RSCCにおける
高エネルギー原子核物理実験データ解析

理化学研究所 放射線研究室 四日市 悟
S. Yokkaichi, Radiation Lab., RIKEN

- めざす物理
- 原子核／素粒子物理実験のデータの特徴
- RSCCでの計算：この一年
- これから

放射線研究室の activity

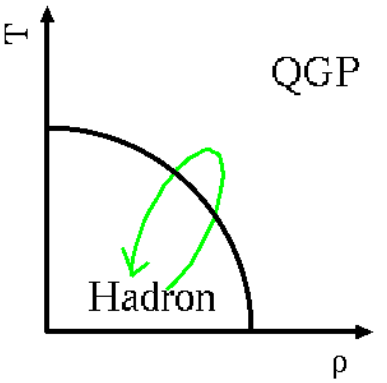
- ハドロン物理 hadronic physics
  - 陽子、中性子、pi, K etc.
    - 3 or 2個のクォーク：閉じ込めの謎
  - Why and how ?
- RHIC/PHENIX at BNL(in USA)
- KEK-PS
- CCJ (computing center in Japan) since 1999

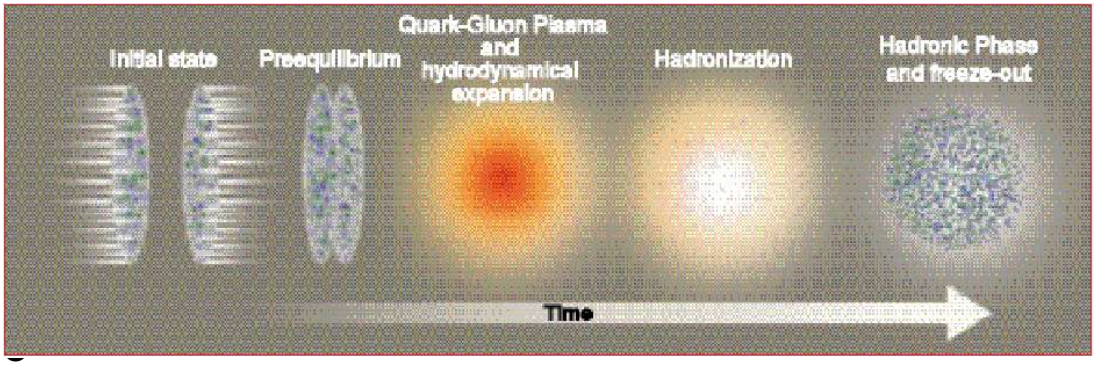
CPU farm
(332 CPU)

tape robot
(600TB)

現在のテーマ

- クォークグルオンプラズマの探索
  - published paper 24(+10),
    ~100 of conference proceedings

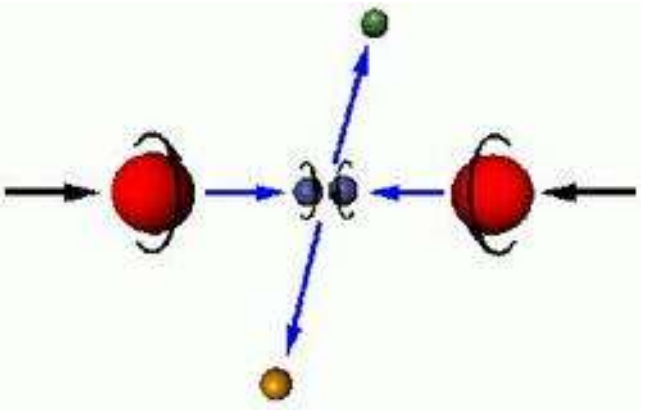




- 核子のスピンの起源の探索
  - first result from 2003 data

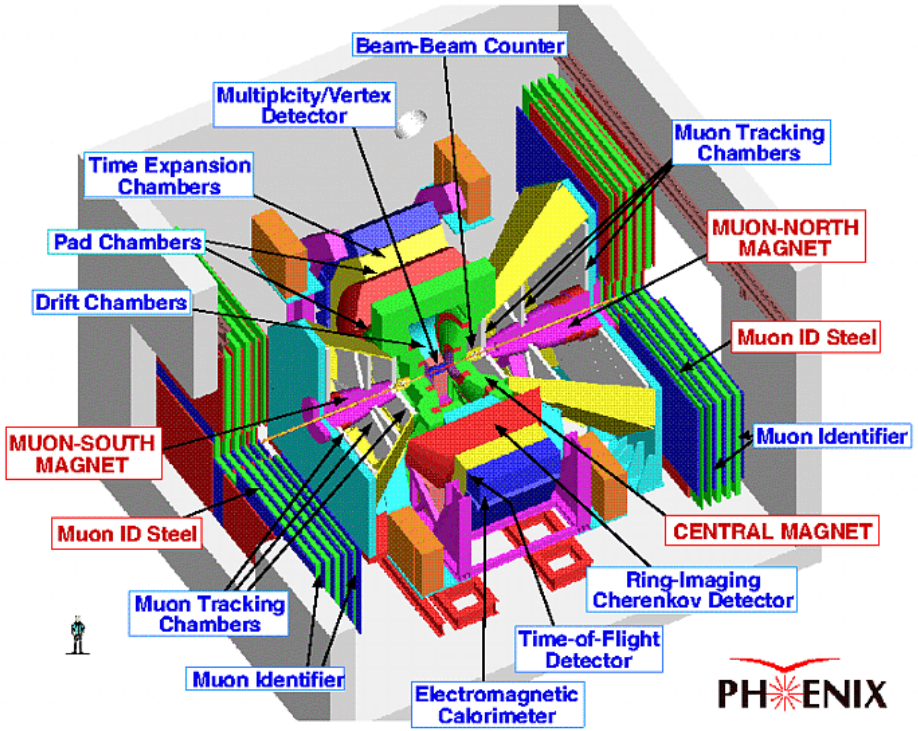
proton spin=1/2=(1/2)ΔΣ+ΔG+L_Q+L_G

ΔΣ : Quark Spin ~ 0.2-0.3

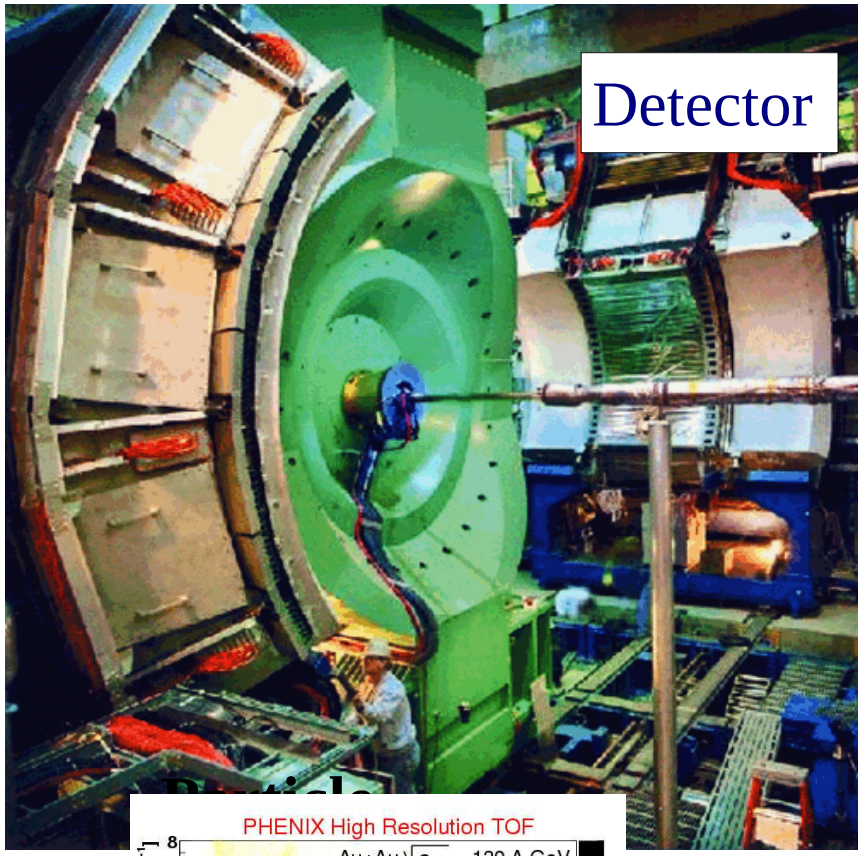


- 原子核中でのカイラル対称性の回復 (KEK-PS)

RHIC/PHENIX experiment

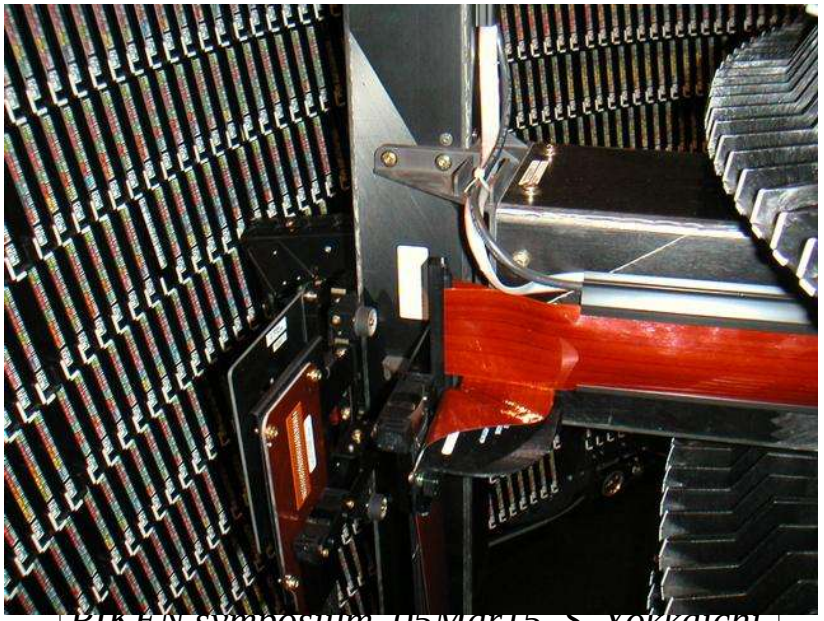


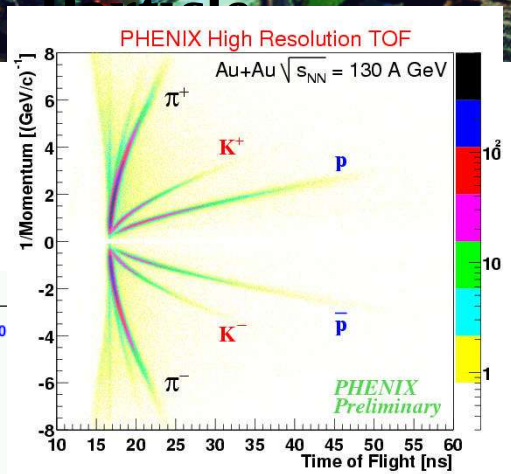
- 2000年より実験開始
  - あと５－１５年は続く
- $\sqrt{s_{NN}}$ = 200 GeV
  - 重イオン衝突(Au-Au)：クォークグルオンプラズマ(QGP)
  - 偏極陽子衝突（p-p）：核子（陽子）のスピン構造

