

# Introduction: NGS Hands-On Workshop

John Fonner and Matt Vaughn  
Computational Biology  
Texas Advanced Computing Center  
The University of Texas at Austin

# Agenda

1:30 – 1:45 Introduction to TACC  
1:45 – 2:30 Break  
2:30 – 5:30 User Environment  
Q/A

Lunch

Introduction to Linux/Unix  
Q/A  
Hands-on Lab

# Agenda

**1:30–1:45** – Introduction to TACC

**1:45–2:30** – Lecture: Now Gen Sequencing

**2:30–5:30** – Hands-on Tutorials and Discussion

- Should have already created a TACC account
- SSH client required if you brought your own laptop

# What is TACC?

- Research center within the University of Texas
- ~100 computational experts and researchers
- Manages computational resources for the scientific community and provides them **free of charge** to investigators on a project basis
- Conducts research on a variety of topics
- Funded by NSF, UT, other federal grants, and traditional PI research grants



# High Performance Computing

## Ranger

- 62,976 cores, 2.0 GHz AMD Opteron
- 504 Teraflops



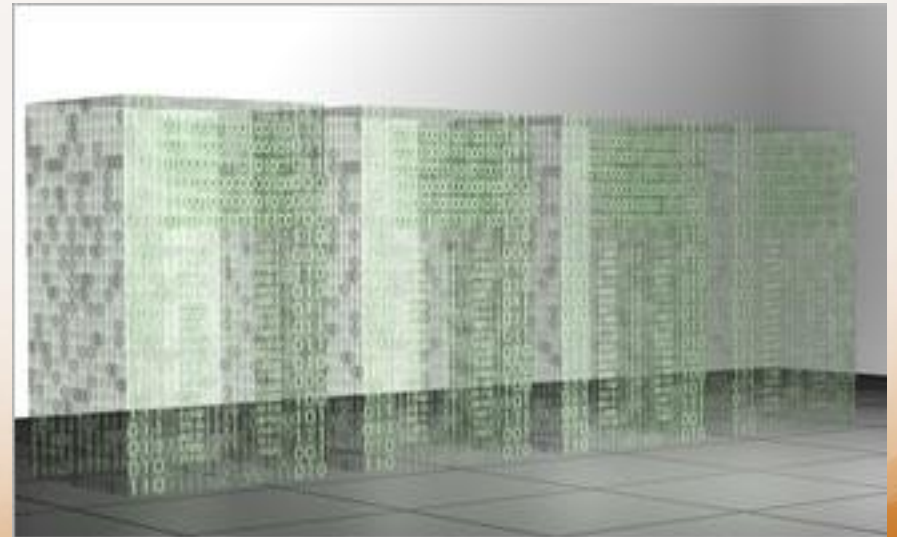
## Lonestar 4

- 22,656 cores, 3.33 GHz Intel “Westmere”
- 302 Teraflops
- GPU and large memory nodes available

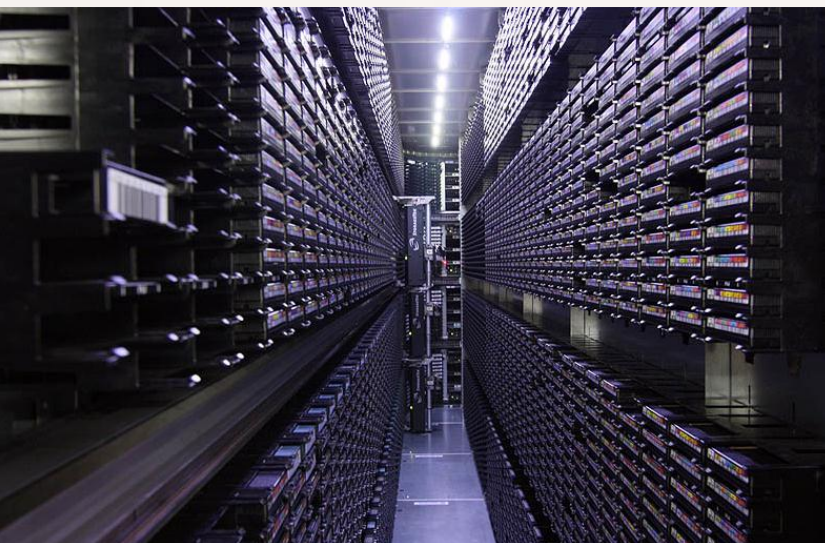
# On the Horizon

## Stampede

- Deployed in 2013
- Intel Xeon E5 processors (not released yet)
- Intel Many Integrated Cores (MIC) co-processors
- 10 petaflops peak performance
  - 2 petaflops from the Xeon processors
  - 8 petaflops from the MIC co-processors
- GPU and large memory nodes will be available



# Data Storage



## Ranch

- 70 petabyte capacity
- Used for long-term storage
- Access provided to users of other TACC resources

