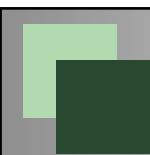


# Do More, Faster: Leveraging Computational Resources in Your Research

Dr. Dirk Colbry  
Institute for Cyber Enabled Research  
Adjunct Faculty, Electrical and Computer Engineering  
Michigan State University

MICHIGAN STATE UNIVERSITY © 2012 Michigan State University Board of Trustees



## Agenda

- Who am I
- What is iCER / HPCC
- Steps to High Performance
- Common classes of computational science problems

MICHIGAN STATE UNIVERSITY



# I am an engineer

- Mechanical Engineering degree from Georgia Tech
  - 3 years as a Mechanical Engineer for Delta Airlines in Atlanta
  - 2 years as a Robotics Engineer for FANUC Robotics in Auburn Hills



# I am a researcher

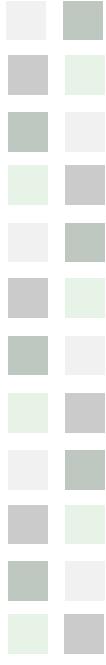
- Ph.D. in CSE from MSU, focusing on image processing and pattern recognition
  - Adjunct Faculty in ECE
  - Research interests include optimizing image analysis workflows for researchers
  - Currently collaborating with researchers in Engineering, Biology, Math and Statistics





# I am a computational consultant

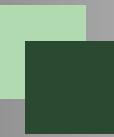
- One-on-one consulting
- HPC Programming
- Proposal Writing
- Training and Education
- Outreach
- Reduce the “Mean time to Science”



## Agenda

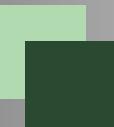
- Who am I
- What is iCER / HPCC
- Steps to High Performance
- Common classes of computational science problems





## What is Advanced Computing Hardware?

- Anything more advanced than your desktop
- Local resources
  - Lab, Department, Institution (HPCC)
- National resources
  - NSF (XSEDE), DOE (Jaguar) , Others
- Commercial Resources (cloud computing)
  - Amazon, Azure, Liquid Web, Others



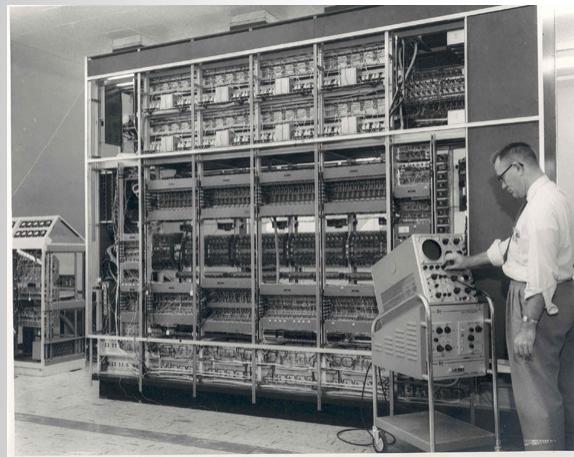
## Why use Advanced Computing Hardware?

- Science takes too long
- Computation runs out of memory
- Run out of disk space
- Need licensed software
- Need advanced interface (visualization)



## 1957 MISTIC Mainframe

- MSU's first mainframe
- Hand built by grad students
  - Dick Reid
  - Glen Keeney



MICHIGAN STATE  
UNIVERSITY



## After MISTIC

- 1957 MISTIC
- 1963-1973 CDC 3600
- 1967 Computer Science Department
- 1968 CDC 6500
- 1971 MERIT
- 1978 Cyber 750
- **2004 HPCC**
- **2009 ICER**

MICHIGAN STATE  
UNIVERSITY



# 2004 MSU HPCC

- Provide a level of performance beyond what you could get and reasonably maintain as a small group
- Provide a variety of technology, hardware and software, that would allow for innovation not easily found

MICHIGAN STATE  
UNIVERSITY

11



## What is iCER?

- Institute for Cyber-Enabled Research
  - Established in 2009 to encourage and support the application of advanced computing resources and techniques by MSU researchers
  - Goal is to maintain and enhance the university's national and international standing in computational disciplines and research thrusts

