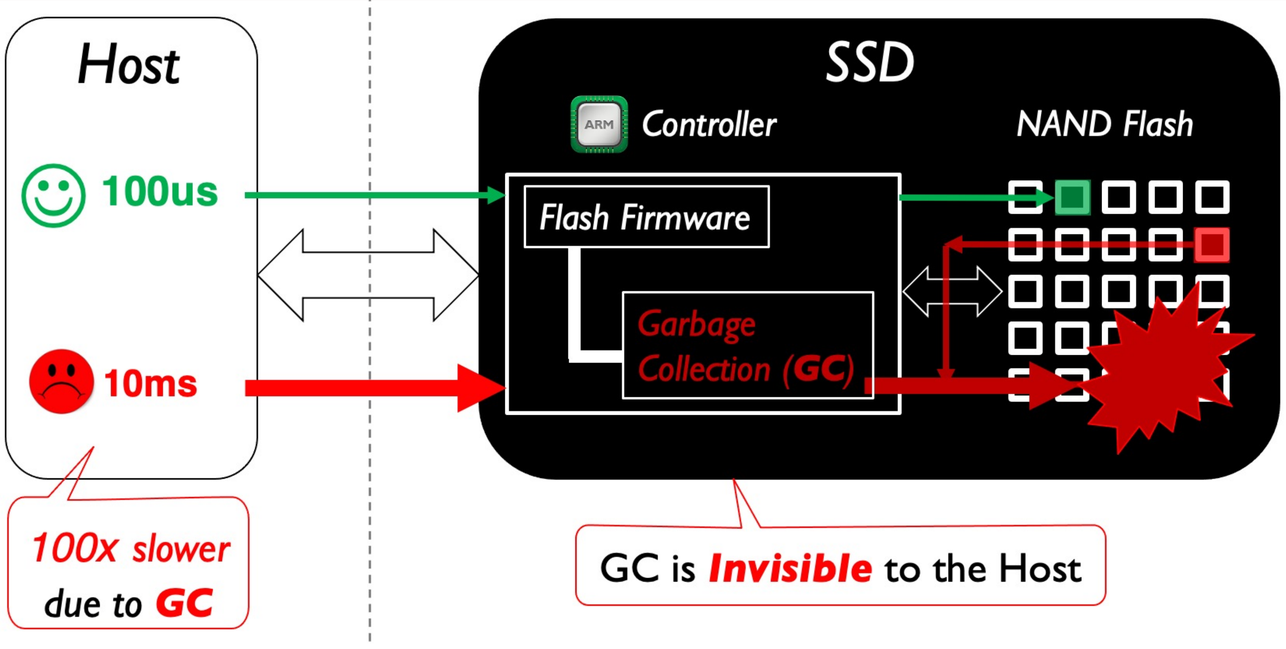
IODA: A Host/Device Co-Design for Strong Predictability Contract on Modern Flash Storage

Huaicheng Li, Martin L. Putra, Ronald Shi, Xing Lin, Gregory R. Ganger, Haryadi S. Gunawi

THE UNIVERSITY OF CHICAGO CMU NetApp

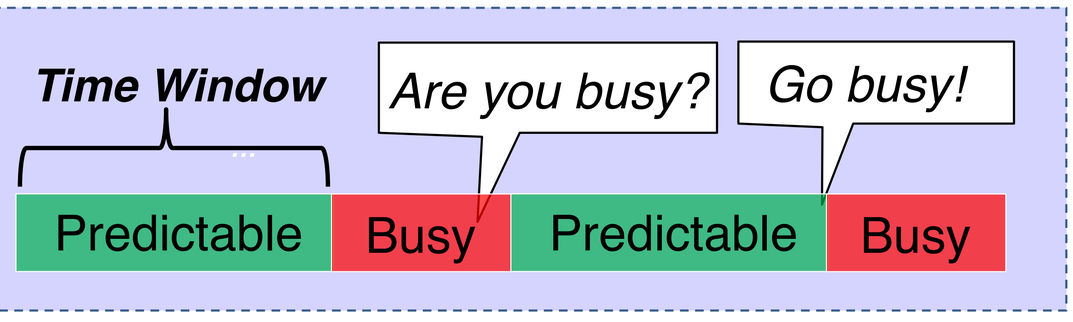
Latency Unpredictability in SSDs



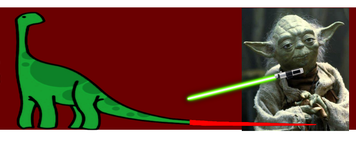
NVMe Predictable Latency Mode (PLM)

PLM is a major leap for host/SSD co-design

- Device status query & toggling
- Predictable/Busy Time Window (TW)

