Fermilab

Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

Dark Energy Survey on the OSG

Ken Herner

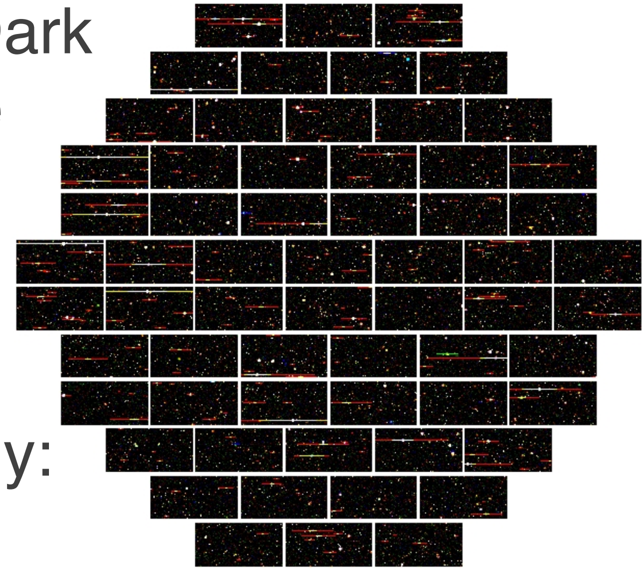
OSG All-Hands Meeting

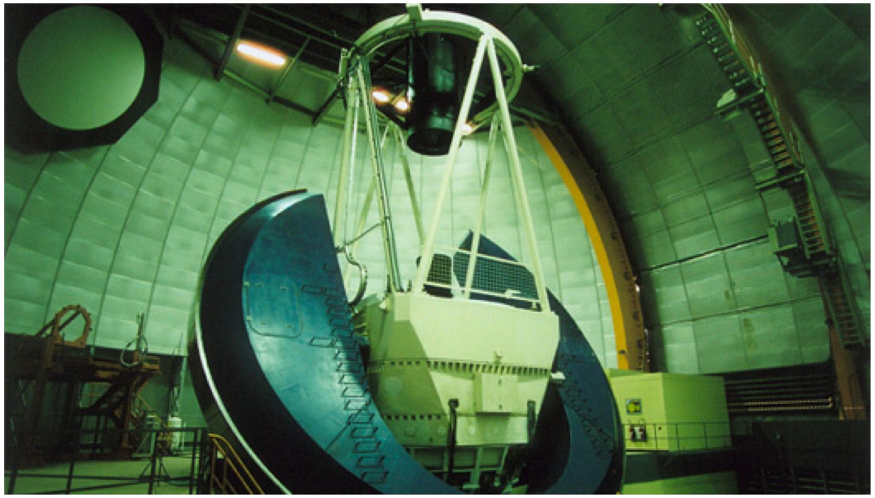
19 Mar 2018

Credit: T. Abbott and NOAO/AURA/NSF

The Dark Energy Survey: Introduction

- Collaboration of 400 scientists using the Dark Energy Camera (DECam) mounted on the 4m Blanco telescope at CTIO in Chile
- Recently finished 5th year of 5-yr mission
  - Expect additional half-year in fall 2018
- Main program is four probes of dark energy:
  - Type Ia Supernovae
  - Baryon Acoustic Oscillations
  - Galaxy Clusters
  - Weak Lensing
- A number of other projects e.g.:
  - Trans-Neptunian/ moving objects

