

PROJECTES DE LABORATORI DE PAP

Curs 2019/20 (Quadrimestre de primavera)

Designing our own HPC-oriented cluster

Rafel-Albert Bros Esqueu

Andrea Querol de Porras

Joan Vinyals Ylla-Català

Pablo Vizcaino Serrano

Facultat d'Informàtica de Barcelona

Índex

1	Introducció	3
1.1	Restriccions del disseny	3
2	Primeres decisions	4
2.1	CPU	4
2.2	Motherboard	5
2.3	Memòria RAM	7
2.4	Xarxa d'interconnexió	8
2.5	GPU	9
3	Anàlisi de possibles configuracions	13
4	Clúster final	15
4.1	Roofline	15
	Apèndix	16
	Apèndix A hola	16

Índex de figures

1	Comparativa entre les diverses CPUs valorades.	5
2	Exemple de nodes que podem trobar a la web. A l'esquerra amb GPUs i a la dreta sense. Tots single-socket.	6

Índex de taules

1	Comparativa d'especificacions per cada processador considerat. . . .	4
2	Comparació inicial d'especificacions de plaques base.	6
3	Comparació entre les diferents configuracions dels nodes.	7
4	Comparació de les diferents memòries estudiades.	8
5	Comparació entre les diferents configuracions dels nodes.	9
6	Característiques i rendiment de les targetes gràfiques estudiades. . .	10
7	Mostra de les diverses combinacions valorades.	14

1 Introducció

En aquest projecte ens proposem dissenyar un clúster d'HPC (*High Performance Computing*) obeint les restriccions establertes en la documentació. Aquestes condicions seran explicades en detall exposant, com no podria ser d'altre manera, les problemàtiques causades de cara a les decisions que caldrà prendre a l'hora d'escollir els diversos components del clúster.

1.1 Restriccions del disseny

En primer lloc, hi ha una restricció molt clara i sense marge per dubtes: el cost màxim per adquirir tots els components del clúster ha de ser d'un milió i mig d'euros. Això requerirà tenir ben present el preu de tots els elements i procurar no superar el límit.

L'espai també queda definit: disposem de dos armaris, cadascun d'ells de 42U i, en definitiva, 84U totals. Aquells nodes que contenen una gràfica ocuparan 2U, mentre que aquells que no en tinguin ocuparan 1U. Cal entendre que a més commutadors major serà el nombre de nodes que podem connectar a la xarxa però, paral·lelament, menor serà el nombre d'unitats que quedaran disponibles pels nodes de computació. Un cop s'hagi establert el nombre d'unitats i nodes disponibles, s'haurà de cercar la millor combinació possible entre nodes amb GPUs i nodes lliures d'elles.

Pel que fa a GPUs, el nostre disseny ve delimitat per la condició que fins a un màxim del 30% dels flops del sistema poden provenir de nodes de computació accelerats, és a dir, de nodes amb GPUs. Tanmateix, el nombre màxim de GPUs dependrà de la placa base escollida, que al seu torn depèn del processador que haguem escollit. En definitiva: les decisions s'hauran de prendre amb peus de plom, sempre mantenint la possibilitat de fer un pas enrere i modificar qualsevol elecció prèvia.

Finalment, no existeix cap restricció específica pel que fa al consum energètic, però evidentment és important intentar maximitzar l'eficiència energètica del clúster.

2 Primeres decisions

En aquest apartat exposarem els diversos components escollits inicialment pel nostre clúster, incloent en tot moment el procés de raonament que hem anat seguint per prendre qualsevol decisió.

Tenint en compte que una elecció es veurà afectada per totes les prèvies, en aquesta primera fase s'han escollit diversos components candidats per a cada aspecte del clúster. Properament, els estudis per elegir els components definitius es mostraran explicats en profunditat a l'apartat d'*Anàlisi*.

2.1 CPU

La primera qüestió a plantejar-nos ha estat quin seria el processador més adient. Per a prendre una decisió de la forma més encertada possible hem cregut oportú elaborar una taula amb diverses possibilitats i comparar les característiques de totes elles, tal com es pot observar a la taula següent:

