



Support for Large Scale HPC and Big Data Science Projects in South Africa: SKA and More

...

25th September 2017

AN INITIATIVE OF:



science
& technology

Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA

MANAGED BY:

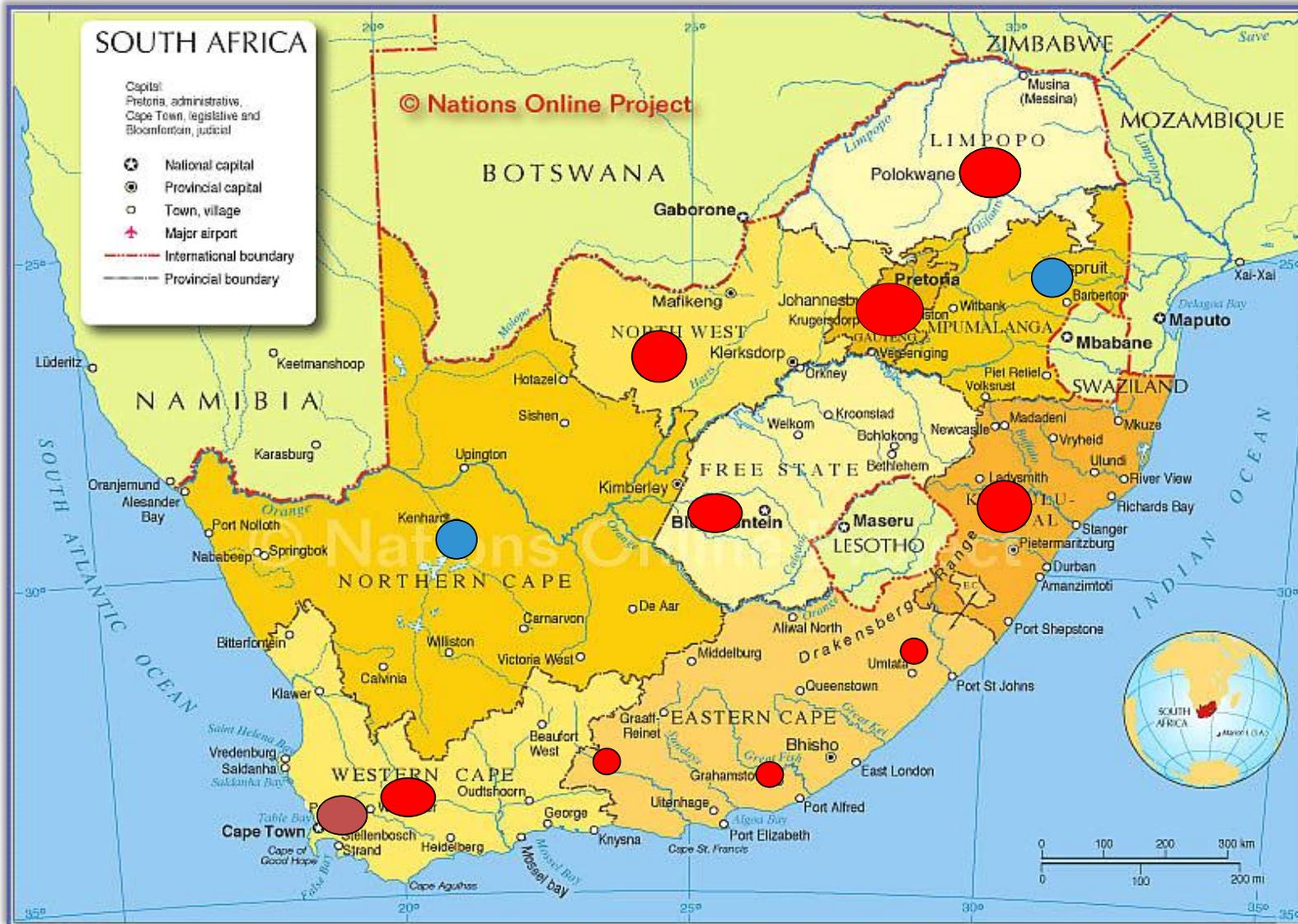


Background: The CHPC in SA

CHPC



User Community: South Africa



- CHPC
- Existing Users
- Future Users

Background

- ❑ CHPC is **national HPC facility** funded by the DST and managed by CSIR
- ❑ It is part of the National Integrated Cyber- Infrastructure System (NICIS)
- ❑ Started operations in **June 2007**
- ❑ Staff complement of 40
- ❑ CHPC has **Research, Technical** and **Operational** divisions
- ❑ Provide HPC resources and services to universities, science councils and industry
- ❑ Provide support to large scale science projects such as H3Africa, CERN, SKA and IPCC
- ❑ Human Capital Development is core to CHPC success

User Community: Engagement Strategy

Special Interest Groups (SIG's) / Research Domains

Advanced Computer Engineering
Astronomy/Cosmology
Bioinformatics
Epidemiology
Chemistry
Computer Science/GRID
Computational Finance
Earth Sciences
Humanities
Physics
Remote Sensing
Materials Science
Visualisation

Activities
**National HPC
Infrastructure for Users**
(Academic + Non-Academic)

Flagship Projects
(Currently referred to as collaborative
projects)

Training Programs
(Workshops, Short-Courses , outreach
activities amounts to more than 80
events to date)

Annual CHPC National Meetings
(10 successful meetings to date
2017 event: 4 – 9 Dec 2017, Durban)

CHPC



Partners

Higher Education Institutions
Government Departments
Science Councils
Research Agencies
Industry
International Collaborators
Policy Makers
Outreach Organisations

Investment in HPC Systems



2011: Tsessebe Cluster
Upgraded 2.9GHz (4032 cores)
61.6 Tflops total (Linpack)

2007: IE
cores
2.5

Hz; 640

90 TB GP File System

4 PB DIRISA
Storage
Unit



2009 : Sun Constellation Cluster
2.9/3.0 GHz; 2684 cores
27 Tflops (Linpack)
480 TB Lustre File System



The Road to PetaFLOP

3 May 2016 **7 March 2017**

System Configuration	Phase 1	Phase 2
Dell PowerEdge C6320 Servers:		
Standard Compute nodes 128GB (64GB) / node	1 008	1 368
2 x Intel Xeon E5-2690 v3 (Haswell) processors (12 Cores Each \Rightarrow 24 cores / node)	24 192	32 832
Dell PowerEdge R930 servers:		
Large Memory Compute Nodes 1024GB / node (FAT nodes)	5	5
4 x Intel Xeon E7-4850 v3 processors (14 Cores Each \Rightarrow 56 cores / node)	280	280
Infiniband FDR 2:1 Blocking (56 Gbps)		
Parallel Storage (Useable) PB	4	4
Total Number of Racks (including Compute, Login, Management and Storage Nodes)	19	24
Centos 7.1 with Bright Cluster Manager and Altair PBS Pro		
Total Linpack Performance (Tflop/s)	783	1029

