The EU DataGrid

How the use of distributed resources can help solve complex problems

David Groep, NIKHEF
davidg@nikhef.nl

Zeger Hendrikse, UvA IvI
zegerh@science.uva.nl

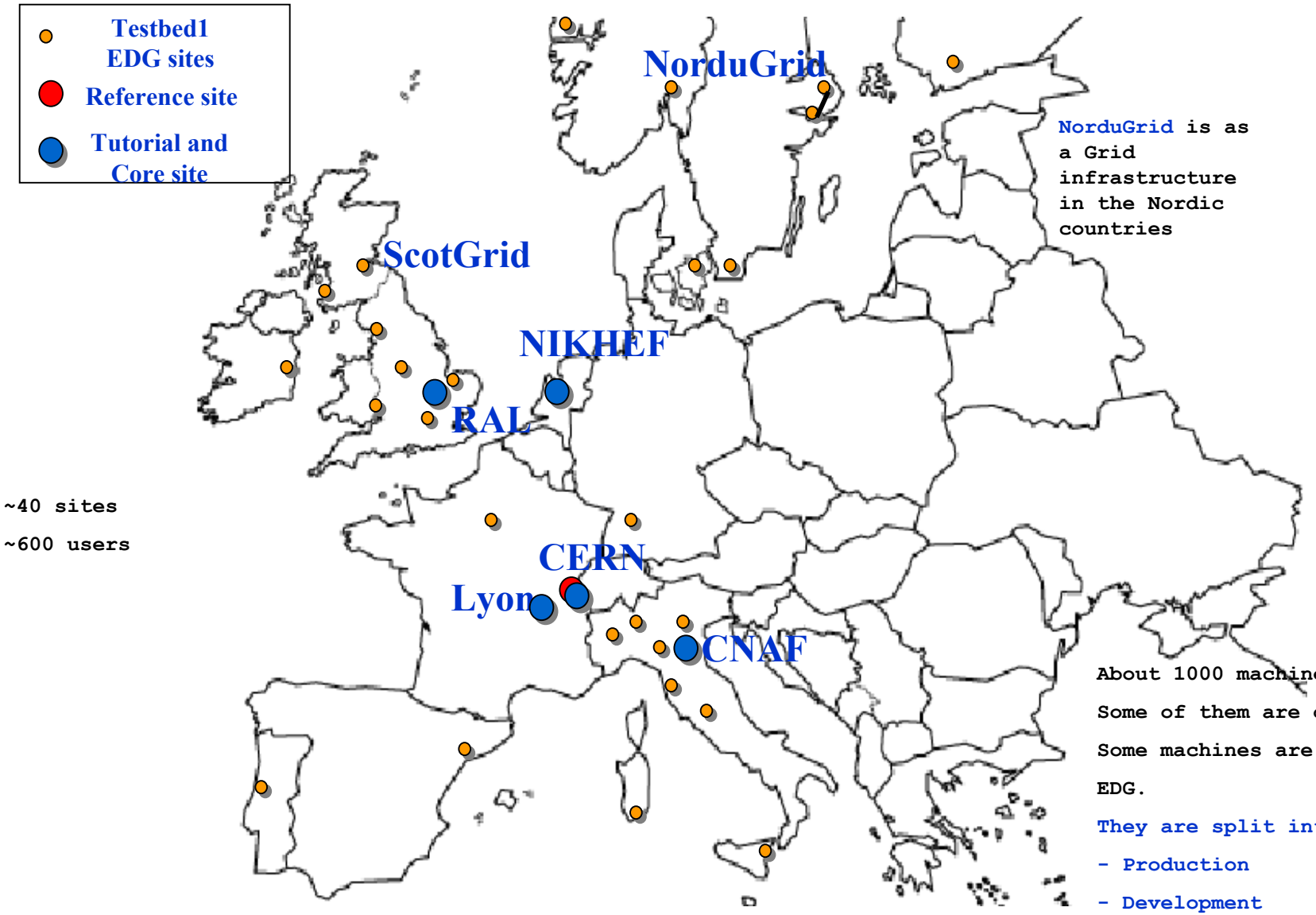
The EU DataGrid Project

- Largest operational grid infrastructure
- 21 partners in EDG project
- More partners in related projects (LCG, CrossGrid)

- Specifically looking at *deployable* grids
  - 40 sites, 600 users
  - Interconnected with high-speed networks

- Europe-wide resource distribution … you'll see in the demo

Current EU DataGrid Testbed



NorduGrid is as a Grid infrastructure in the Nordic countries

~40 sites

~600 users

About 1000 machines are available.

Some of them are dual processor.

Some machines are only used for EDG.

They are split into:

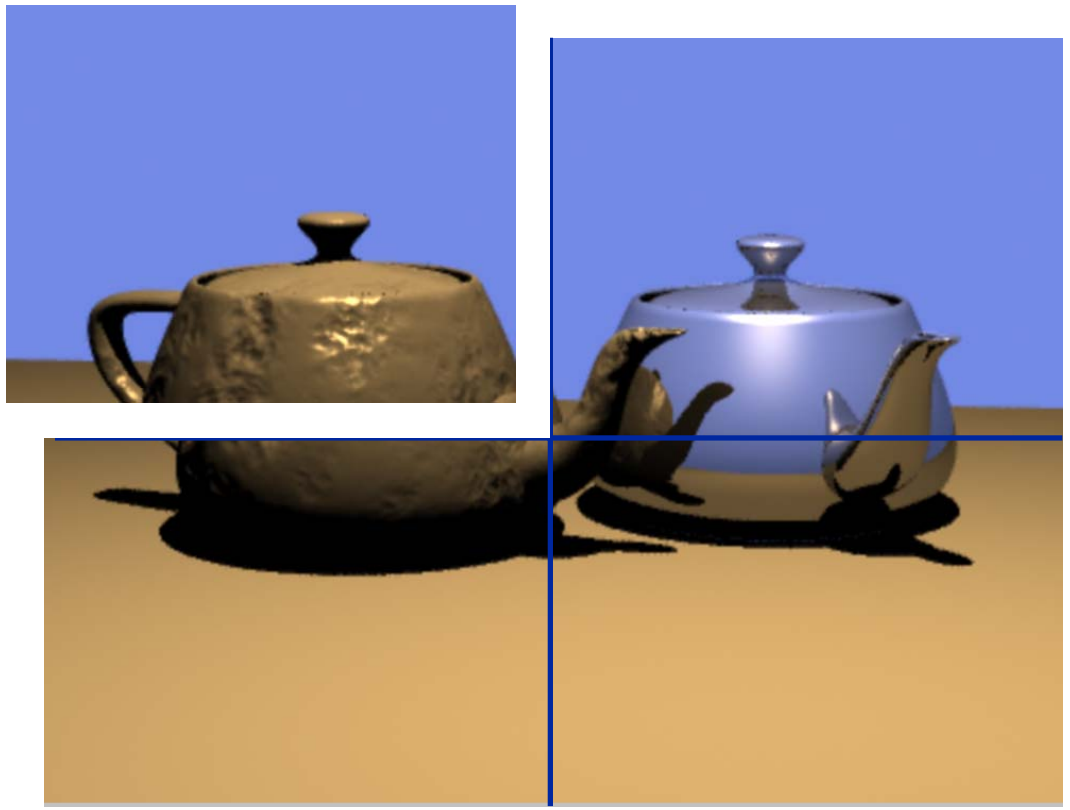
- Production
- Development

Services Used in this Demonstration

- User Interfaces
  - NIKHEF Netherlands
- Resource Brokers
  - Imperial College, UK; CCIN2P3 Lyon, FR; CNR Roma, IT
- Information Services
  - RAL, UK
- Compute services
  - Anywhere in Europe, e.g. NIKHEF, Glasgow, Barcelona, Catania, …
- Storage Elements
  - *not used in this demo*

Our (simple) example problem

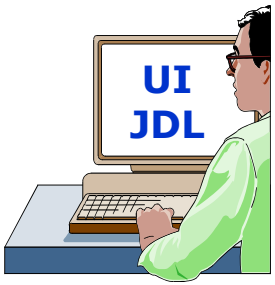
- Render a complex image
- *e.g. for use in animation movies*
- using many resources
  (to get speedup)



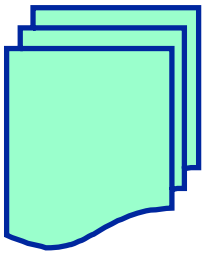
- Rendering package: **Rendrib**
- Method
  - Split image in NxN parts
  - Distribute N^2 jobs over the Grid
  - Combine parts on the
    User Interface afterwards
- Using the EDG Application Test Bed

This job does not need any staged input data, but more realistic jobs will

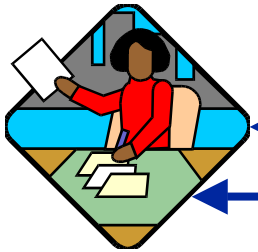
The job flow through the system



Replica Catalogue (RC)



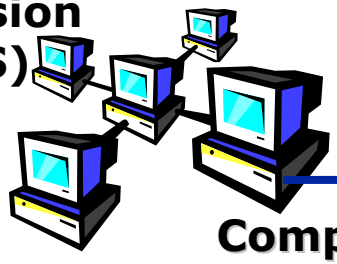
Information Service (IS)



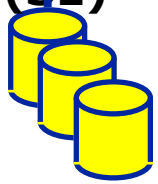
Resource Broker (RB)

Logging & Bookkeeping (LB)

Job Submission Service (JSS)

