CENTER FOR
HIGH THROUGHPUT
COMPUTING

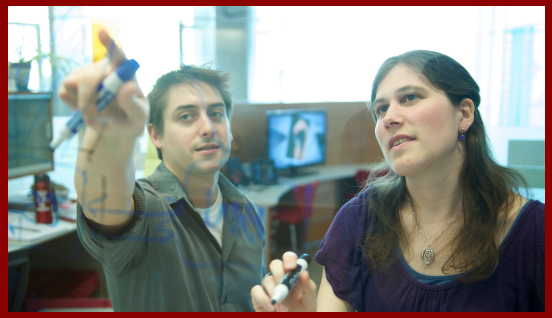
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Accelerating Research Transformations

- **Development of distributed computing technologies like HTCondor**
- **Computing infrastructure for campus research**
- **Research Computing *Facilitation***

CENTER FOR HIGH THROUGHPUT COMPUTING

serving computational research across campus



In the Last Year	Quick Facts
400	Million Hours Served
280	Research Projects
60	Departments

chtc.cs.wisc.edu

Agile, Shared Compute Systems

heterogeneous hardware, optimized compute modes

CHTC Pool

GPUs

high-memory

single-core

multi-core

MPI

CENTER FOR HIGH THROUGHPUT COMPUTING

Agile, Shared Compute Systems

high-throughput computing

>10,000

CPU hrs/day/user

CHTC Pool

GPUs

high-memory

single-core

multi-core

MPI

CENTER FOR HIGH THROUGHPUT COMPUTING

HTCondor

300,000+ jobs/day

400,000 hrs/day

Agile, Shared Compute Systems

elasticity, high-utilization resource sharing

CENTER FOR HIGH THROUGHPUT COMPUTING

HTCondor

Agile, Shared Compute Systems

"submit locally, run globally"

