SDC
STORAGE DEVELOPER CONFERENCE
SNIA ■ SANTA CLARA, 2014

Method to Enhance the Performance of a Storage Array System with SSD Drives

Dr. M. K. Jibbe, Technical Director, NetApp
Kuok Hoe Tan, Senior QA Engineer, NetApp

Abstract

- As the enterprise adoption of SSD in the datacenter continues to accelerate, the need to further understand SSD characteristics with various enterprise workload has become more critical to the success of an all flash storage platform. For the NetApp EF-Series of all SSD storage arrays, we have been studying and analyzing different methods on how our customers can fully realize overall application performance gains after deploying EF-Series arrays into their datacenters by enhancing how we approach array performance optimization to guarantee consistency low latency I/O performance. In this presentation we will be covering the various workload which were benchmarked, the results of the various optimization done, and their impact to the overall performance of the NetApp E-series arrays

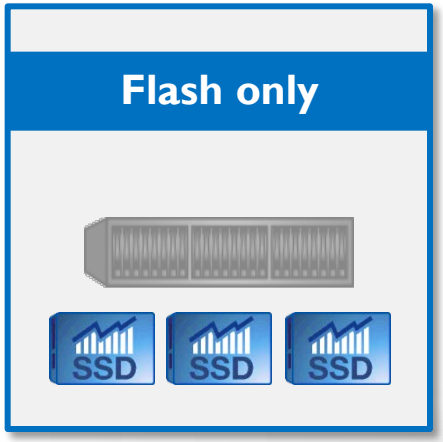
Agenda

- Performance Criteria are
  1. Queue Depth
  2. Dual-active paths from each controller to every SSD
  3. Dynamically disabling of full stripe writes
     a) Disabling of write caching under certain conditions
  4. Dynamic Disk Pool (DDP)
  5. Processor Swap
  6. Multicore
  7. Best Practices: Design the volumes to establish a high I/O rate with a low latency
     a) Volume count (Small),
     b) Volume capacity (large Capacity), and
     c) Volume RAID level (DDP or 8+2 RAID 6, others)

Flash-Accelerated Storage

Hybrid Array

- Best mix of $/IOPS and $/GB
  - Good average latency
  - Leverages HDDs for capacity
- Ideal for consolidated workloads



All-Flash Arrays

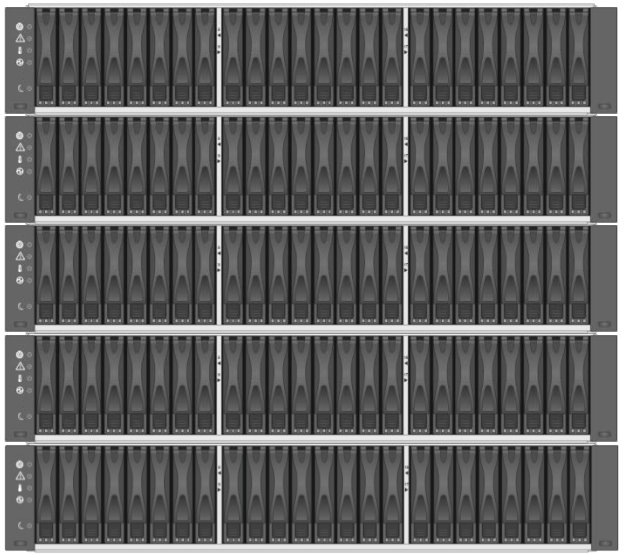
- Predictable ultralow latency
- Extreme IOPS and throughput
- Ideal for performance-driven apps

Leveraging All-Flash Where it Counts

- Transactional workloads (OLTP)
- Analytics (OLAP)
- Virtual Desktop Infrastructure (power users)
- Decision support systems
- Customer service systems
- Technical computing
- Web services

