JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

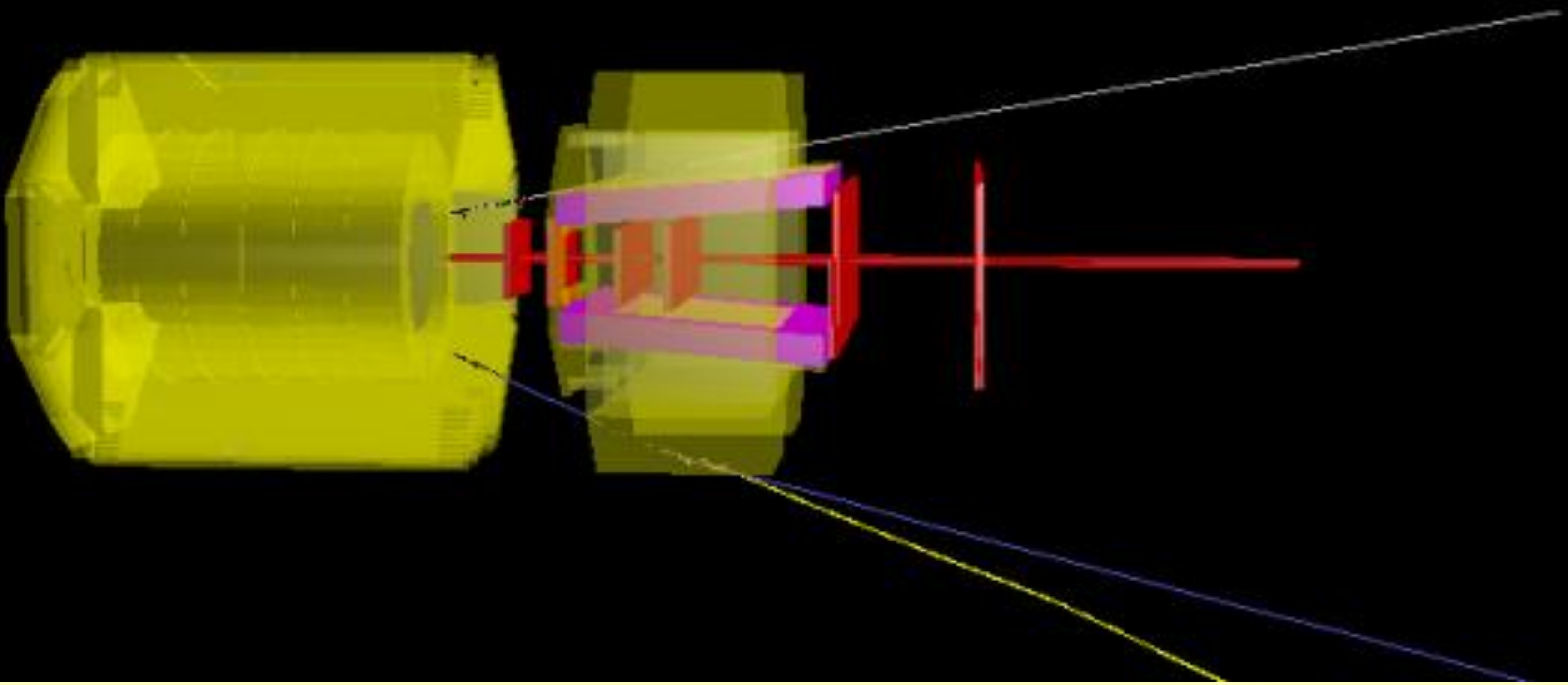
Plans for Online Computing at Panda

Sören Lange, II. Physikalisches Institut, Universität Giessen

Workshop on fast Čerenkov detectors
May 11-13, 2009, Giessen, Germany

Outline

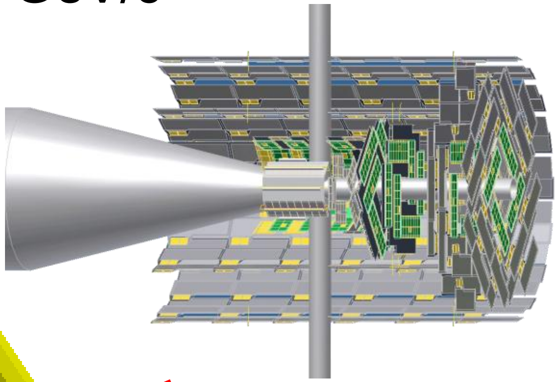
- **What must Panda Online Computing achieve?**
  - **Data rates for DAQ**
  - **Data rates for Offline Computing**
- **The Compute Node (ATCA based FPGA platform)**
  - **Algorithms (Track Finder, Ring Finder)**
- **Experiences for Online Computing:**
  The STAR Level-3 Trigger
- **A new approach:**
  Graphics Adapters as Coprocessors

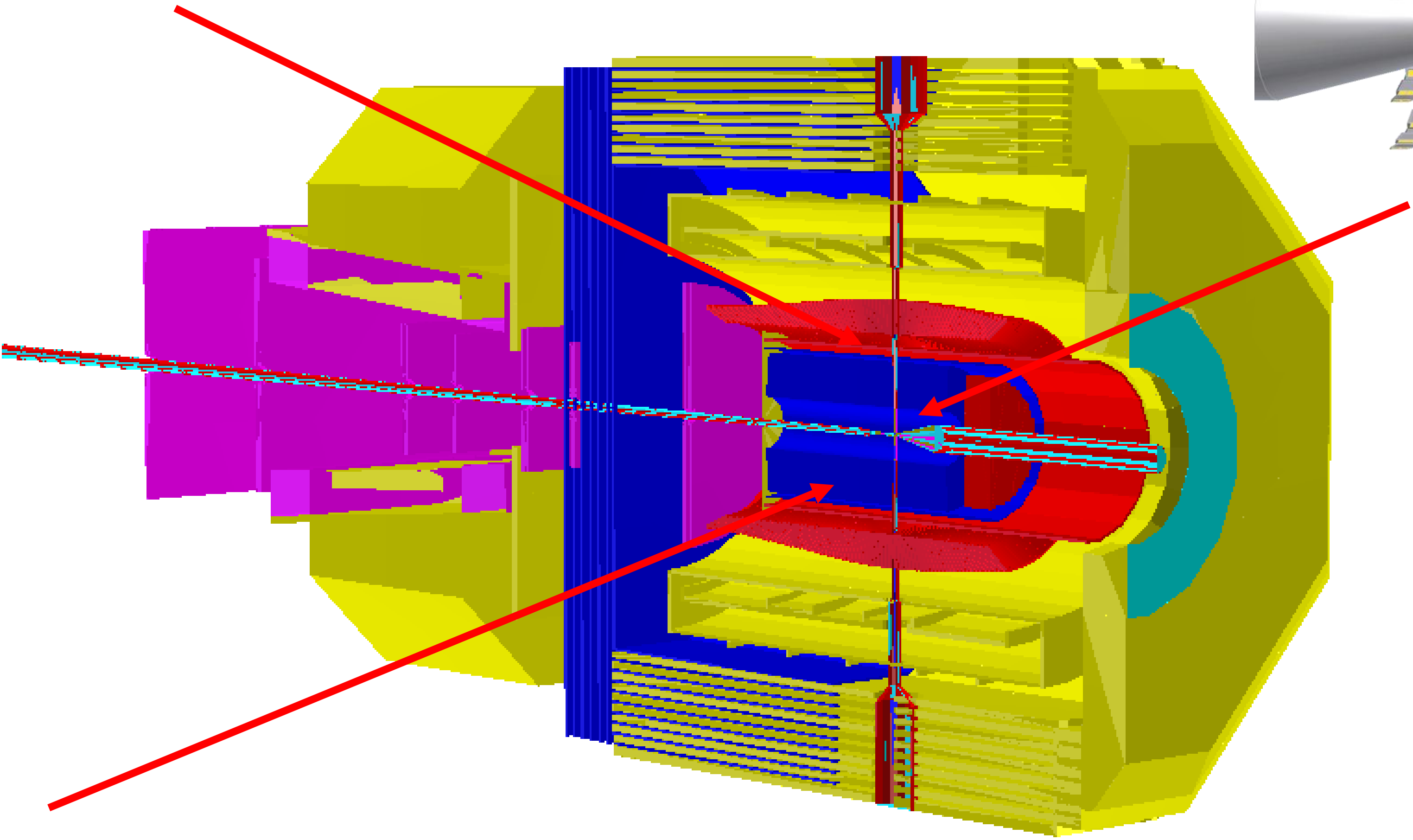


$\bar{p}p$ and $\bar{p}A$, fixed target, $p_{beam} \leq 15$ GeV/c

Panda

DRC, 16 quartz (n=1.47) Bars
d=1.7 cm, R=48 cm





MVD
10^7 pixels
100x100 μm^2
$7 \cdot 10^4$ strips

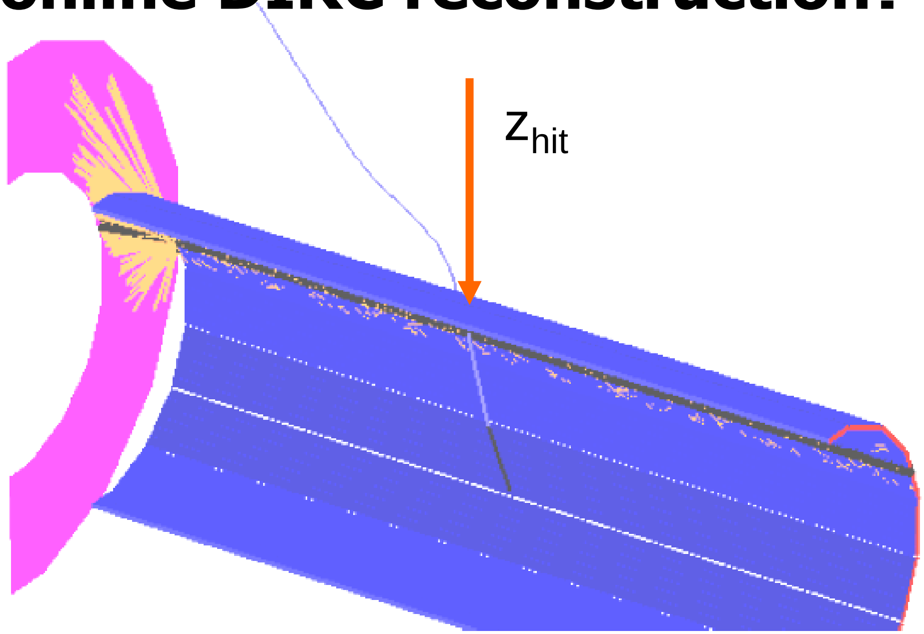
TPC (Time Projection Chamber), R=15-41 cm, 135 padrows, 135,169 pads of 2x2 mm^2
STT (Straw Tube Tracker), 4100 tubes, 15 double layers, R=1 cm, L=1,5 m, axial or skewed

What would we need for online DIRC reconstruction?

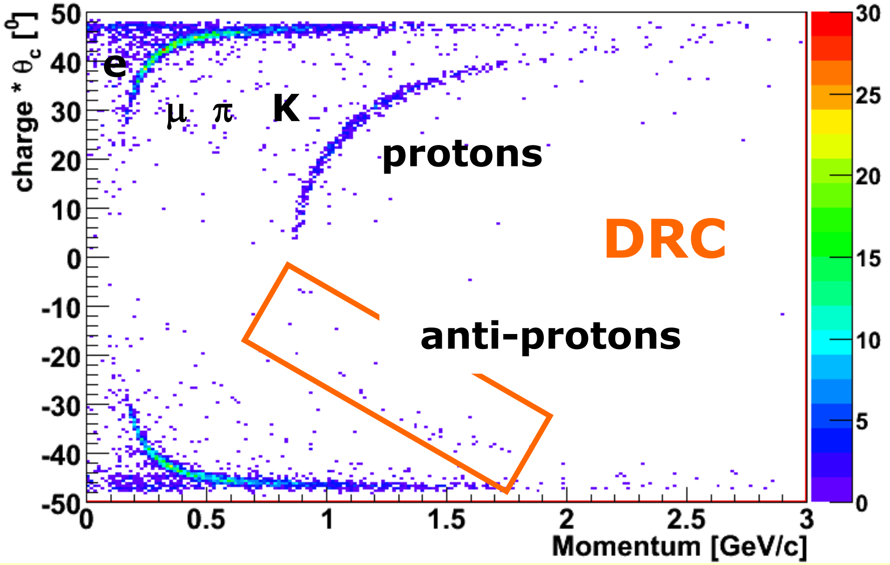
P_{track}

$\vartheta_c = f(R_{ring}, z_{hit})$

z_{hit} is extrapolated hit position of track into DIRC



10000 events, DPM, p_{beam}=6 GeV/c



We need:

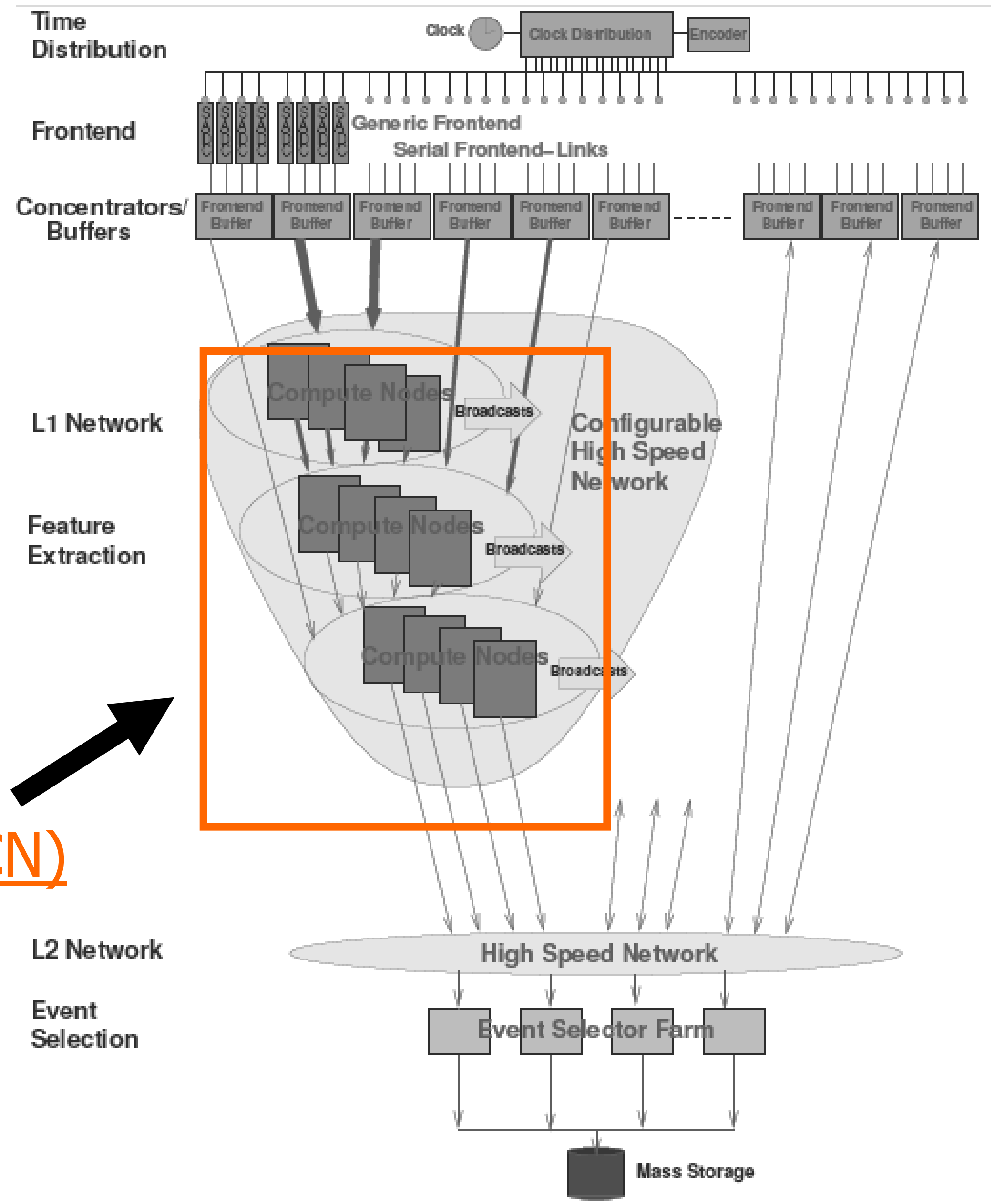
online **ring finder**
→ example from Hades

online **track finder** and **track fitter**
→ example from Panda MC

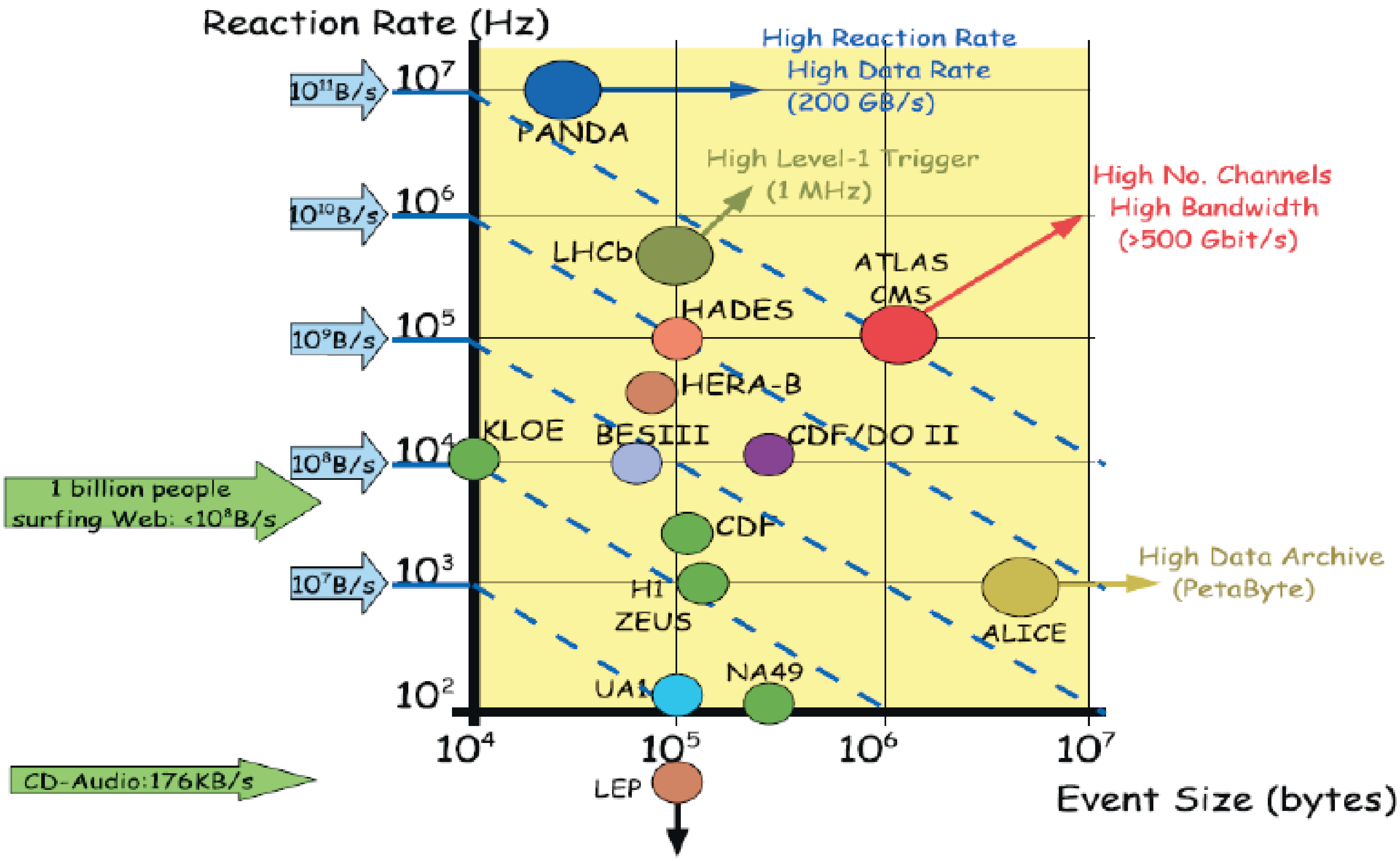
online track **extrapolation** to DIRC
→ example from STAR

Panda DAQ

- $\leq 3 \cdot 10^7$ events per 1 s
- Event size 4-20 kB
- No hardware triggers
- Digitization on frontend
- All raw data are streamed to DAQ
- Data reduction with 3 levels of Compute Nodes (CN)
- Requires intelligent reduction algorithms in VHDL on FPGAs



Bandwidth Requirement



(Document on Offline Computing, to be submitted soon)

The Computing Model of

$\overline{\mathrm{P}}$ANDA

(AntiProton Annihilations at Darmstadt)

$\overline{\mathrm{P}}$ANDA Collaboration

$\simeq$ 400 authors

The $\overline{\mathrm{P}}$ANDA experiment at the future facility, FAIR, will extensively study the strong interaction by providing precision data to test the theory of Quantum Chromodynamics. Intense beams of anti-protons together with a multi-purpose and compact detection system will be the basis for these experiments.

