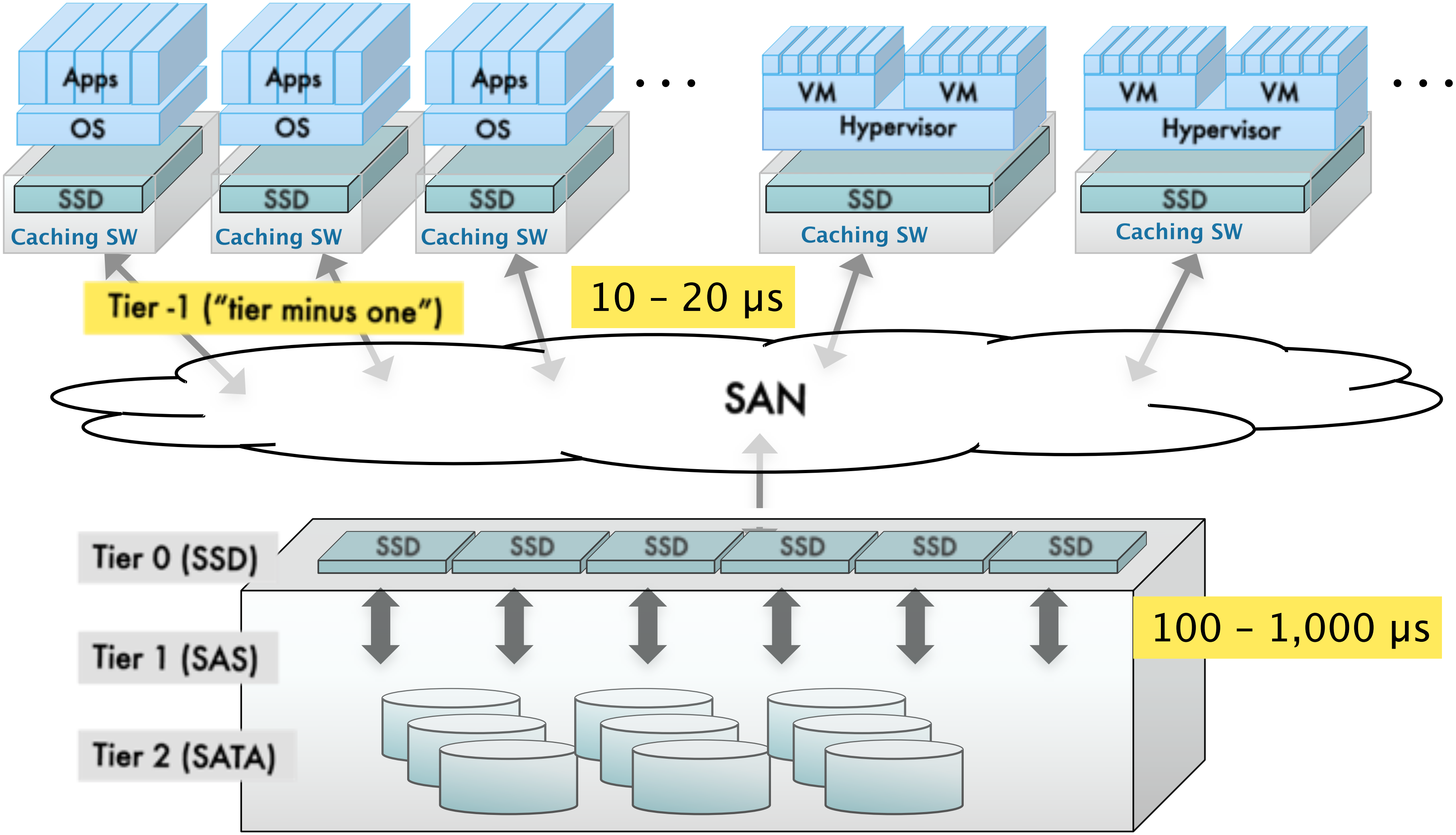
Flash as a Cache: Challenges & Opportunities

Ted Sanford
CEO, FlashSoft

Agenda

- Technical Choices & Challenges
  - What is “tier minus one” caching?
  - How can it be implemented?
  - When is it useful?
  - Key design considerations
- Use Cases & Performance Impact
- Architectural Considerations
  - Accelerating a stand-alone server
  - Distributed cache for server cluster
  - Extending System memory
  - Multi-level solid state cache design
  - Managing IO resources

What is T –1 caching?



