

Programma Computational Science

**NWO Exacte Wetenschappen
Postbus 93138
2509 AC Den Haag
tel. 070-3440.811
fax 070-3440.861
e-mail: kas@nwo.nl**

23 mei 2000

0. Samenvatting

Computational Science is het wetenschapsgebied dat zich bezighoudt met

- de opslag en bewerking van gegevens met behulp van computers ten behoeve van praktische wetenschappelijke problemen op het gebied van grootschalig rekenen;
- het onderzoeken en ontwikkelen van generieke methoden en technieken voor combinatorische bewerkingen en rekenmodellen van complexe wetenschappelijke fenomenen.

Kenmerkend voor Computational Science is de *wisselwerking* tussen beide aspecten, in dit rapport aangeduid als *applicatiegericht* en *funderend* onderzoek. Doel van het programma is deze wisselwerking in het Nederlandse Computational Science-onderzoek te versterken door middel van een gerichte investering. Het programma beoogt in wetenschappelijk opzicht significante voortgang te bewerkstelligen bij het bestuderen, begrijpen en voorspellen van het gedrag van complexe (natuurlijke) systemen door middel van de ontwikkeling en het gebruik van gekoppelde dynamische modellen.

Het programma verschaft financiering voor de uitvoering van een aantal grote multidisciplinaire projecten (subsidie tot maximaal Mf 1,5) en een aantal kleinere projecten (subsidie tot maximaal Mf 0,75). Bij de kleinere projecten zal ongeveer de helft hoofdzakelijk gericht zijn op applicaties en de andere helft op funderend onderzoek. Voor programmabrede activiteiten (de organisatie van workshops en de inrichting van een software-uitwisselings-platform) wordt maximaal Mf 1 uitgetrokken. Het totale benodigde budget is Mf 15 tot 20, waarvan voor de periode tot en met 2004 via de begroting van het Gebiedsbestuur Exacte Wetenschappen thans tussen de Mf 7 en Mf 8,5 beschikbaar is.

Het programma zal een looptijd hebben van zes tot acht jaar. De te subsidiëren onderzoeksvoorstellen zullen in twee rondes ingewacht en geselecteerd worden. De eerste oproep voor indiening van voorstellen wordt in de zomer van 2000 verwacht. De tweede oproep zal niet eerder dan in 2002 gedaan worden.

Het programma is mede geïnspireerd door het prioriteitsprogramma Massaal Parallel Rekenen (MPR) en het onderzoek gefinancierd door de stichting High Performance Computing and Networking (HPCN). Het programma wil fungeren als een forum voor de verdere ontwikkeling van Computational Science. Het zal in eerste instantie gericht zijn op de disciplines Astronomie, Fysica, Informatica en Wiskunde. Samenwerking met andere disciplines (zoals Chemie en Biologie) en met nationale en internationale initiatieven op aangrenzende gebieden (zoals Bio(moleculaire) Informatica) behoort echter tot de mogelijkheden.

1. Karakteristiek van het vakgebied Computational Science

1.1 Achtergrond

Voor de oplossing van complexe problemen is er naast theorie en experiment altijd behoefte geweest aan andere wetenschappelijke methoden, zoals simulatie (dit is het onder gecontroleerde condities nabootsen van (aspecten van) de werkelijkheid, bijvoorbeeld aan waterbouwkundige modellen ten behoeve van de Deltawerken), numerieke analyse, statistische beschrijving en testen, en (recentelijk) beeldbewerking en visualisatie. Onderzoekers hebben vanaf het begin de computer gezien als een gereedschap bij het gebruik van deze methoden. Zo werd de eerste geheel elektronische computer (de ENIAC) midden jaren '40 gebouwd met als enig doel de baan van een afgeschoten granaat nauwkeurig te kunnen berekenen.

De technologische push

Nieuwe rekenkundige ontwikkelingen, zoals in vector- of parallel computing, leiden tot nieuwe oplosmethodieken en de daarbij behorende algoritmen. Deze leiden weer tot mogelijkheden voor nieuwe toepassingen, zowel in de wetenschap als in maatschappelijke sectoren, zoals de industrie.

In feite werd daarmee een model gemaakt van de fysische grootheden die de granaatbaan bepalen. De beschikbaarheid van de computertechnologie gaf onderzoekers de *push* dit type modellen te maken.

Ingewikkelder verschijnselen vergden echter aanzienlijk meer rekenkracht. Van het onderzoek ging dan ook een sterke *pull* uit naar de ontwikkeling van computertechnologie waarmee beter aan de wetenschappelijke behoefte voldaan kon worden. Deze wisselwerking tussen technologische *push* en wetenschappelijke *pull* is kenmerkend voor de ontwikkeling van het grootschalig rekenwerk.

De wetenschappelijke pull

De ontwikkelingen aan de *pull*-kant verlopen langs een keten met als schakels:

1. analytische hulpmiddelen schieten tekort voor de verificatie van een theorie;
2. toepassingen van bestaande of ontwikkeling van nieuwe technieken om het probleem op te lossen;
3. generalisering van nieuw ontwikkelde technieken;
4. kennisuitwisseling.

1.2 Het Prioriteitsprogramma Massaal Parallel Rekenen

In Nederland is de ontwikkeling van het grootschalig rekenen de afgelopen jaren gestimuleerd door middel van het prioriteitsprogramma Massaal Parallel Rekenen (MPR). Het programma werd vanaf 1992 voorbereid, het startte in 1994 en wordt in de loop van 2000 afgerond. Het doel van MPR was de in Nederland aanwezige expertise op het gebied van massaal parallel rekenen en de expertise op verschillende toepassingsgebieden te bundelen en te versterken. Er werden twee hoofdlijnen van onderzoek voorzien: stromingsleer en veel-deeltjes-systemen. Het MPR-programma bestond uit vijf multidisciplinaire clusters en twee projecten, waarin onderzoek naar discipline applicaties werd samengebracht met dat naar generieke methoden en technieken. De gemeenschappelijke noemer voor alle clusters en het daarbinnen uitgevoerde onderzoek was het gebruik van computers met een groot aantal processoren, de zogenaamde massaal parallelle machines.

