Product Brief

Data Center
Intel® Optane™ Persistent Memory 200 Series

Achieve Greater Insight from Your Data with Intel® Optane™ Persistent Memory

Intel Optane Persistent Memory 200 Series Delivers 32 Percent More Bandwidth on Average[1] with up to 6 TB Total Memory per Socket[2]

Over 90 percent of enterprises are in the midst of their digital transformation journey[3] to become data-centric businesses. Thus, they will need to capture, analyze, and secure increasing amounts of data. Data growth is accelerating the demand for performant, data-intensive computing. As the demand for compute grows, memory capacity in the system typically needs to scale along with it. Large pools of DRAM help accelerate computing with low latency, but DRAM is limited in capacity, is volatile, and expensive. DRAM is becoming one of the most expensive components in modern server bills of materials. Alternatively, block storage is large, cheap, and persistent, but slow to bring data to the CPU.

Intel® Optane™ persistent memory (PMem) bridges the gap with an innovative memory technology. This new memory delivers a unique combination of affordable large capacity and support for data persistence. With 3rd Gen Intel® Xeon® Scalable processors and Intel Optane PMem 200 series workloads can optimize performance and cost by creating a 2-tier hierarchy in memory and storage. Supported for these processors on 2-socket and 4-socket platforms, PMem is helping to turn more data into actionable insights.

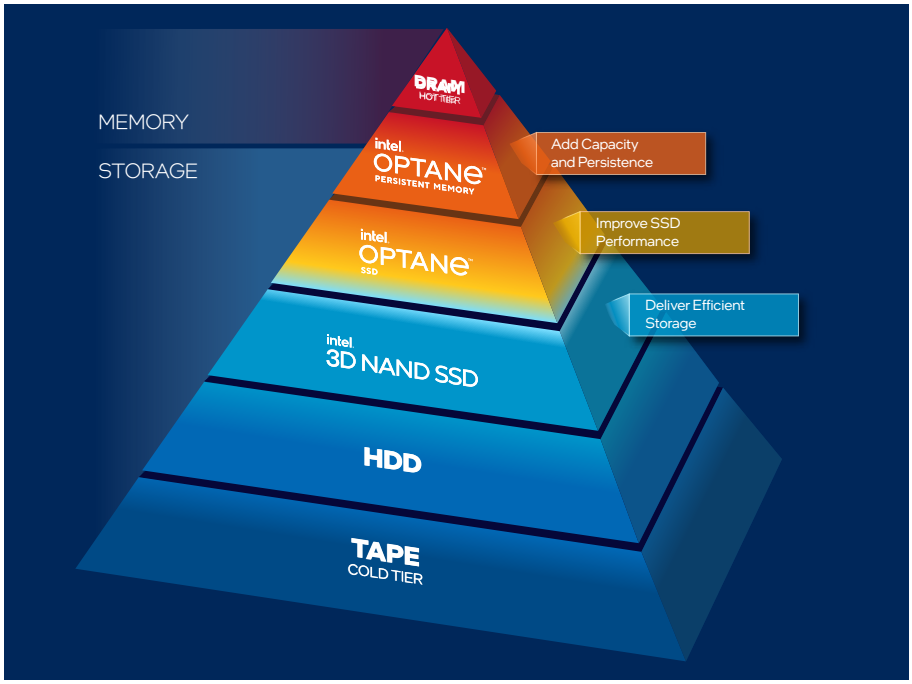


Figure 1. Intel® Optane™ persistent memory enables hierarchical architectures for high-performance, large memory computing.

