

PLEC DE PRESCRIPCIONS TÈCNIQUES

1. Objecte

L'objecte d'aquest contracte és l'ampliació progressiva del clúster de computació per al Center of Brain Cognition de la Universitat Pompeu Fabra a emprar en els projectes de recerca HYPERSTIM (High-dimensional electrical stimulation for visual prosthesis) finançat pel fons EUROPEAN INNOVATION COUNCIL AND SMES EXECUTIVE AGENCY, núm 101071015 i NEMESIS (Neurological Mechanisms Of Injury, And Sleep-Like Cellular Dynamics) finançat pel fons EUROPEAN RESEARCH COUNCIL EXECUTIVE AGENCY, núm 101071900.

L'adquisició del nou sistema ha d'incloure:

- Servidors amb capacitat de supercomputació GPU.
- Les llicències de programari.
- La garantia del maquinari.
- El suport i actualització del programari.
- Cablejat i altres components per a la connexió del sistema.
- Instal·lació, posada en marxa i formació al personal tècnic.

Codi CPV: 48820000-2 Servidors

2. Requeriments tècnics

Els licitadors han de presentar una proposta de maquinari que compleixi com a mínim amb els requisits que es resumeixen en l'annex 4 i que a continuació es desenvolupen:

2.1. Sistema de càlcul

Tots els components presentats han d'estar certificats pel propi fabricant del model en qüestió.

La proposta ha de tenir el mateix model de CPU a tots els servidors, el mateix model de dimm de RAM i la mateixa quantitat de memòria RAM total del server, el mateix model de GPU i els mateixos discs SSD.

Cadascun dels servidors (node) de càlcul inclòs a l'oferta haurà de tenir, com a mínim, les següents característiques:

2.1.1. CPU

Mínim de 2 micro-processadors amb un total mínim de 32 cores per processador a una velocitat mínima de 2,1 Ghz i 60MB de memòria caché.

Per compatibilitat amb l'actual clúster, els processadors han de tenir arquitectura de 64 bits (x86-64), pertanyent a la família Sapphire Rapids d'Intel (4ª generació d'Intel).

No seran vàlids entorns que precisin de refrigeració líquida.

Mínim 60 MB de cache

2.1.2. Emmagatzematge

Cada node tindrà com a mínim 2 discs durs SSD o NVMe iguals, de mínim 480GB, de qualitat Enterprise/Server i configurats en RAID-1.

El RAID s'ha de construir per hardware, no sent vàlides solucions RAID per software.

En cas que la placa base no permeti la creació/gestió de RAIDs, s'haurà d'afegir una controladora de disc que ho permeti.

El menú de gestió dels RAIDs ha d'aparèixer a la BIOS o en un menú que es mostri en temps d'arrencada.

Solucions on l'eina de gestió estigui fora de l'arrencada quedaran descartades.

2.1.3. Memòria

Tenir un mínim de 512 GB on els xips de memòria siguin de mínim 32 GB i iguals entre sí.

La memòria, que ha de ser de tipus ECC, ha d'estar equilibrada, es a dir, ha de tenir els mateixos DIMMs de memòria a cada canal de comunicació entre core i memòria (accés simètric).

La memòria ha de ser DDR5 amb freqüència mínima de 4800 Mhz i la més alta que el micro-processador permeti.

2.1.4. GPU

Cada servidor han d'incloure un a mínim d'una targeta GPU.

S'han de connectar a un port PCIe 4x.

Cal que tingui tecnologia CUDA per compatibilitat amb el clúster actual

El model de GPU ha de complir:

<i>mínim de memòria de GPU:</i>	<i>48GB GDDR6 with ECC</i>
<i>mínim ample de banda de memòria:</i>	<i>864GB/s</i>
<i>interconnexió: PCIe Gen4 x 16:</i>	<i>64GB/s bidireccional</i>
<i>mínim FP32 teraFLOPS:</i>	<i>91,6</i>
<i>mínim TF32 Tensor Core teraFLOPS:</i>	<i>183* 366**</i>
<i>mínim BFLOAT16 Tensor Core teraFLOPS:</i>	<i>362.05* 733**</i>
<i>mínim FP16 Tensor Core:</i>	<i>362.05* 733**</i>
<i>mínim FP8 Tensor Core:</i>	<i>733* 1,466**</i>
<i>mínim Peak INT8 Tensor TOPS:</i>	<i>733* 1,466**</i>
<i>mínim Peak INT4 Tensor TOPS:</i>	<i>733* 1,466**</i>
<i>Potència consum màxima:</i>	<i>350 W</i>

** En operacions normals*

***En operacions amb dispersió*

2.1.5. Servidor

El node ha de ser de màxim 2Us.

Dur dos connectors amb targeta de xarxa 10Gb SFP+ o compatibles.



Disposar de port IPMI per administració remota i gestió de sistema integrat out-of-band. En cas de tenir llicència, haurà d'incorporar una amb validesa perpètua i sense costos addicionals futurs.

Ha de ser "enrackable" en l'armari descrit a l'apartat 2.3: Armari RACK. També s'acceptarà si es proporciona els adaptadors necessaris per poder-lo enrackar a l'armari descrit en l'apartat 2.3

El node comptarà amb almenys una targeta HDR 200Gb/s (connector QSFP56) compatible amb els equips Infiniband descrit en l'apartat 2.1 Infiniband.

Les fonts redundants amb certificació Platinum hauran de suportar la càrrega de totes les GPUs funcionant al 100%. En aquest cas d'avaria d'una font, el servidor haurà de seguir funcionant en el supòsit de tenir totes les GPUs treballant al 100%.