De activiteiten binnen het MPR-programma werden afgestemd met die van de stichting High Performance Computing and Networking (HPCN). Deze mede door de ministeries van Economische Zaken en Onderwijs, Cultuur en Wetenschappen gefinancierde stichting

Massaal Parallel Rekenen (MPR)

- Looptijd: 1994 – 2000.
- Totaal budget: Mf 10, waarvan Mf 6 voor onderzoek en Mf 4 voor apparatuur.
- Totaal aantal betrokken onderzoekers: 36 postdocs, 4 promovendi, 4 overig personeel.
- Het onderzoek werd uitgevoerd in vijf multidisciplinaire clusters en twee projecten. De clusters waren: (1) Parallele I/O & Imaging, (2) Computational magneto-fluid dynamics, (3) Portable parallel simulation of crystal nucleation and growth, (4) The discreteness-continuity dichotomy in individual-based population dynamics using massively parallel machines, (5) Scalable algorithms for quantum molecular dynamics simulations. De twee projecten waren: (1) Imaging en (2) Genome Computing.

begeleidde, parallel aan de looptijd van MPR, 25 projecten waarin wetenschap en industrie samenwerkten.

Tijdens de uitvoering van het MPR-programma is een schat aan ervaring opgedaan met multidisciplinaire samenwerking tussen ontwikkelaars van generieke methoden en onderzoekers in applicatiegebieden. Het programma voor

Computational Science dat in dit rapport wordt voorgesteld, bouwt voort op deze infrastructuur.

1.3 Definitie van Computational Science

Sameh et al. definiëren Computational Science als volgt:

“Computational Science seeks to advance science and engineering disciplines through better understanding of advanced computers and computational models, as well as advancing the state of the art in computer architecture, system software, and algorithm design through better understanding of science and engineering applications.”¹

Een wezenlijk kenmerk van Computational Science is de *wisselwerking* tussen funderend en applicatiegericht onderzoek: de applicatiegebieden formuleren de problemen waarvoor in funderend onderzoek methoden en technieken worden ontwikkeld die een essentiële bijdrage leveren aan mogelijke oplossingen. Bij het oplossen van problemen gaan binnen Computational Science theorievorming, experimentatie en simulatie hand in hand. Schematisch kan dit volgens Sloot als volgt worden weergegeven²:

“Computational Science brings together:
- Application
- Algorithm
- Architecture
_____ into a Problem Solving Environment”

Computational Science heeft zich ontwikkeld tot een volwaardige wetenschappelijke discipline, naast de experimentele en de theoretische disciplines. Zowel het klassieke experimentele als het analytisch-theoretische onderzoek profiteren van de inzichten uit het computationele onderzoek. In de praktijk zijn de drie onderzoeksparadigma's moeilijk van elkaar te scheiden. Illustratief in dit verband is dat de theoretische doorbraak van onze Nobelprijswinnaars 't Hooft en Veltman vooraf gegaan werd door sturend computationeel werk van Veltman.

¹ Sameh et al, “Computational Science and Engineering”, in *ACM Computing Surveys*, Vol. 28, No. 4, December 1996.

² Sloot, P.M.A., “The Renaissance of Computing”, theme presentation at ASCI'99 (15-17 June 1999).

1.4 De wetenschappelijke uitdaging voor een programma Computational Science

In de jaren '80 leidde de toegenomen computerrekenkracht samen met het ontwerp van efficiënte en robuuste algoritmen en het beschikbaar komen van betere mogelijkheden voor dataverwerking tot de ontwikkeling van zogenoemde “*large scale*-modellen”. Deze modellen leveren voor geïsoleerde fysische (of chemische of biologische etc.) verschijnselen realistische en experimenteel verifieerbare resultaten op. Met *large scale*-modellen kan bijvoorbeeld vloeistofstroming in een regelmatig kanaal worden bestudeerd. Kenmerkend voor *large scale*-modellen is dat zij redelijk goed begrepen en opschaalbaar zijn en een tamelijk uniforme structuur hebben. Grofweg kunnen twee groepen van *large scale*-modellen worden onderscheiden, namelijk discrete modellen (zoals veel-deeltjes-modellen) en continue modellen (zoals in de vloeistofdynamica). Discrete modellen worden veelal gebruikt om van systemen het “lokale” gedrag met korte tijd- en lengteschalen te bestuderen. Continue modellen worden daarentegen vooral ingezet voor het “globale” gedrag op grotere tijd- en lengteschalen.

Maar natuurlijke systemen vertonen, zelfs wanneer ze beschreven worden in termen van intrinsiek eenvoudig formuleerbare vergelijkingen, grillig en zeer complex gedrag. Zo blijkt het bijvoorbeeld veel lastiger te zijn dan aanvankelijk werd gedacht om de macroscopische systemen die we in de praktijk van alledag waarnemen, op te bouwen uit hun elementaire componenten. Onderzoekers die een nauwkeuriger benadering willen geven van het gedrag van dergelijke complexe systemen, stuiten op een reeks van fundamentele problemen, zoals:

- toegenomen complexiteit van modellen, door verschillende tijd- en lengteschalen;
- de noodzaak om meer details (“lokaal” gedrag) te simuleren teneinde het “globale” gedrag van het verschijnsel beter te begrijpen;
- de noodzaak om meer dan één fysisch verschijnsel (stroming, warmte, meerdere stoffen, chemische reacties, straling, etc.) te modelleren;
- de noodzaak van het gebruik van realistische domeinstructuren (grillige randen).

