WELCOME TO INTRODUCTION TO SCIENTIFIC PROGRAMING

KEVIN SCHMIDT, SUSAN LINDSEY, and CHARLIE DEY

Us

Kevin Schmidt
Research Associate
HPC Software Tools Group

kschmidt@tacc.utexas.edu

Kevin joined TACC in 2018 as a Research Associate in the High Performance Computing Group. Prior to his position at TACC, Kevin studied materials design using electronic structure theory, taught several engineering courses at the University of Nevada-Reno (UNR), and was the HPC Administrator for the engineering compute cluster at UNR.

Us

Susan Lindsey
Technical Information Coordinator
User Services

slindsey@tacc.utexas.edu

Susan is TACC's Technical Information Coordinator. As such, she is responsible for gathering and presenting timely, consistent and accurate technical information to the supercomputing user community. She will also be evaluating and implementing emerging technologies related to the dissemination and presentation of technical content.

Susan comes to TACC after eight years at the San Diego Supercomputer Center where she researched and programmed on a variety of computational biology projects.

Us

Charlie Dey
Director, Training And Professional Development
User Services

charlie@tacc.utexas.edu

Charlie is the Director of Training and Professional Development with the User Services group at TACC with a background in web development and scientific computing. Charlie's responsibilities at TACC include organizing, developing content, and building curriculums for TACC's academic course selection taught in conjunction with several departments at the University of Texas at Austin, as well as for TACC's professional development and educational training.

Prior to joining TACC, he worked as a Senior Application Developer for the Carle Foundation, and as a computer science instructor at Parkland College in Champaign, IL. He was also a member of a specialized application development team at the University of Illinois and has also been a contracted research consultant for NASA Ames Research Center, studying computational immunology and bioinformatics.

TACC AT A GLANCE

- Research center located at UT Austin
- ~160 Staff (~70 PhD scientists, ~20 students)
- Funded by UT System, NSF (85% external grants)

- **Users:** >10,000 on 2,300 active projects across all fields
- **Partnerships:** UT Research Cyberinfrastructure (UTRC), Extreme Science and Engineering Discovery Environment (XSEDE), Industry, International

Mission: "To enable discoveries that advance science and society through the application of advanced computing technologies."

TACC AT A GLANCE

- **Capacity and Infrastructure:**
  - A billion compute hours per year
  - 5 billion files, 50 petabytes of data
  - Hundreds of public datasets
  - 10 MW data center

- **Systems and Services:**
  - High performance computing (HPC), high throughput computing (HTC), large scale data storage, cloud computing, visualization
  - Portals and gateways, web service APIs, rich software stacks
  - Consulting, curation and analysis, code optimization, training and outreach

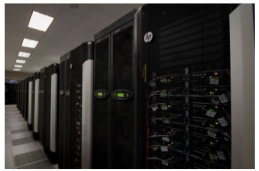
TACC SYSTEMS AT A GLANCE

STAMPEDE

HPC, Visualization, Data Analysis, Data Intensive Computing

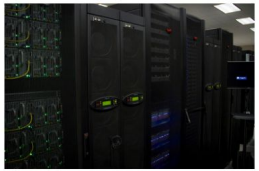
LONESTAR 5

HPC, Remote Visualization



MAVERICK

Interactive Visualization, Data Analytics



WRANGLER

Data Analysis, Data Management

CHAMELEON

Cloud Computing Testbed

HIKARI

Sustainable Supercomputing

VISLAB

Advanced Visualization Resources and Consulting

STALLION

Tiled-Display System

LASSO

Multi-Touch Display

JETSTREAM

Self-Service Cloud System

FABRIC

Alternate Computer Architectures

DISCOVERY

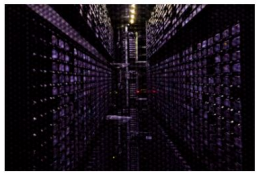
Testbed Cluster

RODEO

Cloud Computing, Storage

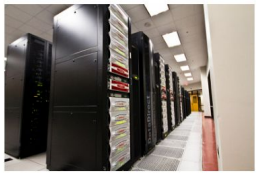
CORRAL

Storage, Data Management



RANCH

Mass Archival Storage



STOCKYARD

Global File System

CATAPULT

A Reconfigurable Architecture for Large Scale Machine Learning

RUSTLER

Data Intensive Computing

HIGH PERFORMANCE COMPUTE CLUSTERS

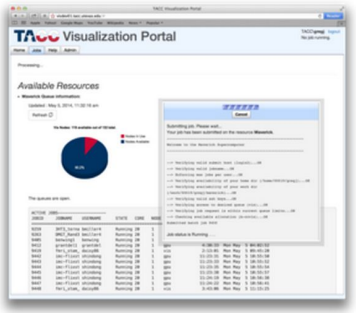
- **Stampede2**
  - 4200 KNL nodes, 1736 Intel Xeon nodes; ~18 PF peak performance
  - >600 TB RAM; 21 PB lustre filesystem
  - XSEDE / UT System

- **Lonestar5**
  - 1,282 nodes; 30K compute cores; ~1.2 PF peak performance
  - Large mem, GPU, hyperthreading
  - **UT System only**

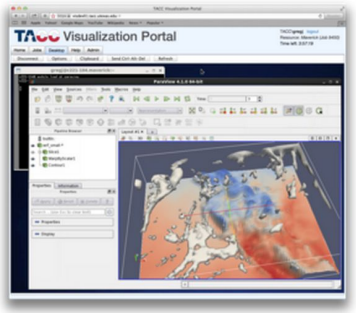
VISUALIZATION AND GPGPU

- **Maverick**
  - 132 NVIDIA Tesla K40s
  - Remote visualization via VNC
  - GPGPU calculations
  - Machine learning / deep learning

