Introducing the Intel® Optane™ SSD P1600X Series for Boot Drives

This Intel® SSD is optimized for cost-efficient capacity, high performance, and excellent reliability, providing an ideal data center boot drive solution.

Authors
Allison Goodman
Sr. Principal Engineer

Jacek Wysoczynski
Sr. Manager of Product Planning

Alessandro Goncalves
Optane Solution Architect

Esmond Chang
Optane Solution Architect

Table of Contents

Executive Summary

Data center servers are usually equipped with SATA NAND-based SSDs as boot drives. But these drives may raise concerns about performance, reliability, and cost, especially for today's data-hungry, mission-critical workloads. Keep reading to find out how using an Intel® Optane™ SSD P1600X Series as a boot drive can increase data center efficiency through right-sized capacity, excellent endurance, and up to 4.4x better read, write, and read/write performance than a SATA SSD.[1] The Intel Optane SSD P1600X Series may also act as a RAID 1 boot volume with a storage RAID card or a host bus adapter (HBA).

Overview of Boot Drives and Interface Protocols

Boot drives play a critical role in loading the OS during system power-on. They also provide file system access during day-to-day operations. When choosing a boot drive, system architects look for three things: cost efficiency, performance, and reliability. If a boot drive fails, the server will crash. In a 2020 study, 29 percent of respondents said that the hourly cost of a critical server outage was more than USD 1M,[2] and another 48 percent put that cost somewhere between USD 300,000 and USD 1M. For mission-critical applications, OS-loading and file system access latency can be a concern. And as economic pressures continue to mount in the data center, getting optimal performance and reliability for an affordable price is always the goal.

Boot drives come in a variety of protocol interfaces and form factors. The majority of today's boot drives are SSDs that use the SATA interface. This protocol has been around for more than a decade and runs on an older AHCI controller specification that supports HDDs. However, when paired with devices built with flash memory media technology used in SSDs, performance is bottlenecked by the SATA interface due to the protocol overhead that was created to support the rotating platter of HDDs.

While SAS SSDs offer faster data transfer rates than SATA SSDs (see Figure 1), they are not commonly used as boot drives. Today, the best option for boot drives is the high-performance host-controller Nonvolatile Memory Express (NVMe) protocol running on top of a Peripheral Component Interconnect Express (PCIe) interface. The NVMe interface was designed specifically for flash memory and can greatly improve SSD performance compared to SATA protocols—providing up to 6.5x better throughput compared to the latest SATA generation.[3]

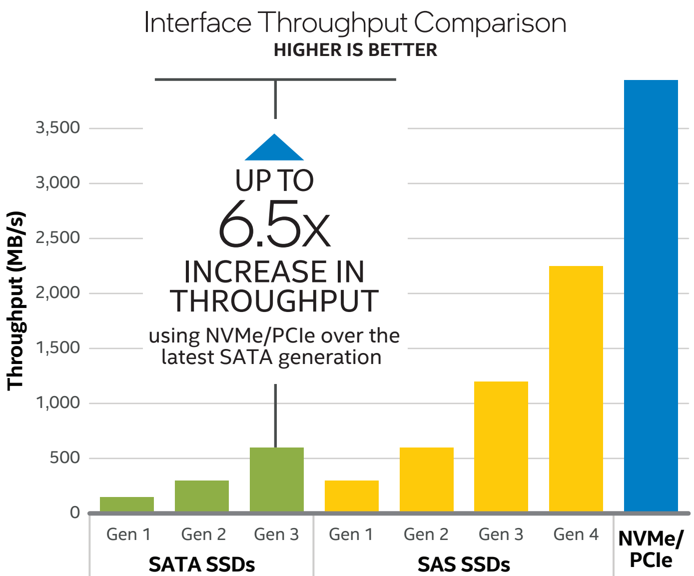

Figure 1. Comparison of SSD throughput by interface type.[3]

Many data center system architects are recognizing the benefits of NVMe SSDs. The transition from SATA or SAS to NVMe is quickly progressing in the storage industry, with the majority of SSDs moving to the NVMe interface in 2022, as shown in Figure 2.

Boot drives typically come in two form factors: M.2 and 2.5-inch. The M.2 form factor SSD was designed to be thin and lightweight (smaller than the U.2 form factor) and has gained popularity in client, embedded, and server markets where space in the system is constrained.

