SURF

Een sprong voorwaarts in wetenschap

De roep om laagdrempelige toegang voor Nederlandse onderzoekers tot 's wereld beste High Performance Computing (HPC).

Managementsamenvatting

Laagdrempelige toegang voor Nederlandse onderzoekers tot 's wereld beste High Performance Computing (HPC) faciliteiten is van doorslaggevend belang voor hun internationale concurrentievermogen. Door laagdrempelige toegang tot (pre-)exascale HPC-systemen kunnen wetenschappelijke toepassingen worden opgeschaald in resolutie en probleemgrootte, komen nieuwe onderzoeksvragen binnen bereik, kunnen grenzen echt worden verlegd, en kan internationaal zo goed mogelijk worden samengewerkt. Science cases van tien vooraanstaande Nederlandse onderzoekers maken dit duidelijk en ondersteunen deze noodzaak.

Nederlandse deelname in EuroHPC project- en infrastructuurconsortia is daartoe cruciaal, en strategisch van belang.

Dit leidt tot vier strategische aanbevelingen voor investeringen in kennis en infrastructuur:

- Wetenschappelijk concurrentiepositie: zorg dat wetenschappers Europees kunnen blijven participeren, door structurele financiering voor matching van EuroHPC R&I projecten, en versterking van de kennispositie van de Nederlandse onderzoekers in relatie tot het EuroHPC werkprogramma en gerelateerde onderdelen in de werkprogramma's van Digital Europe en Horizon Europe
- Aansluiting op het internationale HPC-ecosysteem: blijf als Nederland volwaardig onderdeel van de Europese HPC-infrastructuur door een investering van 10 Mio in EuroHPC-faciliteiten
- Springplank: geef het toegangsbeleid tot Europese HPC-faciliteiten vorm via de NWO commissie Wetenschappelijk Gebruik Supercomputers (WGS)
- Kennis en Kunde: zet in op capacity building met de domeinwetenschap via flankerend beleid op communityvorming, kennisdisseminatie en expertise opbouw via onder meer de thematische DCC's

Inhoudsopgave

1 Inleiding

De EuroHPC Joint Undertaking stelt de Europese Unie en de deelnemende landen in staat hun inspanningen te coördineren en hun middelen te bundelen om van Europa een wereldleider op het gebied van supercomputing te maken. Doelstelling van het partnerschap is om technologieën en applicaties te ontwikkeling waarmee exaflop toepassingen kunnen worden bereikt. Dit zal de Europese wetenschappelijke excellentie en industriële kracht stimuleren, de digitale transformatie van zijn economie ondersteunen en tegelijk zijn technologische soevereiniteit waarborgen.[1]

Nederland behoort tot de eerste zeven landen die de EuroHPC declaration - de voorloper van de Joint Undertaking tussen de lidstaten en de Europese Commissie - hebben ondertekend. Vooralsnog neemt Nederland niet deel in één van de consortia voor de implementatie van (pre-)exascale systemen. De langetermijnstrategie High Performance Computing 2019 stelt dat SURF samen met gebruikers moet bezien of een investering noodzakelijk is voor hun behoeftevoorziening. Eind 2020 is er door Nederlandse wetenschappers een brandbrief aan de Ministeries van OCW en EZK gestuurd waarin de noodklok wordt geluid over de terugloop van de deelname van Nederlandse wetenschapers in EuroHPC gerelateerde onderzoeksprojecten.

In deze notitie wordt de noodzaak voor Nederlandse deelname in EuroHPC project- en infrastructuurconsortia duidelijk gemaakt. Deze wordt ondersteund door science cases van tien vooraanstaande Nederlandse onderzoekers. Op basis hiervan wordt een viertal strategische aanbevelingen gedaan voor investeringen in kennis en infrastructuur.

2 Wetenschappelijke behoefte

2.1 Blijvend wetenschappelijk concurrentievermogen

In de moderne wetenschap is de toegang tot digitale faciliteiten waaronder ook High Performance Computing (HPC) faciliteiten van doorslaggevend belang voor het concurrentievermogen van onderzoekers. Mondiaal gezien loopt Europa achter, met name bij de Verenigde Staten, China en Japan[2]. Met EuroHPC maakt Europa de benodigde inhaalslag. Het is cruciaal voor Nederlandse onderzoekers in deze inhaalslag niet achterop te raken en bij te kunnen blijven met de

[1] https://eurohpc-ju.europa.eu/discover-eurohpc-ju

[2] https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_3

veranderingen in hun wetenschapsgebied en zo laagdrempelig mogelijk toegang te hebben tot 's werelds beste HPC-faciliteiten.

2.2 Wetenschap zonder grenzen

Wetenschappelijke uitdagingen kennen geen grenzen; het plafond tot waar onderzoekers gebruik kunnen maken van High Performance Computing (HPC) faciliteiten daarmee ook niet. Alle onderzoekers die in de in deze notitie opgenomen science cases aan het woord komen, zeggen baat te hebben bij het opschalen van hun applicaties naar een grotere probleemgrootte zoals bij het gebruik van hogere resoluties en bij grotere snelheid van het verkrijgen van de uitkomsten. Daarnaast inspireren nieuwe faciliteiten tot nieuwe onderzoeksvragen of het verleggen van grenzen daarvan. De optimale omvang van faciliteiten wordt daarmee in belangrijke mate bepaald door strategische overwegingen in relatie tot het beschikbare budget.

2.3 Internationale samenwerking op eigen kracht

Eigen kennis, toegang tot faciliteiten en financiële middelen zijn voor Nederlandse onderzoekers noodzakelijk om als sterke aantrekkelijke samenwerkingspartij in Europese samenwerkingsverbanden gezien te worden. Geavanceerde wetenschappelijke toepassingen en resultaten komen tot stand in internationale samenwerkingsverbanden, waarbij consortialeden onderling vergelijkbare wetenschappelijke en technische kennis of een eigen complementair specialisme inbrengen. Zonder eigen inbreng van hoogwaardige HPC-expertise of faciliteiten is een onderzoeker daarbij afhankelijk van de welwillendheid van het consortium.

Voor de EuroHPC Joint Undertaking geldt dat naast de bijdrage van de Europese Commissie, in vergelijking met het reguliere kaderprogramma, hogere investeringseisen aan deelnemende landen voor projectmatching en infrastructuurvoorzieningen worden gesteld. Leidende wetenschappers signaleren op dit moment een terugloop in de deelname van Nederlandse onderzoekers in Europees gefinancierde HPC-projecten, mede als gevolg van de matchingseisen in de EuroHPC R&I projecten.

3 Strategische overwegingen

3.1 Goed vestigingsklimaat vereist goede faciliteiten

Nederland is een kenniseconomie. Voor het vestigingsklimaat van bedrijven is de aanwezigheid van hoogwaardige expertise van doorslaggevend belang. Met participatie in onderzoeksprojecten wordt kennis opgebouwd en talent opgeleid. Een deel van dit talent vindt ook zijn weg op de arbeidsmarkt buiten het academisch domein. Een goede wetenschappelijke concurrentiepositie en de daarbij behorende faciliteiten, zijn van direct belang om wetenschappelijk talent aan te blijven trekken en vast te blijven houden.

3.2 Blijvende aansluiting op het Internationale HPC ecosysteem

Het ecosysteem van High Performance Computing faciliteiten vormt een continuüm. Met sterke lokale faciliteiten in Delft, Leiden, Groningen en Nikhef, en een state-of-the art nationale supercomputer beschikt Nederland over een excellent HPC-ecosysteem. De basis is hiermee gezond, maar ontoereikend in schaalgrootte om alle wetenschapsvragen te kunnen beantwoorden.

Naadloze aansluiting op internationale faciliteiten is cruciaal voor de opschaling van applicaties uit de wetenschapsgebieden. Nederlandse onderzoekers zijn historisch gezien succesvol in het meedingen naar rekentijd op de internationaal toegankelijke voorzieningen die via PRACE op basis van wetenschappelijke excellentie wordt verdeeld. Zij hebben zich daarvoor kunnen - maar vooral ook moeten - voorbereiden op de nationale voorzieningen.

3.3 Springplank voor grootschalig gebruik

Nederlandse onderzoekers kunnen nu via een internationaal reviewproces toegang krijgen tot EuroHPC (pre-)exascale faciliteiten. De langetermijnstrategie HPC gaat ervan uit dat wanneer Nederlandse onderzoekers via deze route toegang kunnen verkrijgen tot EuroHPC-faciliteiten, een eigen investering in een systeem niet noodzakelijk is. Dit model is gebaseerd op wetenschappelijke excellentie, maar kent echter een aantal beperkingen:

- de toegang van nationale onderzoeksprioriteiten tot (pre-)exascale-voorzieningen is afhankelijk van doorlooptijden en concurrentie in het internationale aanvraagproces;
- onderzoekers hebben minder gelegenheid tot voorbereiding van hun applicaties, en zijn daardoor minder kansrijk in het internationaal reviewproces.

Onderzoekers moeten kunnen aantonen dat hun toepassingen op een EuroHPC-schaal supercomputer effectief kunnen werken.
De mogelijkheden voor Nederlandse onderzoekers om zich voor te bereiden op het gebruik van EuroHPC-faciliteiten zijn nu echter beperkt. Lidstaten die zelf in een EuroHPC-systeem hebben mee-geïnvesteerd kennen die beperkingen niet, omdat zij zelf over het toegangsbeleid van hun aandeel gaan. Een groot deel van de Europese landen heeft reeds een

investering in één van de pre-exascale systemen gedaan.[3] Voor de Nederlandse situatie is dit vergelijkbaar met de huidige verdeling van rekentijd via NWO die onderzoekers onder meer gebruiken in voorbereiding op PRACE aanvragen. Een Nederlands aandeel in een EuroHPC-faciliteit heft beide bovengenoemde beperkingen op en vormt zo in beide opzichten een aanvulling op de nationale HPC-faciliteiten.

3.4 Investeren in kennis en kunde

Om grootschalige rekensystemen binnen een acceptabel investerings- en energiebudget te houden, bestaat een steeds groter deel van de faciliteiten uit nieuwe, accelerator-based hardware-technologieën. Dat geldt zeker ook voor de EuroHPC-systemen. Het gebruiksprofiel van Nederlandse onderzoekers kent voor een aanzienlijk deel nog conventionele rekentechnologieën. Zonder maatregelen dreigt een gat te ontstaan tussen het gebruik van Nederlandse onderzoekers en de ontwikkeling van toekomstige exascale HPC-faciliteiten.

Een aantal maatregelen wordt nu reeds genomen. In Snellius is het aandeel accelerator-based hardware significant toegenomen en via het promising applications programma krijgen onderzoekers ondersteuning van SURF bij het inzetten van nieuwe technologieën en methoden voor hun toepassingen. Investeringen in faciliteiten gaan hand in hand met investeringen in kennis over de modernisering en opschaling van applicaties, en zijn dus cruciaal om ervoor te zorgen dat ook in de toekomst een veel groter deel van de Nederlandse wetenschappelijke toepassingen effectief gebruik kan maken van (pre-) exascale faciliteiten.

3.5 Investeringsmomentum

In 2021 is gestart met de realisatie van de LUMI en Leonardo pre-exascale systemen (en de MeluXina, HPC Vega, Karolina en Discoverer petascale-systemen). Nederland heeft daarmee geen invloed meer op de samenstelling van deze systemen en de systemen zijn inmiddels op weg in hun economische levensduur van vijf tot zeven jaar. De langetermijnstrategie HPC beveelt aan dat een investering bij voorkeur gebeurt in een vroeg stadium van ontwerp, zodat de voorkeuren van Nederlandse onderzoekers daarin kunnen worden meegenomen. Daarmee ligt het voor Nederland voor de lange-termijn voor de hand om te investeren in één van de twee toekomstige exascalesystemen.

Tegelijkertijd moet het wetenschapsveld zich kunnen voorbereiden op het gebruik van exascalesystemen. De exascalesystemen komen naar verwachting pas 2024-2025 in productie. Een korte-termijn investering in

---

[3] https://eurohpc-ju.europa.eu/discover-eurohpc-ju#ecl-inpage-211

een bestaande pre-exascale voorziening kan in 2022 resulteren in het beschikbaar komen van rekentijd, om alvast ervaring op te doen met rekentoepassingen en opschaling vanaf de nationale supercomputer richting pre-exascale en exascale.