Intel Optane Persistent Memory 200 Series

Intel Optane PMem 200 series coexists side-by-side with system memory, occupying existing DRAM slots. It delivers an average of up to 32 percent more memory bandwidth than the previous generation.[1] Available in 128 GB, 256 GB, and 512 GB modules, it offers both large capacity and persistence that enable new platform architectures to:

- Help accelerate large-memory computing by keeping more data closer to the CPU
- Accelerate restart times with reduced I/O by persisting data in memory and not require reloading from storage
- Reduce power consumption for large-memory nodes

3rd Gen Intel Xeon Scalable processors and Intel Optane PMem 200 series create a powerful foundation for optimizing platforms for multiple types of workloads. With 3rd Gen Intel Xeon Scalable processors for 2-socket and 4-socket platforms, each memory channel can support one Intel Optane PMem 200 series module (six channels on 4-socket platforms and eight channels on 2-socket platforms). Each module draws only a maximum 15 watts of power, offering high capacity without high power demand and creating large-memory platforms to power data-centric businesses.

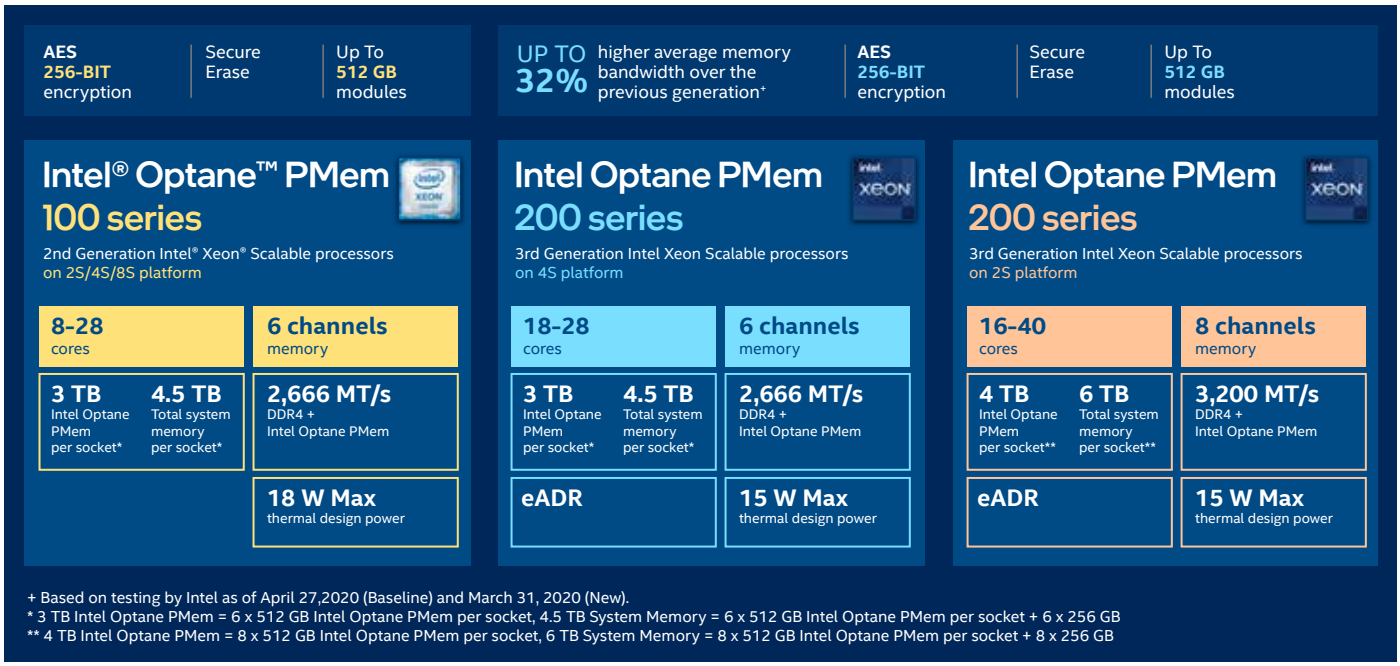


Figure 2. 3rd Gen Intel® Xeon® Scalable processors and Intel® Optane™ PMem 200 series create a powerful foundation for large-memory computing.

Data-intensive and compute-intensive workloads, from cloud to databases, in-memory analytics, virtualized infrastructure, content delivery networks, and more, can easily take advantage of large-scale and persistent memory. Enabling these large memory pools helps speed time to insights that may form the basis of critical business decisions, cost savings, and new revenues.

