

# The KLOE computing environment

**M. Moulson – INFN/Frascati  
for the KLOE Collaboration**

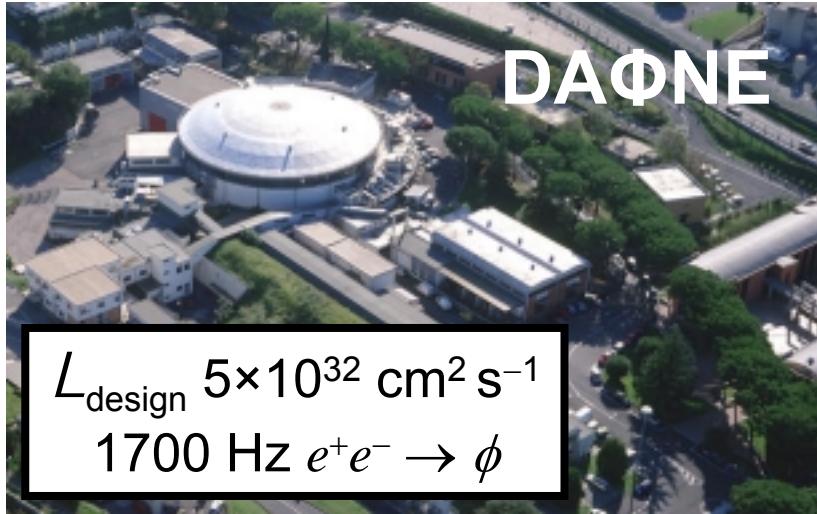


Nuclear Science  
Symposium

Portland, Oregon, USA • 20 October 2003

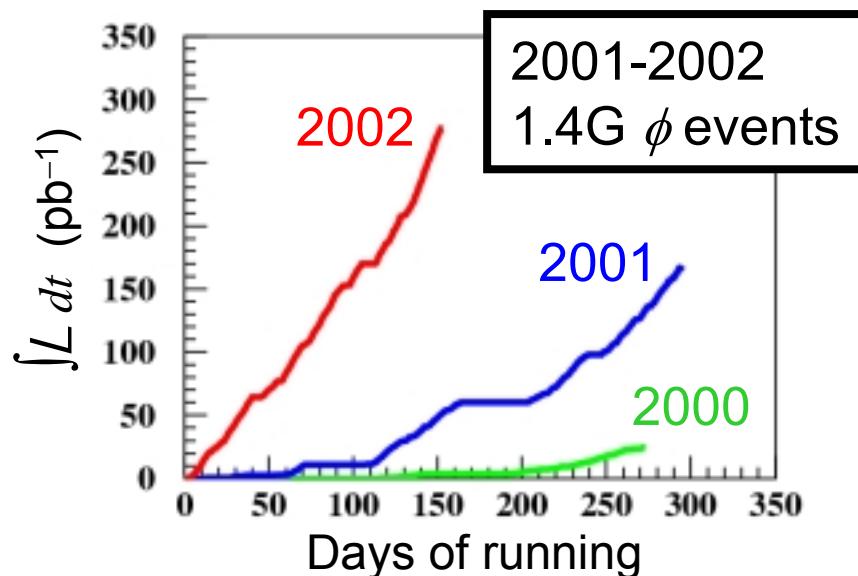
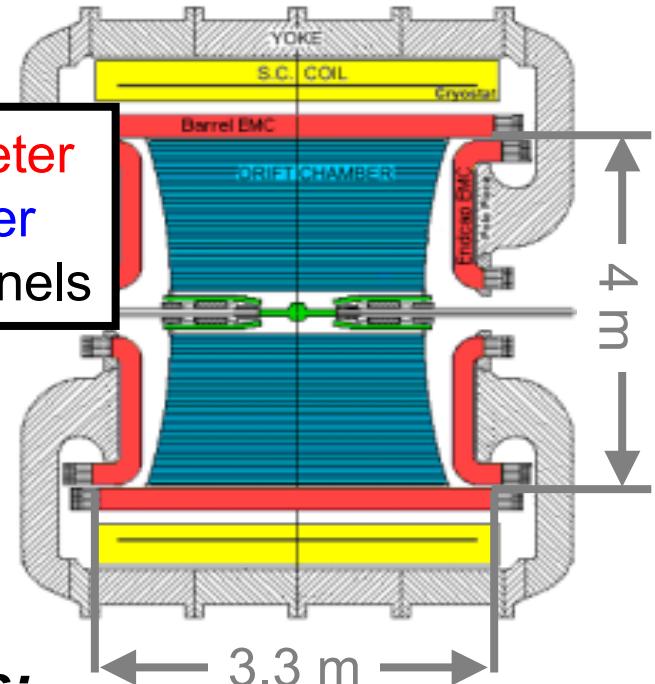


