Asynchronous I/O Stack: A Low-latency Kernel I/O Stack for Ultra-Low Latency SSDs

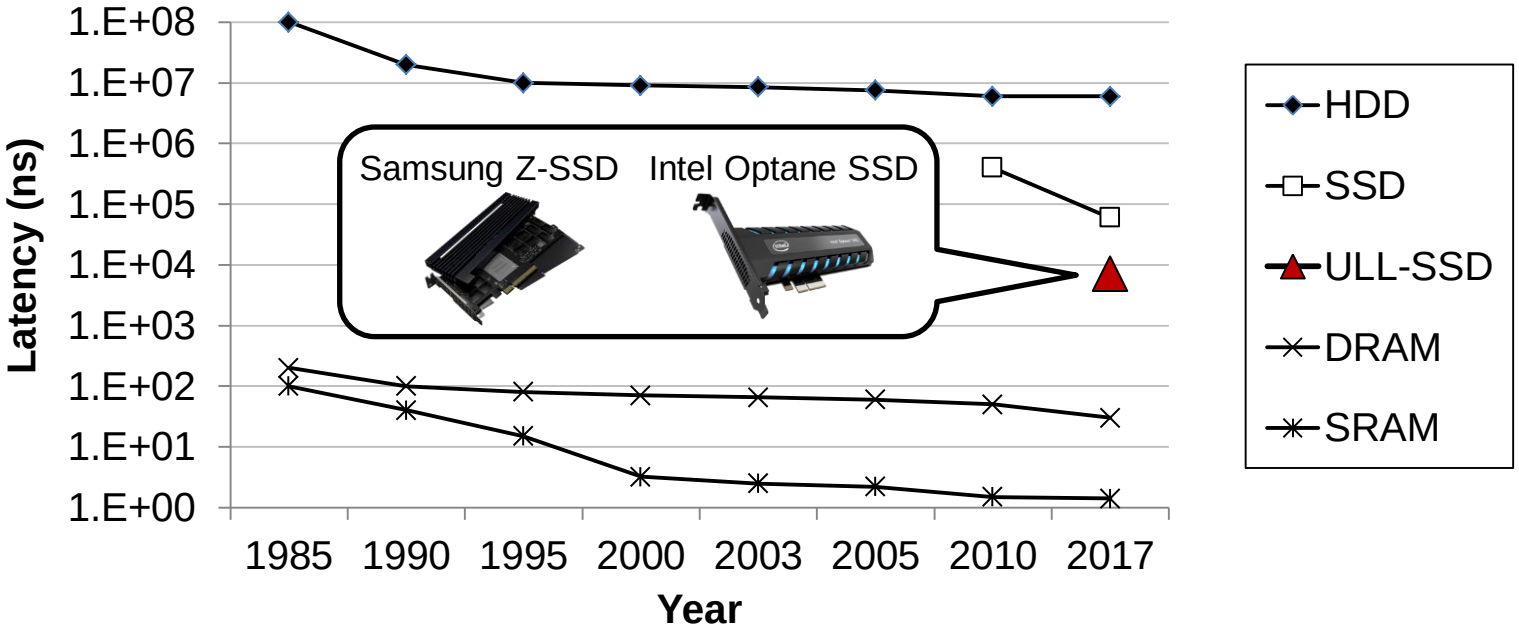
Gyusun Lee[1], Seokha Shin[1], Wonsuk Song[1], Tae Jun Ham[2], Jae W. Lee[2] and Jinkyu Jeong[1]

Sungkyunkwan University (SKKU)[1]

Seoul National University (SNU)[2]

