

Recerca i supercomputació al segle XXI

Mateo Valero

Director del Barcelona Supercomputing Center-
Centre Nacional de Supercomputació

La supercomputació evoluciona gràcies a diversos factors. El primer ha estat la invenció del tub de buit. El segon és aconseguir que els transistors siguin cada cop més petits. Durant seixanta anys, els tecnòlegs han estat capaços de duplicar la densitat dels transistors cada divuit mesos i, actualment, ho aconsegueixen fer cada dos anys. I el tercer factor és la capacitat que tenen els arquitectes de computadors de duplicar la velocitat dels processadors. Tots aquests elements han fet que, actualment, el processador de qualsevol telèfon mòbil sigui més potent que qualsevol dels processadors que hi havia al món quan l'home va arribar a la Lluna.

Els arquitectes de computadors utilitzem les tècniques que va utilitzar Henry Ford per a fabricar cotxes: una cadena de muntatge. El fabricant de la companyia Ford Motor Company volia fer cotxes i va dividir la fabricació de cotxes en etapes. Així, una cadena podia tenir un cotxe en cada etapa. Nosaltres fem els processadors d'aquesta manera. Quan diem que un Pentium funciona a 3 GHz vol dir que s'ha aconseguit que les instruccions viatgin d'una etapa a una altra a tres mil milions d'operacions per segon i que els processadors tinguin la capacitat d'executar quatre instruccions alhora. Un Pentium d'aquestes característiques, que funciona a 3 GHz i executa quatre instruccions per cicle, realitza unes 10^{10} operacions per segon.

Els experts que es dediquen a plegar proteïnes a través de la simulació necessiten 10^{22} operacions per plegar una proteïna. Un processador normal trigaria 10^5 anys per acabar tots els càlculs que es necessiten. Per això, encara que els processadors vagin molt ràpid, hi ha aplicacions que requereixen una potència molt més gran que la que té un processador. Fer una pel·lícula digital és fàcil perquè el nombre d'operacions que s'ha de realitzar es pot dividir fàcilment. Però, en canvi, hi ha aplicacions en les quals els processadors han d'intercanviar resultats parcials cada mil·lèsima de segon per a poder avançar en l'algoritme. I per a fer-ho es necessita una potència de càlcul i un mecanisme que permeti que els processadors es comuniquin a una velocitat molt alta. Això ho permeten superordinadors com el MareNostrum, que té deu mil processadors i es comporta com un processador que és deu mil vegades més ràpid. Aquests superordinadors són màquines molt ràpides que estan construïdes amb molts processadors, molt ràpids en si mateixos però que, gràcies a un ràpid mecanisme de comunicació entre ells, s'aconsegueix, si es programa de manera adequada, tenir un sol processador avançat segons la llei de Moore.

L'any 1994, un processador podia fer una operació per segon, actualment tenim la capacitat de fer dos-cents vuitanta bilions d'operacions per segon (280 Tflop/s). Encara que la potència de càlcul hagi augmentat molt, l'única forma de crear superordinadors és col·locant molts processadors i fent que treballin conjuntament. El BlueGene és el supercomputador més ràpid del món, amb el qual nosaltres, des de Barcelona, hem treballat en cinc projectes amb IBM. En cada xip del BlueGene hi ha un parell de processadors, col·locats en targetes i en seixanta-quatre armaris, que tenen en el seu conjunt una potència de càlcul en punta de 367 Tflop/s.

Aquests supercomputadors permeten executar les aplicacions que els investigadors necessiten per a investigar en el funcionament dels sistemes vius i trobar nous fàrmacs i nous materials, dissenyar millor els avions i els cotxes, estudiar l'evolució de l'Univers i predir el canvi climàtic. Junt amb la teoria i l'experimentació, la supercomputació constitueix el tercer suport de la ciència i de l'enginyeria. Per a dissenyar avions, per exemple, necessitem la teoria, així com l'experimentació, però actualment es fa imprescindible la supercomputació. Gràcies a aquest suport, es poden simular els projectes abans de veure'ls i els científics poden millorar la seva ciència i predir-ne els resultats. Un exemple és el projecte Blueblein, que s'està realitzant des de Suïssa, i que intenta descodificar l'enginyeria del cervell; en aquest projecte s'utilitza el supercomputador per simular com funciona el cervell. La col·laboració entre la teoria, l'experiment i la simulació permet l'obtenció de resultats que es realimenten i permeten fer avançar la ciència.

