NVRAMOS 2016

Over-Provisioning and Performance of Mobile Storage

SKhynix

2016. 10. 21

Contents

SK hynix

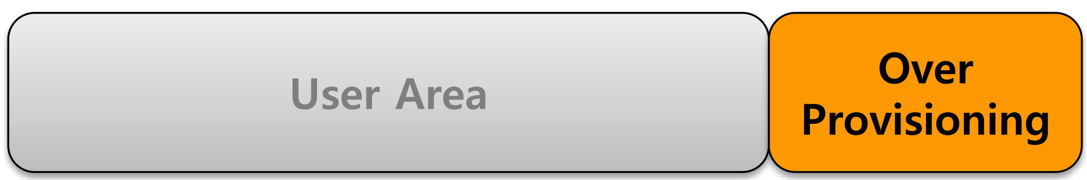
Introduction to Over-Provisioning

- **It is a special region inside flash memory storages**
  - Required due to the inability to overwrite flash before erasure
  - Required due to operate the user area

- **Over-Provisioning Rate (%)**

$$\left(\frac{Physical\ Capacity\ - User\ Capacity}{User\ Capacity}\right) * 100 = \text{Over Provisioning } Rate\ (\%)$$

  - 128GiB physical flash capacity vs. 128GB user capacity
    - About 7% of physical capacity is used for over-provisioning



Introduction to Over-Provisioning

- **A portion of storage capacity held in reserve**
  - Garbage collection (the major use)
  - Controller firmware metadata (small %)
  - Reserved for bad block (small %)
  - Additional features such as data protection

Over-Provisioning in SSD

- **Level 1 : 7.37%**
  - The capacity of *HDDs* and *SSDs* is often measured in GB (power of 10)
  - The capacity of *memory* is represented in Gibibyte (power of 2)
  - Internal over-provisioning given by unit difference

    Ex) GiB (2^30) – GB(10^9)

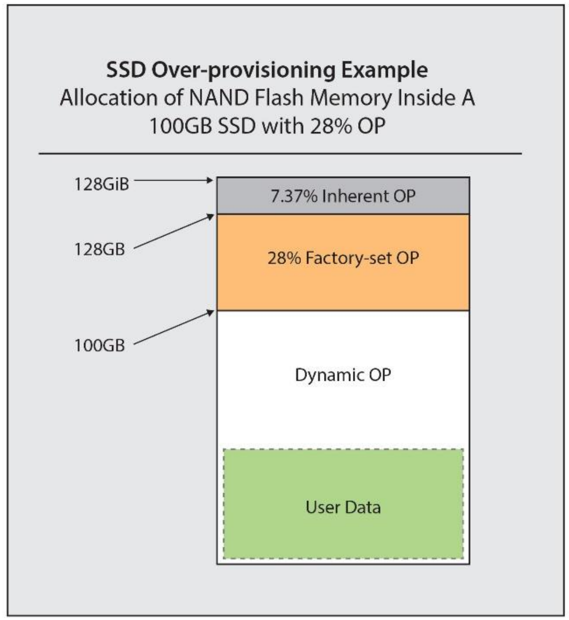
- **Level 2 : 0, 7, 28%**
  - SSDs are factory-set with a 2nd OP level

    to improve performance and life time

    Ex) 128GB, 120GB, 100GB

- **Level 3 : Dynamic OP**
  - The user space that is not yet occupied by user data

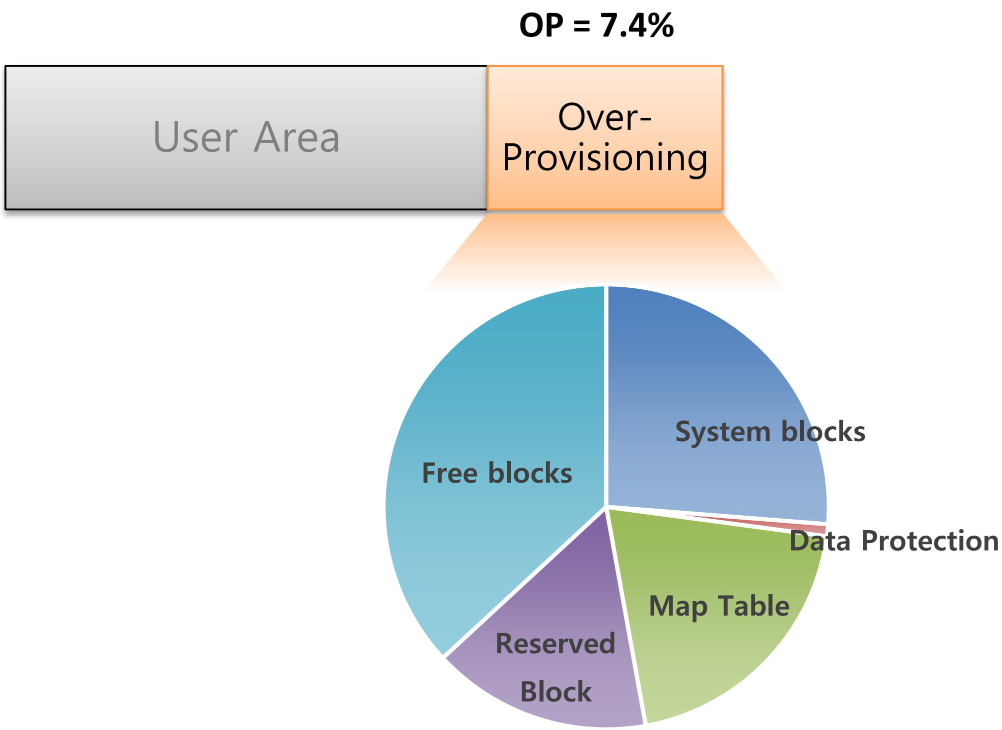
    can be automatically used as over-provisioning area

