



MCIT

SPECTRUM SCALE (GPFS) AT NYU LANGONE HEALTH

Michael Costantino, *HPC Systems Administrator*

John Speakman, *Senior Director, Research IT*

Martin Ossowski, PhD, *Director, HPC*



NYU GPFS Configuration

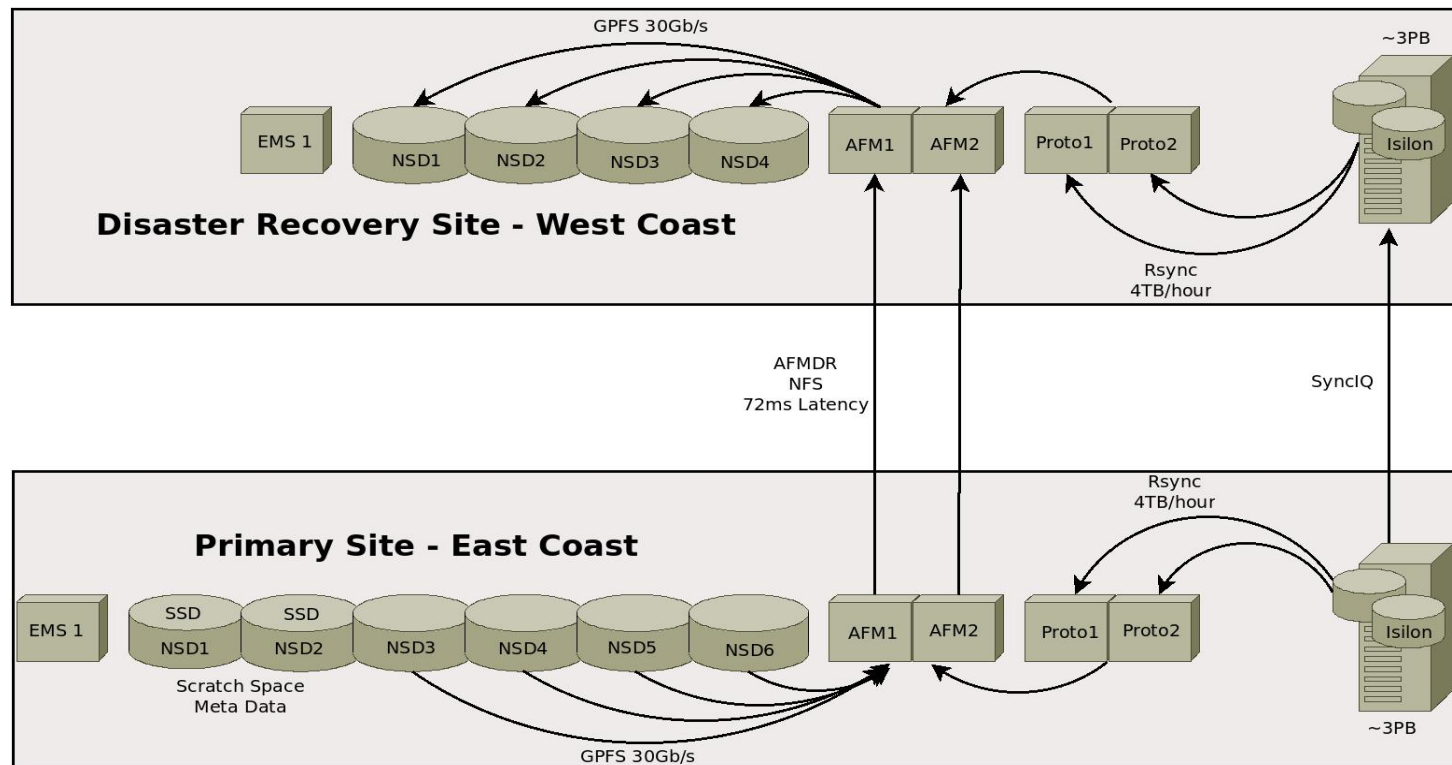
Environment

- Versions: ESS=5.3.0, GPFS=5.0.0
- Server Types:
 - 2 x GS2S = SSD
 - 4 x GL4S = HD
 - 2 x Protocol Nodes
 - 2 x AFM Gateway Nodes
- Network: Mellanox EDR 100Gb IB
- 1 x Partially Replicated Disaster Recovery Cluster

Media and Storage Pools

- 668 x 8TB 72.K Drives
- Total Raw Capacity: 5344 TB
- Data Pool = 4PB 8+2p
- 48 x 15 TB SSD Drives
- Total Raw Capacity: 720 TB
- Meta data Pool: 32TB 3WayReplication
- Scratch Pool: 420TB 8+2p