# The KLOE experiment at DAΦNE



KLOE

EM calorimeter  
Drift chamber  
23000 channels

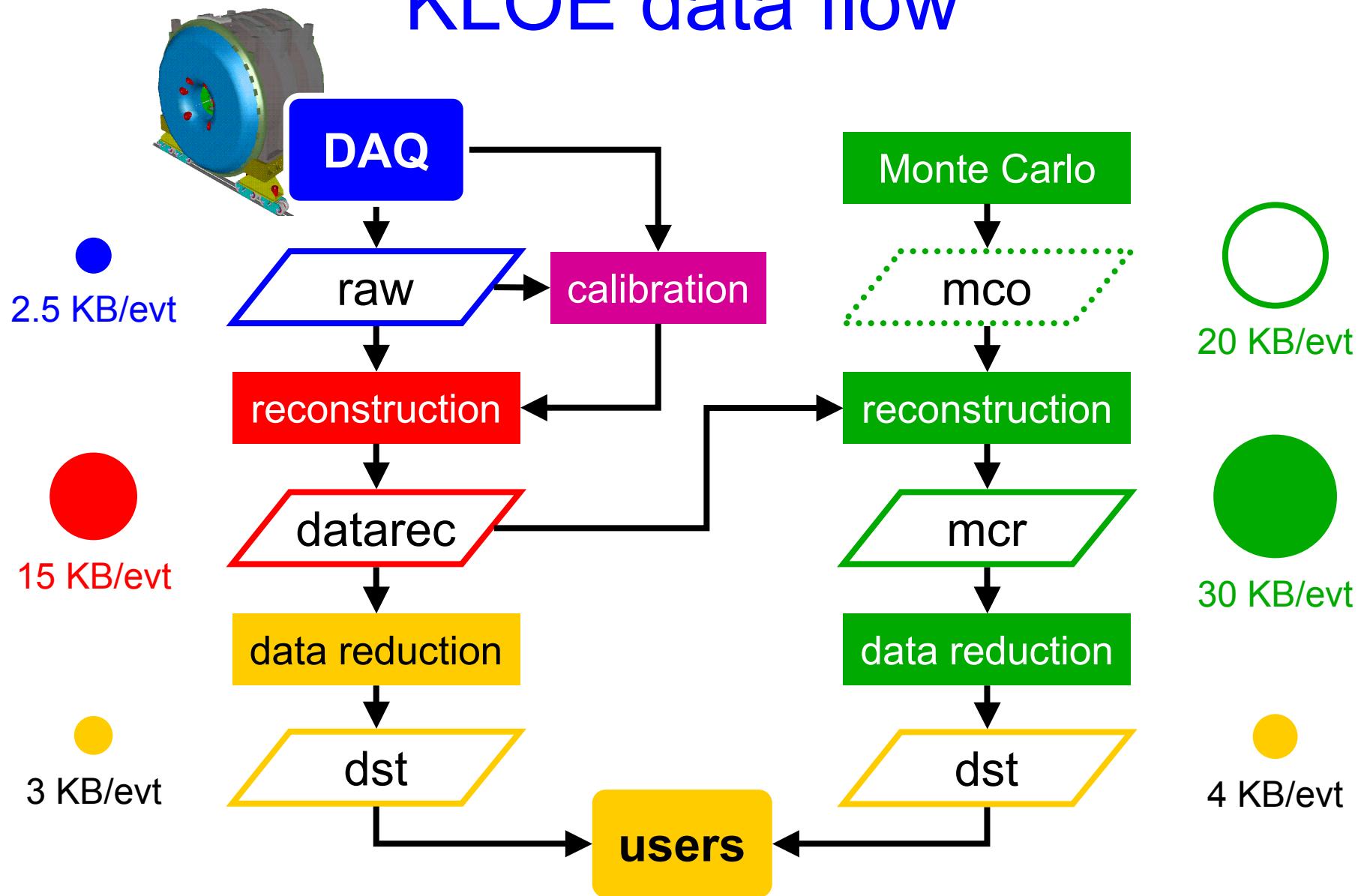


**Physics objectives:**

