scale out
- Storage Technology
  - Data Services
  - Replication

Just a Bunch Of Disks (JBOD)

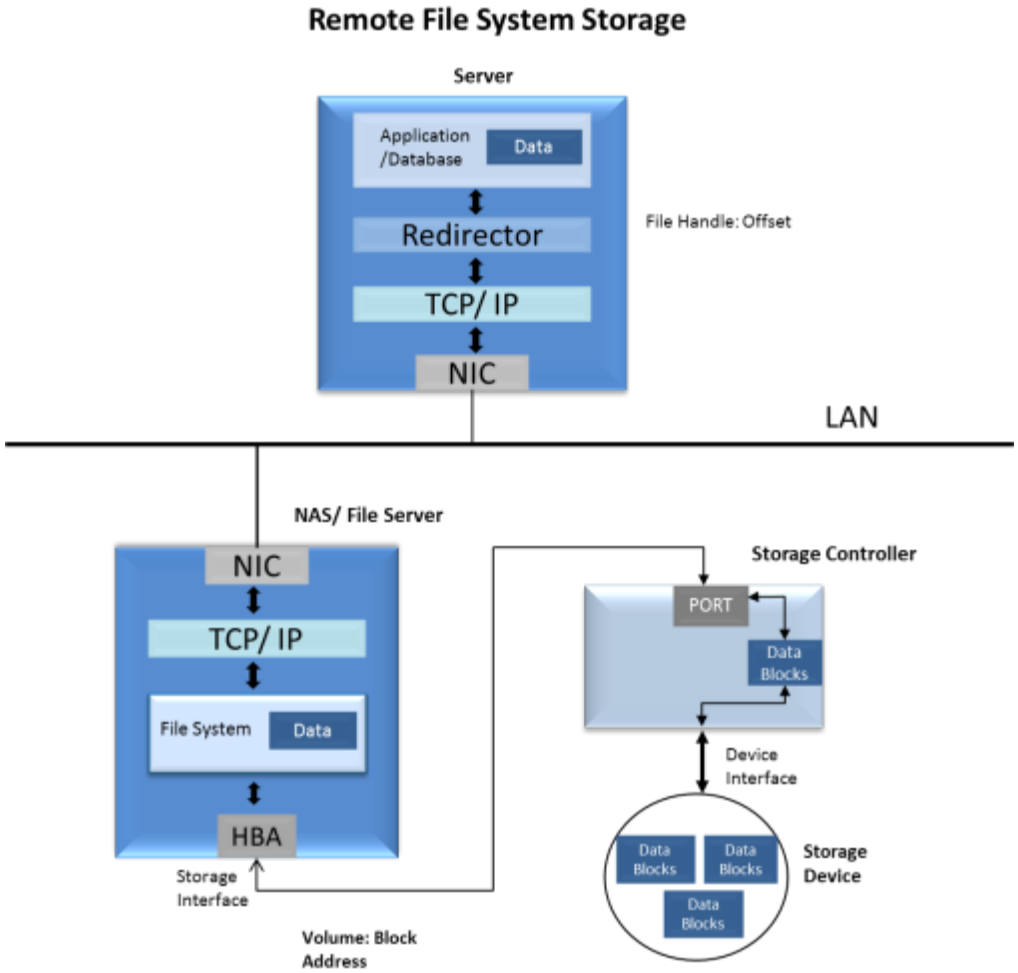
- Multiple disk drives
- No RAID functionality
- Can be used as
  - Individual disk drives
  - One logical volume (spanning)
- No redundancy
- No performance enhancements
- Minimal to no data protection