Processador	Freq (GHz)	SIMD (bits)	SIMD units	Nuclis	Watts	Preu (Dòllars)	GFlops/ Socket	GFlops/ Watts	GFlops/ Dòllar
ThunderX2 CN9980	2.20	128	2	32	180	1795	563.20	3.13	0.314
ThunderX2 CN9978	2.20	128	2	30	175	1600	528.00	3.02	0.330
AMD EPYC 7702P	2.00	256	2	64	200	4783.99	2048.00	10.24	0.428
AMD EPYC 7502P	2.50	256	2	32	180	2493.99	1280.00	7.11	0.513
AMD EPYC 7742	2.25	256	2	64	225	7581.02	2304.00	10.24	0.304
AMD EPYC 7601	2.20	256	2	32	180	3,388.34	1126.40	6.26	0.332
AMD EPYC 7542	2.90	256	2	32	225	3840.88	1484.80	6.60	0.387
Intel Xeon Platinum 8253	2.20	512	2	16	125	3115	1126.40	9.01	0.362
Intel Xeon Platinum 8256	3.80	512	2	8	105	7007	972.80	9.26	0.139
Intel Xeon Platinum 8260	2.40	512	2	24	165	4702	1843.20	11.17	0.392
Intel Xeon Platinum 8276	2.20	512	2	28	165	8716	1971.20	11.95	0.226
Intel Xeon Platinum 8280L	2.70	512	2	28	205	13012	2419.20	11.80	0.186
Intel Xeon Platinum 8284	3.00	512	2	28	240	15460	2688.00	11.20	0.174

Taula 1: Comparativa d'especificacions per cada processador considerat.

Per a major precisió i comoditat a l'hora d'analitzar els diversos processadors candidats, els principals criteris valorats per a prendre aquesta elecció, tal com es pot veure a la figura 1, han estat els GFlops/Socket, els GFlops/Watt i els GFlops/Dòllar (sent els GFlops normalitzats respecte la mitjana).

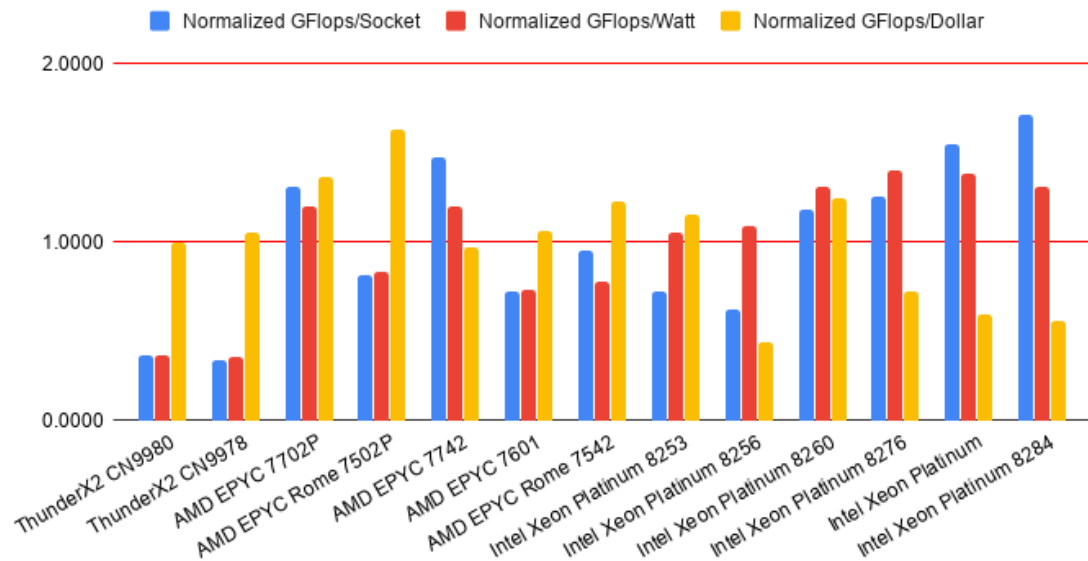


Figura 1: Comparativa entre les diverses CPUs valorades.

Després d'estudiar els resultats, per tenir un major marge de maniobra de cara a decisions futures hem pre-seleccionat dos processadors: l'*AMD EPYC 7702P* [1] i l'*AMD EPYC Rome 7502P* [2], el primer per mostrar un excel·lent equilibri entre les tres principals característiques estudiades, i el segon per oferir el GFlop més barat sense menystenir excessivament els altres dos aspectes principals.