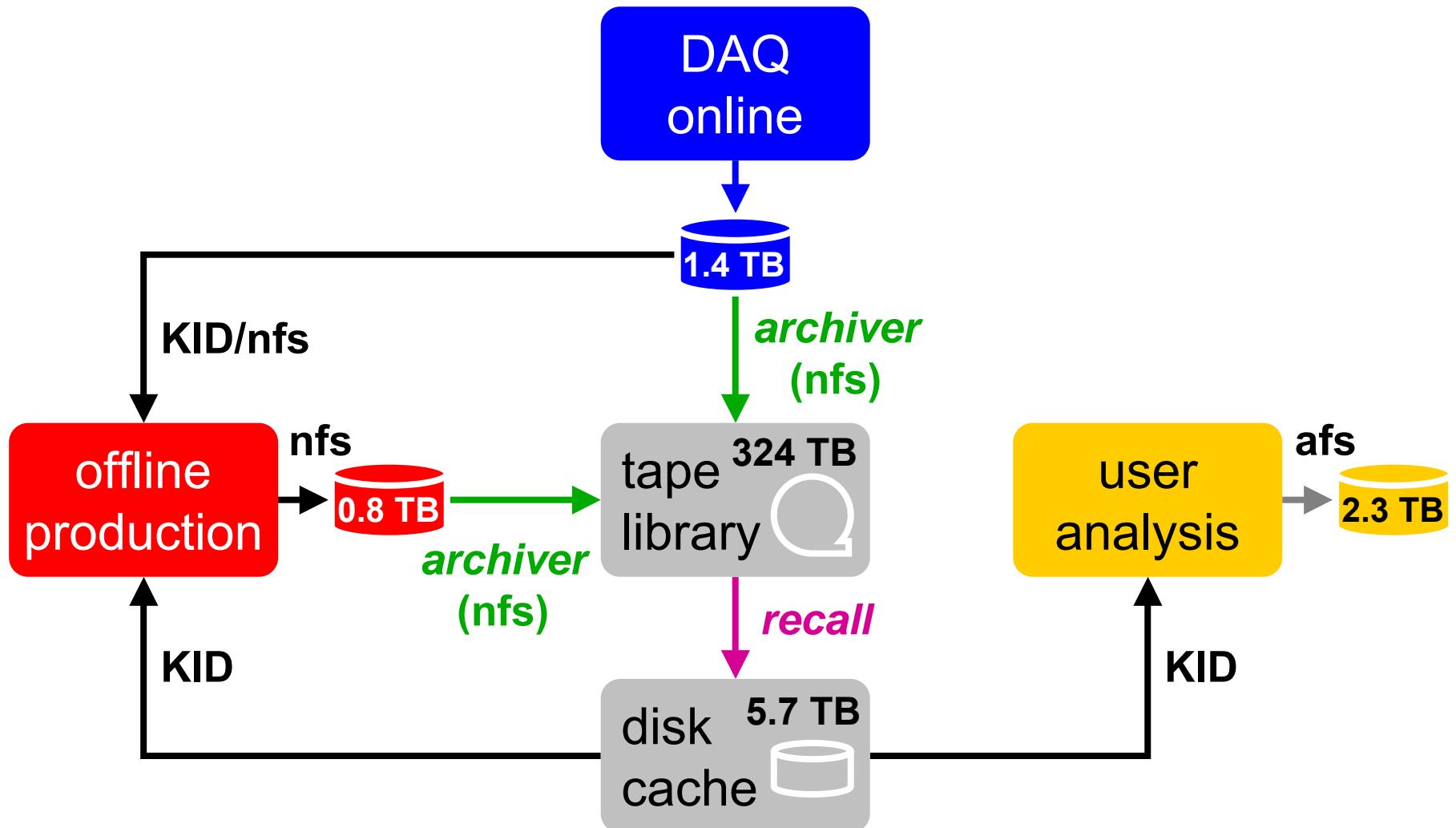
- $CP/CPT$  symmetries in  $K_S K_L$  system
- $K^\pm$ , radiative  $\phi$ ,  $\eta$  decays
- $\sigma(e^+e^- \rightarrow \text{hadrons})$

