The (WLCG) Common Computing Readiness Challenge – CCRC'08

Jamie.Shiers@cern.ch

~ ~ ~

HEPiX, 8th May 2008

Overview

- CCRC'08 – What's it all about?
- Specific Computing Challenges
- Specific **Grid** Computing Challenges
- Specific **WLCG** Computing Challenges
- What have we learned – so far...
- Outlook...

CCRC'08 - BACKGROUND

Introduction (January)

- 2008 promises to be a very busy – and hopefully very rewarding year
- This is the year in which collisions in the LHC are planned and will exercise **all** aspects of the **WLCG** Computing Service at **all** sites and for **all** supported experiments **concurrently**
- **The goal of the CCRC'08 exercises is to understand where we stand with respect to these needs and to identify and fix any problems as rapidly as possible**
- **'No surprises' would be a surprise in itself!**
- We must **assume** that not everything will work as expected / hoped and work as efficiently as possible towards **solutions**
- The (very) recent past has numerous examples where 'show-stoppers' "disappeared" almost overnight!

CCRC'08 – Background

- For many years, the LHC experiments have been preparing for data taking

- On the Computing side, this has meant a series of "**Data Challenges**" designed to verify their computing models and offline software / production chains
- To a large extent, these challenges have been independent of each other, whereas in reality, we (**almost all sites**) have to support (**almost all experiments**) simultaneously

💣 **Are there some bottlenecks or unforeseen couplings between the experiments and / or the services?**

☝ **There certainly is at the level of support personnel!**

A Comparison with LEP…

- In **January 1989**, we were expecting e^+e^- collisions in the summer of that year...
- The “MUSCLE” report was 1 year old and “Computing at CERN in the 1990s” was yet to be published (**July 1989**)

☝ **It took quite some time for the offline environment (CERNLIB+experiment s/w) to reach maturity**

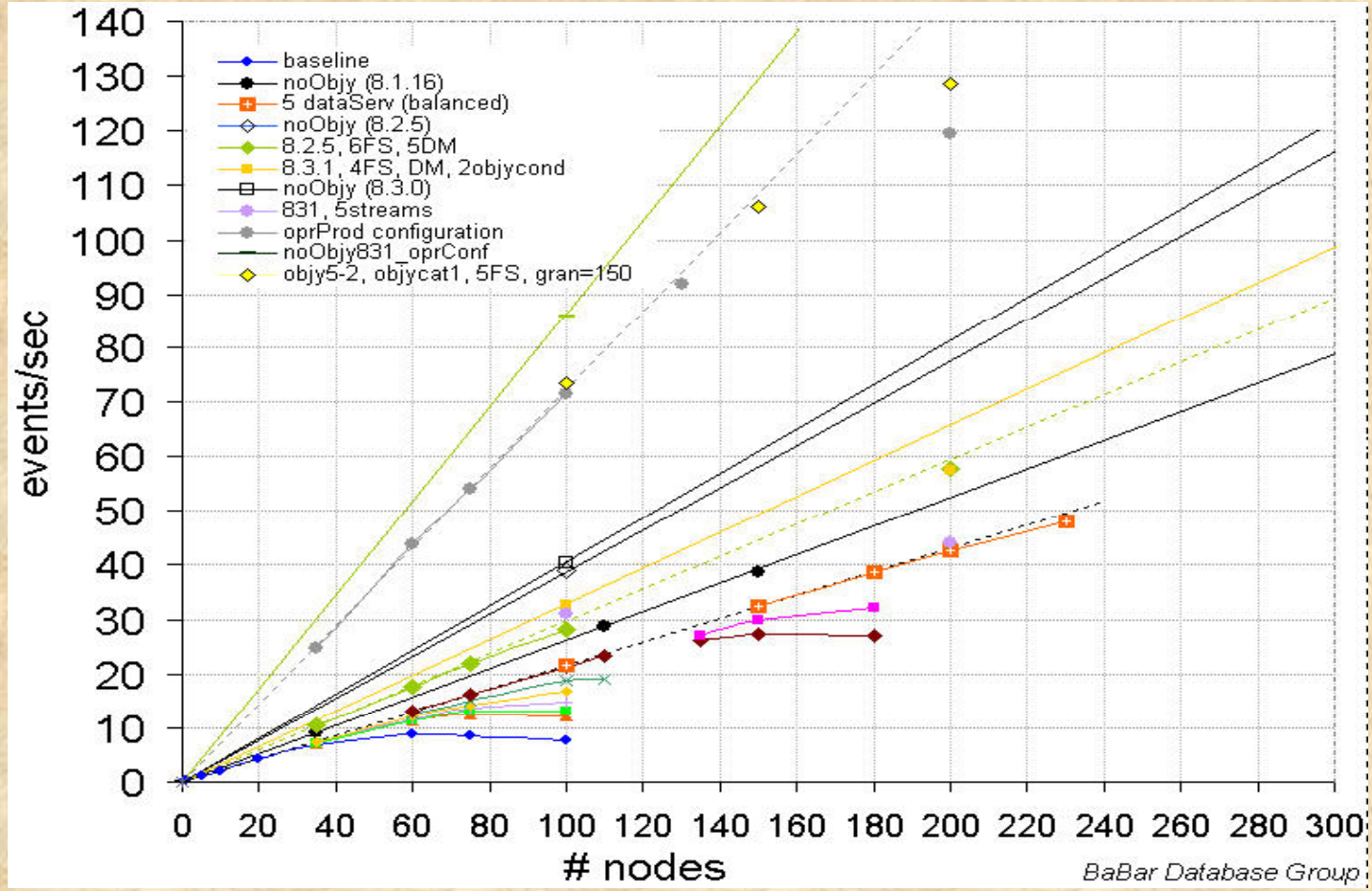
- Some key components had not even been designed!
  - Not one single line of “FATMEN” had been written by this stage!

