Radian®
MEMORY SYSTEMS

Zoned Namespaces (ZNS)

Cooperative Flash Management

Targeting compliance with the forthcoming Zoned Namespaces (ZNS) specification from the NVM Express™ industry standards organization, Radian's Zoned Namespaces are idealized, configurable Flash SSD zones that can be associated with namespaces. Based upon the award winning Symphonic™ Cooperative Flash Management technology, Radian's Zoned Namespaces abstract vendor specific NAND attributes but present host systems with sequential write zones of 'Idealized Flash'. Routine Flash management processes are simplified, and cooperative between the device and the host to provide superior determinism, parallelism, write amplification and tail latencies that cannot be achieved with conventional Flash Translation Layers (FTL).

In addition to zone sizes being user configurable, these zones can be factory configured to support different types of memory, ranging from NV-RAM to SLC and TLC today, and for SCM and QLC in the future.

Cooperative & Configurable Zones of *'Idealized Flash'*

***'All Firmware'* SSD implementation is**
less OS dependent and optimal for SPDK

Optional factory configurations for mixing zones of different memory types (NV-RAM, SCM, SLC, TLC, QLC, etc.) on the same SSD

Comprised of NAND Erase Units (blocks), these zones are subsets of physically, performance-isolated regions of memory, where the regions are configurable in size and can be associated with namespaces. Zones appear as a range of contiguous LBAs accessible via conventional addressing through the NVMe command set. Certain SMR zone commands, such as 'Zone Report' and 'Zone Reset', are supported as part of extensions to the conventional NVMe command set.

Iso-boxes & Namespaces

Iso-boxes are user configurable, discrete performance-isolated regions based upon NAND dies and channels that can be associated with namespaces.

'Idealized Flash'

Radian's Zones are comprised of NAND Erase Units (blocks) that come from the same iso-boxes. Geometry emulation abstracts NAND geometry and vendor-specific attributes, presenting the host with zones of *'Idealized Flash'*.

Host Data Placement

Sequential write zones are presented to the host as a contiguous range of LBAs and host data placement is preserved through to the media in that zone. An optional Delegated Move operation enables the host to command the device to transfer data between zones.

Configurability

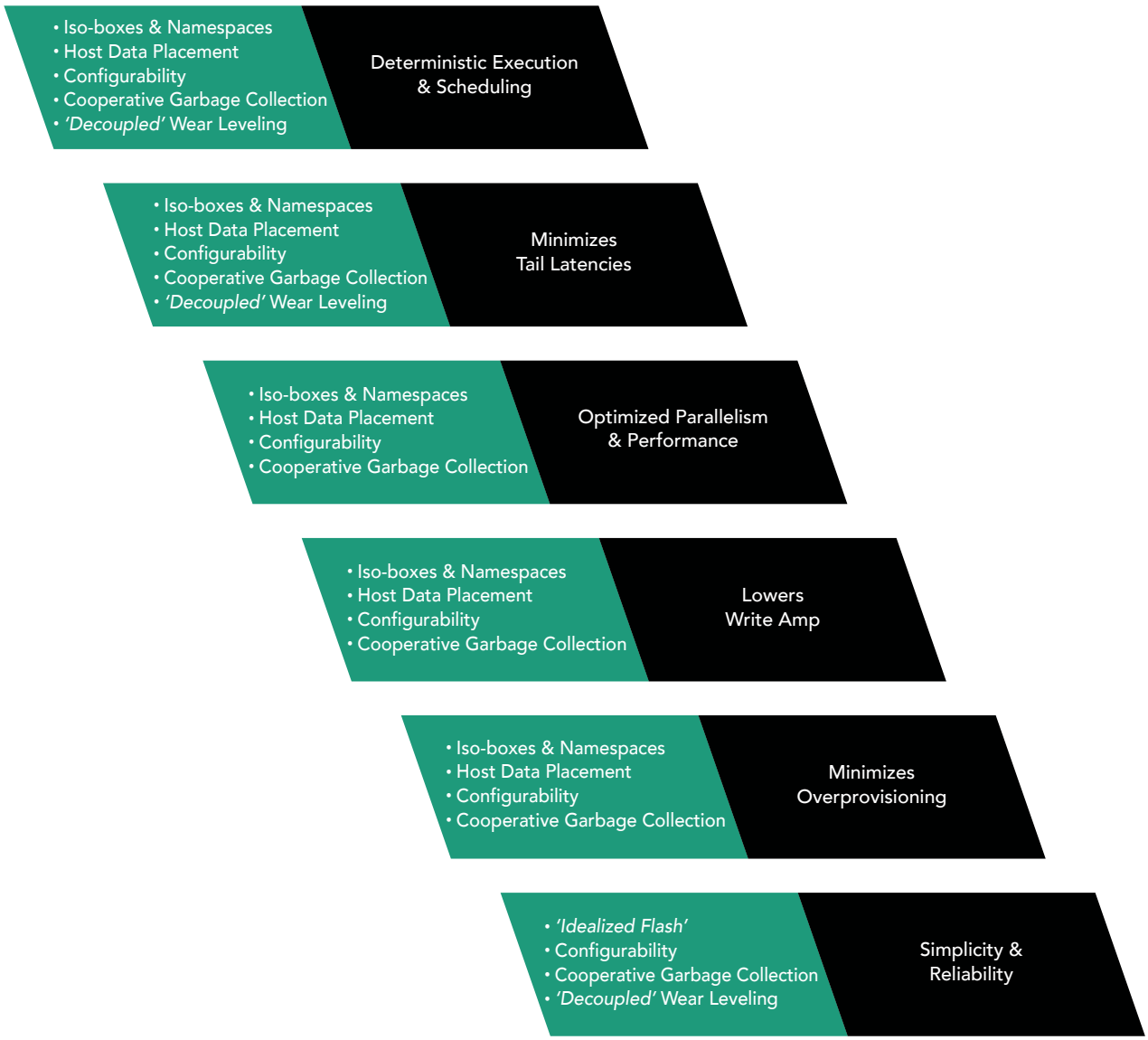
Radian's Zones, iso-boxes, and write stripes are all user configurable.

