New Data Storage Model for H1

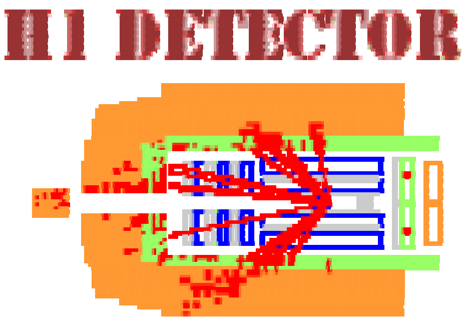
Thomas Benisch, DESY

CHEP 2000
February 7−11, 2000
Padova, Italy

- Introduction
- Shortcomings of the present Data Storage Model
- New Data Storage Model
- Benchmark Results
- Conclusions

Introduction (1)

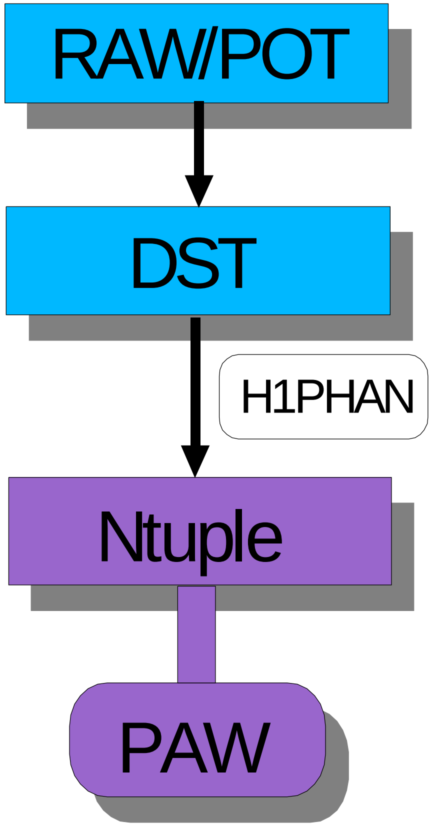
- H1 is an experiment at the ep collider HERA at DESY in Hamburg (Germany)
- H1 started collecting data in 1992
- H1 and HERA are performing substantial upgrades in the year 2000
- increased demands on data storage and data handling



Introduction (2)

- the H1 collaboration has decided to adopt the ROOT framework for
  - data storage } ⇒ *this talk*
  - physics analysis
  - event display } ⇒ *see talk of U. Berthon*

Present Data Storage Model



- RAW Raw Data
- POT Production Output Tape
  raw + reconstructed events
- DST Data Summary Tape
  reduced set of information for physics analysis

BOS/FPACK

ZEBRA

Data Structure and Data Access

- data storage is based on
  - memory management system BOS (Bank Organisation System)
  - I/O package FPACK
- Entity−Relationship model is used as a basis for data structures
- data are stored in so−called BOS banks
  - described by a data definition language (DDL)
  - contain closely related data
- I/O package FPACK
  - machine independent
  - common user interface for all operating systems
  - network server for remote file access

Shortcomings of the Present System

- The current H1 analysis environment is rather inhomogenous
  - data storage based on BOS/FPACK
  - physics analysis is done with PAW
  - event display uses graphics package LOOK
- events must always be read and written completely, reading of single variables is not possible
- event selection is done by using index files, based on a 32 bit classification word (predefined, static!)

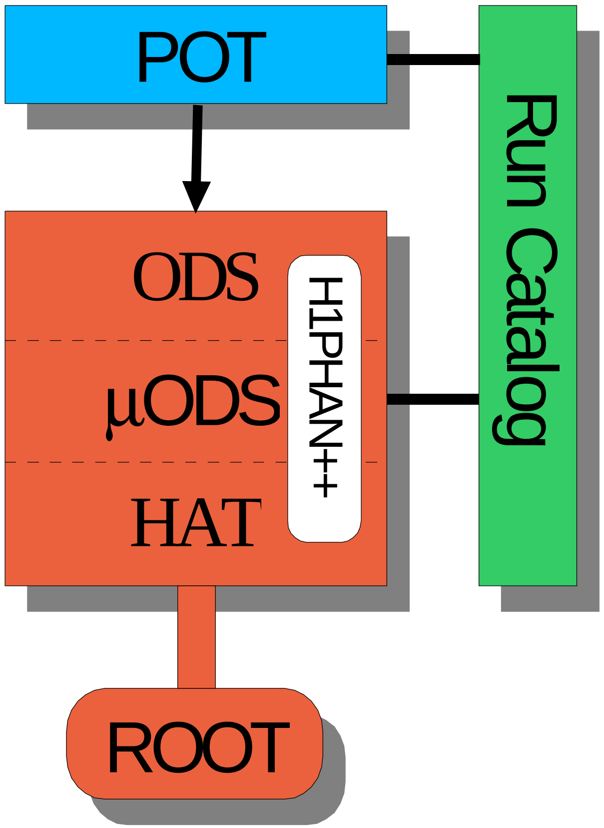
New Technology for Data Storage: ROOT

- increased demands on data handling after the upgrade in the year 2000
- H1 has chosen a new technology to improve data access: ROOT
- important features of ROOT I/O:
  - high I/O efficiency
  - sequential and direct data access techniques
  - possibility of splitting event data into several streams
  - access to part of the event data only
  - support of networking
  - built−in gzip−type compression algorithm

