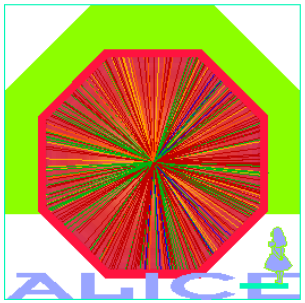


Offline Calibration and Alignment DB

L.Betev

HLT workshop

Geneva, December 8, 2004

Outline

- Technology choices
- Geometry database and the Geometrical Modeller
- Calibration database
- Relation and connectivity with the other ALICE and external DBs
- Common issues
- DB design parameters – call for user requirements
- Offline roadmap

Offline technology choices

- Data object store technology is already done – ROOT
- All offline alignment and calibration files are read-only
- Read-write tables are needed for the file and metadata catalogue – this is already implemented in AliEn (gLite) and is handled transparently for the user
- Interactions with external databases will (preferably) be handled through APIs
- In the cases where the above is not possible: through old-fashioned manual data insertion (only suitable for numbers that very rarely change)

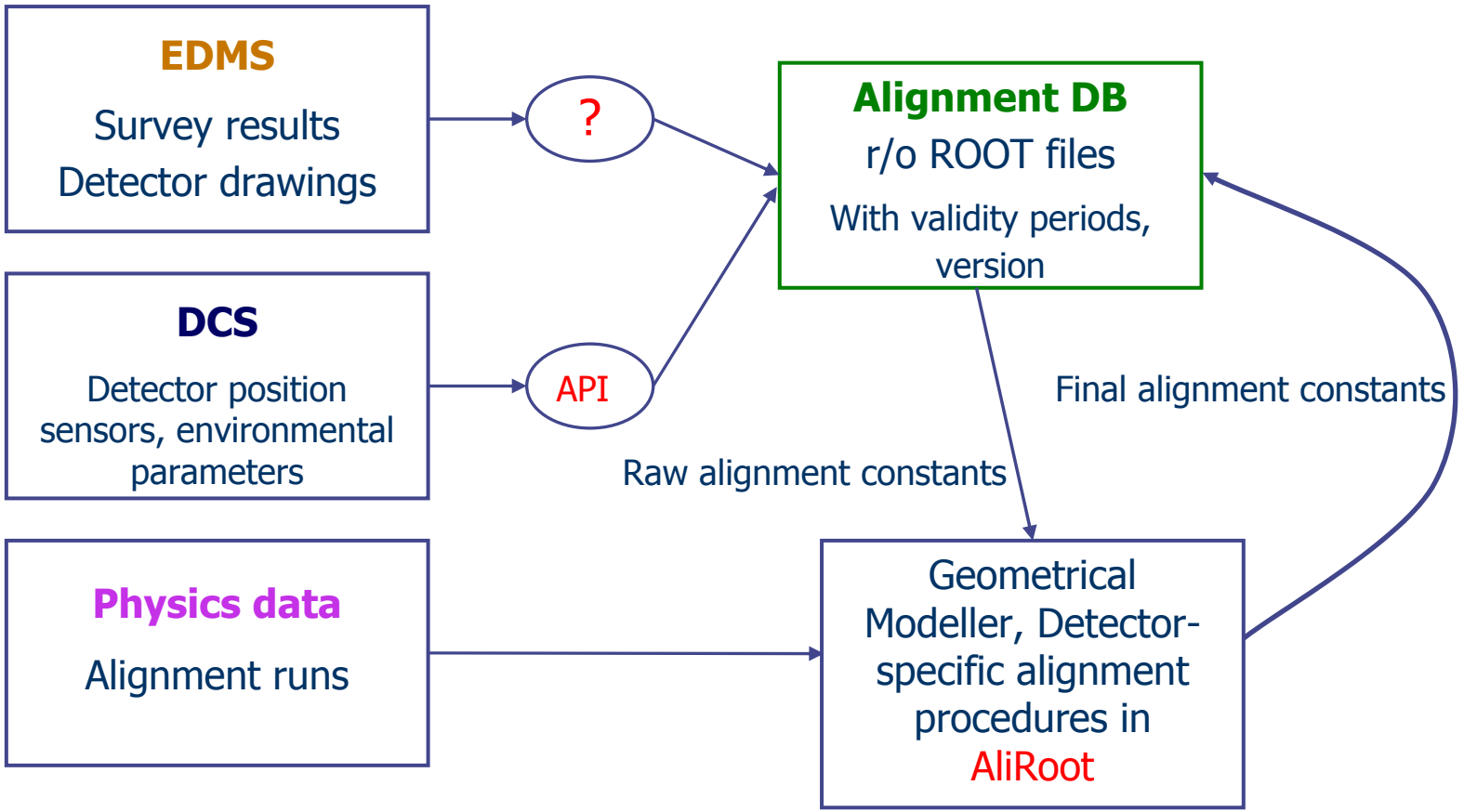
- **Summary: all calibration and alignment data will be stored in ROOT files with the proper tags, described in the AliEn (gLite) metadata catalogue:**
  - **validity period – run, event, time**
  - **version**

