GridPP

UK Computing for Particle Physics

ALICE, Others and the other Others

GridPP31 – Imperial College
24th September 2011

j.coles@rl.ac.uk

Jeremy Coles

ALICE

T2K

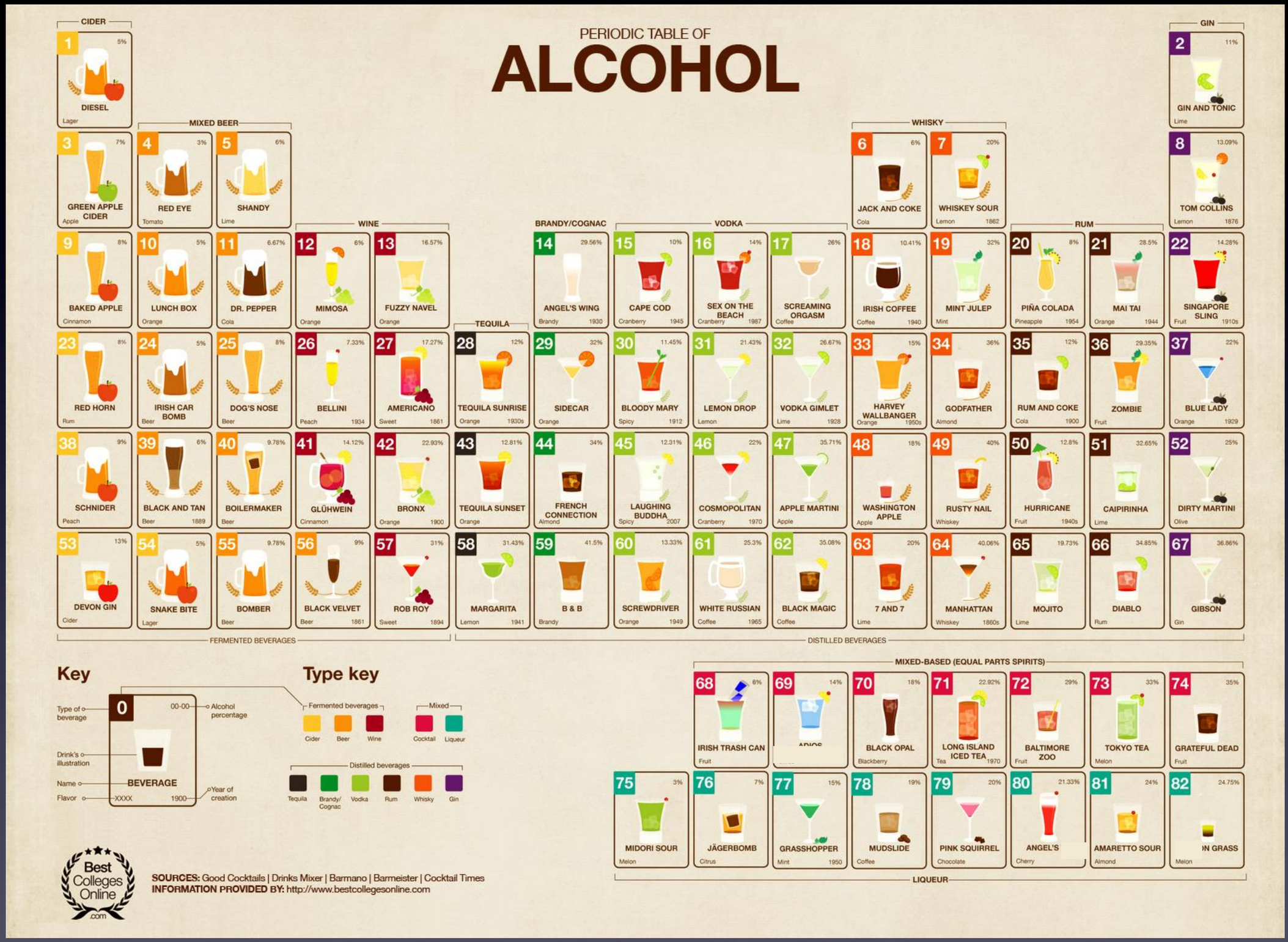
NA62

HyperK

WeNMR

GridPP
UK Computing for Particle Physics

Probably the only way to get your attention…



The question and core….

