DaVinci

The LHCb Physics Analysis Application

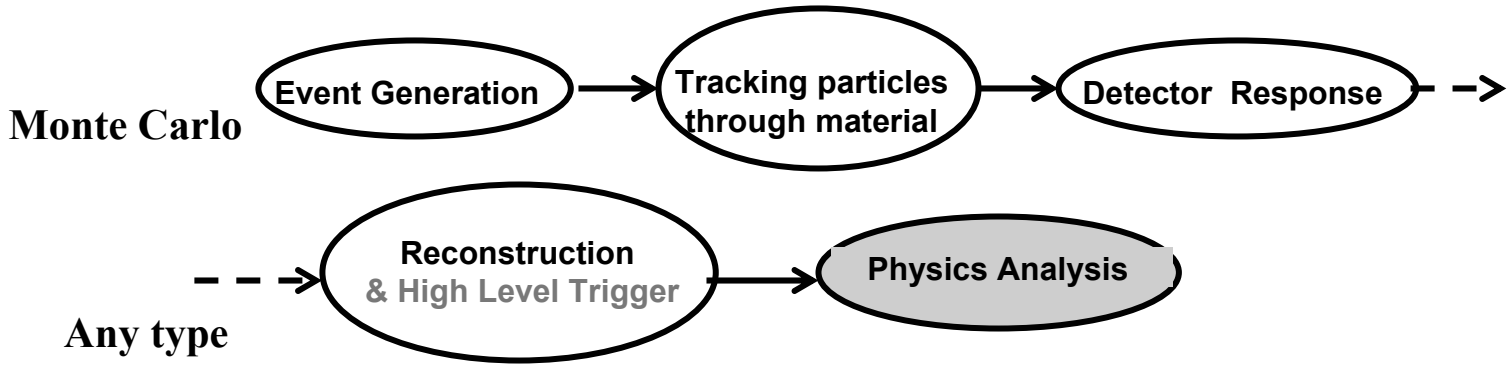
ATLAS Software Workshop
September 24th 2003
G. Corti / CERN

Outline

- Scope of DaVinci
- DaVinci architecture
  - A framework "on top" of Gaudi and LHCb core software
  - Relationship with Brunel, the reconstruction application
- DaVinci specific features and functionalities
  - Main components
    - The Physics Event Model
      - Relationship with reconstruction data
    - Basic and High Level tools
  - Analysis of Monte Carlo data
  - Input(s) & Output(s)
    - DST, stripped DSTs, Event tag collections, AODs, n-tuples
- Experience and conclusions

Scope of DaVinci

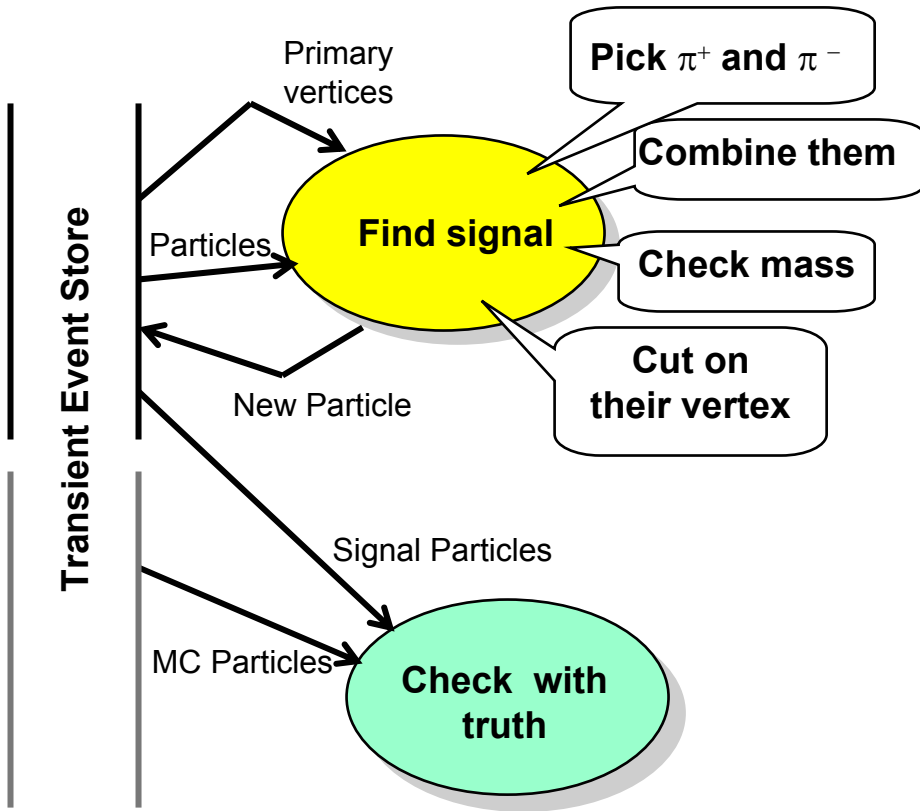
- In LHCb typical phases of Particle Physics software processing are "encapsulated" in various applications
  - Each is a producer and/or consumer of data for other stages



