

Pliego de prescripciones técnicas que regulan la adquisición de una infraestructura de servidores para servicios virtualizados y de cálculo con soporte GPU (actuación que forma parte de la inversión 2 “*Fortalecimiento de las capacidades, infraestructuras y equipamientos de los agentes del SECTI*” del componente 17 del PRTR)

Barcelona, en la fecha de la firma

DOCUMENTO TRADUCIDO DEL ORIGINAL REDACTADO EN CATALÁN, EN CASO DE CONFLICTO ENTRE AMBOS DOCUMENTOS PREVALECE LA REDACCIÓN EN IDIOMA CATALÁN

Validado por,	Aprovado por,

1.	Antecedentes	5
2.	Objetivo y alcance del proyecto	5
3.	Características técnicas del equipamiento	6
4.	Instalación y configuración de los equipos	8
5.	Garantía de soporte	10
6.	Prestaciones que optimizan el rendimiento de los equipos	12
7.	Duración del proyecto y plazos de entrega	13

1. Antecedentes

El Consorci de Serveis Universitaris de Catalunya (CSUC) tiene una larga tradición en la prestación de servicios de computación de altas prestaciones a la comunidad investigadora. Este servicio tiene como objetivo mejorar la eficiencia de todo el sistema, permitiendo a los grupos de investigación disponer de un hardware de alto rendimiento, gestionado por un equipo de soporte experto y que se mantiene actualizado con las últimas tecnologías disponibles en el mercado.

Todos estos servicios se apoyan en la infraestructura tecnológica disponible en el CSUC. En cuanto a la arquitectura de servidores, esta está orientada a la nube y se basa en un diseño de abstracción en tres capas: virtualización, almacenamiento y hardware.

Este modelo proporciona una arquitectura sólida que permite entornos complejos, de gran volumen, flexibles y redundantes; así como una base tecnológica homogénea y, por tanto, eficiente y sostenible. Además, permite una evolución continua que garantiza la calidad y disponibilidad de los servicios.

2. Objetivo y alcance del proyecto

En el marco de las actuaciones previstas para dar respuesta al crecimiento de los servicios digitales y científicos, se hace necesario disponer de una infraestructura común que permita ampliar la capacidad de virtualización e incorporar nuevos recursos especializados para la visualización de simulaciones en entornos de cálculo avanzado.

Para alcanzar estos objetivos, se prevé la adquisición de una infraestructura de servidores de alto rendimiento que pueda dar servicio de manera integrada tanto a las aplicaciones virtualizadas como a los entornos de cálculo científico para la ejecución de cargas de trabajo de inteligencia artificial (IA), con un enfoque específico en tareas de inferencia. Estos procesos, que incluyen la aplicación de modelos de aprendizaje automático para el análisis de datos en tiempo real, el procesamiento de lenguaje natural o la visión por computador, exigen un hardware especializado, como las unidades de procesamiento gráfico (GPU), para garantizar una operativa eficiente y con baja latencia.

Esta infraestructura debe permitir:

- Disponer de la capacidad de proceso necesaria para absorber un crecimiento significativo de carga de trabajo a medio plazo.
- Mejorar la flexibilidad, escalabilidad y fiabilidad de los servicios virtualizados y de cálculo, integrarse de manera transparente con el entorno existente y permitir futuras ampliaciones de forma sencilla.

- Dar respuesta a requisitos técnicos que incluyen el uso de GPUs de altas prestaciones para tareas de visualización científica, simulaciones complejas y ejecutar modelos de inferencia de manera eficiente.

Esta licitación debe incluir el suministro, la implantación y la puesta en marcha de un nuevo equipamiento destinado a la ampliación de la plataforma de Infraestructura como Servicio (IaaS). Comprende la entrega del hardware, el software y el licenciamiento asociado, así como todos los servicios necesarios para su completa integración y funcionamiento dentro de la infraestructura tecnológica existente, garantizando una solución llave en mano.

El equipamiento actual se basa en CPU Intel, y por criterios de conocimiento y estandarización se requiere que los nuevos equipos incorporen esta misma tecnología.

