SAND2008-2512P

Informatics Architectures

Jonathan Berry
Scalable Algorithms Department

April 14, 2008

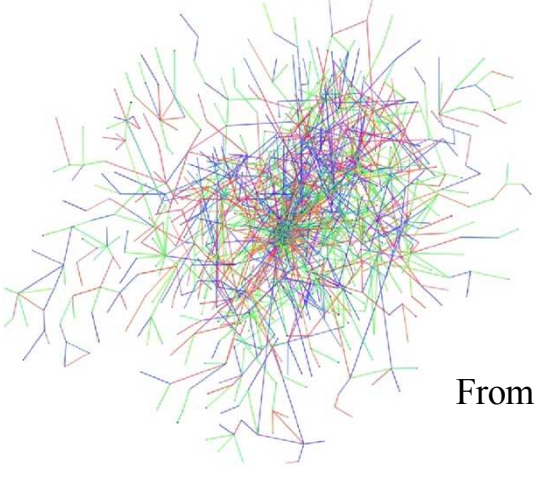
Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Informatics Datasets Are Different

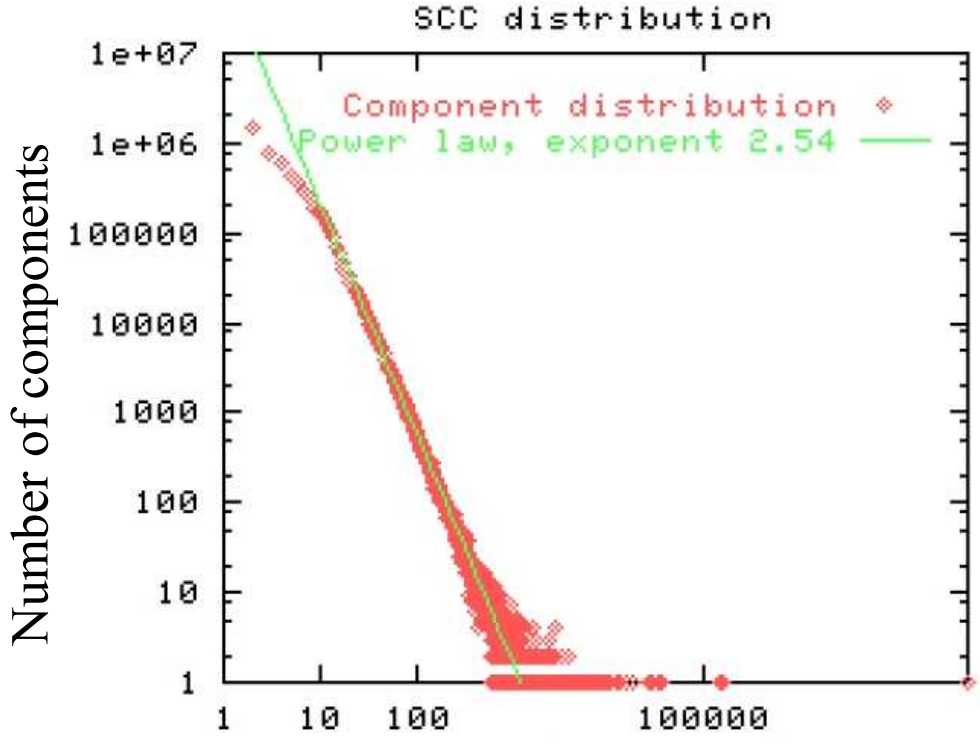
Informatics: The analysis of datasets arising from "information" sources such as the WWW (not physical simulation)

Motivating Applications:

- Homeland security
- Computer security (DOE emphasis)
- Biological networks, etc.



From UCSD '08



Broder, et al. '00

"*One of the interesting ramifications of the fact that the PageRank calculation converges rapidly is that the web is an expander-like graph*"

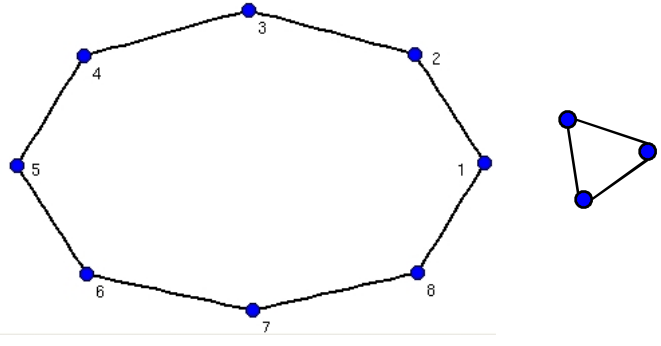
Page, Brin, Motwani,Winograd 1999

Primary HPC Implication: Any partitioning is "bad"

