

Using HTCondor Glideins to Run in IceCube Heterogeneous Resources

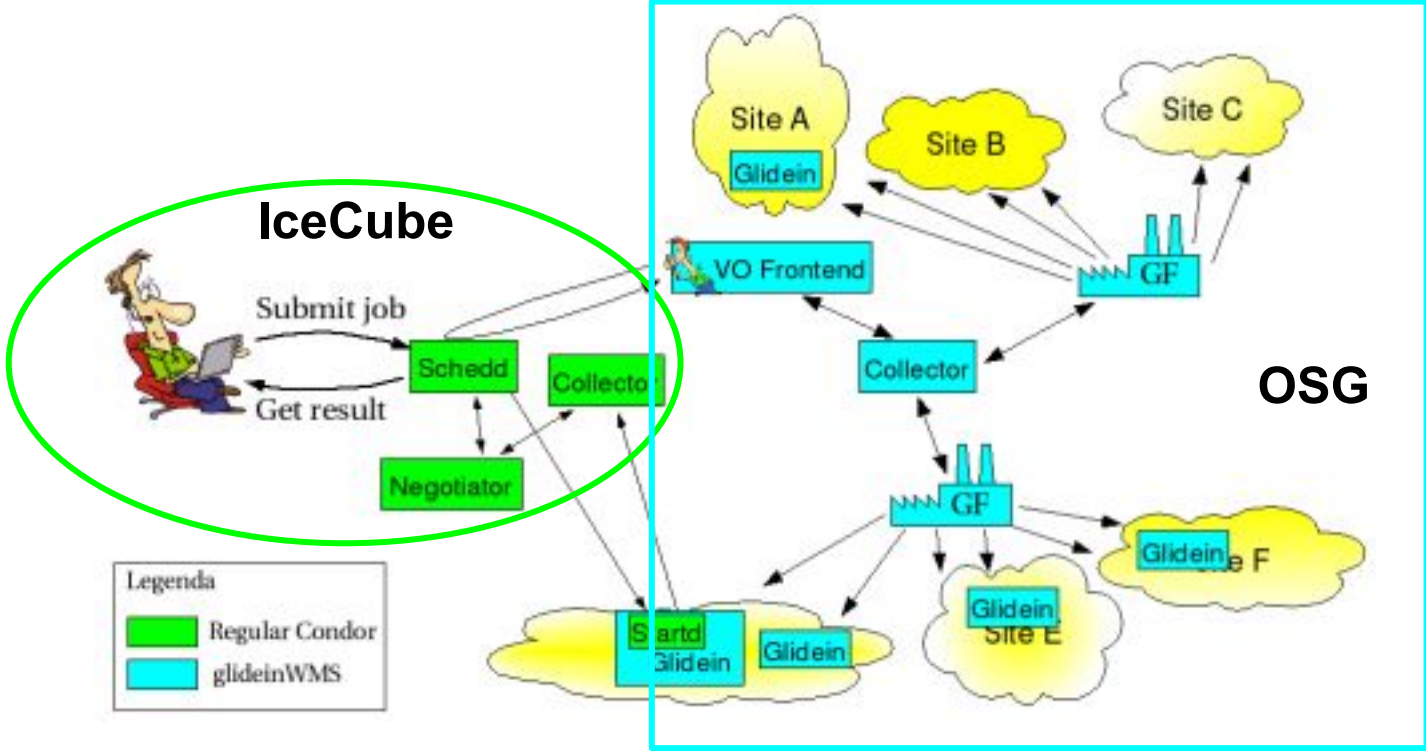
David Schultz
IceCube



Overview

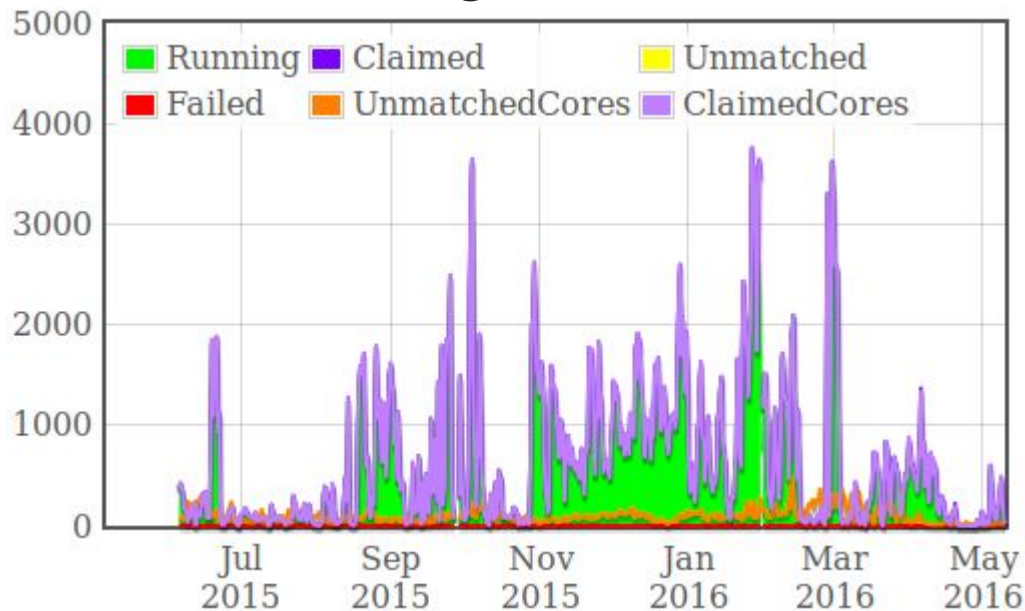
- Grid sites: GlideinWMS
- Non-grid sites: pyglidein
- Various resource types:
CPUs, GPUs, large memory

