

CCSD Ultra-scale Computing Initiative

LDRD Information Meeting

Agenda –

- Computer Science & Math
- Computational Science
- Knowledge Discovery
- Information/Cyber Security

FY09 LDRD
Ultrascale Computing Initiative
Computer Science and Mathematics
Point of Contact: Al Geist

In FY09 ORNL has two new crosscutting initiatives

1. Understanding Climate Change Impacts: Energy, Carbon and Water
2. Emerging Science and Technology for Sustainable Bioenergy

Math and CS proposals targeted to either of these areas should be submitted to the new crosscutting initiative

And not to the Ultrascale call.

ASCR call LAB 08-19 is out. *“Software Development Tools for Improved Ease of use of Petascale Systems”* proposals in this area should be submitted to DOE not to Ultrascale LDRD.

FY 2009 Exciting year for CS and Math

For the past five years the CS and Math part of the LDRD call has focused on helping birth the Leadership Computing Facility.

(topics were practical, short horizon, more development than research)

Because of this the CS and Math proposals did not compete well against the Ultrascale computational science proposals, which had no restrictions on proposing visionary, high impact research.

Starting in FY2009 there is no longer a restriction on the CS and Math LDRD proposals with respect to the LCF.

By the same token, the bar goes up. We are looking for exciting, high impact, high visibility CS and Math research proposals that will give ORNL a national reputation as a leader in this CS or Math area.

New Strategic Focus Areas

At a recent off-site strategic planning meeting a list of CS and math focus areas were generated.

Criteria for choosing these focus areas:

- Useful for practically all applications
- Will Benefit all three of our sponsors (DOE, DoD, NSF)
- Can be leveraged through ORNL's new software centers
 - Institute for Advanced Architectures and Algorithms
 - Extreme Scale Software Center

Authors of FY2009 LDRDs in CS and Math should consider proposals that fall in one of the following four focus areas.

CS and Math Focus Area #1

Data Analytics

- Becoming national leaders in developing the math and CS algorithms and create practical packages to analyze huge volumes of experiment, simulation, and/or sensor data.
- Consider cases where data may be dynamic, high-dimensional, huge volume, and potentially distributed. Consider cases where data may have high error rates, be corrupted, or compromised.
- New data analysis algorithms that may in turn be used to design or guide experiments or simulations in a feedback loop.

Focus #1 The bottom line is proposing math or CS research that helps turn huge volumes of data into information and helps turn this information into knowledge.

CS and Math Focus Area #2

Programming Models

- Due to changes in architecture we will see in the next 5 years there is the opportunity to become a leader in the development of math algorithms and CS techniques to handle
 - multi-core and many-core chips
 - heterogeneous cores and hybrid architectures (accelerators)
 - fault tolerance
 - more productive methods to program such systems
 - becoming a leader in the next generation of MPI

Focus #2 The bottom line is proposing math or CS research that improves the performance, and or productive use of future multi-core, heterogeneous architectures.

CS and Math Focus Area #3

Multi-scale Application Techniques

Can include lots of things we could develop a leadership reputation in.
Here are a couple ideas feel free to propose your own:

- New parallel math algorithms to handle multi-scale in time as well as algorithms to handle multi-scale space problems
- Composable applications - the ability to build large applications by plugging together smaller pieces into larger and larger (more complex) simulations or analysis packages

Focus #3 The bottom line is proposing math or CS research that would build a national reputation at ORNL for handling some aspect of multi-scale applications.

CS and Math Focus Area #4

Emerging/frontier technologies for computing

- What lies ahead in the post-CMOS era?
- CMOS runs out of atoms around 2015 (devices only 6 atoms wide)

Focus #4 The bottom line is proposing math or CS research that moves beyond incremental advances to explore radical innovations in hardware and architecture of computing and communication systems that use new methods, such as photonics, optics, and nano-scale devices.

Final Words of Advice

Do not propose an LDRD that extends or supplements your existing CS or Math research. That is not the purpose of LDRD funds.

Make a strong case for the impact of your proposed research either to science or in funding to ORNL

- Don't just say DOE apps will really love my work and DOE, DoD, and NSF will want to give me bags of money. Blah, blah, blah.
- To make a strong case you have to be specific, do your homework.
- Consider LDRD funds like Venture Capital and you are trying to make a pitch to review panel of venture capitalists.

Good Luck!

FY09 LDRD
Ultrascale Computing Initiative

Computational Sciences
Point of Contact: Jeff Nichols

Energy Systems Simulation

High priority on the development of predictive capabilities in the area of energy systems simulation in programs such as the Advanced Fuel Cycle Initiative (AFCI), the Fusion Simulation Program (FSP) and the International Thermonuclear Experimental Reactor (ITER).