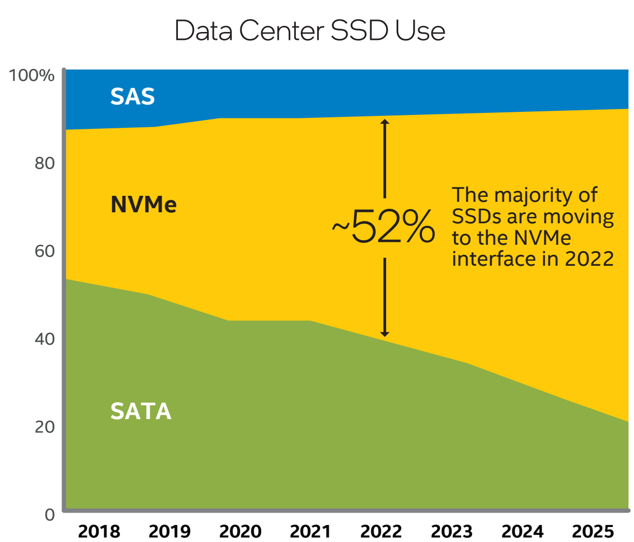

Figure 2. Data centers are transitioning away from SATA SSDs to NVMe SSDs.[4]

A Better Boot Drive: Intel® Optane™ SSD P1600X Series

The Intel® Optane™ SSD P1600X Series is a new data center-optimized M.2 NVMe PCIe SSD in 58 GB and 118 GB capacities that delivers the high I/O performance, low latency, and small capacity desired in an ideal boot drive solution. This SSD supports NVMe1.1 on a PCIe Gen3x4 interface and delivers throughput up to 1.8 GB/s to provide high I/O performance for data access.[5] It comes in a M.2 22x80mm form factor that can easily plug in to a server board without an extra cable (unlike the U.2 form factor, which must connect to the backplane using a cable). The slim design and direct connection make the **Intel Optane P1600X Series** an excellent choice in data centers.

For the three key selection criteria in boot drives, the Intel Optane SSD P1600X Series provides clear advantages over SATA-based NAND SSD boot drives:

- **Cost efficiency.** The capacity of the Intel Optane P1600X Series is optimized to serve as a boot drive, so money isn't wasted on excess capacity.
- **Performance and value.** With faster boot, OS installation, and file access, mission-critical applications get what they need when they need it, with excellent performance-per-dollar (IOPS/$). The performance of the Intel Optane SSD P1600X enables it to also accelerate performance of logging drives.
- **Reliability.** High endurance means data center admins worry less about downtime.

The following sections provide more detail about the advantages of using the Intel Optane SSD P1600X Series as a boot drive.

What Makes Intel® Optane™ SSDs Different from NAND SSDs?

Intel® Optane™ SSDs are built on a unique architecture that allows memory cells to be individually addressed in a dense, transistor-less, stackable design.

The consistently high performance and excellent endurance of Intel Optane SSDs are enabled by the unique Intel® Optane™ Memory media that supports write-in-place and offers several benefits compared to NAND technology:

- Increased performance by avoiding the compute resource overhead associated with garbage collection
- Elimination of the extra writes needed in NAND-based systems
- No write amplification concerns, such as optimizing for sequential writes over random writes and leaving sufficient free space

Pay Only for the Boot Drive Capacity You Need

SATA SSDs have been steadily growing in capacity (see Figure 3).[6] The typical SATA boot drive is now at least 240 GB. However, the most popular data center server operating systems (like Microsoft Windows Server, Red Hat Enterprise Linux, SUSE Linux Enterprise Server, and Ubuntu) don't need 240 GB, or even 128 GB.[7] All these operating systems have system storage requirements of 32 GB or less (see Table 1).[8] That means that system architects are paying for capacity they don't need. Many data centers simply let that excess capacity lie idle. Others may try to partition the boot drive and use excess capacity for data and logging. But as previously described, SATA drives' storage performance is insufficient for today's data-hungry applications and workloads, and SATA or NVMe-based NAND drives don't have the endurance ideal for a logging device.

The Intel Optane SSD P1600X Series is available in capacities of 58 GB and 118 GB, which meet the storage requirements of the major operating systems without over-provisioning.

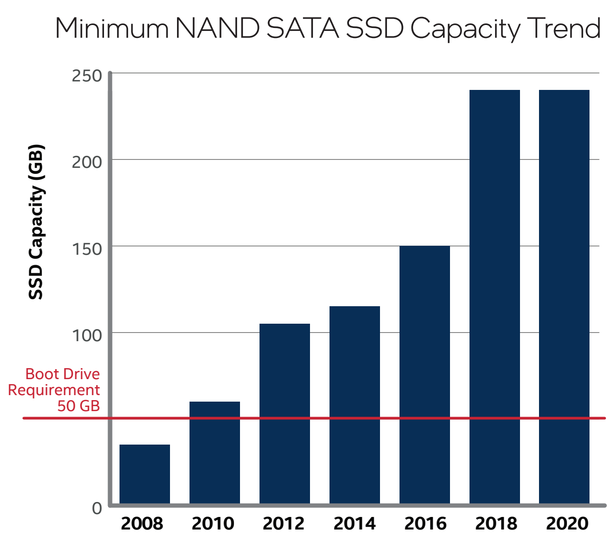

Figure 3. The capacity of NAND SATA SSDs is growing far greater than is required for boot drives.[6,7]

