

# IceProd Supercomputer Mode

How IceCube Production Runs on Firewalled Clusters

David Schultz  
WIPAC, UW-Madison



# Outline

- What is IceCube / IceProd
- Supercomputer Specialties
- Current Usage
- Future Plans

# IceCube Collaboration

**THE ICECUBE COLLABORATION**

**AUSTRALIA**  
University of Adelaide

**BELGIUM**  
Université libre de Bruxelles  
Universiteit Gent  
Vrije Universiteit Brussel

**CANADA**  
SNOLAB  
University of Alberta-Edmonton

**DENMARK**  
University of Copenhagen

**GERMANY**  
Deutsches Elektronen-Synchrotron  
ECAP, Universität Erlangen-Nürnberg  
Humboldt-Universität zu Berlin  
Ruhr-Universität Bochum  
RWTH Aachen University  
Technische Universität Dortmund  
Technische Universität München  
Universität Mainz  
Universität Wuppertal  
Westfälische Wilhelms-Universität  
Münster

**JAPAN**  
Chiba University

**NEW ZEALAND**  
University of Canterbury

**REPUBLIC OF KOREA**  
Sungkyunkwan University

**SWEDEN**  
Stockholms universitet  
Uppsala universitet

**SWITZERLAND**  
Université de Genève

**UNITED KINGDOM**  
University of Oxford

**UNITED STATES**  
Clark Atlanta University  
Drexel University  
Georgia Institute of Technology  
Lawrence Berkeley National Lab  
Marquette University  
Massachusetts Institute of Technology  
Michigan State University  
Ohio State University  
Pennsylvania State University  
South Dakota School of Mines and  
Technology

Southern University  
and A&M College  
Stony Brook University  
University of Alabama  
University of Alaska Anchorage  
University of California, Berkeley  
University of California, Irvine  
University of California, Los Angeles  
University of Delaware  
University of Kansas  
University of Maryland  
University of Rochester

University of Texas at Arlington  
University of Wisconsin-Madison  
University of Wisconsin-River Falls  
Yale University

**FUNDING AGENCIES**

Fonds de la Recherche Scientifique (FRS-FNRS)  
Fonds Wetenschappelijk Onderzoek-Vlaanderen  
(FWO-Vlaanderen)

Federal Ministry of Education and Research (BMBF)  
German Research Foundation (DFG)  
Deutsches Elektronen-Synchrotron (DESY)

Japan Society for the Promotion of Science (JSPS)  
Knut and Alice Wallenberg Foundation  
Swedish Polar Research Secretariat

