A Physics-Optimized Grid Environment for Experiments

Stewart C. Loken

18 August 1999

The Challenges

- Providing rapid access to event samples and subsets from massive data stores, from 100s of Terabytes in 2000 to 100 Petabytes by 2010.
- Transparent access to computing resources, throughout the US, and throughout the world.
- The extraction of small or subtle new physics signals from large and potentially overwhelming backgrounds.
- Enabling access to the data, and to the rest of the physics community across an ensemble of networks of varying capability and reliability, using heterogeneous computing resources.

Achieving a Balance

- Proximity of the data to central computing and data handling resources,
- Proximity of frequently accessed data to the users, to be processed in desktops, local facilities, or in "regional centers",
- Making efficient use of limited network bandwidth; especially transoceanic,
- Making appropriate use of regional and local computing and data-handling,
- Involving scientists and students in each world region in the physics analysis

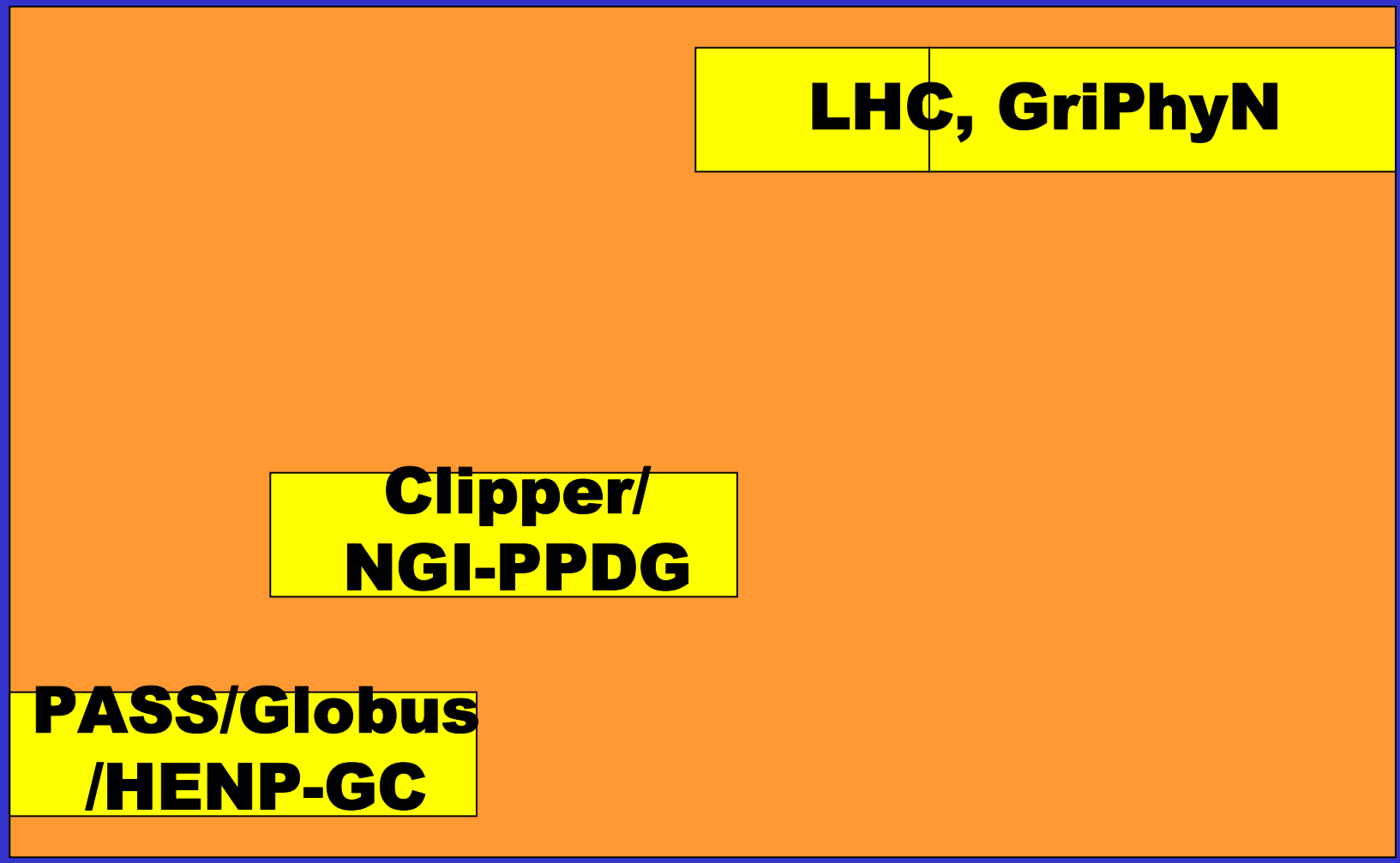
Need for Optimization

- Meeting the demands of hundreds of users who need transparent access to local and remote data in disk caches and tape stores;
- Prioritizing hundreds to thousands of requests from the local and remote communities;
- Structuring and organizing the data, and providing the tools for locating, moving and scheduling data transport between tape and disk and across networks,
- Ensuring that the overall system is dimensioned correctly to meet the aggregate need.

