GridPP

UK Computing for Particle Physics

UNIVERSITY of GLASGOW

UK Computing for HEP

ECFA

7th Nov 2014

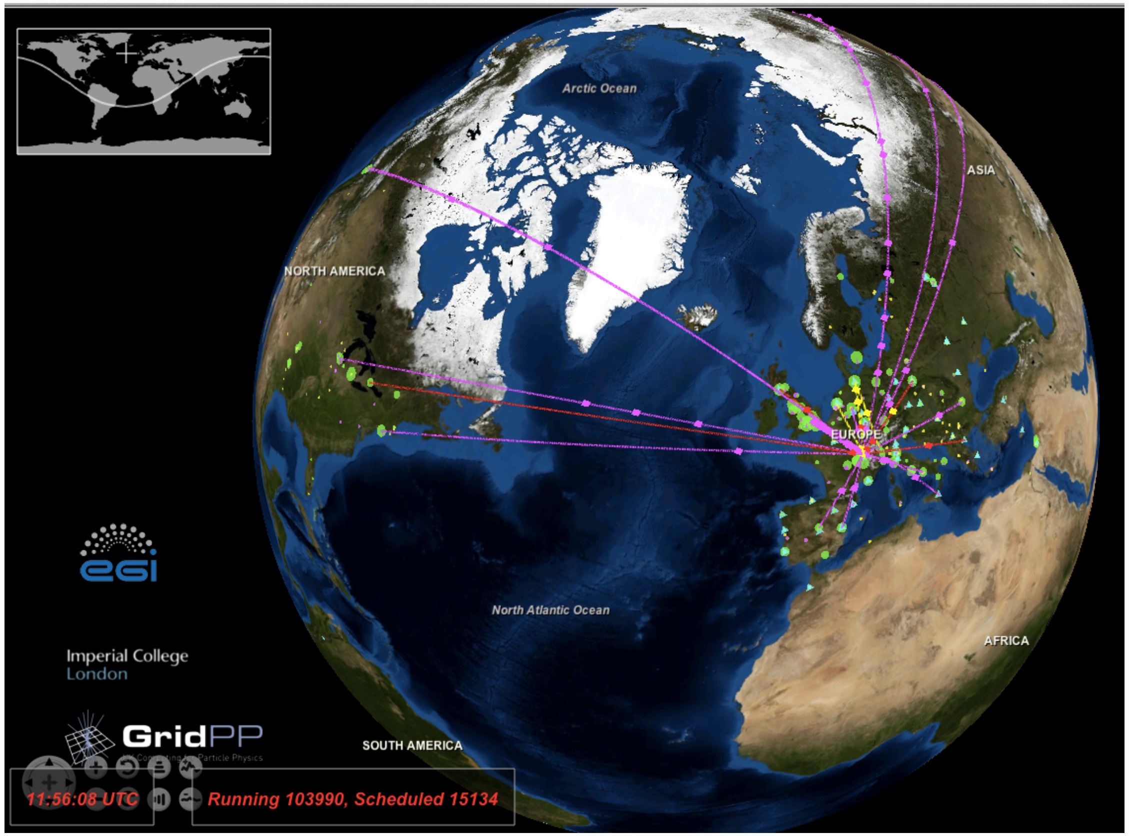
Prof. David Britton

GridPP Project leader

University of Glasgow

Outline

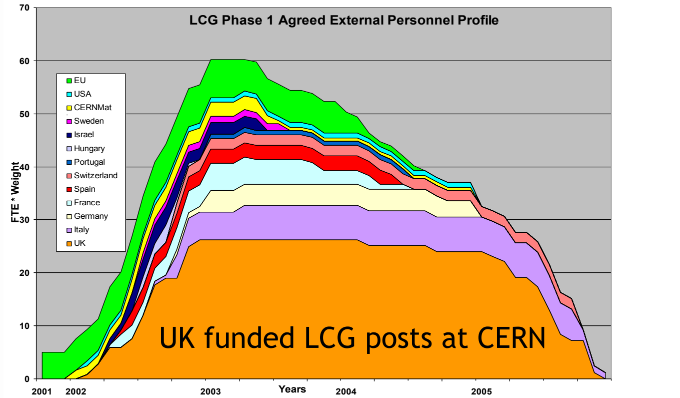
- GridPP:
  - Mission and Structure
  - Resources, Pledges, Utilisation
  - NGI services.
- Future:
  - Requirements
  - Challenges
- Broader Context:
  - UK EcoSystem
  - European Initiatives
- Summary



GridPP Mission

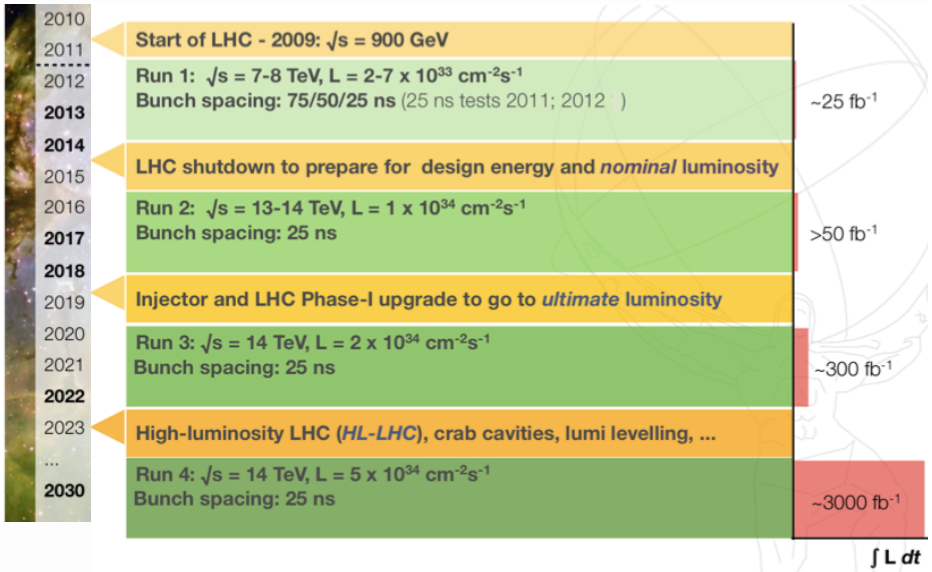
GridPP
UK Computing for Particle Physics

Mission: To deliver resources to the UK and worldwide particle physics community in accordance with the WLCG MOU, by means of a large-scale computing Grid in the UK.



The UK kick-started WLCG in 2002 with a £5.6m investment.

