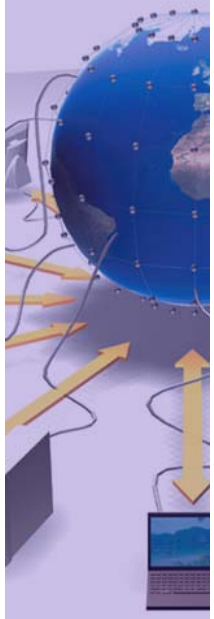
CERN IT Department

GS



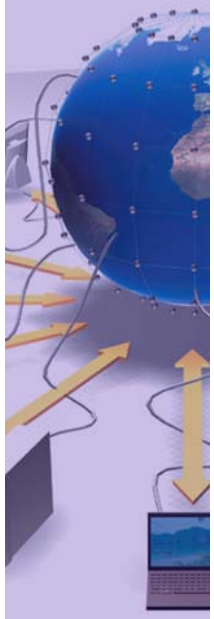
EXPERIMENT SUPPORT DURING CCRC08

D. Bonacorsi, S. Campana, P. Mendez Lorenzo, R. Santinelli

Alice: some important remarks

CERN IT Department

GS

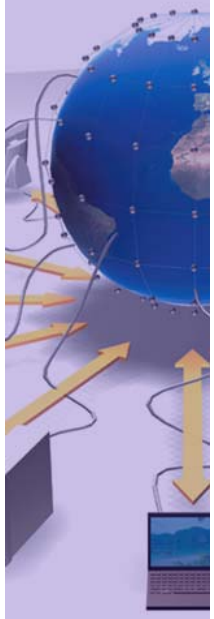


- Alice performs the Run-III exercise of the FDR in June`08
- The notification and alarm system described here belongs to the Run-II exercise (Feb-08) and beyond
- It can be extrapolated to the notification procedure that Alice will use in June`08

Alice Run-II exercise

CERN IT Department

GS

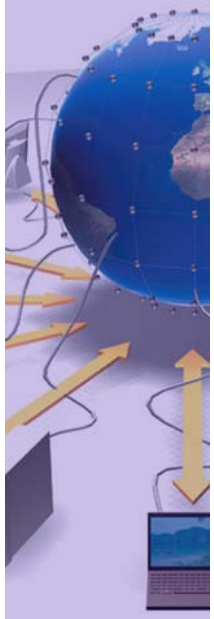


- Alice Run-II exercise: from 15th of February until the 10th of March
- Services stressed at that time
  - VO-boxes at T0 and T1
  - CE, WMS at T0
  - Castor2 at T0
  - FTS T0-T1
  - MSS@T1
  - (See Critical Services talk tomorrow for further details)
- Different notification procedures were used for all these services
- Positive feedback of Alice regarding the different alarm notification systems and the time respond in general

Alice Notification procedures

CERN IT Department

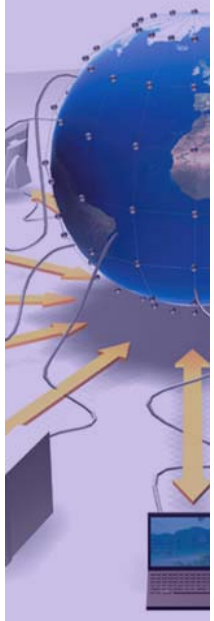
GS



- **e-log**
  - Mostly used to report VO-boxes and FTS issues
  - All notifications defined as **max priority**
    - During the June exercise the procedure will continue in this way
- **SAM alarm system**
  - Use for VOBOXES and local WMS/RB
- **Direct contact with the experts at T0**
  - FTS at the origin (fts-support list), and Castor2 issues
- **Direct contact with the Alice contact persons at the T1s**
- **GGUS**
  - Normally to ensure the tracking and history of the problem
  - In general, the direct contact with the experts is preferred

Who is “CMS” in CCRC/2?

CERN IT Department



Ops teams:

We ran CSA/CRUZET and CCRC thanks to the usual suspects

1. DataOps (we have a central team - bigger than 2)
2. FacilitiesOps (we have a central team - smaller than 1)
3. WLCG sites people !!! (our most precious resource!)

On the CMS side, we use to run nothing without allowing some % of mess

everybody free to contribute (e.g. via HN, or via chat) by spontaneously aggregating knowledge around identified tasks/problems, then dissolving groups back into their own original longish-term activities

e.g. developers are gold to us. Glad if they can pop in in case something specific arises, but we also tend by design to protect them from ops and let them just develop.

e.g. brilliant site people may want to help on general situations/problems, and go beyond their std role as site persons…

All this was constantly encouraged (just trying to fight the entropy increase)

