Cray XMT Overview
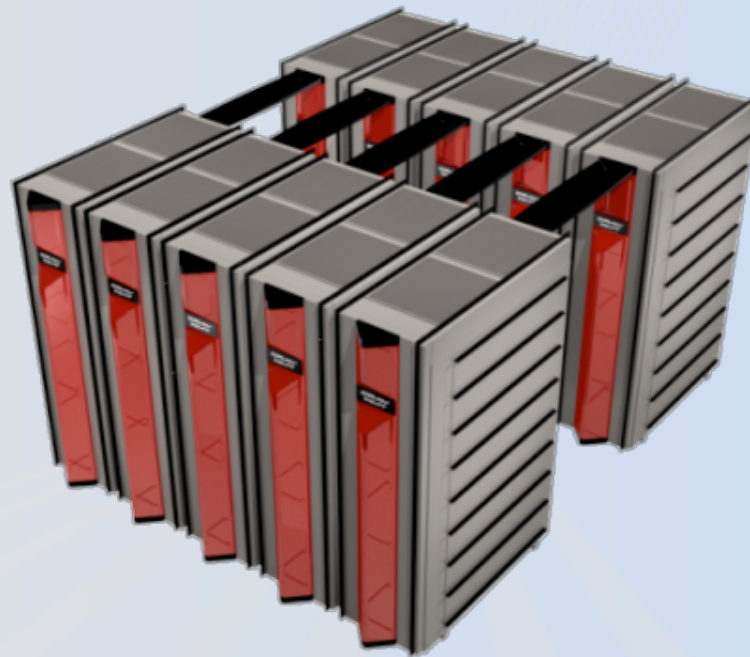
Shoaib Mufti
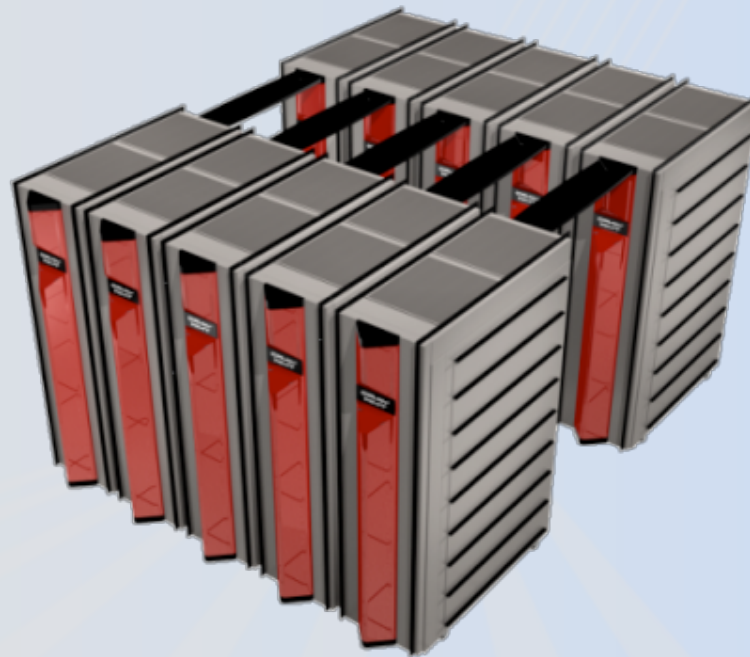Director, Program Management
18 November, 2008

Cray XMT (Extreme Multi-Threading)

- **Product Value Proposition and Overview**
- **Performance Comparisons**
- **Program Status**

Cray XMT (Extreme Multi-Threading)

- **Product Value Proposition and Overview**
- **Performance Comparisons**
- **Program Status**

Large Data Computing – New Challenges

- **Growing size of on-line content and new frontiers in science are creating applications that require processing of a massive amount of unstructured low locality data efficiently**
  - **Challenge: To transform massive amount of unrelated data into real time information that enables vital discoveries and timely decisions. Performance depends on data access patterns (e.g. locality)**
    - **Need a new kind of data-centric computing**