Alignment database

- Contents (educated guess):
  - Raw:
    - Optical survey results – these will be published by the surveyors in EDMS
    - Full detector description (modules, sectors, pads, wires ….) – from drawings, schemes (also partially in EDMS)
  - Final:
    - Results from the alignment procedures – both internal and global detector alignment
  - This is just a preliminary list, which will be expanded as we learn more from the detector groups
  - Not just numbers – tightly coupled with the Geometrical Modeller
- Geometrical Modeller – unique description tool for the detector geometry:
  - Integral part of the alignment procedure, for which it will need:
    - Complete inter sub-detector geometry description and its relation to the common Alice coordinate system

Alignment procedures and relation to external DBs

- Known relations:



Alignment database information sources

- What is existing (partially):
  - Flat EDMS structures for survey results and drawings:
    - We don't know yet what is relevant
    - This will become clear after discussions with detector groups
- Most time-consuming part:
  - ***Alignment procedures*** – we are aiming for an accuracy better than 100 microns in a big part of the detector

Calibration database

- Contents (educated guess):
  - Calibration constants – detector dead channel maps, gain factors, environmental conditions, running conditions .... - these will be gathered from many sources (partially reviewed later in the talk)

- Several sub-detector groups are already implementing code in AliRoot dealing with calibration issues – this is being reviewed
- The calibration DB and procedures are potentially the most complicated problem remaining to be solved in offline

DBs in Alice

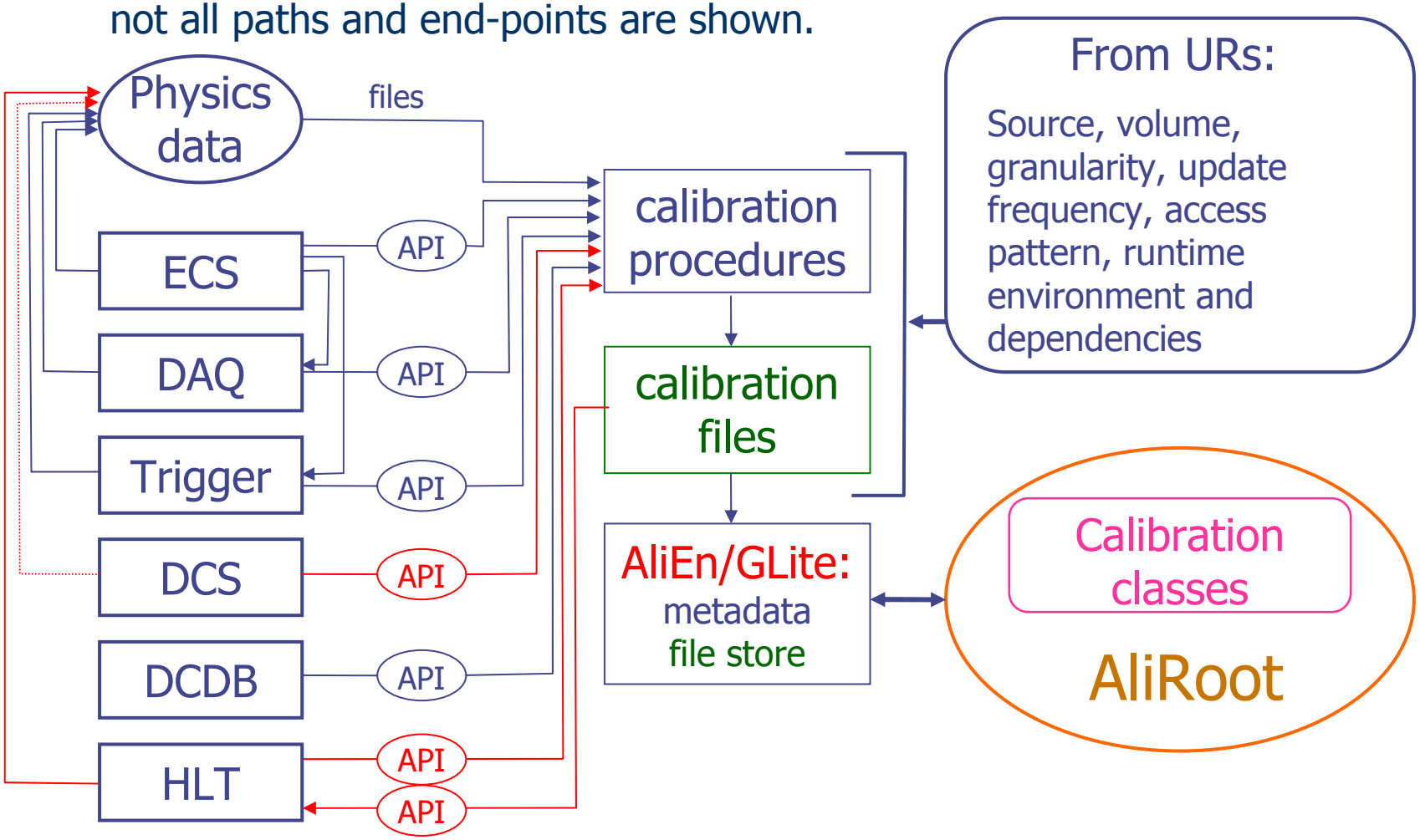
- Condensed table for offline:

