



Lemon + SLS

WLCG Grid Monitoring WG, 25/1/07

Presented by G. Cancio – CERN/IT





Outline



- ◆ Lemon Screenshots
- ◆ SLS Screenshots
- ◆ Sensor information to be provided by M/W developers
- ◆ Real life example (using Lemon)



Lemon Screenshots



Lemon-status



Using a web browser

◆ CC Overview

CERN Monitoring Web Pages - CERN Computing center informations - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites PCFinder Configuration Metrics Error trending Help

Address http://ccs003d.cern.ch/lemon-status/ Go Links

Home Documentation Alarms PCFinder Configuration Metrics Error trending Help

CERN computing center info: overview 10 Oct 2005 Mon 14:16:17

Interactive services
 hosts up: 45 (97.8%)
 hosts total: 46
 av. load: 1.17

Batch services
 hosts up: 1094 (99.6%)
 hosts total: 1098
 av. load: 2.21

Fileserver services
 hosts up: 178 (99.4%)
 hosts total: 179
 av. load: 0.35

Tapeserver services
 hosts up: 69 (95.8%)
 hosts total: 72
 av. load: 0.33

CPU - last day Network - last day

Search host: Virtual Clusters Clusters Racks HW models Virtual Organizations PDUs Power Locations

Last modified by mirsi (CERN IT/FIO-FS), April 01 2005 13:39:27. PHP version: 4.3.4

W3C HTML 4.01 W3C CSS

Local intranet



Lemo

Using a web-

- ◆ CC Overview
- ◆ Clusters and
- ◆ VO's

CERN Monitoring Web Pages - CERN Computing center informations - Microsoft Internet Explorer

Address: http://lemonweb.cern.ch/lemon-status/info.php?vo=vo_cms&detailed=Get%20Detailed%20Information&auto_update=Not%20auto%20update&time=0

Home | Documentation | Alarms | PCFinder | Configuration | Metrics | Error trending | Help

Vo info: vo_cms 10 Oct 2005 Mon 14:10:36

Virtual Organization Information

# of hosts (down):	690 (10)
operating system(s):	2.4.21-32.0.1.EL.cernsmp
# of CPUs (down):	1380 (20)
average up time:	90 days, 17h:49m (boots per host)
hosts down:	lxb0412 , lxb1357

ITCM history

Select from hosts:

Metric Distributions Correlations

Load Percentages

0-0.5	28.0%
0.5-1.0	66.4%
1.0-2.0	2.6%
> 2.0	1.8%
down	1.4%

CPU utilization - last day

User CPU	aver: 875.10m max: 13322.73m min: 510.17m curr: 883.17m
System CPU	aver: 2062.27m max: 7740.94m min: 1339.31m curr: 1931.15m
Nice CPU	aver: 89588.65m max: 95033.09m min: 75539.00m curr: 84086.81m
Idle CPU	aver: 7474.69m max: 14609.52m min: 2949.38m curr: 13099.15m

Network utilization - last day

In	aver: 327.41M max: 568.78M min: 215.65M curr: 566.71M
Out	aver: 20.28M max: 58.17M min: 11.16M curr: 24.58M

Search host: Virtual Clusters | Clusters | Racks | HW models | Virtual Organizations | PDUs | Power Locations

Hosts availability

last day - Hosts up (average)

UP_HOSTS	aver: 693.15 max: 836.43 min: 547.69 curr: 681.92
----------	---

last day - Hosts down (average)

DOWN_HOSTS	aver: 0.00 max: 0.00 min: 0.00 curr: 0.00
------------	---

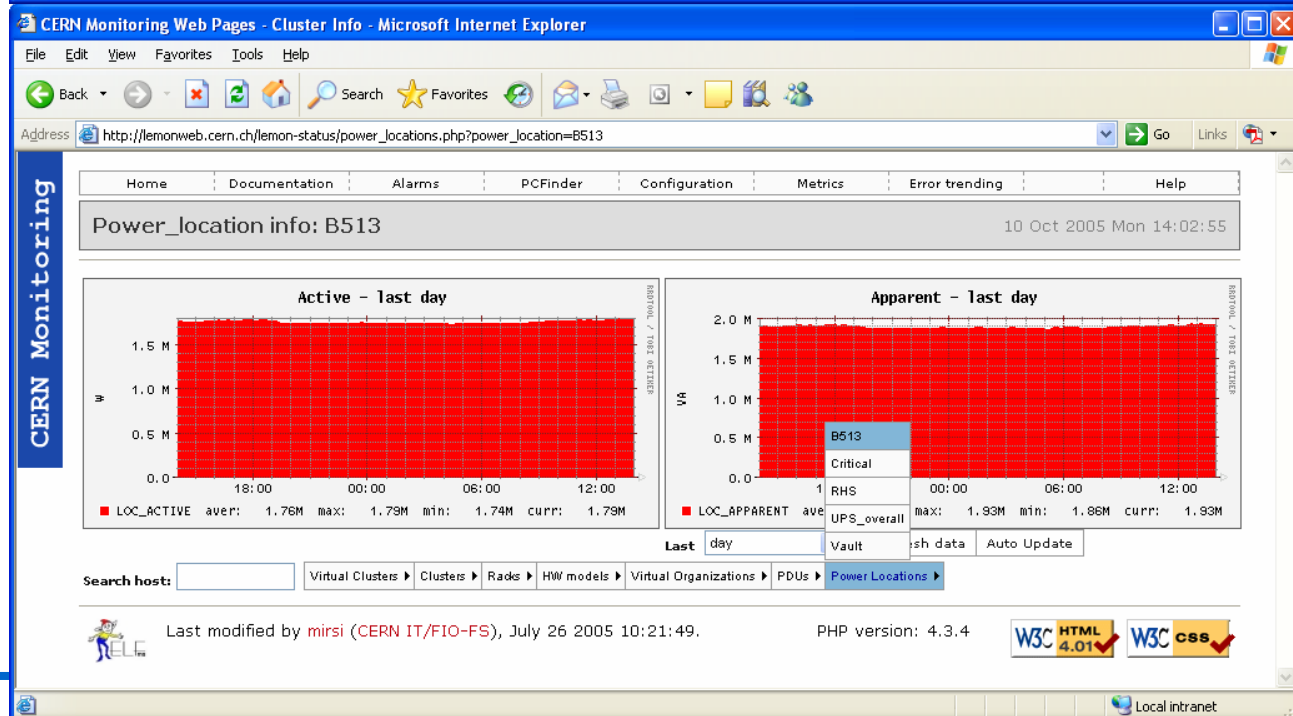
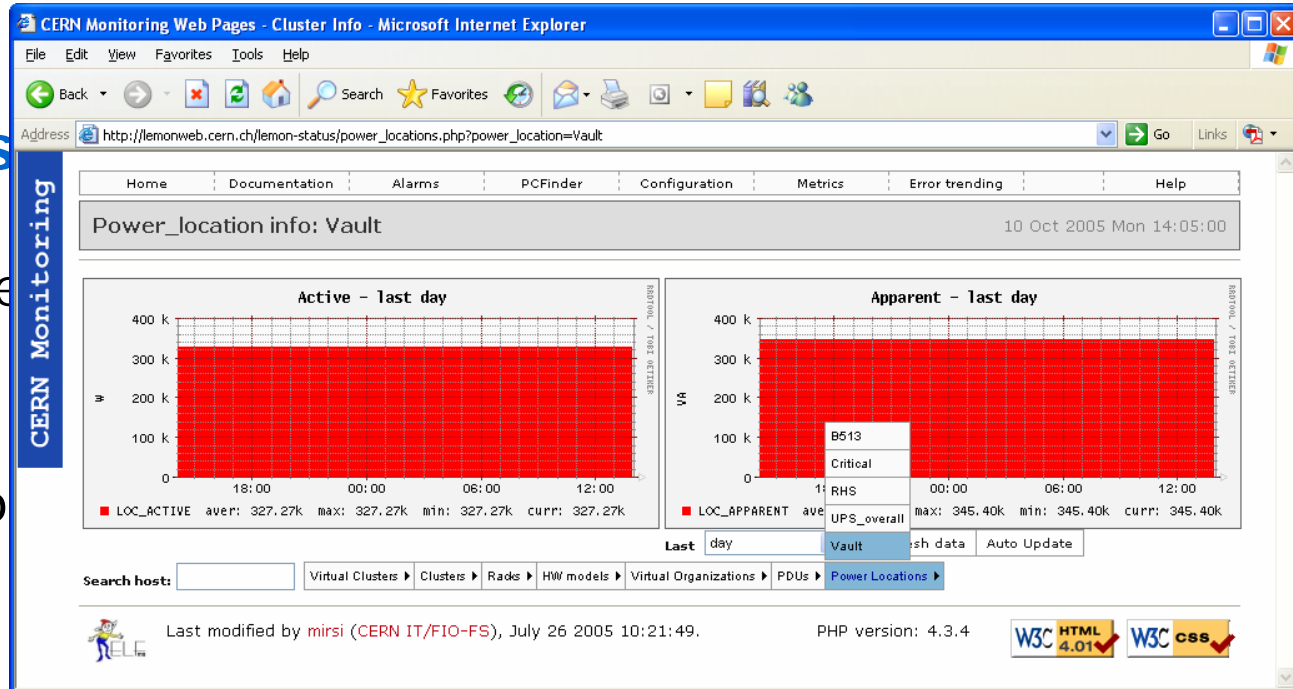




Lemon-s

Using a web-based

- ◆ CC Overview
- ◆ Clusters and no
- ◆ VO's
- ◆ Power



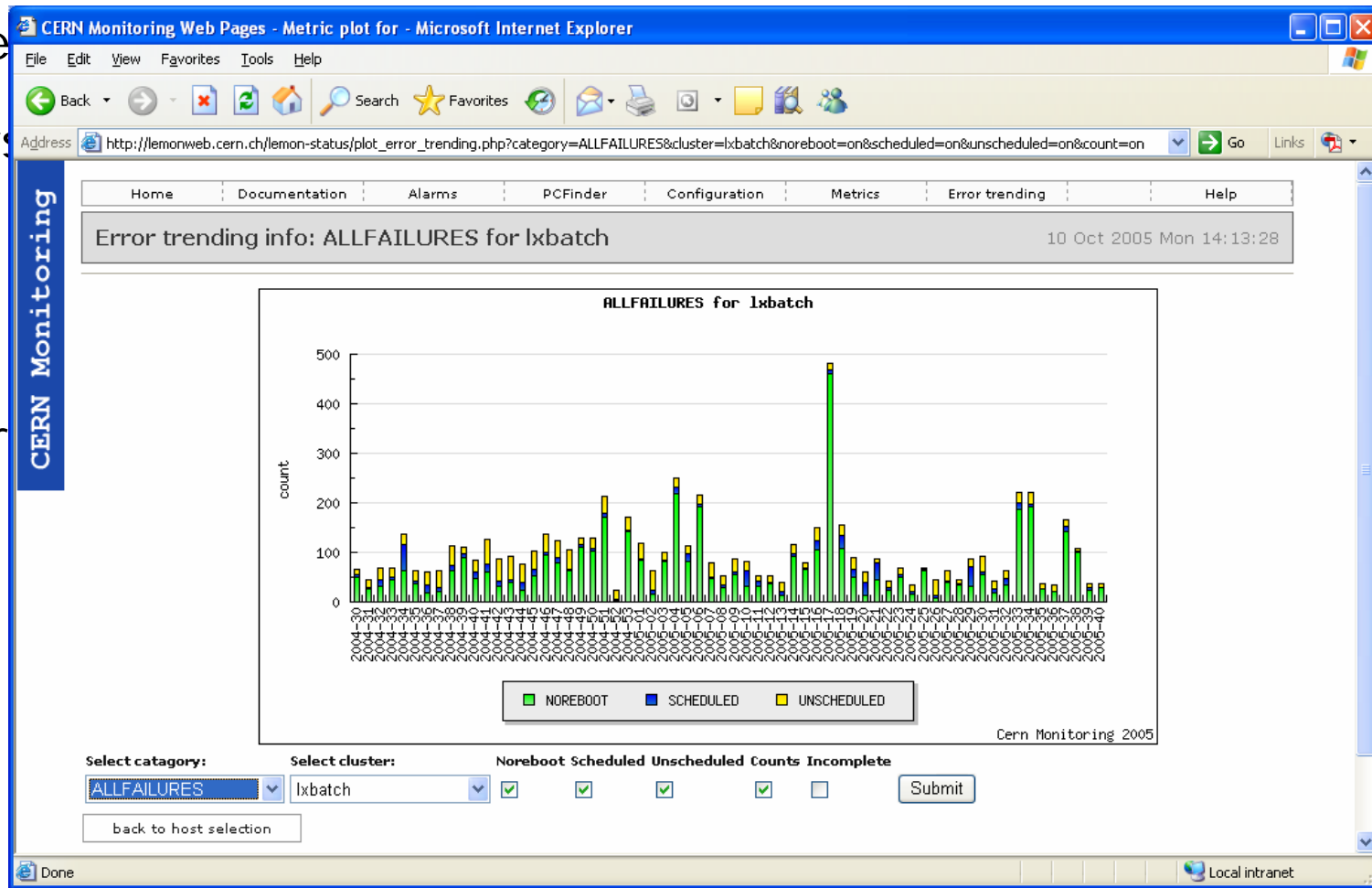


Lemon-status



Using a web-based status display:

- ◆ CC Over
- ◆ Clusters
- ◆ VO's
- ◆ Power
- ◆ Error tr





Lemon-status



Using a web

- ◆ CC Overview
- ◆ Clusters and
- ◆ VO's
- ◆ Power
- ◆ Error trends
- ◆ Batch systems

CERN Monitoring Web Pages - LSF CERN Computing center informations - Microsoft Internet Explorer

Address: http://ccs003d.cern.ch/lsf/

LSF queues info: all queues

10 Oct 2005 Mon 17:56:57

LSF Information

active users: akalinow, alaszlo

Select from queue:

Running jobs per queue

1nd	aver: 326.16	max: 502.01	min: 181.00	curr: 385.20
1nh	aver: 138.26	max: 478.55	min: 30.84	curr: 46.02
1nw	aver: 79.61	max: 84.67	min: 69.00	curr: 76.00
2nd	aver: 568.60	max: 746.82	min: 306.24	curr: 450.05
8nh	aver: 365.20	max: 616.81	min: 177.57	curr: 401.62
8nm	aver: 8.43	max: 55.73	min: 3.60	curr: 5.13
cmsdc04	aver: 0.00	max: 0.00	min: 0.00	curr: 0.00
cmsprs	aver: 47.54	max: 60.00	min: 11.00	curr: 16.00
grid_alice	aver: 7.48	max: 8.00	min: 7.00	curr: 7.00
grid_atlas	aver: 23.01	max: 87.00	min: 6.00	curr: 8.00
grid_cms	aver: 191.21	max: 517.92	min: 104.00	curr: 498.65
grid_dteam	aver: 0.02	max: 1.00	min: 0.00	curr: 0.00
grid_lhcb	aver: 7.18	max: 30.00	min: 6.00	curr: 30.00
grid_na48	aver: 0.00	max: 0.00	min: 0.00	curr: 0.00
grid_unosat	aver: 0.00	max: 0.00	min: 0.00	curr: 0.00
prod100	aver: 153.18	max: 168.00	min: 150.00	curr: 168.00
prod200	aver: 328.17	max: 341.00	min: 265.00	curr: 294.38
prod400	aver: 592.06	max: 772.15	min: 433.00	curr: 579.00

Last

Search user: queue: group:

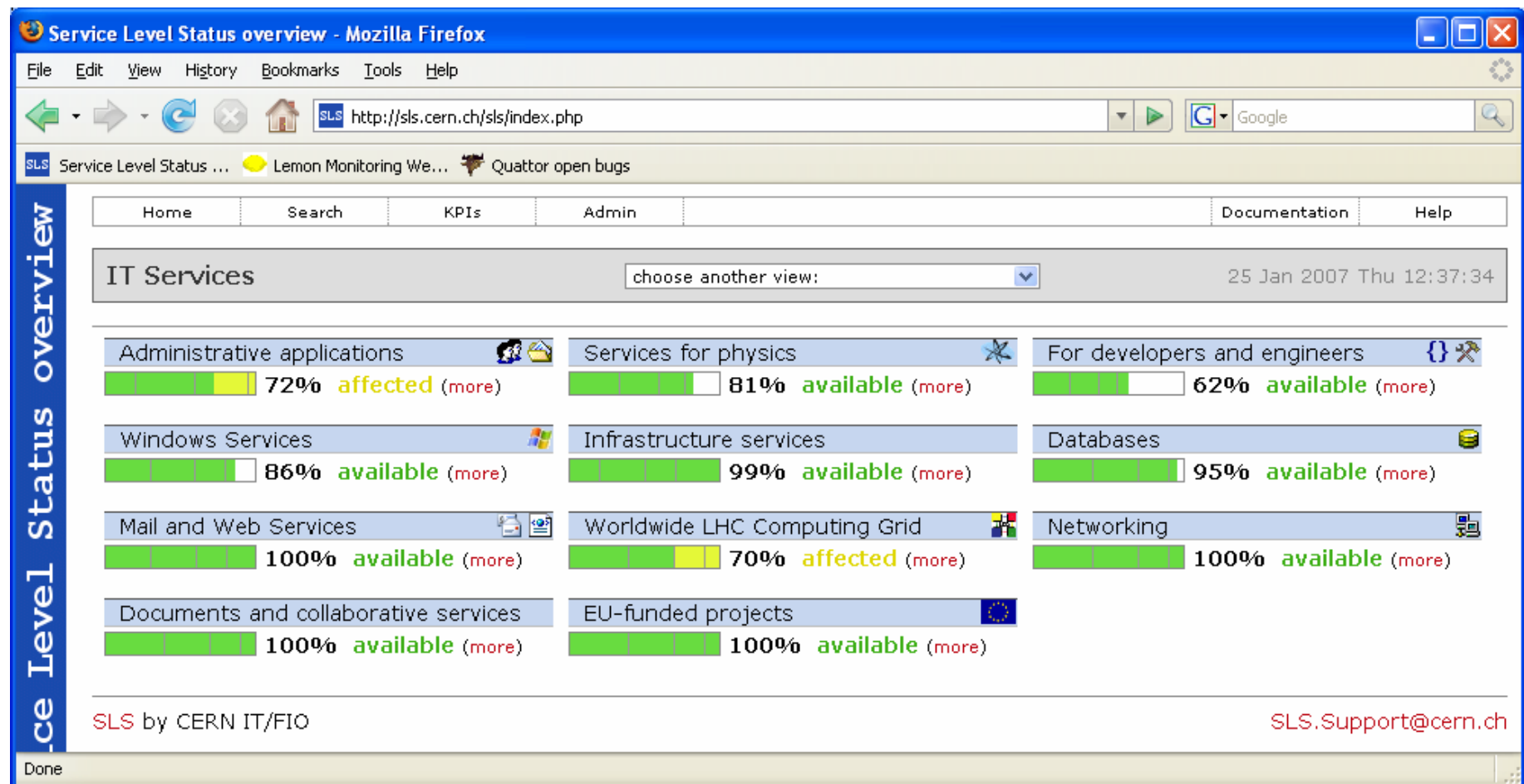
Done Local intranet



SLS Screenshots

Using a web-based status display:

◆ (Meta-)Services Overview



Service Level Status overview - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://sls.cern.ch/sls/index.php

Service Level Status ... Lemon Monitoring We... Quattor open bugs

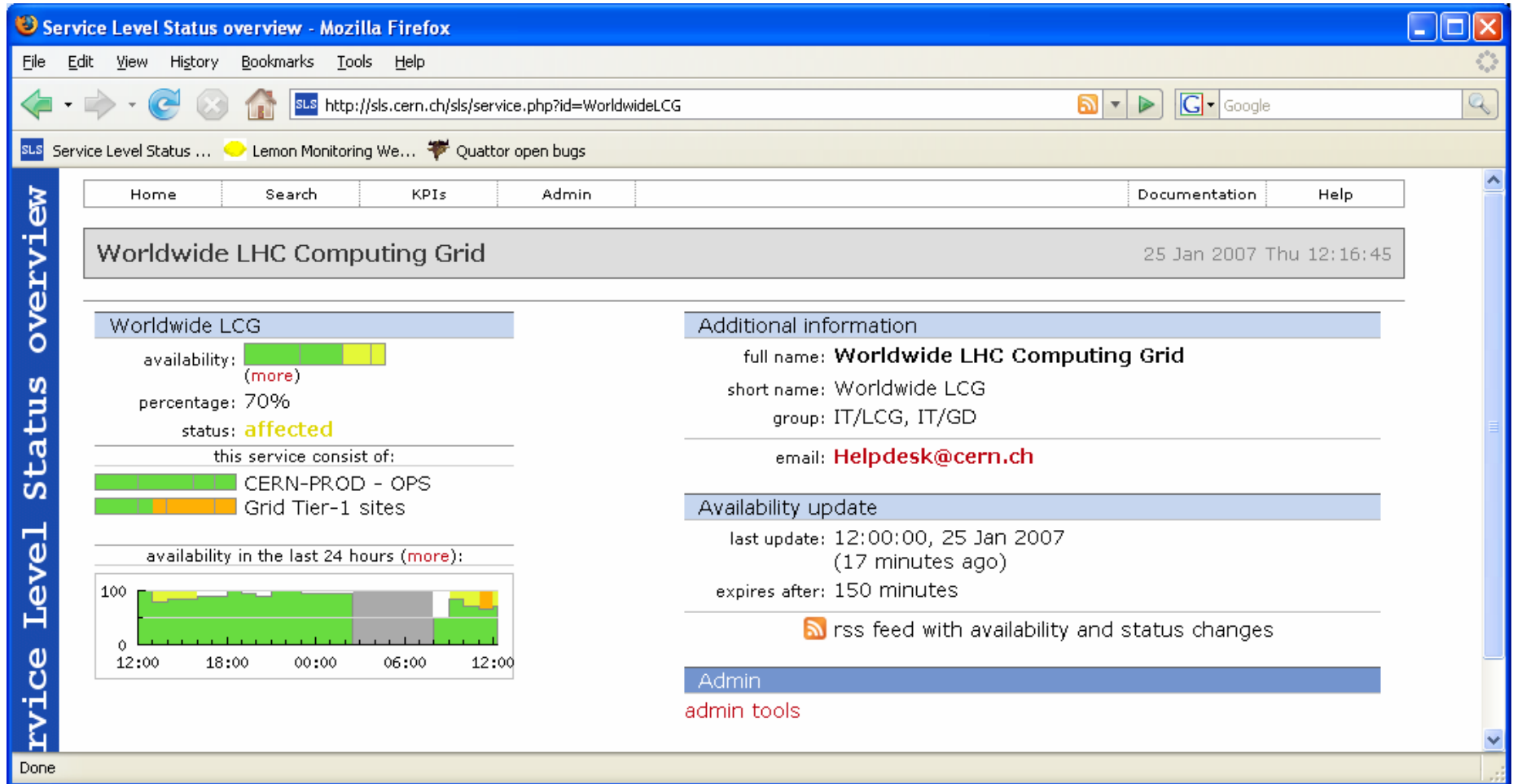
Home Search KPIs Admin Documentation Help

IT Services choose another view: 25 Jan 2007 Thu 12:37:34

Administrative applications 72% affected (more)	Services for physics 81% available (more)	For developers and engineers 62% available (more)
Windows Services 86% available (more)	Infrastructure services 99% available (more)	Databases 95% available (more)
Mail and Web Services 100% available (more)	Worldwide LHC Computing Grid 70% affected (more)	Networking 100% available (more)
Documents and collaborative services 100% available (more)	EU-funded projects 100% available (more)	

SLS by CERN IT/FIO SLS.Support@cern.ch

Using a web-based status display:



Service Level Status overview - Mozilla Firefox

File Edit View History Bookmarks Tools Help


http://sls.cern.ch/sls/service.php?id=WorldwideLCG

Service Level Status ... Lemon Monitoring We... Quattor open bugs

Home Search KPIs Admin Documentation Help

Worldwide LHC Computing Grid 25 Jan 2007 Thu 12:16:45



Worldwide LCG

availability:  (more)

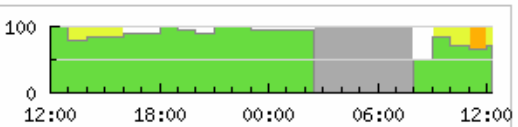
percentage: 70%

status: **affected**

this service consist of:

-  CERN-PROD - OPS
-  Grid Tier-1 sites

availability in the last 24 hours (more):



Additional information

full name: **Worldwide LHC Computing Grid**

short name: Worldwide LCG


group: IT/LCG, IT/GD

email: Helpdesk@cern.ch

Availability update

last update: 12:00:00, 25 Jan 2007 (17 minutes ago)

expires after: 150 minutes

 [rss feed with availability and status changes](#)

Admin

[admin tools](#)

Done



Us

◆

◆

◆

Service Level Status overview

Done

Service Level Status overview

Done

Service Level Status overview - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://sls.cern.ch/sls/service.php?id=LCG-Tier1-OPS

Service Level Status ... Lemon Monitoring We... Quattor open bugs

Home Search KPIs Admin Documentation Help

Grid Tier-1 sites 25 Jan 2007 Thu 12:17:43

Service Level Status overview - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://sls.cern.ch/sls/history.php?id=LCG-Tier1-OPS&more=availability&period=week

Service Level Status ... Lemon Monitoring We... Quattor open bugs

Home Search KPIs Admin Documentation Help

Grid Tier-1 sites 25 Jan 2007 Thu 12:28:27

Service name: **Grid Tier-1 sites**
[back to service details](#)

Menu

see details of:

Service availability

for the period of:

last 24 hours
 last week
 last month

More on availability in the last week

Grid Tier-1 sites
 availability - last 7 days

Statistics:

last week avg: **68% affected**
 max: **100% available** last at 24 Jan 2007, 21:00
 min: **30% degraded** last at 25 Jan 2007, 10:00



Using

- ◆ (Me
- ◆ Drill
- ◆ Mor
- ◆ A sp

Service Level Status overview - Mozilla Firefox

http://sls.cern.ch/sls/service.php?id=LCG-INFN-T1-OPS

Service Level Status overview - Mozilla Firefox

http://sls.cern.ch/sls/history.php?id=LCG-INFN-T1-OPS&more=availability&period=month

Service Level Status ... Lemon Monitoring We... Quattor open bugs

Home Search KPIs Admin Documentation Help

Grid site INFN-T1 for VO OPS 25 Jan 2007 Thu 12:19:07

Service name: **INFN-T1 - OPS**
[back to service details](#)

Menu
see details of:
[Service availability](#)
for the period of:
[last 24 hours](#)
[last week](#)
[last month](#)

More on availability in the last month

INFN-T1 - OPS
availability - last month

Statistics:
last month avg: **82% affected**
max: **100% available** last at 24 Jan 2007, 23:00
min: **0% not available** last at 25 Jan 2007, 12:00

SLS by CERN IT/FIO SLS.Support@cern.ch

availability in the last 24 hours (more):

CERN's CASTOR Tape Service 25 Jan 2007 Thu 12:20:21

Service Level Status overview

Service information

full name: **CERN's CASTOR Tape Service**
 short name: CASTOR Tape Service
 group: IT-FIO
 site: CERN

email: **Castor.Support@cern.ch**
 web site: <http://cern.ch/it-dep-fio-ds/Documentation/tapedrive/Welcome...>

service: Vladimir Bahyl
 managers: Hugo Cacote
 Charles Curran

Part of (subservice of):

Services for physics
 IT/FIO services

Subservices

none / not declared

Clusters, subclusters and nodes

cluster **tapeserver** ⚠
 subcluster **tapeserver / 3592b2** ⚠
 subcluster **tapeserver / 994br0**
 subcluster **tapeserver / 994br4** ⚠
 subcluster **tapeserver / 994br5** ⚠
 subcluster **tapeserver / t10kr1** ⚠

Service availability (more)

availability: 
 percentage: 92%
 availability info: Out of 123 tape drives of various types, there are 112 that are fully functional.
 status: **available**
 last update: 12:12:15, 25 Jan 2007 (8 minutes ago)
 expires after: 30 minutes


Service performance

Key Performance Indicators:

- ALICE Tape Space Usage TB: ●
- ATLAS Tape Space Usage TB: ●
- CMS Tape Space Usage TB: ●
- COMPASS Tape Space Usage TB: ●
- LHCb Tape Space Usage TB: ●
- NA48 Tape Space Usage TB: ●
- Other Tape Space Usage TB: ●

Depends on

this service uses:

-  castor

Depended on by

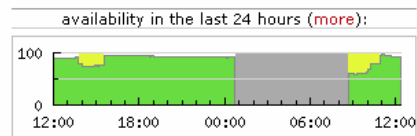
All OK: LHCb Tape Space Usage (430.66TB) was lower than maximum expected/ target level (437.21TB)

Additional service information (more)

3592B2:	95
994BR0:	90
994BR4:	92
994BR5:	100
T10KR1:	82

 [rss feed with availability and status changes](#)

how is availability measured or estimated:
 CERN's CASTOR Tape Service is considered fully available if at least 80% of all tape drives in all production device groups are running normally. However, if there is at least 1 device group with availability lower than 50%, the service will be marked as affected.



Above is the percentage availability of each tape drive device group.



SLS



Using

◆ (Me

◆ Drill

◆ Mor

◆ A sp

◆ Servic

◆ Other

Service Level Status overview

Service Level Status overview - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://sls.cern.ch/sls/?view=vo

Home Search KPIs Admin Documentation Help

Services for each Virtual Organization

choose another view: 25 Jan 2007 Thu 12:24:59

Service Level Status overview - Mozilla Firefox


File Edit View History Bookmarks Tools Help

http://sls.cern.ch/sls/service.php?id=ServicesForATLAS

Home Search KPIs Admin Documentation Help

Services for ATLAS 25 Jan 2007 Thu 12:25:24







Services for ATLAS

availability:  (more)

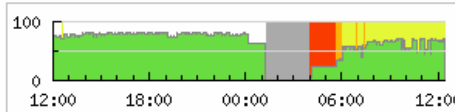
percentage: 70%

status: **affected**

this service consist of:

-  AFS Service for ATLAS
-  ATLAS RAC database
-  Castor2ATLAS
-  LXBUILD (ATLAS)
-  LXGATE (ATLAS)
-  Special Facility (ATLAS)

availability in the last 24 hours (more):



Additional information

full name: **Services for ATLAS**

vo: ATLAS


email: Helpdesk@cern.ch

web site: <http://cern.ch/Atlas...>

Availability update

last update: 12:20:56, 25 Jan 2007
(4 minutes ago)

expires after: 60 minutes

 [rss feed with availability and status changes](#)

Admin

[admin tools](#)



Service Level Status overview - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://sls.cern.ch/sls/service.php?id=Administrative


Service Level Status ... Lemon Monitoring We... Quattor open bugs

Home Search KPIs Admin Documentation Help

Administrative applications

25 Jan 2007 Thu 12:31:01
















Administrative applications

availability:  (more)

percentage: 72%

status: **affected**

this service consist of:

-  EDH system
-  AISMEDIA
-  AIS Common Login
-  Warehouse (BAAN)
-  CET system
-  Contracts (CFU)
-  Recruitment (ERT)
-  Human Resources (HR)
-  HR Management (HRT)
-  PPT system
-  PPT for LHC
-  Foundation application
-  WOS system
-  XOTHER
-  CRA

Additional information

full name: **Administrative applications**

group: IT/AIS

site: CERN


email: **AIS.Support@cern.ch**

web site: <http://cern.ch/ais/>

Availability update

last update: 12:21:30, 25 Jan 2007
(10 minutes ago)

expires after: 60 minutes

 [rss feed with availability and status changes](#)

how is availability measured or estimated:
Availability of administrative applications provided by CERN IT/AIS group is measured as a percentage of successful SiteScope test of these applications.

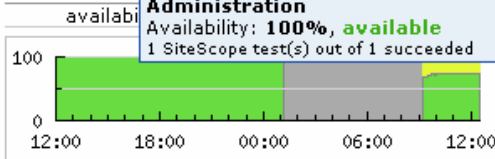
Admin

[admin tools](#)

Computing Resources Administration

Availability: **100%**, **available**

1 SiteScope test(s) out of 1 succeeded



SLS by CERN IT/FIO SLS.Support@cern.ch

Done

Service Level Status overview

- ◆ Us
- ◆ (
- ◆ [
- ◆ M
- ◆ A
- ◆ S
- ◆ C
- ◆ C



Service availability and status



Service fully (100%) available



Service available in 95%, still marked as *fully available*

- above the highest threshold



Service available in 87%, marked as *affected*

- below the highest threshold



Service available in 50%, marked as *degraded*

- below the medium threshold



Service available in 13%, marked as *not available*

- below the lowest threshold



Service info expired, update not available



Different status thresholds mean different status for services with the same availability



(more at <http://cern.ch/SLS/help.php>)



Information to be provided by MW developers



Monitoring MW services



- ◆ For monitoring a Grid Service comprehensively, it is important to look at the service **both** 'externally' (observing the service and its environment) and 'internally' (talking to the service).
 - **Externally:**
 - Related to the service itself: log files, processes
 - Related to the node(s) hosting the service: server load, free/used space on partitions/directories
 - Changes in processes, log file formats, etc. should be avoided as this implies changing the monitoring setup as well simultaneously with the new software rollout!
 - **Internally:**
 - Typically requires writing a service-specific sensor.
 - M/W developers should be encouraged to provide test applications (a la voms-ping), if possible emulating typical user behaviour. The output of that test application should be machine-parsable.
 - Ideally, availability numbers (SLS-like: 0 to 100) should be produced as well by that test application.
- ◆ It is convenient to identify the nature of a metric:
 - performance metrics (e.g. transactions per second)
 - exception metrics (e.g. transactions/s < minimum, service process dead)
- ◆ Exceptions: what needs to be done, who needs to be informed
- ◆ Recovery: Under which circumstances can a recovery action be started, and how?



Monitoring MW services: metric details



- ◆ General:
 - How often does a metric need to be sampled?
- ◆ Log files:
 - Name(s) e.g. /var/log/myservice.log, /var/log/myservice.log.*(gz|Z)
 - Format. Use a per-line format, starting with a timestamp. The timestamp should follow a standard format (e.g. syslog-like, %M %D %H:%M:%S)
 - Regular expression indicating an event (performance/exception): e.g. "(failed|Timed out)"
 - How far back one needs to look in the log file? 5m, 2h, 1d, since last boot? (May depend on sampling rate).
- ◆ Processes:
 - Name of process(es) to be checked
 - uid (root? Service uid?)
 - PPid (e.g. init=1)
 - # of processes: a single one, >=1, >=1 && <= 500?
- ◆ File system:
 - Space free/used on a file system (e.g. /var, /castor/data/*)
 - Space used below a directory
 - MD5 checksum of (read-only) files
 - File presence, size, timestamp/age



Monitoring MW services: metric details



◆ “Internal” monitoring

- What command to run (not: API to code against!) for the service test application?
- Structure of output:
 - Availability metrics (overall availability: 0 to 100)
 - Performance metrics (list of key-value pairs)
 - Detailed information (text)

◆ Exceptions and Recovery actions

- Based on the collected metrics, what (combination of) metrics indicate an exception?
- What should be done with the exception? Who should be informed?
- Is there a recovery procedure? Can the recovery procedure be executed automatically before any humans intervene?
 - What executable to run
 - Timeout
 - Max # of runs / time window in case problem persists



Example: LFC



Real example: GridLFC (thanks to James ;-).

Checking the service from 'outside' using standard sensors:

◆ Processes

- [Exception] lfc daemon running processes: Should be exactly 1 (name: "lfcdaemon"; uid: 17700, ppid: 1 (init))
 - Raise exception if the above is not true
- [Exception] Same for lfc-dli process

◆ Log files

- [Exception] `/var/log/lfc/log`: Check for Oracle errors.
 - Date format: `%m/%d %T`
 - Regexp to search for: `ORA-(?! (02396 | 03113 | 12547 | 12514 | 03114))`
 - Raise exception if true more than 0 times in the last 30 minutes
- [Performance] `/var/log/lfc/log`: How many LFC (read/write/delete/total) operations in the last 5 minutes
 - Regexp to search for:
 - total: `NS098 -`
 - write: `NS098 - (creat|mkdir|setfsizereg|symlink|addreplica)`
 - (similar for read and delete)



Example: LFC



- ◆ Specific service checks, talking directly to the service using a new sensor:
 - [Exception] lfc.readEntry: check we can do a stat on a LFC file (status: integer)
 - Read /grid/ops. Raise exception if failure
 - [Exception] lfc.writeEntry: check we can write a new directory into a given LFC directory (status: integer)
 - Create /grid/ops. Raise exception if failure
 - [Exept+Perform] lfc.timeReaddir: return how long it takes to do a readdir on a given LFC (time: integer)
 - Read /grid/ops. Raise exception if takes > 10s
 - [Performance] lfc.activeConnections: count the number of TCP connections to the LFC port in ESTABLISHED state (count: integer)



Example: LFC



Recovery action: Try to restart the service if the following is true:

- LFC daemon not running OR cannot read /grid/ops OR cannot write /grid/ops
- Recovery action: Run

```
/usr/bin/killall -9 lfcdaemon; sleep 1; /sbin/service lfcdaemon start
```

 - Try this a maximum of 3 times, if the error condition above is still true after 3 attempts, give up, inform human operators