High Performance Computing Working Group:

The Future of eScience and High Performance Computing for all Australian Astronomers

July 2010

Working Group Members:

Darren Croton (Chair)
Matthew Bailes
Lindsay Botten
Geraint Lewis
Tara Murphy
Brian Schmidt
Peter Quinn

Astronomy Australia Ltd.

Astronomy Australia Limited http://astronomyaustralia.org.au

Contents

Executive summary

eScience is an emerging research frontier that uses high performance systems and networks to exploit complex data that would otherwise be unmanageable in a smaller research setting. The Australian Government has identified eScience as an essential component of our future research directions and has invested heavily in establishing an eResearch structure to enable the Australian scientific community.

Australia's next generation flagship telescopes - SkyMapper, ASKAP and MWA - will return unprecedented volumes of data that will require processing, tagging and storage before analysis can begin to extract the science. This will require optimal use of our High Performance Computing centres, such as the NCI National Facility at ANU and the Pawsey High Performance Computing Centre for SKA Science in Perth. The data challenge will be so great that raw supercomputing power alone will not be enough. A carefully planned and executed "National Integrated Astronomy Data Fabric" solution to our coming data bottleneck will be essential.

This Data Fabric should be layered with high performance hardware and software as a foundation, database middleware to appropriately tag and structure the data, and high-level Virtual Observatory (VO)-enabled tools and interfaces through which the astronomy community can efficiently access and manipulate the data. In particular, a lack of funding for middleware database experts is identified as a weakness in the current data strategy and is specifically flagged for attention.

The layers of the Data Fabric can, and should, be designed to integrate data from all wavelengths equivalently, seamlessly federating our most important radio and optical datasets for use by the astronomy community. This spirit of multi-wavelength cooperation has already begun under the recently funded ARC Centre for Excellence for All-Sky Astrophysics (CAASTRO), and we expect CAASTRO to play a role in the data initiative described here. To implement the National Integrated Astronomy Data Fabric, the astronomy community must partner with the appropriate Government eResearch initiatives.

1. What is HPC and eScience?

eScience is enabled by High Performance Computing

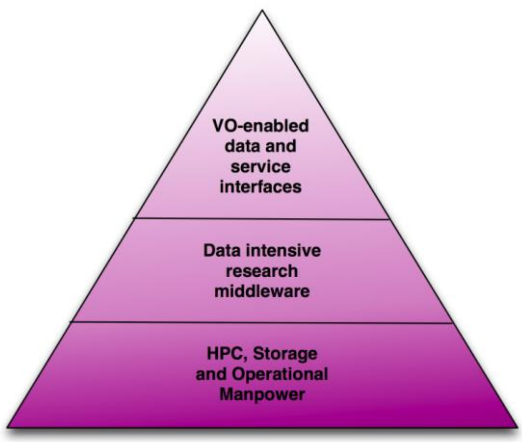
eScience employs High Performance Computing (HPC), often in highly distributed network environments, to understand and solve complex science problems. eScience encompasses data acquisition and management, simulations and modelling, and data analysis and visualisation. HPC enables eScience.

In modern eScience, the "high performance" requirements apply to processing power (computing), storage and networks. Different areas of astronomy eScience have different dependencies on one or more of these components. For example, radio astronomy data flow and processing are dominated by I/O performance limits whereas computational astrophysics is often limited by memory size and processing speed. When considering the future of eScience in astronomy, we need to take account of a range of high performance computing needs, and design and operate our high performance systems appropriately.

The definition of High Performance Computing is constantly evolving. Moores' Law, the demonstrable principal that the computing power per dollar invested doubles approximately every 1.5 years, drives the supercomputer applications of yesteryear onto the desktops of today. Here we restrict our discussions to infrastructure that requires facilities typically in excess of what a small department could reasonably supply, expected to cost hundreds of thousands to many millions of dollars. Our definition of HPC encompasses both hardware and software, and the human expertise to implement, optimise and maintain it for use by the community.

The eScience pyramid

eScience infrastructure can be grouped into three distinct layers linking the eResearcher to the high performance resources. At the base of the resource pyramid is the high performance hardware and software (processing, storage and networking) that must be operated and maintained by dedicated human resources. The second tier consists of essential middleware – a data and service fabric that is closely connected to database technologies for organising and servicing queries to and from the base level high performance systems. The top tier is the user interface technologies, standards and tools needed for scientists to interface to, and explore, the structured content delivered through the bottom two layers.