Launch of the HPC system

CHPC
CENTRE FOR HIGH
PERFORMANCE COMPUTING



- **161 on TOP500**
- **Awarded the fastest supercomputer in Continent**



science
& technology
Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA

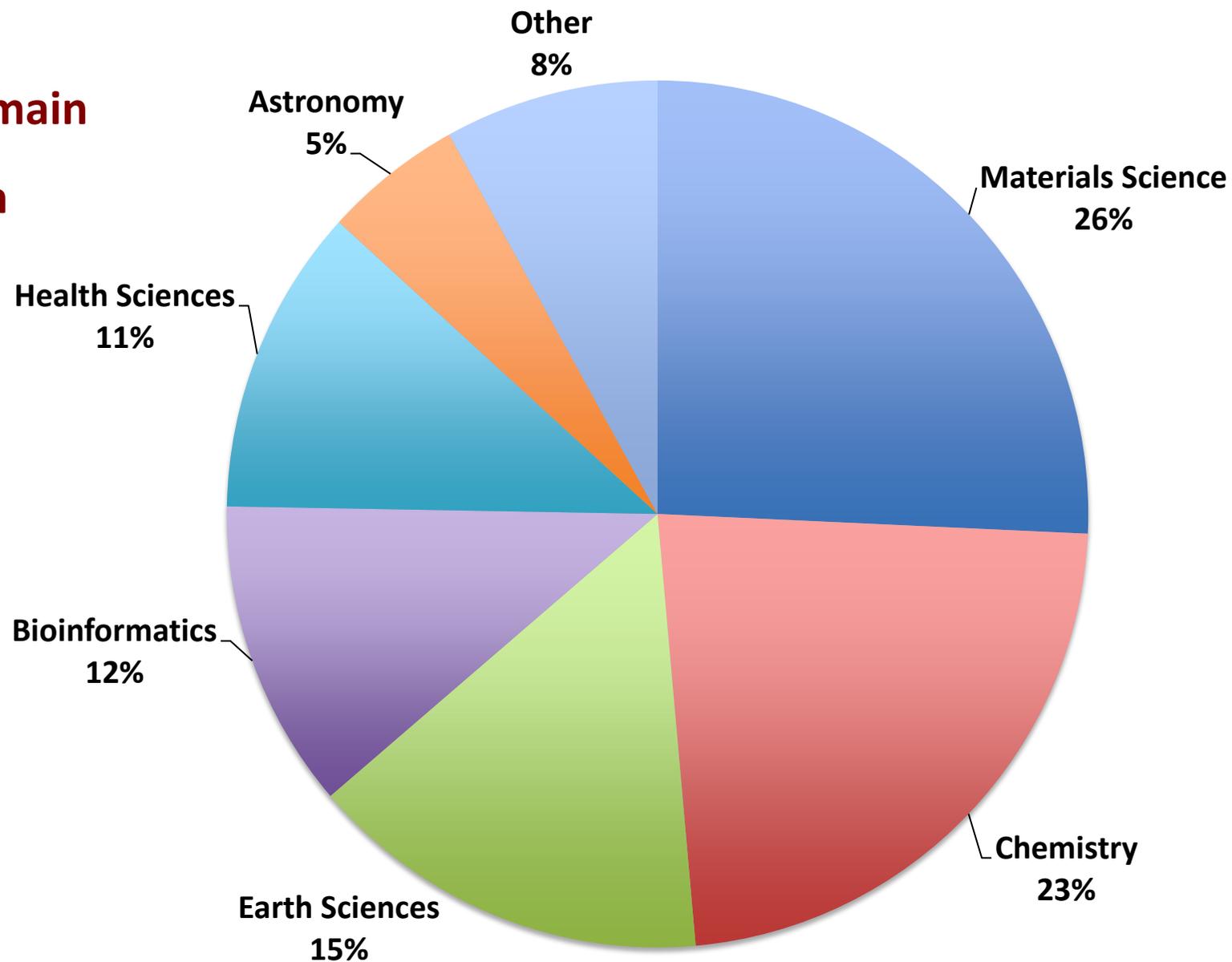
CSIR
our future through science

Resources at the CHPC

CPU Hours:

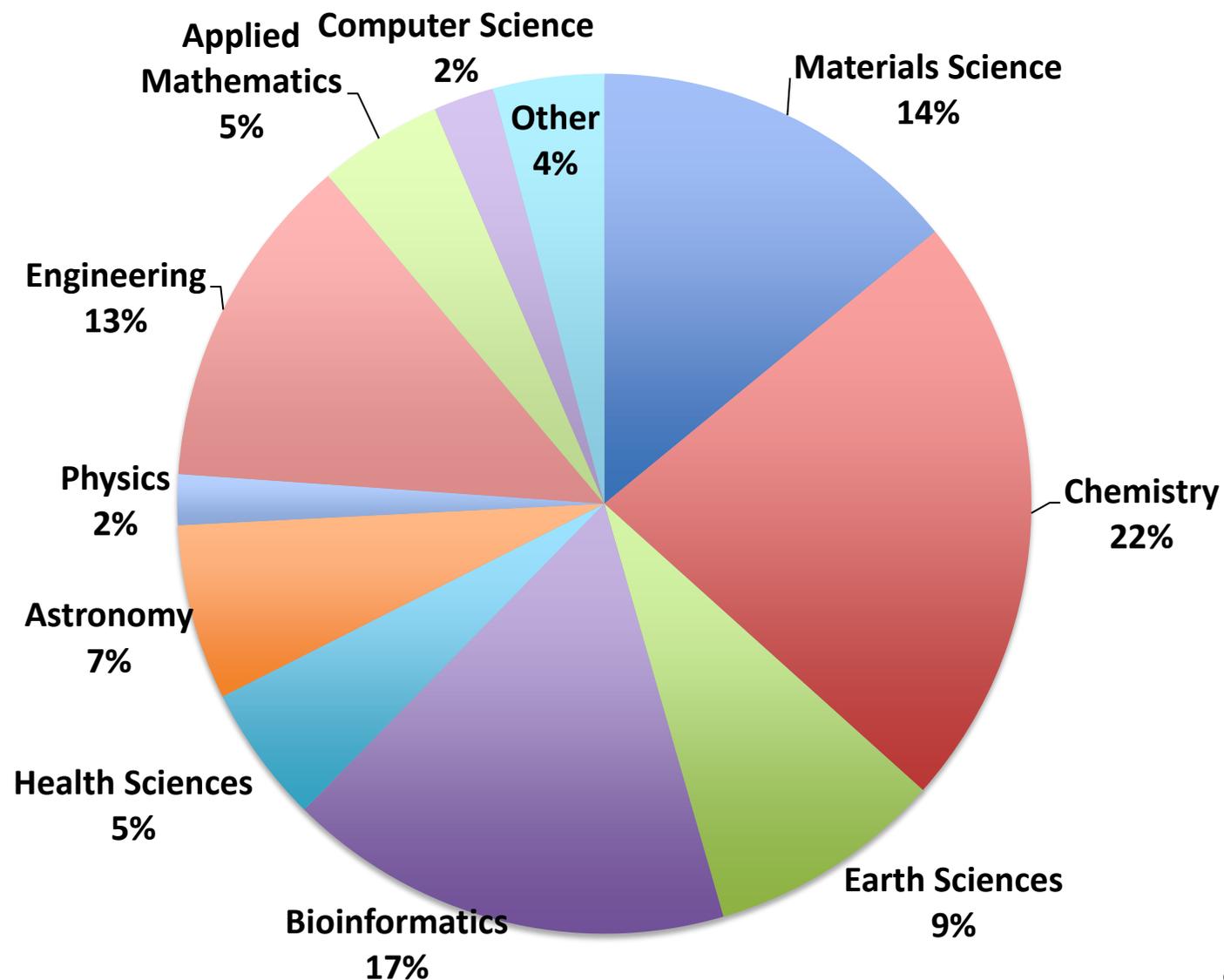
Science Domain

Distribution



Who is using the CHPC?

Research Programs: Research Domain Distribution (210 Programs)



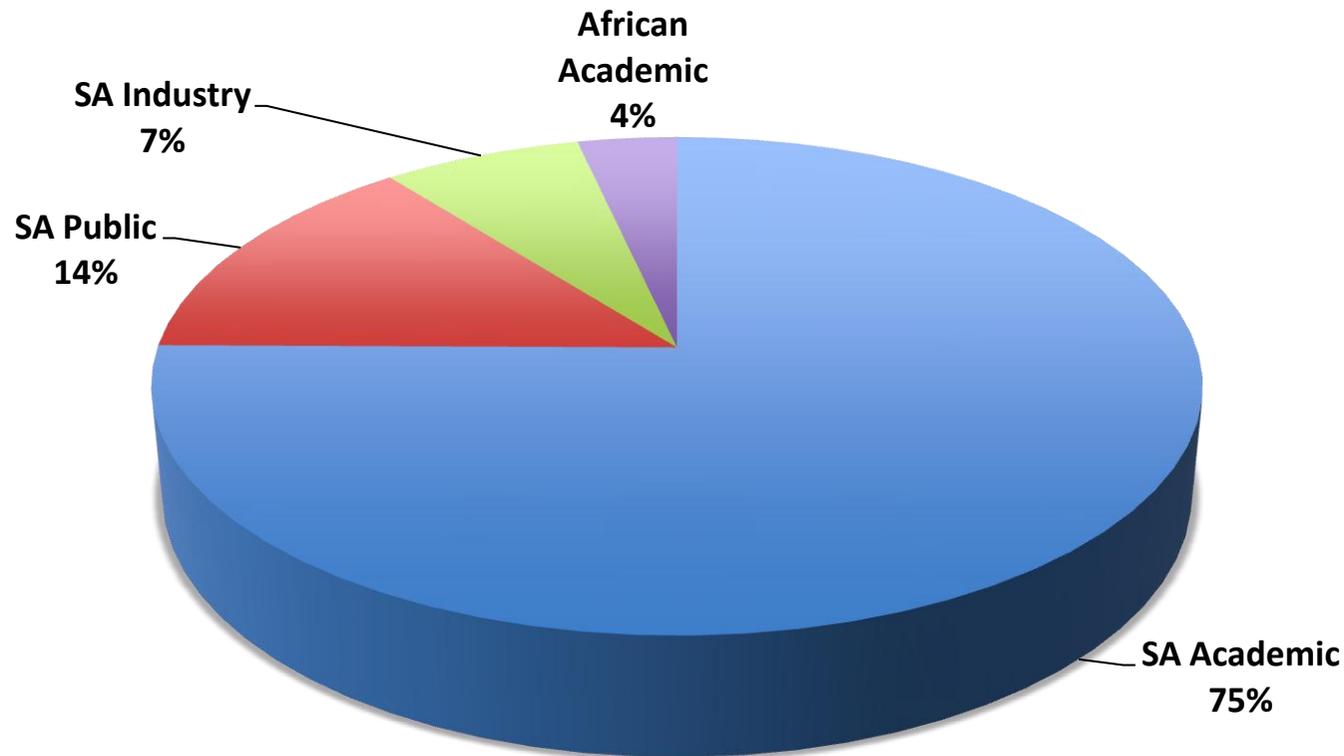
Who is using the CHPC?

