



YottaYotta

High Performance NetStorage

Storage Grid and Distributed Block Systems

**Geoff Hayward, Director of Advanced
Technology**

P. 780-989-6800

Email: ghayward@yottayotta.com

NASA/IEEE MSST 2004

**12th NASA Goddard/21st IEEE Conference on
Mass Storage Systems & Technologies**

The Inn and Conference Center

University of Maryland University College

Adelphi MD USA

April 13-16, 2004



Outline

1. Concepts

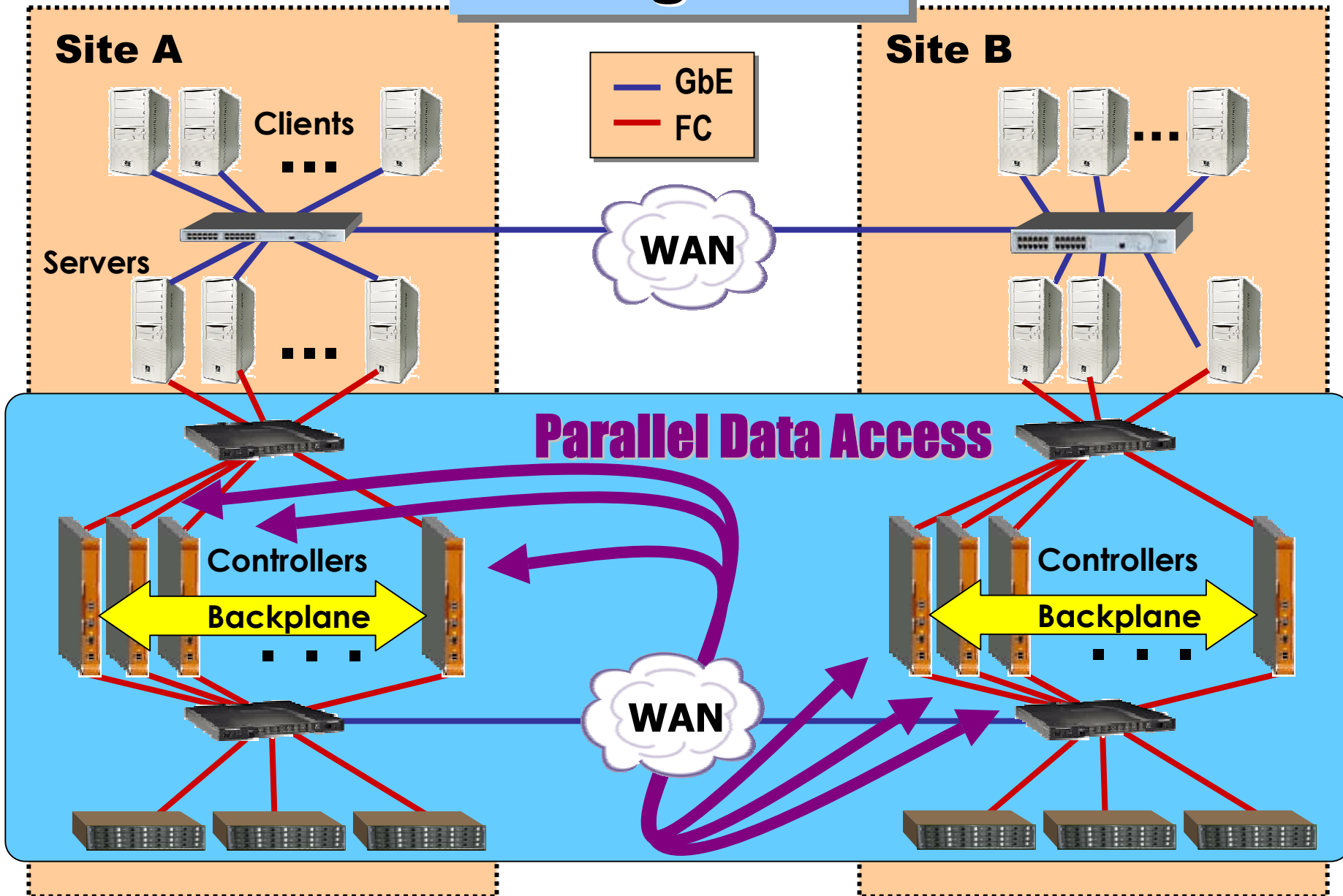
- Storage Grid
- Distributed Block Systems

2. Recent Trials and Work in Progress

- High-speed bulk transfer
- Geo-clustered file systems
- Geo-clustered databases
- Geo-HSM

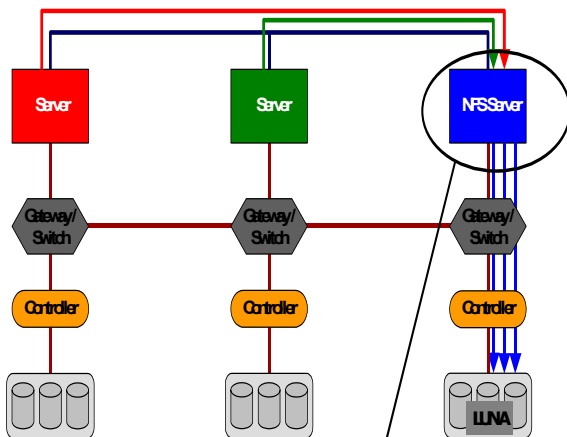


Storage Grid



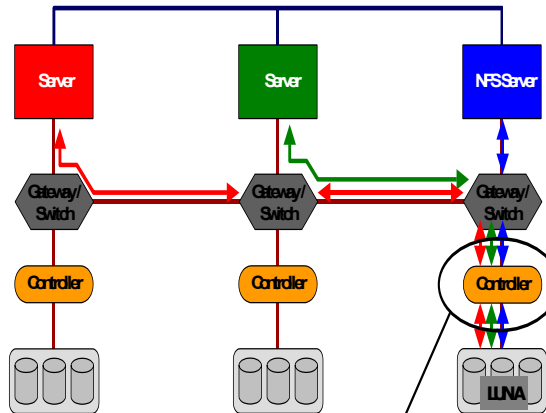
Distributed Block System

NFS



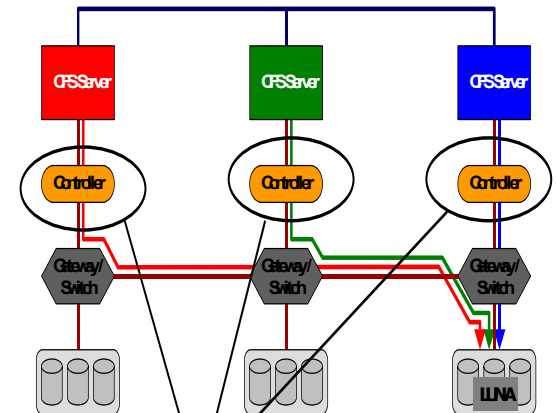
I/O through NFS Server

CFS / Serial Block System



Serial I/O

CFS / Dist. Block System



Parallel Controller Paths

Distributed Block System

- Maintains Distributed block-level cache coherence
- Cache sharing
- Parallel paths to storage
- Fail-over
- Global logical name space (for disks)
- Localizes data and meta-data where it is needed

Data and Meta-data Localization Strategies

1. Geographic caching
2. Data migration according to usage
3. Distributed meta-data directory: segments migrate according to usage
4. Geographic redundancy groups
 - Site-aware I/O routing, fail-over, and incremental re-build
5. Parallel transfer and pre-fetch across multiple optical paths

Applications

- High-speed data movement
- Geo-clustered file systems
- Heterogeneous storage sharing
- Site-aware fail-over
- Geo-clustered databases
- Geographic HSM

Geo-Clustered File System Trials

September-November 2003

Partners



YottaYotta

High Performance NetStorage



CANARIE INC.