Het begint duidelijk te worden dat het voor de modellering van complexe systemen noodzakelijk is inzicht te verwerven in de *interactie* tussen “lokale” en “globale” verschijnselen en de verschillende tijd- en lengteschalen waarop deze zich afspelen. De verschillende niveau's die in een complex systeem kunnen worden onderscheiden, moeten dus aan elkaar *gekoppeld* worden en wel zó, dat de voortdurende interactie daarin tot uitdrukking wordt gebracht. De wetenschappelijke uitdaging is daarom te komen tot *gekoppelde dynamische modellen* om het gedrag van complexe systemen te bestuderen, te begrijpen en te voorspellen. Het gaat dan dus om modellen waarin aandacht is voor de effecten op micro-, meso- en macroschaal en hun onderlinge koppeling. Dergelijke modellen kunnen gekarakteriseerd worden als “*large scope*-modellen”. Voor het ontwerp, begrijpen en optimaliseren van deze *large scope*-modellen is Computational Science nodig.

1.5 Doelstellingen van het programma

De algemene doelstellingen van het programma Computational Science kunnen als volgt samengevat worden:

1. het bewerkstelligen van een significante voortgang bij het bestuderen, begrijpen en voorspellen van het gedrag van complexe (natuurlijke) systemen door middel van de ontwikkeling en het gebruik van gekoppelde dynamische modellen, zowel in applicatiegericht als funderend opzicht;
2. het overdraagbaar maken van de nieuw ontwikkelde modellen en software;
3. het versterken van het kennisnetwerk en het verder bevorderen van kennisontsluiting en kennisoverdracht van onderzoeksresultaten op het gebied van Computational Science;

Het programma Computational Science operationaliseert deze algemene doelstellingen door middel van de volgende activiteiten:

- A. bevorderen van wisselwerking tussen applicatiegericht en funderend onderzoek;
- B. tot stand brengen van concrete samenwerkingsverbanden tussen applicatiegericht en funderend onderzoek;
- C. stimuleren van het verder ontwikkelen en overdraagbaar maken van modellen en software;
- D. vormen van een forum dat moet fungeren als nationale en internationale ontmoetings- en oversteekplaats voor Computational Science-onderzoek.

2. Het programma Computational Science

Het centrale wetenschappelijke doel van het programma Computational Science is significante voortgang te boeken bij het bestuderen, begrijpen en voorspellen van het gedrag van complexe (natuurlijke) systemen door middel van de ontwikkeling en het gebruik van gekoppelde dynamische modellen. Essentieel daarbij is de interdisciplinaire wisselwerking tussen enerzijds modelgebaseerd rekenwerk dat ingezet wordt om bepaalde domeingebonden problemen op te lossen (*applicatiegericht onderzoek*) en anderzijds de ontwikkeling van (nieuwe) generieke methoden en technieken voor rekenkundige en combinatorische bewerkingen (*funderend onderzoek*). De wisselwerking bestaat hieruit dat binnen dit Computational Science-programma aan applicatiegericht onderzoek altijd funderend onderzoek gekoppeld zal moeten zijn, terwijl het funderende onderzoek nooit los gezien kan worden van de applicaties. In dit hoofdstuk worden de uitdagende problemen binnen het applicatiegerichte en het funderende onderzoek behandeld, en wordt aandacht besteed aan de noodzaak tot het maken van overdraagbare *tools*.

2.1 *Applicatiegericht onderzoek*³

Het programma Computational Science richt zich vooralsnog op applicatiegebieden zoals omvat door de Exacte Wetenschappen binnen NWO. Uitgangspunt voor het programma is: Er zijn onopgeloste problemen in de Fysica, de Astronomie en de Informatietechnologie die een grote uitdaging vormen voor die disciplines als geheel. Aan het oplossen van deze problemen kan een Computational Science-aanpak een belangrijke bijdrage leveren. Een representatieve keuze van uitdagende problemen is de volgende:

- *Turbulentie*
Turbulentie is en blijft vooralsnog één van de grote uitdagende problemen van de Natuurwetenschappen. Door de recente vooruitgang in rekenkracht van supercomputers is er nu meer kans op echt nieuw inzicht, omdat één van de grote grondproblemen van turbulentie het in rekening brengen is van de verschillende grootteschalen waarop de vloeistofdynamica zich afspeelt. De mogelijkheid hierbij chemische reacties en verbranding mee te nemen maakt dit gebied tot uiterst relevant voor zowel industriële toepassingen als milieuproblematiek.
- *Vorming van structuur in het Heelal*
Een belangrijk onderzoeksthema in de hedendaagse Astronomie is het vraagstuk van de vorming van sterrenstelsels. Aan de vorming van sterrenstelsels ligt een uiterst complex samenspel van tal van fysische processen ten grondslag. Bij de evaluatie van de astrofysische scenario's spelen computerberekeningen een essentiële rol als primaire onderzoeksmethode. De onstuimige toename van computercapaciteit heeft daarbij al geleid tot fundamentele doorbraken. Met name veel-deeltjes-simulaties hebben hierbij een belangrijke plaats ingenomen. De laatste jaren maakt de computercapaciteit het ook mogelijk de vorming van sterrenstelsels en sterren aan te pakken, binnen de context van door de gravitatie gevormde structuren. Een eerste aanzet van tamelijk ruwe hydrodynamische berekeningen heeft al tot aanzienlijke aanscherping van ons begrip geleid. Nu het besef bestaat dat het gaat om een complex samenspel van gravitatie, gasdynamica, stralingsprocessen en *radiative transfer*, stervorming, supernova-explosies en chemische verrijking, is de trend gezet naar de exploratie en exploitatie van steeds veeleisender *large scope*-modelleringen.
- *De fysica van plasma's*
De dynamica van plasma's, beschreven met de Magnetohydrodynamica (MHD) en berekend met *Computational Magneto-Fluid Dynamics* (CMFD), is van groot belang voor laboratoriumexperimenten, maar ook voor astrofysische waarnemingen op vrijwel alle schalen en in een grote diversiteit aan kosmische omgevingen, aangezien een aanzienlijk deel van de baryonische materie in het heelal

