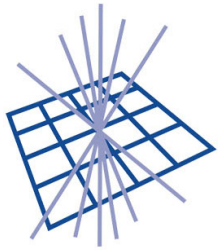
Integration of the ATLAS Tag Database with Data Management and Analysis Components

Caitriana Nicholson
University of Glasgow
3rd September 2007
CHEP, Victoria, B.C.

UNIVERSITY of GLASGOW



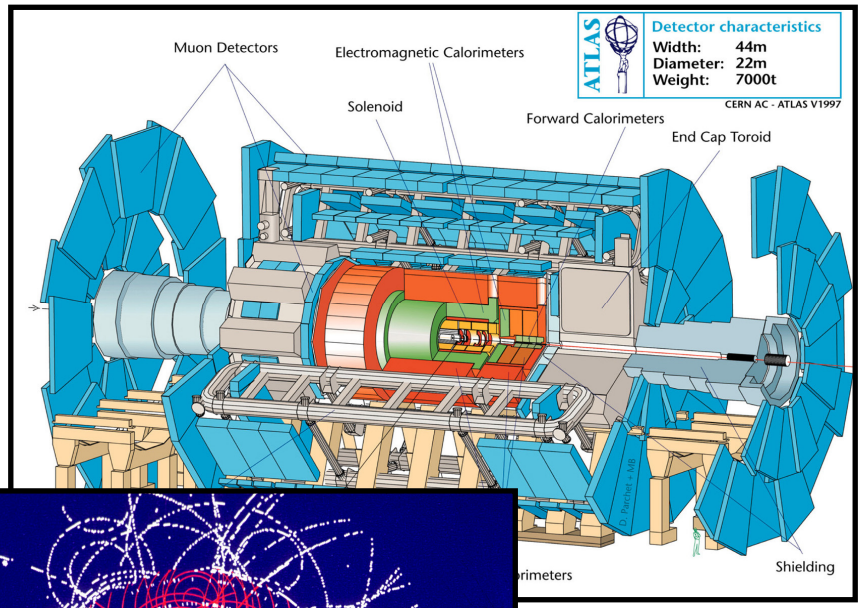
GridPP
UK Computing for Particle Physics

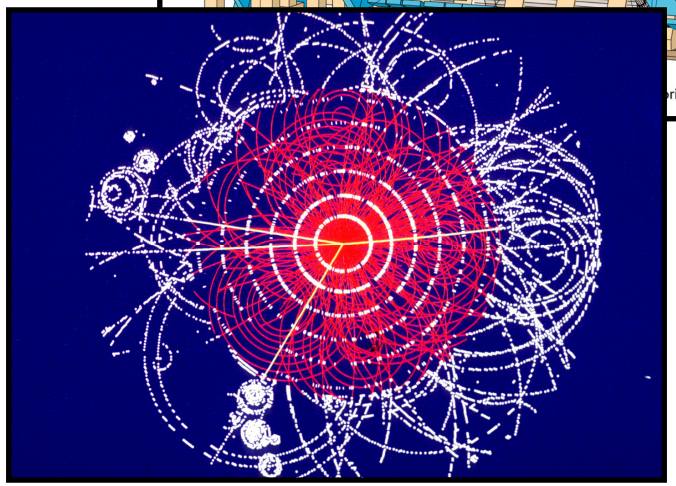
Overview

- Introduction
- ATLAS event-level metadata
- The ATLAS Tag Database
- Distributed Data Management & Analysis
- The ATLAS Tag Navigator Tool
- Some performance measurements
- Conclusions & further work

ATLAS Event Rates 101

- ATLAS raw data rate after triggers: 200 Hz
- Raw event size: 1.6 MB
- Processed to ESD (1 MB), AOD (100 kB)
- 2×10^9 events/year
- Selectivity for Higgs event: ~1 in 10^{13}