MICHIGAN STATE  
UNIVERSITY



# Bigger Science

- The goal of iCER is NOT:
  - Kflops / second
- Instead, the goal of iCER IS:
  - KSciences / second
- Doing More Science, Faster
  - Reducing the “Mean time to Science”
- iCER is designed to help researchers do their science and when appropriate scale them up to one of the national labs



MICHIGAN STATE  
UNIVERSITY



## Hardware

High speed network interconnect  
 • MPI jobs  
 • High Speed Parallel Scratch Space



### General Purpose Graphics Cluster

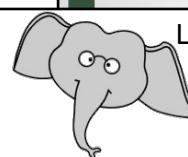


512 core, 128 node cluster installed in 2005. Each node contains four 2.2 GHz AMD Opteron cores, 8 GB of RAM, and 146 GB of local disk.

2005

2010

2010



Large Capacity “FAT” Nodes  
 • Up to 2TB of RAM  
 • Up to 64 cores

2011

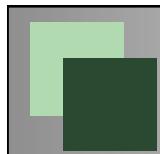
Common OS Image  
 • RHEL6.0  
 • Compile once  
 • Run anywhere



1  
 612GB - 2TB RAM nodes  
 -64 core 2.66 GHZ Xeon processors.

MICHIGAN STATE  
UNIVERSITY





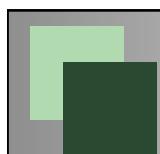
# Hardware Summary

**FREE\***

- Over 3110 Nodes
- Over 23072 CPU cores
- Over 363 TB of scratch memory space
- 2TB Shared memory machines
- 50GB backed up home directory space
- GPGPU cluster with 64 Tesla Nodes
- High Throughput condor cluster
- Specialized Bioinformatics VMs

MICHIGAN STATE  
UNIVERSITY


  
ICER

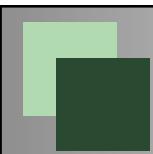


# Available Software

- Center Supported Development Software
  - Intel compilers, openmp, openmpi, mvapich, totalview, mkl, pathscale, gnu...
- Center Supported Research Software
  - MATLAB, R, fluent, abaqus, HEEDS, amber, blast, ls-dyna, starp...
- Customer Software
  - gromacs, cmake, cuda, imagemagick, java, openmm, siesta...
  - For a more up to date list, see the documentation wiki:
    - <http://wiki.hpc.msu.edu/>

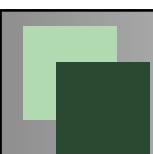
MICHIGAN STATE  
UNIVERSITY


  
ICER



## Accounts

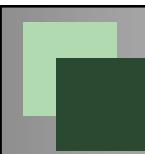
- Each user has 50Gigs of backed-up personal hard drive space
  - /mnt/home/username/
- Users have access to 363TB of high speed parallel scratch space
  - /mnt/scratch/username/
- Shared group space is also available upon request



## Buy-In Opportunities

- Every 6-12 months
- We will maintain your computers for you
- Researchers get exclusive use of their nodes within 4 hours of submitting a job
- Buy-in jobs will automatically overflow into the general resources