³ In deze paragraaf is dankbaar gebruik gemaakt van de nota *Computational Physics in Nederland* (FOM, december 1999), geschreven door De Leeuw en Goedbloed.

zich in de plasmafase bevindt. In dit opzicht is er overeenkomst met de gewone CFD (verkregen in de limiet van verdwijnend magnetisch veld), zodat sommige numerieke technieken uit de CFD ook in de CMFD bruikbaar zijn. Echter, de grote uitdaging is de beschrijving van de niet-lineaire interactie van stromende plasma's met de magnetische geometrie. Dit betreft zowel ongewenste disrupties van laboratoriumplasma's als de voorspelling van coronale massa-uitstoot van de zon die tot magnetische stormen in de aardse magnetosfeer leiden ("*space weather*"). Iedere ontwikkeling in dergelijk onderzoek is direct profijtelijk voor velerlei andere onderzoeksgebieden, zoals roterende plasma's, stellaire plasma's, de rol van magneetvelden in de vorming van sterren en de zeer ingewikkelde situaties die in kernen van sterrenstelsels resulteren in de intrigerende manifestaties van *Active Galactic Nuclei*.

- *Mesoscopische materiaalkunde/fysica*

Bij scheurvorming in een vaste stof worden chemische bindingen gebroken die kwantummechanisch behandeld dienen te worden. Iets verder weg van het scheurfront kan de beweging van de atomen met klassieke mechanica beschreven worden, terwijl ver van het scheurfront elasticiteitstheorie kan worden gebruikt om het systeem te beschrijven. Al deze niveaus zijn echter gekoppeld en een uniforme beschrijving van het scheurproces vergt dus een koppeling tussen een kwantummechanisch model, gebaseerd op bijvoorbeeld dichtheidsfunctionaaltheorie, een klassiek model (moleculaire dynamica) en een continuümmodel (eindige elementen). Een geünificeerd model voor scheurvorming met bijbehorende simulaties vormt een van de grote uitdagingen van de materiaalkunde. Ook bij de bestudering van complexe materialen en vloeistoffen, zoals polymeren, micro-emulsies en hybride materialen, is er sprake van een veelheid aan fysische wetten en vergt dit een koppeling tussen de verschillende modellen voor structuur en dynamica. Mogelijke technieken hierbij zijn moleculaire dynamica, gekoppeld aan mesoscopische dynamica met rooster Boltzmann modellen of Dissipative Particle Dynamics.

2.2 *Funderend onderzoek*

Binnen het funderende gedeelte van Computational Science staan onderzoekers voor de opgave om technieken te ontwikkelen waarnaar binnen het applicatiegerichte onderzoek vraag is. In het kader van het in paragraaf 1.4 besproken *large scale*-onderzoek zijn in de afgelopen jaren binnen tal van disciplines methoden, technieken en (software-) instrumenten ontwikkeld waarmee complexe systemen bestudeerd kunnen worden. Voor verdere ontwikkeling hiervan is in met name de Informatica en de Wiskunde funderend onderzoek noodzakelijk. Net als het applicatiegerichte onderzoek staat het funderende onderzoek voor een aantal wetenschappelijk uitdagende problemen. Een representatieve keuze van niet geheel nevenschikte thema's hieruit is de volgende⁴:

- *Simulatie en modelleren*

De kern in het Computational Science onderzoek wordt gevormd door de algoritmen waarmee de berekeningen uitgevoerd worden. Verdere ontwikkeling van geschikte algoritmen, zowel wat betreft snelheidsverbetering, diversiteit qua probleemgrootte en domeingebied als nauwkeurigheid, is daarom een wezenlijk onderdeel in het programma. Te denken valt daarbij aan het ontwerp en de analyse van poly-algoritmen (bijvoorbeeld gekoppelde Euler-Lagrange modellen) en datastructuren waarmee de koppeling tussen modellen en domeinen van ongelijksoortige aard gerealiseerd kunnen worden en algoritmen waarmee het onderscheid tussen globaal en lokaal gedrag zichtbaar gemaakt kan worden.

⁴ Ook in het NWO EW-programma "Netwerken" (dat naar verwachting eveneens in 2000 van start zal gaan) zal funderend onderzoek gedaan worden aan een aantal van de genoemde problemen.

- *Stochastiek en Kansrekening*

Het beschikbaar komen van steeds krachtiger computers en de ontwikkeling van efficiënte algoritmen zoals Markov chain Monte Carlo hebben ertoe geleid dat het nu mogelijk is hoogdimensionale, complexe, niet-lineaire systemen te bestuderen. Men denke bijvoorbeeld aan modellen voor de interactie tussen deeltjes in een vloeistof of gas of tussen de objecten in een beeld, aan grafische modellen en aan de bevrijding van geconjugeerde a priori-verdelingen in de Bayesiaanse statistiek. Monte Carlo-methoden -- op basis van Markov-ketens of soms van differentiaalvergelijkingen -- maken het mogelijk simulaties uit te voeren, de modelparameters te schatten, betrouwbaarheidsintervallen te construeren, of kansen op en gemiddelde waarden van saillante kenmerken te schatten. Recentelijk is gebleken dat in sommige gevallen zelfs exacte simulatie mogelijk is. Ook zijn stochastische technieken nuttig bij het berekenen van de optimale waarde in vele combinatorische problemen. Gezien de grote verscheidenheid aan technieken, is het de uitdaging bij een gegeven applicatie het meest geschikte algoritme te vinden.