>200,000

CPU hrs/day/user

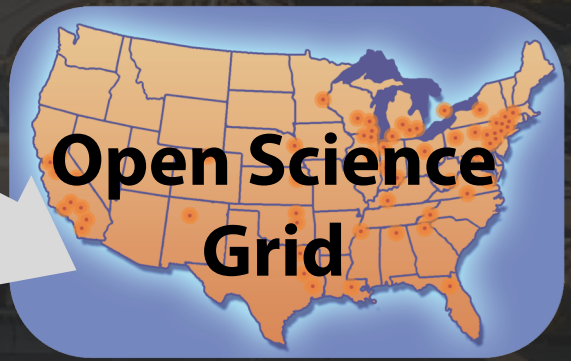
Campus Grid

CHTC

pool

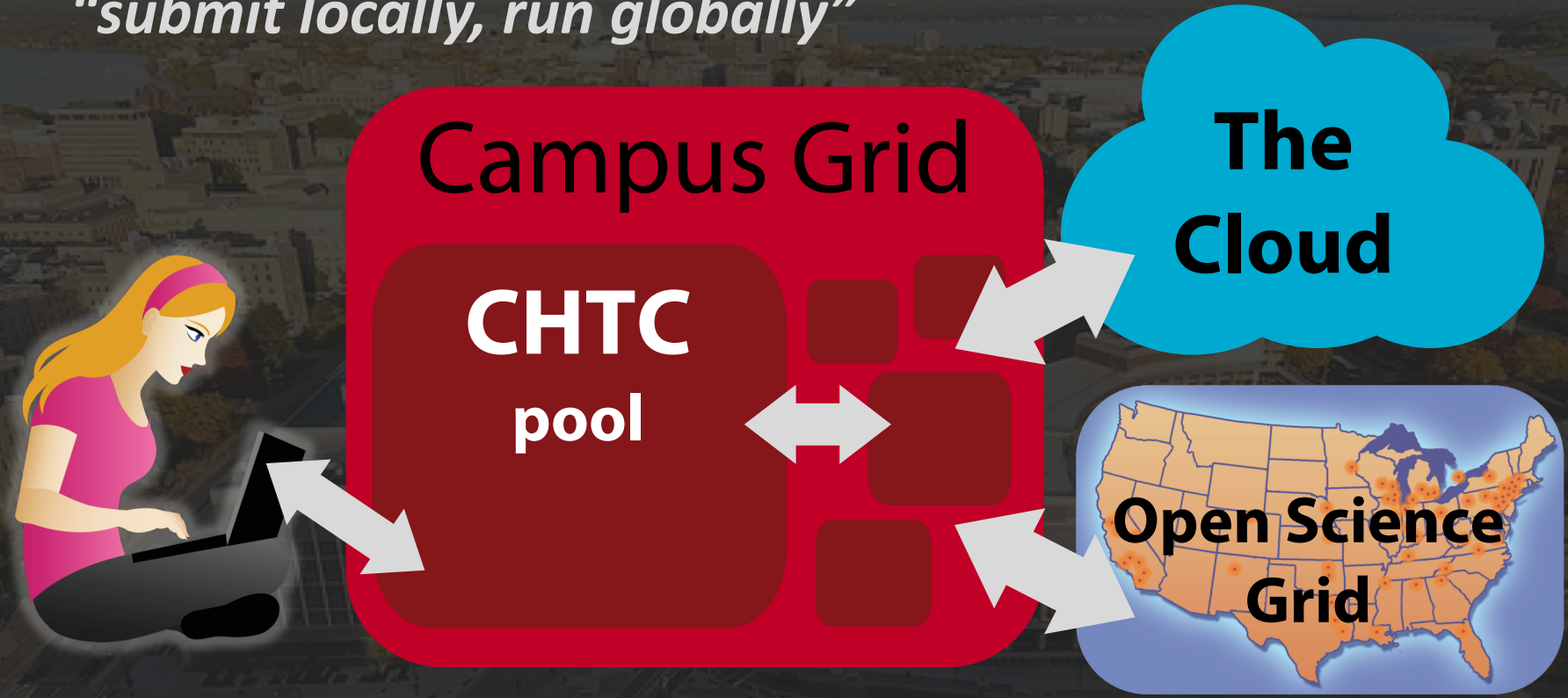
>1 million hours

delivered per day



Agile, Shared Compute Systems

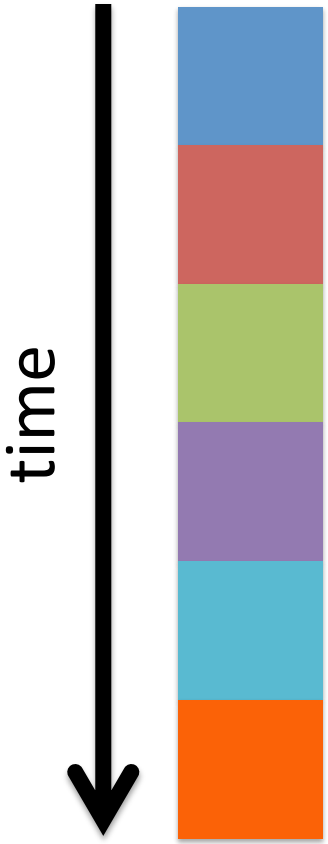
"submit locally, run globally"



Moving beyond the desktop

What many programs look like:

- *Serial* execution, running on one processor (CPU core) at a time
- Overall compute time grows significantly as individual tasks get more complicated (long) or if the number of tasks increases
- ***How can you speed things up?***



