

Do we still need IO schedulers for low-latency disks?

Caeden Whitaker, Sidharth Sundar, <u>Bryan Harris</u>, Nihat Altiparmak

Computer Systems Lab

Computer Science & Engineering Department

UNIVERSITY OF LOUISVILLE

HotStorage '23 (Boston), July 9, 2023

Outline

- Introduction
- Latency costs
- Energy efficiency costs
- Conclusion?





With ever greater parallelism



System software affects energy efficiency



Disk IO scheduling for HDDs



Mechanical

Sequential

Reads/writes **single** location at a time Favors long contiguous regions

Reordering necessary for performance Efficient actuator arm movements

Disk IO scheduling for SSDs



Solid state

Internally Parallel

Reads/writes **multiple** locations at a time Some optimizations for contiguous regions

FTL layer hides internals

Only controller knows physical placement Host has reduced ability to reorder for performance

(See also open-channel, ZNS)

Disk IO scheduling for ULL SSDs



Ultra low latency (<10 µs)

Parallel

- Reads/writes **multiple** locations at a time
- Favors long contiguous regions

IO schedulers — Linux *blk-mq*

- "none" FIFO per *blk-mq* queue
- **mq-deadline** soft latency deadlines per request
- **kyber** token-based balancing between domains (read, write, discard, other)
- **bfq** per-application bandwidth budgets and heuristics to detect and prioritize real time tasks



Experimental setup





Latency cost

Scheduler	Read	% diff.	Write	% diff.
io_uring with defaults				
none	12.52	_	16.55	_
mq-deadline	13.39	6.9%	17.65	6.6%
kyber	13.44	7.3%	17.24	4.2%
bfq	15.01	19.9%	19.43	17.4%
<i>io_uring</i> with performance				
none	7.81	_	12.19	_
mq-deadline	8.32	6.5%	12.83	5.3%
kyber	8.33	6.7%	12.90	5.8%
bfq	9.41	20.5%	14.18	16.3%

HotStorage'23

Median latencies for 1 million 4 KB requests,

single proc., QD=1



Workloads

Microbenchmarks (fio using io_uring)

- Single-tenant (single proc., QD=1–128)
- Multi-tenant (1–128 threads)
- Random reads, writes, and 50-50 mix.

Macrobenchmarks (RocksDB — randomread)

- Random reads
- 1 to 64 threads



Scheduling hurts performance



Scheduling hurts energy efficiency



HotStorage'23

LO

RocksDB



Do we still need IO schedulers?

Scope

- Focus on **ULL** performance, pointing to the future
- Evaluated **performance** and **energy efficiency**

No!

- Significantly impairs ULL performance
- Scheduling impairs the energy efficiency of a system.

Maybe?

• What if fairness or QoS is valued?



Questions?

