

¿Cuál es la situación de las grids en el mundo?

Raúl Hazas Izquierdo
Dirección de Telemática
CICESE

Coordinador Grids/Supercómputo CUDI



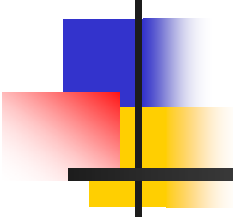


Prólogo

Hablaremos de algunos de los problemas que enfrentan los desarrolladores de grids y cómo han tratado de resolverlos. Mostraremos casos de estudio para tres áreas importantes:

- El manejo de líneas de espera y calendarización (Globus, condor-g, etc.)
- La manera de enviar tareas para su ejecución (Ninf-g, Nimrod, etc.)
- Las precauciones que se toman en cuanto a seguridad (IGTF, pkIris, etc.)





Looking Beyond Globus Grids

Arun Krishnan





Globus

- Positives
 - Seems to have finally decided on a set of standards
 - Has a nice name.
 - Hmmm...
 - Lots of success stories eg. NIMROD/ ROCKS
 - Security Aspects / Certificates etc..
- Drawbacks
 - Very flaky code
 - Implementation is a pain
 - Maintenance is even more painful





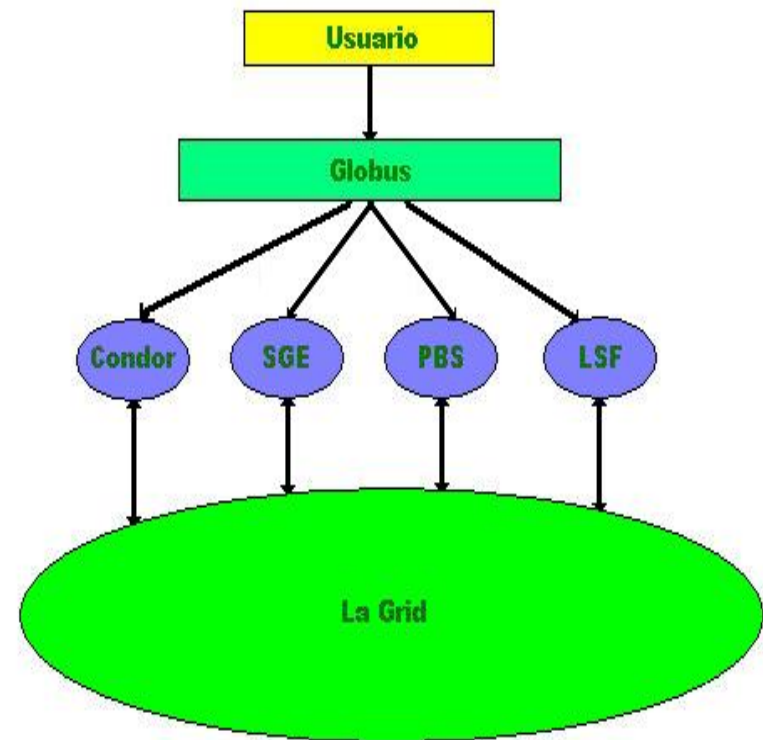
PRAGMA Resources

- Currently, Globus seems to be a de-facto middleware
- Can we extend this?
- Or should we stick to one standard?
- Let the discussions begin.
 - Don't throw the baby out with the bath water ?
 - Let applications team decide on which middleware to use..
 - What don't we like about Globus? Design? Implementation?
 - Stability for today or tolerance for tomorrow?
 - Webservices for tomorrow?
 - Can we feedback to GGF from the PRAGMA community?
 - Challenge to the resource group: How can we add new capabilities to the PRAGMA testbed?



Manejo de líneas de espera y calendarización

- Las tareas pueden enviarse a través de Globus a los sistemas manejados por Condor, SGE, PBS y LSF



	Condor 6.6.3	SGE 5.3	PBS Pro 5.4	LSF 6.0
Disponibilidad	Descargable en la red	Descargable en la red	Producto comercial	Producto comercial
Soporte para Microsoft Windows	Parcialmente	No	Si	Si
Soporte para GUI	No	Si	Si	Si
Tareas en tandas	Si	Si	Si	Si
Tareas interactivas	Si	Si	Si	Si
Tareas en paralelo	Si	Si	Si	Si
Capacidad de reservación de recursos	No	Si	Si	Si
Checkpointing	Si	Si	Si	Si
Migración de tareas	Si	No	Si	Si
Calendarización por prioridad	Si	Si	Si	Si
Calendarización previamente delimitada	No	Si	Si	Si
Acoplamiento entre recursos y demanda	Si	Si	Si	Si
Manejo del flujo de tareas	Si	Si	Si	Si



Layered Programming Model/Method

Portal / PSE

GridPort, HotPage,
GSDK, Grid PSE Builder,
etc...



Easy but
inflexible



High-level Grid Middleware

MPI (MPICH-G2, PACX-MPI, ...)
GridRPC (Ninf-G, NetSolve, ...)



MPI

Low-level Grid Middleware

Globus Toolkit



Primitives

Socket, system calls, ...

