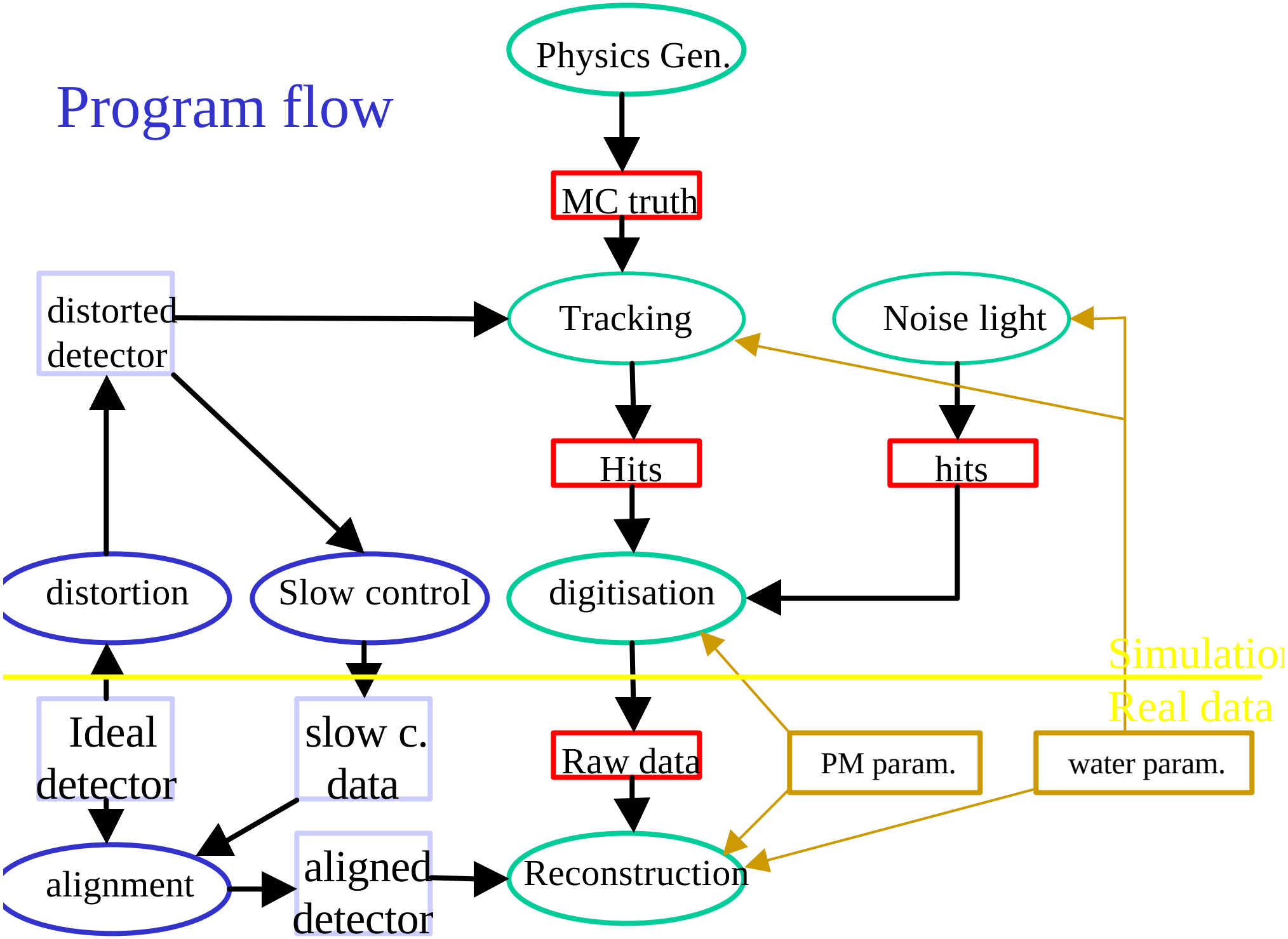
Le OFFLINE dans le projet ANTARES

Jurgen Brunner- Cargèse 2001

Program flow



Physics generators

- Most complex for atm. Muonshowers

  primary flux - atm. Shower - muon propagation

- Mostly inherited from other projects

  KORSIKA (F77) (Kaskade) HEMAS (F77) (Macro)

