

Hybrid Solid State Drives for Improved Performance and Enhanced Lifetime

Yongseok Oh Eunjae Lee Jongmoo Choi

Donghee Lee Sam H. Noh



University of Seoul
{ysoh, kckjn97, dhl_express}
@uos.ac.kr



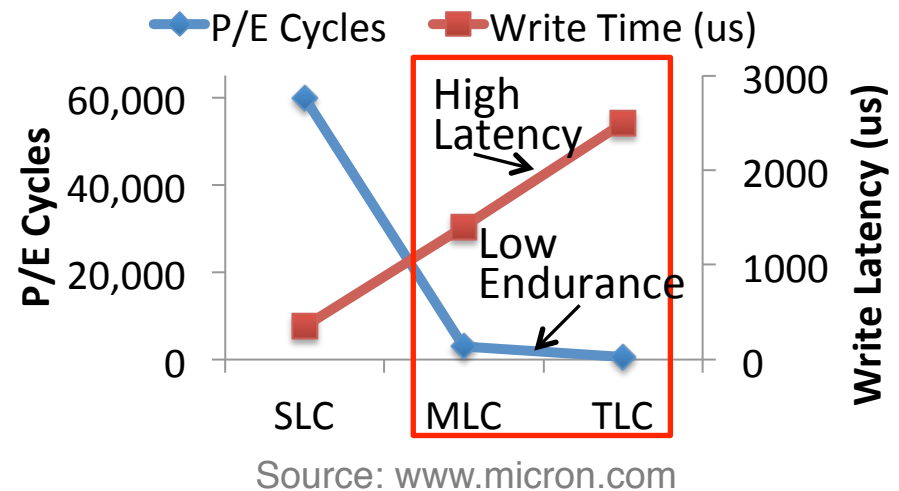
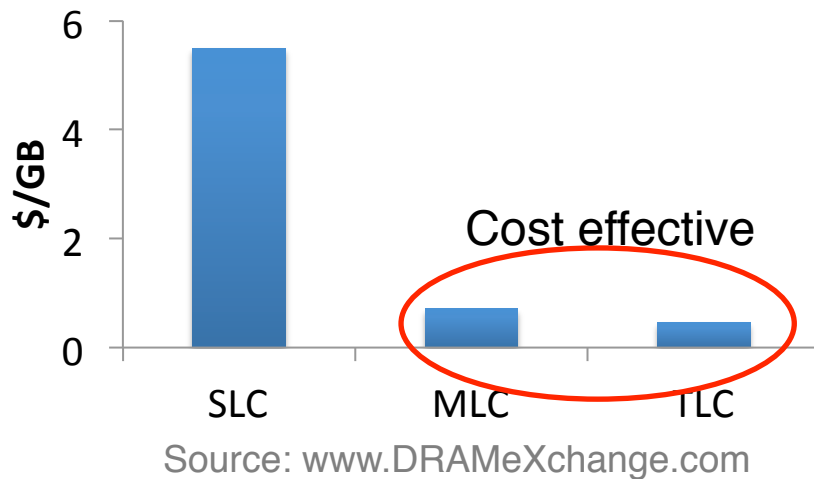
Dankook University
choijm@dankook.ac.kr



