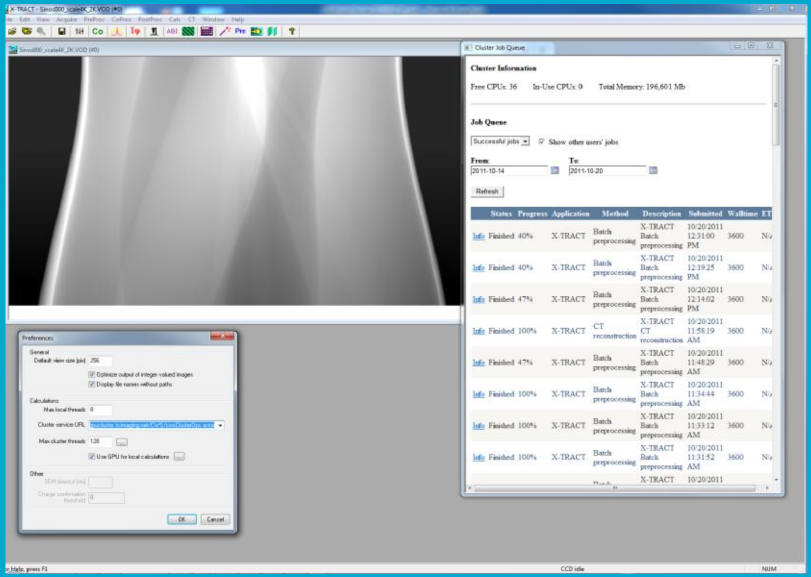


X-ray imaging software tools for HPC clusters and the Cloud

Darren Thompson | Application Support Specialist
9 October 2012

NeAT Remote CT & visualisation project

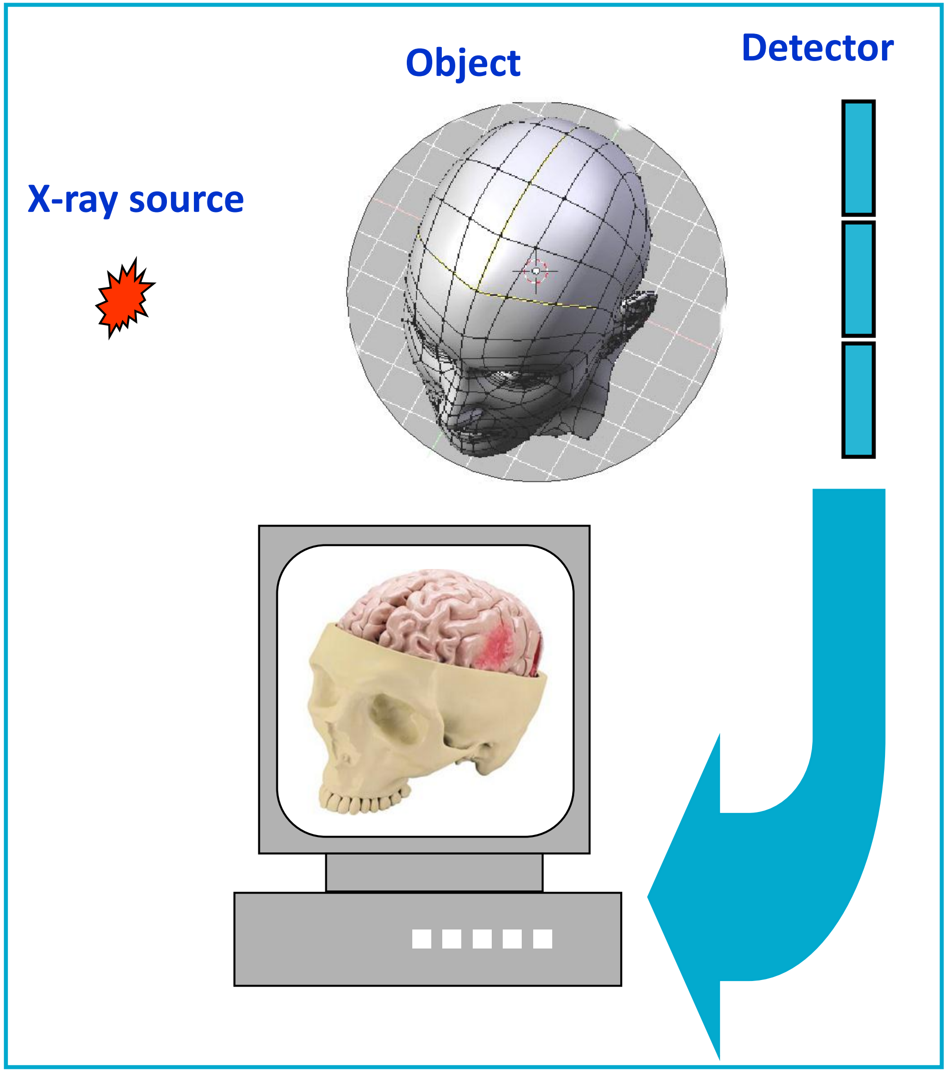
- Aim: to develop a user-friendly service for remote 3D CT reconstruction, modelling and visualization running on CPU/GPU computer clusters at the Australian Synchrotron ("MASSIVE") and at the micro-CT facility at the Australian National University (ANU).
- Team members: CSIRO, Australian Synchrotron, Australian National University, Victorian e-Research Strategic Initiative (VeRSI), Victorian Partnership for Advanced Computing (VPAC).
- Collaborators: Monash Uni e-Research Centre, Monash Uni Centre for Synchrotron Science.
- Start date: 20 August 2009; End date: 30 June 2011