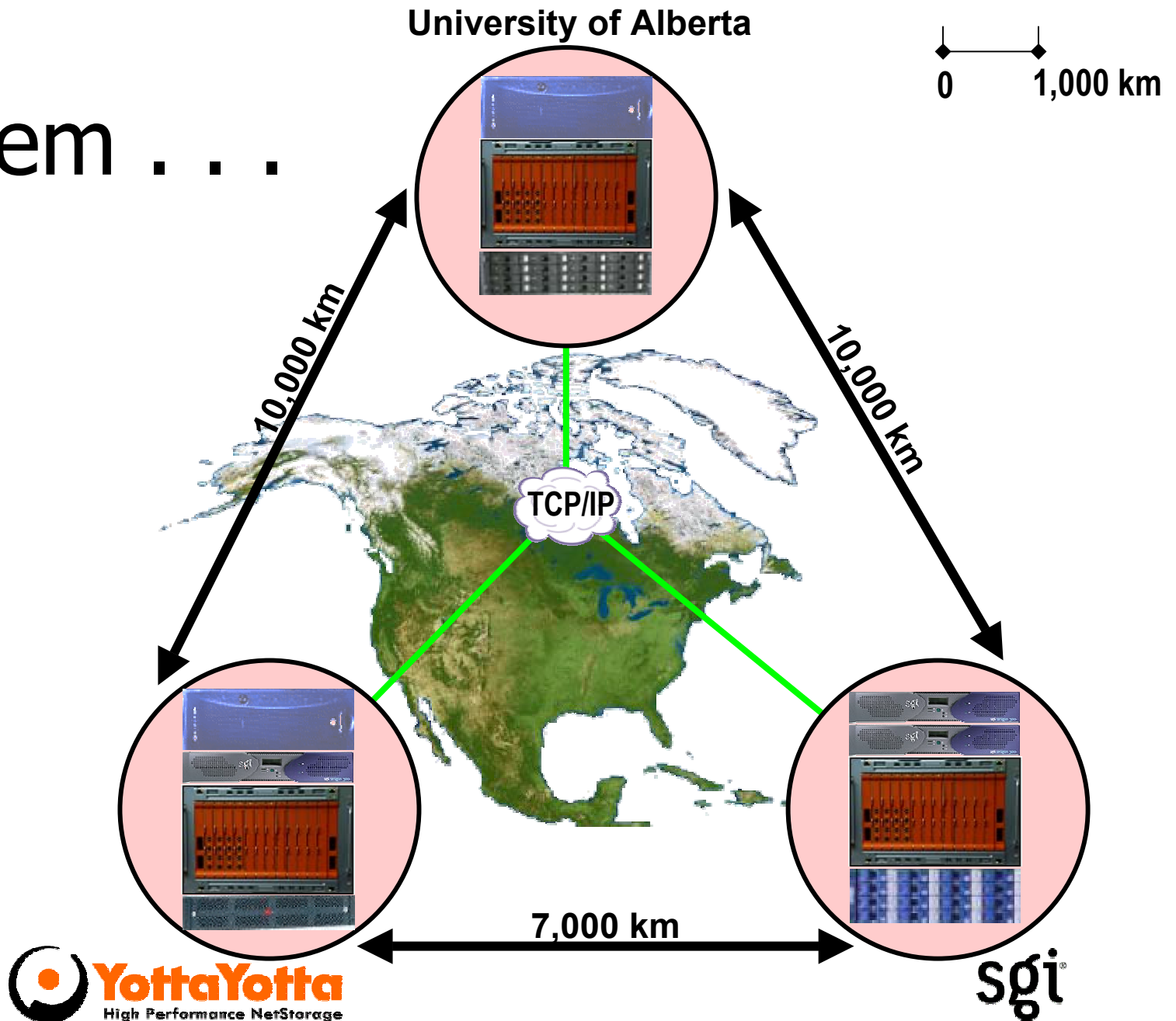
netera
ALLIANCE

STARLIGHT SM



UNIVERSITY OF
ALBERTA

The System . . .



