Computing at D0

Focus on D0-France Contribution

Patrice Lebrun

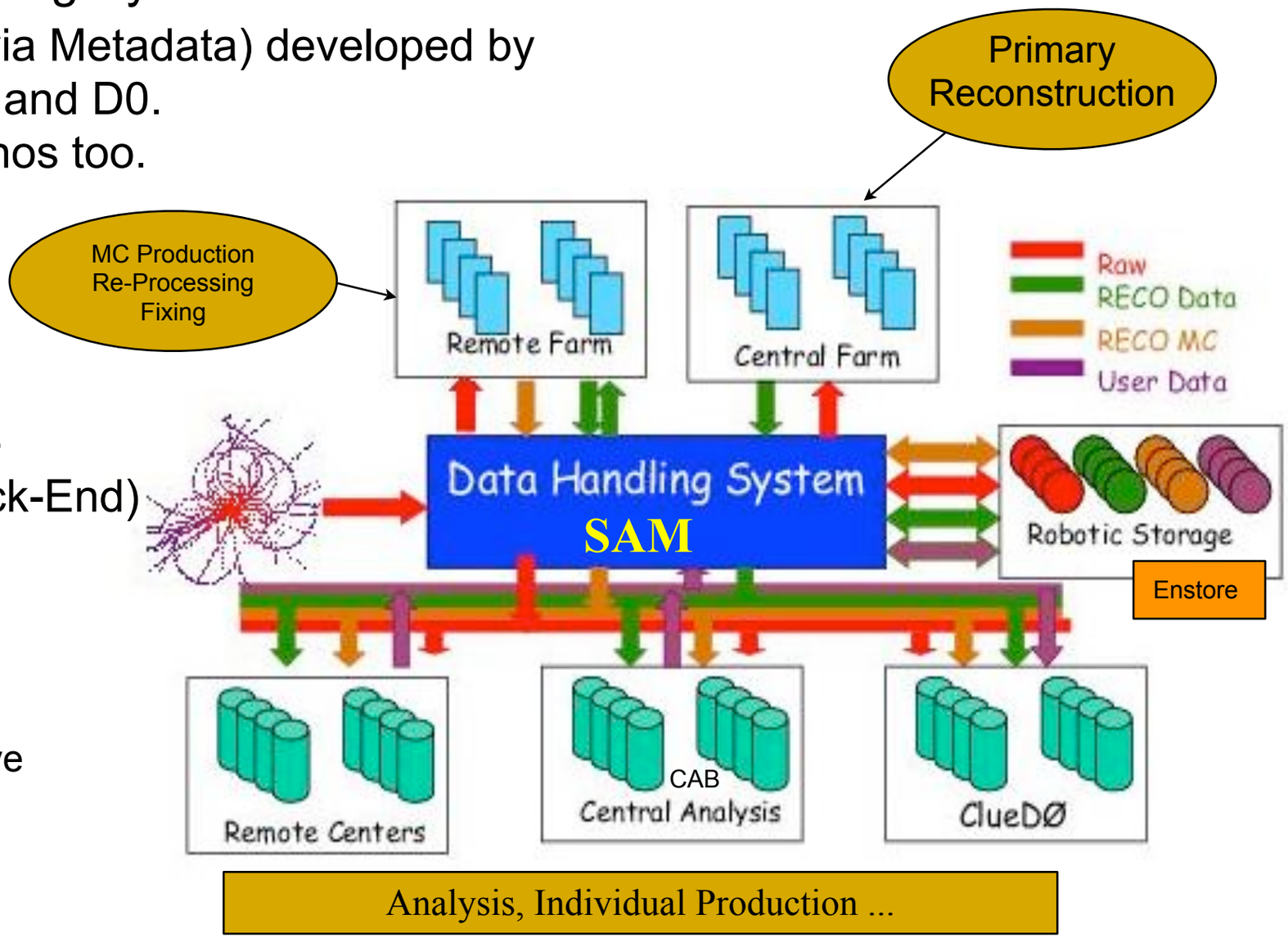
October 13, 2008 at LPNHE
10-25 Years of D0 france

Introduction

- Tevatron :
  - Less data than LHC but still PBs/experiment
    - Total 4.65 Billion raw events in Run II (reconstruction is ~80 Million events behind)
      - Production rate this year 2008 : 4.6 M events/day, O(1 TB/ day)
    - Total Storage (as today): 3.8 PB
  - Need to have a very Powerful and Robust Computing System
    - Computing Model based on distributed data handling and job grid technology
  - D0 is a Running Experiment
    - Makes the Computing Model Evolution not easy to do
      - No GRID computing available at the beginning of RUN II
  - Very large number of analysis
    - Data management is very important: Task of Common Sample Group
      - MC requests management
      - Data and MC data management
      - Definition of Common Analysis Tools and File Format
  - D0-France has large impact in the D0 computing