NeAT Remote CT & visualisation project

- Services for rapid CT simulation and reconstruction from large datasets
- Services for efficient transfer of large datasets to and from the remote computational facilities
- Services for remote 3D visualization and collaboration
- Projected impact: the service is expected to increase the productivity of Imaging & Medical Beamline at the AS by at least 25%, which equates to an estimated of ~$590,000 per year.

Principles of X-ray Computed Tomography (CT)

- Typical experimental set-up involves an X-ray source, a sample on a rotation stage and a 2D position-sensitive detector
- Images of the sample are collected at many different rotation angles spanning 180 or 360 degrees
- Acquired images are processed in a computer to produce a 3D representation of the internal structure of the sample



CT's "challenges"

- Computationally intensive!
  - CT Reconstruction is $O(N^4)$
- Data intensive!
- Infeasible to compute/store large datasets from Synchrotrons and lab equipment on single PC's
- On the plus side – highly parallelizable!
  - GPUs

CT reconstruction speed-up using GPUs

Volume	CPU (1 thread)	CPU (4 threads)	CPU+GPU (1 thread)	CPU+GPU (4 threads)
1024^3 voxels	9h 6' 2"	2h 25' 9" (3.76×)	5' 42" (95.8×)	2' 56" (186×)
2048^3 voxels	161.7h	40.5h (3.995×)	1h 18' 14" (124×)	41' 53" **(232×)**

Typical CT reconstruction data sizes

N / M*	N^2 float (projection / slice)	NM float (sinogram)	N^2M float (all sinograms)	N^3 float (all slices)
1k / 720	4 MB	2.8 MB	2.8 GB	4 GB
2k / 1,440	16 MB	11¼ MB	22½ GB	32 GB
4k / 2,880	64 MB	45 MB	180 GB	256 GB
8k / 5,760	256 MB	180 MB	1.4 TB	2 TB
16k / 11,520	1 GB	720 MB	11¼ TB	16 TB

* N is the linear size of a projection/slice
M is the number of projections

CSIRO

X-TRACT

- Windows GUI application for X-ray image processing with > 10 years development and refinement
- Supports both local & remote (Windows HPC cluster) processing modes
- Specialized X-TRACT console “worker” application executes on cluster compute nodes
- Implements multiple levels of parallelization, cluster nodes -> CPU cores & threads -> GPUs (CUDA)
- CWS (Cluster Web Services) ASP.NET webservice based system developed to as the external gateway between X-TRACT clients and Windows HPC clusters.
- Free 30 day trial accounts available at https://ts-imaging.net/Services/SignUp.aspx