3.6 Visie op de bekostiging

De Rijksoverheid heeft in 2018-2019 geen middelen vrijgemaakt om te participeren in de opzet van de pre-exascale (en petascale) systemen. In de langetermijnstrategie HPC wordt benadrukt dat een investering in EuroHPC niet moet worden geprioriteerd ten koste van de nationale supercomputer en dat er geen ruimte is binnen de reserveringen die momenteel voor de nationale HPC behoefte worden gemaakt. Een derde (jaarlijks € 2 miljoen) van de deze reserveringen wordt reeds bijeengebracht door de Nederlandse kennisinstellingen. De strategie stelt dat de hoogte van een investering in EuroHPC primair zou moeten worden bepaald door de investeringsbereidheid van de beoogde gebruikers.

In de voorgaande paragrafen wordt geïllustreerd dat buiten de directe meerwaarde voor gebruikers, een deelname in een EuroHPC-systeem een versterkend effect heeft op het gehele HPC-ecosysteem. Deelname in EuroHPC heeft daarmee kenmerken van een collectieve voorziening, welke niet enkel kunnen leunen op de investeringsbereidheid van de directe vraag.

De eerdergenoemde brandbrief van onderzoekers is indicatief, dat binnen het huidig financieringsbestel er voor de wetenschap geen ruimte is voor extra investeringen in EuroHPC. Kennisinstellingen worden geconfronteerd met een groeiende vraag naar lokale reken- en datavoorzieningen in alle wetenschapsgebieden. Investeringen in digitalisering legt een steeds groter beslag op het budget van kennisinstellingen. Deze investeringen van instellingen in lokale voorzieningen en de kennis en expertise die daarbij hoort, zijn cruciaal voor een gezonde basis van het HPC-ecosysteem. Deelname in EuroHPC zou niet ten koste mogen gaan het vermogen van instellingen om tegemoet te komen aan deze groeiende vraag.

De regiekosten en doorlooptijd van het organiseren van de vraaggeoriënteerde bekostiging waar de langetermijnstrategie op dit moment op is geënt, past niet bij de slagkracht die het momentum van ontwikkelingen als EuroHPC vragen. Om deze investering te realiseren is het binnen het huidig kader noodzakelijk nieuwe financieringsbronnen aan te boren.

4 Aanbevelingen

1. Wetenschappelijke concurrentiepositie: Zorg dat wetenschappers Europees kunnen blijven participeren

Voor wetenschappers is het belangrijk dat zij internationaal competitief blijven en de internationale aansluiting niet verliezen.

Zorg dat middelen die nu eenmalig voor de matching voor EuroHPC R&I projecten door het Rijk beschikbaar zijn gesteld, een blijvende faciliteit wordt voor de duur van de EuroHPC Joint Undertaking. De 50% matching overstijgt in de praktijk de draagkracht van kennisinstellingen en kennispartners uit de private sector. De investering door het Rijk in een matching levert meer middelen uit het EuroHPC werkprogramma op en vorm zo een multiplier voor Nederlands HPC-gefaciliteerd onderzoek.

Versterk daarnaast de kennispositie van de Nederlandse onderzoekers in relatie tot het EuroHPC werkprogramma en gerelateerde onderdelen in de werkprogramma's van Digital Europe en Horizon Europe. Veldpartijen moeten onderling zorgen voor meer coherentie tussen participatie van Nederlandse onderzoekers in het Europees HPC landschap en de matchingbehoefte naar de overheid. Zo ontstaan zwaartepuntvorming van Nederlands HPC-gefaciliteerd onderzoek in Europa en betere benutting van de beschikbare matchingmiddelen. Structurele monitoring van calls, de organisatie van een gesprekplatform voor kennisbundeling, consortiumvorming en destillatie van algemene behoefte vormen een aanvulling op de eigen inspanningen van individuele onderzoekers en kennisinstellingen.

2. Aansluiting op het internationale HPC-ecosysteem: Blijf onderdeel van de Europese HPC ontwikkeling door een investering van 10 Mio in EuroHPC-faciliteiten

Door een investering creëren we:

1. Directe toegang tot EuroHPC-faciliteiten voor Nederlandse wetenschappers, onafhankelijk van het op wetenschappelijke excellentie gebaseerde toegangsproces;
2. Kennis en technologische oplossingen om onderzoekers te helpen met het opschalen van applicaties en data in het gefedereerd HPC-ecosysteem;
3. Invloed op de ontwikkeling van het gefedereerd HPC-ecosysteem;
4. Een extensie op de accelerator-based capability van de nationale supercomputer Snellius;
5. Bestendiging en mogelijk uitbreiding/versteviging van de bestaande Europese samenwerkingsverbanden op het gebied van

HPC. Op dit moment is Nederland/SURF een gewaardeerd partner. Als we niet mee-investeren, verliezen we deze positie die de afgelopen 15 jaar zorgvuldig is opgebouwd.

De optimale investeringsstrategie op dit moment, is daarom een combinatie van een grotere investering in één van de nog te vormen exascaleconsortia, en een beperkte investering in één van de bestaande pre-exascaleconsortia.

10 Mio vormt de ondergrens voor een betekenisvolle investering in een EuroHPC consortium[4], maar blijft tegelijkertijd in verhouding met het investeringsbedrag van de Nationale Supercomputer van 20 Mio. 10 Mio investering resulteert in de helft meer rekencapaciteit die samen de Nationale Supercomputer voor Nederlandse onderzoekers beschikbaar zal zijn. Hiervan wordt geïnvesteerd:

- 8 Mio in een exascaleconsortium. Deze bevinden zich in een vroeg stadium van ontwerp en een investering levert lange termijn samenwerking op, maar deze systemen komen pas omstreeks 2024-2025 beschikbaar.

- 2 Mio in een pre-exascaleconsortium. Deze systemen zijn nu al operationeel zodat per direct door Nederlandse onderzoekers ervaring opgedaan kan worden ter voorbereiding op het effectief gebruik van de exascale systemen. Daarnaast resulteert dit eerder in de tijd in een versterking van de Nederlandse positie in de vorming van Europees HPC-ecosysteem.

De verdeling van 8 Mio om 2 Mio is de huidige inschatting. De uiteindelijke afweging van deze verdeling moet plaats vinden op basis van waar-voor-geld van de investeringen na gesprekken met de consortia.

Het investeringsbudget in EuroHPC-systemen moet daarbij op termijn structureel worden, in aanvulling op de langetermijnstrategie High Performance Computing.

3. Springplank: Geef het toegangsbeleid vorm met de commissie Wetenschappelijk Gebruik Supercomputers

Een Nederlands aandeel in EuroHPC is een onderdeel van de nationaal toegankelijke computersystemen, zoals de nationale supercomputer en dataverwerkingsfaciliteiten. De toegang tot deze faciliteiten wordt

[4] Investeringsbedrag LUMI: € 202 miljoen; LEONARDO: € 120 miljoen; verwacht Exascale: € 500 miljoen.

verdeeld door middel van de Rekentijd Nationale Computersystemen NWO call. Om de totale Nederlandse gebruiksvraag te overzien en de beschikbare capaciteit optimaal te verdelen, ligt het voor de hand ook een EuroHPC aandeel via dit mechanisme te verdelen.
Kijk door toevoeging van EuroHPC capaciteit goed naar de verdeling van aanvragen over de nationale supercomputer (tier-1) en EuroHPC (tier-0). In het huidig toegangsbeleid worden aanvragen van rekentijd beperkt tot ca 10% van de jaarlijkse economische capaciteit. De capaciteit in EuroHPC kan gebruikt worden voor zeer grote rekentijdvragen en rekentijdaanvragen die een andere systeemarchitectuur vereisen.

4. Kennis en Kunde: Zet in op Capacity Building met de domeinwetenschap

Door onderzoekers te stimuleren om gebruik te maken van nieuwe rekentechnologieën kunnen zij optimaal gebruik blijven maken van toekomstige faciliteiten en blijven zij internationaal concurrerend.

De ontwikkeling van rekentoepassingen vindt plaats in wetenschappelijke domeinen als onderdeel van lopend onderzoek. Het stimuleren en adopteren van nieuwe rekentechnologieën kan daarom alleen in samenwerking met de domeinwetenschap zelf plaatsvinden.
We pleiten daarom voor een domeingerichte aanpak waarbij programmatisch de inzet vanuit domeinwetenschap bij de kennisinstellingen, de digitale infrastructuur bij SURF en het Netherlands eScience Center en financieringsinstrumenten van NWO wordt gebundeld, om in een tijdsbestek van vijf jaar - binnen de levensduur van het huidige tier-1 systeem - doorbraken te forceren. Het naderend einde van Moore's Law vereist dat opschaling van toepassingen nieuwe aanpakken vraagt, met accelerators, machine learning-benaderingen en in de verdere toekomst quantum computing.

Hierbij ligt de focus niet op de onderzoeksvraag zelf, maar op de applicaties, om zo de potentie tot verduurzaming over onderzoeksprojecten, vakgroepen en landsgrenzen heen te bevorderen via samenwerking en generalisatie.

Met flankerend beleid op communityvorming, kennisdisseminatie en expertise opbouw via onder meer de thematische DCC's, kan in samenwerking met instellingen en discipline-specifieke infrastructuren gewerkt worden aan bestendiging en olievlekwerking vanuit de voorlopers.

5 Use Cases

1. *'Our scientific infrastructure should be prepared for all kinds of extreme events like a pandemic or the reversal of the Gulf stream'*

Prof. dr Detlef Lohse

2. *'We need access to Exascale computing in the Netherlands in order to stay relevant in fundamental science'*

Prof. dr. Simon Portegies Zwart

3. *'Innovation not only means pushing the boundaries in science, but also in the infrastructures you set up'*

Prof. dr. ir. Alfons Hoekstra

4. *'Access to Exascale computing will lead to a new paradigm in climate research'*

Prof. dr. ir. Henk Dijkstra

5. *'The more HPC capacity becomes available, the more we can predict'*

Dr. Jocelyne Vreede

6. *'If we in the Netherlands want to maintain our innovation power, we must provide top talent with top tools'*

prof.dr.ir. Jos Benschop, ASML

7. *'Computational modelling from the atomic or molecular level is traditionally a strong field in the Netherlands. We now need Exascale HPC infrastructure to compete'*

prof. dr. Siewert-Jan Marrink

8. *‘This competition is not about cleverness or ideas. It is all about the ability to perform experiments’*

Prof. dr. Cees Snoek

9. *‘The life sciences are at the beginning of a data analysis explosion; access to Exascale HPC is needed as part of a strong distributed data infrastructure’*

Prof. dr Wiro Niessen

10. *‘It would be beneficial for our Dutch knowledge economy if we can strengthen our expertise of software and algorithms with Exascale HPC and European collaboration’*

Dr Menno Genseberger, Deltares

1.

There are huge challenges for mankind where fluid dynamics can majorly contribute to solutions. Fluid dynamics is used to solve issues in the areas of Health, Energy, High Tech Industry, and Climate. For example on the latter, fluid dynamics studies precursors that could warn us that the Golf stream is changing directions. In the High Tech Industry, it enables techniques with which microchips can become even smaller than they are now.

Because of the complex models that are used, HPC plays an important role in all of the fluid dynamics research. Access to exa-scale computing via EuroHPC means that the models can become more detailed, and for instance enable research on how ventilation can help to avoid COVID-19 contagion via aerosols in offices.

‘Our scientific infrastructure should be prepared for all kinds of extreme events like a pandemic or the reversal of the Gulf stream’

Prof. dr Detlef Lohse

HPC plays an important role in all of our research

‘I have been Professor for Fluid Dynamics in Twente for 23 years now. My Physics of Fluids group is interested in turbulence and multiphase flows and micro- and nanofluid dynamics, including molecular dynamic simulations. All these subjects are extremely demanding in terms of computer power, like in most other fields of research in which fluid dynamics play a role.

There are huge challenges we face nowadays where fluid dynamics can majorly contribute to solutions. For instance, you can use fluid dynamics to solve issues in the areas of Health, Energy, Climate and also for High Tech Industry. No matter whether we study the behaviour of a single drop of liquid or large-scale turbulence, we always use complex modelling. So you can imagine that HPC plays an important role in all of our research.’

