Building a High IOPS Flash Array: A Software-Defined Approach

Weafon Tsao Ph.D.

VP of R&D Division, AccelStor, Inc.

Myth 1:
Σ High-IOPS SSDs = High-IOPS All-Flash Array

Clarification

SSDs are not the only hardware component of all-flash array

- Network adaptors connect user applications with SSDs
- Processors glue SSDs and network adaptors together

Myth 2: Off-the-shelf SW + SSD = High IOPS Flash Array

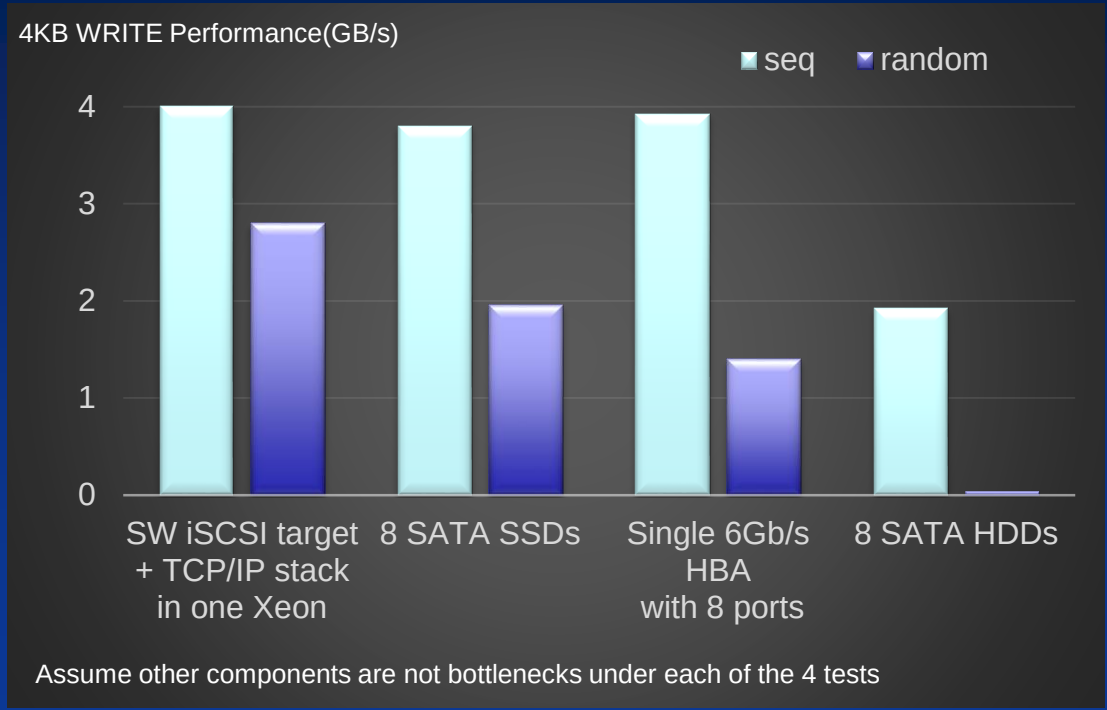
Clarification

Off-the-shelf SW/Technology may become the bottleneck

- Data redundancy
- Volume management/snapshot
- HA/scale out
- De-dup/compress

The Root of Challenge: Random IO

- Random IO cannot be merged into a big chunk
  - ⇨ cause large overheads
  - ⇨ impact the HW/SW design