The $\overline{\mathrm{P}}$ANDA experiment is challenged with the production of PBytes of data per year, together with computing intense operations to carry out the data analysis. The final results need to fulfill the requirements of being extremely reliable together with an optimum precision. The data and the planned computing environment need to be accessible for a large research community, even beyond the $\overline{\mathrm{P}}$ANDA collaboration itself.

This document presents the first version of a computing model for $\overline{\mathrm{P}}$ANDA. It describes a preliminary scheme of the data processing, a first estimate on the capacity requirements (CPU, storage, and network), our plans for distributed computing, and an overview of the requirements for software and its development. Furthermore, this document summarizes the wishes and requirements for a high-end data analysis scheme, which study was made prior to this more complete document.

The Panda Computing Model: Panda Offline Computing Requirements

- Rate to tape: 25 kHz
  x 175 days = 3.78 10^{11} events on tape per 1 year
- 4 kB for micro-DSTs
  = 1,512,000 Gbyte ($\geq$ 1 PetaByte)
  $\simeq$ 378,000 DVD´s per year
- Estimate for 1^{st} year of Panda data taking
  - incl. MC, raw data, hit data etc.
    $\rightarrow$ 11.5 PetaByte per year
  - $\simeq$ 2000 quad core CPUs
    for reconstruction, analysis and MC production
  - $\simeq$ 3.7 Mill. Euro for CPU´s and disk storage
    (only in the 1^{st} year)

Summary

Panda <u>Online</u> Computing must perform full reconstruction to achieve a data reduction factor:

$3 \cdot 10^7$ events/s raw data

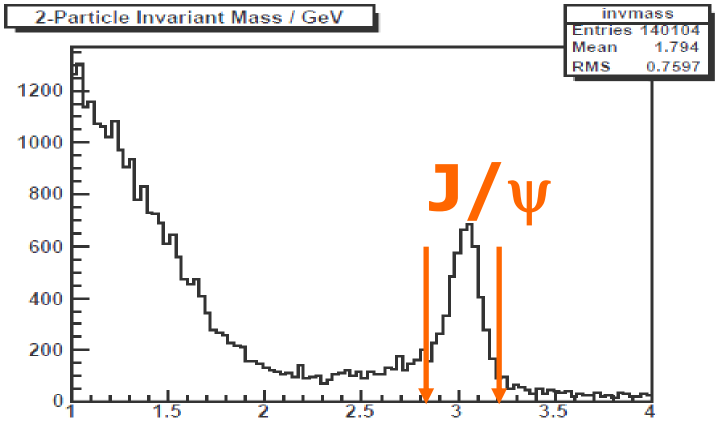
↓ factor $\simeq$ 1200

$25 \cdot 10^3$ events/s to tape

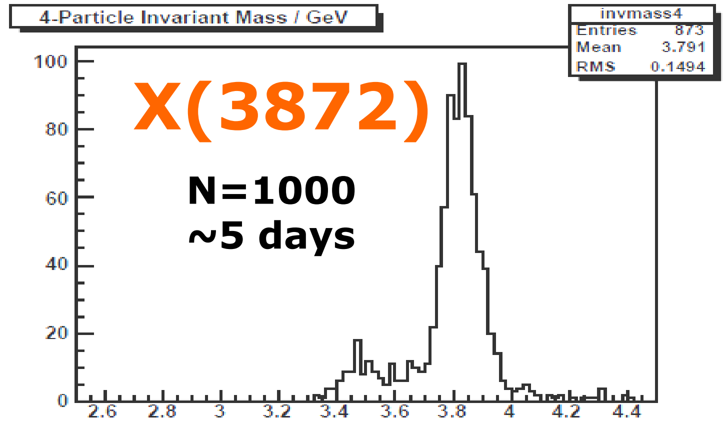
Why does Panda have no hardware L1 trigger?

- Background Events have very similar signature to signal events
- signal example: charmonium, $\bar{p}p \to \bar{c}c$ (X)
- background: $\bar{p}p \to \bar{u}u, \bar{d}d, \bar{s}s$ (X)
- there are no simple L1 trigger criteria (e.g. N_{track}) to distinguish signal and background
- we have to calculate invariant mass online on the Compute Nodes

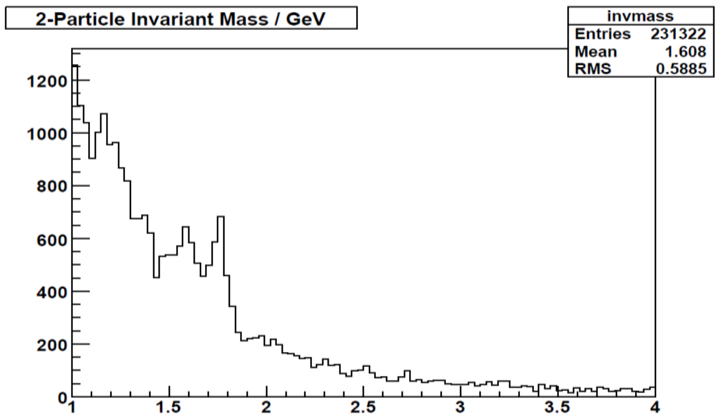
An Example for Panda Physics: X(3872)



2-Particle Invariant Mass / GeV



4-Particle Invariant Mass / GeV



2-Particle Invariant Mass / GeV

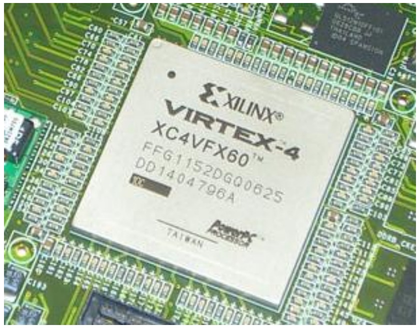
Background: Dual Parton Model,
p_{pbar}=6.991 GeV/c,
N=10^4 Events

Expected Yield: $\sigma\cdot BR \simeq$ 250 pb, N~200 events per day

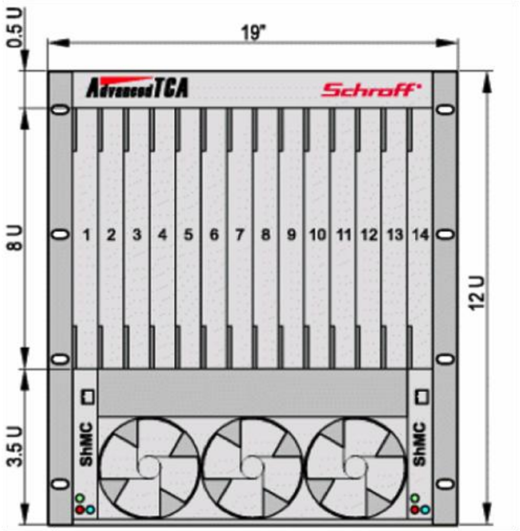
The Compute Nodes

(ATCA based FPGA Platform)

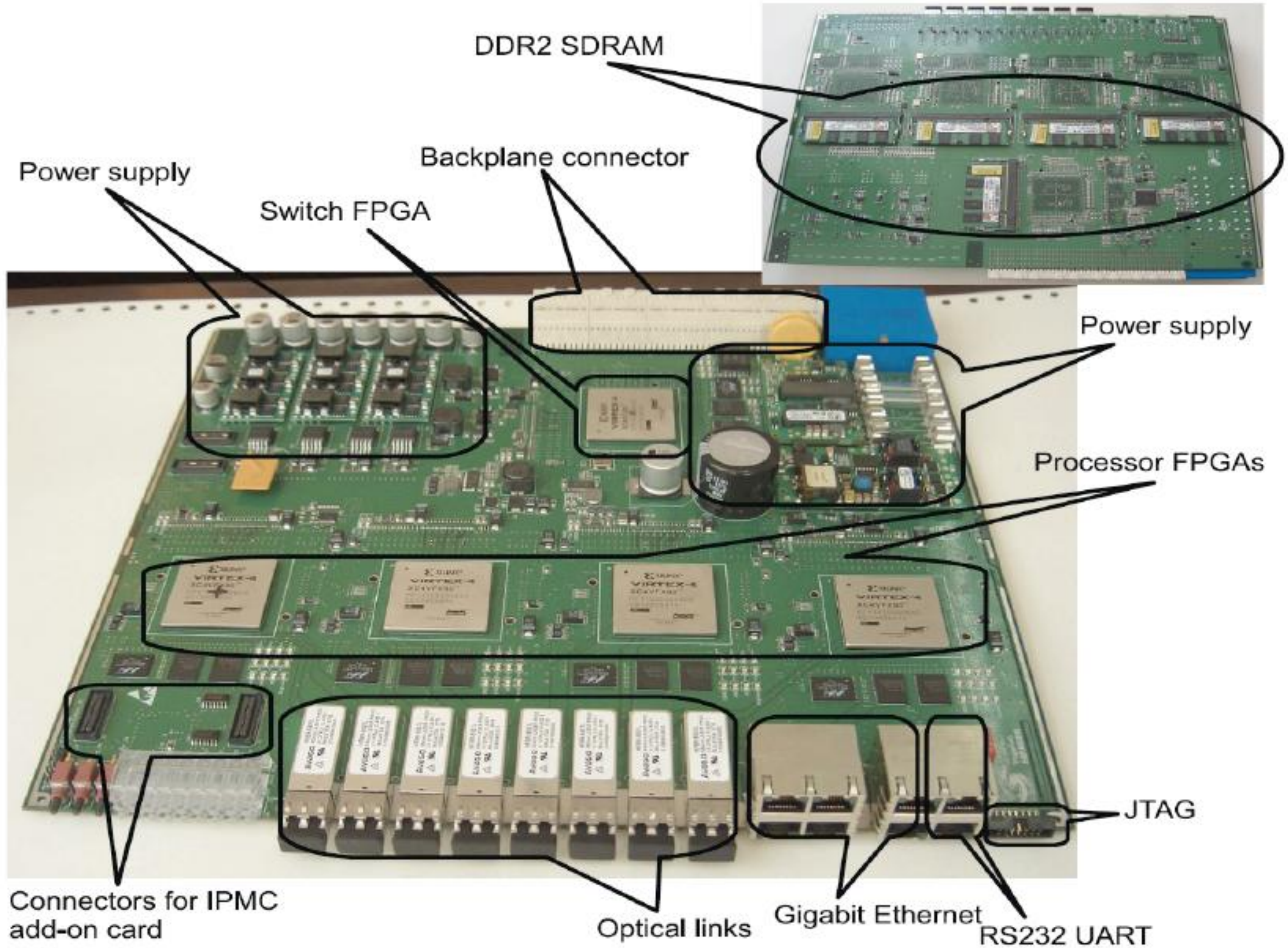
Compute Node (CN) Concept



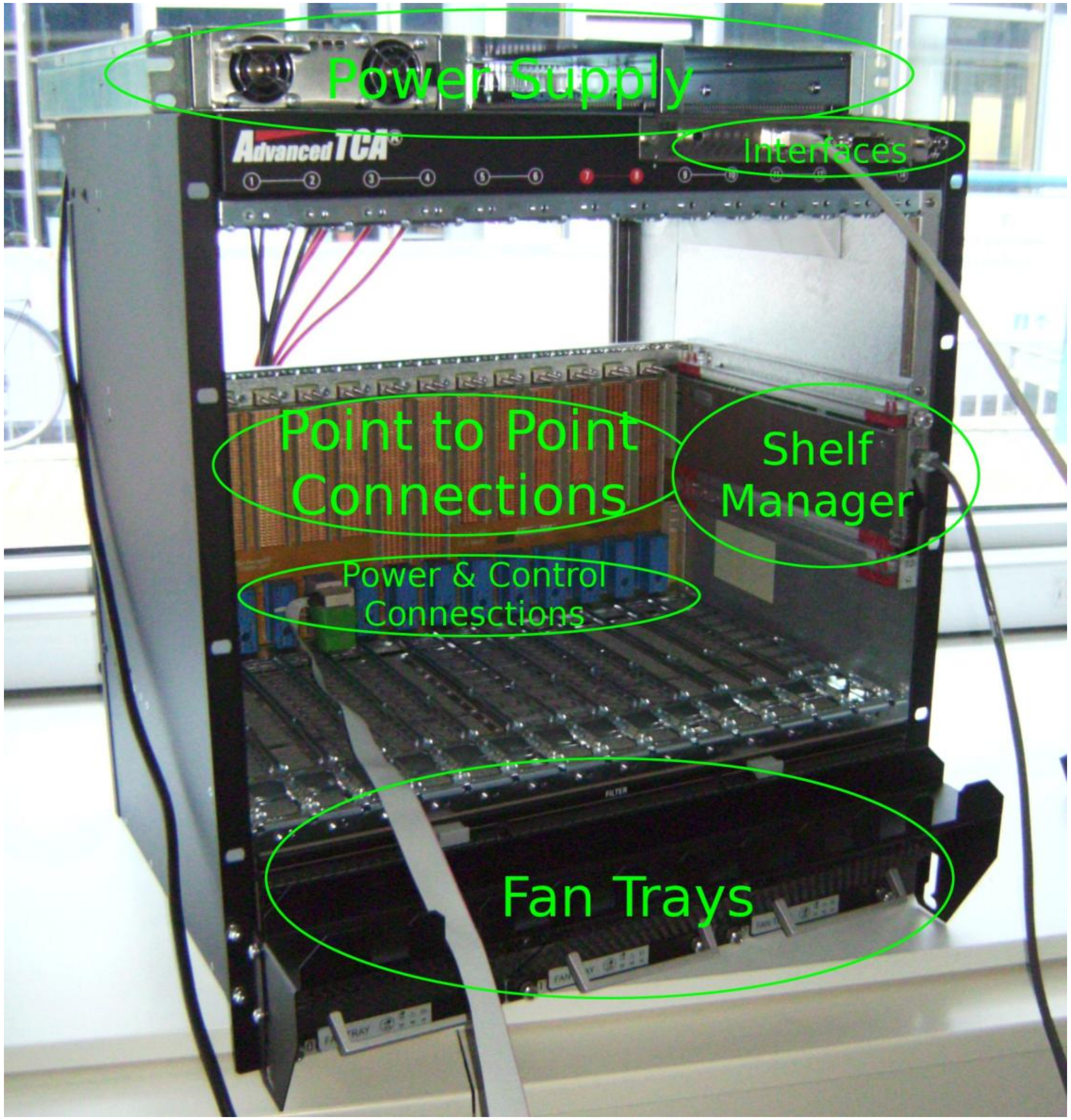
- 5 x VIRTEX4 FX-60 FPGAs
  - each FPGA has 2 x 300 MHz PowerPC
  - Linux 2.6.27 (open source version)
  - algorithm programming in VHDL (XILINX ISE 10.1)
- ATCA (Advanced Telecommunications Computing Architecture) with <u>full mesh backplane</u>
  point-to-point connections on backplane
  from each CN to each other CN,
  i.e. no bus arbitration
- 8 optical links
  (connected to RocketIO at FPGA)
  $\leq$ 6.5 Gbits/s
- 5 x Gigabit Ethernet
- ATCA management (IPMI)
  by add-on card



Compute Node (CN) Version 1.0



ATCA Shelf



Power Supply
Interfaces
AdvancedTCA
Point to Point Connections
Shelf Manager
Power & Control Connesctions
Fan Trays

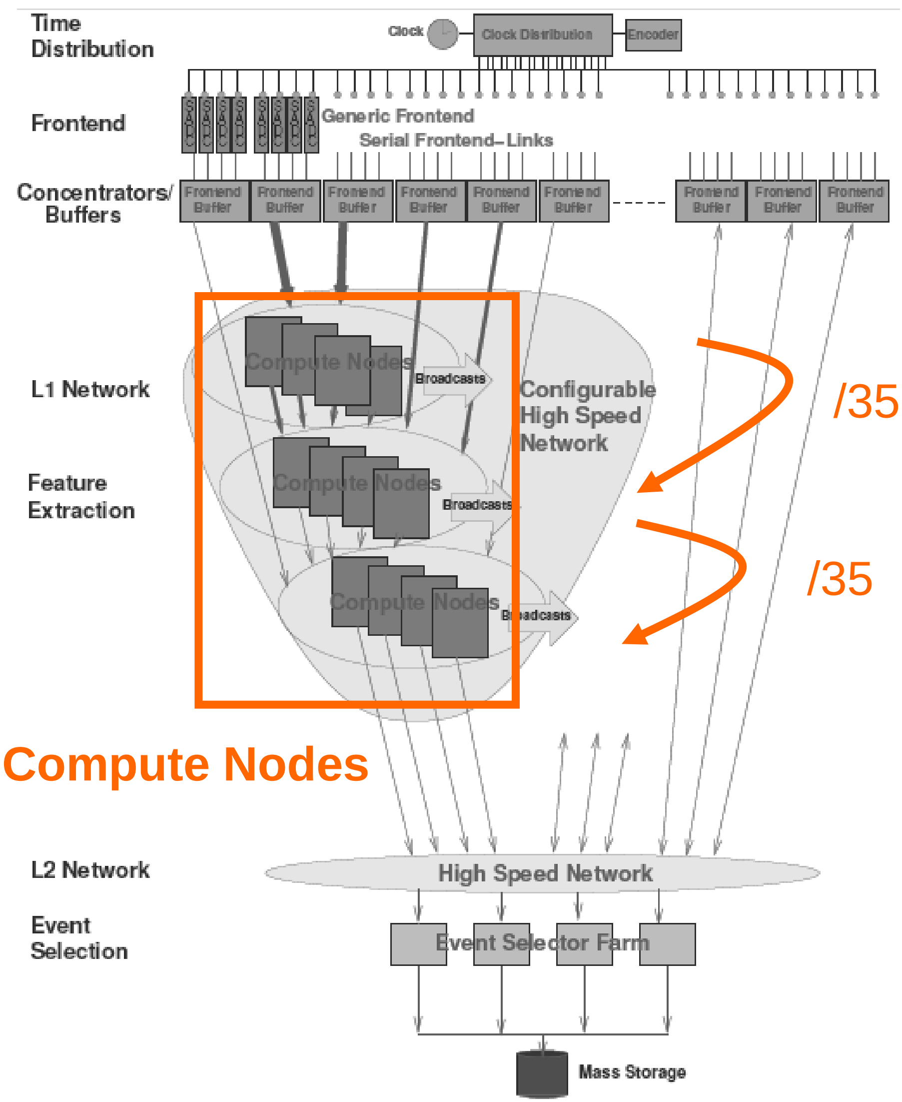
No hardware triggers

All raw data streamed

„Level 1“

„Level 2“

„Level 3“



Total Panda raw data
30 MHz x 14 kByte
= 420 Gbyte/s

Input to 1 ATCA:
14 CN x 8 optical links
= 91 Gbyte/s

Output of 1 ATCA:
14 CN x 5 GB Ethernet
= 2.6 GByte/s

factor 35 reduction required

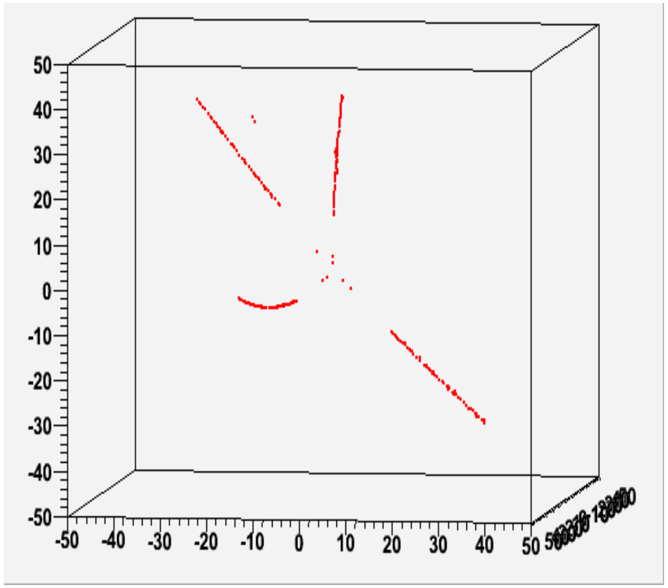
Algorithms for the Compute Nodes

(FPGA Programming in VHDL)