- Group and individual analysis proceeds from the output of the reconstruction. The framework for selections of events and physics analysis is called DaVinci
  - Provides common basis for writing physics analysis code
  - An application bringing together different libraries from Gaudi, LHCb Core software and providing physics analysis specific code

An LHCb physics analysis job

- LHCb primary goal is to make precision measurement of CP violating decays in the B system
  - Many different b decay modes with small BR
- In a typical analysis (ex: $B^0 \rightarrow \pi^+\pi^-$) the first step is to find interesting signal events by applying cuts on the characteristics of the decay
  - ParticleID, vertices, mass of combinations, topologies
- Compare with the truth for Monte Carlo data



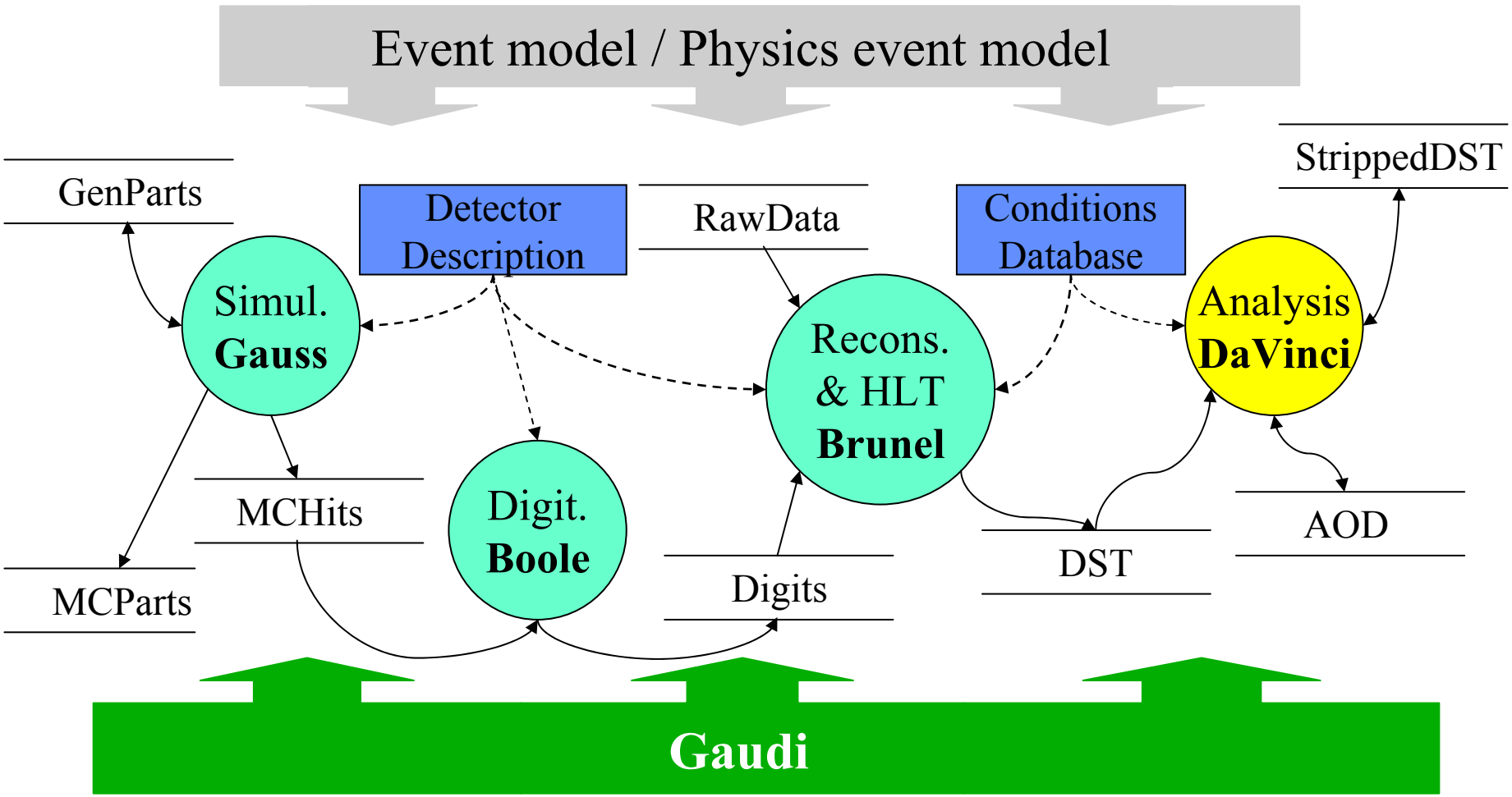
DaVinci and Gaudi

- DaVinci (as all LHCb applications) is based on the Gaudi framework
  - It follows the Gaudi design choices and evolution
    - Separation between "data" and "algorithms"
    - Separation between "transient" and "persistent" representations of data
    - "End-user" writes in few specific places (*Algorithms*, *Tools*)
    - Well defined component "interfaces" ("*Physics tools*")
  - Uses common services that it needs to provide a "generic" application functionality
    - JobOptions Service
    - Message Service
    - Persistency Services
    - Histogram Service
    - N-tuple Service
    - Particle Properties Service
    - Event Data Service
    - …
