Maximizing Performance/cost of SSD Composed of Memory-type and Storage-type SCMs

Reika Kinoshita, Chihiro Matsui and Ken Takeuchi

Chuo University, Japan

Outline

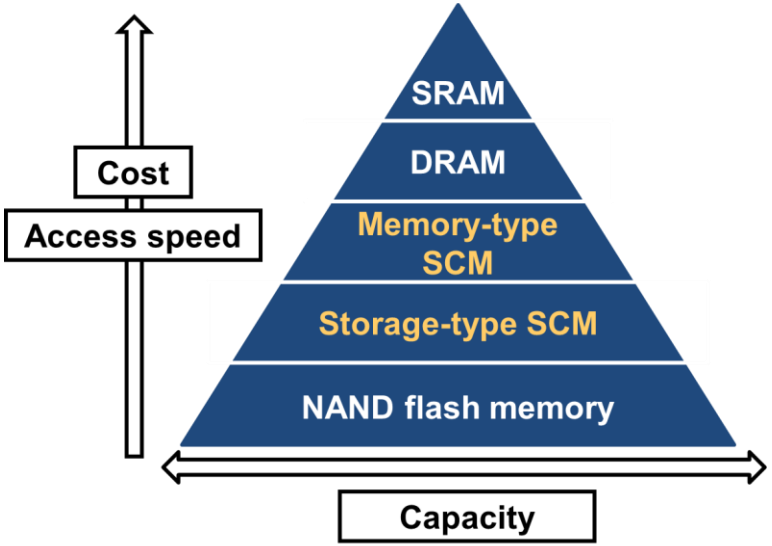
1. Introduction
2. Data Management Algorithm and Evaluation Setup
3. Optimal Composition of Multi-SCM SSD
4. Wear-leveling Policy for Multi-SCM SSD
    1. Performance effect of wear-leveling
    2. Optimal wear-leveling applying cases
5. Conclusion

Outline

1. **Introduction**
2. Data Management Algorithm and Evaluation Setup
3. Optimal Composition of Multi-SCM SSD
4. Wear-leveling Policy for Multi-SCM SSD
    1. Performance effect of wear-leveling
    2. Optimal wear-leveling applying cases
5. Conclusion

Storage Class Memories (SCMs)

- SCMs can be categorized into 2 types;
  - Memory-type SCM (M-SCM) [1] and Storage-type SCM (S-SCM) [2]

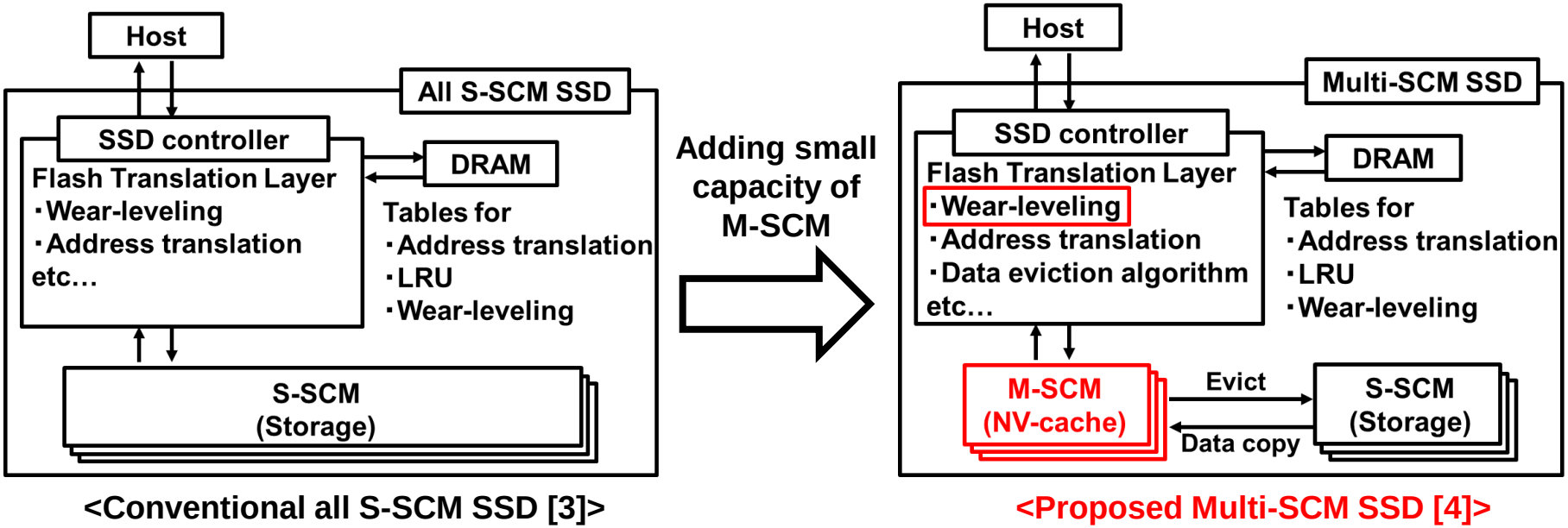


	M-SCM	S-SCM
Access unit	Sector (512 Bytes)	
Write/Read latency	100 ns / 100 ns	1 us / 1 us
Bit cost (TLC NAND flash = 1)	20	6
Endurance	10^9	10^5
Candidates	MRAM	ReRAM PRAM 3D XPoint

[1] J. J. Kan et al., TED, vol. 64, no. 9, pp. 3639-3646, Sep. 2017.
[2] Intel Optain Technology, https://ark.intel.com/content/www/us/en/ark/products/97159/intel-optane-ssd-dc-p4800x-series-1-5tb-1-2-height-pcie-x4-3d-xpoint.html

Proposed Multi-SCM SSD [4]

- M-SCM is used as non-volatile (NV) cache of S-SCM



<Conventional all S-SCM SSD [3]>

<Proposed Multi-SCM SSD [4]>

[3] T. Onagi *et al.*, *JJAP*, vol. 54, no. 4S, pp. 04DE04-1-04DE04-7, 2015.
[4] R. Kinoshita *et al.*, *NVMTS*, 2018, pp. 1-6.