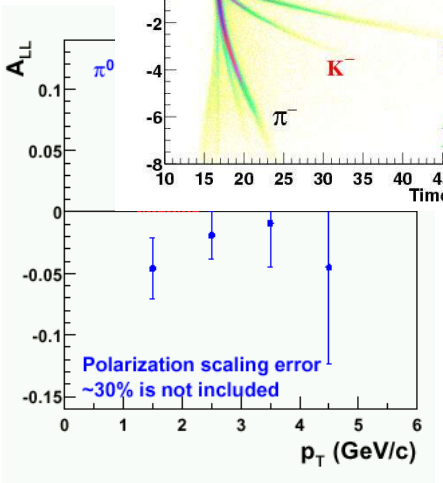


cpu server

tape robot





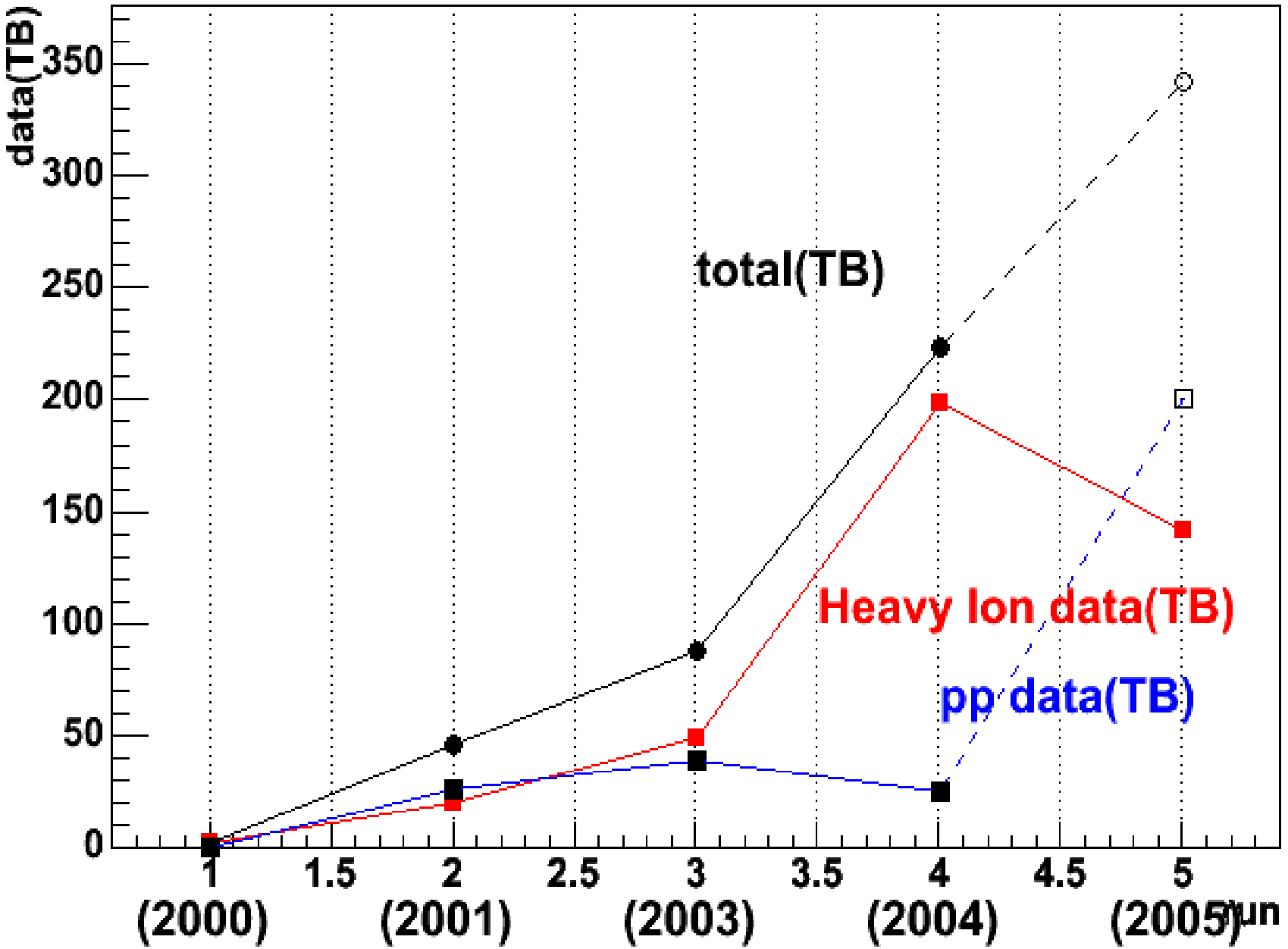


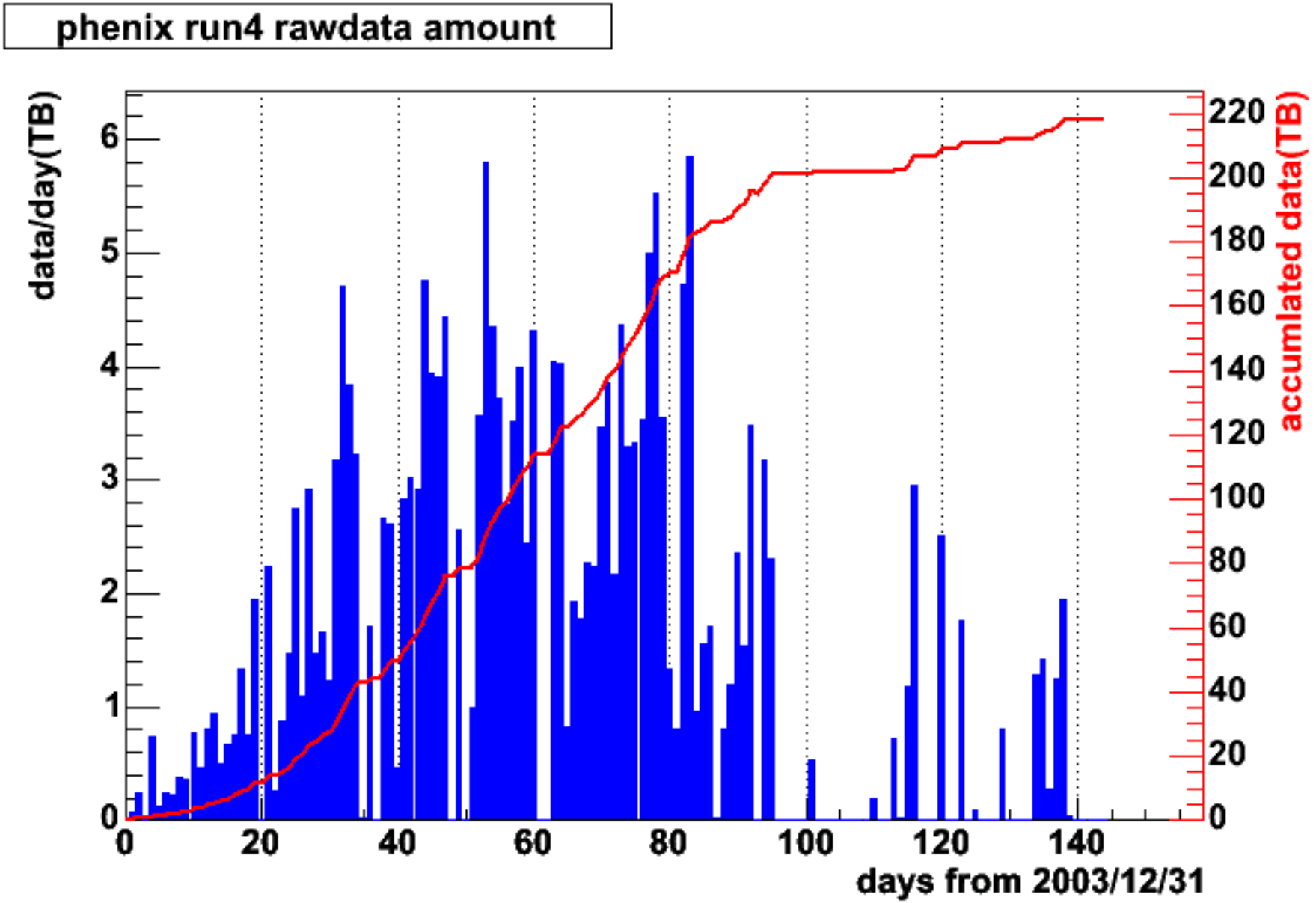
result

原子核／素粒子物理実験のデータの特徴

- Event構造
  - 相関をもたない構造のくりかえし
  - たくさんのCPUにふりわけても相互間の通信はない
  - ただし dataの読み書き／DBserverとの通信が一点に集中しがち
- Data量／file数が大量
  - 220TB/13万file (PHENIX Run4(2004))
    - 300KB/event -> 700M event
  - Job数も大量

