STORAGE INDUSTRY
SUMMIT

Realizing the Benefits of the Convergence of Storage and Memory

JANUARY 20, 2015, SAN JOSE, CA

Reza Bacchus

Hewlett Packard

Options Chief Technologist

NVDIMM: Super-Charging the Data Center

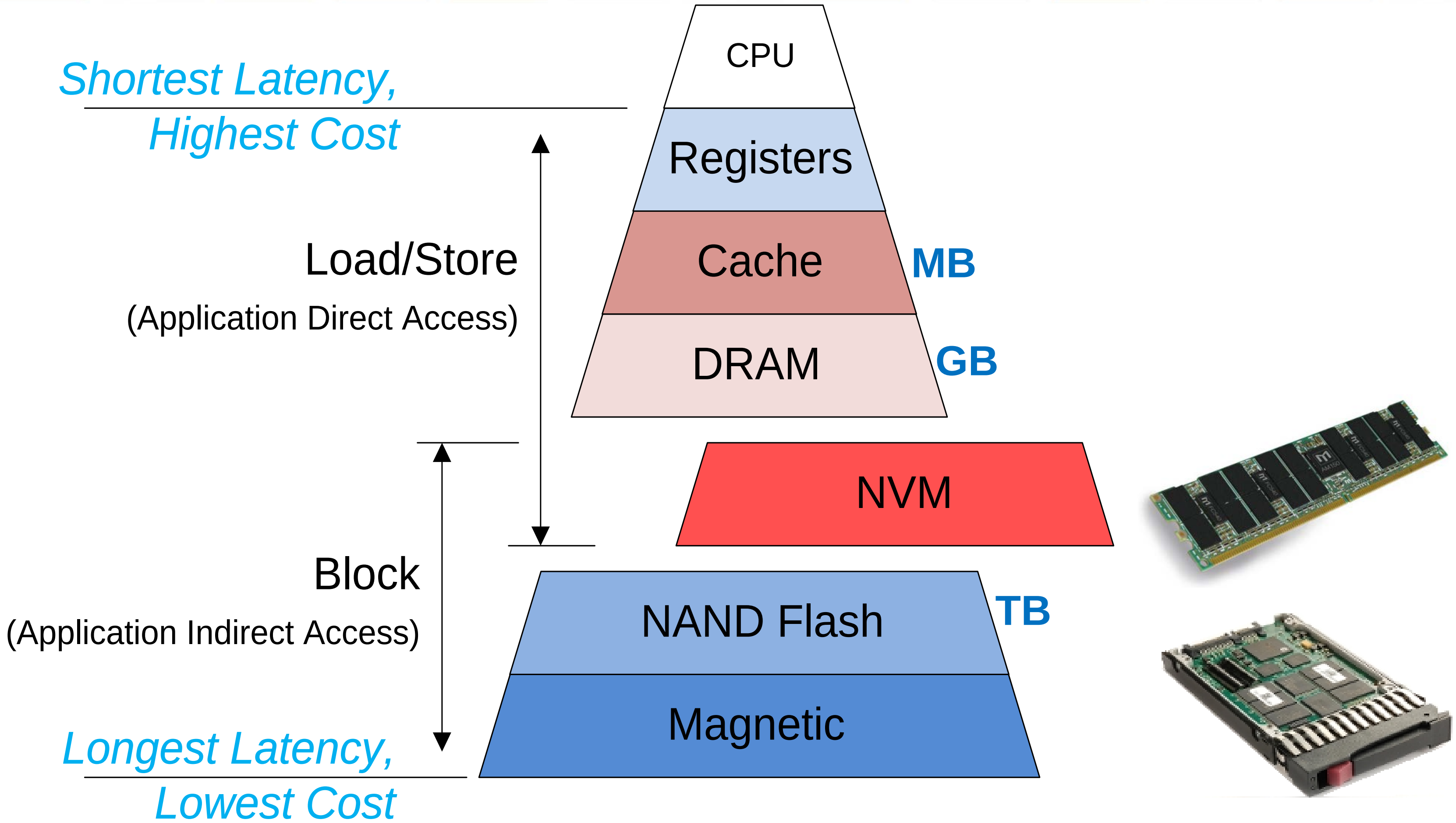
SNIA™

Solid State Storage Initiative

Agenda

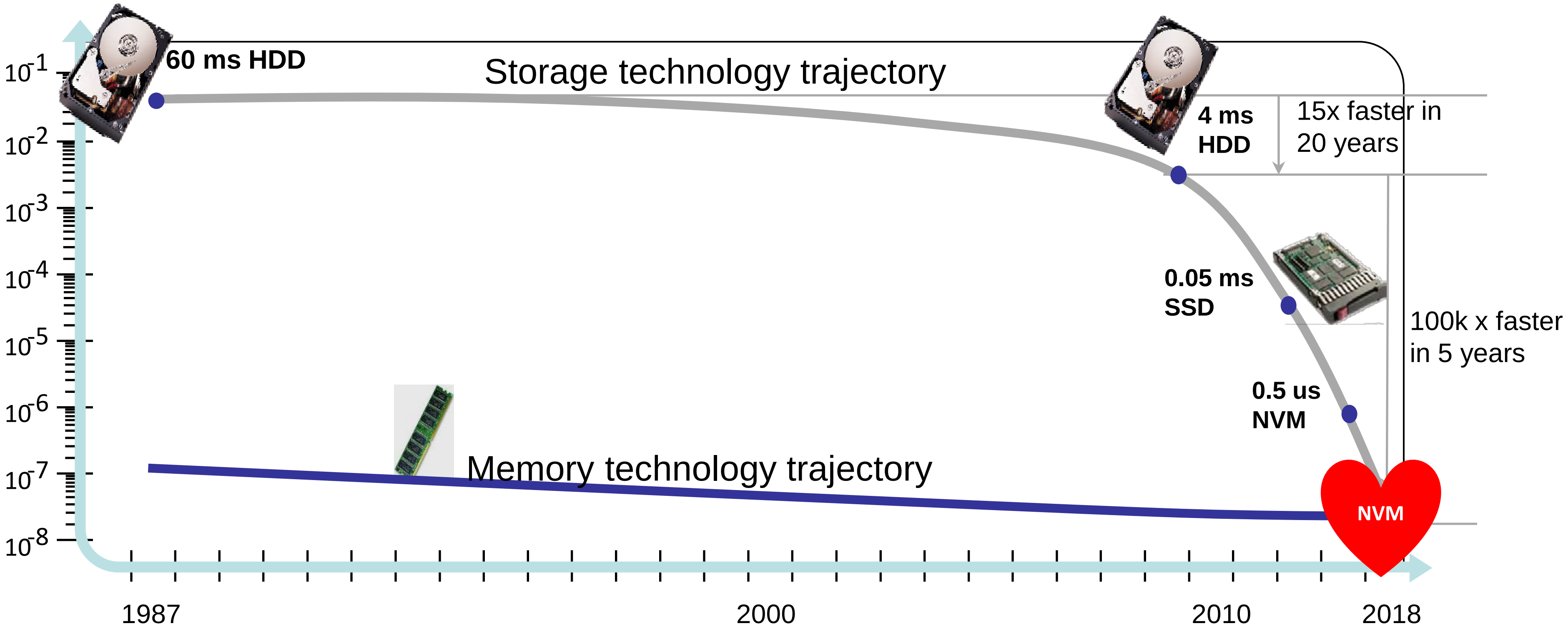
- Memory Hierarchy
- Latency trends
- NVDIMM Types
- Accelerating Applications
- The Evolution of NVDIMMs

Memory Hierarchy



Each level is an accelerator for the next lower level

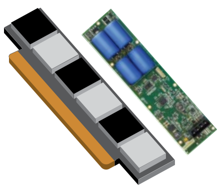
Latency Trends

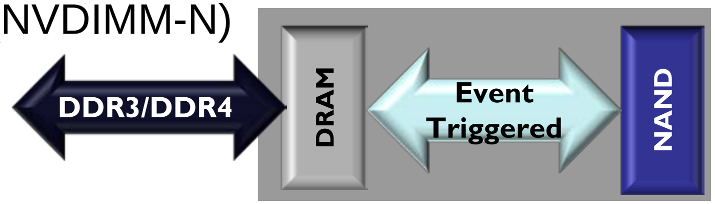


Timing is everything – Invest Now!

NVDIMM Types

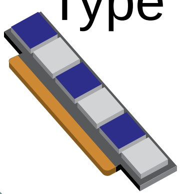
Type 1 (NVDIMM-N)

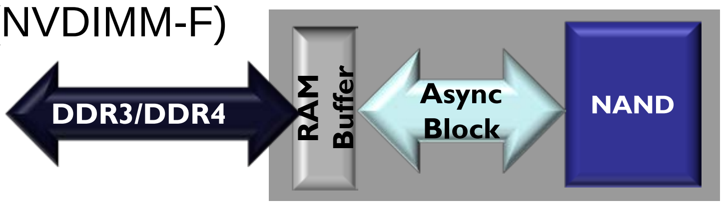




- Access Methods -> Load/Store, Emulated Block
- Capacity = DRAM (10's GB)
- Latency = DRAM (10's of nanoseconds)
- Energy source for backup

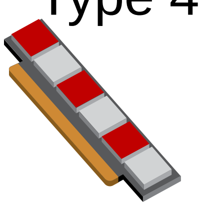
Type 3 (NVDIMM-F)

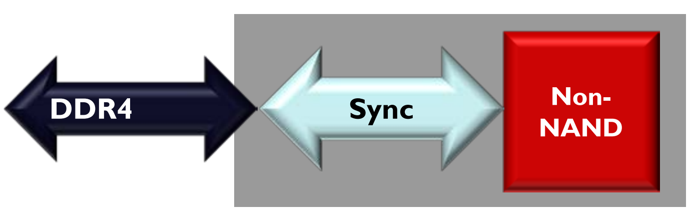




- Supported -> Emulated Block
- Capacity = NAND (100's GB – 1's TB)
- Latency = NAND (10's of microseconds)

Type 4





- Supported -> Load/Store, Emulated Block
- Capacity = NVM (100's GB – 1's TB)
- Latency = NVM (100's of nanoseconds)

Customer Value by Workload

SNIA
Solid State Storage Initiative

Low Cost → **High Capacity** → **Persistency**

In Memory Database: Journaling, reduced recovery time, Ex-large tables

Traditional Database: Log acceleration by write combining and caching

Enterprise Storage: Tiering and caching without an auxiliary power source

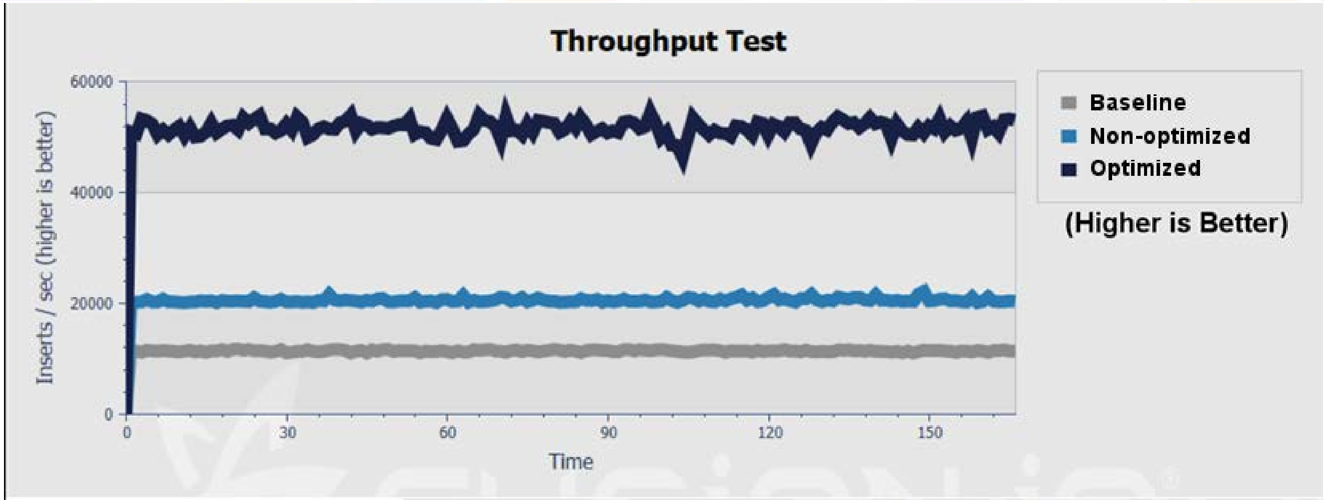
Virtualization: Higher VM consolidation with greater memory density