Solution Statement

Build a high random IOPS all-flash array

with rich enterprise storage features

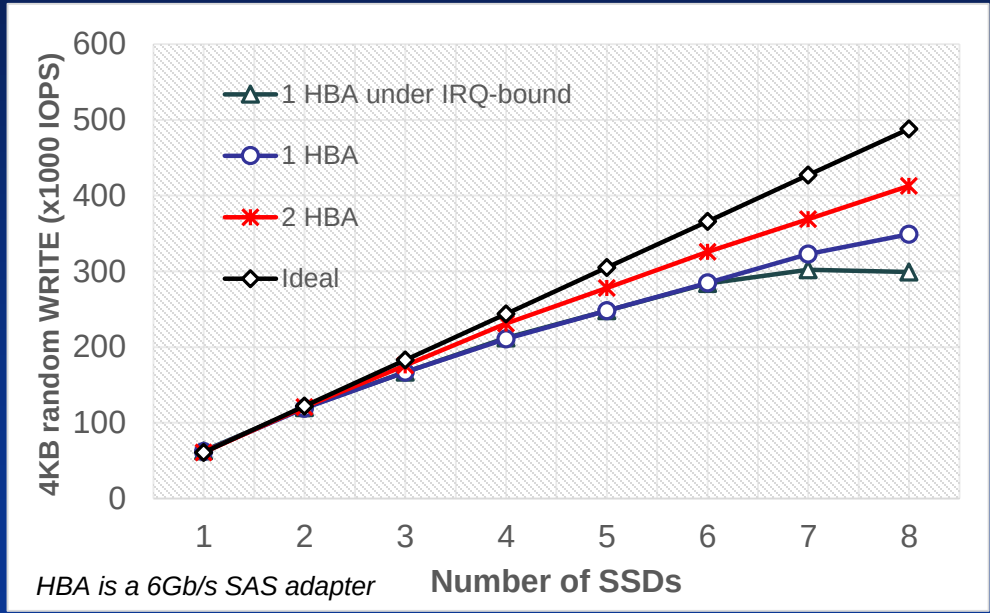
on commodity hardware?

Dimensions to Focus

A. Leverage latest generation hardware or hardware-offloading
B. Consume all the hardware resources
   - Considerate configuration
   - Multi-thread programming
C. Minimize overheads on
   - RAID5
   - Snapshot
   - HA/scale out
   - …

A. Leverage the Latest Generation of HW or HW-offloading

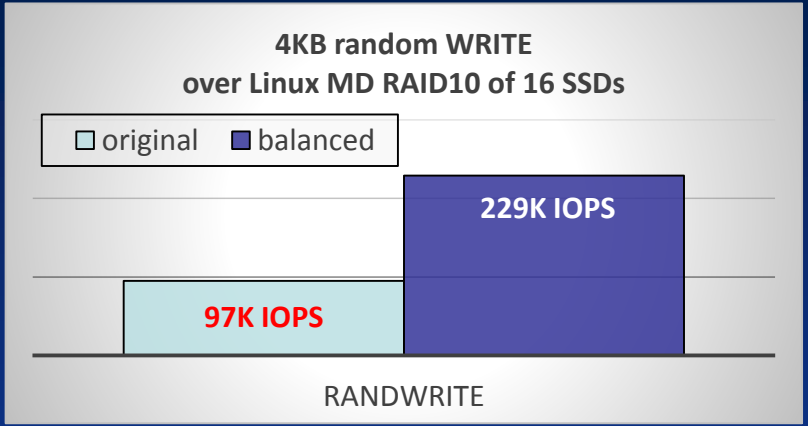
- NVMe, RDMA
- The performance of commodity hardware is limited.
- Commodity hardware are mainly designed for mainstream market with manufacturing cost consideration.



B. To Consume All The Hardware Resources

Considerately balance SW/HW IRQs on

- HBA
- Network: Ethernet / FC / InfiniBand

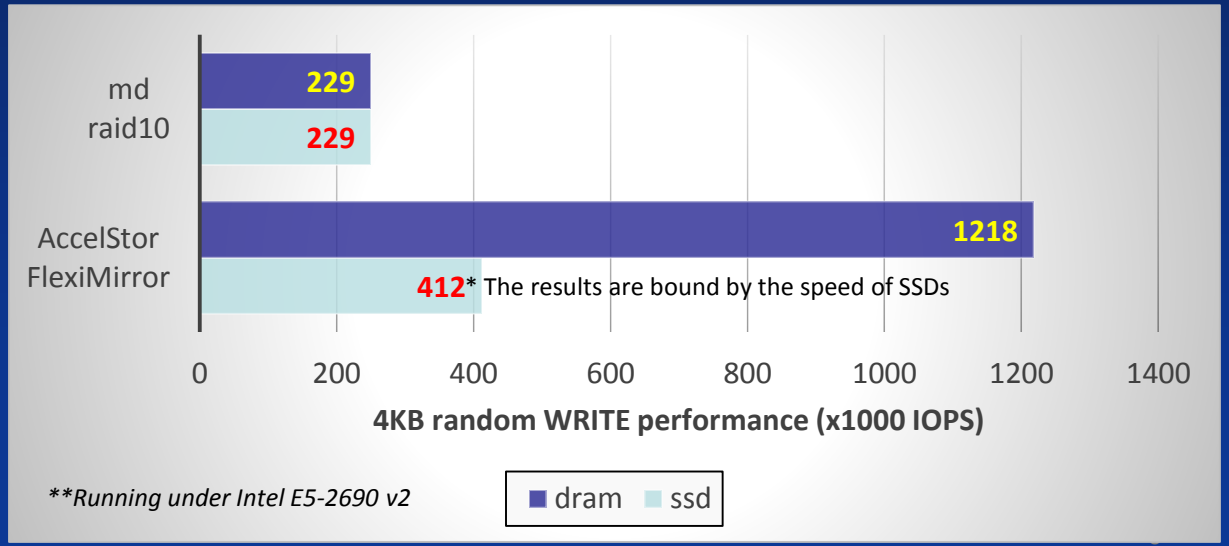



```
09:31:33 AM  CPU    %usr   %nice    %sys %iowait    %irq   %soft  %steal  %guest   %idle
09:31:34 AM  all    0.02    0.00    3.21    8.23    0.00    2.26    0.00    0.00   86.27
09:31:34 AM    0    0.00    0.00    0.00    0.00    0.00  100.00    0.00    0.00    0.00
09:31:34 AM    1    0.00    0.00   12.00    0.00    0.00    0.00    0.00    0.00   88.00
09:31:34 AM    2    0.00    0.00   12.87    0.00    0.00    0.00    0.00    0.00   87.13
09:31:34 AM    3    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
09:31:34 AM    4    0.75    0.00    0.00    0.00    0.00    0.00    0.00    0.00   99.25
```