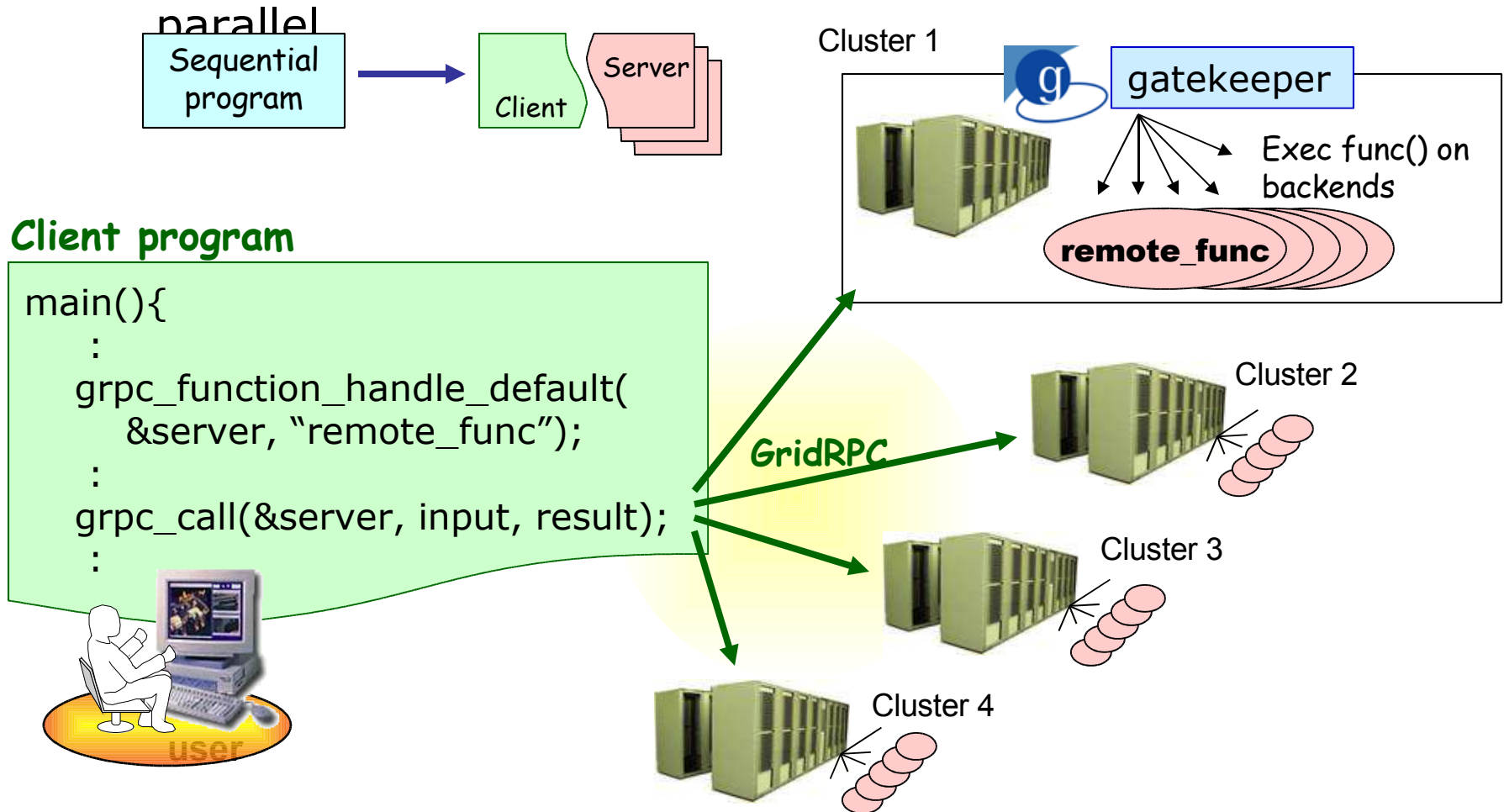
Difficult
but flexible



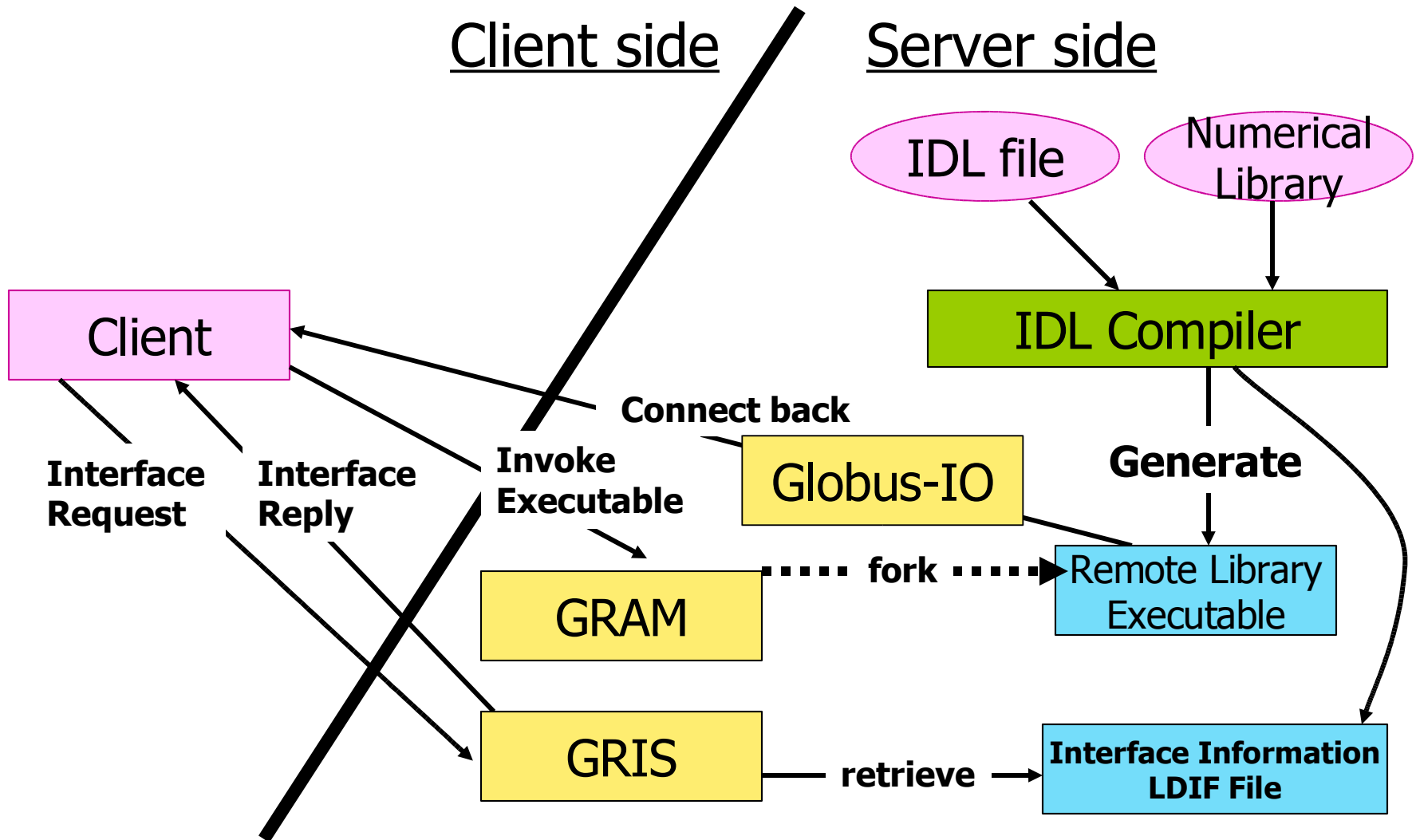
Overview of GridRPC

■ One of the programming models

- ▶ Execute partial calculations on multiple servers in parallel



Architecture of Ninf-G



Direction of Implementing FT

■ Execute our application as long as possible

- ▶ Along with the routine-basis experiments (Daily use of the Grid) of PRAGMA project
- ▶ On the Asia Pacific Grid testbed operated by PRAGMA / ApGrid project
