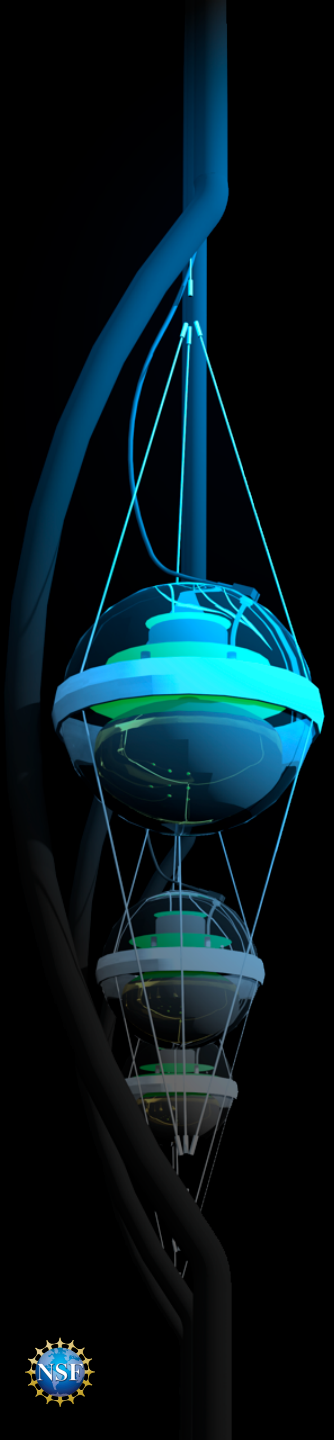
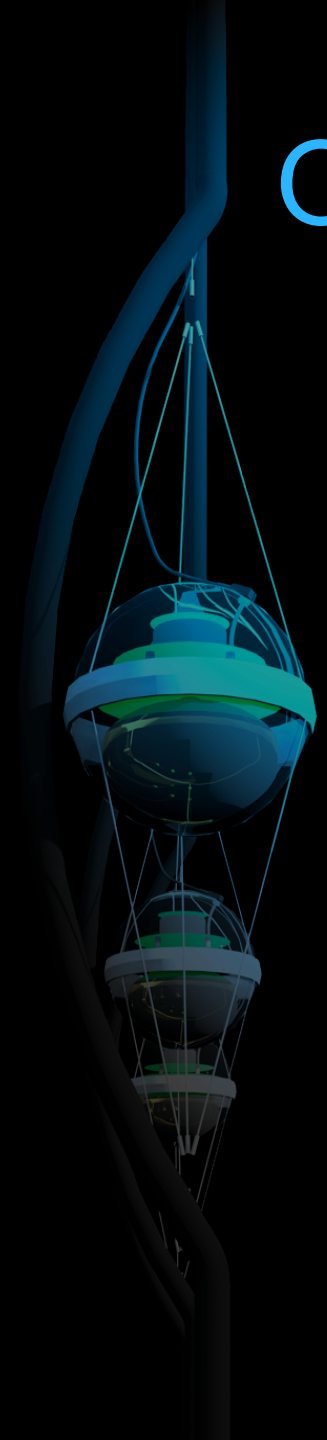


