CGG’s Storage Subsystem For Seismic Processing

T. Barragy, 3/2017

Passion for Geoscience
CGG Seismic Processing Challenges

- **CGG is a service company**
  - We are very cost sensitive
  - The type of processing is quite specific vs O&G majors
  - The data flows are specific to CGG proprietary SW

- **Data growth rates**
  - Double every 5 years

- **Data capacity needs are high**
  - Need 10’s of PB online at any given time

- **Data access needs are modest**
  - Modest bandwidth needs
  - Large transfer (bytes) dominated
  - Latency tolerant (throughput computing)
  - Specific types of datasets can be very hot (cacheable)
CGG Data Center

- **Seismic OS**
  Data & Job management, File & Utility services
  Hosted on commodity x86 HW

- **Storage Islands**
  Unix host + 10GbE + direct attach arrays of 250TB raw
  Several hundred in total, ~ 20 PB
  Data accessed via I/O proxies running on Unix host

- **CPU Compute Nodes**
  Commodity x86, GbE, HDD for scratch
  ~ 2 PFLOP (SP)

- **GPU Compute Nodes**
  Commodity x86 + GPU, GbE, SSD for scratch
  ~ 20 PFLOP (SP)

*Optimized for Small MPI Job Throughput & CGG Workflows*
Storage Island: Dell MD Series

- **R620-A & R620-B**
  - One logical SI (storage island) per host
  - Hosts are active:active with failover
  - Also runs proxy daemons serving data to compute nodes

- **Each logical SI consists of**
  - A single (Posix) filesystem
  - Linux md stripe across 3 volumes, each RAID6/8+2

- **Do Not Lose Data, But Keep Cost Down**
  - Dual 10GbE connectivity
  - Dual port SAS drives, controllers

- **Later versions**
  - Add drive arrays, add logical SI
  - Each logical SI is in its own Linux Container
Data Growth

- Enormous increase in data volumes over ~ 20 years
  Increased geophone density, array size, source offset, sample rates

- Raw survey size doubles every ~5 years

- Intermediate datasets are ~10x of raw datasets
  Raw datasets ~100TB
  Intermediate datasets ~ 1PB
Data Lifetime

- Neither Enterprise nor HPC (checkpoint)
  - Enterprise might live years
  - Checkpoints might live a few hours
  - 30% lives 1 months or less
  - 50% lives 1 – 4 months

- Average lifetime ~ a few months
  - Signal Processing: Read / Write chain
  - Velocity Modelling: Read / Write loop, low res datasets
  - Final Migration: Read / Write, high res datasets
Data Temperature

- **Monte Carlo Profile of /proc**
  - Covers 6 months / 50K samples
  - 300 second sample windows

- **Average Temp – ‘Cool’**
  - Average bandwidth is 1/10 of HW peak
  - Individual datasets can be very hot (cacheable)
  - R:W ratio is ~ 2:1 on average

- **Motivated adding arrays to SI**
  - Brings CPU, disk, network into better balance

Note: axis scale is obfuscated
Data Access

- **Linux Block Layer Trace Method**
  Reconstruct temporal trace to get user level request size

- **Request Sizes**
  50% of requests are small – latency / seek dominated
  50% of requests are large – data transfer dominated

- **Request Bytes Read**
  99% of bytes read are for large multi-MB requests

- **Is an ‘object store’ feasible?**
  Needs POSIX file system support
  Needs to be reasonably efficient for small files / requests
  Underlying object sizes of a few MB are suggested

Note: axis scale is obfuscated
Object Store

- **Motivation**
  - Primarily cost
  - Secondary is improved data protection & etc

- **A Work In Progress**
  - Currently testing one implementation
  - DSS7000 target building block

- **Object Store Characteristics**
  - Increased latency (5x) for individual requests
  - Excellent (linear) scaling with client load

- **HW RAID Characteristics**
  - Low latency at low client load
  - Performance degrades with increasing client load

Note: axis scale is obfuscated
Thank you
Backup & Older
Marine Seismic Processing

- **Processing Flow**
  - **Signal Processing**
    - Sorting / Binning, Filtering / Conditioning / De-Noising
    - Several months of calendar time, significant interactive analyst time
  - **Velocity Modelling**
    - Migration on coarser or localized mesh
    - Many iterations, adjust velocity model each time
    - Several months of calendar time, lots of interactive analyst time
  - **Final Migration**
    - Produces the subsurface image
    - Several weeks of calendar time

- **Reverse Time Migration**
  - Forward wave eqtn on initial pulse (source)
  - Backward wave eqtn on receiver data
    - Cross correlate these two wavefields in time dimension
    - Yields an image in x,y,z for a single source & receiver group
  - Sum images over source / receiver groups
    - Yields the final 'stacked' image in x,y,z

\[ \frac{\partial^2 u}{\partial^2 t} - c^2 \nabla^2 u = f \]
Storage Profile

- **Data Lifetime: Mixed**
  - 80% < 4 months
  - 70% > 1 month

- **Data Temperature: On Average, Cool**
  - SI average load is 1/8 of HW capability
  - Individual datasets can be very hot (cacheable)

- **Data Access: Bimodal Request Size**
  - 50% ‘small’ - latency / seek dominated
  - 50% ‘large’ - bandwidth / transfer dominated
CGG: O&G Service Segment

- Sercel
  - Started in 1956
  - Produces seismic sensors

- Marine/Land
  - Started in 1931
  - Acquires seismic data

- Subsurface Imaging
  - Processes & interprets data
  - Very cost sensitive
  - Every compute job is revenue
Acquisition: Marine Surveys Over Time

~1995 'NAZ'

~2000 'MAZ'

~2005 'WAZ'

~2010 'FAZ'

StagSeis / BroadSeis
An Integrated Geoscience Company

Full range of products and clear market leadership onshore, offshore and downhole:
- Technology leadership
- Large installed base
- A cornerstone for CGG integrated solutions

Full range of seismic and other geophysical methods for acquisition:
- Marine
- Land
- Multi-Physics
- Seabed*

Subsurface Imaging
- GeoConsulting
- GeoSoftware
- Multi-Client & New Ventures
- Data Management Services

*Seabed Geosolutions joint venture 60%-owned by Fugro and 40% by CGG
Optimal low frequencies are achieved by reducing the zero Hz ghost notch as much as possible for both source and receivers. The deeper the hydrophone tow depth the better the low frequency signal. Improved low frequency signal-to-noise. No compromise on high frequencies.