

## Executive Summary

In August 2019, the Australian astro-computing working group undertook a survey of computing usage by the Australian astronomical community. Combining the responses to this survey with public data and data provided by the NCI, Pawsey, and OzStar computing centres provides a snapshot of the current usage and needs of the community. Our survey shows that the Australian astro-computing community uses approximately 130 million CPU-hours per year of computer time in Australian facilities, and approximately 150 million CPU-hours per year on non-Australian machines. This time supports approximately 100 researchers, spread across most of the major astronomical research institutions in Australia. The largest Australian users also have significant non-Australian resources, and the primary use of non-Australian facilities is for projects that are too large to be carried out using any facilities available in Australia. Large computing users report that the lack of resources available in Australia compromises their ability to play significant roles in large international collaborations, and to carry out the scientific investigations they wish to complete.

## Data Sources

The primary source of data for this report is a survey carried out in August 2019. The full survey form, including all questions asked, may be found at <https://docs.google.com/forms/d/1uliMI31jomxUsySvqGKmfuii7l68qk8X9btxu9Cg818/>. The survey was directed specifically at research group leaders, so we did not request responses from each individual student or postdoc in a group; many of the individual responses stated specifically that they were responding on behalf of a large team. The survey received 28 responses in total, covering most of the major astronomical research institutions in Australia. Respondent affiliations are listed in the table below. Note that the survey specifically did not look at processing requirements associated with ASKAP or similar CSIRO-operated facilities, so the data collected exclude data processing usage. However, we do include data processing associated with OzGrav and similar gravity-wave work.

Institution	Number of respondents
Australian National University	4
CSIRO	1
ICRAR / University of Western Australia	4
Macquarie University	1
Monash University	5
Swinburne University of Technology	2
University of Melbourne	3
University of New South Wales, Canberra	1
University of Queensland	2
University of Southern Queensland	1
University of Sydney	2
University of Tasmania	1
Western Sydney University	1

In addition to the survey, we collected information on computing awards through the National Computational Merit Allocation Scheme (<https://ncmas.nci.org.au/2019/outcomes>), the ANU Merit Allocation Scheme (<https://anumas.nci.org.au/2019/outcomes>), partner share allocations by the Pawsey Supercomputing Centre (provided directly by Pawsey), and on OzStar (drawn from a report by Jarrod Hurley together with the ADACS Q4 Project Report for FY2019). We use these data to supplement the survey responses, but we are careful not to double-count allocations, i.e., if a survey respondent reports computing usage consistent with the time reported to have been awarded from the public data, we do not count that allocation twice.

### Time Usage within Australia

We first examine computer time used on facilities within Australia. The table below reports total usage on Australian facilities, broken out by facility and by institution. In each case the number we report is the number of CPU-hours used in a single year. Self-reported quantities have been cross-checked against public allocation information where possible, and time allocated via the NCMAS process by people who did not respond to the survey has been added. Divisions between facilities are somewhat approximate, since some survey respondents reported total usage and which facilities they used, but did not report exact amounts on each facility. The quantities reported are lower limits, since they do not include usage by researchers who did not respond to the survey, and who received time via processes that do not produce public reports. Also note that the sum of usage by institution is slightly smaller than the sum by facility because Pawsey only provided aggregate usage information for all of astronomy, not usage broken out by institution or PI. Thus we could only allocate usage on Pawsey to a particular institution if those users who were allocated the time self-identified through the survey. Similarly, we only had access to OzStar usage for Q4, so we assumed that the institutional allocation throughout the year matched that in Q4; thus some researchers will be undercounted and some overcounted on OzStar, depending on the division of their usage across the year.

Facility	Total Time Used (kCPU-hours)
NCI/Raijin	77600
Pawsey/Magnus and Pawsey/Zeus	20350
OzStar	24500
Other / In-house machines*	10600
<b>Total</b>	<b>133050</b>

\* Includes university-owned clusters and state facilities

Institution	Total Time Used (kCPU-hours)
Australian National University	64525
Monash University	17310
ICRAR / University of Western Australia	12310
Swinburne University of Technology	11014
University of Sydney	5050
University of Melbourne	2580

University of New South Wales, Canberra	2324
University of Southern Queensland	2300
University of Tasmania	2300
CSIRO	1524
Macquarie University	1397
University of Queensland	196
Curtin University	98
Western Sydney University	25
<b>Total</b>	<b>123152</b>