Block Storage controller System

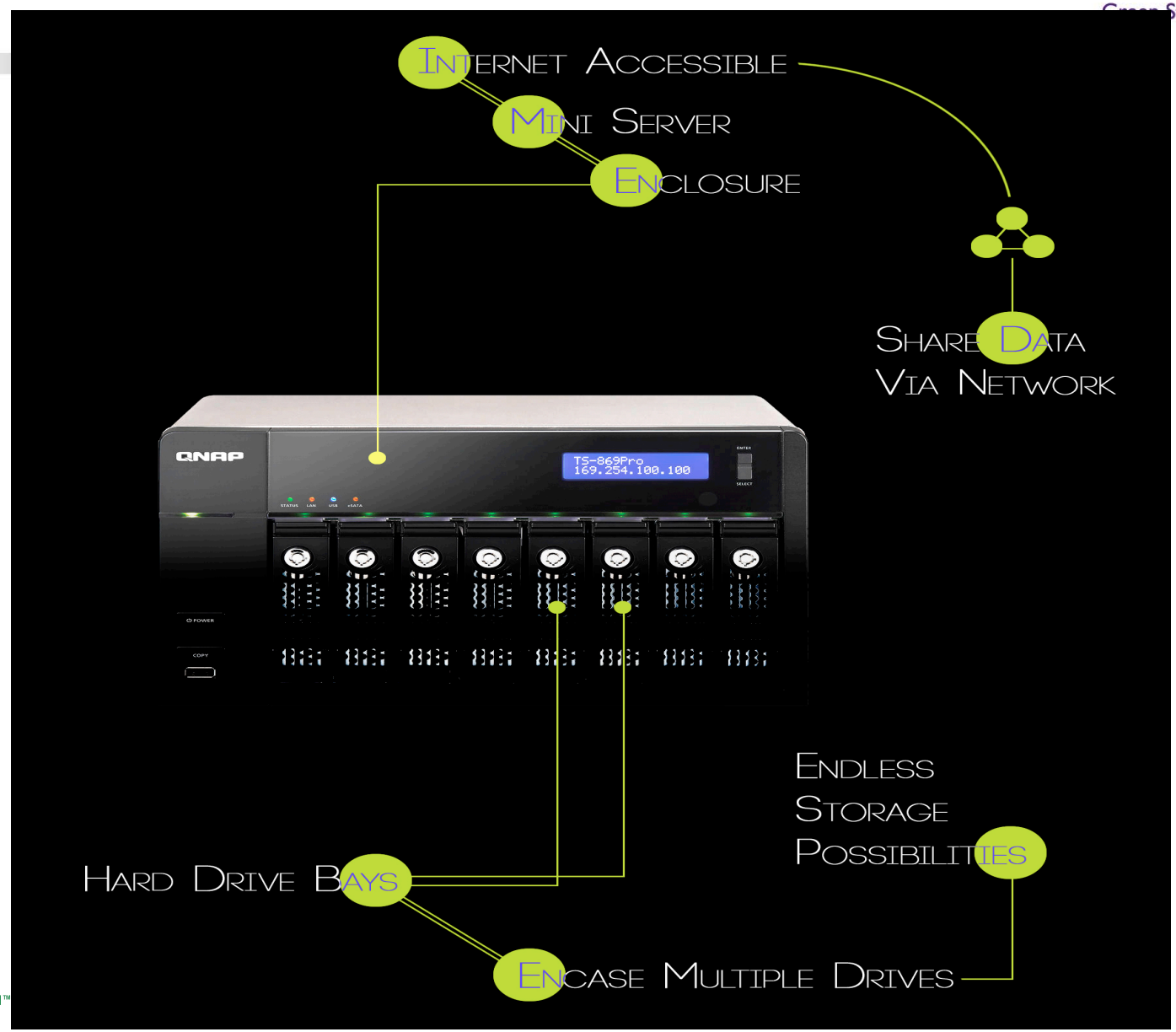


NAS

- Application writes file data to mounted file system
- File redirected over network
- NAS / File server takes file and converts to block I/O
- Data written to device as data block

