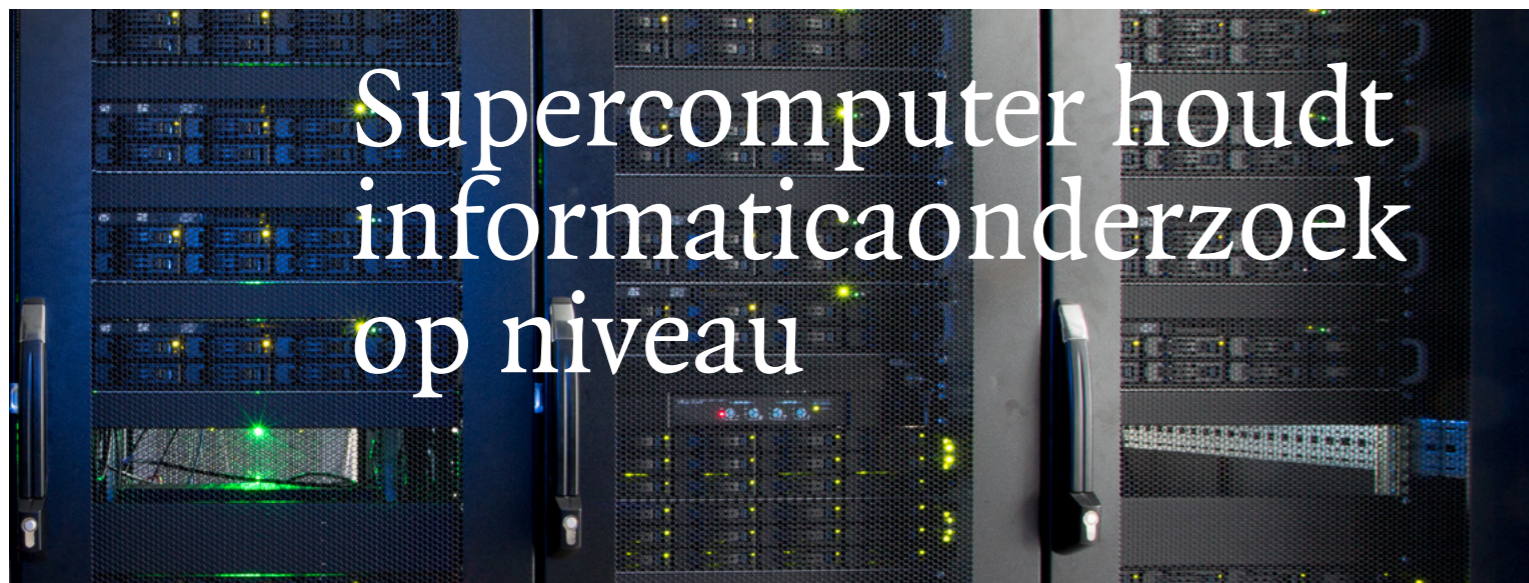


Al meer dan honderd proefschriften verlieten de drukker dankzij de supercomputer DAS. De nieuwste versie, DAS-5, zorgt ervoor dat het Nederlandse informaticaonderzoek op internationaal topniveau blijft. De nadruk van het onderzoek zal liggen bij big data, cloud computing en many core processoren. Door Enith Vlooswijk



Henri Bal zet een Amsterdams accent op als hij Cruijf napraat: 'Elk nadeel heb z'n voordeel.' De hoogleraar Computersystemen van de Vrije Universiteit heeft het over de relatief kleine omvang van de Nederlandse informatica. Deze maakt het mogelijk om het computersysteem voor de experimenten elke vier jaar compleet opnieuw op te zetten. 'Het Grid5000-systeem in Frankrijk is tien keer zo groot. Het zou veel te duur zijn om dat eens in de vier jaar in zijn geheel te vervangen.'

Henri Bal: 'Je kunt DAS-5 zien als een experimenteerspeeltuin voor informatici'



Subsidie voor apparatuur De supercomputer DAS-5 is grotendeels bekostigd vanuit NWO-Middelgroot, een subsidie bedoeld voor investeringen in vernieuwende apparatuur of dataverzamelingen van internationaal belang. Daarnaast betaalden deelnemende universiteiten en onderzoeksinstellingen mee aan DAS-5. Dit zijn de Vrije Universiteit, de Universiteit van Amsterdam, de Technische Universiteit Delft, de Universiteit Leiden, Astron, MultimediaN, COMMIT/ en het eScience Center. Samen maken zij deel uit van ASCI (Advanced School for Computing and Imaging). Het optische netwerk tussen de zes clusters is door SURFnet opgezet.

Dat geldt niet voor DAS, de gedistribueerde supercomputer waar het Nederlandse informaticaonderzoek mede zijn hoge niveau aan dankt. Afgelopen zomer werd alweer de vijfde versie in gebruik genomen. Het systeem bestaat uit tweehonderd machines, verspreid over zes clusters. Qua omvang is dat niet indrukwekkend – de top 500 supercomputers ter wereld bevat systemen van honderden keren zo groot. Toch is DAS-5 uniek in de wereld, vertelt Bal. 'In de meeste landen liften informaticaonderzoekers mee op supercomputers die nooit zijn bedoeld voor informaticaonderzoek. Ons systeem is daar juist specifiek voor gemaakt. Onderzoekers kunnen bijvoorbeeld wijzigingen aanbrengen in de hardware voor hun experiment; applicatiegebruikers willen juist een stabiel systeem.'

Dat dit verschil essentieel is, beaamt ook Alexandru Iosup, universitair hoofd-docent aan de Technische Universiteit Delft. Zonder de supercomputer zou veel van zijn onderzoek niet mogelijk zijn.

'Informatica is in steeds grotere mate een experimentele wetenschap,' vertelt hij. 'Bij mijn onderzoek behelst het aan de praat krijgen van een technologie de helft van het project. Hoe maak je een softwaresysteem efficiënt en schaalbaar en werkt de theorie ook in de praktijk? DAS-5 biedt een heel flexibele onderzoeksomgeving, waar we allerlei gedistribueerde experimenten kunnen uitvoeren.'

Revolutie

Bij de samenstelling van DAS5 hadden de ontwerpers specifieke onderzoeksrichtingen voor ogen. 'Het informaticalandschap verandert heel snel,' vertelt Bal. 'Elke nieuwe DAS-generatie moet aansluiten bij de behoeften van dat moment. De laatste jaren voltrekt zich een revolutie op het gebied van cloud computing en processoren. Er zijn al many core processoren met duizenden

rekenkernen en supercomputers die miljoenen dingen tegelijk kunnen doen. Het probleem is: de software die erop draait is hier meestal niet geschikt voor, dus hoe programmeer je die? Vergelijk het met het managen van duizenden mensen, in plaats van een paar. Dat is niet gemakkelijk. Daar richt ons onderzoek zich op.'

Ook is het systeem zo gebouwd, dat het veeleisend onderzoek op het gebied van big data ondersteunt. DAS-5 heeft meer geheugen en diskruimte om niet alleen zwaar rekenwerk, maar ook grote datastromen te verwerken. Dat is bijvoorbeeld nodig bij onderzoek naar efficiënte graph processing tools: technologieën om op grote schaal relaties tussen verschillende eenheden te analyseren. Het is een van de onderzoeksvelden van Iosup.

'Graphs zijn relaties, bijvoorbeeld sociale relaties tussen mensen, maar ook relaties tussen moleculen,' legt hij uit. 'De meest gebruikte tools voor big data werken op dit gebied echter erg slecht. Wij zoeken naar nieuwe tools, waarmee we extreem grote hoeveelheden informatie kunnen verwerken op een efficiënte manier. Zodat graph processing niet alleen voor de Facebooks op deze wereld beschikbaar komt, maar ook voor kleine en middelgrote bedrijven in Nederland.'

Alexandru Iosup: 'DAS-5 biedt een heel flexibele onderzoeksomgeving, waar we allerlei gedistribueerde experimenten kunnen uitvoeren'



DAS-5 is beschikbaar voor alle leden van ASCI (Advanced School for Computing and Imaging) die aan de supercomputer hebben bijgedragen, plus hun onderzoekspartners. 'Je kunt DAS-5 zien als een experimenteerspeeltuin voor informatici', zegt Bal. 'Dit in tegenstelling tot de grote Nederlandse supercomputer bij SURFsara, waar bijvoorbeeld veel klimaatonderzoekers dagenlang hun modellen op laten lopen. Wij zoeken wel naar manieren om simulaties te versnellen, die vervolgens kunnen worden toegepast bij SURFsara.'

Honderd proefschriften

In totaal kostte DAS-5 ongeveer 1,4 miljoen euro. In vergelijking met andere supercomputers is dat een heel bescheiden prijs, zegt Bal trots. 'De afgelopen versies hebben al meer dan honderd proefschriften opgeleverd. Zonder DAS hadden al die promovendi hun eigen hardwaresysteem moeten opzetten. Alleen al het systeembeheer zou veel duurder zijn geweest dan alle DAS-systemen samen.'

Het aantal onderzoeksprojecten op DAS stijgt de laatste jaren flink. Bij de oplevering lag er al een lijst klaar van 50 projecten van promovendi en postdocs. Ook masterstudenten profiteren van DAS-5, vertelt Iosup. 'Ze kunnen met de supercomputer krachtige experimenten doen die vaak resulteren in goede publicaties. DAS maakt het voor ons mogelijk om internationaal succesvol te wedijveren.' I/O

DAS-5

De nieuwste supercomputer van de onderzoeksschool ASCI bestaat uit tweehonderd machines, verspreid over zes clusters die zijn geplaatst in vijf onderzoeksinstellingen: ASTRON, de Technische Universiteit Delft, de Universiteit van Amsterdam, de Universiteit Leiden, en de Vrije Universiteit. Tussen de verschillende computerclusters zijn supersnelle dataverbindingen aangebracht. Elk cluster heeft een 1 gigabit per seconde verbinding met het reguliere internet. Daarnaast bestaat de mogelijkheid om via SURFnet lichtpaden op te zetten: rechtstreekse dataverbindingen tot 100 gigabit per seconde tussen twee punten in het netwerk, buiten het reguliere internetverkeer om.

Meer informatie: <http://www.cs.vu.nl/das5/>