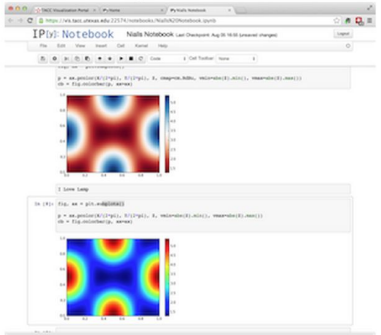
Job Submission



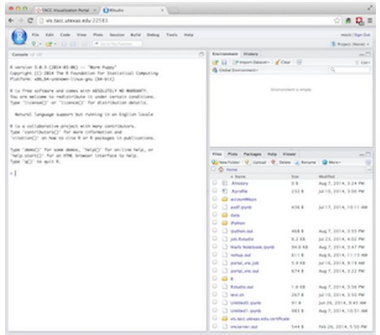
VNC Visualization Session



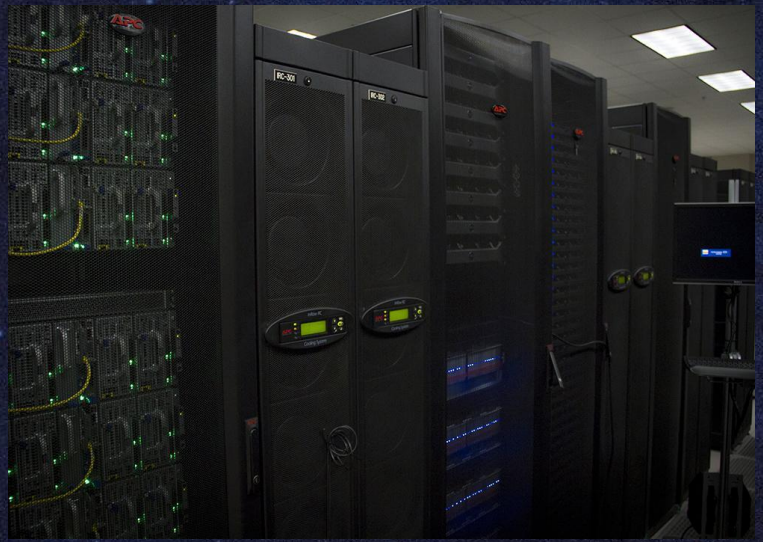
iPython Notebook



RStudio



DATA INTENSIVE COMPUTING



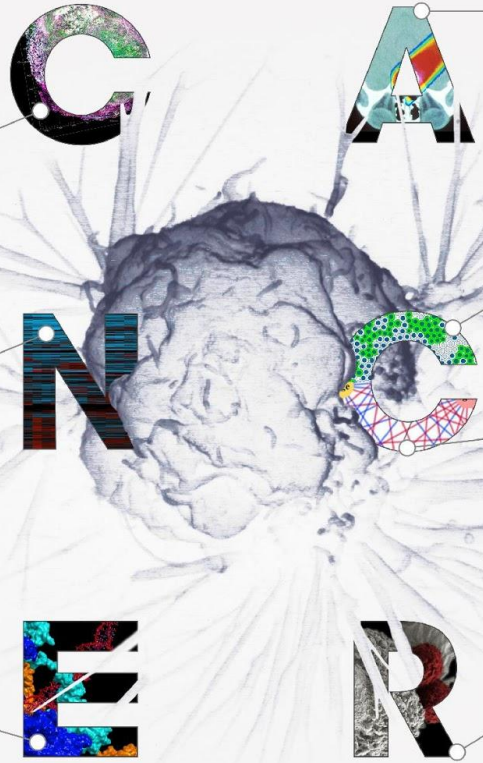
- **Wrangler**
  - >3,000 processor cores for analytics
  - 10 PB storage system; 600 TB DSSD flash storage
  - Aggregate bandwidth >1 TB/s
  - **Only allocable through XSEDE at this time**

DATA AND COLLECTIONS REPOSITORY

- **Corral**
  - 11 PB geo-replicated storage (Austin / Arlington)
  - 5 TB free to all UT System principal investigators
  - $118 per TB per year after the first 5 TB
  - Project data sharing / collection hosting

The Texas Advanced Computing Center accelerates basic and applied cancer research to help save lives.

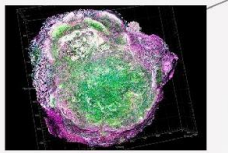
fighting



with supercomputers

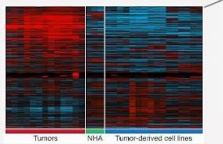
Computer Modeling

Researchers use advanced computing to model tissues, cells and drug interactions, and to design patient-specific treatments and identify new medicines.



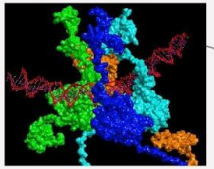
Big Data Analysis

Supercomputers allow researchers to find patterns in genomes and among patient outcomes to pinpoint risks and target treatments.



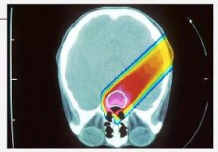
Molecular Dynamics Simulations

Simulating protein and drug interactions at the atomic level enables scientists to understand cancer and design more effective therapies.



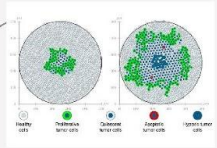
Quantum Calculations

Exploring how proton and x-ray beams interact with DNA on the quantum level helps explain why radiation treatments work and how they can be optimized.



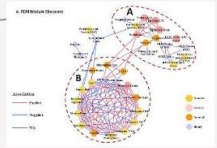
Trial Design

Researchers use TACC's advanced computers to design clinical trials that can determine the combination of dosages that will be most effective.



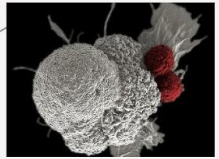
Clinical Planning

Supercomputers can test thousands of potential treatments in advance to help decide which one will work best.