  - DaVinci takes care of data reading and integration of packages that are of common use.

DaVinci and LHCb core software

- All the LHCb Event Data Model and Detector Description are shared with the other applications
  - Both common "infrastructures" and detector specific
  - All data for the different stages of the processing are available directly (same file) or indirectly (via references)
    - Detailed understanding of reconstruction
    - Comparison with Monte Carlo truth
  - Algorithms providing functionality specific to different tasks can be shared, moved from / to DaVinci
    - HLT are being developed in DaVinci (easier to work on DSTs) to be moved to Brunel
    - New tuning of particle ID in DaVinci after better understanding of DST data

Relationships between Applications

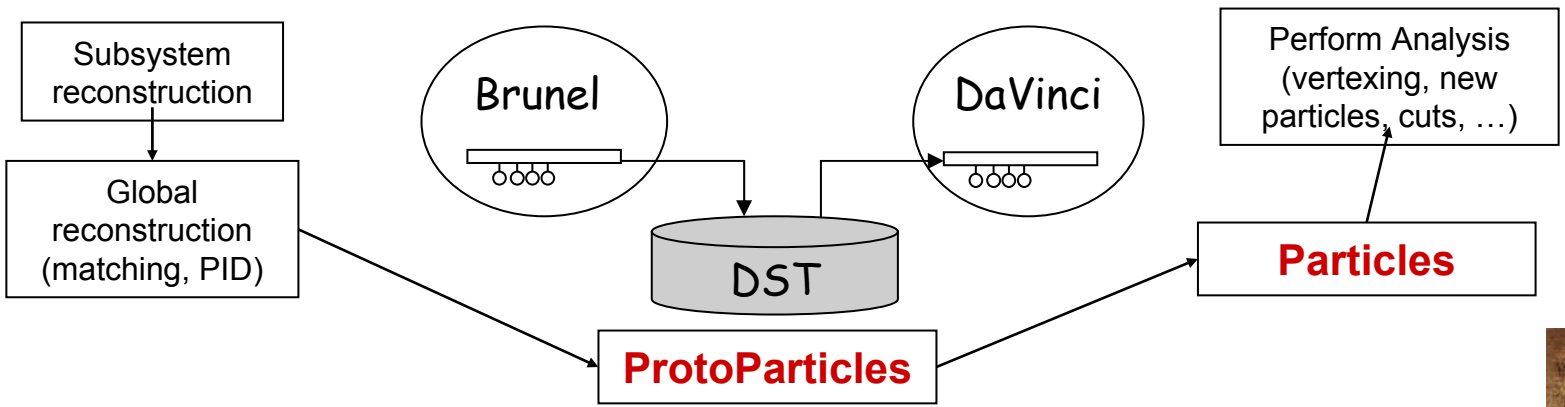


Interface with Brunel

- **Brunel writes a OODST file**
  - **Sub-system specific simple objects**
    - **Tracks, Calorimeter clusters, Muon system and RICH particle ID objects**
  - **Global reconstruction objects (ProtoParticles)**
    - **All available reconstructed information for a "physical" measurement**
- **DaVinci reads the OODST**
  - **Makes the "almost standard" reconstruction objects that were not done in the reconstruction, rerun with latest tuning (learnt from last Data Production)**
  - **"Particles" are produced from reconstruction objects with a specified particle ID hypothesis, interesting decays are "built" via selection and combination of particles**

