



Datasets on the GRID

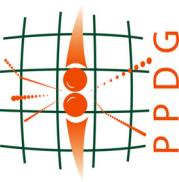
David Adams

PPDG All Hands Meeting

Catalogs and Datasets session

June 11, 2003 BNL

Introduction

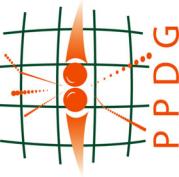


- **Contents**
 - What is a dataset?
 - Dataset properties
 - Logical *vs.* physical
 - Dataset operations
 - Implementation

- **More information**

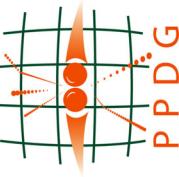
- Dataset page (little out of date) at
 - <http://www.usatlas.bnl.gov/~dladams/dataset>
- Document with content similar to this talk
 - <http://www.usatlas.bnl.gov/~dladams/dataset/griddataset.doc>

What is a dataset?



- A dataset is a collection of data that belong together
- Important example for HEP is the event dataset
 - Data organized as a collection of events
 - Other realms may have records equivalent to events
 - Each event/record to be processed in the same way
- Use cases:
 - Select a dataset
 - Of all the event datasets in my experiment, which meet my criteria?
 - Process a dataset
 - For each event in my event dataset, run my application and apply some task
 - Natural to do distributed production by splitting the dataset along event boundaries

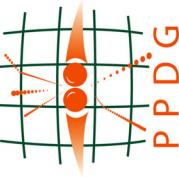
Dataset properties



- Datasets have the following properties
 - 0. Identity
 - 1. Content
 - 2. Location (of the data)
 - 3. Mapping (content to location)
 - 4. Provenance (of the dataset)
 - 5. History (of the production)
 - 6. Metadata (describing dataset)
- Details follow
- All the above are “metadata”
- Dataset also has “data”

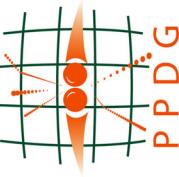


Dataset properties (cont)



- Identity
 - Index or name to distinguish datasets
- Content
 - Tells what kind of data is carried by the dataset
 - E.g. for HEP event data
 - List of event identifiers
 - Content for each event (raw, clusters, tracks, jets, ...)
 - Location
 - Most interesting example is a collection of logical files
 - Could also be
 - physical files
 - DB
 - data store such as POOL

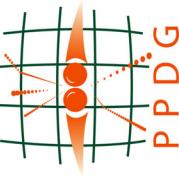
Dataset properties (cont)



- **Mapping**
 - For each piece of content, where (in the location) is the associated data
 - For collection of logical files, which file and where in the file
- **Provenance**
 - Parent datasets
 - Transformation
 - Applied to parent dataset to obtain this dataset
 - Sufficient to construct dataset (virtual data)
- **History**
 - Production history beyond provenance
 - Division into jobs, which compute nodes, time stamps, ...



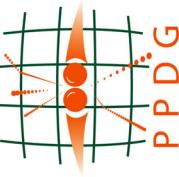
Dataset properties (cont)



- Metadata
 - Additional data characterizing the dataset
 - For example
 - Motivation for dataset
 - E.g. starting point for Higgs searches
 - Results of dataset evaluation
 - E.g. approved for basis of publications
 - Other global properties
 - E.g. integrated luminosity for an event dataset



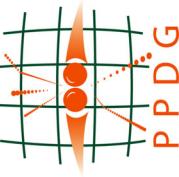
Logical vs. physical



- Distributed file management expressed in terms of logical files and physical replicas
- Useful to introduce the same decomposition for datasets
 - Data in a dataset may be replicated to produce another physical representation (replica) of a dataset
 - Common example is after event selection
 - Selected event ID's plus original dataset or
 - Replicate selected objects to define the new (physical) dataset
 - These are two replicas of the same logical dataset
 - For a dataset made up of logical files
 - File replication does not imply a new physical dataset
 - Moving data to a new collection of logical files is a new physical dataset



Logical vs. physical (cont)



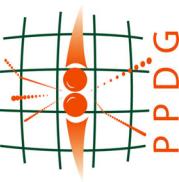
Property	Logical	Physical
Identity	X	X
Content	X	
Location		X
Mapping		X
Provenance	X	
History		X
Metadata		X

Dataset properties are associated with either the physical or logical view of a dataset

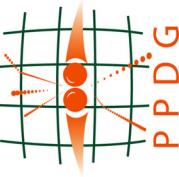


Logical vs. physical (cont)

- Our use cases break down along logical/physical properties
- Dataset selection uses content, provenance and metadata, the logical properties
- Job processing requires content, location and mapping, i.e. must make use of the physical properties