Artificial Intelligence

AI on high-performance computers can uncover relationships among complex cellular networks and reverse-engineer interventions.

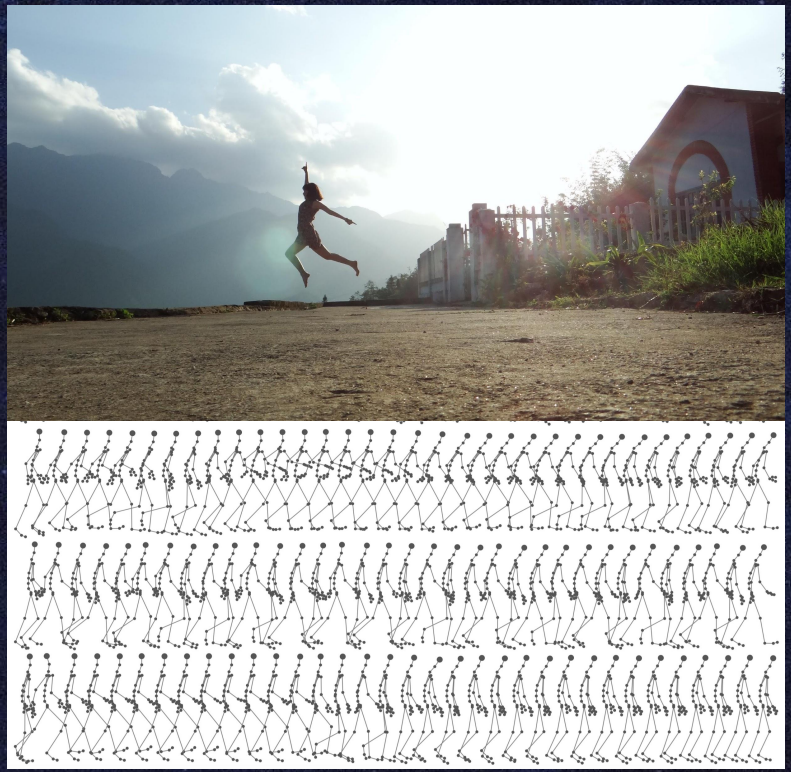


TACC
TEXAS ADVANCED COMPUTING CENTER

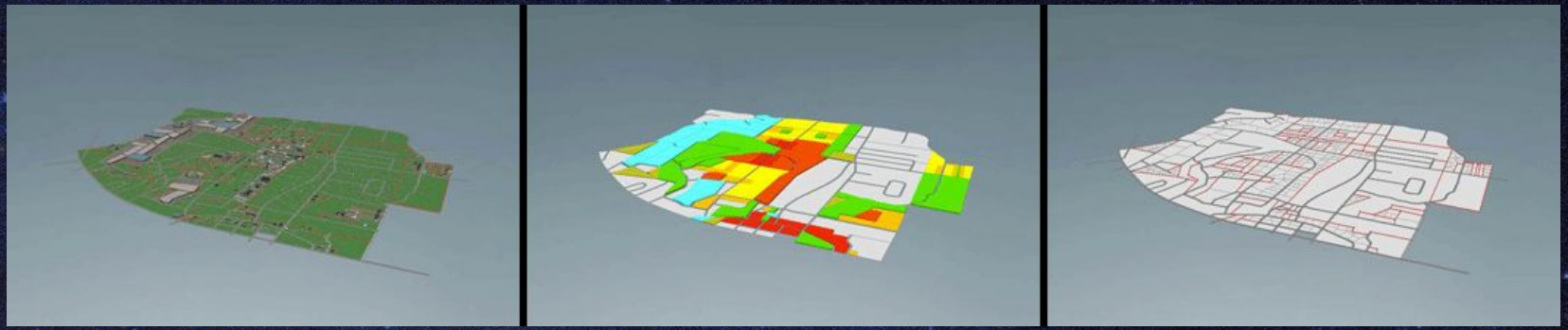
ADVANCED COMPUTING FOR SOCIAL CHANGE

- Use visualization and data analytics to increase awareness
- Engage students in a social change challenge
- Increase the participation of students historically underserved in STEM

A DANCE WITH ALGORITHMS



- XSEDE resources help researchers merge art and science to create systems that can understand and produce human-quality movement
- Uses deep learning techniques and relies on the collaboration of graduate students with artists to create new algorithms
- *On a single computer, running our algorithms would take years, on medium-sized resources months, but using XSEDE, we can train some of most complex models within 24 hours.* - Philippe Pasquier, professor and researcher, Simon Fraser University



- How will the Central Texas region evolve over the next 20 years?
- Newly developed suite of analytics tools developed by TACC for understanding of the impacts of various development patterns, such as:
  - What types of housing and business development should be planned?
  - How concentrating growth can maximize a community's infrastructure?
  - Where natural resources should be preserved?
  - How can communities promote better health?
  - How the region can ensure all segments of the community have access to education and jobs?