Table 1. Popular OS Capacity Recommendations

Operating System	Boot Drive Capacity Recommendation
Red Hat Enterprise Linux v7 and v8	20 GB
SUSE Linux Enterprise Server v12 and v15	32 GB
Ubuntu 20 LTS	25 GB
Microsoft Windows Server 2019	32 GB

Boost Boot Drive Performance and Value

Performance can be measured by throughput (higher is better) and by latency (lower is better). The Intel Optane SSD P1600X Series provides consistent performance (both throughput and latency) across all the read-heavy, write-heavy, and mixed random read/write workloads that are common for boot drives. The consistent performance can help substantially reduce both average storage latency and the infrequent—but potentially harmful—longer latencies that occur under heavy system load.

Excellent throughput and IOPS/$

Random read/write with small block size is common when accessing the OS file system. As shown in Figure 4, this Intel Optane SSD delivers more than 4.4x the 4 KB random read/write performance, with a similar increase in IOPS/$, compared to a SATA-based NAND SSD.[9] High I/O performance is important in speeding up file system access after the server has booted and the OS has been installed.

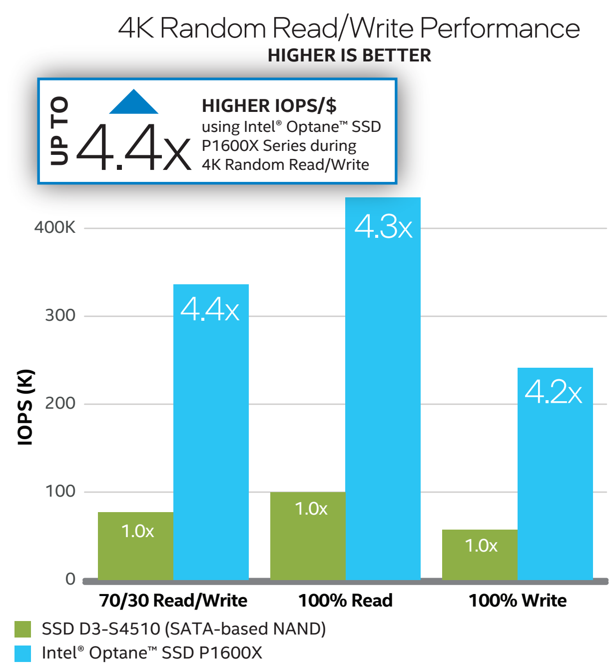

Figure 4. For all workload types, the Intel® Optane™ SSD P1600X significantly outperforms the SSD D3-S4510, with much better IOPS/$.

Shortened boot and install processes

Our tests showed that compared to a typical SATA-based NAND boot drive, the Intel Optane SSD P1600X Series shortened the overall system boot time by up to 7 percent and improved the OS installation time up to 12 percent.[10]

Consistently low latency

The OS on the boot drive handles all application and file access. Mission-critical workloads can't afford long response times. The average read latency for the Intel Optane SSD P1600X Series is 7 µs, which is 4.8x lower than 34 µs for the SSD D3-S4510. In comparison, the average write latency for the Intel Optane SSD P1600X Series is 10 µs, which is 3.6x lower than the 36 µs for the SSD D3-S4510.[11]

Get More Out of Your Boot Drive

The unique write-in-place and high endurance qualities (see Improve Boot Drive Reliability) make the Intel Optane SSD P1600X Series an ideal choice when optimizing boot drive value, because it supports multiple workloads and mixed usages.

As mentioned earlier, NAND drives do not have the necessary endurance to serve as logging devices. In contrast, the Intel Optane SSD P1600X Series is a six drive write per day (DWPD) device with 3x higher endurance than the SSD D3-S4510.[12] With such high endurance, this drive can be partitioned so that excess capacity not needed for boot drive processes can be used to run a write-intensive workload such as metadata logging.

