

Expediente 1/2015: Contratación documentalmente simplificada, sujeta a la regulación armonizada, por el procedimiento abierto y tramitación ordinaria, del suministro de un sistema supercomputador para el Centro de Supercomputación de Galicia (Finis terrae II), cofinanciado por el fondo europeo de desarrollo regional 2007-2013

Respuestas a las consultas técnicas recibidas acerca de los pliegos

1) ¿Los 180KW de potencia límite indicados en el pliego son sólo para el suministro eléctrico del cluster ó incluyen la refrigeración? ¿Este tope de consumo es referenciado al 100% de carga ó con carga equivalente a correr el Linpack?

Los 180KW son para el suministro del cluster, solo para la parte IT, es decir incluyendo los servidores y todos los equipos hardware suministrados, corriendo el Linpack y con todos los componentes encendidos. No hay que incluir la refrigeración.

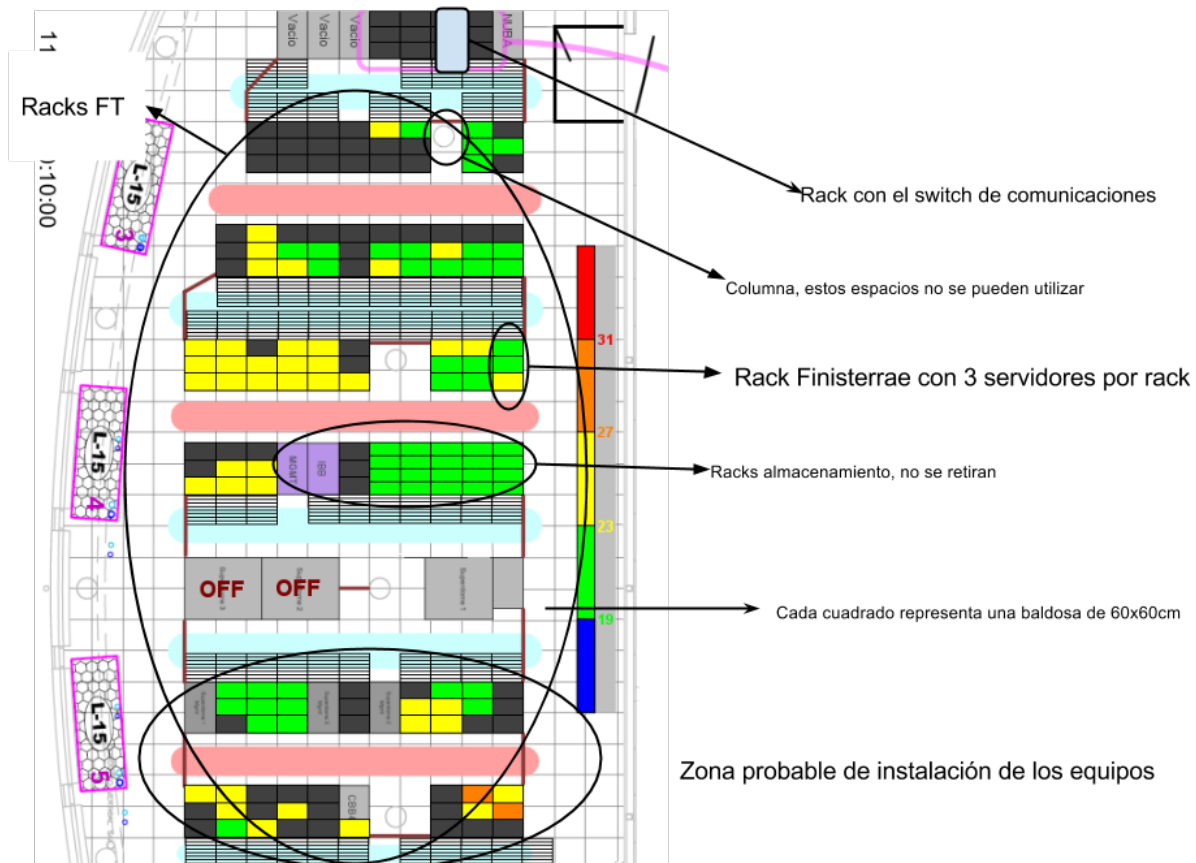
2) ¿El CESGA cuenta con suelo técnico en el CPD? ¿El cableado actual y preferencia para el cluster a suministrar será por el suelo?

Actualmente el CESGA cuenta con suelo técnico con altura suficiente y sistema rejiband para el cableado. No tendremos preferencia por el cableado que sea por suelo o por techo, pero si no es por el suelo tendrían que suministrarse los materiales y realizarse la instalación necesaria para completarla.