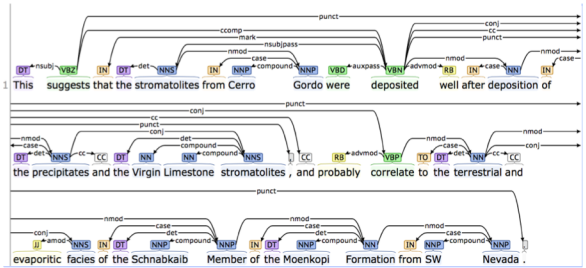
High-Throughput Computing

- Parallelize!
- Independent tasks run on different cores

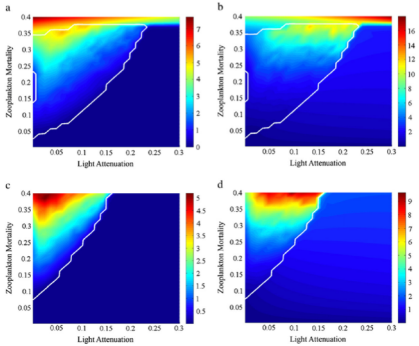
An HTC Analogy

HTC and the World's Largest Cake

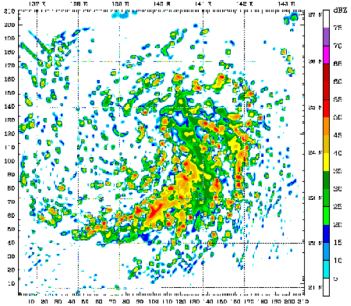
HTC Examples



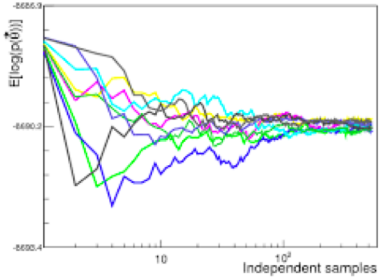
text analysis (most genomics …)



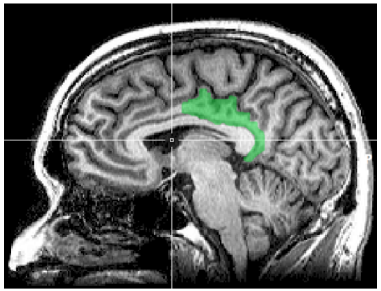
parameter sweeps



multi-start simulations



statistical model optimization
(MCMC, numerical methods, etc.)