EF550 Product Specifications

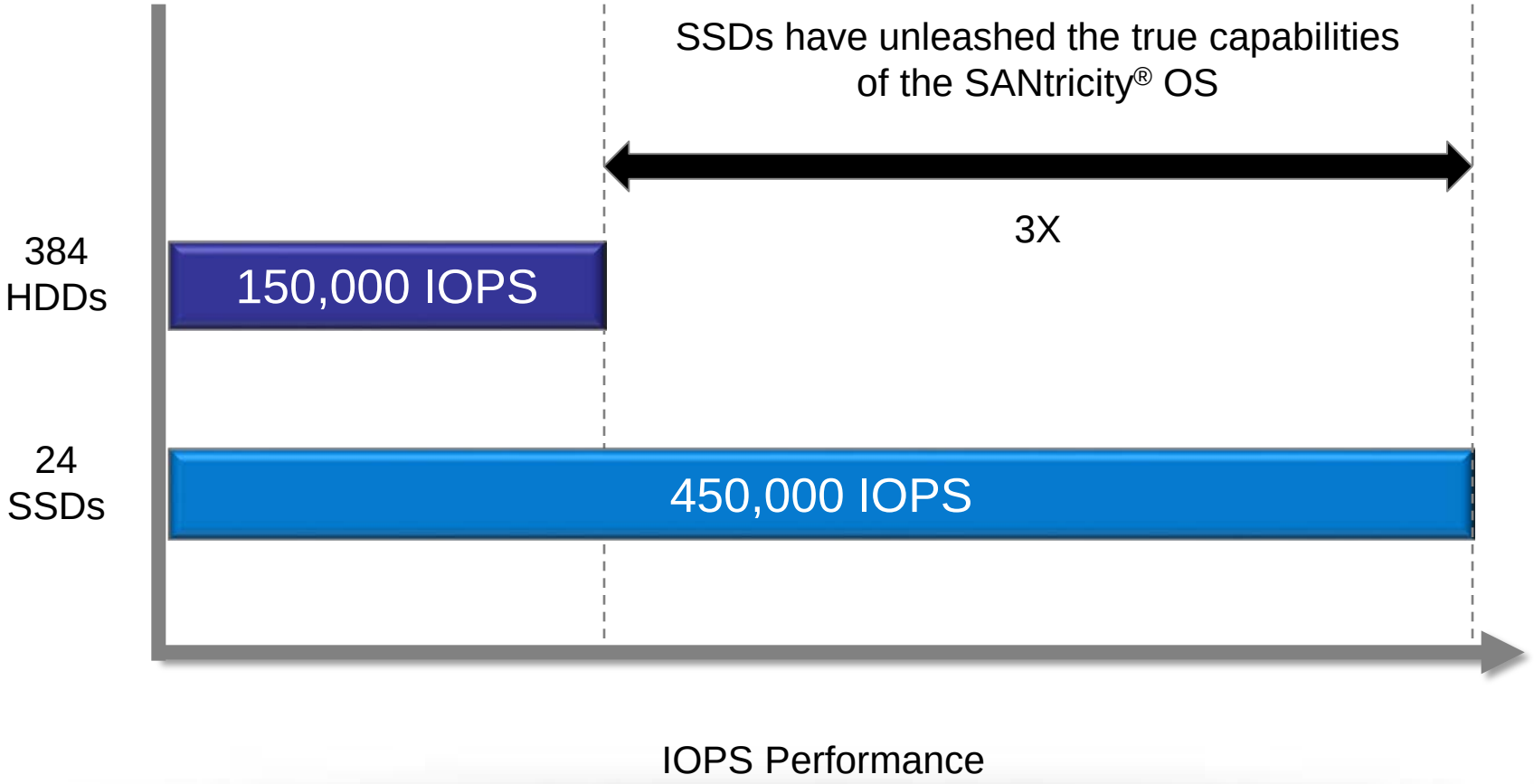
Min / Max base models	▪ 6 x 400GB SSDs (2.4TB) ▪ 24 x 1.6TB SSDs (38.4TB)
Expandable	▪ Up to 120 SSDs in single or multiple SSD increments
Max capacity	▪ 192TB (120x1.6TB)
SSDs	▪ 400GB, 800GB, 1.6TB SSD
I/O interface options	▪ (8) 16Gb FC ▪ (8) 10Gb iSCSI ▪ (8) 6Gb SAS ▪ (4) 40Gb IB
Storage OS	▪ SANtricity® 11
Key Features	▪ Dynamic disk pools ▪ Thin provisioning ▪ Snapshot™ copies ▪ Volume copy ▪ Remote mirroring



Performance:

- Burst: 900,000 IOPS
- Sustained: over 400,000 IOPS @ <1ms latency
- Sustained: up to 12GB/s

Designed for Performance



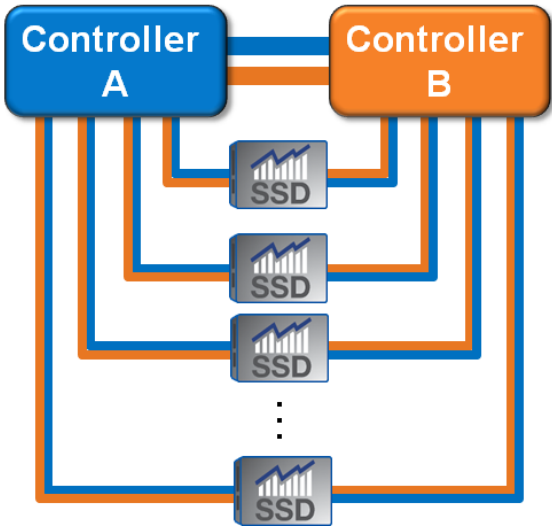
Basic Characteristics of EF Series

- The EF-Series storage arrays are highly configurable, providing many tuning options that allow customers to specifically address their particular workloads.
  - RAID levels 0, 1, 10, 5, and 6 are fully supported.
  - Dynamic disk pools, a fully featured RAID implementation based on RAID 6, offers exceptional rebuild performance over traditional RAID configurations with its own unique integrated hot sparing, thin provisioning support, and I/O load balancing capabilities via specialized distributed data and parity locations.
  - Segment size and cache block size are also customer-configurable options.
  - Each array features enterprise-level premium features as well, including snapshots, volume copies, remote volume mirroring, and persistent cache backup in the event of a power failure.

