

IBM

## Alluxio - CAPI Flash Integration for Big Data Frameworks



Kavana N Bhat & Shajith Chandran Power Systems, IBM





## Big Data Ecosystem Today







## Alluxio (formerly Tachyon)

- Memory Speed Virtual Distributed Storage
- Enables Virtualized Data Across Multiple Types of Storage







## Why use Alluxio?

Memory-speed data sharing across jobs and frameworks

Data survive in memory after computation crashes



Off-heap storage, no GC

### HDFS / Amazon S3





## Storage Tier Hierarchy





¢.



## Automatic Data Migration

- Data can be evicted to lower layers if it is "cooling down"
- Data can be promoted to upper layers if it is "warming up"







## **CAPI: Coherent Accelerator Processor Interface**











## CAPI vs. I/O Device Driver: Data Prep







## Alluxio – CAPI Integration Choices

- Modify Alluxio to use the new user-level CAPI Flash block APIs
  - Requires changes to Alluxio
  - Alluxio-specific Implementation
  - > Non-reusable
- Create a User-Space Filesystem for CAPI Flash
  - Generic Implementation
  - Provides standard interfaces for file operations No changes to Alluxio





## CAPI Userspace Filesystem (USFS)

- The userspace filesystem is primarily contained in a library(libcflsh\_usfs), that runs on top of the CAPI flash block library API.
- □ Provides analogs of all major filesystem APIs (open, close, read, write, aio\_read, aio\_write etc).
- Provides wrapper layer that allows applications to work without modifications with use of the library preload feature.
  - > Wrapper intercepts filesystem calls from libc intended for libcflsh\_usfs and routes them internally.



#### Library Preload for existing Application POCs





## Alluxio with CAPI USFS







## Alluxio Performance with CAPI



Alluxio with Traditional FS on Flash





## Alluxio Performance with CAPI



Alluxio with CAPI USFS on Flash





## References

- Alluxio Project: <u>www.alluxio.org</u>
- IBM CAPI: <u>http://ibm.biz/powercapi</u>
- IBM CAPI User Space Block Library: https://github.com/open-power/capiflash
- What is CAPI?
  - http://lt.be.ibm.com/stg/ltu48342
  - For Partners: <u>http://www.ibm.com/services/weblectures/dlv/partnerworld/ltu48850</u>
- SuperVessel Website: <u>https://ptopenlab.com/cloudlabconsole/#/</u>
- Contact: kavana.bhat@in.ibm.com, shajithchandran@in.ibm.com





## **Questions?**





## **Thank You**



Prior to CAPI, an application called a device driver to utilize an FPGA Accelerator.

The device driver performed a memory mapping operation.







#### With CAPI, the FPGA shares memory with the cores















# Works

CAPI

CAPI Developer Kit Card



**POWER8** Processor





POWER8 Processor



## CAPI solution flow





Power Systems