Non-Academic Public

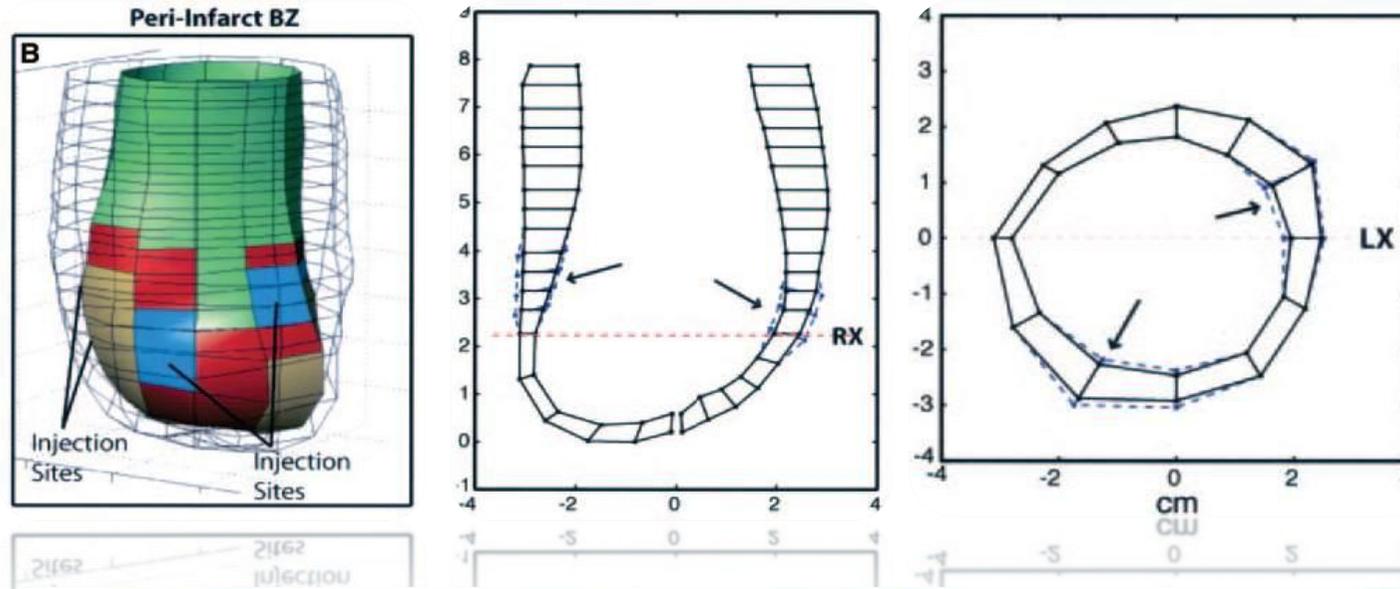
- CSIR
- SAWS
- ARC
- NZG
- SAAO
- SANBI
- Sugar Institute
- NICD
- NECSA

Research Programmes

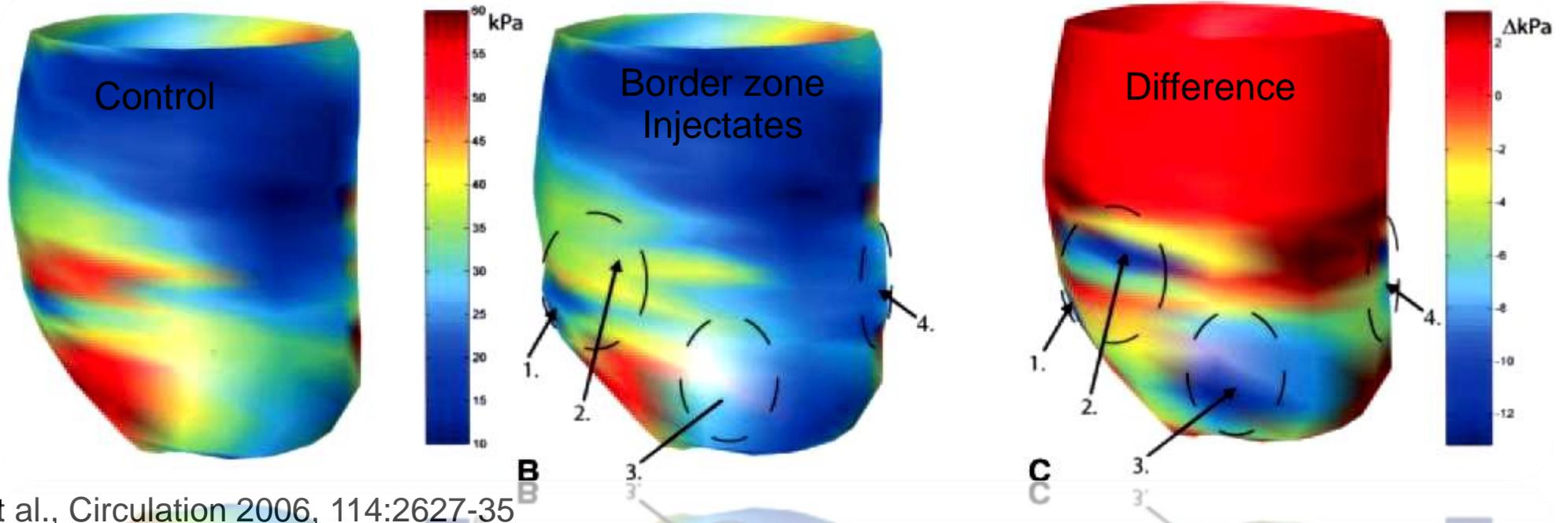
Distribution



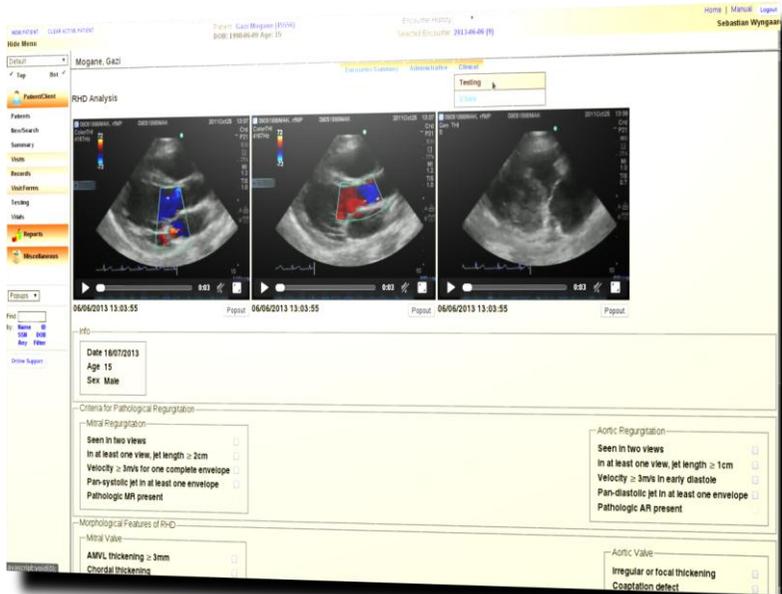
Injectate Mechanics



Reduction of
Border Zone
Fibre Stress



Rheumatic Heart Disease Progress



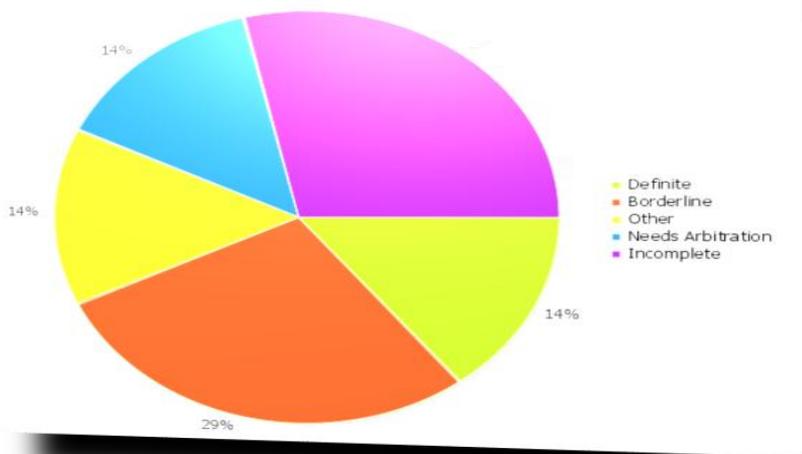
The Portal helps medical practitioners to analyse and diagnose echocardiograms for RHD.

The system automatically tracks statistical data on the number of patients processed and catalogs diagnostic results.