- *Analyse van onvolledige gegevens en statistische inverse problemen.*

Bij het doorrekenen van grootschalige problemen doet zich regelmatig het probleem voor dat de uitgangsgegevens niet zonder meer geschikt zijn om het model met voldoende voorspellingskracht in werking te zetten. Een bekend voorbeeld is weersvoorspelling, waarbij de begintoestand, bijvoorbeeld druk en temperatuur, voor een berekening niet voldoende fijnmazig gemeten kan worden, dan wel dat de beschikbare data niet uniform op hetzelfde tijdstip bekend zijn. Dit verschijnsel doet zich bij vrijwel elk realistisch voorspellingsmodel voor. Een andere situatie is dat metingen, in een database, niet allemaal aan hetzelfde statistische model voldoen, simpelweg omdat ze met verschillende apparatuur of onder verschillende omstandigheden tot stand zijn gekomen. Ook komt het voor dat data door meet- of transmissiestoringen (satellietwaarnemingen) deels verloren zijn gegaan of zijn verontreinigd en men alleen informatie heeft over een interval of een gebied waartoe de waarneming behoort. Statistische technieken zijn nodig om ontbrekende gegevens aan te vullen en de verdeling van de niet-observeerbare grootheden te schatten, dan wel om het effect van benaderingstechnieken te kunnen evalueren. Dit vereist interactie met modelmatige gegevens, met algoritmische technieken, en met informatie over het uiteindelijke rekenmodel. De zogenaamde "interval censurering" is een statistische techniek die recent object van studie is geworden. Zowel het berekenen van schatters als het ontwikkelen van de verdelingstheorie is hierbij veel moeilijker. Meestal kunnen de schatters alleen via computer-intensieve iteratieve algoritmen berekend worden en over de eigenschappen van de zo berekende schatters is nog weinig bekend. Deze techniek behoort tot de zogenaamde "statistische inverse problemen", net als bijvoorbeeld deconvolutie, die gebruikt wordt om uit waarnemingen van sommen van stochastische variabelen de achterliggende individuele verdelingen op te sporen.

- *Beeldverwerkende technieken*

Dankzij de ontwikkeling van snelle grafische apparatuur en krachtige computers, alsmede de ontwikkeling van nieuwe methoden, technieken en gereedschappen is het gebruik van visualisatietechnieken gemeengoed geworden voor het analyseren en presenteren van de grote hoeveelheden gegevens die door simulaties geproduceerd worden. Het tonen van de relatie tussen modelparameters en uitkomsten, en het vergelijken van gemeten en gesimuleerde gegevens kan op een aantal manieren worden aangepakt. In de eerste plaats door middel van automatisch herkennen van patronen en structuren, ook wel bekend als *feature extraction* en simplificatie. Hiertoe kunnen methoden uit onder meer de beeldbewerking, de statistische analyse en ook numerieke simulatie worden ingezet. In de tweede plaats, door middel van alternatieve presentaties van deze (al dan niet gesimplificeerde) gegevens. Hierbij kan worden gedacht aan iconificatie en het gebruik van technieken uit informativisualisatie, maar ook aan het gebruik van *Virtual Reality*-technieken, zoals stereoprojectie. In de derde plaats, door het ontwikkelen van nieuwe technieken voor de gebruikersinterface, met name voor het verhogen van het abstractieniveau van communicatie tussen gebruiker en systeem en het automatisch genereren van visualisaties. In de vierde plaats, door simulatie en modelvorming sterker te koppelen aan visualisatie, zodat de parameterruimte kan worden verkleind. De uitdaging in al deze oplossingsrichtingen is daarbij de vaak voor een toepassingsgebied specifieke techniek te generaliseren tot generieke methoden die voor een scala van toepassingen inzetbaar zijn.

2.3 *Overdraagbaarheid van methoden uit applicatiegericht en funderend onderzoek*

Kenmerkend voor wetenschappelijk onderzoek waarbij modellen en software worden ontwikkeld is, dat zodra het vernieuwende gedeelte achter de rug is, er onvoldoende tijd, menskracht en expertise overblijft om de generieke modellen en software verder te ontwikkelen tot handzame en herbruikbare *tools*. De bedoeling van dit programma Computational Science is om deze stagnatie tegen te gaan en ruimte te bieden aan de verdere ontwikkeling hiervan. Het programma heeft de ambitie bij te dragen aan de verzameling modules, eventueel bijeengebracht in software-'bibliotheken'. Applicatiegerichte onderzoekers krijgen zo de beschikking over software-instrumenten die zij kunnen inpassen in hun eigen rekenmodellen, die op hun beurt weer kunnen worden gekoppeld aan verdere verwerkingsmogelijkheden (visualisatie, statistische bewerking, gevoeligheidsanalyse, rekenfout-afschattingen, etc). Een fysicus zou bijvoorbeeld een stroming moeten kunnen zien veranderen onder invloed van de hem interesserende modelaanpassingen. Zo zou het bijvoorbeeld veel eenvoudiger moeten zijn om de complexiteit van een model te veranderen, door convectieve termen toe te voegen, over te gaan van een rotatiesymmetrisch model naar een niet-symmetrisch model, van twee naar drie dimensies etc. Of hij zou zich met behulp van geavanceerde statistische technieken een goede indruk moeten kunnen verschaffen van de zwakke kanten van een model. Zo kan er gericht over nieuwe experimenten en theorievorming gedacht worden en leidt het gebruik en de schakeling van welbegrepen technieken weer tot vooruitgang in het Computational Science-onderzoek.

