

Making ATLAS Central Services Reliable

B. Koblitz - CERN-IT



Overview



- What are the Atlas central services?
- Changes in the software development cycle for more reliable services
- Reliable hardware setups
- Monitoring: Noticing when things went wrong.
- Shifts, Expert on Call: How to fix things.



Atlas Central Services



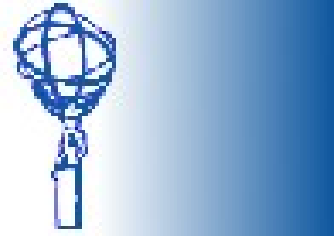
- Tier 0: Processing of data coming from the pit, Registration in Grid
- Production System: Database, Panda
- Site Services: Transfers out of CERN, within Clouds
- Central Catalogues: Location of datasets, subscriptions of datasets to sites
- AMI: Dataset metadata
- ... Many services like LFC, ProdDB (Oracle) I am not going to talk about here.



Towards Production



- Big changes for Atlas Distributed Computing in last year: Split into Development and Operation teams
 - Services are run by operations team, get releases from the developers
 - Quattorized base installation of service nodes, applications are installed via apt
- Development changed focus
 - Software releases with integration and testing cycle
 - Switch to RPM-based releases
 - Redesign: Changed to Oracle as Central Catalogue backend



