SNIA DSI '16 DATA STORAGE INNOVATION CONFERENCE
Innovation in Storage Products, Services, and Solutions
June 13-15, 2016 | Marriott San Mateo | San Mateo, CA

Increasing SSD Performance and Lifetime with Multi-stream Technology

Changho Choi, PhD
Principal Engineer

Memory Solutions Lab.
Samsung Semiconductor, Inc.
Samsung Electronics, Ltd.

Disclaimer

This presentation is intended to provide information concerning SSD technology. We do our best to make sure that information presented is accurate and fully up-to-date. However, the presentation may be subject to technical inaccuracies, information that is not up-to-date or typographical errors. As a consequence, Samsung does not in any way guarantee the accuracy or completeness of information provided on this presentation. Samsung reserves the right to make improvements, corrections and/or changes to this presentation at any time.

The information in this presentation or accompanying oral statements may include forward-looking statements. These forward-looking statements include all matters that are not historical facts, statements regarding the Samsung Electronics' intentions, beliefs or current expectations concerning, among other things, market prospects, growth, strategies, and the industry in which Samsung operates. 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. Samsung cautions you that forward looking statements are not guarantees of future performance and that the actual developments of Samsung, the market, or industry in which Samsung operates may differ materially from those made or suggested by the forward-looking statements contained in this presentation or in the accompanying oral statements. In addition, even if the information contained herein or the oral statements are shown to be accurate, those developments may not be indicative developments in future periods. The information is provided as a general understanding and not directly representing any product.

Agenda

- NAND flash characteristics
- Multi-stream
  - Multi-stream concept
  - Multi-stream system architecture
  - Multi-stream operation
- Performance benefit
- Stadards
- Summary
- Q&A

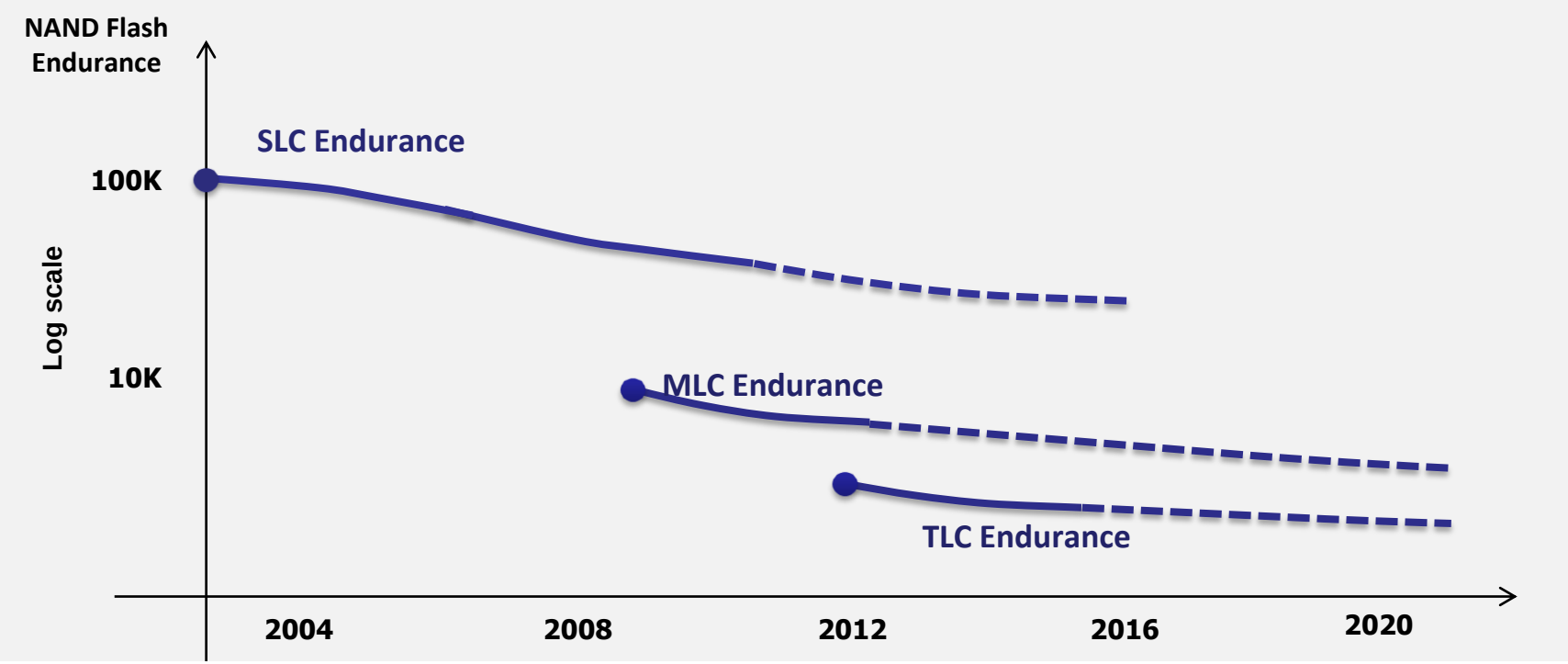
NAND Flash Characteristics

- Operation
  - Read/Program/Erase
- Operation unit
  - Read/Program: Page
  - Erase: block (= multiple pages)
- Out-of-place update
  - In-place update(=overwrite) NOT allowed
    - Invalidate overwritten data
  - Page MUST be erased before programming(writes)
    - Program/Erase (P/E) cycles
  - Need garbage collection operation