💣 **Major changes in the computing environment were about to strike!**

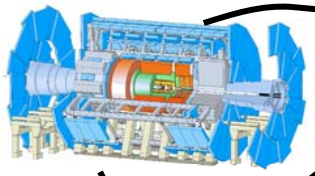
- We had just migrated ***to*** CERNVM – the Web was around the corner, as was distributed computing (SHIFT)
- (Not to mention OO & early LHC computing!)

☝ **HEPiX was two years from being born!**

Startup woes – BaBar experience



Data Handling and Computation for Physics Analysis



detector

event filter (selection & reconstruction)

reconstruction

raw data

event summary data

processed data

event reprocessing

batch physics analysis

analysis

analysis objects (extracted by physics topic)

interactive physics analysis

event simulation

simulation

LCG Service Hierarchy

Tier-0 – the accelerator centre

- Data acquisition & initial processing
- Long-term data curation
- ➢ **Data Distribution to Tier-1 centres**

Tier-1 – "online" to data acquisition process → high availability

- Managed Mass Storage – → grid-enabled data service
- ➢ **All re-processing passes**
- Data-heavy analysis
- National, regional support

Tier-2 – ~100 centres in ~40 countries

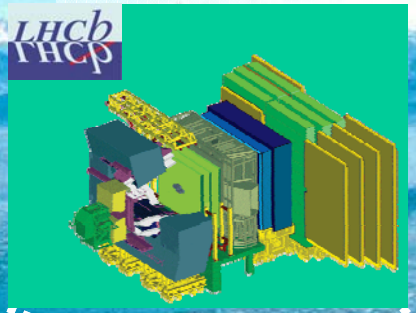
- Simulation
- End-user analysis – batch and interactive
- ➢ **Services, including Data Archive and Delivery, from Tier-1s**