## Corral

- 1.2 petabyte capacity
- Designed for hosting data collections



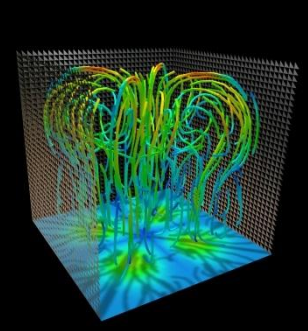
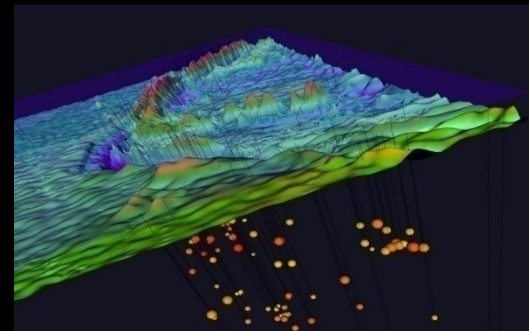
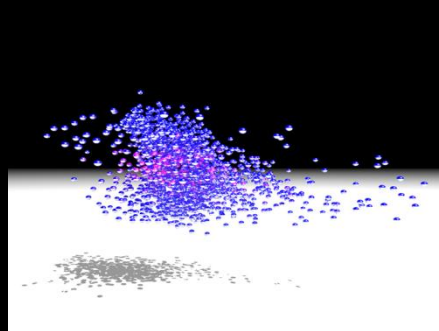
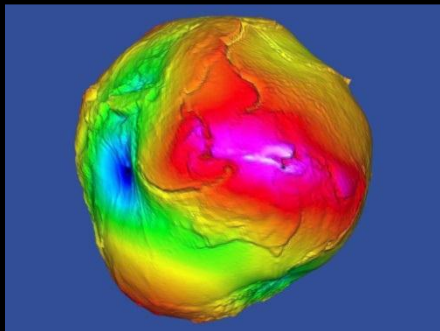
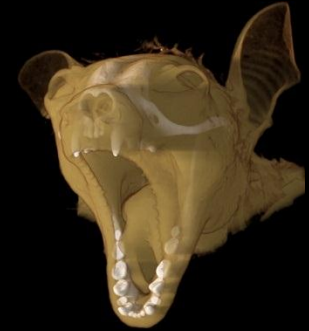
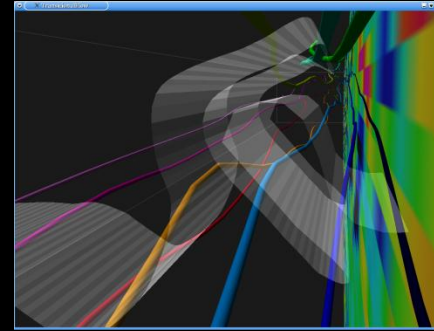
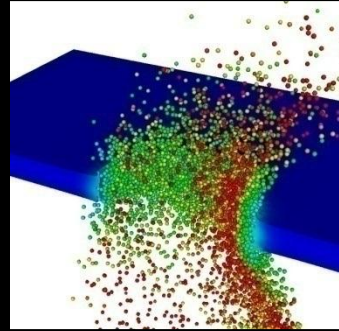
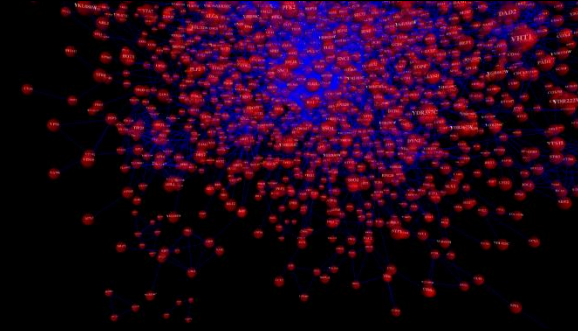
# On the Horizon

## Research Data Repository

- Complete in December, 2011
- Multi-Petabyte, fully replicated, high-performance storage
- Will support both web and file system access
- 5TB should be available to all UT System researchers on request
- Larger quantities will eventually require some review and/or financial contribution



# Visualization and Rendering



# TACC Computational Biology Group

**Goal: Establish TACC as a leading center for ENABLING computational biology research.**

- Dynamic group of 9 scientists with complimentary backgrounds in structural biology, genomics, mathematics, cell biology, computer science
- Biological research represents ~25% of all computation on TACC systems
- Foster collaborative efforts with investigators wanting to develop/implement HPC for their work

# Key Bio Projects and Collaborations

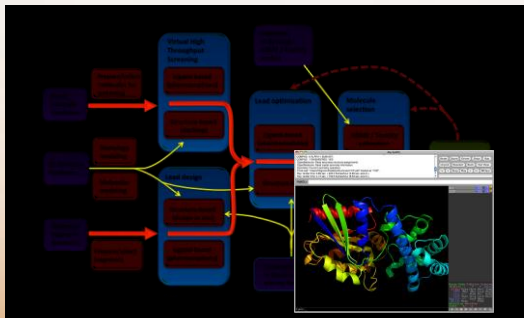
- Team members serve as PI/co-PI on several large externally funded projects totaling over \$52M
- **iPlant**: \$50M NSF project to build CI for “Grand Challenges” in the Plant Science Community
- **Epigenetic profiling**: two projects (~\$1.3M) seeks to understand the biological basis for how traits are passed on through imprinting and developmental signaling
- **Virtual Drug Screening**: projects aims to allow reserachers to identify new, effective drug candidates by using computers to simulate and predict ligand (drug) and receptor (protein) binding interactions.