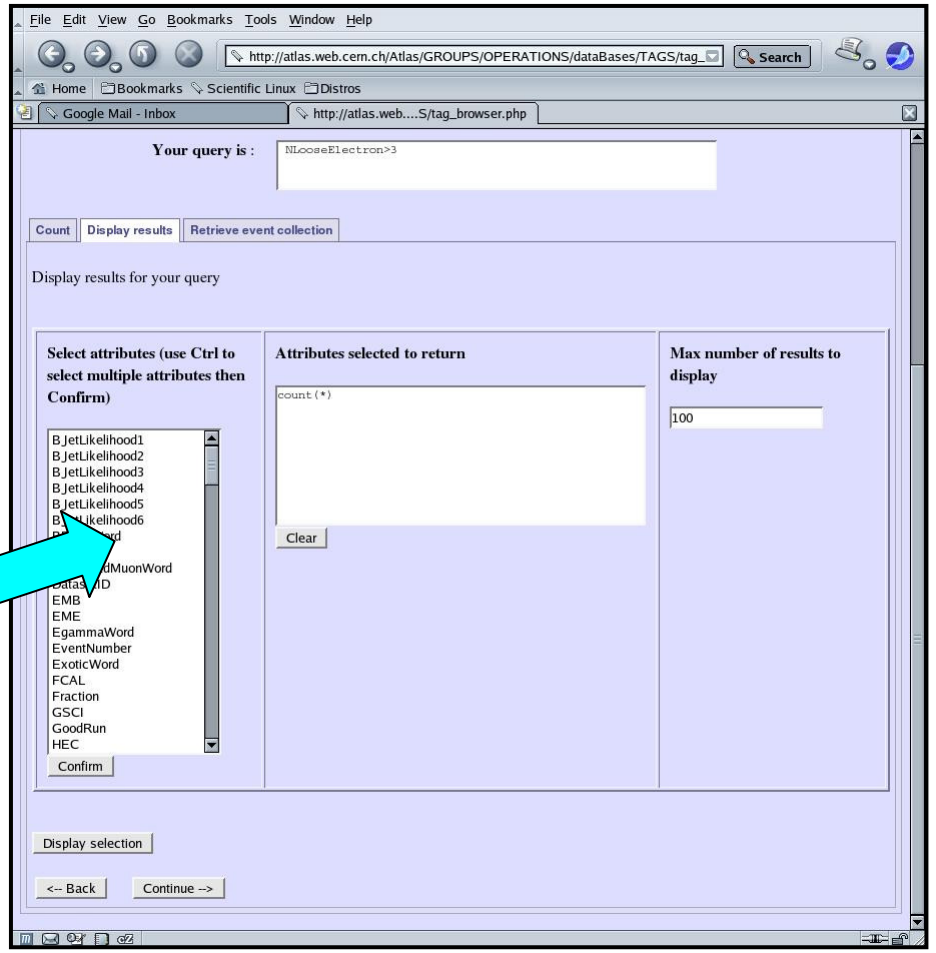
ATLAS Event-level Metadata

- "Tags" – summary physics data for events
- Allows:
  - efficient *selection* of interesting events
  - direct *navigation* to these events
- 2 formats
  - ROOT files: useful as *indices* to event
  - Relational Database: useful for *querying*
- 1 kB/event, 6 types of attribute stored
  - Includes pointers to AOD, ESD and RAW data files

The Tag Database

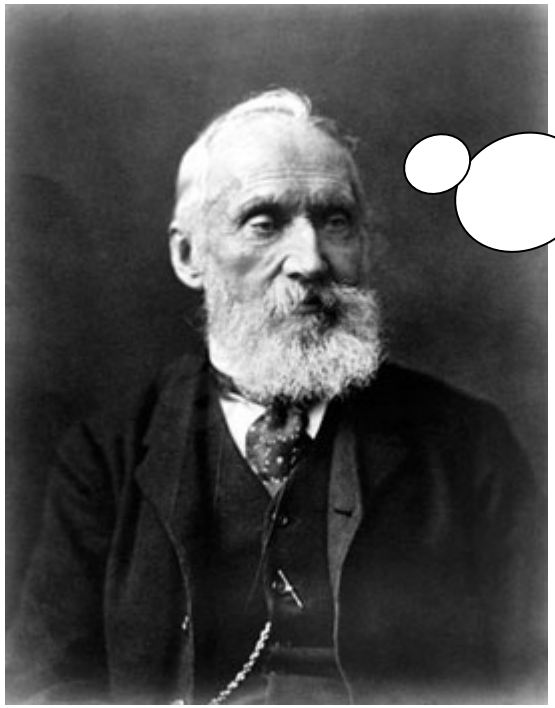
- Generated from file-based tags which are produced at reconstruction
- Global Oracle database at CERN
- Replication to other sites (Oracle / MySQL) under investigation

- Series of test databases deployed
  - Largest: 1 TB
  - Most realistic: 2 GB (+ indices)
- Prototype Web Query Browser



Talk #161 – **"Building a Scalable Event-Level Metadata System for ATLAS" -** has details of performance tests

