ADVANCED COMPUTING IN SCIENCE & ENGINEERING

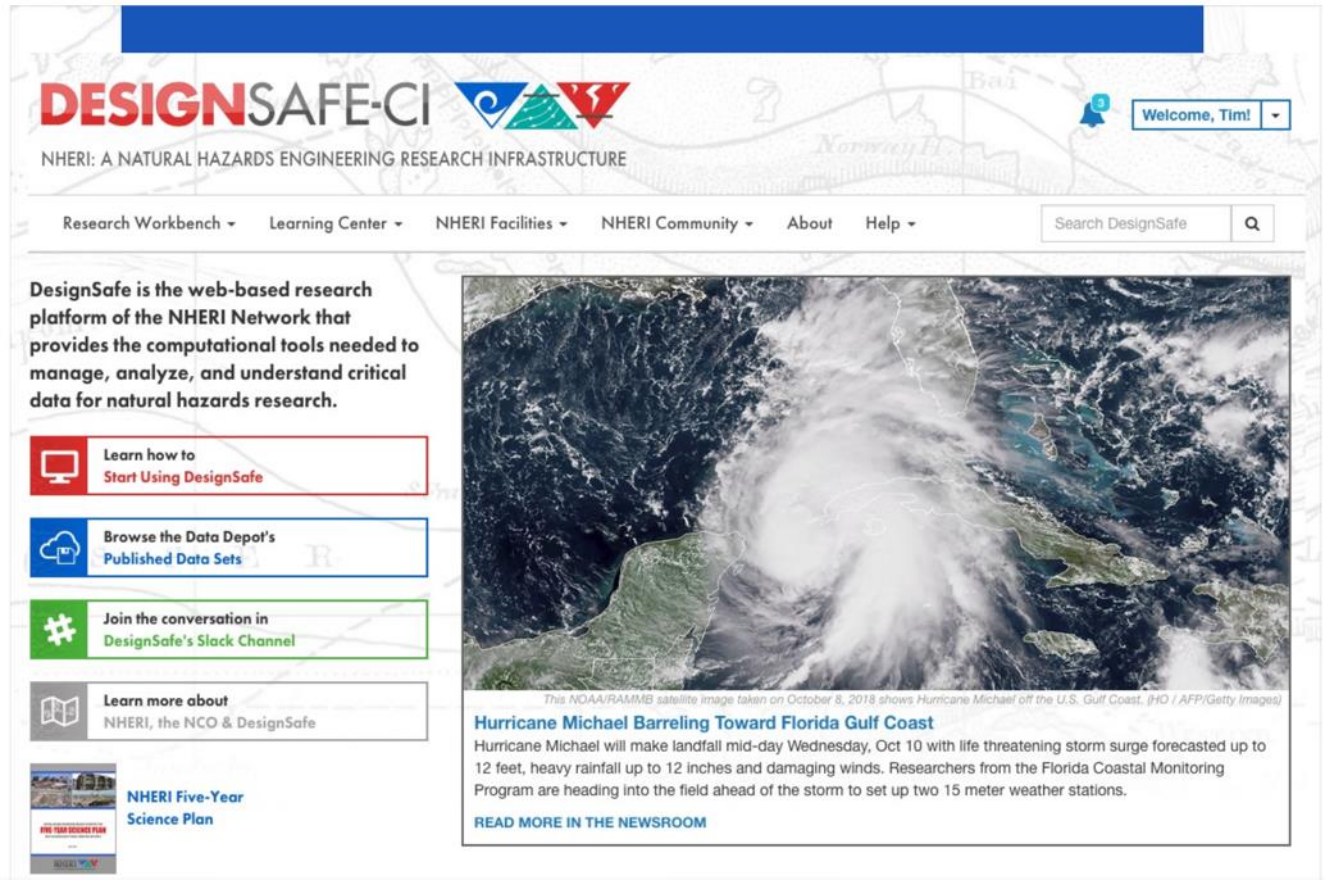
John West
Texas Advanced Computing Center
October 2019

During the 2017 hurricane season 17 named storms caused more than 100 direct fatalities, caused $265 billion in damages, and disrupted millions of lives.

