

PLIEGO DE PRESCRIPCIONES TÉCNICAS QUE RIGEN EL CONTRATO DE SUMINISTRO, INSTALACIÓN Y PUESTA EN MARCHA DE EQUIPAMIENTO INFORMÁTICO PARA EL CENTRO DE PROCESAMIENTO DE DATOS (CPD) DE LA PLATAFORMA DE SECUENCIACIÓN GENÓMICA DENTRO DEL PARC CIENTÍFIC DE BARCELONA (Expediente Núm. 09/09)

Introducción

Este documento describe las características técnicas del equipamiento informático para la plataforma de secuenciación, centro que albergaría inicialmente 5 secuenciadores y con la posibilidad de ampliación a 10.

Cada uno de los secuenciadores son unos equipamientos que son capaces de producir por cada experimento un conjunto de 1 TB de datos cada 3 días. Dicha información en bruto deberá ser almacenada y procesada posteriormente por diversos procesos informáticos para la realización de las investigaciones pertinentes.

Para tal efecto se describe el equipamiento informático para dar soporte a la plataforma secuenciación y su trabajo:

- **Espacio de disco de 1 PB neto**, el cual podrá ser accedido de forma concurrente y óptima por el resto de equipamiento informático. En dicho espacio se almacenarán desde los datos crudos generados por los secuenciadores como cualquier dato producido por la investigación del mismo.
- **Máquina de memoria compartida** para el pre y post proceso de los datos de los secuenciadores. **Este servidor queda fuera del ámbito de este documento pero se mencionan para tenerlo en cuenta en los cálculos.**
- **Cluster de memoria distribuida** para la ejecución de códigos paralelos o masivamente paralelos necesarios para la explotación e investigación de los datos generados por los secuenciadores
- **Servidores web u otros servicios** de publicación de resultados obtenidos con los cálculos realizados. Estos servidores quedan fuera del ámbito de este documento pero se mencionan para tenerlos en cuenta en los cálculos.

Se distinguirán a nivel conceptual 3 redes que interconectarán los diversos elementos informáticos:

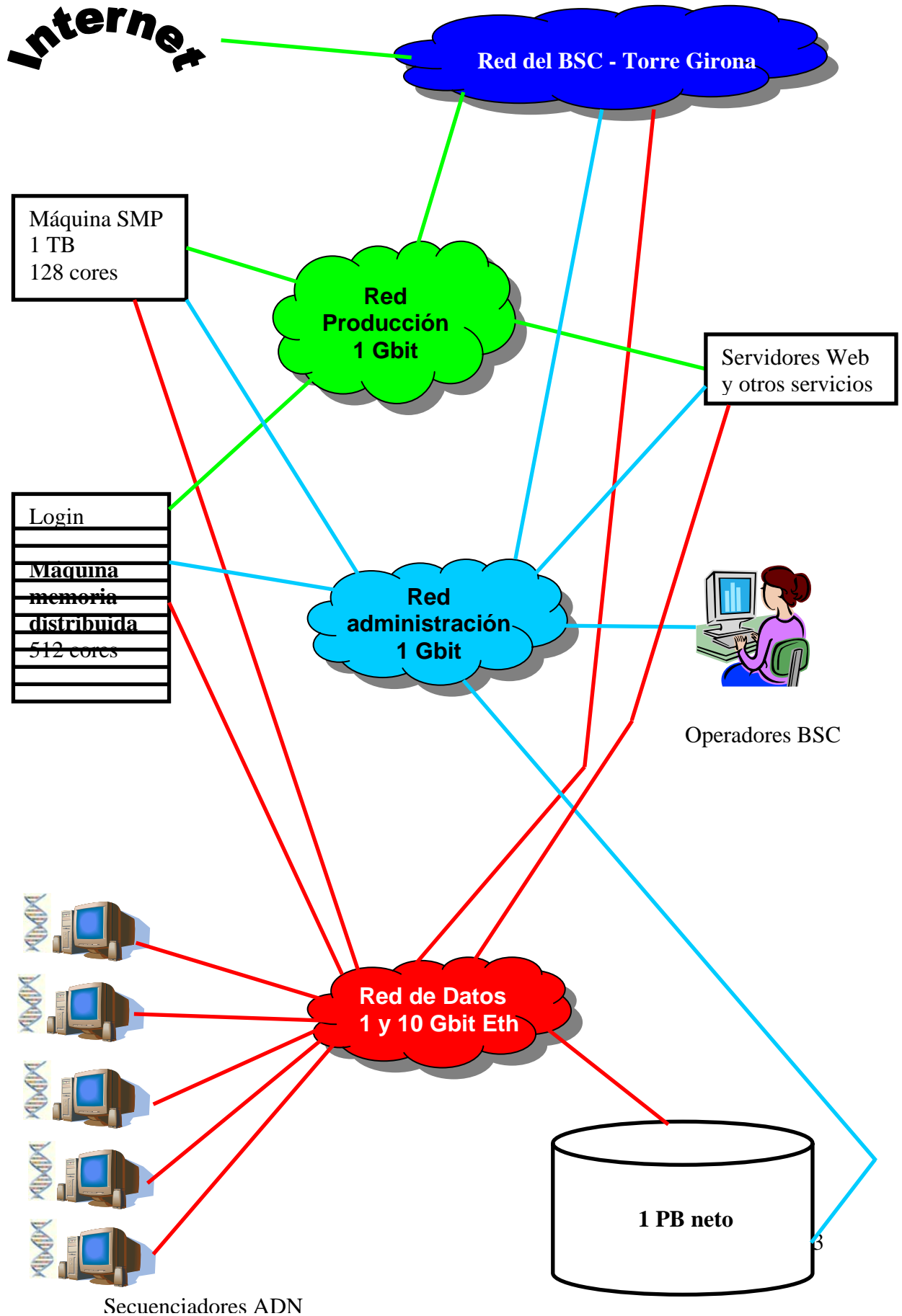
- **Red de datos:** Es aquella red la cual es usada de forma exclusiva para proporcionar el acceso al disco compartido de 1 TB con un alto rendimiento y ancho de banda al resto desde los equipamientos informáticos. La tecnología elegida para esta red es principalmente 1 y 10 Gbit Ethernet.
- **Red de Producción:** Es aquella red por la cual los investigadores realizarán las conexiones para acceder a los equipamientos informáticos y por donde los servidores que ofrezcan servicios hacia Internet publicarán dichos servicios. Es la única de las 3 redes que dispondrá conexión hacia o desde el exterior.

- **Red de Administración:** Red accesible sólo para los administradores informáticos de la infraestructura, dicha red sólo será accesible desde las redes internas del BSC o a través del servicio del gateway del BSC. Esta red servirá para la administración out-of-band de todos los equipamientos informáticos aquí descritos, los usuarios e investigadores no tendrán acceso a esta red.

Toda conectividad hacia o desde el exterior de esta infraestructura a nivel de red se realizará mediante los equipamientos que dispone el BSC en el edificio de capilla del Campus Nord de la UPC. De tal manera, en el apartado de equipamiento de red se incluye la ampliación de equipos necesarios a instalar en capilla del BSC para poder realizar este requisito.

Todo equipo se deberá entregar llaves en mano y completamente integrado en la infraestructura definida.

A continuación se muestra un esquema descriptivo de la infraestructura que se describe más concretamente punto por punto a posteriori:



Requisitos Técnicos de Máquinas Plataforma de Secuenciación

Por cada entrada con requisitos de las tablas siguientes se han clasificado como:

R – Representa que lo anunciado es un requerimiento que se debe cumplir en la solución presentada

D – Representa un requerimiento deseable a tener, se valorarán positivamente aquellas soluciones que lo incorporen

Cluster de memoria distribuida (Lote 1)

El cluster de memoria distribuida estará formado por tres tipos de máquinas:

- Un pequeño grupo de nodos que funcionarán como máquinas de login.
- Resto de nodos que serán los nodos de computación del cluster.
- Conjunto de servidores para la gestión del cluster.