2.4 *Bundeling van inspanningen: het Computational Science-programma als forum*

De meerwaarde van het Computational Science-programma is gelegen in de bundeling van inspanningen die momenteel zowel in het applicatiegerichte als het funderende onderzoek verricht worden. Geïnspireerd door grote buitenlandse programma's zoals het Amerikaanse ASCI (*Accelerated Strategic Computing Initiative*) en voortbouwend op de ervaringen binnen het prioriteitsprogramma MPR en de HPCN-projecten (zie hoofdstuk 1) dient het Computational Science-programma het kader te vormen voor overdracht van kennis op bredere basis, net zoals dit in het verleden geschiedde in de experimentele wetenschappen door experimenteerervaring vast te leggen in algemeen bruikbare instrumentatie. Aldus kan voorkomen worden dat Computational Science-onderzoek zich te geïsoleerd met concrete toepassingen dan wel funderingen bezighoudt. Het is dus van cruciaal belang dat het Computational Science-

programma zal fungeren als forum, d.i. als ontmoetings- en oversteekplaats van onderzoekers en onderzoek, theorieën en praktijkervaringen, ideeën en software-modulen.

3. Vormgeving van het programma Computational Science

3.1 Te subsidiëren activiteiten

Binnen het programma Computational Science zullen drie activiteiten gefinancierd worden:

1. Grote projecten met een maximale omvang van Mf 1,5.
2. Kleine projecten met een maximale omvang van Mf 0,75
3. Programmabrede activiteiten. Hiervoor is in totaal Mf 1 beschikbaar.

De voorziene verdeling van de middelen over de drie activiteiten is globaal als volgt:

programmabrede activiteiten: Mf 1		
funderend onderzoek	onderzoek in samenwerking	applicatiegericht onderzoek
kleine projecten 25% van de middelen	grote projecten 50% van de middelen	kleine projecten 25% van de middelen

Het is daarbij cruciaal een goed onderscheid te maken tussen *wisselwerking* en *samenwerking*. In het *hele* programma Computational Science, dus ook in het applicatiegerichte en het funderende deel, staat, uit de aard van het onderzoeksgebied, *wisselwerking* centraal. Dit betekent dat bij de beoordeling van funderende projecten mee zal wegen in hoeverre er een vraag is vanuit een applicatie, en dat bij de beoordeling van applicatiegerichte projecten in de beoordeling mee zal spelen of de doelstelling is om tevens qua generieke technieken vooruitgang te boeken. Daadwerkelijke *samenwerking* tussen onderzoeksgroepen met een verschillende achtergrond (theoretisch, experimenteel, computationeel) of afkomstig uit verschillende disciplines zal bij de grote projecten een eis zijn. Bij de kleine projecten zal dat een pré maar geen noodzakelijke voorwaarde zijn. Onder “daadwerkelijke samenwerking” moet in dit verband worden verstaan: een gemeenschappelijke vraag- of probleemstelling en een onderzoeksaanpak die erop gericht is zoveel mogelijk gezamenlijke onderzoeksresultaten te behalen. In concreto zal deze samenwerking neerslaan in bijvoorbeeld gezamenlijke publicaties of de ontwikkeling van methoden, algoritmen of software die in onderzoeksgroepen uit meer dan één discipline gebruikt kunnen worden.

Grote projecten

- Voorwaarde voor financiering van een groot project is aantoonbare actieve interdisciplinaire samenwerking. De interactie tussen de deelnemende onderzoekers moet meerwaarde opleveren voor ‘breed’ gebruik, hetzij binnen de ‘eigen’ deelprojecten, hetzij programmabreed. Bij elk groot project moet minstens een informaticus of een wiskundige betrokken zijn. Bij projecten met een subsidie-omvang van meer dan Mf 1 zal, teneinde verankering op langere termijn te waarborgen, de onderzoekinstellingen gevraagd worden het meerdere er zelf bij te leggen. Voor een project van bijvoorbeeld Mf 1,4 wordt een matching van Mf 0,4 in termen van nieuw aan te stellen personeel of instrumentatie verwacht. Hiermee komt de totale omvang van een dergelijk project op Mf 1,8.
- Te financieren: wetenschappelijk onderzoek uitgevoerd door medewerkers in tijdelijke dienst en inzet van wetenschappelijke programmeurs. Per project of projectuitvoerder kan

materieelkrediet worden agevraagd dat software, reiskosten en kleine projectspecifieke apparatuur kan omvatten.

- Trefwoorden voor grote projecten: Sterke gerichtheid op significante vooruitgang op het gebied van de Computational Science. Sterke interdisciplinaire interactie tussen applicatiegericht en funderend onderzoek.