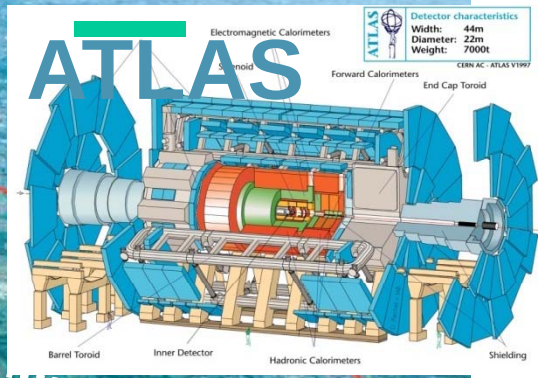
les.robertson@cern.ch

LHC: One Ring to Bind them…

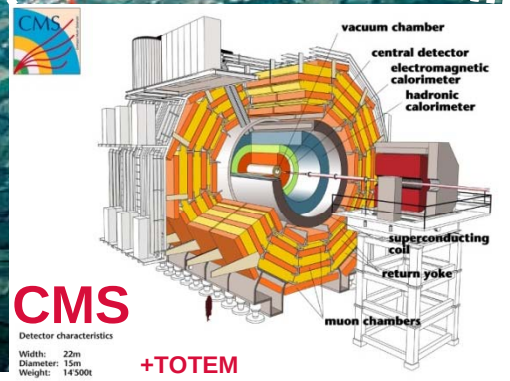
pp, B-Physics,
CP Violation

LHC : 27 km long
100m underground

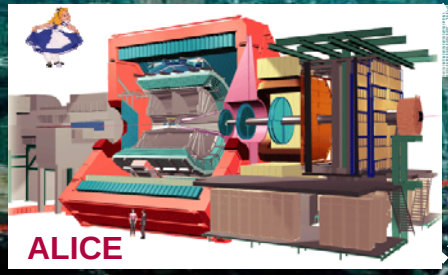




General Purpose,
pp, heavy ions

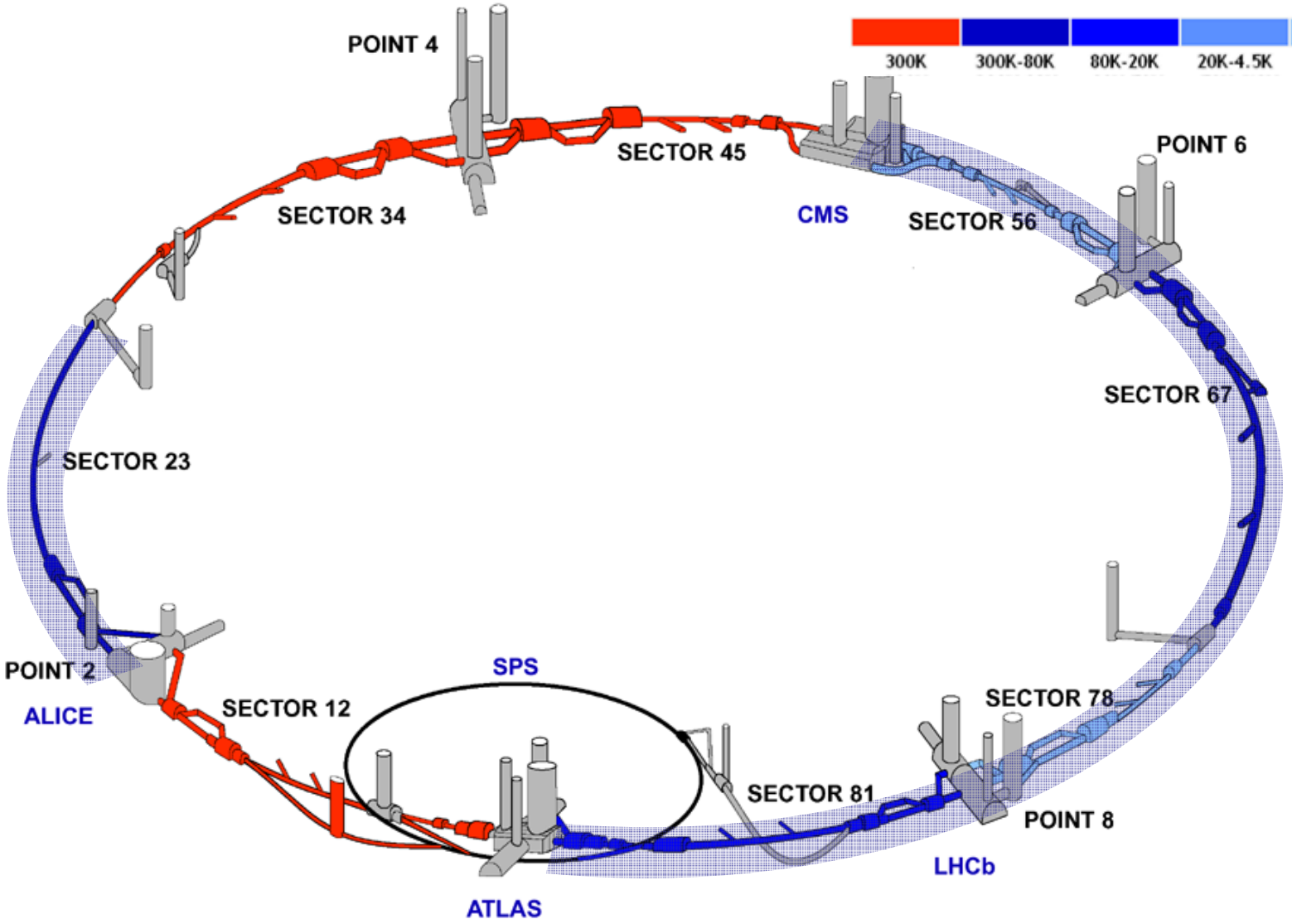


+TOTEM



Heavy ions, pp

Status of the LHC



LHC Computing is Complicated!