New Data Storage Model (1)

- event data are split into several parts according to frequency of their access
- multi−level hierarchical storage system
- higher level contains less data per event
- access to partial event data will be essential for many physics analyses

New Data Storage Model (2)

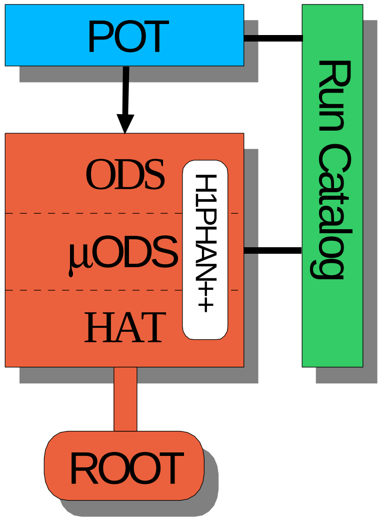


- POT   Production Output Tape
- ODS   Object Data Store
- µODS   micro Object Data Store
- HAT   H1 Analysis Tag

BOS/FPACK

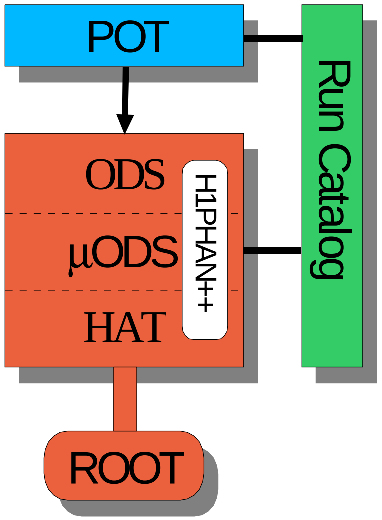
ROOT

New Data Storage Model (3)



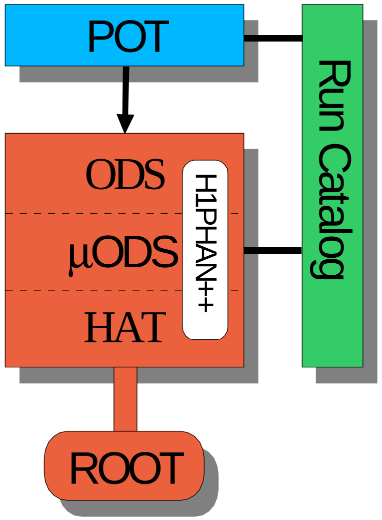
- POT   Production Output Tape   BOS/FPACK
  - contains raw and reconstructed data as now
  - no change to software producing the POTs necessary, e.g. Fortran based reconstruction program
- ODS   Object Data Store   ROOT
  - contains all standard objects for physics analysis, e.g. tracks, clusters, etc.
  - corresponds to current DSTs
  - existing event model is not changed
  - direct mapping of BOS banks into objects using BOS2OOP
  - each bank is written to a seperate branch
  - backward compatibility

New Data Storage Model (4)



- μODS   micro Object Data Store   ROOT
  - contains information on the particle level and their 4−vectors, e.g. electron candidates, jet properties, energy flow, PID probabilities, etc.
  - full expert knowledge of all physics working groups
  - filling code will be part of new physics analysis package H1PHAN++
    - will be written in C++ in an object−oriented way
    - based on algorithms of old physics analysis package

New Data Storage Model (5)



- HAT H1 Analysis Tag ROOT
  - tag database for fast and dynamic event selection
  - contains about 100 kinematic event variables, e.g. Q^2, x, y, etc.
  - experience from present tag database (Objectivity/DB)
- Run Catalog MySQL/Oracle
  - access to different parts of the event
  - prototype will be implemented in MySQL
  - Oracle as alternative