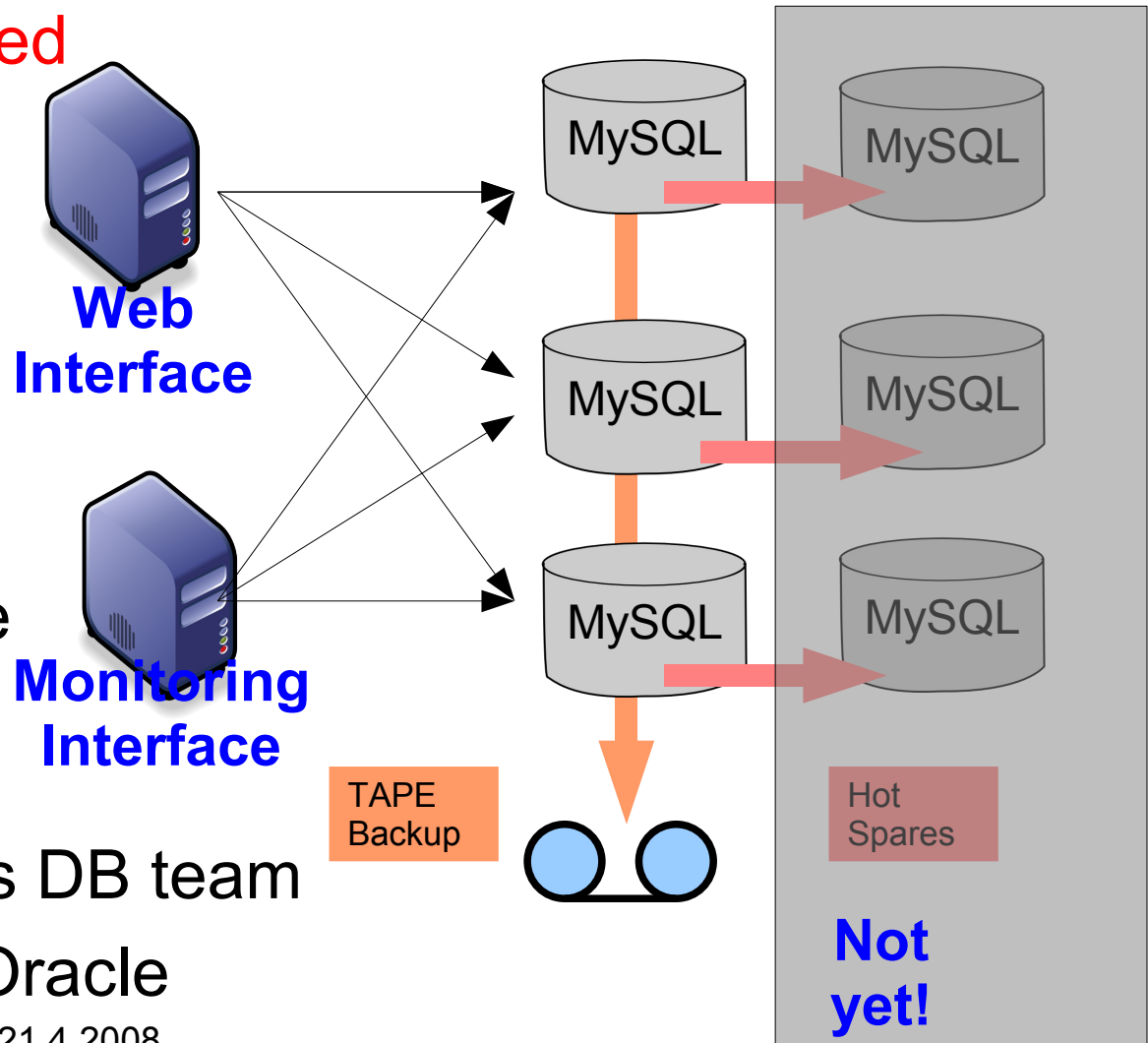
Some Example Service Setups



Panda



- **Panda: Pilot-based job submission system**
- Keeps Info on Pilots, jobs in MySQL DBs
- MySQL DBs crucial: hot spare and regular backup to HSM
- Operated by Atlas DB team
- Long-term goal: Oracle

