Rapporteur talk- "Grid Access"

February 2008

H. Kornmayer

* NEC Laboratories Europe, IT Research Division, NEC Europe Ltd.

Grid Access Session

- Thanks to all speaker:

Session 1:

- g-Eclipse – Easy access to Grid infrastructures
  Dr. Harald KORNMAYER (NEC Laboratories Europe)
- Grid website vulnerabilities and the GridSite Security Framework
  Dr. Andrew MCNAB (UNIVERSITY OF MANCHESTER)
- Portals for Earth Science
  Mr. Julian LINFORD (ESA)
- CRAB, the CMS tool to allow data analysis in a distributed environment
  Federica FANZAGO (CERN-CNAF)

Session 2

- Exploitation path of Interactive European Grid on User Communities
  Dr. Jesus MARCO DE LUCAS (Instituto de Fisica de Cantabria CSIC)
- RINGrid: conceptual design of the remote instrumentation systems
  Dr. Marcin LAWENDA (Poznan Supercomputing and Networking Center)
- XtreemOS: A Grid Operating System Providing Native Virtual Organization Support
  Dr. Yvon JéGOU (INRIA)
- Deploying e-Infrastructures for new User Communities in DORII project
  Dr. Meyer NORBERT (PSNC)

And to my co-convener E. Laure

What is the dream of Grid Access?

Easy and scalable VO management

Efficient, secure, reliable application execution & ease of use and programming

access Grid infrastructures independent of the middleware

✓ Data integration from different sources

❑ User friendly Grid environment with integrated visualization

MPI application support

❑ Interactive job submissions

Integration of instrumentation

All done with a singlelogin click for a user with X.509.

Simple to install

user-friendly tool client – server architecture

hiding Grid and CMS infrastructure details

without knowing all details!!

• Grid user plays different roles

– Portals are in the critical path between User and Middleware

no computer skills needed

browse the distributed laboratory space

remote instrumentation services

What is Grid Access?

- Grid access Tools
  - Command Line tools
    CRAB
  - Web Portals
    Portals for Earth Science   GridSite (Security)
  - Rich Clients
    g-Eclipse
- Operating System
  XtreemOS
- Grid infrastructure
  int.eu.grid
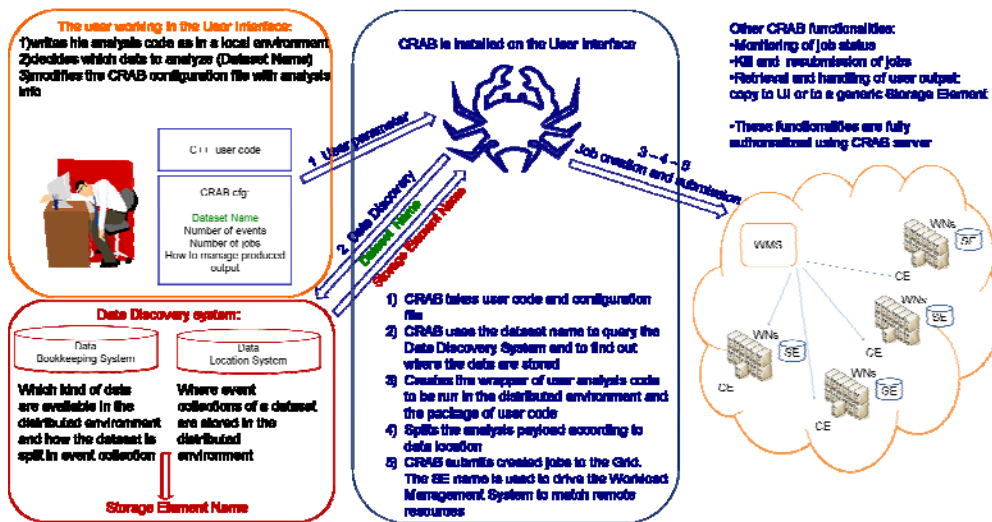- Remote Instrumentations
  RinGrid   DORII

Command like tools - CRAB

CRAB – a tool for CMS (HEP)

- store and manage this huge quantity of data
- assure data access to physicists of CMS collaboration
- assure enough computing power for data analysis and simulation
- guarantee resources and data availability to allowed users

