

PLEC DE PRESCRIPCIONS TÈCNIQUES

1. Objecte

L'objecte del present plec de prescripcions tècniques és definir les condicions tècniques que regiran el **subministrament d'equipament en l'àmbit de càlcul d'altres prestacions (HPC) i emmagatzematge massiu de dades científic-tecnològic destinat a un servei comú d'investigació de la Universitat Pompeu Fabra (en endavant, UPF)**.

El subministrament objecte d'aquesta licitació es divideix en tres lots, d'acord amb les necessitats del servei comú d'investigació de la Universitat Pompeu Fabra (UPF) en l'àmbit del càlcul d'altres prestacions (HPC) i l'emmagatzematge massiu de dades.

Els lots són els següents:

- Lot 1: Ampliació de la cabina d'emmagatzematge de l'entorn de HPC per a la UPF.
- Lot 2: Nodes de càlcul basats en tecnologia GPU en una arquitectura tradicional de HPC per a la UPF.
- Lot 3: Nodes de càlcul basats en tecnologia CPU en una arquitectura tradicional de HPC per a la UPF.

Les adquisicions del lot 1 i 2 venen finançades en el 50 % de l'Ajut en Infraestructura convocatòria 2024 del "Subprograma Estatal de Infraestructuras y de Equipamiento Científico-Técnico [objetivo específico 7 del Plan Estatal (OE7), destinado a desarrollar, mantener e invertir en infraestructuras científicas y tecnológicas que contribuyan a crear un ecosistema de I+D+I internacionalmente competitivo], del Ministerio de Ciencia, Innovación y Universidades" i la resta per recursos propis de la Universitat.

La solució proposada ha de disposar d'un certificat del fabricant que acrediti la capacitat de l'empresa licitadora en relació amb els seus productes tant en la comercialització com en la instal·lació.

Tot el hardware ofert serà nou (en cap cas refurbished).

El subministrament comporta el **lliurament, instal·lació i configuració fins a deixar els equips plenament operatius** incloent-hi la formació necessària al personal tècnic de la UPF.

2. Requeriments tècnics:

Les especificacions tècniques mínimes per a cada lot es presenten en **taules estructurades** amb un sistema d'identificació segons el següent criteri:

- **R (Requeriment):** Requeriment tècnic **obligatori**. El seu incompliment suposarà la **desqualificació automàtica** de l'oferta.
- **D (Desitjable):** Requeriment tècnic **no obligatori**, però **valorat positivament** en la fase de valoració tècnica de les ofertes.

A continuació es detalla l'objecte i finalitat de cada lot:

2.1 Lot 1: Ampliació de la cabina d'emmagatzematge

Aquest lot té per objecte el subministrament, la instal·lació i la posada en funcionament d'una ampliació de la infraestructura d'emmagatzematge científica de la Universitat Pompeu Fabra (UPF), destinada a donar suport als serveis d'investigació i als fluxos de treball intensius en dades.

L'ampliació s'haurà d'integrar tècnicament amb la infraestructura existent, garantint-ne la compatibilitat, l'operativitat i la continuïtat del servei. Aquest sistema ha d'incloure:

- El maquinari necessari
- Les llicències del programari nou
- L'actualització de les llicències ja existents, si escau.
- Suport i actualització del programari.
- Cablejat i interconnexió dels nous sistemes amb els ja existents.
- Instal·lació, posada en marxa i formació del personal tècnic.

2.1.1 Descripció del sistema d'emmagatzematge actual:

El sistema d'emmagatzematge actual va ser adquirit en el marc de l'expedient [UPF-2022-0059](#) i constitueix una plataforma dissenyada per allotjar grans volums de dades associades a projectes científics d'alta intensitat computacional.

Aquesta infraestructura dona servei a diversos entorns de càlcul i gestió de dades, incloent-hi el sistema HPC SCAR. Qualsevol ampliació haurà de **garantir la plena integració i interoperabilitat amb l'arquitectura existent**.

El sistema d'emmagatzematge del clúster SCAR (que respon a l'acrònim Servei de Càlcul d'Alt Rendiment) es basa en una arquitectura de dos nivells (tiers):

- **Tier 1:** Integrat per una cabina d'alt rendiment (ESS 3500) que ofereix 1 PB d'espai útil en discos SAS HDD, amb les metadades hostatjades en discos NVMe per garantir un accés ràpid i eficient.
- **Tier 2:** Destinat al tractament de dades fredes (*cold data*), integrat amb la cabina principal i amb una capacitat addicional d'1 PB. Aquest segon nivell està basat en les especificacions de l'expedient [UPF-2025-0019](#).

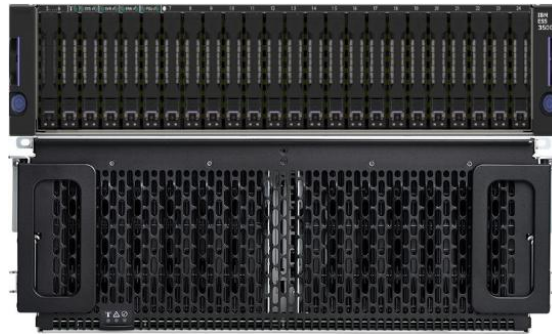
Ambdós tiers estan connectats a una xarxa de baixa latència. D'acord amb les polítiques d'optimització, ús i accés a les dades, el sistema permet la reubicació transparent de la informació entre els diferents nivells d'emmagatzematge, sense afectar l'experiència de l'usuari.

La solució utilitza com a sistema de fitxers el **IBM Spectrum Scale (GPFS)**, oferint una única infraestructura d'emmagatzematge per a tot l'entorn HPC, amb els avantatges següents:

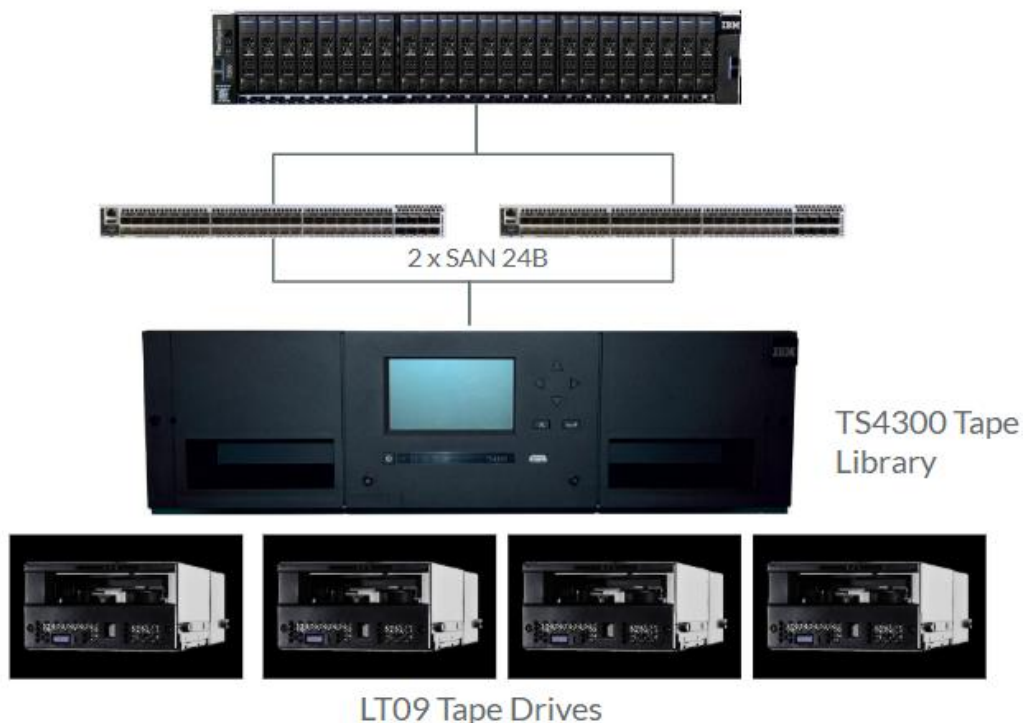
- **Escalabilitat:** Permet el creixement progressiu per donar resposta a futures necessitats sense comprometre el rendiment.
- **Eficiència:** Optimitza els temps d'accés a les dades mitjançant la combinació de diferents tecnologies de discos.
- **Seguretat:** Inclou funcionalitats d'auditoria, traçabilitat i registre d'accés als fitxers.
- **Integració:** Garanteix la interoperabilitat amb la infraestructura HPC actual i amb altres sistemes associats.

Detallem cadascuna de les parts.

Tier 1: Cabina d'alt rendiment



- Cabina d'alt rendiment (ESS 3500) amb 12 (de 24) discos NVMe de 15.36 TB, on s'hostatgen les metadades.
- Safata d'expansió (4U102) amb 102 (de 102) discos HDD de 14TB, on s'hostatgen les dades.
- Té quatre connexions de baixa latència amb el switch infiniband descrit a l'Annex1.#1.
- Té 2 connexions amb el switch S4128F-ON, Annex1.#2.



Un **cold data** (o dades fredes) fa referència a informació que s'accedeix de manera poc freqüent, però que cal conservar durant llargs períodes de temps per motius legals, històrics i/o de recerca. Es prioritza la capacitat a la velocitat d'accés a la dada.

Tècnicament, un servidor amb discos ràpids (flash) actua com a memòria cau (*cache*) abans que la informació sigui transferida als calaixos de cintes (amb cintes estàndard LTO9 de 18 TB de tamany). La dada es manté

disponible mentre no es requereixi l'espai que ocupa, independentment de si ha estat moguda o no al sistema de cintes.

Aquest sistema disposa de quatre unitats de cinta que poden operar en paral·lel, assolint un rendiment agregat de fins a 1,2 GB/s en mode natiu (4,3 TB/h) i fins a 3 GB/s amb compressió (10,8 TB/h).

2.1.2 Objectiu:

Incrementar la capacitat d'emmagatzematge del sistema existent amb un mínim de **2 PB nets addicionals**, assegurant compatibilitat i integració amb la infraestructura actual.

Presentem les especificacions tècniques mínimes que el lot ha de complir en forma de:

- *Solució d'emmagatzematge*. Taula estructurada amb sistema d'identificació de la referència segons sigui R (requerit) o D(desitjable).
- *taula 1: descripció de la solució d'emmagatzematge* amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Manteniment / Suport / Interlocució*:Taula estructurada amb sistema d'identificació de la referència segons sigui R (requerit) o D(desitjable).

a. Solució d'emmagatzematge

Ref.	Descripció
R1	Augmentar la capacitat de la cabina d'alt rendiment (ESS 3500) en 2 PB útils*. * capacitat útil és la que queda disponible després de restar els sistemes de redundància i l'espai reservat per spare
R2	Afegir un mínim de dos unitats d'expansió compatibles amb el sistema existent (ESS 3500 i 4U102) per satisfer R1.
R3	- Cada unitat d'expansió ha d'estar equipada amb discos SAS HDD de mínim 14 TB cadascun. - Els discos han d'ocupar tot l'espai disponible de la unitat d'expansió, assegurant així la màxima utilització de l'equipament.
R4	Cal garantir la integració perfecta amb la infraestructura existent i mantenint els estàndards de rendiment i fiabilitat requerits.
R5	Es requereixen llicències de programari que permetin la gestió i l'operació eficient de les dades en el sistema d'emmagatzematge expandit. Aquestes llicències han de ser plenament compatibles amb les ja recollides a R7, Spectrum Data Management, garantint la integració total i el funcionament coherent del sistema.
R6	- Les quatre boques infiniband de les que disposa la cabina d'alt rendiment (ESS 3500) ja estan ocupades i connectades al switch Infiniband descrit a l'Annex 1.#1 - Cal afegir 4 noves connexions, amb les òptiques i fibres adients per arribar al segon switch infiniband. - El nou switch es troba al passadís contigu a una distància superior als 5 metres i inferior als 10 metres.
R7	Per a garantir la compatibilitat amb els productes ja existents , és necessari incorporar les llicències per a la gestió de dades dissenyada per a entorns complexos i d'alt volum, la suite Spectrum Data Management , desenvolupada per IBM i actualment integrada dins del nostre entorn de còmput.
D1	Es valorarà incloure en la solució 12 nous discos NVMe, omplint així la cabina ESS 3500 fins al màxim (24 discos), de la mateixa tipologia dels existents, 15.36 GB, i oferint un espai NVMe addicional de més de 95 TB útils.
D2	- Es valorarà afegir mòduls d'expansió sense drives amb 40 cintes estàndard LTO9 de 18 TB al tier 2. per

	incrementar l'espai d'arxiu on emmagatzemar la dada freda. - Es podran afegir fins a tres mòduls d'expansió, cadascú amb les cintes comentades al punt anterior.
R8	No cal incloure cap Rack. La solució s'integrarà en els racks existents.
R9	En el moment de presentar l'oferta els licitadors han d'omplir la <i>taula 1: descripció hardware de la solució d'emmagatzematge</i> , on s'especifiquen els valor mínims que cal complir i on s'indican els valors ofertats.

taula 1: descripció hardware de la solució d'emmagatzematge

Concepte	Valor Mínim	Valor Ofertat
característiques de la solució		
Tamany del disc SAS HDD	14 TB	
Nombre de safates d'expansió	2	
Espai total útil afegit	2 PB	
Nombre de nous ports Infiniband	4	
Mínima connectivitat d'òptica	QSFP56 (200 Gbps)	
Llargària de la fibra	10 metres	

b. Manteniment / Suport / Interlocució

Ref.	Descripció
R1	Garantia legal: Tant els nous elements afegits com els ja existents que ho requereixin quedaran sota la cobertura d'un manteniment/garantia de maquinari/programari de 5 anys.
R2	- En aquest termini l'adjudicatari serà responsable de les faltes de conformitat que existeixin en el moment del lliurament i es manifestin dins del termini. - L'empresa adjudicatària haurà de garantir la disponibilitat de peces de recanvi originals o equivalents (directament o a través d'altres agents autoritzats) per tot el cicle de vida previst de l'equipament, que com a mínim serà de 10 anys a partir de que els equips adquirits es deixin de fabricar.
R3	El manteniment/garantia començarà a partir de la data de posada en marxa de l'equipament, prèvia acceptació per escrit de la Universitat Pompeu Fabra.
R4	Garantia addicional obligatòria: - L'adjudicatari està obligat a prestar en un termini mínim de 5 anys una garantia comercial addicional que compren els treballs de manteniment que garanteixin el correcte funcionament de l'equipament durant el termini de garantia comercial i considerats per mal funcionament que no sigui per un mal ús, incloent els recanvis, mà d'obra, desplaçaments, transport i dietes. - El recanvi de qualsevol element avariats es realitzarà sempre utilitzant material original o certificat pel mateix fabricant.
R5	Revisions de firmware: - L'adjudicatari haurà de recomanar els nivells de firmware necessaris de l'equipament hardware i deixar-ho en aquestes versions. - Durant el període de manteniment, l'adjudicatari ha de proporcionar suport per a l'actualització del firmware de tots els components del sistema, amb una freqüència mínima d'una vegada a l'any i sempre que sigui necessari per garantir la seguretat de l'equipament subministrat.
R6	Durant el període de garantia, el licitador estarà obligat a proporcionar el següent suport de software o hardware (SLA)

	- Suport remot 24x7. Temps de resposta en remot: 1 hora. - Suport in situ 24x7. Temps de resposta in situ: 8 hores, que és l'estàndard de l'ENS de la UPF.
R7	En aquells casos en què l'actuació ho requereixi, serà necessari comptar amb el suport oficial de tercer nivell del fabricant per garantir la correcta resolució de les incidències o actualitzacions.
R8	Posada en marxa: - La instal·lació i posada en marxa serà on-site sota la supervisió de l'equip UPF. En posteriors iteracions es permetrà la connexió externa a l'entorn. - L'equip tècnic encarregat de la instal·lació hardware haurà de disposar de la formació i les capacitats tècniques per a la realització d'aquest tipus d'instal·lacions. - Una vegada acabada la instal·lació física s'haurà de comprovar que es compleixen els requeriments operatius establerts en aquest plec, com redundància elèctrica i correcta repartició a les fases elèctriques, adequació de la infraestructura, gestió correcta del flux d'aire, cablejat òptim, ben pentinat i correctament identificat, entre d'altres.
R9	Interlocució: - S'ha de proveir un sol punt d'avís i resolució d'incidències de qualsevol component que formi part de la solució durant el cicle de vida del projecte. - L'empresa adjudicatària haurà de disposar dos interlocutors: un de comercial i un altre de tècnic.
R10	En el moment de presentar l'oferta els licitadors han de presentar un document digital que reculli: - Descripció general dels components de la solució. - Esquemes físics de les connexions que es realitzaran. - Versions de firmware que s'instal·larà. - Planificació del projecte: Amb la seqüència de fases del projecte, les activitats que es duren a terme per a la posada en marxa del lot i el corresponent cronograma temporal. Així mateix, la memòria ha d'explicitar els diferents blocs del projecte, les accions concretes associades a cada fase i els mecanismes de comunicació i coordinació establerts entre l'equip tècnic del licitador i la Universitat Pompeu Fabra.
R11	Formació: - Es farà una transferència de coneixements sobre el funcionament i actualització de la infraestructura, la seva configuració i les principals tasques de manteniment i gestió. Serà presencial en almenys dos sessions amb una durada estimada total de 8 hores i per a uns 8 tècnics informàtics de la Universitat. - El material de la formació es lliurarà al final de les jornades.
R12	- S'ha de garantir que l'empresa licitadora disposa d'un nivell d'acreditació preferent platinum o equivalent dins del programa del fabricant.
D1	Es valorarà aportar un o dos perfils tècnics amb certificació d'especialista per part del fabricant amb almenys un any d'experiència en projectes similars. Actuaran tant en la fase d'instal·lació i formació, com durant el període de garantia del producte per a les tasques de manteniment i suport tècnic que es requereixin.

2.2 Lot 2: Nodes de càlcul basats en tecnologia GPU

Aquest lot té per objecte el **subministrament, instal·lació i posada en funcionament de nodes de càlcul amb tecnologia GPU**, que s'integraran en el clúster d'alt rendiment **SCAR** de la Universitat Pompeu Fabra (UPF), amb l'objectiu de reforçar els serveis d'investigació que requereixen una alta capacitat de càlcul i processament paral·lel de dades.

Els nous nodes formaran part d'un entorn **HPC clàssic**, amb suport per a la **paral·lelització de tasques**, **interconnexió de baixa latència** i **alt ample de banda**, accés a **emmagatzematge compartit** mitjançant un

sistema de fitxers paral·lel (**GPFS**), i gestió de recursos amb **Slurm** (que respon a l'acrònim de Simple Linux for Resource Management), que és un planificador de treballs de codi obert i d'ús majoritari en entorns de còmput d'alt rendiment (HPC). La integració amb la infraestructura actual haurà de ser total, garantint **compatibilitat, operativitat i continuïtat del servei**.

Segons consta a l'Annex 1: Elements hardware, l'actual commutador de baixa latència disposa de 4 ports de connexió lliures, insuficients per connectar els **mínim 10 nous nodes GPU** previstos. Per aquest motiu, aquest lot inclou també el **subministrament d'un nou commutador Infiniband de baixa latència**, amb **capacitat suficient per interconnectar tots els nodes nous amb la cabina d'emmagatzematge d'alt rendiment** (descrita al Lot 1) i amb la resta de la infraestructura HPC.

Sense aquest nou commutador, **no és possible posar en funcionament els nous nodes**, atès que no es podria garantir una connectivitat adequada ni complir amb els requisits de rendiment, latència i escalabilitat del sistema.

La solució haurà d'incloure, com a mínim:

- El **maquinari necessari** per als nodes GPU.
- Les **llicències de programari**, si escau.
- El **cablejat i interconnexió** amb la infraestructura existent.
- El **commutador Infiniband de baixa latència** amb ports suficients.
- La **formació tècnica** per al personal responsable de la seva gestió.

Presentem les especificacions tècniques mínimes que el lot ha de complir en forma de:

- *Nodes de càlcul GPU*: Taula estructurada amb sistema d'identificació de la referència segons sigui R (requerit) o D(desitjable).
- *taula 2 - Descripció dels nodes de càlcul GPU*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Commutador de baixa latència*: Taula estructurada amb sistema d'identificació de la referència segons sigui R (requerit) o D(desitjable).
- *Taula 3 - Descripció dels commutadors de baixa latència*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Operacional: Descripció dels requeriments operacionals relacionats amb els components del clúster*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Software: Compatibilitats de clúster i nodes*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Manteniment*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Taula 4 - Manteniments i garanties*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.

a. Nodes de càlcul GPU:

Ref.	Descripció
R1	Proveir un mínim de 24 targetes GPU del mateix model. S'hauran de presentar instal·lades en nodes GPU que descriurem en el punt següent.
R2	Es consideraran tres tipus de nodes de càlcul: - node GPU: node dimensionat amb les màximes GPUs que pot hostatjar. - node GPU al 50%*: exactament el mateix node (mateixa memòria, mateixes fonts, mateix discos,

	mateixa CPU) del punt anterior, però amb la meitat de les GPUs que pot admetre. - node de login *Els nodes GPUs al 50% s'ampliaran amb el mateix model de GPU en una fase posterior i fora d'aquesta licitació. Per tant, no poden existir restriccions per part del/s fabricant/s per afegir noves GPUs en el/s slot/s disponible/s a tal efecte. La solució actual així mateix ha de garantir que en cas que s'afegeixin noves GPU no es perd la garantia de la solució aquí licitada, per tant el hardware d'aquesta licitació ha de permetre el manteniment de la garantia de fabricant amb la incorporació de GPU posteriorment.
R3	En cas que els nodes estiguin agrupats en un bastidor: - La interfície de gestió a la xarxa de consoles (out-of-band) podrà ser compartida per tots els nodes del clúster. - En cas de tenir un commutador intern per alguna de les altres xarxes (gestió o baixa latència) haurà de seguir complint el nombre d'interfícies per node.
R4	Nodes: - Els busos dels sockets han de proporcionar l'ample de banda màxim admès per la família de processadors. - Els processadors han de ser idèntics i els busos d'interconnexió entre sockets simètrics entre si. - El node ha de ser tipus Rackmount, és a dir, compatible amb rack estàndard.
R5	Interfícies de xarxa: - Interfície per a connexió a xarxa de consoles (out-of-band) per administració remota. Ha de permetre com a mínim fer power on/off, accedir a la consola, monitoritzar l'entorn (temperatura, consums...), detectar alertes, identificar avaries hardware, control de versions de firmware... En cas que realitzar aquestes accions calgui llicència, haurà d'incorporar-la amb validesa perpètua i sense costos addicionals futurs. Connexió RJ45. - Interfície per a connexió a una xarxa de baixa latència, compatible amb l'equipament descrit a l'apartat Annex 1.#1. - Interfície per a connexió a xarxa de servei del clúster, compatible amb l'equipament descrit a l'apartat Annex 1.#2
R6	Fonts d'alimentació: - Els nodes han de tenir fonts plenament redundants. Les fonts d'alimentació han de suportar la càrrega pic de totes les GPUs. En cas d'avaría d'una font, el servidor haurà de seguir funcionant en el supòsit de tenir totes les GPUs treballant al 100%. - El node GPU al 50% haurà de tenir les mateixes fonts d'alimentació que les del node GPU. - El node login ha de disposar de fonts d'alimentació redundants i dimensionades per l'equip. Totes les fonts han d'estar certificades per a alta eficiència energètica (Titanium, equivalent o superior)
R7	CPU: - Per compatibilitat amb l'actual clúster, els processadors han de ser Intel x86-64 amb compatibilitat amb les instruccions de processador AVX-512. - No seran vàlides les refrigeracions líquides.
R8	Disc: - Mínim 1 disc dur pel sistema operatiu. - Mínim 1 disc dur ràpid per a dades locals.
R9	GPU: - Cal que tinguin tecnologia CUDA per compatibilitat amb el clúster actual i per tant basat en arquitectura Nvidia.
R10	Memòria: - Els DIMMs de memòria han de ser de la mateixa velocitat i tamany entre si. - La configuració de memòria presentada haurà d'estar equilibrada des de/cap a tots els cores d'un mateix

	socket a la memòria - La freqüència d'accés a memòria haurà de ser la més alta que la família de processadors ofertats permet. - El mínim de memòria del node GPU respectarà la ratio 256 GB de RAM per cada GPU. - El mínim de memòria del node GPU al 50% respectarà la ratio 512 GB de RAM per cada GPU. - El mínim de memòria del node de login serà de 256 GB de RAM
R11	Node de login: - Ha de complir R4, R5, R6, R7, R8, R10 d'aquest apartat - Ha de disposar d'una interfície extra d'accés a la xarxa de servei per a configurar un LACP.
R12	En el moment de presentar l'oferta els licitadors han d'omplir la <i>Taula 2 - Descripció dels nodes de càlcul GPU</i>, on s'especifiquen els valors mínims de la proposta i els ofertats

