Architectural Considerations and Performance Evaluations Of Shared Storage Area Networks At NASA Goddard Space Flight Center

Hoot Thompson, Curt Tilmes, Robert Cavey, Bill Fink, Paul Lang and Ben Kobler

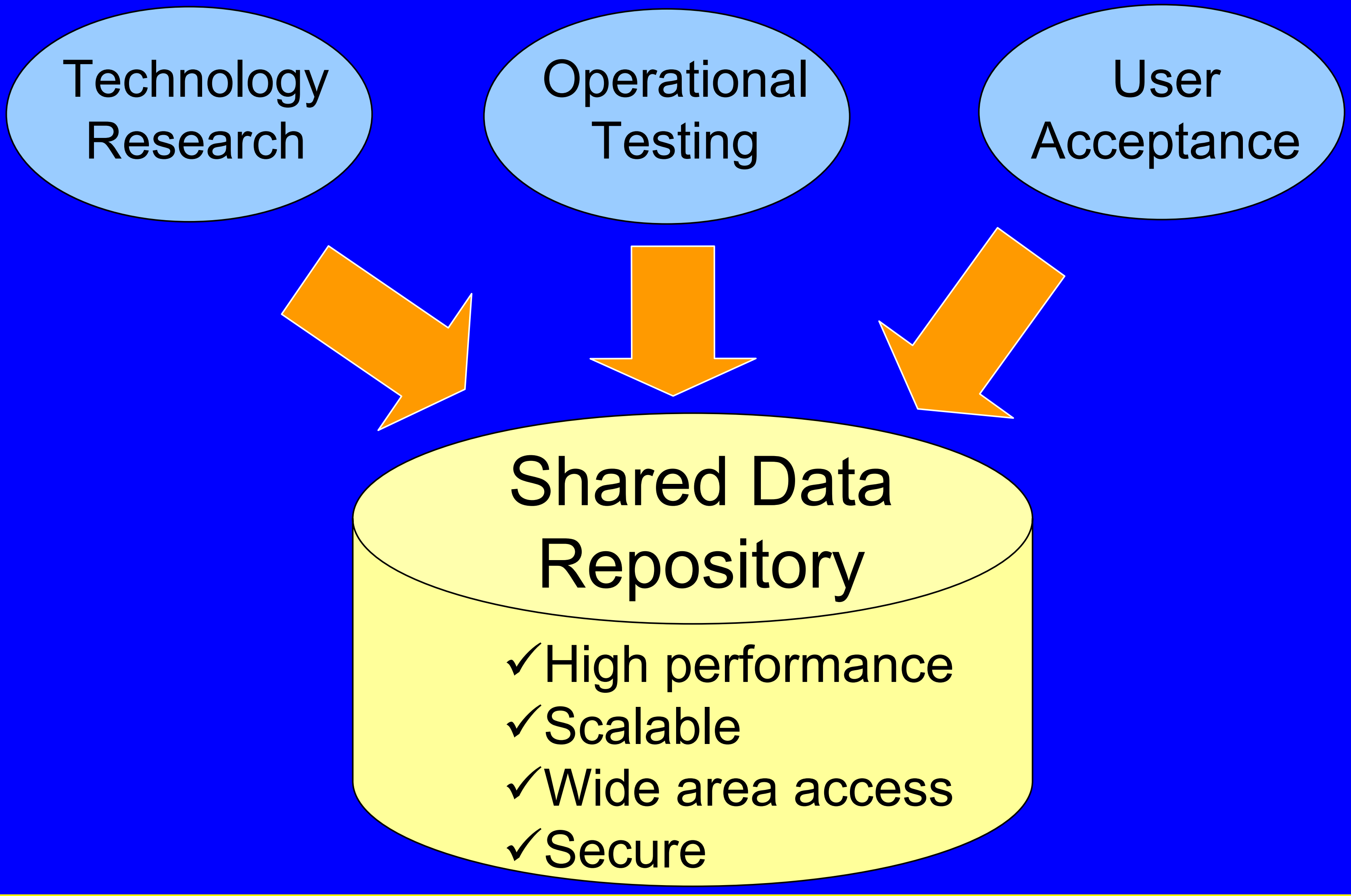
hoot@ptpnow.com
Patuxent Technology Partners, LLC
www.ptpnow.com

The Twentieth IEEE/Eleventh NASA Goddard Conference
on Mass Storage Systems and Technologies

Introduction – Some Trends

- Storage consolidation continues to be a popular architectural construct – Storage Area Networks (SAN)
- Minimizing costs while expanding storage capacity and providing shared access is a driver
- Fibre Channel connected storage remains dominant at least in the high performance space
- Existing IP infrastructures versus ‘still’ expensive Fibre Channel represent a potentially attractive mechanism for connecting users to storage at the block level

