The Grid Workloads Archive
Alexandru Iosup, Dick Epema (TUDelft) Catalin Dumitrescu (UChicago) Hui Li, Lex Wolters (ULeiden) Mathieu Jan (INRIA)

"Easy access to grid workload traces."

Why do we need grid traces?
- Understand how real grids are used
- Address the challenges facing grid resource management research and practice
- Develop and test grid resource management solutions
- Perform realistic simulations

The GWA, a potential solution
The Grid Workloads Archive (GWA) is at the same time a workload data (trace) repository and a community center for resource management and scheduling in grids.

- Standard data format (GWF)
  - Share traces with the community
  - Use extensions for specific modeling aspects
  - Text-based, easy to parse for custom tasks
  - Additional SQL-compatible data format (GWF-SQLite)

- Automated trace analysis
  - Provide ready-to-use tools to the community
  - Promote results availability and comparability

- Automated trace ranking
  - Help non-experts with their trace selection process

Community effort
We have already received contributions from a number of grid owners and/or operators, which we gratefully acknowledge and thank.

We need your help!
We are looking for one-time contributors, that is, contributors who agree to share their old grid traces with the community, and, most importantly, for monthly contributors, that is, contributors who agree to send current trace data monthly.

Use GWA traces in your work!

Timeline
- 2006
  - GWA idea
  - Mar
  - Dec
- 2007
  - GWA draft
  - GWA online
  - 10 traces online
  - 20 traces online
  - Mar
  - Dec
- 2010
  - March

http://gwa.ewi.tudelft.nl/