

**iVEC & Pawsey Centre  
Infrastructure and Resources**

Charles Schwartz

iVEC Technical Operations Director

7<sup>th</sup> May 2013



## Outline

- iVEC Mission
- Overview of the Pawsey Centre in Western Australia
- Pawsey system deployment schedule
- iVEC & Pawsey supercomputing infrastructure

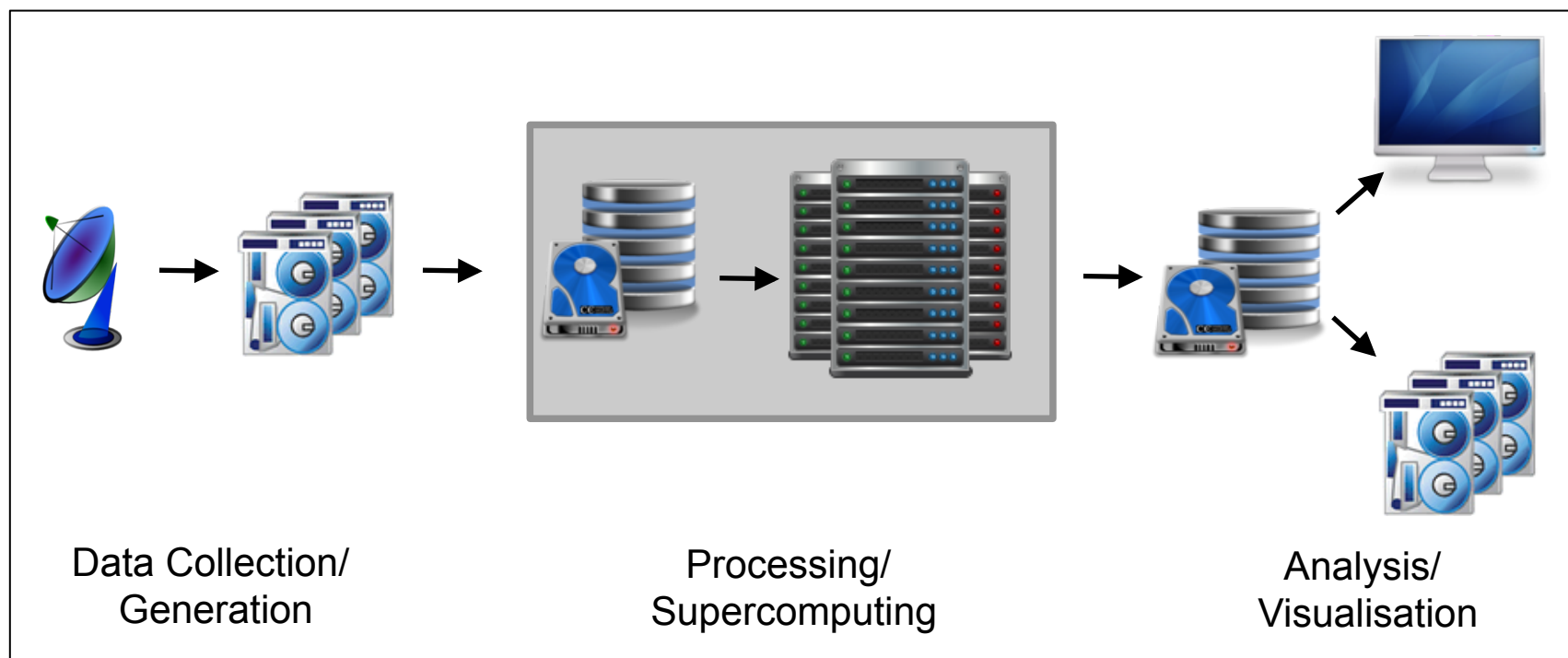
## iVEEC Overview

- Unincorporated joint venture
  - CSIRO
  - Curtin University
  - Edith Cowan University
  - Murdoch University
  - The University of Western Australia
- CSIRO as centre agent



## iVEC Mission: Building a Science Engine

- iVEC strives to build useful computing environments that accommodate the researcher's workflow(s)