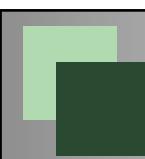




# More than Just Hardware

- Education
- Grant Writing / Support
- Advancing Computational Science

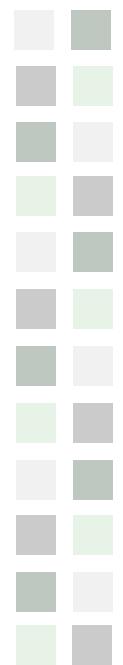
MICHIGAN STATE  
UNIVERSITY

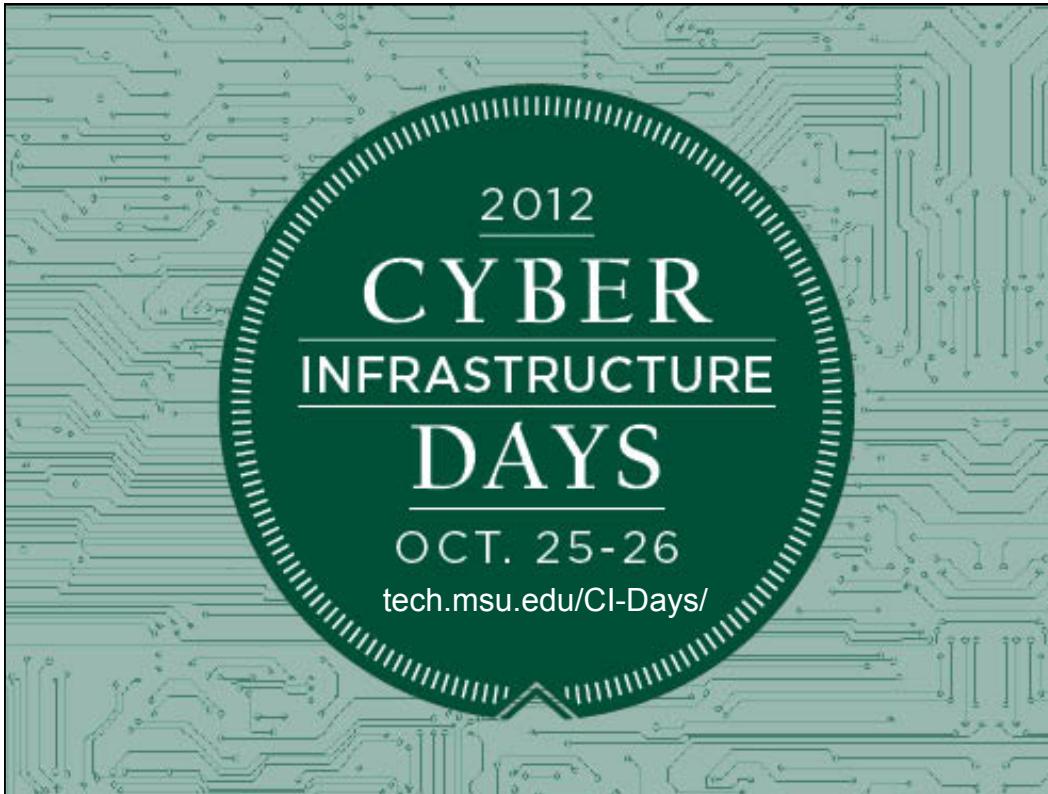


# Grant Writing / Support

- Buy-in opportunities
- $\frac{1}{2}$  time post-doc support
- Part time HPCC Programmer Support
- Support finding research collaborations
- News about funding opportunities
- Technical Expertise

MICHIGAN STATE  
UNIVERSITY





# MSU Seminars in Research and Instructional Technology

*Dec 18 and 19, 2012*

- Two days of no-cost seminars to faculty and graduate students on technology topics
    - Morning sessions run from 8:30 to 11:30 am
    - Afternoon sessions run from 1:30 to 4:30 pm
    - Lunch is provided that will feature guest speakers on instructional technology
    - Introduction to HPC
    - Advanced HPC



<http://train.msu.edu/faculty/seminars/>



# We are here to help

- [www.hpcc.msu.edu/contact](http://www.hpcc.msu.edu/contact)
  - Questions
  - Schedule Consultations
  - Code Reviews
  - Programming help
  - Hardware Purchasing
  - Help with Grants
  - Support for Grants

MICHIGAN STATE  
UNIVERSITY



# Agenda

- Who am I
- What is iCER / HPCC
- **Steps to High Performance**
- Common classes of computational science problems

MICHIGAN STATE  
UNIVERSITY



# Steps in Using Resource

- Connect to Resource
- Determine required software
- Transfer required input files and source code
- Compile programs (if needed)
- Test software/programs on a developer node
- Write a submission script
- Submit the job
- Get your results and write a paper!!