High-Performance Computing: Check point acceleration and/or elimination

Other**:** Object stores, unstructured data management, Genomics, archiving

NVDIMMs Accelerate a Rainbow of Applications

NVDIMM Acceleration



Baseline:

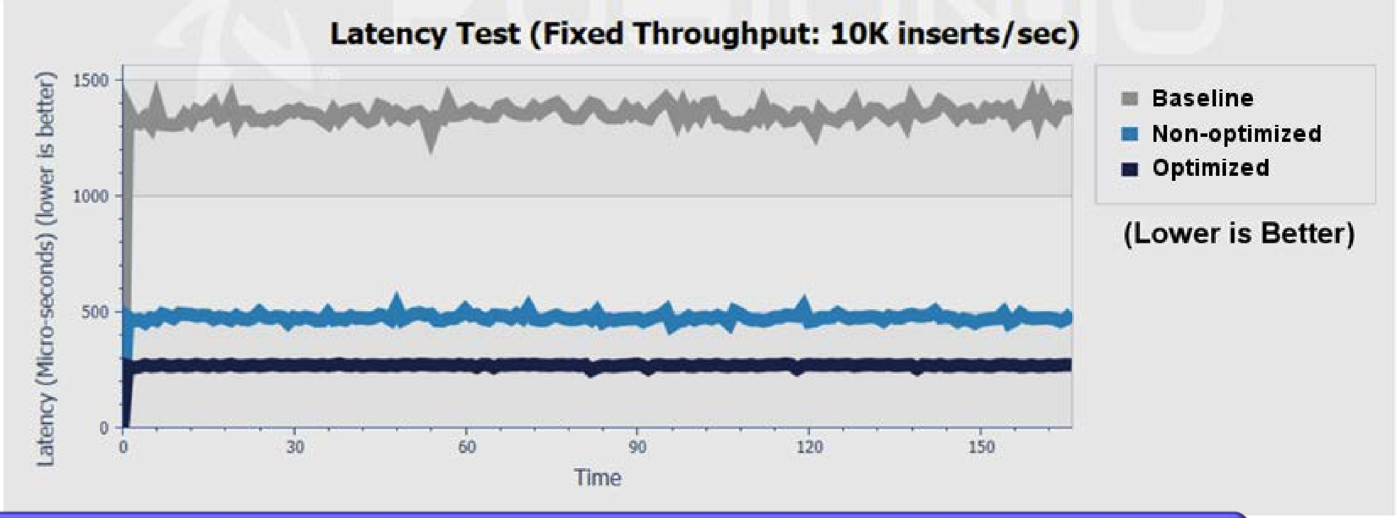
MySQL on FusionIO

Non-optimized:

Baseline + HP NVDIMM + FusionIO + standard MySQL

Optimized:

Non-optimized + Flash-Aware MySQL



4x Acceleration in Throughput and Latency

The Evolution of NVDIMMs

- Origin as accelerator for cache on storage array controllers
  - Battery backed
- Type-1 accelerating database and HPC
  - Eliminating bottlenecks in journaling
- Type-3 accelerating High Frequency Trading
  - High IOPs
- Type-4 accelerating search, data manipulation, initialization
  - MicroController on NVDIMM
- And Beyond! FPGA on NVDIMM
  - Programmable on-the-fly
  - Accelerate dynamically changing workloads

NVDIMMs Accelerating Applications in the Future

Thank You!

RAM DISK Block IO Performance

Measured to the SNIA PTS

NVM Summit
January 20, 2015
St. Claire Hotel, San Jose CA

Eden Kim, CEO

Calypso Systems, Inc.

CALYPSO Systems

RAM Disk Block IO Performance

Non Volatile Memory – being advanced by NVMP TWG / NVDIMM SIG

- Memory Mapped Load/Store
- NVDIMM Block IO Type N / Type F
- Faster than NAND Flash Based SSDs

In Memory Block IO Storage – data on a Linux RAM Disk

- RAM Disk performance shows top end of Type N NVDIMM
- RAM Disk Block IO performance is sensitive to settings
- NVDIMM Block IO may approach RAM Disk performance

Linux RAM Disk is compared to Traditional Classes of NAND Flash SSD

Test Set Up

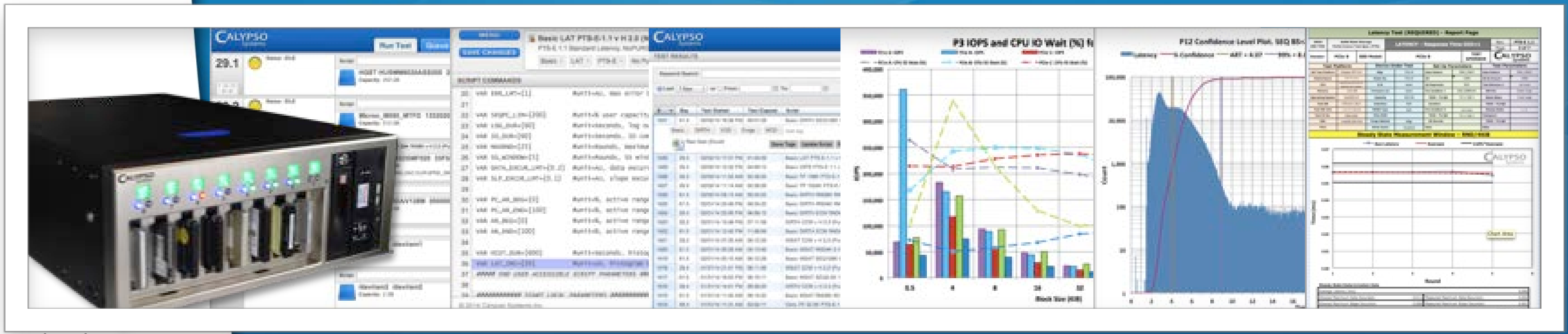
Hardware: PTS Reference Test Platform

- Intel S2600 COE Gen 3
- Dual Xeon 8 core, 3.2Ghz E5 2687W
- 32 GB (4GB x 8) DDR3 1600 ECC

Software:

- OS – CentOS 6.5
- Linux RAM Disk Block IO Driver 2.6.32-431.11.2.EL6.x86_64
- Test Software - Calypso CTS BE 1.9.216-eL6

Test Methodology: SNIA SSS Performance Test Specification v 1.1.1



Pick up a copy of this chart

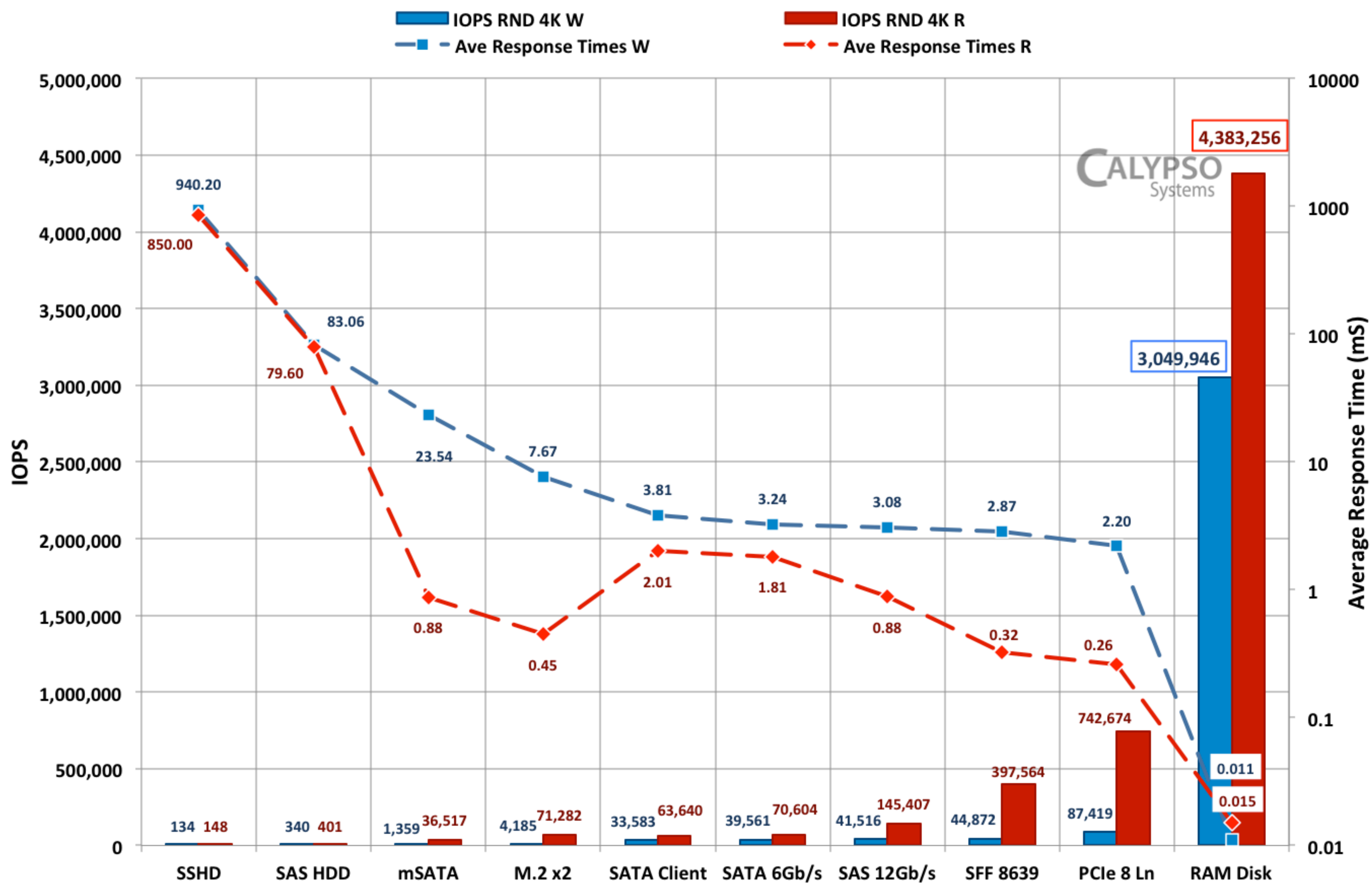
in Review 2014 - Summary Performance Comparison by Storage Class

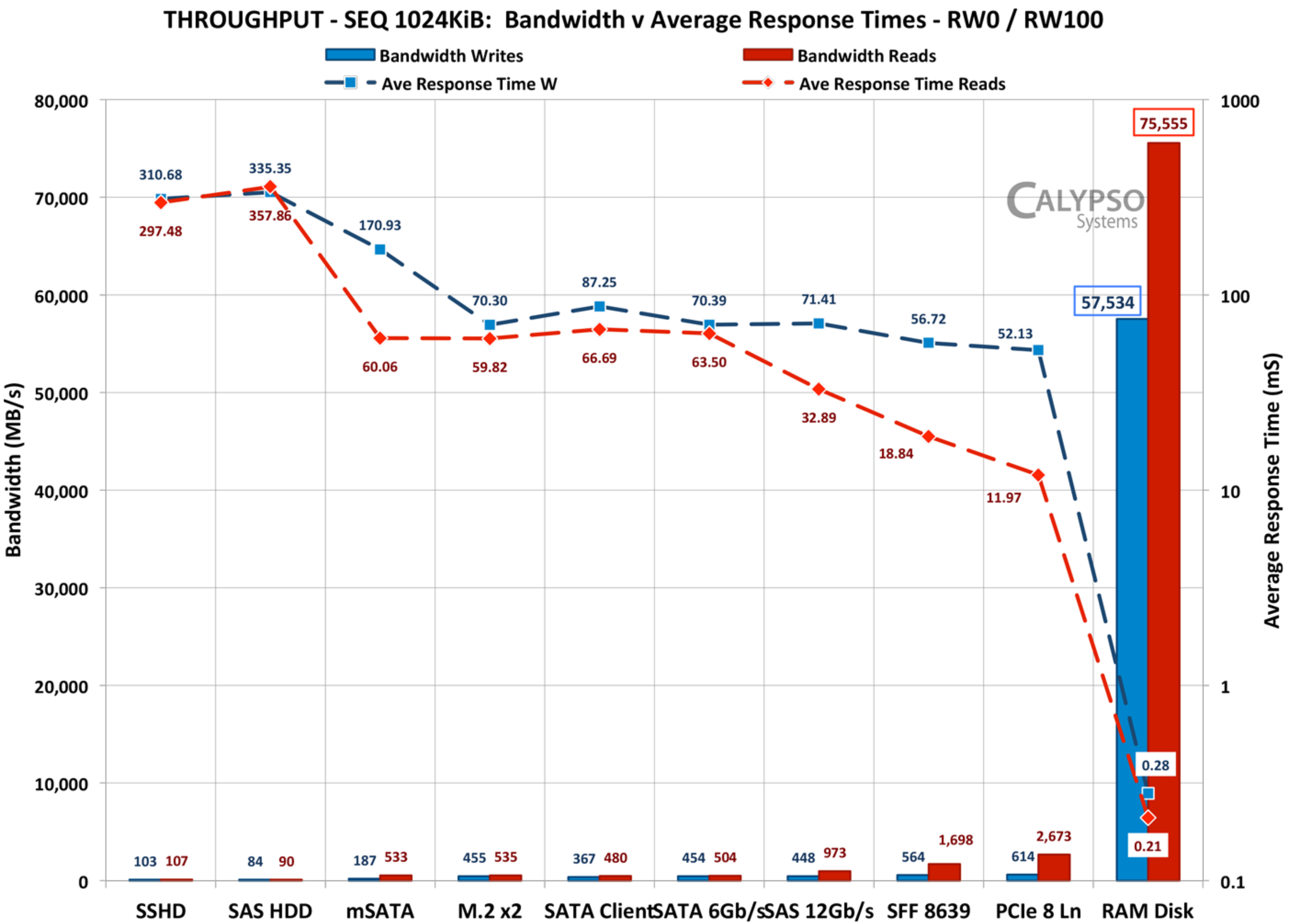
Storage Class				IOPS FOB PTS WSAT - T4Q32	IOPS Steady State PTS IOPS - T2Q16 / T4Q32			Bandwidth PTS Throughput - T1Q32		Response Time PTS Latency - T1Q1	
	Category	Device Type	Capacity	RND 4KiB 100% W	RND 4KiB 100% W	RND 4KiB 65:35 RW	RND 4KiB 100% R	SEQ 1024KiB 100% W	SEQ 1024KiB 100% R	RND 4KiB 100% W Ave	RND 4KiB 100% W Max
HDD & SSHD											
1	SSHD	7,200 RPM 2.5" SATA Hybrid	500 GB	134	134	131	148	107 MB/s	103 MB/s	18.54 mSec	40.63 mSec
2	SAS HDD	15,000 RPM 3.5" SAS HDD	80 GB	350	340	398	401	84 MB/s	90 MB/s	55.39 mSec	97.28 mSec
CLIENT SSDs											
3	mSATA	mSATA 1.8" MLC	128 GB	45,743	1,359	1,926	36,517	187 MB/s	533 MB/s	0.74 mSec	543.41 mSec
4	M.2 x2	M.2 x2 2280 MLC	512 GB	61,506	4,185	9,532	71,282	455 MB/s	535 MB/s	0.29 mSec	24.99 mSec
5	SATA Client	SATAIII 2.5" MLC	200 GB	54,788	33,583	50,708	63,640	367 MB/s	480 MB/s	0.06 mSec	11.95 mSec
ENTERPRISE SSDs											
6	SATA 6Gb/s	SATA 6Gb/s 2.5" eMLC	800 GB	57,422	39,561	46,072	70,604	454 MB/s	504 MB/s	0.05 mSec	0.22 mSec
7	SAS 12Gb/s	SAS 12Gb/s 2.5" MLC	800 GB	97,950	41,516	72,342	145,407	448 MB/s	973 MB/s	0.05 mSec	11.84 mSec
8	SFF 8639	SFF 8639 4 lane 2.5" MLC	700 GB	149,512	44,872	166,002	397,564	564 MB/s	1,698 MB/s	0.01 mSec	0.38 mSec
9	PCIe 8 Lane	PCIe 8 Lane Edge Card MLC	1400 GB	159,926	87,419	236,227	742,674	614 MB/s	2,673 MB/s	0.01 mSec	0.56 mSec

All measurements taken on the RTP 3.0 CTS 6.5 Reference Test Platform pursuant to the SNIA PTS-E 1.1.
NOTE: Thread and Queue settings for PTS IOPS are T2Q16 for HDD/SSHD & Client SSDs and T4Q32 for Enterprise SSDs.

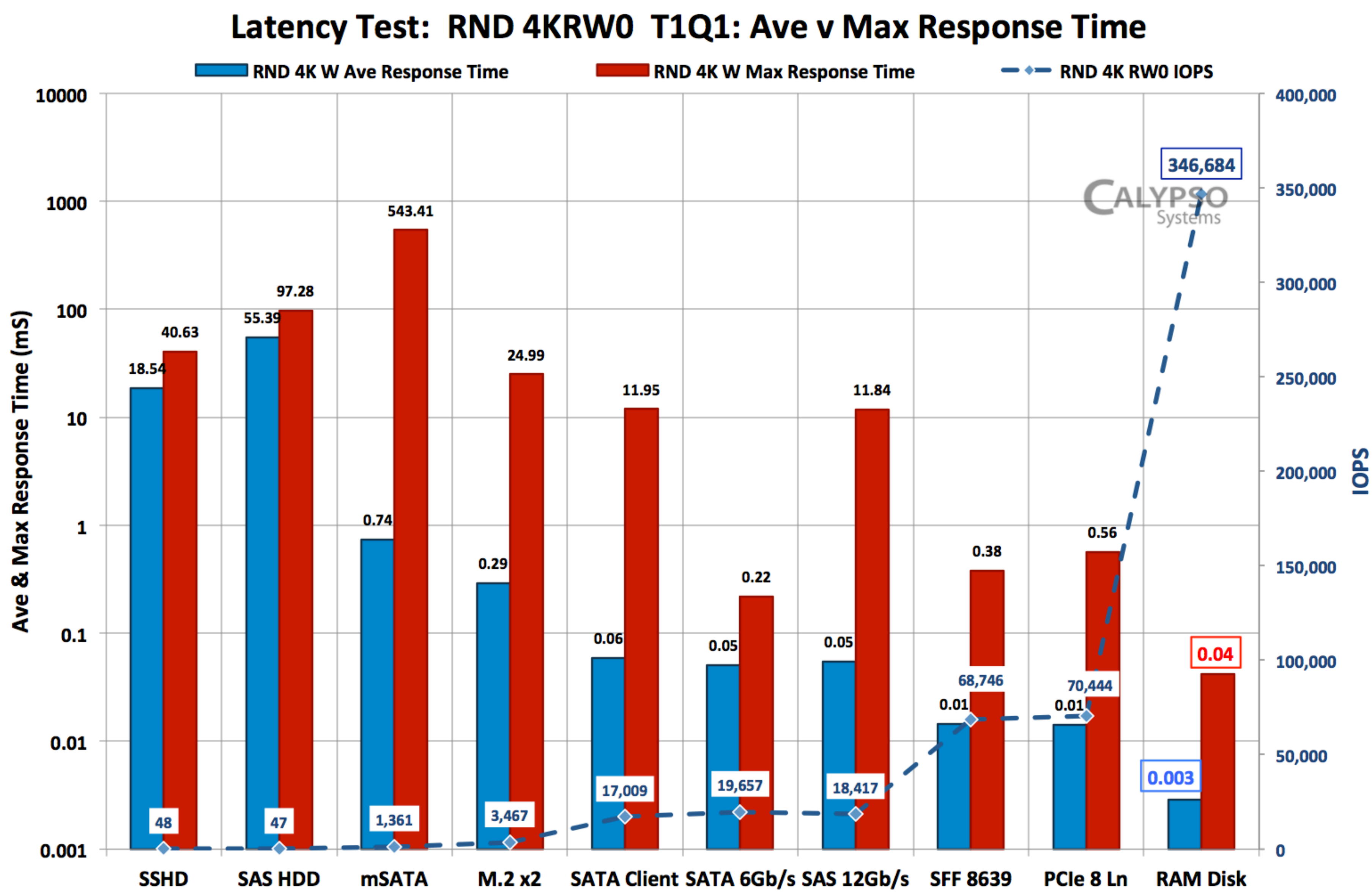
CALYPSO Systems

IOPS v Average Response Times for RND 4K RW0 / RW100





CALYPSO Systems



Calypso Systems

RAM Disk Application Demand Intensity

Effects on Performance Measurement

- NAND Flash: Pre Conditioning, Steady State, Write History
- RAM Disk: Outstanding IO, CPU cores, Memory Channels

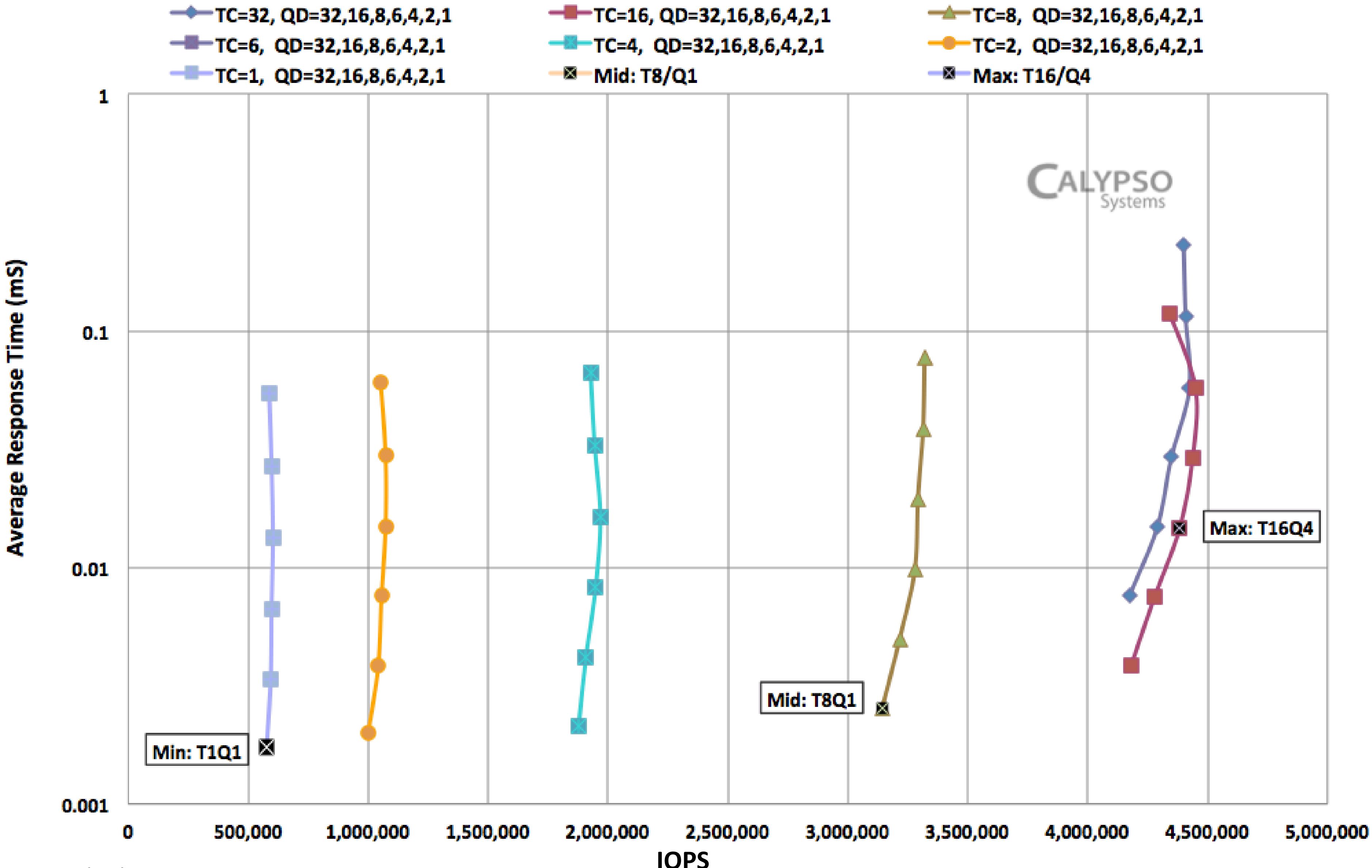
Mapping RAM disks for Outstanding IOs (OIO) by Application Workload

- Vary Thread Count and Queue Depth
- Measure OIO by Response Time and IOPS (Demand Intensity)

Confidence Level Plot Compare (CLPC)

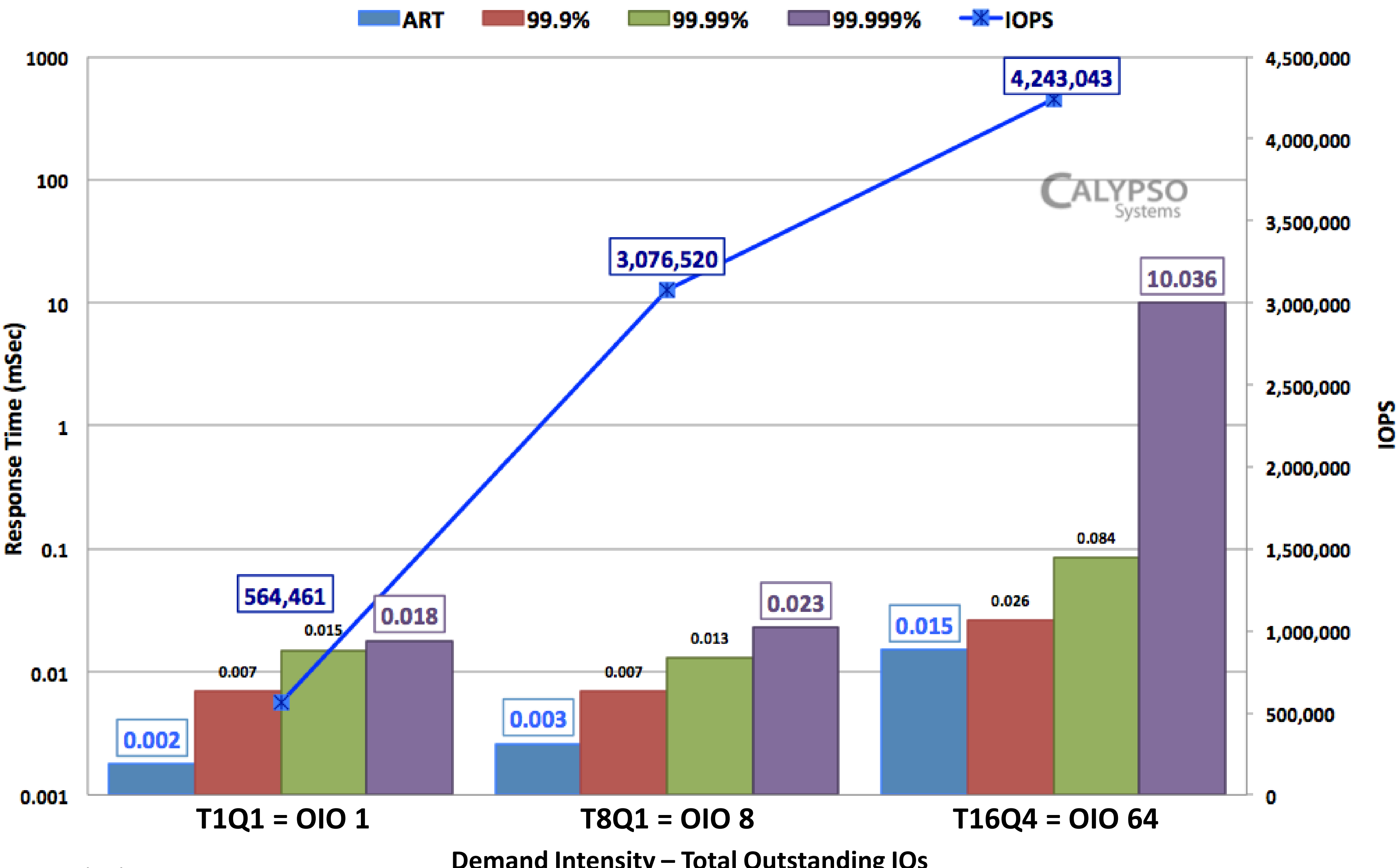
- Optimal OIO – Performance at different Thread Count x Queue Depth
- IOPS
- Response Time Quality of Service (QoS) – “5 9s” percentile response times
- Response Time Ceiling – Maximum response time allowed by application

Demand Intensity – RAM Disk: OIO x ART & IOPS



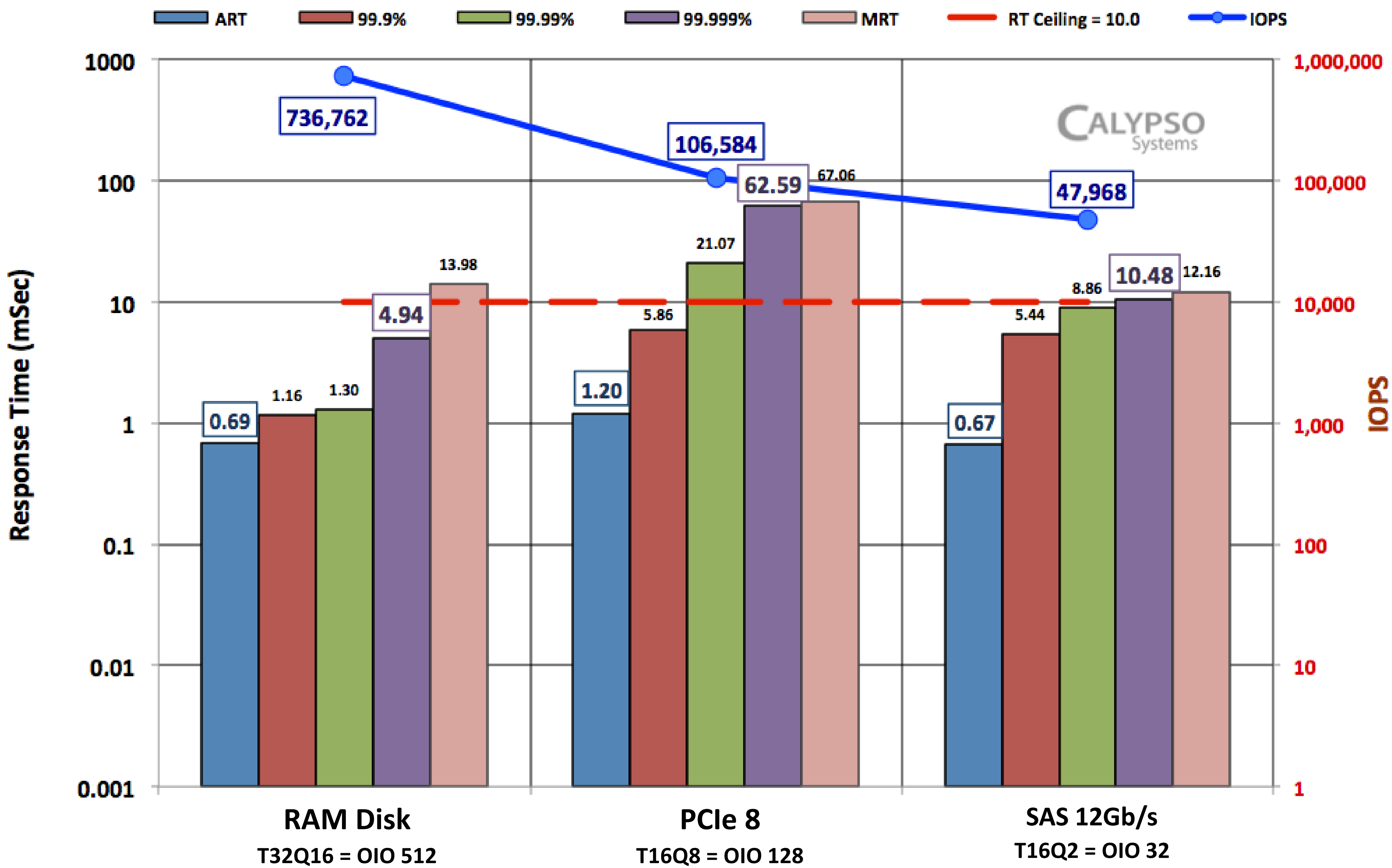
CALYPSO Systems

CLPC 'QoS' for a Single RAM Disk – RND 4KIB 100% Reads



CALYPSO Systems

Comparison db OLTP: RAM Disk – PCIe x8 – SAS 12Gb/s



CALYPSO Systems

TAKE AWAYS

RAM Disk Block IO Performance is much higher than NAND Flash

NVDIMM RAM Block IO can approach the level of RAM Disk Block IO

RAM Disk Block IO Performance Depends on Settings

Applications can run much faster with RAM Disk and/or NVDIMM SSD

RAM Disk / NVDIMM SSD offer new Storage Tiering Opportunities

감사합니다 Natick

Grazie Danke Ευχαριστίες Dalu Obrigado

Thank You Köszönöm

Tack

Спасибо Dank Gracias

谢谢 Merci Seé

ありがとう

For more information, contact Calypso Systems, Inc.

info@calypsotesters.com www.calypsotesters.com

Calypso Systems, Inc.

2255 Martin Ave., ste H

Santa Clara, CA 95050

408.982.9955

info@calypsotesters.com

Calypso Systems

NVDIMMS in Enterprise Storage Arrays drive performance

Tom McKnight, Vice President of Hardware Platform @ Nimble Storage

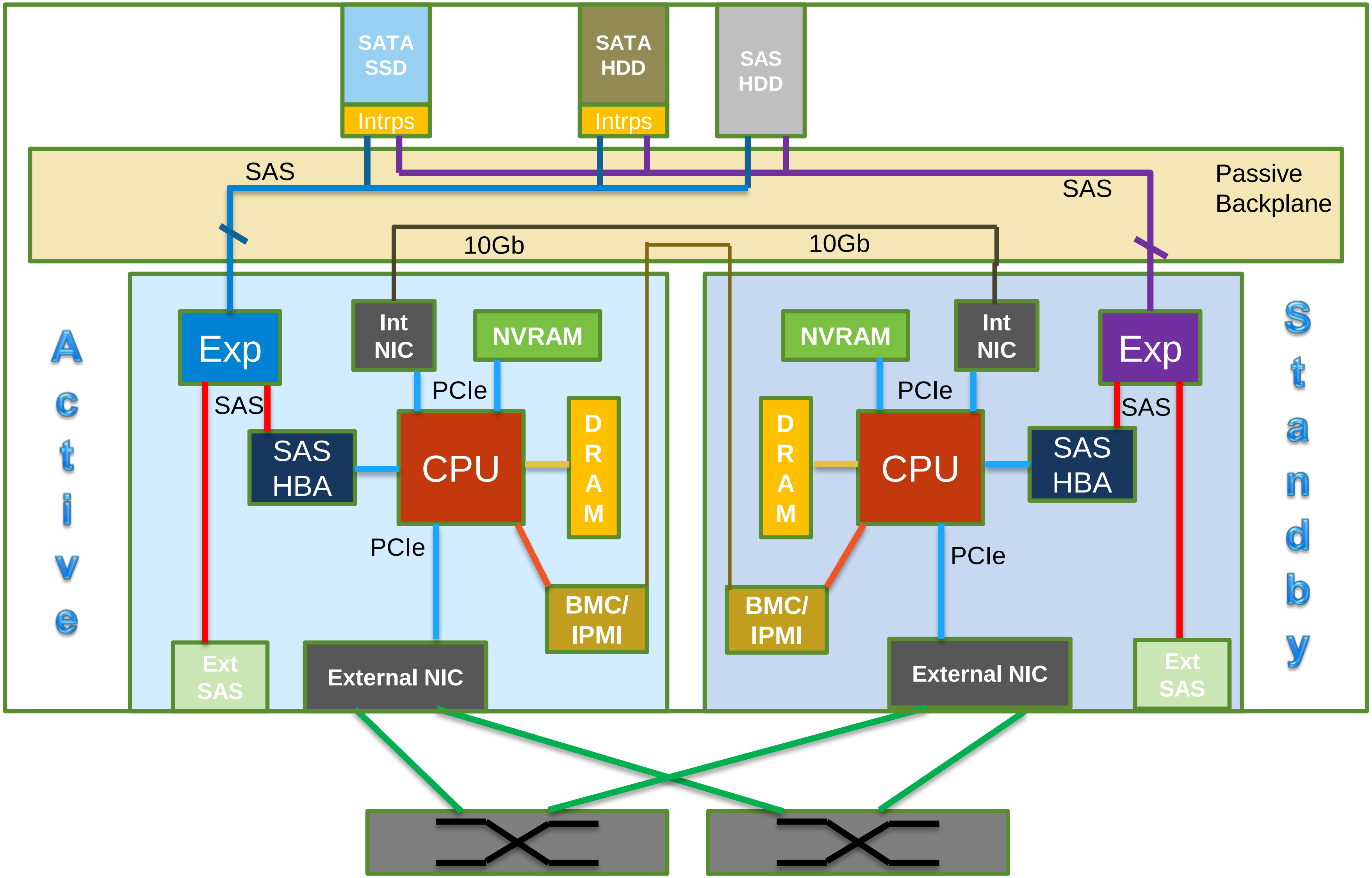
Preface

- NVDIMMs combined with PCIe NTB's have enabled Integrated Enterprise Storage Platforms to achieve significant performance improvements *(> 4X Write IOP latency improvement !!)*