La gestión y orquestación de esta infraestructura se realiza mediante la plataforma de gestión de nubes OpenNebula. La tecnología de virtualización subyacente en todos los nodos de computación del clúster es el hipervisor KVM (Kernel-based Virtual Machine), que se ejecuta sobre una distribución empresarial de Linux.

El adjudicatario será responsable de proveer una solución integral que cumpla todos los requisitos especificados en este documento, desde el suministro de los equipos hasta la validación final de su funcionamiento, pasando por la instalación física y la configuración lógica.

3. Características técnicas del equipamiento

A continuación se presentan los requisitos técnicos mínimos para los servidores con GPU. Todos los equipos y componentes suministrados deberán ser completamente nuevos, sin uso previo, y pertenecer a las series de producción vigentes del fabricante en el momento de la presentación de la oferta. No se aceptará material reacondicionado, descatalogado ni en situación de fin de vida (End-of-Life), con el objetivo de garantizar la fiabilidad y la continuidad del servicio.

Asimismo, todos los servidores deberán ser del mismo fabricante, modelo y configuración para asegurar la homogeneidad de la plataforma y facilitar su gestión y mantenimiento. El número mínimo de servidores requerido es de seis (6), sin perjuicio de que puedan considerarse ampliaciones futuras en función de las necesidades del servicio.

Requisitos para cada servidor

Los requisitos mínimos para cada servidor son:

- 2 procesadores físicos por servidor, basados en la arquitectura x86-64.
- Cada procesador deberá disponer de un mínimo de 64 núcleos físicos (dando como resultado un total mínimo de 128 núcleos físicos por servidor).

- Basado en procesadores de arquitectura x86 de 64 bits, de la generación “Granite Rapids” o “Emerald Rapids” de Intel, orientados a la virtualización.
- La memoria principal debe ser como mínimo de 1,5 TB, distribuida de manera equilibrada entre todos los sockets y canales de memoria disponibles y de tipo DDR5 ECC Registrada (RDIMM).
- El formato del chasis deberá ser rack de 2U. Este formato debe ser capaz de alojar todos los componentes internos requeridos, incluyendo el número especificado de tarjetas gráficas, garantizando una refrigeración adecuada para todos ellos bajo carga máxima.
- Soportar virtualización asistida por hardware.
- Ser compatibles con el hipervisor KVM.

Requisitos para el subsistema de disco

- Disponer de dos discos locales para el sistema operativo de tipo NVMe con controladora por hardware en configuración RAID1.
- La capacidad mínima de espacio en disco deberá ser de al menos 1,92 TB.
- Los discos deberán ser de tecnología *enterprise*.

Requisitos para la conectividad

Cada servidor deberá disponer de seis interfaces de red independientes: dos para datos (NFS); dos para acceso a servicios; una para la administración de los propios nodos; y una para el acceso a la gestión mediante procesadores de servicio fuera de banda.

- Las dos interfaces de la red de datos deberán ser de tipo Gigabit Ethernet a 25 Gbps.
- Las dos interfaces de la red de acceso a los servicios deberán ser de tipo Gigabit Ethernet a 25 Gbps.
- Las dos interfaces para la administración de los servidores y la de acceso a la gestión fuera de banda deberán ser de tipo Gigabit Ethernet a 1 Gbps.

La oferta deberá incluir todos los elementos de red necesarios para poder integrar el sistema en el equipamiento de red existente, garantizando su compatibilidad y redundancia (conectores SFP de 25 Gbps en ambos extremos y cableado).

Requisitos para las GPU

- Los servidores deben tener una GPU y posibilidad de ampliarse hasta 2 en cada servidor.
- Las GPU deben tener soporte para vGPU tanto de cómputo como de gráficos.
- Las GPU deben tener un rendimiento mínimo en FP16 y FP32 de 91,6 TFLOPS.
- Las GPU deben tener un rendimiento en INT8 de aproximadamente 1,4 TFLOPS.
- Las GPU deben contar con al menos 48 GB de memoria, y ser del tipo GDDR6 con soporte ECC.
- Las GPU deben disponer de una interfaz de conexión PCI Express 4.0 x16.

