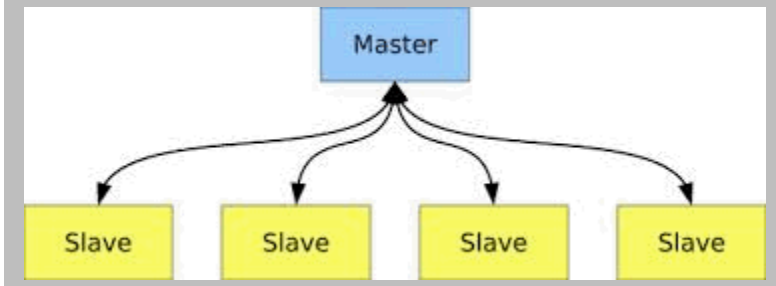
SAND2014-19596C



Cluster-Based Approach to a Multi-GPU CT Reconstruction Algorithm

Laurel J. Orr, Edward S. Jimenez, and Kyle R. Thompson

Controlling, Data Monitoring and Concurrent Processing

IEE Nuclear Science Symposium 2014

November 11th, 2014

U.S. DEPARTMENT OF ENERGY | NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Image Source: www.4nsi.com/assets/files/3dscan-nsi.pdf; www.flickr.com/photos/chrisdag/3707257190/; mit-crpg.github.io/openmc/methods/parallelization.html

What is Computed-Tomography?

- Computed-Tomography (CT) takes a series of X-ray images around an object to generate a 3D approximation of the object's interior and exterior structure
- Input: X-Ray images and scan geometry
- Output: series of cross-sectional image planes which approximate the volume