Dentro del cluster se definirán 3 redes internas sólo visibles desde el propio cluster:

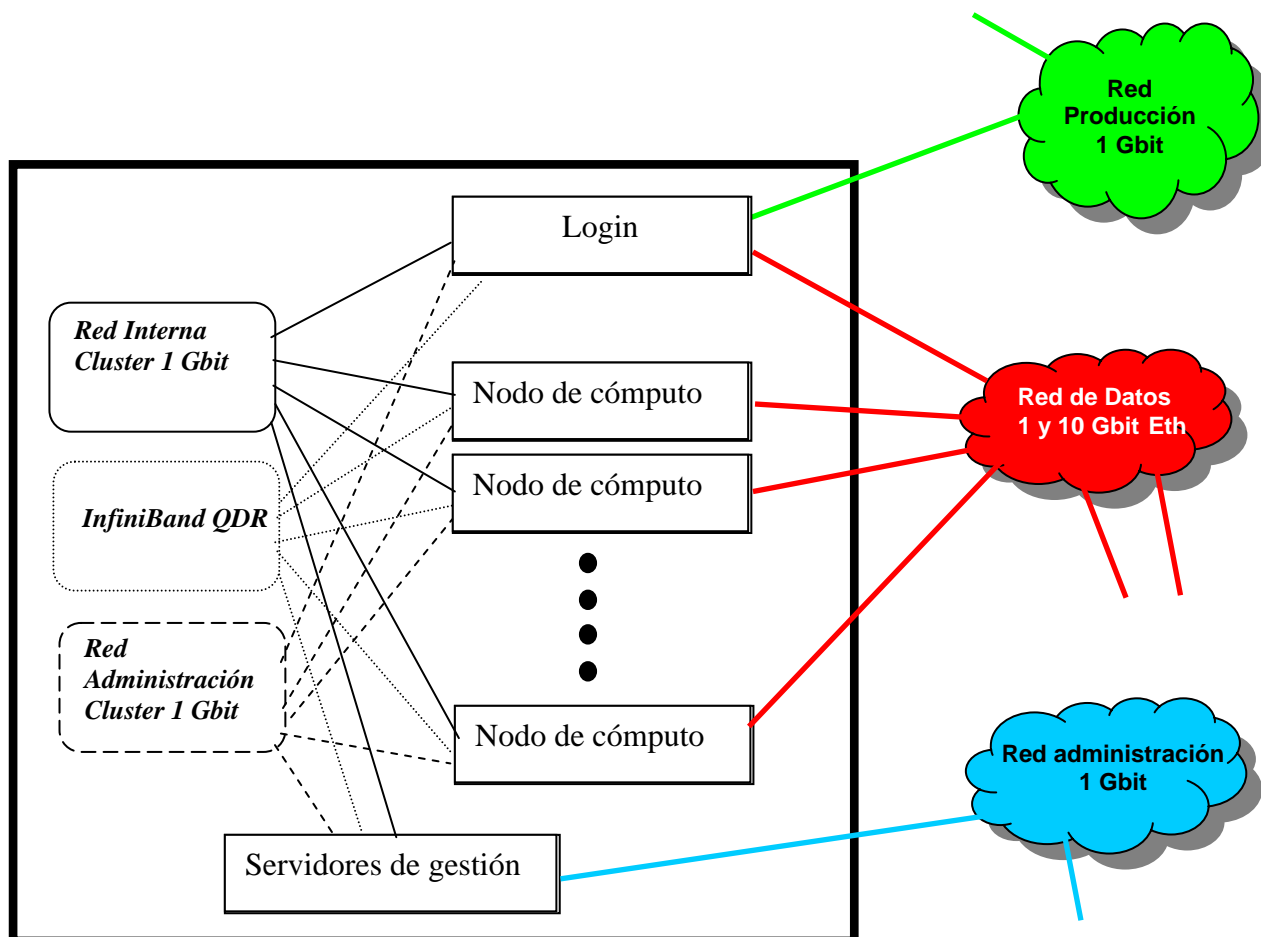
- Red interna producción del cluster (tecnología 1 Gbit).
- Red de alto rendimiento interna (tecnología InfiniBand QDR).
- Red de administración interna del cluster.

