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

ICRAR

International
Centre for
Radio
Astronomy
Research

THINK BOTTOM UP

CASE STUDY

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

THINK BOTTOM UP

Enabling local tech entrepreneurs is a core part of ICRAR's work. The Centre is helping WA start-up Think Bottom Up to prepare their unique and innovative database engine for the world stage.

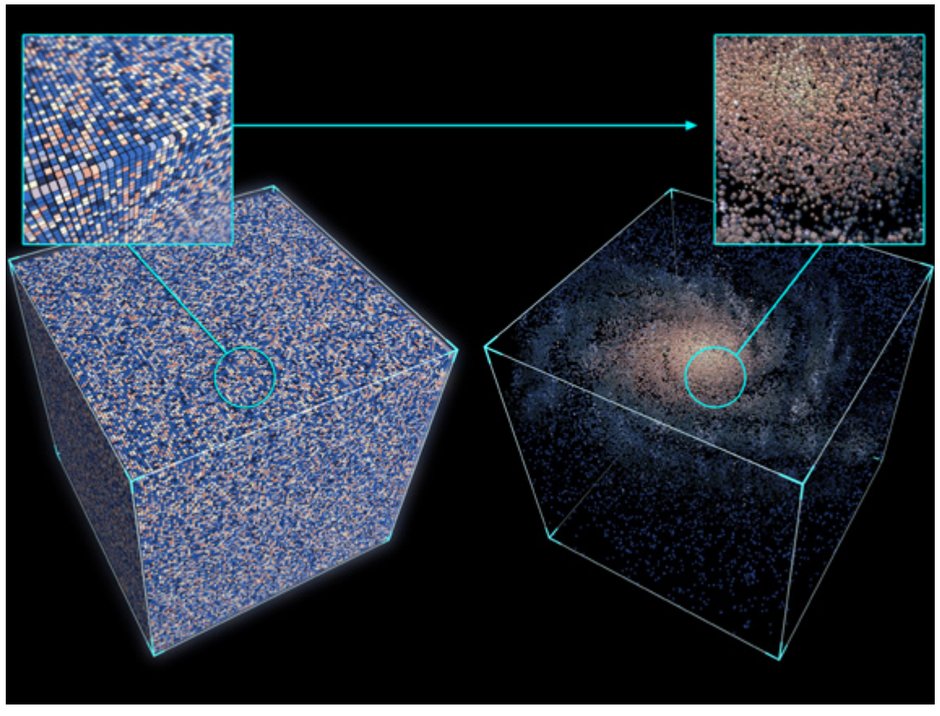
When a connection in the Netherlands pointed ICRAR at Perth software developers Think Bottom Up, the Centre's head of Data Intensive Astronomy Professor Andreas Wicenec knew there must be something worth following up. Professor Wicenec tracked down the small company and found they had a pioneering database engine that could be harnessed for big data work.

"The Perth tech company is helping ICRAR to build a computing platform for the future."

Now, the Perth tech company is helping ICRAR to build a computing platform for the future. Their database engine has been employed in the development of DALiuGE, a prototype execution framework capable of managing unprecedented amounts of data. It is being developed for the largest radio telescope in the world, the SKA, which is expected to produce around 250 petabytes of data each year during the initial phase—an amount of data equivalent to a stack of books eight million kilometres high.

Despite the company's state-of-the-art technology, Think Bottom Up consists of just three people. Their technically sophisticated database engine has been used in the background for several projects but is not yet a product in its own right. "Think Bottom Up is interested in taking this database engine and making it a full product," Professor Wicenec said. "At the end of the day they want to sell it but that requires a bit of work and also some new applications because what they've done until now is mostly in the mining sector. In order for a database to be successful, they need to have a bigger user base."

Astronomy is at the cutting-edge of data intensive science, with radio telescopes producing unparalleled volumes of information. And ICRAR's big data experts were instantly intrigued by what Think Bottom Up had developed. "The technology behind Think Bottom Up's database engine is very interesting," Professor Wicenec said. "We thought we'd need something like this, especially for the DALiuGE framework, so we talked to them and decided to do a small project with the company to test the waters." This proof-of-concept work was a small image reconstruction—similar to a Rubik's cube. The successful study gave ICRAR all the information it needed to engage Think Bottom Up to work on a much bigger project that could be integrated into DALiuGE.



Right A visual representation showing how Think Bottom Up's database engine is employed by ICRAR's DALiuGE execution framework to support the processing of large volumes of astronomical data.