Challenge:

How do you keep a natural disaster from becoming a disaster for society?

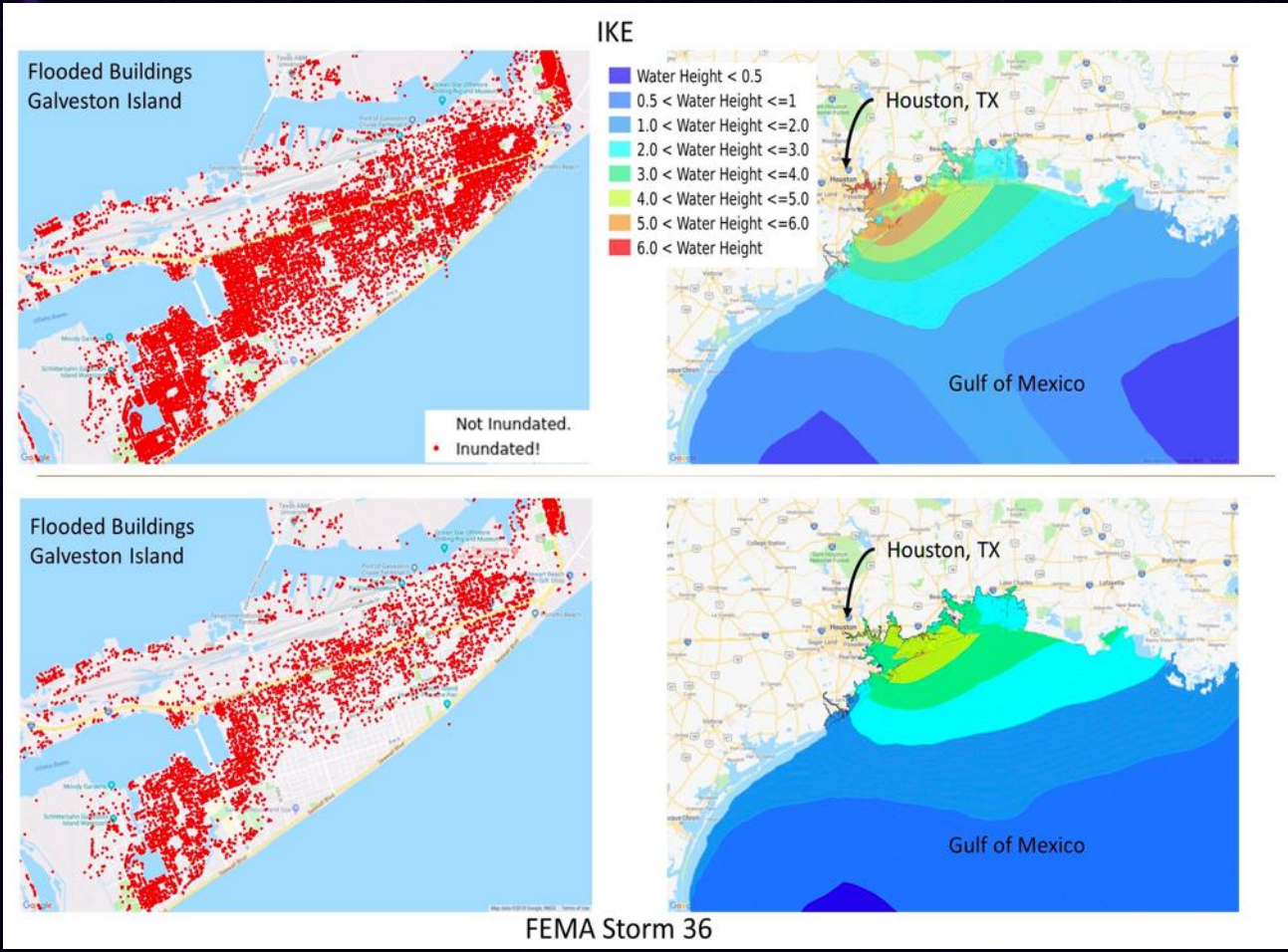
A web-based platform that provides computational tools to manage, analyze, and understand critical data for natural hazards research and response.



