www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Recursos de Supercomputación en BSC-CNS, RES & PRACE 2013

EXCELENCIA SEVERO OCHOA

Tenerife, 29 de mayo 2013

Visión ESFRI del servicio HPC en Europa


BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación
EXCELENCIA
SEVERO
OCHOA
BARCELONA
SUPERCOMPUTING CENTER

BSC-CNS: Constitución

- Es el Centro Nacional de Supercomputación,
  - Consorcio constituido por:

  - Coordinador de la **Red Española de Supercomputación.**
  - Centro miembro del proyecto **PRACE**.
  - Participante en múltiples proyectos tales como Human Brain Project, Montblanc y proyectos industriales.

BSC-CNS: Misión

« La misión del BSC-CNS es **investigar, desarrollar** y **gestionar la tecnología** para facilitar el **progreso científico**.

- Proporciona **soporte de la supercomputación para la investigación externa** al BSC-CNS.
- Realiza I+D en **Ciencias de la Computación, Ciencias de la Vida** y **Ciencias de la Tierra**.

BSC-CNS: Dpt. Ciencias de la Computación

Arquitectura de computadores e interfaz de SO

Arquitecturas heterogéneas

Arquitectura de computadores paralelos

Herramientas de análisis de rendimiento

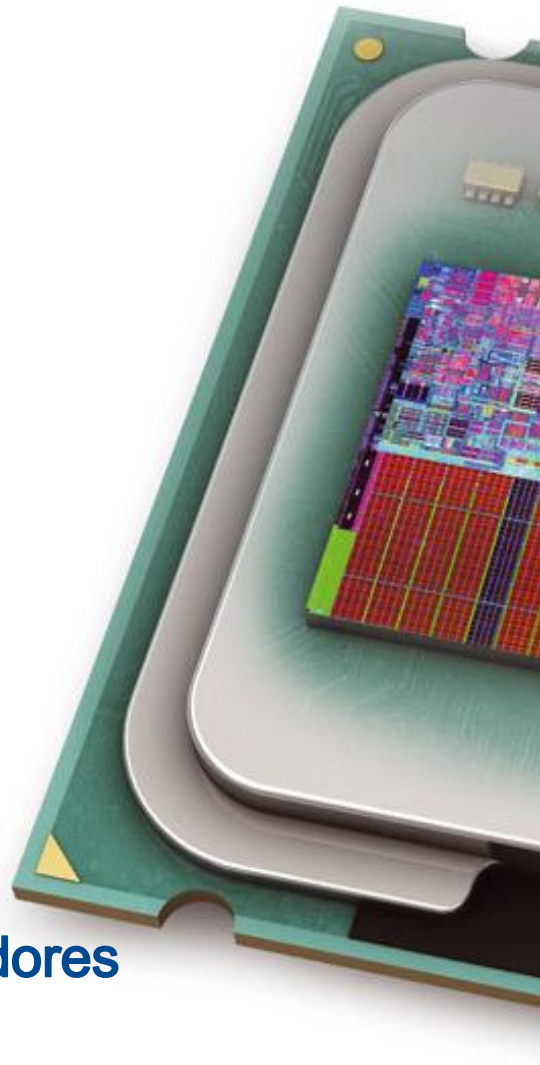
Modelos de programación

Sistemas y clusters Grid

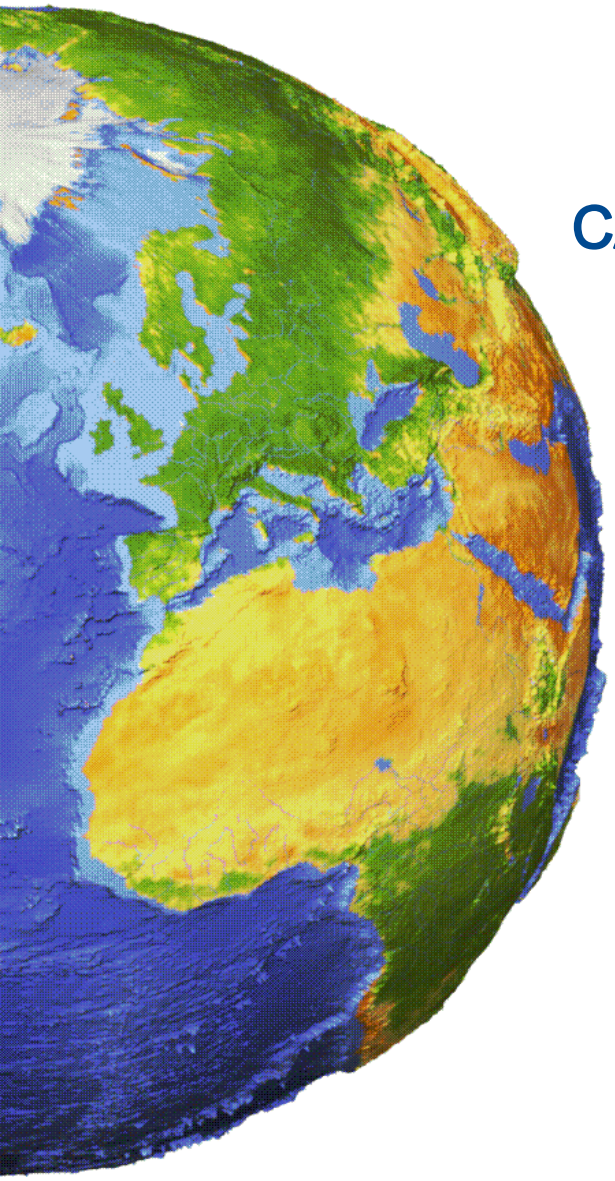
Sistemas autónomos y plataformas de e-Business

Sistemas de almacenamiento

Arquitectura no convencional de redes y computadores



BSC-CNS: Dpt. Ciencias de la Tierra



CALIOPE Sistema de Pronóstico de la CALIdad del aire Operacional Para España (y Europa)

