



PARQUE CIENTÍFICO MURCIA

Uso del supercomputador Ben Arabí

CENTRO DE
SUPERCOMPUTACIÓN

José Ginés Picón López
Técnico de aplicaciones

Murcia a 2 de Febrero de 2012



Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- Acceso al Sistema*
- Espacio de Almacenamiento Permanente*
- Espacios de Ejecución
- Entorno de Ejecución*
- Sistema de Colas LSF*
- Ejemplos*
- Transferencia de Ficheros*



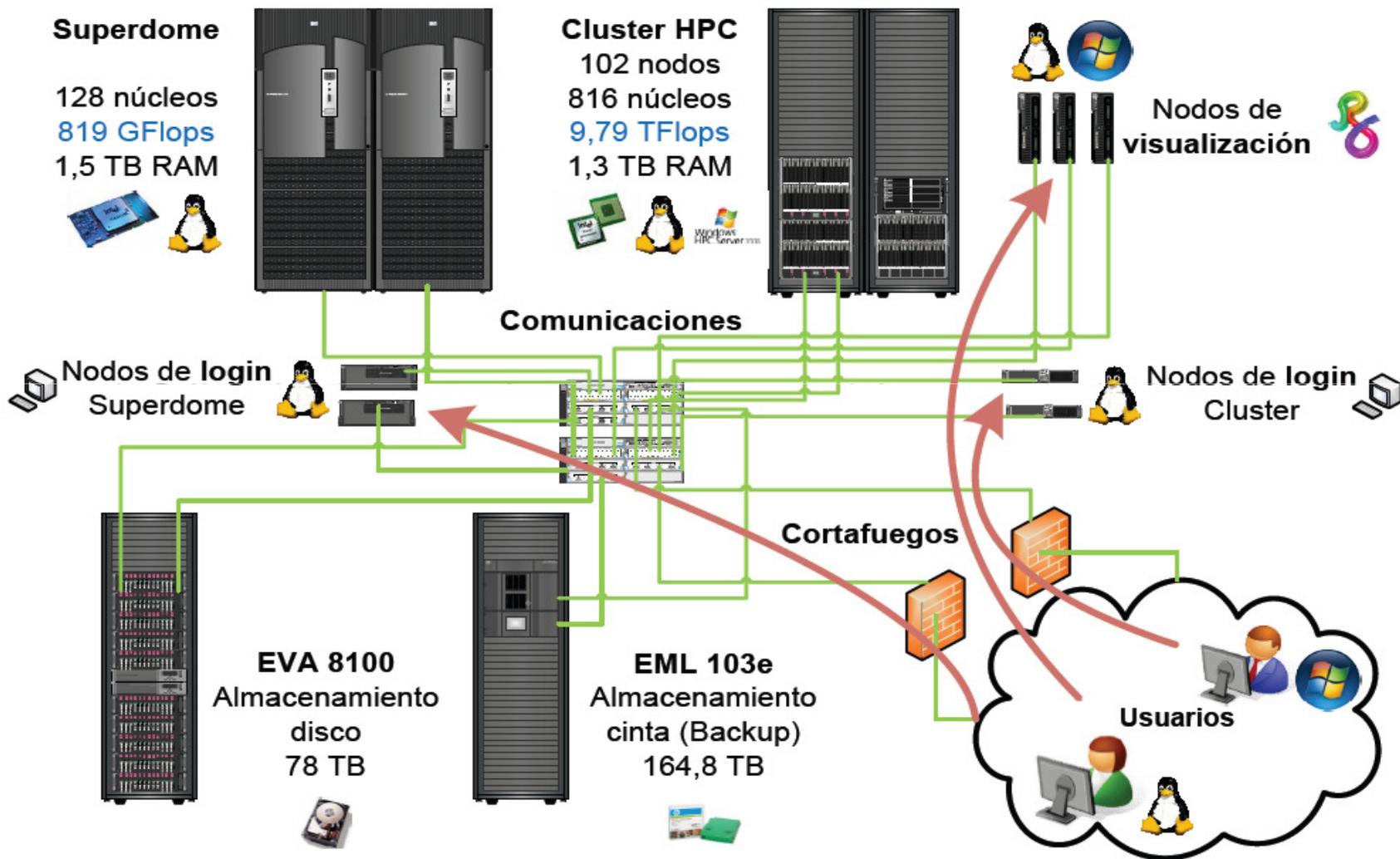
Uso del supercomputador Ben Arabí

- **Descripción de la Arquitectura**
- Acceso al Sistema*
- Espacio de Almacenamiento Permanente*
- Espacios de Ejecución
- Entorno de Ejecución*
- Sistema de Colas LSF*
- Ejemplos*
- Transferencia de Ficheros*



PARQUE CIENTÍFICO MURCIA

Ben Arabí Supercomputer



Unión Europea
Fondo Europeo de
Desarrollo Regional

"Una manera de hacer Europa"