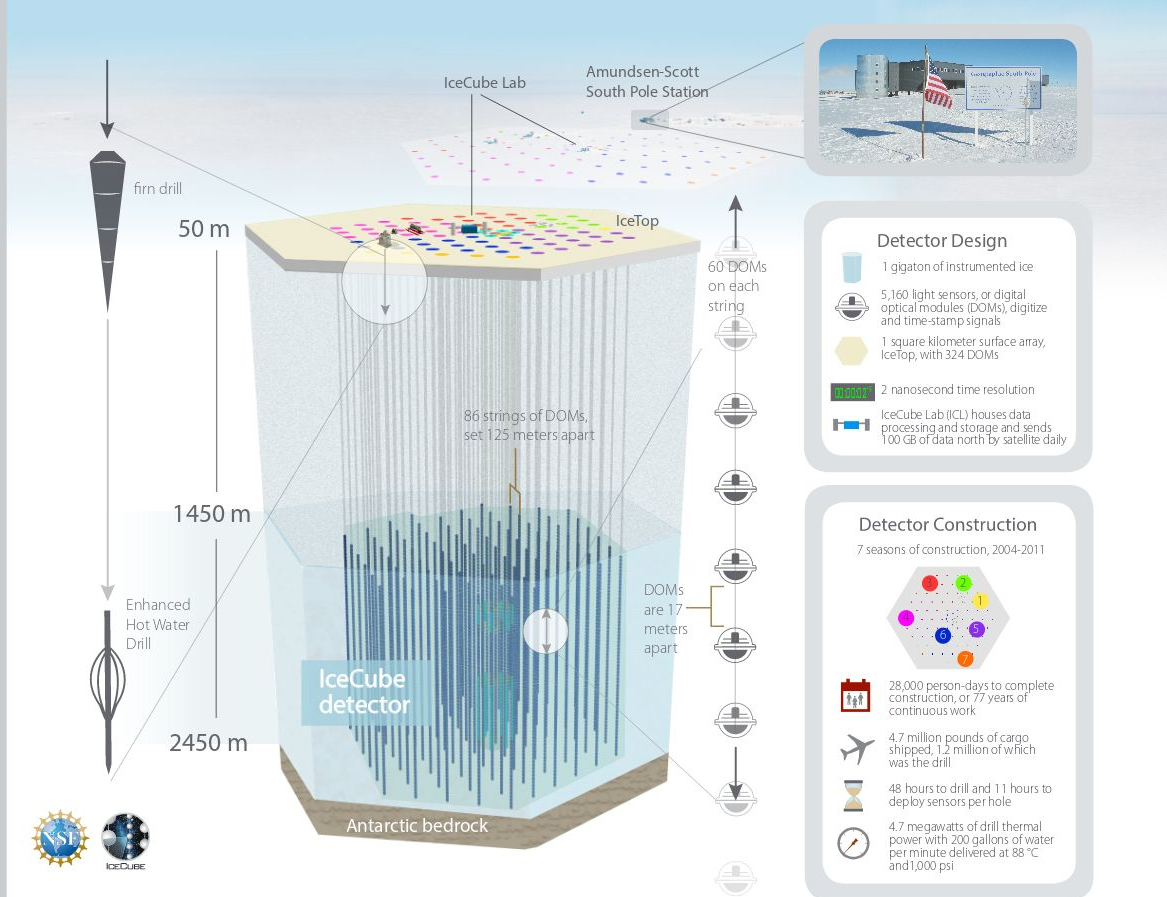
The Swedish Research Council (VR)  
University of Wisconsin Alumni Research Foundation (WARF)  
US National Science Foundation (NSF)