When this worked, it’s Ops people merit

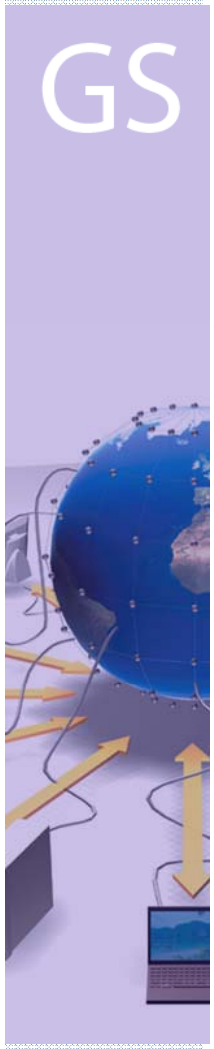
When this did not work, it’s coordination’s fault

Coordination:

Daniele Bonacorsi (FacilitiesOps), Lothar Bauerdick (DataOps)

Info flow on CCRC within CMS

CERN IT Department



Dynamic info/doc based on twikies (CSA, CRUZET, CCRC, …)

Main CCRC twiki: https://twiki.cern.ch/twiki/bin/view/CMS/CMSCCRC08

Visit links therein, there are many nested…

You may find everything there indeed, though my feeling is that for a challenge anyway summing up twikies yields to no knowledge

there is always something missing, which you knew whom to ask in urgent ops within CCRC times, you did that, and *never* updated the twiki back…

IMPORTANT: twikies are maintained by operators of each sub-test

Check there to get fresh info, and give credit to the real workers!

Discussion through HN and CMS meetings/calls

- Daily 30' CSA/CCRC calls (similar to WLCG ones, just 30' after)
- DataOps meetings on Tuesdays
- CCRC (or PADA) Computing meetings on Thursdays
- FacilitiesOps meetings on Fridays

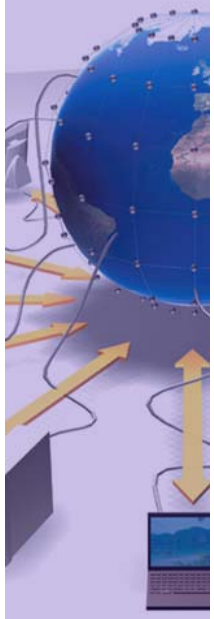
Highlights from CCRC daily ops for CMS at

https://twiki.cern.ch/twiki/bin/view/CMS/CCRC08-Phase2-OpsElog

How to communicate problems, and feedback

CERN IT Department

GS



Concerning daily communication - as foreseen -
- CSA+CRUZET+CCRC/2 was a more complex and time-consuming exercises for CMS wrt CCRC1 “alone” in Feb
  - Less time, so (unfortunately) less care in wide and good communication, less care in documenting on a daily basis, etc
    - Less elogs, less tickets, more phone calls, fast actions… and more coffees.

HN/mail was the main source of communication with sites
- GGUS + savannah soon after (could/should have been used more)
- ELOG usage much less than in CCRC/1

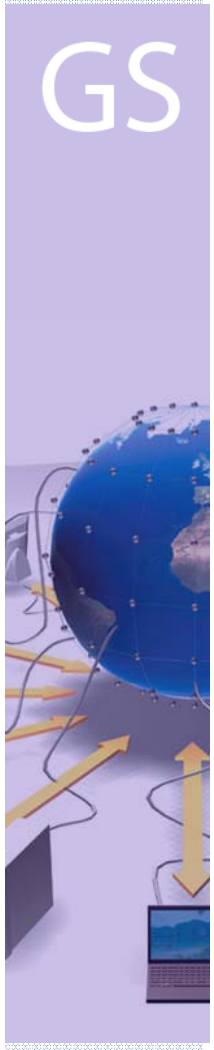
Hottest topics (e.g. Castor issues, actions after power cut, …) were addressed via cms.support ml, or phone/meetings, or raised at daily WLCG calls

Support from Grid Ops team and CERN was in general very good

Just one remark from CMS: you need to be more pedantic in notifying ALL interventions on critical services, regardless of the criticality, including those expected to be transparent…

Your remarks to CMS ?

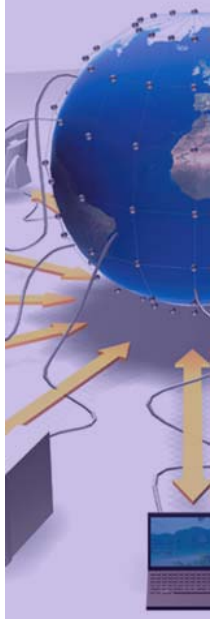
Information Sharing



- Extensive usage of e-books + GGUS.
  - LHCb has setup its own e-book for internal problem tracking http://lblogbook.cern.ch/CCRC08/ that has to be added to GGUS and to CCRC'08 e-book
    https://prod-grid-logger.cern.ch/elog/CCRC'08+Logbook/
- Internal daily operational meeting for discussing issues and planning activities
- Daily Meeting with Service Providers.
  - Bidirectional flow of the information VO←→ Service Providers
- Duplication of information/problems at different places.
  - Different level of details for different flavors of problems
  - Different target (LHCb specific / Service oriented)
  - Different usage of the information
  - Fully embraced the approach: "*A problem is not a problem if GGUS hasn't been open*"