Región de Murcia



GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA E INNOVACIÓN





PARQUE CIENTÍFICO MURCIA

Descripción de la Arquitectura I





PARQUE CIENTÍFICO MURCIA

Descripción de la Arquitectura II

Características del nodo de memoria compartida Superdome

Capacidad	819 Gflops
Procesador	Intel Itanium-2 Dual-Core Montvale
Nº de núcleos	128
Memoria Compartida	1,5 TB DDR-2
Memoria Caché	18 MB L3
Frecuencia de reloj	1,6 Ghz
Discos de trabajo temporal	40 x 146 GB SAS = 5,84 TB



PARQUE CIENTÍFICO MURCIA

Descripción de la Arquitectura III

Características del Cluster

Capacidad	9,72 Tflops
Procesador	Intel Xeon Quad-Core E5450
Número de nodos	102
Nº de núcleos	816
Memoria/Nodo	32 nodos de 16 GB y 70 de 8 GB
Memoria/Core	3 MB (6 MB compartidos entre 2 núcleos)
Frecuencia de reloj	3 Ghz



PARQUE CIENTÍFICO MURCIA

Red Infiniband

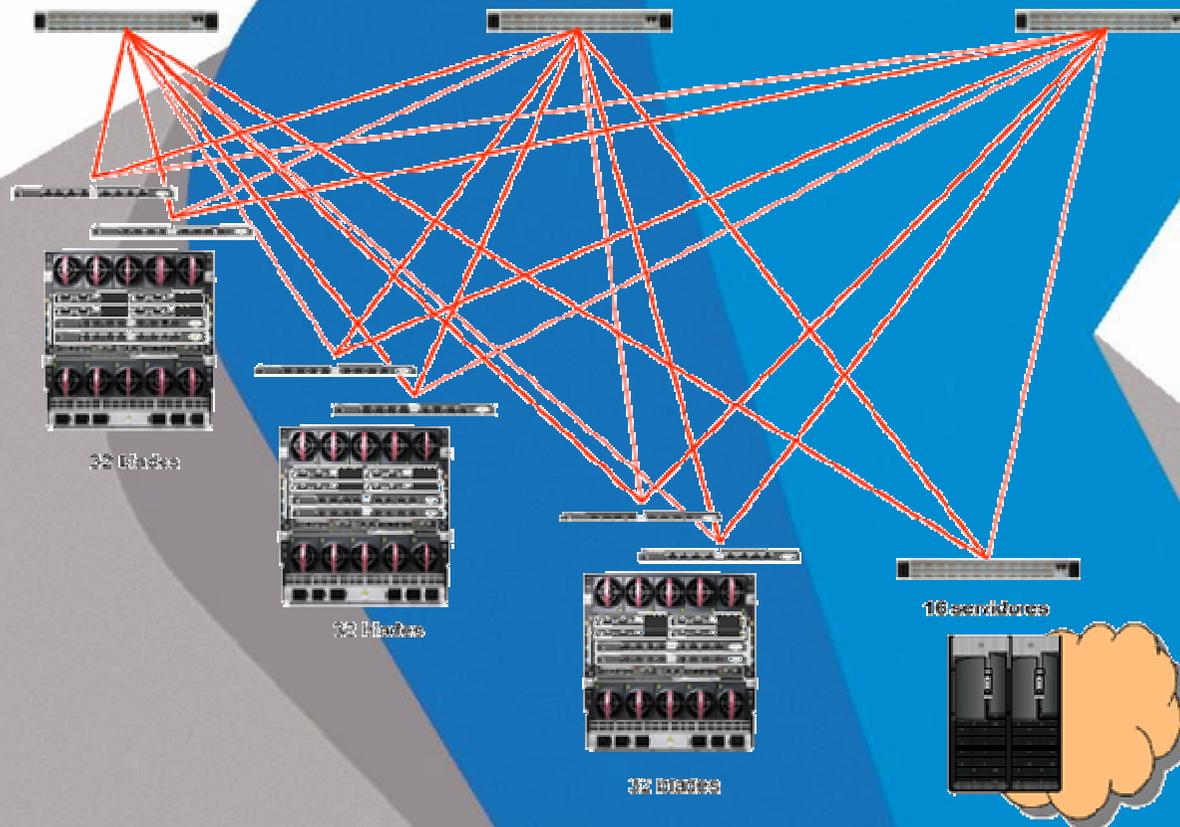
Red cálculo de fibra óptica

