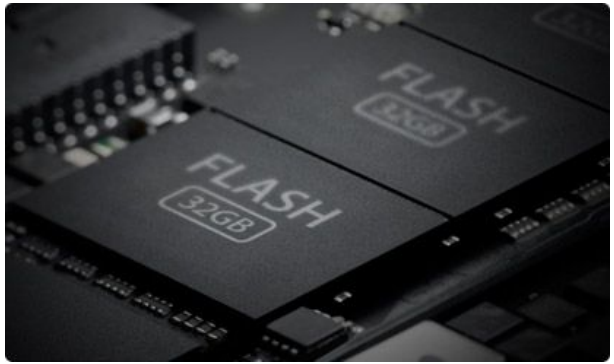
Understanding NVMe Zoned Namespace (ZNS) Devices

Nick Tehrany and Animesh Trivedi
Contact: N.Tehrany@student.tudelft.nl
Homepage: https://nicktehrany.github.io/
June 8-10
Compsys'22

Flash Storage

SSD



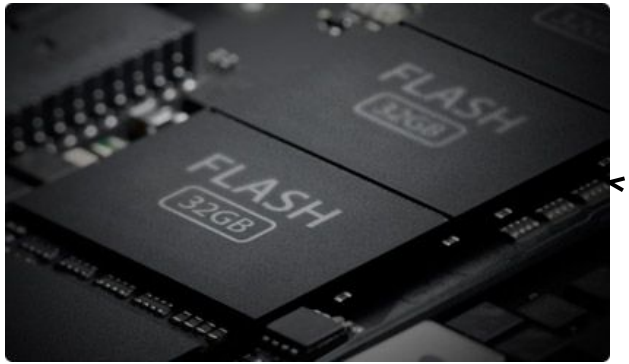
FAST

- Multiple GB/s Bandwidth
- Millions IOPS

1. Nitin Agrawal, Vijayan Prabhakaran, Ted Wobber, John D. Davis, Mark Manasse, and Rina Panigrahy. *Design tradeoffs for SSD performance.* 2008. *USENIX ATC'08.*

Flash Storage

SSD



FAST

- Multiple GB/s Bandwidth
- Millions IOPS

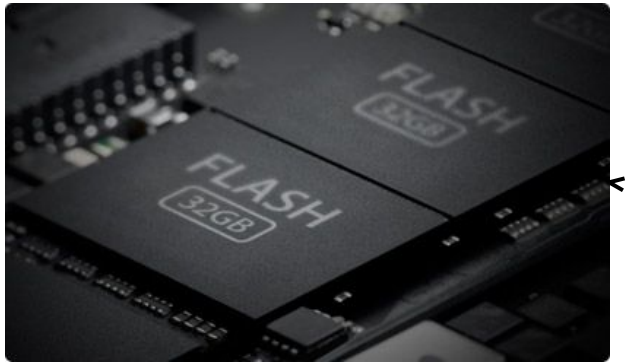
Flash Chip

- µ-second r/w latency
- **cannot** overwrite must erase

1. Nitin Agrawal, Vijayan Prabhakaran, Ted Wobber, John D. Davis, Mark Manasse, and Rina Panigrahy. *Design tradeoffs for SSD performance.* 2008. *USENIX ATC'08.*

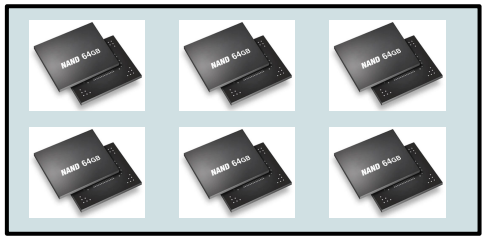
Flash Storage

SSD



FAST

- Multiple GB/s Bandwidth
- Millions IOPS

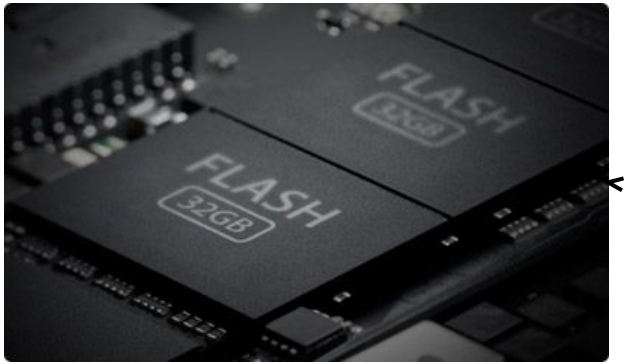


- Package flash chips together
- Erase multiple at once

1. Nitin Agrawal, Vijayan Prabhakaran, Ted Wobber, John D. Davis, Mark Manasse, and Rina Panigrahy. *Design tradeoffs for SSD performance.* 2008. *USENIX ATC'08.*

Flash Storage

SSD



FAST

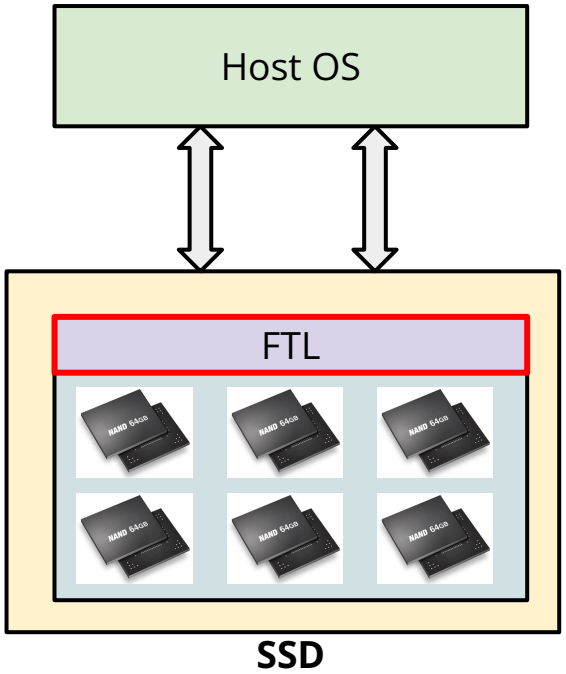
- Multiple GB/s Bandwidth
- Millions IOPS

