Performance Studies of Scientific Applications on High Performance Computing Cluster with Big-Memory Footprint

By Derrick Lee, Garrett Lord,
Jesse Martinez, and Christopher Moore

Mentors: HB Chen, Parks Fields, Benjamin McClelland, Alfred Torrez

Computer Systems, Cluster, and Networking Summer Institute 2011

New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

Outline

- Big Memory Footprint
  - Introduction
  - Applications
  - Devices Used
  - Types of Benchmarks
  - Results from Benchmarks
  - Future Development
  - Conclusion

New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

Introduction

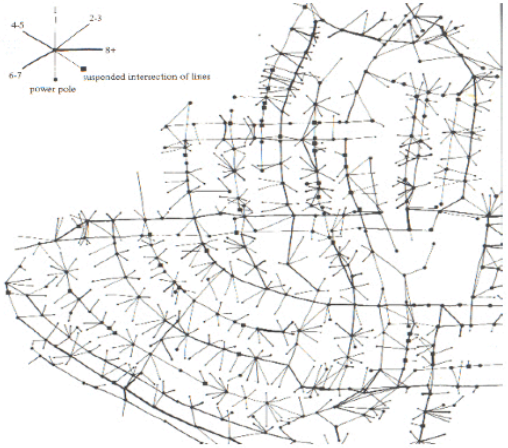
- We have conducted a sequence of benchmark tests that are related to the LANL Data Intensive Super Computing (DISC) project
- Virtualized Aggregation System
- The vSMP system will be well-suited to large memory problems.

New Mexico CONSORTIUM

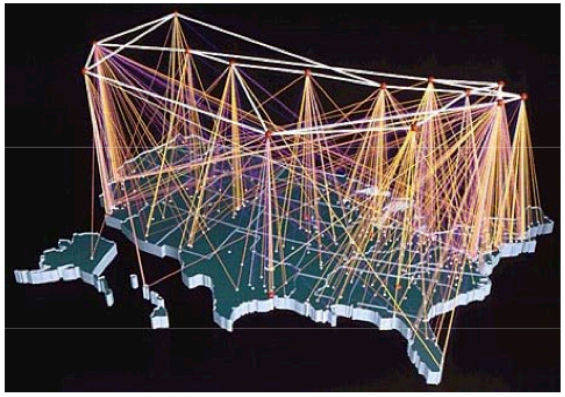
Los Alamos NATIONAL LABORATORY EST. 1943

Graph problems arise from a variety of sources

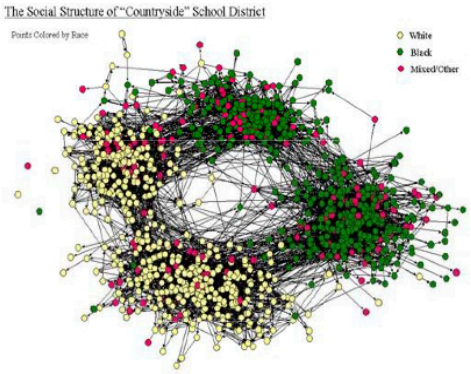
Power Distribution Networks



Internet backbone



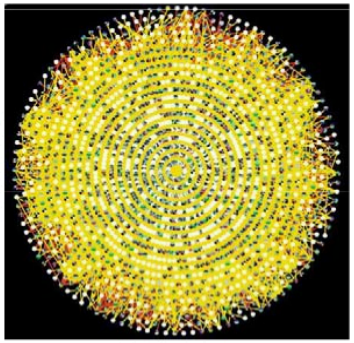
Social Networks



Graphs are everywhere!