A Physicist Use Case



Query Tag Database

Get list of events

Find correct files on grid

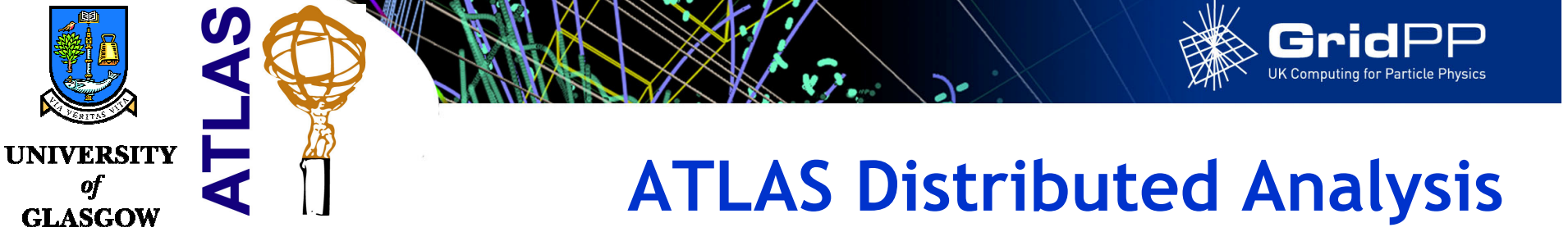
Run distributed analysis

Get output!

ATLAS Distributed Data Management

- 3 grids used by ATLAS: LCG, OSG, NDGF
- Data movement and cataloguing by Distributed Data Management system: DQ2
- Uses *dataset* as unit of data handling
- Dataset = group of files + metadata
- See talk #64: "**Managing ATLAS data on a petabyte-scale with DQ2**"

Challenge: Tag Database ignorant of datasets



ATLAS Distributed Analysis

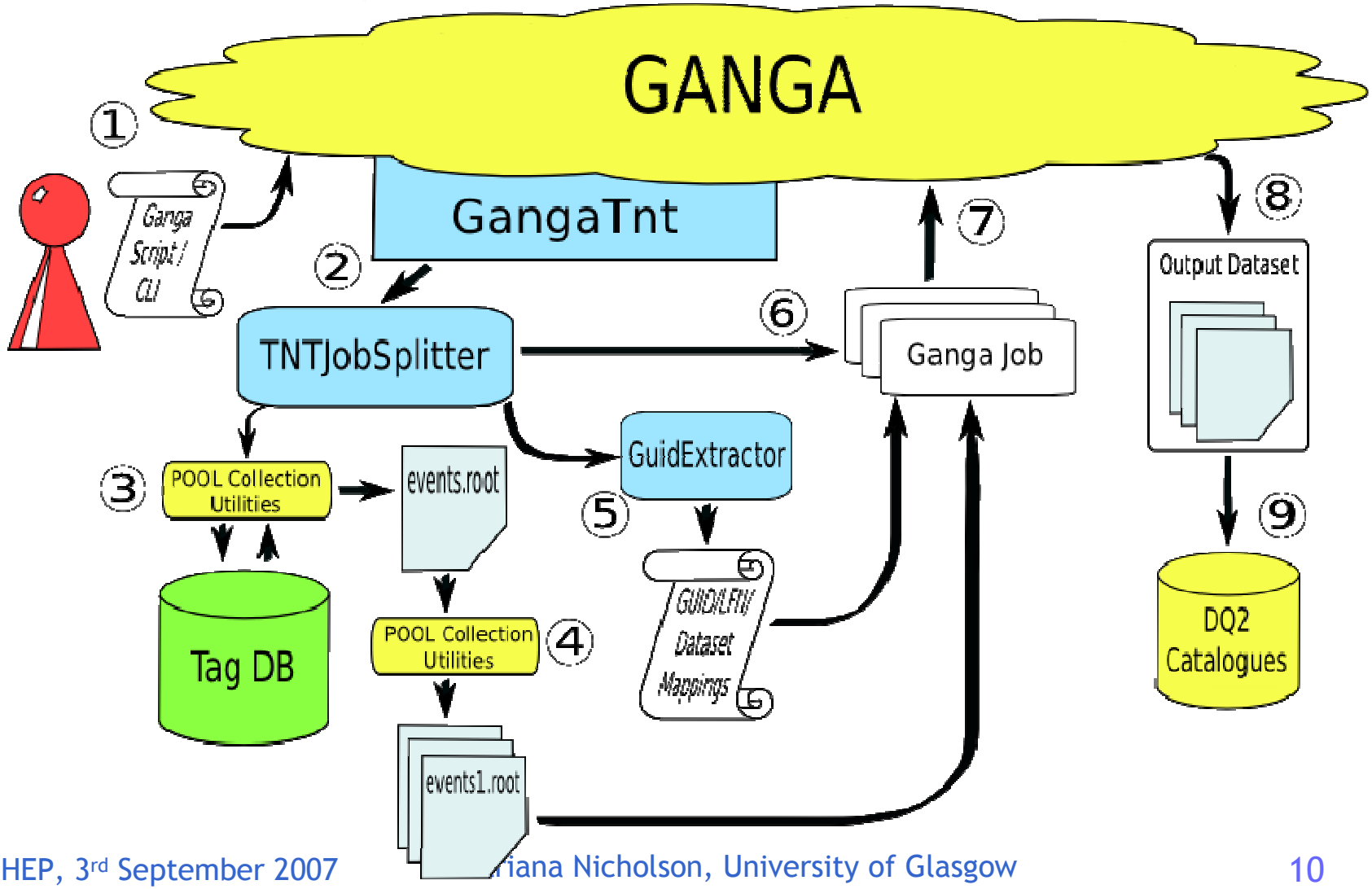
- Use grid computing for analysis, hide complexity from users
- Tools developed include
  - PanDA (Talk # 167)
  - GANGA (Talks #146, 287)
- Tag Database initially integrated with GANGA