(multi-)image and
sample analysis

Research Computing *Facilitation*

accelerating research transformations

proactive engagement

personalized guidance

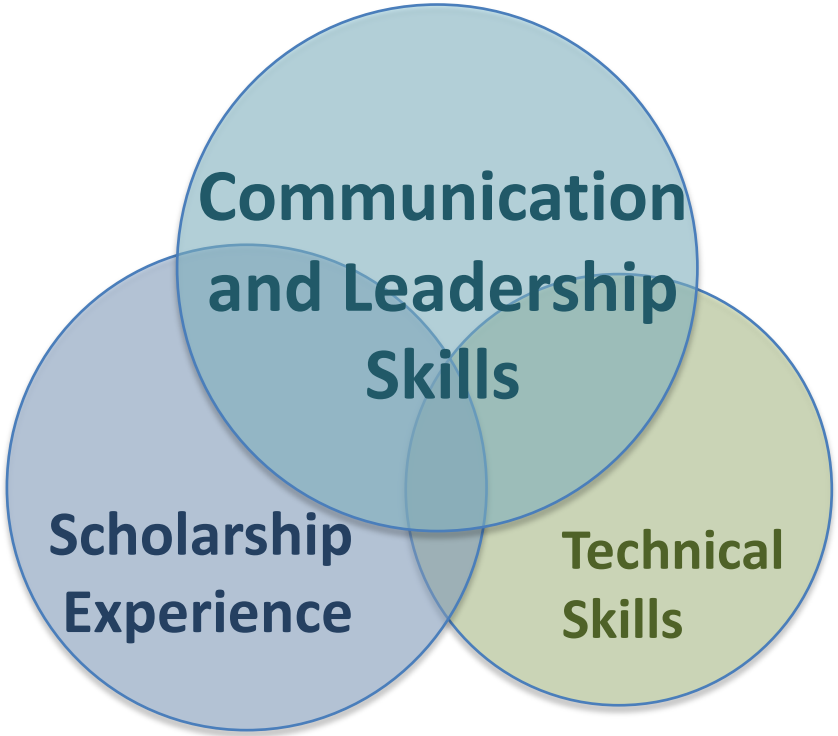
teach-to-fish training

technology agnostic

collaboration liaising

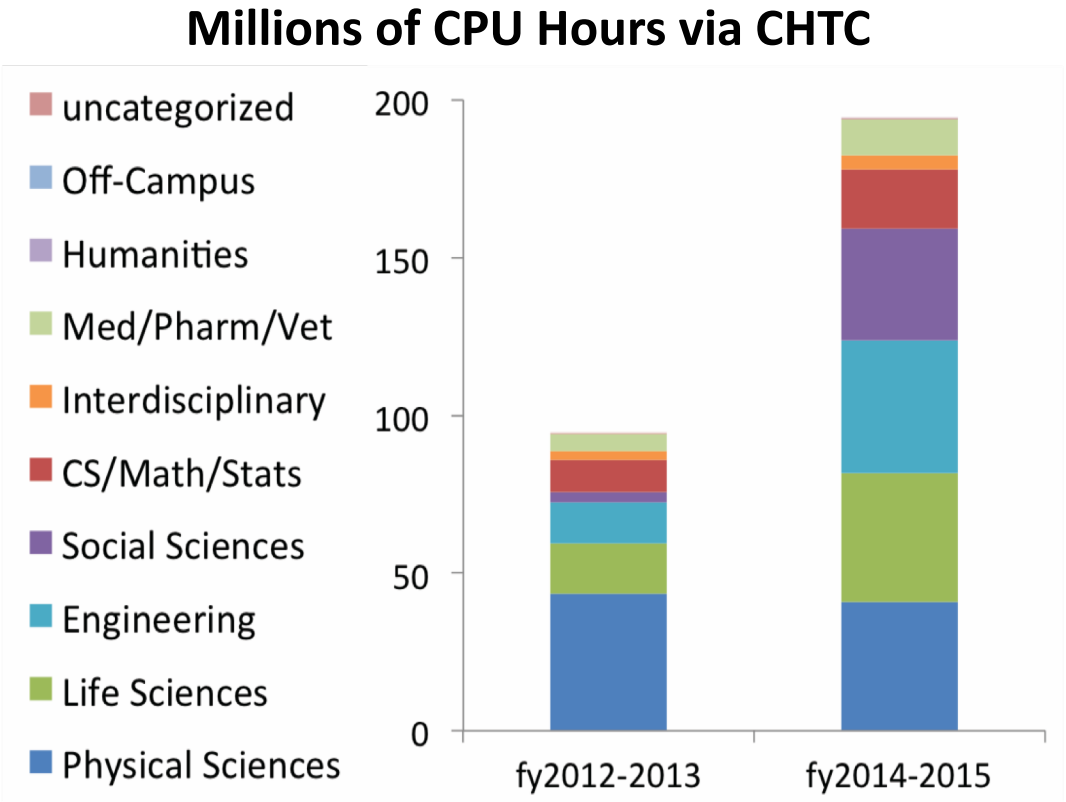
upward advocacy

Research Computing Facilitators



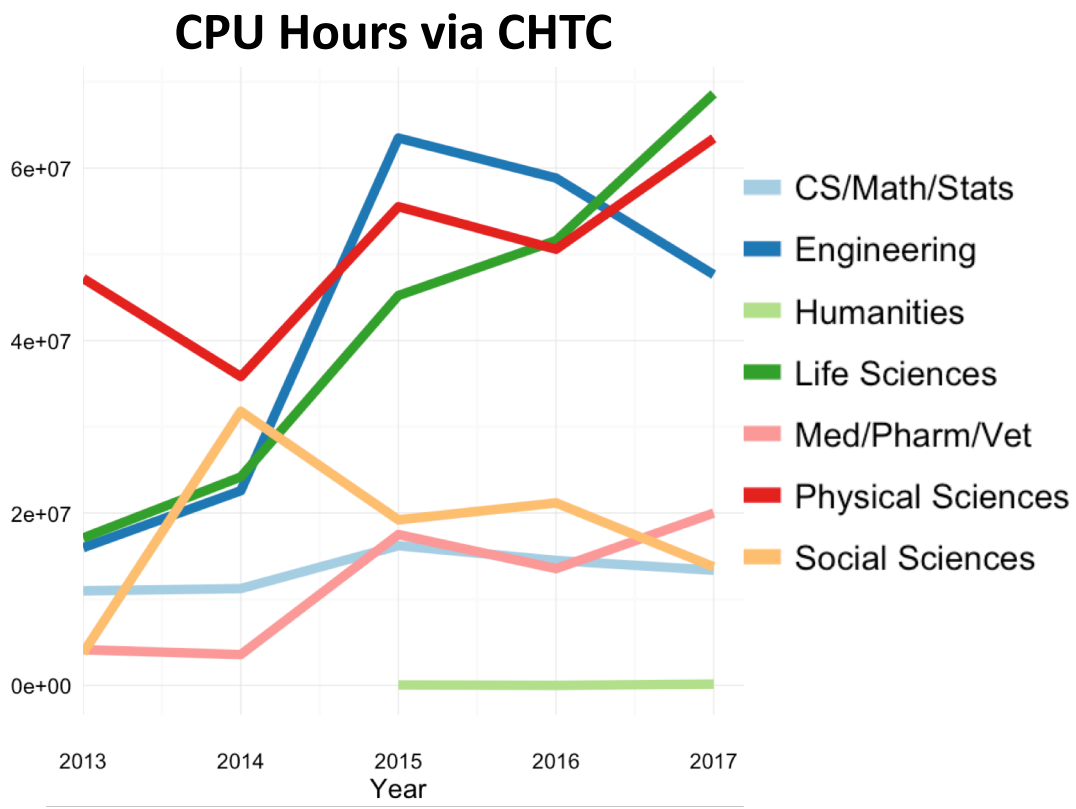
designated, dedicated, deliberate

Impact Across Domains



Facilitators hired:
Jan 2013, Nov 2014

Impact Across Domains



In 2018	
Research Groups Supported	280
Researcher Consultations	495
Office Hours Visits	420

Marshfield Clinic

High-Throughput Machine Learning from Electronic Health Records

Ross Kleiman

Paul Bennett, Scott Hebbring, Charles Kuang, Peggy Peissig, Michael Caldwell, David Page, Finn Kuusisto, Ron Stewart

Predicting All Diagnoses

- Prior work: Individual disease models
- How well can we predict **all** diagnoses?

- **Given:** All EHR data
- **Do:** Learn model to predict each ~~disease~~ ICD-9 Code

- Build a high-throughput machine learning pipeline

>100 Years of Computing!

KinderMiner

Output Rank

1. Compute article count contingency tables
2. Filter terms by one-sided Fisher Exact test
3. Sort terms by $\frac{Key\ Phrase\ \&\ Term}{Term\ Total}$

Example

NANOG + "Embryonic stem cell" + 2004

NANOG	Term	⌐ Term	Total
Key Phrase	15	2,012	**2,027**
⌐ Key Phrase	44	17,010,295	**17,010,339**
Total	**59**	**17,012,307**	**17,012,366**

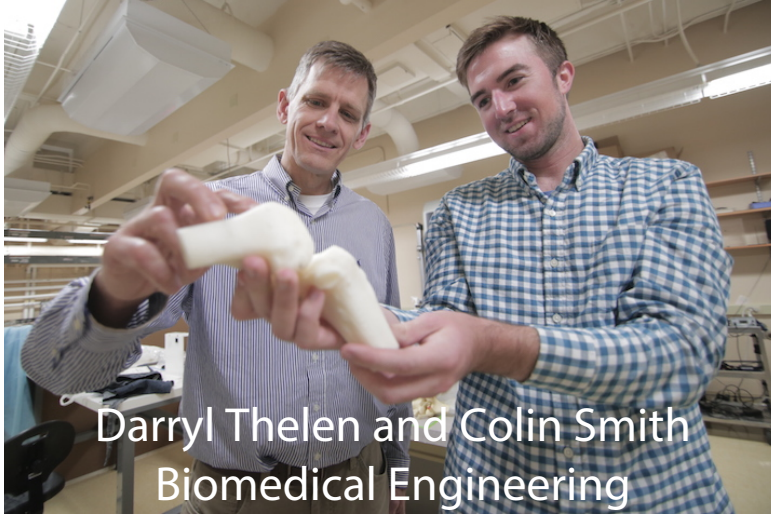
One-sided FET p: 5.219e-46 **Sort Ratio:** $\frac{15}{59} = 0.254$