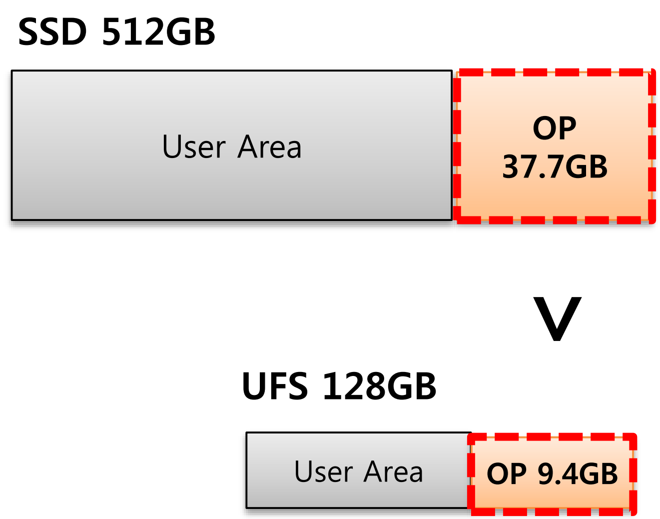


https://en.wikipedia.org/wiki/Write_amplification#Over-Provisioning

OP Restrictions in Mobiles: ① Smaller Over-Provisioning

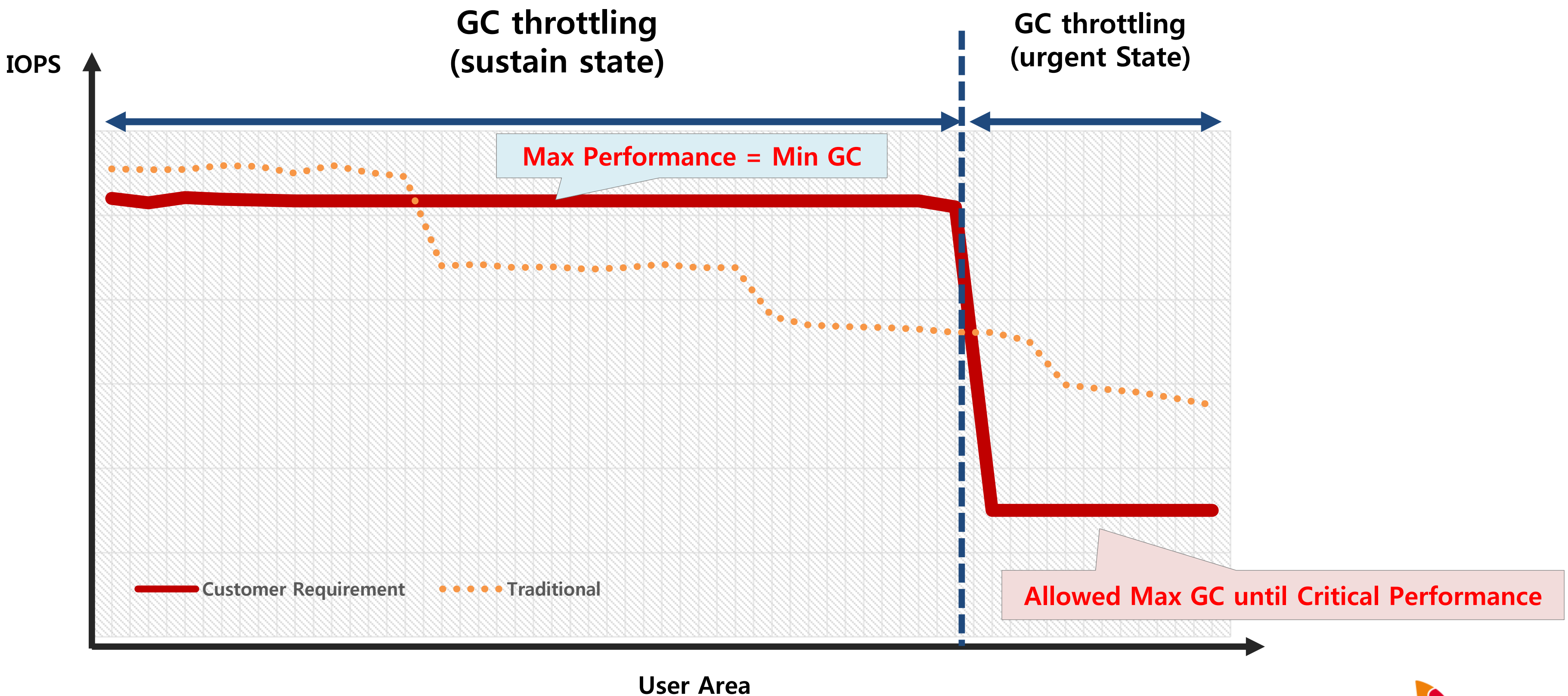
- Level 2 and 3 OP are not allowed in mobiles
- If the size of system area is almost the same,
  the remaining space (free blocks) for GC is less in mobiles than in SSDs