Computing Model

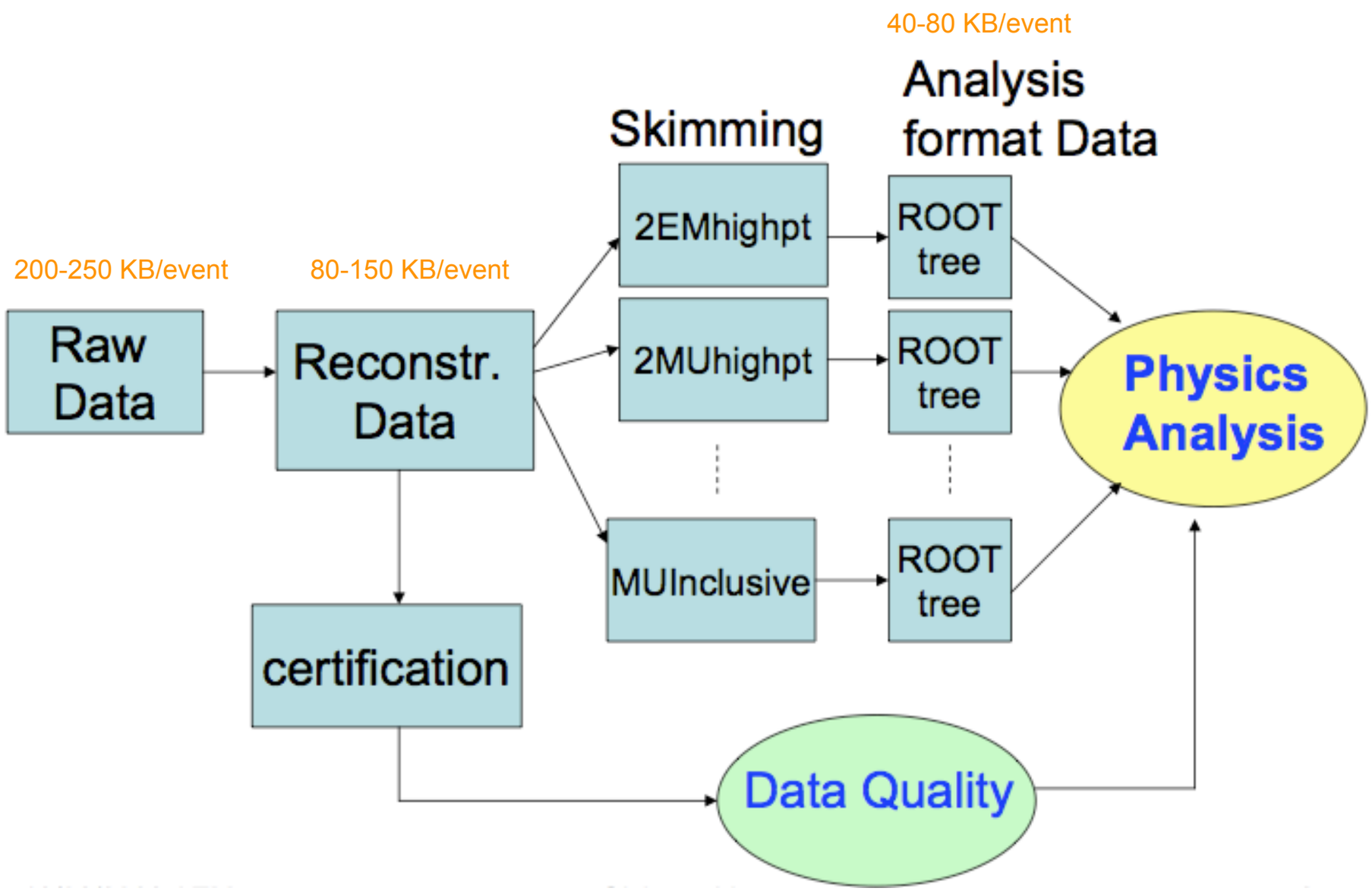
- The Distributed Data Handling System is core of this Model
  - SAM (Sequential Access via Metadata) developed by FNAL-Computing Division and D0.
    Now used by CDF and Minos too.



- 2 Kinds of Cluster at FNAL
  - CAB (Central Analysis Back-End)
    - Use in Batch with high speed access to Data
  - ClueD0
    - Owned by institutes and used essentially for interactive analysis and development.

- Data Processing centralized at FNAL
- Data Re-Processing at FNAL and in the Remote Sites
- MC Production and fixing produced by Remote sites Only