Sandia National Laboratories

Informatics Algorithms Are Different As Well

Connected Components: *find groupings of vertices such that all vertices within a group can reach each other*



> "*The single largest performance bottleneck in the distributed connected components algorithm is the effect of poor vertex distribution…Several methods…have been implemented but none has been successful as of yet.*"
>
> D. Gregor, from Parallel Boost Graph Library documentation on connected components

S-T Connectivity: *find a short path from vertex S to vertex T*

Single-Source Shortest Saths (SSSP): *from a given vertex, find the shortest paths to all other vertices*

> *"[in power law graphs] there is a giant component…of size O(n)"*
> Aiello, Chung, Lu, 2000

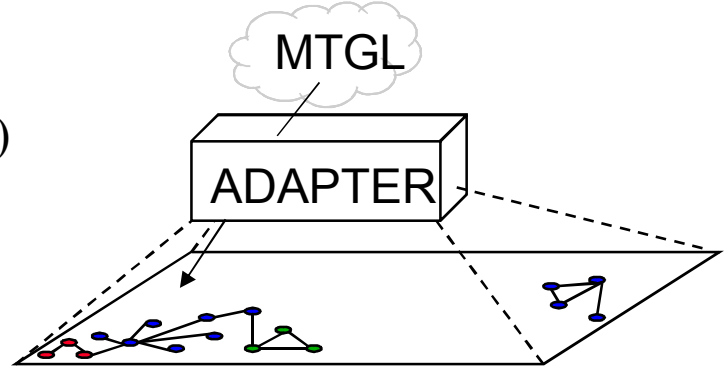
Informatics Problems Demand New Architectures

Distributed Memory Architectures	Massively Multithreaded Architectures	Key Issues
Fast CPU (~3GHz)	Slow CPU (~200-500MHz)	Power, concurrency
Elaborate memory hierarchy	Almost no memory hierarchy	Is cache justified?
Memory per-processor, partitioned	Global address space	Can you partition?
Operating system for threading, synchronization	Hardware for threading, synchronization	How fine-grained is your data interaction?
Programming paradigm is standardized (MPI)	Programming paradigm is machine-specific (mta-pe)	Portability, debuggability

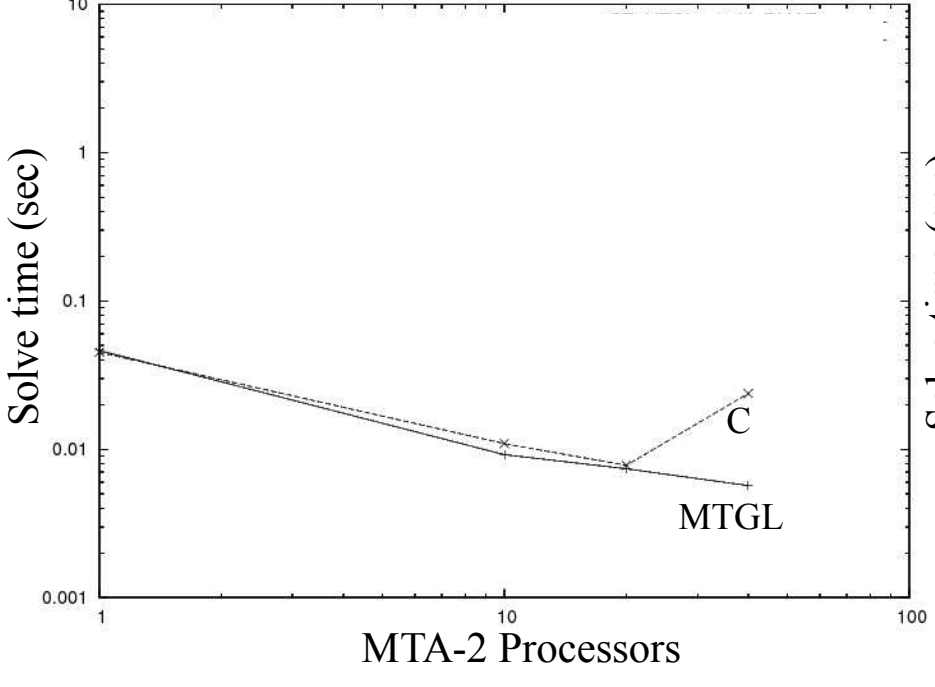
Multithreaded architectures show promise for informatics problems, but more work is necessary…

