INTEL® VIRTUAL RAID ON CPU RAID SCALING

February 26, 2019

Legal Disclaimers

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.

Intel does not control or audit the design or implementation of third party benchmark data or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmark data are reported and confirm whether the referenced benchmark data are accurate and reflect performance of systems available for purchase.

Intel® VROC RAID Scaling

Scaling Summary: Intel® VROC is the primary solution for NVMe* RAID. One key metric used to assess RAID solutions is the ability of that solution to scale performance. Scaling is measured by how much incremental performance is achieved as more storage devices are added to a RAID array.

Intel® VROC scales performance with NVMe SSDs in RAID0 and RAID5 configurations. How many drives are used, how they are connected, and the workload used can impact scaling results. This guide shows actual Intel® VROC scaling data and lists some tips to optimize storage performance for scaling.

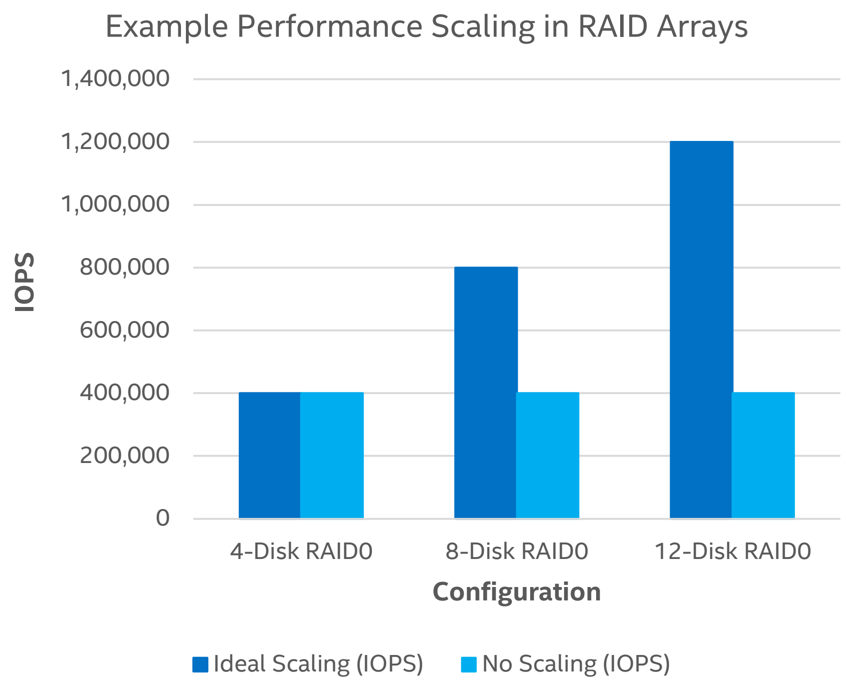
Document Scope

This presentation is meant to show that Intel® Virtual RAID on CPU scales performance with NVMe* SSDs. It will present some example performance results and explain what system characteristics may limit scaling results.

The goal is to use actual data to demonstrate the advantages gained by NVMe SSDs with Intel® VROC. It is important to note that specific configuration details will impact performance. This document uses a broadly accepted workload for generalized results, and may not exactly demonstrate the results of a specific use case. Instead, this document should be used as a guide to run similar tests and get applicable details for a specific installation.

*Other names and brands may be claimed as the property of others

What is Scaling in RAID?



RAID levels that utilize striping (e.g. RAID0/5) are able to improve storage performance by doing more work in parallel as more devices are added to the array. This is called scaling. Two examples of scaling are:

- "Ideal Scaling" - each additional device adds incremental performance to the RAID array equal to the pass-thru performance of the device
- "No Scaling" – each additional devices added to the array has no impact on performance.

RAID solutions will fall somewhere in-between "Ideal Scaling" and "No Scaling." The closer to "Ideal Scaling" a RAID solution can deliver, the better it utilizes the available storage media and leverages IT investment for business results.