OP Restrictions in Mobiles: ② Performance Requirement

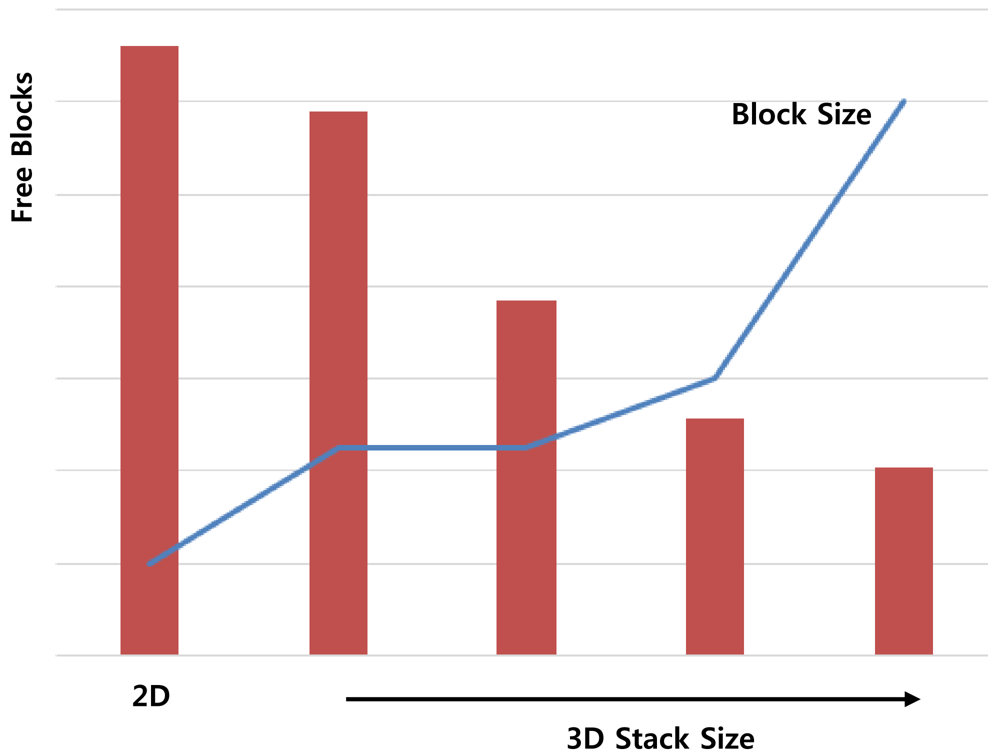
- Customer requirement
  - Mobile storage should provide a required fraction of the sustain (steady) state performance even when the user space runs out
- The performance is directly effected by over-provisioning size



SK hynix

OP Restrictions in Mobiles: ③ Block Size Increase

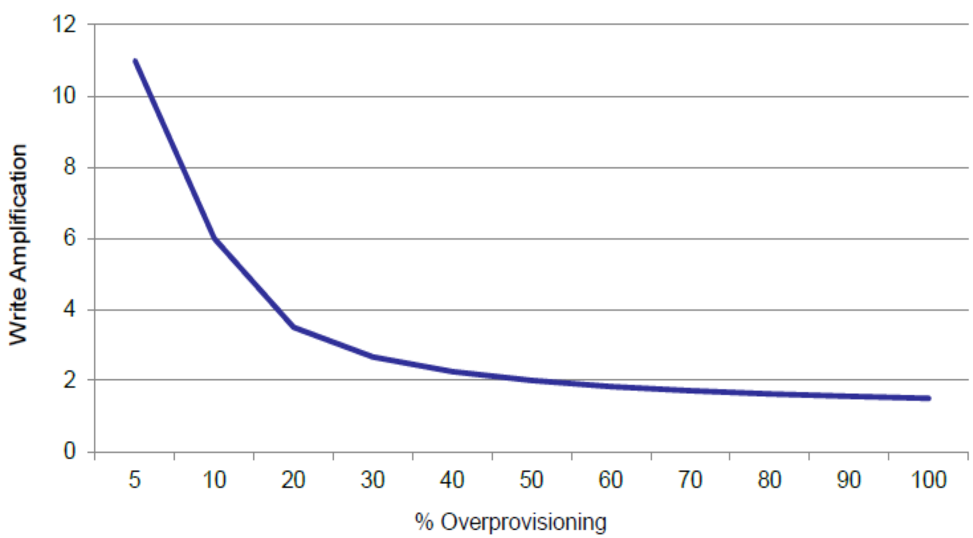
- The size of a block increases as the stack of 3D NANDs becomes higher
- Larger blocks may not be efficient when storages use superblock-based mapping
- Assuming the number of system blocks are fixed, less space will be left free as the block size increases

