

EGI InSPIRE HUC SA3 – Services for Heavy User Communities of DCIs

Jamie.Shiers@cern.ch

Experiment Support Group, CERN IT

April 2010, Uppsala

Overview

- This talk explains **which communities** are targeted by the work that is foreseen and outlines **mechanisms** whereby its progress can be tracked
- It examines how the **potential** benefits – within individual communities, between them (e.g. **common** tools and services) as well as to the more general DCI community and **beyond** – can be achieved
- **Goal:** start the discussion – **not** present a plan

Which Communities?

- The communities identified as Heavy Users Communities (HUCs) within the EGI InSPIRE proposal are:
 - High Energy Physics (HEP)
 - Life Sciences (LS)
 - Astronomy and Astrophysics (A&A)
 - Computational Chemistry and Materials Sciences and Technologies (CCMST)
 - Earth Sciences (ES)
 - Fusion (F)
- There are tools and services that are (typically) of use (today) to more than one community within this set, together with the potential for wider adoption also beyond

1. How to best organize the work amongst these communities?

2. How to maximize benefit to others?

Which Tools & Services?

- **Specific:** e.g. Ganga
- **Frameworks:** e.g. dashboard(s)
- **Generic:** workflows / schedulers
- Within each of these areas – as well as those not explicitly mentioned – our experience tells us that there can be **much** to gain from **collaboration**
 - But it requires an **investment** and time to pay off
 - Several examples of this included as “BACKUP” slides...
- **Services:** services for HEP is a significant component (PMs) of this activity
 - This has been an area of major investment / progress over the duration of EGEE I/II/III
 - What can be re-used by other communities?
 - What can be generalized?

Milestones & Deliverables

What	When
Report HUC Contact points and the support model (CSC)	M: 1
Capabilities offered by the HUCs to other communities (TCD)	D: 4
HUC Software Roadmap (INFN)	M: 4
Services for High Energy Physics (CERN)	M: 4
Training and dissemination event for all shared services (TCD)	M: 4
Sustainability plans for the HUC activities (CNRS)	D: 9
HUC Software Roadmap (INFN)	M: 10
Annual Report on the HUC Tools and Services (CERN)	M: 11
Hydra service deployment on a multi-servers configuration (CNRS)	M: 12
Integration of the VisIVO server with the production infrastructure (INFN)	M: 12
Medical Data Manager release (CNRS)	M: 12

How to Measure Success?

- ✓ **Directly** – by agreeing up-front goals and metrics with our communities and evaluating them on a regular (e.g. quarterly) basis
 - Light-weight community-focused reporting which feeds into EGI QRs (felt useful – even necessary – for HEP...)
 - WLCG has on-going service reports based on KPIs (initially weekly but now ~twice monthly) & management reporting (quarterly / twice-yearly)
- ✓ **Indirectly** – by the “impact” this funding is intended to achieve
 - Also a measure of the above – the supported communities should (preferably) actively acknowledge the contribution(s) to their goals
 - e.g. “The scientific potential of the LHC” & equivalent
- ✓ **And beyond** – in terms of timescale & “community” in the sense of Science & Society
 - Most over-used example is the Web: what has been the value to the global economy and human-kind from this LEP-era development?
 - A (**THE?**) key motivation for international and cross-disciplinary funding

Sub-activities: HEP

- 60PM: Dashboards ([Julia Andreeva](#))
- 60PM: Ganga/Diane ([Massimo Lamanna](#))
- 200PM: VO services ([Maria Girone](#))
+60 PM via INFN
- CERN-based activity hosted in IT Experiment Support group
 - Leader [Jamie Shiers](#), Deputy Maria Girone
 - (both have specific roles in WLCG Service activities)
- **Regular discussions with LHC computing managers (within the experiments) to agree priorities, assignments, results, future needs etc.**
- Also CERN + experiments + INFN to ensure overall priorities understood and agreed, in particular for VO services

Light-weight Reporting

VO / project	Purpose	Man. Contacts	Tech. Contacts	Goals	Achievements	Support requested	Time Spent	Issues & Concerns
SIXT	Simulation of particle bunches in LHC							
...								
ILD	International Linear Detector studies							
...								
PART-NER	Hadron therapy							
Enviro-Grids	Black Sea							
...								
ATLAS	Main detector / collaboration at LHC							
...								

- The sum of the non-LHC + non-HEP VOs supported by “Experiment Support” is quite significant!
- Some minimal level of accounting is needed

SERVICE ISSUES

Some key service issues from the viewpoint of WLCG

WLCG – Strengths

- Key strengths – that have allowed us to deliver and to adapt – are:

1. The WLCG Operations Model;

2. Personnel – at sites and in the experiments;

3. The Service itself.

WLCG Operations Report – Structure

KPI	Status	Comment
GGUS tickets	No alarms; normal # team and user tickets	No issues to report
Site Usability (SAM-based)	Fully green	No issues to report
SIRs & Change assessments	None	No issues to report