Fluid dynamics for Climate

‘My group studies turbulence in the atmosphere and turbulence in the ocean. A major challenge is to predict the developments of turbulence using different models. The hierarchical structure of many different length scales in a turbulence flow requires a huge amount of computational time. And to make it even more complex, in the ocean you have a phenomenon called double diffusive convection, which means that the turbulence is not only driven by heat differences, but also by differences in salinity.

Take the Gulf Stream, for instance. At the moment, it goes from the warm Caribbean Sea, which is also very salty because of the evaporation of water, to the cold Arctic Ocean, which is less salty because of all the melted ice. These two driving forces oppose each other. On the one hand you have a density difference due to salinity. On the other hand a density difference due to temperature. The outcome is that the stream flows in one direction. The key question is now: is it possible that this Gulf Stream will change direction when something changes in this temperature/salinity balance?

This would be a major catastrophe for civilizations in Europe. It would mean that it gets much, much colder than it is now. Within model systems, we are looking for any precursor that could warn us in advance, and we want to understand the statistical features of a flow reversal like that. This is an example of how Fluid Dynamics can help overcome one of the major challenges we are facing nowadays. We need major HPC power to do so.'

Fluid dynamics for Energy

'Another example is in the energy field: next to modelling how turbulence drives wind turbines, it is important to solve issues in the transport of liquefied gases. This happens under high pressure. Since the temperature is close to boiling, bubbles can form. Companies want to better understand how these liquefied gasses flow to avoid this. Therefore you must develop numerical methods and numerical schemes to calculate the turbulence and such, using HPC. This is needed to transport these gasses as safe and efficient as possible.'

Fluid dynamics for Health

'Ventilation is crucial in the context of the Covid-19 pandemic. At the moment it is possible to direct numerical simulation of the flow of air with two people in a room, with some ventilation. What we want to study is how to adapt ventilation to get rid of aerosols efficiently, in particular when these people start to move. At the moment, this is borderline possible; it takes a long time to calculate these models. But you can imagine that there is real value to be found there if we scale the HPC capacity up to the exa-scale. We could get a better understanding of how to best ventilate offices in order for people to safely get back to the work place.'

Fluid dynamics for High Tech Industry

'The Dutch company ASML builds huge machines to allow for mass production of patterns on silicon, which is used by all of the world's leading chipmakers. Their lithography solutions allow microchips to get smaller and smaller. They use droplets of liquid tin with a high performance laser to create plasma. They want to know how to do this in an efficient way using models, so fluid dynamics is crucial for them and we have a great partnership. However, calculating these models at the moment costs months. It would be very beneficial if we can take calculations like that into the Exa-scale dimension.

The fluid dynamics of droplets is also important in the professional printing industry. For instance, you don't want droplets of ink to mix. We have a long-term working relationship with printer manufacture Canon Production Printing, for improving inkjet printing. Again, there is a lot of industrial value to gain there if we can scale up. And if it becomes possible to take this even further to 3D printing, there is a lot to gain. I'm pretty sure that that within the next 20 years, it will be possible to 3D print organs. The kidney, for example, has only 20 different cells. Skin can already be printed today as the structure is not so complicated. This brings us back to Fluid dynamics for Health, and again, all this is only possible if the right HPC capacity is available.'

EuroHPC can accelerate findings enormously

'When using models, you want to go towards systems with more realistic geometries like rough walls on the bottom of the ocean or landscape. Or interaction of turbulence with the vegetation. For instance to determine how to place larger amounts of wind turbines in an optimal way, either at sea or on land, you could take into account how the structure of a forest influences the wind turbines as well. The sheer amount of calculations that we need, ask for exa-scale computing power.

Overall, if we can use the planned EuroHPC instances for fluid dynamics like in the examples I mentioned, we could scale up the complexity of the different models. Therefore we can get to solutions on a completely different scale much faster. For instance, if we could scale up the complexity of our ventilation models, to be able to simulate several people in a room with more complicated geometries, and different elements to determine the best way to ventilate, that would be very beneficial to society. The ideas are there, but it is non-trivial to execute them. For that we need major, major computational power.'

EuroHPC is important for the Netherlands

'Having good high performance computing facilities available for a diverse array of research topics is of enormous importance for a country like the Netherlands. We should not only invest in our local infrastructure, but also in the EuroHPC so that researchers can access an infrastructure that will be 100 times larger than the national infrastructure. The benefits are enormous, not only in the field of fluid dynamics, but in other fields as well.

All our projects require major computational skills. How to handle, transfer, store and visualise vast amounts of data is also a challenge. We are very happy with SURF, who do a great job in advising, training and helping us. That is a huge difference compared to just 'buying' HPC capacity in for instance the US or China. Our Chinese colleagues say they might have the hardcore HPC power there, but not the infrastructure to use it in such an efficient way. This infrastructure must be kept updated and improved continuously.

We need long term thinking to really be a resilient society. Our scientific infrastructure should be prepared for all kinds of extreme events like a pandemic, or like the reversal of the Gulf stream.

For instance, all of the sudden, because of COVID-19, there is now an enormous interest in our research on aerosols and ventilation. No one saw that coming. A lesson learned is that you need diversification to be prepared for the unknown. We must ensure that our research infrastructure is flexible enough to be able to effectively react to the grand challenges that are ahead. Investing in having enough High Performance Computing available is a very crucial part of this.'

Quotes:

'You can use Fluid dynamics to solve major problems in the areas of Health, Energy, Climate and in the High Tech Industry'

'The Dutch company ASML builds machines which enable the mass production of patterns on silicon. These machines are used by all of the world's leading chipmakers'

'If we can use the planned EuroHPC instances, we could scale up the complexity of our models. Therefore we can get to societal and industrial solutions on a completely new scale, and much faster as well'

'We need long-term thinking to really be a resilient society'

'A lesson learned from this pandemic is that you need diversification to be prepared for the unknown'

Prof. dr Detlef Lohse studied in Kiel and Bonn and did his PhD at the University of Marburg. His PhD thesis, in 1992, was about fully developed turbulence. He joined the University of Twente in 1998 as a Full Professor of Physics of Fluids. Lohse received many prizes, like the Max Planck Medal, the Balzan Prize, the George Batchelor Prize, the Fluid Dynamics Prize of the American Physical Society, the Dutch 'Simon Stevin Meester' and the Spinoza Prize. He is a member of the Royal Netherlands Academy of Arts and Sciences (KNAW), the German Academy of Sciences 'Leopoldina' and of the American National Academy of Engineering.

cc0

more pics available via Detlef

2.

In fundamental science such as astrophysics, most of the recent breakthroughs would not be possible without HPC. With access to Exascale HPC, simulations that include all hundred billion stars in the galaxy are suddenly within reach. The formation of the solar system or the evolution of the galaxy can be explored.

World leading expertise in algorithmic development is already available. It is now of the highest importance that the Netherlands invest in Exascale HPC, to give researchers a chance to maintain their skills and stay on the forefront in fundamental science.

'We need access to Exascale computing in the Netherlands in order to stay relevant in fundamental science'

Prof. dr. Simon Portegies Zwart

'I'm professor of numerical star dynamics at the university of Leiden. What I do is fundamental science. We use supercomputers to explain the age of the universe, the formation of the solar system, its future, the structure of the galaxy and the measurements of the cosmos and the emergence of life.

You can't study the universe in the laboratory, but black holes can collide with each other in a super computer. When we have access to Exascale computing it will be possible to simulate all the stars in the galaxy: a hundred billion stars. We have the software to do it. We have the expertise. If we can combine that with EuroHPC power, there is so much we can do! Understanding of the formation of the solar system or the evolution of the galaxy will suddenly be in reach.'

Breakthroughs that need HPC
'A lot of recent scientific breakthroughs were only possible because of huge HPC power. Think of the recent 'picture' of the black hole in the M87 galaxy. Think of gravitational wave detections; all completely impossible without supercomputers. Actually, I would say that more than 90 percent of the papers in fundamental fields of science would be impossible without some form of HPC, nowadays.

Still, the six 6m wide telescopes, 5000 metres high up in the Atacama Desert in Chile get a lot of attention. It is fantastic to have them high up there, but without the large HPC instances next to them, these telescopes would be useless. Another example is Gaia, the satellite that 'maps' approximately one billion stars. That is a 200 million euro instrument. Apparently that is what the public is ready to pay for this type of research. Then why not buy a supercomputer that you can use to know the position of the hundred billion stars in the galaxy, so a hundred times as many stars?

It is worrisome that Europe is roughly 10 years behind in technology compared to Japan and the United States. And to make it worse is that in my perspective, the Netherlands are another 10 years behind the Swiss or the Germans here in Europe. That is devastating. And the longer we wait with real investments, the more we will be lagging behind.'

Given the ideal hardware, we can basically solve every technical problem
'Having the ideal software, and given the ideal hardware that is available nowadays, I truly believe that we can basically solve every technical problem. From unimportant, annoying things like teapots that drip to the energy efficiency of cars, all the way up to fundamental science.

We for instance try to understand how stars form in groups. That is relevant for us humans, because the Sun and Earth were formed in such a group. If we better understand how stars form groups, we will better understand how the solar system was formed and how the earth is formed. To give you an example, I would like to understand what the chances are that Earth will be hit by an asteroid. Earth is habitable planet and it's quite important to keep it habitable. Asteroids are notoriously dangerous for life on planets; as the dinosaurs have noticed. So it is very important to understand the probability that this happens again, and when this could happen.'

We can't lose the expertise to work on the highest level of computing in the Netherlands
'If you are world leading in algorithmic developments, like we are, you need to build up HPC skills and expertise in your team. Therefore, you have to be able to give researchers the opportunity to work on different levels of HPC instances.

It is just like driving: you first have to know how to drive a normal car, before you can drive a sports car. And if you want to drive a formula 1 car you have to re-adjust everything again. So as a student, you learn to program on a laptop on a single core. As a grad student, you start learning about multiple nodes. Then you begin to work on the local HPC, before using more powerful national supercomputers, all the way up to for instance the EuroHPC. Every time you scale up, your calculations scale up with the same order of magnitude. Everything has to be redesigned, which requires expertise. We have to ensure that Dutch researchers get an opportunity to learn to do that, in order to keep up with others who do world class research on the Exascale.

Moreover, with EuroHPC, you can cooperate with other scientists who operate on the Exascale level. You can learn and benefit from each other. This is equally important to stay relevant.'

We could deliver home grown world-class research
'Japan has top notch HPC facilities which are completely based on Japanese technology. Costs for this were around 200 million euros. The Unites States and

China spend a lot of money on HPC as well. Compared to those countries we are not even in the mid-range anymore.

However, at SURF we have one of the best maintained supercomputers in Europe. Not the fastest, not the most powerful at all, but the whole combination of disk facilities, storage facilities, internet facilities and support is excellent. If we can combine that level of service with EuroHPC Exascale power and the right type of applications, and we are able to ensure our scientists are well skilled to work with these HPC instances, that would be an ideal combination. Then we can ensure we are no longer lagging behind other countries. We could deliver our own home-grown World-Class research again.'

Quotes:

'You can't study the universe in the laboratory, but black holes can collide with each other in a supercomputer'

'A lot of recent scientific breakthroughs would not have been possible without huge HPC power'

'It is worrisome that Europe is roughly 10 years behind in technology compared to Japan and the United States'

Visualisation of a simulation that includes 1 billion stars: only 100 million stars are visible but not the dark matter.

3.

Professor Hoekstra's team at the Computational Science Lab works on multiscale models for hemodynamics, as well as on computational environments for in-silico clinical trials. Hoekstra sees great benefits in High-Performance Multi Scale Computing: thousands of instantiations of coupled simulations that can easily be executed to enable parameter sweeping or uncertainty quantification at scale, using HPC.

If EuroHPC would enable Multi-Scale Computing on the exa-scale level it would save an enormous amount of money and time. For instance when validating in silico clinical trials that replace real people with 'virtual' ones, to avoid having to run costly and time consuming clinical trials. Not only can the results be put to use by European companies, it can also, for instance, reduce the complications when placing stents in arteries and in the treatment of strokes.