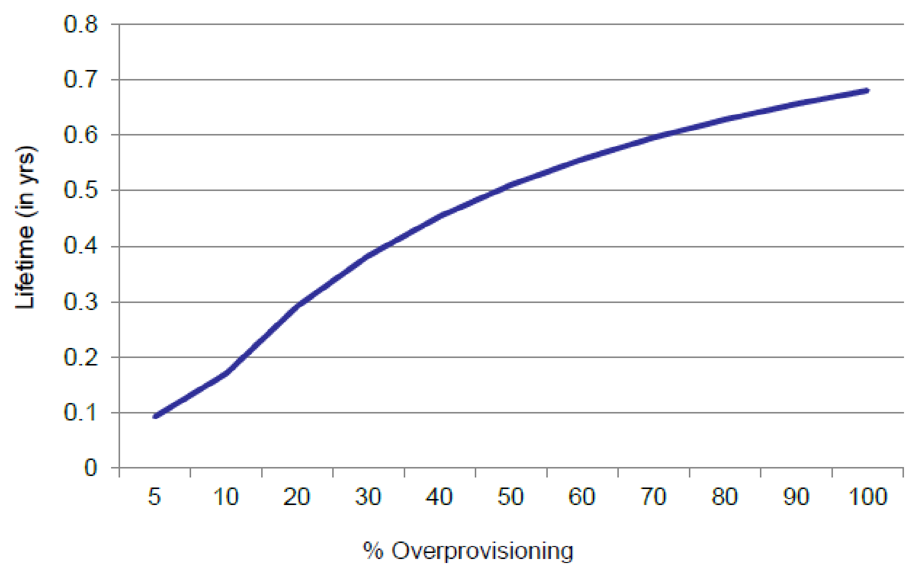


* If the density of all product is same,

Relation between OP and Performance

- **Larger OP space helps to improve performance**
  - Higher write performance and lower write amplification
  - Longer life time (endurance)

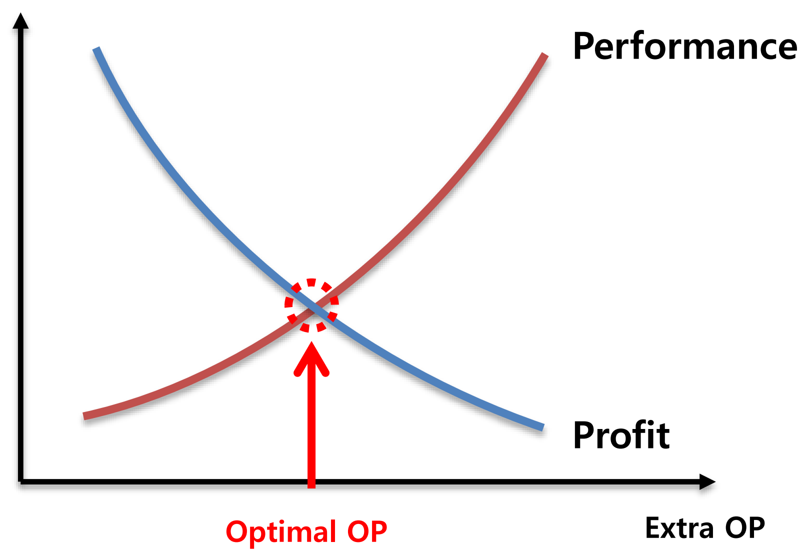




Reference : "Overprovisioning in All-Flash Arrays", Bill Radke, Skyera, Inc, 2013, Flash Summit