While most people would naturally cite tier one and three in their list of infrastructure needs, the middle tier (data query and organisational fabric) is often overlooked. As modern astronomy database sizes grow to billions of objects with thousands of attributes, we face the same (or greater) data management challenges as Google and large financial systems. To meet these requirements we need both high performance middleware solutions and the experienced people to design, deploy and operate it for the astronomy community.

It cannot be overstated that people play a critical role in developing and operating eScience infrastructure at all three layers of the resource pyramid. Without a funding mechanism that is persistent and can sustain operational manpower in terms of availability and curation of resources, it will not be possible to support the real time and frontline research requirements placed upon Australian astronomical facilities at the national and international level.

The government's eScience research structure

The Australian government has a two-fold investment in eScience infrastructure in Australia:

- The development of the Platforms for Collaboration components of the National Collaborative Research Infrastructure Strategy (NCRIS), providing a total of $82M between 2007 and 2011.
- The development of eResearch components of the Super Science Initiatives, providing a further $312M between 2009 and 2013.

Both investments underline the critical importance of eResearch (and specifically eScience) infrastructure to future research competitiveness. This investment *"is intended to enhance research collaborations, assist researchers to manage massive data sets, and provide supercomputing and analysis tools that enable Australian researchers to tackle the complex, national and global issues needed to secure Australia's future."*[1]

The ongoing development of the eResearch landscape is co-ordinated by the Australian eResearch Infrastructure Council across four key areas of functionality. Each area involves an identifiable community of practice with its own specialised expertise:

Function	Organisation	Funding	Purpose
Data Federation ($75M)	Australian National Data Service (ANDS)	NCRIS and Super-Science	Establish an Australian data commons to support the data federations needed by 21st century research (e.g. metadata tagging, data "mining", data curation).
Research Workflows ($69M)	Australian Research Collaboration Service (ARCS)	NCRIS and Super-Science	Develop and deploy collaboration tools and services (physical or virtual) that support research teams and research workflows accessing instrument, computing and data resources nationwide.
Advanced Models ($156M)	NCI and Pawsey Super-computing Centres	NCRIS and Super-Science	Move towards peta-scale supercomputing in support of climate, earth systems, science and water research, and provide a new peta-scale computing resource targeting SKA and other modelling demands in nanotechnology, biotechnology, geo informatics, engineering, atomic physics, chemis-

[1] https://www.pfc.org.au/bin/view/Main/WebHome

			try, and mineralogy.
Seamless Reach ($39M)	Australian Access Federation (AAF) and AARNet	DIISR and Super-Science	To provide sector wide researcher authentication services supporting single sign on functionality (AAF). To provide research enabling advanced network services (AARNet).

The astronomy community has been well supported by the "Advanced Models" component of the Government's eResearch infrastructure, through the funding of both the NCI Facility at ANU and the Pawsey High Performance Computing Centre for SKA Science in Western Australia (discussed below). This sets a solid HPC foundation for the eScience pyramid, upon which the upper layers can solidly be built. However, as we will later argue, there is a need for the astronomy community to more closely work with the other Government eResearch components (or equivalent), namely ANDS and ARCS, to implement middleware database (most importantly) and VO solutions. Only then will we be fully equipped to manage and exploit the coming massive data products that will be collected from the SkyMapper, ASKAP and MWA telescopes.

2. The current state of astronomy HPC and eScience

Who uses High Performance Computing?

Astronomers are active users of High Performance Computing. Within the Australian community three primary groups can be identified:

- Theoretical/Numerical Astrophysics: Numerical simulations of astrophysical processes (e.g. cosmology, galaxy evolution, star formation) often require vast CPU and memory resources to reach the required level of accuracy. Numerical astrophysics is a rapidly expanding area within the Australian community, with several new hires in the past few years distributed across a number of Universities. Theorists commonly use Swinburne's "Green Machine" supercomputer or their own local facilities to perform simulations. Theorists tend to run either "embarrassingly parallel" codes, where multiple instances of the same application run independently, or more complicated message-passing codes, that take multiple processors and high speed interconnects to increase the dimension of the problems they can address. HPC facilities must be designed to accommodate both uses.

- Radio Astronomy: The radio community has a wide range of computing needs: collecting and preprocessing data straight from the telescope, transporting and post-processing data for signal analysis, and storing for later retrieval and analysis. Within radio astronomy, the major supercomputer users have been pulsar astronomers at Swinburne and the Australia Telescope National Facility (ATNF), and for Very Long Baseline Interferometry (VLBI) where the supercomputer acts as a correlator. Radio HPC needs can vary widely. For example, the data volumes coming from the Australia Telescope Compact Array (ATCA) can be easily handled by a Linux workstation as the number of baselines is very small (15). In contrast, HPC clusters are routinely used at the Parkes radio telescope to process vast volumes of live radio data.

