

THE ALICE ONLINE-OFFLINE FRAMEWORK FOR THE EXTRACTION OF CONDITIONS DATA

COMPUTING IN HIGH ENERGY AND NUCLEAR PHYSICS

Prague, Czech Republic, 21-27 March 2009

Alberto Colla, Jan Fiete Grosse-Oetringhaus,
Chiara Zampolli

for the ALICE Collaboration



Outline

*The **Shuttle** is the ALICE Online-Offline software framework dedicated to the extraction of conditions data – calibration and alignment – during data taking, running detector specific procedures called **preprocessors***

Outline:

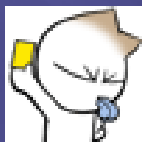
- Motivations
- Shuttle flow
- Detectors Preprocessors
- Monitoring of the Shuttle
- The Shuttle Test SetUp
- The Shuttle in Production
- Remarks, Summary and Conclusions

Motivations

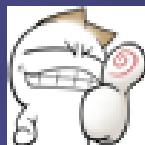


ALICE Conditions Data

- During data-taking, the **18 ALICE detectors** interact with five online systems – **DAQ** (Data Acquisition), **DCS** (Detector Control System), **ECS** (Experiment Control System), **HLT** (High Level Trigger), **Trigger**
- The result of data taking are raw data, stored in the Grid
- From the raw data, the **conditions data** – alignment, calibration, environment – needed for reconstruction should be extracted
 - An extra pass over the raw data would be necessary before reconstruction
→ problematic, too expensive in terms of computing resources
- Moreover, not all the conditions data can be retrieved from the raw



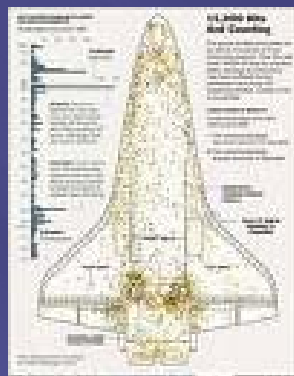
*Conditions data should be extracted at the same time
as data taking*



The Shuttle Concept

- Given the number of subdetectors and of online systems, the need for a common infrastructure to coordinate the extraction, the gathering, and the publication of the condition data came out
- Online machinery (DAQ, DCS, ECS, HLT, Trigger) are shielded to the external world by firewall
- The conditions data for reconstruction should reside in an easily accessible place
- The aim was to create a robust framework that allows temporary failures of all online system as well as malfunctional code without losing data