- Muon propagation

  PropMu (Macro) Mum (Baikal)

Detector Monte Carlo

- Geant 3.21 (F77) with adaption for muon processes

  above 100 TeV (ionisation, pair production, bremsstrahlung)

- Geant 4 (C++) Test phase, should finally replace

  Geant 3.21

- km3 (F90, F77) Includes photon scattering in
- water ; based on prefilled tables - very fast

Reconstruction/Calibration

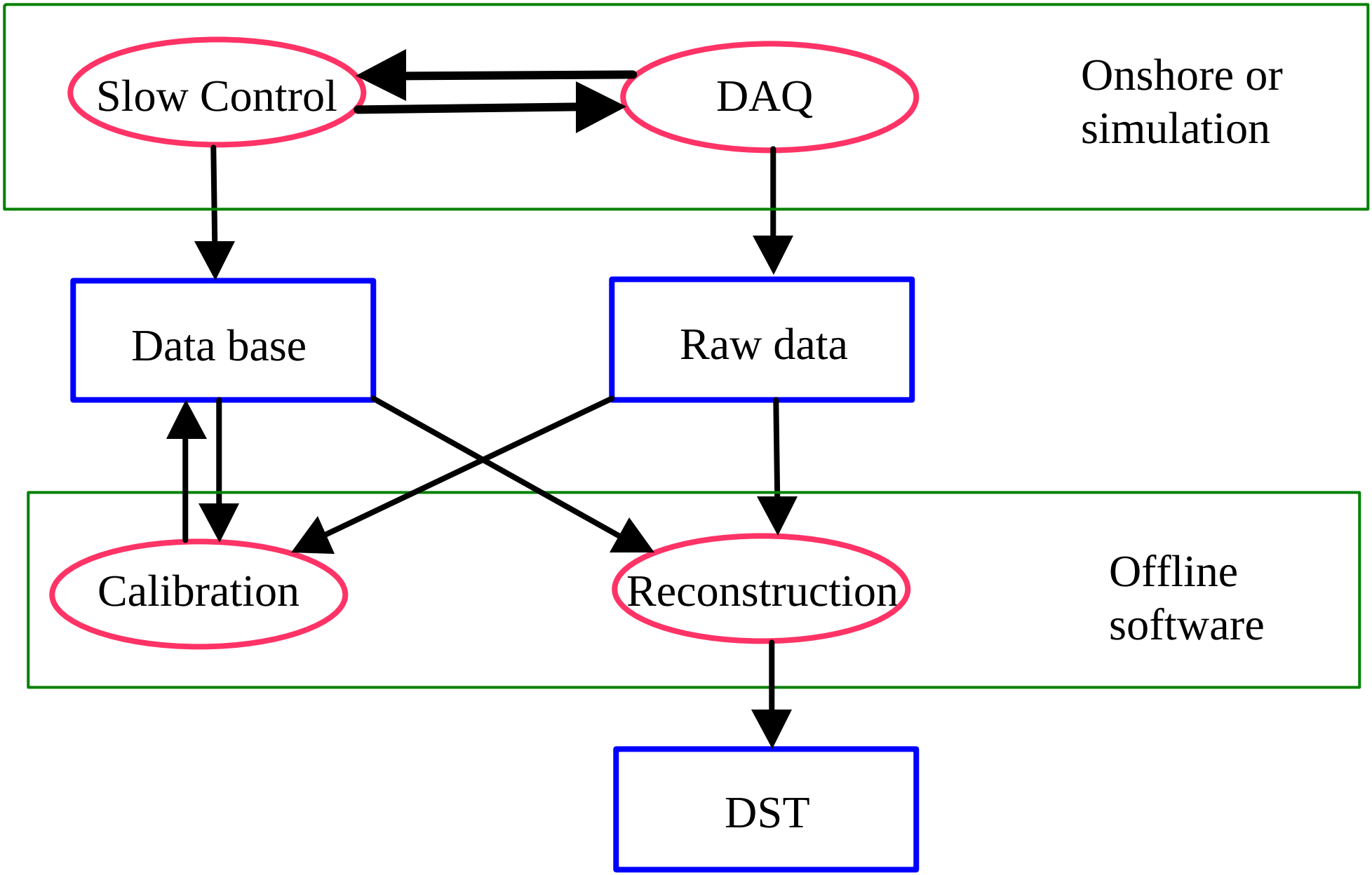
- Antcc C++ Common Framework of base classes, e.g.

