High Throughput Computing in the Service of Scientific Discovery

Miron Livny
John P. Morgridge Professor of Computer Science
Center for High Throughput Computing
University of Wisconsin-Madison

MORGRIDGE INSTITUTE FOR RESEARCH

HT CENTER FOR HIGH THROUGHPUT COMPUTING

WISCONSIN UNIVERSITY OF WISCONSIN-MADISON


UNIVERSITY OF WISCONSIN–
MADISON

Some Numbers

ENROLLMENT: 43,820

- Undergraduate, graduate and professional students from 50 states and more than 120 countries
- Faculty and staff: 22,038
- Budget: $3 Billion
- Sponsored Research: $1.2 Billion
- 936-acre lakefront campus

Horizontal Markets vs. Vertical Markets

Businesses that operate in a horizontal market system seek to appeal to a wide demographic that's not really niche. For example, a reseller of general office furniture is probably not going to target (sell to) other companies that specialize in office furniture. Rather, they're going to target all types of businesses that maintain offices – accounting firms, travel agencies, insurance agencies, etc. Their market is anyone who needs office furniture.

Vertical marketing tries to attract a very niche demographic. For example, this could include a manufacturer of solar panel technology. These types of firms usually sell their goods to solar contractors and installers. In other words, those they sell to are usually businesses that compete against one another.

Verticals solve problems

When it comes to Computing, (most) domain researchers look for an (end-to-end) vertical solution

- Solution typically consists of layers of software that interface the researcher to the computing resources
- Solution harness processing, storage and communication capacity offered by hardware
- Adoption depends on effectiveness

Horizontals Provide Capabilities

(most) Computer Scientists develop horizontal frameworks and technologies that offer capabilities

- Serve as a layer of a vertical solution
- Breadth increases value/merit of (computer) science
- Best if evaluated experimentally as part of verticals with engaged end users and on production environments

MIT News

MIT will reshape itself to shape the future, investing $1 billion to address the rapid evolution of computing and AI — and its global effects. At the heart of this effort: a $350 million gift to found the MIT Stephen A. Schwarzman College of Computing.

Photo: Christopher Harting

MIT reshapes itself to shape the future

MIT News Office
October 15, 2018

The College will:

- reorient MIT to bring the power of computing and AI to all fields of study at MIT, allowing the future of computing and AI to be shaped by insights from all other disciplines;

The UW-Madison Center for High Throughput Computing (**CHTC**) was established in 2006 **to bring the power of Distributed High Throughput Computing (HTC) to all fields of study, allowing the future of Distributed HTC to be shaped by insight from other disciplines**

MORGRIDGE INSTITUTE FOR RESEARCH
CENTER FOR HIGH THROUGHPUT COMPUTING
WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

Harmonize Verticals and Horizontals (Square the Circle)

I am using Approximate Bayesian Computation method to figure out how mice evolved on the island and the genetic basis for their gigantism. The method involves extensive coalescent simulations, especially when I am trying to do it on the genomic scale, which requires very large amount of computer resources (CPUs). Without CHTC, I am not able to conduct this project at all. I will have to scale back the project significantly, which would reduce the power to make meaningful inference. So CHTC is essential to my project.

Design, Develop, Deploy, Operate and Evaluate (DDDOE)

CHTC is committed to put the best vertical solutions in the hands of UW-Madison researchers

- Upper layers are domain specific selected/provided by the researcher
- Lower layers (and hardware) operated by **CHTC**
- (best?) Middle layers selected and operated by **CHTC**
- Partner in the entire life-cycle of the solution
- Facilitate translation of (computing) capabilities into new/different research methods/directions

Advancing Computing Capabilities

- **CHTC** has been developing widely adopted (includes  ) Distributed High Throughput Computing (DHTC) software tools (**HTCondor**) that are based on novel frameworks and cross different Computer science areas
- **CHTC** pioneered Research Computing Facilitation (**RCF**)
- In the past 12 month **CHTC** delivered to the more than 250 projects across the UW campus more than 1K RCF hours and close to 400M core hours

