CHPC
CENTRE FOR HIGH PERFORMANCE COMPUTING

The Centre for High Performance Computing

The Role of the CHPC within eResearch in South Africa

Dr Werner Janse van Rensburg
24 November 2014

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

CSIR
our future through science

Outline

- Background
- User Community
- Application Research Areas
- Training Initiatives
- HPC Platforms
- Acknowledgements

Background

- CHPC is **national HPC facility** funded by the DST
- Managed by the **CSIR** via the Meraka Institute (ICT)
- Started operations in **June 2007** and based in Cape Town
- Currently hosts the **largest HPC systems** on **African** continent
- **Staff** consists of research, technical, support and studentships
- Totalling a number of ~**30** employees

Background

- CHPC **mandate**: ⇒ Provide **computational resources** to SA research community

  ⇒ **Support research** user groups (public and private)

  ⇒ Support **HCD** through training + funding

VISION

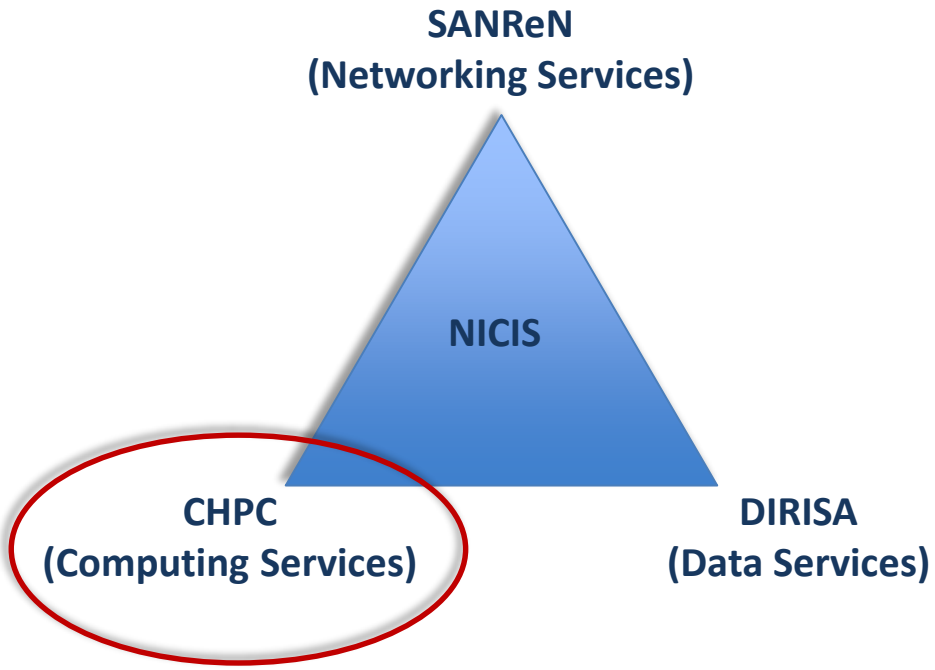
- 'An accomplished and preferred **partner** for HPC Solutions'

MISSION

- 'Provide world-class HPC that **enables** cutting-edge **research** with high impact on the South African Economy'

Background: NICIS

- **NICIS**: National Integrated Cyber Infrastructure System
- Framework **report**[1] issued to DST – made public in June 2014
- Contains **recommendations** for sustainable NICIS in SA