2.2 Motherboard

Una vegada escollits dos processadors com a contrincants, hem començat a mirar quines plaques bases hi havia al mercat. Primerament, les hem anat buscant de manera individual i les que hem seleccionat es poden veure a la taula següent:

Placa mare	Nuclis Màx.	Storage Speed	DIMS Màx.	USB 3.0	PCIe 4.0	Preu (Euros)
H12DSU-iN [3]	64	6 Gbps	32	5	1x16, 1x32, 1x40	Coming soon
H12DST-B [4]	64	6 Gbps	16	2	3x16, 1x24	-
H12SSW-NT [5]	64	6 Gbps	8	7	1x16, 1x32	1356.53 [6]
H12SSW-iN [7]	64	6 Gbps	8	7	2x32	-
H12SST-PS [8]	64	6 Gbps	8	2	3x16	3959.86 [9]

Taula 2: Comparació inicial d'especificacions de plaques base.

Com es pot veure, les diferències principals són el número de DIMS de memòria i els PCI-Express. No obstant, una vegada fetes les comparacions inicials, ens hem adonat que per algunes d'elles era excessivament complicat trobar el preu. Per exemple, els llocs on trobàvem els preus no és el mateix que on trobàvem les especificacions, per tant en alguns no estem segurs de que els preus siguin del tot realistes.

Buscant els preus ens hem trobat amb una altra web [10] que ens permet crear un node seleccionant nosaltres els seus components. Hem vist que els tipus de node es diferencien inicialment entre aquells amb GPU i aquells sense GPU. Dins d'aquests dos grups, trobem també els dual-socket i els single-socket.

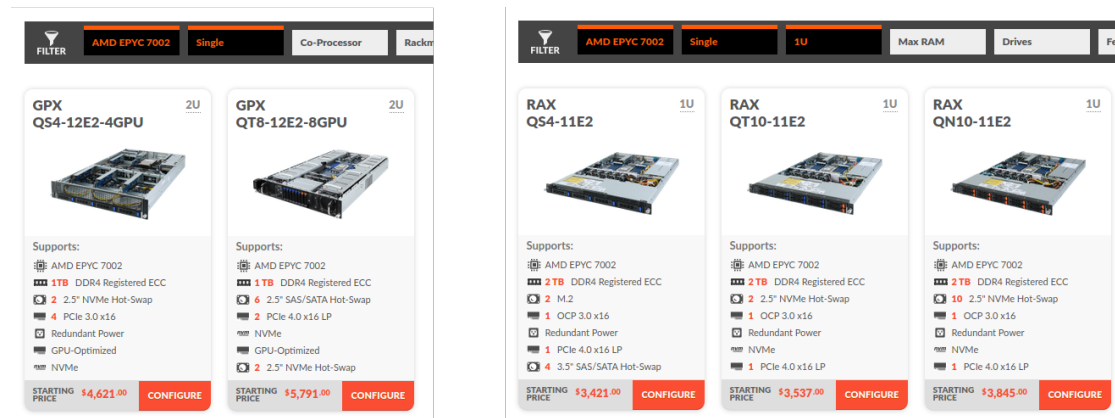


Figura 2: Exemple de nodes que podem trobar a la web. A l'esquerra amb GPUs i a la dreta sense. Tots single-socket.

Com que el processador *AMD EPYC 7702P* té 64 nuclis i l'*AMD EPYC Rome 7502P* en té 32. En cas d'escollir un node amb suport per GPU, ocuparem 2 Us

en compte d'una de sola. En el cas dels nodes amb dual-socket, es poden suportar fins a 8 GPUs mentre que amb els single-socket només la meitat, 4.

En la següent taula es veuen aquestes possibles combinacions. En el cas d'utilitzar un node amb GPUs no s'ha inclòs el preu de la pròpia GPU, ja que això dependrà de quina escollim més endavant. Per la mateixa raó, tampoc s'han exposat els GFLOPs de les versions amb GPU. En aquests preus tampoc s'inclouen els costos de les memòries i les targetes de xarxa.

Configuració	Nuclis	GFlops CPU	GPUs	Us	Preu (Euros)	GFlops/ Euros
Dual amb AMD EPYC Rome 7502P	128	2560	-	1	9393	0.256
			8	2	12266	-
Single amb AMD EPYC 7702P	128	2048	-	1	7542	0.251
			4	2	9461	-

Taula 3: Comparació entre les diferents configuracions dels nodes.

Veiem que tot i que la versió amb dual-socket ens proporciona més GFlops, l'eficiència GFlops/euro és millor per la versió amb single-socket. Per tant, en aquest punt encara no podem decantar-nos per cap de les dues. Una vegada haguem considerat quines GPUs utilitzem, quina xarxa tenim, etc., podrem valorar si tenim suficient pressupost com per escollir la opció amb més GFlops. A més, quan les GPUs entrin en joc també s'haurà de tenir en compte que amb l'opció dual-socket en podem posar el doble per cada node.