Data Flow

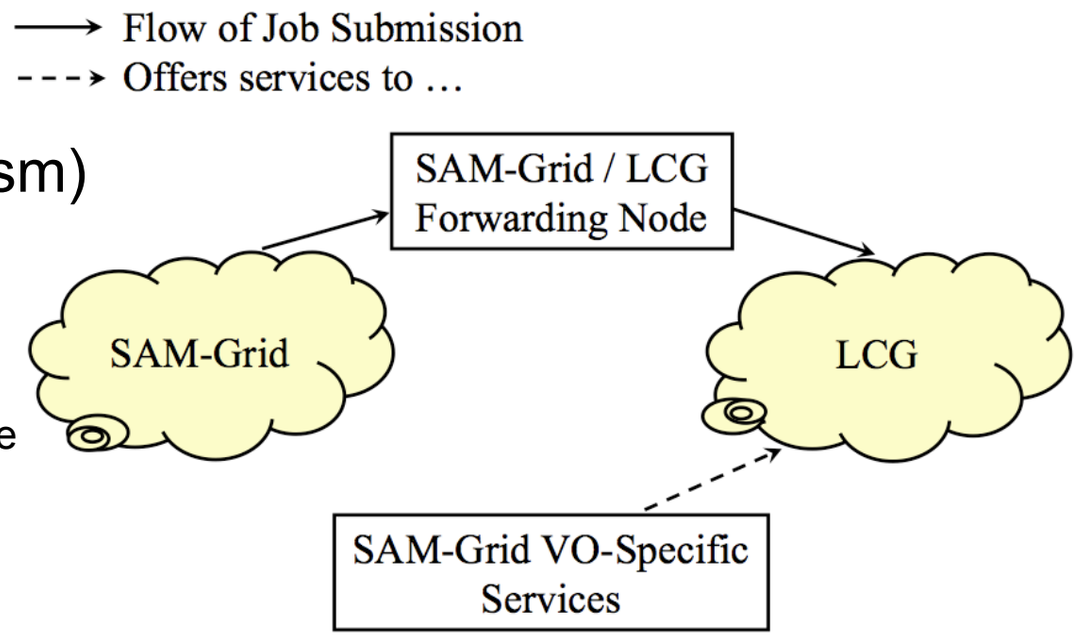

40-80 KB/event
Analysis format Data
Skimming
200-250 KB/event
80-150 KB/event
Raw Data
Reconstr. Data
2EMhighpt
2MUhighpt
MUInclusive
ROOT tree
ROOT tree
ROOT tree
Physics Analysis
certification
Data Quality

Computational GRID

- D0RUNJOB (at the worker node level)
  - handles job workflow management
  - Makes Metadata

- SAMGRID (D0 Grid)
  - JIM (job information and monitoring) + SAM = SAMGRID
    - Two SAM Stations at CCIN2P3 available

- GRID interoperability (forwarding mechanism)
  - SAMGRID with:
    - LHC Grid (LCG)
      - First test in 2003 at CCIN2P3
      - 1 SAM Station dedicated at CCIN2P3 for this purpose (ccin2p3-grid1)
    - US Open Science Grid (OSG)



Re-Processing and Fixing

- Special tasks needed after a best understanding of the detector and/or after better algorithms of reconstruction
  - Reprocessing from RAW data
  - Fixing runs from reconstructed (TMB)

- Fixing of MC has been also done
  - Exclusively taken in charge by D0-France and produced at CCIN2P3

- Reprocessing is performed by Remote Sites
  - A large part done by D0-France at CCIN2P3
    - Good performance due to HPSS (High Performance Storage System) facility where data has been pre-staged
  - Has been done for Years 2003, 2005 and 2007
  - Short History of these data reprocessing in the next slides ...

Year 2003

- First Re-Processing, not “real computing grid” yet
- 100 M events done remotely (20%)
- ~25 TB of data transferred
- D0-France uses a local bookkeeping based on Oracle DB (used for MC production)
  - Very efficient resubmission of failed jobs
  - Easy control and fine tuning of jobs to fit with the CC infrastructure

Farm	# Events (millions)	# CPUs (1 GHz PIII)
CCIN2P3	36	160
UK GridPP	23	270
GridKa	21	200
WestGrid (CA)	12	1000
NIKHEF	7	400

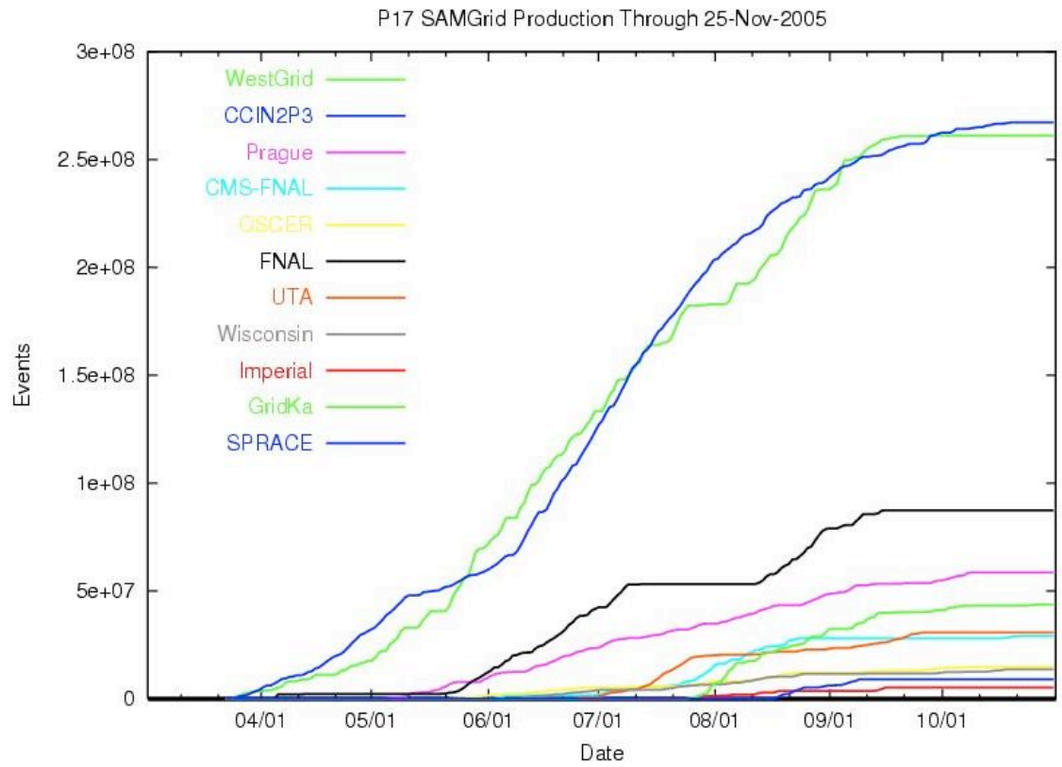
Year 2005

- Re-Processing all Data
- Only SAMGRID used

Declared Available Resources Total **3430 CPUs** (1 GHz PIII)
Total # Events to be reprocessed **986.2 M**

…few examples …. not all sites!