When is T -1 caching useful?

- IO must be a significant bottleneck
  - CPU not fully utilized
  - Sufficient memory available (esp for VMs)
- Data must have hot spots
  - The 80-20 rule
  - The 90-90 rule – hot spots are always changing
  - Goal: matching the economic value of the data to the appropriate IO resource

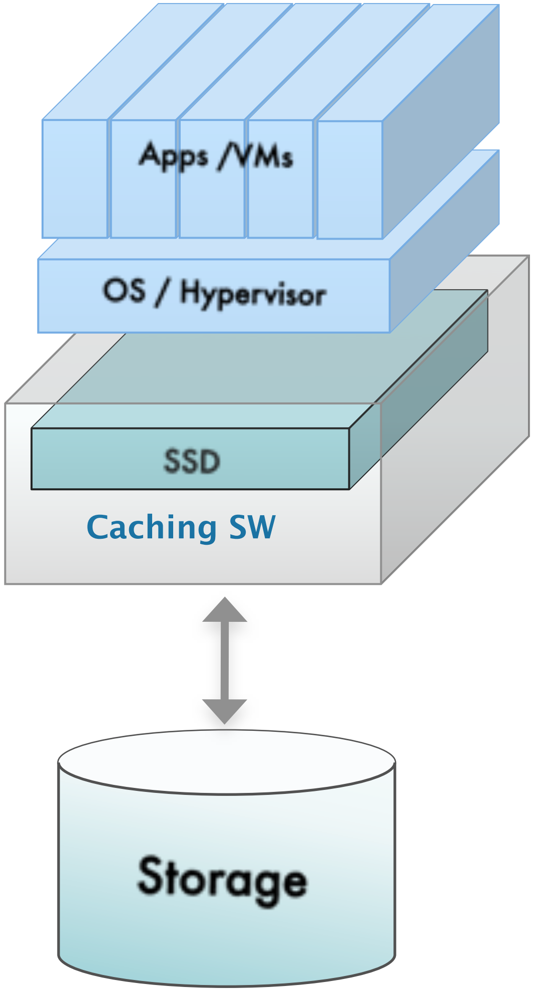
Key Design Considerations

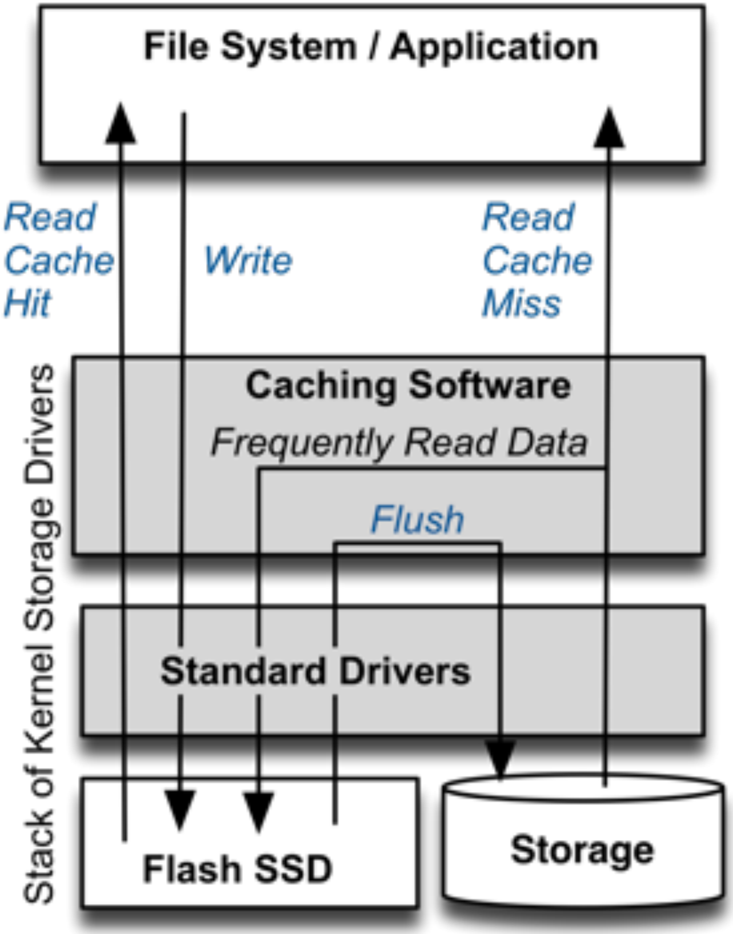
- Designing cache optimized for flash architecture
  - Traditional buffer cache design
  - Circular buffer, log-structured cache
- Minimizing server overhead
  - Memory
  - CPU
- Read-only vs. Read-Write
- Accommodating diversity of SSD designs
  - Size
  - Speed
  - Block architecture
  - Controller