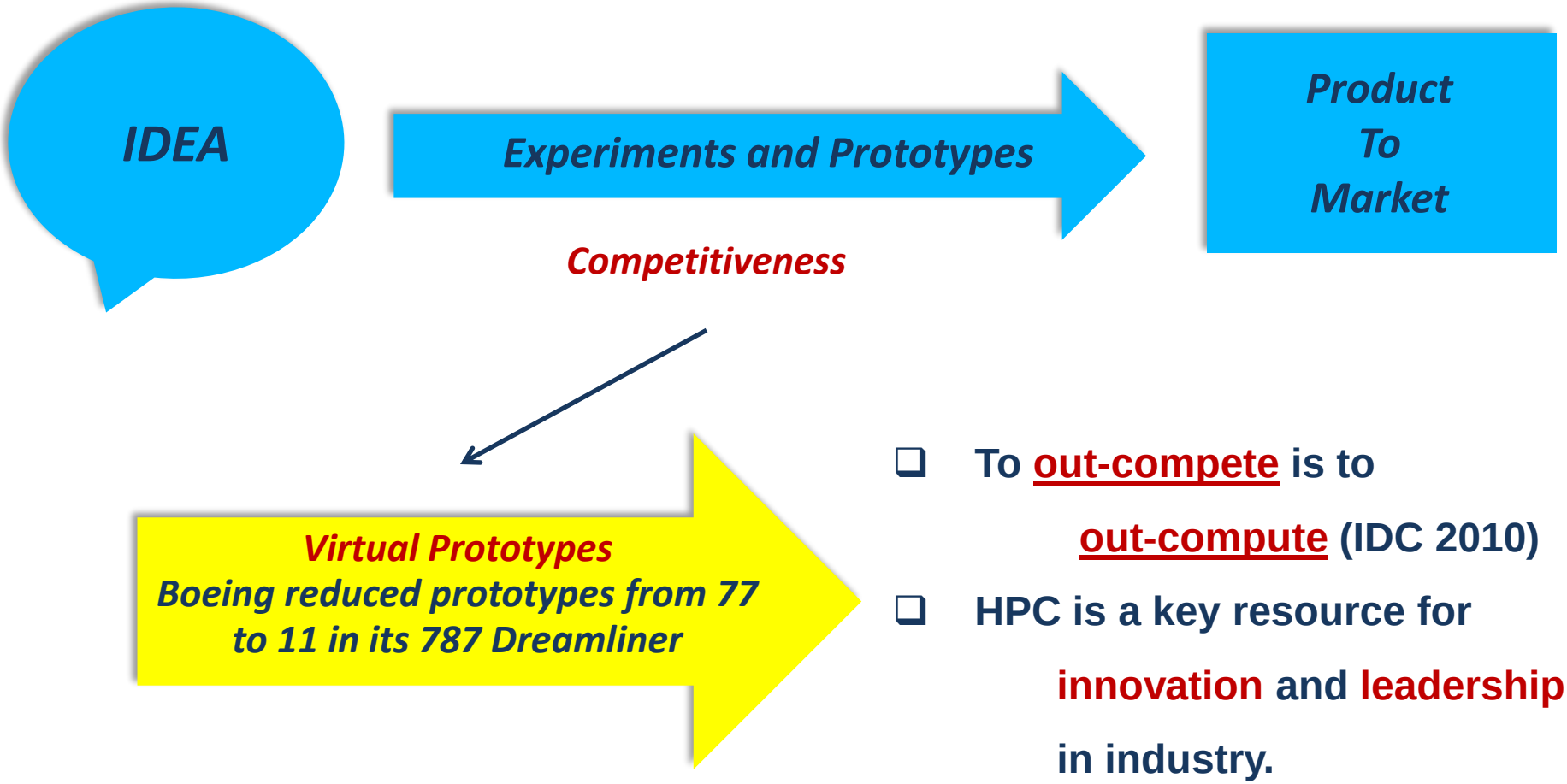
eResearch Africa Presentation:
NICIS
Prof Colin Wright
9:30 – 10:00
Tuesday, 25 November 2014

[1] http://www.dst.gov.za/index.php/resource-center/cyber-infrastructure2

Background: Rationale for Investment

- HPC is central to most modern **research-based developments**
- HPC globally defines **competitiveness** of countries
- South Africa has some **key industries** that drives its own **economic development**
- Established track record of success of **HPC impact** in other countries

Background: Why HPC?



- To out-compete is to out-compute (IDC 2010)
- HPC is a key resource for innovation and leadership in industry.

User Community

2013/2014 Financial Year:

- CHPC had **305** active users
- HPC systems utilisation: **91%**
- **>50 million core hours** with 368 281 jobs successfully completed
- Achieved with **83.8% uptime** (5.2% scheduled downtime)
- Centre **enabled 60** peer reviewed **publications** by users
- **51** completed **postgraduate degrees** utilising CHPC resources

User Community: Engagement Strategy

CENTRE FOR HIGH PERFORMANCE COMPUTING

Special Interest Groups (SIG Research Domains

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

Activities

Flagship Projects
(About 16 Projects. Total of about 60 students and postdoctoral fellows.)