phenix rawdata size





- 最大 5-6TB/day
  - DAQ : 100MB/sec

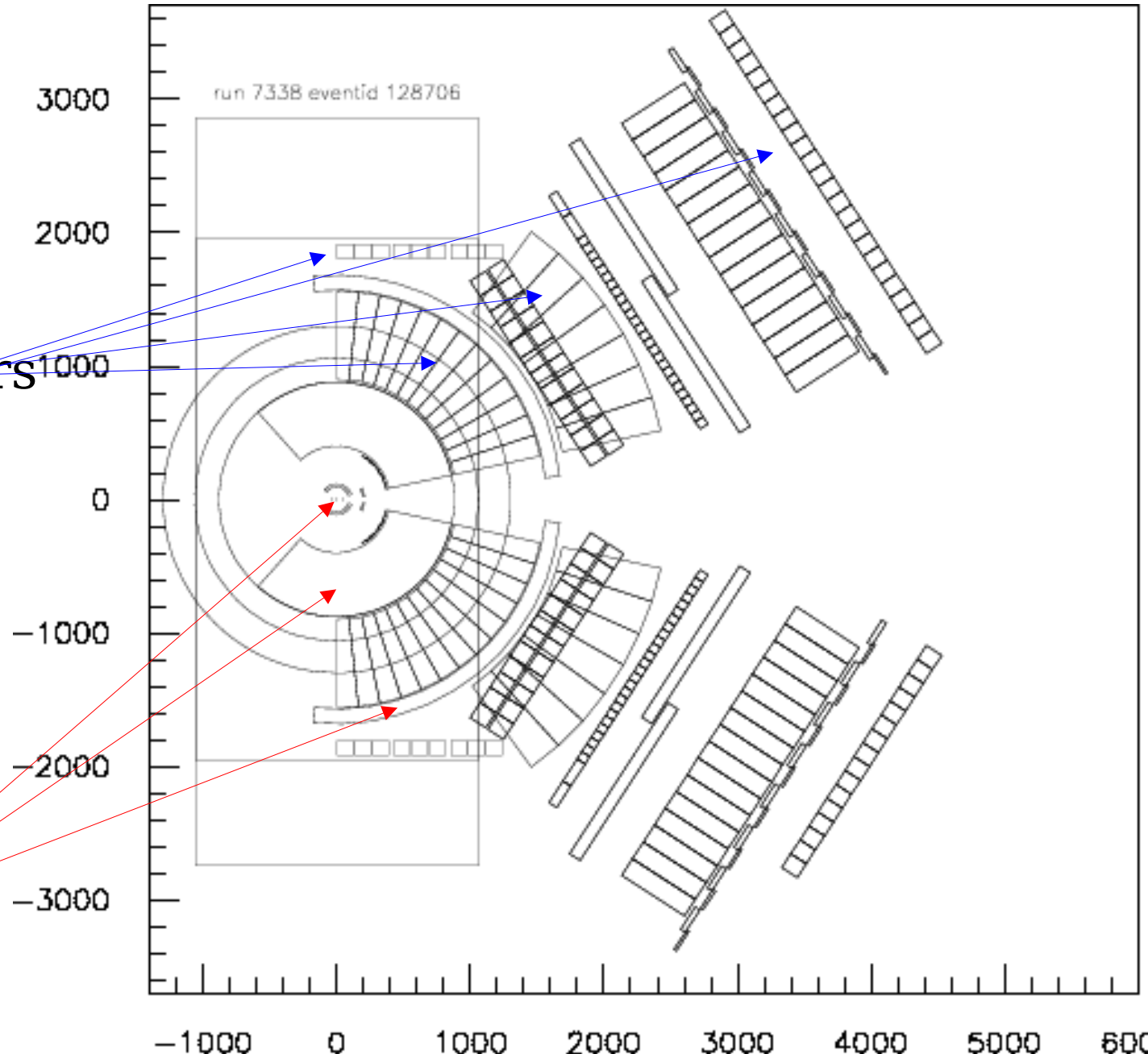
データ解析の流れの例

- E325 at KEK-PS
  - vector meson measurement
  - using electron pair

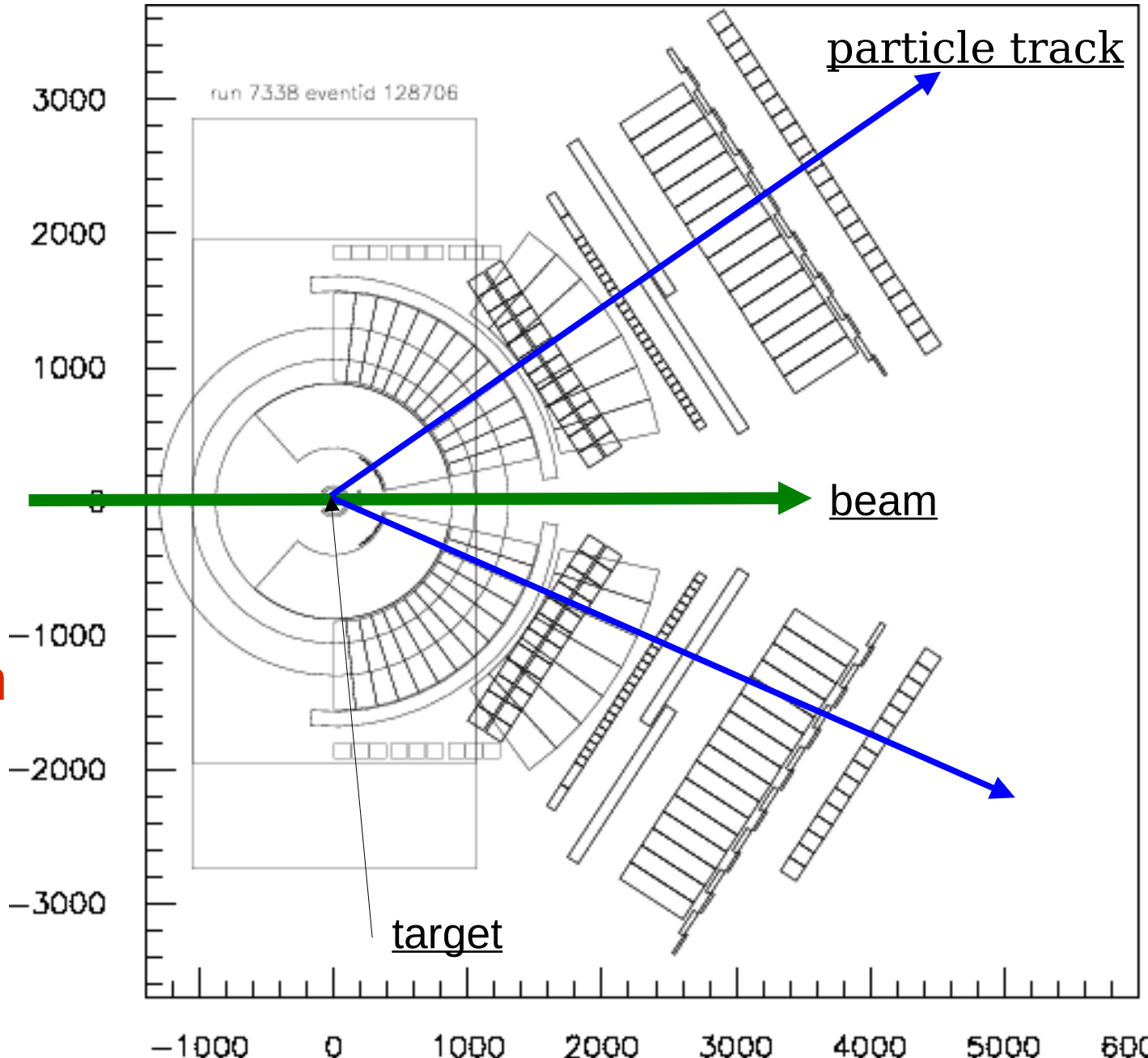
データ解析の流れ：検出器平面図

electron ID counters
(電子検出器)

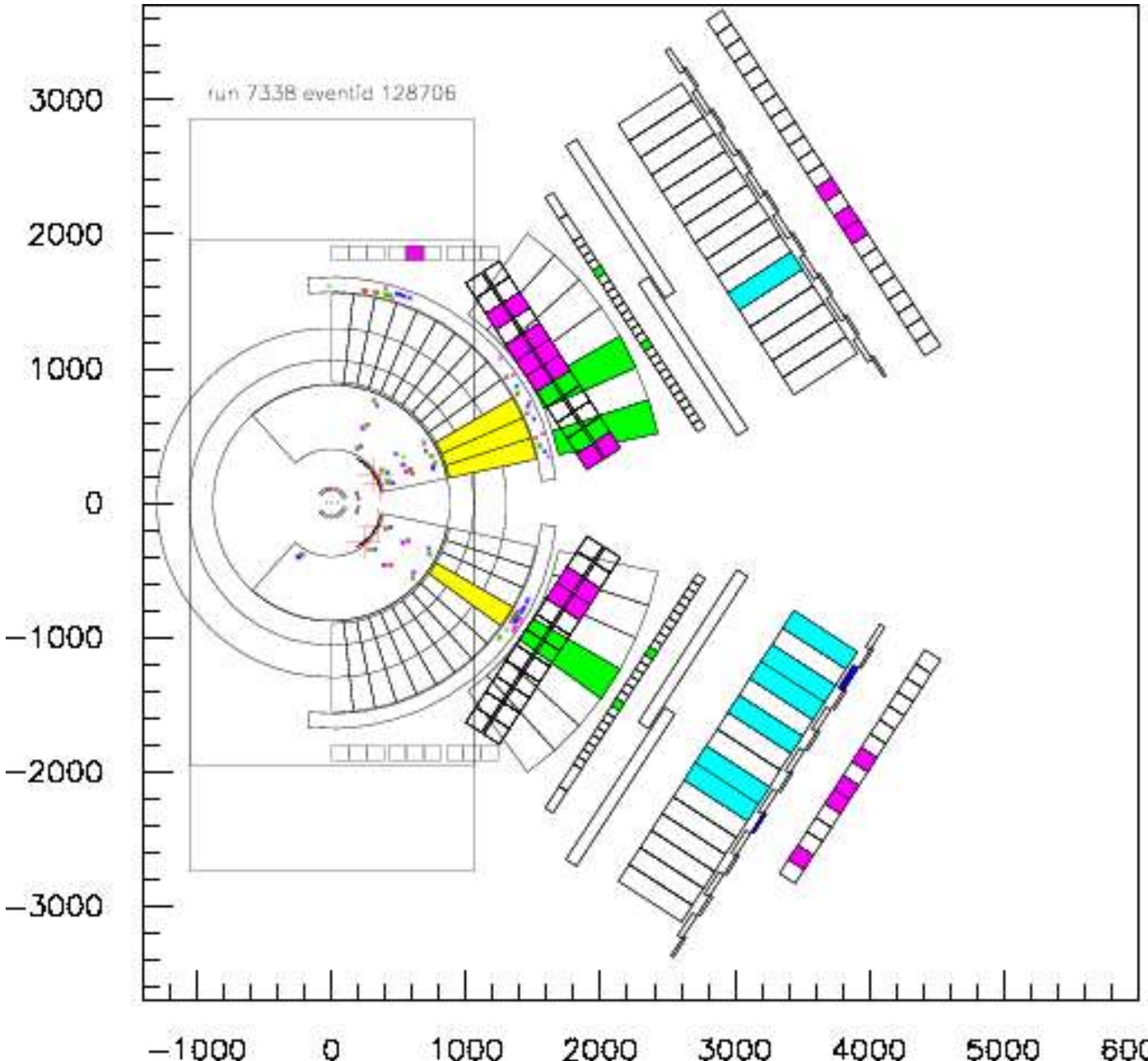
Tracker
(飛跡検出器)