The new series of PMem is compatible with the software ecosystem already established for previous-generation Intel Optane persistent memory. Thus, migrating to or adding new systems built on 3rd Gen Intel Xeon Scalable processors with Intel Optane PMem 200 series is seamless and transparent to the software designed for the previous generation. With the industry-standard persistent memory programming model, developers can build simpler and more powerful applications to prepare their data center investment for the future.

Affordable Large Capacity

Intel Optane PMem 200 series enables more value to be extracted from larger data sets and increases the utility of each server. In-memory databases can access more data at DRAM-like speeds, and workloads processing massive data sets, such as scientific or data warehousing and analytics, can work continuously without repeatedly loading and storing data locally. Additionally, Intel Optane PMem can offer greater memory capacity per socket than DRAM for virtualized data center infrastructures. PMem leaves more headroom for virtualizing future workloads requiring larger memory capacity rather than having to run those demanding workloads on bare metal. When deployed, Intel Optane PMem 200 series can enable you to consolidate and reduce your server footprint, leading to lower software licensing costs, reduced power consumption, and other operational efficiencies.

Secure Data at Rest for Enhanced Data Protection

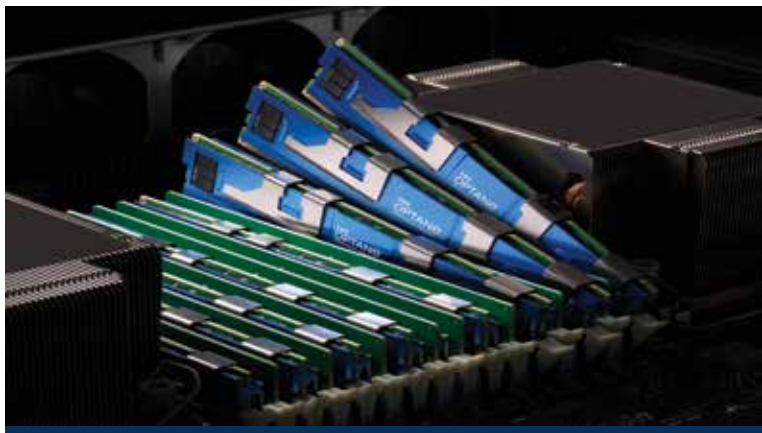
With Intel Optane PMem 200 series, data at rest is better protected by strong, industry-standard security measures. All data is encrypted by industry-standard 256-bit Advanced Encryption Standard (AES-256). With hardware-based cipher key processing, encryption is transparent to application software. Thus, no software code changes are needed when adding PMem. Hardware-enabled encryption on 3rd Gen Intel Xeon Scalable processors results in strong, industry-standard data security with low impact on performance.

Data Persistence and Automatic Cache Flushes Enhance Operational Efficiencies

Unlike DRAM, Intel Optane PMem 200 series retains data during a planned or unplanned restart, avoiding time-consuming data reloads. Maintaining data in PMem, means less down time, fewer losses from system outages, and increased operational efficiency.

As described in the persistent memory programming model, applications that manage data structures in persistent memory routinely call cache-flushing commands to move data stored in the CPU cache to persistent memory. Waiting for these cache flushes to complete can reduce performance. Applications can now avoid those waits altogether with extended asynchronous DRAM refresh (eADR), a new platform feature available with Intel Optane PMem 200 series with 3rd Gen Intel Xeon Scalable processors. The application automatically detects eADR and skips cache flushing, knowing the system will do it automatically, even if it experiences a system crash or power failure. This means application “Stores” are considered persistent as soon as they are visible to the application, which brings “lock-free” programming to persistent memory. eADR helps free up resources and improves performance.

Using eADR with Intel Optane PMem requires platform hardware support via the ability to have additional stored energy available to allow cache flushes to occur on power failure. This coupled with the application checking for the enablement of eADR and following the persistent memory programming model will ensure your solution is enabled fully for eADR support.