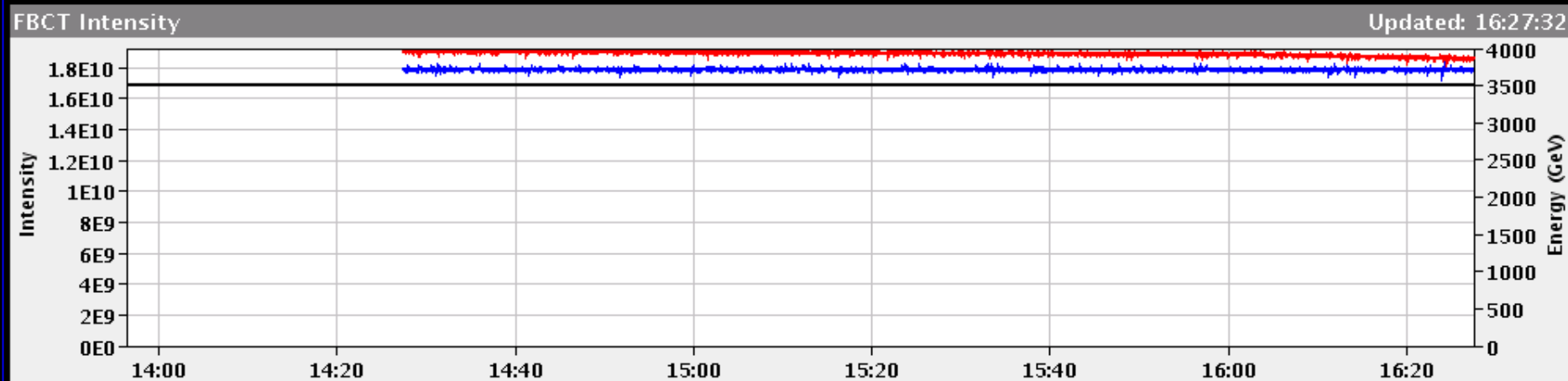
KPI	Status	Comment
GGUS tickets	Few alarms; normal # team and user tickets <i>and/or</i>	Drill-down
Site Usability	Some issues <i>and/or</i>	Drill-down
SIRs & Change assessments	Some	Drill-down

KPI	Status	Comment
GGUS tickets	Many alarms / other tickets	Drill-down
Site Usability	Poor	Drill-down
SIRs & Change assessments	Several	Drill-down

Normally; **Sometimes;** **Rarely / Never**

PROTON PHYSICS: STABLE BEAMS

Energy: 3500 GeV I(B1): 1.45e+10 I(B2): 2.10e+10



Comments 12-04-2010 16:06:39 :

**** STABLE BEAMS ****

LUMI SCAN FINISHED FOR ALL EXPS

!!!"ADJUST" Handshake around 16:30!!!

Squeeze at 5 p.m.

BIS status and SMP flags

B1

B2

Link Status of Beam Permits

true

true

Global Beam Permit

true

true

Setup Beam

true

true

Beam Presence

true

true

Moveable Devices Allowed In

true

true

Stable Beams

true

true

LHC Operation in CCC : 77600, 70480

PM Status B1

ENABLED

PM Status B2

ENABLED

WLCG – Weaknesses

- The service is not perfect – it is [still] improving but there are a number of problems that clearly characterize it
- **Given our past experience, it does not appear likely that these weaknesses will be solved anytime soon**
- 1. **Delays** in introducing new services or service versions;
- 2. **Chronic** problems – which either come and go or simply persist “for ever” – a cause of much frustration
- 3. **Problems** in data / storage management and data access – these deserve a discussion of their own

General Observations

- Running on the grid has been relatively smooth during and after data taking
- Data distribution was normally quick around all sites
- Reprocessing ran smoothly at T1s
- Analysis has been working well at T2s
- There have been many minor problems, but these have mostly been resolved quickly by sites
- So we would like to say thank you to all sites for their efforts and stability



Jet Event at 2.36 TeV Collision Energy

2009-12-14, 04:30 CET, Run 142308, Event 482137
<http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html>

Opportunities (External)

- There have been many developments in recent years that have occurred in parallel to the build-up of WLCG
- In some cases there is active R&D – or even exploitation – of these technologies / methodologies
- We can “ignore” them for a short period – but not forever
- And we should not – we have much to benefit from them but the impact may be significant!
- The list could include clouds, virtualization, multi-core, storage hardware, filesystems, access protocols, data management techniques, ... virtually open-ended!

WLCG – Threats (External)

Potential Threat	Potential Impact (-ve)
<p>Funding problems for sites or countries;</p> <p>Funding problems for projects (e.g. ROSOE)</p>	<p>Resources at sites;</p> <p>Support issues: ROSCOE would have provided resources at CERN for Experiment Integration & Operations support, Dashboard support and Distributed Analysis Support at several sites (IC, Bham, Oslo, FZU, INFN + non-LHC VOs: FAIR, ILC, ...)</p> <p>For CERN alone the loss is 4-5 funded FTEs over 3 years</p>
<p>Move from EGEE to EGI & NGIs – many “sub-threats”, e.g. in handling of releases and pre-production, user support, basic operations tools & portals</p>	<p>Instability; loss of needed functionality; extra cost (multi-party negotiations?) – complications introduced by distance...</p> <p>We have been preparing for this transition for several years and are “largely” independent...</p>

WLCG Service Summary

↑ **Massive** increase in service usage & deployed resources

= Sum of resources (at any one time) at different tiers ~constant

↘ **Flat** or **decreasing** support load

= Sum of tickets at different tiers ~constant

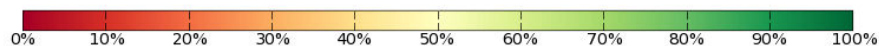
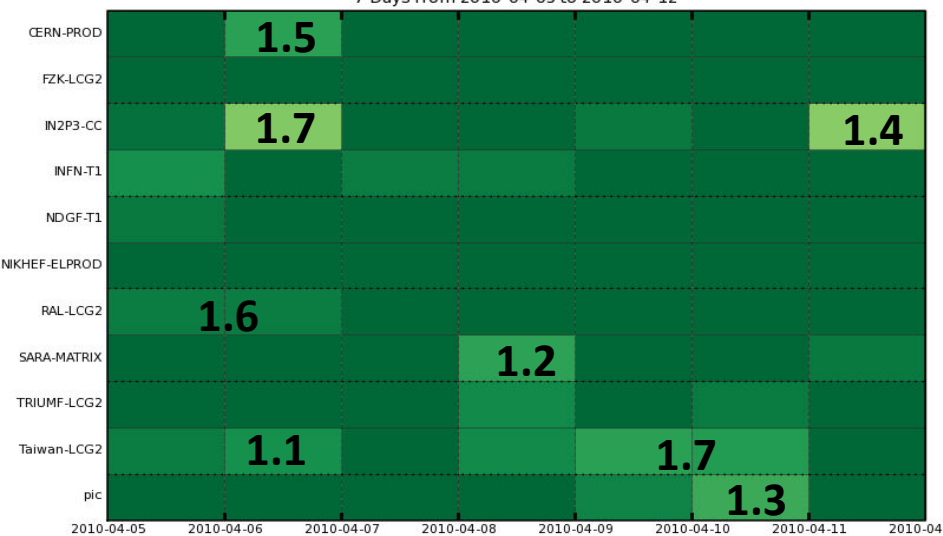
➤ **Service is now fully delivering:**