The reports show that the Australian astronomical community uses approximately 130 MCPU-hours per year on Australian facilities. Approximately 60% of this time is allocated on NCI, with OzStaz and Pawsey providing 15-20% each, and a collection of state- and institutionally-owned machines providing 5-10%. ANU is the single largest institutional user, followed by Monash, ICRAR/UWA, and Swinburne, and Monash. The results that a large fraction of the time goes to researchers at ANU, and that NCI provides a majority of the computing resources, are not unrelated: ANU researchers typically receive  $\sim 2/3$  of their time via ANU partner share and only  $\sim 1/3$  via open time allocated through NCMAS. Thus ANU's share of national open time is only slightly larger than that of ICRAR, Monash, or Swinburne. The large difference shown in the table above is due to ANU's partner share in NCI.

### Time Usage on Non-Australian Facilities

The survey also asked respondents to report their time usage per year on machines outside of Australia. We summarise the responses received, divided by non-Australian region and by Australian institutional affiliation, in the tables below.

Region	Total Time Used (kCPU-hours)
North America	10010
Europe	136015
Asia	1000
<b>Total</b>	<b>147025</b>

Institution	Total Time Used (kCPU-hours)
University of Queensland	60000
Australian National University	46000
ICRAR / University of Western Australia	31000
University of Melbourne	7000
Monash University	3010
Swinburne University of Technology	15
<b>Total</b>	<b>147025</b>

We see that usage on non-Australian facilities is approximately 150 MCPU-hours per year, or roughly 110% of total usage on Australian facilities. Thus, more than half of the computing done by Australians is not done in Australia. There are a few patterns in the non-Australian usage that are important. First, the great majority of the non-Australian usage is in large allocations; the single largest non-Australian allocation reported is 60 MCPU-hours on European machines (to Pat Scott's group at UQ), the second-largest is 30 MCPU-hours on European and UK machines (Aaron Ludlow's group at ICRAR/UWA), the next two largest are both > 20 MCPU-hours, and so forth. Taken together, individual allocations of >15 MCPU-hours constitute >95% of overseas usage. Allocations of this size are simply unavailable on Australian facilities; as noted above, the single largest astronomy allocation in Australia is 22 MCPU-hours, and 2/3 of this is ANU partner share. In 2019, the single largest allocation via NCMAS, the flagship computing program in Australia, in *any field of science and engineering* was 11 MCPU-hours.

Second, with the sole exception of Pat Scott's group, the groups with the largest Australian allocations also have the largest non-Australian allocations. Taken together, these two results indicate that non-Australian usage is coming from mostly from the largest Australian users, who typically have allocations on non-Australian machines that are as large or larger than their Australian resources.

### **How Many Researchers Does this Time Support?**

The survey asked each respondent to estimate how many people, in addition to themselves, made use of the awarded time either directly, via running computations using the allocated time, or indirectly, via usage and analysis of computational results (e.g., post-processing of simulations). Adding up the responses indicates that approximately 100 students and postdocs make use of supercomputer time directly or indirectly, in addition to the 28 direct respondents to the survey.

### **Does CPU Availability Limit Science?**

The survey asked respondents to comment on how / whether they would use additional CPU time, and what science it would enable. There is a diverse group of responses, but a common theme is that CPU time does represent a significant limiting factor for many groups' science, and that lack of CPU resources in Australia limits the participation of Australians in large international collaborations. Lack of Australian facilities also means that only Australians who have access to time overseas via non-Australian collaborators are able to engage in certain types of science. The following quotes taken from the survey responses illustrate the general tenor of the comments:

- "We are currently using of order 20 million CPU-h per year in Australia (NCI) and anticipate that this demand will at least increase by a factor of 3 in the next year. We are already using up our full allocation per quarter in one month, i.e., in 1/3 of the time."

- “I currently do not lack access to resources, but rather lack access to Australian resources. My science goals are therefore not being limited/restricted, but it is a sad state of affairs that I am generally forced to access CPU time for my work through international collaborations.”
- “Whilst nearly a third of the core members of GAMBIT are located in Australia, none of the major computing for our 13 papers to date has been done in Australia. This has limited the overall profile and impact of the Australian groups both within the Collaboration and on the international scientific stage.”