Purpose of This Work

- ✓ **Optimal M-SCM capacity** is determined by total SSD cost and SSD performance
  - **Maximize performance/cost** of Multi-SCM SSD

- ✓ **Optimal wear-leveling** is investigated to boost SSD performance
  - Boost SSD performance **without reducing endurance lifetime** of Multi-SCM SSD

Outline

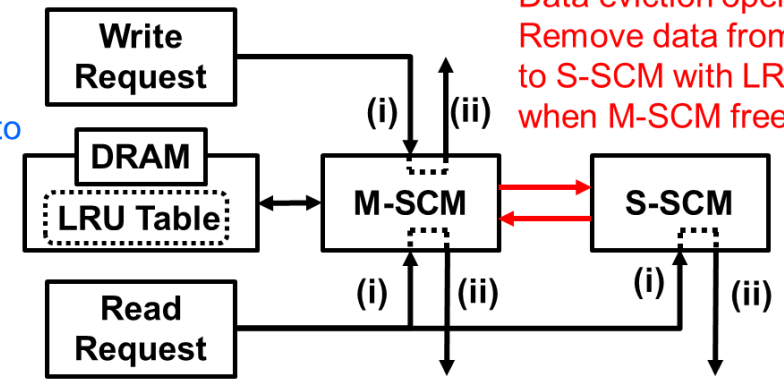
1. Introduction
2. **Data Management Algorithm and Evaluation Setup**
3. Optimal Composition of Multi-SCM SSD
4. Wear-leveling Policy for Multi-SCM SSD
    1. Performance effect of wear-leveling
    2. Optimal wear-leveling applying cases
5. Conclusion

Data Management Algorithm

- Non-volatile Write-back (NV-WB) algorithm stores frequently accessed data in M-SCM [5]

Write operation:
(i) Write request and update least recently used (LRU) table
(ii) Write completed message to host

Read operation:
(i) Read request and update LRU table
(ii) Read completed message to host



Data eviction operation:
Remove data from M-SCM to S-SCM with LRU order when M-SCM free space < 20%

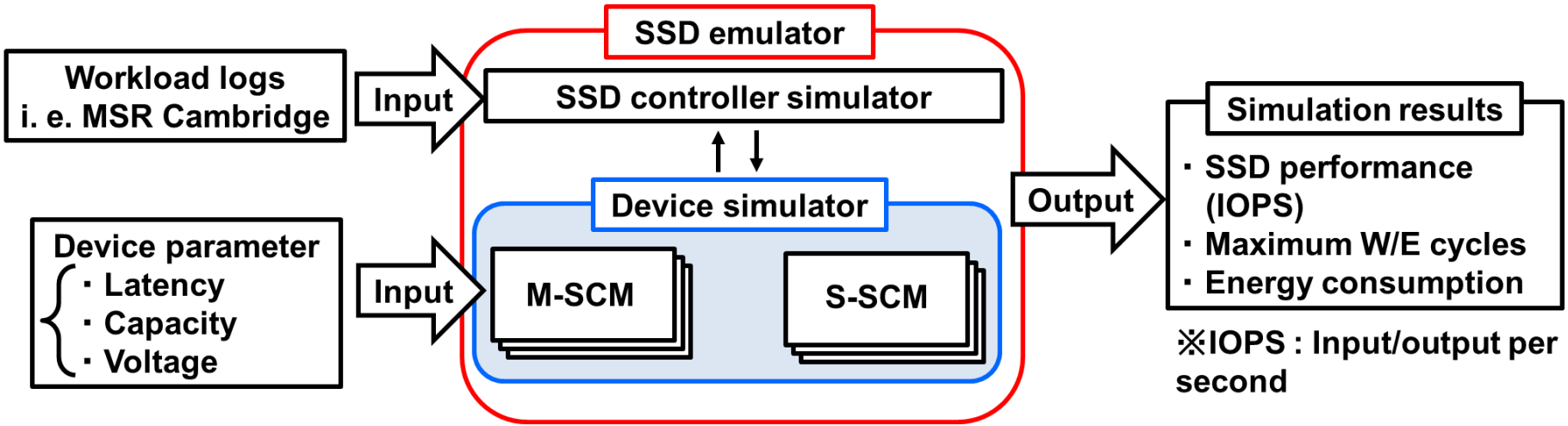
Write back operation:
Write back data to M-SCM when access count to S-SCM reaches threshold

Write and read operation : Sector unit (512 Bytes)
Data movement : Page unit (16 KBytes)

[5] S. Okamoto *et al.*, *IMW*, 2015, pp. 157-160.