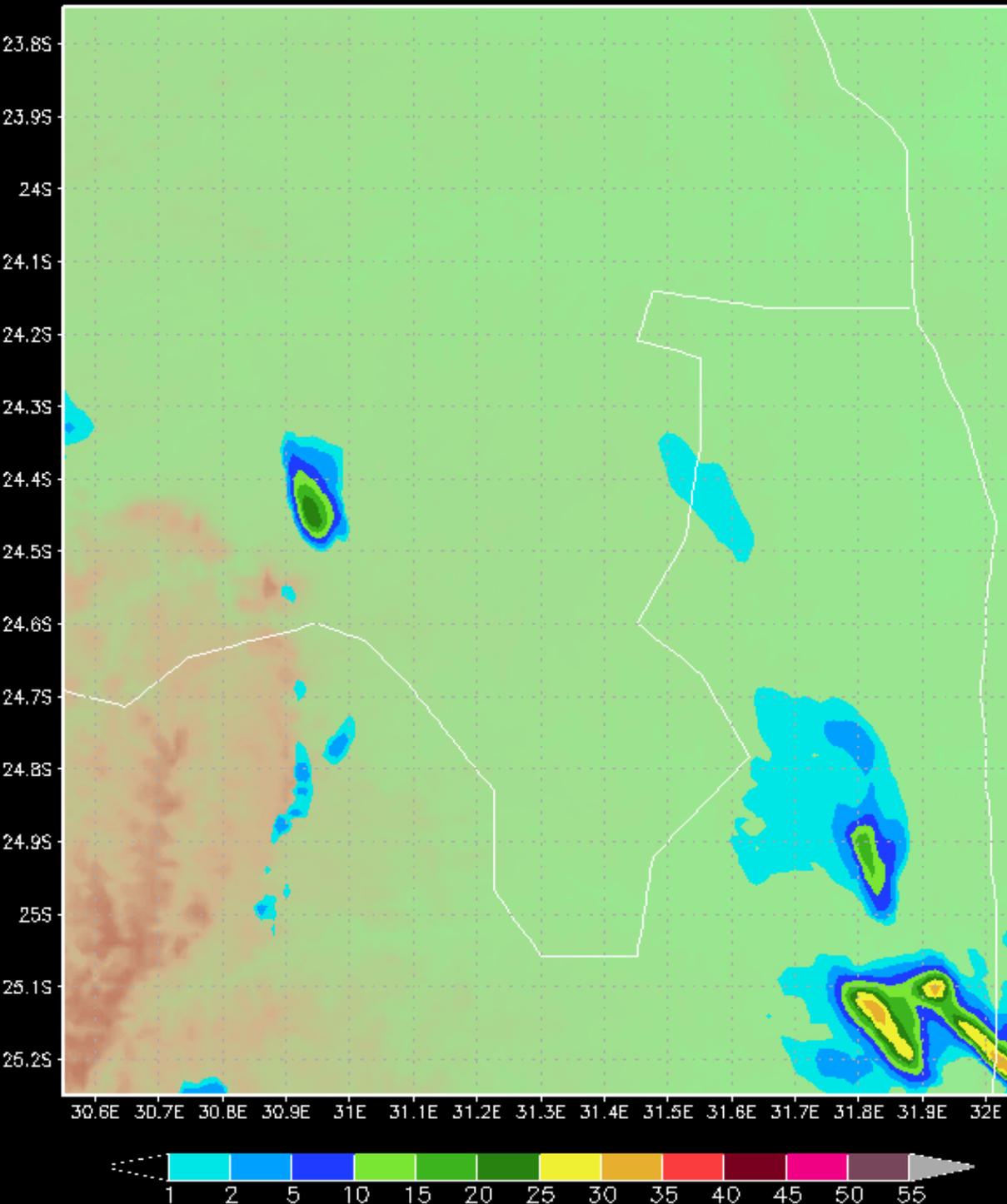
There are currently over 2 000 patients registered on the system, and 1 800 of them processed.

Raw images for the patients (DICOM) is contained in 200 000 files and account for 4TB storage in DIRISA.

Additional 3000 patients will be added to the system to extend this portal for the rest of South Africa.



CCAM 1km res Skukuza
00Z26JAN2010



- Simulation period:
2009-2012
- Multiple nudging: ERA
reanalyses
 - 75km to
 - 8km
 - (1300kmx1300km
domain size)
 - 1km (150kmx150km)

Human Capital Development

Dedicated User Training Program (Winter school, Introductory Programming School HPC Ecosystem, SKA Readiness Program, Domain specific workshops etc..)

Training undergraduate students in HPC

Three stage program:

(1) Winter school: **22 teams**, (2) CHPC National Meeting: **10 Teams**, (3) ISC: **1 Team** representing SA

**ISC'13
Champions**



Leipzig

**ISC'14
Champions**



Leipzig

**ISC'15
2nd Place**



Frankfurt

**ISC'16
Champions**



Frankfurt

**ISC'17
2nd Place**

Frankfurt

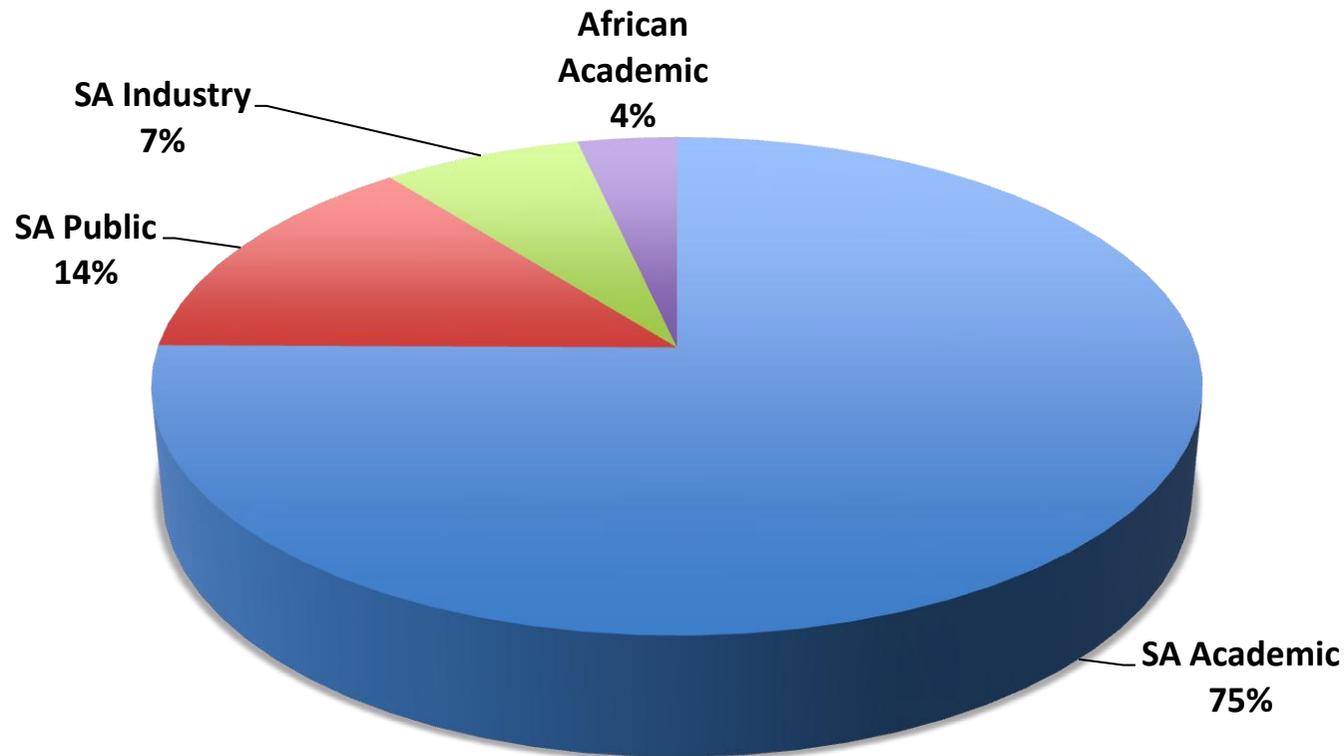
Who is using the CHPC?

SA Industry

- Mintek
- Eskom
- Transnet
- De Beers Marine
- Johnson Matthey
- Hatch
- Roche
- Inqaba Biotec
- MTech
- eScience