**ICECUBE**  
SOUTH POLE NEUTRINO OBSERVATORY

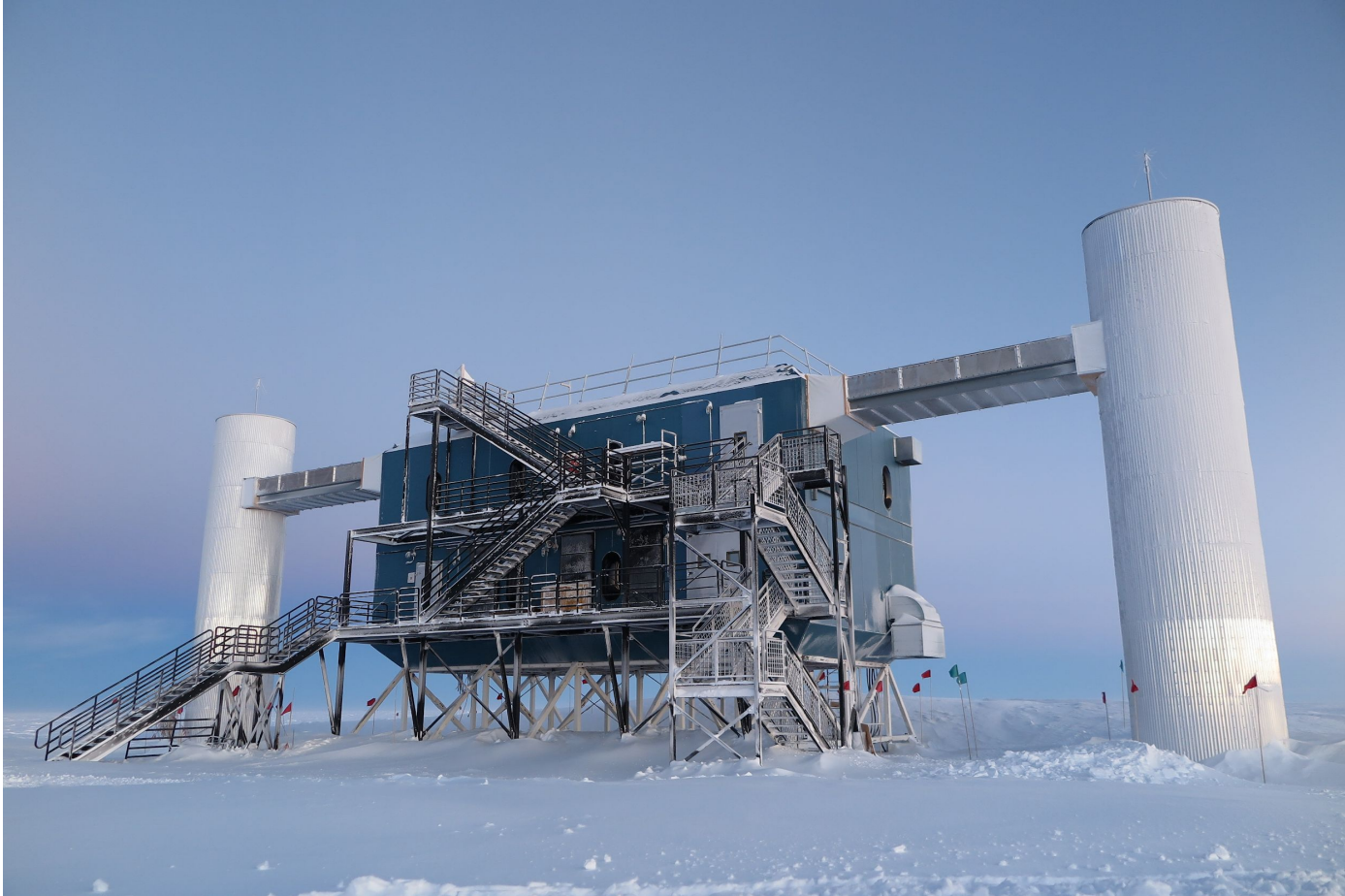
icecube.wisc.edu

# IceCube Neutrino Observatory

## The IceCube Neutrino Observatory Design and construction

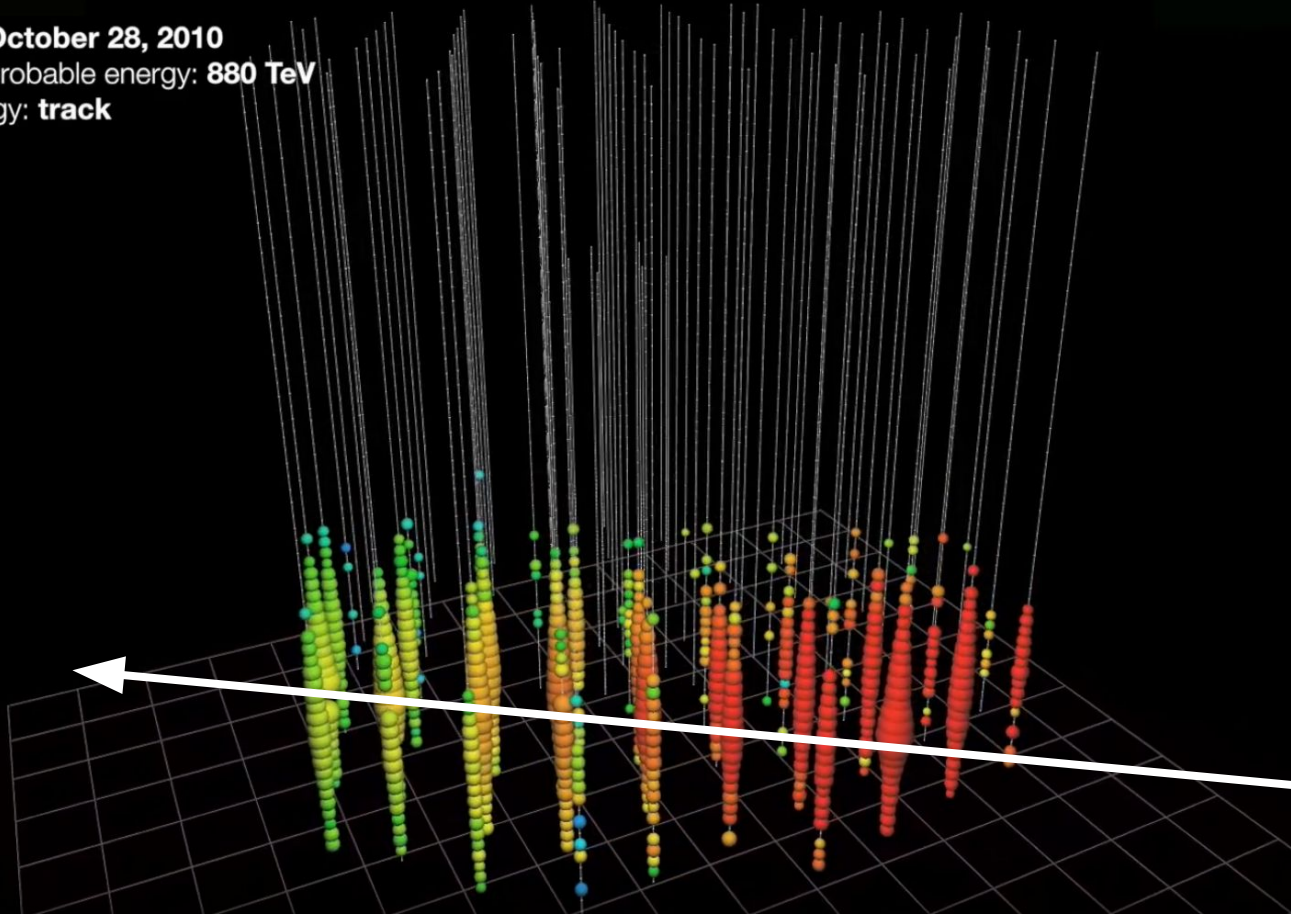


# IceCube Neutrino Observatory



# IceCube Neutrino Observatory

date: **October 28, 2010**  
most probable energy: **880 TeV**  
topology: **track**



# What is IceProd

---

## Data provenance

- Configuration - which software, what versions, when/where it ran, ...

## Dataset submission

- Monitor job status, resource usage
- Retry failed jobs - resubmit with different requirements

# Supercomputer Challenges

## Networking

- No external access from worker nodes
- Special “data mover” nodes separate from submit nodes

## Solution:

- Stage data using “data mover” nodes
  - Automated to happen before job submit, after job completion
- Edit job config to point to local paths before submission



# Supercomputer Challenges

---

## OS support

- Special flavors of operating systems not compatible with existing builds
  - Some cases where OS reports it is RHEL 7, but all libraries are different versions