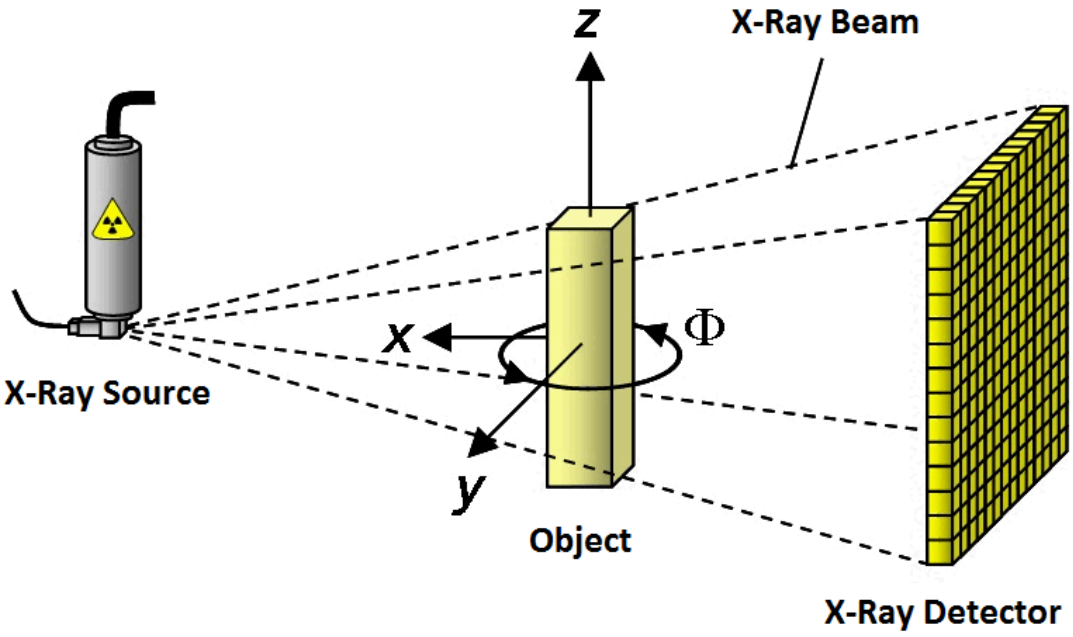


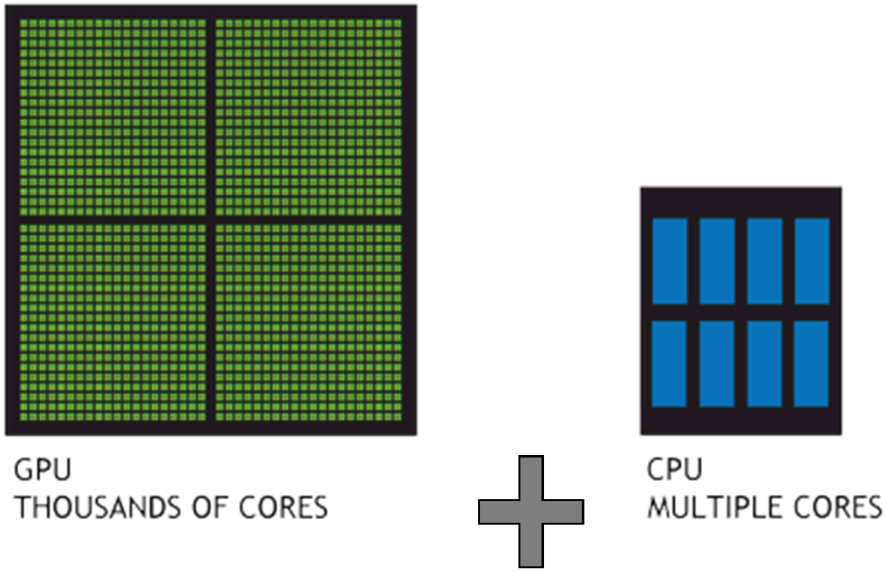
Image Source: http://www.x-ray-worx.com/joomla/index.php/en/microfocus-x-ray-tubes-for-non-destructive-material-testing-of-x-ray-worx/special-x-ray-tubes-for-ct/industrial-computed-tomography.html

Graphics Processing Units (GPUs)

Sandia National Laboratories

Specialized processors originally designed for image processing and visualization applications

- Massively parallel architecture
- Fast read-only caches
- API extensions



GPGPU combines the processing power of CPUs and GPUs to greatly increase performance of parallel problems

