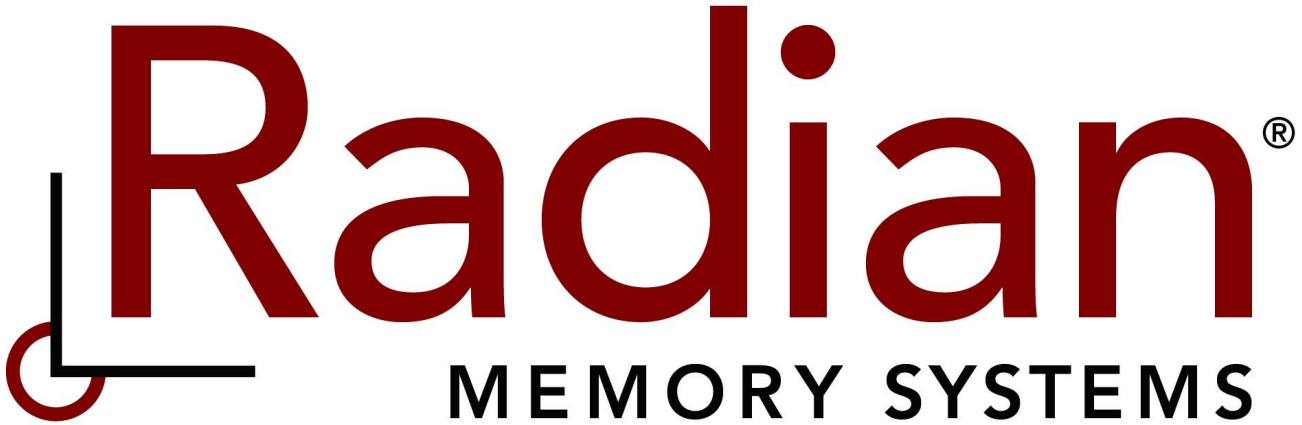


ZNS SSD Demo
RMS-350 with RocksDB / ZenFS and gzbd-viewer

Bob Varney
Radian Memory Systems, Inc.
August 2020

Legal

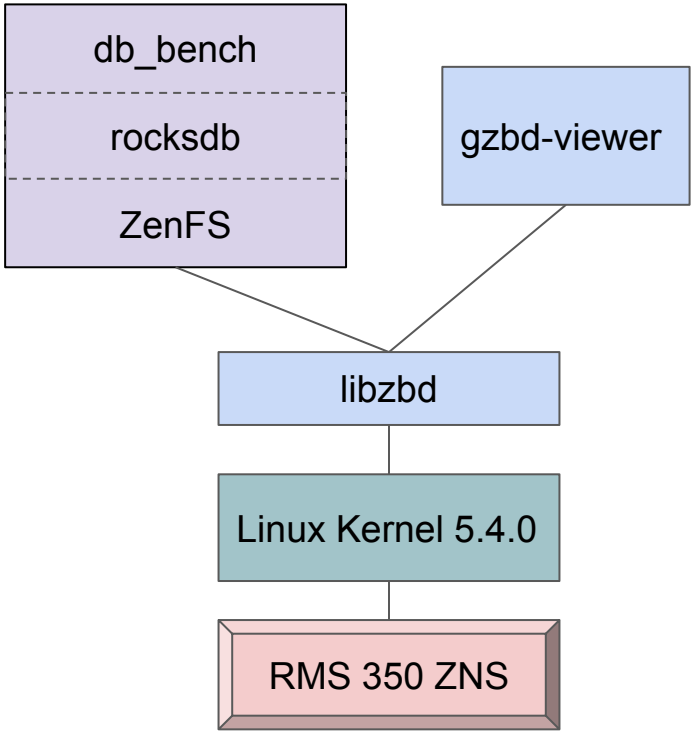
This document and the information contained herein are the property of Radian Memory Systems, Inc. Statements may be subjective in nature and the company makes no guarantees regarding such statements. All marks are the property of their respective owners.

SSD

RMS-350 ZNS SSD

- Compatible with NVM Express™
  Zoned Namespaces Command Set 1.0, TP-4053*
- Configurable Zone Sizes
  - “Narrower” zones = 144MB each
  - “Wider” zones = 9GB each
- Raw Device Level Performance
  - Sequential Read Throughput = 3,424 MB/s
    (100% 128K reads QD=32 under fio)
  - Sequential Write Throughput = 2,636 MB/s
    (100% 128K writes QD=32 under fio)

System



- Rocksdb [1] with modifications to ZenFS [2] (on-the-fly filesystem creation)
- Modifications to libzbd [3]
  - NVMe ZNS commands operable in Kernel 5.4
  - Support contiguous addressing
- No modifications to gzbd-viewer (other than to link modified libzbd)
- Demo only uses 32-64 zones of 4TB RMS-350
- Run on Fedora, on Intel i7 3.6GHz processor with 8 cores and 16G memory

References:

1. Rocksdb, https://rocksdb.org
2. ZenFS, https://github.com/facebook/rocksdb/pull/6961
3. libzbd, https://github.com/westerndigitalcorporation/libzbd

Results Reported By **db_bench** (MB/s)

Benchmark	**1 x db_bench** **Standard Rocksdb**	**1 x db_bench** **Standard Rocksdb** **(No WAL)**	**1 x db_bench** **Rocksdb + ZenFS** **48 Zones (144M each)**	**1 x db_bench** **Rocksdb + ZenFS** **48 Zones (9G each)**	**1 x db_bench** **Rocksdb + ZenFS** **48 Zones (9G each)** **(No WAL)**
fillseq	254.7	483.8	32.6	229.1	800.8
fillrandom	243.4	302.1	-	183.2	678.7
overwrite	228.7	328.1	-	200.1	624.0

More Results Reported By **db_bench** (MB/s)

Benchmark	**“Whole”** **1 x db_bench** **Rocksdb + ZenFS** **1 x 64 Zones 9G Each** **(No WAL)**	**“Single Half”** **1 x db_bench** **Rocksdb + ZenFS** **1 x 32 Zones 9G Each** **(No WAL)**	**“Double”** **2 x db_bench** **Rocksdb + ZenFS** **2 x 32 Zones 9G Each** **(No WAL)**	**“Double Quiet”** **2 x db_bench** **Rocksdb + ZenFS** **2 x 32 Zones 9G Each** **(No WAL)**
fillseq	830.3	836.1	980.7*	1097.1*
fillrandom	737.7	746.9	801.1*	875.4*
overwrite	733.5	744.1	797.5*	864.8*