SSD Emulator [6]

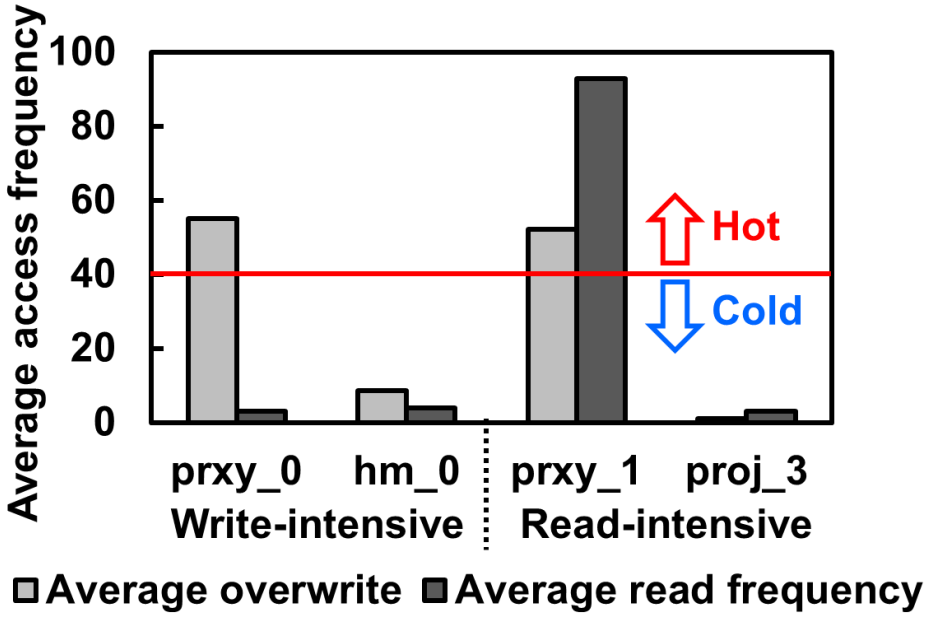
- Transaction Level Modeling emulator is used to evaluate SSD performance and endurance



[6] H. Fujii *et al.*, *VLSI Circuit Dig. Tech. Papers*, 2012, pp. 134-135.

Workloads [5, 7]

- For performance evaluation, 4 workloads are used



Workload	Userdata size (GB)	Category
prxy_0	0.98	Write-hot
hm_0	21.67	Write-cold
prxy_1	4.51	Read-hot
proj_3	5.90	Read-cold

- Average access frequency :
  - Average overwrite
    = Total write data size / User data size
  - Average read frequency
    = Total read data size / User data size
- Hot/Cold : Large/Small number of overwrite or re-read requests

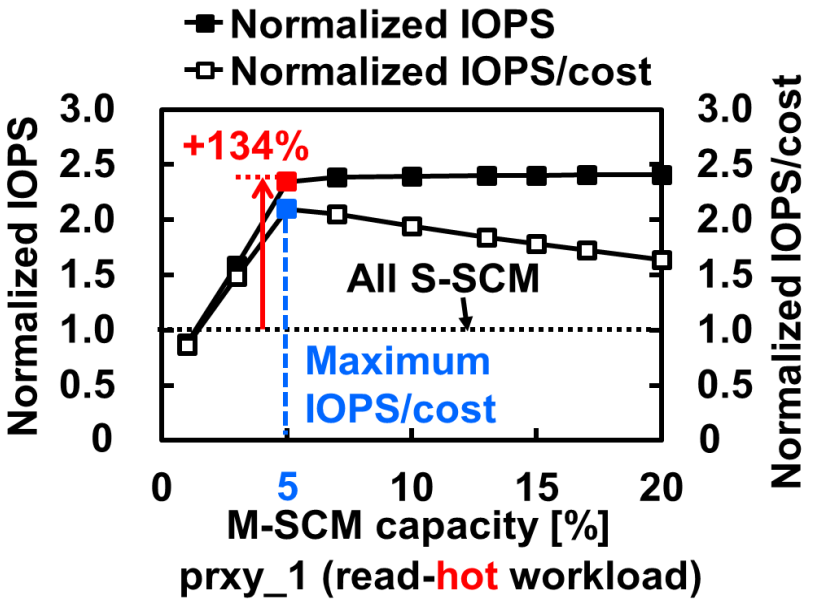
[5] S. Okamoto *et al.*, *IMW*, 2015, pp. 157-160.
[7] MSR Cambridge, http://iota.snia/org/traces/388.

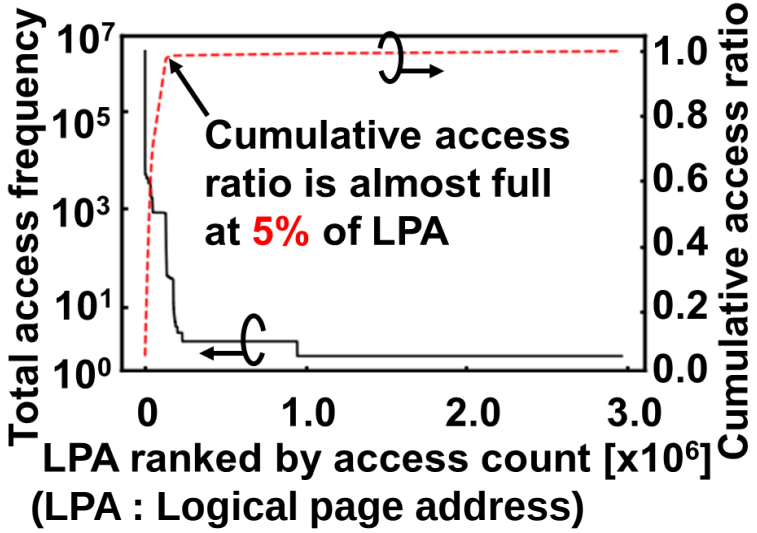
Outline

1. Introduction
2. Data Management Algorithm and Evaluation Setup
3. **Optimal Composition of Multi-SCM SSD**
4. Wear-leveling Policy for Multi-SCM SSD
    1. Performance effect of wear-leveling
    2. Optimal wear-leveling applying cases
5. Conclusion

Performance and Cost Results for Hot Workload (prxy_1)