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

Annual HPC Conference
(7 successful meetings with the support from international collaborators. Integration with Main Stream Conferences)

CHPC

Partners

ducation Institutions
ent Departments
ouncils
Agencies

l Collaborators
rs
ganisations

User Community: South Africa

CENTRE FOR HIGH PERFORMANCE COMPUTING

- CHPC
- Existing Users
- Future Users

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

CSIR
our future through science

User Community: Africa

- **SADC HPC Framework** developed and adopted by Ministerial Committee Meeting in Mozambique on the 19th June 2014
- HPC Research **white papers** in agriculture and health
- CHPC is providing HPC resources to **African students** (RSA collaborator required for access)
- **RANGER** Program ⇒ Botswana and Tanzania
- Student cluster challenge **to Africa** planned
- In collaboration with **SKA**, assist the **African partners**

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

CSIR
our future through science

User Community: Industry

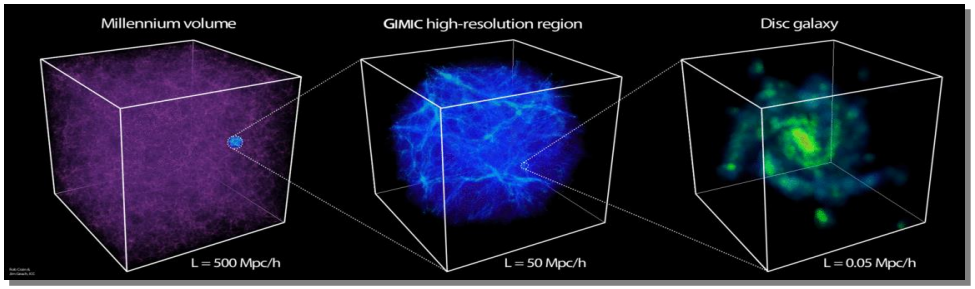
SASOL
De Beers
Johnson Matthey
ESKOM
GLENCORE XSTRATA
ECI-JV
NNR
TOTAL
MINTEK
HATCH

CHPC Industry Advisory Council Forum:

- Develop long term **partnership with industries** within the country.
- Provide **30%** of the **resources** for this program to avoid unfair competition with academia.
- **SLA's** signed based on unique **requirements** and needs of specific industrial partner.
- Initiated an **Industrial Advisory Council Forum** in 2011 to identify mutual **benefits** and **challenges** for different industries.

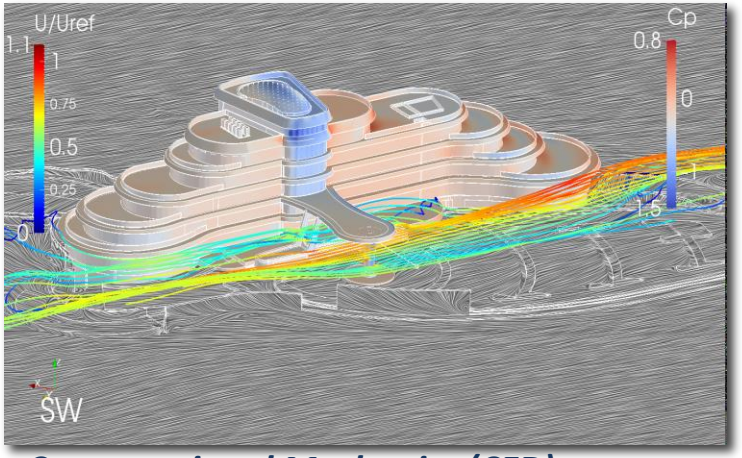
Application Research Areas

CHPC
CENTRE FOR HIGH PERFORMANCE COMPUTING



Astronomy/Cosmology (Universe/Galaxy Evolution)

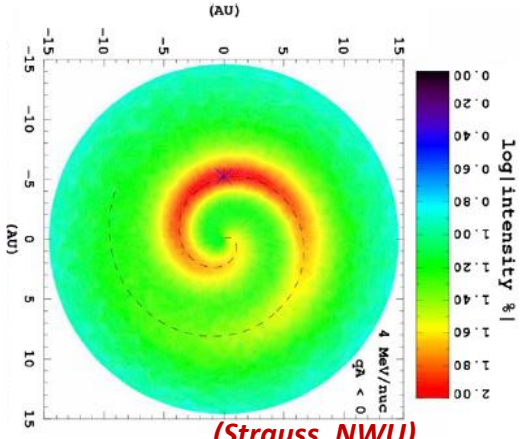
(Courtesy Cress and Cunnama)



Computational Mechanics (CFD)

(Example: Building aerodynamics)

