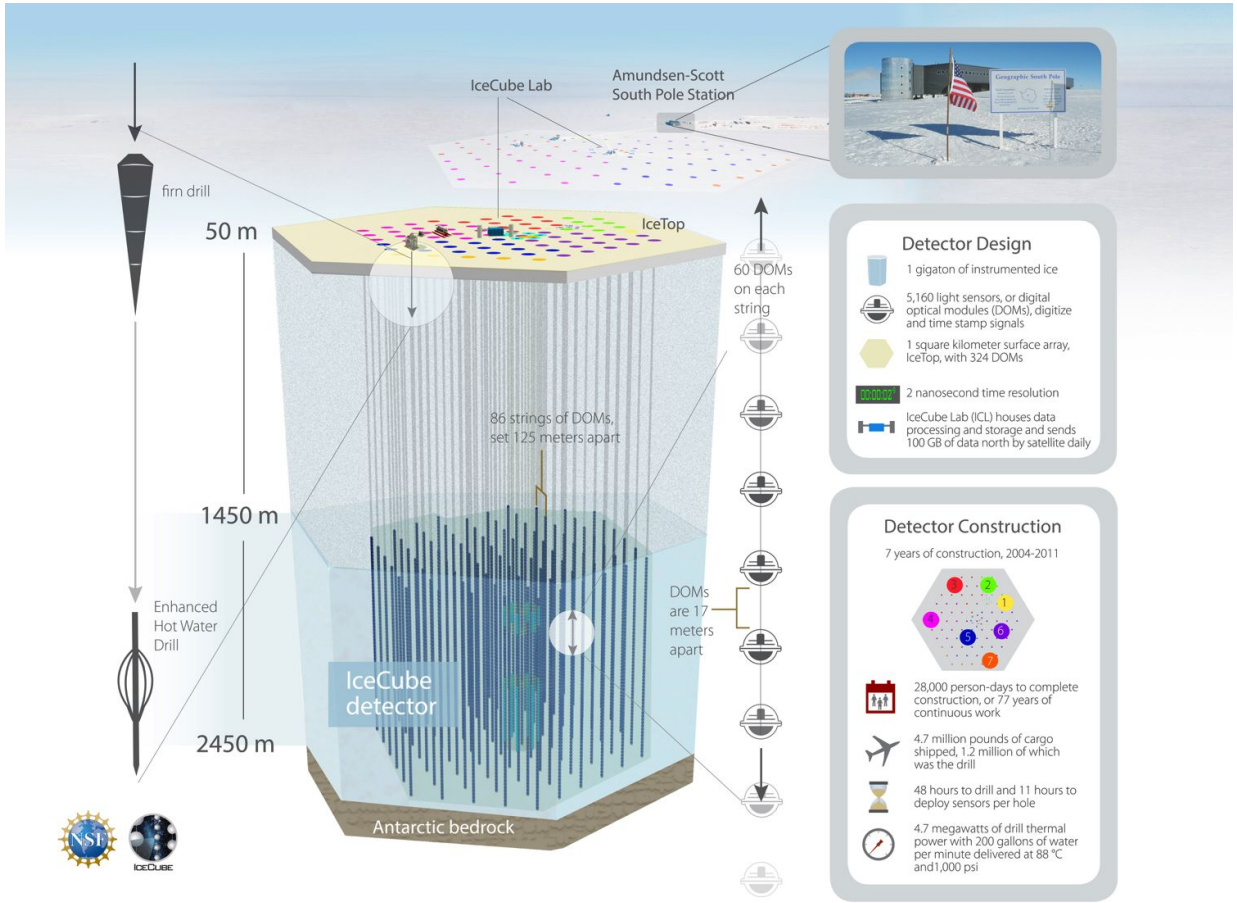
Distributed Computing in IceCube

The IceCube Neutrino Observatory



5160 Photomultipliers instrumenting 1km^3 of antarctic ice

Construction from 2004 to 2011 - huge engineering challenge

… since then, working like a swiss clock

Taking data 24/7

>**99%** uptime

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
Friedrich-Alexander-Universität Erlangen-Nürnberg
Humboldt–Universität zu Berlin
Ruhr-Universität Bochum
RWTH Aachen
Technische Universität Dortmund
Technische Universität München
Universität Münster
Universität Mainz
Universität Wuppertal

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 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