Modelado Atmosférico y Predicción climática

Modelado global de polvo mineral

Transporte de polvo mineral

Transporte de ceniza volcánica

Estudios de impacto ambiental

BSC-CNS: Dpt. Ciencias de la Vida

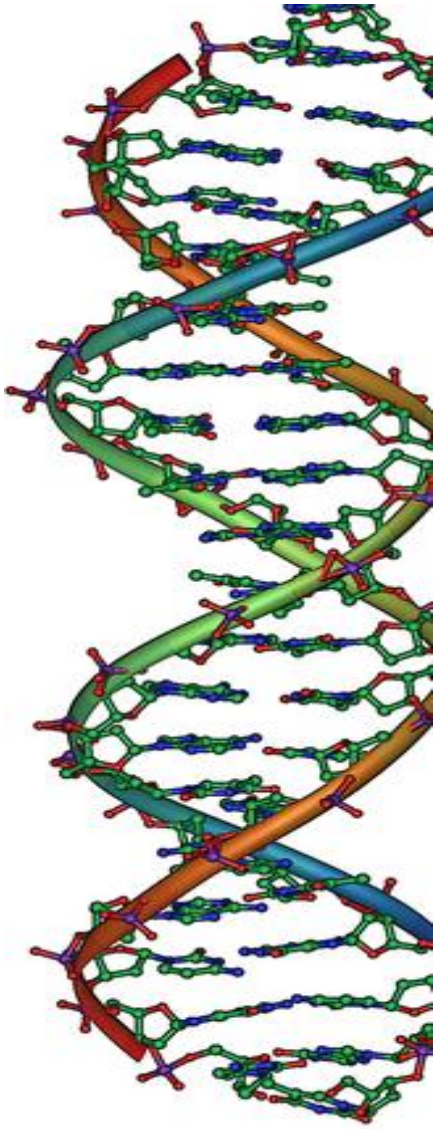
Bases estructurales de interacción proteína-proteína

Análisis de genomas y redes para modelar enfermedades, sistemas y evolución de organismos

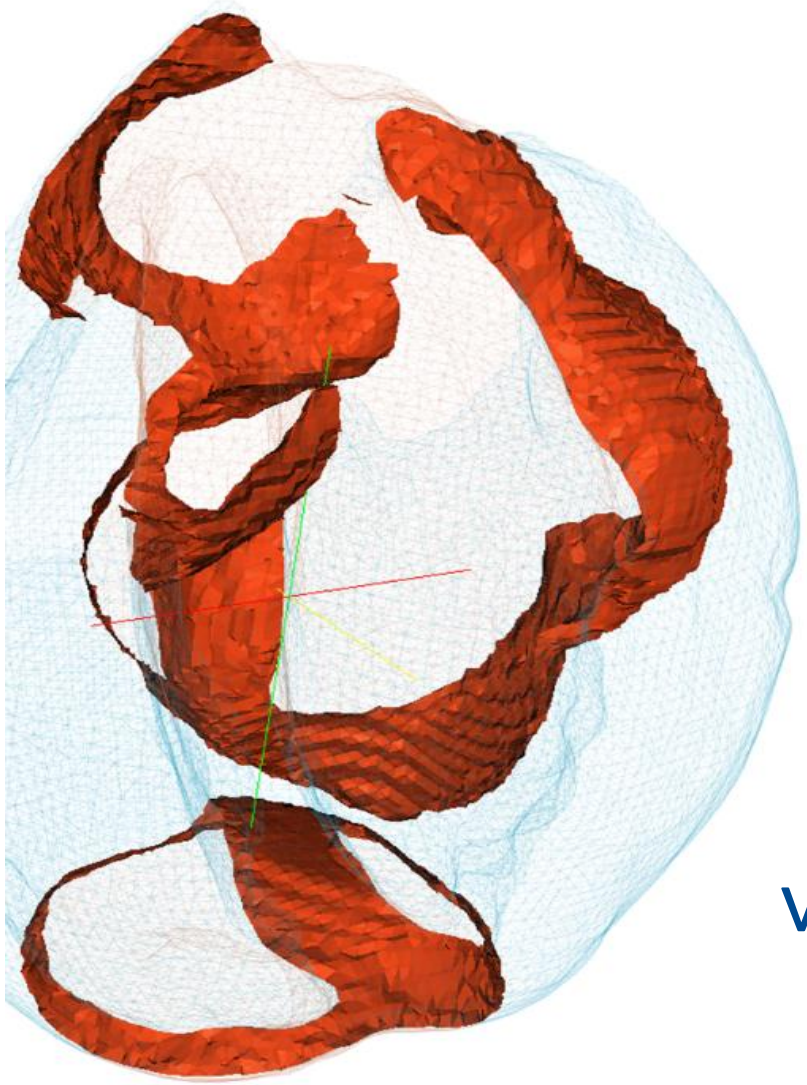
Modelado atómico (y electrónico) de bioquímica y biofísica de proteínas

Diseño de fármacos

Modelado micro y mesoscópico de macromoléculas



BSC-CNS: Dpt. Aplicaciones para Ciencia e Ingeniería



Dinámica de fluidos computacional

Mecánica de solidos computacional

Representación gráfica de seísmos

Simulaciones sociales

Flujos geofísicos

Paralelización de código

Optimización de código

Visualización y post-procesado de datos

BSC-CNS Microsoft Centro de Investigación

BSC Microsoft Research Centre

- Inauguración del centro en 2008
- Centrado en el desarrollo de aplicaciones con Memoria Transaccional, herramientas para facilitar el desarrollo de dichas aplicaciones y la propuesta de implementaciones escalables de Memoria Transaccional Híbrida.

Además:

- Investigación en hardware eficiente para el soporte de sistemas de ejecución de nuevos lenguajes de programación, a la vez que su sincronización y fiabilidad, Procesadores Vectoriales de bajo consumo y OmpSs@Barrelfish.

Colaboración Intel-BSC Exascale Lab

- Inicio de actividades en 2011

- Objetivo: Enfrentarse a los retos del camino hacia Exascale

Eficiencia, variabilidad, memoria, escalado (concurrencia) y complejidad respecto a jerarquía y heterogeneidad.

- Principales líneas de investigación:
  - OmpSs: modelo de programación para algoritmos y sistemas de ejecución responsables de asignarse a los recursos (autotuning dinámico, resiliencia, reducción y equilibrio de carga)
  - Desarrollo de herramientas de análisis: monitorización y modelización del rendimiento potencia, predicción de estrategias de el distribución de tareas.
  - Implementación de aplicaciones y algoritmos que contemplen la asincronía y la complejidad (MPI híbrido/OmpSs).

