Towards A Scalable, Resilient, and Efficient Data Service for Exascale Computing

Michael Brim, Tonglin Li, Sarp Oral, Geoffroy Vallee, Feiyi Wang, Scott Atchley

Oak Ridge National Laboratory

1st Joint International Workshop on Parallel Data Storage and Data Intensive Scalable Computing Systems (PDSW-DISCS'16)

MONDAY, NOVEMBER 14, 2016
SALT LAKE CITY, UT
ExCDS: the Exascale Computing Data Service

• Ease data management for application and I/O middleware developers
• Improve center-wide I/O performance and load balance
• Better support for variety of workloads
  – Checkpoint-Restart
  – Data analysis (in-situ, co-analysis, mining, deep learning, etc.)
  – Data visualization and computational steering
• Versatility to accommodate variety of data storage technologies
ExCDS Architecture

• Integrate with applications

• Deploy data service across entire storage hierarchy
  – Service process per node
  – Connected via scalable overlay

• Data service capabilities:
  – Monitoring: host, network and storage
  – Data placement & movement
  – BB and PFS behavior tuning
  – Application I/O aggregation and scheduling
User API and Semantics

• Namespace
  – Attach/Detach
  – Set attributes

• Object Collection
  – Create/Delete
  – Set attributes
  – Open/Close
  – Commit (versions)

• Object
  – Create/Delete
  – Read/Write
Service Components

IPC – domain socket and shmem

User Request Manager

Request Handlers

Data Mover

Smart I/O
(Data placement and Movement Decision Engine, Distributed Commit)

Data Manager

Remote Forwarding Layer

Storage Layer

Metadata Layer

Monitoring

MELT

POSIX FS

Burst Buffer

KVS

SNOflake Overlay
References & Contact Information

• References
  – MELT/ SNOflake
  – SmartIO

• Contact
  – ExCDS project: excds@ornl.gov