- **Graph theory is ideal to implement algorithms to process a massive amount of unstructured data for these new applications**
  - **Challenge: Conventional distributed and shared memory systems do not perform well in analyzing data represented by large complex graphs as the data cannot be prefetched or cached**
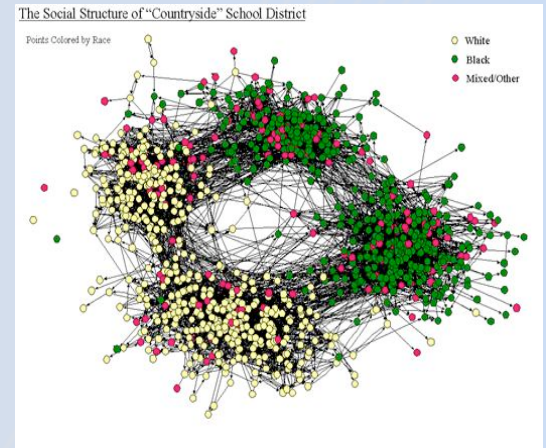   - **Need a new system to process these large graphs efficiently**
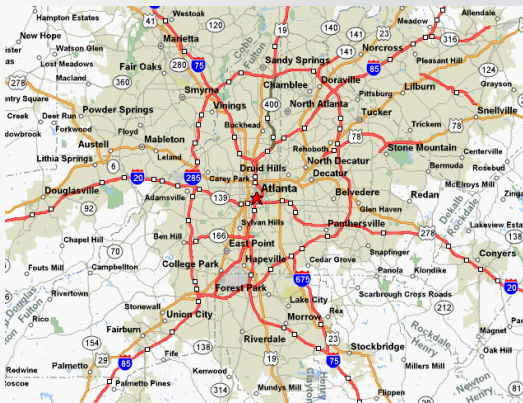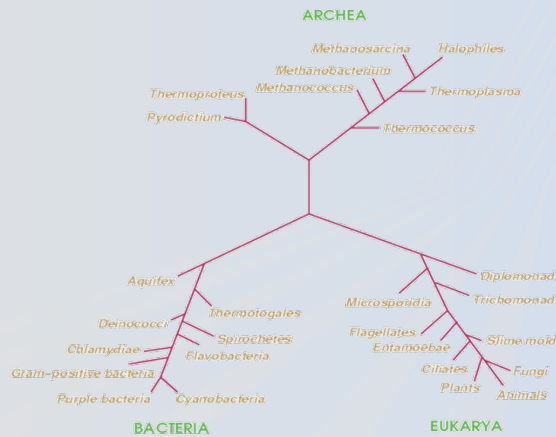
Power Distribution Networks
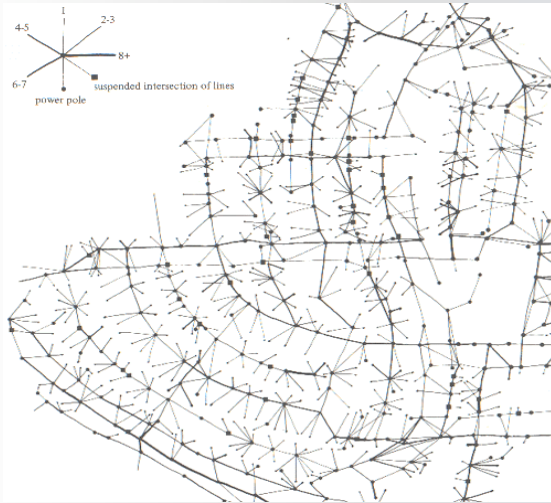
Internet backbone
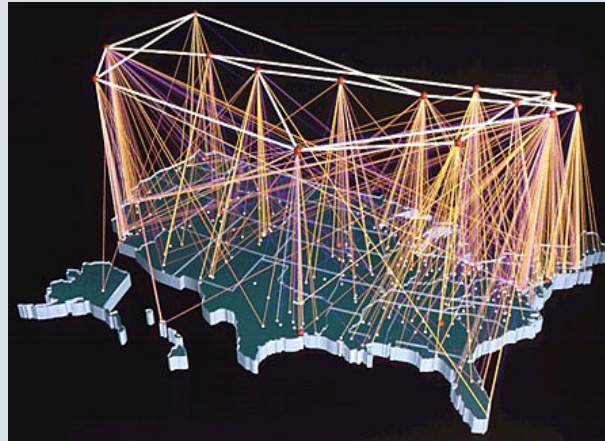
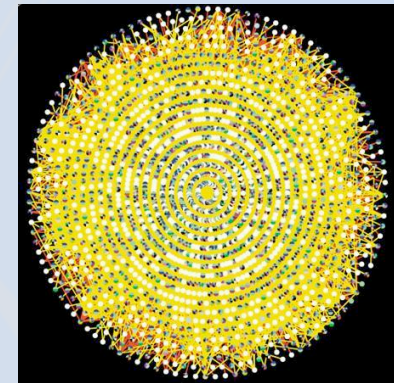**Social Networks**

