

LA-UR-12-23586

Approved for public release; distribution is unlimited.

Title: GlusterFS One Storage Server to Rule Them All

Author(s): Boyer, Eric B.  
Broomfield, Matthew C.  
Perrotti, Terrell A.

Intended for: Computing and Information Technology Student Mini Showcase, 2012-08-02  
(Los Alamos, New Mexico, United States)



Disclaimer:

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes.

Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.



# GlusterFS

## *One Storage Server to Rule Them All*

### **Team Members:**

Matthew Broomfield (New Mexico Tech)

Eric Boyer (Michigan Tech)

Terrell Perrotti (South Carolina State University)

### **Mentors:**

David Kennel (DCS-1)

Greg Lee (DCS-1)

### **Instructors:**

Dane Gardner (NMC)

Andree Jacobson (NMC)



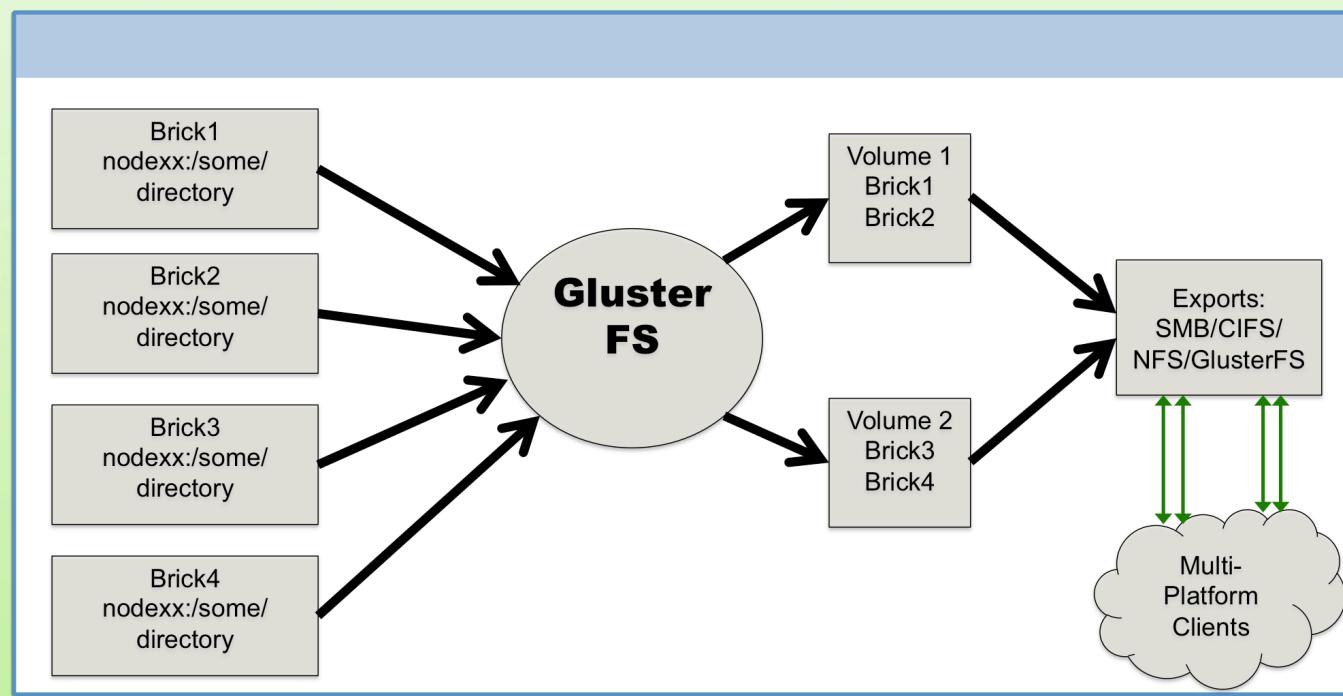
# Outline

- Introduction to GlusterFS
- Services
- Administration
- Performance
- Conclusions
- Future Work



# What is GlusterFS?

- GlusterFS is a Linux based distributed file system, designed to be highly scalable and serve many clients



# Why Use GlusterFS?

- No centralized metadata server
- Scalability
- Open Source
- Dynamic and live service modifications
- Can be used over Infiniband or Ethernet
- Can be tuned for speed and/or resilience
- Flexible administration