MICHIGAN STATE  
UNIVERSITY



# Types of problems

- CPU bound
  - Lots of computing (think simulation)
- Memory bound
  - Requires lots of memory (think genomics)
- I/O bound
  - Requires lots of data (think astronomy)

(many problems fall in more than one category)

MICHIGAN STATE  
UNIVERSITY



# Steps to High Performance

Note: Every application is different

1. Analyze your code
  - Profilers (gprof, vtune, tau)
  - Debuggers / memory trackers (gdb, totalview)
2. Optimize calculations
  - Trade memory for time (i.e., never do the same calculation twice)
3. Find ways to parallelize
  - Look for loops
  - Find iterations independent from each other
  - Determine how much information needs to be transferred

MICHIGAN STATE  
UNIVERSITY

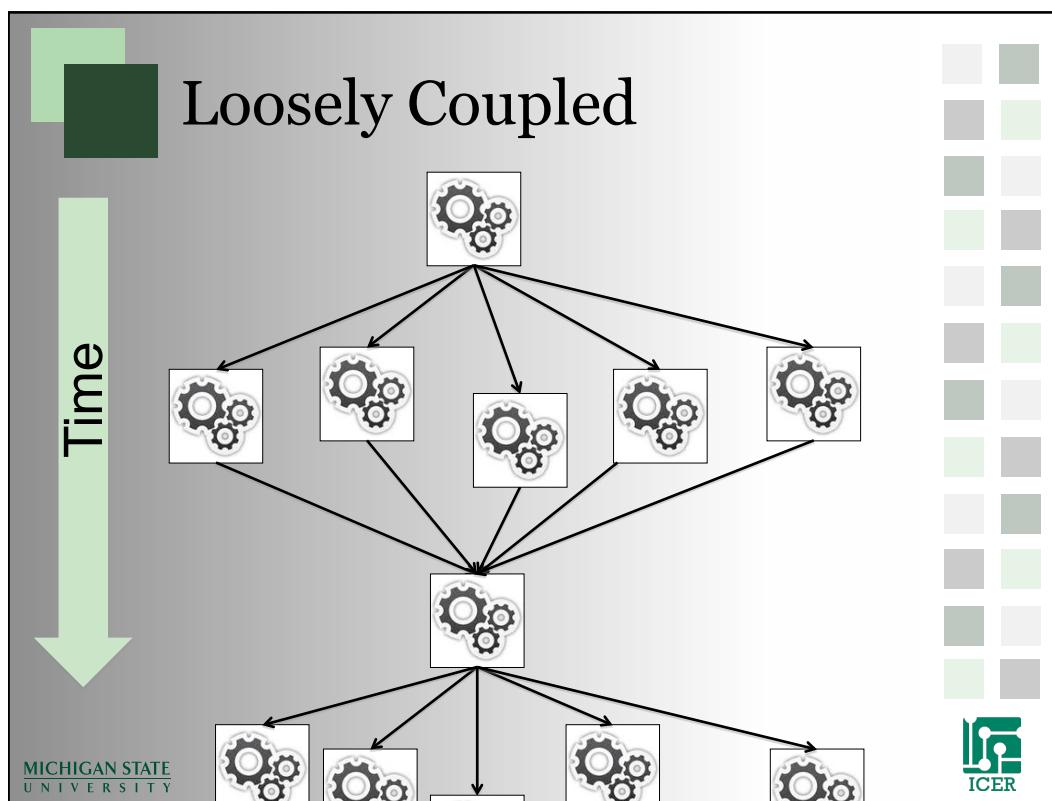
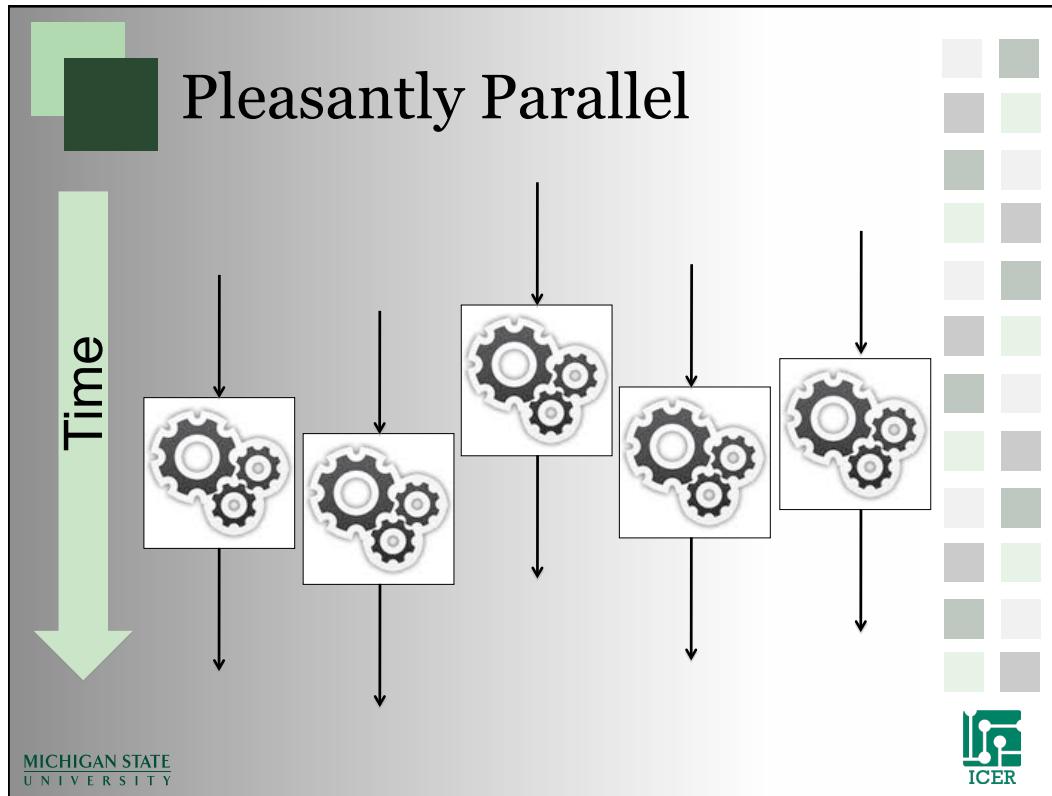