**All  $\phi$  events interesting**

# KLOE data flow



# Data handling



# Data handling

## File database: IBM DB2

Tracks location and status of several million files

## Data-handling interface: *KLOE Integrated Dataflow (KID)*

Access to data regardless of file location via URL's

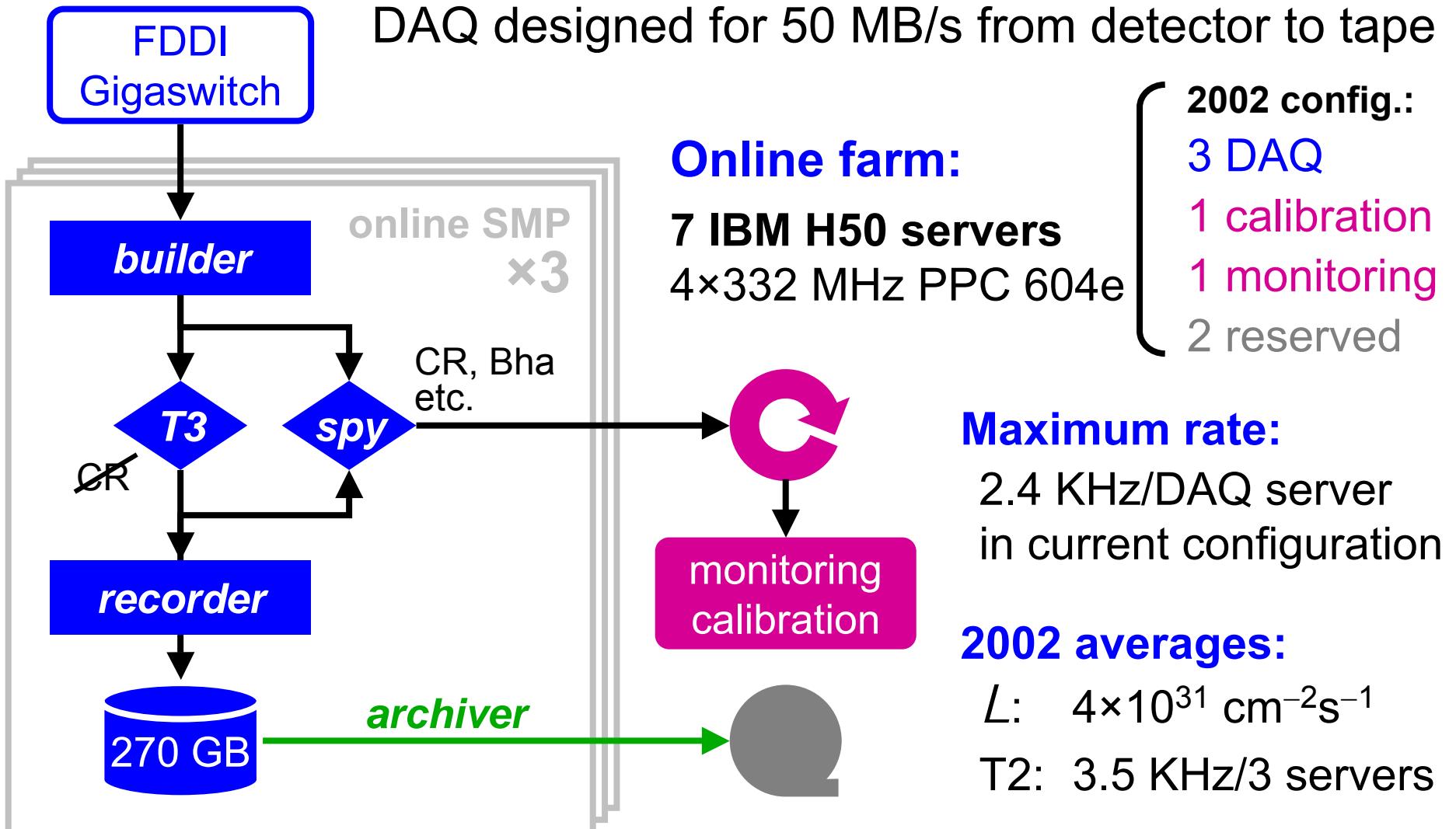
KID URL's can incorporate SQL query for file database:

`dbraw:run_nr=23015`

`dbdatarec:run_nr=23015 and  
stream_code='ksl'`

`dbraw:run_nr between 23000 and 24000 and  
analyzed is not null`

# Online computing



# Offline computing

## **Hardware environment:**

**16 to 19 IBM B80 servers:** 4×375 MHz POWER3

**8 Sun E450 servers:** 4×400 MHz UltraSPARC-II

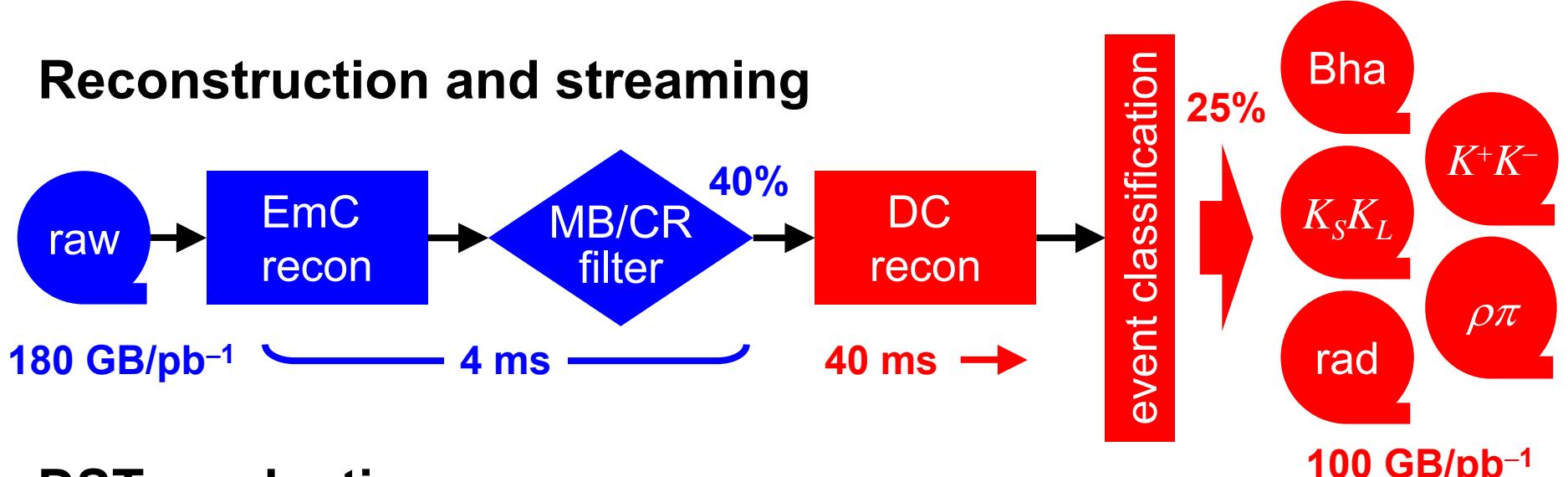
**800 GB** NFS-mounted disk for I/O staging