Hongik University
samhnoh@hongik.ac.kr

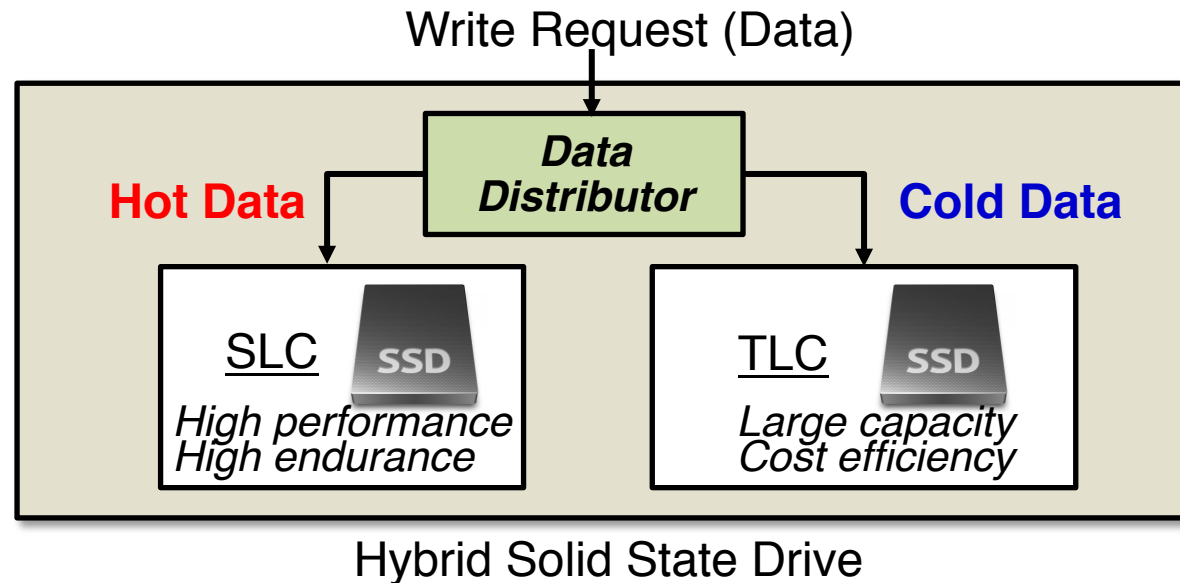
Introduction

- SSDs are being increasingly deployed
- MLC(2bits) and TLC(3bits) flash chips are considered
 - Pros: cost effectiveness
 - Cons: high write latency, low endurance
- **Hybrid Approach** is needed (e.g., SLC+MLC SSDs, SLC+TLC SSDs,)



Hybrid SSD Architectures

- Existing works
 - Li-Pin Chang [ASPDAC'08]
 - Sungjin Lee et al. [ATC'09]
 - Soojun Im et al. [SAC'09]
 - Muthukumar Murugan et al. [MASCOTS'12]
- Data distribution as a typical approach
 - Impact on performance and lifetime

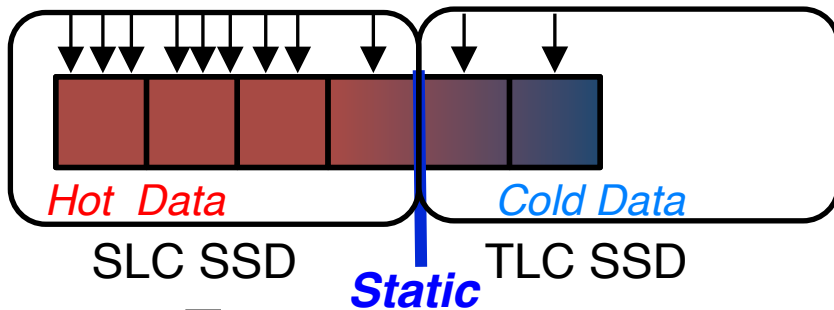


Problem and Motivation

Static Data Distribution

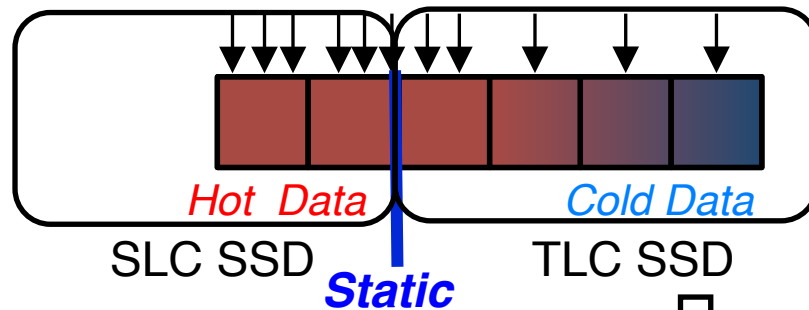
- Distributed based on amount of data
- How to determine optimal amount of hot data?
 - More hot data or less hot data to SLC SSD?