Requisitos del software a incluir

- Todos los servidores deben entregarse con la última versión disponible LTS de Ubuntu Server 24.04.
- Todos los servidores deberán contar con las licencias necesarias y activadas para soportar instrucciones nativas de virtualización en el ámbito de CPUs con KVM (OpenNebula y Oracle Virtualization Manager).
- Todos los servidores deben entregarse con la última versión disponible del software de virtualización de GPUs.
- Cada GPU deberá incluir la suscripción NVIDIA AI Enterprise Essentials (NVAIE) o una solución superior del mismo fabricante que incluya todas las funcionalidades, en modalidad EDU con una duración mínima de 5 años. Esta suite de software es esencial para dar soporte a cargas de trabajo de inteligencia artificial e inferencia en un entorno virtualizado sobre KVM. La suscripción debe incluir el servicio de soporte estándar 8x5 (8 horas al día, 5 días a la semana). Esta suscripción debe garantizar el acceso a nuevas versiones del software, actualizaciones de seguridad y controladores, así como al servicio de soporte técnico.

Requisitos de las fuentes de alimentación y gestión remota

- Cada servidor deberá estar equipado con dos (2) fuentes de alimentación redundantes (1+1) y hot-swap.
- La potencia de cada fuente deberá ser suficiente para alimentar el servidor con todos sus componentes (incluidas las 2 GPU) funcionando a máxima carga.
- El servidor deberá incorporar un módulo de gestión remota avanzado con un puerto de red 1GbE dedicado e independiente de la red de datos.
- Este módulo deberá permitir una gestión completa fuera de banda a través de una interfaz web HTML5 segura, sin dependencias de Java o ActiveX.
- Las funcionalidades mínimas requeridas son: control remoto de encendido, apagado y reinicio; consola KVM-sobre-IP; montaje de soportes virtuales (Virtual Media) para la instalación remota de sistemas operativos; monitorización y alerta del estado de todos los componentes de hardware (temperaturas, voltajes, ventiladores, estado de discos y fuentes); y actualización remota de firmware.

Con el fin de que el equipo suministrado cumpla con unos mínimos de garantías de respeto por el medio ambiente, el servidor debe estar catalogado en España en el registro EPEAT, en lo relativo a su consumo energético y características medioambientales. Este dato debe poder verificarse en la página web de EPEAT, en la categoría de servidores (<https://www.epeat.net/search-servers>).

4. Instalación y configuración de los equipos

La empresa adjudicataria ejecutará las tareas de configuración que resulten de las reuniones con los responsables del CSUC, concretando los términos de las mismas. Finalmente, se elaborará una documentación exhaustiva de las configuraciones aplicadas una vez se hayan llevado a cabo, entregándola en soporte papel y electrónico.

Los trabajos a realizar incluyen:

- Instalación física en las instalaciones del CSUC del hardware y cableado (también para la electrónica de red, del servidor al *patch pannel* y del *patch pannel* a los equipos de electrónica de red), así como de los correspondientes conectores. Antes de la recepción de los equipos, el adjudicatario deberá revisar que las instalaciones técnicas (instalación eléctrica, sistemas de refrigeración, etc.) sean las adecuadas para el equipamiento ofertado y sugerir, en caso de detectarse deficiencias, aquellas mejoras o modificaciones que considere necesarias para el funcionamiento óptimo del sistema. El adjudicatario deberá proporcionar todos los componentes y material necesario para la instalación y conexión, cableado, conectores, etc.
- Configuración de los nuevos sistemas. Una vez se entreguen los equipos, el adjudicatario deberá proporcionar la asistencia técnica necesaria para la instalación y puesta en marcha del sistema, para la configuración y optimización y para la integración de los equipos en el entorno de trabajo del Centro.
 - Actualización del *firmware* de todos los componentes (BIOS/UEFI, controladora RAID, tarjetas de red, fuentes de alimentación, GPU y módulo de gestión remota) a la última versión estable y certificada por el fabricante.
 - Configuración de la BIOS/UEFI, habilitando parámetros clave como las tecnologías de virtualización (Intel VT-x/AMD-V) y IOMMU (VT-d) para la virtualización de dispositivos.
 - Configuración de los módulos de gestión remota (BMC/iLO/iDRAC) con los datos de red y credenciales de acceso que indique el organismo.
 - Instalación del sistema operativo base anfitrión (distribución Linux acordada con el organismo) junto con todas las dependencias necesarias para su funcionamiento.
 - Instalación y configuración de los controladores (drivers) de las GPU NVIDIA en el ámbito anfitrión.
- Especificaciones del cableado:
 - Fibra óptica:
 - Puertos de datos: cableado de color azul de 3 metros.
 - Puertos de servicio: cableado de color verde de 3 metros.
 - Cableado de cobre: deberá ser cables PATCHSEE de 3 metros.
 - Etiquetado: Tanto el cableado de corriente como el de red deberán estar claramente etiquetados en ambos extremos, indicando el punto de origen y el punto de destino.