2001 GridPP1 – *From Web to Grid*
2004 GridPP2 – *From Prototype to Production*
2007 GridPP2+ *(6-month extension)*
2008 GridPP3 – *From Production to Exploitation*
2011 GridPP4 - *Computing in the LHC era*
2015 GridPP4+ *(One year extension)*
2016 GridPP5 – *Computing beyond the Higgs*

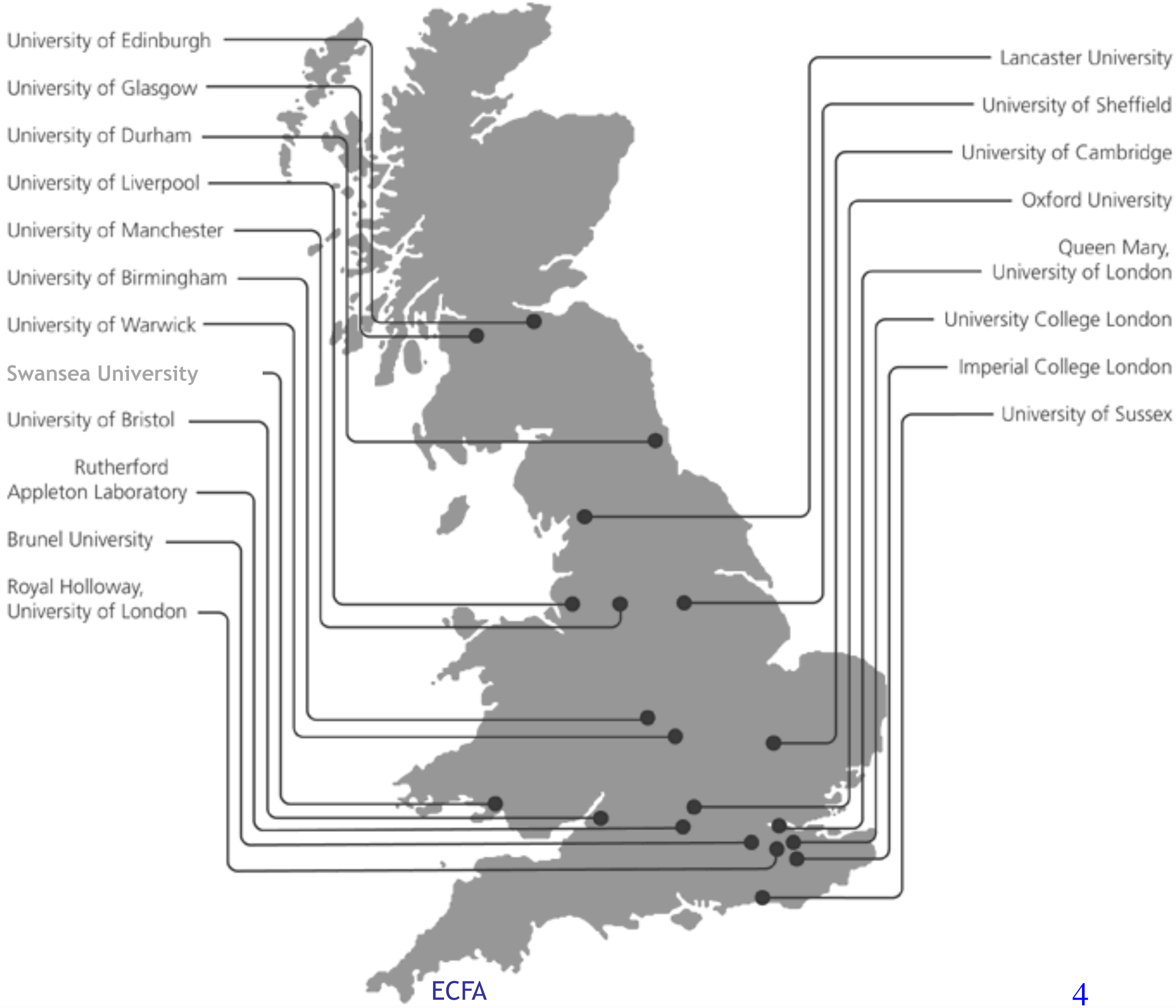


GridPP Sites

GridPP UK Computing for Particle Physics

GridPP is a collaboration of 19 UK universities, and the Rutherford-Appleton Laboratory (STFC).

It has built and operates a distributed computing Grid across the UK for particle physicists.



Janet6 Network

GridPP
UK Computing for Particle Physics

Janet6 rolled out 1 year ago: Optical 40/100G national academic network.

- Long term – 10+ years for fibre leases
- Transmission equipment refresh after ~7+ years
- Roadmap to next capacity increment per wavelength (400G ?)
- More flexible optical architecture
- Direct management of optical infrastructure
- Huge capability for additional capacity
- Continued integration of Regions

GridPP Resources

GridPP
UK Computing for Particle Physics

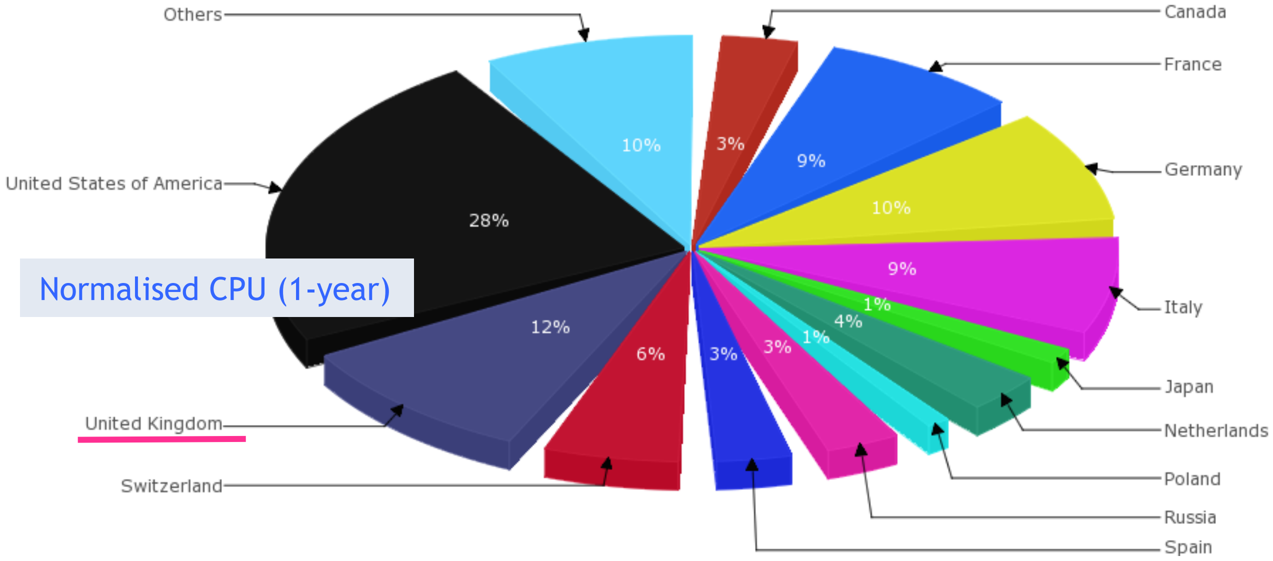
RAL Tier-1:

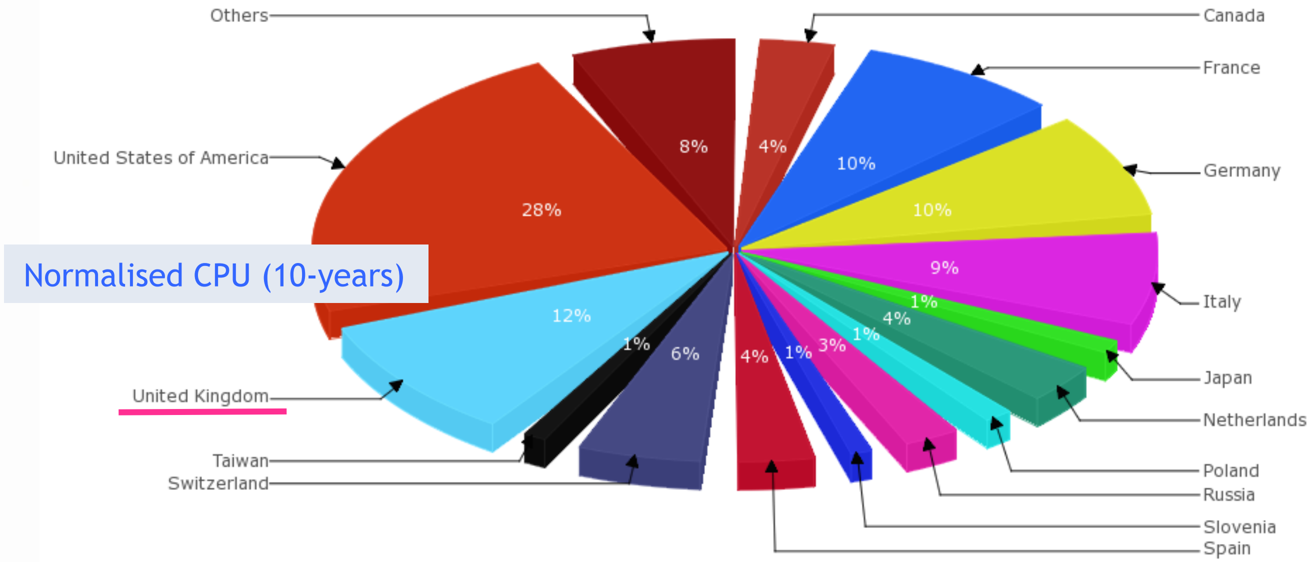
- 13,500 Logical CPU
- 10 PB Disk
- 13 PB Tape

Distributed Tier-2:

- 50,000 Logical CPU
- 20 PB Disk

Together, consistently delivered 12% of WLCG resources over the last decade.





Pledges and Delivery

UK pledges and delivers the UK authorship fraction of the (scrutinised and approved) Global LHC resource requirements.

- 10% of ATLAS (12.5% of Tier-1 country authors).
- 4% of CMS (7% of Tier-1 country authors).
- 19% of LHCb (31% of Tier-1 country authors).
- 2% of ALICE.

UK delivery exceeds all these number due to additional non-GridPP funded resources contributed by the participating institutes and due to highly-efficient operations.

atlas VO CPU Efficiency by REGION and DATE.
ALL VOs. November 2013 - October 2014.

The following table shows the distribution of CPU Efficiency grouped by REGION and DATE.

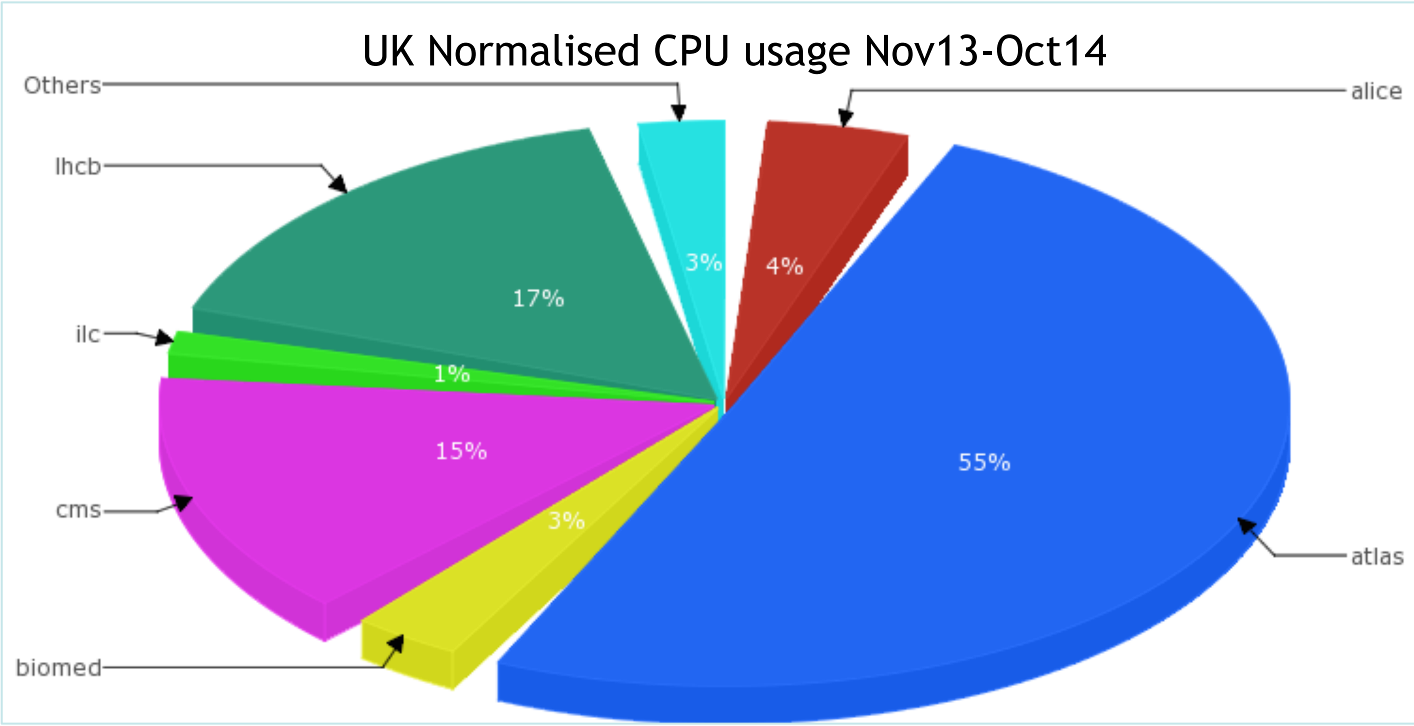
CPU Efficiency (%) by REGION and DATE													
REGION	Nov 13	Dec 13	Jan 14	Feb 14	Mar 14	Apr 14	May 14	Jun 14	Jul 14	Aug 14	Sep 14	Oct 14	Total
CA-TRIUMF	95.0	92.9	93.1	92.8	112.2	102.5	96.4	95.9	126.8	117.7	105.2	164.1	104.4
CH-CERN	93.3	89.3	89.6	72.9	28.3	88.4	87.5	92.3	92.7	94.1	92.7	77.0	81.7
DE-KIT	98.2	92.4	102.0	100.8	114.2	103.1	106.4	103.8	143.6	139.0	129.4	176.5	110.5
ES-PIC	95.8	94.9	95.9	94.4	95.1	96.9	95.5	95.8	153.7	121.3	130.8	134.1	100.5
FR-CCIN2P3	95.7	145.1	102.9	101.3	110.8	107.5	95.1	94.9	120.0	118.2	116.9	435.4	109.4
IT-INFN-CNAF	95.1	89.9	86.1	86.2	91.4	93.7	88.7	94.6	94.3	99.2	99.9		91.4
NDGF	61.0	58.1	67.3	40.8	72.2	83.8	91.1	73.1	84.2	77.8	88.1	83.9	73.6
NL-T1	92.2	96.0	96.7	94.3	97.8	99.6	95.7	94.7	100.3	102.7	106.4	112.0	98.2
NRC-KI-T1	96.7	90.8	92.2	98.2	96.0	91.9	89.5	89.7	93.3	91.3	90.2	82.9	91.9
TW-ASGC	82.8	94.1	111.7	114.9	137.8	163.9	97.6	94.8	94.6	91.4	95.2	89.3	106.2
UK-T1-RAL	97.1	93.5	119.1	123.0	123.4	106.0	91.3	93.5	130.0	125.1	107.9	168.2	111.3
US-T1-BNL	90.5	81.3	73.4	88.5	86.1	81.5	89.1	92.2	89.4	81.2	82.2	81.3	84.8
Total	**93.2**	**92.4**	**93.4**	**93.9**	**91.0**	**96.2**	**92.2**	**93.9**	**109.0**	**105.4**	**100.2**	**117.7**	**96.8**