Top 10 projects from latest 24 hours report from the CHTC

Fm:	2019-10-23	Total	
To:	2019-10-24	Hours	%Pool
90	Projects	1,133,729	100.0%
1	CMS	327,130	28.9%
2	IceCube	286,665	25.3%
3	Physics_Yavuz	51,781	4.6%
4	BMRB	33,795	3.0%
5	Math_Sukiennik	28,233	2.5%
6	EdPsychology_Kaplan	27,898	2.5%
7	OSG	26,954	2.4%
8	Bacteriology_Rey	25,688	2.3%
9	ChemBioE_Abbott	24,532	2.2%
10	NutritionalSciences_Parks	18,690	1.6%

Research Computing *Facilitation*

accelerating research transformations

proactive engagement

personalized guidance

teach-to-fish training

technology agnostic

collaboration liaising

upward advocacy

Summary

Discovery of astrophysical neutrinos - new era of Neutrino Astronomy started.

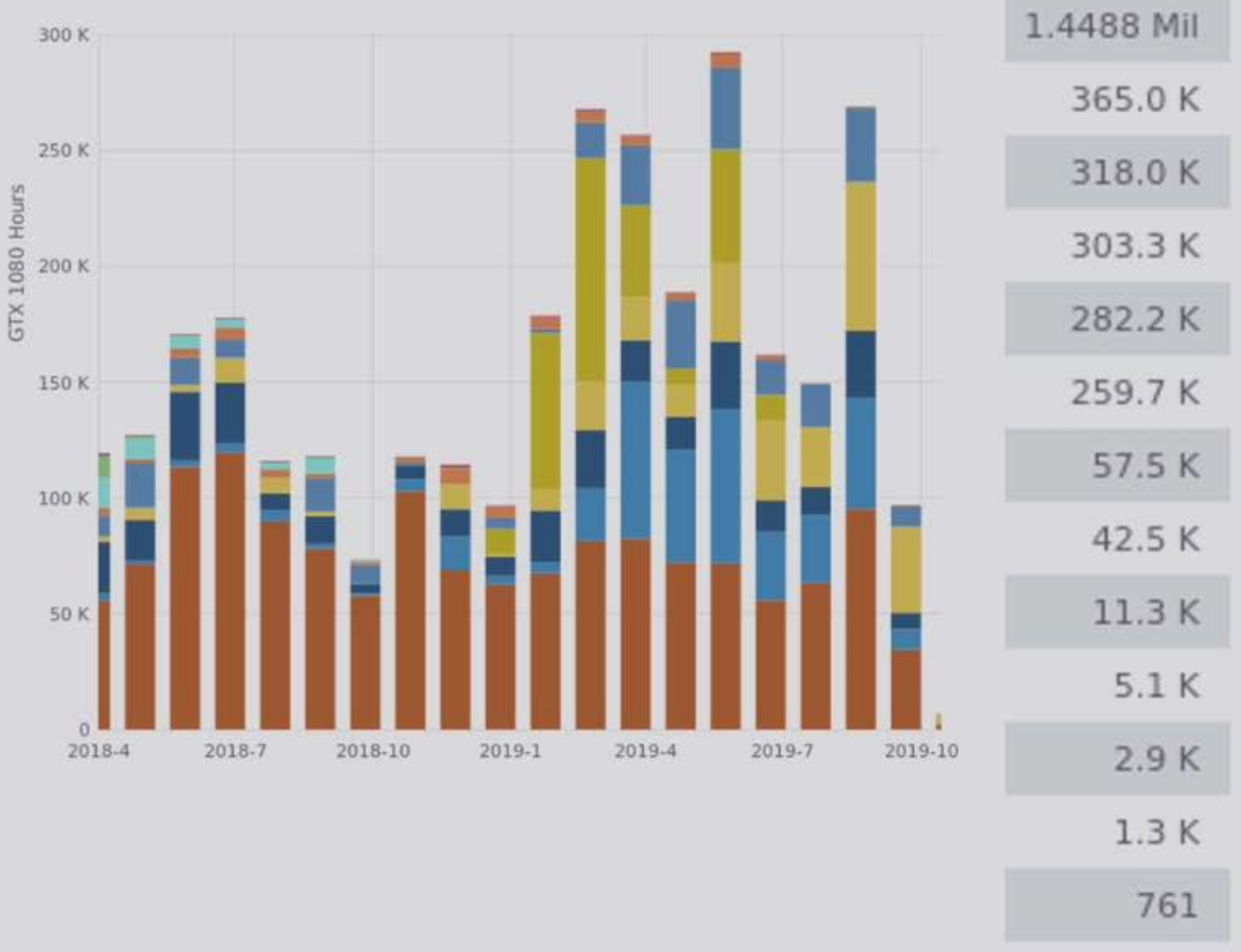
Beyond 5 years after construction ended, IceCube keeps improving its uptime