- Optical Astronomy: In the last decade or so the majority of Australian astronomical supercomputing has been undertaken by theorists and radio astronomers. However, optical astronomers are now expanded their need for computing power as the size and complexity of their data has increased. Galaxy surveys conducted at the Anglo-Australian Telescope (AAT), such as WiggleZ and 6dFGS, use local HPC for data analysis and theoretical modelling. ANU's SkyMapper telescope is surveying the southern sky and collecting an unprecedented 500 TB of data that will require processing, storage and subsequent analysis using the NCI National Supercomputer Facility at ANU.

HPC resources tied to national telescope facilities

Each national telescope facility tends to host its own computing resources, customised for the needs of the telescope and for the data it acquires. Our national telescopes are key to many large Australian led international projects and must be resourced appropriately. The HPC needs of these facilities have traditionally ranged from modest to very high.

Existing national telescope facilities with HPC

- Parkes and Narrabri: The ATNF currently operates a number of major Australian telescope facilities, including the Australia Telescope Compact Array (ATCA) at Narrabri and the 64 metre Radio Telescope at Parkes. Both Parkes and Narrabri have similar HPC systems in place. The Parkes telescope has two moderate clusters in operation for real-time and offline processing of voltage data. The

APSR cluster was funded by Swinburne University of Technology and is used in three main ways: (A) For real-time and offline Very Long Baseline correlation, coordinated by the CSIRO Astronomy and Space Science (CASS) unit and Curtin Universities. (B) For real-time processing of spectral data - these data are both written to LTO tapes on site and streamed to the Swinburne supercomputer via the Gb link. (C) For real-time coherent dedispersion of packetised data from the Digital Filterbank v3 (DFB3). In addition, a smaller Infiniband-linked cluster exists that employs GPUs for real-time coherent dedispersion of a single 400 MHz band. CASS has deployed a large quantity of high performance disk at Parkes primarily for Very Long Baseline Interferometry. The HPC cluster at Narrabri is of similar dimension to that at Parkes and with significant disk space, funded by the Curtin University of Technology.

- AAT: Our national optical facility is the Anglo-Australian Telescope (AAT), a 3.9 metre telescope that hosts the AAOmega multi-fibre spectrograph. This instrument and its predecessor, the two degree field spectrograph, have been key in the success of a number of large galaxy survey programs of international impact: GAMA, WiggleZ, 6dFGS, 2dFGRS. The AAT provides minimal computing facilities, enough for telescope operations, data processing, and sufficient resources to support the storage and analysis requirements of the staff at the Observatory. For more intensive HPC needs, data is taken offsite by the astronomer to be processed elsewhere.

Telescope facilities coming online now

- ASKAP: The Australian Square Kilometre Array Pathfinder (ASKAP) will be a next-generation radio telescope incorporating novel receiver technologies and leading-edge ICT systems. ASKAP will be a world-class telescope in its own right as well as being a pathfinder instrument for the full Square Kilometre Array. It will comprise an array of 36 antennas each 12m in diameter, capable of high dynamic range imaging and using wide-field-of-view phased array feeds. The telescope will produce about 75TB data per day. To process the observations will require approximately a petaflop of processing power for full angular and spectral resolution using the Pawsey HPC Centre for SKA Science, and about one tenth of that for decreased angular resolution in spectral line observations. The storage for observed data and derived science products could reach about 100PB/year at Pawsey, but in practice some data will be discarded to fit within a realistic budget for storage cost.

- MWA: The Murchison Widefield Array (MWA) is a low frequency radio interferometer being built at the Murchison Radio-astronomy Observatory (MRO), alongside ASKAP. The MWA has been designated an SKA Precursor by the SKA Program Development Office and science goals for the instrument will focus primarily on the early Universe epoch of reionisation, solar and heliospheric science, galactic and extragalactic radio astronomy and radio transients. The MWA will use substantial HPC facilities, both for on-site real-time imaging and calibration processing and for large-scale image data storage and post-observation processing. A GPU-based cluster of approximately 64 nodes will provide the real-time imaging and calibration. The output of this system will be transported via high-speed optical fibre to Perth at a rate of approximately 2 Gb/s, where the images will be stored at the Pawsey HPC Centre for SKA Science. The volume of annual achievable data from the MWA will be a fraction of a PB, initially substantially less but increasing as the instrument matures.

