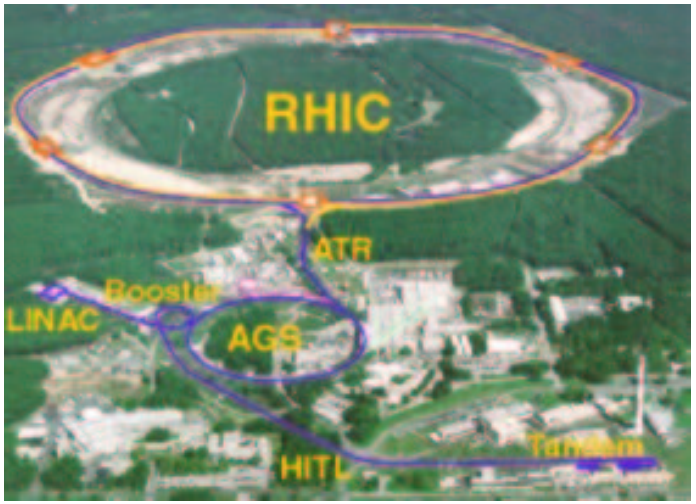
Data processing in Nuclear/Particle Physics using Accelerator and PC cluster

Satoshi Yokkaichi
Radiation Lab., RIKEN

- Introduction to data processing in this field
- RIKEN-CCJ at Wako campus
- Physics result from CCJ and prospect

Radiation Lab, RHIC/PHENIX and RIKEN-CCJ

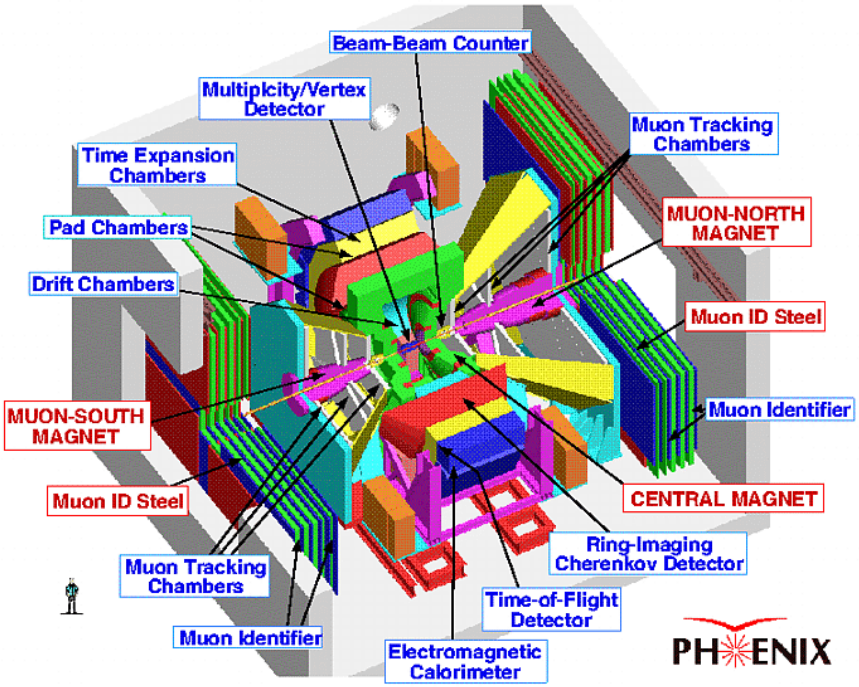
- Hadronic Physics
  - proton, neutron, pi, K etc.
  - consist of 3(2) quarks:containment
  - Why and how ?
- RHIC/PHENIX at BNL(in USA)
- CCJ (computing center in Japan) at Wako



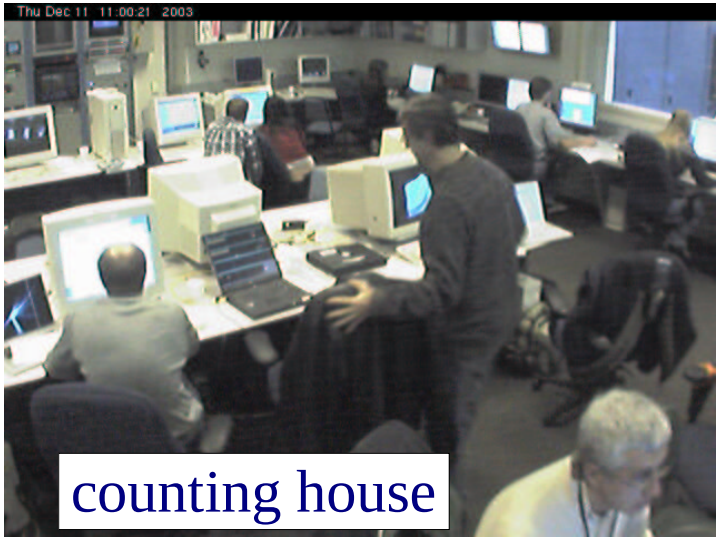
tape robot
(600TB)

CPU farm
(332 CPU)

RHIC/PHENIX experiment

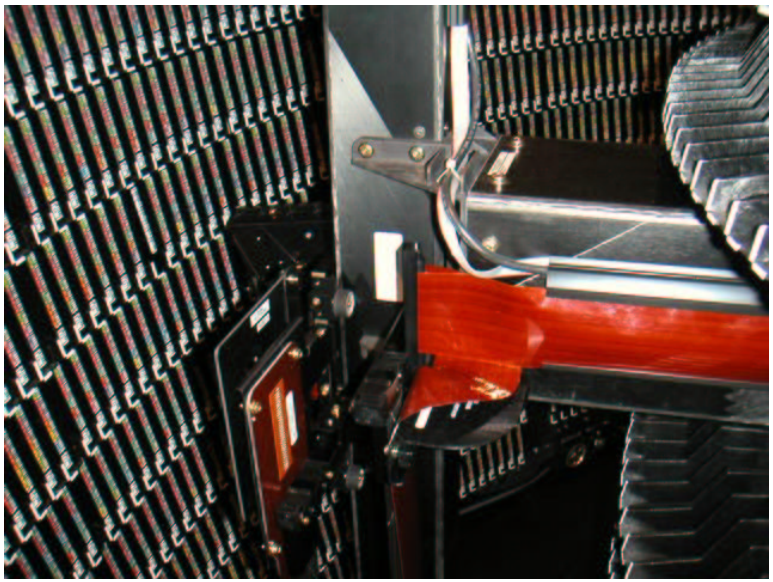


- data taking started in 2000
- $\sqrt{s_{NN}}$ = 200 GeV
  - Heavy Ion collision (Au-Au) Quark Gluon Plasma
  - polarized proton collision Nucleon spin structure