What technologies and services (not hardware levels) are envisaged as needed by your VO for the period 2015-2020?

- Central site and VO information handling (GOCDB & Ops portal);
- Authorisation/authentication (currently via CA issued certificates and VOMS);
- Ability to distribute data;
- Ability to cataologue data.

- You may be aware of trends or needs developing in your VO that will need to be addressed, for example provision of a job submission interface (such as offered by DIRAC/ganga), direct access to files (e.g. via WebDAV support) or WN root access (as may be possible with a cloud implementation). You may foresee an increased demand for MPI capabilities or improved data locality systems (e.g. hadoop) perhaps with stronger access control, use of whole-node queues or an easy way to link existing job submission frameworks to GPU farms. If so please let me know.

- There is a general need to explore implementing more federated cloud resources, and within GridPP we have a working group doing feasibility studies and testing. This already raises questions, for example, about the provision of and distribution of VM images. We can foresee a possible need for a VM storage and publishing service. If your VO has looked into these technologies your current conclusions (and possible service needs) would be very relevant input.

The question and core….

Diesel - Raw

What technologies and services (not hardware levels) are envisaged as needed by your VO for the period 2015-2020?

- Central site and VO information handling (GOCDB & Ops portal);
- Authorisation/authentication (currently via CA issued certificates and VOMS);
- Ability to distribute data;
- Ability to cataologue data.

- You may be aware of trends or needs developing in your VO that will need to be addressed, for example provision of a job submission interface (such as offered by DIRAC/ganga), direct access to files (e.g. via WebDAV support) or WN root access (as may be possible with a cloud implementation). You may foresee an increased demand for MPI capabilities or improved data locality systems (e.g. hadoop) perhaps with stronger access control, use of whole-node queues or an easy way to link existing job submission frameworks to GPU farms. If so please let me know.

- There is a general need to explore implementing more federated cloud resources, and within GridPP we have a working group doing feasibility studies and testing. This already raises questions, for example, about the provision of and distribution of VM images. We can foresee a possible need for a VM storage and publishing service. If your VO has looked into these technologies your current conclusions (and possible service needs) would be very relevant input.

ALICE Planning

- ALICE has been having in-depth discussion the evolution of the computing model through Runs 2 and 3
  - see eg GDB presentation 2013-09-11- P. Buncic
- Requirements set by expected data-taking capabilities for Run 2 and Run 3 are different
  - They affect both capacity and the way we do things
- Try to move towards solutions in synergy with other experiments to ease support issues
  - eg currently transitioning from Torrent distribution to CVMFS

ALICE Planning

Beer - Fermented

- ALICE has been having in-depth discussion the evolution of the computing model through Runs 2 and 3
  - see eg GDB presentation 2013-09-11- P. Buncic
- Requirements set by expected data-taking capabilities for Run 2 and Run 3 are different
  - They affect both capacity and the way we do things
- Try to move towards solutions in synergy with other experiments to ease support issues
  - eg currently transitioning from Torrent distribution to CVMFS

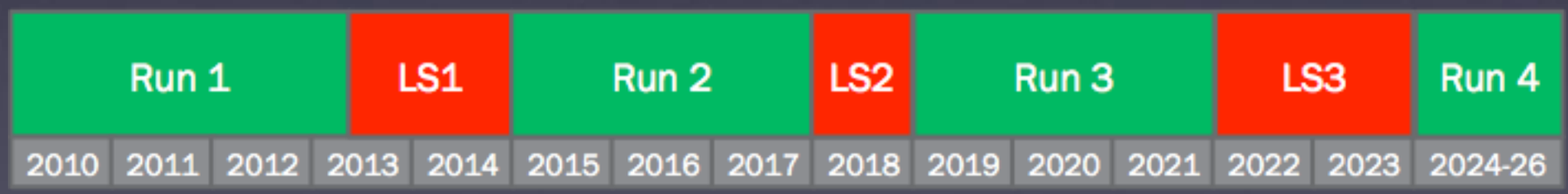
ALICE Plans: Runs 2 & 3

- Run 2 Targets
  - 4-fold increase in Pb-Pb instantaneous luminosity
  - improved readout (factor 2)
  - Existing mix of services, at least until mid-point of Run 2, with little evolution

- Run 3 Targets
  - 100x increase in interaction rate
  - Detector upgrades, continuous readout (1.1 TB/s)
  - Need online reconstruction to achieve data compression
  - Reconfigure services