Efficient data placement increases performance
with reduced garbage collection overhead

NAND Flash Characteristics (Cont'd)

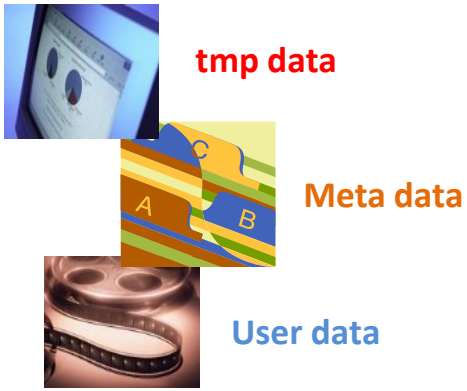
- Limited number of Program/Erase cycles

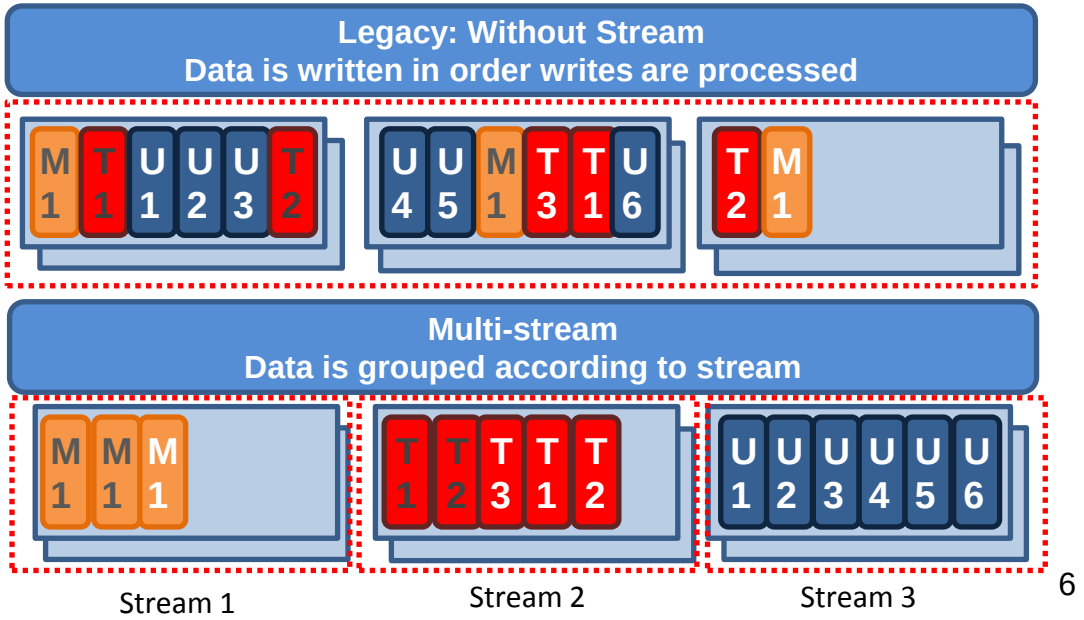


Efficient data placement increases lifetime of SSD(endurance)