IceCube, HTCondor, and GPUs

Steve Barnet

University of Wisconsin – Madison

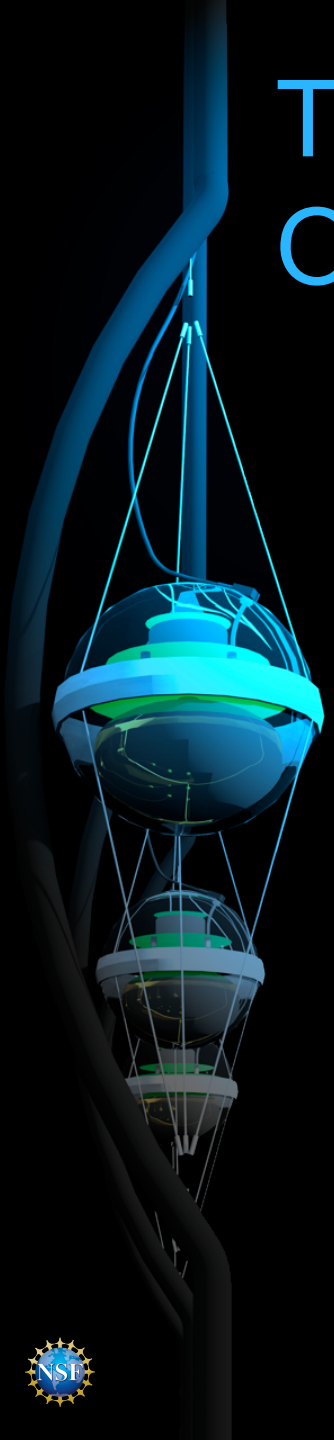
Condor Week 2014



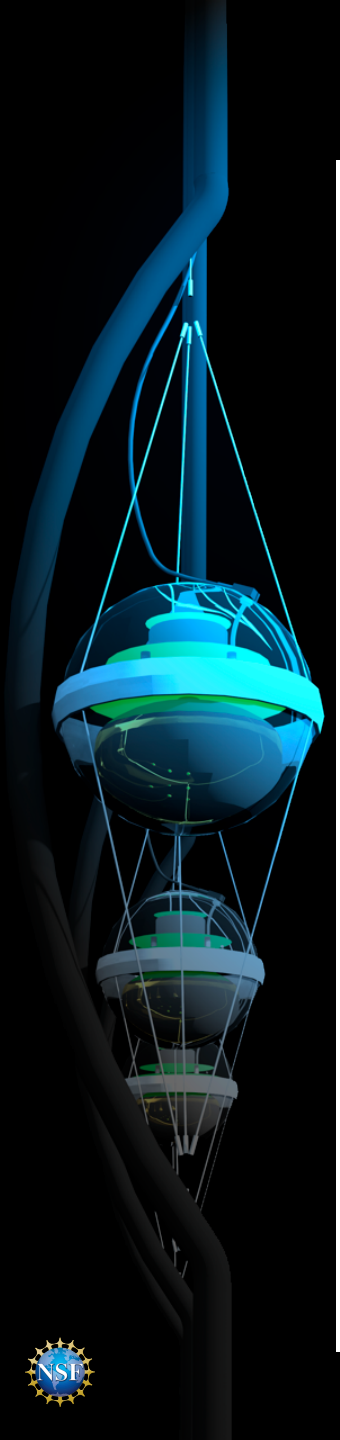
Overview

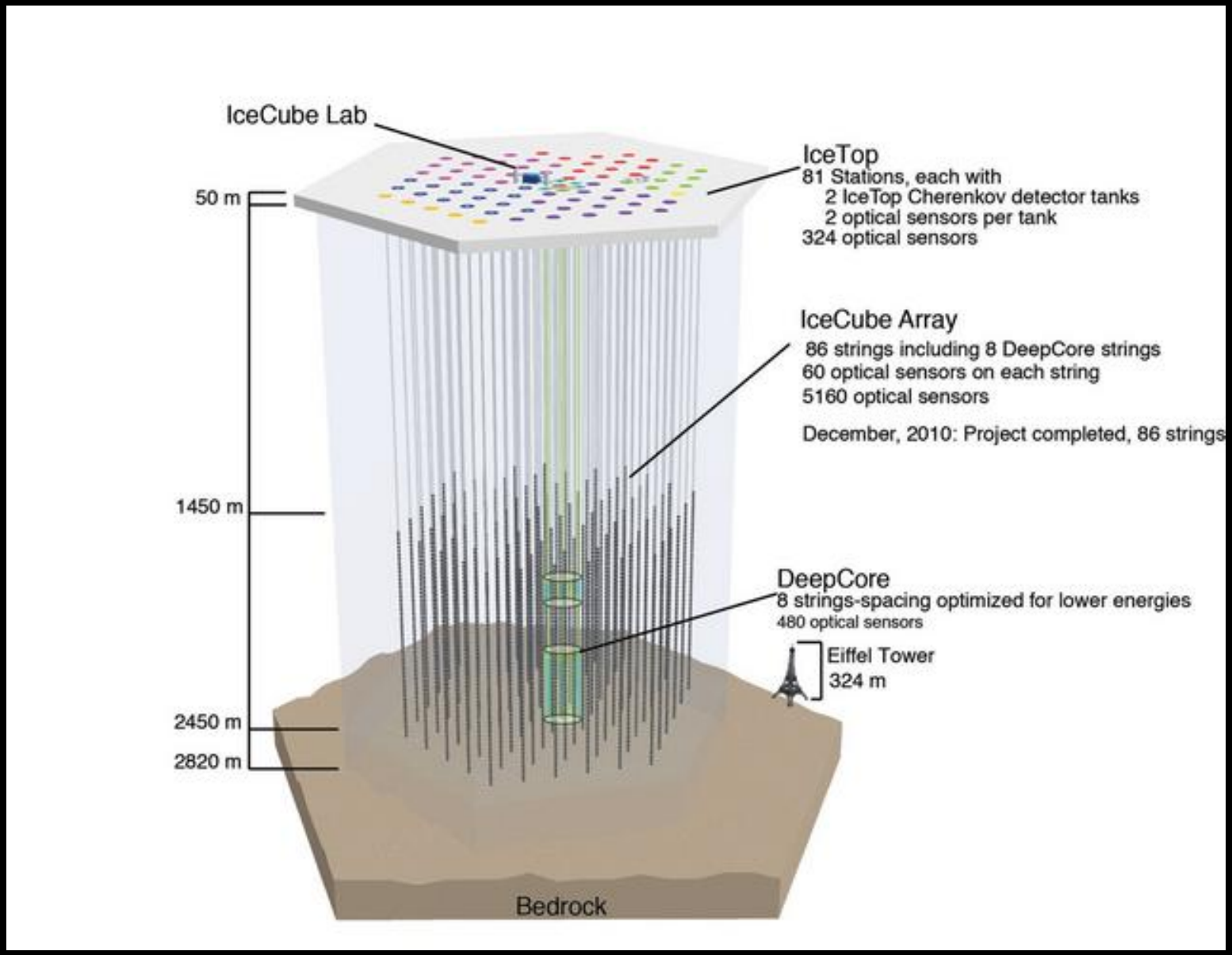
- The Basics of IceCube
- GPUs – The Motivation
- Evolution
- Summary