We Are Developing The MultiThreaded Graph Library

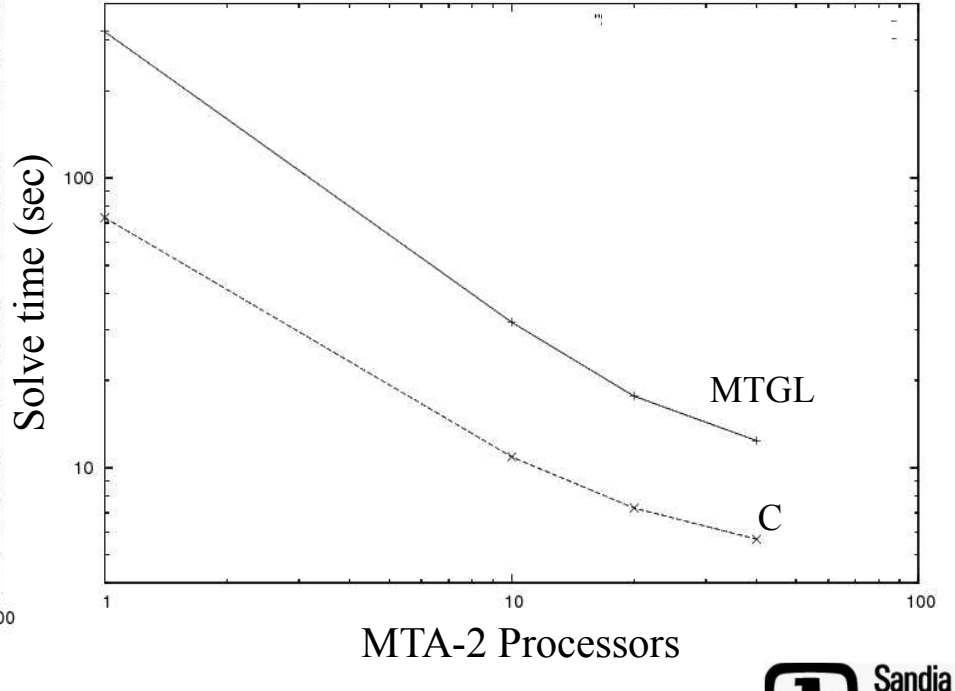
- Enables multithreaded graph algorithms
- Builds upon community standard (Boost Graph Library)
- Abstracts data structures and other application specifics
- Hide some shared memory issues
- Preserves good multithreaded performance



S-T connectivity scaling (MTA-2)



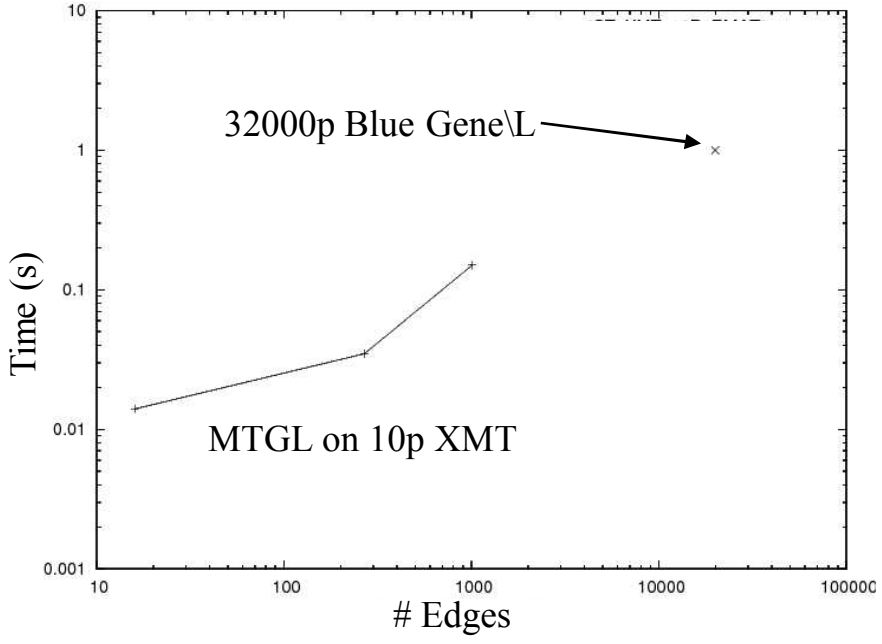
SSSP scaling (MTA-2)