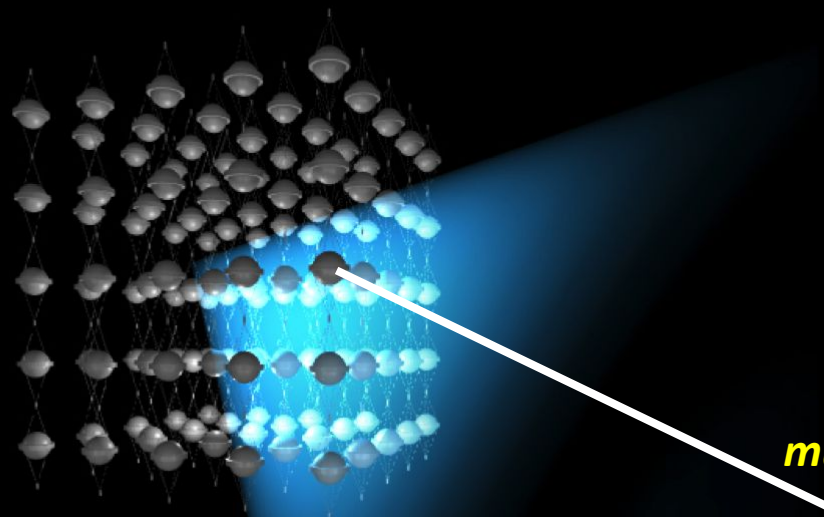
FUNDING AGENCIES

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

When a neutrino hits a water nucleus head-on it generates a nuclear reaction.

The emitted charged particles produce light.



muon

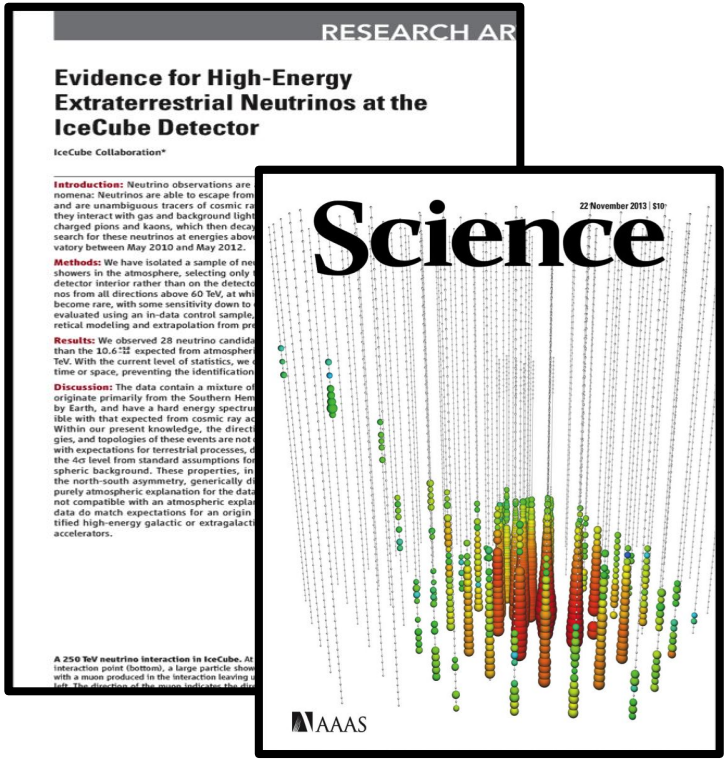
interaction

neutrino

Dark and transparent media

Lots of light sensors

Nov 2013 - astrophysical neutrinos discovery



IceCube trigger ~ 3000 kHz, every year:

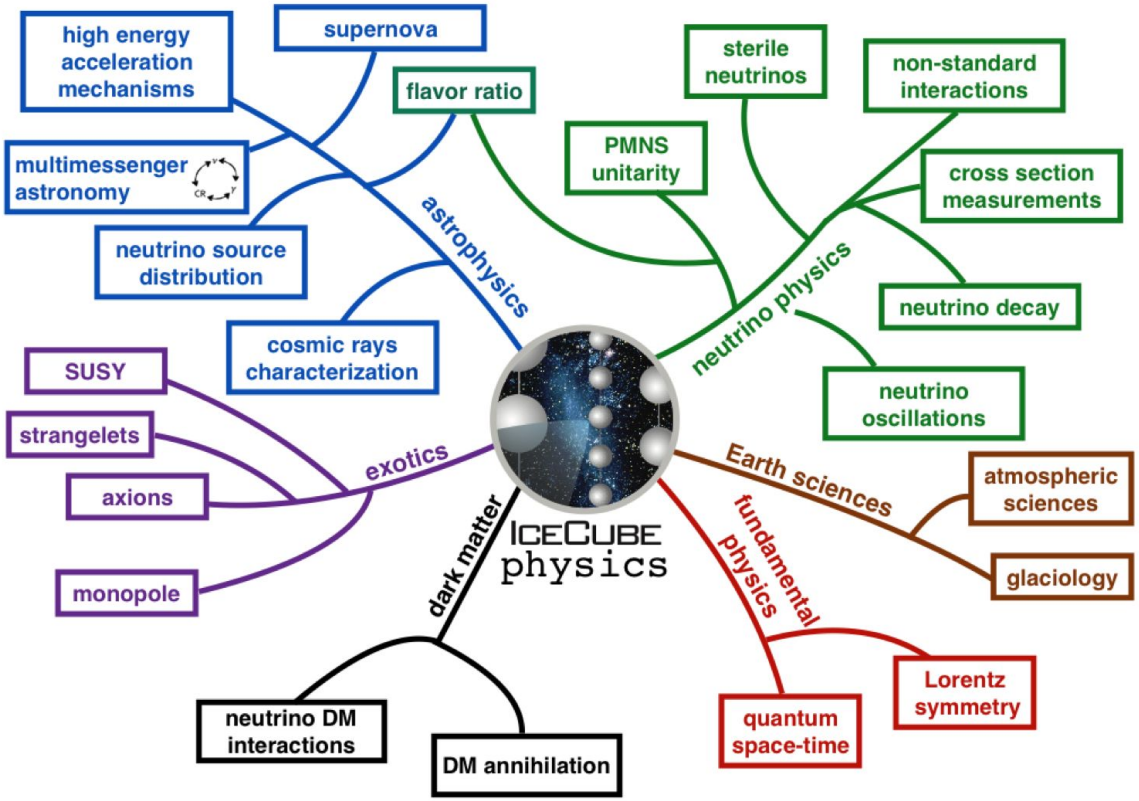
- ~ 100 billion (mostly) background atmospheric muons
- ~ 100,000 atmospheric neutrinos
- ~ 10-15 astrophysical neutrinos

Neutrino signal events need to be distinguished from a background of downgoing atmospheric muons based on the pattern of emitted Cherenkov light.

With ~ 7 years of data taken, IceCube is transitioning from discovery to precision measurement phase.