Multi-stream

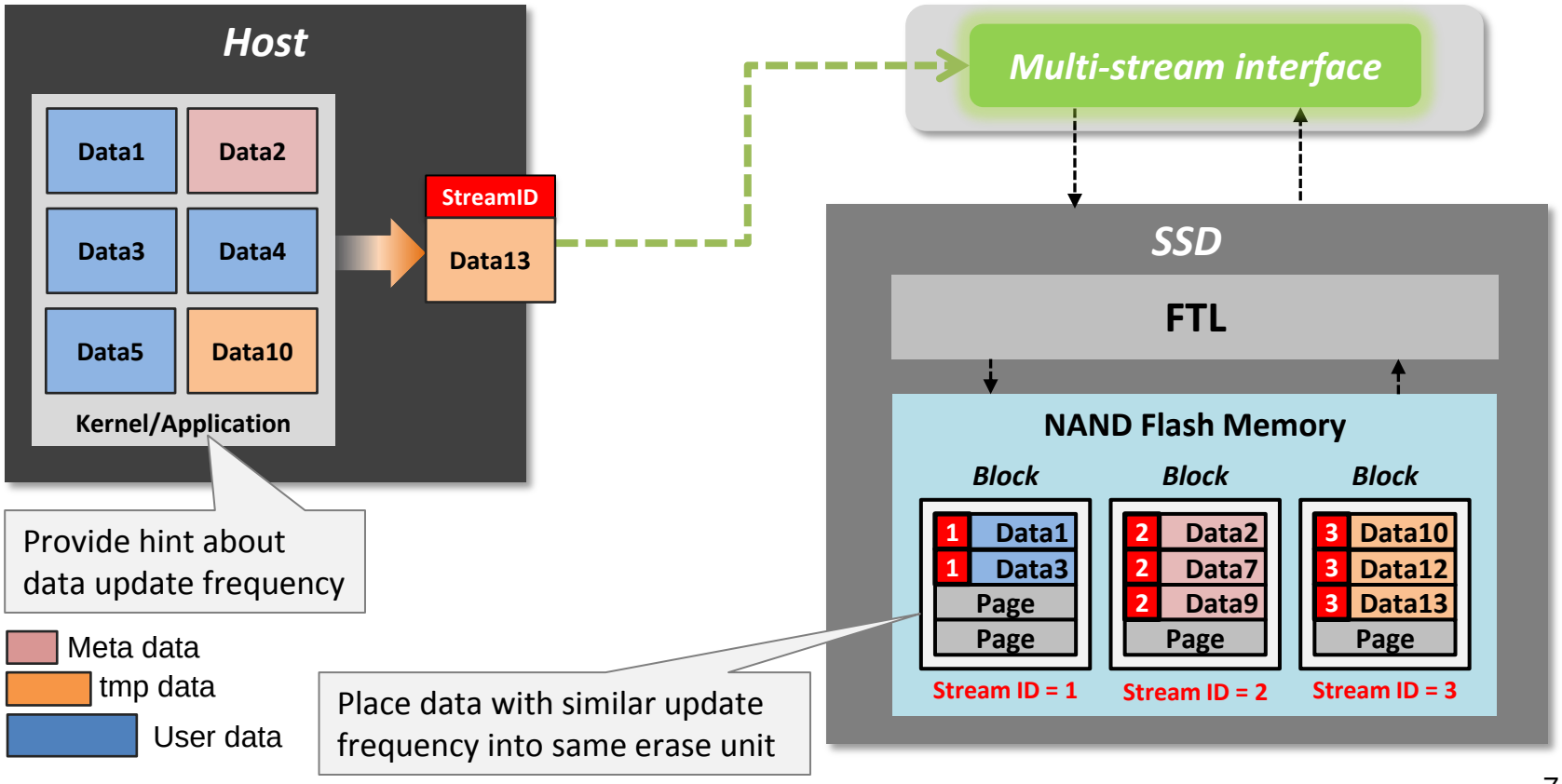
- Provide better endurance, improved performance, and consistent latency
  - Allow host to associate each write operation with a stream
  - All data associated with a stream is expected to be invalidated at the same time (e.g., updated, trimmed, unmapped, deallocated)
  - Align NAND block allocation based on application data characteristics(e.g., update frequency)