The IceCube Neutrino Observatory

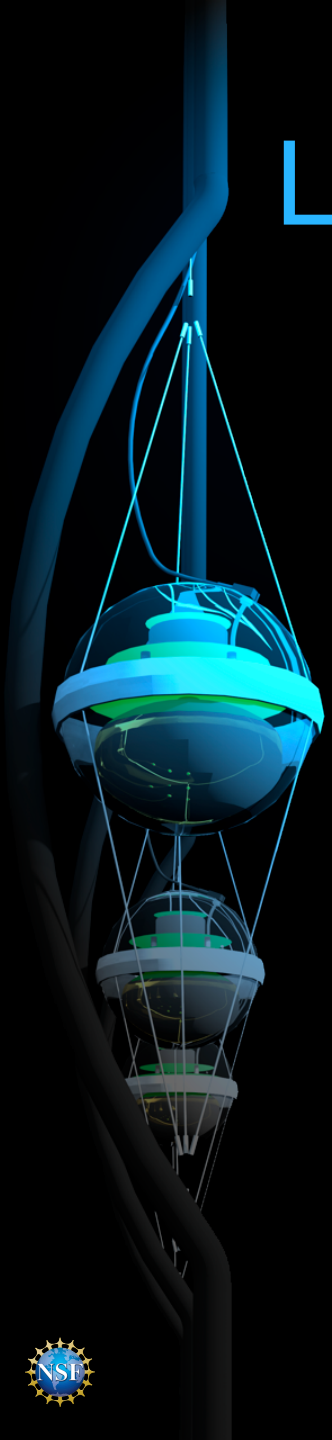


- A kilometer scale neutrino detector
- Located at geographic South Pole
- Detects Cherenkov light from neutrino interactions



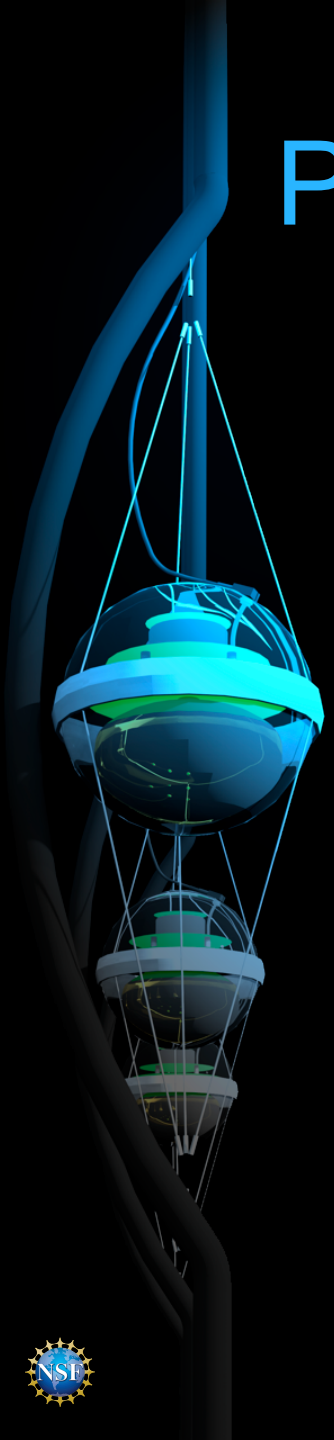

IceCube Lab
IceTop
81 Stations, each with
2 IceTop Cherenkov detector tanks
2 optical sensors per tank
324 optical sensors
50 m
IceCube Array
86 strings including 8 DeepCore strings
60 optical sensors on each string
5160 optical sensors
December, 2010: Project completed, 86 strings
1450 m
DeepCore
8 strings-spacing optimized for lower energies
480 optical sensors
Eiffel Tower
324 m
2450 m
2820 m
Bedrock

