ICRAR

International Centre for Radio Astronomy Research

Australia

Activities and Planning for SKA Regional Centres

Peter Quinn

Curtin University

THE UNIVERSITY OF WESTERN AUSTRALIA

Government of Western Australia
Department of the Premier and Cabinet
Office of Science

Australian activities

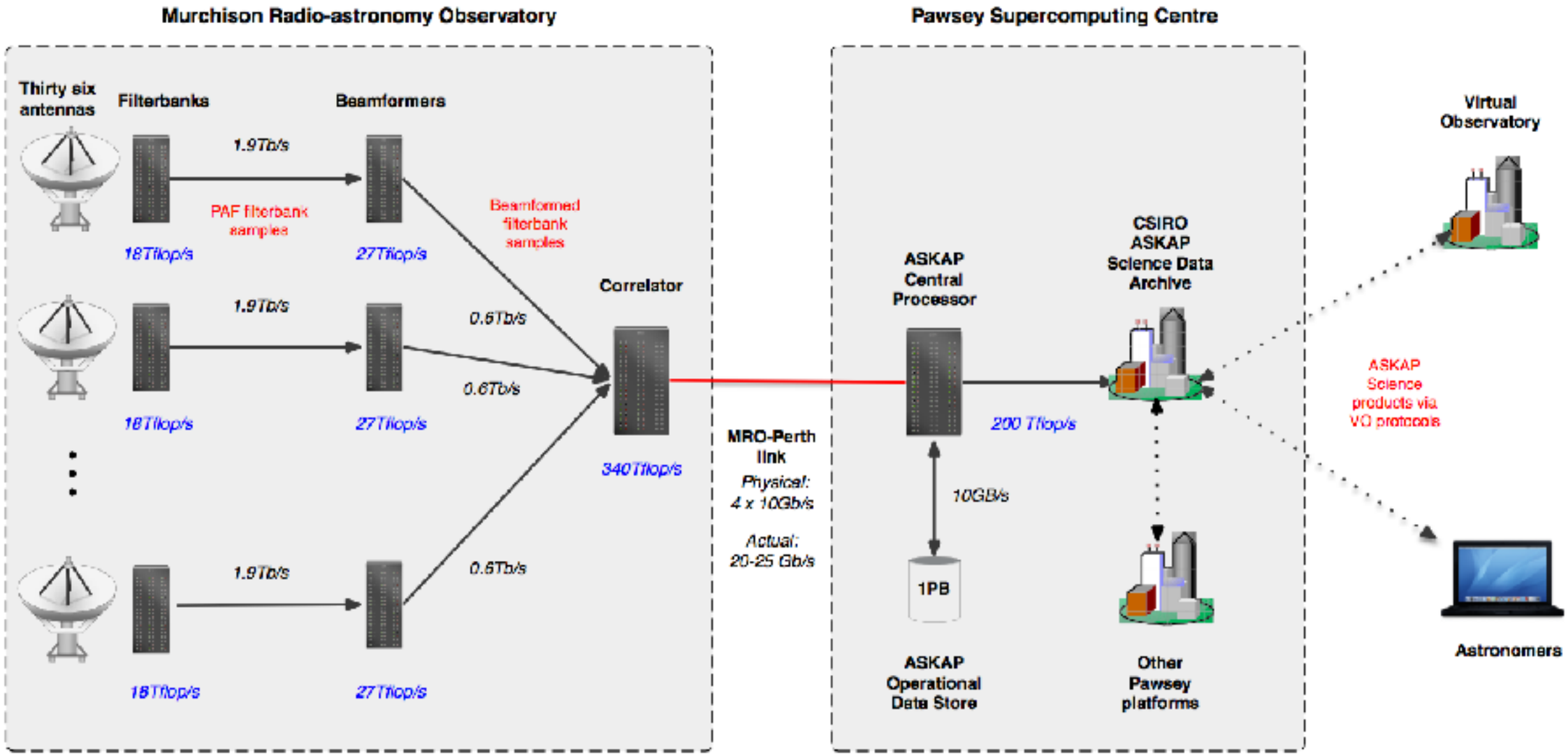
- AENEAS members
  - CSIRO
  - AARNET

- Australian SKA Regional Centre planning

- The ERIDANUS project & joint Australia-China activities

- Results from SDP and SRC related research and development at ICRAR

ASKAP data


Murchison Radio-astronomy Observatory
Pawsey Supercomputing Centre
Thirty six antennas
Filterbanks
Beamformers
1.9Tb/s
PAF filterbank samples
18Tflop/s
27Tflop/s
Beamformed filterbank samples
Correlator
0.6Tb/s
340Tflop/s
MRO-Perth link
Physical: 4 x 10Gb/s
Actual: 20-25 Gb/s
ASKAP Central Processor
200 Tflop/s
10GB/s
1PB
ASKAP Operational Data Store
CSIRO ASKAP Science Data Archive
Other Pawsey platforms
Virtual Observatory
ASKAP Science products via VO protocols
Astronomers

CSIRO ASKAP Science Data Archive

CASDA provides long-term archiving of and access to the large scientific datasets taken by ASKAP.