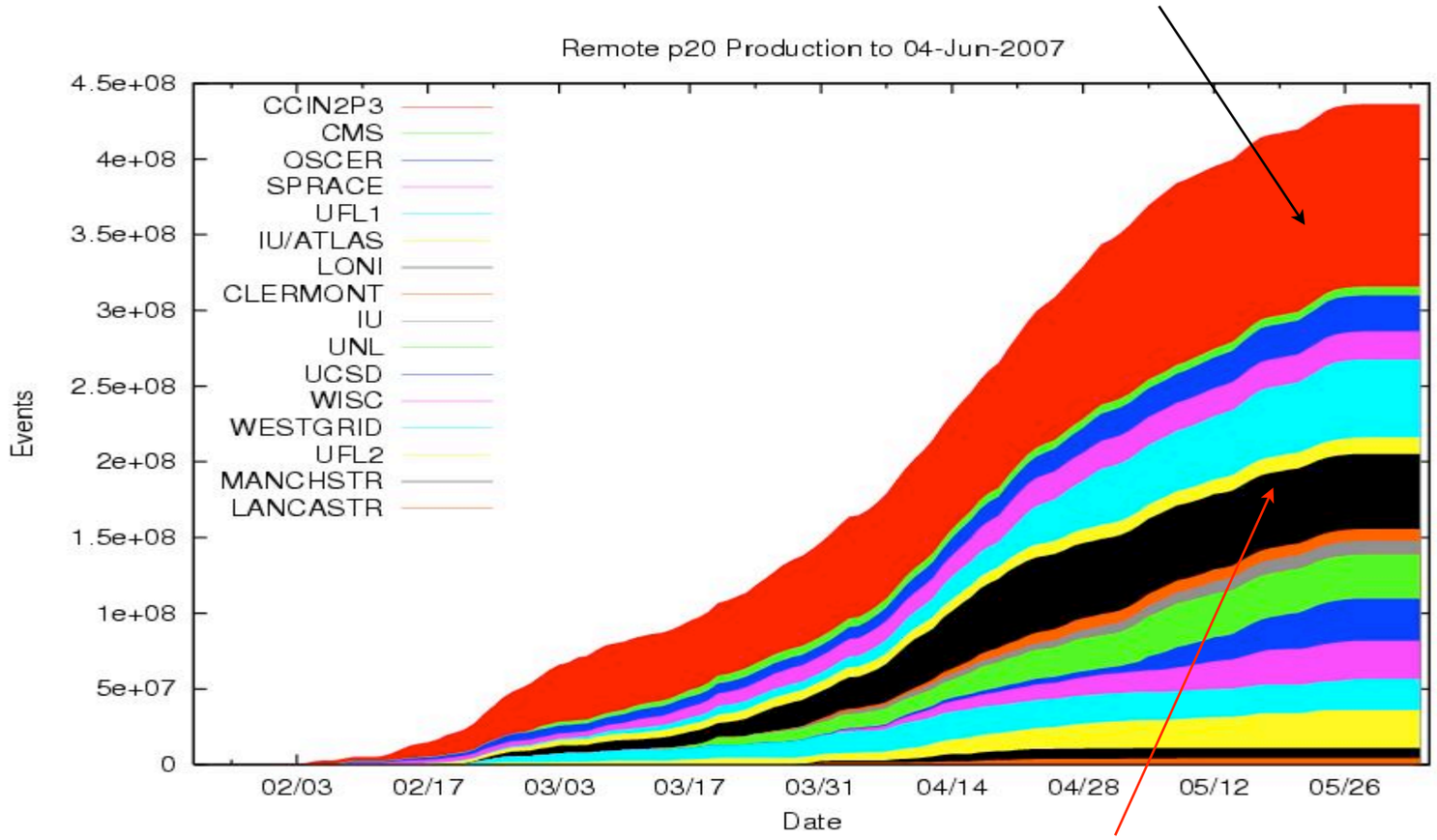
Institution	Available Resources		# Events Reprocessed		QF
UK (4 sites)	750	(21.9 %)	3.2M	(0.3 %)	0.01
WestGrid Vancouver	600	(17.5 %)	261.0M	(26.5 %)	1.51
GridKa Karlsruhe	500	(14.6 %)	39.0M	(4.0 %)	0.27
CCIN2P3	**400**	**(11.7 %)**	**267.3M**	**(27.1 %)**	**2.32**
FNAL	340	(9.9 %)	218.7M	(22.0%)	2.22
FZU-GRID Prag	200	(5.8 %)	54.9M	(5.6 %)	0.97
CMS-Farm FNAL	100	(2.9 %)	29.2M	(3.0 %)	1.03



Year 2007

- Re-Processing of data taken after the upgrade of detector in Summer 2006
- OSG and LCG used (about 20 sites)
- SAMGRID used at CCIN2P3
- Coordination done by D0-France.