X-TRACT

Core functionality

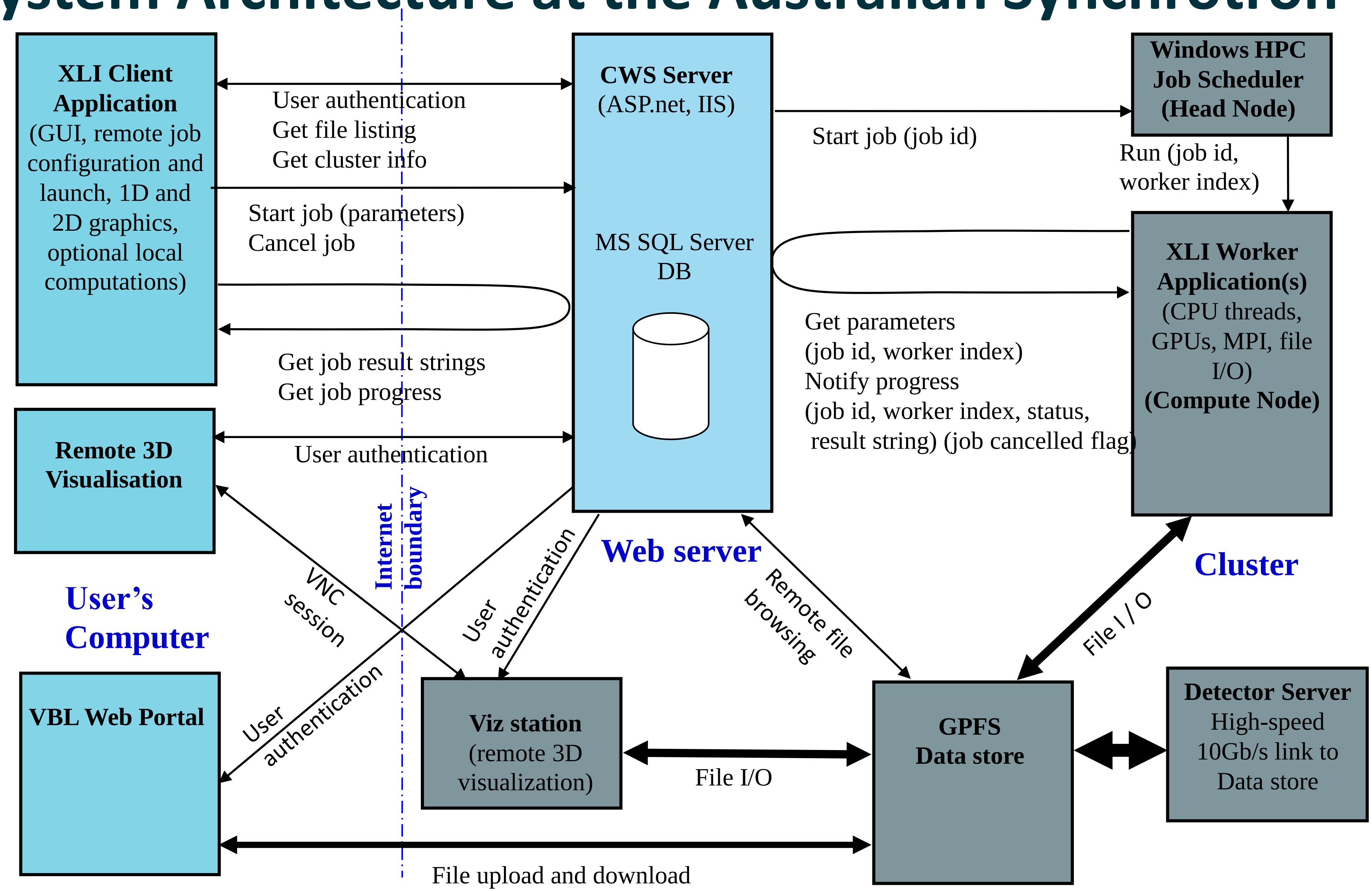
- Pre-processing
  - Dark current correction, flat field correction, CCD defective pixel replacement, image drift correction, beam hardening and sinogram creation
- Co-processing
  - Automated image matching of magnification and drift or rotation and drift for pairs of images
- Phase retrieval
  - >20 algorithms for phase and/or amplitude extraction from in-line X-ray images
- Convolution/Deconvolution
  - Image filtering, super-resolution, estimating X-ray source size and spatial resolution