Optimization parameters: SSD-Specific Functionality

- SANtricity® proactively tracks the wear life of each SSD and issues an alert if defined thresholds are reached
  - Average erase count over 80%
  - Spare blocks remaining under 20%

Media type:	Solid State Disk
Interface type:	SAS Serial Attached SCSI (SAS)
Drive path redundancy:	OK
Wear life monitoring:	Enabled
Average erase count:	0%
Spare blocks remaining:	100%



- Dual-active paths from each controller to every SSD
  - EF-Series drive-side I/O driver has been optimized to provide I/Os to both ports of an SSD drive simultaneously.
  - Whereas HDDs perform best when accepting I/O through one drive port at a time, despite being dual ported, due to the need to avoid disk-head thrashing,

Optimization Parameters: Queue Depth

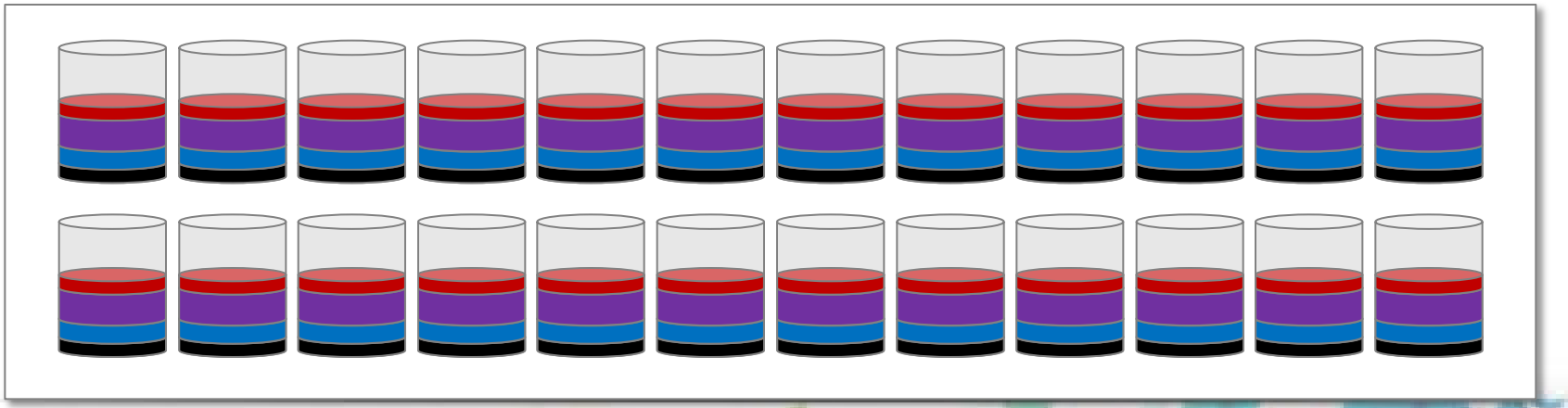
- SSD queue depth has been increased 4-fold, to 64.
  - Max queue depth of 128
    - 64 per path
- An increase in queue depth means that each SSD can handle up to 4 times the number of I/O operations from a software perspective.
- In a single-enclosure 24-drive system,
  - An EF550 is capable of issuing 1,536 open back-end drive I/Os at once,
  - An E5500 (Non SSD) is capable of only 384 I/Os

Optimization parameters: Full Stripe Write

- E-Series products use Full Stripe Write Through which is a special mode if the hosts I/O fit the following profile:
  - Aligns properly on stripe boundaries and
  - Stays that way long enough for CFW to complete a learning cycle
  - Applies when CME (Cache Mirroring Enable) is enabled.
  - Lets CFW write through the full host stripe instead of mirroring the data
  - Does later a full stripe write to empty the cache.
- If the host doesn't fit strict alignment and size limits compared to the array segments
  - The FSWT feature turns off and use regular Cache management
    - Much slower than mirroring data and writing later after CFW has full stripes to dump
- The learning mode is very invasive when trying to run performance tests not directly related to FSWT.
- Pathological cases can be contrived to force the learning cycle to turn on and off the FSWT, making a big variation in performance.

Dynamic Disk Pools Overview

- DDP dynamically distributes data, spare capacity, and parity information across a pool of SSDs
  - All drives are active (no idle hot spares)
  - Spare capacity is available to all volumes
- Data is dynamically recreated/redistributed whenever pools grows or shrinks



Optimization parameters: Processor Swap

OLTP-A results
- 17% IOP performance improvement from a 20% faster CPU
- Same improvement achieved at highest and target (<1ms) response time levels

Test #	Performance Level	Processor	IOP's	ART	Array CPU	IOP Increase
1	<1ms RT	2.0GHz	115K	0.937ms		
2	Max IO Rate	2.0GHz	145K	2.236ms		
3	<1ms RT	2.4GHz	135K	0.967ms		17%
4	Max IO Rate	2.4GHz	170K	2.129ms	97%	17%

