LHCC Comprehensive Review

SC3 Report
& SC4 Plans

Jamie Shiers, LCG Service Manager

Agenda

- Goals of SC3

- Preparation Steps

- Experience of Throughput Tests

- Lessons Regarding LCG Production Services

- SC4 Planning

- Summary

SC3 Goals

- Much more than just a throughput test!

- More data management services:
    - SRM required at all sites
    - Reliable File Transfer Service based on gLite FTS
    - LFC file catalog deployed as required
        - Global catalog for LHCb, site local for ALICE + ATLAS
        - Some enhancements agreed for CMS (essentially identical to LHCb requests…)
        - (Ratified through Task Forces)
    - Other services as per BSWG

- More sites:
    - All Tier1s took part – this was better than foreseen!
    - Many Tier2s – now above 20 sites, covering most regions. This too is working well!

- All experiments:
    - Clear goals established together with metrics for measuring success
    - List of issues will be summarised by Nick Brook – many issues already resolved

- Throughput targets:
    - 50% higher than SC2 but using SRM & FTS as above (150MB/s to disk at T1s)
    - 60MB/s to tape at Tier1s (following disk – disk tests)
    - Modest T2->T1 targets, representing MC upload (3 x 1GB file / hour)

SC3 Service Summary

- Services identified through combination of Baseline Services Working Group, Storage Management Workshop and 1-1 discussions with experiments
    - Timeline of BSWG & service setup lead time did not allow to wait for 'final report' before starting to implement services
- For new services (LFC, FTS), two flavours established at CERN
    - 'Pilot' – to allow experiments to gain experience with functionality, adapt their s/w to interfaces etc.
    - 'SC3' – full production services
    - ➢ This separation proved useful!
- New services for sites: LFC (most sites), FTS (T1s), SRM (DPM, dCache at T2s)
- Support lists established for these services, plus global 'catch-call'
    - Clear that this needs to be re-worked as we move to WLCG pilot
        - ▫ A proposal on this later…
- 'SC3' services being re-deployed for full production
    - Some of this work was done during end-Oct / early Nov intervention
- List of Services by site will be covered in SC4 planning presentation

SC3 Throughput Tests

➢ Unfortunately, July Throughput Tests did not meet targets

▪ Compounded by service instability

▪ Continued through to the end, i.e. disk – disk, disk – tape and T2 – T1 components

▪ Spent most of August debugging and fixing

▪ dCache workshop held in DESY identified concrete actions / configurations / dCache improvements

▪ Improvements also in CASTOR SRM & gLite FTS

▪ All software upgrades now released – deployment to be scheduled

➢ Disk – disk rates obtained in July around 1/2 target, without stability!

SC3 Throughput: Disk & Tape

**Disk target:
150MB/s/site
1GB/s (CERN)**
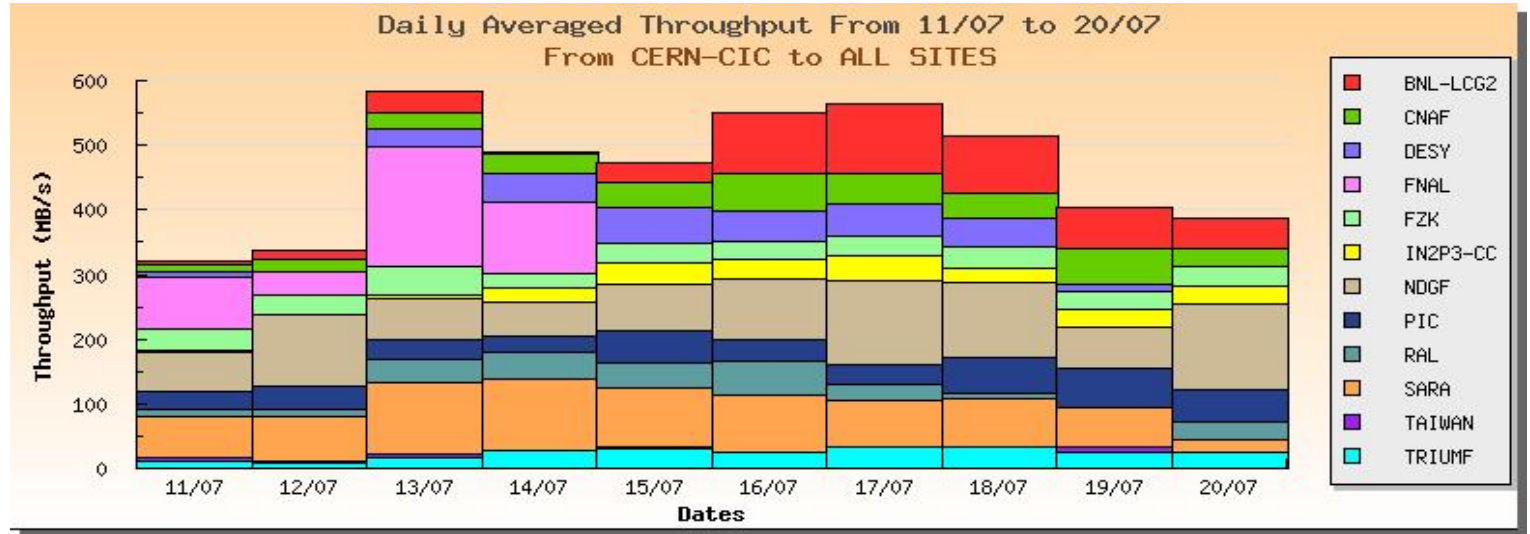
**Tape target:
60MB/s/site
(Not all sites)**

James Casey
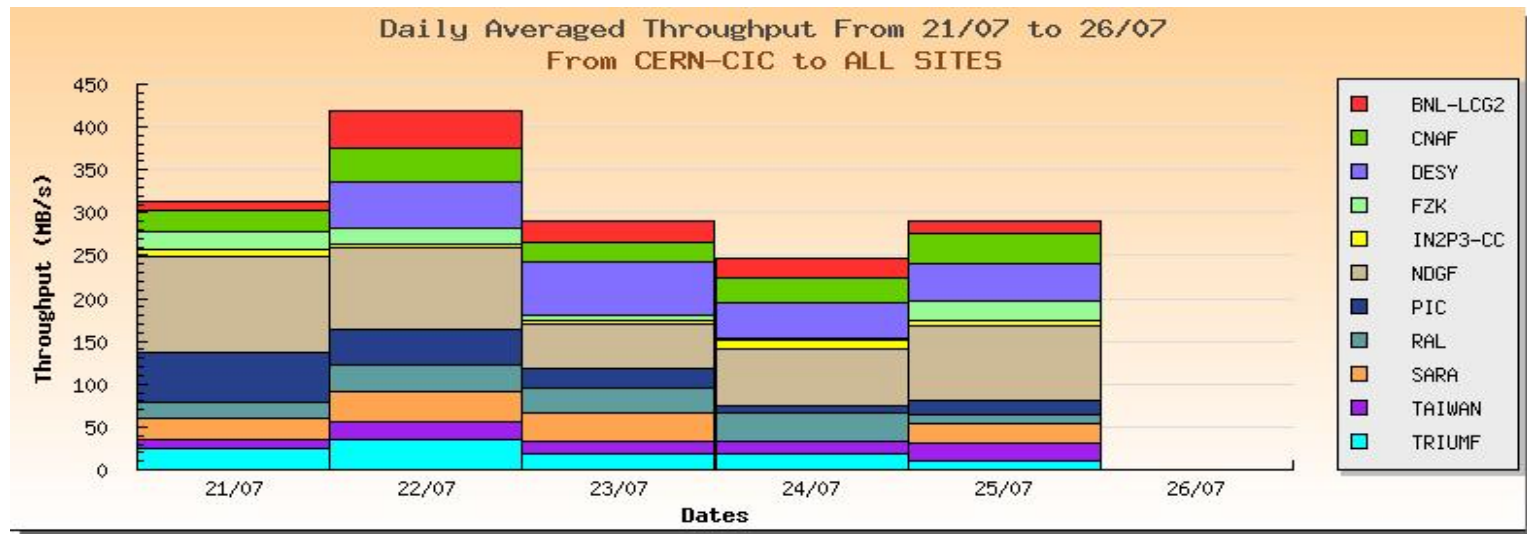
Results of SC3 in terms of Transfers

- Target data rates 50% higher than during SC2
- All T1s (most supporting T2s) participated in this challenge
- Transfers between SRMs (not the case in SC1/2)
- Important step to gain experience with the services before SC4

| Site	MoU Target (Tape)	Daily average MB/s (Disk)
ASGC	100	10
BNL	200	107
FNAL	200	185
GridKa	200	42
CC-IN2P3	200	40
CNAF	200	50
NDGF	50	129
PIC	100	54
RAL	150	52
SARA/NIKHEF	150	111
TRIUMF	50	34

Rates during July throughput tests. Better single-site rates since, but need to rerun tests…

For this we need dCache 1.6.6(+) to be released/deployed, latest FTS (now), network upgrades etc.