El nostre grup de recerca va començar als anys vuitanta, a Barcelona, desenvolupant, amb els nostres recursos, multiprocessadors. Dissenyàvem plaques amb 8-12 microprocessadors amb els quals vam guanyar algun concurs. Gràcies a aquesta recerca i a la qualitat del grup, empreses com Intel o Compact es van establir a Barcelona i van donar-nos suport per

a continuar desenvolupant arquitectura de computadors d'una manera multidisciplinària. El nostre grup treballa tots els nivells del programari, des de l'aplicació fins a l'últim bit. Dissenyem xips, targetes, sistemes petits i intermedis i fa quinze anys que treballem en un concepte, que per a nosaltres ja és antic, com el GRID.

Actualment, som cent vint persones treballant a la Universitat Politècnica de Catalunya, al Barcelona Supercomputing Center-Centre Nacional de Supercomputació (BSC-CNS). Un equip constituït per catedràtics d'universitats espanyoles, europees i americanes, que atrau investigadors d'arreu del món i que ha permès la formació de quatre-cents doctors. Actuem, així, com a pols d'atracció de nous cervells i investigadors.

El grup ha motivat a Europa la recerca en arquitectura de computadors, que era inexistent, i actualment coordina una xarxa d'excel·lència en aquesta àrea on es desenvolupa no només recerca, sinó diferents activitats, com l'Escola d'Estiu, que durant una setmana reuneix dos-cents doctorands europeus amb els millors professors i experts mundials en supercomputació. Desenvolupem i participem en molts projectes europeus i investiguem amb les empreses més importants del món del maquinari i del programari. Per exemple, vam ser el primer equip que va treballar en arquitectura de computadors amb Microsoft.

L'any 1984 vam aconseguir que la Generalitat de Catalunya, el Ministeri d'Educació i Ciència i la Universitat Politècnica de Catalunya (UPC) creessin el primer centre espanyol en recerca de tècniques de computadors paral·lels, el CEPBA. Gràcies a tota la recerca que vam desenvolupar, vam aconseguir associar-nos amb IBM i crear el CEPBA-IBM Research Institute (CIRI), un centre de recerca mixt que va treballar a la UPC durant quatre anys i que ens va permetre investigar a un alt nivell en arquitectura de computadors i programari de computadors paral·lels.

El CIRI va catalitzar més de quaranta milions d'euros en projectes europeus i la nostra recerca va permetre la difusió en l'àmbit europeu de la cultura del paral·lelisme, i vam coordinar durant tres anys diferents projectes: el PCI-PACOS (Parallel Computing Initiative Spain). Vam col·laborar amb més de cinquanta empreses europees, espanyoles i catalanes de diversos àmbits. Amb Ferrari, per exemple, vam treballar en tècniques de simulació perquè el vehicle anés molt ràpid, però que no volés. Volkswagen, Hesperia i Iberdrola són altres empreses amb les quals vam col·laborar.

A Catalunya, també va ser molt important la iniciativa de la Generalitat de Catalunya, que va crear l'any 1991 el Centre de Supercomputació de Catalunya (CESCA), com un centre de servei i com una eina per a millorar la recerca dels investigadors catalans. De l'any 1994 al 1998 la Generalitat va coordinar el CESCA i el CEPBA, i es va crear el C4, que va estar actiu durant quatre anys.

Com a conseqüència de l'experiència, de la recerca i de la transferència de tecnologia, va néixer la idea amb IBM de construir a Barcelona un supercomputador, el MareNostrum, que alberga el Barcelona Supercomputing Center-Centre Nacional de Supercomputació,

creat l'any 2005 pel Ministeri d'Educació i Ciència, la Generalitat de Catalunya i la UPC, les mateixes institucions que van crear el CEPBA.