State emergency workers tackle Hurricane Harvey-related crises at the Texas DPS Emergency Operations Center (EOC). Credit: *Bob Daemmrich for the Texas Tribune*

"I want to know a list of addresses that have been flooded above their electrical outlet height. You have a super-computer… you can do that, right?"

When you bring together storm surge simulations, real-time weather observations, elevation maps, home constructions records, and advanced computing, you can target resources for mobilization.



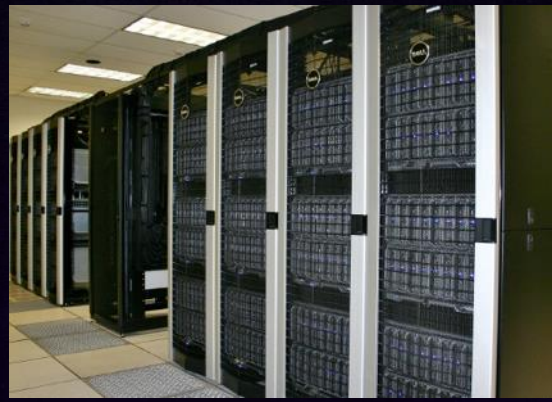
THIS IS THE MOTIVATION FOR SUPERCOMPUTING

Using advanced computers, mathematical models, sophisticated software, and specialized expertise to solve problems that cannot otherwise be solved in the real world:

- Too big
- Too small
- Too expensive
- Too dangerous

TACC LAUNCHED IN JUNE, 2001 AFTER EXTERNAL REVIEW

- Original HPC effort launched by Hans Mark as System initiative in 1986 at $30M
- In 2001, budget of $600k staff of 12 (some shared).
- 50GF computing resource (1/200,000th of the current system).



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 staff

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

- Frontera is the #5 ranked system in the world – and the fastest at any university in the world.
- Fastest primarily Intel-based system
- Highest ranked Dell system ever.
- Frontera and Stampede2 are #1 and #2 among US Universities (and Lonestar5 is still in the Top 10).

- Frontera consumes 6 Megawatts of Power at Peak
- Measured HPL power; **59+KW/rack, 5400KW from compute nodes**
- Direct water cooling of primary compute racks
- Oil immersion Cooling (GRC)
- Solar, Wind inputs

Thomas McConnell

THE KNOWLEDGE NEEDED TO SOLVE REAL PROBLEMS PRODUCTIVELY IS (ALWAYS) BEING CREATED

SOME OF THE CHALLENGES OF MODERN SYSTEMS

On Stampede2

- Each processor has 68 cores.
- Each core runs 4 hardware threads
- Each instruction can operate on 8 operands at a time (512-bit vectors).
- 68*4= 272. 272*8 = 2,176 ; your code needs to process 2,176 instructions *at a time* to run well on *one processor* of Stampede2.

TACC AT A GLANCE

Personnel

185 Staff (~70 PhD)

Facilities

12 MW Data center capacity
Two office buildings, Three Datacenters, two visualization facilities, and a chilling plant.

Systems and Services

>Seven Billion compute hours per year
>5 Billion files, >100 Petabytes of Data, NSF Frontera (Track 1), Stampede2 (XSEDE Flagship), Jetstream (Cloud), Chameleon (Cloud Testbed) system

Usage

>15,000 direct users in >4,000 projects, >50,000 web/portal users, User demand 8x available system time. Thousands of training/outreach participants annually

MODERN COMPUTATIONAL SCIENCE

Simulation
Computationally query our
mathematical models of the world

Machine Learning/AI
Computationally query our
data sets
(depending on technique,
also called deep learning)

Analytics
Computationally analyze our
experiments
(driven by instruments that produce
lots of digital information)

