# LHCb Computing Model and updated requirements

John Harvey

# LHC Computing Review

- Three technical review panels
  - World wide analysis
  - Software
  - Management and resources
- First meeting of each panel took place end March
- LHCb presentations available on web
  - http://lhcb.cern.ch/computing/Steering/Reviews/Hommann-2000/default.htm
- Draft LHCb Technical Notes being prepared (see web)
  - Baseline Computing model
  - Answers to SPP questions
- Invite feedback
- Next meetings April 14th (Panel 3), May 12th (Panels 1,3)

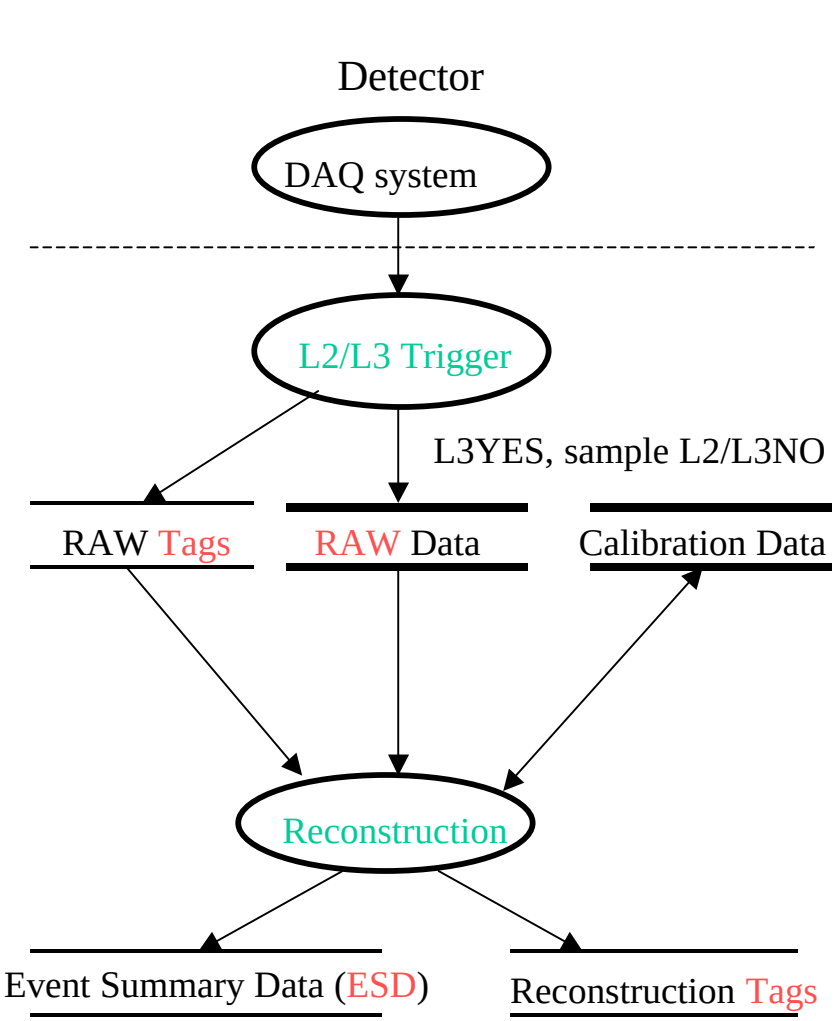
# Panel 1 : World-wide analysis

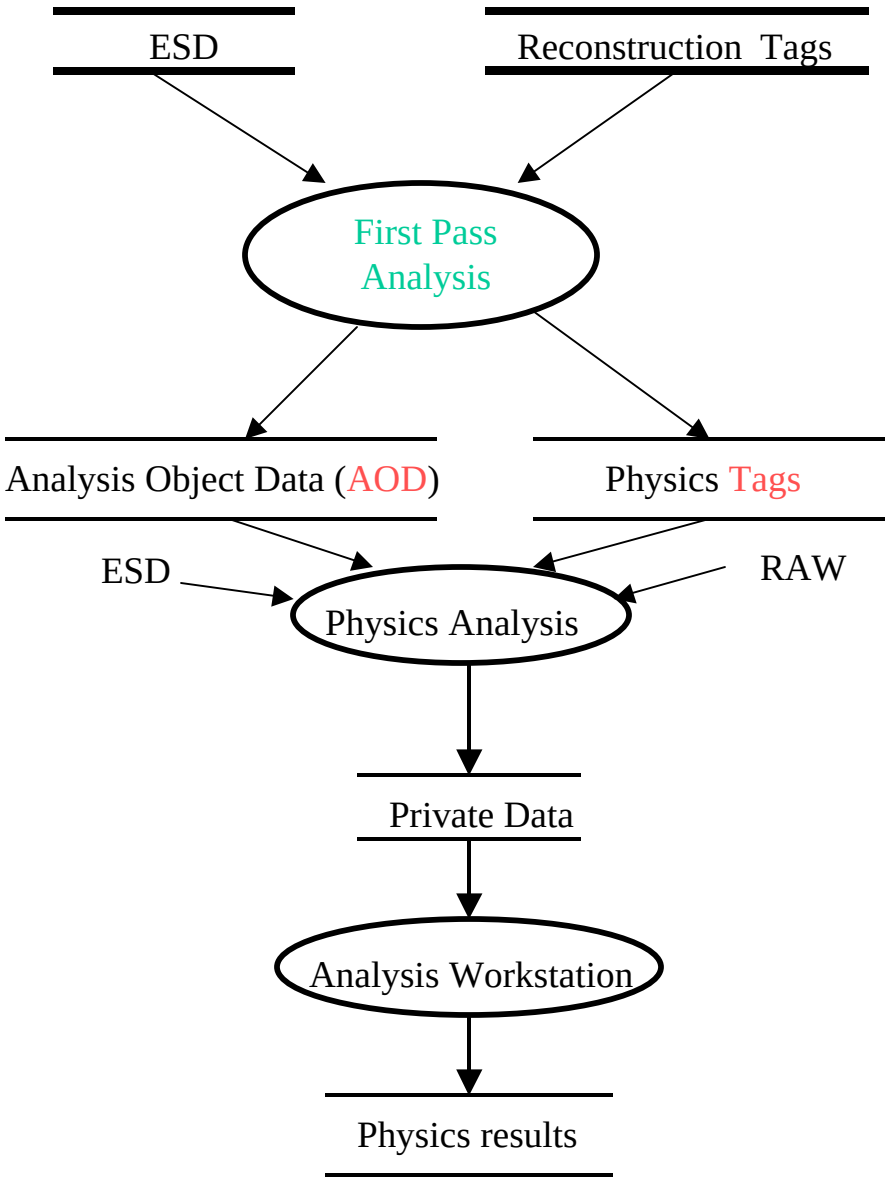
- Give a description of baseline computing model, indicating how it differs from MONARC generic model.
- What is the current tentative distribution of the different regional centres?
- Give a tentative assessment for data storage and management and for processing power required at the different centres (now and at the different stages towards LHC start-up).
- Explain your current understanding of your policy regarding magnetic tapes versus disks.
- Given the current status of the implementation of the above model, is the experiment involved/interested in the current GRID test activities?
- Do you plan Mock Data Challenges or equivalent test-beds to validate your model and provide intermediate services between now and LHC start-up?
- What are the resources available/planned for the implementation of the above model (again staggered between now and LHC start-up)?

# Panel 3 : Management and Resources

- Compare computing needs in terms of hardware, software, organization, and manpower, to produce a plan to be presented to the funding agencies.
- Produce an "Interim version of the Memorandum of Understanding " indicating responsibilities of institutions
  - ➢ Include responsibilities for writing and maintenance of the software as has been done for construction of the detector.
- First meeting focused on first 6 months of data-taking:
  - ➢ Prepare a description of the computing systems needed to calibrate the subdetectors and the related strategy to get from raw data to first "simple" physics analysis.
- "world wide physics analysis" left to a following meeting.