- Package flash chips together
- Erase multiple at once

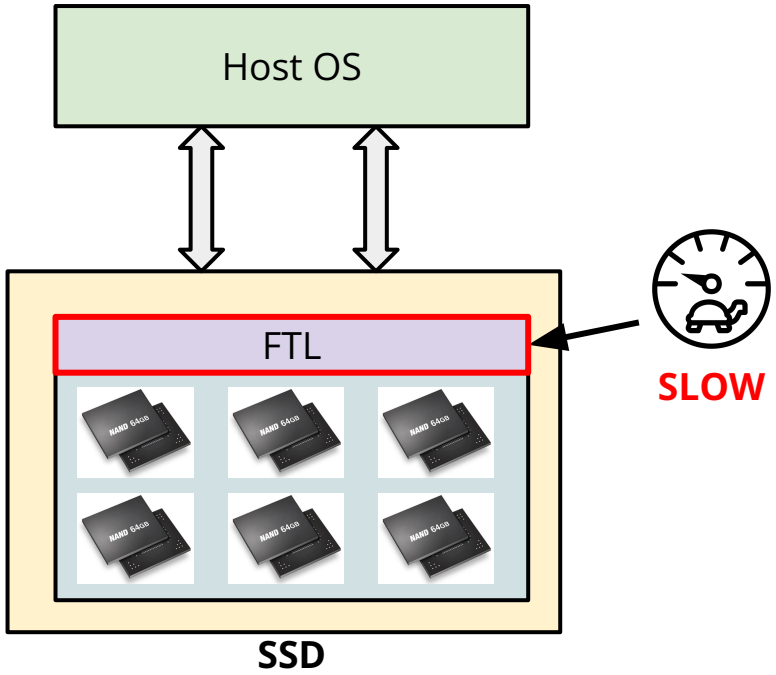
GC

1. Nitin Agrawal, Vijayan Prabhakaran, Ted Wobber, John D. Davis, Mark Manasse, and Rina Panigrahy. *Design tradeoffs for SSD performance.* 2008. *USENIX ATC'08.*