- A rich physics program ahead
- Transition from discovery to precision measurement phase

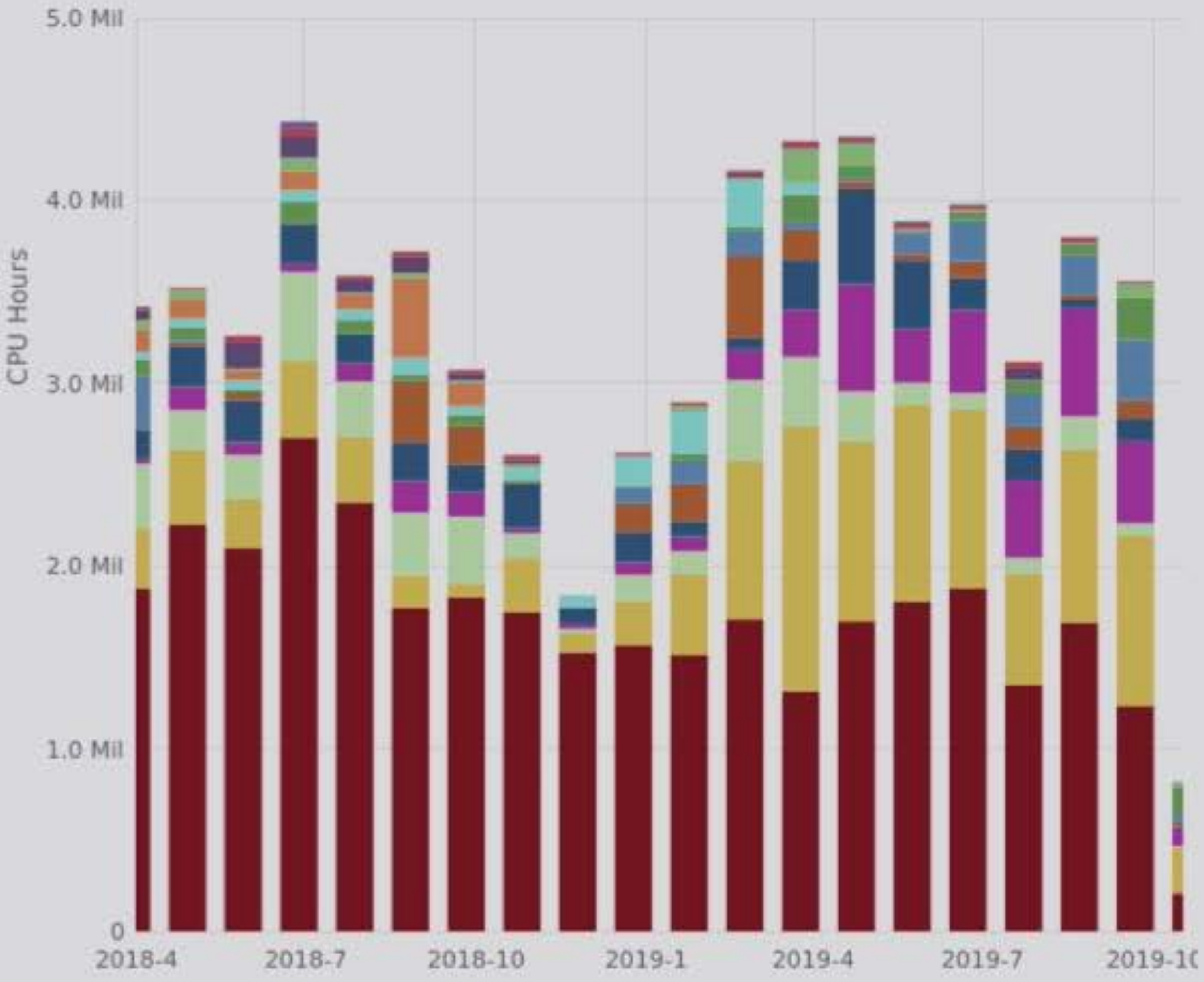
Simulation is essential - light propagation in the ice & related systematics