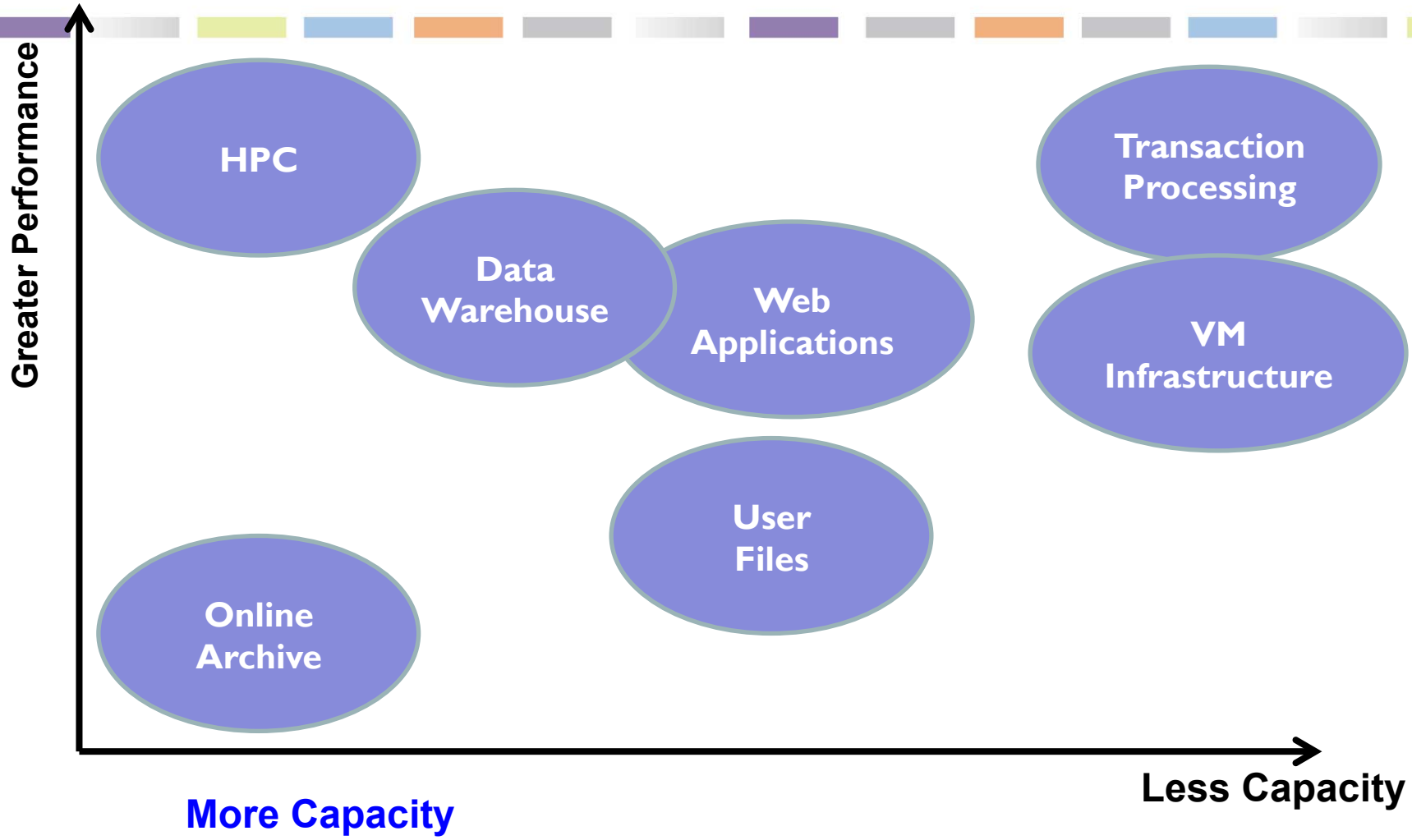


NAS Storage System



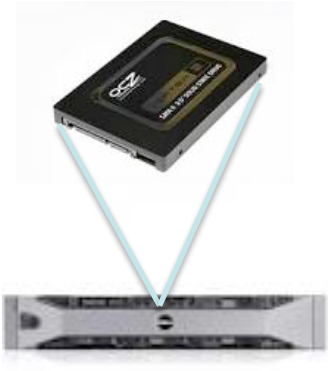
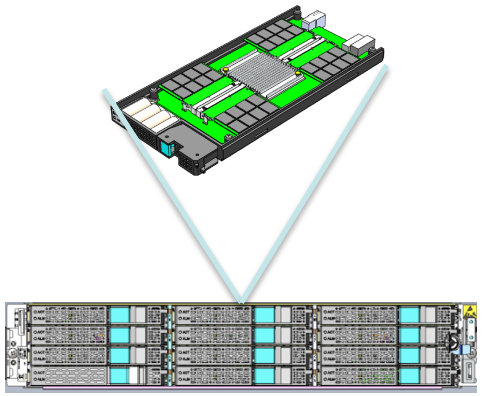
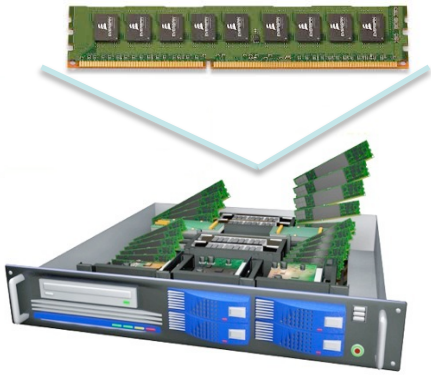
SOLID STATE ARRAYS