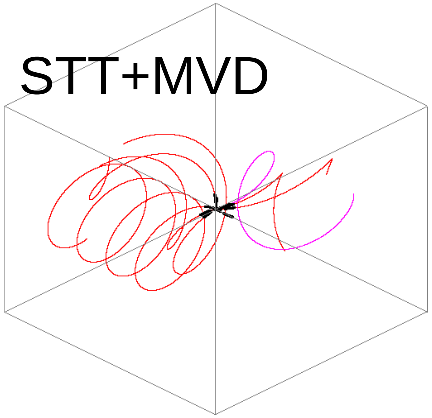
Algorithms for the Compute Node
Panda Track Finder and Track Fitter for FPGA

- **STT (Straw Tube Tracker) +**
  MVD (Micro Vertex Detector)
  helix, 30+(5-7) hits per 1 track
- **Field B_z=2 T**
  (TOSCA Field Maps)
- **Step #1: Conformal Map**
- **Step #2: Hough Transform**
- **Example:**
  10 $\mu^{\pm}$ Tracks p=1 GeV/c

digitized XYZ coordinates / cm



approach:
helix track
parametrization



Example X(3872) $\rightarrow J/\psi\ \pi^+\ \pi^-$
TPC digitization, MVD digitization

David Münchow, Giessen

Development of FPGA Track Finder for Panda STT and MVD
Step #1: Conformal Map

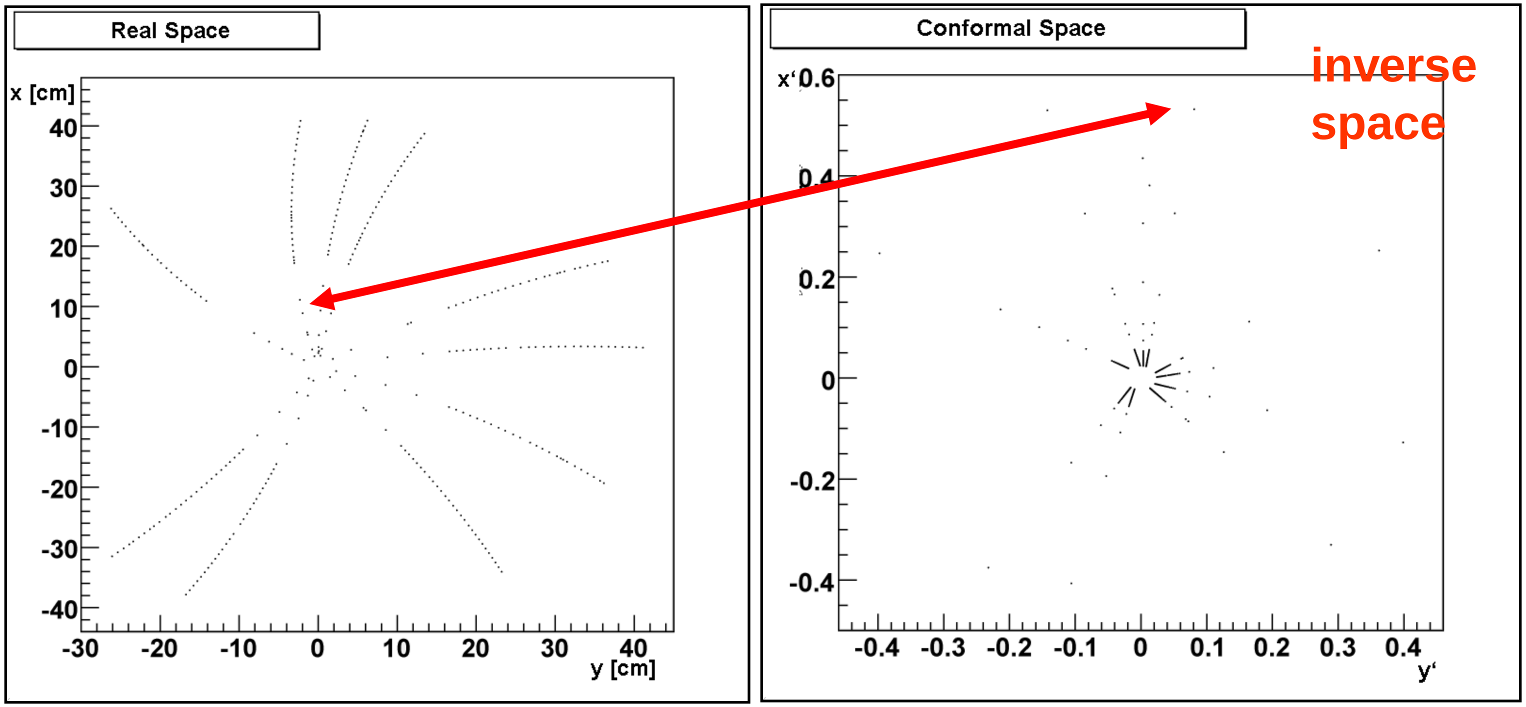
real space

$$x' = \frac{x - x_0}{r^2}$$

conformal space

$$y' = \frac{y - y_0}{r^2}$$

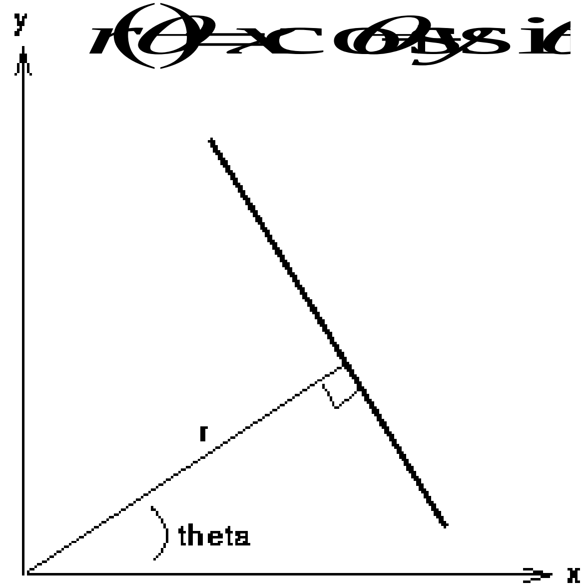
$$r^2 = (x - x_0)^2 + (y - y_0)^2$$



Development of
FPGA Track Finder for Panda STT and MVD
Step #2: Hough Transform
(work ongoing)

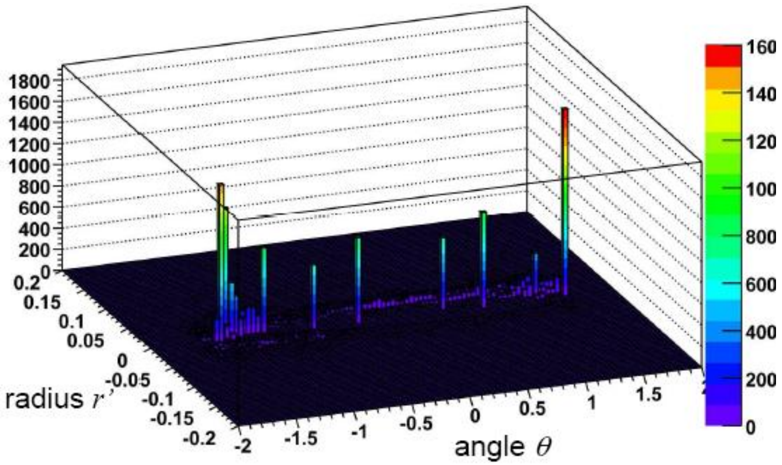
$$r(\theta) = x\cos\theta + y\sin\theta$$

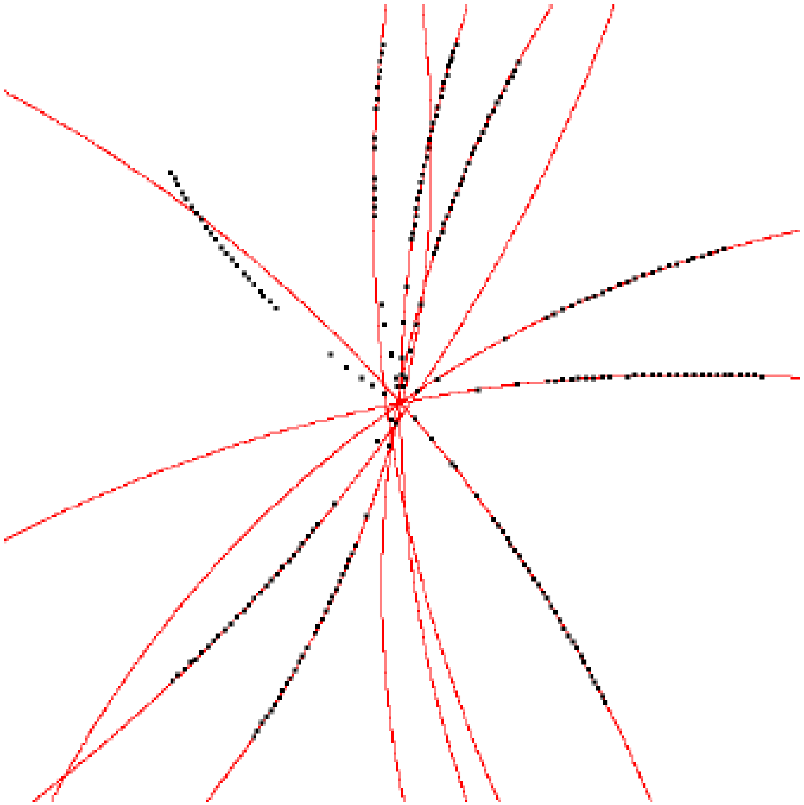
- **fix point arithmetics (instead of floating point)**
- **24 bit (in division and multiplication 48 bit)**
- **Hough space 512×512**
- **lookup table for sinus(): 128 x 16 bit**



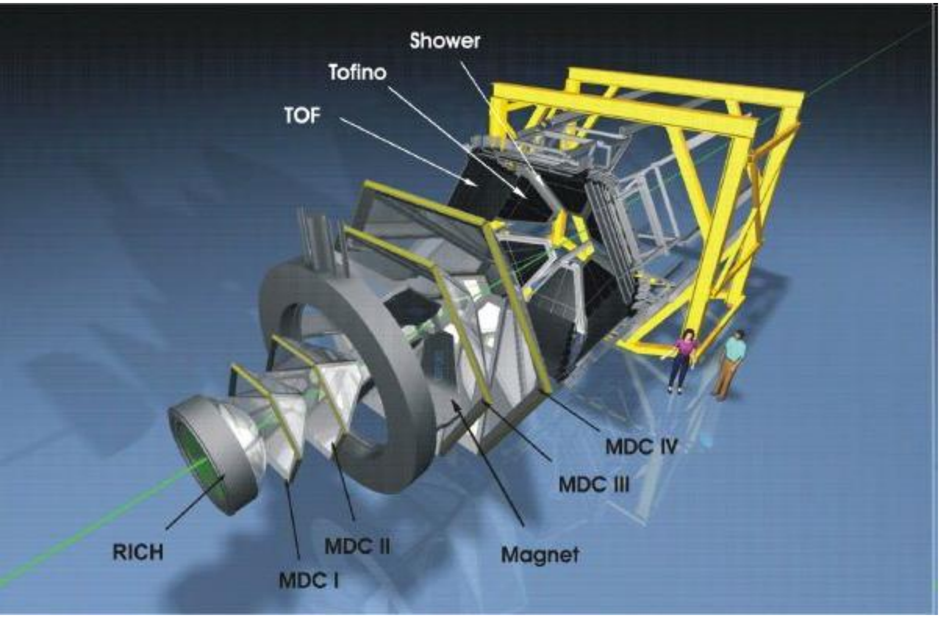
Development of
FPGA Track Finder for Panda STT and MVD
Step #2: Hough Transform
(work ongoing)

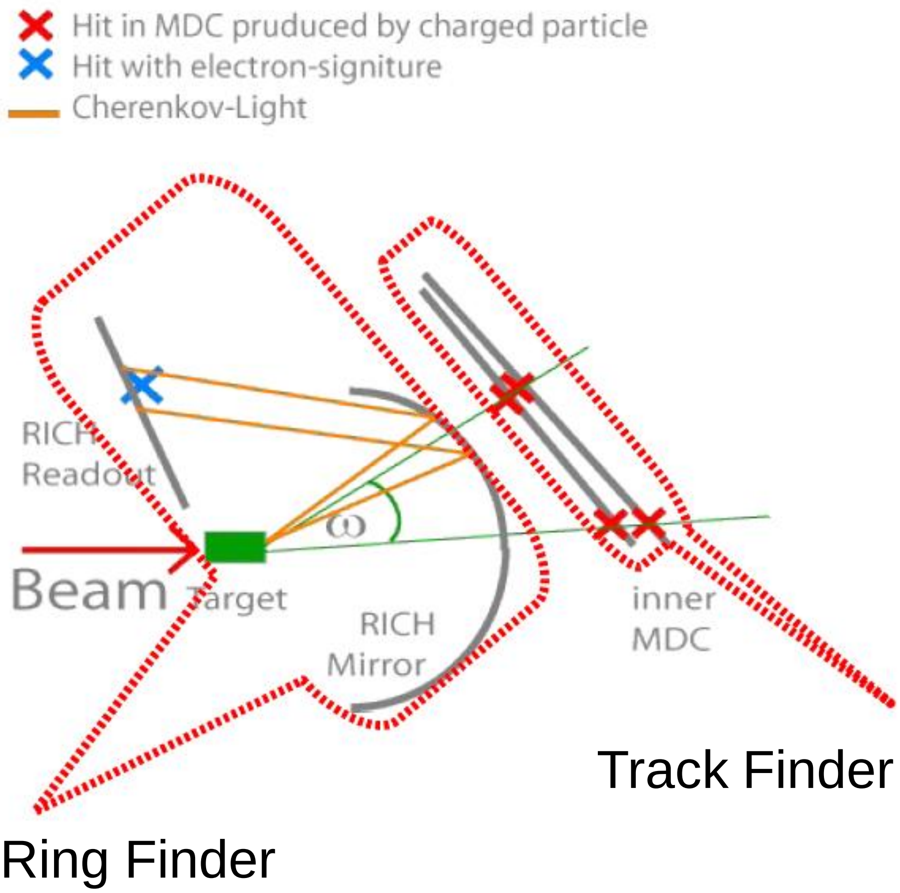
Preliminary result
from algorithm emulation
for 10 $\mu^{\pm}$ tracks, p=1 GeV/c





HADES RICH Ring Finder Algorithm Development incl. FPGA implementation (work ongoing)





Johannes Roskoss, Giessen

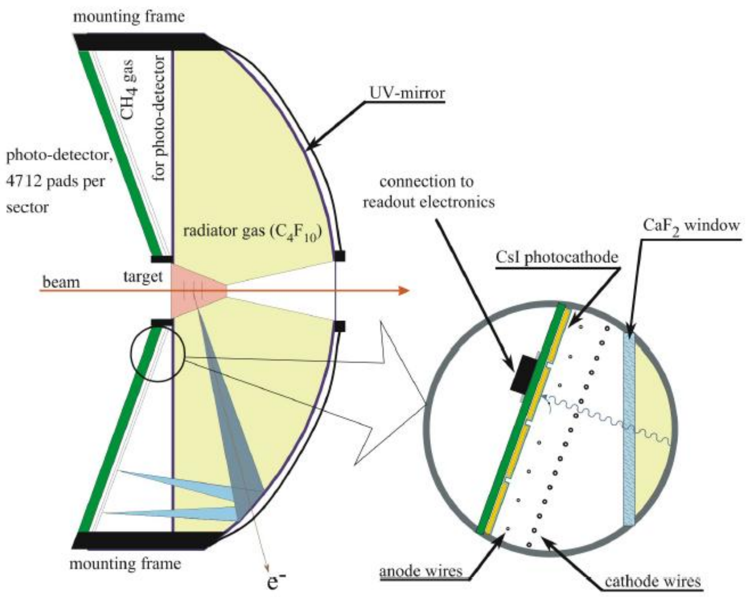
HADES RICH Ring Finder Algorithm Development incl. FPGA implementation (work ongoing)

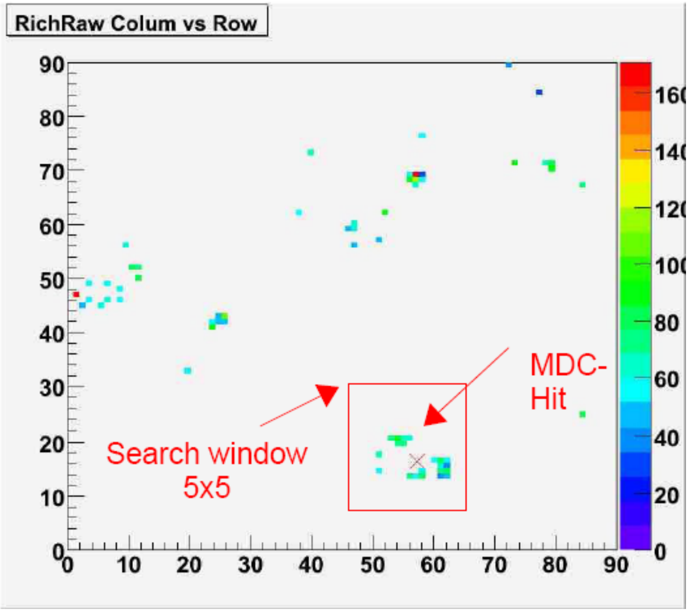
55,296 pads

of different shapes

13x13 pad search region

inner and outer veto rings



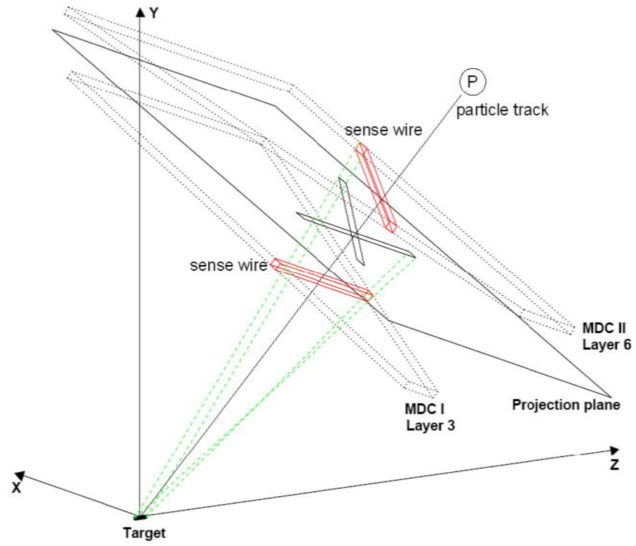


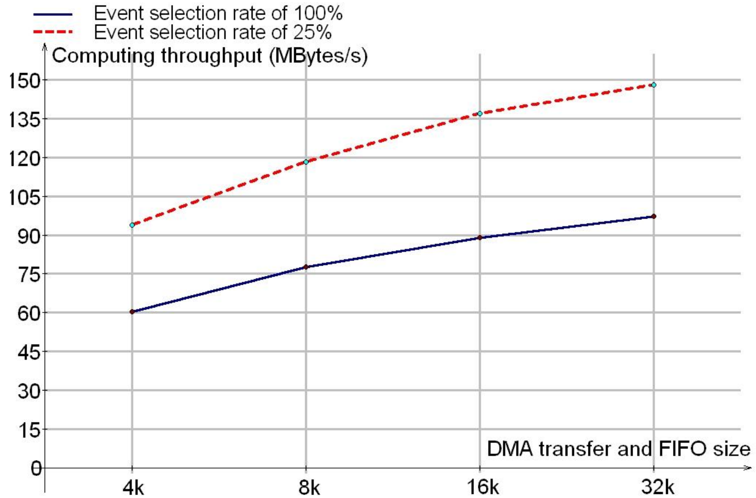
incl. track extrapolation
match ring to drift chamber track
(i.e. straight line approximation,
$\lesssim$ 12 wires per track)