# Computational Biology at TACC



Support core set of software tools

Provide training – in person and online



Collaborate to better leverage advanced computing for biology

# Comp Bio Software Stack

- TACC maintains and supports a core set of biology related applications that:
  - Have high impact and broad appeal to Life Science research
  - Are computationally intensive and/or parallelized
  - **Are highly requested by the user community**

# Installed Bio Software Stack

## Molecular Dynamics & Structural Biology

---

NAMD  
Amber  
GROMACS  
Desmond  
VASP  
LAMMPS  
APBS  
NWChem  
GAMESS  
AutoDock  
Vina

## Genomics & Bioinformatics

---

BLAST+	Velvet
mpiBLAST	Abyss
HMMER	Trinity
MAFFT	BWA
MUSCLE	Ssake
R	SOAP de novo
BioPerl	AMOS
FASTX-Toolkit	Maq
Picard	Bowtie
SAMtools	Cufflinks
SHRiMP	TopHat

# How To Get Time on TACC Systems?

- NO COST for using TACC systems for academic research
- Time on TACC systems is acquired by submitting an allocation request:

**<https://portal.tacc.utexas.edu>**

**<https://www.xsede.org>**

- Open to all UT academic researchers through the TACC Portal allocations process
  - Up to 500,000 SUs on Ranger and Lonestar
  - Above 500K, apply through XSEDE
- Resources are allocated on a project basis to a single principal investigator (PI)

# TACC User Portal

Easy to get started! Easy to find information

The screenshot displays the TACC User Portal interface. At the top left is the TACC logo and the text 'TACC USER PORTAL'. On the top right, there is a 'Login' link and a 'Welcome, Guest' message. Below the logo is a horizontal navigation menu with buttons for 'Home', 'News', 'Resources', 'Allocations', 'Documentation', 'Training', 'Consulting', and 'About'. A secondary navigation bar contains links for 'Welcome', 'New User?', 'Forgot Password?', 'Account Migration', and 'Migration FAQ'. The main content area is divided into two columns. The left column, titled 'Welcome', features a 'New TACC user?' section with a link to request an account, a login form with 'Username:' and 'Password:' fields, a 'Keep me logged in' checkbox, and a 'Login' button. Below the form is a 'Forgot your password?' link and a paragraph of introductory text. The right column, titled 'User News', lists several announcements with dates and times, including 'Fifth Annual Scientific Software Days', 'Longhorn System Maintenance 5/10', 'Ranch System Maintenance 5/3', 'TACC IT System Maintenance 4/30', and 'Lonestar System Maintenance 5/3'. A 'view all user news' link is provided at the bottom of this section. Below the news is a 'System Status' section containing a table with columns for 'Name', 'Status', 'Load', and 'Jobs'. The table lists 'Ranger' and 'Lonestar' systems with their respective status indicators and job counts.

**TACC USER PORTAL**

Home News Resources Allocations Documentation Training Consulting About

Welcome New User? Forgot Password? Account Migration Migration FAQ

**Welcome**

**New TACC user?**  
Click here to request a TACC account.

Username:

Password:

Keep me logged in

[Forgot your password?](#)

While continuing to provide access to user news, training, consulting, documentation, and information on TACC's computational, visualization, and storage resources, the TACC User Portal is your interface to a user management system at TACC designed to provide you with streamlined management of your accounts, projects, and allocations.

Features of the TACC User Portal include:

- Single account for all resources

**User News**

**(Reminder) Announcement: Fifth Annual Scientific Software Days**  
Thu, 05 May 2011 17:13:55 CDT

**Longhorn System Maintenance 5/10**  
Tue, 03 May 2011 12:44:22 CDT

**Ranch System Maintenance 5/3**  
Thu, 28 Apr 2011 16:26:35 CDT

**TACC IT System Maintenance 4/30**  
Tue, 26 Apr 2011 15:48:39 CDT

**Lonestar System Maintenance 5/3**  
Tue, 26 Apr 2011 12:27:56 CDT

[view all user news](#)

**System Status**

Name	Status	Load	Jobs
Ranger	↑	<div style="width: 100%;"></div>	328R 451Q 164O
Lonestar	↑	<div style="width: 100%;"></div>	250R 279Q 85O

# Questions?

[jfonner@tacc.utexas.edu](mailto:jfonner@tacc.utexas.edu)