Una vez los equipos se vayan instalando en los bastidores (sin estar en funcionamiento), el adjudicatario deberá retirar todo el embalaje asociado al equipamiento instalado de las instalaciones del CSUC ese mismo día.

El adjudicatario deberá proporcionar al personal técnico del CSUC la documentación necesaria para la correcta administración, configuración y explotación de la infraestructura suministrada. Esta documentación deberá entregarse en formato digital e incluirá, como mínimo, los siguientes elementos:

- Manual de administración de los equipos, que servirá como guía para las tareas de configuración, explotación y soporte posterior. Este manual deberá incluir:
 - Descripción general del sistema, de sus componentes y de la arquitectura global.
 - Esquemas gráficos detallados de la distribución física y la interconexión de los diferentes componentes.
 - Descripción de los principales parámetros utilizados para la configuración del sistema.
 - Procedimientos básicos de administración y explotación, incluyendo:
 - Puesta en marcha y apagado seguro de todos los componentes, incluyendo el apagado eléctrico.
 - Creación y gestión de archivos de diagnóstico en caso de caída del sistema.
 - Gestión de recursos virtuales y asignación de GPUs a los entornos virtualizados.
 - Configuración y uso de las funcionalidades de visualización remota con soporte GPU.
 - Recomendaciones para la monitorización del rendimiento y la gestión de cargas.
 - Procedimientos de mantenimiento preventivo y actualización de firmware y drivers.
 - Resolución de incidencias habituales y buenas prácticas de uso eficiente de las GPUs en entornos de cálculo científico.

El manual deberá entregarse en catalán, en formato PDF, y podrá incluir anexos técnicos, esquemas, capturas de pantalla y enlaces a documentación oficial del fabricante.

5. Garantía de soporte

Los equipos ofertados deberán incluir un período mínimo de **garantía de soporte de cinco años** para todo el hardware por parte del fabricante y el software suministrado.

El nivel de servicio no podrá ser inferior al siguiente:

- **Incidencia de prioridad alta:** cuando, a causa de la misma, se produzca una afectación total o una reducción del rendimiento estándar en un 15 % de los servicios ofrecidos por el CSUC. El tiempo de respuesta desde su inicio debe ser como máximo de 2 horas laborables y el tiempo de resolución como máximo de 8 horas laborables.
- **Incidencia de prioridad media:** cuando los logs o mensajes de error de los equipos indiquen que se ha producido una alteración no esperada de su comportamiento o bien cuando se haya producido algún error que deba corregirse. El tiempo de respuesta desde su inicio debe ser como máximo de 8 horas laborables y el tiempo de resolución como máximo de 24 horas laborables.

- **Incidencia de prioridad baja:** cuando sea informativa. El tiempo de respuesta desde su inicio debe ser como máximo el siguiente día laborable y el tiempo de resolución como máximo de 5 días laborables.

En caso de producirse una incidencia de prioridad alta, el CSUC podrá solicitar al adjudicatario informes puntuales sobre incidencias concretas, que deberán estar disponibles en un plazo máximo de 3 días laborables.

El incumplimiento de estos plazos se penalizará según lo dispuesto en el apartado O del Cuadro de características del Pliego de Cláusulas Administrativas Particulares.

Entendemos como tiempo de atención el tiempo desde que se notifica una incidencia hasta el inicio de las tareas necesarias para su resolución, y como tiempo de resolución el tiempo invertido en su resolución desde el momento en que se atiende hasta que realmente queda resuelta.

Adicionalmente, también se requiere:

- Cobertura horaria 24x7x365.
- Soporte telefónico ante incidencias y consultas con respuesta inmediata.
- Tareas preventivas: todas aquellas prestaciones necesarias para mantener los equipos solicitados en unos niveles de disponibilidad y funcionamiento óptimos, así como el ajuste de los elementos que permitan mantener dichos niveles. Entre otras, deberán contemplarse las operaciones de comprobación, diagnóstico y seguimiento de las prestaciones de los equipos, según las especificaciones del fabricante.
- Gestión proactiva por parte del licitador de la solución propuesta para la resolución de posibles incidencias de hardware con desplazamiento al CPD.
- Revisión y actualización (si es necesario y/o por motivos de seguridad) del sistema operativo, firmware, mejoras de ingeniería, etc., que se realizará de acuerdo con la planificación establecida con el personal técnico del CSUC. Las revisiones del entorno y actuaciones asociadas se realizarán anualmente. Las actualizaciones por motivos de seguridad se realizarán tan pronto como esté disponible la nueva versión del sistema operativo o firmware que solucione el problema en la plataforma del fabricante.
- La empresa adjudicataria deberá proponer, dentro de las condiciones de garantía, unas tareas preventivas y correctivas del equipamiento suministrado en los términos que establece el pliego.
- Disponer de un sistema de seguimiento de incidencias y consultas mediante una base de conocimiento (preferiblemente consultable vía internet). En caso de que el licitador quiera proponer esta alternativa, deberá describir brevemente las funcionalidades principales de la misma.

Los licitadores deberán especificar en su oferta los perfiles de los técnicos que llevarán a cabo el servicio de soporte, así como su experiencia previa en instalaciones y servicios de soporte con características similares a las del objeto del presente pliego.

El licitador deberá describir en detalle los tipos de soporte disponibles y las condiciones de contratación de cada uno de ellos una vez finalizado el período de garantía, así como sus condiciones contractuales, que deberán quedar detalladas en la oferta económica del sobre C. Concretamente, será necesario especificar la política del licitador respecto a averías y reparaciones del equipamiento, la jornada laboral para incidencias y el tiempo de respuesta previsto para su resolución.

Cabe señalar que, durante el período de garantía, todos los costes de mantenimiento deben estar incluidos en el precio ofertado por el equipamiento.

La oferta deberá describir en detalle cómo se llevarán a cabo la instalación y las condiciones de garantía posteriores a la instalación.

Al finalizar la instalación, el proveedor deberá entregar un inventario completo de todo el equipamiento suministrado, que deberá incluir, como mínimo, la siguiente información para cada elemento:

- Descripción del equipamiento
- Número de serie
- Part number o código del fabricante
- Importe de compra (precio unitario)

Los elementos comunes de la instalación (como cableado, conectores, accesorios menores, etc.) podrán agruparse bajo conceptos genéricos, indicando siempre la cantidad y el coste total asociado.

Este inventario es necesario para que todo el material quede correctamente inventariado y contabilizado financieramente, y se considera un requisito imprescindible para el cierre del proyecto.

6. Prestaciones que optimizan el rendimiento de los equipos

Los licitadores podrán ofrecer prestaciones que optimicen el rendimiento de los equipos y que deberán incorporarse en el sobre que corresponda en cada caso. Estas son:

- Número de servidores. (Sobre B)
- Modelo del procesador. (Sobre B)

- Incrementar la memoria RAM por servidor. (Sobre B)
- Garantías de respeto por el medio ambiente. (Sobre B)

7. Duración del proyecto y plazos de entrega

Los licitadores presentarán una propuesta de cronograma, para cada lote, que distinguirá claramente cada una de las fases previstas para la ejecución del proyecto en lo que respecta a la entrega y puesta en marcha de todo el hardware y el software, que incluirá como mínimo los siguientes hitos:

- Revisión de las instalaciones técnicas.
- Entrega de los equipos.
- Instalación y configuración del sistema.
- Plan de aceptación y pruebas del equipamiento.
- Plan de formación, si procede.

El equipamiento deberá estar instalado y en producción como máximo el 15 de mayo de 2026.