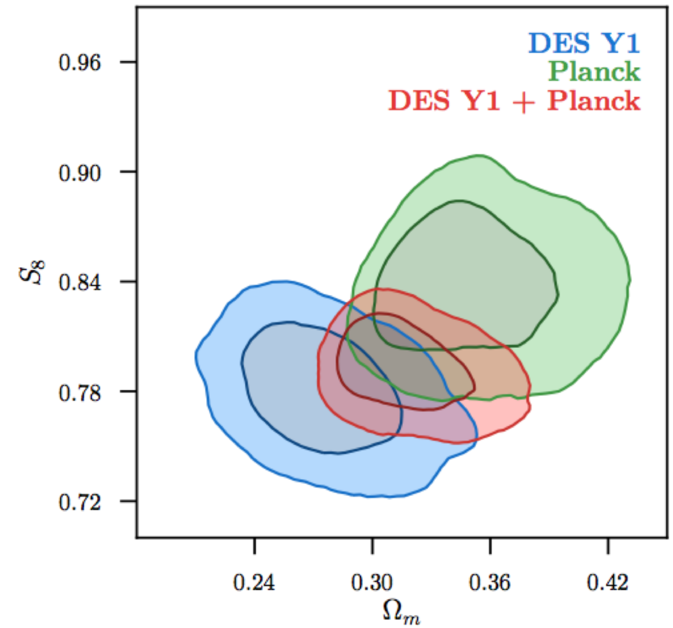


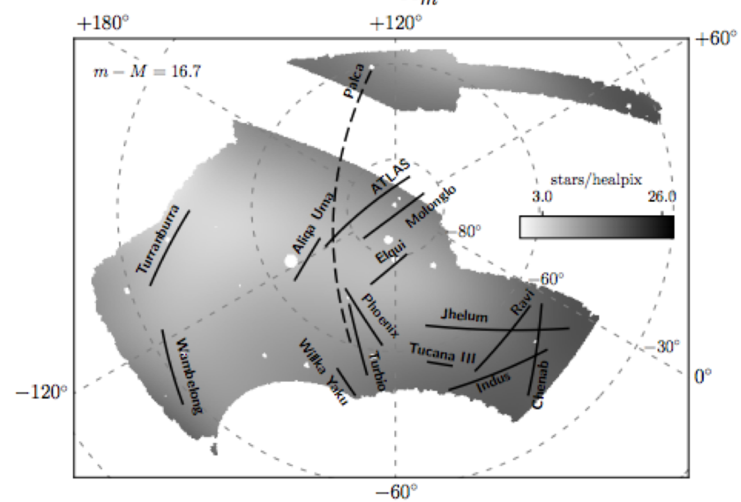


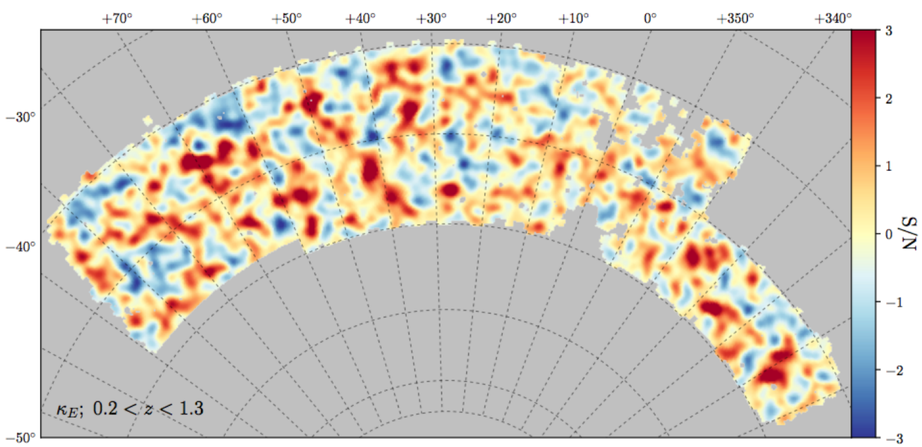
Recent DES Science Highlights (not exhaustive)

THE DARK ENERGY SURVEY

- Most accurate DM map to date
- Oldest supernova observed to date
- Competitive results on cosmological parameters
- Recent public release of first 3 years of data
- New stellar streams in Milky Way
- Optical Follow-up of GW triggers



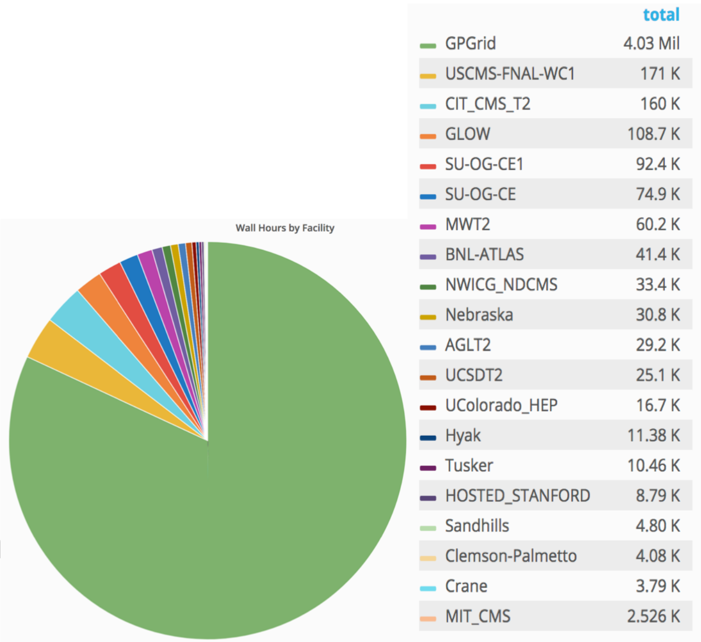




Overview of DES computing resources

THE DARK ENERGY SURVEY

- About 3 dozen rack servers, 32-48 cores each, part of FNAL GPGrid but DES can reserve them. used for nightly processing, reprocessing campaigns, and deep coadds (64+ GB RAM)using direct submission from NCSA.
- Allocation of 980 "slots" (1 slot = 1 cpu 2 GB RAM) on FermiGrid, plus opportunistic cycles
- OSG resources (all sites where Fermilab VO is supported)
- Regularly running at NERSC
- Various campus clusters
- Individuals use FNAL Wilson (GPU) Cluster
  - Difficult to run at scale due to overall demand
- By the numbers:
  - 3/2017 – 3/2018: 4 M hours; 81% on FermiGrid
  - Does not count NERSC/campus resources
  - Does not count NCSA->FNAL direct submission



