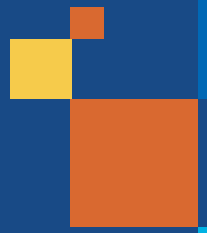
Intel® Virtual RAID on CPU (Intel® VROC) Detailed Comparison to RAID HBA



intel®

Notices and Disclaimers

Notices & Disclaimers

Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

Purpose:

Broad categorical comparison of Intel VROC (Integrated RAID) vs HW RAID HBAs on features, performance, latency, CPU% and power usage.

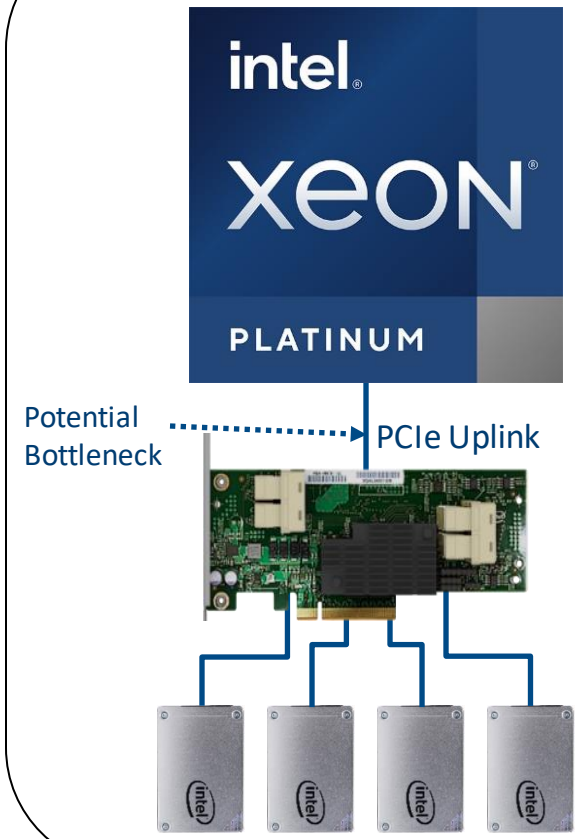
Agenda:

1. Architecture and Feature Comparison
2. Key findings
3. Intel® Optane™ SSD Comparisons
4. Test Configuration Details
5. Pass-thru Mode (No RAID) Comparison
6. RAID0/1/5/10 Performance Results
7. Detailed RAID0/5 Review (Latency, CPU%, Power)

Architecture and Feature Comparison

Intel® VROC vs RAID HBA

Legacy RAID Architecture



RAID HBA

Product:
- MegaRAID 9560-16i

Category:
- HW RAID

PCIe Generation:
- Gen. 4

Storage Uplink:
- x8 PCIe Lanes

Drives:
- 4 SSDs

Intel VROC

Product:
- Intel VROC

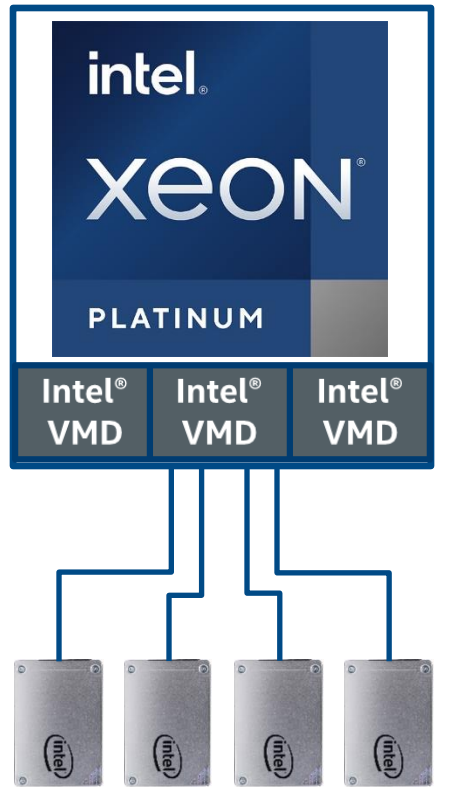
Category:
- **Integrated RAID**

PCIe Generation:
- Gen. 4

Storage Uplink:
- **X4 PCIe per SSD**

Drives:
- 4 SSDs

Intel® Xeon® Scalable Processor



Intel® VROC onboards RAID HBA functionality onto Intel® Xeon® CPUs[1]

1-Intel VROC and Intel VMD are available on all generations (Gen. 1, 2 and 3) and SKUs (Bronze, Silver, Gold, and Platinum) of Intel Xeon Scalable Processor

Intel® VROC vs RAID HBA

Major RAID Features	HW RAID	VROC	Intel® VROC Comment
Error Handling/Isolation	√	√	Both architectures isolates SSD error/event handling to reduce OS crash/reboot
Reliable data storage	√	√	Enterprise data protection, even when power loss occurs
Boot support	√	√	Redundant system volume = less down-time/crashes
In-band Management Tools	√	√	Various UEFI, GUI, and CLI Utilities for each
Out-of-band RAID Config.	√	X	Intel VROC has OOB on roadmap for upcoming releases
Full NVMe SSD x4 Bandwidth	X	√	Intel VROC + Intel VMD allows full x4 access to SSDs, no HW Uplink
RAID Processing Location	On HBA	On Intel® Xeon	Uses powerful Intel® Xeon® CPU to RAID the fast NVMe* SSDs. Better scaling for heavy workloads (see Detailed CPU Review)
Supported RAID Levels	0/1/5/6/10/50/60	0/1/5/10	RAID6/50/60 not needed for perf./AFR of NVMe SSDs
Write back cache	DRAM + BBU	Integrated Caching + Intel® Optane™ SSD	Replace DRAM WB Cache + BBU with persistent Intel® Optane™ media
SED Key Management	On HBA	Platform Integrated	Intel VROC uses platform protocols and remote KMS to manage keys
Idle Power[1]	577W	562W	Tested 15W reduction in Idle Power Usage with Intel VROC

See backup for configuration details. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks..

Key Findings

Summary (Highlights)[1,2]

1. Intel VROC has **compelling features to replace RAID HBA**, plus a roadmap to fill any gaps (OOB)
2. Intel VROC is the only RAID solution that scales with the Intel Optane SSD solution to deliver extraordinary performance (**Over 5.6M IOPS!**)
3. Intel VROC performance for all RAID levels is equal or better than RAID HBA (**↑ Performance, ↓ Latency**)
4. Intel VROC can improve resource utilization by removing the HBA and related choke points (**↓ CPU Usage, ↓ Power**)
5. Intel VROC has a scalable, integrated design that is better designed for NVMe SSDs (**↑ IOPS/CPU Core, ↑ IOPS/W**)

See backup for configuration details. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks..