Johannes Roskoss, Giessen

Additional Compute Node Algorithm Development all incl. FPGA implementation (work finished)

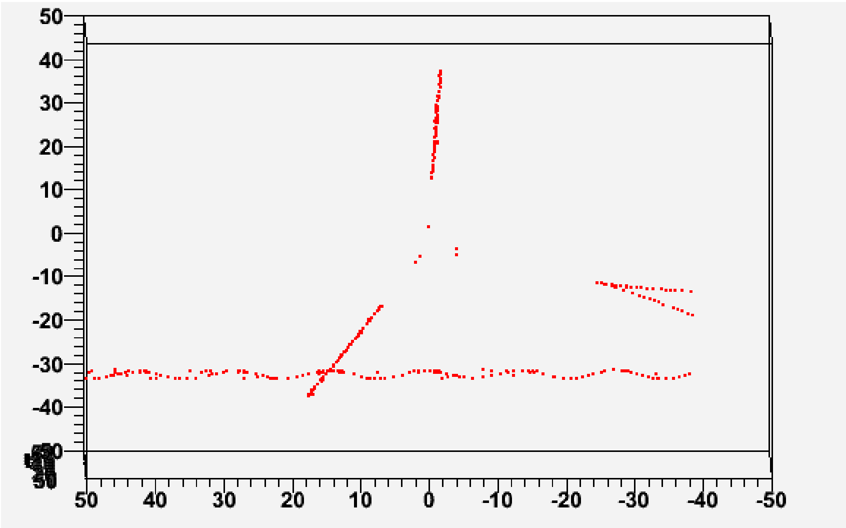
- **Ming Liu – HADES track finder**
  - **straight line, ≤12 wires per track**
  - **compare speed to C program running on Xeon 2.4 GHz (as software reference)**
  - **different** wire **multiplicities N_{wire}=10, 30, 50, 200, 400 fired wires of 2110**
  - **speedup of 10.8 – 24.3 (compared to reference)**
- **Shuo Yang – HADES event selector**
  - **read HADES binary events from DDR2 memory**
  - **decode, issue accept/reject, discard or write back to DDR2**
  - **achieved ≃ 100 MB/s for 32 kB fifo and DMA size (for 100% event accept)**



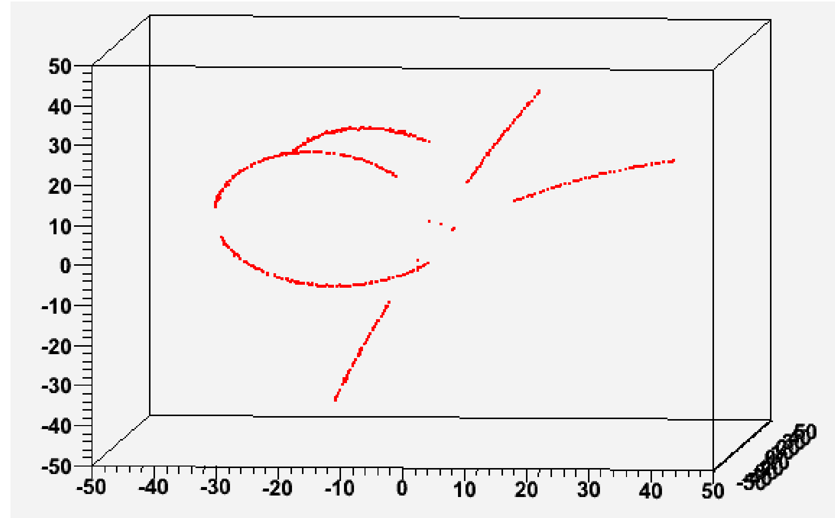


Quality of Panda online track finding?
X(3872)→ J/ψ π^+ π^- Events in Panda MVD+TPC, Fake Tracks

XYZ coordinates / cm

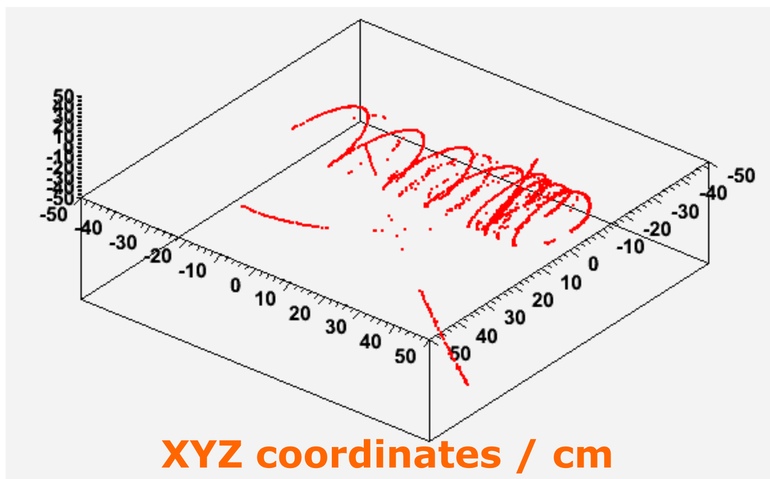


XYZ coordinates / cm



Background Tracks from material

Simulations with PandaRoot 2.0



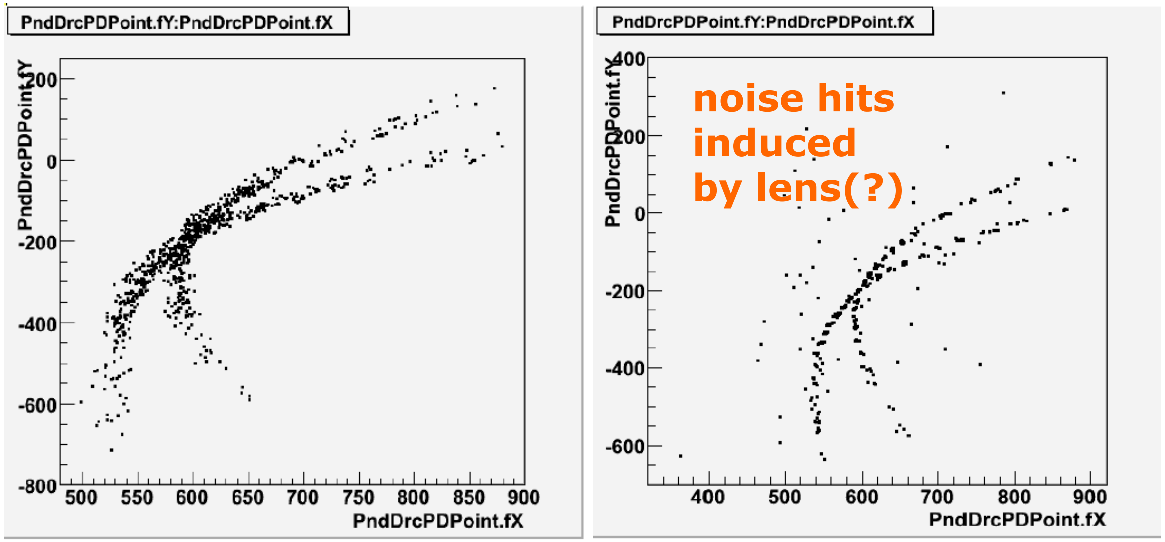
Beampipe Interaction

Quality of Panda online ring recognition?
Ring Patterns with Background
Example: implementation of a lens system

Lens shape calculated by external optics program

Simulations with PandaRoot 2.0

No Lens **Lens**



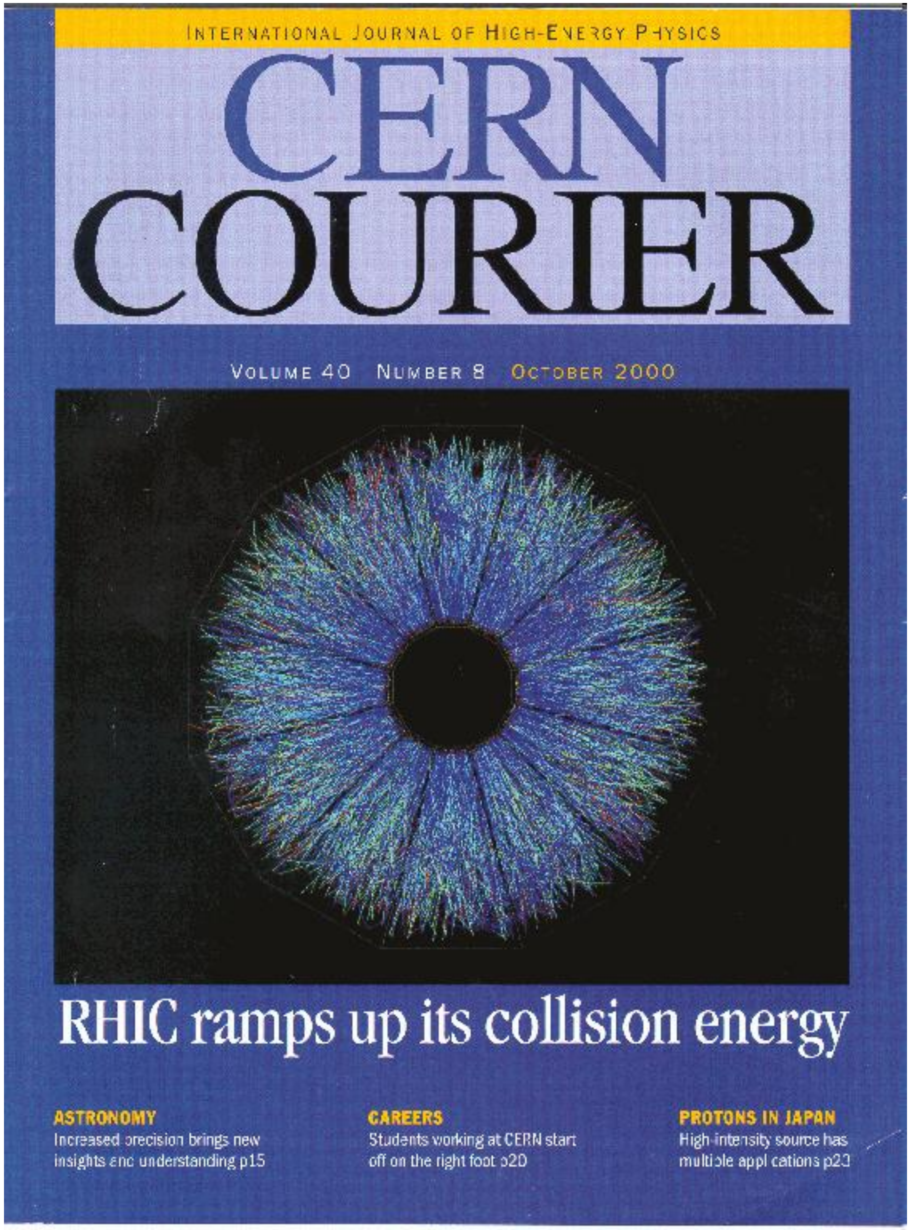
Kaon, p=0.6 GeV/c

C. Schwarz, A. Cecchi

Online Computing at STAR

Level-3 Trigger for the STAR TPC
(1999-2004)

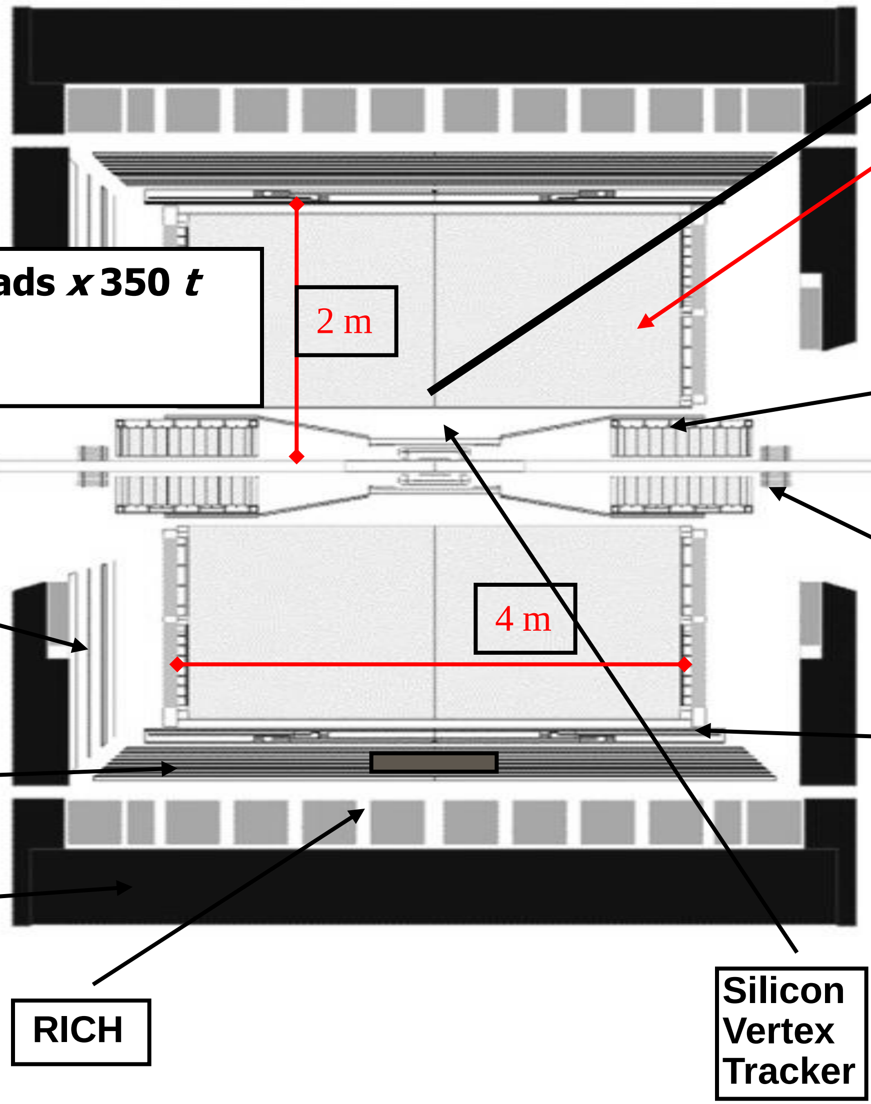
Realtime Track Finder and Track Fitter ($t \leq 120$ ms)
using Conformal Map

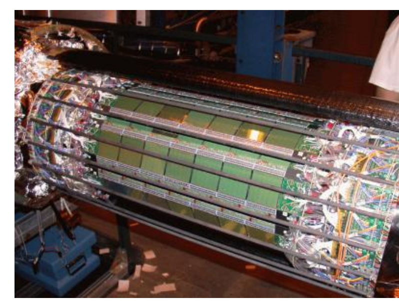

INTERNATIONAL JOURNAL OF HIGH-ENERGY PHYSICS
CERN
COURIER
VOLUME 40 NUMBER 8 OCTOBER 2000
RHIC ramps up its collision energy
ASTRONOMY
Increased precision brings new insights and understanding p15
CAREERS
Students working at CERN start off on the right foot p20
PROTONS IN JAPAN
High-intensity source has multiple applications p23

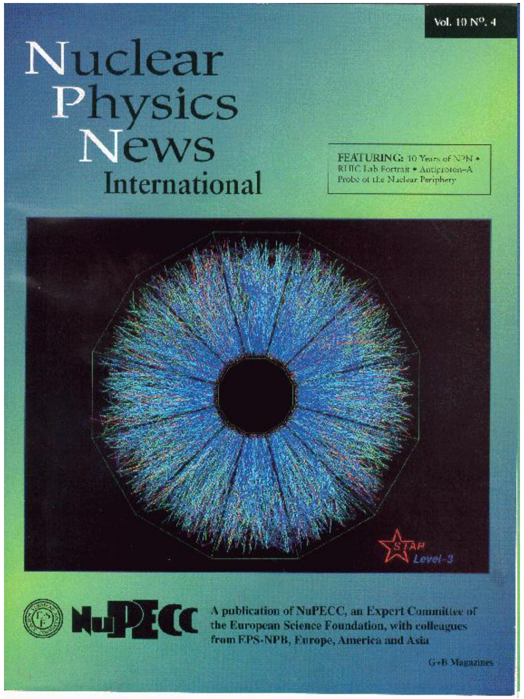
1st RHIC Collisions, July 2000

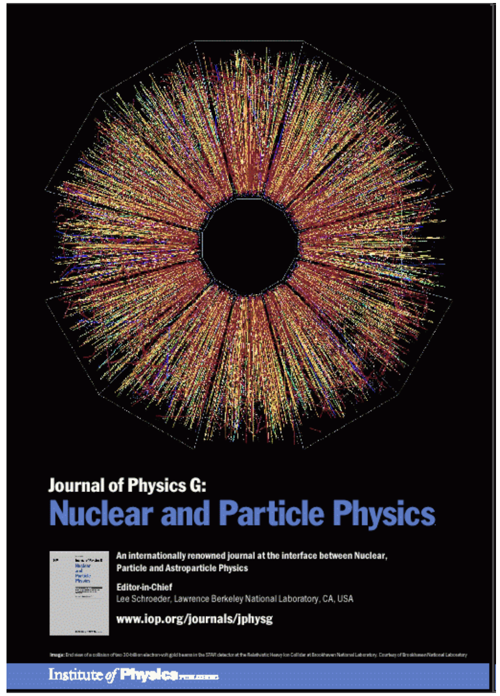
The STAR Experiment

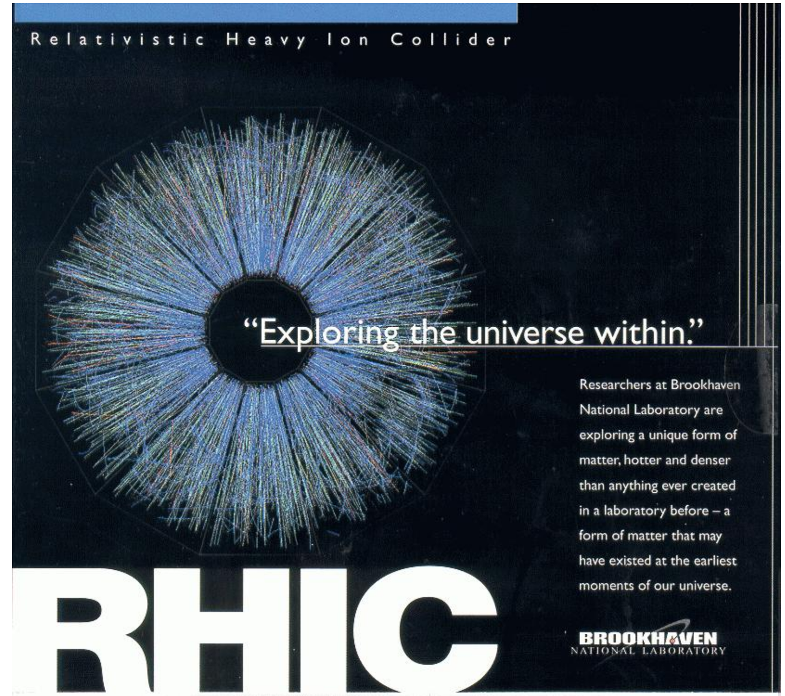
24 sectors *x* 5692 rϕ pads *x* 350 *t* bins
= 47,812,800 pixels






Vol. 10 N°. 4
Nuclear Physics News International
FEATURING: 10 Years of NPN • RHIC Lab Portrait • Antiproton-A Probe of the Nuclear Periphery
STAR Level-3
NuPECC
A publication of NuPECC, an Expert Committee of the European Science Foundation, with colleagues from EPS-NPB, Europe, America and Asia
G+B Magazines


Journal of Physics G:
Nuclear and Particle Physics
An internationally renowned journal at the interface between Nuclear, Particle and Astroparticle Physics
Editor-in-Chief
Lee Schroeder, Lawrence Berkeley National Laboratory, CA, USA
www.iop.org/journals/jphysg
Institute of Physics


Relativistic Heavy Ion Collider
"Exploring the universe within."
Researchers at Brookhaven National Laboratory are exploring a unique form of matter, hotter and denser than anything ever created in a laboratory before – a form of matter that may have existed at the earliest moments of our universe.
RHIC
BROOKHAVEN NATIONAL LABORATORY

ELSEVIER

Nuclear Instruments and Methods in Physics Research A 453 (2000) 397–404

NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A

www.elsevier.nl/locate/nima

The STAR level-3 trigger system

J.S. Lange[a,*], C. Adler[a], J. Berger[a], M. Demello[b], D. Flierl[a], J. Landgraf[c], M.J. LeVine[c], A. Ljubicic Jr.[c], J. Nelson[d], D. Roehrich[e], J.J. Schambach[f], D. Schmischke[a], M.W. Schulz[g], R. Stock[a], C. Struck[a], P. Yepes[b]

[a] *University of Frankfurt, August-Euler-Straße 6, D-60486 Frankfurt, Germany*
[b] *Rice University, Houston, Texas 77251, USA*
[c] *Brookhaven National Laboratory, Upton, New York 11973, USA*
[d] *University of Birmingham, Birmingham B15 2TT, United Kingdom*
[e] *University of Bergen, Allegaten 55, 5007 Bergen, Norway*
[f] *University of Texas, Austin, Texas 78712, USA*
[g] *University of Heidelberg, Philosophenweg 12, 69120 Heidelberg, Germany*

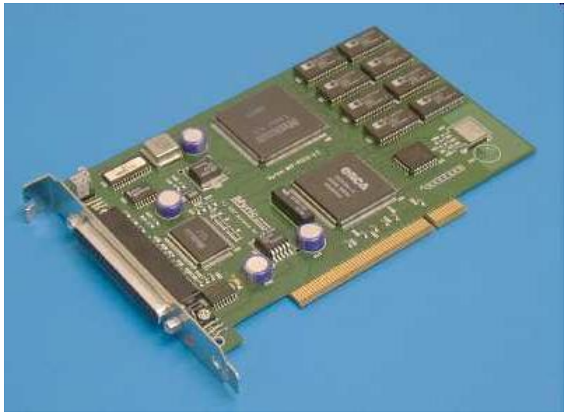
Accepted 20 June 2000

Abstract

The STAR level-3 trigger is a MYRINET interconnected ALPHA processor farm, performing online tracking of $N_{track} \geq 8000$ particles ($N_{point} \leq 45$ per track) with a design input rate of $R = 100$ Hz. A large-scale prototype system was tested in 12/99 with laser and cosmic particle events.