More hot data stored in SLC SSD



*Increased GC cost,
Shortened Lifetime
of SLC SSD*

Less hot data stored in SLC SSD



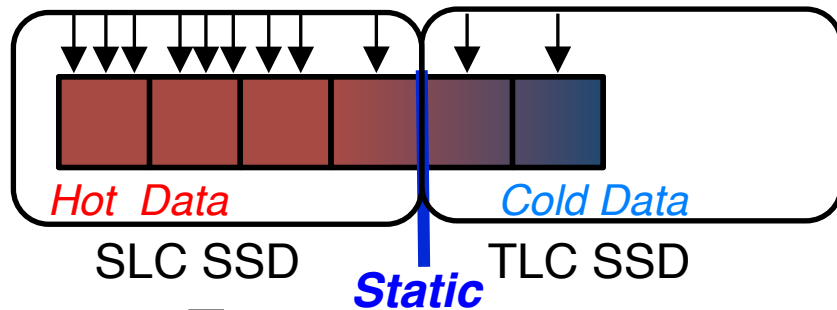
*Increased GC cost,
Shortened Lifetime
of TLC SSD*

Problem and Motivation

Static Data Distribution

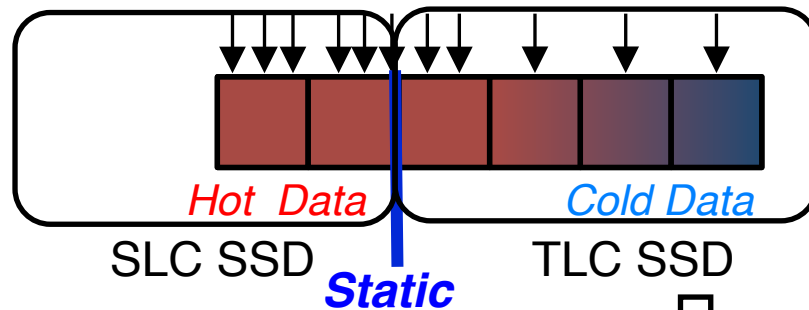
-
- Q: How do we adjust the distribution?
- A: Analytic Model based Dynamic Approach

More hot data stored in SLC SSD



*Increased GC cost,
Shortened Lifetime
of SLC SSD*

Less hot data stored in SLC SSD



*Increased GC cost,
Shortened Lifetime
of TLC SSD*

Analytic Model Based Dynamic Approach

- **Perf** scheme

- Re-distributes hot and cold data for the performance
- Employs analytic model to find the hot and cold data distribution
 - ◆ Determines optimal sizes of hot and cold data
- **Considers GC cost and workload pattern**