Application Landscape for Solid State Storage



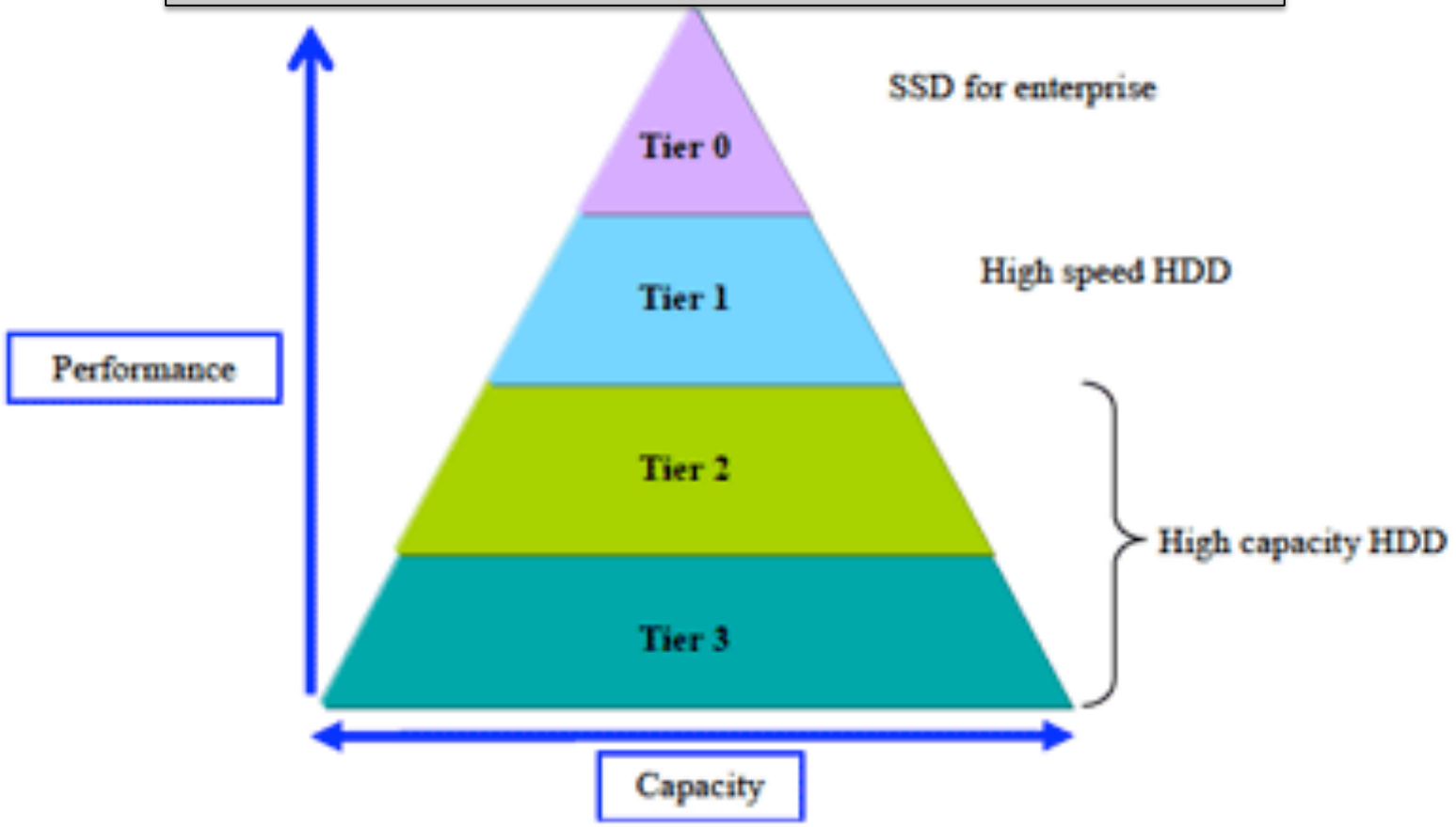
Solid State Storage (Flash) Arrays

SNIA Green Storage Initiative

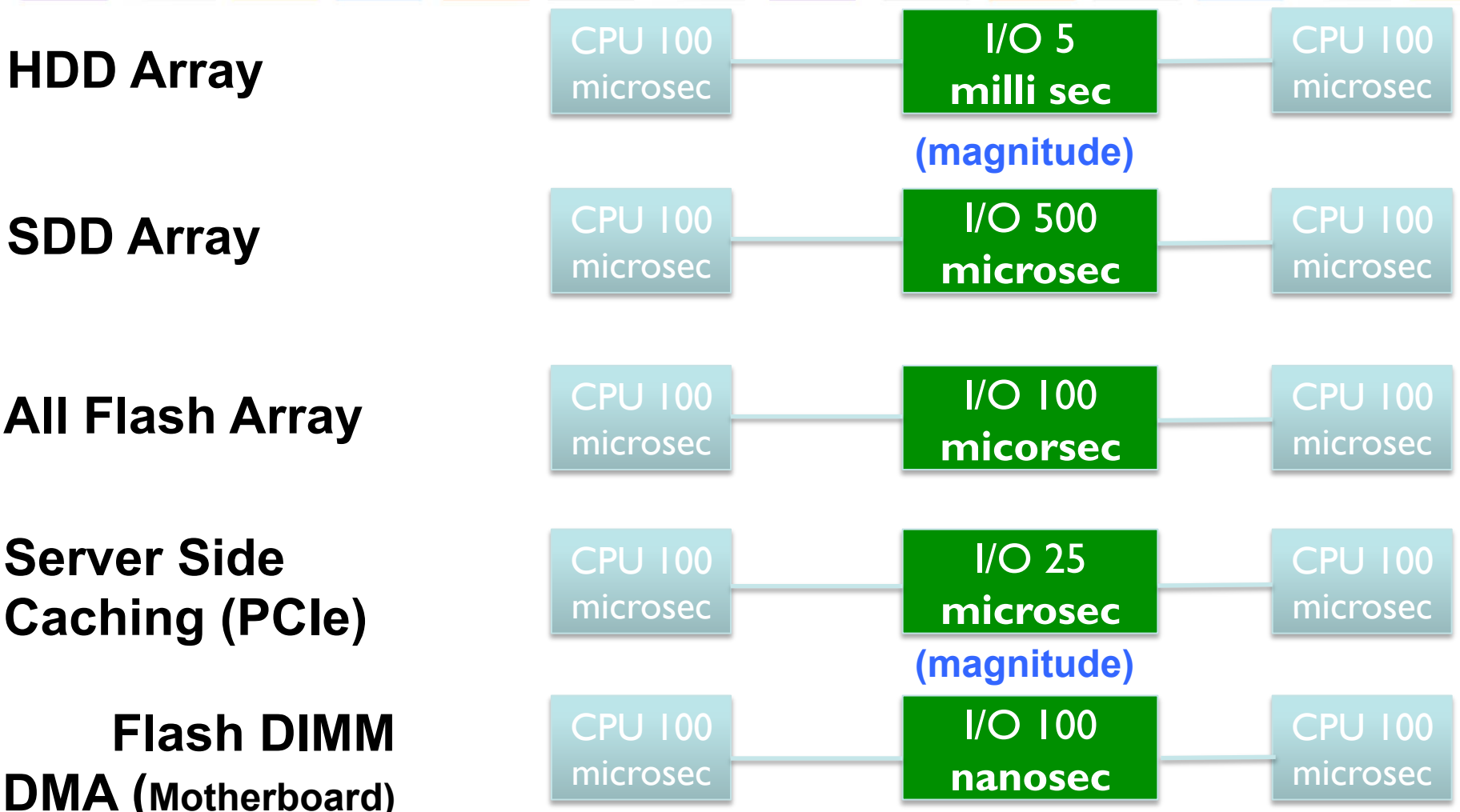
SSD	SSD	Flash Module	Flash DIMM (or SSD)
Traditional Storage Systems - Disk replace - Hybrid - Dynamic Tiering	Re-Architect Controller - Disk Replace - Dynamic Tiering	Flash Optimized Controller - Disk Replace - Dynamic Tiering	Architected for Flash - Virtual Tiering
			