Research Programmes

Distribution



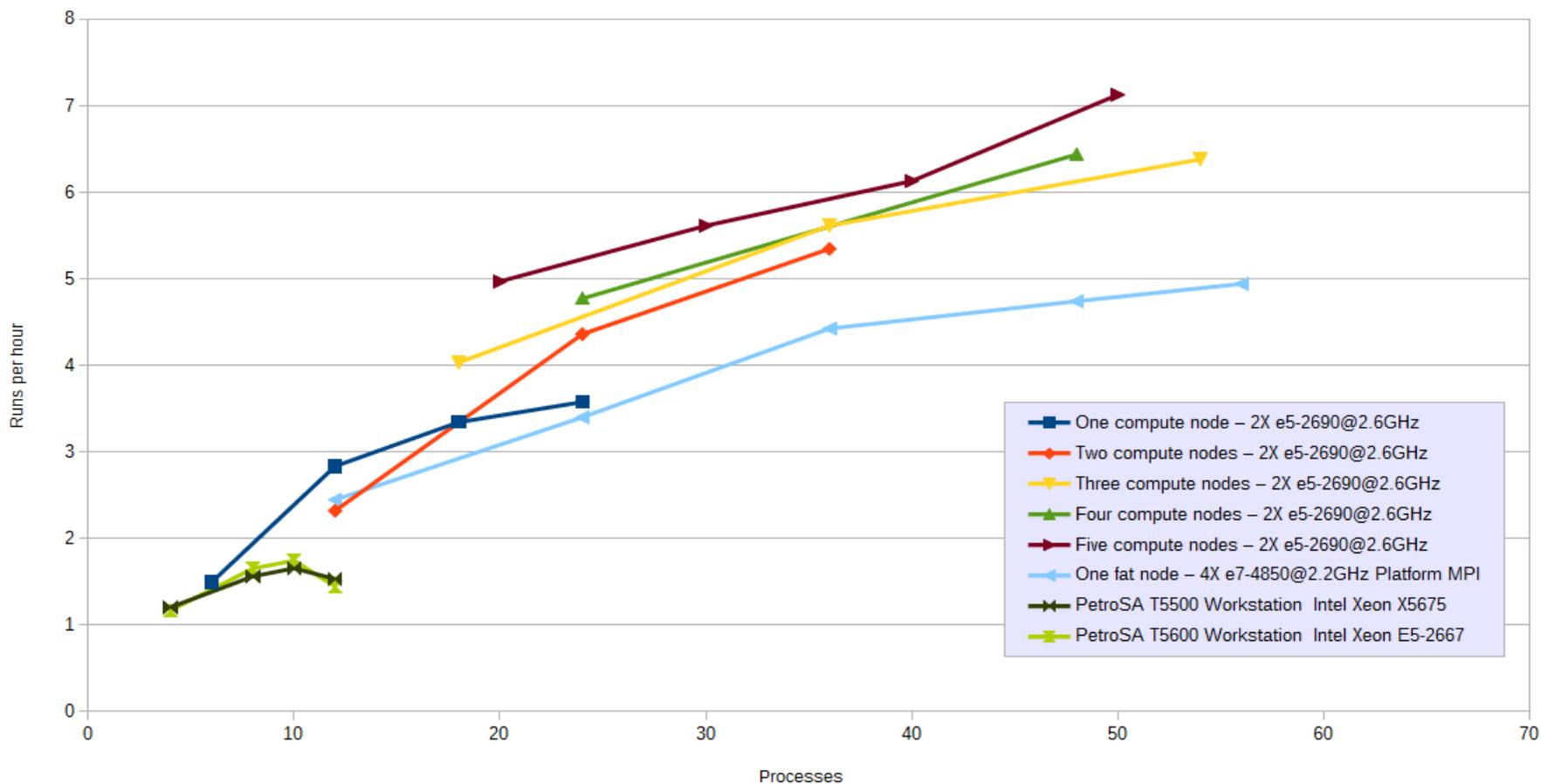
Industry Partners



Industry Support

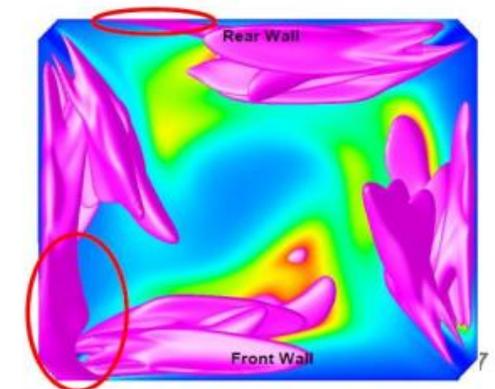
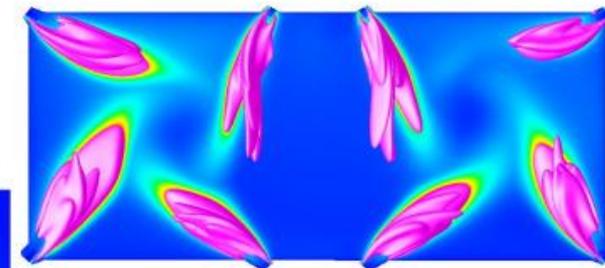
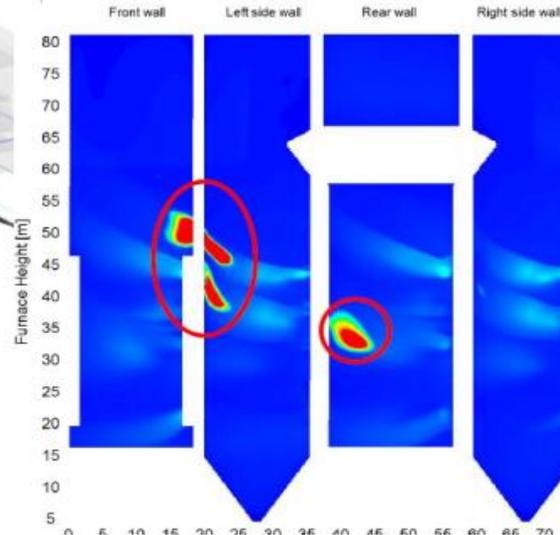
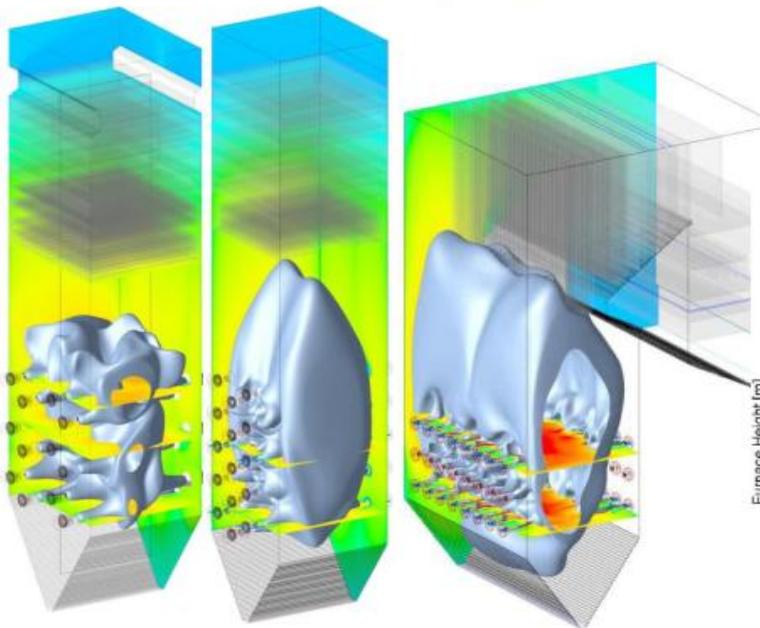
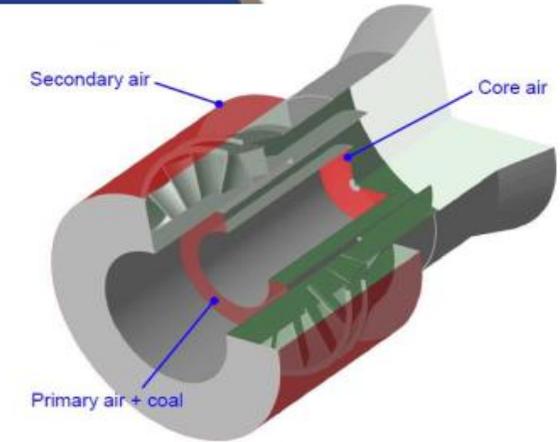
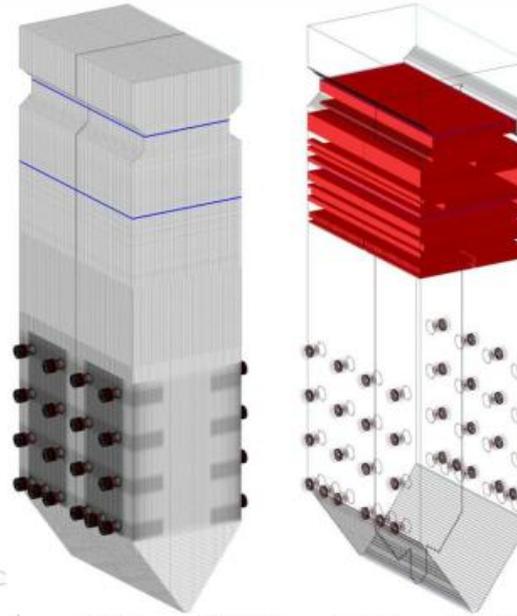
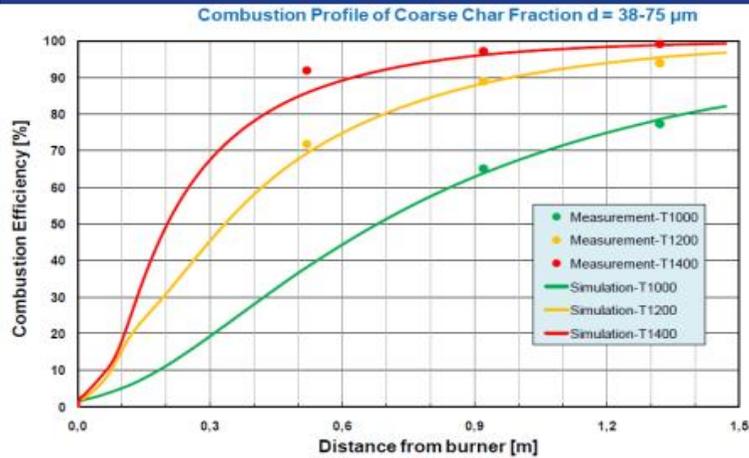
PetroSA Eclipse Benchmark Case