L'avantatge del MareNostrum respecte als altres superordinadors existents és que es va construir amb tecnologia *commodity*, que permet comprar les peces de manera senzilla, com si anéssim a comprar-les a un supermercat, i muntar-les. Volíem que aquesta màquina estigués entre els cinc primers llocs del món i que passés el test dels Top 500 el novembre del 2004. Aquest test el fan tots els supercomputadors dues vegades a l'any per conèixer la velocitat de càlcul i se'ls fa resoldre un sistema d'equacions.

El MareNostrum va costar dotze milions d'euros. IBM ens el va vendre a un preu molt econòmic per la relació existent durant tants anys. El preu incloïa tres anys de manteniment gratuït.

El BSC-CNS està integrat al mapa espanyol d'Infraestructures Científiques i Tecnològiques singulars, i és l'única que tenim a Catalunya, fins que arribi el sincrotró. Com a gran instal·lació té un reglament d'ús. Hi ha un comissió formada per quaranta-quatre experts de quatre àmbits de la ciència que reben tots els projectes que sol·liciten utilitzar el MareNostrum i avaluen i decideixen quins projectes utilitzen el supercomputador.

El MareNostrum, ubicat a Torre Girona, una capella desacralitzada, va començar oficialment el 25 d'abril de l'any 2004 amb vint-i-cinc projectes i, actualment, en aquest centre treballem 175 investigadors en els camps de la ciència de computadors, ciències de la vida i ciències de la Terra.

El MareNostrum té vuit gigabytes de memòria, dos discos i 2.580 *blades* (plaques) que es comuniquen entre si. Cada *blade* o placa està composta per dos processadors. De cada *blade* surt una fibra òptica que es connecta amb una sèrie de *racks*, on hi ha centrals de telefònica que permeten que qualsevol parella de processadors pugui intercanviar informació en quatre microsegons i a una velocitat de quatre gigabytes per segon. És el prototip més gran de xarxa d'interconnexió del món. Diversos mòduls de 1.024 processadors es connecten a un primer nivell i un segon nivell permet que qualsevol processador es pugui connectar amb qualsevol altre.

El MareNostrum és un processador deu mil vegades més ràpid que un ordinador de taula, i en algunes aplicacions s'aconsegueix anar encara més ràpid. El MareNostrum té cinc petabytes (PB) de cintes, que el converteix en el magatzem més gran d'informació d'Europa. Alguns investigadors, com els que dissenyen les ales dels avions, produeixen moltes dades que guardem al MareNostrum. Aquests cinc PB és l'1 % de tota la informació que es va produir al món fa tres anys, segons un estudi elaborat a la Universitat de Berkeley.

El MareNostrum està instal·lat dins d'una urna de vidre de 16 x 9 metres. Sota l'urna, estan instal·lats tots els cables d'alimentació, els tubs de l'aigua i les xarxes de connexió. Fora de la capella hi ha diversos transformadors d'1,5 megabytes, que el mantenen a la temperatura necessària. Una de les coses que pot fer que no hi hagi superordinadors més potents en el

futur serà el cost energètic d'aquestes màquines. L'han qualificat com el supercomputador més maco del món, ha aollit filmacions de pel·lícules i l'han visitat més de dotze mil persones.

Al test Top 500 del mes de novembre del 2004, el MareNostrum va ser el primer d'Europa i el quart del món. Des de llavors ha estat el superordinador més ràpid en un centre públic de recerca. El ordinadors que se situen en els primers llocs estan instal·lats en centres privats, algun dels quals és militar. En l'últim rànquing que s'ha fet, el MareNostrum se situa com el cinquè superordinador més ràpid del món.

Els superordinadors són una manera d'amplificar la llei de Moore. El arquitectes de computadors hem duplicat la velocitat dels processadors i quan n'ajuntem molts és com si hi hagués una nova llei de Moore. En dades, l'evolució de les tecnologies fa que el superordinador que ocupa el primer lloc del Top 500, en sis o set anys baixi al número cinc-cents de la llista. El nostre ordinador personal de la versió 2006 hauria estat l'any 1994 el supercomputador número cinc-cents de la llista. I això continua evolucionant.