Orders of Magnitude Speed Up

Image Source: http://www.nvidia.com/object/what-is-gpu-computing.html

Industrial Reconstruction Limitations

Sandia National Laboratories

PROBLEM

Industrial CT datasets are on the order of Gigabytes

+

On high performance CPU, reconstruction can take days

+

100-Megapixel X-ray images (Terabyte datasets) may soon be possible

=

Infeasible Reconstruction

SOLUTION

Parallelize reconstruction utilizing GPUs

+

Modularized algorithmic approach to reconstruction by Orr and Jimenez*

+

High performance computer

=

Infeasible Reconstruction

NEW PROBLEM

Reached limit of one node's capabilities

+

Hardware specifications unrealistic for some users

=

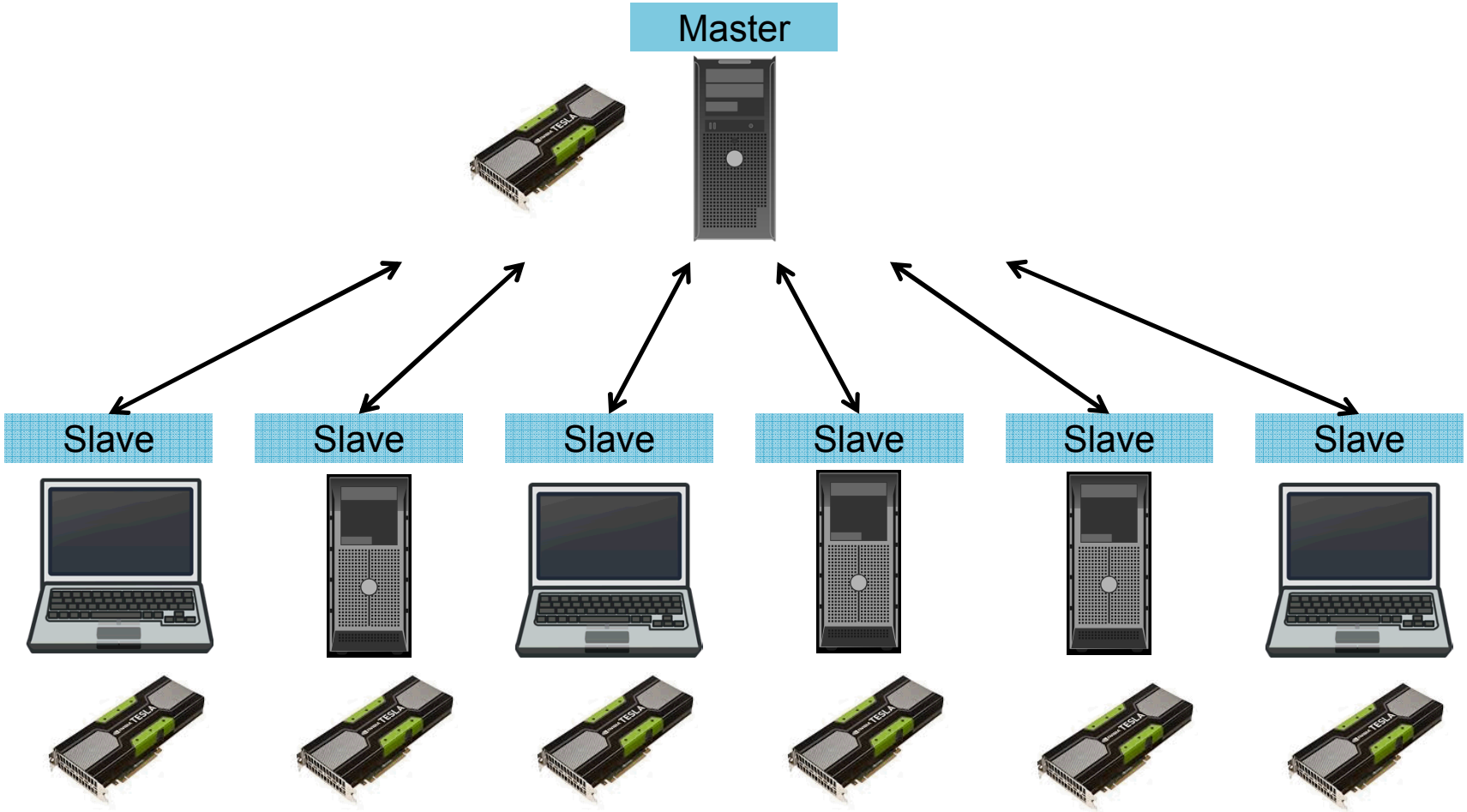
Still Infeasible Reconstruction

SOLUTION ?