Dimensions 149X118X128 - 2.25 Million cells



Industry Partner Example

Combustion CFD Modelling

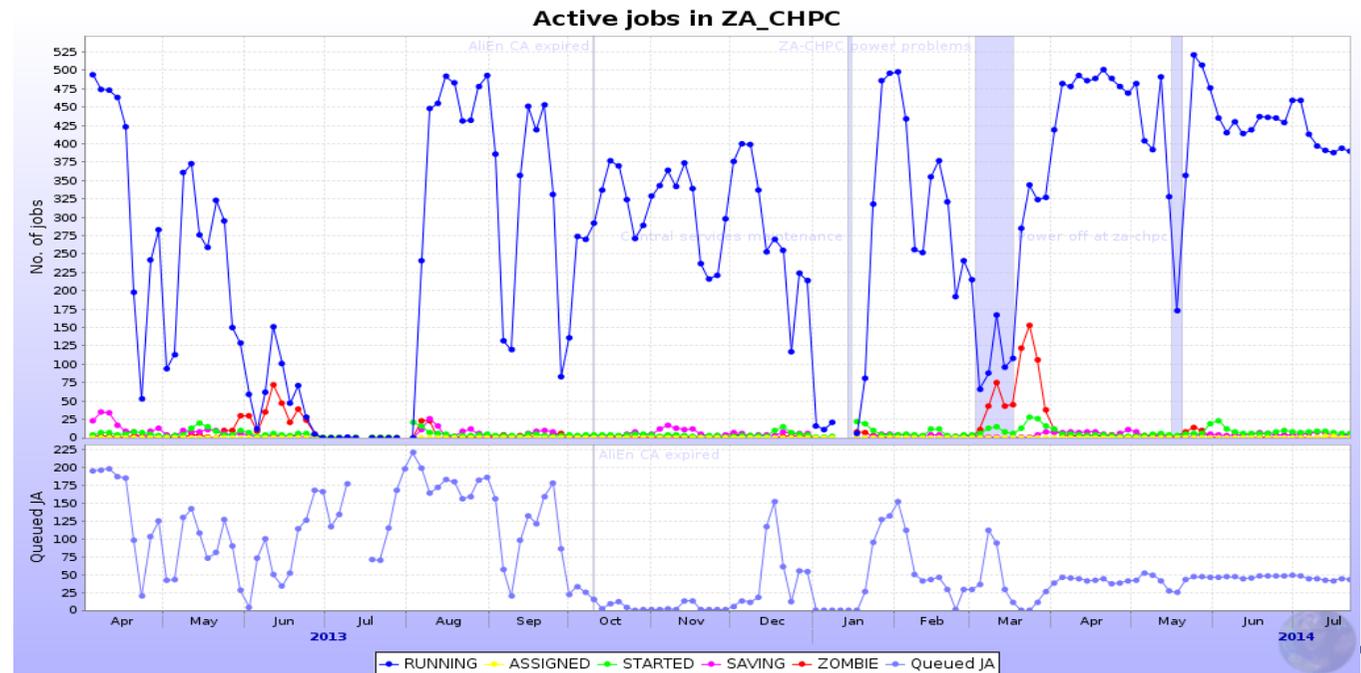


2017/03/07



Support for CERN Experiments

- Dedicated cluster for the CERN experiments. Mainly ALICE and ATLAS.
- Currently providing support for 2 400 jobs/day (increase in international bandwidth will increase the # of jobs).
- Plans to move to WACS cable for better international connectivity.
- Officially signed the MOU for a Tier-2 facility in April 2015.
- Plans to expand the service for Tier-1 in future.



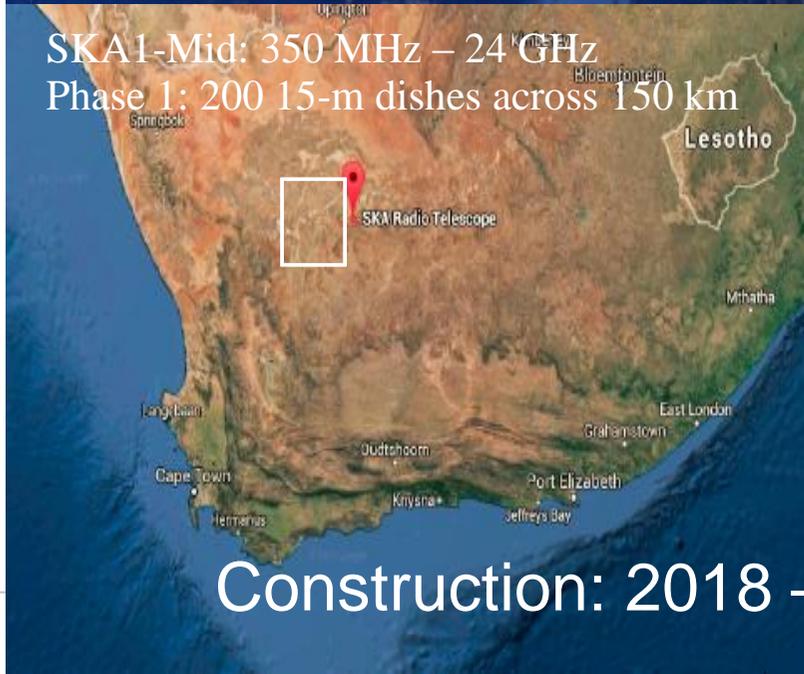
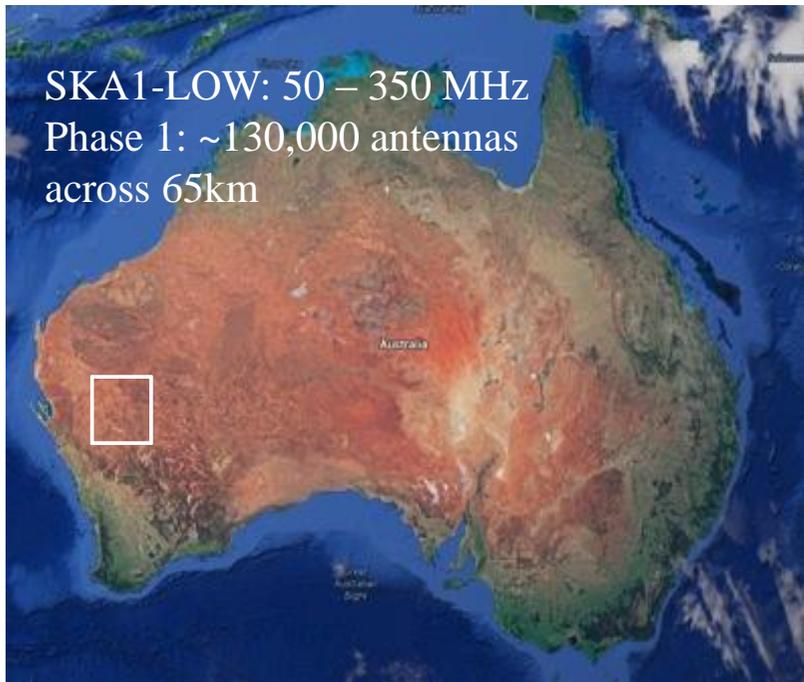
SKA Bi

Outcome announce



- Originally SKA South Africa was administered directly by the Department of Science and Technology (DST) in South Africa;
- In 2005 the National Research Foundation was requested by DST to take over the administration of SKA South Africa;
- SKA SA and Hartebeesthoek Radio Astronomy Observatory is currently merging into one national facility.

SKA1: HQ in UK; telescopes in AUS & RSA



Construction: 2018 – 2024; Cost cap: €675M

Square Kilometre Array Phase

3 sites: 2 telescopes + HQ
1 in a nutshell
= 1 Observatory

Design Phase: > €170M; 600 scientists + engineers

Phase 1

Construction: 2018 – 2024

Construction cost cap: €675M (inflation-adjusted)

Operations cost: under development (see below)

MeerKat integrated

Observatory Development Programme (€20M/year planned)

SKA Regional centres out of scope of centrally-funded SKAO.

Phase 2: start mid-2020s

~2000 dishes across 3500km in Africa

Major expansion of SKA1-Low across Western Australia

SKA Global organisation

- Australia (DoI&S)
- Canada (NRC-HIA)
- China (MOST)
- India (DAE)
- Italy (INAF)
- Netherlands (NWO)
- New Zealand (MED)
- South Africa (DST)
- Sweden (Chalmers)
- UK (STFC)



- Full members
- SKA Headquarters host country
- SKA Phase 1 and Phase 2 host countries

Interested Countries:

- France
- Germany
- Japan
- Korea
- Malta
- Portugal
- Spain
- Switzerland
- USA

Contacts:

- Mexico
- Brazil
- Ireland
- Russia



- African partner countries (non-member SKA Phase 2 host countries)

SKA Science Drivers – the history of the universe

Cosmic Dawn
(First Stars and Galaxies)

Testing General Relativity
(Strong Regime, Gravitational Waves)

Galaxy Evolution
(Normal Galaxies $z \sim 2-3$)

Cradle of Life
(Planets, Molecules, SETI)