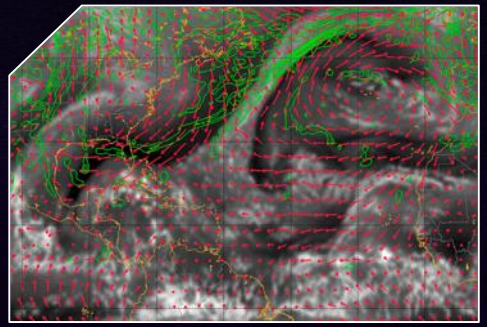
WHAT SUPERCOMPUTERS ARE FOR…

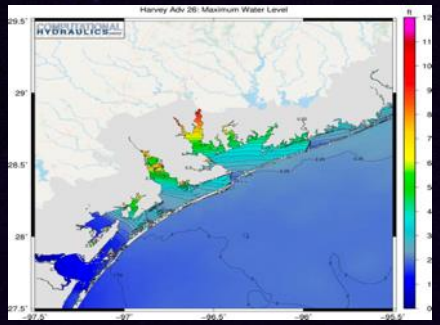
BLACK HOLE IMAGING

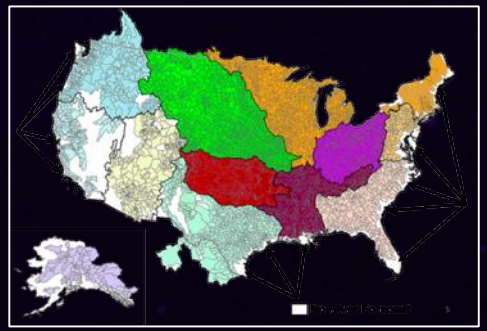
- In May, you may have seen the wide announcement (e.g. front page of NY Times) about the first successful imaging of a black hole corona (M87)
- A hugely complex project with massive data and eight different telescopes.
- Four TACC projects contributed compute time, software, or expertise to this.
  - Simulation runs of estimated M87 mass on Stampede 5 years ago.
  - Much of the new sims on Stampede2.
  - Design of the Cloud workflow on IU/TACC Jetstream system
- Much of the production data analysis was run on the Google Cloud (with pipelines designed in collaboration with TACC team).

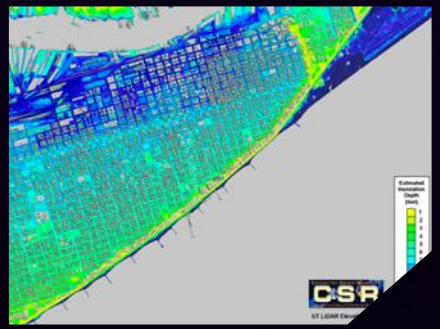
WHAT SUPERCOMPUTERS ARE FOR…

HURRICANE DORIAN









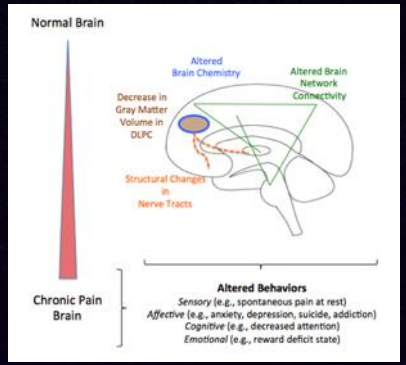
- Next Generation Storm Forecasting (with Penn State)
- Storm Surge Modeling (with Clint Dawson UT Austin)
- Preliminary river flooding and inundation maps (David Maidment UT Austin)
- Remote Image Integration and Assimilation (Center for Space Research, UT Austin)