GSFC SAN Pilot Initiative



Technology Research

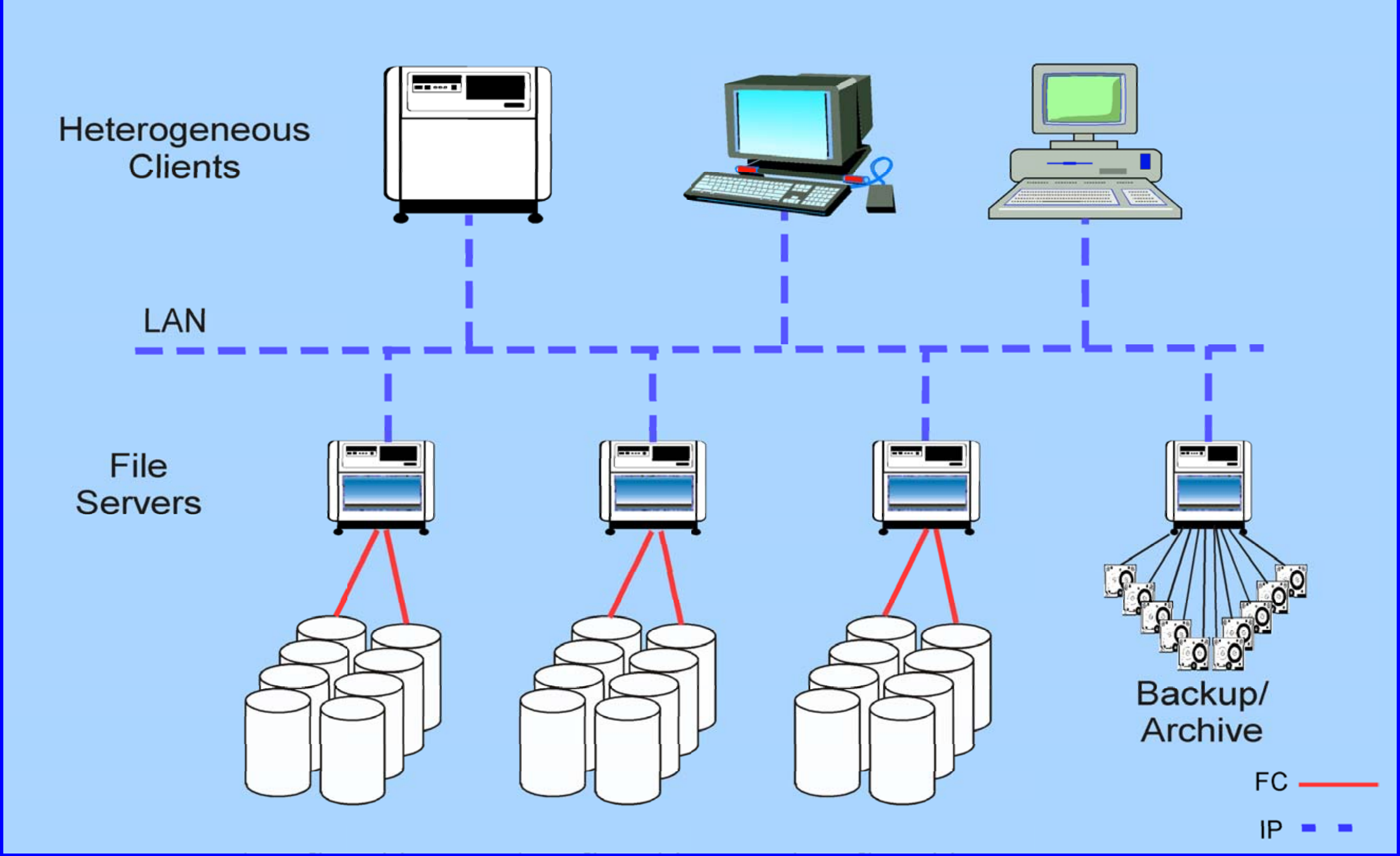
Operational Testing

User Acceptance

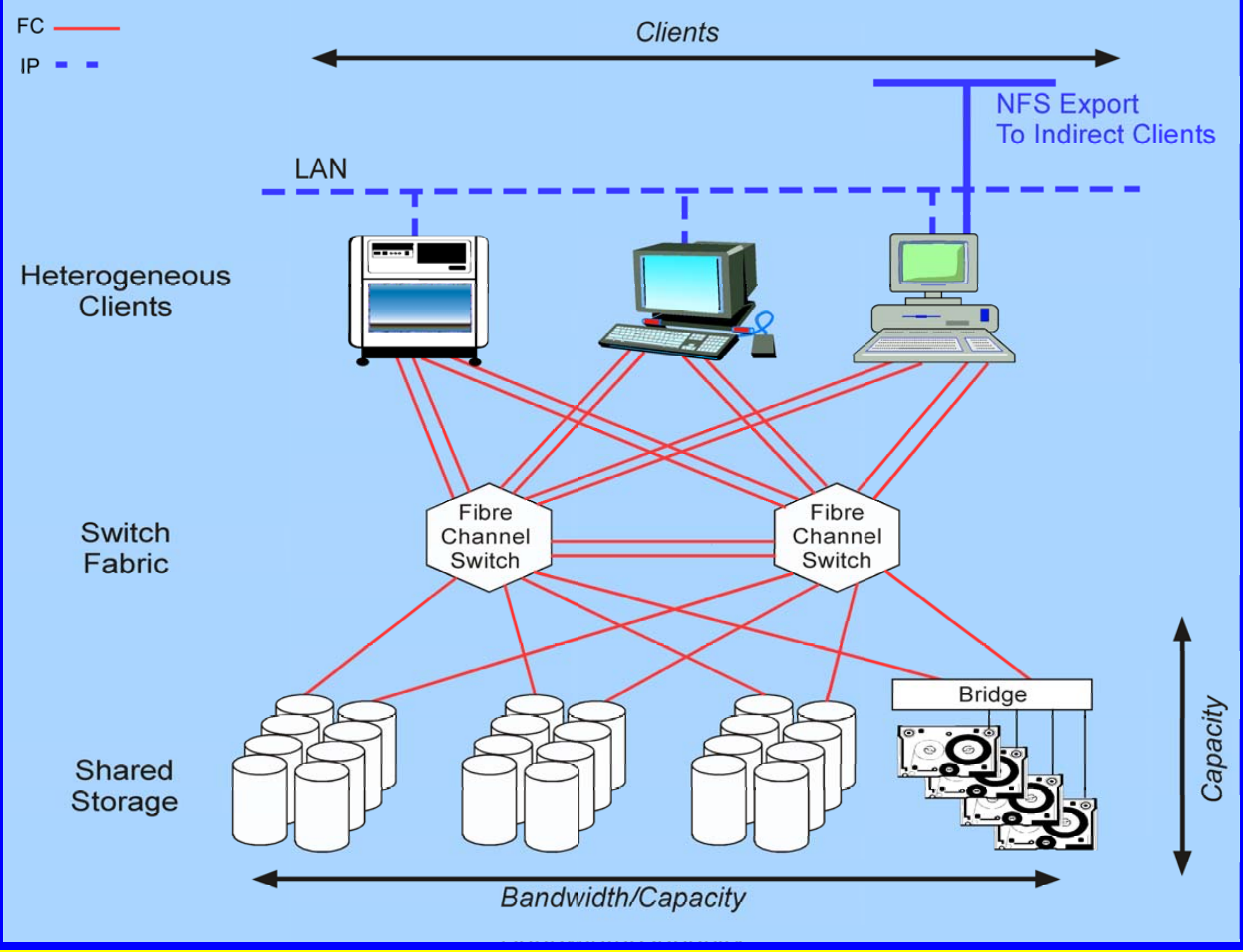
Shared Data Repository

- ✓High performance
- ✓Scalable
- ✓Wide area access
- ✓Secure

By Way of Review –Traditional Infrastructure



Storage Area Network Infrastructure



Some Definitions

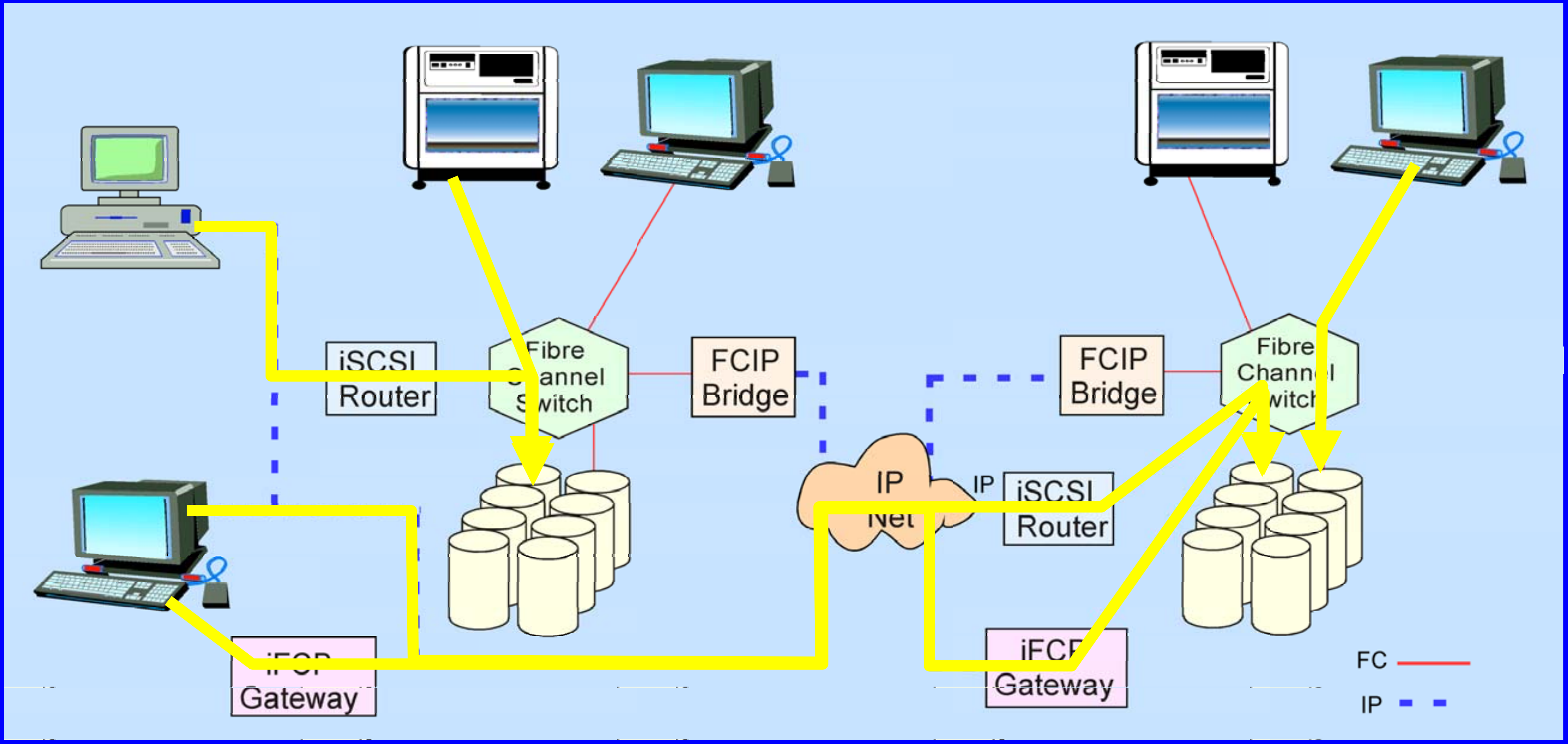
- Fibre Channel
  - Industry standard high-speed SCSI transport technology
  - 1 or 2 Gbit/sec, 10 Gbit/sec coming
- Internet SCSI (iSCSI)
  - *"represents a light switch approach to storage networking"*
- Fibre Channel Over IP (FCIP)
  - *"means of encapsulating Fibre Channel frames within TCP/IP specifically for linking Fibre Channel SANs over wide areas"*
- Internet Fibre Channel Protocol (iFCP)
  - *"gateway-to-gateway protocol for providing Fibre Channel fabric services to Fibre Channel end devices over a TCP/IP network"*