Click here for a CSV dump of this table
Click here for a Extended CSV dump of this table
Click here for XML encoded data

Key: 0% <= eff < 50%; 50% <= eff < 60%; 60% <= eff < 75%; 75% <= eff < 90%; 90% <= eff < 100%; eff >= 100% (parallel jobs)

VOs Supported

- camont
- cernatschool.org
- comet.j-parc.jp
- constellation.stfc.ac.uk
- earthsci.vo.gridpp.ac.uk
- epic.vo.gridpp.ac.uk
- eresearchsouth.ac.uk
- gridpp
- hyperk.org
- ltwo
- mapper-project.eu
- mice
- mott2.org
- na62.vo.gridpp.ac.uk
- neiss.org.uk
- neurogrid.incf.org
- nsccs.ac.uk
- nw-grid.ac.uk
- oxgrid.ox.ac.uk
- pheno
- ralpp
- scarf.rl.ac.uk
- skatelescope.org
- snoplus.snolab.ca
- t2k.org
- ukmhd.ac.uk
- ukqcd.vo.gridpp.ac.uk
- vo.helio-vo.eu
- vo.landslides.mossaic.org
- vo.londongrid.ac.uk
- vo.northgrid.ac.uk
- vo.scotgrid.ac.uk
- vo.southgrid.ac.uk
- westfocus.westminster.ac.uk
- wrgrid.org.uk



Beyond the LHC

Grid Production Tools

What was learned

- found one hole in the coverage for low momentum muons from pion decays, deflected at large angle by the spectrometer magnet. A small muon detector was added outside the previous acceptance region
- found one hole in the coverage for high momentum pions slightly escaping the beam pipe downstream the hadron calorimeters and muon detector. A small hadron calorimeter was added just outside the beam pipe
- the digitisation and reconstruction software was improved after training on the previous MC production and comparison with test run data
- background studies have been refined trigger studies have been started
- some weaknesses in the MC has been identified and corrected, both from the software and the physics point of view

Two serious gaps in the detector acceptance that would have compromised the measurement

NA62 gLite-based Grid Production System

Up and running [illegible] 5 UK sites and one Bulgaria[illegible] production/simulations round [illegible] decay channels, 22,675 runs, 19,469 files produced, 200,000 cumulated CPU hours, 29 TB of data on CASTOR.

Data produced was used for

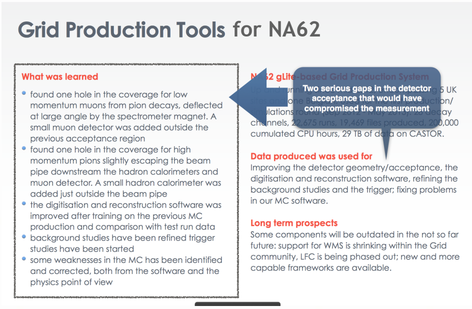
Improving the detector geometry/acceptance, the digitisation and reconstruction software, refining the background studies and the trigger; fixing problems in our MC software.

Long term prospects

Some components will be outdated in the not so far future: support for WMS is shrinking within the Grid community, LFC is being phased out; new and more capable frameworks are available.

Beyond the LHC

GridPP
UK Computing for Particle Physics



Grid Production Tools for NA62

What was learned

- found one hole in the coverage for low momentum muons from pion decays, deflected at large angle by the spectrometer magnet. A small muon detector was added outside the previous acceptance region
- found one hole in the coverage for high momentum pions slightly escaping the beam pipe downstream the hadron calorimeters and muon detector. A small hadron calorimeter was added just outside the beam pipe
- the digitisation and reconstruction software was improved after training on the previous MC production and comparison with test run data
- background studies have been refined trigger studies have been started
- some weaknesses in the MC has been identified and corrected, both from the software and the physics point of view

Two serious gaps in the detector acceptance that would have compromised the measurement

cumulated CPU hours, 29 TB of data on CASTOR.

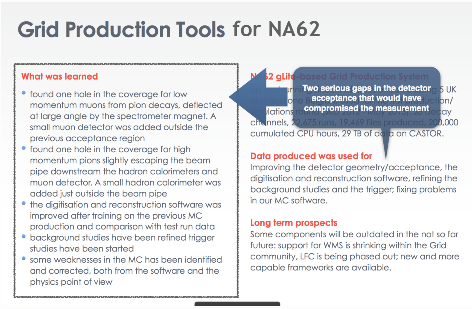
Data produced was used for

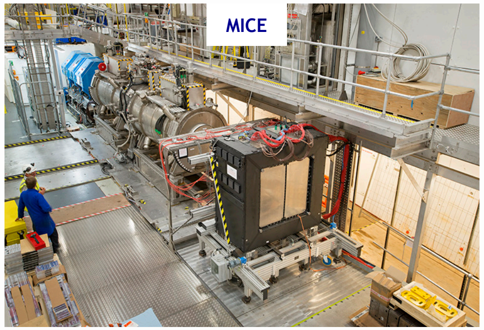
Improving the detector geometry/acceptance, the digitisation and reconstruction software, refining the background studies and the trigger; fixing problems in our MC software.

Long term prospects

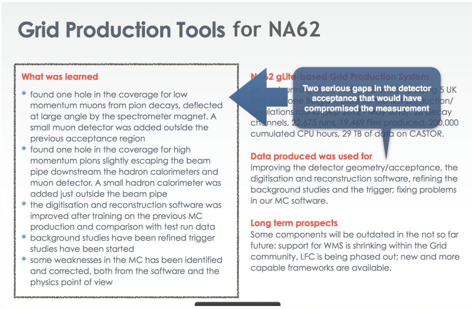
Some components will be outdated in the not so far future: support for WMS is shrinking within the Grid community, LFC is being phased out; new and more capable frameworks are available.