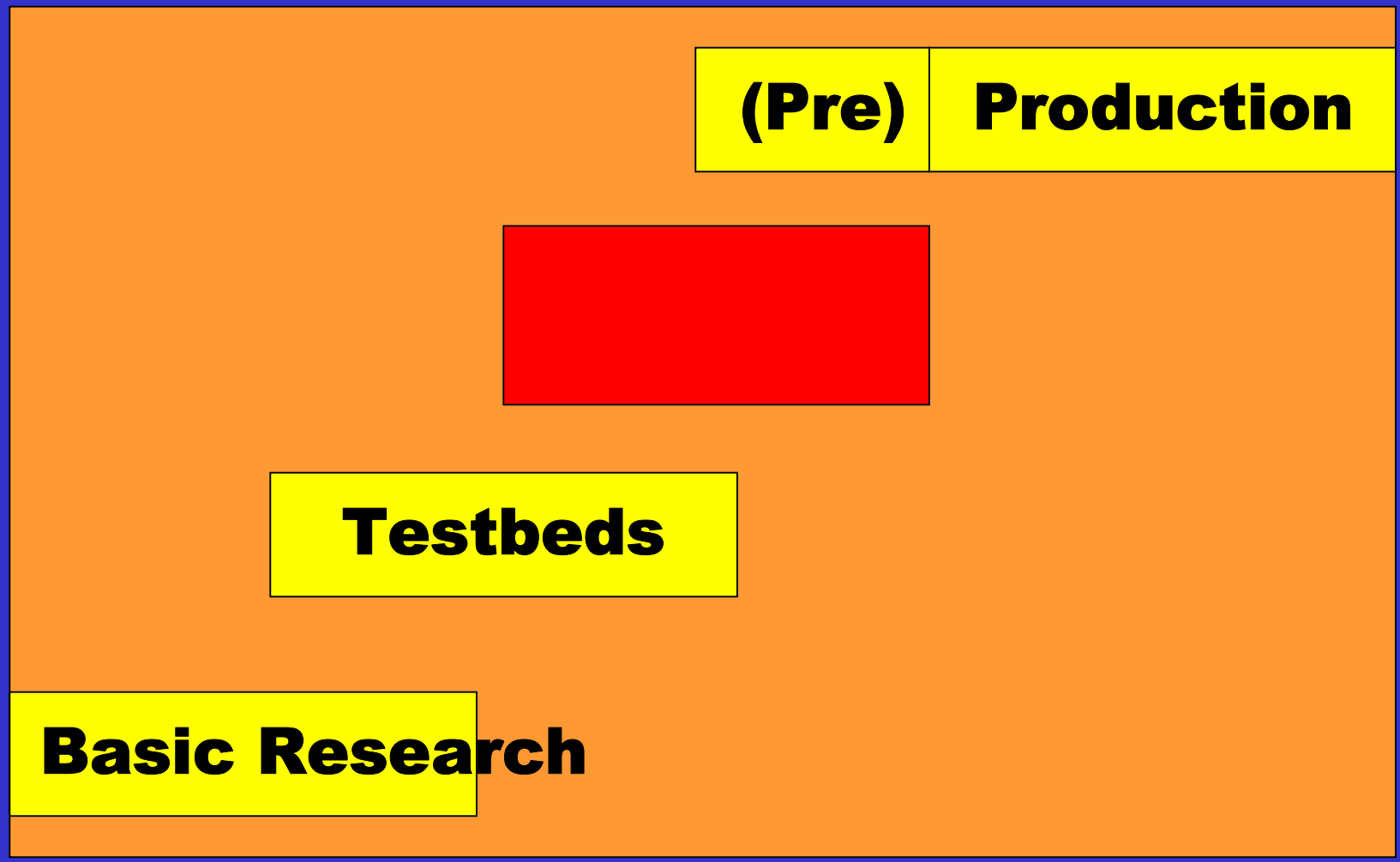
Multiple Tiers

- Tier 0: The central facility at the experimental site (CERN in the case of the LHC experiments)
- Tier 1: The national regional centers (such as the US center at FNAL for US CMS and at BNL for US ATLAS)
- Tier 2: A center in one region of a large country (e.g. the US centers proposed by the GriPhyN collaboration)
- Tier 3: The computing resources of a single university group
- Tier 4: An individual workstation.