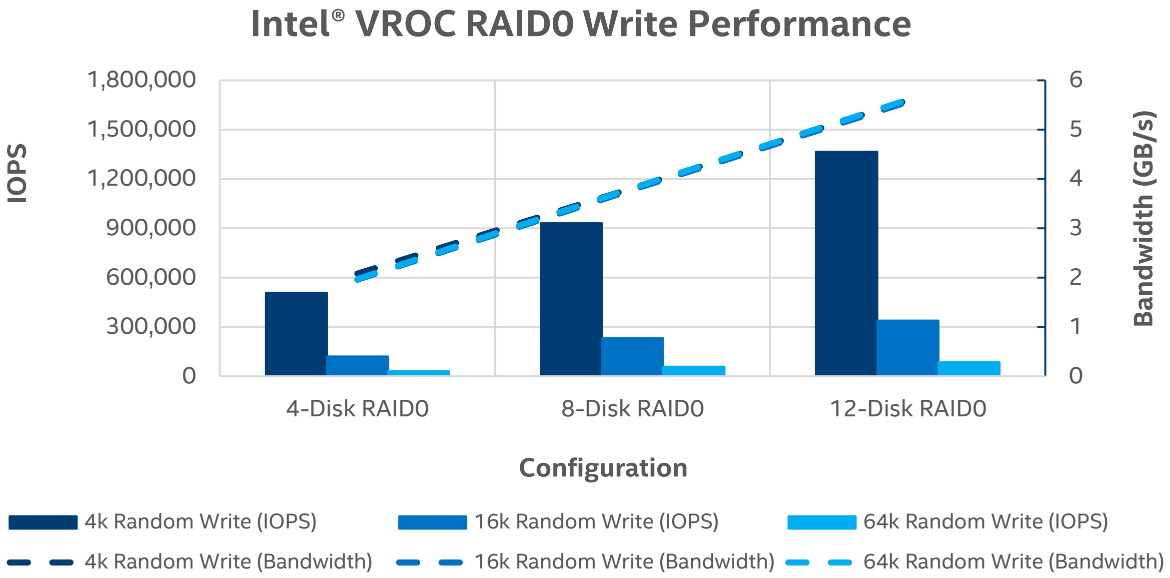
Intel® Virtual RAID on CPU Scaling Assessment

To evaluate how well Intel® Virtual RAID on CPU scales, the below configuration and workloads will be observed:

- Platform: S2600WFT with 2x Intel® Xeon® Platinum 8170 CPUs
- SSDs Used: 4TB Intel® SSD DC P4510
- RAID Levels: RAID0 and RAID5
- # of Drives: 4, 8, and 12 drives
  - 4 and 8 Drive arrays are on 1 Socket, 12 Drive arrays span both Sockets
- Workload:
  - 32 Threads, 128 IODepth
  - 4k, 16k, and 64k Random Reads and Writes

*Other names and brands may be claimed as the property of others

Intel® VROC RAID0 Write Scaling Results[1]



Intel® VROC with RAID0 Writes is an example of good RAID scaling.

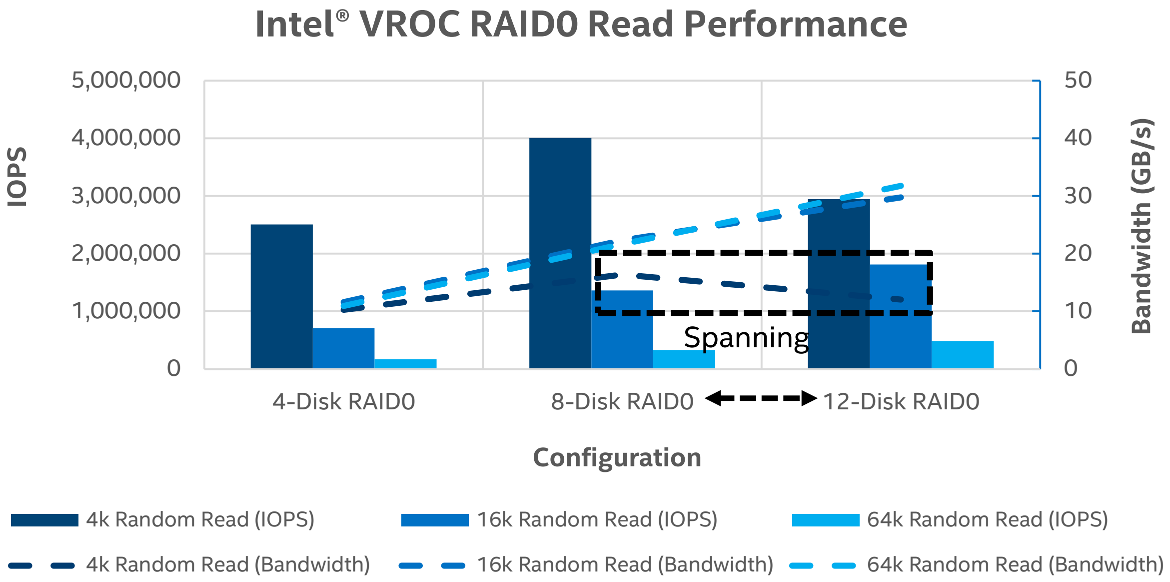
- Peak Performance: 5.6 GB/s (12-Disk, all block sizes)
- Performance scales from 4 to 8 to12 SSDs and increases by about the same amount
- Each chart will show scaling at 3 different block sizes. In some cases, large block sizes have better scaling.