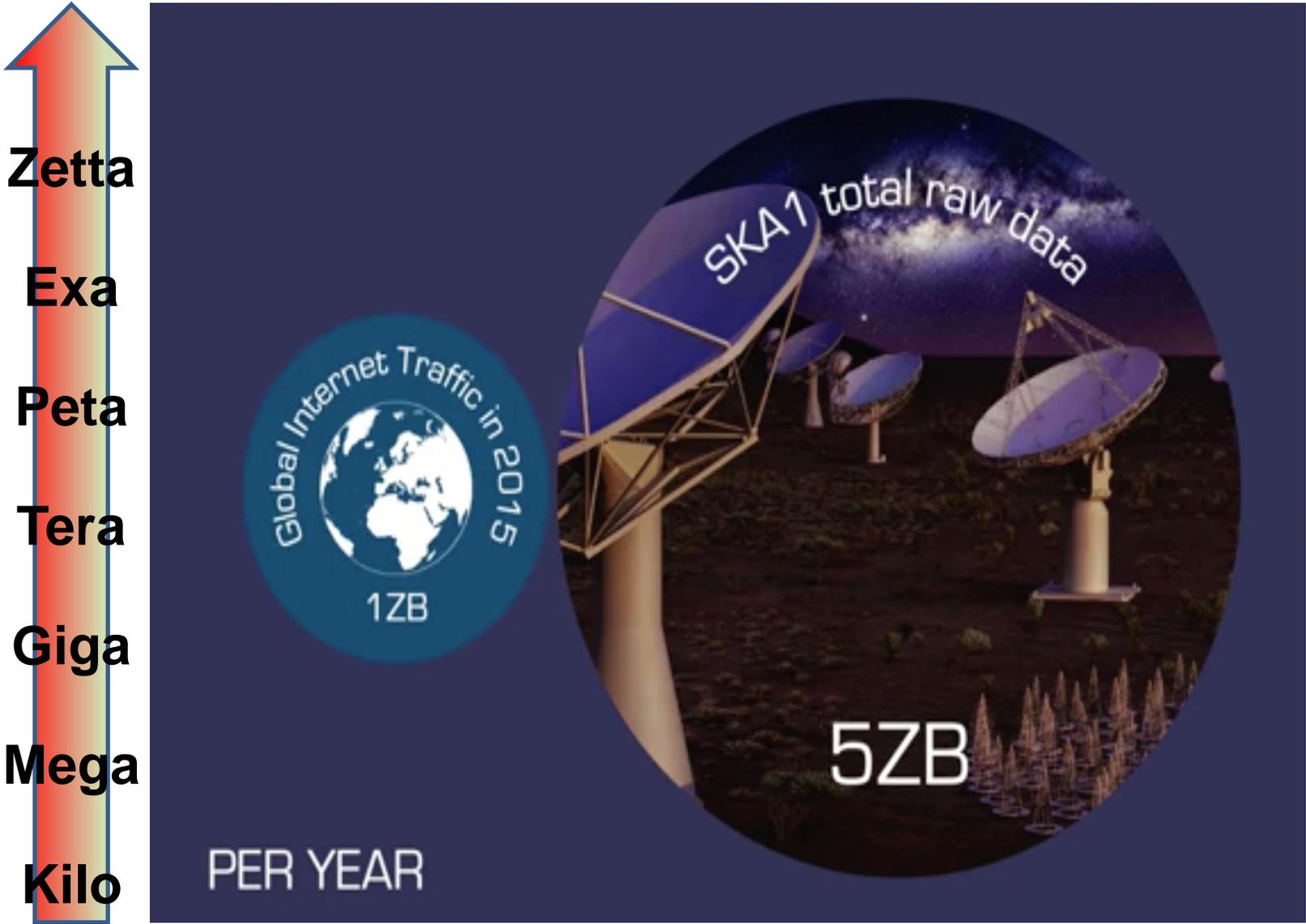
Cosmology
(Dark Energy, Large Scale Structure)

Cosmic Magnetism
(Origin, Evolution)

Exploration of the Unknown

Extremely broad range of science!

SKA Big Data challenge

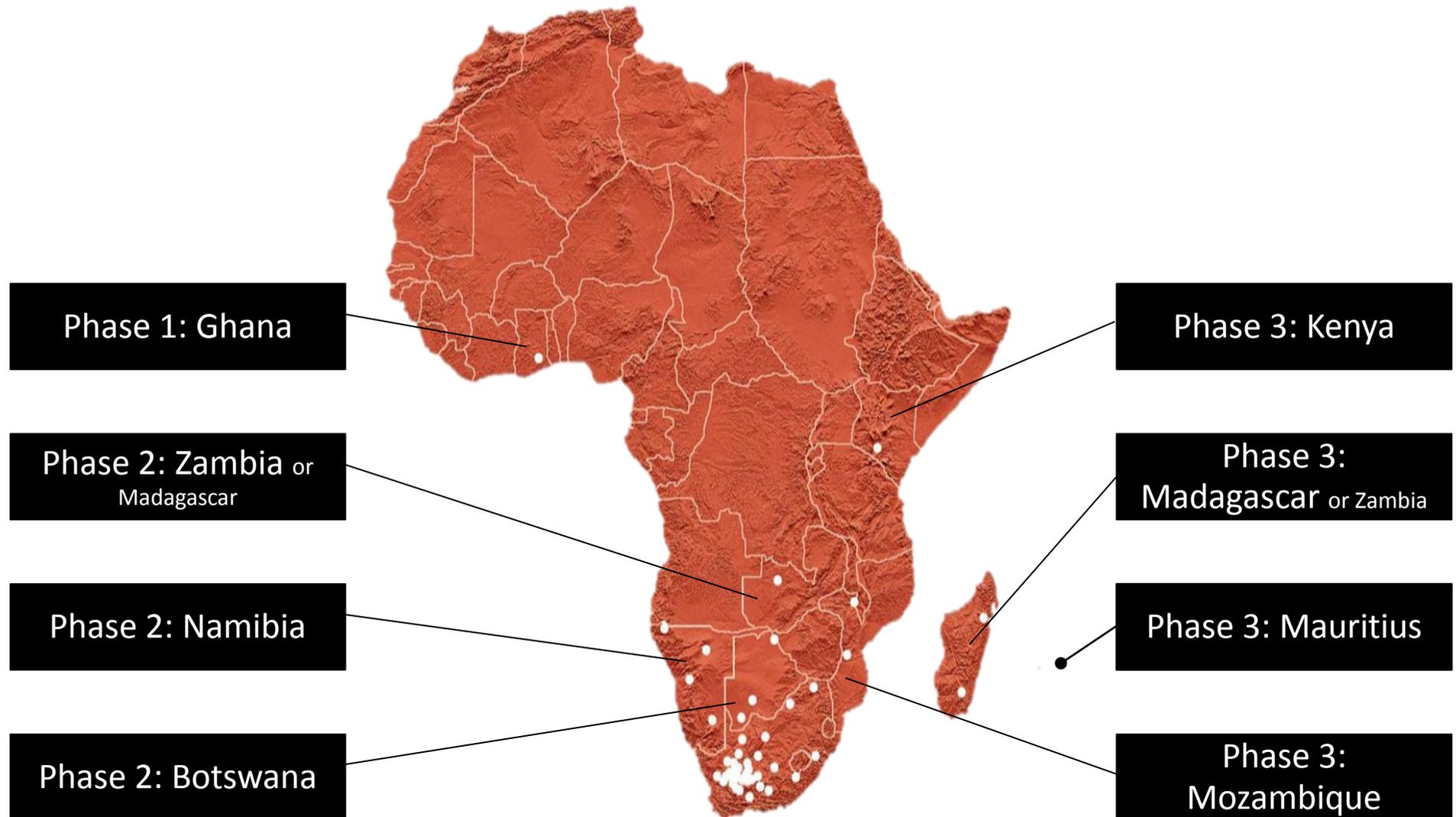


SKA precursors



The AVN project

African VLBI Network of Telescopes in Africa - ensuring readiness for SKA2



SKA in South Africa (SKA SA)

A South African Radio Astronomy Observatory

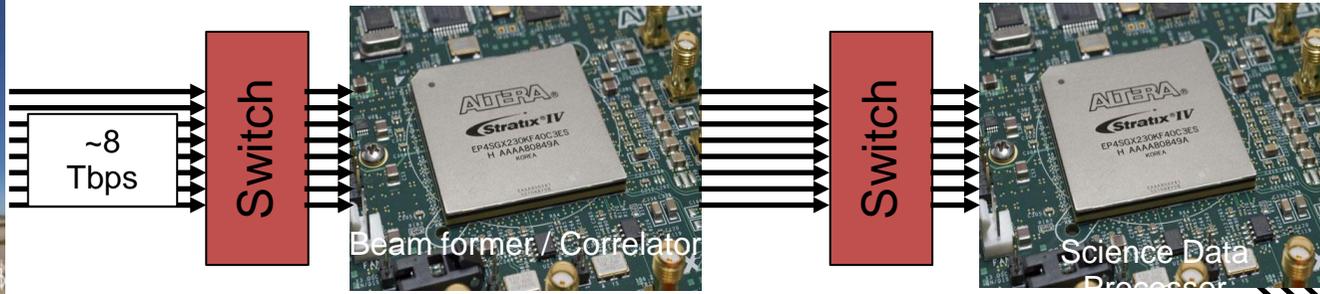
- Until 31 March 2017 the SKA SA programme was structured around specific projects with project teams defined and recruited on the basis of the needs of the specific engineering projects and interventions;
- With the SKA SA & HartRAO merger, the newly formed entity will become an National Research Foundation facility and will be re-structured in order to maximise efficient use of resources;
- Current engineering projects and initiatives within SKA SA include:
 - The global SKA telescope project;
 - The MeerKAT telescope project;
 - The African Very Long Baseline interferometry network of telescopes (AVN) project(s);
 - The Hydrogen Epoch of Reionisation (HERA) telescope project;
 - Big Data initiatives;
 - Infrastructure for MeerKAT, SKA and other telescope or geodesy projects.