Cooperative Garbage Collection

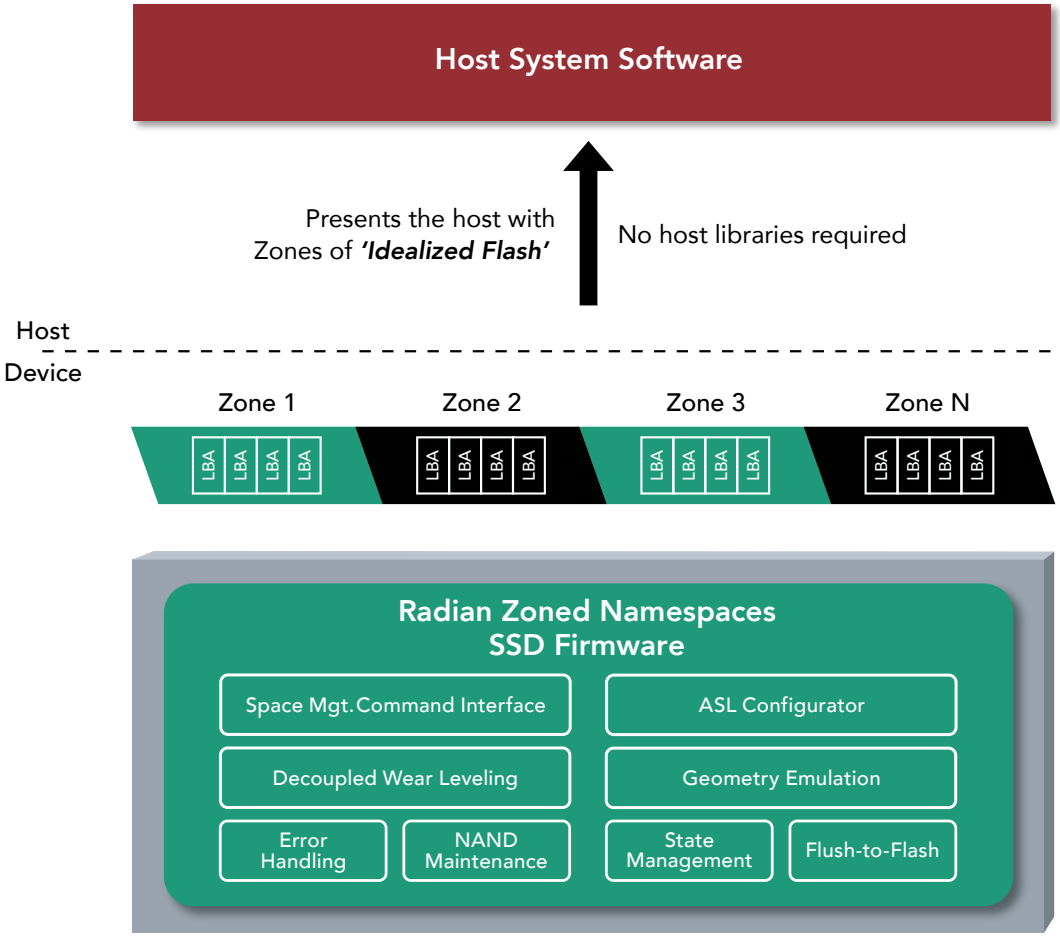
After relocating valid data, zones are erased by the host issuing a 'Zone Reset' command to the SSD for the target zone, whereby the device immediately erases that zone.