The Shuttle framework has been developed



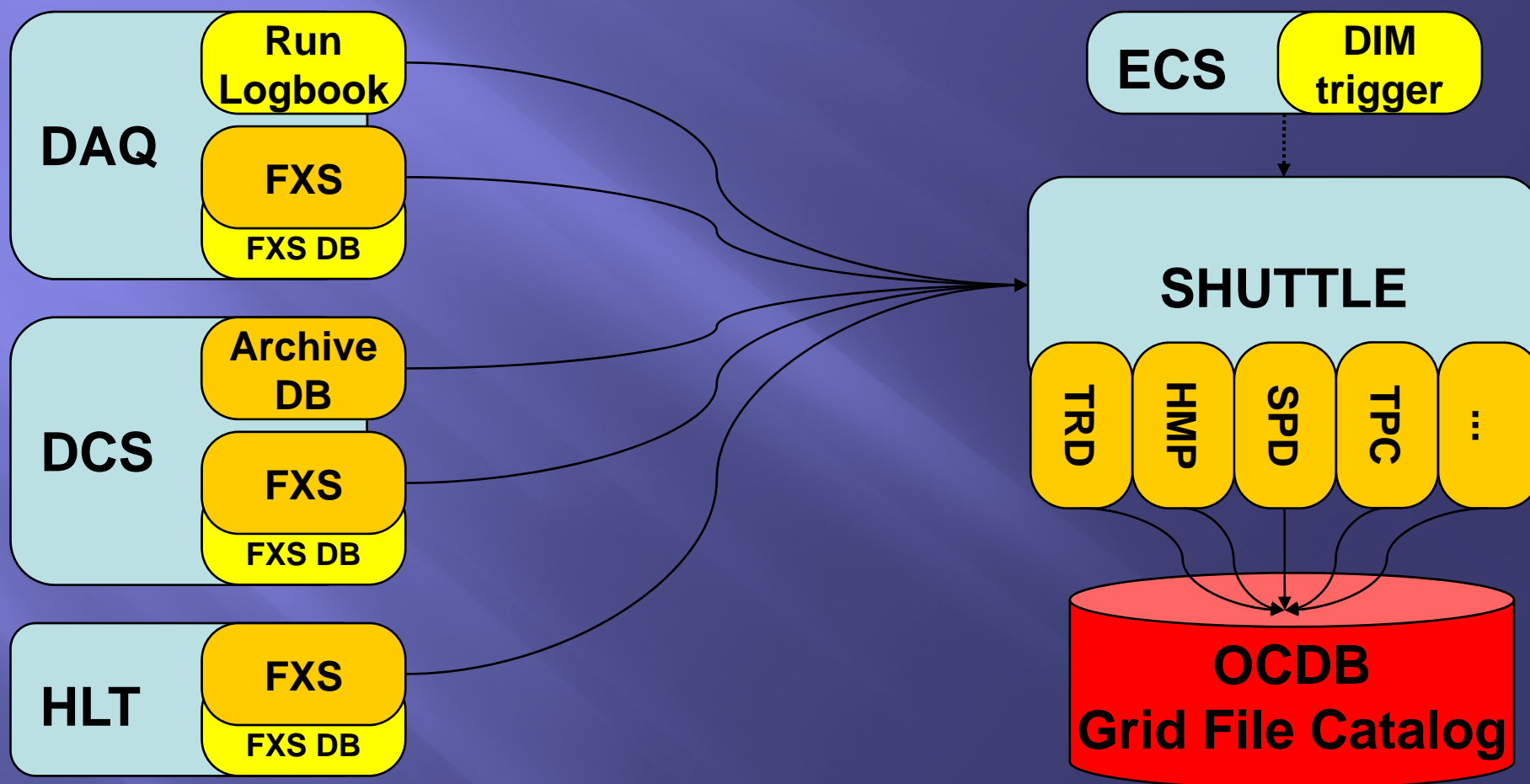
The Shuttle Flow



What the Shuttle does

- ~ **Copying** of data in any format produced by the online systems DAQ, DCS and HLT for each subdetector;
- ~ **Preprocessing** of the data, e.g. consolidation, fitting;
- ~ **Reformatting** to ROOT format;
- ~ **Storing** in the Offline Conditions DataBase (OCDB);
 - ~ Not a relational database, but a set of AliEn entries pointing to ROOT files stored in various SE of the Grid
- ~ Indicating that a given **run** has been **processed**, which is a precondition to starting the reconstruction.

The Shuttle General Schema

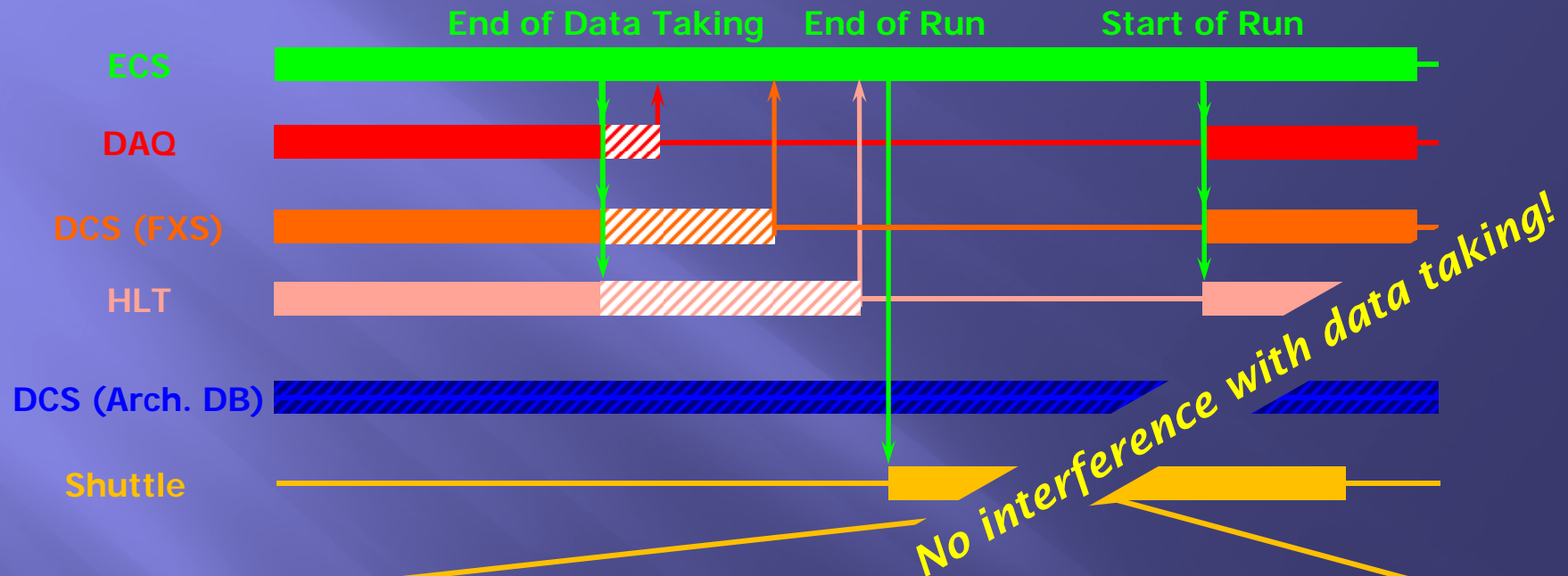


No alternative system to extract data (especially online calibration results) between data-taking and first reconstruction pass!

The Shuttle Program Flow

- Reading of the configuration parameters from an **LDAP database**
- **Triggering** of the Shuttle by ECS at every end-of-run
- Access to the **ECS logbook**, to retrieve run information (start of run, end of run...) and the list of active detectors
- Access to the **online systems** to retrieve the conditions data
 - parameters monitored and archived by DCS
 - parameters created during data-taking by dedicated Detector Algorithms (DA), written in files saved in dedicated File Exchange Servers (FXS) stored by the detectors during the run
- Triggering of the detector-specific codes – **preprocessors** – in charge of extracting the conditions data
- Storing of the conditions data in the **OCDB**
 - Used in reconstruction
- Storing of the **Reference Data**
 - OCDB-like database, not used in reconstruction

Sequence Diagram



Loop over all detectors (+ GRP and HLT)
Registration of conditions data files in AliEn
Interfaces with info providers

ACORDE	EMCAL	GRP	HLT	HMPID	FMD	ITS(*)	MUON(**)	PHOS(***)	PMD	T0	TOF	TPC	TRD	V0	ZDC
--------	-------	-----	-----	-------	-----	--------	----------	-----------	-----	----	-----	-----	-----	----	-----