Test Configuration Details

Test Configuration Details (Optane)

Legacy RAID HBA

4 x 400GB Intel Optane P5800X SSDs

- Write Spec: 1,150,000 IOPS
- Read Spec: 1,500,000 IOPS

Tested Configurations:

- Single Drive Performance
- 4x Drives pass-thru in parallel (no RAID)
- 4x Drive RAID0/5/10
- 2x Drive RAID1

Workload Details:

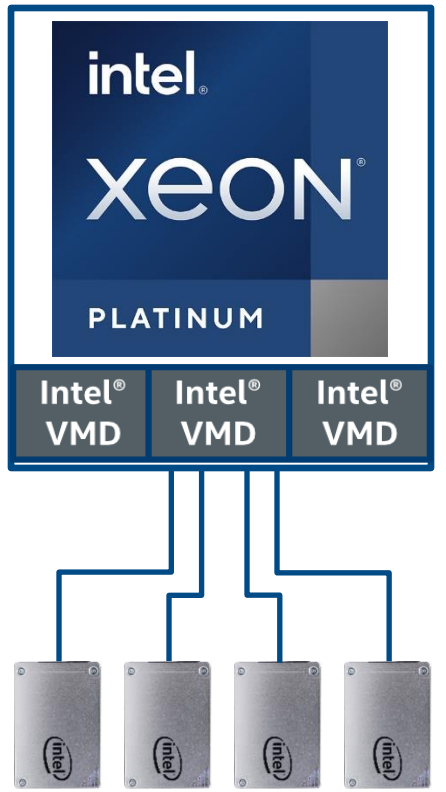
- 4k Random: 70/30 R/W
- 16 Threads, 16 IODepth

Metrics

- Performance: IOPS
- Bandwidth: MB/sec
- Latency: µsec
- CPU Usage*: Effective Intel Xeon Cores used

Data on Slides 11-14

Intel® Xeon® Scalable Processor



*CPU Usage measured as total platform CPU % consumption, includes workload generation, storage stack (RAID) usage, and background activity
Measured as "Cores Used" = CPU% report out * # cores on system (64 cores)

Test Configuration Details (NAND)

Legacy RAID HBA

4x 3.84TB Intel D7 P5510 SSDs

- Write Spec: 170,000 IOPS
- Read Spec: 700,000 IOPS

Tested Configurations:

- Single Drive Performance
- 4x Drives pass-thru in parallel (no RAID)
- 4x Drive RAID0/5/10
- 2x Drive RAID1

Workload Details:

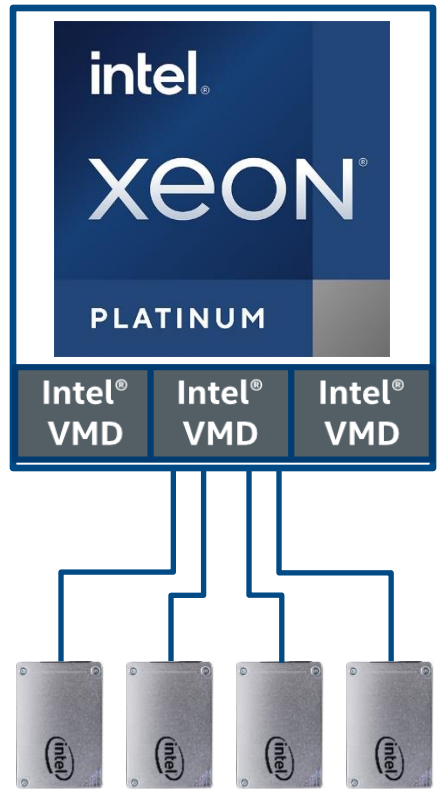
- 4k Random: 100% Reads, 70/30 R/W, 100% Writes
- 1 Threads, 1 IODepth (Isolate Storage Path)
- 16 Threads, 64 and 256 IODepth (Peak performance)

Metrics

- Performance: IOPS
- Power: Watts (Idle and under load)
- Latency: µsec
- CPU Usage*: Effective Intel Xeon Cores used

Data on slides15-28

Intel® Xeon® Scalable Processor



*CPU Usage measured as total platform CPU % consumption, includes workload generation, storage stack (RAID) usage, and background activity
Measured as "Cores Used" = CPU% report out * # cores on system (64 cores)

Intel Optane Comparisons

RAID Levels Performance Comparison[1]

Intel® Optane™ SSDs: 16 Thread, 16 IODepth: 70/30 R/W

Intel VROC achieves up to 5.6 million IOPS with RAID0 on mixed workloads

Intel VROC has up to:

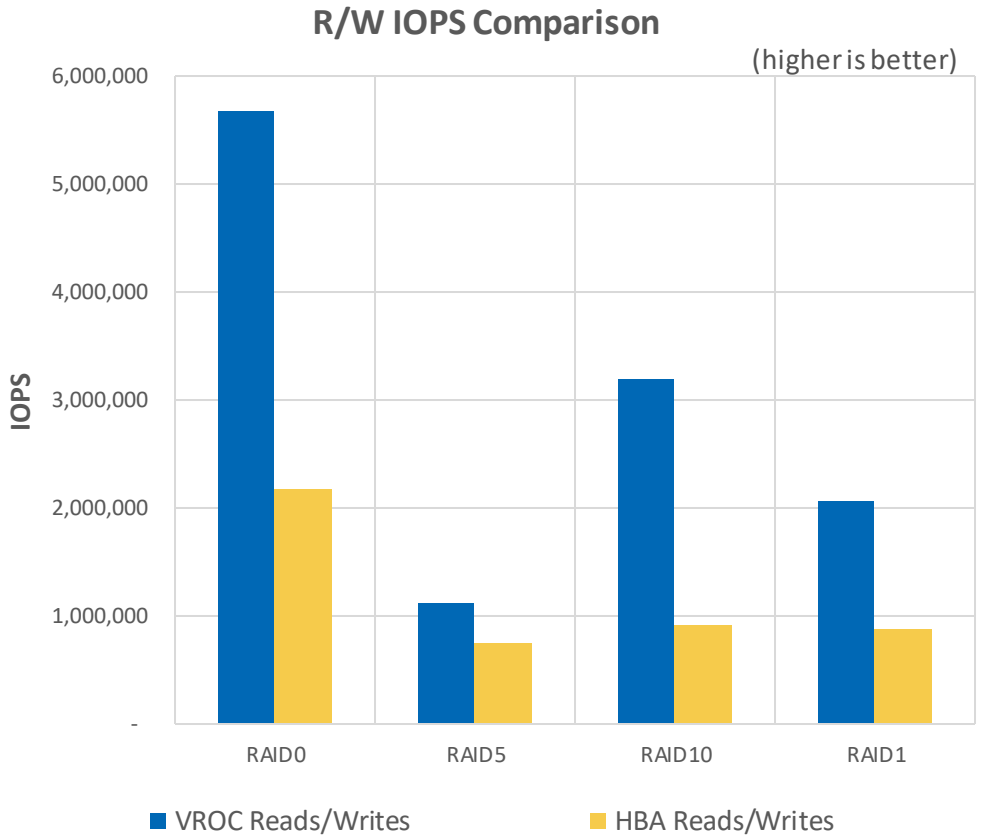
161% more IOPS on RAID0

50% more IOPS on RAID5

248% more IOPS on RAID10

138% more IOPS on RAID1

Intel VROC RAID5 > HBA RAID10 performance



See backup for configuration details. Results may vary

RAID0 Simultaneous Read/Write Comparison[1]