Grid sites: GlideinWMS



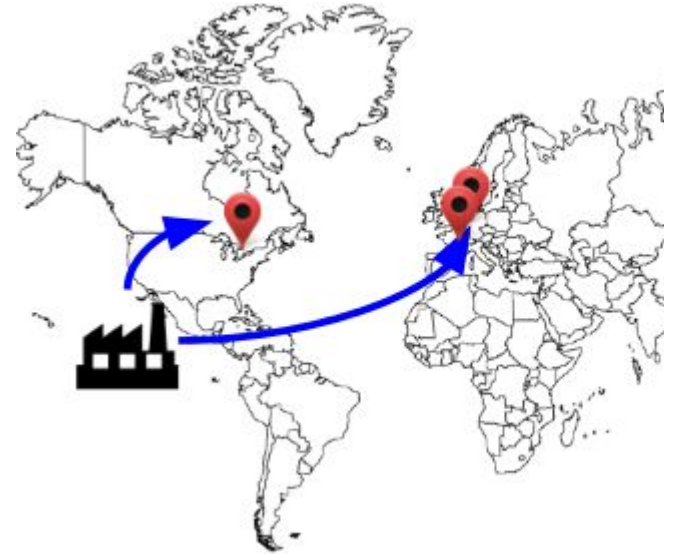
Grid sites: GlideinWMS

Since 2013, IceCube has used the GLOW VO on the Open Science Grid, through CHTC

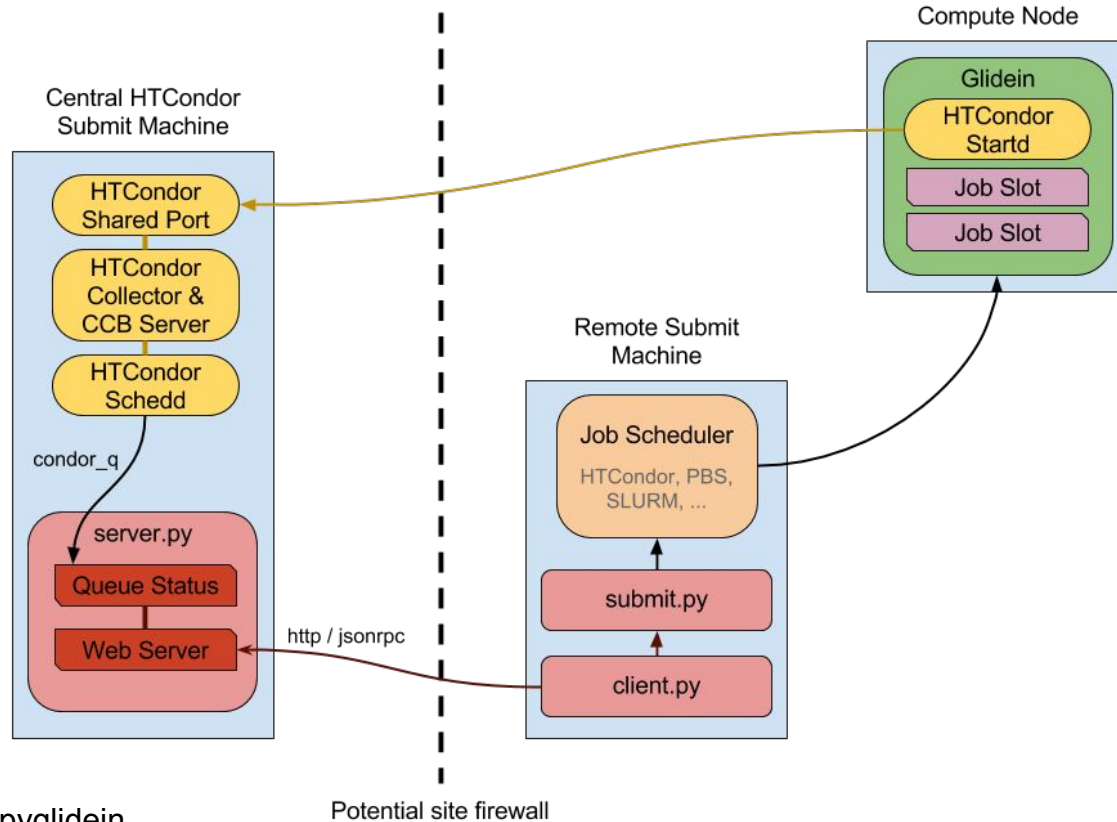


Grid sites: GlideinWMS

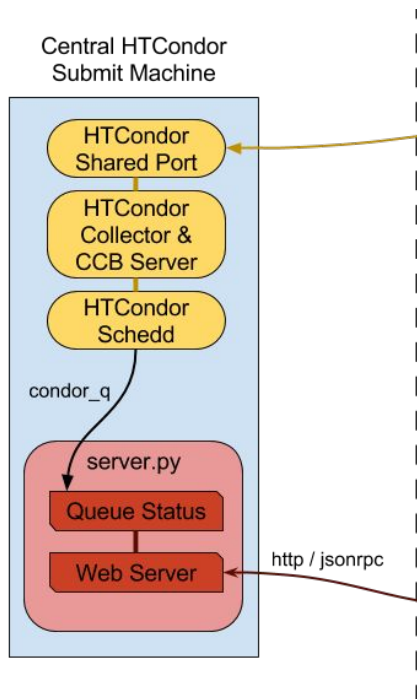
- Moving to icecube VO
- Still leveraging CHTC / OSG
- Adding more sites:
 - Germany
 - Canada
 - Other IceCube grid sites



Non-grid sites: pyglidein



Non-grid sites: pyglidein



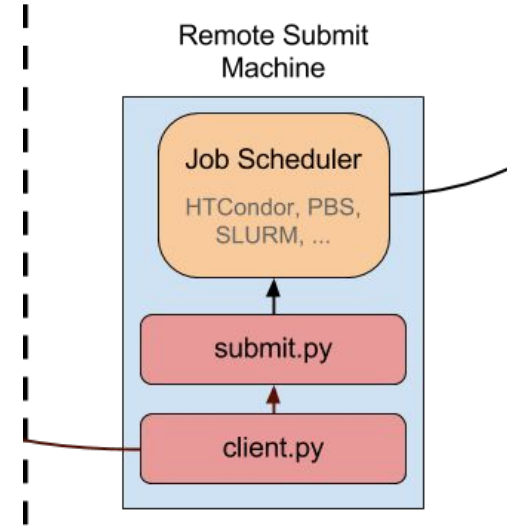
- Standard HTCondor server
 - ▶ Shared port and CCB to make networking easier
- **server.py** user script
 - ▶ Query HTCondor every X minutes
 - ▶ Aggregate idle job resource requests
 - ▶ Present requests via http / jsonrpc

Non-grid sites: pyglidein

- **client.py** user script
 - ▶ Query **server.py** for requests
 - ▶ Check local queue for # idle
 - ▶ Submit new requests

- **submit.py**

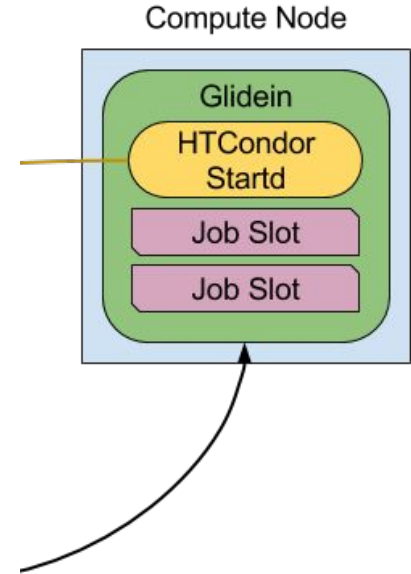
- ▶ Handles abstraction of different job schedulers



Non-grid sites: pyglidein

– Glidein job

- ▶ Get resources allocated by scheduler
