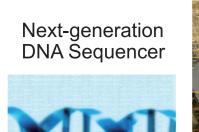
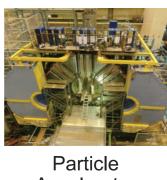


System Software for Post Petascale Data Intensive Science

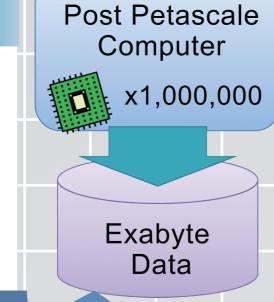
Background

Data-intensive Science









Objective

Development of

System Software

for Data-intensive Computing to promote Data-intensive Science

This research is supported by JST/CREST



Research Topics

Runtime System

File System Kernel Driver

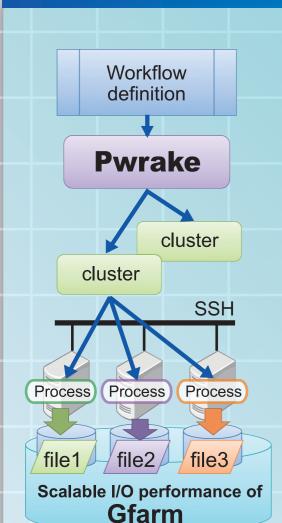
Distributed File System

File System Kernel Driver is developed at The University of Electro-Communications

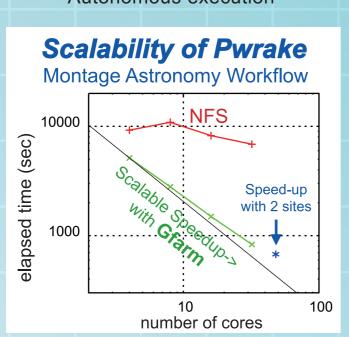


Runtime System

Pwrake: Scalable Workflow System

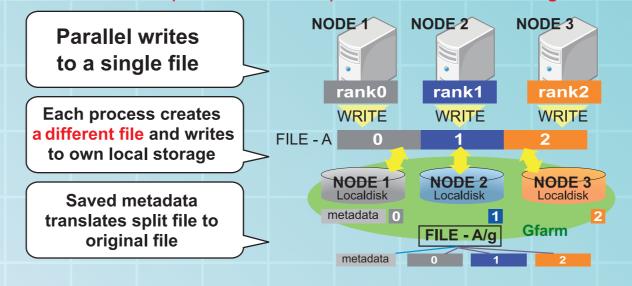


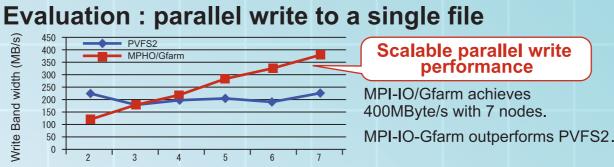
- Workflow Language : Rake Allows users to write complex scientific workflows
- Scalable I/O Performance by: Adaptive scheduling File locality
- Graph partitioning • For the use of 1 million cores: Hierarchical management Autonomous execution



MPI-IO / Gfarm

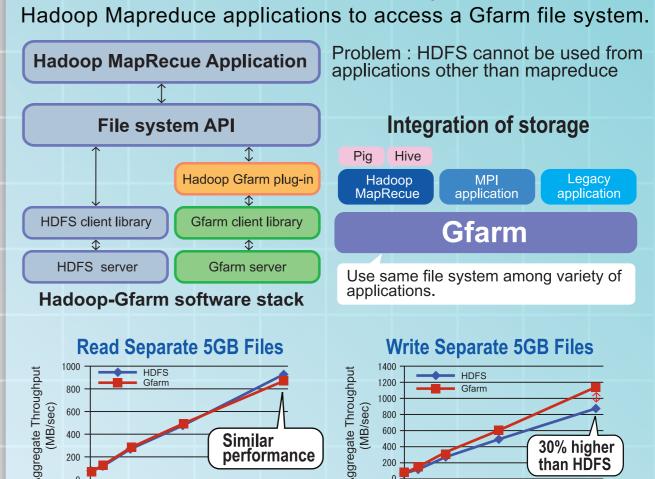
MPI-IO/Gfarm is an MPI-IO implementation for the Gfarm file system designed to achieve scalable parallel I/O performance. Problem: Poor performance of parallel writes to a single file





Hadoop / Gfarm

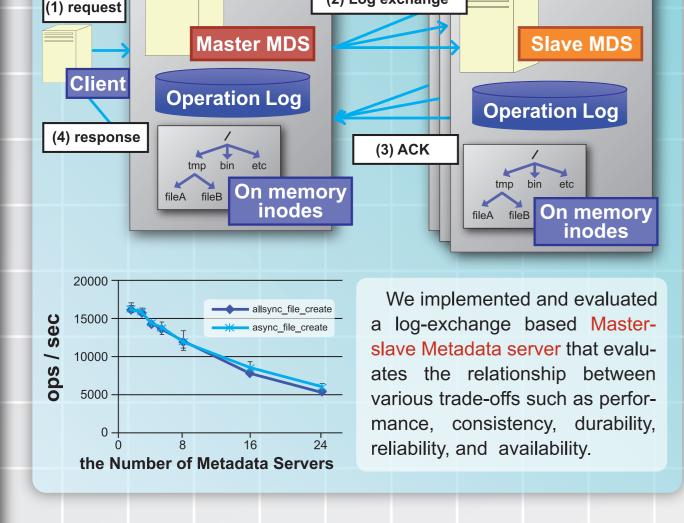
We have developed a Hadoop-Gfarm plugin which enables