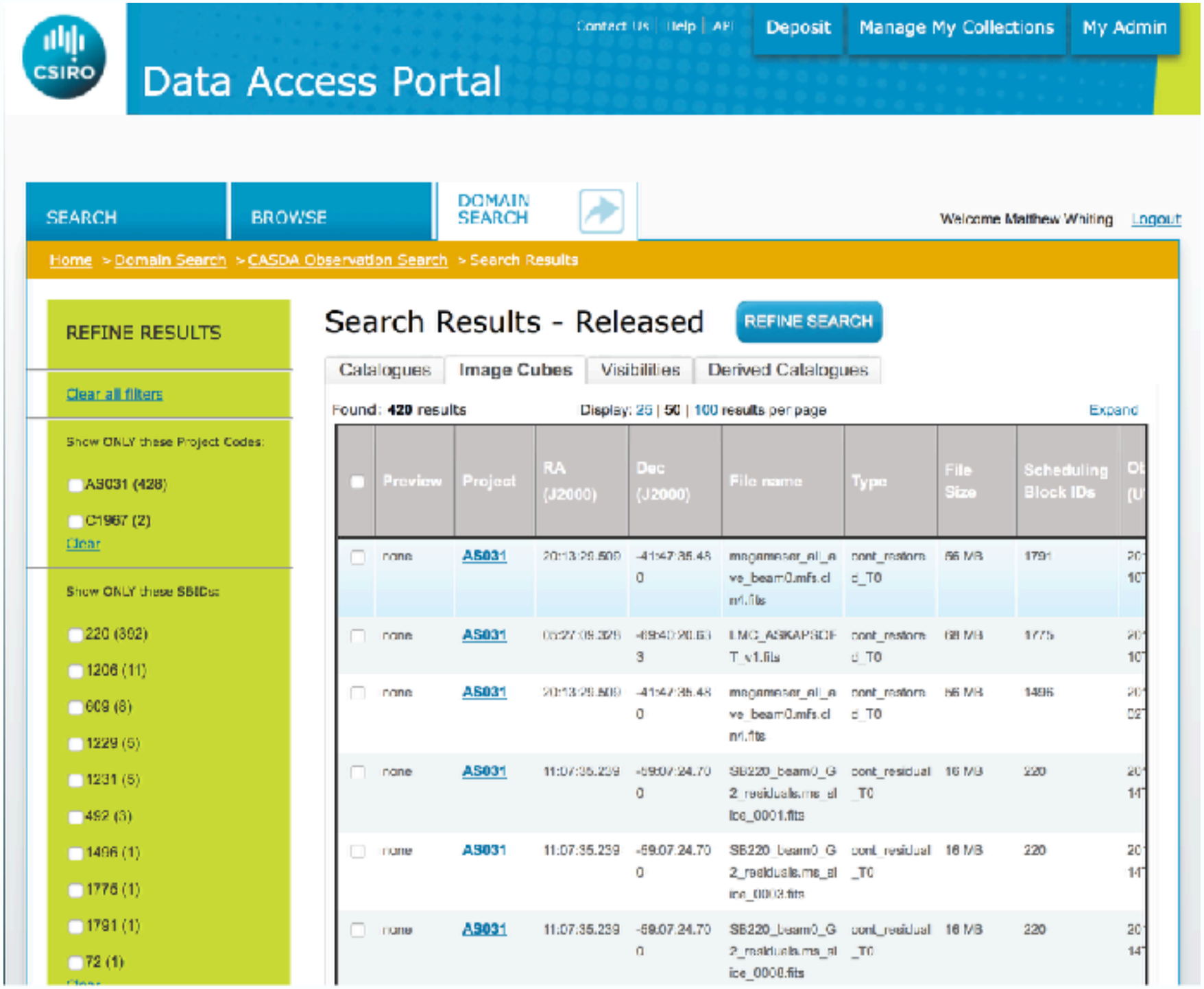
Functionality includes:

- Long term storage of ASKAP science data products
- Searches and data access via web (CSIRO Data Access Portal) and Virtual Observatory services
- Validation of ASKAP observations
- Upload of value-added science catalogues and image cubes
- Digital Object Identification (DOI) for all datasets
- Archive administration

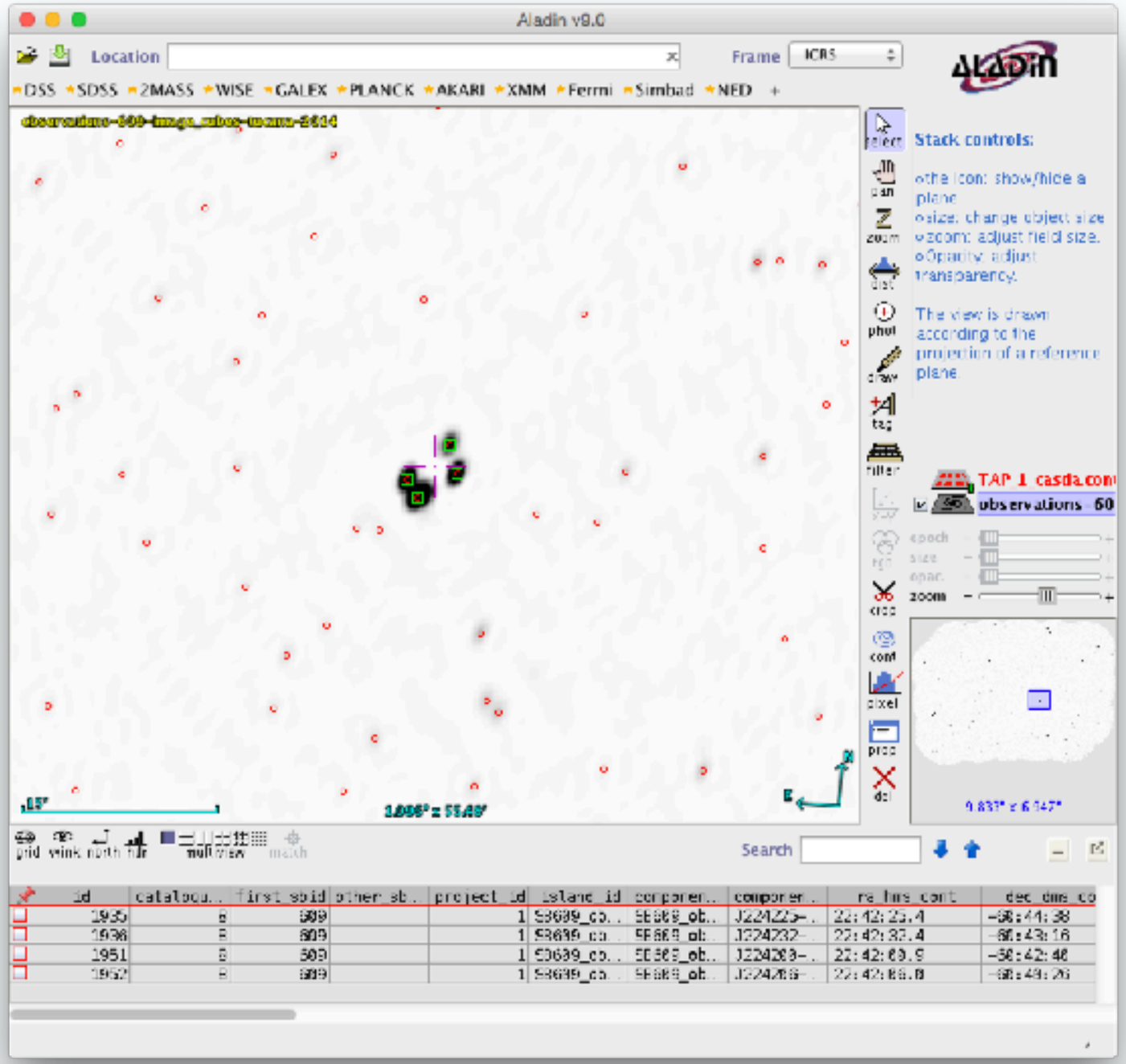
Approach:

