HARP DATA AND SOFTWARE MIGRATION FROM Objectivity TO ORACLE

Authors: A.Valassi, D.Geppert, M. Lübeck, K. Nienartowicz, M. Nowak (CERN IT-DB, Geneva, Switzerland), E. Tcherniaev (CERN PH, Geneva, Switzerland), D. Kolev (University of Sofia, Bulgaria)

MOTIVATION

End of Objectivity support at CERN in 2003

SCOPE

1 - Migrate HARP data from Objectivity to the new persistency solutions based on Oracle

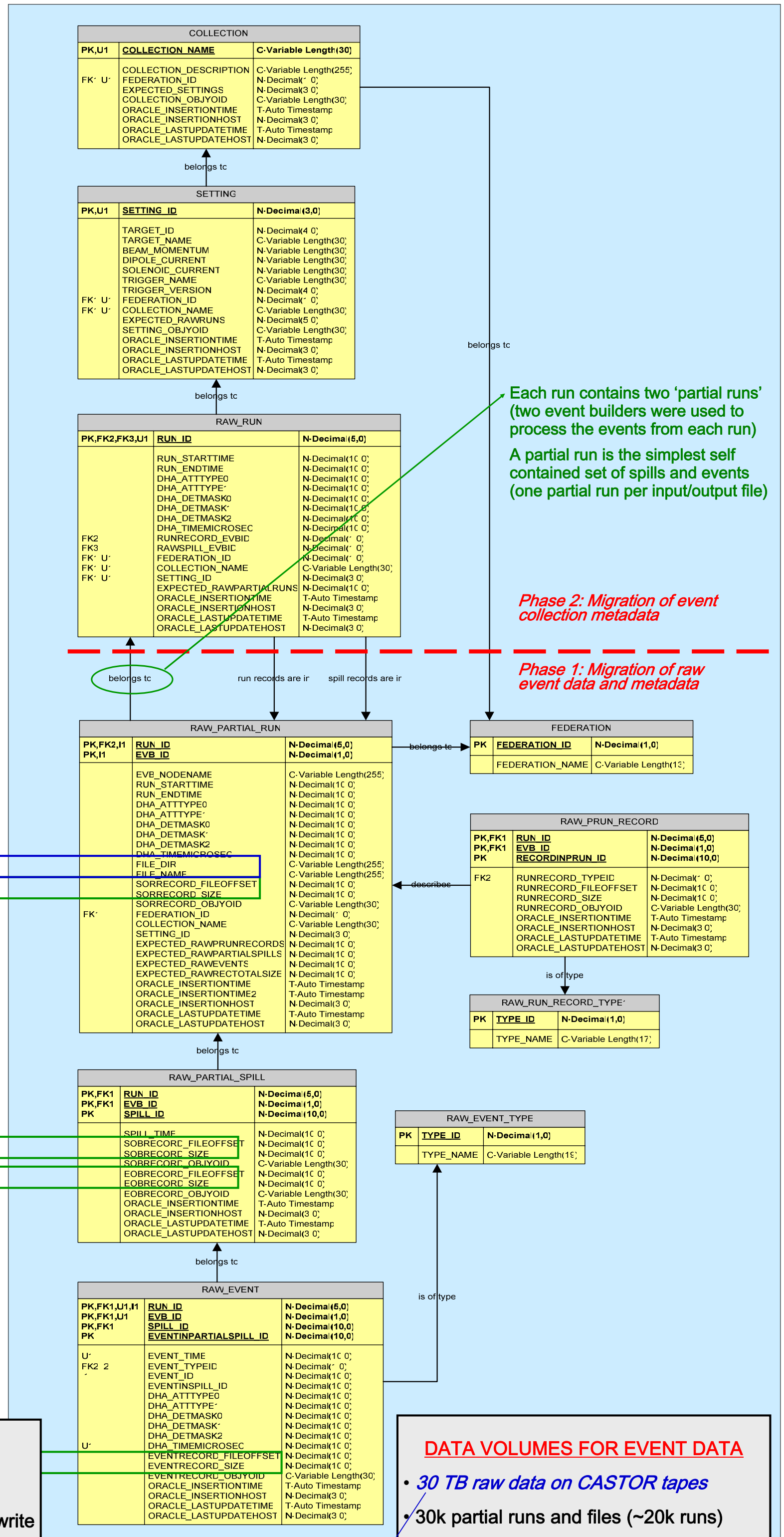
2 - Develop read-only Oracle software to allow event processing in the HARP Gaudi framework

THREE PHASES

1 - Raw event data migration

2 - Event collection metadata migration and development of Oracle navigation software