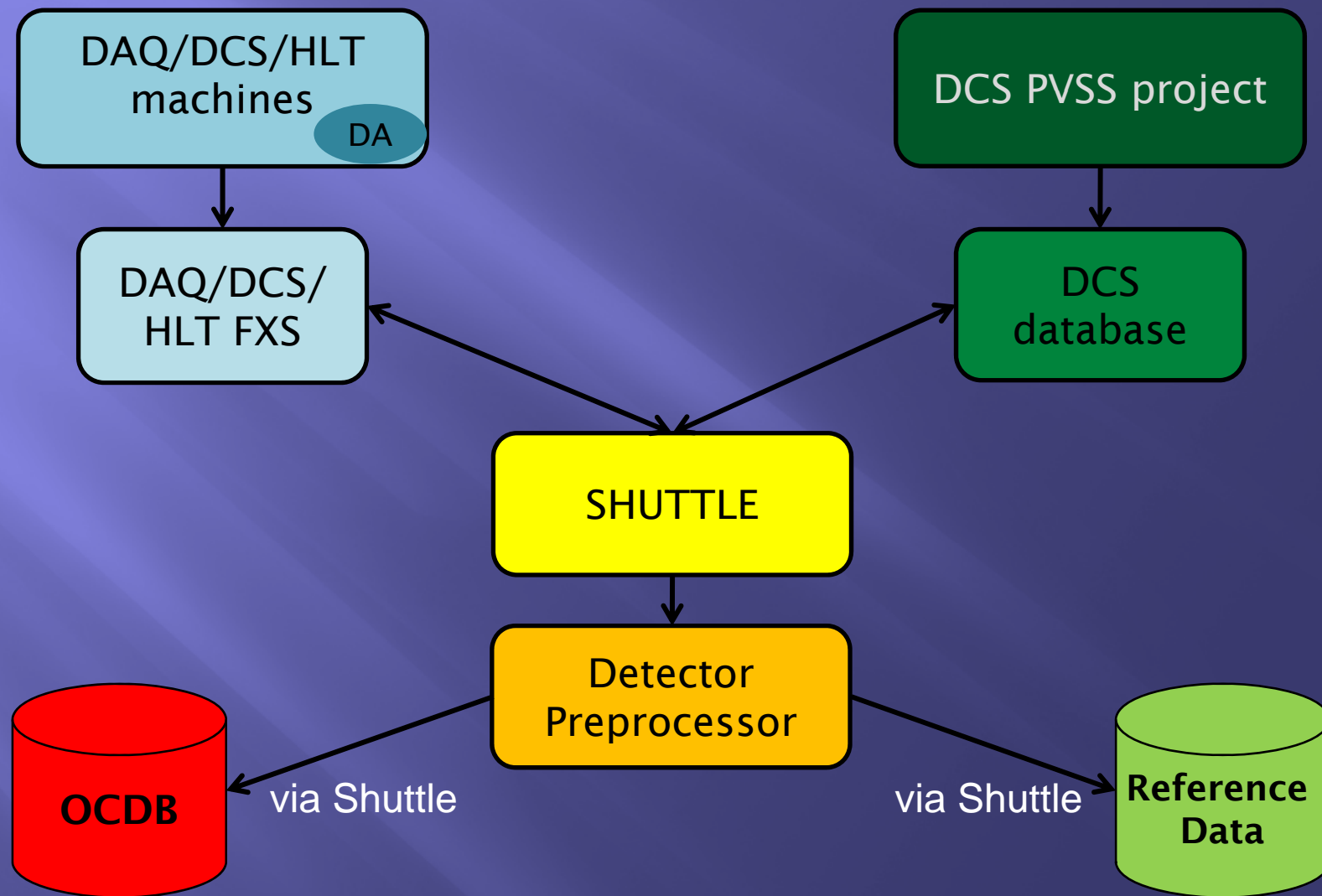
 Detector preprocessors

(*) ITS = SPD + SDD + SSD

(**) MUON = MCH + MTR

(***) PHOS = PHS + CPV

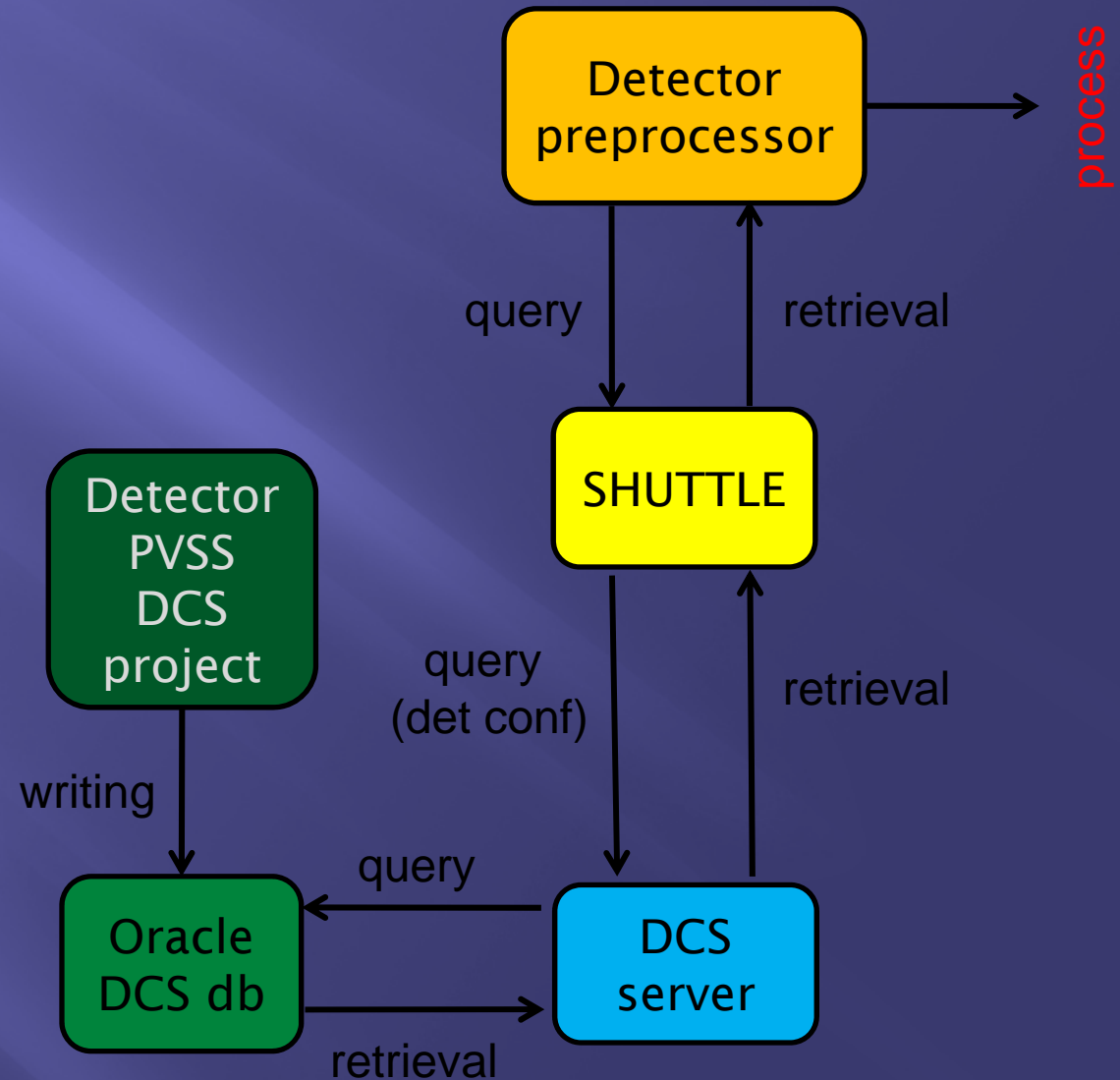
The Shuttle Data Flow – Schema per Detector



Shuttle Retrieval of Conditions Data – DPs

DCS DP's:

- Conditions parameters monitored and archived by DCS
 - Stored in an Oracle DB
 - Each associated to a **Data Point** (DP) name (and optionally, an alias name)
 - Stored value \leftrightarrow timestamp of the storage
- Dedicated server-client communication protocol used in the Shuttle



Shuttle Retrieval of Conditions Data – DAs

- ~ Conditions Data from **DAs**
 - ~ **Detector Algorithms** (DAs) are program that runs on the online systems – DAQ, DCS, HLT – producing conditions data that are then stored in any type of files in the corresponding File Exchange Server (FXS)