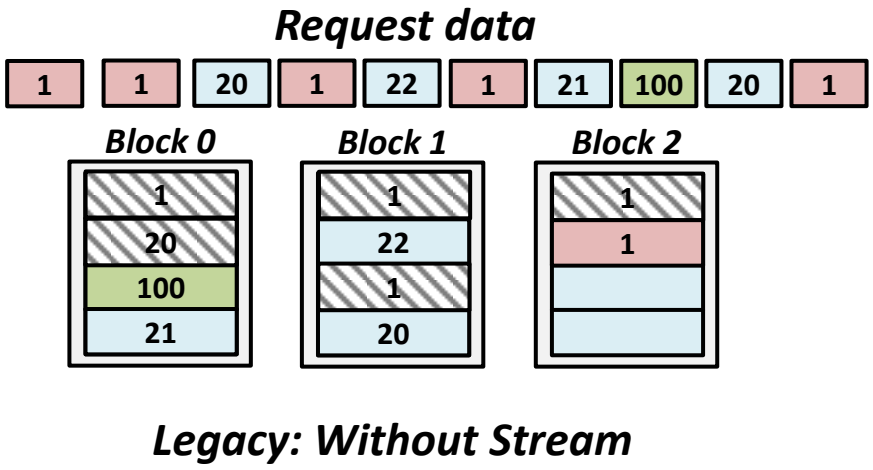
Multi-stream Operation

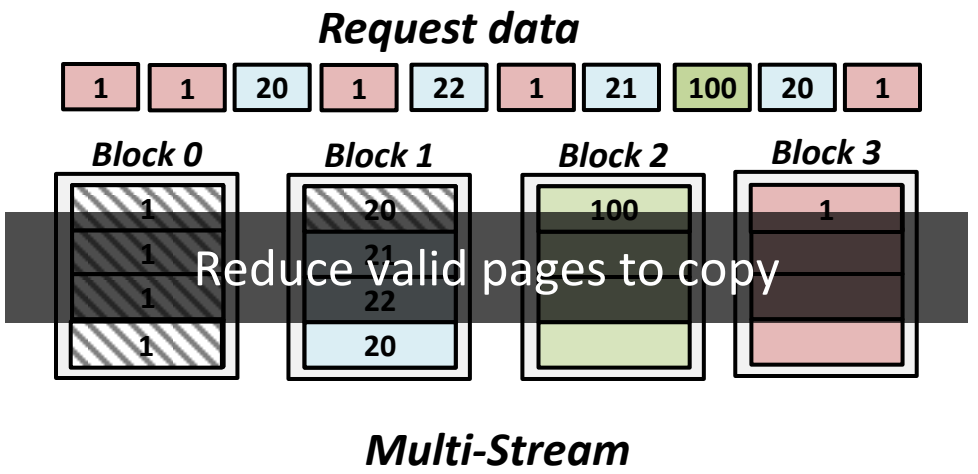
- Mapping data with different update frequency to different streams



Operation Example

- Efficient data placement with multi-stream
  - Reduce GC overheads -> better performance and lifetime!





For effective multi-streaming,
proper mapping of data to streams is important!

FIO Performance Measurement System

- Hardware
  - Quad Core Intel i7-4790 CPU 3.60GHz
  - 16GB memory
- Software
  - Ubuntu 14.04 LTS, v4.0.3 Kernel with multi-stream patch
  - FIO 2.2.5 with multi-stream patch
- Device
  - Multi-stream enabled NVMe SSD

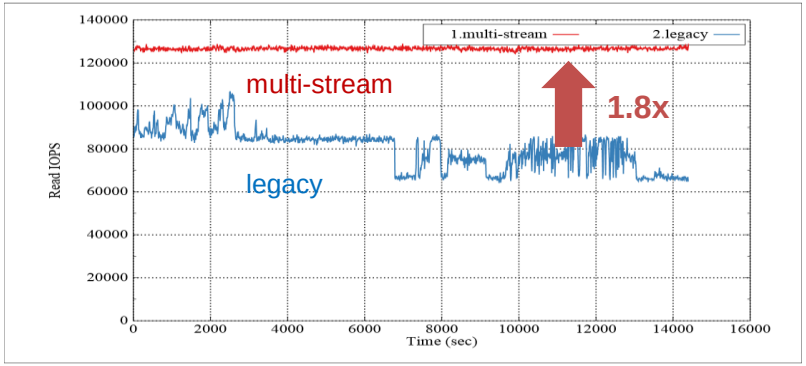
Performance Measurement Configuration

- Four sequential writes jobs 1+ random read job
  - Different data lifetime: 1x, 10x, 33x, 55x
- Precondition
  - 2 hours with four-write jobs

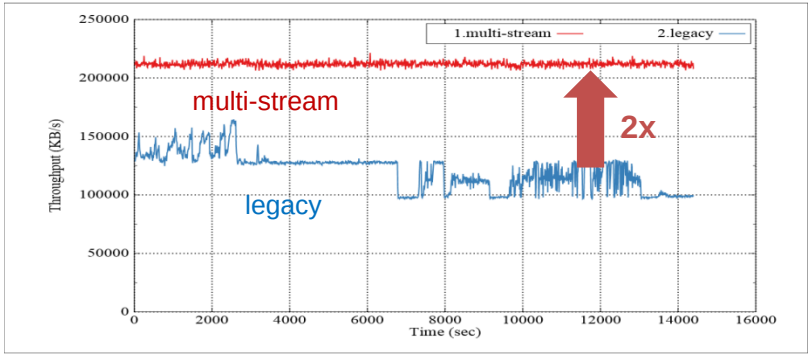
Four Streams – Read/Write(70%/30%)

- Reads
  - Jobs: 6
  - Block size: 4k
  - Iodepth: 64
- Writes
  - Jobs: 4
  - Block size: 128k
  - Iodepth: 1

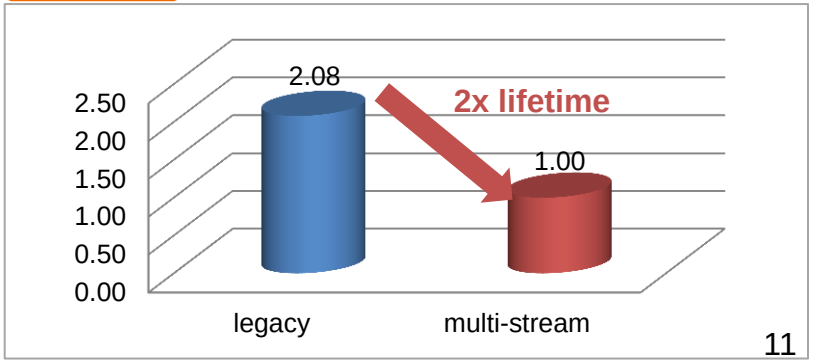
Read IOPS



Write Throughput



WAF



RocksDB

- Embedded NoSQL database
  - Storage directly attached application servers
  - Persistent Key-Value Store
- Optimized for fast storage (e.g., SSD)
- Uses Log Structured Merge(LSM) Tree architecture
- Server workloads