Analytical Model: C_{HYSSD}

← *Read the paper
for detailed derivations*

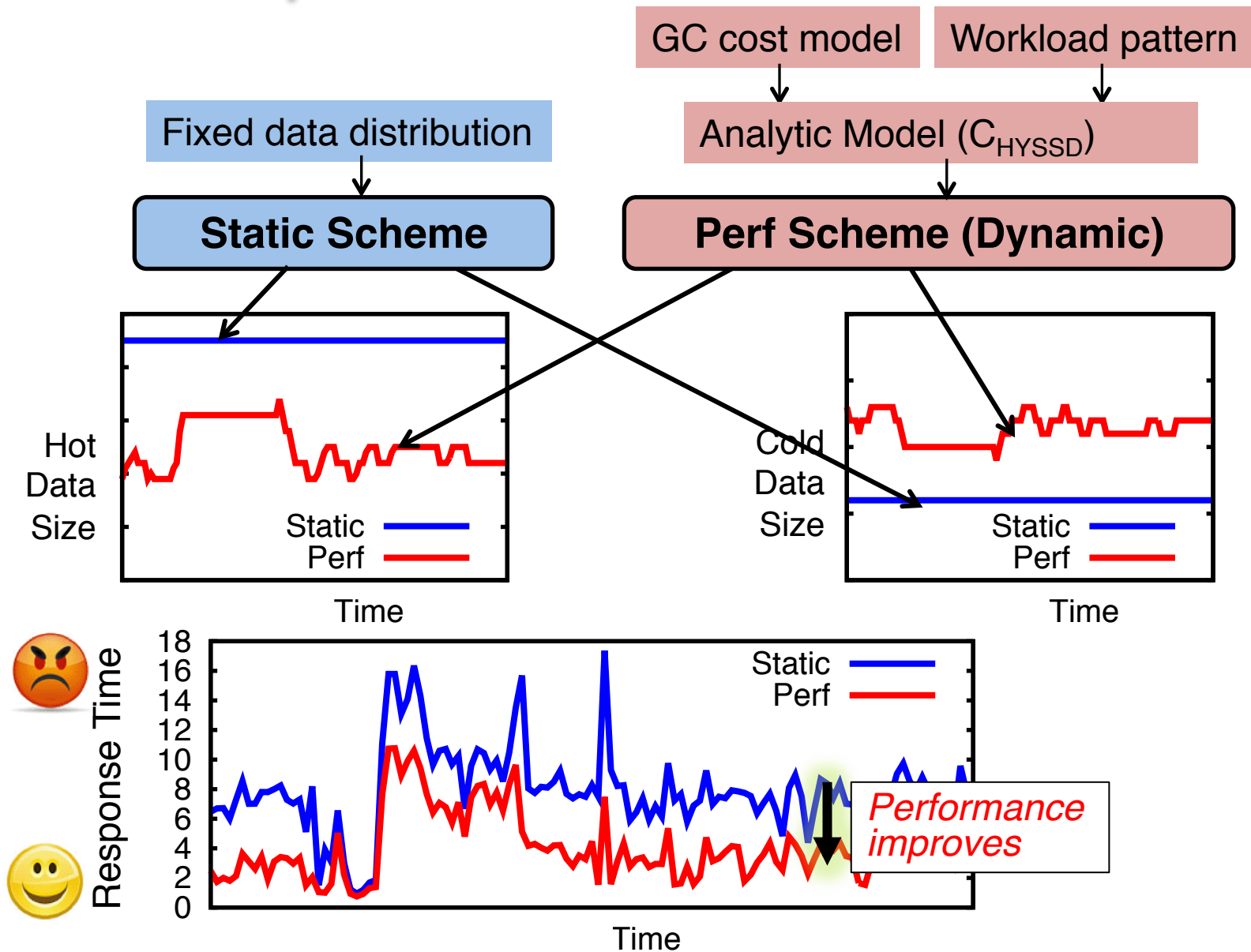
- **Life** scheme

- Calibrates performance optimized distribution of Perf scheme
- Uses relative remaining lifetime metric

$$L_{SLC} = \frac{E_{SLC}^{LEFT}}{E_{SLC}^{USED}} \text{ (SLC Lifetime)} \quad L_{TLC} = \frac{E_{TLC}^{LEFT}}{E_{TLC}^{USED}} \text{ (TLC Lifetime)}$$