 - ~ The Shuttle locates in the **FXS** the files requested by the detectors for that particular run, and retrieves them via secure copy (scp)

*see talk by V. Chibante Barroso,
Tue 24th, parallel session 1B*

Detectors' Preprocessors



The Detectors' Preprocessors

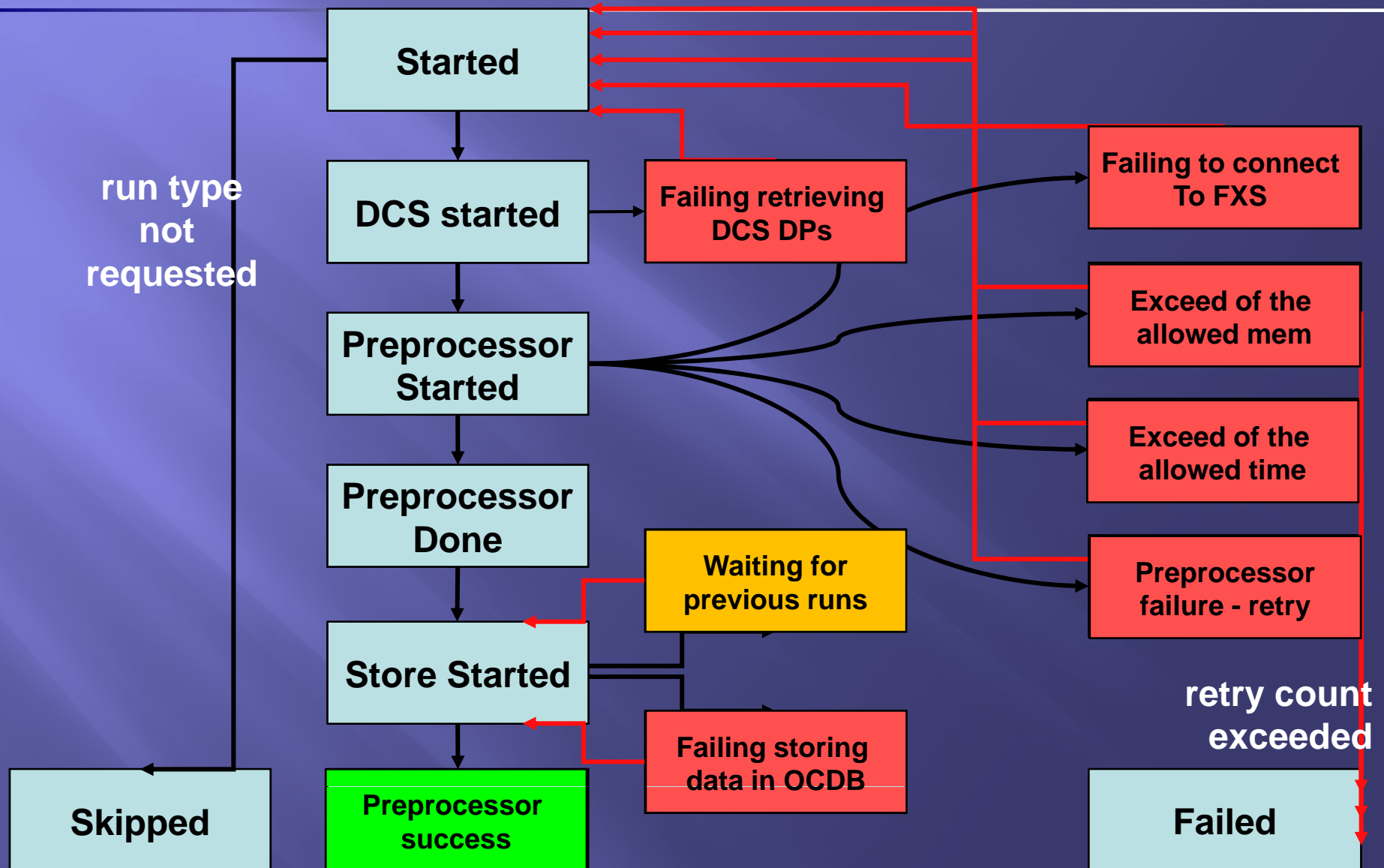
- One **Preprocessor** per detector
- Preprocessors started by the Shuttle according to the ECS logbook
- Preprocessors run **sequentially**, but **independently**
- Preprocessors process the conditions data (from the DCS archive, and the DAQ/DCS/HLT FXS):
 - **Query** the conditions data
 - **Reformat/process** the conditions data, if necessary
 - **Store** the conditions data in the OCDB