Leverage the Power of Multi-core CPU

Multi-thread programming is necessary

PID	%usr	%system	%guest	%CPU	CPU	Command
687	0.00	100.00	0.00	100.00	10	md0_raid10
750	0.00	100.00	0.00	100.00	21	aw0
751	0.00	100.00	0.00	100.00	22	aw1
752	0.00	100.00	0.00	100.00	23	aw2
753	0.00	2.00	0.00	2.00	33	pidstat



C. Minimize Overheads on Data Protection

Overhead: Parity check update

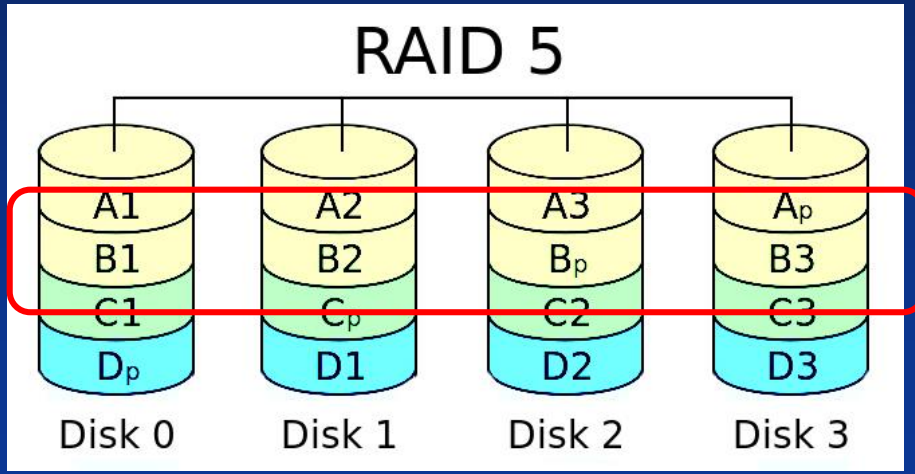
Solution: Write buffer? Low hit ratio for random WRITE

