



**EGEE Conference
Istanbul, september 25, 2008
Dick Epema
Delft University of Technology**

Main contributor: Alexandru Iosup

Other contributors:

- Hui Li and Lex Wolters, Leiden University, The Netherlands
- Catalin Dumitrescu, Mathieu Jan, Shanny Anoep, TU Delft



Contents

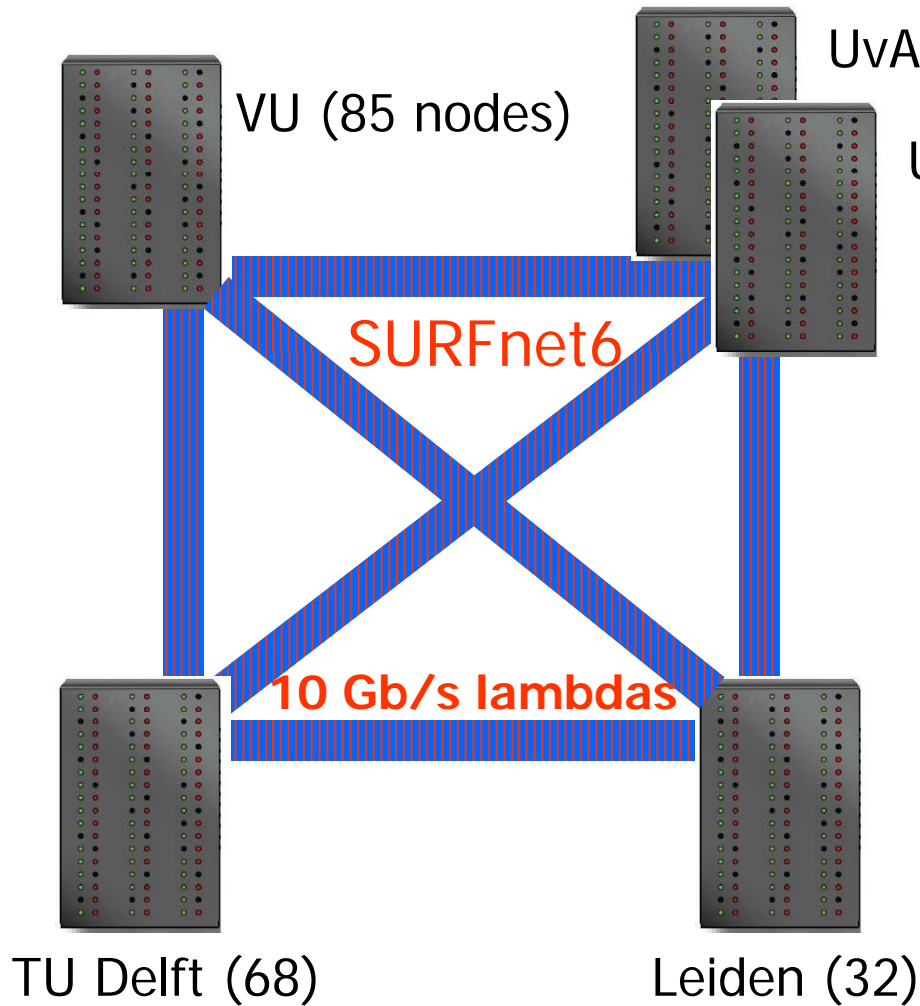
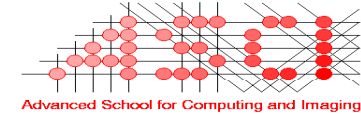
- The Context of Grid Research in Delft
- Why a Grids Workloads Archive (GWA)?
- Requirements for a GWA
- Design of the GWA
- Contents of the GWA
- Conclusions

Context 1: The VL-e project

- A 22-MEuro grid project in the Netherlands (2004-2009)
- **Aim:** design, build, and test a **Virtual Lab** for **digitally enhanced science (e-science)** experiments (no in-vivo or in-vitro, but **in-silico** experiments)
- University partners:
 - University of Amsterdam
 - Vrije Universiteit Amsterdam
 - Delft University of Technology
- Research institutes:
 - CWI (Math&CS), NIKHEF and AMOLF (Physics), ...
- Companies:
 - Unilever (e-food), Philips (e-health), IBM (e-*) , ...