- Strongly rely on distributed computing
  - Benefiting a lot of common areas with LHC: CVMFS, opportunistic access to WLCG sites ...
  - Infrastructure based in HTCondor components - user interface is HTCondor
- GPUs a critical platform for IceCube

Normalized GPU hours per Institution



CPU hours per Institution



total
34.082 Mil
11.275 Mil
4.355 Mil
4.212 Mil
3.705 Mil
2.035 Mil
1.835 Mil
1.297 Mil
1.274 Mil
1.036 Mil
735 K
663 K
413 K
40 K
18 K
801
458

IceCube
South Pole Neutrino Observatory

IceCube neutrinos point to long-sought cosmic ray accelerator

A blazar is a giant elliptical galaxy with a massive, rapidly spinning black hole at its core. A signature feature of blazars is that twin jets of light and elementary particles, one of which is pointing to Earth, are emitted from the poles along the axis of the black hole's rotation. This blazar is situated in the night sky just off the left shoulder of the constellation Orion and is about 4 billion light years from Earth.

From: **HONG ZHAN** <hzhan3@wisc.edu>
Date: Mon, Jun 3, 2019 at 14:03
Subject: Thanks for all help and for more in the future!
To: LAUREN A MICHAEL <lmichael@wisc.edu>
CC: Christina Koch <ckoch5@wisc.edu>

Hi Lauren and Christina,

I would like to share some exciting news that I just got this 12A structure of our protein of interests recently with the help from you two on the HTC. I eagled to make it better. I am sorry I did not get this for the HTCondor presentation, but it will be for next year workshop if I still could be invited:D

If ok, could we make a meeting together sometime in July?

[12A of our structure with potentially closed protein motifs fitting in our EM density map]

Best,
Hong

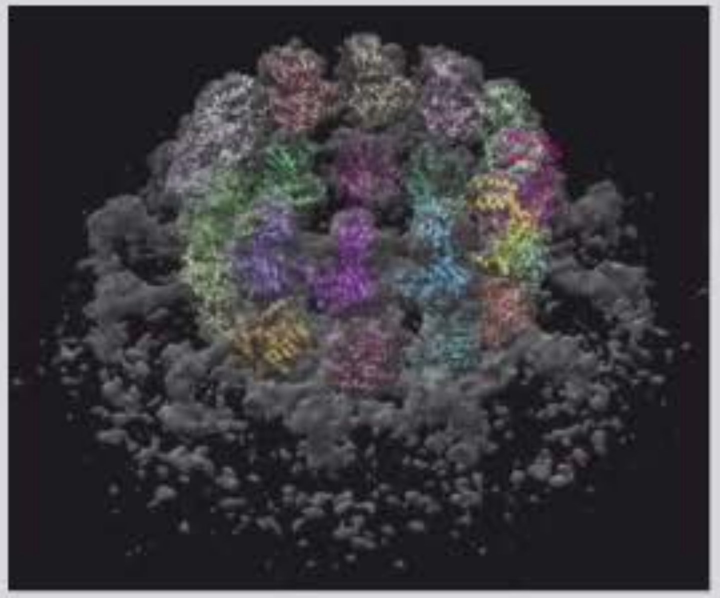
Post Doctoral Research Associate

Area:
John W. and Jeanne M. Rowe Center for Research in Virology

Cell Phone:
(608) 265-9734

Email:
hzhan3@wisc.edu

I am interested in the ultrastructure of proteins and protein complex. My goal is to use cyro electron microscope to decipher virus proteins' structure to understand their role in virus infection and RNA replication process.