Netera



Cisco
4912

Nbase WDM

CANARIE



Nbase WDM

CA*Net4

Cisco
15454

Juniper
T640

Abilene

StarLight

Force10
E600

Router

SciNet

YY
Booth

SGI
Booth

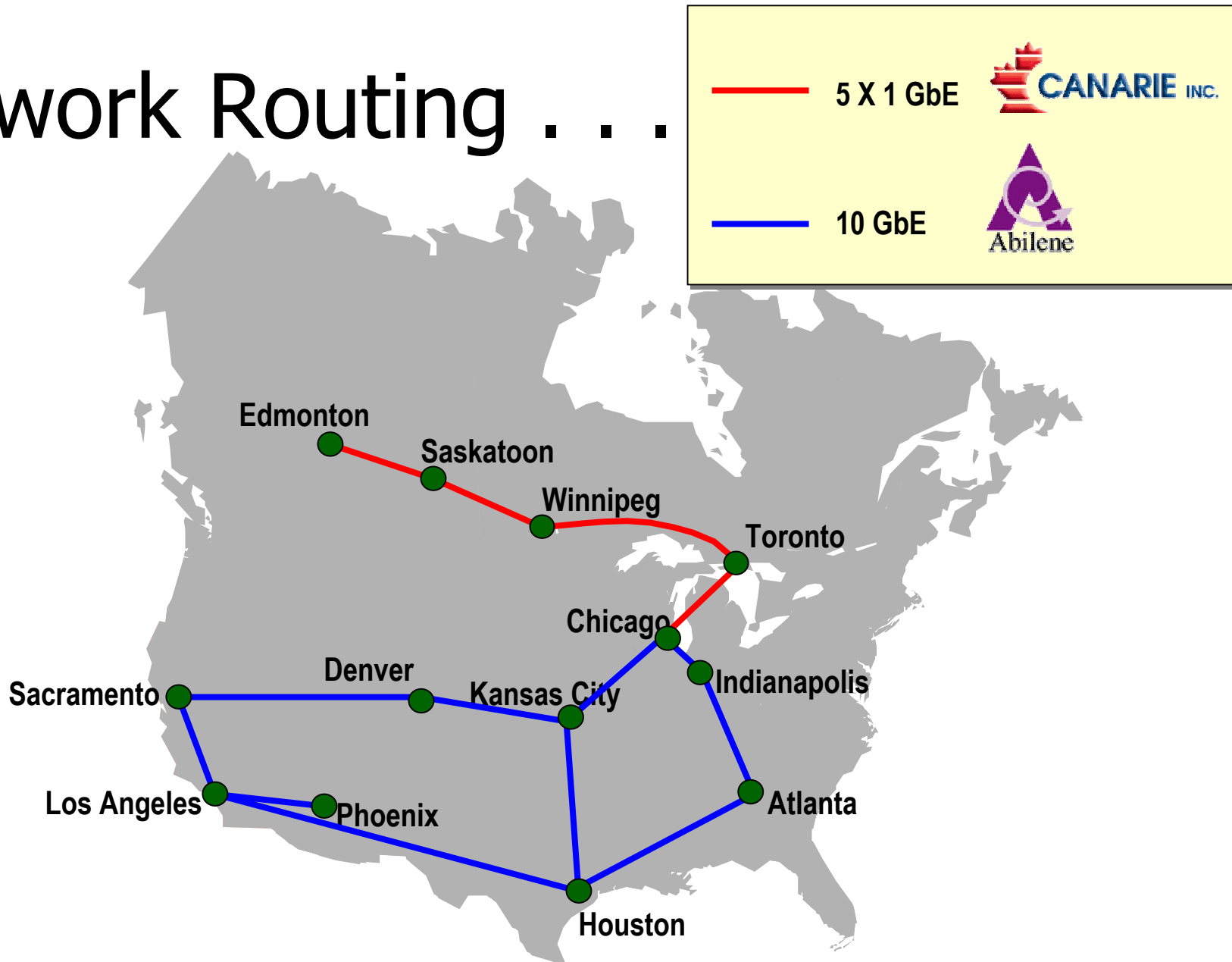
WDM

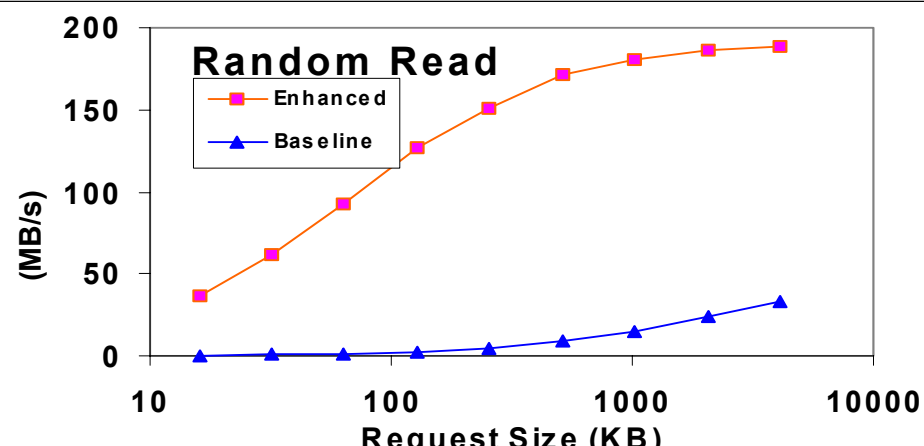
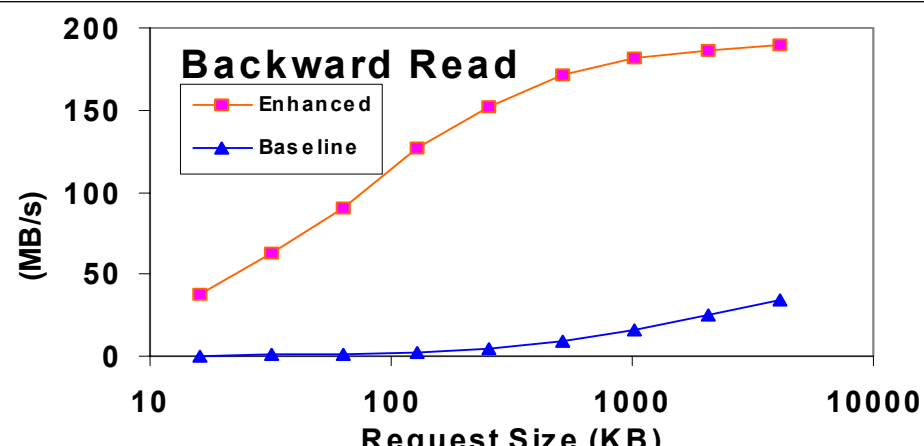
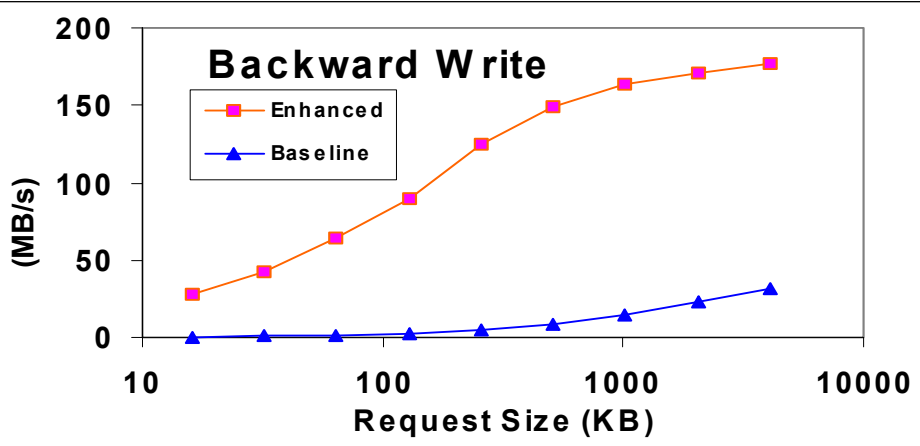
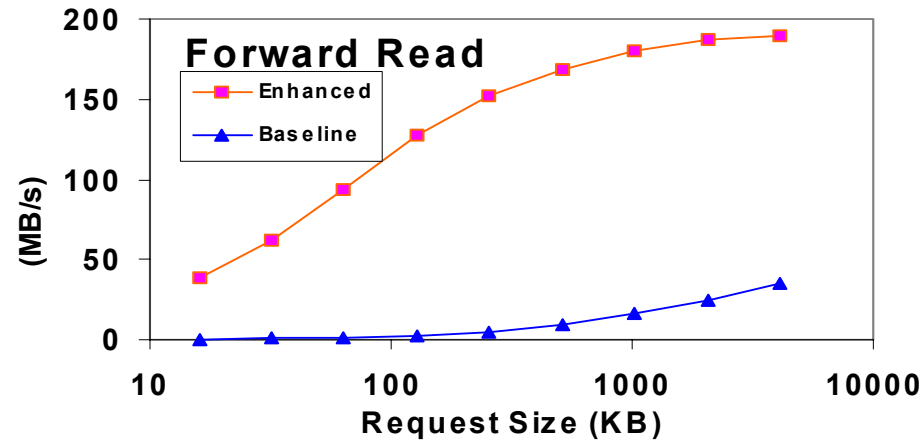
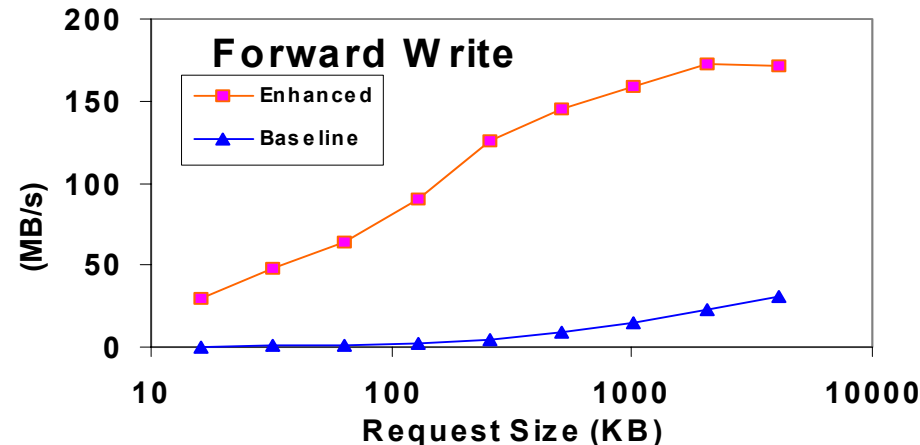
OC 192

10 GbE

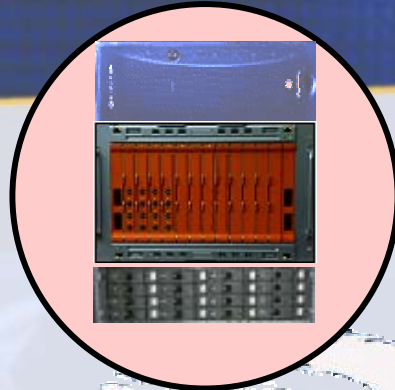
1 GbE

Network Routing . . .