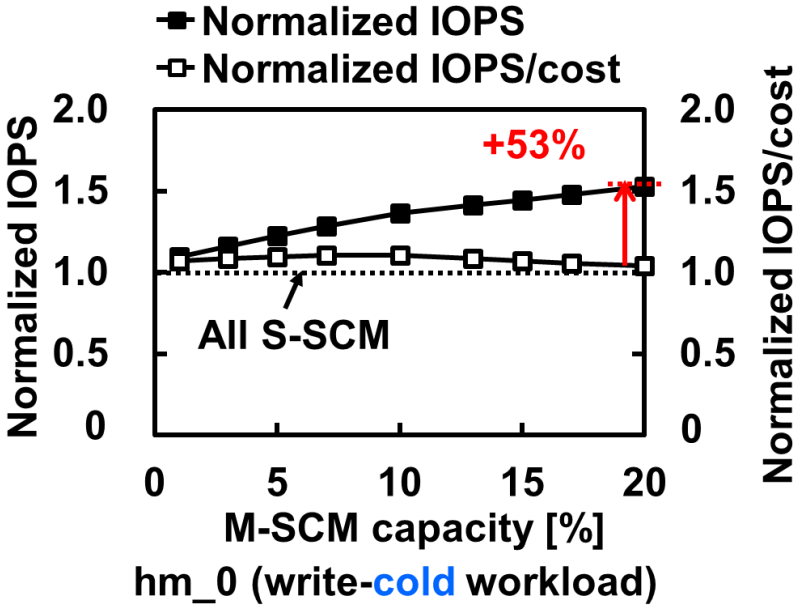
- Accesses from host concentrate on small part of all LPA
  → Small capacity of M-SCM improves SSD performance

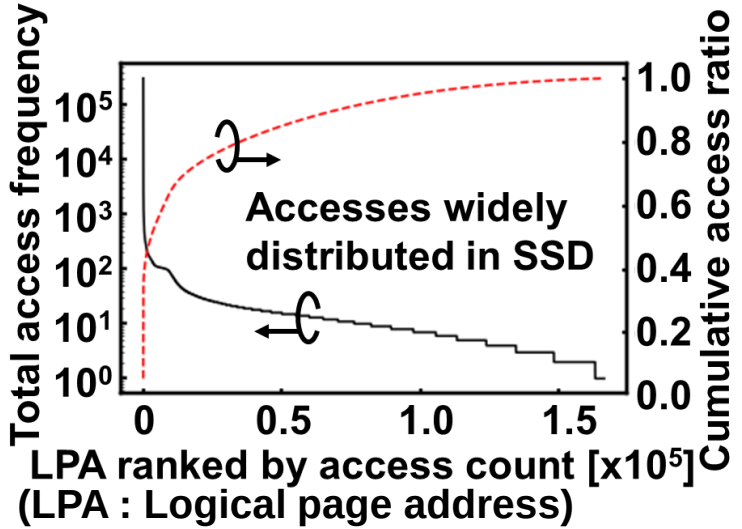




Performance and Cost Results for Cold Workload (hm_0)

- All S-SCM SSD is suitable because access does not concenter on particular addresses for cold workloads



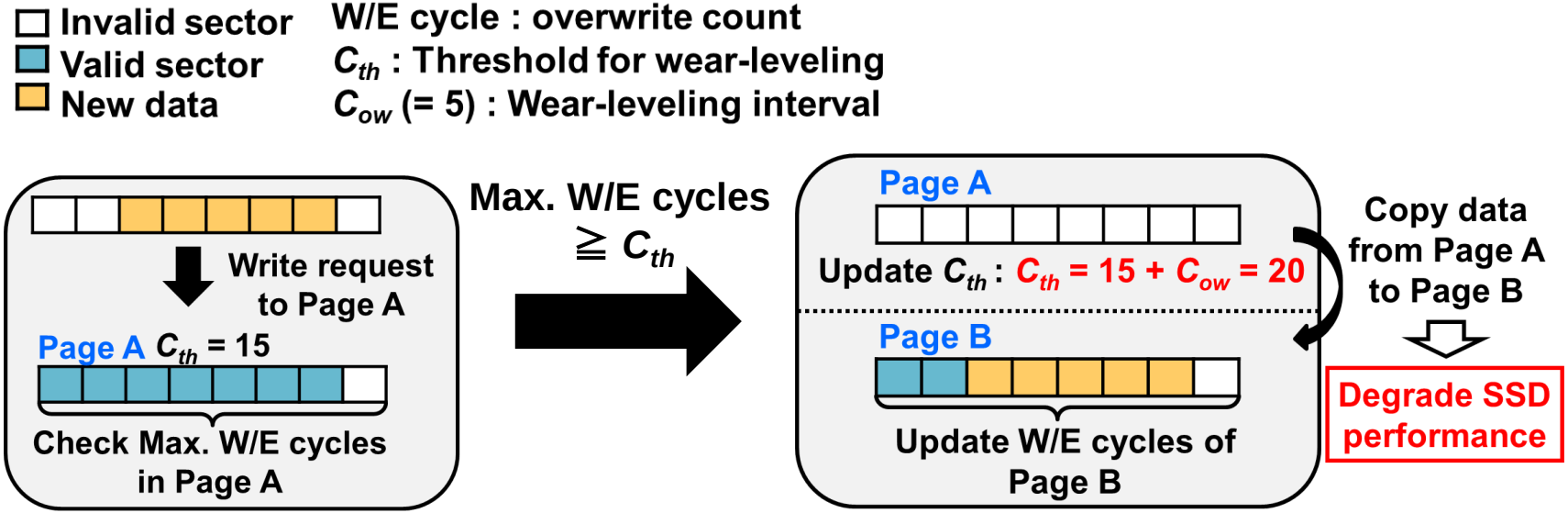


Outline

1. Introduction
2. Data Management Algorithm and Evaluation Setup
3. Optimal Composition of Multi-SCM SSD
4. **Wear-leveling Policy for Multi-SCM SSD**
    1. **Performance effect of wear-leveling**
    2. **Optimal wear-leveling applying cases**
5. **Conclusion**

Problem of Wear-leveling

- Data copy by wear-leveling decreases SSD performance



[3] T. Onagi *et al.*, *JJAP*, vol. 54, no. 4S, pp. 04DE04-1-04DE04-7, 2015.

