



#### Performance Characteristics of the Lustre File System on the Cray XT5 with Respect to Application I/O Patterns.

#### Lonnie D. Crosby

#### National Institute for Computational Sciences



# **Application Performance**

- Computation
  - -2.3 GHz Quad Core AMD Opteron
- Inter-process Communication

   SeaStar2+ 3D-Torus
   HyperTransport
- Memory Capacity
  - 8 16 GB
- File I/O
  - Lustre File System



Compute the Future

#### COMPUTE Efuture

# A Bigger Picture: Kraken XT5

- Computation Nodes
   8253 (608 TF)
- Object Storage Servers
   48 (30 GB/s)
- Object Storage Targets
   336 (2.4 PB)





# Factors which affect file I/O.

- Who is performing I/O and when?
   Number of processes which perform I/O.
- How is I/O performed in the Application?
  - I/O Rate from processes.
  - File access pattern.
- How is I/O handled by the Lustre file system?
  - File striping pattern.





# File Striping in Lustre

Stripe Count





• 32 MB per OST (32 MB – 5 GB) and 32 MB Transfer Size





Single Writer Read Performance





Single Writer Transfer vs. Stripe Size





- Conclusions
  - Striping a single file over multiple OSTs does not substantially improve performance if a single process is performing I/O.
  - Performance can be limited by either transfer or stripe sizes.



- Important Considerations
  - Data locality
  - Data Continuity
- Parallel file Structure













Single Shared File (32 Processes) 1 GB and 2 GB file





#### File Per Process

#### • 128 MB per file and a 32 MB Transfer size





#### 32 MB per process and 32 MB Transfer size

----POSIX Write (MB/s) ----MPIIO ----HDF5 **Proccesses** 

Single Shared File Write Performance



# Conclusions

- Important Considerations
  - Data Locality and Continuity
- Parameters
  - Restrict a process to one OST.
  - Utilize sufficiently large transfer and stripe sizes.
  - Consider the layout of parallel files.
- Limitations manifest themselves at large process counts.



# Subsetting I/O

Lustre

- Advantages
  - Decreases number of files.
  - Increases the volume of I/O
- Disadvantages
  - Communication and memory costs.
- For shared files
  - Increases number of files