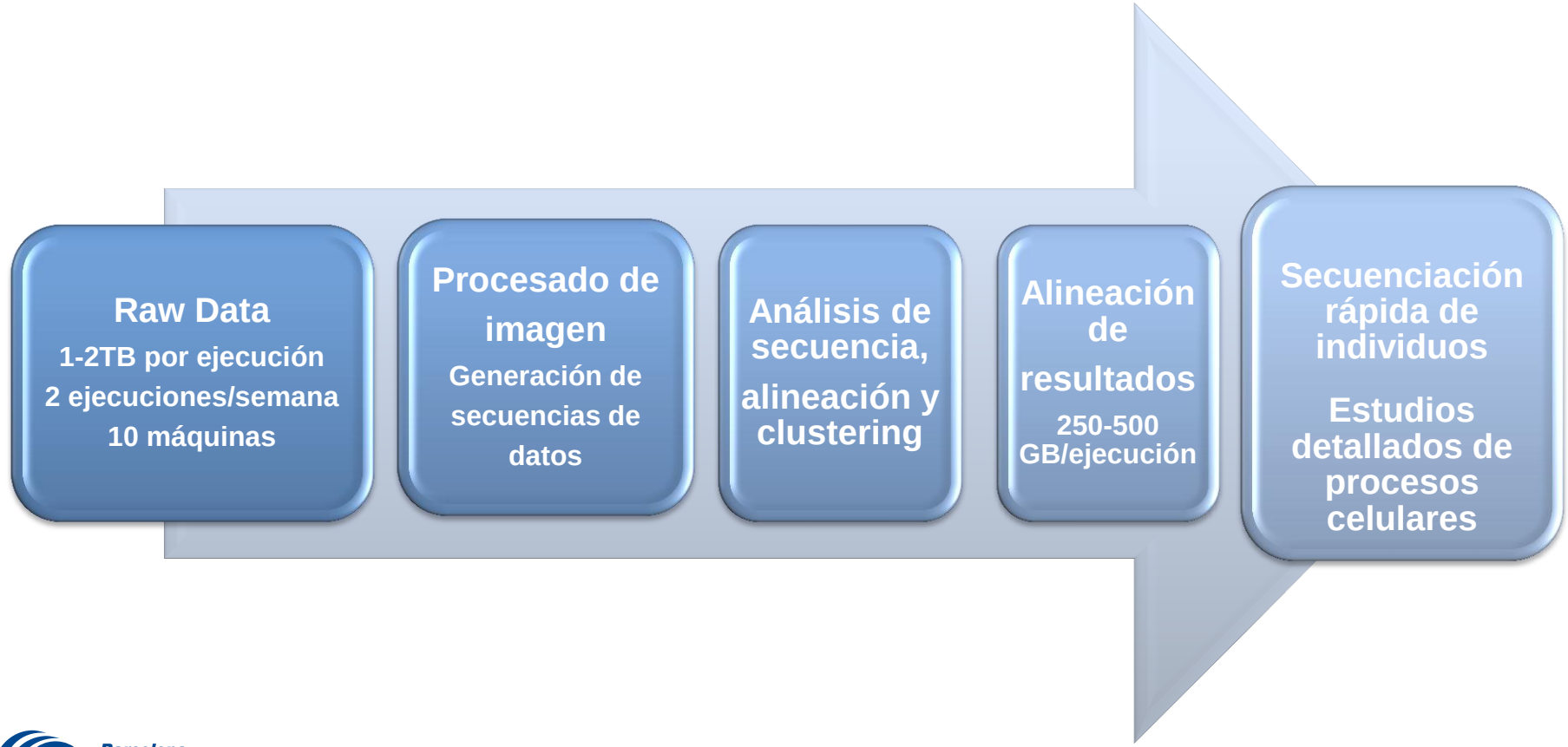
BSC-CNS Centro de Excelencia CUDA

- Desde 2010 es Centro de Investigación CUDA
- Desde 2011 reconocido como Centros de Excelencia NVDIA.
  - Modelos de programación: entorno de ejecución GMAC (Global Memory for Accelerators) , OmpSs para clusters híbrido CPU/GPU y auto-vectorización.
  - Desarrollo de aplicaciones
  - Crear programa de educación y escuela de verano para formación en CUDA, OpenCL y starSs.

BSC-CNS Soporte a CNAG

- Soporte a nivel de HPC, IT y datos para la secuenciación de nueva generación.