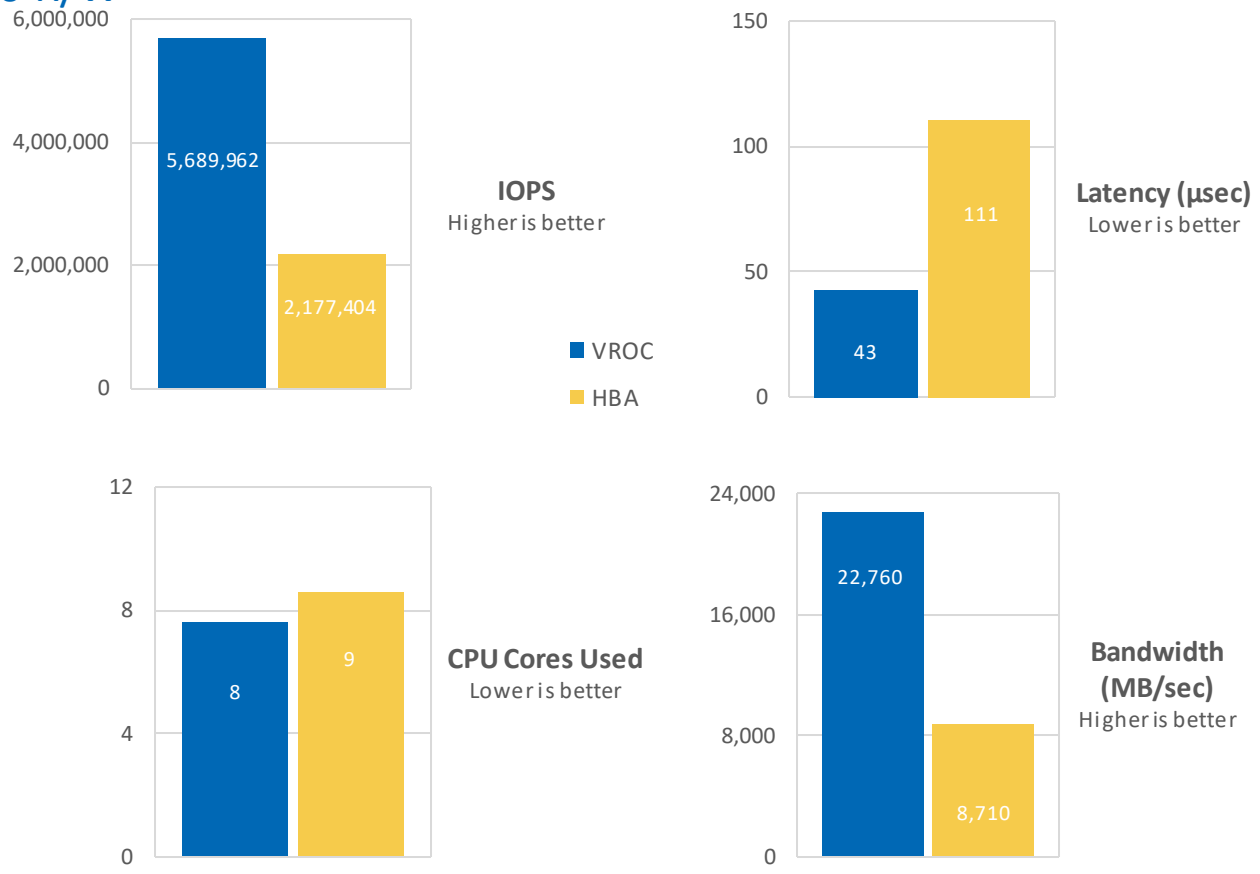
Intel® Optane™ SSDs: 16 Thread, 16 IODepth: 70/30 R/W

Intel VROC RAID0 reads/writes provides:

- ↑ **IOPS**
- ↓ **Latency**
- ↓ **CPU Usage**
- ↑ **Bandwidth**

RAID0 provides higher performance metrics but with lower resource usage (CPU)

Up to 161% more Read/Write IOPS
Up to 61% lower latency



See backup for configuration details. Results may vary

RAID5 Simultaneous Read/Write Comparison[1]

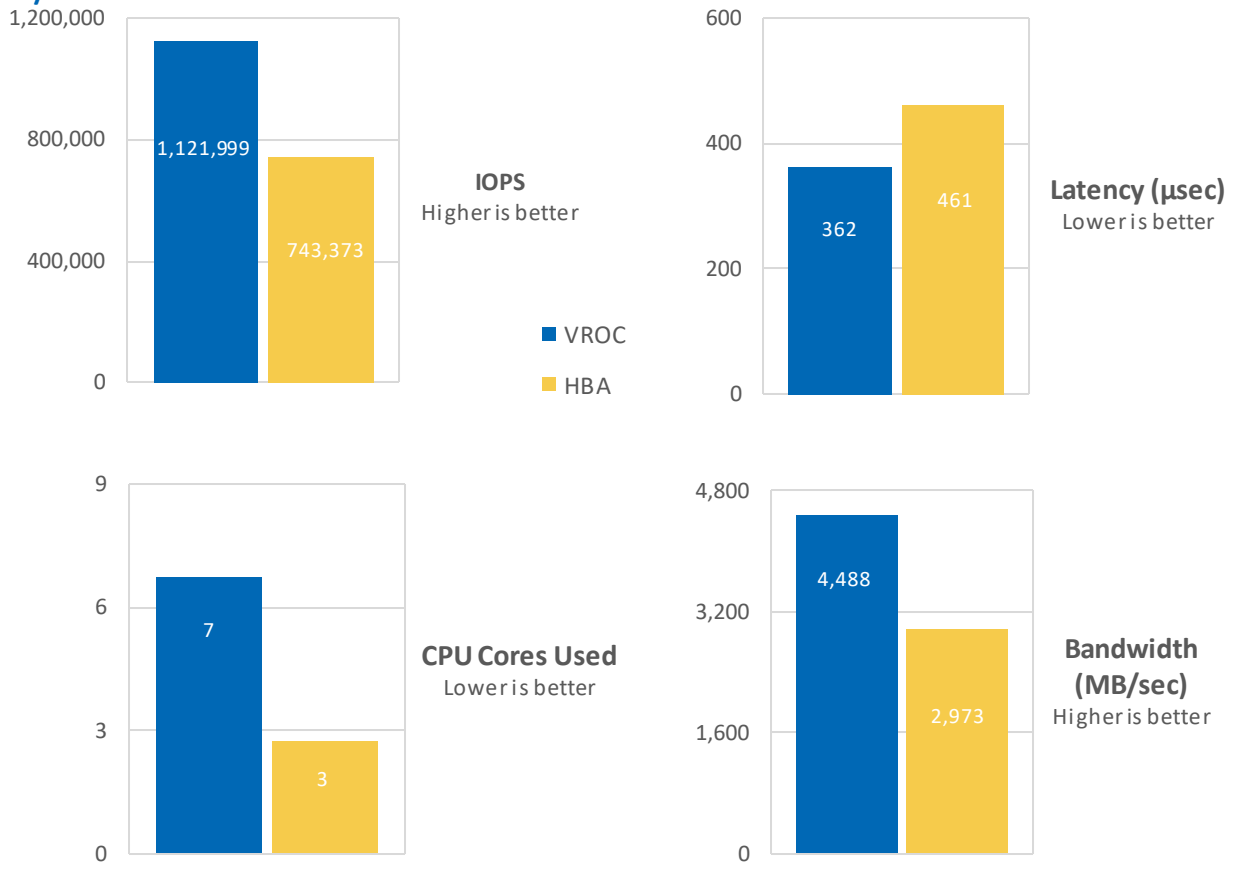
Intel® Optane™ SSDs: 16 Thread, 16 IODepth: 70/30 R/W