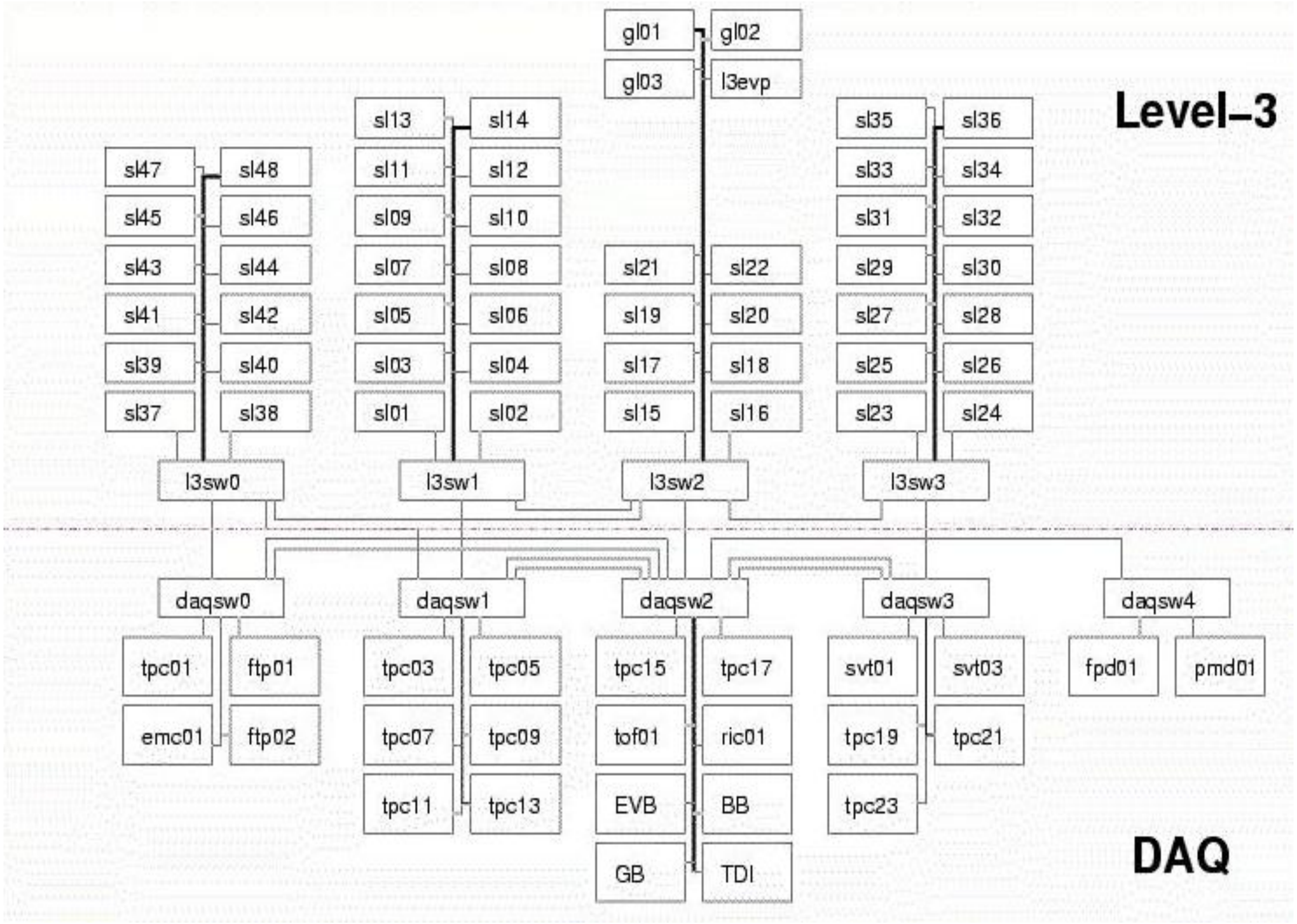
STAR Level-3 Trigger System

- realtime full event reconstruction (dE/dx, invariant mass)
- TPC cluster finder on 432 Intel i960 (VxWorks)
- TPC track finder on 48 ALPHA DS-10 (Linux)
- $t \leq 120$ ms
  per 1 central Au+Au
  N_{Track}=4,500
  $N_{Cluster}$=130,000
- MYRINET
  DMA
  $\simeq$74 Mbytes/s
  latency $\leq 50\ \mu$ s
- Programming in C/C++ and partially assembler for sqrt()

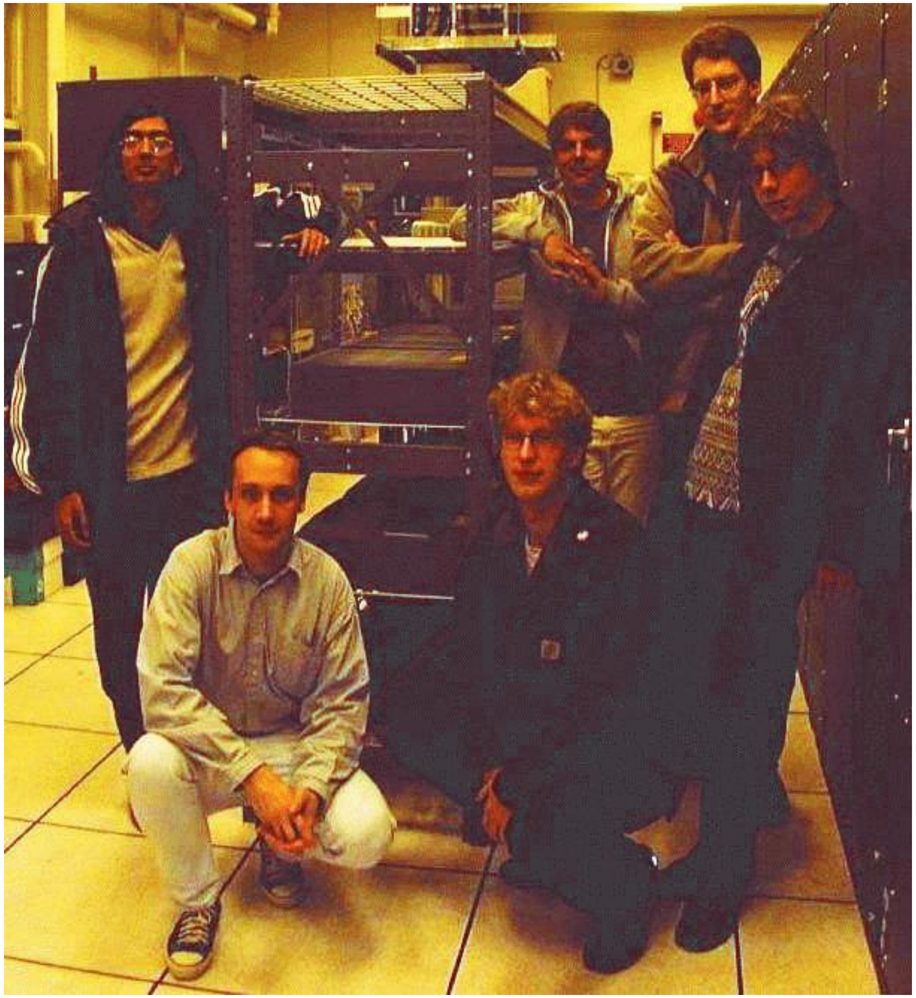


- **trigger rare events:**
  - **high-p_T anti-protons**
  - **Upsilon $\rightarrow e^+e^-$**
  - **Anti-^{3}He/^{4}He**