CC-IN2P3: 120M events …. 27%



LCG: 19M…4.3 %

Clermont, Manchester, Lancaster

MC Production

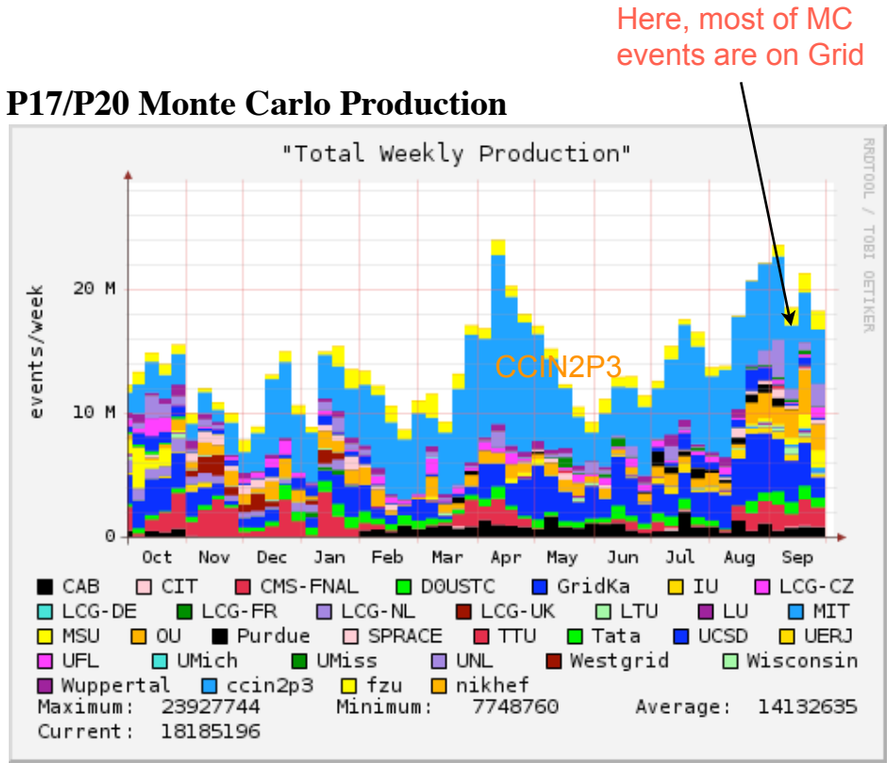
- D0-France is involved in many levels of it

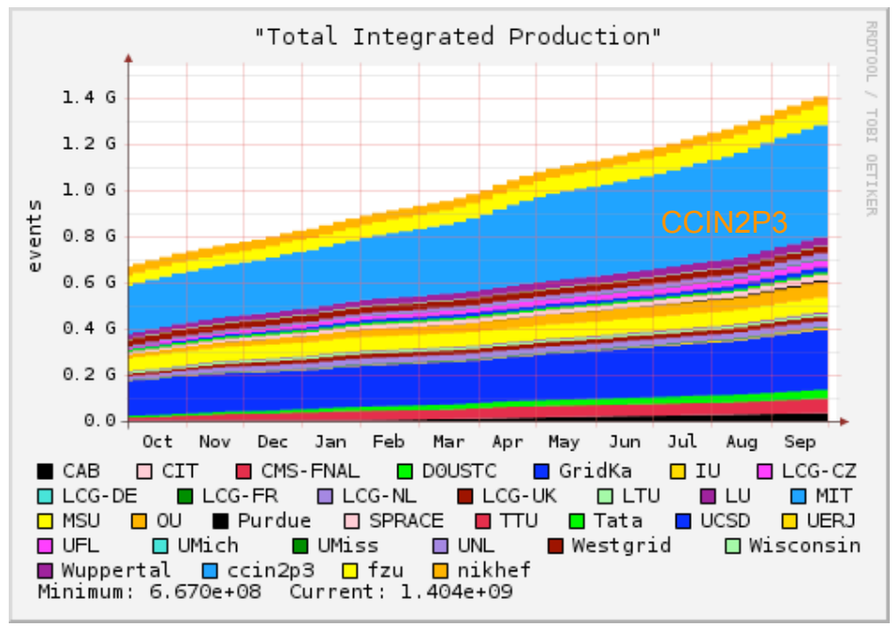
- MC Generator Group
  - Management and integration of generators in D0 Framework
  - Definition of Default Parameters
  - Follow the needs of Physics Groups
- MC Requests Coordination
  - Definition of priorities with the Physics Coordinators
  - Checking and Approval of Requests
    - Rules correctly applied
  - Creation (submission) of Requests
    - About 100 requests / week
- Samples Access
  - MC Catalog
  - Request Id List Maker (REILM) to find requests for a given process.
  - Metadata definition