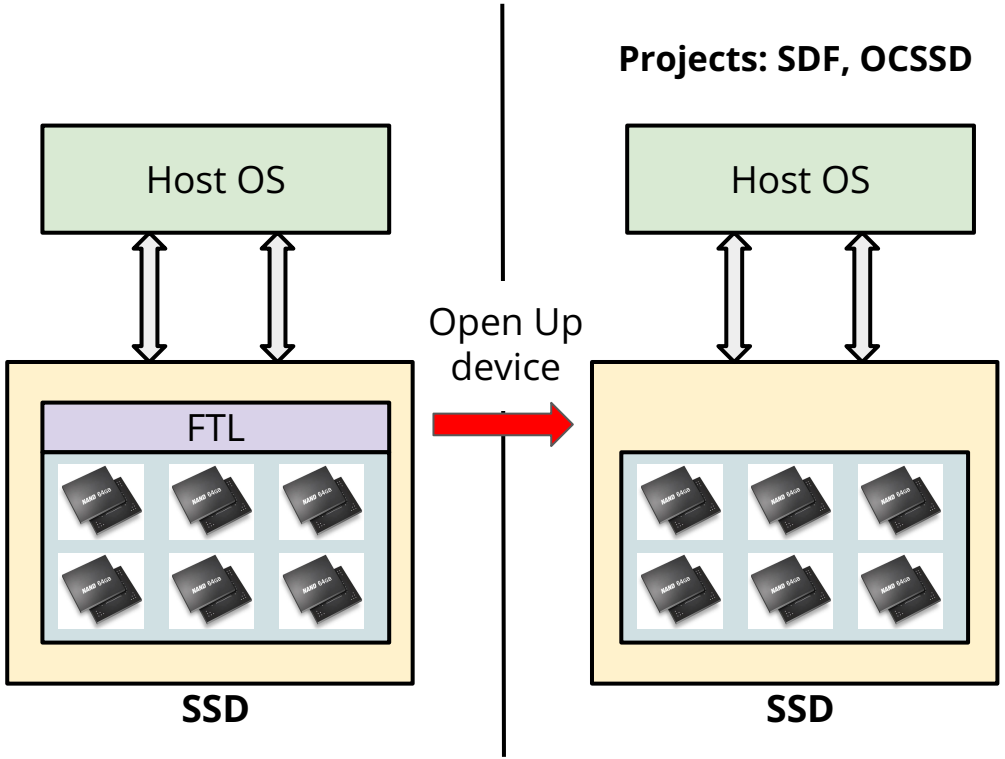
Flash Storage Integration



Flash Storage Integration

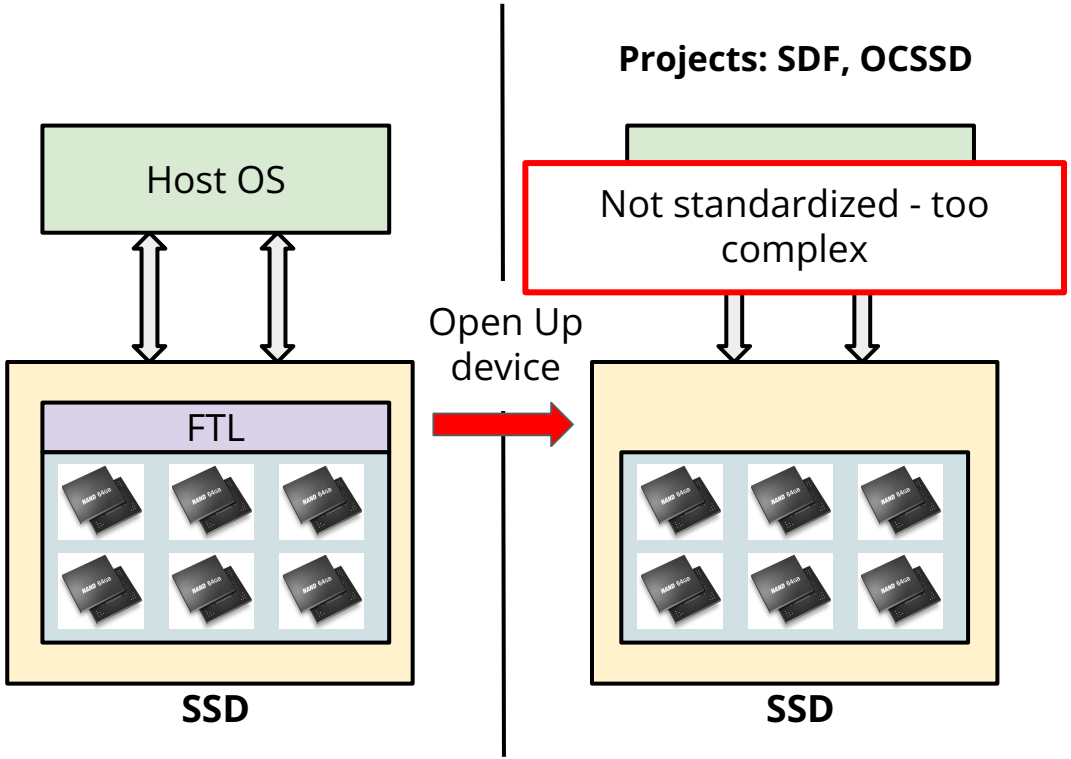


Flash Storage Integration



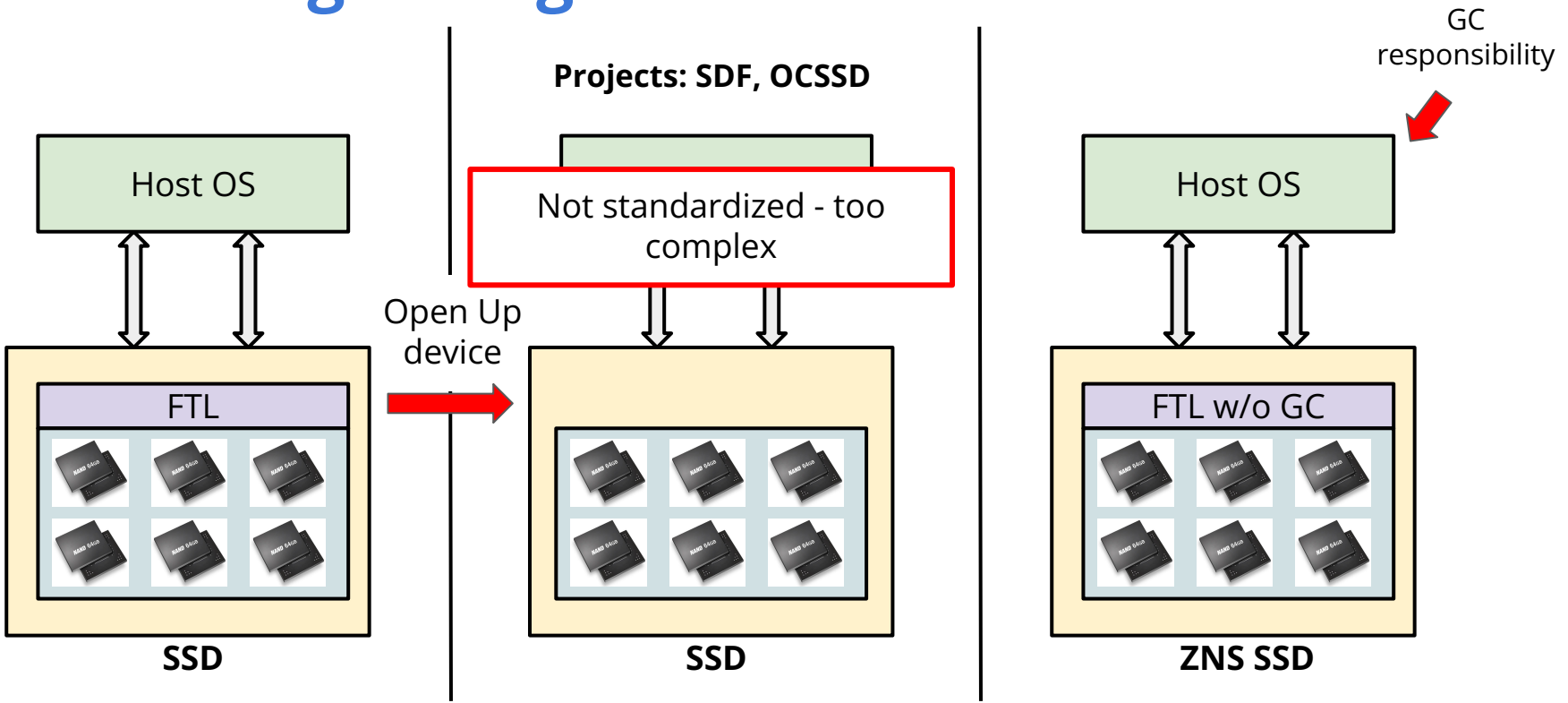
1. Ouyang, Jian, et al. "SDF: Software-defined flash for web-scale internet storage systems." Proceedings of the 19th international conference on Architectural support for programming languages and operating systems. 2014.
2. Picoli, Ivan Luiz, et al. "Open-Channel SSD (What is it Good For)." CIDR. 2020.