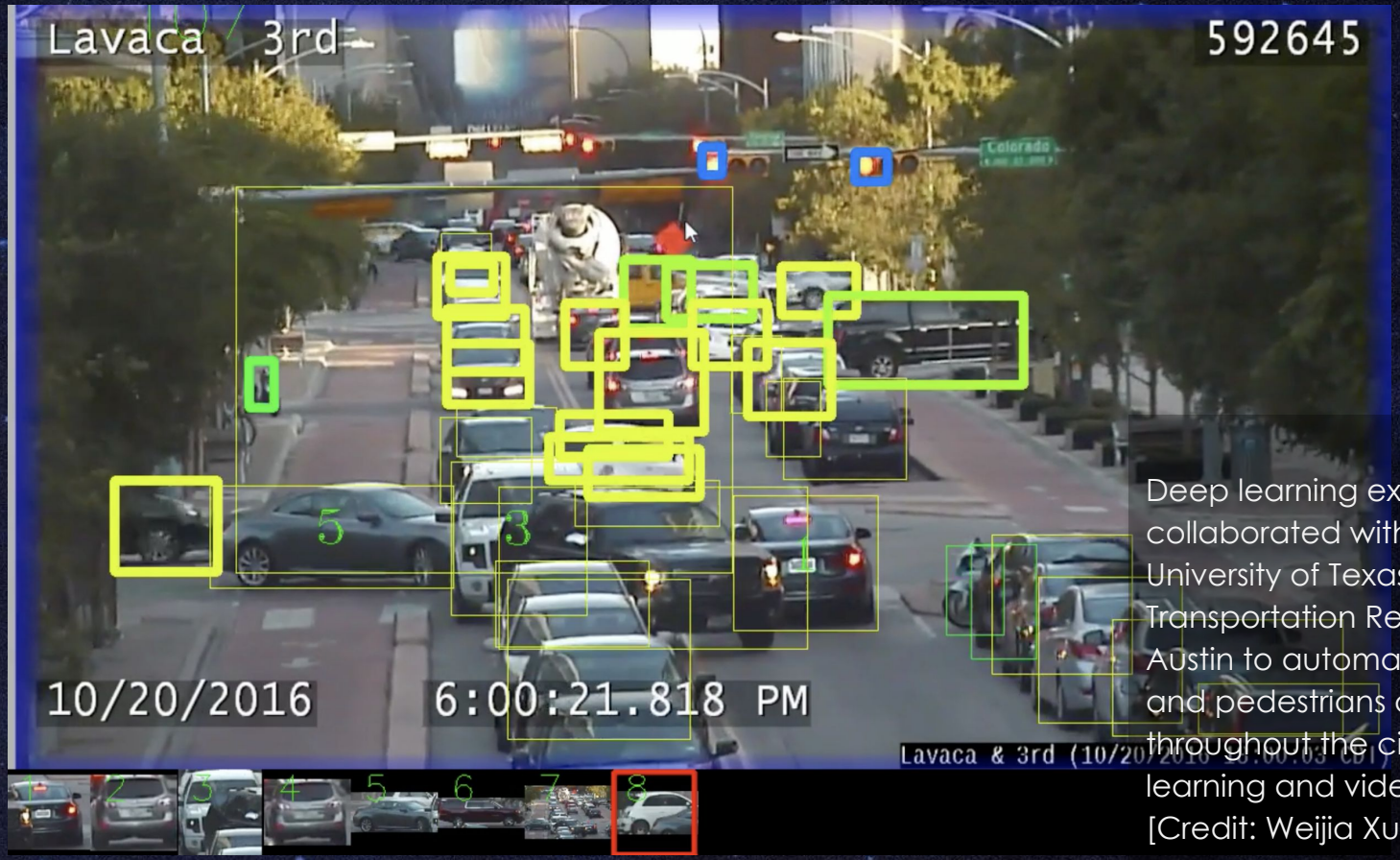
CITY AND REGIONAL PLANNING IN 3D

TAXATION AT THE TOP

- Economist use Stampede supercomputer to assess consequences of increasing tax rate on top earners
- "We need (TACC) to efficiently compute optimal individual decision rules about working, learning and saving, and then simulate the behavior of millions of model households, using these rules, in order to determine economic aggregates and equilibrium behavior,"
- Badel and Huggett found that a spike in the tax rate would have adverse affects on potential high earners.

USING SUPERCOMPUTERS TO ILLUMINATE THE RENAISSANCE

- Who knew whom in Renaissance Britain?
- Analyze writers work using machine learning, graph inferences, and web development to reconstruct and communicate social networks of Britain from about 1500 to 1700.
- *Once you employ computational techniques you can start to assemble relationships at a much greater scale. This is something no human could ever have in their head. By putting this together and making it available for the scholarly community we hope that we're facilitating a new way of doing scholarship that allows for a full appreciation of these historical networks.* - Christopher Warren, Associate Professor of English, Carnegie Mellon University



Deep learning experts from TACC collaborated with researchers at the University of Texas Center for Transportation Research and the City of Austin to automatically detect vehicles and pedestrians at critical intersections throughout the city using machine learning and video image analysis. [Credit: Weijia Xu, TACC]

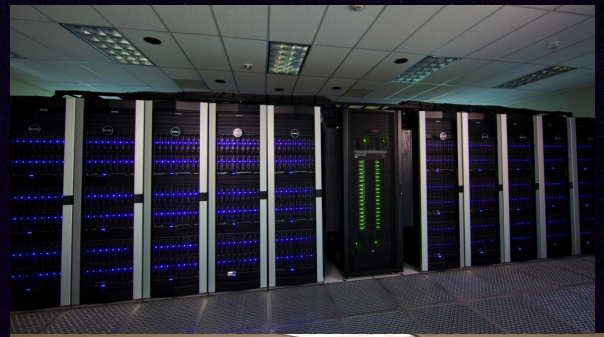
"COMPUTING FOR THE ENDLESS FRONTIER"

Dan Stanzione
Executive Director, Texas Advanced Computing Center
Associate Vice President for Research, The University of Texas at Austin