Beyond the LHC





Beyond the LHC


Grid Production Tools for NA62

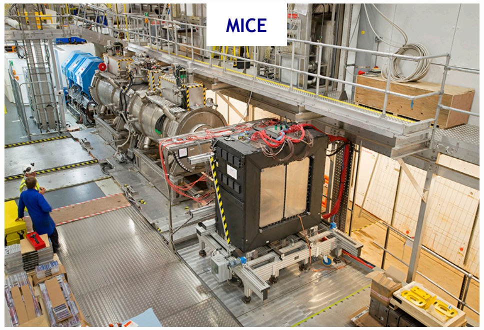
What was learned
- found one hole in the coverage for low momentum muons from pion decays, deflected at large angle by the spectrometer magnet. A small muon detector was added outside the previous acceptance region
- found one hole in the coverage for high momentum pions slightly escaping the beam pipe downstream the hadron calorimeters and muon detector. A small hadron calorimeter was added just outside the beam pipe
- the digitisation and reconstruction software was improved after training on the previous MC production and comparison with test run data
- background studies have been refined trigger studies have been started
- some weaknesses in the MC has been identified and corrected, both from the software and the physics point of view

Two serious gaps in the detector acceptance that would have compromised the measurement

cumulated CPU hours, 29 TB of data on CASTOR.

Data produced was used for
Improving the detector geometry/acceptance, the digitisation and reconstruction software, refining the background studies and the trigger; fixing problems in our MC software.

Long term prospects
Some components will be outdated in the not so far future: support for WMS is shrinking within the Grid community, LFC is being phased out; new and more capable frameworks are available.




BBC | Sign in
NEWS SCIENCE & ENVIRONMENT
Home | World | UK | England | N. Ireland | Scotland | Wales | Business | Politics | Health | Education | Sci/Environment

30 May 2012 Last updated at 18:04

Flavour changer: Genome could enhance tomato taste

By Matt McGrath
Science reporter, BBC World Service

The successful sequencing of the tomato genome will lead to tastier varieties within five years say scientists.

They believe that the elusive flavour of home grown tomatoes will by then be widely available in supermarkets.

Writing in the journal Nature, the researchers say the genetic information could reduce the need for pesticides.

Can the genome sequence pave the way for a return of the home grown taste

The authors believe the genome will also boost conventional breeding techniques over genetic modification.

Global Contributions

Global Services used by WLCG and EGI:

- Operates/maintains the GOCDB (semi-static database of site capabilities).
- Operates/maintains the APEL database (dynamic database of usage accounting information).

International Leadership:

- Leadership role in security policy development and international coordination.
- Leadership of the EGI CSIRT (operations security group in Europe).
- Leadership of WLCG security vulnerabilities group.
- Leadership of the IPv6 HEPiX working group.
- Co-chair WLCG Operations coordination.
- Chair many working groups (e.g. Cloud Traceability; Data management; wn+SL6 migration; perfSONAR; IPV6 deployment).

GridPP
UK Computing for Particle Physics

Current Funding Status

- GridPP5 proposal submitted in early 2014 for a 4 year project; could not be funded at that time.
- One year extension granted to GridPP4 for 2015-2016.
- Procurement is well underway that will deliver the April 2015 hardware pledges.
- The GridPP4 extension and will enable procurement for hardware to meet the April 2016 pledges.
- The GridPP5 proposal needs to be revised/re-submitted in early 2015 to cover funding of hardware and manpower over the period April 2016 – 2019.

STFC Programme Proposal

The UK Grid for Particle Physics Collaboration

GridPP
UK Computing for Particle Physics

University of Birmingham
University of Bristol
Brunel University
CERN, European Organisation for Nuclear Research
University of Cambridge
University of Durham
University of Edinburgh
University of Glasgow
Imperial College London
Lancaster University
University of Liverpool
University of Manchester
University of Oxford
Queen Mary, University of London
Royal Holloway, University of London
Rutherford Appleton Laboratory
University of Sheffield
University of Sussex
University of Warwick
Swansea University
University College London

Contacts

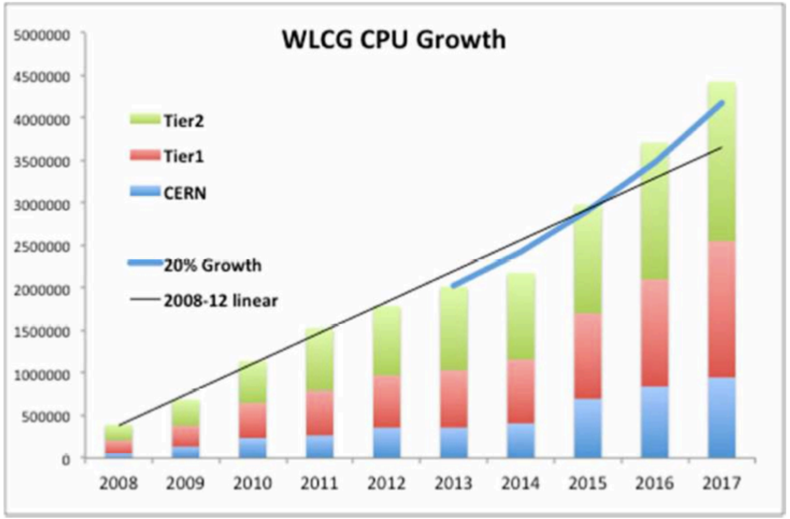
Prof. David Britton –David.Britton@glasgow.ac.uk (Project Leader)
Prof. Steve Lloyd – S.L.Lloyd@qmul.ac.uk (Collaboration Board Chairman)
Mr. Peter Gronbech – p.gronbech1@physics.ox.ac.uk (Project Manager)

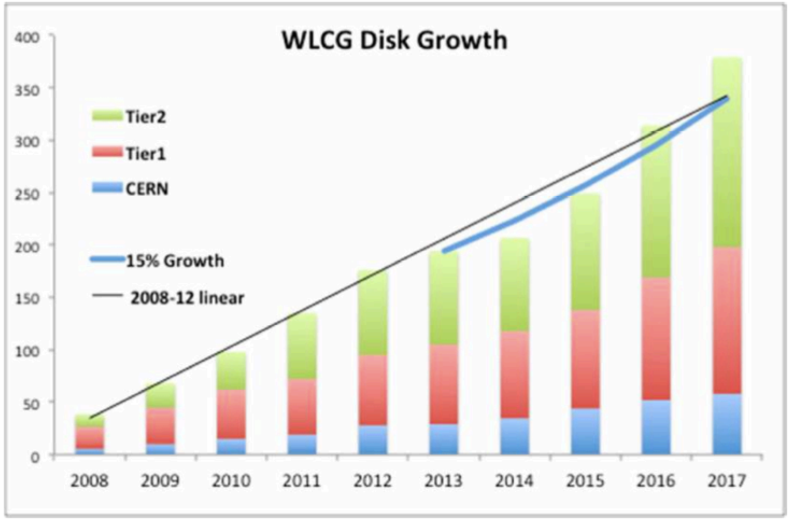
Abstract

This document contains input from GridPP to the Strategic Review Panel. We respond to the request for "Flat-Cash" and "50% of Flat-Cash" scenarios to cover the provision of Grid Computing for the LHC experiments and other groups, for the period April 2015 to March 2019. This period covers the LHC Run-2 where an increase in computing resources of a factor of 2.5x is required to exploit the last two decades of investment in the LHC programme.