- ✓ **Enabling more & better science;**
- ✓ **Encouraging and facilitating inter-disciplinary and international collaboration(s);**
- ✓ **Tangible, long-term benefits to science & society.**

ATLAS

Site Availability using WLCG_SRM2

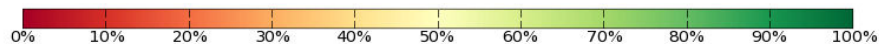
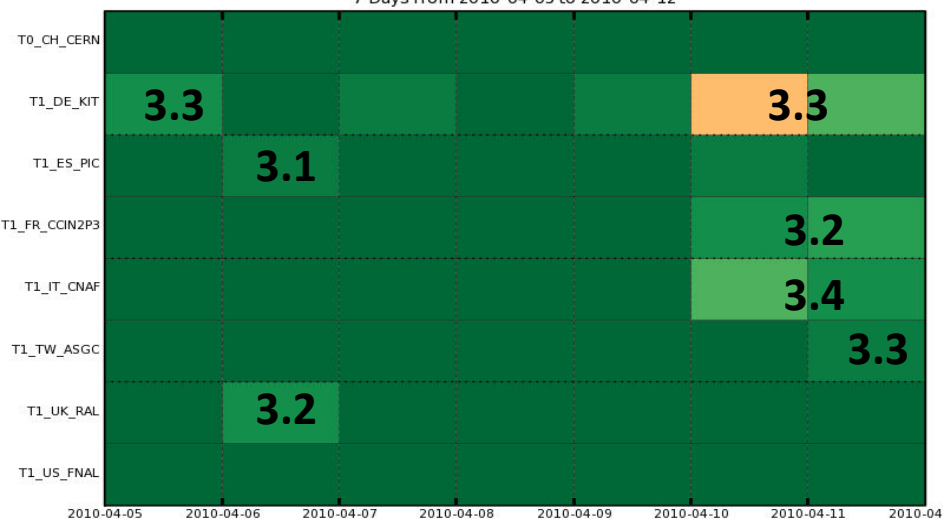
7 Days from 2010-04-05 to 2010-04-12



CMS

Site Availability

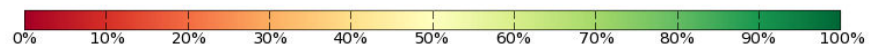
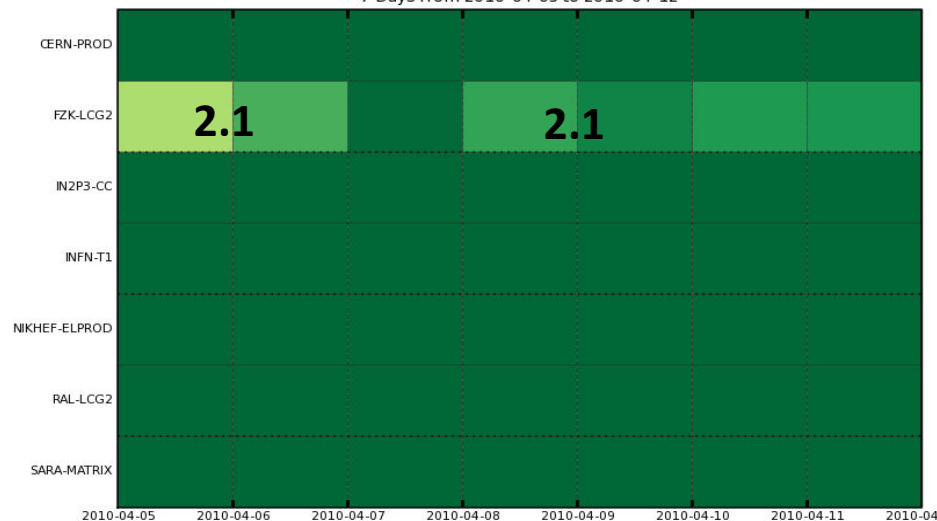
7 Days from 2010-04-05 to 2010-04-12



ALICE

Site Availability using WLCG Availability (FCR critical

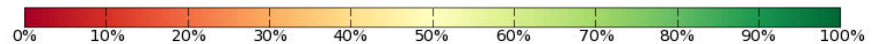
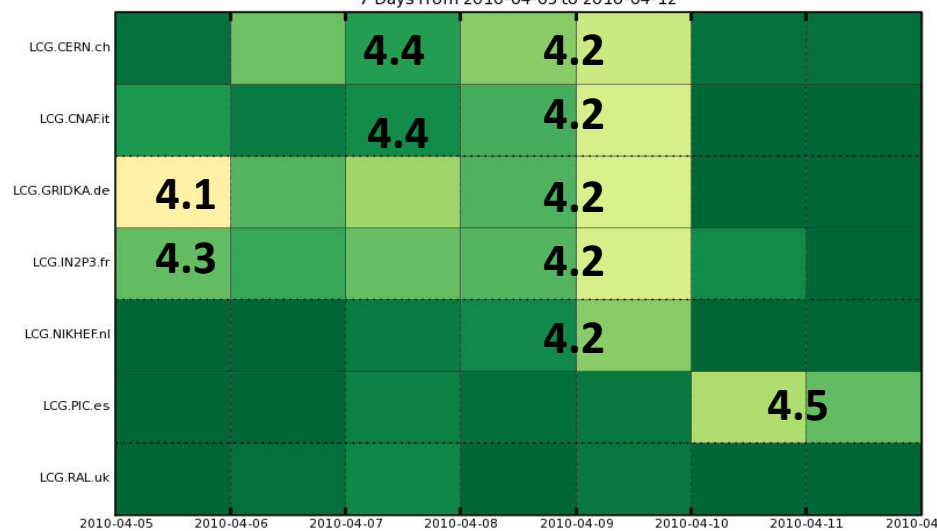
7 Days from 2010-04-05 to 2010-04-12



