



# Continuous Service at CC-IN2P3

**Fabio Hernandez**  
fabio@in2p3.fr

LCG Management Board, August 29th, 2006

# ▶ Contents



CCIN2P3



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE

- Context
- On-call service
  - Organisation
  - Mode of operation
  - Tools
- Questions



- Round-the-clock site operation
  - Public holidays, Christmas and new year days included
- Office time
  - Monday to Friday: 8h-18h
- Operator presence
  - Monday to Friday: 7h30-21h
  - Saturday, Sunday: 10h-17h45
- Guardian presence
  - Monday to Friday: 18h-8h
  - Saturday, Sunday and public holidays: 24h

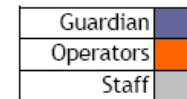
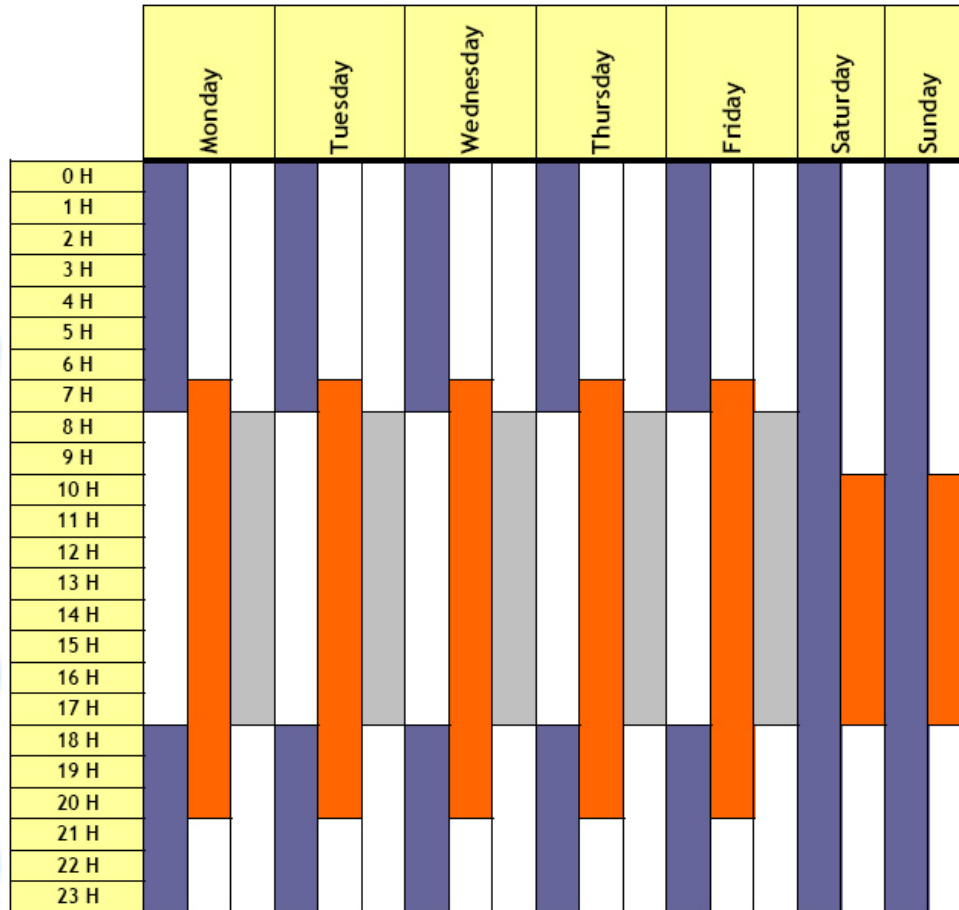
# On-site presence



CCIN2P3



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE



# ▶ On Call Service



- Goal
  - Maintain the core site services in the best possible level of operation during non-office hours
    - *Tactic: automatically detect the dysfunctional component as early as possible and (if needed) trigger human intervention*
    - *Note that the help desk during non-office hours is a best effort service*
- On call service is under the responsibility of the Operations team
  - Coordination, planning, documentation, monitoring tools, web portal, etc.
- Rotation
  - 1 engineer on duty during a full week (Thursday to Thursday)
  - 1 cycle every 4 months
    - *currently around 20 individuals are members of this team*
  - Covers the period not covered by office hours neither by the operators

# ▶ On Call Service (cont.)



CCIN2P3



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE

- Intervention of the on-call engineer can be triggered by
  - Site's guardian
    - *For power, cooling, fire or flooding incidents*
  - On-site operator
    - *If he cannot fully handle an incident*
  - Alarming system through text messages sent to the mobile phone
    - *Sent as a result of a detected abnormal situation requiring immediate human intervention*
- On-call person also performs active monitoring
  - Regularly connects to the site for monitoring the operational status of the core services
  - Reads e-mail and browse tickets opened by end users
- The main role of the on-call person is:
  - To identify the affected service/component and assess the critical level of the incident
  - To isolate it or to trigger immediate or delayed intervention by the experts
  - To coordinate the intervention until the incident is completely handled