  Hit, Track, TrackFitter, FitStrategy

  Detector, OM, String, GPSTime etc.

- Several algorithms developped in this frame
- 3DTrackFit, SingleStringFit, AcousticCalibration
- Development continues

Data Flow : Share between database and sequential data file



Use of CC-IN2P3 Resources

CPU / Tapes / Disk

Database / Accounts

CC-IN2P3 is THE main computer centre of Antares

- all raw data will be stored there
- software repository under CVS
- main run/detector database installed there (Oracle)
- Web server http://antares.in2p3.fr
- Mailing lists under listserv

Development of CPU use

(Lyon units)

Year	1998	1999	2000	2001
Hours	142,000	134,000	230,000	290,000 (June)

Main platform: Linux (gcc, g77, g++)

Increasing CPU need (larger collaboration)
(up to now only Monte Carlo)
typical job: small I/O bandwidth but long (300,000 sec)
Since 2000:

Largest MC production of atmospheric muon showers transferred to RAL
However: Reconstruction at Lyon
(2001: several new algorithms have been tested)

Common Disk Space

GROUP_DIR	36 Gbyte
THRONG_DIR	2 Gbyte

Software

Compiler: gnu (egcs) g77, gcc, g++ (STL)

commercial: F90

library: cernlib, clhep, root, geant4

commercial: naglib

Tapes

Year	1997 -1999	2000	2001 (estim.)
3490	271	90	150
DLT (import)		4	10
9840		5	10
Total(Gbyte)	270	350	700

2000: 350 Gbyte tape space used
2001: 700 Gbyte tape sspace needed
most small files (100 Mbyte)

Accounts for an international collaboration

Last 3 years: Antares has substantially grown

1997: CPPM, CEA, Valencia, Oxford, Sheffield
1998: Birmingham, IRES, GRPHE, Moscou (52 accounts)
1999: NIKHEF, Bologna, Genova (81 comptes, 27 non-IN2P3)
2000: Catania, Bari, Roma (101 comptes, 39 non-IN2P3)

Collaboration size now: 169 persons, 87 physicists

Data base

Antares will use one main database for the following type of informations:

slow control
detector geometry
Calibration
Book keeping

Choice: Oracle

Master copy at Lyon
Slave copy at Sablette (shore station)
Update once per night

Network

From 2003 on:
Problem of data transfer between Sablettes and Lyon

atm. muons:	30 Hz trigger rate
noise:	1500 Hz trigger rate
event size:	4 kbytes

Output DAQ: 0.1-6 Mbytes/sec
per year: 3-150 Tbyte
Local DLT + manuel transfer or transfer via network ?
(autonomy 1 month - 1 day with robot 8 DLT)