Notes:

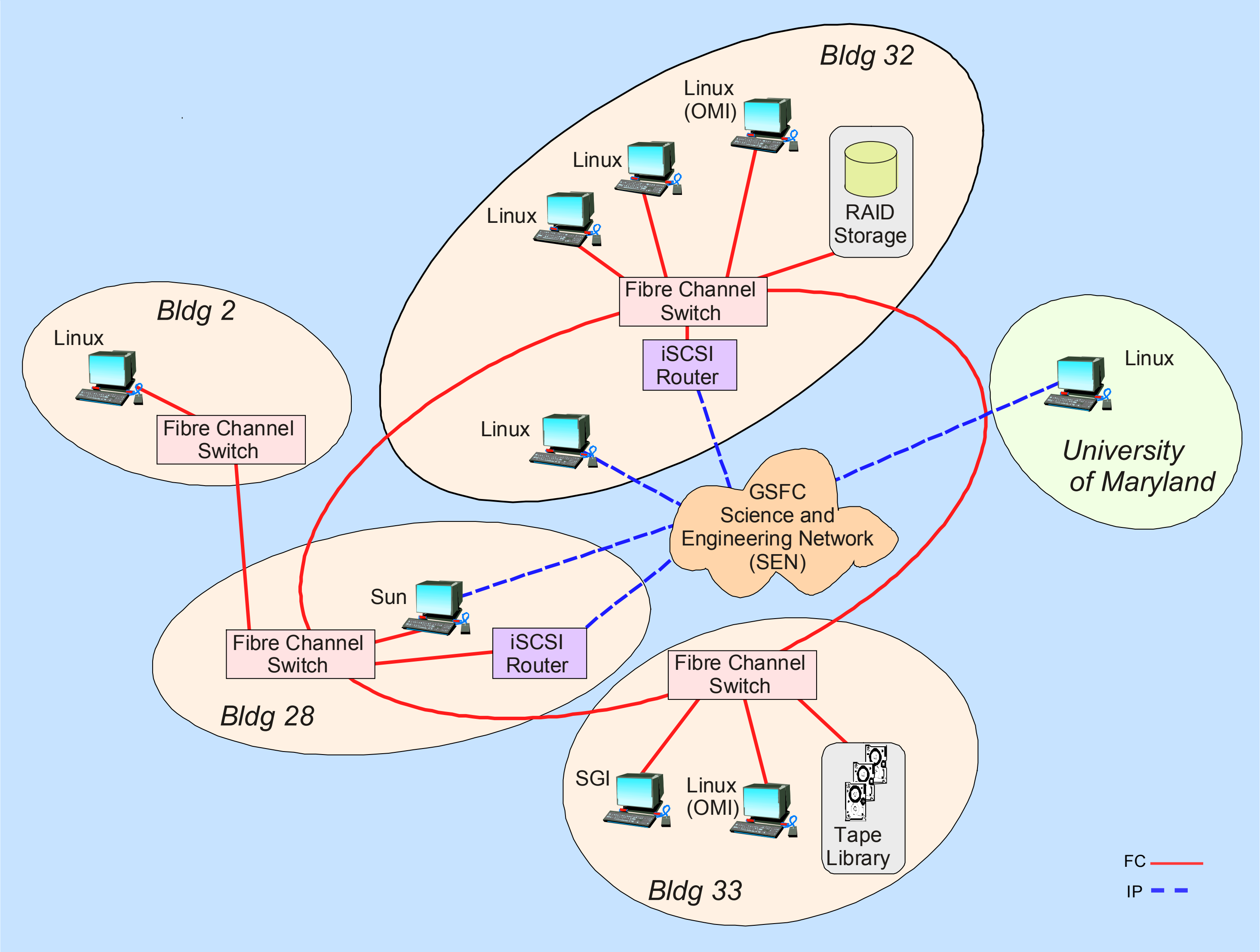
➢Standards largely the work of Internet Engineering Task Force (IETF) – www.ietf.org

➢Definitions from "IP SANs – A Guide to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks" by Tom Clark

IP-Based SAN Connections



GSFC Pilot SAN Configuration



GSFC IP Testing

- iSCSI in comparison to native Fibre Channel
- Primarily Linux based machines
  - Sun platform also ‘looked at’
  - SGI currently does not support iSCSI
  - Windows® not the platform of choice
- TCP off-load engine (TOE) card
  - Evaluated on Windows machine

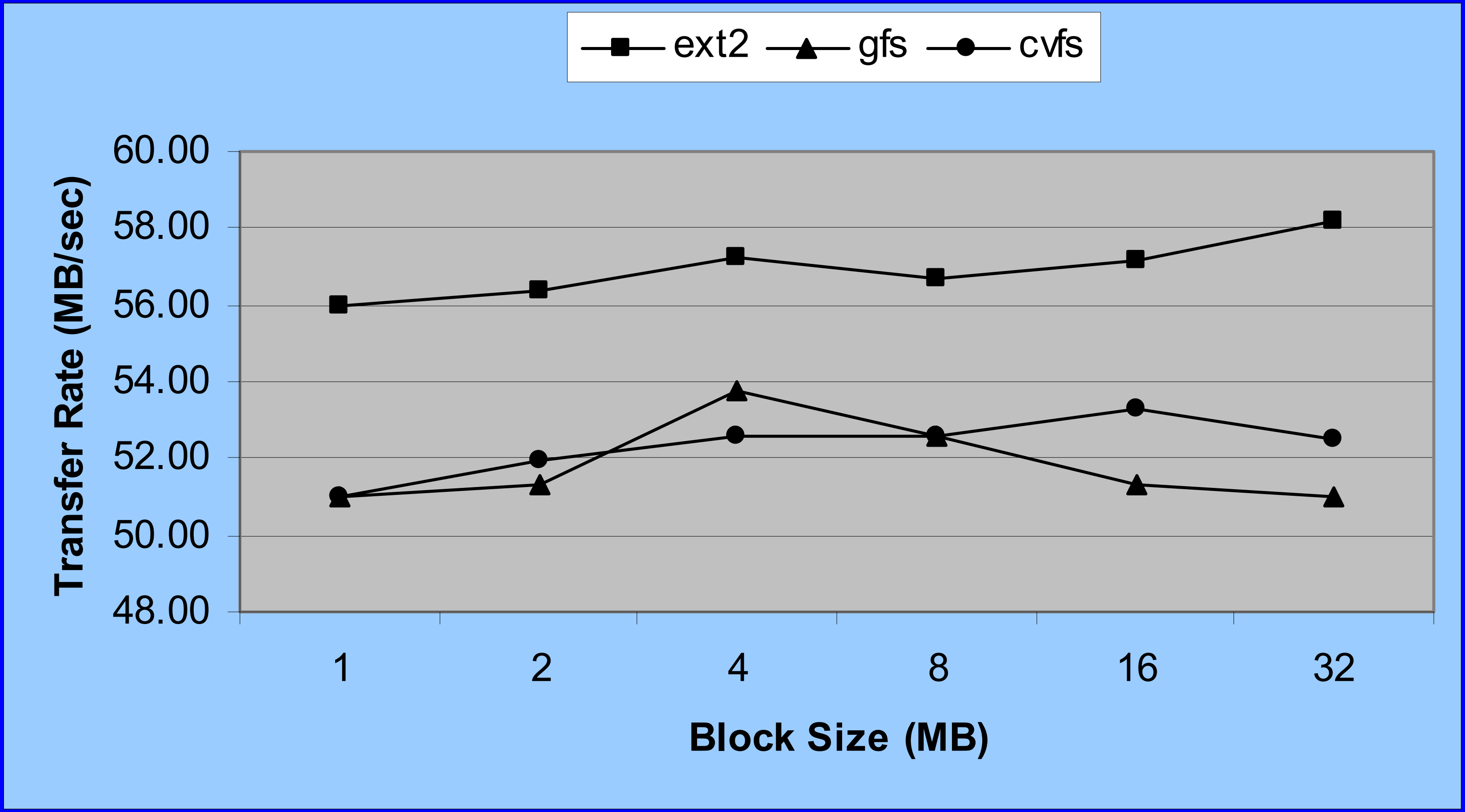
Benchmarks

- Large file, large block sequential transfers
  - *lmdd (http://www.bitmover.com/lmbench/lmdd.8.html)*
- Small file, transaction oriented tests
  - *bonnie++ (http://www.coker.com.au/bonnie++)*
  - *Postmark (http://www.netapp.com/tech_library/3022.html)*
- Application testing
  - Composite generation from MODIS data by U of MD staff
- Different file systems
  - ext2fs (native Linux file system)
  - Global File System (GFS, Sistina)
  - CentraVision™ File System (CVFS) now StorNext File System (ADIC)