Use Cases

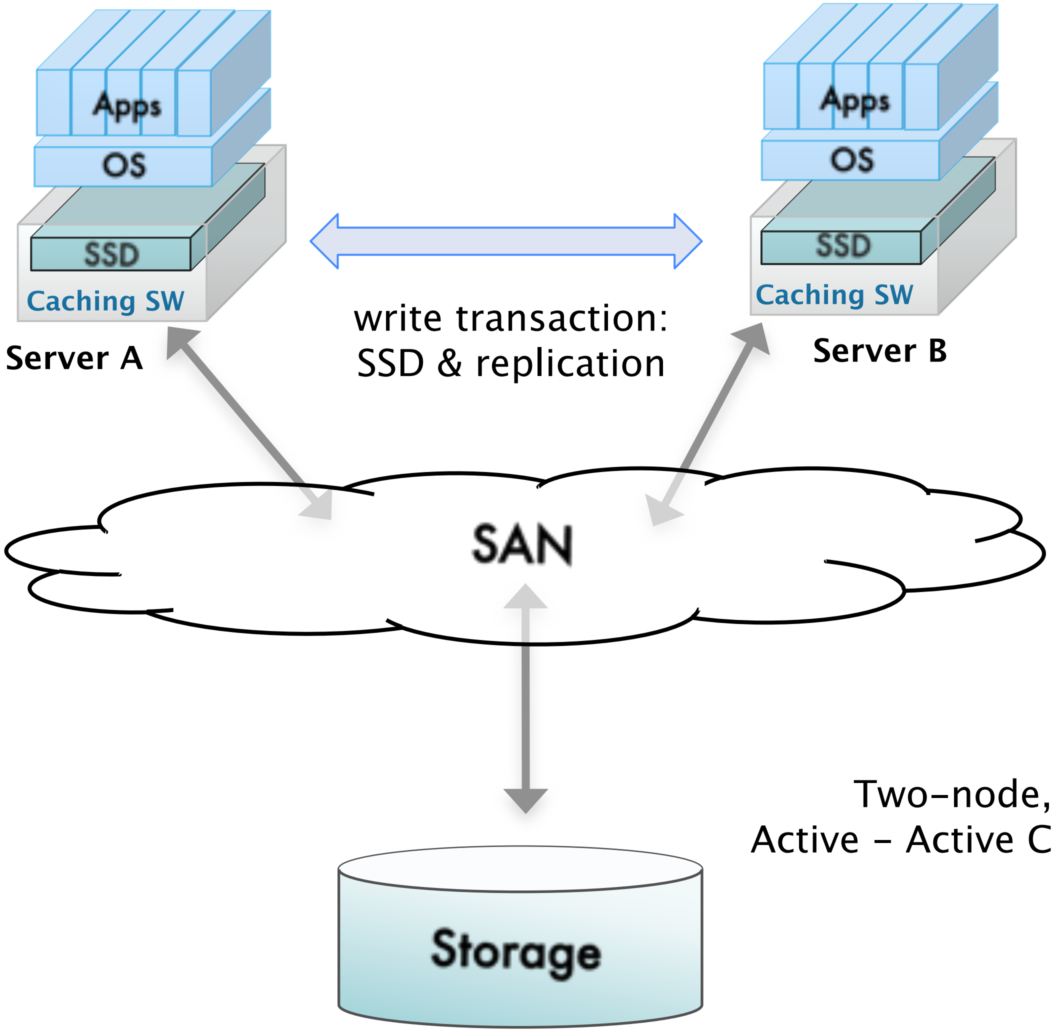
- Accelerating applications on stand-alone and cluster servers
  - Windows
  - Linux
  - Virtualization
- Applications most likely to benefit
  - Databases: Analytics, BI, ERP
  - Websites: Social, Search, SaaS, E2.0