Graphs are everywhere!

Ground Transportation

Tree of Life

Protein-interaction networks

Our First 100 Million

- by [Mark Zuckerberg](#)
  Tuesday, August 26, 2008 at 9:21a

- We hit a big milestone today -- **100 million people** around the world are now using Facebook. This is a really gratifying moment for us because it means a lot that you have decided that Facebook is a good, trusted place for you to share your lives with your friends. So we just wanted to take this moment to say, "thanks."

- **Tipping Point:** Malcolm Gladwell, in *The Tipping Point*, identifies three personality types that play central roles in epidemic/viral spread: Connectors, Mavens, and Salespeople. **We can identify, for example, Connectors who are people who bridge between social communities.**

The problem is too important to ignore…

- Protect the world against pandemic spread of disease, e.g. Avian flu
- Unravel the mysteries of the HIV virus
- Discover proteins implicated in cancer
- Gain intelligence from massive social networks
- Detect anomalies in billions of credit transactions
- Conserve energy by power transmission and distribution analysis

How do we process these Graphs?

- **Challenges**:
  - Runtime is dominated by latency
    - Random accesses to global address space
    - Perhaps many at once
  - Essentially no computation to hide memory costs
  - Access pattern is data dependent
    - Prefetching unlikely to help
    - Usually only want small part of cache line
 - Potentially abysmal locality at all levels of memory hierarchy

- Desired Features:
  - Low latency / high bandwidth
    - For small messages!
  - Latency tolerant
  - Light-weight synchronization mechanisms
  - Global address space
    - No graph partitioning required
    - Avoid memory-consuming profusion of ghost-nodes
    - No local/global numbering conversions
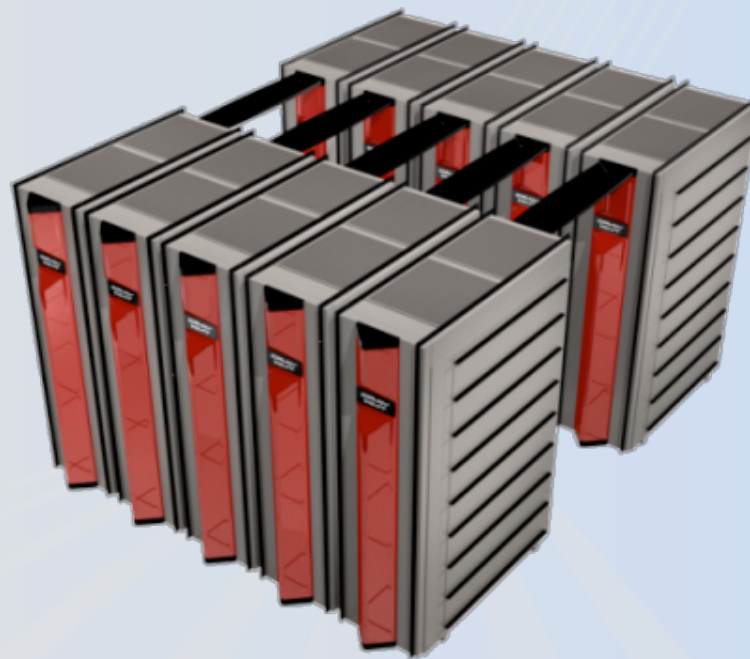 - One machine with these properties is the **Cray XMT**