2.3 Memòria RAM

Una de les restriccions del disseny es que hem de tenir 2GB de RAM per cada core. Tenint en compte que tenim 64 cores (ja sigui amb dual-socket o amb single-socket), necessitarem 128GB de RAM.

Hem decidit avaluar tres combinacions diferents per arribar a aquests 128GB. La primera és amb 16 DIMMs de 8GB, la segona amb 8 DIMMs de 16GB i per últim 4 DIMMs de 32GB. És més econòmic tenir menys DIMMs amb memòries de més capacitat, però limita el paral·lisme per accedir-hi. Un altre aspecte a tenir en compte és que quants menys DIMMs utilitzem, més ens podrem expandir en un futur comprant més memòria.

També hem comparat les opcions de comprar memòria ECC (feta expressament per servidors) o NO-ECC. La diferència de preu és notable però creiem que val la pena, considerant que el pressupost no és ajustat. No hem aconseguit trobar memòries ECC de 3200Hz. En la taula següent es pot observar les memòries que hem comparat:

Memòria RAM	ECC	GB	Freq. (Hz)	#	Latència	Preu (€) (1u)	Preu (€) (total)
Kingston KSM29RS8 [11]	SI	8	2993	16	CL21	64.9	1038.4
Kingston KSM29RS4 [12]	SI	16	2993	8	CL21	122.9	983.2
Kingston KSM29RD4 [13]	SI	32	2993	4	CL21	242.9	971.6
Kingston HyperX Impact [14]	NO	8	3200	16	CL20	64.99	1039.84
Kingston HyperX Impact [15]	NO	16	3200	8	CL20	107.99	863.92
Kingston HyperX Fury Black [16]	NO	32	3200	4	CL16	169.0	676

Taula 4: Comparació de les diferents memòries estudiades.

Hem decidit decantar-nos per la versió amb 8 DIMMs de 16GB. Creiem que té un bon balanceig entre paral·lelisme, possible extensió i preu. Com estem dins el pressupost, hem escollit la versió ECC.

2.4 Xarxa d'interconnexió

Ja que tenim un nombre relativament reduït d'unitats, i volem dedicar tantes unitats com sigui possible a nodes de còmput, volem, per tant reduir el nombre d'unitats que dediquem als commutadors. Per aconseguir aquesta fita, necessitem que cada commutador tingui el màxim nombre de ports per unitat. Com a efecte col·lateral d'aquesta estratègia, ens veiem amb un sistema que pràcticament es veu forçat a utilitzar una topologia de xarxa de tipus malla.

Pel que fa a distribuïdors, ens vam centrar en Mellanox Technologies. A la pàgina web [17] de la companyia, vam trobar commutadors InfiniBand, de diferents preus i velocitats (com més ports més velocitat). A la tala 5 hem recollit els diferents models candidats a ser utilitzats en el nostre sistema.

La quantitat de commutadors que utilitzem en el nostre sistema es un factor clau, que marca, tant el preu, com el rendiment, d'aquest. D'altre banda, però, afecta diferent als dos parametres, anteriors en funció dels altres components. Per tant

no podem determinar encara quina sera la millor configuració. No obstant si que podem reduir el rang de configuracions, per tal de que l'anàlisi sigui més àgil.

Ja que utilitzem una topologia de tipus malla completa, el nombre de connexions C a nodes de comput que podra tennir la xarxa ve marcat per el nombre de ports P del que disposa cada commutador aixi com el nombre de commutadors N que tingui el sistema. $C = N \times (P - N + 1) = -N^2 + N(P + 1)$

-
- Per el cas del model MSX6018F-1SFS, ja que tenim 18 ports, el nombre de nodes que podem connectar-hi, ve determinat

Com es pot veure a la taula 5, hem considerat només commutadors de Mellanox. Tots aquests commutadors ocupen només una unitat.

Model	Ports	Speed (Gb/s)	Preu (\$)
MSX6012F-2BFS[18]	12	56	9309.00
MSX6018F-1SFS[19]	18	56	14791.00
MSB7800-ES2F[20]	36	100	25633.00
MQM8700-HS2F[21]	40	200	29629.00

Taula 5: Comparació entre les diferents configuracions dels nodes.

2.5 GPU

Finalment, queda plantejar-nos l'últim element pel nostre clúster: les unitats de processament gràfic o GPUs. Un cop més, el procediment ha estat buscar diverses GPUs candidates i comparar les seves característiques bolcant-les en una taula, tal com es pot veure a la taula 6.