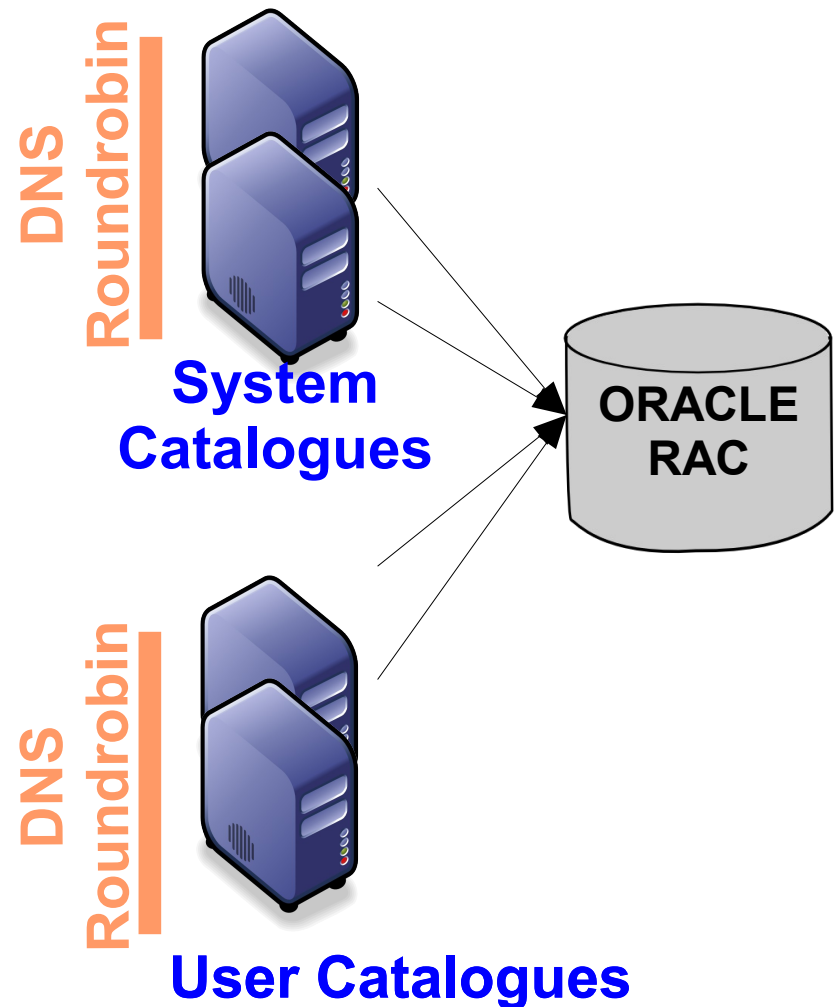




Central Catalogues



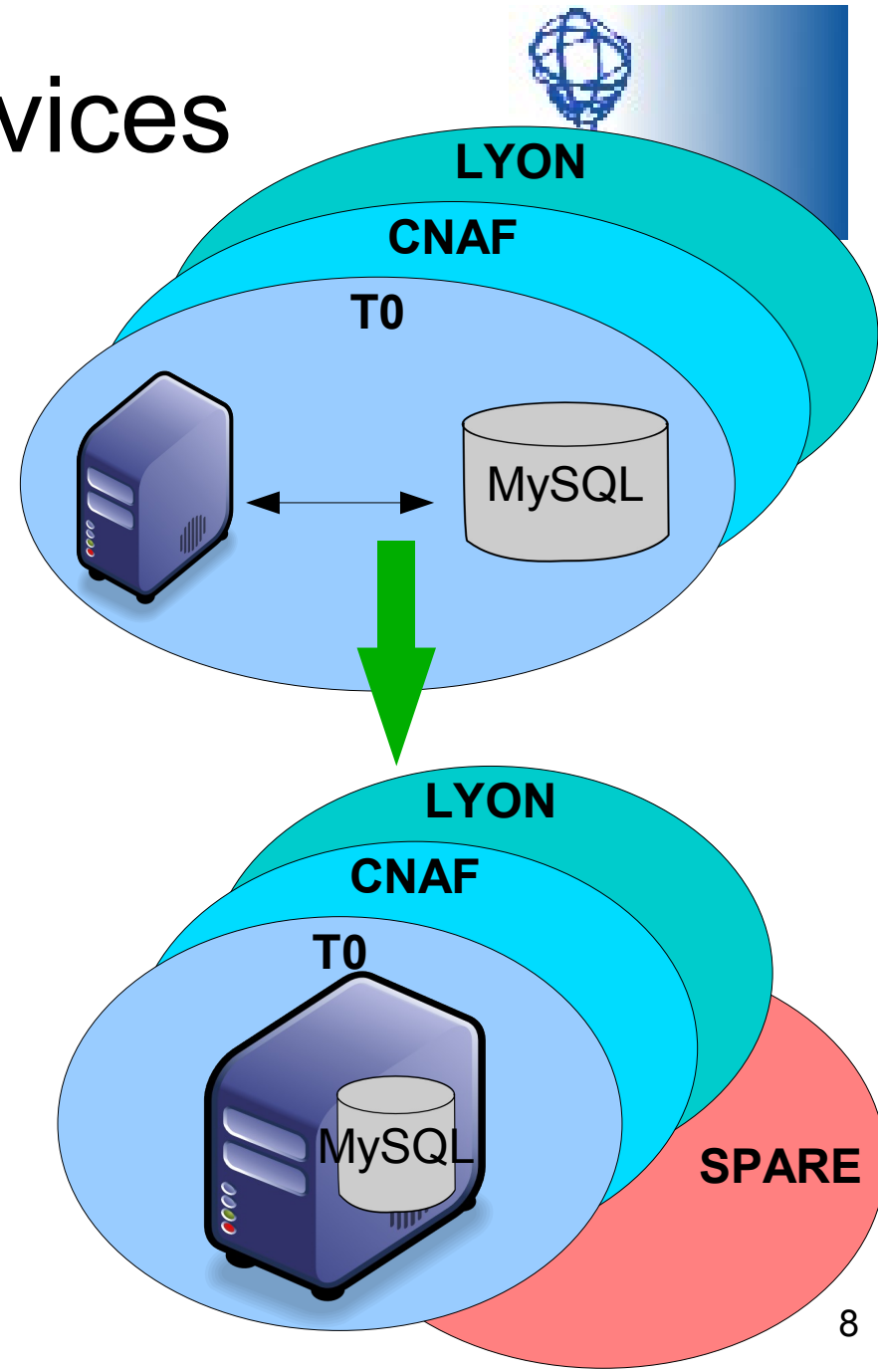
- **Central store of dataset locations, subscriptions**
- Separate systems for users and central services
- DNS round-robin used
- Moving to more powerful batch-type nodes
- Implementing actuators which take broken head-nodes out of DNS setup





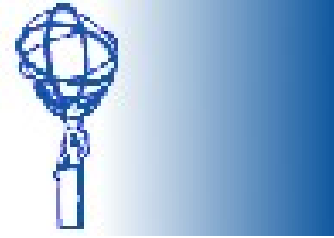
Site Services

- **Aka VO-Boxes:**
Transfer of files from T0 and within Cloud
- Moved from 2 boxes with SS and MySQL to one (more powerful) box
- MySQL DB only holds transient state
- Have separate testbed, will have integration testbed