2X	5X	5X-10x	10X

Tiered Storage

Solid State Technology can work well in Tiered Storage Designs

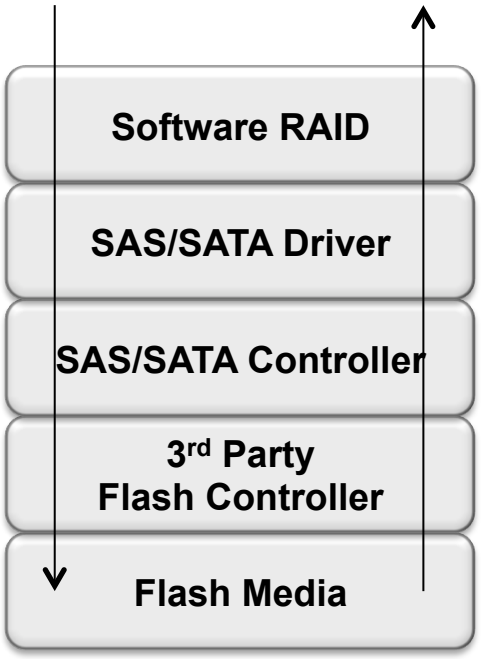


I/O Request Performance



Solid State Storage I/O Stacks

- New storage system architectures for all solid state
- Internal design to handle memory speed
  - PCIe Internally
  - DMA to handle DRAM backup
- Manage technology characteristics
  - writes, cell sparing
  - Done with custom Flash controllers
  - Wear Leveling



Modern All Flash Arrays vs. HDD Arrays

Modern Flash Arrays

- Wear Leveling
- Garbage collection
- Metadata Management
- Self-healing techniques
- Inline Data deduplication
- Inline Compression
- SSD = 3-5 watts per device
- Heat = 10 – 20 BTU/h per device