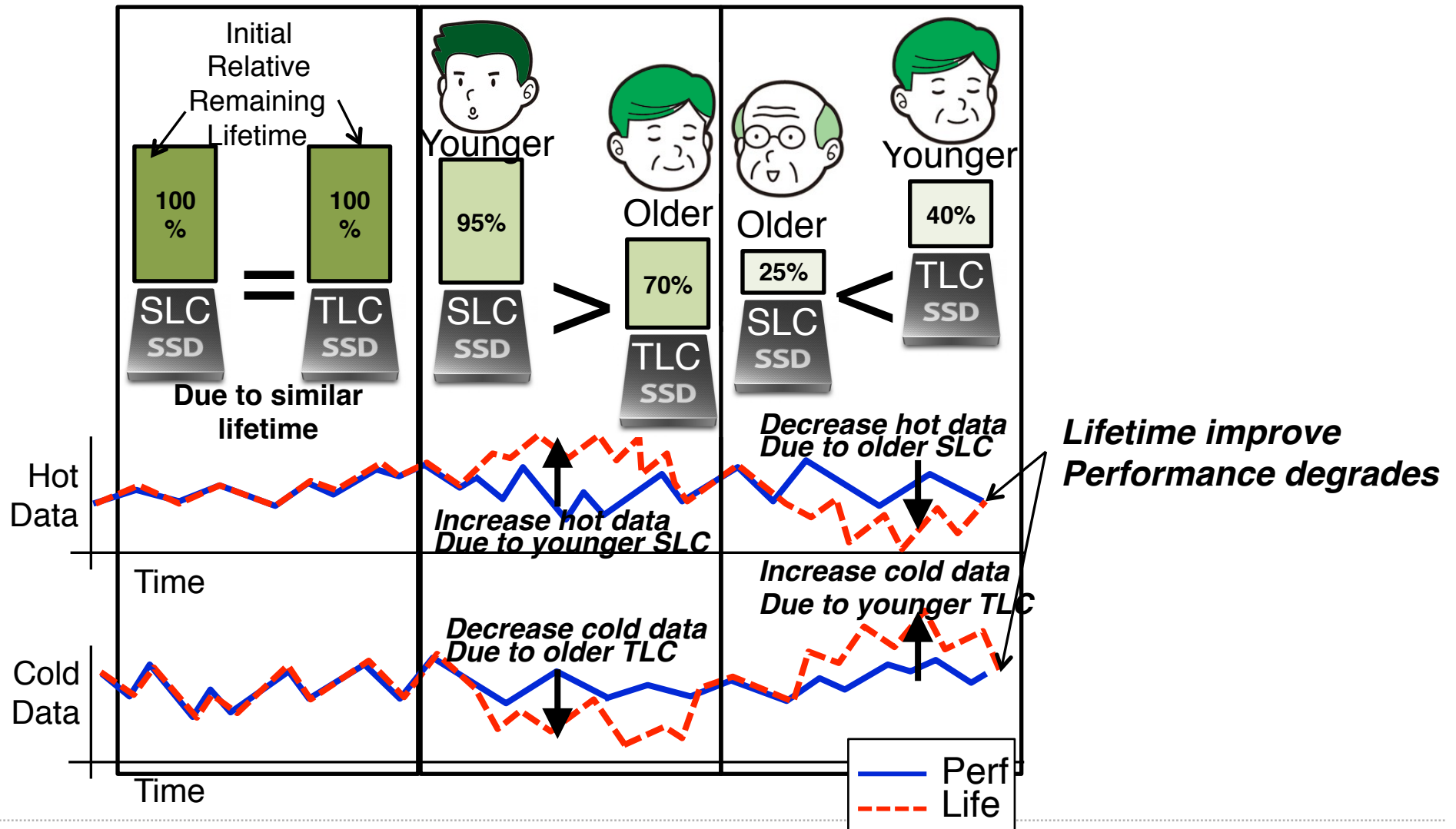
- ◆ $E_{SLC, TLC}^{LEFT}$: the remaining erase cycles
- ◆ $E_{SLC, TLC}^{USED}$: the used erase cycles