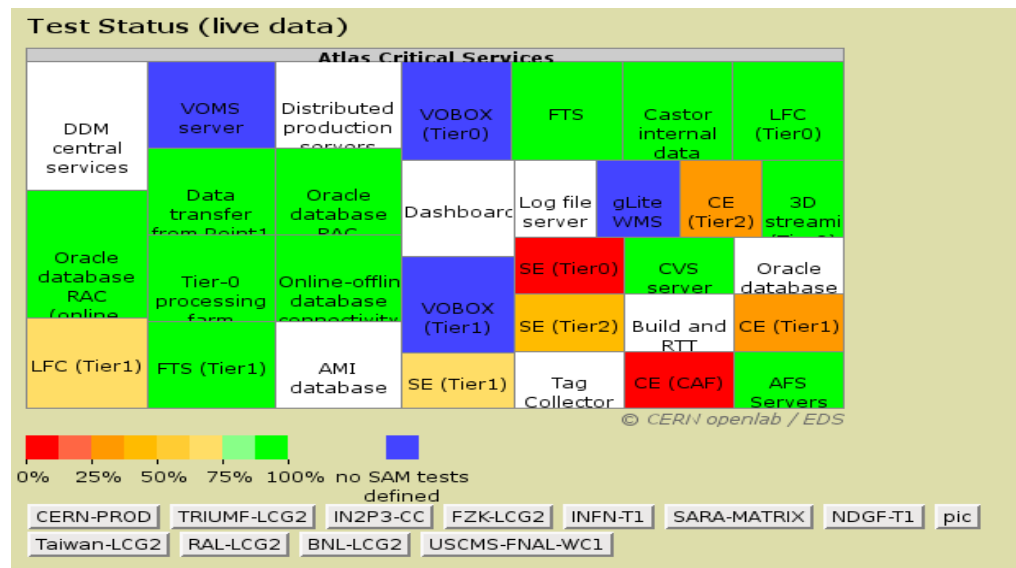
Monitoring



- **What** are the Services doing?

