

Clusters Linux, Grids Computacionales y el proyecto EELA

Genghis Ríos (grios@pucp.edu.pe)
Dirección de Informática Académica
PUCP



Pontificia Universidad Católica del Perú

Toda la vida adelante.

Agenda

- Clusters Computacionales
- Grids Computing
- EELA



Pontificia Universidad Católica del Perú

Toda la vida adelante.

Clusters Computacionales



Pontificia Universidad Católica del Perú

Toda la vida adelante.

Cluster Computacionales

- Se requiere gran capacidad de cómputo para aplicaciones que van desde simulaciones del medio ambiente, biomedicina, física de partículas, estadísticas, criptografía, etc.
- La potencia de las computadoras es cada vez mayor, pero al mismo tiempo siguen apareciendo retos computacionales más grandes.



Cluster Beowulf

- Nombre que reciben en general las técnicas y procedimientos para usar Linux como cluster computacional.
- Hay librerías y API's para C, C++, Fortran, Java para el desarrollo de aplicaciones paralelas:
 - **MPI** (Message Passing Interface)
 - **PVM** (Parallel Virtual Machine) para el desarrollo de aplicaciones paralelizadas.
- **MPI es la librería más usada, en particular la implementación MPICH.**
- MPI es una herramienta para HPC (High Performance Computing)



Distribuciones Linux para clusters

- OSCAR:
 - <http://oscar.openclustergroup.org/tiki-index.php>
- ROCKS:
 - <http://www.rocksclusters.org>



MOSIX

- Es un conjunto de módulos que actúan básicamente a nivel del kernel Linux, no a nivel de la aplicación.
- Las aplicaciones corren de manera distribuida sobre una gran cantidad de equipos.
- No es necesario hacer cambios en las aplicaciones.
- Herramienta HTC (High Throughput Computing)



MOSIX

- Página del proyecto MOSIX
 - <http://openmosix.sourceforge.net>



- En particular hay una distribución basada en Knoppix llamada Cluster Knoppix
 - <http://bofh.be/clusterknoppix>



MOSIX vs Beowulf

- Si se trata de resolver un problema que implique la ejecución de miles de procesos pequeños e independientes entonces tal vez sea más sencillo resolverlo con MOSIX.
- Si se trata de un solo proceso que necesita una gran cantidad de cómputo entonces sería mejor el uso del esquema Beowulf lo cual implica el desarrollo de la aplicación respectiva y adaptarla a las librerías de clustering, es decir, las MPI o PVM.