Intel VROC RAID5 reads/write provides:

- **↑ IOPS**
- **↓ Latency**
- ↑ CPU Usage*
- **↓ Bandwidth**

*RAID5 uses 4 more cores but delivers up to 380K additional IOPS

Up to 50% more Read/Write IOPS
Up to 50% more Bandwidth



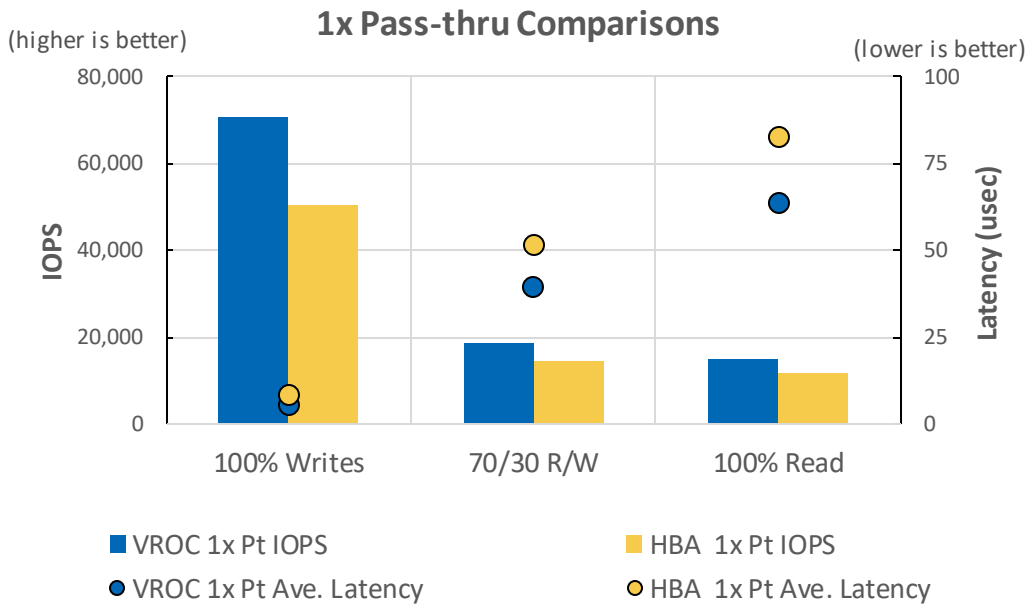
See backup for configuration details. Results may vary

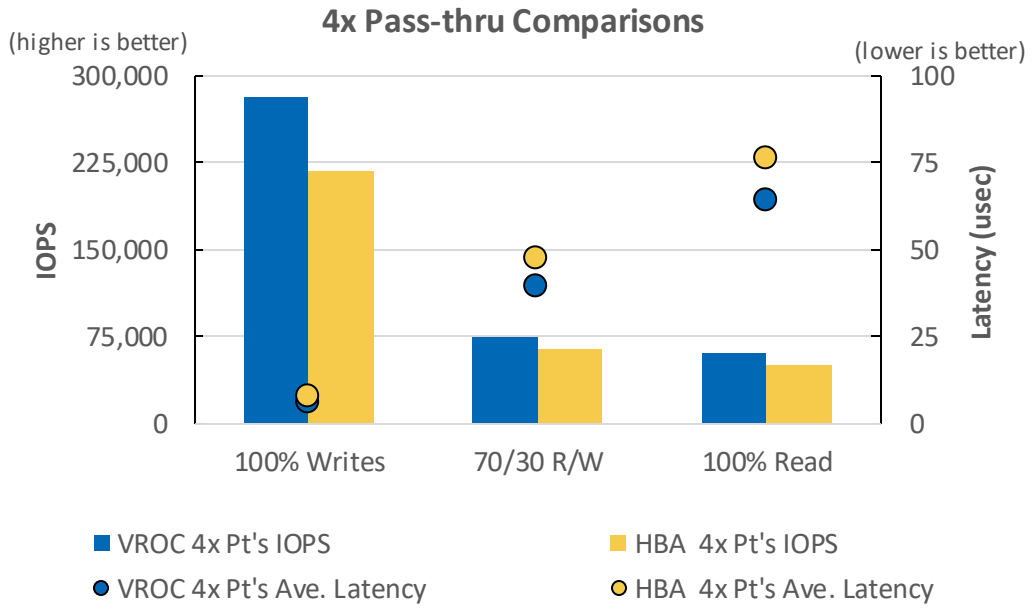
NAND SSD Comparisons

Pass-thru Mode (No RAID) Comparison

Low Workload, Pass-Thru Comparison[2]

NAND SSDs: 1 Thread, 1 IODepth





Intel VROC provides unimpeded access to storage for lower latency I/0

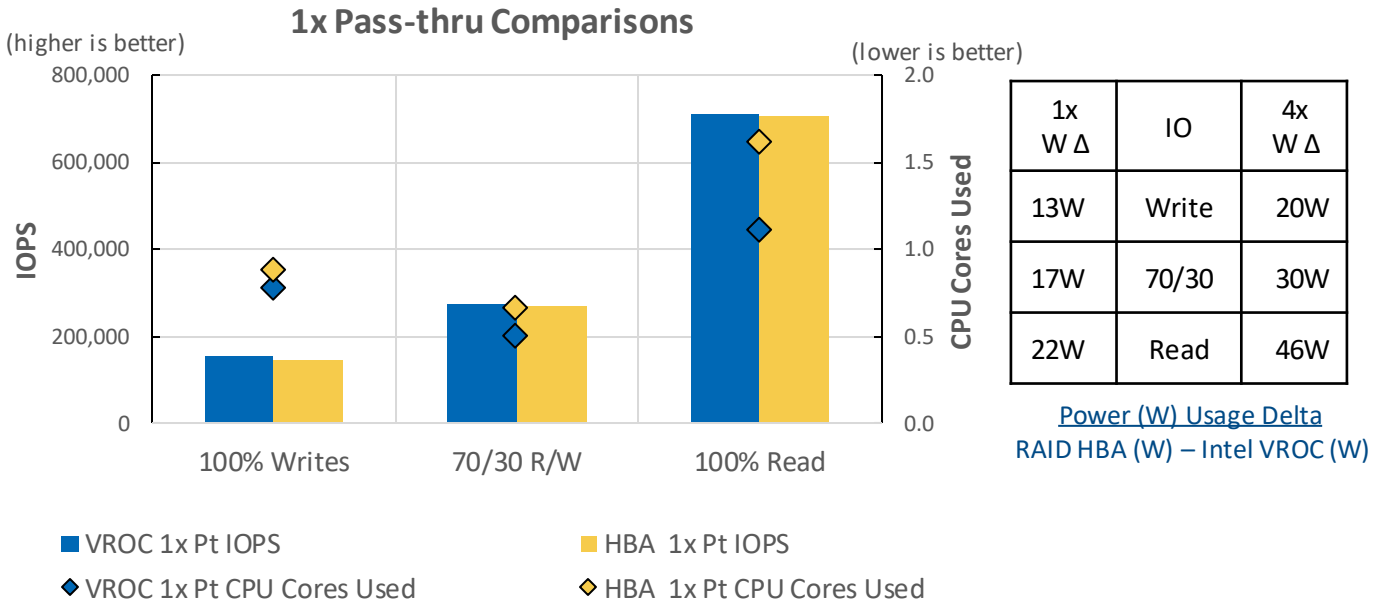
- Single Drive, 100% Write: {**40% IOPS ↑, 32% Latency ↓**}
- Single Drive, 100% Read: {**29% IOPS ↑, 23% Latency ↓**}