La recerca del BSC-CNS s'estructura en les àrees d'astronomia, espai i ciències de la Terra; biomedicina i ciències de la vida; física i enginyeria, i química i ciència i tecnologia dels materials. El MareNostrum ha donat suport a uns dos-cents projectes i ha assolit un nivell d'ocupació del 100 %.

Des del començament, el superordinador ha estat connectat a la xarxa europea DEISA, que connecta els onze supercomputadors més ràpids a través de la Red Geant. Aquesta xarxa permet que qualsevol investigador europeu pugui utilitzar qualsevol dels onze supercomputadors fàcilment.

El BSC-CNS atrau investigadors d'arreu del món. Des de l'any 1991, més de quatre-cents investigadors europeus a Barcelona han fet estades de tres mesos i col·laboren amb els nostres investigadors. A més, actualment, més de vint-i-cinc investigadors del centre que no han estat formats a Espanya formen part del planter del centre.

Des del BSC-CNS, també contribuïm a integrar Espanya en les grans infraestructures que vertebraran la recerca europea en el futur. Dins del setè programa marc, el MareNostrum és el representant espanyol i està en bona posició per a ser una de les seus, juntament amb França, Alemanya i Gran Bretanya. L'objectiu és cada quatre anys, i amb finançament europeu, canviar i millorar un dels supercomputadors.

El MareNostrum es va crear el novembre de 2004 i ara fa uns mesos, el novembre del 2006, s'ha renovat amb la introducció de dos processadors en cada una de les 2.560 plaques que incorpora, la qual cosa li ha conferit més potència de càlcul. El nou MareNostrum ha costat onze milions d'euros i és el doble de ràpid que el que vam comprar l'any 2004. D'aquesta manera, en poc temps hem tingut dos MareNostrum i per vint-i-tres milions d'euros. Els components substituïts del MareNostrum han tornat a tenir vida en sis centres de supercomputació espanyols (Madrid, Cantàbria, Saragossa, València, Màlaga i Canàries), que integren la xarxa espanyola de supercomputació, creada pel Ministeri d'Educació i

Ciència per donar suport a la comunitat científica. El BSC-CNS serà el coordinador d'aquesta xarxa. Amb la renovació, el BSC-CNS és el primer centre que ha tingut set entrades al Top 500.

En l'àrea de l'arquitectura de computadors, dirigida pel professor Jesús Labarta, estem investigant amb Microsoft el disseny del programari i el maquinari paral·lel que s'executarà en els futurs computadors personals. Aquests tindran d'aquí a deu anys centenars de processadors molt ràpids en xips de 4 cm² i seran tan potents com qualsevol dels trenta-un armaris del MareNostrum. IBM i Intel han garantit crear cada divuit mesos processadors més ràpids. Però això ja no és possible, perquè els arquitectes de computadors ja no sabem fer processadors més ràpids. Davant d'aquesta nova situació, hi ha diverses solucions. Una és multiplicar. On abans posàvem un processador, ara en posem dos, quatre, vuit o setze. Aquesta nova realitat planteja a empreses com Microsoft preparar el seu programari per poder executar les aplicacions. Amb aquesta finalitat, el BSC-CNS ha signat un acord d'un milió de dòlars amb Microsoft per a investigar com el seu programari paral·lel, el del futur, en el qual treballen més de quinze mil persones, pot executar-se en plataformes *multicore*, amb més d'un processador dins del xip. Amb IBM treballem en el disseny d'un supercomputador cent mil vegades més potent que el MareNostrum. Tindria cent petaflops i podria estar llest en quatre anys.

Per al projecte Desafío Español Copa América 2007, estem desenvolupant programes per fer prediccions locals de vent amb molta precisió. Aquests programes són útils per als esports de vela, però també per a la planificació de la producció elèctrica dels parcs eòlics. Amb l'empresa Gas Natural s'ha realitzat un estudi de l'impacte ambiental sobre la qualitat de l'aire produït pel trànsit de ciutats com Barcelona i Madrid. Amb Airbus ajudem a dissenyar millor l'aerodinàmica dels futurs avions i amb Repsol fem simulacions amb el superordinador per trobar petroli al golf de Mèxic.

