Current and Future directions for HPC at Pawsey Supercomputing Centre

Ugo Varetto and Mark Gray

The Pawsey Supercomputing Centre is an unincorporated joint venture between

CSIRO, Curtin University, ECU Edith Cowan University, Murdoch University, The University of Western Australia

and proudly funded by

NCRIS National Research Infrastructure for Australia, An Australian Government Initiative

Government of Western Australia

PAWSEY supercomputing centre

Pawsey Supercomputing Centre

PAWSEY
supercomputing centre

Supporting Australian Researchers

1673 researchers

179 projects

46 staff

MAKING TOMORROW HAPPEN, TODAY

through

- supercomputing
- data
- visualisation
- training & consulting

National Reach

PAWSEY supercomputing centre

35,712 cores, 1.09 PFLOPS,
Aries dragonfly interconnect

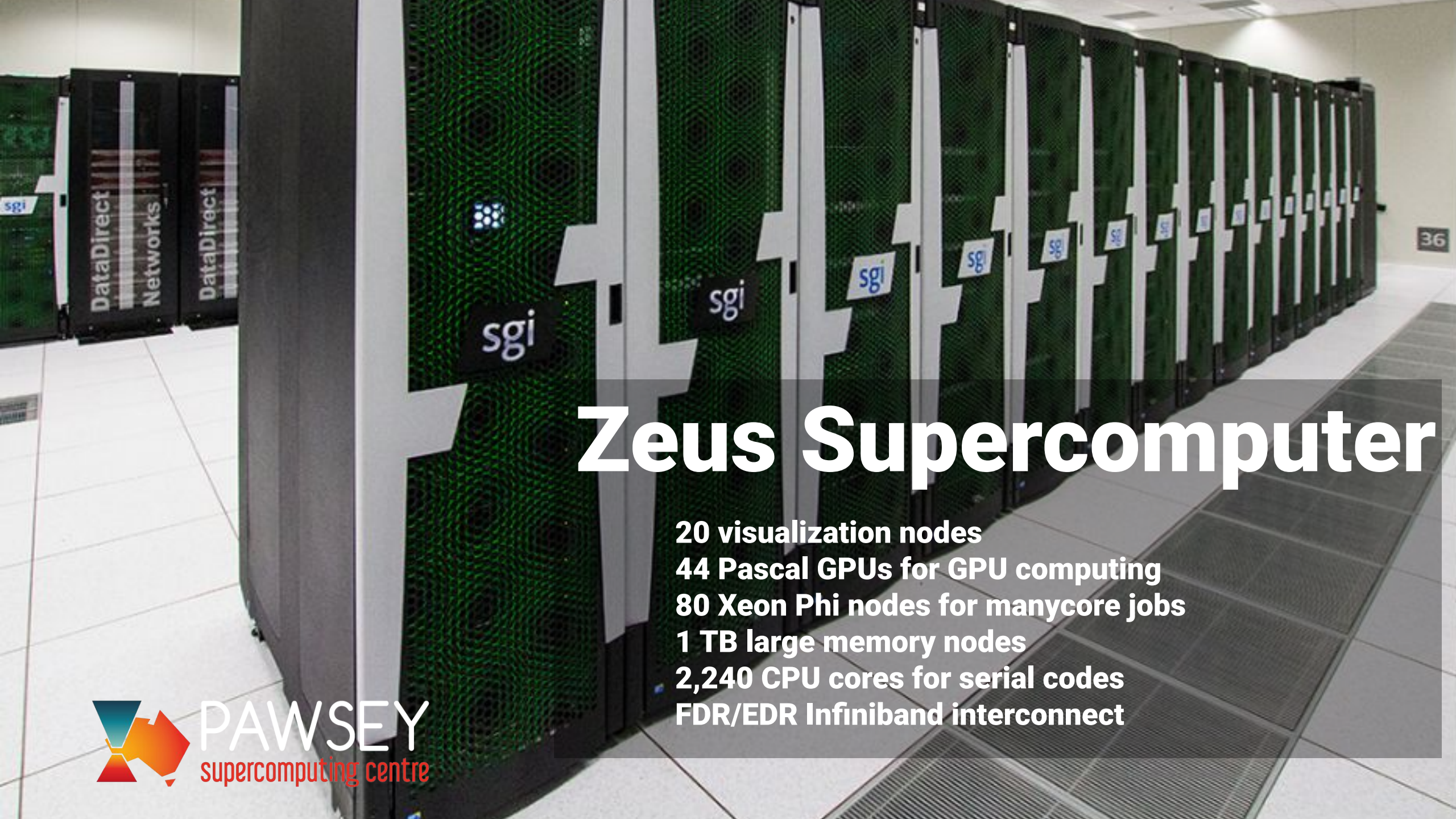
Magnus Supercomputer

PAWSEY
supercomputing centre

9,440 CPU cores
64 K20X GPUs
Aries dragonfly interconnect

Galaxy Supercomputer

PAWSEY
supercomputing centre



Zeus Supercomputer

20 visualization nodes
44 Pascal GPUs for GPU computing
80 Xeon Phi nodes for manycore jobs
1 TB large memory nodes
2,240 CPU cores for serial codes
FDR/EDR Infiniband interconnect

