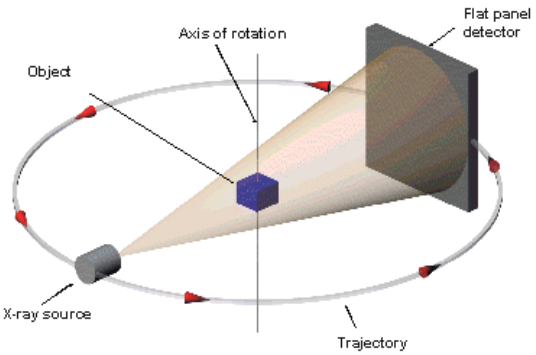
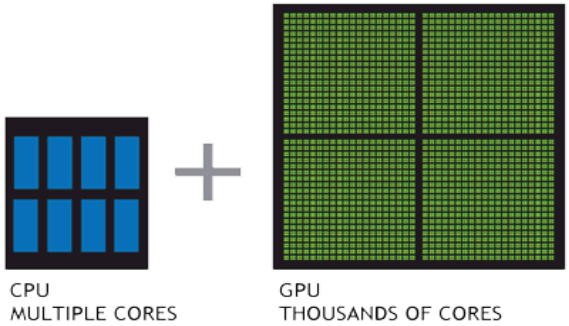
SAND2017-8517C





Optimization of Input Parameters to Improve Big Data Computed Tomography Reconstruction Performance

Celia J. Flicker[1,2], Srivathsan P. Koundinyan[1,3], Dr. Edward S. Jimenez[1]

1. Sandia National Laboratories

2. New Mexico Institute of Mining and Technology

3. Stanford University

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Computed Tomography (CT)

Sandia National Laboratories

- Enables volumetric cross-sectional imaging of an object
  - *Input:* projection images from multiple angles
  - *Reconstruction Technique:* Three-dimensional FDK (Feldkamp)
  - *Output:* Object slices and planes
  - *Applications:*
    - Medical imaging
    - Failure inspection
    - Material identification

Computed Tomography (CT)

- Enables volumetric cross-sectional imaging of an object
  - *Input:* projection images from multiple angles
  - *Reconstruction Technique:* Three-dimensional FDK (Feldkamp)
  - *Output:* Object slices and planes
  - *Applications:*
    - Medical imaging
    - **Failure inspection**
    - **Material identification**

Industrial CT

Sandia National Laboratories

Datasets of sizes larger than 100GB

Standard CPU reconstruction techniques [$O(n^4)$ or $O(n^3 log(n))$]