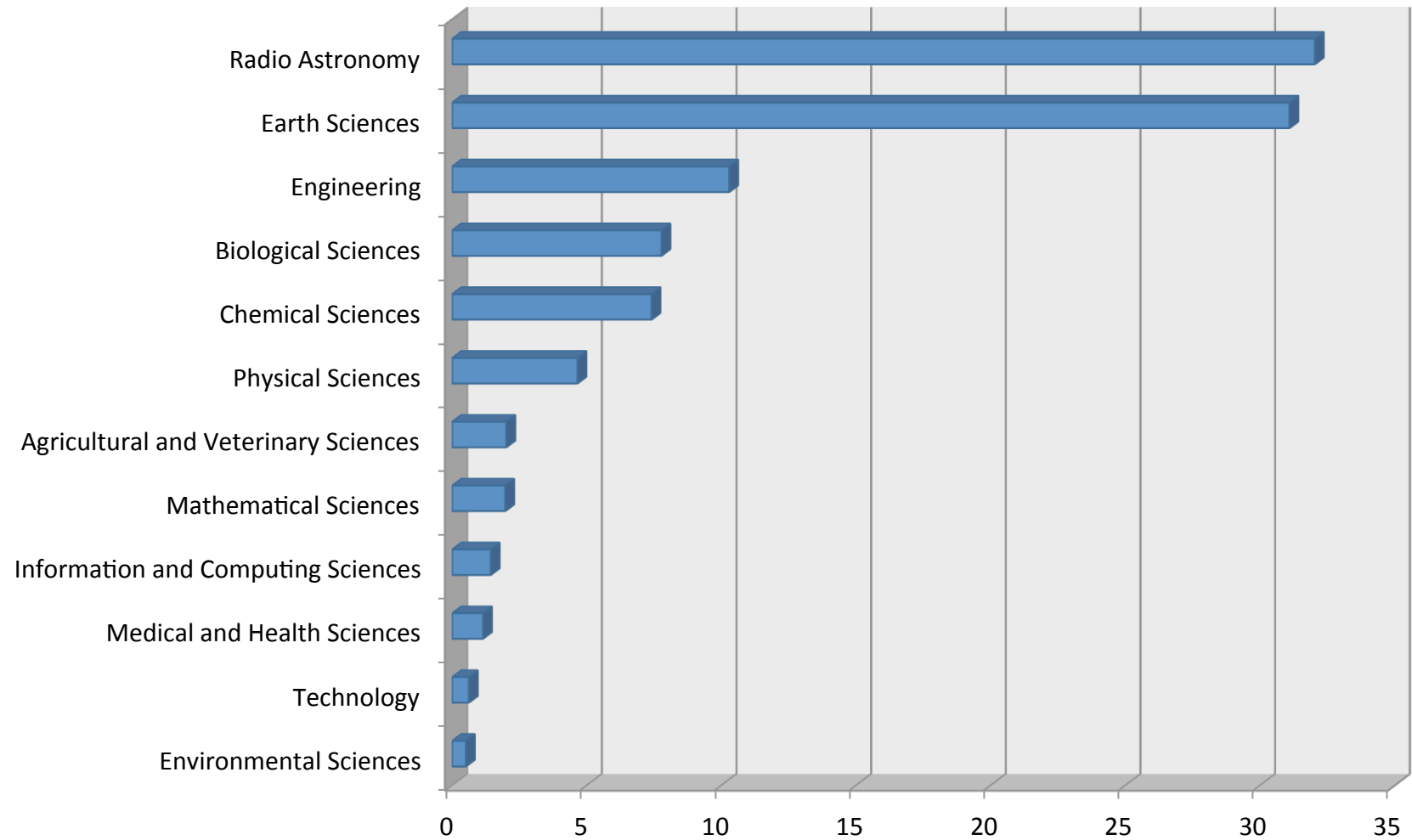


## Uptake Programs

- Alongside leading-edge facilities, we need to support researchers
- Five uptake programs help researchers to upscale their computing and their ambitions
  - Education
  - Supercomputing Technology and Applications (STAP)
  - Industry and Government Uptake (IGUP)
  - eResearch
  - Visualisation