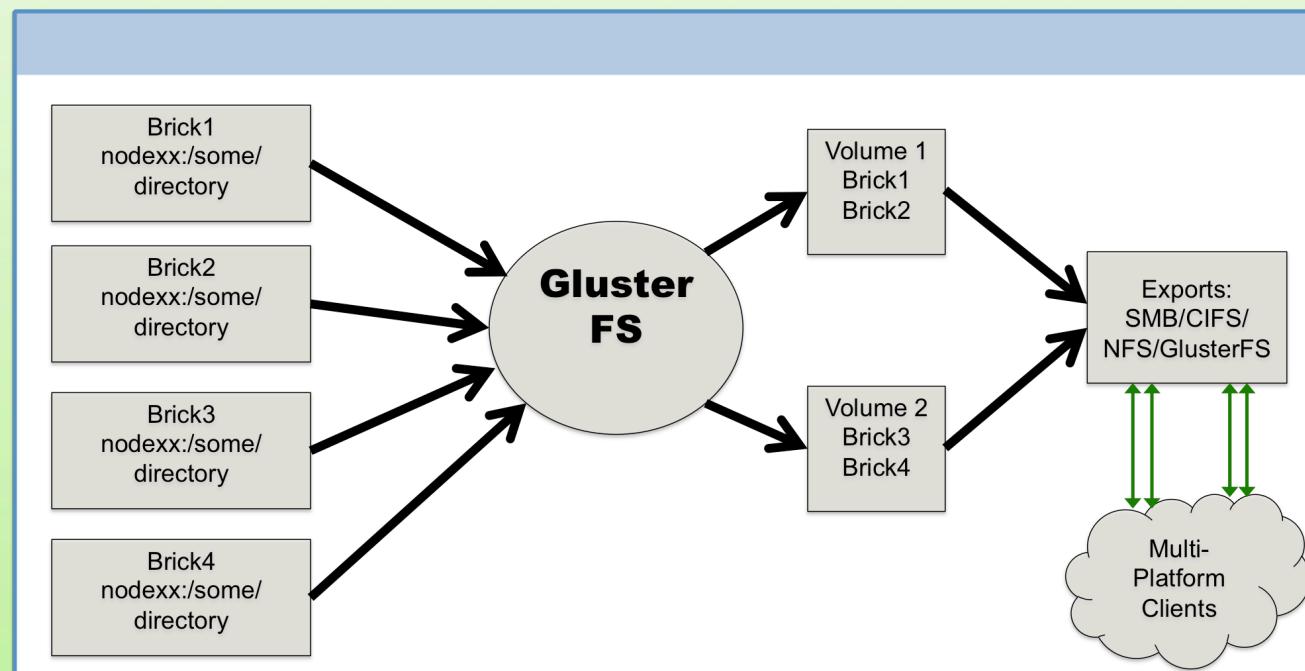
# Where Is It Useful?

- Enterprise environments
  - Virtualization
- High Performance Computing (HPC)
- Works with Mac, Linux, and Windows clients



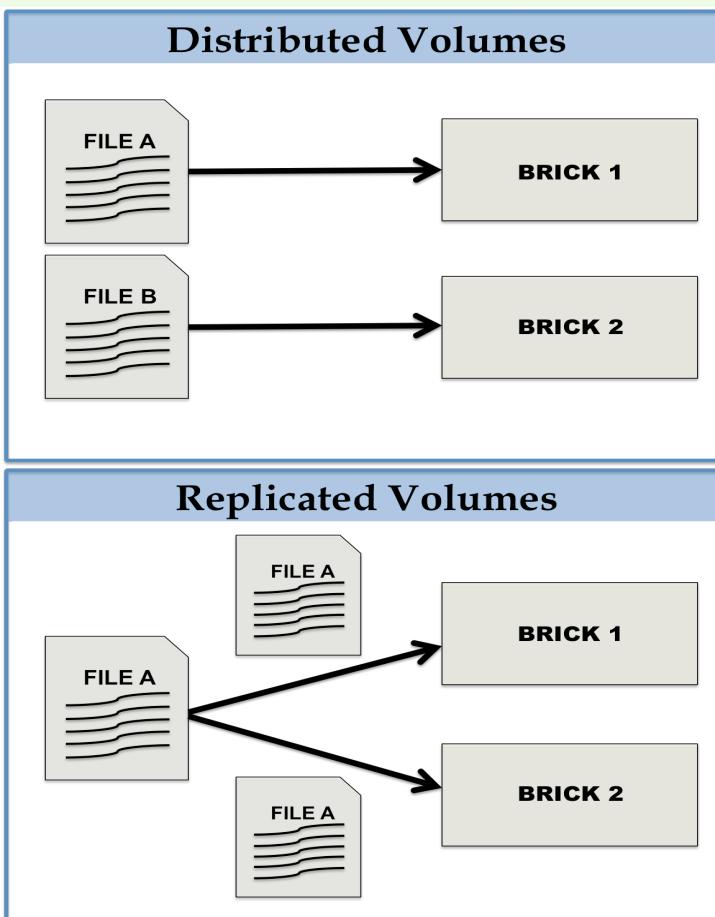
# How Does It Work?