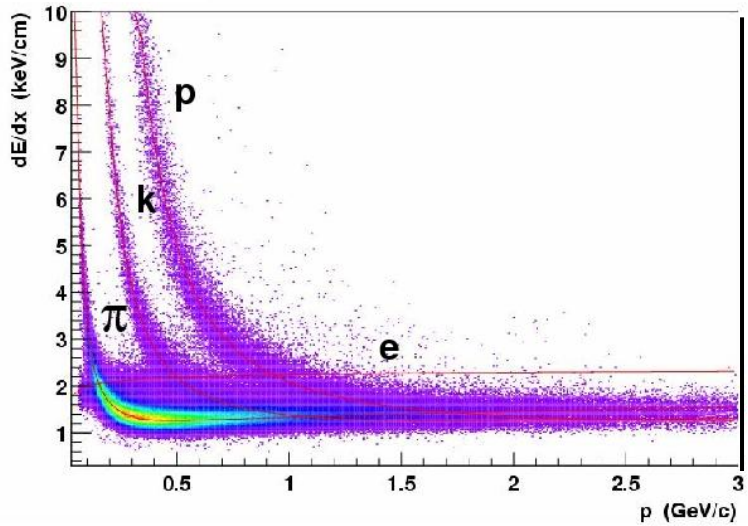
STAR Level-3 Trigger System

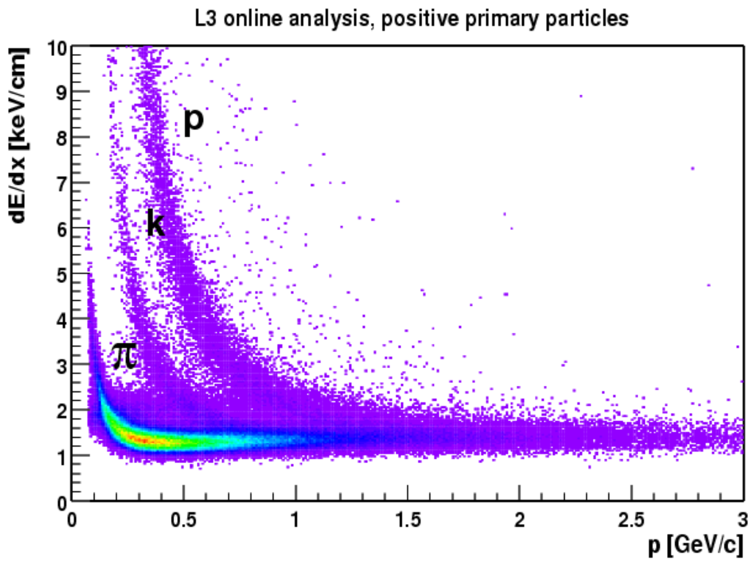


The STAR Level-3 Trigger System
University of Frankfurt, 1999-2004



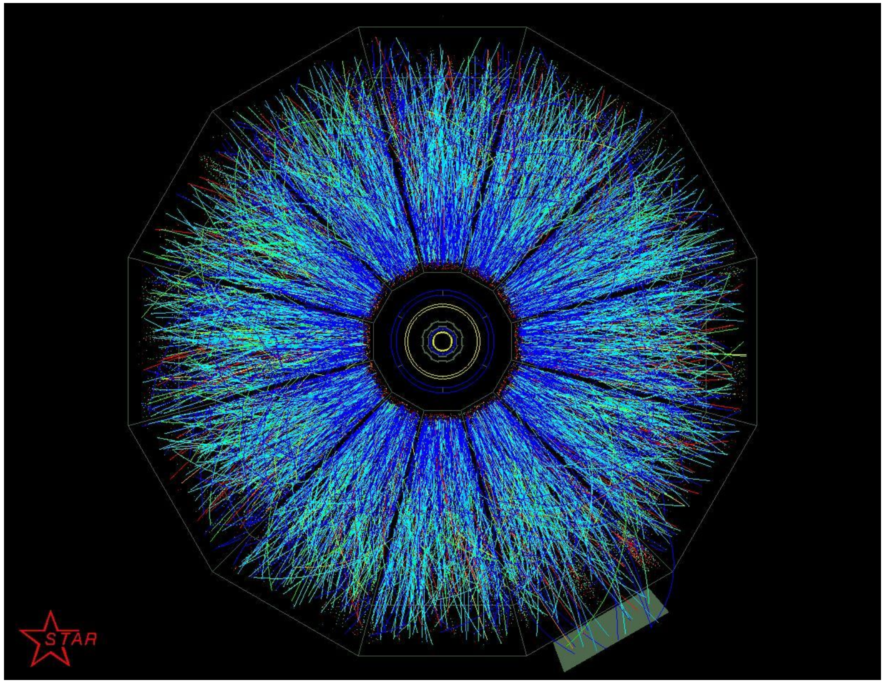
dE/dx offline analysis

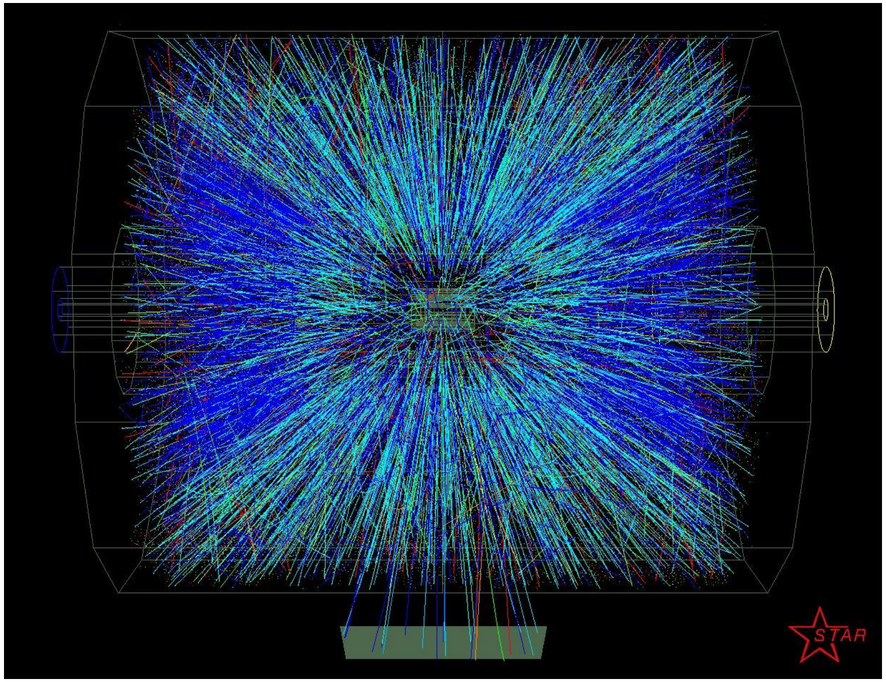




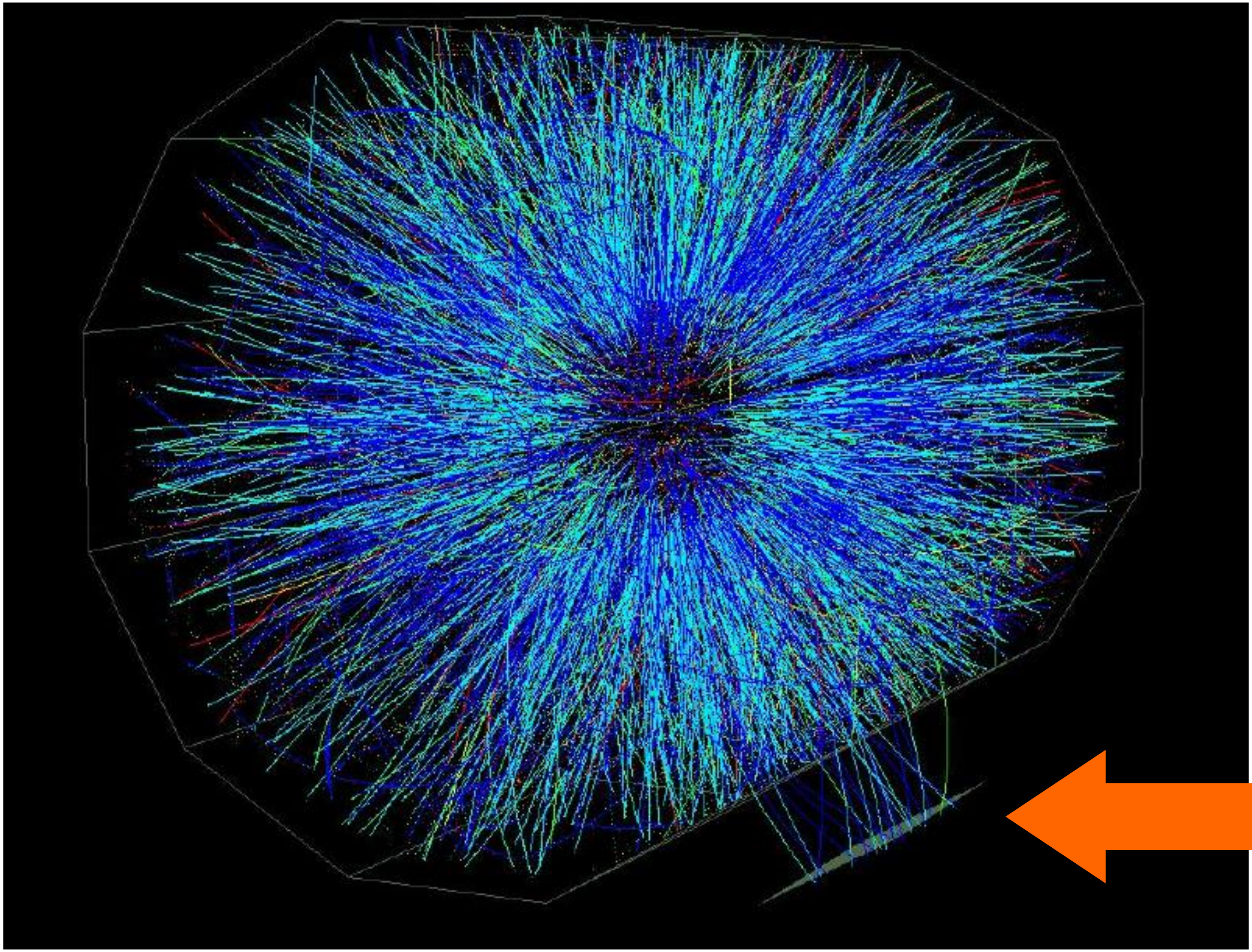
dE/dx online analysis

Pid in 200 ms with L3 online analysis





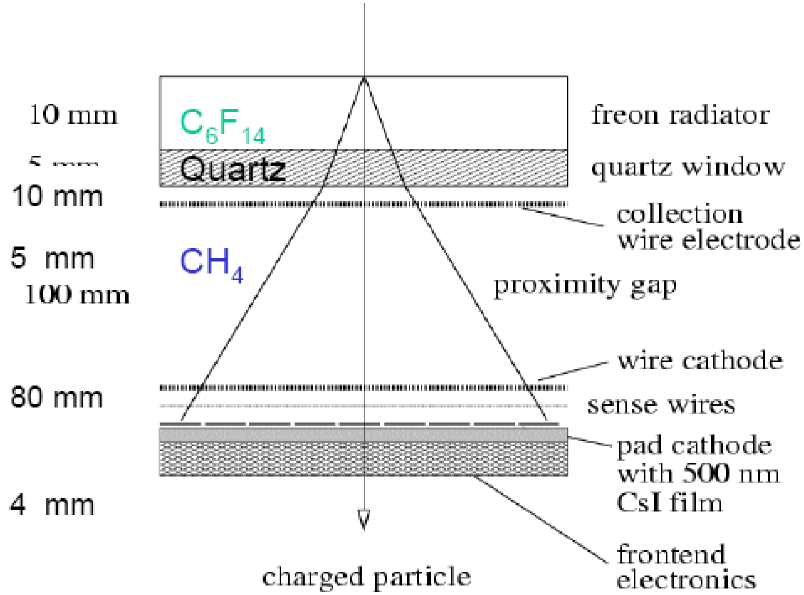
Au+Au collision at $\sqrt{s}$=200 GeV

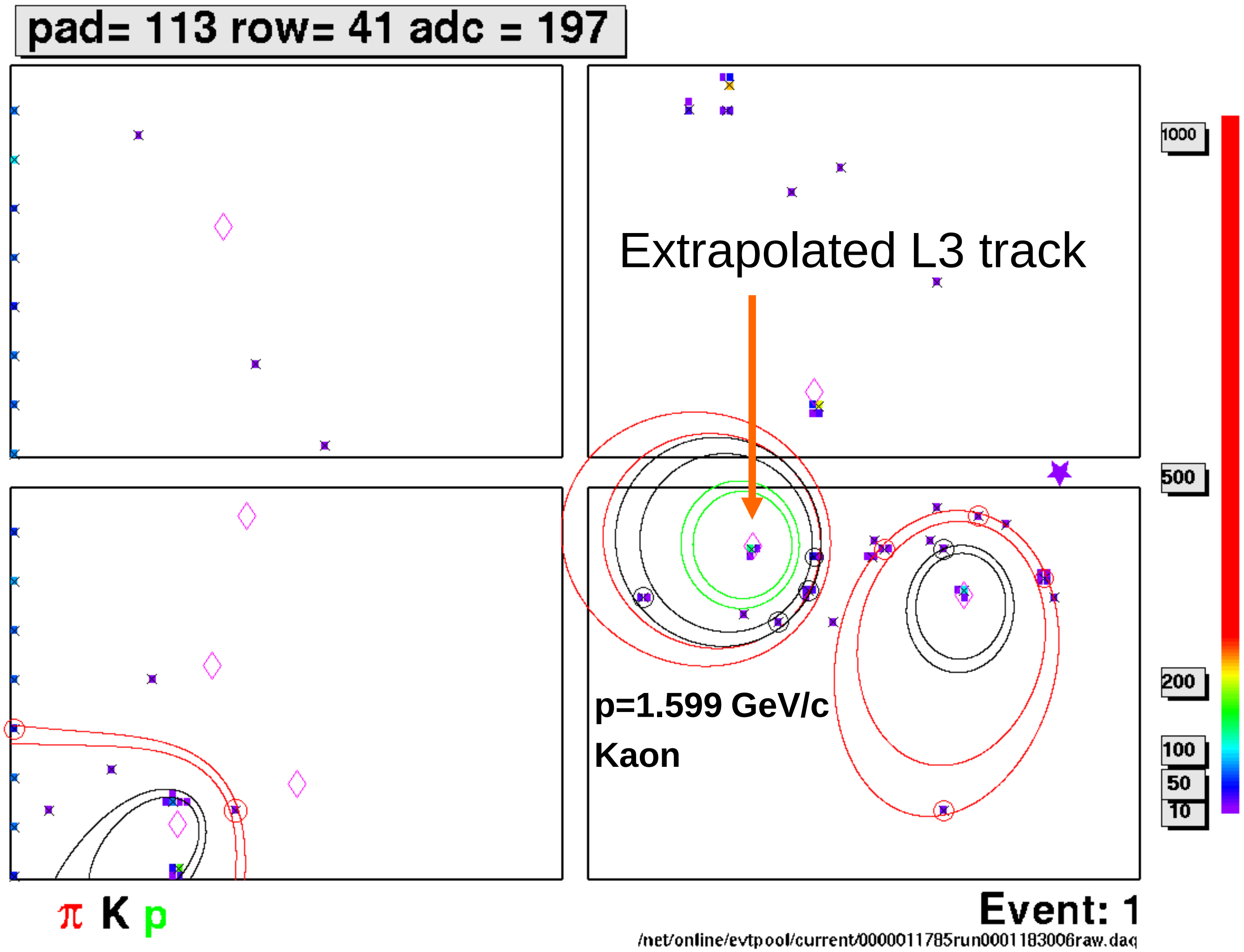


?

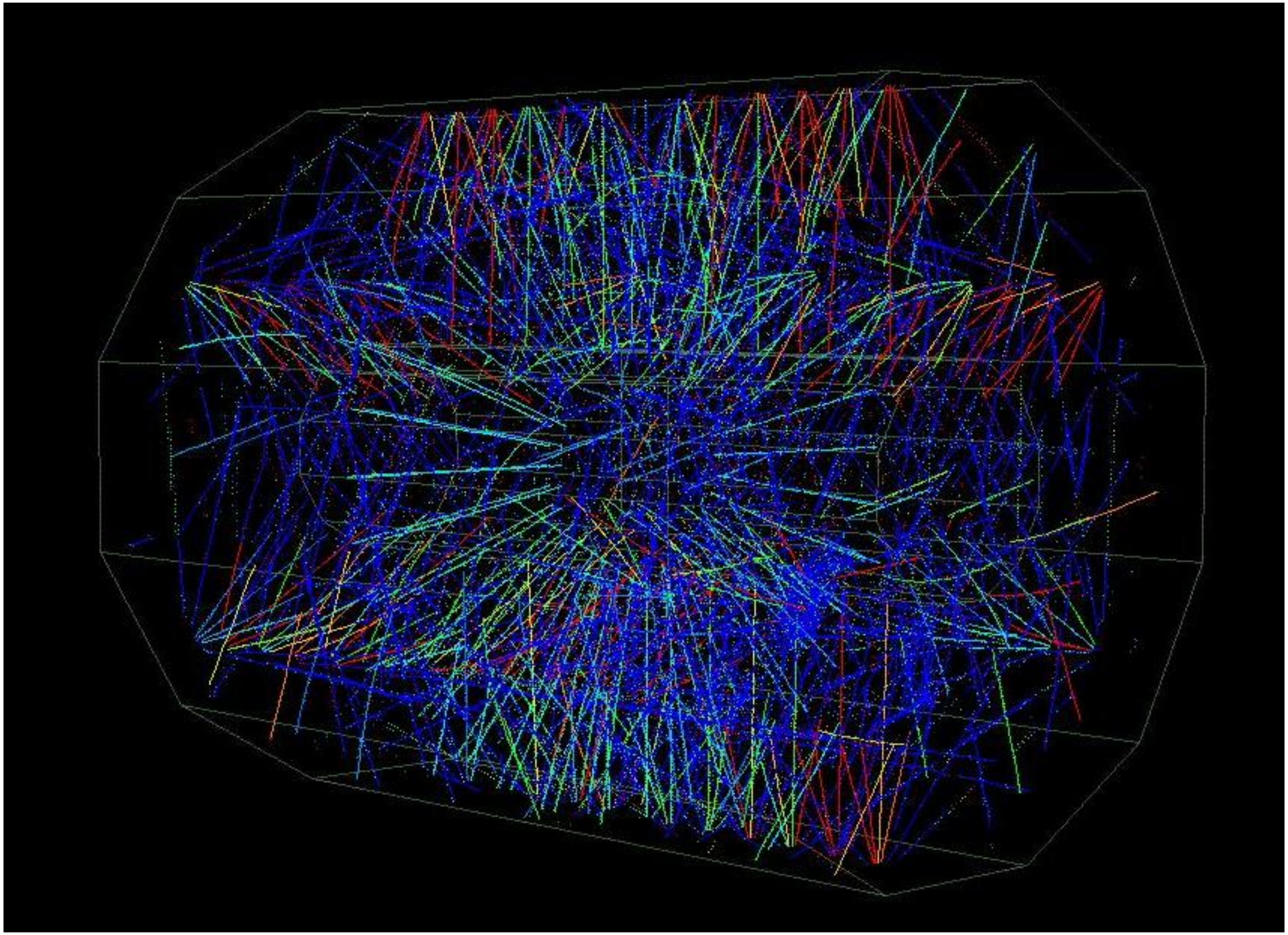
The STAR RICH Detector (1999-2001)

- Area 1 m^2
- C_6 F_{14} Liquid Radiator
- CsI Photo Cathode (d=0.3 μm)
- MWPC, 15,360 pads, 8.0x8.4 mm^2
- K/π identification for p<3 GeV/c
  proton/anti-proton indentification
  for p<5 GeV/c
- required Level-3 Trigger
  track extrapolation
  for high p_T tracks (p_T>3 GeV/c)
  because of small acceptance

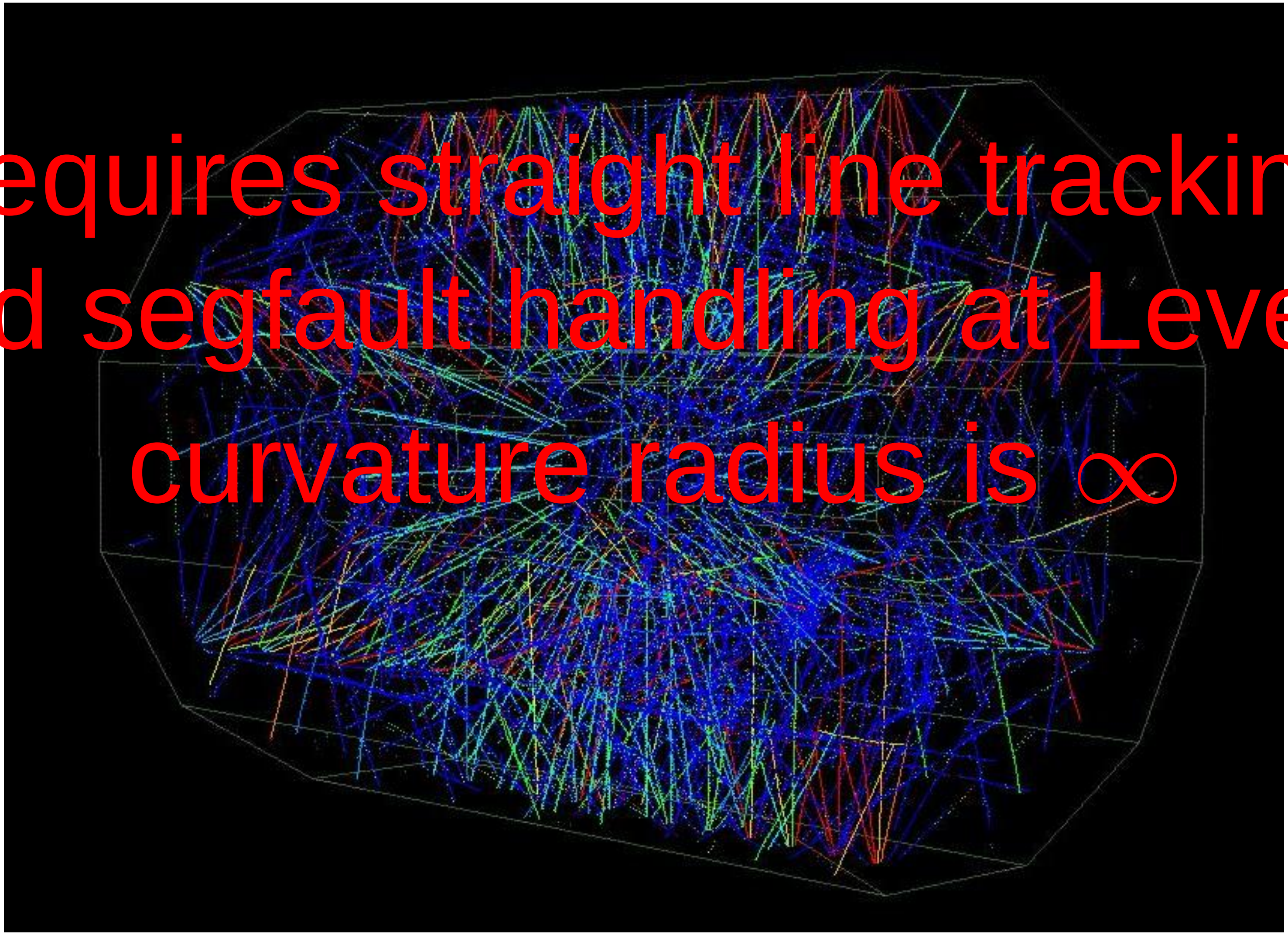



pad= 113 row= 41 adc = 197
Extrapolated L3 track
p=1.599 GeV/c
Kaon
π K p
Event: 1
/net/online/evtpool/current/0000011785run0001183006raw.daq
1000
500
200
100
50
10

Laser Events for TPC Drift Velocity Measurement

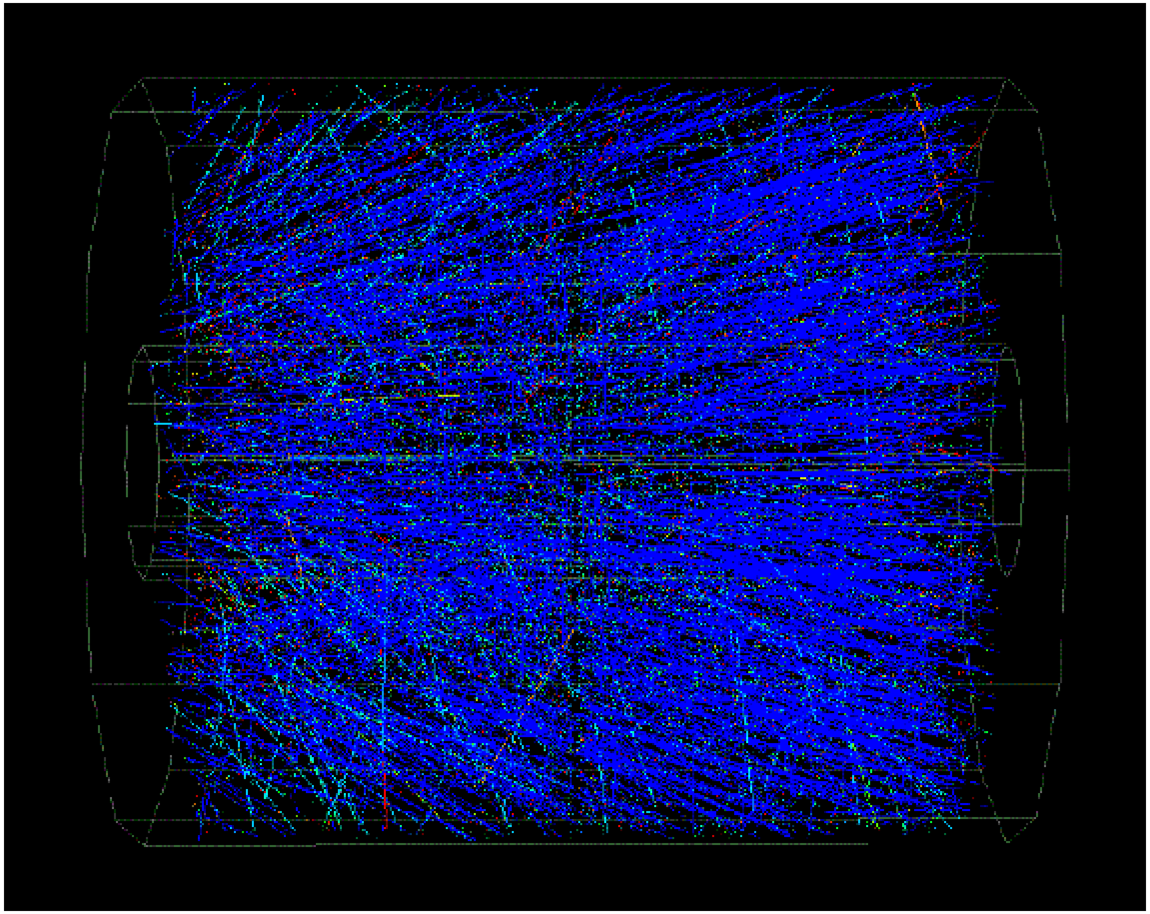


Laser Events for TPC Drift Velocity Measurement



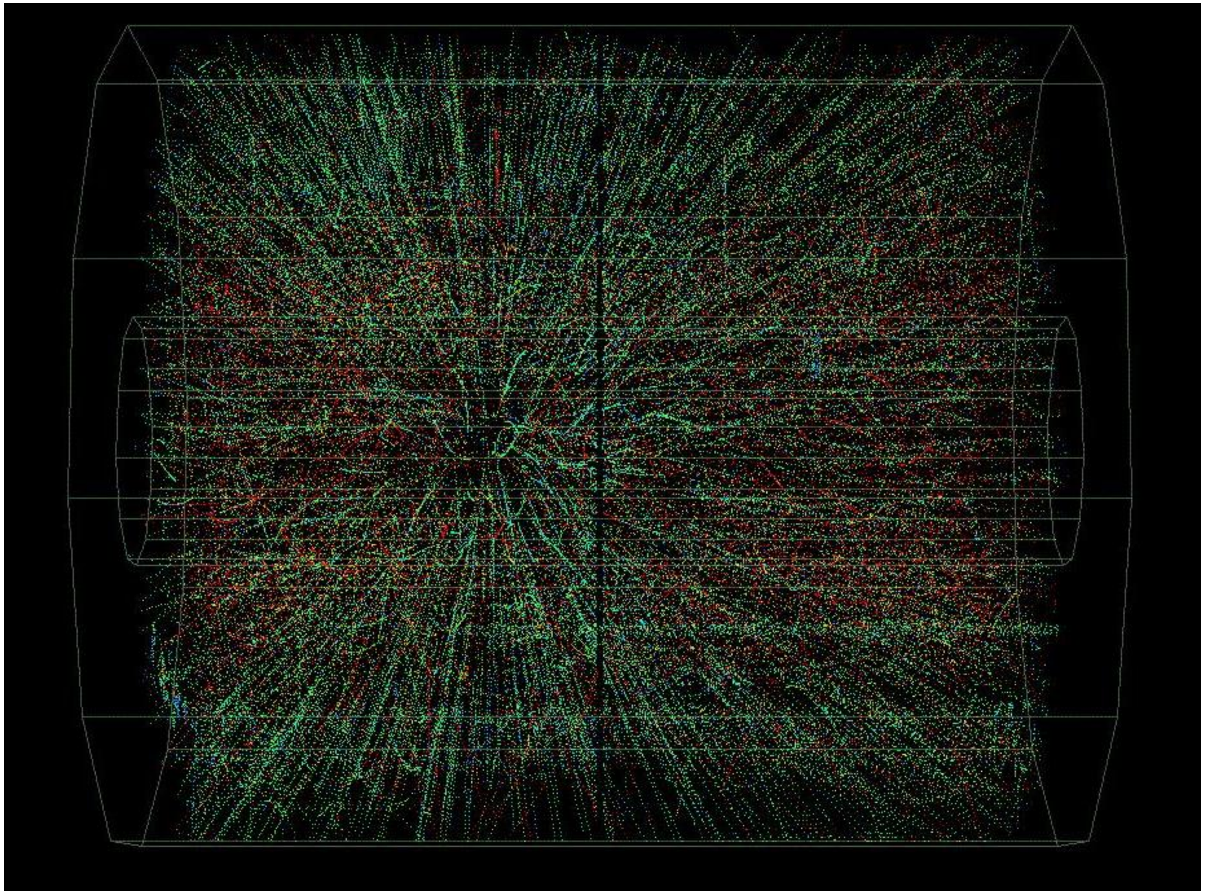
requires straight line tracking
and segfault handling at Level-3
curvature radius is ∞