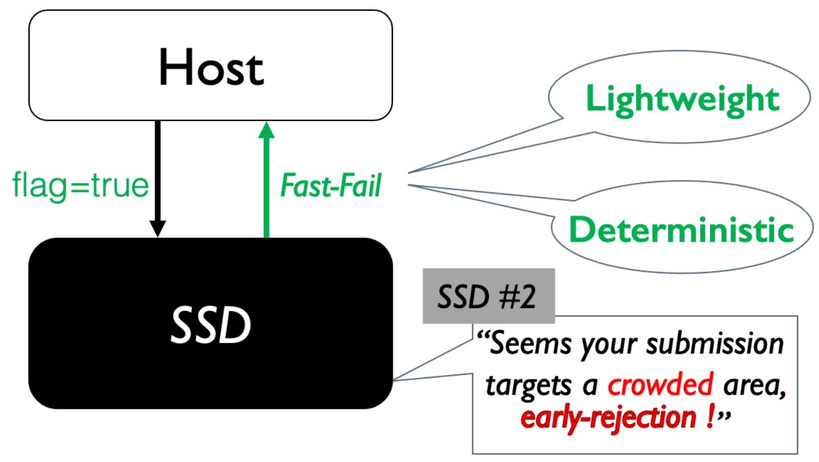


IODA's Journey for Always-Predictable Latencies



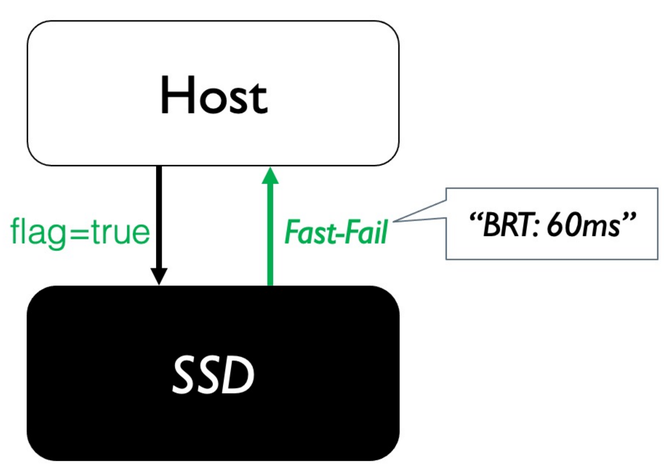
IOD_1: Predictable Latency Flagged I/Os

"Fail-if-Slow": SSDs fast-fail I/Os contending with GC



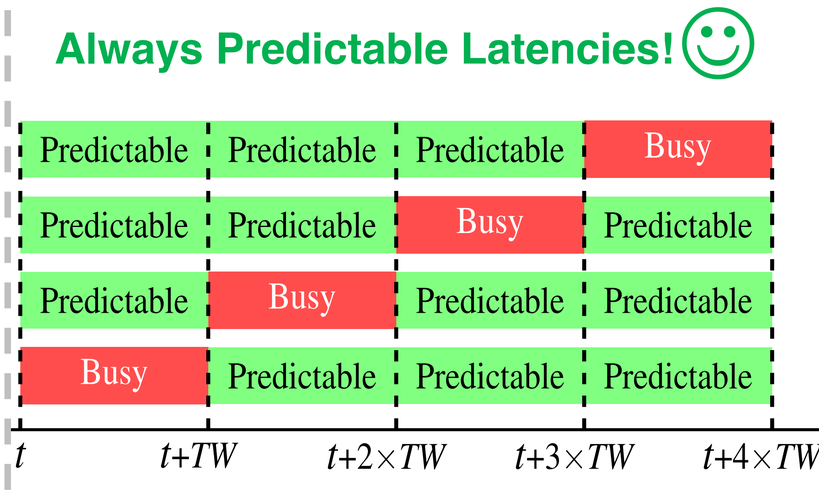
IOD_2: Busy Remaining Time (BRT)