"Our profile within Western Australia and the wider international community has been raised by being associated with this important and highly visible international project"

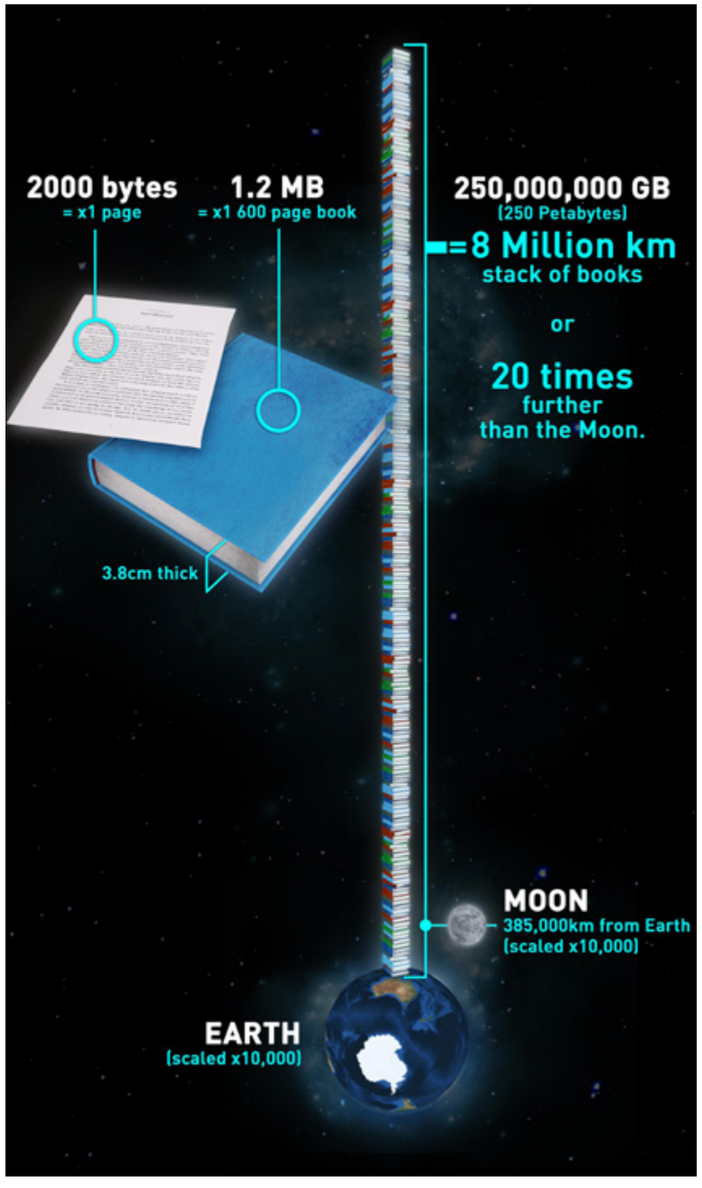
LACHLAN PARTINGTON
THINK BOTTOM UP DIRECTOR

DALiuGE's power is in its ability to allow researchers to run up to tens of millions of tasks in just a few hours using thousands of compute nodes with minimal overheads. Imagine a Rubik's cube with tens of millions of sub-cubes, rather than just 27. The goal is to spend as little time as possible managing the process and as much time as possible doing the actual work. Think Bottom Up's technology has been put to use in the framework doing what's called data life cycle management. It is a part of the system that monitors the flow of data 'drops' throughout their life cycle—from creation and initial storage to the time when they become obsolete and are deleted.

Think Bottom Up director Lachlan Partington said the Science Data Processor being designed for the SKA represents an extreme example of high performance and massively distributed computing. "It is the ideal test case for our technology," he said. "The partnership with ICRAR has also provided important marketing benefits to Think Bottom Up. Our profile within Western Australia and the wider international community has been raised by being associated with this important and highly visible international project."

Unlike typical databases such as Microsoft SQL Server or Oracle DB, which rely on a central engine, Think Bottom Up's product is a completely distributed database and is 'eventually consistent'. "If you want to have a fully distributed system, it can't be immediately consistent," Professor Wicenec said. "So if you change some information on one server, the same information will not immediately be changed on the other servers in the system, but only after a short amount of time. Our application, as well as many other applications, can live with such a situation, if it gets consistent eventually."

The story is just one example of ICRAR creating opportunities for local industry to work on a project with global visibility. With data processing for the SKA being designed by a consortium of prestigious institutions around the world, Think Bottom Up would likely never have got a seat at the table without ICRAR. Instead the start-up's signature database engine has found a place in an endeavor that will go down as one of the greatest data challenges of our time. "It's a really, really good product and the people in the company are technically absolutely outstanding," Professor Wicenec said.



Right During the project's initial phase, the SKA is expected to produce an amount of data equivalent to a stack of books eight million kilometres high!