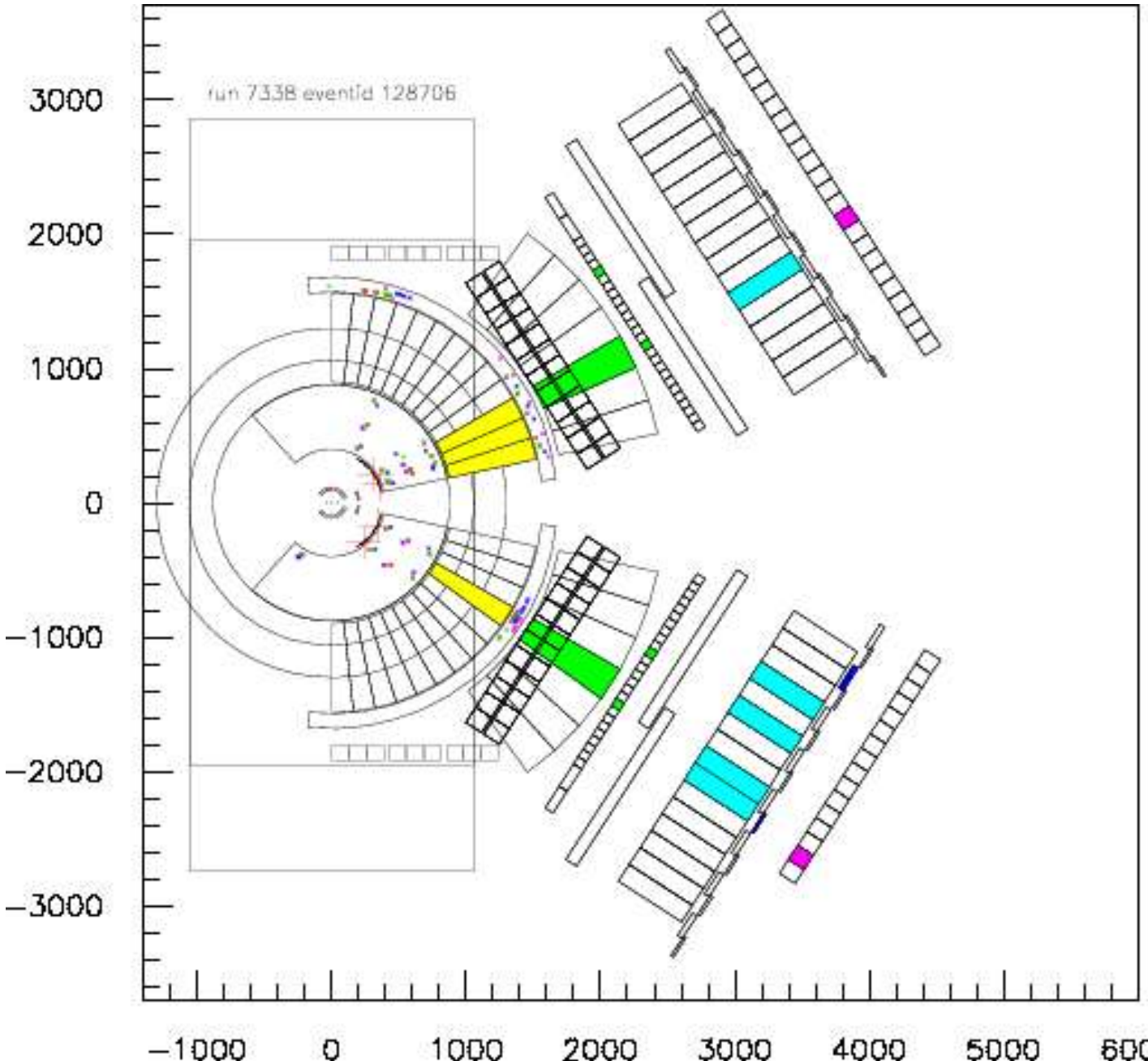
- Experiment
  - electron pair measurement
  - 3.5K channel
- raw data
  - from the digitize circuit
  - 2KB/event
  - 1.3TB/yr(month
  - need decoding &channel mapping



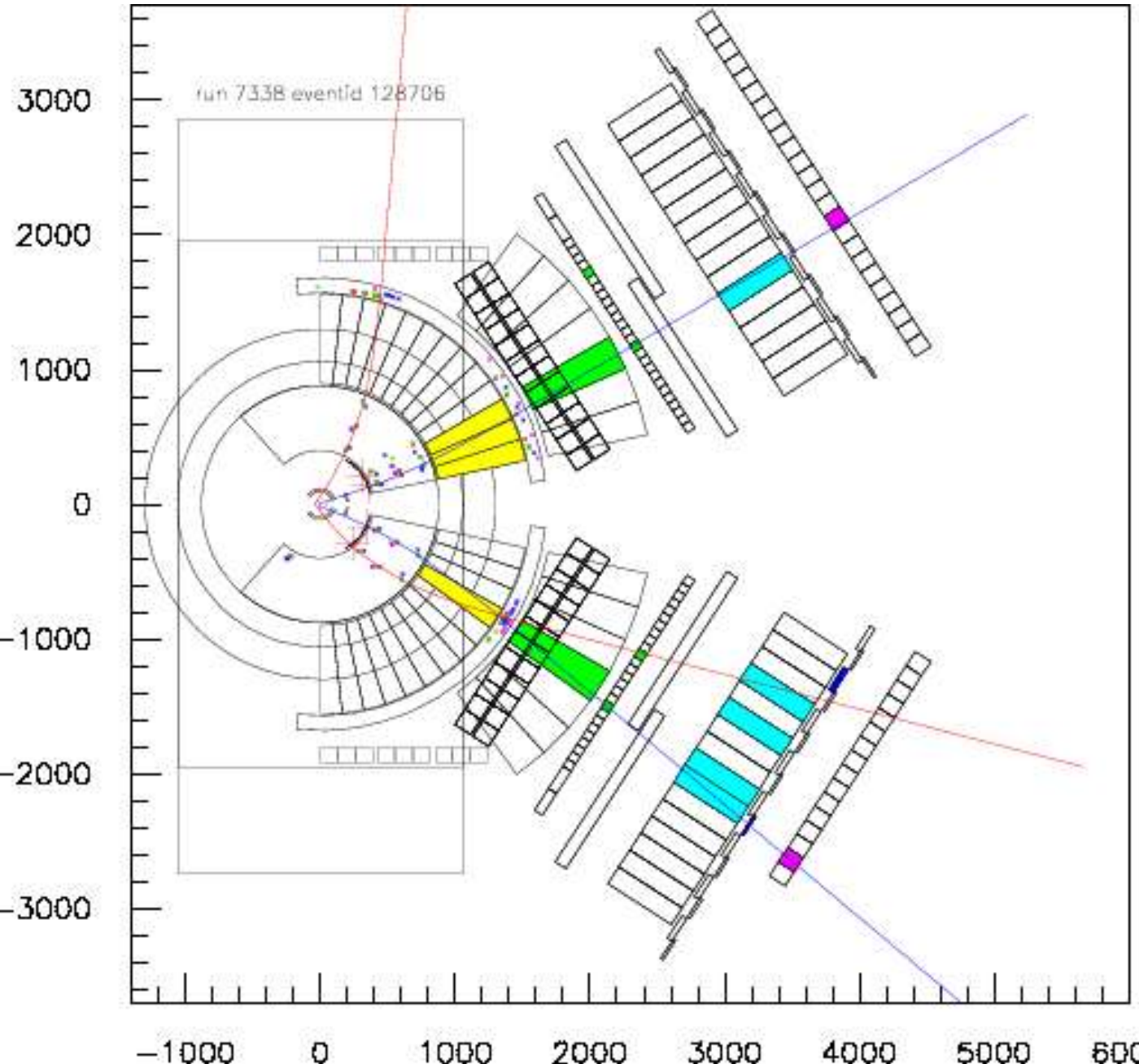
- raw data
  - channel mapping
  - detector hit information



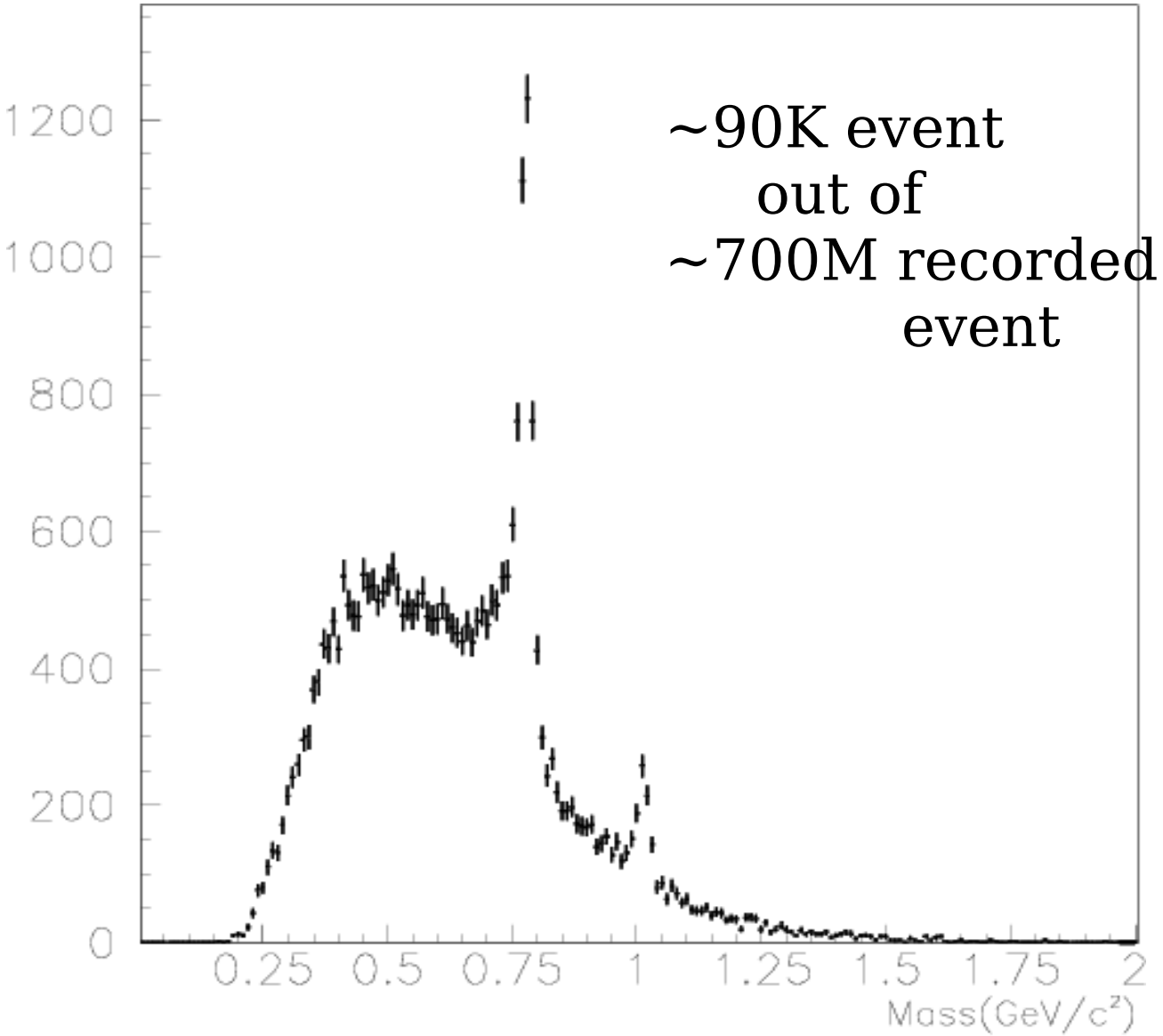
- raw data
- calibration
  - tracker hit position
  - pulse hight-energy relation in EMCal
  - etc.