Cray XMT

- Architected for large-scale data analysis
- Exploits thousands of parallel threads accessing large irregular datasets
  - Hardware supports 128 concurrent threads per processor; runtime software supports "oversubscription"
  - Scalable to over 8000 sockets and 1M threads
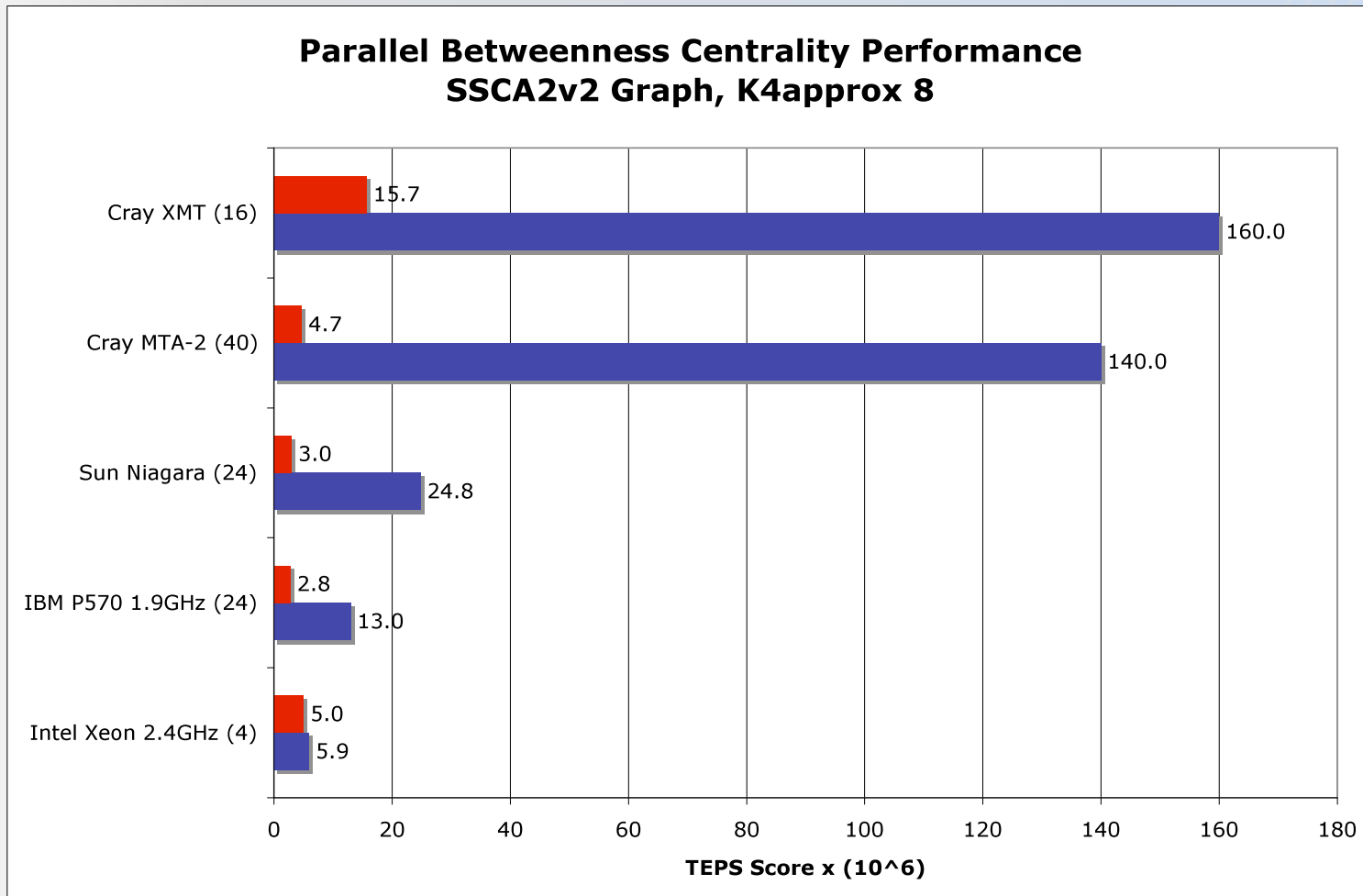 - Scalable to 128 terabytes of shared memory
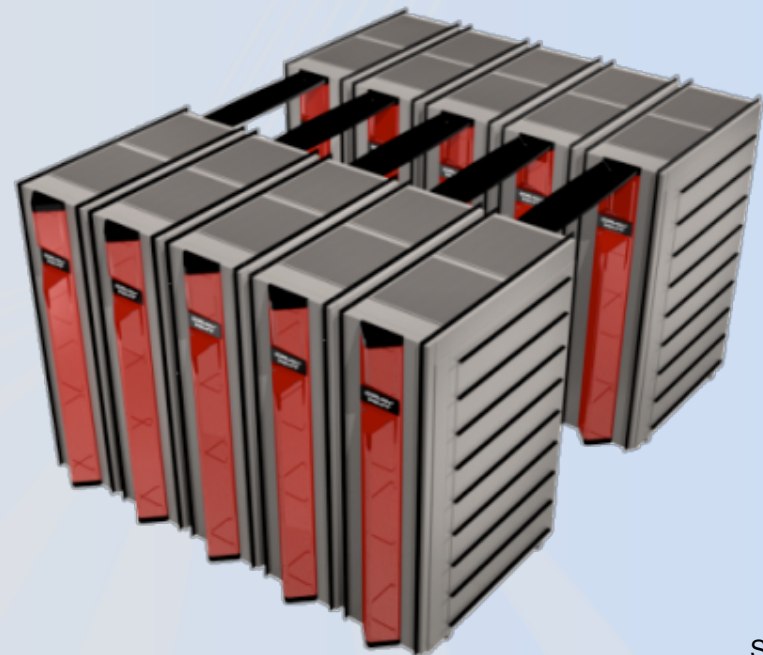
Cray XMT (Extreme Multi-Threading)