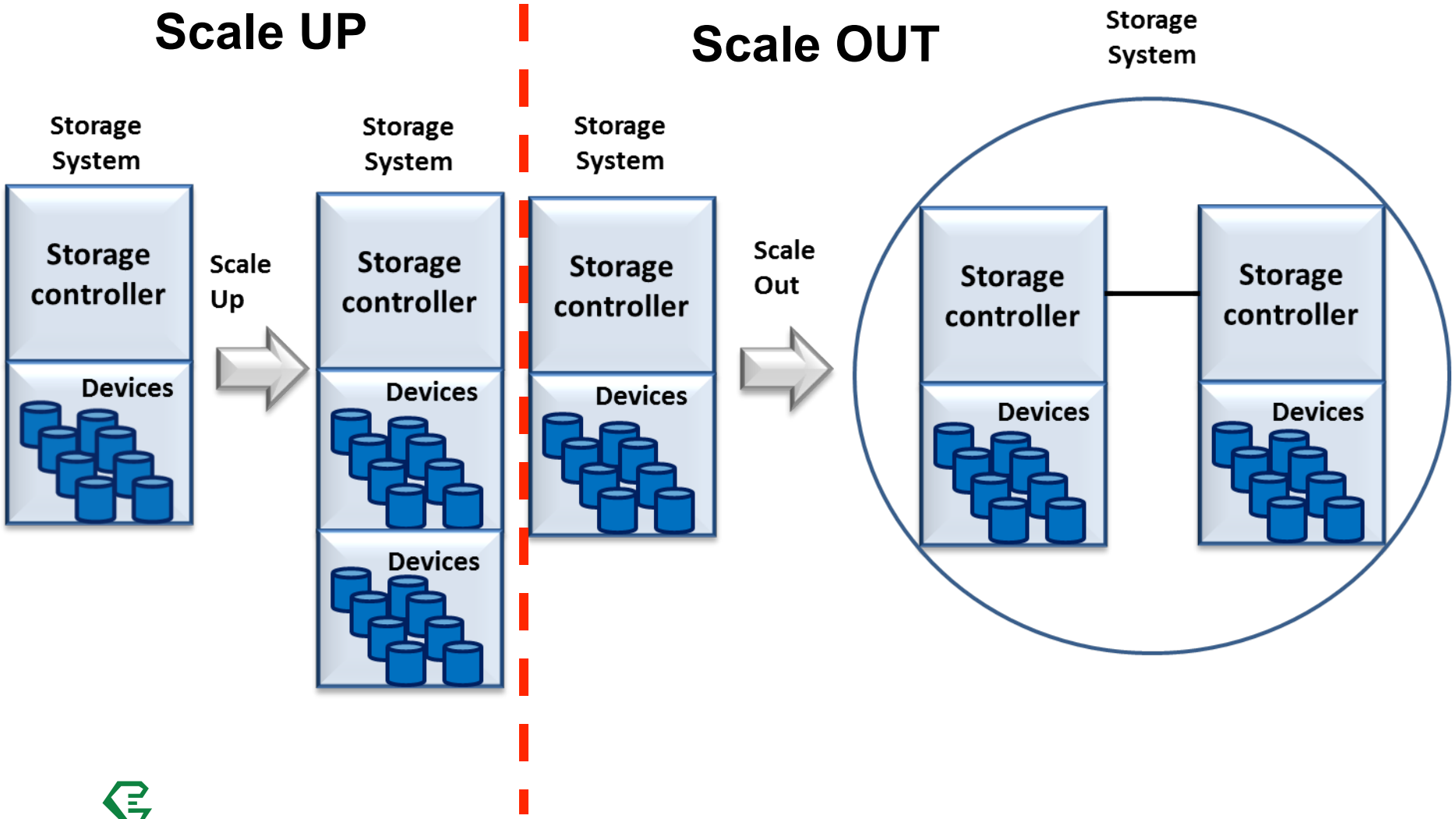
Traditional HDD Arrays

- Rotational Latency
- Seek Times
- Mechanical parts
- Controllers designed to handle HDD
- HDD = 8-12 watts per drive
- Heat 30 – 40 BTU/h per drive

Storage Technology

SCALE OUT VS SCALE UP

Scale Out vs. Scale Up



Scale Out Storage: NAS

- Implemented using distributed file system usually
  - Clustered hardware
  - Global namespace across nodes
- Key implementation points
  - Linkage between nodes – InfiniBand, Ethernet:10Gb
  - I/O balancing between nodes
  - Capacity balancing between nodes
  - Coherency – across distance
  - Switching requirements

Scale Out Storage: Block Storage

- Two approaches used
- Multiple controller cards to common backend storage device pools
  - Backplane connected typically
  - Normally associated with high-end enterprise systems
- Federation of separate controller nodes (included with NAS systems in some cases as integrated unified storage)
  - Complexities in cache coherency and I/O routing
  - Vendor differences

Storage Technology

DATA SERVICES

Considerations for Solid State Storage Arrays

Data Services Management	Investment Protection	Support
Data Reduction - Deduplication - Compression - Thin Provisioning	**Self-healing techniques (Reliability)**	**Hypervisor** - VMware vSphere - MS Hyper V
Replication - Local (writable) - Remote (Future)	**Hardware Redundancy (Availability)**	**Scale out and Clustering**
Management - Non-disruptive upgrades - REST APIs	**Serviceability**	**Application & OS**

Storage Technology

REPLICATION

Why Replicate Data?

- Migrate data at same or different location
  - New technology or added capacity
  - Consolidations
  - Mergers
  - Workload balancing
- Application testing
- Data exchange for warehousing / mining
- Business continuance, Disaster Recovery, Data Protection
  - Motivated by:
    - Business economics
    - RPO and RTO
    - Legal requirements

Replication Technologies

- Point-in-Time Copy – Snapshot / Flashcopy
  - Space efficient copy – only changed data
  - Variations in storage system implementations
- Remote replication
  - Synchronous – continuous transfer of all data changed, operations wait until transfer completes
  - Asynchronous – data sent but operation continues without waiting for transfer completion
  - Asynchronous periodic – data sent periodically, most often is incremental snapshot

Types of Point-in-Time Copy

- Snapshot
  - Copy-on-write - only changed data is copied
  - Redirect-on-write – writes to new location
  - Various implementations use pointer manipulations, side files, or dynamically allocated space from storage pool
- Cloning
  - Makes a complete copy of data
  - Clone may be continuous or split and later resynced
- Variations are differentiation points for vendors

Remote Copy Approaches

- Synchronous
  - Storage system-based or through fabric appliance
  - Data must be stored at remote site before application can resume
    - Significant impact on performance limits distance
  - Application stops if I/O can't complete
    - May stall application
- Asynchronous
  - Initially host software based, but system based is now available
  - Application continues before data is stored on remote site
  - Delayed write represents a window of "risk"
  - continuous or periodic transmission of data
- Multi-hop or cascading
  - Combination of synchronous and asynchronous with intermediate storage system