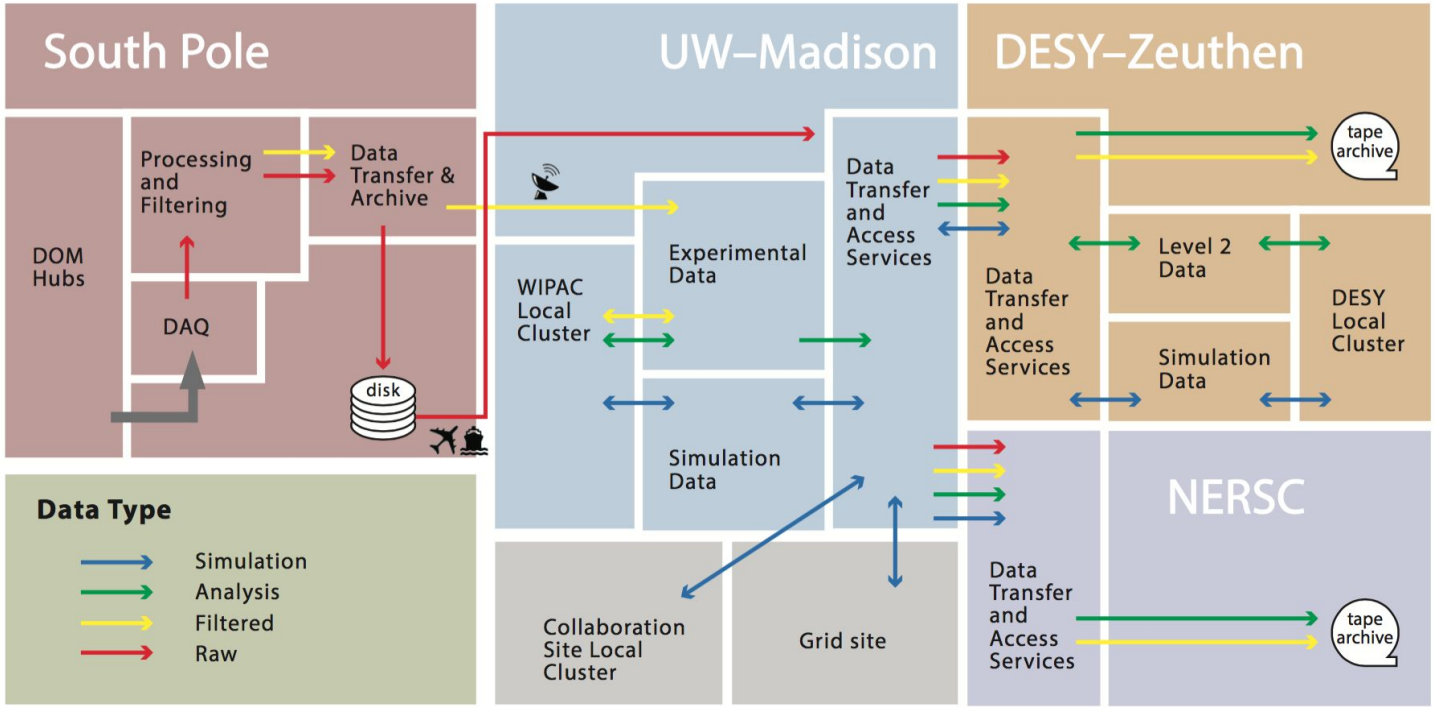
- Understanding of systematics is key
- An important one: **light propagation** through the km^3 antarctic ice block

A broad physics program



C.Argüelles

IceCube Data Flow



Data archival policies

Data type	Subtype	Growth (TB/yr)	DESY-ZN tape	NERSC tape
Experimental	Raw	300		yes
	SNraw	30		yes
	Ancillary	5		yes
	SuperDST	70	yes	yes
	Filtered	40	yes	yes
	Level2	100	yes	yes
	Level3	100		yes
Simulation	Level2	400		
	Level3	100		yes
	Photon tables	10		

~700 TB/year to NERSC archive
~200 TB/year to DESY archive

DESY and NERSC Tier1 centers

- pledge automated tape resources for long term archive
- MoU in place

IceCube distributed computing - history

Pre-2014 setup:

- HTCondor-flock to UW clusters and GLOW VOFrontend (glideinWMS infrastructure)
- IceCube simulation framework doing local submissions at ~20 sites