cnag
centre nacional d'anàlisi genòmica
centro nacional de análisis genómico



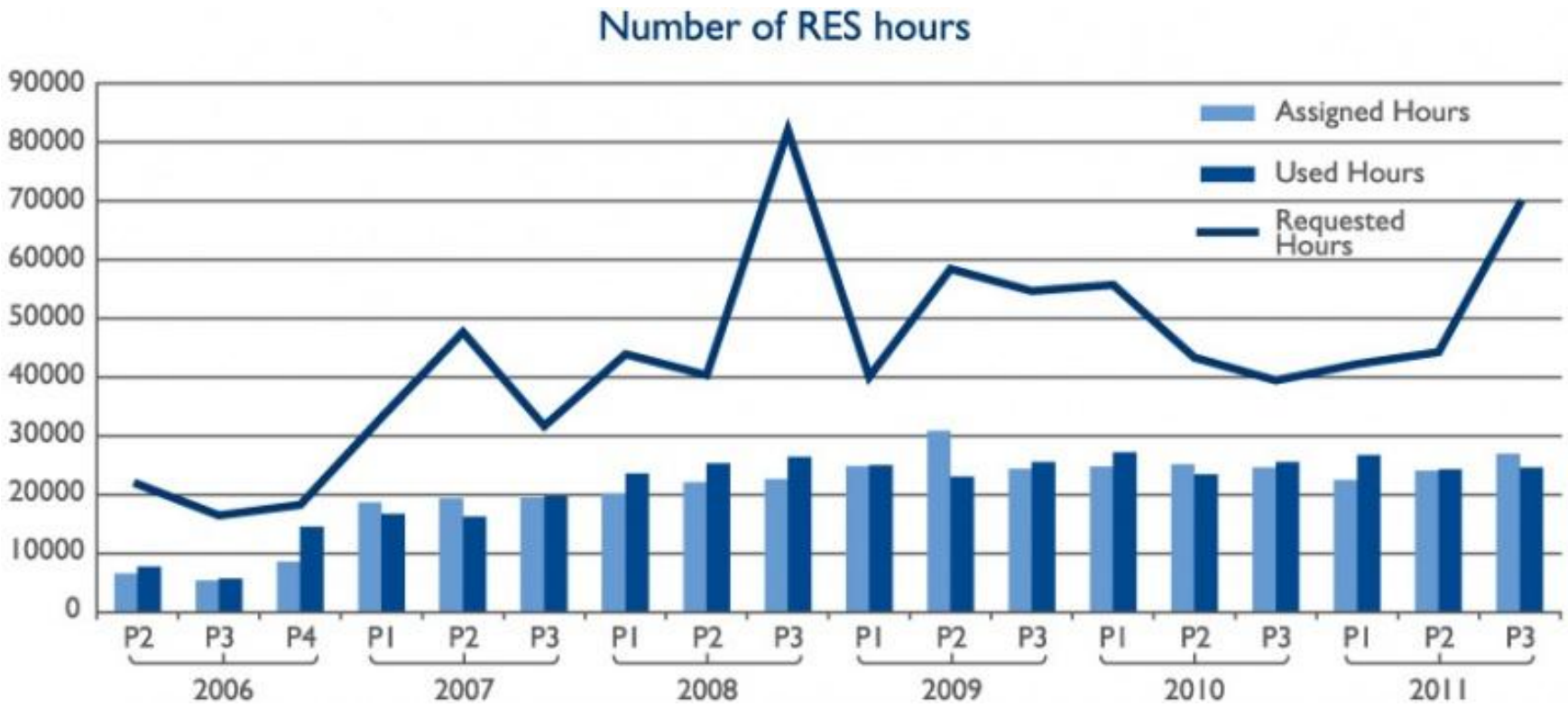
RED ESPAÑOLA DE SUPERCOMPUTACIÓN

Red Española de Supercomputación

- Desde 2006, la **RES** es una infraestructura virtual distribuida consistente en la interconexión de **supercomputadores**, que en régimen de trabajo compartido, gestionan su capacidad de cálculo y proporcionan **servicio a investigadores españoles**.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Volumen de proyectos con acceso a RES



Servicios principales ofrecido por la RES

- Principalmente: Optimización de código
  - Mejora de la paralelización y escalabilidad
  - Optimización de Entrada/Salida
  - Portabilidad de código (Arquitectura, procesador, aceleradores, GridSuperscalar, StarSs, …)
  - Depuración de código (Totalview, Paraver, DDT…)

Pero también:

- Almacenamiento de datos
- Actividades de test
- Programas de movilidad
- Soporte técnico
- Difusión científica

Más servicios ofrecido por la RES

Además:

- Reuniones de usuarios
- Cursos para usuarios
- Seminarios científicos
- Cursos para los equipos técnicos de la RES
- Colaboración para seminarios y conferencias:
  - Compartiendo experiencia y proporcionando expertos
  - Realizando cursos de formación específica
  - Co-financiando la celebración de eventos para promover la ciencia

Suscripción en http://www.bsc.es/hpc-events-trainings.xml

MareNostrum3

- **Racks de Cómputo: 36 IBM iDataPlex** *(actualmente 25)*
  - Cada Rack dispone de 84 IBM dx360 M4 nodos de cómputo:
    - 2x SandyBridge-EP E5-2670 2.6GHz/1600 20M 8-core 115W
    - 8x 4G DDR3-1600 DIMMs (2GB/core)
    - 500GB 7200 rpm SATA II local HDD
- **Nodos de cómputo en rack de gestión: 4x IBM dx460 M4**
- **Nodos de cómputo: 3028** *(actualmente 2104)*
  - 48.448 Intel cores *(actualmente 33664)*
- **Memoria: 94.62 TB** *(67.32 TB today)*
  - 32GB/node
- **Peak performance: 1.0 Pflop/s** *(actualmente 0.7 Pflop/s)*
  - Node performance: 332.8 Gflops
  - Rack Performance: 27.95 Tflops
- **Red Infiniband FDR10 con topologia non-blocking Fat Tree**
- **Consumo energético estimado: 1.08 MW** *(0.7 MW today)*
  - Rack Consumption: 28.04 kW/rack (nominal bajo HPL)

Hardware de MareNostrum2 vs MareNostrum3

		MN2	MN3
Cómputo	Cores/chip	2	**8**
	Chip/nodo	2	**2**
	Cores/nodo	4	**16**
	Nodos	2560	**3028**
	Total cores	10240	**48448**
Rendimiento	Frecuencia	2,3	**2,6**
	Gflops/core	9,2	**20,8**
	Gflops/node	36,8	**332,8**
	Total Tflops	94,2	**1000,0**
Memoria	GB/core (GB)	2	**2**
	GB/node (GB)	8	**32**
	Total (TB)	20	**96,89**
Interconexión	Topologia	**Non- blocking Fat Tree**	
	Latencia (µs)	4	**0,7**
	Bandwidth (Gb/s)	4	**40**
Almacenamiento	(TB)	460	**2000**
Consumo	(KW)	750	**1080**

Software de MareNostrum2 vs MareNostrum3

	MN2	MN3
Sistema Opertaivo	SLES10 SP2	SLES11 SP1
Sistema de gestión	DIM	xCAT
Sistema de ficheros	GPFS	
Gestor de recursos/Scheduler	SLURM/MOAB	Loadleveler
Libreria MPI	MPICH-GM/MX	IBM PE
Sistema de monitorización	Ganglia/BSCtoolkit	xCAT
Compiladores	GCC, IBM XL	GCC, Intel Cluster Suite
Herramientas de rendimiento	BSC, PAPI	
Librerias	IBM ESSL	Intel MKL

Incluyendo: - Las licencias para IBM Intel HPC Stack
- Las UFM y OFED stacks soportadas por Mellanox.
- Licencia & soporte SLES + IBM Support line para Linux

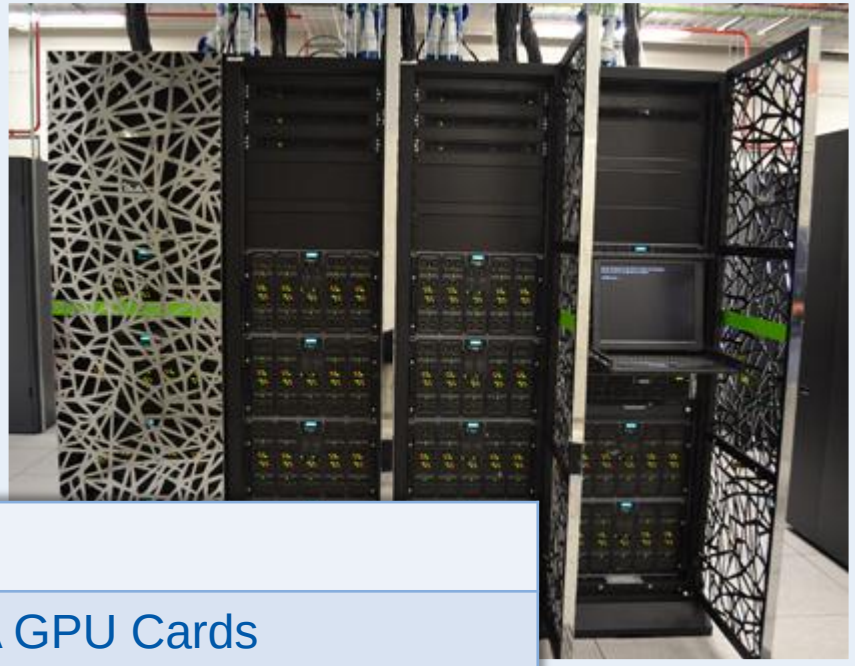
Recursos de MareNostrum2 vs MareNostrum3