- SkyMapper: SkyMapper is the Australian National University's new 1.3m telescope that is starting to undertake an unprecedented survey of the Southern Sky. The richness and scale of this 500

Terabyte dataset is expected to attract thousands of users and millions of queries from across the globe. High Performance Computing is central to the ability to analyse and serve this dataset. Each day's data will be transferred via a Gigabit link to the NCI Supercomputing Facility in Canberra (see below) where it will be archived using the mass data storage facility. A linked computational facility to the mass-storage system will process the data, creating a 10 Terabyte database in the process. This NCI supported database will have the ability to serve key scientific information of more than a billion objects across the southern sky, as well as provide access to more than 500 Terabytes of raw and reduced images to the astronomical community.

Dedicated supercomputing facilities and support - current

Aside from the HPC resources available through the use of national observing facilities, there are currently two other main sources of community supercomputing available to Australian astronomers. These are commonly used by theorists and by observers for additional data processing/analysis:

- The NCI National Facility at the Australian National University: The Sun Constellation supercomputer at ANU has 11,936 cores (8 cores/node; Nehalem) with at least 24Gb memory/node, a 500Tb file system and Infiniband interconnect. Its theoretical peak performance is 140 TFlops. Access to the supercomputer is available to the astronomy community through a proposal process. Ambitious projects with potential high impact results may be given special allocations of time after review. In 2010-11 NCI is contributing 1,000,000 CPU hours to the astronomy community under its Specialised Support program which is also funding two staff located at the NCI National Facility and at Swinburne University. In addition, NCI is supporting the ARC Centre of Excellence, CAASTRO, with a further 4,000,000 hours per annum together with a substantial allocation of storage.

- The "Green Machine" at Swinburne University: The Centre for Astrophysics and Supercomputing hosts a 10 TFlop supercomputer (theoretical peak) with 1,280 cores (8 cores/node; Clovertown), 16Gb memory/node, a 250Tb file system and Infiniband interconnect. Access has historically been made to experts in the community via request. For two years starting 2010 Swinburne is donating 1,000,000 CPU hours/year to the astronomy community. Swinburne will be investing $3M in updating the Green Machine in 2011, which includes $1M (EIF funded) for the gSTAR GPU supercomputer (see below).

Local HPC resources provide astronomers with a second line of computing power, usually available based on institutional affiliation. Examples most related to the astronomy community include:

- iVEC: Astronomers at iVEC affiliated Western Australian universities and CSIRO have access to four complementary supercomputers: (1) 512-core SGI Altix XE. This is a general purpose Linux cluster of commodity Intel Xeon-based servers, connected by Infiniband. (2) 328-core Cray XT3. This is a specialised AMD Opteron based system, with a proprietary low-latency SeaStar network arranged in a Torus topology. (3) 192-core SGI Altix. This is a specialised "shared memory machine" Linux system based on Intel Itanium2 processors, with a proprietary low latency NUMA-link interconnect and 384 GB of memory. (4) 160-core SGI Altix XE. This is a general purpose cluster of commodity Intel Xeon-based servers, connected by Infiniband. In addition, iVEC has a Sun/Oracle peta-scale data store for storage and dissemination of large data sets. iVEC is building a separate SAN-attached virtual server infrastructure for storing databases such as those used for searching metadata, where it is not appropriate to store them on a hierarchical storage system.

- VPAC: The Victorian Partnership for Advanced Computing (VPAC) overseas HPC hardware and software, advanced visualisation and collaboration tools, and grid resources to support the Victorian research community, and has expertise across a broad range of sciences and related technologies. Astronomers at universities affiliated with VPAC have access to a range of computing options, the most powerful of which is the "Tango" supercomputer. Tango hosts 888 cores (8 cores/node; Shanghai), with 32Gb memory/node and Infiniband interconnect. VPAC has recently installed a small 5-node GPU cluster attached to Tango, serving partly as a pathfinder for gSTAR (below) and partly as a test bed for further GPU upgrades. VPAC manages a number of university supercomputers, including Green at Swinburne, and provides dedicated HPC support for them to the wider user community.

Other machines inhabit Universities throughout the country, available to staff and their collaborators. Although much science can be undertaken on such clusters, "big science" projects typically require astronomers to seek out our national facilities or go overseas.

Dedicated supercomputing facilities and support - future

In the past year the Australian Government and its eScience branches have funded a number of astronomy initiatives.