3) ¿Los transceivers de la electrónica de Red del Cesga donde se conectarán los puertos 10GbEth del cluster son provistos por el CESGA?
Sí.

4) ¿Cuál es la ubicación en el CPD (para calcular la distancia de los cables)?

Se adjunta un plano con la distribución actual del Finis Terrae y la ubicación del rack en el que se encuentra el switch al que se conectarían los puertos de 10 GbEth.



5) ¿Hay una preferencia del CESGA de ubicación en la sala para el cluster a suministrar? ¿Es posible elegir libremente por el suministrador la ubicación del mismo contando tanto con el espacio libre disponible como con las huellas ocupadas por el actual Finisterrae I con el fin de encontrar la posición más idónea mirando tanto al suministro eléctrico como a la refrigeración más óptima del nuevo cluster?

La ubicación definitiva se planificará y consensuará durante la instalación en función del número de racks y sus características para determinar la mejor posición en el CPD del centro.

6) ¿Es aceptado ofertar para los FAT Nodes servidores con dos procesadores Haswell con memoria DDR4 cuya escalabilidad máxima es 1,5TB de RAM en lugar de los 2TB de RAM indicados en el pliego?

El pliego exige memoria de 768GB, y que se valorará que sea ampliable a 2TB (es decir, no es requisito).

7) ¿Hay alguna preferencia de distribución Linux?
Preferiblemente RedHat-like.

8) Listado/inventario de equipos de Finisterrae I a valorar desinstalación y destrucción.

142 RX7640 con 142 MSA50 (48 racks) y 3 Superdomes (9 racks)

9) ¿Es posible obtener planos actualizados de ocupación del CPD, así como de instalaciones (eléctrica, refrigeración y agua)

Sí, se adjunta un plano con la ubicación actual de los racks de Finis Terrae y las climatizadoras (unidades L15 en el plano)

Plano colocación:

https://docs.google.com/drawings/d/1vq0f4XtJ_Qb4nGe3UBODyDmGJ9I_I87b_Az3vRuNWHU/edit?usp=sharing

10) Arranque de los nodos. En dos párrafos de la RFP se dice que los nodos podrán arrancar desde la red o desde disco. El arranque en red, ¿es para arrancar el nodo en modo "diskless"? o ¿es para solo para arrancar desde el gestor de cluster, pero usando el Sistema Operativo del disco interno?

Normalmente utilizamos el arranque por red para instalar los nodos por red, pero no descartamos el utilizarlos en modo diskless.

11) Según el plano del CPD suministrado, hay un elemento marcado como "rack con el switch de comunicaciones". ¿Es esa la ubicación de los conmutadores a los que se refiere el punto 2.3.3 del Anexo VII, "Conexión a la red del CESGA"? Si es así, ¿cuál es la distancia que separa este armario de los situados en la parte inferior izquierda del plano?

Sí, el "rack con el switch de comunicaciones" es al que se refiere en el punto 2.3.3 del Anexo VII y la conexión a la red del CESGA. Las distancias se pueden calcular teniendo en cuenta que cada cuadrícula es una baldosa de 60x60cm, es decir, 25 baldosas (1500cm) en un sentido y 8 baldosas (480cm) en el otro.

12) ¿Cómo se realiza actualmente la evacuación del aire del pasillo caliente? ¿Existe algún tipo de contención para facilitar dicha evacuación?

Los pasillos fríos están cerrados, los pasillos calientes abiertos y el aire caliente circula libremente hasta las climatizadoras L15 que se encuentran en un extremo del CPD (figuran en el plano enviado a la izquierda).

13) Monitorización:

¿Está CESGA familiarizado con alguna solución de Monitorización y gestión del mundo Open?
P.ej. Nagios / Ganglia.

Utilizamos actualmente Nagios y Ganglia.

14.) el SW actual de monitorización, ¿que alcance tiene?

Disponemos de un software de monitorización de los equipos basado en Nagios y Ganglia y otro de la infraestructura que monitoriza temperaturas, consumos eléctricos, estado de las enfriadoras, temperatura del agua, estado de los chillers y SAls, conexión a grupo electrógeno, en líneas generales

15) Gestor de Carga:

El GridEngine desplegado en la actualidad ¿cual es, el open Source basado en SUN, o el Open Source basado en UNIVA?

Es el open source basado en SUN.

16) Tarjeta iDRAC

¿Es necesario Direccionamiento de consola?, el usuario puede ver en su pantalla desde el arranque del sistema hasta el S.O.

Como indica el pliego, el sistema debe incluir una red independiente que permita acceder a las facilidades de consola virtual de los nodos. Además dispondrán de un sistema de administración remoto que permita como mínimo realizar el encendido y apagado del clúster, acceder a la consola, ...