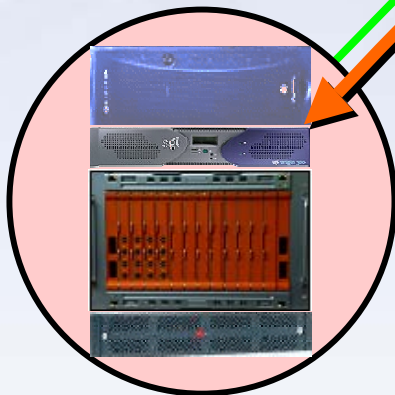


University of Alberta

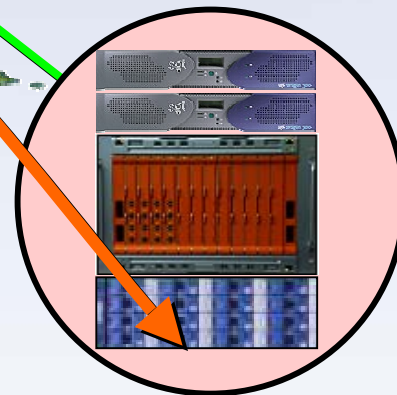


Geographically-aware
Data access

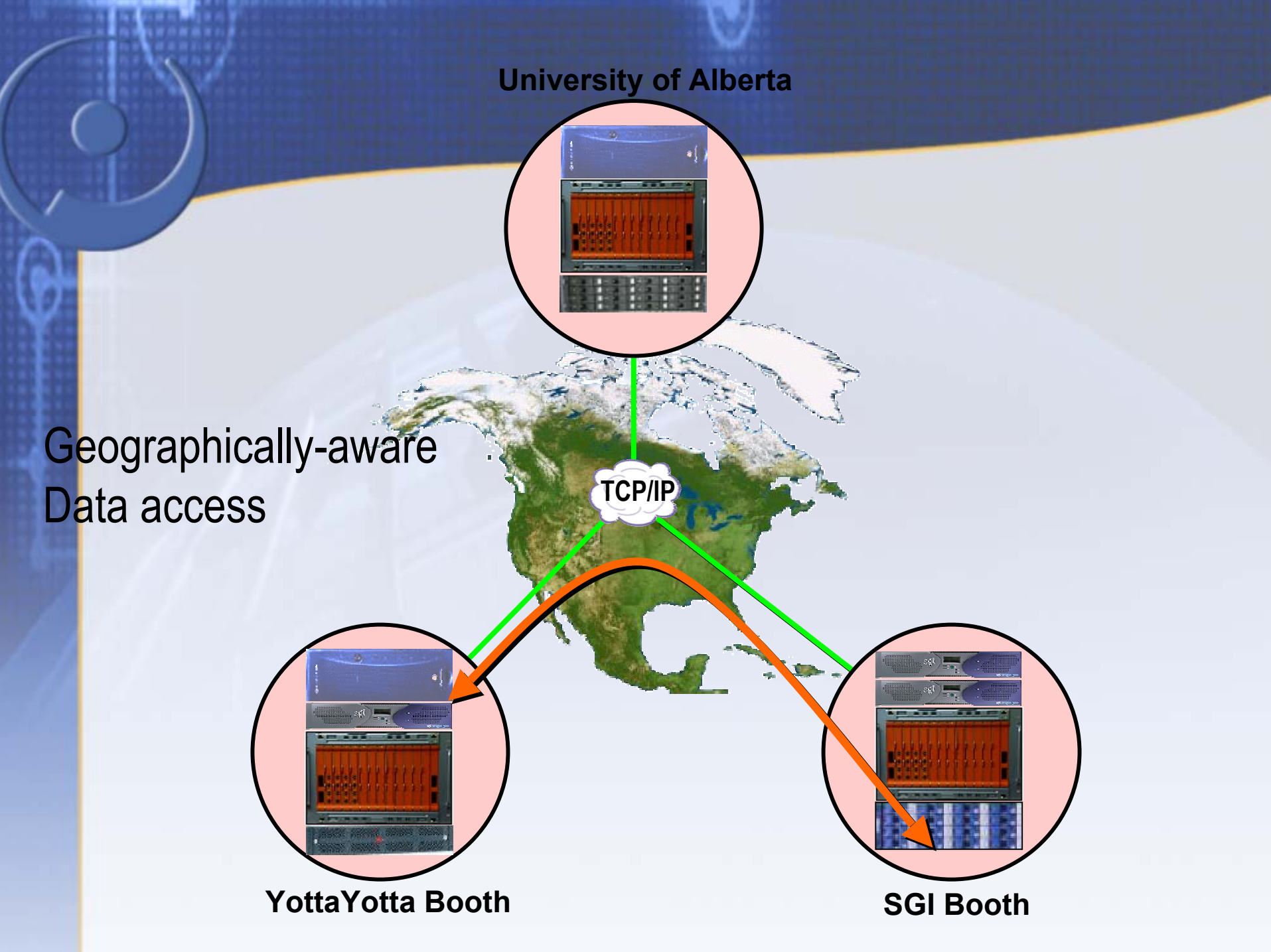
TCP/IP



YottaYotta Booth

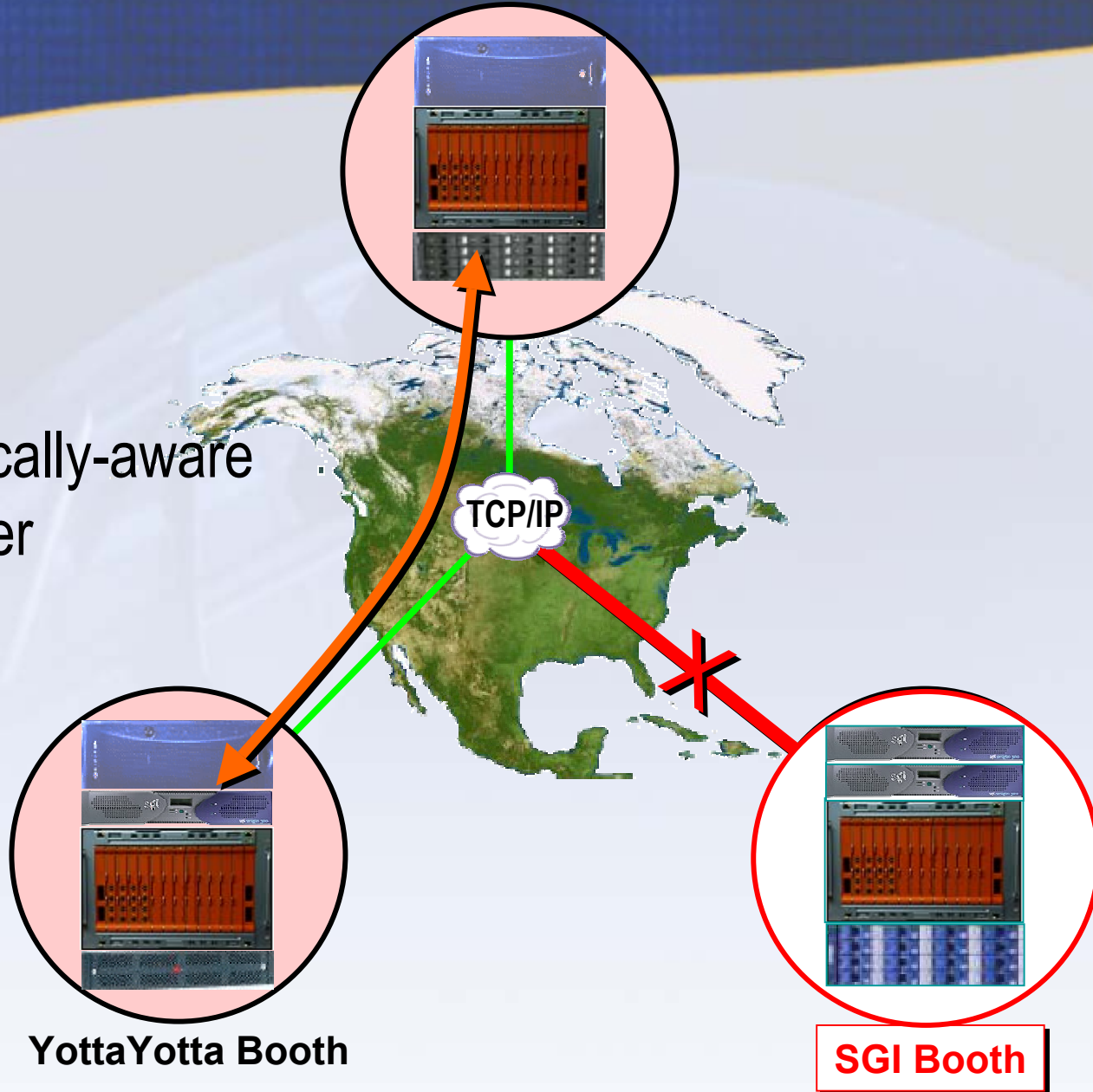


SGI Booth

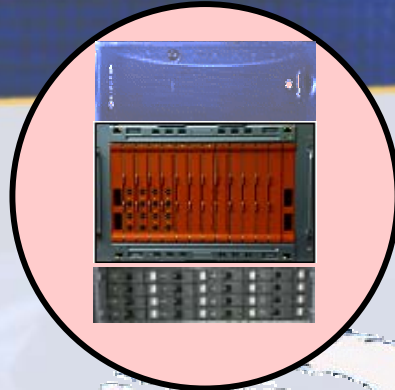


University of Alberta