Por otra parte el cluster se deberá poder conectar a las redes externas definidas en el esquema inicial siguientes:

- Red de producción: La máquinas de login deberán estar conectadas a la red de producción del centro para que los usuarios pueden acceder al cluster.
- Red de datos: Todos los nodos del cluster deberán poder acceder a la red de datos para poder montar el disco centralizado.

A parte de los nodos de login y nodos de cómputo el cluster deberá incluir todos aquellos servidores para alojar los servicios de funcionamiento del propio cluster.

Aquí se dispone de un posible esquema con lo descrito:



Ref	Descripción
Hardware	
R	Cluster de máquinas de memoria distribuida con un mínimo de 512 cores, la máquina deberá tener un mínimo de 4 GB por core
D	Se valorará la mejora de memoria principal por core
R	Los procesadores deberán soportar operaciones en coma flotante de 64 bits
D	Se valorará la tecnología y la frecuencia de acceso a los módulos de memoria
R	Mínimo de 5.5 TeraFlops pico de rendimiento
D	Se valorará la potencia de cálculo extra ofertada en TeraFlops.
R	Cada nodo deberá disponer de un espacio local de almacenamiento de unas 5 veces la memoria principal por core (mínimo 10 GB por core) que inicialmente sólo será visible desde cada nodo y se usará para datos temporales.

Ref	Descripción
R	El cluster deberá venir con una red de interconexión interna de InfiniBand QDR (40 Gbit/s, donde cada nodo del cluster deberá poder disfrutar de este ancho de banda.
R	El cluster deberá incorporar todo aquel hardware necesario para disponer de una red de 1 Gbit Ethernet que interconectará todos los nodos del cluster que servirá como red interna y privada del cluster.
R	Cada nodo del cluster deberá disponer de un mínimo de 2 interfaces Gbit Ethernet (1 para la red interna del cluster y otra de reserva)
R	El cluster deberá estar conectado a la red de datos de tal manera que disponga aproximadamente un ancho de banda de 20 a 25 MB/s por core. La tecnología de la red de datos es 10 Gbit Ethernet tipo de conexión multimodo con conector LC.
R	El cluster deberá incluir todos aquellos servidores, equipamiento o redes que permitan la administración del cluster de forma out-of-band, sin que interfiera con las redes o la producción del cluster. Entre otras cosas este hardware deberá hacerse responsable de la gestión de imágenes de sistema operativo, servicios básicos para el cluster como DHCP, NTP, DNS, ...; el sistema de colas, etc.
D	Se valorará que la arquitectura presentada pueda ser extendida en capacidad de cálculo fácilmente
Software	
R	El sistema operativo deberá ser UNIX like y compatible con el X/Open Standard POSIX 1003 (IS/IEC 9945). Se valorará que la selección de sistema operativo sea Linux.
R	El cluster deberá incorporar un software de clustering que permita: <ul style="list-style-type: none"> - Una única imagen de sistema operativo para los nodos de computación que pueda ser mantenida y que los cambios se distribuyan de forma automática o semi-automática a todo el cluster. - Arranque y parada de los nodos de computo - Obtención de información de los nodos del cluster: Consola, temperatura, configuración de dispositivo de boot, etc. - Añadido, eliminación o sustitución de nodos del cluster
D	Se valorará que la solución de clustering sea una solución 'Diskless nodes', donde el sistema operativo no reside en la máquina local de cada nodo del cluster.
R	Con el sistema operativo se debe incluir todo el entorno de programación para la arquitectura de la máquina, como mínimo deberá incluir C, C++, Java y Fortran. A parte del entorno de programación Open-Source proporcionado, se deberá de proporcionar el entorno de programación específico para la arquitectura proporcionada: <ul style="list-style-type: none"> - por ejemplo: Intel Compilers. IBM Compilers, etc. según el caso.
R	Se deberán de proporcionar las librerías numéricas proporcionadas por el fabricante de los procesadores debidamente optimizadas para dicha arquitectura.
R	También se deberá proporcionar los compiladores, librerías y/o las herramientas necesarias para el uso paralelo de la arquitectura mediante paradigmas estándares

Ref	Descripción
	como OpenMP ó MPI. Para OpenMP deberá soportar la versión 2.5 estándar. Por otro lado, se deberá soportar completamente el estándar MPI versión 1.2 para la red de interconexión del cluster en este caso IB QDR.
D	Se valorará el soporte de de MPI versión 2.1 (excluyendo el tratamiento dinámico de procesos)
R	Se deberá incluir un software de sistema de colas que permita el envío de trabajos batch a la máquina
D	Se valorará la inclusión de herramientas de profiling y debugging de aplicativos secuenciales y paralelos
Requerimientos operacionales	
R	La solución deberá incorporar monitorización de todo el sistema para detectar: <ul style="list-style-type: none"> - El buen estado del sistema - Rendimiento o carga de funcionamiento - Alarmas y avisos por email y otros métodos de cualquier eventualidad / problema
R	La máquina deberá disponer de un sistema de monitorización de la temperatura de tal manera que provoque la parada controlada del sistema en caso de temperatura muy alta.
R	El consumo máximo de la solución presentada deberá ser inferior a 30 Kw
D	Se valorará positivamente el mínimo consumo eléctrico. Indicar el máximo y el típico consumo en Kw. de la solución.
R	Realizar un cambio o añadir un fichero a todos los nodos de computo no debería tardar más de 1 minuto
D	Aislamiento de privilegios, se valorará la inclusión de una red privada de administración del cluster aislada de la red interna de producción del cluster
R	El sistema deberá ocupar 2 racks Standard de 42U
R	Los racks deberán de incorporar refrigeración líquida de puertas traseras que eliminen en un mínimo del 80% del calor generado. Un valor superior al 80% será valorado positivamente.
Mantenimiento y soporte	
R	Garantía y soporte de 3 años en todos los componentes (hardware y software), con una respuesta en 4 horas dentro de las horas de oficina (10:00 – 17:00) y con un servicio de soporte de Next Business Day.
D	Se valorará la extensión de la garantía / mantenimiento
R	El proyecto de instalación incluirá la comprobación del buen funcionamiento, integración y óptimo rendimiento de la solución con el resto de equipos informáticos aquí anunciados. En tal caso, se exigirá el trabajo en equipo con las empresas instaladoras del resto de equipos aquí descritos, para la solución de cualquier problema de incompatibilidad que surja.

Ref	Descripción
R	Se proporcionará (dentro de período de garantía): <ul style="list-style-type: none"> - Acceso a todos los software upgrade (incluyendo sistemas operativos y firmware) de todos los componentes de la solución - Punto único de soporte para el aviso de problemas e incidencias de cualquier componente que componga la solución
D	Se valorará el soporte pro-activo, notificando y recomendado subidas de versión tanto de software como de firmware.
R	Se deberá entregar al final de la instalación una documentación digital en la que se describa: <ul style="list-style-type: none"> - Descripción general de los componentes de la solución - Esquema de conexionado e IPs - Valores de configuración empleados - Explicación del proceso de instalación y tareas realizadas - Explicación procedimientos para: Puesta en marcha, y disaster recovery
R	Se deberán ofrecer formación durante la instalación de la solución, que cubran: <ul style="list-style-type: none"> - Conceptos básicos - Administración básica y procedimientos básicos de configuración - Optimización de la solución - Solución de problemas
D	Se valorará la inclusión de cursos oficiales sobre los componentes de la solución.

Disco centralizado (Lote 2)

Nota: GB equivale a Gigabytes. Gbit equivale a Gigabit. PB equivale a Petabyte. TB equivale a Terabyte.

Ref	Descripción
Hardware	
R	Sistema de almacenamiento con una capacidad neta de 1 PB de espacio. Se valorará si más espacio es ofertado.
R	Los discos deben estar protegidos por RAID hardware
R	El sistema debe incorporar todo aquel hardware como servidores, switches, SAN, etc. para que la solución permita el acceso concurrente y paralelo desde el resto de máquinas definidas en este documento, sin pérdida de rendimiento.
R	La solución deberá ser escalable de tal manera que pueda ser capaz de ser incrementada fácilmente en un factor de 2x tanto en capacidad como en velocidad de acceso.
R	Posibilidad de crear filesystems con tamaño de PBs o cientos de TBs.
R	Cualquier componente de la solución deberá estar redundando de tal manera que no exista un punto único de fallo. Se deberá poder apagar cualquier componente de la solución y el sistema seguir en producción.
R	Rendimiento mínimo de 5 GB/s tanto en lecturas como en escrituras a la solución de forma agregada. La tecnología de conexión hacia la red de datos para proporcionar este rendimiento deberá ser 10 Gbit Ethernet con un tipo de conexión multimodo con conector LC.
D	Se valorará la mejora en el ancho de banda agregado de la solución ofertada.
Software	
R	La solución debería incorporar un filesystem paralelo que permita su montaje de forma concurrente en el resto de máquinas de este documento, como futuras máquinas: <ul style="list-style-type: none"> - Máquina memoria compartida - Cluster de memoria distribuida - Ordenadores de los secuenciadores
R	El cliente del filesystem paralelo deberá estar soportado en las principales distribuciones Linux sobre las principales arquitecturas: x86, x86_64, ia64, ppc64.
R	El filesystem paralelo ofertado debería cumplir: <ul style="list-style-type: none"> - Escritura de datos de forma inteligente para el balanceo de carga entre los discos. - Acceso a los datos mediante POSIX desde los clientes - Debería ser capaz de soportar el montaje desde miles de clientes - Lock de ficheros a nivel de bloque y gestión distribuida - Posibilidad de exportar el filesystem paralelo mediante protocolos como NFS o CIFS

Ref	Descripción
D	Se valorará si el filesystem permite definir políticas de ILM para el tratamiento de datos como por ejemplo para optimizar el borrado automático de ficheros antiguos, etc.
R	El filesystem deberá soportar un número de decenas de miles de ficheros por directorio sin degradación de rendimiento
R	La red de conexión entre los clientes y los servers del sistema de ficheros paralelo será de 10 Gbit Ethernet, el software de la solución deberá soportar y ser capaz de aprovechar dicha red de datos.
R	La solución deberá permitir configurar diversos espacios lógicos o filesystems con características de acceso diferentes.
Requerimientos operacionales	
R	La solución deberá incorporar monitorización de todo el sistema para detectar: <ul style="list-style-type: none"> - El buen estado del sistema - Rendimiento o carga de funcionamiento - Alarmas y avisos por email y otros métodos de cualquier eventualidad / problema
R	La solución debe asegurar la consistencia de los datos cuando son accedidos o guardados en el filesystem paralelo y global.
D	Se valorará la integración del filesystem paralelo con el software de backup: Tivoli Storage Manager usado en el BSC.
R	El consumo máximo de la solución presentada deberá ser inferior a 15 Kw
D	Se valorará positivamente el mínimo consumo eléctrico. Indicar el máximo y el típico consumo en Kw. de la solución.
R	Deberá ocupar menos de 2 racks estándar de 42 U.
R	Los racks deberán de incorporar refrigeración líquida de puertas traseras que eliminen en un mínimo del 80% del calor generado. Un valor superior al 80% será valorado positivamente.
Mantenimiento y soporte	
R	Garantía y mantenimiento de 3 años en todos los componentes (hardware y software), con una respuesta en 4 horas dentro de las horas de oficina (10:00 – 17:00) y con un servicio de soporte de Next Business Day
D	Se valorará la extensión de la garantía/mantenimiento
R	El proyecto de instalación incluirá la comprobación del buen funcionamiento, integración y óptimo rendimiento de la solución con el resto de equipos informáticos aquí anunciados. En tal caso, se exigirá el trabajo en equipo con las empresas instaladoras del resto de equipos aquí descritos, para la solución de cualquier problema de incompatibilidad que surja.
R	Se proporcionará (dentro de período de garantía): <ul style="list-style-type: none"> - Acceso a todos los software upgrade (incluyendo sistemas operativos y