Flash Storage Integration



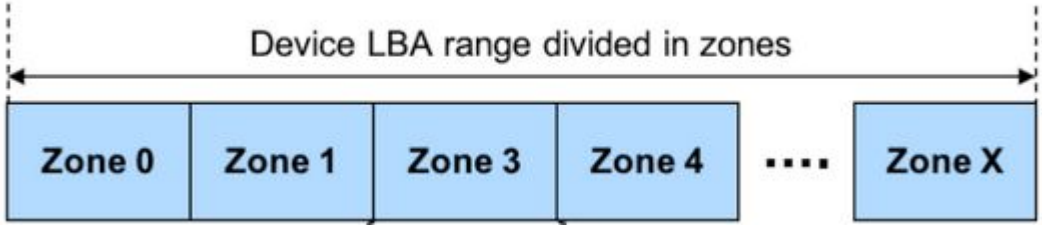
1. Ouyang, Jian, et al. "SDF: Software-defined flash for web-scale internet storage systems." Proceedings of the 19th international conference on Architectural support for programming languages and operating systems. 2014.
2. Picoli, Ivan Luiz, et al. "Open-Channel SSD (What is it Good For)." CIDR. 2020.

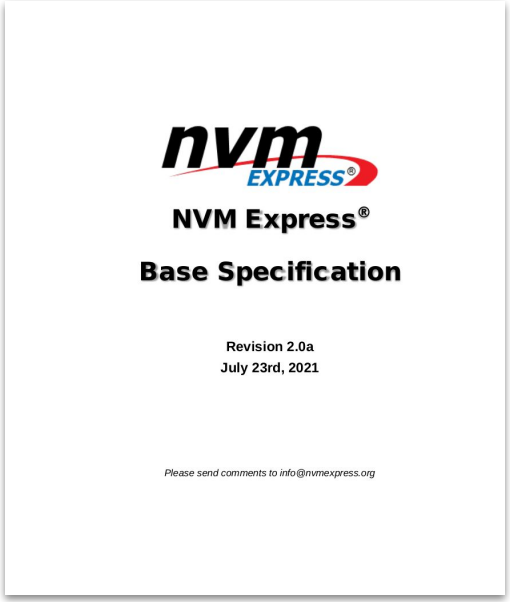
Flash Storage Integration



1. Ouyang, Jian, et al. "SDF: Software-defined flash for web-scale internet storage systems." Proceedings of the 19th international conference on Architectural support for programming languages and operating systems. 2014.
2. Picoli, Ivan Luiz, et al. "Open-Channel SSD (What is it Good For)." CIDR. 2020.

Zoned Namespace (ZNS) Devices

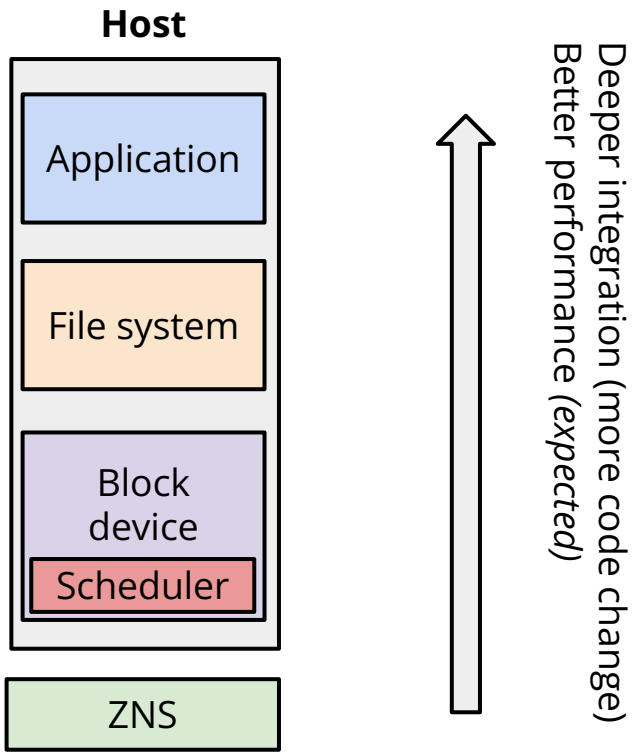




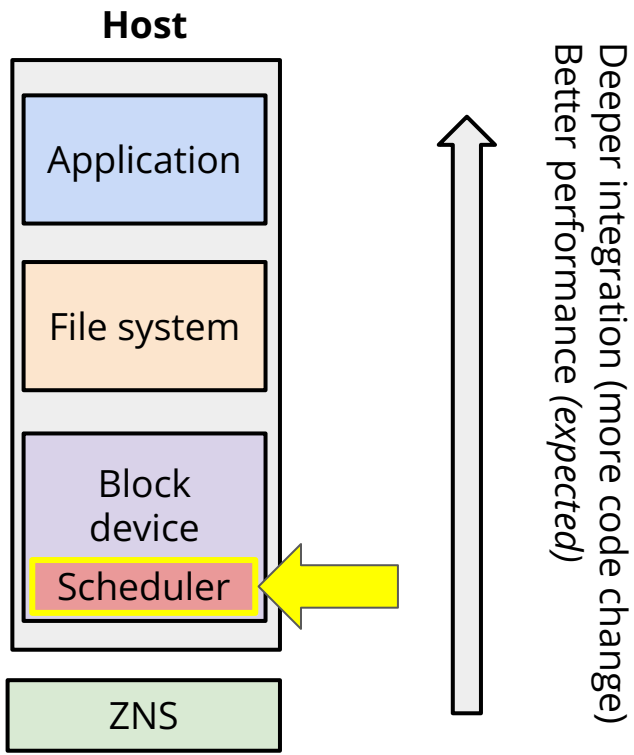
- Append-only sequential writes
- Zone erase before overwrite
- 1 outstanding write per zone