Geographically-aware
Site fail-over

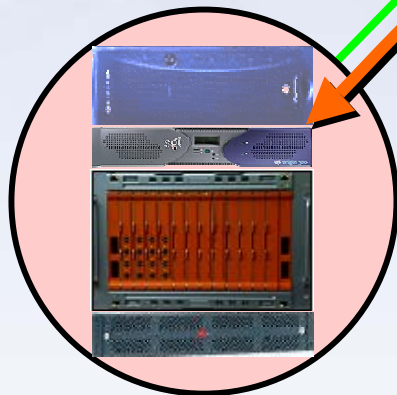


University of Alberta

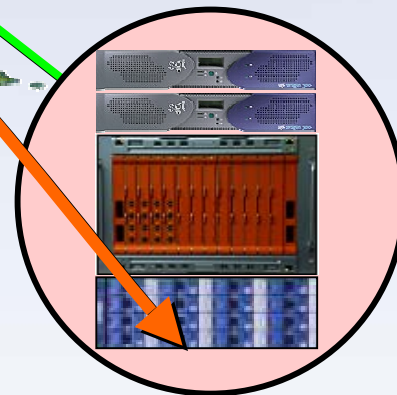


Geographically-aware
Automatic fail-back

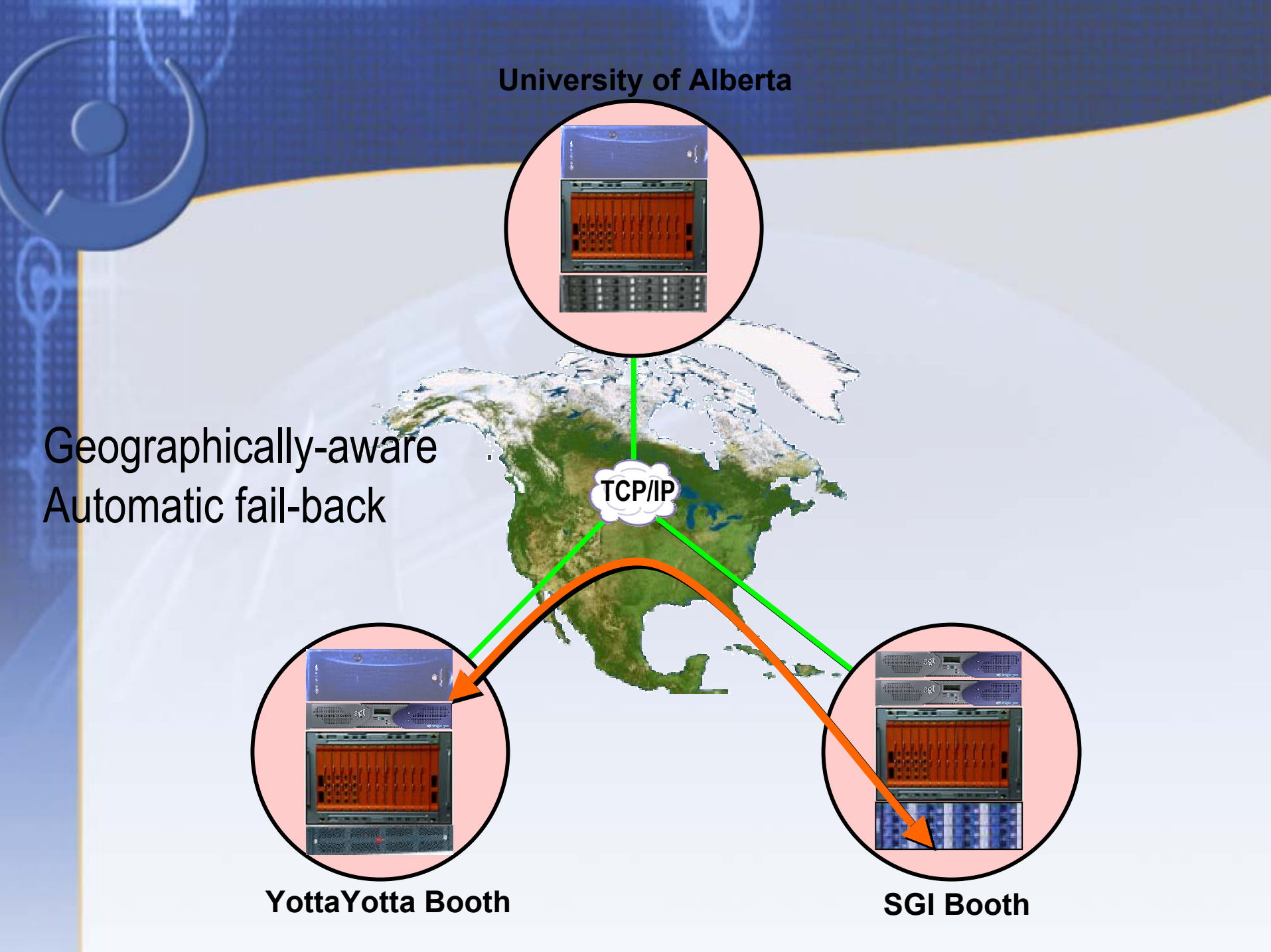
TCP/IP



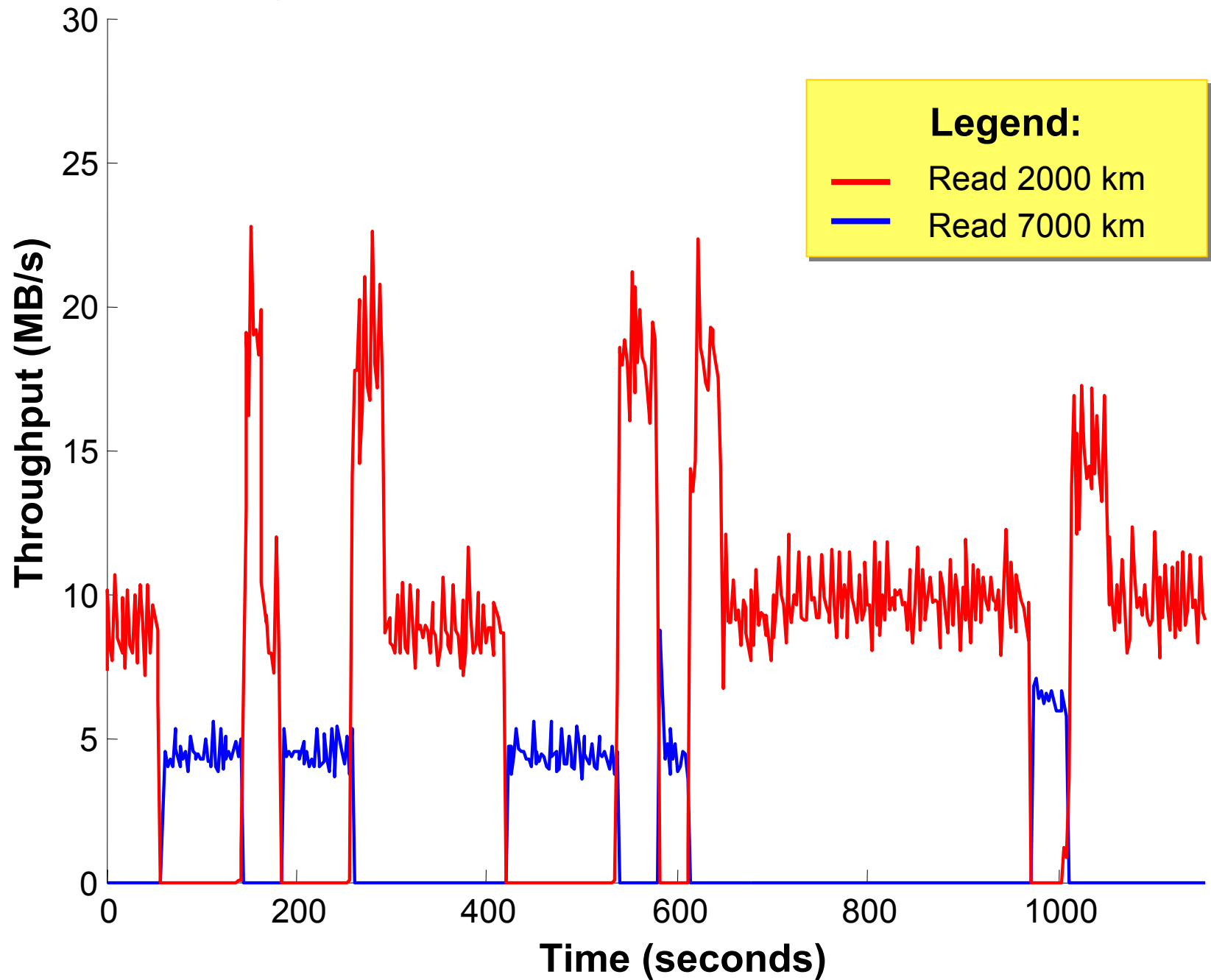
YottaYotta Booth