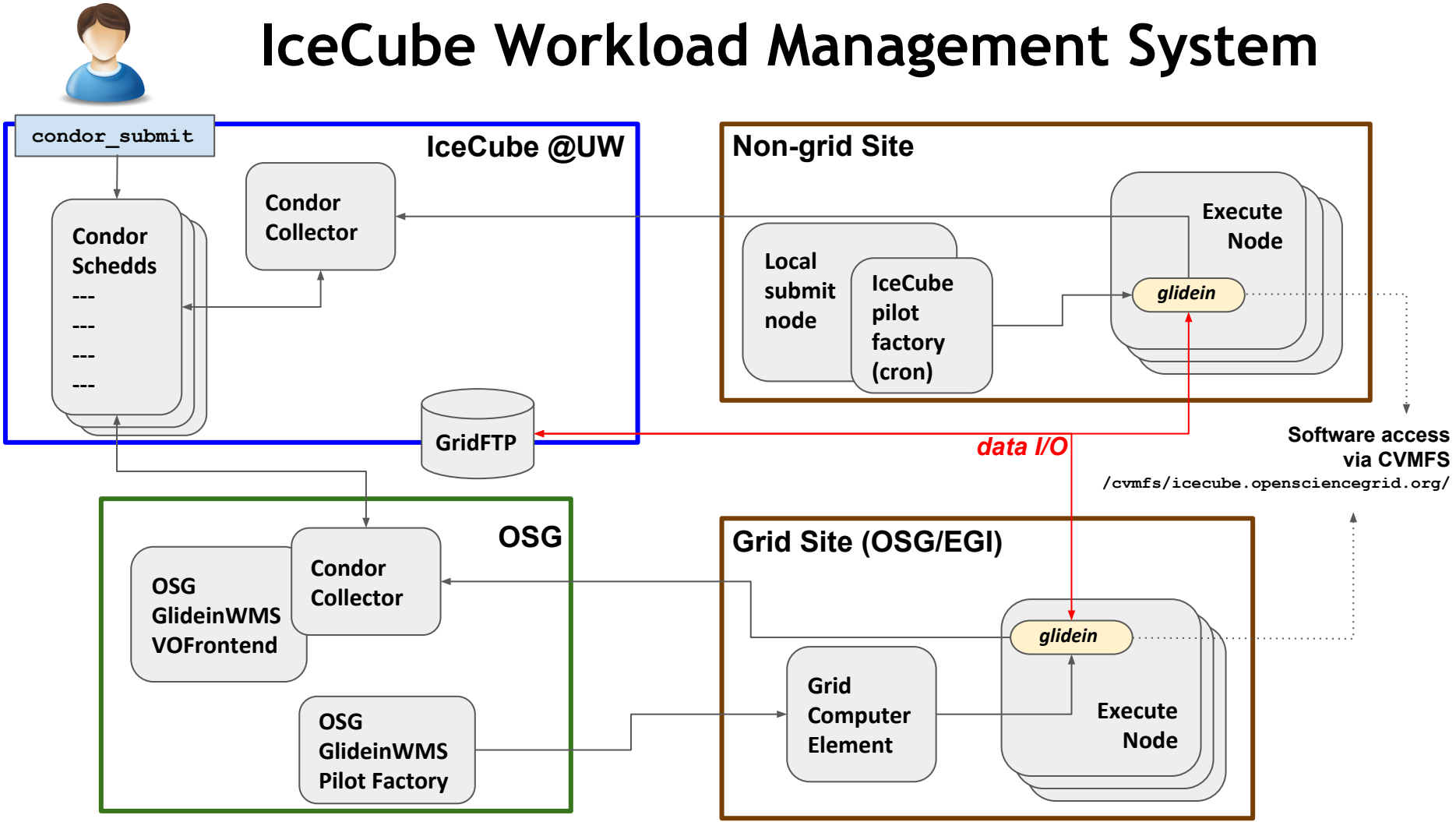
2014 - 2015

- Icecube VO configured in VOFrontend ⇒ some EGI and CA sites via OSG glideins
- IceCube simulation framework doing local submissions at ~10 sites

2016 - now

- Keep using OSG/glideinWMS and Icecube VO for “Grid enabled” sites
- Lightweight HTCondor glidein system for non-Grid sites (codenamed “pyglidein”)

IceCube Workload Management System



Grid sites - access via the Icecube VO and glideinWMS

IceCube Sites

DESY
Dortmund
Aachen
Wuppertal

CA-Toronto
CA-McGill
Manchester
Brussels

Notable opportunistic OSG Sites
(thanks!)