## Solution:

- Containers!
- Run everything in the standard OSG WN container

# Supercomputer Challenges

---

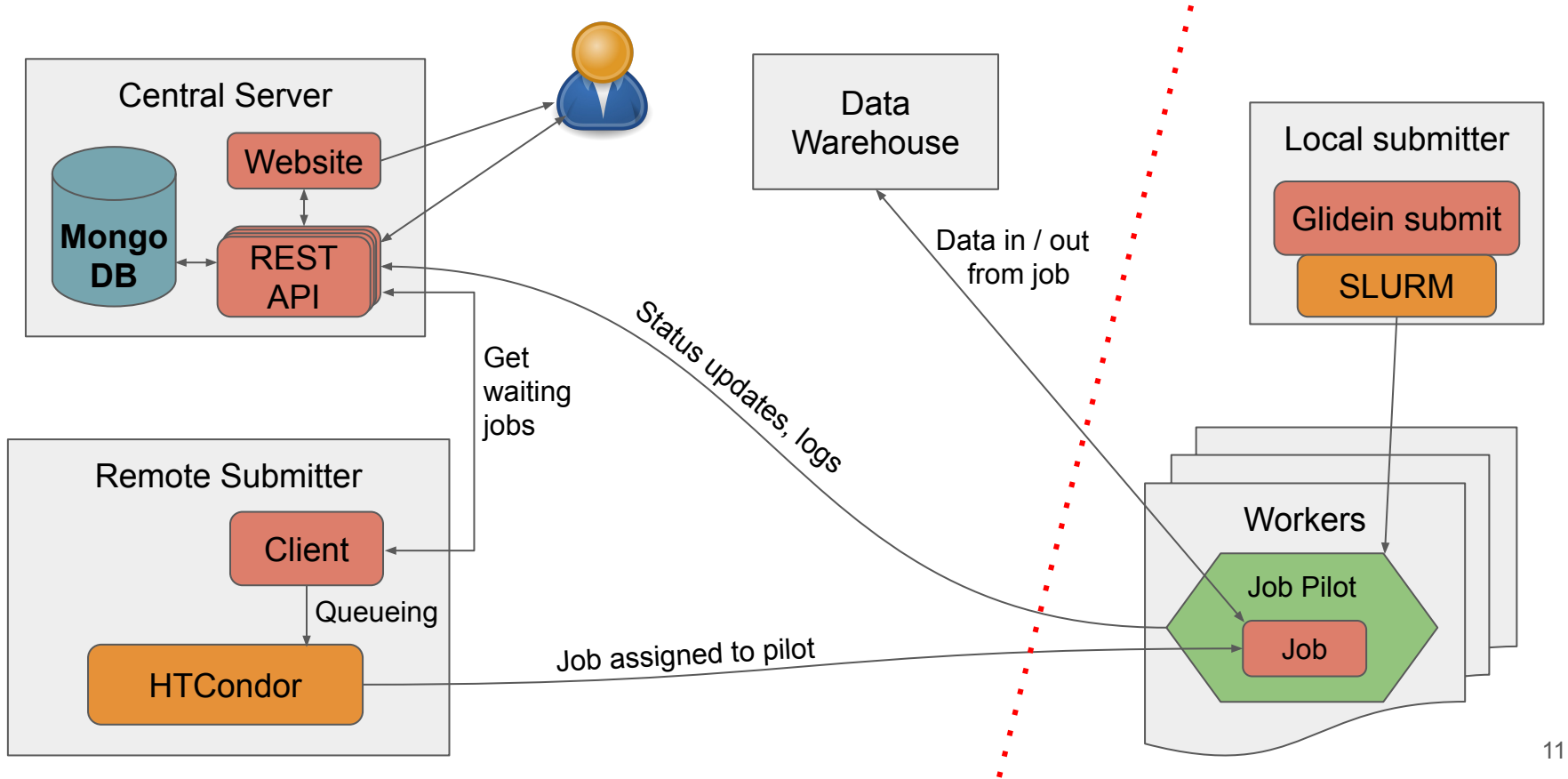
## CVMFS support

- Some sites still don't support CVMFS

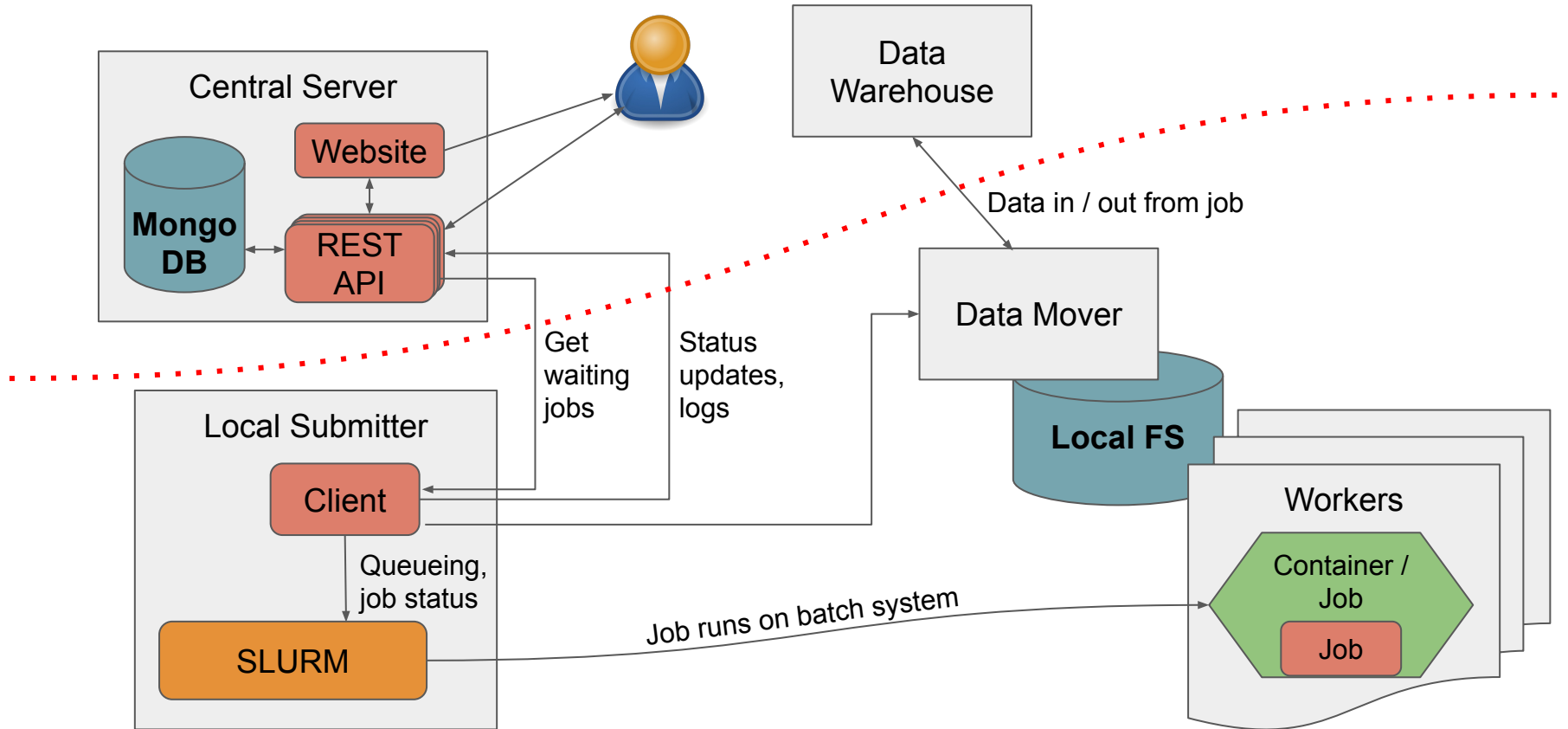
## Solution:

- Containers again!
- Rsync CVMFS nightly to local FS, then bind mount into container

# Typical Production Architecture



# Supercomputer Architecture



# Current Sites

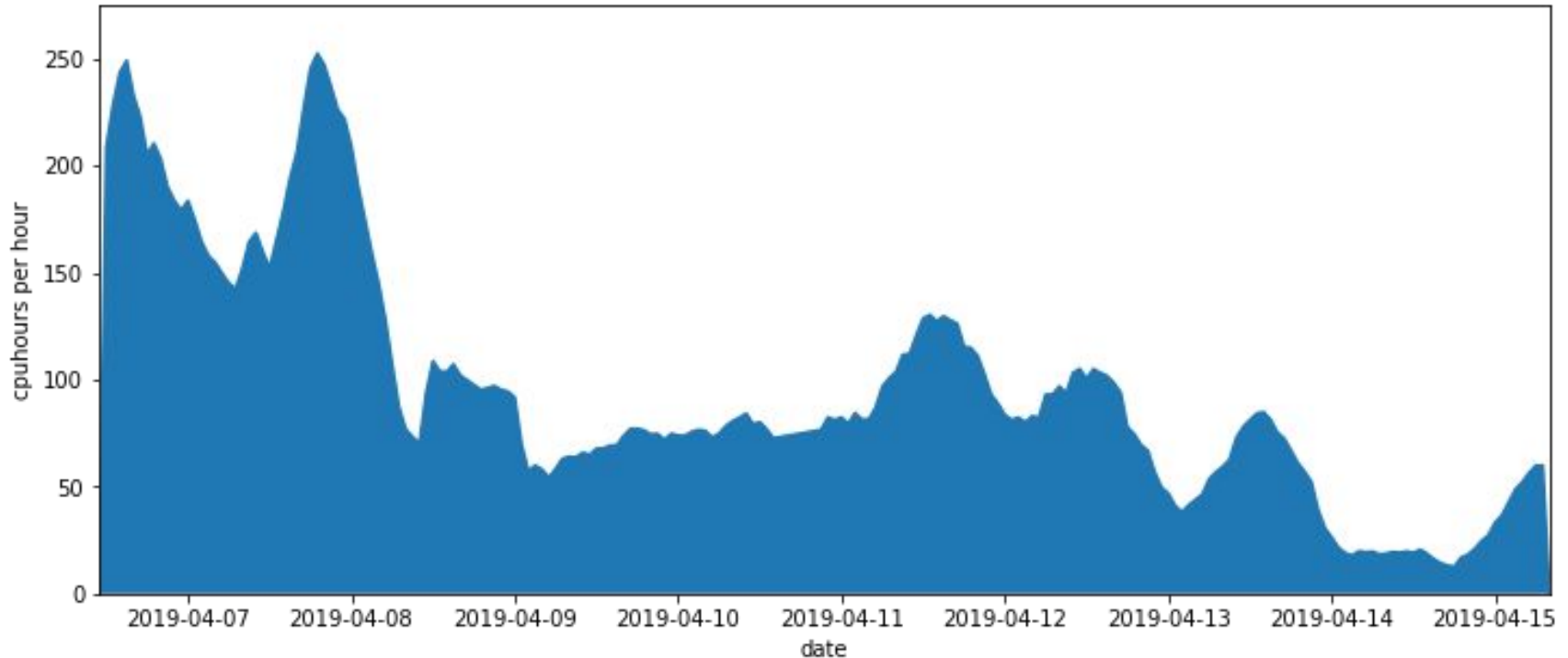
---

## Graham in Compute Canada

- Part of IceCube's Compute Canada allocation
- “Small” share, about 100 CPUs continuously
- Running since April 2019