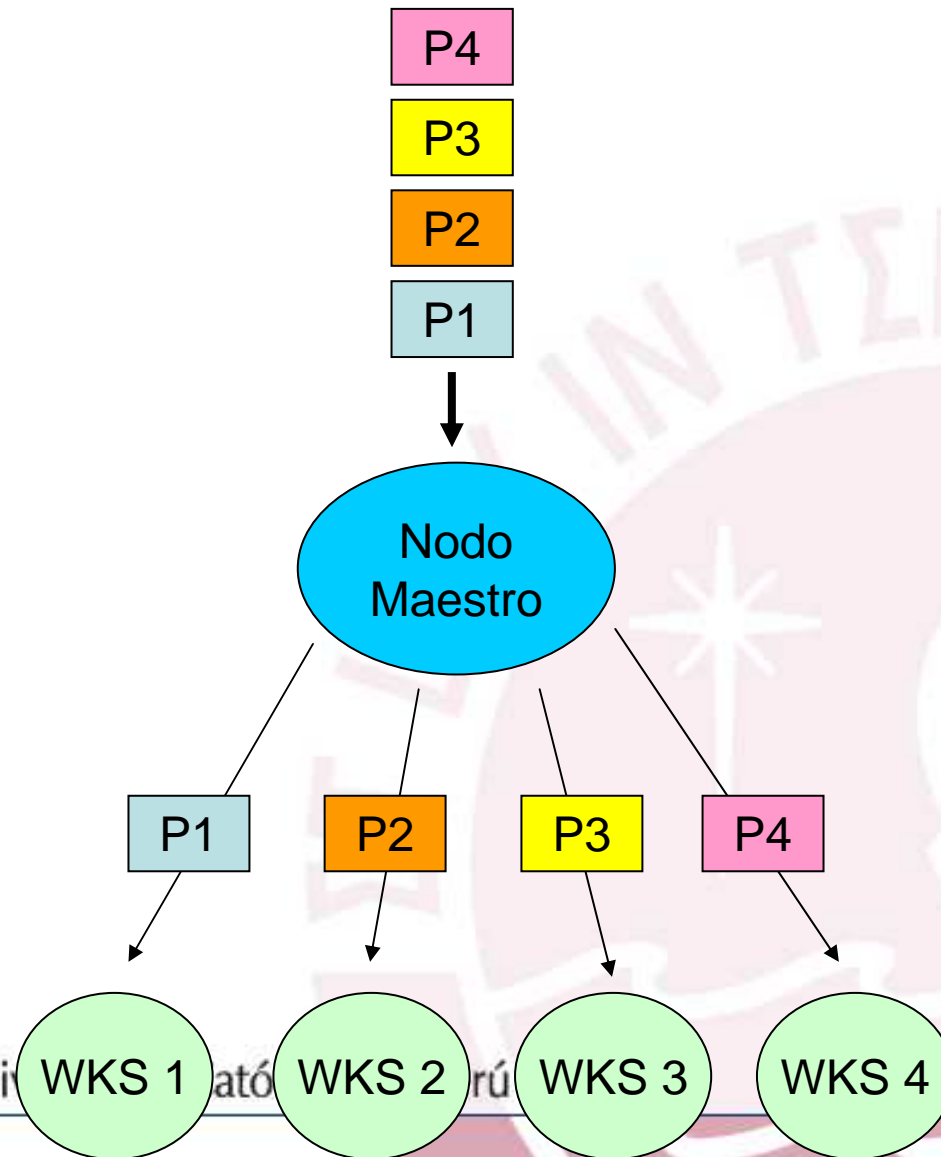


MOSIX vs Beowulf

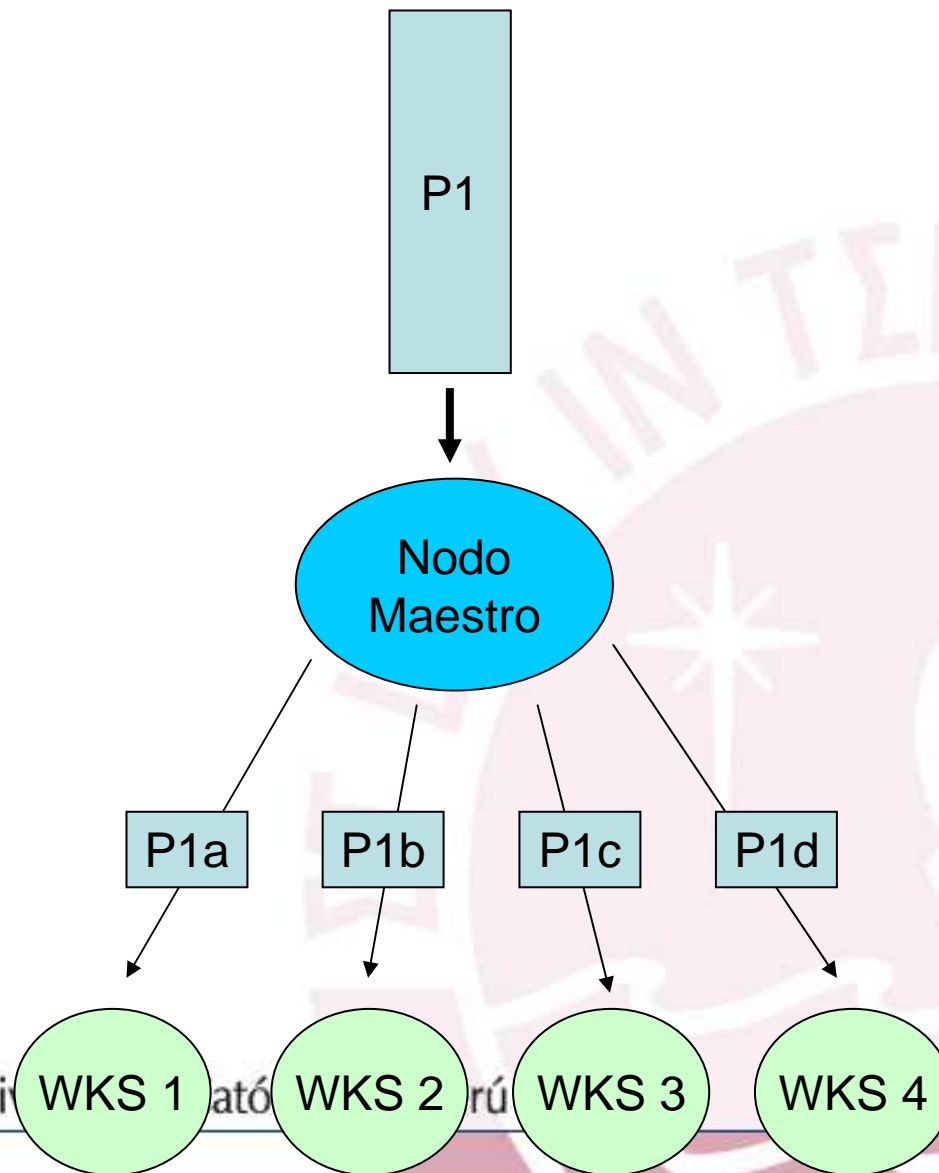
- HTC se usa para trabajos en los que el tiempo no es importante, se interesa en poder disponer de los recursos ociosos de la red.
- HPC se usa para trabajos en los que el tiempo es importantísimo, se interesa en poder optimizar el tiempo de ejecución de las aplicaciones.
- MOSIX es para HTC.
- Beowulf es para HPC.