X-TRACT

Core functionality

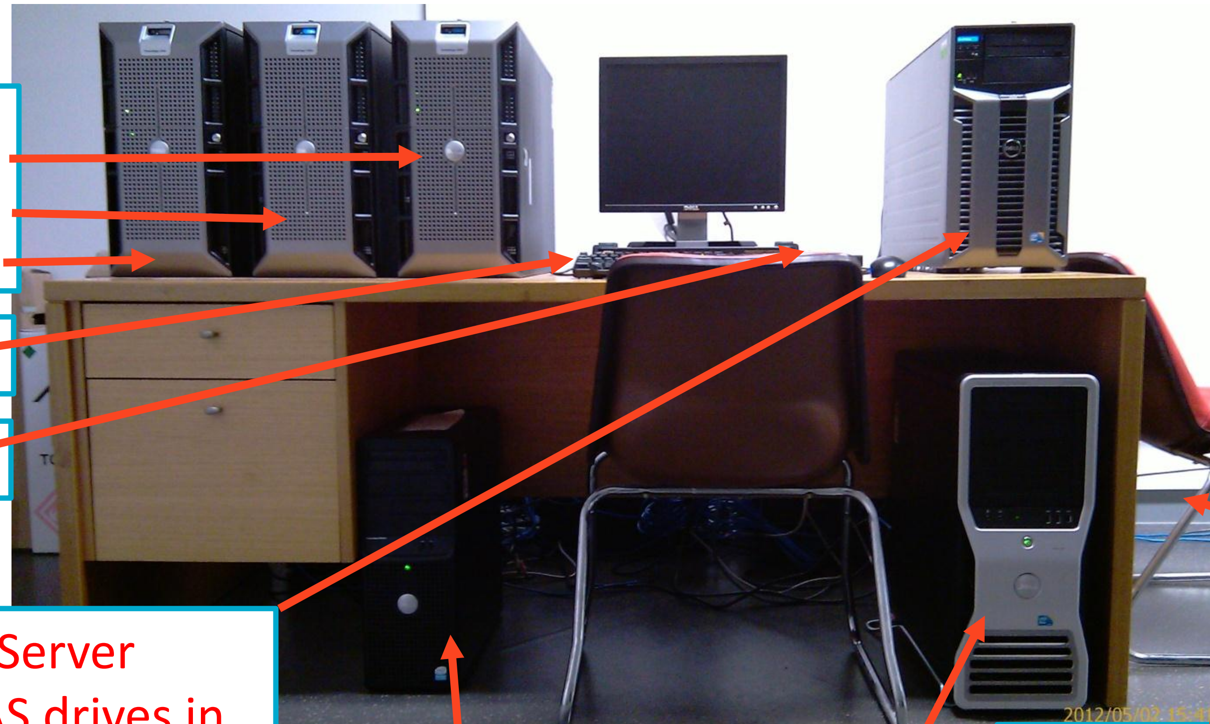
- Image Calculator
  - >50 major operations (e.g. for summation, division, rotation, padding, interpolation, etc, of images, Fast Fourier Transform, evaluation of Kirchhoff integrals, simulation of Poisson and Gaussian noise, spatial filtering, etc.)
- CT-Reconstruction
  - Parallel-beam FBP, Iterative parallel-beam (new), and cone-beam FDK reconstruction algorithms, simulation of CT projections. GPU and CPU implementations.
- ABI (analyser-based phase retrieval)
  - Simulation of ABI images and multiple methods of amplitude/phase reconstruction from experimental ABI images.
- OMNI Optics
  - Simulation of multiple phase-contrast imaging modalities