Piggyback ***BRT***: reconstruct data from less busy SSDs

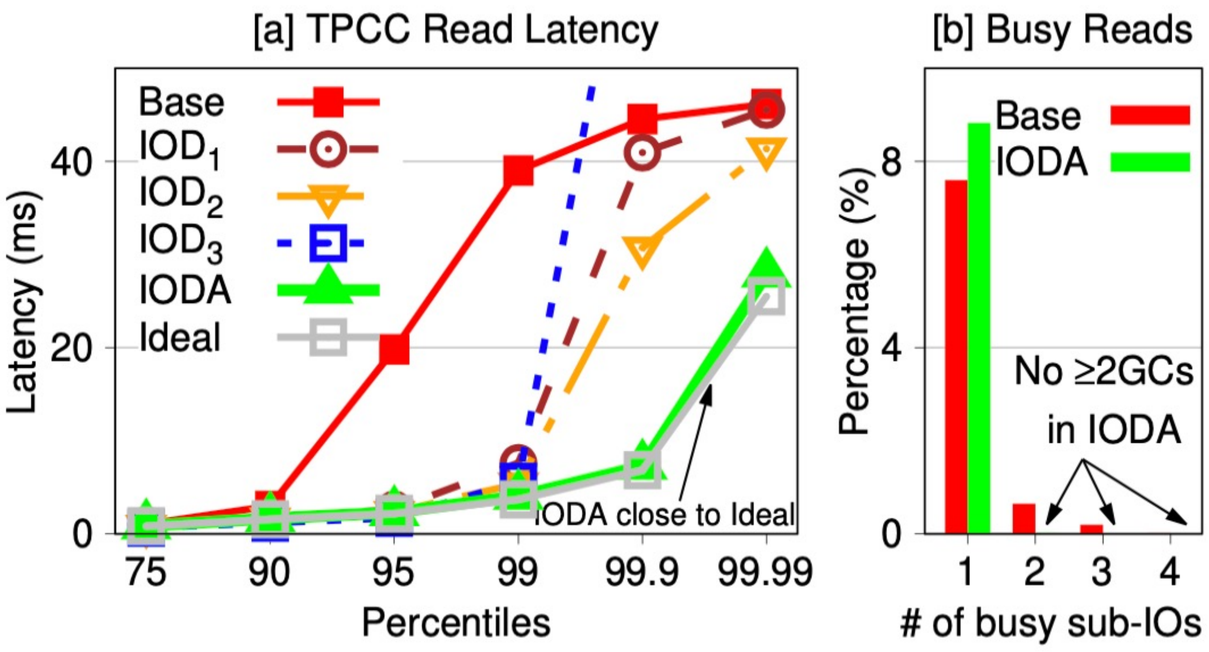


IODA Busy Latency Windows

IODA TW : Taking turns for GC; 1 busy SSD at a time!



IODA Evaluation Results

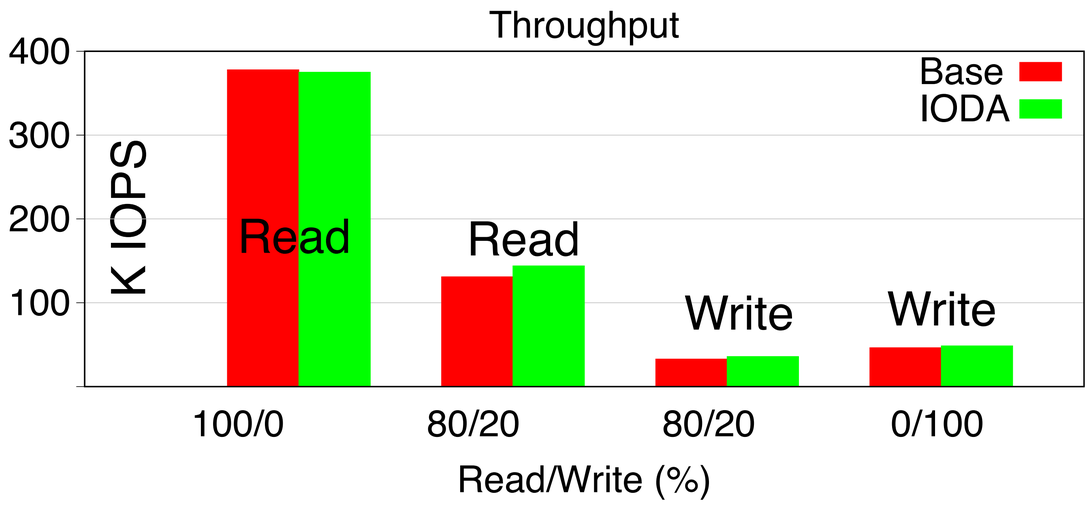


IODA is close to Ideal between P95 and P99.99

IODA successfully eliminates concurrent GCs

Comparison of IODA to state-of-the-art

	IODA	Preemption	Partitioning	Speculation	Suspension	Coordination	TTFLASH	Prediction
Determinism	✓	✗	✓	✗	✗	✗	✓	✗
Throughput	✓	✗	✗	✗	✓	✓	✗	✓
Transparency	✓	✓	✓	✓	✓	✓	✓	✗
Deployment	✓	✓	✓	✓	✓	✓	✗	✓



IODA does not sacrifice the array's throughput!

Appears in the ACM SIGOPS 28th Symposium on Operating Systems Principles (SOSP '21)