Optimal Over-Provisioning in Mobile

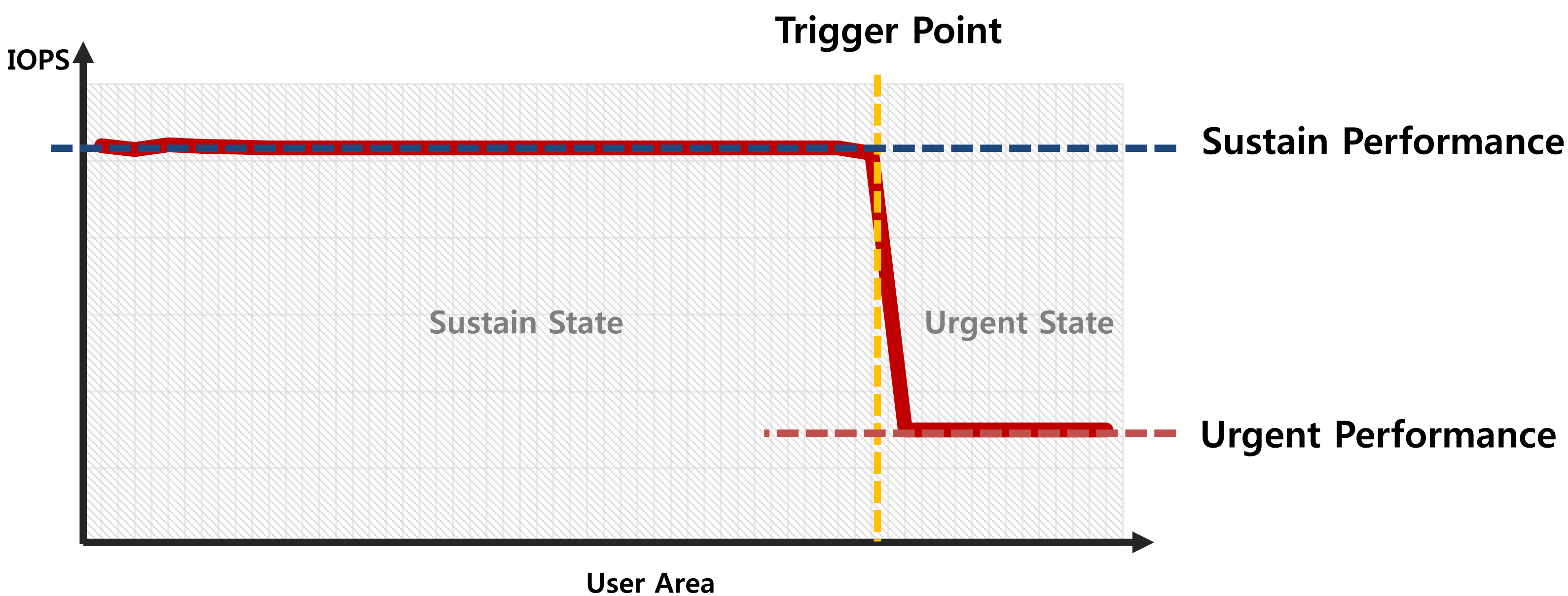
- **In summary, we have following limitations in mobile storages**
  - Used only internal over-provisioning (the level 1 of SSD)
  - Performance behavior specified by customers
  - Decreased internal over-provisioning as the stack of 3D NAND grows
- **Additional OP besides internal OP could be provided with extra cost**
  - How to find the optimal amount of additional OP space while satisfying the performance requirement given by the customer?



SK hynix

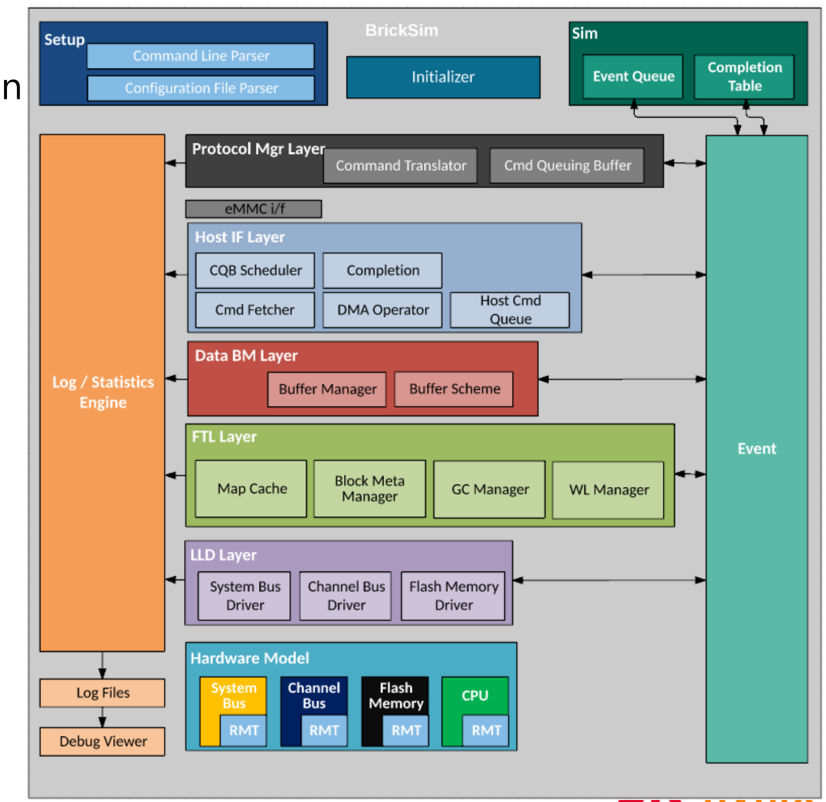
Evaluations : Case Studies

- **Test Cases**
  - Case #1 : Performance variation depending on OP amount in urgent state
  - Case #2 : Analysis of relationship between performance and OP
  - Case #3 : Exploration for optimal performance



