Persistent Memory Region

What is a Persistent Memory Region?

A Persistent Memory Region (PMR) is an "optional area of persistent memory that is located on the NVMe™ device, that can be read with standard PCIe® memory reads/writes." This is defined by the NVMe™ specification (see our paper on the NVMe™ Specification 2.0) which does not include use cases. Persistent memory (PMEM) could be volatile memory, like DRAM, or non-volatile media including storage class memory (SCM), or some other type of fast media, with the requirement that data can be flushed on power loss. This means that non-volatile PMEM must be backed up by power loss protection (PLP). One development in the server space is the use of persistent memory modules (PMM) such as Intel's Optane in order to reduce reliance on DRAM while improving storage performance.

Related to PMR is the controller memory buffer (CMB), as defined under the NVMe™ specification, which offers an alternative to having queues and other data for direct memory access (DMA) in host memory by acting as a PCIe® base address register (BAR). Other PCIe® devices can have BARs, for example consumer graphical processing units (GPUs) recently introduced a resizable BAR which allowed the central processing unit (CPU) to directly access the full extent of the GPU's video memory (VRAM) or frame buffer. While PMR exposes the memory (e.g. DRAM on the SSD) as general-purpose memory the CMB option presents "some or all of the DRAM as a PCIe® BAR." One advantage of this is the ability to use the memory as a peer-to-peer DMA (P2PDMA) device which can reduce CPU overhead with NVMe™ over Fabrics (NVMe-oF™).

Mode of Operation

A PMR is memory-mapped after enumeration, allocated to the host based on driver-determined capability. This allows for memory-mapped I/O (MMIO) access through the PCIe® bus. PCIe® writes are sent as posted writes, that is without acknowledgement, while reads are done end-to-end from the PMR to the host's CPU. If power loss occurs the PMR gets flushed to non-volatile media (flash) so the data can be restored to the PMR at the next power-on state. PMR can be standardized with a programming model API for memory access which includes, for example, block-based legacy support.

Advantages

As mapped, the PMR is byte-addressable but can be built on top of block-addressable storage like flash SSDs – this increases device flexibility. The ability to utilize existing SSD volatile memory (DRAM) for persistent memory means a potential reduction in total system RAM load. This is particularly useful as memory with PLP will not lose data on power loss. Many enterprise SSDs have dual ports for redundancy which also increases the reliability of a PMR. SSDs can also be designed with more DRAM to better leverage the PMR feature including sharing it with flash translation layer (FTL) duties, especially as the required metadata DRAM load for SSDs decreases with transitions to key:value (KV) and zoned namespaces (ZNS) – see our documents on these features.

One of the primary advantages is that PMR access is handled over PCIe®, which is a mature and reliable interface. The ability to use volatile memory means that the PMR has very low latency relative to traditional storage which makes it a superior choice for many workloads (see below). Using a PMR is, in the least, simpler and faster than developing a NVMe™ I/O command queue and waiting for acknowledgement of completion. It's also possible to aggregate PMRs over a number of devices which includes all of the options found with NVMe™, for example with NVMe-oF™. This makes it easier to integrate and use within an existing framework.

Applications

PMRs have a ton of potential uses in enterprise, oriented primarily on workloads that benefit the most from persistent, low latency storage. This includes logging for software RAID, NOSQL, NOSQL, and MySQL databases. It's also useful with journaling – that is, tracking file system changes not yet committed. It can be used as a buffer for caching systems, for example being utilized when combining/coalescing writes. It's also useful space for metadata storage which tends to be latency-sensitive data, for example with larger or slower storage systems. PMRs are also useful as a staging space for de-duplication – that is, the elimination of duplicate data – as well as for compression/decompression. A PMR can additionally be utilized for remote direct memory access (RDMA) transactions, or the accessing of data between different hosts, as can be combined with P2PDMA.

Summary

PMRs allow for more efficient use of hardware resources through existing, reliable standards. The potential reduction in CPU and system memory load is particularly useful, especially as the use of low latency memory for certain workloads increases storage flexibility. The enterprise market already utilizes memory hierarchies including new forms of persistent memory, like phase change (PCM) and memristor technology, but PMRs can also use volatile DRAM typically found on flash-based SSDs if they have PLP – which they often do.

Capacious storage tends to be block-based but PMRs are byte-addressable directly to host CPUs. This usage is expanded further with in-memory applications and acceleration – that is, compute directly in memory – which is increasingly used in the Internet of Things (IoT) and Edge computing (see our documents).

PMRs reduce the amount of system memory needed while also increasing the reliability of especially volatile memory. This is done over the mature PCIe® bus – including as a BAR with CMB – and also falls under the NVMe™ specification. The combination of speed and redundancy makes PMRs particularly good for specific workloads: logging, journaling, buffering/caching, metadata, staging, and RDMA. P2PDMA is already be using in other areas, like gaming – with GPUs and soon, storage, for example with DirectStorage – and in enterprise, being able to leverage existing standards to expand resources over a larger network. In the long term this means increased flexibility and efficiency with an overall reduction in overhead.

Our SSD Solutions

PCIe™ - Our ED1 Series is a powerful, high performance SSD made for edge storage applications. It comes in M.2 and U.2 form factors.

SATA - Our ER2 SSD Series delivers affordability and performance with superior random read/write speeds of up to 90,000/45,000 IOPS. It comes in M.2 and 2.5" form factors.

Please contact our Solid State Storage Technology Corp. expert for more information.

*Specifications and features are subject to change without prior notice. Images are samples only, not actual products.

Request Full Specs Sheets

ABOUT US

A subsidiary of KIOXIA Corporation, **Solid State Storage Technology Corporation** is a global leader in the design, development, and manufacturing of digital storage solutions. We offer a comprehensive lineup of high-performance customizable SSDs for the Enterprise, Industrial, and Business Client markets. With various form factors and interfaces, our SSD solutions help businesses simplify their storage infrastructures accelerating variable workloads, improving efficiency, and reducing total cost of ownership.

Learn more at www.ssstc.com