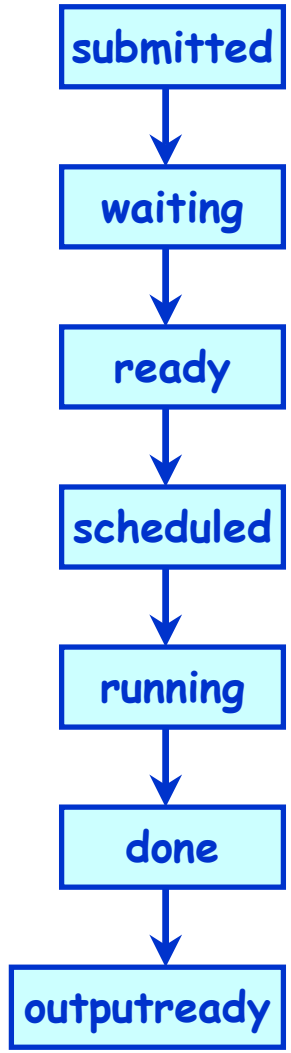


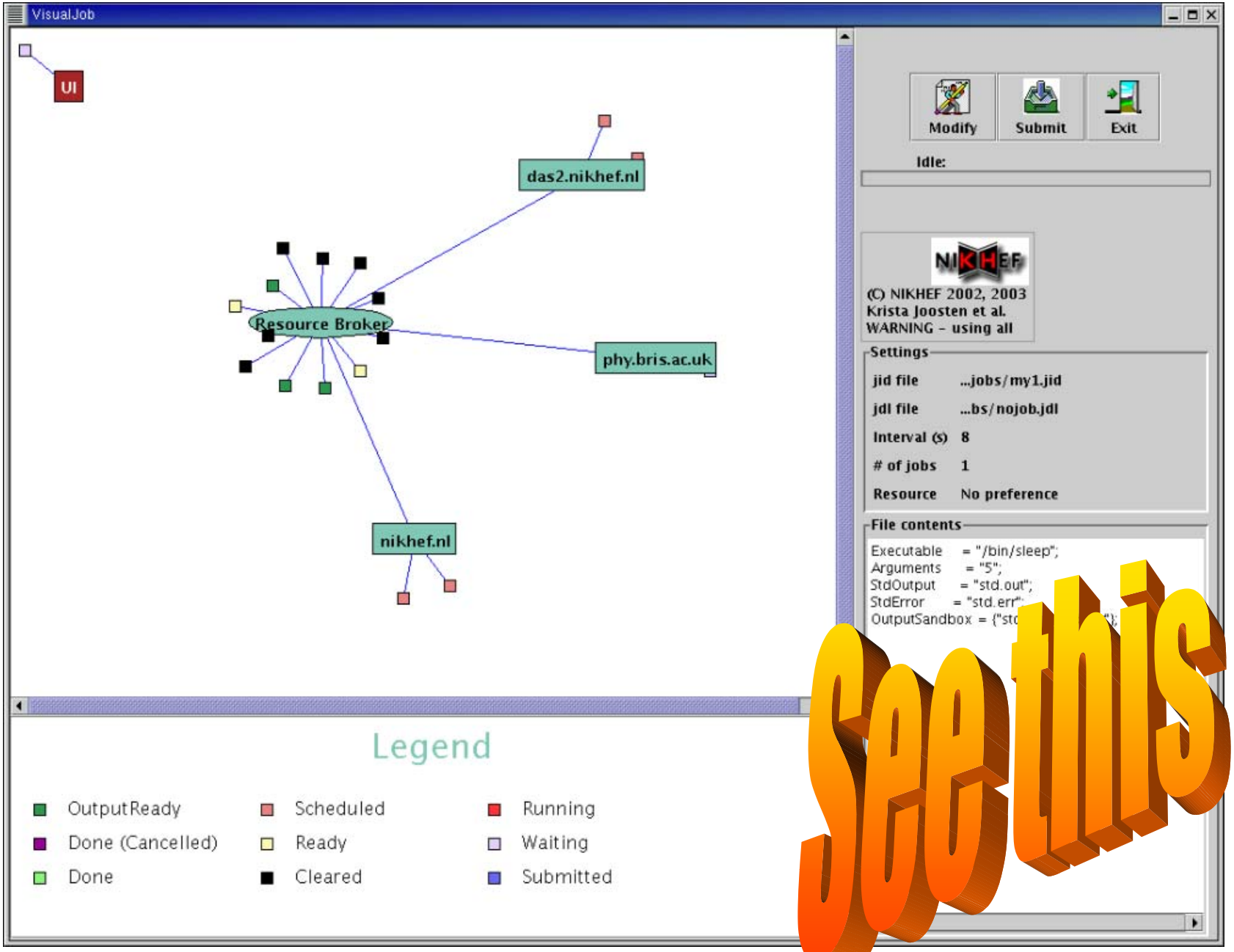
Compute Element (CE)



Storage Element (SE)

Job Status




VisualJob
UI
das2.nikhef.nl
Resource Broker
phy.bris.ac.uk
nikhef.nl
Modify
Submit
Exit
Idle:
NIKHEF
(C) NIKHEF 2002, 2003
Krista Joosten et al.
WARNING - using all
Settings
jid file ...jobs/my1.jid
jdl file ...bs/nojob.jdl
Interval (s) 8
# of jobs 1
Resource No preference
File contents
Executable = "/bin/sleep";
Arguments = "5";
StdOutput = "std.out";
StdError = "std.err";
OutputSandbox = {"std
Legend
OutputReady
Done (Cancelled)
Done
Scheduled
Ready
Cleared
Running
Waiting
Submitted

see this live

What else could we have shown?

- This demonstration showed only “simple” jobs
- There are more enhanced EU DataGrid capabilities
  - **Intelligent data replication**
  - **Job migration to data**
  - **Inclusion of network costs**
- Scales to many users in a production environment

The EU DataGrid project web	*www.edg.org*
DutchGrid Platform	*www.dutchgrid.nl*
DutchGrid Forum	*www.dutchgridforum.org*
For other grid projects, see	*www.gridstart.org* *www.enterthegrid.com*

Acknowledgements

- Roberto Puccinelli
- Heinz Stockinger
- Flavia Donno
- *… and all site managers & ITeam*