See appendix for footnotes

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.

Performance results are based on testing as of January 11, 2019 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

Intel® VROC RAID0 Read Scaling Results[1]

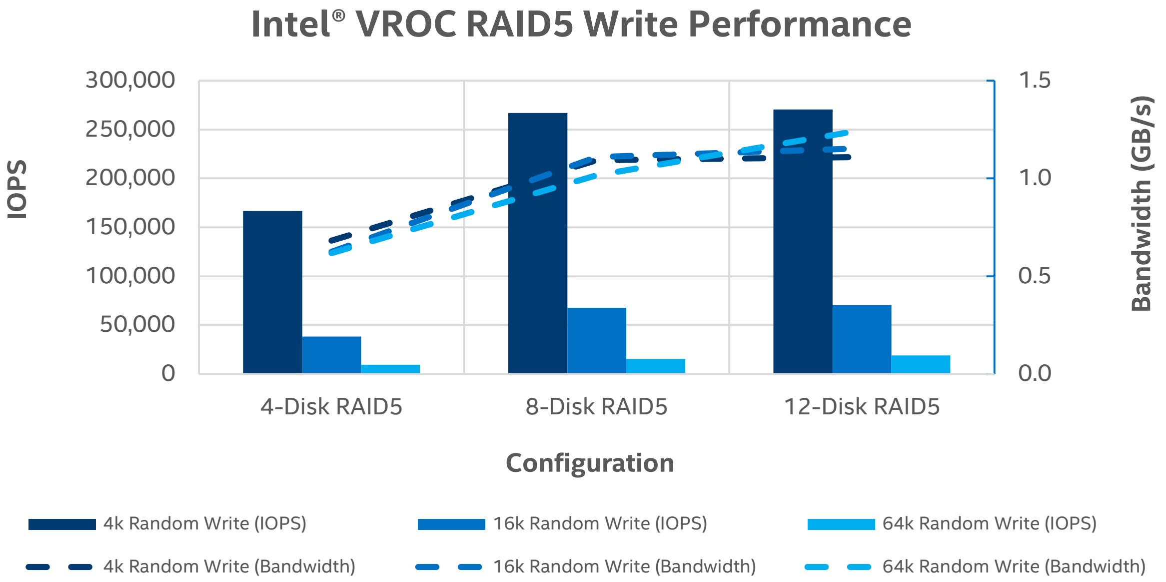


Intel® VROC with RAID0 Reads scales in some situations, but is doing much more work than RAID0 Writes:

- Peak Performance: 32 GB/s (12-Disk, 64k Block Size)
- Performance scales with 16k and 64k block sizes but is limited at 4k
- 12 Disk Scaling may be limited due to spanning CPUs with high IOPS

See appendix for footnotes

Intel® VROC RAID5 Write Scaling Results[1]