 - ⊗ Unstable network in the Asia, Less practical experiments
 - ⊗ What kinds of faults happens? How often?
- ▶ Repeat the execution while improving the program

■ Development issues

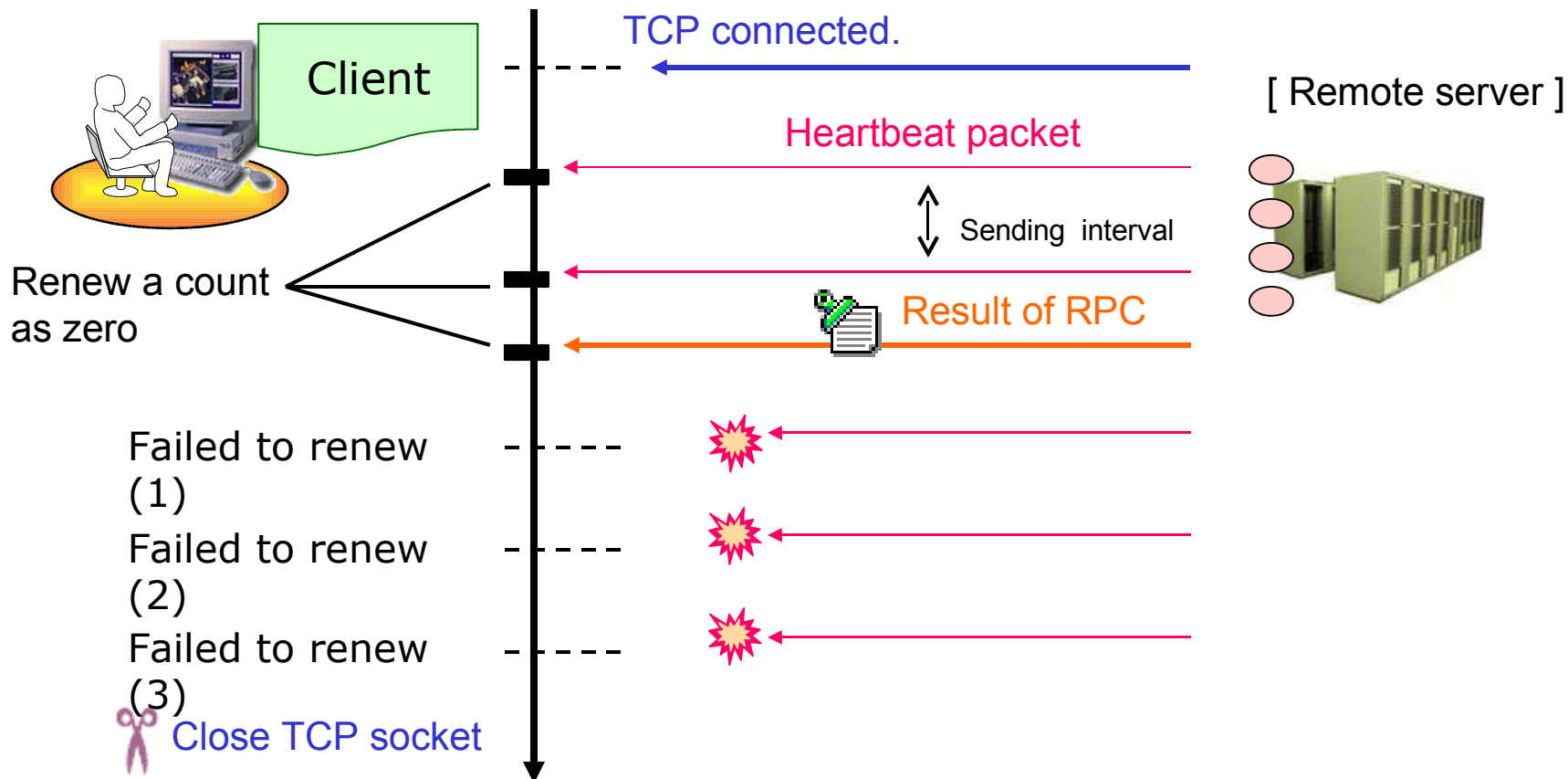
- ▶ Application should continue calculation without down servers. A failed RPC should be performed on another server on another lived node.
- ▶ Down servers should be restarted after a fault is resolved.