- Agile Scrum software development

Data Access Portal interface



Virtual Observatory interface



Results: ASKAPsoft: First 36 beam image

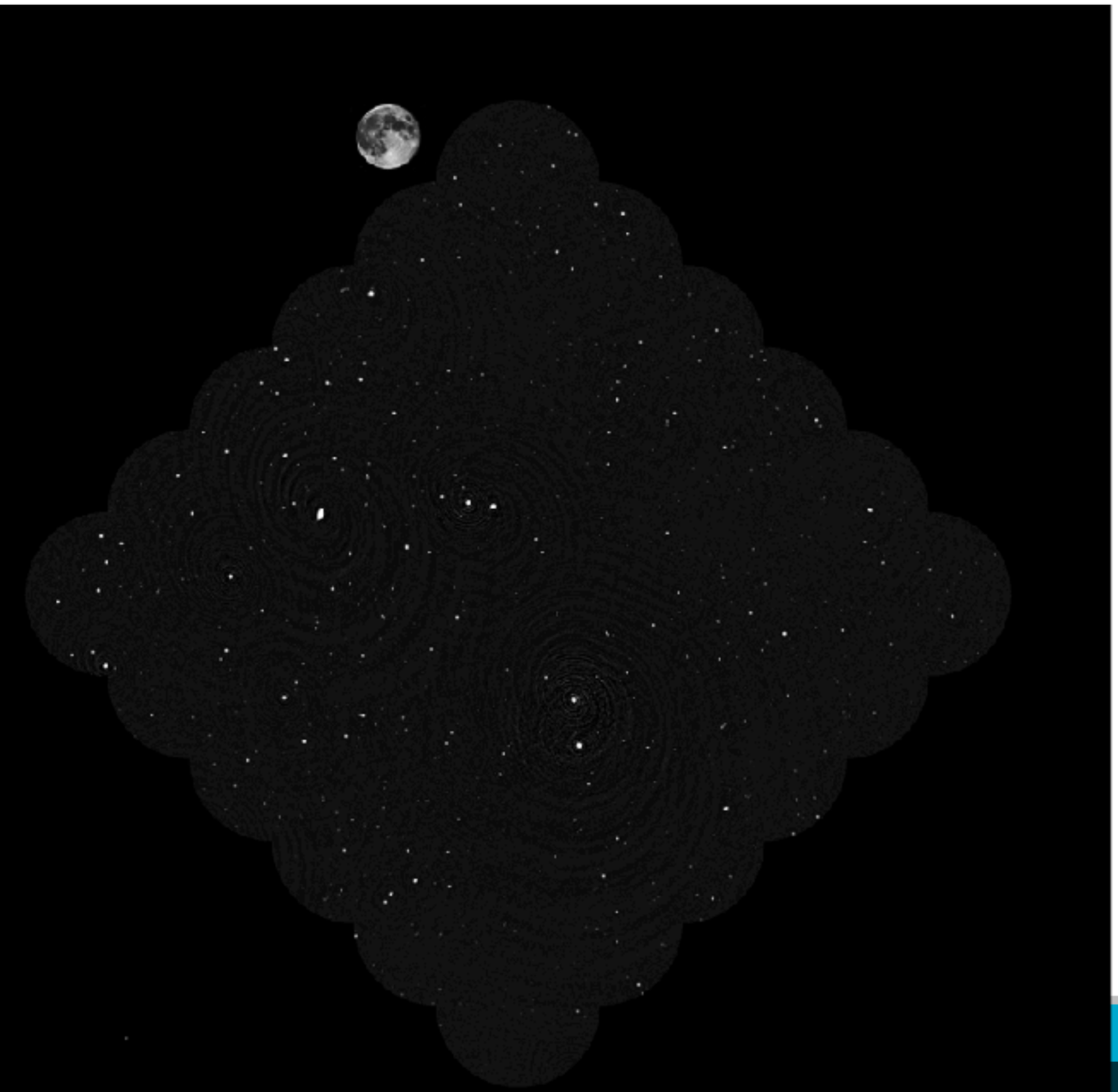
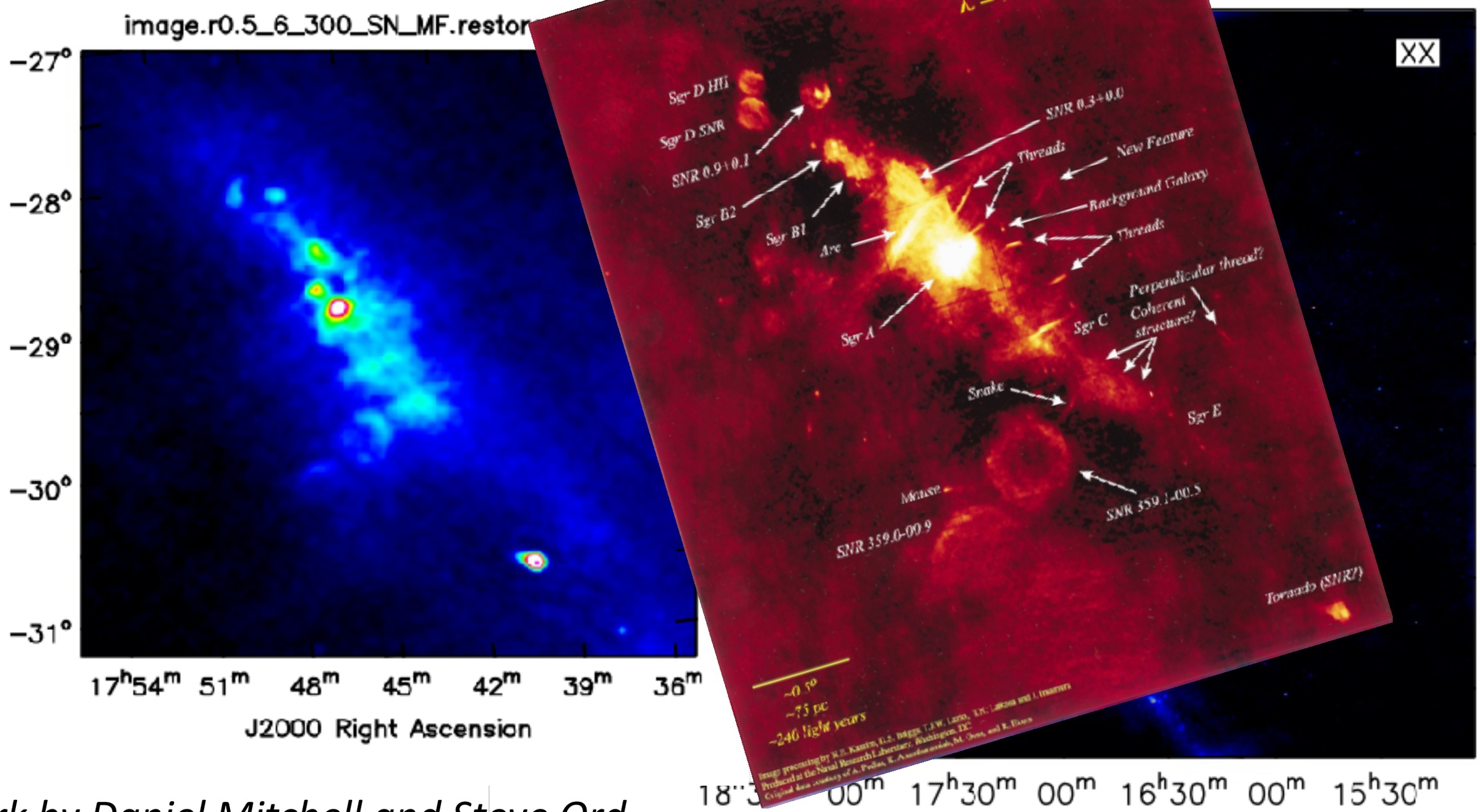