Taula 2 - Descripció dels nodes de càlcul GPU

Concepte	Valor Mínim	Valor Ofertat
característiques de la targeta GPU		
Model de la GPU:	—	
Mínima memòria GPU (GB)	48	
Màxim consum elèctric (W)	350	
cores CUDA	18.000	
Memory Bandwidth (GB/s)	864	
Thermal (tipus de ventilació)	Passiva	
FP32 TFLOPS (Rendiment en precisió simple)	91	
característiques comunes a TOTS els nodes (login, node gpu i node gpu al 50%)		
Nombre de xips o sockets per node	2	
Nombre de processadors per cada socket	1	
Model de processador	—	
Velocitat de transmissió de dades entre processadors dins un mateix node (GT/s o Gb/s)	—	
Cores independents per processador (sense Hyper-threading)	32	
Velocitat de rellotge base (freqüència nominal de cada core)	2,0 GHz	
Estàndard de memòria	DDR5	
Freqüència de funcionament de memòria	—	
Capacitat d'emmagatzematge local per a Sistema Operatiu (GB)	480	
Tecnologia d'emmagatzematge local per a Sistema Operatiu	SSD	
Capacitat d'emmagatzematge local per a dades (Tb)	1	
Tecnologia d'emmagatzematge local per a dades	NVMe	
Interfícies 1Gb	1	
Interfícies de baixa latència	1	
Ample de banda a xarxa de baixa latència (Gbps)	200	
node GPU		
Alçada mínima (per node)	2U	
Interfícies Ethernet 10/25 SFP28 (per node)	2	
Transceivers del node	SFP+	
Nombre de DIMMs (per node)	—	
Tamany del DIMM	—	
TOTAL de memòria RAM (per node)	1.024 GB	
TOTAL de número de nodes	2	

TOTAL de GPUs (per node)	4	
node GPU al 50%		
Alçada mínima (per node)	2U	
Interfícies Ethernet 10/25 SFP28 (per node)	2	
Transceivers del node	SFP+	
Nombre de DIMMs (per node)	—	
Tamany del DIMM	—	
TOTAL de memòria RAM (per node)	1.024 GB	
TOTAL número de nodes	8	
TOTAL GPUs (per node)	2	
node login		
Alçada màxima (per node)	1U	
Interfícies Ethernet 10/25 SFP28 (per node)	2	
Transceivers del node	SFP+	
Nombre de DIMMs (per node)	—	
Tamany del DIMM	—	
TOTAL de memòria RAM (per node)	256 GB	
TOTAL número de nodes	1	

b. Commutador de baixa latència

Ref.	Descripció
R1	El switch InfiniBand haurà de ser compatible amb el model descrit a l'Annex 1-#1. La finalitat d'aquest nou switch serà: - Suportar les connexions de tots els nodes d'aquesta lot. - Dotar a la cabina de baixa latència, alta disponibilitat infiniband, ja establerta en el lot 1. - Agregar els dos switchos infiniband per oferir major ample de banda, redundància i balanceig de trànsit (ECMP / multipathing).
R2	La velocitat de connexió mínima amb la que negociaran els nodes de càlcul serà la marcada pel switch existent, HDR 200Gb/s
R3	Ha de ser compatible amb velocitats anteriors (Backward compatible)
R4	Connectivitat entre switchos: - Anirà ubicat a Passadís 2. - Cal el cablejat necessari per interconnectar amb el switch descrit a Annex 1-#1. - La distància entre passadissos és inferior a 10 metres i superior a 5 metres. - Cal aconseguir 400 Gbps de connexió agregada entre switchos, emprant dos ports i dos línies punt a punt, o dos de 200 Gbps sense agregació en cas que no sigui possible.
R5	El Centre de Processament de Dades (CPD) està organitzat en mode CAC (Cold Aisle Containment), és a dir, tancant el passadís fred i mantenint l'aire fred dins d'aquesta àrea. Fet aquest aclariment, els ports (boques) de connexió del switch infiniband hauran d'estar orientades a la part de darrera del rack, zona calenta, i mantenir la mínima distància Infiniband entre servidor i switch.
R6	La ventilació ha de ser C2P (cooler-to-port), fent entrar l'aire fred per la part trasera del switch i sortir calent pel lloc dels ports
R7	En el moment de presentar l'oferta els licitadors han d'omplir la <i>Taula 3 - Descripció dels commutador</i>

	de baixa latència, on s'especifiquen els valors mínims de la proposta i els oferts
--	--

Taula 3 - Descripció dels commutador de baixa latència

Concepte	Valor Mínim	Valor Ofertat
característiques de la targeta GPU		
Permet stackat amb model QM8790?	SI	
Backward compatible	SI	
Mínima velocitat dels ports (sense fer servir splitter cables)	200 Gbps	
Nombre de Non-blocking ports (boques) disponibles	40	
Mínim nombre d'enllaços entre switchos.	2	

c. Operacional: Descripció dels requeriments operacionals relacionats amb els components del clúster.

Ref.	Descripció
R1	No cal incloure cap Rack. La solució s'integrarà en els racks existents. Són racks de 19" estàndard de 42U i es disposarà de dos armaris / rack lliures per a equilibrar la solució.
R2	No cal incloure cap PDU. El licitador informará de l'alimentació elèctrica emprada (C13, C19 Schuko...) pel servidor, ho detallarà a l'informe quantes boques elèctriques i connector cal a cada PDU.
R3	En el moment de presentar l'oferta els licitadors han de presentar un esquema frontal amb l'ocupació dels racks dels diversos equipaments presentats a la solució. Es detallarà el hardware ofertat amb les Us que ocupa cada component de la solució. S'afegirà el detall del cables / fibres que surten de cada rack per a cadascuna de les xarxes definides.
R4	Tots els nodes de còmput i la resta de components de la solució han de disposar de fonts d'alimentació redundants N+N. Que es connectaran a les PDUs corresponents redundades i a la vegada a quadres elèctrics diferents.
R5	Tant al passadís 1 com al passadís 2, els commutadors Infiniband i els de xarxa gestió i serveis estan físicament al mateix rack amb màxim 2U de separació entre ells.
R6	Cablejat - La solució de nodes es distribueix físicament en dos racks independents, amb una disposició equitativa del nombre de nodes entre ambdós racks. - Tot el cablejat dels nodes allotjats al rack més proper al rack dels commutadors serà de mínim 3m - Tot el cablejat dels nodes allotjats al rack més distant del rack dels commutadors serà de mínim 5m - Cal incloure les òptiques i els transceptors adients per aquestes interconnexions, tant en xarxa de baixa latència com en xarxa de servei. La correcta selecció i instal·lació dels cables i transceivers serà responsabilitat de l'adjudicatari, el qual haurà de garantir l'adequació tècnica de tots els components per assolir les prestacions requerides pel sistema.
R7	Etiquetat: - S'hauran d'incloure totes les tasques de cablejat i etiquetat dels components segons el disseny acordat entre la UPF i l'empresa licitadora i adaptat a les normes que fixi el finançament rebut. - La instal·lació del cablejat (fibra, elèctric, coure...) ha de quedar correctament "pentinat" i endreçat. - Tot component de la solució ha d'anar degudament etiquetat, per a ser identificat de forma única segons nomenclatura que s'estableixi entre la UPF i l'empresa instal·ladora.
R8	Els nodes de còmput hauran d'haver passat un burn-in test a fàbrica o a les instal·lacions del licitador per a evitar els DOA (Dead on Arrival).