Common

Analysis specific



DaVinci specific features

- Two main common aspects specific to DaVinci
  - The Physics Event Model
    - How to represent Particles, Vertices and their relationships
    - How to connect to the Reconstruction data objects
    - How to connect to the Monte Carlo data
    - Follow the LHCb Event Model conventions
      - Use utility classes and services
      - Gaudi Object Description
      - Programming language independent (XML based)
      - Automatic code generation of header files, serialisers and dictionary
  - Tools of general utilities are provided
    - "Basic" tools shared by all (most) physics analysis algorithms
      - For vertexing and manipulating the data
    - "High-level" tools and algorithms
      - Primary Vertex finder
      - Flavour tagging
      - Physics (Pre-)selections
    - Based on the Gaudi AlgTool and Algorithm classes

Physics Event Model

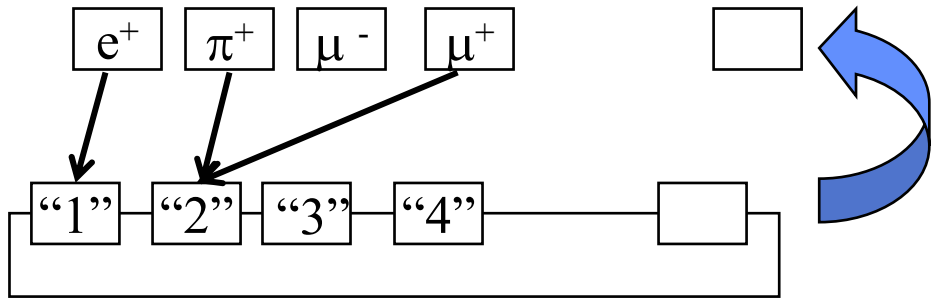
- A Physics analysis algorithm interacts mainly with two type of objects: Particles & Vertices
  - **Particle:**
    - contains the physical information about a "particle" and can be made starting from a ProtoParticle object or combining other Particle objects via composition tools
    - detector independent
  - **Vertex:**
    - contain the physical information for a vertex and can be made via vertexing tools as well as the links to the Particles used to make it

- Decay trees are represented through the relationship between particles and vertices

Relationship with reconstruction data

- **Particles & ProtoParticles**
  - **A ProtoParticle represent what the detector measured with all possible hypothesis which will help in deciding the nature of a particle**
    - Combines all sub-systems available information
    - End product of the reconstruction
  - **A Particle can be made starting from a ProtoParticle choosing ONE of the possible particleID hypothesis**
    - Different particles can originate from the same ProtoParticle