- **Address the system-scale operation issues of modern and future energy systems that are needed to model advanced energy production and distribution systems.**
 - higher spatial resolutions,
 - higher temperature operating regimes,
 - longer simulation times,
 - nonlinear phenomena, and
 - higher model dimensionality
- **All relevant related issues, e.g., efficiency, fuel cycles, safety, accident scenarios, and materials effects are of interest.**
- **New areas of simulation needed to design the devices beyond ITER and to minimize the lead time for commercialization of fusion – materials and fusion engineering.**
- **Program development opportunities similar to that in FSP – project definition and design, prototype projects, leading to full-scale simulation.**

Computational Systems Biology/Medicine

The focus is on capability computational systems biology/medicine and any associated data and/or data center needs.

- **Examples of capability computing topics within computational systems biology/medicine include:**
 - data-science applications with very large data sets,
 - stochastic regulatory and protein interaction networks, and
 - large-scale biophysical simulations, such as molecular dynamics simulations of lipid membranes and membrane channels, and ribosomal interactions and dynamics.
- **When scaled to humans, the integration of systems genetics/genomics and systems biology creates the entity we refer to as systems medicine.**

Nanoscience

The development and application of new molecular-scale modeling, including classical, quantum, and hybrid classical-quantum approaches, for solutions and interpretations to the most challenging problems in biology, chemistry, physics, materials science, nanotechnology, and biotechnology.

- **Computationally intensive, numerical implementations of fundamental molecular-level descriptions,**
- **Simulations of a size and duration sufficient to be predictive in areas such as materials synthesis, catalysis, and biology,**
- **Development and application of these methods for world-class science problems in molecular biology, biotechnology, chemistry, nanoscale science and technology, and materials science.**

Good Luck!

FY09 LDRD Ultrascale Computing Initiative

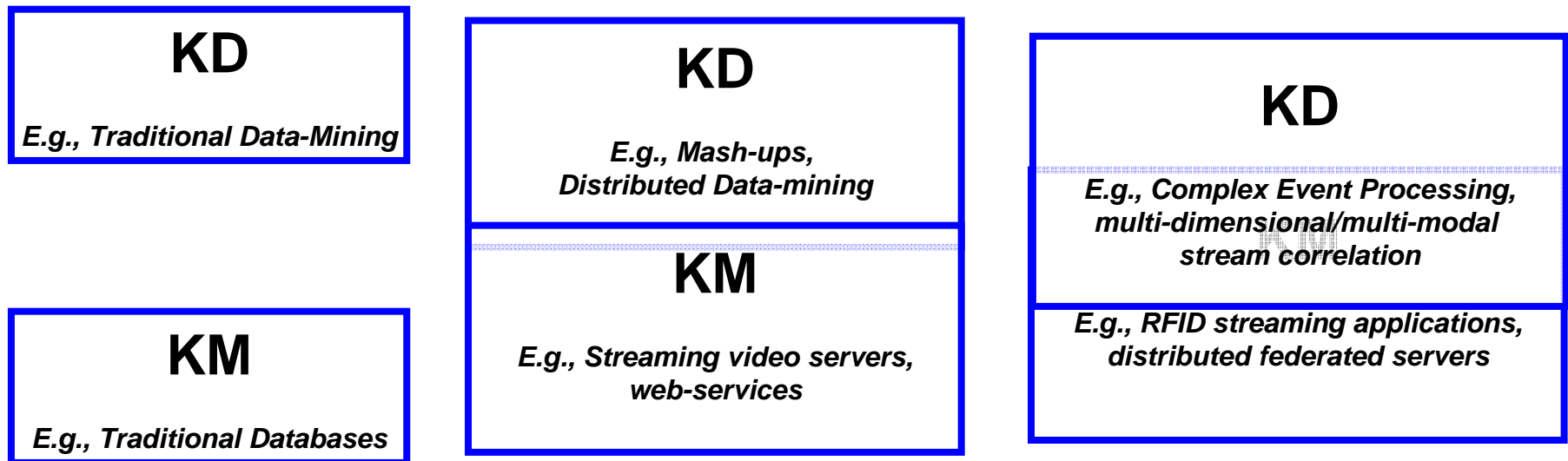
Knowledge Discovery Point of Contact: Brian Worley

ORNL Is Committed to the Knowledge Discovery Agenda

- Entire Research Division Focused on Knowledge Discovery
- Appropriate Resources: HPC, Networking, MRF, JICS
- LDRD Initiative in Knowledge Discovery
- Programmatic efforts well-aligned with this science agenda



Knowledge management and knowledge discovery are becoming indistinguishable processes.