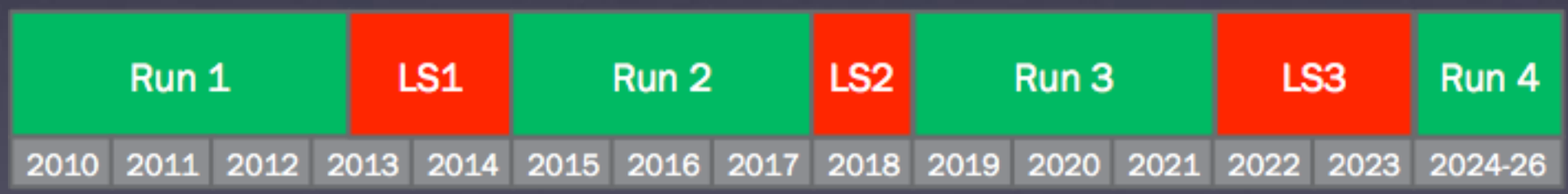
ALICE Request Details

- Disk storage
  - move away from CASTOR for pure disk SE, would prefer EOS
- Grid Gateway
  - Submission to CREAM, CREAM-Condor, or Condor
  - No ARC, not supported in ALiEn



ALICE Request Details

- Authentication. xrootd plugin for file access on SEs
- Simulation. If GEANT4 on GPU is working we aim to be ready to use any GPU provision

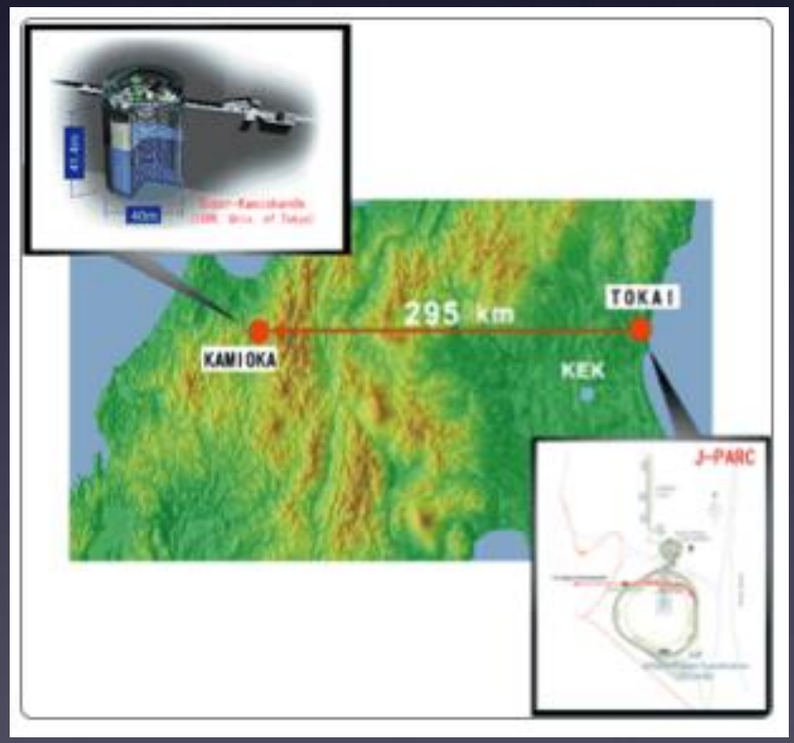


ALICE Request Details

- More flexible Cloud-like processing, opportunistic use
- Basing virtualization strategy around CernVM family of tools
- Studying different use cases
  - HLT Farm, volunteer computing, on-demand analysis clusters