MORGRIDGE INSTITUTE FOR RESEARCH

CENTER FOR HIGH THROUGHPUT COMPUTING

WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

Hunting Gravitational Waves with HTCondor

Carsten Aulbert
Max Planck Institute for Gravitational Physics
Albert-Einstein-Institute
Hannover branch

2017-06-08

Thank you HTCondor Team

- not sure in which manner LIGO would have made detection without HTCondor and its ecosystem
- after 15+ years of usage it is our solid basement/paradigm of computing
- new features actually *are* useful — quite unlike other large software projects
- had no time to address these, but would like to learn/discuss how *you* tackle problems
  - user education about cluster/condor usage,
  - storing multi PB data sets, handling literally billions of files,
  - automate "everything",
  - **reproducible science** (Singularity + more?)

(HT)Condor was first deployed in production at the UW-Madison Computer Sciences Department in 1985

Claims for “benefits” provided by Distributed Processing Systems

P.H. Enslow, *“What is a Distributed Data Processing System?”* Computer, January 1978

– High Availability and Reliability

– High System Performance

– Ease of Modular and Incremental Growth

– Automatic Load and Resource Sharing

– Good Response to Temporary Overloads

– Easy Expansion in Capacity and/or Function

Definitional Criteria for a Distributed Processing System

P.H. Enslow and T. G. Saponas ""*Distributed and Decentralized Control in Fully Distributed Processing Systems*" Technical Report, 1981

- Multiplicity of resources
- Component interconnection
- **Unity of control**
- System transparency
- **Component autonomy**

Unity of Control

All the component of the system should be **unified** in their desire to achieve a **common goal**. This goal will determine the rules according to which each of these elements will be controlled.

MORGRIDGE INSTITUTE FOR RESEARCH

HT CENTER FOR HIGH THROUGHPUT COMPUTING

WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

Component autonomy

The components of the system, both the logical and physical, should be **autonomous** and are thus afforded the ability to refuse a request of service made by another element. However, in order to achieve the system's goals they have to interact in a **cooperative** manner and thus adhere to a common set of policies. These policies should be carried out by the control schemes of each element.

The Structure of the "THE"-Multiprogramming System

Edsger W. Dijkstra
Technological University, Eindhoven, The Netherlands

A multiprogramming system is described in which all activities are divided over a number of sequential processes. These sequential processes are placed at various hierarchical levels, in each of which one or more independent abstractions have been implemented. The hierarchical structure proved to be vital for the verification of the logical soundness of the design and the correctness of its implementation.

Presented at an ACM Symposium on Operating System Principles, Gatlinburg, Tennessee, October 1–4, 1967.

On the second page of the paper …

Our first major mistake was that for too long a time we confined our attention to "a perfect installation"; by the time we considered how to make the best of it, one of the peripherals broke down, we were faced with nasty problems. Taking care of the "pathology" took more energy than we had expected, and some of our troubles were a direct consequence of our earlier ingenuity, i.e. the complexity of the situation into which the system could have maneuvered itself. Had we paid attention to the pathology at an earlier stage of the design, our management rules would certainly have been less refined.

The second major mistake has been that we conceived and programmed the major part of the system without giving more than scanty thought to the problem of debugging it. I must decline all credit for the fact that this mistake had no serious consequences—on the contrary! one might argue as an afterthought.

CERN 92

Flock

Global Scientific Computing via a Flock of Condors

Miron Livny

Computer Sciences Department
University of Wisconsin — Madison
Madison, Wisconsin
{miron@cs.wisc.edu}

Flock

MISSION

Give scientists effective and efficient access to large amounts of cheap (if possible free) CPU cycles and main memory storage

Flock

THE CHALLENGE

How to turn existing privetly owned clusters of *workstations*, *farms*, *multiprocessors*, and *supercomputers* into an efficient and effective Global Computing Environment?

In other words, how to minimize wait while idle?