RocksDB Performance Measurement Configuration

Hardware environment

Processor/Memory Details	Operating System	SSD Details
Processor Dual Socket: Intel(R) Xeon(R) CPU E5-2640 v3 @ 2.60GHz. **Total Logical CPU :** 32 **Total memory :** 128 GB	**Distro :** Ubuntu 14.04.1 LTS, **Kernel :** 3.19.0-11-generic with multi-stream patch **Arch :** x86_64	**SSD :** Multi-stream enabled NVMe SSD

Software

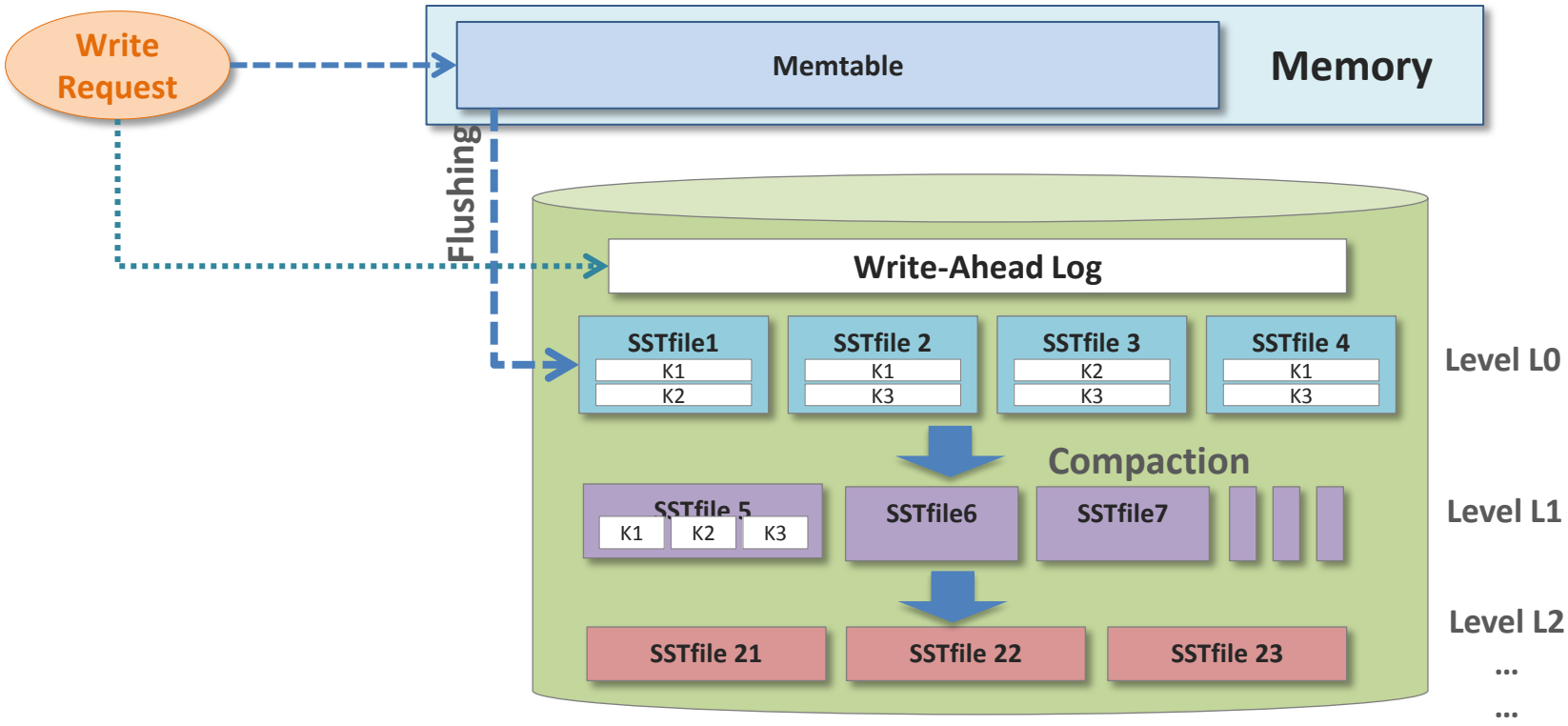
Software	Functionality	Version/Remarks
RocksDB	Persistent Embedded Key Value Store	Modified to add Multi-stream support
YCSB	Yahoo Cloud Benchmark Tool	0.1.4
SSDB-Rocks	Provides an interface to RocksDB for YCSB	1.6.6

Workload

Parameter	Value
Read/Update Mix	50%/50%
Pre-inserted records	370 Million
No of Operation	1.2 Billion
No of YCSB Threads	32

