Challenges of Real-time Processing in HPC Environments – the ASKAP experience

Eric Bastholm | Team Leader
19 October 2017
Outline

• Australian SKA Pathfinder (ASKAP)
• Some specific computing challenges encountered
  • Disk Performance
  • Application Performance
  • Process Isolation
• Lessons learned
• Takeaways
ASKAP: Australian Square Kilometre Array Pathfinder
ASKAP

- 36-antenna interferometer
- Located in radio quiet zone “outback”
- Supercomputer for data processing
- 10,000 cores, 4 GB/core, 200 TFLOPs Peak
- Data Ingest ~ 2.8 GB/s = ~ 10 TB/h
- Cyclic disk buffer, deleting data post processing
- 5 PB data products per year

image credit: Alex Cherney / terrastro.com
Challenges of Real-time Processing in HPC Environments

Eric Bastholm
Challenges of Real-time Processing in HPC Environments

Eric Bastholm
Challenges of Real-time Processing in HPC Environments

Eric Bastholm
ASKAP

Challenges of Real-time Processing in HPC Environments

Eric Bastholm
Computing Challenges:
Computing Challenges

• Disk performance
• Application performance (memory and messaging)
• Process isolation
Computing Challenges
Computing Challenges:
Disk Performance
Disk Performance

• High performance parallel Lustre file system
• 1 PB designed for high throughput (10 GB/s)
• Single thread performance – not so much
• Performance can vary a lot depending on
  • Other users
  • System parameters
  • Legacy code (non-parallel)
Disk Performance

Management Server (MGS)
Metadata Server (MDS)

Management Target (MGT)
Metadata Target (MDT)

Co-located MGS and MDS share storage

Lustre clients

Ethernet or InfiniBand Network

Object Storage Servers (OSSs)

OSS 1

OSS 2

doc.lustre.org
Disk Performance

![Image of Disk Performance graphs]

- Chart 1: RMSink writing times per integration
- Chart 2: Distribution
Disk Performance

Challenges of Real-time Processing in HPC Environments

Eric Bastholm
Disk Performance
Disk Performance
Computing Challenges: Application Performance
Application Performance

• Locking in messaging (MPI) code between parallel processes
• Low memory transfer rate
• Causes buffer overruns
• Lost data
• Write more files from separate processes
Application Performance

Ingest durations + buffers

- IngestD BufferUsagePercent
- IngestD MSWritingDuration
- IngestD VHCornerTurnDuration
- IngestD SourceTaskDuration
- IngestD ProcessingDuration
- Ingest36 ProcessingDuration
Application Performance
Computing Challenges:
Process Isolation
Process Isolation

• Isolate the critical processes as much as possible
• Dedicate resources they need
• Control read/write access to disk
• Reduce OS jitter by minimising installed image
• Pure isolation not possible in a shared environment
Process Isolation
Lessons Learned:
Lessons Learned

• Don’t assume that the platform is infallible just because it’s high tech, BIG, and FAST
• Identify possible performance issues
• Prototype at a granularity necessary to enable the important design decisions – walk before you try to run
• Use realistic data sets
• Establish good working relationship with the platform providers
• Develop monitoring and reporting early
• Develop good testing early
Takeaways:
3 Takeaways

- Prototype and integrate often, don’t implement final solution in one leap
- Monitor everything
- Work with platform provider – make them part of the team
We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site.

Thank you

CSIRO Astronomy and Space Science
Eric Bastholm
Team Leader ASKAP SDP

+61 8 6436 8505
eric.bastholm@csiro.au
www.csiro.au/Research/Astronomy