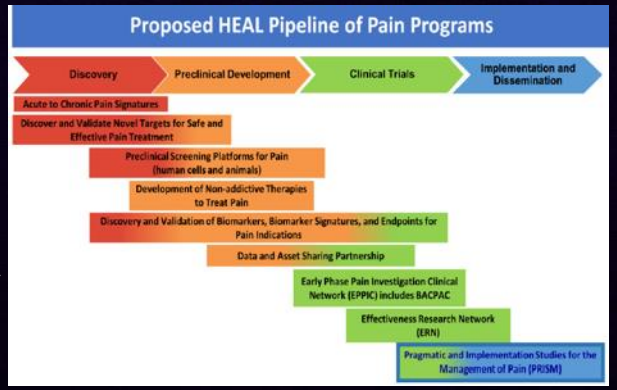
WHAT SUPERCOMPUTERS ARE FOR…

THE OPIOID CRISIS

A2CPS: Acute to Chronic Pain Signatures Program

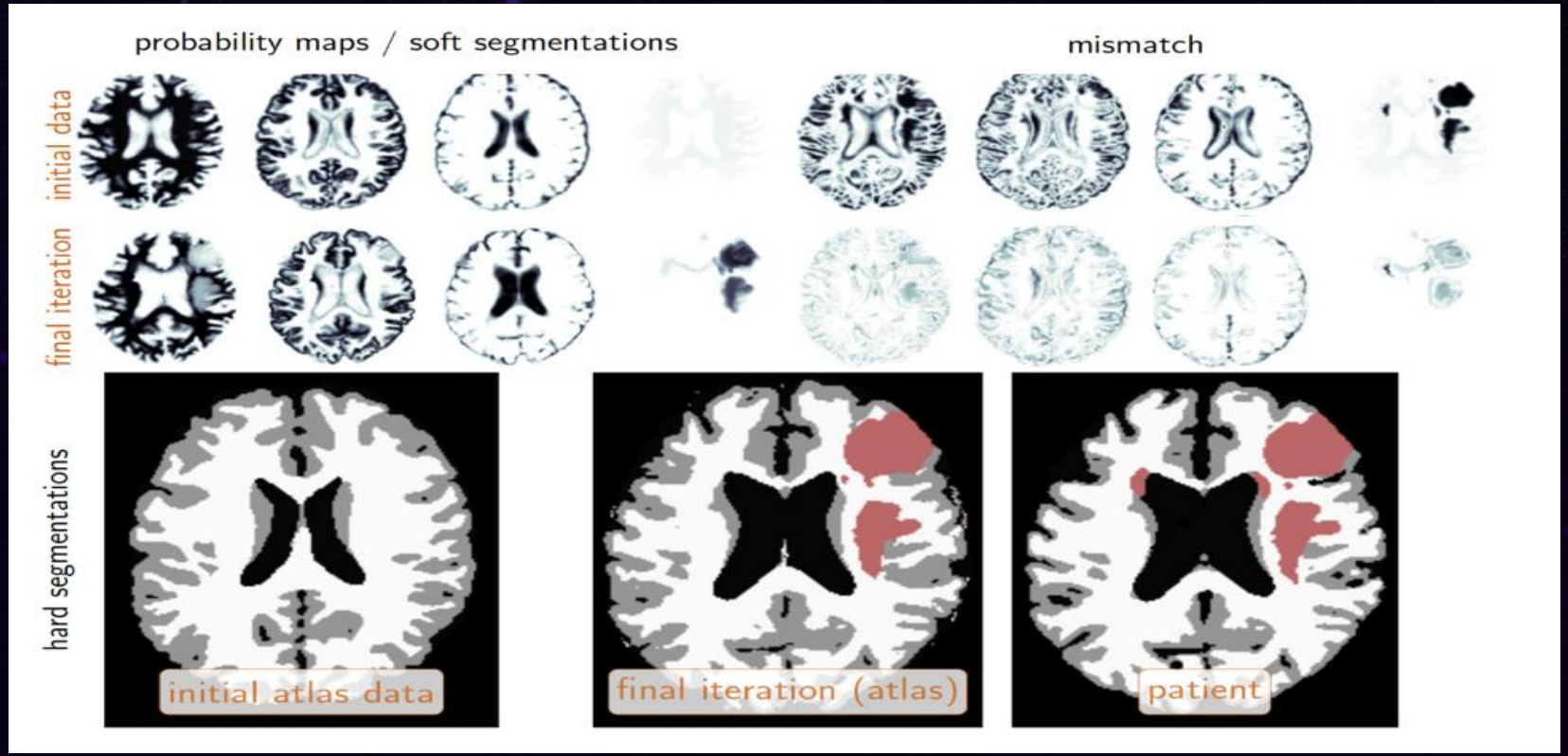
- Identify biosignatures (sets of biomarkers) that predict addiction to medicines.
- Catalyze development of new drugs that people won't get addicted to.
- Data Integration and Resource Center (DIRC) for Common Fund Acute to Chronic Pain Signatures Program





Goal of DIRC: Integrate the efforts of all components of the A2CPS and serve as a community-wide nexus for protocols, data, assay and data standards, and other resources generated by the A2CPS Program.