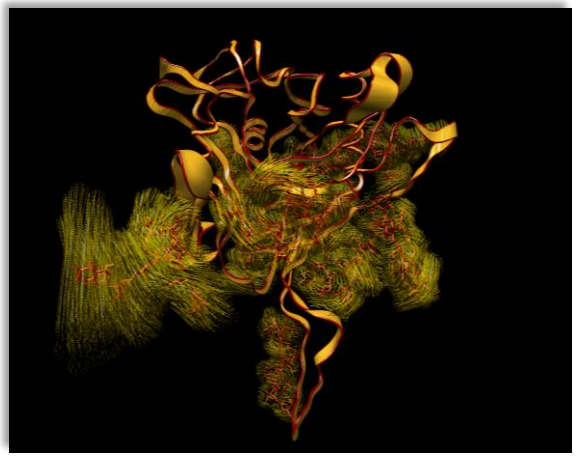
(ECI-JV)



(Strauss, NWU)

Computational Space and Astrophysics

Bioinformatics
(Example of simulated movement of HIV-1 gp120 protein and attached sugar chains)



(Image courtesy Natasha Wood, SANBI, UWC)

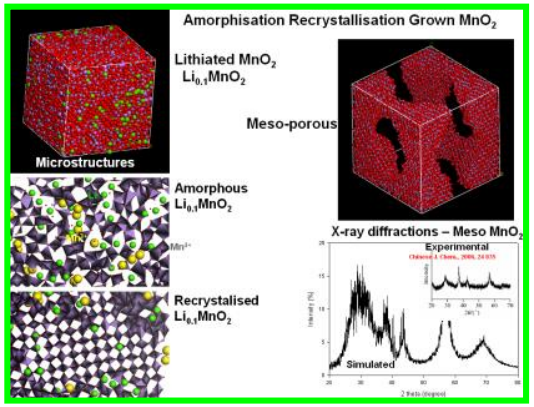
Applied Mathematics and Computational Methods

Application Research Areas

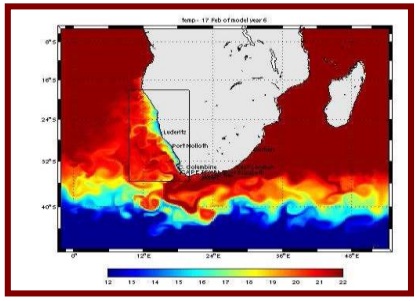
CHPC
CENTRE FOR HIGH PERFORMANCE COMPUTING

Advanced Computer Engineering (ACE)

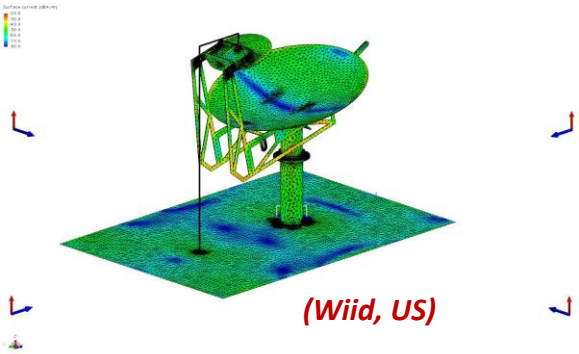
Bioinformatics Service Platform (BSP) – Successfully Implemented



Energy Security *(Ngoepe, UL)*
Mineral Beneficiation
Drug discovery
Materials Science

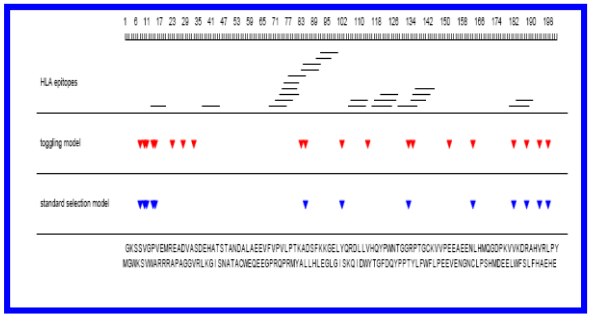


Climate Modelling and Weather Prediction.