PAWSEY
supercomputing centre

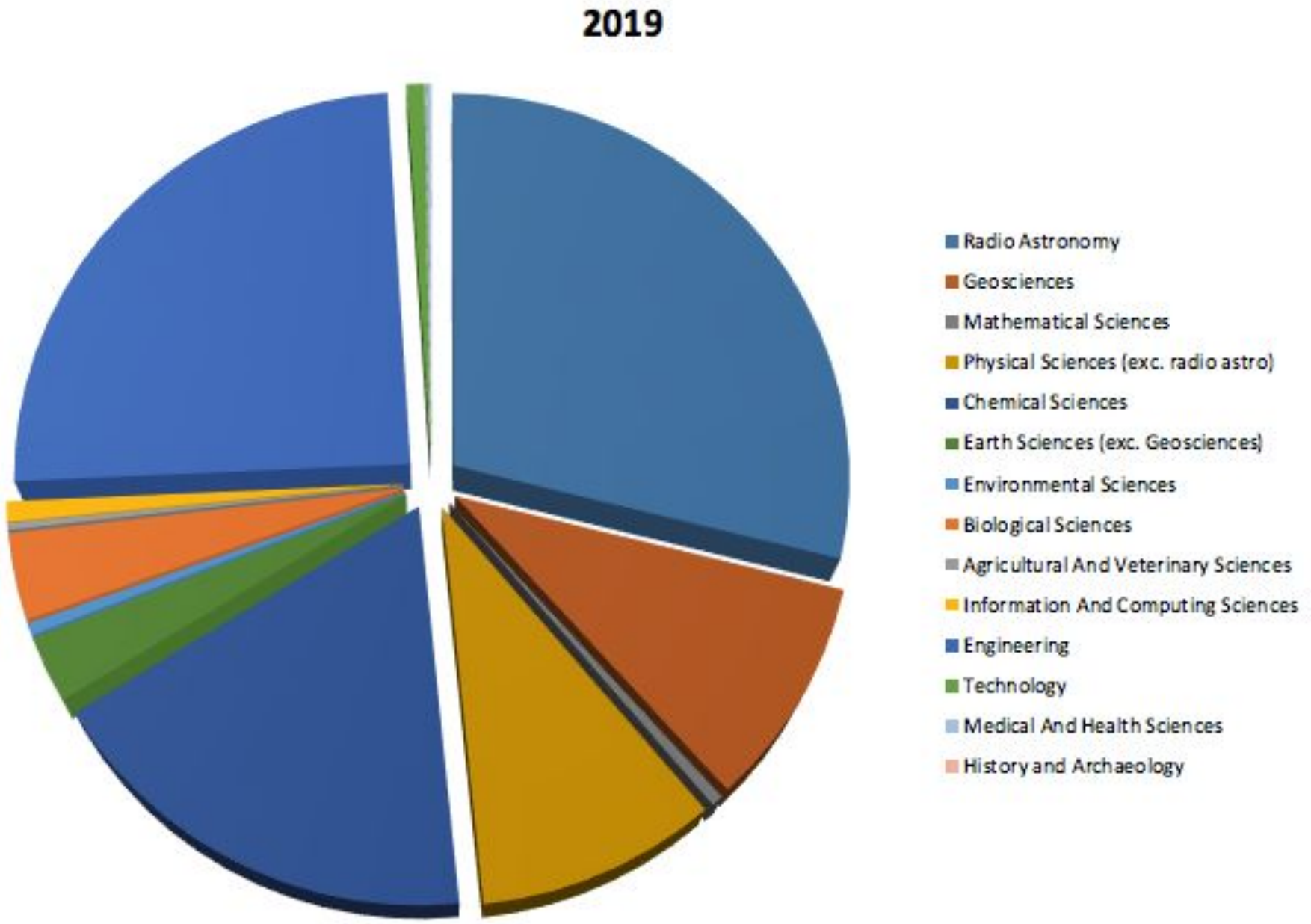

3000 Cores, OpenStack,
Sahara, Volta GPUs
Nimbus Research Cloud
PAWSEY
supercomputing centre

65 PB Migrating Disk and Tape

Data Storage

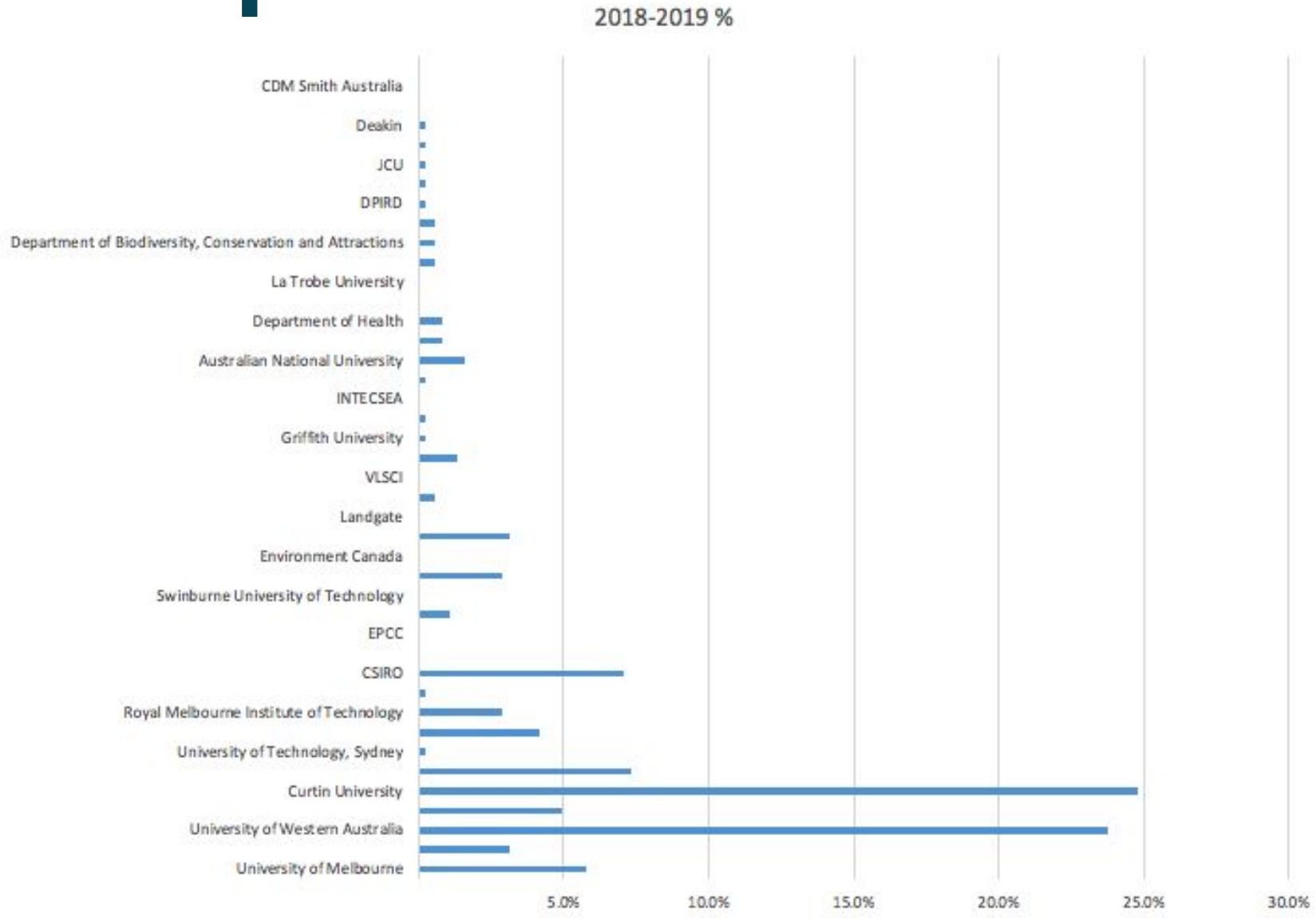
PAWSEY
supercomputing centre

Usage by science domain



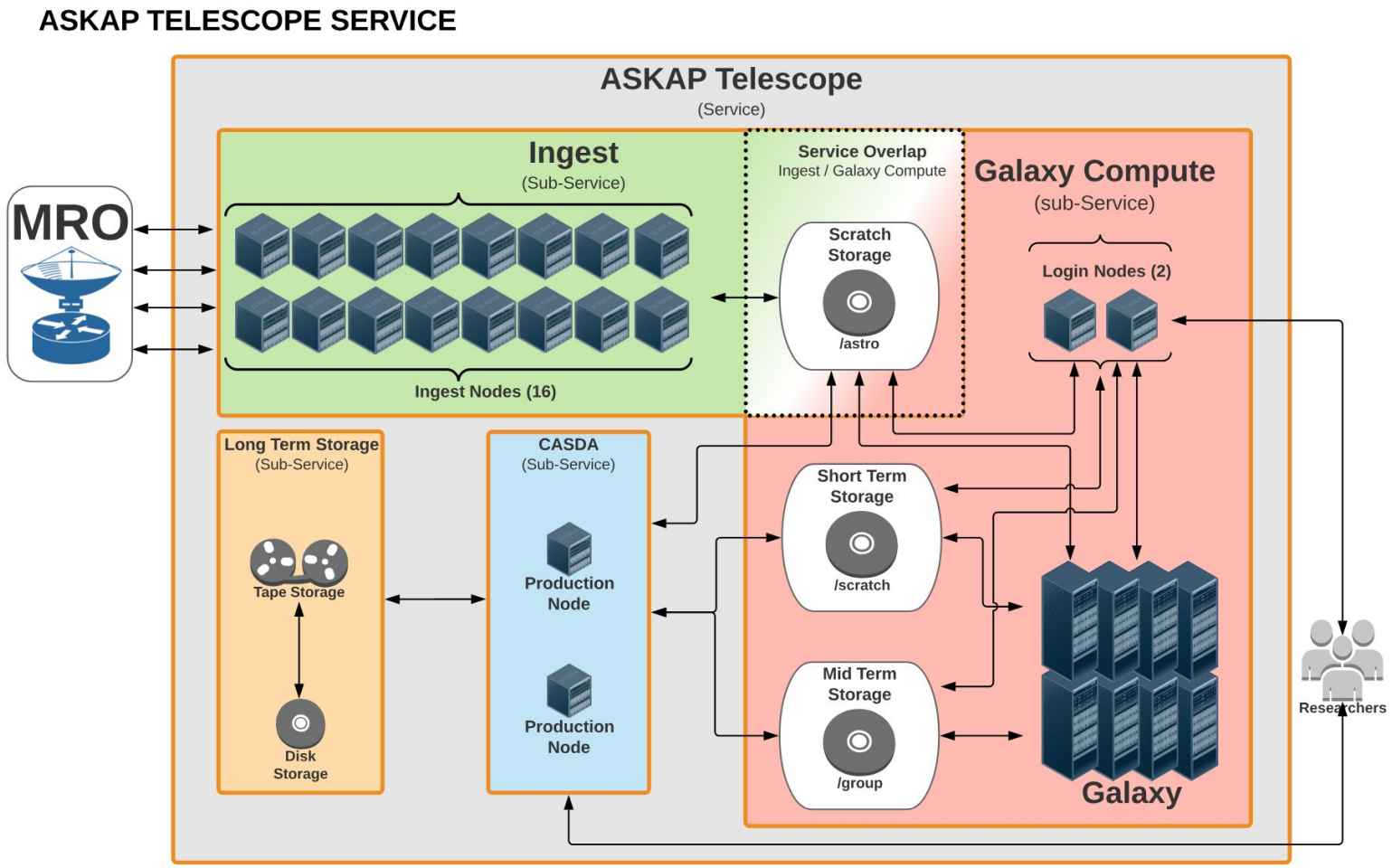
PAWSEY supercomputing centre

Allocation per institution



PAWSEY supercomputing centre

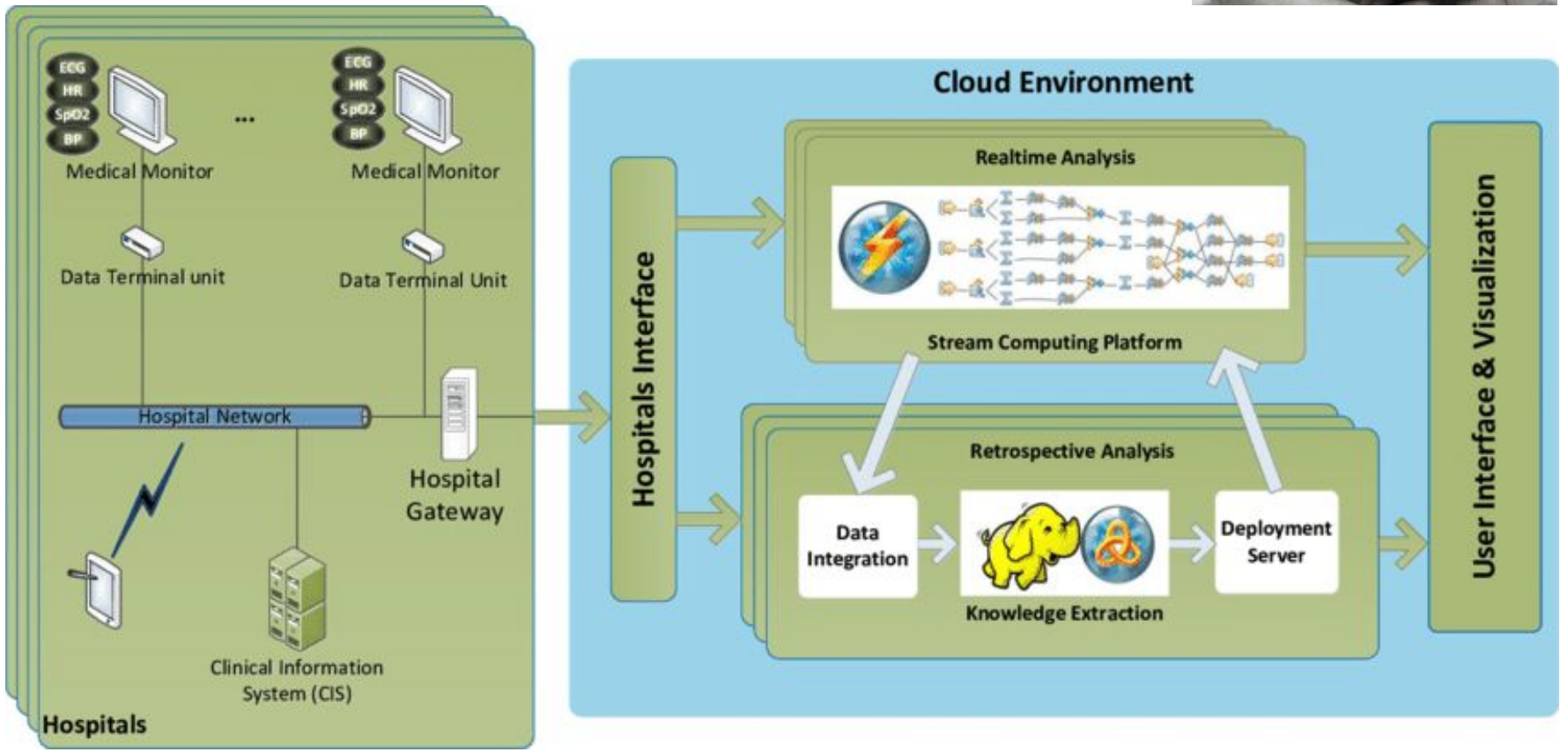
Real-time data ingest