# ▶ On Call Service (cont.)



- Tools
  - **Dedicated account**
    - *For e-mail and for interactive intervention on critical machines/services*
    - *Appropriate privileges granted to this account*
  - **Dedicated mobile phone**
    - *With access to personal phone numbers of the entire site's staff*
    - *Some level of organization is needed to know who to call when for each service*
  - **Instant messaging**
    - *Jabber-based private server managed by the site*
    - *For coordinating remote interventions by experts*
  - **Broadband internet connection from home for people likely to remotely intervene on the site**
    - *Provided by any commercial ISP*
    - *Connection bill is paid by the individual and reimbursed by the site*
    - *Site-managed VPN in place*
  - **An internal web portal giving access to a collection of monitoring tools**

# ▶ On Call Service (cont.)



- Separate on-call people (and procedures) for facility-related incidents
  - Cooling, power, fire, flooding, etc.
- Operating model proven effective
  - In place since 1997
  - Evolution from a more simple model implemented during the mainframe era



# Monitoring & Alarming



CCIN2P3



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE

- Collection of web-based tools for having a quick overview of the operational status of the services, including
  - Number of queued and running jobs over the last 24 hours
  - Details on (probably abnormally) early-ended jobs
  - List of jobs that look stuck (from the batch system point of view)
  - Global status of core services like batch, AFS, HPSS, dCache, cartridge library, ...
  - Connectivity status of the site
  - ...
- Home-grown tool for centralizing messages from every service/machine
  - It is the main message bus
  - Web-based interface with links to hints associated to the message and the actions to be triggered (if any)
    - *Filtering by severity, provenance, timestamp or contents of the message*
    - *The knowledge base is constantly updated*
    - *The same tool is also used by the operators*
  - NGOP-based alarm system extracts information from this message repository and triggers alarms (if needed)
    - *SMS messages, e-mail*

# ▶ Introducing new services



CCIN2P3



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE

- Introducing new services (like grid components) in our system is greatly facilitated if we have
  - Automatic ways to remotely query/detect the operational status of the service
  - An interface to interact with the service
    - *start, stop, drain, shutdown, ...*
    - *Standardizing this interface (protocol and actions) for all the grid services would be extremely helpful*
  - Standardized formats and locations of log files

# Questions/Comments



CCIN2P3



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE



# Backup Slides (1/2)

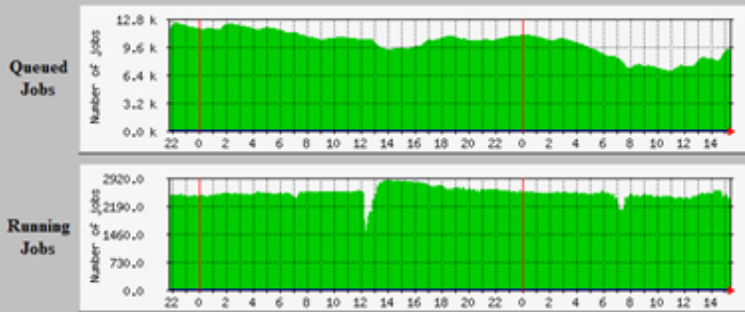


CCIN2P3



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE

Queued And Running Jobs : Daily statistic



RLS Web Console

Start date: 2006-08-28 Start time: 00:00:00  
 End date: 2006-08-28 End time: 23:59:59  
 Severity: greater than WARNING Filter: none  
 Tag: contains \*  
 Hostname: contains \*  
 Message: contains \*  auto reload  operator

Field selection

user/group  hostname  PID  severity  
 tag/group  message  received date  sent date

clear reset display

111 messages found - 2006/08/28 from 00h00 to 23h59

|        |          |         |         |  |
|--------|----------|---------|---------|--|
| 024600 | 01:28:00 | SIMONE  | ALERT   | alert, "/var/core" is over limit (100%)  |
| 024600 | 03:17:16 | AFSSURV | ERROR   | afsd is OK on ccr00000.in2p3.fr  |
| 024600 | 04:21:00 | SIMONE  | ALERT   | alert, "/var/core" is over limit (100%)  |
| 024600 | 07:07:49 | HPSS    | WARNING | HAJR CODE 2079 RC=11 Data copy operation failed during Stage: Resource temporarily unavailable |
| 024600 | 07:11:02 | HPSS    | WARNING | HAJR CODE 2079 RC=11 Data copy operation failed during Stage: Resource temporarily unavailable |
| 024600 | 09:32:30 | SIMONE  | ALERT   | alert, "/var/core" is over limit (93%)   |
| 024600 | 10:13:34 | AFSSURV | ERROR   | afsd is OK on ccr00000.in2p3.fr  |
| 024600 | 10:41:55 | RXSTAT  | ALERT   | Cache manager failed on ccr00000.in2p3.fr  |
| 024600 | 10:41:55 | RXSTAT  | ALERT   | Cache manager failed on ccr00000.in2p3.fr  |
| 024600 | 10:41:55 | RXSTAT  | ALERT   | Cache manager failed on ccr00000.in2p3.fr  |
| 024600 | 10:41:55 | RXSTAT  | ALERT   | Cache manager failed on ccr00000.in2p3.fr  |
| 024600 | 10:41:55 | RXSTAT  | ALERT   | Cache manager failed on ccr00000.in2p3.fr  |
| 024600 | 10:46:35 | FPSTAT  | ALERT   | Ping failed : ccr00000.in2p3.fr  |
| 024600 | 10:54:30 | SIMONE  | ALERT   | alert, "/var/core" is over limit (100%)  |