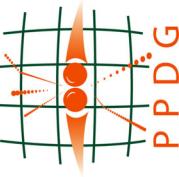


Dataset operations



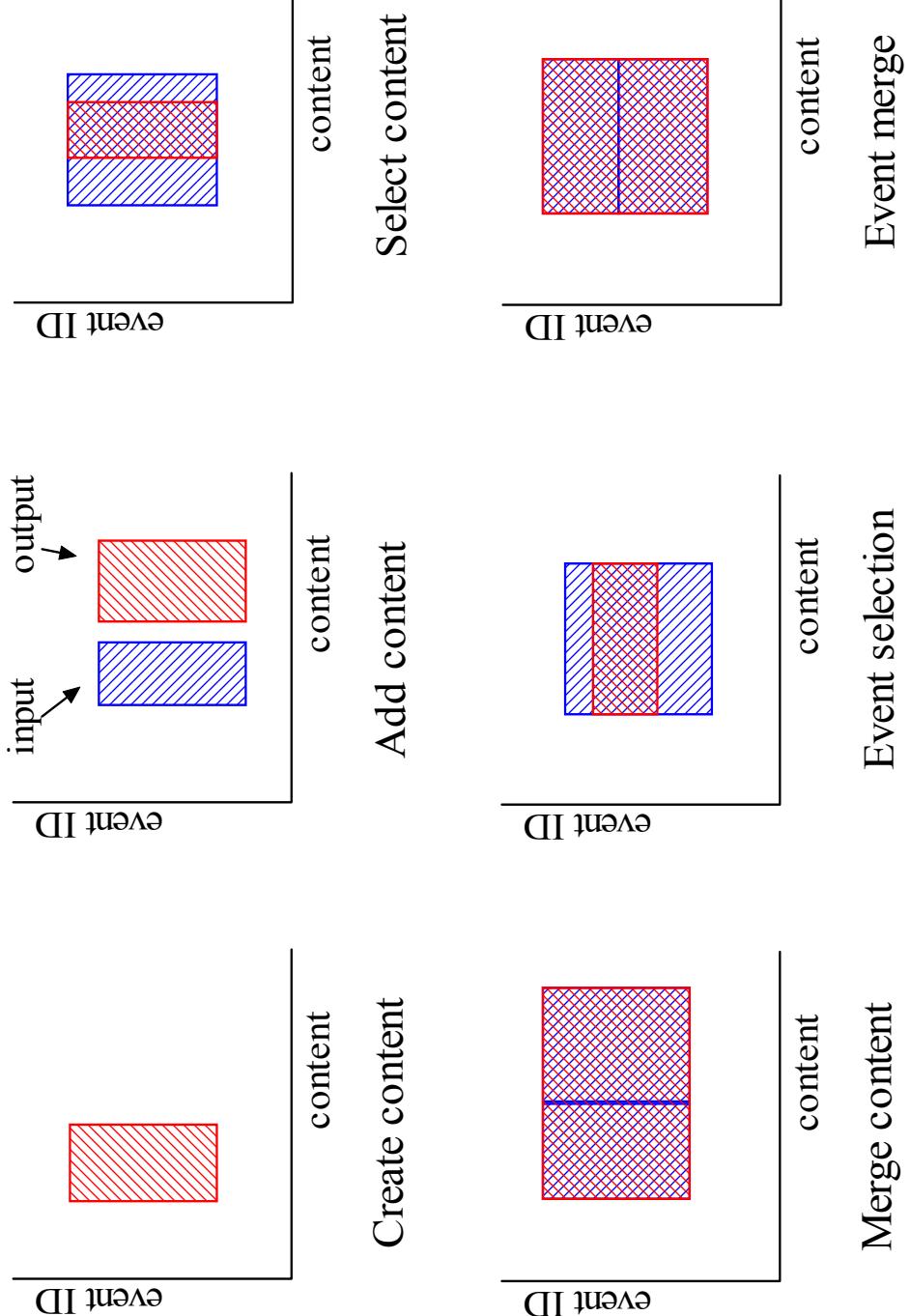
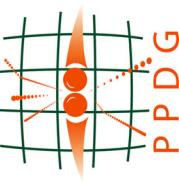
- Next consider some of the operations that datasets should support
 - These include
 - Content selection
 - Splitting
 - Merging
 - Content selection
 - User can create a new dataset by selecting content from an existing dataset
 - For an event dataset, there are two categories
 - Event selection
 - Apply an algorithm which flags which events to accept
 - Event content selection
 - e.g. keep tracks, drop jets

Dataset operations (cont)



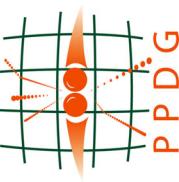
- **Splitting**
 - Distributed processing requires that the input dataset be split into sub-datasets for processing
 - Natural to split along file boundaries
 - For event datasets, the split is along event boundaries
- **Merging**
 - Combine multiple datasets to form a new dataset
 - For event datasets, there are again two dimensions
 - Datasets with the same event content and data for different events may be combined
 - E.g. data from two different run periods
 - Datasets with the same events and different content may be combined
 - E.g. raw data and the corresponding reconstructed data
 - Or reconstructed data and the conditions data used in its construction

Dataset operations (cont)



Event dataset operations

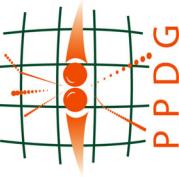
Implementation



- The dataset properties span many realms
- Different properties will be stored in different ways
 - Properties used for selection naturally reside in relational DB tables
 - Content, provenance, metadata
 - Properties required for end applications on worker nodes might expect properties to reside in files
 - Content, location mapping
 - Some property data will likely be replicated
 - both RDB and files
- Identification of the different properties is a first step in implementing datasets



Implementation (cont)



- Interface for schedulers
 - Central problem is how to split dataset
 - Account for
 - Data location
 - Compute cycle locations
 - Matching these
- Interactive analysis (*fast response*) is a special challenge
- How can different dataset providers share a scheduler?
- Options:
 - Common dataset interface which provides the required information
 - Logical file mapping or
 - Composite structure (sub-datasets)
 - Each provider also provides a service to do the splitting