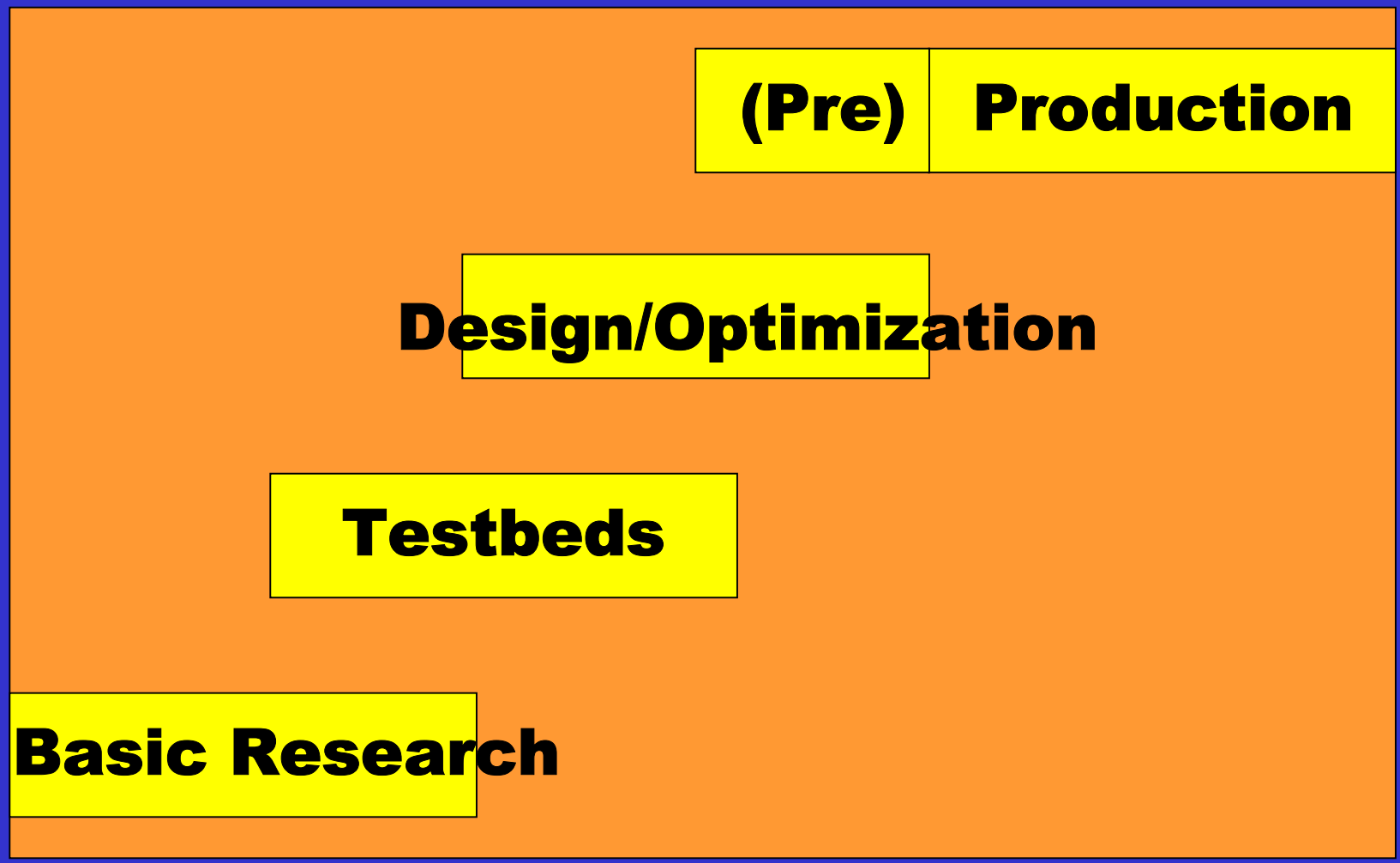
Technology Projects



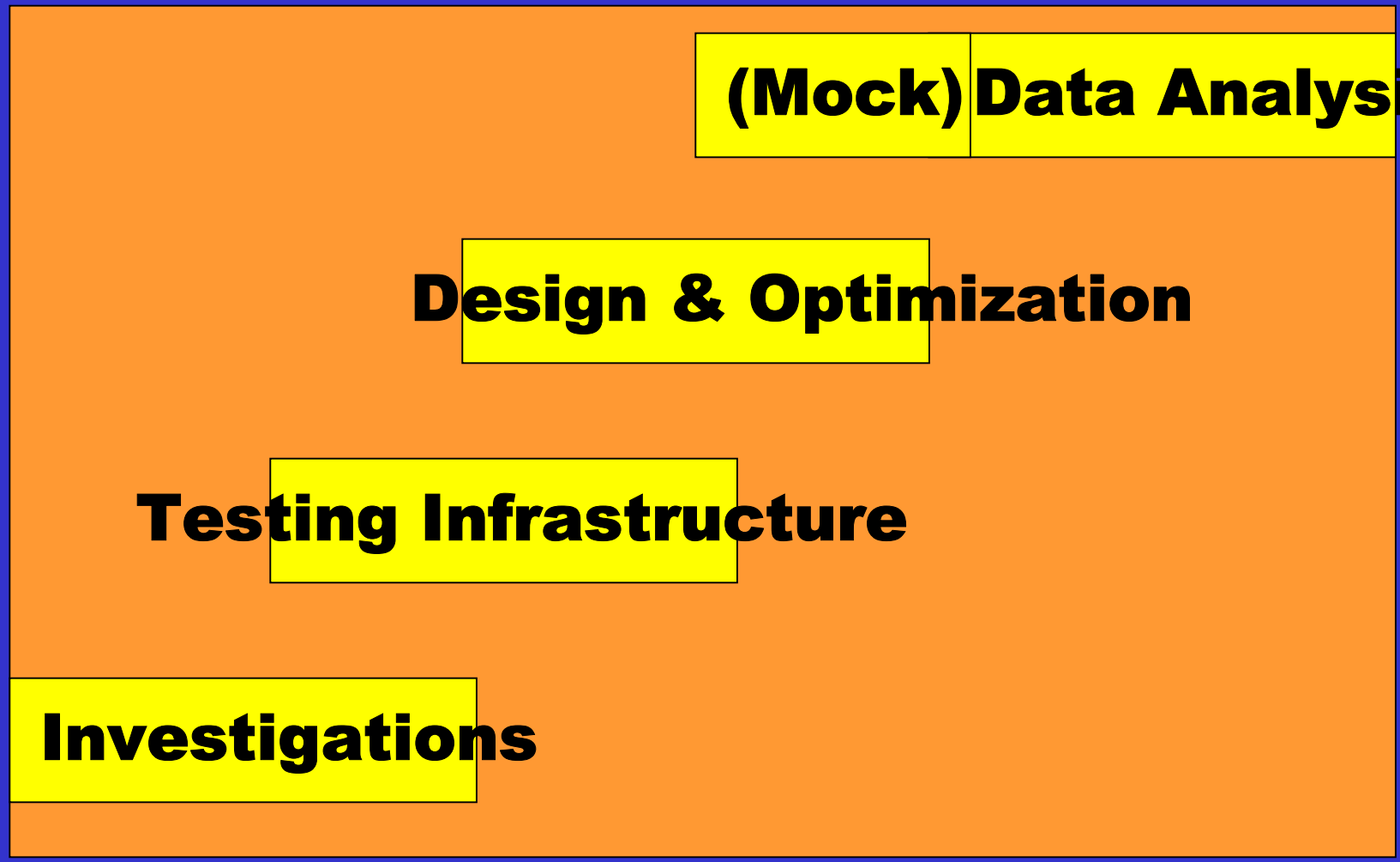
The Gap


(Pre)
Production
Testbeds
Basic Research
1995
2000
2005
2010

LHC Timeline



Activities


(Mock) Data Analysis
Design & Optimization
Testing Infrastructure
Investigations
1995
2000
2005
2010

Program Goal

- A Coordinated Approach to the Design and Optimization of a Data Analysis Grid for HENP Experiments

Initial Focus

- Simulation of a Computational Grid Hierarchy
- Instrumentation of a Grid Prototype
- Optimization of HENP Systems
- Coordination of U.S. HENP Distributed Data Management Projects

Deliverables

- Modular Simulation System of facilities, networks, analysis strategies and network use
- Implementations of full-scale simulations for ATLAS and CMS
- Toolset for Optimizing Systems
- Evaluation Criteria
- Strategies and Guidelines

Relationships

- Builds on Particle Physics Data Grid, MONARC
- Utilizes results of GIOD, Clipper, Globus, Nile and Condor
- Will work with Fermilab, BNL et al. In study and development of Regional Centers
- Will be a critical for NSF project for a Grid Physics Network (LHC/LIGO)

What is a Grid?

- **New paradigm for large-scale computational science**
  - Transparent application access to remote data and systems
  - Supports moving, visualizing, and analyzing large data sets
  - Analogous to electric power grid
- **Aggregates existing computational infrastructure**
  - Computers
  - Mass storage systems
  - Experimental apparatus
  - Visualization devices

Appeal of the Grid

- Pervasive availability of computational and storage facilities
- The infrastructure and services to support distributed analysis and collaboration
- Unprecedented support, participation and investment

Timing is Everything

- This is the moment to invest in this effort
- Significant contributions by many key people who are now engaged in the HENP problem e.g.
  - Ian Foster
  - Miron Livny
  - Bill Johnston
  - Brian Tierney