	MN2	MN3
Total cores	10240	48448
Total de horas de cores en 1 año (x 365 x 24)	89.702.400	424.369.440
Usable Total core hours in a year (82%)	73.555.968	347.982.940
Horas para PRACE (70%)	0	243.588.058
Horas para BSC (20% x 30%)	14.711.193	20.878.976
Horas para RES (80% x 30%)	58.844.774	83.515.905
Horas para cada Call de la RES (:3)	19.614.924	27.838.635

1h Sandy Bridge (MN3) = 3h Power PC (MN2)

La Oferta a la RES aumenta en un factor 4,2

MinoTauro

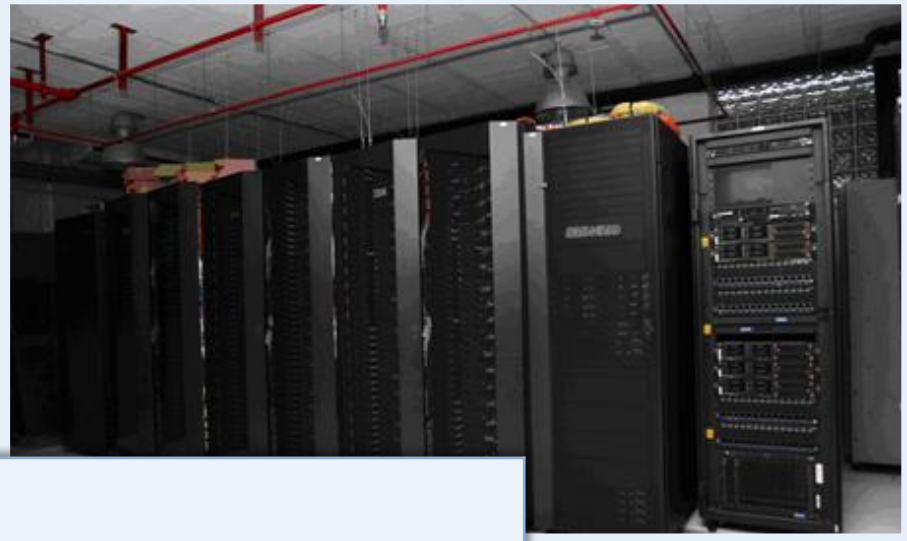


Rendimiento pico	185,78 TFLOPS
Procesador	256 M2090 NVIDIA GPU Cards 256 Intel E5649 2,53 GHz 6-Cores
Memoria	3 TB
Disco	31 TB
Redes	Infiniband QDR, 10 GbE
Sistema	ReddHat Linux

Magerit2

Rendimiento pico	103,4 TFLOPS
Procesador	3.920 IBM Power7 3.3
Memoria	8700 GB
Disco	190 TB
Redes	Infiniband, GbE
Sistema	Linux

Altamira



Rendimiento pico	52 TFLOPS
Procesador	316 Intel Xeon CPU E5-2670 2,60GHz
Memoria	10112 GB
Disco	14 TB
Redes	Infiniband
Sistema	Scientific Linux 6.2

UC UNIVERSIDAD DE CANTABRIA

IFCA Instituto de Física de Cantabria

LaPalma2

Rendimiento pico	9,4 TFLOPS
Procesador	1.024 IBM PowerPC 970 2.3GHz
Memoria	2 TB
Disco	14 + 10 TB
Redes	Myrinet, GbE, 10/100
Sistema	SUSE Linux

Tirant2

Rendimiento pico	18,8 TFLOPS
Procesador	2048 IBM PowerPC 970 2.3GHz
Memoria	2 TB
Disco	56 + 40 TB
Redes	Myrinet, GbE, 10/100
Sistema	SUSE Linux

Atlante

Gobierno de Canarias

itc

Rendimiento pico	3,1 TFLOPS
Procesador	336 IBM PowerPC 970 2.3GHz
Memoria	672 GB
Disco	3 + 90 TB
Redes	Myrinet, GbE, 10/100
Sistema	SUSE Linux

Picasso

Rendimiento pico	63 TFLOPS
Procesador	82 AMD Opteron 6176, 96 Intel E5-2670 56 Intel E7-4870 32 GPUS Nvidia Tesla M2075
Memoria	21 TB
Disco	600 TB Lustre + 260 TB
Redes	Infiniband, GbE
Sistema	SUSE Linux

Memento

Rendimiento pico	25,8 TFLOPS
Procesador	3072 AMD Opteron 6272 a 2.1GHz
Memoria	12,5 TB
Disco	36 TB Lustre
Redes	Infiniband, GbE
Sistema	Scientific Linux

Solicitud de acceso a RES

Información de convocatorias, condiciones y solicitud en:
www.bsc.es/RES

Información a aportar en la petición de acceso:

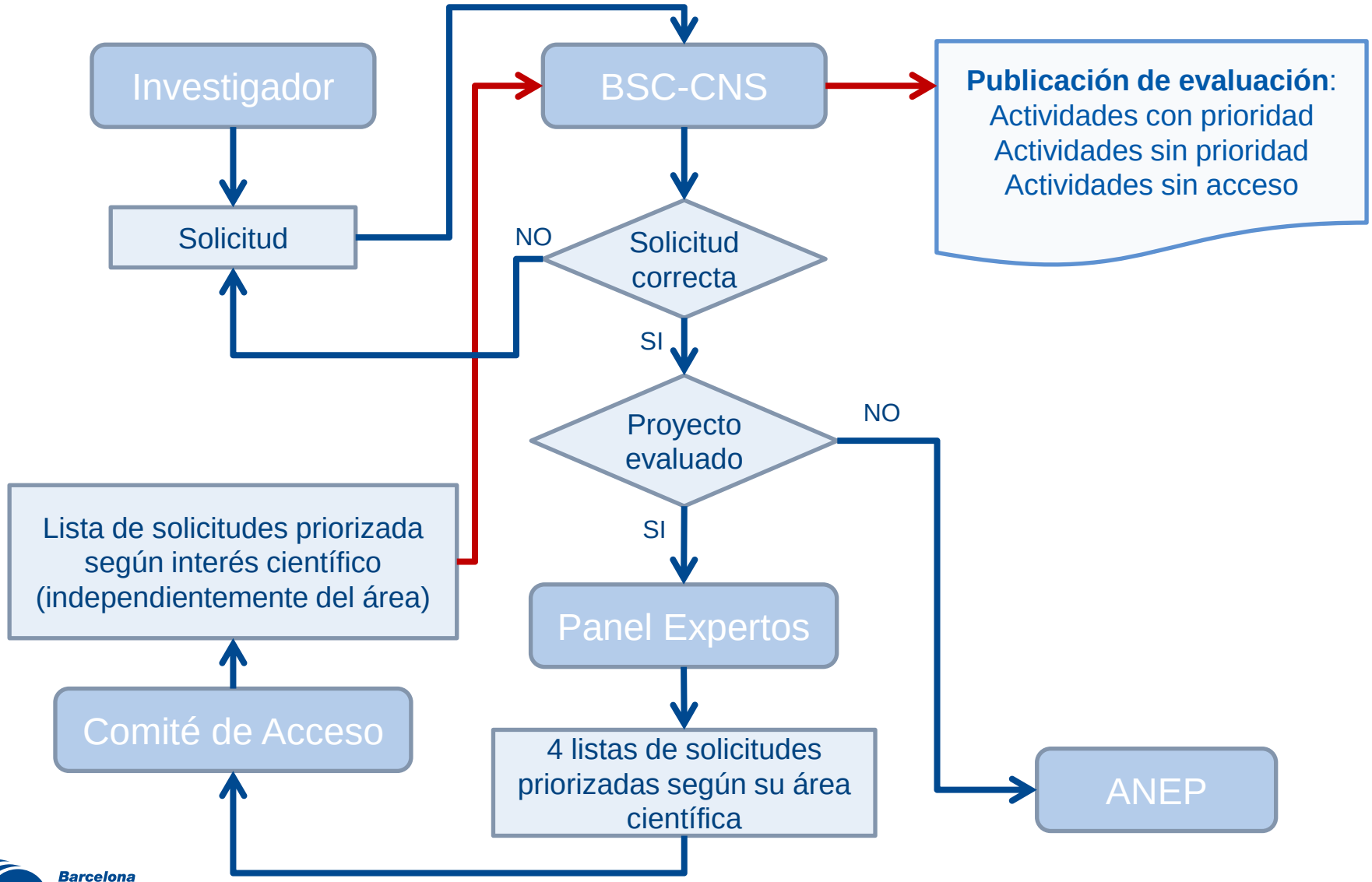
- Título de la actividad a desarrollar
- Descripción del proyecto
- Librerías numéricas y software necesario
- Descripción del equipo de investigación
- Cantidad de recursos necesarios
- Resumen del objetivo de la actividad para su publicación

Comité de Acceso y Paneles de Expertos

- Paneles de expertos
  - 10 científicos y un coordinador en cada uno de los 4 paneles :
    - Matemáticas, Física e Ingeniera
    - Química y Ciencia y Tecnología de los Materiales
    - Biomedicina y Ciencias de la Tierra
    - Astronomía, Espacio y Ciencias de la Tierra
- Comité de Acceso
  - 1 experto en gestión de innovación
  - 1 representante de la ANEP
  - 1 experto en supercomputación de la RES pero externo al BSC-CNS
  - 1 experto en supercomputación del BSC-CNS

Los componentes de los Paneles de Expertos y del Comité de Acceso son nombrados por el MINECO.

Proceso de evaluación de acceso a la RES



CURES, Comité de Usuarios de la RES

Historia

El CURES se estableció en febrero de 2010

Composición

- 8 científicos (2 de cada área de la RES) que hayan sido Investigadores Principales de actividades con acceso a la RES.

Propósito

- Transmitir la opinión e interés de los usuarios y asesorar al coordinador de la RES sobre los servicios y recursos disponibles.
- Promover el uso efectivo de los recursos de la RES compartiendo información sobre la experiencia de los usuarios y sugiriendo futuras líneas de investigación.

Difusión científica

- Website de BSC-CNS
- Informe anual del BSC-CNS y la RES
- Difusión de artículos en las revistas más prestigiosas
- Presentación de investigaciones en congresos nacionales e internacionales : Ibergrid, Supercomputing, International Supercomputing

PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

PRACE: Infraestructura para la investigación

- PRACE AISBL
  - Inauguración en Barcelona en 2010, sede en Bruselas
  - 24 miembros representantes de 20 países europeos
  - Miembros anfitriones: Francia, Alemania, Italia y España

- Financiación garantizada entre 2010 y 2015
  - 400 M€ de los miembros anfitriones
  - 70 M€ de EC FP7 para su desarrollo e implementación (Concesión INFSO-RI-211528 y 261557)