Create a high-performance, large-capacity persistent memory tier with Intel® Optane™ PMem 200 series to improve key data center metrics such as:

- Increased CPU utilization and utility of each server
- Increased application throughput with more memory capacity
- Increased VM density or support for more services and users
- Support for larger in-memory databases without prohibitive price tags
- Improve business resilience for systems with critical data
- Consolidate server footprint, reduce software licensing costs, and optimize return on your enterprise investment

Persistent Memory Addresses Many Key Challenges

Intel Optane PMem 200 series addresses many of the computing challenges that data centers face today. These challenges include high DRAM costs for large-memory nodes, data protection during outages and maintenance, emerging workloads that take advantage of hierarchical memory architectures, and more.

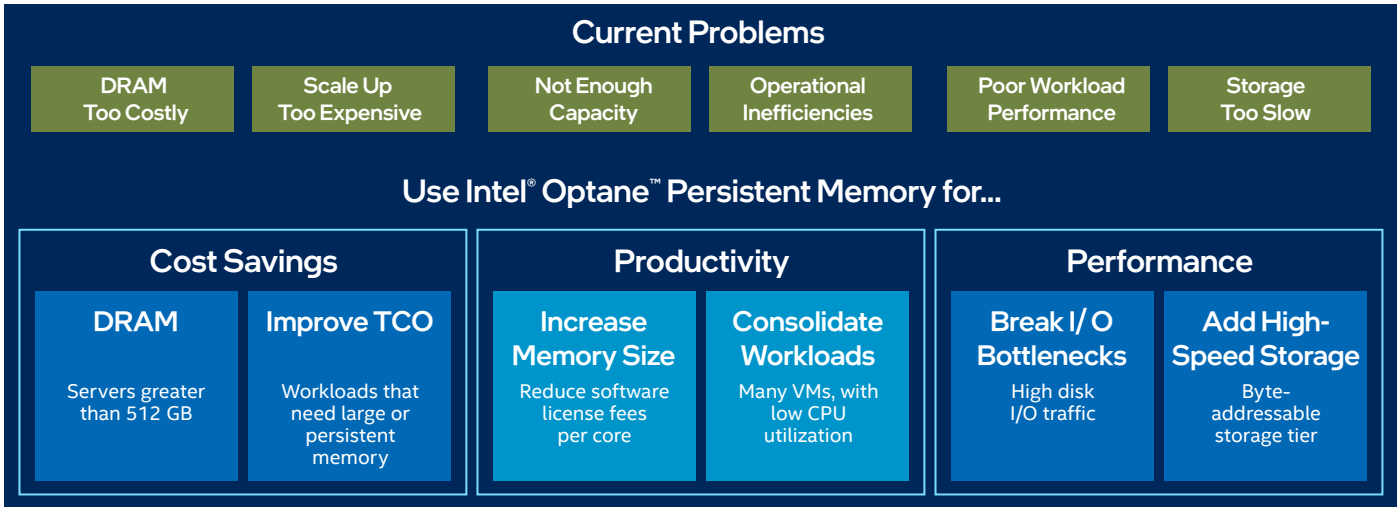


Figure 3. Intel® Optane™ PMem 200 series solves several key challenges in computing today.

Optimize Workloads with 2-Tier Memory and Storage Hierarchies

Intel Optane PMem offers system architects and application developers new options to create hierarchical memory and storage tiers to address data performance and capacity challenges. A hierarchical approach allows application developers to optimize the resources of the platform for data access and transport. Programmers can leverage the speed and proximity of technologies closest to the CPU, while taking advantage of the capacity available in the system.

For a two-tier memory model, low-latency DRAM offers exceptionally fast performance, while Intel Optane PMem 200 series creates large memory capacity to store and protect data with DRAM-like speed. Depending on the application, PMem used as the large memory capacity can be persistent or volatile.