Test Environment

- **Simulator : BrickSim**
  - Develop by SKhynix and an university lab as industry-academy collaboration
  - Objective
    - Integrated simulation of HW/SW sub-systems
    - Architecture and algorithm exploration
    - Easy extension of SW algorithms
    - Event-driven simulation framework for fast simulation
  - Key Features
    - Scheduler backbone based on event queue
    - Modular and layered components
    - HW components and SW algorithms to be easily modifiable
    - SW algorithm can be changed using wrapper function



Test Environment (with Default Settings)

- **HW Configuration**
  - 64GB NAND Flash
- **Algorithms**
  - The urgent state is triggered when free blocks are about N% of total block (N = 3 by default)
  - The performance of urgent state is K% of sustain state (K = 30 by default)
- **Workload**
  - TIO (Thread IO) Trace
    - Full user area write : repeat the sequence (Sequential Write 1G -> Random Write 200M)
- **Block Usage**

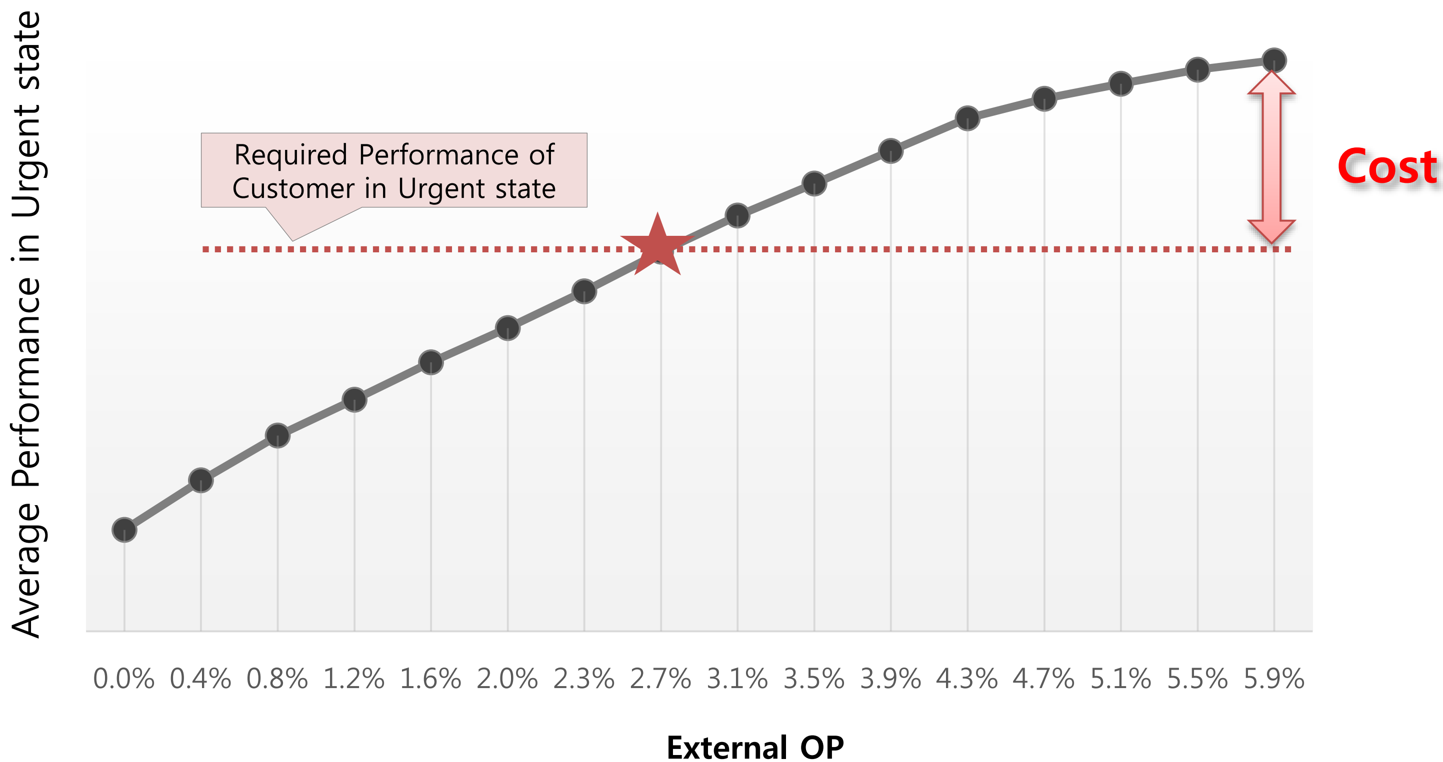
Main Block : Total 1024 Blocks (100%)		
User Blocks		93%
Internal Over-Provisioning	Free Blocks	2%
	System Blocks	5%

