



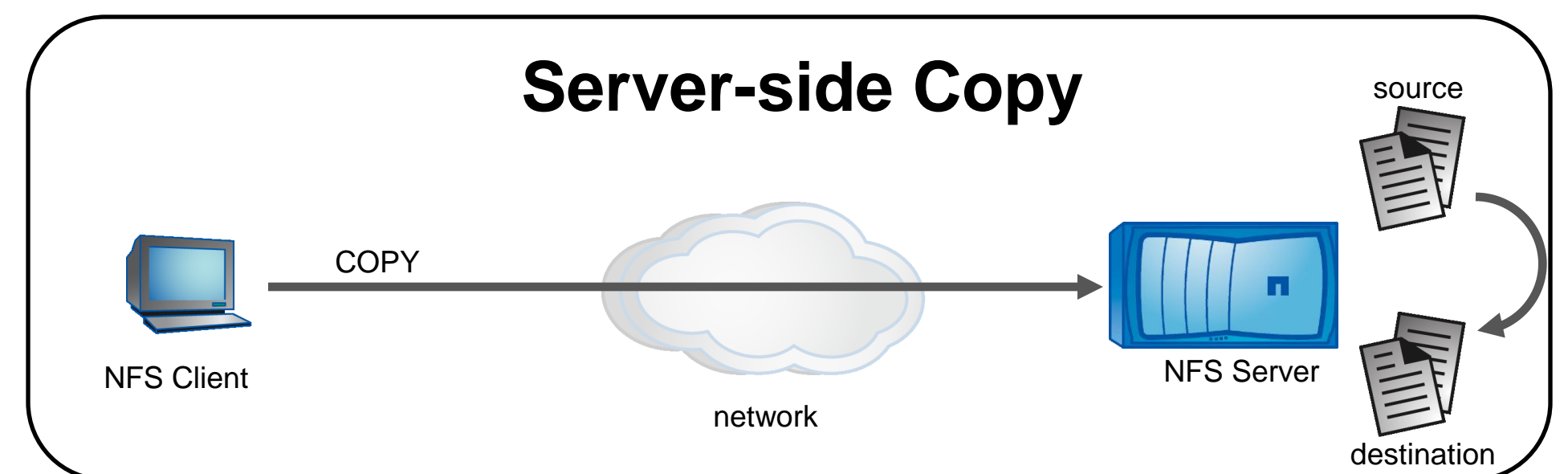
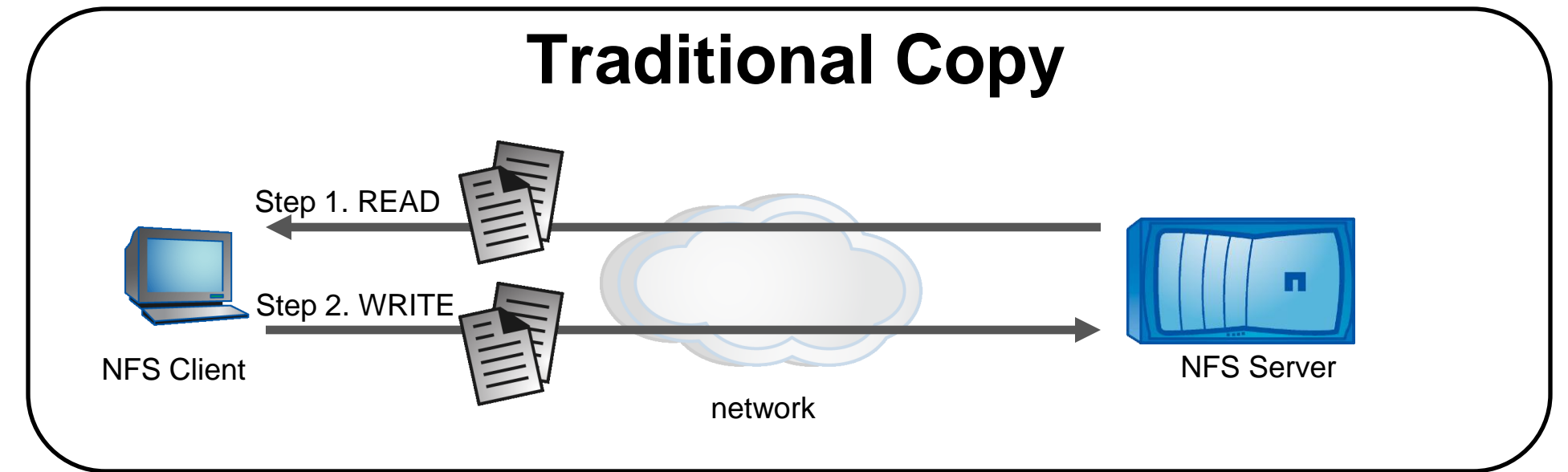
Accelerating NFS with Server-side Copy