System Architecture at the Australian Synchrotron



CSIRO

CSIRO TBI "Minicluster"



3 x Dell PowerEdge 2900 compute nodes (2 x quad-core Xeon CPU in each)

1 x 6-port KVM switch

1 x 8-port Gigabit switch

1 x Dell PowerEdge T710 File Server node (6 x 125 GB 15k RPM SAS drives in RAID 0 configuration)

1 x Web server

5 x APC 1500 Smart UPSs

In total:
36 Xeon CPU cores across 4 compute nodes
192 GB of RAM across 4 compute nodes

1 x Dell Precision 7500 GPU node (2 x 6-core Xeon CPUs, 2 x NVidia GTX470 GPUs)

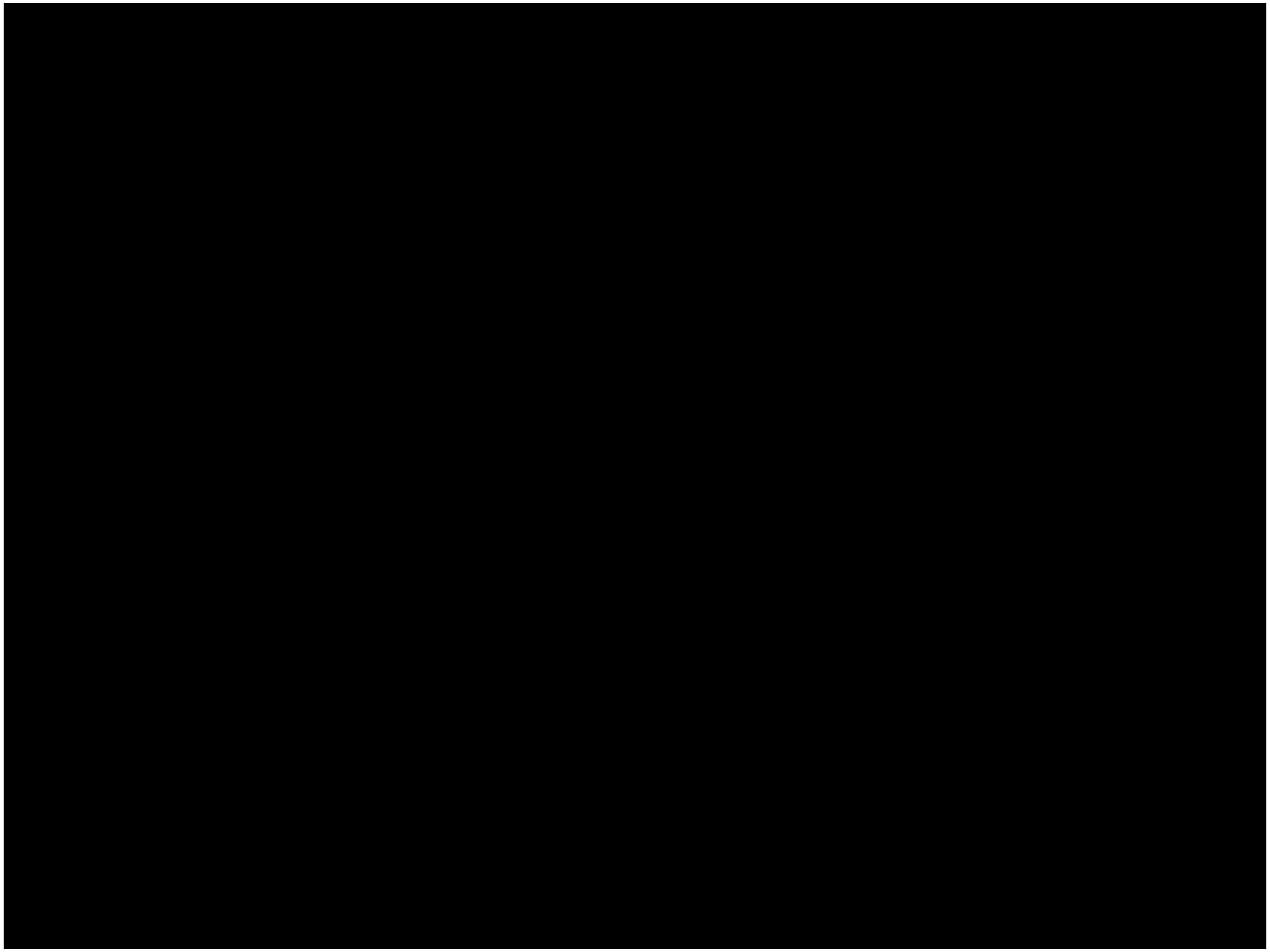
Large Windows HPC cluster deployments

CSIRO Bragg Cluster	MASSIVE-1 cluster at the Australian Synchrotron
• 2048 Sandy Bridge CPU cores • 128 GB RAM per node (16,384 GB RAM total) • 384 NVIDIA M2070 GPUs • 40 Gb/s Infiniband Interconnect • 80 TB high-performance local storage (HNAS)	• 504 Xeon CPU cores • 48 GB RAM per node (2,016 GB RAM total) • 84 NVIDIA M2070 GPUs • 4x QDR Infiniband Interconnect • 58 TB of fast access parallel file system (GPFS)

Deployment & operational considerations

- Requires externally accessible webserver (Microsoft IIS) and SQL server for hosting CWS
- Remotely accessible file storage system, ie FTP
- Site specific CWS customizations for required for non-standard Windows deployments
- Minimize overall system complexity – unify where possible
- Coexistence with existing and new systems & policies