History of a problem

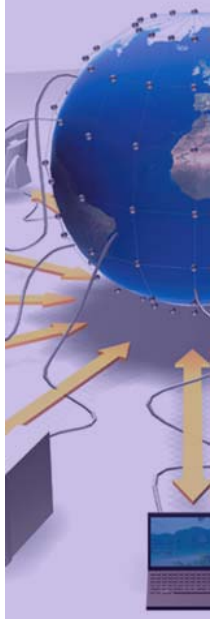
GS



- The Experiment Monitoring System is the first element of the chain.
- The operator watches at the quality of the activities and realizes some sites is badly behaving.
- The operator raises the alarm by logging in the internal logbook the observation or reporting at the internal ops meeting
- The expert/support drills down into details by using all tools he can rely on (see critical services talk) and isolates the problem
- He upgrades the logbook, opens a GGUS (also cc'ed to sysadmin directly) , fills up the service e-book in case it is a service related problem
- The expert follows the problem interacting with sysadmin/service providers at several meetings or/and through several channels.
- The expert keeps update the experiment by updating e-books and reporting to daily meetings
- The problem is solved, the GGUS is closed the entries on the various e-books are also updated.

Support received

GS



- No major complaints from LHCb about the support received. The reaction is usually extremely fast (few exceptions included)

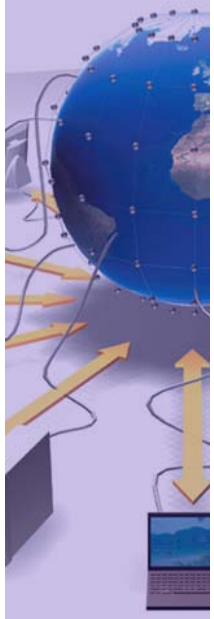
  BUT…

- The time to get the problem solved is sometimes longer than expected for problems felt as critical.
- LHCb application specific problems often need debugging “in situ” and would require access on the WNs by an LHCb expert but this is usually not possible, thus increasing the time for understanding problems
- Efficiency in fixing problem at the site still too strongly relying on local contact person. LHCb is a small collaboration and can’t rely on a army of people like other VOs
- Problem are too often spotted by the VO while they should have been trapped by local site fabric monitoring tool (gsidcap server down, dcap door dead)

ATLAS notification procedure

CERN IT Department

GS

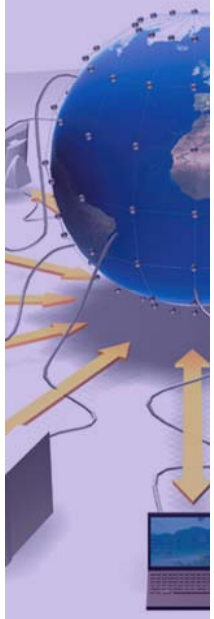


- For CCRC08, ATLAS used **elog** as primary placeholder for problem tracking
  - There is also ATLAS elog for internal issues/actions
- Beside elogs, email is sent to **cloud mailing list** + atlas **contact at the cloud**
- In addition GGUS ticket is submitted
  - For traceability
- ATLAS follows a strict regulation for ticket severity
  - TOP PRIORITY: problem at T0, blocking all data export activity
  - VERY URGENT: problem at T1, blocking all data import
  - URGENT: degrade of service at T0 or T1
  - LESS URGENT: problem at T2 or observation of already solved problem at T0 or T1
- Shifters (following regular production activities) use **GGUS** as main ticketing system

ATLAS notification procedure

CERN IT Department

GS

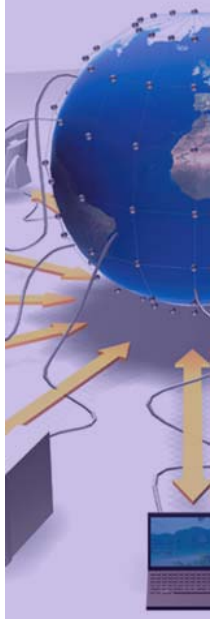


- Issues and actions are discussed in daily ATLAS meetings
  - Useful to have day-by-day
- Also the Alarm Mailing list at CERN has been used
  - Saturday morning problem at T0
  - From the ATLAS perspective it worked
  - Internally, looks like it took a different route than expected
- Got used to look at the IT service status board
  - I find it useful
- Very hard to follow the EGEE broadcasts for site downtimes
  - Also difficult to visualize the relevant ones in GOCDB
  - Now ATLAS dashboard collects the RSS feeds, parses, displays
- ATLAS is happy with support and problem follow up at sites during CCRC08
  - Very fast response and problem fixing

More consideration from ATLAS

CERN IT Department

GS



- ATLAS share some of the comments of other VOs
  - Some infrastructure problems should be trapped by fabric monitoring
  - It is important to notify to the VO all type of intervention to critical services, does not matter how transparent it is
- In addition, sites and EGEE operations should try to make sure ATLAS SAM critical test do not fail at the site
  - We are talking about 2 test:
    - A job van be submitted successfully to the CE
    - A file can be copied successfully to the SE
  - Basic functionalities