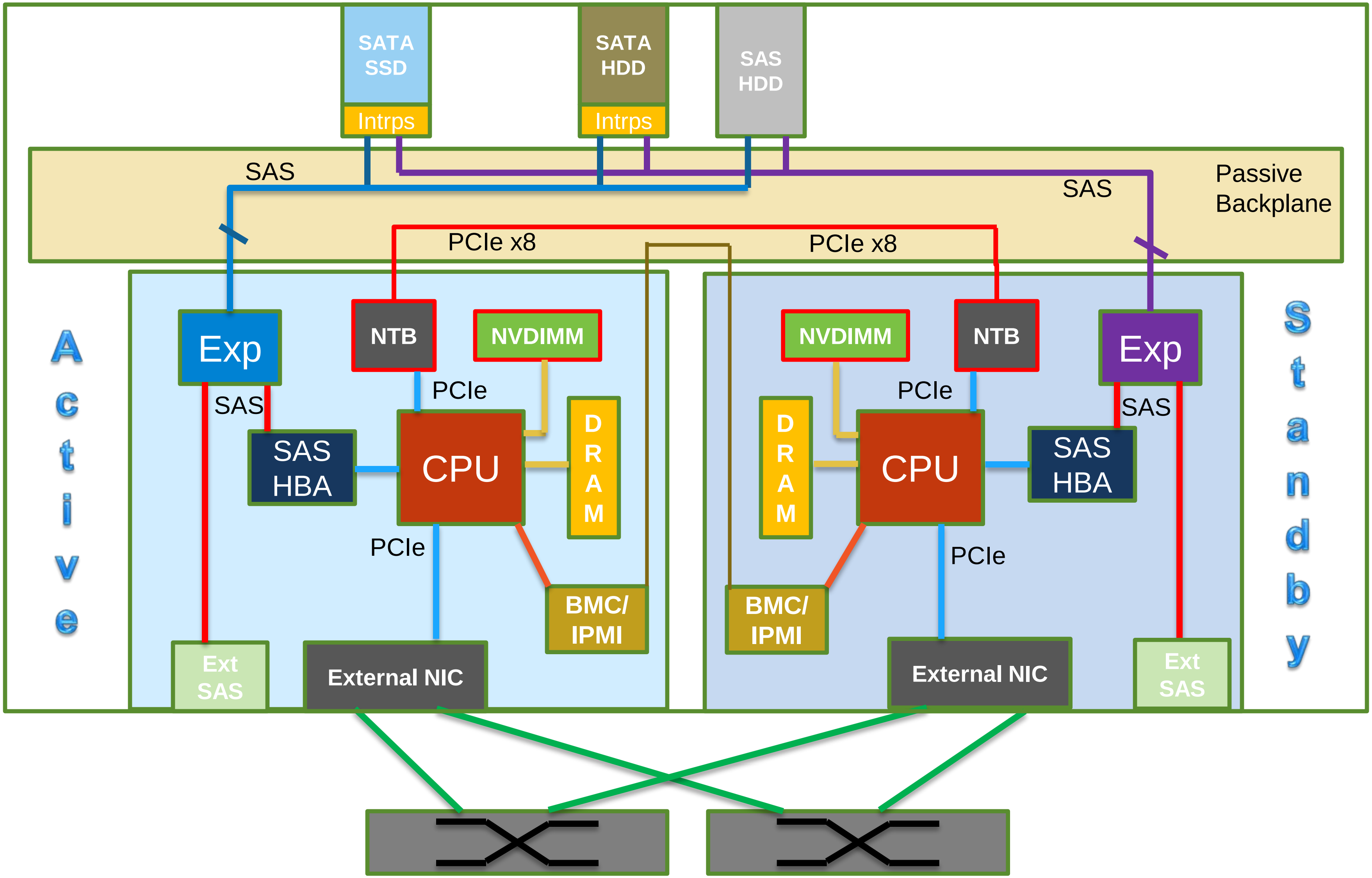
- Background:
  - An Integrated Enterprise / Fault Tolerant Storage Platform that implements software based RAID must preserve transient write data in the event of a power failure or hardware component failure. Data loss is NOT acceptable.
  - To insure fault tolerance the transient data must be replicated to the peer / standby controller before the write operation can be acknowledged. This inherent mirroring latency will heavily impact the systems maximum write performance (typically measured in IOPS).

Lets compare the write performance of a legacy storage architecture against a new NVDIMM enabled architecture!

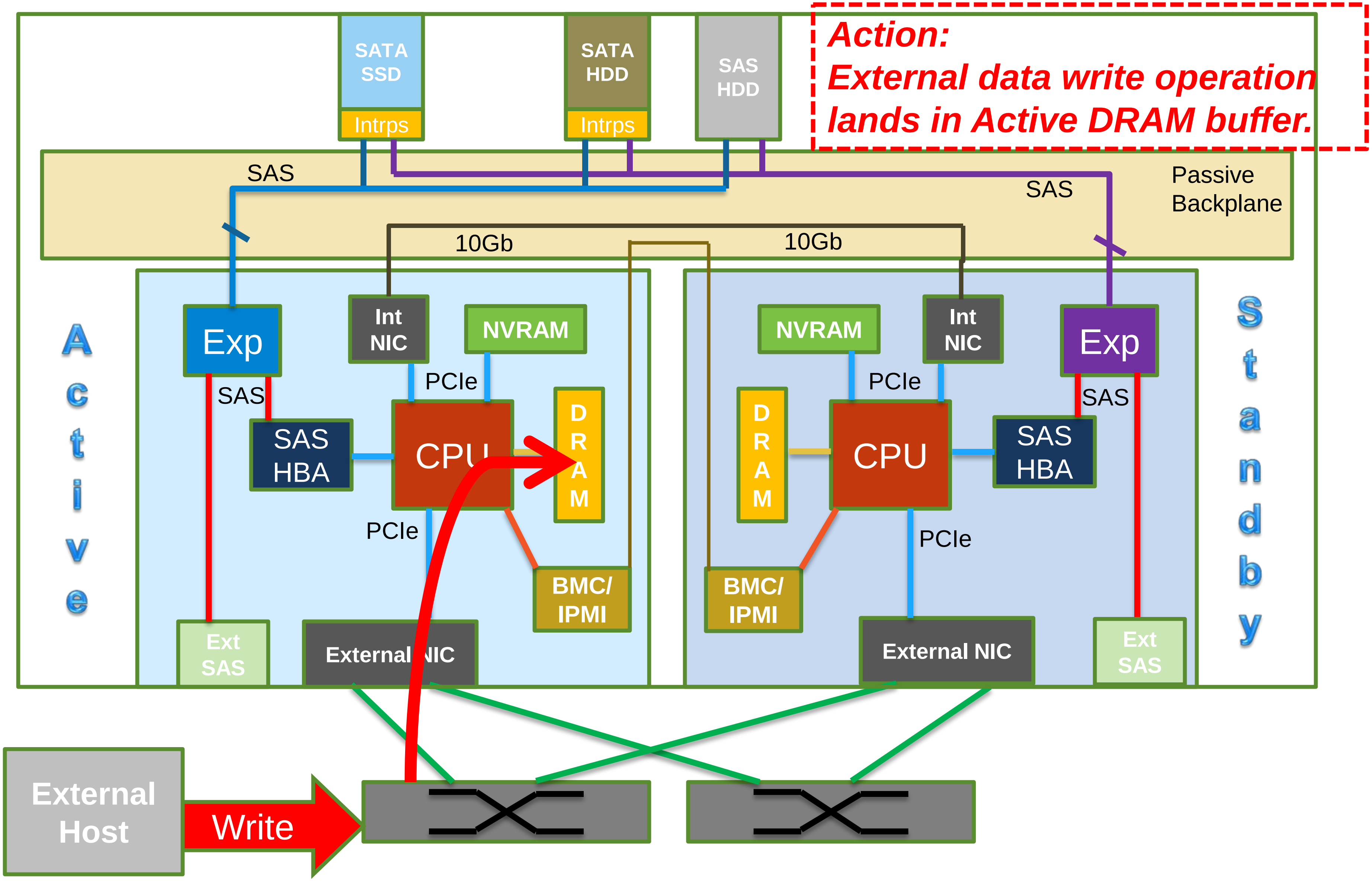
Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards



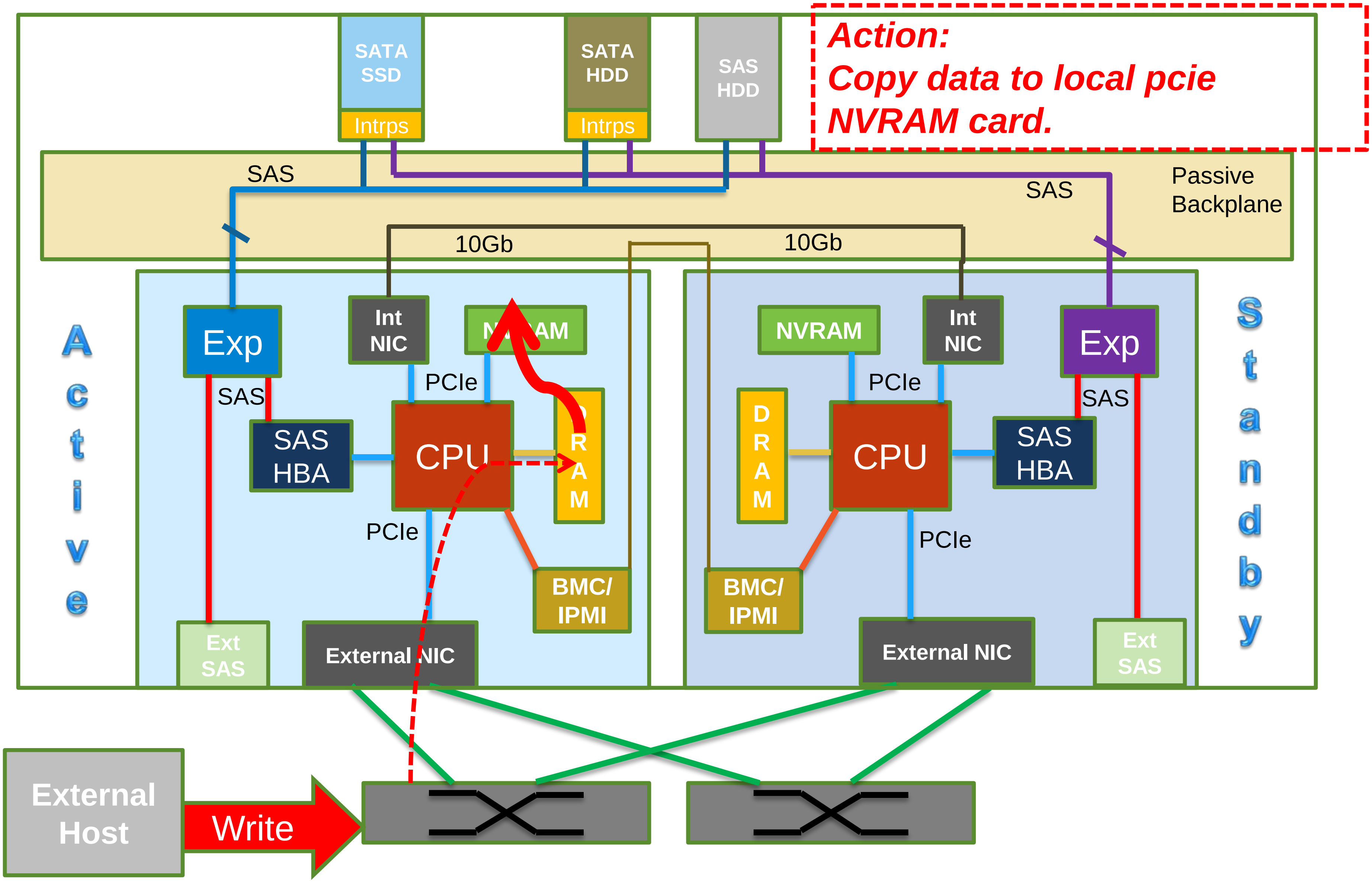
Next Generation Integrated Enterprise Storage Array – Mirror NVData to / from NVDIMM via PCIe NTB