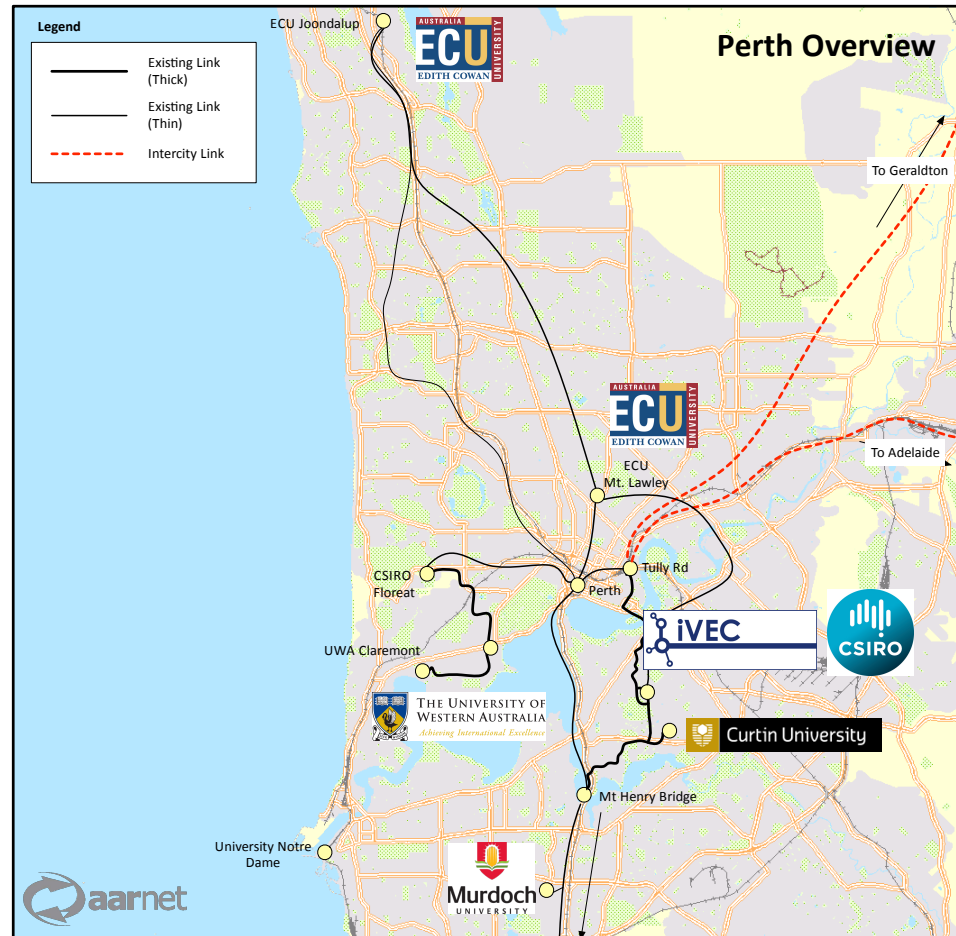
## iVEC Community



## iVEC Timeline

- Pawsey Project officially launched Aug 2009
- Stage 1 composed of two systems
  - Epic – operational mid 2011
  - Fornax – operational early 2012
- Stage 2 composed of systems in Pawsey Centre
  - Magnus – anticipated online by July 2013
  - RTC – anticipated in production by Nov 2013
  - Magnus 2 – anticipated online by June 2014

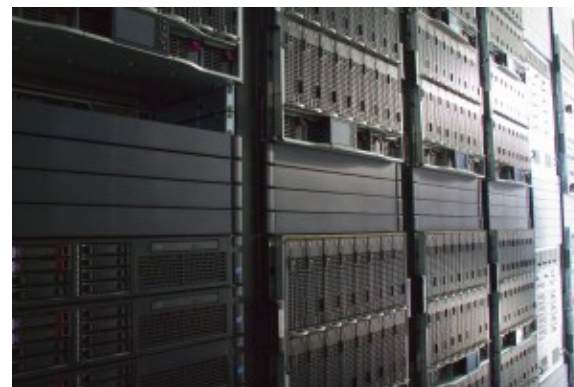
# iVEC Network





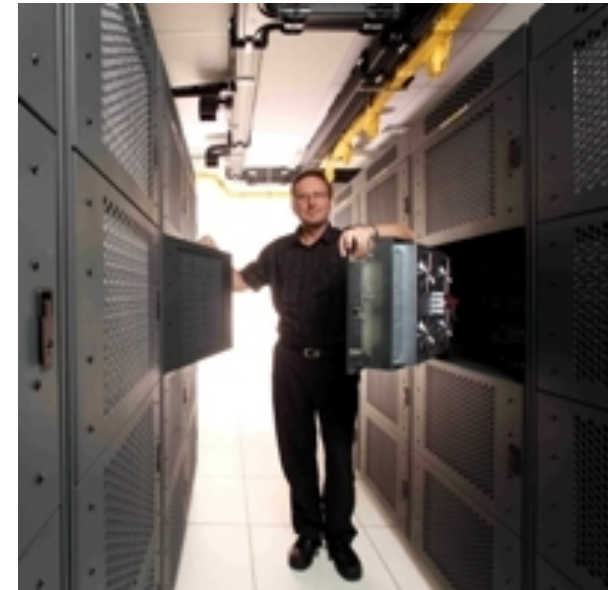
## Stage 1A – Epic @ Murdoch

- HP Performance Optimised Datacentre
  - 87 TFLOPs peak
  - Debuted at #88 in Top500
  - 9600 cores @ 2.8 GHz
  - 800 total nodes
  - 12 cores/node
  - 24 GB/node
  - QDR IB connectivity