James Lentini, Anshul Madan, Trond Myklebust

NetApp™

Motivation

- Traditionally an NFS client copies a file by
 - Reading from a source file on the server
 - Writing to a destination file on the server
- Server-side Copy** allows an NFS client to efficiently copy a file by sending a single message.
 - Saves client resources** processing network packets, context switching in/out of application, and copying data.
 - Saves network resources** by decreasing traffic on the network
- Server-side copy is a proposed NFSv4.2 feature



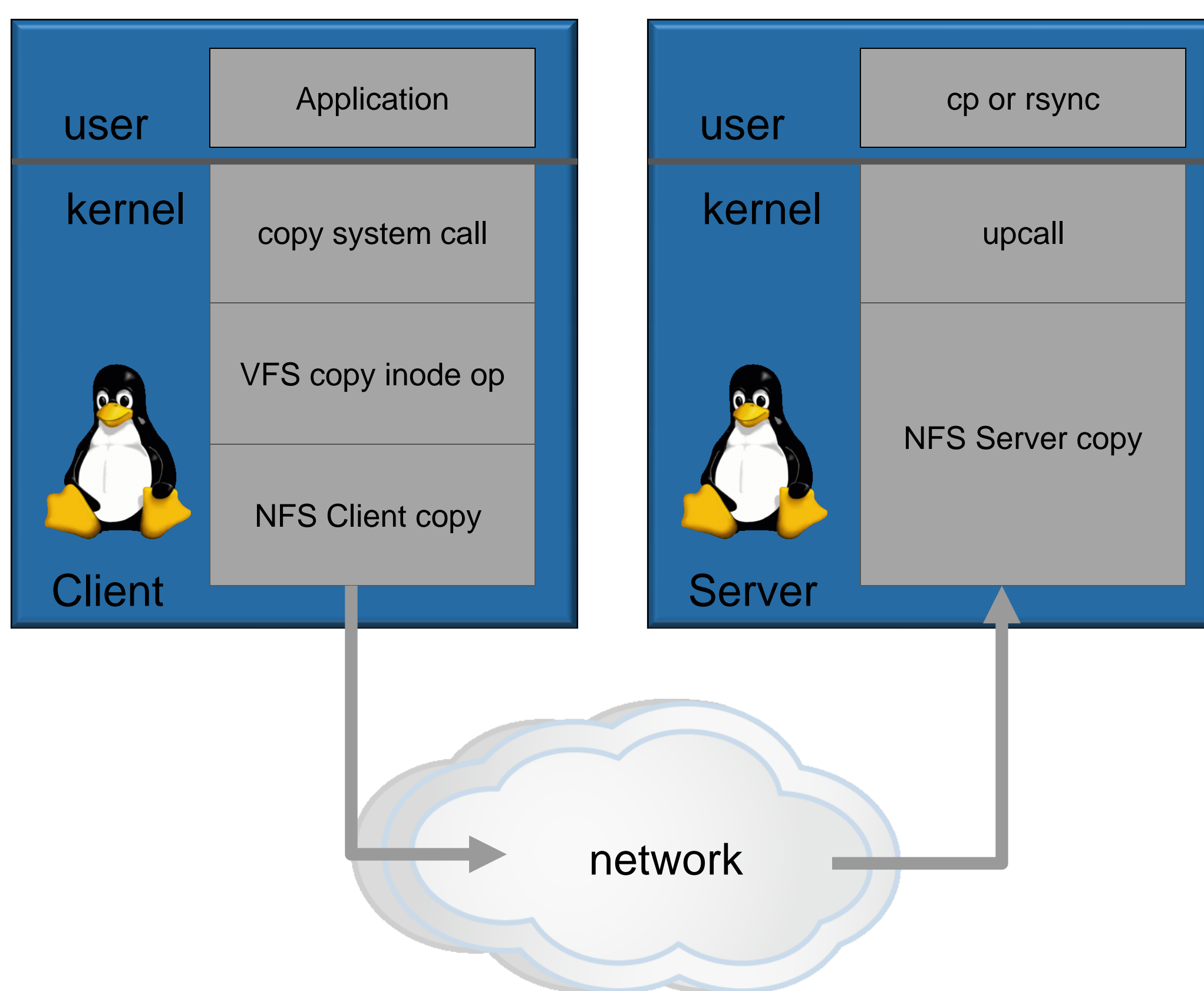
Applications

- Virtualized Environments:** allows a hypervisor to efficiently **backup**, **clone**, or **migrate** a VM's virtual disk when it is an NFS file
- File Restore:** the contents of a backup can be copied into the active file system
- Zero Copy Clone:** exposes time and space efficient file clone operation, requires support in the server's filesystem
- Network Write Deduplication:** allows a client-side cache to eliminate write operations
- And, in general, anytime data is copied from one location to another.