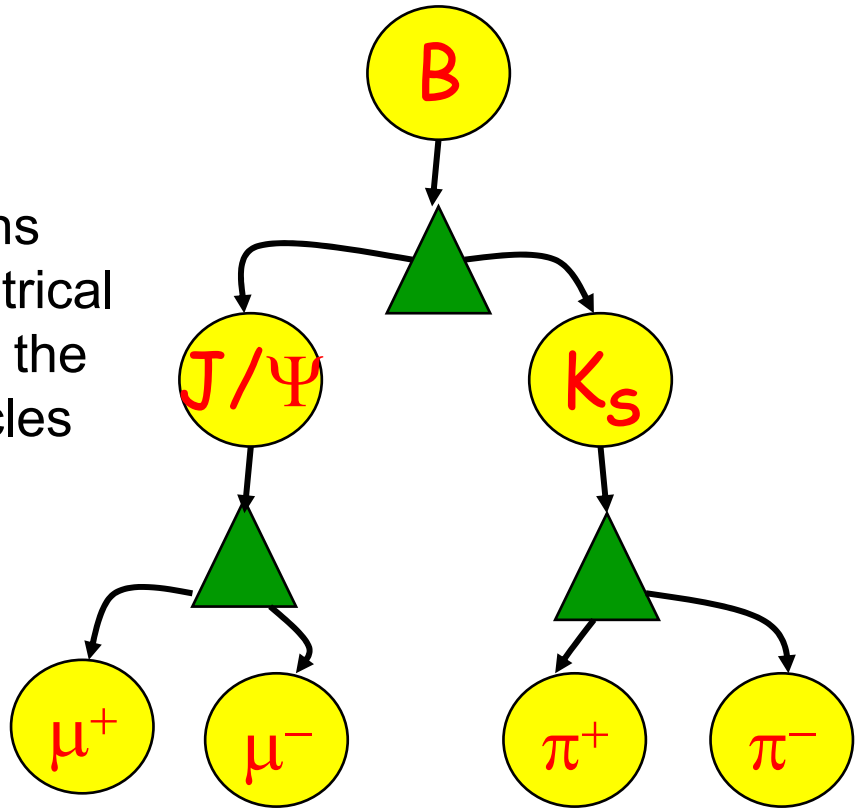
For ex: all pions with CL > 40% and all muons where muonID has the highest CL



  - **ProtoParticles are the bridge between Physics analysis and Reconstruction**

Decay Tree representation

Represent Decay tree through
Particle(parent)-Vertex(decay)-Particles(products) relationship



A vertex contains both the geometrical information and the link to the particles used to make it