- Individual nodes export bricks ( directories ) to GlusterFS
- GlusterFS combines bricks into virtual volumes



# GlusterFS Volume Types

- Different GlusterFs volume types:



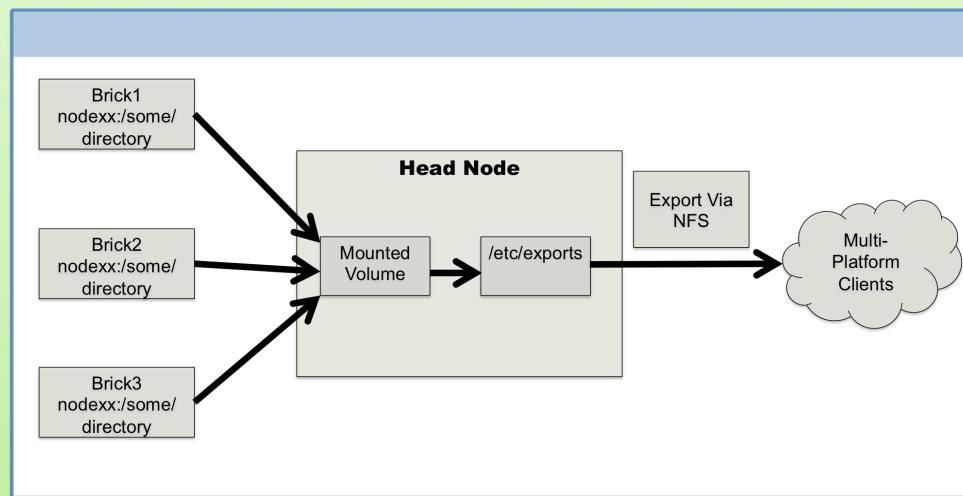
# Data Control

- GlusterFS has a built in quota service and uses POSIX ACLs for user control
- POSIX ACLs
  - Can set individual users or group permissions
- Quotas
  - Given via directory
- Both apply to Mac and windows clients (NFS/ SAMBA)



# Exporting

- GlusterFS volumes can be exported via NFSv3
  - POSIX ACLs are lost when exporting directly via NFS
  - Enable POSIX ACLs by mounting via GlusterFS and exporting via NFS
- SAMBA allows Windows users to modify NTFS permissions on files



# Data Support

- Snapshots
  - We used the rsnapshot utility to enable snapshots
  - Use cron jobs to specify snapshot intervals and locations
- Auditing
  - The auditd utility can be used in conjunction with GlusterFS
  - Shows detailed file interactions
- GlusterFS built-in logging support
  - Performance
  - Diagnostics
  - Events (Warnings, Errors, General Information)



# Administration

- GlusterFS has an intuitive CLI
  - Allows for quick volume tuning, shrinking, and expanding while the system is online and available
  - Easily integrated into current infrastructure
- Pitfalls
  - Latency induced when mounting and exporting
  - GlusterFS mounting/unmounting occasionally hung
  - Metadata is distributed, thus harder to remove