Performance Optimization



Lifetime Optimization

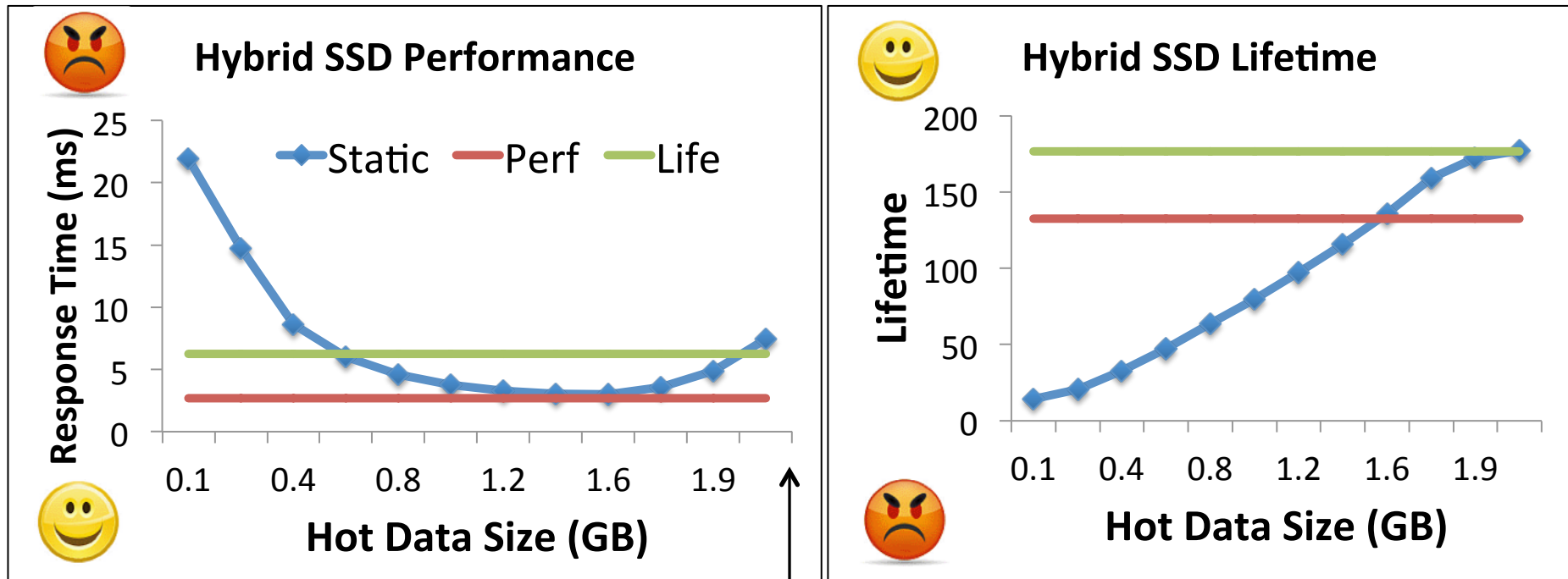
- Maximizing the lifetime of the hybrid SSD
 - Evenly utilize the lifetime of SSDs



Experimental Setup

- Hybrid SSD Simulator: DiskSim + SSD Extension
- Various schemes
 - **Static**: fixed size of hot and cold data
 - **Dynamic: Perf, Life**
- Configurations
 - ST-SSD: 2GB SLC + 16GB TLC SSDs
 - MT-SSD: 2GB MLC + 16GB TLC SSDs (in the paper)
- Usable address space: 16GB
 - Hot data size cannot be larger than capacity of SLC SSD
- Realistic workload traces: Exchange and Financial (in the paper)

