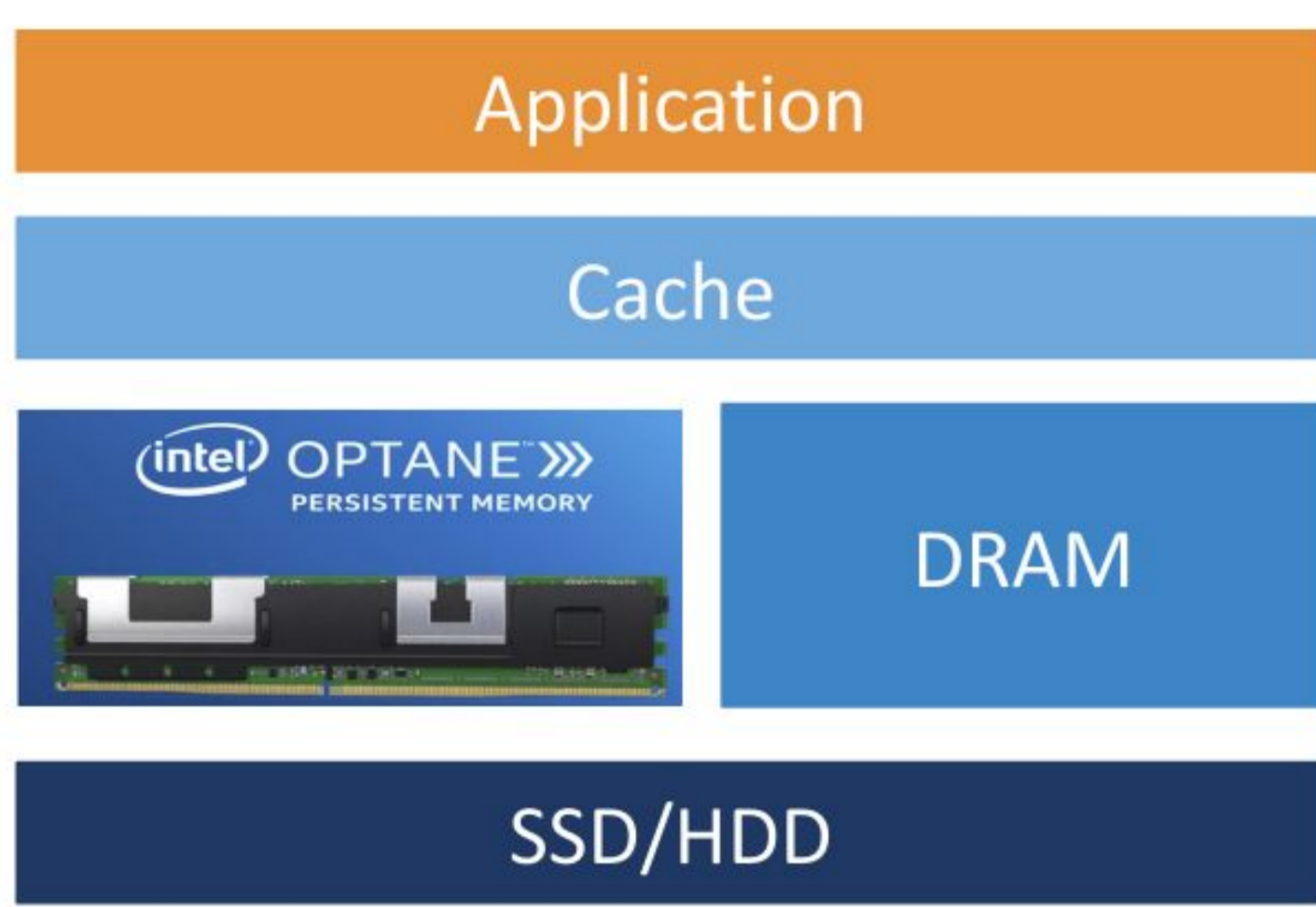


Zhuque: Failure is Not an Option, it's an Exception

George Hodgkins*, Yi Xu*, Steven Swanson, Joseph Izraelevitz
*co-first author

Persistent Memory (PMEM)



PMEM characteristics

- Byte-addressable.
 - Persistent over power failures.
 - Delivers DRAM-class latency and bandwidth.
- So
- PMEM enables an application's in-memory data to live beyond its lifetime.