Ground Transportation

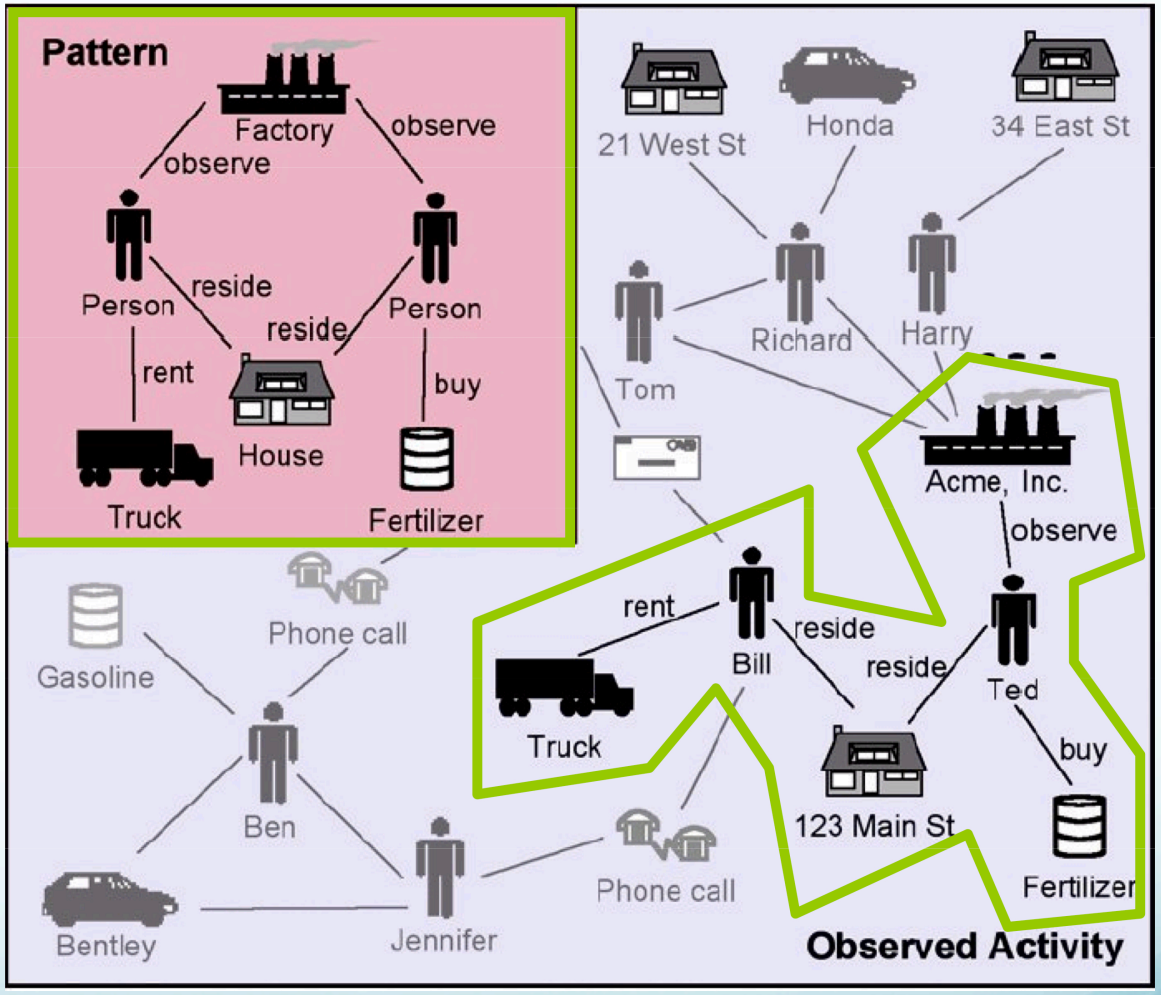
Tree of Life



Protein-interaction networks

Application – Homeland Security: Terrorist Networks

- Certain activities are often suspicious not because of the characteristics of a single actor, but because of the interactions among a group of actors
- Interactions are modeled through a graph abstraction where there entities are represented by vertices, and their interactions are the directed edges in the graph