(Wiid, US)

Meerkat/SKA Dish Design
Computational Mechanics for Astronomy



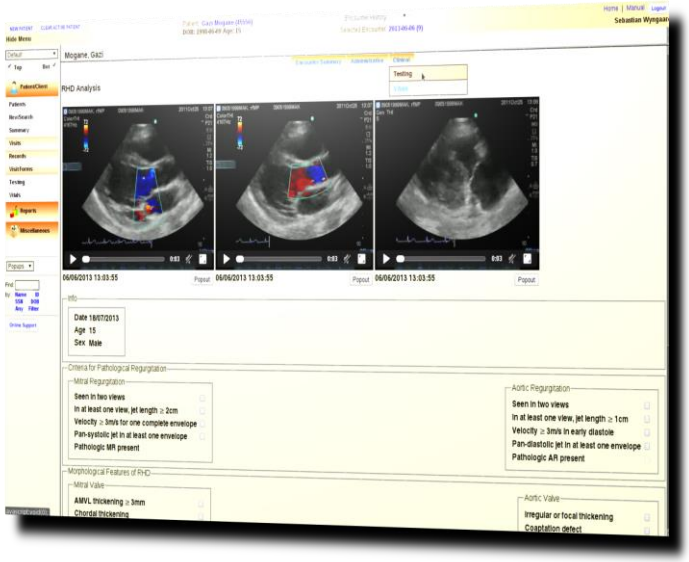
HIV Mutation
Brain Imaging
Cardio-Vascular

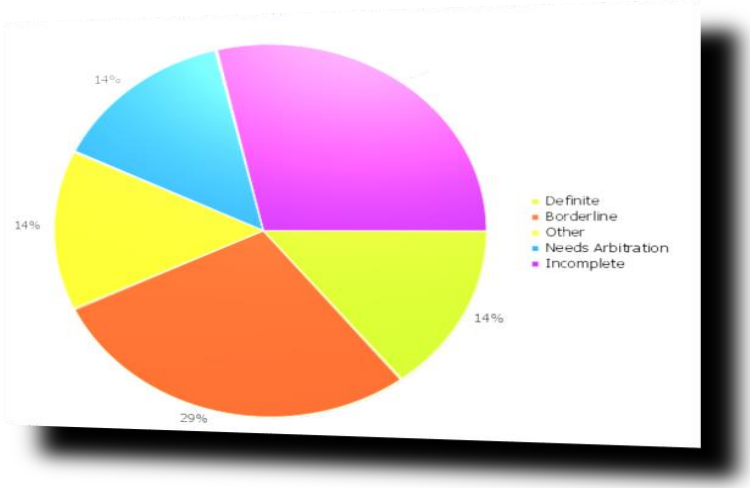
Visualation Analytics
Animation
Movie industry

Application Research Areas

Rheumatic Heart Disease Portal

CENTRE FOR HIGH PERFORMANCE COMPUTING





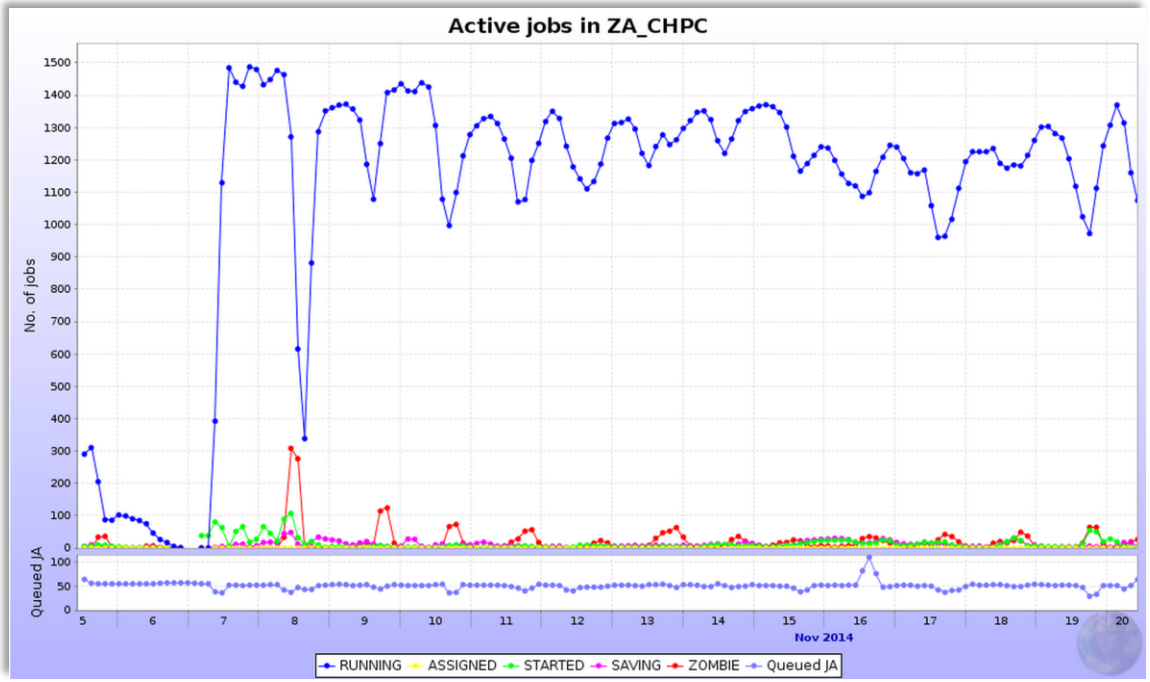
- **Medical practitioners** to analyse and diagnose echocardiograms for **RHD**
- **Tracks statistical data** – number of patients processed and catalogs diagnostic results.
- Currently **>2000 patients registered** of which 1800 processed.
- **Raw images** for the patients (DICOM) is contained in **200 000 files** and account for 4TB storage in DIRISA.
- Additional 3000 patients will be added – **Extend** this **portal** for the rest of SA