Context 2: DAS-3



Operational: oct. 2006

272 AMD Opteron nodes
792 cores, 1TB memory
More heterogeneous:
2.2-2.6 GHz
single/dual core nodes
Myrinet-10G (excl. Delft)
Gigabit Ethernet

Context 3: The KOALA Co-Allocating grid scheduler

- Main goals:
 1. **processor co-allocation**: spread jobs across clusters
 2. **data co-allocation**: move large input files to the locations where the job components will run prior to execution
 3. **load sharing**: in the absence of co-allocation
 4. **run alongside local schedulers**
- **KOALA**
 - is written in Java
 - is middleware-independent
 - has been deployed on the DAS2 in september 2005
 - currently runs on the DAS3



Goals of the GWA

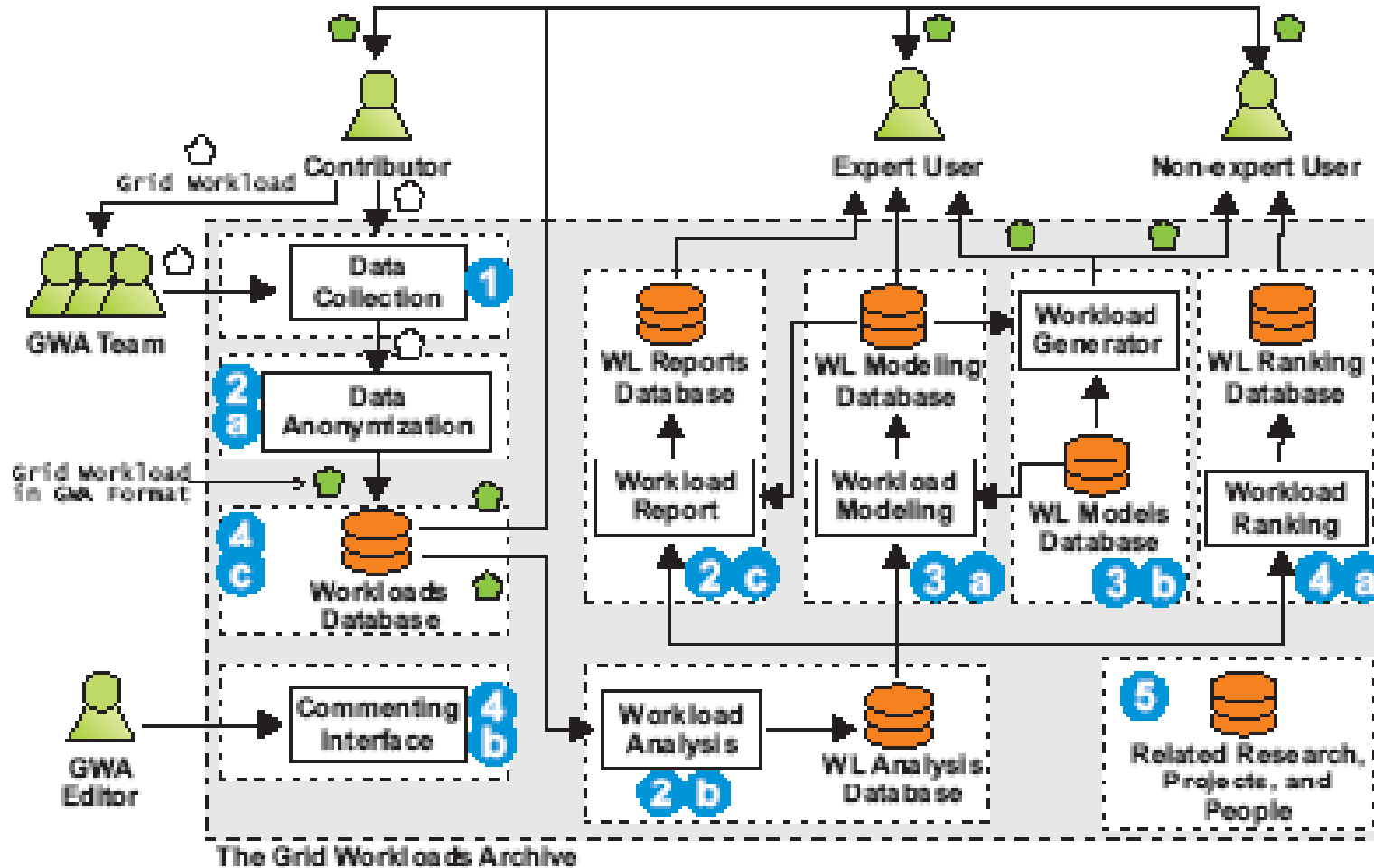
- **Research in grid resource management:**
 - **how are grids actually used?**
 - realistic input to performance experiments (=simulations and actual runs in real systems)
 - reproducibility of performance experiments
 - comparability of performance experiments
- **Aiding in grid design and dimensioning**
- The GWA is modeled after Dror Feitelson's **Parallel Workloads Archive (PWA)**

Requirements for a GWA

1. Tools for collecting grid workloads
 - common grid workload format
2. Tools for grid workload processing
 - proper way of anonymization
 - statistical analysis+modeling
3. Tools for using grid workloads
 - generate synthetic workloads
4. Tools for sharing grid workloads
 - trace ranking and selection
5. Community building tools
 - bibliography
 - list of projects and people using the GWA

The GWA is not simply a trace repository!!