1. Western Digital Corporation. *NVMe Zoned Namespaces (ZNS) SSDs.* https://zonedstorage.io/docs/introduction/zns

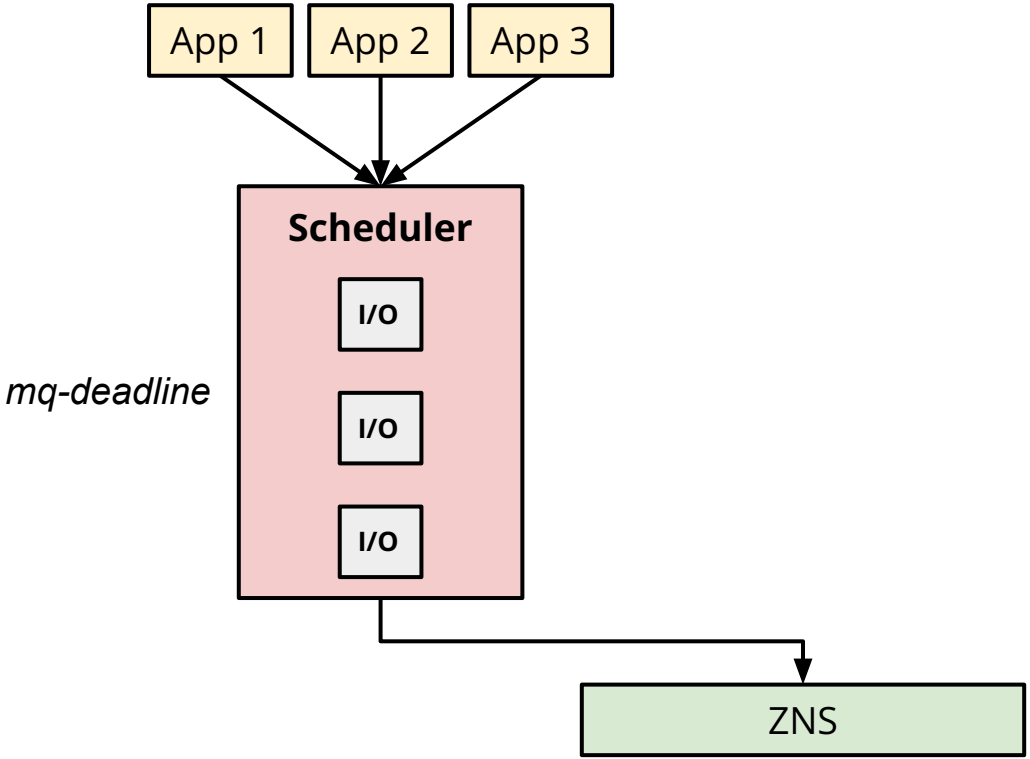
ZNS Integration - Tradeoff



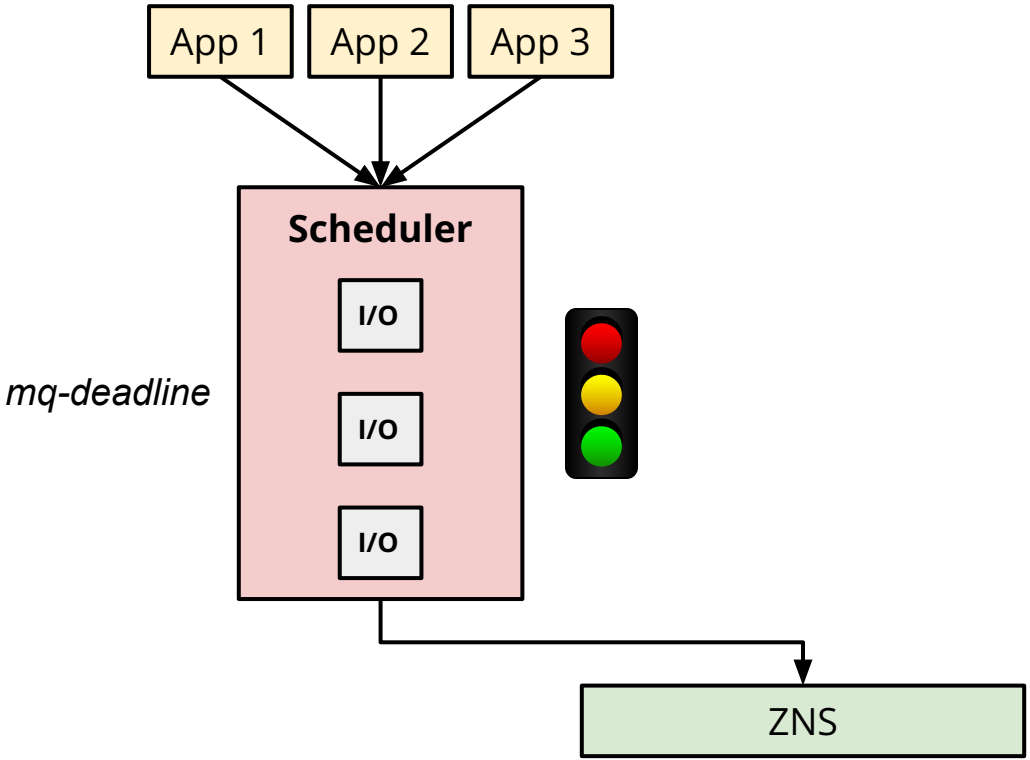