 - Environment variables from `submit.py`
 - Auto-sense for assigned GPU(s)
- ▶ Pass resources to HTCondor Startd



Non-grid sites: pyglidein

- Started in 2015
 - Simple, non-optimized, yet ran 20% of production
- Can be deployed in minutes by a non-expert
- Because we host it, updates are fast
 - GPU errors at a new site fixed in a day
 - Latest parrot version needed for our OpenCL code

Non-grid sites: pyglidein

- Several collaboration sites have small, local clusters
 - Pyglidein gives them a way to contribute in a non-monetary way

Non-grid sites: pyglidein

- Used for IceCube supercomputer allocations through XSEDE:
 - Comet (>10,000 GPU hours used so far)
 - Bridges (coming soon)

Heterogeneous Resources

- IceCube jobs need (variously):
 - Large memory
 - Large scratch disk
 - GPUs

Heterogeneous Resources

- HTCondor partitionable/dynamic slots

- ▶ A regular single slot:

Glidein - 1 CPU, 2GB	
Slot - 1 CPU, 2GB	

- ▶ PBS high memory:

4 CPU, 10GB			
1 CPU, 6GB	1	1	1

- ▶ Whole node:

24 CPU, 64GB, 2 GPU						
1 CPU, 2GB, 1 GPU	1 CPU, 2GB, 1 GPU	1 CPU, 10GB	1	1	...	1

Questions?