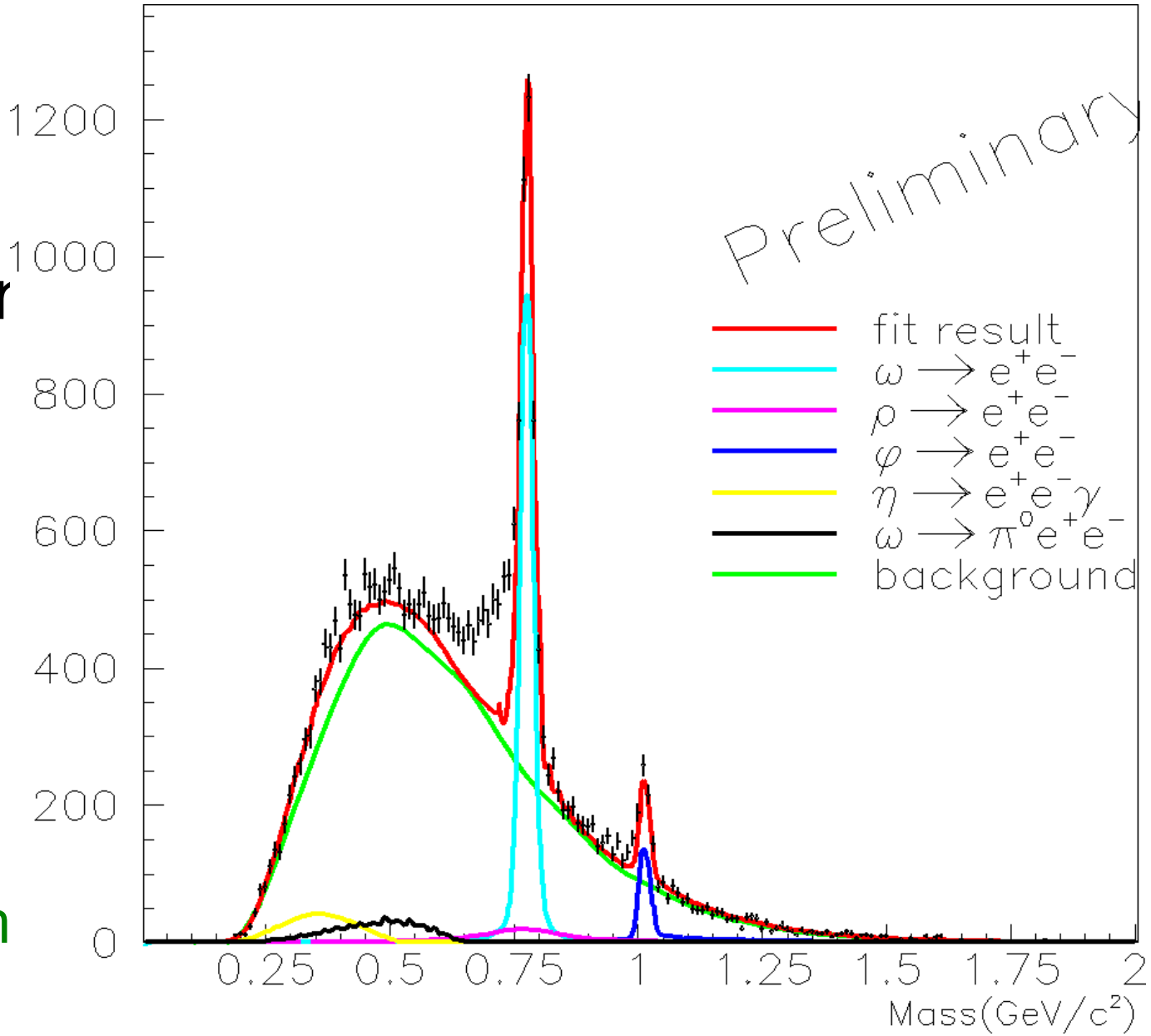
- raw data
- calibration
- track/event reconstructio
  - momentum
  - Particle ID
    - blue:electro
    - red : other
  - invariant mass of eletron pair



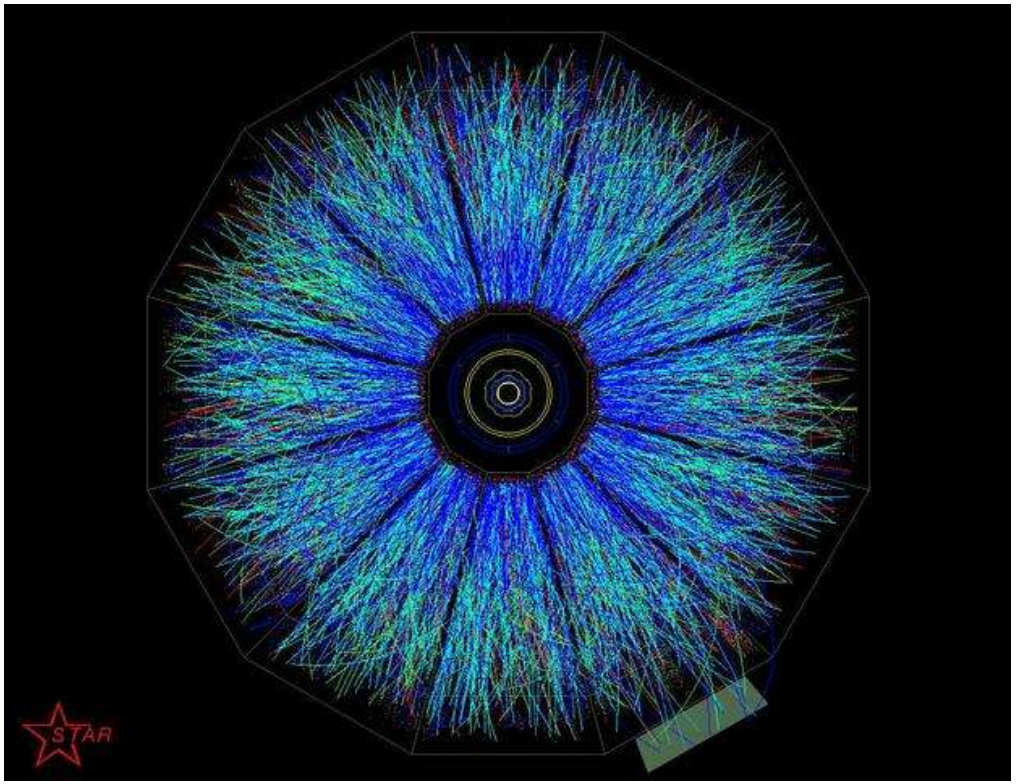
- raw data
- calibration
- event reconstruction
- mass spectra
  - ee invariant mass
  - accumlate events



- raw data
- calibration
- event reconstruction
- mass spectra
  - fitting
  - number
  - efficiency correction
- Physics:
  - cross section



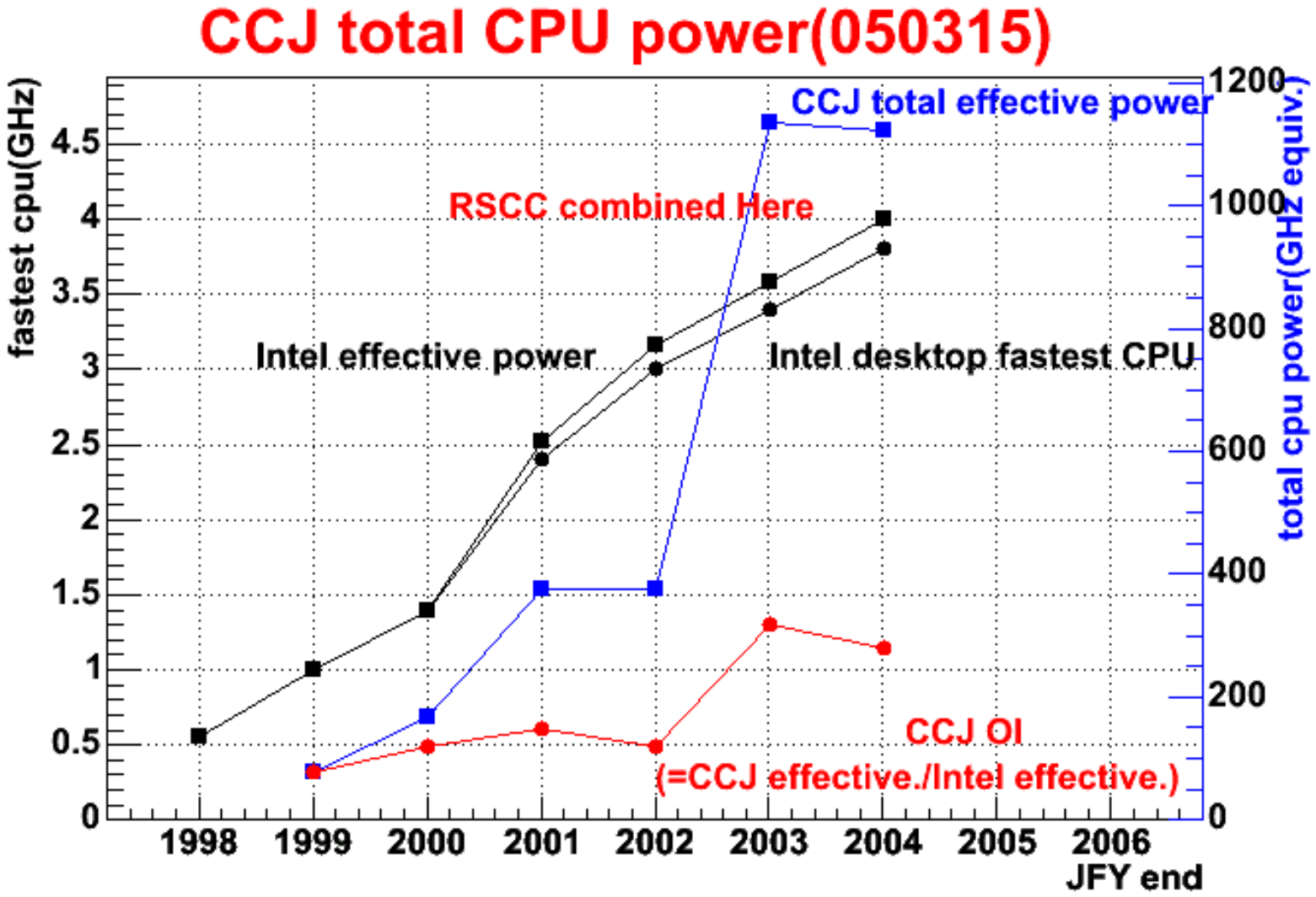
event reconstruction in RHIC exp.