Future Requirements

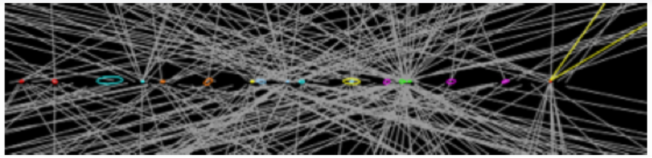
- Experiments resource planning is shaped by flat-cash constraints.
- Computing models have evolved significantly for Run-2 based on experience:
  - Increased use of network.
  - Decrease in number of copies of data.
  - Better design/integration reduces CPU and intermediate data requirements.

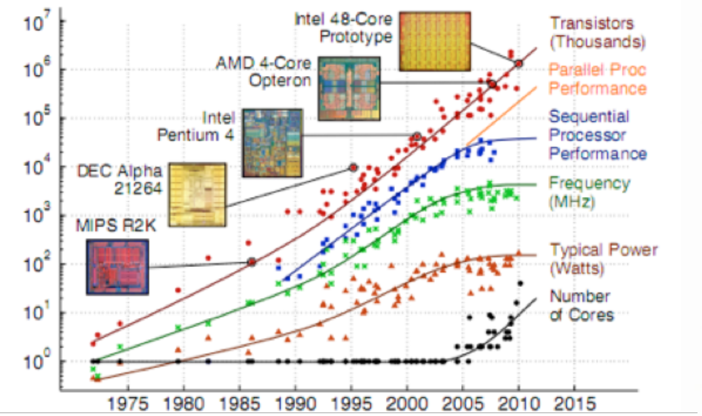


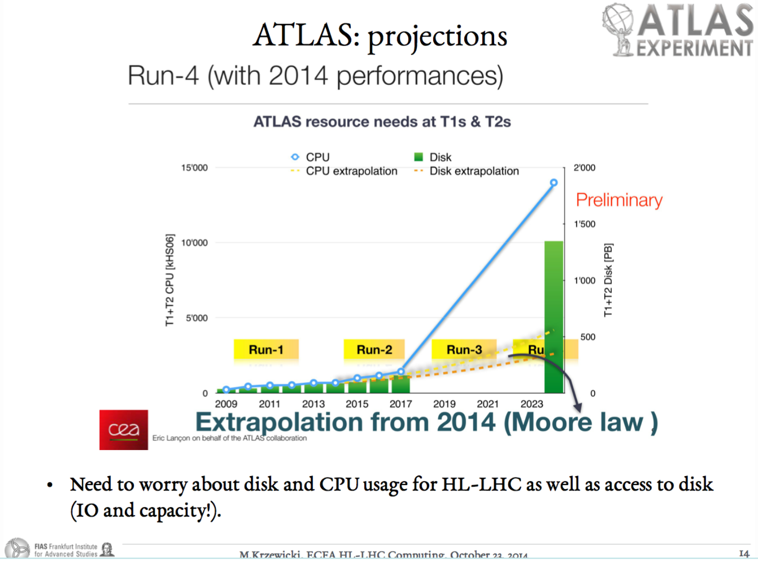


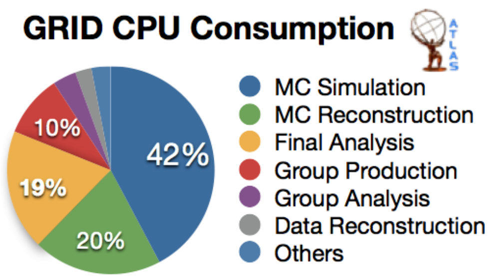
Future Challenges

- Run-2 data is expected to increase in both volume (x2-3) and complexity (PU → 50);
- Increasingly difficult to make use of the technical advances that are maintaining Moore's Law growth.
- Increasingly hard to find big gains in the computing models (copies of data are now minimal; efficiencies are high; etc.)
- Preparations for beyond Run-2 must start.
- Millions of lines of legacy code need to be parallelized; Memory-use and disk-access needs optimisation; Monte-Carlo simulation strategies need to make more flexible use of fast simulation; etc.



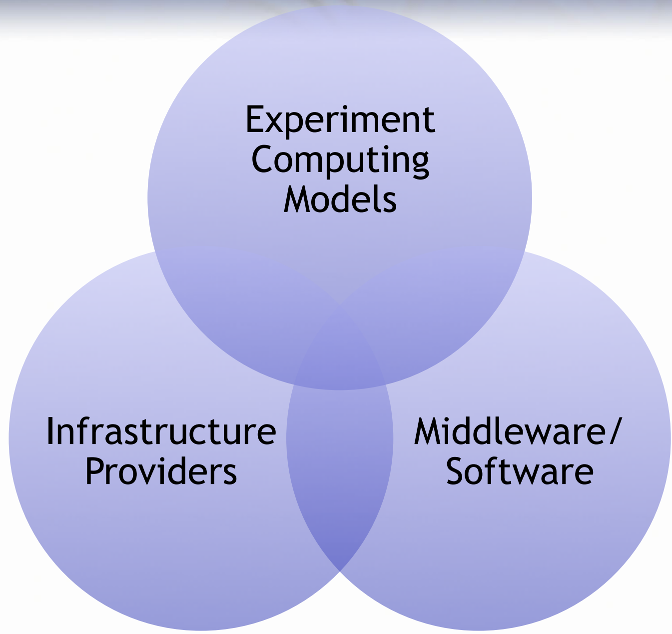






Addressing Challenges

- Require a continued close partnership between those developing the Experiment computing models, the middleware/software developers, and the infrastructure developers/operators.
- Long-term planning and large-scale testing are required.
- Manpower and hardware resources are required above those needed to just to operate the service.



The UK is active in:

- Preparing for IPV6.
- Security policy development.
- Testing large-scale deployment of CEPH as a successor to CASTOR for disk storage.
- Support and development of the DPM project for Tier-2 disk management.
- Cloud developments – leading CMS developments; major roles on ATLAS and LHCb.
- Development of operational security models for Cloud technology.
- Leading role in FIM4R activity towards use of federated identities.
- VAC and VCycle – simplifying Grid sites using cloud technologies.

