Fast Integration and Furious Performance with Zoned Flash drives

Robert Lercari
V.P. Engineering
Radian Memory Systems, Inc.
Legal

This presentation and/or accompanying oral statements by Radian Memory Systems, Inc. ("RMS") representatives collectively, the “Presentation”, is intended to provide information concerning Flash memory technology. While RMS strives to provide information that is accurate and up-to-date, this Presentation may nonetheless contain inaccuracies or omissions. As a consequence, RMS does not in any way guarantee the accuracy or completeness of the information provided in this Presentation.

This Presentation may include forward-looking statements, including, but not limited to, statements about any matter that is not a historical fact; statements regarding RMS’ intentions, beliefs or current expectations concerning, among other things, market prospects, technological developments, growth, strategies, and the industry in which RMS operates; and statements regarding products or features that are still in development. By their nature, forward-looking statements involve risks and uncertainties, because they relate to events and depend on circumstances that may or may not occur in the future. RMS cautions you that forward looking statements are not guarantees of future performance and that the actual developments of RMS, the market, or industry in which RMS operates may differ materially from those made or suggested by the forward-looking statements in this Presentation.

In addition, even if such forward-looking statements are shown to be accurate, those developments may not be indicative of developments in future periods.
Agenda

• Cooperative Flash Management (CFM) and Zoned Flash
• Zone Configuration and Address Space Layout (ASL)
• Solutions for Strict Write Pointers
• ‘Apples to Apples’ Performance Benchmark: Zoned Flash vs. FTL SSD
Cooperative Flash Management (CFM)

Redistribution of Flash Management between Host/Device

- Data Placement
- Leverages host segment cleaning for Garbage Collection
- Scheduling

Host System Software

- Wear Leveling
- NAND Maintenance
- Maintains device state
- Idealized Flash
- Configurable Addressing
- Offload process execution

Standard NVMe API plus vendor specific extensions
Cooperative Flash Management (CFM)

2014
Symphonic™ v1
Demo’d CFM to leading Flash Fabs

2015
Symphonic™ v2
Won FMS

2017
SMR Drive
Zoned Flash

2018
Implementation for OC2

2018
Symphonic™ v3
‘All Firmware’ implementation

2019
Zoned Flash

Zoned Flash
Zoned Flash

- Idealized Flash
- ASL Configurator
- Decoupled Wear Leveling and NAND Maintenance
- Back Channel*
- Delegated Move offload*
- Zone Append*
- Relaxed Write Pointer

*Optional feature
Configurability

Address Space Layout (ASL)

- Performance
- Endurance
- Capacity

Namespace

Iso-Region
Dies form discrete, physically isolated regions

Iso-Box
One or more iso-regions that can be associated with a namespace
• Zones and Application Segments: Write Amp

• Write Stripes: Bandwidth/Latency

NAND Erase Units (blocks) from dies from within the same iso-region

NAND pages from within Erase Units (blocks) within a zone.

• Zone Report command
Strict Write Pointer

- NAND requires sequential programming
- Tangled Ordering
- Performance Impact
Zone Append

**Pros**
- No Strict Write Pointer requirement
- Overcomes NAND addressing anomalies, geometry or vendor specific attributes
- No FTL L2P storage requirements
- .1%, 1GB mapping space for 1TB capacity

**Cons**
- Modifications to host system software
- New consistency models
- Potential latency impact

Radian’s Zone Append can support multiple, concurrent append request/completions

Host System Software

1. **Host sends data and specifies zone**
2. **SSD determines LBA in designated zone and provides it to host**
3. **Host updates mapping table**
Relaxed Write Pointer

- Overcomes Tangled Ordering if host attempts to write sequentially
- No modifications to host software, no new consistency models or additional latency
- Minimal SSD memory (not 0.1% like L2P tables)
Testing Zoned SSDs

Zoned Flash U.2 NVMe SSD

- 2TB -16TB TLC Flash
- Two different NAND vendors/fabs
- User NV-RAM
- Single or Dual Port

How to test...

- Garbage Collection
- Endurance and Data Retention (e.g., JESD219 workloads)
- HA/Fault Tolerance (e.g., dual port, shorn writes)
- Performance and Comparative Performance Testing

Radian Block Translation Layer

- Provides support for in-place overwriting of zones (Conventional Zones)
- Log Structured design serializes random overwriting workload
- Performs segment cleaning (garbage collection) with Zone Reset
Apples to Apples Comparison

Identical Silicon
- Same SSD Processor
- Same Flash Array
  3D TLC NAND
  Dies/Package
  # of Channels
  # of Packages/Channel
  4.6TB Raw capacity

- Same DDR4 array
  DDR4
  # of Devices

Zoned Flash
U.2 NVMe SSD
RMS-350

FTL
U.2 NVMe SSD
System Test Configuration

Application Workload

- 70/30 Mix
- 4K Random Read
- 4K Random Write
- SSD Queue Depth = 32
  - 4 worker threads
  - IOD = 8/thread

- Performs segment cleaning (garbage collection)
- Creates Log-on-Log

Measuring at the system level

- Emulates typical SDS and All-Flash Array storage stacks
- Serializes random, overwriting workload
# Overprovisioning (OP)

<table>
<thead>
<tr>
<th></th>
<th>FTL SSD</th>
<th>Zoned SSD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advertised User Capacity</strong></td>
<td>3.23TB</td>
<td>3.23TB</td>
</tr>
<tr>
<td><strong>Total OP</strong></td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td><strong>LS Host Free Space</strong></td>
<td>13%</td>
<td>27%</td>
</tr>
<tr>
<td><strong>Advertised Device Capacity</strong></td>
<td>3.84TB</td>
<td>4.49TB</td>
</tr>
<tr>
<td><strong>SSD Internal OP</strong></td>
<td>17%</td>
<td>3%</td>
</tr>
<tr>
<td><strong>Raw Capacity</strong></td>
<td>4.62TB</td>
<td>4.62TB</td>
</tr>
</tbody>
</table>

- **Advertised User Capacity**
- **Total Overprovisioning**
- **LS Host Free Space**
- **Advertised Device Capacity**
- **SSD Internal Overprovisioning**
- **Raw Device Capacity**
• 70/30 Mix
• 4K Random Read
• 4K Random Write
• SSD Queue Depth = 32
  4 worker threads
  IOD = 8/thread
• Total Overprovisioning = 30%
• Single Namespace

99.99% Latency Over Time

99.99% Latency @ IOPS

FTL SSD

Radian Zoned SSD

>75ms Delta @ 250K IOPS
- 70/30 Mix
- 4K Random Read
- 4K Random Write
- SSD Queue Depth = 32

99.99% Latency @ IOPS
Single Namespace, 25% OP and 30% OP

99.99% Latency @ IOPS
Sixteen Namespaces, 30% OP

Log-on-Log
+ Noisy Neighbors

---

FMS 2019
Santa Clara, CA