## **Software environment:**

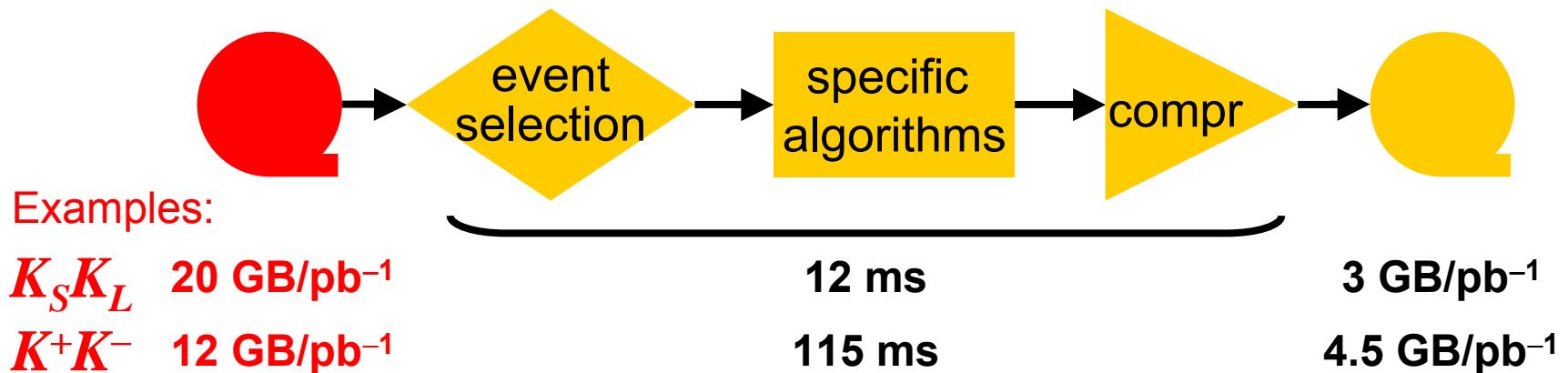
**ANALYSIS\_CONTROL** with KID interface

# Offline computing

## Reconstruction and streaming



## DST production



# Monte Carlo production

**Simulated event samples statistically comparable to data**

$\phi \rightarrow \text{all}$	450 pb <sup>-1</sup> at 1:5 scale	260M events
$\phi \rightarrow K_S K_L$	450 pb <sup>-1</sup> at 1:1 scale	430M events

**Each run in data set individually simulated**

- $\sqrt{s}$ ,  $\mathbf{p}_\phi$ ,  $\mathbf{x}_\phi$ , background, dead wires, trigger thresholds...

**GEANT 3.21-based simulation**

**Inclusion of accidental activity from machine background**

- Extracted from  $e^+e^- \rightarrow \gamma\gamma$  events in data set
- Inserted run-by-run to match temporal profile of data