'Innovation not only means pushing the boundaries in science, but also in the infrastructures you set up'

Prof. dr. ir. Alfons Hoekstra

High-Performance Multi Scale Computing
'Typically, when given additional computational power, scientists want to dive deeper into a complex problem. There are many reasons to scale up the complexity of the problems that we study, or to speed up processes so that the time to solution is more effective.

However, I'm involved in a series of projects on what I call High-Performance Multi Scale Computing and multiscale computing pattern. In such pattern, replica computing, you don't use the given extra HPC capacity to add complexity. You load thousands of jobs simultaneously instead; jobs that can also be executed in a more modest environment.

There is value to be found there, next to the more traditional use of the available HPC, because it makes your outcomes more robust. For instance, it allows to put error bars on the outcomes using Uncertainty Quantification techniques that rely on executing thousands of runs with slightly different settings of parameters and inputs. It opens up a world of further possibilities.

Over the last four years, I participate in a project called INSIST *IN Silico trials for treatment of acute Ischemic STroke*. One aspect of this project is to work on using computational techniques to mimic real clinical trials to test new or better devices. Clinical trials are very costly and time consuming procedures. Therefore there are huge benefits to using in silico clinical trials to augment or enhance the real trials. In these trials “virtual” patients are treated with “virtual” treatments.

The European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) request that you validate models you use for these in silico trials. The only way to do so is by running very extensive uncertainty quantification campaigns. In terms of computing, this means running thousands and thousands of models. At the moment we are doing this using the national HPC instances at SURF. However, this would be an excellent use of the new EuroHPC instances; the costs of validating in silico trials would reduce significantly on the exa-scale level.

I would be really happy if we can show that through EuroHPC we are able to validate these biomedical models quickly and cheaper. This raises our competitiveness. Not only can the results be put to use by European companies, we can also help reduce the complications in for instance the treatment of strokes as a result.’

Run thousands of simulations in order to understand what happens
‘Another project I’m working on is using suspension simulations. This project is all about understanding the collective motion of millions of red blood cells. With simulating the behaviour of platelets close to vessel walls we can understand the behaviour in microfluidic devices, which we can than apply in clinically relevant settings. We also look at the probability of thrombosis.

For instance, lots of people in the Netherlands have clogged arteries. Placing a stent, a small medical device, in a coronary artery is a standard procedure. However, it turns out that there is a kind of wound healing response that occurs after a small percentage of procedures. This response can result in the potential re-closing of the artery. Since this procedure is carried out so often, there are lots of people who suffer as a consequence. You want to avoid that.

We are now capable of doing simulations in the order of one cubic millimetre of blood simulating the behaviour of all individual red blood cells, using the compute power that is offered by the current Dutch national supercomputer. We still need to run thousands and thousands of these types of simulations, again to understand the impact of variations in parameters of the model, but also to be able to extract constitutive equations to be used in coarse grained models. It would be ideal if we had access to exa-scale computing for this.’

Train surrogate models on the exa-scale HPC
‘Also, as a next step, we can run surrogate models on the fly. Each time you finish running a model you have data on which you can train a surrogate. Instead of running 10.000 simulations in one shot, you can first run 500. You use the data acquired from that run to train your network with another run of 500. Then, at some point your surrogate network reaches a level fidelity where you can run that surrogate instead of the real one. That would be excellent news for those that cannot access Tier one or Tier zero machines themselves: to run these

surrogates we don't need that kind of computing power. It would mean we only have to use Tier-0 or -1 if we want to change little things in the set up and we need to retrain the surrogate model. That would be highly efficient use of the EuroHPC instances.'

Investing in EuroHPC means having a voice

'If you want to innovate, you not only have to push the boundaries in the science that you do, but also in the machines and in the infrastructures you that you set up. Innovations on the exa-scale will trickle down immediately. Next to the examples I mentioned, all kind of Tier -0 innovations, such as new algorithms, will be used in Tier-1 and Tier-2 HPC instances not much later.

The Netherlands should not only invest in the EuroHPC to get access in terms of computing time and power. Joining also means being able to co-design how to use HPC on the exa-scale level. We already know that different scientific domains all have the same type of computational patterns of how different components interact with each other, even though they all use different models and software. Hopefully, when joining the EuroHPC, we create opportunities to explore this further.

There is lots of potential to learn from each other. We really need a Schengen zone for HPC, if you ask me. We could for instance expand our work on the coupling of machine learning, the flexibility of computing, trial new multi scale surrogate models. We would have the capacity to adapt to new paradigms.

I'm convinced you will also see an interesting merge of the current data driven systems and the kind of mathematical models based computing we use in High-Performance Multi Scale Computing. For our next innovative steps we need to do both. The EuroHPC can enable this, and therefore it can open up all kind of fantastic new possibilities.'

Links:

Computational Science Lab

Quotes:

'We need to run tens of thousands of simulations to make our outcomes more robust'

'The costs of validating In Silico models would reduce significantly on the exa-scale level'

'We really need a Schengen zone for HPC'

'Joining EuroHPC also means being able to co-design the EuroHPC'

Bio:
Prof. dr. ir. Alfons G. Hoekstra is full professor Computational Science and Engineering at the University of Amsterdam. He is also the director of the Informatics Institute at that university. His research focusses on multiscale modelling, actionable simulations, and high performance computing. His research is driven forward by and applied in the biomedical domain. His main interests are currently in multiscale modelling of hemodynamics with applications in cardiovascular diseases, and in the development and accreditation of in-silico trials.

Photo:Dirk Gillisen

4.

There is a great need for scenarios of what our actions in the next eight years mean for climate on the longer term. New, high-resolution climate models are already delivering high-impact breakthroughs. Exascale systems would reduce the current turnaround times (of about a year) for typical simulations with these new models enormously. This would allow the community to come up with urgently needed results faster than ever: from changes in the Gulf Stream to regional sea level changes. When studying climate change, there is no time to lose.

‘Access to Exascale computing will lead to a new paradigm in climate research’

Prof. dr. ir. Henk Dijkstra

There is a urgent need for plausible scenarios of climate change

‘I'm a professor of dynamical oceanography. In this field we study ocean circulation and the climate system from a physical point of view. We use a hierarchy of different climate models and HPC is required for most of these.

There is a urgent need for plausible scenarios of climate change and its regional impacts up to the year 2100. What is going to happen to precipitation in Europe, in particular over the Netherlands? How fast will the sea ice melt in the Arctic? Will that affect the path of the Gulf Stream, which would have consequences to everyone living in Europe?’

An international breakthrough based on HPC

‘Recently, using a very high-resolution climate model, our group found out that resolving ocean eddies strongly affects projected global mean sea-level rise in the coming 100 years. This value is about 25% lower than found by currently used climate model simulations in which eddies are not captured. This was the result of a few years of work, and could not have been done without major HPC power.

If we can run these very high-resolution models on Exascale systems which are 500 times more powerful than Cartesius that we use now, we can decrease the current turnaround time of a few years enormously. This will lead to a new paradigm on climate system behaviour, and far better projections of climate change, including regional effects.

Everybody in the European climate research community wants that kind of HPC machines for years already. We know the climate models we use scale very well. Knowing that our political and societal actions in the coming years will heavily influence the future of all of us, there is no time to lose.'

Complex climate models
'To get an idea of how complex some of our used climate models are, you can think of weather prediction models, to which you add extra detailed components of the climate system such as the ocean, land and ice. You then integrate these models in time to determine the future state of the climate system (atmosphere, ocean, land and ice). So that is significantly more complex than weather prediction models. And instead of 14 days predictions, climate models make projections on a time scale of a hundred years.

Next to that, the computational effort needed depends on the spatial resolution; the size of the grid boxes used to represent the physical balances in the climate system. Most of the current models that the Intergovernmental Panel on Climate Change uses in their upcoming 5-yearly report have 100 km^2 boxes in the ocean. In my group we work with models with boxes of 10 km^2 instead. Therefore, smaller-scale ocean processes can be taken into account and ocean eddies can be represented and, as a result, climate projections become far more accurate.

These are very expensive simulations, which at the moment take about a year to run, and years to analyse. If we are able to run these models on the Exascale level, we will be able to narrow down possible scenarios of climate change and we can do so much faster as well. This is why we so urgently need access to Exascale HPC.

Next to that, Exascale HPC will make it possible to better incorporate important components of the climate system, such as the ocean marine biosphere which is essential for CO_2 regulation. The ocean has so far taken up about 25 percent of the CO_2 emissions of humans, but that can change under climate change. To incorporate such components in the very high-resolution models is unfeasible on the HPC systems that are currently available: one such simulation would take about 100 years. Once the limitation in compute power is lifted, we can learn a lot more. It is of the utmost importance to know the development of the marine carbon cycle; changes that happen now can have dramatic consequences in the future.'

Collaboration can strengthen the position of the Netherlands
'Collaboration on the EuroHPC level could not only strengthen the position of Europe as a whole, but also the position of the Netherlands. For instance, I'm also involved in a Dutch research consortium that deals with salt intrusion in the Dutch delta from the Rhine and the Meuse (Maas) River. In the drought of 2018, salt from the North Sea came up the rivers all the way to Gouda. Salt in our rivers can spoil our drinking water. Vegetation dies so agriculture is affected hugely as well. This is a very direct threat to the Netherlands. So we look at what should be done do to keep the salt out of the Dutch rivers.

Because we have such high resolution models we can actually make an estimate of the chances of droughts in the summer, enhanced rainfall in the winter and the impact on the saltiness of the water in the Rhine area. This is an eight million euros project. This kind of project does not only have a direct impact for the Netherlands; knowledge about these kind of water related issues is also used by Dutch companies all over the world.

As said, the available HPC power is a limiting factor at the moment. If we can access Exascale instances, we will add so much more to projects like this that we are involved in. We have the high-resolution models. We have highly trained experts in our team. We could bring out research to the next level and come up with urgently needed results for policy making faster than ever. As said, when studying climate change, there is no time to lose.'

Quotes

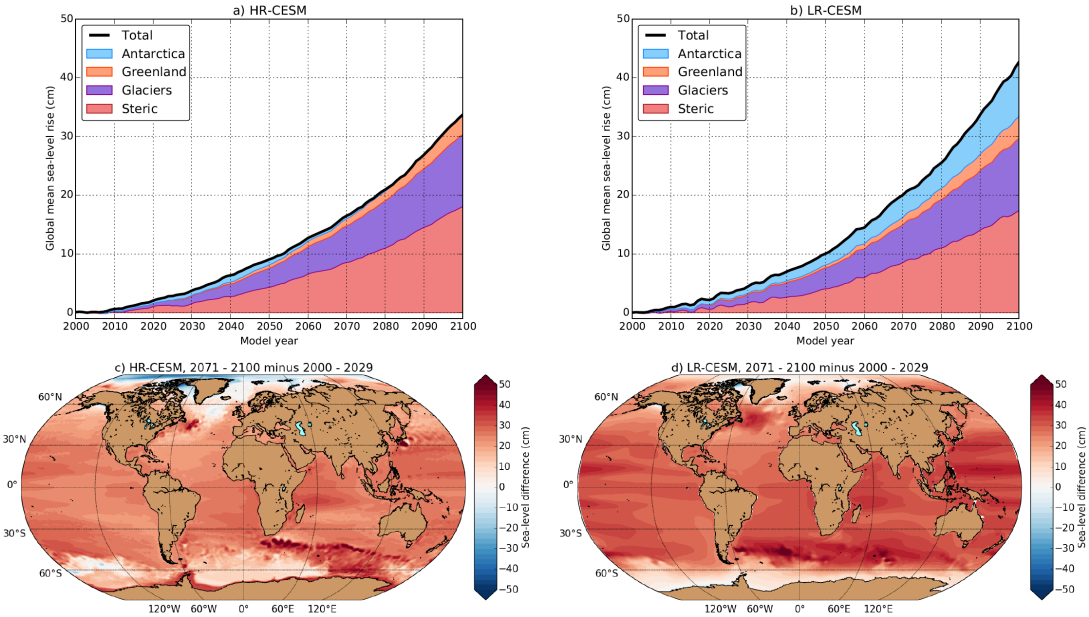
'Access to Exascale computing will lead to a new paradigm in climate research, and far better scenarios of climate change over the next decades'

'Knowing that our political and societal actions in the coming years will heavily influence the future of all of us, there is no time to lose'

'Knowledge about these kind of water related issues is used by Dutch companies all over the world'

Bio:
Henk A. Dijkstra is professor of Dynamical Oceanography at the Institute for Marine and Atmospheric Research Utrecht and director of the Centre for Complex Systems Studies within the Department of Physics of Utrecht University. He was trained as an applied mathematician and held positions at the University of Groningen, Cornell University, and Colorado State University. His main research interests are on climate variability, in particular climate transitions and climate change with a focus on the role of the oceans. Since 2002, he is a full member of the Royal Netherlands Academy of Arts and Sciences and since 2009 he is a Fellow of the Society for Industrial and Applied Mathematics. In 2005, he received the Lewis Fry Richardson medal from the European Geosciences Union.

private photo



Caption to the figure: (available in PDF)

(A and B) Contributions to global mean sea-level rise over the 101-year period for the very-high resolution climate model and the low-resolution climate model, respectively; the black curve indicates the total global mean sea-level rise. (C and D) The local sea-level difference between the time mean over years 2071–2100 and time mean over years 2000–2029 for both type of models (source: van Westen and Dijkstra, Science Advances, 7, eabf1674, (2021).

5.

Dr. Jocelyne Vreede studies molecular processes such as how different conditions influence how the enzyme lactase performs its function. Because of the sheer amount of atoms needed to predict their behaviour, their calculations require HPC. If Exascale HPC power becomes available, predictions will become more accurate. The interest from industry would grow accordingly.

Vreede is working in a very active field of research that looks at how to use machine learning in extracting the relevant information from molecular systems. The long term goals of the projects she is working on are to get to better treatment for lactose intolerant people, or to be able to screen for potential cancer medications.

'The more HPC capacity becomes available, the more we can predict'

Dr. Jocelyne Vreede

Molecular processes that are relevant for living things

'My specific domain is biochemistry. I look at molecular processes that are relevant for living things. For instance, one of my current projects looks at how the enzyme lactase performs its function. Lactase breaks down the lactose in food so your body can absorb it. People who are lactose intolerant can't break down or digest lactose. This leads to unpleasant symptoms such as bloating, diarrhoea, and gas. They can take extra lactase, for instance in pills, to break down lactose in food. We look at what the lactase enzyme does (stick together, bind lactose or other sugar molecules) at different conditions (pH, salt concentration, temperature). If we understand better how different conditions influence how lactase works, we can improve it, so it works better over a wider range of conditions

Lactase is an enzyme, which is a protein, and therefore it is built up from amino acids; around a 1.000 of them. Each of these amino acids consists of about 20 atoms, which means each lactose enzyme consists of about 20.000 atoms. In our computer models we include each of these. And to add to that, we look at systems which include several lactases, dissolved in water. So we're looking at a system of around 200.000 or 300.0000 atoms. And that is just the initial model.'

A completely different scale

‘The next step is to look at how this model changes with time. That is what we need the HPC resources for. If you have a snapshot of the system with the atoms at certain positions, the next possible step is to calculate what forces they exert on each other. All these atoms, they feel each other. They exert forces on each other. And when something exerts a force on something else, that means that it will change velocity. We can calculate what their new positions will be, taking very small steps in time.

The physics behind it is not very complicated. We use algorithms that have been developed over centuries. However, with the compute power that is available to us now, we can use those algorithms on a completely different scale. We calculate new forces, and based on that we recalculate following new positions, and so on.’

Because of the sheer amount of atoms that we look at, these calculations cannot be done without HPC. To give you an idea, the steps in time we use is femtoseconds: one millionth of one billionth of a second. Let's say we want to calculate the behaviour of lactases for a nanosecond, that is one million femtoseconds. That would take years to calculate at the moment. Anything with proteins, or DNA, or membranes, automatically leads to this sort of exponential calculations.’

The dream of looking at larger complex structures

‘I’m working in a very active field of research; how to use machine learning in extracting the relevant information from molecular systems. One of our long term goals would be to look at a protein system and be able to predict the effect of mutations. An even longer term goal would be to predict the structure. That's sort of a holy grail; that you can you have an amino acid sequence and you can predict the three dimensional structure of the protein at various conditions.

With more computational power, you could also start simulate a larger structure of a whole complex of proteins. For instance, the process of DNA replication; that process includes a lot of different proteins and nucleic acids that all are assembled in a specific way. It would be really cool if you could simulate that because DNA replication is one of the essential processes in life, and is the basis of biological heredity. EuroHPC could make this dream of looking at larger complex structures come true.’

Predictions

‘The algorithms get better, the descriptions get more accurate. We are now able to do simulations that last long enough to gain good statistics, so we can improve the description of the interaction between the atoms. We can also speed up certain processes. The more HPC capacity becomes available, the larger the molecular systems and longer the time scales become that we can process in the amount of available processing time. Therefore, the more we can predict. So the HPC power that EuroHPC can bring, would open up all kind of new possibilities for our research, which is excellent.’

Industry
‘When there is exa-scale computer power available, our predictions become even more accurate. Because the predictions get better and better, there is an growing interest for the work that I do from industry. I have two projects in which compagnies are involved, for instance the Lactase project I described earlier. Once you have developed all the right algorithms, computationally screening for something like potential drug molecules is much cheaper than testing them all in the lab.

There are definitely benefits for Dutch industry if the Netherlands co funds the EuroHPC exa-scale systems. For instance, if you partner up with a Dutch university, that university will have access to this new Euro HPC power in combination with someone like me with the specific biochemical knowledge. That is really powerful.’

The long term goal of these predictions
‘Having access to the highest level of computations can help a lot to understand the processes that are going on in all kind of living things. We are not working on issues that you can solve in five years but we are working on the first steps, which is to know how it all works.

I'm also looking at a protein that is involved in the development of tumours. Ultimately, we can use the data on mutations we collect to get to screen for potential cancer medications. On the longer term, my research will have impact on the UNESCO Sustainable Development Goal ‘Good health and well-being’. Certain other of my projects can help to reach ‘Zero Hunger’. Others in my field are working on ‘Affordable and Clean Energy’, for instance in projects that look at water splitting to produce hydrogen for fuel.

We really have to think about accelerating our computational research now. There is so much potential. Participation in Euro HPC can definitely help taking important steps in my field. I’m really curious what the future might bring.’

Quotes:

‘Because of the sheer amount of atoms that we look at, these calculations cannot be done without HPC’

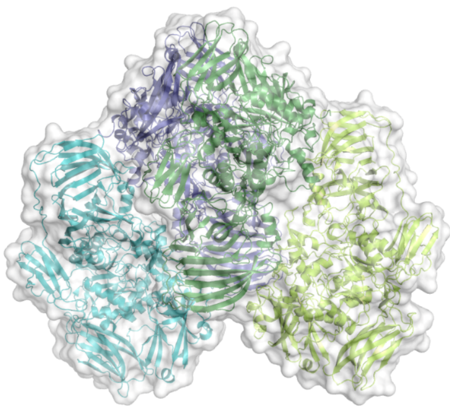
‘The steps in time that we use is femtoseconds: one millionth of one billionth of a second’

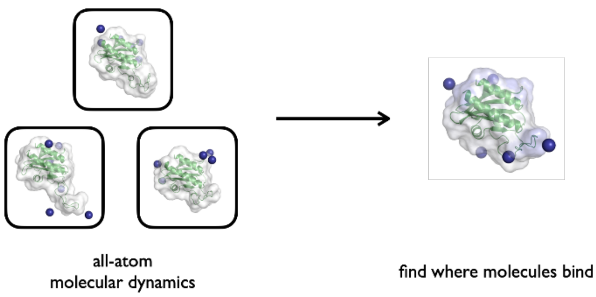
‘When exa-scale computer power becomes available, our predictions become even more accurate. When predictions get better and better, there is an growing interest for the work that I do from industry’

‘We are not working on issues that you can solve in five years. We are working on the first steps, which is to know how it all works’

Biography Jocelyne Vreede

Jocelyne started out as a student in chemistry at the university of Amsterdam in 1996. Already from the beginning, she liked the chemistry of life, but the lab work did not suit her. In 1998 she discovered she was really good with computers, and in 2001 she received her MSc degree in chemistry, specialized in molecular simulation. Jocelyne stayed at UvA, to do a PhD in a microbiology group with the focus on studying signaling processes in unicellular organisms. She performed advanced molecular simulation studies on blue light receptors. In 2007 Jocelyne received her PhD, while already working as a postdoctoral researcher in the computational chemistry group. In 2008 Jocelyne received a veni grant to investigate DNA binding proteins. Since 2011 Jocelyne is an assistant professor in the Computational Chemistry group at the van 't Hoff Institute for Molecular Sciences at the University of Amsterdam. Currently, she manages a research group of around 10 people, focusing on conformational transitions in proteins and DNA.





Caption: The enzyme lactase

6.

Dr.ir. Jos Benschop is Senior Vice President Technology at ASML, which is headquartered in Veldhoven. ASML produces lithography solutions to mass produce patterns on silicon. ASML is an innovation leader in the semiconductor industry. All major chip makers use their technology, and in 2020, ASML spent €2.2 billion on R&D alone.

Next to Senior Vice President Technology at ASML, Benschop is a member of the Advisory council for science, technology and innovation (AWTI) advising the Dutch government and parlement.

‘If we in the Netherlands want to maintain our innovation power, we must provide top talent with top tools’

Dr.ir. Jos Benschop

No direct impact for ASML

‘For ASML, the direct impact of the decision whether or not the Netherlands join the EuroHPC endeavour is none. However, the indirect impact is large. Not only for us at ASML, but for the Netherlands as a whole.

At ASML, we provide our customers with everything they need – hardware, software and services – to mass produce patterns on silicon, allowing them to increase the value and lower the cost of semiconductor chips. Our customers include all of the world's leading chipmakers. They have their own HPC compute power at their disposal. If we need HPC ourselves, we are likely to use supercomputers in the US. We are also based in San Jose, California where we, for instance, study computational lithography.

However, we do work very closely together with academics. For them access to Exascale computing is of the utmost importance. We work with the Lawrence Livermore National Lab in the US but also with the Russian Institute of the Academy of Science to simulate our laser produced plasma source, to name a few. These academics look at certain scientific challenges, often building understanding of phenomena that we don't quite understand.’

World Class scientist in the Netherlands

‘I simply work with the best in the world; we cannot afford to work with second class. Luckily, there are many scientists working at Dutch universities that are world class. I love to work with them. It is of critical importance for us to also have talent in our own local network. Even though we are competing globally, we make a difference locally.

For instance, we work with professionals like Prof Detlef Lohse from Twente University on fluid dynamics in different ways and forms. He, as well as academics at the Advanced Research Center Nanolithography, located in Amsterdam, look for example at the impact of a high energy laser pulse on a liquid. These are very challenging things to simulate.

So even though we don’t need any HPC capacity ourselves, it is critical that we can work with academics that are used to work with high level HPC and that understand our issues.
In order for the Netherlands to be attractive to worldwide academic talent - I’m talking about people with a deep understanding of the phenomena that occur in all we encounter- it needs to give access to state-of-the-art tools and infrastructure.’

Both companies and academics are competing on a worldwide scale
‘Companies like ASML are competing on a worldwide scale. Our competition comes from Japan and our customers, such as Intel, Samsung and TSMC, are mainly based in the USA and Asia.

Academics are also competing on a global scale; there is a worldwide competition for talent. The Netherlands can save small money by not joining a consortium like the EuroHPC. However, that would not only mean Dutch universities lose access to top high-performance computing. It would also mean that the Netherlands will become less attractive for top talent in the longer run.

Moreover, we are too small a country to compete with big guys like China and the U.S., so we have to team up thoroughly in Europe. ‘A good neighbour is better than a faraway friend’ goes the saying. That is true in innovation. We should not be afraid of international mobility, not at all. We do need to have something to offer, in order not to end up on the wrong side of the equation.’

Solving the challenges to society together with industry
‘Another point I would like to highlight is that it is important to invest in science and technology to solve the bigger challenges to humanity. This is sometimes depicted as opposing a healthy business climate. I am convinced that you can address critical societal challenges and make money doing so. Future industry will be to a large extent based on providing solutions to major societal challenges, such as clean energy and the climate and health crises we are facing.