- PHENIX and STAR exp. reconstructed event
  - 300K channel, 300KB/event, 200TB/yr
  - (3.5K ch, 2KB/ev, 1.3TB/yr at KEK-PS E325)
- ~ 5 sec/event @ PentiumIII 1GHz
  - ~4500 CPU days for 77M event(~18TB) (Run2 AuAu v3 : 2001年データ実績)

RSCC/CCJの連携(1)

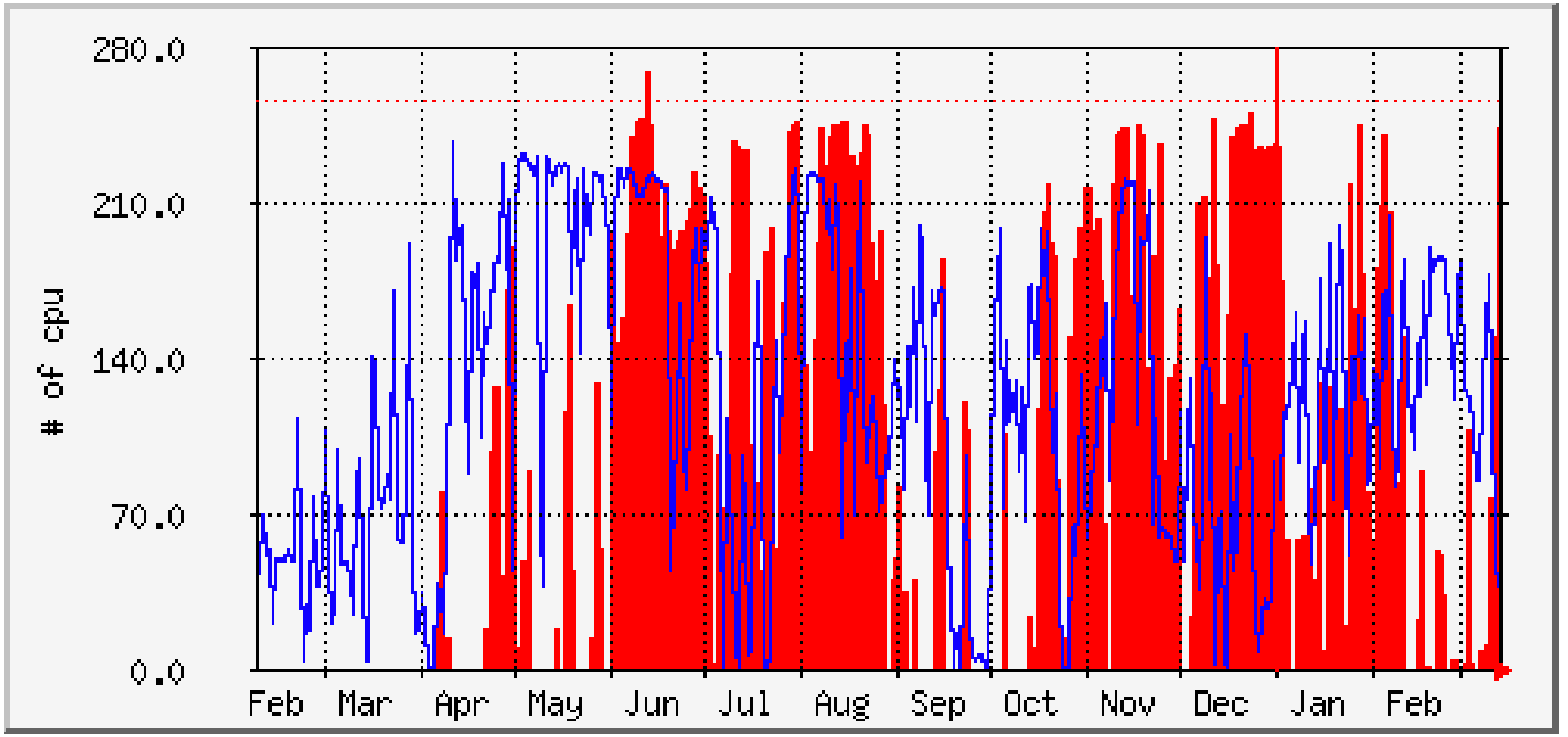
- CCJ：332 CPU ＋ RSCC pc2c 256 CPU
  - Power は ３倍に。



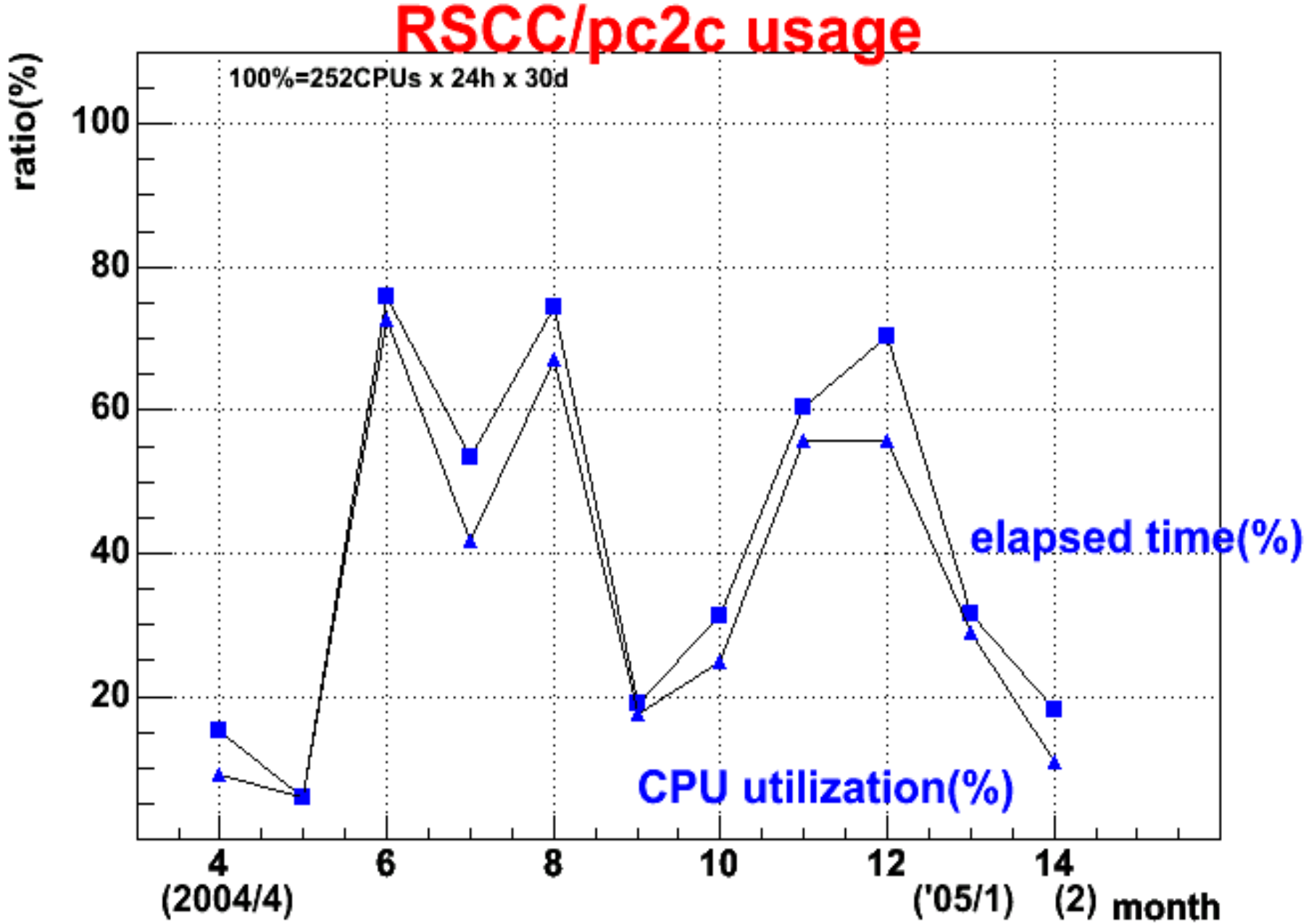
RSCC/CCJの連携(2)