- Despite high-level diagrams (next), the Computing TDRs and other invaluable documents, it is very hard to maintain a complete view of all of the processes that form part of even one experiment's production chain
- **Both detailed views of the individual services, together with the high-level "WLCG" view are required…**
- It is ~impossible (for an individual) to focus on both…

☝ Need to work together as a team, sharing the necessary information, aggregating as required etc.

➢ **The needed information must be logged & accessible!**

- (Service interventions, changes etc.)

☝ **This is critical when offering a smooth service with affordable manpower**

Early enthusiasts discuss LHC Computing...

CCRC'08 – Motivation and Goals

What when:

- LHC is operating and experiments take data?
- All experiments want to use the computing infrastructure simultaneously?
- The data rates and volumes to be handled at the Tier0, the Tier1 and Tier2 centers are the sum of ALICE, ATLAS, CMS and LHCb as specified in the experiments computing model

- Each experiment has done data challenges, computing challenges, tests, dress rehearsals, ... at a schedule defined by the experiment

- This will stop: we will no longer be the master of our schedule... .... Once LHC starts to operate.

- We need to prepare for this ... together ....

A combined challenge by all Experiments should be used to demonstrate the readiness of the WLCG Computing infrastructure before start of data taking at a scale comparable to the data taking in 2008.

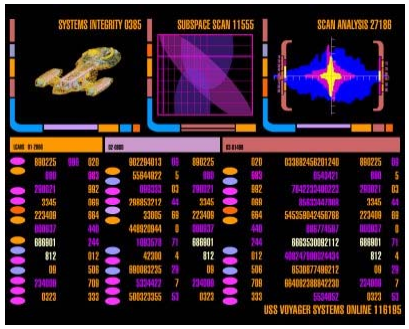
This should be done well in advance of the start of data taking on order to identify flaws, bottlenecks and allow to fix those.

We must do this challenge as WLCG collaboration: Centers and Experiments

COMPUTING CHALLENGES

So What is the Grid?

- According to Mike Vetterli, the new chair of the WLCG Collaboration Board, the Grid is
- "**The Star Trek Computer**"



- Another vision is that it should work as "**the Ultimate Application Accelerator**"

☹ **On the down-site, one must also consider the complications of a vast, distributed system…**

- Issues related to the deployment and operation of such a system must be carefully balanced against the benefits...

Why a Grid Solution?

- The **WLCG** Technical Design Report lists:

1. Significant costs of [providing] maintaining and upgrading the necessary resources … more easily handled in a distributed environment, **where individual institutes and … organisations can fund local resources … whilst contributing to the global goal**

2. … no single points of failure. Multiple copies of the data, automatic reassigning of tasks to resources… facilitates access to data for all scientists independent of location. … round the clock monitoring and support.

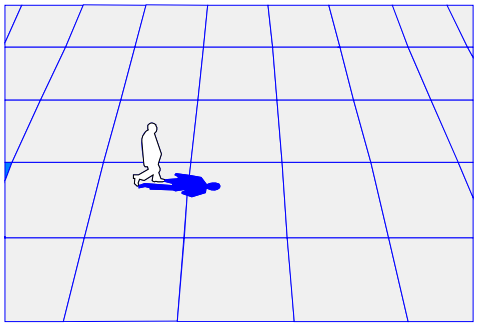
GRID CHALLENGES

Grid Computing in 3 Easy Steps

- Today there are many definitions of *Grid computing*:

- **"A distributed system is one in which the failure of a computer you didn't even know existed can render your own computer unusable."**

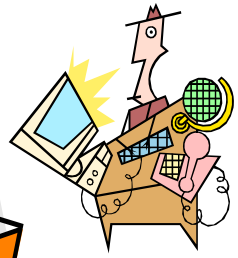
Leslie Lamport



➢ **… Some sort of Distributed System at least…**

  - that crosses Management / Enterprise domains

EDG Interfaces



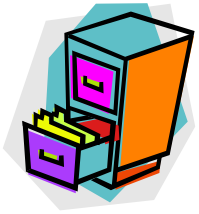
Application Developers

Local Application

Local Database

Scientists

Certificate Authorities



System Managers

File Systems

User Accounts

Grid Application Layer

Job Management | Data Management | Metadata Management | Object to File Mapping

Collective Services

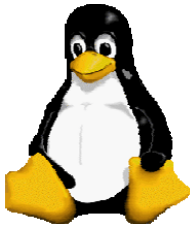
Information & Monitoring | Replica Manager | Grid Scheduler

Underlying Grid Services

SQL Database Services | Computing Element Services | Storage Element Services | Replica Catalog | Authorization Authentication and Accounting | Service Index