  Complemento de ~60 M€ por los miembros de PRACE

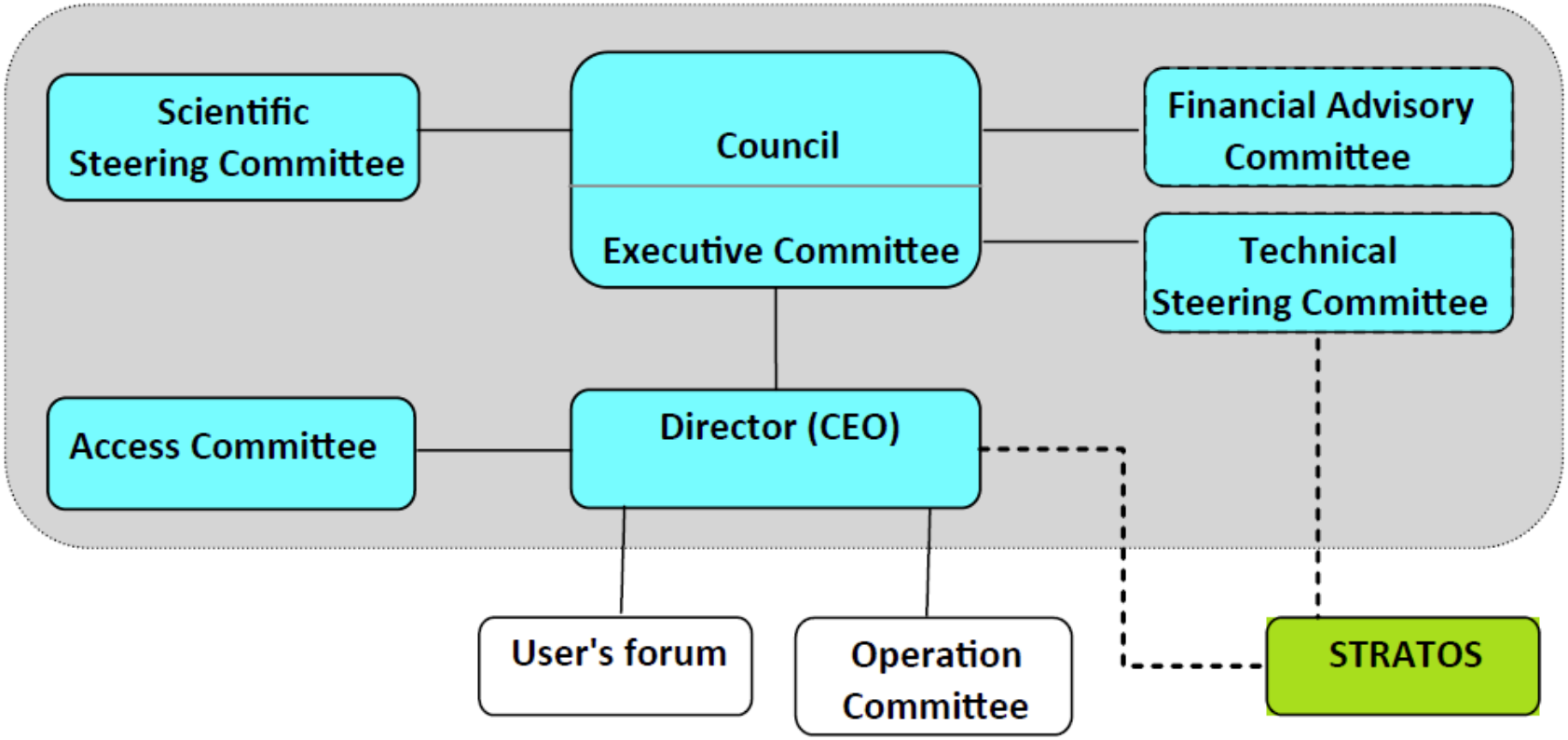
PRACE: Objetivos

- Desarrollo y provisión de una infraestructura a nivel europeo, que permita a la comunidad científica, incluída la industrial, acceso a sistemas High-end Computing o Tier-0.

- Gestión de la coordinación entre Tier-0, centros nacionales (Tier-1) y centros regionales (Tier-2) y facilitar así el contacto con sus usuarios.

- Provisión y racionalización del acceso a infraestructuras, a través de la evaluación y concesión a expertos científicos.

PRACE: Organización



PRACE: Scientific Steering Committee

- Consta de un máximo de 21 miembros.
- Opina en todas las materias científicas o técnicas.
- Miembros seleccionados por el Consejo en base a una lista de candidatos preparados por el SSC.
- Vinculación durante 2 años (con un máximo de 2 renovaciones)
- Propone los miembros del Comité de Acceso.

Richard Kenway (UK, particle physics), Chair
Jose M. Baldasano (Spain, environment)
Kurt Binder (Germany, statistical physics)
Paolo Carloni (Italy, biological physics)
Giovanni Ciccotti (Italy, statistical physics)
Dann Frenkel (Netherlands, molecular simulations)
Sylvie Joussaume (France, environment)
Ben Moore (Switzerland, astrophysics)
Gernot Muenster (Germany, particle physics)
Risto Nieminen (Finland, materials)
Modesto Orozco(Spain, life sciences)
Maurizio Ottaviani (France, plasma physics)
Michelle Parrinello (Switzerland, chemistry)
Olivier Pironneau (France, mathematics)
Thierry Poinsot (France, engineering)
Simon Portegies Zwart (Netherlands, astrophysics)
Kenneth Ruud (Norway, chemistry)
Wolfgang Schroeder (Germany, engineering)
Luis Silva (Portugal, plasma physics)
Alfonso Valencia (Spain, bioinformatics)

CURIE, Bull Bullx cluster (GENCI TGCC/CEA, Francia)

- Compuesto por 3 particiones:
- Fat node: 360 nodos con 32 cores por nodo para un rendimiento pico de 105 Tflops
- Thin node: 5040 blades con 16 cores por nodo, rendimiento pico de 1.5 PetaFlops
- Nodos híbridos: 144 blades con 8 cores escalares y 2 GPU por nodo, con un rendimiento pico de 200 TeraFlops

Más información en: http://www-hpc.cea.fr/en/complexe/tgcccurie.htm

FERMI, IBM BG/Q (CINECA, Italia)

- Arquitectura: 10 BGQ Frame con 2 MidPlanes
- Rendimiento pico: 2.1 PFlop/s
- Front-end Nodes OS: Red-Hat EL 6.2
- Compute Node Kernel: lightweight Linux-like kernel
- Procesador: IBM PowerA2, 1.6 GHz
- Nodos de computo: 10.240 con 16 cores, un total de 163.840 cores
- RAM: 16GB / node; 1GB/core
- Red interna: Interfaz de red con 11 enlaces en 5D Torus
- Espacio de disco: más de 2PB de scratch

Más información en: http://www.hpc.cineca.it/content/ibm-fermi-user-guide

HERMIT, Cray XE6 (GCS HLRS, Alemania)

- Rendimiento pico de 1 Petaflops
- Está diseñado para mantener un rendimiento sostenido de las aplicaciones y una alta escalabilidad.
- Compusto por 3552 nodos dual socket equipado con procesadores AMD Interlagos con un total de 113664 cores.
- Memoria principal en nodos con 32GB o 64GB

Más información en: http://www.hlrs.de/systems/platforms/crayxe6-hermit

JUQUEEN, IBM Blue Gene/Q (GCS Jülich, Alemania)

- En su etapa final tendrá un rendimiento pico de 5.87 Petaflops.
- Consiste en 28 racks, cad rack con 1024 nodes (16394 cores)
- Memoria principal de 458 TB

Más información en: http://www.fz-juelich.de/ias/jsc/juqueen

SUPERMUC, IBM System x iDataPlex (GCS LRZ, Alemania)

- Rendimiento pico de 3 Petaflops
- 155656 cores en 9400 nodos de cómputo
- >300 TB RAM
- Infiniband FDR10 interconnect
- 4 PB of NAS-based almacenamiento de disco permanente
- 10 PB of GPFS-based almacenamiento de disco temporal
- >30 PB almacenamiento en cinta
- Potentes sistemas de visualización
- Alta eficiencia energética