Percent of Data that is Real-Time

Our largest set of projects is research and development related to the collection, analysis, and dissemination of sensor data.

- **Interdiction, detection, emergency response**
 - Mobile, Transportation Corridors, Ports, Military Bases
- **Real-Time Data Management**
 - Collection, Dissemination, Archiving
- **Pre-deployment analysis**
 - Cost, Performance Prediction, Risk vs Benefit
- **Wide-area ubiquitous sensing, actuation, and deployment**
 - Orchestrating the functionality across a large system of distributed sensors/processors (eg Electric Grid, Autonomous robotic systems)



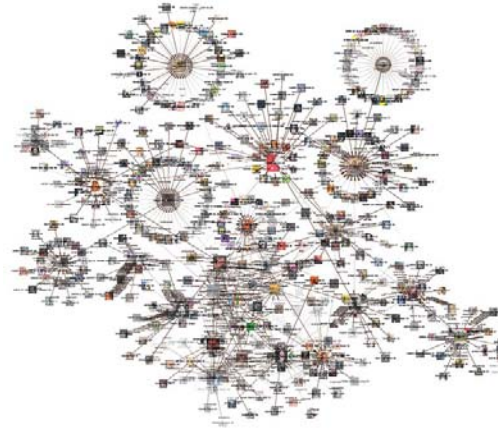
- **Cross-agency and cross-administrative boundary data-sharing and interoperability**
 - Standards and policies
- **Net-Centric Services**
- **Security, Access Controls**