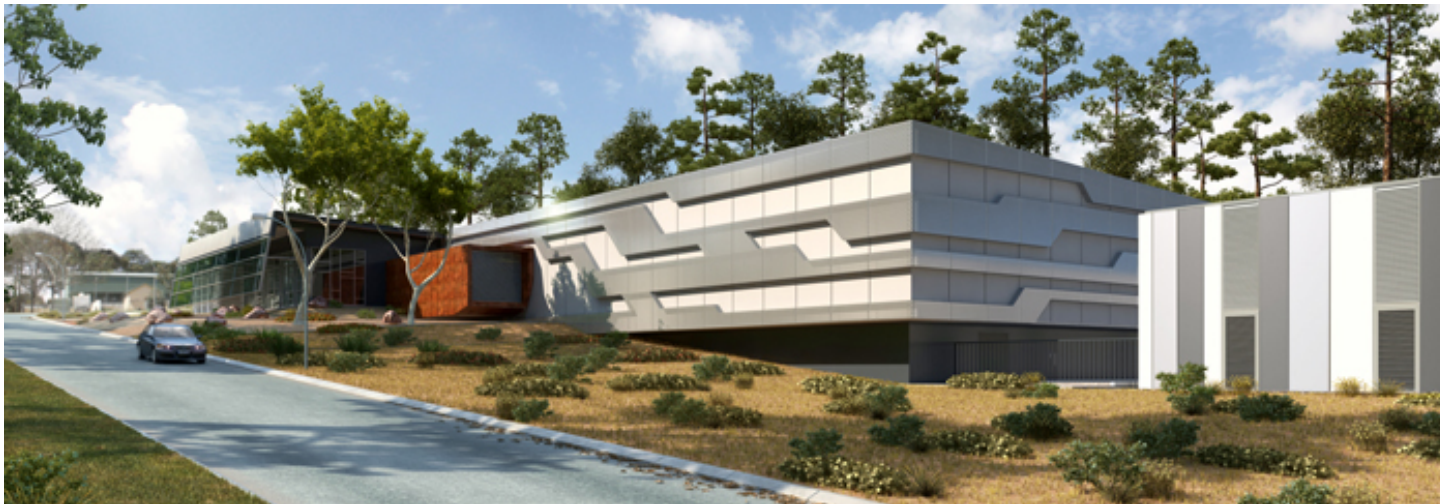
## Stage 1B – Fornax @ UWA

- SGI GPU cluster
  - 62 TFLOPs peak
  - 1152 cores @ 2.66 GHz
  - 96 nodes
    - 12 cores/node, 72 GB/node
  - 96 NVIDIA Tesla C2075 GPUs
- >500 TB Lustre file system
- 7 TB local storage per node
- Dual-rail IB interconnect for high throughput



## Pawsey Centre

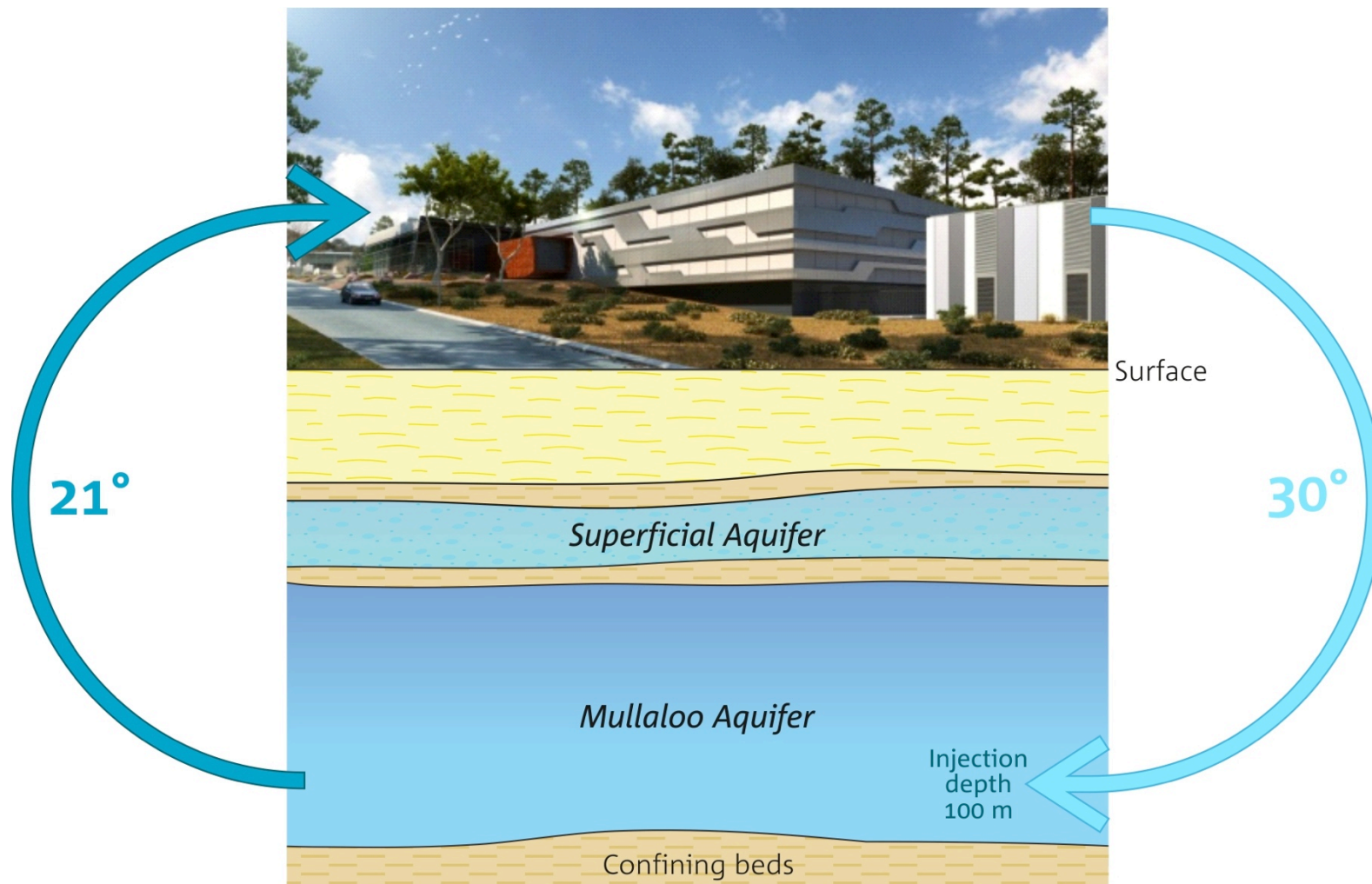
- Purpose-built supercomputing centre collocated at ARRC in WA
- Announced in May 2009 under Super Science Initiative
- Built by CSIRO, operated by iVEC