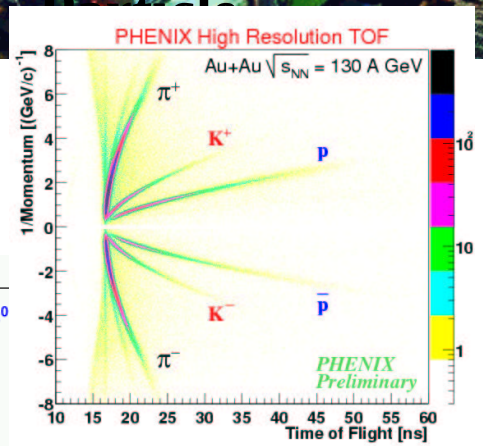


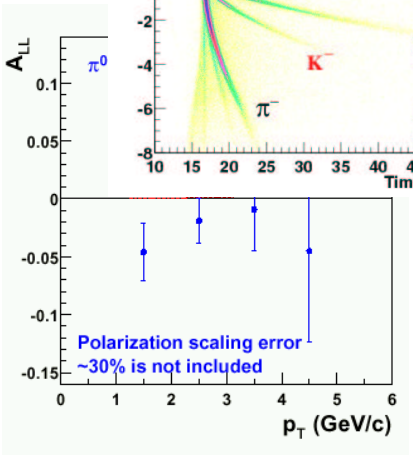
cpu server

tape robot



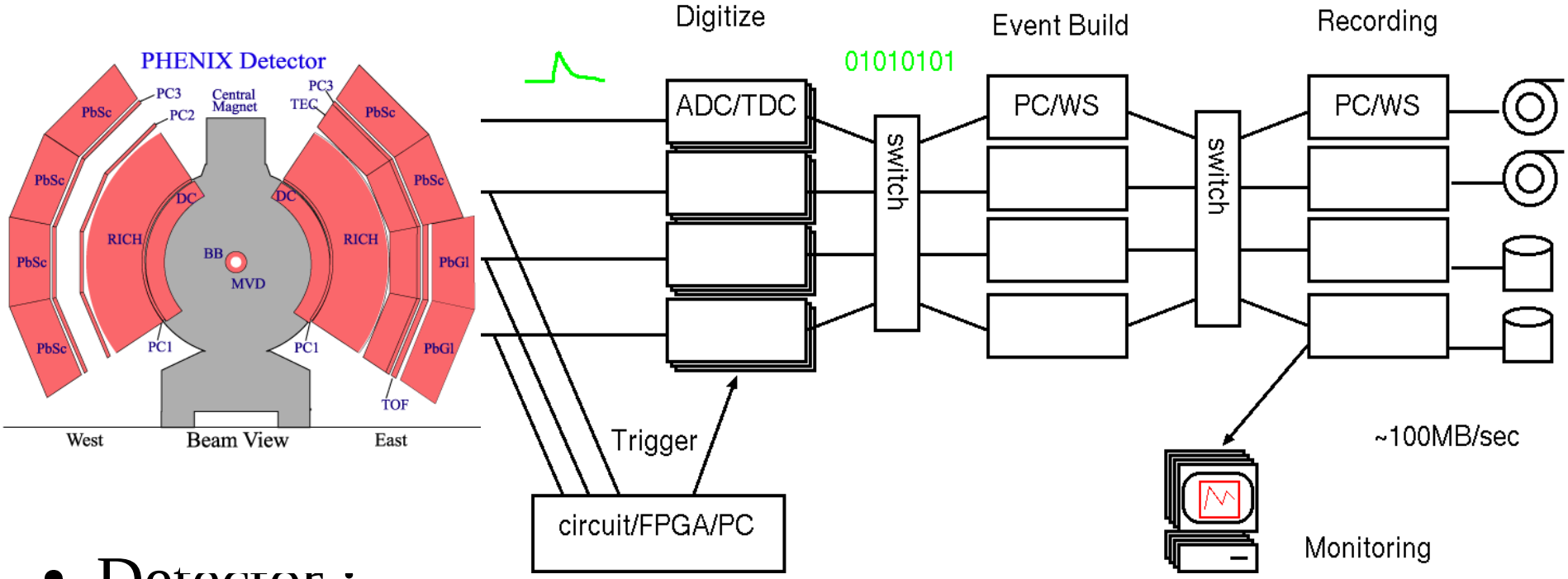
New horizon of HPC, 03Dec15, S. Yokkaichi





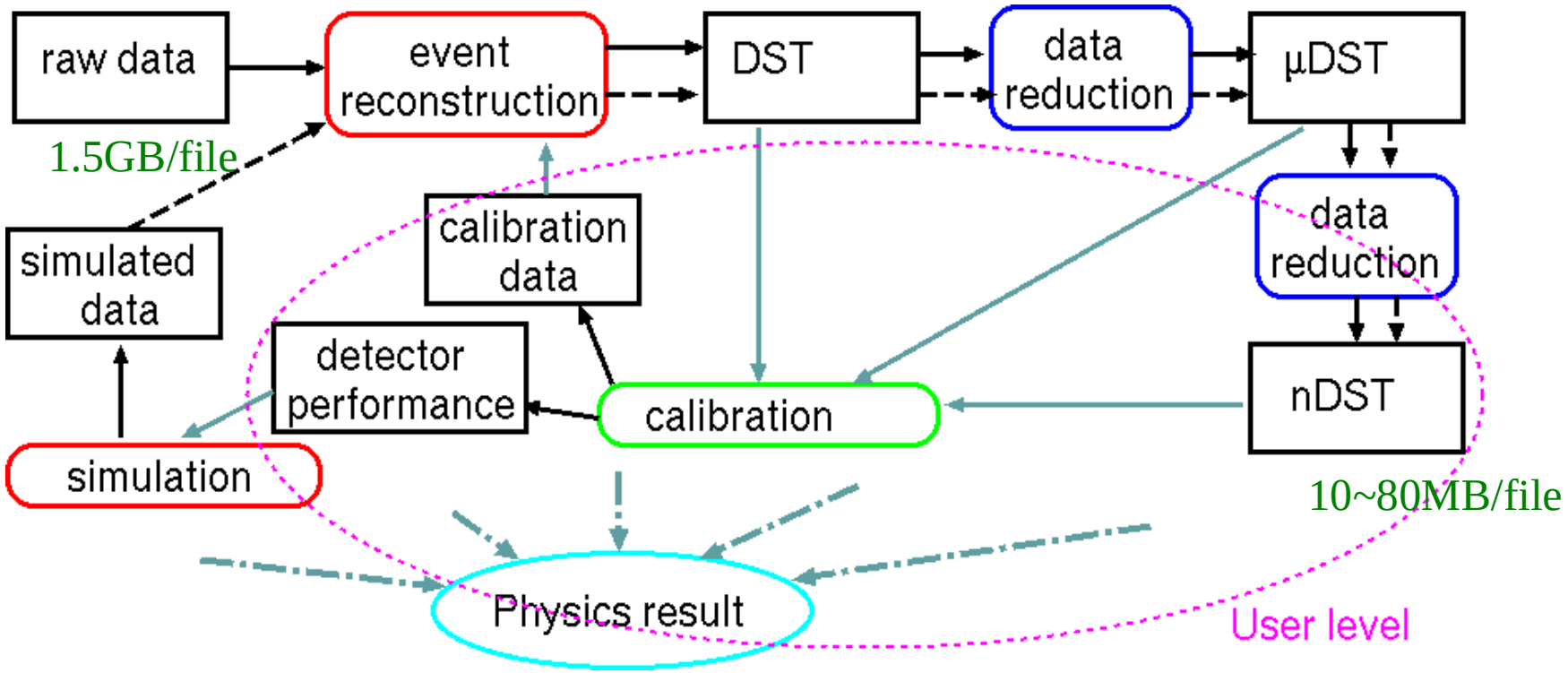
result

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, ~88TB/year (Run-3(2003))