Challenge: GANGA supported file-based tags, not relational tags

The Tag Navigator Tool

- Developed to meet these challenges
- Standalone version: Python wrapper around existing grid tools
  - Limited to LCG
- GangaTnt: plugin for GANGA
  - Modular GANGA design gave easy integration
  - Allows access to GANGA job handling
  - Allows access to other GANGA plugins

GangaTnt


GANGA
GangaTnt
1
Ganga Script / CLI
2
TNTJobSplitter
3
POOL Collection Utilities
events.root
Tag DB
4
POOL Collection Utilities
events1.root
5
GuidExtractor
GUID/LFN/ Dataset Mappings
6
Ganga Job
7
8
Output Dataset
9
DQ2 Catalogues

Some Performance Measurements

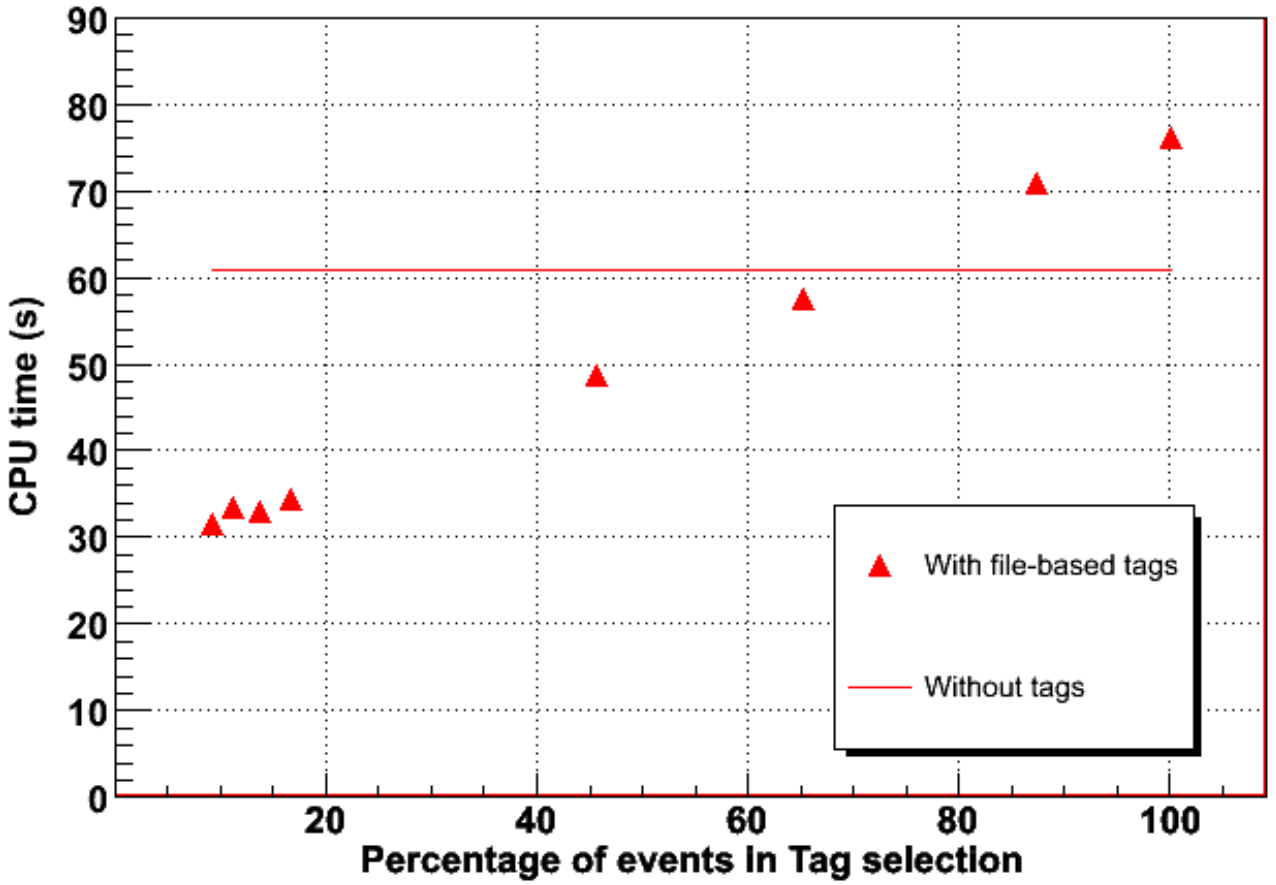
- Simple tests to give *initial* understanding of performance
  - Much more work needed!
- Example Z → e,e analysis
  - Reads electron objects from persistent storage and reconstructs invariant mass to get Z peak
- Comparison of analysis without tags, with file-based tags, with relational tags
- AOD cut (~10%) :