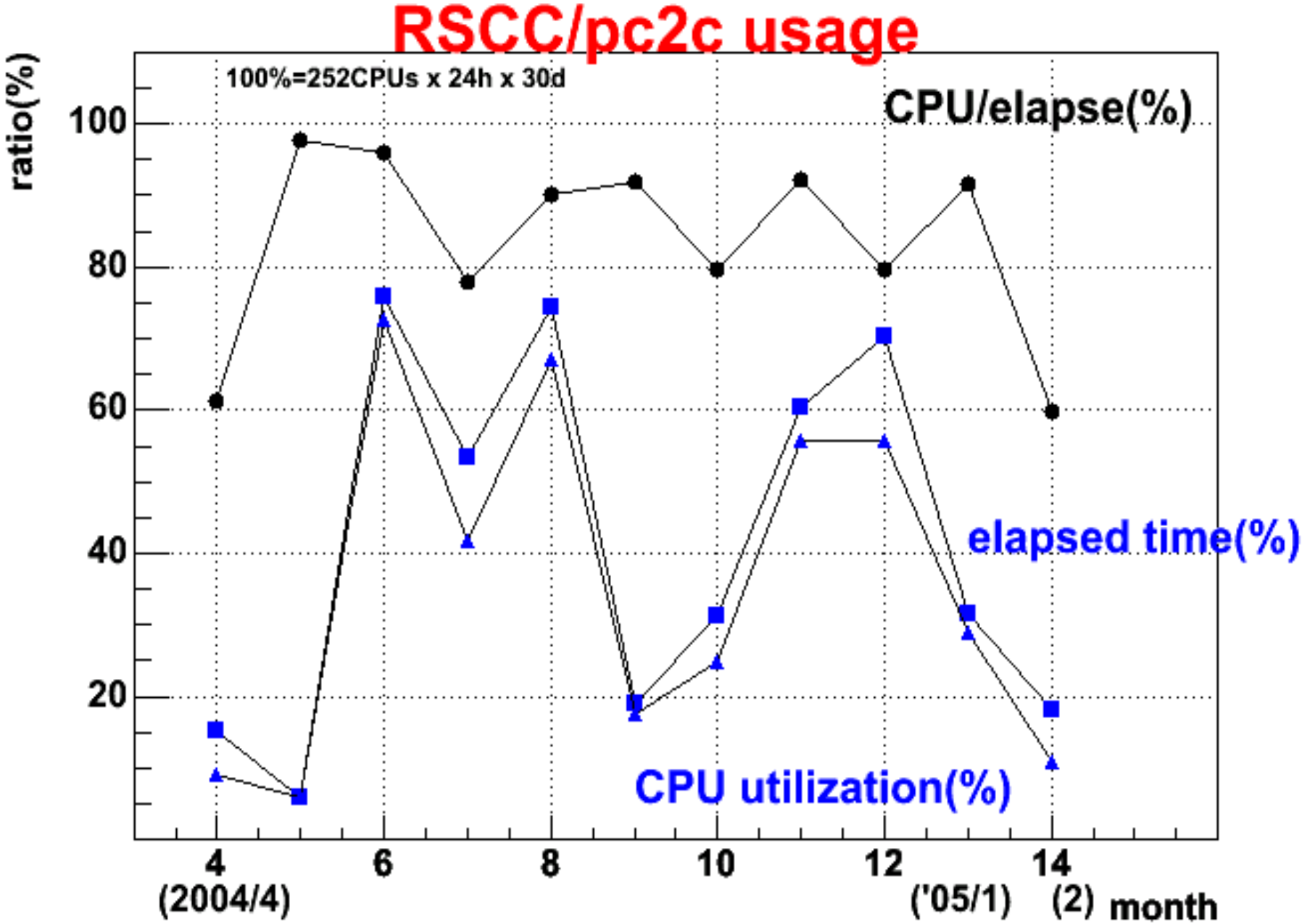
- PHENIX software 環境の共有
  - OS (Linux RH80->SL3)
  - NFS : ライブラリ 70GB
  - DB server (PostgreSQL)
  - 30TB raid
  - HPSS
- Batch queueing system
  - NQS -> LSF
  - Jobの特徴
    - ２~３時間 x 1000 ~ 10000 job
    - User level Analysis / MC

RSCC/CCJ usage



- CPU
  - RSCC/CCJ
- 6月くらいから本格使用


RSCC/pc2c usage
100%=252CPUs x 24h x 30d
ratio(%)
100
80
60
40
20
4
(2004/4)
6
8
10
12
14
('05/1)
(2)
month
elapsed time(%)
CPU utilization(%)


RSCC/pc2c usage
100%=252CPUs x 24h x 30d
ratio(%)
100
80
60
40
20
CPU/elapse(%)
elapsed time(%)
CPU utilization(%)
4
(2004/4)
6
8
10
12
14
('05/1)
(2)
month

これからの一年

- 4~6月 : ~200TB pp data
  - 5TB/day (= 60MB/sec) x 40 days
  - Networkで運ぶ予定
    - 1Gbps : RIKEN-BNL
    - 1TB/day で10TBくらいの実績はある(bbftp)
    - テスト中 (GridFTP, 7-800Mbps=~100MB/sec)
- 200TBの処理 (event再構成)
  - CCJ (200CPU + I/O 50MB/sec) : 72days
  - RSCC + I/O 100MB/sec : 30days
- CCJ
  - サーバ (NFS,etc) 系の更新
  - data転送

これから 5年後

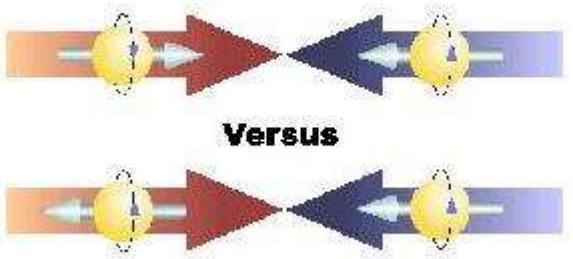
- データ量は増える一方
  - 1 PB /year from PHENIX ?
    - LHC(CERN) : 1PB/year at 2007
- networkデータ転送
  - 5-10Gbps 1GB/sec = 86TB/day
- I/Oの分散
  - ただ読むだけで 1PB/70days = 160 MB/sec
    - これだけなら現行HPSSでいける。しかし、
    - 時間をCPUと 1:1でわけあうと 320 MB/sec
    - 10:1 1.6 GB/sec
  - (CPU+disk) farm using GRID

Back up

Spin Crisis...核子スピンの起源の謎

proton spin=1/2=(1/2)ΔΣ+ΔG+L_Q+L_G

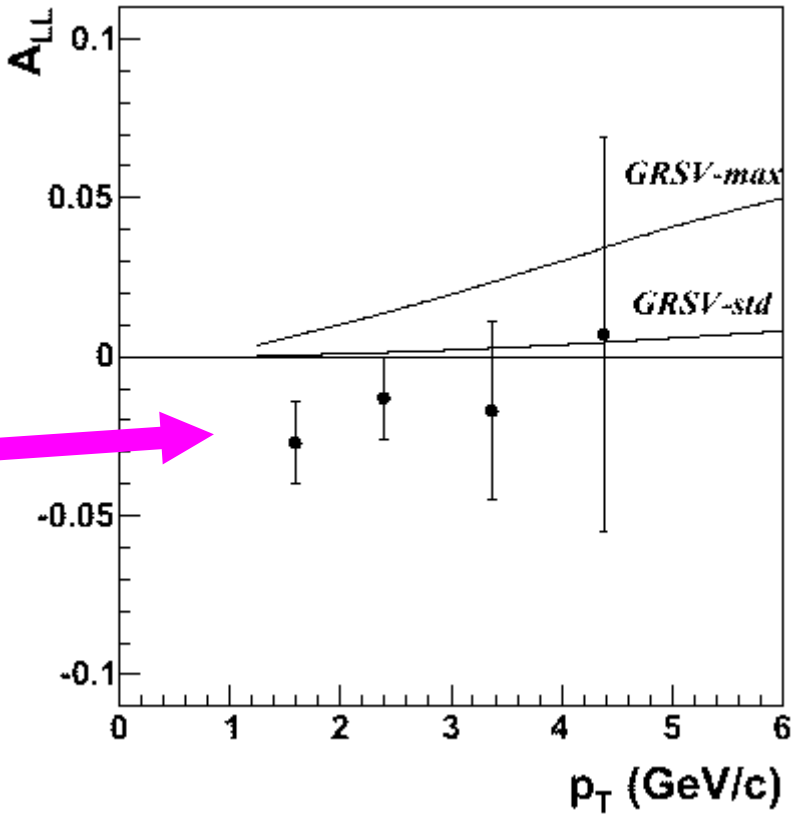
ΔΣ : Quark Spin ~ 0.2-0.3



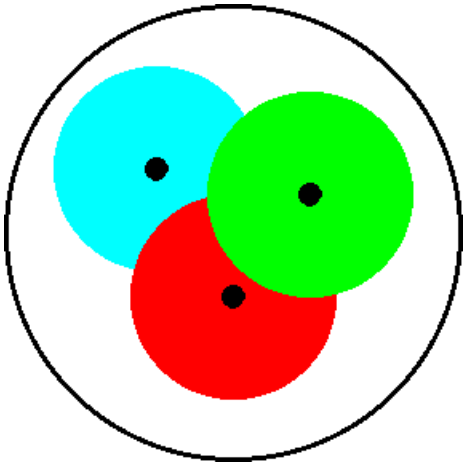
only pol. proton collider can measure the gluon spin ΔG

- result of pi0 asymmetry (from 2003 data)
  - large negative value is not explained usual pQCD theory
  - contribution of gluon is very small ??
  - data accumulation should be done in 2005- run !

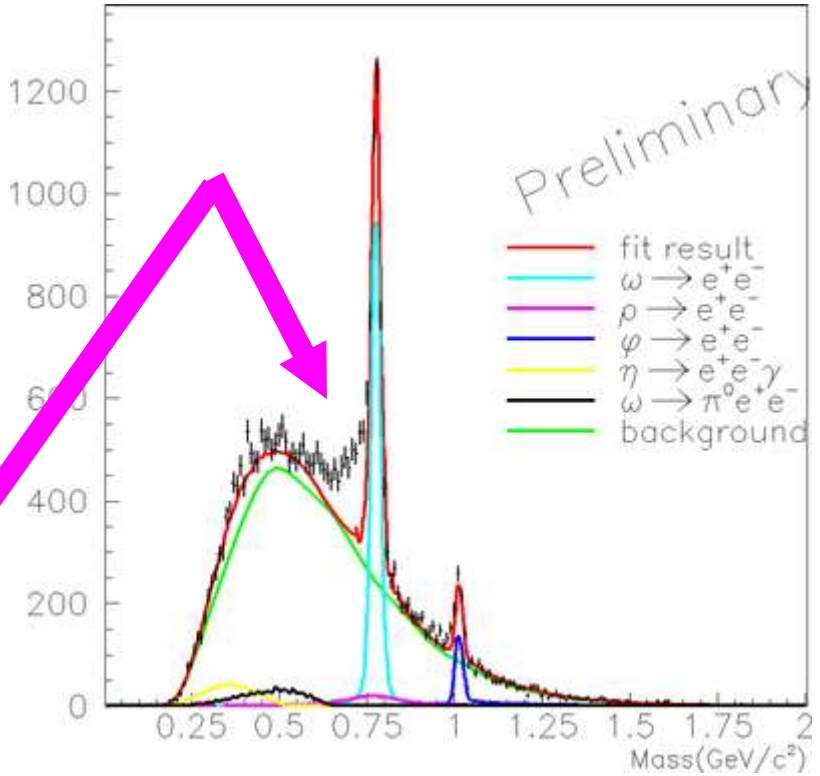
PRL93(04)202002



Chiral symmetry restoration in nuclear matter

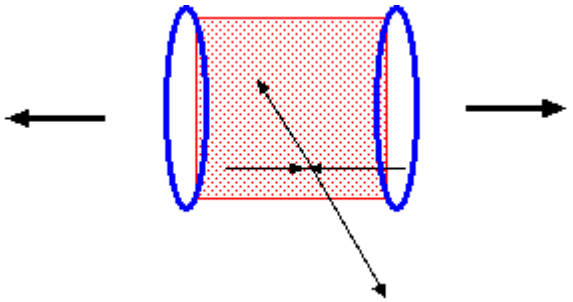


- vacuum structure in QCD
  - chiral symmetry is dynamically broken
    - origin of (constituent) quark and hadron mass
  - meson modification in hot and/or dense matter
- ee decay of vector mesons (ρ,ω,ϕ) in nuclei at 12GeV p+C/Cu reaction at KEK
  - ρ/ω meson modification is observed