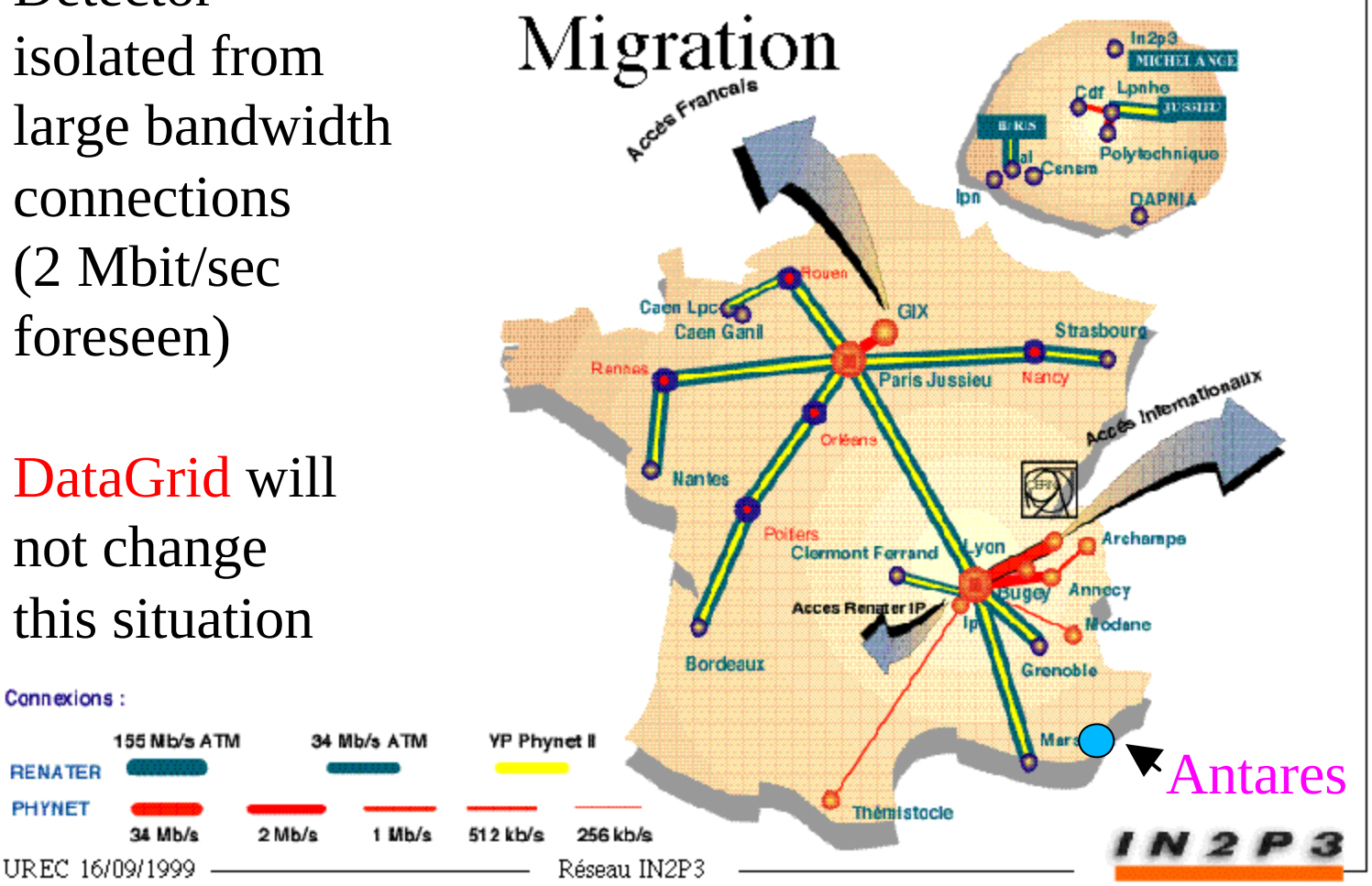
Optional:
Use advanced technology for first pass offline filter
DataGRID