In the Open Zettabyte file system (OpenZFS), when the system has synchronous writes to storage, write completion notifications are delayed until the file system receives acknowledgement from the storage device. This delay results in poor overall system write performance. To address this issue, ZFS Intent Log and Separate ZFS Intent Log (ZIL and SLOG) act as write cache prior to the synchronization operation. Fast and persistent storage is critical so that the ZIL can improve the overall write performance. The excess capacity of Intel Optane P1600X can be partitioned as a cache drive to accelerate the logging transactions of the SSD in the system tasked with handling these operations. Our tests revealed that using a 10 GB partition of an Intel Optane SSD P1600X Series as a FreeBSD ZFS ZIL/SLOG cache device improves synchronous write performance by up to 207 percent (see Figure 5) and cuts latency by up to 67 percent of the transfer rate for a 4 KB transfer size (see Figure 6).[13]

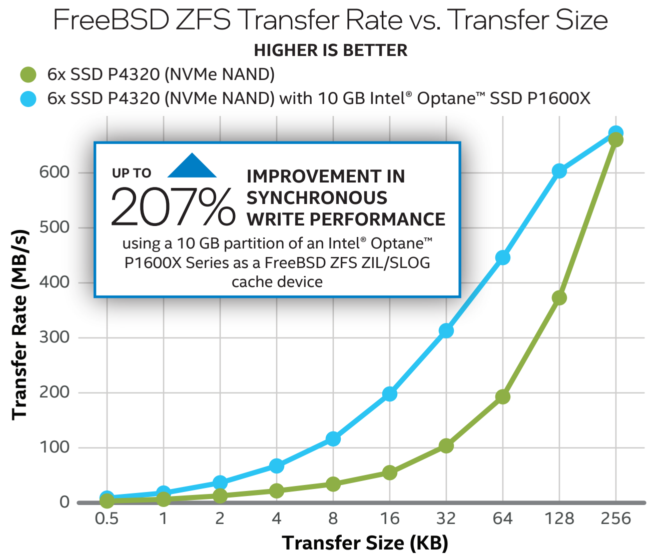

Figure 5. Using a partitioned Intel® Optane™ SSD P1600X Series as a logging device boosts transfer rate in FreeBSD ZFS.

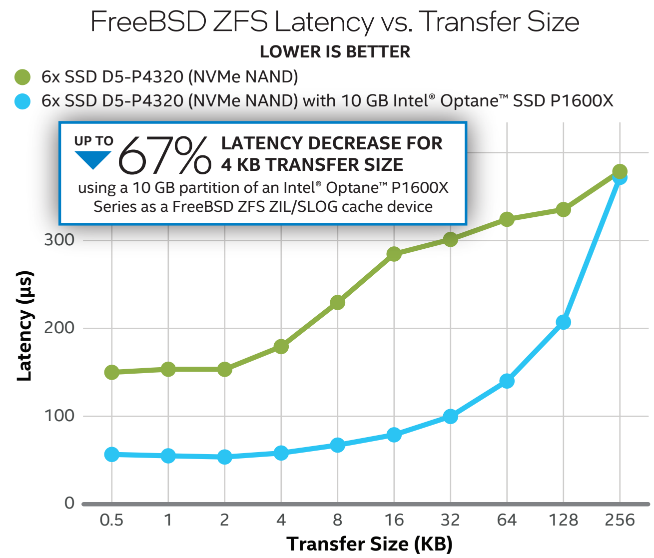

Figure 6. Using a partitioned Intel® Optane™ SSD P1600X Series as a logging device keeps latency low in FreeBSD ZFS.

Improve Boot Drive Reliability

The Intel Optane SSD P1600X Series has valuable endurance-related advantages that set it apart from NAND SSDs, especially considering the long lifetime and high reliability needs of the boot drive. For example, the Intel Optane SSD P1600X Series has up to 4.6x higher endurance than an SSD D3-S4510.[14] The Intel Optane SSD P1600X Series' high endurance contributes to improved operational costs: improved endurance can translate directly into lower total cost of ownership without sacrificing performance.

In addition, the write-in-place architecture means significantly less firmware management and buffering in the data path. As a result, unplanned power loss is not as disruptive to Intel Optane SSDs as it is to NAND SSDs, reducing error rates that require rebuilds, or that result in partial data loss or drive failures.

The Intel Optane SSD P1600X Series meets and exceeds the reliability requirements compared to NAND SSDs based on the JESD218 standard. Examples include the following:[15]

- UBER smaller than 1 sector per 10^17 bits read
- MTBF of 2 million hours
- Six DWPD
- Five-year warranty

The Intel® Memory and Storage Tool (Intel® MAS)[16] drive management software provides a graphical user interface to configure and retrieve data on the Intel Optane SSD P1600X Series. This tool enables the execution of NVMe format commands, parsing device log data, and dumping NLogs, event logs, and SMART (Self-Monitoring, Analysis and Reporting Technology) information.