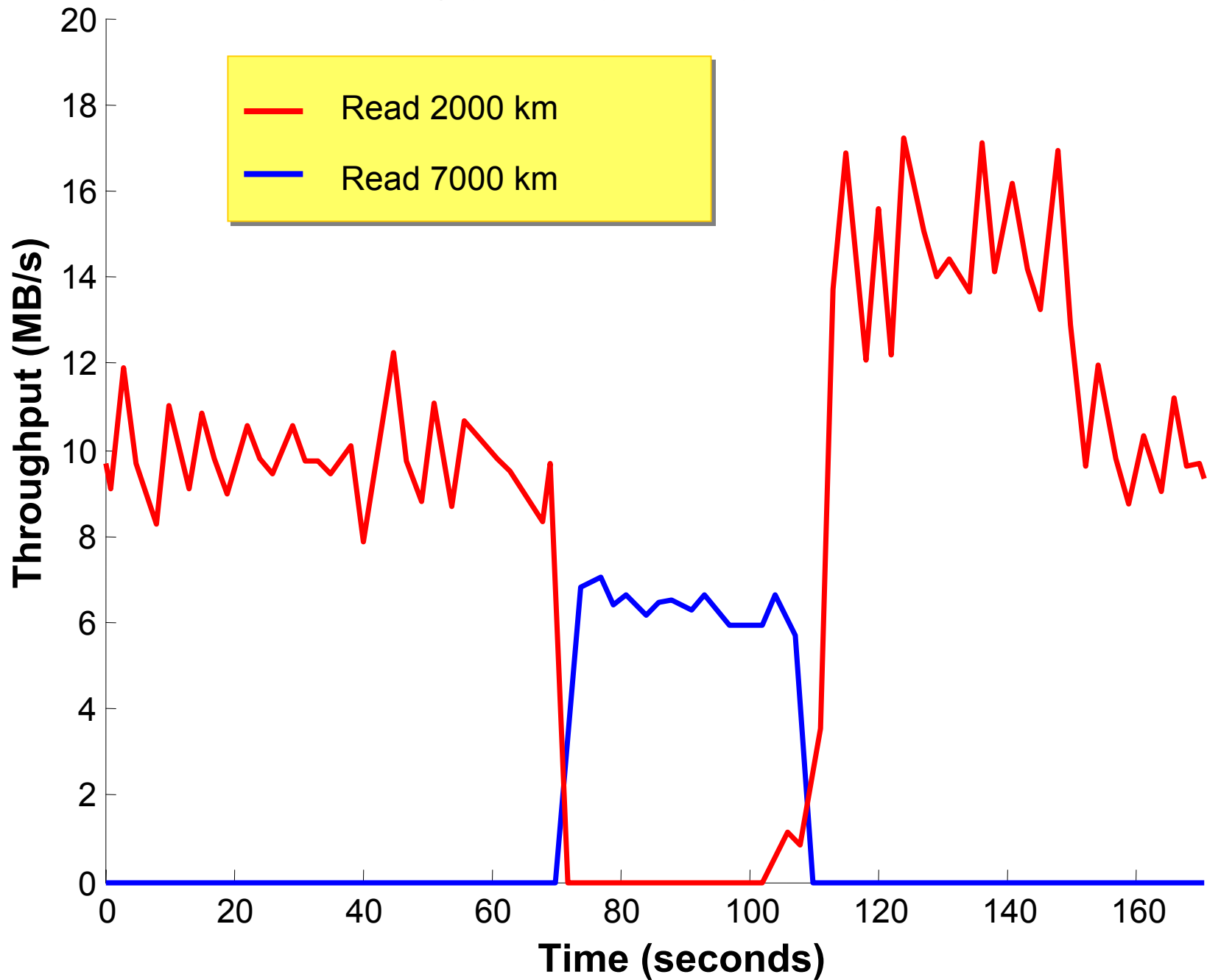
SGI Booth

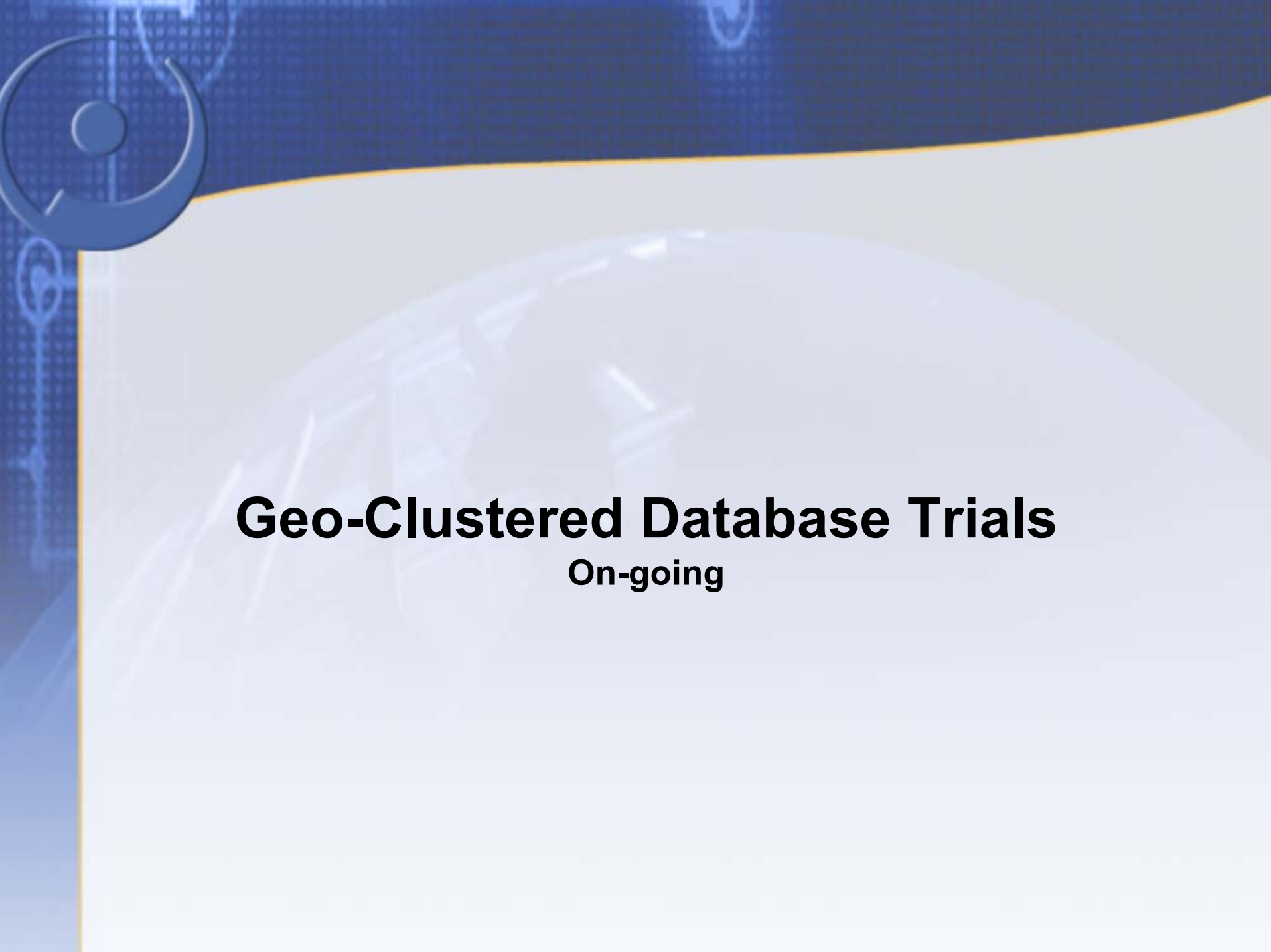


Site Fail-over between 2000 km and 7000 km



Site Fail-over Detailed View

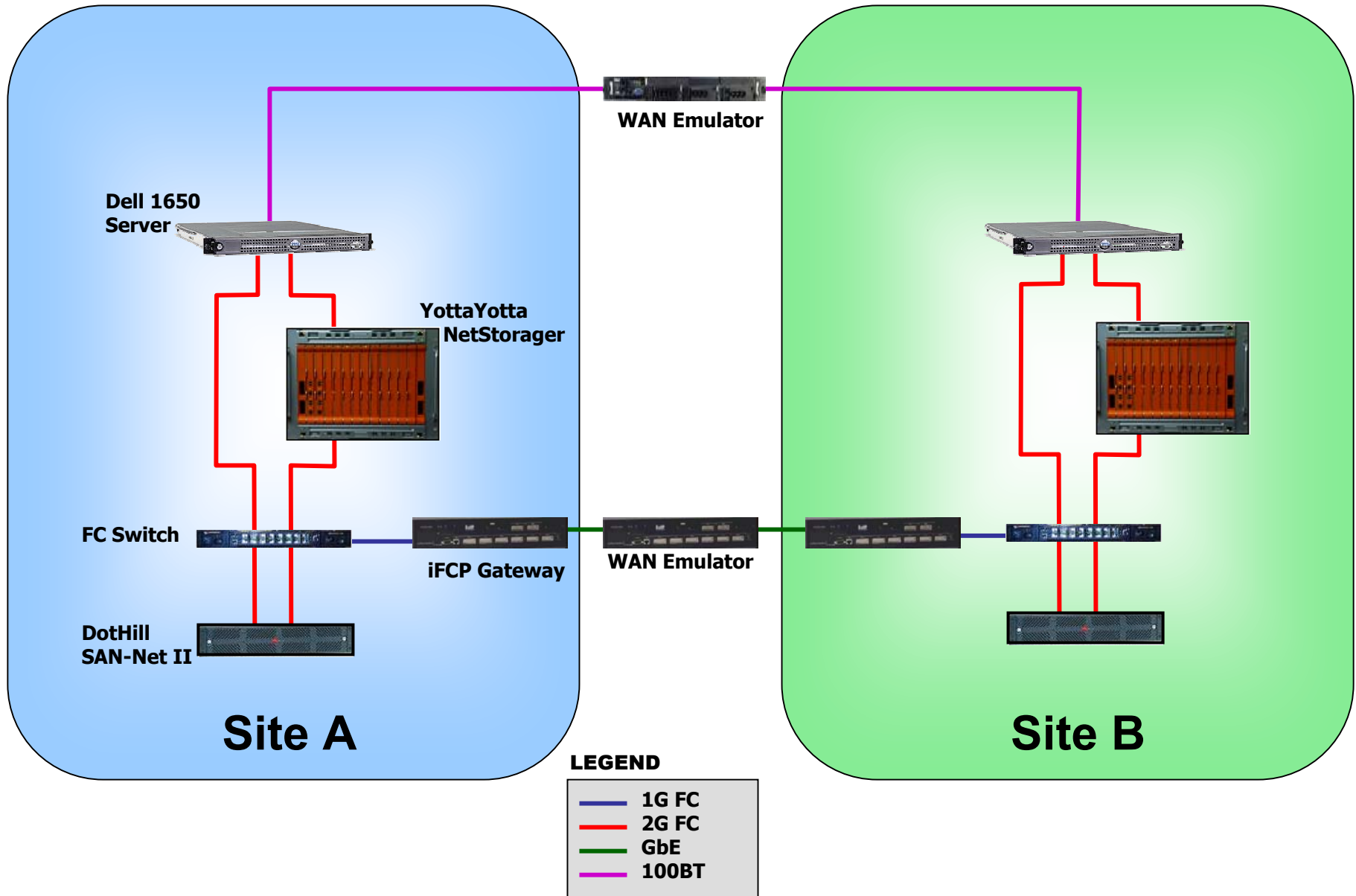




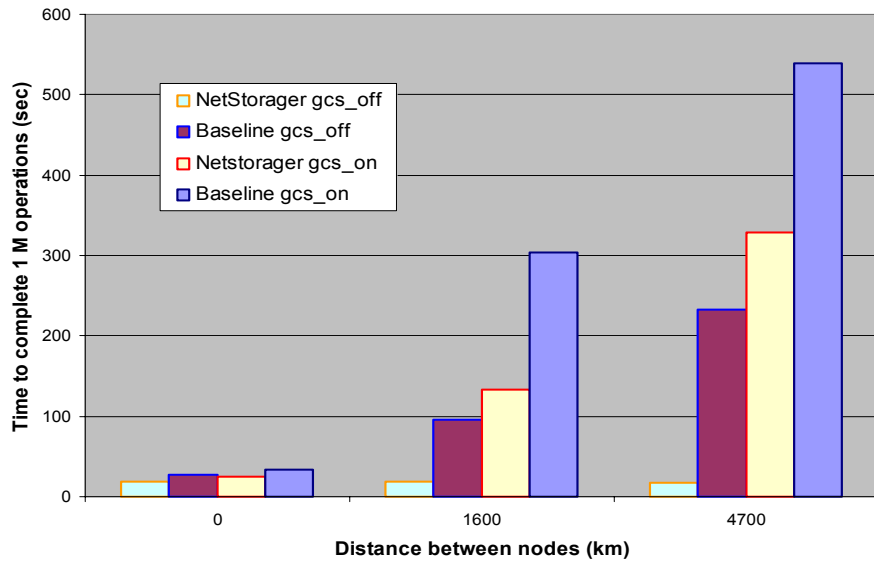
Geo-Clustered Database Trials

On-going

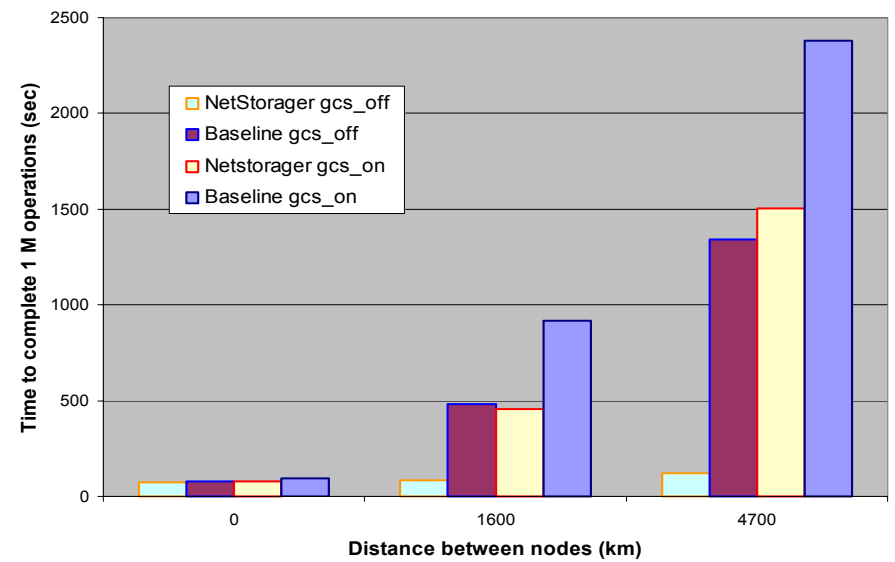