17) Red Ethernet nodos de cálculo

Solo solicitan 2 puertos máximo 1GbE con soporte PXE, ¿es correcto? Por el tipo de características de la red pág. 74 parece que si.

En los nodos de cálculo es necesario suministrar al menos 2 puertos con soporte PXE y conector RJ-45

18) Nodos FAT y Transfer

No se solicita red IB ¿Es correcto?

Los nodos FAT y Transfer deben tener los mismos requisitos que los mínimos de los nodos de cálculo salvo que se especifique lo contrario, por tanto también deben incluir la tarjeta de red de baja latencia Infiniband o con prestaciones equivalentes.

19) Características de los servidores aceleradores

No se especifican características de Procesador y RAM de este sistema, solo la GPGU en este caso NVIDIA K40 y el Xeon Phi 7120P ¿Podemos tener alguna aproximación?

Los nodos aceleradores deben tener los mismos requisitos mínimos de los nodos de cálculo en cuanto a procesador y memoria RAM, tal y como se indica en el pliego

20) Red de Aplicación, Red de Gestión sobre todo para la Red de Aplicación, ¿es altamente valorable la opción 10GbE para esta red?

10 GbE no es un requerimiento salvo para los nodos que se indican en el pliego. La influencia de esta característica en la valoración global de la solución no puede determinarse sin conocer el global de la solución.

21) Garantía

Solo se solicita NBD y se valora 2 puntos ¿es correcto?.

Como indica el pliego, el tiempo de respuesta desde la recepción del aviso del incidente y la presencia efectiva del técnico en las instalaciones del CESGA para incidentes críticos será como máximo al día siguiente laborable. Las mejoras en este apartado se valoran según el criterio del plan de implantación y garantía, cuya máxima puntuación es de 2 puntos.

22) En el dimensionamiento del almacenamiento del proyecto de CESGA necesitaríamos saber valdría con una config que proporcione 20GB/s en lecturas y se quede en 8.7GB/s en escrituras, o es necesario/imprescindible los 20GB/s escrituras

En el pliego se indica que el sistema de almacenamiento paralelo de altas prestaciones debe proporcionar al menos 20 GByte/s sostenido y demostrable mediante IOZONE/IOR desde los nodos de cálculo en lectura y escritura.

En cuanto al software:

23) Gestor de despliegue y management del cluster que prefiere el cliente

No tenemos preferencia.

24) Gestor de colas que desean que se instale.

No estamos cerrados a ninguna solución que pueda aprovechar eficientemente los recursos, si bien actualmente trabajamos con Gridengine, y también consideramos que Slurm es una buena solución.

25) Benchmark lozone:

¿ Hay un fichero, distributed_benchmark_all.sh, puesto a disposición de los licitadores para su descarga desde el CESGA ?

Se ha subido el fichero a la plataforma de contratación.

26) ¿Se medirá el rendimiento del sistema de fichero paralelo servido mediante red Ethernet y protocolo NFS al resto de los servidores del centro, dentro las pruebas de rendimiento benchmarks.?

En los benchmarks no se hace referencia a esta prueba, aunque será una funcionalidad que al estar incluida en el pliego podrá formar parte de las pruebas de aceptación:

"Deberá conectarse mediante red Ethernet y protocolo NFS al resto de los servidores del centro. Para ello deberán proporcionarse los elementos necesarios para obtener un ancho de banda de 1 GByte/s desde fuera de la solución mediante clientes NFS."

"Pruebas de funcionalidad. Se realizarán durante un período de 48 horas y deberán completarse de modo satisfactorio en un tiempo máximo de 14 días, a contar a partir de la puesta a disposición del CESGA de los equipos. Las pruebas verificarán el cumplimiento de las funcionalidades requeridas por el CESGA en el presente pliego y de las mejoras ofertadas por el adjudicatario."

27) Tenemos una duda con respecto a las siguientes características indicadas en el cuadro de características técnicas que tenemos que rellenar:

Sistema de ficheros jerárquico (en caso de soportar)
Políticas soportadas por el sistema de ficheros jerárquico
Snapshot
Clone
Soporte de réplicas
Software de copia remota

En el pliego de características técnicas no hace referencia a la obligatoriedad de cumplirlas, pero queremos asegurarnos y verificar que son opcionales.

"Las ofertas deberán contemplar la retirada de Finis Terrae I, puesto que será totalmente desconectado tras la puesta en producción de los nuevos equipos." (Página 82)

Las dudas son:

28) Entendemos por retirar que el adjudicatario puede/debe de llevarse Finisterrae I al completo? (Nodos Rx7640, Superdome, Comunicaciones Ethernet, Infiniband, Racks, etc...). En caso afirmativo, se requiere algún tipo de certificado al respecto?