- Pawsey ($80M): The most significant recent HPC investment by the Government is the Pawsey High Performance Computing Centre for SKA Science, to be located in Western Australia. The specs for this new machine have not yet been made available but are expected to place it in the top 20 machines in the world by 2013 (i.e. Peta-scale). A significant fraction of the Pawsey Centre resources (25%) will be dedicated to support radio astronomy (data collection, processing, storage, and modelling), with a focus on MWA and ASKAP. It will serve as a HPC pathfinder for the Square Kilometre Array that is expected to have supercomputing needs at the Exa-scale.

- gSTAR ($1M): The GPU Supercomputer for Theoretical Astrophysics (gSTAR) is a specialised computing cluster to be based at the Centre for Astrophysics and Supercomputing at Swinburne University starting mid-2011. gSTAR is expected to have a theoretical peak performance of 600 Tflops of computing power and will be connected to a peta-scale storage system. The machine will open approximately 40% of Swinburne's total supercomputing resources to the Australian astronomy community, enabling a range of science applications, including LIGO data analysis, next generation N-body and hydrodynamic simulations, simulation of MWA data and epoch of reionisation (EoR) pipeline development.

- NCI support personnel (ANU/Swinburne; $600k total): The NCI is funding two HPC two year support positions, one based at the NCI facility at ANU, and the other in the Centre for Astrophysics and Supercomputing at Swinburne University. The two positions will commence late-2010 and will provide expertise to the community in parallel and GPU computing, mass data processing and analysis, and data visualisation.

- NCI will contribute its support for the ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO) from 2012 through a substantial allocation of resources on its new petascale system, the infrastructure for which is being procured through one of the Super Science initiatives.

During its one-year tenure, this working group was involved in securing the latter two of the above HPC-related initiatives. In addition, the working group has been promoting a new National Integrated Astronomy Data Fabric initiative, described below. This initiative has included the organisation of a national data workshop (6th May 2010) to assess community HPC and eScience needs (with a second to follow in September 2010), and a series of white papers written by the leaders of the most HPC intensive radio and optical projects.

Notable mentions

- VLSCI: The Victorian Life Sciences Computation Initiative (VLSCI) is a $100M government funded (national and state) supercomputing facility of peta-scale proportions, with a focus on servicing the life sciences, including biology and medical research. It will be housed and operated by the University of Melbourne and VPAC, with phase one beginning 2010. We expect limited or zero general access to this machine for the wider astronomy community. However, those affiliated with the University of Melbourne itself are expected to get some time based on internal University allocations.

- CSIRO: CSIRO operates a new 1024 core GPU-based supercomputer, housed in Canberra, with a theoretical peak performance of 256 Tflops. The GPU supercomputer supports computational and simulation science research within CSIRO, including space science and astronomy.

3. What's missing? Working group recommendations

The Government has implemented a series of eScience initiatives that support Australian eResearch and within which astronomers can exploit Australia's existing and future telescope and HPC investments. Throughout our one-year tenure, the HPC working group has consulted with the community to construct a big picture view of the HPC and eScience needs and wants of astronomers within Australia. From this process we have identified two major areas of need in the currently funded national infrastructure, which must be addressed if we are to maximally exploit the significant investments already made. These are (1) the need for a national integrated astronomy data fabric, linking theoretical, radio and optical astronomy data resources, and (2) the need to invest more heavily in "human" infrastructure, to enable the use of increasingly complex HPC and eScience systems by all Australian astronomers. We discuss both of these below and provide recommendations. We also give a "watch list" of secondary HPC-related issues that may require future attention.

The need for a National Integrated Astronomy Data Fabric

If the Australian astronomical community is going to maximise its scientific return from new astronomical survey facilities like SkyMapper, MWA and ASKAP, then we will have to fund, build and operate three layers of eScience infrastructure that interface these facilities to the community. These layers, introduced previously as parts of the eScience pyramid, are:

- High performance hardware and software facilities for the processing and storage of astronomical data, including operational manpower, user support and curation services.
- A data intensive research middleware that joins database systems, high performance storage and high performance computing into a services architecture.
- A VO-enabled interface layer that connects the wider multi-wavelength research community to the underlying services and database middleware.

These three layers form what we call the *national integrated astronomy data fabric*.

(A) HPC/Data Facilities

With data flow rates of hundreds of Gb/s, annual data accumulation rates of 10 PB and processing power in the 100 TFlops range, even SKA-pathfinder class facilities warrant HPC-class resources for their day-to-day operation. The traditional notion of a batch-oriented, non-real-time supercomputer centre will have to be married to the needs of real-time data flow if national-scale HPC resources are to form an operational part of modern radio and optical telescopes. The management of real-time data, the curation of data sets, the quality control of routine processing, and the support of scientific users are all increasingly important services to be provided by a HPC facility in addition to the more "traditional" needs of large scale HPC users.