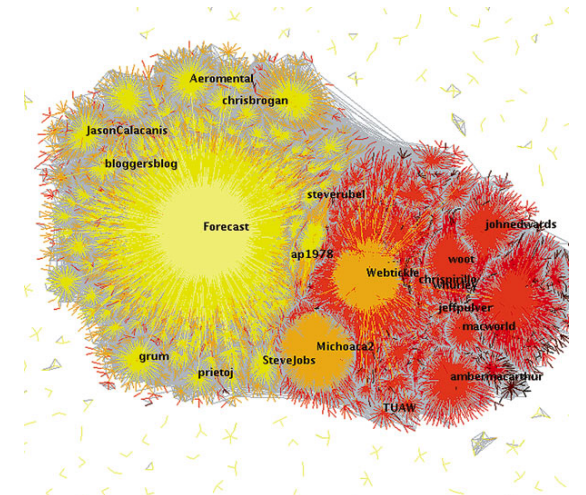
Social Networks Analysis



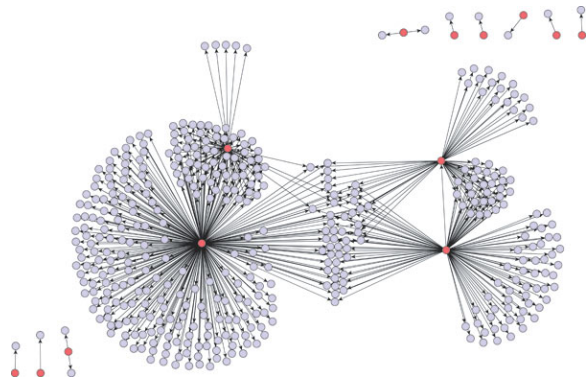
Blogosphere



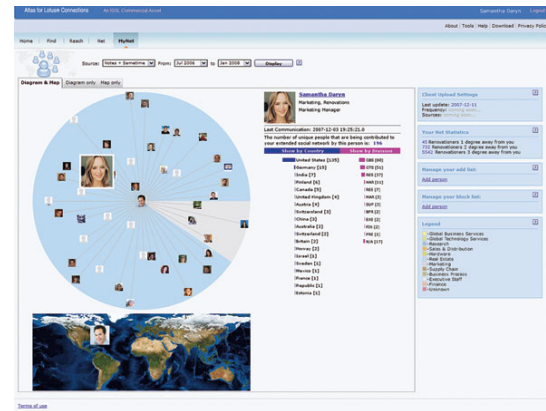
Comment Flow



Twitter Social Network



Viral Marketing

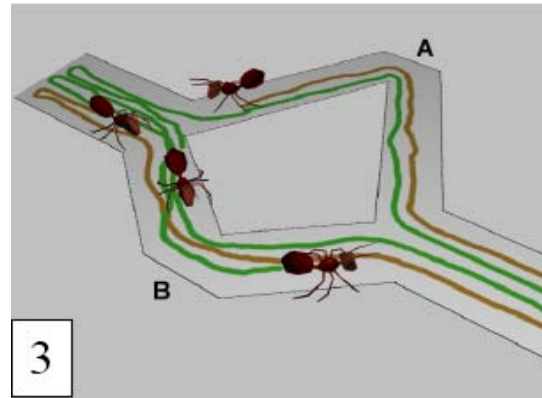


Workplace Networks

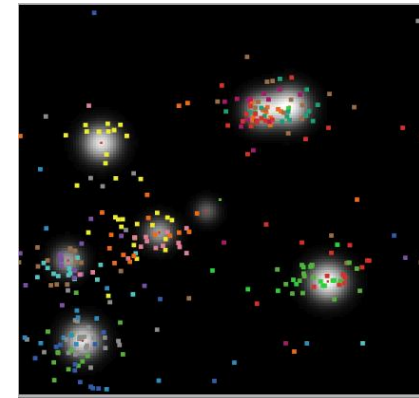
Emergent Behavior in Social Systems



Birds flocking



Ant pathways



Insurgent relationships

Agent-based simulations

Discrete-event simulations

Cultural Based Data Evaluation

- **Language context and meaning**
- **Image context and meaning**
- **Information fusion across languages**
- **Content extraction**
- **Disambiguation of phrases and terms**

Virtual Worlds to Explore Social Behaviors



Second Life – Linden Lab

Education

Tourism



Collaboration

Shopping

Interviews

Managed by UT-Battelle for the
U. S. Department of Energy



Good Luck!

FY09 LDRD
Ultrascale Computing Initiative

Information and Cyber Security
Points of Contact: Scott Studham, Chris Kemper

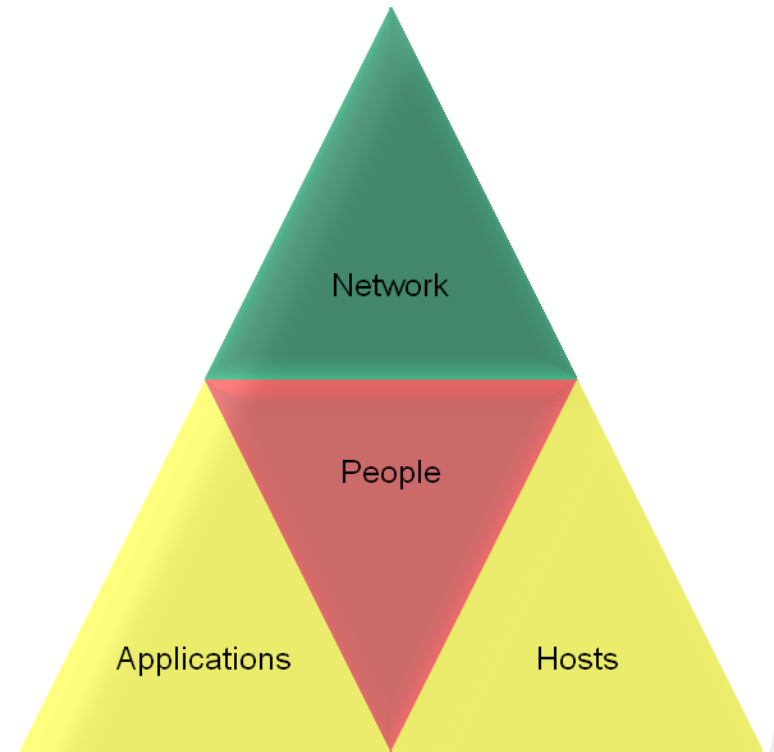
State of Cyber Security

The cyber attack environment we live in includes:

- **Sophisticated Nation States**
- **12,000,000 SPAM/Phishing emails per month. For the ones that get through, users have a 10% “click rate”.**
- **Organized crime interested in personally identifiable information.**

Current attacks are targeting hosts and people.