If you want to write B2

$B_p^{new} = B2^{old}$ xor B_p^{old} xor $B2^{new}$

W R R W

1W ⇨ 2R + 2W

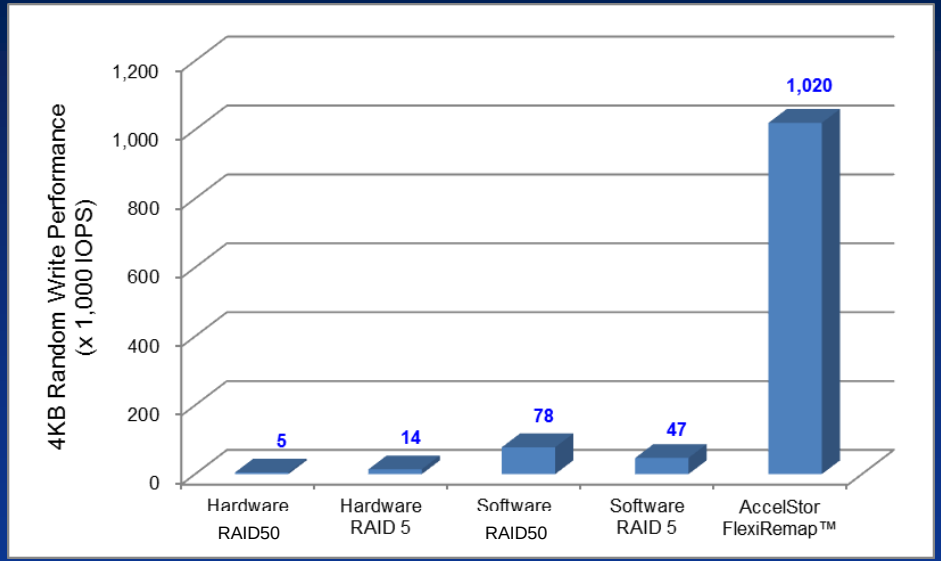


FlexiRemap Technology:
1M IOPS with Space-Efficient Data Protection

Compared to conventional RAID configuration:

- More efficient and effective redundancy without performance and lifespan penalty for SSDs

- Automatic workload redirection upon SSD failure



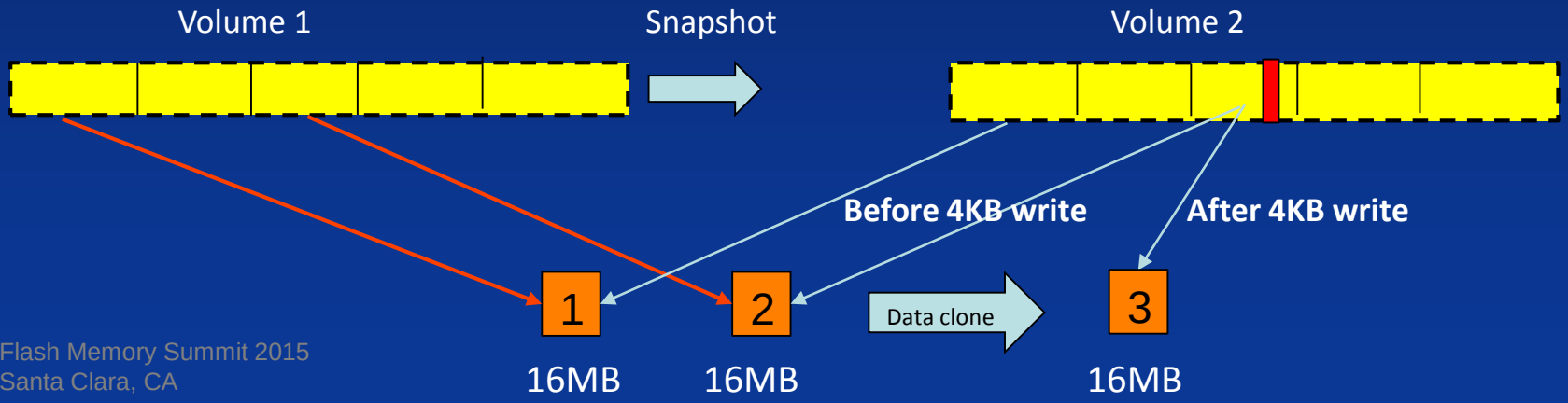
- 20 SSDs (55K IOPS)
- RAID50: 10 SSDs per group
- FlexiRemap: 10 SSDs per group