FRONTERA
TACC
NSF
TEXAS

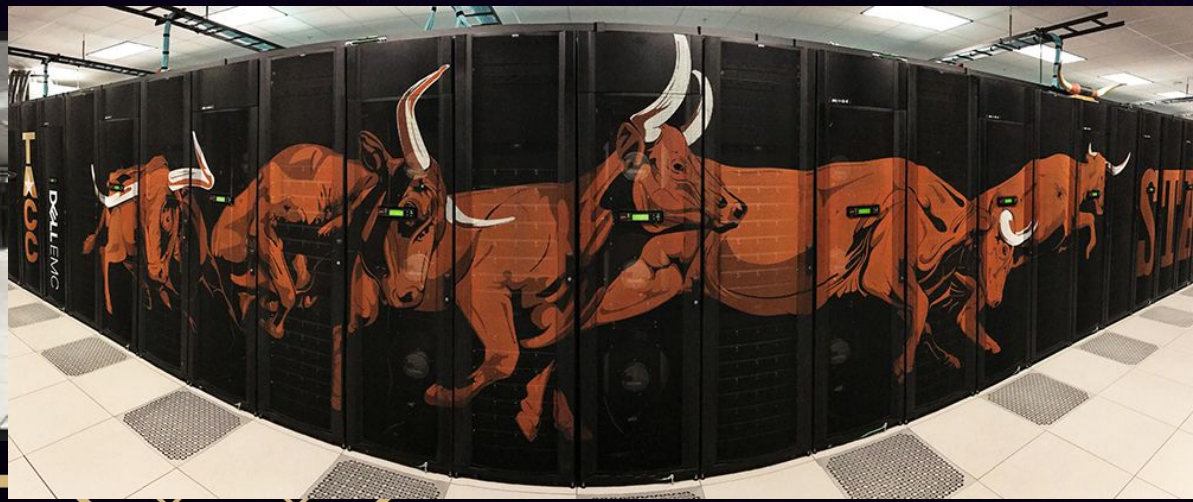
RAPID GROWTH FROM THEN TO NOW…

- 2003 – First Terascale Linux cluster for open science (#26)
- 2004 – NSF funding to join the Teragrid
- 2006 – UT System Partnership to provide Lonestar-3 (#12)
- **2007 - $59M NSF award – largest in UT history – to deploy Ranger, the world's largest open system (#4)**
- 2008 – funding for new Vis software and launch of revamped visualization lab.
- 2009 - $50M iPlant Collaborative award (largest NSF bioinformatics award) moves a major component to TACC, life sciences group launched.
  - In 2009, we reached, 65 employees.



NOW, A WORLD LEADER IN CYBERINFRASTRUCTURE

- 2010, TACC becomes a core partner (1 of 4) in XSEDE, the TeraGrid Replacement
- 2012, Stampede replaces Ranger with new $51.5M NSF Award
- 2013, iPlant is renewed, expanded to $100M
- 2015, Wrangler, first data intensive supercomputer is deployed.
- 2015, Chameleon cloud is launched
- 2015, DesignSafe, the cyberinfrastructure for natural hazard engineering, is launched.
- 2016 Stampede-2 awarded the largest academic system in the United States, 2017-2021.



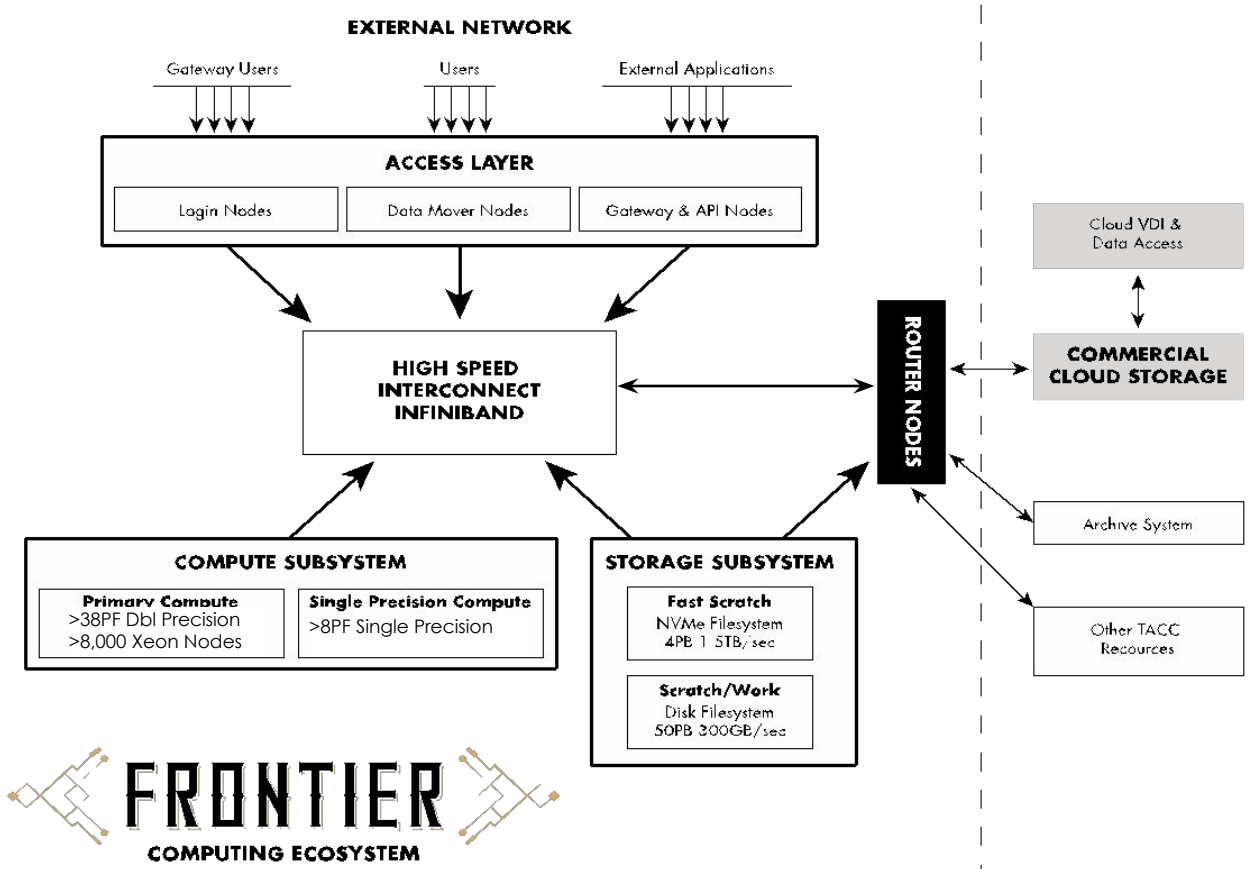
FRONTERA SYSTEM --- PROJECT

- Deploy a system in 2019 for the largest problems scientists and engineers currently face.
- Support and operate this system for 5 years.
- Plan a potential phase 2 system, with 10x the capabilities, for the future challenges scientists will face.