offline analysis - PC cluster's role-



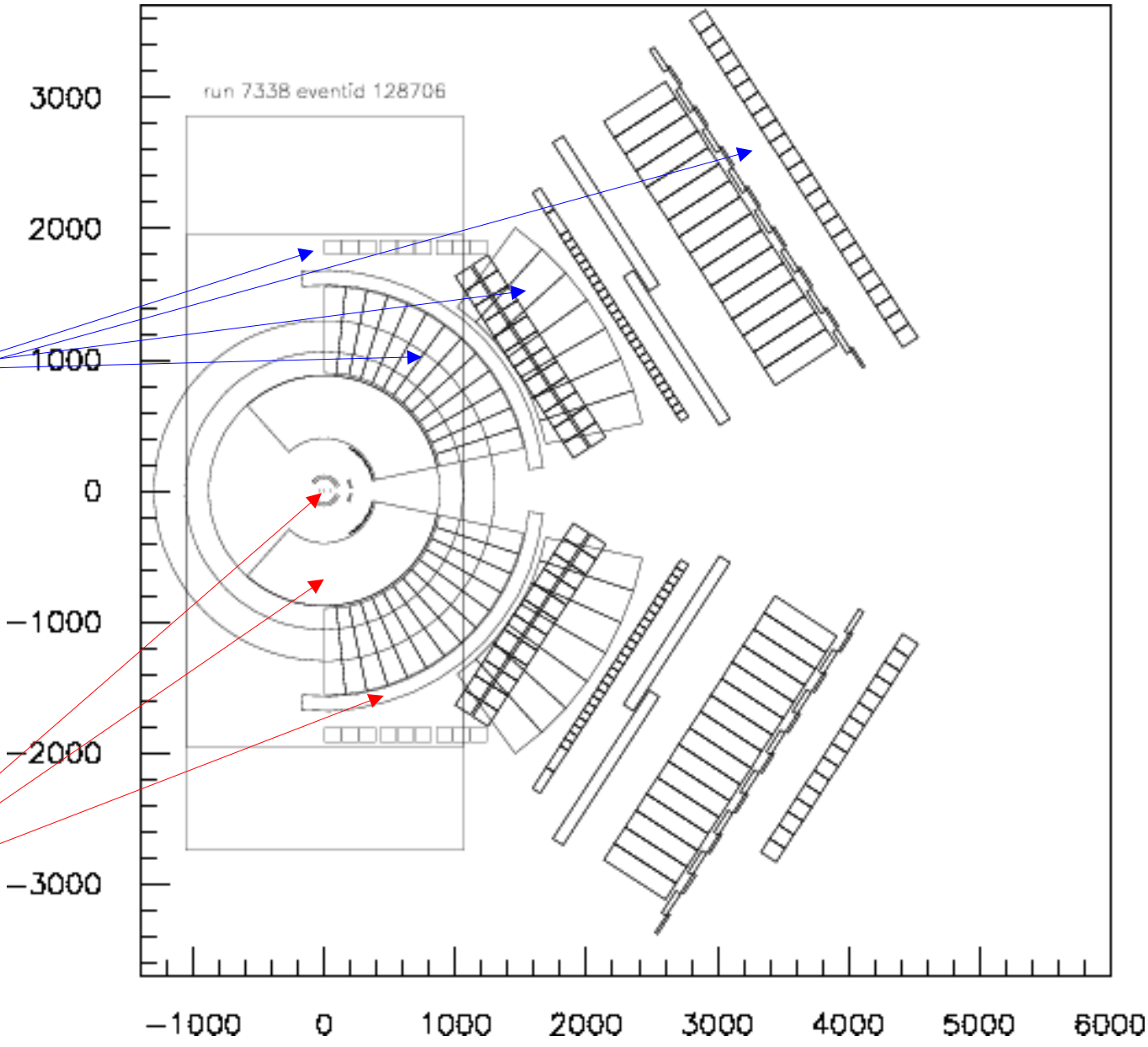
- CPU：event reconstruction/simulation, I/O：data reduction
  - Particle rack reconstruction under magnetic field :~90%
- event by event structure
  - no parallel calc. needed
- collaboration level work/ user level work

Example of data analysis flow

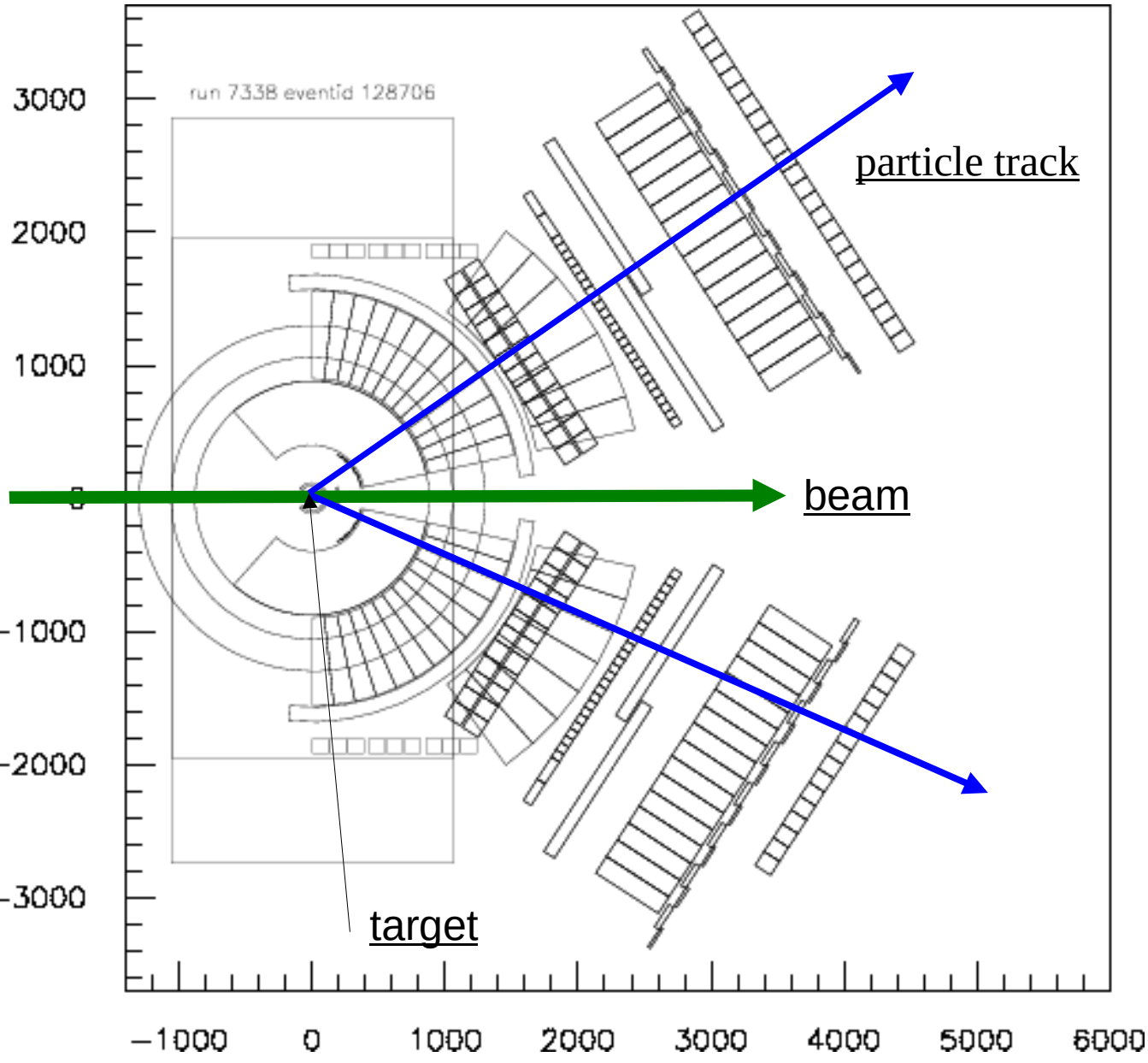
- Experiment at KEK-PS
  - vector meson measurement
  - using electron pair

