

# Distributed Data Analysis in ATLAS

US ATLAS core/grid software meeting

David Adams

BNL

May 6, 2002

# Contents

- Definition
- Request processing
- Production *vs.* analysis
- Applications

# Definition

## Distributed

- User and data at different sites
  - In general data may distributed over multiple sites

## Analysis

- Interactive configuration and response
- View histograms, correlation plots, events
- Partial results for long requests (1 min?)

## Data

- Summary (tuples) to EDO's

# Request processing

## Configuration

- Specify input data
- Define transformation
  - Output includes histogram and tuples

## Execution

- Locate data at sites
- Divide and assign jobs to sites
- Monitor progress
- Collect output (including partial results)

# Production vs. analysis

## Commonalities

- Distributed production and analysis share a great deal of (GRID & ATLAS) infrastructure
- May have common applications
  - *E.g. athena*

## Resource allocation

- Either dedicated analysis resources or
- Give priority to analysis activities
  - Suspend production when an analysis request is received

# Applications

## Athena

- Allows access to arbitrary EDO's through StoreGate

## ROOT

- Popular HEP tool for data analysis
- Keep the user interface
- Extend for distributed processing over the grid