- CCIN2P3 Production Farm
  -

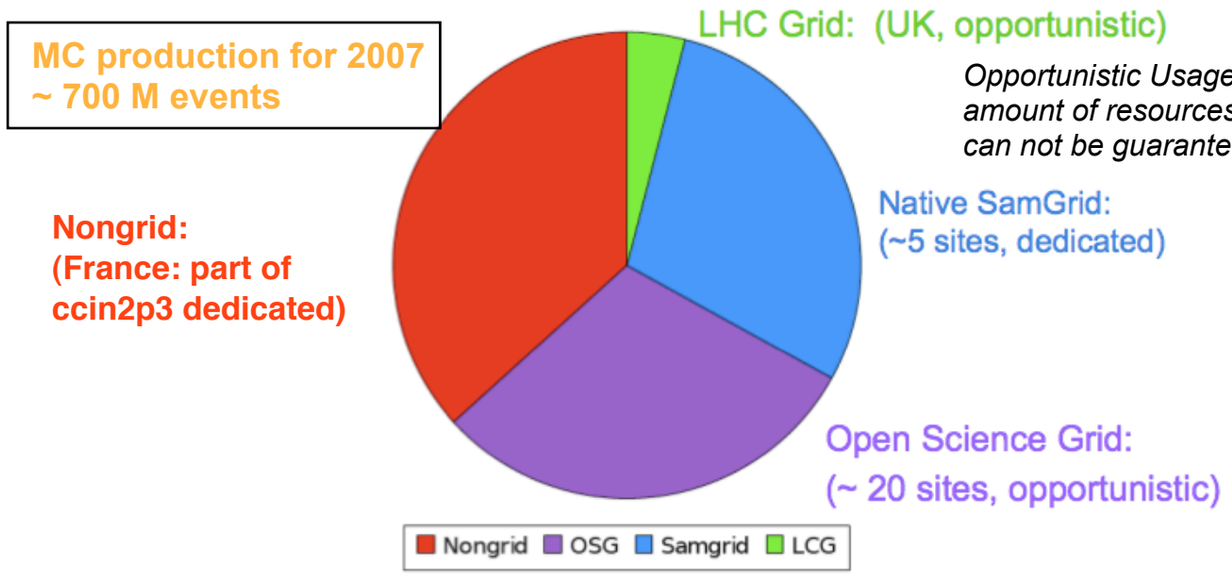
MC Production in Farms

- A large part of events are produced at the CCIN2P3 by D0-France
  - On about 15 M of events produced per week , about more than 5 M are produced by D0-France
- Production at CCIN2P3 is out of GRID
  - very large adaptability due to
    - A local Oracle DB to manage jobs and files produced before to declare and store files at FNAL
    - Files are stored locally in HPSS (permanently or temporally)
      - CCIN2P3 is the only Farm having a very large Storage System
    - Due to this many very specifics requests are produced





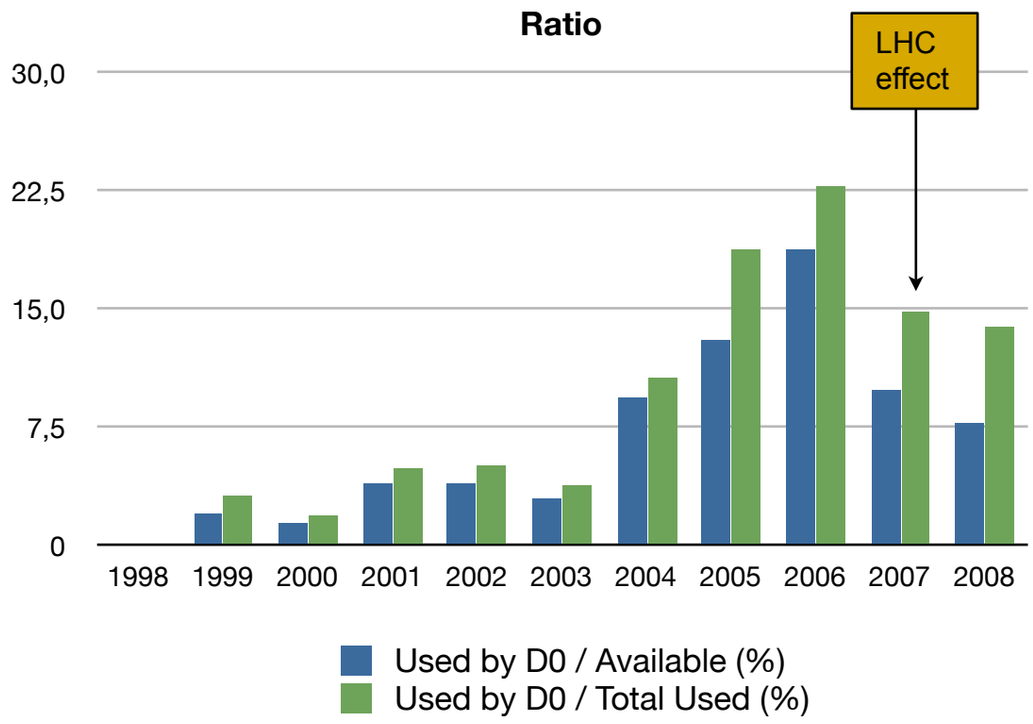
since September 5, 2005 to October 5, 2008.