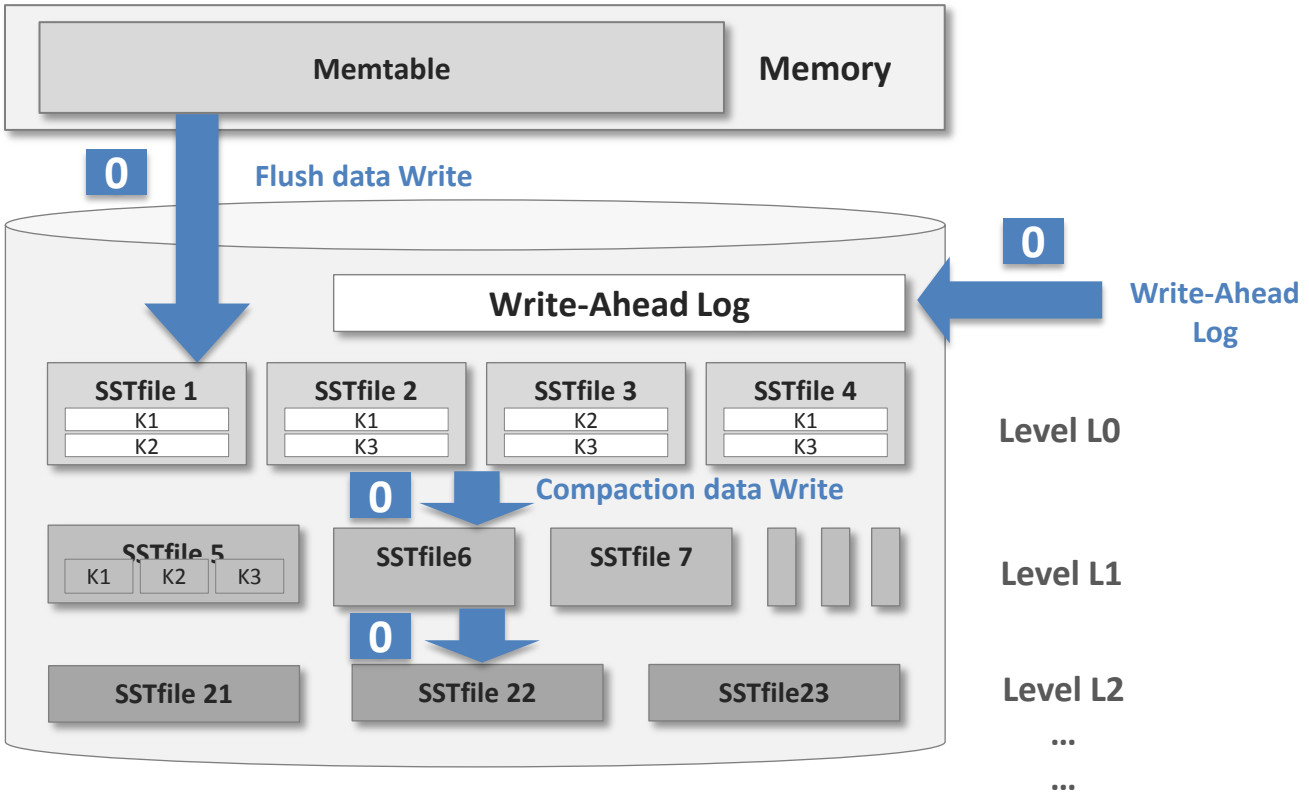
RocksDB Architecture

- Level-tiered compaction



RocksDB in Legacy SSD

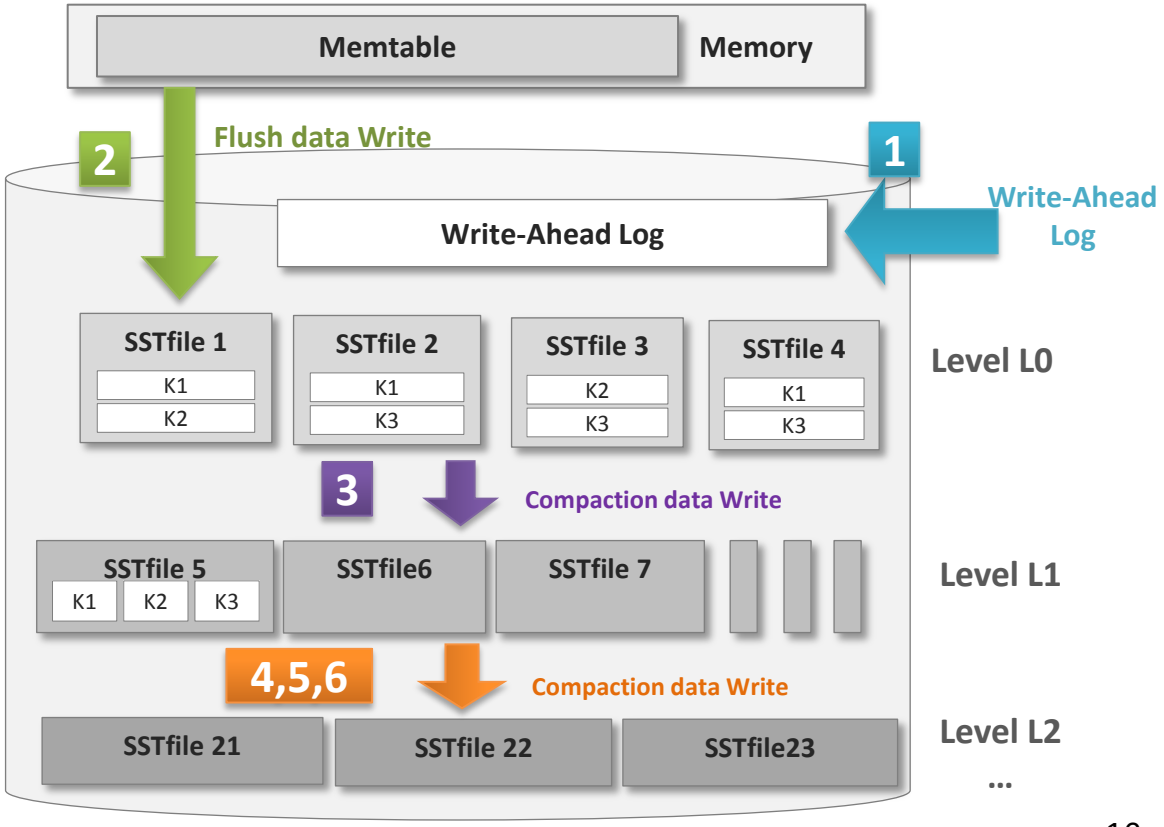
- Legacy SSD (same as a single stream ID case)



