

**PLIEGO DE PRESCRIPCIONES TÉCNICAS PARA LA CONTRATACIÓN, POR PROCEDIMIENTO NEGOCIADO CON PUBLICIDAD, DEL SUMINISTRO DE AMPLIACIÓN Y ACTUALIZACIÓN DEL SISTEMA DE DISCOS Y FICHEROS ADQUIRIDO POR EL BSC-CNS EN EL AÑO 2010 MEDIANTE EL EXPEDIENTE DE CONTRATACIÓN NÚMERO CONSU02010003OP.**

El Barcelona SuperComputing Center – Centro Nacional de Supercomputación requiere la ampliación de su sistema de almacenamiento para HPC que está proporcionando servicio desde finales de 2010. Dada la evolución de la tecnología de almacenamiento, esta actualización de la solución del 2010, permite integrar servicios de Big Data en procesadores más próximos a los datos/discos, e integrar los mismos servicios a través de diferentes redes de interconexión (Infiniband/10GE Ethernet).

Estos nuevos sistemas de almacenamiento, orientados a entornos Exascale y Big Data, permiten demostrar la escalabilidad tanto a nivel de capacidad de almacenamiento como del rendimiento de los diferentes sub-modulos sobre los que se basa.

Se propone actualizar el sistema de ficheros actual del BSC, aprovechando el máximo los equipos existentes, con una tecnología nueva, orientada a HPC Exascale y BigData. Con los elementos hardware sobrantes de la actualización se propone instalarlos en el sistema de archivo activo para ampliar su capacidad de almacenamiento.

Se describe a continuación las partes componentes objeto del concurso.

### **Actualización sistema almacenamiento HPC del BSC a tecnologías Exascale**

#### **Características generales a cumplir:**

- Instalación de una capacidad de almacenamiento de mínimo 2.5 PB neto, basado en discos duros y declustered raid, con un rendimiento aproximado de 50 GB/s
- El sistema tendría de ser modular capaz de crecer en el futuro en rendimiento y capacidad hasta superar los 10PB y unos 200 GB/s en 5 racks
- Sistema de almacenamiento basado en flash o SSD con una capacidad de 50 TB netos, para el uso de metadatos de los filesystems y “Burst buffer” de los filesystems
- Sistema basado en GPFS compatible con el almacenamiento actual HPC del BSC
- Los nuevos componentes de datos y metadatos deberán poder conectarse al switch existente de 10 Gbit Ethernet, y mediante Infiniband a MareNostrum
- Los elementos de almacenamiento propuesto se deberán instalar en 3 racks standard 42”
- El proyecto deberá incorporar todas las tareas necesarias para la integración de esta actualización con los componentes actuales del almacenamiento para HPC del BSC y su funcionamiento óptimo
- Los 3 data building blocks actuales del almacenamiento se deberán mover e integrar en el archivo activo del BSC, ampliando su capacidad de 4.7PB a 6.7 PB

## EXPEDIENTE CONSU02015015NG

- Se deberá incluir cualquier ampliación o cambio de los elementos de red, switches, GBICs, cables, fibras necesarios para llevar a cabo todos los cambios anunciados. (1/10 Gbit Ethernet , Infiniband, etc.)
- A parte del almacenamiento se requiere la compra de un conjunto de servidores de análisis de datos basados en tecnología Power8, dichos servidores irán en un rack a parte disponible en dependencias del BSC.