Se trata de retirarlo y llevarlo a algún almacén del CESGA? En este caso necesitaríamos ubicación del mismo.

Se trata de que el ganador del concurso desmonte y retire a su cargo el equipamiento del actual Finis Terrae que no vaya a ser reutilizado en el nuevo Finis Terrae por el ganador del concurso. Necesitaremos un certificado de achataamiento del equipamiento retirado con el cual el CESGA dará de baja de su inventario esos equipos.

Las normas de reciclaje de equipos informáticos obsoletos, determinan las actuaciones a realizar para obtener el mejor aprovechamiento de los residuos resultantes.

29) En ambos casos de posible retirada, dispone el CESGA de las “baldosas” que cubran los huecos que dejen los racks a retirar?

Actualmente disponemos de 35 baldosas enteras para utilizar.

30) En el párrafo:

2.5.1- Software de Sistema

El sistema operativo de todos los elementos excepto del sistema de almacenamiento paralelo debe ser Linux. Las versiones implementadas han de ser soportadas por el fabricante de los equipos que se suministren.

¿Por versión soportada se entiende que debe ser un sistema operativo que simplemente este validado por el fabricante y que con dicho SO instalado se asegure que la maquina funcione en un momento determinado de la vida de la misma con la versión instalada y se atiendan todas las incidencias HW que se produzcan en el sistema?

El proveedor debe dar soporte de todo el HW y SW suministrado, con un punto único de contacto, durante el período de garantía.

31) ¿O adicionalmente a esto ha de darse soporte a las incidencias/consultas relativas al propio sistema operativo?

El sistema operativo está incluido en el suministro y por tanto debe estar soportado.

32) ¿Soporte también a los drivers que puedan llegar a necesitarse en un momento determinado?

Los drivers deben de estar soportados por el sistema operativo suministrado. Las actualizaciones y nuevas versiones del sistema operativo que puedan requerir actualizaciones de los drivers deberán analizarse entre las partes. En el caso de actualizaciones (y versiones) del sistema operativo para resolver incidentes críticos, se deberá suministrar una actualización de los drivers que funcione con esta actualización. En el caso de otras actualizaciones (y versiones), no será necesario suministrar el driver actualizado cuando la actualización sea motivada por el CESGA.

33) Ha de darse soporte no solo a la versión que se instale sino también a las versiones futuras de dicho SO (como mínimo durante los 4 años del contrato)? Han de proporcionarse las nuevas versiones de dicho sistema operativo?

El soporte a nuevas versiones del sistema operativo estará vinculado a los casos mencionados en el caso anterior: para la resolución de incidentes críticos o por consenso entre las partes

34) Entendemos que dicho soporte es para el sistema en completo, es decir, no solo nodos de cálculo sino también, almacenamiento red de interconexión, software de monitorización y gestión suministrado, etc... es correcto?

Sí, el proveedor debe dar soporte de todo el HW y SW suministrado, con un punto único de contacto, durante el período de garantía.

35) En el formulario hay un apartado referente a la eficiencia energética, los GFlops son los sostenidos ó los GFlops pico. No lo indica pero imagino que serán los sostenidos, ¿es correcto?

Eficiencia energética (GFlops/Kw)			
-----------------------------------	--	--	--

El consumo eléctrico es el de todo el sistema o sólo de las máquinas que computan para obtener el LINPACK. Imagino que todo el sistema, es correcto?

Los GFlops se refieren a los GFlops pico teóricos.

En la pestaña de la solución global las características del cluster de computación no deben de incluir el sistema de almacenamiento paralelo, por tanto, la eficiencia energética se calcula como los GFlops pico teóricos agregados de todos los nodos de cálculo dividido entre el consumo máximo de la solución sin el sistema de almacenamiento paralelo

Todos los demás datos de esa pestaña deben de cubrirse sin incluir el sistema de almacenamiento paralelo

En la pestaña de los nodos de cálculo la eficiencia energética se calcula como los GFlops pico teóricos del servidor dividido entre el consumo máximo del servidor.

36) Entiendo que podemos aprovechar las lineas electricas de cierto numero de racks (previa retirada de los mismos) y que estas lineas son dos líneas trifásicas de 32A por rack con conectores CTAC (o sucko, creo que los trifásicos son los rojos). Es así?

Sí, actualmente tenemos dos tomas trifásicas de 32A por rack con conectores CTAC que se pueden reaprovechar.

37) No creo que sea un problema, pero nos podríais confirmar la carga máxima soportada por el forjado y suelo técnico?

El forjado y suelo técnico podría soportar hasta 2.500Kg/m²

Santiago de Compostela, 24 de marzo de 2015

José Ignacio López Cabido

Subdirector Técnico de la Fundación CESGA