Caching in a stand-alone server





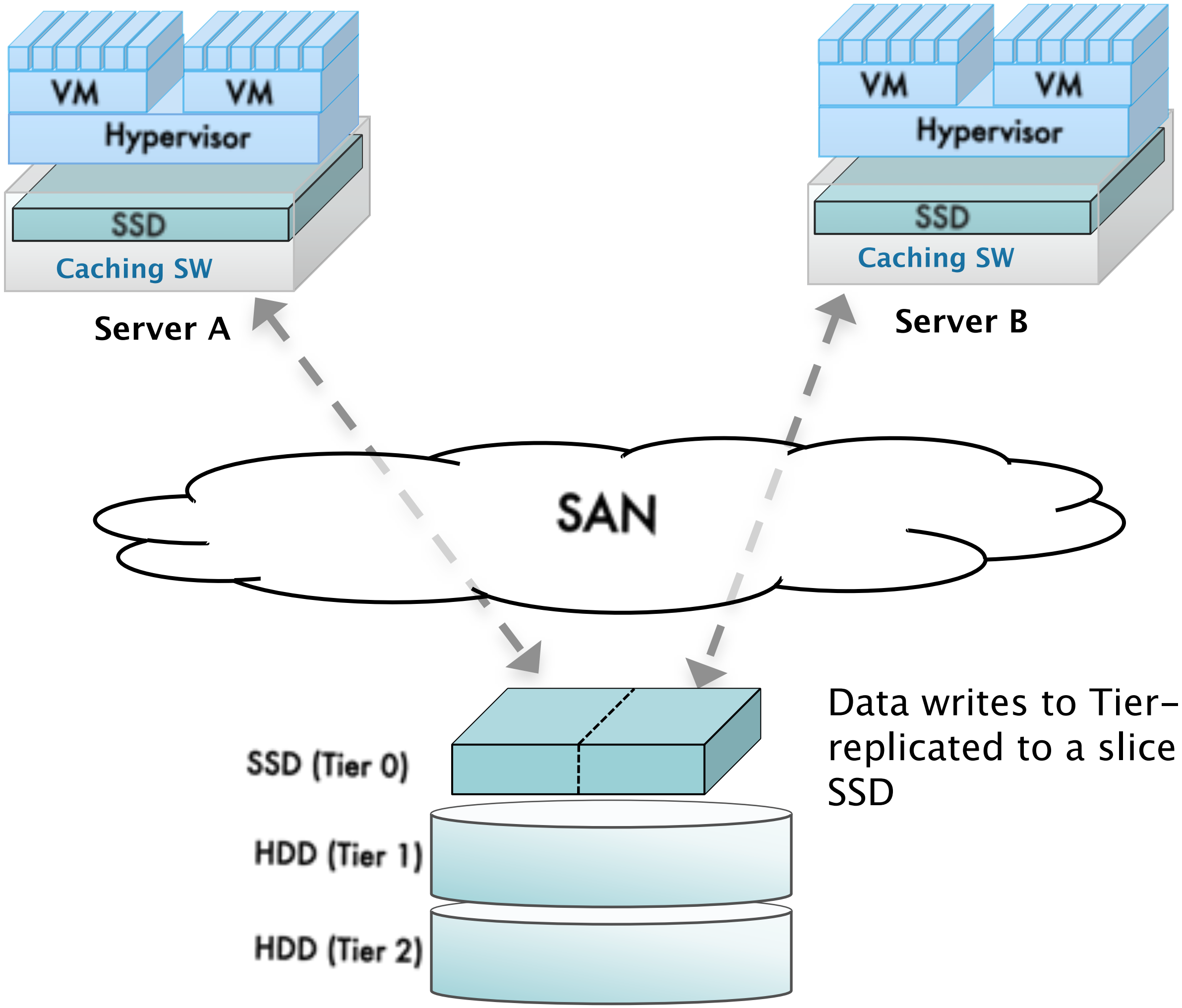
Cluster: Horizontal Replication



write transaction:
SSD & replication

Two-node,
Active – Active Cluster

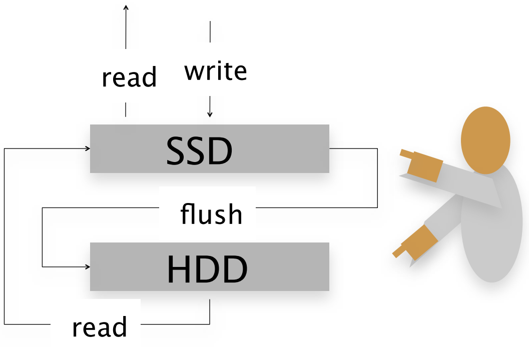