## Descripción general de la actualización:

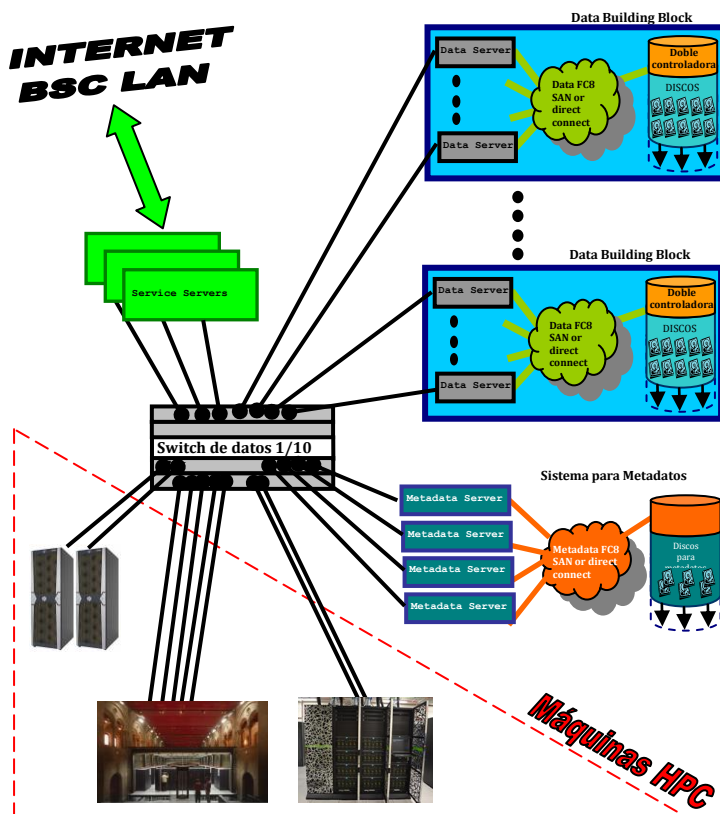
Se propone actualizar el sistema de ficheros actual del BSC, aprovechando el máximo los equipos existentes, con una tecnología nueva, orientada a HPC Exascale y BigData. Con los elementos hardware sobrantes de la actualización se propone la ampliación del sistema actual de archivado activo.

### Sistema de almacenamiento HPC BSC inicial

El sistema de almacenamiento adquirido en 2010, basado en tecnología GPFS de IBM se compone de los siguientes componentes hardware:

- 3 Data building blocks para el almacenamiento de datos basados en RAID hardware
- 1 Metadata building block para el almacenamiento de los metadatos de los sistemas de ficheros GPFS
- Un equipo de red con tecnología 1 Gbit y 10 Gbit Ethernet para la interconexión de dicho almacenamiento con todos los nodos HPC del BSC
- 8 servidores de servicio que albergan los servicios asociados a datos como son backup de los datos, transferencia de datos, exportaciones via NFS y otros protocolos.

A continuación mostramos una figura con los diversos componentes de la solución actual de almacenamiento HPC del BSC:



Este almacenamiento está instalado en la capilla donde reside MareNostrum ocupando 5 racks, ofreciendo la capacidad de 1.9 PB netos con un rendimiento de 15 GB/s.

*Propuesta de mejora almacenamiento*

Se propone en la mejora del almacenamiento la actualización de los data building blocks basados en tecnología de hardware RAIDs por data building blocks basados en tecnología declustered raids.

Estos nuevos sistemas de almacenamiento de datos basados en declustered RAIDs, deberán proporcionar un almacenamiento de 2.5 PB mínimo neto y un rendimiento aproximado de 50GB/s, para poder soportar las nuevas aplicaciones Exascale y BigData. Dicho almacenamiento debería ser modular y poder crecer tanto en rendimiento como en capacidad de almacenamiento.

Estos sistemas de almacenamiento se deberán conectar e integrar al sistema ya existente, concretamente se deberá poder integrar con:

- Equipo de red 1Gbit/10 Gbit Ethernet
- Sistema de metadatos
- Servidores de servicio

Además la actualización permitirá conectarse mediante la red Infiniband a MareNostrum para mejorar el rendimiento de acceso desde este superordenador.

A parte del sistema de almacenamiento de datos, se propone también la instalación e integración de un sistema de almacenamiento basado en Flash, SSD o tecnología similar, con un tamaño mínimo de 50 TB netos que se usará como el nuevo sistema de metadatos y “Burst Buffer” para las aplicaciones con alta demanda de entrada/salida.