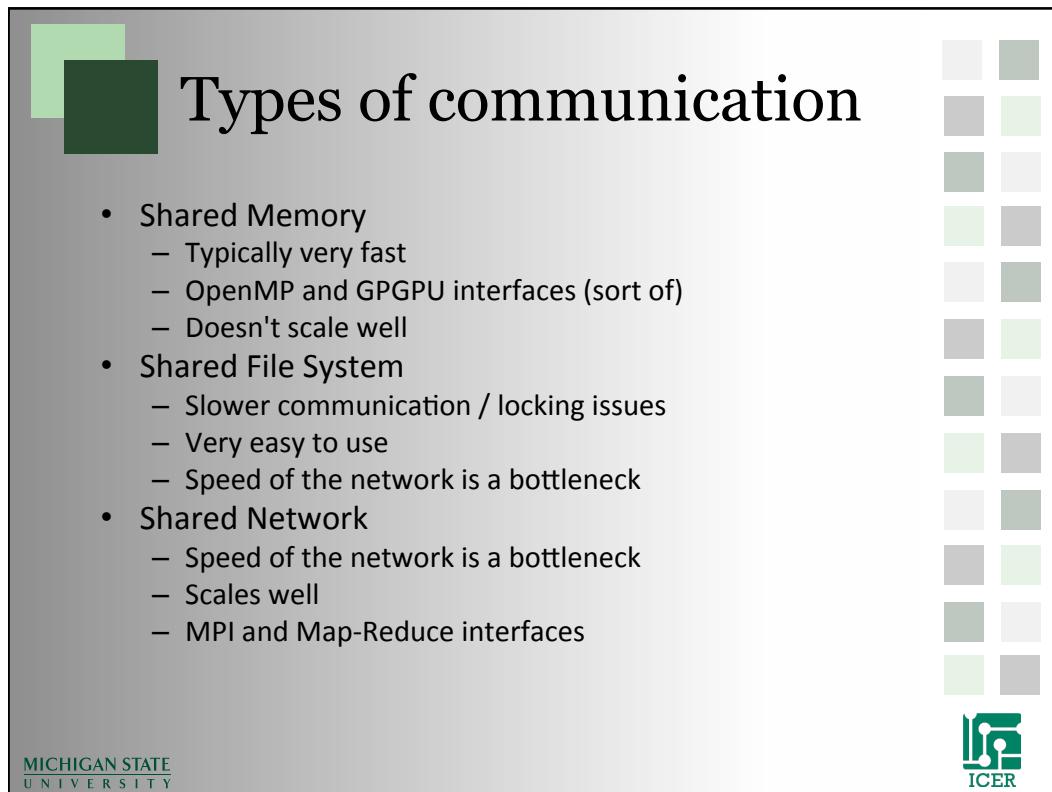
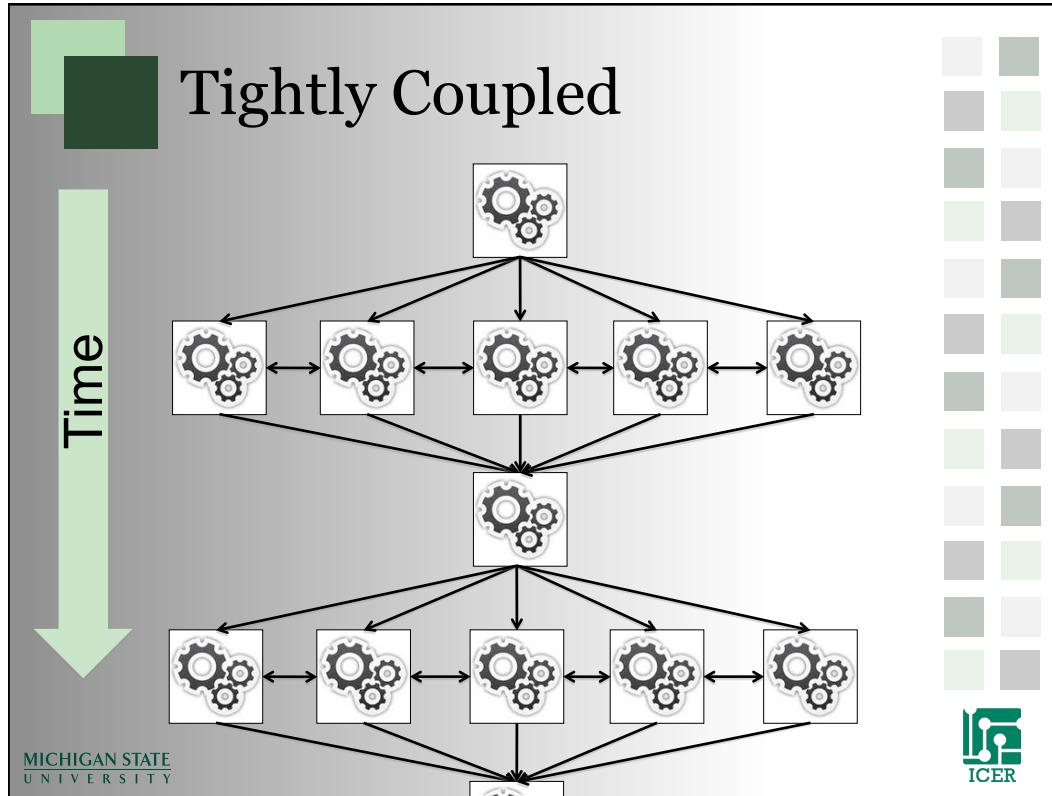
# How much Communication?