En l'àmbit de les ciències de la vida treballa un grup de quaranta persones liderades pel professor Modesto Orozco. La recerca de nous fàrmacs o el plegat de proteïnes són alguns dels treballs de l'equip, entre els quals destaca la creació d'una base de dades de dues-centes proteïnes no només amb imatges sinó també amb dinamisme. Aquest mapa és la base per a tota la comunitat científica que treballa en la recerca de nous fàrmacs.

En l'àrea de les ciències de la Terra, liderat pel professor José María Baldasano, s'han realitzat models de la qualitat de l'aire. Es poden consultar les previsions de la qualitat de l'aire per als propers tres dies a Barcelona i a Madrid. I no fallen. S'han realitzat també models de la dispersió de la sorra del Sàhara, per a la Unió Europea, i s'estan executant aplicacions per conèixer el model de canvi climàtic dels propers cinquanta anys.

Una altra solució a la impossibilitat de crear processadors més ràpids és orientar l'ús dels transistors, com fa el Cell. Un xip que té IBM fa dos anys, de 2,35 cm², amb 235 milions de transistors i nou processadors, un de general i vuit acceleradors específics. Sony utilitza

aquests xips per a la Playstation 3 (el desenvolupament d'aquest xip va costar un milió d'euros) i Toshiba el fa servir per a la televisió d'alta definició.

Intel també ha anunciat la creació d'un xip de vuitanta processadors que comercialitzarà d'aquí a quatre o cinc anys. Un dels problemes del xip és que, encara que tingui molta potència de càlcul, l'has d'alimentar. Per solucionar-ho, s'han creat tecnologies en 3D que permeten que la memòria per a alimentar els xips sigui vertical.

La barrera del teraflop de 10^{12} operacions en un xip de 4 cm^2 va ser superada per primera vegada l'any 1991, amb deu mil pentiums que tenien vuitanta-cinc armaris i ocupaven 200 m^2 . Aquesta potència, gràcies a la recerca, la tindrem als nostres ordinadors personals en quatre o cinc anys.

Els supercomputadors del futur seran com el MareNostrum però amb molts més processadors. Això complicarà més la programació. Per això, tots els investigadors que tinguin un programari per a setze processadors i no facin un esforç per fer servir un programari paral·lel, veuran com el seu programari anirà més lent per falta de potència.

Una altra solució a la impossibilitat de crear processadors més ràpids serà combinar processadors opteron generals amb acceleradors. IBM està construint una màquina per al Los Alamos National Laboratory que combina processadors opteron i xips Cell.

La supercomputació ha canviat la manera de fer ciència. Abans els experiments s'apuntenaven en un paper, actualment qualsevol recerca produeix una quantitat molt gran de dades i d'informació, que s'ha de processar i es pot emmagatzemar gràcies a la supercomputació. A més, el treball d'equips multidisciplinaris, d'altres àmbits de la ciència i d'arquitectura de computadors, permeten millorar la ciència. Actualment, el BSC-CNS treballa amb enginyers de Madrid per a desenvolupar un *software* per nanotecnologia. És important que treballin junts els experts en algorismes i els que controlen l'arquitectura per a crear un programari que pugui utilitzar un ampli espectre d'experts en nanotecnologia. A més, tenim projectes amb vint-i-tres grups espanyols per desenvolupar conjuntament tècniques informàtiques, en els quals hi ha intercanvi d'informació que permet que el resultat del treball conjunt interdisciplinari sigui diferent.

Experts de renom internacional consideren que la supercomputació ha assolit actualment el paper que les matemàtiques van tenir per a desenvolupar la física. Per això, qualsevol àrea experta necessita el suport de la supercomputació, i si no desenvolupa noves tècniques basades en informàtica, no podrà avançar.

En aquest context, el MareNostrum ja s'ha convertit en un centre de creació de recerca i un pol d'atracció de la recerca al sud d'Europa.