# Current Sites

## Graham in Compute Canada



# Current Sites

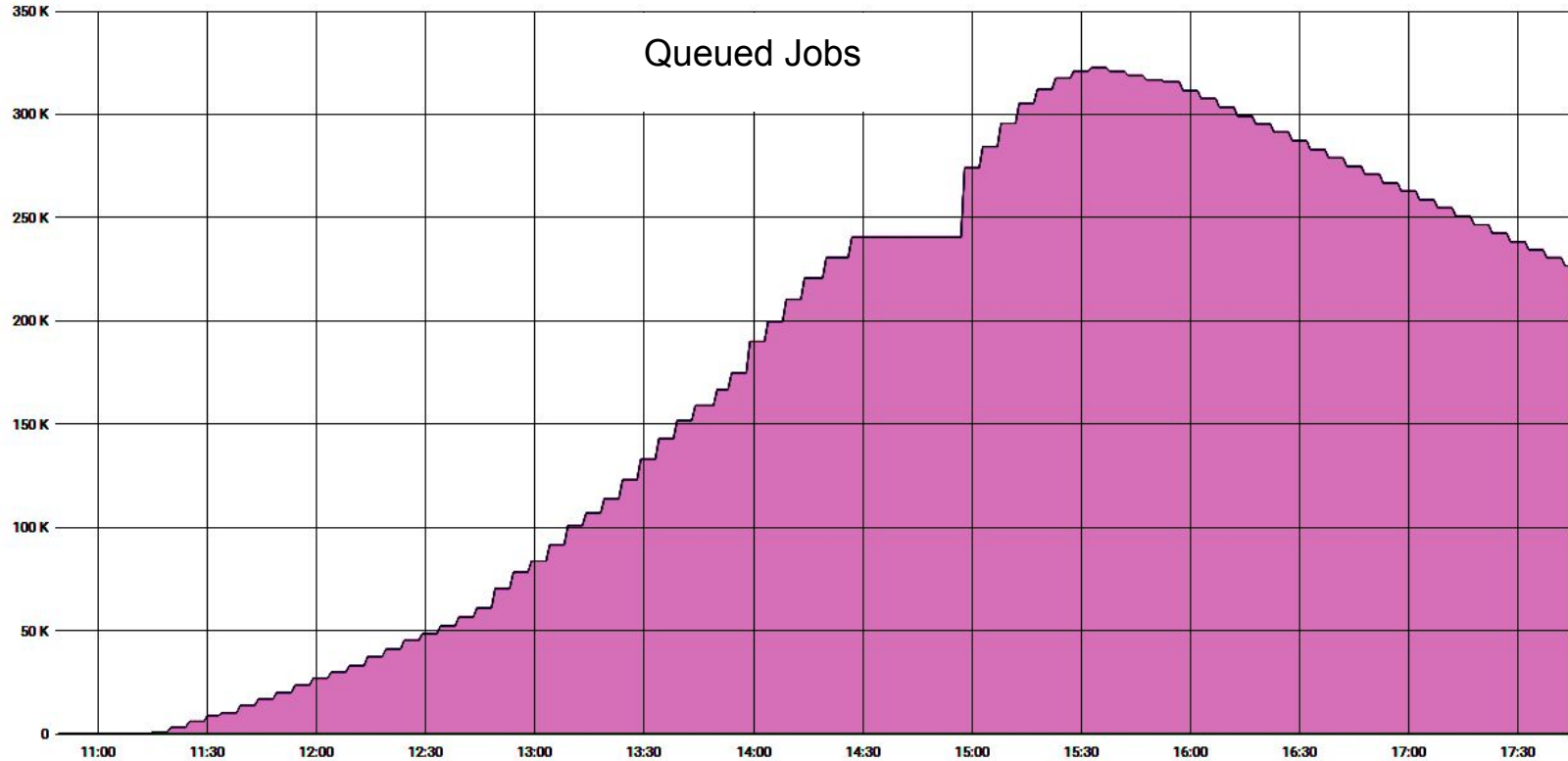
---

## Exascale Cloud Demo

- NSF EAGER grant with SDSC, Internet2
- Using 3 major cloud providers: AWS, Google, Azure
- Goal: get as close to an exaflop as possible in a single condor pool, on cloud resources today
- Like a supercomputer:
  - No direct I/O – our servers can't sustain that rate
  - No pilots – infrastructure can't handle the load