- **Product Value Proposition and Overview**
- **Performance Comparisons**
- **Program Status**

SSCA2 TEPS Performance Comparison

Parallel Betweenness Centrality Performance
SSCA2v2 Graph, K4approx 8

Cray XMT (16): 15.7 (Single Processor), 160.0 (All Processors)
Cray MTA-2 (40): 4.7 (Single Processor), 140.0 (All Processors)
Sun Niagara (24): 3.0 (Single Processor), 24.8 (All Processors)
IBM P570 1.9GHz (24): 2.8 (Single Processor), 13.0 (All Processors)
Intel Xeon 2.4GHz (4): 5.0 (Single Processor), 5.9 (All Processors)
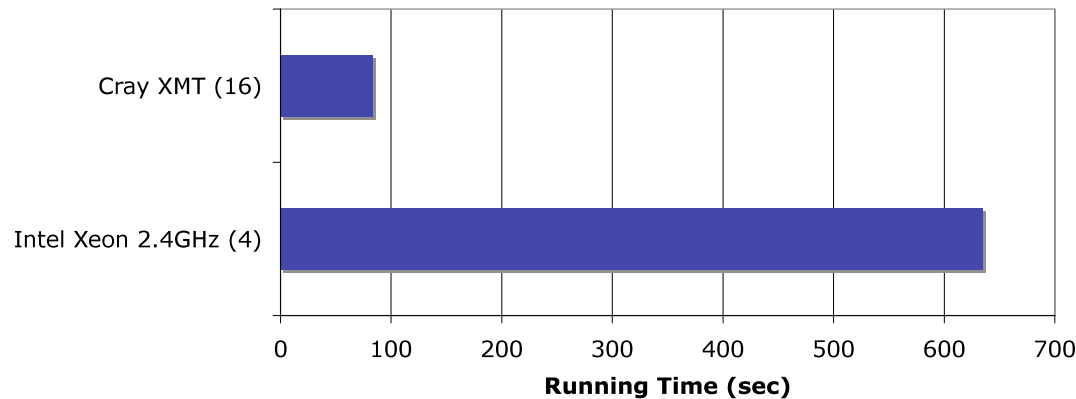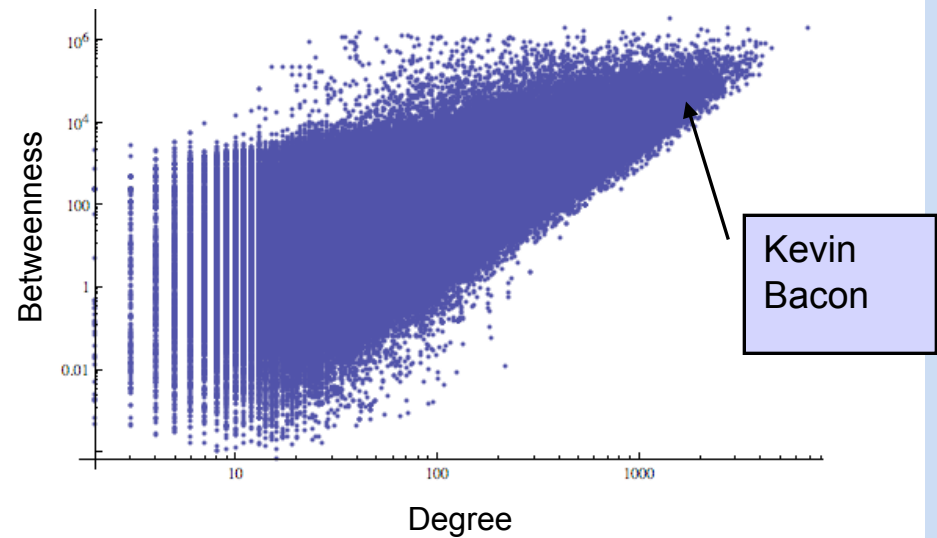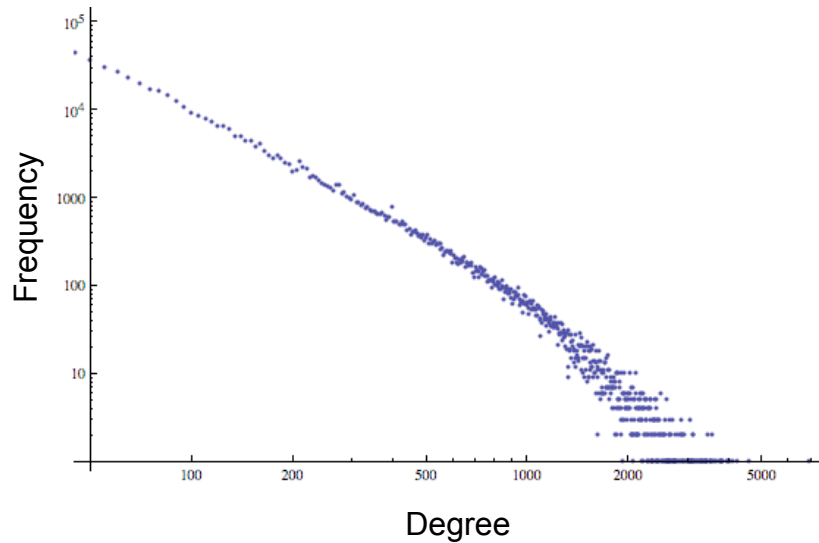
TEPS Score x (10^6)