January '06 (<CHEP)

Post-debugging

- Now can achieve same rate as before with fewer sites
  - Still need to add in other sites, and see what the new upper limit is
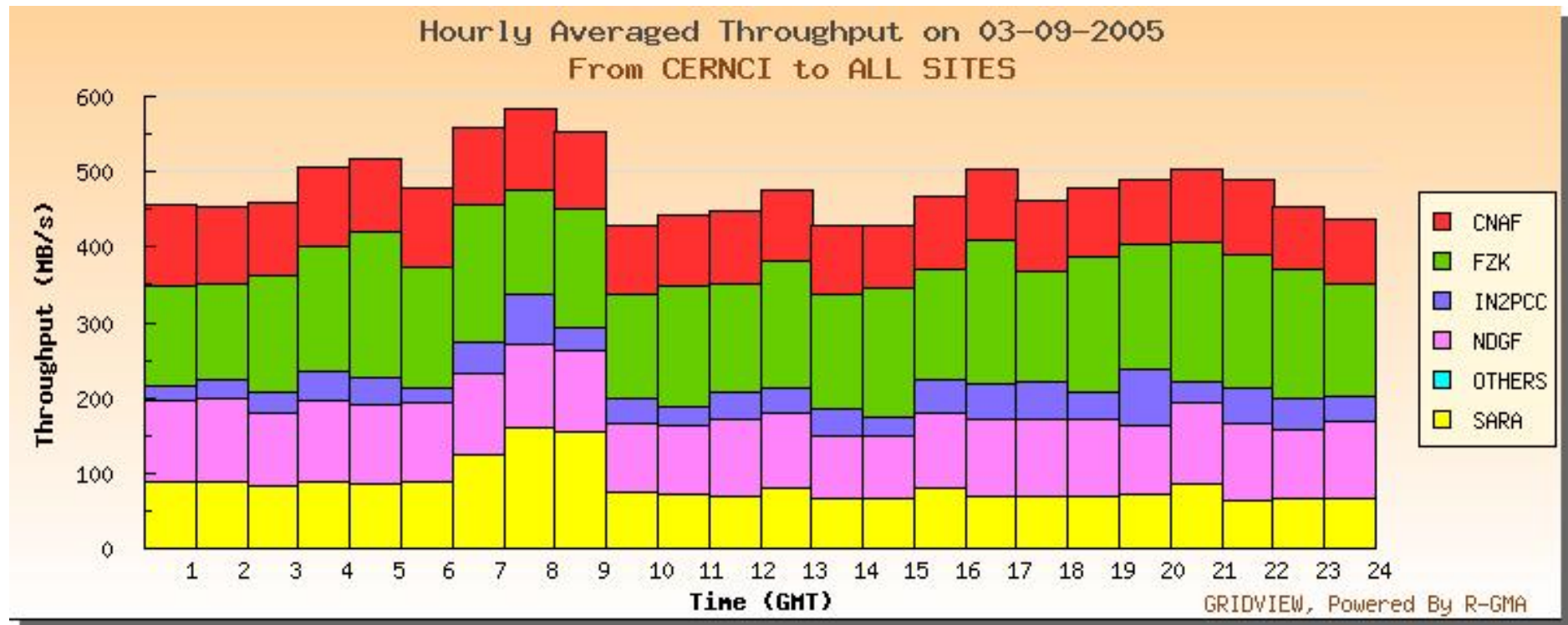
Pre-Requisites for Re-Run of Throughput Tests

- Deployment of gLite FTS 1.4 (srmcp support)
  - ✓ Done at CERN in recent intervention

- dCache 1.6.6 (or later) release and deployed at all dCache sites.
  - ✓ Released – some sites already planning upgrade

- CASTOR2 clients and CASTORSRM version 2.2.8 (or later) at all CASTOR sites (ASGC, CNAF, PIC).

- Upgrade to CERN internal network infrastructure.
  - Partly done – remainder during Christmas shutdown
  - ➢ N.B. intend to keep Grid running over Xmas! (Close to last chance…)

- 10Gbit/s network connections at operational at the following sites:
  - IN2P3, GridKA, CNAF, NIKHEF/SARA, BNL, FNAL

dCache – the Upgrade (CHEP 2006)

- *For the last two years, the dCache/SRM Storage Element has been successfully integrated into the LCG framework and is in heavy production at several dozens of sites, spanning a range from single host installations up to those with some hundreds of TB of disk space, delivering more than 50 TB per day to clients. Based on the permanent feedback from our users and the detailed reports given by representatives of large dCache sites during our workshop at DESY end of August 2005, the dCache team has been identified important areas of improvement.*

- *This includes a more sophisticated handling of the various supported tape back-ends, the introduction of multiple I/O queues per pool with different properties to account for the diverse behaviours of the different I/O protocols and the possibility to have one dCache instance spread over more than one physical site.*

- *… changes in the name-space management as short and long term perspective to keep up with future requirements.*

- *… initiative to make dCache a widely scalable storage element by introducing dCache, the Book, plans for improved packaging and more convenient source code license terms.*

- *Finally I would like to cover the dCache part of the German e-science project, d-Grid, which will allow for improved scheduling of tape to disk restore operations as well as advanced job scheduling by providing extended information exchange between storage elements and Job Scheduler.*

Disk – Disk Rates (SC3 Repeat)

Centre	ALICE	ATLAS	CMS	LHCb	Target Data Rate MBytes/sec
Canada, TRIUMF		X			50
France, CC-IN2P3	X	X	X	X	150
Germany, GridKA	X	X	X	X	150
Italy, CNAF	X	X	X	X	150
Netherlands, NIKHEF/SARA	X	X		X	150
Nordic Data Grid Facility	X	X	X		50
Spain, PIC Barcelona		X	X	X	100
Taipei, ASGC		X	X		100
UK, RAL	X	X	X	X	150
USA, BNL		X			150
USA, FNAL			X		150

- These are the nominal data rates capped at 150MB/s

SC3 Summary

- Underlined the complexity of reliable, high rate sustained transfers to be met during LHC data taking

- Many issues understood and resolved – need to confirm by re-run of Throughput exercise

- We are now well into the Service Phase (Sep – Dec)

- Collaboration with sites & experiments has been excellent

- We are continuing to address problems as they are raised

- Together with preparing for SC4 and WLCG pilot / production

- The experiment view will be presented later by Nick Brook…

End of Part I

SC4 Introduction

- Many additional Use Cases to be covered
  - Partial list next… Full list to be established by CHEP, using workshop…

- Data rates go up to full nominal rates
  - Disk – Disk in April; Disk – Tape in July

- Additional Tier2 sites
  - Target of 20 in April; 40 in September

- Service Level targets as per MoU

- Service Level Monitoring

- Stream-line Operations and User Support

- Step by step planning – write things down as they become clear / agreed!