Funding: While HPC hardware is naturally supported under the Government's "Advanced Models" scheme (currently through NCI and Pawsey), alternative funding channels (e.g. ARCS or ANDS) may be needed to address some of the new unexpected or untested hardware and software data challenges, and also the increasingly complex service-level HPC facility needs (through e.g. ARC, LEIF, University). Operational costs and the like are often not included in federal-level support and must be funded through State Government or universities.

(B) Database Middleware

The primary scientific output from a survey telescope is a list of detections. These detections satisfy complex selection criteria executed via optimised algorithms running on high performance, and often real-time, dedicated computing systems trawling through large amounts of digital data. This list could be short (e.g. a unique or transient event) or it could contain billions of sources. With each detection there is a collection of defining and extracted information (metadata) that will enable the source to be assessed, classified and compared to other objects in other lists. The number of metadata terms (columns in a spreadsheet) can easily range from tens to tens of thousands, each term itself containing large amounts of data.

The efficient and rapid finding, extracting and comparing of objects stored as arrays of metadata has been the driving force behind the development of large commercial database systems that lie at the heart of the banking industry, airlines, online commerce and search engines. The use of database technologies and languages has only recently become commonplace in the astronomical community.

The Sloan Digital Sky Survey (SDSS) is one of the first public large-scale astronomical survey projects to make extensive use of database technology. The SDSS project provides an SQL (standard database query language) interface to users who wish to access their database of 350 million objects. This database has supported 70 PhD projects, 2000 articles in refereed journals and 70,000 citations since the first data release in 2003.

As survey projects grow from SDSS-scale to the billions of objects (with thousands of billions of pieces of metadata) expected from telescopes like the Large Synoptic Survey Telescope (LSST) and the SKA, new and efficient database technologies will be essential to extract the maximum scientific return from the large investments made.

Funding: This kind of middleware falls most naturally under the Government funding umbrella of both “Data Federation” (ANDS), to optimise the data structures through metadata tagging, and “Research Workflows” (ARCS), to create the hardware and software systems that enable efficient and accurate cross-referencing of metadata and access. It cannot be understated that such expertise is not a natural attribute of the average astronomer, and the community must seek out and employ skilled specialists to implement and maintain each of our precious datasets at this critical level.

(C) VO-enabled interfaces

Databases are fed by data archives. The size of digital archives being developed in the international community across a range of physical sciences, and the associated computing power needed to fill them, now represents a significant fraction of the cost of research infrastructures. This means that resources have to be tiered and centralised and users will be required to remotely access data and services on data.

Astronomy has taken on a leading role in defining the systems and standards needed to enable remote access and services across a large range of data types throughout the efforts of the International Virtual Observatory Alliance (IVOA). Though IVOA standards, archives can now be designed that are interoperable across multiple projects and wavelengths and which can be uniformly accessed through a new generation of VO-enabled tools. The VO standards sit above the specifics of individual archives. The database technologies and the individual hardware and software systems of a given data centre are effectively hidden by VO interfaces but the research effectiveness of VO queries depends critically on the efficiency and performance of the underlying database and computing architectures.

Funding: Enabling federated multi-wavelength data access most naturally falls under both “Data Federation” and “Research Workflows” in the Governments eResearch landscape (ANDS and ARCS): to design a

VO-based system that can be used by both optical and radio astronomers which conforms to the appropriate VO standards and utilises the appropriate VO-enabled tools. This layer of the data fabric should focus on unifying existing data sets and integrating new data from the SkyMapper, ASKAP and MWA telescopes for community access.

In the past three years, Australian astronomy has been very fortunate by receiving significant investments of research facilities funding from Federal and State Governments. In particular, we have received more than $300 million for the funding of ASKAP, the MWA, and the Pawsey HPC Centre for SKA Science. Our task now is to connect the data flow from the telescopes to the processing and storage resources of centres like Pawsey, and then to enable research on those data resources by the Australian and international astronomical community. While we may be currently “hardware rich”, we are currently “middle-ware” and “wet-ware” (i.e. people) poor. Our challenge is to design, build and operate all three levels of the research infrastructure pyramid identified above.

The principal missing ingredients for a national data fabric are:

1. the middleware infrastructure, focusing on the database technology that links the low-level HPC to the high-level user interface,
2. the human expertise to implement and optimise each level of the infrastructure pyramid, and
3. the ongoing infrastructure support (hardware, software and human) for data curation, needed to maintain the data resources for users into the future.