electron ID counters

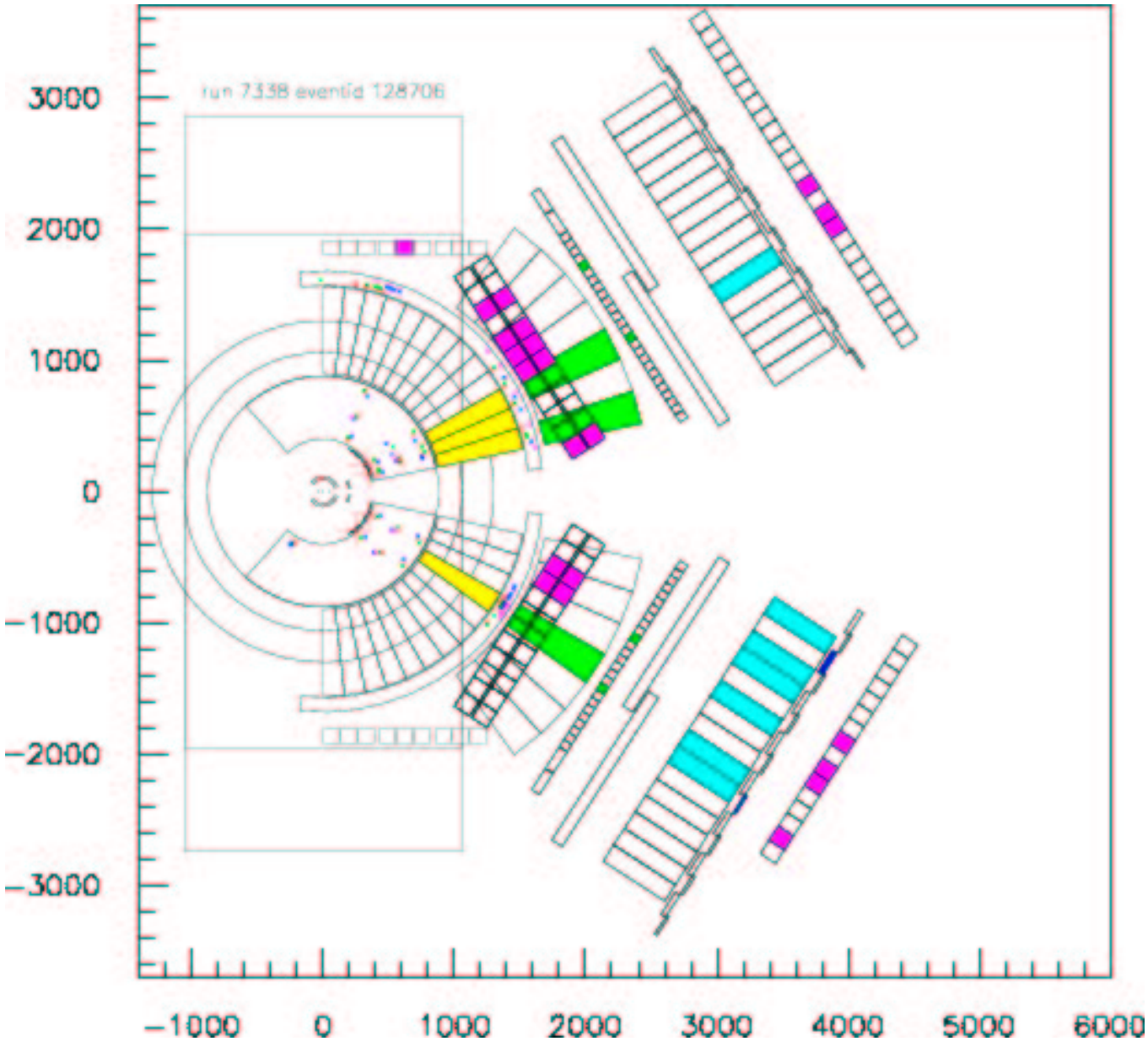
tracker



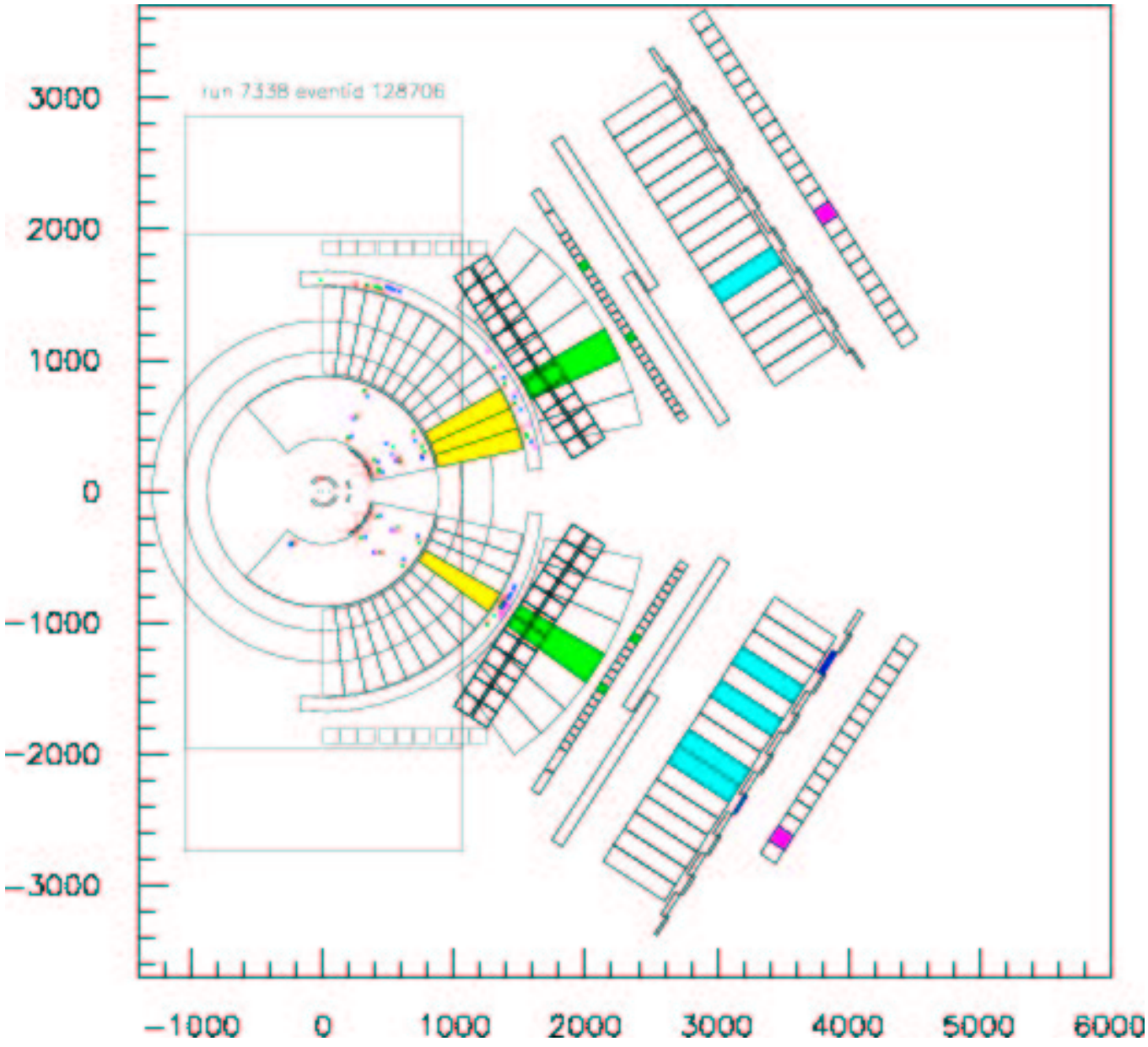
- Experiment
  - electron pair measurement
  - 5K channel
- raw data
  - from the digitize circuit
  - 2KB/event
  - 1.3TB/yr(month)
  - need decoding &channel mappin