MOSIX



Cluster Beowulf



Grid Computing



Pontificia Universidad Católica del Perú

Toda la vida adelante.

Grid Computing

- Metodología de trabajo en red con arquitecturas y protocolos bien definidos para el uso compartido de recursos informáticos (cómputo y almacenamiento).
- Permite agrupar virtualmente equipos de todo el mundo.
- Necesario controles de acceso y seguridad.



Grid Computing

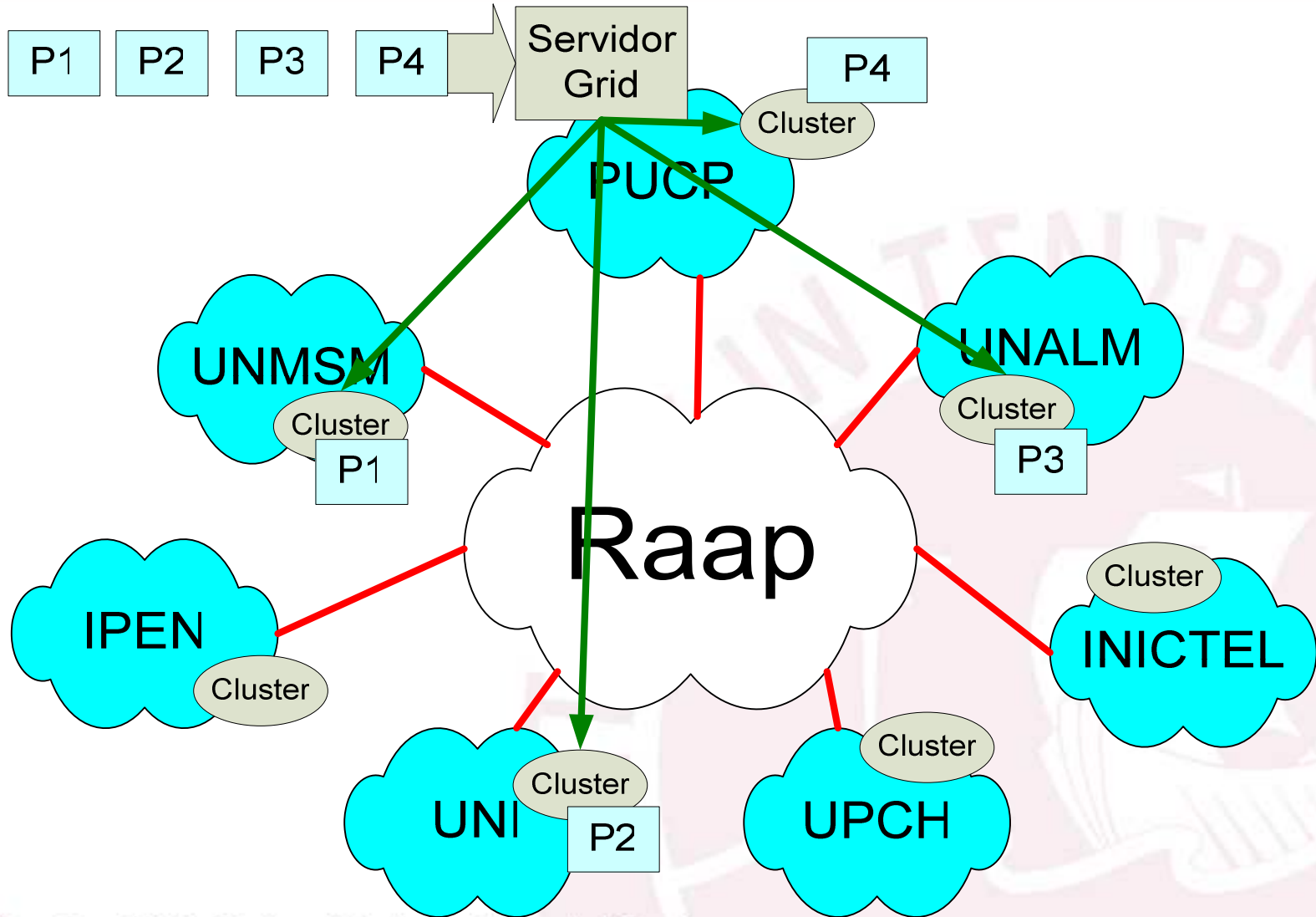
- Analogía con el servicio eléctrico:
 - *“No es necesario que cada uno tenga un generador de corriente en casa, es simplemente tener un tomacorriente y un gran proveedor de energía”.*
- La Grid permite que instituciones de menos recursos tengan acceso a poder computacional de forma remota, o que diversas instituciones puedan unir sus recursos computacionales para obtener uno más poderoso