**MC DST's to provide convenient user interface**

# Analysis environment

## Production of histograms/Ntuples on analysis farm:

4 to 7 IBM B80 servers + 2 Sun E450 servers

DST's latent on 5.7 TB recall disk cache accessed by KID

Batch processes managed by IBM LoadLeveler

Output to 2.3 TB AFS cell accessed by user PC's

## Final-stage analysis on user PC/Linux systems

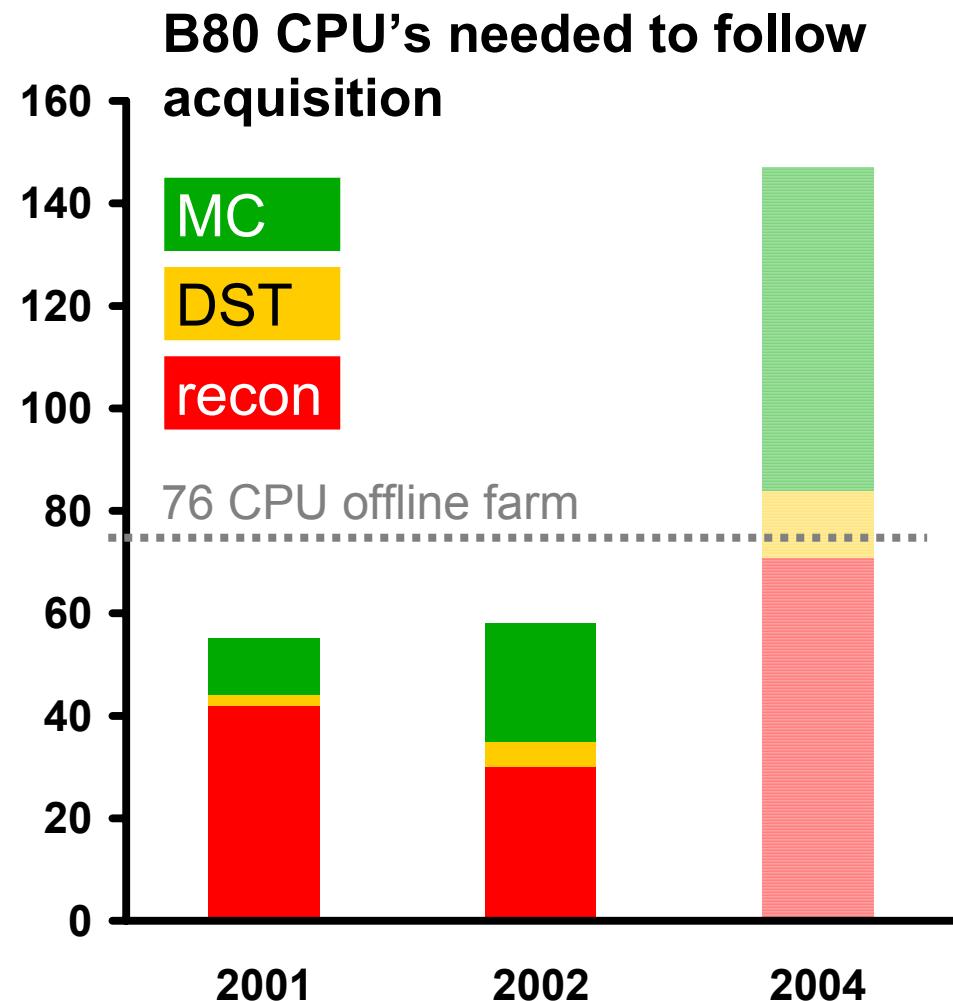
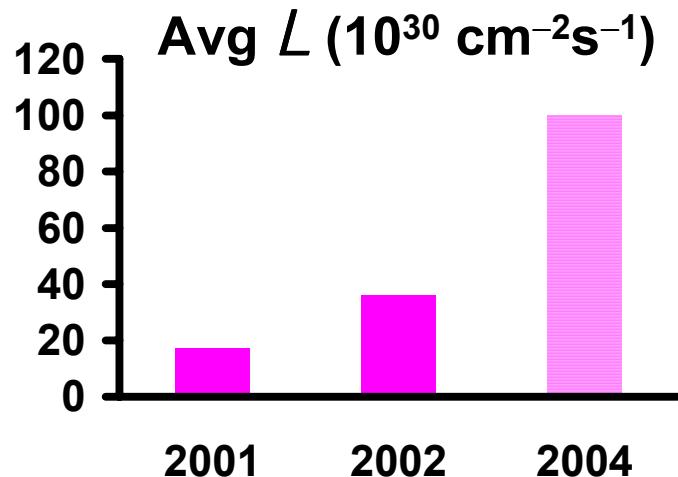
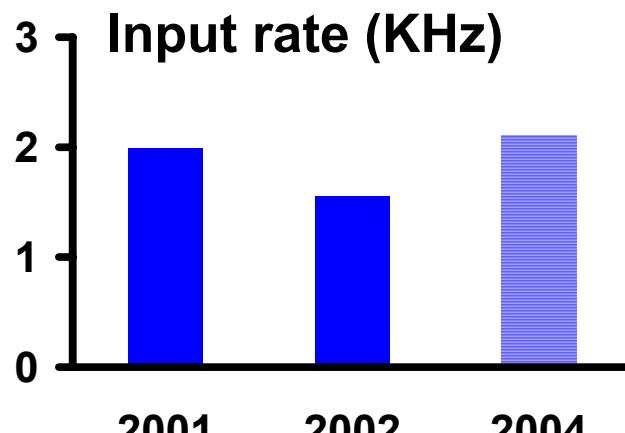
### Analysis example:

700M  $K_S K_L$  events, 1.4 TB DST's

6 days elapsed for 6 simultaneous batch processes

Output on order of 10-100 GB

# CPU power requirements



# Mass storage requirements

## Installed hardware:

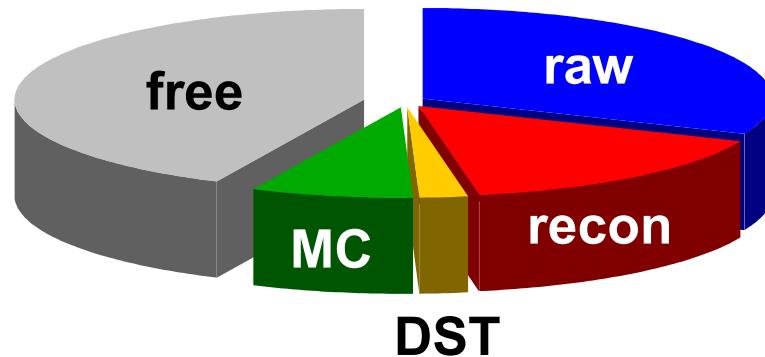
**5.7 TB recall disk cache (DST)**

**IBM 3494 tape library**

- 12 Magstar 3590 drives (14 MB/s)
- 60 GB/cartridge
- 5400 slots, 2 accessors
- Tivoli Storage Manager

**Max. capacity: 324 TB**

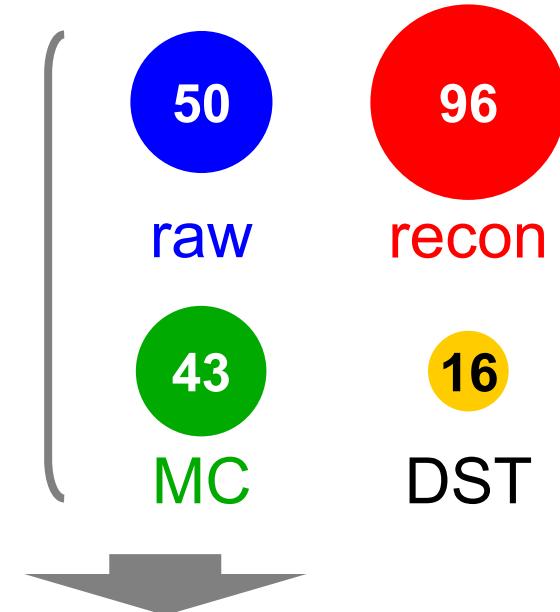
**In use: 185 TB**



## Predicted needs for 2004:

$2 \text{ fb}^{-1}$  at  $L = 1 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$

Total:  
205  
GB/pb $^{-1}$



410 TB tape storage  
16 TB DST disk cache (50%)

# Upgrades for 2004

**Additional servers for offline farm: ~80 B80 equivalents**

10 IBM p630 servers:  $4 \times 1.45$  GHz POWER4+

**Additional disk space: ~20 TB for DST cache and AFS cell**

**Additional IBM 3494 tape library: 300 TB**

Magstar 3592 drives: 300 GB/cartridge, 40 MB/s

Initially 1000 cartridges with space for 3600 (1080 TB)

2 accessors, 6 drives, remotely accessed via FC/SAN interface

**KLOE computing environment flexible and scalable  
Will be capable of handling  $10 \text{ pb}^{-1}/\text{day}$  in 2004**

# Additional information

# KLOE computing resources