Distributed File System

(2) Log exchange

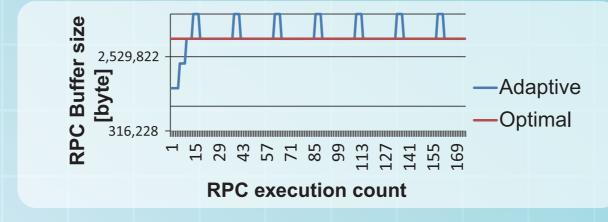
Reliable Distributed Metadata Servers



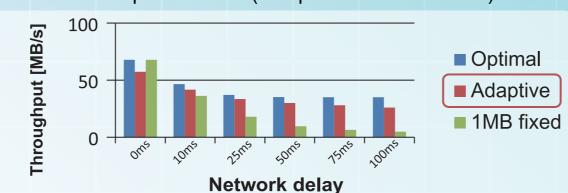
Optimization of Remote File Access

Dynamic optimization method for remote file access systems under the high network latency environment.

Change the transferring block size dynamically:

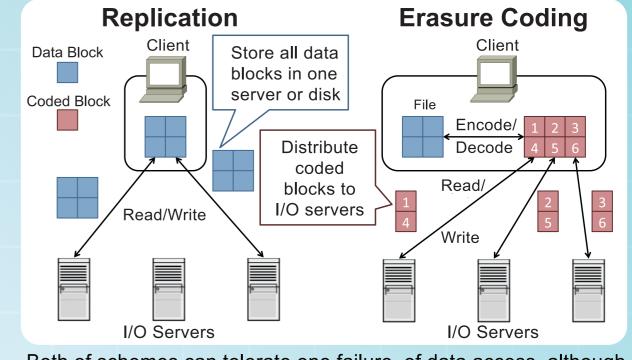


Performance improvement: (Adaptive vs. 1MB fixed)



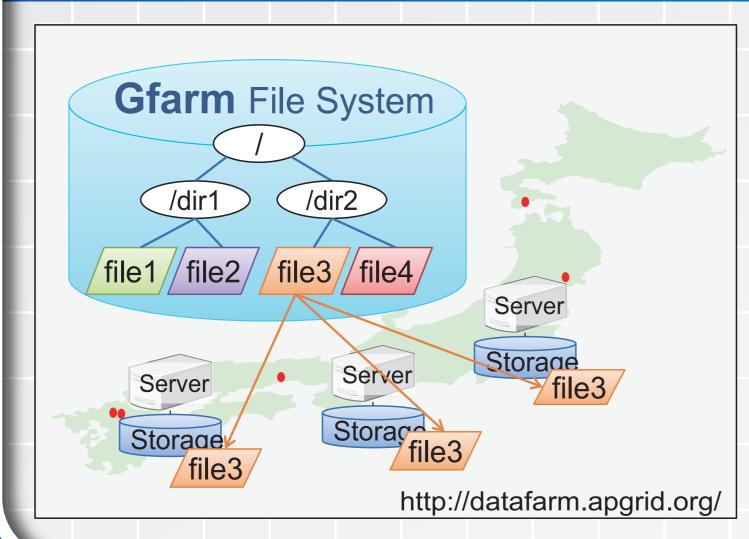
Redundant Data Storage Method

We plan to propose a redundant data storage approach through the use of erasure coding scheme, while keeping high I/O performance

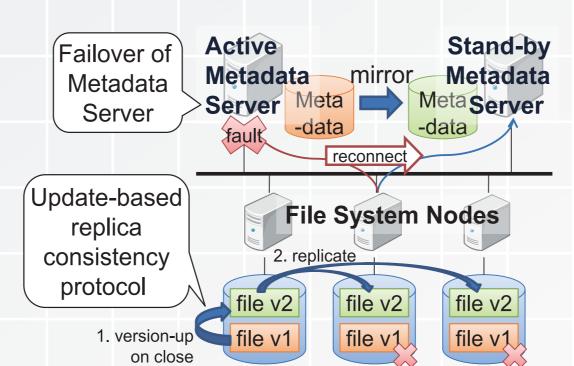


Both of schemes can tolerate one failure of data access, although capacity overhead of replication is high (at least 2 copy of whole data blocks). On the other hand, if we adopt erasure coding, •capacity overhead is smaller than replication. computation costs of encoding/decoding are needed.

Gfarm: Wide-area Distributed File System



Gfarm Technology for Failover and File Replication



3. remove after replicate



This research is supported by the **RENKEI** (Resources linkage for e-Science) project sponsored by MEXT of Japan.

RENKEI aims to federate (= renkei) e-Science communities through research and development of middleware technologies. RENKEI website: http://www.e-sciren.org/

NTT Communications

METI Next Generation Green IT infrastructure: Accountability for Cloud Computing

As part of a 3 year effort to create highly reliable and accountable cloud storage platforms (PI: NTT Communications), The Ministry of Economy, Trade and Industry(METI) has awarded The University of Tsukuba a grant to research exascale cloud storage infrastructure technologies capable of federating thousands of individual clouds.