New Mexico CONSORTIUM

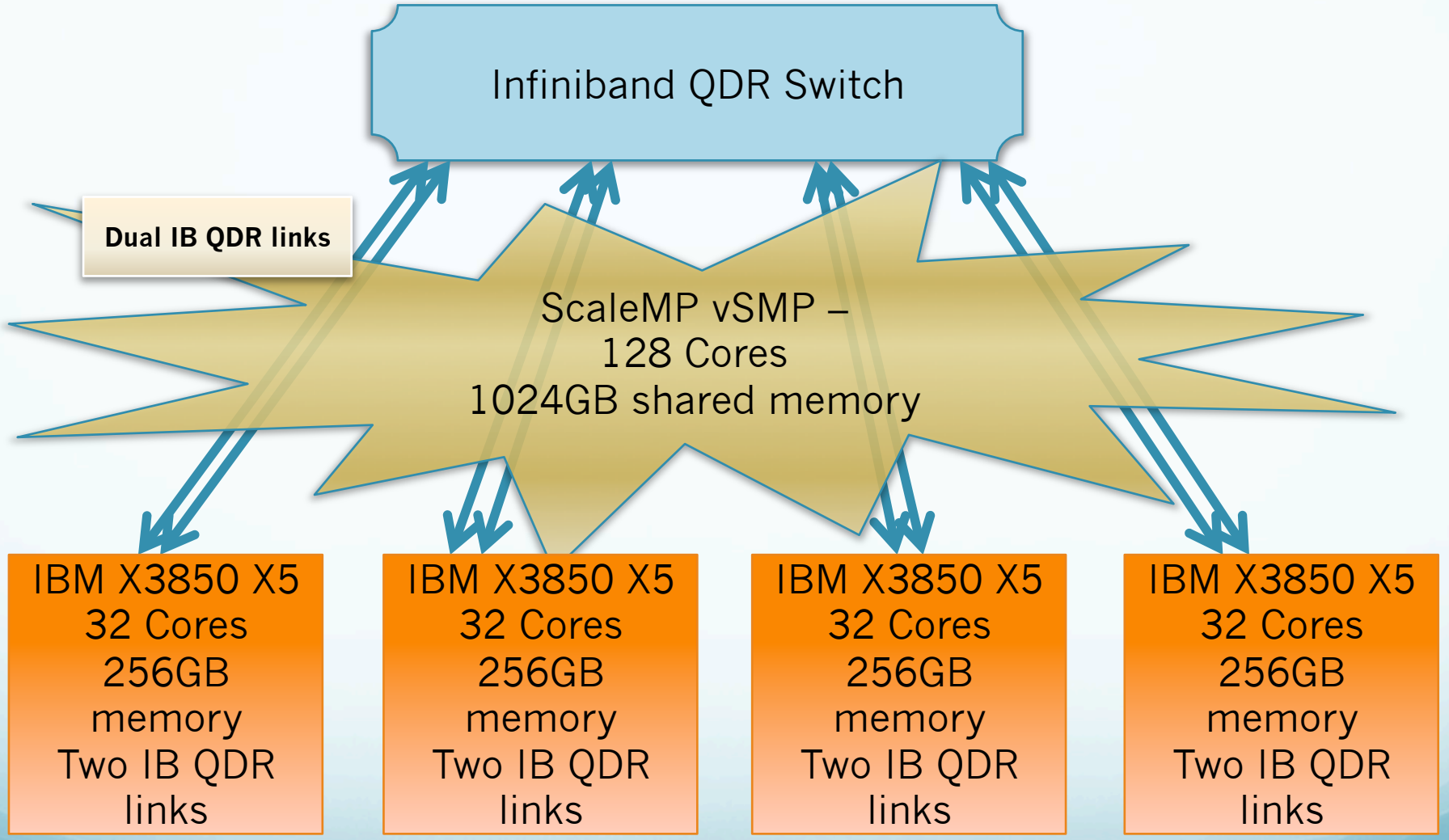
Georgia Tech College of Computing

Los Alamos NATIONAL LABORATORY EST. 1943

Devices

- 4x IBM x3850 X5 servers
  - RAM: 256 GB per Node
  - Total: 4x 256 GB = 1TB
  - 128 Cores
  - ScaleMP Virtual Aggregation
  - vSMP Foundation (i.e. Versatile SMP)
- SGI UV 100/1000 system – 128 cores, 512GB memory, 2.67GHz/24MB E7- 8837 processors, remote testing account provided by SGI

Test Bed Setup

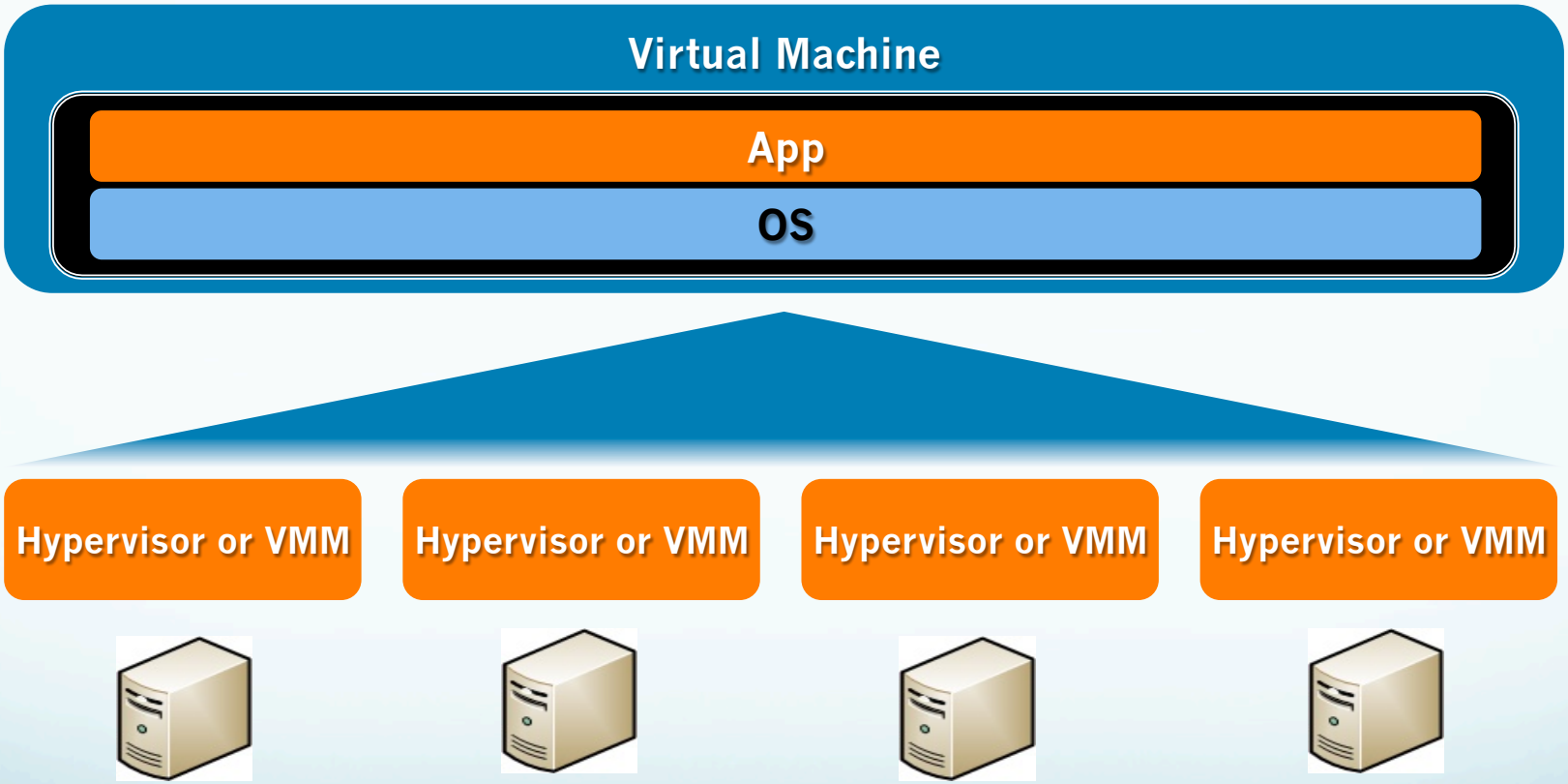


New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

AGGREGATION

Concatenation of physical resources

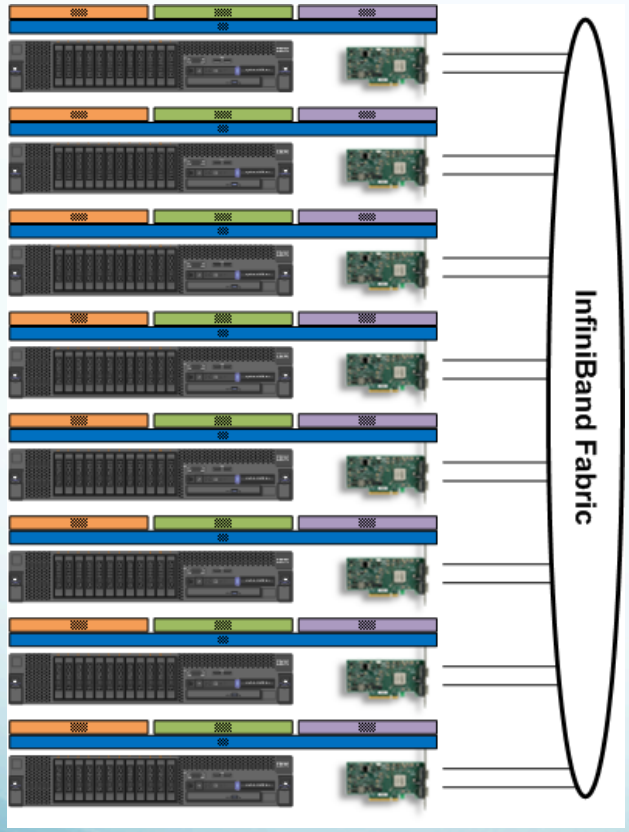


New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

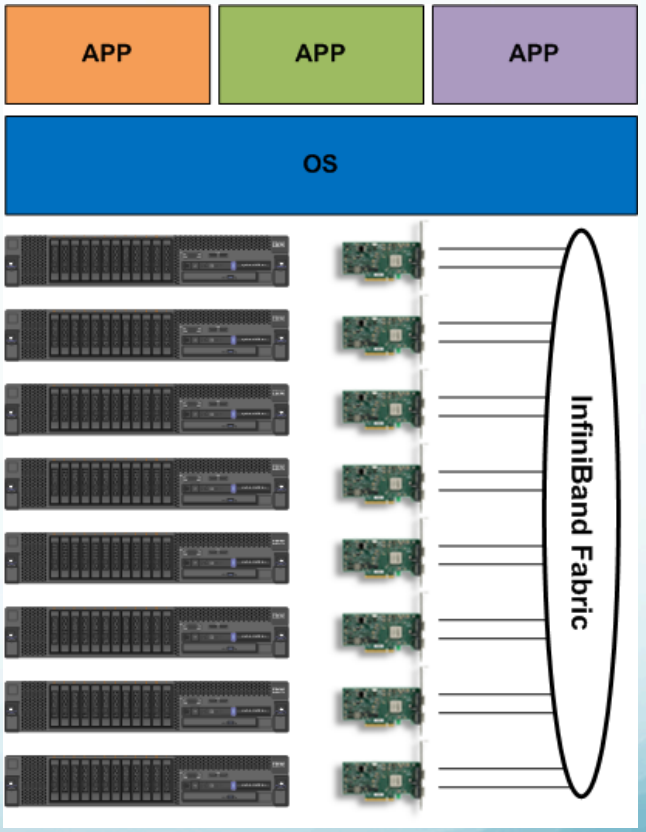
How It Works

Multiple Computers with Multiple Operating Systems



ScaleMP™

Multiple Computers with a Single Operating System



ScaleMP™

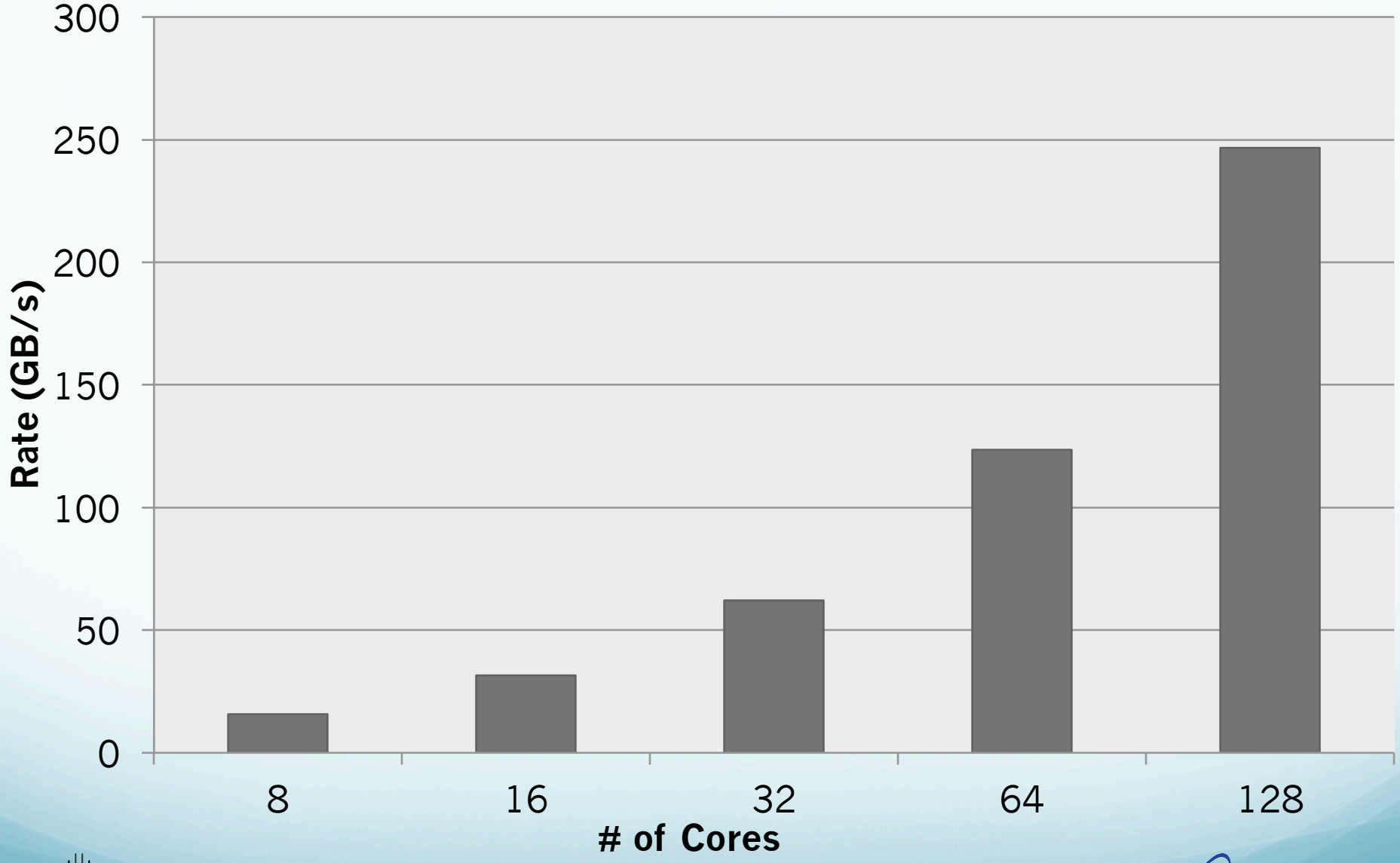
New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