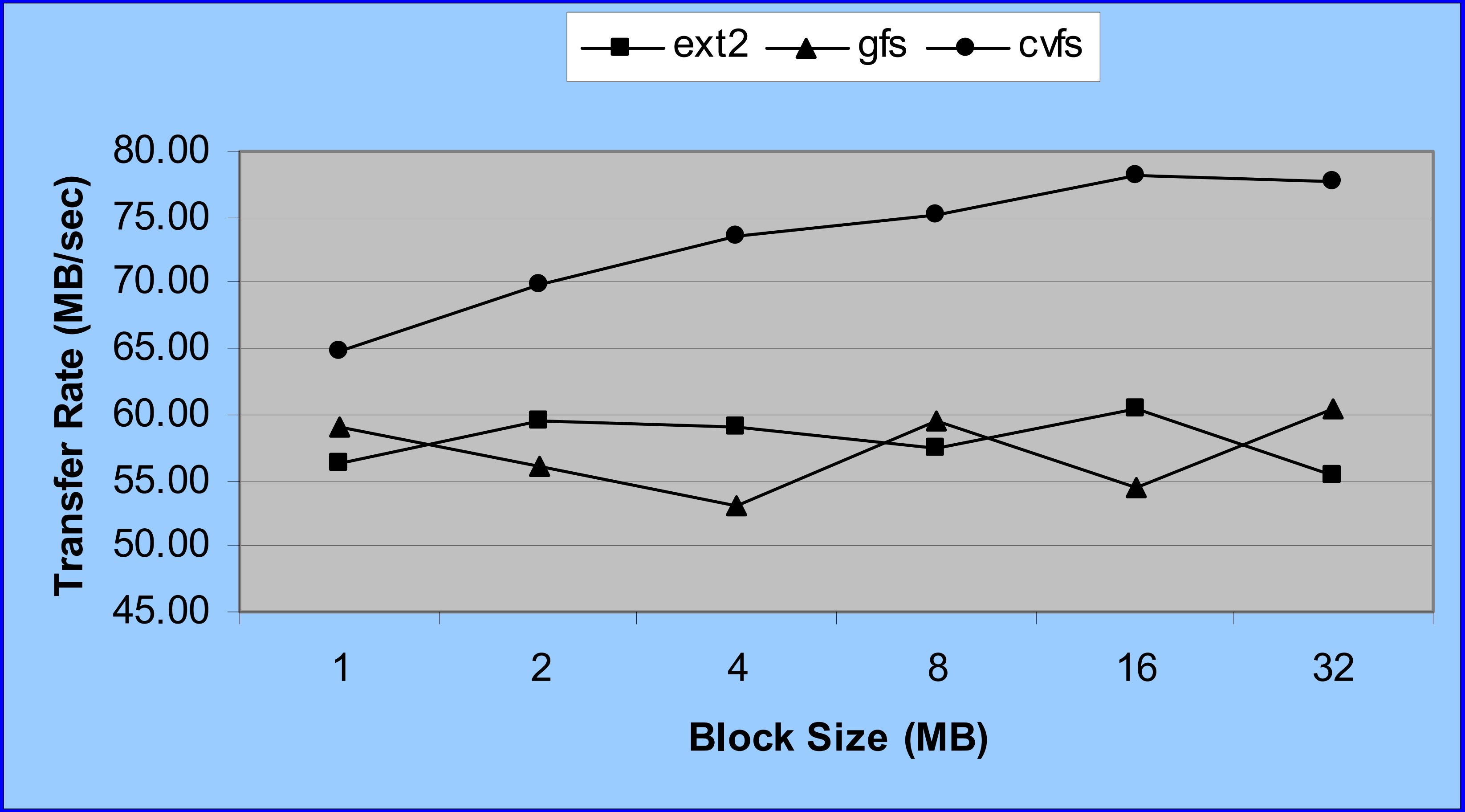
Notes:

➢Benchmark numbers for the most part are 'out of the box' results and should be viewed as representative not definitive.