3 - Condition data and software migration

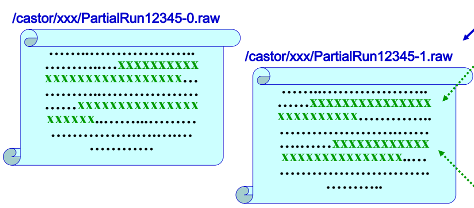


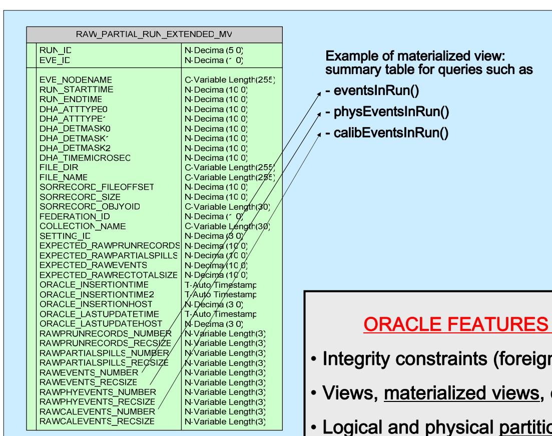
PHASES 1 AND 2 (EVENT DATA): HYBRID DATA STORE

Raw data in flat files (DATE format)

Metadata in Oracle

Similarities to the POOL data model





Example of materialized view: summary table for queries such as
- eventsInRun()
- physEventsInRun()
- calibEventsInRun()

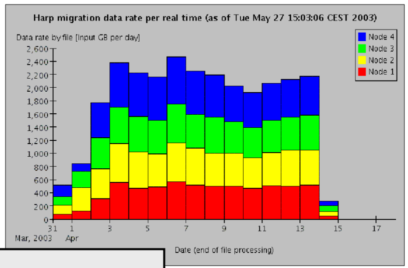
ORACLE FEATURES USED

- Integrity constraints (foreign keys)
- Views, materialized views, query rewrite
- Logical and physical partitioning for all tables and materialized views (separate tablespaces for different table partitions)
- Indexes (local/global, b-tree/bitmap)
- C++ binding: OCCI client library (use *SQL bind variables* and *bulk operations*)

DATA VOLUMES FOR EVENT DATA

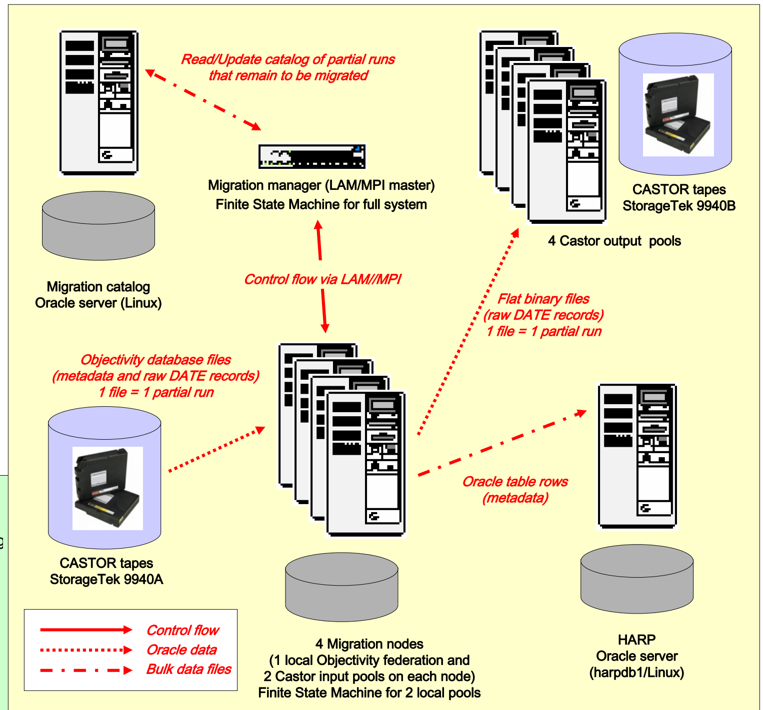
- *30 TB raw data on CASTOR tapes*
- 30k partial runs and files (~20k runs)
- 800M events (rows in Oracle event table)
- *100 GB event metadata in Oracle*
- In addition: 100 GB Oracle indexes

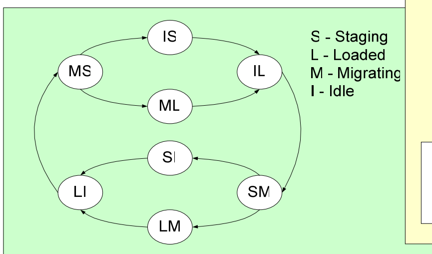
The idea of storing also the DATE binary records as BLOBs in the Oracle database had been initially considered… but what would you gain? The content of the raw records is not meant to be queried using SQL anyway!