The combination of DRAM with persistent memory:

- Allows fast in-memory computing on massive data sets.
- Allows for consolidating more virtual machines on a platform in a virtual environment with large performant capacity.
- Reduces risk of losing critical data when the application is persistent-memory aware.
- Speeds time-to-solution from large computations where intermediate results can be persisted and reloaded for final analysis.

With a two-tier storage model, Intel Optane PMem 200 series, utilized as a performance tier, delivers fast, byte-addressable access to most frequently-accessed data. Other technologies, such as SSDs, offer slower access to warm data storage in a capacity tier.

These hierarchical memory and storage architectures made possible by Intel Optane technology help optimize speed, latency, capacity, and cost. Choosing the right combinations for each application can help optimize systems and their workloads, from cloud to databases, in-memory analytics, virtualized infrastructure, content delivery networks, and more.

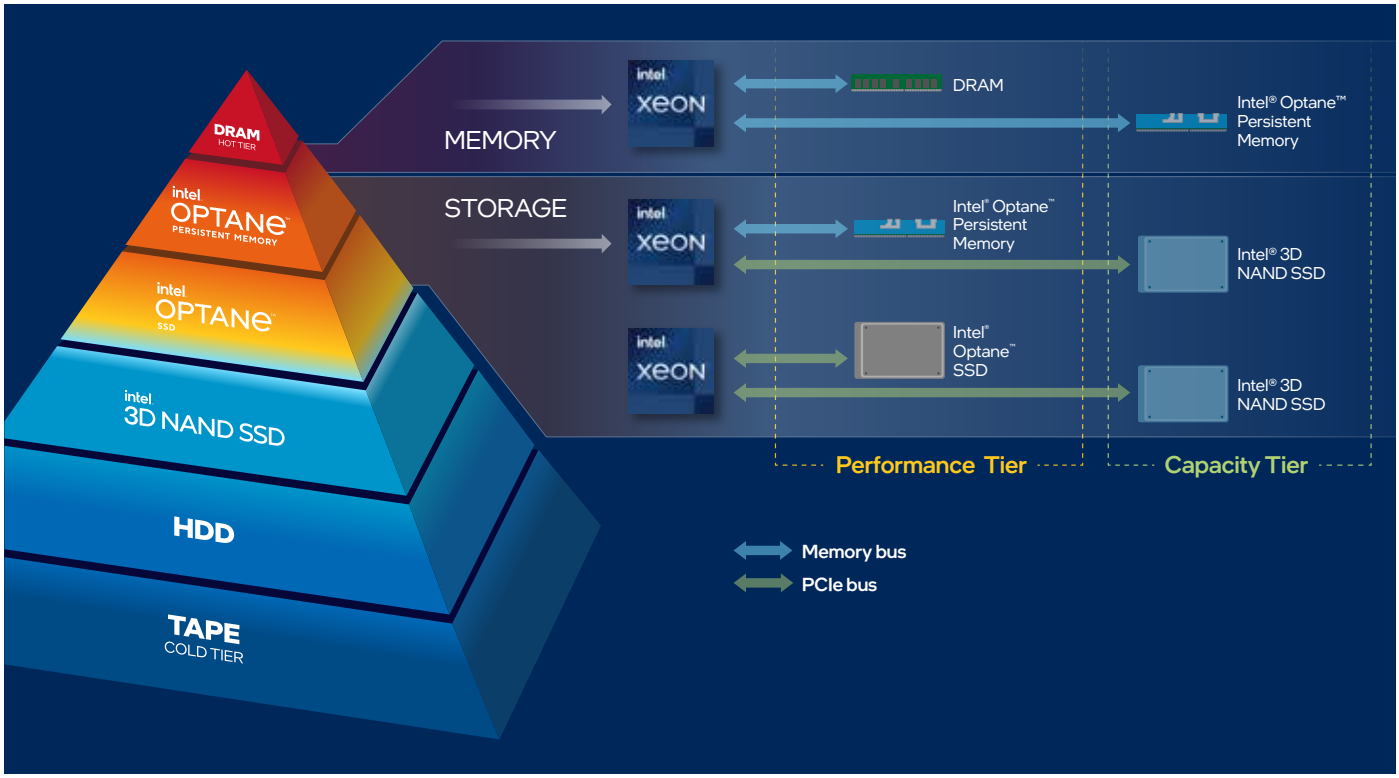


Figure 4. Intel Optane Technology creates multi-tiered memory and storage hierarchies to enable optimized workloads.

