NATIONAL CLIMATE AGENCY SELECTS VSM FOR ARCHIVING

VSM manages large scale HPC climate data, archiving from Lustre filesystem

Versity

In this case study, learn about how VSM helped a national climate agency meet demanding high performance computing (HPC) archive performance requirements. The system currently archives at over 12GB/s to tape, moving one PB of data per day for one of the most active high performance archive solutions on the planet.

THE CHALLENGE

A national climate agency runs a variety of numerical and observation simulation experiments to accomplish their mission to improve weather forecasting and to gain greater understanding of the climate on a large temporal scale. Through their simulation endeavors the agency both receives and produces massive amounts of large file data that is used as part of their HPC workflow.

The file data must be available for their Cray HPC system, but because storing the entire data collection in the parallel filesystems is cost prohibitive, they needed a low cost archive solution featuring both capacity and performance. The solution needed to ingest files from their Lustre parallel filesystem quickly and needed to write 2 copies to tape at a speed of at least 10GB/s under everyday production conditions.

THE SOLUTION

The national climate agency selected Versity's VSM product for their archiving system. The industry leading software was an ideal fit for their performance, capacity, and total cost of ownership requirements.
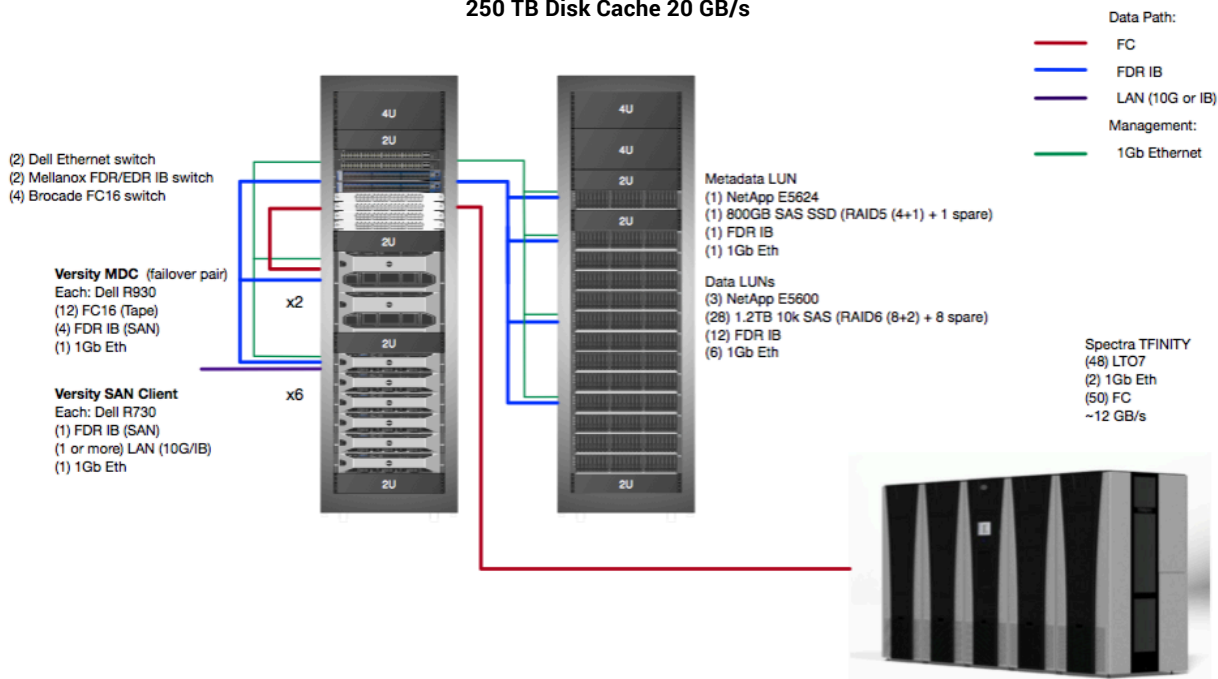
SOLUTION FOCUS

- Archiving from Lustre
- Climate and HPC sector
- High performance to tape

CHALLENGE

- High speed ingest
- Write multiple copies
- Write at least 10GB/s to tape under everyday production workloads

SOLUTION

- Versity Storage Manager (VSM)

KEY BENEFITS

- Extreme performance to tape
- Fast ingest
- Rich policy manager
- Sole focus on archiving

Environment Snapshot:

- **35 PB Capacity**
- **850 TB - 1 PB write throughput per day**
- **12 GB/s**
- **2 Servers (active/passive)**
- **6 SAN Clients**
- **48 LTO7 Drives**
- **Millions of Files**

With 48 LTO-7 tape drives on a Spectra TFinity series library and demanding ingest

DEPLOYMENT

48 LTO7 drive 12 GB/s

250 TB Disk Cache 20 GB/s

Data Path:
FC
FDR IB
LAN (10G or IB)
Management:
1Gb Ethernet

(2) Dell Ethernet switch
(2) Mellanox FDR/EDR IB switch
(4) Brocade FC16 switch

Versity MDC (failover pair)
Each: Dell R930
(12) FC16 (Tape)
(4) FDR IB (SAN)
(1) 1Gb Eth

x2

Versity SAN Client
Each: Dell R730
(1) FDR IB (SAN)
(1 or more) LAN (10G/IB)
(1) 1Gb Eth

x6

Metadata LUN
(1) NetApp E5624
(1) 800GB SAS SSD (RAID5 (4+1) + 1 spare)
(1) FDR IB
(1) 1Gb Eth

Data LUNs
(3) NetApp E5600
(28) 1.2TB 10k SAS (RAID6 (8+2) + 8 spare)
(12) FDR IB
(6) 1Gb Eth

Spectra TFINITY
(48) LTO7
(2) 1Gb Eth
(50) FC
~12 GB/s

and staging requirements, they required a product that was capable of scale and performance. They selected a company that understands serial devices and how to use them efficiently. Versity not only met the the agency's performance requirements, but exceeded the archive to tape speed requirement by over 2GB/s for a total of 12GB/s in performance. The data moves in and out of the archive on a regular basis, moving up to 1 PB of data per day.

Versity enabled the national climate agency to deploy the solution on commodity hardware. Their data cache was affordable and met their stringent workload and performance requirements. With Lustre mounted on several VSM nodes, the cache configuration is capable of ingesting data from their Lustre system at over 20GB/s. As files are no longer needed on Lustre they are transferred to the VSM cache where policies are automatically applied and archive copies are sent out to tape. The files are then recalled (staged) when HPC jobs are run that require the files to be moved back into the Lustre system.

Automatic tiering and a single interface for all archive data made archiving faster and easier for the national climate team. The agency appreciates the way cache space is managed automatically, relieving worry that the filesystem may get too full, freeing IT staff time for other tasks.

About the climate agency: The agency provides the best possible services in forecasting monsoons and other weather/climate events for its host nation. They also help with science and technology for exploration and development of ocean resources, and play an important role for Antarctic/Arctic and Southern Ocean research.

ABOUT VERSITY

Versity Software is the leading independent large archive software company. Redefined business practices set us apart from the legacy big company archive vendors. Discover what it's like to love your archive solution!

415.723.0949
info@versity.com
versity.com
@versitysoftware

100 Pine St, Suite 1250
San Francisco, CA 94111