Timeout mechanism

■ Example of heartbeat

► Users' configuration: Sending interval = 60 sec, Max count = 3

@ Timeout seconds will be $60 \times 3 (= 180)$.



PRAGMA/ApGrid Testbed

■ **Total 8 countries / 10 sites / 104 nodes / 210 CPUs**



SDSC, USA



64 CPUs



AIST, Japan



28 CPUs



UNAM, Mexico



6 CPUs



TITECH, Japan



8 CPUs



NCHC, Taiwan



16 CPUs



KISTI, Korea



16 CPUs



KU, Thailand



16 CPUs



Bioinformatics
Institute

BII, Singapore



16 CPUs



UNIVERSITI SAINS MALAYSIA

USM, Malaysia



32 CPUs



NCSA, USA



8 CPUs

Statistical Results for 3 months

Cumulative results

- ▶ # of executions by 2 users: 43
- ▶ Execution time (Total) : 1210 hours (50.4 days)
 - (Longest) : 164 hours (6.8 days)
 - (Average) : 28.14 hours (1.2 days)
- ▶ Total # of RPCs : 2,500,000
- ▶ Total # of RPC failures : 1,600
 - @ Error ratio : 0.064 %

Major faults

- ▶ Unstable networks between client and server
 - @ Packet drop, Fluctuating throughput, TCP disconnection
- ▶ Server node down
 - @ Due to heat, electricity, HDD and NFS problem, and moving

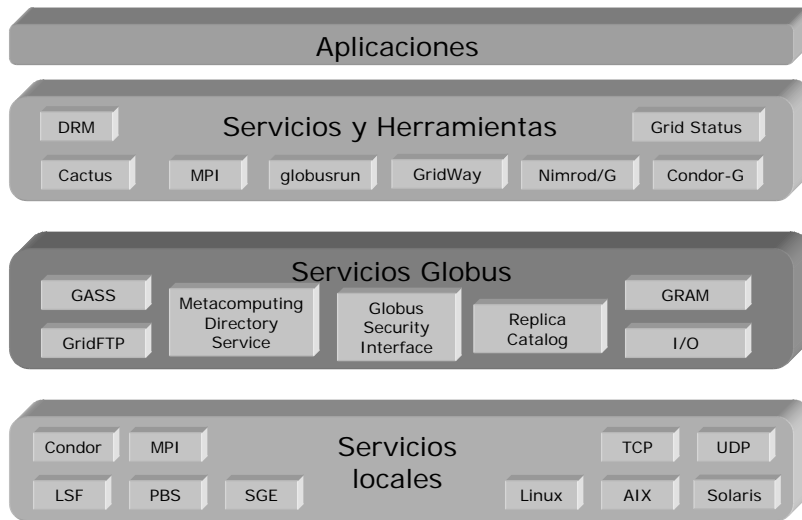


Necesidades

- Algunos componentes aún no cuentan con la madurez necesaria para correr aplicaciones de manera
 - sostenida
 - rutinaria
- Se requiere contar con
 - un supercalendarizador
 - una mayor capacidad de tolerancia a fallas