LHCb

Site Availability using LHCb Critical Availability

7 Days from 2010-04-05 to 2010-04-12



ATLAS

- 1.1 ASGC:** disk server problem, fixed
- 1.2 SARA-MATRIX:** short scheduled SE downtime
- 1.3 PIC:** Transfer errors, SRM down for over 2 hours, more gridftp doors in production solves the problem
- 1.4 IN2P3:** SRM down (a lot of activity during the night). SRM server restarted, OK
- 1.5 CERN:** File transfer problem: failed to contact on remote SRM, caused by a BDII problem, fixed
- 1.6 RAL:** failing transfer attempts, was immediately understood, the file was successfully transferred
- 1.7 ASGC, IN2P3:** SRM tests failures (timeouts)

ALICE

- 2.1 FZK:** Temporary SAM Test Failure for VOBOX-User-Proxy-Registration Test

CMS

- 3.1 PIC:** file access problem: repeated errors trying to access one file during the 355 rereco preproduction
- 3.2 RAL, IN2P3:** temporary SRM test failures (timeouts)
- 3.3 KIT, ASGC:** MARADONA errors
- 3.4 CNAF:** SRM tests failures

LHCb

- 4.1 GRIDKA:** banned for shared area issue. In progress
- 4.2** Setup the CE-sft-vo-swdir test with a more aggressive operation on the shared area (it does not propagate properly the information of the killing process because running out of time to the final SAM test results (will be fixed soon))
- 4.3 IN2P3:** restart of the LHCb services running on LHCb VO box due to wrong version of dirac
- 4.4** Some router instabilities occurred, which could have affected CNAF, IN2P3; network outages on the LHCb ONLINE DB
- 4.5 PIC:** file access test failures: Error in ROOT macro opening/reading the file

InSPIRE SA3: Next Steps...

- **Representatives** of the different communities – in terms of the services that will be provided through this work-package – need to **meet** and **agree** on the above and other key issues
- Much of this can be virtual: phone, video, collaborative tools, knowledge bases etc, but **F2F** meetings – planning, reviewing, technical workshops etc. – should also be considered
- **Where possible building on existing commitments / events ??**
 - I guess we will all be going to Amsterdam quite often this year – a good reason to go somewhere else?
 - I would like to see things happening << summer 2010 – even if some important aspects (grant agreements etc.) may still be underway [**needed for WLCG!**]
- **Use this week! As many formal / informal discussions as possible**
- Some **events** are already being discussed around the time of the EGEE III final review: an occasion for a follow-up discussion? (At least amongst the key representatives mentioned above?)

Specific Ideas & Plans

- Within the HEP community we have ideas and plans for how to move this forward – specifically addressing the needs of our community (LHC-focused) but also collaborations using **Ganga**, **Dashboards** and others
 - These are along the lines of what is discussed above which could be input for wider discussions
- Rather than spend time on these here, I would prefer to use most of the remaining time on open issues
- Specifically, I would like to ensure that representatives of all of the communities in SA3 meet (immediately after this?) and schedule **further discussions** and **concrete actions**
 - This week and beyond...

Looking Further Ahead...

- EGI conference in September – assume an SA3 session
- WLCG workshop in July – service oriented – possibility to add additional tracks
 - Cross-fertilization between different disciplines should be a priority
- Topical workshops: June on storage and data access; end year / early 2011?

Summary

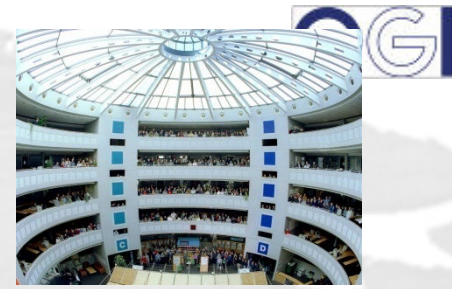
- We are at the beginning for a multi-year, multi-disciplinary, multi-national project with a strong focus – at least in this WP – on service
- **We have much experience in this domain but still a long way to go...**
- **Collaboration** is a **proven** mechanism for providing solutions that are of wider applicability than those developed within single domains – are typically (much) **better** for it!



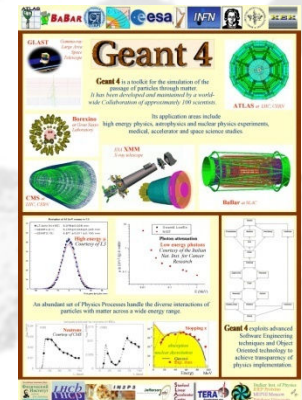
BACKUP

COLLABORATION – MOTIVATION

Collaboration(s)



- The CERN Program Library
- GEANT4 (RD44)
- The (Worldwide) Web



CERNLIB

- The **CERN** Program Library is a large collection of general purpose programs maintained and offered in both source and object code form on the **CERN** central computers. Most of these programs were developed at **CERN** and are therefore oriented towards the needs of a **physics research** laboratory.