TPC Gap Events, Example #1

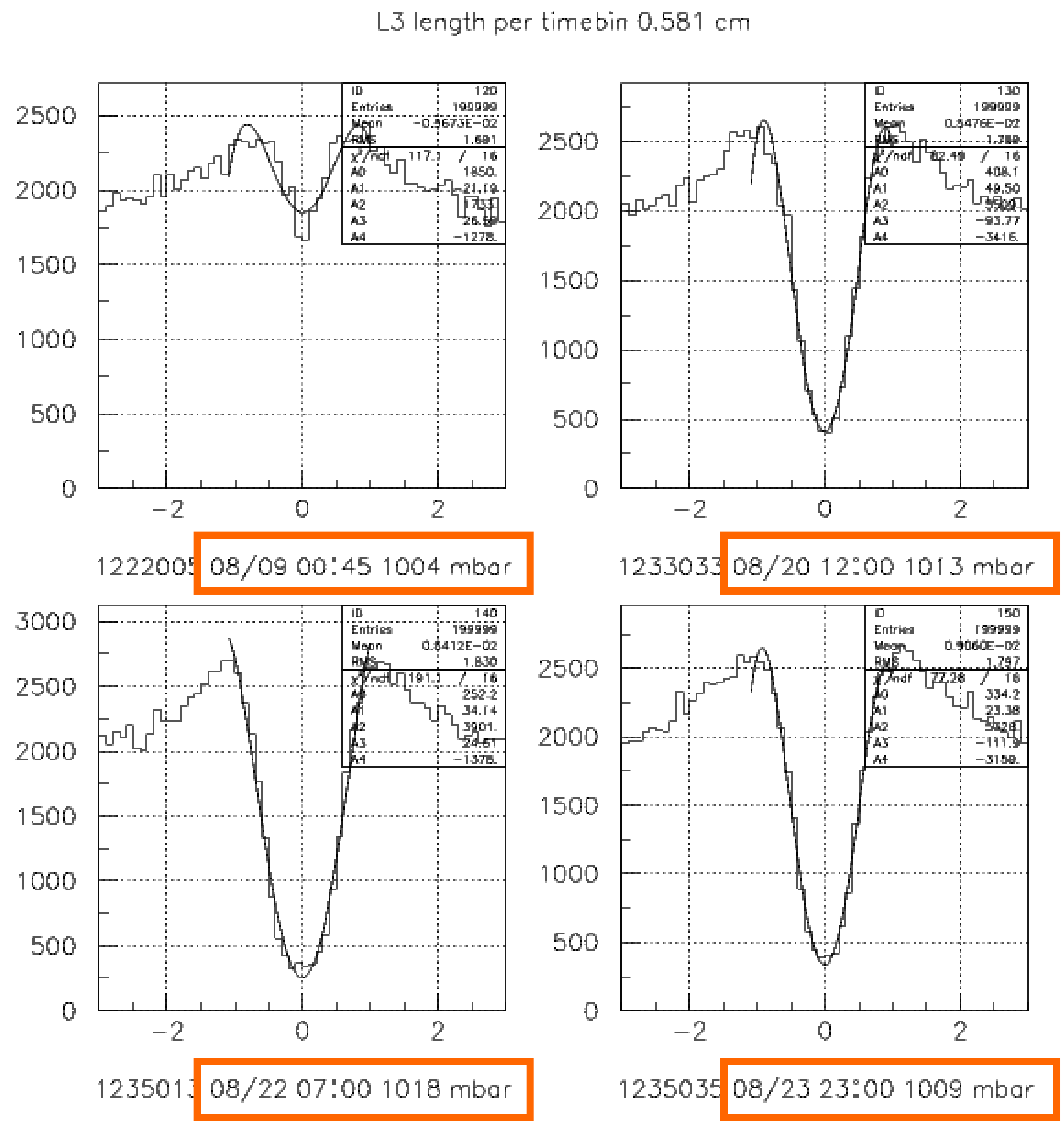


Beam pipe interaction at z=−4 m.
Gap at z=0 is visible, because triggered for different beam bunch crossing

TPC Gap Events, Example #2

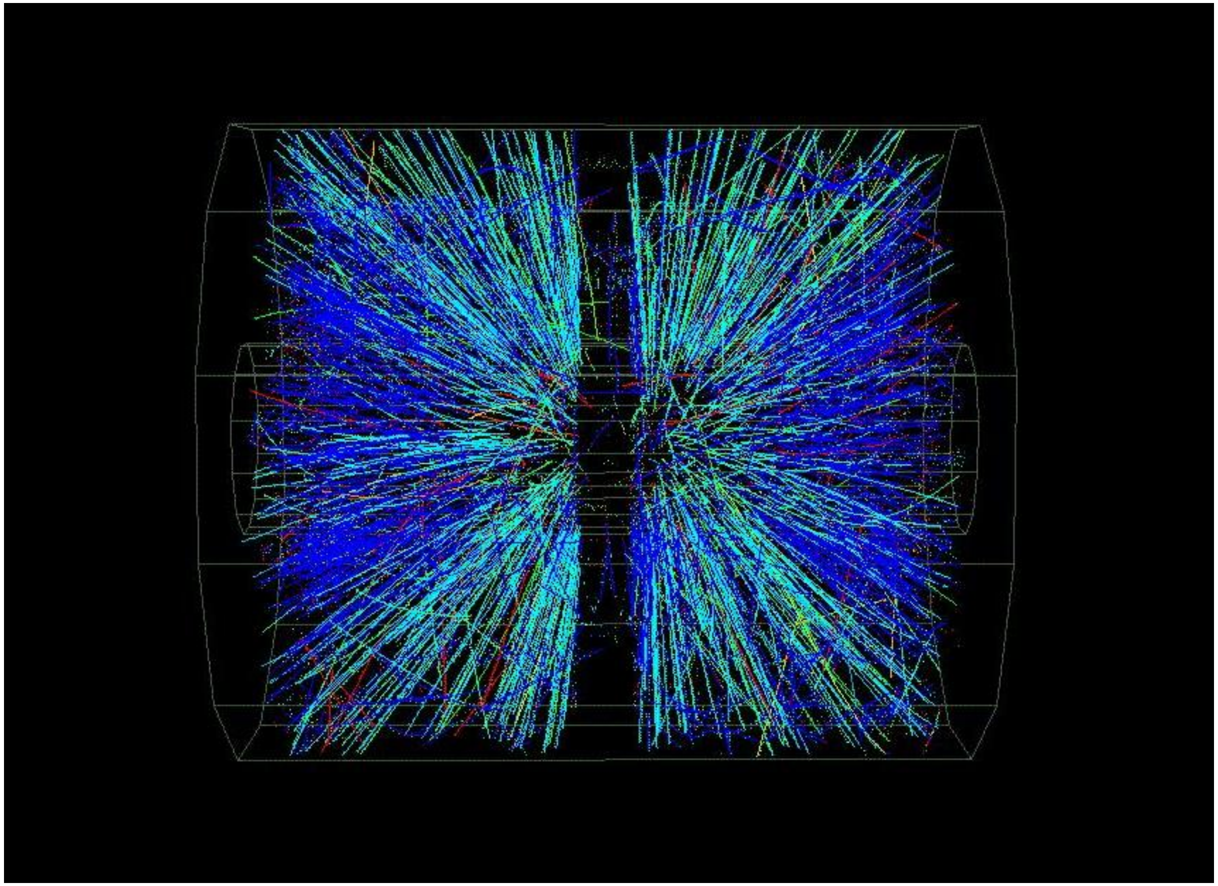


Gap @ z=0 due to v_{drift} change
i.e. weather change (atmospheric pressure p=1004 $\rightarrow$ p=1018 mbar)



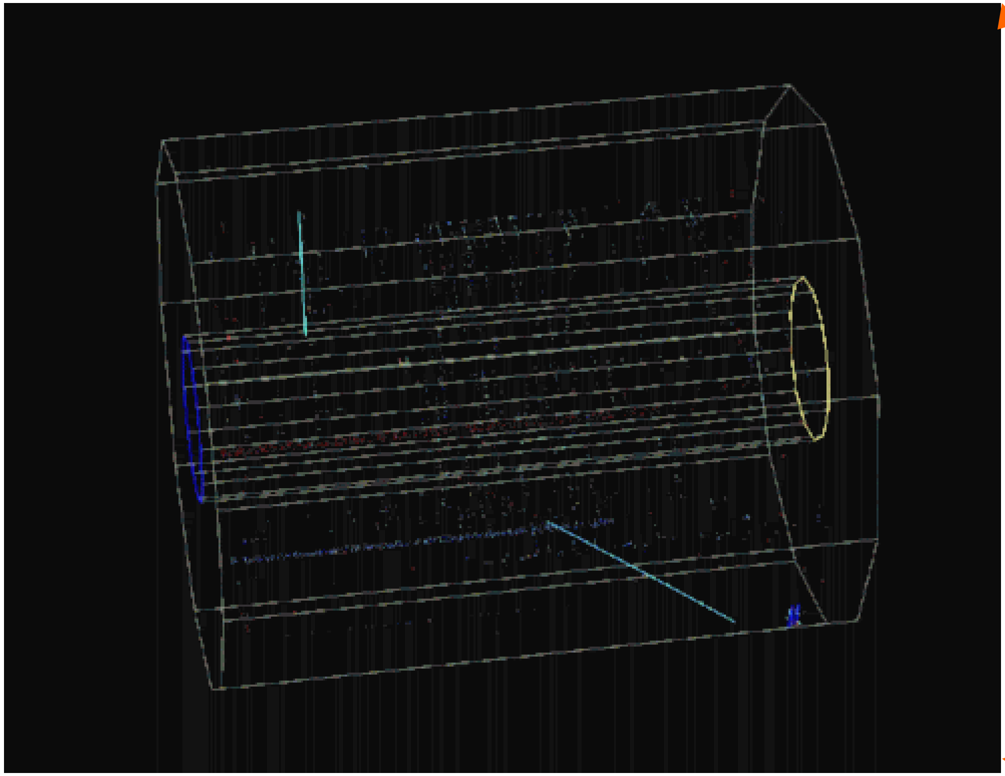
L3 length per timebin 0.581 cm
1222005 08/09 00:45 1004 mbar
1233033 08/20 12:00 1013 mbar
1235013 08/22 07:00 1018 mbar
1235035 08/23 23:00 1009 mbar

TPC Gap Events, Example #3



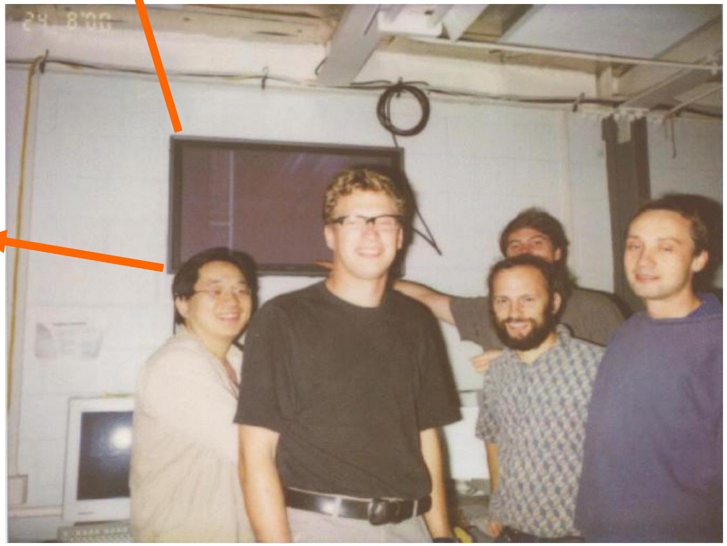
pile-up event
high multiplicity event + peripheral collision event ($t \simeq 5$ μs earlier)

Ultra-Peripheral Au+Au Collision



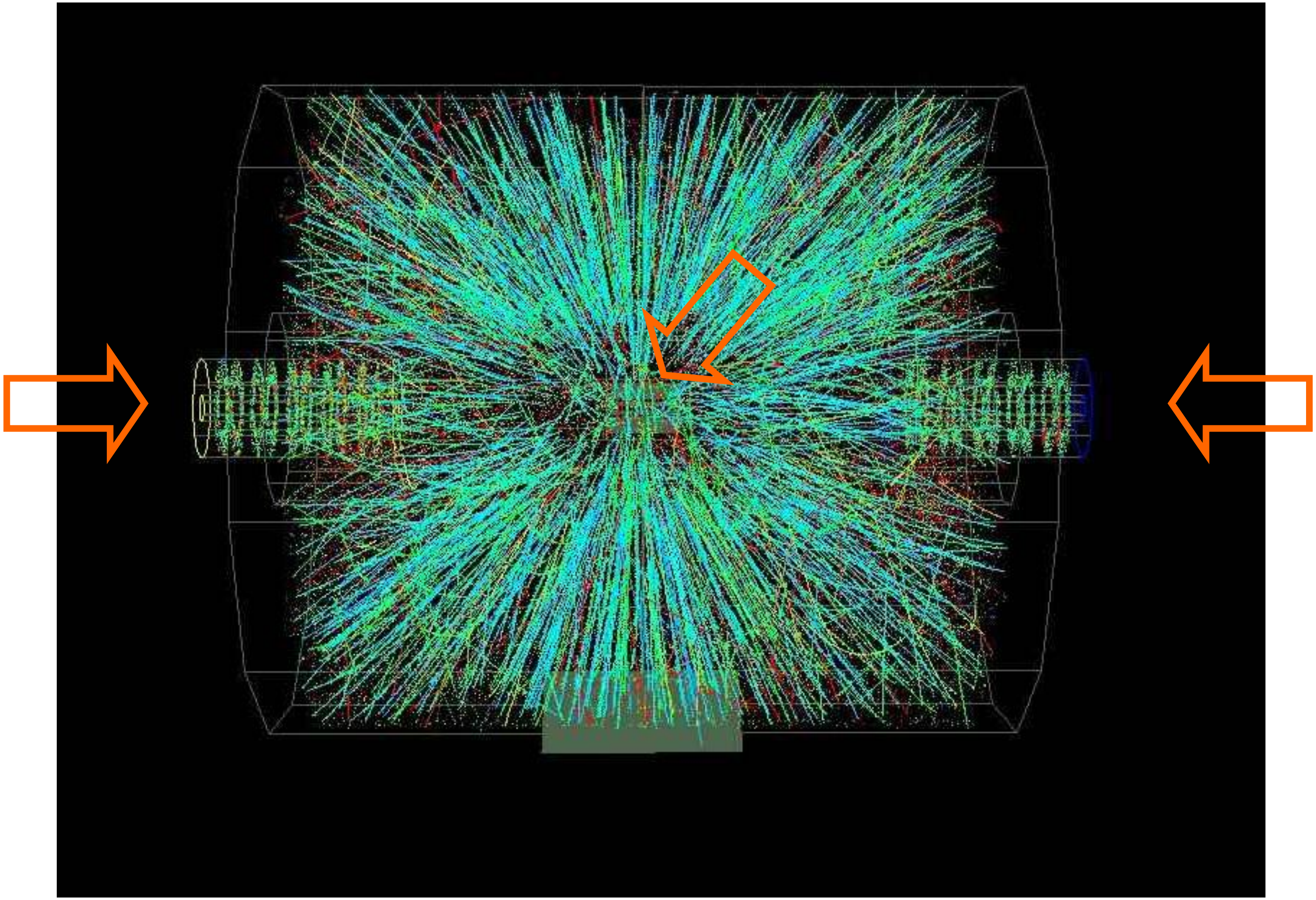
Au + Au -> Au* + Au* + ρ^o

Level-3 Trigger
on $\rho^0 \to \pi^+ \pi^-$
invariant mass



1st event 08.24.2000

More Detectors used in Level-3 Reconstruction: Forward TPC and Silicon Drift Detector (2003)

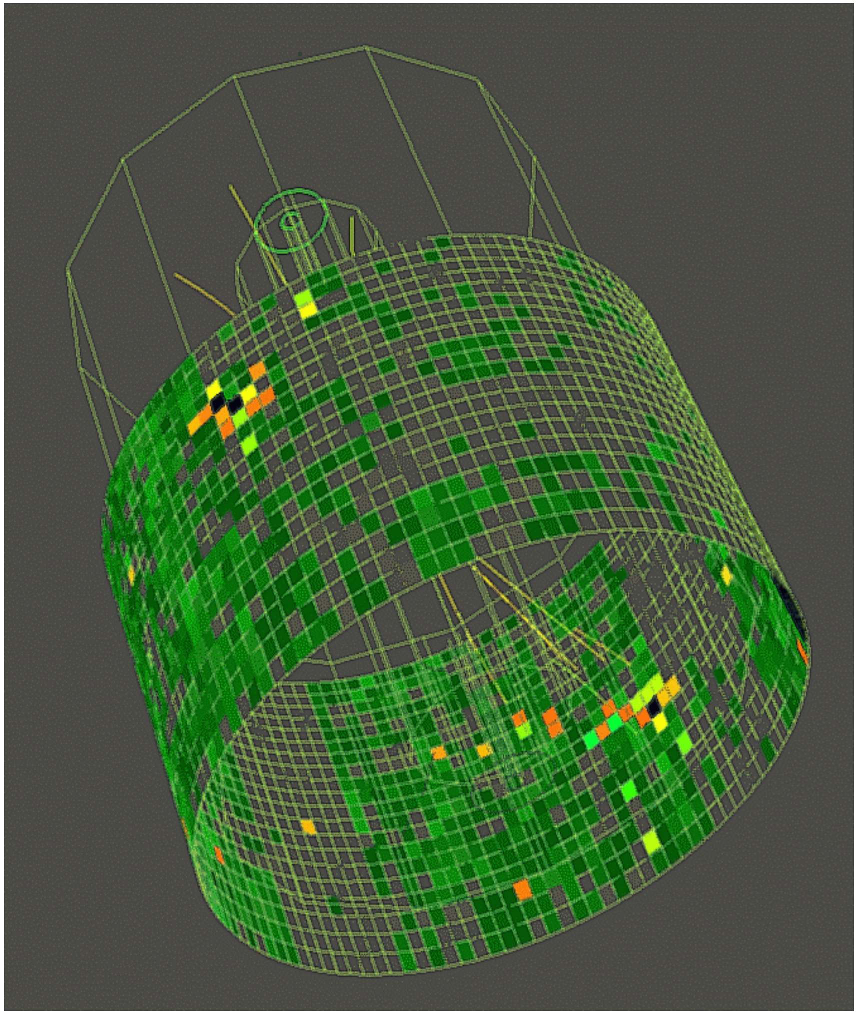


More Detectors used in L3 Reconstruction: Electromagnetic Calorimeter (2004)