Location, Location, Location



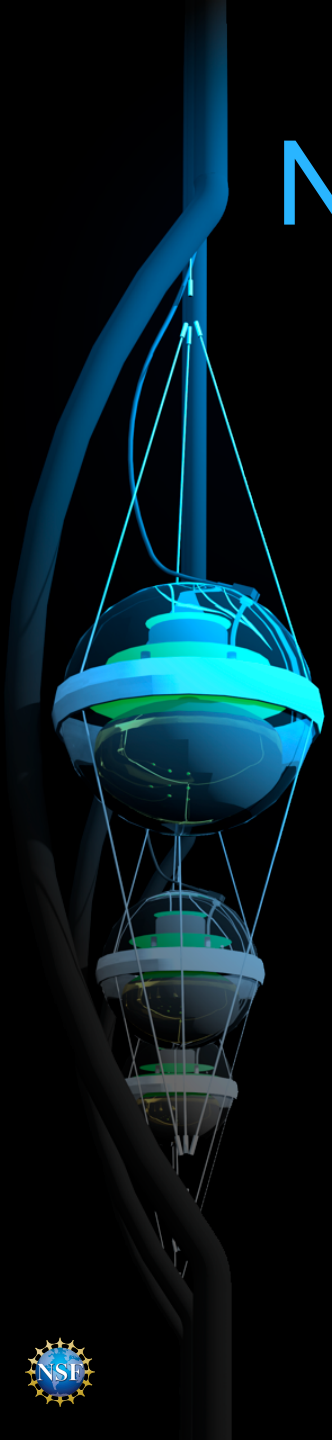
- Why the South Pole?
- Lots of ice – a great detection medium
- The ice is very clear
- Thick ice sheet – sensors deep enough to provide significant background reduction

Propagation of Light

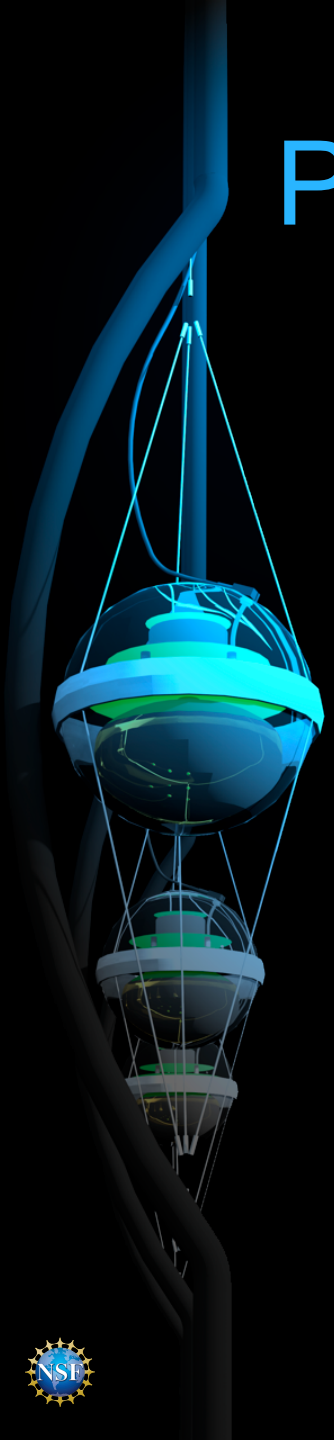


- A particle creates light when it interacts with the ice
- At what time did light hit a module?
- How much light hit the module?
- Shows us the track of the particle
- Provides information about the energy of the particle
- Clues as to the type of particle

Need to Model the Ice

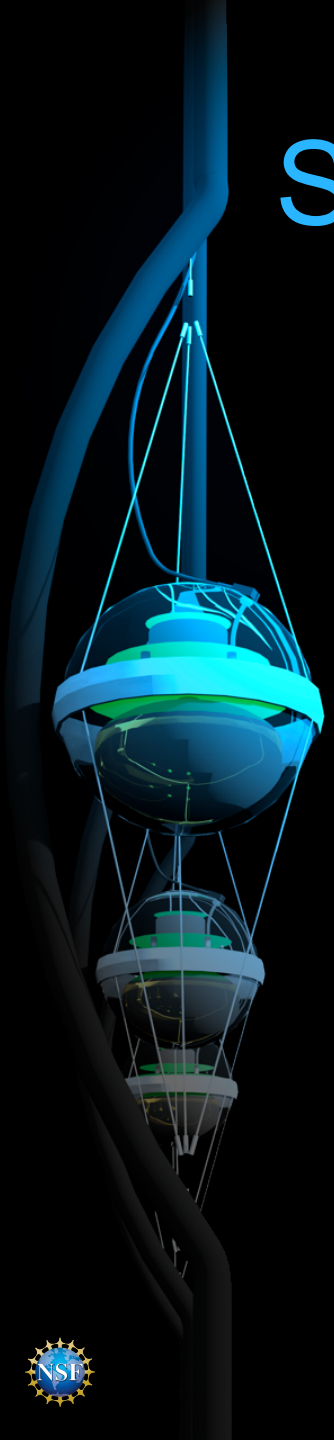


- Given an interaction point, how does the light propagate through the ice?
- Ice scatters and absorbs light, so do a bit of math – voila!
- Actually, rather a lot of math (or a little math a lot of times)
- Early efforts (starting 15 years ago) – approximate and create lookup tables