Python

Issues:
- scalability
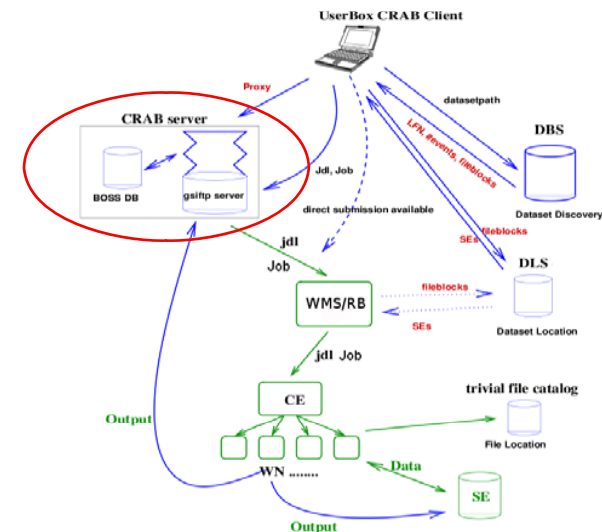- reliability
- automating common CLI actions
  - submission, error handling, …

From "standalone" (client side) tool

to a client server architecture

Empowered by Innovation **NEC**

Portals – Earth Science

Earth Science

Data integration from different sources

- **-- Satellite & in-situ instruments**
- **-- Global, multi-dimensional coverage**
- **-- Long time series and historic data - increasing**
- **-- Explosions of data - the "Data Deluge"**

Scientists and instruments widely scattered across geographical and organizational boundaries

- **-- Requirement for largescale computing networks**
- **-- Large infrastructure building is ongoing**

GRID model for largescale "loosely distributed" computing

→ DEGREE study

Approach

- Focus on four specific areas:
  1. Grid Application Families, Requirements and Test Suites
  2. Grid Job Control and Workflow Management
  3. Grid Data Management
  4. Portals for Grid, SOA and eCollaborations
- For each area
  - --Establish state-of-art in ES and other science communities
  - --ES Requirements gathering & analysis
  - --Gap analysis
  - --Provide Key requirements & Recommendations for input to ES Grid Roadmap
- Generate ES Grid Roadmap
- Disseminate results to wide ES and Grid Communities

Portals – Earth Science (II)

Portal Requirements

- Generic requirements
  - Interoperability between different Grid MW & infrastructure
  - Reliability & QoS
  - Guaranteed fast turnaround
  - Standard "off-the-shelf" tools for integrated Grid Security and User management
  - Dynamic content authoring, addition of customized services, registration of available resources
  - User support, tutorials
- ES specific requirements
  - Strong emphasis on Metadata and Data, its Discovery and Access
  - Integration of heterogeneous distributed services (Grid & Geo-services, OGC)
  - Tools & interfaces readily useable by ES Scientist
  - Facilitate integration with ES web services
  - Interoperability with ES data catalogues

Gap analysis for portals:

- Focus must be on ES functionality, grid working as back-end does not have to be visible.

- Higher level components targeted to ES

- Big emphasis on Metadata and Data, its Discovery and Access

- Graphical interfaces for different kinds of ES data, activated by data type

- Standard "off-the-shelf" tools for integrated Grid Security and User Management

- Light-weight grid service interfaces to grid functionality for easy integration and/or mash-ups creation

Empowered by Innovation **NEC**

Portal Security - GridSite

- Portals are websites, security is an issue
  - Sessions are manage by cookies
    - Stealing cookies enables front door attacks
- CSRF – Cross Site Request Fogery
  - From the user's side, there isn't much you can do about CSRF: you click the fake button on the attacker website, whoosh! you're on the victim website, the damage is done
- CSRF and the Grid
  - Websites use X.509 == cookies that last a year

- GridSite
  - Available in release 1.6
  - Offers a concept and software solution to make you Website CSRF secure

Empowered by Innovation NEC

Portals – Summary

- Domain specific CLI tools exist
  - Not so easy to port
    - → New domains build them again and again
- Many portals exist using different technology
  - "Too many portals exist!"
  - Application specific
  - Middleware specific
    - → there is no "quasi standard" solution/framework

Hard for new users to choose the right one?
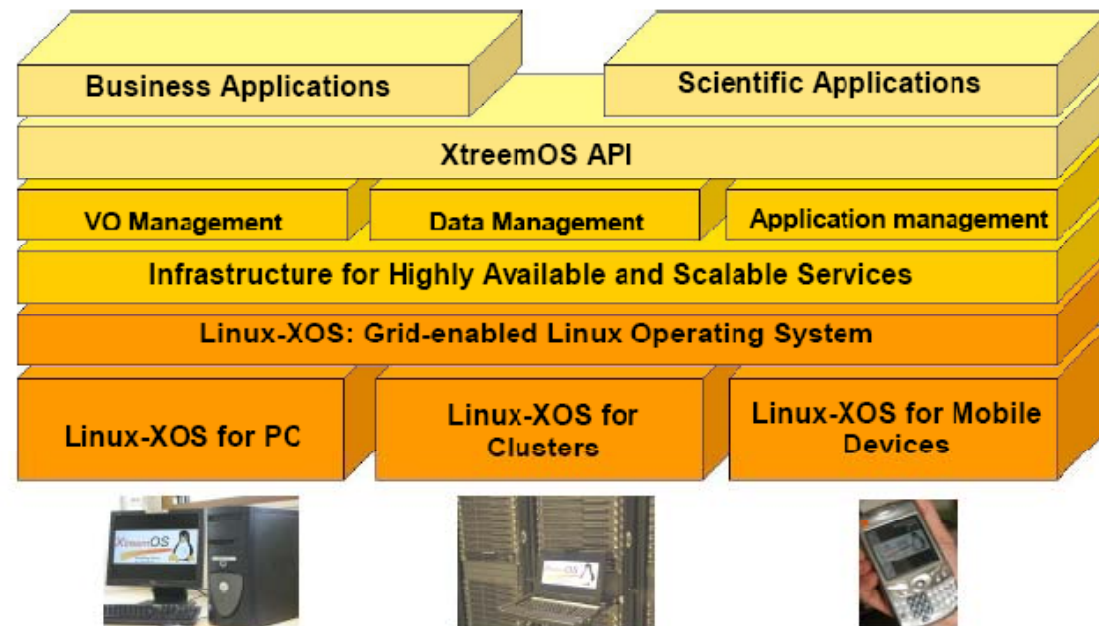
ROI, TCO, ...

Rich Clients – g-Eclipse

- g-Eclipse has the potential for the client side "browser" of Grid infrastructures
  - Middleware independent
  - OS independent
  - reliable
  - extensible

- Built on top of the extensible Eclipse eco system
  - Open Source changes the process of software development
  - Eclipse has 70% market share in tooling market

  - Important: Eclipse offers a Open Source development process for industry to deliver in time and quality
    - Planned, Transparency, Quality, Agile methods

- G-Eclipse build not just a single access tool,

  But a generic framework/eco system to build you own tool

- Grid user plays different roles
  - Grid applications user
  - Grid resource provider and operator
  - Grid application developer
  - ….

- Grid user acts in different contexts
  - Virtual Organizations
  - Projects
  - …
-

Grid operating system - XtreemOS

- XtreemOS
  A Grid Operating System Providing Native Virtual Organization
  Support
- → easy and scalable VO management
- → efficient, secure, reliable application execution
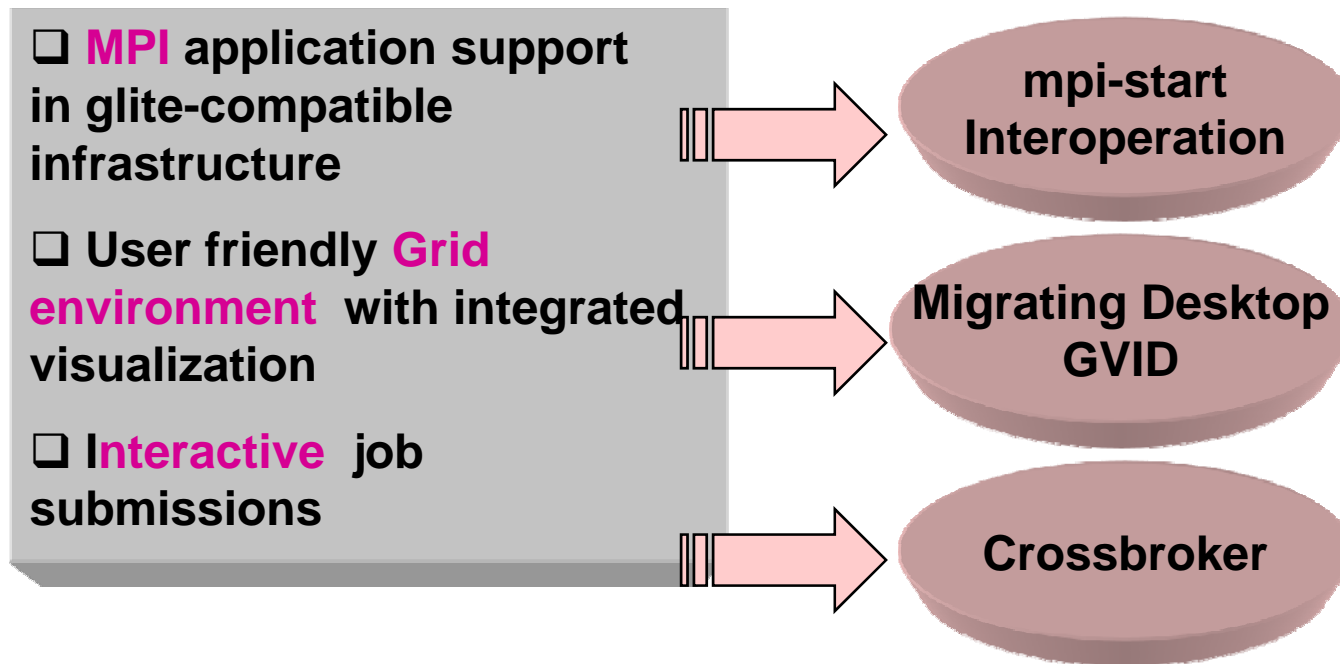- → ease of use and programmi

Challenges:
- Aim at adding fundamental functionalities to be embedd in Linux for secure applicat execution in Grids
- Secure resource management in highly dynamic Grids
- Provide a simple API with POSIX

Implementation ongoing → 1st release this summer

Grid Infrastructure – int.eu.grid

□ **MPI** application support in glite-compatible infrastructure

□ User friendly **Grid environment** with integrated visualization

□ **Interactive** job submissions

mpi-start Interoperation

Migrating Desktop GVID

Crossbroker

· Support for different types of parallel jobs
  – MPI implementations: Open MPI, PACX-MPI, MPICH and MPICH-G2
· Support for interactive jobs

• The int.eu.grid infrastructure runs in parallel with the EGEE one with additional features (compatible/interoperable)
• component integration with RESPECT

Service offer to communities/users based on a written SLA
Price estimations are available

i2g SERVICES ON THE GRID – Beta

January, 2008

i2g Grid Services Offer includes:

- i2g Try & See Service (i2g TRY)
  - for new users, or users with limited requests
- i2g Storage Service (i2g STORAGE)
  - for handling medium to large storage requests
- i2g CONSOLIDATION Service (i2g CONSOLIDATION)
  - long term storage or archiving
- i2g Batch Service (i2g BATCH)
  - to support execution of high throughput jobs
- i2g Interactive Service (i2g INTERACTIVE)
  - immediate execution, steering and visualization
- i2g MPI Service (i2g MPI)
  - supporting parallel programs
- i2g Interactive MPI Service (i2g INTERACTIVE MPI)
  - interactive execution of parallel programs
- i2g PACX-MPI Service (i2g PACX-MPI)
  - distributed parallel programs to use all available resources

The i2g Consortium is launching a Beta Offer for providing a suite of services oriented to support e-Science work for research communities in academy and industry.
The offer is supported by a pan-European distributed e-Infrastructure providing a significant computing and storage capacity able to satisfy medium-size demands (typically 100 processors and 10 Terabytes).
It is specially designed to make demanding and interactive computing in the Grid easier for developers and friendly for final users.

Service offer

Price

SLA

EXAMPLE OF i2g SERVICE AGREEMENT

Updated January 28, 2008

PLEASE READ CAREFULLY - THIS IS A BINDING CONTRACT

THIS **i2g** SERVICES CUSTOMER AGREEMENT ("AGREEMENT" OR "**i2g** SERVICES CUSTOMER AGREEMENT") IS A BINDING AGREEMENT BETWEEN **i2g** CONSORTIUM ("i2g") AND YOU AND, IF APPLICABLE, THE COMPANY, COLLABORATION, OR OTHER LEGAL ENTITY YOU REPRESENT (COLLECTIVELY, "YOU"). THIS AGREEMENT INCORPORATES BY REFERENCE (1) THE PRIVACY NOTICE ("PRIVACY NOTICE") AND (2) THE CONDITIONS OF USE ("CONDITIONS OF USE") POSTED ON WWW.i2g.EU, AS THEY MAY BE MODIFIED BY US FROM TIME TO TIME.