Kleine projecten

- Kleine projecten kunnen monodisciplinair geïnitieerd worden, maar hun horizon wordt wel bepaald door toepassing of inpassing in *large scope*-modellen. Daadwerkelijke interdisciplinaire samenwerking is niet strikt noodzakelijk, maar wordt wel aanbevolen. Kleine projecten kunnen een accent hebben op applicatiegericht of op funderend onderzoek.
- Te financieren: wetenschappelijk onderzoek uitgevoerd door medewerkers in tijdelijke dienst, zoals promovendi en postdocs alsmede een bijdrage aan de inzet van eventuele wetenschappelijke programmeurs. Per project of projectuitvoerder kan materieelkrediet worden agevraagd dat software, reiskosten en kleine projectspecifieke apparatuur kan omvatten.
- Bij de beoordeling van de kleine projecten zal, naast de wetenschappelijke kwaliteit, worden gekeken naar de applicatiegerichte componenten in voorstellen voor funderend onderzoek en funderende componenten in applicatiegerichte voorstellen.
- Trefwoorden voor applicatiegerichte kleine projecten: Applicatiegerichte projecten waarbij (de noodzaak voor) ontwikkeling van generieke *enabling tools* een cruciale rol speelt (bijvoorbeeld de inpassing van visualisatie-tools of interactieve simulatie-tools in een bestaand computermodel).
- Trefwoorden voor funderende kleine projecten: generieke *enabling tools*, gericht op inpassing in *large scope*-modellen. *Problem solving environments*.
- Trefwoorden voor beide: Mogelijkheid voor specialistische projecten waarin rekenmodulen van velerlei aard (bijvoorbeeld numerieke software, visualisatie-tools, fysisch georiënteerde applicaties, statistische technieken) op overdraagbare wijze voor algemener gebruik worden ontwikkeld en vastgelegd.

Programmabrede activiteiten

- Gedurende de looptijd van het programma wordt er ieder jaar een workshop gehouden voor alle onderzoekers die bij het project betrokken zijn en voor andere belangstellenden. Voor de organisatie hiervan is voor de totale looptijd van het programma ca. Mf 0,4 beschikbaar.
- De inrichting van een platform voor de uitwisseling, opslag, het onderhoud en de documentatie van software. Voor dit doel zal een website worden ontwikkeld. Voor de totale looptijd van het programma ca. Mf 0,6.

De inrichting van het demonstratieplatform is binnen het programma een apart project, dat, ten behoeve van het hele programma, na competitie bij één van de deelnemende groepen zal worden ondergebracht.

3.2 Benodigde en beschikbare middelen

Een Computational Science-programma met de bovengenoemde doelstellingen en met als deelnemers onderzoekers uit de disciplines Astronomie, Fysica, Informatica en Wiskunde moet

een totale omvang hebben van ca. Mf 15 tot Mf 20.⁵ Uitgaande van twee subsidierondes waarin projecten met een looptijd van vier à vijf jaar worden uitgezet, is de voorziene looptijd van het programma zes tot acht jaar. Op de begroting van het Gebiedsbestuur Exacte Wetenschappen is tussen de Mf 7 en Mf 8,5 beschikbaar voor de periode tot en met 2004. Dit is voldoende om het programma van start te laten gaan. De programmacommissie zal o.a. tot taak moeten krijgen additionele financiering te vinden. Een mogelijk aanhaken van onderzoekers uit andere dan de vier genoemde disciplines zal moeten leiden tot een grotere benodigde omvang van het programma.

3.3 *Benodigde en beschikbare computerfaciliteiten*

In brede kring wordt de behoefte gevoeld aan een onderzoeksinfrastructuur die het Nederlandse onderzoekers mogelijk maakt een vooraanstaande rol te spelen in de overgang van het ontwerp *large scale*- naar *large scope*-modellen. De Nederlandse uitgangspositie daarvoor is goed. Nederland beschikt over één van de snelste computernetwerken ter wereld en zal in 2000 een nieuwe nationale massaal parallelle computer met een rekencapaciteit van ongeveer 1 Teraflop en om en nabij 1 Terabyte direct adresseerbaar geheugen aanschaffen.

Deze computer wordt aangeschaft en beheerd onder auspiciën van de NWO-stichting Nationale Computerfaciliteiten (NCF), waarmee in het programma nauw zal worden samengewerkt daar waar Computational Science-projecten van de supercomputer gebruik wensen te maken.

Overigens beoefent men Computational Science niet alleen op een grootschalige Teraflop-machine met zeer veel geheugen (al is de grootschalige centraal gefinancierde faciliteit noodzakelijk om aansluiting bij het internationale toponderzoek te behouden), maar ook op werkstations en alles wat daar tussenin ligt (de zogenaamde piramidestructuur). Bovendien vereisen verschillende algoritmen of programma's vaak een verschillende computerarchitectuur voor een efficiënte verwerking. Een breed scala aan computers met verschillende architectuur is daarom wenselijk. Er bestaat echter niet de indruk dat er zich wat dit betreft momenteel schrijnende tekorten voordoen. Binnen het programma zullen daarom geen middelen voor infrastructuur gerelateerde apparatuur beschikbaar worden gesteld.

3.4 *Globaal tijdpad*

De gedachten gaan uit naar een programma dat zes tot acht jaar loopt. De middelen worden in twee rondes uitgezet. De projecten zullen een gemiddelde looptijd hebben van vier à vijf jaar. Grote projecten, die gericht zijn op verankering van het onderzoek in het universitaire bestel, kunnen langer duren. Voor de eerste ronde zal een oproep tot indiening van voorstellen zomer 2000 uitgaan. De gehonoreerde projecten kunnen dan in de loop van 2001 beginnen.

Het tijdstip van de tweede ronde zal o.a. afhangen van de beschikbare middelen en de ervaringen opgedaan met de eerste ronde. De oproep voor indiening van voorstellen voor de tweede ronde wordt niet voor 2002 verwacht.

3.5 *Bestuurlijke vormgeving*

De voortgang en de wetenschappelijke kwaliteit van het programma zal worden bewaakt door een programmacommissie bestaande uit deskundigen afkomstig uit de bij het programma betrokken disciplines. De programmacommissie werkt onder verantwoordelijkheid van het NWO-Gebiedsbestuur Exacte Wetenschappen, dat tevens de budgethouder is. De programmacommissie richt, in samenspraak met het Gebiedsbestuur, zelf haar werkzaamheden in. Zij heeft de vrijheid om desgewenst deskundigen van buiten de commissie bij haar

⁵ De Leeuw en Goedbloed stellen in hun in voetnoot 3 reeds genoemde nota dat alleen al voor het vakgebied *Computational Physics* tussen de Mf 9 en Mf 12 nodig zou zijn.

werkzaamheden te betrekken. De programmacommissie zal zich voor de dagelijkse en secretariële werkzaamheden laten bijstaan door een programmacoördinator.