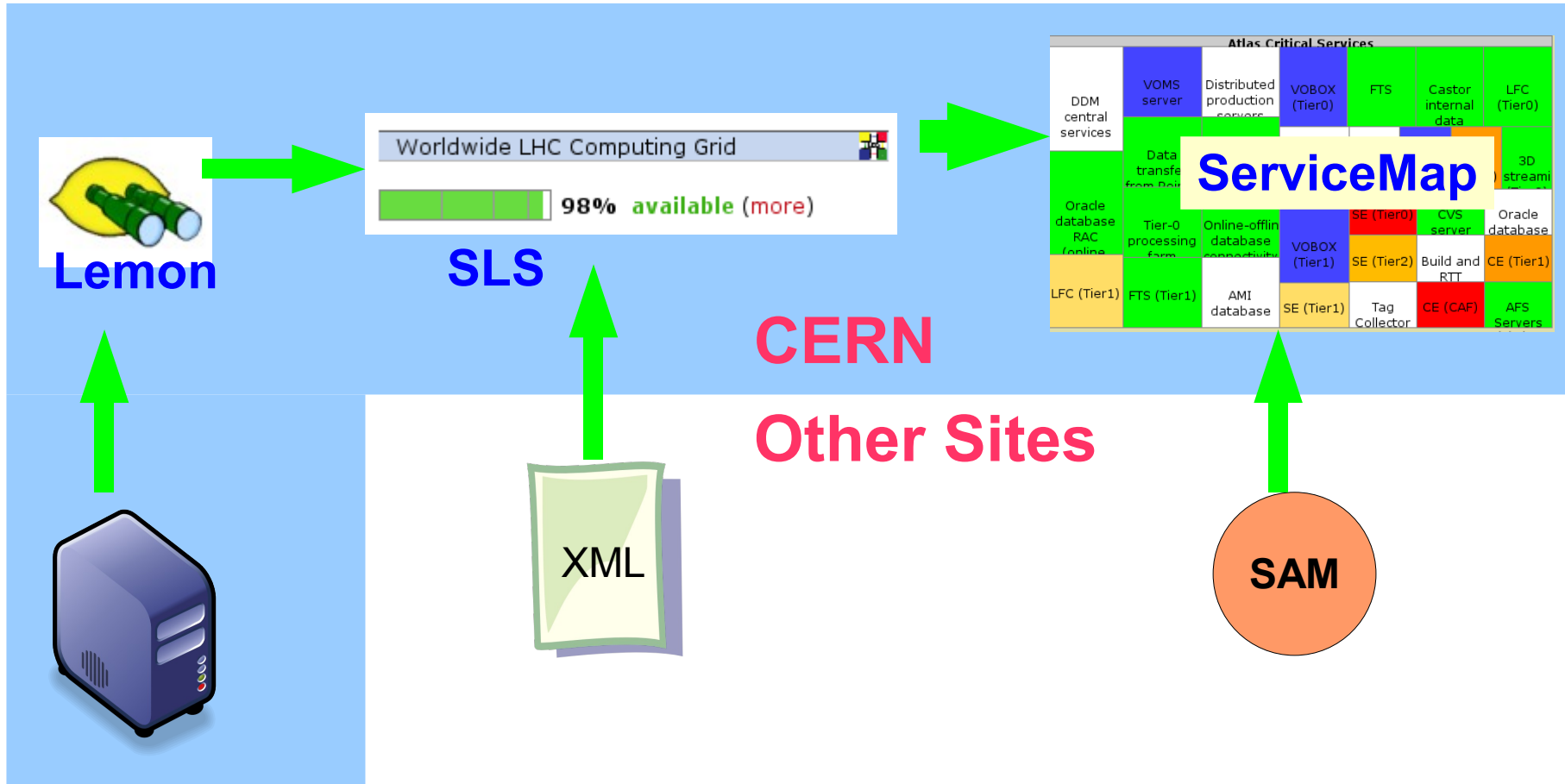


- **How** are the Services doing?





Collecting Monitoring Info



**CERN
Other Sites**



Lemon: Example



- Definition of a sensor for the site services, the exception tries to restart the service before sending an alarm:

```
# Check DQ2 Site-Services are running
# by checking their log-files are updated
"/system/monitoring/metric/_4031" = nlist(
  "name", "dq2_ss_timestamp_check",
  "descr", "DQ2 Site Services",
  "class", "file.sslmtime",
  "param", list("file", "/tmp/dq2.log"),
  "period", 60, # One sample in 60 seconds
  "smooth", nlist("typeString", false,
    "maxdiff", 0.0, "maxtime", 600),
  "active", true,
  "latestonly", false,
);
```

```
"/system/monitoring/exception/_30145" = nlist(
  "name", "dq2_ss_wrong",
  "descr", "DQ2 Site Services wrong",
  "active", true,
  "latestonly", false,
  "importance", 2,
  "alarmtext", "dq2_ss_wrong",
  "correlation",
    "(4031:1 > 600) || (4031:1 == -1)",
  "actuator", nlist("execve",
    "/etc/init.d/dq2-siteservices restart",
  "maxruns", 3,
  "timeout", 0,
  "silent", true,
  "active", true)
);
```



SLS: AMI



The screenshot shows a web browser window titled "Service Level Status overview - Iceweasel" with the URL <https://sls.cern.ch/sls/service.php?id=ATLAS-AMI>. The page displays the following information:

- Service information:**
 - full name: **ATLAS-AMI**
 - group: IT-GS
 - site: CERN
 - web site: <http://ami.in2p3.fr/AMI/>
 - manager: **Alessandro Di Girolamo**
- Service availability (more):**
 - availability:
 - percentage: 100%
 - status: **available**
- Additional service information (more):**
 - ccami01.in2p3.fr: 100
 - ccami02.in2p3.fr: 100
- Part of (subservice of):** none / not declared
- Subservices:** none / not declared
- Clusters, subclusters and nodes:** none / not declared
- Depends on:** none / not declared
- Depended on by:** none / not declared