t2k.org service requirements

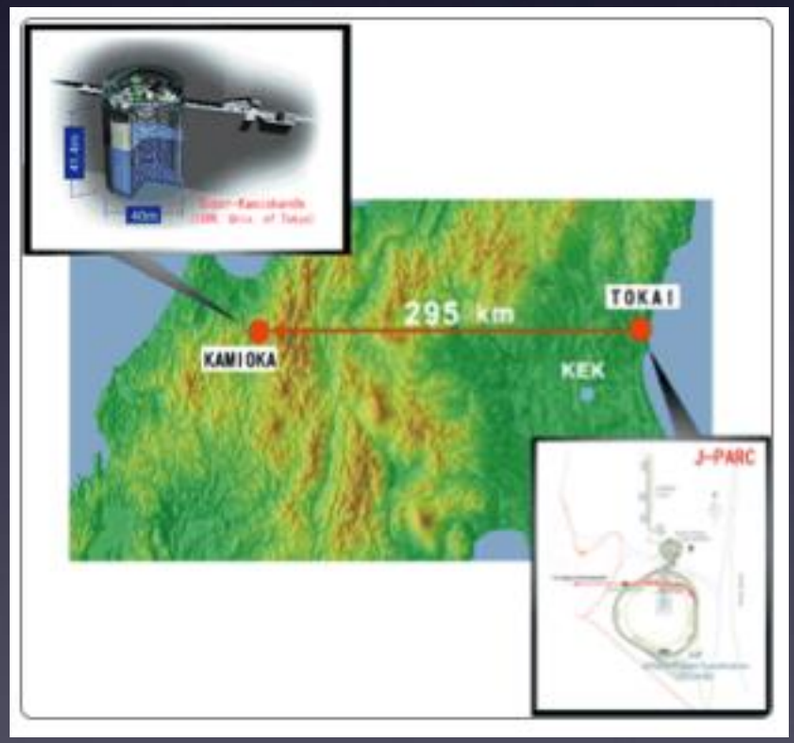
- Exclusively rely on off-the-shelf EMI/gLite tools
  - Wrapped in python home brew
  - No additional frontends (DIRAC, GANGA, etc)
- Exclusively rely on WMS for job submission
  - T2k.org friendly WMSs at RAL and Imperial
  - Requires compatible myproxy server (RAL)
  - Lcg-cp for job I/O
- Exclusively rely on LFC for data distribution and archiving
  - If it isn't in the LFC, we don't know about it
  - FTS2/3 used for file transfers, separate registration
- Can continue the above approach indefinitely
  - IFF the services continue to be developed and supported
  - Exploring DIRAC and GANGA but lack significant manpower
    - Some questions remain over e.g. DIRAC documentation



t2k.org service requirements

Tequila sunrise - Distilled

- Exclusively rely on off-the-shelf EMI/gLite tools
  - Wrapped in python home brew
  - No additional frontends (DIRAC, GANGA, etc)
- Exclusively rely on WMS for job submission
  - T2k.org friendly WMSs at RAL and Imperial
  - Requires compatible myproxy server (RAL)
  - Lcg-cp for job I/O
- Exclusively rely on LFC for data distribution and archiving
  - If it isn't in the LFC, we don't know about it
  - FTS2/3 used for file transfers, separate registration
- Can continue the above approach indefinitely
  - IFF the services continue to be developed and supported
  - Exploring DIRAC and GANGA but lack significant manpower
    - Some questions remain over e.g. DIRAC documentation



t2k.org continued…

- Certification – local CAs and UK NGS
- VOMS – all VO admin via voms.gridpp.ac.uk
- GGUS – primary facility for site/VO/development support
- NAGIOS – (Oxford) real-time site monitoring
- JISC-mail – TB-Support, GRIDPP-Storage
- EGI-portal – VO requirements, VO info, accounting etc.

cvs
ncurses-devel
libX11-devel
libXft-devel
libXpm-devel

libtermcap-devel

libXext-devel

libxml2-devel

- * Longstanding feature request for FTS to register transfers in LFC (GGUS 89038)