Fermilab
Nebraska
CIT_CMS_T2
SU-OG
MWT2
BNL-ATLAS

Non-Grid sites - access via pyglidein

IceCube Sites

DESY
Mainz
Dortmund
Brussels
Uppsala

CA-Toronto
CA-Alberta
CA-McGill
UMD
MSU
Delaware
Tokyo

XSEDE

Comet
Bridges
XStream

CVMFS

/cvmfs/icecube.opensciencegrid.org/

- Started in Aug 2014
- Using OSG/EGI Stratum1s since ~ Oct 2014

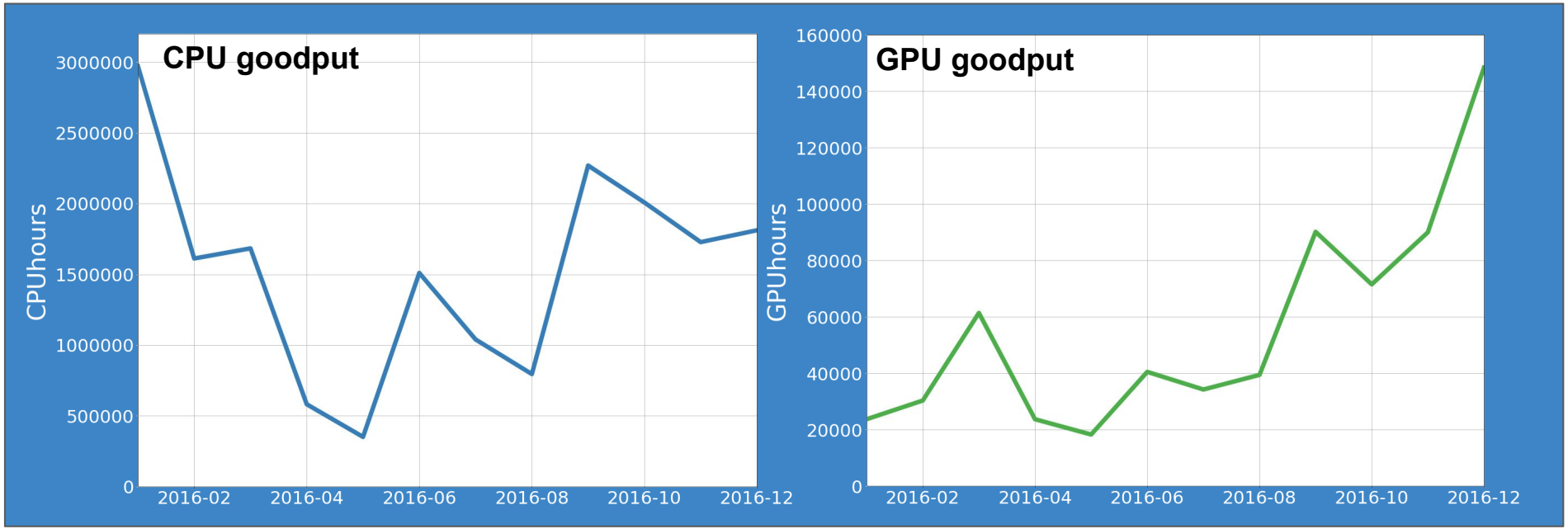
Stats

- Total file size: 300 GB
- Spool size: 45 GB
- Num. files: 2.9 M

Yearly growth

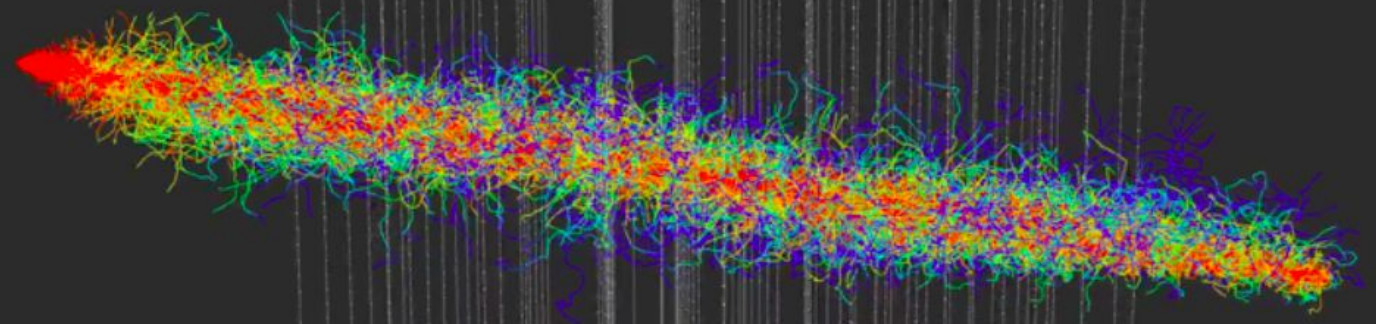
- Total file size: 120 GB
- Spool size: 10 GB
- Num. files: 1.2 M

Grid Usage Totals



CPU: 18.3 Mhours GPU 650 Khours Badput ~20%

GPUs: direct photon propagation



GPUs are ideal for the workload

- One muon generates >10^7 detectable photons - billions of muons in the data
- scattering model is simple (scatter, absorb, change ice layer or hit a sensor)
- Simulate each photon with an independent thread
- Only interrupt parallelism when a photon hits a sensor and signal needs to be stored (very rare!)

GPUs are O(~100) faster than CPUs

IceCube GPU Cluster

Good news: photon propagation code is single precision

⇒ consumer-grade GPUs are OK

Not so good: GPUs still a rare beast, not easy to find accessible GPU clusters out there.

⇒ needed to build an in-house sizeable cluster.