Operational Modes

Intel Optane PMem 200 series has multiple operating modes:[4]

Memory Mode delivers large memory capacity without application changes and with performance close to that of DRAM, depending on the workload. In Memory Mode, the CPU memory controller sees all of the Intel Optane PMem 200 series as volatile system memory (without persistence). The CPU uses DRAM as a fast cache to the Intel Optane PMem. In Memory Mode, data in the modules is protected with a single encryption key that is discarded upon power down, making the data inaccessible.

Memory Mode's large capacity enables more VMs and more memory per VM at a lower cost compared to DDR4 DIMMs.[5] Workloads that are I/O bound can also benefit from using Intel Optane PMem 200 series in Memory Mode, because the larger capacity supports bigger datasets at a lower cost compared to DDR4 DIMMs. With increased capacity, there is greater VM, and application density, which increases the utilization of 3rd Gen Intel Xeon Scalable processors.

App Direct Mode enables large memory capacity and data persistence for software to access DRAM and persistent memory as two separate pools of memory. In App Direct Mode, software and applications that are enabled for the industry standard persistent memory programming model have the ability to talk directly to PMem. Direct access reduces the complexity in the stack and takes full advantage of byte-addressable persistent memory with cache coherence, which extends the usage of persistent memory outside the local node and provides consistent low latency, supporting larger datasets.

App Direct Mode can also be used with standard file APIs to access the same persistent memory address space (called Storage over App Direct) without any modifications to the existing applications or the file systems that expect block storage devices. Storage over App Direct presents Intel Optane PMem as high-performance block storage, without the latency of moving data to and from the I/O bus.

In App Direct Mode, data is encrypted using a key stored on the module in a security metadata region, which can only be accessed by the Intel Optane PMem 200 series controller. The modules are locked at power loss, and a passphrase is needed to unlock and access the data. If a module is repurposed or discarded, a secure cryptographic erase and DIMM over-write operation keeps data from being accessed.

Intel Optane Persistent Memory 200 Series in the Data Center: Delivering Real Value Today[1]