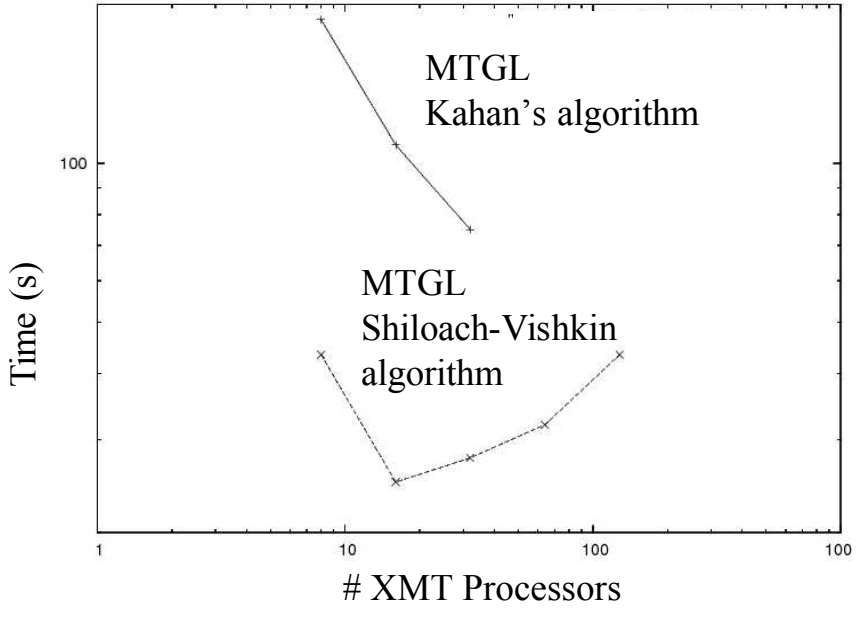
Sandia National Laboratories

Initial Algorithmic Impacts of MTGL on XMT Are Promising

- **S-T Connectivity**
  - Gathering evidence for 2005† prediction
  - 128P XMT can handle up to 10 billion edges
  - This plot show results for ≤ 1 billion



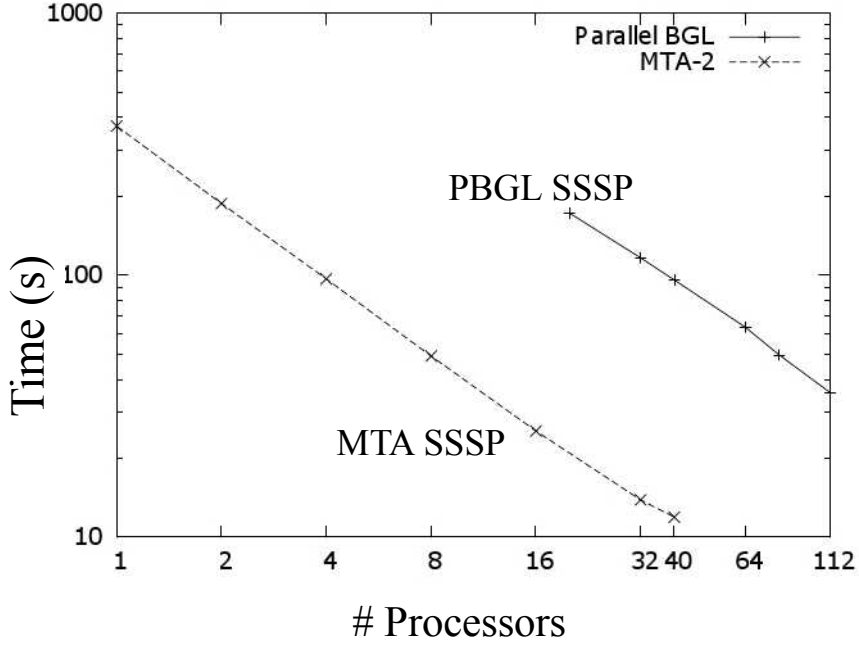
- **Connected Components**
  - Simple SV is fast, but hot-spots
  - Multilevel Kahan algorithm scales (but XMT data incomplete)
  - No current competitors for large power-law graphs