# Backup Slides (2/2)



CCIN2P3



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

**Current Network Status**  
Last updated: Mon Aug 22 22:28:08 CEST 2006  
Updated every 30 seconds  
Version: 3.0.0.0.0.0.0.0.0.0  
Logged in as centre

View Metrics For This Host  
View Notifications For This Host  
View Service Status Details For All Hosts

**Display Filters:**  
Host Status Types: All  
Host Properties: Any  
Service Status Types: Recovery / OK  
Service Properties: Any

| Host        | Service                        | Status | Last Check          | Duration       | Uptime | Output Information   |
|-------------|--------------------------------|--------|---------------------|----------------|--------|--|
| ccicgcell02 | check_ldap_authentication      | OK     | 06-26-2006 22:34:20 | 14 10s 20m 42s | 5/16   | LDAP Authentication was successful <a href="#">[details]</a>   |
| ccicgcell02 | check_ldap                     | OK     | 06-26-2006 22:15:16 | 14 10s 20m 42s | 5/16   | 220 ccicgcell02.HQ31.N GridPP Server 1.12 OSSAP type Gridware/01 eva-2.8.2 (gr32dmg, 100200000-42) ready |
| ccicgcell02 | check_ldap_authn               | OK     | 06-26-2006 14:58:53 | 09 46 46s 23s  | 3/16   | gridconnect is up-to-date (last update: Aug 30) <a href="#">[details]</a>                                |
| ccicgcell02 | check_ldap                     | OK     | 06-26-2006 22:15:17 | 14 10s 20m 42s | 5/16   | This server is for a CI - QueueName=ccicgcell02HPSS = 3/16   |
| ccicgcell02 | check_ldap_authn               | OK     | 06-26-2006 14:57:31 | 09 46 47s 26s  | 3/16   | SUCCESS: verification is valid until Sat 27 09 25 16 2007 GMT  |
| ccicgcell02 | check_ldap_authentication_ldap | OK     | 06-26-2006 21:57:46 | 04 34 37m 22s  | 1/16   | gridconnect seems to be working <a href="#">[details]</a>  |
| ccicgcell02 | check_ldap_authn_ldap          | OK     | 06-26-2006 22:15:16 | 14 10s 20m 41s | 5/16   | 0 packets transmitted, 0 received, 0% packet loss, time 4020ms   |

7 Matching Service Entries Displayed

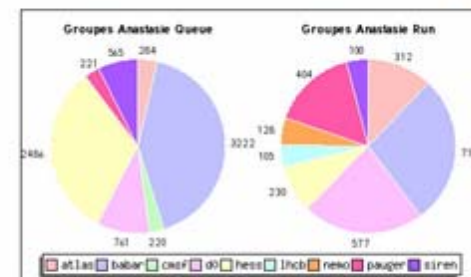
- SQS
- pistonJobs 22:38
- resources 22:41
- anastasiaJobsGrille 22:31
- jobsBio/NoeBio 22:38
- anastasiaClasses 22:38
- jobsGrille/NoeGrille 22:31
- pistonGroupes 22:38
- courbesJobs
- anastasiaGroupes 22:38
- SERVICE
- etatServices 22:37
- HPSS
- drivesHpss 22:30
- rfcioComesions 22:38
- DIVA
- drivesDivs 22:40
- + ACSLS
- LENS UTILILES
- VIDER

**JOBS GRILLE NON GRILLE**



**ETAT DES JOBS ANASTASIE PAR GROUPES**

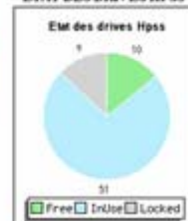
Entrez le seul limite des jobs en queue : 200 , running : 100



**ETAT GENERAL DES SERVICES**

|                    |                       |                   |                    |
|--------------------|-----------------------|-------------------|--------------------|
| SEMPER : available | ANASTASIE : available | AFS : available   | ORACLE : available |
| TEST : available   | HPSS : available      | TSM : available   | PISTOO : available |
|                    |                       | XTAGE : available | ACSL : available   |

**ETAT DES DRIVES HPSS**



**ETAT DES DRIVES ACSLS**