Altas Prestaciones

Interconecta todos los nodos de Ben Arabí

Topología Infiniband

— 2x líneas IB
— 3x líneas IB





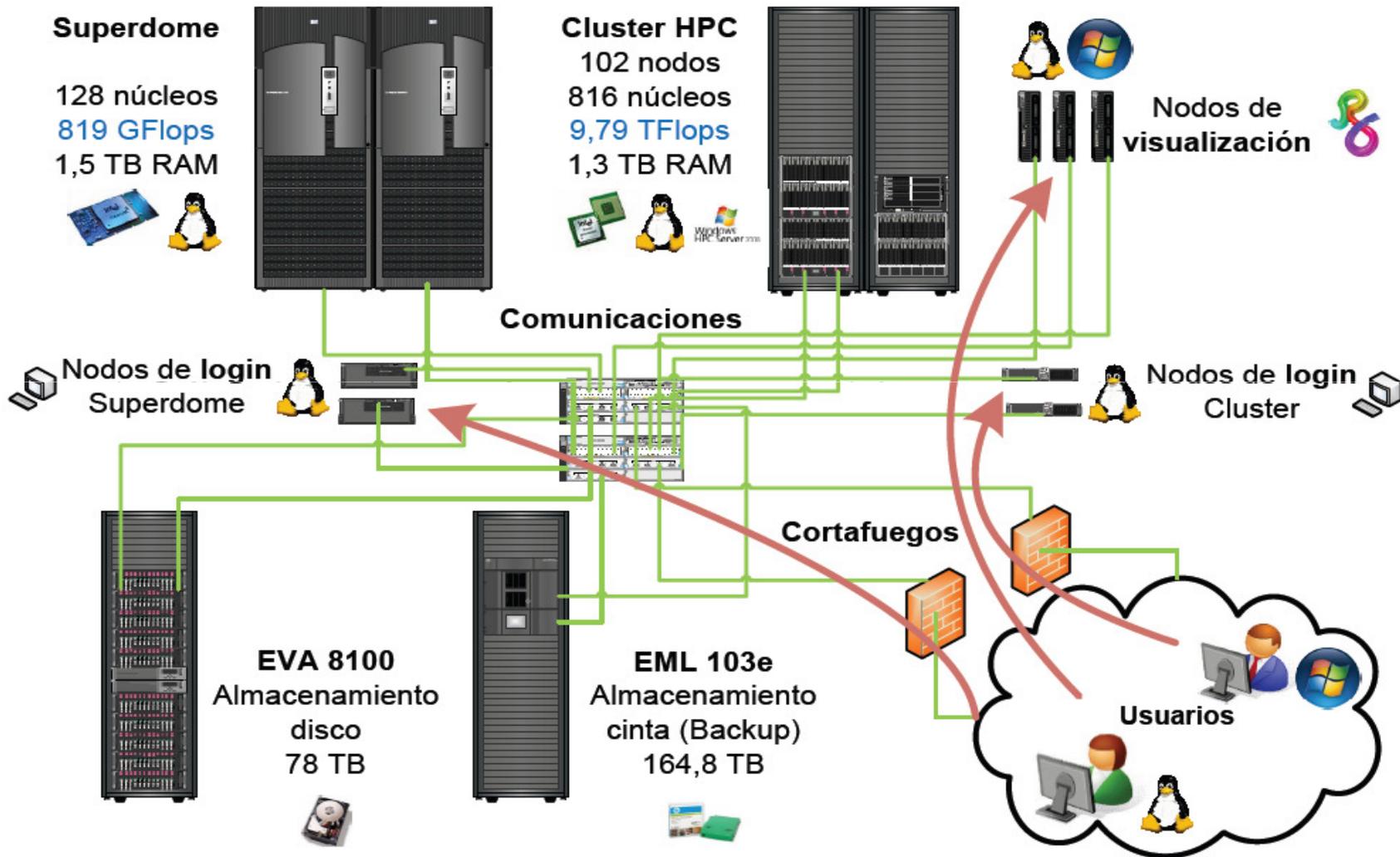
Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- **Acceso al Sistema***
- Espacio de Almacenamiento Permanente*
- Espacios de Ejecución
- Entorno de Ejecución*
- Sistema de Colas LSF*
- Ejemplos*
- Transferencia de Ficheros*



PARQUE CIENTÍFICO MURCIA

Ben Arabí Supercomputer



Unión Europea
Fondo Europeo de Desarrollo Regional

"Una manera de hacer Europa"



Región de Murcia



GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA E INNOVACIÓN





Acceso al Sistema I

- ❑ Acceso a través de los nodos de login via ssh.
 - **ben.fpcmur.es** para acceder al sistema superdome
 - **arabi.fpcmur.es** para acceder al cluster.

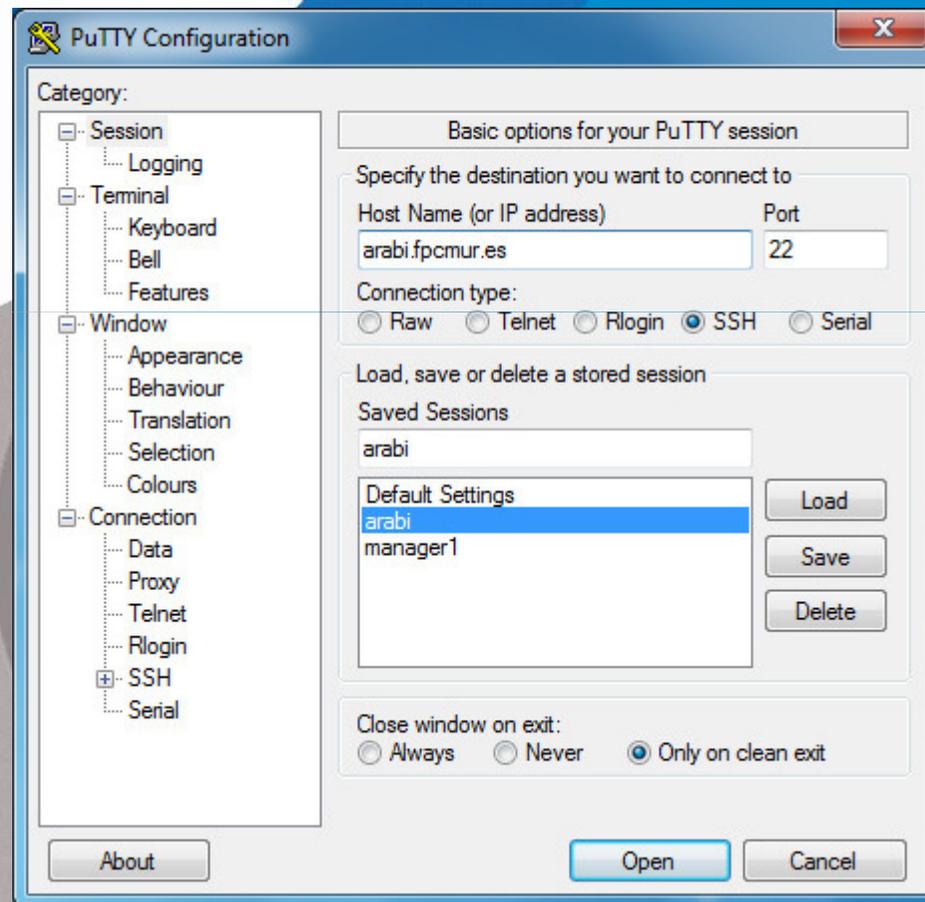
- ❑ Desde Linux o Mac:

```
$ ssh usuario@<nombre_maquina>
```