Additional block : Variable Blocks	
External Over-Provisioning	0~6%

Case #1 : Performance Variation depending on OP amount

- **Performance variation from entering the urgent state to 100% occupation of user area**
- **The performance increases as the OP rate grows**
- **But, once the performance requirement from customer is satisfied, no more OPs are needed**

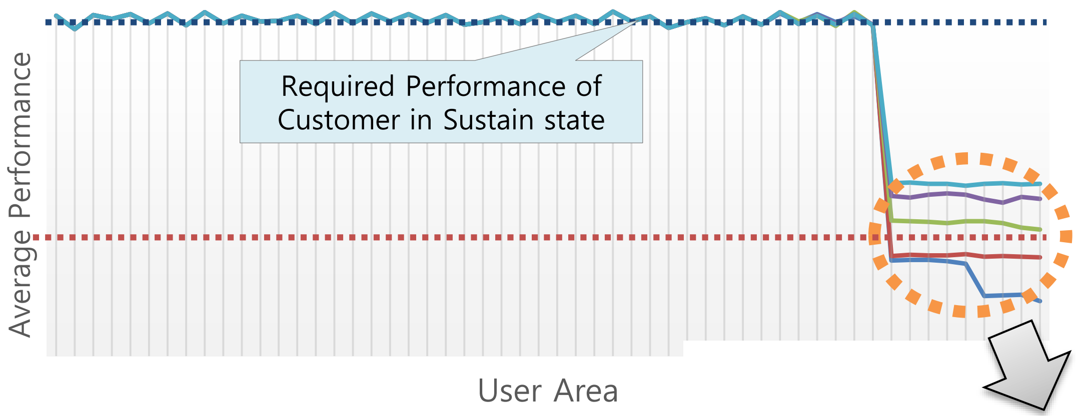
*** Random Write Test**

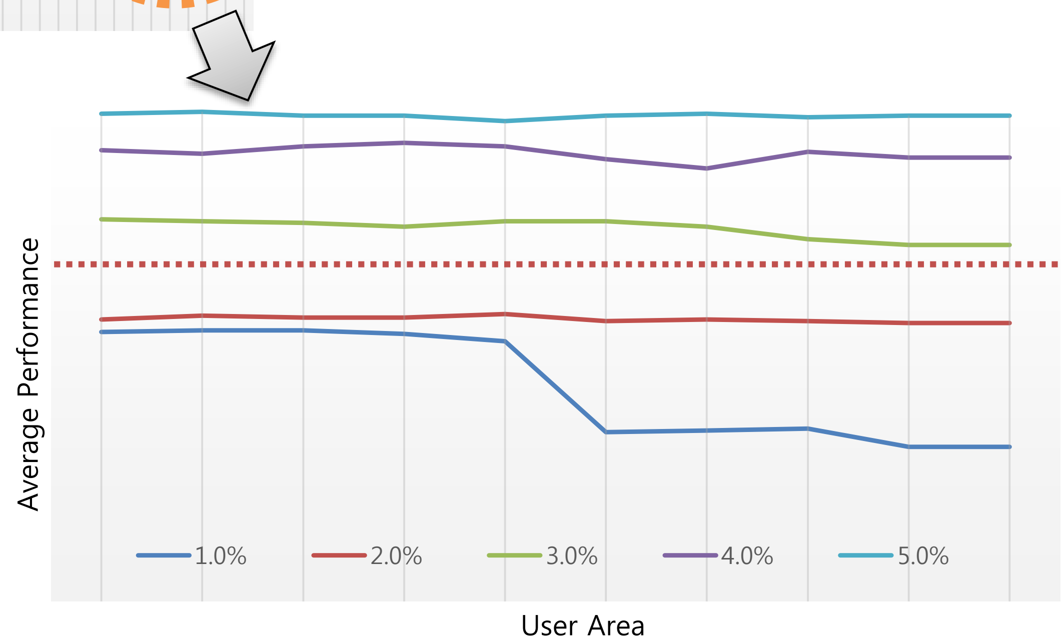


SK hynix

Case #2 : Analysis for relation of Performance and OP

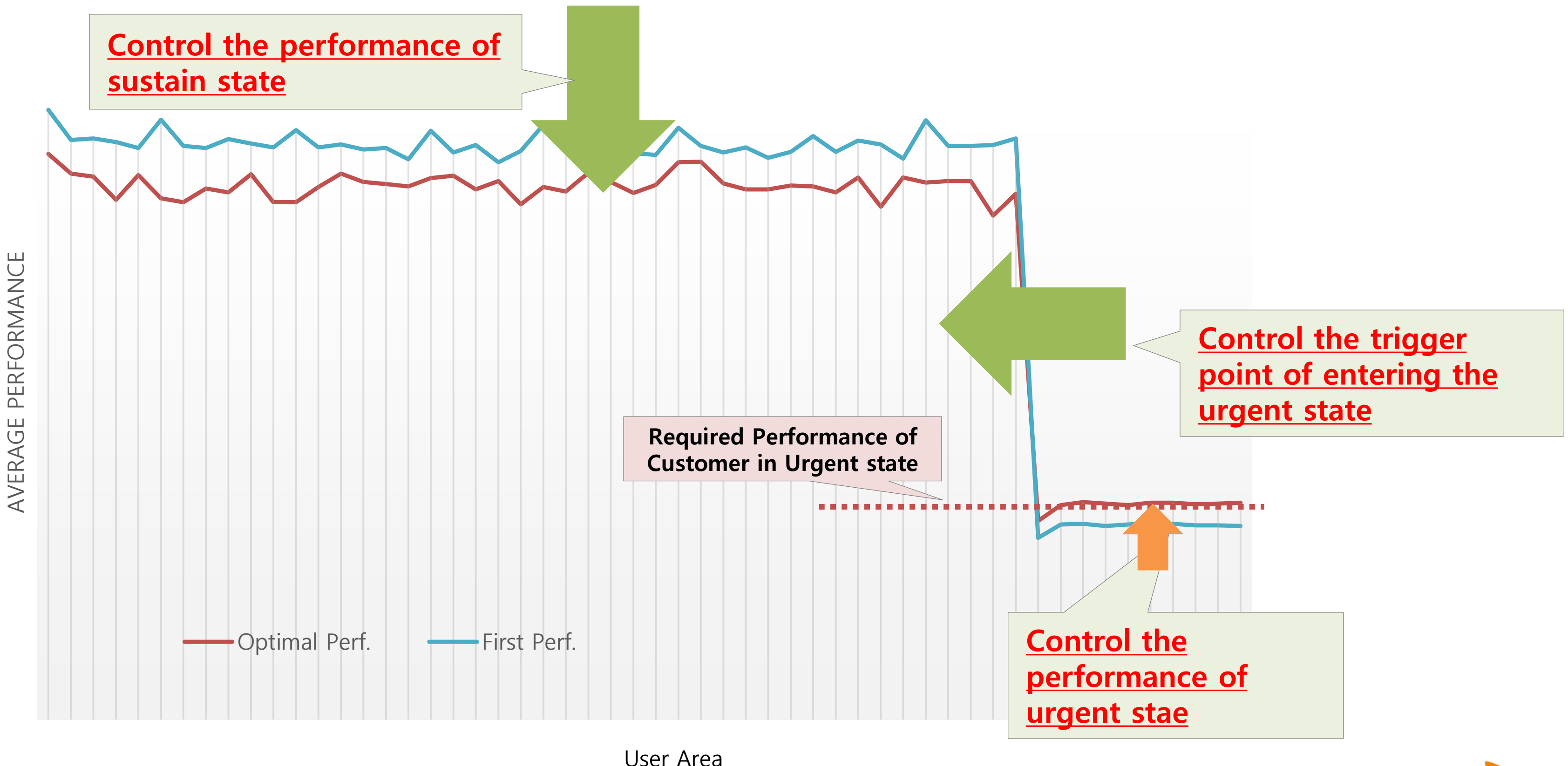
- Once the sustain-state performance meets the requirement, the OP amount is varied to see the performance in urgent state
- In the urgent state, there is 1.6% of performance gain observed as OP increases by 1%





Case #3 : Exploration for Optimal Performance in Sustain State

- If the performance in urgent state does not meet customer requirement, it could be improved
  - by controlling the performance in sustain state or
  - by early entering the urgent state



SK hynix

Summary

- **Smaller over-provisioning area in mobile than in SSD**
  - Either a factory-set OP or a dynamic OP is not used
  - GC operations should work efficiently in urgent state to meet performance requirement
  - Smaller internal OP than in SSD assuming the number of system blocks is almost the same
  - System data consumes more blocks in OP as the 3D stack grows
- **An optimal OP rate can be decided by considering the performance of sustain and urgent state**
  - Higher OP is better for high performance
  - In mobile, larder over-provisioning helps to increase performance but it will be saturated under customer performance requirement
  - When OP is fixed, the performance could be tailored by controlling the performance in sustain state or by entering the urgent state early

Thank you

References

1. "NAND Storage Simulator for Architecture Exploration and Performance Evaluation," SKhynix, 2016.
2. "Understanding SSD Over Provisioning", Kent Smith, LSI, 2012.
3. https://en.wikipedia.org/wiki/Write_amplification#Over-Provisioning, Wikipedia
4. SSD Over-Provisioning And Its Benefits, Seagate