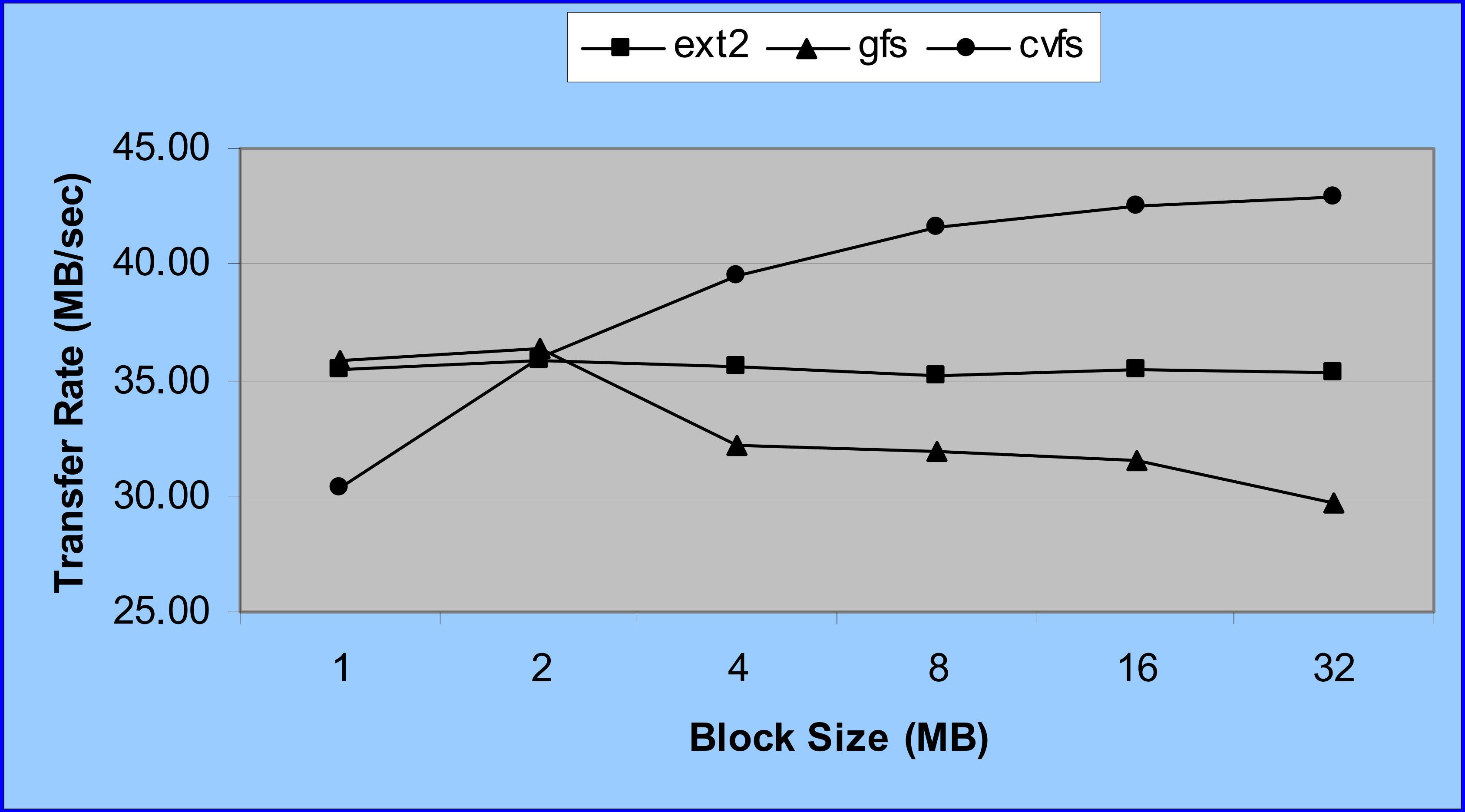
Linux Fibre Channel Connected – Writes



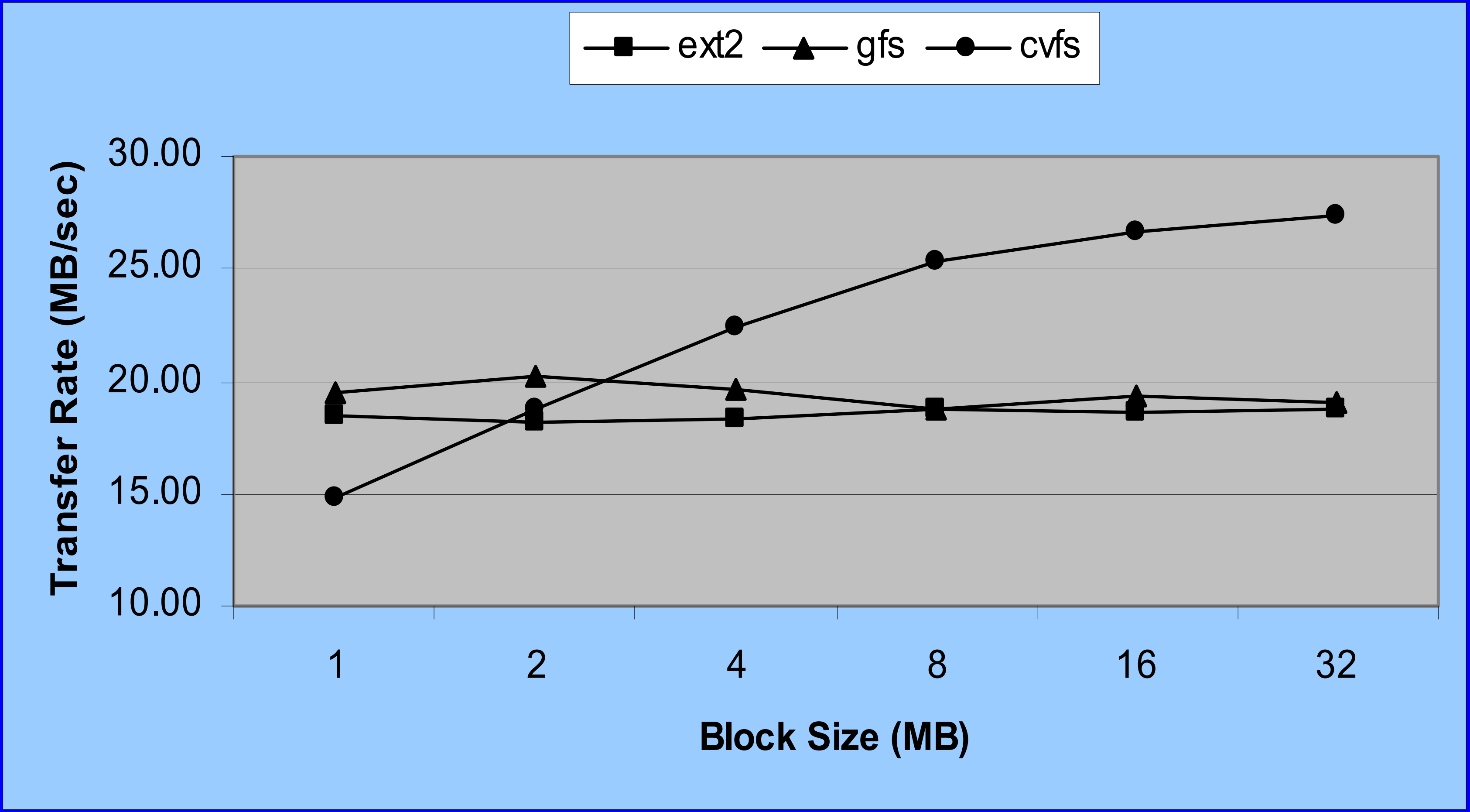
Linux FC Connected – Reads



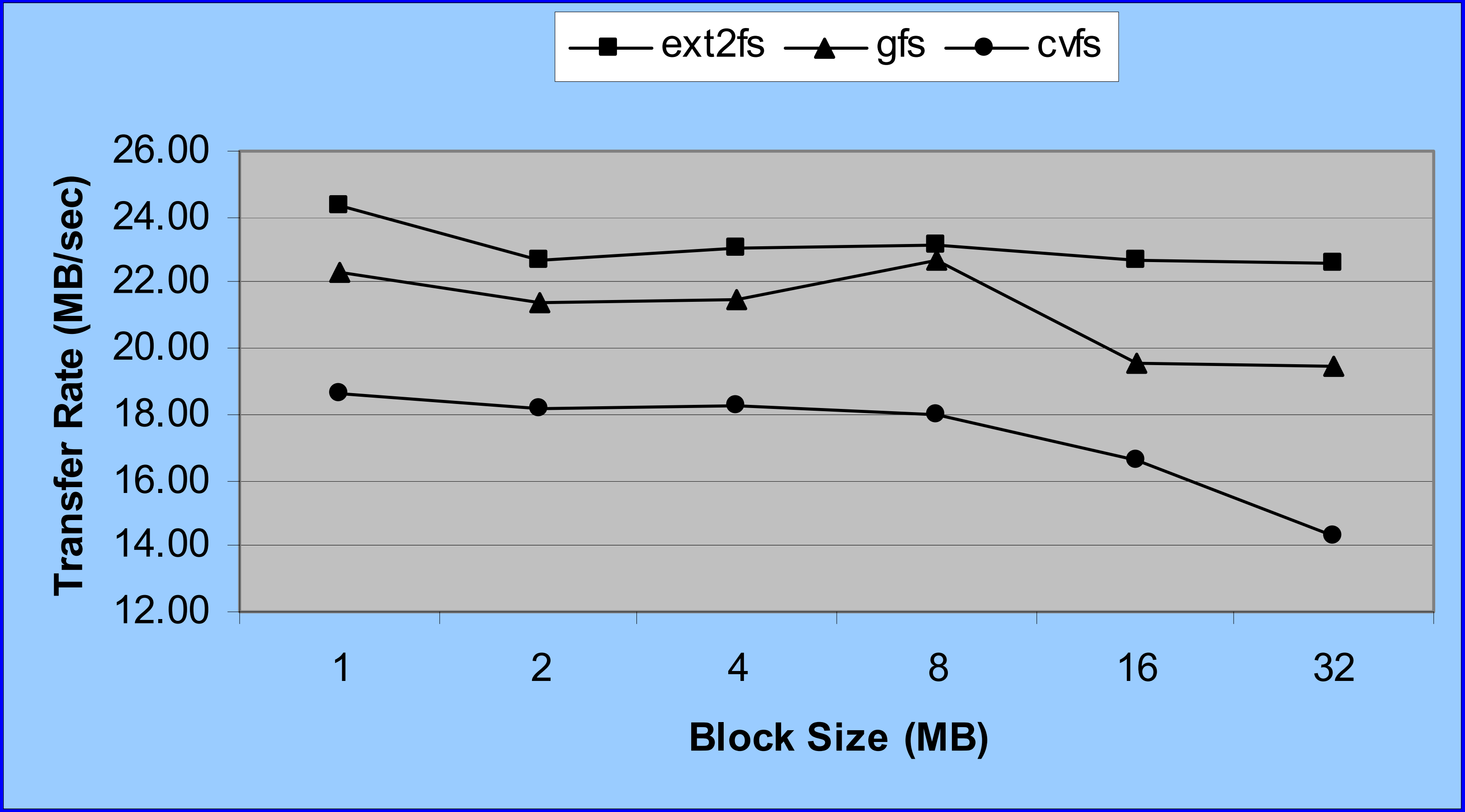
Linux iSCSI Connected – Writes



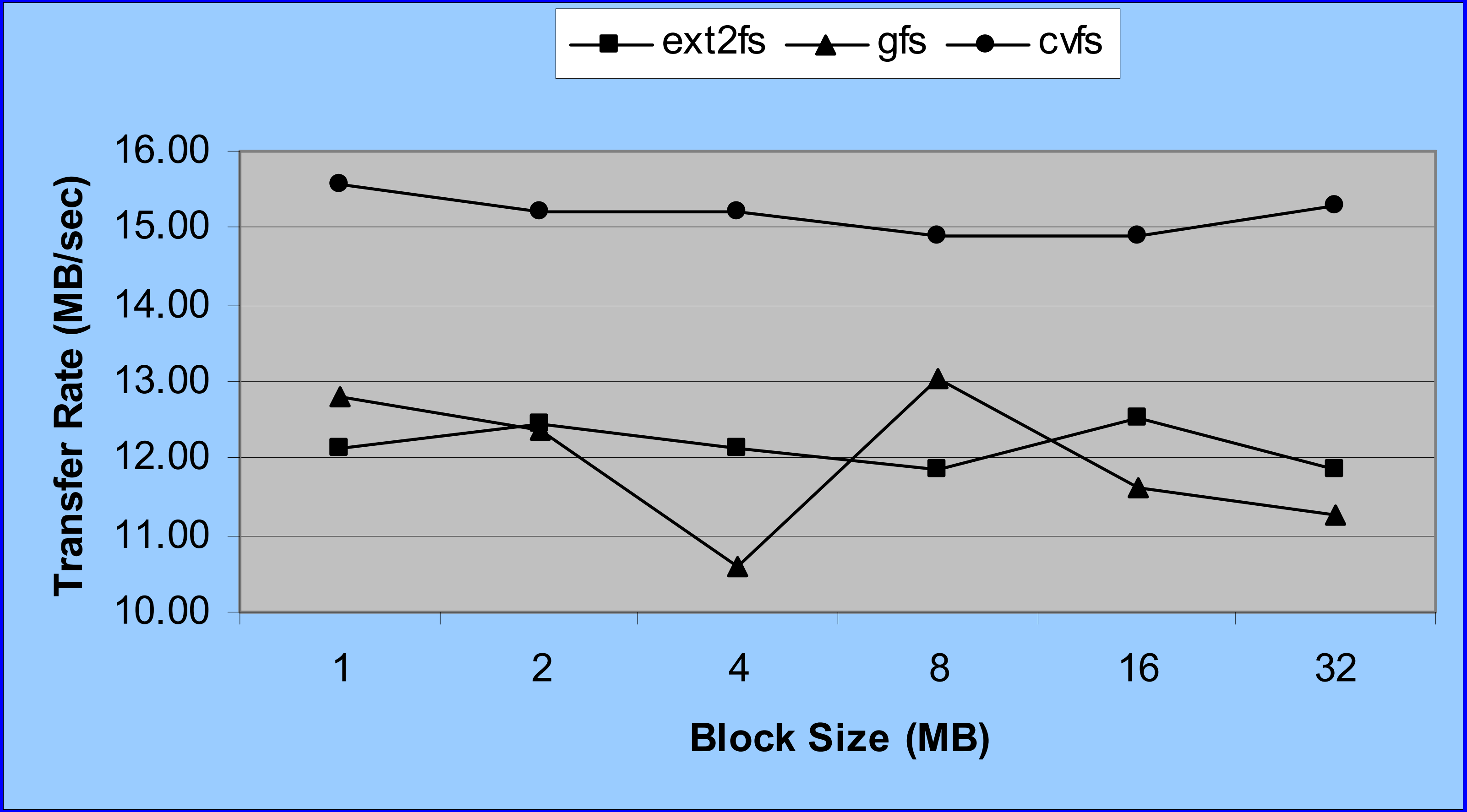
Linux iSCSI Connected – Reads



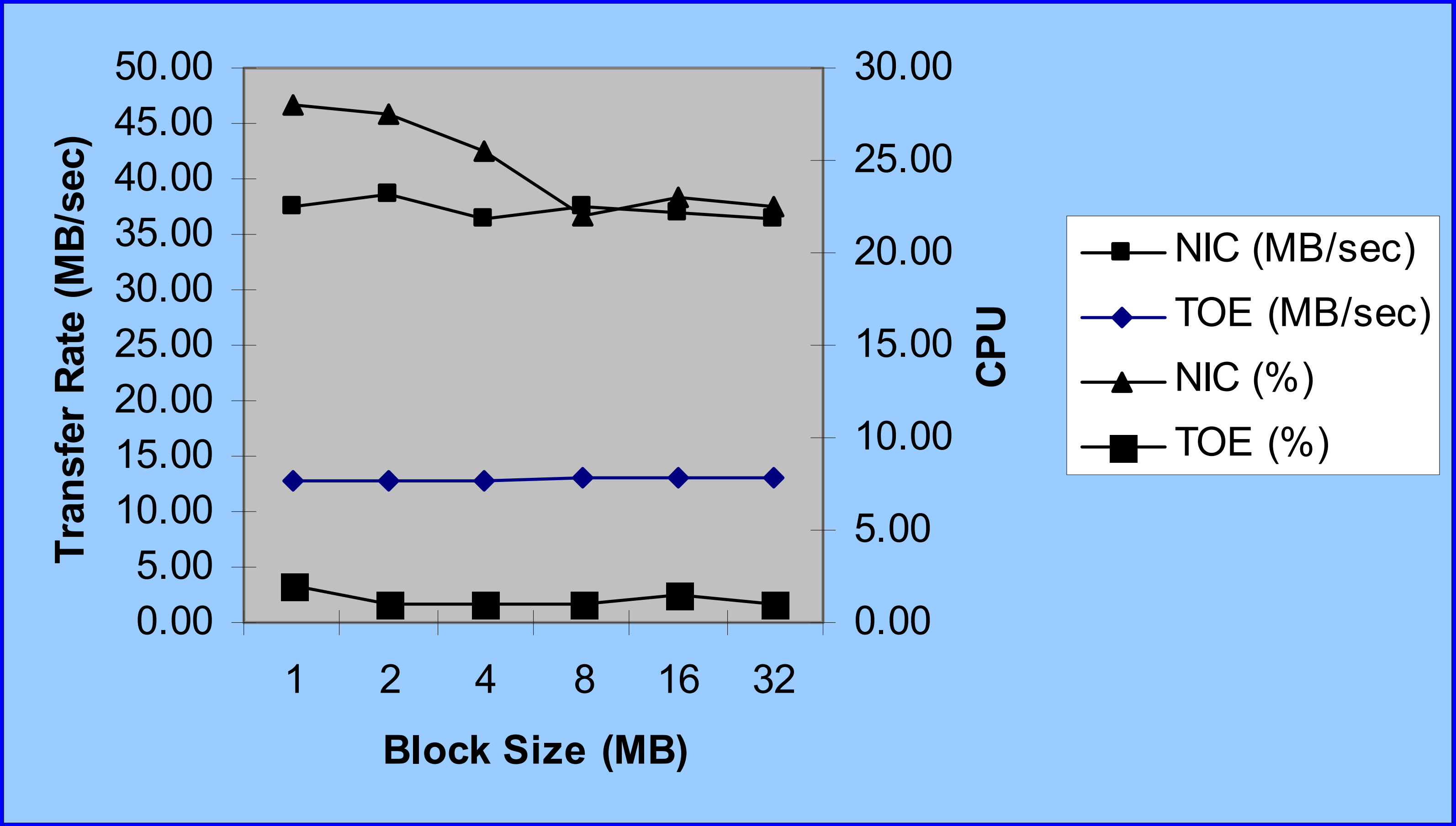
U of MD Linux iSCSI Connected – Writes



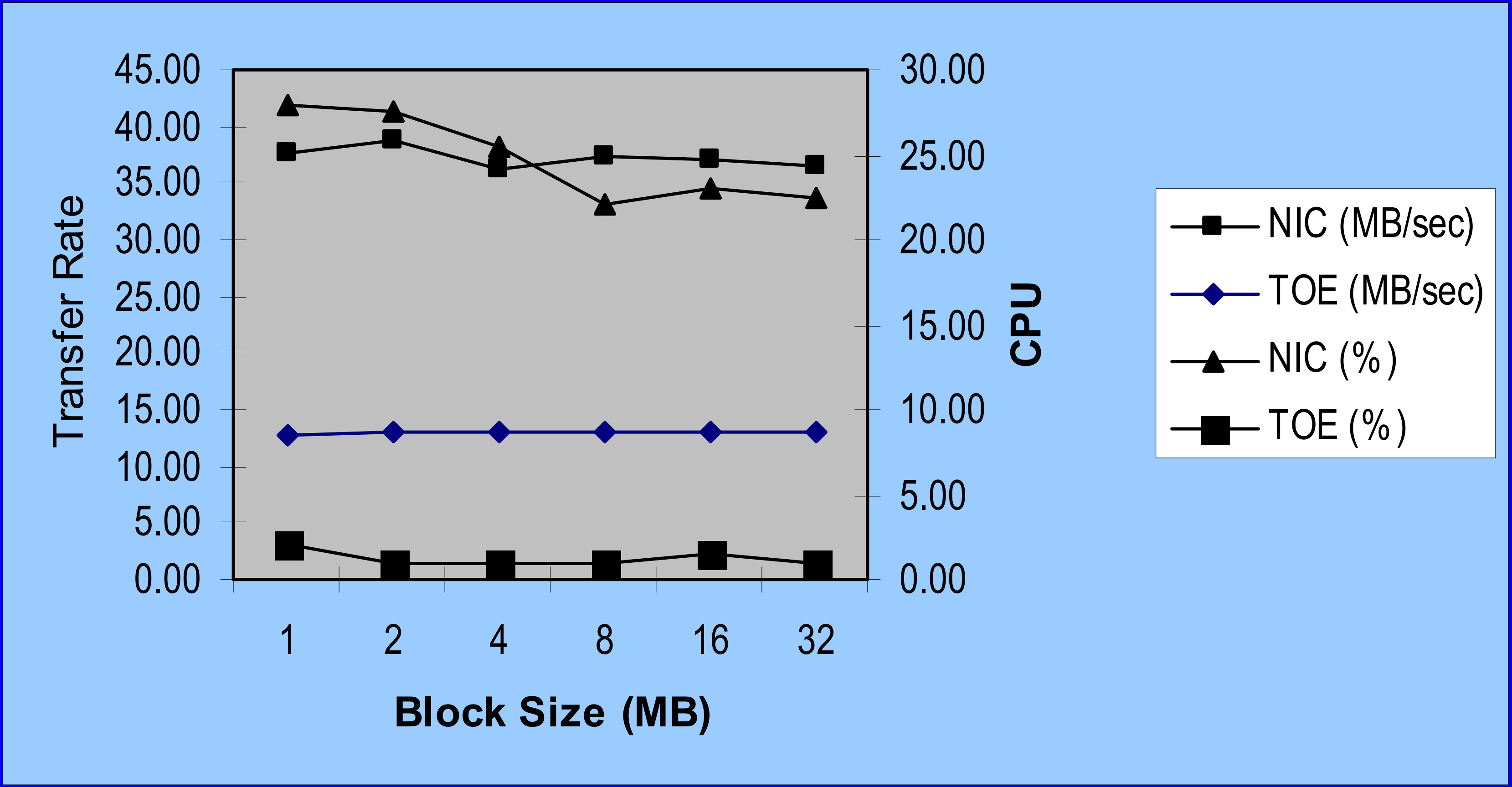
U of MD Linux iSCSI Connected – Reads



iSCSI TOE Performance – Writes



iSCSI TOE Performance – Reads



Bonnie++ and Postmark Results

- Compared various combinations of clients and file systems for small file, random operations