Model		NVIDIA Tesla V100 [22]	NVIDIA RTX 2080 TI [23]	NVIDIA Quadro RTX 8000 [24]	NVIDIA Tesla T4 [25]	AMD Radeon Pro WX 8200[26]	AMD Radeon Instinct MI50 [27]	AMD Radeon Instinct MI25 [28]
Memòria	Freq (MHz)	1246	1350	1395	585	1200	1200	1400
	Type	hbm2	gddr6	gddr6	gddr6	hbm2	hbm2	hbm2
	GB	16 o 32	11	48	16	8	32	16
	Bandwidth (GB/s)	900	616	672	320	512	1024	484
Interconnexió	Interface	PCIe 3.0	PCIe 3.0	PCIe 3.0	PCIe 3.0	PCIe 3.0	PCIe 3.0 & 4.0	PCIe3.0
	Lanes	16	16	16	16	16	16	16
	Bandwidth (GB/s)	32	32	32	32	32	32	32
	Power (W)	300	250	260	70	230	300	300
FP64	Preu (\$)	7,399.00 [29]	1,159.99 [30]	4,902.00 [31]	2,199 [32]	969.99 [33]	4,438.00 [34]	9,156.65 [35]
	GFlops	7,235.58	420.20	509.80	254.40	672.00	6,865.92	768.00
	GFlops/W	24.12	1.68	1.96	3.63	2.92	22.89	2.56
	GFlops/\$	0.98	0.36	0.10	0.12	0.69	1.55	0.08
FP32	GFlops	14,469.12	13,772.80	16,701.44	8,336.38	11,008.00	13,731.84	12,584.96
	GFlops/W	48.23	55.09	64.24	119.09	47.86	45.77	41.95
	GFlops/\$	1.96	11.87	3.41	3.79	11.35	3.09	1.37
FP16	GFlops	28938.24	27545.60	33402.88	66693.12	2201.60	27463.68	25169.92
	GFlops/W	96.46	110.18	128.47	952.76	9.57	91.55	83.90
	GFlops/\$	3.91	23.75	6.81	30.33	2.27	6.19	2.75

Taula 6: Característiques i rendiment de les targetes gràfiques estudiades.

En el mercat actual, hi ha principalment dos desenvolupadors de targetes gràfiques en mercat: NVIDIA i AMD. La recerca de GPUs s'ha centrat en models d'aquestes dues marques. Tradicionalment NVIDIA ha tingut la principal quota de mercat en l'entorn de la supercomputació, però últimament AMD està desenvolupant productes per a aquest mercat molt competitiu amb NVIDIA.

S'han buscat dos tipus de models per cada desenvolupador: un amb molta potència de càlcul i l'altre amb uns GFlops inferiors. L'explicació és que així ens podíem plantejar dos escenaris, un per cada tipus de model. Els models amb molts GFlops els vam buscar pel cas de tenir només uns pocs nodes amb moltes GPUs, és a dir, concentrant la potència de càlculs. L'altre tipus de model, amb menys GFlops per unitat, es van escollir pel suposat escenari de tenir la potència de càlcul de la GPU en més nodes, o fins i tot en tots.

Aquests dos escenaris presenten avantatges i inconvenients. En el cas de tenir pocs nodes amb molts GFlops concentrats:

- Avantatges:

- Menys nodes de 2U i més d'1U. Això permetria tenir més cores en el clúster i per tant més nodes per a escalar aplicacions o executar diverses aplicacions sense haver d'esperar recursos disponibles.
- Inconvenients:
 - Els usuaris hauran de compartir nodes per poder accedir a les GPUs.
 - En el cas que una aplicació utilitzi MPI i CUDA/openCL, es veurà limitada l'escalabilitat en nombre de ranks MPI perquè si tots llencen workloads de gràfiques, hauran d'estar en nodes amb gràfiques disponibles.

En l'escenari de tenir més nodes amb targetes gràfiques, les avantatges i desavantatges passen a ser els inconvenients i les avantatges del cas anterior, respectivament.

Una diferència important entre NVIDIA i AMD és el model de programació amb el que es pot programar. NVIDIA té un model de programació propietari, anomenat CUDA [36]. És relativament senzill de programar amb aquest model. Per això, moltes aplicacions que estan optimitzades per a utilitzar gràfiques s'han programat amb CUDA i no amb OpenCL.