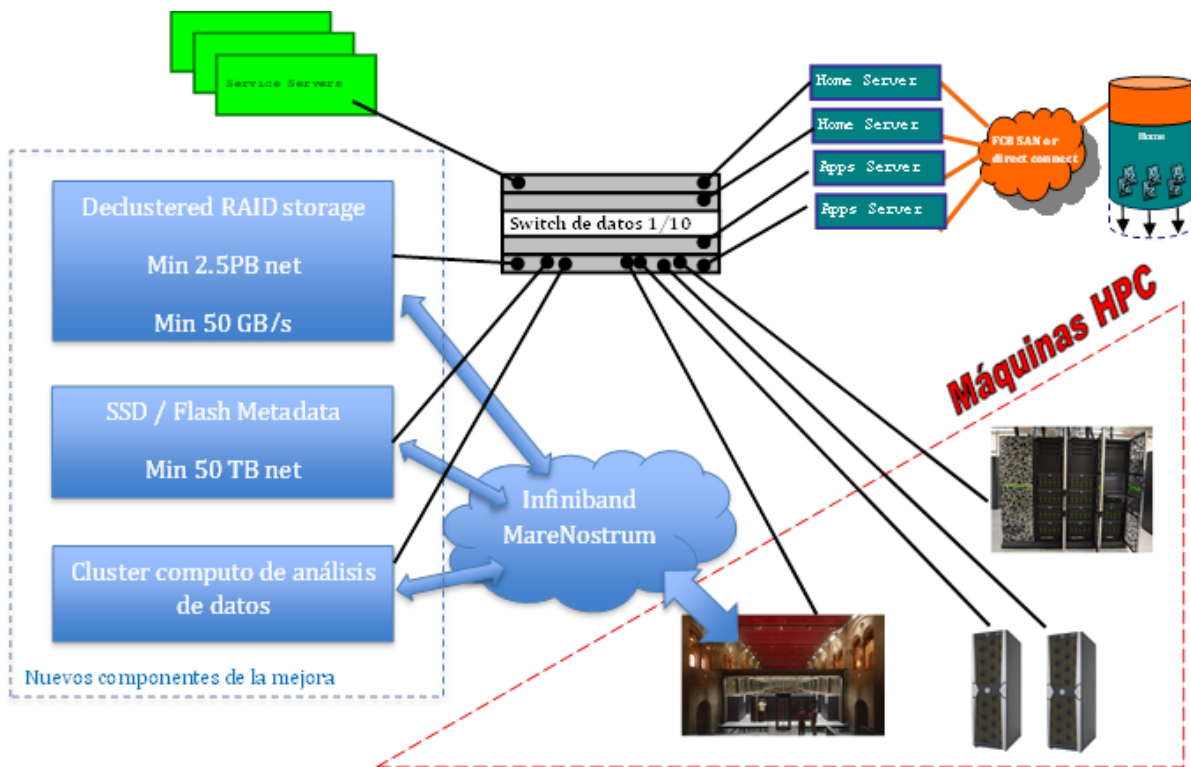
Durante el proyecto de actualización el sistema de metadatos inicial se reconfigurará de forma progresiva para que su rol cambie de “sistema de metadatos” a servir los filesystems de “home” y “apps” de todos los supercomputadores del BSC.

Durante esta integración se deberá realizar la migración de datos entre las dos tecnologías minimizando el tiempo de parada de servicio.

Una vez los datos estén migrados, los 3 “data building blocks” basados RAID tradicionales sobrantes se integraran en el almacenamiento del Archivo Activo adquirido por el BSC en 2013. Dicho almacenamiento permitirá subir de los 4.7 PB a 6.7 PB de capacidad para el archivado de datos científicos. Dentro del proyecto de la ampliación también estarán incluidas las tareas de dicha integración.

Aparte de estas mejoras en almacenamiento se adquirirán cierto número de nodos para añadir a los servidores de servicio para añadir servicio de cómputo basados en análisis de datos y BigData.

La siguiente figura muestra un esquema del sistema de almacenamiento después de la actualización propuesta:



**Especificación de las características generales y datos valorables:**

Se presenta a continuación un conjunto de requerimientos y condiciones valorables de la ampliación propuesta.

En las siguientes tablas se identifican los campos por las letras R y D, cuyo significado es:

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

D- Representa un requerimiento deseable a tener y que se valorará positivamente aquellas soluciones que lo incorporen.