FRONTERA SYSTEM --- HARDWARE

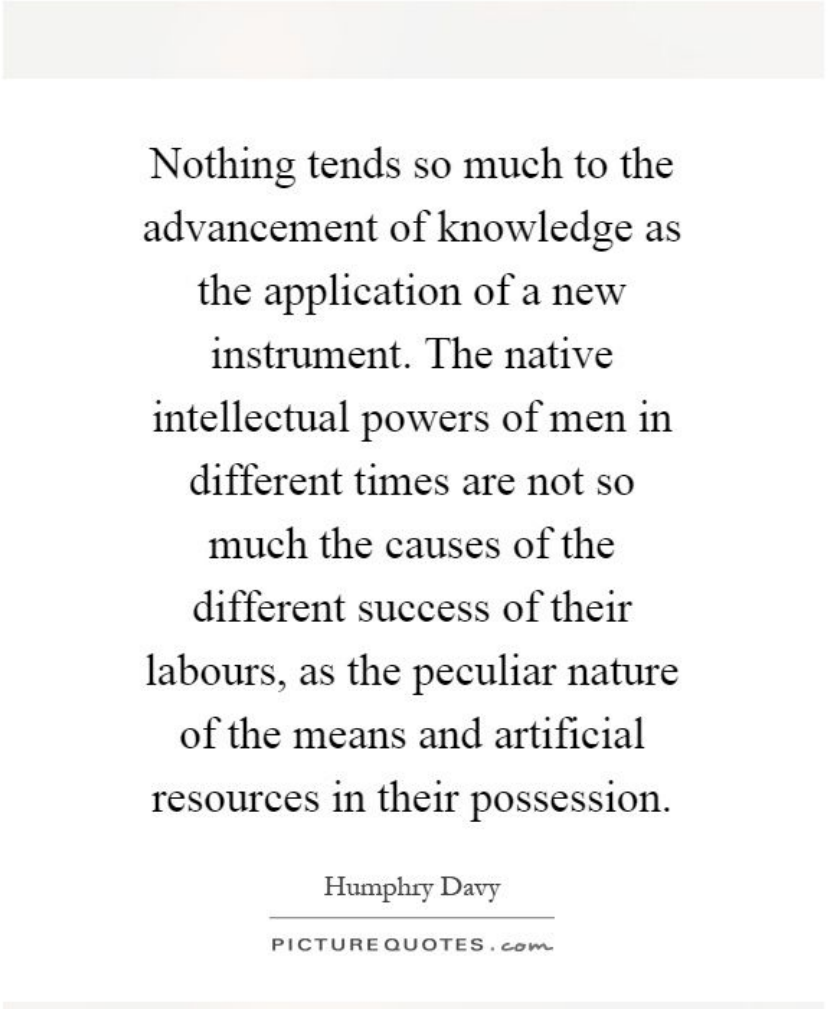
- Primary compute system: DellEMC and Intel
  - 35-40 PetaFlops Peak Performance (Next Generation Xeon processors)
- Interconnect: Mellanox HDR and HDR-100 links.
  - Fat Tree topology, 200Gb/s links between switches.
- Storage: DataDirect Networks
  - 50+ PB disk, 3PB of Flash, 1.5TB/sec peak I/O rate.
- Single Precision Compute Subsystem: Nvidia
- Front end for data movers, workflow, API

SYSTEM OVERVIEW



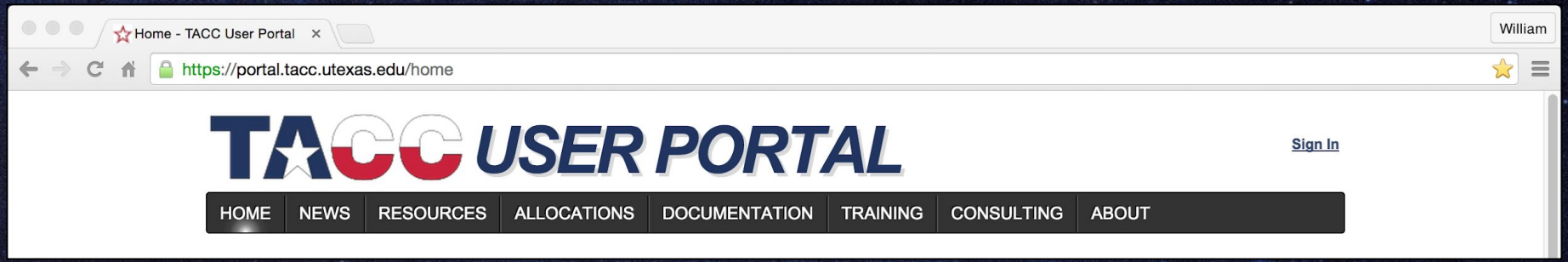
FRONTERA

- **Humphry Davy, Inventor of Electrochemistry, 1812**
- (Pretty sure he was talking about our machine).