# Dataflow Model



# Real Data Storage Requirements

| | | | |
|---|---|---|---|
| Length of period | 120 days | $10^7$ secs | |
| LHC duty cycle | 50% | | |
| Event rate stored | 200 Hz | $10^7$ per day | $10^9$ per year |
| RAW data size | 100 kB/event | 1 TB/day | 100 TB/yr |
| ESD data size | 100 kB/event | 1 TB/day | 100 TB/yr |
| AOD data size | 20 kB/event | 0.2 TB/day | 20 TB/yr |
| TAG data size | 1 kB/event | 0.01 TB/day | 1 TB/yr |

# Real Data CPU Requirements

| | | | |
|---|---|---|---|
| L2 trigger CPU | 0.25 SI95sec/event | @40 kHz | 10,000 SI95 |
| L3 trigger CPU | 5 SI95sec/event | @5 kHz | 25,000 SI95 |
| Reconstruction CPU | 250 SI95sec/event | @200 Hz | 50,000 SI95 |
| First Pass Analysis | 5 SI95/event | $2.10^8$ in 2 days | 5000 SI95 |
| User analysis at RC | 20 SI95/event | | 10,000 SI95 |
| User analysis CERN | 20 SI95/event | | 20,000 SI95 |

- 1 SI95 = 40 MIPS
- Today 1 PC ~ 10 SI95
- By 2005 130 SI95/cpu (low cost), 250 SI95 (high end servers)

# User Analysis Requirements

- Assume that physicist performs a production analysis and requires a response time of 4 hours
- The ~$10^7$ events tagged by first pass analysis are scanned and candidates selected (0.25 SI95 /event)
- The selected candidates are subjected to analysis algorithm (20 SI95 / event)
- Total installed cpu power needed calculated assuming:
  - ~140 physicists actively doing analysis
  - each submits 1 job / day ( NB. many short jobs as well)
  - analysis distributed over a number of regional centres (~5) and assume ~20 physicists at each Regional Centre, ~40 at CERN
  - Assume 0.3 x $10^7$ events selected for algorithm on average
  - 10,000 SI95 at each Regional Centre, 20,000 SI95 at CERN

# Simulation Requirements - Signal Events

- CPU power to simulate $10^7$ B -> D*$\pi$ events in 1 year
  - assume need to simulate 10 times real data sample ($10^6$)
  - installed capacity needed is 100,000 SI95

| Step | Number of events | Cpu time/evt | Total cpu power |
|---|---|---|---|
| Generator | $10^{10}$ | 200 SI95sec | $2.\ 10^{12}$ SI95sec |
| GEANT tracking | $10^9$ | 1000 SI95sec | $10^{12}$ SI95sec |
| Digitisation | $10^9$ | 100 SI95sec | $10^{11}$ SI95sec |
| Trigger | $10^9$ | 100 SI95sec | $10^{11}$ SI95sec |
| Reconstruction | $10^8$ | 250 SI95sec | $2.5\ 10^{10}$ SI95sec |
| First Pass analysis | $10^7$ | 20 SI95sec | $2.\ 10^8$ SI95sec |