Oracle 9iRAC Windows Cluster Test Bed



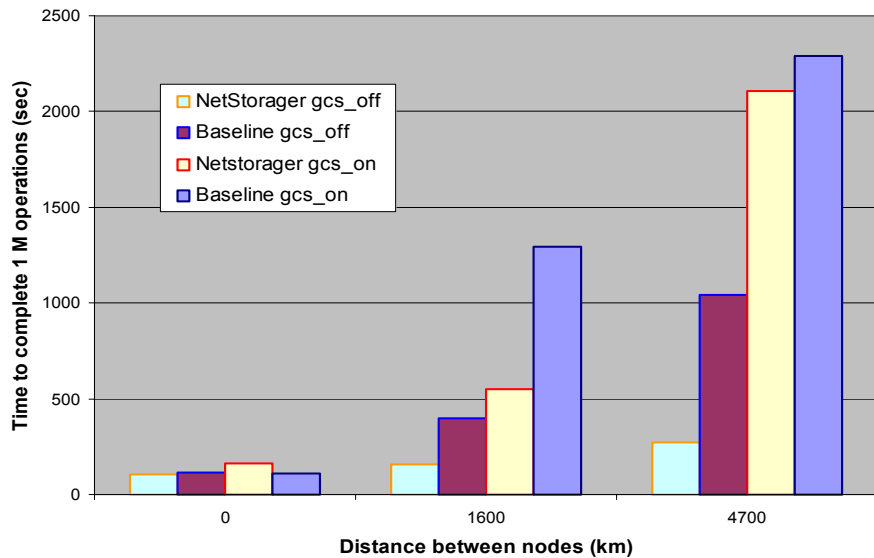
SELECT Operations to Remote LUN



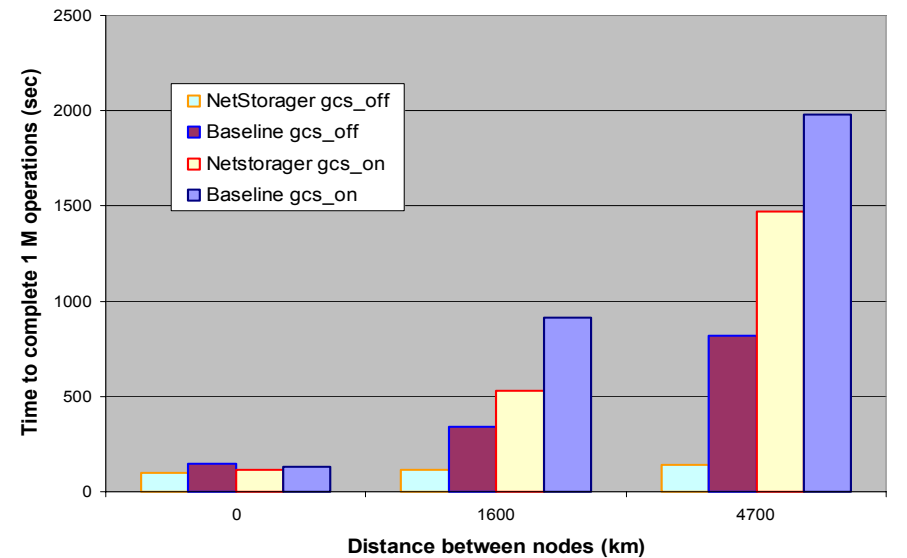
UPDATE Operations to Remote LUN



DELETE Operations to Remote LUN



INSERT Operations to Remote LUN

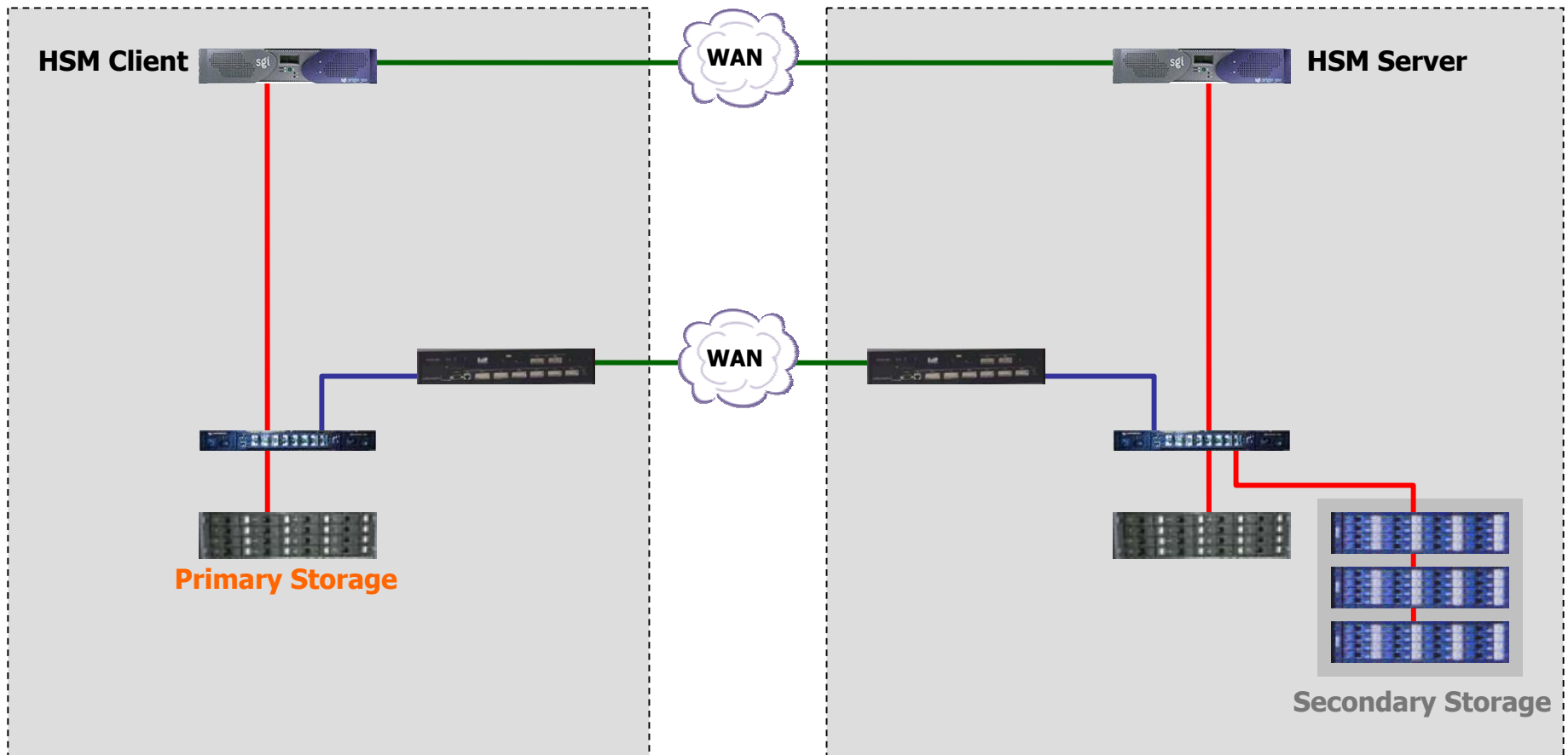




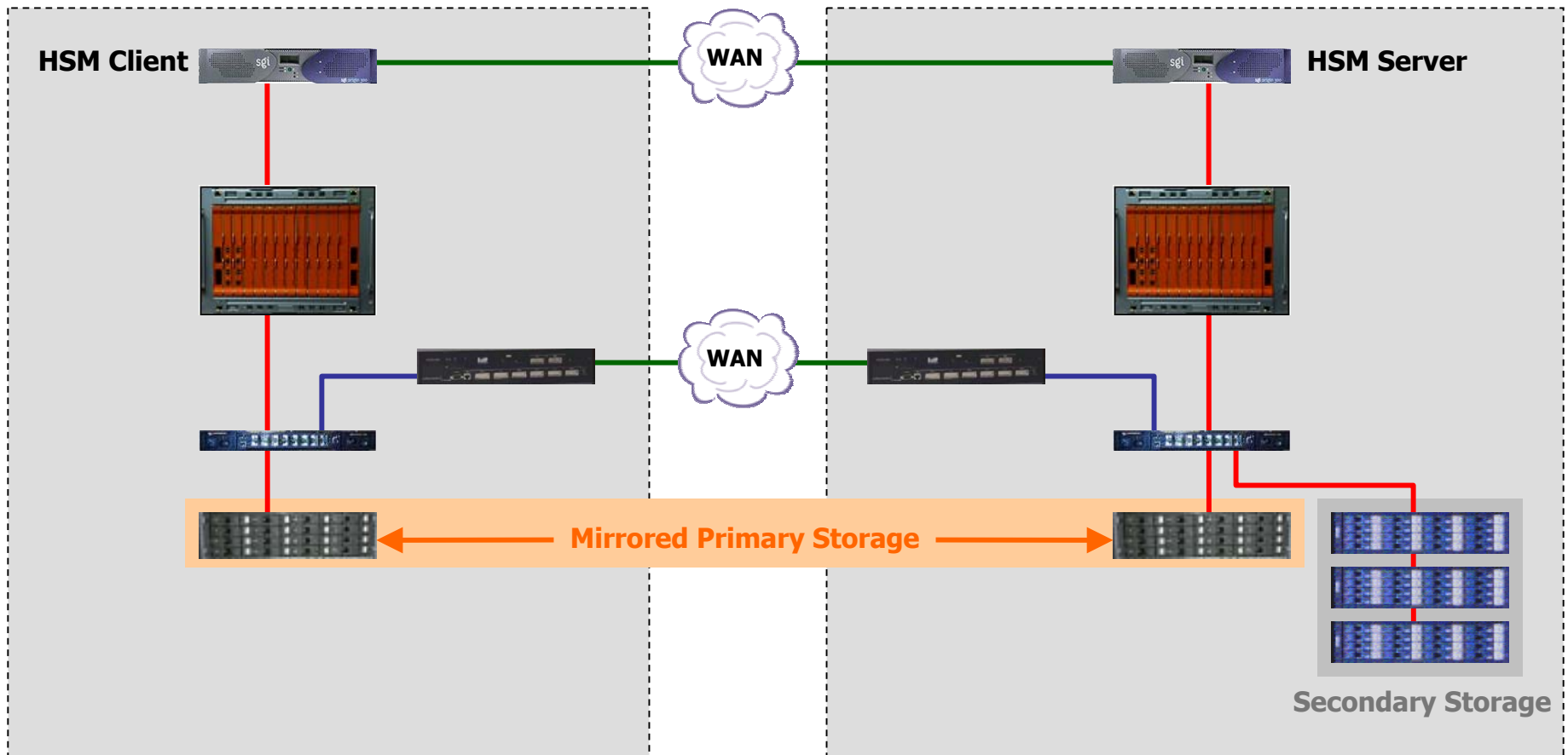