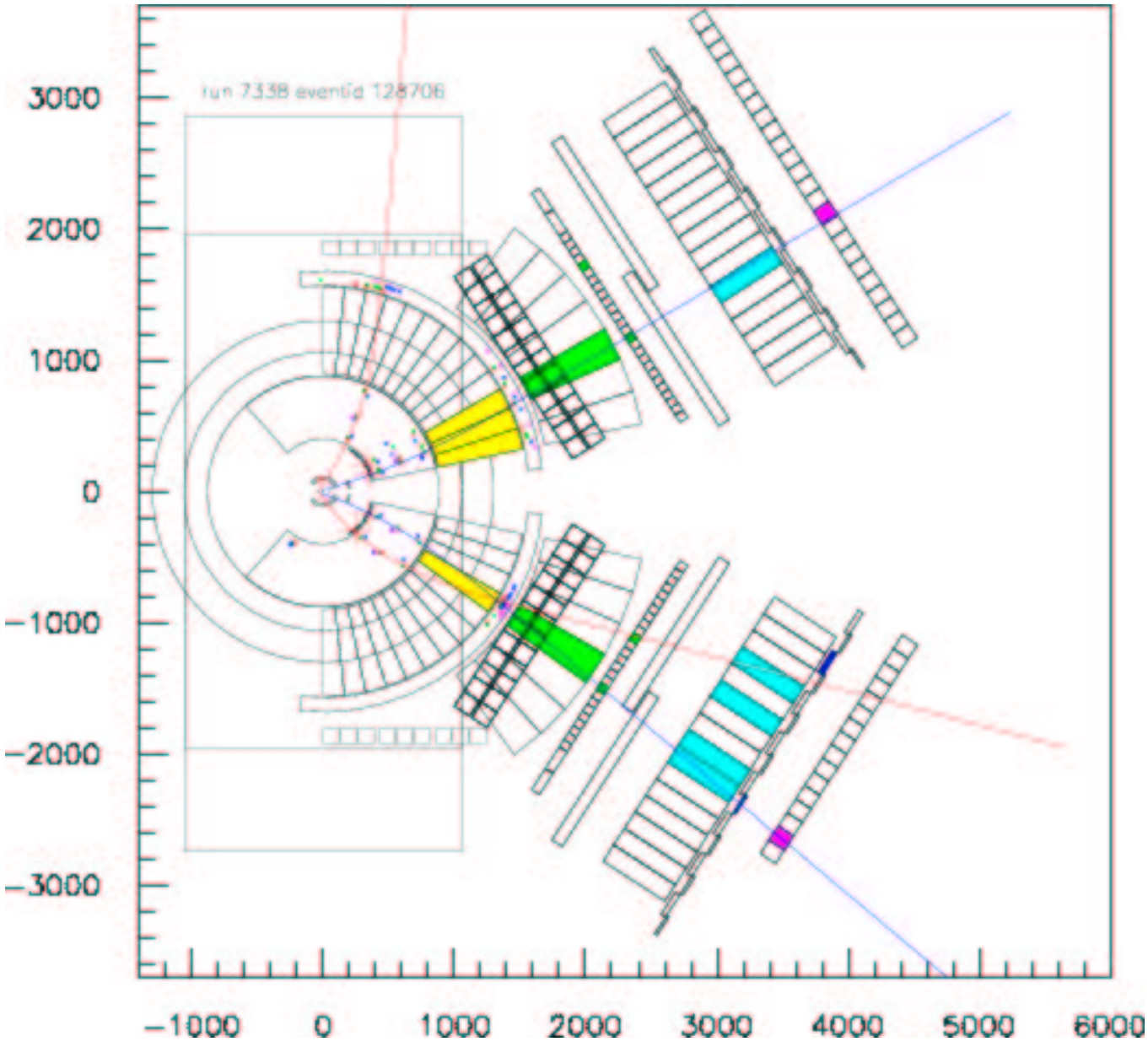
- raw data
  - channel mapping
  - detector hit information



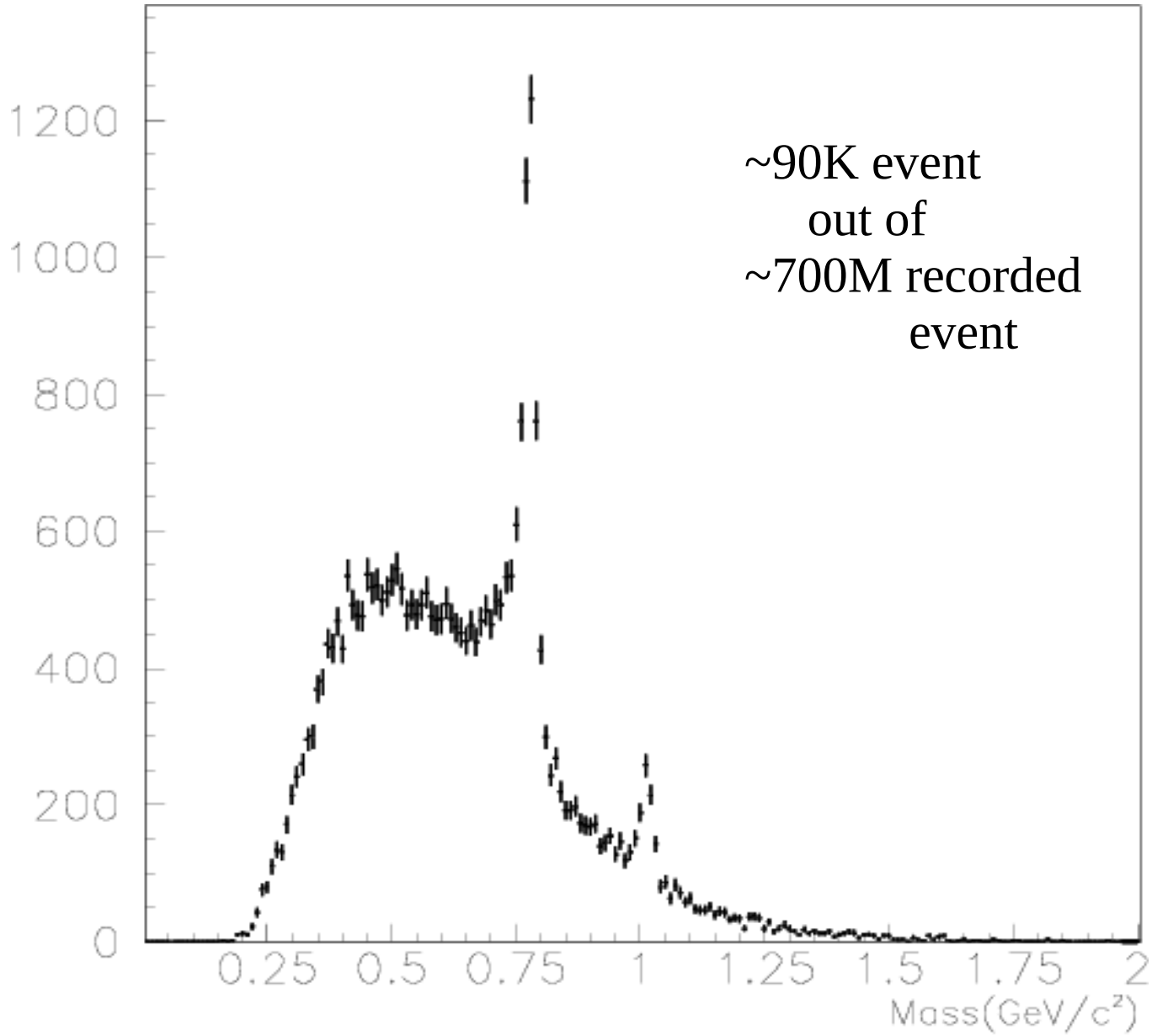
- raw data
- calibration
  - tracker hit position
  - pulse hight-energy relation in EMCal
  - etc.