Image credit: Wasim Raja

- Continuum image with 9 antennas at 939.5 MHz
- Processing resembles an early-science experiment
- Each beam calibrated separately
- Individual deconvolution of different beams
- Only ASKAPsoft used

Proof-Of-Concept: Simple Pipeli r MWA



Work by Daniel Mitchell and Steve Ord

AENEAS – WP4

- Build on CSIRO's work since the mid-2000s with long-haul, high-bandwidth VLBI data transfers and e-VLBI:
  - Australian Long-Baseline Array,
  - EXPReS and NEXPReS EU funded projects (both concluded).
- Work closely with the Australian NREN, AARNet (Australian Academic and Research Network) to build capacity and conduct performance testing.
- CSIRO and AARNet have built and operated the long-haul network to connect the MRO to the world:
  - ~850km (about 390km fibre built as part of establishing the MRO),
  - running multiple 100Gbit/s wavelengths (MRO-Perth),
  - Provides and ideal real-world testbed for long-haul, long distance data transfers.
- Planning to undertake performance testing from April/May 2017 (depending on availability of hardware), initially within Australia and then Internationally.

Australian planning for SRCs

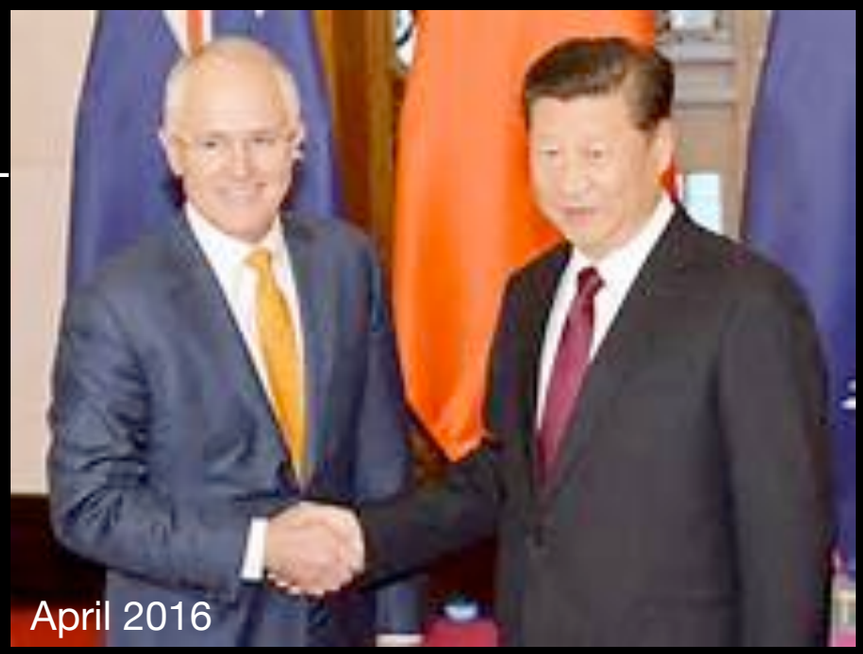
- Australian and New Zealand SKA Coordination Committee (ANZSCC) SKA Regional Centre Working Group (ASRCWG):
- Formed - first two meetings 9 August and 12 December 2016
- Members: Chair: Peter Quinn (**ICRAR**), **ASKAO:** David Luchetti, **CSIRO**: Douglas Bock/David McConnell, SKA Operations: Sarah Pearce, ICRAR: Andreas Wicenec, **CAASTRO**: Naomi McClure-Griffiths, **AAL**: Yeshe Fenner, **AARNET:** Peter Elford, Pawsey Centre: Neil Stringfellow, ANZSCC SAC: Carole Jackson, **ACAMAR**: Lister Staveley-Smith, **ASKAIC**: Geof Heydon, **New Zealand**: Willem van Straten, Melanie Hohnston-Hollitt
- Terms of Reference:

Australian planning for SRCs

1. To define the Australian SKA Regional Centre scope, opportunities, requirements and potential funding opportunities, in line with SKAO Operational planning and the SKAO Board's response to the recommendations of the SKA Data Flow Advisory Committee
2. To explore opportunities to expand the Australian SKA Regional Centre activities to include collaboration with similar activities in China, New Zealand and the broader Asia-Pacific Region with a particular focus on business case development and precursor enabled technological and scientific programs
3. To initiate a detailed study (ERIDANUS) of the data and processing requirements and costs within an Australian SKA Regional Centre in support of Australian Survey Science with the SKA precursors and SKA-1

China-Australia strengths

- **SCIENCE :**
  - ACAMAR (**A**ustralian-Chin**A** **C**onsortiu**M** for **A**strophysical **R**esearch) alliance - optical/IR , radio, theoretical and Antarctic astronomy
  - Future: FAST/ASKAP/MWA surveys, VLBI, Kunlun
- **TECHNOLOGY**:
  - CSIRO Multi-beam technologies for FAST
  - Low frequency technology - 21CMA, MWA+
  - Antennae design
- **DATA:**
  - Joint effort on SKA SDP systems
  - SDP prototypes - FAST, MWA and ASKAP
  - NGAS data flow systems for FAST
  - Pipelines on FAST
  - Developing regional VO capabilities
  - Capabilities and opportunities to develop regional solutions for SKA-class data flows and processing