- Benchmarks highlighted file system design differences more so than iSCSI characteristics
  - ext2fs inodes on disk
  - GFS central lock manager with locally cached inodes
  - CVFS centralized metadata function
- Ext2fs performed the best then GFS then CVFS
- Performance best for FC client, U of MD worst
- Complete set of numbers available in paper

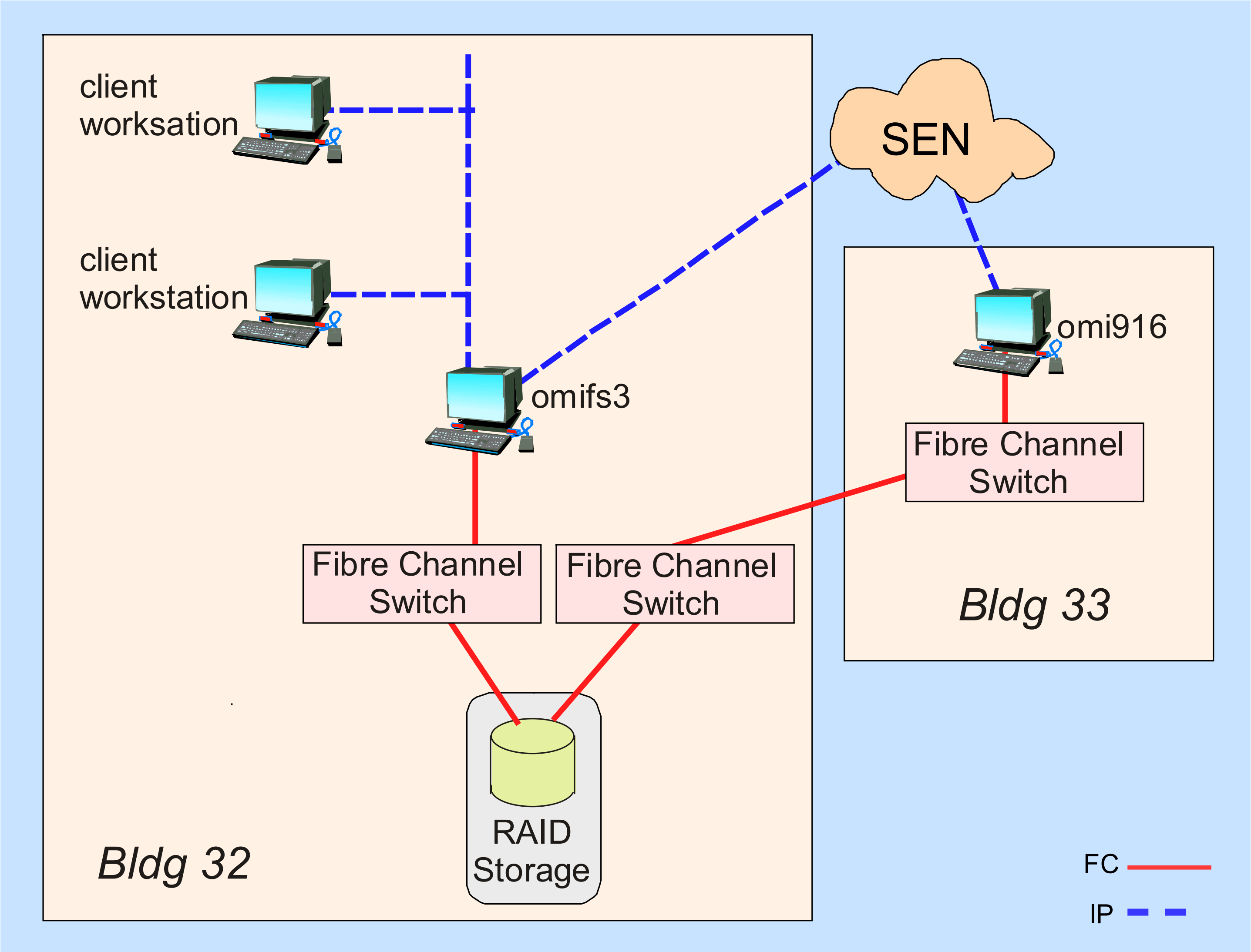
Application Testing – U of MD

- Purpose:
  - Demonstrate operational value of shared data repository
  - Generate MODIS composite data at U of MD using storage located at GSFC ~ 6 miles away
- Test results:
  - Local ext2fs dataset
    - > 1hr 45 min
    - Excluding ftp overhead time of 45 min
  - GSFC resident dataset
    - gfs > 2 hr 8 min
    - cvfs > 3hr 15 min

Ozone Monitoring Instrument System

- Purpose:
  - Study datasets from the Total Ozone Mapping Spectrometer (TOMS) instruments and Solar Backscatter Ultraviolet (SBUV) instruments
  - Prepare for Ozone Monitoring Instrument (OMI)
- Architecture:
  - Multi-building FC and NFS connected clients
  - CVFS SAN file system
  - DataDirect storage
  - Brocade switches
- Opinions to date:
  - Shared SAN storage performance comparable to local
  - NFS performance marginal but expected to improve