Class	Update frequency	Access frequency	Access pattern	Size	Technology
Geometry & Alignment	Every alignment cycle	Reconstruction / simulation job	Write once read many	Mega Bytes	ROOT
Calibration	Every Run (?)	Reconstruction / simulation job	Write once read many	Tera Bytes	ROOT
File, Job and Output catalogue	Every time a file is created/removed	Each time a file is read/deleted	Many writers and readers	Giga Bytes	AliEn/gLite (distributed RDBMS)

- DCS:
  - Application scope: configuration of systems and devices (modules and channels), front-end configuration (busses, thresholds); Archiving of monitored detectors and devices parameters
  - Size: millions of records, Tera bytes
- HLT:
  - Application scope: mini-DST like TAG/ESD database for physics studies and offline event selection
  - Size: up to 10^9 events and 30TB per year
- DCDB:
  - Application scope: use by individual sub-detector groups and integration, repository and flow management for modules, components and their test data, cables, racks
  - Size: millions of records, Tera bytes
- ECS:
  - Application scope: inclusion/exclusion of sub-detectors to a partition
  - Size: small number of small records
- DAQ:
  - Application scope: parameter repository and resources assignment to DAQ tasks: configurations (current and stored), run parameters (current and stored)
  - Size: possibly large number of small records
- Trigger:
  - Application scope: repository for trigger classes (input to CTP), definition of trigger masks
  - Size: large number of small records

External relations and DB connectivity

- Relations between DBs not final, some access methods not included, not all paths and end-points are shown.



API – Application Program Interface

Common issues

- User interface:
  - AliEn/gLite API is under development
  - Metadata catalogue has to be further enriched with functionality to match the needs of the alignment/calibration tasks
- Level of data replication:
  - How much data has to be replicated in the Offline DBs:
    - Detector geometry/survey – entirely
    - From DCS, DAQ, Trigger, ECS, HLT, DCDB: (all/partial)
  - The above is with regards to the accessibility of data and reduction of dependencies during the data reconstruction/analysis phase. It will influence the physical design (hardware) and may limit the choice of RDBMS
  - Full replication is impossible and not necessary (DCS is potentially huge!) – have to decide on what is important and make sure the access to this data is fast and uninterrupted

- Ultimate goal: to reduce the problem of accessing different DB sources to an already (to a large extent) solved problem of accessing annotated data sets from the AliEn (gLite) file catalogue user interface:
  - May not be fully achievable
  - Still, we have to reduce the amount of external DB dependencies to a minimum
- Responsibilities:
  - Offline internal – tight coupling between the AliRoot framework, AliEn/gLite and detector group code
  - Good understanding of the DB relations between offline and DCDB, ECS, DAQ and trigger systems:
    - Have to iron out the technical details
  - DCS depends on the tools and components delivered by JCOP and PVSS – have to make sure they take into account the requirements of the offline (these will become more clear after the UR are gathered and studier)
  - HLT – discussion on elements of common interest started

Call for user requirements

- Essential parameters of the Alignment and Calibration DBs:
  - Access pattern and frequency
    - Only offline reconstruction?
  - Typical record size
    - Very difficult to estimate from first principles
  - Frequency of updates
    - We expect the majority of the updates to be infrequent
    - This probably will not be the case
  - Validity range parameters
    - Tags: Run, event, time, version – what else?

- These all have to come from the user community – a call for user requirements has been sent

Offline DB roadmap

Action Item	Date completion
Collection and analysis of user requirements	End 2004
Prototyping, first version	January 2004
Alignment and calibration workshop – discussion of the UR, calibration and alignment procedures	February 2005
Test of functionality, integration with AliROOT, PDC05	May 2005
Performance tests, scalability tests	December 2005
Ready for full functionality tests during PDC 2006	January 2006

Conclusions

- We are starting to work on the offline alignment/calibration databases
- Basic technology choices are already done – this will help enormously and we have good understanding how to use it and what has to be improved
- To understand fully the content and parameters of the databases we need the input from the detector groups – the gathering of information is advancing
- The most time-consuming task will be to develop the alignment/calibration algorithms:
  - Powerful tool – the Geometrical Modeller – will be the base for the ALICE alignment procedures
  - Calibration – some detector groups have already started to develop internal procedures
- Discussion among the detector experts is essential:
  - The time frame is extremely tight:
  - We cannot afford to have divergent/incompatible development paths: early on we have to agree on a common line
- The discussion with DAQ, HLT, Trigger, ECS, DCS and DCDB groups, developing DBs relevant for the offline have to be further expanded and the DB relation paths fully understood