Fast MRAM Write Buffers Make I/O Determinism Practical

Rizwan Ahmed
Everspin
IO Determinism is a Major Pain Point for Hyperscalers
Read Latency Challenge in Mixed Work Load

90% Random 4K Read, 10% 4K Write Latency Distribution

Source: “Solving Latency Challenge With NVM Express SSDs at Scale” - Chris Peterson, Facebook. FMS 2017
Current SSD Challenges That Can Be Addressed by STT-MRAM

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limited buffer size causes poor performance specifically QoS</td>
<td>Capacitors take too much board space limiting SSD capacity</td>
</tr>
<tr>
<td>Capacitors are leading cause of drive failures</td>
<td></td>
</tr>
<tr>
<td>In 3D TLC, sequential read performance can be impacted if buffer size is small</td>
<td>Longer time to market—PFail test is rigorous</td>
</tr>
<tr>
<td>Customers want more streams for better performance &amp; write amplification</td>
<td></td>
</tr>
</tbody>
</table>
MRAM Enables Bigger Write Buffer

- Capacitors limit write buffer size
- Only 4-20MB write buffer size in typical enterprise SSDs
- STT-MRAM can provide 64MB or more write buffer size
- Up to 2 STT-MRAMs placement for M.2
- Up to 4 STT-MRAMs placement for U.2
Significant QoS Improvement Using STT-MRAM

- Mixed workload QoS is critical performance metric
- Host Reads are blocked by
  - Host writes
  - Garbage collection
  - Erase operations
- Bigger write buffer absorbs more writes
- Significant QoS improvement expected as a function of MRAM buffer size

Exceedance chart for 70/30 R/W Workload

Improvement is based on buffer size
Capacitors are Leading Cause of Drive Failures
STT-MRAM Improves Product Reliability

- Capacitors are #1 cause of drive failure
- Redundant caps needed to meet product life spec
- STT-MRAM eliminates the need for capacitors
- MRAM has higher temperature operating range (0c-85c)
STT-MRAM Enables Higher SSD Capacity

- Capacitors take up lot of board space
- Optimize board layout by removing capacitors
- Up to 4 additional NAND placements possible
- Up to 25% additional SSD capacity possible
Faster Time to Market Using STT-MRAM
Study by Ohio State University and HP labs on “Robustness of SSD under Power Fault”

- Complicated code for managing power fail scenarios
- Lots of corner cases to validate
- Simplifying power fail code leveraging MRAM
- Leads to faster product qualification

- Tested 15 drives from 5 different vendors
- 13 drives failed for power fault

STT-MRAM Improves Sequential Read Performance in 3D NAND

- 2D NAND to 3D NAND transition requires bigger buffer size
  - Upper & lower page programmed together
- Data layout on NAND is not optimal with current buffer size
  - Sequential LBAs are written to upper and lower page
  - 2 NAND reads are needed to complete one host command
- Bigger NAND die capacity will continue to increase buffer size requirement
- 25-30% Sequential read improvement possible with optimal write buffer size
STT-MRAM Enables Higher Number of Streams Improving Performance & Write Amplification

- NVMe streams provide significant performance improvement
- Number of streams is limited by buffer size due to RAID/XOR operations
- Bigger buffer size can support more streams, resulting in better performance and WA

Source: “Multi-Stream Write SSD” - Changho Choi, Samsung. FMS 2016
IBM® Selected Everspin MRAM to Power Their FlashCore® Modules

Now Shipping in the IBM FlashSystem® 9100 Array
SSDs with STT-MRAM will have:

- Better QoS & seq. read performance
- More reliable design
- Higher capacity
- Faster time to market

Available now
Thank you.