**A. Descripción Hardware****A1. Almacenamiento de Datos**

Ref	Descripción
R1	Sustitución de los building blocks actuales por un nuevo sistema modular basado en declustered raids con una capacidad mínima de 2.5 PB neta y un rendimiento alrededor de los 50 GB/s por Infiniband y 8 GB/s por 10 Gbit Ethernet.
D2	Se valorará la mejora del almacenamiento proporcionado y la mejora en rendimiento proporcionado
R3	La solución propuesta deberá ser basada en módulos, los cuales todos ellos deberán ser idénticos y que sean capaces de escalar por encima de 10PB y por encima de 200 GB/s en 5 racks
R4	Todos los elementos de la actualización deberán ser redundantes, y no deberán tener un único punto de fallo. Concretamente todos los elementos deberán llevar fuente redundantes pudiendo funcionar con la mitad de ellas sin pérdida de rendimiento.
R5	El almacenamiento deberá proveer las conexiones necesarias para conectarse al sistema de red del actual almacenamiento (Force10 E1200i 10 Gbit Ethernet) y para conectarse al superordenador MareNostrum mediante la red Infiniband sin sobre-suscripción (no bloqueante). Cada módulo ofrecido de “declustered raid” deberá como mínimo proporcionar: <ul style="list-style-type: none"> <li>- 4 conexiones 10 Gbit Ethernet</li> <li>- 8 conexiones Infiniband FDR (no bloqueantes)</li> </ul>
D6	Se valorará la mejora de conexiones proporcionadas por módulo de declustered raid
R7	Se deberá de proveer todos aquellos servidores y elementos necesarios (hardware/software) para la gestión de los nuevos elementos, o la reconfiguración de los existentes. Se priorizará el uso de elementos

Ref	Descripción
	administrativos existentes antes de la inclusión de nuevos.
R8	Los servidores proporcionados deberán disponer de una metodología de administración out-of-line: consola remota, power on y off, monitorización de alarmas y temperaturas, etc.
R9	Los servidores proporcionados que sirvan este almacenamiento, deberán disponer de potencia de cálculo libre para poder soportar operativas de BigData, como pasa en las plataformas NSD basadas en Power8.
R10	El almacenamiento de datos del actual sistema de almacenamiento se deberá ir reconfigurando de forma progresiva para que sea el nuevo sistema que sirva los filesystems de projects y scratch de todos los supercomputadores del BSC.

## A2. Metadatos

Ref	Descripción
R1	Se deberá proveer de un almacenamiento basado en flash o SSD que albergará los metadatos de los filesystems GPFS en la actualización del almacenamiento HPC del BSC. Dicho almacenamiento deberá tener un mínimo de 50 TB netos disponibles. Deberá ser capaz de proporcionar un rendimiento de IOPS entre 300 mil y 1 Millón IOPS y un rendimiento de metadatos (bloques de 4KB) entre 4 y 20 GB/s.
D2	Se valorará la mejora del almacenamiento proporcionado y la mejora en rendimiento proporcionado
R3	Todos los elementos del almacenamiento flash deberán ser redundantes, y no deberán tener un único punto de fallo.
R4	La solución propuesta deberá ser basada en módulos, los cuales todos ellos deberán ser idénticos, la solución ofertada debe permitir escalar tanto en rendimiento como en capacidad.
R5	La solución de flash deberá permitir su configuración en bare-metal, pudiendo instalar el sistema operativo de forma directa en los servidores como la configuración del almacenamiento de forma nativa.
R6	El almacenamiento de metadatos deberá proveer las conexiones necesarias para conectarse al sistema de red del actual almacenamiento (Force10 E1200i 10 Gbit Ethernet) y para conectarse al superordenador MareNostrum mediante Infiniband sin sobre-suscripción (no bloqueante). Cada módulo de metadatos deberá de tener un mínimo de: <ul style="list-style-type: none"> <li>- 4 conexiones 10 Gbit Ethernet</li> <li>- 4 conexiones Infiniband FDR (no bloqueantes)</li> </ul>

D7	Se valorará la mejora de conexiones proporcionadas por cada módulo de metadatos basado en flash proporcionado
R8	Los servidores que formen parte de estos componentes deberán disponer de una metodología de administración out-of-line: consola remota, power on y off, monitorización de alarmas y temperaturas, etc.
R9	El almacenamiento de metadatos del actual sistema de almacenamiento se deberá ir reconfigurando de forma progresiva para que sea el nuevo sistema que sirva los filesystems de home y apps de todos los supercomputadores del BSC.