- ❑ Desde Windows tendremos que utilizar un cliente ssh como PuTTY:(<http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html>)



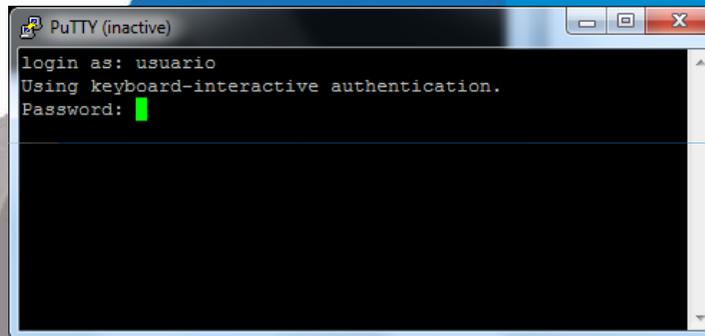
Acceso al Sistema II





Acceso al Sistema III

- ❑ Pulsamos el botón open e introducimos nuestro usuario y password:



```
PuTTY (inactive)
login as: usuario
Using keyboard-interactive authentication.
Password: █
```

- ❑ En el nodo de login se puede:
 - Compilar código fuente
 - Enviar trabajos a la cola
 - Chequear el estado de los trabajos
 - Borrar trabajos de la cola
 - Crear y editar ficheros



Acceso al Sistema IV

- ❑ La contraseña debe cumplir una serie de restricciones:
 - Debe tener como mínimo 8 caracteres
 - Debe contener al menos una mayúscula
 - Debe contener al menos una minúscula
 - Debe contener al menos un dígito
 - Debe ser una contraseña NUNCA ANTES establecida para el usuario



PARQUE CIENTÍFICO MURCIA

Acceso al Sistema V

Using username "usuario".

Using keyboard-interactive authentication.

Password: **CONTRASEÑA_PROVISIONAL**

Using keyboard-interactive authentication.

Warning: password has expired.

Password: **CONTRASEÑA_PROVISIONAL**

Using keyboard-interactive authentication.

New Password: **CONTRASEÑA_NUEVA**

Using keyboard-interactive authentication.

Repeat New

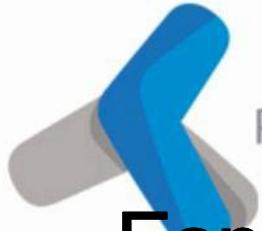
Password: **CONTRASEÑA_NUEVA**

usuario@arabi1:~>



Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- Acceso al Sistema*
- **Espacio de Almacenamiento Permanente ***
- Espacios de Ejecución
- Entorno de Ejecución*
- Sistema de Colas LSF*
- Ejemplos*
- Transferencia de Ficheros*



Espacios de Almacenamiento Permanente I

- Espacio personal y privado en `/users/<USUARIO>`, compartido tanto por superdome como por cluster.
- Pensado para almacenar código fuente, ejecutables y ficheros de trabajo diario.
- Capacidad de 10 GB por defecto.
- Copia de seguridad regularmente.
- Adicionalmente existe un espacio accesible por todos los miembros del proyecto en `/project/<PROYECTO>`



Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- Acceso al Sistema*
- Espacio de Almacenamiento Permanente*
- **Espacios de Ejecución**
- Entorno de Ejecución*
- Sistema de Colas LSF*
- Ejemplos*
- Transferencia de Ficheros*



Espacios de Ejecución I

- ❑ Cada sistema dispone de un espacio de alto rendimiento para el almacenamiento temporal en tiempo de ejecución.
 - /scratch: En el cluster
 - /scratchSD: Para Superdome

- ❑ Al final de la ejecución debe mover los datos a algún espacio de almacenamiento permanente.



Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- Acceso al Sistema*
- Espacio de Almacenamiento Permanente*
- Espacios de Ejecución
- **Entorno de Ejecución***
- Sistema de Colas LSF*
- Ejemplos*
- Transferencia de Ficheros*



Entorno de ejecución I

- ❑ Se gestiona a través del paquete **modules**.
- ❑ Permite modificación del entorno de forma dinámica.
- ❑ Comandos útiles:
 - *module avail*
 - *module load <aplicación>*
 - *module list*
 - *module unload <aplicación>*
 - *module show <aplicación>*

Aplicaciones Científicas Disponibles



Renderizado 3D



Multiphysics HPC

CFD (Fluent & CFX) HPC



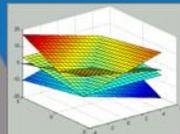
Método de elementos finitos



Mecánica de fluidos

ATLAS

Álgebra lineal



Química cuántica



Análisis estadístico y gráfico



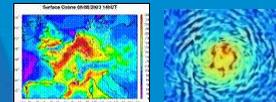
Dinámica molecular



Cálculos matemáticos

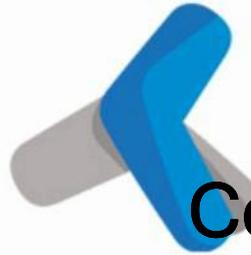
CHIMERE, MM5 y WRF

Cálculos meteorológicos y climáticos



LISTADO COMPLETO en web:

http://www.cesmu.es/computacion/index_aplicaciones.aspx



Compiladores, MPI y OpenMP en Ben Arabí

Compiladores:

- GCC 4.1.2 (sistema), 4.4.2, 4.5.1
- Intel 11.1

MPI:

- Open MPI 1.4.1 para GCC e Intel
- Intel MPI 3.2.1, 3.2.2, 4.0.0
- Platform MPI 8.0

Ben Arabí Supercomputer	MPI	OpenMP	MPI + OpenMP
Xeon Cluster (Arabi)	SI	SI (hasta 8 hilos)	SI (hasta 8 hilos por proceso)
Itanium Superdome (Ben)	SI	SI	SI



Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- Acceso al Sistema*
- Espacio de Almacenamiento Permanente*
- Espacios de Ejecución
- Entorno de Ejecución*
- **Sistema de Colas LSF***
- Ejemplos
- Transferencia de Ficheros*



Sistema de Colas LSF I

- Generalmente, los trabajos requieren más recursos que los disponibles para procesos interactivos.
- Es necesario un sistema para la gestión y planificación.
- LSF (Load Sharing Facility).
- Los trabajos se envían a la cola a través de un script o mediante línea de comandos.
- El comando más importante es **BSUB**.



PARQUE CIENTÍFICO MURCIA

Sistema de Colas LSF II

Opciones más comunes del comando BSUB

-J nombre_trabajo	Asigna un nombre al trabajo
-u email	Indica la dirección de correo
-B	Envía un correo al empezar el trabajo
-N	Envía un correo al finalizar el trabajo
-e fichero_error	Redirige stderr al fichero especificado
-o fichero_salida	Redirige stdout al fichero especificado
-q nombreCola	Especifica la cola que se va a usar
-n num_core	Especifica el número de cores



Sistema de Colas LSF III

- ❑ Envío de trabajos:
 - $\$ bsub < nombre_script$

- ❑ Consulta de colas disponibles:
 - $\$ bqueues$

- ❑ En general, las colas siguen la siguiente nomenclatura

[Tipo de sistema]_[Nº de cores]_[Nº de horas]

- ❑ Por ejemplo, la cola arabi 128x96h nos permite ejecutar en 128 cores del cluster durante 96 horas.



Sistema de Colas LSF IV

Monitorización del estado de los trabajos:

- *\$ bjobs*

PEND: esperando en la cola para ser atendido.

RUN: enviado a un host y ejecutandose.

Consulta de la salida del trabajo en ejecución:

- *\$ bpeek <jobid>*

Eliminar un trabajo:

- *\$ bkill <jobid>*



Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- Acceso al Sistema*
- Espacio de Almacenamiento Permanente*
- Espacios de Ejecución
- Entorno de Ejecución*
- Sistema de Colas LSF*
- **Ejemplos***
- Transferencia de Ficheros*



Ejemplo I (Secuencial)

```
#!/bin/bash
#BSUB -J nombre_trabajo           # Nombre del trabajo
#BSUB -o salida.%J.out            # Nombre del fichero de salida
#BSUB -e error.%J.err            # Nombre del fichero de error
#BSUB -q nombre_cola              # Nombre de la cola utilizada

# Cargamos Entorno Necesario
source /etc/profile.d/modules.sh
module load nombre_modulo_aplicacion

# Se ejecuta el trabajo
comando_aplicacion
```



Ejemplo II

- ❑ El envío de los scripts al supercomputador se realiza con el comando `bsub`:

`bsub < nombre_script`

- ❑ También es posible el envío mediante línea de comando:

`$ bsub -J nombre_trabajo -e error.%J.err -o salida.%J.out -q nombreCola "module load nombre_modulo_aplicacion; comando_aplicación"`