Wear-leveling Cases

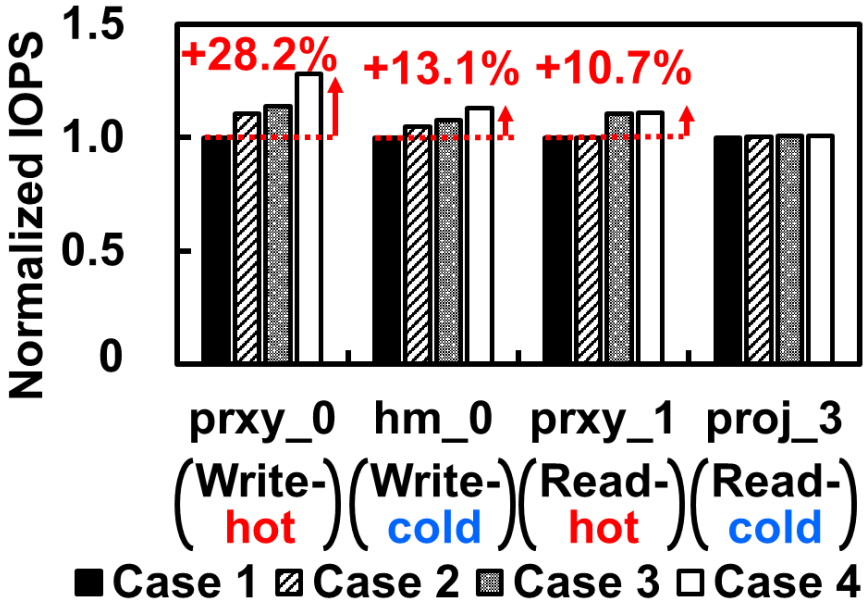
- Wear-leveling is omitted from M-SCM or S-SCM

Case	M-SCM	S-SCM
1	w/ wear-leveling	w/ wear-leveling
2	w/ wear-leveling	w/o wear-leveling
3	w/o wear-leveling	w/ wear-leveling
4	w/o wear-leveling	w/o wear-leveling

	M-SCM	S-SCM
Write/Read latency	100 ns / 100 ns	1 us / 1 us
Capacity	5% of SSD capacity	SSD capacity – M-SCM capacity
Endurance	10^9	10^5
Default C_{ow}	5	

Performance Results

- Case 4 achieves best SSD performance
- However, Case 4 is not realistic for SSD endurance lifetime



<Wear-leveling Cases>

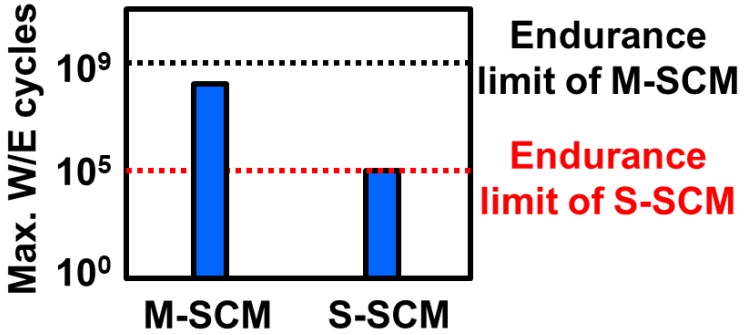
Case	M-SCM	S-SCM
1	w/ wear-leveling	w/ wear-leveling
2	w/ wear-leveling	w/o wear-leveling
3	w/o wear-leveling	w/ wear-leveling
4	w/o wear-leveling	w/o wear-leveling

Outline

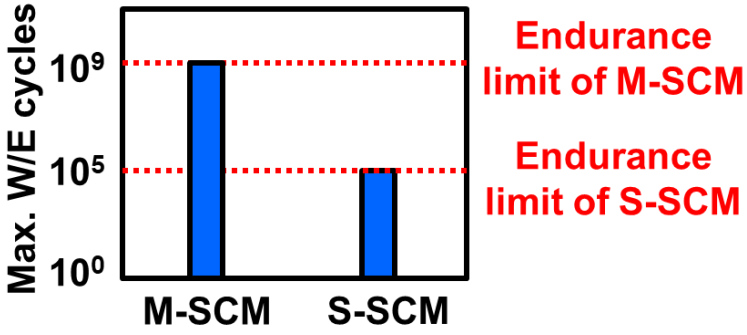
1. Introduction
2. Data Management Algorithm and Evaluation Setup
3. Optimal Composition of Multi-SCM SSD
4. **Wear-leveling Policy for Multi-SCM SSD**
    1. **Performance effect of wear-leveling**
    2. **Optimal wear-leveling applying cases**
5. Conclusion

Difference of Endurance Limits between M-SCM and S-SCM

- When Max. write/erase cycles of M-SCM or S-SCM reach endurance limit, SSD cannot be used
  - M-SCM endurance limit : **10^9** cycles
  - S-SCM endurance limit : **10^5** cycles