# Pawsey Centre

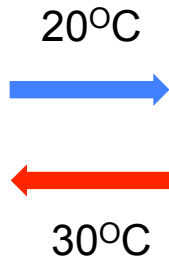
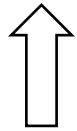


## Ground water cooling @ Pawsey Centre



# Ground water cooling @ Pawsey Centre

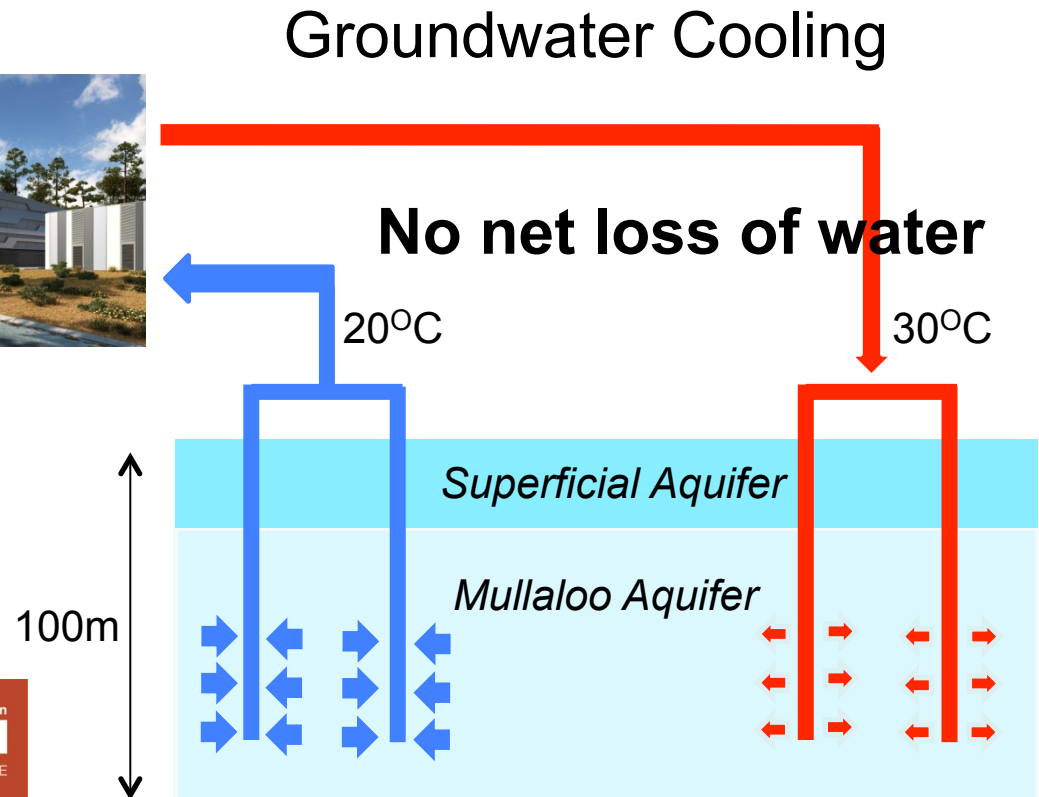
Substantial  
potable water  
consumption



