ICRAR
E info@icrar.org
T +61 (08) 6488 7766
www.icrar.org

ICRAR
International Centre for Radio Astronomy Research

TIANHE-2

CASE STUDY

ICRAR International Centre for Radio Astronomy Research
Curtin University
THE UNIVERSITY OF WESTERN AUSTRALIA
THE GOVERNMENT OF WESTERN AUSTRALIA

TIANHE-2

Working on the frontier of big data innovation attracts partnerships with the best researchers and the best facilities on the planet. ICRAR has leveraged strong ties to China to access Tianhe-2, the second fastest supercomputer in the world.

ICRAR was able to use Tianhe-2 to run a prototype part of the software system to manage data from the SKA telescope. This software is known as the SKA Science Data Processor, or the 'brain' of the telescope. It will process raw observations of the very first stars and galaxies to exist in the Universe more than 10 billion years ago, and prepare the data for analysis by astronomers from around the world.

ICRAR's access to this exceptional facility was possible because of the Centre's reputation for pioneering excellence in data science and strong international partnerships. It has worked hard to forge meaningful connections in the Asian region in particular, and has a close relationship with China.

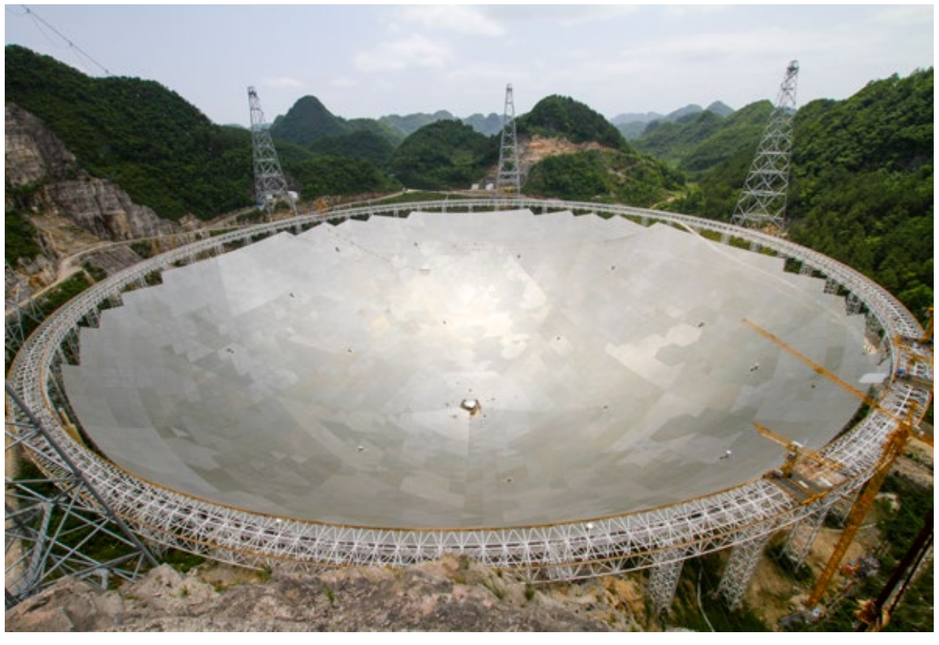
As well as Tianhe-2, ICRAR collaborates on the FAST telescope—a Chinese mega-science project to construct the largest filled-aperture radio telescope in the world. The Centre hosts Chinese students who come to study at ICRAR through the Chinese scholarship scheme, as well as postdoctoral fellows, and is part of the Australia-ChinA ConsortiuM for Astrophysical Research (ACAMAR).

"This world-leading facility is capable of performing quadrillions of calculations per second across 16,000 computer nodes."

The successful deployment of the prototype SKA software on Tianhe-2 in 2016 was the result of a collaboration between ICRAR and Shanghai Astronomical Observatory. This relationship with China is growing rapidly and will continue to be a major part of ICRAR in the future, with China and Australia working closely on a regional centre concept for the Asia-Pacific region.

ICRAR's director of data intensive astronomy Professor Andreas Wicenec said partnerships such as that with Tianhe-2 were incredibly important. "ICRAR's data intensive astronomy team is leading the international effort to overcome the immense challenges associated with the flow of data within the SKA," he said. "This world-leading facility is capable of performing quadrillions of calculations per second across 16,000 computer nodes. It's helping us learn to manage unprecedented amounts of information as we prepare for the big data challenges of the future."

Professor Wicenec said the prototype software run on Tianhe-2 was developed at ICRAR. It can be used to handle information from any field, pushing the boundaries of big data management. "It's really about getting the right algorithms in place—once that's done we can apply it to any type of data we choose," Professor Wicenec said.



Cover Chinese Supercomputer Tianhe-2. Image Credit: Prof. Yutong Lu.

Right Five-hundred-metre Aperture Spherical Telescope (FAST) in the southwestern province of Guizhou. Credit: Prof. Andreas Wicenec/ICRAR.

Left Dave Pallot, Associate Professor Chen Wu and Professor Andreas Wicenec from ICRAR's data intensive astronomy team.

The prototype software provides the control and monitoring environment to execute millions of tasks, consuming and producing millions of data items on many thousands of individual computers. This is the scale of processing required for six to 12 hours of observations with the SKA.

Professor Wicenec said the system is "data activated", meaning individual data items are wrapped in an active piece of software that automatically triggers the applications needed to process it. "Whenever a data item is ready, that's triggering the next task—the task is not running idle, waiting for anything," he said.

Excitingly, the framework can be used on any large data set, even outside of astronomy. "There's nothing specific to astronomy about it at all, that's the beauty of it," Professor Wicenec said. "It's a completely generic framework and we've already used it for other things. It's really about getting the right algorithms in place—once that's done we can put whatever we like in there."

The prototype was initially run on 500 compute nodes of Tianhe-2 and then extended to 1000 nodes.

The next step is to ramp up the number of individual items being deployed and then increase the number of compute nodes to what is expected for the SKA computer.

Professor Wicenec said the system is now running up to 12 million items. "Then we'll run between 50 and 60 million items on 5,000 nodes," he said.

"We're on track for SKA-class computing because of the system we've developed for Tianhe-2"

PROFESSOR ANDREAS WICENEC
ICRAR

Tianhe-2 director Professor Yutong Lu said the most important part of the project was the co-design and co-optimisation of SKA data processing software set and supercomputers such as Tianhe-2. "We're preparing for the faster computers in a few years from now," he said.

This work was carried out as part of the Science Data Processor work package for the SKA, which is led by the University of Cambridge.

ABOUT TIANHE-2

World's fastest supercomputer from 2013 to 2016, currently holds the title of the world's second fastest supercomputer, behind China's Sunway TaihuLight.

Housed at the National Super Computing Center in Guangzhou, China.

Contains 16,000 computer nodes, each with 88 gigabytes of memory.

Developed by a team of 1300 scientists and engineers.

Built by China's National University of Defense Technology (NUDT) in collaboration with the Chinese IT firm Inspur.

Used for simulation, analysis and government security applications, according to the NUDT.