Flock

APPROACH

Use wide-area networks to transfer batch jobs between Condor systems

- Boundaries of each Condor system will be determined by physical or administrative considerations

Flock

TWO EFFORTS

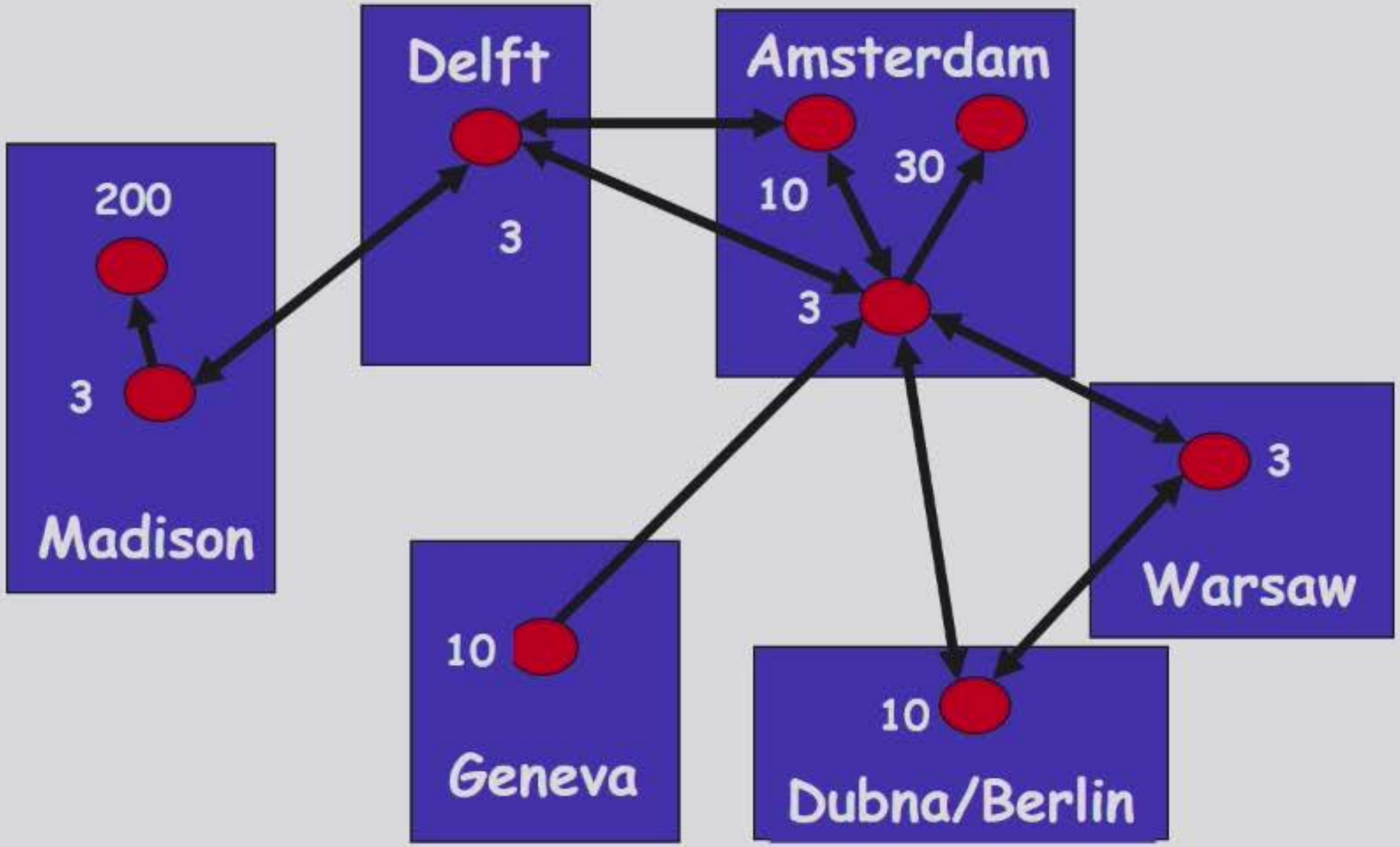
☐ ***UW CAMPUS***

Condor systems at Engineering, Statistics, and Computer Sciences

☐ ***INTERNATIONAL***

We have started a collaboration between CERN-SMC-NIKHEF-Univ. of Amsterdam, and University of Wisconsin-Madison

1994 Worldwide Flock of Condors



D. H. J Epema, Miron Livny, R. van Dantzig, X. Evers, and Jim Pruyne, "*A Worldwide Flock of Condors : Load Sharing among Workstation Clusters*" *Journal on Future Generations of Computer Systems*, Volume 12, 1996