If, as a country, you have a solid pipeline from top fundamental science to top business, you can work on new solutions, have great impact on society while generating employment and boosting your economy. This all starts with ensuring a solid basis. I'm deeply convinced that if we in the Netherlands want to maintain our innovation power, if we want to have a thriving ecosystem including fundamental science we must be willing to provide top talent with top tools such as Exascale High-Performance Computing.'

Quotes:

'It is critical for ASML to work with academics that are used to work with state-of-the-art HPC and that understand our issues'

'If, as a country, you have a solid pipeline from top fundamental science to top business, you can work on new solutions, have great impact on society while generating employment and boosting your economy'

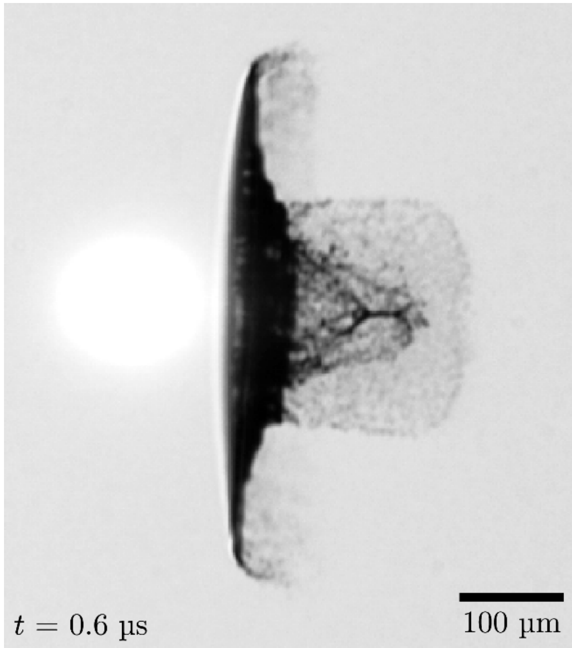


Photo: Randy Meijer

van een gemeten fysisch fenomeen wat een ongekende rekenkracht vraagt mochten we het ooit willen simuleren: impact van een korte laserpuls op een tin druppeltje (relevant voor onze EUV bron)

7.

In molecular dynamics HPC is used to simulate interactions between molecules. Computational modelling on the atomic or molecular level is traditionally a strong field in the Netherlands.
In molecular dynamics HPC is used to simulate interactions between molecules. Computational modelling on the atomic or molecular level is traditionally a strong field in the Netherlands.
Industrial partners are very keen on using this methodology and the way science is connected to various companies is already quite good in the Netherlands.

Recently, the group of prof. Marrink was the first worldwide to simulate a mitochondrion, a sub-part of a cell. They now want to model an entire biological cell with all the molecular detail. That is not yet possible with the current SURF resources but access to Exascale HPC can make this become a reality.

'Computational modelling from the atomic or molecular level is traditionally a strong field in the Netherlands. We now need Exascale HPC infrastructure to compete'

prof. dr. Siewert-Jan Marrink

We want to model an entire biological cell
'We use supercomputers to study molecular dynamics. We firstly capture interactions between molecules in computer simulations to get a fundamental understanding of how complex cellular processes occur at a molecular level. The next step is to make predictions: if you change certain properties of the molecules, you can expect certain related behaviour.

We want to model an entire biological cell with all the molecular detail because that would give the most realistic picture to build on further research. That means billions of molecules interacting in one model. Then we want to predict how that system behaves during timesteps of a few femtoseconds; one millionth of one billionth of a second, which is the intrinsic timescale of molecules. We want to go up to the millisecond and second time scale ranges or even beyond. With the current resources, we are not able of doing that yet. With Exascale HPC, I'm sure we can.'

The first to simulate a mitochondrion worldwide

‘Even though we can’t model an entire cell yet, we have recently simulated a smaller sub-part of a cell, namely a mitochondrion, using the Cartesius HPC instance at SURF. We were the first to do so, worldwide. This part of the cell is responsible for the whole energy household: it's the factory which converts the elements we eat into energy that can be used by other parts of the cell. This organelle is already quite large. It consists of millions of individual molecules and it has a very complex geometry with many folded substructures.

If you have a model of an entire biological cell, you want to probe this over time scales that are comparable to microscopic timescales like seconds, minutes, hours. But for the moment, reaching the millisecond time scale would already be an enormous breakthrough for this kind of systems. That is the holy grail in our field. If you, for instance, want to combat diseases, it is essential that you understand at a molecular level how certain processes occur. This is exactly what we try to unravel with our simulations.’

Screening millions of drugs

‘If you would have really strong computational resources, you could also go to high throughput applications and molecular dynamic simulations. Then you can, for instance, try to design a drug that binds to a certain protein. Typically you start with some drugs that might be effective, and then you play around with them, to see if you find any drugs that are even more strongly binding to a protein.

The potential drug library is almost infinite. You should be able to start screening millions of drugs. There are so many chemical compounds. If you can simulate the binding of millions and millions of potential drugs to your protein targets or to multiple protein targets, in a high throughput fashion, that would mean a huge step forward. And if you can tie this to machine learning approaches, you can go even further.’

We are ready for the next generation of High-Performance Computing

‘Certain groups, especially in the United States, already have a lot of computational power available. They often do things that we also want to do. Unfortunately, we currently can’t, because we don’t have access to that kind of HPC. We are collaborating with some groups in the States. Sometimes we can hook up to their systems, but it is essential for Europe to have competing Exascale HPC infrastructure available.

If Exascale computing becomes available, the methods will be in place to make efficient use of that. We, like other groups, are working on methods to allow these kind of high throughput simulations. We foresee that this will be needed if we have enough HPC power available.

We already started pilot cases where we try to use our simulations to look at protein drug binding in a very efficient way, with the clear prospect of doing this in a high throughput fashion. There is also a lot of software development ongoing to ensure we are ready for the next generation of High-Performance Computing. Not by us, but we know it will become available. This means we are capable of benefiting from the new successful era in HPC.'

Industrial partners

'The type of research that we're doing has very many applications. Industrial partners are very keen on using our methodology. Computational modelling from the atomic or molecular level is traditionally a strong field in the Netherlands. The way science is connected to various companies is already quite good in the Netherlands, so the pipeline from fundamental science to industry is there.

There are many examples of research groups that work with industry. For instance in the construction of synthetic cells. A lot of medicine is being produced by certain engineered cells that can then make certain compounds that are otherwise very hard to make in the lab. If you can control and build synthetic cells, you can then also make them produce any type of product. You can use them as test cells for doing experiments, see how drugs would affect them, for instance.

And if we go back to this high throughput design, this can also be applied to material science. There are so much possible materials that can be constructed using molecular building blocks. Experimentalists are getting better and better at constructing materials on the nanoscale where they can make very detailed layers of different molecules and control how they all fit together. This gives rise to all sorts of materials with very unique properties.

Examples like this show that possibilities are infinite if we can take this to the next level. And it is our experience, that if you start up new things, you will encounter surprises that you have not been able to predict. That is where the interesting science really emerges. This can lead to great opportunities. If we are the first worldwide to make a synthetic cell, that would have an enormous impact, not only scientific but also for industry. For all this, we firstly need access to high level computational power.'

Quotes:

'Sometimes we can hook up to HPC systems in the US via partnerships, but it is essential for Europe to have competing Exascale HPC infrastructure available'

'The way our field of science is connected to various companies is quite good in the Netherlands. The pipeline from fundamental science to industry is there.'

Siewert J. Marrink received his Ph.D. in chemistry (1994) from the University of Groningen, the Netherlands. Following postdoctoral appointments at the Max Planck Institute of Tuebingen (Germany) and Australian National University (Australia), since 2005, he has been full professor at the University of Groningen, heading a research group in molecular dynamics. He is also the director of the Berendsen Center for Multiscale Modeling (University of Groningen) and holds an ERC Advanced grant. His main research interest is on multiscale modeling of (bio)molecular processes, with a focus on unraveling the organization principles of cells. To this end, his group established the popular Martini coarse-grained model.

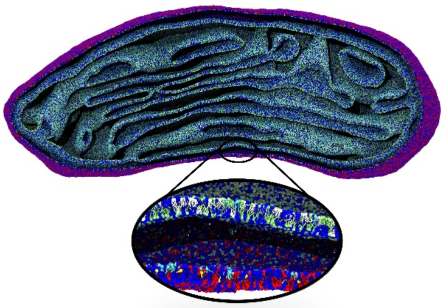


Photo: Simulated Mitochondrion CC BY Pezeshkian, W., König, M., Wassenaar, T.A. *et al*.

8.

If it becomes easier to scale up computations in the life sciences, breakthroughs with a direct impact on people's individual lives will become evident. And the more large and diverse datasets can be combined, the more precise the predictions of the chances of someone becoming ill can get. Therefore, access to Exascale HPC is needed as part of a strong distributed data infrastructure for the life sciences.

Only countries that invest in their data infrastructure make it into the World top 10 in terms of income per capita and quality of life, 20 years from now. So the Netherlands have to ensure their data infrastructure is as world class as their infrastructure of railroads, highways, rivers and airports.

In artificial intelligence, powerhouses such as Google, Facebook and Chinese companies set the standards, since they are processing enormous amounts of data. In the Video and Image Sense lab of the University of Amsterdam, prof. Snoek's team makes sense of video and images with artificial and human intelligence. The lab studies computer vision, deep learning and cognitive science.

Snoek foresees that the field of AI needs to be ready for opportunities that are still unknown, but that will require huge amounts of compute power. Ensuring there is a solid European research data infrastructure available for AI will solve a lot of competitiveness and ethical issues they face nowadays.

'This competition is not about cleverness or ideas. It is all about the ability to perform experiments'

Prof. dr. Cees Snoek

A focus on exploratory research
'The group I'm heading is called the Video and Image Sense lab. We make sense of video and images with artificial and human intelligence. The lab studies computer vision, deep learning and cognitive science. Because of the large amounts of video material we process, High-Performance Computing is a natural element of our research. Next to that, our machine learning and deep learning algorithms require GPUs (Graphics Processing Units) with as much memory as possible.

We were one of the first groups in the Netherlands doing computer vision based research with high performance computing. Very early on we switched from CPU to GPU based computing, even before the deep learning revolution started. At the moment we use GPU computing for everything we do with deep neural networks. That is our main bottleneck when trying to stay in line with our competition, so to say. The current pre-exascale systems and the future exascale systems from EuroHPC feature enormous amounts of GPUs (>10,000), so for us it would be highly beneficial to access this type of systems.

In artificial intelligence you have powerhouses such as Google, Facebook and Chinese companies. They have huge facilities and can attract a lot of research talent. Universities cannot compete with that anymore. Therefore, we focus on exploratory research, the kind that is less likely to succeed, and therefore not of the highest priority for these companies. This still requires huge amounts of computation. We have to report our experimental findings on the data sets that are the standard in the field, and much of the research agenda in our field is determined by researchers that have affiliations at Facebook, Google, Amazon et cetera. They work with enormous data sets.'

You have to be ready for unknown opportunities

'The demand for HPC in our field will only grow these coming years. For instance, the Internet of Things will become omnipresent. More and more visual data will be produced. Think of self-driving cars. They have six cameras which record for hours. If you want to use this data for your research project, it will require a huge amount of computing and storage.

Many things that we will be working on in the coming decade are beyond our imagination now. We have to be ready for new opportunities. Even though these are still unknown, we do foresee that they will require huge amounts of compute power. If we are not prepared, we have to stand by and look at how China and the US walk away with the spin-off of our ideas.

This competition is not about cleverness, nor about ideas. It is all about the ability to perform experiments. For instance, recently a new neural network architecture came to light: Transformer. So now every other day a paper appears online from either Google, Facebook or Amazon with a new variant of this Transformer architecture, simply because they can run experiments at their facilities, with which we cannot compete.

I have to send students to Amazon, because that is the only way for them to do their experiments. Somehow policy makers in Europe think that somebody else will take care of the innovation in ICT, after which the solutions can be bought. But we should not forget to innovate, and to have solid data infrastructure available. For this we also need Exascale HPC.'