Ref	Descripción
	firmware) de todos los componentes de la solución - Punto único de soporte para el aviso de problemas e incidencias de cualquier componente que componga la solución
D	Se valorará el soporte pro-activo, notificando y recomendado subidas de versión tanto de software como de firmware.
R	Se deberá entregar al final de la instalación una documentación digital en la que se describa: <ul style="list-style-type: none"> - Descripción general de los componentes de la solución - Esquema de conexionado e IPs - Valores de configuración empleados - Explicación del proceso de instalación y tareas realizadas - Explicación procedimientos para: Puesta en marcha, y disaster recovery
R	Se deberán ofrecer formación durante la instalación de la solución, que cubran: <ul style="list-style-type: none"> - Conceptos básicos - Administración básica y procedimientos básicos de configuración - Optimización de la solución - Solución de problemas
D	Se valorará la inclusión de cursos oficiales sobre los componentes de la solución.

Equipamiento de redes (Lote 3)

Ref	Descripción
Hardware	
Red de datos	
R	Switch modular con fuente redundante, supervisora redundante y sistema de ventilación redundada
R	Soporte de inserción / retirada de tarjetas “en caliente” y sin interrupción del servicio sobre el resto del hardware
D	Se valorará la posibilidad de crecimiento en slots sin necesidad de comprar otro chasis
R	Mínimo de 20 puertos 10 Gbit Ethernet multimodo con GBIC asociado
D	Se valorará la inclusión de más puertos 10 Gbit Ethernet multimodo con GBIC asociado
R	4 puertos 10 Gbit Ethernet monomodo con GBIC asociado para su conexión hacia el BSC, para poder cubrir una distancia de 5 Km.
R	48 puertos Ethernet UTP 10 / 100 / 1000
R	Mínimo 20 fibras de 15 metros multimodo con un extremo LC y el otro extremo en función del GBIC ofertado
R	4 fibras monomodo de 5 metros con un extremo en función del GBIC ofertado y el otro extremo conector SC cuadrado para conectar a patch panel
R	Requerimientos de funcionalidades: <ul style="list-style-type: none"> - Filtrado y políticas de nivel 2 - Listas de control de acceso - SSHv2 - Soporte VLAN Privadas - Soporte 802.1Q - Gestión completa a través de SNMPv1, 2 y 3 - Protección de bucles de Spanning-tree - Soporte al menos de 256 VLANs simultáneas - Soporte de redundancia de enlaces (Duplicación de enlaces hacia los distintos elementos de red) protocolos a usar (LACP o EtherChannel) - Capacidad de filtrado de BPDUs de STP
D	Se valorará la inclusión de 2 fibras más de cada tipo
D	Se valorará la inclusión de GBICs en stock extras a los requeridos en el concurso
D	Se valorará el soporte de IPv6 y la capacidad de realizar tareas de routing
Red de producción y administración	
R	Switch modular con fuente redundante y supervisora redundante
R	96 puertos 10 / 100 / 1000 Ethernet UTP

Ref	Descripción
R	4 puertos fibra 1 Gbit Ethernet monomodo con GBIC asociado
R	4 fibras monomodo de 5 metros con conector según el GBIC ofertado por un lado y el otro extremo conector SC cuadrado para conectar a patch panel
R	Requerimientos de funcionalidades: <ul style="list-style-type: none"> - Filtrado y políticas de nivel 2 - Listas de control de acceso - SSHv2 - Soporte VLAN Privadas - Soporte 802.1Q - Gestión completa a través de SNMPv1, 2 y 3 - Protección de bucles de Spanning-tree - Soporte al menos de 256 VLANs simultáneas - Soporte de redundancia de enlaces (Duplicación de enlaces hacia los distintos elementos de red) protocolos a usar (LACP o EtherChannel) - Capacidad de filtrado de BPDUs de STP
D	Se valorará la inclusión de 2 fibras más de cada tipo
D	Se valorará la inclusión de GBICs en stock extras a los requeridos en el concurso
D	Se valorará el soporte de IPv6 y la capacidad de realizar tareas de routing
Equipamiento de red para conexión con el BSC	
R	Switch modular con fuente redundante, supervisora redundante y sistema de ventilación redundada
R	Soporte de inserción / retirada de tarjetas “en caliente” y sin interrupción del servicio sobre el resto del hardware
D	Se valorará la posibilidad de crecimiento en slots sin necesidad de comprar otro chasis
R	Módulo para conectar al menos 4 puertos de 10Gbit
R	2 GBIC 10 Gbit Ethernet de fibra monomodo, para cubrir una distancia de 5 Km.
R	2 GBIC 10 Gbit Ethernet de fibra multimodo
R	2 fibras monomodo de 3 metros con conector según el GBIC ofertado por un lado y el otro extremo conector SC cuadrado para conectar a patch panel
R	2 fibras multimodo de 2 metros con conector según el GBIC ofertado por un lado y el otro extremo conector SC cuadrado para conectar a patch panel
R	Requerimientos de funcionalidades: <ul style="list-style-type: none"> - Filtrado y políticas de nivel 2 - Listas de control de acceso - SSHv2 - Soporte VLAN Privadas - Soporte 802.1Q - Gestión completa a través de SNMPv1, 2 y 3 - Protección de bucles de Spanning-tree - Soporte al menos de 256 VLANs simultáneas - Soporte de redundancia de enlaces (Duplicación de enlaces hacia los distintos