AFMDR and Data Migration



Replication and Migration Challenges

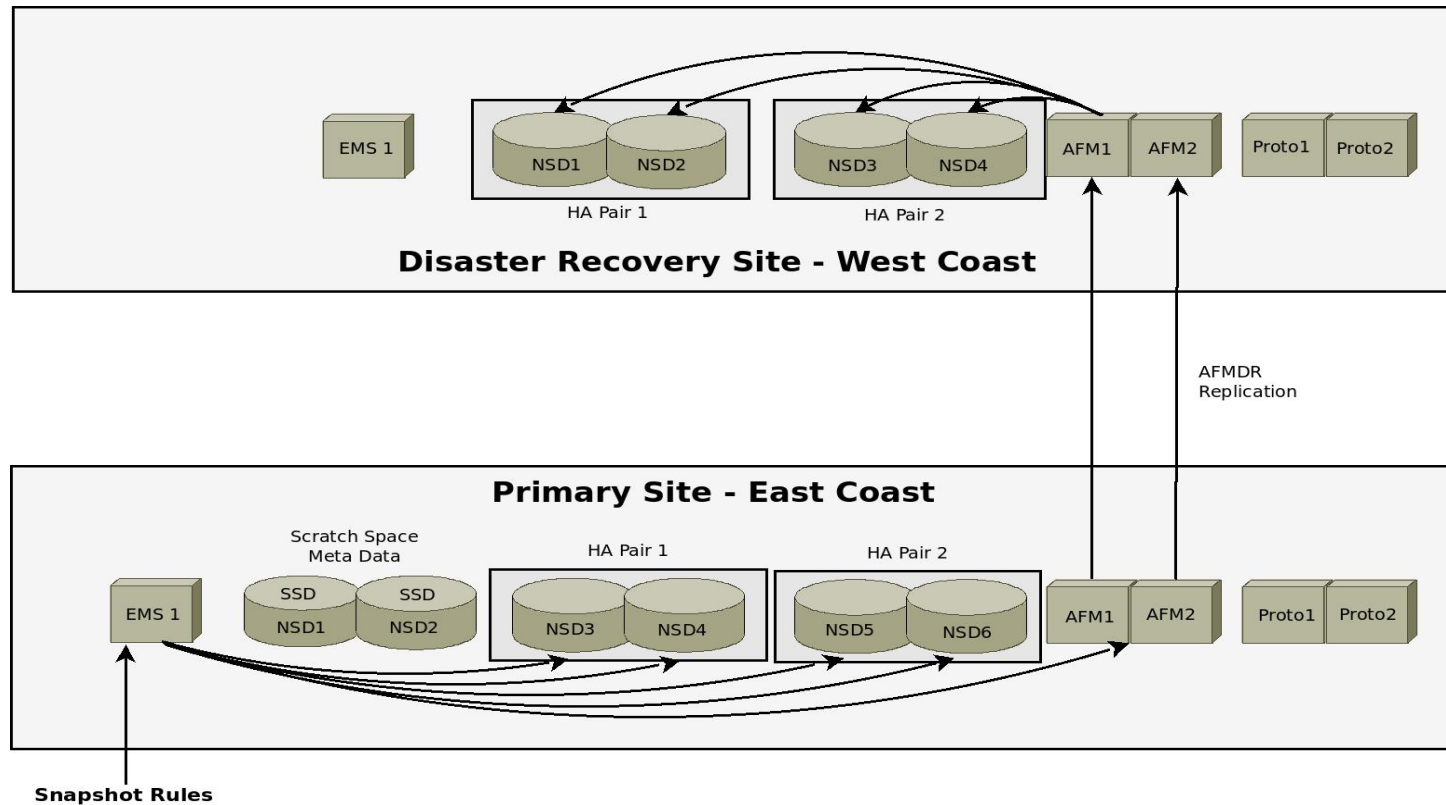
Constraints

- 72ms latency between sites
- NFS only over high latency network
- 100 million files per fileset
- Number of independent filesets
- No dependent filesets
- Time stamp sensitivity
- ID mapping
- Migration of 3PB of data

Optimizations & Solutions

- TCP Window Size for NFS
- Parallel Rsync with ID mapping > 3.1.2
- afmnumflushthreads
- Strict quotas on home with filesets per lab
- out of band replication initialization
- ID Mapping
- Migration of 3PB of Data

Snapshots



Snapshots

- HA pairs for redundancy
- Symlinks to sync snapshot schedule across all nodes
 - cron.d symlink to ces_root cronjob
- AFMDR replication to sync snapshots across clusters
- Snapshot Script Arguments:
 - File System, Fileset, duration
- Snapshot Wrapper Arguments:
 - Script Arguments + Primary Snapshot Node
- Snapshot Scripts on ces_root File System
- Snapshots balanced across NSD servers
- Incremental Snapshot schedule
- Reduction in administrative effort

Upcoming Adventures

- Spectrum Archive
- Quota Alerting
- Bright Health Scripts
- Auditing with Spectrum Scale 5.0.1



MCIT

QUESTIONS?





THANK YOU