* See SPIE Optical Engineering + Applications 2013 and NSS-MIC 2013

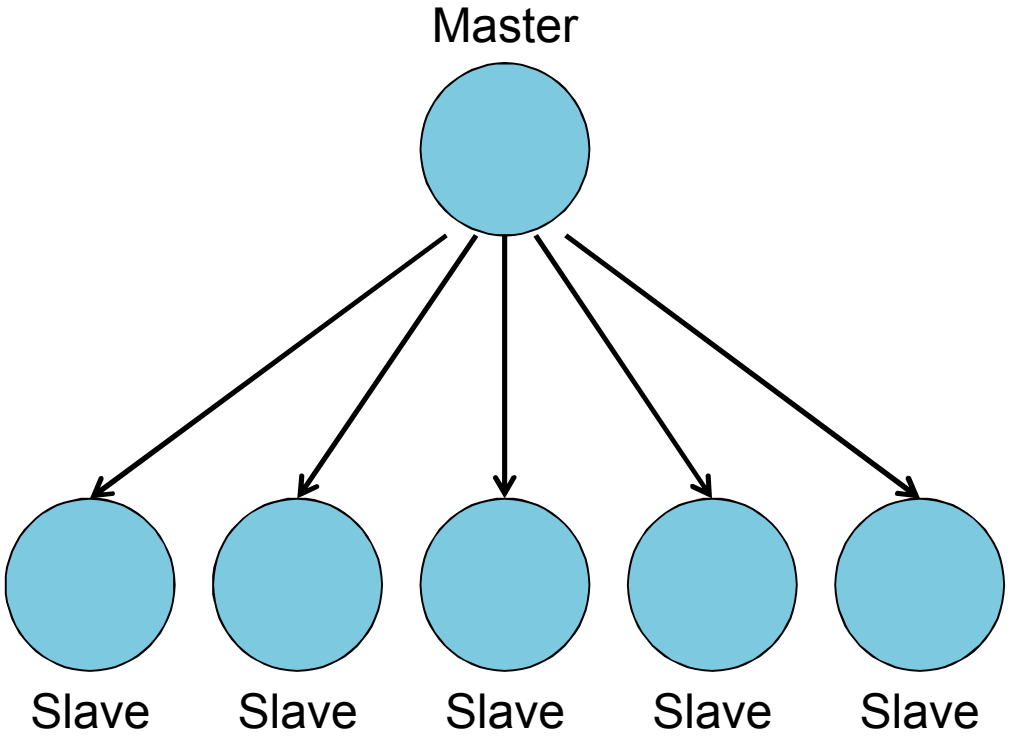
Solution

Goal: improve performance and accommodate a variety of workstations and resources (exploit GPU-capable systems already in NDE lab)