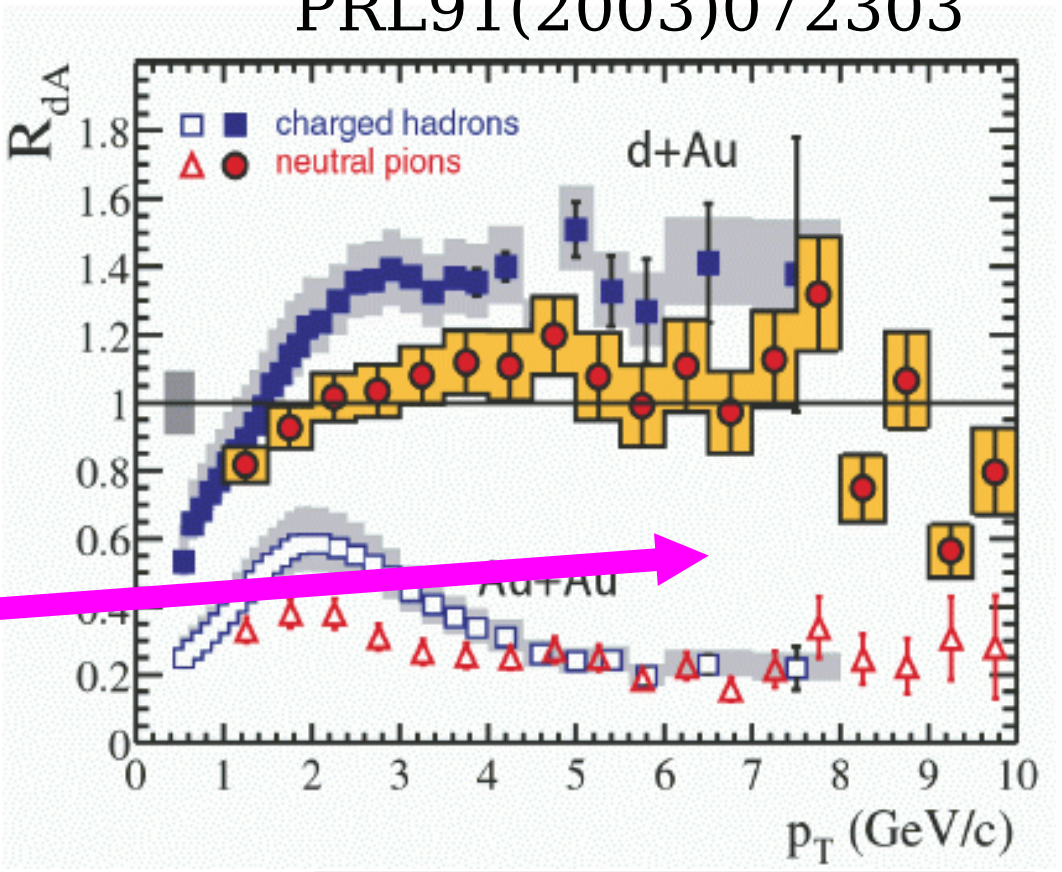


Suppression of High Pt particle
-Jet quenching in Quark Gluon Plasma-

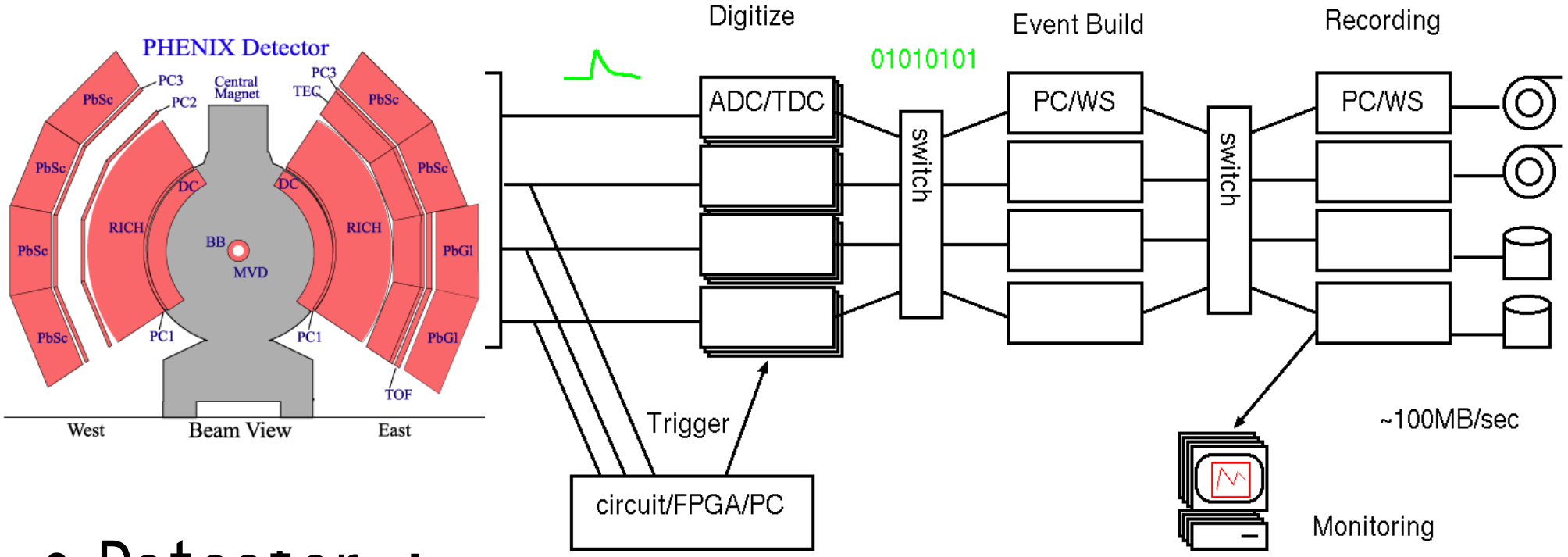
- Prediction : Jet is suppressed in QGP
  - parton energy loss in matter
- pion production cross section (normarized by pp data) in $\sqrt{s_{NN}}$ =200GeV
  - Au+Au : suppressed in high Pt, in comparison with d+Au
  - centrality dependence also supports QGP picture



PRL91(2003)072303

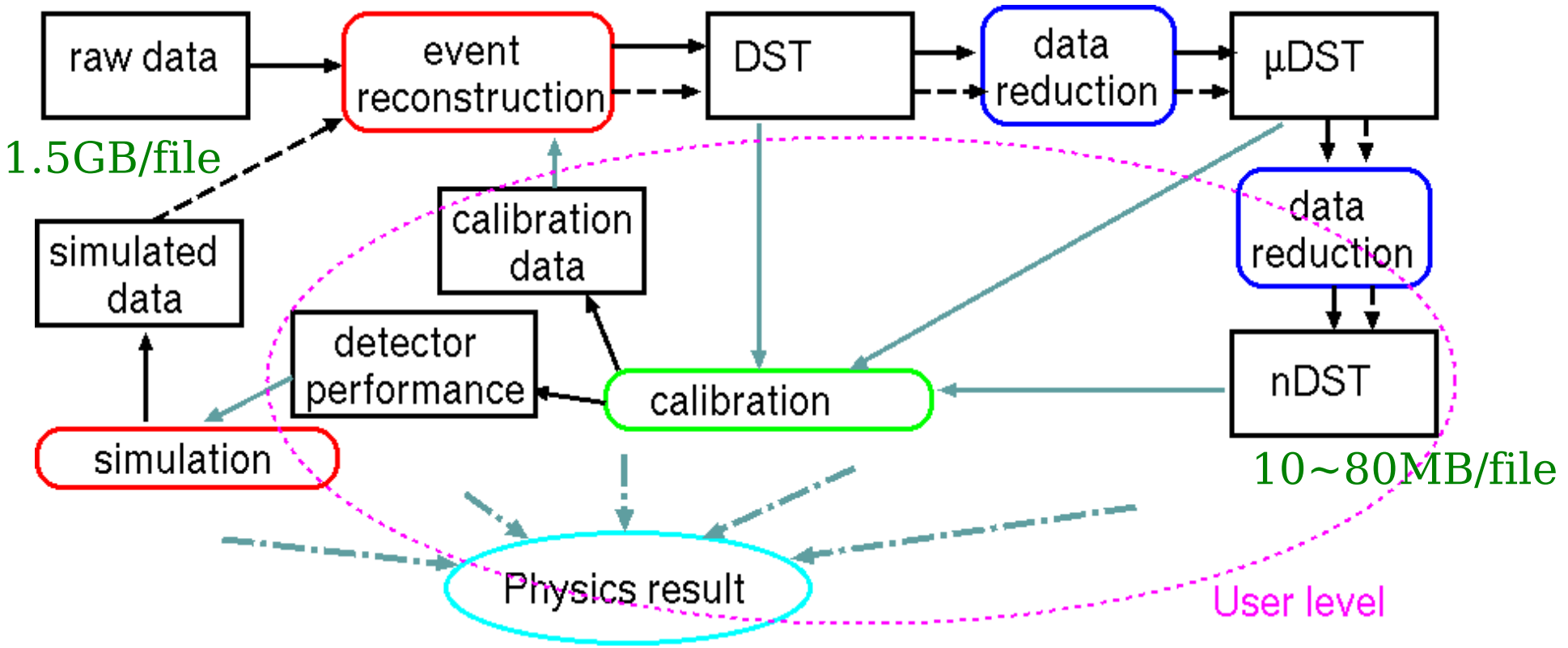


Data flow in accel. Exp.-online-



- Detector :
  - Energy, position, timing etc. of particles from reaction point
  - 300K channel @ PHENIX
- Event : set of data for one reaction
- data size : ~300KB/event, ~220TB/year (Run-4 (2004))

offline analysis – PC cluster's role–



- CPU：event reconstruction/simulation, I/O：data reduction
  - Particle track reconstruction under magnetic field :~90%
- event by event structure : no parallel calc. needed
- collaboration level work/ user level work