April 2016

Data Canals and Rivers

FUNDED: 15 July

Advanced European Network of E-infrastructures for Astronomy with the SKA

ERIDANUS

Exascale Research Infrastructure for Data in Asian-Pacific astroNomy Using the SKA

Launched Nov 2016

eridanus.net.au

ERIDANUS - open invitation

- **The ERIDANUS Project is a three year, bottom-up, design study commencing in 2017, aimed at deploying prototype data intensive research infrastructure and middleware, between and within Australia and China, capable of addressing SKA-class data and processing challenges. The Project will respond to challenges identified by the SRCCG and will work in a coordinated and collaborative manner with the European Aeneas (Advanced European Network for E-infrastructures for Astronomy with the SKA) project.**

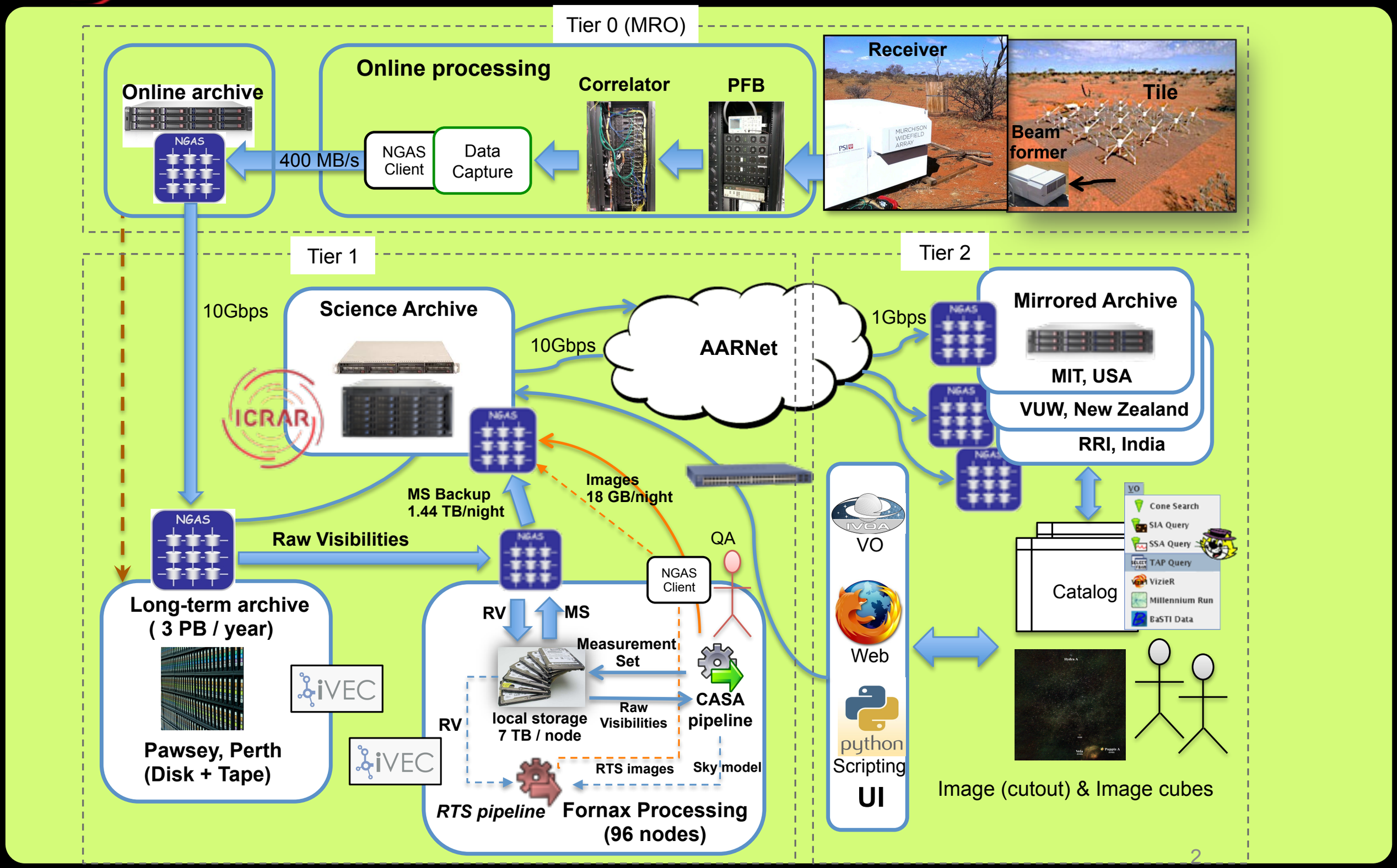
- **Founded by ICRAR and SHAO, now includes CSIRO and in discussions with others (e.g. CADC Canada)**

- **OPEN INVITATION to others to discuss joining**

- **Australia-China workshop on SKA Big Data Challenges - Shanghai (April 9-13, 2017)**