Nearly all, however, are of a general mathematical or data-handling nature, applicable to a wide range of problems.

- The library contains several thousand subroutines and complete programs which are grouped together by logical affiliation into several hundred program packages. 80% of the programs are written in Fortran and the remainder in assembly code, or C usually with a Fortran version also available. The language supported is currently Fortran.
- [Wikipedia entry](#)

GEANT4 (né RD44)

- **Geant4** (for **GEometry ANd Tracking**) is a platform for "the simulation of the passage of particles through matter," using Monte Carlo methods. It is the successor of the GEANT series of software toolkits developed by CERN, and the first to use Object oriented programming (in C++).
- Its development, maintenance and user support are taken care by the international Geant4 Collaboration.
- Application areas include high energy physics and nuclear experiments, medical, accelerator and space physics studies. The software is used by a number of research projects around the world.
- Wikipedia entry

The World Wide Web

- Often abbreviated as **WWW** and commonly known as **The Web**, is a system of interlinked [hypertext](#) documents contained on the [Internet](#). With a [web browser](#), one can view [web pages](#) that may contain [text](#), [images](#), [videos](#), and other [multimedia](#) and navigate between them by using [hyperlinks](#).
- Using concepts from earlier hypertext systems, British engineer and computer scientist [Sir Tim Berners Lee](#), now the Director of the [World Wide Web Consortium](#), wrote a proposal in **March 1989** for what would eventually become the World Wide Web.^[1] He was later joined by Belgian computer scientist [Robert Cailliau](#) while both were working at [CERN](#) in [Geneva, Switzerland](#). In 1990, they proposed using "HyperText [...] to link and access information of various kinds as a web of nodes in which the user can browse at will",^[2] and released that web in December.^[3]
- "The World-Wide Web (W3) was developed to be a pool of human knowledge, which would allow collaborators in remote sites to share their ideas and all aspects of a common project." ^[4] If two projects are independently created, rather than have a central figure make the changes, the two bodies of information could form into one cohesive piece of work.

The WLCG Operations Model...

- Centred around the daily [Operations call](#) at 15:00 Geneva time
 - These calls started prior to CCRC'08 and have been key to ensuring good information flow, action on tickets and other problems: peer-to-peer in all respects
- There have been a number of simple optimizations this year: now all experiments provide pre-reports that are pasted into the daily minutes and displayed during the meeting
 - Comments & site reports added online: meeting minutes typically available by close-of-meeting! Issues at T0/T1/T2 discussed...
- Regular reporting using KPIs to Management Board – these reports too have been optimized!
 - The recently established [Tier1 Service Coordination](#) meeting also plays an important role addressing issues on a longer timescale (held every two weeks), such as FTS 2.2.3 roll-out...

WLCG Service – The KPIs

- Include:
 1. Drill-down on any anomalies during the reporting period
 - E.g. incidents that triggered a Service Incident Report – typically due to an outage or degradation longer (or of the order) of those listed in the MoU (“maximum delay in responding to operational problems”...)
 2. Summary of GGUS tickets and drill-down on each and every alarm ticket
 - Almost all alarms are tests: we compare against the WLCG Critical Service targets in the following table
 3. Site Status Reports – using the experiment-specific SAM tests
 - Annotated with an explanation of all issues
- An analysis of Change Assessments – coming from discussion triggered by ATLAS at January GDB



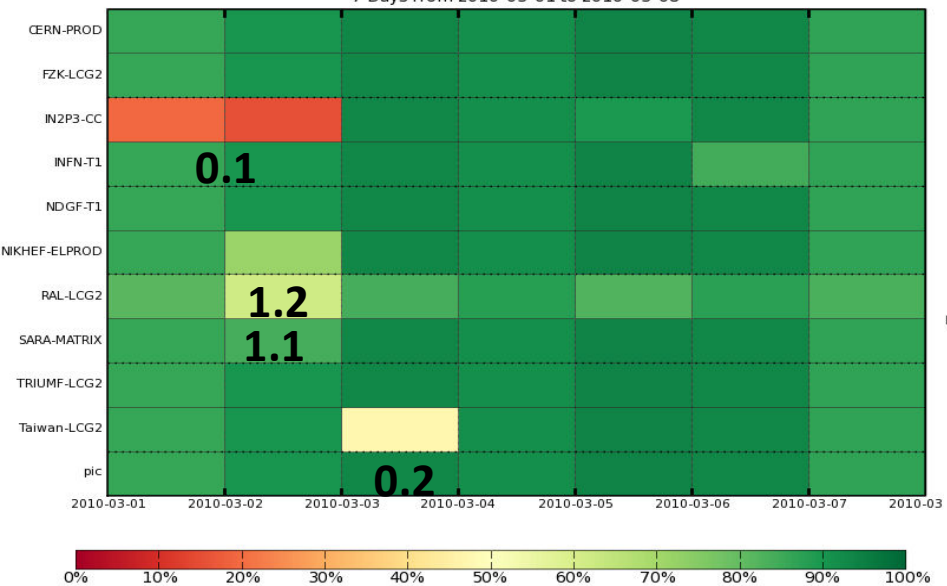
Critical Service Follow-up

- Targets (not commitments) proposed for Tier0 services
 - Similar targets requested for Tier1s/Tier2s
 - Experience from first week of CCRC'08 suggests targets for **problem resolution** should not be too high (if ~achievable)
 - The MoU lists targets for responding to problems (12 hours for T1s)
- ¿ **Tier1s: 95% of problems resolved <1 working day ?**
- ¿ **Tier2s: 90% of problems resolved < 1 working day ?**
- **Post-mortem triggered when targets not met!**