# Total Cost of Operation

- Open Source
- Can use commodity hardware
- Can use 1 or 10Gbps Ethernet



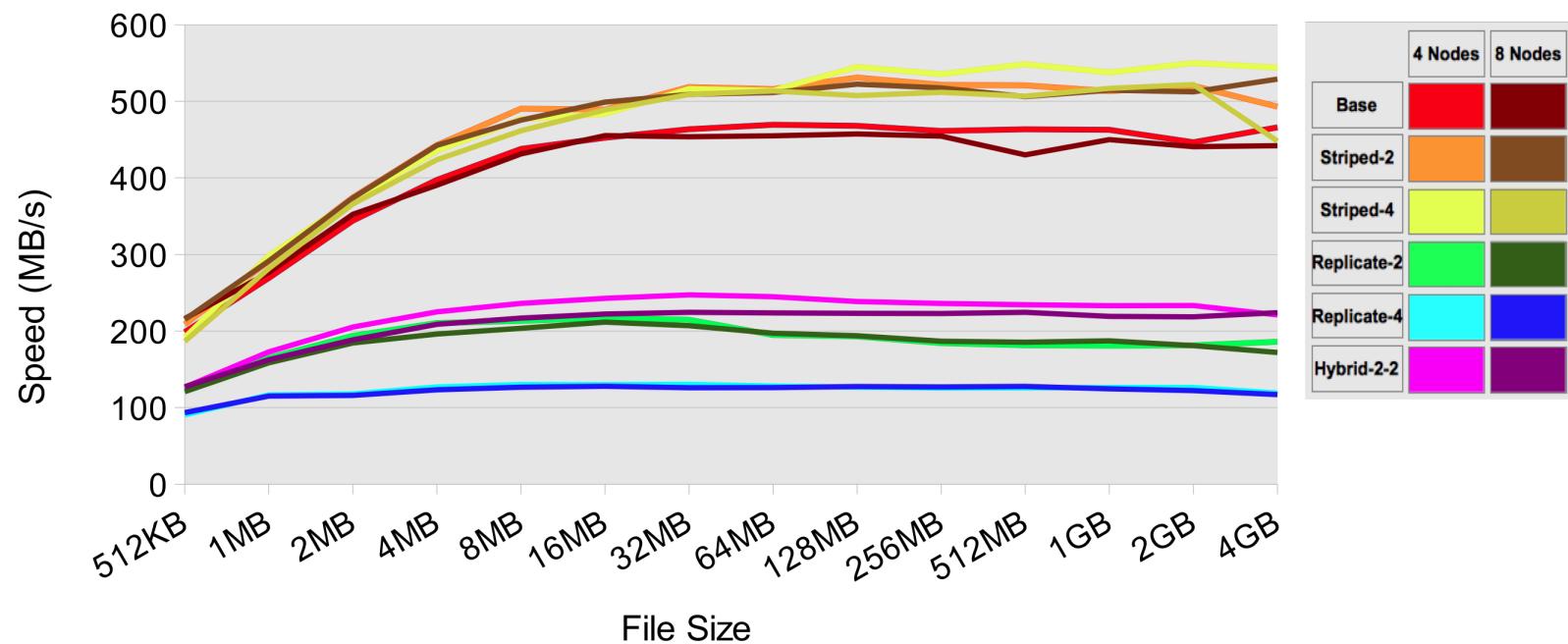
# Performance Testing Key



# Write Performance

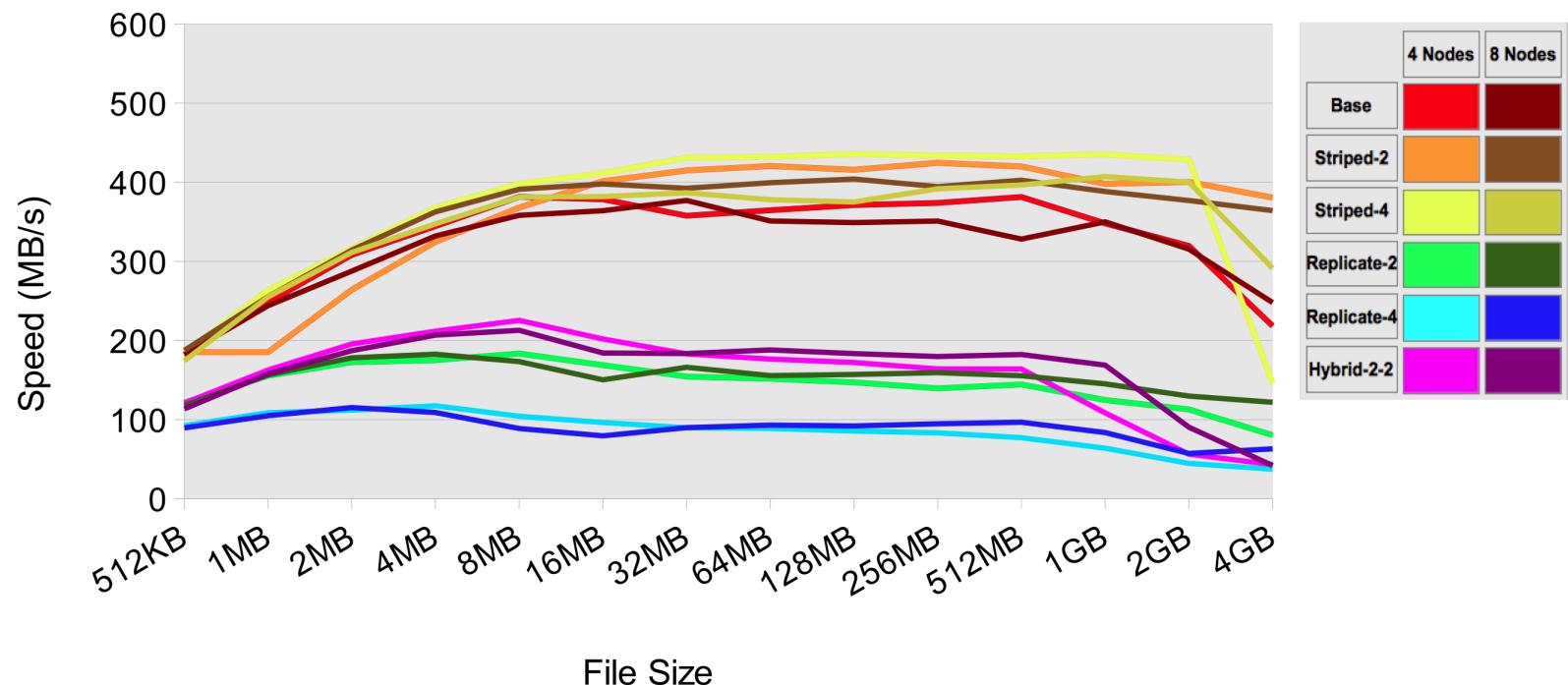
## GlusterFS Write Speeds with 1 User in Parallel