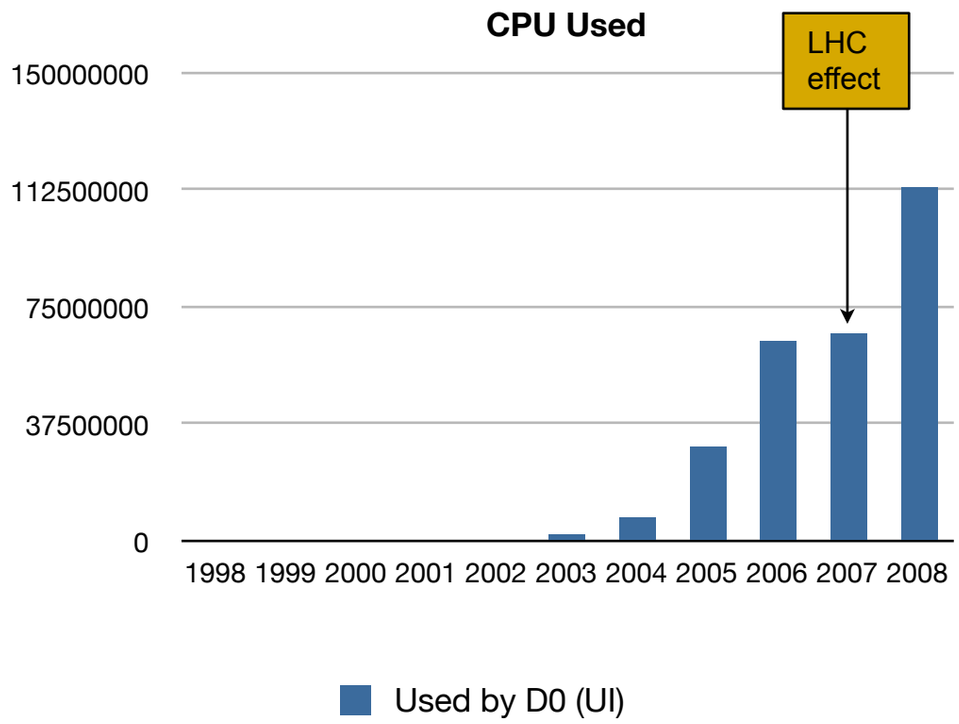


CPU Consumption at CCIN2P3

- CPUs used at CCIN2P3 by D0 (Analysis is about 5% of this amount)
  - D0 is the first user of CIN2P3 with BABAR in 2004-2006, with ATLAS in 2007-2008.

Years	Used by D0 (UI)	Available (UI)	Total Used (UI)	Used by D0 / Available (%)	Used by D0 / Total Used (%)
1998	307,3	2202062,7	1659963,1	0,01	0,02
1999	77478,7	3724670,2	2461400,2	2,08	3,15
2000	171887,4	12013792,4	9124054,3	1,43	1,88
2001	897183,1	22862516,8	17938302,0	3,92	5,00
2002	1113976,1	28289448,5	21608683,9	3,94	5,16
2003	2029229,7	68610205,1	53544225,2	2,96	3,79
2004	8334910,9	88181008,7	77983379,1	9,45	10,69
2005	30880713,9	235081298,8	164543521,4	13,14	18,77
2006	64283696,2	340552672,3	282273506,0	18,88	22,77
2007	67164677,3	683014352,0	450392867,4	9,83	14,91
2008	114059576,6	1458368396,9	821554838,7	7,82	13,88

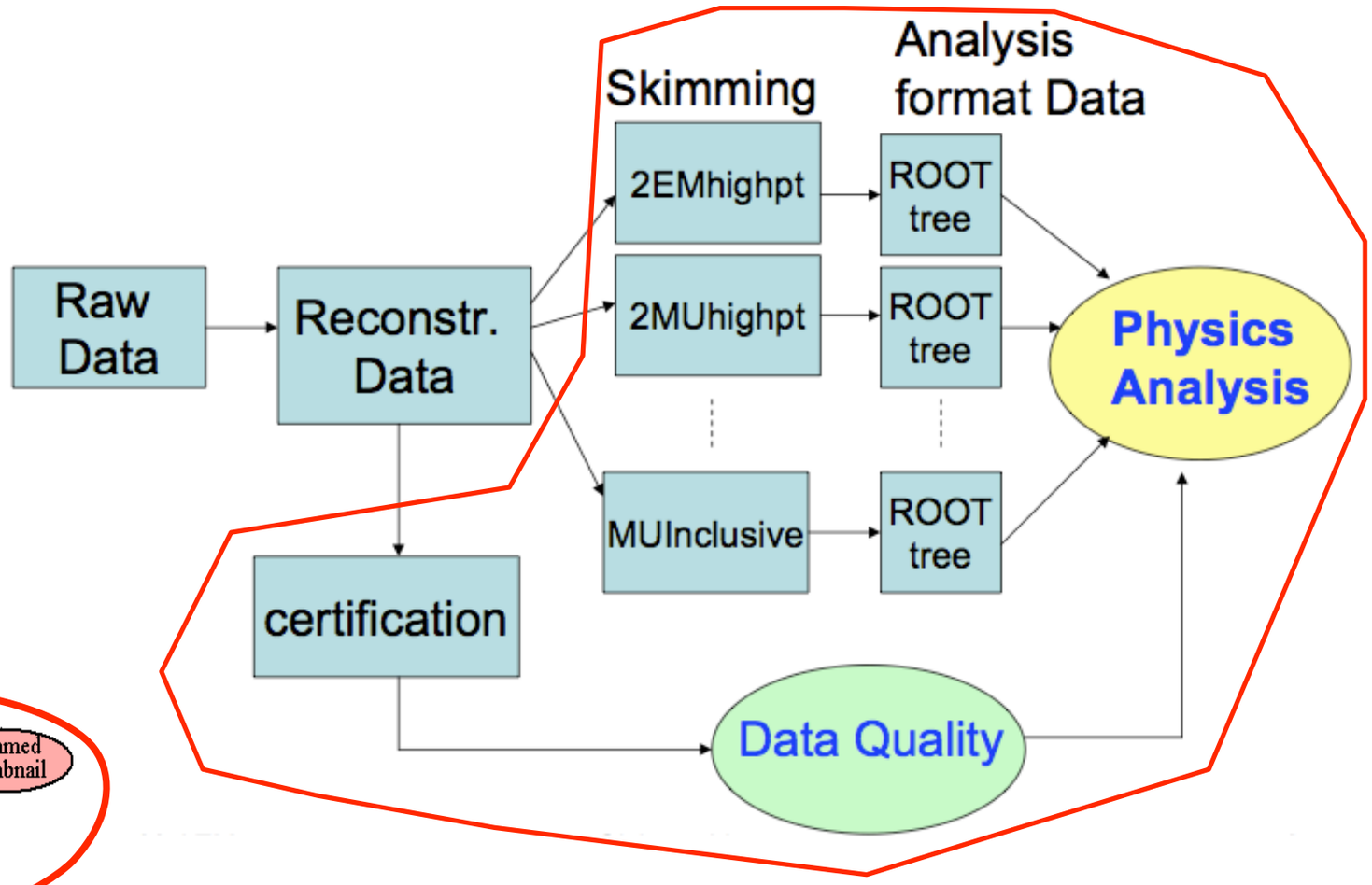


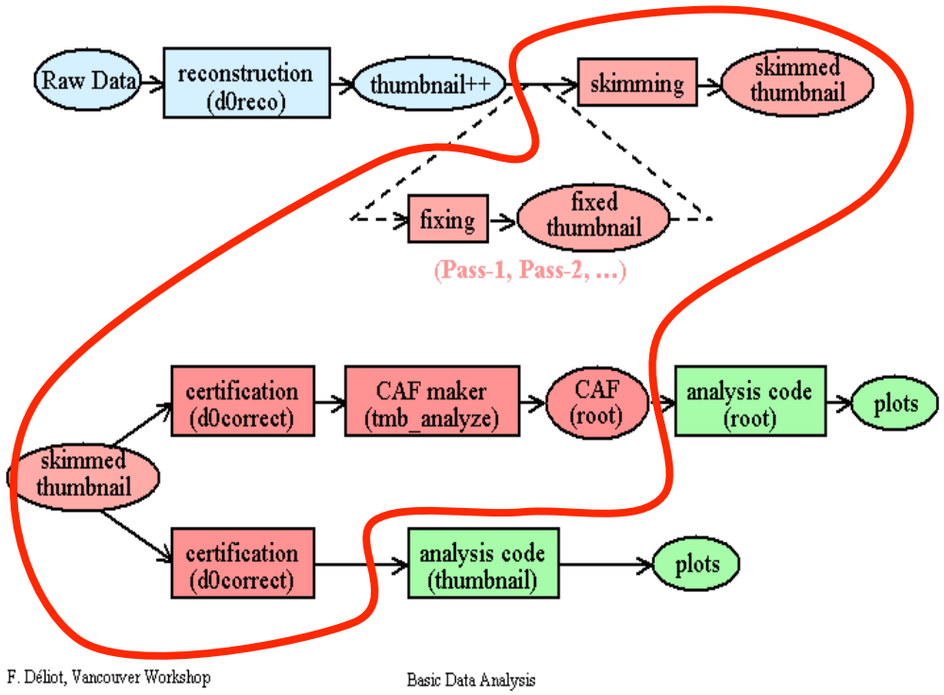


About Common Sample Group

- Group created in 2003 with the following goals
  - Maintain and document official D0 samples that are of general interest for physics analyses. […]
  - Maintain and document commonly used Monte Carlo samples.
  - Use the common-samples group meeting as a platform for discussing issues of these samples."
- One convener of this group from D0-France (Saclay) since its creation
- Number of charges is large
  - Data fixing (post-processing to rerun part of the reconstruction, done in the past)
  - Definition of the common data skims and datasets based on physics objects
    - inclusive skims: events with 1 electron, 1 muon, jets, …
  - Applying standard corrections/quality criteria (d0correct):
    - definition of what is a good muon, electron, applying JES
  - Definition and implementation of the common root format for the analyses (CAF: done in the past)
  - Production of the CAF files for all produced MC
  - Development of the standard root based tools to do analyses:

CSG and Data Flow





F. Déliot, Vancouver Workshop Basic Data Analysis

Some words about Analysis

- Mainly, analysis are done at FNAL on the central Analysis Back-End (CAB) and Clued0 facilities at FNAL
- About 5% of resources used by D0 at CCIN2P3 are used for analysis by some french groups
  - For structural reason is not easy but possible

- Analysis on GRID
  - A prototype is been developed by d0-IN2P3
    - SAM not used here

Conclusion

- Contribution of D0-France to D0 computing is important
  - Small group of D0-France laboratories of dedicated people:

    LAL Orsay, IPN Lyon and CPPM Marseille

- Thanks to the close collaboration with FNAL-Computing Division and D0 experts
  - For instance, at the beginning HPSS interface to SAM has been very important for us (many thanks to Igor Terekhov) then many other collaborations

- Thanks to CCIN2P3 for their help and cooperation

- Up to the end of D0 experiment, is very important
  - to keep CCIN2P3 as a large computing Center for D0
    - Need to live with LHC experiments
  - To maintain our cluster of machines at FNAL for Analysis.