SC4 Planning – Step by Step

- Initial goal was to have a workshop Sep / Oct 2005
  - Discussed at June workshop and July PEB / GDB

- Insufficient response to go ahead – retreat to CHEP w/s

- Planning documents covering: (attached to agenda page)
  - MoU responsibilities & target data rates
  - Services & Service levels
  - Throughput testing – focusing on initial data export

- Others will be written as things become clear
  - SC4 covers *all* offline Use Cases of the experiments

- See list of Use Cases for discussion at CHEP
  - As much to be documented / explained up-front as possible

SC4 Use Cases – Establish by CHEP!

Not covered so far in Service Challenges:

- T0 recording to tape (and then out)

- Reprocessing at T1s

- Calibrations & distribution of calibration data

- HEPCAL II Use Cases

- Individual (mini-) productions (if / as allowed)

Additional services to be included:

- Full VOMS integration

- COOL, other AA services, experiment-specific services (e.g. ATLAS HVS)

- PROOF, xrootd, … (analysis services in general…)

- Testing of next generation IBM and STK tape drives

SC4 Workshop:

1. *Data Management*
2. *Grid Services*
3. *Expt. Use Cases*

Iterating on agenda and attendance.

From July PEB/GDB

SC4 Preparation

- The main technical problems and how we plan to address them

- Identifying Additional Use Cases

- Service Coordination, Operation and User Support

WLCG – Major Challenges Ahead

1. Get data rates at all Tier1s up to MoU Values
   - Stable, reliable, rock-solid services
   - **We are currently about 1/2 the target level, without including tape**

2. Re-deploy Required Services at Sites to meet MoU Targets
   - Measured, delivered Availability; maximum intervention time etc.
   - Ensure that the services provided match the experiments' requirements

➢ **T0 and T1 services are tightly coupled!**
   - Particularly during accelerator operation

- Need to build strong collaborative spirit to be able to deliver required level of services
  - And survive the inevitable 'crises'…

- (These are not the only issues – just the top two!)

SC4 – Transfer Rate Plan

- Split Transfer Tests into Separate Steps

  1. Rerun of SC3 Throughput in January 2006
  2. Tier0 – Tier1 "loop-back" tests to new tape h/w by March 2006
     - Target is twice maximum nominal rate, i.e. 400MB/s
  3. Tier0 – Tier1 transfers at full nominal rate (disk – disk) April 2006
  4. Tier0 – Tier1 transfers scaled to current h/w (disk – tape) April 2006
  5. Tier0 – Tier1 transfers at full nominal rate (disk – tape) July 2006
     - Needs to be coordinated with site acquisition plans

- Identify additional data flows & rates and establish corresponding milestones

  - There is already material on this in the TDRs and in a number of presentations by the experiments
  - Need to clearly explain these together with Tier1s / Tier2s
    - Sites often have 'local' experts!
  - Pre-CHEP workshop has one day dedicated to this! (10 – 12 February, Mumbai)

- We are also working proactively with the sites on Throughput issues
  - Using all available opportunities! e.g. FZK workshop, GridPP15, 1er Colloq.FR

Disk – Disk Rates (SC4 part 1)

Centre	ALICE	ATLAS	CMS	LHCb	Rate into T1 (pp) MB/s
ASGC, Taipei	-	8%	10%	-	100
CNAF, Italy	7%	7%	13%	11%	200
PIC, Spain	-	5%	5%	6.5%	100
IN2P3, Lyon	9%	13%	10%	27%	200
GridKA, Germany	20%	10%	8%	10%	200
RAL, UK	-	7%	3%	15%	150
BNL, USA	-	22%	-	-	200
FNAL, USA	-	-	28%	-	200
TRIUMF, Canada	-	4%	-	-	50
NIKHEF/SARA, NL	3%	13%	-	23%	150
Nordic Data Grid Facility	6%	6%	-	-	50
Totals	-	-	-	-	1,600

These are the nominal values based on Computing TDRs with rates weighted by agreed resource allocation / VO.

April 2006

Disk – Tape Rates (SC4 part 1)

Centre	ALICE	ATLAS	CMS	LHCb	Target Data Rate MB/s
Canada, TRIUMF		X			50
France, CC-IN2P3	X	X	X	X	75
Germany, GridKA	X	X	X	X	75
Italy, CNAF	X	X	X	X	75
Netherlands, NIKHEF/SARA	X	X		X	75
Nordic Data Grid Facility	X	X	X		50
Spain, PIC Barcelona		X	X	X	75
Taipei, ASGC		X	X		75
UK, RAL	X	X	X	X	75
USA, BNL		X			75
USA, FNAL			X		75

■ Still using SRM 1.1 & Current Tape Technology?

Disk – Tape Rates (SC4 part 2)

Centre	ALICE	ATLAS	CMS	LHCb	Rate into T1 (pp) MB/s
ASGC, Taipei	-	8%	10%	-	100
CNAF, Italy	7%	7%	13%	11%	200
PIC, Spain	-	5%	5%	6.5%	100
IN2P3, Lyon	9%	13%	10%	27%	200
GridKA, Germany	20%	10%	8%	10%	200
RAL, UK	-	7%	3%	15%	150
BNL, USA	-	22%	-	-	200
FNAL, USA	-	-	28%	-	200
TRIUMF, Canada	-	4%	-	-	50
NIKHEF/SARA, NL	3%	13%	-	23%	150
Nordic Data Grid Facility	6%	6%	-	-	50
Totals	-	-	-	-	1,600

Have to ramp up to twice this rate prior to April 2007! (See LCG TDR).

July 2006

WLCG – Major Challenges Ahead

1. Get data rates at all Tier1s up to MoU Values
    - Stable, reliable, rock-solid services
    - **We are currently about 1/2 the target level, without including tape**

2. (Re-)deploy Required Services at Sites to meet MoU Targets
    - Measured, delivered Availability, maximum intervention time etc.
    - Ensure that the services delivered match the experiments' requirements

➢ **T0 and T1 services are tightly coupled!**
    - Particularly during accelerator operation

▪ Need to build strong collaborative spirit to be able to deliver required level of services
    - And survive the inevitable 'crises'…

Site Components - Updated

- **Each T1 to provide 10Gb network link to CERN**
- **Each site to provide SRM 1.1 interface to managed storage**
  - All sites involved in SC3: T0, T1s, T2s.

- T0 to provide File Transfer Service; also at named T1s for T2-T1 transfer tests
  - Named Tier1s: BNL, CNAF, FZK, RAL; Others also setting up FTS
  - CMS T2s being supported by a number of T1s using PhEDEx

- LCG File Catalog – not involved in Throughput but **needed for Service**
  - ALICE / ATLAS: site local catalog
  - LHCb: central catalog with >1 R/O 'copies' (on ~October timescale)
    - IN2P3 to host one copy; CNAF? Taiwan? RAL?
  - CMS: evaluating different catalogs
    - FNAL: Globus RLS, T0+other T1s: LFC; T2s: POOL MySQL, GRLS, …

- T2s – many more than foreseen
  - Running DPM or dCache, depending on T1 / local preferences / support
  - [Support load at CERN through DPM / LFC / FTS client]

- **Work still needed to have these consistently available as services**

Services & Service Levels

- List of services that need to be provided by each site is now clear
    - Including any VO-specific variations…

- For SC4 / pilot WLCG none of these services are new
    - Expect to see some analysis-oriented services coming later…
    - Maybe prototyped at some 'volunteer' T2s, e.g. DESY, CALTECH, Padua, .. ?