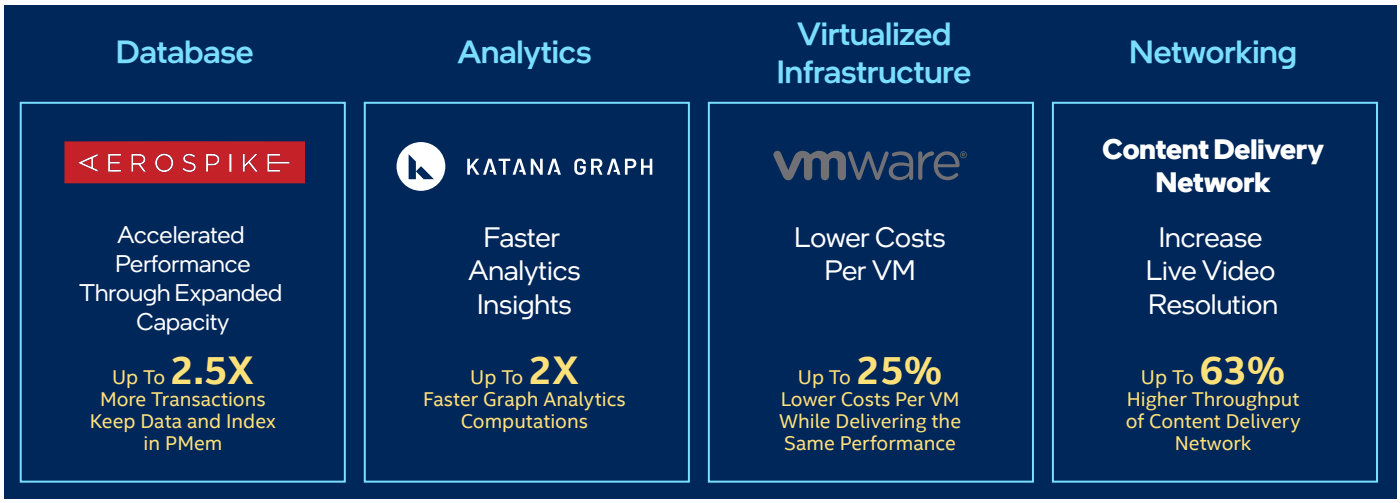


Figure 5. Intel® Optane™ persistent memory 200 series boosts performance across a wide range of enterprise applications.

Drive Application Innovation and Explore New Data-Intensive Use Cases

With Intel Optane PMem 200 series, developers have direct, byte-addressable load/store access to large memory. They can drive innovation and capabilities using the same persistent memory programming model introduced with the first generation of PMem. Rapid adoption is easy. Customers are able to take full advantage of PMem capabilities with a growing global ecosystem that includes the following:

- ISVs and OSVs
- Virtualization solution providers
- Database and enterprise solution vendors
- Data analytics vendors
- Open source solution providers
- Cloud Service Providers
- Hardware OEMs
- Standards bodies, such as the Storage Network Industry Association (SNIA), ACPI, UEFI, and DMTF

Programming Model

The software interface for using Intel Optane persistent memory was designed in collaboration with dozens of companies to create a unified programming model for the technology. The Storage Network Industry Association (SNIA) formed a technical workgroup, which has published a specification of the model. This software interface is independent of any specific persistent memory technology and can be used with Intel Optane PMem 200 series or any other persistent memory technology.

The model exposes three main capabilities:

- The management path allows system administrators to configure persistent memory products and check their health.
- The storage path supports the traditional storage APIs, where existing applications and file systems need no change. They simply see the persistent memory as very fast storage.
- The memory-mapped path exposes persistent memory through a persistent memory-aware file system. Thus, applications have direct load/store access to the persistent memory. This direct access does not use the page cache like traditional file systems. It has been named DAX by the operating system vendors.

The PMDK, available at https://pmem.io, provides libraries meant to make PMem programming easier. Software developers only pull in the features they need, keeping their programs lean and fast on PMem. These libraries are fully validated and performance-tuned by Intel. They are open source and product-neutral, working well on a variety of PMem products. The PMDK contains a collection of open source libraries, which build on the SNIA programming model. The PMDK is fully documented and includes code samples, tutorials, and blogs. Language support for the libraries exists in C and C++, with support for Java, Python, and other languages in progress.

Turn Data from a Burden to an Asset

Intel Optane PMem 200 series is the next-generation of a groundbreaking memory technology innovation. Deployed with 3rd Gen Intel Xeon Scalable processors, this technology can transform critical data workloads—from cloud to databases, in-memory analytics, virtualized infrastructure, content delivery networks, and more.