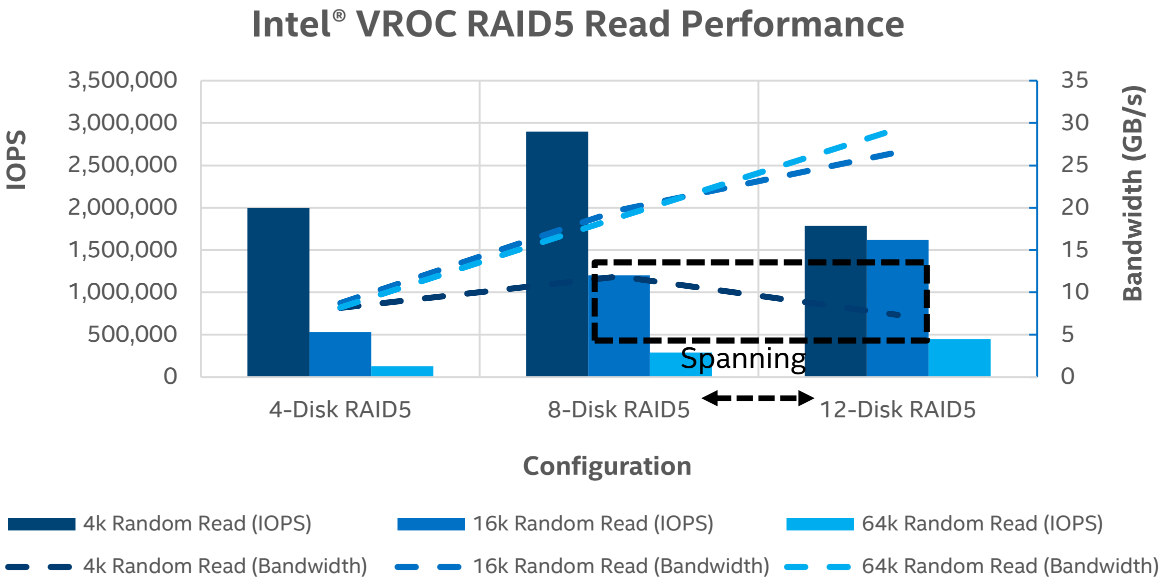


Intel® VROC with RAID5 Writes scales partially. Due to parity calculations, total throughput is lower than RAID0:

- Peak Performance: 1.25 GB/s (12-Disk, 64k Block Size)
- Block size has minimal impact on performance at these transfer rates

See appendix for footnotes

Intel® VROC RAID5 Read Scaling Results[1]

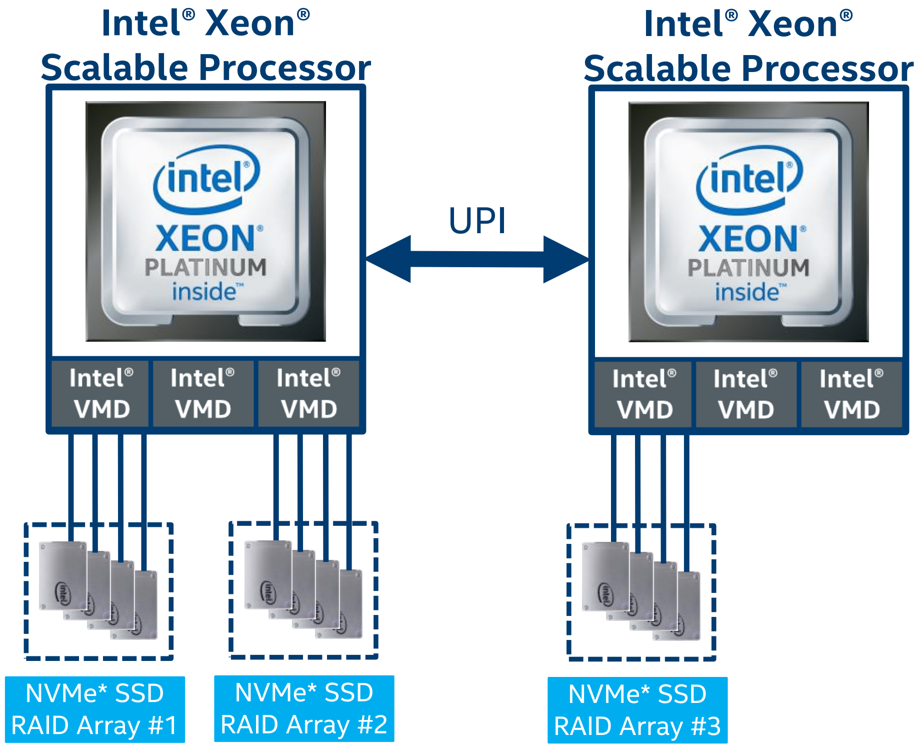


Intel® VROC with RAID5 Reads has a similar performance profile as RAID0 Reads:

- Peak Performance: 29 GB/s (12-Disk, 64k Block Size)
- Performance scales with 16k and 64k block sizes but is limited at 4k
- 12 Disk Scaling may be limited due to spanning CPUs with high IOPS