Ref	Descripción
	<p>elementos de red) protocolos a usar (LACP o EtherChannel)</p> <ul style="list-style-type: none"> - Capacidad de filtrado de BPDUs de STP
R	<p>Ampliación para el equipo de red del BSC Cisco 6500:</p> <ul style="list-style-type: none"> - Tarjeta de 24 puertos SFP, modelo: CEF720 24 port 1000mb SFP WS-X6724-SFP - 2 GBIC Cisco 10Gbit Ethernet XENPAK -SR (fibra multimodo) - 4 GBIC Cisco SFP GLC-LH 1 Gbit Ethernet de fibra monomodo
R	<p>2 fibras multimodo de 3 metros con conector según el GBIC ofertado por un lado y el otro extremo conector SC cuadrado para conectar a patch panel</p>
R	<p>4 fibras monomodo de 3 metros con conector según el GBIC ofertado por un lado y el otro extremo conector SC cuadrado para conectar a patch panel</p>
D	<p>Se valorará la inclusión de GBICs en stock extras a los requeridos en el concurso</p>
D	<p>Se valorará la inclusión de 2 fibras más de cada tipo</p>
Mantenimiento y soporte	
R	<p>Garantía y soporte de 3 años en todos los componentes (hardware y software), con una respuesta en 4 horas dentro de las horas de oficina (10:00 – 17:00) y con un servicio de soporte de Next Business Day.</p>
D	<p>Se valorará la extensión de la garantía / mantenimiento</p>
R	<p>El proyecto de instalación incluirá la comprobación del buen funcionamiento, integración y óptimo rendimiento de la solución con el resto de equipos informáticos aquí anunciados.</p> <p>En tal caso, se exigirá el trabajo en equipo con las empresas instaladoras del resto de equipos aquí descritos, para la solución de cualquier problema de incompatibilidad que surja.</p>
R	<p>Se proporcionará (dentro de período de garantía):</p> <ul style="list-style-type: none"> - Acceso a todos los software upgrade (incluyendo sistemas operativos y firmware) de todos los componentes de la solución - Punto único de soporte para el aviso de problemas e incidencias de cualquier componente que componga la solución
D	<p>Se valorará el soporte pro-activo, notificando y recomendado subidas de versión tanto de software como de firmware.</p>
R	<p>Se deberá entregar al final de la instalación una documentación digital en la que se describa:</p> <ul style="list-style-type: none"> - Descripción general de los componentes de la solución - Esquema de conexionado e IPs - Valores de configuración empleados - Explicación del proceso de instalación y tareas realizadas - Explicación procedimientos para: Puesta en marcha, y disaster recovery
R	<p>Se deberán ofrecer formación durante la instalación de la solución, que cubran:</p> <ul style="list-style-type: none"> - Conceptos básicos - Administración básica y procedimientos básicos de configuración - Optimización de la solución

Ref	Descripción
D	- Solución de problemas Se valorará la inclusión de cursos oficiales sobre los componentes de la solución.

Barcelona, a 21 de septiembre de 2009.

D. Dídac Ramírez i Sarrió
Presidente del Parc Científic de Barcelona