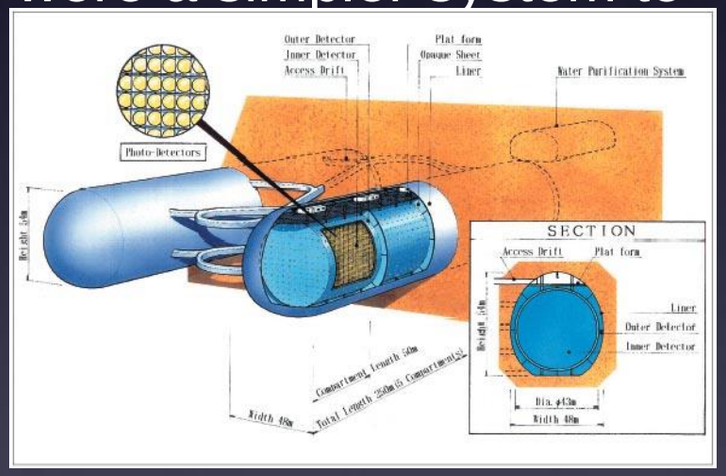
HyperK

- Hyper-Kamiokande, it is also called T2HK (Tokai to Hyper-Kamiokande), however the VO is just hyperk
- The experiment is expected to start to take data in 2023, so in 10 years from now.
- Up to then we expect to take care of MC simulation and process the data from test beams or just lots of cosmics from a prototype.
- Technologies: we want to be at the forefront of the new technologies. The experiment is so much in the future that we will need to be projected towards what will be the latest technologies by then. We would like to make use of Cloud technologies (if possible)
- Currently yes we need certificate management, but if there were a simpler system to use
- Yes we have a need for VO management and access
- many would be an favour of that.
- Used and like iRODS as a mechanism to distribute data.
- The ability to catalogue data would be nice
- Automatic job submission is very much needed.
- Started to look at the Cloud set up in the Summer at KEK.
- Direct file access would be useful. For the cloud we would also like fast access to mounted filesystems (as fast as native access)
- I think for HK we want to move to use the Cloud as soon as reasonably possible.

HyperK

Tequila sunset- Distilled

- Hyper-Kamiokande, it is also called T2HK (Tokai to Hyper-Kamiokande), however the VO is just hyperk
- The experiment is expected to start to take data in 2023, so in 10 years from now.
- Up to that we expect to take care of MC simulation and process the data from test beams or just lots of cosmics from a prototype.
- Technologies: we want to be at the forefront of the new technologies. The experiment is so much in the future that we will need to be projected towards what will be the latest technologies by then. We would like to make use of Cloud technologies (if possible)
- Currently yes we need certificate management, but if there were a simpler system to use
- Yes we have a need for VO management and access
- Used and like iRODS as a mechanism to distribute data.
- The ability to catalogue data would be nice
- Automatic job submission is very much needed.
- Atarted to look at the Cloud set up in the Summer at KEK.
- Direct file access would be useful. For the cloud we would also like fast access to mounted filesystems (as fast as native access)
- I think for HK we want to move to use the Cloud as soon as reasonably possible.



- We have set up a MC production system that has been demonstrated at scale with our 2012/13 production campaign. We are about to start another production campaign (in the time-frame of the next 4-8 weeks) which will take several months. We will no doubt wish to continue and increase this work as NA62 starts to take data next year (so we assume that services we use now will work in the future).

- We plan to work with NA62 to develop a computing model for data-taking, reconstruction, and analysis. This will happen in the next few months. We intend to have a workshop where we invite experts from the LHC experiments and try and get a consensus as to which bits of their computing models we can best re-use. Therefore, the requirements on GridPP are not known in detail, except that they are highly likely to be a subset of the existing LHC services.

Fuzzy navel - Fermented

- We have set up a MC production system that has been demonstrated at scale with our 2012/13 production campaign. We are about to start another production campaign (in the time-frame of the next 4-8 weeks) which will take several months. We will no doubt wish to continue and increase this work as NA62 starts to take data next year (so we assume that services we use now will work in the future).

- We plan to work with NA62 to develop a computing model for data-taking, reconstruction, and analysis. This will happen in the next few months. We intend to have a workshop where we invite experts from the LHC experiments and try and get a consensus as to which bits of their computing models we can best re-use. Therefore, the requirements on GridPP are not known in detail, except that they are highly likely to be a subset of the existing LHC services.

EGI federation testbed technologies and services, future

GridPP
UK Computing for Particle Physics

EGI Core Infrastructure

Service Availability
SAM

Configuration DB
GOCDB

Accounting
APEL

External Services

Certification Authorities

ID Management Federations

AAI

VOMS Proxy

User Interfaces

OCCI Clients
rOCCI; WNoDeS-CLI
SlipStream,
VMDIRAC,
CompatibleOne

CDMI Clients
Libcdmi-java

Resource Providers

VOMS proxy	OpenNebula OpenStack StratusLab WNoDeS Syneffo, others	SSM
OCCI server		Vmcatcher Vmcaster
CDMI server	Marketplace client	LDAP