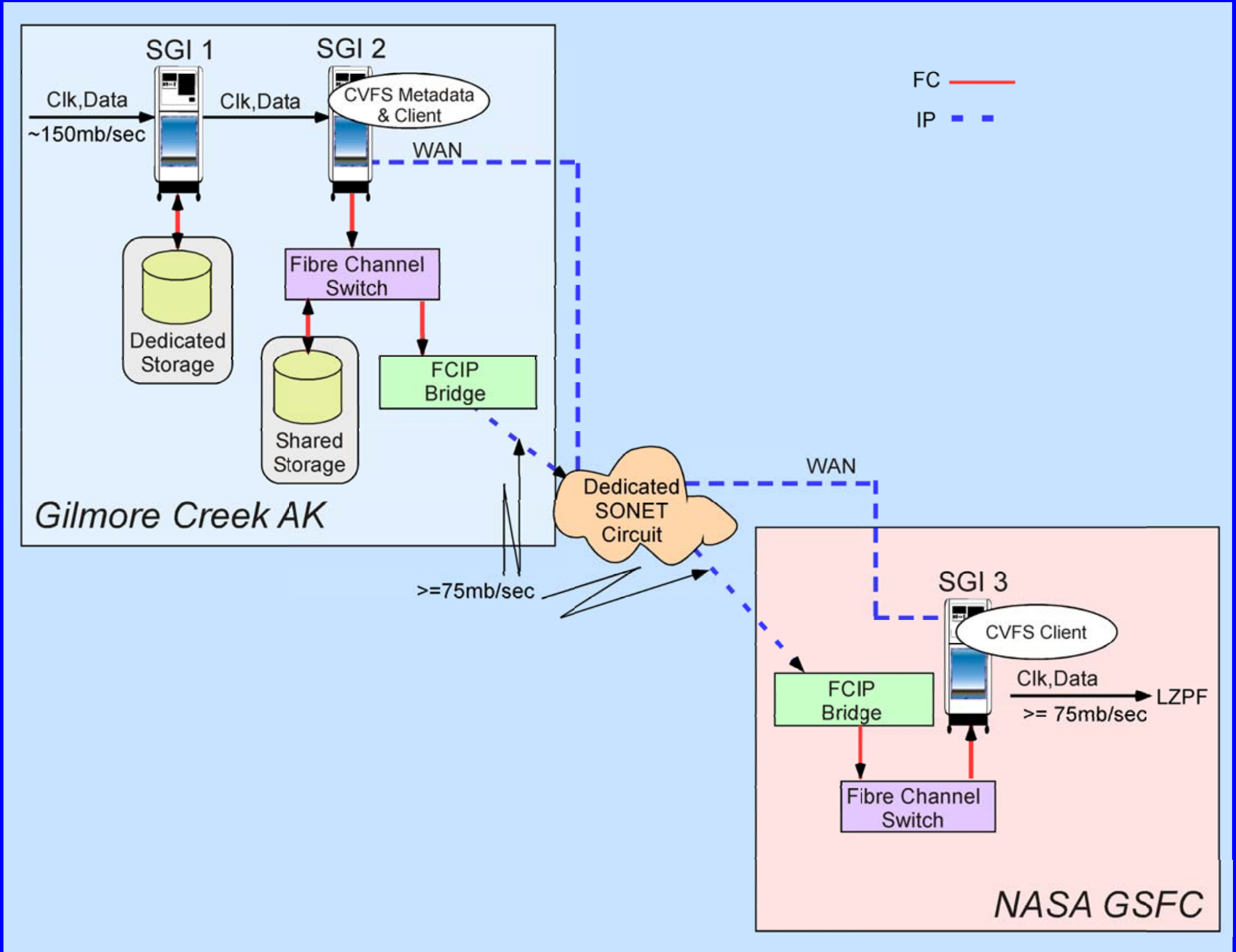
OMI Architecture



Alaska-to-GSFC SAN File System

- Objective:
  - Leverage IP technology for moving data between Alaska and GSFC
- Trades:
  - Standard FTP
  - SAN file system involving Alaska and GSFC
- Results to date – inconclusive, a work in progress:
  - First set-up used product designed for E-port expansion
    - Not well suited for extending block device over the required distance
    - Bandwidth was unacceptable
  - Tests used delay simulator and loopback
  - Additional tests planned using different equipment

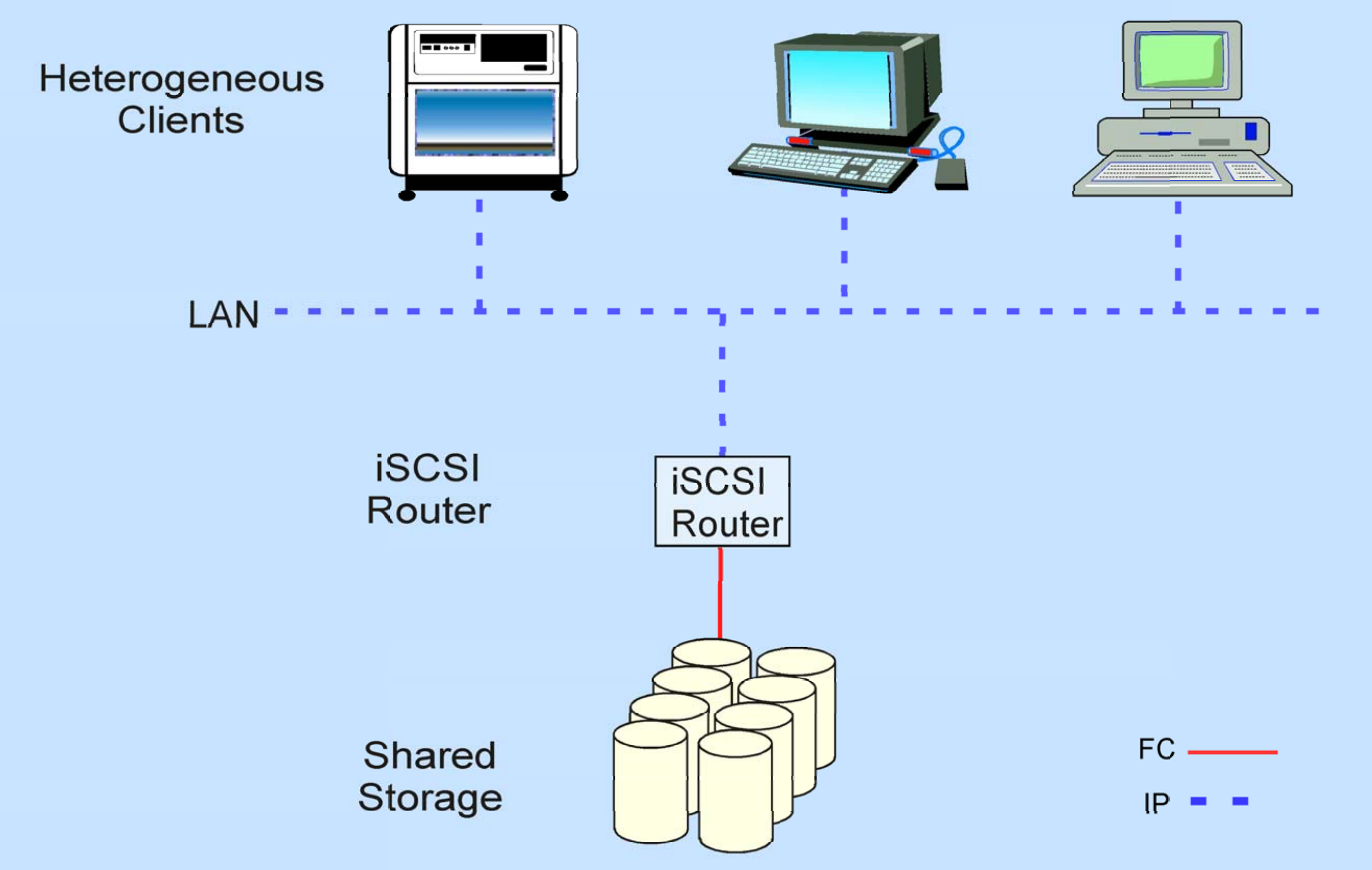
Alaska SAN Architecture



Conclusions

- IP is more than just a viable SAN technology
  - Painless to implement and test
  - Poised to have a dramatic impact
  - Market has yet to play out the options - iSCSI, FCIP and/or iFCP
  - Vendor commitment still forming
  - Standards:
    - iSCSI passed
    - FCIP and iFCP are in work
- Gain in flexibility offsets bandwidth loss for potentially a large category of users
  - Easy to envision a SAN constructed completely of iSCSI connected clients

Simple iSCSI SAN



Plans For Additional Testing

- iSCSI
  - Jumbo frames
  - Distance limit testing
- 2Gbit Fibre Channel and Fibre Channel trunking
- Geographically distributed SAN file systems
- FCIP and iFCP equipment
- SAN management software
- Security products
- Etc.

Build a happy operational base!

Acknowledgements

- J Patrick Gary, NASA GSFC
- George Uhl, NASA GSFC
- Tino Sciuto, NASA GSFC
- Charlene DiMiceli, University of Maryland
- Fritz McCall, University of Maryland
- Mike Smorul, University of Maryland