Handling I/O

Typically, master node handles all data transfer

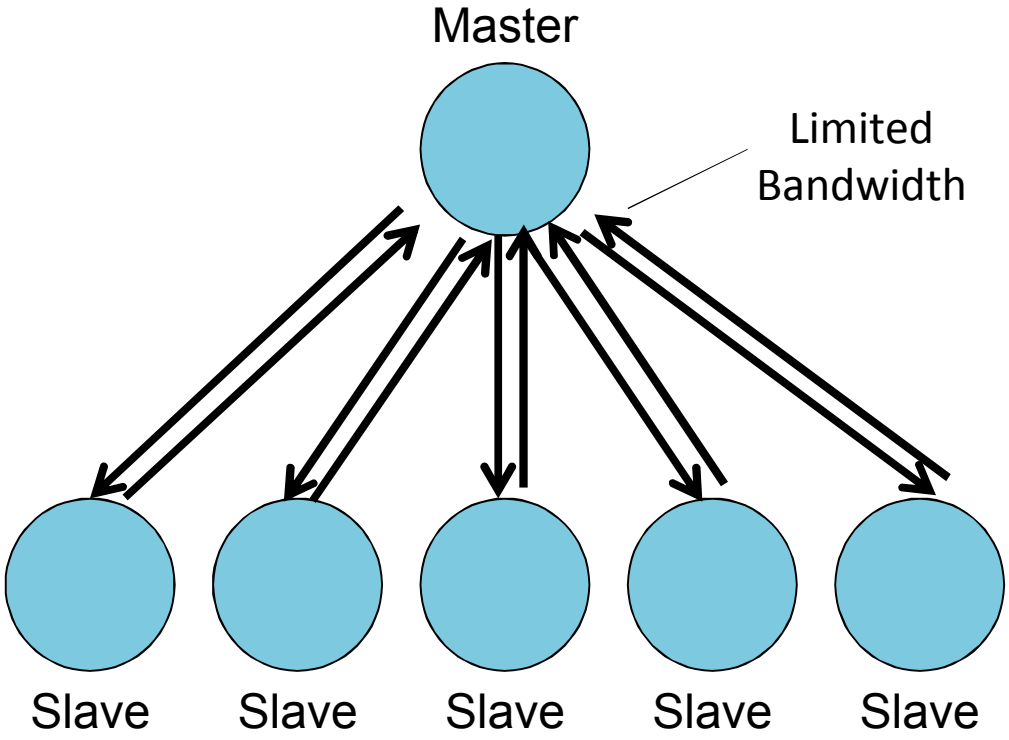


Initially, data transfer to slaves is unhindered and efficient

Beginning of Reconstruction

Handling I/O

Typically, master node handles all data transfer



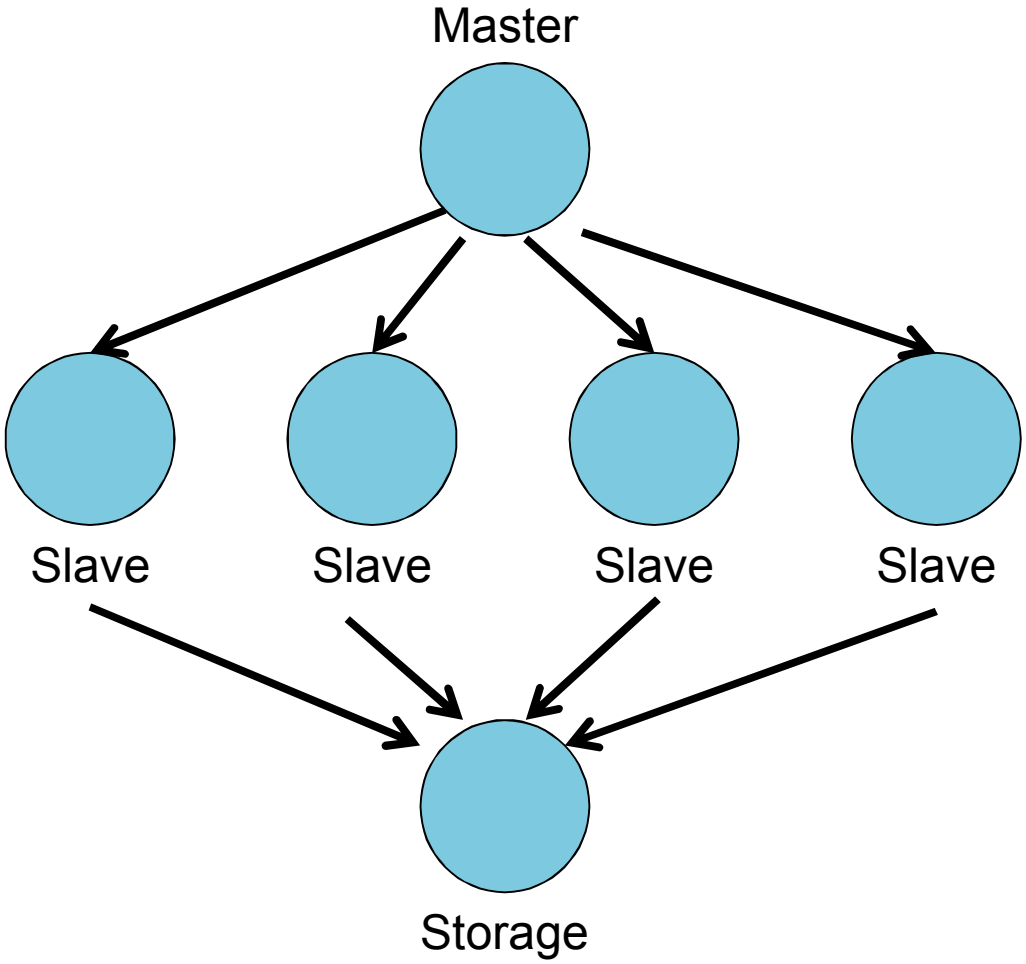
When slaves finish reconstructing subvolumes and start transferring data back to the master, the data I/O becomes congested.

GPUs sit idle waiting for more data which hurts performance.

Middle of Reconstruction

Handling I/O

Split data I/O between two nodes



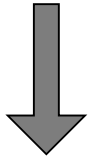
Specify one node to handle all data transfer to the slaves and another to handle data transfer from the slaves.

Load Balancing

Sandia National Laboratories

For most efficient GPU usage, must take into account:

- Transfer in contiguous blocks of X-rays/reconstructed slices
- Heterogeneous cluster: different node compute capabilities
- Irregularity in ratio of X-rays images to reconstructed slices



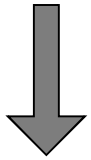
Cannot uniformly distribute input across nodes!

