

pNFS: What and Why

59th IETF – Seoul

March 4, 2004

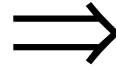
pNFS - Parallel NFS

- NFS Extensions for Parallel Storage
- Workshop @ UMich CITI, December 2003
- Explored NFSv4 extensions to meet scalability needs, including:
 - Parallel NFS request routing / file virtualization
 - Extend block or object “layout” information to clients, enabling parallel direct access

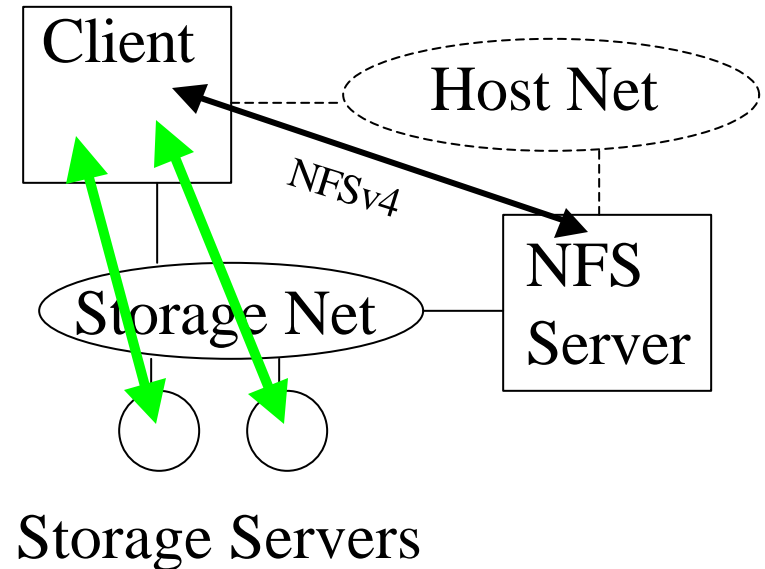
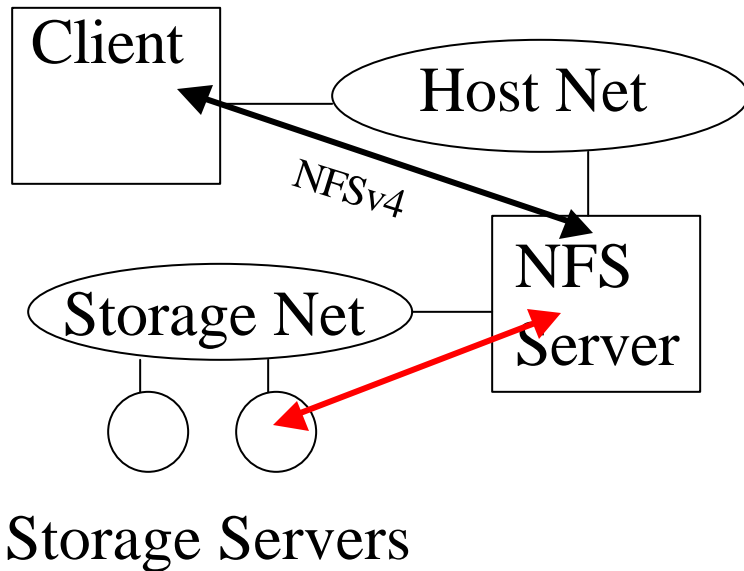
NFS Bandwidth Scalability

- NFS limitations
 - Single NFS server has limited bandwidth & CPU
 - Multiple NFS servers are disjoint storage islands
- Cluster computing driver
 - Through parallelism, cluster demand far outstrips traditional NFS server
- Goal
 - Harness multiple storage servers to provide parallel access to single files and filesystems

Now



Goal



- **Now**: requested data moves through NFS server
- **Goal**: reply from NFS server enables parallel access to diverse storage servers (NFS, iSCSI, FC, ...)

Example Applications

- High Performance Computing (HPC) clusters
- Seismic data processing
- Rendering farms
- Data mining
- Life Sciences
- Compile farms
- Electronic Design Automation

Today's Proprietary Solutions

- SAN (IBM, EMC, Sun) and Object (IBM, HP, Panasas) file systems:
 - Allow clients to bypass metadata server to access data, but
 - ⇒ Each system is doing this differently with little or no standardization
 - ⇒ Solutions have to provide kernel code in diverse OS's
- Scalable/cluster NFS servers:
 - Are standards compliant and facilitate scalable aggregate load, but
 - ⇒ Don't provide scalable/parallel bandwidth to a single file

Why work on pNFS in the IETF?

- Cluster/grid computing is a growing trend
 - HPC applications strong growth market
 - Data center technology gearing up for more cluster/grid use
- Increasingly, storage data is transported in IETF protocols
 - NFS, iSCSI, NFS-RDMA
- NFS is best candidate for a scalable bandwidth standard
 - Wide acceptance and use of NFS, integrated security
 - NFS interoperability track record
 - NFSv4 has mechanisms for clients to act more autonomously
 - NFSv4 minor versioning may suffice for incremental functionality
- A standards venue for interoperable efforts
 - Strengthen NFS as most useful open standard file system
 - NFS and iSCSI symbiosis

Community Interest

- December 2003 U. Michigan workshop: Call for ideas
 - Whitepapers at www.citi.umich.edu/NEPS
 - U. Michigan CITI, EMC, IBM, Johns Hopkins Univ., Network Appliance, Panasas, Spinnaker Networks, Veritas, ZForce
- Problem Statement Internet-Draft published
 - `draft-gibson-pnfs-problem-statement-00.txt`
- Upcoming BOF at Usenix FAST '04

Summary

- **Bandwidth Scalability Problem:**
 - Data intensive apps & clusters rapidly scaling up demand
- **Current state of the art:**
 - Clustered NFS today inadequate to scale single file and file system bandwidth
 - Non-NFS file systems allow clients direct access to storage device, but are proprietary extensions
- **Desired solution:**
 - NFSv4 extensions enabling clients to access data in parallel storage devices of multiple standard flavors