Single drive performance improvements scales to multiple drives

See backup for configuration details. Results may vary

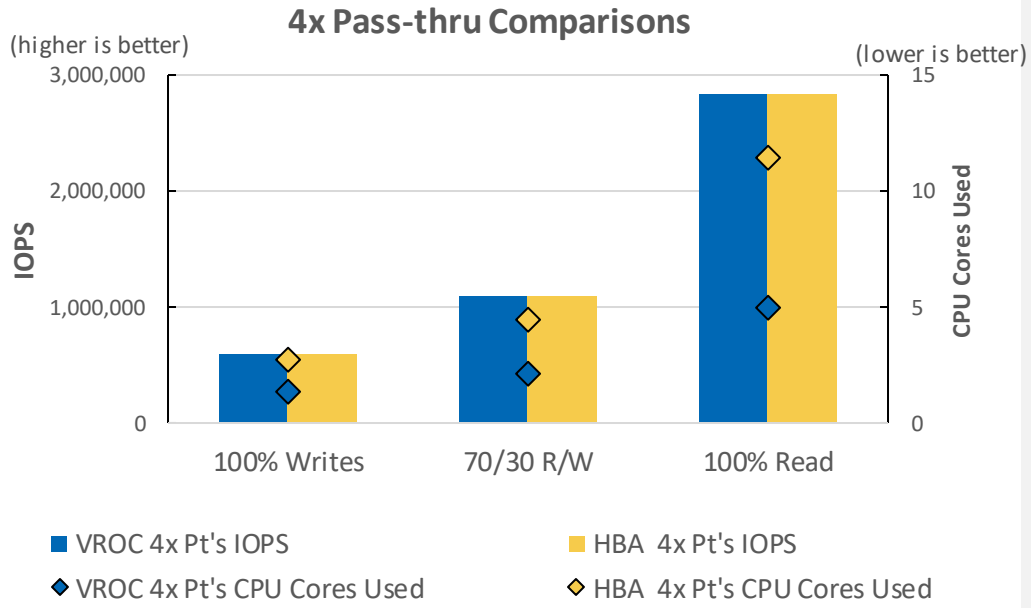
Peak Performance, Pass-Thru Comparison[2]

NAND SSDs: 16 Thread, 64 IODepth



1x W Δ	IO	4x W Δ
13W	Write	20W
17W	70/30	30W
22W	Read	46W

Power (W) Usage Delta
RAID HBA (W) – Intel VROC (W)



Higher workloads saturate the storage on both solutions

- Latency differences are masked, performance becomes equivalent

Other architecture differences are exposed: Power and CPU usage

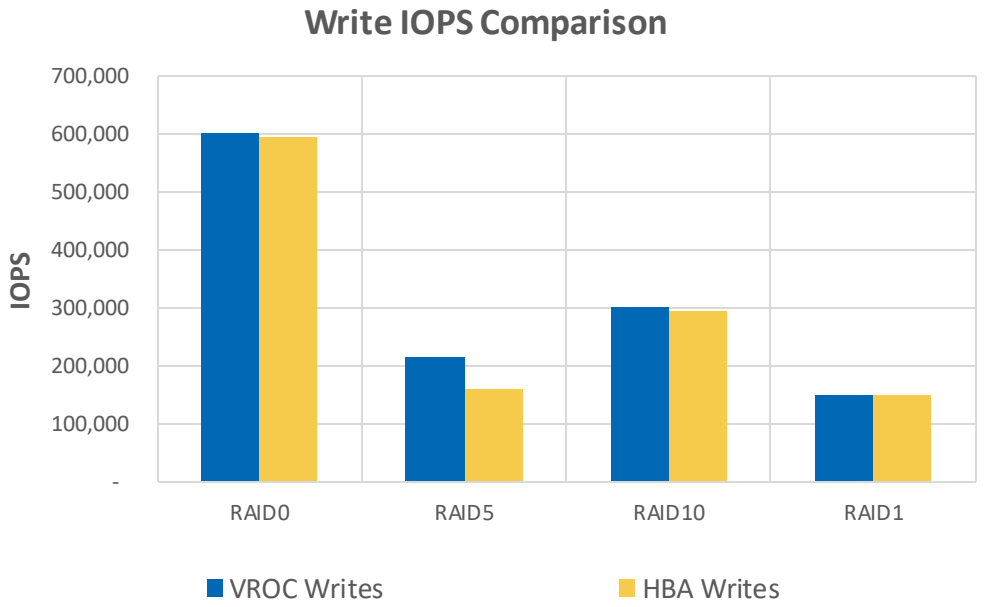
- Additional HBA power draw creates positive WΔ; **Intel VROC ↓ Power**
- RAID HBA on card processing is oversaturated by larger workloads; **Intel VROC ↓ CPU Usage** (See detailed CPU Review)

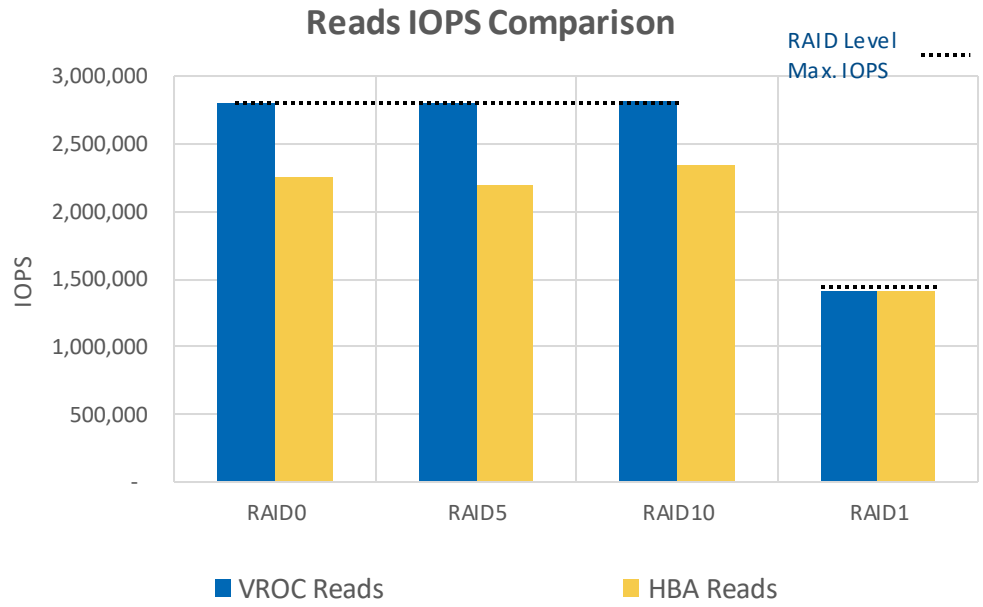
See backup for configuration details. Results may vary