- Pleasantly parallel
  - No communication required
- Loosely coupled
  - Typically sync at regular intervals
- Tightly coupled
  - Constant communication

MICHIGAN STATE  
UNIVERSITY







# Agenda

- Who am I
- What is iCER / HPCC
- Steps to High Performance
- Common classes of computational science problems

MICHIGAN STATE  
UNIVERSITY



# Example Problems

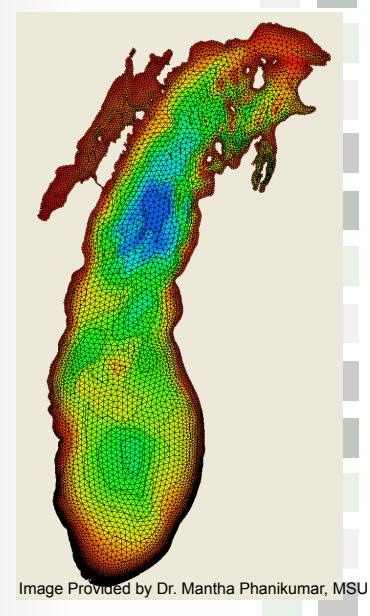
- Boundary Simulations
- Data Analysis
- Search

MICHIGAN STATE  
UNIVERSITY



## Example: Boundary simulations

1. Divide a 2D or 3D simulation space into a grid of cells
2. Define information that is transferred at the boundary of the cells
3. Simulate the dynamics of the cell during a time interval
4. Repeat steps 2 and 3



MICHIGAN STATE  
UNIVERSITY



## Boundary Simulations

- Fluid dynamics
  - Finite element analysis
  - Molecular dynamics
  - Weather
  - Etc.
- O'Shea
- System of PDE (Partial Differential equations)
  - Mathematically equivalent to inverse of a matrix

MICHIGAN STATE  
UNIVERSITY



# Boundary Simulation

- Tightly to loosely coupled
- Typically solved with MPI
- PDE solutions available for GPU and OpenMP



MICHIGAN STATE  
UNIVERSITY

Turk, Smith, O'Shea



# Example: Data Analysis

1. Input data file
2. Find features,  
search or filter  
data in some  
way
3. Output Results

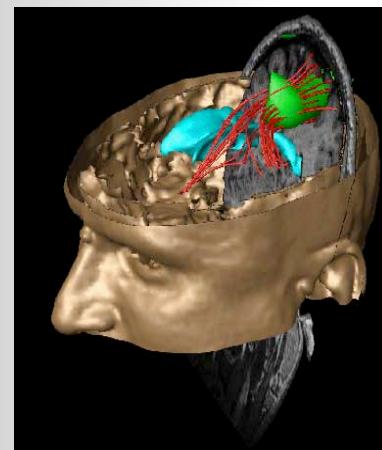


Image from OpenDX

MICHIGAN STATE  
UNIVERSITY



# Data Analysis

- Computer vision tasks
  - Bioinformatics
  - Astrophysics
  - Etc.

**MICHIGAN STATE  
UNIVERSITY**

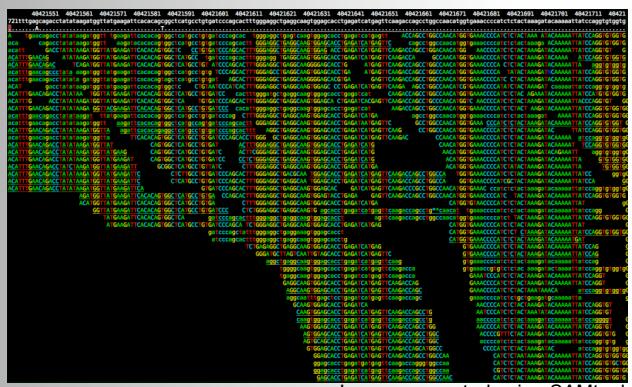


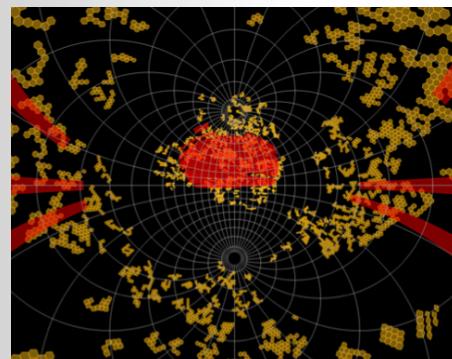
Image generated using SAMtools



# Data Analysis

- Loosely coupled
  - Bulk of computation is typically pleasantly parallel
  - Can be I/O bound