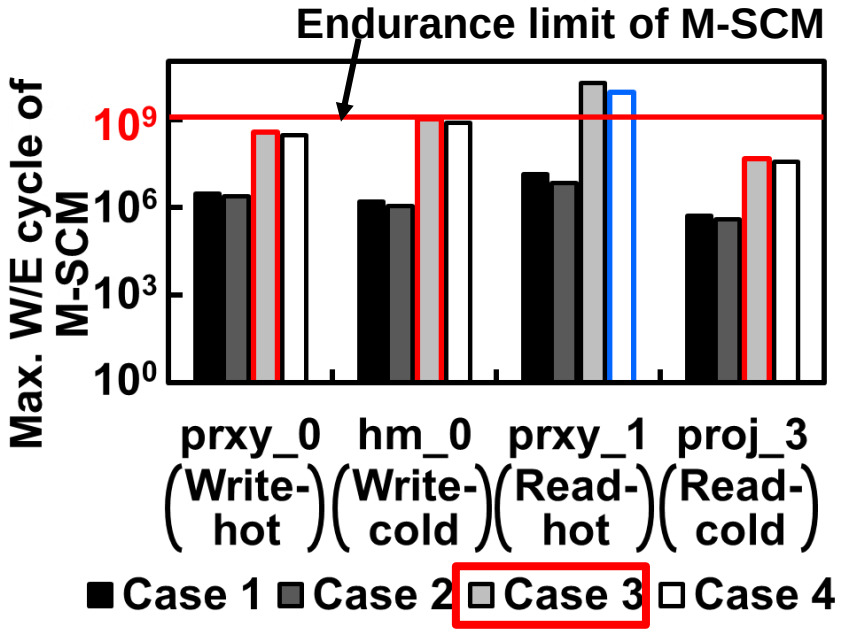
Still usable for M-SCM but not for Multi-SCM SSD



Ideal condition: Both SCMs reach their limits simultaneously

Endurance Limits for M-SCM

- Maximum W/E cycles of M-SCM when Max. W/E cycles of S-SCM exceed endurance limit and **S-SCM no longer works**



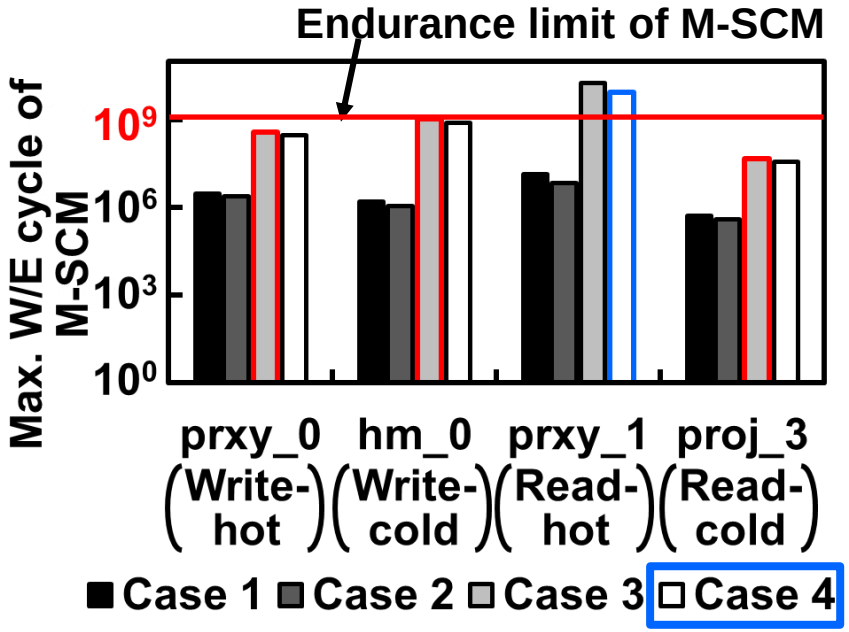
<Wear-leveling Cases>

Case	M-SCM	S-SCM
1	w/ wear-leveling	w/ wear-leveling
2	w/ wear-leveling	w/o wear-leveling
3	w/o wear-leveling	w/ wear-leveling
4	w/o wear-leveling	w/o wear-leveling

Case of closest to M-SCM limit and best for most workloads

Endurance Limits for M-SCM

- Maximum W/E cycles of M-SCM when Max. W/E cycles of S-SCM exceed endurance limit and **S-SCM no longer works**