The GWA: Overview of the Design



The GWA: Standard Workload Format

- Extends the PWA format with grid-related information, e.g.:
 - submission site
- Document on GWA web site describing the details
- Is extensible
- Traces also available in SQLite format (for easy extraction/querying)

The GWA: Trace ranking and selecting

- Purpose: to help users/researchers select an appropriate trace
- Ranking with respect to six properties
- Every trace is categorized

Category	Sample	*	**	***	****	*****
System Sites	-	1	2-5	6-10	11-20	>20
System Cores	0-100	101-1k	1k-5k	5k-10k	10k-25k	>25k
No. Users	0-50	51-100	101-200	201-500	0.5k-1k	>1k
No. Jobs	0-15k	15k-100k	100k-200k	200k-500k	500k-1M	>1M
Utilization	0-10%	11-20%	21-40%	41-60%	61-75%	>75%
Reported work	0	1	2-5	6-10	11-20	>20

The GWA: Contents (1)

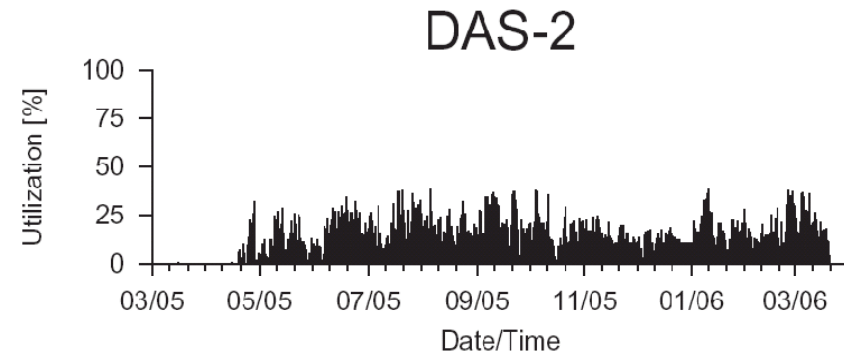
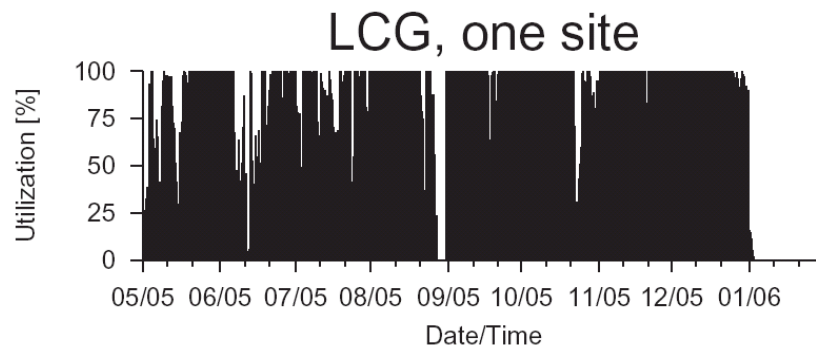
- **Production Grids:** e.g., LCG, Grid3, TeraGrid
- **Academic Grid:** DAS2, Grid'5000
- **Features (as of mid 2007):**
 - long traces (6+ months),
 - active environments (500,000+ jobs per trace, 100s of users)
 - in total: 2000 users, 7 million jobs, 13 operational years, 130 sites, 10,000 CPUs
- For every trace, there is a **summary report** (and a script generating it)

The GWA: Contents (2) (sample)

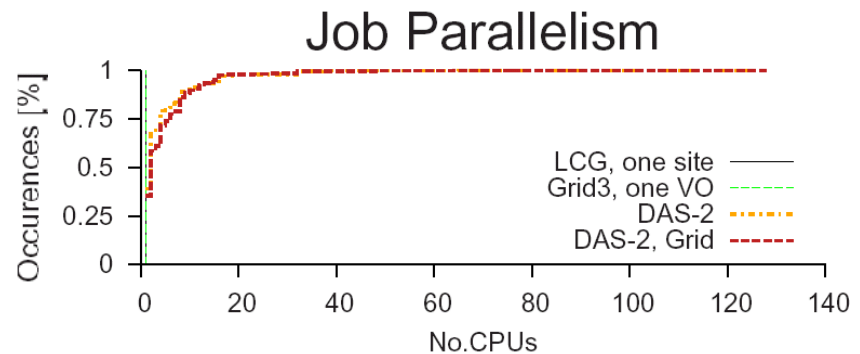
system	period	sites	CPUs	jobs
DAS-2	02/05-03-06	5	400	602K
Grid'5000	05/04-11/06	15	2500	951K
Nordugrid	05/04-02/06	75	2000	781K
LCG	05/05-01/06	1	880	1.1M
TeraGrid	08/05-03/06	1	96	1.1M