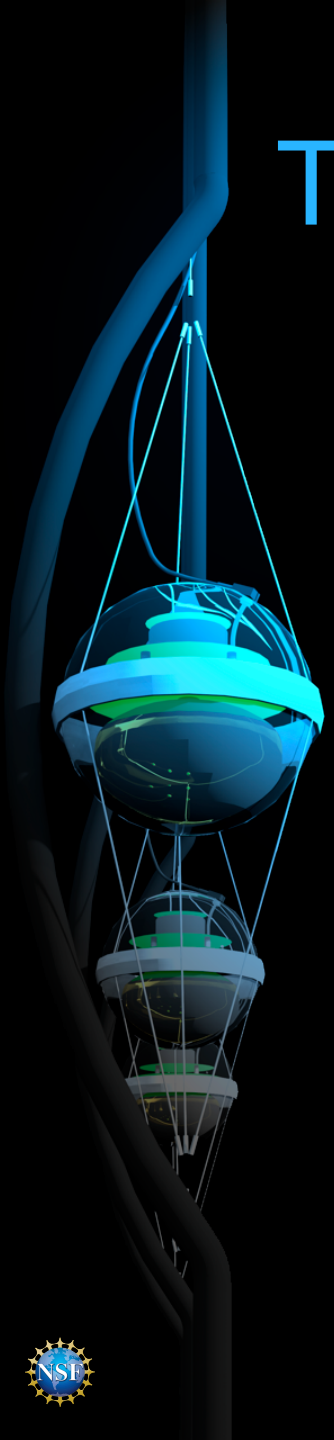
Problems

- The ice has structure
  - Dust layers
  - Tilt
  - Each parameter adds complexity/size to the tables
- Tables introduce (large) systematic errors



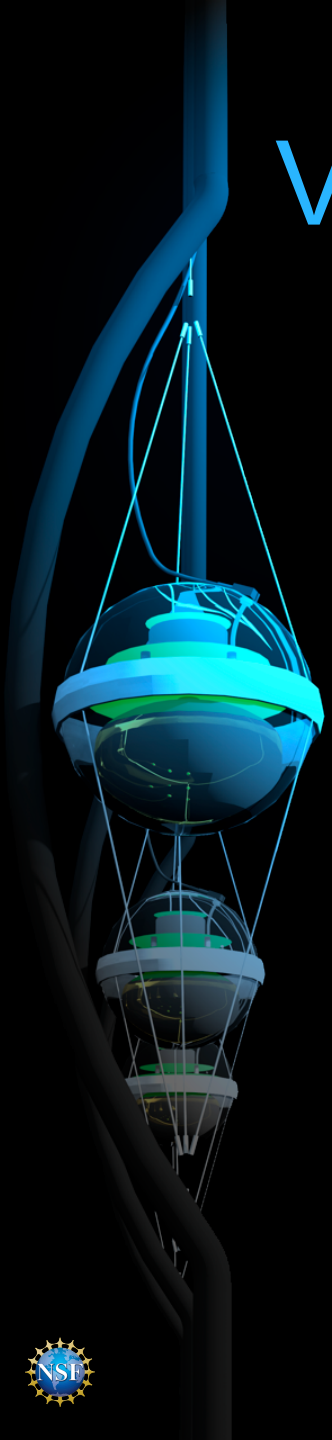
Solution – Direct Propagation

- Simulate the track of each photon through the ice
- Computationally intensive
  - Even a low energy event generates a few million photons
  - Event rate is ~ 3000 events/s
- Serious efforts in this direction began in 2009



The Story So Far

- Can we do this with CPUs?
  - Yes! Early code written in C++ ~ 2009
  - Also assembly (1.25x-.1.37x speed-up)
- Followed closely by GPU implementations
  - nVidia/CUDA based
  - 150x initial speed-up vs CPU versions