GETTING STARTED AT TACC

THE TACC USER PORTAL



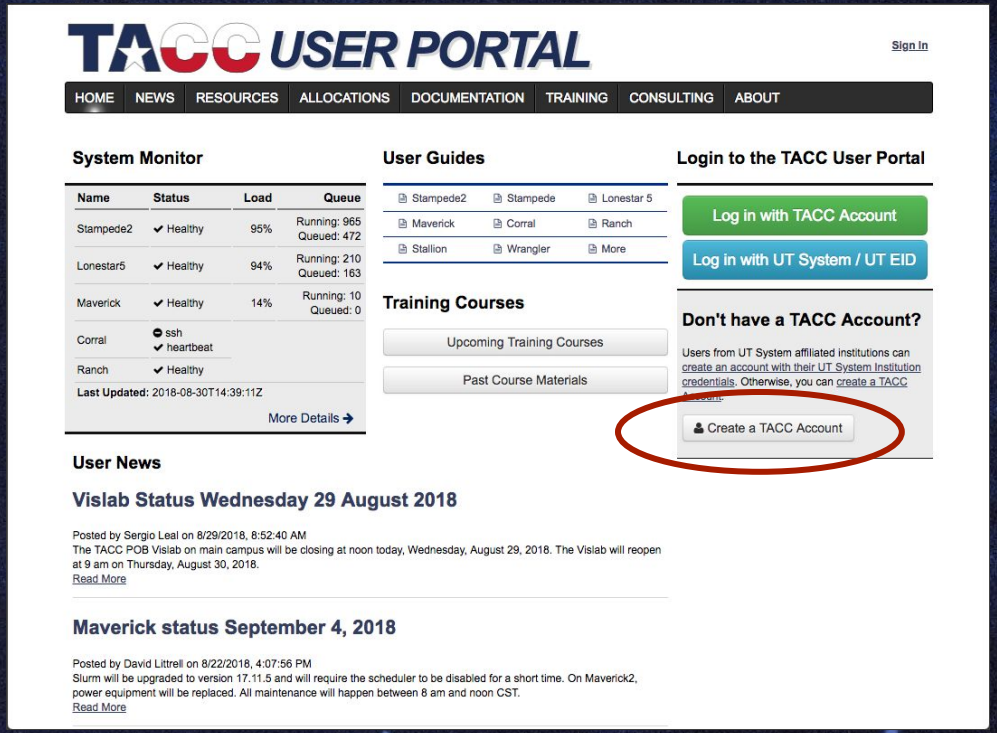
- Create accounts
- Request allocations
- Extensive system documentation
- Online training and workshops
- Software installation – availability / help
- File tickets for assistance

https://portal.tacc.utexas.edu

PROJECTS AND ALLOCATIONS

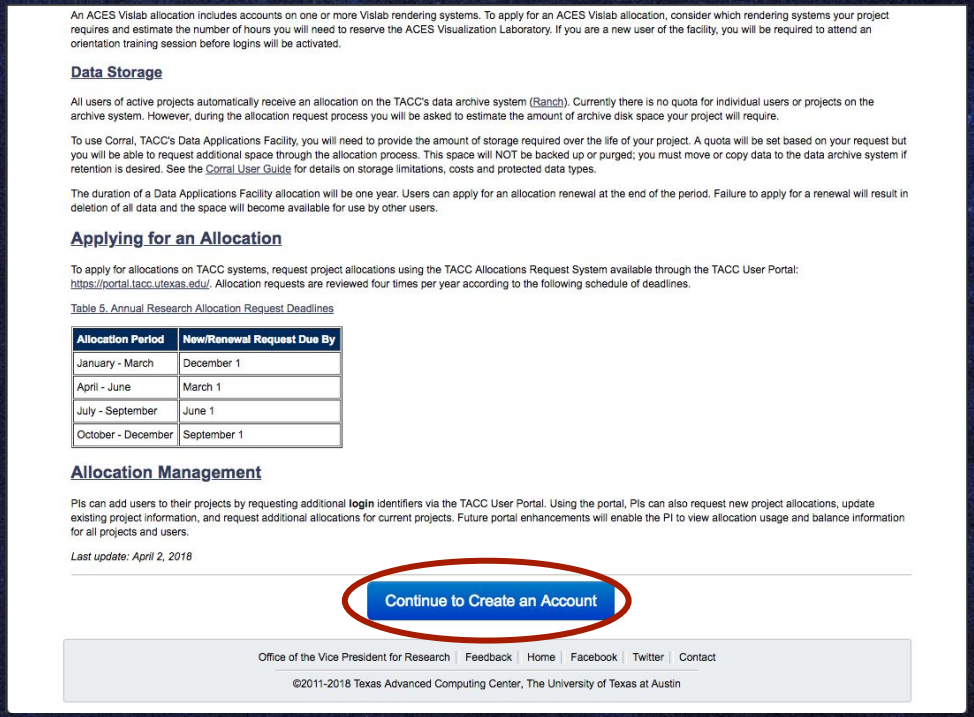
- **Project:** A group of faculty, students, and / or staff working on a common, specific research goal. Each project has one designated PI.
- **Allocation:** A specific allotment of resources (including CPU hours and / or storage space) to be shared by all users associated with a project. Each project may have multiple allocations.
- **All PIs must go through Project & Allocation system to start using TACC resources.**

SIGNING UP FOR AN ACCOUNT



- Click "Create a TACC Account"
- DO NOT CLICK THE BLUE BUTTON!

SIGNING UP FOR AN ACCOUNT



An ACES Vislab allocation includes accounts on one or more Vislab rendering systems. To apply for an ACES Vislab allocation, consider which rendering systems your project requires and estimate the number of hours you will need to reserve the ACES Visualization Laboratory. If you are a new user of the facility, you will be required to attend an orientation training session before logins will be activated.