The UK Ecosystem

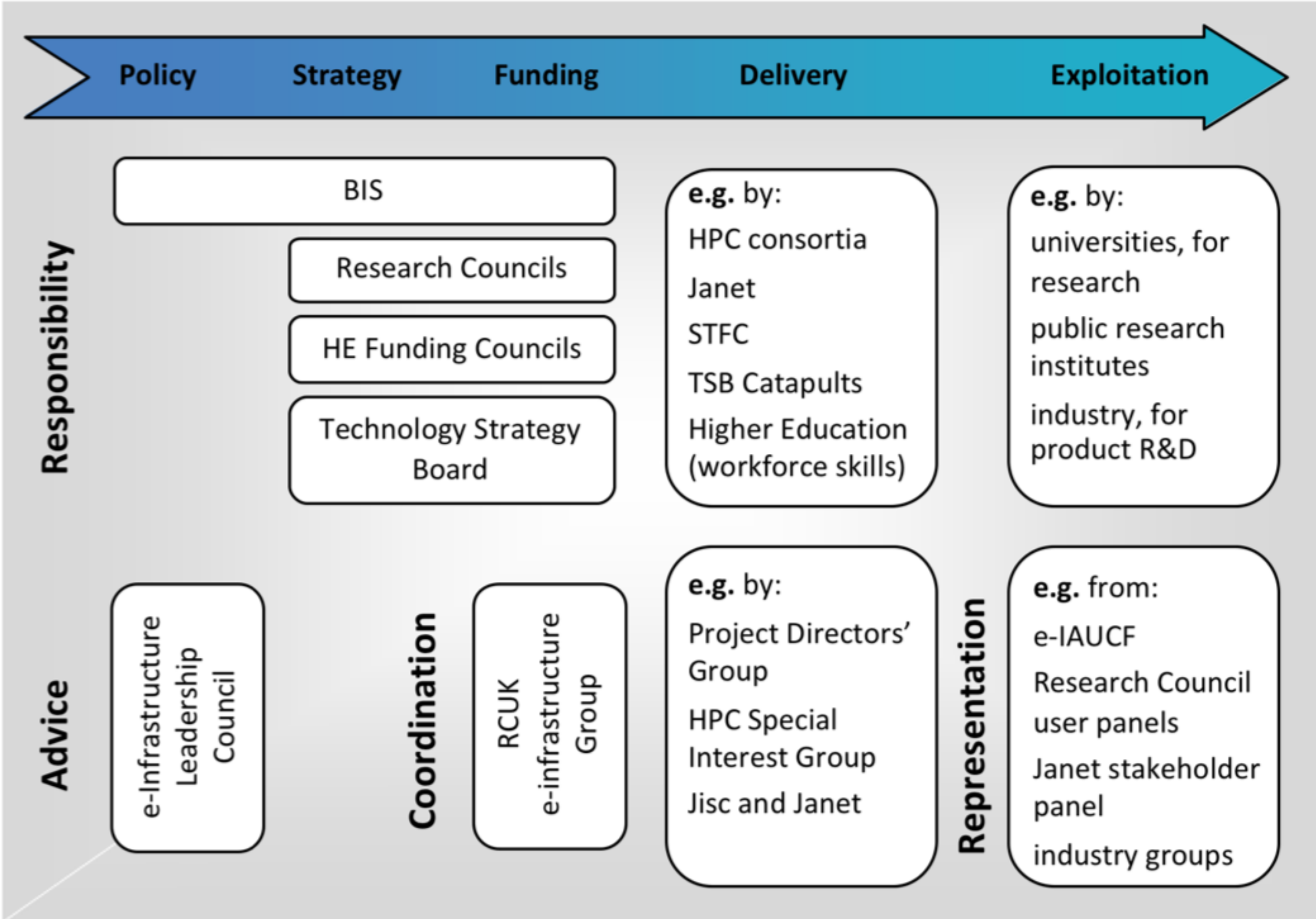
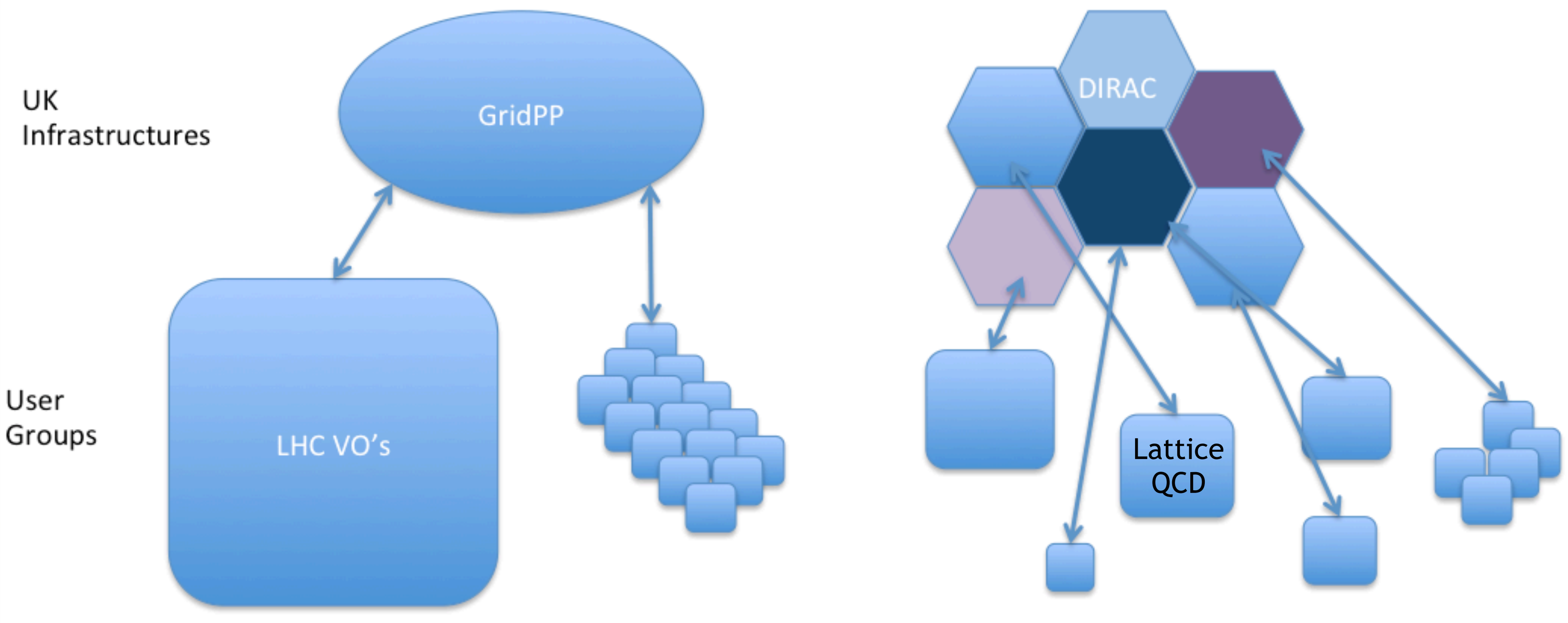


Figure 8: The e-infrastructure pipeline

DIRAC – Sister Project



Highly tuned bespoke infrastructures for large projects designed for optimal performance.

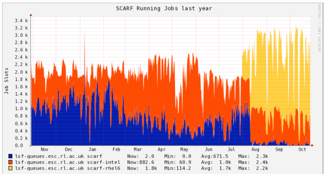
versus

Reducing barriers to entry; the long-tail of science; sharing resources for efficiency/flexibility

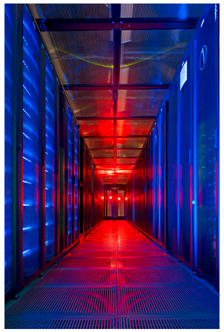
GridPP
UK Computing for Particle Physics

Scientific Computing Department

The GridPP Tier-1 for the LHC



SCARF - STFC local HPC cluster supporting STFC scientists, collaborators and facility user from SC, ISIS, CLF and others including RAL Space and Diamond



The JASMIN "super-data-cluster"

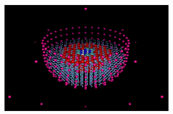
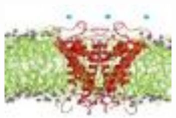
Bringing the compute to the data for all NERC sciences.