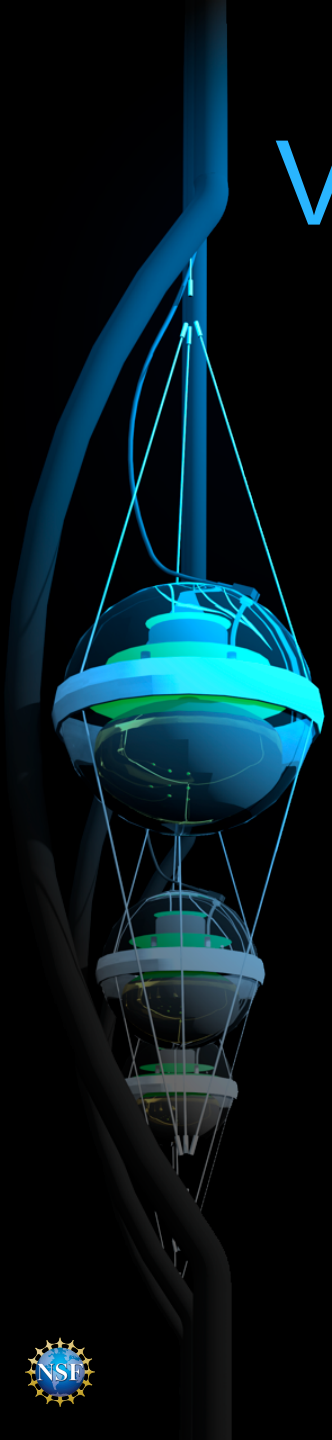


Version 1

- Initial system – consumer GPUs
  - Custom-built gamer desktops
  - Proof of concept
  - No HTCondor

Version 2

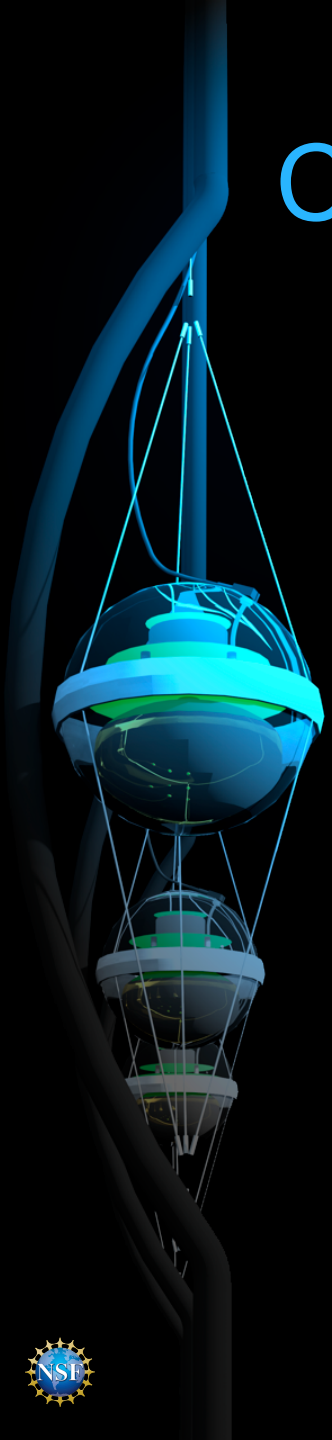
- GZK-9000 cluster
- Commercial GPUs
  - 48x nVidia M2070 GPUs
  - High density, increased scale
  - Hosted at CHTC
  - HTCondor managed



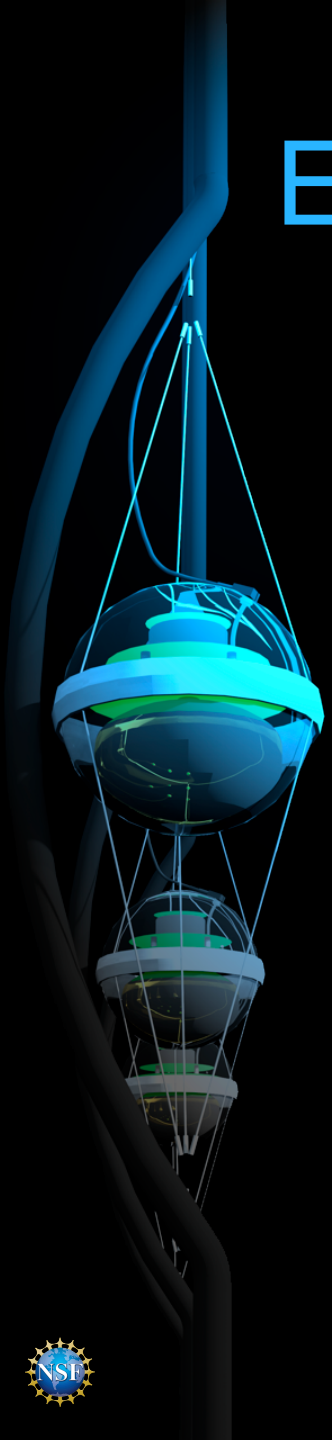
Version 3

- Consumer grade GPUs
  - Cheaper (*much* cheaper – 4x)
  - Split pool – 32x nVidia GTX-690, 32x ATI Radeon 7970
  - SuperMicro based – integrated by re-seller
  - HTCondor managed

Cheap GPUs – What you lose



- No Free Lunch …
- … or ECC either (maybe)
- … and limited Double Precision performance
- Detailed monitoring
  - Temp & Fan speed mostly
  - No direct utilization reporting