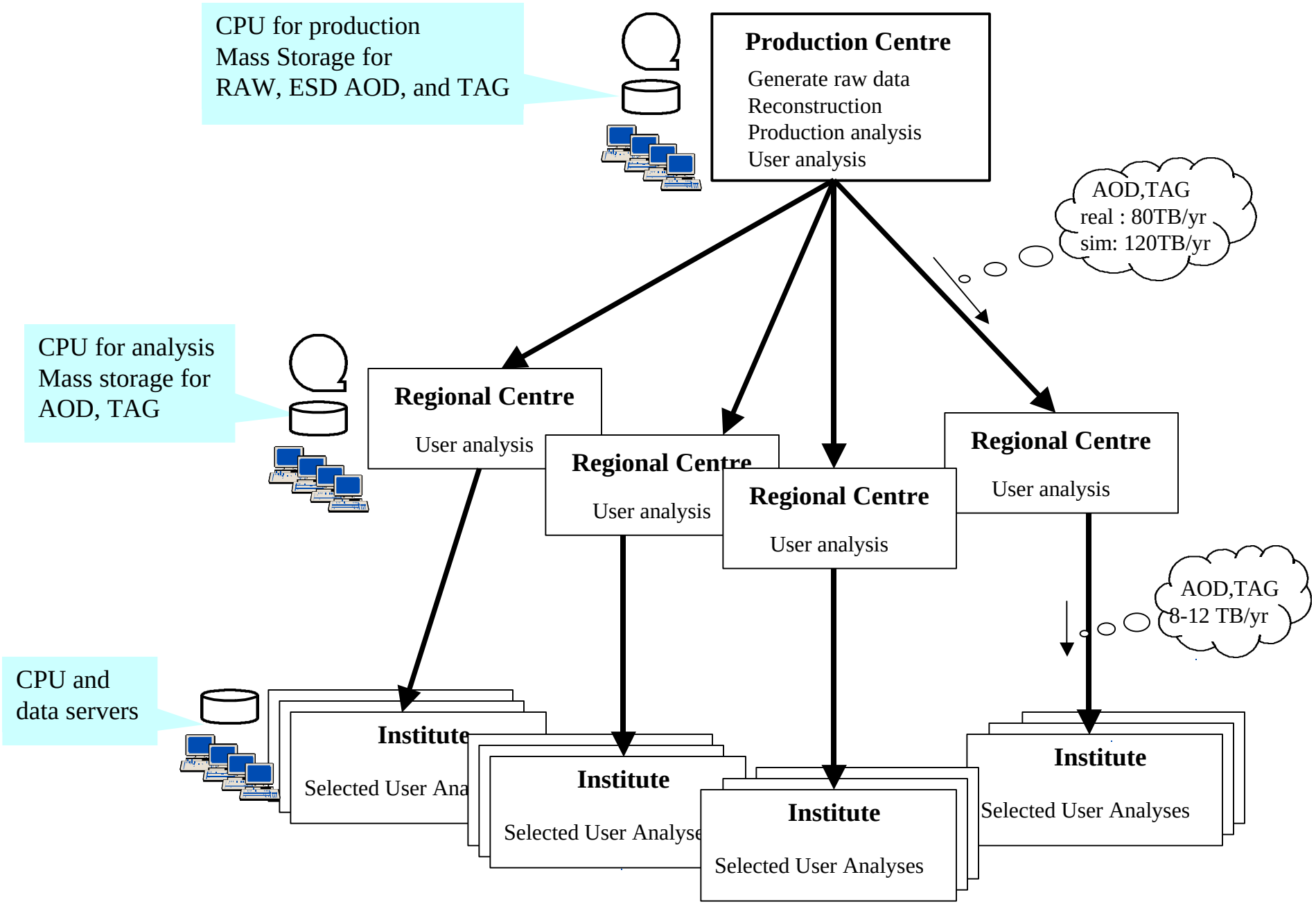
# Simulation Requirements - Background

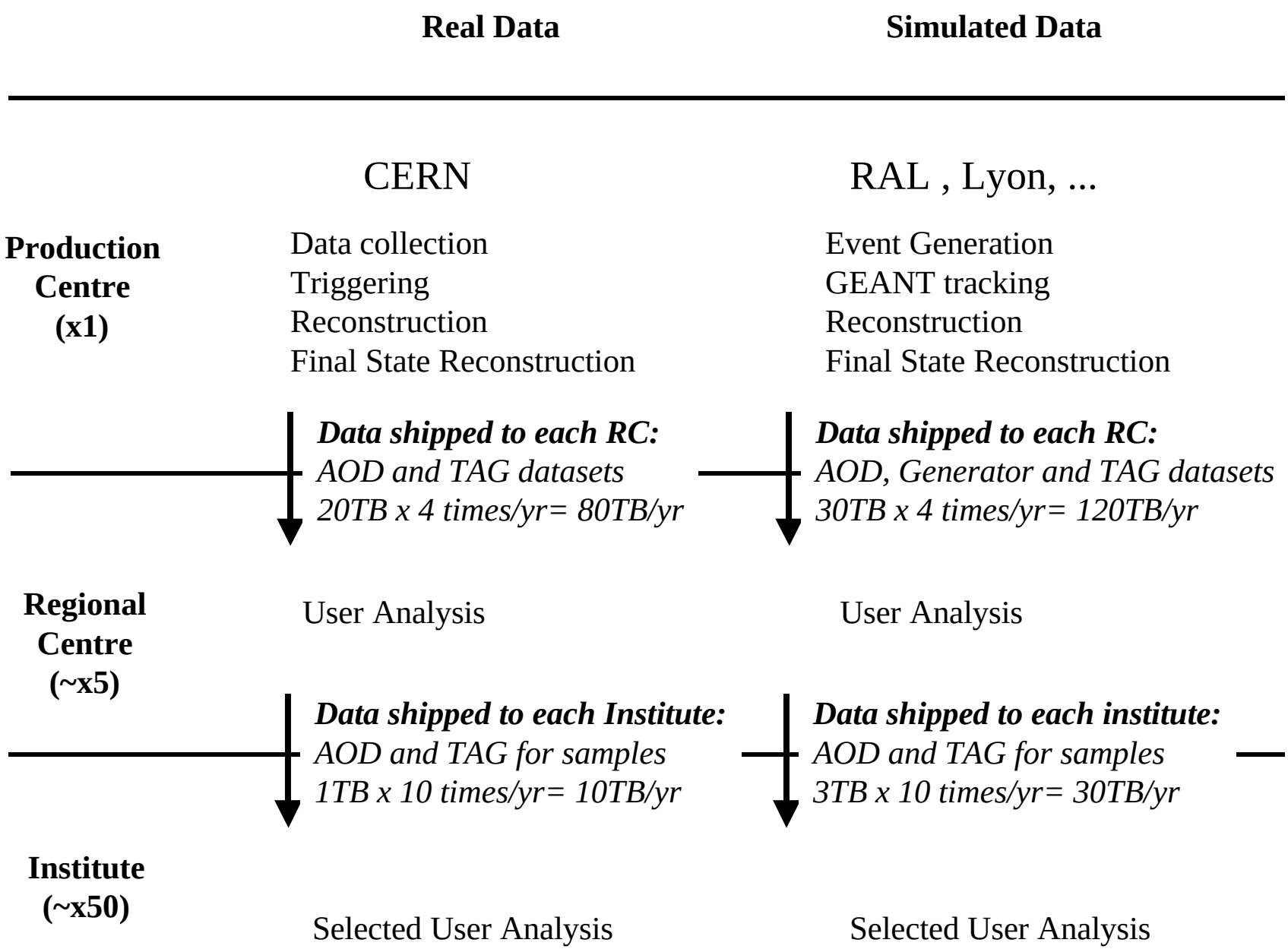
- $10^5$ $b\bar{b}$ inclusive events in detector every second
- ~100 Hz are recorded in real data
  - ➢ trigger efficiency $10^{-3}$
- If need as many to be simulated then need to generate, track, digitise and trigger $10^{12}$ bb inclusive events/yr and $10^9$ will have to be reconstructed
  - ➢ corresponds to 3. $10^{14}$ SI95 sec/yr i.e.10,000,000 SI95
- Obviously need to study ways of optimising background simulation
  - ➢ store and reuse data produced at generator level (storage!)
  - ➢ optimise generation step without biasing physics
  - ➢ focus on background particularly dangerous for a specific physics channel
  - ➢ aim to reduce requirements by > 1 order of magnitude
- Assume 400,000 SI95 required