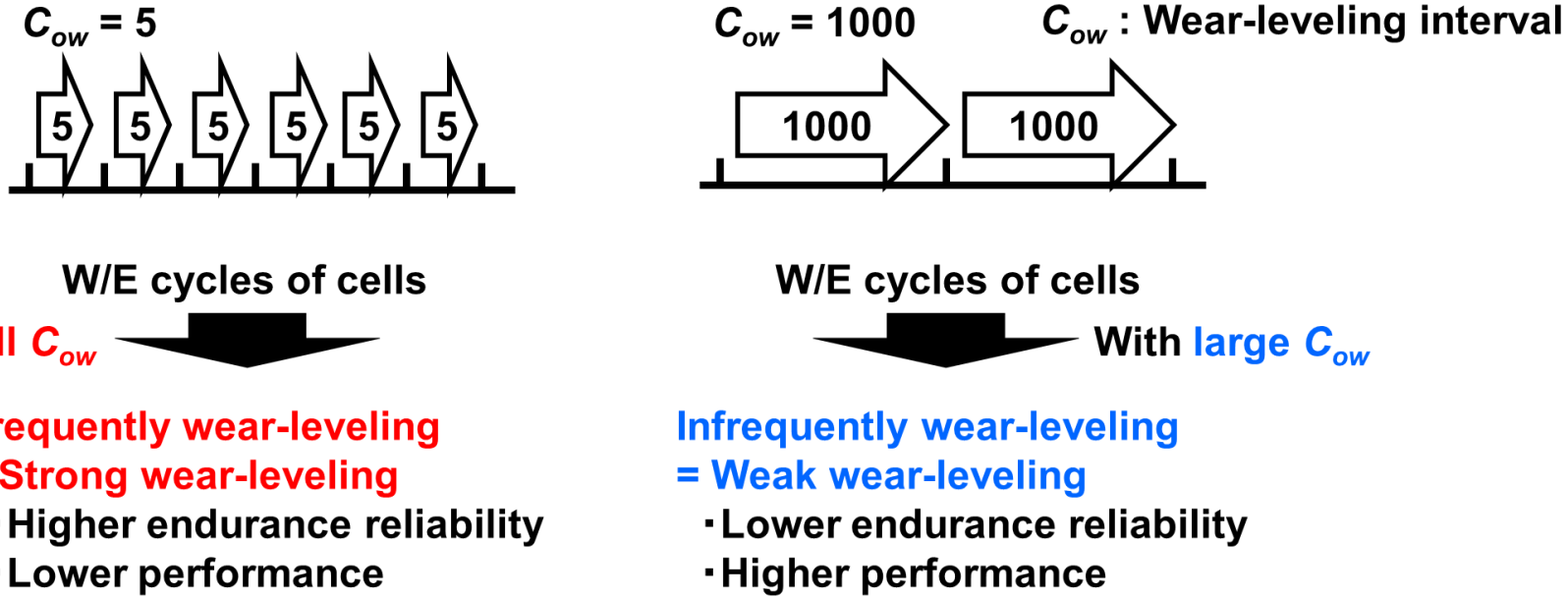
<Wear-leveling Cases>

Case	M-SCM	S-SCM
1	w/ wear-leveling	w/ wear-leveling
2	w/ wear-leveling	w/o wear-leveling
3	w/o wear-leveling	w/ wear-leveling
4	w/o wear-leveling	w/o wear-leveling

Best Case for prxy_1
but over M-SCM endurance limit
➡ **Weak wear-leveling**

Weak Wear-leveling

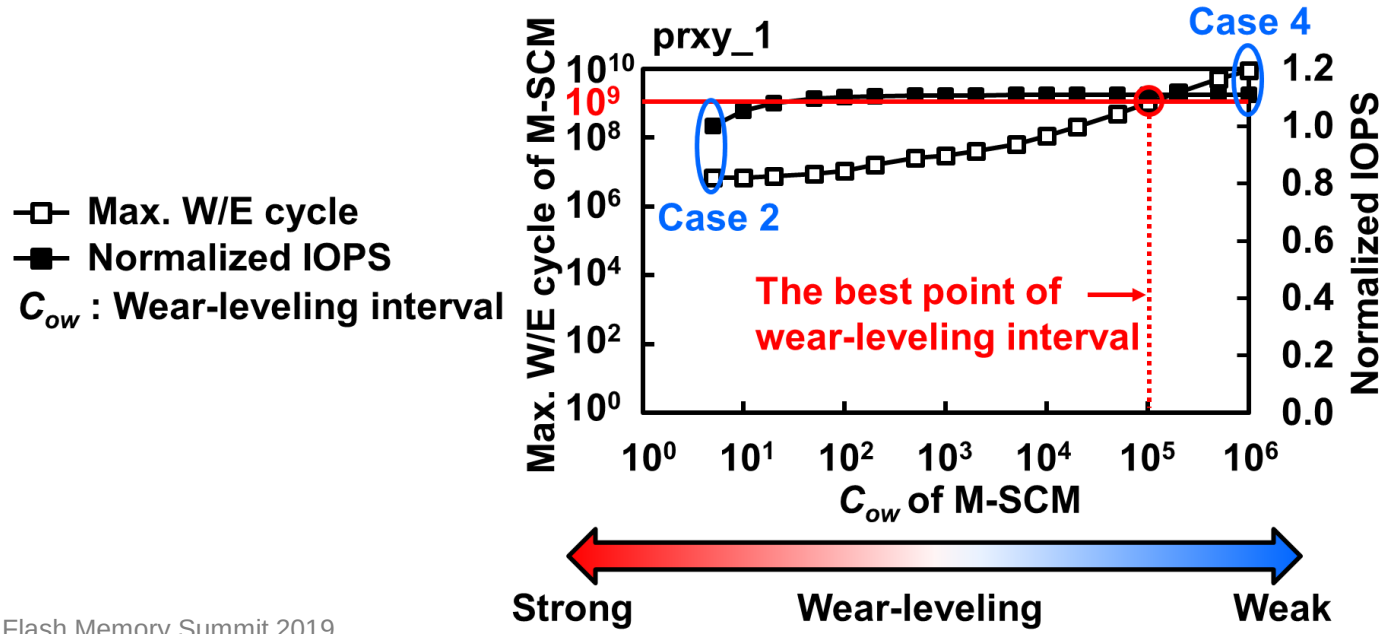
- To enhance system performance, wear-leveling frequency is decreased (Default C_{ow} is set to 5)



[2] T. Onagi *et al.*, *JJAP*, vol. 54, no. 4S, pp. 04DE04-1-04DE04-7, 2015.

Weak Wear-leveling for prxy_1

- Infrequently wear-leveling with large interval is applied for M-SCM



Outline

1. Introduction
2. Data Management Algorithm and Evaluation Setup
3. Optimal Composition of Multi-SCM SSD
4. Wear-leveling Policy for Multi-SCM SSD
    1. Performance effect of wear-leveling
    2. Optimal wear-leveling applying cases
5. **Conclusion**

Conclusion

- Small capacity of M-SCM enhances SSD performance efficiently for hot workloads
- Wear-leveling is not needed to high endurance M-SCM for most workloads

	prxy_0 (Write-hot)	hm_0 (Write-cold)	prxy_1 (Read-hot)	proj_3 (Read-cold)
Optimal M-SCM capacity (IOPS improvement from all S-SCM SSD)	8% (+104%)	10% (+36%)	5% (+134%)	M-SCM is not needed
Optimal wear-leveling (IOPS improvement)	Only S-SCM (+14%)	Only S-SCM (+7.6%)	Only M-SCM with weak wear-leveling (+10.7%)	Only S-SCM (+0.6%)