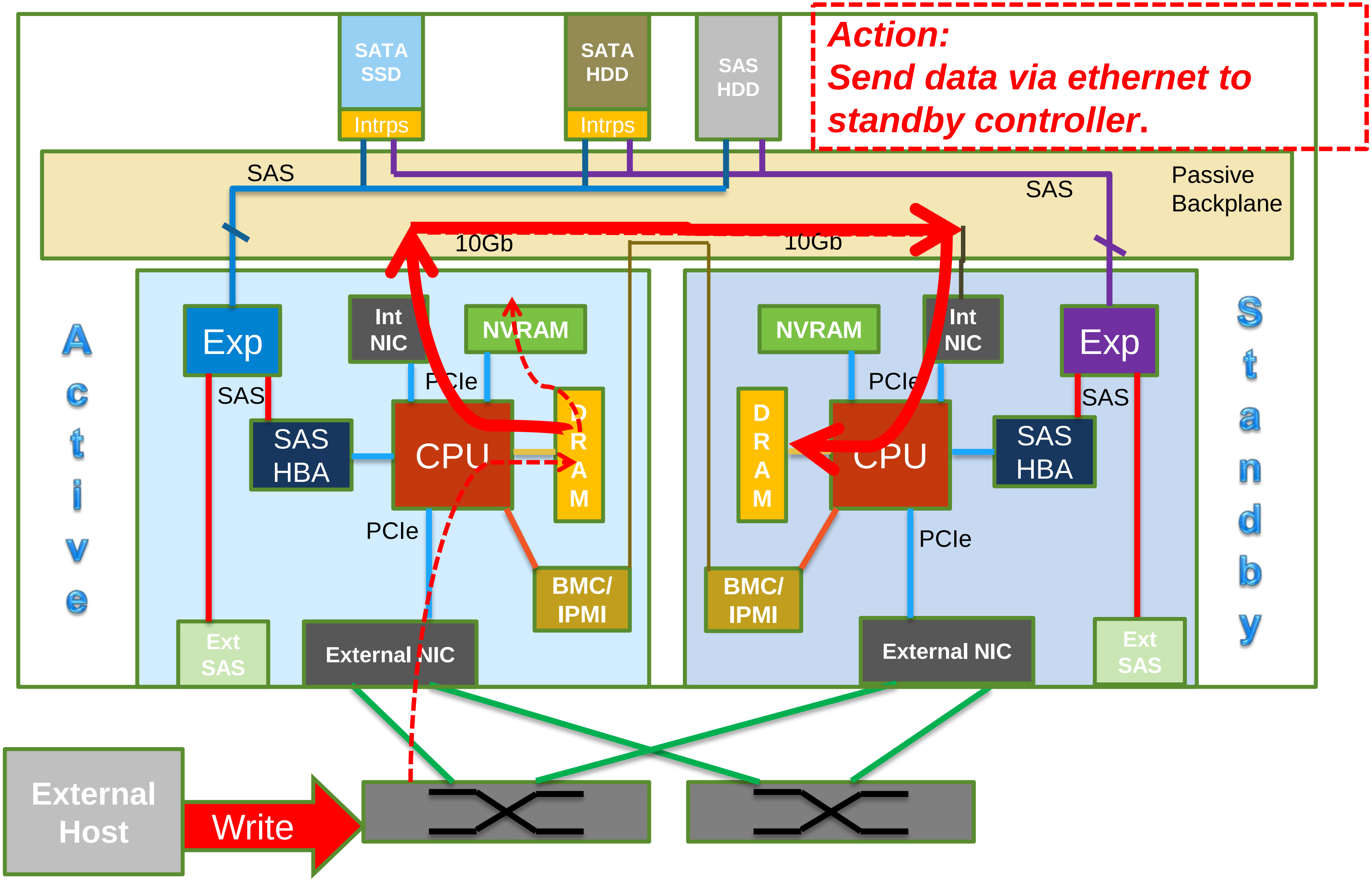
Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards



Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards

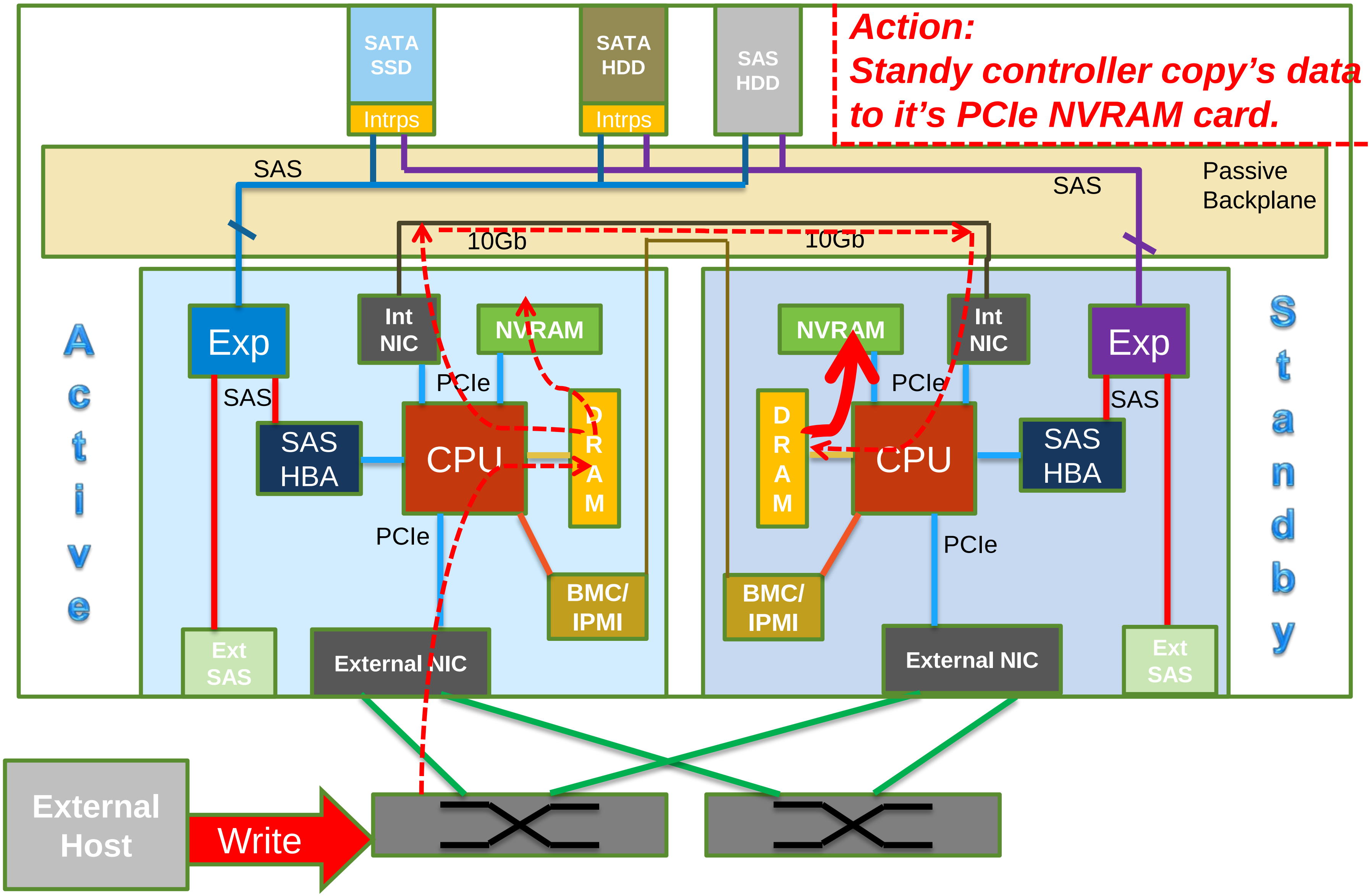


Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards

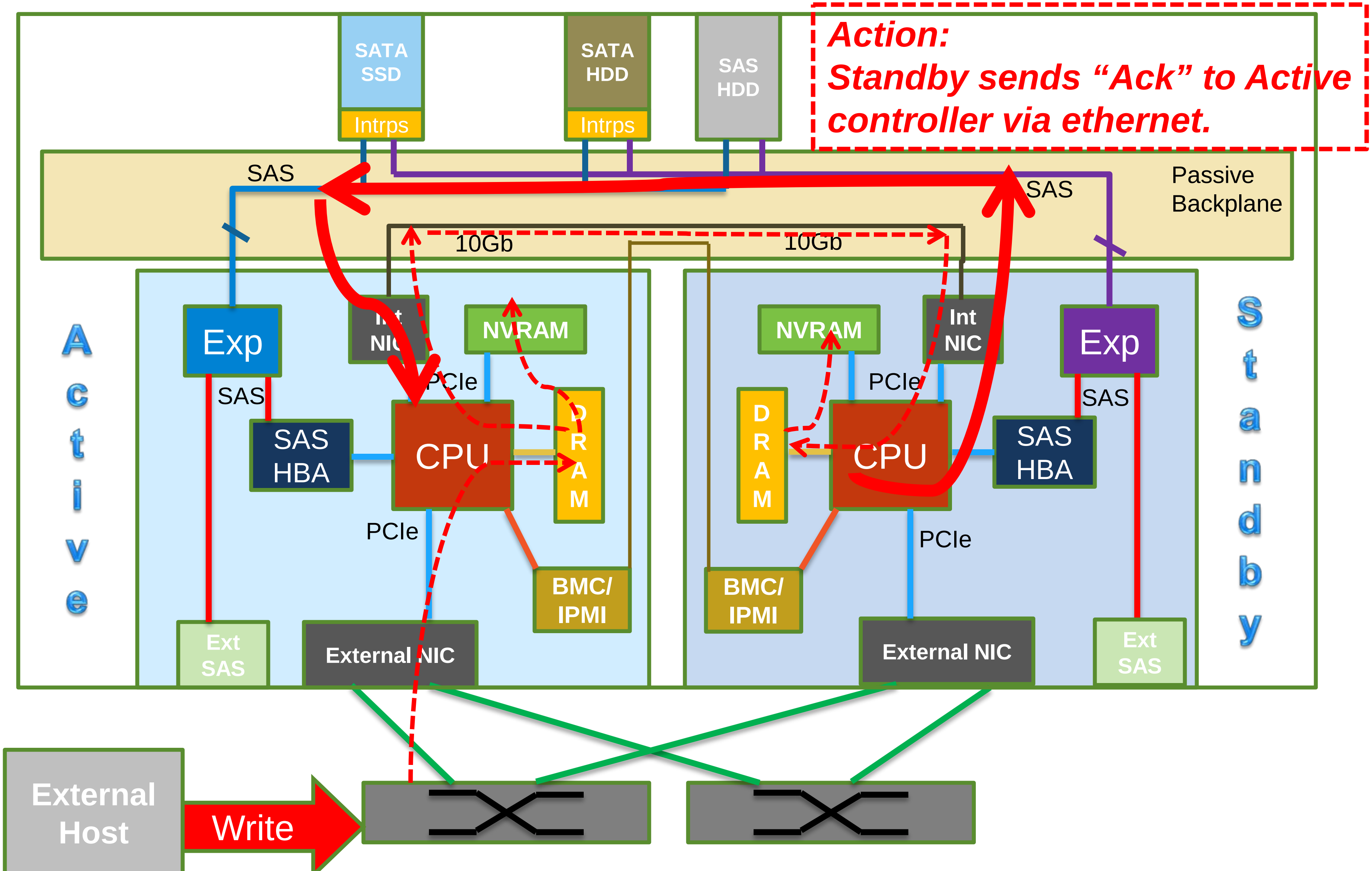


Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards

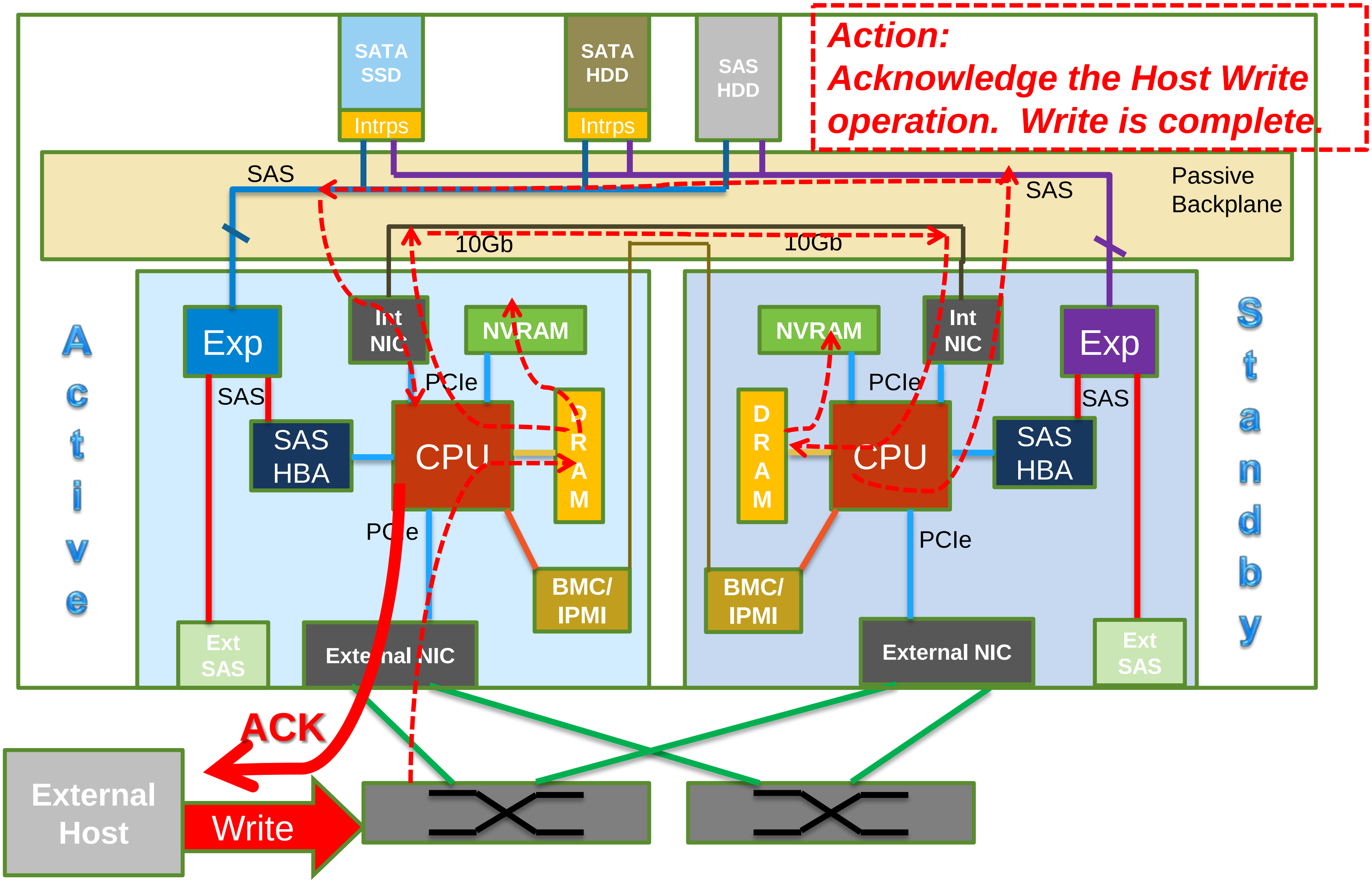
Action:
Standy controller copy's data to it's PCIe NVRAM card.



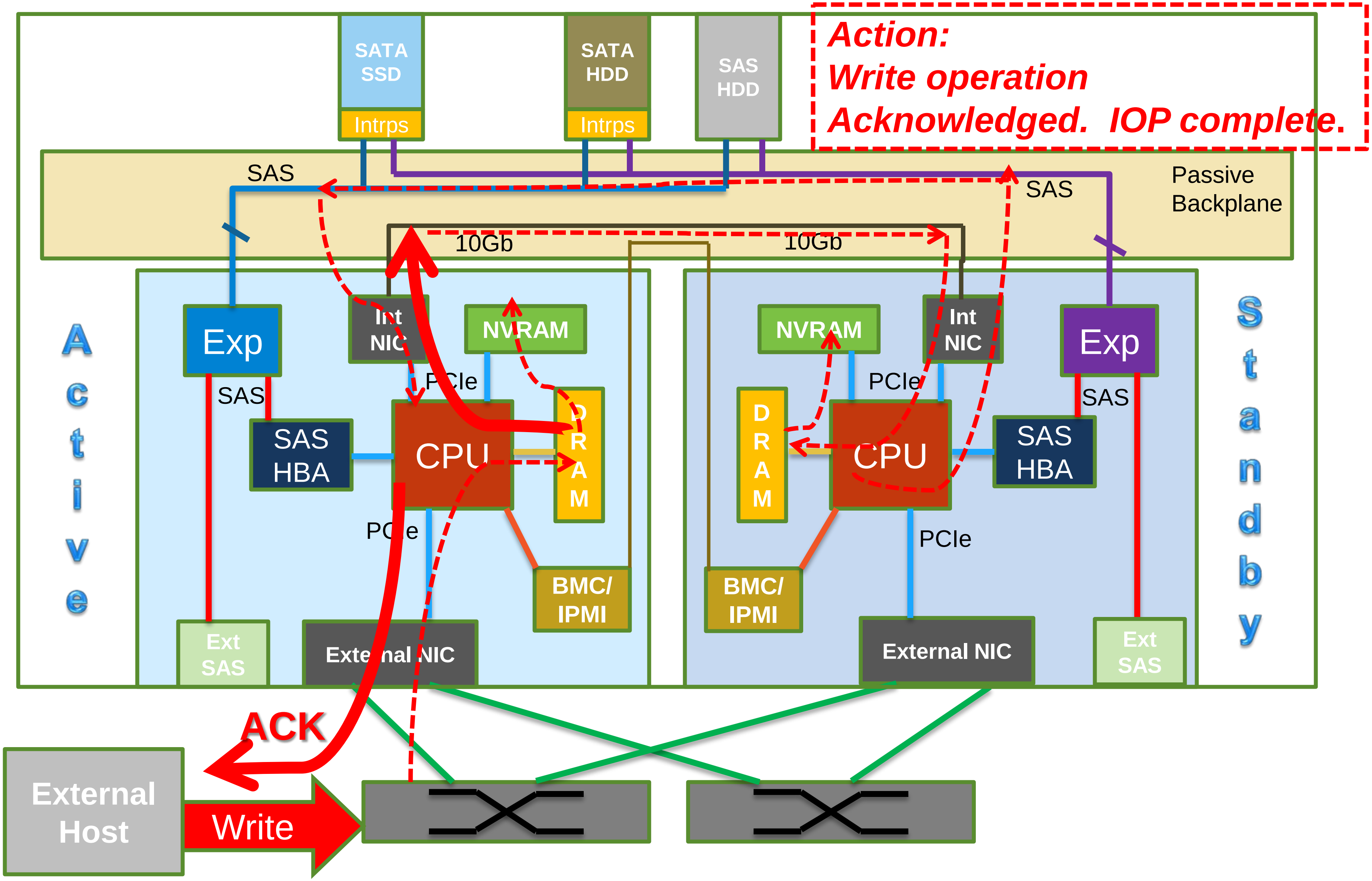
Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards



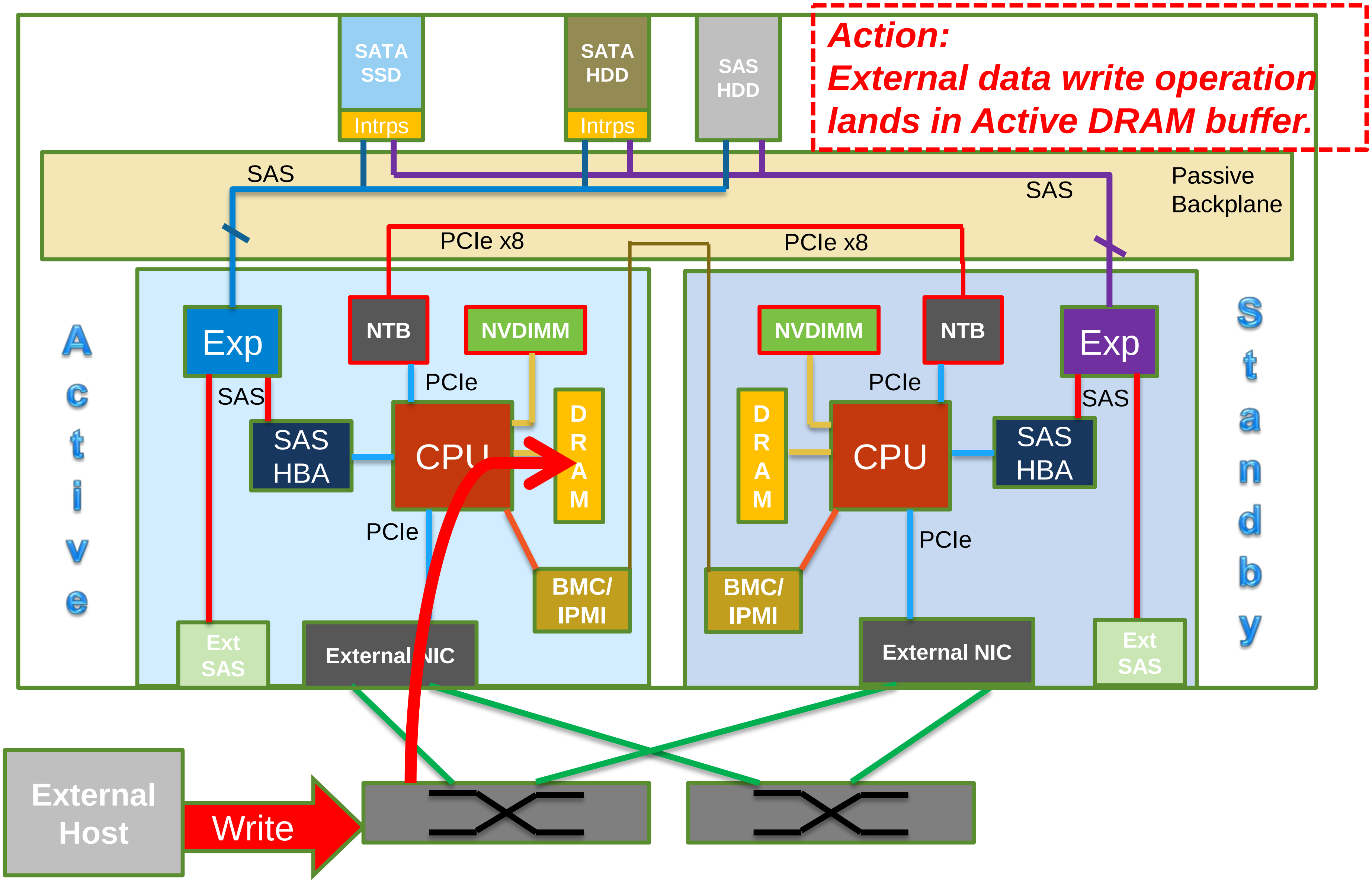
Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards



Legacy Integrated Enterprise Storage Array – Mirror NVData with Ethernet & PCIe NVRAM Cards