CERN Site Report

Andrei Dumitru
CERN IT Department
14th of October 2019

BATCH, HPC and CERNMegabus

Batch migration to CERN CentOS 7 (99%)
LSF retired now

- HTCondor in production for all batch

In 1996 I introduced the distinction between High **Performance** Computing (**HPC**) and High **Throughput** Computing (**HTC**) in a seminar at the NASA Goddard Flight Center in and a month later at the European Laboratory for Particle Physics (CERN). In June of 1997 HPCWire published an interview on High Throughput Computing.

```
HIGH THROUGHPUT COMPUTING: AN INTERVIEW WITH MIRON LIVNY               06.27.97
by Alan Beck, editor in chief                                           HPCwire
===============================================================================

  This month, NCSA's (National Center for Supercomputing Applications)
Advanced Computing Group (ACG) will begin testing Condor, a software system
developed at the University of Wisconsin that promises to expand computing
capabilities through efficient capture of cycles on idle machines. The
software, operating within an HTC (High Throughput Computing) rather than a
traditional HPC (High Performance Computing) paradigm, organizes machines
```

High Throughput Computing requires automation as it is a 24-7-365 activity that involves large numbers of jobs and computing resources

FLOPY ≠ (60*60*24*7*52)*FLOPS

100K Hours*1 Job ≠ 1 H*100K J

MORGRIDGE INSTITUTE FOR RESEARCH

HT CENTER FOR HIGH THROUGHPUT COMPUTING

WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020

AUTHORS

Committee on Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science in 2017-2020; Computer Science and Telecommunications Board; Division on Engineering and Physical Sciences; National Academies of Sciences, Engineering, and Medicine

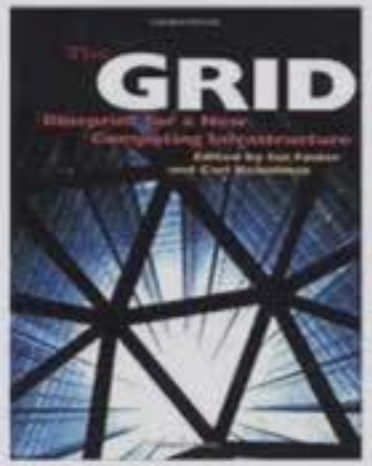
"... many fields today rely on **high-throughput computing** for **discovery**."

"Many fields **increasingly** rely **on high-throughput computing**"

Mechanisms hold the key

The most important lesson our HTC experience has taught us is that in order to deliver and sustain high throughput over long time intervals, a computing environment must build its resource management services on an integrated collection of robust, scalable and portable mechanisms. Robustness minimizes down time whereas scalability and portability increases the size of the resource pool the environment can draw upon to serve its customers. As will be argued

The Grid: Blueprint for a New Computing Infrastructure
Edited by Ian Foster and Carl Kesselman
July 1998, 701 pages.



1992

Learn

What Did We Learn From Serving a Quarter of a Million Batch Jobs on a Cluster of Privately Owned Workstations

Miron Livny

Computer Sciences Department
University of Wisconsin — Madison
Madison, Wisconsin
{miron@cs.wisc.edu}

Learn

User Prospective

- Maximize the capacity of resources accessible via a single interface
- Minimize overhead of accessing remote capacity
- Preserve local computation environment

Submit locally

(queue and manage your resource acquisition and job execution locally)

and run globally

(acquire and use any resource that is capable and willing to run your jobs)

MORGRIDGE INSTITUTE FOR RESEARCH

HT CENTER FOR HIGH THROUGHPUT COMPUTING

WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

desktop

Est.
2003

floor

desktop

Est.
2003

Est.
2003

Est.
2003

The current (young) generation of researchers transitioned from the desk/lap top to the Jupyter notebook

- Researcher "lives" in the notebook
- Bring Python to the HTC environment – bindings and APIs
- Bring HTC to Python – the HTMap module
- Support testing and debugging of HTC applications/workflows in the notebook

Agile, Shared Computing

"submit locally, run globally"