These needs apply equally to radio and optical astronomy, and we now have a unique opportunity to coordinate and join the essential infrastructure for both. This infrastructure would naturally focus around the strong regional resources being developed at the NCI (optical) and Pawsey Centre (radio). High performance links provided through existing or new NBN resources would then couple both major resource centres to key research universities and departments with multi-wavelength research coordination provided by the ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO).

Recommendation #1:
Leveraging the regional eScience investment at NCI and Pawsey, and with the astronomical multi-wavelength eResearch coordination provided by the recently funded ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO), Australia needs to build an astronomical data fabric that links high performance resources through appropriate data middleware and networks to create new opportunities for discovery by Australian researchers based on data flowing from telescopes like SkyMapper, ASKAP and MWA.

Increased support for "human" infrastructure

Astronomers deal with some of the most fundamental question in the Universe, and large telescopes and HPC systems are often employed to answer them. In today’s astronomy, scientific answers are almost always extracted from observational data using the computer as a processing tool. This does not necessarily require the use of HPC; simple (or cleverly phrased) questions often only need simple datasets to answer and this can be done with minimal computing power. However more complex questions, or those that need the statistics of large datasets to answer, are not so fortunate. Hence, HPC has become a necessary tool in the modern astronomers toolkit, and those without the requisite HPC skills are gradually being left behind as the size and complexity of the data increases.

The ability to use modern HPC takes a number of forms. First and foremost are the skills to code in low and high-level programming and scripting languages in a way that is optimised for many-core (or "embarrassingly" parallel) supercomputing. Equally important is the knowledge to write efficient I/O routines for fast reading and writing of large data files (the computer will sit idle if the data cannot be read and written at least as fast as it can be processed). The ability to visualise data output is becoming more important as data becomes multi-dimensional, as is the knowledge to use database languages like MySQL to mine large datasets, either locally or online. Computer architectures are changing, and GPU technology is becoming increasingly common in high performance facilities. Astronomers face a great challenge in porting their existing codes to keep pace with the new opportunities such architectures bring.

Due to the rapid rate at which HPC has advanced in recent years, and the reality that astronomers cannot be "experts at everything", there is a growing realisation that the community requires additional support personnel in targeted areas who have the necessary skills to bridge the knowledge gap for those who need it.

The future challenge of astronomical supercomputing is, therefore, not necessarily access to computational facilities, but rather to the expertise required to migrate into, and to take significant advantage of, the next generation of supercomputers. Specialist expertise in HPC is located in a small fraction of the astronomical community, and then in individuals whose knowledge is focused upon computational aspects related to their research projects. Furthermore, most expertise is centred upon CPU-based supercomputers, with virtually no significant expertise in the programming of the architecturally more challenging GPU-based systems. Therefore, to fully take advantage of new investments in computing hardware, it is essential that there is sufficient investment to build a knowledge base of computational expertise that is accessible to the community, allowing efficient migration onto and between high performance computing environments.

Funding: The need for specialist support personnel, with a focus on "ground-level" HPC science by astronomers (programming, code optimisation and porting, etc), falls somewhat between the gaps in the Government's eResearch infrastructure scheme. Such support positions will probably need to be sought opportunistically, through ARC, LEIF or University-level funding schemes.

Recommendation #2:
Australian astronomers are far more likely to have difficulty getting access to experts with HPC programming experience than they are getting access to enough HPC hardware. We recommend that AAL look for opportunities for investment in HPC support personnel for the community.

General recommendations and needs: keeping pace with the future

We conclude our recommendations by highlighting a number of related key areas that will require constant attention in the coming years.

Storage

The cost of data storage has plummeted dramatically over the past twenty years. A 1TB disk is now less than $100, whereas in 1990 1GB disks were $10,000. This has created fantastic opportunities, and ambitious astronomers are continually planning surveys with increasingly insatiable appetites for data storage. Petabytes are likely to be the norm and Petabyte programs are already being undertaken at the Parkes telescope.

Need #1:
Astronomers will require secure storage for Petabyte datasets. By secure we mean that the data should be backed up and have disaster-recovery plans. Surveys underway or planned will cost millions of dollars to undertake, and are a potential rich source of discoveries going into the future. It is important that National Facility providers factor into the cost of surveys the price of data curation.