Future datasets of sizes larger than 1TB

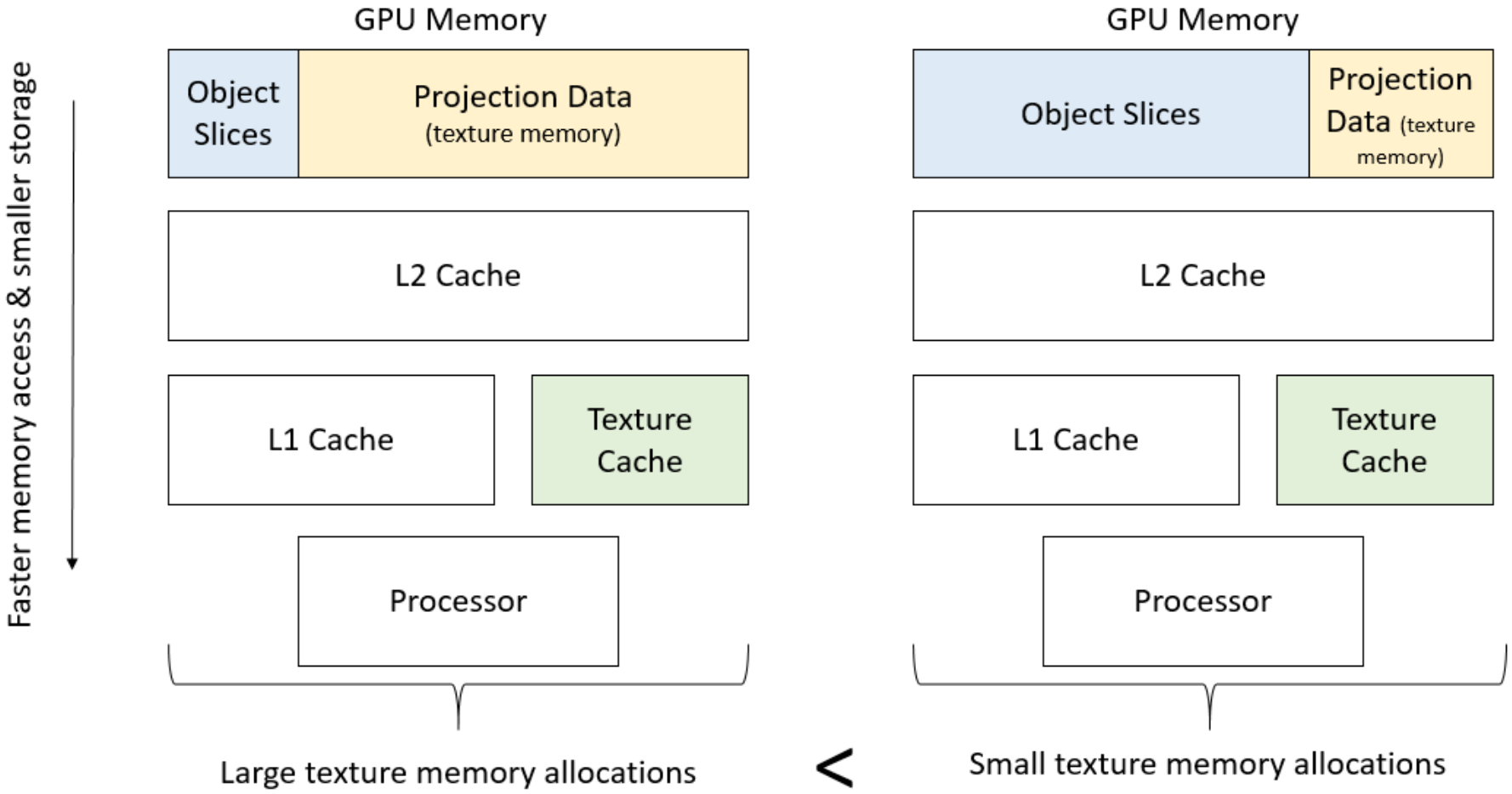
Reconstruction time on the order of weeks!

Graphical Processing Units (GPUs)

Sandia National Laboratories

- Originally built for gaming and CAD industry
- Key features:
  - Massively parallel architectures
  - Fast read-only memory
  - Compatible with multi-threaded CPU computing
- Reconstruction times may still span few days for industrial CT due to:
  - GPU memory limitations
  - GPU memory management complications
- Prior work utilizing GPUs for reconstruction optimized for small-scale datasets (~1 GB)

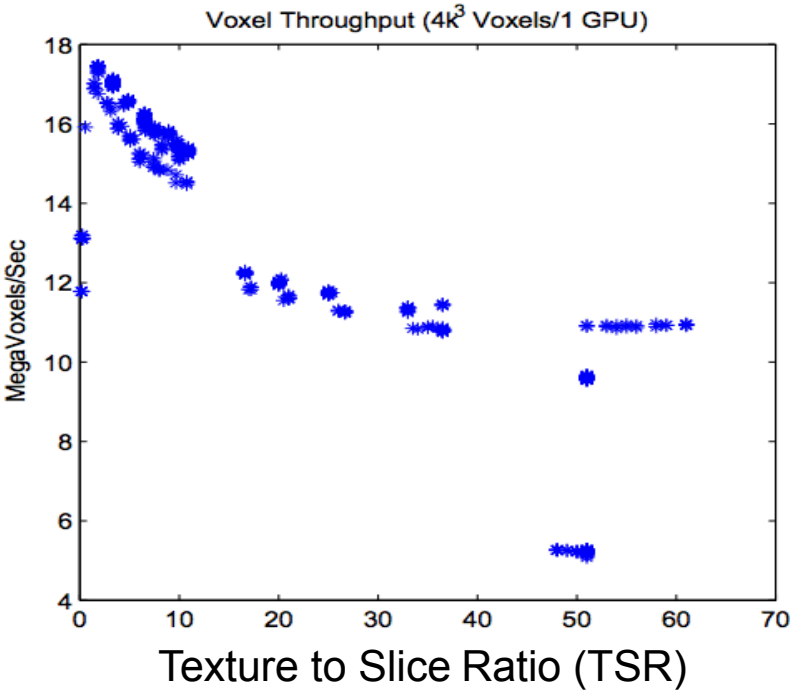
Irregular CT Reconstruction

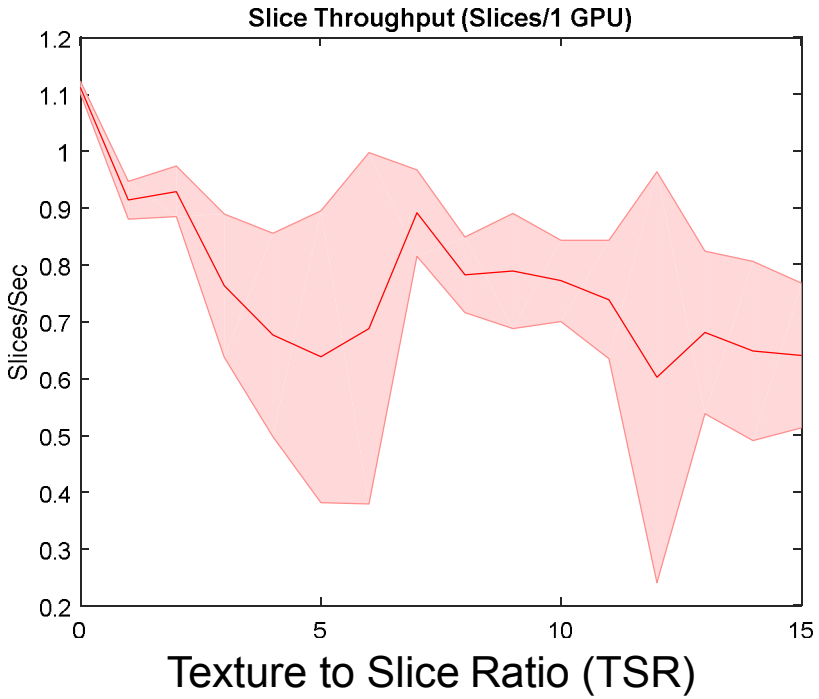


Max slice depth largest number of slices stored in GPU memory	**Texture-to-Slice Ratio (TSR)** (Memory allocated to projection data) / (Memory allocated to object slices)	**Memory fraction** fraction of total GPU memory available for computation

Irregular CT Reconstruction

- Performance analysis with constant max slice depth and memory fraction, and varying TSR:

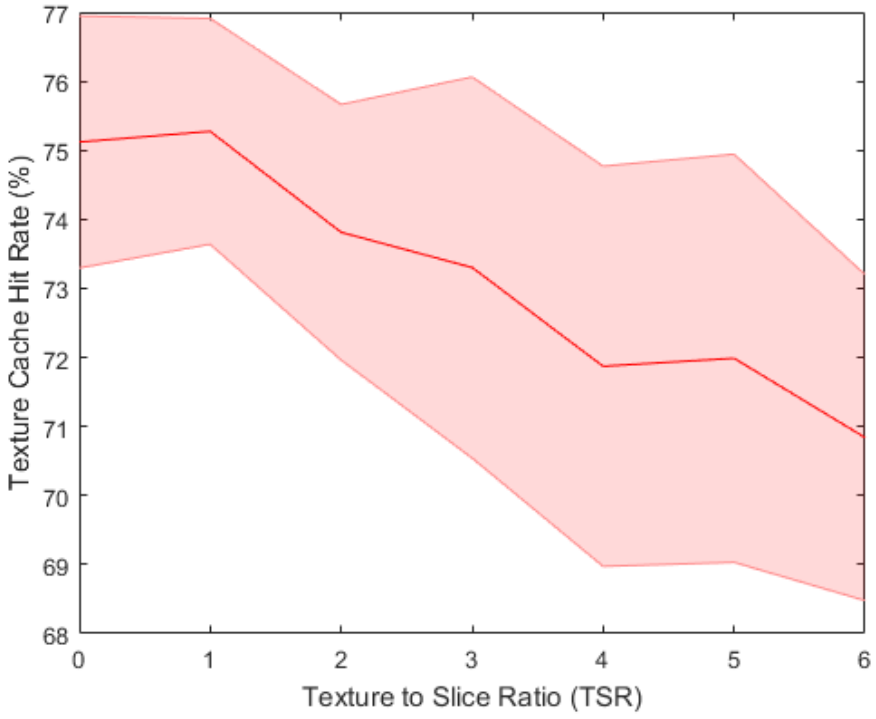


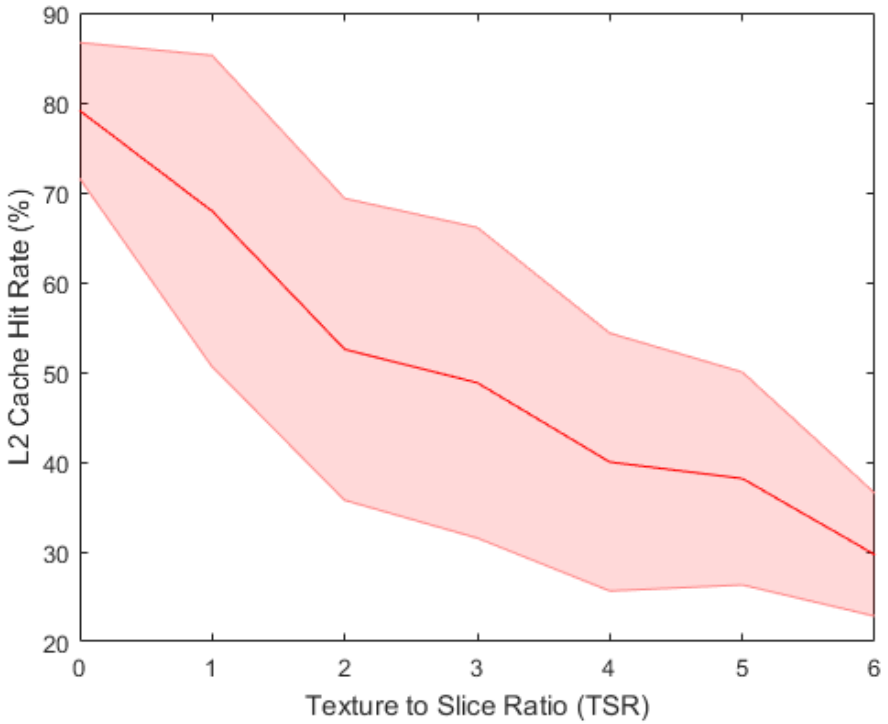


Smaller TSR = improved computation time

Irregular CT Reconstruction

- Performance analysis with constant max slice depth and memory fraction, and varying TSR:



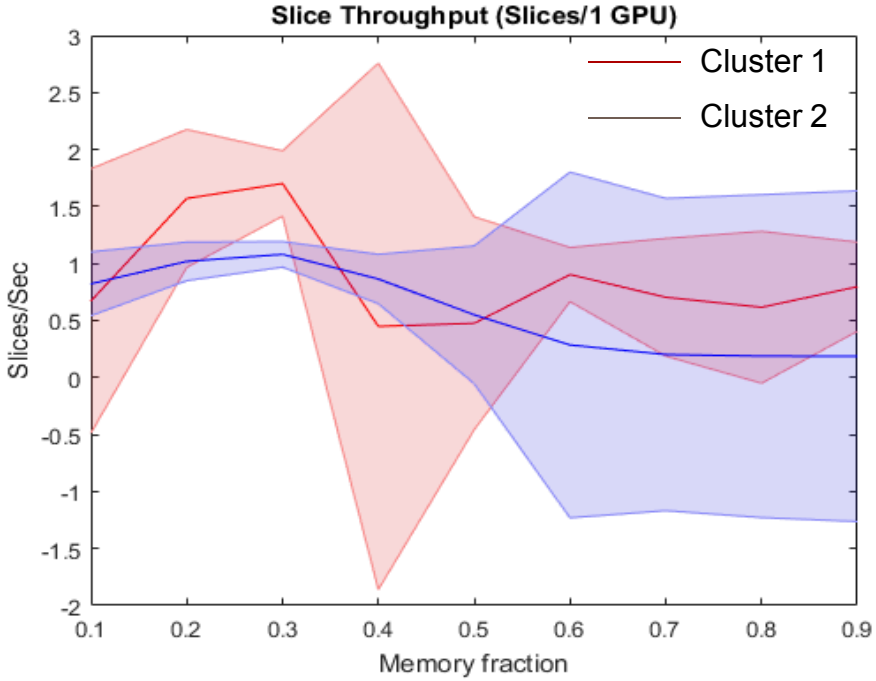


Smaller TSR = improved cache hit rates

Performance on Different Architectures

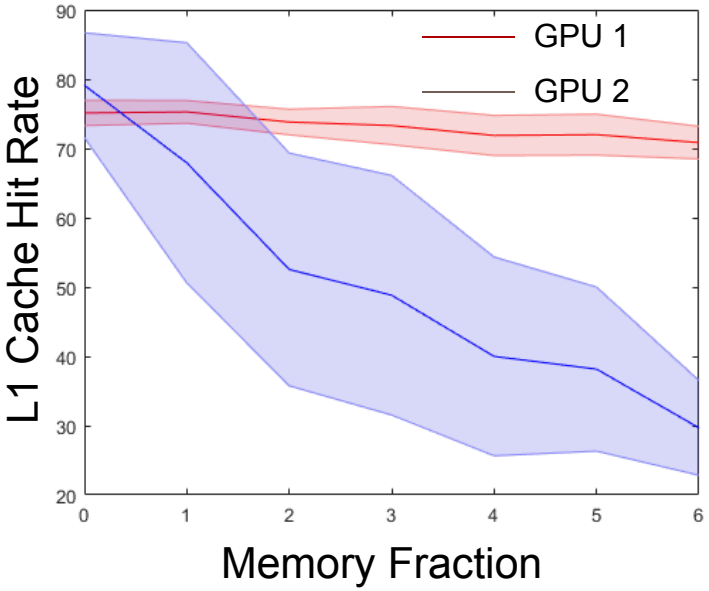
- Performance analysis with constant TSR and max slice depth, and varying memory fraction across two different computing architectures:
  - **Cluster 1:** Tesla K80 with 12 GB of memory
  - **Cluster 2:** Tesla P100-PCIE with 16 GB of memory

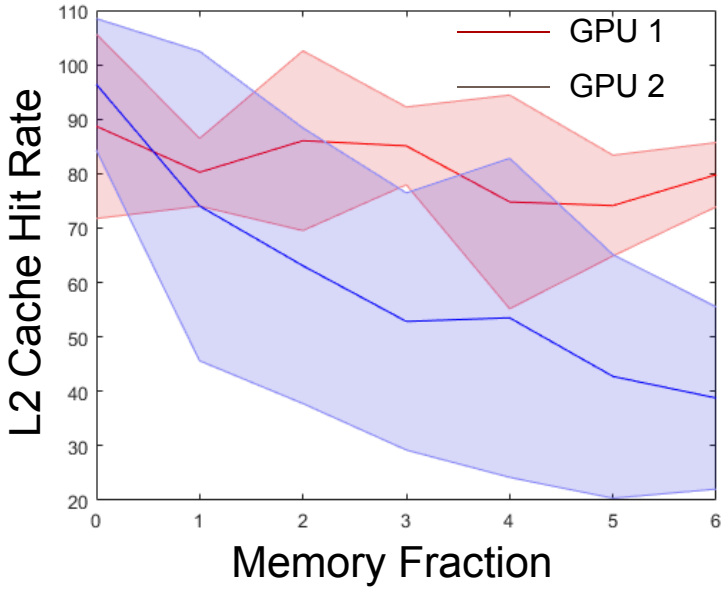
Cluster 1 outperforms cluster 2 for certain memory fractions



Performance on Different Architectures

- Performance analysis with constant TSR and max slice depth, and varying memory fraction across two different computing architectures:
  - **GPU 1:** GTX 1080 with 8 GB of memory
  - **GPU 2:** Kepler with 2 GB of memory





GPU 2 has better performance than GPU 1 for certain memory fractions

Optimization of Input Parameters

Sandia National Laboratories

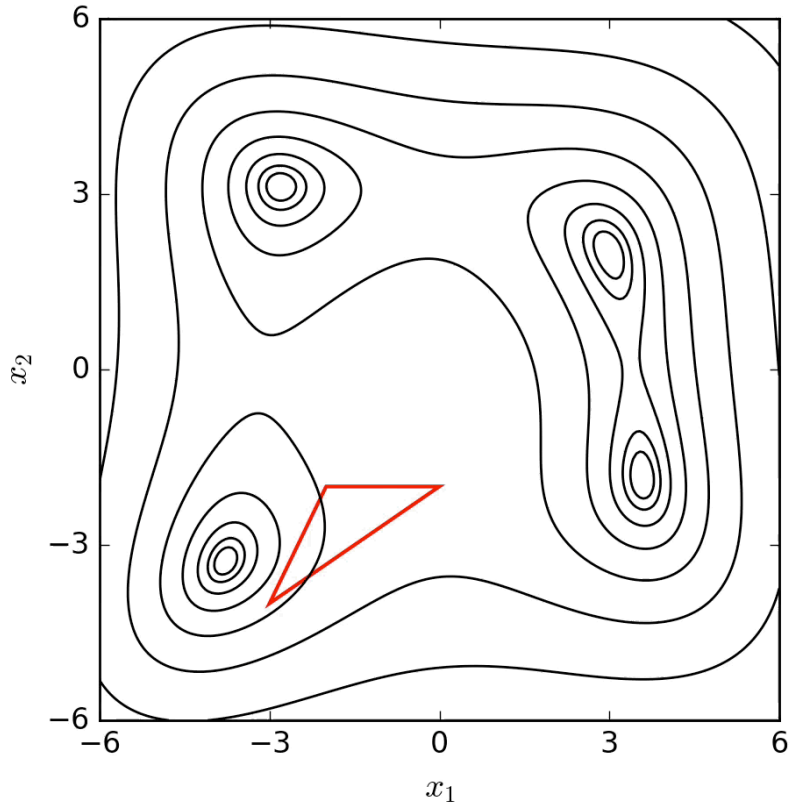
Texture-to-Slice Ratio (TSR)

Max slice depth

Memory fraction

Automated Optimization of Input Parameters

- The best input parameters for a given dataset size and computing architecture can be identified using the Nelder-Mead optimization technique



Conclusion

- Demonstrated
  - Intuitive parameters suboptimal
  - GPU performance
  - Effects of parameters
- Future work
  - Data size and GPU architecture
  - Parameter optimization