  2 electrons, electron $p_T > 20$ GeV, $|\eta| < 2.5$

Local Tests: single file

- AOD and Tag files on local disk (CERN lxplus)
- 5000 events
- Run Athena on whole file
  - Without tags
  - With varying percentage of events pre-selected with tag file

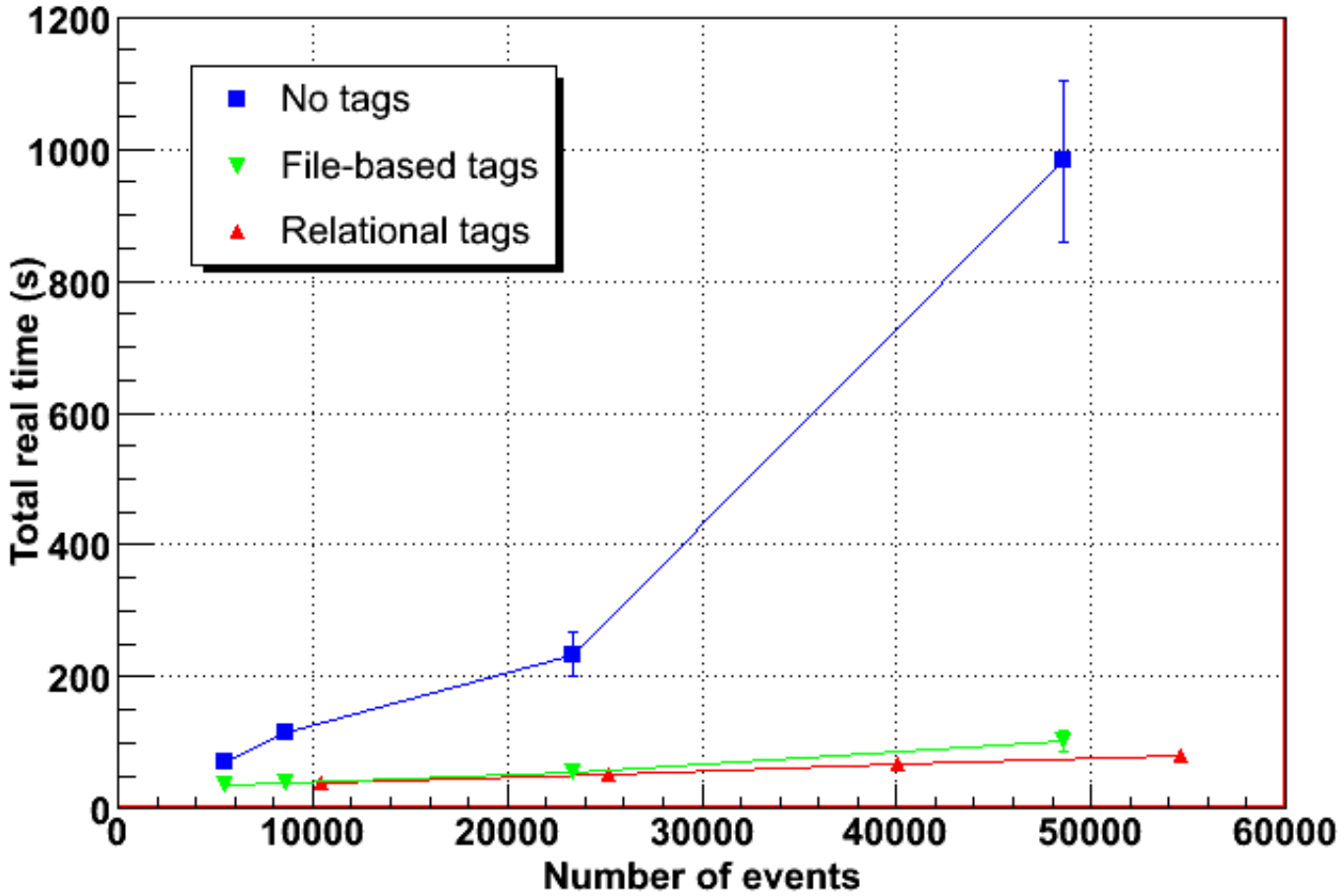
Local Tests: single file results