Storage Performance Trends

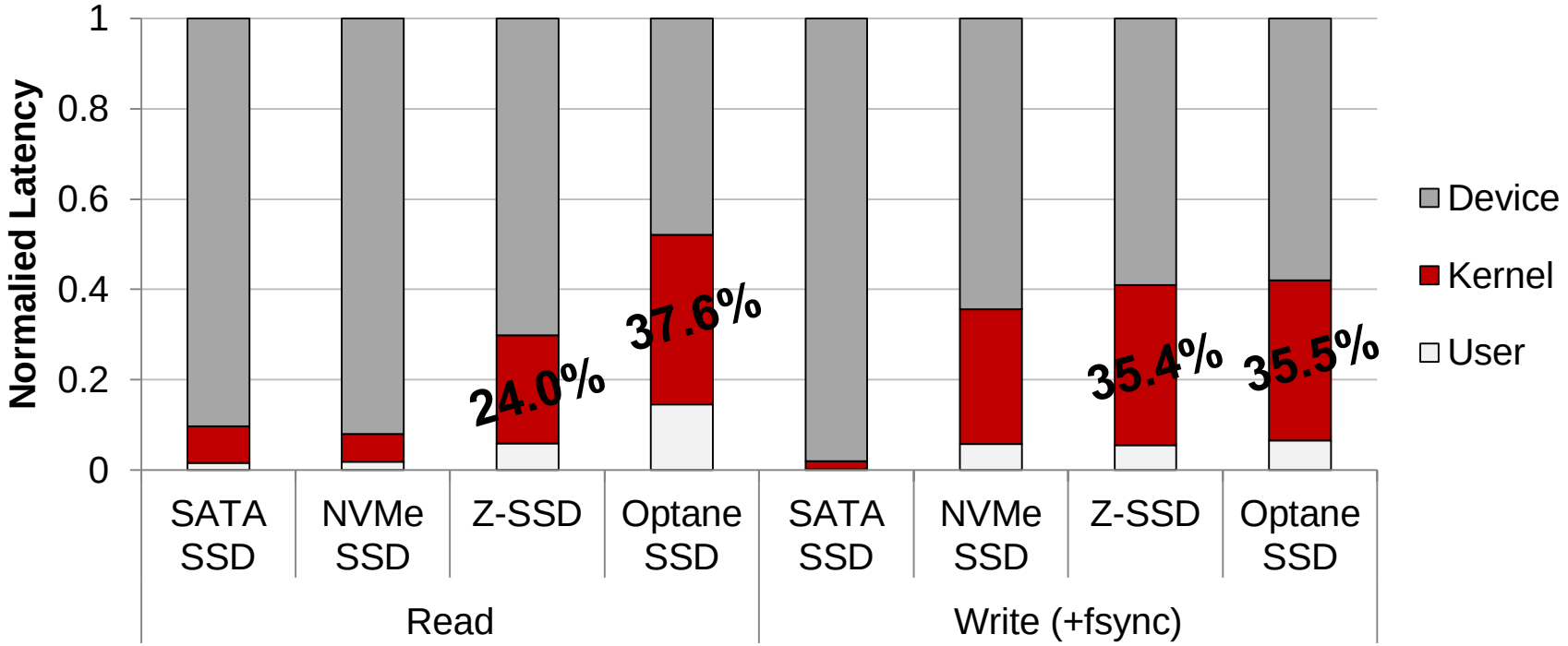
- **Emerging ultra low-latency SSDs deliver I/Os in <u>a few *µs*</u>**



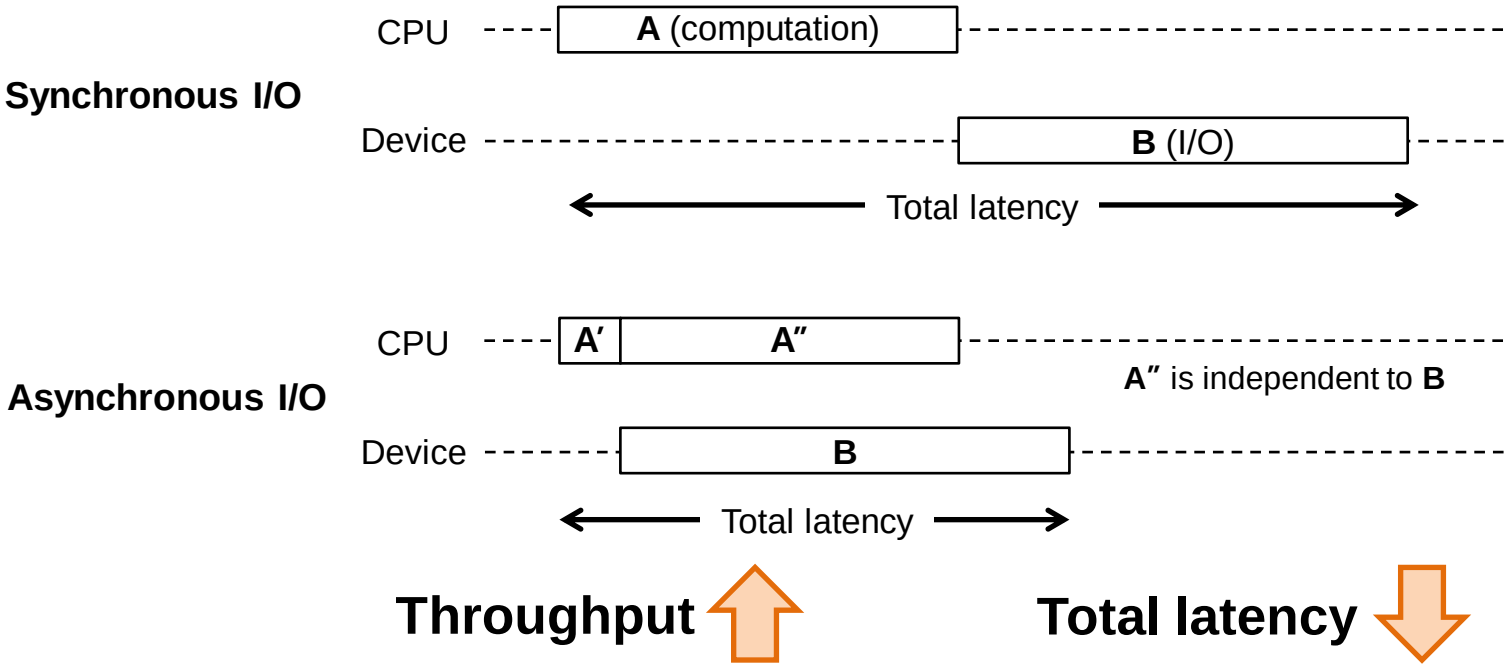
Source: R. E. Bryant and D. R. O'Hallaron, Computer Systems: A Programmer's Perspective, Second Edition, Pearson Education, Inc., 2015

Overhead of Kernel I/O Stack

- **Low-latency SSDs expose the overhead of kernel I/O stack**

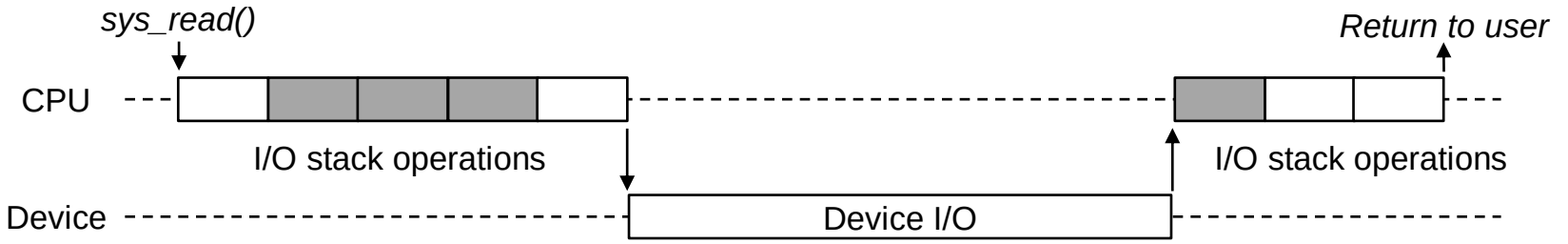


Synchronous I/O vs. Asynchronous I/O

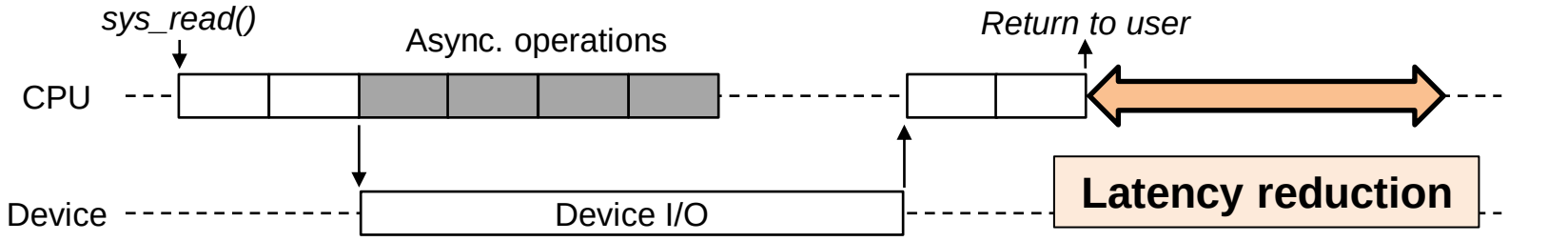


Our Idea: apply asynchronous I/O concept to the I/O stack itself

Read Path in Asynchronous I/O Stack

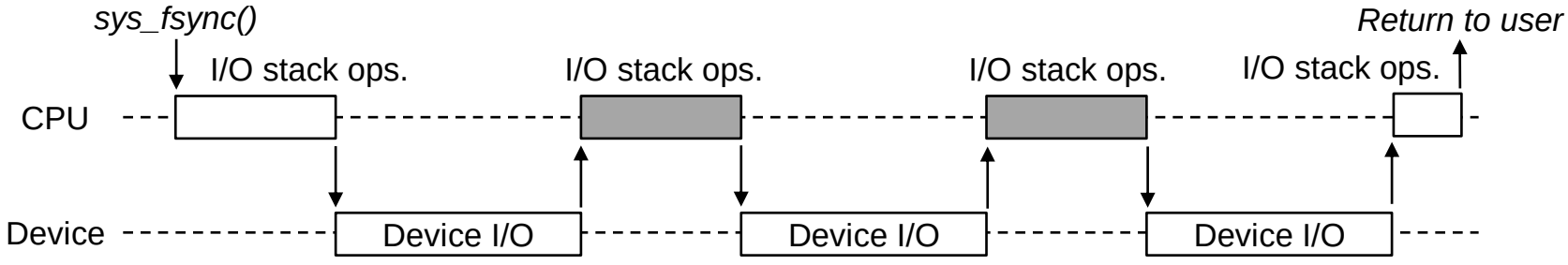


Vanilla I/O Stack

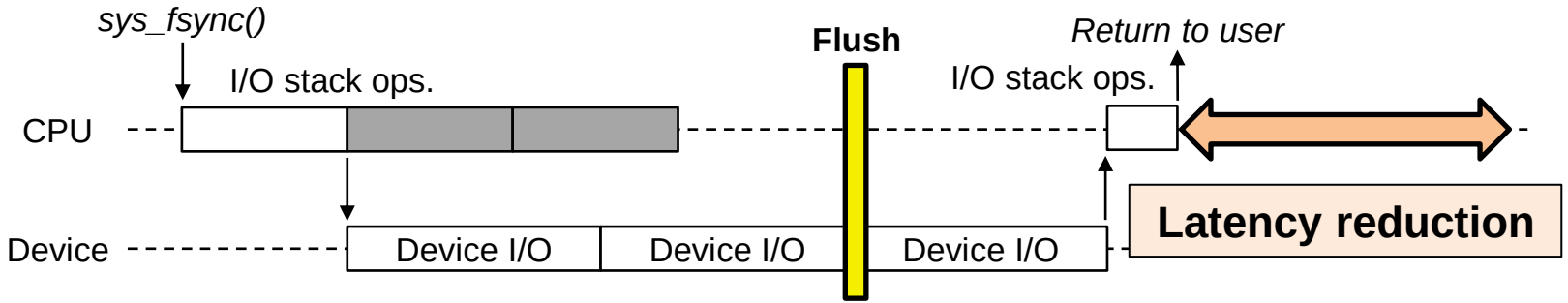


Asynchronous I/O Stack

Write(+fsync) Path in Asynchrnous I/O Stack



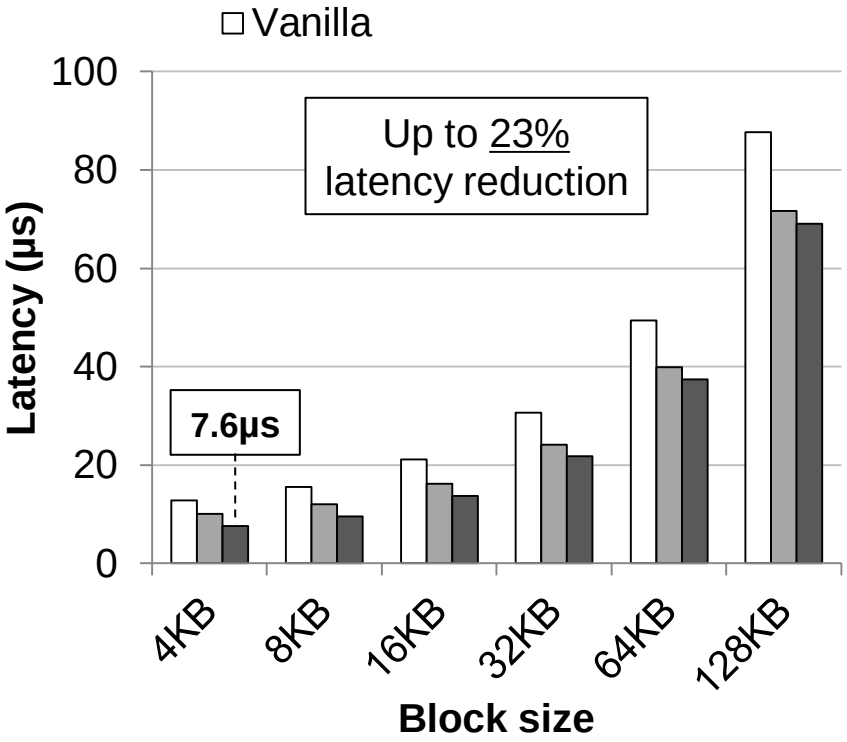
Vanilla I/O Stack



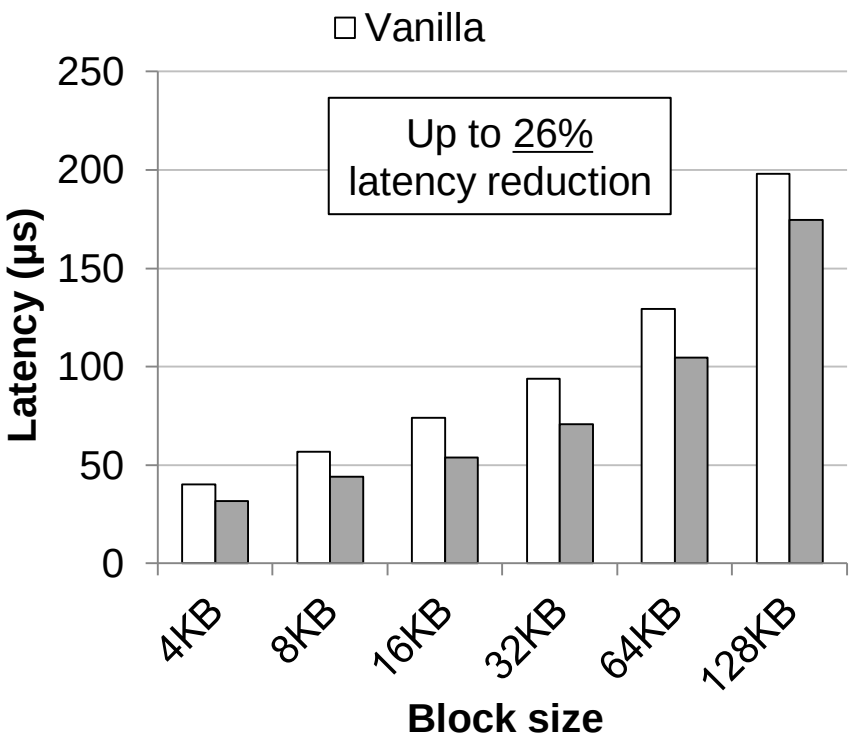
Asynchronous I/O Stack

FIO Performance (on Optane SSD)

- **Random read**



- **Random write (+fsync)**



Main Talk of Asynchronous I/O Stack

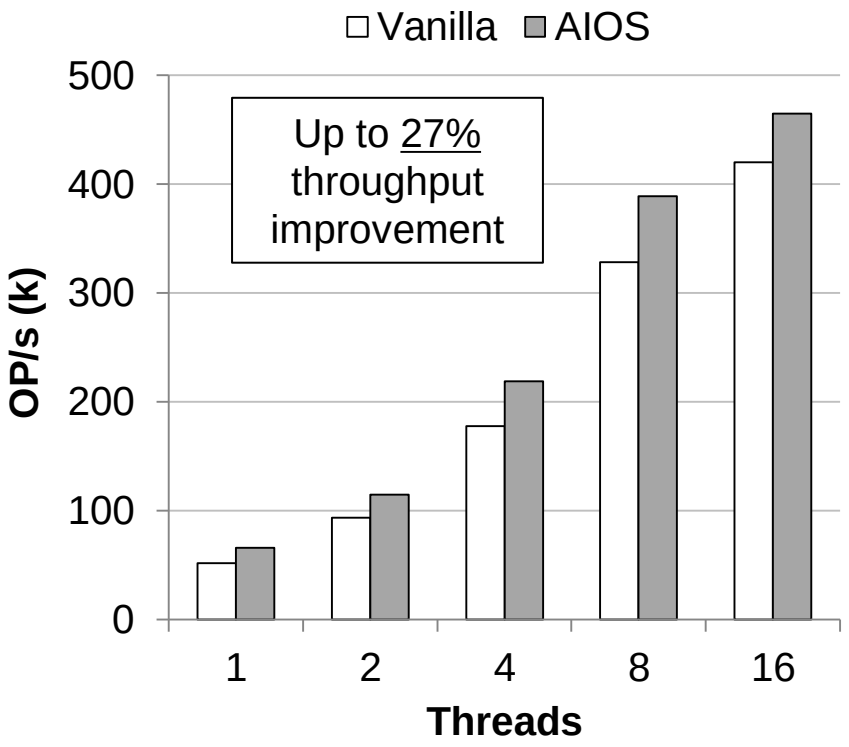
- **Detailed analysis of read/write I/O stack**
- **Asynchronous I/O stack**
  - Proposed read and write(+fsync) paths in detail
- **Lightweight block I/O layer**
  - Low-latency block I/O service for ultra-low latency SSDs
- **More performance measurements with various workloads**

USENIX ATC 2019, Thursday July 11
Track Ⅰ Exotic Kernel Features, 3:50pm

Extra Slides

Real-world Workload Performance

- **RocksDB DBbench readrandom**



- **Filebench varmail**