Grid Computing

- El principal objetivo es hacer un uso más eficiente de los recursos computacionales de la red asignándoles trabajos a los equipos más ociosos.
- De las 24 horas del día sólo una fracción del tiempo una computadora. La Grid localiza los equipos menos usados y les asigna tareas.
- No todos los problemas se pueden resolver con Grid, sólo son algunos los que encajan en éste esquema.





Grid Computing

- Links

- <http://www.gridcomputing.com>
- <http://publib-b.boulder.ibm.com/cgi-bin/searchsite.cgi?query=grid+computing+>



Grid Computing

- Necesario un ***middleware***, es decir, una interface que se ubica entre los recursos finales (las computadoras) y la aplicación de usuario.
- Algunos middleware para grid son:
 - *Globus Toolkit*
 - *Sun Grid Engine*
 - *LCG*
 - *G-Lite*



Condor



Condor
High Throughput Computing

- Aplicación que establece un protocolo de comunicación entre los “nodos esclavos” y el “servidor maestro”.
- Permite el pase de comandos para la ejecución remota de aplicaciones.
- Herramienta HTC (High Throughput Computing)
- Usado por Globus, LCG y G-Lite.

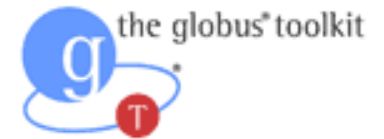


Grid Engine

- Fácil instalación y administración.
- Gratuito (pero no libre).
- Actualmente está en la versión 6.0
- Soporta diversos sistemas operativos, incluyendo Linux (recomendable RedHat).
- Los clientes pueden ser Windows.
- Puede interactuar con el Globus Toolkit (servicio de pago).
- <http://gridengine.sunsource.net/>
- <http://www.sun.com/software/gridware/>



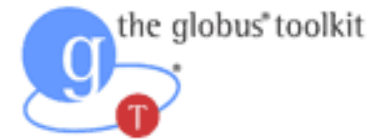
Globus Toolkit



- IBM, HP, Intel y Sun contribuyen a la estandarización de protocolos y librerías abiertas que permitan crear una Grid a escala mundial mediante el Globus Toolkit.
- Sólo trabaja sobre Unix/Linux.
- Globus es considerado estándar "de facto" para aplicaciones Grid.
 - <http://www.globus.org/>
 - <http://www-unix.globus.org/toolkit/>
 - <http://www.globus.org/research/applications/>



Globus Toolkit



- Es una mezcla de C con Java. Va por la versión 4.
- Globus sigue el estándar OGSI (Open Grid Services Infrastructure)
- Problemas que se pueden resolver con Globus están en:

<http://www.globus.org/research/applications/>

- IBM tiene su propio instalador (sólo funciona sobre Linux RHAS)

<http://www-106.ibm.com/developerworks/grid/library/gr-develop/>

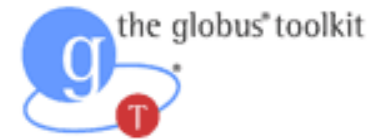
- El sistema de videoconferencia para redes avanzadas AccessGrid utiliza Globus.