MGTL on XMT sets performance standard for informatics problems

A Recent Comparison With PBGL Finds Efficiency Gap

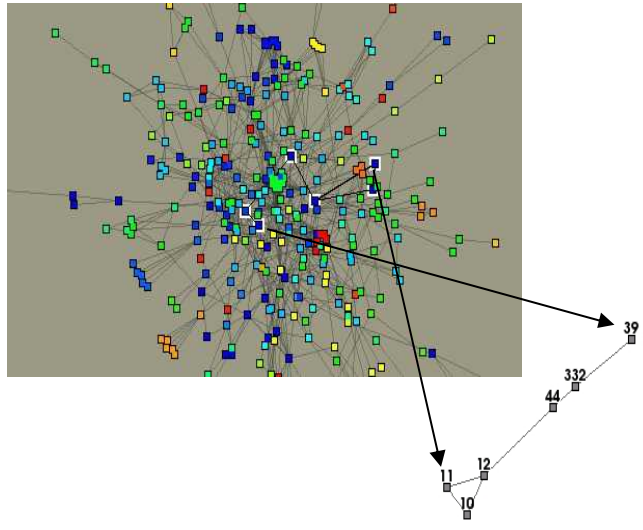
- **Parallel Boost Graph Library (PBGL)**
  - Run Boost GL on clusters
  - Some graph algorithms can scale on some inputs
- **PBGL - MTA Comparison on SSSP**
  - PBGL SSSP can scale on *non-power law graphs*
  - We compared to a pre-MTGL C code on the MTA-2
  - 1 order of magnitude raw speed
  - 1 order of magnitude processor efficiency

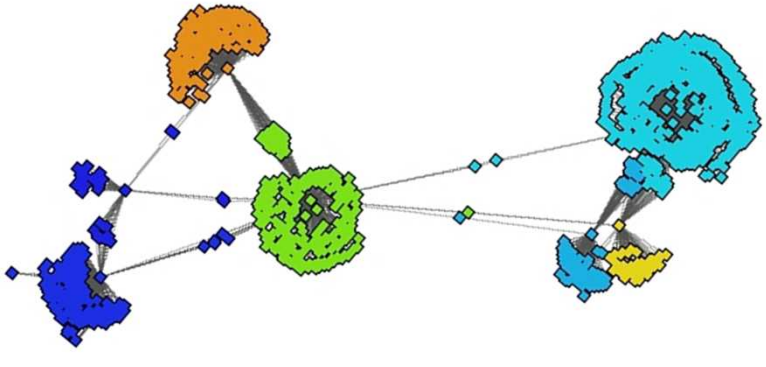


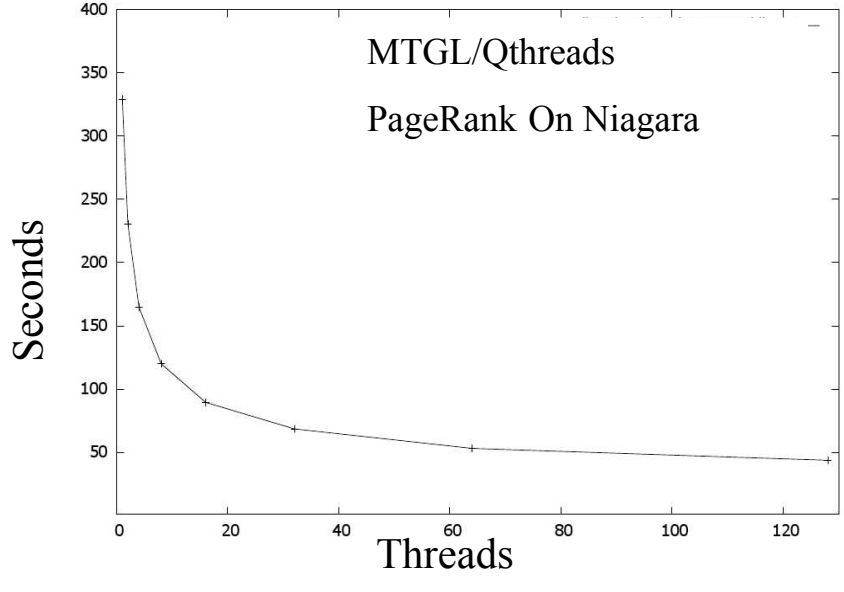
Even when distributed memory approaches scale, massively multithreaded approaches are currently faster and more efficient.

The MTGL Is Having An Interdisciplinary Impact

- **Algorithms/architectures/visualization integration**
  - Sandia architects profiled MTGL to predict performance on XMT
  - Titan visualization framework uses MTGL
  - Qthreads/MTGL → X-caliber driver application
- **Scalable facility location on MTA-2**
  - Based on expertise gained in EPA sensor placement WFO project
  - Applications to community detection, sensor placement, …







Sandia National Laboratories

Impact of HPC Informatics Activities

- **New Work**
  - Graph informatics and the MTGL has become a new business area
- **LDRD**
  - The Networks Grand Challenge LDRD is building on MTGL's success
- **Industry**
  - 2005 WFO project helped justify the Cray XMT
- **Scholarly community**
  - 3 algorithms track papers 1st MTAAP (2007)
  - Opening talk at 2nd MTAAP (2008)
  - IEEE CiSE Special Issue on Combinatorial Computing
  - DIMACS shortest paths challenge
  - Indiana University collaboration: Parallel Processing Letters, BGL refactor