Long-term Vision

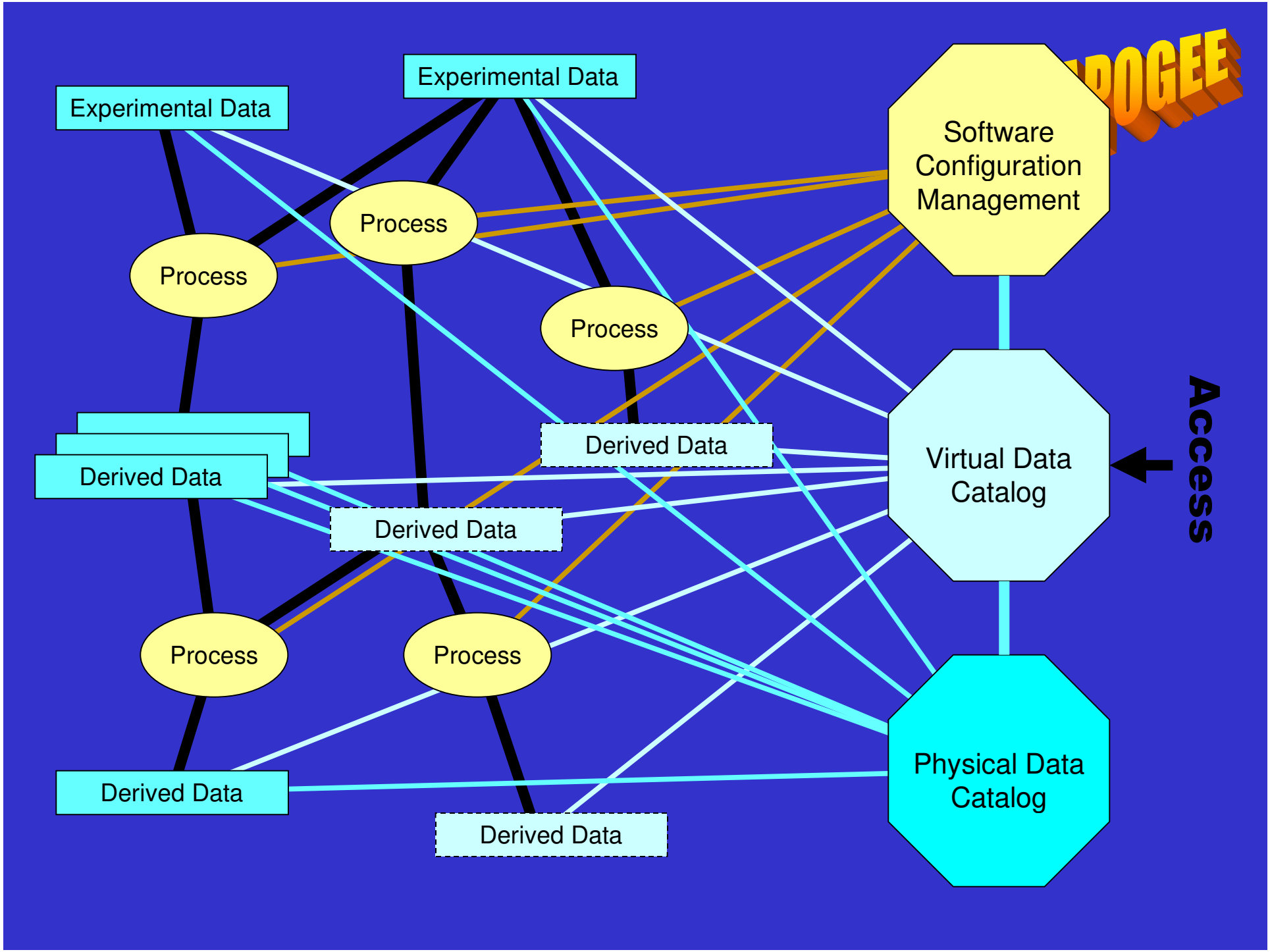
The HENP Data Grid will be the foundation of a longer-term plan, building on the PPDG and earlier projects, and leading into the production-scale Grid hierarchy that will include the LHC Tier 1 Centers and the Tier2 regional centers proposed to NSF in the GriPhyN project.

Three Major Phases

- Support for distributed transparent read/write access to high-volume data
- Support for wide-area distributed computing system where data-intensive tasks move automatically
- Support for agent technology and "virtual data"

Virtual Data

- All derived data (objects, collections or files) are defined in terms of
  - input data
  - operation performed
- Zero, one or multiple copies of derived data may be cached by the Data Management System
- The user need not be concerned with:
  - number and location of cached copies;
  - location at which calculations are performed.


Experimental Data
Experimental Data
Process
Process
Process
Derived Data
Derived Data
Derived Data
Process
Process
Derived Data
Derived Data
Software Configuration Management
Virtual Data Catalog
Physical Data Catalog
Access

HENP Data Management

Rigorous Approach to System Design and Optimization:

- Instrumentation at all levels, *e.g.*
  - Mass storage management
  - Network services
  - Cache management
  - Query optimization
- Modeling and simulation with predictive power.

Conclusion

- The proposed work fills a critical gap that exists in on-going and future projects
- This proposal builds on the results of basic Computer Science R&D and leverages investment in testbeds
- This work will create an optimized design for the HENP facilities of the future