Blueprint

- Start the production of a detailed document on data preservation
  - Gets in details of the individual projects
  - Proposes working directions
  - Define the organization

- Destination: funding requests
  - For individual need (Babar from SLAC, HERA from DESY, Tevatron from DOE/Fermilab)
  - For common projects

- Method: split in chapters, break in subgroups, get together for common sessions
  - Leave CERN with version 0.2

Blueprint of
Data Preservation in High-Energy Physics

Study Group for Data Preservation and
Long Term Analysis in High Energy Physics

http://dphep.org

Abstract

Data from high-energy physics (HEP) experiments are collected with significant financial and human effort and are mostly unique. At the same time, HEP has no coherent strategy for data preservation and re-use. An proposes an International Organization devoted to the data preservation in high-energy physics. The organization is structured around an inter-experimental Study Group supervised by the International Committee for future Accelerators (ICFA). The present document presents the motivation for such an organization, collects example of initiatives at experiment level, defines common R&D projects and draw the main lines of the

Content

- Chapter 1: Executive Summary and General remarks

  – (ICFA document)

- Chapter 2: Physics Motivation: detailed cases, simulations

- Chapter 3: Experiments archival models

- Chapter 4: DPHEP Common projects

- Chapter 5: DPHEP Organization

Chapter 2: The Scientific Potential of the Data Preservation in High Energy Physics

- Prospects for the utility of the whole enterprise

- Past: JADE+DIS+find others.

- Examples of "Gedanken experiments" on archived data
  - I.e. where are the non-experimental weak points
  - Quantify what it needs to re-do the analysis
  - What the result would look like if some errors are reduced

- Give 2-3 striking examples from each experiment:
  - A Measurement (strong coupling, branching ratio, jet cross sections)
  - A search for new physics
  - A complementary analysis or a combination with a future/existing project

- A summary of policies/practices in other fields

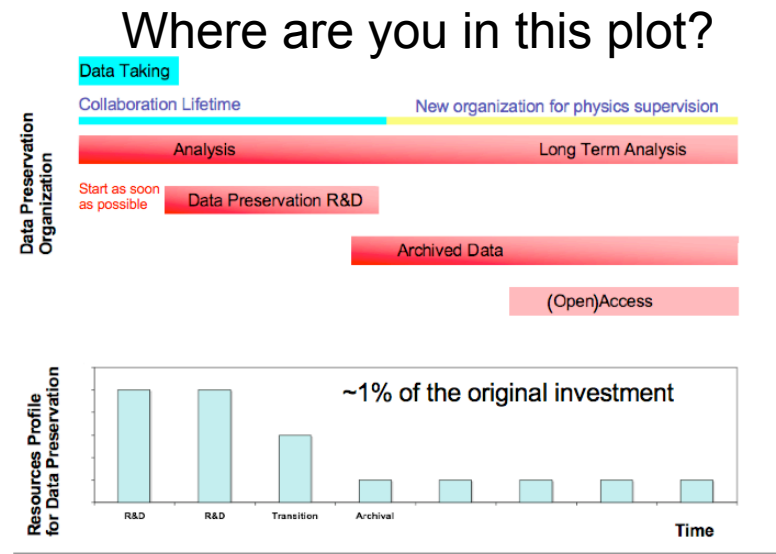# Chapter 3: Experiments DP projects
A: project/hardware/resources

- The present computing model

- The archival system (description)

- The resources

  - Person power for R&D

    - Within the experiment

    - External needs

  - Hardware requirements and Computing Center issues

    - Archival strategy: what "bits" will be "stored"

    - Access to data policy

    - Security

  - Data archivist requirements: daily/monthly/yearly tasks

Chapter 3: Experiments DP projects
B: governance, international scene

- Organisation of the Physics Supervision (connection to DPHEP)
- Timelines, transition
- Costs/benefits statement
- Interface to DPHEP
  - (input to Chapter 5 too)

- A more global aim: produce a template for an experiment to be part of DEPHEP (see later)

Chapter 4: Inter-experiment R&D projects

- DPHEP-RD1: Technology for Data Preservation
  - Virtualization
  - Multi-core
  - Documentation parsing techniques (*)
  - Validation Suites?
  - Super-structure for archival systems?
- DPHEP-RD2: Outreach
  - Standard formats, tools, communication techniques
- DPHEP-RD3: Extension of the public documentation
  - Figures, Ntuples, Notes and internal legacy (*)

Chapter 5: DPHEP

- **Modus Operandi**
  - management definitions, meetings, reports

- **Connections levels**
  - Experiments

    Supervisions committees => DPHEP
  - Laboratories

    Data Archive Team=> DPHEP
  - Connections from external HEP contributors

  - Connections to other fields

    NB: IVOA HEP contact person: Fabio Pasian, connect more organizations?)

A long term organization of HEP experiments
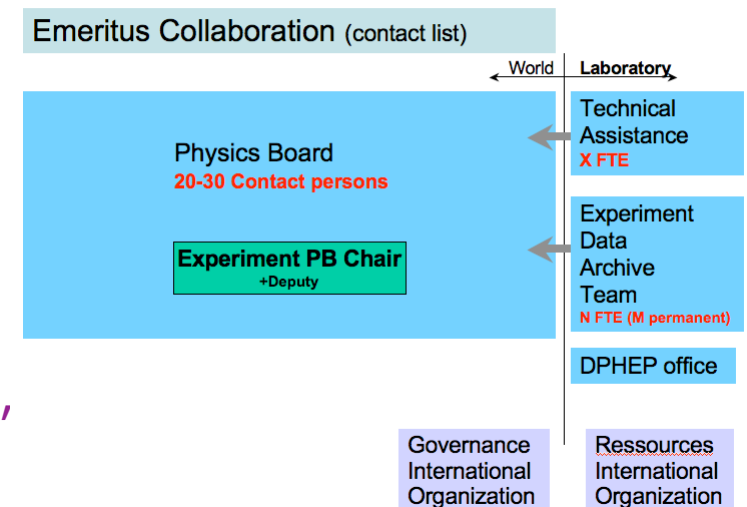
Chapter 5: DPHEP Organization

- Define/propose the funding model

  – Possible contributors

  – Investigate funding programs

    USA

    Europe