'Decoupled' Wear Leveling

Routine wear leveling and NAND maintenance are performed internally by the SSD in a coherently aligned manner, where the device only initiates cooperative requests to the host if required data movement could conflict with other I/O access.



Accessing a Radian Zoned Namespaces (ZNS) SSD follows the same host/device model as previous Radian Symphonic SSDs. This enables the host to control data placement while the SSD abstracts lower level media management, including geometry and vendor-specific NAND attributes (*'Idealized Flash'*). Flash management processes such as garbage collection, wear leveling, and NAND maintenance are executed by the device, under cooperative host control, and hence performed deterministically.



- Configurable, cooperative zones of *'Idealized Flash'* are presented to the host as contiguous LBAs and discovered through the 'Zone Report' command (NVMe vendor extension).
- As part of garbage collection, hosts are responsible for selecting valid data and a relocation destination on a different zone, either performing a copy/write operation directly or using Radian's optional Delegated Move command (NVMe vendor extension) that delegates the data transfer to the device.
- Zones are cleaned (erased) through the use of a 'Zone Reset' command (NVMe vendor extension) that is issued by the host to the device, or via a zone aligned NVMe deallocate command.
- By default, routine wear leveling and NAND maintenance (data retention, scrubbing, error handling) are performed internally by the device without requiring host involvement or interfering with host latencies.
- The device initiates a cooperative request to the host if additional wear leveling or other NAND maintenance is required that could conflict with host I/O access latencies.

Configurable Zones = Minimum Write Amp

Minimizing write amplification is becoming more critical with each new generation of NAND Flash. As NAND die capacities continue to grow larger, unnecessary traffic from write amp has a greater impact on latencies. And as process nodes shrink and bits per cell increase, such as with QLC, endurance is further diminished, potentially to a point where excess write amp prevents use of the technology for many applications.

Most modern host systems are non-overwriting, and many are based on log structured architectures where a contiguous address range, known as a segment, is cleaned as part of the space reclamation process. Typical SSDs have FTLs, which are also log structured and have the equivalent of a segment that is cleaned as part of garbage collection. At the system level, this creates a challenge known as 'log on log', where each layer is independently, redundantly cleaning and likely on different segment lengths and distributions. The result is additional write amp and unpredictable latency spikes.

With Radian's Zoned Namespaces, the host continues to be responsible for cleaning but the SSD does not have a redundant cleaning log, eliminating the 'log on log' write amplification. Of equal importance, Radian's Zones are user configurable. The ***Address Space Layout (ASL) configurator*** enables users to configure the SSD zones to match the host file system's segment size. This configurability minimizes modifications to host system software and associated integration efforts, while also eliminating the write amplification that would otherwise occur if the host segments and SSD zones were not aligned.

Deterministic Performance

Cooperative Garbage Collection

Unlike conventional FTL SSDs that clean (garbage collect) random address ranges without warning, Radian's Zones are cleaned deterministically.

As part of its normal space reclamation process, the host selects which zones (segments) to clean. Relocation of valid data is either performed directly by the host with a copy/write operation, or using Radian's optional Delegated Move operation where the host specifies the destination address and commands the device to perform the data transfer. Then the host simply issues a 'Zone Reset' command to the device, and the device immediately erases the associated zone without introducing any new or complex scheduling requirements. This enables SSD garbage collection to be deterministic and prevents unforeseen latency spikes.