Experiences

- Getting these stable takes time
  - For us, true for commercial & consumer alike
- Another software stack to manage
  - GPU drivers
  - SDK & execution libraries
- Consumer card market extremely variable
  - Availability
  - Many card vendors (eVGA, Asus, XFX, PNY, …)
  - AMD (ATI) is not as "into it" as nVidia

HTCondor Config

- Start with static config
- STARTD config :

```
SLOT_TYPE_3 = cpus=1, mem=2000
NUM_SLOTS_TYPE_3 = 1
SLOT3_HAS_GPU = TRUE
SLOT3_GPU_NAME = "nVidia GeForce GTX 680"
SLOT3_GPU_DEV = 1
SLOT3_STARTD_ATTRS = HAS_GPU GPU_DEV GPU_NAME

USER_JOB_WRAPPER = /etc/condor/gpu_job_wrapper.sh
```

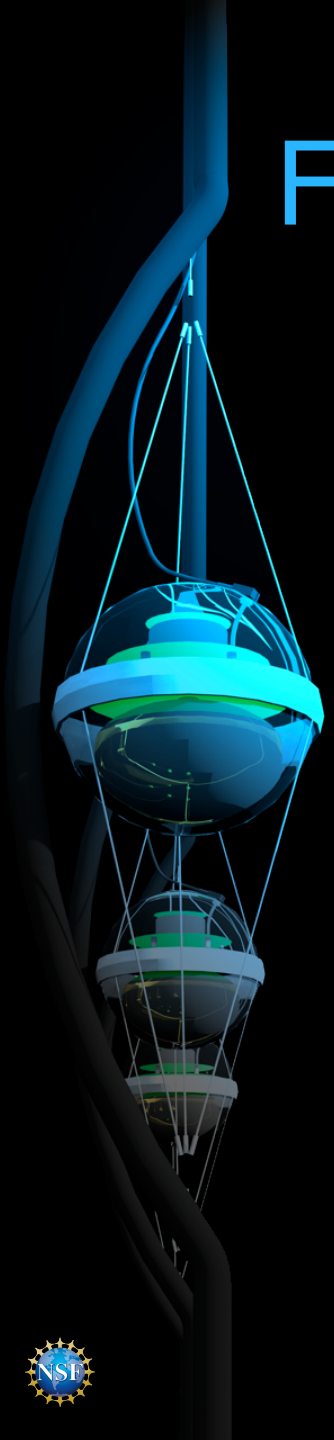
HTCondor Config

- STARTD cont'd:
- Wrapper script:

```
#!/bin/bash
gpu_dev=$(awk -F ' = ' '/^GPU_DEV = /{print
$2}' $_CONDOR_MACHINE_AD)
if [ -n "$gpu_dev" ]; then
    export CUDA_VISIBLE_DEVICES=$gpu_dev
    export COMPUTE=:0.$gpu_dev
    export GPU_DEVICE_ORDINAL=$gpu_dev
fi
exec "$@"
```

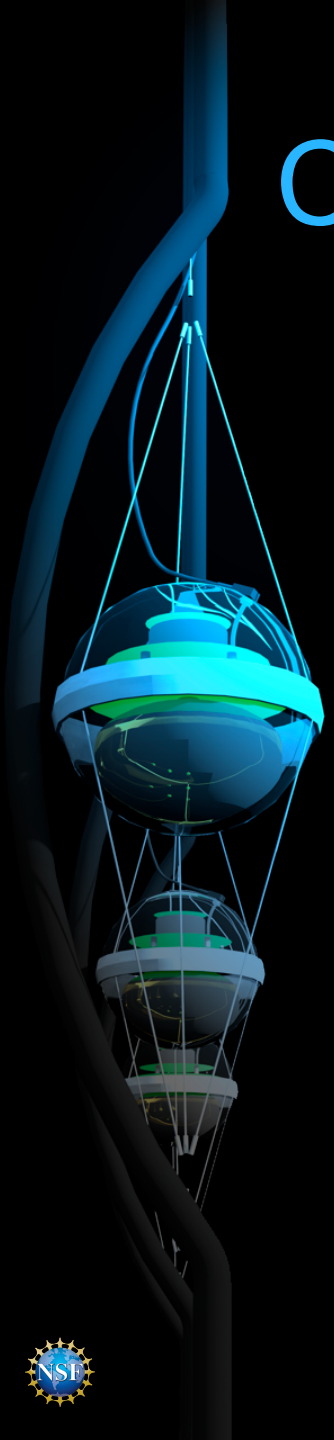
Submit side

- GPU allocation is managed with Accounting Groups
- A job requests the GPU accounting group
  - `+AccountingGroup="gpu.$ENV(USER)"`
- CUDA jobs must require CUDA:
  - `requirements = HAS_CUDA`
- OpenCL jobs can run anywhere, so no special requirements