Cluster: Vertical Replication API



Data writes to Tier-1 SSD replicated to a slice of the Tier 0 SSD

Multi-level solid state cache design

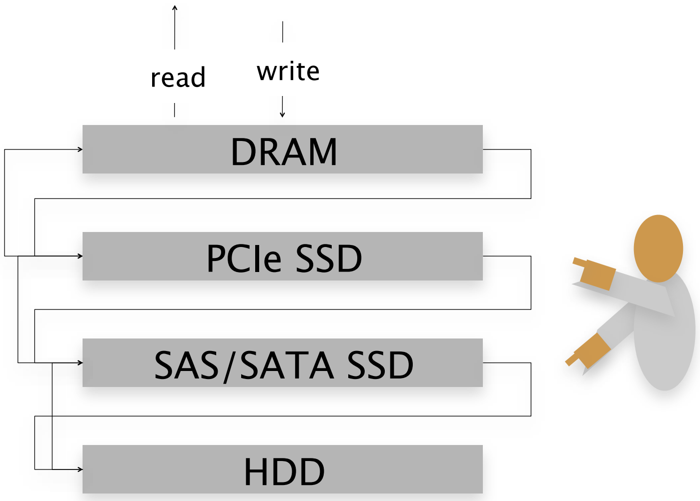
Today:



Two Tier caching

- allocates data to either SSD or HDD

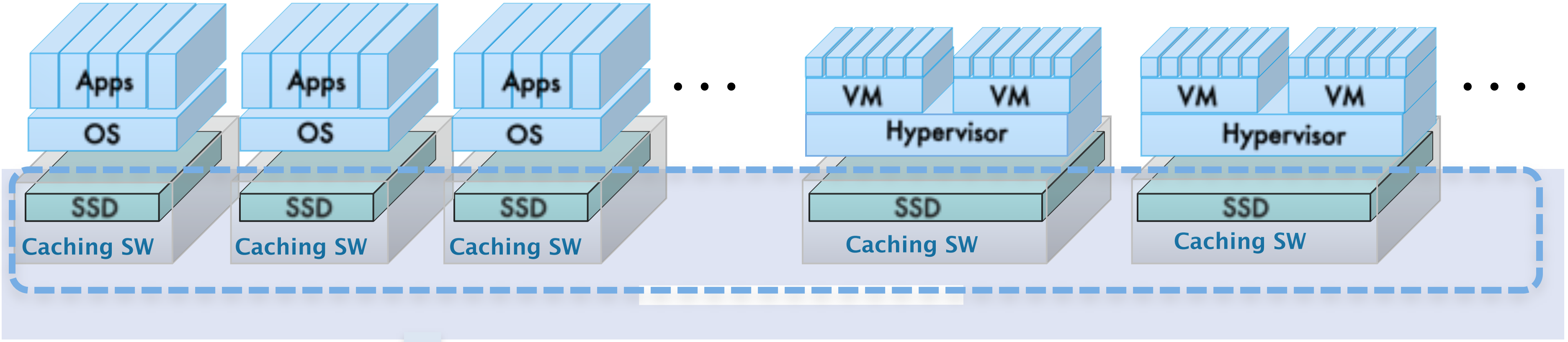
Future:

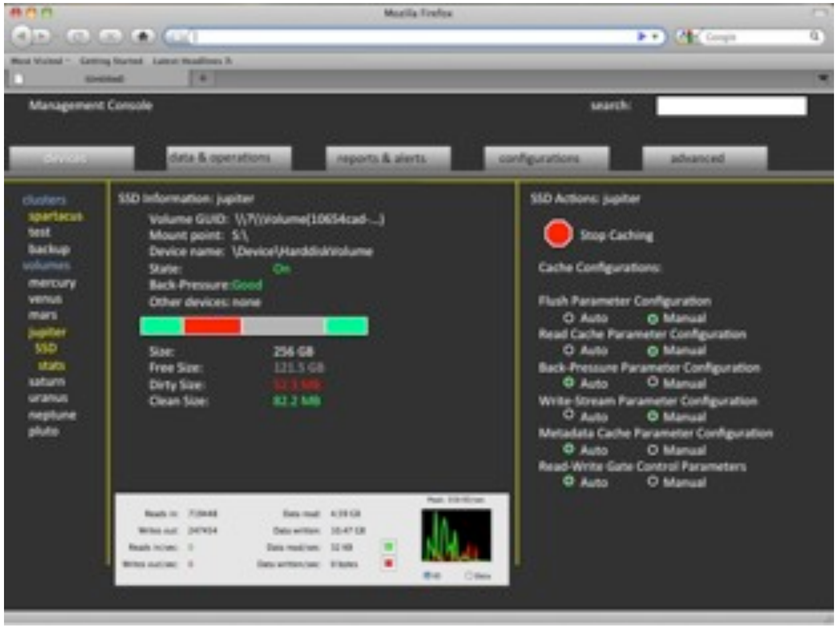


Multi-tier caching

- allocates data to DRAM, PCIe SSD, SAS/SATA SSD, or HDD
- Extensible to NVMe, PCM

Managing IO resources





SSD Information: jupiter

Volume GUID: \\?\Volume{10654cad-...}
Mount point: S:\
Device name: \Device\HarddiskVolume
State: On
Back-Pressure:Good
Other devices: none

Size: 256 GB
Free Size: 121.5 GB
Dirty Size: 32.5 MB
Clean Size: 82.2 MB

Reads in: 719448 Data read: 4.19 GB
Writes out: 247414 Data written: 10.47 GB
Reads in/sec: 0 Data read/sec: 32 KB
Writes out/sec: 0 Data written/sec: 0 bytes

SSD Actions: jupiter

● Stop Caching

Cache Configurations:

Flush Parameter Configuration
○ Auto ● Manual
Read Cache Parameter Configuration
○ Auto ● Manual
Back-Pressure Parameter Configuration
● Auto ○ Manual
Write-Stream Parameter Configuration
○ Auto ● Manual
Metadata Cache Parameter Configuration
● Auto ○ Manual
Read-Write Gate Control Parameters
● Auto ○ Manual

Flash Memory SUMMIT

Thank You