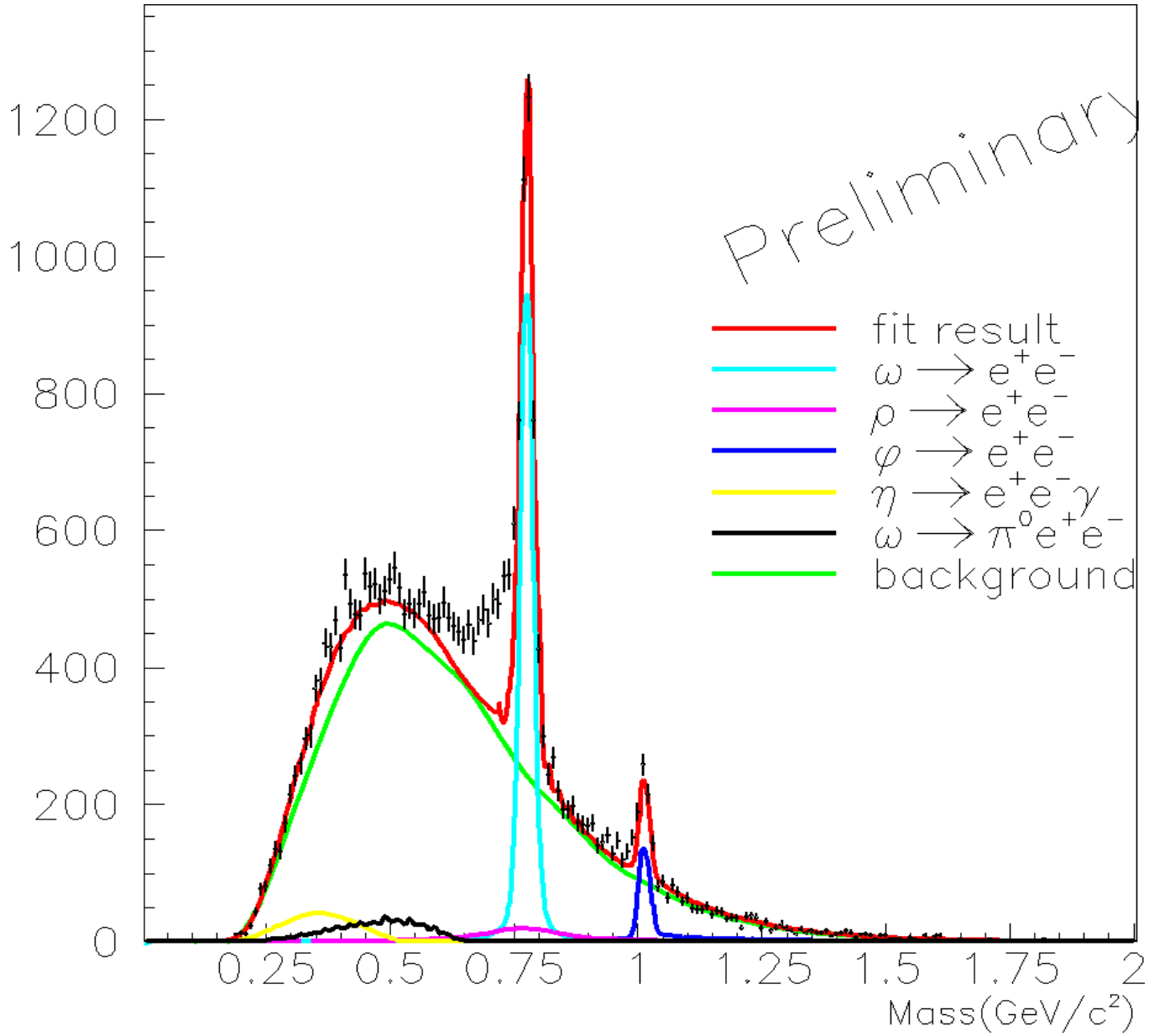
- raw data
- calibration
- track/event reconstruction
  - momentum
  - Particle ID
    - blue:electron
    - red : other
  - invariant mass of eletron pair



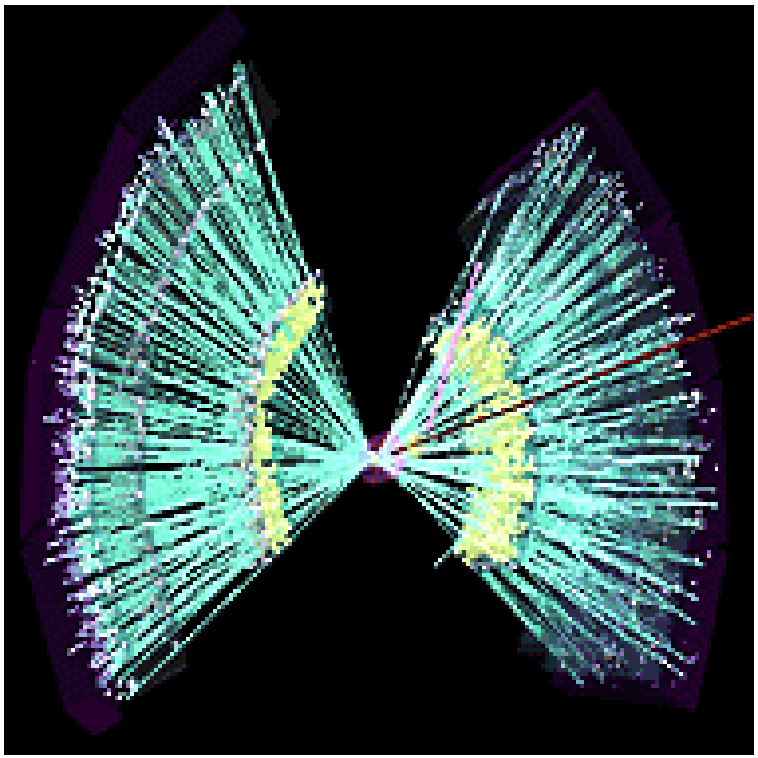
- raw data
- calibration
- event reconstruction
- mass spectra
  - ee invariant mass
  - accumlate events

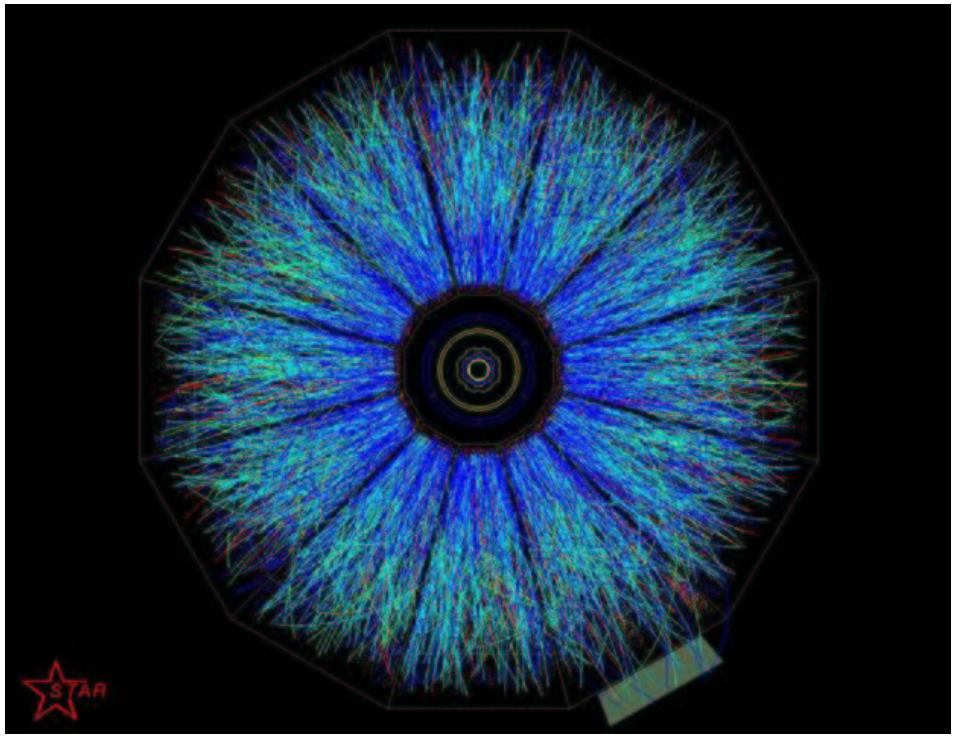


- raw data
- calibration
- event reconstruction
- mass spectra
  - fitting
  - number
  - efficiency correction
- Physics:
  - cross section



event reconstruction in RHIC exp.