Cooling Towers

## Groundwater Cooling

No net loss of water

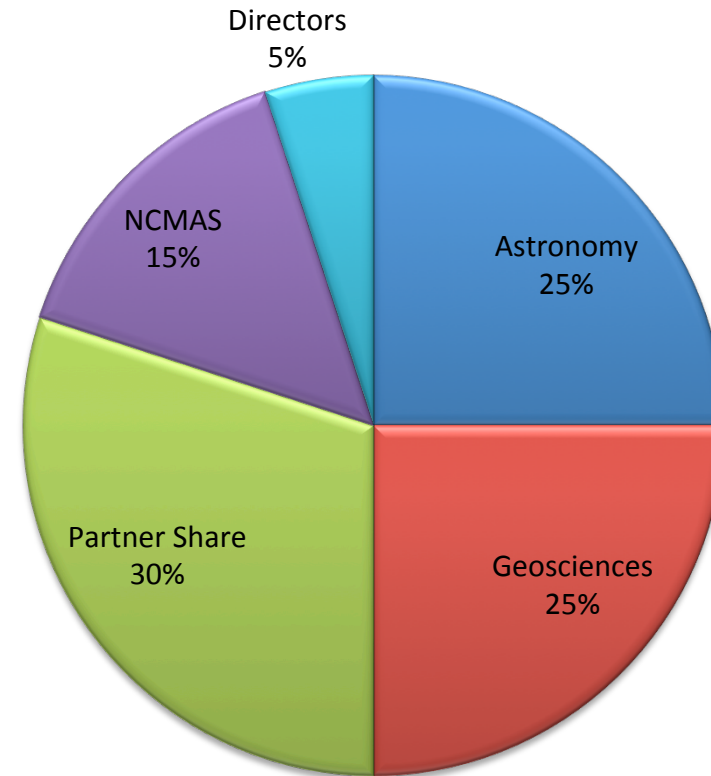


## Pawsey Centre - Green Initiatives

- Use of 22°C water for supercomputer cooling
- Pioneering research into groundwater cooling for data centre
- Use of photovoltaic cells on front of building
  - Approximately 5-10 kW of power generation
- Installation of photovoltaic cells on roof
  - Approximately 140 kW of power generation
- Total high watermark power draw: ~ 1.9 MW for phase 1 system

## Pawsey Centre

- System allocations
  - 25% Radio-astronomy
  - 25% Geosciences
  - 30% iVEC Partner Share
    - 15% National Merit Share
      - aka NCMAS
  - 5% iVEC Director





## Pawsey Centre – Radio-astronomy

- Two primary radio-astronomy projects
  - **Australian Square Kilometre Array Pathfinder (ASKAP) Project**
    - High-frequency
    - 36 antennae in Murchison Radio Observatory (MRO)
      - ~800 km north of Perth
    - 4× 10 GigE links into Pawsey from MRO
    - 5 PB/year
  - **Murchison Widefield Array (MWA)**
    - Low-frequency
    - 1× 10 GigE link into Pawsey from MRO
    - 3 PB/year