Handling of the Detectors' Preprocessors

At every run, for every preprocessor:

- ~ **Forking** of the Shuttle process
 - ~ the child process executes the current preprocessor
 - ~ the parent process executes the main Shuttle, monitoring the child
 - a crash of one preprocessor does not affect the main process
 - no memory leaks can occur
- ~ Storing by the Shuttle of the **processing status** (error recovering)
- ~ Monitoring of **memory consumption** → killing of the child process in case it exceeds a defined critical threshold
- ~ Monitoring of the **processing time** → killing in case it times out
- ~ If failing, a certain number of later **retries** is allowed

Preprocessor Status Flow



Monitoring of the Shuttle



MonALISA Monitoring

- ~ The status of the Shuttle processing is monitored using the **MonALISA** system
- ~ Every change in the Shuttle/detector preprocessors states are sent to the MonALISA service, and visualized by the ALICE MonALISA repository:

<http://pcalimonitor.cern.ch/shuttle.jsp?instance=PROD>

*see talk by R. Voicu,
Thu, 26th, parallel session 4
see poster by C. Grigoras,
Thu, 26th, 067*

time

Access to the history of the status

time

08.13 18 3.09

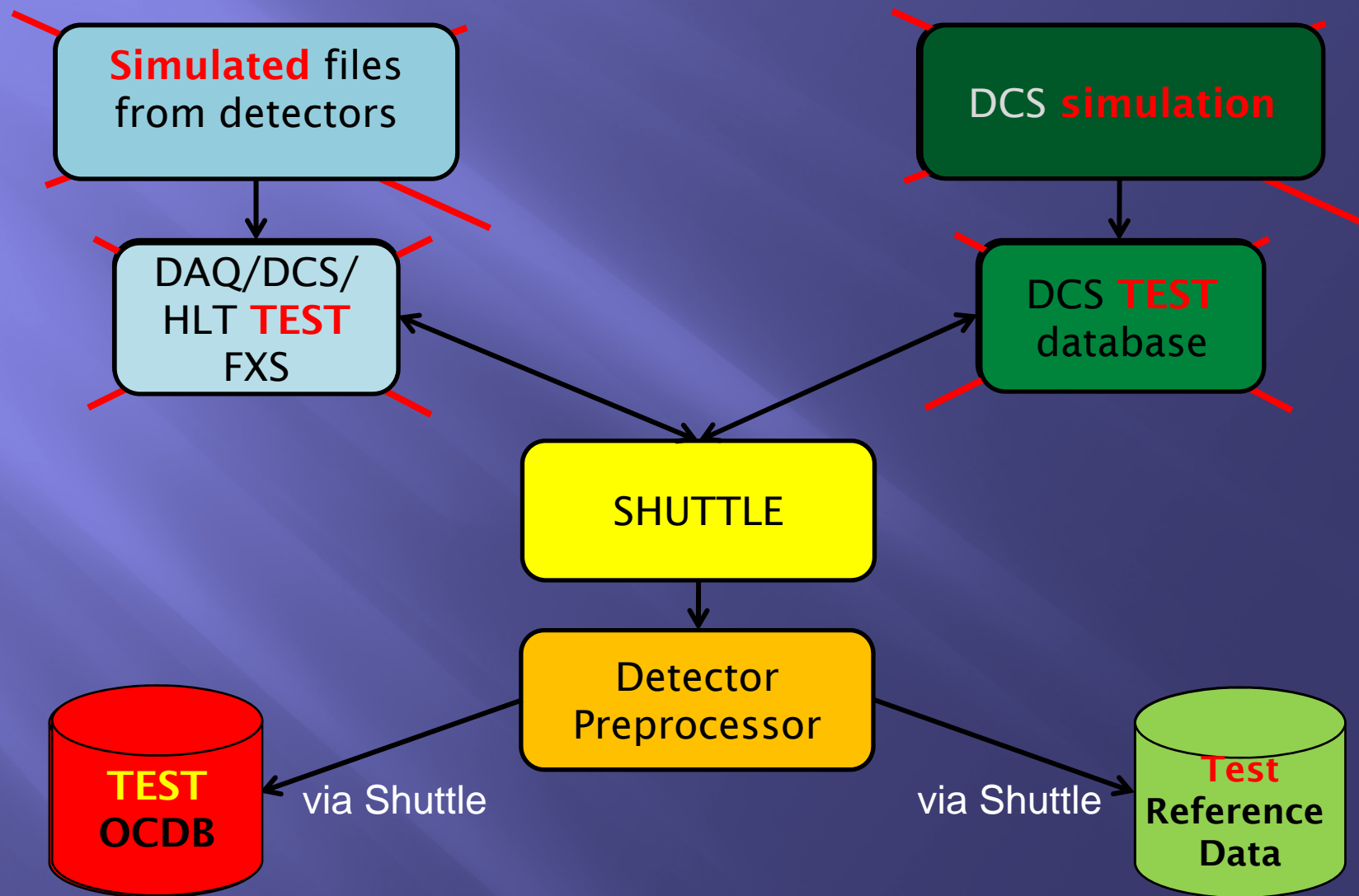
The Shuttle Test SetUp



Shuttle Test Setup

- ✿ In order to check the current status of the detectors preprocessors, a Shuttle **Test Setup** has been developed, in parallel with the production (on real data taking) setup
 - ✿ A *different LDAP configuration* is used for the Shuttle
 - ✿ *Artificial runs* are created
 - ✿ A *dedicated MonALISA* monitoring web page is filled
 - ✿ Working with the *latest available AliRoot* (ALICE Offline software for simulation, reconstruction and analysis)
 - ✿ *Failover tests* included

The Test Shuttle Data Flow – Schema per Detector

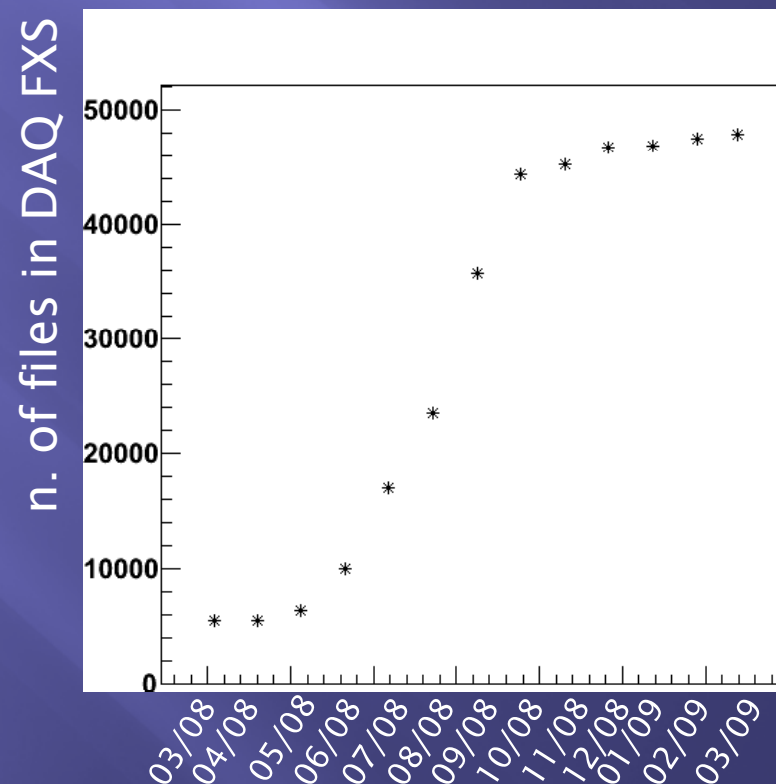


The Shuttle in Production



Shuttle in Production

- The Shuttle is working smoothly since December 2007
- 44346 runs successfully processed since then
- Figures presented hereafter from Feb '08 up to Mar '09 (cosmics, calibration, LHC data)