'Decoupled' Wear Leveling NAND Maintenance

In conventional FTLs, the wear leveling process is often integrated with garbage collection processes and algorithms. The Radian Zoned Namespaces SSD performs *'Decoupled'* Wear Leveling and *'Decoupled'* NAND maintenance. These are cooperative, memory controller-owned processes performed by the device and effectively decoupled from the host's garbage collection algorithm.

When the host issues a 'Zone Reset' as part of its aforementioned space reclamation and cleaning process, the Radian Zoned Namespaces SSD will internally, concurrently perform wear leveling and data scrubbing on that same zone in a coherently aligned manner that does not interfere with other host-directed I/O accesses.

Because the host, or a host FTL, is in control of garbage collection and likely to be log structured, writes will inherently tend to level wear. When this is not adequate and the device determines that additional wear leveling or NAND maintenance data scrubbing is required, the Radian Zoned Namespaces SSD will use its Back Channel, an out of band communication path, to initiate a request to the host that certain ranges of data be moved or refreshed.

The device continues to escalate these requests until the host responds. If the host does not respond, the device will eventually proceed with the necessary data movement which may briefly interfere with other I/O, but the latency conflict is one that has been forecasted to the host and is hence deterministic. The host can also poll the device to request this information in advance, taking it into account as part of its routine cleaning and data relocation processes.

Similarly, the Radian Zoned SSD performs bad block management by transparently remapping erase units (bad blocks) from erase units held in reserve. This swapping of erase units is again handled deterministically and typically without impacting host latencies.

ASL Configurator = Optimized Scheduling

ASL Configurator

Radian's Address Space Layout (ASL) configurator enables user configuration of different size zones, write stripes, iso-regions, and iso-boxes. An iso-region is a physically discrete, performance-isolated region based upon NAND dies and channels. An iso-box consists of one or more iso-regions and can be associated with a namespace. In addition to being based upon capacity, these isolated regions can be configured based upon characteristics in terms of endurance, I/O bandwidth, predictable I/O latency, cleaning policies, deterministic scheduling or other combinations of desired metrics via parameterized descriptions. Radian's Zones are comprised of NAND Erase Units (blocks) that come from the same iso-regions. And write stripes are formed from a number of NAND pages from within that zone.

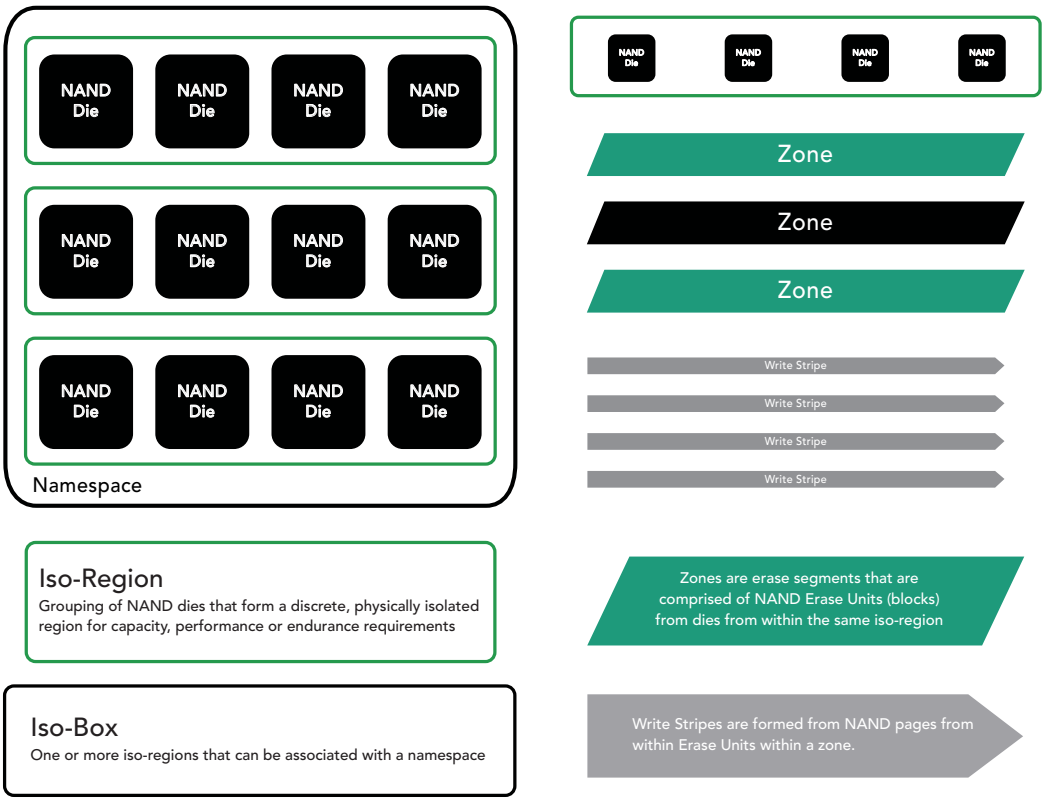
Radian's ***Geometry Emulation*** virtualizes the topology of the NAND, including geometry and vendor-specific attributes, to present the host with zones of *'Idealized Flash'* while maintaining symmetric alignment through to the physical memory. The ASL configurator utilizes hierarchical address virtualization to enable users to configure those zones to best match their system requirements.