Geo-HSM Trials

On-going

Baseline Geo-HSM Configuration

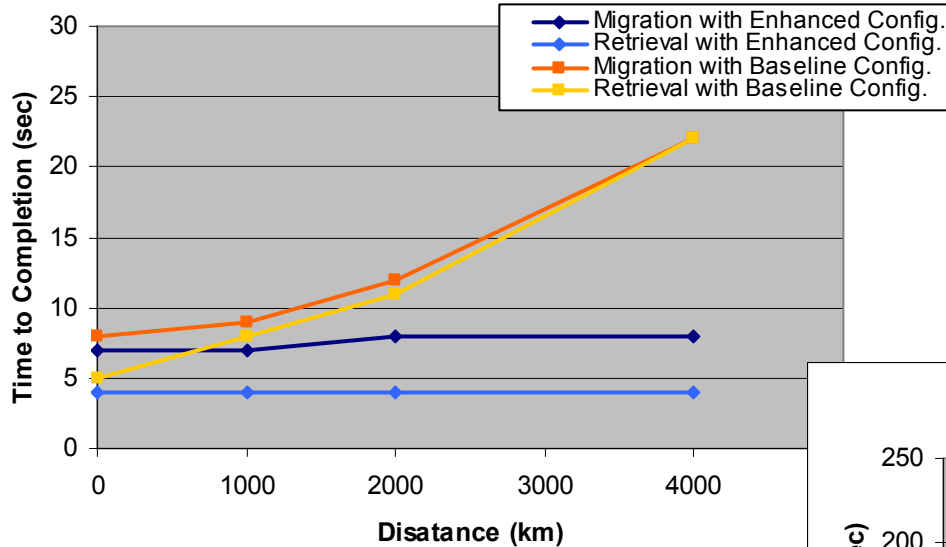


Enhanced Geo-HSM Configuration

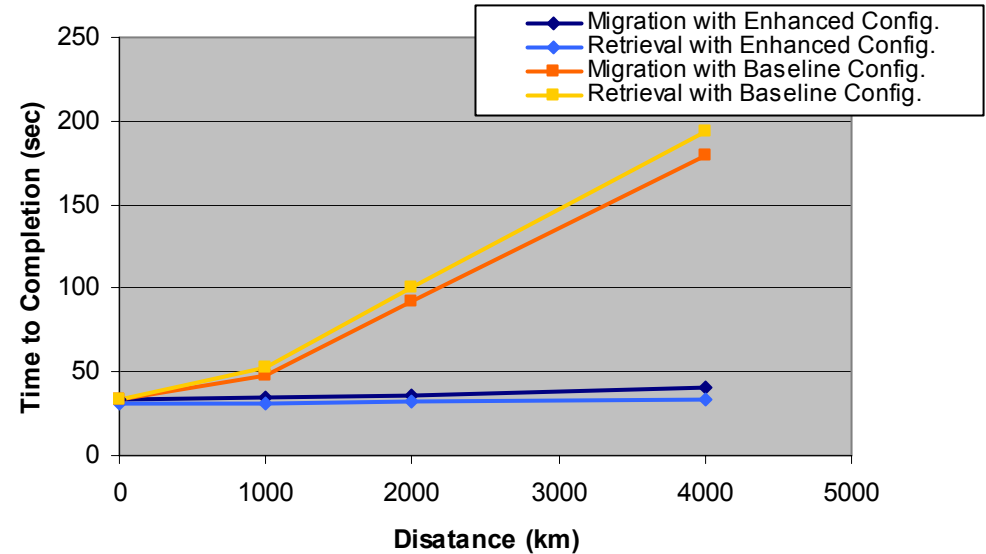


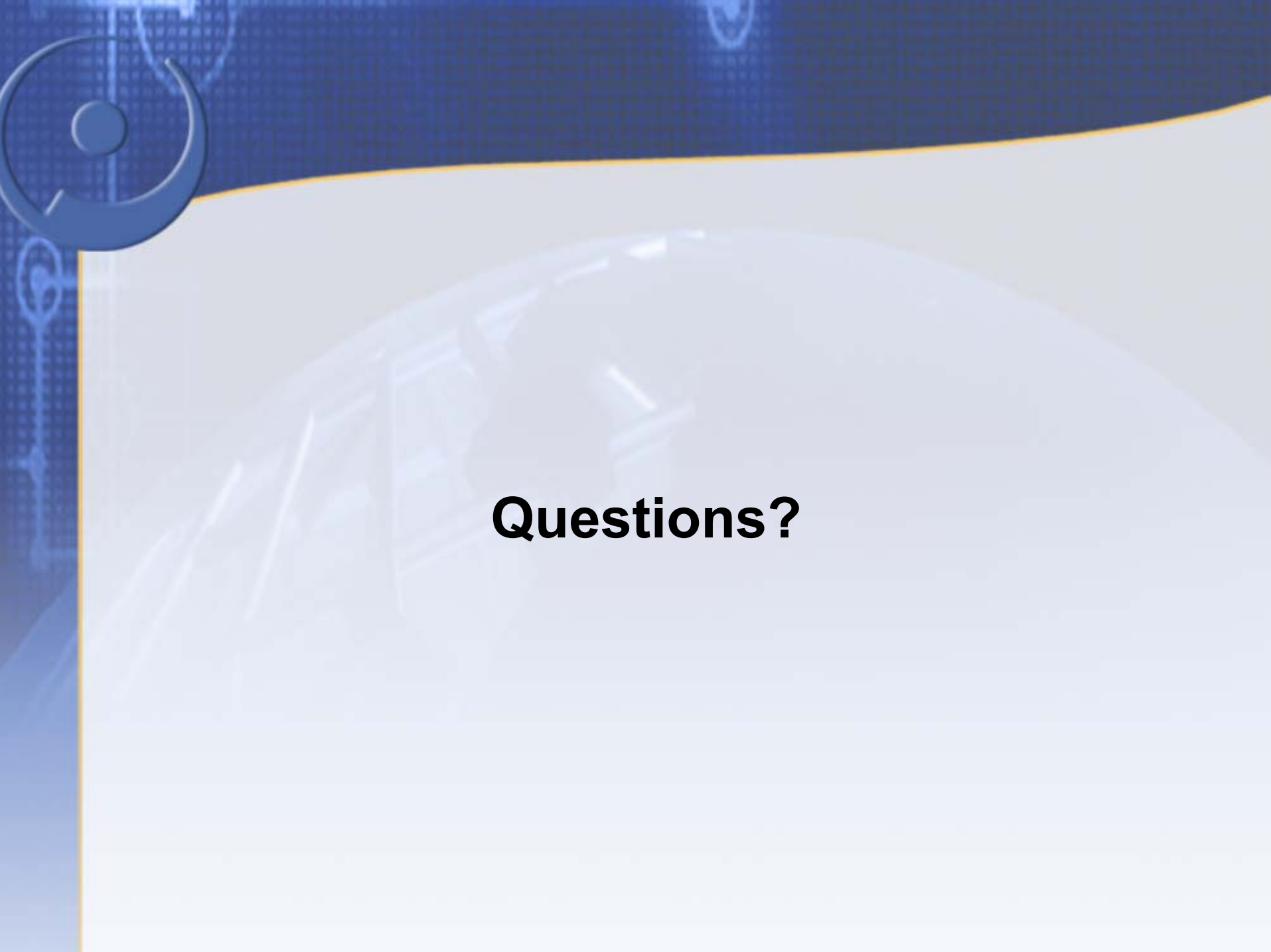
Performance Results

10x10MB Files DMF Access



10x100MB Files DMF Access





Questions?