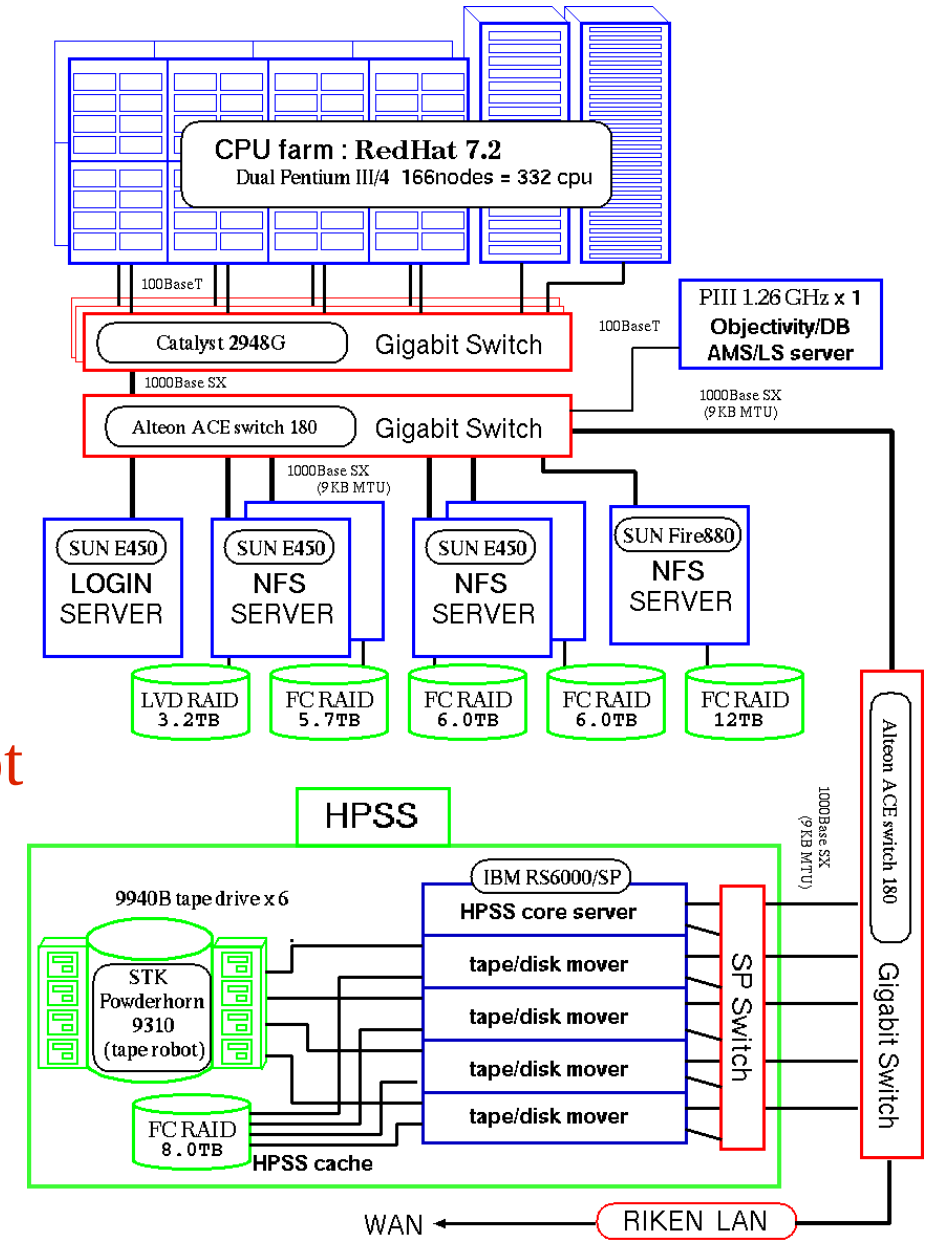




- PHENIX and STAR exp. reconstructed event
- ~ 5 sec/event @ PentiumIII 1GHz
  - ~4500 CPU days for 77M event(~18TB) (Run2 AuAu v3)

CCJ current configuration

- since 1999
- calc nodes :dual P3/P4x166
  - Linux(RH7.2 -> RH8)
- data server : SUN WS x5
  - 33TB raid disk
- Tape storatge: HPSS v4.5
  - 600TB tape in STK tape robot
  - 150TB of data in HPSS
  - tape duplication facility at RCF@BNL ... for transport
- Gbit Ether / Jumbo Frame

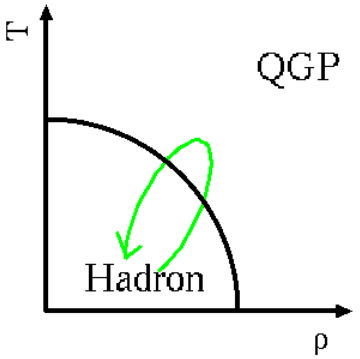


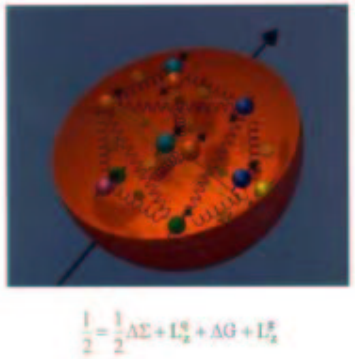
CCJ operation

- Most major offsite computing center in PHENIX
- Tape handling capability : large scale data processing
  - DST produciton -> transfer to RCF@BNL
- data transport by tape and network
  - 50TB transported in this year (10TB by net, 3~4MB/s)
- Usage
  - official DST production / simulation
  - user level analysis
    - nDST production/analysis (including calibration)/simulation
    - I/O bound job
- LSF : batch queueing system
  - 5000 jobs are submitted by one user , running 3~20 hours

Highlight of physics output from CCJ

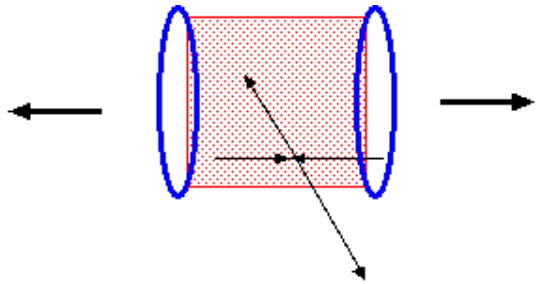
- Quark Gluon Plasma Search at RHIC
  - published paper 15(+6), ~100 of conference proceedings
- Measurement of Nucleon spin structure using polarized proton at RHIC
  - first result from 2003 data
- Detection of chiral symmetry restoration in nuclear matter at KEK-PS