FedCloud Core Services

Image Metadata
Marketplace

Appliance Repository

Information System
Top-DBII

Monitoring
Nagios + community probes

VO
PERUN server

EGI federation testbed technologies and services, future

Zombie – Distilled… but mixer

EGI Core Infrastructure

Service Availability
SAM

Configuration DB
GOCDB

Accounting
APEL

External Services

Certification Authorities

ID Management Federations

AAI

VOMS Proxy

User Interfaces

OCCI Clients
rOCCI; WNoDeS-CLI
SlipStream,
VMDIRAC,
CompatibleOne

CDMI Clients
Libcdmi-java

Resource Providers

VOMS proxy	OpenNebula OpenStack StratusLab WNoDeS Syneffo, others	SSM
OCCI server		Vmcatcher Vmcaster
CDMI server	Marketplace client	LDAP

FedCloud Core Services

Image Metadata
Marketplace

Appliance Repository

Information System
Top-DBII

Monitoring
Nagios + community probes

VO
PERUN server

From EGI VOs

- IBM: New platforms and application structures. Open industry APIs.
- We-NMR (http://www.wenmr.eu):
  - Work through portals (600 registered users)
  - GPU access would be interesting (molecular dynamics simulations)
  - Transparent access to HPC, HTC, various m/w)
  - Ways to recover job data where queue limit exceeded
  - Resources/queue slots adapted to the application - priorities adapted to requested time.
  - Data sharing a la Dropbox (without x509)
  - Persistent storage and file identification (e.g. Xenodo)
  - File transfer- data repositories (without putting on SE and requiring cert)
  - No need for user interfaces - too much depends on application.
- Integration needs: seamless access to PRACE & XSEDE.
- Flexible Virtual Research Environments and integration with commercial offerings
- GlobusOnline.eu transfer services
- Support with application porting

From EGI VOs

French Connection - Distilled

- IBM: New platforms and application structures. Open industry APIs.
- We-NMR (http://www.wenmr.eu):
  - Work through portals (600 registered users)
  - GPU access would be interesting (molecular dynamics simulations)
  - Transparent access to HPC, HTC, various m/w)
  - Ways to recover job data where queue limit exceeded
  - Resources/queue slots adapted to the application - priorities adapted to requested time.
  - Data sharing a la Dropbox (without x509)
  - Persistent storage and file identification (e.g. Xenodo)
  - File transfer- data repositories (without putting on SE and requiring cert)
  - No need for user interfaces - too much depends on application.
- Integration needs: seamless access to PRACE & XSEDE.
- Flexible Virtual Research Environments and integration with commercial offerings
- GlobusOnline.eu transfer services
- Support with application porting

Commercial provider perspective

- Services
  - Offer a full integration pathway – and automated integration testing
  - Fix current services – buggy cloudmon & Nagios does not work on load-balanced services
  - Provide a penetration testing service (Openstack developers…)
  - Market place based job submission system (with comparible SLAs)
  - More transparent and better pricing models (e.g. select precise requirements)
  - Data distribution services (sync)
  - Image libraries
- Technologies
  - Distributed block storage
  - Increase use of SSDs
  - Less virtualisation and more containerisation
  - Adoption of GPGPUs when libraries more polished

P.S Commercial costs will drop from p/GB to almost zero if/when SJ links established. Charges are data out only. Can now efficiently and automatically rebuild resources in 15 minutes.

A commercial provider joining an academic cloud federation:

Commercial provider perspective

Black Magic - Distilled – Hard to swallow.

- Services
  - Offer a full integration pathway – and automated integration testing
  - Fix current services – buggy cloudmon & Nagios does not work on load-balanced services
  - Provide a penetration testing service (Openstack developers…)
  - Market place based job submission system (with comparible SLAs)
  - More transparent and better pricing models (e.g. select precise requirements)
  - Data distribution services (sync)
  - Image libraries
- Technologies
  - Distributed block storage
  - Increase use of SSDs
  - Less virtualisation and more containerisation (go native)
  - Adoption of GPGPUs when libraries more polished

P.S Commercial costs will drop from p/GB to almost zero if/when SJ links established. Charges are data out only. Can now efficiently and automatically rebuild resources in 15 minutes.