BY SENDING THE "ACCEPT" FOR THIS AGREEMENT OR ACCEPTING ANY MODIFICATION TO THIS AGREEMENT IN ACCORDANCE WITH SECTION 2 BELOW, YOU AGREE TO BE BOUND BY THE TERMS AND CONDITIONS OF THIS AGREEMENT. IF YOU ARE ENTERING INTO THIS AGREEMENT ON BEHALF OF A COMPANY OR OTHER LEGAL ENTITY, YOU REPRESENT THAT YOU HAVE THE LEGAL AUTHORITY TO BIND THE LEGAL ENTITY TO THIS AGREEMENT, IN WHICH CASE "YOU" SHALL MEAN SUCH ENTITY.

ACCOUNTING RECORD FOR VIRTUAL ORGANIZATION *RESERVOIR*
Period: 1-Feb-2008 to 29-Feb-2008
Company/Collaboration: *ECOHYDROS SL*
Contact person: *FCO (fco [at] ecohydros.com)*
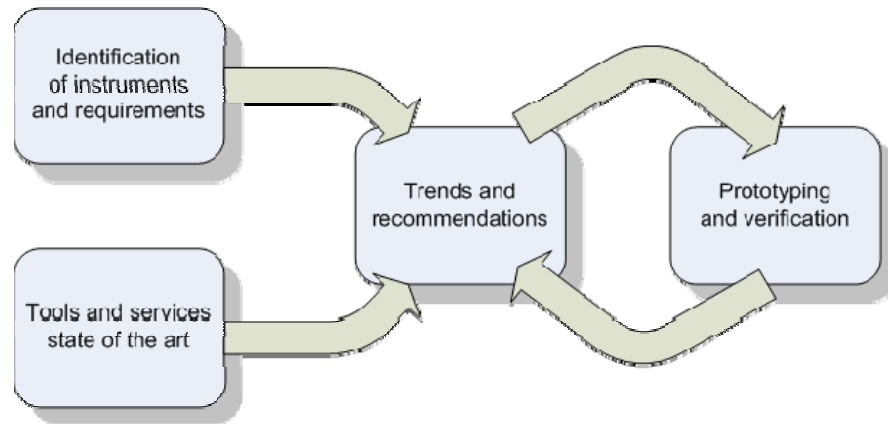Users: *FCO,ALEX,AGUS,CARLOS*

Service	Initial Estimation	Actual usage	Fee	Total Cost	Description of Use
TRY	100HR 10GB	87HR 2GB	0€	0€	Test first executions of SMDR
STORAGE	10000GB	780GB	1€	780€	Store maps for Cuerda del Pozo
CONSOLIDATION	5000GB	300GB	1€	300€	Archive maps + output
BATCH	10000HR	7532HR	0,1€	753€	Compute 10m grid for Itoiz
INTERACTIVE	1000HR	466HR	0,5€	233€	Display maps
MPI	1000HR	371HR	0,2€	74€	Compute 5m grid for Cuerda del Pozo
INTERACTIVE MPI	5000HR	298HR	1€	298€	6 pretests for 5m grid for Cuerda del Pozo
PACX-MPI	5000HR	728HR	0,2€	155€	Compute at 2m grid in 2 days
TOTAL				2593€	

Remote Instrumentation – RINGrid & DORII

Need/Goal:
Build an integrated scientific
infrastructure including
• remote sensors
• remote instruments
• computing (Grid, HPC)

Identification of instruments and requirements

Tools and services state of the art

Trends and recommendations

Prototyping and verification

• Demand (network, storage, …) is lower than expected, but often varying!
• IPv6, SIP, interactivity, instrumented workflows, …

Implementation of the RINGrid results with existing tools/frameworks:
(CrossBroker, g-Eclipse, GVid, gLite (EGEE), …)

Three exemplary domains:
Earthquake,
Environment,
Synchrotron and free electron lasers

Empowered by Innovation **NEC**

Summary Grid Access

- For communities, tools and infrastructure exists!
- Communities have often specific requirements
-

  - Generic CLI and Portals are missing!
    - A portal is too cheap, so wheels are reinvented often.
  - The wilder approach of g-Eclipse by using an existing reliable platform can be a blueprint for further developments
    - It is worth to look what happens beyond your community!
    - Collaboration in tools instead of competition!
  - Access includes SLA documents for users and cost estimates!
  - For scientists the Grid is just a part of a bigger integrated scientific infrastructure!!

Empowered by Innovation

NEC