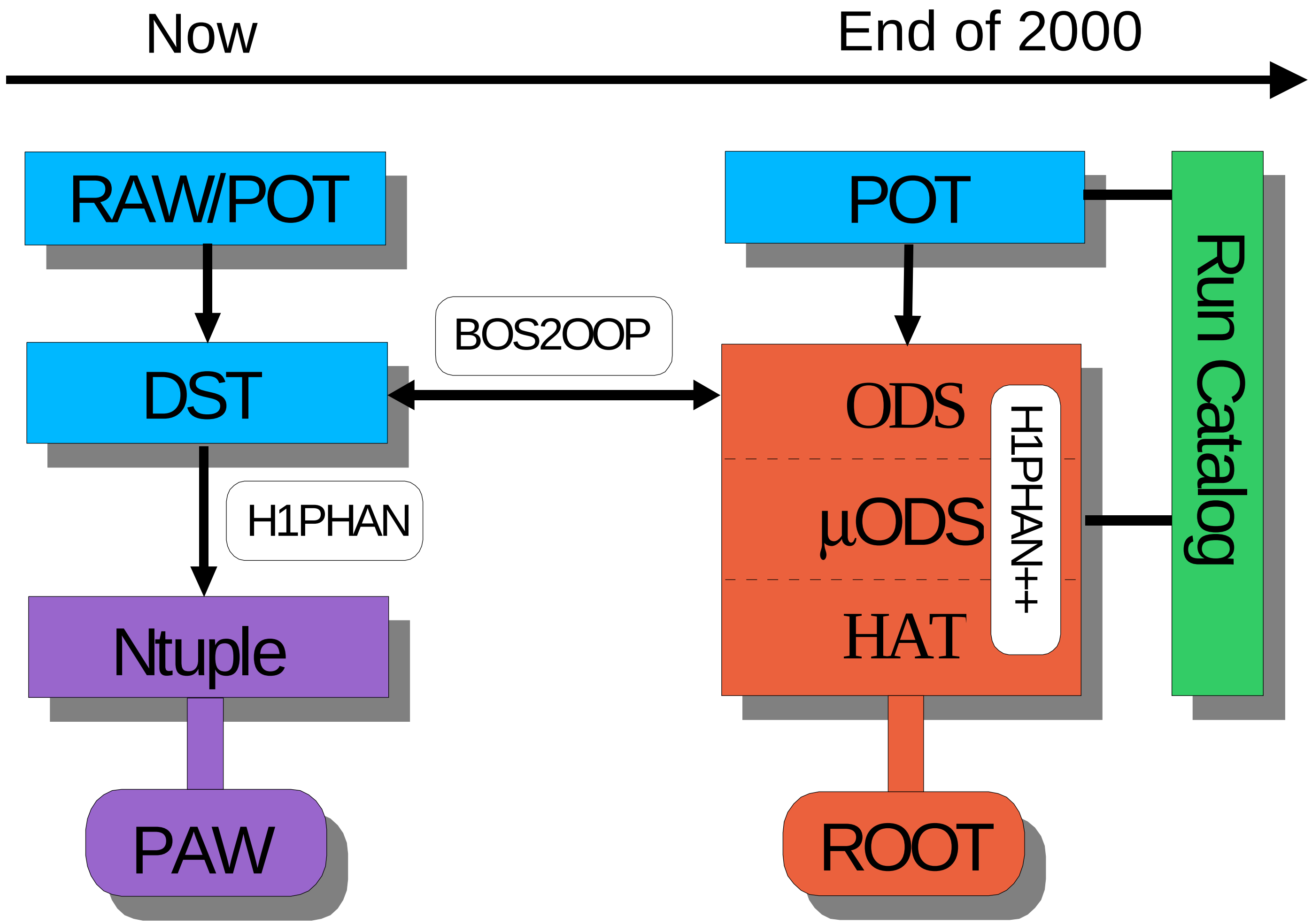
Data Volumes

H1 data volumes after the luminosity upgrade

in the year 2000

storage level	storage media	event size (kB)	data volume (GB/year)
POT	tape	200	10000
ODS	disk	15	500
μODS	disk	3	100
HAT	disk	0.5	15

Timeline for Implementation



Benchmark Results (1)

- ROOT I/O performance comparisons were already done by
  - ROOT team
    - ROOT versus ZEBRA
    - ROOT versus Objectivity/DB
  - CDF
    - ROOT versus YBOS

⇒ significant speed−up in data access time with ROOT I/O

Benchmark Results (2)

- H1 compared ROOT and FPACK
  - convert FPACK DST file into ROOT format
    - ROOT file factor 3 smaller for compression level 2
    - ROOT file factor 1.3 bigger for no compression
  - read whole event data
    - ROOT about a factor 4.6 slower than FPACK
    - no optimization done yet
  - read only part of the event data
    - no results from H1 available yet, but significant speed improvements expected

Conclusions

- the new H1 data storage model has been presented
- the choice of ROOT will provide an efficient solution for data handling and storage after the upgrade in the year 2000