Current IceCube GPU cluster at UW-Madison:

- 32 AMD 7970
- 32 Nvidia GeForce GTX 690
- 256 Nvidia GeForce GTX 980
- 56 Nvidia GeForce GTX 1080

Supercomputers

2016: XSEDE “research” allocation awarded in 2 GPU-enabled systems:

Comet at SDSC: 5,543,895 SUs allocated

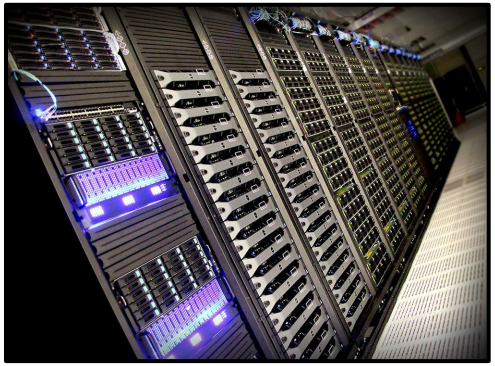
36 nodes with 2x NVIDIA K80 each

Bridges at PSC: 512,665 SUs allocated

16 nodes with 2x NVIDIA K80 GPUs each
+32 nodes with 2x NVIDIA P100 GPUs each

XStream: 5,000 SUs test allocation

65 nodes with 8x NVIDIA K80 each



2017: Plan to renew XSEDE allocation - target XStream, largest XSEDE GPU system

- Also considering DOE supercomputers - e.g. TITAN: LOTS! of GPUs

DESY Tier1 centre

18-May-2015

Statement of Work
DESY
1 October 2015 – 31 December 2019

IceCube Maintenance and Operations

This amendment is to exhibit A of the Memorandum of Understanding for IceCube Maintenance and Operations effective January 1, 2015 between the Institutions of the IceCube Collaboration and the Board of Regents of the University of Wisconsin System.

Tier 1 datacenter at DESY

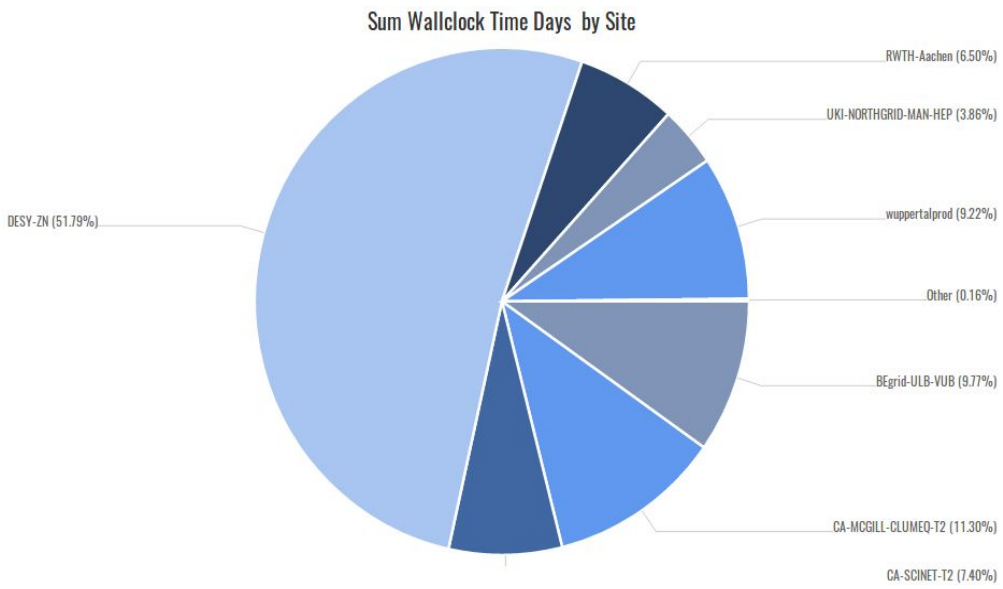
DESY agrees to act as a Tier-1 datacenter for the IceCube collaboration. The services DESY will provide to the IceCube collaboration are described in detail in the following paragraphs:

DESY Tier1 centre (cont.)

Resources (Grid and non-Grid)

CPUs	20kHS06
GPUs	80 (nVidia Kepler)
Storage (dCache)	1.4PB

EGI grid computing last year (per grid site)



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
  - Learning the specifics of federating GPU clusters
  - Also turning our view to tapping on Supercomputer resources.