Protocol Design

- Supports regular files, not directories
 - Simplifies the protocol, no recursive directory copies
 - Directory copies can be synthesized using multiple directory creates and file copies
- Synchronous and asynchronous operating modes
 - Asynchronous mode for long duration copies
 - Server decides which mode to use
- Copies on a single server or between two servers
 - Inter-server copy pulls data from source to destination
- Supports partial (sub-regions of file), guarded (exclusive create), and metadata (file attributes) copies
- Secured using RPCSEC_GSSv3

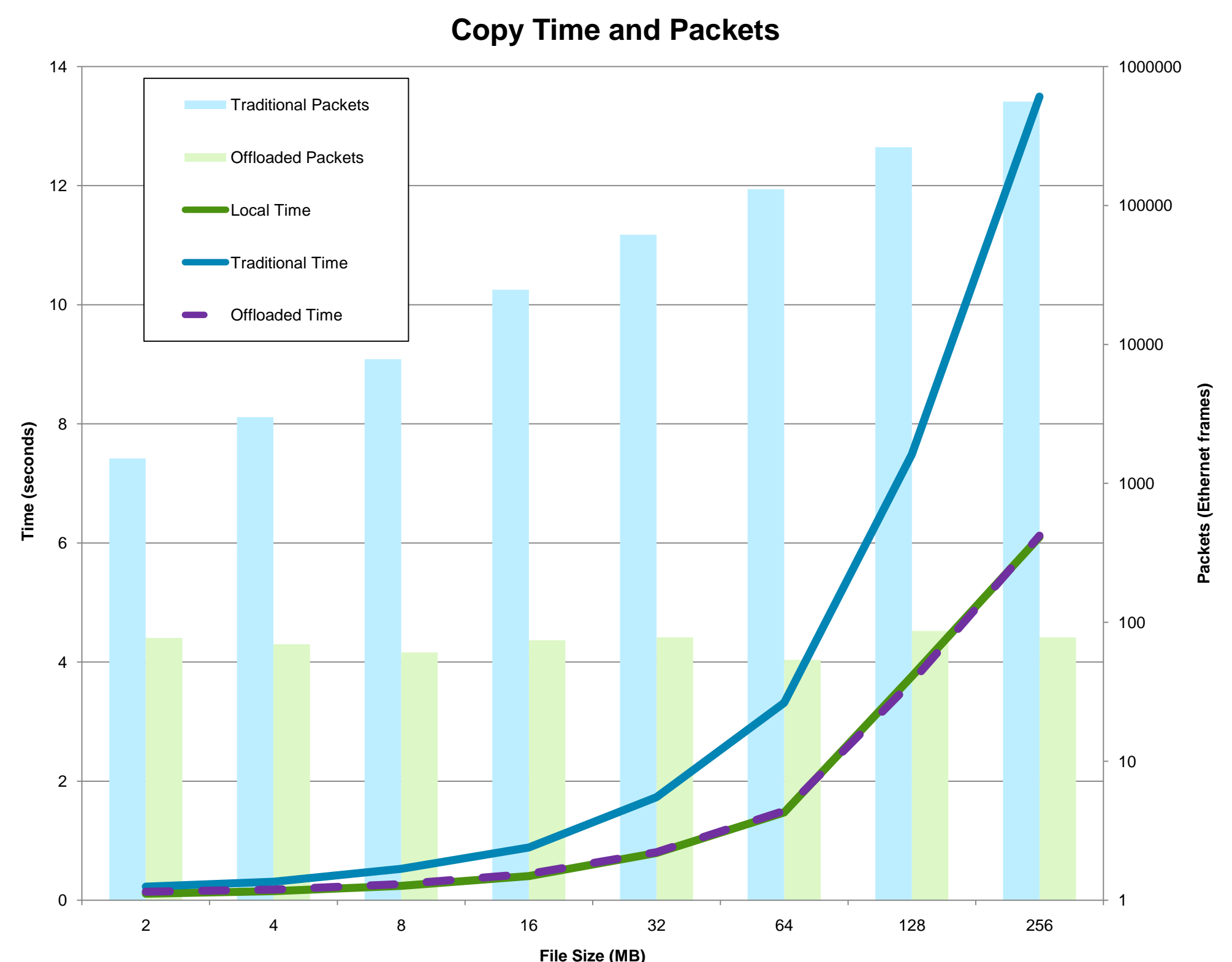
Implementation



- Modified Linux 2.6.34. Added:
 - Synchronous copyfileat() system call
 - vfs_copy() inode operation to VFS layer
 - NFS client and server COPY RPC
- NFS protocol implementation supports whole file, synchronous mode, intra-server copy

Results

Reduces network traffic, CPU interrupts, kernel space CPU execution time, and client memory cache.



- Client/Server: dedicated, point-to-point 1 Gbps network, EXT4 export, dual core 1.8 Ghz CPUs, 4 GB RAM
- Local Copy: cp command on server
- Traditional NFS Copy: cp command on client
- Offloaded NFS Copy: copyfileat() on client