En canvi, OpenCL és un model de programació obert per sistemes heterogenis [37]. La diferència amb CUDA és que no només pot llançar workloads de feina a gràfiques, sinó que a qualsevol element de comput del sistema com altres cores o FPGAs (Field-Programmable Gate Array). A més, és més complicat de programar. Aquest model de programació paral·lela no és tant senzill d'utilitzar com CUDA i és menys utilitzat.

Les targetes gràfiques d'NVIDIA poden executar tant CUDA com OpenCL, però la performance amb OpenCL és inferior a la que es pot obtenir amb CUDA. En canvi, les GPUs d'AMD es poden programar amb OpenCL i no amb CUDA. Això representa una diferència important, ja que aplicacions que només estiguin programades amb CUDA no podran fer ús de GPUs AMD.

D'aquest component no farem cap preselecció perquè dependrà de diversos factors quina targeta gràfica escollim. Aquests factors són:

- Processador: en funció del processador tindrem més GFlops o menys, per tant els GFlops procedents d'acceleradors seran majors o menors, sempre mantenint-nos dins del límit de que màxim el 30% dels GFlops del clúster poden provenir d'acceleradors.

- Xarxa d'interconnexió: la quantitat de nodes que es puguin connectar ens permetrà tenir més o menys nodes de 2U, és a dir, nodes amb GPUs. Per exemple, si tenim una xarxa amb pocs ports i queden Us al rack sense connexions, es pot optar per utilitzar nodes que ocupin més espai (nodes amb gràfiques), aprofitant al màxim les connexions i l'espai.

3 Anàlisi de possibles configuracions

Arribats a aquest punt, hem deixat moltes eleccions de components al aire ja que depenen unes de les altres. Per a poder valorar-les correctament, hem decidit provar-les totes amb l'ajut de fulls de càlcul que es poden trobar a l'annex.

Les combinacions de components que tenim són les següents:

- Nodes single-socket (cpu 7702) o dual-socket (cpu 7502)
- Nodes amb o sense gràfiques
- 7 gràfiques diferents
- 3 commutadors diferents amb 4 possibles combinacions de xarxa

Per a cada una de les xarxes (nombre i tipus de commutador) s'ha estudiat quina és la combinació de nodes sense gràfica (1U) i amb gràfica (2U) que ens maximitza els TFLOPS, complint la restricció del percentatge dels TFLOPS que poden representar de les gràfiques. S'ha considerat que els nodes que tenen gràfica ocupen tots els possibles slots per aquestes, i així s'optimitza millor l'espai. En la següent taula es poden veure els resultats resumits (és a dir, per cada combinació de nodes dins la xarxa s'ha mostrat només la que otorga més TFLOPS), però si es vol veure totes les possibles combinacions es pot anar a l'annex.

		[Commutadors] x [Ports]			
		6x18	7x18	3x36	3x40
Single-Socket AMD 7702	Nodes 1U	72	71	75	75
	Nodes 2U	3	3	3	3
	Gràfiques	NVIDIA tesla V100	AMD Radeon Instinct MI50	NVIDIA tesla V100	NVIDIA tesla V100
	TFlops	213.594	208.345	219.594	219.594
	Preu (€)	847263	826255.8	892893.6	973131.6
	Gflops/€	0.258	0.258	0.252	0.231
Dual-Socket AMD-7502	Nodes 1U	74	73	77	77
	Nodes 2U	2	2	2	2
	Gràfiques	AMD Radeon Instinct MI50	NVIDIA tesla V100	AMD Radeon Instinct MI50	AMD Radeon Instinct MI50
	TFlops	270.46	258.16	277.96	277.96
	Preu (€)	983231.2	1007763	1034778.8	1034778.8
	Gflops/€	0.282	0.262	0.275	0.255

Taula 7: Mostra de les diverses combinacions valorades.

En primer lloc, analitzem la taula en el eix horitzontal. S'observa que els TFLOPS que ens dona utilitzar 3 commutadors de 36 ports i 3 de 40 són els mateixos. Això es perquè amb la primera opció ja tenim suficients ports com per emplenar els racks amb tots els nodes que hi caben. La millora d'utilitzar els commutadors de 40 ports és que tenen una velocitat més bona, 200Gbits en comptes de 100. El preu puja, però ens ho podem permetre.

Per altra banda, observem que amb 6 commutadors de 18 ports aconseguir una eficiència Gflops/€ més bona, tot i que no arribem a tants TFLOPS com amb altres opcions.

Analitzant la taula en l'eix vertical, veiem que per qualsevol configuració la versió amb Dual-Socket ens dona més TFLOPS amb una eficiència econòmica millor. Això pot ser degut en part a que els nodes amb gràfiques poden tenir-ne fins a 8 per node en comptes de 4 com en el cas del Single-Socket. Per tant, podem tenir més gràfiques utilitzant menys Us, deixant-ne més disponibles per a nodes sense gràfiques (1U).

4 Clúster final

4.1 Roofline

Apèndix

A hola

hi

Referències

- [1] Amd epyc 7702p: *Newegg Business*. [en línia]. Disponible a <https://www.neweggbusiness.com/product/product.aspx?item=9b-19-113-583>. Últim accés maig 2020.
- [2] Amd epyc 7502p: *Newegg Business*. [en línia]. Disponible a <https://www.neweggbusiness.com/product/product.aspx?item=9siv1dsa223\241&bri=9b-19-113-589>. Últim accés maig 2020.
- [3] H12dsu-in. Diponible a <https://www.supermicro.com/en/products/motherboard/H12DSU-iN>. Últim accés Maig 2020.
- [4] H12dst-b. Diponible a <https://www.supermicro.com/en/products/motherboard/H12DST-B>. Últim accés Maig 2020.
- [5] H12ssw-nt. Diponible a <https://www.supermicro.com/en/products/motherboard/H12SSW-NT>. Últim accés Maig 2020.
- [6] H12ssw-nt price. Diponible a https://www.lambda-tek.es/Supermicro-AS-1014S-WTRT~sh/B43663422&origin=gbaseES15.4?gclid=Cj0KCQjw-_j1BRDkARIsAJcfmTEzkltnhxZZGr08mxNuXdXE7JggwVY8yQ-5h5KjuhvcuzriN4lMWtYaAsw2EALw_wcB. Últim accés Maig 2020.
- [7] H12ssw-in. Diponible a <https://www.supermicro.com/en/products/motherboard/H12DSSW-iN>. Últim accés Maig 2020.
- [8] H12sst-ps. Diposnible a <https://www.supermicro.com/en/products/motherboard/H12SST-PS>. Últim accés Maig 2020.
- [9] H12sst-ps. Disponible a https://www.google.com/aclk?sa=l&ai=DChcSEwiujOfg1bXpAhVK-VEKHfiIBI8YABAGGgJ3cw&sig=AOD64_1yf0YWbq7mlwEoe-mRhdVSr_IBZA&ctype=5&q=&ved=0ahUKEwifg-Lg1bXpAhUQxYUKHZLjCWYQsEMIDg&adurl=. Últim accés Maig 2020.
- [10] Thinkmake. Disponible a <https://www.thinkmate.com/systems/servers>. Últim accés Maig 2020.
- [11] Kingston ksm29rs8. Disponible a <https://www.alternate.es/Kingston/KSM29RS8-8MEI-m%C3%B3dulo-de-memoria-8-GB-DDR4-2933-MHz-ECC-Memoria-RAM/html/product/1538082?>. Últim accés Maig 2020.

- [12] Kingston ksm29rs4. Disponible a <https://www.alternate.es/Kingston/KSM29RS4-16MEI-m%C3%B3dulo-de-memoria-16-GB-DDR4-2933-MHz-ECC-Memoria-RAM/html/product/1538086?> Últim accés Maig 2020.
- [13] Kingston ksm29rd4. Disponible a <https://www.alternate.es/Kingston/KSM29RD4-32MEI-m%C3%B3dulo-de-memoria-32-GB-DDR4-2933-MHz-ECC-Memoria-RAM/html/product/1538088?> Últim accés Maig 2020.
- [14] Kingston hyperx impact. Disponible a <https://www.pccomponentes.com/memoria-ram-kingston-hyperx-impact-sodimm-ddr4-3200mhz-8gb-cl20>. Últim accés Maig 2020.
- [15] Kingston hyperx impact. Disponible a <https://www.pccomponentes.com/memoria-ram-kingston-hyperx-impact-sodimm-ddr4-3200mhz-16gb-cl20>. Últim accés Maig 2020.
- [16] Kingston hyperx fury black. Disponible a <https://www.pccomponentes.com/kingston-hyperx-fury-black-ddr4-3200mhz-32gb-cl16>. Últim accés Maig 2020.
- [17] Mellanox technologies. en [en línia]. Disponible a <https://www.mellanox.com/>. Últim accés Maig 2020.
- [18] Mellanox msx6012f-2bfs switchx-2 based fdr infiniband. en [en línia]. Disponible a <https://store.mellanox.com/products/mellanox-msx6012f-2bfs-switchx-2-fdr-infiniband-1u-switch-12-qsfp-ports-2-power>.html. Últim accés Maig 2020.
- [19] Mellanox msx6018f-1sfs switchx-2 fdr infiniband. en [en línia]. Disponible a <https://store.mellanox.com/products/mellanox-msx6018f-1sfs-switchx-2-fdr-infiniband-1u-switch-18-qsfp-ports-1-power>.html. Últim accés Maig 2020.
- [20] Mellanox msb7800-es2f switch-ib 2 based edr infiniband. en [en línia]. Disponible a <https://store.mellanox.com/products/mellanox-msb7800-es2f-switch-ib-2-based-edr-infiniband-1u-switch-36-qsfp28-ports>.html. Últim accés Maig 2020.
- [21] Mellanox mqm8700-hs2f quantum hdr infiniband. en [en línia]. Disponible a <https://store.mellanox.com/products/mellanox-mqm8700-hs2f-quantum-hdr-infiniband-switch-40-qsfp56-ports-2-power-supply>.html. Últim accés Maig 2020.

- [22] Nvidia tesla v100: *TechPowerUp*. en [en línia]. Disponible a <https://www.techpowerup.com/gpu-specs/tesla-v100-pcie-16-gb.c2957>. Últim accés Maig 2020.
- [23] Nvidia rtx 2080 ti: *TechPowerUp*. en [en línia]. Disponible a <https://www.techpowerup.com/gpu-specs/geforce-rtx-2080-ti.c3305/>. Últim accés Maig 2020.
- [24] Nvidia quadro rtx 8000: *TechPowerUp*. en [en línia]. Disponible a <https://www.techpowerup.com/gpu-specs/quadro-rtx-8000.c3306>. Últim accés Maig 2020.
- [25] Nvidia tesla t4: *TechPowerUp*. en [en línia]. Disponible a <https://www.techpowerup.com/gpu-specs/tesla-t4.c3316>. Últim accés Maig 2020.
- [26] Amd radeon wx 8200: *TechPowerUp*. en [en línia]. Disponible a <https://www.techpowerup.com/gpu-specs/radeon-pro-wx-8200.c3303>. Últim accés Maig 2020.
- [27] Amd radeon instinct mi50: *TechPowerUp*. en [en línia]. Disponible a <https://www.techpowerup.com/gpu-specs/radeon-instinct-mi50.c3335>. Últim accés Maig 2020.
- [28] Amd radeon instinct mi25: *TechPowerUp*. en [en línia]. Disponible a <https://www.techpowerup.com/gpu-specs/radeon-instinct-mi25.c2983>. Últim accés Maig 2020.
- [29] Nvidia tesla v100: *NewEggBussiness*. en [en línia]. Disponible a <https://www.neweggbusiness.com/product/product.aspx?item=9siv13zb8p1031>. Últim accés Maig 2020.
- [30] Nvidia rtx 2080 ti: *NewEggBussiness*. en [en línia]. Disponible a <https://www.neweggbusiness.com/product/product.aspx?item=9siv18m9g85353>. Últim accés Maig 2020.
- [31] Nvidia quadro rtx 8000: *NewEggBussiness*. en [en línia]. Disponible a <https://www.neweggbusiness.com/product/product.aspx?item=9b-1ft-000p-003f5>. Últim accés Maig 2020.
- [32] Nvidia tesla t4: *NewEggBussiness*. en [en línia]. Disponible a <https://www.neweggbusiness.com/product/product.aspx?item=9siv06watj0315>. Últim accés Maig 2020.

- [33] Amd radeon wx 8200: *NewEggBussiness*. en [en línia]. Disponible a <https://www.neweggbusiness.com/product/product.aspx?item=9b-14-105-110>. Últim accés Maig 2020.
- [34] Amd radeon instrinct mi50: *Lambda-Tek*. en [en línia]. Disponible a <https://www.lambda-tek.es/AMD-100-506143~sh/B43704341>. Últim accés Maig 2020.
- [35] Amd radeon instrinct mi25: *Lambda-Tek*. en [en línia]. Disponible a <https://www.lambda-tek.es/HPE-Q1K38A~sh/B42127391>. Últim accés Maig 2020.
- [36] Cuda: *NVIDIA*. en [en línia]. Disponible a <https://developer.nvidia.com/cuda-zone>. Últim accés Maig 2020.
- [37] Opencl: *Khronos*. en [en línia]. Disponible a <https://www.khronos.org/opencl/>. Últim accés Maig 2020.