Table-Less Controllers for Persistent Memory in Datacenter Applications

Bernard Shung, Ph.D.
President and Co-Founder, Wolley Inc.
Outline

• Introduction to Storage Class Memory (SCM)
• Address Indirection Table (AIT) in NAND FTL
  • Why it is not a good fit for SCM
• Table-less Controller Architecture for SCM
  • Technology feasibility – In-place write
  • Architecture challenges – “Start-Gap” not good enough
• Results of Wolley’s Table-less SCM Controller
Storage Class Memory (SCM)

- This storage organization can be thought of as a pyramid:

  Interface & Speed like Memory
  Capacity & Non-volatility like Storage
Address Indirection Table (AIT) is commonly employed in NAND FTL today
  - To address finite write endurance and wear leveling
  - Table is about 0.1% of storage size
  - Two forms of implementations
    - Stored in DRAM – faster, need SPOR handling
    - Stored in NAND – slower, need to store in SLC
Table-based Controller Not Scalable for Memory Applications

<table>
<thead>
<tr>
<th></th>
<th>SSD</th>
<th>NVDIMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>1 TB</td>
<td>256 GB</td>
</tr>
<tr>
<td>Access Unit</td>
<td>4 KB</td>
<td>64 B</td>
</tr>
<tr>
<td>Table Size</td>
<td>1 GB (0.1%)</td>
<td>16 GB (6.4%)</td>
</tr>
</tbody>
</table>

- If the table is stored in DRAM, then the super-cap won’t fit inside the DIMM!
  (A proven challenge of NVDIMM-N)
Table Inside SCM?

- Lower performance though penalty smaller with SCM than NAND
- >6.4% storage overhead to avoid the table being “wear-leveling bottleneck”
- Data write and Table write not ”atomic” – potential synchronization issue
  - This was why eMMC/UFS firmware was tougher
Concept of A Table-Less Controller

- Physical Address = F (Logical Address, Controller State)
- The physical address is "computed" from the logical address and the state of the controller, hence no table is needed
- The controller state needs to be maintained in the NV domain to recover from power loss
Why Table-Less Controllers Possible for SCM but not for NAND?

- NAND cannot write “in-place” but SCM can
- NAND always writes to a new, erased block, and cannot write to the same location as last time
  - $\text{PA} = F(\text{LA, State})$ cannot produce different PA given the same LA unless State is changing very rapidly
  - Table is the best solution for NAND
- For SCM, since in-place write is doable, the table-less architecture is possible
“Start-Gap” Table-Less Controller (IBM 2009)

• Divide into many structures like above, each using N+1 location to store N values
• For each structure, \( X = LA + \text{Start} \mod (N+1) \), If \( X \neq \text{Gap} \) then \( PA = X \), else \( PA = X + 1 \)
• \{Start, Gap\} of all structures constitute the State of the controller
• Move Gap when there are 100 writes to the structure; Move Start when it is hit by Gap
Start-Gap is Not Good Enough

• Start-Gap is simple, authentically table-less, but not implementable
  • The Gap cannot move very fast for synchronization and performance concerns
  • A location is a “sitting duck” until the Gap passes it through
  • If the structure is large (large N), the sitting duck issue is more severe
  • If the structure is small (small N), there will be too many structures
  • Do the exercise with write endurance = $10^6$, and 4G total locations
Calling for Architecture Innovation

• Problem Statement:
  Finding a mapping function, $PA = F(LA, State)$, such that
  1. Meets device write endurance constraint
  2. Minimizes state transition synchronization
  3. Has manageable complexity

• Reward: higher performance, lower cost, new “SCM translation layer” de-facto standard

• Existence Proof: Wolley has come up with one solution
Wolley’s Table-Less Controller Prototype

- SCM emulated by DRAM – additional delay added in FPGA to emulate SCM latency
- PCIe Gen3x8 interface supporting memory (64B) mode and NVMe SSD (4KB) mode
  - 64B random IOPs: 16M read, 10M write
  - 4KB random IOPs: 770K read, 700K write
- SPOR firmware tested extensively to prove robustness
- Gen1 with 8GB capacity; Gen2 with 256GB capacity
Handling SCM Device Issues

• Controller needs to handle other SCM device issues
  • Bad location management, ECC, Read/Write disturb, Drift, etc.
• Most of these issues are orthogonal to LA -> PA address mapping, but some are related
  • For example, table-based or table-less architecture has an impact on bad location management
• Based on our research, we believe table-less controllers are capable and effective to handle these SCM device issues
Summary

• SCM *should* and *could* employ a new controller
  • Table-based architecture does not scale for memory applications
  • SCM device characteristics enables new controller architecture
• Table-less SCM controllers could achieve higher performance and lower cost
• Innovation needed to find a good table-less controller
  • Existing Start-Gap controller is analyzed to show design criteria
• Wolley has a promising table-less SCM controller prototype
  • Please come to see our demo in Booth #806
Thank You!

cbshung@wolleytech.com

www.wolleytech.com