Increased interest from industry

'There is a heavy increase in the interest in the algorithms that we develop, also from industry. We do a lot of 'applied projects'. For example, we work with Schiphol Airport which has more than 3000 cameras at their facility. At the moment, these images are all manually monitored and they are used retroactively mostly. If we can help automate the monitoring process, even partially, that footage can be used in a completely different way. We also work with TomTom on high definition, three dimensional maps that will look like a movie. These maps are constructed using data from camera streams, radar streams et cetera.

Those projects are not only about developing a new algorithm, but also about applying an algorithm at a large scale. For this you need the type of HPC power that EuroHPC can deliver.

There is also a lot of societal gain in our projects. For instance, we are involved in a project that uses robotics with cameras to inspect trash underwater. You can use camera footage to monitor climate change, and I'm also involved in a start-up, where we monitor elderly care patients during the night using cameras. If they fall out of the bed, an alert is sent to the nurse.'

Exascale HPC will change our modus operandi

'If Exascale HPC would become available, it would really change our strategy, especially if that would also mean we also get access to large-memory GPUs. Another bottleneck in our projects is transferring data to the computer, because we work with large video datasets so often. So, we also need a good pipeline to the resources, in order to avoid months of transferring data to an instance. If that is the case, we can start to answer more complex questions. Answering these complex problems means we can create more impact for industry and society.

At the moment Europe is spearheading ethics in the use of technology and data, which is a good development. However, you have to offer solid state-of-the-art data infrastructure that is in line with these regulations. Access to EuroHPC will change our modus operandi and open up our research projects. At the moment, the computational demand is often too large to be able to start a research project. That means we are limiting our researchers in developing their intuition and creativity because we cannot provide them with the right compute infrastructure. They try to use free US credits and stuff like that, to be able to perform their experiments. That is not sustainable. The Netherlands joining an European project like EuroHPC, ensuring there is a solid research data infrastructure available for AI, will solve a lot of competitiveness and ethical issues we face nowadays.'

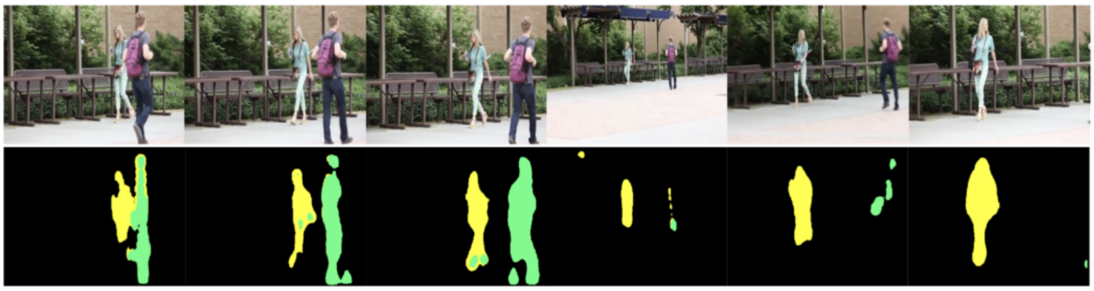
Quotes:

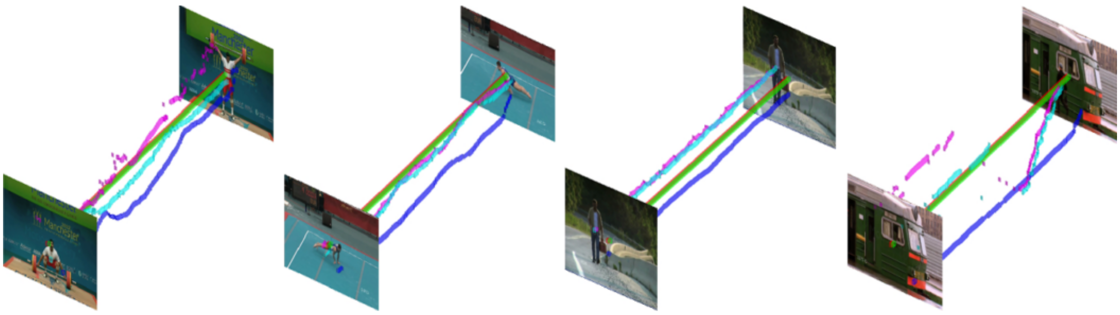
'At the moment we are limiting our researchers in developing their intuition and creativity because we cannot provide them with the right compute infrastructure.'

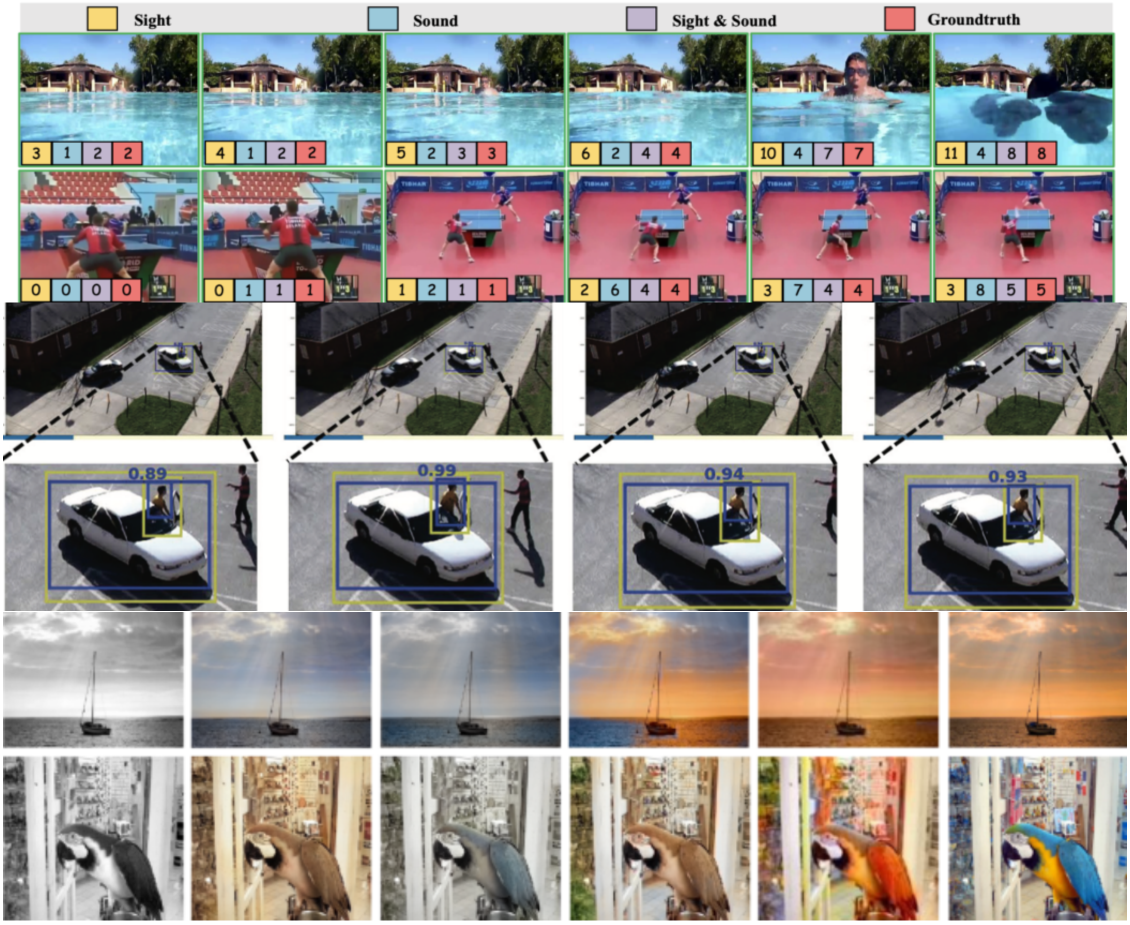
'We work with a company like TomTom on high definition, three dimensional maps that look like a movie'

'If we are not prepared, we have to stand by and look at how China and the US walk away with the spin-off of our ideas.'

Cees Snoek is a full professor in computer science at the University of Amsterdam, where he heads the Video & Image Sense Lab. He is also a director of three public-private AI research labs: QUVA Lab with Qualcomm, Atlas Lab with TomTom and AIM Lab with the Inception Institute of Artificial Intelligence. At University spin-off Kepler Vision Technologies he acts as Chief Scientific Officer. Professor Snoek is also the director of the master program in Artificial Intelligence and co-founder of the Innovation Center for Artificial Intelligence.






Sight
Sound
Sight & Sound
Groundtruth
0.89
0.99
0.94
0.93

9.

If it becomes easier to scale up computations in the life sciences, it will enable research that will yield breakthroughs with a direct impact on people's individual lives. For example, by increasing the size and diversity of datasets that are analyzed with sophisticated AI techniques or modelling tools, the more precise we may be able to predict the chances of someone becoming ill, or being responsive to a certain treatment. Therefore, access to Exascale HPC is needed as part of a strong distributed data infrastructure for the life sciences.

Only countries that invest in their data infrastructure make it into the top countries in the world in terms of income per capita and quality of life, 20 years from now. Consequently, The Netherlands have to ensure their data infrastructure is as world class as their infrastructure of railroads, highways, rivers and airports.

'The life sciences are at the beginning of a data analysis explosion; access to Exascale HPC is needed as part of a strong distributed data infrastructure'

Prof. dr Wiro Niessen

'In the life sciences, we both study fundamental life science questions and conduct applied research, e.g. to develop new diagnostic and prognostic tools. We need a very good distributed data infrastructure including access to Exascale HPC for both.

I'm convinced we are only at the beginning of an explosion in data analytics in the life sciences. If improved data infrastructures enable largers and more diverse datasets to be coupled, and if it becomes even easier to scale up computations, this will result in breakthroughs with a real direct impact on people's individual lives.

In recent years, the life sciences have become more computationally intensive because the information we gather per individual becomes increasingly rich. Think of genomics data, imaging data, clinical data, sensor data, but also data that people collect themselves through wearables. Next to that, we bring more data collections together. This also means that our models, e.g. to characterize health status, or to study a certain disease, become increasingly complex. With these model, we can address more complex questions in the life sciences than ever, which is really exciting. But as you can imagine, our demand for high level HPC is growing accordingly.

In many of the research projects I am involved in, we want to know the chances of a person becoming ill, or the chances that he or she responds to treatment in a certain way. Therefore, we have to understand the relationship between different variables and measurements that describe an individual, and what might happen to that individual. Relevant factors can be genetic liability of disease, someone's lifestyle, and a current set of clinical data. Owing to the advances in the AI field, we model the relation between these factors and relevant clinical outcomes increasingly with neural networks, which will have a large number of variables that you optimize. Optimizing such networks are highly, computationally intensive tasks.'

The benefits of being able to use large combined datasets
'The larger and more diverse the datasets that we use become, the more likely it contains the information that is needed to predict if someone will become ill. Therefore, we need a strong distributed infrastructure: the data we study comes from hospitals, and sometimes even individuals, from all over the world.

My research group has closely worked with researchers of the Rotterdam Study, a long term prospective cohort study in which about 15.000 people from Rotterdam have been followed since around 1990. To optimally exploit the rich imaging and genetic data in this cohort, we already need intensive computing. Next to that we connect with other population studies worldwide, further increasing the size of experiments and the need for Exascale HPC.

Currently, most of the datasets that we use in our research come from western countries. It is important to realize that for a fairer healthcare system, we have to be able to sample representative datasets. If you develop a diagnostic or prognostic tool based on MRI-scans, you have to be aware that most scans that we have access to are from a certain part of the world. If most of the testing of a prognostic tool is done on a non-representative part of the population, the result will be biased. Therefore, we have to increase our efforts to use more diverse datasets and datasets from certain less developed parts of the world. The more different people we include in the study, the richer the data. This is also important for rare diseases. The larger the datasets, the higher the chances you will find patterns you can use to learn about a specific rare condition.'

The direct benefits of precision health

'The benefits of these data driven projects are aplenty. First, we would like to enable a paradigm shift from treatment to prevention, by being able to assess earlier who is at risk. Second, if someone enters the healthcare system, we want to improve their clinical outcome, by selecting the right treatment and also avoiding overtreatment. For instance, in Oncology our aim is to use MR data to predict for which prostate cancer patients watchful waiting is better suited, because the chances of the tumour becoming more aggressive are very slim, and for which patients additional diagnostics is needed such as a biopsy. Similar research is performed in e.g. the cardiovascular disease domain and other disease areas.