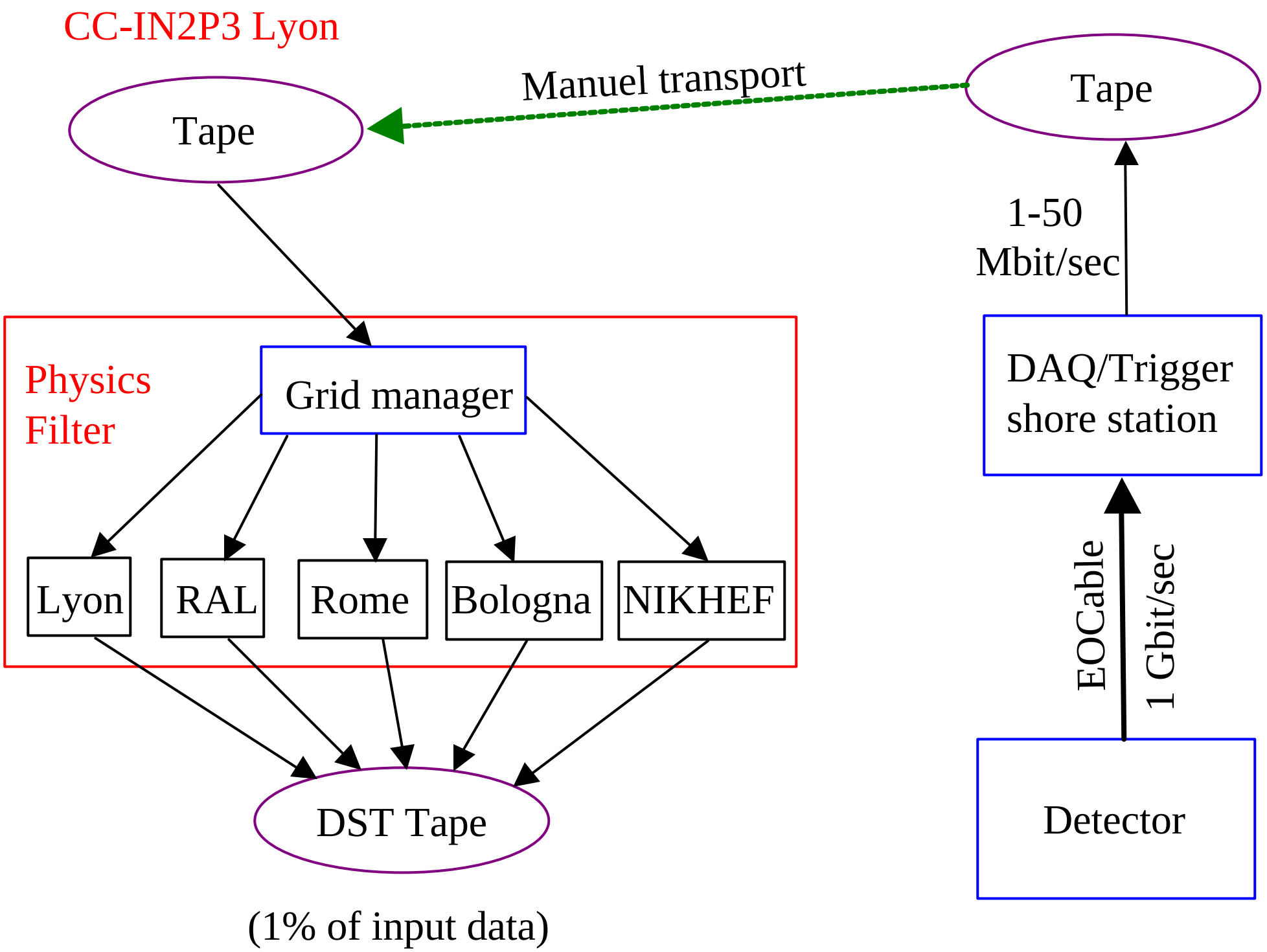
ANTARES

Data Repository and Main Computing Center: CC-IN2P3Lyon

Detector isolated from large bandwidth connections (2 Mbit/sec foreseen)

DataGrid will not change this situation




CC-IN2P3 Lyon
Tape
Manuel transport
Tape
1-50
Mbit/sec
Physics
Filter
Grid manager
DAQ/Trigger
shore station
Lyon
RAL
Rome
Bologna
NIKHEF
EOCable
1 Gbit/sec
DST Tape
Detector
(1% of input data)

CPU Time Constraints

- 10^9 -$5x10^{10}$ events per year
- Filter needs track reconstruction

(non-linear minimisation): estimation

4 Lyonsec/event

- Per year 10^6-$5x10^7$ Lyon-hours

(1998-2000 Antares CPU use: 140,000 Lh)

- Too much for single batch farm

Application for Data Grid

Characteristics of Antares Filter

- Input stream: time ordered "events"
- No complicated hierarchical structure of event stream
- Easy to parallelise on basis of "events" or bigger time slices ("runs")
- No multi-user interface needed (main production)
- High bandwidth connection between Grid -farms needed
- Ideally shore station included as well
- output files small, can be copied where ever needed