Zones, Write Stripes, & Scheduling

NAND Flash arrays have an inherent tradeoff between performance and efficiency constraints. Configuring zones and write stripes to be wide and shallow will increase bandwidth, but will also increase write amplification and collisions that induce latency spikes. Alternatively, configuring zones and write stripes to be narrow and deep will reduce write amplification and latency spikes, but will also reduce frontier bandwidth.

Beyond performance and efficiency constraints, optimized system scheduling should be taken into consideration. While presenting 4TB, 8TB, or more as a single SSD creates challenges around deterministic performance, configuring a single SSD as hundreds of individual block devices or namespaces can create other challenges, including significant complexities in terms of optimizing host scheduling.

Radian's ASL Configurator provides different parameterized profiles that optimize Address Space Layout to help address the challenges of complex scheduling, along with the sizing tradeoffs associated with zones and write stripes. This includes the ability to configure isolated regions of variable sizes, with different ASL profiles, within a Radian SSD to obtain the configuration that best matches the application requirements.



Strict Write Pointers

NAND Flash memory requires programming the media sequentially. Hosts that access the media directly on a SSD are therefore required to write to the relevant addresses (e.g., within a zone) sequentially. For this reason, zoned drives employ what is known as a Strict Write Pointer, requiring that the write pointer always points to the next available address on the media, and that the LBA of arriving writes must match the current write pointer of the associated zone.

Tangled Ordering

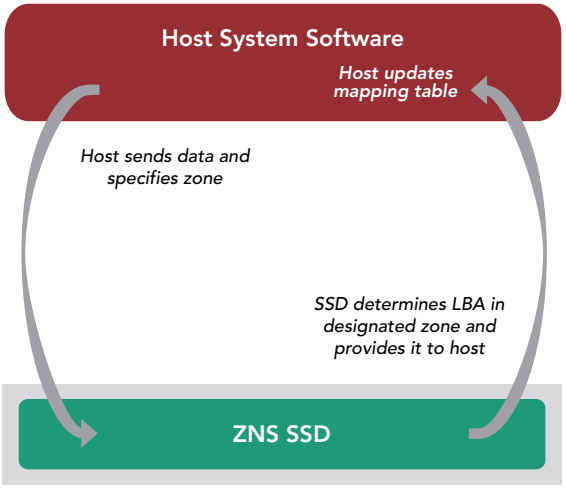
Various transport protocols support packet reordering to improve flow control and other link transmission requirements. Tangled Ordering occurs when host system software issues sequential write operations for LBAs to a storage device, but as the LBAs travel through the system, over different chip sets and link protocols, they do not arrive at the storage device's receive buffer in exactly the same sequential order that they were issued.

On devices constrained with strict write pointers, this can result in a write error, requiring out-of-order writes be reissued, thereby negatively impacting performance. Sequential ordering can be guaranteed through the use of single I/O depths or issuing all writes synchronously, but this approach will again negatively impact performance.