Time Interval	Issue (Tier0 Services)	Target
End 2008	Consistent use of all WLCG Service Standards	100%
30'	Operator response to alarm / call to x5011	99%
1 hour	Operator response to alarm / call to x5011	100%
4 hours	Expert intervention in response to above	95%
8 hours	Problem resolved	90%
24 hours	Problem resolved	99%

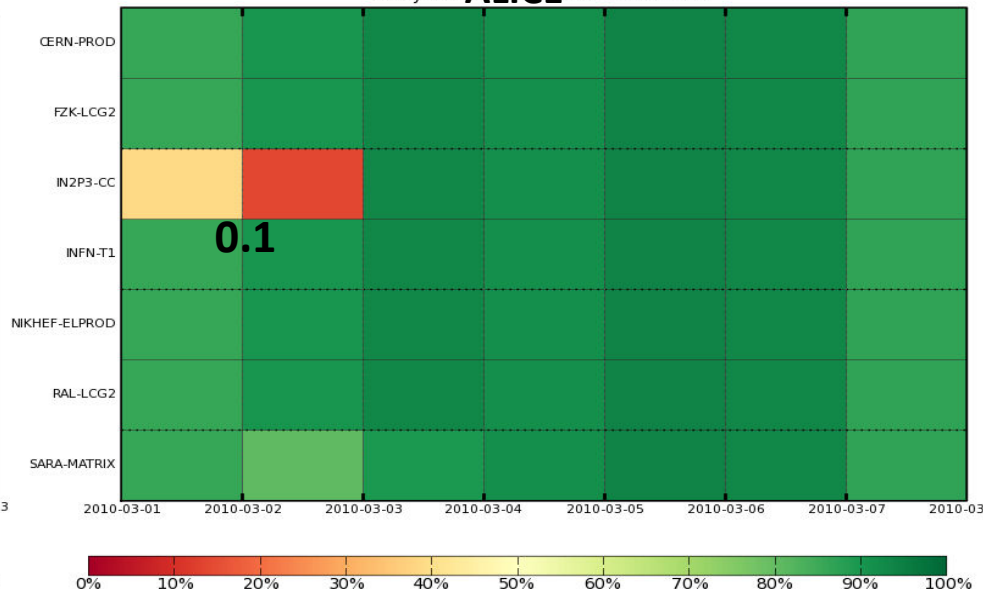
Site Availability using WLCG_SRM2

7 Days from 2010-03-01 to 2010-03-08



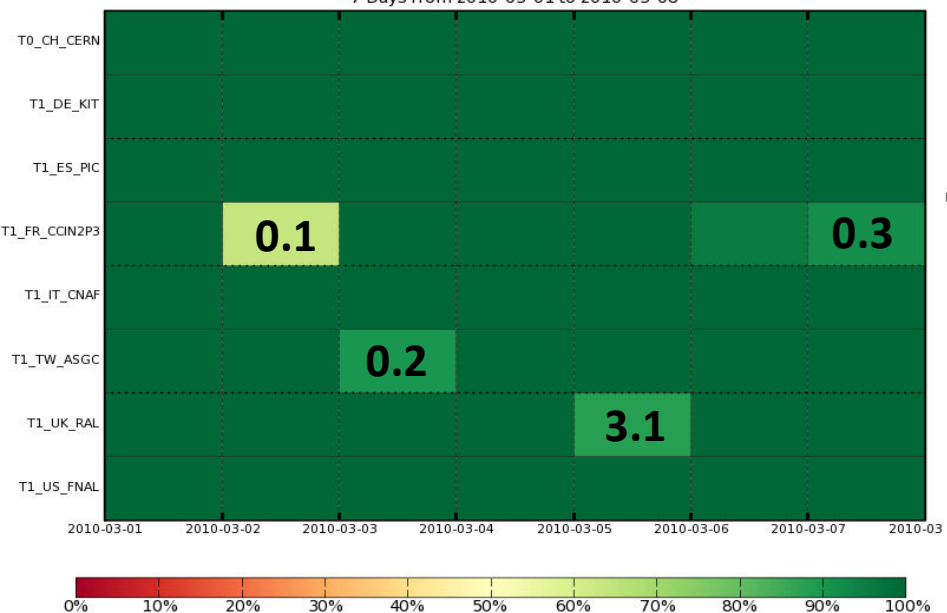
Site Availability using WLCG Availability (FCR critical)

7 Days from 2010-03-01 to 2010-03-08



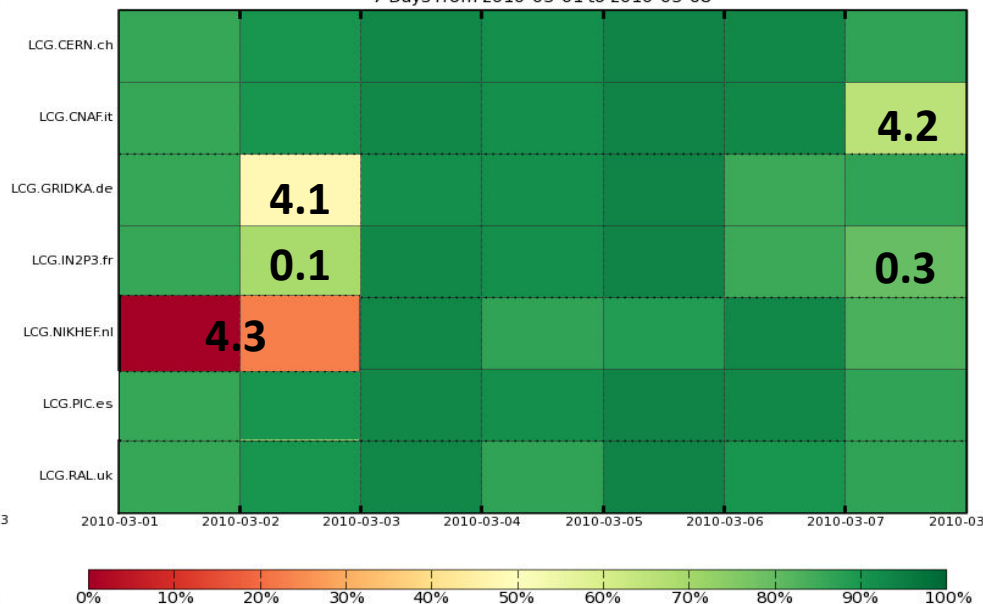
CMS Site Availability

7 Days from 2010-03-01 to 2010-03-08



LHCb Site Availability using LHCb Critical Availability

7 Days from 2010-03-01 to 2010-03-08



Analysis of the availability plots



COMMON FOR THE ALL EXPERIMENTS

0.1 IN2P3: Planned outage for maintenance of batch and mass storage

0.2 TAIWAN: Scheduled downtime Wednesday morning. Most services got recovered quickly except for lfc and FTS due to some oracle block error, a 2-hour unscheduled downtime was created for this. These 2 services got recovered at 14:15

0.3 IN2P3: SAM SRM tests failures, disappeared after ~4 hours (problems with BDII. Known performance problem of SL5 BDII)

ATLAS

1.1 RAL: SRM overload (tests hitting 10 minutes timeouts). Two ATLASDATADISK servers out with independent problems

1.2 NIKHEF: Problem with one disk server (seem to be due to an Infiniband driver. Kernel timeout values need to be increased)

ALICE

Nothing to report

CMS

3.1 RAL: Temporary tests failures due to deployment of the new version of the File Catalog

LHCb

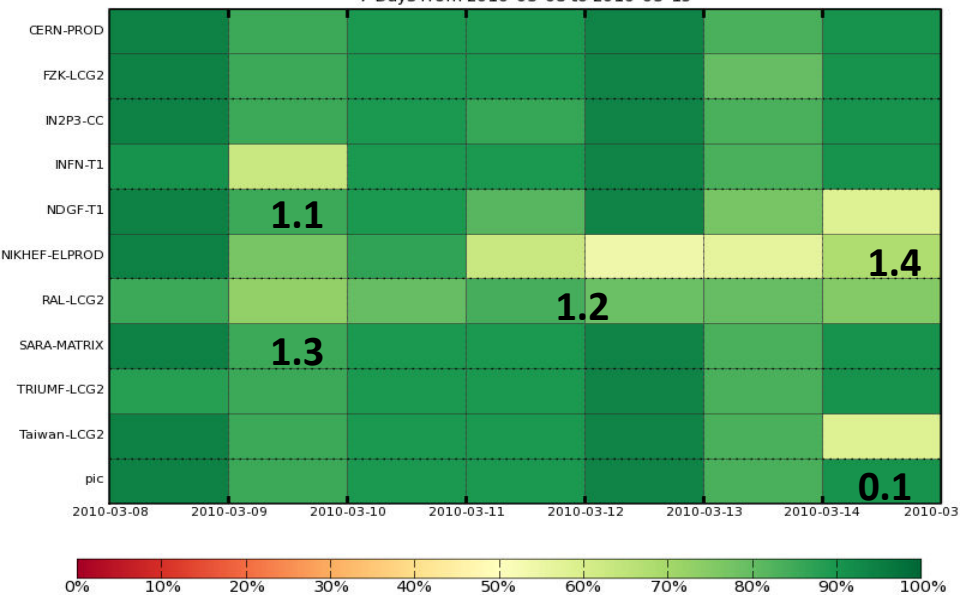
4.1 GRIDKA: SQLite problems due to the usual nfslock mechanism getting stuck. Restarted the nfs server

4.2 CNAF: Problems with local batch system, investigating

4.3 NIKHEF: The critical File Access SAM test failing has been understood by the core application devs as some libraries (libgsitunnel) for slc5 platforms not properly deployed in the AA.