Using The dd Command



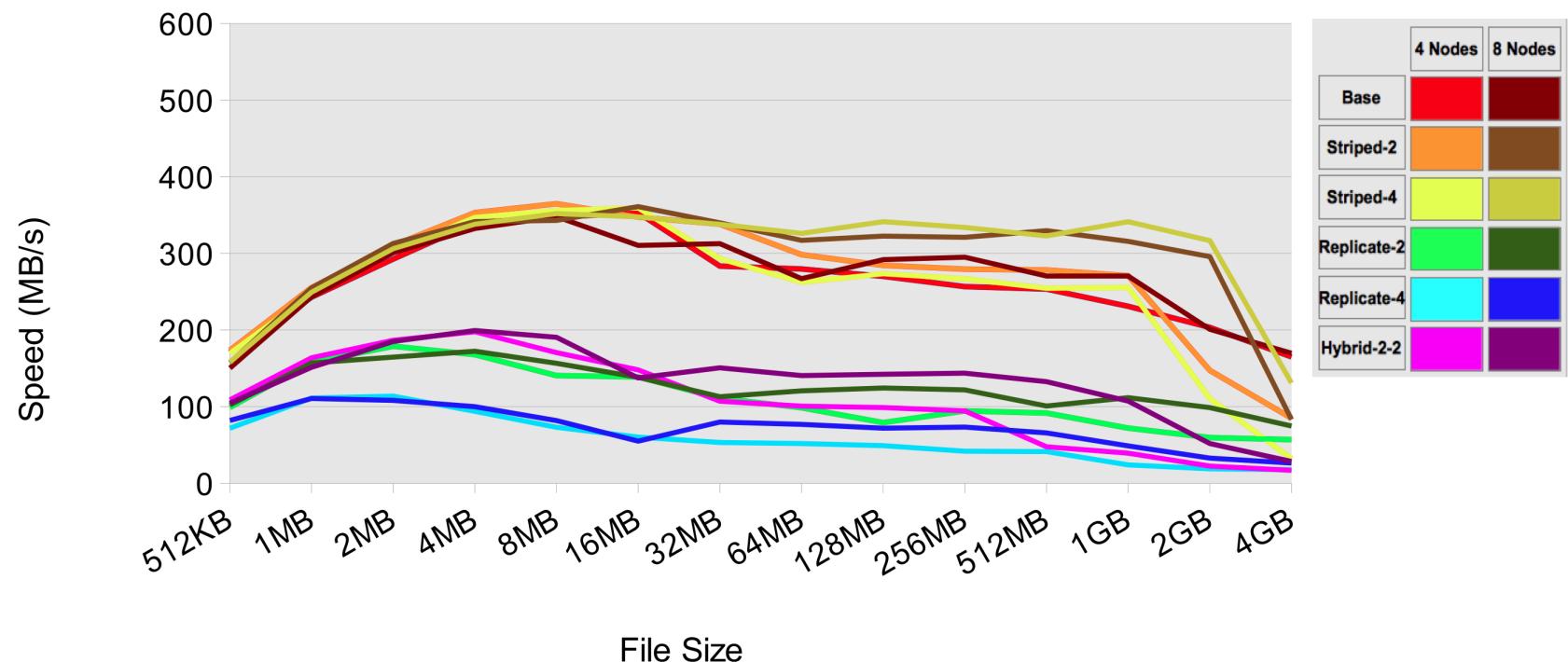
# Write Performance

GlusterFS Write Speeds with 4 Users in Parallel  
Using The dd Command



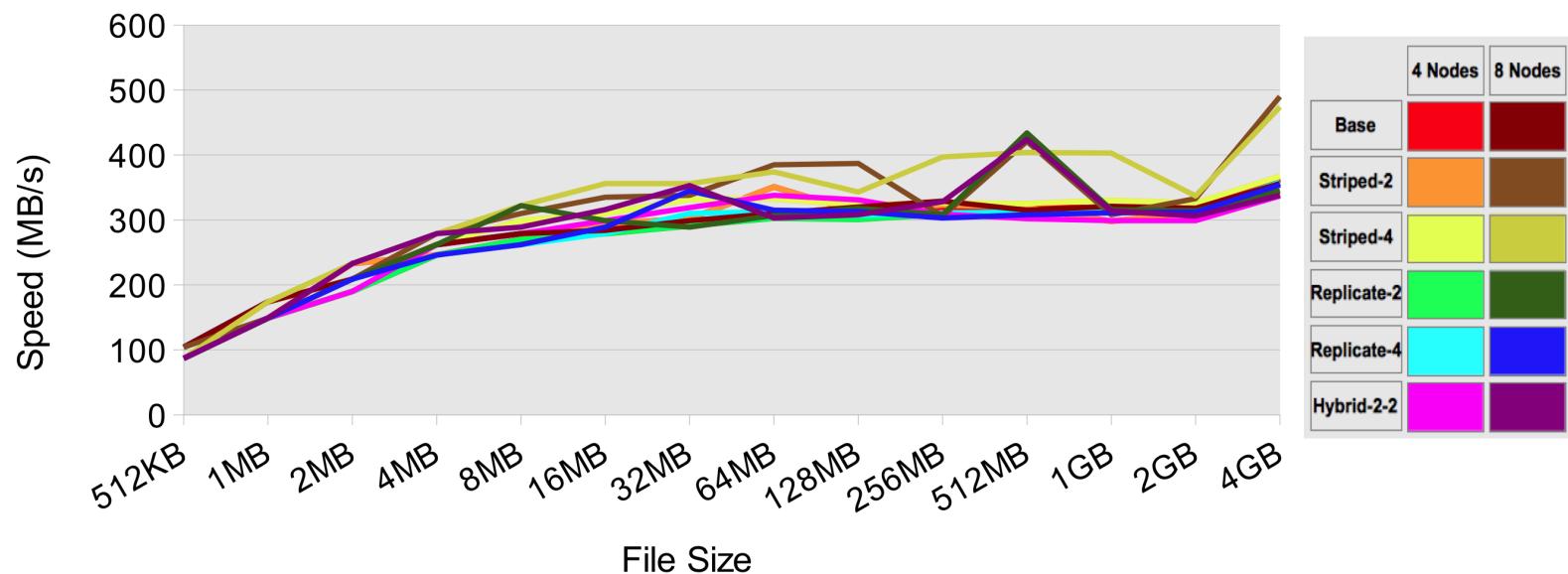
# Write Performance

GlusterFS Write Speeds With 8 Users in Parallel  
Using The dd Command



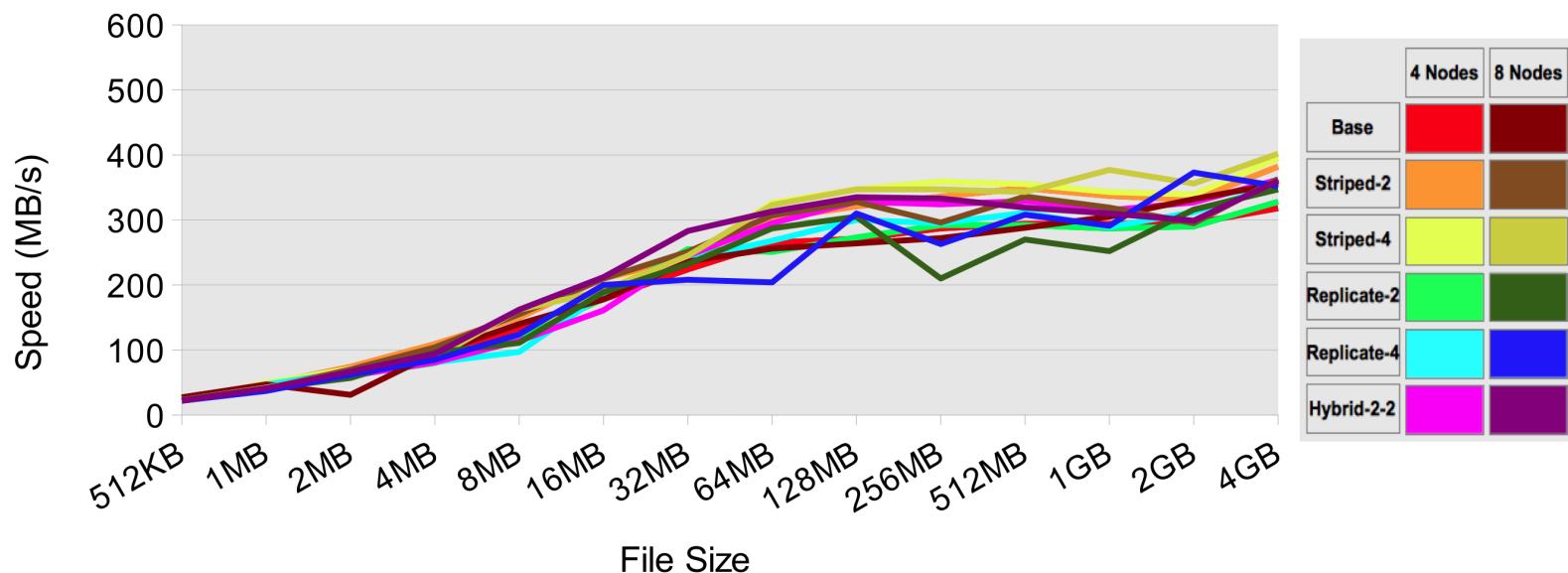
# Read Performance

GlusterFS Read Speeds with 1 User  
Using the dd Command