PHASE 1: RAW EVENT DATA

- Data migrated: *30 TB raw data in 15 days*
- Reuse tools and hardware from previous COMPASS migration (300 TB raw data)
- HARP specific preparation, cleanup and consistency checks: 2-3 months

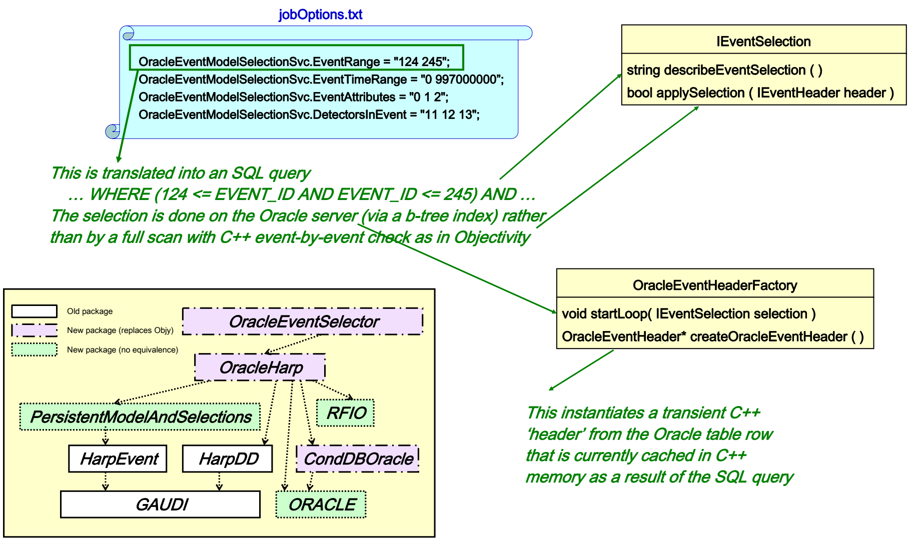


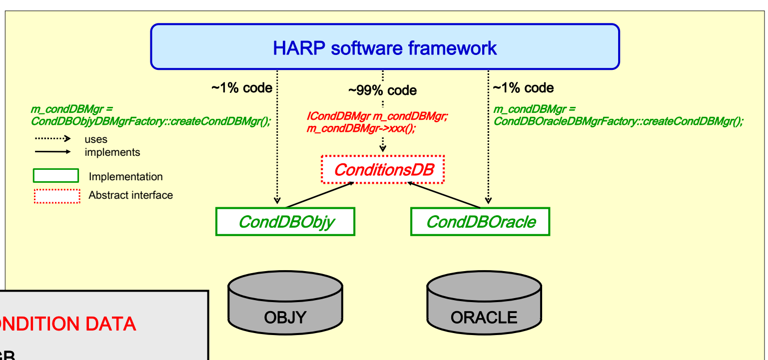


Finite State Machine for the two local pools on a migration node (optimize use of input tape drives and local CPU)

PHASE 2: EVENT COLLECTION METADATA

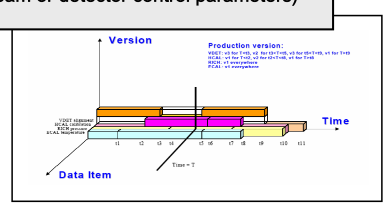
- Data volume: < 1 GB
- Complex data model: ~ 4-5 months
- Relational schema design (see diagram)
- Migration application for event collections also checks for internal consistency when 'merging' two partial runs
- Implementation of read-only navigation and selection criteria (performance optimization by the use of bind variables, result set caching and materialized views)
- Re-engineer event selection (move most algorithms to technology-independent PersistentModelAndSelections, tested also in adapted version of Objectivity code)
- Detailed comparison of Oracle and Objectivity results





CONDITIONS DATABASE API

- Metadata for condition data: data item name, validity interval [since,till], version
- Actual condition data encoded in binary records (BLOBs): strings (calibration file names), streamed vectors of numbers (beam or detector control parameters)



PHASE 3: CONDITION DATA

- Data volume: < 10 GB
- Fast and rather painless: ~ 1 month
- No design needed, solution available: use Oracle implementation of the same abstract interface that was implemented in Objectivity
- Data migration: via simple dedicated tools (partly using the API, partly bypassing it)
- Software for read-only data access: only minor changes with respect to Objectivity

SUMMARY

- Large data migration successfully completed!
- Hybrid data store with Oracle metadata navigation
- Data migrations may also be needed at LHC!

Many thanks to:

IT-DB, IT-DS, IT-ADC, Castor team, S. Giani, I.Papadopoulos