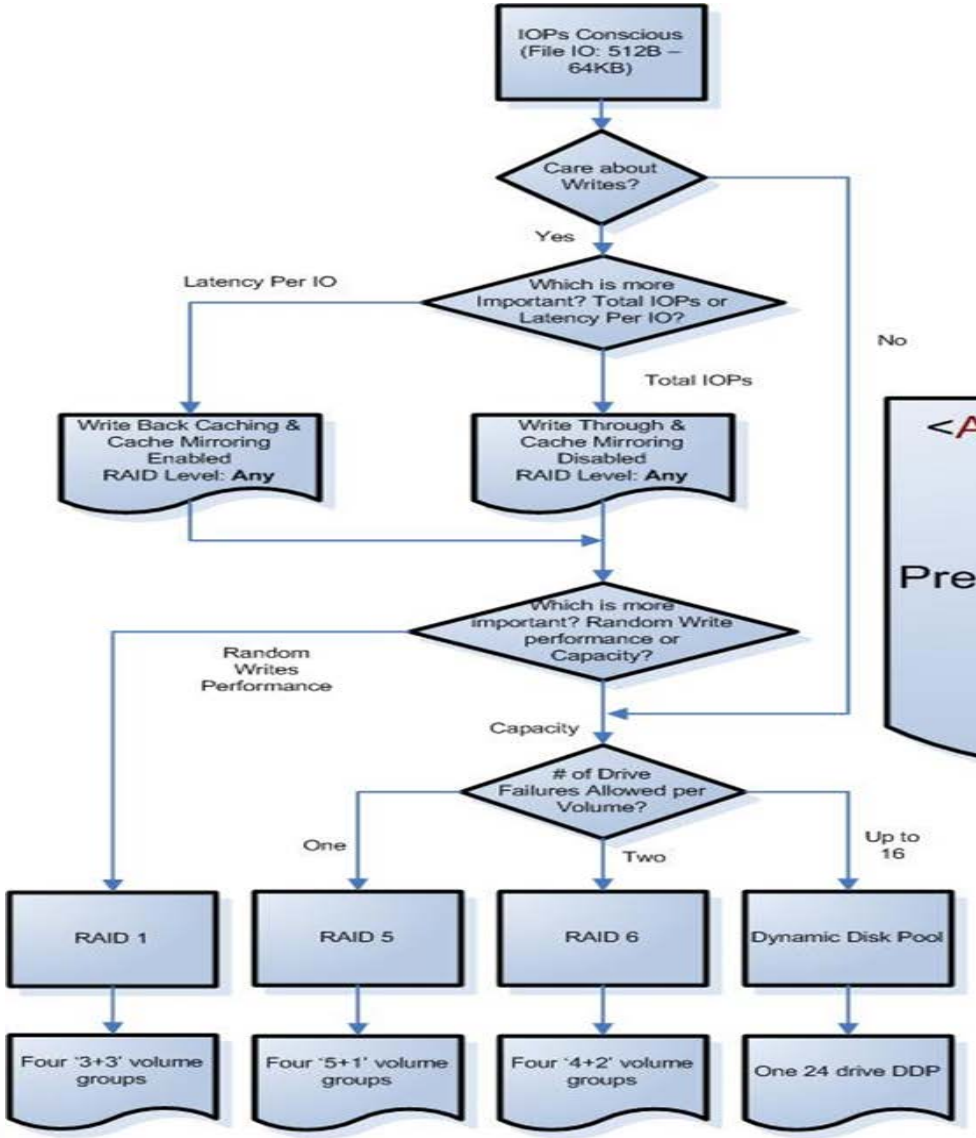
Optimization parameters: Multicore

- EF560 controller has the software stack is moving toward
  - Multiprocessing architecture using multiple cores
    - to bring to bear more computing resources
    - to take advantage of the greater I/O servicing capabilities of SSDs
- New multicore architecture has the following characteristics:
  - Handles RAID mapping and Cache requests in one core
  - Handle I/O traffic on both the front end and the back end of a given controller, as well as other parts of the software stack with additional cores
  - Splits the I/O path into pieces, so that the computing resources of all utilized cores are heavily leveraged during the course of a single I/O
  - Allows each part of the I/O path to grow corresponding to resource use without involving the resources of the other core
- For example, a higher-intensity front-end I/O activity will no longer affect the computing power available to the RAID and cache portions of the I/O path.

Best Practices: Design the volumes to establish a high I/O rate with a low latency

- Key Storage Array and Volume Settings for IO Performance
  - **RAID Level Selection**
  - **Segment Size**
  - **Cache Block Size**
  - **Cache Read Ahead / Read Ahead Prefetch Feature**
  - **Write-Through (WT) Caching**
  - **Non-Volatile Write-Back (WB) Caching**
  - **Data Integrity Features**
    - **Pre-Read Redundancy Feature**
    - **Data Assurance (T10-PI) Feature**
- **Typical IO Workload Characteristics**
- **Host Consideration & IO Tool**
  - Type of initiator
  - Here is a list of considerations for the initiator
    - number of Host Bus Adapter (HBA)
    - Initiator PCI Slot type where HBA is inserted
    - Queue Depth of the HBA
    - Failover driver load balancing (if applicable)
    - IO tool (eg. IOGEN) & parameters
      - Queue Depth
      - IO Size of 4K
      - Random IO
      - Read/Write Ratio