Total ~90 GB

OCDB Size

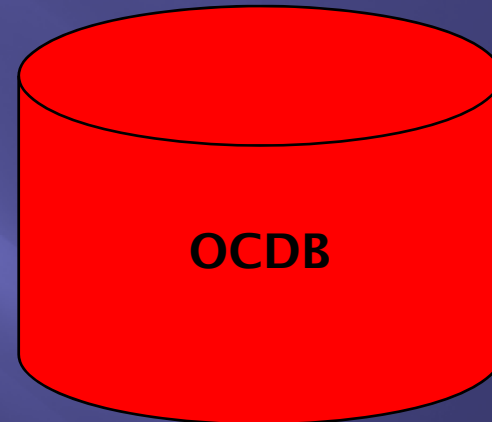
2008 data taking

~ OCDB

~ 37000 files

~ Total size: 5.2 GB

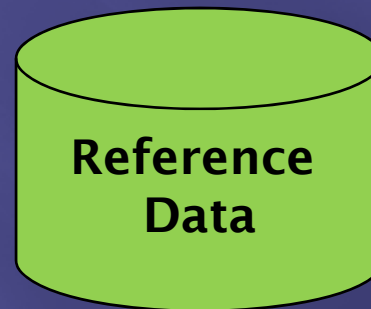
→ *Much smaller than DAQ FXS*



~ Reference Data

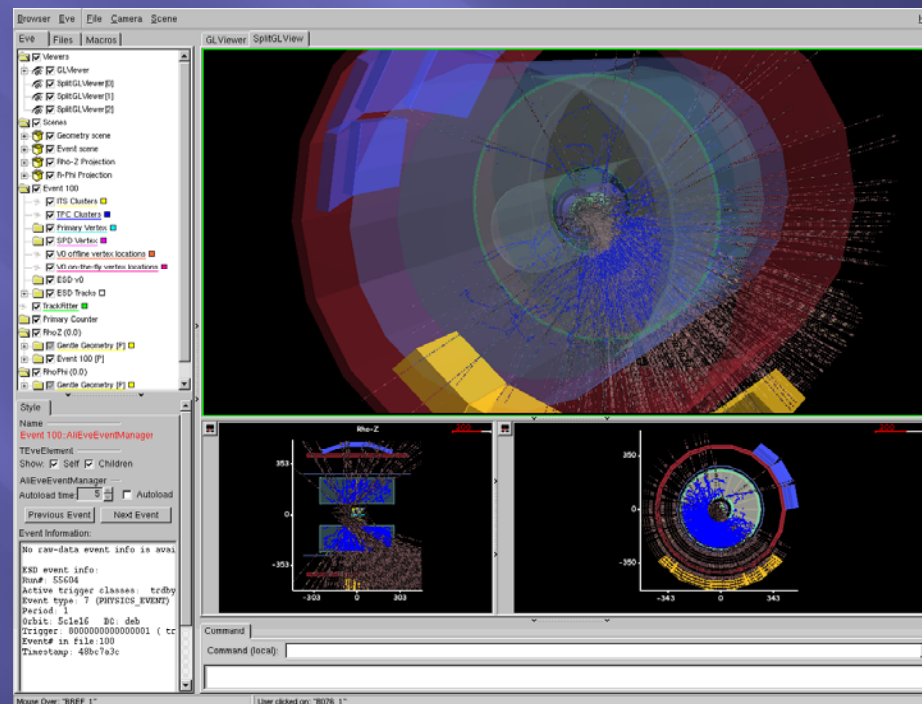
~ 9500 files

~ Total size: 3.2 GB



Quasi-Online Reconstruction

- The **quasi-online reconstruction** reconstructs events directly during a run and visualizes them in the control room
 - **Fast analysis** of the data being taken
 - **Fast visualization** using AliEve
- The ALICE Quasi-Online Reconstruction accesses to information needed to steer the reconstruction **via the Shuttle**



02/09/2008

Remarks, Summary and Conclusions



Feedback and Problems

- ~ During the development of the framework, and during data taking, **problems** arose, **issues** came up, **fixes** were put in place, new **features** and **improvements** were implemented:
 - ~ Failover tests added
 - ~ Mail to subdetector in case of failure
 - ~ Memory monitoring
 - ~ Check on the file checksum when retrieving from FXS
 - ~ Proper handling of runs according to ECS logbook
 - ~ Selective run type processing
 - ~ Possibility to test DAs on real data, w/o writing OCDB data
 - ~ ...

Summary and Conclusions

- The **Shuttle** – ALICE Online-Offline Framework for the Extraction of Conditions Data – serves as the only possible link between the data produced in the online systems – DAQ, DCS, ECS, HLT, Trigger – and the experiment subdetectors
- The detectors' procedures for the processing of the conditions data and the gathering of the conditions data needed for reconstruction – the **preprocessors** – are handled sequentially and independently by the Shuttle
- The Shuttle is also in charge of storing the condition data in the ALICE **Offline Conditions DataBase** (OCDB)
- A monitoring service – **MonALISA** – is used to follow the processing
- A parallel **test setup** allows debugging of the detectors' code
- Starting **from Dec 2007**, the Shuttle has been **running** smoothly on cosmics runs and LHC data, collecting many conditions data
- The overall experience and interaction with the online systems and the detectors has shown a high **efficiency** and an extreme **stability** of the framework



Back-Ups

Preprocessor Exit States

- **FXS Error**: when a connection to one of the FXSs failed → the preprocessor will be run again at next trigger
- **Preprocessor Error**: the preprocessor failed processing the current run → a predefined number of retries is allowed
- **Preprocessor TimeOut**: the preprocessor exceeded the allowed processing time → a predefined number of retries is allowed
- **Preprocessor OutOfMemory**: the preprocessor exceeded the allowed memory usage → a predefined number of retries is allowed
- **Preprocessor Done**: the processing ended successfully → no retry needed ☺

Other Possible Error States

- **DCS Error**: if the retrieving of the DCS data from the DCS archive fails → the preprocessor will be run again at next trigger
- **Store Error**: if there was an error in transferring the conditions data to the OCDB → the preprocessor will be run again at next trigger
- **Store Delayed**: if the validity of the object is infinite and some previous runs are still open → the preprocessor will be run again at next trigger

Reference DB

- ~ It is an OCDB-like database, where to store what could be useful for a “manual” debugging of the physics data
- ~ The data stored in the Reference DB are NOT used during reconstruction
- ~ While the OCDB objects are replicated to the various Grid sites to make them easily accessible, the Reference data are not

The LDAP Shuttle Database

- ~ The Shuttle configuration parameters are read from an LDAP database
 - ~ Global configuration:
 - ~ Main Shuttle running parameter (e.g. maximum number of retries for a preprocessor to be run)
 - ~ ECS logbook address
 - ~ DCS query general parameters
 - ~ MonALISA monitoring page settings (see later)
 - ~ Online systems configurations:
 - ~ FXS specific information
 - ~ Detector configurations:
 - ~ Detector preprocessor base parameters
 - ~ DCS configuration (i.e. DCS AMANDA server and DPs list)
- ~ The configuration information are accessed via a dedicated class in AliRoot