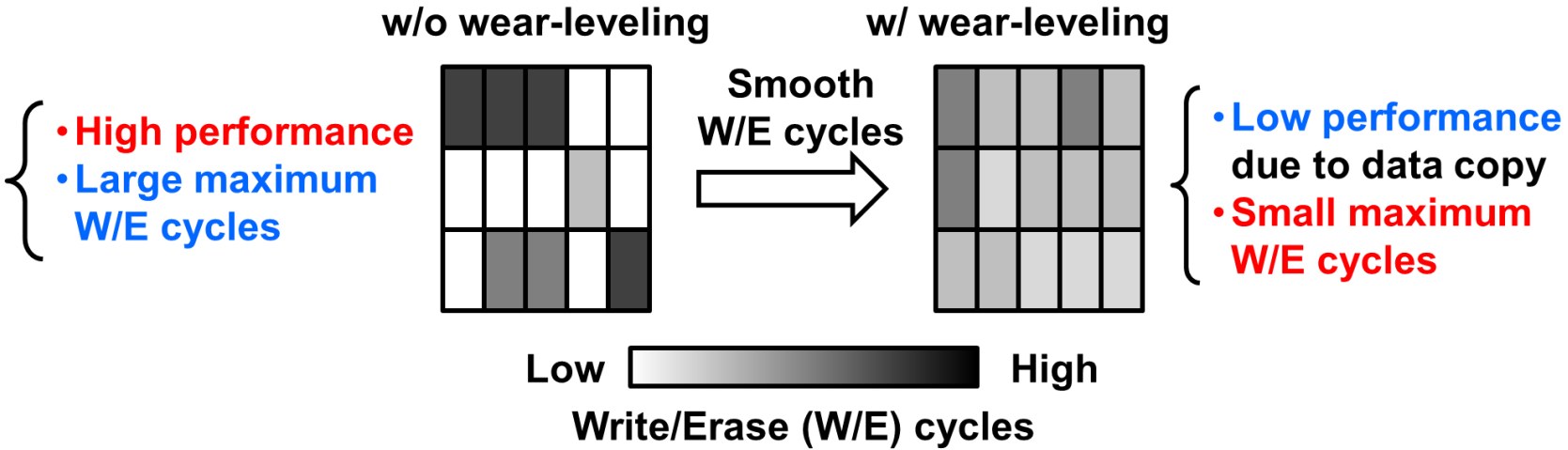
Thank you for kind attention

This work is based on results from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO)


Flash Memory Summit

Wear-leveling

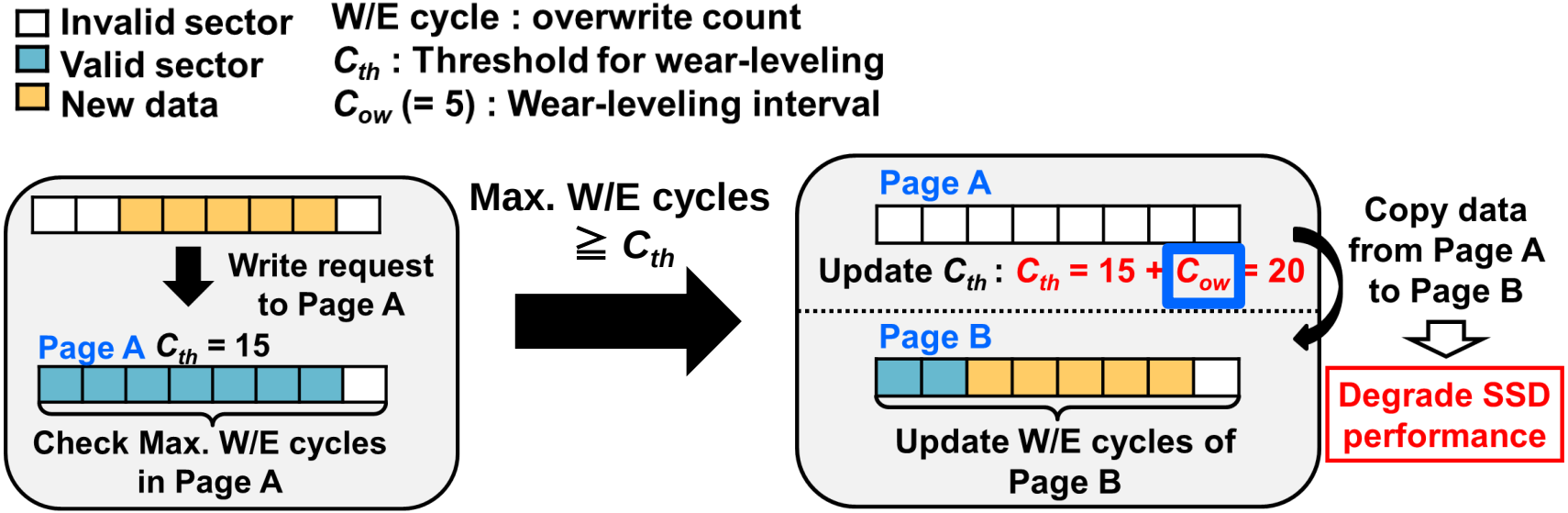
- To extend endurance lifetime of SSD, wear-leveling technique is applied to SCMs
- Data copy by wear-leveling decreases SSD performance



[2] T. Onagi *et al.*, *JJAP*, vol. 54, no. 4S, pp. 04DE04-1-04DE04-7, 2015.

Problem of Wear-leveling

- Data movement by wear-leveling decreases SSD performance



[2] T. Onagi *et al.*, *JJAP*, vol. 54, no. 4S, pp. 04DE04-1-04DE04-7, 2015.