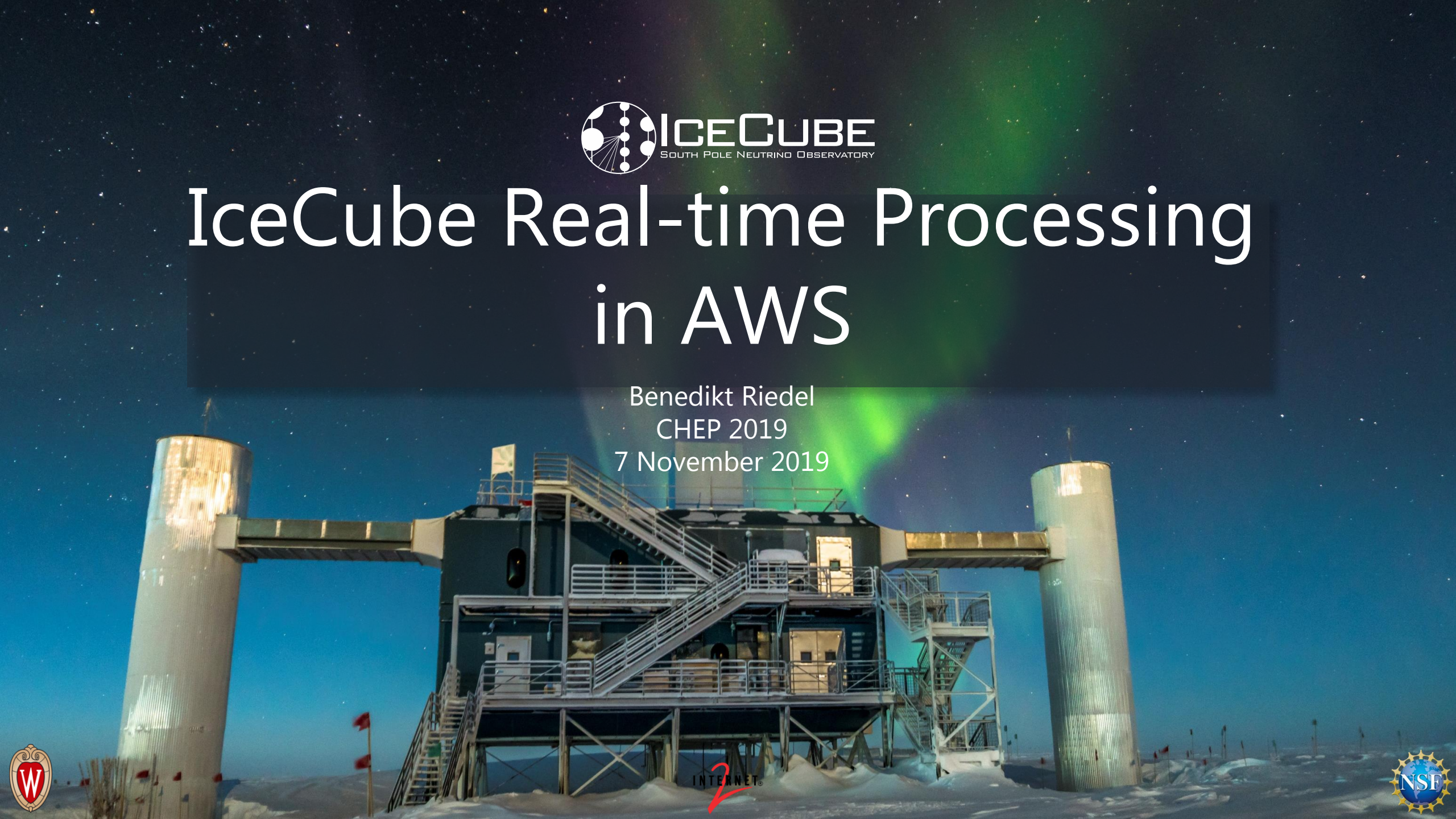


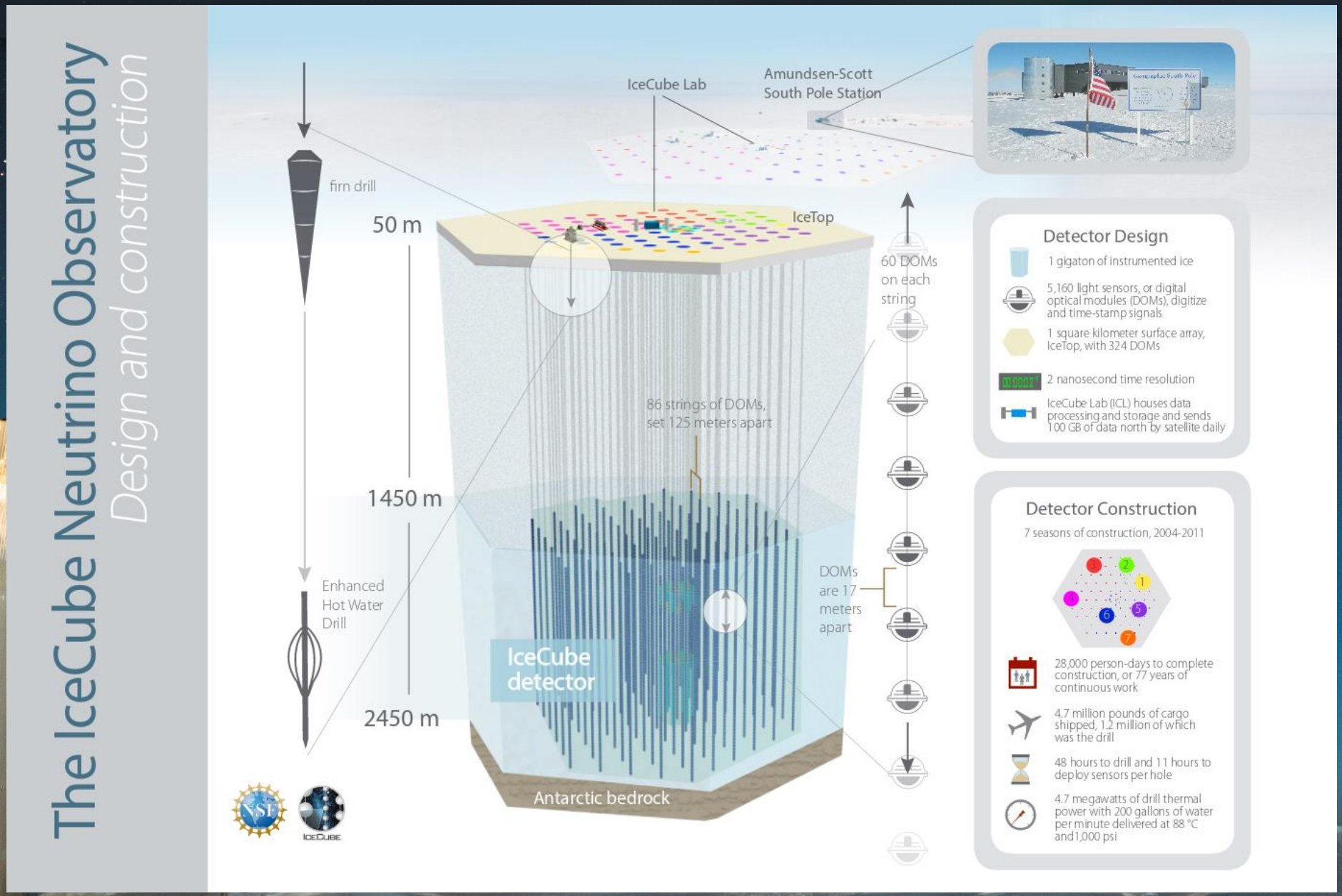
IceCube Real-time Processing in AWS

Benedikt Riedel

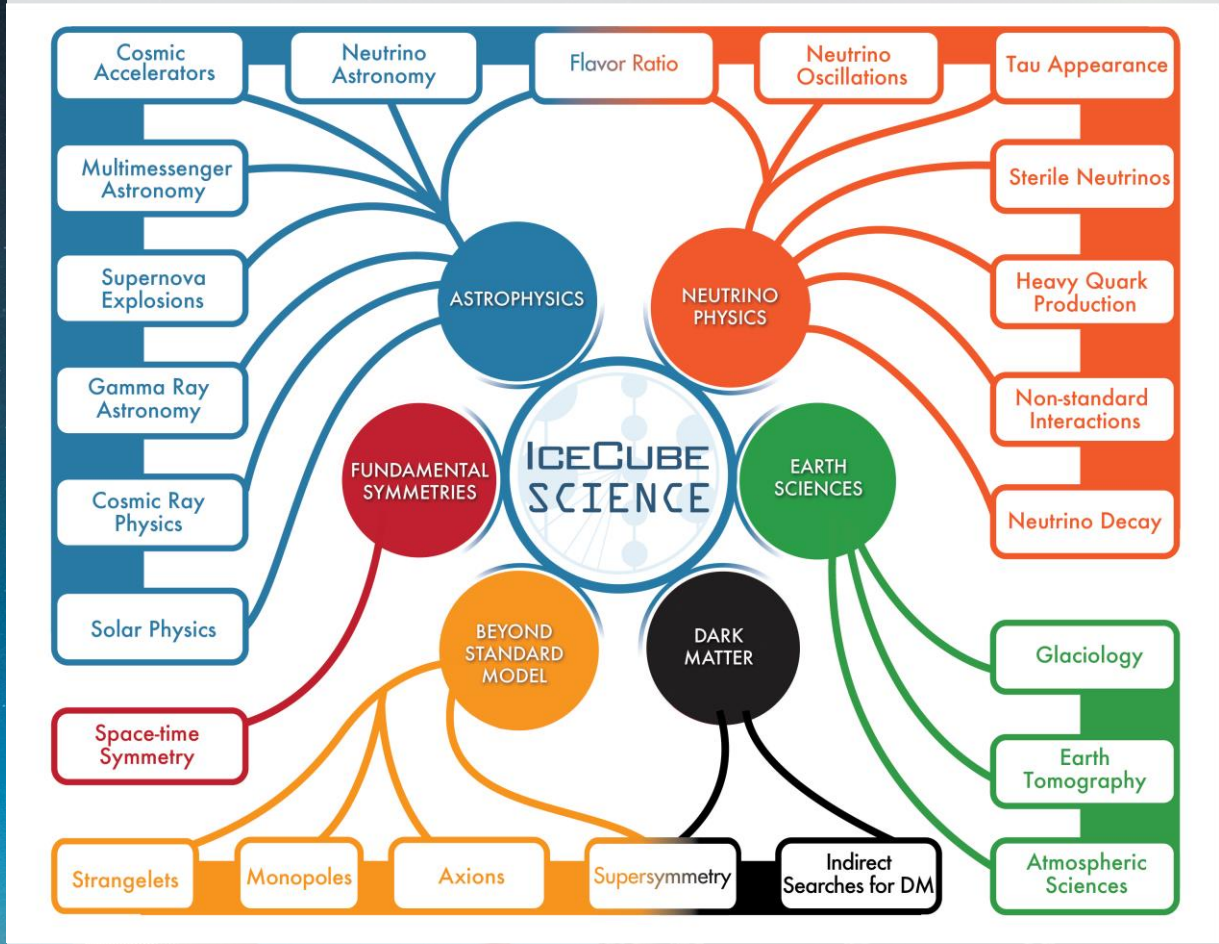
CHEP 2019

7 November 2019

IceCube

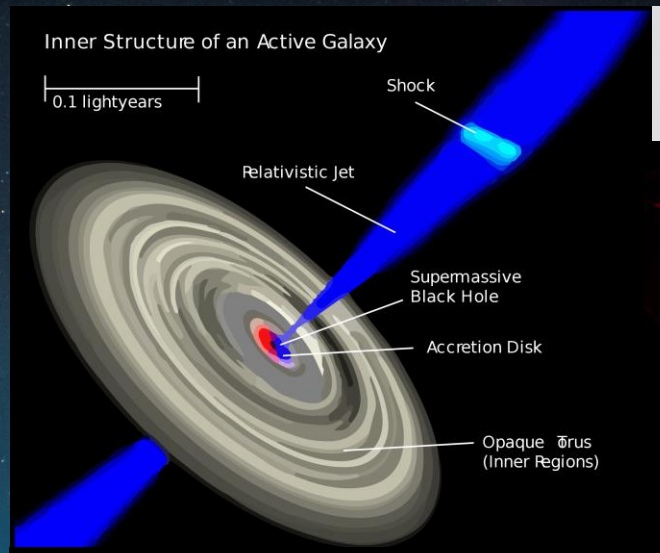

The IceCube Neutrino Observatory
Design and construction
IceCube Lab
Amundsen-Scott South Pole Station
firn drill
50 m
IceTop
60 DOMs on each string
86 strings of DOMs, set 125 meters apart
1450 m
Enhanced Hot Water Drill
DOMs are 17 meters apart
IceCube detector
2450 m
Antarctic bedrock
Detector Design
1 gigaton of instrumented ice
5,160 light sensors, or digital optical modules (DOMs), digitize and time-stamp signals
1 square kilometer surface array, IceTop, with 324 DOMs
2 nanosecond time resolution
IceCube Lab (ICL) houses data processing and storage and sends 100 GB of data north by satellite daily
Detector Construction
7 seasons of construction, 2004-2011
28,000 person-days to complete construction, or 77 years of continuous work
4.7 million pounds of cargo shipped, 1.2 million of which was the drill
48 hours to drill and 11 hours to deploy sensors per hole
4.7 megawatts of drill thermal power with 200 gallons of water per minute delivered at 88 °C and 1,000 psi

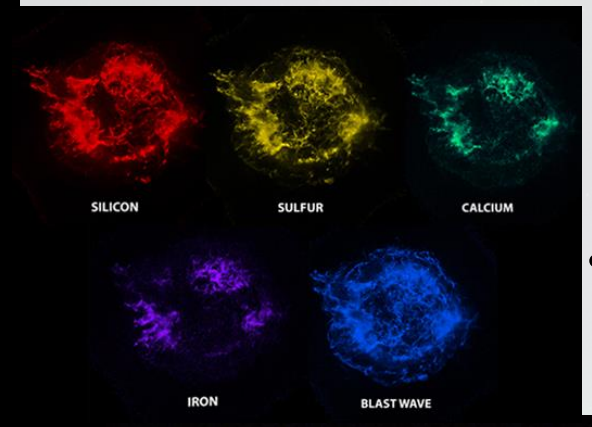
IceCube – What drives us?

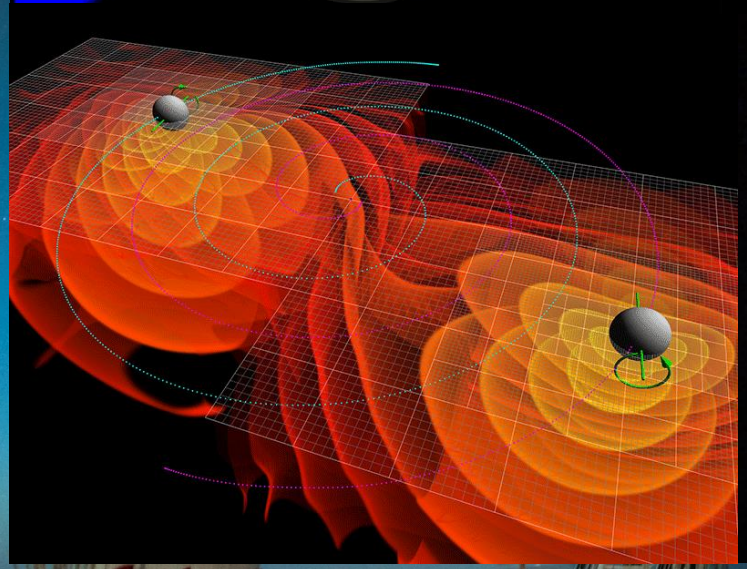


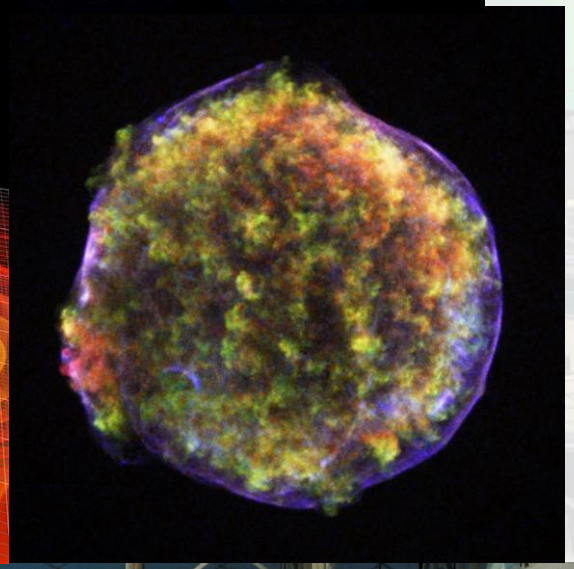
- Novel instrument in multiple fields
- Broad science abilities, e.g. astrophysics, particle physics, and earth sciences
- Lots of data that needs to be processed in different ways
- Lots of simulation that needs to be generated

Multi-Messenger Astrophysics



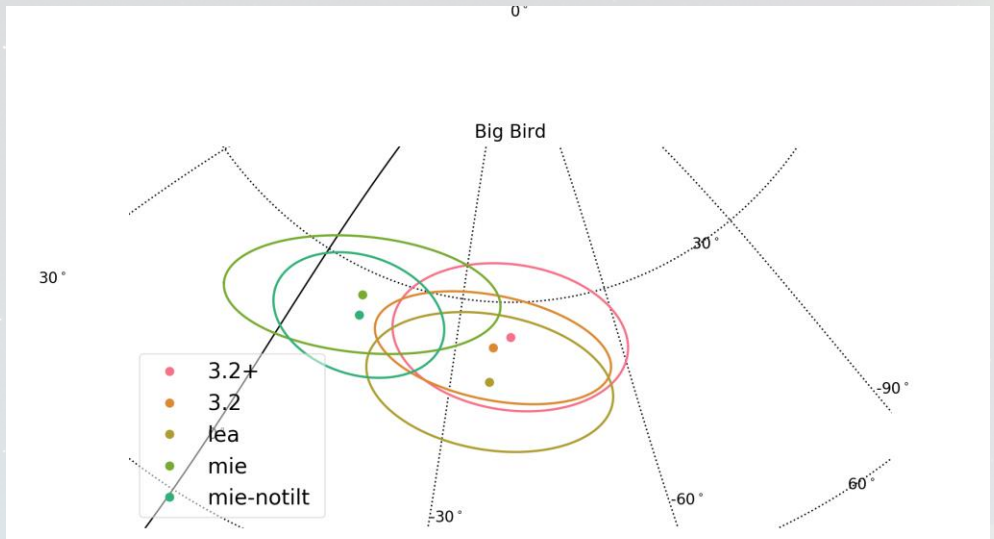


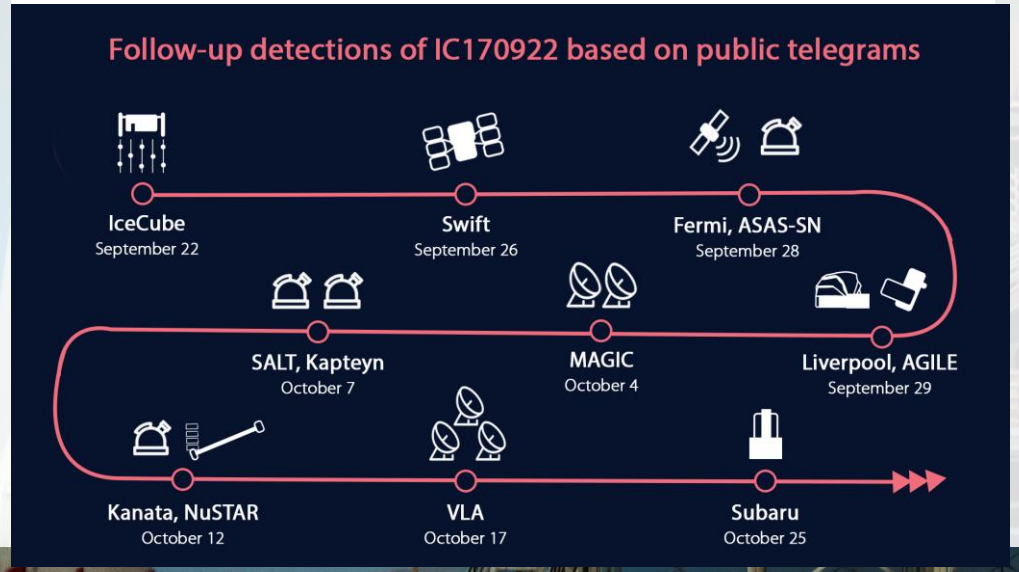




- Observing astrophysical phenomena in using multiple astrophysical "messengers"
  - Transient phenomena – mergers of compact objects, flaring objects, supernovae, etc.
  - Continuous sources – Pulsars, AGN, etc.
- What is a "messenger"?
  - Traditionally
    - Electromagnetic (EM) emission – Radio, IR, Optical, X-Ray, Gamma
    - Cosmic Rays – Particles
  - Today
    - EM emission – Same as before with a focus on Gamma and Optical
    - Cosmic Rays – Particularly extremely high energy (EeV / particle)
    - Neutrinos
    - Gravitational waves (GW)

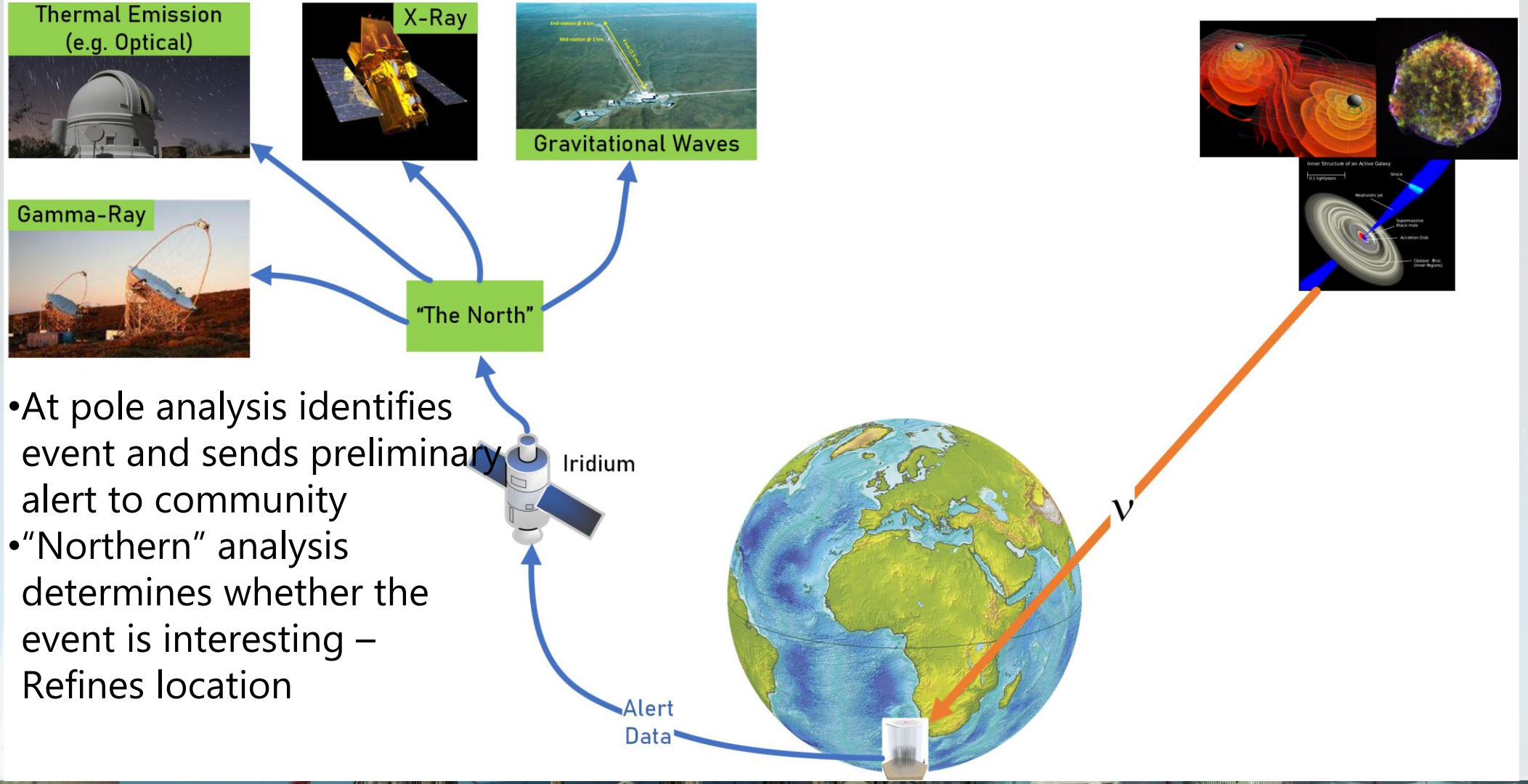
Multi-Messenger Astrophysics – IceCube





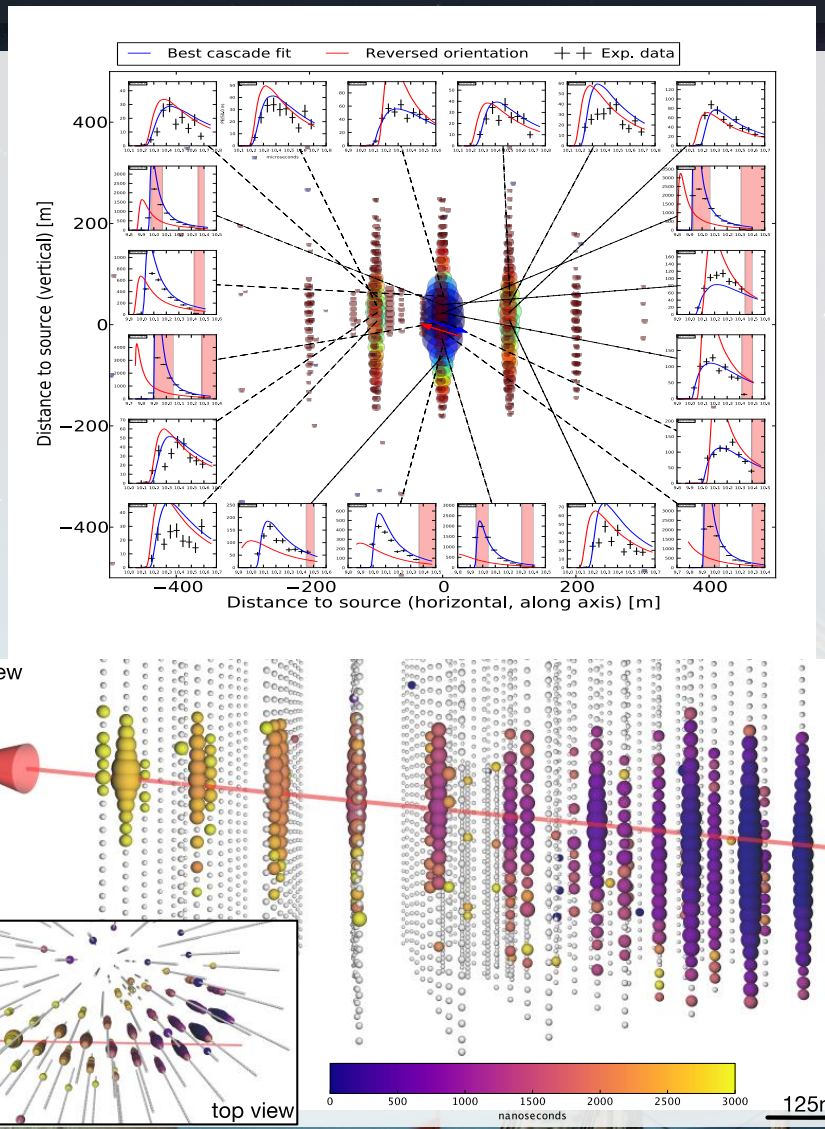
- Want to alert the community at large about interesting events
- Majority telescopes have a minute area of the sky they cover
- Need to be as precise and fast as possible, else wasting valuable telescope time or will miss source (transient sources)
- Optical model of ice is very important factor
- IceCube alerted community on 22 Sept 2017 about a muon neutrino coming from close to Blazar TXS 0506+056
- Follow-up by multiple telescopes

Multi-Messenger Astrophysics - Alert



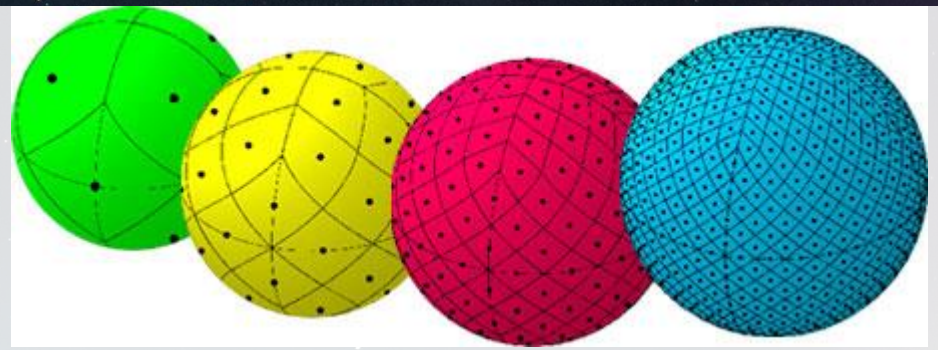
- At pole analysis identifies event and sends preliminary alert to community
- "Northern" analysis determines whether the event is interesting – Refines location

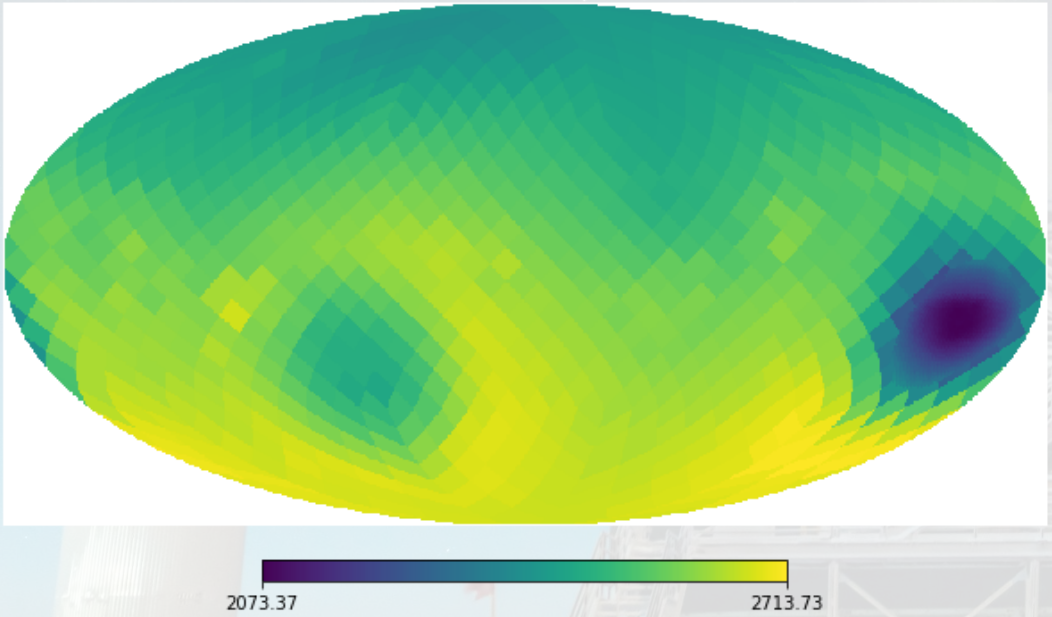
Reconstruction



- "Northern" processing is mostly refining the location on sky
- IceCube has developed a multitude of energy and directional reconstructions
- Trade-off between speed and accuracy
  - At pole
    - All events: Speed >>> Accuracy
    - "MMA" events: Speed > Accuracy
  - "North"
    - "Physics" events: Speed >> Accuracy
    - "MMA" events: Accuracy > Speed
    - "Really MMA" events: Accuracy >> Speed
- Big question is still the ice optical model – Not all reconstructions support

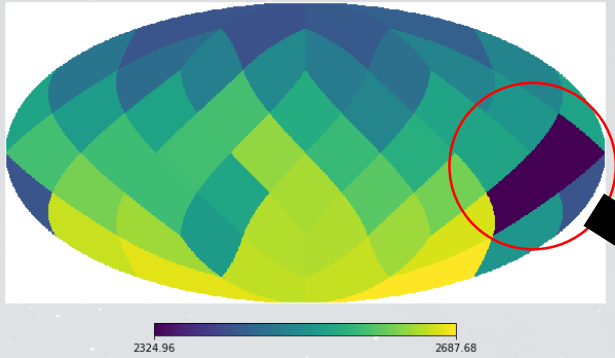
Reconstruction – Accuracy

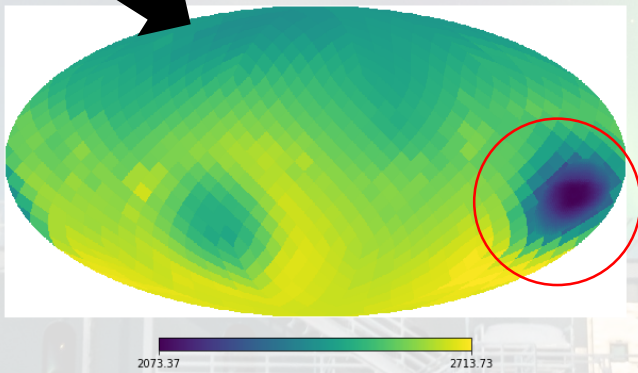


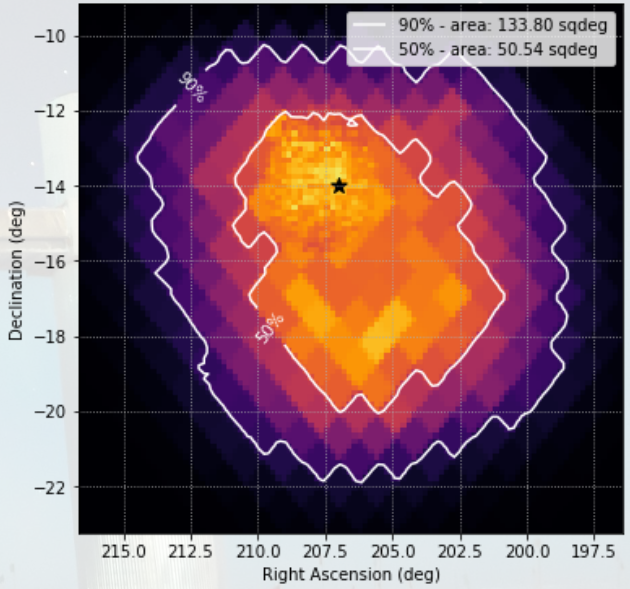


- Most accurate directional reconstruction comes by scanning across the sky
  - Split sky into constant surface area pieces
  - Test each directional hypothesis against likelihood
  - Create directional likelihood map
  - Gives most probable direction and error
- Each hypothesis calculation is independent – Easy to split up workload across O(1000[000]) or cores

Current Scan



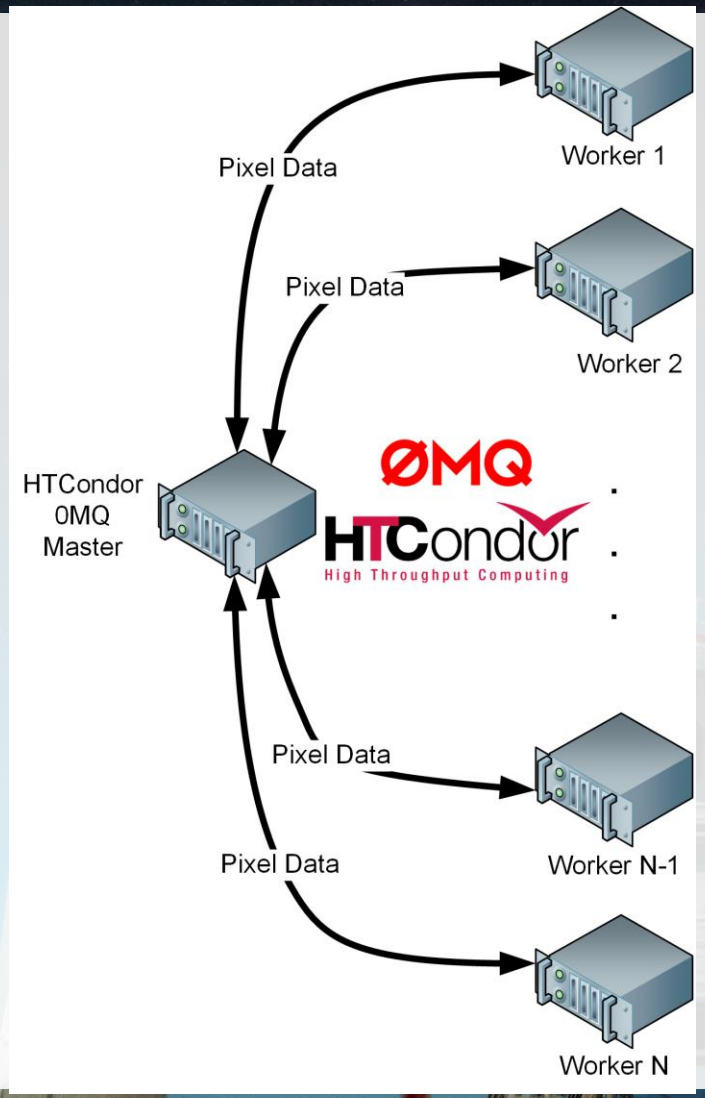




- Start with a rough estimate
- Zoom in on “hot” region with more detailed scan
- Step through scans until we have reach detector resolution

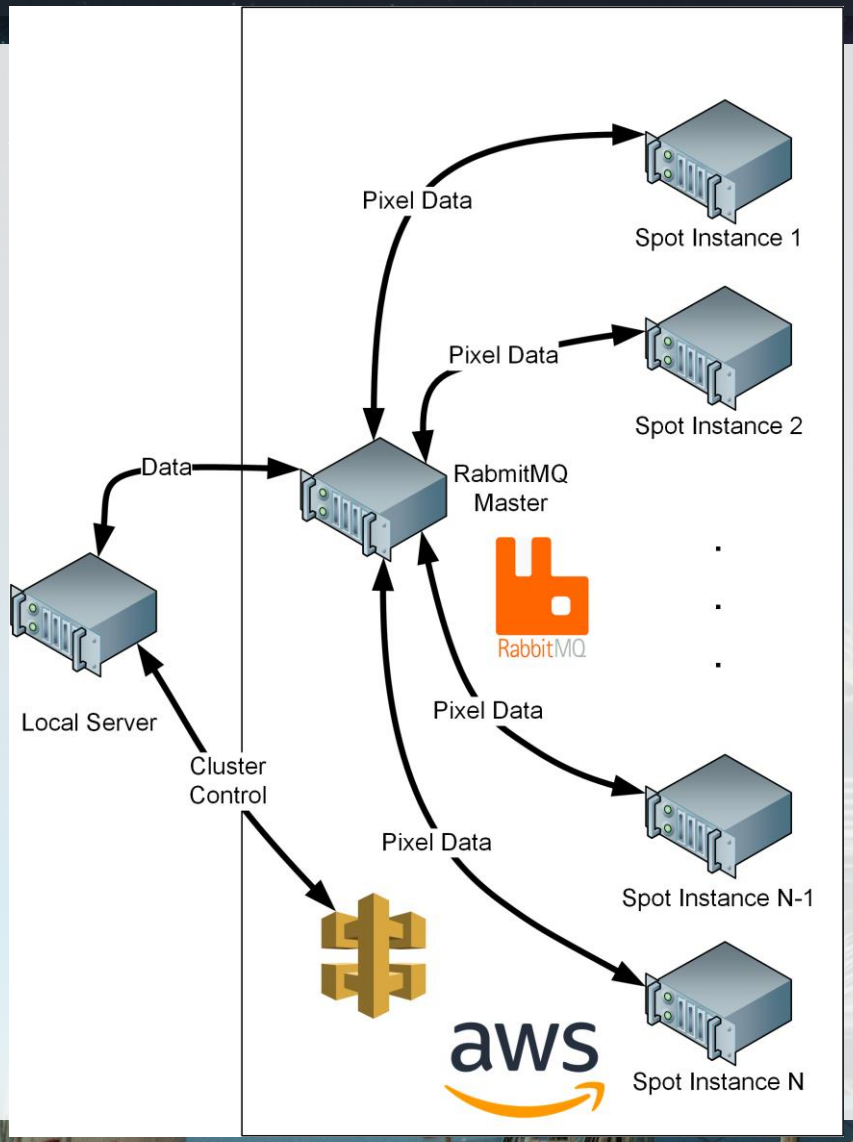
- Need to repeat scan to get systematic errors
- Shift in systematic space and re-scan
- Mostly useful for internal studies

Current Scan – Distributing Work



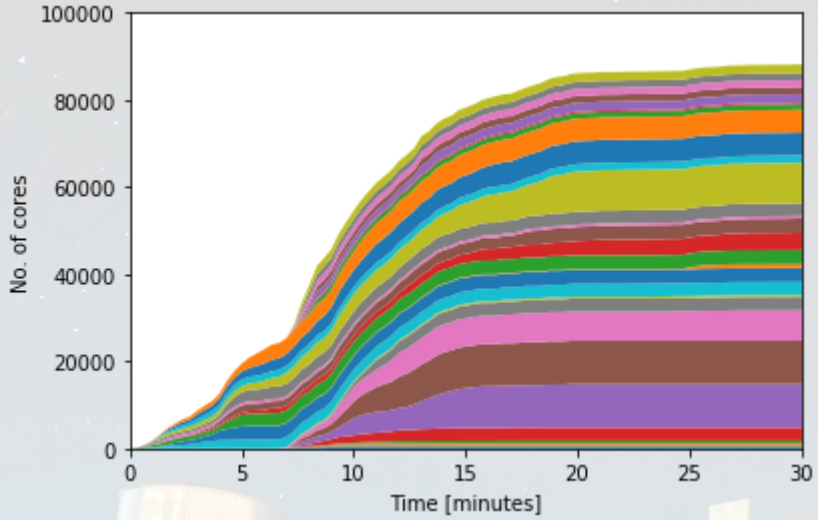
- Master-worker setup
- Worker resource requirements are the same
- Master makes decision about next scan
- Data communication via ZeroMQ
  - Easy to use and setup – Experience in other applications
  - Scaling an issue
    - Issue with communicating over the grid – Firewall issues
    - Over 2000 cores it can't keep up
- Using HTCondor for scheduling workers

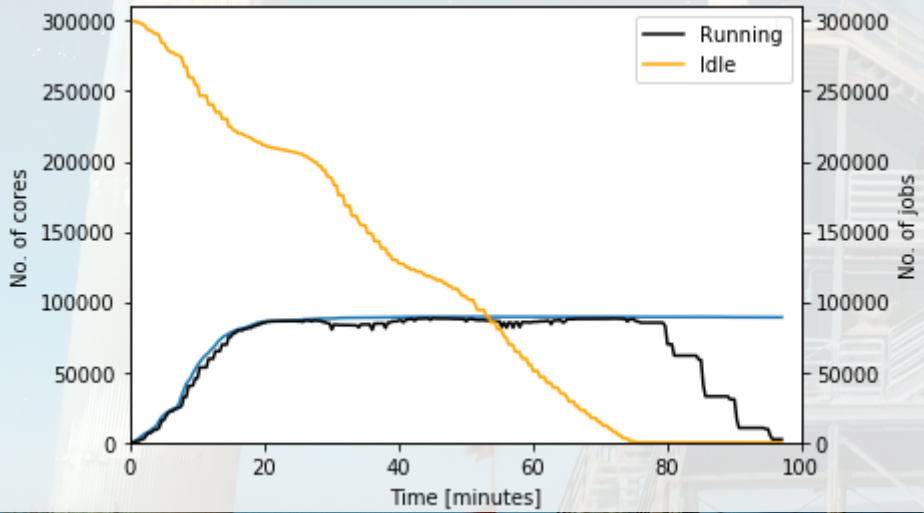
Cloud Scan – Distributing Work



- Cloud/AWS gives us the ability to scale quickly – "Burst" into cloud
- Reduce time to result!
  - Go to detector resolution immediately
  - 3000 core hours needed in O(10 min)
- Data movement via RabbitMQ
  - Container deployment – Avoid Erlang/BEAM knowledge/headache
  - Many Source, many sinks
  - Higher scaling than ZeroMQ - > 5000
  - Planning on replacing most Monte Carlo production data data movement
- Scaling in a single cloud zone limited by AWS – Current spot scaling limit 5000
- Multi-zone/multi-core needed to get to large scaling
  - Orchestration layer for data to reduce transfer cost
  - Are there sufficient multi-core instances in a zone?

Cloud Scaling Tests – EAGER





- Can scale to > 80k cores across 3 cloud and 28 zones within ~20 min
- Processed 300k jobs within 100 minutes
- Need multi-cloud, multi-zone deployment to reach this level
- Used HTCondor to do testing and matching
- Question whether this would work in master-worker setup with RabbitMQ
- Funded through EAGER for SC19 demo
- For other results from testing see talks regarding cloud networking

Summary

- MMA follow-up requires massive amounts
- AWS/cloud can provide means to reduce time to result for “bursty” applications – MMA follow-up ideal
- Data distribution for “bursty” application an issue
  - ZeroMQ doesn’t scale well
  - RabbitMQ scales
  - Cross-zone data transfers are costly
- Multi-zone vs. multi-core instances an open question


Thank you!
Questions?