Single Processor   **All Processors**

courtesy of David Bader, GA Tech
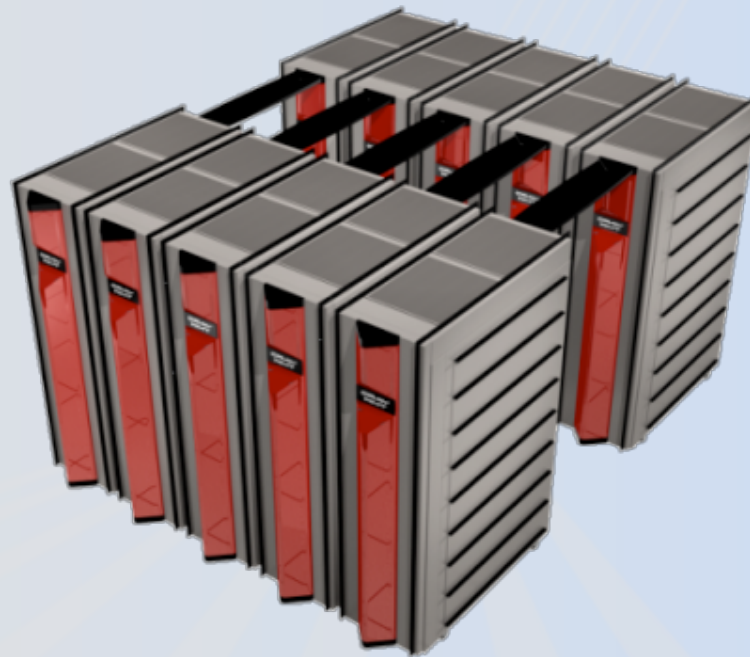
IMDB Movie Actor Network

An undirected graph of 1.54 million vertices (movie actors) and 78 million edges. An edge corresponds to a link between two actors, if they have acted together in a movie.

courtesy of David Bader, GA Tech

Cray XMT (Extreme Multi-Threading)

- **Product Value Proposition and Overview**
- **Performance Comparisons**
- **Program Status**

Program Status

- **Great progress during the last year!**

- **Reached greater scaling**
  - 256P system is stable at Cray
  - 512P build underway

- **Shipped multiple revenue systems**
  - Over 200 processors in the field
  - Systems are very stable at customer sites

- **Formed new alliances**
  - CASS-MT
    - Focus on XMT applications and system software
    - PNNL, DOD, Georgia Tech, Sandia National Labs, Cray
  - NSF computing research Infrastructure for multi-threading
    - Research platform is XMT
    - Univ. of Notre Dame, Univ. of Delaware, UC Santa Barbara, CalTech, UC Berkeley, Sandia National Labs
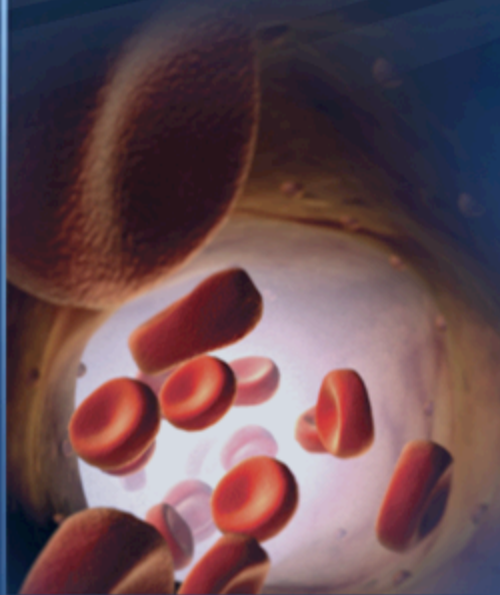
Program Status -- Continued

- **Training**
  - A formal training course will be offered to XMT users in Q1 2009
  - Offering both advanced and beginner training classes
  - Accepting enrollments
  - Please enroll if you are interested

Acknowledgements

- David Bader -- Georgia Tech

- Kamesh Madduri -- Lawrence Berkeley National Laboratory

- Daniel Chavarria -- PNNL

Shoaib Mufti (shoaib@cray.com)