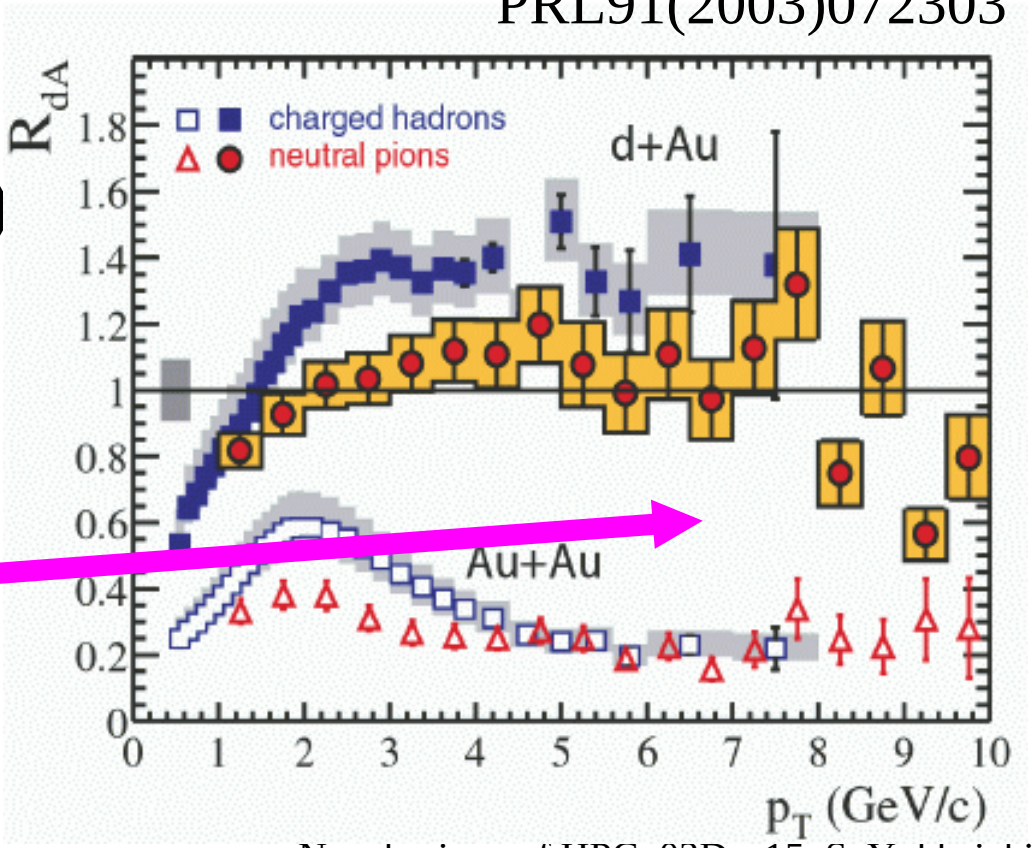


Suppression of High Pt particle
-Jet quenching in Quark Gluon Plasma-

- 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



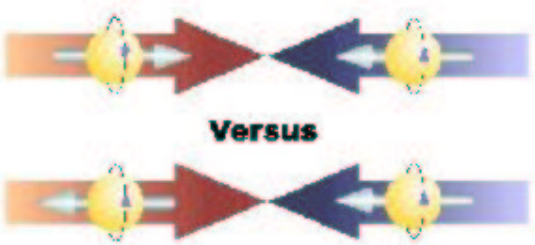
PRL91(2003)072303



Spin Crisis...Origin of the nucleon spin

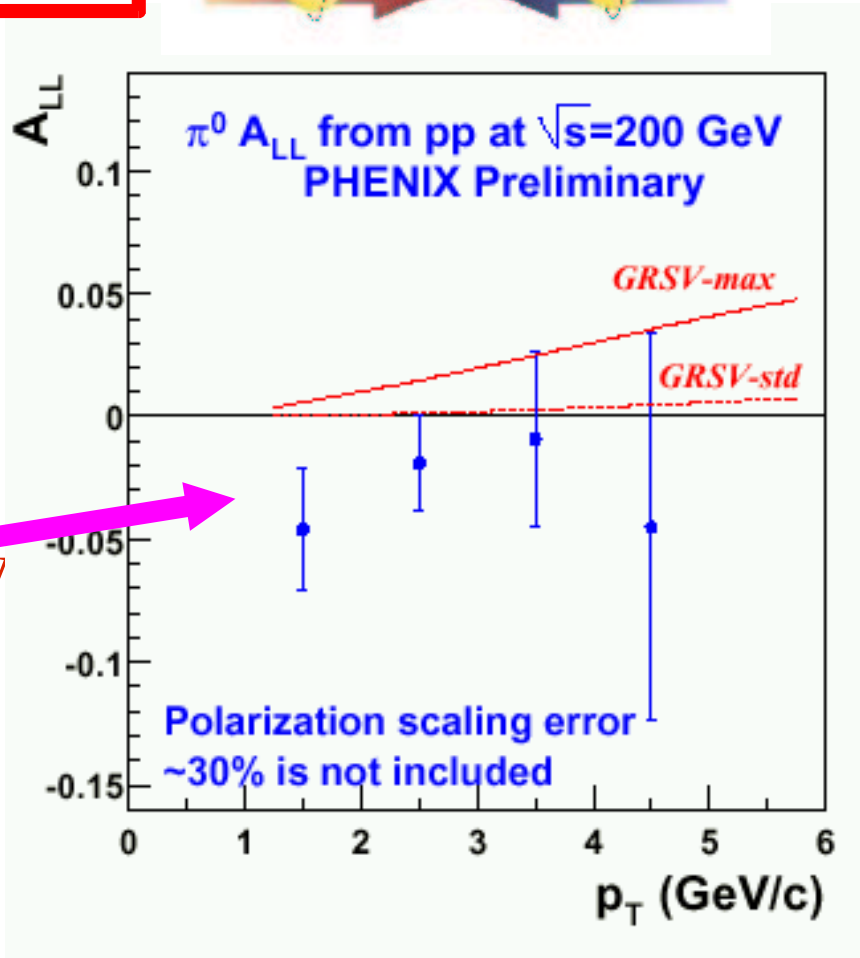
proton spin=$1/2=(1/2)\Delta\Sigma+\Delta G+L_Q+L_G$

$\Delta\Sigma$: Quark Spin ~ 0.2-0.3



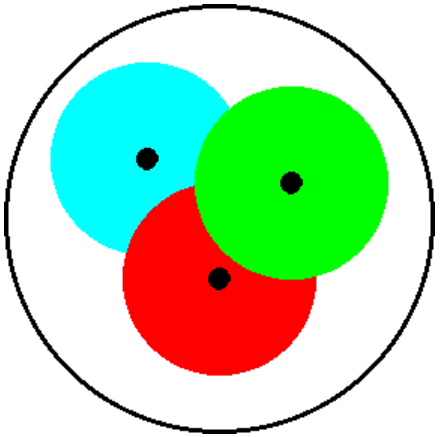
only pol. proton collider can measure the gluon spin ΔG

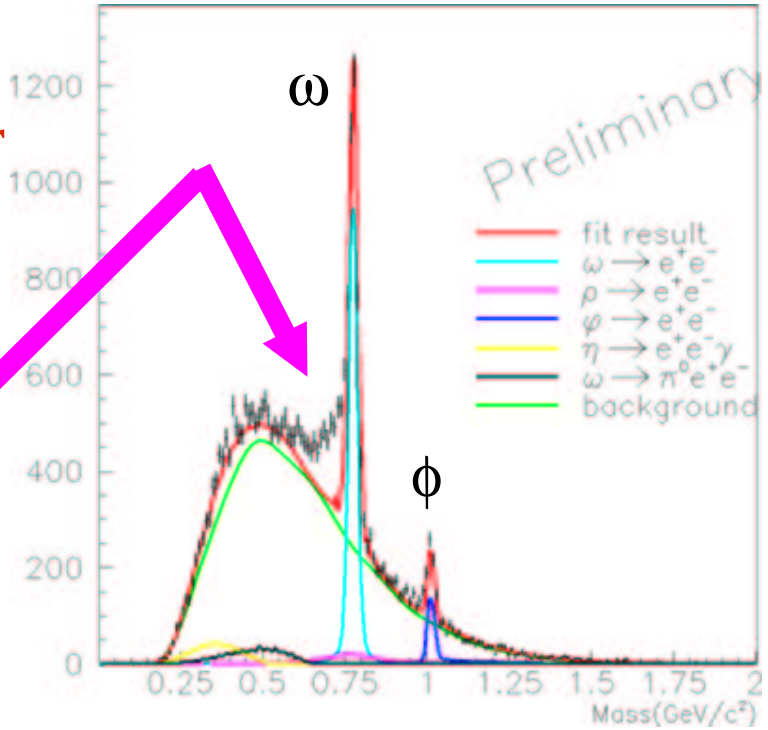
- preliminary result of pi0 asymmetry (from 2003 data)
  - large negative value is not explained usual pQCD theory
  - contribution of gluon is very small ??
  - data accumulation should be done in next run !



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





Prospect of physics at RHIC and RIKEN new system

- Run-4 : 03Nov ~ 04 May
  - QGP search in Au-Au collision
    - J/psi, light vector meson in lepton channel
  - Pol. accel. test for spin physics
- Run-5 : 04 winter~05 summer
  - polarized pp collision for spin physics
    - $\pi 0$ asymmetry problem... enough statistics
    - other channel for gluon polarization
    - in 2~3 years : new detector introduction is planned

- more and more data should be analyzed (100~200TB/year)
  - 2-3 times as large CPU power as now, even if 1/10 of RIKEN new system is available

New horizon of HPC, 03Dec15, S. Yokkaichi

Summary

- In Nuclear/Particle Physics using accelerator, data production rate is growing to 100TB/year ~ 1PB/year
  - GRID technology may be key issue in data recording/distributing/analysis
- PC cluster is used online and especially offline analysis
- Hadronic physics at RHIC
  - QGP & Nucleon spin structure : in progress
    - clue has been obtained
  - CCJ & new system is very powerfull tool in data analysis