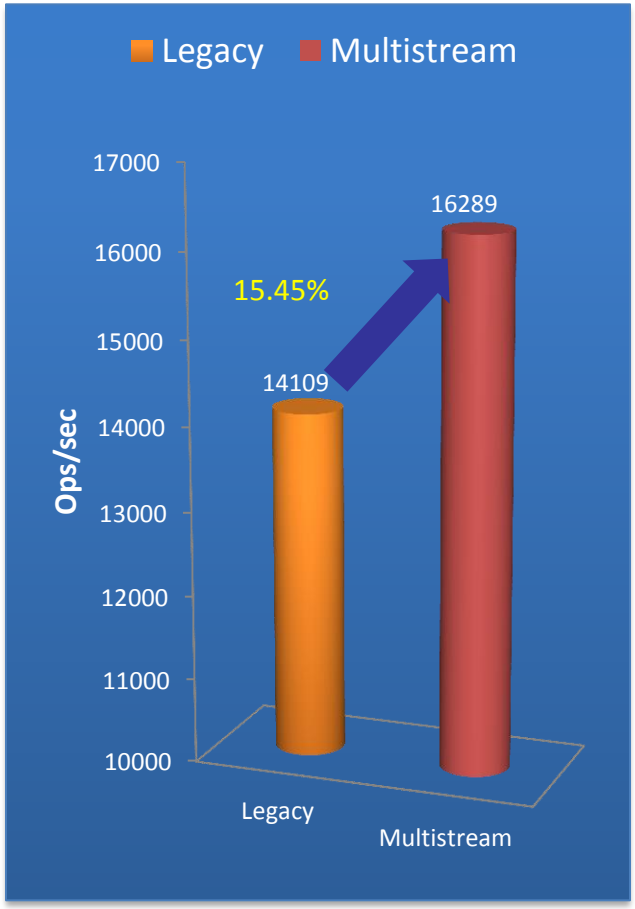
RocksDB with Multi-stream

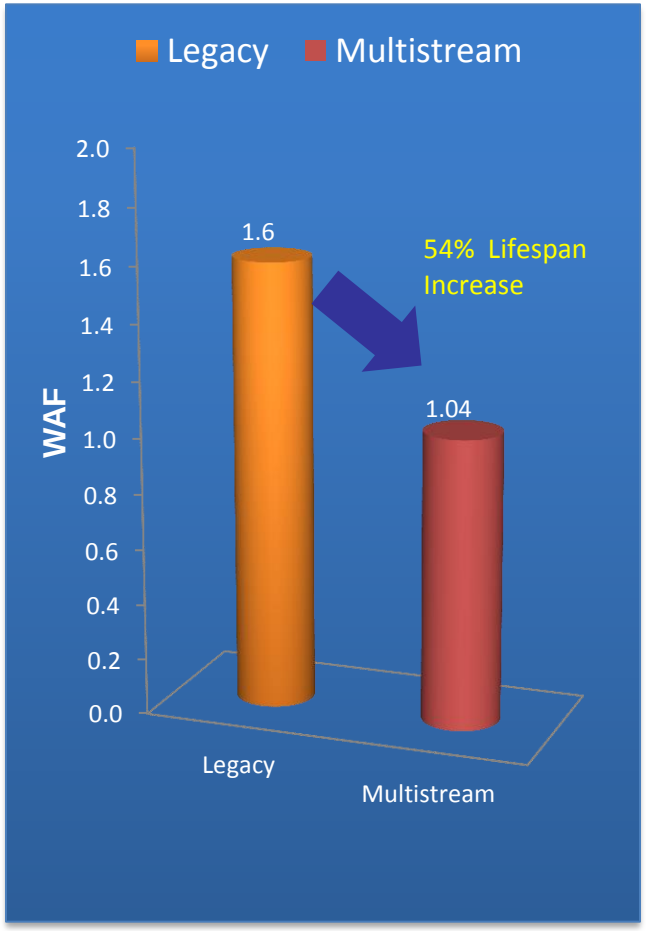
- Assign stream IDs according to SSTfile levels

Stream ID	Level
1	Write-Ahead Log
2	L0
3	L1
4, 5, 6	L2
7, 8, 9	L3
10,11,12	L4
13, 14	L5
No data on Level 6 due to dataset size	L6

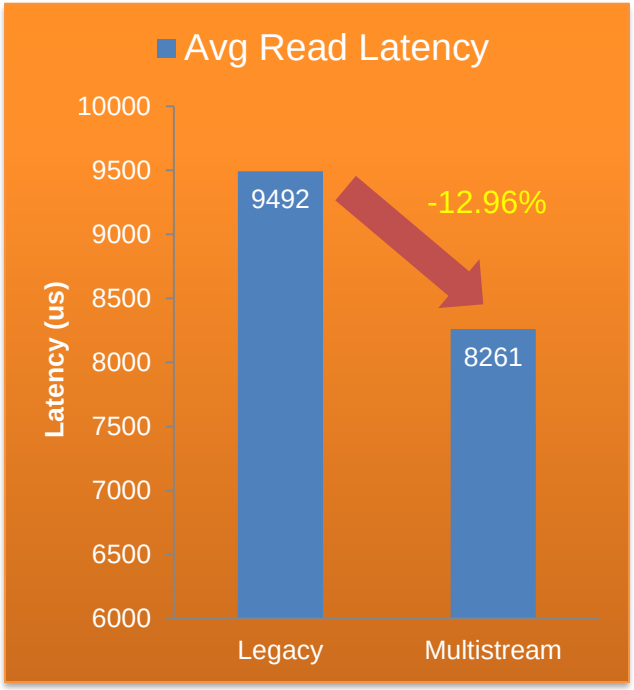


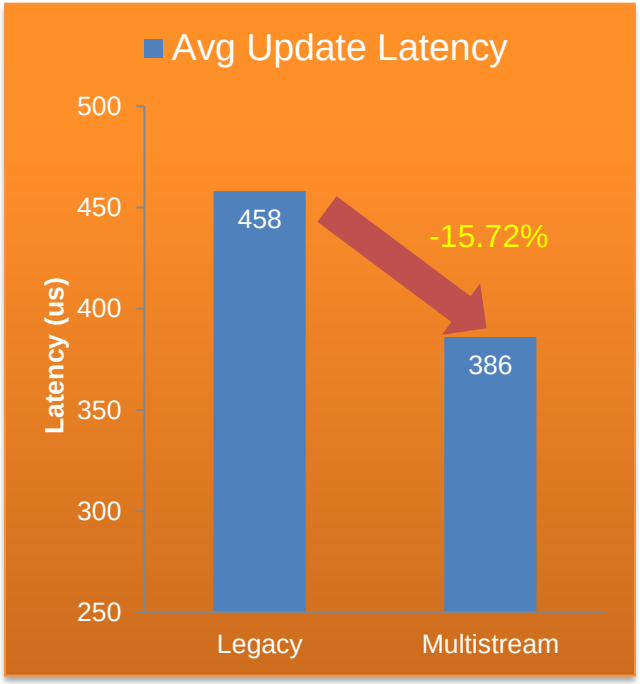
RocksDB:
15%+ Performance and 54% Lifespan





RocksDB: 10%+ Better Average Latency





Standards

- SCSI/SAS: Completed in May, 2015
  - Standard spec:

    http://www.t10.org/cgi-bin/ac.pl?t=f&f=sbc4r10.pdf

- NVMe: standardization ongoing
- SATA: standardization ongoing

Summary

- With multi-stream, SSDs can be more efficiently used for
  - Consistent better performance
  - Better endurance (=better SSD lifetime)
- With multi-stream
  - FIO: more than 2x SSD lifetime in addition to the decent I/O performance enhancement
  - RocksDB: more than 50% SSD lifetime as well as more than 15% I/O performance improvement