The B candidate has access to the full decay

Vertices are external to Particles

Stable particles have null pointers to their decay vertex

Always a vertex to hold the relationship

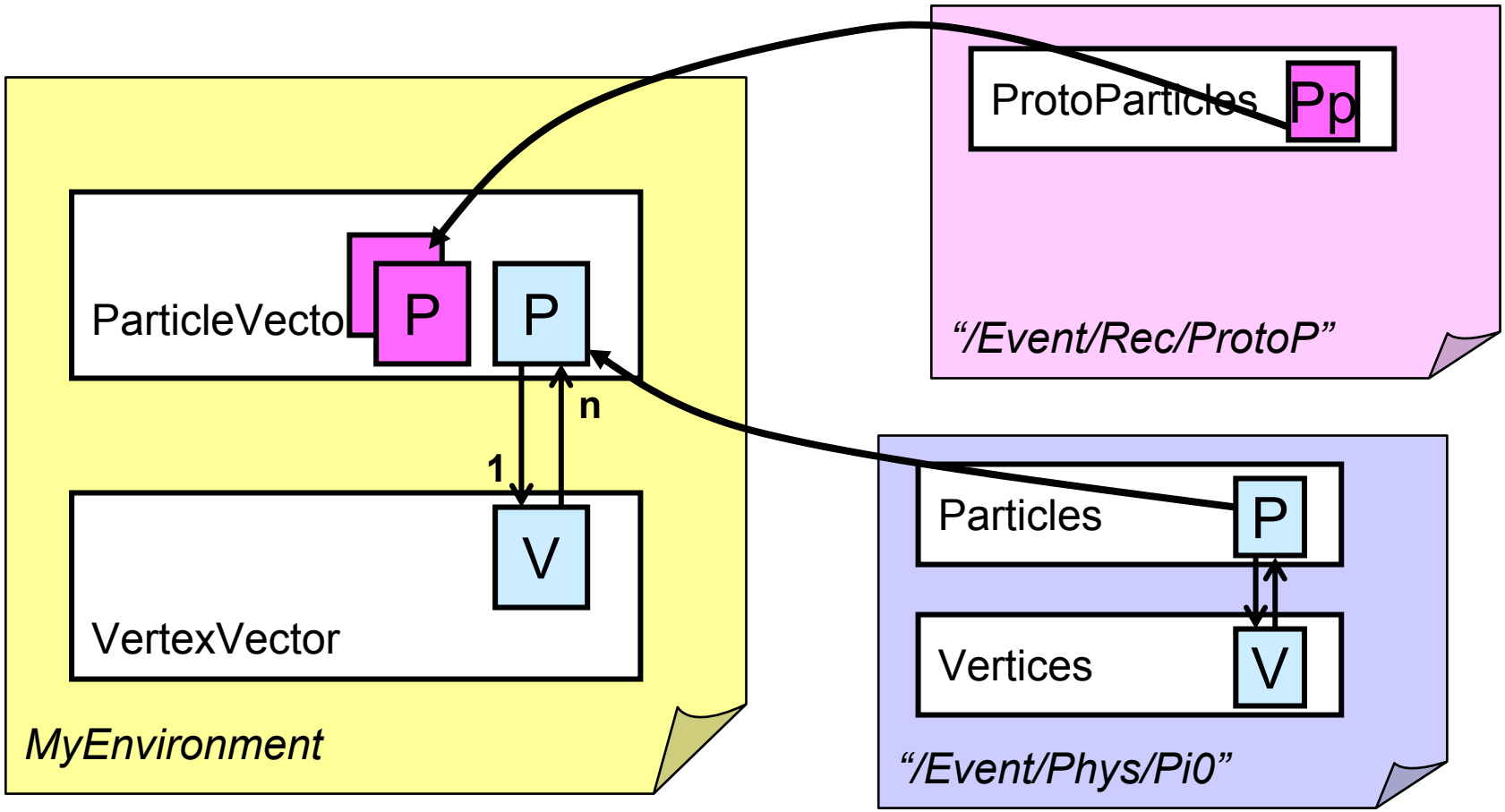
"Basic" Physics analysis tools

- DaVinci provides different category of "basic" tools, performing specific *simple* tasks as self contained as possible
  - Tools related to Particle "management"
    - ***PhysDesktop*** **to shield the end-user from interacting with the Transient Event Store in retrieving "old" Particles, making new ones and saving them**
    - ***ParticleMakers*** **to make Particles from ProtoParticles based for example on Combined ParticleIDs**
    - ***ParticleFilter*** **to organize Particles in containers based on various picking criteria**
    - **…**
  - Vertexing and topological tools
    - **Geometrical and mass constrained vertex fitters**
    - **Tools to calculate geometrical information (Impact Parameters, Distance of closest approach)**
    - **Pointing constraints**
    - **Tools to *transport* a Particle to a given position**
    - **…**
  - They are "all" necessary to have a complete physics analysis

Management of Physics data in the event store

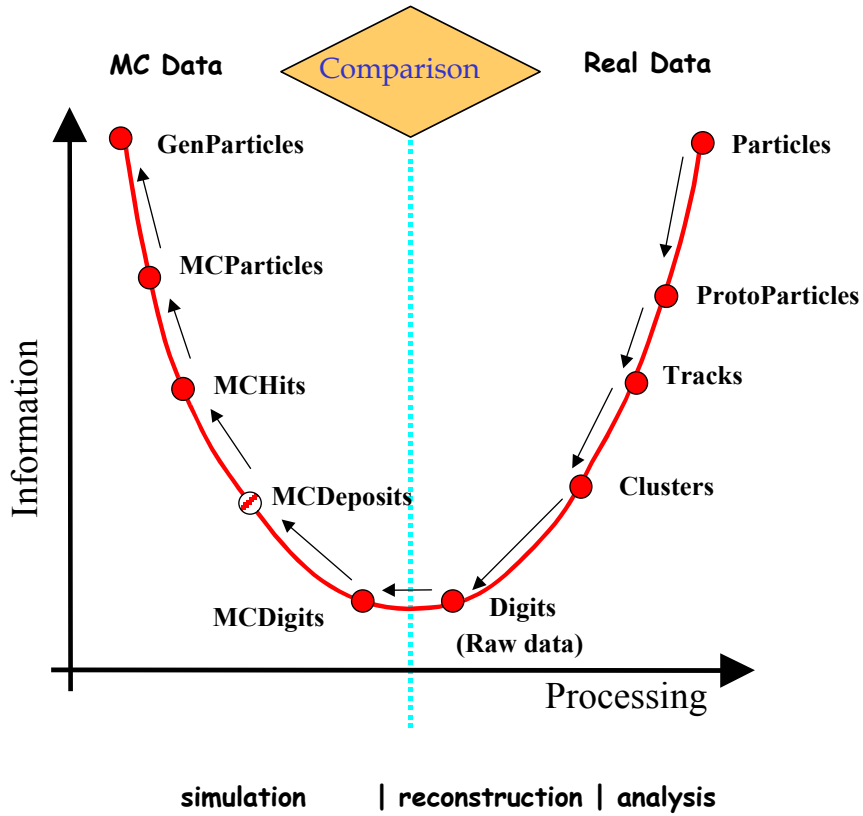
- The PhysDesktop hides the complexity of interacting with the Transient Event Store, of converting ProtoParticles to Particles and maintains the working environment
  - Particles are specific to an analysis
    - Not all selection need all type of Particles
  - Simple Particles are "picked" from ProtoParticles as needed
    - Delegates to a ParticleMaker that "knows" how to transform from the different type of objects with various criteria (ex. CombinedPIDParticleMaker)
  - Particles can be available from previously run algorithms
    - Ks, pi0, J/Psi, D+, …..
    - They are in different locations in TES
  - For an analysis all Particles (and Vertices) should be collected together
    - The PhysDesktop gets the initial Particles&Vertices in local containers with all identified Particles (from ProtoParticles or from previous processing) and eventual Vertices
  - New Particles and Vertices are produced
    - Only "good candidates" will be saved for further processing
    - Has a set of methods for saving the whole or parts of the working environement
  - Both the PhysDesktop and the ParticleMakers are configured via Job Options