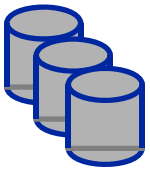
Fabric services

Resource Management | Configuration Management | Monitoring and Fault Tolerance | Node Installation & Management | Fabric Storage Management

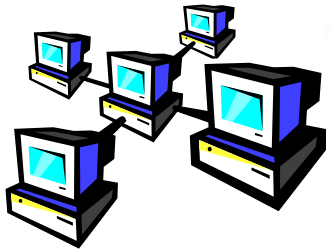


Operating Systems

Mass Storage Systems
HPSS, Castor



Storage Elements



Computing Elements

Batch Systems
PBS, LSF

Background

- 100% of my **Grid** experience relates to the deployment and delivery of **Production Services**
- This started already in the days of EDG with the Replica Location Service and its deployment at CERN and some key (**WLCG**) Tier1 sites
- In the then-current **WLCG** Computing Model, the EDG-RLS was a critical component which, if unavailable, meant:

➢ **Running jobs could not access existing data**

➢ **Scheduling of jobs at sites where the needed data was located was not possible**

☹ **The Grid – if not down – was at least seriously impaired…**

- **This was taken into account when designing the service deployment strategy & procedures – a taste of things to come!**

What's Special about "the Grid"

From the viewpoint of a "large consumer":

- ☺ **Grids have proven to be an excellent way of federating resources across computer centres of varying sizes into much larger quasi-homogeneous infrastructures.**
- ☺ **This matches well with the needs of international science, allowing resources at participating institutes to meet the needs of the entire collaboration.**
- ☺ **This in turn adds value to the individual sites, leading to a positive feedback situation.**

[Legendary] Grid Classification…

- Grid Computing (potentially) offers value to a wide range of applications, broadly classified as follows:

➢ **Provisioned**
  - Large scale, long term “Grand Challenge”
  - e.g. LHC (“space microscopes”), space telescopes, ….

➢ **Scheduled**
  - Require large resources for short periods
  - Far too expensive to provision for a single ‘application’
  - Not (always) time critical – disaster response?

➢ **Opportunistic**
  - Which includes the above but also other areas which are less “real time”

- You can find **numerous** examples of “Mission Critical” applications in each of these categories (e.g. EGEE User Forum!)
  - “Mission Critical” as in “Life or Death”

WLCG CHALLENGES

CCRC'08 – Objectives

- Primary objective (**next**) was to demonstrate that we (sites, experiments) could run together at 2008 production scale
  - **This includes testing all "functional blocks":**
    - **Experiment to CERN MSS; CERN to Tier1; Tier1 to Tier2s etc.**

- Two challenge phases were foreseen:
  1. **February :** not all 2008 resources in place – still adapting to new versions of some services (e.g. SRM v2.2) & experiment s/w
  2. **May:** all 2008 resources in place – full 2008 workload, all aspects of experiments' production chains

💣 N.B. failure to meet target(s) would necessarily result in discussions on **de-scoping!**

- Fortunately, the results suggest that this is not needed, although **much** still needs to be done before, and during, May!

Service Availability Targets

- The WLCG Memorandum of Understanding defines:
  - **The services that a given site must provide (Tier0, Tier1, Tier2);**
  - **The availability of these services (measured on an annual basis);**
  - **The maximum time to intervene in case of problems.**

- Taken together, these service availability targets are somewhat aggressive and range from 95% to 99% for **compound** services, e.g.
  - **Acceptance of raw data from Tier0**
  - **Data-intensive analysis services, including networking to Tier0**

- Such 'services' involve many sub-services, e.g. storage services, catalog and metadata services, DB services, experiment-specific services etc.

- Major concerns include both **scheduled** and unscheduled interventions – must design all elements of the service correspondingly
  - **Hardware configuration; procedures & documentation; middleware**

WLCG Tier1 Services[1]

i. acceptance of an agreed share of raw data from the Tier0 Centre, keeping up with data acquisition;

ii. acceptance of an agreed share of first-pass reconstructed data from the Tier0 Centre;

iii. acceptance of processed and simulated data from other centres of the WLCG;

iv. recording and archival storage of the accepted share of raw data (distributed back-up);

v. recording and maintenance of processed and simulated data on permanent mass storage;

vi. provision of managed disk storage providing permanent and temporary data storage for files and databases;

vii. provision of access to the stored data by other centres of the WLCG and by named AF's as defined in paragraph X of this MoU;

viii. operation of a data-intensive analysis facility;

ix. provision of other services according to agreed Experiment requirements;

x. ensure high-capacity network bandwidth and services for data exchange with the Tier0 Centre, as part of an overall plan agreed amongst the Experiments, Tier1 and Tier0 Centres;

xi. ensure network bandwidth and services for data exchange with Tier1 and Tier2 Centres, as part of an overall plan agreed amongst the Experiments, Tier1 and Tier2 Centres;

xii. administration of databases required by Experiments at Tier1 Centres.

- All storage and computational services shall be "grid enabled" according to standards agreed between the LHC Experiments and the regional centres.

[1] WLCG Memorandum of Understanding (signed by each T0/T1/T2)

Problem Response Time and Availability targets Tier-1 Centres

Service	*Maximum delay in responding to operational problems (hours)*			*Availability*
	Service interruption	*Degradation of the service*		
		> 50%	*> 20%*	
Acceptance of data from the Tier-0 Centre during accelerator operation	12	12	24	99%
Other essential services – prime service hours	2	2	4	98%
Other essential services – outside prime service hours	24	48	48	97%

les.robertson@cern

How We Measured Our Success

- Agreed up-front on specific targets and metrics – these were 3-fold and helped integrate different aspects of the service (CCRC'08 wiki):
  1. Explicit "**scaling factors**" set by the experiments for each functional block: discussed in detail together with sites to ensure that the necessary resources and configuration were in place;
  2. Targets for the lists of "**critical services**" defined by the experiments – those essential for their production, with an analysis of the impact of service degradation or interruption **(WLCG Design, Implementation & Deployment standards)**
  3. WLCG "**Memorandum of Understanding**" (MoU) targets – services to be provided by sites, target availability, time to intervene / resolve problems ...

☝ **Clearly some rationalization of these would be useful – significant but not complete overlap**

LESSONS LEARNED

Main Lessons Learned

☺ **Generally, things worked reasonably well…**

➢ **Still improvements in communication are needed!**

- Tools still need to be streamlined (e.g. elog-books / GGUS), and reporting automated
  - ¿ **Service dashboards – should be in place before May…**
- F2Fs and other meetings working well in this direction!

- **Pre-established metrics extremely valuable!**
  - As well as careful preparation and extensive communication!

Now in continuous production mode – this will continue – as will today's infrastructure & meetings

What Did We Achieve? (High Level)

- Even **before** the official start date of the February challenge, it had proven an extremely **useful** focusing exercise, in helping understand missing and / or weak aspects of the service and in identifying **pragmatic** solutions
- Although later than desirable, the **main bugs** in the middleware were fixed (just) in time and many sites upgraded to these versions
- The deployment, configuration and usage of SRM v2.2 went **better** than had predicted, with a noticeable **improvement** during the month
- Despite the high workload, we also **demonstrated** (most importantly) that **we can support** this work with the available manpower, although essentially no remaining effort for longer-term work (**more later…**)
- If we can do the same in May – when the bar is placed much higher – we will be in a **good position** for this year's data taking
- However, there are certainly **significant concerns** around the available manpower at all sites – not only today, but also in the longer term, when funding is unclear (**e.g. post EGEE III**)

Recommendations – mid-February!

- ✓ **To improve communications with Tier2s and the DB community, 2 new mailing lists have been setup, as well as regular con-calls with Asia-Pacific sites (time zones…)**
- Follow-up on the lists of "**Critical Services**" must continue, implementing not only the appropriate monitoring, but also ensuring that the WLCG "standards" are followed for
  Design, **I**mplementation, **D**eployment and **O**peration
  - **We are having some (small) success in convincing others in the Grid….**
- **Clarify reporting and problem escalation lines (e.g. operator call-out triggered by named experts, …) and introduce (light-weight) post-mortems when MoU targets not met**
- ➢ **We must continue to improve on open & transparent reporting, as well as further automations in monitoring, logging & accounting**

- **We should foresee "data taking readiness" challenges in future years – probably with a similar schedule to this year – to ensure that full chain (new resources, new versions of experiment + AA s/w, middleware, storage-ware) is ready [maybe also July 2008!]**

Timeline

- In 2010, the LHC will reach design luminosity
- In 2010, EGEE III will terminate
- It is inconceivable that we:
  a. Don't run the LHC machine
  b. Run the LHC machine without a computing infrastructure
  c. Run the computing infrastructure without Grid operations

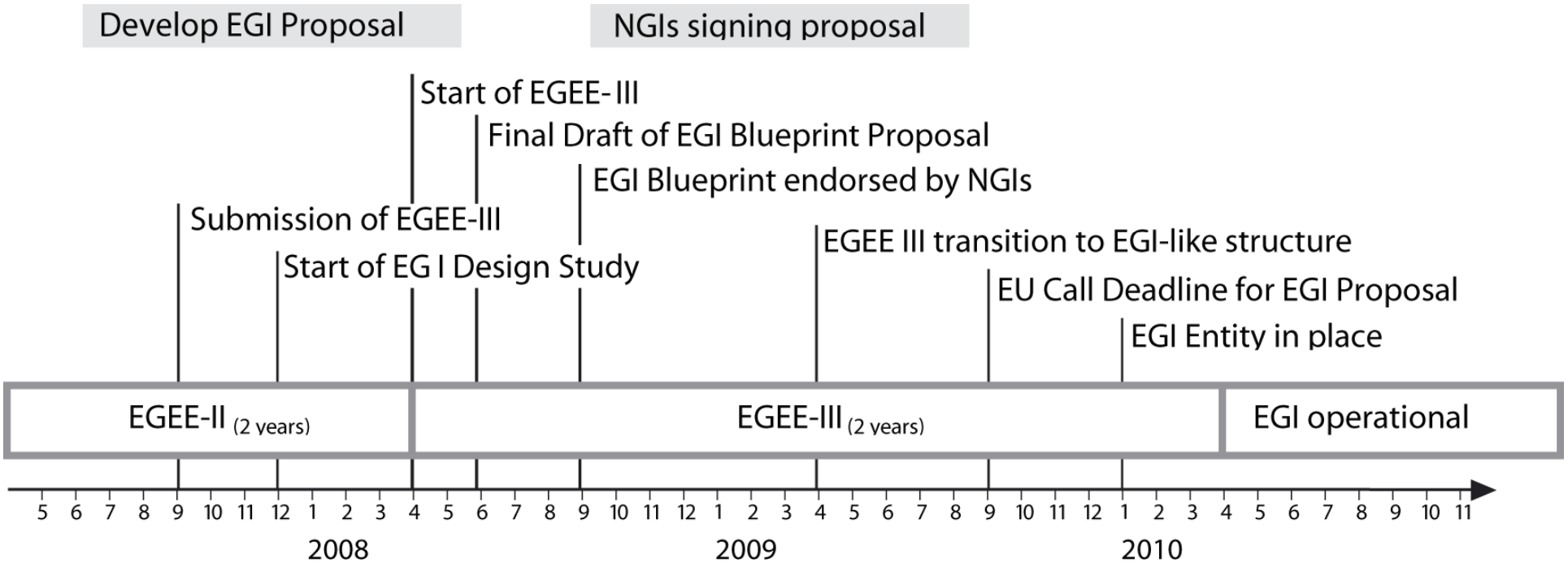
➢ This is required for other mission critical applications that are dependant on this infrastructure

☝ The transition to the new scenario must be
  a. On time
  b. Non-disruptive

- This is a fundamental requirement – it is not an issue for discussion

From the EGI_DS DoW…

- **The establishment of EGI is guided by two basic principles:**
  1. **Build on the experience and successful operation of EGEE and related projects**
  2. **Make EGI operational before EGEE III ends**



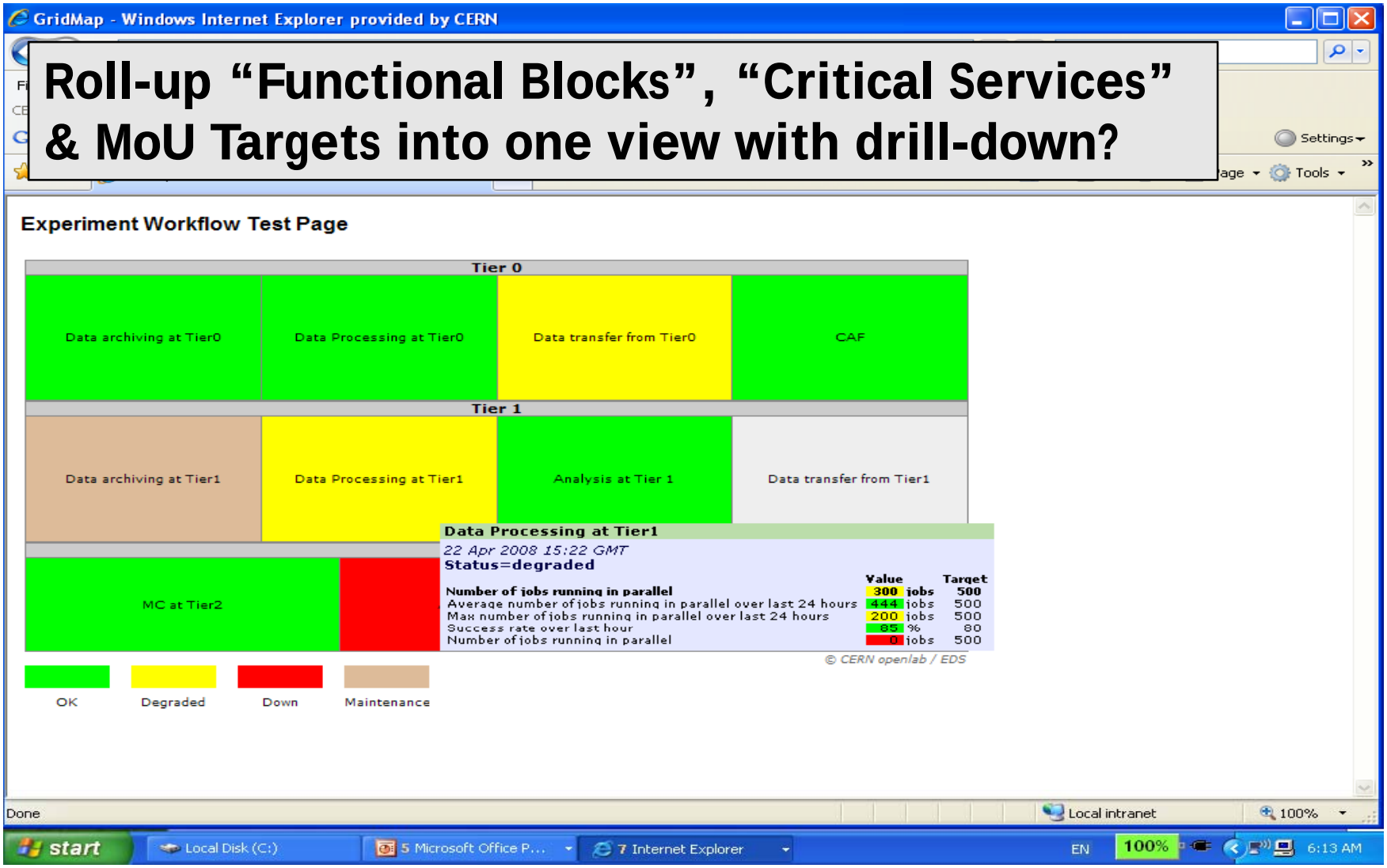
Critical Service Follow-up

- 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 **/ alarm e-mail**	99%
1 hour	Operator response to alarm / call to x5011 **/ alarm e-mail**	100%
4 hours	Expert intervention in response to above	95%
8 hours	Problem resolved	90%
24 hours	Problem resolved	99%

Gridmap for monitoring of the workflows of the Experiments



Outlook – 2008

- **Stable baseline service with (small) incremental improvements throughout the year**
- Detailed post-mortem of May CCRC'08 in June (12-13)
- Still need to converge on additional service issues:
  - SRM v2.2 MoU addendum, monitoring improvements, better use of storage resources (file size, batching of requests, ...)

➢ **It is inevitable that both the May run and first pp data taking will bring with them new issues – and solutions!**

- Continued (late) deployment of 2008 resources over several months
- A further CCRC'08 phase in July should be expected to test these resources and any further service enhancements

☺ **First pp data taking in 2nd half of the year…**

Outlook – Beyond

- **We can still expect some significant increases in resources, plus changes in some of the services / experiment-ware, for the foreseeable future**
- Some examples of possible changes:
  - Computing Models – based on real-life experience;
  - LCG OPN (recent proposal rejected);
  - Operations infrastructure – move from EGEE III to EGI(?)
  - ...
- **Expecting everything to work without a large-scale test of all experiments and all sites is contrary to our experience**
- **As previously discussed (February) a "readiness challenge" for each year's data taking should be foreseen**
- And scheduled to be consistent with needs of reprocessing!

The WLCG Experience

- There can be **no** doubt that the close physical proximity / inter-mingling people from the different projects (LCG, EGEE, related, …) has been **extremely** beneficial during the deployment and hardening phases of **WLCG**
- This is clearly not **scalable** to large numbers of application communities and may well be in **contradiction** with a **"sustainable (long-term?) e-instructure"**

CCRC'08 – Conclusions

- The **WLCG** service is running (reasonably) **smoothly**
- The functionality **matches** what has been tested so far – and what is (known to be) required

☛ **We have a good baseline on which to build**

- (Big) **improvements** over the past year are a good indication of what can be expected over the next!
- (Very) detailed **analysis** of results compared to up-front metrics – in particular from experiments!

The Grid

The Power of 3