PAWSEY supercomputing centre

Real-time data ingest

ARTEMIS: A neonatal Internet of Things.



PAWSEY supercomputing centre

Capital Investment Status

- Currently at the start of a $70 million capital refresh
- Approximately 2/3 for compute, 1/3 for storage and network
- A few small procurements soon out to market (storage, network, cluster, cloud, visualisation)
- Collaborations started or about to start with a number of institutions

Future of HPC

- Unified model:
  - From HPC, Cluster, Cloud Converged architecture
  - Software defined infrastructure
- Science from data
- Machine intelligence:
  - Machine learning
  - AI-augmented scientific simulations
- Synergies: solve large scale science problems collectively (e.g. SKA, LHC)

Future of HPC

- Integrated workflows
- New scientific domains:
  - Genomics
  - Phenomics
  - Medical Informatics
  - Advanced Imaging (including medical)
- Dynamic, on-demand, real-time job scheduling
- Security, handling of sensitive data

Challenges

- Our scientific community is a diverse group
- Communication
- Data intensive science
- Quasi real time HPC
- On demand compute – may need to “burst”

Challenges

- Storage: how can we support growth?
  - Funding
  - Physical space
  - Bandwidth
- Data privacy and security
- Heterogeneous workloads
- Data Management
  - no scalable data management solution currently implemented
- Development
  - Support for flexible, on-demand, workloads beyond simple batch scheduling
- Quality of service
  - Regressions
  - Uptime
  - Consistency of environment and applications across maintenance sessions
  - Consistent performance

Solutions (?)

- **Storage**
- **Heterogeneous workflows**
- **Data Management**
- **Development workloads**
- **Security**
- **Quality of Service**

- **Hybrid cloud**
- **Federation**

- **Workflow mgmt.**
- **Interactive Vis.**

- **RUCIO**
- **Meerkat**

- **SchedMD/Elastic workloads**

- **Smart contracts/blockchain**

- **Re-frame**
- **Ganglia**
- **Containers**
- **Secondment**

PAWSEY supercomputing centre

Questions?

PAWSEY
supercomputing centre