Pontificia Universidad Católica del Perú

Toda la vida adelante.

Globus Toolkit



- Manual de instalación GT 4
 - <http://garl.serc.iisc.ernet.in/docs/2005/interns/pushparajan-report.pdf>
- La distribución ROCKS permite una fácil instalación del GT 4.
- Para el desarrollo de aplicaciones paralelas se hace uso del MPICH-G2



LCG



- LHC Computing Grid
- Large Hadron Collider (LHC) es un acelerador de partículas del CERN.
- Middleware desarrollado para el proyecto EGEE (Enabling Grids for E-Science in Europe)
- Última versión es LCG-2.
- Usa GT 2 y Condor.
 - <http://lcg.web.cern.ch/LCG/activities/middleware.html>



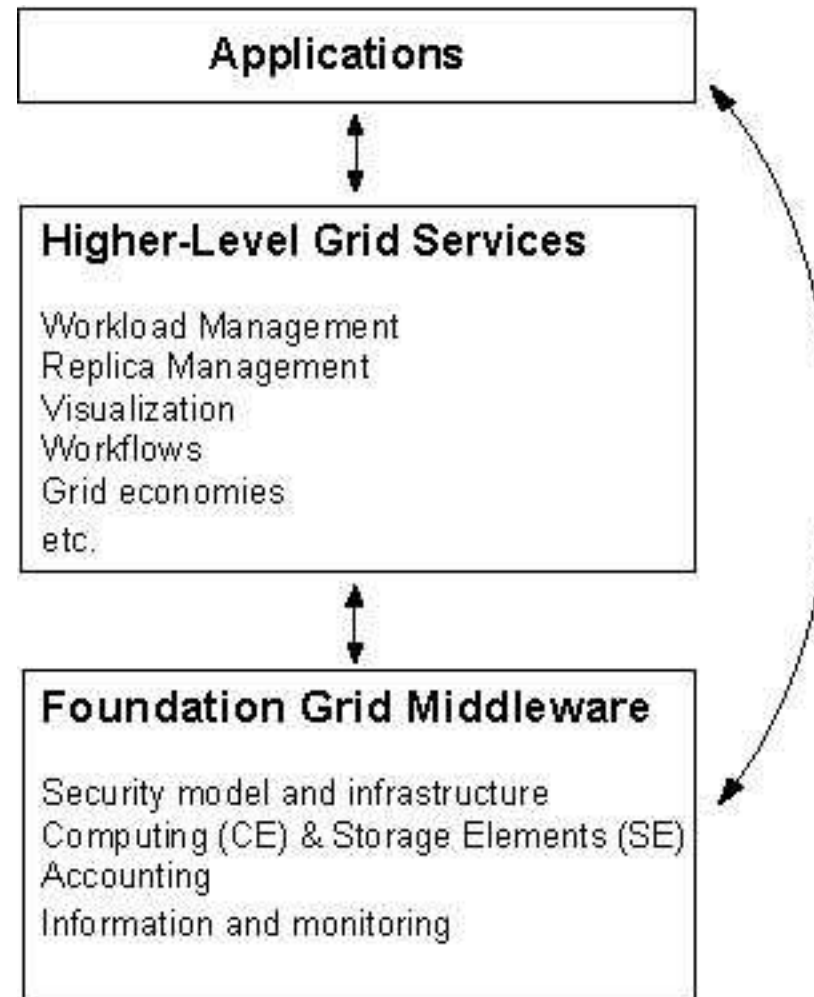
g-Lite



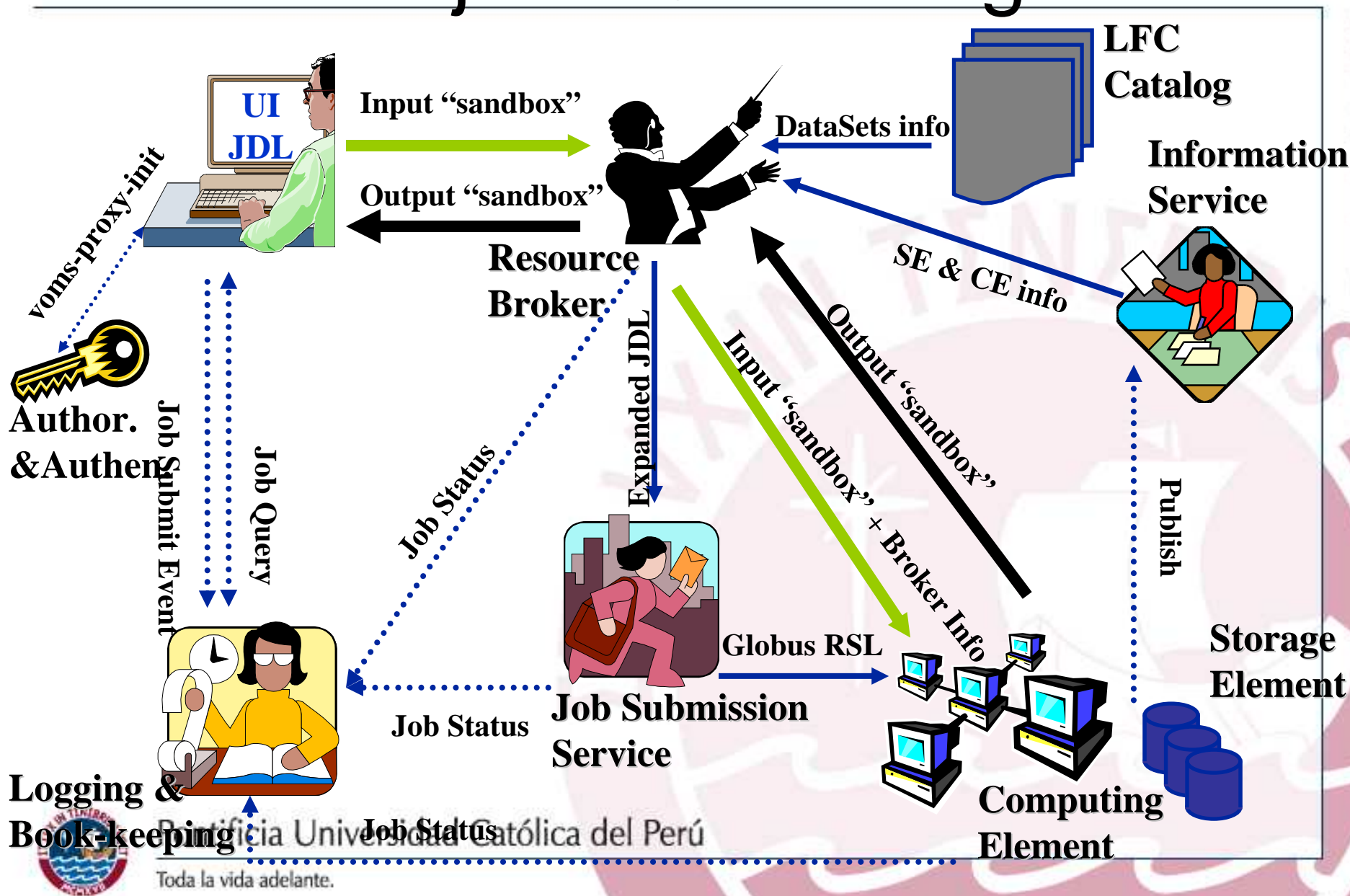
- Es el sucesor del LCG-2, desarrollado también por el EGEE.
- Usa también GT2 y Condor.
- Actualmente disponible g-Lite 3.0
 - <http://glite.web.cern.ch/glite/>
- La distribución Scientific Linux incluye G-Lite
 - <https://www.scientificlinux.org/>
- Posee un portal web para facilidad de los usuarios llamado Genius (Grid Enabled web eNvironment for site Independent User job Submission)
 - <https://genius.ct.infn.it>