- The service list at CERN has been classified based on impact of service degradation / unavailability
    - Draft classification for Tier1s and Tier2s also exists & sent to GDB (August)

- A check-list has been produced and the Critical Services are being re-deployed target end-2005
    - Must provide operator procedures, support contacts etc etc

- We will measure service availability at all sites and report regularly
    - Results visible through Web used for daily operations purposes

Service Level Definitions

Class	Description	Downtime	Reduced	Degraded	Availability
C	Critical	1 hour	1 hour	4 hours	99%
H	High	4 hours	6 hours	6 hours	99%
M	Medium	6 hours	6 hours	12 hours	99%
L	Low	12 hours	24 hours	48 hours	98%
U	Unmanaged	None	None	None	None

Tier0 services: C/H, Tier1 services: H/M, Tier2 services M/L

Service	Maximum delay in responding to operational problems			Average availability measured on an annual basis	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outside accelerator operation	24 hours	48 hours	48 hours	n/a	98%

Tier0 Services

Service	VOs	Class
SRM 2.1	All VOs	C
LFC	LHCb	C
LFC	ALICE, ATLAS	H
FTS	ALICE, ATLAS, LHCb, (CMS)	C
CE	All VOs	C
RB		C
Global BDII		C
Site BDII		H
Myproxy		C
VOMS		H→C
R-GMA		H

Services at CERN

- Building on **'standard service mode**

1. First level support: operations tea
   - Box-level monitoring, reboot, ala

2. Second level support team: Grid D
   - Alerted by operators and/or alar
   - Follow 'smoke-tests' for applicati
   - Identify appropriate 3rd level sup
   - Responsible for maintaining and i
   - Two people per week: complementary to Service Manager on Duty
   - Provide daily report to SC meeting (09:00); interact with experiments
   - Members: IT-GD-EIS, IT-GD-SC
   - Phone numbers: 164111; 164222

3. Third level support teams: by service
   - Notified by **2nd level** and / or through operators (by agreement)
   - Should be called (very) rarely… **(Definition of a service?)**

*Big on-going effort
in this area:*
- *Services being reimplemented*
- *Merge of daily OPS meetings*
- *Service Coordination meetings*
- *Con-calls with sites*
- *Workshops*
- *etc.*

- *Goal is all Critical Services ready by Christmas*
- *(This means essentially all…)*

TierO Service Dashboard

An evaluation for each product within the four primary task areas:

1. <u>Requirements</u> – covers the infrastructure requirements with regard to machines, disks, network;

2. <u>Development</u> – covers from software creation and documentation to certification and delivery to the installation teams;

3. <u>Hardware</u> – covers the procurement, delivery, burn in, physical installation and base operating systems;

4. <u>Operations</u> – covers the administration, monitoring, configuration and backup of the service to the levels requested.

Tim Bell

Operations Checklist

- 2nd level support organisation defined (who to call when there is a problem with the application or middleware)
- Mechanism to contact 2nd level organisation
- Response time for 2nd level organisation
- List of machines where service is running defined
- List of configuration parameters and their values for the software components
- List of processes to monitor
- List of file systems and their emergency thresholds for alarms
- Application status check script requirements defined
- Definition of scheduled processes (e.g. cron)
- Test environment defined and available
- Problem determination procedures including how to determine application vs middleware vs database issues
- Procedures for start/stop/drain/check status defined
- Automatic monitoring of the application in place
- Backup procedures defined and tested

Tier0 Service Coordination

- Progress on re-implementing services monitored at fortnightly LCG Service Coordination Meeting
  - http://agenda.cern.ch/displayLevel.php?fid=654

- Area updates provided by area coordinators on Wiki prior to meeting

- Meeting remains crisp, focussed and short
  - Typically less than one hour…

- Target is to get all Critical services re-implemented by year-end

Tier0 Services

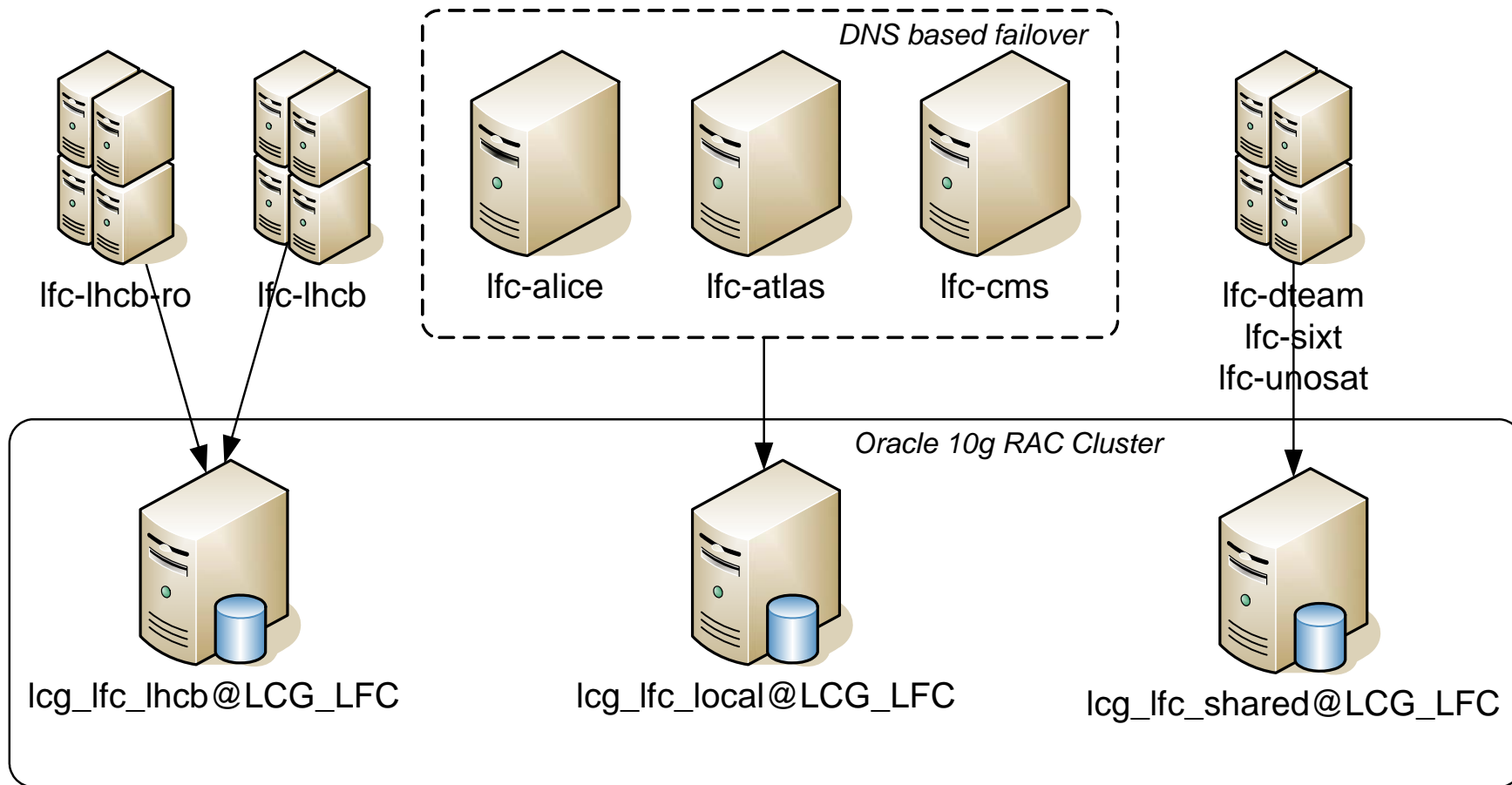| Service	VOs	Class
SRM 2.1	All VOs	C
LFC	LHCb	C
LFC	ALICE, ATLAS	H
FTS	ALICE, ATLAS, LHCb, (CMS)	C
CE	All VOs	C
RB		C
Global BDII		C
Site BDII		H
Myproxy		C
VOMS		H→C
R-GMA		H

CERN BDII Production Deployment Layout
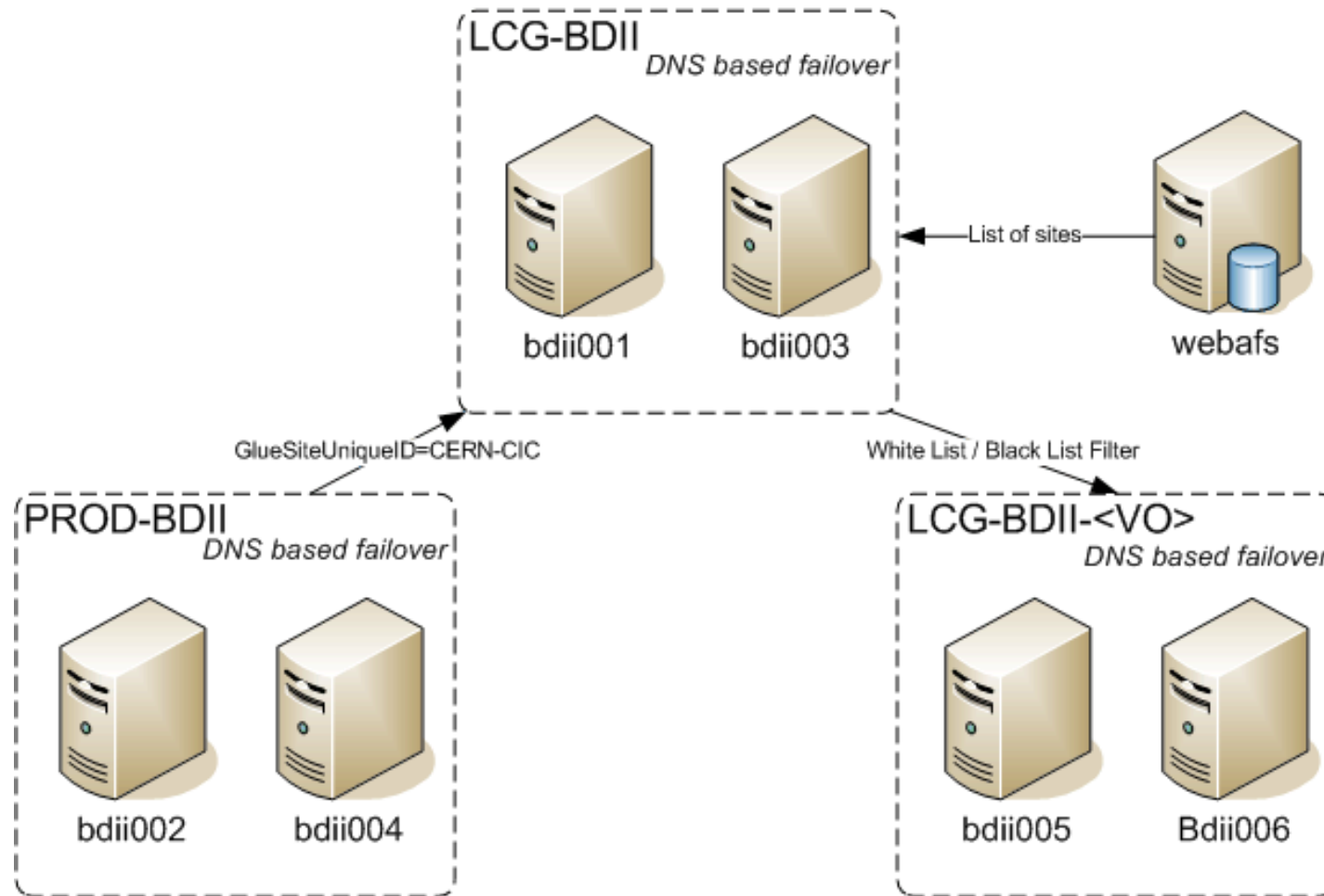
- The Global <u>BDII</u> which provides a world wide view of the <u>BDII</u> data on the grid
- The site GIIS which provides a consolidated view of the various GRIS servers on the CE and SE.
- A vo-specific <u>BDII</u> which is a view on the Global <u>BDII</u> with the inclusion of the VO white and black listing of sites

CERN PX Production Deployment Layout

myproxy

Linux HA based Master/Slave

px001
(master)

px002
(slave)
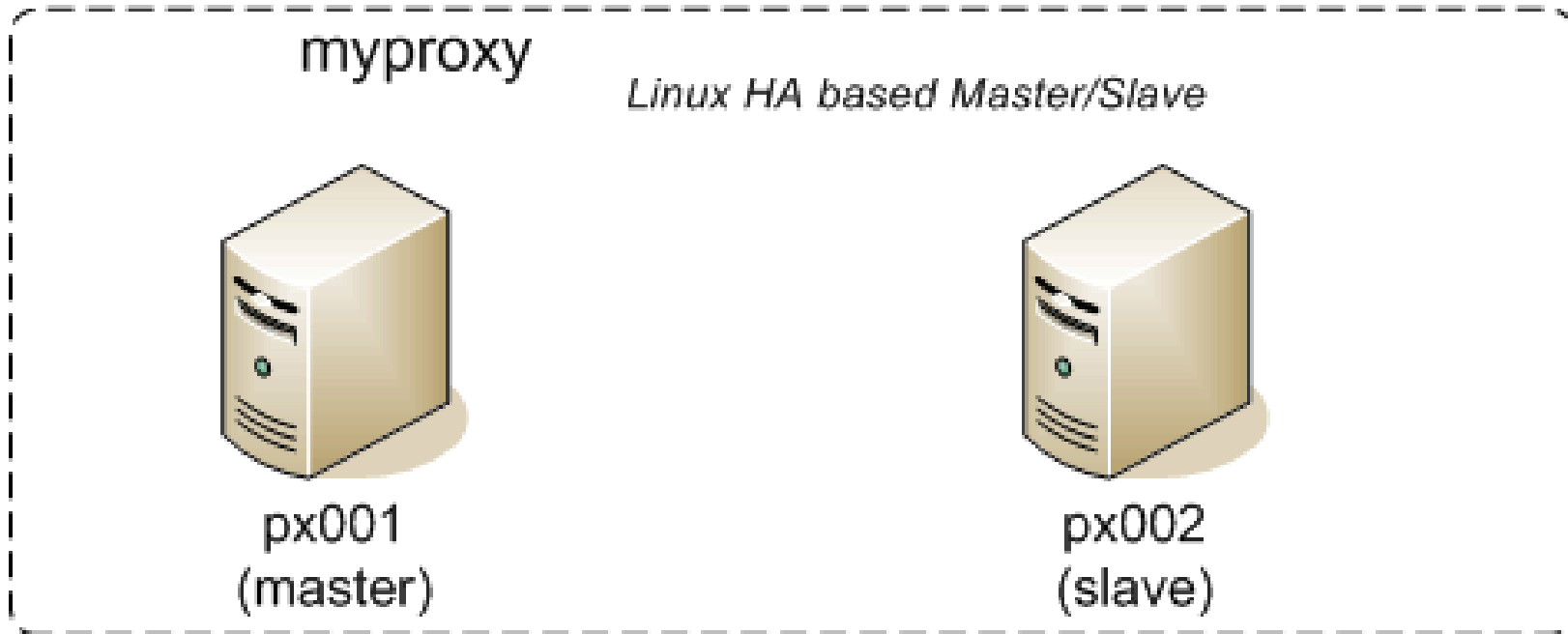
- Master/Slave set up using Linux-HA and shared IP service address
- Master stores data in /var/proxy and replicates using myproxy_replicate to slave in /var/proxy.slave
- Master rsync's data from /var/proxy to the slave /var/proxy directory
- The slave myproxy server is started in slave mode to read from /var/proxy.slave (i.e. read-only mode)
- In the event of master failure as detected by Linux-HA, the daemon is stopped on the slave and then restarted with the read-write copy from /var/proxy

LFC Production Deployment Layout

DNS based failover

lfc-lhcb-ro   lfc-lhcb   lfc-alice   lfc-atlas   lfc-cms   lfc-dteam
lfc-sixt
lfc-umosat

Oracle 10g RAC Cluster

lcg_lfc_lhcb@LCG_LFC   lcg_lfc_local@LCG_LFC   lcg_lfc_shared@LCG_LFC

27th October 2005

- Failover both at middle and database tiers

WLCG and Database Services

- Many 'middleware' components require a database:

  - dCache – PostgreSQL (CNAF porting to Oracle?)
  - CASTOR / DPM / FTS* / LFC / VOMS – Oracle or MySQL
  - **Some MySQL only: RB, R-GMA#, SFT#**

- Most of these fall into the 'Critical' or 'High' category at Tier0
  - See definitions below; T0 = C/H, T1 = H/M, T2 = M/L

- Implicit requirement for 'high-ish service level'
  - (to avoid using a phrase such as H/A…)

- At this level, no current need beyond site-local+ services
  - Which may include RAC and / or DataGuard
  - [TBD together with service provider]
    - Expected at AA & VO levels

Required Tier1 Services

Service	VOs	Class
SRM 2.1	All VOs	H/M
LFC	ALICE, ATLAS	H/M
FTS	ALICE, ATLAS, LHCb, (CMS)	H/M
CE		H/M
Site BDII		H/M
R-GMA		H/M

Many also run e.g. an RB etc. Current status for ALICE (hidden)

ALICE RBs in SC3 Production (for ex.)

- CERN:
  - gdrb01.cern.ch:7772
  - gdrb02.cern.ch:7772
  - gdrb03.cern.ch:7772
  - gdrb07.cern.ch:7772
  - gdrb08.cern.ch:7772
  - gdrb11.cern.ch:7772
  - lxn1177.cern.ch:7772
  - lxn1186.cern.ch:7772
  - lxn1188.cern.ch:7772

- SARA:
  - mu3.matrix.sara.nl:7772

- NIKHEF:
  - bosheks.nikhef.nl:7772

- GridKA:
  - a01-004-127.gridka.de:7772

- RAL:
  - lcgrb01.gridpp.rl.ac.uk:7772

- CNAF:
  - egee-rb-01.cnaf.infn.it:7772
  - gridit-rb-01.cnaf.infn.it:7772

- SINICA:
  - lcg00124.grid.sinica.edu.tw:7772

Tier1 MoU Availability Targets

Service	Maximum delay in responding to operational problems			Average availability measured on an annual basis	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outside accelerator operation	24 hours	48 hours	48 hours	n/a	98%
All other services – prime service hours[1]	2 hour	2 hour	4 hours	98%	98%
All other services – outside prime service hours	24 hours	48 hours	48 hours	97%	97%

[1] Prime service hours for Tier1 Centres: 08:00-18:00 in the time zone of the Tier1 Centre,
during the working week of the centre, except public holidays and other scheduled centre closures.

Required Tier2 Services

| Service	VOs	Class
SRM 2.1	All VOs	M/L
LFC	ATLAS, ALICE	M/L
CE		M/L
Site BDII		M/L
R-GMA		M/L

There are also some optional services and some for CIC/ROC and other such sites (this applies also / more to Tier1s…)

Measuring Service Availability

- Will be measured using standard tests run from the Site Functional Test framework

- Will start by regular tests, frequency matched to Service Class
    - i.e. Critical components will be tested every hour
    - High every 4 hours etc.

- This means that interruptions shorter than sampling frequency may be missed
    - But will be supplemented by logs and other information…

- More complex jobs, including VO-specific ones, can / will be added
    - e.g. transfer of data from Tier0 – Tier1 is higher-level function closer to MoU responsibilities

# Measuring computing resources availability - status

- Based on SFT jobs sent to all sites at least once per 3 hours
 - More frequent submissions if needed
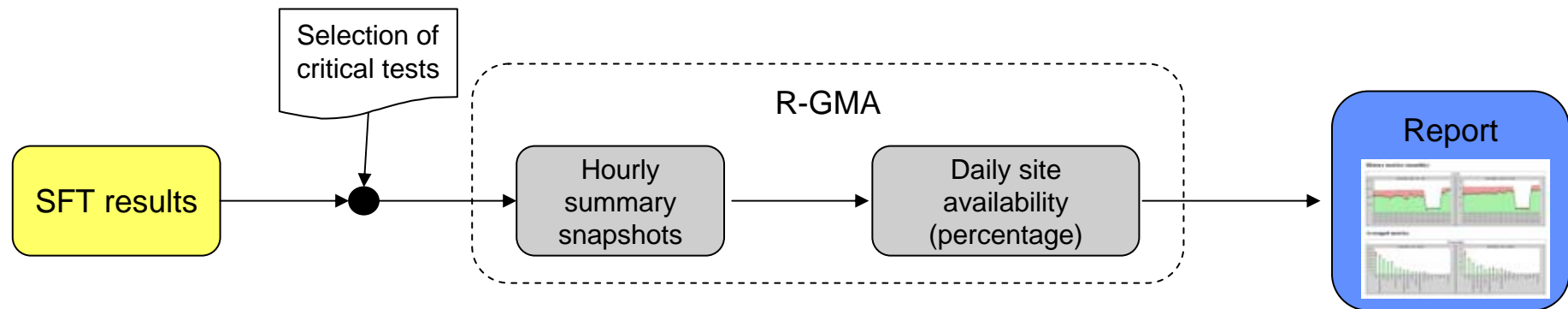
- Measurements stored and archived in R-GMA
  - Currently MySQL but Oracle foreseen
- Aggregated by region (ROC) and for the whole grid
- Current report shows only regional aggregation but "per site" view will be available soon
  - Data is already there
- Additional metric: availability multiplied by published amount of CPUs
  - "Good" resources vs. potential resources
- No direct testing of storage resources
  - Indirect testing - replica management tests

Piotyr Nyczyk

Measuring computing resources availability - graphs

History metrics (monthly)

Averaged metrics

Piotyr Nyczyk
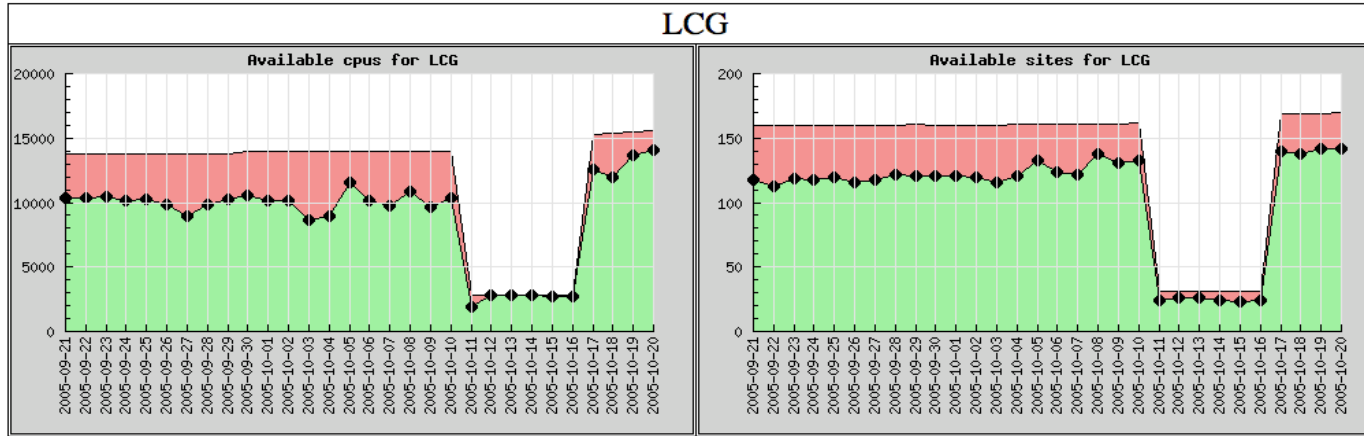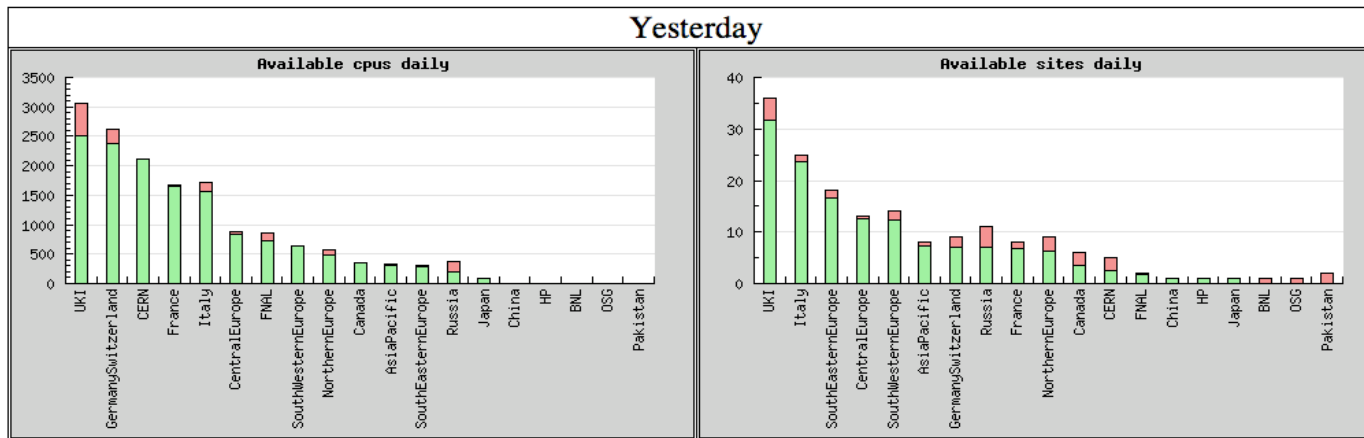
TierO Services – Status of Monitoring

Service	Responsible	Class
SRM 2.1	Dave Kant	C
LFC	LFC support	C
LFC	LFC support	H
FTS	FTS support	C
CE	Monitored by SFT today	C
RB	Dave Kant (partially done)	C
Global BDII	Tbd (Gstat) Min Tsai	C
Site BDII	Done (Gstat) Min Tsai	H
Myproxy	Maarten Litmaath	C
VOMS	Valerio Venturi	H→C
R-GMA	Lawrence Field	H

WLCG – Major Challenges Ahead

1. Get data rates at all Tier1s up to MoU Values
   - Stable, reliable, rock-solid services
   - <u>We are currently about 1/2 the target level, without including tape</u>
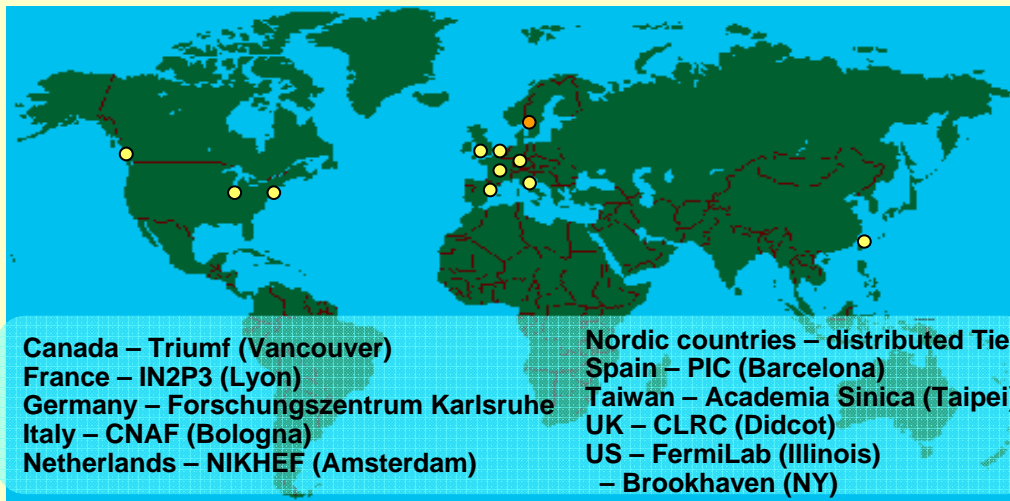
2. (Re-)implement Required Services at Sites so that they can meet MoU Targets
   - Measured, delivered Availability, maximum intervention time etc.
   - Ensure that the services delivered match the experiments' requirements

➢ <u>**T0 and T1 services are tightly coupled!**</u>
   - Particularly during accelerator operation

▪ Need to build strong collaborative spirit to be able to deliver required level of services
   - And survive the inevitable 'crises'…

LCG Service Hierarchy

Tier-0 – the accelerator centre

- Data acquisition & initial processing
- Long-term data curation
- Distribution of data → Tier-1 centres

Canada – Triumf (Vancouver)
France – IN2P3 (Lyon)
Germany – Forschungszentrum Karlsruhe
Italy – CNAF (Bologna)
Netherlands – NIKHEF (Amsterdam)

Nordic countries – distributed Tier-1
Spain – PIC (Barcelona)
Taiwan – Academia Sinica (Taipei)
UK – CLRC (Didcot)
US – FermiLab (Illinois)
**   – Brookhaven (NY)**

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

- Managed Mass Storage –
      → grid-enabled data service
- Data intensive analysis
- National, regional support
- Continual reprocessing activity

Tier-2 –   ~100 centres in ~40 countries

- Simulation
- End-user analysis – batch and interactive

Les Robertson

Tier2 Sites – Target is 20 (April) / 40 (July)

Site	ALICE	ATLAS	CMS	LHCb
Bari	X		X	
Catania	X			
Bologna			x	
Legnaro			x	
Pisa			X	
Rome			X	
Catania	X			
GSI	X			
Torino	X			
DESY			X	
CIEMAT+IFCA			X	
jinr	x			
itep	x		x	x
sinp			x	
mano			x	x
TAIWAN NCU			X	
IC			X	
Caltech			x	
Florida			x	
Nebraska			X	
Purdue			x	
UCSD			X	
Wisconsin			X	

This is not an official list!

We should easily(?) meeting April target! But need to measure service delivered!
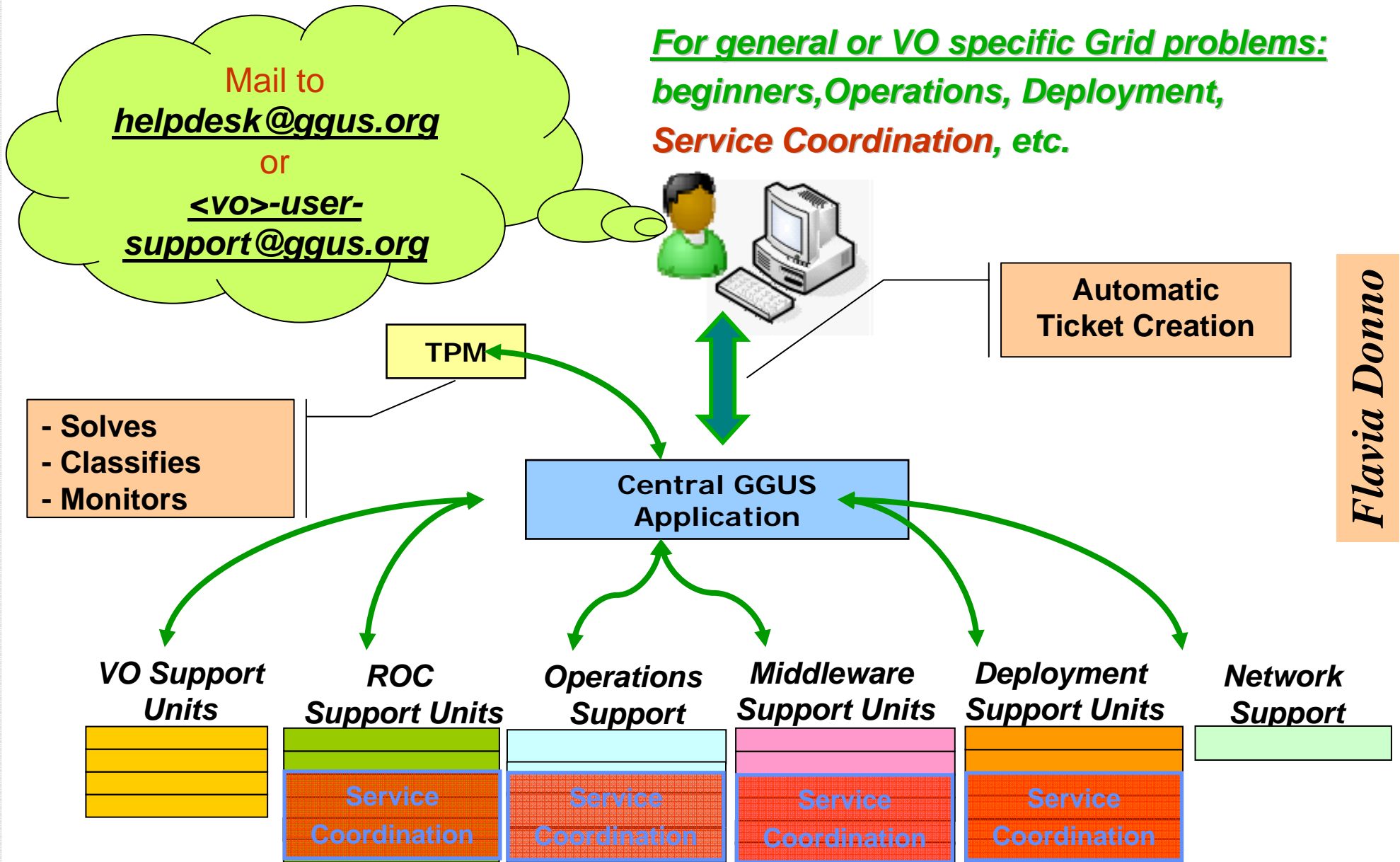
Operations Goals

- Take active role in EGEE and joint EGEE-OSG operations workshops (and any others that are relevant...)

- Joint responsibility for COD 7 workshop agenda? (Jan 17-18, Barcelona)

- Started understanding how Grid operations & Tier0 operations can interact

➢ Weekly con-call with sites still useful (experiments represented)

- Ramp-up use of standard infrastructure, improving as needed

- Goal: MoU targets automatically monitored using Site Functional Tests prior to end-2005

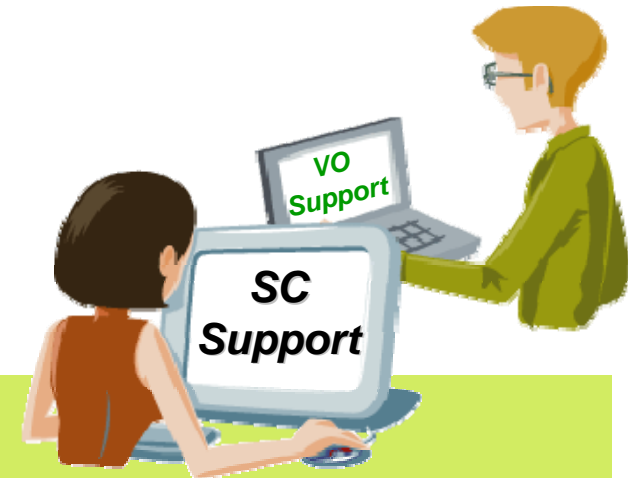➢ This will provide required basis on which to build Grid User Support

User Support Goals

- As services become well understood and debugged, progressively hand-over first Operations, then User Support, to agreed Grid bodies

- Target: all core services well prior to end-September 2006 milestone for the Production WLCG Service

- Propose: identify an experiment prepared to test this now

- ATLAS is the obvious candidate...

Service Coordination – GGUS Support Workflow

The LHC Computing Grid – (The Worldwide LCG)

Mail to
helpdesk@ggus.org
or
<vo>-user-support@ggus.org

For general or VO specific Grid problems:
beginners, Operations, Deployment, Service Coordination, etc.

Automatic Ticket Creation

Flavia Donno

TPM

- Solves
- Classifies
- Monitors

Central GGUS Application

VO Support Units

ROC Support Units

Operations Support

Middleware Support Units

Deployment Support Units

Network Support

Service Coordination

Service Coordination

Service Coordination

Service Coordination

Service Coordination - GGUS Schedule

VO
Support

SC
Support

- *The plan*:
  - Need to define special Category Type for Service Coordination
  - Need to define special support units in addition to what already there
  - Prepare/Update user/site documentation for supporters and users
  - Train Supporters
  - Make public announcement of system availability
  - Work with VOs to use/improve current implementation

- *The schedule*:
  - The GGUS ticketing system will be ready in 1 week from now
  - Documentation ready in about 2 weeks
  - Supporters trained while doing the job for the first 2 weeks by a supporting team
  - ATLAS can act as guinea pig
  - 1st of December 2005 system running in production with full support for Service Coordination

Flavia Donno

WLCG Service Coordination

- Fortnightly Service Coordination meetings held at CERN
  - Almost all work prepared beforehand

- Weekly con-calls will possibly be split into two (but seem to work well):
  1. Experiment usage of WLCG Services (what's happening, what's coming…)
  2. Services Issues (move to operations meeting?)

- Quarterly WLCG Service Coordination Meetings
  - All Tier1s, main Tier2s, … minutes, agenda etc, material circulated in advance…
  - 1st is December 20th at CERN

- Bi-annual Service workshops
  - Possibly early June? Proposal of Northern Sardinia… Joint with COD 9??
  - Should have updated picture of LHC schedule and initial operations by then…

- Thematic workshops, site visits as required
  - Each Tier1 visited once per quarter(?)
    - Combined with other events where appropriate
  - Regular 1-1 Video Meetings

WLCG – Major Challenges Ahead

1. Get data rates at all Tier1s up to MoU Values
   - This is currently our biggest challenge – by far
   - Plan is to work with a few key sites and gradually expand
   - (Focus on highest-data rate sites initially…)

2. (Re-)deploy Required Services at Sites so that they meet MoU Targets
   - Tier0 will have all services re-deployed prior to SC4 Service Phase (WLCG Pilot)
   - Plans are being shared with Tier1s and Tier2s, as will be experience
   - LCG Service Coordination team will be proactive in driving this forward
   - A lot of work, but no major show-stopper foreseen

3. Understand other key Use Cases for verification / validation
   - Many will be tested by experiment production
   - Which should be explicitly tested as dedicated "Service Tests"?

How do we measure success?

- By measuring the service we deliver against the MoU targets

    - Data transfer rates
    - Service availability and time to resolve problems

- By the "challenge" established at CHEP 2004:

    - [The service] *"should not limit ability of physicist to exploit performance of detectors nor LHC's physics potential"*
    - *"…whilst being stable, reliable and easy to use"*

- Preferably both…

    - Actually I have a 3rd metric but I'm saving that for CHEP

Service Coordination Team

- James Casey
- Flavia Donno
- Maarten Litmaath
- Harry Renshall
- Jamie Shiers

+ other members of IT-GD, IT in general, sites, experiments…

Conclusions

- A great deal of progress in **less than one year…**

- Which is what we have left until **FULL PRODUCTION**

- Focus now is on **SERVICE**

- Service levels & functionality (including data transfers) defined in WLCG MoU

- A **huge** amount of work by **many** people… Thanks to **all!**