The PhysDesktop



Analysis of Monte Carlo data

- In the Analysis of Monte Carlo events necessary to compare what is reconstructed with the generated Monte Carlo ("the truth")
  - Keep the two worlds as separate as possible
  - Associator tools to connect Particles with MCParticles
    - Chi square agreement on kinematics
    - Following links to originating objects and their association



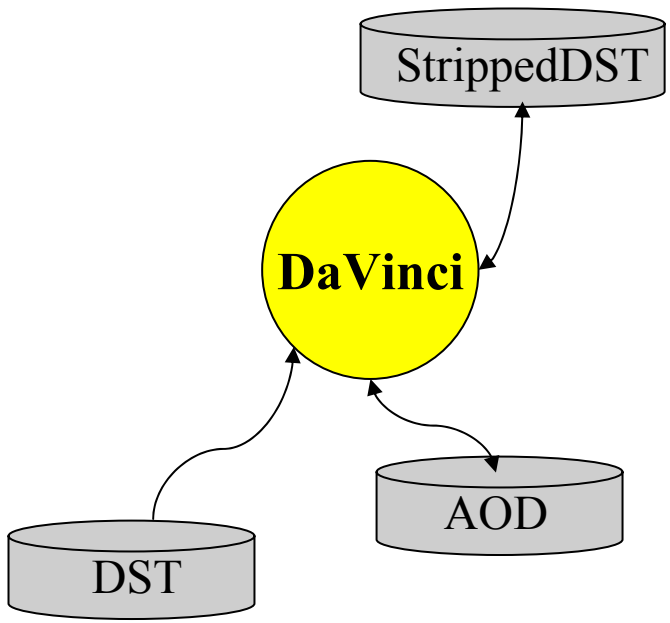
- Utility tools
  - MCDecayFinder
    - Finds an MCParticle decaying as specified in a string in jobOptions

      J/psi(1S) → mu+ mu- {gamma,} (optional gamma)

  - Debugging tool
    - to print a tree, a single particle, the whole event both for MCParticles/MCVertices and Particles/Vertices

Input(s) & Output(s)

- DaVinci can be (is) used for various types of analysis jobs with different inputs and producing different outputs
  - The first stage of the analysis uses DST as input
  - Production of stripped DSTs (containing selected events) are made with pre-selections algorithm integrated in DaVinci
  - Some end-user (depending on their needs and taste) write
    - AOD
    - AOD+DST of selected events
    - Event tag collections
    - N-tuples (HBOOK, ROOT)
    - Histograms (HBOOK, ROOT)