### A3. Servidores Análisis de datos

Ref	Descripción
R1	<p>Se deberán de proveer un mínimo de 2 servidores para análisis de datos basados en arquitectura Power8 para poder integrar cálculos de análisis de datos y Big Data.</p> <p>Cada servidor deberá como mínimo cumplir las siguientes características:</p> <ul style="list-style-type: none"> <li>- 20 cores por server</li> <li>- 256 GB RAM</li> <li>- 1 tarjeta PCIeexpress NVMe SSD con un mínimo de 1.5 TB de capacidad</li> <li>- 2 discos NL-SAS de 1 TB mínimo</li> <li>- 1 tarjeta NVIDIA K80 ó 2 tarjetas NVIDIA K40</li> </ul> <p>Todos los servidores ofrecidos deberán tener las mismas características técnicas.</p>
D2	Se valorará la inclusión de más servidores o las mejoras técnicas mínimas antes indicadas de cada uno de ellos.
R3	<p>Cada servidor deberá tener las siguientes interfaces de red para conectarse con el almacenamiento HPC (actual + ampliación):</p> <ul style="list-style-type: none"> <li>- Tarjeta Infiniband FDR para conectarse a la red Infiniband de MN3 y ampliación almacenamiento</li> <li>- Mínimo de 2 interfaces Gigabit Ethernet</li> <li>- Una interfaz 10 Gbit Ethernet para conexión a la red GPFS (Force10 E1200i)</li> <li>- Interfaz Ethernet de gestión out-of-line</li> </ul>
R4	Cada servidor deberá ofrecer los buses independientes suficientes para poder soportar las conexiones a las diversas redes que se describen anteriormente, sin ser ningún factor limitante. Deberá venir con una configuración óptima (DIMMs instalados usando todos los buses posibles) para la realización de operaciones Entrada/Salida.
R5	Cada server deberá disponer de puertos PCI-Express libres suficientes para la futura instalación de GPUs u otros componentes



R6	Cada servidor deberá disponer de un sistema de administración remoto (out-of-band), el cual debería permitir como mínimo: poder realizar el power on/off, coger la consola, monitorización del entorno (Temperatura, consumo, ...), generación de alarmas por SNMP, detección de problemas hardware/firmware, etc.
R7	Durante el año de garantía de estos servidores (duración de garantía) se deberá de proveer la actualización/sustitución de dichos servidores por unos basados en nuevas tecnologías de procesador (Power8+, Power9, Power9+, etc.) sin ningún coste. Las actualizaciones deberán como mínimo mantener o mejorar las características de los servidores que sustituyen.
R8	Se deberá incluir todo aquel cableado necesario para la conexión de estos servidores al actual cluster para su óptimo funcionamiento.

#### A4. Switches y Redes

Ref	Descripción
R1	Se deberán de proveer todos aquellos elementos de red necesarios para la ampliación tanto del almacenamiento HPC del BSC, como la ampliación del Archivo Activo. Dicha necesidad puede venir de la configuración final, como de cualquier fase intermedia de la ampliación necesaria debida a la copia de datos, etc.
R2	Se entiende, por ejemplo, y sin estar limitado a esta lista, como elementos de redes necesarios a proveer: <ul style="list-style-type: none"> <li>- Switches Gigabit Ethernet</li> <li>- GBICs 10 Gbit Ethernet o linecards del equipo Force10 E1200i del BSC</li> <li>- Cables cobre / Fibras multi o mono modo (1/10 Gbit Ethernet)</li> <li>- Switches Infiniband</li> <li>- Modulos Infiniband para los IB core de MareNostrum 3</li> <li>- Cables Infiniband FDR</li> </ul>
R3	Cualquier componente extra a aportar deberá ser completamente compatible e integrable en los componentes de red ya existentes de las diversas soluciones (Almacenamiento HPC como Archivo Activo).
R4	Todos los links 10 Gbit Ethernet salientes de los nuevos componentes (datos y metadatos) a integrar en el almacenamiento HPC deberán ser line-rate (sin sobresuscripción) a la hora de conectarlos a los switches 10 Gbit Ethernet.
R5	Una vez los nuevos componentes de datos y metadatos entren en producción, MareNostrum 3 deberá ser capaz de usar la red Infiniband para conectarse al nuevo almacenamiento. Se puede considerar la reducción de los uplinks de 10 Gbit Ethernet de MareNostrum a los switches Force10 E1200i a la mitad una vez esto pase.