- **See eridanus.net.au for updates**

MWA data flow



Trials

Execute a radio astronomy survey science process in the Cloud with AWS

CHILES

CHILES, the COSMOS HI Large Extragalactic Survey

1002 hours JVLA B array

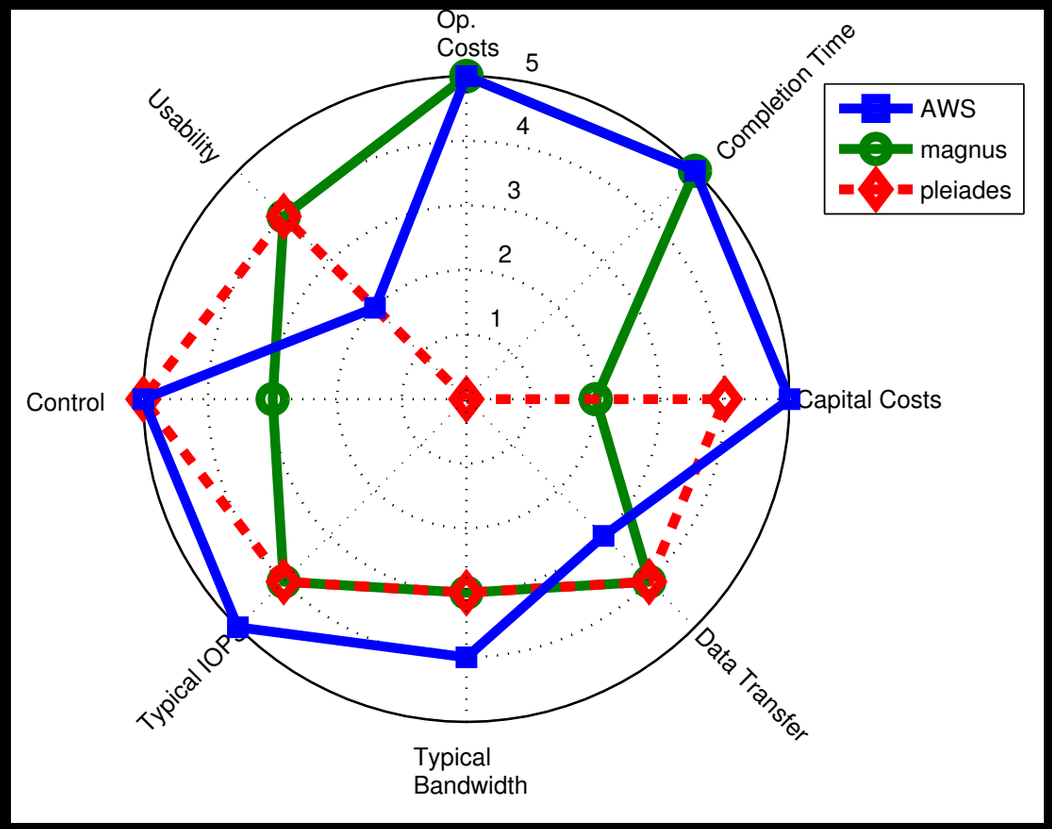
Jacqueline van Gorkom, Ximena Fernandez, Kelley Hess, D.J. Pisano, Kathryn Kreckel, Emmanuel Momjian, Attila Popping, Tom Oosterloo, Laura Chomiuk, Marc Verheijen, Patricia Henning, David Schiminovich, Matthew Bershady, Eric Wilcots, Nick Scoville

Computing efforts

Conventional Cluster (pleiades) 5 nodes each node has 2x Intel Xeon X5650 2.66GHz CPUs (6 cores / 12 HTs)	Enough computing power, however it would take weeks
Super computer (MAGNUS) Cray XC40 - 24 cores per node 2.6GHz Intel Xeon E5-2690V3 64GB per Node 35,712 cores available 3PB of storage #58 in the world	
AWS Whatever we wanted r3.xlarge 16 cores 122GB Ram	amazon web services™

Environment comparison

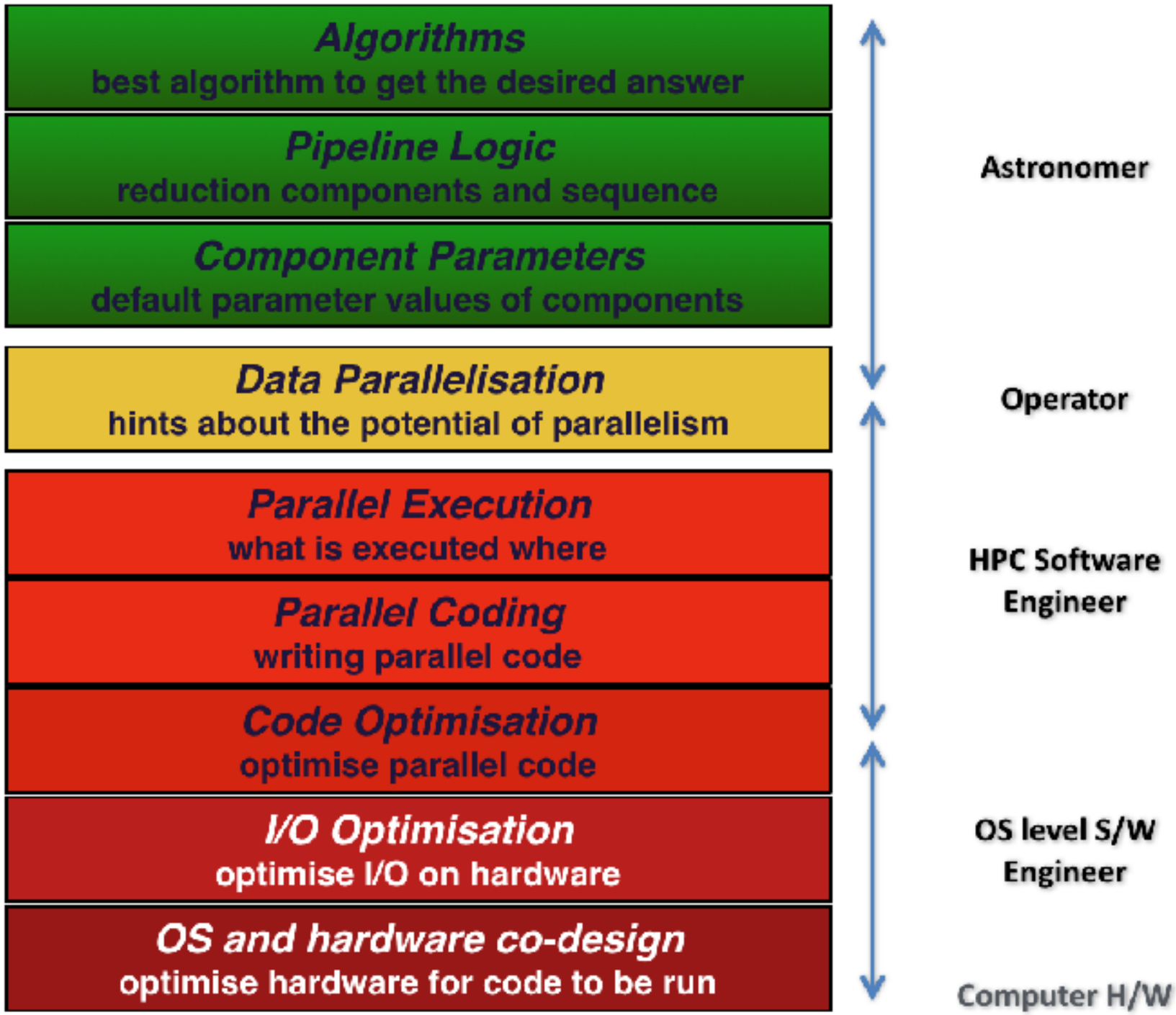
Consideration	*AWS*		*Magnus*		*Pleiades*	
Completion Time	96hr	5	110hr	5	1,060 hr (est.)	0
Capital Costs	$0	5	$340,000	2	$50,000	4
Operational Costs	$2,000	5	$3,240	5	-	0
Data Transfer	1Gb (high variance)	3	10Gb	4	10Gb	4
Typical Bandwidth	∼300MB/s	4	∼100MB/s	3	∼100MB/s	3
Typical IOPS	∼1,000	5	∼100	4	∼100	4
Control	Root Access	5	Limited Access	3	Root Access	5
Usability	Python/Boto	2	Python	4	Python	4
Product ($\Pi/5^8$)		0.15		0.07		0