Intel Optane Persistent Memory 200 Series Data Sheet

Product Family	**Intel® Optane™ Persistent Memory 200 Series**					
Compatible Processor	**3rd Gen Intel® Xeon® Scalable processors on 2-socket and 4-socket platforms**					
Form Factor	Persistent Memory Module					
SKU*	128 GB		256 GB		512 GB	
User Capacity*	126.7 GB		253.7 GB		507.7 GB	
Platform Capacities	4S systems: 3 TB PMem + 1.5 TB DRAM per socket (4.5 TB total) per socket 2S systems: 4TB PMem + 2 TB DRAM per socket (6 TB total) per socket					
MOQ	4	50	4	50	4	50
MM#	999HGR	999HGZ	999HH0	999HH1	999HH2	999HH3
Product Code	NMB1XXD128GPSU4	NMB1XXD128GPSUF	NMB1XXD256GPSU4	NMB1XXD256GPSUF	NMB1XXD512GPSU4	NMB1XXD512GPSUF
Model String	NMA1XXD128GPS		NMA1XXD256GPS		NMA1XXD512GPS	
Technology	Intel® Optane™ Technology					
Limited Warranty	5 years					
AFR	≤ 0.44					
Endurance 100% Writes 15W 256B	292 PBW		497 PBW		410 PBW	
Endurance 67% READ; 33% Write 15W 256B	224 PBW		297 PBW		242 PBW	
Endurance 100% Write 15W 64B	73 PBW		125 PBW		103 PBW	
Endurance 67% READ; 33% Write 15W 64B	56 PBW		74 PBW		60 PBW	
Bandwidth 100% Read 15W 256B	7.45 GB/s		8.10 GB/s		7.45 GB/s	
Bandwidth 67% Read; 33% Write 15W 256B	4.25 GB/s		5.65 GB/s		4.60 GB/s	
Bandwidth 100% Write 15W 256B	2.25 GB/s		3.15 GB/s		2.60 GB/s	
Bandwidth 100% Read 15W 64B	1.86 GB/s		2.03 GB/s		1.86 GB/s	
Bandwidth 67% Read; 33% Write 15W 64B	1.06 GB/s		1.41 GB/s		1.15 GB/s	
Bandwidth 100% Write 15W 64B	0.56 GB/s		0.79 GB/s		0.65 GB/s	
DDR Frequency	Up to 2666 MT/s (4-socket systems); Up to 3,200 MT/s (2-socket systems)					
Operating Modes	App Direct only (4-socket systems); App Direct & Memory Mode (2-socket systems)					
Max TDP	15W					
Temperature (Tjmax)	≤ 83°C (85°C shutdown, 83°C default) media temperature					
Temperature (Tambient)	48°C @ 2.4m/s for 12W					
Temperature (Tambient)	43°C @ 2.7m/s for 15W					

*GB = 10^9

Bandwidths are +/- 3%

Learn more at intel.com/optanepersistentmemory

intel®

Notices & Disclaimers

[1] www.intel.com/PerformanceIndex

[2] Up to 6 TB on 2-socket systems. Up to 4.5 TB on 4-socket systems.

[3] Source: IDC 2020 MaturityScape Digital Transformation

[4] Requires 3rd Gen Intel Xeon Scalable processors to deliver Intel Optane persistent memory 200 series capabilities.

[5] Lower overall total cost of ownership (TCO).*
*Supports for more than 1.2x more VMs than the previous generation. Based on testing by Intel as of April 27, 2020 (Baseline) and March 31, 2020 (New).
Baseline: 1-node, 2x Intel® Xeon® Platinum 8270 processor on 2S platform with 768GB (DDR 24 slots / 32GB / 2666) total memory, ucode 0x4000014 running Windows Server 2019 RS5-17763, and OLTP Cloud benchmark, storage is 7 x Samsung PM963 M.2 960GB, 4x Intel SSDs S4600 (1.92TB), 1xIntel X520 SR2 (10Gb). Tested by Intel, on 31 Jan 2019.

New: 1-node, 2x Intel® Xeon® Platinum 8270 processor on 2S platform with 192, 1024GB (DDR 12 slots / 16 GB / 2666 + PMem 8 slots /128GB / 2666) total memory, ucode 0x4000014 running Windows Server 2019 RS5-17763, and OLTP Cloud benchmark, storage is 7x Samsung PM963 M.2 960GB, 4x Intel SSDs S4600 (1.92TB), 1x Intel X520 SR2 (10Gb). Tested by Intel, on 31 Jan 2019.

Intel® Optane™ persistent memory pricing & DRAM pricing as of March 2021. Pricing referenced in TCO calculations is provided for guidance and planning purposes only and does not constitute a final offer. Pricing guidance is subject to change and may revise up or down based on market dynamics. Contact your OEM/distributor for actual pricing.

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.

0321/BW/HBD/PDF ♻ Please Recycle 343429-002US