R6	Cualquier switch (Ethernet/Infiniband) nuevo que se requiera para la ampliación deberá venir con fuente de alimentación redundante y cumplir con las mismas características técnicas de los switches actuales que complementaran.
R7	Se deberá de proveer un switch Gigabit Ethernet con 48 puertos cobre 1 Gbit Ethernet y 4 puertos uplink a 10 Gbit Ethernet a instalar en el rack IB8 para crear una red de management con los servers de análisis de datos ofertados a los cuales se integrarán los 2 servidores actuales Power8 que ya dispone el BSC. De cada servidor Power8 conectará como mínimo cables (HMC y administración SSH). Este switch deberá tener las fuentes y los ventiladores modulares, (se deben poder cambiar en caliente).
R8	Se deberán de proveer los componentes de red necesarios para poder conectar los servidores de Análisis de datos a: <ul style="list-style-type: none"> <li>- Red Infiniband de MareNostrum3 y ampliación de almacenamiento</li> <li>- Red 10Gbit Ethernet del almacenamiento</li> </ul>
R9	Se deberá reconfigurar la Infiniband de MareNostrum para albergar los nuevos componentes y poder separar lógicamente el tráfico de MPI con el de IO del nuevo almacenamiento para poder monitorizarlo por separado.
D10	Se valorará el dejar un 5% de puertos libres en todos los switches afectados por esta ampliación para futuras ampliaciones.

## B. Operacional

Ref	Descripción
<b>Requerimientos operacionales</b>	
R1	Los 3 racks deberán venir incluidos con la solución y deberán venir con las PDU's adecuadas para conectar todos los equipos de la solución y proporcionar redundancia en la circunstancia de la caída del 50% de las PDUs de cada rack.
R2	Se deben proveer 3 racks con toda la ampliación instalada (a excepción de los servidores de análisis de datos)
R3	Los racks deberán incorporar refrigeración por puerta trasera que elimine como mínimo el 80% del calor generado.
R4	Tanto los racks como las puertas traseras deberán ser compatibles con los racks actuales del sistema de almacenamiento HPC del BSC, de tal manera que no haga falta ninguna actuación de la instalación eléctrica o del sistema de refrigeración por agua por puerta trasera.
R5	El peso de cada rack no deberá nunca superar más de 2000 Kg. x m2 y un límite de consumo eléctrico por rack de 15 kW
R6	Se deberá presentar en la documentación un esquema con la ocupación de los racks de los diversos equipos presentados en la solución. En ella se deberá claramente especificar el hardware ofertado para cumplir los requerimientos mínimos y las ampliaciones indicadas y el espacio ocupado, como las U's que ocupa cada componente de la solución. También se deberán de especificar por cada tipo de rack el número de cables/fibras que sale de cada rack para cada una de las redes definidas.
R7	Todos componentes de la solución deben disponer de fuentes de alimentación redundadas
R8	Se deberá presentar esquema de conexionado eléctrico interno de cada tipo de rack de la solución. Mostrando la redundancia de todos los elementos a nivel de alimentación a través de diferentes grupos de PDU
D9	Se valorará el uso mínimo de U's de la solución
R10	Todo conexionado entre racks se deberá realizar a través del falso suelo, no se permitirá la tirada de cables entre racks colindantes o por la parte superior del rack. El cableado dentro del rack deberá ser ordenado y nunca salir del espacio que determina la planta del rack.
R11	Los componentes deberán 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.
R12	Se exigirá en la instalación cableado (fibra, eléctrico, cobre Ethernet, etc.) ordenado, elegante y vistoso debido a que quedará a la vista. Se deberá seguir el código de colores de cableado usado en la solución actual.
R13	Todo componente de la solución (rack, server, switch, cable, fibra, ...) deberá

Ref	Descripción
	ir debidamente etiquetado, para ser identificado físicamente de forma única según nomenclatura establecida por cada una de las soluciones ya en producción. En los cables y fibras se deberá indicar origen y destino de la conexión.
R14	Los servidores de Análisis de datos se deberán enrackar en el rack IB8 de MareNostrum 3 y deberán ocupar un máximo de 24Us
R15	Se deberá de proveer de una puerta trasera de refrigeración por agua para el rack IB8 y su instalación y puesta en marcha
R16	La solución deberá incluir el montaje en racks de toda la solución, además de la recogida de todos los materiales sobrantes de la instalación.
R17	Los 3 racks de datos salientes del almacenamiento actual se deberán de enviar y poner en marcha en el CPD del PCB para poder integrarlos con el archivo activo.
R18	El proyecto deberá incluir cualquier trabajo logístico derivado de las acciones del plan de la actualización, explicado más adelante, como puede ser por ejemplo: Transporte de racks, movimiento de racks dentro de los CPDs, recableado, purgado y puesta en marcha de puertas traseras, conexionado eléctrico de racks, etc.
R19	La ampliación se debe entregar llave en mano y completamente en producción siguiendo cada una de las fases especificadas en el plan de actualización