X-TRACT in action on MASSIVE cluster

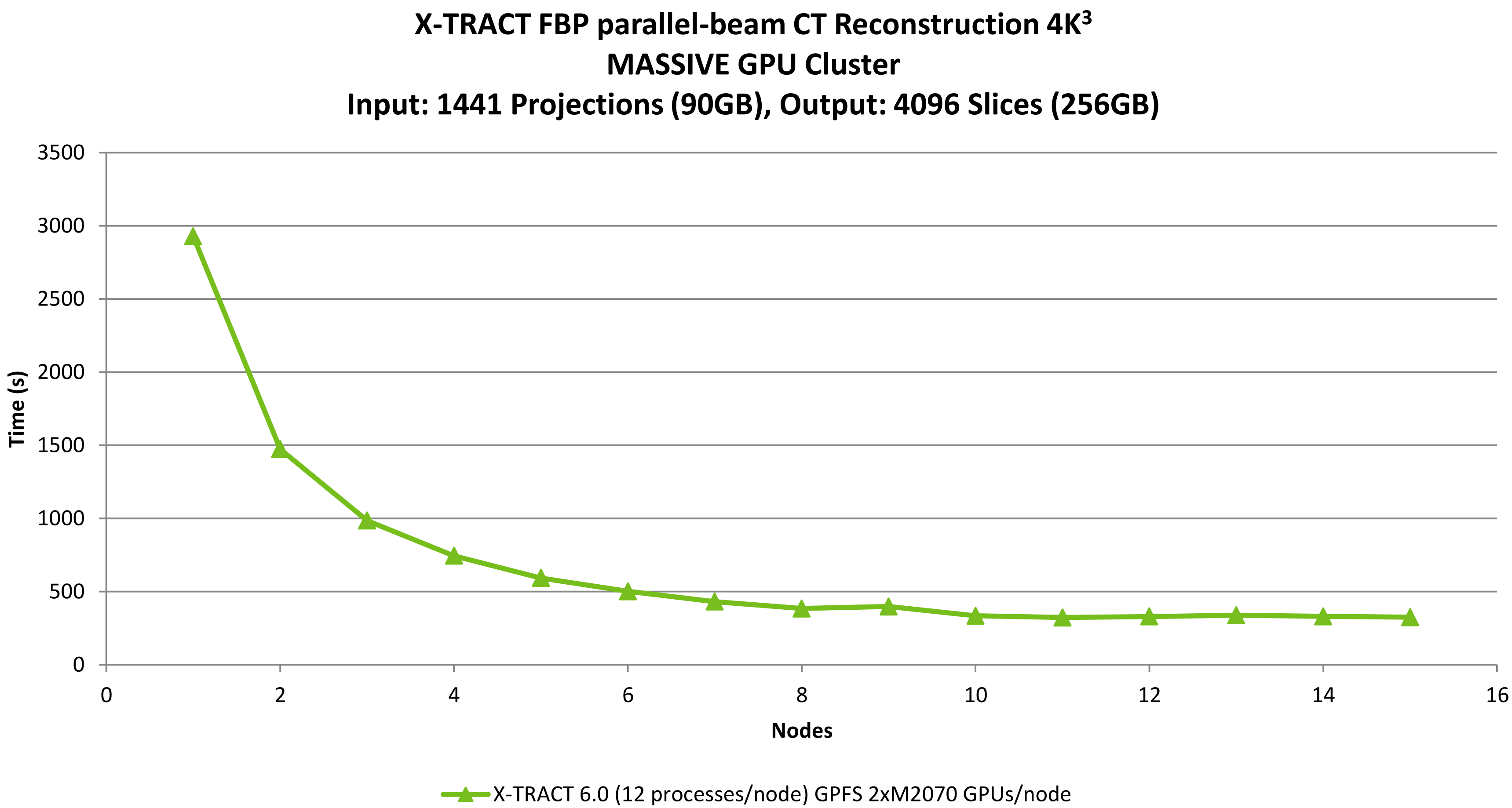


2048^3 reconstruction volume

Input: ~11GB

Output: ~32GB

Performance



Current development and future directions

- Actively developing a “streamlined” CT UI specifically for AS users
- Regular engagement with AS to improve/refine system components
- Deployment of Visualization server

- Currently porting X-TRACT’s computational backend to platform independent code
  - Demand and desire to support both Windows HPC & linux clusters
  - Windows only client and CWS retained
  - CPU only, GPU FBP CT reconstruction & pre-processing modules implemented for linux, ~10% done

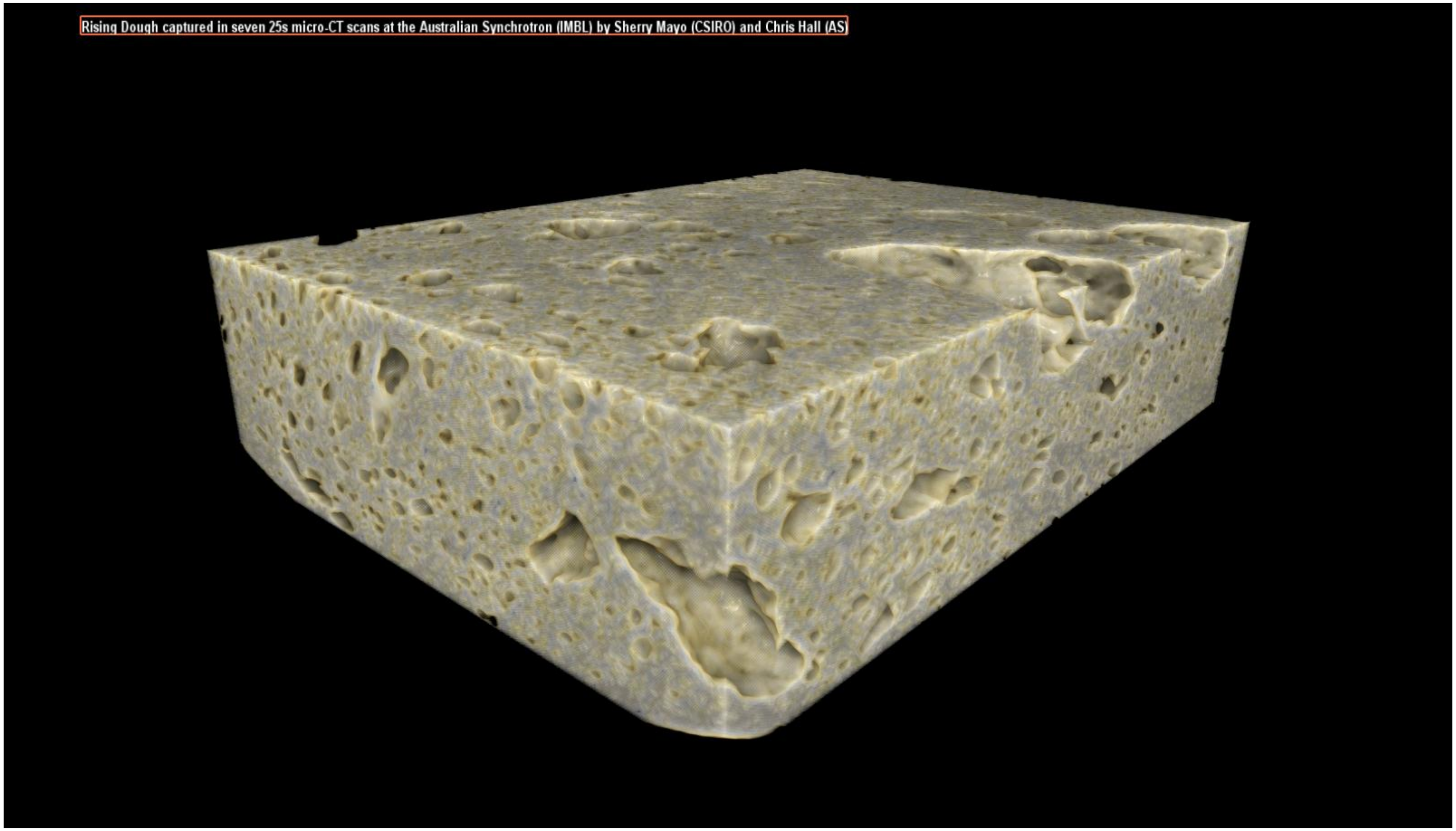
Current development and future directions...

- Explore the use of "volume" based data storage - i.e. HDF5.
  - Adoption/development of tomography storage/metadata standards
- Investigate addition of X-TRACT pre-processing to acquisition phase
  - Refined processing pipelines
  - Possible elimination of IO steps
- Integration of fast local and remote/distributed "volume rendering" into X-TRACT UI
  - Input region of interest selection
  - Output viewing

X-TRACT on the Cloud

- Developed a “proof-of-concept” Azure X-TRACT implementation
  - Adapted “backend” cluster components for use with Azure
  - Web UI for data upload/download and job creation and monitoring
- CSIRO NeCTAR Cloud Based Image Analysis and Processing Toolbox project
  - 2 year project
  - Allow for the construction and executing of imaging “workflows”
  - Newly developed linux modules to provide CT tools
  - Investigating use of Galaxy for first prototype
  - Possible use of CSIRO Bragg cluster and/or MASSIVE for GPUs and high-speed interconnect??

4D CT – Rising dough



7 sequential scans, each 25s long, ~5 min apart. Sample ~1.3cm, voxel size ~13um

Our Team

- Tim Gureyev, CSIRO CMSE
- Yakov Nesterets, CSIRO CMSE
- Darren Thompson, CSIRO IM&T ASC
- Alex Khassapov, CSIRO IM&T

- Special thanks to Sherry Mayo, CSIRO CMSE

Thank you

IM&T ASC
Darren Thompson
Application Support Specialist

t +61 3 9518 5940
e darren.thompson@csiro.au
w www.csiro.au