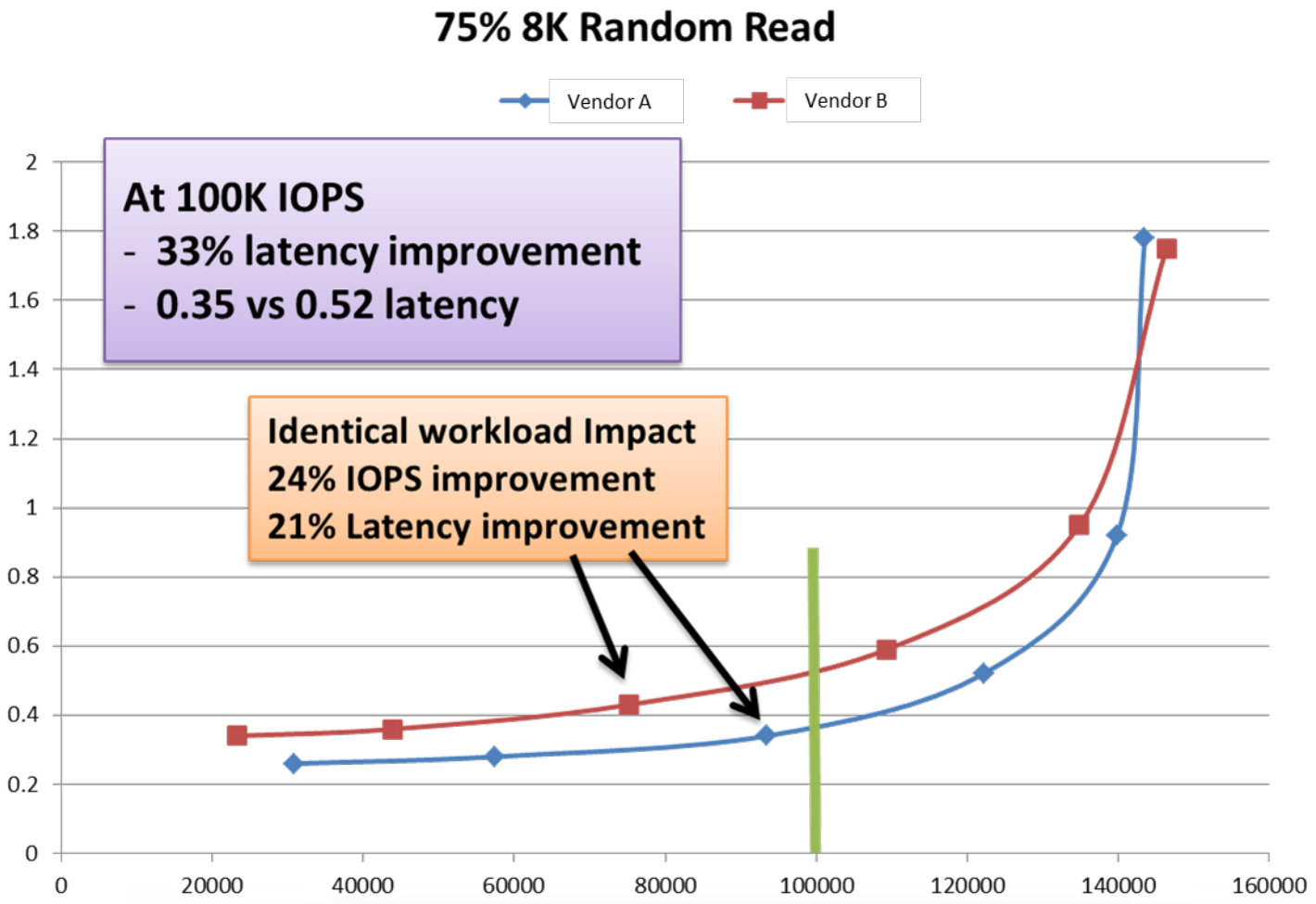
Best Practices: Design the volumes to establish a high I/O rate with a low latency



<Applies to both Read or Write>
Segment Size: **128K**
Cache Block Size: **32K**
Read Pre-fetch: **Off**
Pre-Read Redundancy Check: **Off**
Data Assurance (PI): **Off**
Device Type: **Raw**