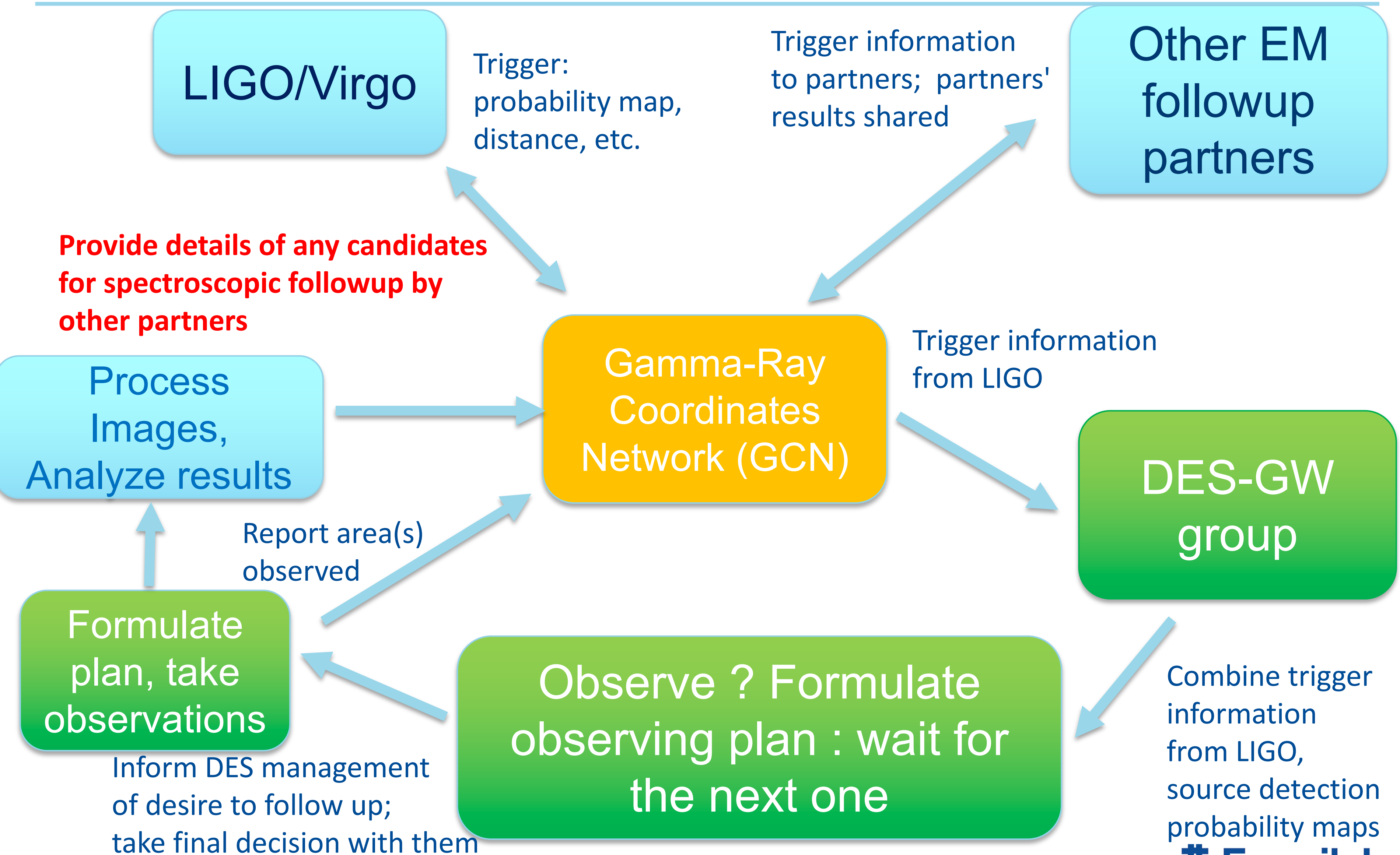
Fermilab

StashCache adoption

- Started using StashCache in late spring 2017
  - Used mostly for image calibration files (flats, sky subtraction, etc.) in single-epoch processing (performed on raw image); typically tens of MB per file and a dozen or so per job
  - Also used for small on-the-fly star catalog is GW followup processing in regions outside of normal DES areas.
- Slightly difficult to quantify efficiency improvement, but wall time of affected jobs now shorter
  - Extremely important for LIGO follow-up when one can't wait for standard pipeline processing (usually 6 hours after image taken)

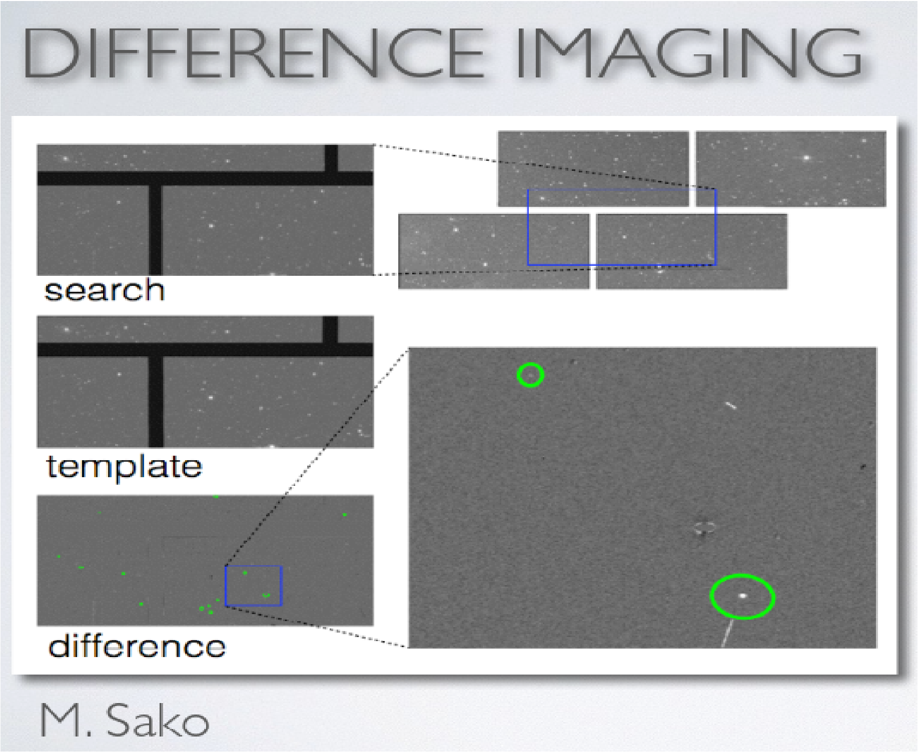
Overview of GW EM followup

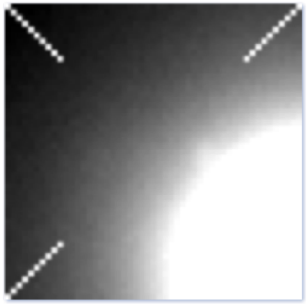
THE DARK ENERGY SURVEY



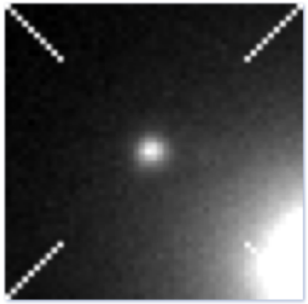
Difference Imaging Software (GW Follow-up and TNOs)

- Originally developed for Supernova studies
- Several ways to get GW events
- DES is sensitive to neutron star mergers or BH-NS mergers (get an optical counterpart), core collapse
- Main analysis: use “difference imaging” pipeline to compare search images with same piece of sky in the past (i.e. look for objects that weren’t there before)

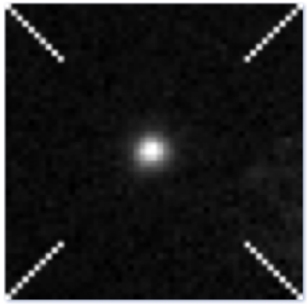




Template, z-band



Search, z-band
20170817

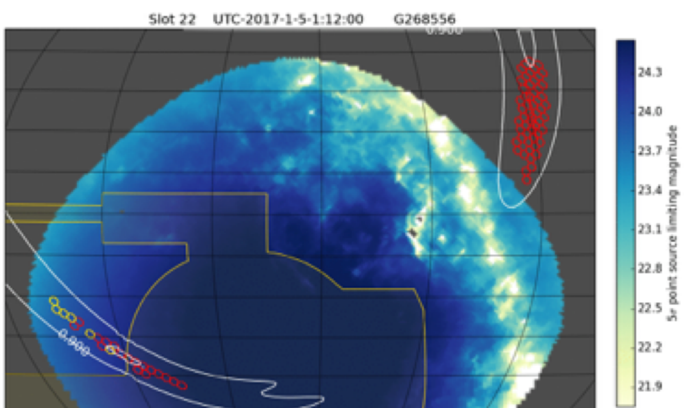


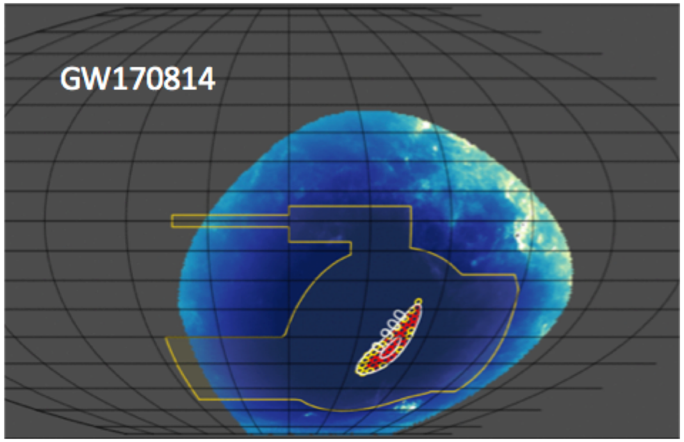
Difference, z-band
20170817

Image analysis pipeline

- Each search and template image first goes through "single epoch" processing (few hours per image). About 10 templates per image on average (some overlap of course)
  - Now uses 4 cpus and 3.5 – 4GB memory; up to 100 GB local disk. Increased resources don't hurt as much because memory/core actually went down.
- Once done, run difference imaging (template subtraction) on each CCD individually (around 1 hour per job, 2 GB RAM, ~50 GB local disk)
- Totals for first GW event: about 240 images for main analysis *59 CCDs per image three nights = 5000 CPU-hours for diffimg runs needed per night
  - Addition of VIRGO means 6-10x smaller search areas
- File I/O is with FNAL dCache

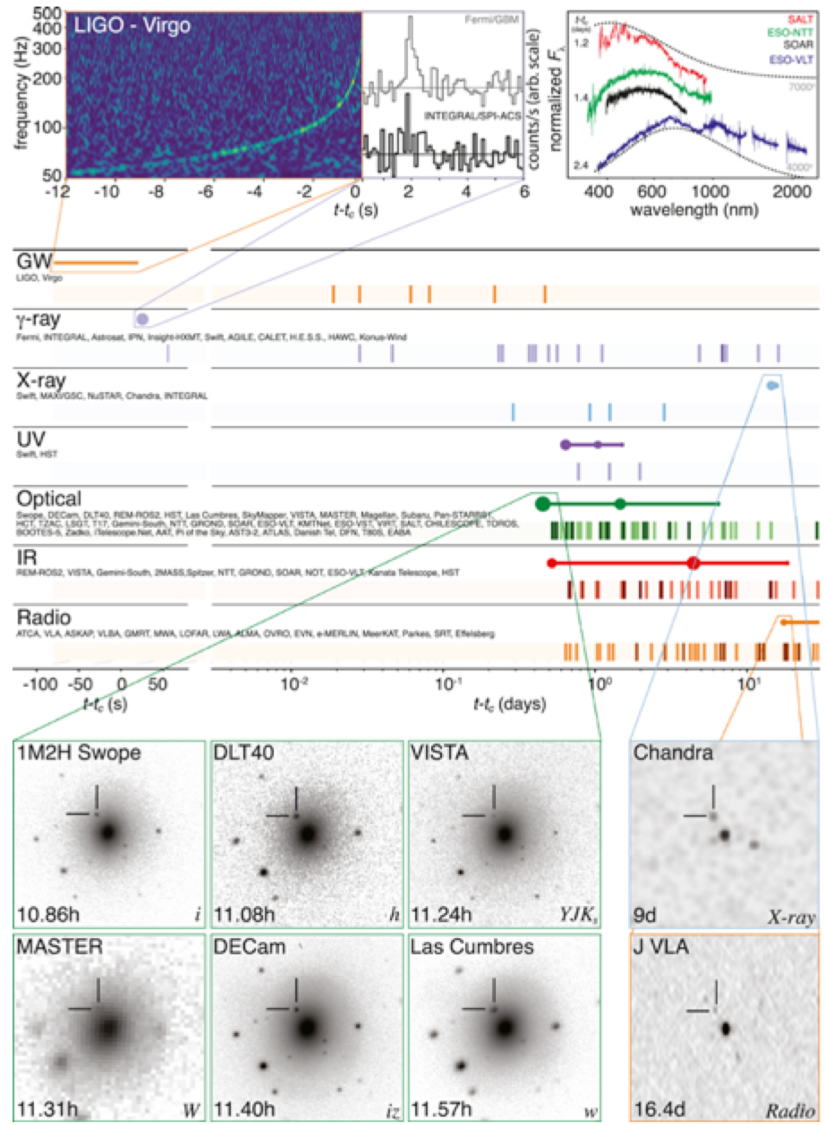
GW170104
(no VIRGO)





The Need for Speed

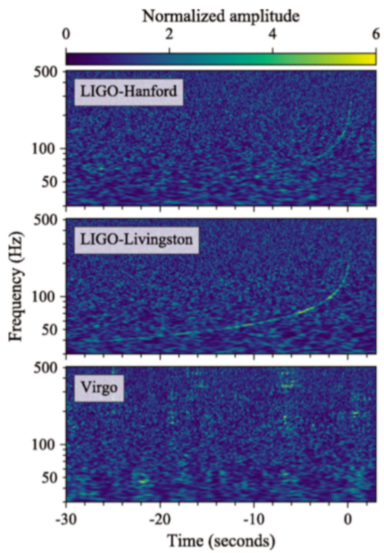
- 6k CPUs is not that much in one day, but one can't wait a long time for them. Want to process images within 24 hours (15 is even better) **allowing DES to send alerts out for even more followup while object is still visible.** First event was over a longer period.
  - **For next LIGO run, aim for 1 hour image processing turnaround**
- Necessitates opportunistic resources (OSG); possibly Amazon/Google at some point if opportunistic resources unavailable
  - Did a successful AWS test last within FNAL HEPCloud demo

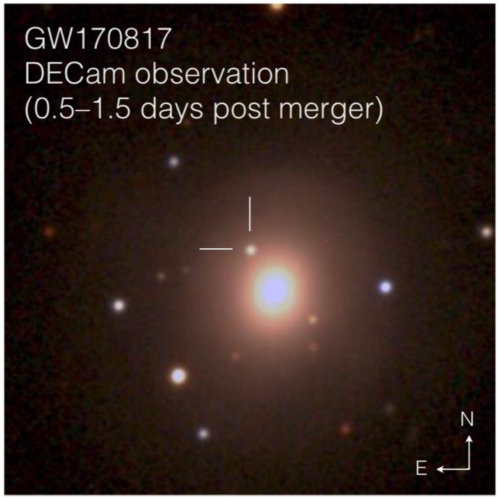


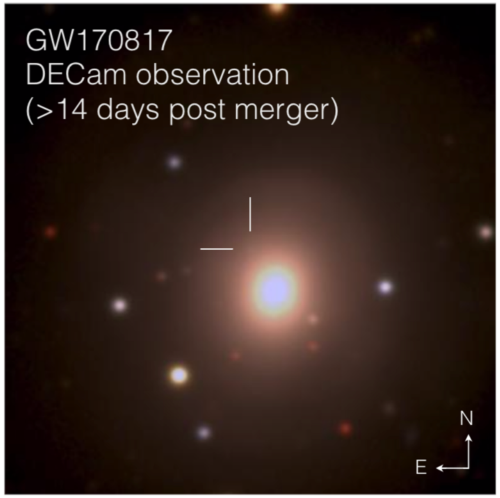
LVC et al., ApJL 848:L12 (2017)

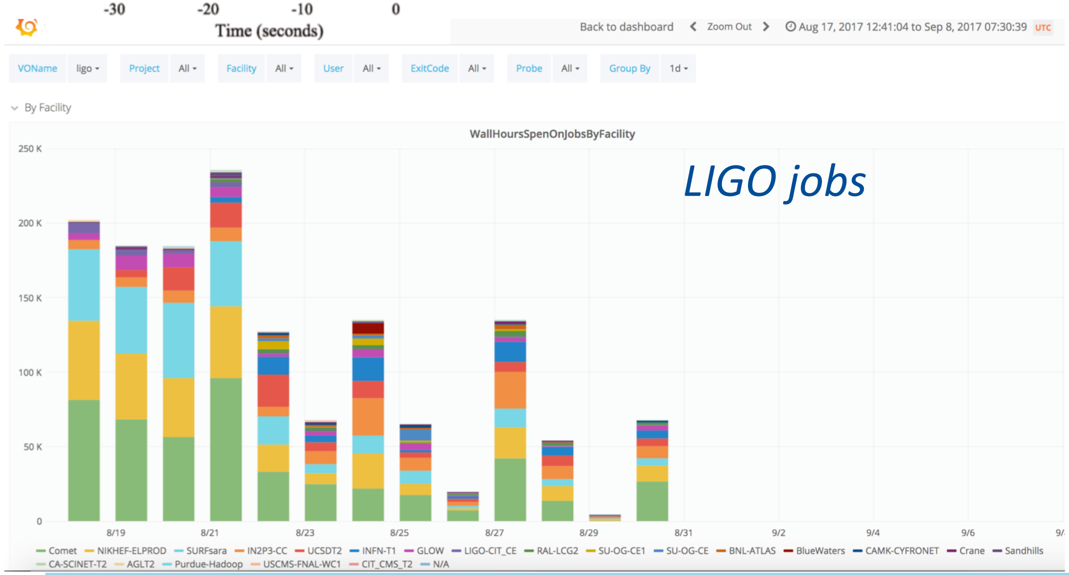
A bit on GW170817

- LIGO was no stranger to OSG...

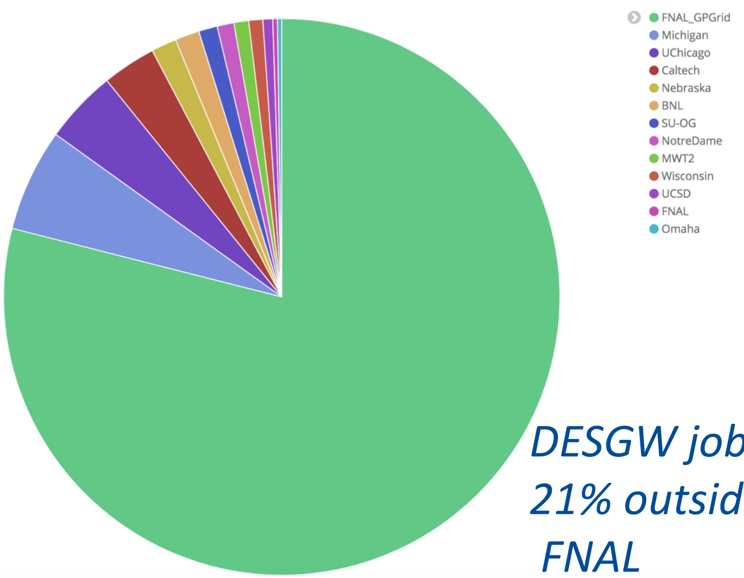








LIGO jobs



DESGW jobs 21% outside FNAL

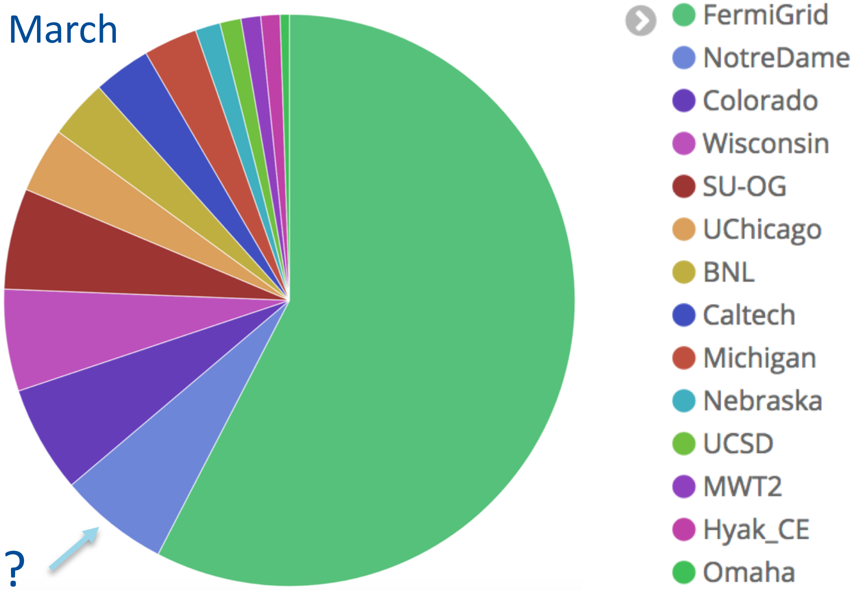
- ...Neither were some EM partners

Recent DESGW work

- Currently doing a re-analysis of all BBH events followed up so far (GW170814, GW170104, GW151226,GW150914)
- Re-partitioned DB, improved fake point sources, packaged outputs (output tarball saves factor of 3-5 space and factor ~50 for file transfer counts)
- Much higher OSG fraction lately (tweaked job requirements)

8 March – 15 March

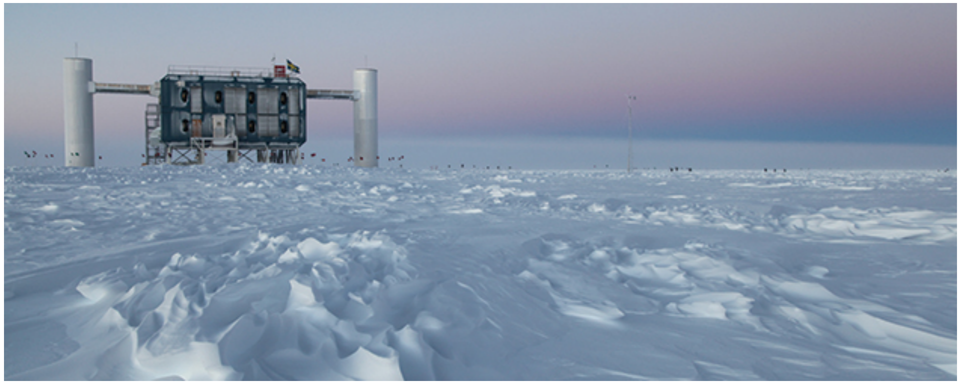
OSG helps insulate from high FermiGrid demand for summer confs.



Spring break effect?

Additional optical follow-up campaigns

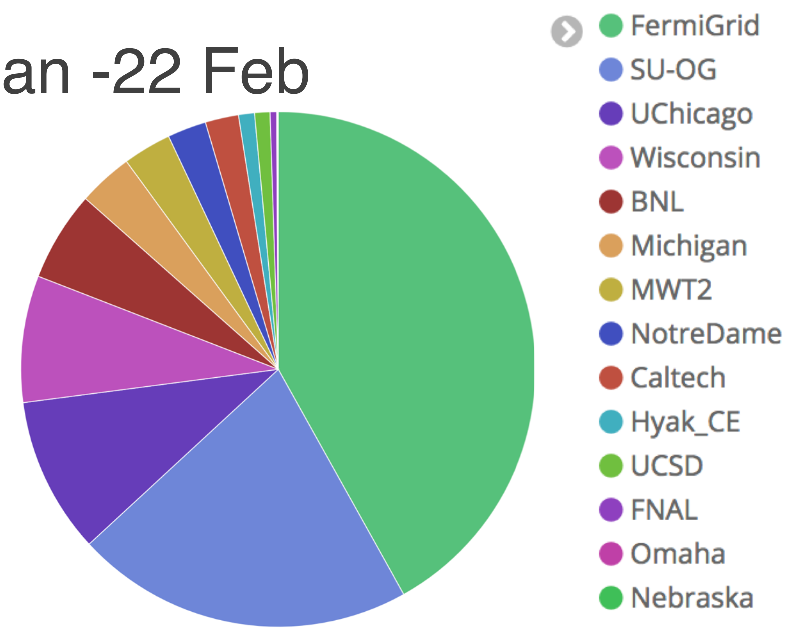
- DES has an agreement with IceCube to do follow-up observations
  - Search for astrophysical sources of ultra-high energy (PeV) neutrinos, e.g. core-collapse supernova
  - IceCube candidates typically localized to a much smaller area on the sky (within one DECam pointing) than LIGO evts
- Followed up two IceCube alerts so far; analysis ongoing
- Uses same pipeline as LIGO followup



Nightly difference imaging

THE DARK ENERGY SURVEY

- Run pipeline on all WS images; no specific target(s)
  - Any number of motivations: TNO searches, Planet 9, IceCube followup, untriggered KNe, the unknown...
- Candidates in 1-2 days (wait for standard image reductions to finish first; run diffimg during day after images taken)
- Ran in RA "windows" this year, 15 Jan -22 Feb
  - Typically 40-50% of images
  - g,r,i,z bands
- Y5.5: Plan to scale up to all images
- **Agreement: program should not affect standard operations**
  - More than half of jobs run offsite



Additional Workflows on the OSG

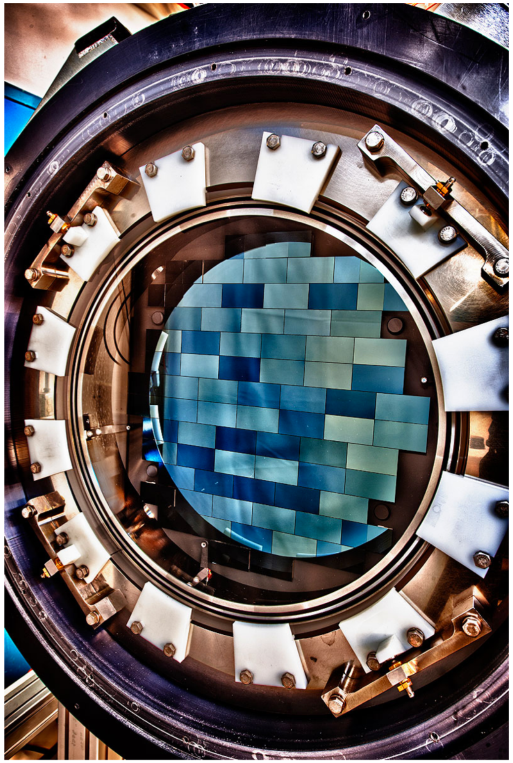
- Several new workflows now OSG-capable
  - SN analysis: < 2 GB memory, long run times (24 h) for MCMC calculations. Highly efficient though
  - Simulations (some fit into usual 2 GB RAM slots)
  - Other workflows require 4-8 GB memory; being run at FNAL right now on dedicated nodes. Not a requirement but difficult to get such high-mem slots in general at a high enough rate
- Other workflows include:
  - Balrog/MOF: galaxy simulation used for detection efficiency studies in cluster analyses, MOF= Multi-object fitter
    - **Different stages of the workflow fit better with different solutions (HPC vs. HTC)**
  - COSMOSIS (cosmological parameter estimation): http://arxiv.org/abs/1409.3409

OSG benefits to DES

- When it works, it's great!
- Biggest issues are the usual pre-emption, network bandwidth
  - Most DES workflows (at least so far) are **very** I/O limited: some workflows transfer several GB of input
  - Still have to copy a lot of images around (currently most SW doesn't support streaming)
- HPC resources may make more sense for other workflows (though having easy ways to get to them is really nice!)
  - Some analyses have MPI-based workflows. Works well when able to get multiple machines (not set up for that right now)
- Strong interest in additional GPU resources. DES will work with FIFE expts on common tools for OSG GPU access

Summary

- Lots of good science coming out of DES right in multiple areas, expect one more half-year of observations
- OSG is and will be an important resource provider for the collaboration
- Opportunistic resources are critical for timely GW candidate follow-ups and TNO searches (i.e. Planet 9)
- Trying to get additional workflows on to OSG resources now
- Very interested in additional non-HTC resources (MPI and GPUs especially.) OSG could be a great resource provider here



Credit: Raider Hahn, Fermilab

BACKUP

DESGW Program Motivations

- **Cosmology**
  - Standard sirens (the GW-equivalent of standard candles)
- **Astrophysics**
  - First observations of NS-NS, NS-BH mergers
  - Evolution of binary systems and their environment
  - Origin of r-process elements in the Universe: what are the dominant r-process sites?
  - Neutron Star equation of state
  - Potential for discovery of new astrophysical phenomena
- **Physics of space-time**
  - Time of flight experiments (including neutrinos!)
  - Tests of General Relativity

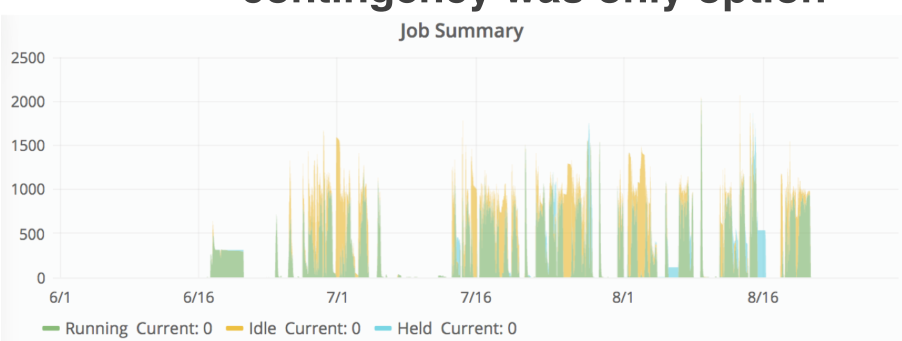
Dataflow and Day-to-Day Operations With Grid Resources

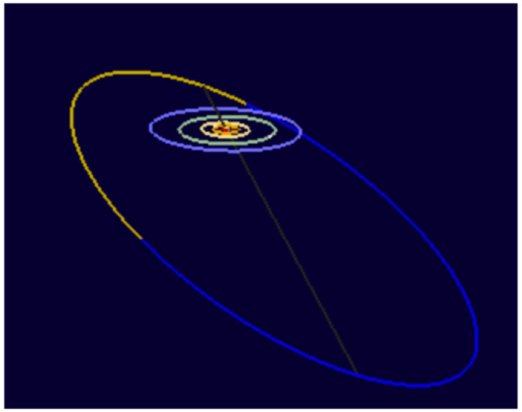
- Dedicated ground link between La Serena and main archive at NCSA (transfer is a few minutes per image)
- Nightly processing occurs at FNAL
  - Submitted from NCSA to FNAL GPGrid cluster via direct condor submission
  - Reprocessing campaigns (additional corrections, etc.) underway at FNAL

Dwarf Planet Discovery

THE DARK ENERGY SURVEY

- DES researchers found dwarf planet candidate 2014 UZ224 (currently nicknamed DeeDee)
  - D Gerdes et al., https://arxiv.org/abs/1702.00731
- All code is OSG ready and is basically the same as SE + diffimg processing with very minor tweaks
  - After diffimg identifies candidates then other code makes "triplets" of candidates to verify that the same thing's seen in multiple images
  - Main processing burst was July-August when FNAL GPGrid was under light load, so >99% of jobs ended up on GPGrid
    - Required resources would have exhausted NERSC allocation; **FNAL w/OSG as contingency was only option**



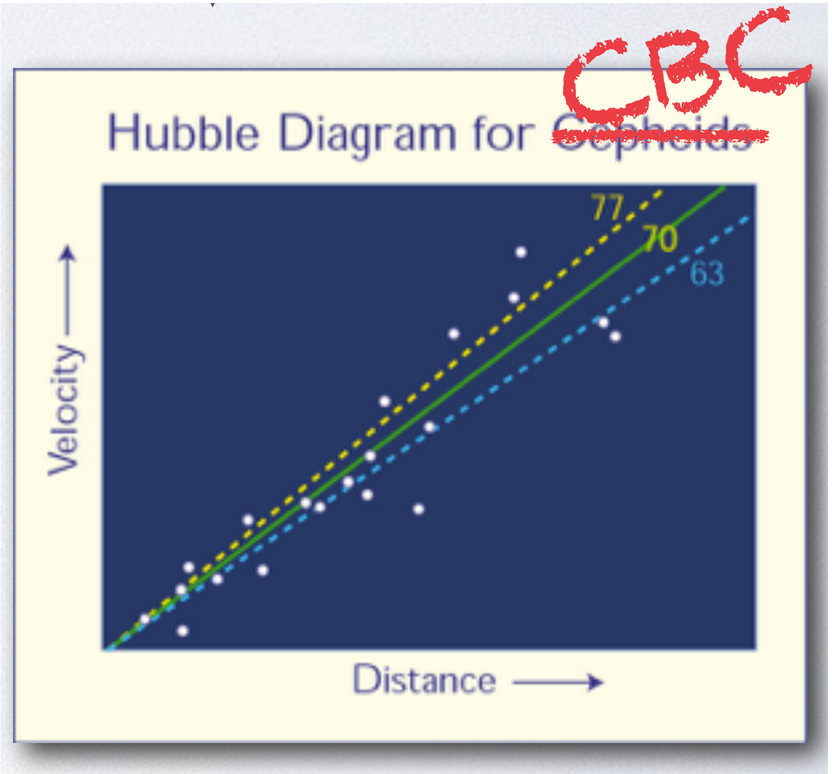


Made at Minor Planet Center site:
http://www.minorplanetcenter.net/db_search/show_object?utf8=%E2%9C%93&object_id=2014+UZ224

Fermilab

Motivation for Optical follow-up of GW events

- The "golden channel" is merger of two neutron stars, with the GW component detected by LIGO and the EM component detected by a telescope
- If one can observe both the GW and EM component, it opens up a lot of opportunities

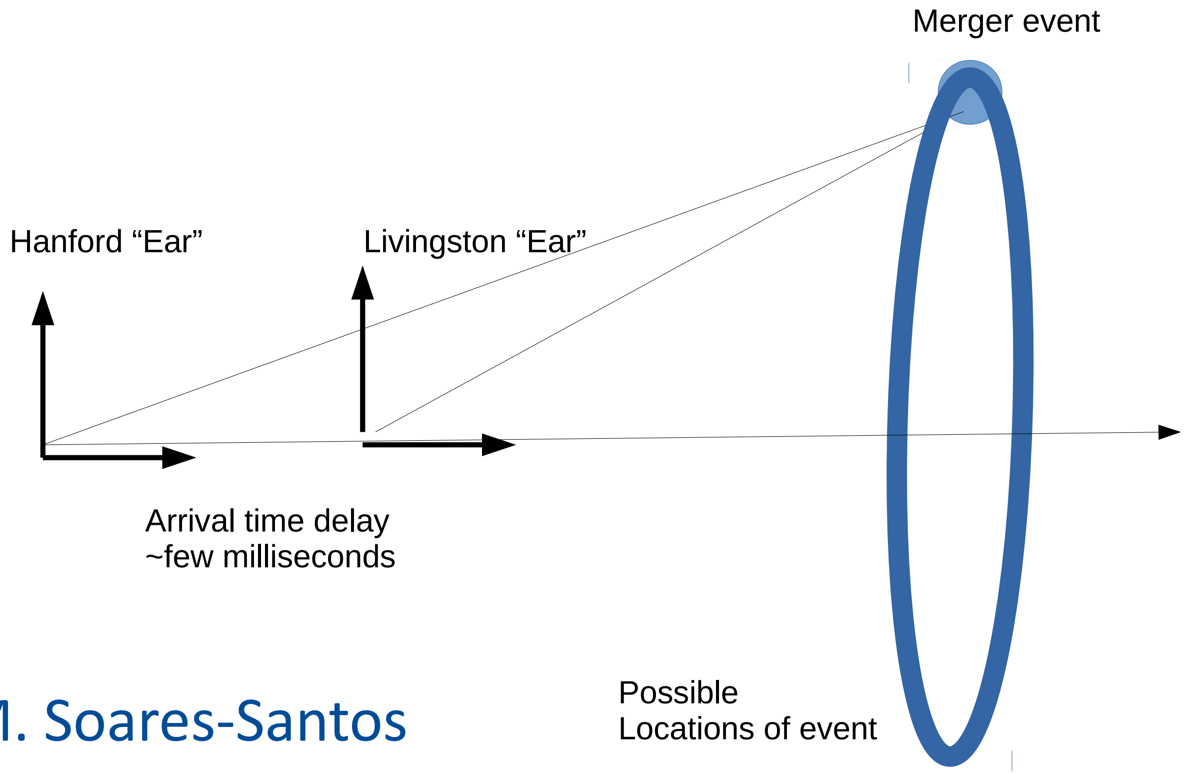


GW gives distance
EM counterpart gives redshift
(from host galaxy)
Together they give a new
way to measure Hubble parameter

CBC = Compact Binary Coalescence

Event Localization

- Similar to how our ears work
- With 2 detectors area can still be hundreds of sq. deg.
- With Virgo detector, would be localized to few tens of sq. deg.



M. Soares-Santos