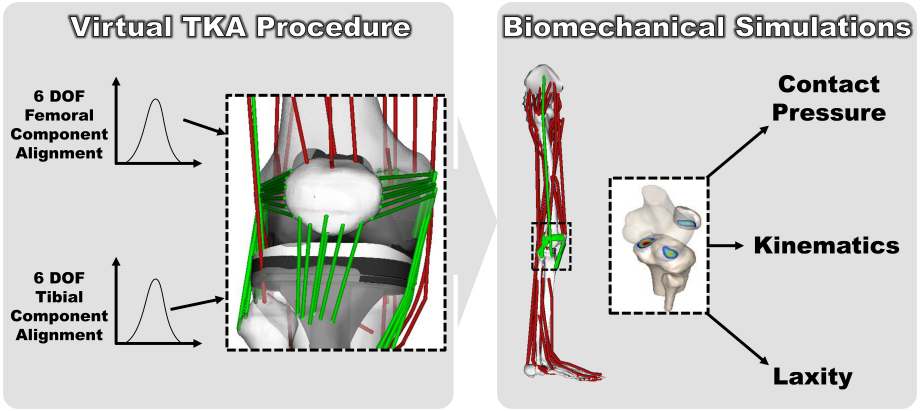
MORGRIDGE INSTITUTE FOR RESEARCH

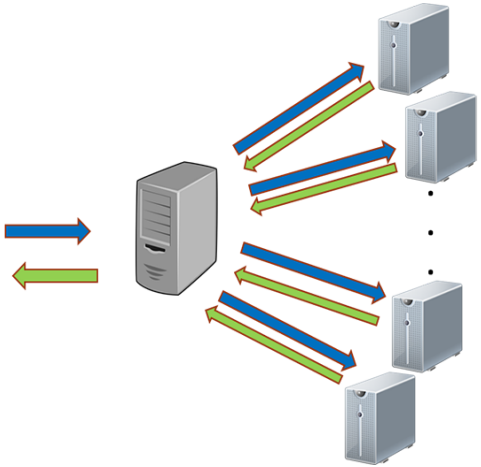
ABOUT RESEARCH OUTREACH NEWS

High-throughput computing plays pivotal role in knee biomechanics research

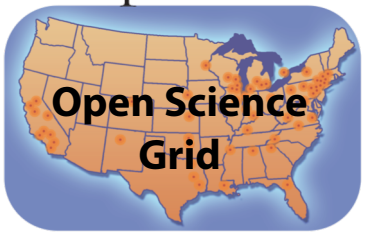
by Ashley Osbourne | February 24, 2017







OpenSim



CCBCDeepDive Digital Research Library (CCBC-D3) Pilot Project

The Cooperative Children's Book Center (CCBC), known internationally for groundbreaking work on children's literature, is teaming with data analytics experts from UW–Madison's Computer Sciences Department to study how race, ethnicity, gender, plot, location and character types vary across children's books. Project members are: **KT Horning**, director of the School of Education's CCBC; **Brenda Spychalla**, co-CIO for the School of Education; **Miron Livny**, a world renowned UW–Madison computer scientist; and **Lauren Michael** of the Wisconsin Institute for Discovery.

LEARNING CONNECTIONS, SCHOOL OF ED

School of Education
UNIVERSITY OF WISCONSIN–MADISON

THE CAP TIMES

EDITOR'S PICK TOPICAL

'Open the barn doors, Hal!': Artificial intelligence could one day run a dairy farm

ERIK LORENZSONN | The Capital Times | eriki@madison.com | @eriklorenzsonn

May 4, 2017

UNIVERSITY of WISCONSIN MADISON

CALS NEWS
COLLEGE OF AGRICULTURAL & LIFE SCIENCES

UW team's "virtual dairy farm brain" aims to help farmers make smarter decisions

August 22, 2017 |

UW dairy scientists are no strangers to data management, but wrangling so many streams of disparate data in real time requires a specialized skill set. That's why they're collaborating with the UW's Center for High Throughput Computing.

chtc@cs.wisc.edu

chtc.cs.wisc.edu > How To > **Get Started**

lmichael@wisc.edu

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON