Introduction to GRID computing

Stefano Cozzini

DEMOCRITOS
DEmocritos MOdeling Center for Research In aTOmistic Simulation INFM

This first introduction

- What is GRID computing ?
- GRID computing: a short history
- The elements of grid computing:
  - Middleware
  - Applications
- examples of GRID project/infrastructures

A first definitions

- *A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities.*
  - Carl Kesselman,Ian Foster in "The Grid: Blueprint for a New Computing Infrastructure" 1998
- *Grid computing is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations"*
  - Carl Kesselman,Ian Foster in "the anatomy of the grid" 2000

A GRID checklist (Ian Foster 2002)

- a Grid is a system that:

  1) coordinates resources that are not subject to centralized control ...

     (Otherwise, we are dealing with a local management system.)

  2)..using standard, open, general-purpose protocols and interfaces...

     (Otherwise, we are dealing with an application specific system.)

  3) ...to deliver nontrivial qualities of service.

     (It should meet complex user demands, so that the utility of the combined system is significantly greater than that of the sum of its parts.)

A few concepts in GRID COMPUTING

- Resources are locally managed and controlled
- Different resources can have different policies and mechanism
  - Computing resources managed by different batch system
  - Different storage system on different node
  - Different policies granted to the same user on different resources on the GRID
- Dynamic nature:
  - Resources and users can change frequently

history of GRID computing...

(http://www.cio.com/archive/080105/grid_sidebar_two.html)

- 1960s: Distant Relatives
  - In 1965, the developers of an operating system called Multics (Multiplexed Information and Computing Service, an ancestor of Unix) presented a vision of "computing as a utility," which is similar to grid computing today
- 1970s: The Birth of Grid
  - According to Grid.org, when computers were first linked by networks, the idea of harnessing unused CPU cycles was born. A few early experiments included a pair of programs called Creeper and Reaper that ran on the ARPAnet (the precursor to the Internet).

history of GRID computing...

(http://www.cio.com/archive/080105/grid_sidebar_two.html)

- 1980s: Grid Refined
  - Scientists used grid computing to connect multiple workstations, which allowed them to work on complicated math problems and software compilations, utilizing idle CPUs to reduce processing times.
- 1996:Free Grid!
  - The Globus Alliance formed to conduct R&D for the technology, standards and systems that form the grid. Alliance members eventually produced open-source software: Globus Toolkit

history of GRID computing...

(http://www.cio.com/archive/080105/grid_sidebar_two.html)

- 1997:The First on the Net
  - Distributed.net became the first general-purpose grid-computing network on the Internet, according to Grid.org. Distributed.net eventually brought thousands of people together to crack cryptographic challenges in a distributed environment.
- 1999: SETI, Phone Home
  - The SETI@home project launched at the University of California at Berkeley. It uses Internet-connected computers in the Search for Extraterrestrial Intelligence. Anyone who has an Internet connection and some spare CPUs can participate by running a free program that analyzes radio telescope data.

history of GRID computing...

(http://www.cio.com/archive/080105/grid_sidebar_two.html)

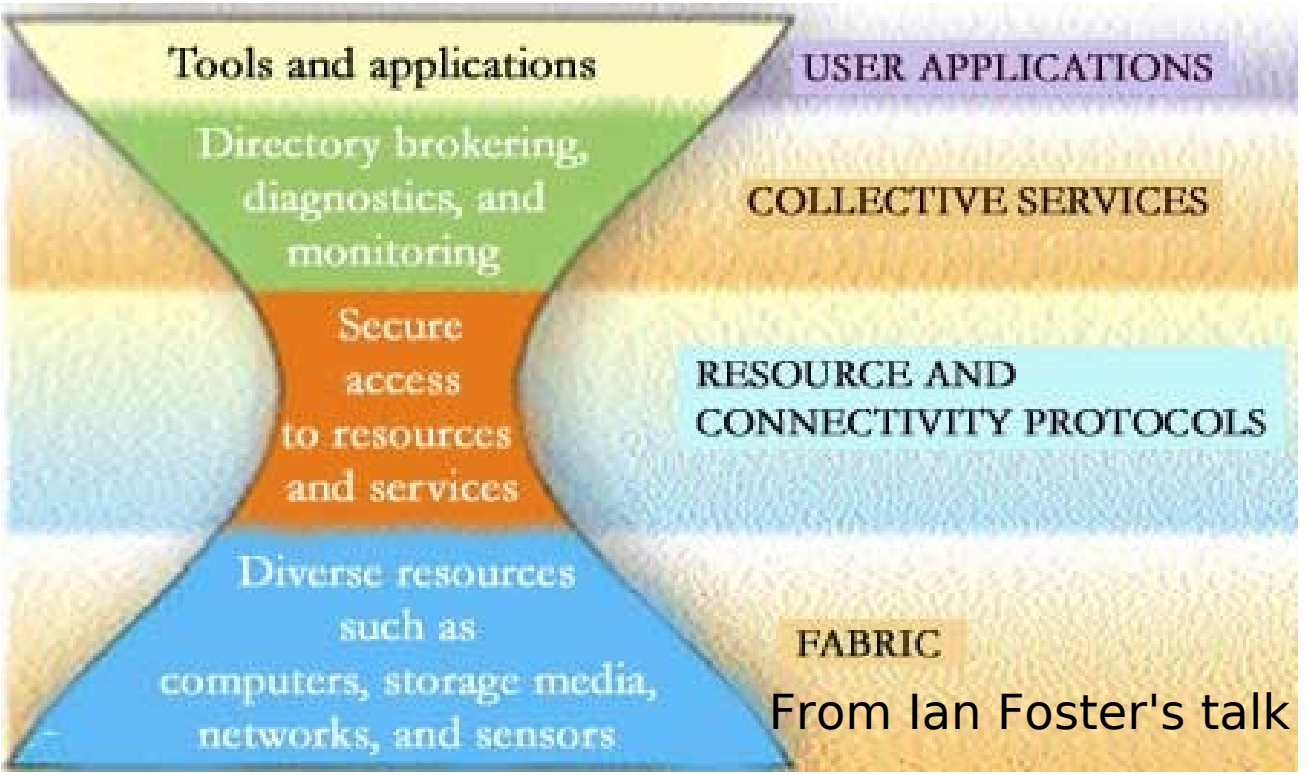
- 2001:Global Grid Forum starts
  - The first Global Grid Forum meeting was held in March 2001. Since then, GGF has produced numerous standards and specifications documents and held over a dozen additional events around the world.
- 2002-2005
  - Large European/American/International Projects on GRID computing
  - Many commercial/industrial grids..

The elements of a GRID infrastructure

- Hardware/Resources
  - Made available from different sites geographically distributed
  - CPU/Storage/Instruments/DB
- Software:
  - Something that links together all these resources: the middleware
  - Some applications to use the computational power made available
- People:
  - Who maintain the Grid
  - Who use the GRID

GRID middleware

- Middleware is "the software layer that lies between the operating system and the applications "

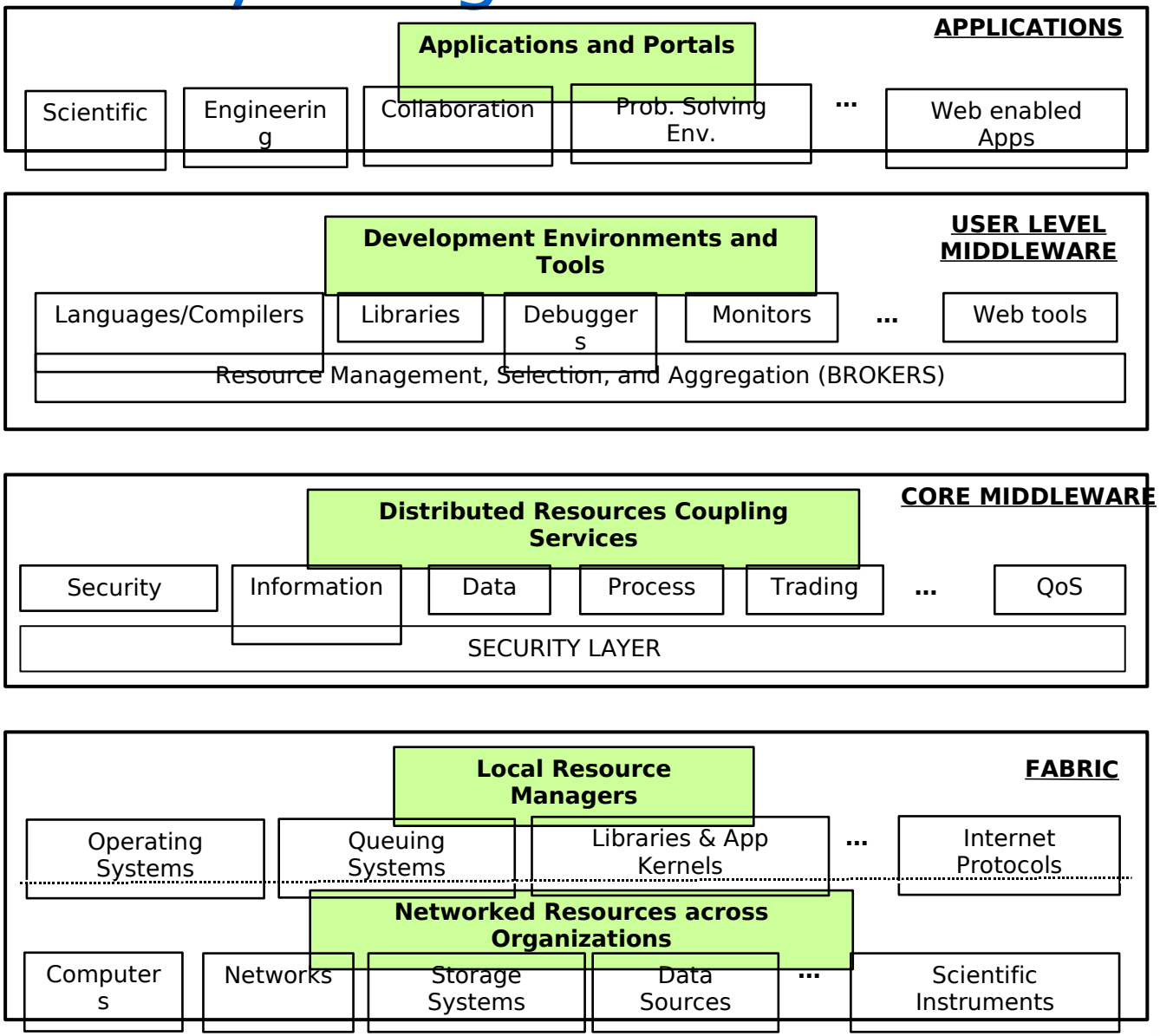


From Ian Foster's talk

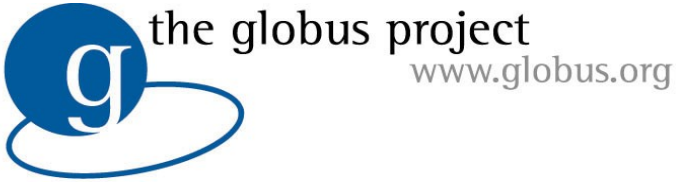
Basic elements of the grid middleware:

- Security
- Resource management
- Data management
- Information Services

Layered grid architecture..



Middleware available



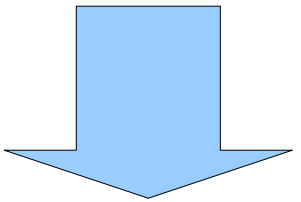
- Globus Toolkit (Argonne+ISI)
- LCG/Glite (from EU projects)
- Gridbus (Melbourne)
- Unicore... (Germany)
- And many other...

Applications for GRID computing..

- Computation intensive
  - Interactive simulation (climate modeling)
  - Large-scale simulation and analysis (galaxy formation, atomistic simulations)
  - Engineering (parameter studies, optimization model)
- Data intensive
  - Experimental data analysis (e.g., H.E.P.)
  - Image & sensor analysis (astronomy, climate)
- Distributed collaboration
  - Online instrumentation (microscopes, x-ray) Remote visualization (climate studies, biology)
  - Engineering (large-scale structural testing)

Common factor..

The size and/or complexity of the problem requires that people in several organizations collaborate and share computing resources, data, instruments



VIRTUAL ORGANIZATIONS

Virtual Organization

- Distributed resources and people
- Linked by networks, crossing admin domains
