Welcome to Notre Dame!

CCL Workshop 2013

www.nd.edu/~ccl/workshop/2013
The Cooperative Computing Lab
University of Notre Dame

http://www.nd.edu/~ccl
Our goal:

Make it easy to scale up real applications from a single laptop to national infrastructure, enabling new discoveries.
The Cooperative Computing Lab

- We *collaborate with people* who have large scale computing problems in science, engineering, and other fields.
- We *operate computer systems* on the O(10,000) cores: clusters, clouds, grids.
- We *conduct computer science* research in the context of real people and problems.
- We *release open source software* for large scale distributed computing.

http://www.nd.edu/~ccl
CCL Team Changes

Ben Tovar, Ph.D.
Principal Research
Software Engineer

Li Yu
Financial Software
Bloomberg

Principal
Hadoop Developer
AT&T Foundry

Dinesh Rajan

Patrick Donnelly

Peter Sempolinski

Casey Robinson

Peter Ivie

Haiyan Meng

Nick Hazekamp
Current Focus Areas

• Right Sizing Systems and Applications
  – I have this workload. How many machines (and what size) do I actually need to run it?

• Scaling Up without Blowing Up
  – 10M tasks on 10K workers?

• Data Intensive Applications
  – Move code to data and share across workers, across applications, across users....

• Evolving with Our Ecosystem:
  – Hadoop, Galaxy, GPUs, iRODS, OSG, XSEDE
dV/dt : Accelerating the Rate of Progress Towards Extreme Scale Collaborative Science

Miron Livny (UW), Ewa Deelman (USC/ISI), Douglas Thain (ND), Frank Wuerthwein (UCSD), Bill Allcock (ANL)

... make it easier for scientists to conduct large-scale computational tasks that use the power of computing resources they do not own to process data they did not collect with applications they did not develop ...
Resource Management Cycle

Allocate Resources

Historical Repository

Observed Resources

Measurement and Enforcement

Exception Handling
Is it an outlier?
CCTools in Education

• UW – Eau Claire: IEEE Cluster 2013 Best Education Outreach and Training Paper:
  – Peter Bui, Travis Boettcher, Nicholas Jaeger, Jeffrey Westphal, “Using Clusters in Undergraduate Research”

• MF and WQ in University of Arizona Advanced Cyberinfrastructure Class (ACIC)

• WQ in Notre Dame Undergraduate Programming Paradigms Class
Your Highlight Here!
Agenda

• **9:00 Introduction**
  – Welcome and Introduction
  – What’s new in the CCTools – CCL Team

• **10:15 Coffee Break**
  – Stash and Skeleton Key on the Open Science Grid, Rob Gardner, University of Chicago
  – Swift+Chirp for Synchrotron Beamline Data Analysis, Justin Wozniak, Argonne National Lab
  – Big Data and Data Science: The View from NIST, Geoffrey Fox, Indiana University

• **Noon - Lunch**
  – Scaling Up CMS Tier-3 Data Processing, Kevin Lannon
  – HotRAD: Harnessing distributed systems for population genetics sampling Scott Emrich
  – Data Management Challenges in the iPlant Collaboration, Nirav Merchant
  – Massively Parallel Molecular Dynamics Using Adaptive Weighted Ensemble, RJ Nowling

• **2:30 - Coffee Break**
  – Weather Data Processing with Makeflow and SLURM, Neil Best
  – Using Work Queue Inside and Outside the Classroom, Peter Bui
  – Systematic development of accurate model potentials using the Work Queue distributed computing environment, Lee-Ping Wang
  – Large Scale Image Processing for Biometrics, Joseph Thompson

• **4:00 Discussion and Adjourn**