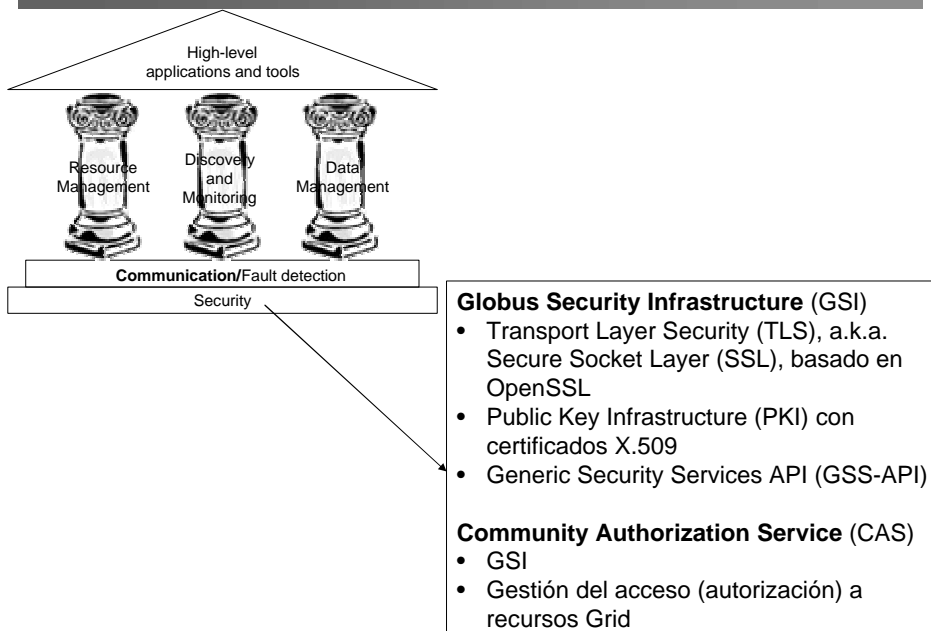
Modelo de Capas



Visión Global de las Componentes de Globus

5/108

Seguridad

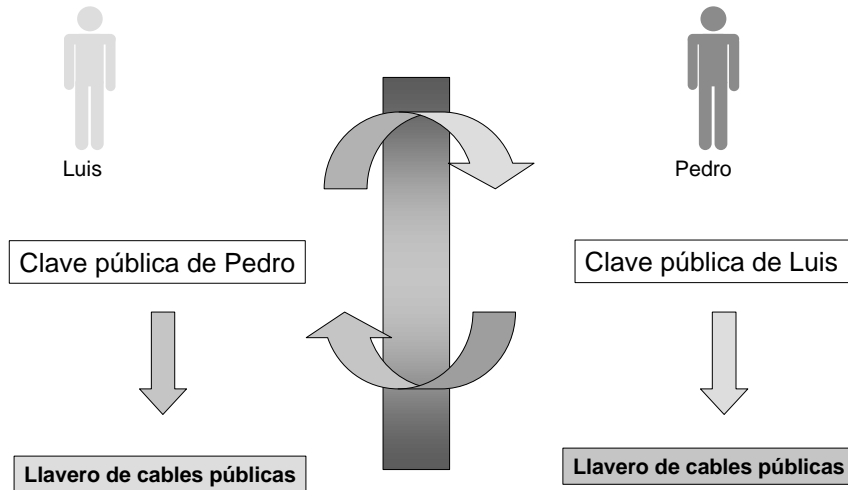


Visión Global de las Componentes de Globus

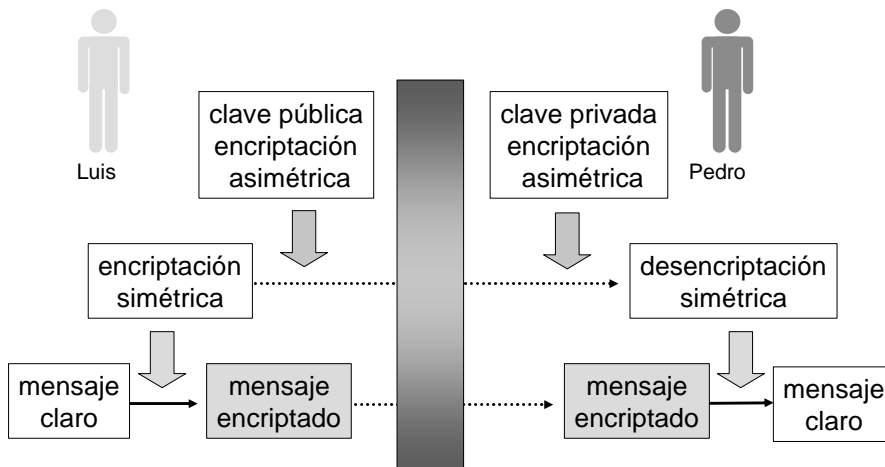
6/108

Encriptación Asimétrica

La primera vez que se comunican deben intercambiarse sus claves públicas



Encriptación por Clave de Sesión



Entidades de Certificación

Proceso

- Un usuario debe crear una **clave privada** y un **certificado** que incluye su **clave pública** además de otra información de identidad (nº de serie del certificado, nombre del usuario, ...)
- El usuario envía el certificado a una entidad de certificación para su firma. Al firmarlo se incluyen nuevos campos como en nombre de la entidad, y además se **firma el certificado con la clave privada de la entidad**.
- La entidad devuelve al usuario el **certificado firmado**
- Cuando el usuario quiera establecer una comunicación segura envía su certificado y de este modo el receptor confirma que el usuario es quién dice ser gracias a que la clave pública (certificado) de la entidad de certificación es conocida
- Cada usuario guarda su clave privada generalmente de forma encriptada con una nueva clave o un PIN.

Certificados Digitales X509

Características de los certificados X509

- Documentos (ficheros) que asocian:
 - Un recurso del Grid o usuario, expresados mediante detalles específicos sobre el propietario
 - Su clave pública.
- *Entidad Certificadora*: Certifica que la clave pública del certificado pertenece a su propietario.
- *Firma Digital*: Garantiza la integridad del certificado así como la relación entre la clave pública y la información que contiene. *La relación entre ambos se debe establecer por medios no criptográficos.*

DN Field:	Abbrev.	Description:
Common Name	CN	Name being certified
Organization or Company	O	Name is associated with this organization
Organizational Unit	OU	Name is associated with this organization unit, such as a department
City/Locality	L	Name is located in this City
State/Province	ST	Name is located in this State/Province
Country	C	Name is located in this Country (ISO code)

Componentes del certificado X509

Versión del Certificado
Número de serie
Algoritmo de Firma del Emisor
Nombre del Emisor
Periodo de Validez
Subject DN (<i>distinguished name</i>)
Clave Pública
ID del emisor
Extensiones
Firma Digital de la CA

environment. Here you, the owner of DN, are authorized by host B to act as a local user on the host B.

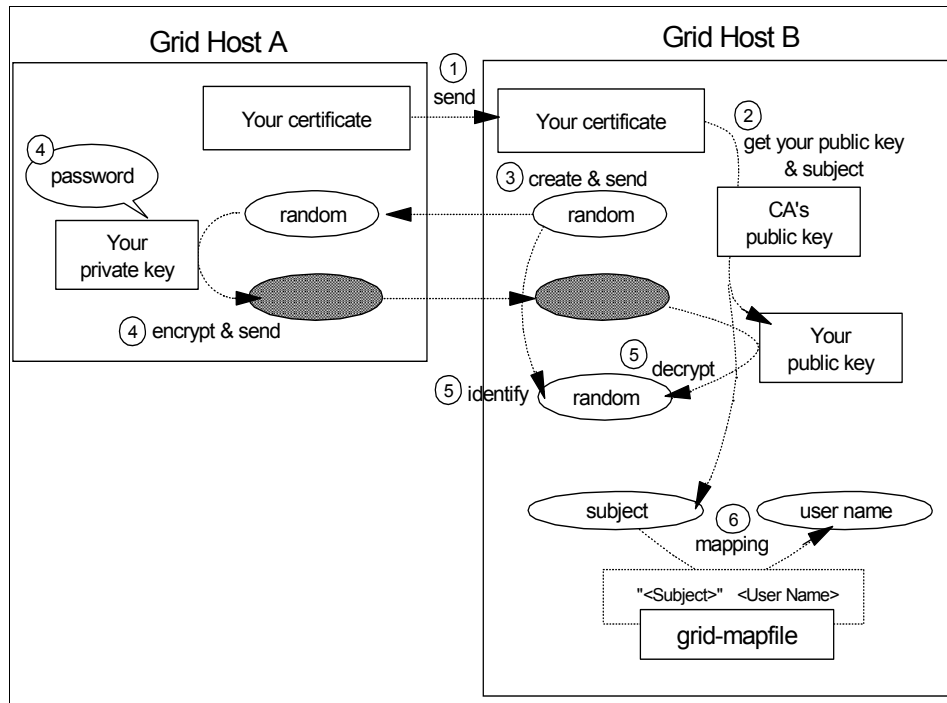


Figure 3-4 Authentication procedure

In grid environments, your host will become a client in some cases, and in other cases, a server. Therefore, your host might be required to authenticate another host and be authenticated by the host at the same time. In this case, you can use the mutual authentication function of GSI. This function is almost the same as explained above, and it proceeds with the authentication steps, and changes the direction of hosts and redoes the procedure.

Briefly speaking, authentication is the process of sharing public keys securely with each other, and authorization is the process that maps your DN to a local user/group of a remote host.

Delegation

Imagine a situation where you distribute jobs to remote grid machines and let them distribute their child jobs to other machines under your security policy. In this situation, you can use the delegation function of GSI, as shown in Figure 3-5 on page 66.

3.5 Potential security risks

Building a PKI environment will provide the necessary services along with the GSI to design a secure grid solution. This, however, does not guarantee that there are not any security risks. Within this section, we will examine some possible vulnerabilities to watch out for during your security design. This is by no means a laundry list for all security vulnerabilities or a cookbook for building a secure infrastructure. The importance of this section is to highlight some vulnerabilities that you may not have been aware of and allow you the option of taking the proper steps to improve their security.

Ultimately, it will be up to you to design, build, and test out your security infrastructure within your grid network. All of the security tools, processes, and policies in the world will not completely secure a networked environment. There is still some risk involved, but hopefully with the right people and tools at your disposal, you can reduce that risk to a negligible level.

3.5.1 PKI vulnerabilities

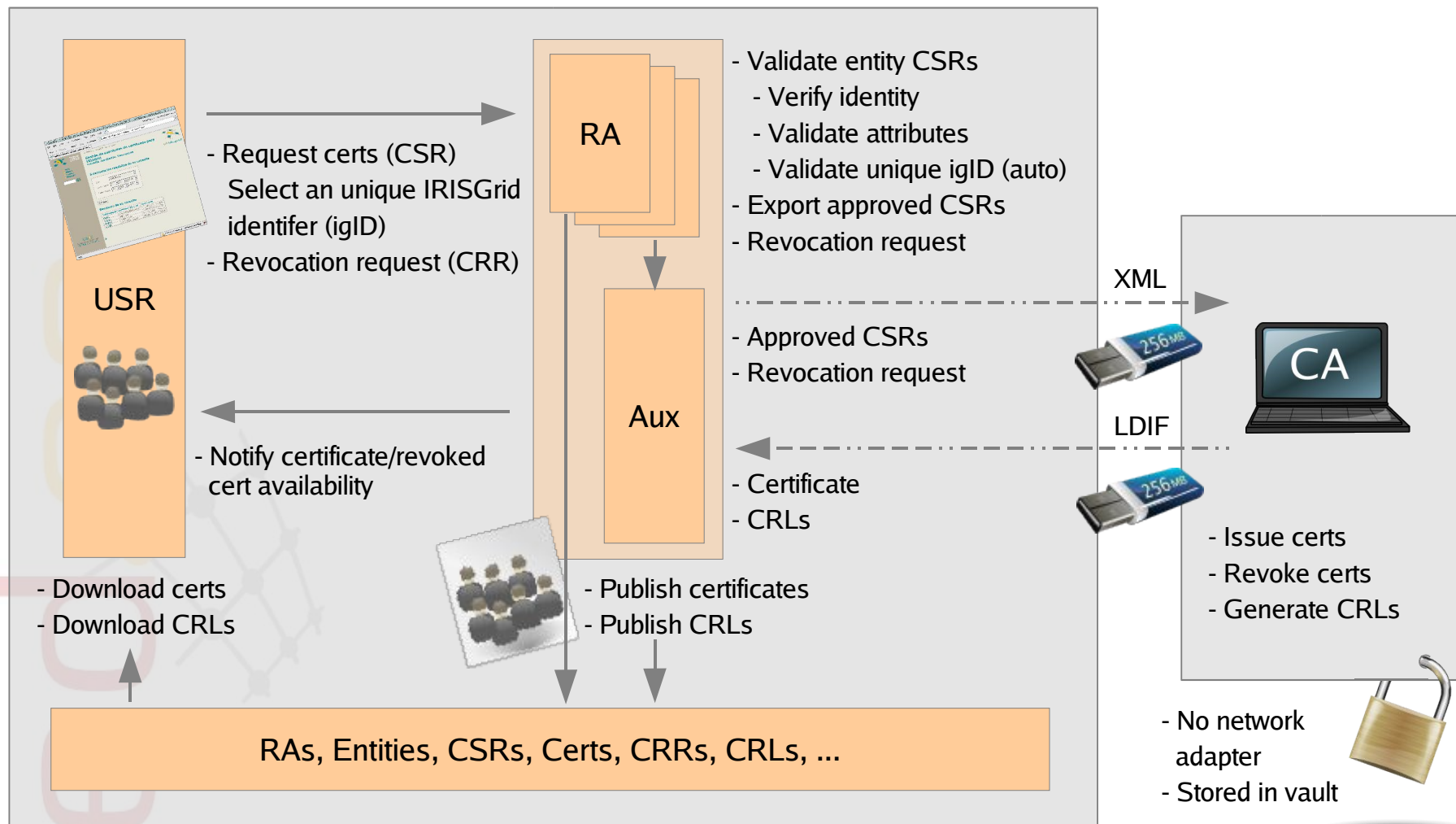
Just because you have built a PKI environment does not mean that your network is completely secure. There are still many vulnerabilities to be aware of. It is necessary to always keep an open mind and understand that with any networked environment there is going to be some risk involved.

Within a PKI environment, you constantly have to worry about the locations of your private keys and thefts of digital certificates. The following areas should be considered when dealing with a PKI environment:

- ▶ Impersonation: Obtaining a certificate through fraudulent means (either user or organization).
- ▶ Theft of private key: Unauthorized use of a private key associated with a valid certificate.
- ▶ Compromise of root CA private key: Using a CA key to sign fraudulent certificates or destroying a private key.
- ▶ Automatic Trust Decisions: Automated trust decisions can also automate fraud.

3.5.2 Grid server vulnerabilities

Any server or workstation that participates in the grid is a potential vulnerability to an external or internal hacker. Knowing this, it is very important to protect and isolate any grid computer from any network or resources that do not need explicit access to the grid. The most common way to isolate or protect your grid computers from unauthorized network access can be done through the use of





**5th EUGRidPMA meeting in Poznan,
September 28-30 2005**



pkIRISGrid

**a PKI for e-science activities
provided by the Spanish NREN RedIRIS**

Javi Masa, RedIRIS



RedIRIS



Trust on the Grid Goes Global

Boston, Mass., USA –

Today (5 October, 2005), users of Grid computing worldwide are a step closer to accessing computers and information in 50 countries and regions¹, from Canada to China, Portugal to Pakistan. The International Grid Trust Federation (IGTF), established this week during the 15th Global Grid Forum (GGF) in Boston, brings together Grid organizations representing Asia, the Americas and Europe that are working towards allowing scientific researchers to identify themselves to any Grid resource in the world with just a single online identity.

IGTF's members issue electronic certificates that allow scientists to use the Grid. The Grids protected by IGTF certificates include over 40,000 computer processors and petabytes of storage - equivalent to over a million DVDs. Making sure the owners of Grids trust each other's security procedures is key to letting researchers access all these resources.

"Living in the information age, access to electronic resources has never been so vital before.", says Christos Kanellopoulos of Aristotle University of Thessaloniki, Greece, and co-chair of the GGF Certification Authorities Operations working group (CAOPS): *"Already today e-scientists can use their certificates to access and use grid enabled resources in any part of the world, making the World Wide Grid a reality. IGTF is a big step towards the dream of bridging the digital divide".*

The IGTF brings grid-oriented organizations around the globe much closer to realizing the promise of grids. Grids aim to harness the power of geographically dispersed computing resources, experimental facilities and research centres. Grid developers' goal is to provide seamless access to all the resources available. However, at present there are many independently operated grids, spread throughout the world, and users able to work on one can't necessarily gain access to the others.

Fundamental to user access is user authentication – making sure that only those users who have the proper credentials are granted access to the resources. While this can be a significant challenge within a grid, achieving agreement on how to provide this level of authentication between grids has been an even bigger challenge. That's where the IGTF takes centre stage. With the establishing of the IGTF, the foundation is laid for building a trusted basis for identity management, and a further step taken towards global interoperability for scientific grids.

The IGTF is a federation of certification authorities or grid policy management authorities (grid PMAs), and the major grid infrastructure projects that together define the policies and standards for grid identity management. Comprising the three regional grid policy management bodies, the Asia Pacific Grid PMA (APGridPMA), the European Policy Management Authority for Grid Authentication in e-Science (EUGridPMA) and the Americas GridPMA (TAGPMA), the federation today has 61 members and covers 50 countries and regions.

The new federation builds on the strong foundations laid in Europe by the EUGridPMA, which established the common baseline for identity providers that is considered trustworthy by an increasing number of resource centres and service providers. These same guidelines were also adopted by the APGridPMA and the TAGPMA, who at the same time enriched the federation with innovative services for quickly bootstrapping new centres in the Grid, and integration of the Grid with the scientists' home organisations.

The three EU e-Infrastructure projects Enabling Grids for e-Science (EGEE), the Distributed European Infrastructure for Supercomputing Applications (DEISA), and the South-Eastern European Grid-enabled e-Infrastructure Development (SEE-GRID), as well as the Nordic European Grid (Nordugrid), were the first to join the common trust domain, and the strong support from the e-Infrastructure Reflection Group at the European policy level further accelerated the building of the federation.

Also the US-based Open Science Grid (OSG) and TERAGrid projects, the ApGrid and PRAGMA projects in the Asia Pacific, and the world-wide LHC Computing Grid (LCG), base their authentication on the certificates issued by the IGTF affiliated certification authorities. Regional and national programmes that collaborate on a global scale also leverage the IGTF foundations today.

"Recently, the number of organizations involved in large scale regional and international Grid projects in the Asia Pacific region has been dramatically increasing," said Yoshio Tanaka (AIST, Tokyo, Japan), chair of the Asia Pacific Grid PMA. *"There is a strong demand for establishing trust federation with production Grid projects in Europe and Americas. The IGTF accelerates the emergence of a globe-wide Grid infrastructure".*

Leveraging both national and international support from a variety of sources, the members of the federation are able to provide high-quality credentials – called certificates – at no cost to the scientists. Key members of federation, like The DoEGrids Certificate Authority operated by the U.S. Department of Energy's ESnet, and the Grid-FR CA operated by the French national research council CNRS, ensure that no scientists are "left out in the rain", and act as a catch-all for communities like the LCG and EGEE projects with a global network of agents.

Tony Genovese from DOE's Lawrence Berkeley National Laboratory, which manages ESnet for DOE, says: *"By establishing IGTF, we are seeing the fruition of a the first steps ESnet and the European Grid took back in February 2002 when a researcher at Fermilab used an authenticating certificate created by ESnet to successfully transfer files to Imperial College and Lancaster University in the U.K. We did this as part of the pilot for the Particle Physics Data Grid. Once the British sites and Fermilab recognized and accepted each other's certificates, the data transfer went smoothly".*

The IGTF is closely linked to the efforts of the CA Operations Working Group in the Global Grid Forum, whose co-chair Darcy Quesnel of CANARIE is also the chair of the Americas Grid PMA. The working group provides the standard federation documents and the broad consensus between providers and relying parties.

The other important element for enabling a wide trust base is the use of the TACAR repository run by TERENA, the Trans-European Research and Educational Networking Association: a single source for all relying parties to validate their trust infrastructure both for the IGTF and for many other academic identity providers.

The future work of the IGTF will venture into better integration of Grid authentication with other mechanisms. *"The future is going to be with integrated services. Currently grid identity management is usually a separate thing the user needs to think about,"* said David Groep of NIKHEF (Amsterdam, the Netherlands) and chair of the EUGridPMA. *"In the future, single sign-on should integrate grid, network and campus resources in a seamless system. Grid computing in a university classroom is currently hard to do; new services that will emerge based on the IGTF work will alleviate this barrier".*

For more information regarding the IGTF in general, please contact

1. Tony Genovese electronically at tony@es.net or by phone at +1 510 486 4003 (available working hours US Pacific time),
2. David L. Groep electronically at: davidg@eugridpma.org (available working hours Central European Time),
3. Yoshio Tanaka electronically at: yoshio.tanaka@aist.go.jp (available working hours Japanese Standard Time)

¹ The countries and regions currently covered by the IGTF are:
Albania, Armenia, Australia, Austria, Belgium, Bosnia and Herzegovina, Brazil, Bulgaria, Canada, CERN, China, Croatia, Cyprus, Czech Republic, Denmark, Estonia, F. Y. R. O. Macedonia, Finland, France, Germany, Greece, Hong Kong, Hungary, Iceland, India, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Malaysia, Norway, Pakistan, Poland, Portugal, Romania, Russia, Serbia and Montenegro, Singapore, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, Taiwan, The Netherlands, Turkey, United Kingdom, and the United States of America.