- Sharing resources, common goals
- Dynamic
- Fault Tolerant..

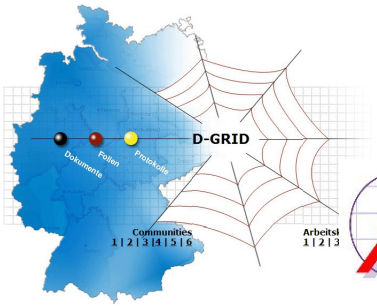
An example

- The Large Hadron Collider (LHC)
- Located at CERN, Geneva Switzerland
- Particle accelerator and collider with a circumference of 27 km
- Scheduled to go into production in july 2007

An Example (2)

- It will generate 10 Petabytes (10^7 Gigabytes) of information per year
- This information must be processed and stored somewhere
- It is beyond the scope of a single institution to manage this problem

A few international GRID project...



ACCESS GRID

Open Science Grid

grangenet

National Grid
NG
SINGAPORE

GRIP
UK
Particle
Physics
P

Grid3
2003

Ninf

iVDgL

CERN
openlab for DataGrid applications
Developing Solutions for the Data-Intensive Science of the Large Hadron Collider

LCG

WestGrid
NORDUGRID
Grid Solution for Wide Area Computing and Data Handling

Versió en Català
English Version
Versión en Castellano
tota la informació sobre el projecte
presentació
descripció tècnica
els serveis oferts
participants
documentació pública
GridCat
grid grid grid grid
com puc participar?
accés al Portal
dintre del grid
Intranet per a investigadors
estat dels serveis actius

INFN
GRID

eGee
Enabling Grids for
E-science in Europe

grid.it
project

Grid Consortium Japan

GridLab

NAREGI
超高速コンピュータ網形成プロジェクト
National Research Grid Initiative
Grid Applications
Grid Middleware
Networking
NAREGI
国立情報学研究所グリッド研究開発推進拠点 NII -The National Institute of Informatics

Tne grid project related to this workshop

International

And eventually: why the name GRID ?

- metaphor for making computer power as easy to access as an electric power Grid.

Computational grid as stable as power grid.. in Italy 23/09/2003