ZNS Integration - Tradeoff



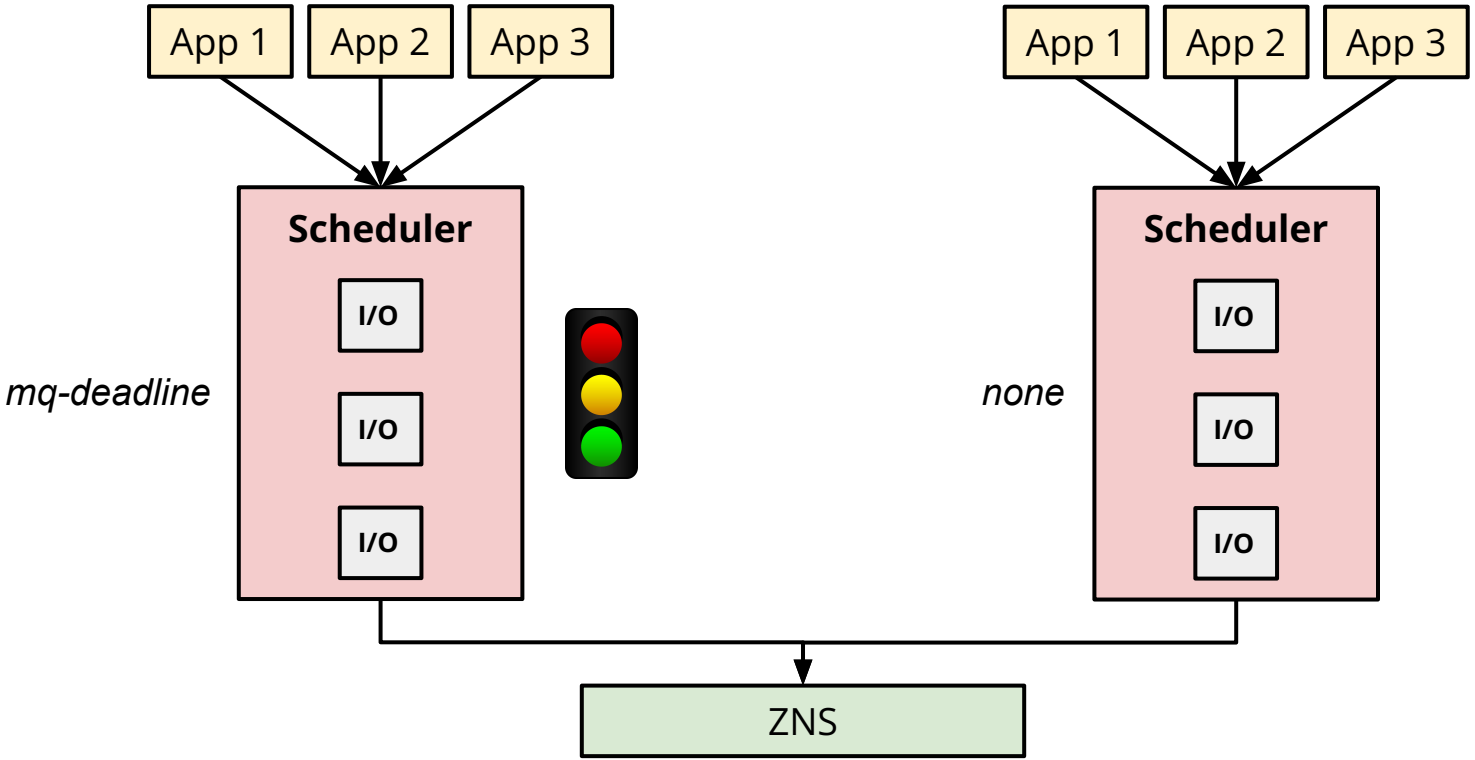
ZNS I/O Scheduling



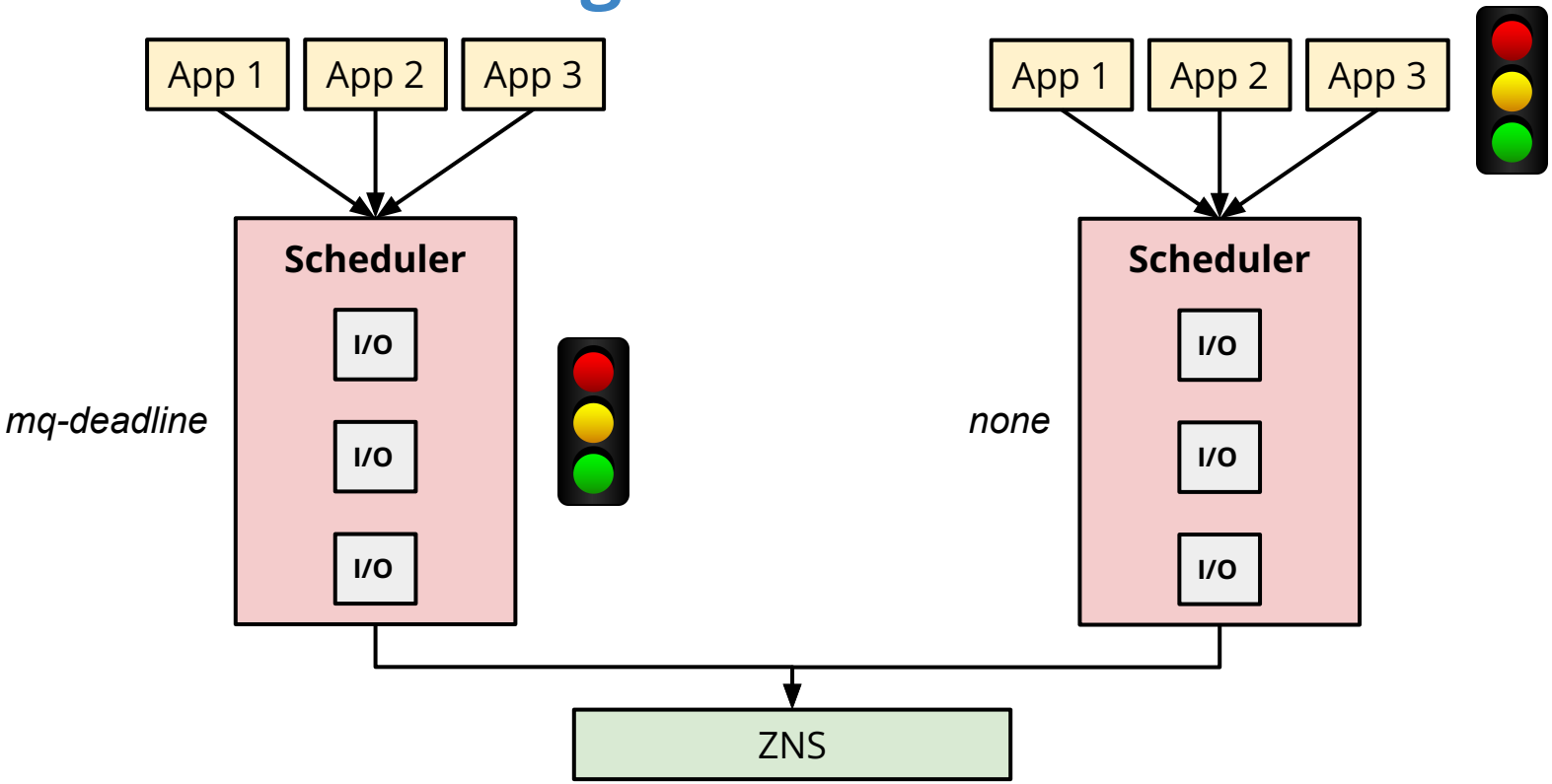
ZNS I/O Scheduling



ZNS I/O Scheduling

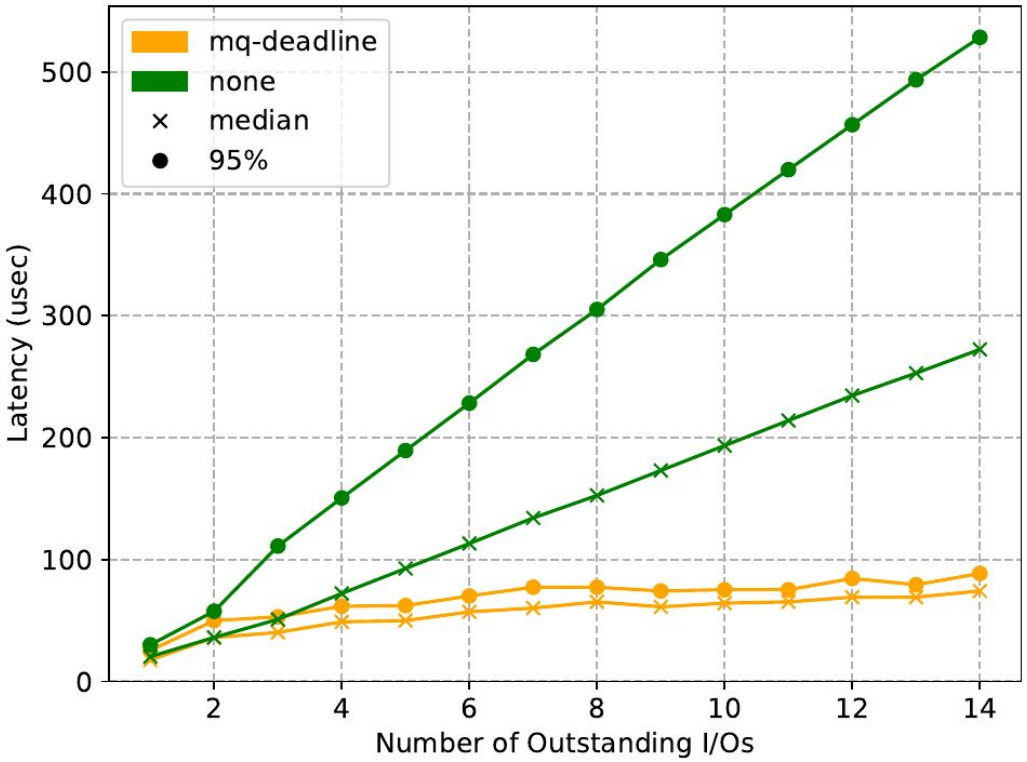


ZNS I/O Scheduling