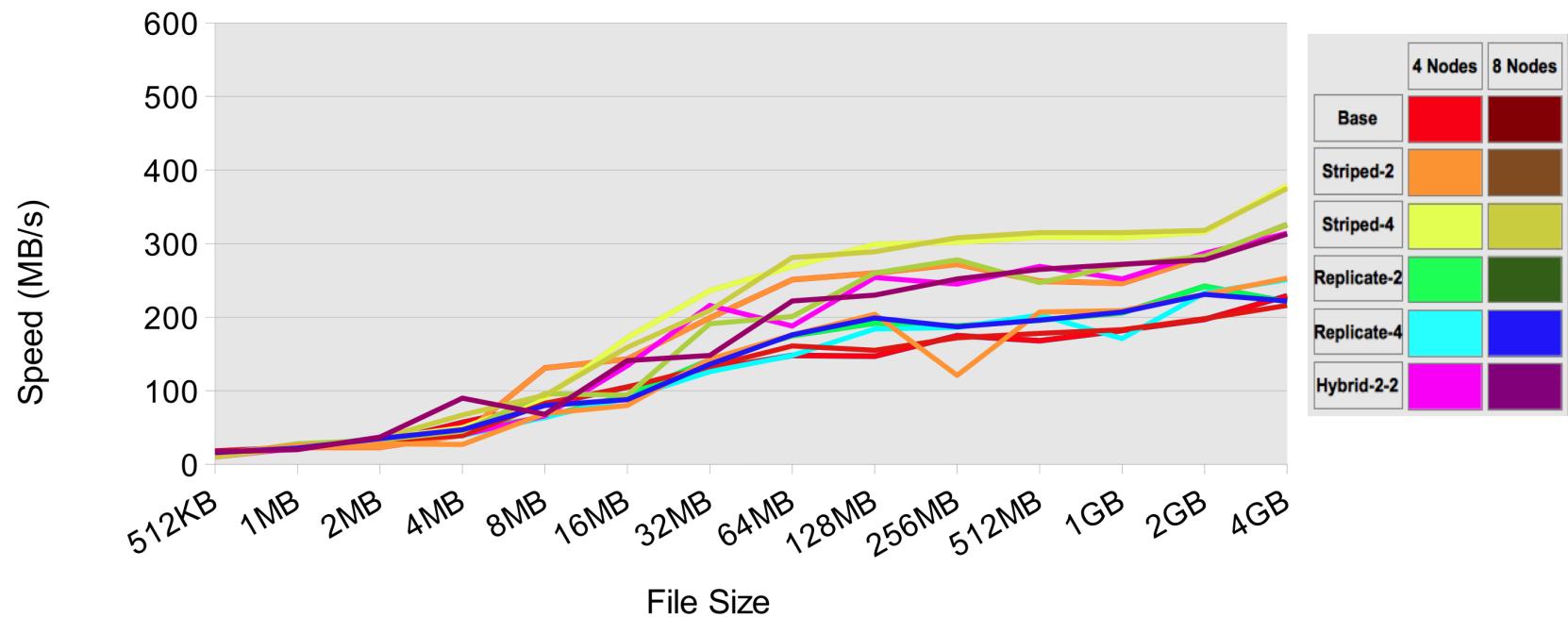
# Read Performance

GlusterFS Read Speeds with 4 Users in Parallel  
Using the dd Command



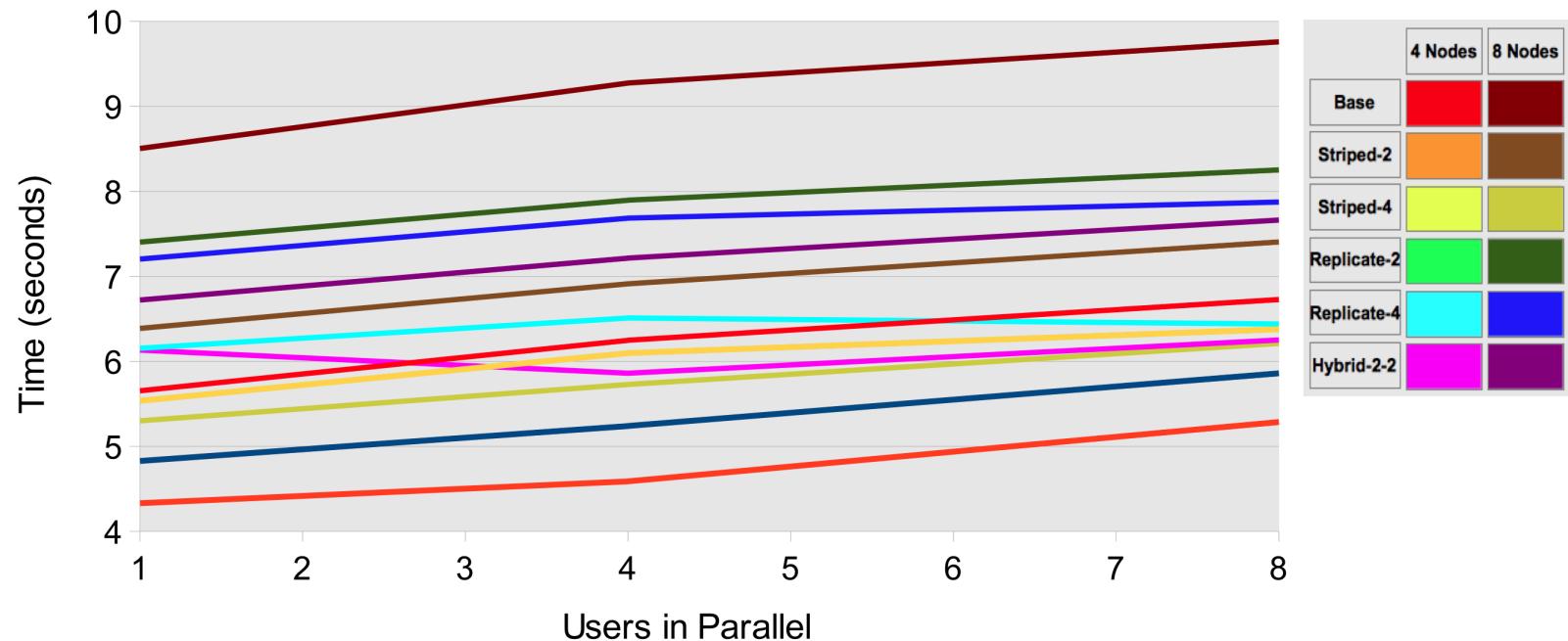
# Read Performance

## GlusterFS Read Speeds with 8 Users in Parallel Using the dd Command



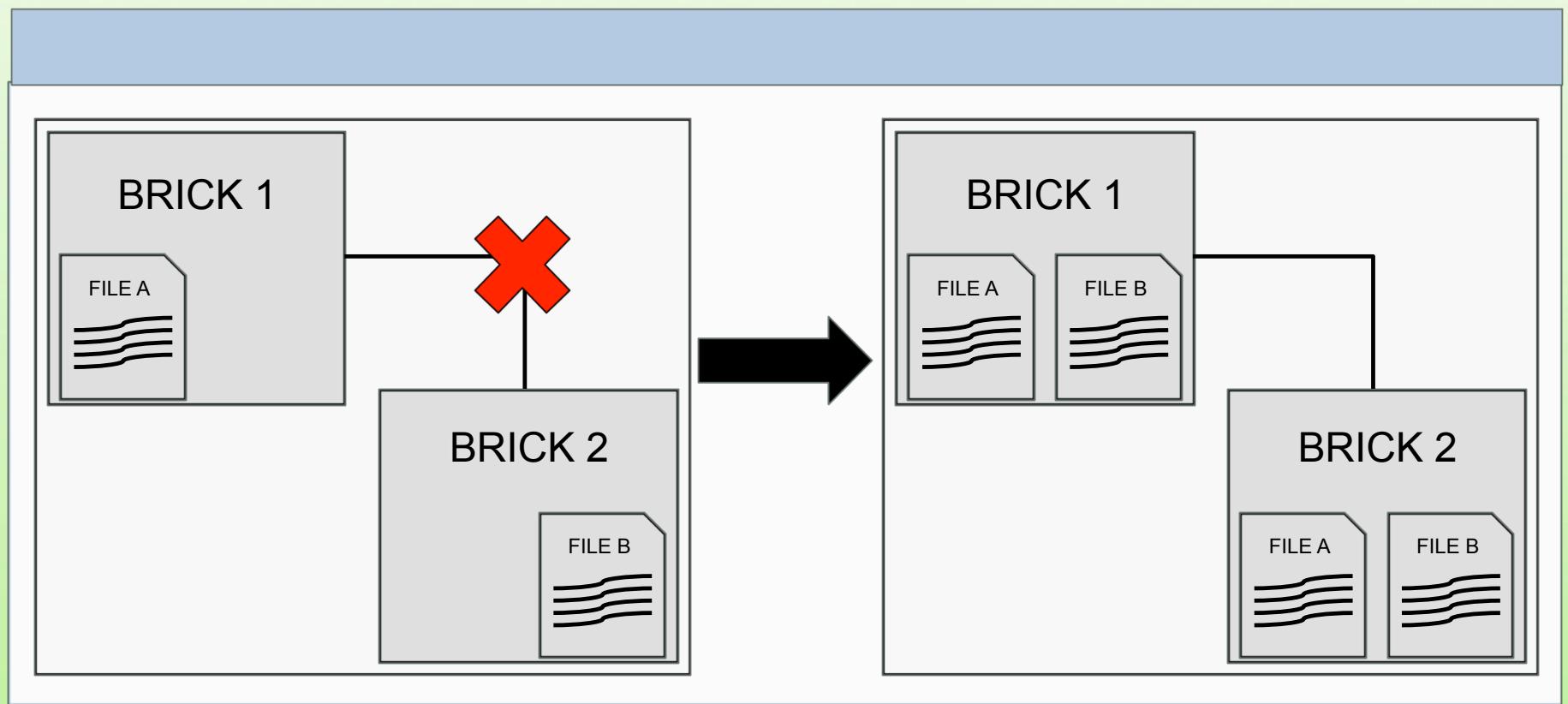
# Ls Performance

Time to ls a Recursive Directory



# Fault Tolerance

- Replicated volumes can self-heal



# Conclusions

- GlusterFS proved to have widespread capabilities as a virtual file system
- Scalability is very dependent upon the underlying hardware
- Lack of built-in encryption and security paradigm
- Best suited in a general purpose computing environment



# Future Research

- GlusterFS over Infiniband
- Geo-replication
- Unified File and Object Storage
- Apache Hadoop
- Scalability for 1000's of nodes
- Using other filesystems on top of GlusterFS
- Testing different RAID types



# Acknowledgments

- We would like to thank:
  - Los Alamos National Laboratory
  - New Mexico Consortium/PRObE
  - National Science Foundation
  - National Nuclear Security Administration
  - Gary Grider and the HPC division
  - Carol Hogsett and Josephine Olivas
  - Our mentors, David Kennel and Greg Lee
  - Our instructors, Dane Gardner and Andree Jacobson



# Questions??