Data Storage

All users of active projects automatically receive an allocation on the TACC's data archive system (Ranch). Currently there is no quota for individual users or projects on the archive system. However, during the allocation request process you will be asked to estimate the amount of archive disk space your project will require.

To use Corral, TACC's Data Applications Facility, you will need to provide the amount of storage required over the life of your project. A quota will be set based on your request but you will be able to request additional space through the allocation process. This space will NOT be backed up or purged; you must move or copy data to the data archive system if retention is desired. See the Corral User Guide for details on storage limitations, costs and protected data types.

The duration of a Data Applications Facility allocation will be one year. Users can apply for an allocation renewal at the end of the period. Failure to apply for a renewal will result in deletion of all data and the space will become available for use by other users.

Applying for an Allocation

To apply for allocations on TACC systems, request project allocations using the TACC Allocations Request System available through the TACC User Portal: https://portal.tacc.utexas.edu/. Allocation requests are reviewed four times per year according to the following schedule of deadlines.

Table 5. Annual Research Allocation Request Deadlines

Allocation Period	New/Renewal Request Due By
January - March	December 1
April - June	March 1
July - September	June 1
October - December	September 1

Allocation Management

PIs can add users to their projects by requesting additional **login** identifiers via the TACC User Portal. Using the portal, PIs can also request new project allocations, update existing project information, and request additional allocations for current projects. Future portal enhancements will enable the PI to view allocation usage and balance information for all projects and users.

Last update: April 2, 2018

Continue to Create an Account

Office of the Vice President for Research | Feedback | Home | Facebook | Twitter | Contact

©2011-2018 Texas Advanced Computing Center, The University of Texas at Austin

- The next page summarizes the project allocation project.
- This blue button you can click

CREATING AN ACCOUNT

page, http://www.tacc.utexas.edu.

TACC Citation

Please reference TACC in any research report, journal or publication that requires citation of any author's work. The recognition of the TACC resources you used to perform research is important for acquiring funding for the next generation hardware, support services, and our Research & Development activities in HPC, visualization, data storage, and grid infrastructure. The minimal content of a citation should include:

> Texas Advanced Computing Center (TACC)
> The University of Texas at Austin

Our suggested acknowledgement is *:

> The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing {HPC, visualization, or storage} resources that have contributed to the research results reported within this paper. URL: http://www.tacc.utexas.edu * Select one or more of the items within the braces, {}. URL: http://www.tacc.utexas.edu

* Select one or more of the items within the braces, {}.

Document Revision History

Date	Sections Affected	Modified By	Description
December 1, 2014		Nathaniel Mendoza	Yearly Review
December 1, 2015		Nathaniel Mendoza	Yearly Review
December 1, 2016		Nathaniel Mendoza	Yearly Review
April 3, 2018	Cryptocurrencies added	Nathaniel Mendoza	Yearly Review

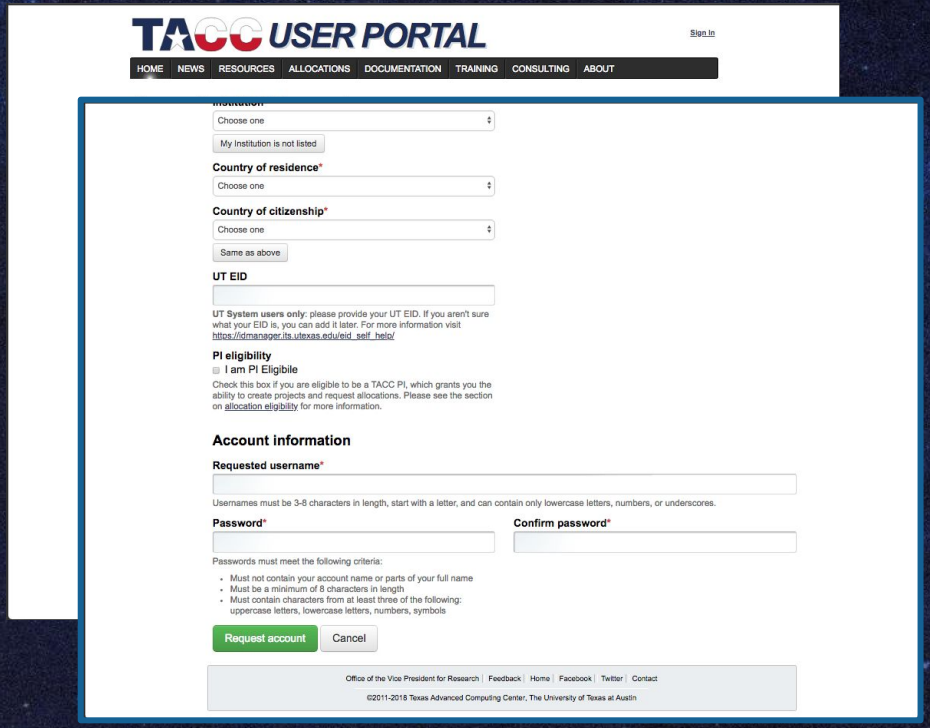
agree to the TACC Acceptable Use Policy | Cancel

Office of the Vice President for Research | Feedback | Home | Facebook | Twitter | Contact

https://portal.tacc.utexas.edu/account-request?p_p_id=createaccount_WAR_createaccountportlet&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1&_createacco...

- Please do not abuse our systems.
- Bro is watching you!

CREATING AN ACCOUNT



- Fill out the form
- Do not check the "PI Eligibility" checkbox
- Continue to setup multi factor authentication

CREATING AN ACCOUNT

- Once your account is created, email Kevin, Susan, and Charlie your username and we will add you to the class project allocation

QUESTION/COMMENTS?

- Kevin Schmidt
  - kschmidt@tacc.utexas.edu
- Susan Lindsey
  - slindsey@tacc.utexas.edu
- Charlie Dey
  - charlie@tacc.utexas.edu