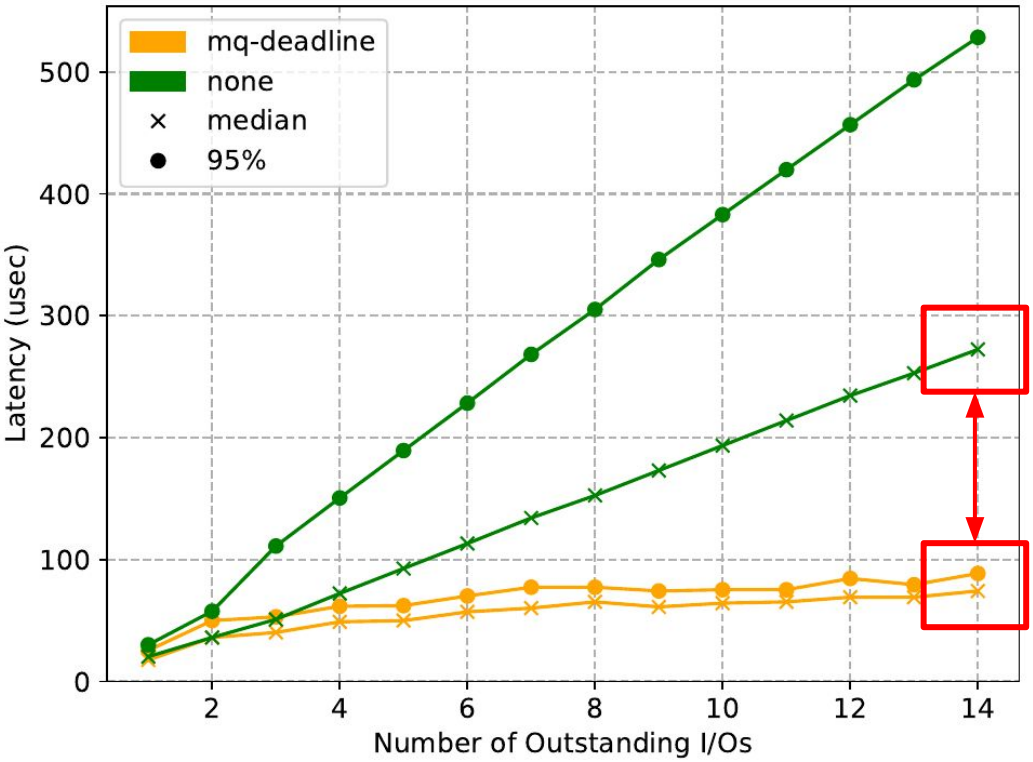


Initial Results

- Write performance of both scheduling configurations:
  - *mq-deadline* - increasing number of outstanding I/Os in a single zone
  - *none* - single I/O per zone and increasing concurrent zones

