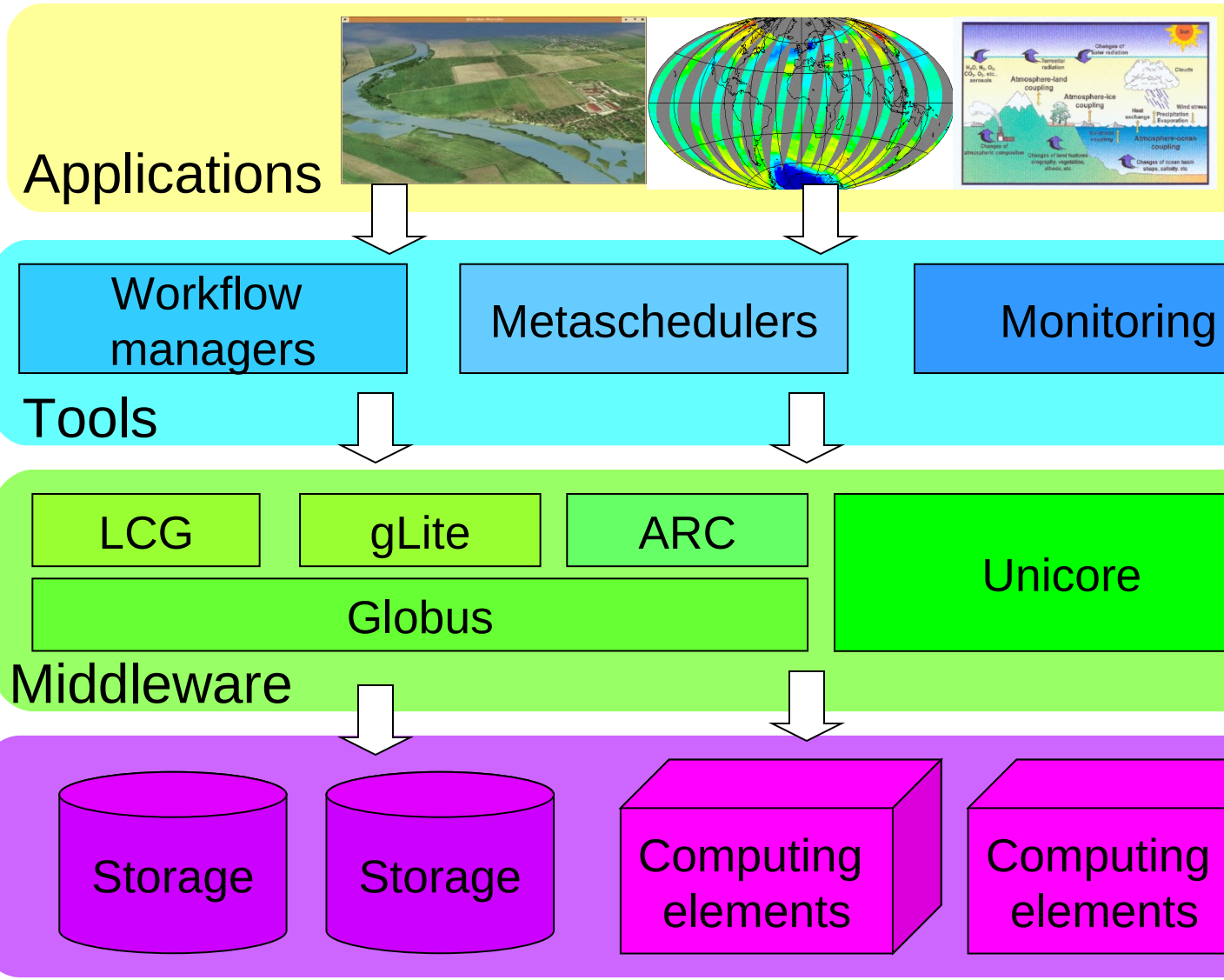
Overview of Grid middlewares and high-level Grid tools

Viet Tran
Institute of Informatics
Slovak Academy of Sciences

Overview of middleware and tools



Middleware versus tools

- **Middleware:**
  - Same middleware over whole infrastructure
  - Installed and managed by admins of sites
  - Optimized for system-wide objectives (e.g. site utilization, …)

- **Tools**
  - Individual for every applications, rich choice
  - Usually installed on single machine
  - Operate with resources in infrastructure via middleware
  - Usually can be installed by users (without root privilege)
  - Optimized for application objectives (e.g. response time, …)