### C. Software

Ref	Descripción
R1	Cualquier servidor debe venir con sistema operativo Linux y compatible con las instalaciones ya existentes tanto en el almacenamiento HPC como en el Archivo Activo. El sistema operativo deberá estar soportado por GPFS.
R2	Se deberá aportar también todo el software necesario para la gestión de todos los componentes que formen la ampliación o de la ampliación de licencias de software de las soluciones actuales.
R3	Con el cluster de Análisis de datos se deberá de proveer de compilador nativo para dicha arquitectura, como mínimo deberá incluir C, C++, Java, Fortran, y OpenCL.
R4	Con el cluster de Análisis de datos se deberán de proporcionar las librerías numéricas (secuenciales y paralelas) proporcionadas por el fabricante de los procesadores debidamente optimizadas para dicha arquitectura.
R5	También se deberá proporcionar los compiladores, librerías y/o las herramientas necesarias para el uso paralelo de la arquitectura mediante paradigmas estándares como OpenMP ó MPI.
R6	Se deberán incluir el software, licencias de Infiniband y las licencias de GPFS para todos los componentes que forman parte de la ampliación (Datos declustered, metadatos, servidores de análisis de datos)

## D. Mantenimiento y soporte

Ref	Descripción
R1	Garantía y soporte de 1 año en todos los componentes (hardware y software), con una respuesta en 4 horas dentro de las horas de oficina (08:00 – 17:00) y con un servicio de soporte de Next Business Day.
D2	Se valorará la extensión de la garantía / mantenimiento tanto en años de duración como cobertura y tiempo de respuesta dentro del horario de oficina.
R3	El proyecto de instalación incluirá la comprobación del buen funcionamiento, integración y óptimo rendimiento de la ampliación en cada una de las fases del proyecto.
R4	Se exigirá un trabajo en equipo con el equipo de sistemas del BSC, para la coordinación en cada una de las fases del proyecto de ampliación y solución de cualquier problema de incompatibilidad que surja en la ampliación.
R5	Se proporcionará (dentro de período de garantía): <ul style="list-style-type: none"> <li>- Acceso a todos los software upgrade (incluyendo sistemas operativos, clientes GPFS y firmware) de todos los componentes de la solución</li> <li>- Punto único de soporte para el aviso de problemas e incidencias de cualquier componente que componga la solución</li> </ul>
R6	Se exigirá soporte pro-activo, notificando y recomendado subidas de versión tanto de software como de firmware de cualquier componente de la solución.
R7	Se deberá entregar al final de la instalación una documentación digital en la que se describa: <ul style="list-style-type: none"> <li>- Configuración en cada una de las fases de la ampliación</li> <li>- Descripción general de los componentes con la configuración final</li> <li>- Esquema de conexionado e IPs</li> <li>- Valores de configuración empleados</li> <li>- Explicación del proceso de instalación y tareas realizadas</li> <li>- Modificación de los procedimientos para: Puesta en marcha, y disaster recovery ya existentes</li> </ul>
R8	Toda la instalación y desarrollo del proyecto se deberá hacer on-site en las instalaciones del BSC bajo la supervisión del grupo de sistemas del BSC. En ningún caso se permitirá el acceso externo o remoto para la configuración o instalación de la solución presentada, a menos que el BSC lo autorice.
R9	Se deberán ofrecer formación durante la instalación de la solución, que cubran: <ul style="list-style-type: none"> <li>- Conceptos básicos</li> <li>- Administración básica y procedimientos básicos de configuración</li> <li>- Optimización de la solución</li> <li>- Solución de problemas</li> </ul>
R10	En la implantación de la solución presentada se exigirá la participación activa y presencial de los expertos de cada uno de los componentes que forman la