## Pawsey Centre – Radio-astronomy – ASKAP

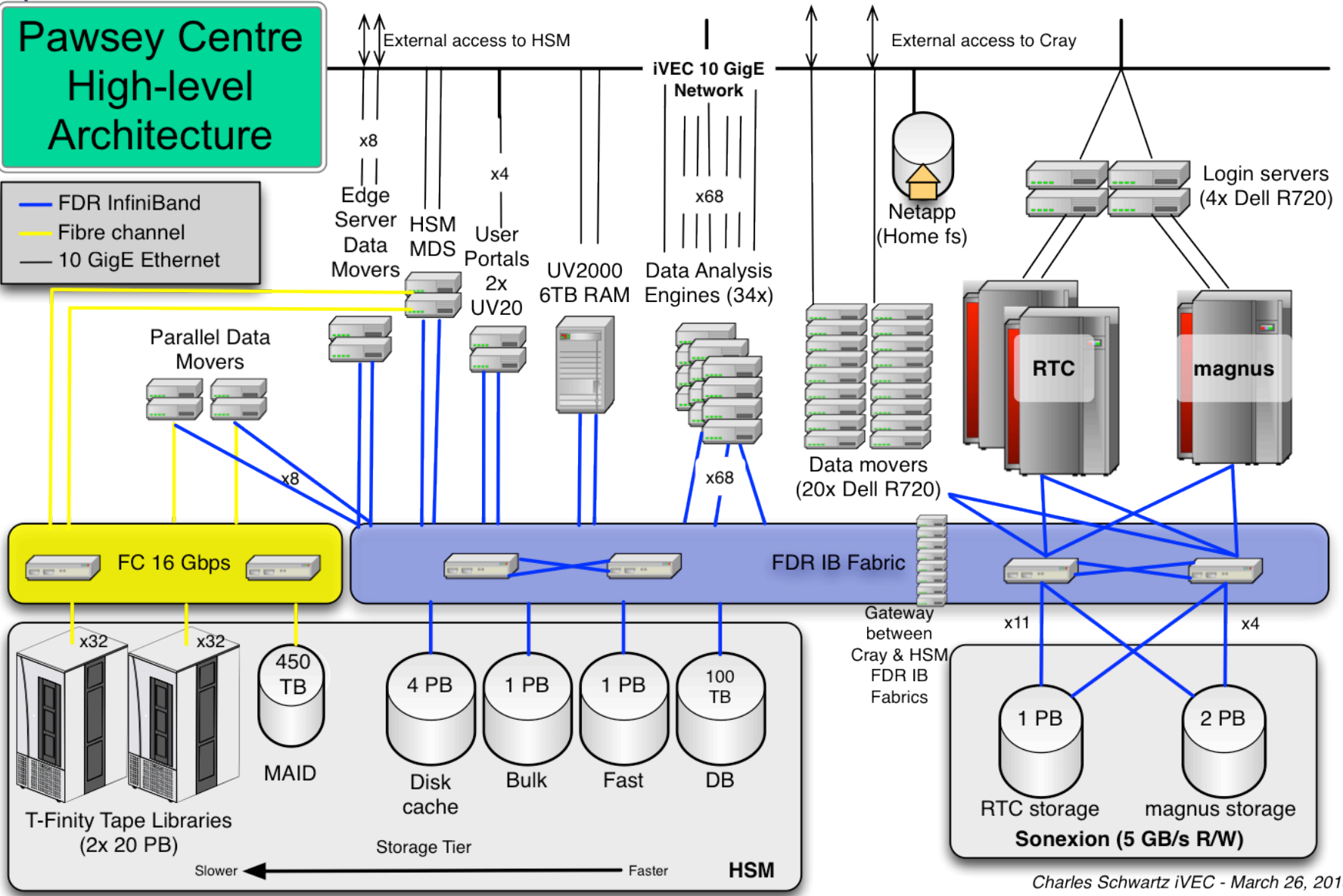
- Real-time data processing in radio-astronomy has not been done before
- Raw observation data is too large to archive or transfer...
  - ~ 220 TB/day of measurement data
  - Only around 10% is transferred to Pawsey
- ...and even the data products are large
  - Spectral line image cube ~ 1 TB
  - 12-hour observation full polarisation continuum visibility data set – 2+ TB

## Pawsey Centre – Radio-astronomy – MWA

- Data correlated at MRO in local facility
  - Powered by solar and diesel fuel – no grid infrastructure!
- Approximately 2 TB/hour in observation data
- Data sharing via NGAS

# Pawsey Centre High-level Architecture

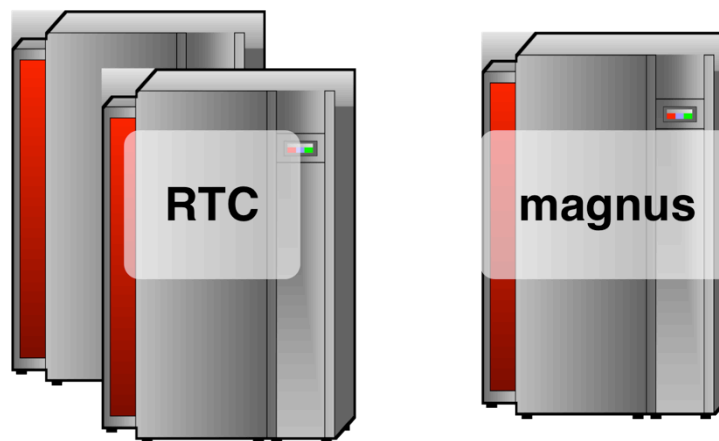
- FDR InfiniBand
- Fibre channel
- 10 GigE Ethernet



## Supercomputers



- 2 Cray XC30 supercomputers



- Magnus
  - 69 TFLOP system for general science
- RTC – real-time computer
  - 200 TFLOP system for ASKAP and MWA production science