Data Transfer Rates

Despite the 100,000-fold decrease in the price of storage capacity since 1990, the transfer rate from hard drives has only increased by a factor of ~50 in the same twenty year period. In other words, the ratio of capacity-to-transfer rate has decreased by three orders-of-magnitude! In practice, this means that although we can afford to store Petabyte data (~150K/PB) it takes an impracticably long time to read it back. If stored on DLT-S4 tape technologies, 5TB/day is the maximum retrieval rate per tape drive. So if an astronomer wants to simply move a Petabyte of data from tape to a supercomputer, a dedicated tape drive would require over six months to complete the task. Astronomers can reduce these times by replacing tape robots with "spinning" hard disk drives, but the relative cost is five to ten times more. A very modern disk array can achieve ~GB/s, making it possible to read back Petabyte datasets in less than a fortnight.

Need #2:
It is important that the data centres to which Australian astronomers have access have sufficient data transfer rates to allow rapid reprocessing of their Petabyte datasets. The data centres will need to be in close proximity to supercomputers, or possess multiple-10Gb links between them and the processing facilities (see below). The breakdown of tape-based and disk-based storage capacity is something that will require careful planning to minimise costs.

Network Access

The Internet has greatly facilitated the use of HPC by external users, who no longer need to be in the same building as the computing facility, or even the same country. But data transport and temporary storage is a constant barrier to effective HPC usage. A user with data wanting to be processed on a distant supercomputer needs both a local data retrieval system (i.e. a tape robot) and a high-speed link.

Need #3:
It is important that the Universities and research centres have high-speed links to the Internet/AARNET, and the ability to unload large datasets without excessive human interactions. The community should try to standardise the media used for data storage and retrieval where possible, and to invest in suitable devices to facilitate data loading/unloading like tape robots.

Data Acquisition

The cost of linking computers for data transport is an important issue to be considered. The annual running cost of leasing a 1Gb fibre from a provider like AARNET is ~$40K per Gb/s link. In a day, a Gb/s link can effectively transport about 5TB at a cost of ~$100. To move a petabyte dataset for analysis at a remote facility would therefore take over six months and cost in excess of $20K.

Need #4:
Australian astronomy would be greatly facilitated by the implementation of a national 10Gb/s (minimum) network between its supercomputers and national facilities.

Power

The enormous investment in computer hardware discussed above (~$300M) will come at a cost. A million dollar supercomputer can easily consume $100K in power per annum, and this is projected to rise if power costs increase faster than inflation in the future. Thus, in power alone, Australia may consume 10% of its hardware investment per annum (i.e. $25M!) just to keep the machines powered and cooled. It is inevitable that the facilities may look to users for contributions to their power bills. With current government policies this is likely to lead to issues for the long-term sustainability of such large machines.

Need #5:
Australian astronomers will need to have input into how National HPC centres are funded, to avoid a loss of their data or restricted access to these facilities. The current investments in hardware are not commensurate with the amount being invested in support staff, power and running costs.

4. Strategic vision for astronomy HPC and eScience in Australia

We summarise this document by describing a strategic vision for astronomical HPC and eScience.

- Australian astronomy is currently "hardware rich", with several new state-of-the-art national telescopes (ASKAP, MWA and SkyMapper) serviced by two multi-million dollar national HPC facilities (the NCI National Facility at ANU and the Pawsey High Performance Computing Centre for SKA Science in Western Australia). These resources will maintain Australian's position as a world leader in radio and optical astronomy throughout the next decade. If Australia is successful in its bid for the SKA the situation will only get better (by a few billion dollars).

- However, the data collectors (telescopes) and data processors (HPC) must effectively communicate if the raw data product is to be transformed into something useful for science. This is a non-trivial task, and one that is unprecedented in Australian astronomy on this scale. Much thought still needs to go into designing the layers of infrastructure (hardware and software) that will be required.

- A key gap in the astronomy eScience strategy is a clear way to fund and then implement the appropriate database middleware that sits between the HPC hardware and higher-level user VO layers. The middleware database is needed to appropriately tag (process) the data and structure it for fast and accurate access at a later time by the community. This is not an area that astronomers typically have expertise in, and specialist help may be required.

- We note that the government has already put in place a mechanism for developing such technology. The community should work more closely with ANDS and ARCS to find innovative solutions to its coming data challenges.

- The database middleware and VO user interface components should be designed to work seamlessly across both radio and optical data, forming what we call the National Integrated Astronomy Data Fabric. This will provide an opportunity to federate our most important datasets and open a new era of integrated multi-wavelength astronomy for the community in Australia. Data federation through this Fabric will enable astronomers to undertake scientific projects not previously possible (at least not with ease). This spirit of multi-wavelength cooperation has already begun under the recently funded ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO).