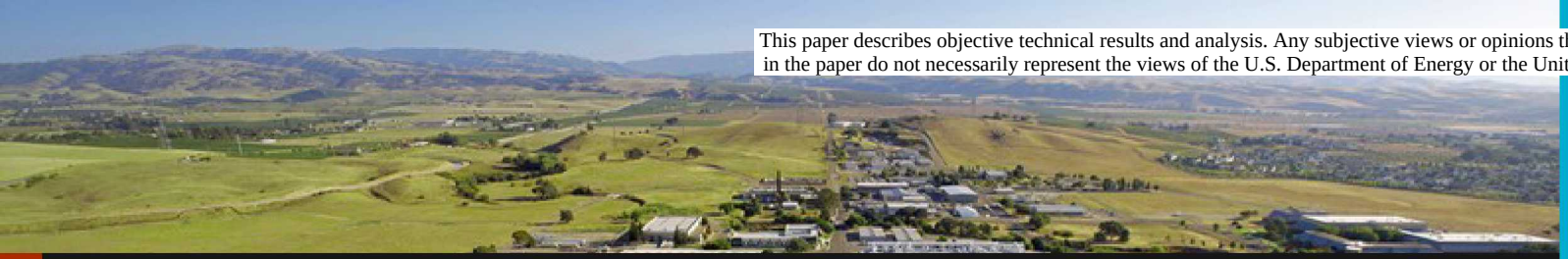
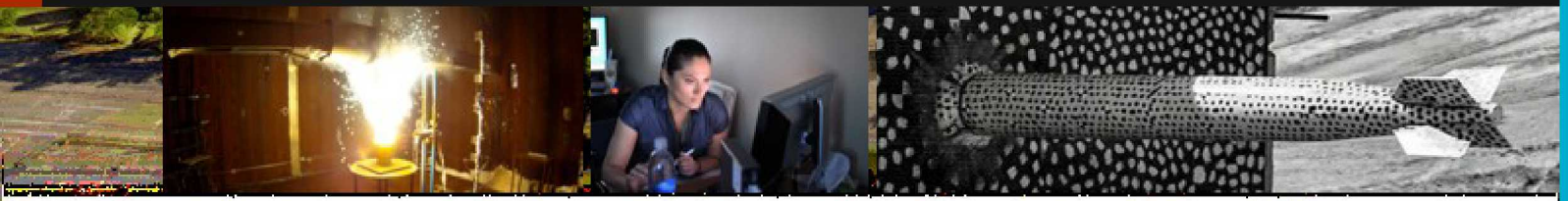


This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

SAND2019-2186C

Computing Infrastructure



PRESENTED BY

Andre Encarnacao

The views expressed here do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

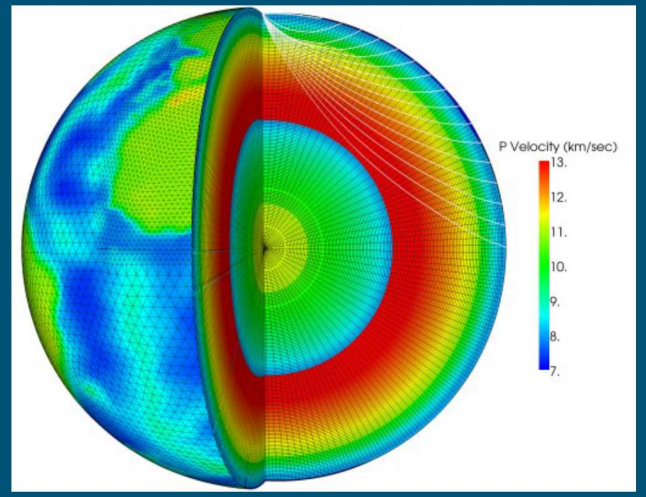
Motivation

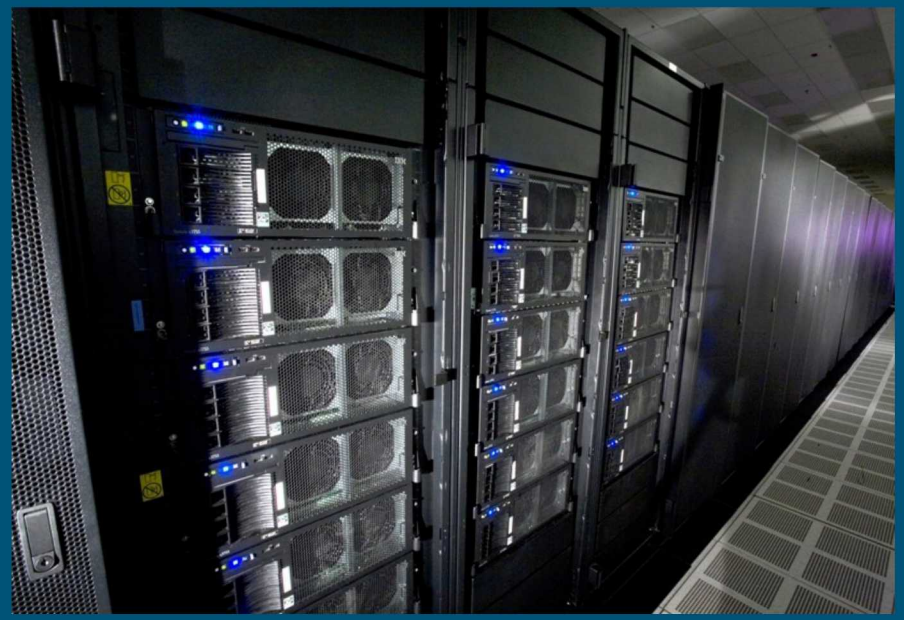
To develop a 3D earth model, the tomographic inversion process must compute travel time predictions > 300 million times (30 iterations * 10 mil observations)

- 1D: Travel times require a simple table lookup (~0.01msec each, 300M = 50 minutes)
- 3D: Travel times require full ray tracing (~0.1sec each, 300M = 347 days)

Parallel Computing

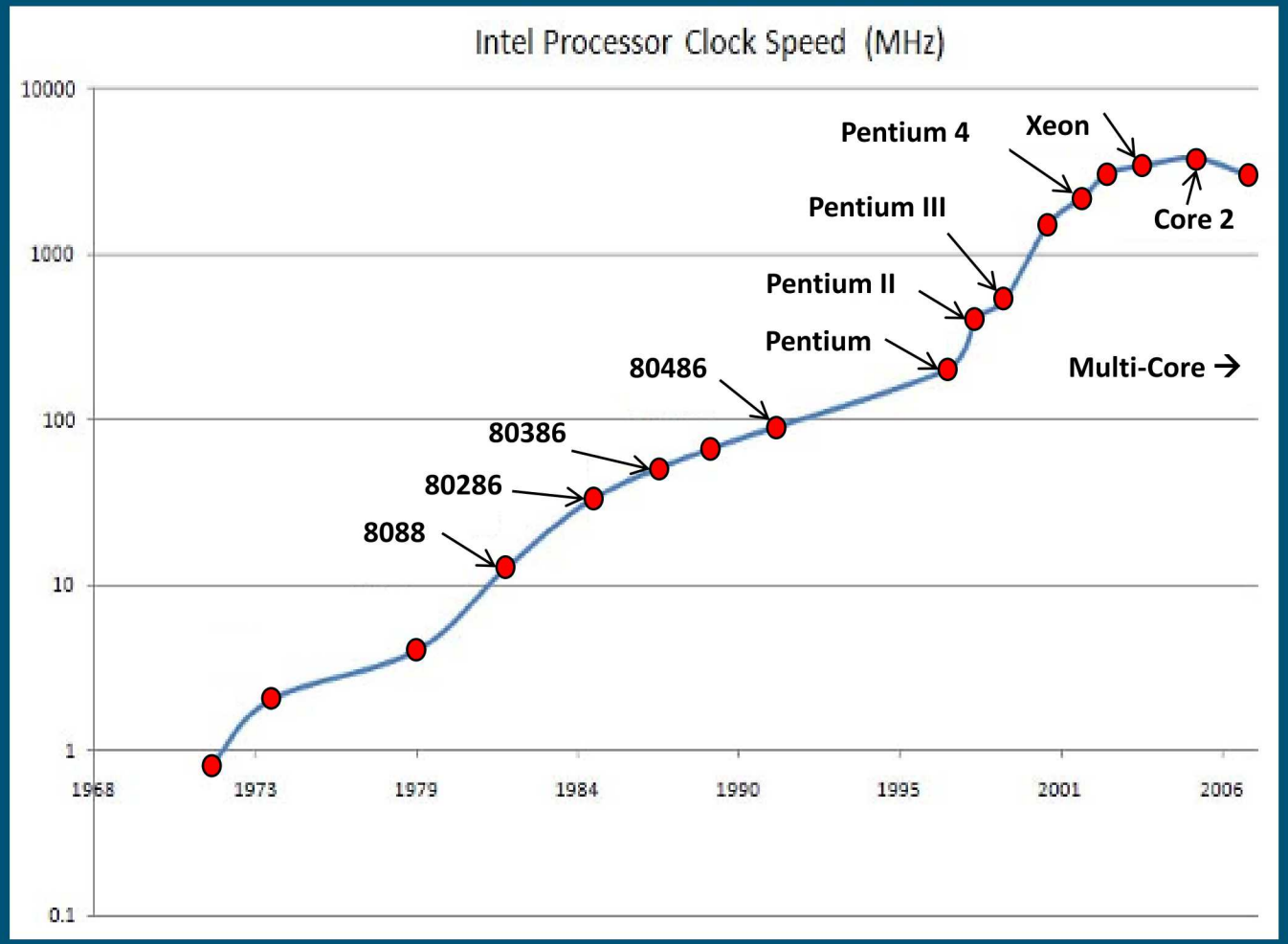
- Supercomputer or commodity hardware?
- This is research - many models and versions of code
- Need a readily accessible solution





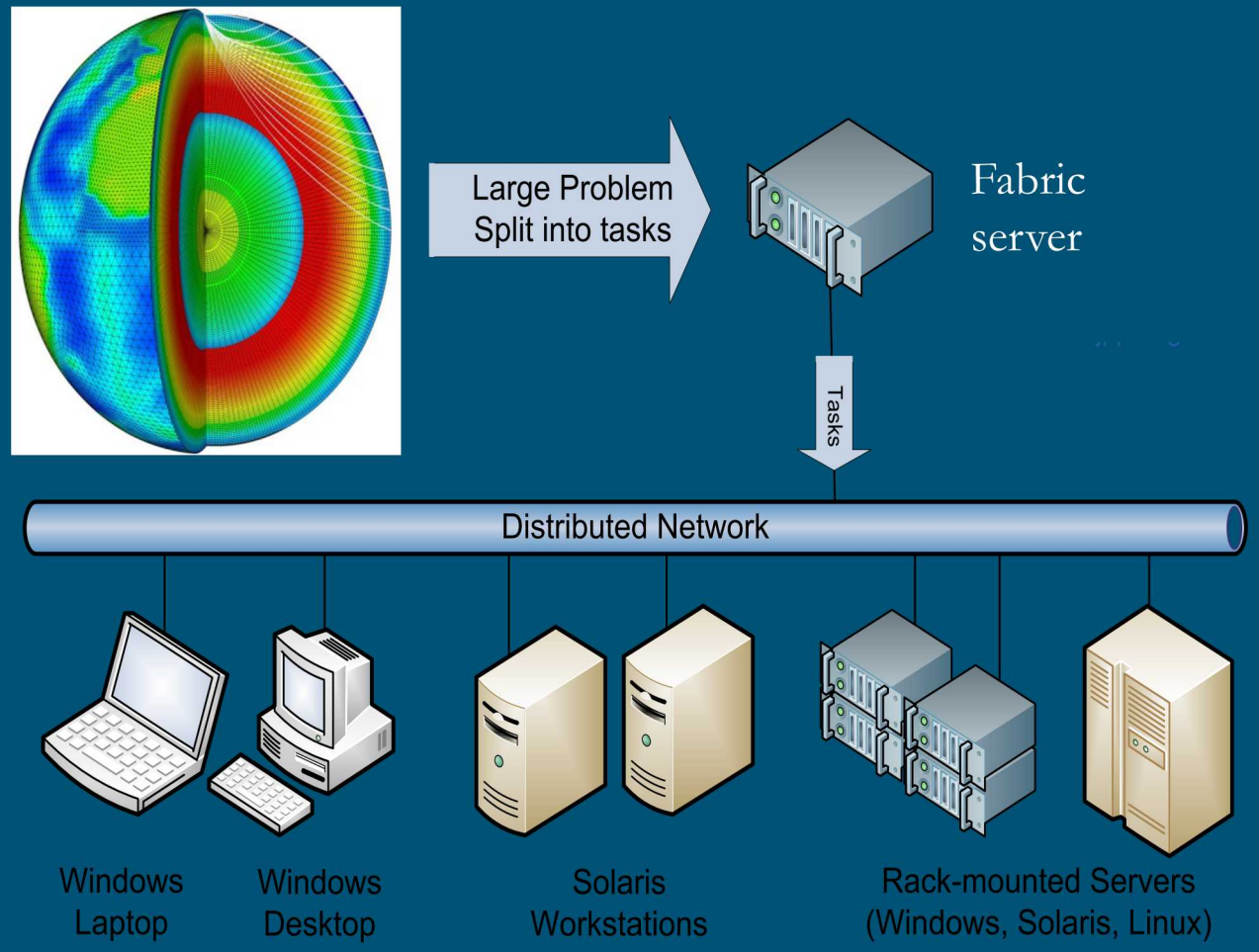
Processor Speed Trends

- Moore's Law is changing
- Decrease in processor clock speeds – Sequential code may actually run slower
- Increase in number of cores / threads – Must program in parallel



Our Approach - Fabric

Platform independent, highly scalable, and shareable

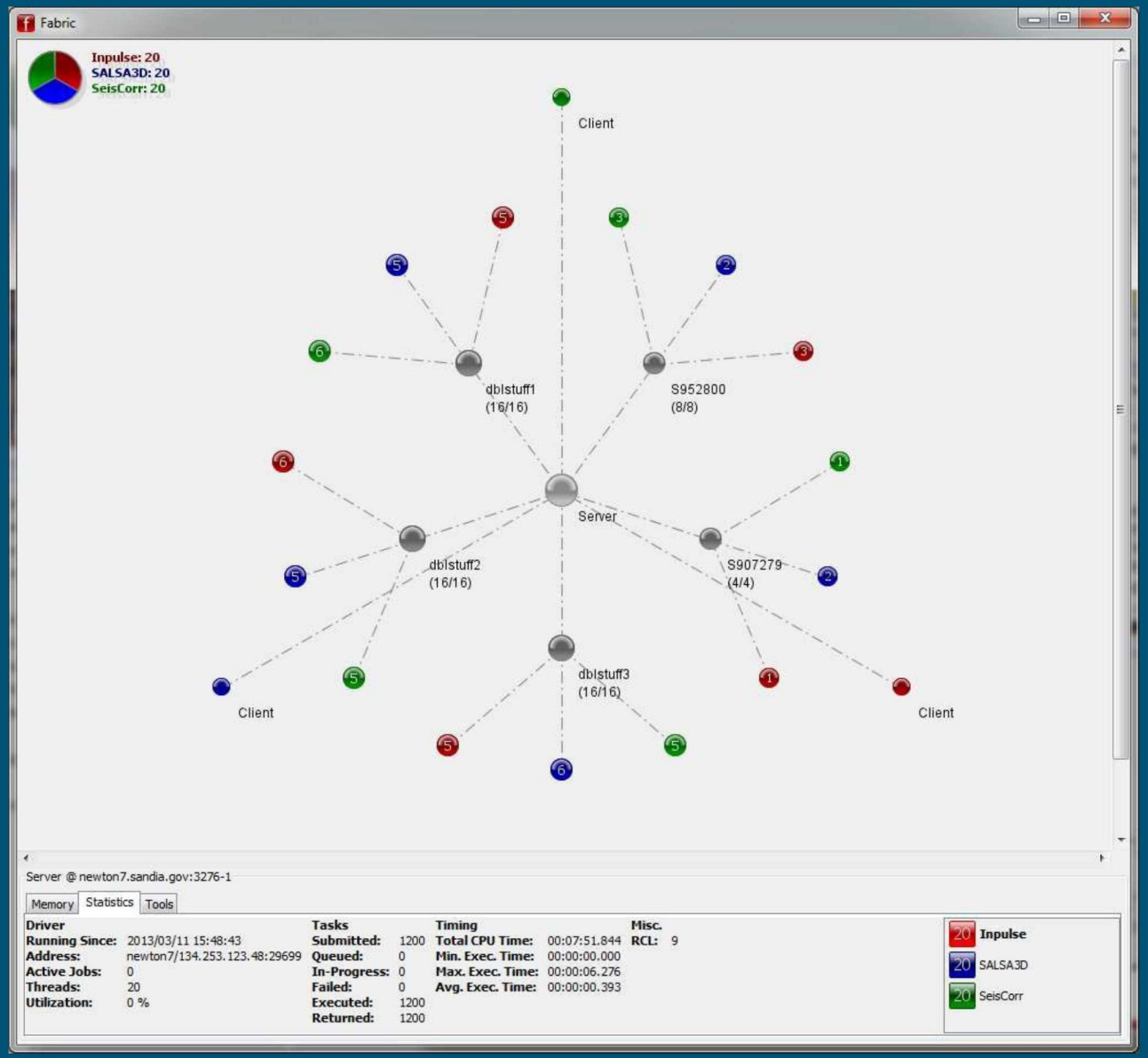


Fabric hosts

Our Linux-based cluster is composed of:

- 8 * 40-core/80-threads, 384GB RAM
- 6 * 24-core/48-threads, 384GB RAM
- 3 * 48-core/96-threads, 768GB RAM each
- 1 shared file server
- 1 shared Oracle DB server
- In total, 1216 processing threads at a time.

Fabric



- Task Parallelism
  - An application is sub-divided into independent tasks
  - Tasks are distributed and processed across different nodes
- Resource sharing
  - Applications can run simultaneously and share compute nodes automatically

Applications and Performance

- Ray Tracing
  - Tomography & event relocation
  - Tomography runs now take less than 24 hours (versus nearly 1 year if run sequentially)

- Covariance matrix inversion
  - Out-of-core parallel
  - Matrix size = N^2 x N^2, where N = # of data nodes in model
    - 250K x 250K matrix = ~500GB
    - Store half (250GB) since matrix is symmetric
    - Divide into ~7500 x 33MB blocks
  - Covariance matrix inversion takes ~12 hours