

September 23-26, 2019 Santa Clara, CA

Zoned Flash SSDs in Advanced Storage Systems

Robert Lercari Radian Memory Systems, Inc.



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
- ASL Configurator
- Idealized Flash
- Delegated Copy-Move
- Back Channel
- Performance Benchmarks

Cooperative Flash Management (CFM)

Redistribution of Flash Management between Host/Device



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

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

Standard NVMe API plus vendor specific extensions

٠

٠

4

SD®





Zoned Flash



Zoned Flash ver. 3.0; RMS-350

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

*Optional feature



6

SD©

Smerf Bridge (optional)

- Zone Block Device (ZBD) to NVMe protocol translation bridge
- Provides support for Zone Report and Zone Reset commands between:
 - Hosts using the ZBD interface
 - Zoned devices using NVMe vendor extensions for Zone Report and Zone Reset
 - Support for multi-drive volumes



Address Space Layout (ASL) Configurator



Iso-Region Dies form discrete, physically isolated regions

Iso-Box One or more iso-regions that can be associated with a namespace

- Performance
- Endurance
- Capacity

8

SD[®]

Address Space Layout (ASL) Configurator

- 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





9

SD©



2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

ር____

MEMORY SYSTEMS

Idealized Flash

- Geometry Emulation
 - Presents idealized NAND
 - Hierarchal Address Virtualization
 - Abstracts vendor-specific attributes
 - Maps Bad Blocks
 - Not a FTL
 - No 0.1% mapping storage requirements
 - Deterministic

Contiguous Addressing





11

SD (®

Strict Write Pointers



Tangled Ordering Write Errors



- Without Idealized Flash (Geometry Emulation turned off)
- · Closed system



2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

SD®

Zone Append

SD©

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







Relaxed Write Pointer

- Idealized Flash enables 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)



SD (9

Variable Capacity & Zone Excursions



- Addressing gaps
- Zones can change capacity due to bad blocks
- · Host software requires modifications

Idealized Flash

	Zoned Na SSD Fir	imespaces rmware	
Space Mgt. Command Interface Decoupled Wear Leveling		ASL Configurator Geometry Emulation	

- No addressing gaps
- Automatic bad block
 replacement
- Zones do not change size/capacity
- · No modifications to host software



Space legt: Command Interface ASL Configurator Decoupled Wear Leveling Geometry Emulation

SD®

	Zoned Na SSD Fir	imespaces rmware	
Space Mgt. Cor	nmand Interface	ASL Configurator Geometry Emulation	
Decoupled 1	Vear Leveling		
Error Handling	NAND Maintenance	State Management	Flush-to-Flash

Idealized Flash

- Idealized Flash overcomes the limitations or need for...
 - Strict Write Pointers
 - Zone Append
 - Variable Zone Capacities and Excursions
- Simplifies integration into existing storage systems
- Minimizes host software modifications when transitioning to different NAND



	Zoned Na SSD Fi	mespaces mware	
Space Mgt. Command Interface Decoupled Wear Leveling		ASL Configurator Geometry Emulation	

Idealized Flash

Idealized Flash overcomes the limitations or need for...



- Simplifies integration into existing storage systems
- Minimizes host software modifications when transitioning to different NAND





Delegated Copy-Move



Delegated Copy-Move Commands

- NVMe vendor specific extension
 - NV-RAM \rightarrow Flash
 - Flash \rightarrow NV-RAM
 - Flash \rightarrow Flash

2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

User NV-RAM

- Multiple use cases and advantages
- Can appear as distinct...
 - Zone(s)
 - Namespaces and/or unique block device
- User Controlled
- Deterministic
- Efficient



19

SD (9



2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

MEMORY SYSTEM



Delegated Copy-Move



Assumes 1.5 Host WA factor Assumes 2.0 FTL SSD internal WA factor = Copied Data

Radian[®]

2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

SD[©]

Disaggregated Storage

DCM avoids host/device copying over the wire



'Back Channel' (*optional*) Wear Leveling & NAND Maint.

- 'Back Channel' optional out-of-band communication path
- Supports various cooperative host/device management activities
- Decoupled Wear leveling and NAND Maintenance managed transparently by device
- Routine management handled transparently in coherently aligned manner
- Additional wear and maintenance signaled by device through Back Channel
- Scheduled by host to avoid unpredictable spikes
- Ultimately enforced by device to maintain warranty

Host Storage Stack





SD©

'Back Channel' (optional) Wear Leveling & NAND Maint.





SD[®]

Block Translation Layer

Enables testing random, overwriting workload



- Log Structured design serializes
 random writes
- Performs segment cleaning (garbage collection) with Zone Reset
- Can enable Conventional Zones



SDC '19 Demo



SD (9



Radian Zoned SSD



Zoned Flash U.2 NVMe SSD RMS-350

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





FTL SSD

FTL U.2 NVMe SSD



2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

SD®

Block Translation Layer



2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

SD @

28

Overprovisioning (OP)

30% Example

	FTL SSD	Zoned SSD				
Advertised User Capacity	3.23TB	3.23TB				
Total OP	30%	30%				
LS Host Free Space	13%	27%				
Advertised Device Capacity	3.84TB	4.49TB				
SSD Internal OP	17%	3%				
Raw Capacity	4.62TB	4.62TB				





• SSD Queue Depth = 32

4 worker threads

• 70/30 Mix

٠

٠

4K Random Read

- Total Overprovisioning = 30%
- Single Namespace

99.99% Latency @ IOPS



2019 Storage Developer Conference. © Radian Memory Systems, Inc. All Rights Reserved.

SD[®]

- 70/30 Mix
- 4K Random Read
- 4K Random Write
- SSD Queue Depth = 32

99.99% Latency @ IOPS

Sixteen Namespaces, 30% OP

Santa Olara.

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