Grid middleware

- **Globus Toolkit 2.x**
  - Provide basic Grid functionalities
    - Security infrastructure: authentication, authorization, delegation (GSI)
    - Data transfer: GridFTP
    - Resource management: GRAM
    - Monitoring
  - Usually used as underlayer of other Grid middleware (LCG, gLite, ARC)

- **Globus Toolkit 4.x**
  - New implementation of Globus
  - OGSA architecture, WSRF standard
  - Many functions from GT2 are re-implemented in WS scheme (e.g. GRAM -> WS-GRAM)
  - Adding new components for new functionalities (meta scheduler GridWay, data access OGSA-DAI, …)

Grid middleware (2)

- **LCG**
  - Support for resource brokering
  - Support for simple workflow (based on DAGMan)
  - Support for virtual organization (VOMS)
  - Used in old EGEE infrastructure, int.eu.grid
  - Slowly replaced by gLite

- **gLite**
  - Web service approach
  - Performance improvement
  - Job collection
  - Job monitoring
  - Current middleware in EGEE

Grid middleware (3)

- **ARC**
  - Using modified GridFTP for job submission
  - Distributed architecture
  - Simple resource brokering on user interface
  - Used in NorduGrid infrastructure

- **Unicore**
  - Provide functionalities for creating Grid from supercomputing centers (security, resource management, monitoring)
  - Also follow OGSA and WSRF
  - Used in DEISA

Workflow managers

- **Manage jobs with data dependence**
  - Usual sequence: data retrieval, pre-processing (data conversion, mining), main computation, post-processing (result analysis, visualization)
  - Combination of several simulation: e.g. meteorology-hydrology-hydraulics
- **A lot of implementations exist: DAGMan, Pegasus, Taverna, Karajan, …**
- **Can be stand-alone or integrated to middleware/portal**
  - Stand-alone: Pegasus, Karajan, GridAnt, …
  - Integrated with middleware/scheduler: DAGMan in LCG/gLite, GridWay
  - Integrated with portals: P-GRADE, K-Wf Grid
- **Most of workflow managers use DAG for representation of workflow**
  - DAGMAn, Karajan, …
  - Other representation: Petri net (K-Wf Grid), BPEL
- **Some have also graphical interface: portal or stand-alone GUI**