Future

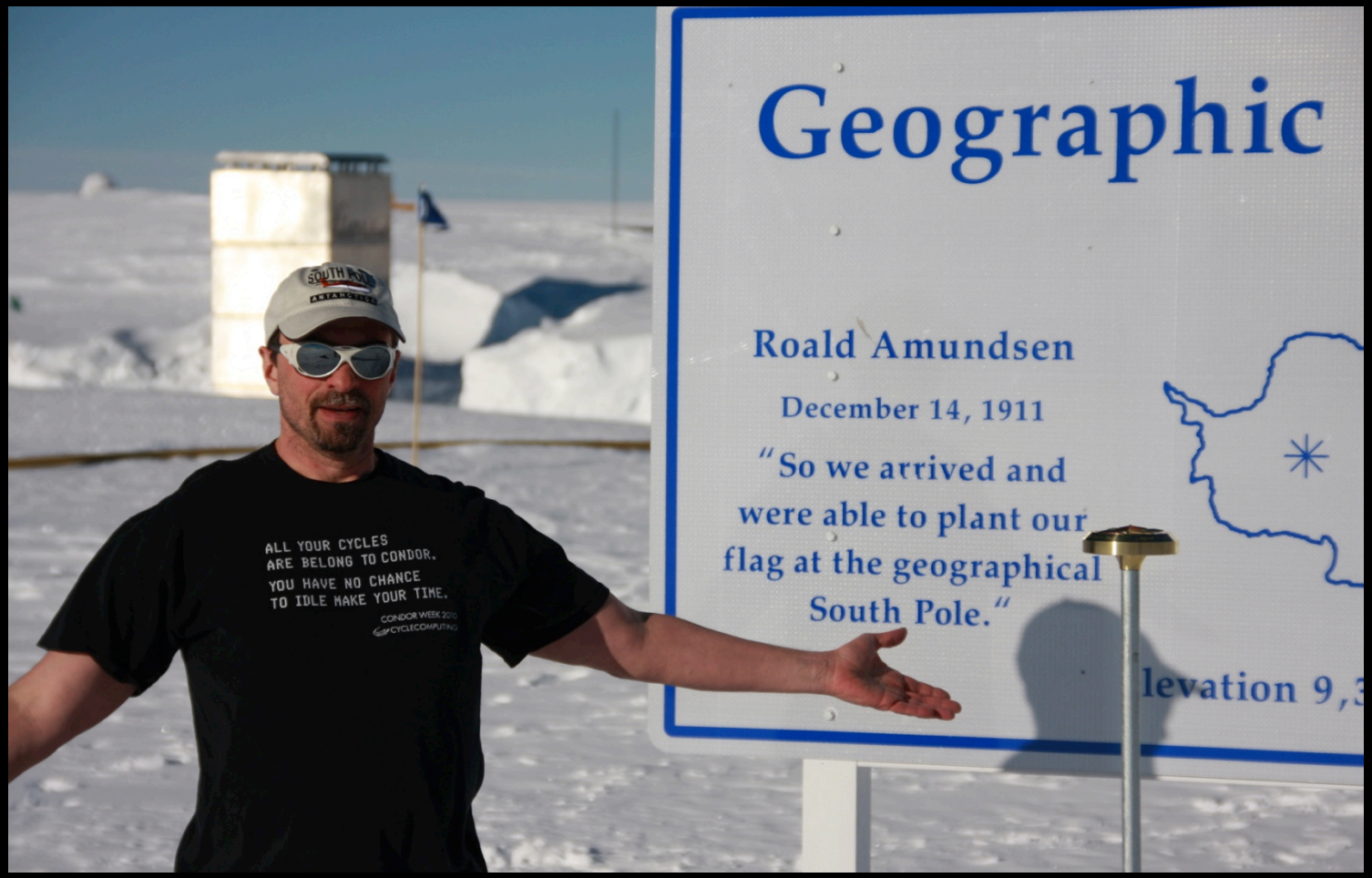
- It's hard to make predictions, especially about the future – Yogi Berra
- Upgrade HTCondor to development series (as we speak)
- Use TJ's new detection features



Credits

- Juan Carlos Diaz-Velez, David Schultz – IceProd framework
- Dmitry Chirkin, Claudio Kopper – GPU simulation development
- Vladimir Brik– GPU wrangling
- CHTC – Hosting the GZK cluster, OSG support, HTCondor GPU efforts (TJ!)
- UW Madison – Fast networks
- NSF – IceCube M&O funding

Questions?


Geographic
Roald Amundsen
December 14, 1911
"So we arrived and were able to plant our flag at the geographical South Pole."
ALL YOUR CYCLES ARE BELONG TO CONDOR.
YOU HAVE NO CHANCE TO IDLE MAKE YOUR TIME.