Datasets: Electronic health records, patient reported outcomes, accelerometer data, sensory testing, CNS imaging, 'omics assays (proteomic, metabolomic, lipidomic, extracellular RNA, transcriptome, array-based gene variants).

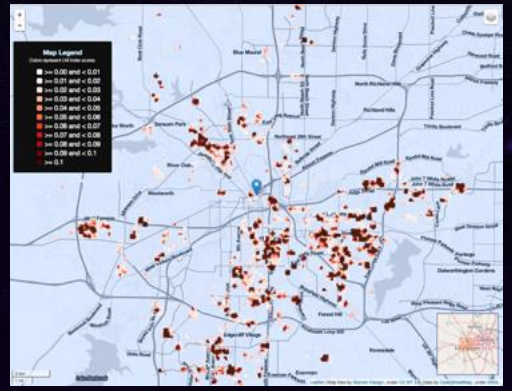


- George Biros from The University of Texas at Austin combines biophysical models of tumor growth with machine learning algorithms for the analysis of Magnetic Resonance imaging data of glioma patients to improve surgical outcomes.

HELPING COMMUNITIES THROUGH PREVENTION

ADDRESSING ISSUES BEFORE THEY BECOME PROBLEMS

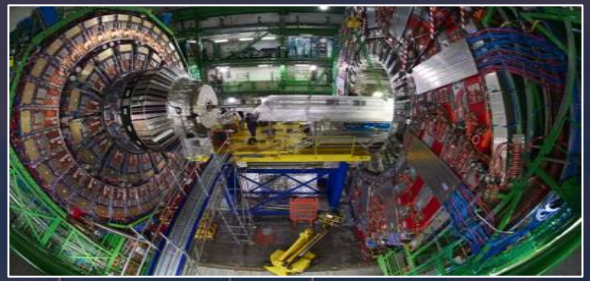
- Using AI to understand and intervene in the deep web which is 500 times what google searches
  - Linguistics studies
  - Arms and Human Trafficking
- Using data to provide positive inputs to reduce potential future problems with 100x spatial resolution
  - Child maltreatment
  - Health Improvements
- Modeling disease and impacts of different intervention plans

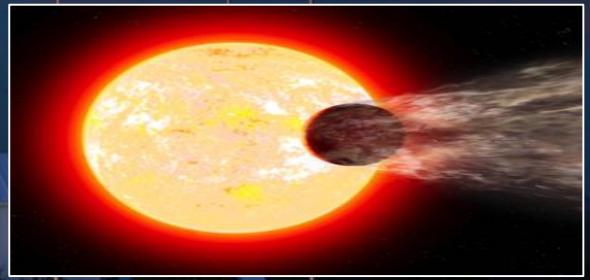


RESEARCHING THE FRONTIERS IN SPACE

MEASURING AND MODELING THE NEWEST DISCOVERIES OF THE UNIVERSE BOTH BIG AND SMALL

- Measuring and understanding never before detected events in our Universe
  - Gravitational Waves
  - Neutrinos from powerful Blazars
- Looking for new and habitable worlds in our Galaxy
  - Detecting and measuring atmospheres on exoplanets
  - Measuring and modeling star systems where habitable exo-planets can exist
- Revealing the nature of gravity, dark matter, and dark energy





TRAINING THE NEXT GENERATION OF RESEARCHERS

- Half- and full-day training event in everything from Linux basics to advanced MPI & Machine Learning
- Full 13 week classes on campus
- Summer institutes
- Code@TACC
- Internships and collaborative research
- SC Cluster Competitions
- Graduate and Undergraduate full-semester courses

THANK YOU

john@tacc.utexas.edu

bit.ly/TACC_STAR

TACC INDUSTRY PARTNERS

- Membership model
- Consulting visits
- Node hour allocation
- Training
- Collaboration events and workshops

As the relationship grows

- Additional allocations
- Technology investigations with TACC researchers
- Partnerships with your customers

Learn More: bit.ly/TACC_STAR