Public social networking sites will increase sophistication of attacks on applications and people



A Cyber Program is as strong as the weakest component

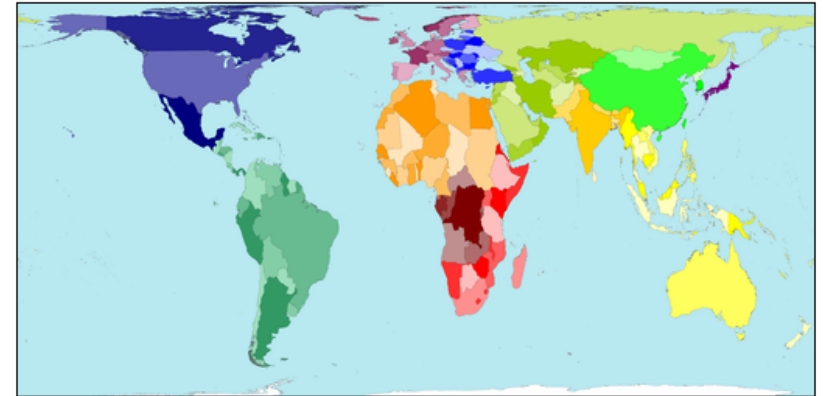
Cyber opportunities

- DOE Cyber R&D Town halls will likely lead to FY09 DOE-SC funding
- OMB Desktop Standard-Federal Desktop Core Configuration
- OMB Trusted Internet Connection (TIC) Initiative
 - June 2008; Comprehensive plan of action/milestones-target completion date
- Counterfeit IT products
- DoD Considering Ban on Personal Use of Internet (1/28/2008)
- Classified Presidential Directive
- System & Software Assurance Guidelines-Information Assurance

< Previous Map

Land Area

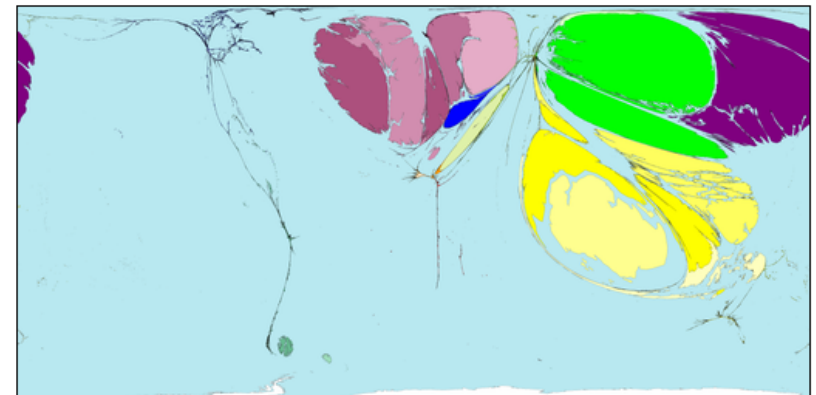
Map No. 1



< Previous Map

Computers Exports

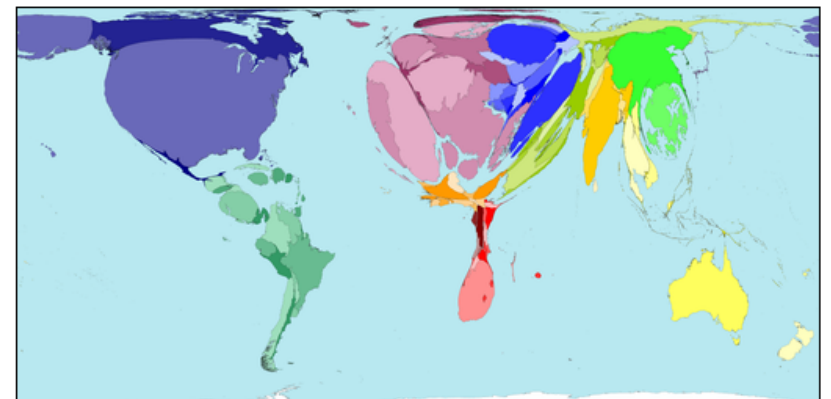
Map No. 91



< Previous Map

Computers Imports

Map No. 92



Leverage ORNL R&D skills to help our customers.

- **Focus on “Defensive Cyber Security”**
 - Identify and evaluate risks and threats
 - Detect/prevent attacks, and develop attack mitigation strategies
 - Improve protections for electronic information, systems, networks and infrastructure
- **Innovation is also sought in developing inherent protections or forensic trails for stored data with the assumption that an adversary has gained network or even physical access.**

Suggested Proposal Areas

- **Advanced risk analysis for complex systems, including HPC systems,**
- **Information analytics,**
- **Mathematical models for detecting and characterizing anomalous network activity,**
- **Predictive modeling of attacker activity,**
- **Technologies and techniques for identifying adversaries across the Internet despite their efforts at anonymity,**
- **Hardware-enabled security for specialized computing environments,**
- **Forensic information fusion for active response,**
- **Visualization in support of cyber-security, real-time situational awareness of attacks in progress including those that have never been seen before, and**
- **Quantum-enhanced security: novel approaches to exploiting the properties of quantum systems for security enhancement.**

Good Luck!