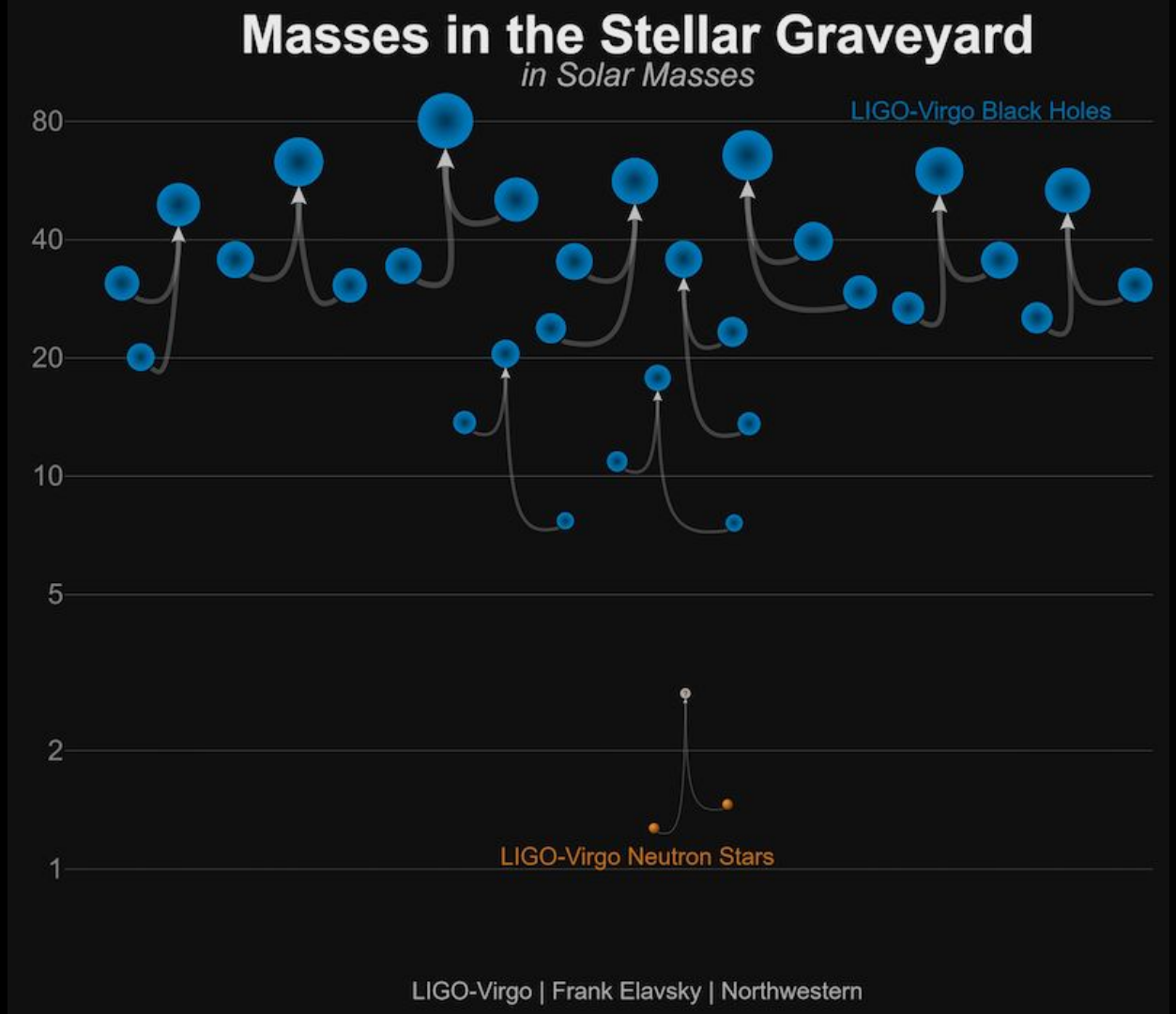
LIGO JupyterLab Server at Cardiff University

(Integrating LIGO into the Cardiff University Supercomputing Wales Hawk cluster)

IRIS Collaboration F2F Meeting - 3/4 April 2019

Paul Hopkins
Cardiff University and LIGO Scientific Collaboration

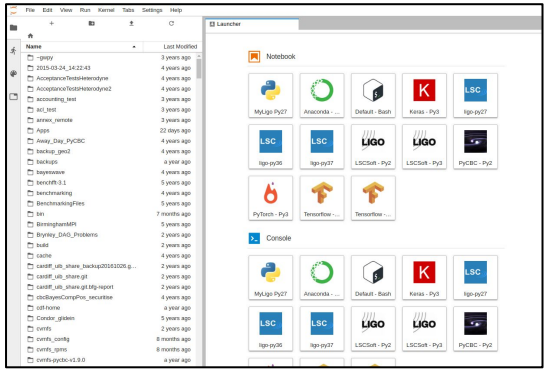

Masses in the Stellar Graveyard
in Solar Masses
LIGO-Virgo Black Holes
80
40
20
10
5
2
1
LIGO-Virgo Neutron Stars
LIGO-Virgo | Frank Elavsky | Northwestern

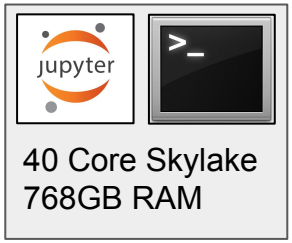
The Cardiff LIGO Data Grid Site

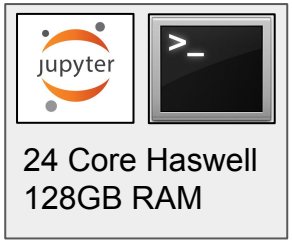
- Cardiff University hosts the UK's LIGO Data Grid site
- One of the 8 main LDG sites
- Provides approximately 10% of LIGO core hours
- Integrated into the Cardiff University clusters:
  - 2013-2019 ARCCA Raven cluster
  - 2019- Super Computing Wales Hawk cluster
- Only LDG site that is not dedicated:
  - Cannot install software into system locations
  - Only minor modifications allowed to compute nodes

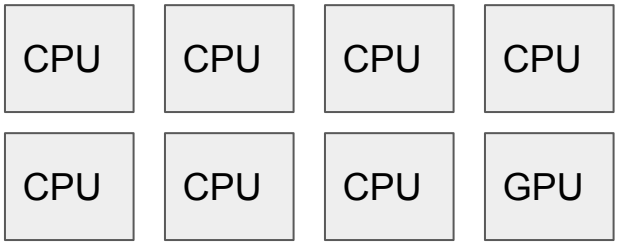
JupyterLab Infrastructure

- Closely integrated with Cardiff HPC Cluster
- Runs directly on LIGO's dedicated login node
- Uses Apache and Shibboleth authentication
- Single-user server also runs on login node:
  - Suited to interactive computing
  - Allows access to the internet
- Potential to use compute nodes via HTCondor
- Access to GPU computing node coming soon
- Deployed using Anaconda









Users, Authentication and Authorisation

- LIGO users added to Raven system with usernames `spxph` or `c123456` ☹
  - LIGO LDAP used to link LIGO account to local user
- SCW Hawk system will use standard LIGO usernames; `albert.einstein`
- Uses LIGO Shibboleth authentication
- SSH access via
  - X509 Certificate
  - Centrally managed SSH key
  - LIGO Password
- Users have 500GB for home directory and also access to other users directories, shared resources etc

Software

- Unlike other LIGO clusters software is not installed into system
- Instead we provide several CVMFS based environments:
  - Singularity images of LIGO Reference OS (CentOS 7) and system software
    - Python 2.7
    - Python 3.4 / Python3.6
  - Newer Conda based environments:
    - Python 2.7
    - Python 3.6 and 3.7
- Users are free to add their own kernels
- System provided and user installed environments available for traditional use

Data

- LIGO proprietary and open data available locally
- Migrating to CVMFS based delivery of data
  - Open Science Grid managed “Stash Cache” co-located at Cardiff
  - Cardiff will cache ~20TB of data on the shared Lustre filesystem
  - Proprietary data uses X509 certificate for authorisation

Thanks