	<p>actualización, para asegurar la configuración y tuning óptimo de los nuevos componentes:</p> <ul style="list-style-type: none"> <li>- Responsables de hardware/ desarrolladores de firmware</li> <li>- Desarrolladores o responsables técnicos del almacenamiento basado en declustered raids</li> <li>- Desarrolladores o responsables técnicos del almacenamiento basado en flash</li> <li>- Desarrolladores o responsables técnicos de GPFS</li> <li>- Desarrolladores o responsables técnicos de Infiniband</li> </ul> <p>Teniendo la posibilidad el personal del BSC poder intercambiar emails de forma directa con dichas personas con el fin de solucionar cualquier problema que surja durante el desarrollo e instalación de la máquina.</p> <p>Si cualquiera de los técnicos expertos designados para la configuración, no se considera que dispone los conocimientos y experiencia adecuada, se deberá de reemplazar por uno que si cumpla con dicha experiencia.</p>
R11	<p>Una vez acabada la instalación se deberá de comprobar que todos los requerimientos de operativa establecidos en este pliego se cumplen. Como por ejemplo:</p> <ul style="list-style-type: none"> <li>- Redundancia eléctrica allí donde se requiera</li> <li>- Eficiencia energética</li> </ul> <p>...</p>
R12	<p>Se deberá comprobar el óptimo funcionamiento de todos los componentes de la actualización y demostrar empíricamente que cumplen los rendimientos (GB/s, IOPS, PFlops, ...) ofertados.</p>
R13	<p>Se deberá entregar toda la documentación de la instalación y administración del sistema en formato Office.</p>

## D. Fases de proyecto

En este apartado se describen a groso modo las fases de la actualización del almacenamiento que este proyecto conlleva, para facilitar el cálculo de material y servicios necesarios.

La lista de tareas aquí descrita no es limitante, cualquier tarea no incluida en esta lista pero necesaria para llevar a cabo los objetivos descritos al principio de este documento estaría incluida dentro del proyecto.

### Fase 1 – Integración datos declustered + metadatos

- Instalación y conexionado de racks con ampliación de datos y metadatos en la parte final de la sala de MareNostrum
- Integración y conexión de los nuevos componentes con el actual almacenamiento y red Infiniband de MareNostrum3
- Reconfiguración de Infiniband de MareNostrum3 para poder monitorizar y separar tráfico I/O y MPI
- Instalación, configuración y optimización de los nuevos componentes por expertos de GPFS y declustered raids
- Actualización o reconfiguración de los servidores de servicios actuales para soportar los nuevos componentes
- Reconfiguración progresiva del antiguo sistema de metadatos para poder servir los filesystems de home y apps
- Inicio de copia/sincronización de datos entre los componentes anteriores y los nuevos. Los técnicos del BSC se harán responsables de la copia/sincronización de datos pero consultarán con el proveedor cual puede ser el mejor mecanismo para realizarlo.

### Fase 2- Puesta en producción de la ampliación

- Movimiento y conexión racks nuevos a primera fila de MareNostrum y recableado definitivo con los componentes del almacenamiento actual y red Infiniband de MareNostrum3
- Retirada de cualquier cableado antiguo del falso suelo que quede en desuso
- Movimiento de los racks de los 3 data building blocks anteriores al final de la sala de MareNostrum y prepararlos para su envío
- Puesta en producción con los componentes de la ampliación: datos declustered + nuevos metadatos + home/apps filesystem servers (antiguos metadatos)

### Fase 3 – Ampliación archivo activo

- Envío y transporte de los 3 data building blocks al CPD del CNAG
- Puesta en posición y conexión racks (eléctrico, sistema refrigeración por agua, redes Ethernet del archivo activo)
- Reconfiguración de los 3 data building blocks para ser integrados como almacenamiento de archivo activo
- Ampliación de la capacidad del archivo activo y puesta en producción



Fase 4 – Instalación cluster de servidores de Análisis de datos

- Enrackado servidores de Análisis de datos + switch en IB8 (se puede hacer en paralelo con la Fase 1)
- Integración con las redes 10 Gbit Ethernet del almacenamiento como con la red Inifinband de MareNostrum 3 (Esto sólo se puede realizar una vez finalizada la Fase 3)
- La configuración de este cluster correrá a nombre del equipo de sistemas del BSC