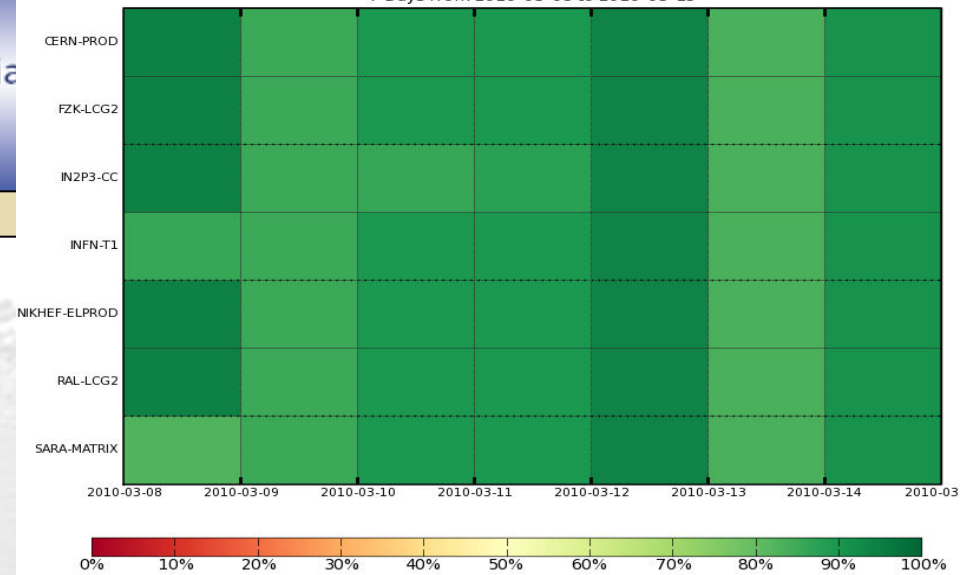
Site Availability using WLCG_SRM2

7 Days from 2010-03-08 to 2010-03-15



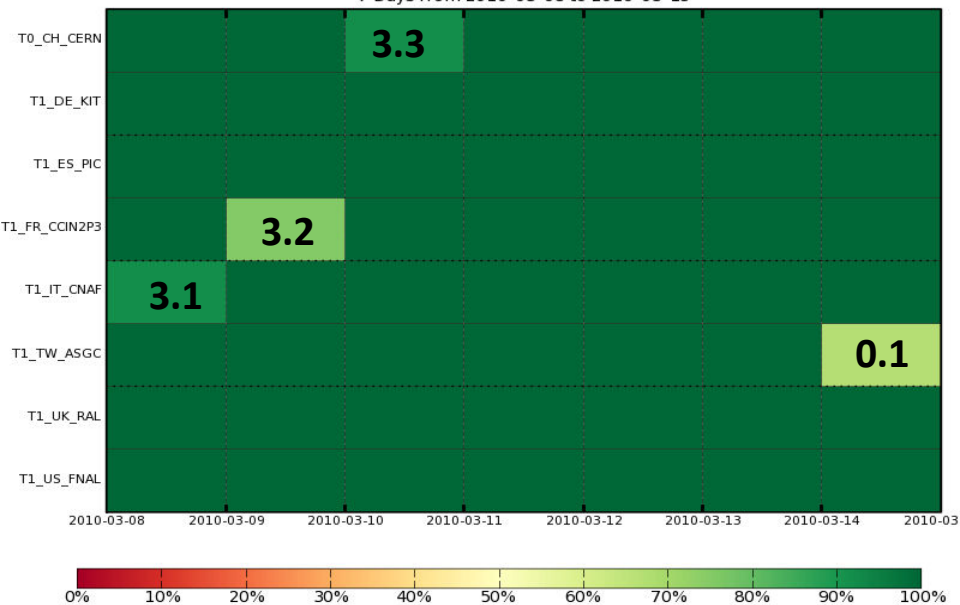
Site Availability using WLCG Availability (FCR critical)

7 Days from 2010-03-08 to 2010-03-15



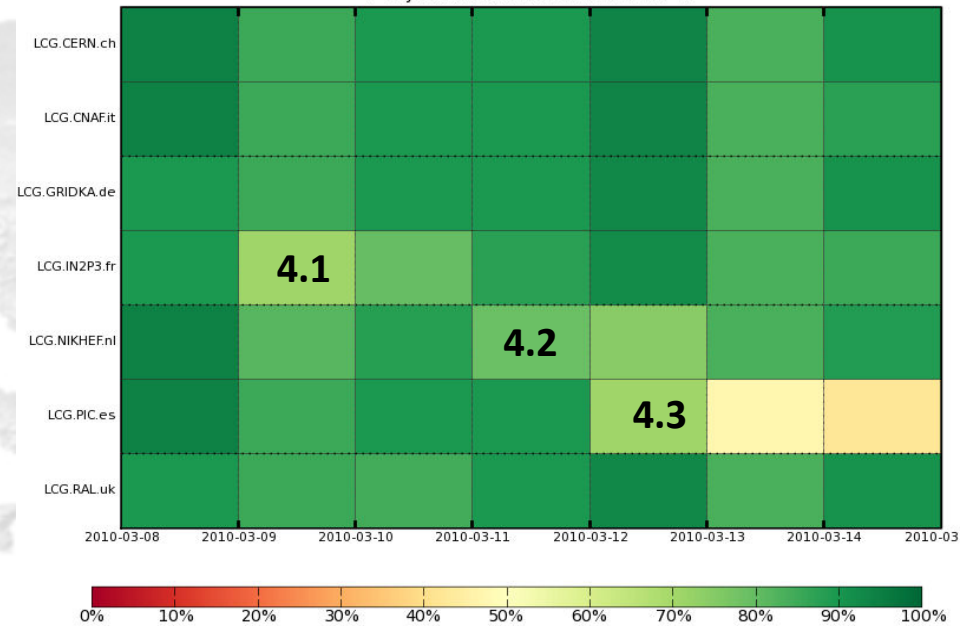
CMS Site Availability

7 Days from 2010-03-08 to 2010-03-15



LHCb Site Availability using LHCb Critical Availability

7 Days from 2010-03-08 to 2010-03-15



Analysis of the availability plots



COMMON FOR THE ALL EXPERIMENTS

0.1 TAIWAN: Unscheduled power cut

ATLAS

1.1 INFN: Enabled checksums for INFN-T1 in FTS. Problems observed, the checksum switched off

1.2 NIKHEF: Unscheduled downtime (FTS is down). Unable to start the transfer agents

1.3 RAL: Disk server out of action, part of ATLAS MC DISK. SRM overload (tests hitting 10 minutes timeouts)

1.4 NDGF: LFC's host certificates have expired, fixed. LFC daemon giving a core dump, under investigation

ALICE

Nothing to report

CMS

3.1 CNAF: CE SAM test failures - LSF master dying

3.2 IN2P3: SRM test failure (authentication problems)

3.3 CERN: Temporary SAM test failure (timeout)

LHCb

4.1 IN2P3: Temporary test failures (software missing)

4.2 NIKHEF: Temporary test failures due to migrating and testing a new test code

4.3 PIC: Application related issues during the w/e on the certification system, accidentally were published in SAM. Experts contacted, fixed

Service Incidents & Change Assessments

<i>Site</i>	<i>When</i>	<i>Issue (as reported in Service Incident Report)</i>
CERN	3 Mar	Replication of LHCb conditions Tier0->Tier1, Tier0->online partially down (18 hours).

▪ Details in next slides

<i>Site</i>	<i>When</i>	<i>Change Assessment & Results</i>
CERN	9 Mar	update SRM-ATLAS to 2.9, likely rollback

- Full Change Assessment [here](#)
- Summary from IT-DSS & ATLAS viewpoints follows