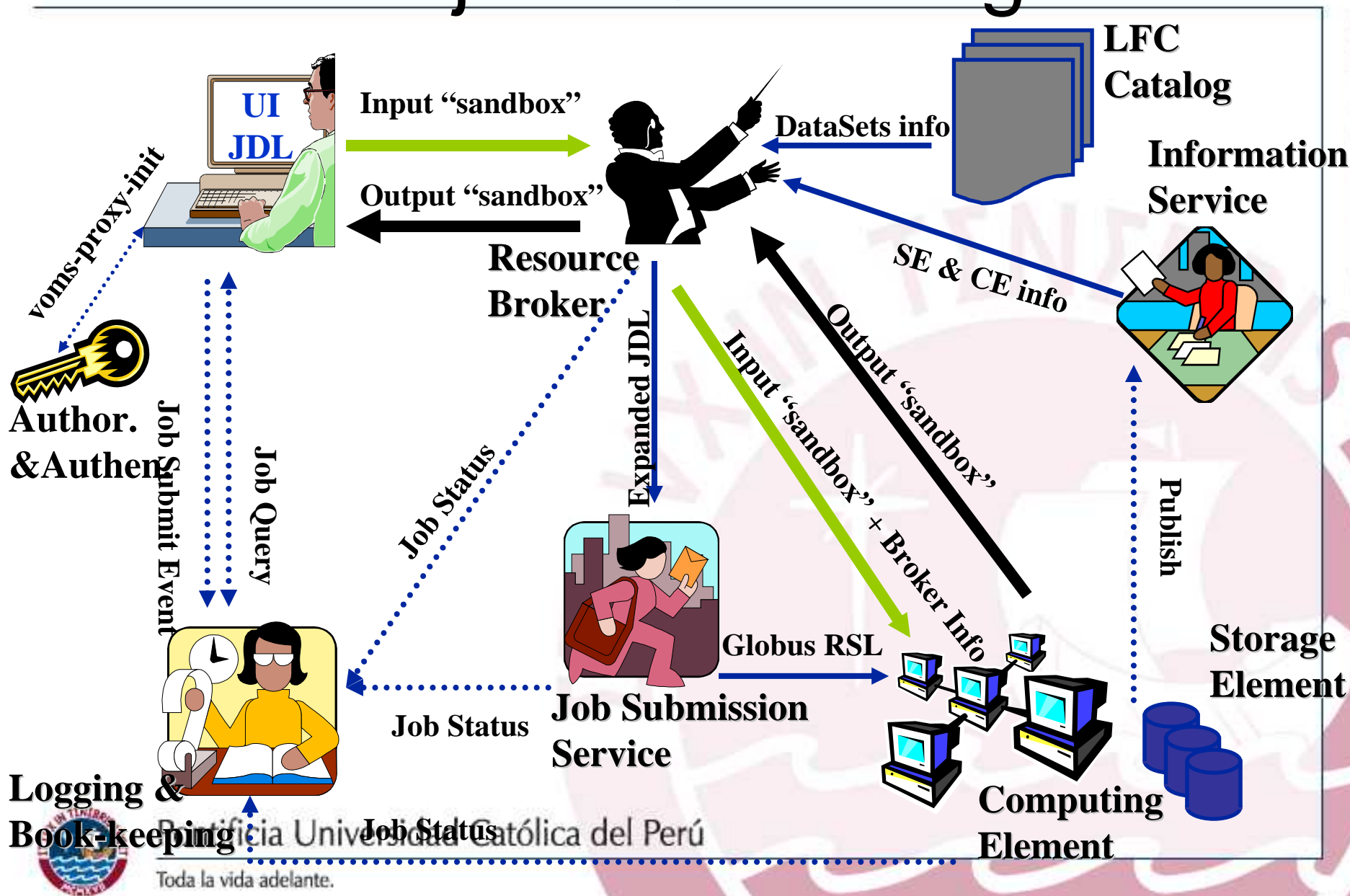
g-Lite



Flujo de Jobs en gLite



Flujo de Jobs en gLite



Aplicaciones en GENIUS

The screenshot displays the GENIUS web portal in a Mozilla browser window. The browser title is "Welcome to the GENIUS INFN GRID Portal - Mozilla" and the address bar shows "https://grid019.ct.infn.it/". The page features logos for INFN (Istituto Nazionale di Fisica Nucleare), EnginFrame, genius, and eGee (Enabling Grids for E-science in Europe). The main content area is titled "Grid Enabled web eNvironment for site Independent User job Submission" and shows a 3D visualization of the ALICE detector. The visualization is split into four views: Previous, Top View, Side View, and Front View. A control panel on the right lists detector components with status indicators: VZERO is on, EMCAL is on, START is on, PMD is on, PHOS is on, MUON is on, FMD is on, TRD is on, ZDC is on, RICH is on, TOF is on, TPC is on, and ITS is on. The browser status bar at the bottom shows the URL "https://grid019.ct.infn.it/infm_grid.xml? uri=//infm_grid/open-desktop".



Pontificia Universidad Católica del Perú

Toda la vida adelante.

EELA



Pontificia Universidad Católica del Perú

Toda la vida adelante.

EELA

- • E-infraestructura shared between Europe and Latin America.
 - <http://www.eu-eela.org/>
- Proyecto EELA creará en 2 años una red humana dedicada a trabajar en Grids (Mallas), e-Infraestructura y e-Ciencia.
- Comenzó su ejecución el 1 de enero de 2006.