Possible Schemes

Scheme 1

- Master copy of all data in Lyon
- Computing shared by local centers (RAL, NIKHEF, INFN)
- Cut data into time slices
- Send copy of certain slice to center with free resources
- Run identical reconstruction code everywhere
- Analysis code accesses reconstruction output, which remains distributed

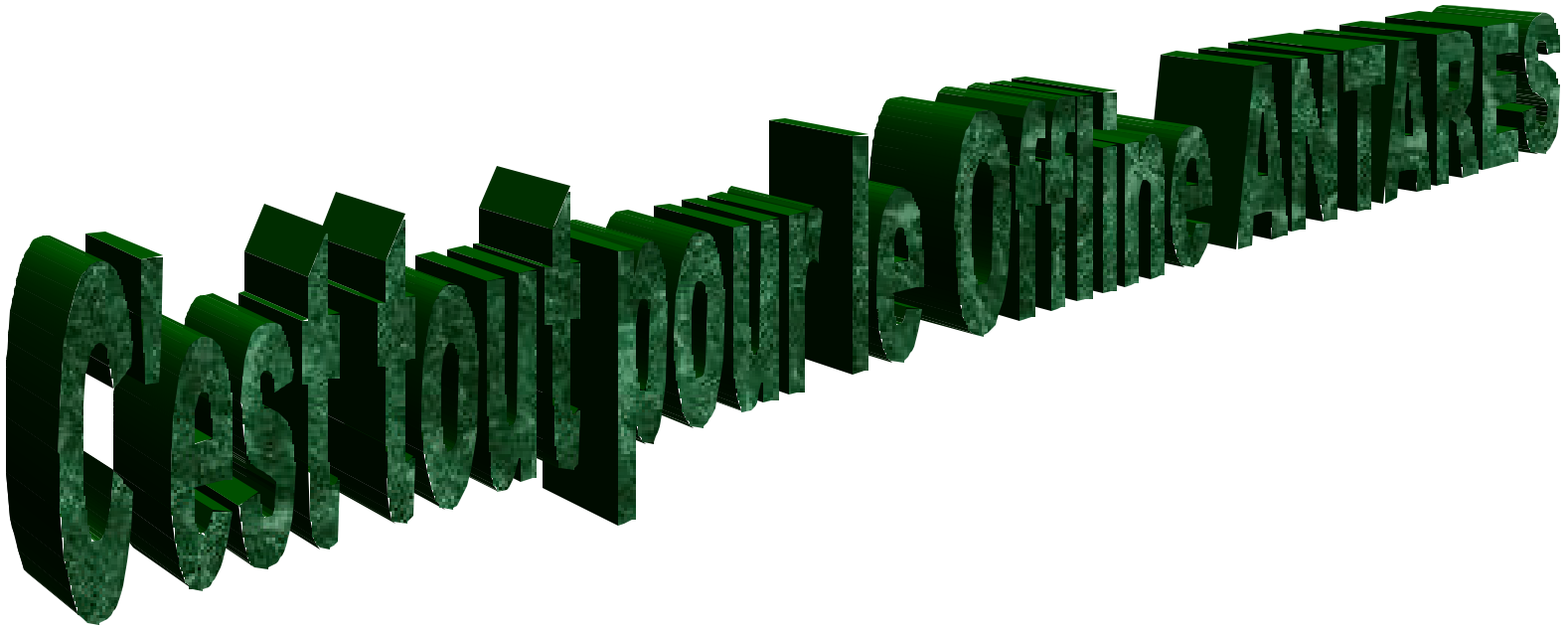
Possible Schemes

Scheme 2

- Master copy of all data in Lyon
- Each local center specializes on one analysis topic (muon-tracks, low energy, supernovae, monopoles, neutralinos)
- A working copy of ALL data is sent to each local center (alternative: direct access to data via network)
- The results of a certain anaysis can be obtained from the specialized center

Schedule

- Offline filter of Antares could be Testbed for DataGrid project
  - Earlier than LHC
  - Data rate factor10-100 smaller
  - relatively simple proram structure
- Bologna group is working on first version of offline filter, no time estimate yet


C'est tout pour le Offline ANTARES