# Current Sites

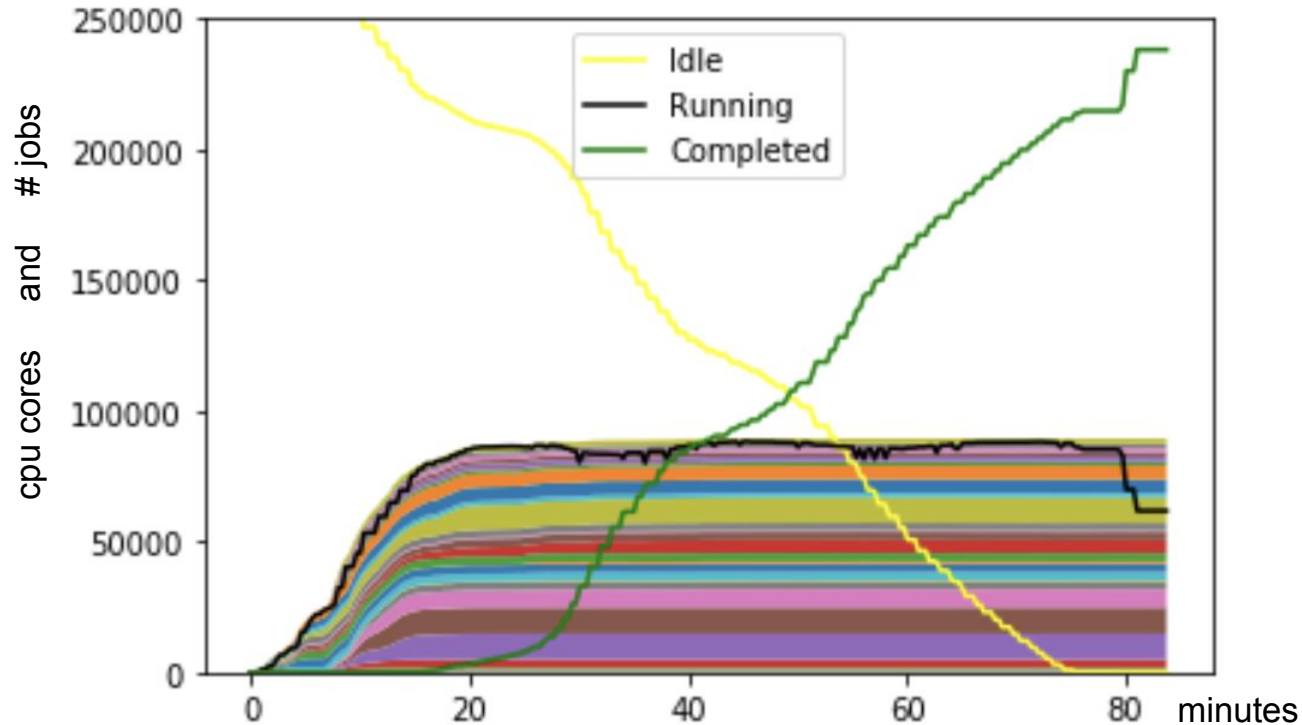
## Exascale Cloud Demo





# Current Sites

## Exascale Cloud Demo



# Future Supercomputer Challenges

## Queueing

- Some sites have policies preferring large MPI jobs over many single-core jobs, or other special things
  - One MPI job can expand to the entire size of the supercomputer
  - Queue thus allows very few jobs per user

## Still working on this

- One idea we've tested: make a local HTCondor pool out of large, long-running jobs (whole machine, 24 hours), and submit physics jobs to that pool

# Conclusions

- IceCube primarily relies on a glidein model with worker node internet access
- Supercomputers disallow this, so a direct submission mode was created
- Works well so far, more sites to be added in the future