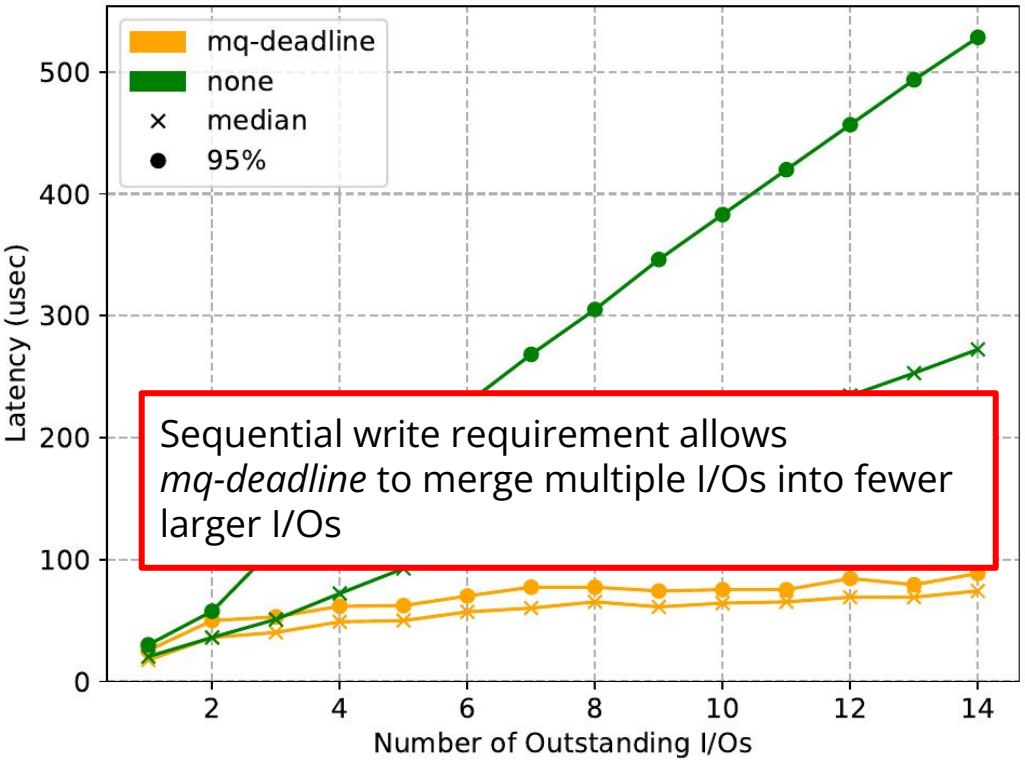


Initial Results



73% lower median latency with *mq-deadline*

Initial Results

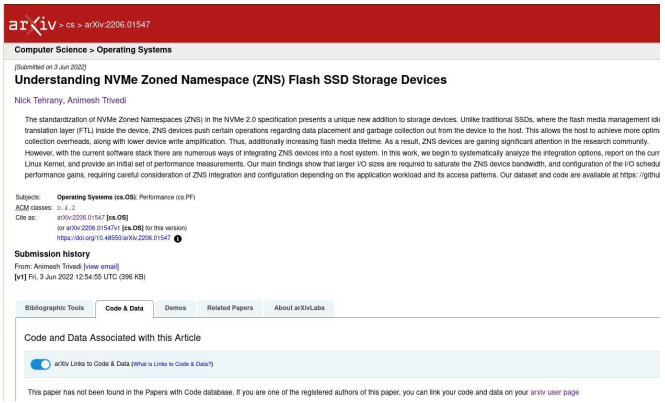


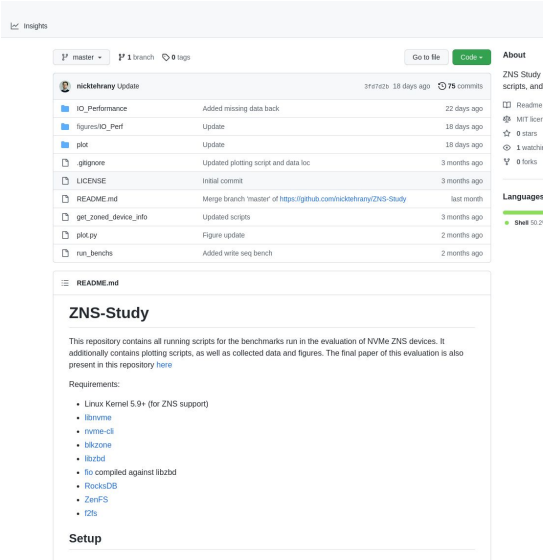
Ongoing Work

- Performance at the varying levels of integration
- GC implications at the different levels
- Concurrent instances of different levels

Summary

- ZNS present unique addition to host storage stack
- Gaining significant attention in research community





https://arxiv.org/abs/2206.01547

https://github.com/nicktehrany/ZNS-Study

Acknowledgements

- This work is generously supported by Western Digital Donations

- Matias Bjørling, Hans Holmberg, and ZNS team at Western Digital provided helpful comments and feedback