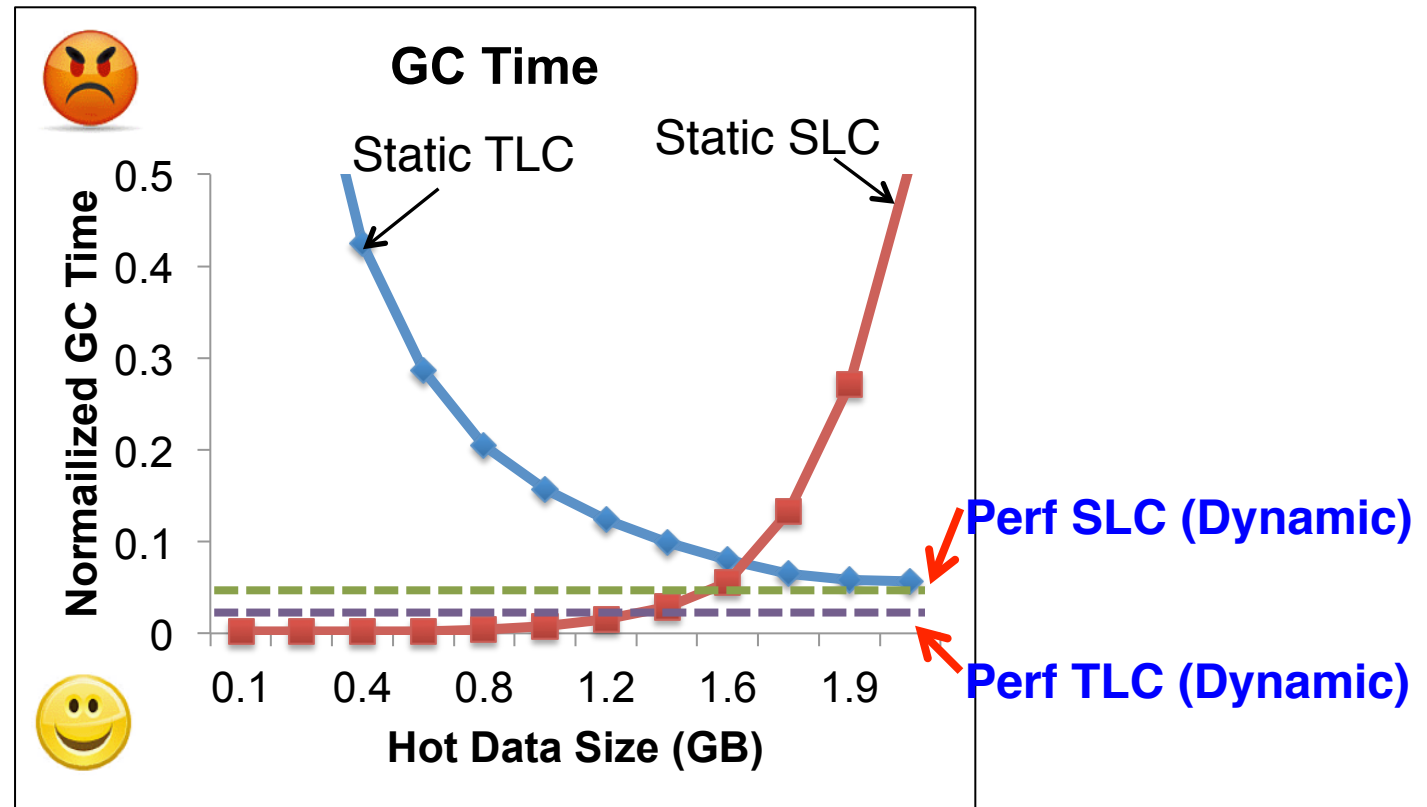
Evaluation Results



Maximum of SLC Capacity (2GB)

- *Static* scheme
 - Response times are varied according to hot data size
- *Perf* scheme achieves the improved performance
- *Life* scheme enhances the lifetime of a hybrid SSD
 - Lifetime metric: $\min(L_{\text{SLC}}, L_{\text{TLC}})$
 - Hybrid SSD is alive when two SSDs are working

Analysis of Result



- Static scheme shows varied GC times along with hot data size
- Perf scheme presents reduced GC time
 - GC cost is considered
 - (Life scheme is omitted due to similar trend)

Concluding Remarks

- Dynamic data distribution schemes
 - Analytical model approach
 - Perf scheme optimizes performance
 - Life scheme enhances lifetime
- Future direction
 - Memory efficient hot-cold data classifier
 - Emerging write amplification models

Thank You!

Hybrid Solid State Drives for Improved Performance and Enhanced Lifetime

Yongseok Oh Eunjae Lee Jongmoo Choi Donghee Lee Sam H. Noh



University of Seoul
{ysoh, kckjn97, dhl_express}
@uos.ac.kr



Dankook University
choijm@dankook.ac.kr



Hongik University
samhnoh@hongik.ac.kr