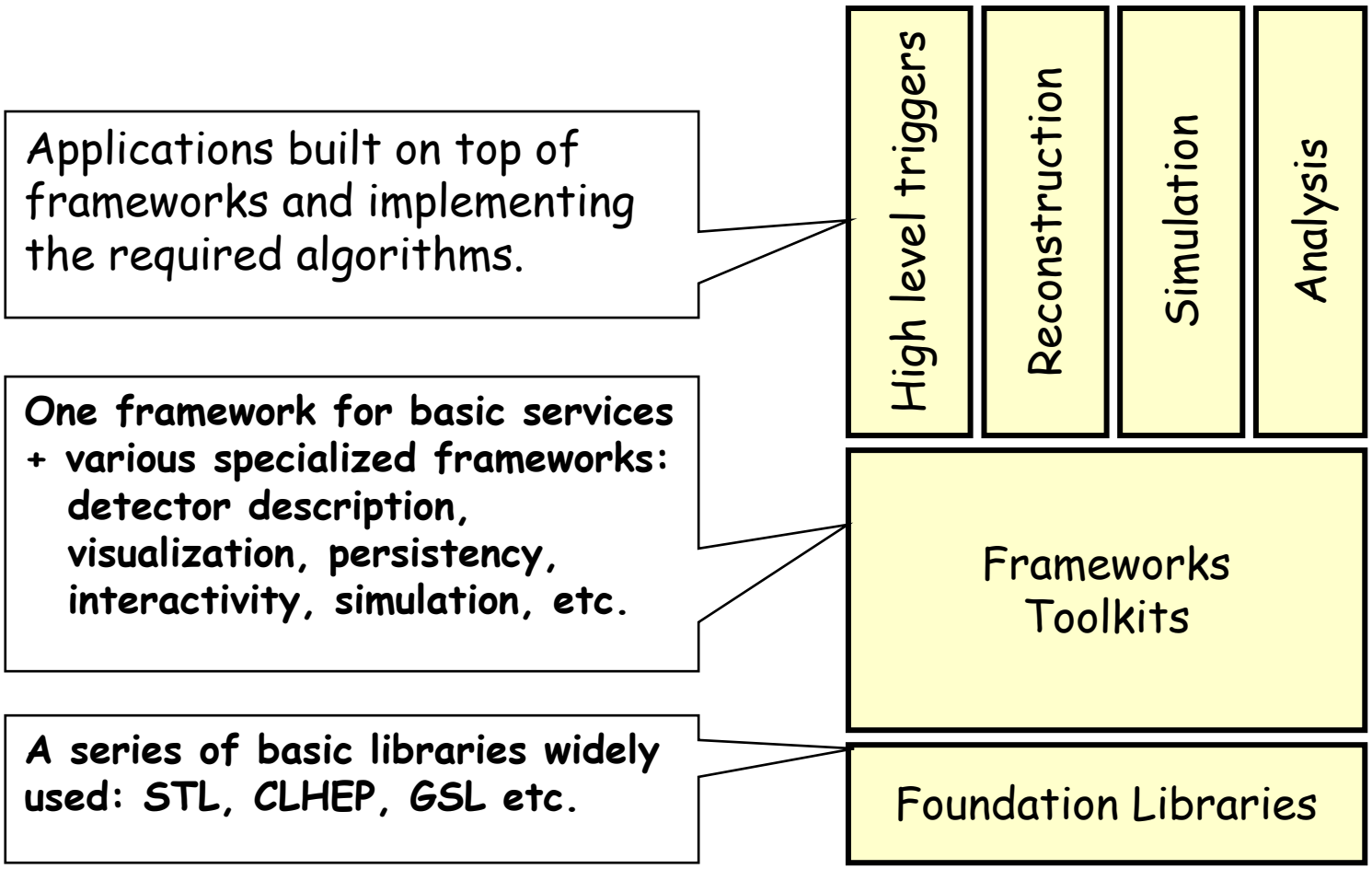
Experience and conclusions

- DaVinci used for Physics studies since July 2002
  - First version completely OO
- Extensively used
  - Provide feed-back to Brunel and production
    - … to understand the data and how further process them
  - Develop new tools and algorithms (physics analysis, HLT)
    - … as long as using data on OO-DST
- All selection studies for Re-optimization and Trigger TDRs performed with DaVinci
  - Provides a "complete" suite of physics tools for selections and evaluation of the results
  - About 20 physics selections mostly developed by Ph.D. students
  - Repository of (Pre)-selection algorithms for use by the all collaboration
  - Pre-selection production on 10M generic b events

Experience and conclusions

- Most users interact with this application. First exposure to LHCb software
  - Must be easy to use and with easy to understand nomenclature
  - Tutorials provided
  - A sorry point is very little documentation available (DoxyGen), mostly people use examples
- DaVinci (both the basic tools and the physics event model) are separated and independent from the LHCb reconstruction
  - Only one connection
- Extensions to the (basic) tools and to the physics event model can be easily accomodated
- Review of DaVinci before end of the year after the extensive use for the TDRs studies
  - Revaluate choices made, maintain software but also go back to the drawing board if necessary

Software Organization



General structure