Más información en: http://www.lrz.de/services/compute/supermuc/systemdescription

PRACE: Tipos de convocatorias

- **Acceso Preparatorio**
  - Destinado a un uso previo para la preparación de solicitud de acceso regular como proyecto
  - Necesidad de revisión técnica
- **Acceso como proyecto**
  - Destinado a investigadores individuales o grupos de investigación
  - Necesidad de revisión técnica y científica
- **Acceso Multianual**
  - Disponible a los proyectos o infraestructura que puedan beneficiarse de los recursos de PRACE
  - Se planea que tenga una duración de 2 años
  - Está en fase de test

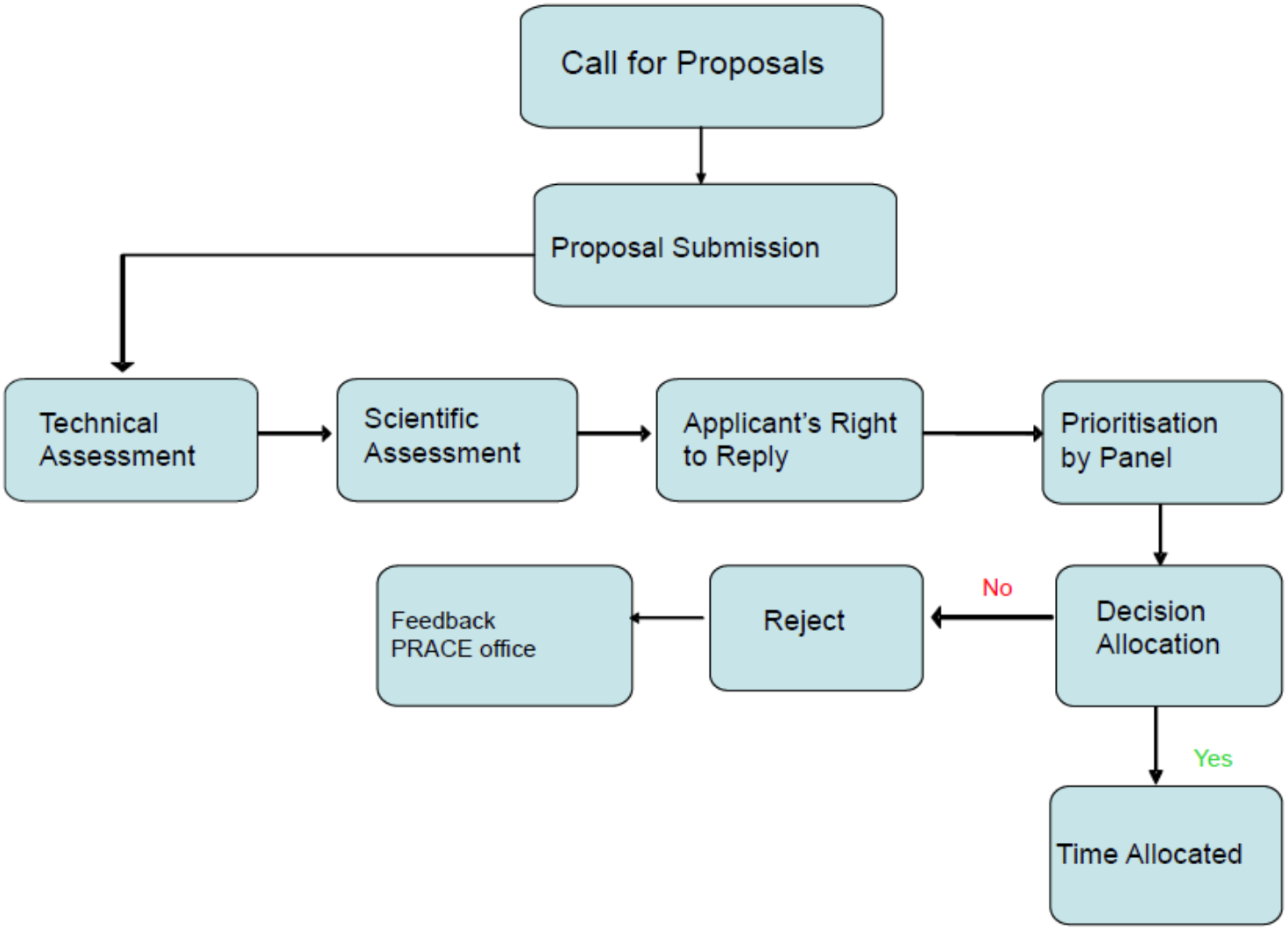
PRACE: Convocatorias vigente

- **Acceso regular**
  - Apertura de la 8ª PRACE Regular Call prevista en Septiembre de 2013.

- **Acceso preparatorio**
  - Continuamente abierto a solicitudes

Solicitud en: https://prace-peer-review.cines.fr

PRACE: Proceso de revisión por pares



PRACE: Criterio para la revisión técnica

- Necesidad de uso del recurso PRACE
- Disponibilidad del software requerido en el recurso solicitado
  - En caso de que el código sea propio del solicitante debe haber sido testeado en cuanto a eficiencia, escalabilidad y adecuación a la infraestructura.
  - Para solicitudes de Acceso como Proyecto deberán enviarse pruebas de los test.
- Adecuación al recurso solicitado
  - La evaluación técnica debe asociar las solicitudes al recursos más adecuado.

PRACE: Criterio para la revisión científica

- Excelencia científica demostrando impacto internacional
- Novedad y cualidades transformadoras
- Relevancia en la convocatoria si se diese un enfoque determinado a la convocatoria (en el caso de realizar convocatorias temáticas)
- Metodologia
- Relevancia de la diseminación en distintos canales y publicaciones
- Sólida estructura de gestión en el proyecto.

PRACE: Asignación de recursos

- El Comité de Acceso hace recomendaciones al PRACE Board of Directors en cuanto a la asignación de recursos.

El Comité de Acceso analiza:

- Informes técnicos y científicos
- Replica del solicitante
- Recursos solicitados

Y produce:

- Una única lista ordenada para cada convocatorial

PRACE: Soporte a través del BSC-CNS

- **Preparación técnica de solicitudes para potenciar su éxito**
- **Pruebas de escalabilidad en instalaciones del BSC-CNS**
- **Ayuda en el acceso y ejecución**
  - Portabilidad de código
  - Transferencia de entrada/salida
  - Acceso a sistema de colas
  - Rendimiento de aplicaciones
- **Ayuda en la transferencia de datos durante y después del acceso**

www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Muchas gracias

Para obtener más información

applications@bsc.es