Eine Einführung ins Grid

Andreas Gellrich

DESY Hamburg
10.06.2008
The Power Grid
The Grid Dream

Mobile Access

Desktop

Visualizing

Supercomputer, PC-Cluster

Data Storage, Sensors, Experiments

Internet, Networks
Grid Computing

Grid Computing is about *virtualization* of *global* resources.

- It is about *transparent* access to globally distributed resources such as data and compute cycles

- A Grid infrastructure consists of *services* to access resources and (of course) of the *resources* itself
  - Opposite to *distributed computing*, Grid resources are *not centrally controlled*
  - Hence it is mandatory to use *standard, open, general-purpose protocols and interfaces*
  - A Grid must *deliver nontrivial qualities of services*

- In general Grid infrastructures are *generic*; without any dependencies of the applications / experiments
Grid Types

- **Data Grids:**
  - Provisioning of transparent access to data which can be physically distributed within *Virtual Organizations* (VO)

- **Computational Grids:**
  - allow for large-scale compute resource sharing within Virtual Organizations (VO)

- **(Information Grids):**
  - Provisioning of information and data exchange, using well defined standards and web services

Jobs are *transient*; data is *persistent*. 
A Virtual Organization (VO) is a dynamic collection of individuals, institutions, and resources which is defined by certain sharing rules.

- A VO represents a collaboration
- Users authenticate with personal certificates (Authentication)
- Users are members of a certain VO (Authorization)
- Certificates are issued by a Certification Authority (CA)

• Grid Infrastructure
  - Core Services (mandatory per VO)
    - VO Membership Services
    - Grid Information Services
    - Workload Management System
  - Resources (brought in by partners (Grid sites))
Overview
Services View

UI

ssh

certs

usercert.pem / userkey.pem

VOMS → Grid Services

grid-mapfile

WMS

JDL

output

LFC

BDII

VO exclusive

SE

CE

GRIS

GIIS

R-GMA

store

WN

WN

WN

WN

WN

WN

SW

Site1

Site2

Site3

world

Grid Services View
VO & Grid Resources @ DESY

- **VOs hosted at DESY:**
  - Global: ‘hone’, ‘ilc’, ‘zeus’
  - Regional: ‘calice’, ‘ildg’
  - Local: ‘desy’, ‘hermes’, ‘icecube’

- **VOs supported at DESY:**
  - Regional: ‘dech’

- **Grid Computing Resources at DESY: (CE)**
  - grid-ce3.desy.de 1006 cores @ 229 hosts
  - hera-ce0.desy.de 124 cores @ 62 hosts (‘hone’, ‘hermes’, ‘zeus’)

- **Grid Storage Resources at DESY: (SE)**
  - grid-se3.desy.de 6TB test system
  - srm-dcache.desy.de O(50 TB) w/ tape backend
  - dcache-se-atlas.desy.de O(200 TB) w/ tape backend
  - dcache-se-cms.desy.de O(200 TB) w/ tape backend
Operations View

UI

ssh

certs

usercert.pem / userkey.pem

VOMS

Grid Services
grid-mapfile

WMS

JDL

LFC

BDII

VO exclusive

CE

Torque/Maui

GIIS

SE

R-GMA

world

WN

SW

DESY-HH

DESY, 10.06.2008
jobs are *transient*
data is *persistent*
The LHC Computing Model

- One bunch crossing per 25 ns
- 100 triggers per second
- Each event is ~1 Mbyte

US Regional Centre
Italian Regional Centre
French Regional Centre
GridKa Regional Centre

DESY ~1 TIPS
Tier 2 Centre ~1 TIPS
Tier 2 Centre ~1 TIPS
Tier 2 Centre ~1 TIPS

Physicists work on analysis “channels”
Each institute has ~10 physicists working on one or more channels
Data for these channels should be cached by the institute server