3. Comunicacions

3.1. Infiniband

S'ha d'incloure tots els cables Mellanox Passive Copper (amb connectors QSFP56) per a la connexió de tots els equips oferts, la llargària ha de ser la suficient per la connectivitat més pentinat entre els diferents punts de connexió. La connectivitat ha de ser compatible amb el switch infiniband descrit en l'annex 2.

3.2. Ethernet

S'ha d'incloure tots els cables 10G SFP+ DAC Cable per a la connexió de tots els equips oferts, dos per equip. La connectivitat ha de ser compatible amb el swich 10G descrit en l'annex 3.

4. Armari Rack

No serà necessari armari RACK sempre que la solució proposada ocupi igual o menys de 14U d'alçada. En aquest cas han de poder ser muntats en un rack amb forat rodó, per tant, les dimensions i les guies han de ser compatibles.

S'hauran d'incloure tantes PDUs com siguin necessàries per a connectar els equips de la proposta a la doble línia elèctrica.

Les PDUs aniran connectades al sistema de distribució elèctric amb entrada a l'armari de corrent. Actualment el CPD disposa de línies monofàsiques de 16 i 32A. Caldrà especificar el nombre i el tipus de línies de corrent necessàries per la seva instal·lació.

5. Traspàs

Al finalitzar la instal·lació i configuració, l'adjudicatari haurà de presentar la documentació tècnica del sistema en qüestió i una sessió de traspàs de coneixements amb les configuracions dutes a terme.

6. Garantia

El termini de garantia serà de 5 anys, a comptar de la data de posada en marxa de l'equipament, prèvia acceptació per escrit de la Universitat Pompeu Fabra. Durant



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'equip que, com a mínim serà de 10 anys a partir dels equips adquirits es deixin de fabricar.

El fabricant ha de disposar del driver i actualitzacions de firmware de la BIOS/UEFI per la plataforma dins dels 10 anys després de l'adquisició.

L'adjudicatari estarà obligat així mateix a prestar en un termini mínim de 3 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. Finalitzat el període de garantia, el cost dels components es facturarà a preu de mercat, si fos el cas.

L'adjudicatari proporcionarà un sistema per a la comunicació i la gestió del seguiment de les incidències, que ha d'explicar de manera explícita en la documentació a presentar amb l'oferta.

7. Interlocució

L'empresa adjudicatària haurà de disposar, un cop formalitzat el contracte, de com a mínim dos perfils d'interlocució, per mantenir contacte durant la fase de preparació de material, lliurament, posada en marxa i gestió de la garantia:

- Un interlocutor comercial
- Un interlocutor tècnic.

8. Condicions de lliurament

El subministrament es dividirà en 2 lliuraments en la forma indicada en el plec de clàusules administratives, per a cada un dels projectes de recerca, ja que tenen períodes d'execució diferenciats tal i com s'explica en la memòria justificativa de la licitació.

Juan Manuel Fuentes Díaz

Responsable de Computació Científica dels Serveis Científic Tècnics de la UPF

Barcelona, 26/9/2023

Annex 1: Sistema actual de computació: nodes de càlcul:

Marca	Model	#	CPU	Sistema Operatiu	Cores	RAM	Ports Ethernet	Ports IB	Targeta IB
Dell	Dell PowerEdge R430-730	26	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	Rocky Linux 8.4	32	256GB	NetXtreme BCM5720 2-port Gigabit Ethernet PCIe	4xFDR	MT27500 Family [ConnectX-3]
Sie	Ladónl - GPU Quadro RTX 4000	2	Intel® Xeon® Gold 6248R	Rocky Linux 8.4	48	1,5TB	Intel Corporation I350 Gigabit Network Connection	4xFDR	Mellanox Technologies MT27800 Family [ConnectX-5]

Annex 2: Sistema actual de computació: Switchos de baixa latència:

Marca	Model	#	Boques totals	Boques lliures	Connectivitat	Observacions
Mellanox	IB-2 SB7800	1	36	4	EDR 100 Gb/s	
Mellanox	Quantum QM8790	1	40	10	HDR 200 Gb/s	

Annex 3: Sistema actual de computació: Switchos ethernet:

Marca	Model	#	Boques totals	Boques lliures	Connectivitat	Observacions
Dell EMC	S4128F-ON	2	56 (2 x 28)	49	10 GBase T	Cada switch té 28 boques a 10Gb/s. Els dos switchos estan configurats en VLT



Annex 4: taula d'especificacions mínimes

Equips: mínim 4 servidors

Especificacions mínimes per servidor	#sockets / cpu	cores	velocitat cpu (Ghz)	RAM (Gb)	Dimm RAM (Gb)	Freq memòria (MT/s)	Discs (Gb)	# GPU
Servidor	2	64	2,1	512	32	4800	2x480	1

Tamany màxim de servidor: 2U

Especificacions mínimes per GPU	Memòria (GB)	FP32 (teraFLOPS)	TF32 Tensor Core (teraFLOPS)	BFLOAT16 Tensor Core (teraFLOPS)	FP16 Tensor Core	FP8 Tensor Core	Peak INT8 Tensor (TOPS)	Peak INT4 Tensor (TOPS)	Ample Banda memòria (GB/s)	Interconne xió PCIe Gen4x16 (GB/s)
gpu	48	91,6	183* 366**	362.05* 733**	362.05* 733**	733* 1,466**	733* 1,466**	733* 1,466**	864	64

* En operacions normals

**En operacions amb dispersió

Potència consum màxima de GPU: 350 W