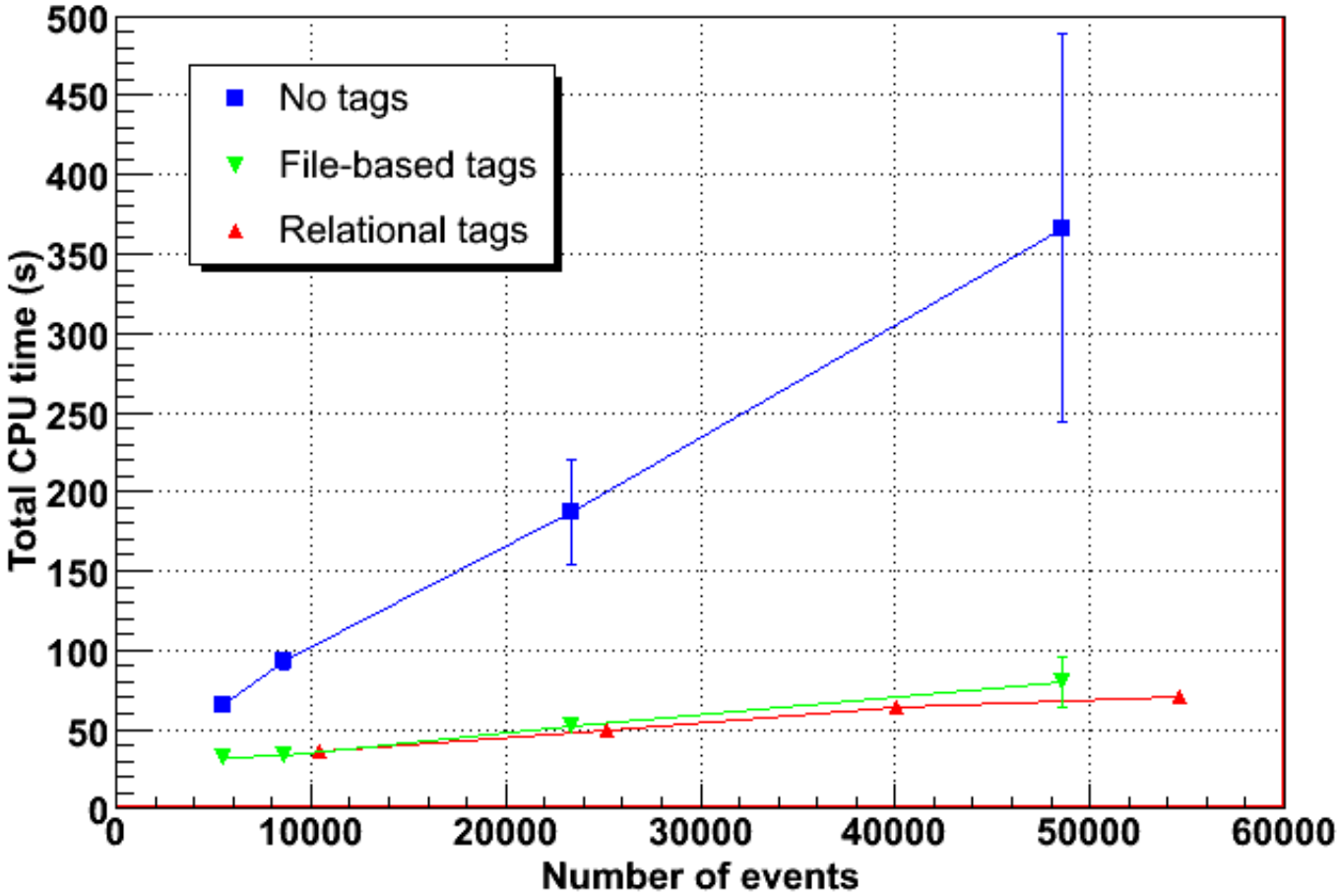
Local Tests: increasing input events

- Increasing number of input files, 4-5000 events each
- All on local disk
- Run Athena on events
  - Without tags
  - With ~10% pre-selection on file-based tags
  - With ~10% pre-selection on Tag Database
    - Query done within Athena, included in analysis time