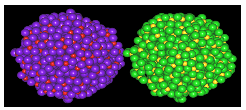
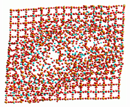
Major funded activities

- 180 staff supporting over 7500 users
- Applications development and support
- Compute and data facilities and services
- Research: over 100 publications p.a.
- Deliver over 3500 training days p.a.
- Systems administration, data services, high-performance computing, numerical analysis & software engineering.
- Expertise across the length scales from processes occurring inside atoms to environmental modelling

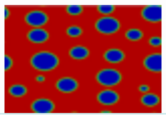
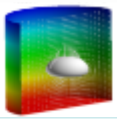
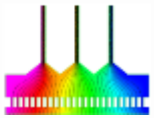
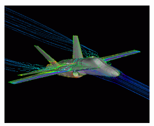
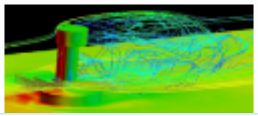
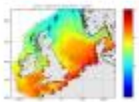
Supporting Data management for STFC facilities

Applications division supporting the CCP program


National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

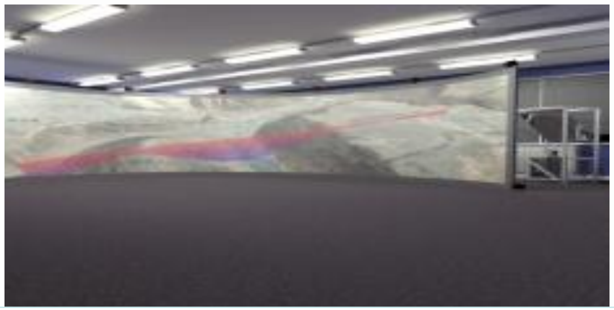
GridPP
UK Computing for Particle Physics

Hartree Centre

Projects and codes developed on state of the art systems:

- BlueGene/Q - Fastest UK machine and world's largest software development platform.
- Over 5 PB disc and 15 PB tape stores.
- iDataplex cluster.
- Data Intensive systems.
- Visualisation System.
- Energy Efficient Computing program.

- Engineering & Manufacturing
  - Vehicle Design & Testing
  - Consumer Electronics Design
  - Consumer Packaged Goods Design
- Environment
  - Weather modelling
- Life Sciences
  - Genomics for better crop yields
- Energy
  - Advanced Battery Cell Design
  - Efficient Well Head Oil extraction
- Financial Services
  - Risk Management
  - Service Modelling



European Context

Our community are involved in many of the new European initiatives:

- EGI Engage (H2020)
- VLData (H2020)
- Zephyr (H2020)
- HEP Software Foundation
- EU-T0 Initiative
- …and more.

GridPP
UK Computing for Particle Physics

Summary

The UK has a long history as a major, and reliable, partner in WLCG supporting all four LHC VOs and over 30 additional groups.

In addition to resources, the UK provides critical services used by the worldwide community and leadership in many areas.

Funding is established until March 2016 but the GridPP5 proposal must be resubmitted by March to cover the remainder of Run-2.

Flat-cash funding is extremely challenging, with the volume and complexity of data increasing against a backdrop of technical evolution away from simple architectures and the need to prepare for beyond Run-2.

Important efforts are underway to reduce the barrier-to-entry for Grids and to integrate Grid resources with other infrastructures.

The UK community is contributing to many pan-European and Global initiatives.

Backup

- Backups

Zephyr

PIC
port d'informació científica

ZEPHYR approach

- Reports on requirements for 2020+ projects exist (HL-LHC, SKA, CTA, Euclid, etc.)
- Organize work using a loose "architecture"
  - API (what do I call to access my data)
  - Control Plane or Storage Virtualization (mediates between APIs and physical data centers)
  - Local Data Center (how resources plug in)
- Transversal items
  - Data Preservation
  - New network features
  - Security and confidentiality

GridPP
UK Computing for Particle Physics

VLData

DIRAC THE INTERWARE

Common

- VLDATA H2020: EINFRA-1 (conce

"The mission of this project is to p
solution for Research Infrastructur
curation, analysis and publicat
primary and scientific data on ex
of RIs in Europe and beyond expec
data samples increases exponentia
the Exa scale. Existing solutions
VLDATA platform for largescal
Transparent integration of all typ
Cloud, Volunteer, HPC, etc., funde
opex, and coming from the public
Various RIs **from different scienti**
sciences or **chemistry** will be va
their concrete use case, achievin
optimization in the cost and
challenge will be ignored. **The co**
be addressed, as well as inte
scientific domains."

- 17 Project

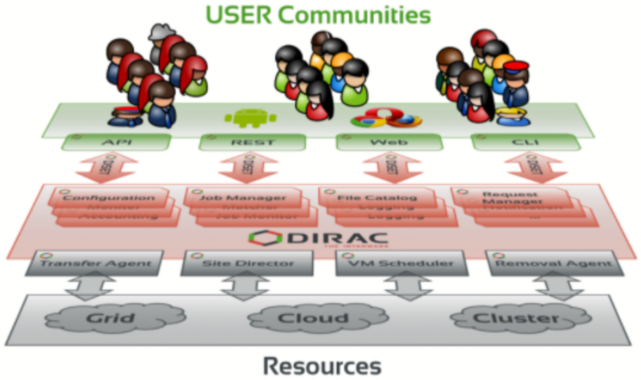
DIRAC
THE INTERWARE

Home News Support Installations DIRAC Publications Corporate Image Search

News

- DIRAC v6r11 release
  11 May 2014
- 4th DIRAC User Workshop registration
  9 Apr 2014
- DIRAC Project Manager Assistant
  8 Apr 2014
- DIRAC Consortium Agreement Signed
  1 Mar 2014
- Current DIRAC production releases
  10 Feb 2014
- DIRAC release v6r10
  20 Jan 2014
- Registration opened for the 4th DIRAC User Workshop
  15 Jan 2014
- The 4th DIRAC User Workshop at CERN
  11 Dec 2013
- New DIRAC Web Portal
  23 Oct 2013
- Current production DIRAC releases
  v6r7p30, v6r8p14

DIRAC (**D**istributed **I**nfrastructure with **R**emote **A**gent **C**ontrol) **INTERWARE** is a software framework for distributed computing providing a complete solution to one (or more) **user community** requiring access to **distributed resources**. DIRAC builds a layer between the users and the resources offering a common **interface** to a number of heterogeneous providers, **integrating** them in a seamless manner, providing **interoperability**, at the same time as an optimized, transparent and reliable usage of the resources.



The Workload Management System with Pilot Jobs introduced by the DIRAC project is now widely used in various grid infrastructures. This concept allows to aggregate in a single system computing resources of different source & nature, such as computational grids, clouds or clusters, transparently for the end users.

Delivery

VO	T1 Pledged (Delivered)	T2 Pledged (Delivered)
ALICE	2% (4.1%)	2% (4.7%)
ATLAS	10% (10.7%)	12.5% (16.1%)
CMS	8% (8.5%)	5% (7.6%)
LHCb	19% (16.7%)	31% (37.7%)

Alice – opportunistic use
ATLAS – T1 >100% efficiency and T2 leverage
CMS – Threshold T1 and T2 leverage
LHCb – UK led the way in prototyping Tier2ds which moved T1 workflows to T2s