# Simulation Requirements - Summary

| | | |
|---|---|---|
| RAWmc data size | 200 kB/event | 200 TB/$10^9$ events |
| Generator data size | 12 kB/event | 12 TB/$10^9$ events |
| ESD data size | 100 kB | 100 TB/$10^9$ events |
| AOD data size | 20 kB/event | 20TB/$10^9$ events |
| TAG data size | 1 kB/event | 1 TB/$10^9$ events |
| CPU power | ~100,000 SI95 signal events | ~400,000 SI95 background events |

- For comparison the experiments quote
  - ALICE (2,000,000), ATLAS/CMS (1,500,000)


CPU for production
Mass Storage for
RAW, ESD AOD, and TAG
Production Centre
Generate raw data
Reconstruction
Production analysis
User analysis
AOD,TAG
real : 80TB/yr
sim: 120TB/yr
CPU for analysis
Mass storage for
AOD, TAG
Regional Centre
User analysis
Regional Centre
User analysis
Regional Centre
User analysis
Regional Centre
User analysis
AOD,TAG
8-12 TB/yr
CPU and
data servers
Institute
Selected User Ana
Institute
Selected User Analyse
Institute
Selected User Analyses
Institute
Selected User Analyses


Real Data
Simulated Data
CERN
RAL , Lyon, ...
Production Centre (x1)
Data collection
Triggering
Reconstruction
Final State Reconstruction
Event Generation
GEANT tracking
Reconstruction
Final State Reconstruction
Data shipped to each RC:
AOD and TAG datasets
20TB x 4 times/yr= 80TB/yr
Data shipped to each RC:
AOD, Generator and TAG datasets
30TB x 4 times/yr= 120TB/yr
Regional Centre (~x5)
User Analysis
User Analysis
Data shipped to each Institute:
AOD and TAG for samples
1TB x 10 times/yr= 10TB/yr
Data shipped to each institute:
AOD and TAG for samples
3TB x 10 times/yr= 30TB/yr
Institute (~x50)
Selected User Analysis
Selected User Analysis