RAID0/1/5/10 Performance Results

RAID Levels Performance Comparison[2]

NAND SSDs: 16 Thread, 64 IODepth





- Intel VROC has **33% more IOPS on RAID5 writes**

Intel VROC Read Performance scales to maximum 4x SSD Spec (~2.8M IOPS RAID0/5/10)

- HBA hits 2.2M IOPS Bottleneck; **Intel VROC delivers up to 27% more IOPS on RAID0/5/10 reads**

See backup for configuration details. Results may vary

Detailed RAID0/5 Review (Latency, CPU%, Power)

RAID0/5 Read Comparison[2]

NAND SSDs: 16 Thread, 64 IODepth

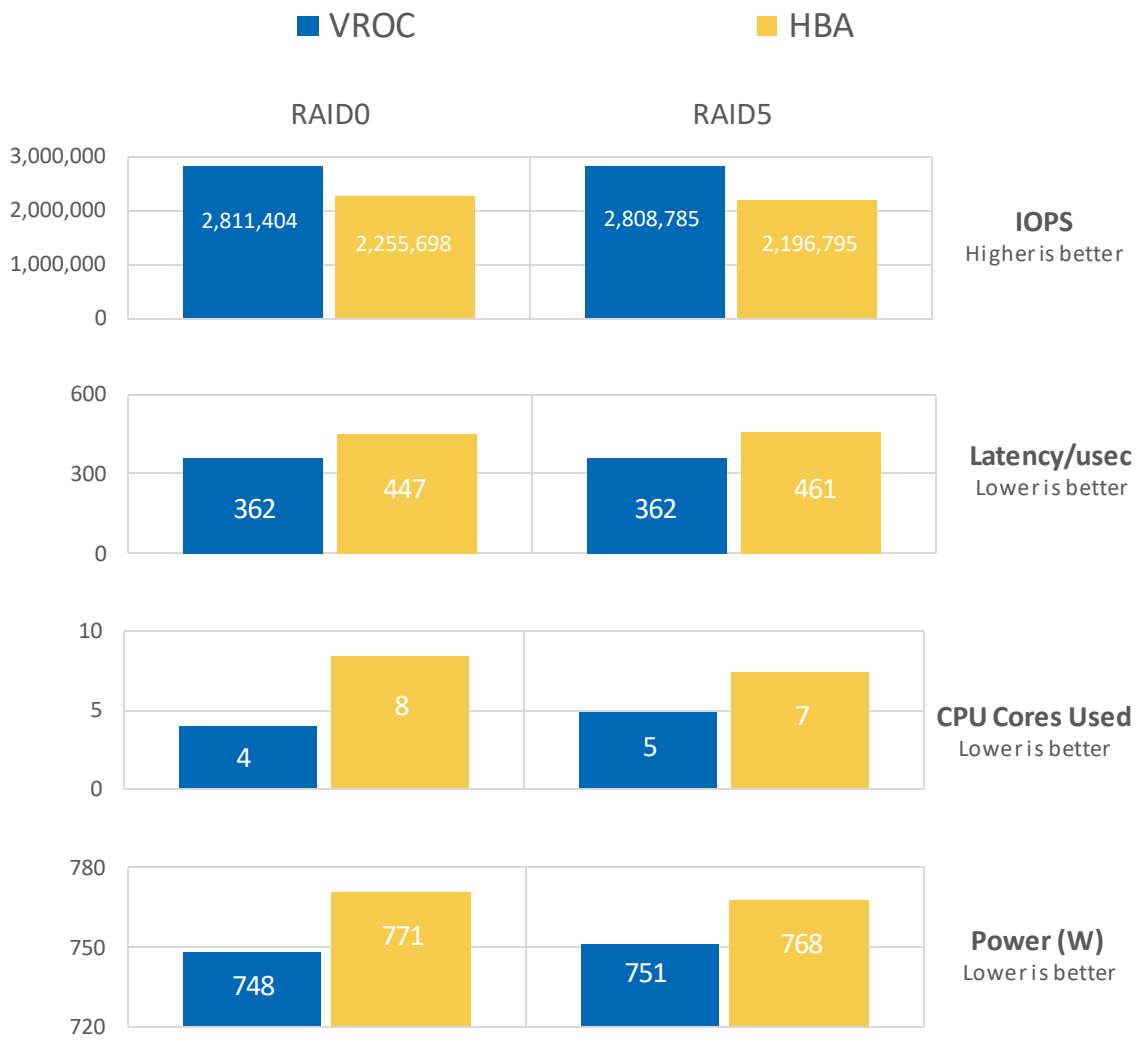
Intel VROC RAID0/5 reads provides:

- ↑ **IOPS**
- ↓ **Latency**
- ↓ **CPU Usage**
- ↓ **Power Consumption**

Integrated RAID is a more effective RAID architecture for NVMe SSDs

Up to 30% more Read IOPS/W

Up to 164% more Read IOPS/CPU Cores Used



See backup for configuration details. Results may vary

RAID0/5 Write Comparison[2]

NAND SSDs: 16 Thread, 64 IODepth

Intel VROC RAID0/5 reads provides:

- ↑ **IOPS**
- ↓ **Latency**

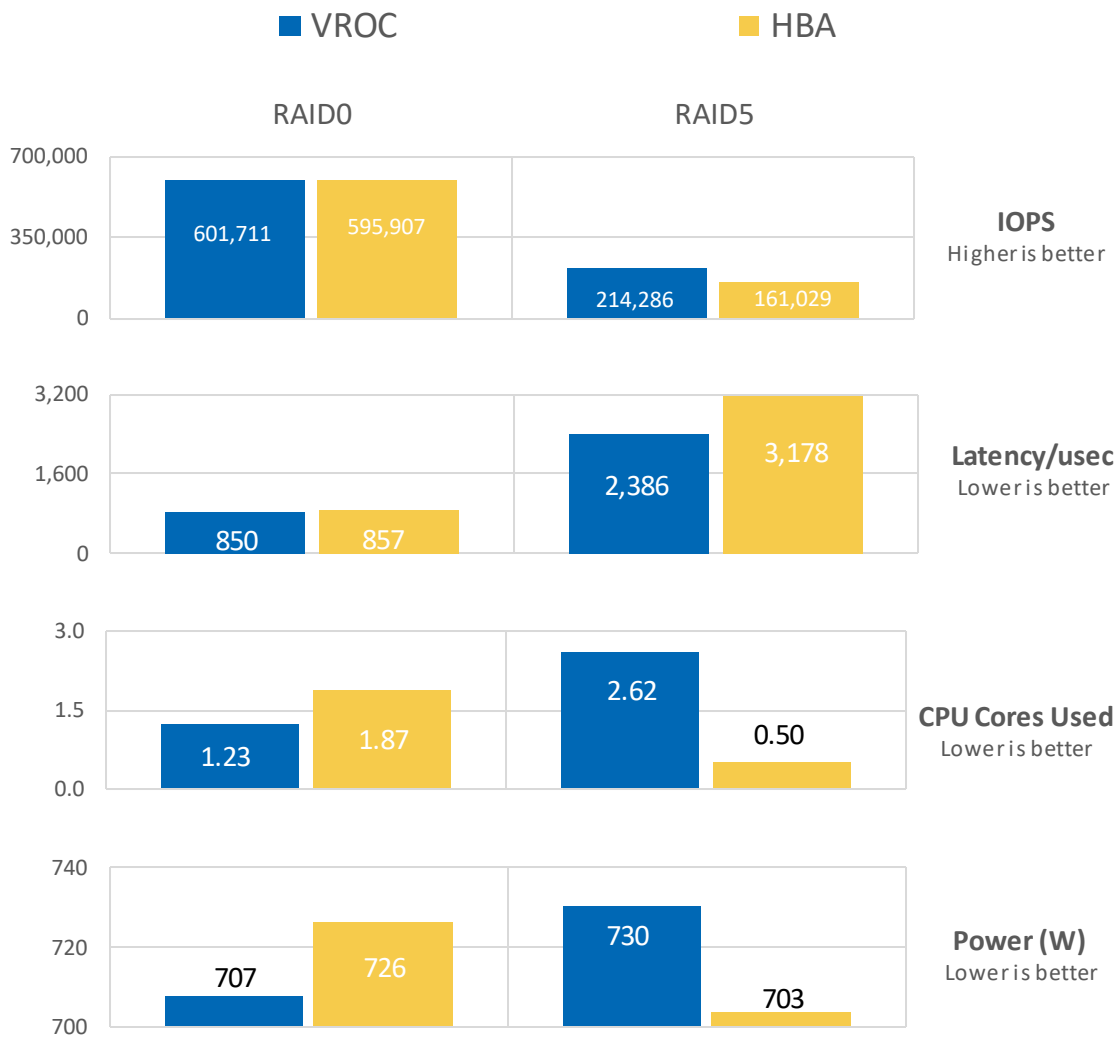
RAID 0 also ↓ **CPU Usage and Power Usage**

RAID5 provides higher performance metrics but with higher resource usage (CPU and Power)....

This is not the whole story

Up to 28% more Write IOPS/W

See 'CPU% Usage Explained' for more



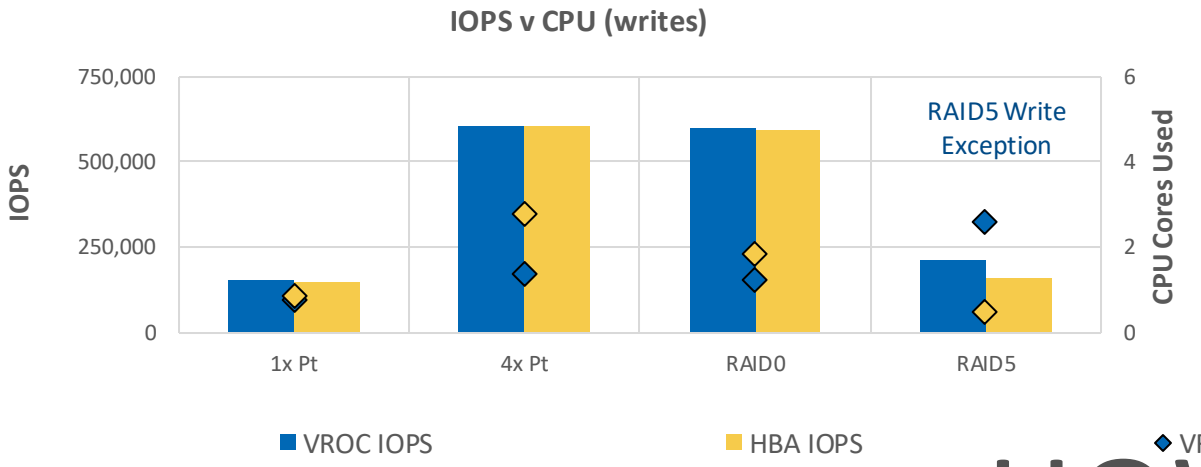
See backup for configuration details. Results may vary

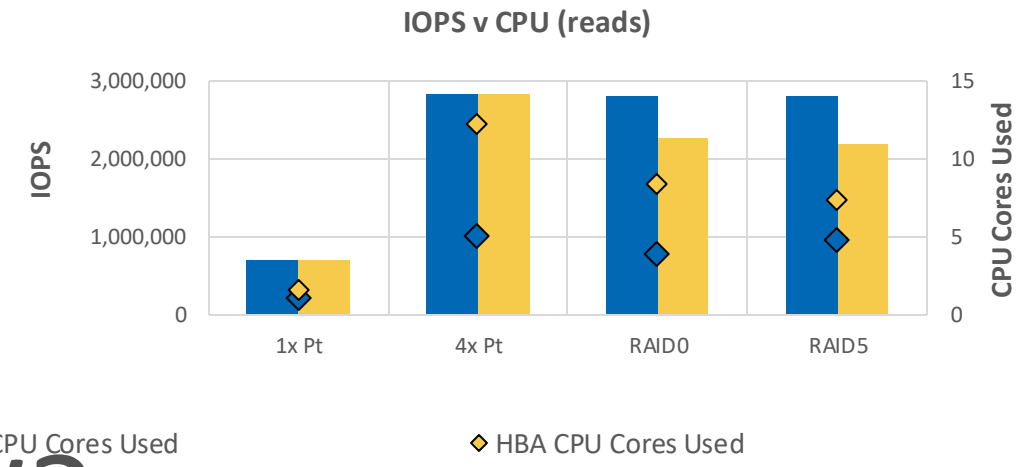
CPU% Usage Explained

CPU% Usage-Perception[2]

Common perception: RAID HBA consumes less host CPU resources due to HBA offload

Reality: **Intel VROC can deliver ↑ performance and consumes ↓ CPU resources!**





HOW?

