# Astronomy NCRIS Project: Gravitational Wave Data Centre

## Summary (August 2019)

## 1 Overview

The Gravitational Wave Data Centre (GWDC) will provide the infrastructure, training and support to enable gravitational wave researchers nationally to lead the discovery of events from the latest data on an international scale and to maximise the scientific impact of these discoveries. Data coverage includes the Laser Interferometer Gravitational wave Observatory (LIGO) and Virgo detectors as well as pulsar timing data from the Square Kilometre Array (SKA) and precursor facilities. Events range from close binary coalescences to gravitational waves from supermassive black holes with live streaming of data to optimised supercomputing facilities in Australia to enable real-time detections. The GWDC is established at lead host institution Swinburne University of Technology alongside the existing Astronomy Data and Computing Services (ADACS) team that provides a generalised service to the national astronomy community.

## 2 Governance

The Gravitational Wave Data Centre Science Advisory Panel (GWDC-SAP) was formed 13/02/2019 with Paul Lasky (Monash) as the inaugural Chair and membership comprising gravitational-wave researchers from Melbourne, ANU, UWA and Swinburne. A representative from ADACS management (Jarrod Hurley) and Secretariat (Yeshe Fenner) also sit on the panel. AAL has appointed an additional member from the wider astronomy community. The first meeting was held 26/02/2019 with subsequent meetings 02/05/2019 and 01/08/2019. The ongoing expectation is for meetings every two months. Activities to date include drafting the terms of reference, agreeing on the GWDC business plan with AAL, and providing scientific advice for hiring GWDC staff.

## 3 Projects and Goals

The activities of the GWDC will broadly fall within five areas:

1. **Systems Support** – ensuring that access to and use of supercomputing resources are optimised for ingesting, processing and understanding gravitational wave data;
2. **Software Development** – creating and maintaining the software infrastructure to support gravitational wave science workflows.
3. **Optimisation** – mapping GW workflows onto the hardware capabilities of OzSTAR and the use of profiling and parallelisation of source codes to deliver maximum efficiency and speed.
4. **Hardware** – acquiring, provisioning and maintaining the hardware required to facilitate the needs of the data centre;
5. **Training** – ensuring that researchers are sufficiently trained in the use of relevant high-performance computing, data portal and associated resources.

The original GWDC proposal identified a number of prioritised large projects which have subsequently been ratified by the GWDC-SAP for immediate focus. These are:

- Optimise the SPIIR time-domain search pipeline (developed at UWA) to process a live stream of the LIGO O3 data using the GPU capabilities of OzSTAR (aiming to detect GW signals before merger to maximise the chance of pointing EM facilities at the source (e.g. MWA, ASKAP, SkyMapper));
- Establish GWCloud to perform and manage rapid parameter estimation of GW events leveraging the existing Bilby parameter estimation software (developed at Monash) and create a data validation portal for non-LIGO scientists to then be developed into a full Virtual Laboratory for gravitational-wave detection (in collaboration with Monash and Melbourne partners);

- Develop pipelines for the SKA-dimension pulsar data emerging from ultra-broadband receivers at SKA-mid's precursor (MeerKAT) and the Parkes (UWL) telescope. The current file format/software (psrFITS/psrCHIVE) is inappropriate for the millions of profiles being generated every hour;
- Establish OzSTAR as a recognised LIGO Tier 3 Data Centre.

Common to both audio and nanohertz gravitational-wave discovery with close synergy to SKA and precursor facilities will be:

- real-time and offline processing of both time-critical and multi-epoch data streams from large-scale international facilities for users around Australia;
- provision of petabyte-scale storage and retrieval facilities in close proximity to supercomputers;
- optimisation of multi-processor, multi-core, multi-compute node supercomputers and mass storage devices with high input rates;
- advanced use of machine (deep) learning/artificial intelligence algorithms for signal detection, astrophysical inference and visualisation;
- provision of online science gateways (data portals and/or virtual laboratories) to facilitate ready access to data for analysis, discovery and verification of published results.

The goals are closely aligned with those of the proposed SKA regional data centres as well as the activities of the Astronomy Data and Computing Services (ADACS) initiative aimed at meeting the data and computing needs of the national astronomy community in general.

# 4 Timeline

| Activity | Date |
|---|---|
| Establishment of the GWDC Science Advisory Panel | February 2019 |
| Commencement of preliminary activities | March 2019 |
| Official project commencement | July 2019 |
| New staff member added to ADACS team to provide 1.0 FTE towards GWDC activities | July 2019 |
| Funding agreement signed | September 2019 |
| Staff recruitment | July-December 2019 |
| LIGO Tier 3 Data Centre Status for OzSTAR | June 2020 |
| OzSTAR storage expansion to minimum 10PB | June 2020 |
| Completion of the first phase of the project | December 2021 |