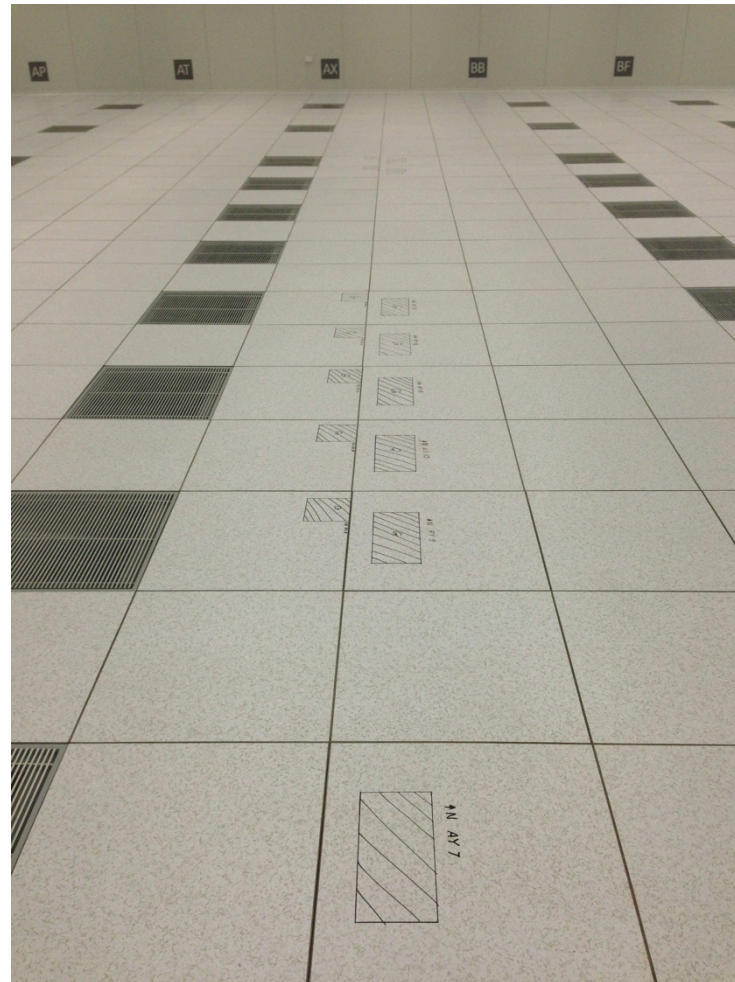
## Magnus

- 69 TFLOP Sandy Bridge system
- 3,328 cores on 208 compute nodes
- 64 GB/node
- 2 PB Sonexion 1600 storage
  - 1 MMU, 9 SSUs
- Delivery starting May 20, 2013
- Operations anticipated ~ July 2013

## RTC (Real Time Computer)

- 200 TFLOP Ivy Bridge system
- 9,440 cores on 472 compute nodes
- 64 GB/node
- 1 PB Sonexion 1600 storage
  - 1 MMU, 6 SSUs

# Supercomputers





## Data Analysis Engines & Visualisation

- 6 TB shared-memory UV2000 for visualisation
  - 4× NVIDIA Quadro
- 34 SGI Rackable 2U servers
  - Virtualisation and visualisation
  - 128 GB RAM
  - NVIDIA Quadros & Keplers
  - 2× 10 GigE, 2× FDR IB



## Data Analysis Engines & Visualisation

- Powered up already
- Commissioning in progress
  
- Also great for drying footwear (until the water cooling is connected)



## High Performance Storage

- 3 PB of Lustre Sonexion storage
  - 2 PB for magnus
  - 1 PB for RTC
- > 6 PB of disk in HSM – IS5600 arrays
  - 5 PB disk cache
    - 1 PB of 10k SAS
    - > 4 PB of 7200 rpm SAS
  - 800 GB SSD storage (*metadata*)
  - DMF



## Archives (or Not High Performance Storage)

- SpectraLogic T-Finity library
  - 40 PB raw capacity, 20 PB usable capacity
  - 64 TS1140 tape drives
  - 4 TB tapes today
- Capacity for up to 100 PB raw
  - 50 PB usable
- 450 TB MAID



## High Performance Network

- *Nearly* end-to-end FDR IB fabric internally
  - Up to 56 Gbps
  - Mellanox hardware
- 10 GigE Ethernet connection to iVEC network
- Fibre channel connectivity to and from tape library and tape drives
  - 2 Gbps

## Software

- Resource management and batch system will be PBS Pro
- Libraries, compilers, system software...
  - Intel & Cray compilers
  - Allinea DDT debugger
  - Cray Linux Environment (CLE) 5

## Magnus – Phase 2

- Phase 2 of Pawsey Centre compute infrastructure expected mid-2014
- Increase total compute capacity to  $\geq 1$  PFLOP
  - More XC30 cabinets, more Lustre storage
  - Power draw expected to go up to  $\sim 2.25$  MW
- Accelerator & CPU balance
  - Xeon Phi, NVIDIA Kepler or Atlas?
  - ~~Intel Haswell CPU?~~

## Magnus – Phase 2

- $\geq 1$  PFLOP Ivy Bridge system
- $\geq 14,560$  cores on 728 compute nodes
- Accelerators – to be determined!
- 64 GB/node
- 6 PB Sonexion 1600 storage total
  - 1 MMU, 9 SSUs, 18 EXPs
- Delivery starting April 2014
- Operations to begin  $\leq$  June 2014



**Thank you!**

- For further information...
  - [www.ivec.org](http://www.ivec.org)