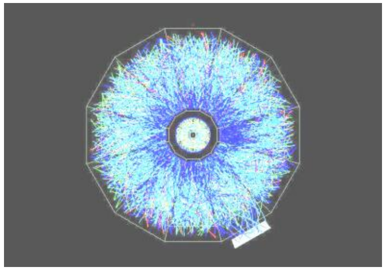
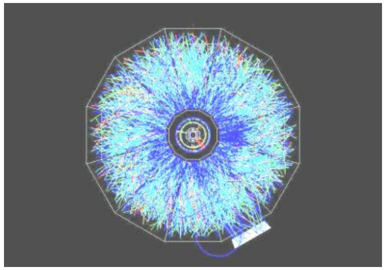
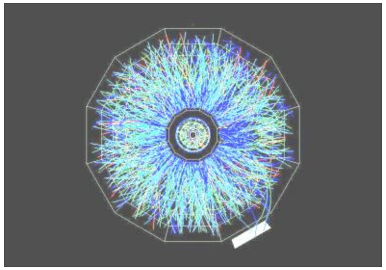
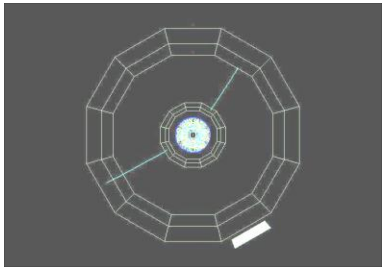
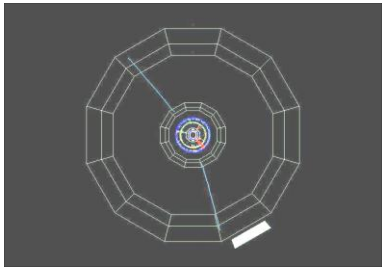
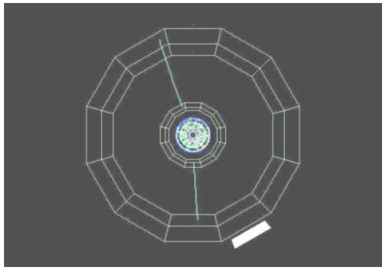
All Calorimeter ADC values send to Level-3 Trigger

matching with extrapolated Level-3 tracks implemented

pp, $\sqrt{s}$=200 GeV
2-jet event
all tracks p_T>1 GeV/c



Upsilon(1s,2s,3s)→e⁺e⁻ candidates with the Level-3 Trigger

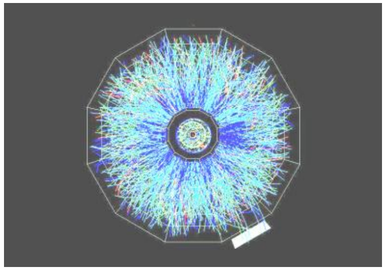
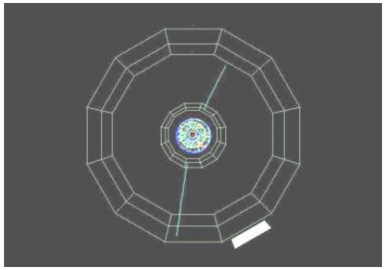







Upsilon->l+l- candidates

Run 2241022
Event #4699
m=10.21 GeV

Run 2243027
Event #1645
m=9.68 GeV

Run 2244029
Event#1570
m=9.25 GeV

$\pi^+ \pi^-$ candidate

Run 2244032 Event #3420
m=9.38 GeV

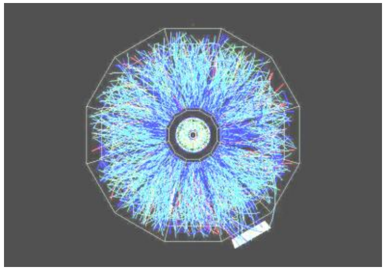
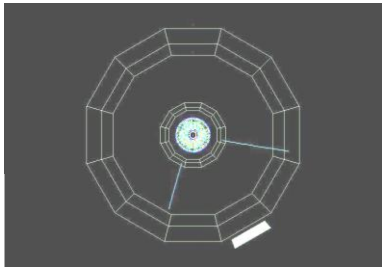
dE/dx<1.4e-5 keV/cm
P_{PID}(lepton)<3%

Au+Au
$\sqrt{s_{NN}}$=200 GeV

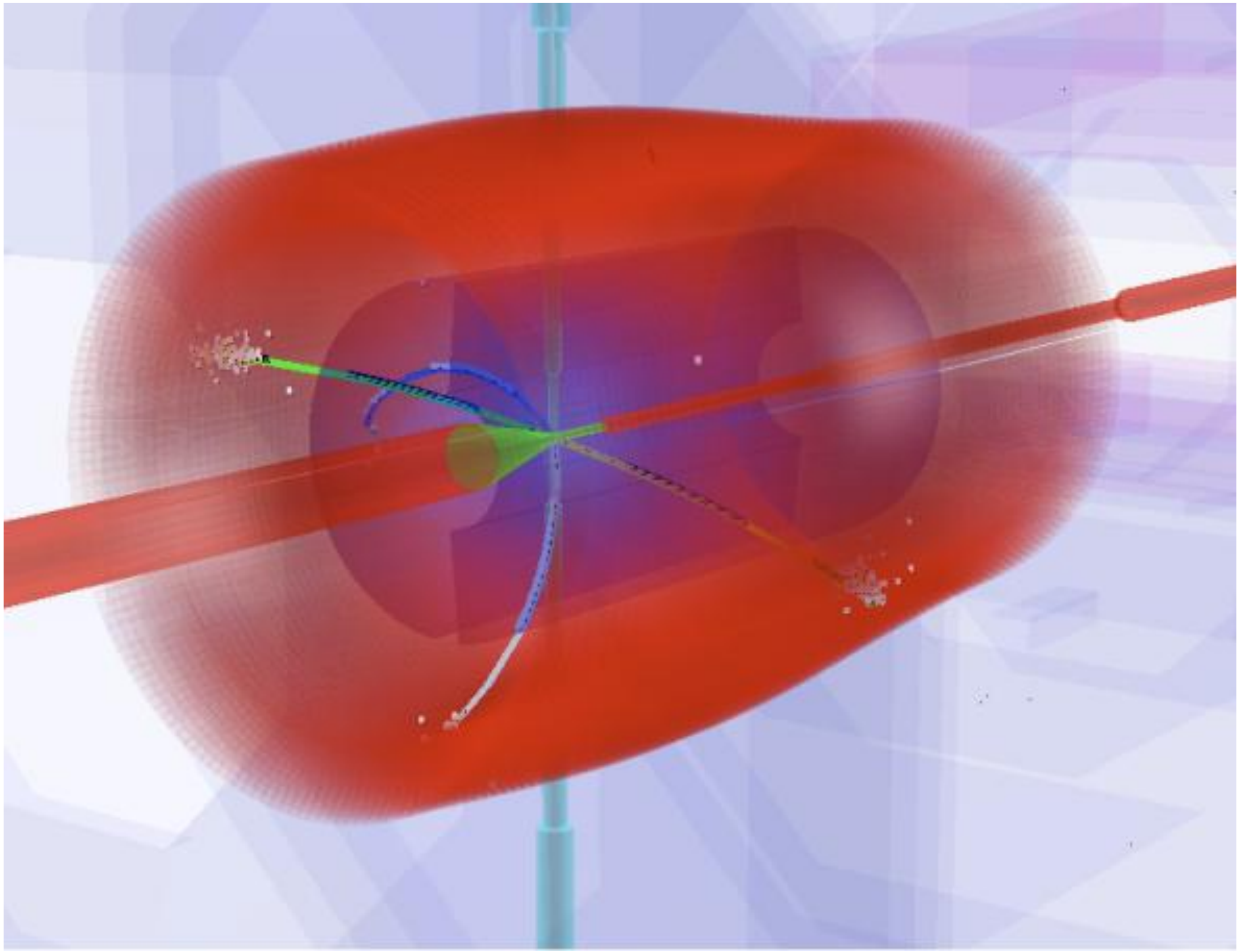
Drell-Yan candidate

Run 2252020 Event #8718
m=10.02 GeV

$\Delta\Phi$ = 78°

Example for Panda (PandaRoot framework) Event Display
$X(3872) \to J/\psi\ \pi^+\ \pi^-$ Event



A new approach:

GPU´s

(Graphics Processing Units

for Co-Processing of Track Fitting)

M. Al-Turany,
F. Uhlig,
R. Karabowicz
(GSI)

GPU Test Bench for Panda: Hardware

- Nvidia C1060 Graphics Adapter
  - # of cores: 240
  - frequency 1.3GHz
  - performance (peak):
    single precision floating point
    933 GFLOPS
    double precision floating point
    78 GFLOPS
- PandaRoot 2.0 Framework
- Conformal Map Track Fitter
- Programming:
  CUDA
  Compute Unified Device Architecture
  for loading code to the GPU

Example for CUDA Code: Matrix Multiplication

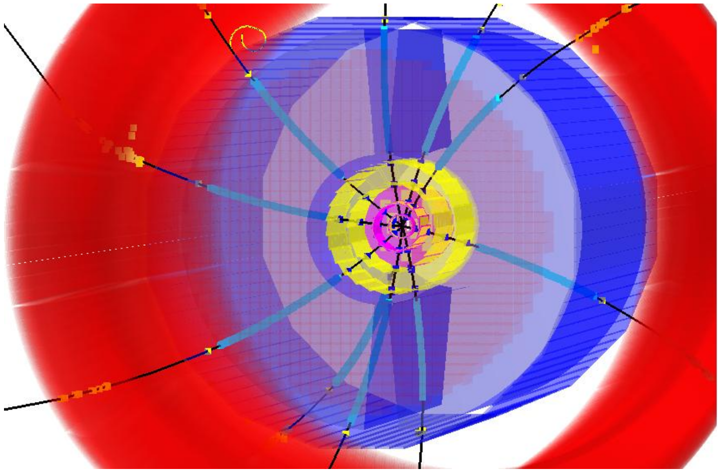
```
// Loop over all the sub-matrices of A and B required to
// compute the block sub-matrix
for (int a = aBegin, b = bBegin;
         a <= aEnd;
         a += aStep, b += bStep) {

    // Shared memory for the sub-matrix of A
    __shared__ float As[BLOCK_SIZE][BLOCK_SIZE];

    // Shared memory for the sub-matrix of B
    __shared__ float Bs[BLOCK_SIZE][BLOCK_SIZE];

    // Load the matrices from global memory to shared memory;
    // each thread loads one element of each matrix
    As[ty][tx] = A[a + wA * ty + tx];
    Bs[ty][tx] = B[b + wB * ty + tx];

    // Synchronize to make sure the matrices are loaded
    __syncthreads();

    // Multiply the two matrices together;
    // each thread computes one element
    // of the block sub-matrix
    for (int k = 0; k < BLOCK_SIZE; ++k)
        Csub += As[ty][k] * Bs[k][tx];

    // Synchronize to make sure that the preceding
    // computation is done before loading two new
    // sub-matrices of A and B in the next iteration
    __syncthreads();
}
```

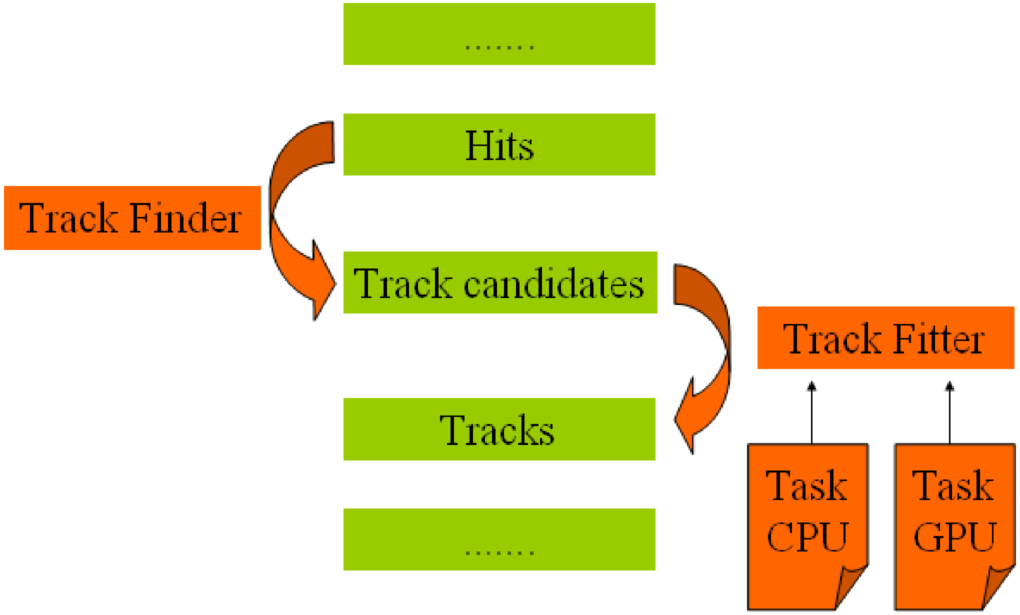
GPU Test for Panda: Software

M. Al-Turany,
F. Uhlig,
R. Karabowicz
(GSI)

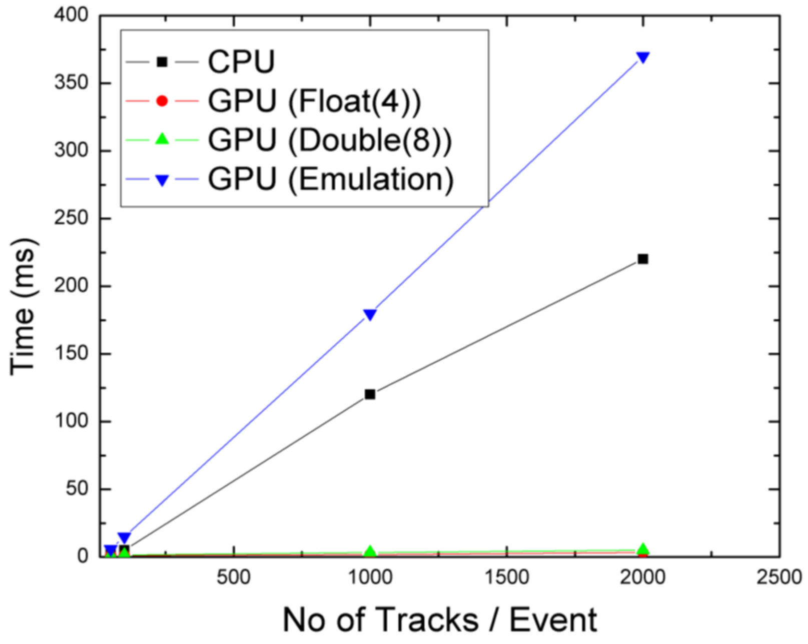
- PandaRoot 2.0 Framework: MVD, TPC, EMC, MDT detectors
- p=1 GeV/c μs
- 50 events for each sample
- 4 samples
  N_{track}=50, 100, 1000, 2000
  primary tracks per event
- ideal track finder
  (copy GEANT tracks
  to the track candidates objects)

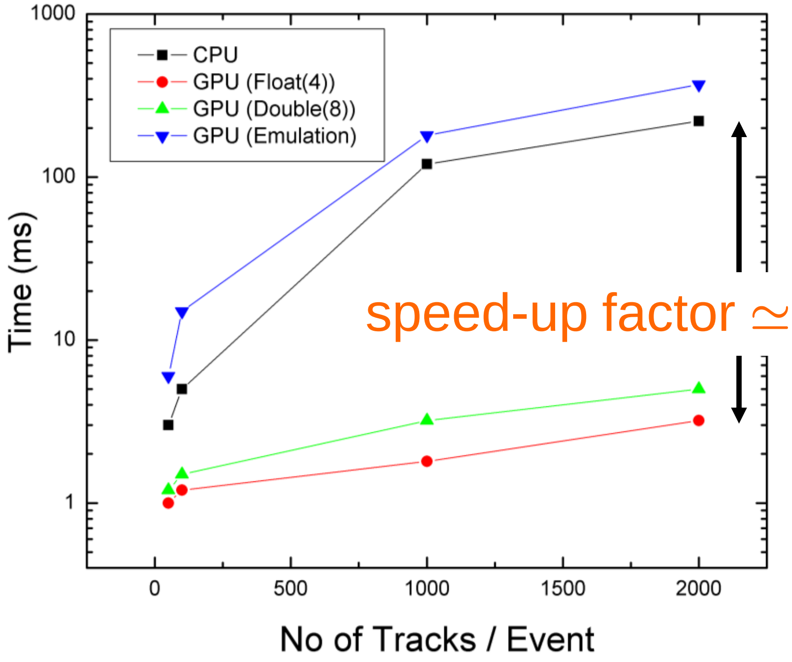
hit data of track candidates are written to shared memory in 32 parallel threads





Track Fitting on CPU and GPU





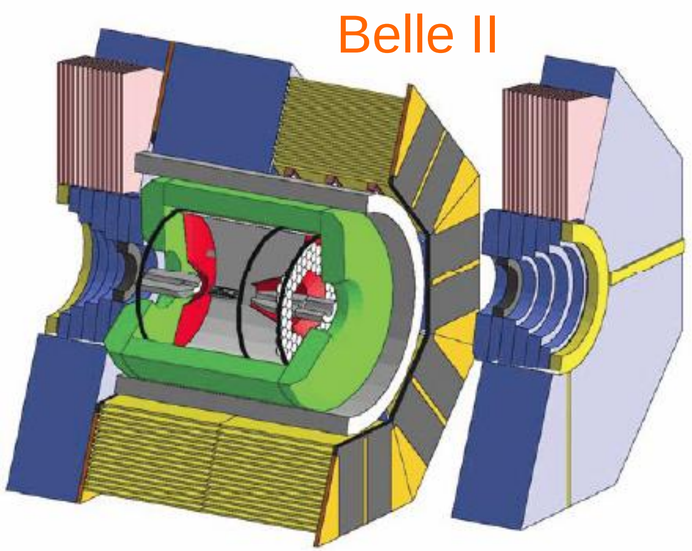
speed-up factor $\simeq$ 68

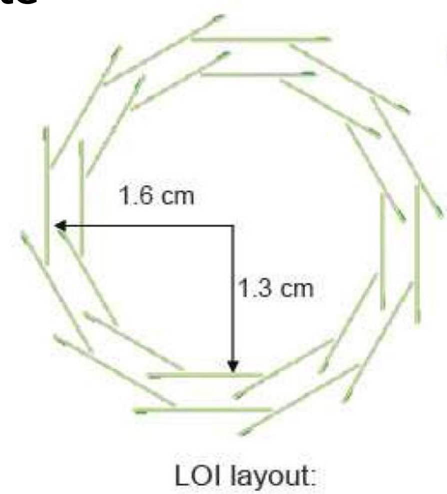
	50	100	1000	2000
GPU (Emu)	6.0	15.0	180	370
CPU	3.0	5.0	120	220
GPU (D)	1.2	1.5	3.2	5.0
GPU (F)	1.0	1.2	1.8	3.2

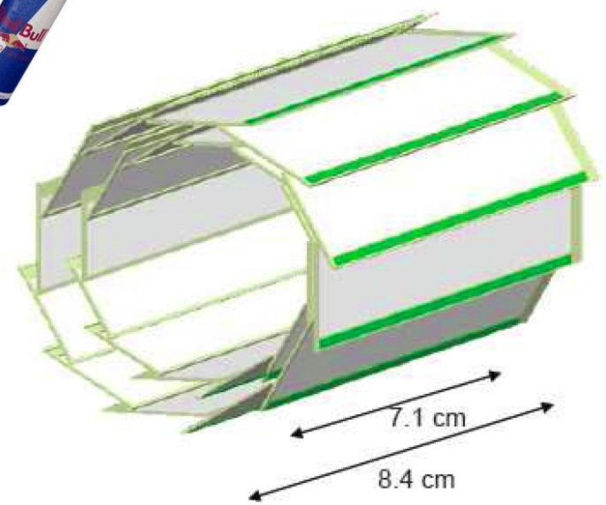
ms

Where can we train for the Panda Online Computing?

- Belle II DEPFET Pixel Detector
  (MPI Munich, Göttingen, Giessen, Bonn, Heidelberg, Karlsruhe)
- readout system:
  1 ATCA Crate with 14 Compute Nodes
- raw data rate
  $\simeq$ 27 Gbytes/s
  = 7 DVD's per 1 s
  $\simeq$ 1/20 of Panda raw data rate
  nice testing ground
- online track finder
  4 SVD hits + 2 pixel hits
- data reduction
  factor 10-100
  required







Summary

$3 \cdot 10^7$ events/s raw data

↓ Panda Online Computing

$25 \cdot 10^3$ events/s to tape

- Requires high bandwidth and high speed computing
  $\rightarrow$ Compute Nodes with ATCA and Virtex-4 FPGA´s
- Online Computing is also required for
  event display, calibration (e.g. TPC drift velocity),
  monitoring of detector stability (e.g. temperature)
  bunch timing, background conditions, pile-up events
  etc. etc.
- GPU´s under testing (maybe for Level-4)
- Lots of work, but also lots of fun ahead of us …