Load Balancing

Sandia National Laboratories

For most efficient GPU usage, must take into account:

- Transfer in contiguous blocks of X-rays/reconstructed slices
- Heterogeneous cluster: different node compute capabilities
- Irregularity in ratio of X-rays images to reconstructed slices



Before reconstruction, each slave sends to master:

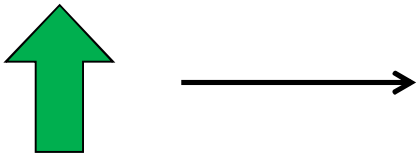
- Number of GPUs
- Ratio of host memory to total GPU memory
- Total number of GPU multiprocessors
- Lowest GPU compute capability

Master then uses non-linear weighting scheme and information on reconstruction size to dynamically partition volume across slaves.

Load Balancing

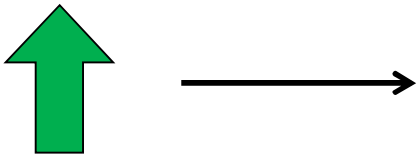
Sandia National Laboratories

Total Number GPUs



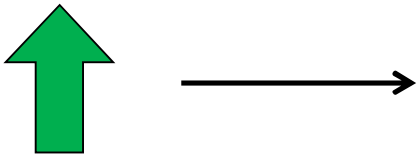
More processing power so is able to reconstruct slices more quickly

Host Mem / Total GPU Mem



More data transfers from host to GPU can occur before host must get more data over the network (non-linear increase in performance)

Total Number GPU Multiprocessors



More CUDA cores to use during computation so able to reconstruct more quickly (less importance than number GPUs)

Lowest GPU compute capability

With higher compute capability, incremental increase in performance power

Load Balancing

Sandia National Laboratories

Before reconstruction, master receives node information and configures cluster to determine what percentage of the reconstructed volume each node should compute.

EXAMPLE (10,000 volume slices to reconstruct, all GPUs same amount free mem)

Node 1	Node 2	Node 3	Node 4	Node 5
Num GPUs	3	3	4	2
Mem Ratio	12	4	12	8
Total MP	32	32	52	12
Lowest CC	2.0	2.0	2.0	2.0
Volume %	29.3	19.5	36.1	15.1

Experimental Set-Up

Sandia National Laboratories

- Cluster
  - 5 nodes: 1 master and 4 compute slaves
  - Nvidia Tesla m2090 (x8), c2070 (x4), and c2050 (x2) GPUs
- Teravoxel synthetic dataset (10,000 100 Megapixel Projections)
- C++ Code
  - Compiled using Visual Studio 2008
  - OpenMP Version 2.0
- CUDA Version 5.0

Results

Sandia National Laboratories

Cluster Configuration (Master is separate node)

	Node 1	Node 2	Node 3	Node 4
Num GPUs	**5**	**1**	**3**	**4**
Mem Ratio	**5.8**	**21.6**	**7.9**	**29.9**
Total MP	**74**	**14**	**46**	**64**
Lowest CC	**2.0**	**2.0**	**2.0**	**2.0**
Volume %	**32.1%**	**10.2%**	**24.3%**	**33.3%**

Results

	Uniform	Weighted
1 Node I/O (Master)	**> 30hrs**	**14.1hrs**

Results

Sandia National Laboratories

Cluster Configuration (Master is separate node)

	Node 1	Node 2	Node 3	Node 4
Num GPUs	**4**	**3**	**3**	**4**
Mem Ratio	**8.4**	**7.4**	**7.9**	**29.9**
Total MP	**56**	**46**	**46**	**64**
Lowest CC	**2.0**	**2.0**	**2.0**	**2.0**
Volume %	**25.4%**	**21.4%**	**20.7%**	**32.3%**

Results

	Uniform	Weighted
1 Node I/O (Master)	**> 30hrs**	**14.2hrs**
2 Node I/O (Master & Node #)	**> 30hrs**	**12.9hrs**

Conclusions

Sandia National Laboratories

- GPUs dramatically improve the performance of industrial CT reconstructions
- Taking a cluster-based approach further improves runtime and allows for reconstruction on less specialized hardware
- Prepares scientific community to efficiently process future-sized datasets
- Future work:
  - Allow the data to be stored across the cluster compared to all on one node
  - Reconstruction on the cloud