SKA Data flow schematic



SKA1-MID



AFRICA

AUSTRALIA

SKA Regional Science Centre



100 PFLOPS

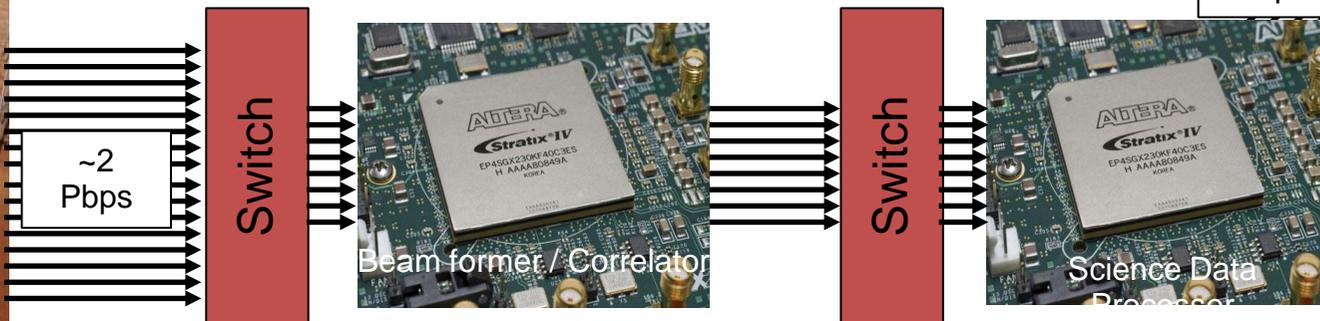
Science Data Archive

~5 Tbps

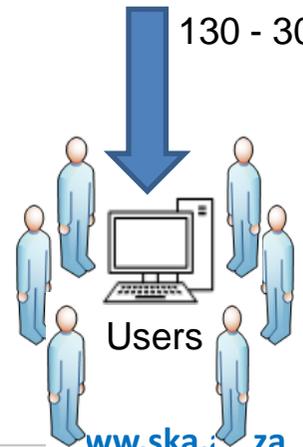
~5 Tbps



SKA1-LOW

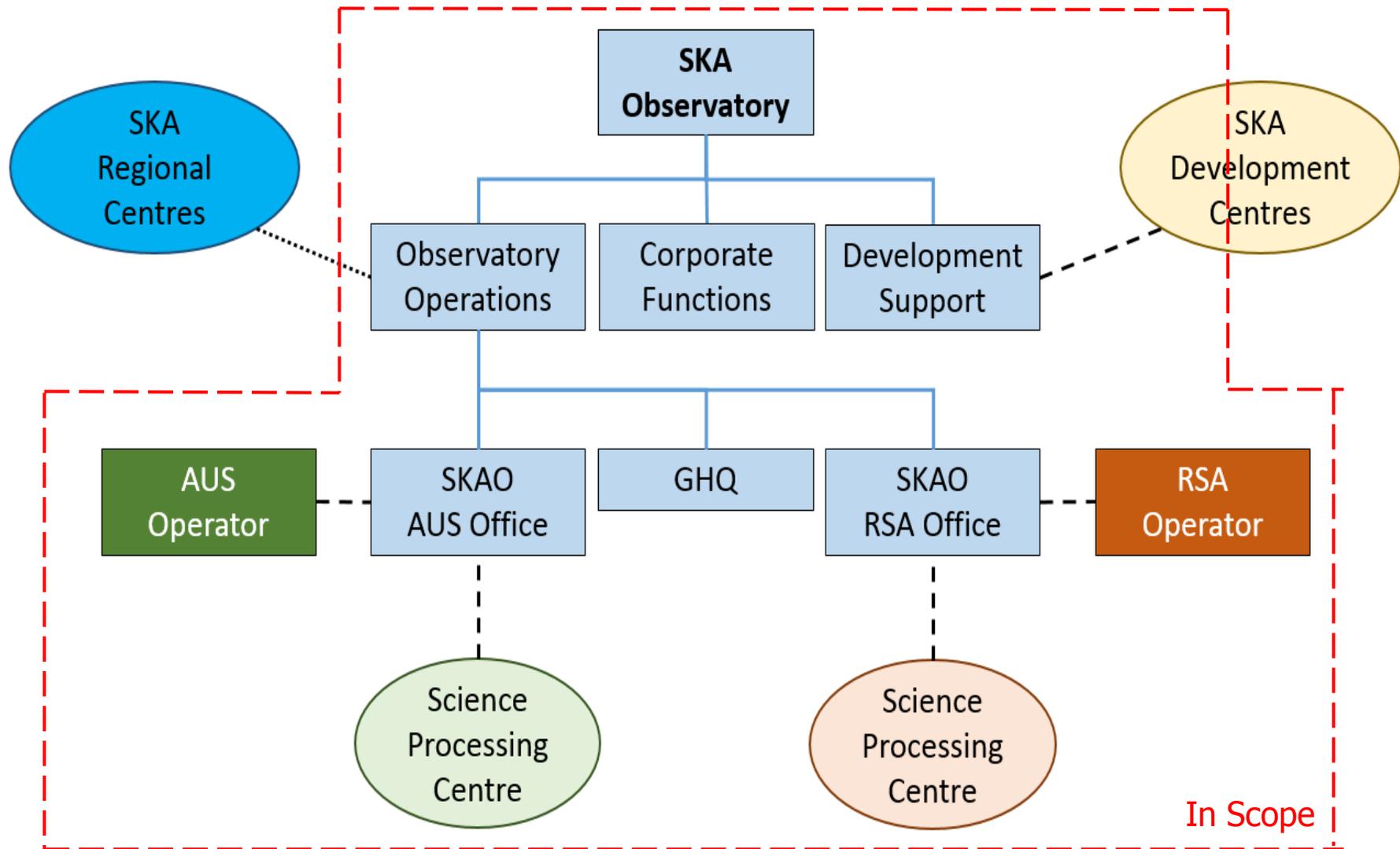


130 - 300 PB/y



Users

SKA1 Scope



----- Service Level Agreements

..... Memorandum of Understanding

2nd Chance for retiring HPC systems



HPC Systems Repurposing



- ❑ Strategy to **repurpose HPC systems** for **local compute** or **training** needs
- ❑ **HPC Ecosystem initiative** ⇒ Provide mid-range **HPC systems** to **universities**
- ❑ One of **key** strategic **initiatives** of the CHPC
- ❑ CHPC initiated the **RANGER Project** – Support to institutions
- ❑ Received resources from other collaborators, like Cambridge.
 - ⇒ **Do not have** HPC facilities
 - ⇒ Need of HPC resources for **learning**

HPC Ecosystem in South Africa

	Operational	Used	Needs Support Equipment To Operate	Power & cooling
University of Fort Hare	Green	Red	Red	Green
University of Kwa-Zulu Natal	Yellow	Yellow	Green	Yellow
University of Venda	Green	Red	Green	Green
University of Witwatersrand	Green	Green	Green	Green
Stellenbosch University	Green	Green	Green	Green
Sol Plaatjie University	Red	Red	Red	Red
North West University	Green	Yellow	Green	Green

- Re-purposed HPC system configured for local computational capacity.
- Emphasize on training system administrators at universities to support local users.
- Partners from UK and USA provides HW and courses.

Progress with HPC in Africa

	Infrastructure			Staff			Development		Status of Hardware		
	Power & cooling	Datacentre ID'ed	Equipment shipped	SysAdmin / Linux	Dedicated time	SysAdmin Course	Projects	Working Group	Operational	Used	Needs Support Equipment To Operate
Mauritius: University of Mauritius (UoM)	3	3	3	2	3	3	2	3	3	3	3
Namibia: University of Science & Technology (NUST)	3	3	3	3	3	3	2	3	3	2	3
Namibia: University of Namibia (UNAM)	3	3	3	3	3	3	2	3	3	2	3
Botswana: University of Botswana (UB)	3	3	3	3	3	3	3	3	3	3	3
Zambia: ZAMREN	3	3	2	3	3	3	3	3	3	3	3
Madagascar: University of Antananarivo (IOGA)	3	3	2	2	2	2	2	2	2	2	2
Ghana:	3	3	3	2	3	2	2	3	2	2	3

Completed

Progressing

Not Started



CHPC National Meeting

2017 CHPC National Meeting: **Pretoria, Velmore Hotel Estate**

3-7 December 2017

www.chpcconf.co.za



Acknowledgements

**SKA –SA and SKA-O
Anita Loots and Rob Adam**

CHPC Team

hsithole@csir.co.za

Thank You.



CHPC
CENTRE FOR HIGH PERFORMANCE COMPUTING