**MICHIGAN STATE  
UNIVERSITY**



## The footprints of the Sloan Digital Sky Survey's 5<sup>th</sup> Data Release and the Galaxy Evolution Explorer's 2<sup>nd</sup> Public Release



## Example: Search

- Randomly generate test candidates
- Evaluate the quality of solution
- Repeat until found

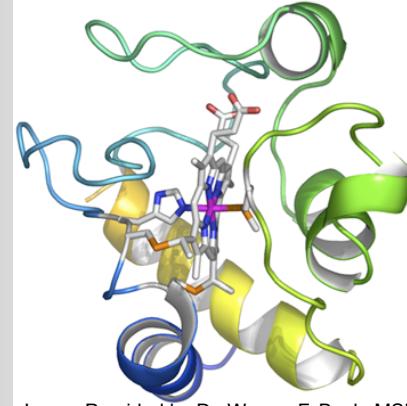


Image Provided by Dr. Warren F. Beck, MSU



## Search

- Evolution (Avida)
- Genetic Algorithms
- RANSAC
- Monte Carlo
- Etc.

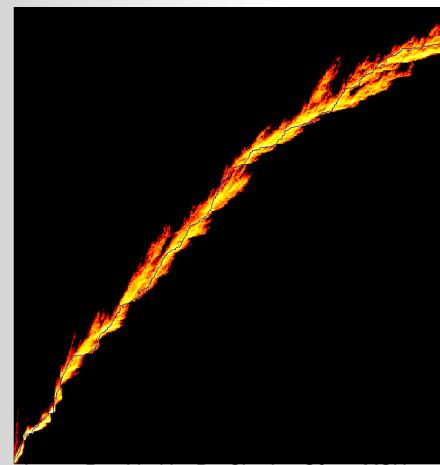
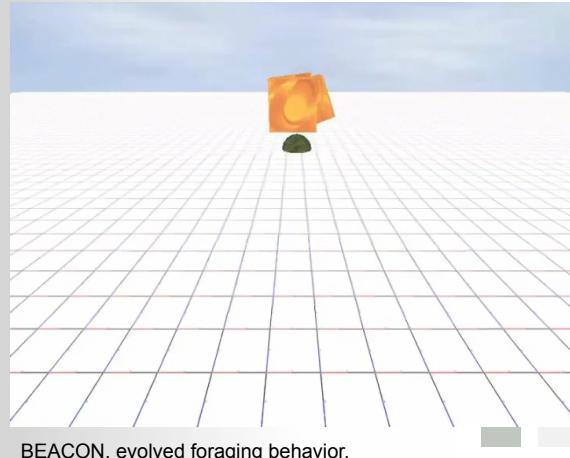


Image Provided by Dr. Charles Ofreia, MSU



# Search

- Pleasantly parallel
- The more the better
- Typically not I/O bound
- Typically not memory bound



BEACON, evolved foraging behavior,  
digital evolution study

MICHIGAN STATE  
UNIVERSITY

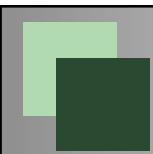


# Summary

- Advanced computational hardware can help you do more science faster
- Advanced computing typically requires knowledge of a primitive interface (command line)
- When parallelizing your computation, think about where the loops are, what needs to be communicated, and where there are bottlenecks
- There are many existing computational resources that can help you get started

MICHIGAN STATE  
UNIVERSITY





# Questions?

iCER Website:  
<http://icer.msu.edu>

User Information on HPCC:  
<http://wiki.hpcc.msu.edu/>



MICHIGAN STATE  
UNIVERSITY