See backup for configuration details. Results may vary

CPU% Usage-Reality Explained[2]

NVMe SSD performance can overwhelm RAID HBA offload design

16 Threads 64 IODepth → 100k's Write IOPS and 1M's Read IOPS

HBA architecture has choke points that can bottleneck performance:

1. Limited PCIe Uplink (x8 PCIe lanes)
2. Fixed amount of RAID processing
3. SCSi-based RAID stack

Full x4 bandwidth per NVMe SSD

Scaled compute on powerful Intel Xeon

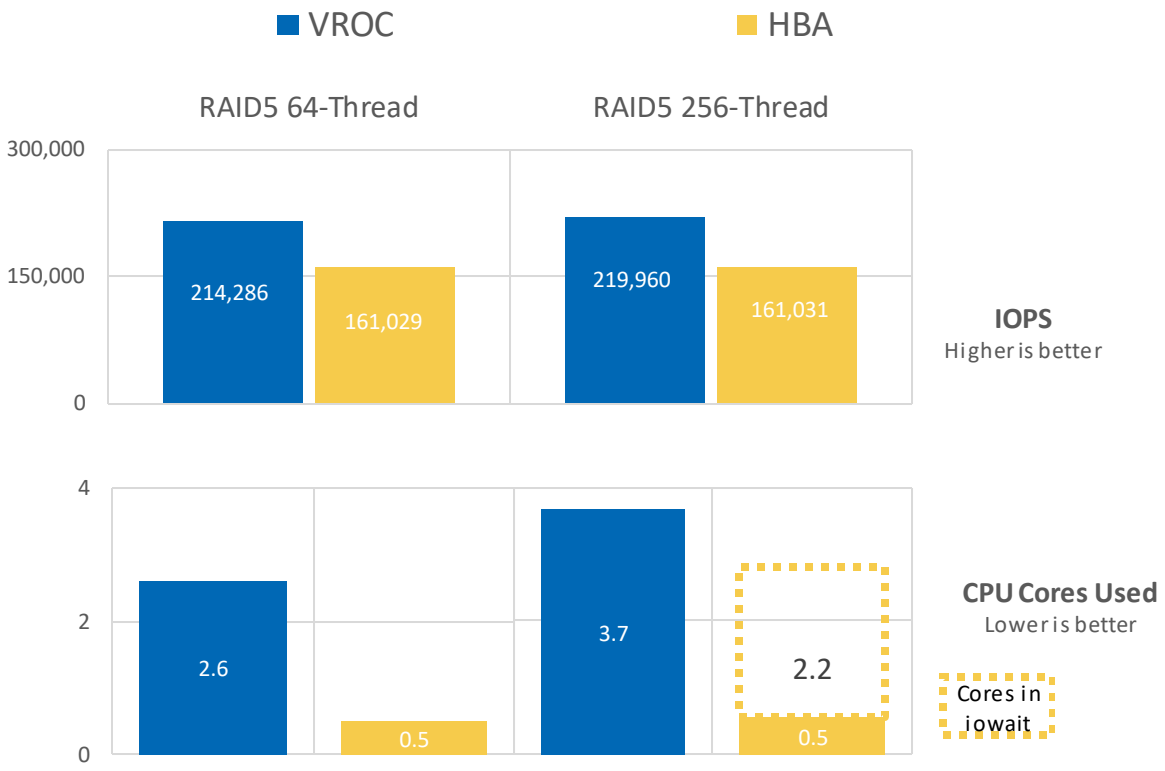
NVMe optimized RAID stack

These limitations cause thrash on CPU%....and can lead to iowait%

See backup for configuration details. Results may vary

iowait% Closer Look[2]

NAND SSDs: 16 Thread, 64 IODepth → 16 Thread, 256 IODepth



RAID5 writes require high CPU%

- Highest of any Intel VROC supported RAID level per IOP

RAID HBA offload generates iowait at higher workloads:

- If limits of HBA architecture are reached (more IO), host CPU usage ramps up in iowait%
- Iowait could be wasted cycles depending on application

Intel VROC is more efficient for RAID5 writes:

No ramping of iowait

Up to 4% more Write IOPS/CPU Cores Used*

*when accounting for iowait%

See backup for configuration details. Results may vary

CPU% Usage-Customer Impact

Server design must plan for **Peak Storage Load**

Peak Storage Load (PSL): Max. IO during data center operation

RAID Solution Response to PSL	RAID HBA	Intel VROC
	• Bottleneck performance • Iowait% ramp and higher latency • Operational Thrash if storage architecture not properly planned	• Scale performance to absorb PSL • Proportionally ramp CPU usage and latency • Mitigate server thrash with fewer CPU cores dedicated for RAID

Intel VROC servers often require fewer CPU cores to handle Peak Storage Load

See backup for configuration details. Results may vary

Backup

Configuration Details

1. Intel VROC vs RADI HBA Comparison (Optane)
System configuration: Beta Coyote Pass M50CYP2SB2U/M50CYP2SBSTD (chassis M50CYP2UR208BPP), 2 x Intel® Xeon® Platinum 8358 CPU @ 2.60GHz, 32 cores each, DRAM 128GB , BIOS Release 04/02/2021, BIOS Version: SE5C6200.86B.0020.P24.2104020811
OS: RedHat* Enterprise Linux 8.1, kernel-4.18.0-147.el8.x86_64, mdadm - v4.1 - 2018-10-01, Intel® VROC Pre-OS version 7.5.0.1152
Storage: Both configurations used 4 x 400GB Intel Optane P5800X PCIe Gen4 U.2 SSDs (Model: SSDPF21Q400GB, Firmware: L0310100) connected to backplane which is connected via SlimSAS cables directly to a Broadcom 9560-16i (x8) card on Riser 2, PCIe slot 1 on CPU2 **BIOS setting**: SpeedStep(Enabled), Turbo(Enabled), ProcessorC6(Enabled), PackageC-State(C0/C1 State), CPU_PowerAndPerformancePolicy(Performance), HardwareP-States(NativeMode), WorkloadConfiguration(I/O Sensitive)
RAID Configurations: 4-Disk RAID0/5/10 and 2-Disk RAID1 with Intel VROC and Broadcom MegaRAID 9560-16i
Workload Generator: FIO 3.25, 16-thread 16-IODepth
Performance results are based on testing as of 6/25/2021 and may not reflect all publicly available updates. See configuration disclosure for details. No product can be absolutely secure.

Configuration Details

2. Intel VROC vs RADI HBA Comparison (NAND)
System configuration: Beta Coyote Pass M50CYP2SB2U/M50CYP2SBSTD (chassis M50CYP2UR208BPP), 2 x Intel® Xeon® Platinum 8358 CPU @ 2.60GHz, 32 cores each, DRAM 128GB , BIOS Release 03/22/2021, BIOS Version: SE5C6200.86B.0022.D08.2103221623
OS: RedHat* Enterprise Linux 8.1, kernel-4.18.0-147.el8.x86_64, mdadm - v4.1 - 2018-10-01, Intel® VROC Pre-OS version 7.5.0.1152
Storage: Both configurations used 4x 3.84 TB Intel® D7-P5510 Series SSDs (Model: SSDPF2KX038TZ, Firmware: JCV10016) connected to internal backplane. With Intel VROC config, backplane connect directly to CPU2 via SlimSAS. With RAID HBA, backplane connect to RAID HBA on Riser 2, PCIe slot 1 on CPU2
BIOS setting: SpeedStep(Enabled), Turbo(Enabled), ProcessorC6(Enabled), PackageC-State(C0/C1 State), CPU_PowerAndPerformancePolicy(Performance), HardwareP-States(NativeMode), WorkloadConfiguration(I/O Sensitive)
RAID Configurations: 4-Disk RAID0/5/10 and 2-Disk RAID1 with Intel VROC and Broadcom MegaRAID 9560-16i
Workload Generator: FIO 3.25, 1-thread 1-IODepth, 16 thread 64/256 IODepth
Performance results are based on testing as of 5/3/2020 and may not reflect all publicly available updates. See configuration disclosure for details. No product can be absolutely secure.