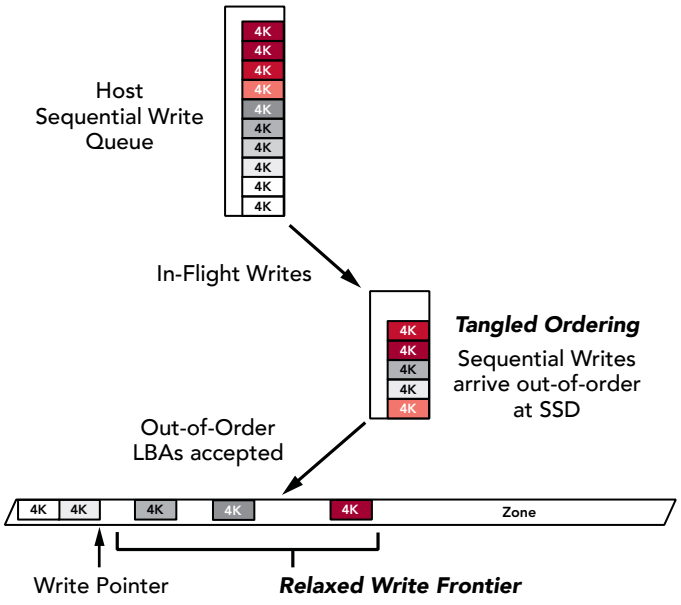
Zoned Append

Zone Append overcomes the challenges with Strict Write Pointers and Tangled Ordering by enabling the device to select the LBA offset into a zone. When issuing write requests, hosts only need to specify the target zone. Upon receiving a write request to a zone, the device appends the data to the current write pointer within the zone, determines the associated LBA, and sends the specific address to the host with a completion status. The host then updates the applicable mapping table. Radian's Zone Append feature provides the ability to run multiple append requests and completions concurrently.

However, Zone Append does introduce a requirement for modifications to host software which vary in complexity depending upon the target architecture. The impact of additional latency and the complexity of maintaining consistency through system failures will also vary depending upon the target software architecture.



Relaxed Write Pointer



Based upon hierarchal address virtualization, Radian's *'Idealized Flash'* supports Relaxed Write Pointers. In addition to overcoming NAND vendor specific attributes, geometry and addressing anomalies, *'Idealized Flash'* relaxes the frontier for the write pointer. This capability enables accepting write operations that are written sequentially by the host, but arrive at the SSD out-of-order (Tangled Writes).

Unlike Zone Append, the Relaxed Write Pointer capability avoids having to introduce modifications to host software, accounting for new system consistency models, or additional latency that could be caused by Zone Append.

Tiered Zones

Radian's Zoned offer a simple, logical model to access different memory types as different zones on the same SSD, from NV-RAM to Storage Class Memory to NAND Flash, and different classes of NAND Flash. Variations of SLC NAND can provide very low latency and high endurance, while TLC NAND provides better capacity and cost efficiencies. The architecture can support factory configured zones designated as SLC, or TLC, or NV-RAM zones today, and other memory types such as QLC, SCM, and specialized ultra low latency SLC variations in the future. Combined with Radian's Delegated Move technology, this enables hosts to readily apply tiering between zones based upon different memory technologies to optimize efficiency trade-offs or cost, performance, capacity and endurance.

'All Firmware' SDF SSD & SPDK

Until now, Software-Defined Flash (SDF) SSDs have either not offered *"Idealized Flash"* or *"Configurable Addressing"*, or, in the case of Radian's SDF SSDs, required proprietary host resident libraries to provide this functionality. Radian's Zoned SSD is the first SDF SSD to provide this functionality completely in device firmware, without requiring any host libraries. This obviates the issues of having to integrate proprietary vendor libraries into host system software and minimizes OS compatibility requirements. The *'All Firmware'* implementation is especially advantageous in SPDK environments which do not require NVMe device drivers or use of the kernel block layer, as existing targets can access the Radian Zoned SSD directly without transitioning through intermediary libraries.

ZBD to NVMe Bridge

Radian offers an optional host library to customers utilizing a zone block device (ZBD) interface. Providing a protocol translation from the zone block device interface to NVMe, this bridge enables system software to access the Radian Zoned SSD as a NVMe block device using a subset of the SMR zone block device commands.

Radian MEMORY SYSTEMS

Radian Memory Systems, Inc.
Tel 818 222 4080 Fax 818 222 4081
sales@radianmemory.com www.radianmemory.com