- accessed range: 80GB
- accessed amount: 40GB
- cache disabled

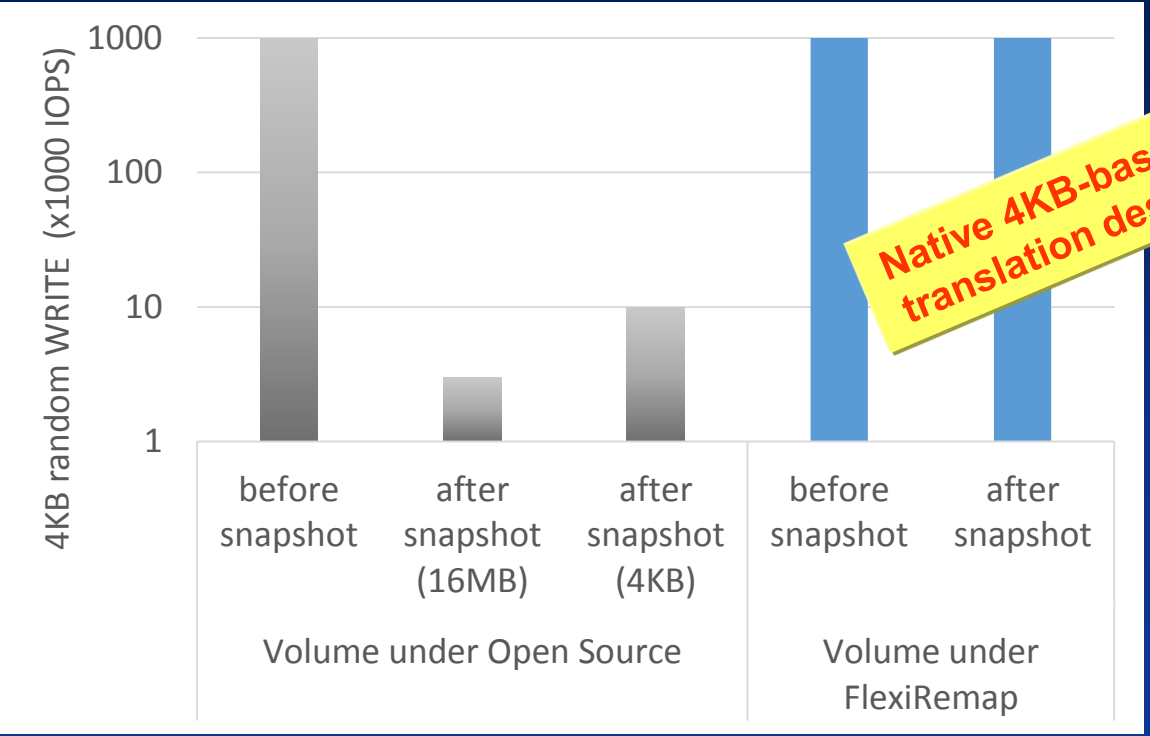
C. Minimize Overhead on Snapshot

- Overhead: small random WRITE
- Solution: small chunk with low overhead

16MB block size -> copy 16MB per 4KB write
4KB block size -> the overhead to allocate new block per 4KB write is too large



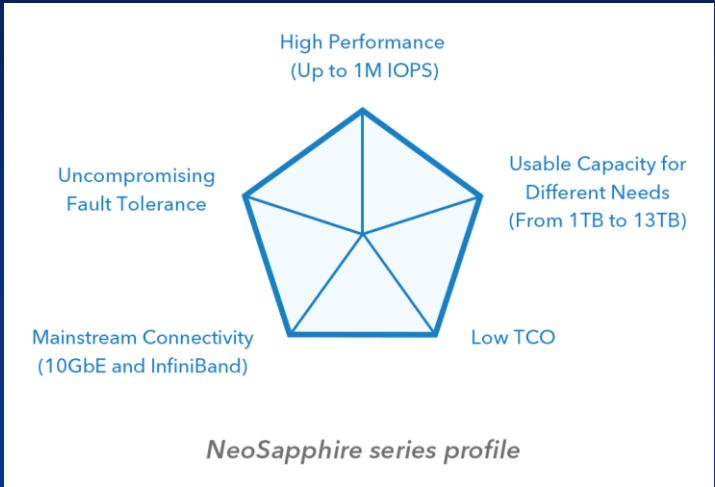
FlexiRemap Technology: Small Chunk with Low Overhead



NeoSapphire All-Flash Array Series

NeoSapphire All-Flash Array Series

- High performance with up to 1M IOPS for 4KB random write
- Fault tolerance and automatic data reconstruction upon drive replacement
- Low TCO with excellent performance and energy efficiency
- Web-based graphical management interface for simplified system setup, health monitoring and management



NeoSapphire series profile

About AccelStor

About AccelStor

Building upon its expertise in software and storage technology, AccelStor is devoted to unleashing the true performance of flash-based storage solutions with a software-defined approach. AccelStor has developed an exclusive FlexiRemap software technology that enables its storage arrays to achieve unparalleled scalability, performance, and efficiency in the same grade as such products.

Core Competence

- Innovations beyond Technical Fluency
- Dedication to Success
- Agility

Management team

- President: Charles Tsai Ph.D.
- Vice President: Weafon Tsao Ph.D.

Website

www.accelstor.com

Email

inquiry@accelstor.com

Calls to Action

- Visit us at booth #810
- Check out our live demonstration
  - AccelStor FlexiRemap technology
  - AccelStor NeoSapphire All-Flash Array
- Tell us about your applications
- Book your samples
- Customization requests are welcome!