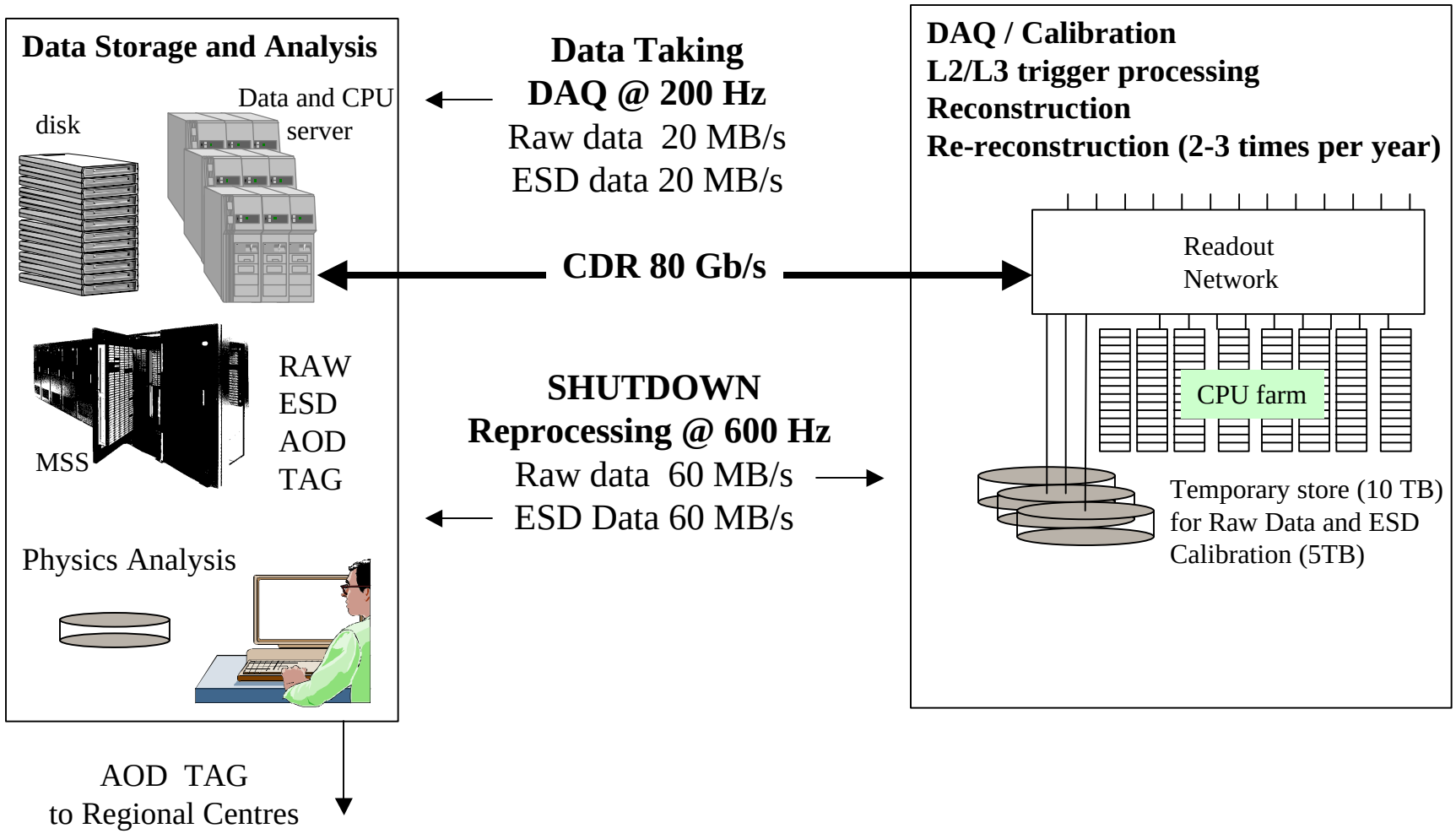
# Computing at CERN

- Run high level triggers and reconstruction on same cpu farm located at LHCb pit
- Send RAW and ESD data over the CDR link (80 Gbps) to computer centre for archive, and first pass analysis
- Maintain local storage at pit in case CDR down
  - ➢ accumulate 2 TB/day, therefore need >10 TB
- Dispatch AOD and TAG etc to regional centres
- During shutdowns, down periods do re-processing of RAW data on the farm in the pit
  - ➢ read RAW back from computer centre
  - ➢ send new ESD from pit to computer centre
  - ➢ full farm available so proceeds at twice the rate
  - ➢ allows 2-3 reprocessings of complete year's data
- Flexible, efficient and maintainable solution

# Compute Facilities at CERN

CERN Computer Centre

Experiment - LHC Pit 8



# Facility at Pit - Requirements

| | |
|---|---|
| CPU Farm | ~100,000 SI95 |
| Disk storage event buffer | > 10 TB |
| Disk storage calibration and secondary data | > 5TB |
| CDR link capacity (80 Gb/s) | 1 Gb/s |

# CERN Computer Centre Requirements

| | |
|---|---|
| RAW data storage | 100 TB/yr |
| Copy RAW data storage | 100 TB/yr |
| ESD data storage | 100 TB/yr |
| AOD data storage | 4 x 20 TB/yr |
| TAG data storage | 1 TB/yr |
| AODmc, Generator storage | 120 TB (30 TB imported 4 times/yr) |
| TAGmc data storage | 4 TB (1 TB imported 4 times/yr) |
| Total data storage | ~500 TB / yr |
| CPU for First Pass analysis | 2000 SI95 |
| CPU for user analysis | 20,000 SI95 |
| WAN for AOD TAG export | 80 TB/yr |
| WAN for AOD TAG import | 124 TB/yr |