See appendix for footnotes

Spanning Proof Point- Configuration Change

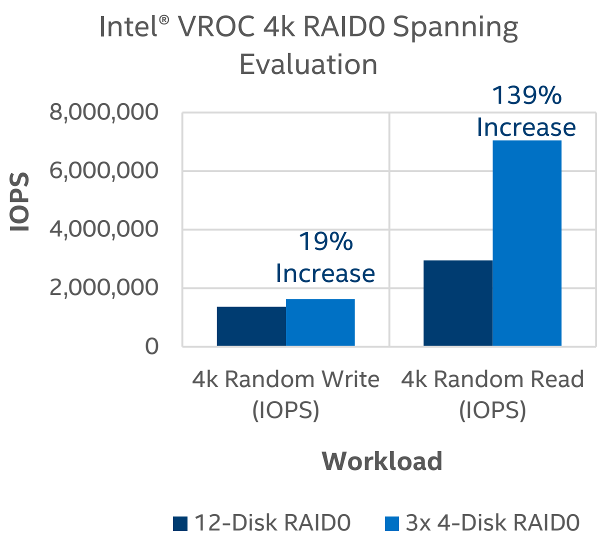


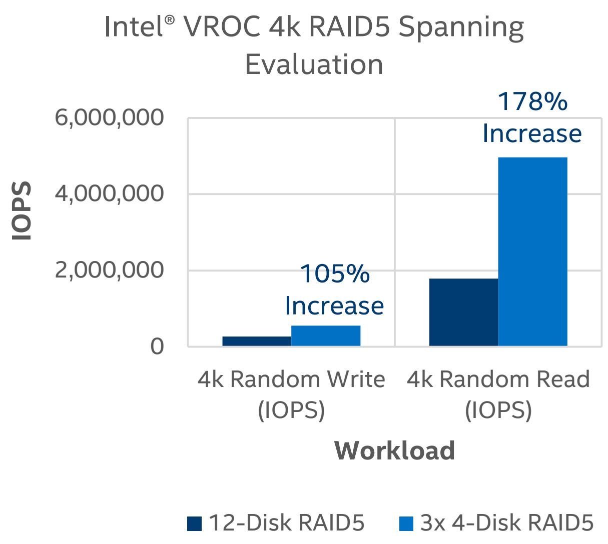
To show how spanning can impact performance on smaller block sizes, the same 4k tests were conducted with 3 separate 4 Disk RAID arrays.

This way, no single RAID array has SSDs on 2 different CPUs.

The aggregate IOP performance of these 3 RAID arrays was compared to the performance of a single 12 Disk RAID array.

Intel® VROC Spanning Proof Point[1]





By comparing a single 12 Disk RAID array with 3x 4 Disk RAID arrays, the effects of spanning can be evaluated:

- IOPS numbers for the 3x 4-Disk arrays is the sum of each arrays individual performance
- Primary impact is on Reads due to higher IOPs, but all workloads see an improvement
- By avoiding spanning, overall system storage performance can be increased for smaller block sizes (4k)

See appendix for footnotes

Conclusion

Intel® Virtual RAID on CPU does scale, but the below configuration details should be considered in order to optimize performance:

- Avoid having single RAID volumes span across CPUs. Break up larger RAID arrays to have multiple arrays on separate CPUs. If larger RAID arrays are needed, use switches to keep them on one CPU. However, the switch uplink may have its own limits on performance.

- Confirm block size of the workload. Sometimes transactions can be the bottleneck, not throughput. Larger block sizes will create fewer transactions, better scaling, and better throughput.

- Confirm workload is large enough to scale performance. Intel® VROC with NVMe* SSDs can handle high bandwidth storage needs, but only if the work is available. Testing should be done with higher Queue Depths than typical benchmark tests in order to saturate many NVMe SSDs. The tests in this document were run with 32 Threads and 128 IODepth.

- Intel® VROC performance in Windows* is limited to ~1M IOPS, which can prevent scaling with smaller block sizes. Using larger block sizes or Linux OS* distributions will have more effective scaling results.

Translation to Real Life…

These results are meant to demonstrate how Intel® VROC scales performance with NVMe* SSDs. The drive counts compared and workload used are a reference point. Results will vary, but consider the below for any installation:

1. Application workloads use variable and inconsistent block/packet sizes. Configuring settings to maximize block size may help with RAID storage performance. However, there may be tradeoffs to other system functions.

2. Scaling to 8+ NVMe SSDs requires a large data stream to saturate all of the storage devices. Hosting multiple VMs/applications on a single server or consolidating storage processes to a single server are examples of usage models that drive up storage bandwidth requirements.

Appendix

1. **System configuration**: Intel® Server Board S2600WFT family, Intel® Xeon® 8170 Series Processors, 26cores@ 2.1GHz, RAM 192GB , BIOS Release 7/09/2018, BIOS Version: SE5C620.86B.00.01.0014.070920180847
OS: RedHat* Linux 7.4, kernel- 3.10.0-693.33.1.el7.x86_64, mdadm - v4.0 - 2018-01-26 Intel build: RSTe_5.4_WW4.5, Intel ® VROC Pre-OS version 5.3.0.1039,
NVMe SSDs: Intel® SSD DC P4510 Series 2TB drive firmware: VDV10131, Retimer
BIOS setting: Hyper-threading enabled, Package C-State set to C6(non retention state) and Processor C6 set to enabled, P-States set to default and SpeedStep and Turbo are enabled
Workload Generator: FIO 3.3, RANDOM: Workers-32, IOdepth- 128, No Filesystem, CPU Affinitized
Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.
Performance results are based on testing as of January 11, 2019 and February 13, 2019 (Intel® VROC Spanning Proof Point) and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

intel®