

The problem: analysis of Large Hadron Collider data

- The Large Hadron Collider is a proton-proton collider at CERN
 - 4 big international experiments measure the collisions (ATLAS, CMS, LHCb, ALICE)
- LHC data in numbers (e.g. ATLAS experiment)
 - 1 billion proton-proton collisions every second
 - Nominal output rate of detector: 68 TB/s -> impossible to record
 - Actual output rate to disk: 1.5 GB/s (reduced via fast identification of “interesting” events)
 - Data rate of up to 100TB per day, for up to 6 months per year, for 10-15 years.
 - Over 3000 physicists want access to any of that data at any time

The solution: the Worldwide LHC Computing Grid

- The LHC Computing Grid is:
 - Distributed computing structure comprising 170 computing centres in 42 countries
 - Experimental data processed and copied to multiple locations
 - Users access the data from their experiment at any location using certificates signed by virtual organisations
 - 2 million jobs run on the Grid every day
- The Manchester Tier-2 centre was the third site to join CERN datagrid in 2001.
 - 4300 job slots and 2.1 PB of storage
 - 90% used for processing ATLAS and LHCb experiments' data...
 - but 10% reserved for other scientific research (i.e. BioMed)
 - Staff have leading roles in LHC experiments strategy for distributed computing and Worldwide LHC Computing Grid coordination