Ejemplo III (OpenMP + MPI)

```
#!/bin/bash
#BSUB -J impiomp_hello
#BSUB -o %J.out
#BSUB -e %J.err
#BSUB -q arabi_formacion
#BSUB -n 16
export THREADS=2
export PPN=1
source /etc/profile.d/modules.sh
module load impi/3.2.2.006
export I_MPI_PIN_DOMAIN=omp
echo -e "==" Serial =="
```

#Número de hilos OpenMP
#Procesos MPI por nodo

```
./impiomp_hello 1
echo -e "\n==" Parallel =="
```

```
impiexec.lsf -ppn $PPN ./impiomp_hello $THREADS
```



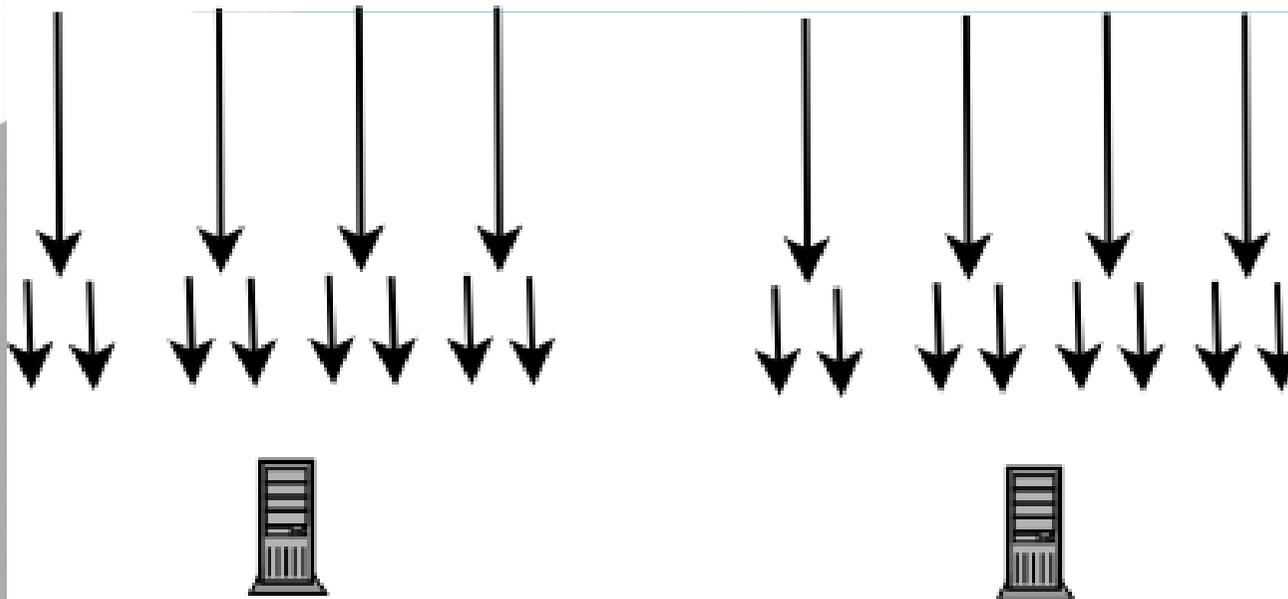
PARQUE CIENTÍFICO MURCIA

Ejemplo IV (OpenMP+MPI)

$n=16$

PPN=4

TREADS=2





Ejemplo V (OpenMP+MPI)

- ❑ Copiamos el directorio `/project/FORMACION/impi_omp` a nuestro directorio de trabajo.
- ❑ Dentro de dicho directorio ejecutamos:

```
module load impi/3.2.2.006  
make  
bsub < bsub.main
```



Uso del supercomputador Ben Arabí

- Descripción de la Arquitectura
- Acceso al Sistema*
- Espacio de Almacenamiento Permanente*
- Espacios de Ejecución
- Entorno de Ejecución*
- Sistema de Colas LSF*
- Ejemplos
- **Transferencia de Ficheros***



Transferencia de Ficheros I

- ❑ En **Linux y Mac** la transferencia se realiza mediante el comando scp:

```
scp [-1246BCpqrsv] [-c cipher] [-F ssh_config] [-i  
identity_file] [-l limit] [-o ssh_option] [-P port] [-S  
program] [[user@]host1:]file1 [...] [[user@]host2:]file2
```

- ❑ Para transferir el fichero fich1 desde nuestro equipo local hacia nuestro espacio de almacenamiento /users/<USUARIO>:

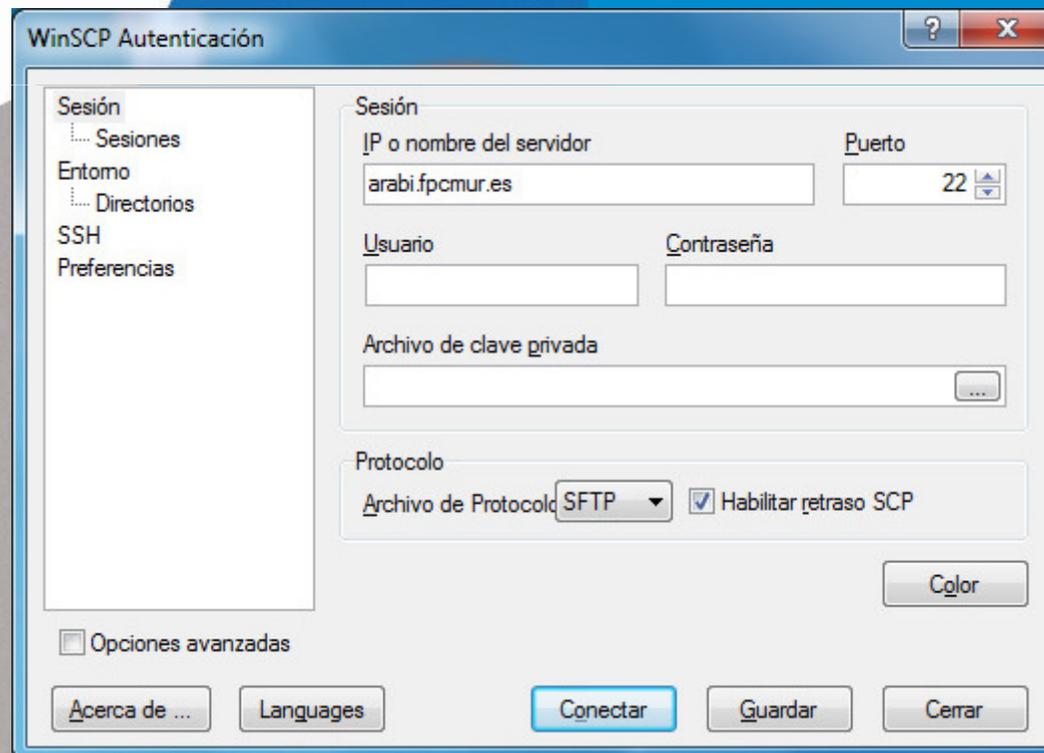
```
$ scp fich1.txt  
<USUARIO>@arabi.fpcmur.es:/users/<USUARIO>/
```



PARQUE CIENTÍFICO MURCIA

Transferencia de Ficheros II

- ❑ En Windows utilizamos winSCP para la transferencia de ficheros: (<http://winscp.net/eng/download.php#download2>).





PARQUE CIENTÍFICO MURCIA

Transferencia de Ficheros III

The screenshot shows the WinSCP interface with two panes. The left pane shows the local Windows file system at 'C:\Users\jpicon\FPCMUR\Documents'. The right pane shows the remote Linux file system at '/users/jpicon'. The remote directory contains various configuration and application folders.

Nombre	Ext	Tamaño	Tipo	Modificado	Permisos
..			Directorio supe...	25/01/2011...	rwxr-xr-x
Mi música			Carpeta de arc...	25/01/2011...	
Mis imágenes			Carpeta de arc...	25/01/2011...	
Mis vídeos			Carpeta de arc...	25/01/2011...	
Default.rdp		1.986	Conexión a Esc...	26/01/2011...	
desktop.ini		402	Opciones de c...	25/01/2011...	
..			Directorio supe...	27/01/2011 12:...	rwxr-xr-x
.ansys			Carpeta de arc...	22/07/2010 9:2...	rwx-----
.cfx			Carpeta de arc...	20/07/2010 12:...	rwx-----
.ChaosGroup			Carpeta de arc...	10/12/2009 12:...	rwx-----
.config			Carpeta de arc...	10/12/2010 14:...	rwxr-xr-x
.dbus			Carpeta de arc...	10/12/2010 14:...	rwx-----
.eclipse			Carpeta de arc...	26/01/2010 15:...	rwx-----
.egg cups			Carpeta de arc...	20/09/2010 12:...	rwxr-x--
.fontconfig			Carpeta de arc...	25/08/2010 10:...	rwxr-xr-x
.fonts			Carpeta de arc...	25/08/2010 10:...	rwxr-xr-x
.gconf			Carpeta de arc...	11/01/2011 13:...	rwx-----
.gconfd			Carpeta de arc...	11/01/2011 13:...	rwx-----
.gnome			Carpeta de arc...	26/08/2010 14:...	rwxr-xr-x
.gnome2			Carpeta de arc...	11/01/2011 13:...	rwx-----
.gnome2_private			Carpeta de arc...	29/12/2009 9:2...	rwx-----
.gstreamer-0.10			Carpeta de arc...	10/12/2010 14:...	rwxr-xr-x
.hpremate			Carpeta de arc...	10/12/2010 14:...	rwxr-xr-x
.idb			Carpeta de arc...	29/01/2010 14:...	rwx-----
.idb_linux_intel64_app			Carpeta de arc...	26/01/2010 12:...	rwx-----
.kde			Carpeta de arc...	25/08/2010 10:...	rwxr-xr-x
.local			Carpeta de arc...	28/12/2009 8:3...	rwx-----

0 B de 2.388 B en 0 de 5

0 B de 3.253 MiB en 0 de 159

F2 Renombrar F4 Editar F5 Copiar F6 Mover F7 Crear directorio F8 Eliminar F9 Propiedades F10 Salir

SFTP-3 0:00:13



PARQUE CIENTÍFICO MURCIA

¿PREGUNTAS?