R9	- La sala compta amb subministrament elèctric per a cada rack, mitjançant dues sortides redundants connectades a dos quadres elèctrics independents. Cada rack disposa de 2 preses CETAC de 32A amb corrent trifàsica. Atès que els nodes s'han de distribuir entre dos racks, es disposarà de 4 preses CETAC de 32A per garantir l'alimentació elèctrica necessària per al lot. - En cas que calgui afegir nous connectors o modificar els existents, aquests ajustos seran responsabilitat del licitador i no suposaran cap cost addicional per a la UPF.
R8	S'han d'incloure tots els elements i components necessaris pel transport i ubicació dels servidors a la sala CPD.

d. Software: Compatibilitats de clúster i nodes.

Ref.	Descripció
R1	El sistema operatiu és linux, Rocky Linux 8.7 en la redacció d'aquest plec, per tant tots els components hardware han de ser compatibles amb aquest sistema operatiu. Un cop formalitzat el contracte, el licitador contactarà amb l'equip UPF per conèixer la versió definitiva del sistema operatiu: si es manté en 8.7 o escala a la versió 9.x.
R2	Qualsevol element hardware (servidors de login, node GPU, node GPU al 50%) ha de suportar l'standard de IPMiv2.
R3	Qualsevol element hardware (servidors de login, node GPU, node GPU al 50%) ha de suportar el sistema de clustering xCAT.

e. Manteniments i garanties:

Ref.	Descripció
R1	Garantia legal: Tant els nous elements afegits com els ja existents que ho requereixin quedaran sota la cobertura d'un manteniment/garantia de maquinari/programari en el termini de 3 anys o el termini superior que ofereixi el licitador en la seva oferta.
R2	- En aquest termini l'adjudicatari serà responsable de les faltes de conformitat que existeixin en el moment del lliurament i es manifestin dins del termini. - L'empresa adjudicatària haurà de garantir la disponibilitat de peces de recanvi originals o equivalents (directament o a través d'altres agents autoritzats) per tot el cicle de vida previst de l'equipament, que com a mínim serà de 10 anys a partir de que els equips adquirits es deixin de fabricar.
R3	El manteniment/garantia començarà a partir de la data de posada en marxa de l'equipament, prèvia acceptació per escrit de la Universitat Pompeu Fabra.
R4	Garantia addicional obligatòria: - L'adjudicatari està obligat a prestar en el termini de 3 anys o el termini superior que ofereixi el licitador en la seva oferta, una garantia comercial addicional que compren els treballs de manteniment que garanteixin el correcte funcionament de l'equipament durant el termini de garantia comercial i considerats per mal funcionament que no sigui per un mal ús, incloent els recanvis, mà d'obra, desplaçaments, transport i dietes. - El recanvi de qualsevol element avariats es realitzarà sempre utilitzant material original o certificat pel mateix fabricant.
D1	Incrementar la garantia del servidor (node GPU) de 3 a 5 anys.
R5	Durant el període de garantia, el licitador estarà obligat a proporcionar el següent suport de software o

	hardware (SLA) - El temps de resposta a la notificació d'incidències haurà de ser màxim 4 hores. - L'horari serà 8x5 (8h al dia / 5 dies a la setmana)
R6	Firmware/ BIOS / IPMI: - L'empresa licitadora haurà de recomanar els nivells de firmware necessaris de l'equipament hardware i deixar-lo en aquestes versions. - S'haurà de presentar la parametrització del firmware / BIOS per atorgar el màxim rendiment dels nodes de còmput. - Els nodes tindran hyperthreading i VT deshabilitat a nivell de BIOS. - Configuració IP de les interfícies de gestió de tots els elements de la solució seguint les indicacions de l'equip UPF. - Configuració de la gestió remota a nivell de xarxa, d'autenticació i comunicació de la sessió, seguint les indicacions de l'equip UPF
R7	La instal·lació del sistema operatiu dels nodes de càlcul serà responsabilitat de l'equip UPF, però l'equip tècnic de l'adjudicatari s'haurà de connectar als equips en cas de mal funcionament o dubtes sobre la configuració tècnica (hardware/firmware).
R8	Interlocució: - S'ha de proveir un sol punt d'avís i resolució d'incidències de qualsevol component que formi part de la solució durant el cicle de vida del projecte. - L'empresa adjudicatària haurà de disposar dos interlocutors: un de comercial i un altre de tècnic.
R9	En el moment de presentar l'oferta els licitadors han de presentar un document digital que reculli: - Descripció general dels components de la solució. - Esquemes físics de les connexions que es realitzaran. - Versions de firmware i BIOS/firmware que s'instal·laran. - Planificació del projecte: Amb la seqüència de fases del projecte, les activitats que es duren a terme per a la posada en marxa dels servidors i el corresponent cronograma temporal. Així mateix, la memòria ha d'explicitar els diferents blocs del projecte, les accions concretes associades a cada fase i els mecanismes de comunicació i coordinació establerts entre l'equip tècnic del licitador i la Universitat Pompeu Fabra.
R10	Posada en marxa: - La instal·lació i posada en marxa serà on-site sota la supervisió de l'equip UPF. En posteriors iteracions es permetrà la connexió externa a l'entorn. - L'equip tècnic encarregat de la instal·lació hardware haurà de disposar de la formació i les capacitats tècniques per a la realització d'aquest tipus d'instal·lacions. - Una vegada acabada la instal·lació física s'haurà de comprovar que es compleixen els requeriments operatius establerts en aquest plec, com redundància elèctrica i correcta repartició a les fases elèctriques, adequació de la infraestructura, gestió correcta del flux d'aire, Cablejat òptim i ben pentinat,...
R11	Una vegada feta la instal·lació física de l'equipament, l'equip tècnic farà la transferència de coneixements respecte l'accés i la configuració tant d' IPMI com de les configuracions de consola via web i la forma futura de realitzar modificacions.
R12	Garanties i manteniments mínims. En el moment de presentar l'oferta els licitadors han d'omplir la <i>Taula 4 - Manteniments i garanties</i>, on s'especifiquen els valors mínims de la proposta i els oferts

Taula 4 - Manteniments i garanties:

Concepte	Valor Mínim	Valor Ofertat
anys de cobertura del manteniment pels elements del lot		
node GPU, GPU al 50% i login (anys)	3	

cadascuna de les unitats GPU (anys)	5	
commutador de baixa latència (anys)	5	

2.3 Lot 3: Nodes de càlcul basats en tecnologia CPU

Aquest lot té per objecte el subministrament, la instal·lació i la posada en funcionament de nous nodes de càlcul amb tecnologia exclusivament CPU, destinats a incrementar la potència de còmput seqüencial i paral·lel del clúster d'alt rendiment SCAR de la Universitat Pompeu Fabra (UPF). L'objectiu és assolir un mínim de rendiment agregat d'escala exaflop (10^{18} operacions en coma flotant per segon, precisió doble) per any que permeti afrontar simulacions numèriques, models de gran escala i fluxos de dades intensius que no depenen de l'acceleració GPU.

Els nodes s'integraran en el mateix entorn HPC clàssic existent, basat en Slurm per a la gestió de recursos, interconnexió de baixa latència i alt ample de banda Infiniband, i accés a emmagatzematge compartit mitjançant sistema de fitxers paral·lel (GPFS) i Slurm. La integració haurà de ser completa per tal de garantir la compatibilitat amb la infraestructura ja desplegada, l'operativitat continuada i la transparència per als usuaris finals.

Tots els nodes han de tenir les mateixes característiques tècniques que siguin descrites en la Taula 5.

Els nous nodes es connectaran directament al switch Infiniband descrit a l'*Annex 1.#1*, que disposa de ports reservats suficients; per tant, no cal subministrar cap nou commutador Infiniband.

La solució a licitar haurà d'incloure, com a mínim:

- Maquinari necessari per als nodes de càlcul basats en CPU.
- Llicències de programari i firmware, si escau.
- Cablejat i elements d'interconnexió compatibles amb la infraestructura existent.
- Formació tècnica per al personal responsable.

Presentem les especificacions tècniques mínimes que el lot ha de complir en forma de:

- *Nodes de càlcul CPU*: Taula estructurada amb sistema d'identificació de la referència dels requeriments.
- *Taula 5: Descripció dels nodes de càlcul CPU*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Operacional: Descripció dels requeriments operacionals relacionats amb els components del clúster*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Software: Compatibilitats de clúster i nodes*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Manteniments i garanties*: Descripció de la solució amb els valors mínims de la proposta i el recull ofert pel licitador.
- *Taula 6: Manteniments i garanties*: Valors mínims de la proposta i el recull ofert pel licitador.

a. Nodes de càlcul CPU:

Ref.	Descripció
R1	Proveir un mínim de 4890 exaFLOPS a l'any de càlcul en CPU seguint la fórmula descrita en l'annex II.

R2	En cas que els nodes estiguin agrupats en un xassís: - La interfície de gestió a la xarxa de consoles (out-of-band) podrà ser compartida per tots els nodes del clúster. - En cas de tenir un switch intern per alguna de les altres xarxes (gestió o baixa latència) haurà de seguir complint el nombre de interfícies per node.
R3	Intel Socket Bus: - Els busos que interconnecten els sockets d'un node hauran d'estar equilibrats i tenir el màxim ample de banda que la família de processadors ofertats permet.
R4	Rack: - El node ha de ser Rackmount (enracable en rack standard). - L' alçada màxima per node en cas de no utilitzar xassís es de 2U. - L' alçada màxima del xassís en cas d'utilitzar xassís és de 3U. - El número màxim de nodes per xassís en cas de xassís es de 4 nodes.
R5	Interfícies de xarxa: - Interfície per a connexió a xarxa de consoles (out-of-band) per administració remota. Ha de permetre com a mínim fer power on/off, accedir a la consola, monitoritzar l'entorn (temperatura, consums...), detectar alertes, identificar avaries hardware, control de versions de firmware... En cas que realitzar aquestes accions calgui llicència, haurà d'incorporar-la amb validesa perpètua i sense costos addicionals futurs. Connexió RJ45. - Interfície per a connexió a una xarxa de baixa latència, compatible amb l'equipament descrit a l'apartat Annex 1#1. - Interfície per a connexió a xarxa de servei del clúster, compatible amb l'equipament descrit a l'apartat Annex 1.#2
R6	Fonts d'alimentació: - Els nodes o xassís han de proporcionar fonts plenament redundants, les fonts d'alimentació han de suportar la càrrega de totes les CPUs del node funcionant al 100%. - En cas d'avaría d'una font, el servidor haurà de seguir funcionant en el supòsit de tenir totes les CPUs treballant al 100%. - En els casos de xassís amb més de dos fonts d'alimentació caldrà implementar una protecció intel·ligent per forçar automàticament a les CPUs del node a entrar en mode ULFM (mode de freqüència ultra baixa per un consum mínim d'energia) quan el controlador de gestió de hardware de la placa base rebí una alerta d'una fallida greu o errors d'una línia de corrent d'alimentació (com pèrdua d'energia, sobretensió, sobrecalentament, problema amb el ventilador, etc). Aquesta funció ha d'evitar en tots el casos a l'apagada dels sistemes. Totes les fonts han d'estar certificades per a alta eficiència energètica (Titanium, equivalent o superior)
R7	CPU: - Per compatibilitat amb l'actual clúster, els processadors han de ser Intel x86-64 amb les instruccions propietàries AMX, SSE4.2, AVX, AVX2 i AVX-512. Aquestes instruccions són imprescindibles ja que permeten una millor utilització dels recursos de càlcul, optimitzant els temps d'execució i millorant el rendiment de les aplicacions que es desenvolupen en entorns HPC, on els requisits de càlcul intensiu i la manipulació de grans volums de dades són habituals. - No seran vàlides les refrigeracions líquides. - Els processadors han de ser idèntics.
R8	Disc: - mínim 1 disc dur pel sistema operatiu.
R9	Memòria: - Els DIMMs de memòria han de ser idèntics, de la mateixa velocitat i tamany entre si. - La configuració de memòria presentada haurà d'estar balancejada des de/cap a tots els cores d'un

	mateix socket a la memòria - La freqüència d'accés a memòria haurà de ser la més alta que la família de processadors ofertats permet. - El mínim de memòria del node CPU respectarà la ratio 6 GB de RAM per cada nucli de la CPU. Es farà el càlcul en unitats de GB dividit per cores del node
R10	En el moment de presentar l'oferta els licitadors han d'omplir la <i>Taula 5 - Descripció dels nodes de càlcul CPU</i> on s'especifiquen els valors mínims de la proposta i els oferts

Taula 5 - Descripció dels nodes de càlcul CPU

Concepte	Valor Mínim	Valor Ofertat
Nombre de xips o sockets per node	2	
Nombre de processadors per cada socket	1	
Model de processador	—	
Velocitat de transmissió de dades entre processadors dins un mateix node (GT/s o Gb/s)	—	
Cores independents per processador (no esté en compte el Hyper-Threading)	86	
ExaFLOPS per any y node (pel càlcul veure l'annex II)	273,92	
Velocitat de rellotge base (freqüència nominal de cada core)	2,0 GHz	
Estàndar de memòria	DDR5	
Freqüència de funcionament de memòria	6400MHz	
Capacitat d'emmagatzematge local per a Sistema Operatiu (GB)	960GB	
Tecnologia d'emmagatzematge local per a Sistema Operatiu	NVMe	
Interfícies 1Gb	1	
Interfícies de baixa latència	1	
Ample de banda a xarxa de baixa latència (Gbps)	200	
Alçada mínima (per node en cas de NO xassís)	2U	
Interfícies Ethernet 10/25 SFP28 (per node)	2	
Transceivers del node	SFP+	
Nombre de DIMMs (per node)	—	
Tamany del DIMM	64GB	
Memòria RAM per nucli	6GB	
TOTAL de nodes	12	

b. Operacional: Descripció dels requeriments operacionals relacionats amb els components del clúster.

Ref.	Descripció
R1	No cal incloure cap Rack. La solució s'integrarà en els racks existents. Són racks de 19" estàndard de 42U i es disposarà de dos armaris / rack lliures per a equilibrar la solució.
R2	No cal incloure cap PDU. El licitador informarà de l'alimentació elèctrica emprada (C13, C19 Schuko...) pel servidor, ho detallarà a l'informe quantes boques elèctriques i connector cal a cada PDU.
R3	En el moment de presentar l'oferta els licitadors han de presentar un esquema frontal amb l'ocupació dels racks dels diversos equipaments presentats a la solució. Es detallarà el hardware ofertat amb les Us que ocupa cada component de la solució. S'afegirà el detall del cables / fibres que surten de cada rack per a cadascuna de les xarxes definides.

R4	Tots els nodes de còmput i la resta de components de la solució han de disposar de fonts d'alimentació redundants N+N. Que es connectaran a les PDUs corresponents redundades i a la vegada a quadres elèctrics diferents.
R5	Tant al passadís 1 com al passadís 2, els commutadors Infiniband i el de xarxa de gestió i serveis estan físicament al mateix rack amb màxim 2U de separació entre ells.
R6	Cablejat - La solució de nodes es distribueix físicament en dos racks independents, amb una disposició equitativa del nombre de nodes entre ambdós racks. - Tot el cablejat dels nodes allotjats al rack més proper al rack dels commutadors serà de mínim 3m - Tot el cablejat dels nodes allotjats al rack més distant del rack dels commutadors serà de mínim 5m - Cal incloure les òptiques i els transceptors adients per aquestes interconnexions, tant en xarxa de baixa latència com en xarxa de servei. - La correcta selecció i instal·lació dels cables i transceivers serà responsabilitat de l'adjudicatari, el qual haurà de garantir l'adequació tècnica de tots els components per assolir les prestacions requerides pel sistema.
R7	Etiquetat: - S'hauran d'incloure totes les tasques de cablejat i etiquetat dels components segons el disseny acordat entre la UPF i l'empresa adjudicatària. - La instal·lació del cablejat (fibra, elèctric, coure...) ha de quedar correctament "pentinat" i endreçat. - Tot component de la solució ha d'anar degudament etiquetat, per a ser identificat de forma única segons nomenclatura que s'estableixi entre la UPF i l'empresa instal·ladora.
R8	Els nodes de còmput hauran d'haver passat un burn-in test a fàbrica o a les instal·lacions del licitador per a evitar els DOA (Dead on Arrival).
R9	- La sala compta amb subministrament elèctric per a cada rack, mitjançant dues sortides redundants connectades a dos quadres elèctrics independents. Cada rack disposa de 2 preses CETAC de 32A amb corrent trifàsica. Atès que els nodes s'han de distribuir entre dos racks, es disposarà de 4 preses CETAC de 32A per garantir l'alimentació elèctrica necessària per al lot. - En cas que calgui afegir nous connectors o modificar els existents, aquests ajustos seran responsabilitat del licitador i no suposaran cap cost addicional per a la UPF.
R10	S'han d'incloure tots els elements i components necessaris pel transport i ubicació dels servidors a la sala CPD.

c. Software: Compatibilitats de clúster i nodes.

Referència	Descripció
R1	El sistema operatiu és linux, Rocky Linux 8.7 en la redacció d'aquest plec, per tant tots els components hardware han de ser compatibles amb aquest sistema operatiu. Un cop formalitzat el contracte, el licitador contactarà amb l'equip UPF per conèixer la versió definitiva del sistema operatiu: si es manté en 8.7 o escala a la versió 9.x.
R2	Qualsevol element hardware ha de suportar l'estàndard de IPMIv2.
R3	Qualsevol element hardware ha de suportar el sistema de clustering xCAT.

d. Manteniments i garanties:

Referència	Descripció
R1	Garantia legal: Tant els nous elements afegits com els ja existents que ho requereixin quedaran sota la cobertura d'un manteniment/garantia de maquinari/programari en el termini de 3 anys o el termini superior que ofereixi el licitador en la seva oferta.
R2	- En aquest termini l'adjudicatari serà responsable de les faltes de conformitat que existeixin en el moment del lliurament i es manifestin dins del termini. - L'empresa adjudicatària haurà de garantir la disponibilitat de peces de recanvi originals o equivalents (directament o a través d'altres agents autoritzats) per tot el cicle de vida previst de l'equipament, que com a mínim serà de 10 anys a partir de que els equips adquirits es deixin de fabricar.
R3	El manteniment/garantia començarà a partir de la data de posada en marxa de l'equipament, prèvia acceptació per escrit de la Universitat Pompeu Fabra.
R4	Garantia addicional obligatòria: - L'adjudicatari està obligat a prestar en el termini de 3 anys o el termini superior que ofereixi el licitador en la seva oferta, una garantia comercial addicional que compren els treballs de manteniment que garanteixin el correcte funcionament de l'equipament durant el termini de garantia comercial i considerats per mal funcionament que no sigui per un mal ús, incloent els recanvis, mà d'obra, desplaçaments, transport i dietes. - El recanvi de qualsevol element avariats es realitzarà sempre utilitzant material original o certificat pel mateix fabricant.
D1	Incrementar la garantia de 3 a 5 anys.
R5	Durant el període de garantia, el licitador estarà obligat a proporcionar el següent suport de software o hardware (SLA): - El temps de resposta a la notificació d'incidències haurà de ser màxim 4 hores. - L'horari serà 8x5 (8h al dia / 5 dies a la setmana)
R6	Firmware/ BIOS: - L'empresa licitadora haurà de recomanar els nivells de firmware necessaris de l'equipament hardware i deixar-lo en aquestes versions. - S'haurà de presentar la parametrització del firmware / BIOS per atorgar el màxim rendiment dels nodes de còmput. - Els nodes tindran hyperthreading deshabilitat a nivell de BIOS. - Configuració IP de les interfícies de gestió de tots els elements de la solució seguint les indicacions de l'equip UPF.
R7	La instal·lació del sistema operatiu del nodes de càlcul serà responsabilitat de l'equip UPF, però l'equip de l'adjudicació s'haurà de connectar amb els equips en cas de mal funcionament o dubtes sobre la configuració tècnica (hardware/firmware).
R8	Interlocució: - S'ha de proveir un sol punt d'avís i resolució d'incidències de qualsevol component que formi part de la solució durant el cicle de vida del projecte. - L'empresa adjudicatària haurà de disposar dos interlocutors: un de comercial i un altre de tècnic.
R9	En el moment de presentar l'oferta els licitadors han de presentar un document digital que reculli: - Descripció general dels components de la solució. - Esquemes físics de les connexions que es realitzaran - Versions de firmware i BIOS/firmware que s'instal·laran.

	- Planificació del projecte: Amb la seqüència de fases del projecte, les activitats que es duran a terme per a la posada en marxa dels servidors i el corresponent cronograma temporal. Així mateix, la memòria ha d'explicitar els diferents blocs del projecte, les accions concretes associades a cada fase i els mecanismes de comunicació i coordinació establerts entre l'equip tècnic del licitador i la Universitat Pompeu Fabra.
R10	Posada en marxa: - La instal·lació i posada en marxa serà on-site sota la supervisió de l'equip UPF. En posteriors iteracions es permetrà la connexió externa a l'entorn. - L'equip tècnic encarregat de la instal·lació hardware haurà de disposar de la formació i les capacitats tècniques per a la realització d'aquest tipus d'instal·lacions. - Una vegada acabada la instal·lació física s'haurà de comprovar que es compleixen els requeriments operatius establerts en aquest plec, com redundància elèctrica i correcta repartició a les fases elèctriques, adequació de la infraestructura, gestió correcta del flux d'aire, Cablejat òptim i ben pentinat,...
R11	Una vegada feta la instal·lació física de l'equipament, l'equip tècnic farà la transferència de coneixements respecte l'accés i la configuració tant d'IPMI com de les configuracions de consola via web i la forma futura de realitzar modificacions.
R12	En el moment de presentar l'oferta els licitadors han d'omplir la <i>Taula 6 - Manteniments i garanties</i> , on s'especifiquen els valors mínims de la proposta i els oferts

Taula 6 - Manteniments i garanties:

Concepte	Valor Mínim	Valor Ofertat
anys de cobertura del manteniment pels elements del lot		
node CPU (anys)	3	

Miquel Martínez Cucala
Coordinador TIC de Recerca del campus Poblenou
de la UPF

Juan Manuel Fuentes Díaz
Responsable de computació científica dels serveis
científico-tècnics del campus Mar de la UPF

Barcelona, 2/6/2025

Annex 1: Elements hardware:

#	Marca	Model	Ubicació	Boques lliures	Boques totals	Connectivitat	Observacions
1	Mellanox	Quantum QM8790	Passadís 1	4	40	HDR 200 Gb/s	Commutador baixa latència
2	Dell EMC	S4128F-ON	Passadís 1	2	28	10Gb SFP+	Commutador xarxa de serveis

Annex II. Càlcul exaFLOPS/any per CPU:

Aquest apartat presenta una metodologia per estimar el volum total d'operacions computacionals que una unitat central de processament (CPU) pot realitzar al llarg d'un any, expressat en exaFLOPS per any (exaFLOPS/any). Aquesta mètrica representa una estimació del rendiment efectiu sostingut sota un conjunt específic de condicions i hipòtesis.

És fonamental entendre les bases sobre les quals es construeix aquest càlcul per interpretar correctament els resultats:

- **Tipus d'Operació:** L'anàlisi es centra exclusivament en operacions de coma flotant de doble precisió (FP64), comunes en càlcul científic i tècnic d'alta precisió.
- **Instruccions Vectorials:** Es pressuposa l'ús intensiu i eficient de les instruccions vectorials AVX (Advanced Vector Extensions, específicament AVX2 o AVX-512 segons la CPU) i de les instruccions FMA (Fused Multiply-Add), que permeten realitzar múltiples operacions per cicle de rellotge.
- **Freqüència de Rellotge:** Per assegurar una estimació de rendiment sostenible en càrregues de treball prolongades, s'utilitza la freqüència base especificada pel fabricant per a cada nucli, en lloc de la freqüència màxima o "turbo".
- **Factor d'Eficiència:** Es reconeix que el rendiment teòric màxim rarament s'assoleix en la pràctica degut a colls d'ampolla (memòria, dependències de dades, etc.) i sobrecàrregues del sistema. Per això, s'aplica un factor d'eficiència pràctica del 80% sobre el rendiment teòric calculat.
- **Base Temporal Anual:** Per normalitzar el càlcul anual i facilitar la planificació, es considera un any estàndard de 360 dies (equivalent a 31.104.000 segons).

Els valors d'exaFLOPS/any obtinguts mitjançant aquesta metodologia ofereixen una eina per comparar la capacitat computacional teòrica ajustada de diferents CPUs sota aquestes condicions estandarditzades. Cal tenir present, però, que el rendiment real obtingut en una aplicació específica dependrà fortament de la naturalesa de l'algorisme, l'optimització del codi, l'amplada de banda i latència de la memòria, i altres característiques del sistema i de la càrrega de treball.

1) Càlcul dels FLOPS per cicle en precisió doble (FP64) mitjançant instruccions vectorials AVX i FMA

Per calcular els FLOPS per cicle en precisió doble (FP64) que pot realitzar una CPU utilitzant instruccions vectorials AVX i FMA, es tindran en compte els següents factors:

Dades necessàries:

- **(A) Amplada del vector** (longitud del registre vectorial):
 - **AVX-512:** 512 bits

- **AVX2:** 256 bits
- **(E) Elements per vector FP64** (cada element FP64 ocupa 64 bits):
 - AVX-512: $512/64 = 8 \text{ elements}$
 - AVX2: $256/64 = 4 \text{ elements}$
- **(U) Unitats d'execució AVX-512 FMA disponibles per nucli:**
 - Depèn del model concret de CPU.
- **(I) Instruccions FMA per cicle per unitat d'execució:**
 - **Constant = 2 FLOP** (cada instrucció FMA realitza simultàniament una multiplicació i una suma, és a dir, dues operacions aritmètiques en una sola instrucció).

Fórmula general:

$$FLOP \text{ per cicle} = E \times U \times I$$

On:

- E: Elements per vector (determinat per AVX-512 o AVX2).
- U: Nombre d'unitats d'execució vectorial FMA (generalment 1 o 2).
- I: Operacions per instrucció FMA (sempre 2).

Exemple:

a) CPU sense AVX-512 (només AVX2) i una unitat d'execució FMA:

- Amplada vectorial: 256 bits
- Elements FP64 per vector: $256 / 64 = 4$
- Unitats FMA disponibles per nucli: $U=1$
- Operacions per FMA: $I=2$

$$FLOP \text{ per cicle (AVX2)} = 4 \times 1 \times 2 = 8 \text{ FLOP/cicle}$$

b) CPU amb AVX-512 i dues unitats d'execució FMA:

- Amplada vectorial: 512 bits
- Elements FP64 per vector: $512 / 64 = 8$
- Unitats FMA disponibles per nucli: $U=2$
- Operacions per FMA: $I=2$

$$FLOP \text{ per cicle (AVX512)} = 8 \times 2 \times 2 = 32 \text{ FLOP/cicle}$$

2) Càlcul dels FLOPS per segon per nucli

Per obtenir la capacitat computacional en FLOPS (operacions en coma flotant per segon) d'un nucli, cal multiplicar els FLOPS per cicle pel nombre de cicles per segon (freqüència del rellotge).

Es farà servir sempre la freqüència base (Base Frequency) publicada pel fabricant, i no la freqüència màxima (Turbo Boost), ja que només la freqüència base garanteix un funcionament sostingut i estable durant períodes de càlcul intensiu, condició indispensable per a l'estimació de rendiment continuat en entorns HPC.

La fórmula és:

$$FLOPS \text{ per nucli} = (FLOP \text{ per cicle}) \times (Freqüència \text{ base en Hz})$$

Exemple:

Suposem un nucli de CPU amb:

- 32 FLOP per cicle (segons el càlcul AVX-512 + FMA),
- Freqüència base de 3,0 GHz.

Sabent que:

- 3,0 GHz = 3×10^9 Hz

tenim:

$$FLOPS \text{ per nucli} = 32 \times 3 \times 10^9 = 96 \times 10^9 \text{ FLOPS} = 96 \text{ GFLOPS per nucli}$$

3) Càlcul dels FLOPS per segon totals per CPU

Per determinar la capacitat computacional teòrica total d'una CPU, cal multiplicar els FLOPS per nucli pel nombre total de nuclis independents disponibles a la CPU (no es té en compte els threads en cas de que la CPU tingui Hyper-Threading).

La fórmula general és:

$$FLOPS \text{ teòrics per CPU} = (FLOPS \text{ per nucli}) \times (\text{Nombre de nuclis})$$

Tanmateix, per obtenir una estimació més realista del rendiment efectiu, **cal aplicar un factor de correcció** que tingui en compte:

- Ineficiències internes del processador (gestió de dependències, latències de memòria, estratègies d'execució fora d'ordre).
- Limitacions derivades de l'ús real de les unitats vectorials (no sempre al 100% d'utilització).
- Overhead de control de tasques, sincronització entre fils i variabilitat de la càrrega.

Per aquest motiu, **es considera una eficiència pràctica mitjana del 80%** respecte al pic teòric de la CPU.

Fórmula ajustada:

$$FLOPS \text{ efectius per CPU} = FLOPS \text{ teòrics per CPU} \times 0,8$$

Exemple de càlcul:

Considerem una CPU amb les següents característiques:

- 32 FLOP per cicle.
- Freqüència base: 3,0 GHz = 3×10^9 Hz.
- Nombre de nuclis: 48 nuclis.

Primer calculem els FLOPS per nucli:

$$FLOPS \text{ per nucli} = 32 \times 3 \times 10^9 = 96 \text{ GFLOPS}$$

Després, els FLOPS teòrics totals per CPU:

$$FLOPS \text{ teòrics per CPU} = 96 \text{ GFLOPS} \times 48 = 4608 \text{ GFLOPS} = 4,608 \text{ TFLOPS}$$

Aplicant el factor d'eficiència del 80%:

$$FLOPS \text{ efectius per CPU} = 4,608 \text{ TFLOPS} \times 0,8 = 3,6864 \text{ TFLOPS}$$

Per tant, el rendiment pràctic previst per cada CPU serà de **3,6864 TFLOPS** en condicions de càrrega intensiva i operació sostinguda.

4) Càlcul dels exaFLOPS per any i CPU

Per obtenir la capacitat computacional anual d'una CPU en termes d'exaFLOPS per any, es multiplica el rendiment efectiu per segon pel nombre total de segons d'activitat anual, i es divideix entre 10^{18} per expressar el resultat en unitats d'exaFLOPS.

En aquest cas, per criteris de normalització en entorns científics i de planificació d'infraestructures, es considerarà un any de **360 dies** en lloc dels 365 naturals.

Segons per any:

$$360 \times 24 \times 3600 = 31104000 \text{ segons/any}$$

Fórmula de càlcul:

$$\text{ExaFLOPS per any i CPU} = [(FLOPS \text{ efectius per CPU}) \times (31104000 \text{ segons})] / 10^{18}$$

Exemple de càlcul:

Amb les dades prèviament calculades:

- Rendiment efectiu per CPU: 3,6864 TFLOPS = $3,6864 \times 10^{12}$ FLOPS.

Substituïm a la fórmula:

$$\text{exaFLOPS/any} = (3,6864 \times 10^{12} \times 31104000) / 10^{18} = 114,78 \text{ exaFLOPS/any}$$

Per tant, **cada CPU assoliria un rendiment efectiu aproximat de 114,78 exaFLOPS per any.**

5) EXEMPLES

A continuació es mostra un exemple del càlcul fet amb diferents CPUs:

Model	Família	Nuclis	Frecuencia bàsica del procesador (Ghz)	Amplada vectorial (AVX-512: 512, AVX2: 256)	Unitats d'execució FMA	FLOPs per cicle	GFLOPS per nucli	ExaFLOPS any
Intel Xeon 6980P	Granite Rapids	128	2	512	2	32	64	203,85
Intel Xeon 6979P	Granite Rapids	120	2,1	512	2	32	67,2	200,66
Intel Xeon 6972P	Granite Rapids	96	2,4	512	2	32	76,8	183,46
Intel Xeon 6960P	Granite Rapids	72	2,7	512	2	32	86,4	154,8
Intel Xeon 6787P	Granite Rapids	86	2	512	2	32	64	136,96
Intel Xeon 6767P	Granite Rapids	64	2,4	512	2	32	76,8	122,31
Intel Xeon 6760P	Granite Rapids	64	2,2	512	2	32	70,4	112,12
Intel Xeon 6747P	Granite Rapids	48	2,7	512	2	32	86,4	103,2
Intel Xeon 6756E	Sierra Forest	128	1,8	256	2	16	28,8	91,73

S'adjunta un full de càlcul a mode d'exemple.