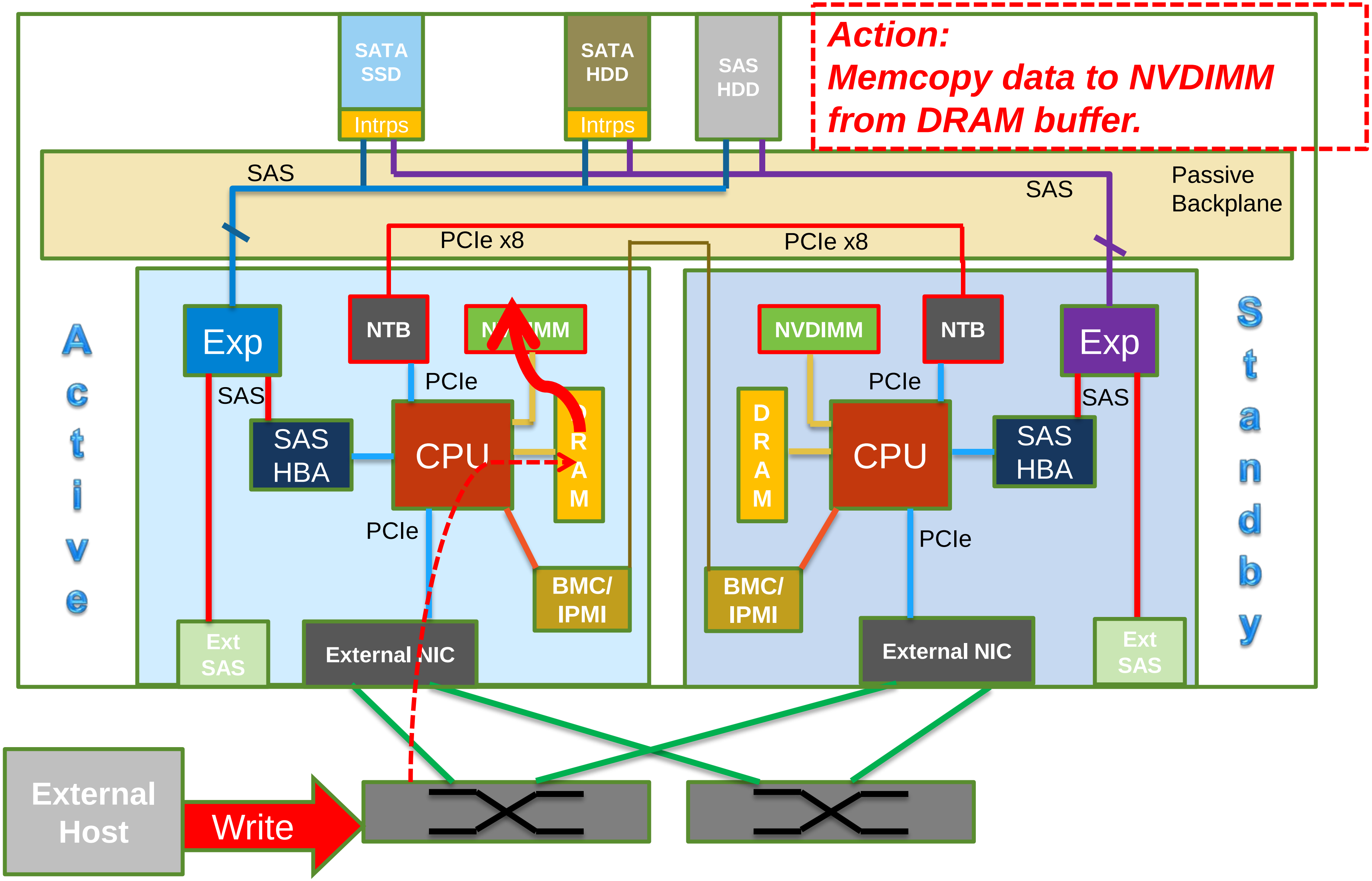
Next Generation Integrated Enterprise Storage Array – Mirror NVData to / from NVDIMM via PCIe NTB



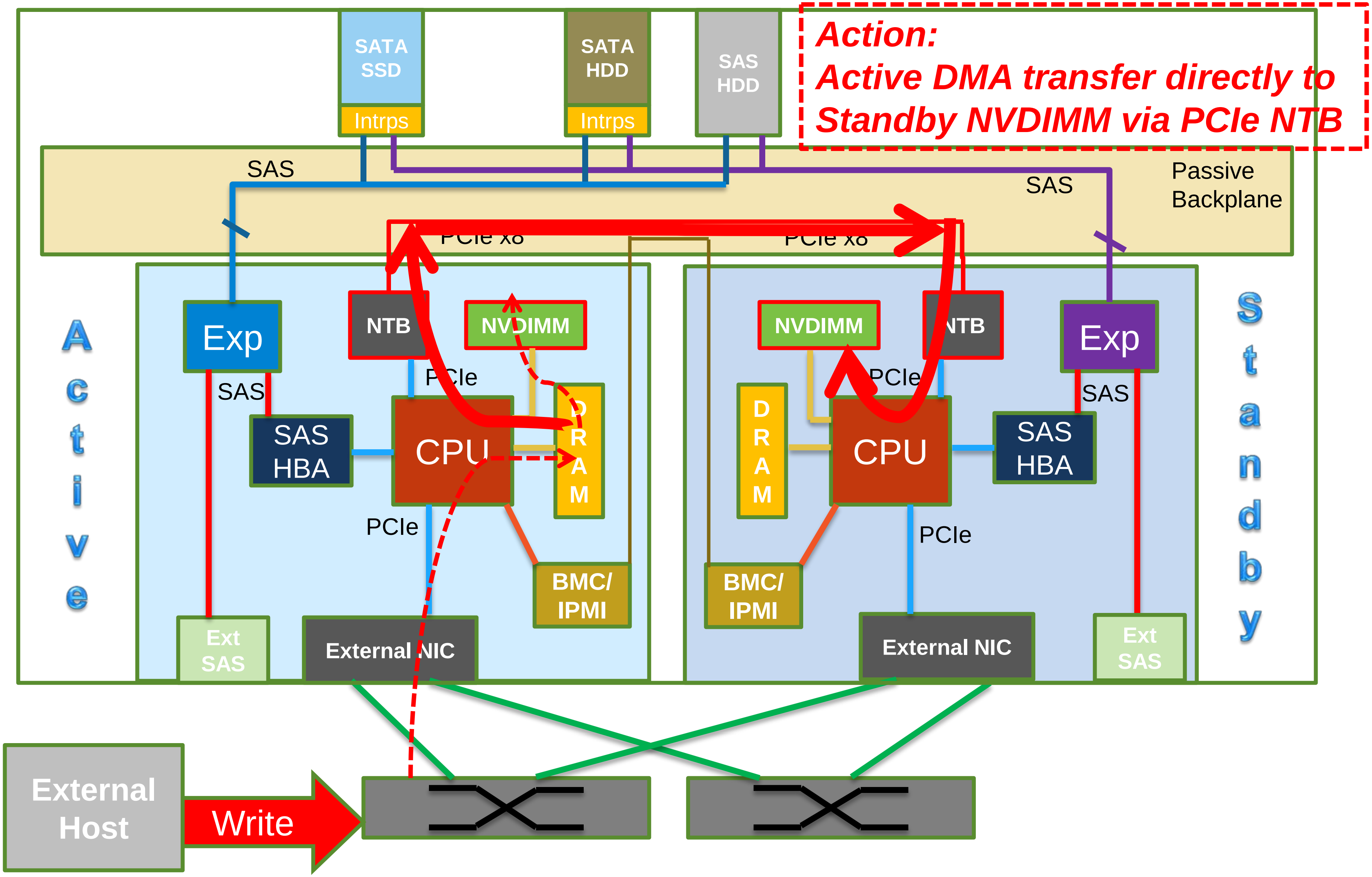
Next Generation Integrated Enterprise Storage Array – Mirror NVData to / from NVDIMM via PCIe NTB

nimblestorage

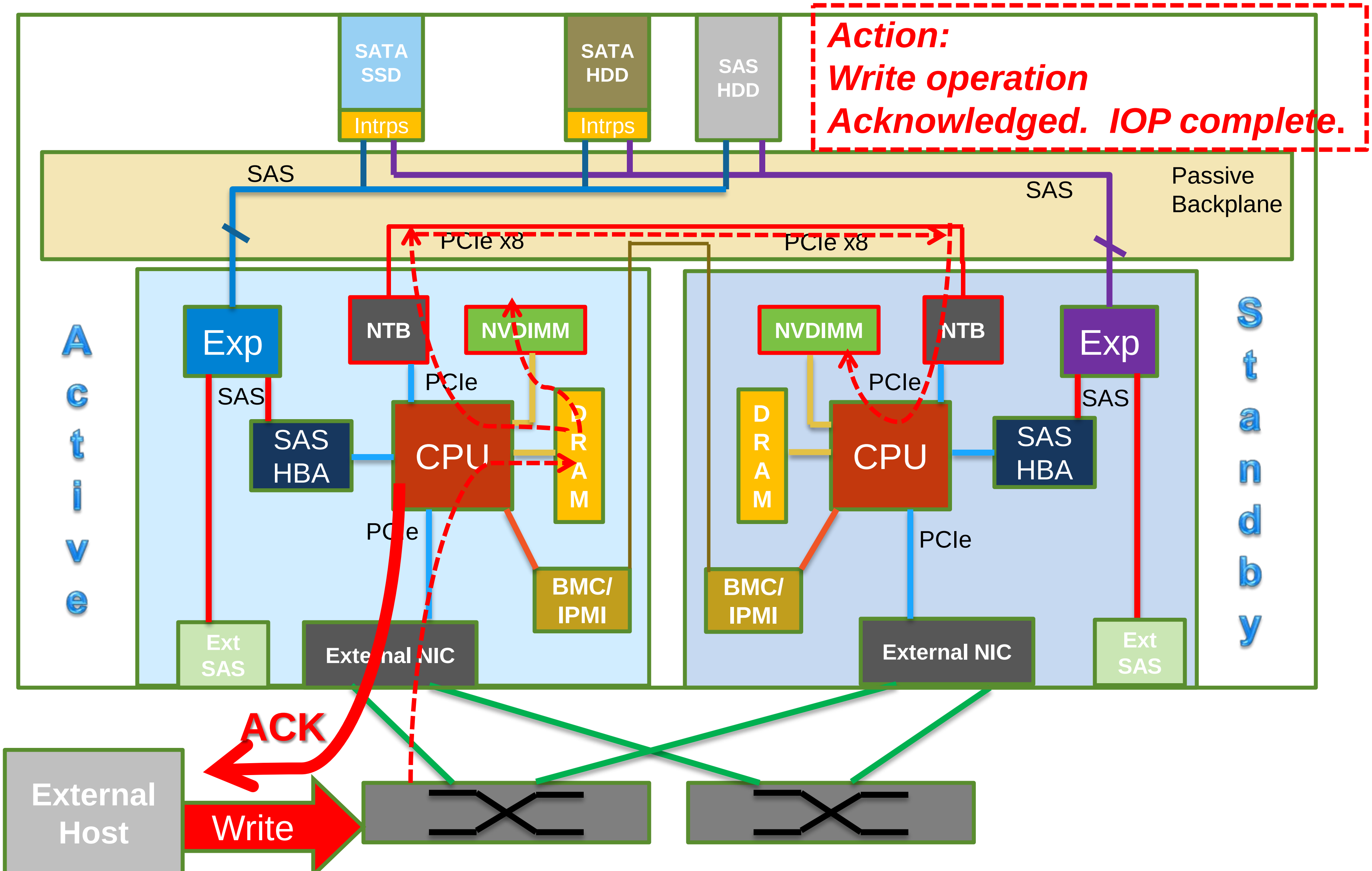
Action:
Memcopy data to NVDIMM from DRAM buffer.



Next Generation Integrated Enterprise Storage Array – Mirror NVData to / from NVDIMM via PCIe NTB



Next Generation Integrated Enterprise Storage Array – Mirror NVData to / from NVDIMM via PCIe NTB



Conclusion

- NVDIMMs combined with PCIe NTB's have enabled Integrated Enterprise Storage Platforms to achieve significant performance improvements *(> 4X Write IOP latency improvement !!)*

And That’s All Folks…..

Thank You!