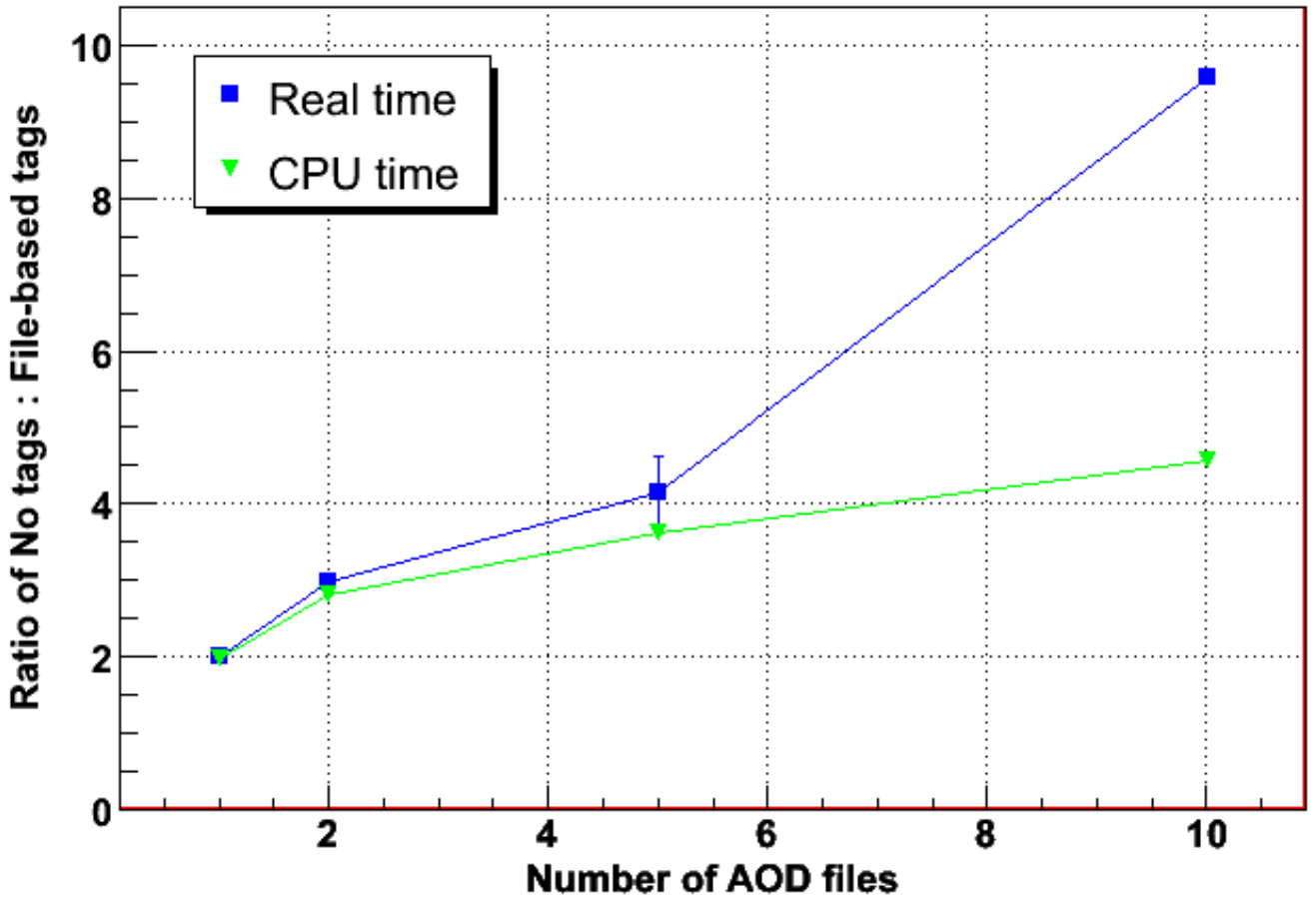
Local Tests: increasing events results



Local Tests: increasing events results



Local Tests: increasing events results



Local Test Summary

- Using tags gives significant improvement in time for tight selections
  - Using tags faster for selectivities < 60%
- As number of input events increases, performance gain from tags increases
- Little difference seen between file-based and relational tags

