Storage Intelligence in SSDs and Standards

Bill Martin, Principal Engineer
Changho Choi, PhD, Principal Engineer

Memory Solutions Lab
Samsung Semiconductor, Inc.
What issues are we addressing?

- Currently hosts have no mechanism to understand the storage device internal features
  - Inefficient operation of background operations
  - Inefficient placement of data
- Current technology requires multiple translation layers
  - Key/Value to Block Storage
- Data and computational processing is not co-located
  - Increased IO traffic
  - Under-utilized compute power in storage device
  - Over utilized compute power in host/storage system
How do we solve these issues?

- Currently hosts have no mechanism to understand the storage device internal features
  - Inefficient operation of background operations
  - Inefficient placement of data
- Current technology requires multiple translation layers
  - Key/Value to Block Storage
- Data and computational processing is not co-located
  - Increased IO traffic
  - Under-utilized compute power in storage device
  - Over utilized compute power in host/storage system
What is Storage Intelligence?

- An interface to provide better collaboration between SSD and storage systems
  - Background operation control
    - Advanced garbage collection
  - Stream operation
    - Stores data with similar lifetime in associated physical locations
- A mechanism to offload performance operations to SSD
  - Object Storage
    - Defines a Key Value Storage API
  - In-Storage Compute
  - Framework for offloading processing to storage device
Background operation control

- Allows a host to control background operations
  - Set background operation mode
  - Start/Stop background operation
  - Retrieve background operation status

- Specifies a time period that the device may perform background operation with minimal impact to system performance

- Why background operation control?
  - IO performance is degrade when background operations occur at the same time as IO
  - Avoids overlap of IO and background operations

- Provides predictable and consistent performance
Predictable & consistent performance

- 3sec Idle Time
- 3sec FIO

Legacy SSD vs. SI-enabled SSD

Background Operation
Control Interface
Stream operation

- Allows host to associate each write operation with a stream
- Device places all data associated with a stream in physically associated locations
- All data associated with a stream is expected to be invalidated at the same time (e.g., trimmed, unmapped)

Why stream operation?
- When different lifetime data is intermixed
  - Garbage Collection overhead increases
  - Write Amplification Factor increases

- Improves system performance
- Improves device endurance
Stream comparison

**Non Stream**
Data is written in order, writes are processed

- Virtual Machine A
  - Database
  - Virtual Machine C

**Stream**
Data is grouped according to stream

- Virtual Machine A
- Database
- Virtual Machine C
Up to 9x performance and 3x SSD endurance

FIO 100% 128K writes with four different lifetime data
Object Storage

- Uses a Key Value Storage model (not block storage model)
- Key value mapping to physical location done by the storage device

Why object storage?
- Translation from Key Value to Block Storage protocol consumes host compute cycles and mapping must be stored in host
- Double logging occurs
- Key Value map may need to be retrieved from storage device at initialization time

- Reduce host compute for Key Value mapping
- Reduce host memory footprint
Object Storage Comparison

Current

Host
- Performs Mapping
- Performs data logging
- Transfers data
- Flushes map to device

Device
- Stores map
- Stores data
- Metadata logging

Object Storage

Host
- Transfers data

Device
- Performs Mapping
- Performs data logging
- Stores data
- Metadata logging
In-Storage Compute

- Offloads host compute to the storage device
- Allows host to download application to device for device processing
- Why In-Storage Compute?
  - High IO traffic caused by reading data, computing, and writing results
  - Unused device compute power and bandwidth
- Reduces IO traffic between storage and host
- Reduces host computing burden
- Enhances application/system performance and power consumption
In-Storage Compute

Current

In Storage Compute

Host
- Retrieves data
- Performs computation
- Generates result

Device
- Returns all data

Host
- Request Computation
- Retrieves result

Device
- Performs computation
- Returns result
Standardization process
Background Operation Control & Stream Operation

- Currently standardized for SCSI
  - Documented in SCSI BLOCK Commands – 4 (SBC-4)
- Proposal being considered for SATA
  - Current proposal f15123r1
  - Expected completion December 2015
- Approved as work item for NVMe
  - Being discussed prior to full NVMe technical group discussions
  - Bring in to NVME technical group in November
  - Expected ratification March 2016
Standardization process
Object Storage & In-Storage Compute

- Object Storage
  - Being developed in SNIA Object Drive TWG
    - Requirements document well developed
    - API document to be started in the near future

- In-Storage Compute
  - Being developed in SNIA Object Drive TWG
    - Requirements document well developed
    - API document to be started in the near future

- Management of IP Drives
  - Being developed in SNIA Object Drive TWG
    - Requirements document well developed
    - Outline of standard started in July
Call for Action

- To get involved in the standardization process contact
  Bill Martin, bill.martin@ssi.samsung.com
- For questions about Samsung’s implementation contact
  Changho Choi, changho.c@ssi.samsung.com
Thank You