Drive Vendor A vs Drive Vendor B – 24 SSDs

8K Random 75% Read / 25% Write

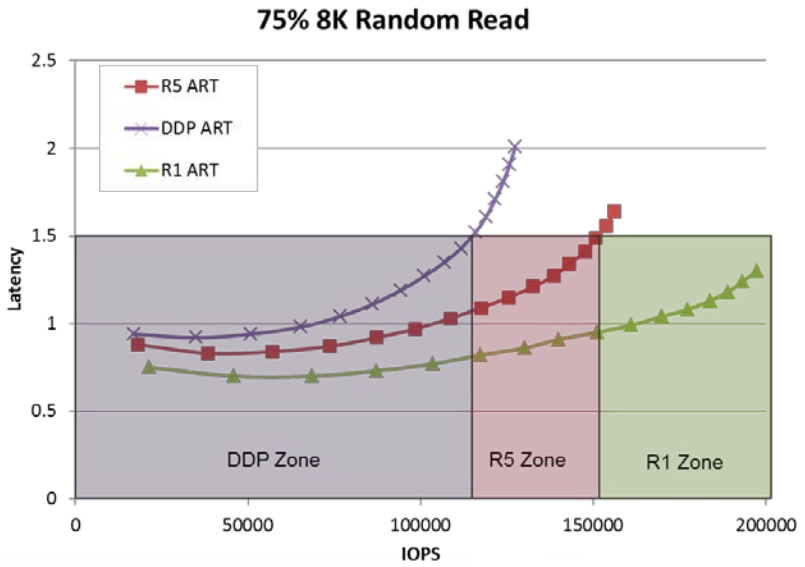


Which Protection Scheme To Pick

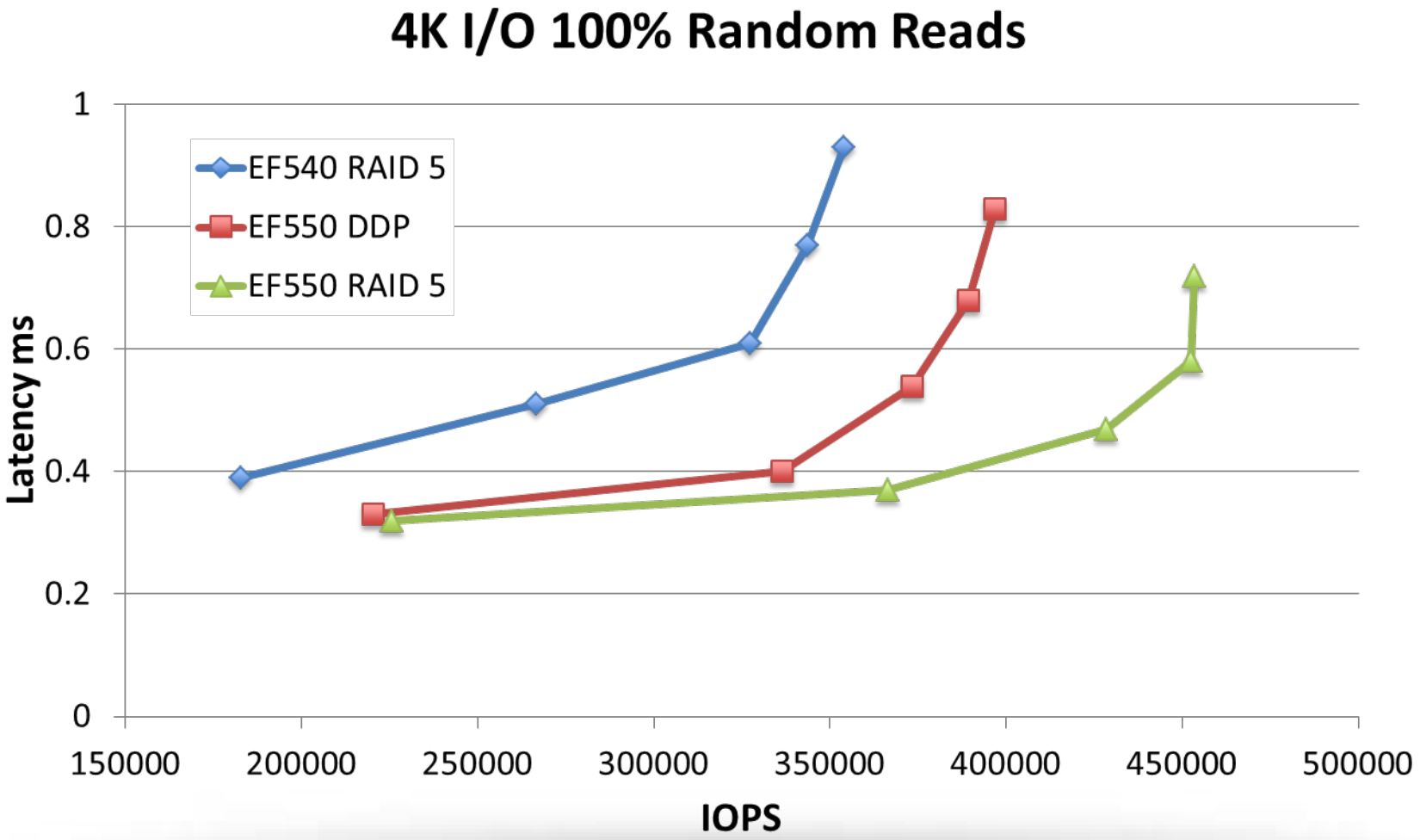
- Look at your customer's requirements
  - Think of the different protection schemes and performance metrics as zones
  - Select highest protection that will meet performance and capacity requirements