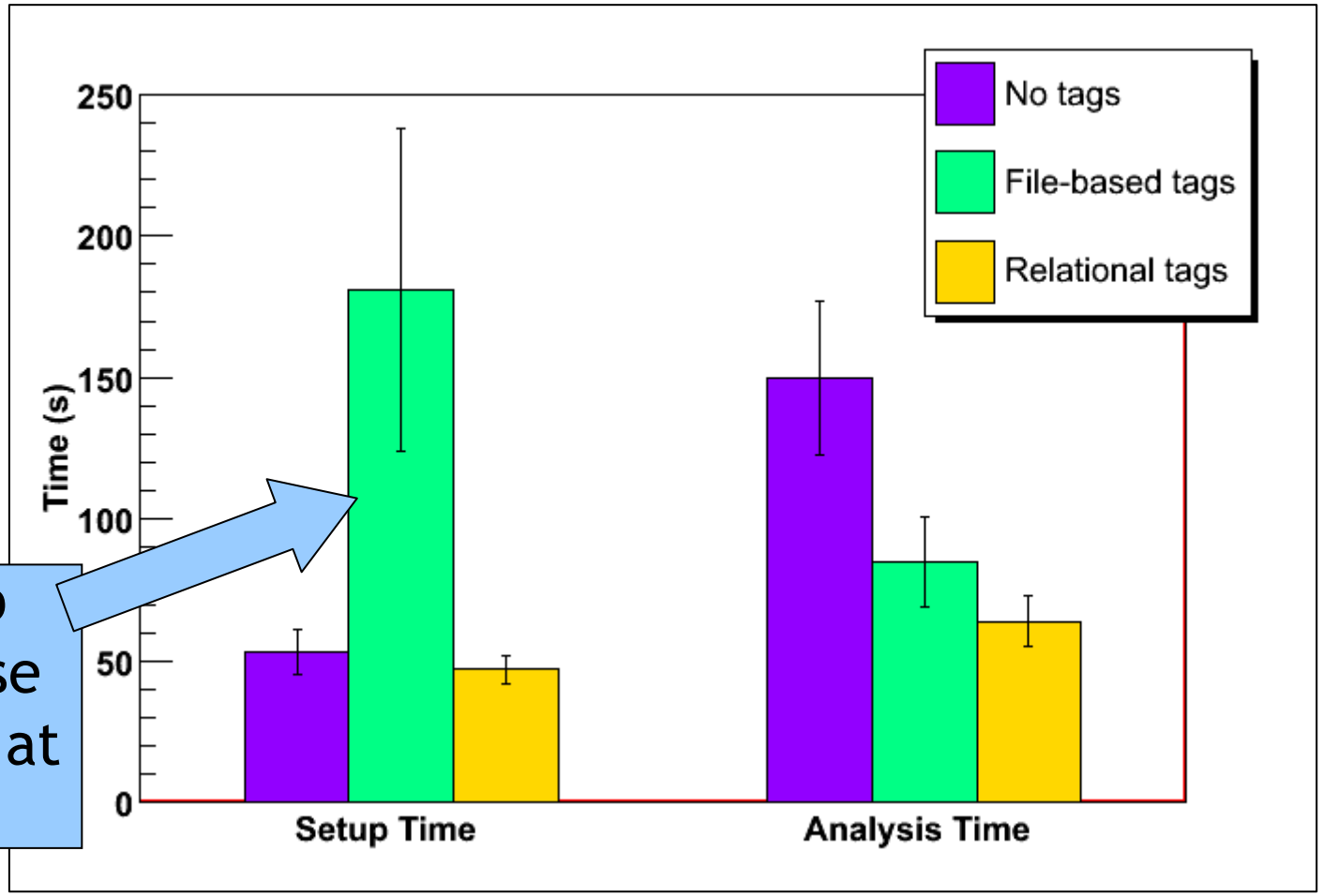
Distributed Analysis Tests

- Same analysis, run on LCG through GANGA
- 2 files from 1 AOD dataset as input
- Jobs sent to sites with the dataset
- No job splitting
- Measured on worker node:
  - Time for setup, including any data fetching (*setup time*)
  - Time for analysis to run (*analysis time*)

Distributed Analysis Tests (ii)

- Run Athena as GANGA application with
  - No tags used: AOD dataset name given to GANGA, job goes to that site
  - File-based tags: AOD and Tag dataset names given to GANGA, job goes to site with AOD
    - *Should* also have Tag dataset there.. but not yet always the case
  - Tag Database: GangaTnt used to run query and find correct dataset

Distributed Analysis Tests: results



Long setup time because tag files not at sites

Distributed Analysis Test Summary

- Setup times similar with and without tags
  - Delay with file-based tags should not be present in future
- Analysis ~ twice as fast with tags
- Little difference between file-based and relational tags
  - GangaTnt query time (few seconds) not included here
- Consistent with local analysis results
  - With larger analyses, impact of using tags will be higher
- GangaTnt and standard GANGA tag use complementary

Conclusions & Future Work

- TNT and GangaTnt enable integration of ATLAS Tag Database with Distributed Data Management and Analysis components
- Initial tests show:
  - 50% cut in analysis time for 10% selection on single file
  - Tags improve performance for selectivity up to 60%
  - Increasing performance gain for tags as input events increase
- Further work needed to understand:
  - Tag use with larger numbers of events
  - Effects of file I/O
  - Differences between file-based and relational tags
- Tag Database will continue to grow and GangaTnt will continue to develop

UNIVERSITY of GLASGOW

ATLAS

GridPP
UK Computing for Particle Physics

Backup Slides

Tag Content

- Budget: 1 kB / event
- 6 groups of attributes:
  - Event quantities: run number, event number, luminosity…
  - Data quality: detector status, “good for physics”…
  - Physics objects: e, μ, τ, jets
  - Physics/Performance Group attributes
  - Trigger information
  - Pointers to event data: AOD, ESD & RAW refs, software version…