GQ

File Filters Help

Search Browse Schema

pcalishuttle01

- dc=cern,dc=ch
 - o=shuttle
 - o=shuttle_prod
 - det=ACO
 - det=EMC
 - det=FMD
 - det=GRP
 - det=HLT
 - det=HMP
 - det=MCH
 - det=MTR
 - det=PHS
 - det=PMD
 - det=SDD
 - det=SPD
 - det=SSD
 - det=T00
 - det=TOF
 - dcshost=alidcsamanda.cern.ch
 - det=TPC
 - det=TRD
 - det=V00
 - det=ZDC
 - name=globalConfig**
 - system=DAQ
 - system=DCS
 - system=HLT

dn	name=globalConfig,o=shuttle_prod,dc=cern,dc=ch
objectClass	top
name	globalConfig
passwdFilePath	/home/shuttle/.shuttle_prod
AlienPath	/alice/data/
daqLbHost	aldaqgw01-gpn.cern.ch
daqLbUser	shuttle
daqLbDB	LOGBOOK
daqLbTable	logbook
shuttleLbTable	logbook_shuttle
runTypeLbTable	logbook_detectors
ppMaxRetries	2
ppTimeOut	1200
ppMaxMem	2097152
dcsTimeOut	1100
nDCSretries	5
DCSQueryOffset	180
DCSDelay	120
monitorHost	localhost
monitorTable	SHUTTLE_PROD
daqLbPort	1444
triggerWait	600
mode	prod
keepDCSMap	0
keepTempFolder	0
shuttleAdmin	chiara.zampolli@cern.ch
amandaAdmin	raffaele.grosso@cern.ch
sendMail	Jan.Fiete.Grosse-Oetringhaus@cern.ch
ShuttleFileSystem	Peter.Chochula@cern.ch
FreeDiskWarningThreshold	1

Apply Add as new Refresh

rpmnew
schema~
schema.copy

Shell - Konsole <2>

Session Edit View Bookmarks Settings Help

```
[wd now: ~]
[pcalishuttle01] /home
[1] 26804
[pcalishuttle01] /home
zampolli@pcalice134:~>xwd -out screenshot.xwd -root
zampolli@pcalice134:~>gimp screenshot.xwd
zampolli@pcalice134:~>xwd -out screenshot.xwd -root
```

2 3 4

GQ Konsole [4] emacs@pcalice134.cern.ch 20.01 17:3.09

Logs

Shuttle log

Detector preprocessor log

CHEP, 23rd Mar 2009

Status History

Operational log for HMP, run# 61990

Event time	Status	Count
07.10.2008 13:49:08	Failed	3
07.10.2008 13:39:08	PPError	3
07.10.2008 13:39:08	PPStarted	3
07.10.2008 13:39:05	DCSStarted	3
07.10.2008 13:38:54	Started	3
07.10.2008 13:28:54	PPError	2
07.10.2008 13:28:54	PPStarted	2
07.10.2008 13:28:52	DCSStarted	2
07.10.2008 13:28:42	Started	2
07.10.2008 13:14:46	PPError	1
07.10.2008 13:14:46	PPStarted	1
07.10.2008 13:14:44	DCSStarted	1
07.10.2008 13:14:34	Started	1

time ↑

Failed preprocessor: 3 retries

Operational log for TPC, run# 61990

Event time	Status	Count
07.10.2008 13:17:47	Done	1
07.10.2008 13:17:43	StoreStarted	1
07.10.2008 13:17:43	PPDone	1
07.10.2008 13:17:40	PPStarted	1
07.10.2008 13:17:36	DCSStarted	1
07.10.2008 13:17:36	Started	1

time ↑

Successful preprocessor: 1 retry

Some Numbers – DCS FXS

starting from mid Feb '08

DCS FXS

Detector	N. of files	Total size (MB)	% files processed
GRP	11177	57.3	36%
TOF	2497	187.7	44%
TPC	76	19.98	1%
TRD	576	967.94	34%

OCDB Size

- ~ 12500 files in LHC08d → 1.4 GB
- ~ LHC08c: 2 GB → 11700
- ~ LHC08b: 4200 → no size available
- ~ LHC08a: 9000 → no size
- ~ Average OCDB file 140 kB
- ~ 1.8 GB estimated for LHC08b + a
- ~ N. Total files 37000 with 5.2 GB

OCDB

- The Offline Condition DataBase is not a relational database, but a list of entries in the AliEn file catalogue that point to root files which are stored in the Grid
- Interface classes part of the ALICE Offline software AliRoot
- Main characteristics of the entries:
 - ROOT objects
 - Run dependency
 - Write-once-read-many (WORM) database
 - Automatically controlled of the versioning

Some Numbers – DAQ FXS

starting from mid Feb '08

DAQ FXS

Detector	n. Files in valid runs	n. of processed files in valid runs	Total size (MB)	% files processed
FMD	449	449	51	100%
GRP	15192	8549	15182	56%
HMPID	9594	4783	1604	50%
MCH	4120	3036	8938	74%
MTR	218	149	16	68%
PHS	1461	536	30246	37%
PMD	435	221	96	51%
SDD	905	621	25045	68%
SPD	4658	3733	363	80%
SSD	1021	925	5484	91%
T0	1420	753	580	53%
TOF	2195	1805	1530	82%
TPC	3469	2975	2115	86%
TRD	1949	395	191	20%
V0	779	37	109	5%
Sum	47865	28967	91576	61%