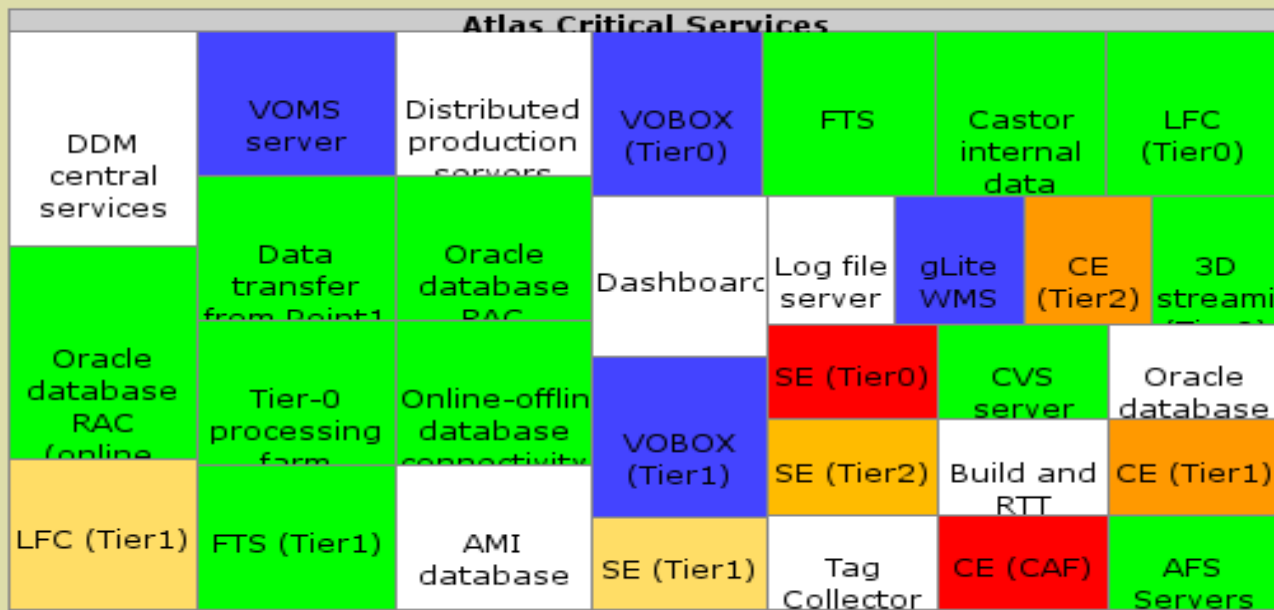
The page also features a navigation menu with links for Home, Search, KPIs, Tags, Admin, Documentation, and Help. A vertical sidebar on the left contains the text "Service Level Status overview". The browser's status bar at the bottom shows "Done" and the address "sls.cern.ch".



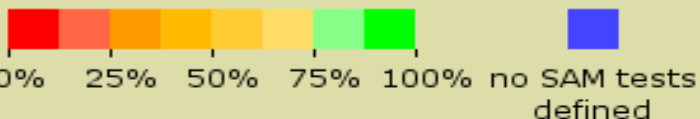
GridMap



Test Status (live data)



© CERN openlab / EDS



- CERN-PROD
- TRIUMF-LCG2
- IN2P3-CC
- FZK-LCG2
- INFN-T1
- SARA-MATRIX
- NDGF-T1
- pic
- Taiwan-LCG2
- RAL-LCG2
- BNL-LCG2
- USCMS-FNAL-WC1