The GWA: Trace Analysis--System-Wide Characteristics

- System utilization is on average 60-80% for **production** Grids, and 10-20% for **academic** Grids

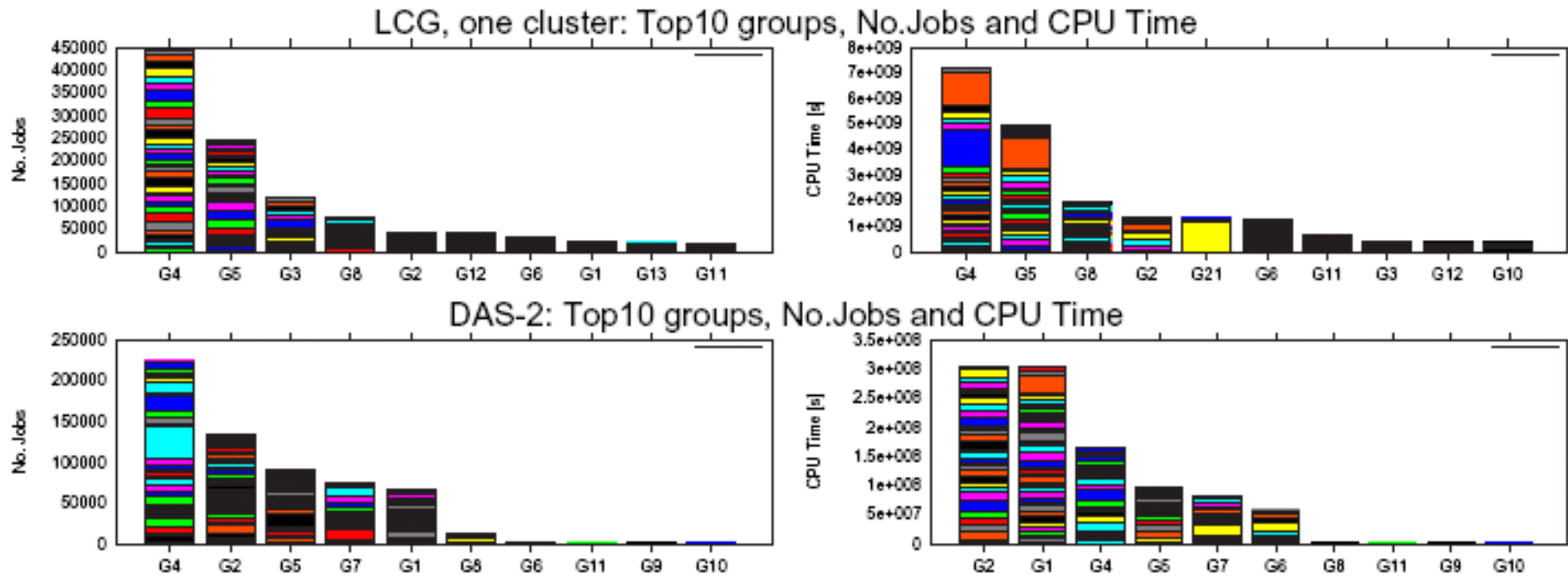


- In some grids, the average job size is **1** (that is, there are **no [!]** tightly-coupled, parallel jobs)



The GWA: Trace Analysis-- VO, Group, and User Characteristics

- Top 2-5 groups/users dominate the workload
- Top groups/users are constant submitters
- The week's top group/user is not always the same



Conclusions

- The GWA is one component of a **larger tool set** (e.g., GrenchMark and DGSim)
- Has been very useful in understanding grid usage
- Does provide **much more** than only traces
- There is some research that does use the GWA
- Obtaining traces is no trivial matter

More information

- **Grid publications PDS group**
 - see PDS publication database at www.pds.ewi.tudelft.nl or www.pds.ewi.tudelft.nl/~epema
- **GWA publications** in Grid2006 and Future Generation Computer Systems (2008)
- **Web sites:**
 - GWA: gwa.st.ewi.tudelft.nl
 - KOALA: www.st.ewi.tudelft.nl/koala
 - DAS3: www.cs.vu.nl/das3
 - VL-e: www.vl-e.nl

Contributions and questions to:
Alexandru Iosup (gwa@tudelft.nl)

Core **GRID**