PARQUE CIENTÍFICO MURCIA

GRACIAS POR SU ATENCIÓN CONTACTA CON NOSOTROS:

Fundación Parque Científico Murcia
CENTRO DE SUPERCOMPUTACIÓN

José Guillén Mercader

Jefe de Proyectos

jguillen@parquecientificomurcia.es

María Eugenia Requena

Directora del Centro de Supercomputación

mrequena@parquecientificomurcia.es

Ctra de Madrid Km 388.

Complejo de Espinardo, Edificio S.

30100 Espinardo, Murcia

T: 968 27 78 30

www.parquecientificomurcia.es

www.cesmu.es



Murcia a 16 de Febrero de 2010



Supercomputador Ben Arabí : Arquitectura III

❑ Sistema de almacenamiento: EVA 8100

- ❖ El sistema almacenará sus datos intermedios y definitivos en un único sistema de almacenamiento en red.
- ❖ Capacidad **78 TB**.



LIBRERÍA DE CINTAS EML 103e



❑ Sistema de backup: LIBRERÍA DE CINTAS EML 103e

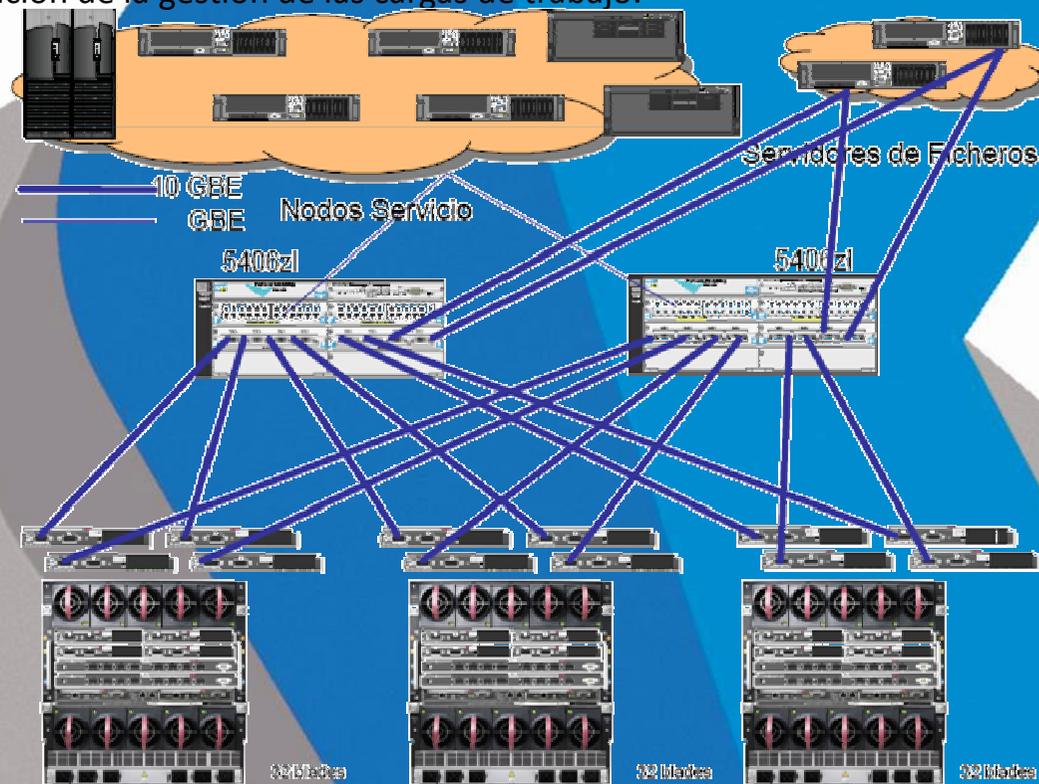
- ❖ El SC dispondrá de un sistema de copias de seguridad que permita el almacenamiento de la información y su posterior recuperación.
- ❖ Capacidad **164,8 TB**.



Supercomputador Ben Arabí : Arquitectura III

❑ Red de comunicaciones y gestión: 10GBE/GBE

Dada la complejidad de las tareas a desarrollar, se requiere un *sistema de gestión centralizado* común a los sistemas de cálculo y almacenamiento, que permita integrar las herramientas de particionamiento y de automatización de la gestión de las cargas de trabajo.





Supercomputador Ben Arabí : Software

❑ Software de Funcionamiento

Adicionalmente a las herramientas, aplicaciones y software necesarios para el correcto funcionamiento del nodo central de memoria compartida, el clúster computacional y los sistemas de almacenamiento y backup requeridos en los apartados anteriores, el sistema dispone de todas aquellas soluciones software necesarias para la su correcta y eficiente explotación.

En este sentido, se incluye:

- ❖ Software de gestión de colas, capaz de gestionar arquitecturas de memoria compartida sobre diferentes sistemas operativos.
- ❖ Software de gestión y optimización de recursos de procesamiento.
- ❖ Software de desarrollo, tales como compiladores, herramientas de optimización y mejora del rendimiento, librerías de paso de mensajes MPI, ...
- ❖ Software para el tratamiento de funciones matemáticas
- ❖ Software científico-técnico open-source