DALiuGE – **D**ata-**A**ctivated **Liu** (流) **G**raph **E**ngine for Harnessing the Astronomical Data **Deluge**

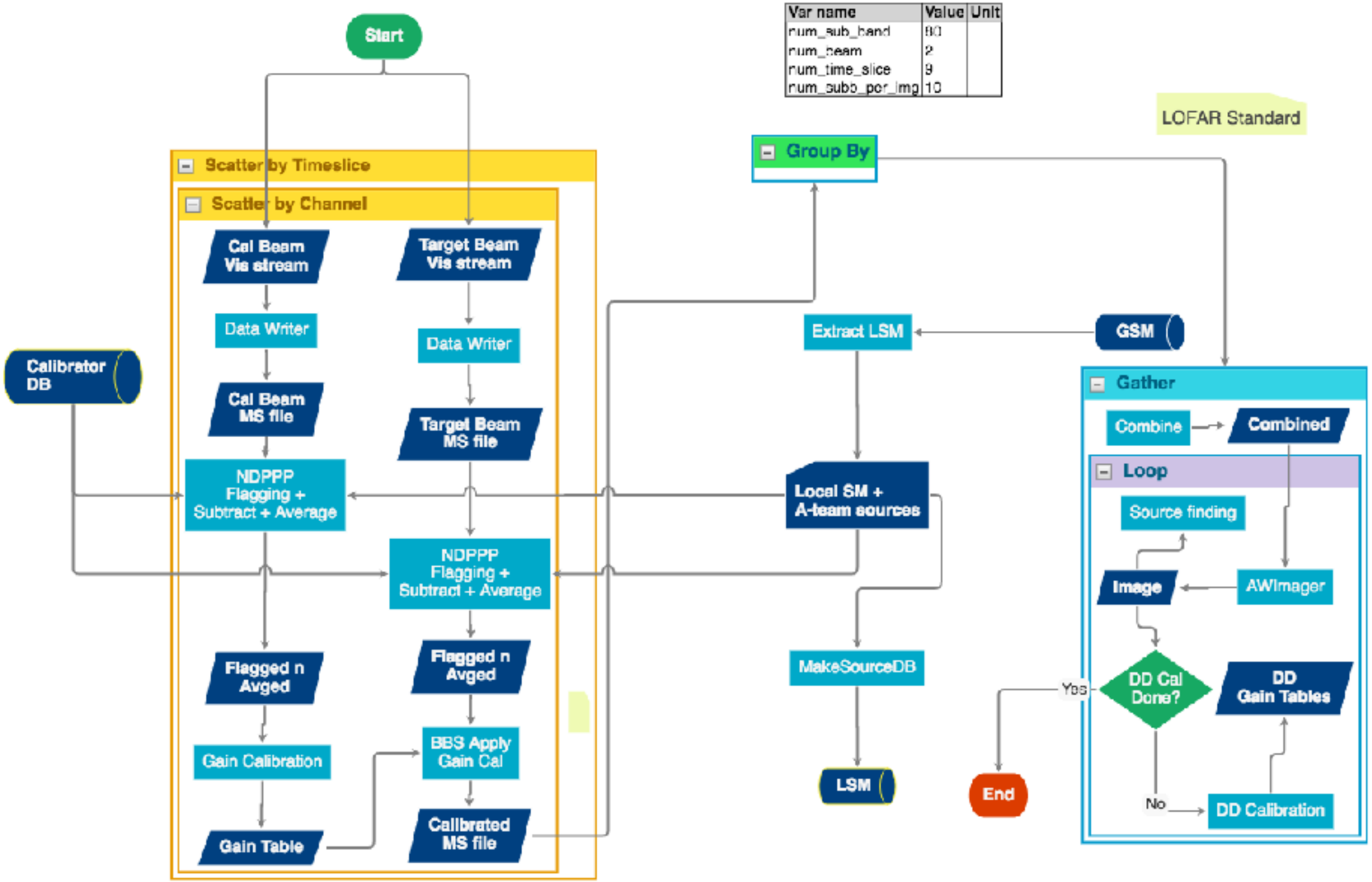
Chen Wu (ICRAR, UWA)

Why DALiuGE (II)