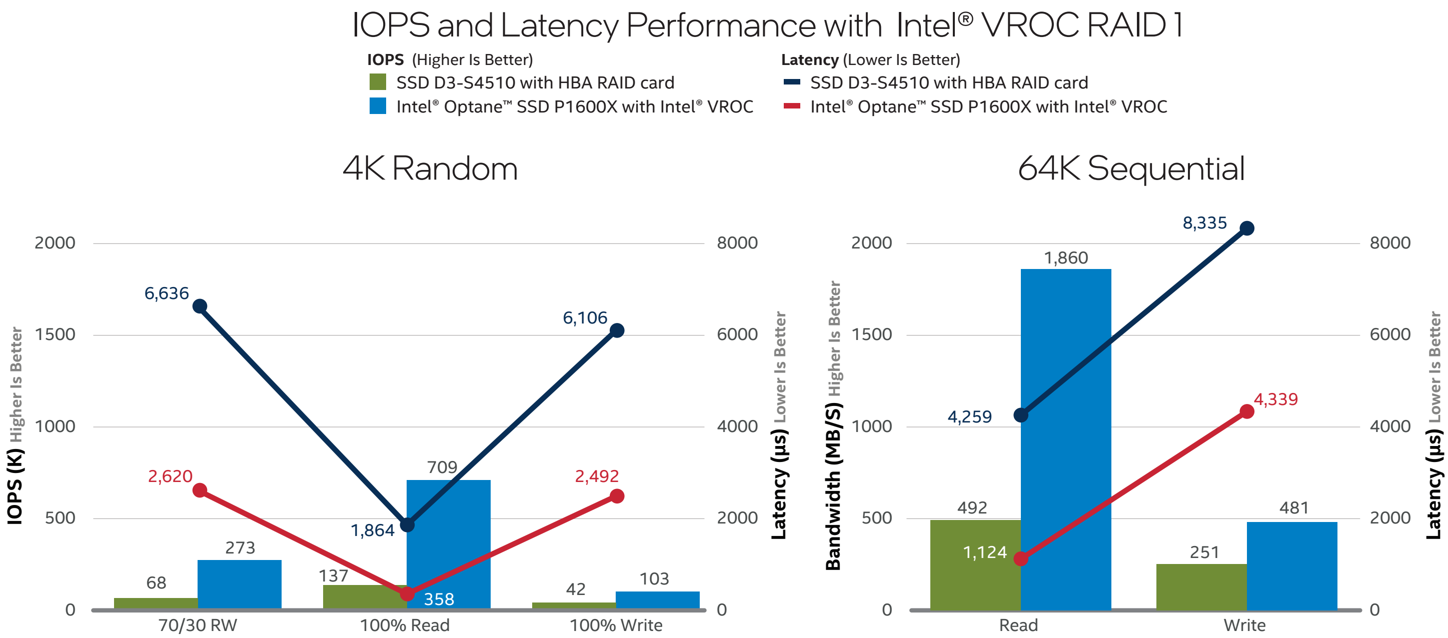

Figure 7. Using Intel® Optane™ SSD P1600X with Intel® VROC shows performance and latency improvement compared to an HBA RAID card solution.[18]

Achieve Bootable RAID 1 While Lowering Cost and Complexity

To help protect against single drive failures on boot drives and to maintain system serviceability, it is common to use a RAID 1 to mirror data on boot drives. The Intel Optane SSD P1600X Series could be used as a RAID 1 boot volume with a storage RAID card or a host bus adapter (HBA). However, Intel also provides a solution that enables customers to use NVMe-based SSDs in RAID 1 mirrored configurations without extra hardware cables or add-in cards.

This solution uses Intel® Volume Management Device (Intel® VMD), which is an Intel® Xeon® Scalable processor feature that enables direct control and management of NVMe SSDs from the PCIe bus without additional hardware adaptors. The combination of Intel VMD with the Intel® Virtual RAID on CPU (Intel® VROC) technology hosts SSDs as a hybrid RAID which presents boot drives as a RAID volume in BIOS for OS installation and boot.

Using Intel VMD and Intel VROC, along with UEFI components, simplifies the data center by avoiding expensive, power-hungry HBAs while also increasing overall solution reliability with fewer components.[17]

When an Intel Optane SSD P1600X and Intel VROC are used together, performance and latency improve compared to a solution that includes an HBA RAID card (see Figure 7).[18]

The Intel Optane SSD P1600X as ZFS ZIL/SLOG could be used in two scenarios with Intel VROC to improve write performance (see Figure 8). In the top scenario, SATA SSDs and a conventional HBA RAID card are used for boot, while the Intel Optane SSD P1600X with the Intel VROC RAID 1 volume is used as the ZFS ZIL/SLOG. However, the bottom scenario is highly recommended because it simplifies infrastructure and therefore provides more value from the Intel Optane SSDs. Here, the Intel Optane SSD P1600X in 2-disk RAID 1 are used for boot, with a partition reserved for use as the ZFS ZIL/SLOG.

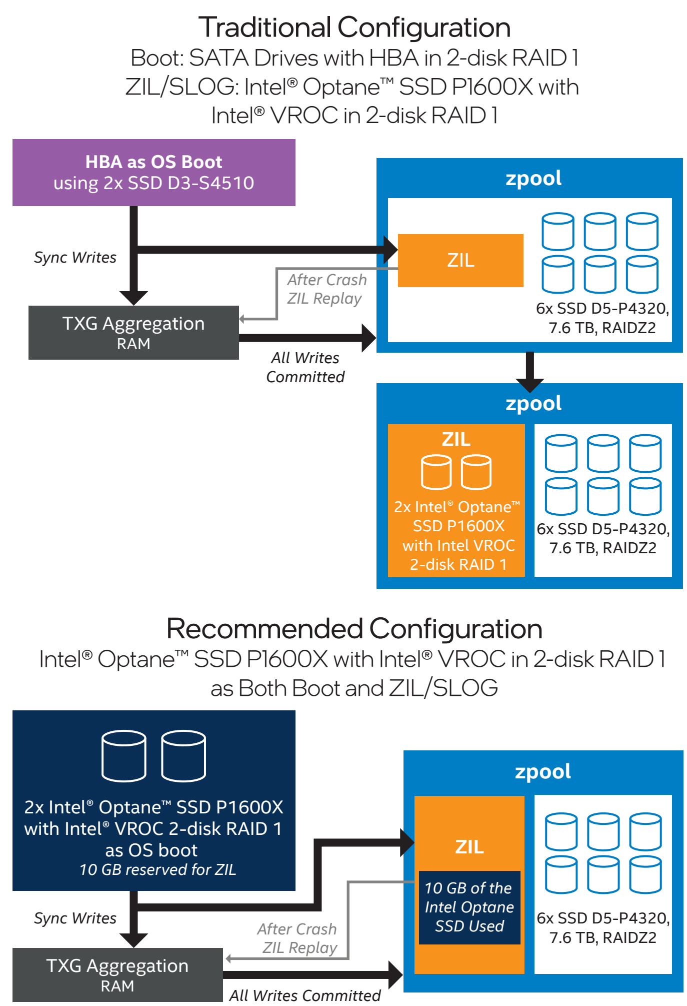

Figure 8. Instead of the traditional configuration, we recommend using Intel® Optane™ SSD P1600X with Intel® VROC as boot, with a 10 GB partition reserved for ZFS ZIL/SLOG.

As shown in Figure 9, using the Intel Optane SSD P1600X as the ZFS ZIL/SLOG provides a performance gain of at least 55 percent in both scenarios.[19]

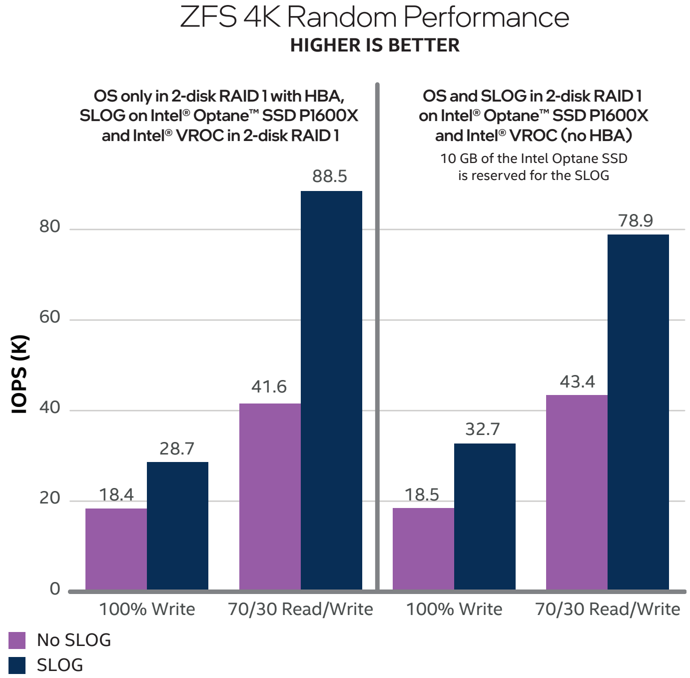

Figure 9. A performance gain of of at least 55 percent can be achieved when using an Intel® Optane™ SSD P1600X as the ZFS ZIL/SLOG.[19]

One of the most important aspects of using a RAID 1 mirror volume is to protect data integrity. Data can be rebuilt when one of the RAID member drives fails. It also allows rebuilding of redundant RAID volumes on the system when a drive failure occurs. However, the data will be at risk when the RAID volume is in degrade mode before the rebuild is complete. The Intel Optane SSD P1600X with Intel VROC provides much faster rebuild performance compared to the conventional HBA RAID card with SATA drives (see Figure 10).[20]

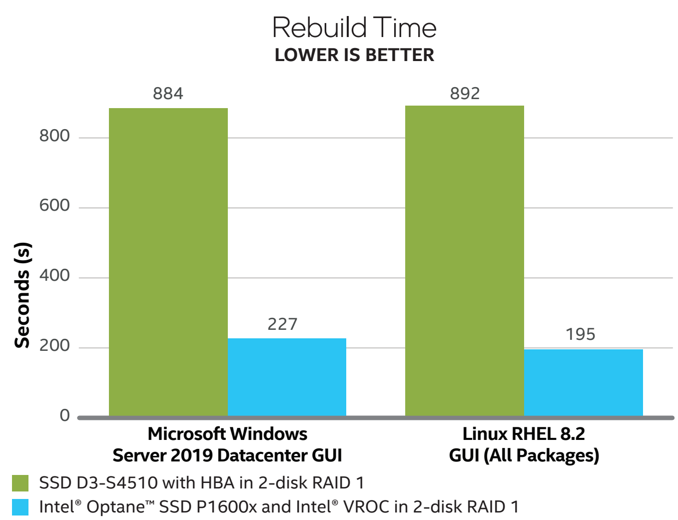

Figure 10. Using an Intel® Optane™ SSD P1600X with Intel® VROC shortens rebuild time, reducing risk of data loss.[20]

Conclusion

The Intel Optane SSD P1600X Series is a data center-optimized M.2 NVMe/PCIe solution built on unique Intel Optane technology. This SSD's high I/O performance and low latency give it a significant IOPS/$ advantage over SATA-based NAND drives. These benefits, combined with right-sized capacity, high-endurance, and manageability, deliver an optimal boot drive solution. The Intel Optane SSD P1600X Series may also act as a RAID 1 boot volume with a storage RAID card or an HBA. In addition, the unique characteristics of Intel Optane SSDs enable extra value by enabling the boot drive to also be used for logging/caching and storage. The Intel Optane SSD P1600X Series is the beginning of a smaller-capacity SSD product line for cost-sensitive deployments.

1 Testing by Intel as of May 28, 2021. **System configuration:** Intel Coyote Pass Platform, 2x Intel® Xeon® Platinum 8368 processor (38 cores, 2.4 GHz), DRAM 256 GB, BIOS Version: SE5C6200.86B.0022.D08.2103221623, 1x Intel® Optane™ SSD P1600X Series 58 GB M.2, 1x SSD D3-S4510 240 GB U.2 (2.5"). **OS:** Red Hat Enterprise Linux v8.2, 4.18.0-193.el8.x86_64, FIO version= 3.26; 4K random R/W, I/O depth=32, numjobs=8. **BIOS setting:** Intel® Hyper-Threading Technology ON, Enhanced Intel SpeedStep® Disabled, Intel® Turbo Boost Technology ON, PackageC-State Processor C6 disabled, HardwareP-States disabled. Pricing data as of June 25, 2021. P1600X and S4510 price is Intel MSRP pricing; other component pricing are estimates based on common industry pricing.

2 Statista, "Global enterprise server hourly downtime cost 2019," Dec 7, 2020. https://www.statista.com/statistics/753938/worldwide-enterprise-server-hourly-downtime-cost

3 SAS, SATA, and NVMe specifications, available at https://sata-io.org, https://www.snia.org, and https://nvmexpress.org/specifications

4 Forward Insights Q1'21 SSD report

5 Intel® Optane™ P1600X Series Product Specification

6 Source: IEEE Technology Paper: Scaling Trends in NAND Flash and Wikipedia

7 While Table 1 shows OS capacity requirements up to 32 GB, other operating systems, such as other flavors of Linux, may require up to 50 GB.

8 For example, see Microsoft Windows Server 2019 server storage requirements: https://docs.microsoft.com/en-us/windows-server/get-started-19/sys-reqs-19 and Red Hat Enterprise Linux storage requirements: https://access.redhat.com/articles/rhel-limits.

9 See endnote 1.

10 Testing by Intel as of May 28, 2021. **System configuration:** Intel Coyote Pass Platform, 2x Intel® Xeon® Platinum 8368 processor (38 cores, 2.4 GHz), DRAM 256 GB, BIOS Version: SE5C6200.86B.0022.D08.2103221623, 1x Intel® Optane™ SSD P1600X Series 58 GB M.2, 1x SSD D3-S4510 240 GB U.2 (2.5"). **OS:** Red Hat Linux 8.2 Server GUI (all packages). Stopwatch to measure between installation start and complete. The boot time is the period between GRUB launch to GUI. **BIOS setting:** Intel® Hyper-Threading Technology ON, Enhanced Intel SpeedStep® disabled, Intel® Turbo Boost Technology ON, PackageC-State Processor C6 disabled, HardwareP-States disabled.

11 SSD D3-S4510 Series product info and Intel® Optane™ SSD P1600X Series

12 Meet the New Data Center-Optimized SSDs

13 Testing by Intel as of May 28, 2021. **System configuration:** Intel Coyote Pass Platform, 2x Intel® Xeon® Platinum 8368 processor (38 cores, 2.4 GHz), DRAM 256 GB, BIOS Version: SE5C6200.86B.0022.D08.2103221623, 1x Intel® Optane™ SSD P1600X Series 58 GB M.2, 6x SSD P4320 7.68 TB. **OS:** FreeBSD* 12.2 ZFS. **BIOS setting:** Intel® Hyper-Threading Technology ON, Enhanced Intel SpeedStep® disabled, Intel® Turbo Boost Technology ON, PackageC-State Processor C6 disabled, HardwareP-States disabled.

14 See endnote 12.

15 See endnote 5.

16 Intel Memory and Storage Tool

17 Integrated Enterprise RAID Optimized for NVMe SSDs

18 Testing by Intel as of October 12, 2021. **System configuration:** Intel Coyote Pass Platform, 2x Intel® Xeon® Platinum 8368 processor (38 cores, 2.4 GHz), DRAM 256 GB, BIOS Version: SE5C6200.86B.0022.D64.210522049, 1x Intel® Optane™ SSD P1600X Series 118 GB M.2 with Intel® Volume Management Device and Intel® Virtual RAID on CPU, 1x SSD D3-S4510 Series (240 GB, SATA, HBA Intel RSP3TD160F, 2-disk RAID 1). **OS:** Red Hat Enterprise Linux v8.2, 4.18.0-193.el8.x86_64, FIO version= 3.19. **BIOS setting:** Intel® Hyper-Threading Technology ON, Enhanced Intel SpeedStep® disabled, Intel® Turbo Boost Technology ON, Package C-State Processor C6 disabled, Hardware P-States disabled.

19 Testing by Intel as of October 12, 2021. **Common configuration:** Intel Coyote Pass Platform, 2x Intel® Xeon® Platinum 8368 processor (38 cores, 2.4 GHz), DRAM 256 GB, BIOS Version: SE5C6200.86B.0022.D64.210522049, 1x Intel® Optane™ SSD P1600X Series 118 GB M.2, 6x SSD P4320 7.68 TB. **OS:** RHEL8.2 with OpenZFS on linux. **BIOS setting:** Intel® Hyper-Threading Technology ON, Enhanced Intel SpeedStep® disabled, Intel® Turbo Boost Technology ON, Package C-State Processor C6 disabled, Hardware P-States disabled. **Configuration 1:** 2x SSD S4510 with HBA as OS boot. 2x Intel SSD P1600X with Intel VROC 2-disk RAID1 as ZIL SLOG. **Configuration 2:** 2x Intel SSD P1600X with Intel® Virtual RAID on CPU 2-disk RAID1 as OS boot and ZIL SLOG.

20 Testing by Intel as of October 12, 2021. **Common configuration:** Intel Coyote Pass Platform, 2x Intel® Xeon® Platinum 8368 processor (38 cores, 2.4 GHz), DRAM 256 GB, BIOS Version: SE5C6200.86B.0022.D64.210522049. **OS:** Windows Server 2019 Datacenter and Red Hat Enterprise Linux 8.2. **BIOS setting:** Intel® Hyper-Threading Technology ON, Enhanced Intel SpeedStep® disabled, Intel® Turbo Boost Technology ON, Package C-State Processor C6 disabled, Hardware P-States disabled. **Configuration 1:** 2x Intel® SSD P1600X 118G with Intel® Virtual RAID on CPU 2-disk RAID 1. **Configuration 2:** 2x SSD D3-S4510 240G Series with HBA 2-disk RAID 1.