A commercial provider joining an academic cloud federation:

Grid components – GridPP29

GridPP

NGS

Phedex agents
DDM

SRM-storage

VO-box

FTS

Squid/cache

S/w repositories

Ganga

Pakiti

GridPP website
GridPP blogs
GridPP twiki

VM image catalogue

Regional dashboard/MyEGI

ARC

CREAM-CE

UNICORE

ARGUS

VOMS

CA

tBDII

myProxy

Security

WMS

UI

MPI

LFC

APEL-Accounting

Helpdesk

Outreach

Nagios

GOCDB

NGI website

OGSA-DAI

Globus WS

VDT/Globus

SRB-storage

UAS

AHE

IaaS Cloud

Viz Services

WS Hosting

Penetration testing

MEG

SARoNGS

Relational DB

NGS website
NGS blogs
NGS twiki

P-Grade Portal

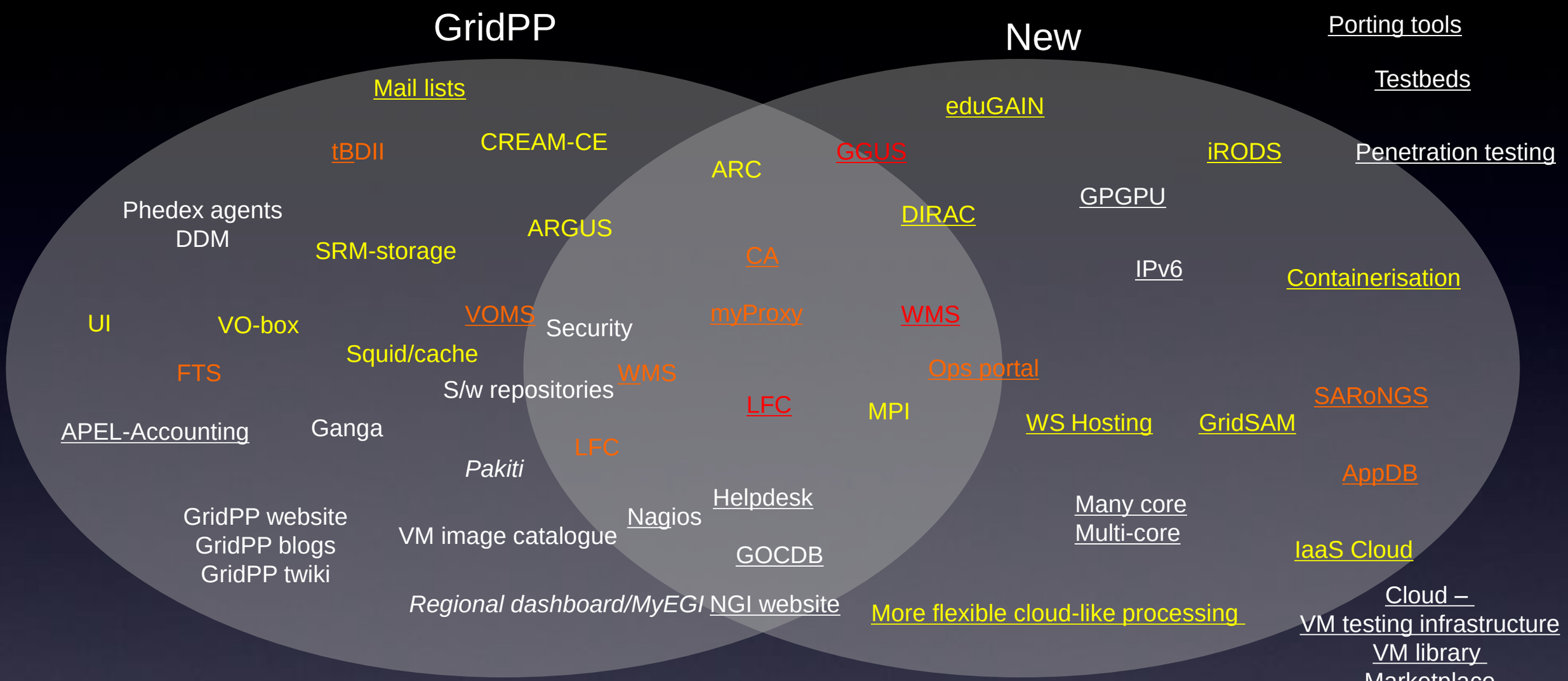
More GridPP specific ⟷ More NGS specific

Site component

Central NGS service

Grid infrastructure

Summary - illustrative



* Existing systems must be maintained/developed