Motivation

Making PMEM-resided data reusable

- Updates to PMEM should NOT violate data consistency across failures.

Existing PMEM Programming Model

- Failure-atomicity libraries: Allow a set of writes be atomically applied to PMEM.
- Whole system persistence: Automatically persist the whole system.

eADR

- A new technology ensures that all writes that reach the cache will be written to PMEM in the event of a power outage.
- Caches are effectively persistent.

Whole Process Persistence (WPP)

An Ideal PMEM Programming Model

- High performance
 - Persistent cache.
 - Limit the scope of persistence to a process (instead of whole system).
- Easy to use.
 - Can run unmodified applications directly on **Zhuque** --- musl-based implementation of WPP.

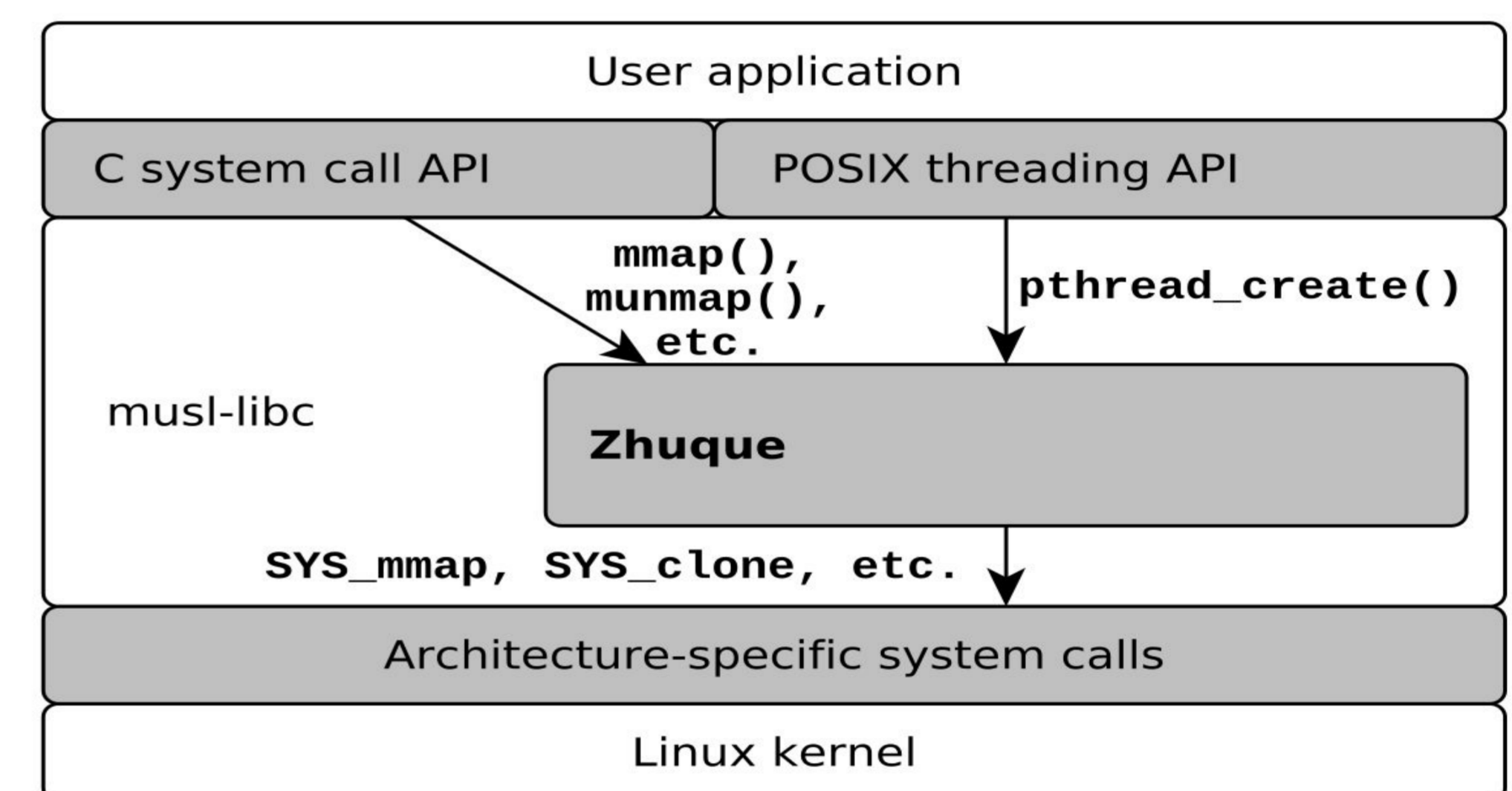
Normal Execution and Failure Recovery

The program runs as usual while Zhuque ensure that all program memory (stack, heap, etc.) resides in persistent memory. Save architectural state to PMEM on kernel entry.

From the application's perspective, power failure is delivered as an asynchronous signal (recoverable exception).