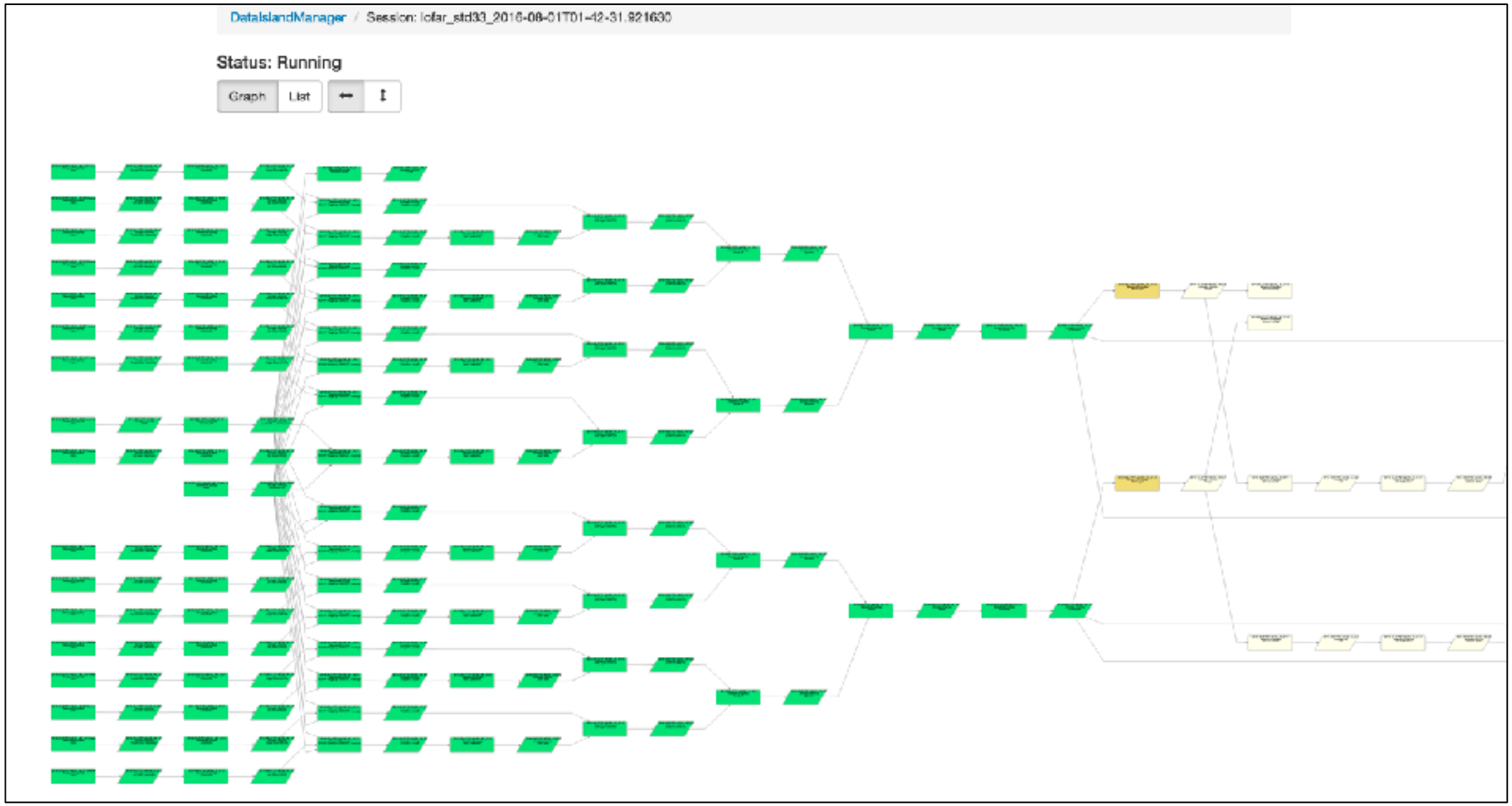


DALiuGE Case study – LOFAR

Var name	Value	Unit
num_sub_band	80	
num_beam	2	
num_time_slice	9	
num_subb_per_img	10	



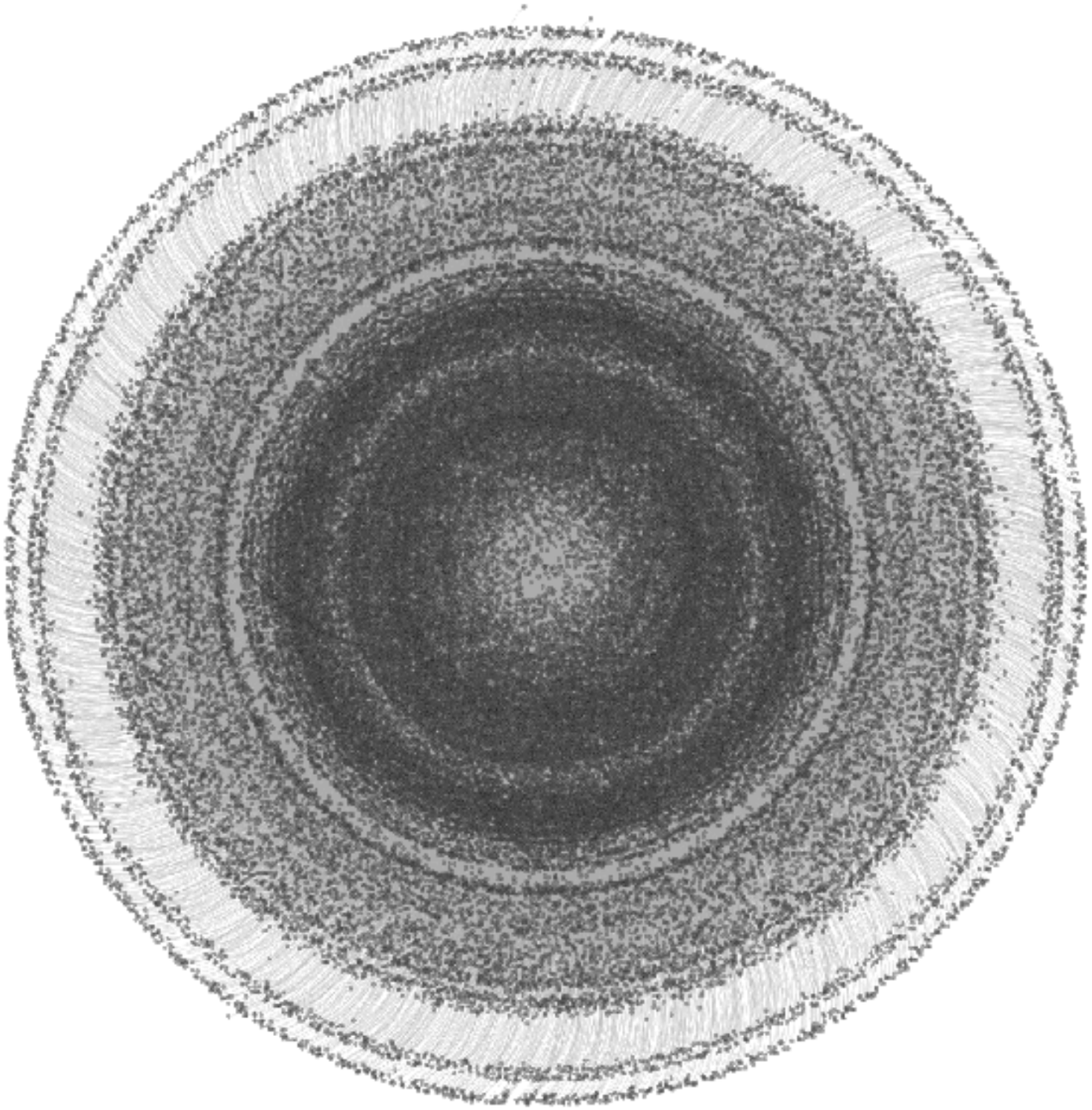
DALiuGE Case study – LOFAR



DALiuGE Case study – LOFAR

70K Drops running on 500 compute nodes at the Tianhe-2 Supercomputer for simulated LOFAR imaging “Dry” run

Gray – Drops not yet started
Yellow – Drops being executed
Green – Drops completed executions
Red – Drops failed



Future

- ERIDANUS deployment of processing and storage prototypes and scaling tests to SKA-1 scale using largest facilities available in Australia and China
  - Focused on precursor needs: ASKAP, MWA and FAST
  - Focused on large-scale surveys: CHILES+, GLASS, IMAGINE, as workflow and CLOUD testbeds
  - Focused on ASKAPsoft, DALiuGE and co-developed technology with industry: Nyriad (NZ) on new processing/storage HW (GPU optimisation)
- Australian SKA Regional Centre requirements definition (with SRCCG and KSPs) and Australian government funding application (2017?)