10/28 CS240 - Rethink Sync

Announcements

Lab 1 is due end of day November 2nd (this Sunday)

No class next Tuesday: Stanford holiday - Democracy Day

For next class (Thursday 11/6)

1. Read: F2FS: A New File System for Flash Storage
2. No reading questions

Paper backgrounds

- Rethink the sync
  - OSDI 2006 - 7th USENIX Symposium on Operating Systems Design and Implementation
  - Award paper
  - Leveraged Speculator - Extend Linux to support speculative execution. Supported roll back.
- What is durability in the paper? How is it related to performance?
- What is a disk cache?
  - write barrier
- What are transactions, ACID properties, and group commit?

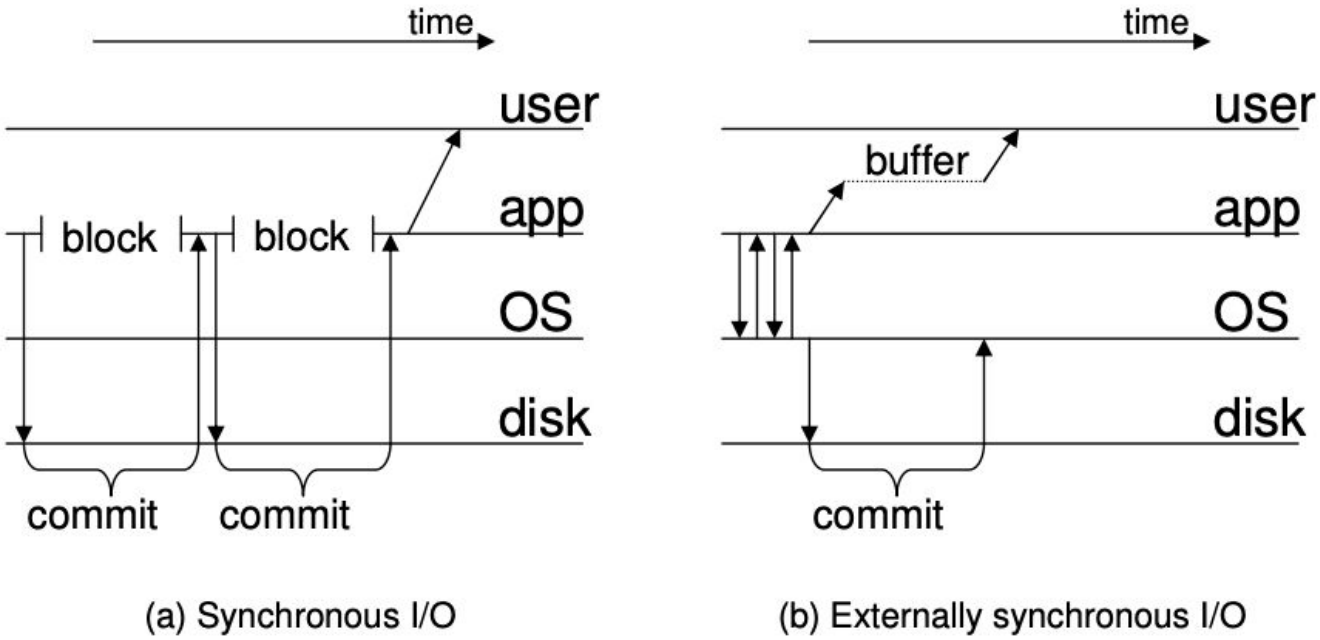
Unix/Linux sync/fsync durability history

- Original Unix - 30 second writeback cache: `sync; sync; sync`
- Linux ext3 fs:
  - Write-ahead logging
    - Ordered vs Journaled
  - Asynchronous
  - Synchronous
- Disk cache
  - write barrier

File system configuration	Data durable on `write`	Data durable on `fsync`
Asynchronous	No	Not on power failure
Synchronous	Not on power failure	Not on power failure
Synchronous with write barriers	Yes	Yes
External synchrony	Yes	Yes

Is this table correct for External synchrony? "Data durable on write"?

Explain externally synchronous I/O



What's hard about externally synchronous implementation

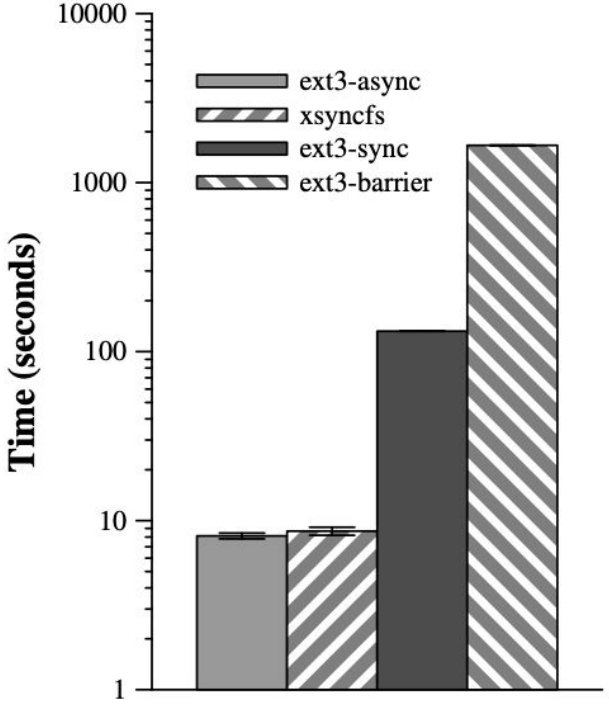
- Casual relationships
- Shared memory?
- `ioctl`?

What is output-triggered commits?

- What is the throughput and latency trade off it attempts to optimize?
- What was eager-commit?

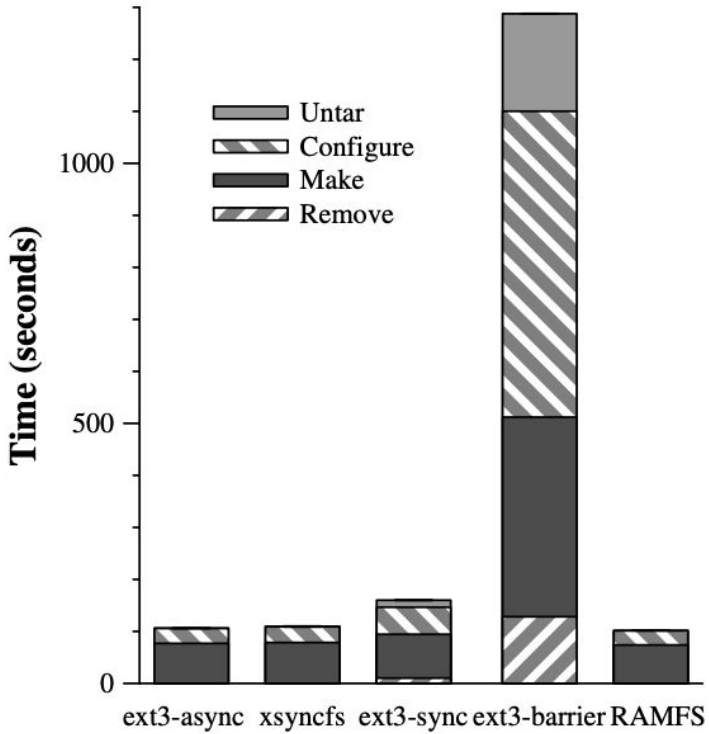
Limitations

- Disk errors: reads, writes
- Temporal expectations
- Multiple file system being modified



This figure shows the time to run the PostMark benchmark — the y-axis is logarithmic. Each value is the mean of 5 trials — the (relatively small) error bars are 90% confidence intervals.

Figure 4. The PostMark file system benchmark



This figure shows the time to run the Apache build benchmark. Each value is the mean of 5 trials — the (relatively small) error bars are 90% confidence intervals.

Figure 5. The Apache build benchmark