Customer requirement
- 8K 75% reads
- Latency <1.5ms
- IOPS ~100,000

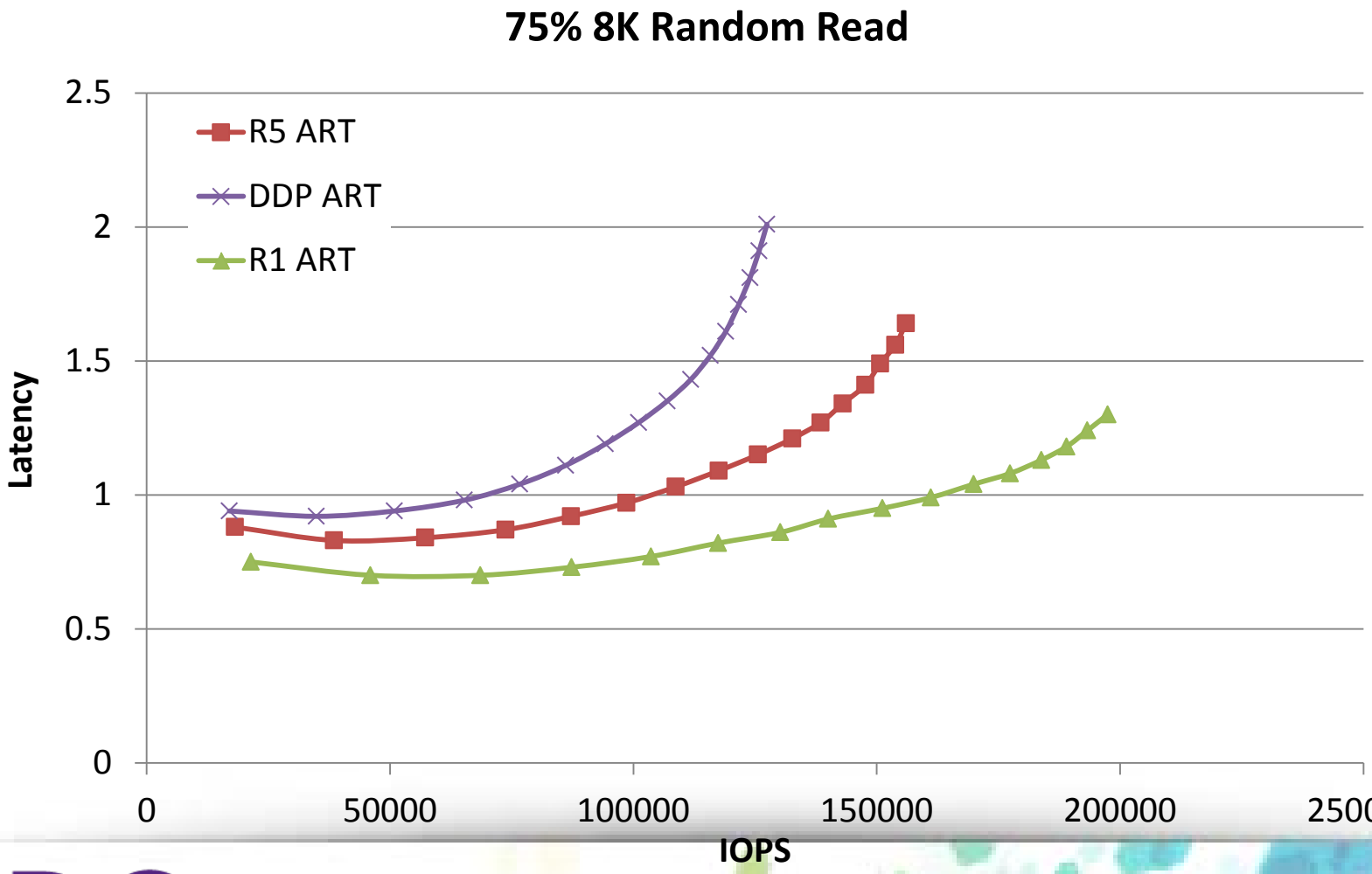
DDP will meet the requirement and provide best protection and ease of use



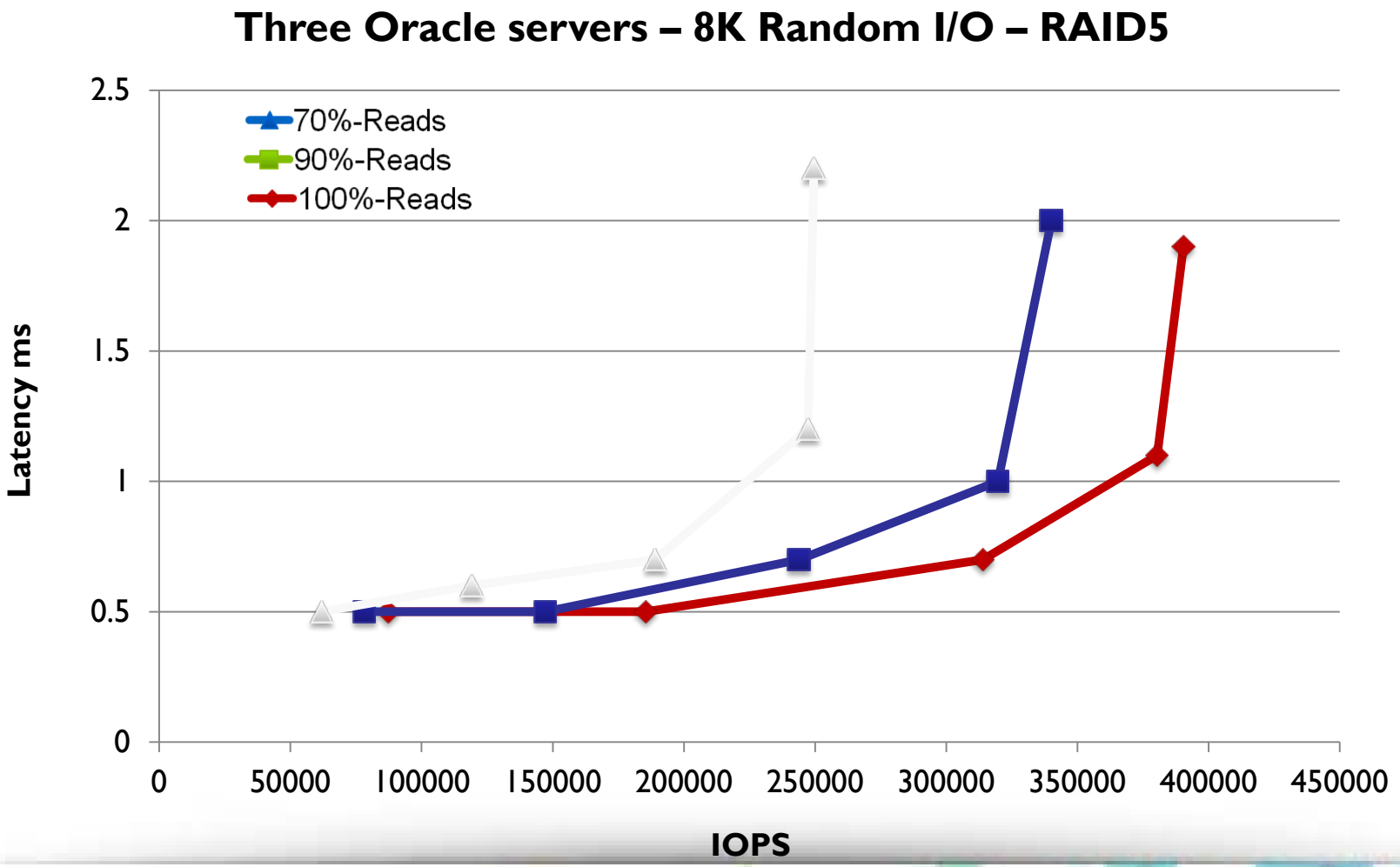
EF-Series Performance Comparison



EF550 Performance: 8K Random 75% Read / 25% Write



EF550 Oracle Performance



Performance Overview

	EF540	EF550
Burst I/O rate – cache reads (512B)	800,000 IOPS	900,000 IOPS
Sustained I/O rate – random disk reads (4K)	300,000 to 330,000 IOPS	400,000 to 450,000 IOPS
Sustained I/O rate – random disk writes (4K)	65,000 IOPS	75,000 IOPS
Sustained throughput – sequential disk reads (512K)	6 GB/s	12 GB/s
Sustained throughput – sequential disk write (CMD 512K)	4 GB/s	10 GB/s
Sustained throughput – sequential disk write (CME 512K)	3 GB/s	6 GB/s

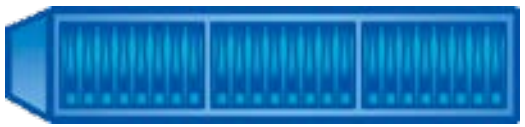
EF550 results using RAID 5

EF-Series: *Go Fast Go Green*

Starting off or Concerned about Datacenter Space

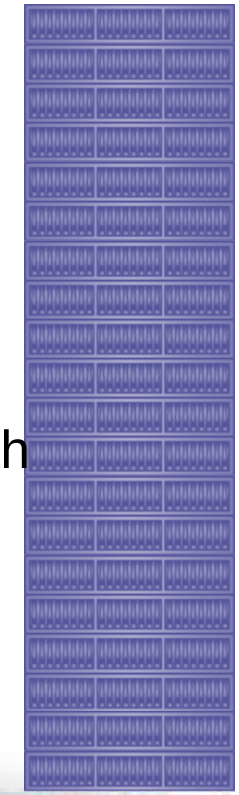
Starter Package

- Single application focus
- EF550
  - 2U
  - Six 400GB SSDs
  - 2.4 TB Raw
  - Over 100K IOPS
  - Nearly 400MB/s Bandwidth
  - Expandable up to 184 TB



Reduce Datacenter Tile Footprint

- Reduce datacenter cost
- Rack configuration
  - 40U rack – single tile configuration
  - 480 SSDs
  - 768TB of Flash
  - Over 9 Million IOPS
  - Over 220 GB/s Bandwidth



Thank you