Our WPP Implementation: Zhuque

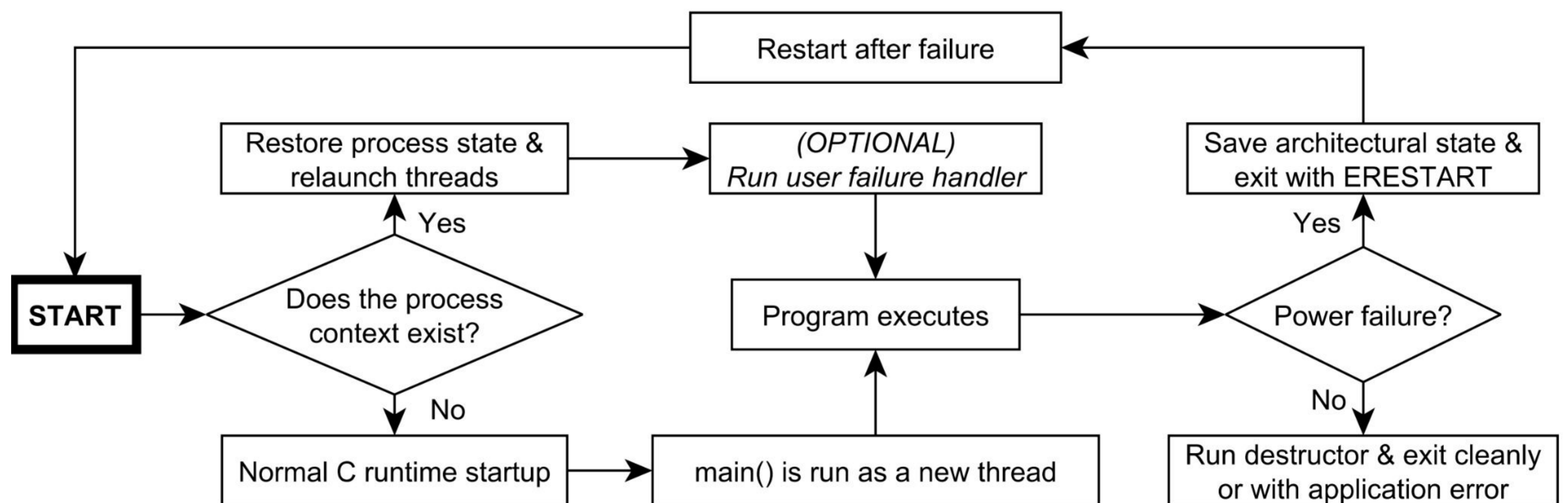
- User applications link to the C APIs provided by musl libc.
- We modify the implementation of the APIs and the arguments passed to the underlying system calls.



Process Lifecycle

Clean Start or Restart After Failure

- Zhuque modifies runtime startup and termination; application code is not modified.



Evaluation

Environment

- Platform with
 - One 20-core Intel Xeon Gold 6230 processor, running at 2.1 GHz.
 - A total of 96 GB of DRAM and 768 GB (6 x 128 GB) of Intel Optane DC Persistent Memory configured in 100% App Direct mode.
- Use Ext4 to manage persistent files and directly access PMEM pages via DAX.

Memcached

- Memcached went through a rewrite of the synchronization framework to use fine-grained locking across seven years of development and over thirty versions.
 - Most current PMEM libraries have strict requirement for the underlying concurrency strategy. These requirements make converting recent versions of memcached to run on PMEM a complicated and difficult process.
 - We run unmodified memcached on **Zhuque**.