Experiences with Monitoring



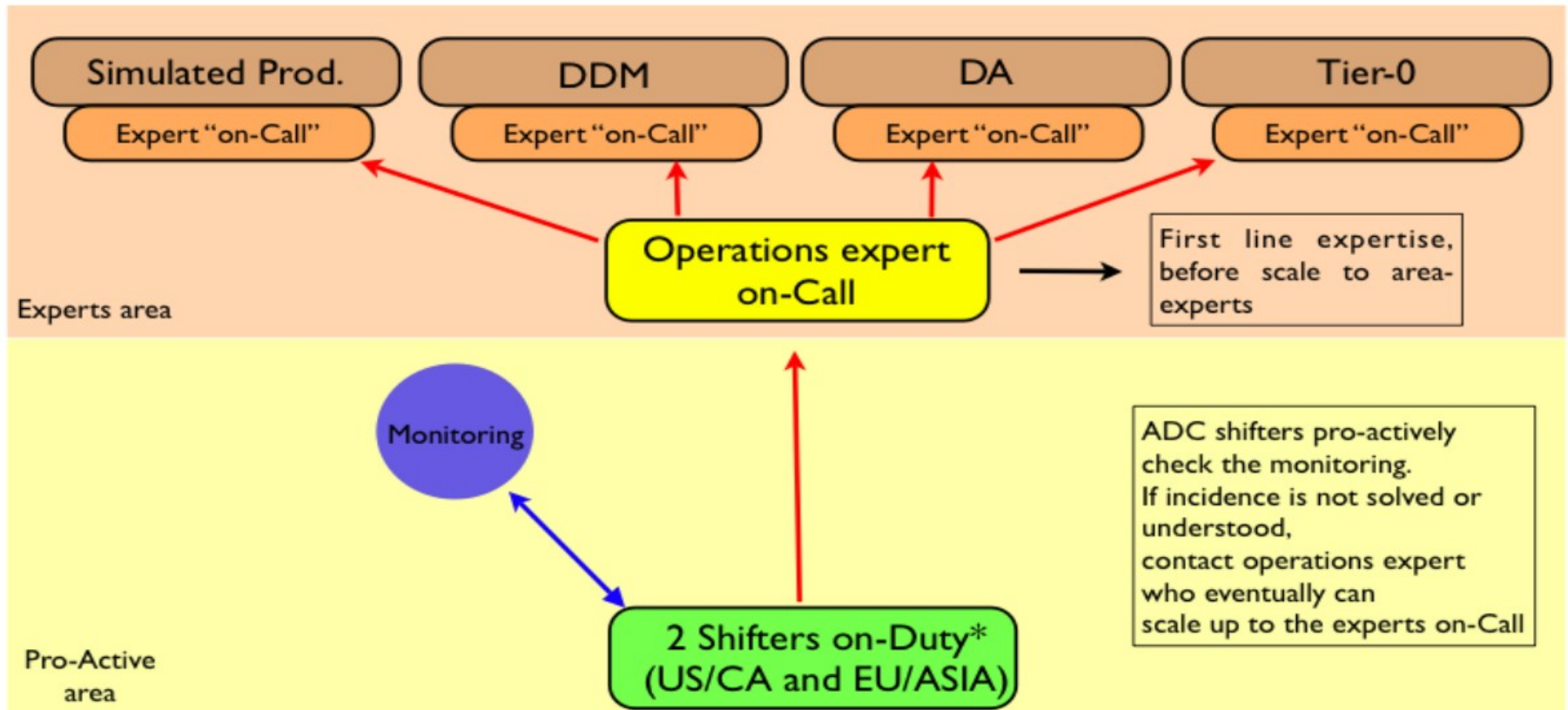
- Very positive feedback from external developers
- New SLS status service was provided by AMI team within hours
- SLS can be used for more than service availability: Show quotas for ATLCAL pool, there
- Currently we provide SLS information via AFS webospace + acron jobs, would like to go to a standalone server