Application Research Areas

Grid Applications

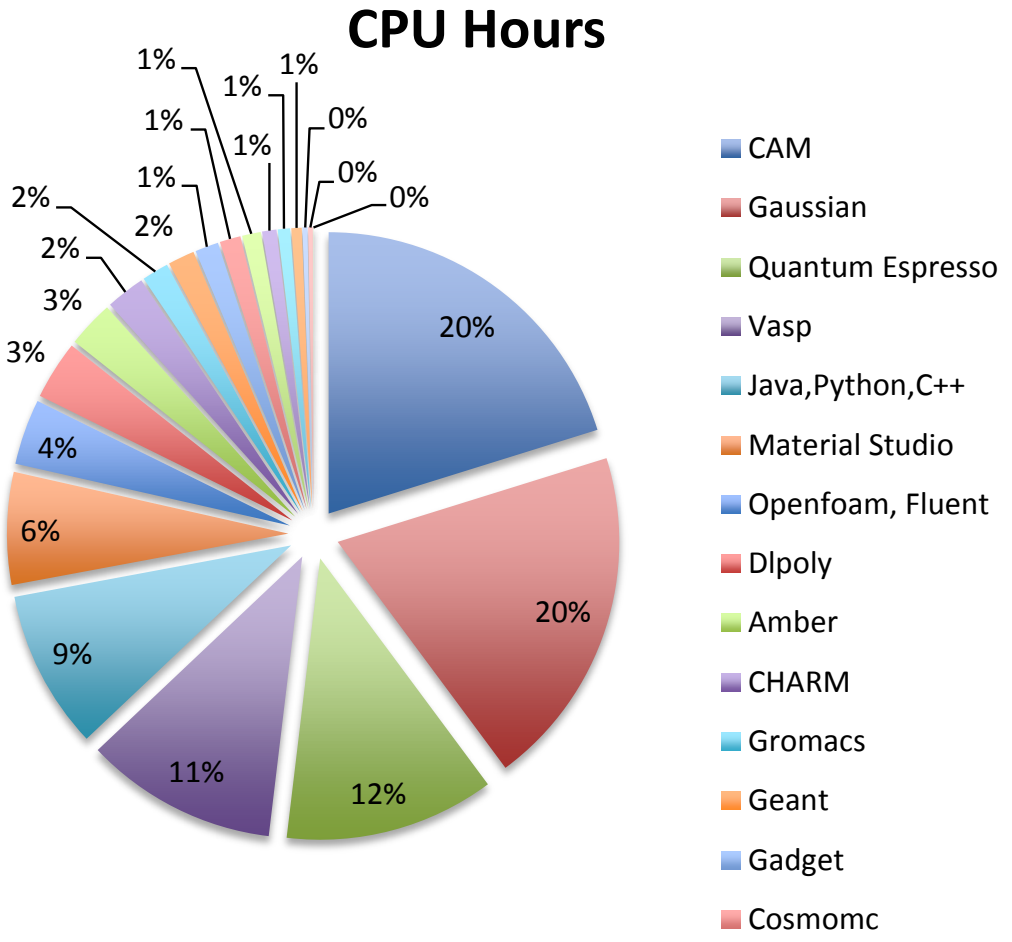
- PowerEdge C8220, Intel E5-Series, **1200 Processors** – Dedicated for **CERN**
- PowerEdge C8220, Intel E5-Series, **816 Processors** – **SAGrid**
- The current system still provides **500 jobs/day** for **ALICE** experiment.