# Simulation requirements 2000-2005

- 2000-2001
  - $10^7$ simulated events/yr for detector optimisation studies
  - prepare TDRs
- 2002-2003
  - $2.\ 10^7$ events/yr for high level trigger studies
- 2004 - 2005
  - start to install and commission large scale facilities
  - start to produce large samples of background events with the final detector description
  - ~$10^8$ simulated events/yr
- >2001
  - use simulation and MDC to test computing model
  - contribute to HEP Application WP of EU grid proposal

# Sizing Estimate for Regional Centre

| | 2000-2001 | 2002-2003 | 2004-2005 | >2005 |
|---|---|---|---|---|
| AOD TAG | | | | 80TB/yr |
| AODmc TAGmc imported | 2TB/yr | 5TB/yr | 20TB/yr | 120 TB/yr |
| CPU analysis | 3000 SI95 | 5000 SI95 | 10000 SI95 | 10000 SI95 |
| RAWmc, ESDmc AODmc TAGmc generated | 5TB/yr | 10TB/yr | 33TB/yr | 333TB |
| CPU mc production | 20000 SI95 | 40000 SI95 | 60000 SI95 | 100000 SI95 |

# Assignment of responsibility

- Understood in LHCb institutes building subdetectors also take responsibility for development and maintenance of software
- The detector TDRs are in preparation now
- MOU after TDRs

# Discussion

- How will our computing needs evolve between now and 2005?
- What regional centres will LHCb use for satisfying these needs? (RAL, CCIN2P3/Lyon, ++...)
- What resources (cpu, storage) will be available for satisfying our simulation needs?
- What is our attitude towards making an MOU for computing? Including software? What timescale?
- What level of engagement should we take in Grid Projects?

# GRID LHCb WP Physics Study

- The total sample of B > $J\Psi/K_s$ simulated events needed is ~10 times the number produced in the real data.
- In one year of datataking we expect to collect and fully reconstruct $10^5$ events, therefore need $10^6$ simulated events.
- The number of events that have to be generated, stored and reconstructed to produce this sample is $10^7$.
- 10% of the ESD data copied for systematic studies (~100 GB).
- The total amount of data generated in this production would be :

| | | | |
|---|---|---|---|
| RAW data | 200 kB/event | x $10^7$ | = 2 .0 TB |
| Generator data | 12 kB/event | x $10^7$ | = 0.12 TB |
| ESD data | 100 kB/event | x $10^7$ | = 1 .0 TB |
| AOD data | 20 kB/event | x $10^7$ | = 0. 2 TB |
| TAG data | 1 kB/event | x $10^7$ | = 0.01 TB |

# Grid LHCb WP - Grid Testbed

- MAP farm at Liverpool has 300 processors, would take 4 months to generate the full sample of events
- All data generated (~3TB) would be transferred to RAL for archive (UK regional facility).
- All AOD and TAG datasets dispatched from RAL to other regional centres, such as Lyon and CERN.
- Physicists run jobs at the regional centre or ship AOD and TAG data to local institute and run jobs there. Also copy ESD for a fraction (~10%) of events for systematic studies (~100 GB).
- The resulting data volumes to be shipped between facilities over 4 months would be as follows :

| | | |
|---|---|---|
| Liverpool to RAL | 3 TB | (RAW ESD AOD and TAG) |
| RAL to LYON/CERN/… | 0.3 TB | (AOD and TAG) |
| LYON to LHCb institute | 0.3 TB | (AOD and TAG) |
| RAL to LHCb institute | 100 GB | (ESD for systematic studies) |