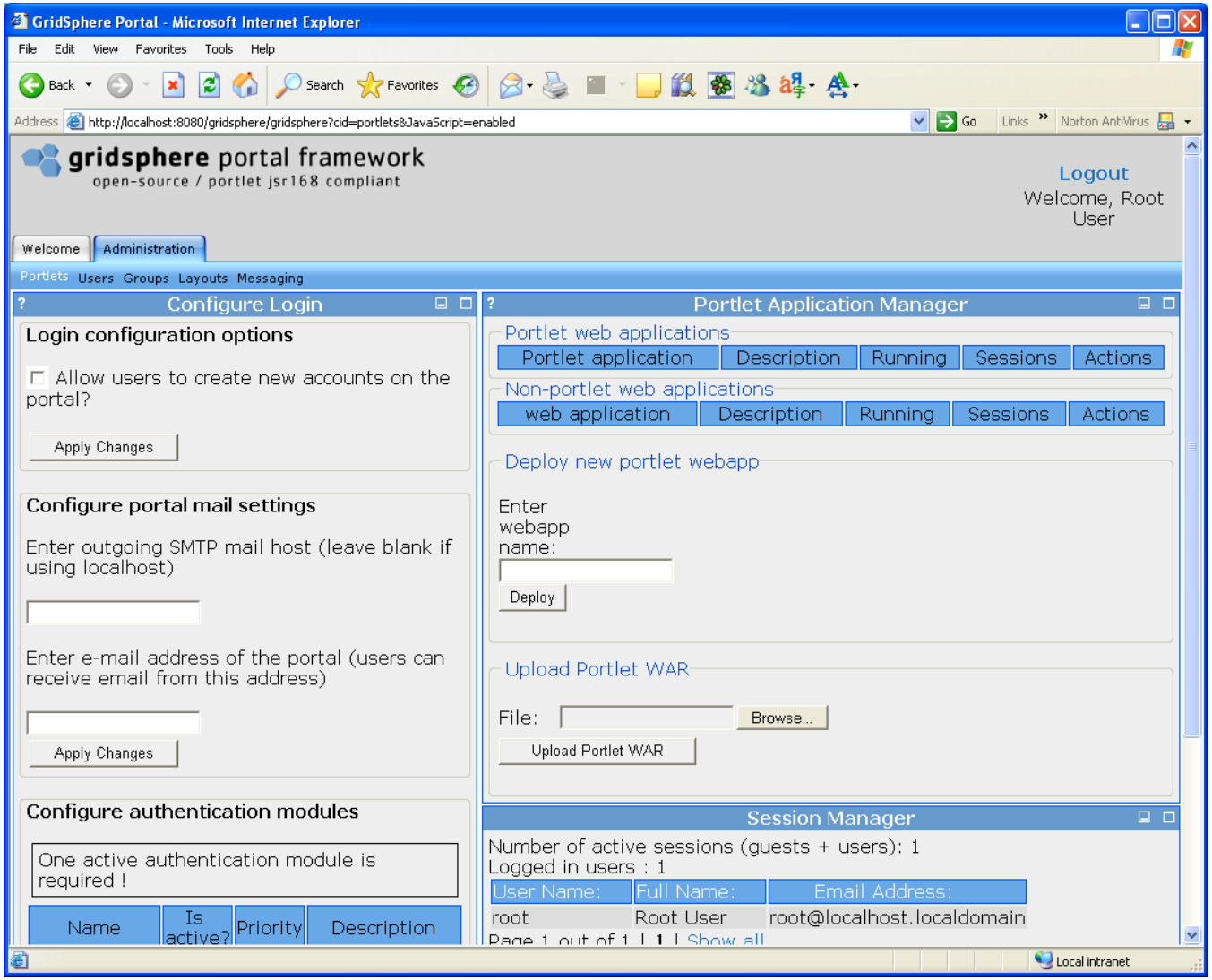
Distributed job management

- **Aim: Manage large number of small jobs**
  - parametric study,
  - data parallel: image processing
- **Working on master-worker scheme:**
  - Worker jobs are submitted to different CEs via middleware
  - Master runs on UI or as stand-alone service and sends tasks to workers
- **Implementations: DIANE, DIET**
- **Advantages:**
  - Can improve response time:
    - Task execution can start when the first worker starts
    - Worker jobs on faster machines will execute larger number of tasks
  - Give partial fault-tolerance
    - If some worker jobs fail, the rest can continue execute tasks

Portals

- **Usually used as User Interface for applications**
  - Accessible from anywhere
  - Hide implementation details
  - Additional security
  - Useful for demonstration, can be combined with visualization
- **Several frameworks and ready-to-use portlets available**
  - Gridsphere/GridPortlets
  - P-GRADE
  - GENIUS
- **Application developers usually customize portals for end-users**

Gridsphere


GridSphere Portal - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://localhost:8080/gridsphere/gridsphere?cid=portlets&JavaScript=enabled
Go Links Norton AntiVirus
gridsphere portal framework
open-source / portlet jsr168 compliant
Logout
Welcome, Root User
Welcome Administration
Portlets Users Groups Layouts Messaging
Configure Login
Login configuration options
Allow users to create new accounts on the portal?
Apply Changes
Configure portal mail settings
Enter outgoing SMTP mail host (leave blank if using localhost)
Enter e-mail address of the portal (users can receive email from this address)
Apply Changes
Configure authentication modules
One active authentication module is required !
Name Is active? Priority Description
Portlet Application Manager
Portlet web applications
Portlet application Description Running Sessions Actions
Non-portlet web applications
web application Description Running Sessions Actions
Deploy new portlet webapp
Enter webapp name:
Deploy
Upload Portlet WAR
File: Browse...
Upload Portlet WAR
Session Manager
Number of active sessions (guests + users): 1
Logged in users : 1
User Name: Full Name: Email Address:
root Root User root@localhost.localdomain
Page 1 out of 1 | 1 | Show all
Local intranet

Overview of Grid tools in EGEE

Miroslav Dobrucký, Viera Šipková
Institute of Informatics
Slovak Academy of Sciences

Information Society and Media

More tools…

GASuC = Grid Application Support Centre (SZTAKI)

- http://www.lpds.sztaki.hu/gasuc

- **P-GRADE portal**
- **GEMLCA – deploy legacy code applications**
  - binary code not modified
- **GridWay – metascheduler**
- **Ganga – job definition & mgmt**
- **Mercury – monitoring**

RESPECT

- http://technical.eu-egee.org

- **i2glogin**
- **GReLC – manage grid databases**
- **IE – instrument element (sensors)**
- **Virtual Control Room – based on GridSphere and Web 2.0**
- **Migrating Desktop**