ADC Shifts

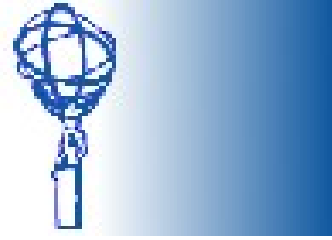


ADCoS Infrastructure





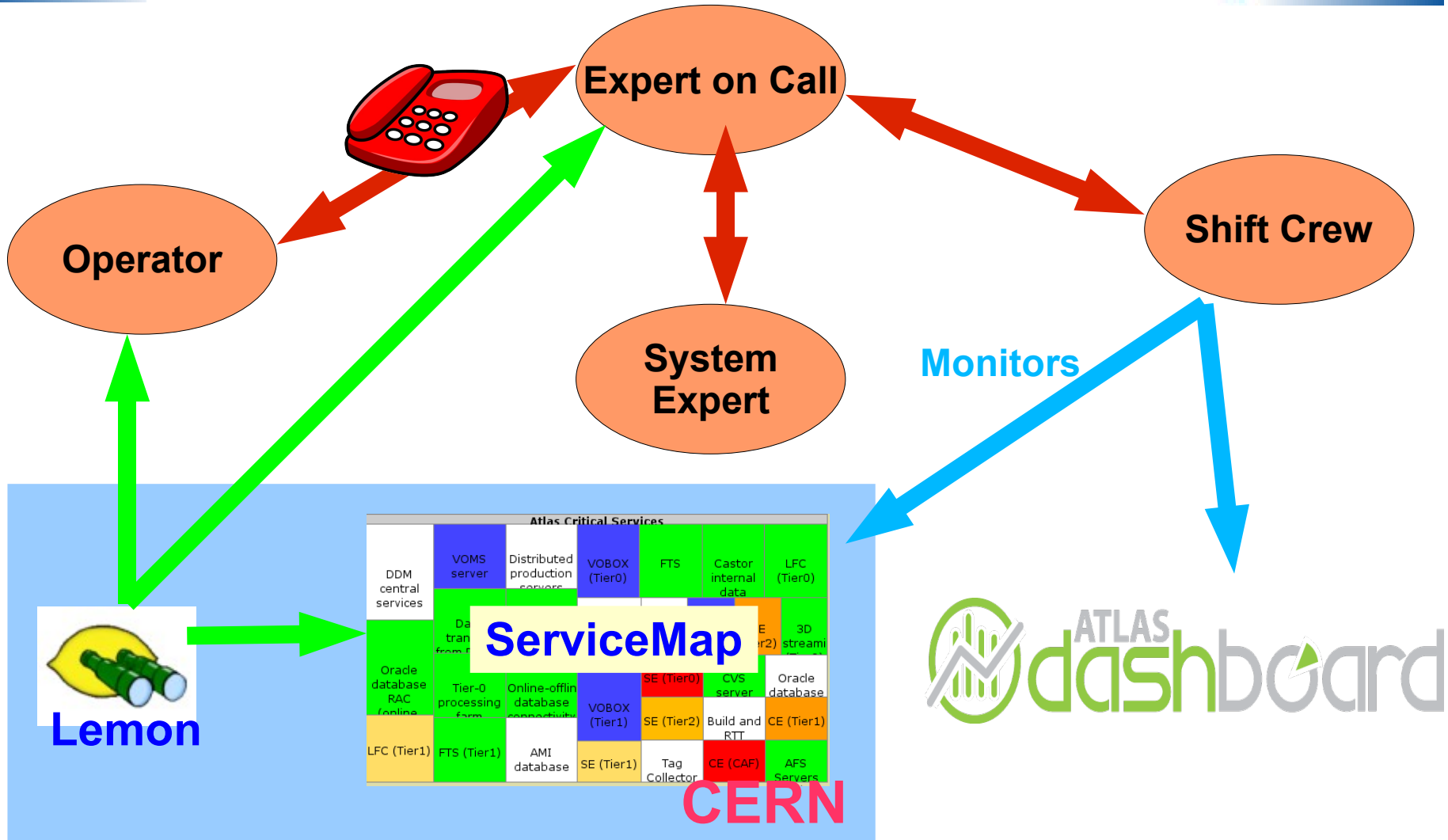
Alarms



- All Alarms at CERN seen by operator and Expert-On-Call (not the shift-crew)
- Operator should have instructions to solve as many problems as possible
- When the operator cannot solve a problem, he will contact the Expert-On-Call
- All communications to the operators or component experts go through the Expert-On-Call



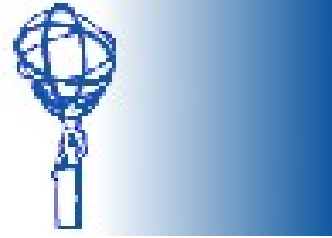
Reacting to Alarms



CERN



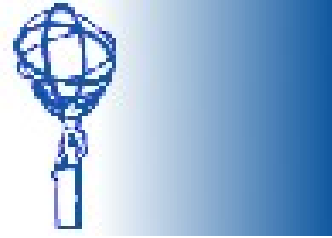
Still a lot to do...



- Several services are still in preproduction stage and will need to be integrated into the production environment
 - Deletion Service
 - Prestaging Service
 - Bamboo
- All these new services need sensors
- There are still many procedures to be written for the shift crews, Expert-On-Call and the operator.



Summary



- In the past year there was a large, concerted effort to make the Atlas Central Services Production-Ready
- Introduced different Development and Operations teams, introduced Release Cycles
- Moving to reliable hardware, quattorized base-installation
- For monitoring the health of the services themselves we depend on Lemon at CERN, SLS and SAM otherwise. This information is complemented by the Dashboard.