- The New system will provide (+-) **2400 jobs/day** for both **ALICE & ATLAS**
- Storage Capacity of **800TB**



Graph courtesy of Sakhile Masoka (CHPC)

Software Applications Profile



- New Resource Manager (**PBSPro**) Jan 2014 and improved resource allocation
- **Accelrys license** 20% CPU utilisation.
- **NEMO code** for ocean modelling **most scalable** with over 1000 CPU's
- GROMACS, NAMD and OpenCV scale well on **GPU architecture**
- In collaboration with **NAG** the user developed codes, DFTB and FDTD, Amber and Quantum Espresso, were **ported in GPU** and scale well.

Training Initiatives

CHPC committed to HCD with various training initiatives with more than 80 different events to date

EXAMPLES:

- **HPC Winter School** in Parallel Programming

  6 Events (2009-2014); 273 postgrad students trained

- **CHPC Introductory Programming School**

  Python and Linux Skills; 4 Events (2011-2013); ~200 postgrads trained

- **Student Cluster Competition** – HPC skills development in country (Undergraduate)
- **SDE Workshop** – Collaboration with NWU
- Training related to the **RANGER** project - 10 system administrators trained
- Studentships, internships and post-doc positions at CHPC
- Special Courses

Training Initiatives: HCD

Student Cluster Competition:

Training students in HPC

- **CHPC Student Cluster Competition**
  - 8 teams of 4 students
  - Winning team entered into ISC Student Cluster Challenge
- **ISC Student Cluster Challenge**
  - Entered twice, won twice

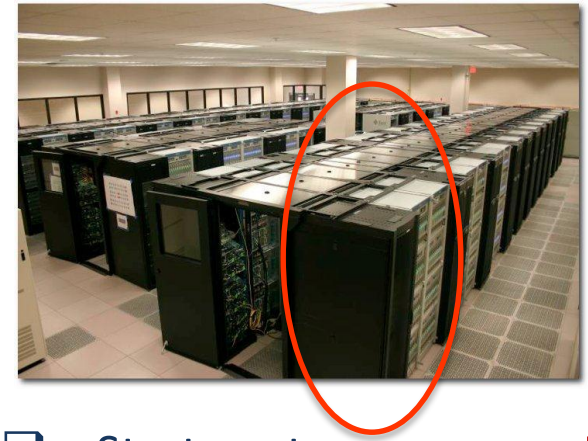
ISC'13 Champions (Leipzig)[1]

ISC'14 Champions (Leipzig)[2]

[1] www.hpcwire.com/2013/06/27/south_african_student_hpc_team_rides_gpus_to_victory_at_isc_2013

[2] www.hpcwire.com/2014/06/26/student-supercomputing-competition-embraces-gpu-power

HPC Ecosystem Initiative



HPC Systems Repurposing

- Strategy to **repurpose HPC systems** out of national production for **local processing** capabilities or **training** facilities
- **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/universities
  - ⇒ **Do not have** HPC facilities
  - ⇒ Need of HPC resources for **learning**

HPC Ecosystem Initiative

RANGER Project

- **RANGER** System at **TACC**
- **82 Racks** of compute hardware
- Each rack **4 Quad Core** AMD CPUs

- **20 Racks** donated to **CHPC** by TACC
- **4 pilot sites** in SA: UFH, UV, UKZN and Wits
- Supported by **DHET**
- **3 International Sites**: Tanzania, Botswana and Zambia
- **ACE Lab** developed a **cluster management suite** from Open Source tools for these systems

HPC lable to date

CHPC
CENTRE FOR HIGH PERFORMANCE COMPUTING

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

2007: IB 640 cores
2.5

90 TB GP File System

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

4 PB DIRISA Storage Unit

Acknowledgements

- Dr Happy Sithole
- Edward Rakate
- Dr Catherine Cress
- Dr Daniel Moeketsi
- Charles Crosby
- Andrew Gill
- Kevin Colville
- David Macleod
- Dane Kennedy
- Dr Anton Lopis
- Sakhile Masoka

- DST
- CSIR
- Meraka Institute
- CHPC User Base