Benchmarks

- STREAM Bandwidth
  - Copy, Scale, Add, Triad
- Graph500
  - Traversed Edges per Second (TEPS)
  - Scales 24-30
  - Cores: 64 and 128

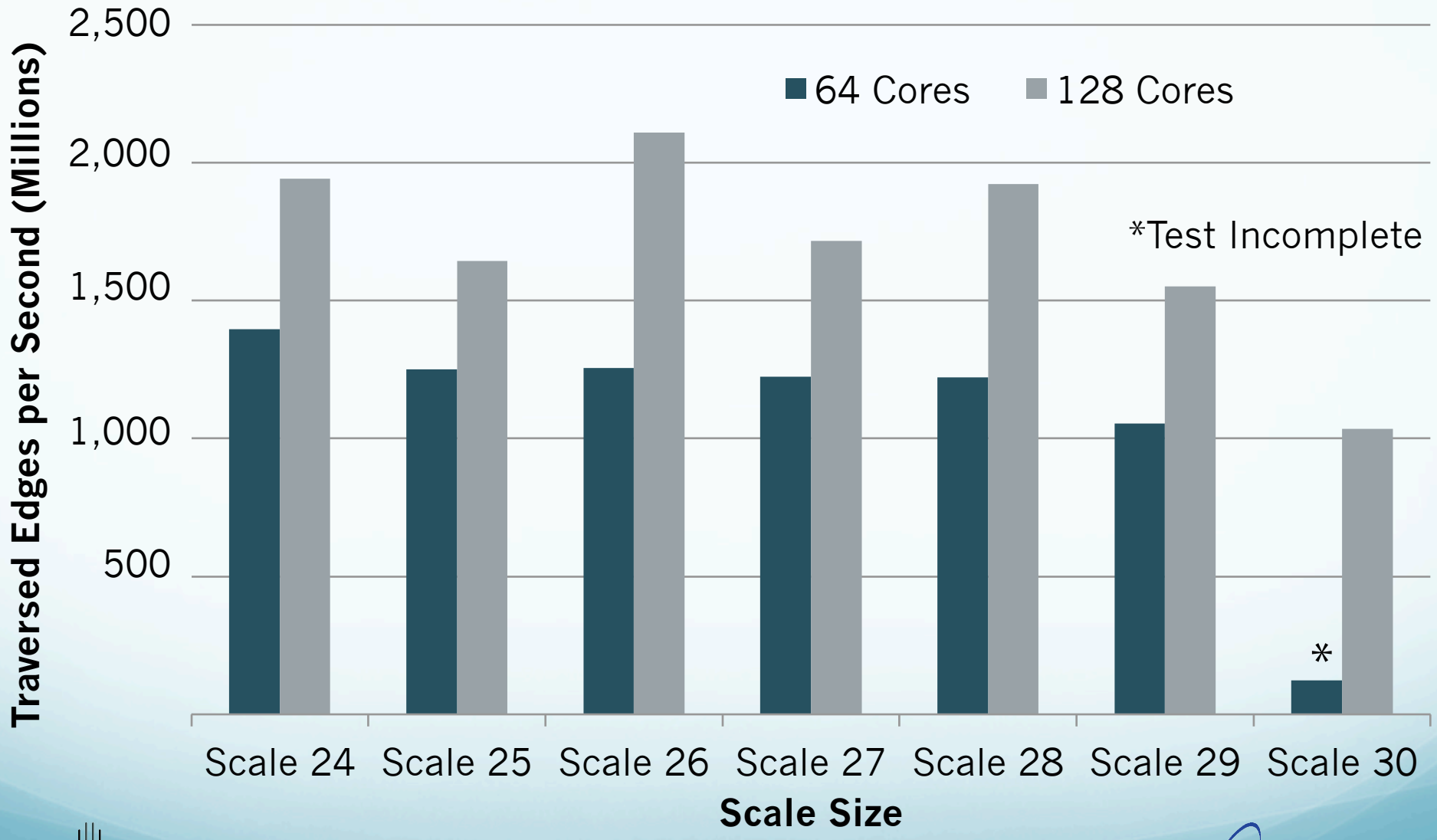
STREAM Results - Triad



New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

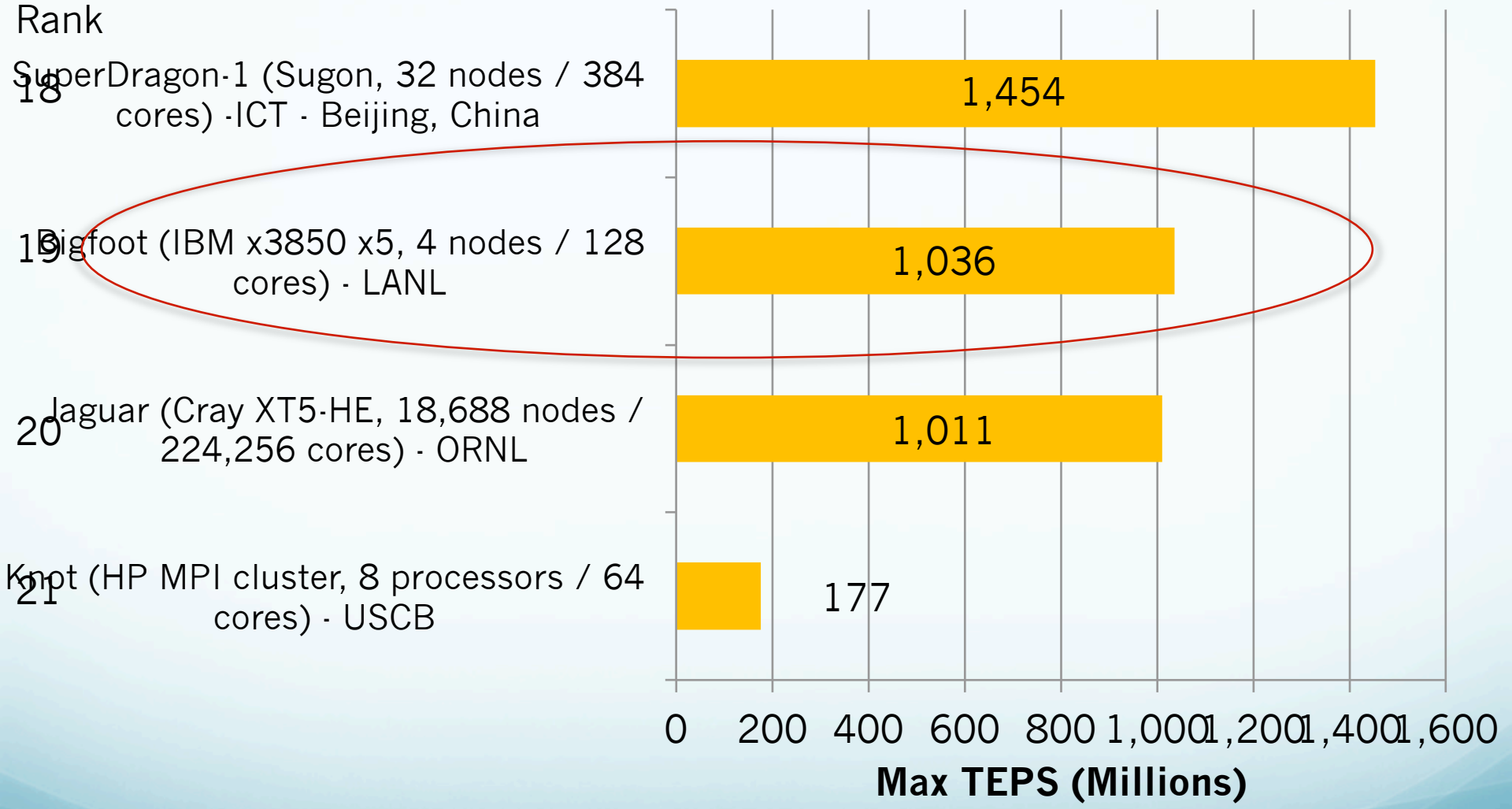
Graph 500 Results
Max. TEPS



New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

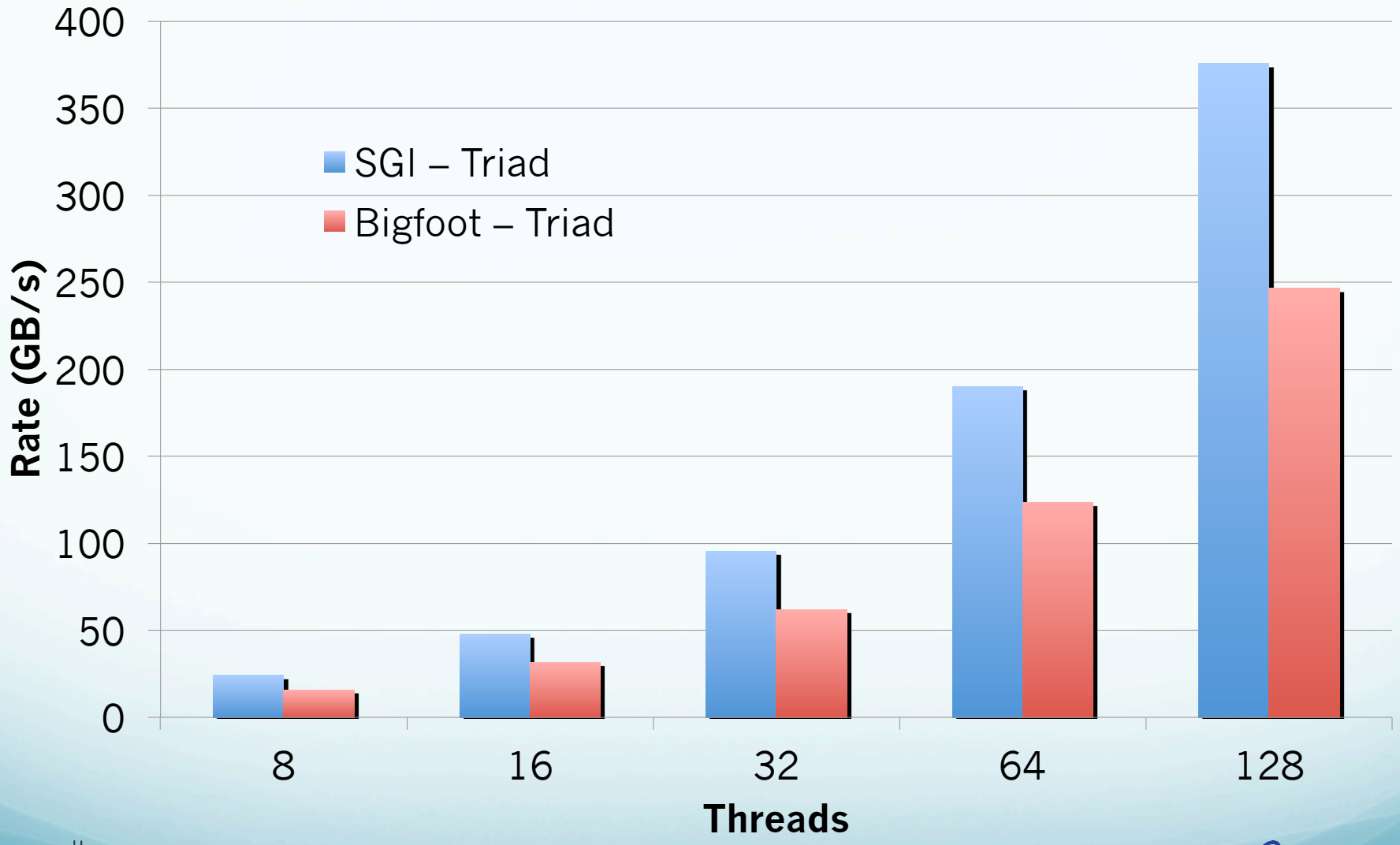
Graph 500 Results Compared to Current Placeholders



New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

SGI UV 100 HW vs Big Memory Footprint SW - STREAM - MPI



New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

Areas for Future Improvement

- Increasing RAM in each Node
- Adding SSDs for Swap File Systems
- Large Scale Graph Generation
- Large Data Analysis for Researchers
- Graph500 testing on SGI UV system
- Comparison of results from test beds
- Summit our Graph500 testing result from the scaleMP/IBM machine to Graph500 for SC11

Conclusion

- Large memory footprints provide performance increase for large data sets
- Graph 500 results showed potential increase for higher results.
- IBM aggregation system is more affordable compared to other machines for its current output

Questions?

New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943