Bijlage 1: Relatie met enkele (inter)nationale ontwikkelingen

Internationaal staat het Computational Science-onderzoek zeer in de belangstelling. Hier zij uitsluitend verwezen naar een drietal grote initiatieven in de Verenigde Staten waarmee vele tientallen miljoenen dollars gemoeid zijn.⁶: het *Accelerated Strategic Computing Initiative*, de *National Computing Science Alliance* en het *Information Technology Research*-initiatief van de *National Science Foundation*.

Maar ook in Nederland is op het gebied van de Computational Science in de afgelopen jaren een aantal initiatieven ontplooid of in voorbereiding genomen. Het voorgestelde Computational Science-programma past in de context van deze initiatieven. We noemen er hier enkele:

Onderzoek

- In het Wetenschaps- en TechnologieCentrum Watergraafsmeer (WTCW, Amsterdam) is het Amsterdam Centrum for Computational Science (ACCS) opgericht. Dit centrum is een samenwerkingsverband tussen Fysica, Wiskunde, Informatica, Biologie, Fysische Geografie en Scheikunde.
- In Eindhoven is in 1999 het Eindhoven Centre for Computational Engineering (ECCE) opgericht waarin Fysica, Wiskunde, Informatica, Scheikundige Technologie, Werktuigbouwkunde en Biomedische Technologie samenwerken op het gebied van computersimulaties.
- Op initiatief van de NWO-gebieden Chemische Wetenschappen (CW) en Aard- en Levenswetenschappen (ALW) heeft een initiatiefgroep de mogelijkheden om te komen tot een programma Biomoleculaire Informatica onderzocht. Op basis van het begin 2000 verschenen rapport zal een programmavoorbereidingscommissie worden ingesteld om de thema's uit te werken.
- Binnen de onderzoeksschool ASCI is in 1999 een onderzoekspilaar Computational Science opgezet, naast de bestaande pilaren Imaging en Computersystemen.

Het ligt voor de hand dat er vanuit het voorgestelde programma nauw zal worden samengewerkt met onderzoekers werkzaam binnen de andere initiatieven, al was het alleen maar omdat onderzoekers vaak bij meer dan één initiatief betrokken zijn.

Onderwijs

- Aan de Universiteit Utrecht bestaat sinds 1997 de mogelijkheid om een volledig studieprogramma Computational Science te volgen. Jaarlijks beginnen 10 tot 15 studenten aan een studie CS.
- De Universiteit van Amsterdam biedt jaarlijks aan 10 tot 20 geselecteerde internationale studenten een tweejarig internationale master-opleiding (MSc) Computational Science aan. Dit multidisciplinaire programma wordt gedragen door het ACCS.
- In september 2000 start aan de Technische Universiteit Eindhoven een masterclass programma "Computational Science & Engineering".
- Binnen de Rijksuniversiteit Groningen is recentelijk voorgesteld binnen elk van de exacte disciplines in het onderwijs een CS-variant op te nemen.

Hoewel het voorgestelde programma uitsluitend betrekking heeft op onderzoek, is het goed te weten dat er binnen het universitaire onderwijs wordt gewerkt aan het kweken van "Nachwuchs". Een langlopend programma als het voorgestelde kan enerzijds veel profijt hebben van de gerichte opleiding van toekomstige onderzoekers, terwijl anderzijds het bestaan van een

⁶ Een overzicht over de internationale stand van zaken is te vinden in de studie die in 1999 verricht is ten behoeve van de aanschaf van een nieuwe nationale supercomputer.

dergelijk programma het voor studenten aantrekkelijker kan maken dit specifieke onderzoeksveld te kiezen.

Bijlage 2: potentieel geïnteresseerde onderzoekers

Er is voldoende belangstelling te verwachten voor een programma van deze omvang. Hieronder volgt een opsomming (in alfabetische volgorde) van de onderzoekscholen waarin zich onderzoekers bevinden die potentieel geïnteresseerd zijn in deelname aan een programma Computational Science.

- Advanced School for Computing and Imaging (ASCI): TUD, UvA, VU, UL
- Centrum voor Plasmafysica en Stralingstechnologie: TUE, UT, UU
- Euler Institute for Discrete Mathematics and its Applications: TUE, TUD, UT
- FOM Instituten: AMOLF, NIKHEF en Rijnhuizen
- Instituut voor Programmatuurkunde en Algoritmiëk (IPA): TUE, VU, UL, UU, KUN, UT
- J.M. Burgerscentrum, Onderzoekschool voor Stromingsleer: TUD, TUE, RUG, UT
- Materiaalkunde Instituut Delft, Eindhoven, Groningen (MIDEG): TUD, TUE, RUG
- Materialen Studie Centrum (MSC): RUG
- Mathematisch Research Instituut (MRI): UU, RUG, KUN, UT
- Nederlandse Onderzoekschool voor Astronomie (NOVA): UvA, VU, RUG, UU, UL
- Onderzoekschool Condensed Matter and Optical Physics: UL, UvA, VU
- Onderzoekschool Logica: UvA, VU, RUG, UU, KUN, KUB
- Stevin Centre for Computational and Experimental Engineering Science: TUE
- Telematica Graduate School (TGS): UT
- Thomas Stieltjes Institute for Mathematics: UL, UvA, VU, EUR, TUD, TUE