Miembros de EELA

- EELA es coordinado por CIEMAT (España) (Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas)
- **Miembros de Europa:** CERN; CSIC, REDIRIS, UC y UPV (España), INFN (Italia) y LIP (Portugal)
- **Miembros de América Latina:** UNLP (Argentina), CEDERJ, RNP, UFF y UFRJ (Brasil), REUNA, UDEC y UTFSM (Chile), CUBAENERGIA (Cuba), UNAM (México), SENAMHI (Perú) y ULA (Venezuela).





EELA - Recursos

- Infraestructura de 20,000 CPU disponibles 24 x 7.
- Cerca de About 5 Petabytes (5 million Gigabytes) of storage.
- Usa LCG y g-Lite como Middleware.
- Para los talleres de entrenamiento se hace uso de la infraestructura Gilda (Grid In Laboratory for Dissemination Activities)
 - <https://gilda.ct.infn.it/>



GILDA



20 sites en 3 continentes



Pontificia Universidad Católica del Perú

Toda la vida adelante.

Tutoriales EELA

- Cronograma
 - <http://indico.eeela.org/categoryDisplay.py?categId=7>
- Cuarto tutorial EELA realizado en Santiago – Chile, el 6 y 7 de Septiembre.
- Material:
 - <http://indico.eeela.org/conferenceTimeTable.py?confId=52>



Estructura del proyecto EELA

4 Work Packages, 10 Tasks, Recursos solicitados de 2 M Euros

WP 1 MANAGEMENT				WP 2 INTEROPERABLE e-INFRASTRUCTURE			
TASK 1.1 MANAGEMENT				TASK 2.1 OPERATION MANAGEMENT			
				TASK 2.2 CERTIFICATION AUTHORITY			
				TASK 2.3 COMPUTING RESOURCES			
				TASK 2.4 NETWORK SUPPORT AND OPERATIONS			
WP 3 GRID ENHANCED APPLICATIONS				WP 4 DISSEMINATION AND TRAINING			
TASK 3.1 HIGH ENERGY PHYSICS				TASK 4.1 DISSEMINATION			
TASK 3.2 BIOMEDICINE				TASK 4.2 TRAINING			
TASK 3.3 ADDITIONAL							



Características de EGEE

- Acceso simplificado (acceso para todos los recursos operacionales que el usuario necesita).
- Computación bajo de manda (acceso rápido a los recursos por medio de su asignación eficiente).
- Acceso pervasivo (accesible de cualquiera localización geográfica).
- Recursos de gran escala (de una escala más grande que de cualquiera centro de computación sólo).
- Compartimento de software y datos (de modo transparente).
- Soporte realizado (aprovechar conocimientos de todos los participantes para ofrecer soporte amplio a aplicaciones clave).



EELA: Objetivos

- Establecer una infraestructura interoperable entre EU y LA.
- Identificar aplicaciones de investigación científica conjunta entre EU y AL.
- Incentivar participación en nuevos proyectos en nivel nacional e internacional.
- Organizar dos eventos generales de alto nivel por año.
- Promover oficinas temáticas especializadas.
- Promover participación en los foros relevantes relacionados a computación en grid.
- Aumentar conciencia sobre computación en grid.



Aplicaciones: GATE

- Geant4 Application to Tomographic Emission (GATE)
 - <http://www-lphe.epfl.ch/GATE>
- GATE es un entorno para la simulación Monte-Carlo de la emisión física de partículas en el campo del desarrollo médico
- The Interest of the LA Community is Led by CUBAENERGÍA



Aplicaciones: BiG

- *BiG: BLAST in Grid* es una interface para acceso a herramientas BLAST en la grid.
- *BLAST (Basic Local Alignment Search Tool)* es un procedimiento bioinformático aplicado a la identificación proteínas compatibles y secuencias de nucleotidos en bases de datos de proteínas y DNA.
- *BLAST* requiere cómputo intensivo.



Pontificia Universidad Católica del Perú

Toda la vida adelante.

EELA



Formulario para nuevas comunidades que usen g-Lite

http://www.eu-eela.org/private/eela_new_communities_form.php

Material de entrenamiento

<http://documents.eu-eela.org/?c=Training+Material&as=0&ln=en>



Pontificia Universidad Católica del Perú

Toda la vida adelante.

¿Preguntas?



Pontificia Universidad Católica del Perú

Toda la vida adelante.

Gracias!!!



Pontificia Universidad Católica del Perú

Toda la vida adelante.