This is what we call precision health: to be able to treat an individual person optimally by learning from the many patients that we've seen before. For instance, we process millions of different image scans per year in the Netherlands. We want to learn from all these to treat the next patient better.'

The life sciences need a distributed HPC infrastructure

'You always need to scale your experiments to the smallest set with which you can solve a problem, and only use the highest level of HPC if needed. However, at the moment we are sometimes forced to reduce the problems we want to solve because of the limit to the amount and type of HPC power that is available. That is limiting our research and therefore the impact of our research, which is regretful.

For the life sciences, we need a distributed high performance computing infrastructure, combined with flexibility in the use of EuroHPC instances. Our vision is to have a portal in which we could simply query what kind of data are available in Europe, or globally. For instance, look at risk factors for a severe Covid-19 trajectory. If you see a few cases in your hospital and you want to know if the people with certain characteristics seem to have a higher chance of getting to the IC, you want to check whether these observations fit in a larger pattern. Ideally, you would go to a portal and query open datasets in order to retrospectively identify a cohort and start from there. To enable this, we are working on harmonizing the way in which we collect and process data, in order to combine them according to existing regulations.'

We want to stay relevant as a country

'What we want is better prevention, better diagnostics, better prognostics and better clinical decision support. The methods will be developed in academia, but they will be brought to the patients by industry.

By making a federated, distributed High-Performance Computing infrastructure available for research, you can start to develop public-private models in which companies in the Netherlands can utilize these resources to develop and validate models before introducing them into the market, in a transparent fashion. This will clearly create a lot of economic activity which enables us to go all the way from very fundamental science to solutions for precision medicine for the individual patient.

At the moment we earn a lot of money by transporting stuff from A to B in the Netherlands, due to our excellent infrastructure and expertise. If we want to stay relevant as a country, we need an equally excellent data and computing network as our railroads, highways, airports and rivers. I truly think that only countries that invest in their data infrastructure and expertise make it into the World top 10 20 years from now, in terms of income per capita and quality of life.’

Quotes:
‘We want to improve the clinical outcome of patients, by selecting the right treatment and also by avoiding overtreatment’

‘By increasing the size and diversity of datasets that are analyzed with sophisticated AI techniques or modelling tools, the more precise we may be able to predict the chances of someone becoming ill’

‘You want to know if people with Covid-19 with certain characteristics have a higher chance of going to the IC’

‘If we want to stay relevant as a country, we need an equally excellent data and computing network as our railroads, highways, airports and rivers’

‘Only countries that invest in their data infrastructure make it into the top countries in the world in terms of income per capita and quality of life, 20 years from now’

10.

Deltares, an independent institute for applied research, is one of the larger players in the world in the field of water and subsurface research. They develop detailed models of rivers, subsurface, reservoirs/lakes, coast lines, and coastal areas, for a lot of applications like flood risk, ground water flow, dike stability, morphology, waves, water quality, and ecology. Their simulation software is used worldwide, by a broad range of organizations like national and regional governmental organizations, universities, consultancy/engineering firms and large dredging companies.

A growing trend of interdisciplinary collaboration with different science domains such as meteorology, hydrodynamics, and climate modelling means there is a need for better connectivity and high level HPC in Europe. The EuroHPC consortium can strengthen their work on a broader scale than just compute power only; it can also ensure the Netherlands stay on the forefront in the development of algorithms and software.

'It would be beneficial for our Dutch knowledge economy if we can strengthen our expertise of software and algorithms with Exascale HPC and European collaboration'

Dr Menno Genseberger

'Deltares is an independent institute for applied research in the field of water and subsurface. We are a non-profit organization based in Delft and Utrecht. I have colleagues from 42 different nationalities. Our applied research is on societal issues related to flood risk, water quality, water resources, subsidence and climate adaptation, in areas all over the globe. Our practical knowledge is used in unique projects like Palm Island in Dubai, or the Delta works in the Netherlands.

Already back in the 1970s, we started to create simulation software next to using real physical scale models, which are time-consuming and costly to build. Moreover, not all physical parameters scale in the same manner, so we started to make models to compute how surface water was flowing, how waves are generated by wind, or how sediments are transported along the river.

The interest from society and industry in our simulation software and models is growing, now that sea level rise, drought and salinization are increasing..

We develop different detailed models of rivers, subsurface, reservoirs/lakes, coast lines, and coastal areas. For a lot of applications like flooding, ground water flow, dike stability, morphology, waves, water quality, and ecology.a diversity of functionality is added, like, wind effects, storm surges, impact of waves on dykes, salt intrusion, subsoil and food chain web processes. It's a lot of physics, but also chemical and biological processes. we need the best high performance computing we can have to do this.'

EuroHPC enables us up to scale up our complex models
'For a long time now we are collaborating with universities on research areas such as hydrodynamics, morphology and computational fluid dynamics/numerical mathematics. These sciences are very well developed in the Netherlands. Out of urgency because we have to be sure that our dykes are high and strong enough to stand storm. A safety assessment of dykes can include thousands of scenarios, and still it should be done both accurate and fast. Similar models are used for operational forecasting of floods, there also we really need HPC to have computations finished in time to act.

Nowadays, our models are highly complex. Calculating changes take several weeks of computer time. With a next generation supercomputer, this would only take a few days.
However, getting access to high level HPC is not only important because of the significant gains in speed. The possibility to scale up our models to complexity is important, and so is the possibility to vary more with parameters such as wind strength and directions and to perform impact assessments of different climate scenarios. Generating higher resolutions, adding functionalities, increasing the simulated time course; it all becomes within reach with extra computing power. New opportunities are also to be found in combining several processes that influence each other, such as currents, waves, wind, sediments, and algae.'

Expertise of software and algorithms are Dutch export products
'For applications in surface water, we use similar models as in oceanographic research, but our models are of smaller, local areas. They are more detailed. Our main models are what is called shallow water solvers, which look at the surface water. Additionally there are process formulations and solvers to simulate morphology, water quality, and ecology. These take into account how sediments or nutrients, which are important for the grow of algae or other species, are transported by the flow of water. This flow can be created by the tides, wind, or by discharges from rivers. These simulations can be used to, for instance, model a whole region around a river suffering from flood related problems, drought periods, or to calculate how rivers and coastal areas will evolve over a few years.

We have used these simulations to learn how to improve the situation around the new Marker Wadden islands, to ensure we have the right amount of light, nutrients and water plants or other species to grow. This knowledge development for Marker Wadden has a great international perspective also.

These models, this kind of expertise of software and algorithms as well, are truly Dutch export products. If we could combine this with Exascale HPC and collaboration on the European level, that would open up further possibilities and cooperation. It would be very beneficial to the Dutch knowledge economy.

Interdisciplinary collaboration is a growing trend from which we can benefit hugely.'

EuroHPC fits into a growing trend of interdisciplinary collaboration
'In the near future, we expect to have even more interactions with science domains such as meteorology and climate modelling. These collaborations are all about enabling different models to interact with each other, and about putting as much knowledge into the different models as possible, learning from each other. For this, we have to enforce our models, and we need to be able to exchange large amounts of data. Next to better connectivity we also need Exascale HPC for these interdisciplinary, innovative projects.

Scaling up is not only about computing power. We need to work on a European scale on a lot of different aspects, for instance on enhancing our algorithms and software. At Deltares, our work is open source and therefore our software is used all over Europe. We really believe dare to share is the future. We therefore hope, the EuroHPC consortium will strengthen this international collaboration on a broader scale than just compute power only.

In the Netherlands our software and our algorithms are our strenght, but we can't work on these kind of developments on our own. It would be helpful if we could have a European roadmap for computational fluid dynamics/numerical algorithms and software development to go with access to Exascale HPC, especially since there is sincere competition coming from the US and China.'

Quotes:
'The research areas we are interested in, such as hydrodynamics, morphology and computational fluid dynamics/numerical mathematics are very well developed in the Netherlands'

'The interest from society and industry in our simulation software and models is growing, now that sea level rise, drought and salinization are increasing and felt'

'New modelling challenges are also to be found in combining several processes that influence each other, such as currents, waves, winds, sediments, and algae'

BIO:
Menno Genseberger joined Deltares in 2002. The first years he focused on the development of simulation software for water related processes. Also he gave lead to the discipline Numerical Methods & Techniques to maintain and improve the numerical quality of simulation software. Gradually, also focus came to the application of the simulation software to surface water flow and water quality of real life applications. This to later strengthen the group of people working on coupled hydrodynamical and transport modelling (both flow, waves, and advection diffusion reaction).

For about 10 years Menno is working on the collaboration between Deltares and SURF/SARA. For that he leaded several European PRACE projects with Deltares and SURF/SARA to improve the Deltares simultation software for surface water for High Performance Computing for real life applications.

Before joining Deltares, Menno Genseberger first studied Mathematics (main subject Numerical Mathematics) and Physics at the University of Amsterdam (including courses of the Masters program Meteorology and Physical Oceanography at Utrecht University). Then, after graduation, he started as a PhD student at Utrecht University and CWI (National Research Institute for Mathematics and Computer Science in the Netherlands). Main subject was the design and analysis of a preconditioner based on domain decomposition for the iterative solution of large scale (linear) eigenvalue problems to enable parallel computing on supercomputers for applications from plasmaphysics, astrophysics, oceanography, and geophysics. Research was carried out within a multidisciplinary project Massively Parallel Programming for Computational Magneto-Fluid Dynamics.

Photo: Deltares

6 Over EuroHPC

6.1 Overzicht

EuroHPC is een Europees partnerschap tussen de Europese Commissie, Europese landen en private partners om een wereldklasse supercomputing ecosysteem in Europa te ontwikkelen. Het partnerschap wordt gevormd door de EuroHPC Joint Undertaking onder Luxemburgs recht.

De investeringen in het partnerschap volgen verschillende lijnen[5].

1. Supercomputer en data-infrastructuur van wereldklasse

In het EuroHPC-werkprogramma wordt geïnvesteerd in drie verschillende klassen van systemen:

   a. petascale (5 systemen op moment van schrijven)
   b. pre-exascale (2 systemen op moment van schrijven)
   c. exascale (2 systemen voorzien aanschaf '23 en ’24 / productie ’24 en ‘25)

Daarnaast bevat het werkprogramma investeringen in (hybrid) quantumtechnologie en in een zogenaamd hyperscale-netwerk, dat de verschillende EuroHPC-systemen met elkaar moet verbinden.

2. Research & Innovation

Via Research & Innovation projecten wordt wetenschappelijk onderzoek naar toepassingen gefinancierd. De eerste en tweede ronde calls vereisen 50% matching. Onderdeel hiervan is het European Processor Initiative dat tot doel heeft Europese processortechnologie te ontwikkelen om afhankelijkheid van Amerikaanse en Aziatische technologie te verminderen.

3. Verbreding van gebruik en kennisontwikkeling

Via het EuroCC project worden nationale competentiecentra gevormd die brede kennisontwikkeling, ondersteuning en oplossingen voor wetenschap en bedrijfsleven moeten stimuleren.

6.2 Europese toegang tot EuroHPC-faciliteiten

De investeringen in EuroHPC-systemen kent een cofinancieringsmodel waarbij 50% van de investering wordt gedaan door de EC en 50% door een consortium van lidstaten. De systemen zijn eigendom van de Joint Undertaking. De beschikbare capaciteit wordt verdeeld naar rato van het aandeel van de investeringen.

[5] https://eurohpc-ju.europa.eu/discover-eurohpc-ju

Lidstaten gaan zelf over de allocatie van de aan hen toebedeelde capaciteit.
Het EC-aandeel van de rekentijd op de EuroHPC wordt verdeeld op basis van wetenschappelijke excellentie vergelijkbaar met het PRACE-model. Hierbinnen is echter zeer beperkte capaciteit beschikbaar voor ontwikkeling en benchmarking van applicaties. Daarmee zijn onderzoekers beperkt in hun mogelijkheden wetenschappelijke excellentie en efficiënt gebruik van een exascale-systeem te onderbouwen.
Verder biedt het excellentiemodel geen mogelijkheden om nationale prioriteiten onder te brengen.

6.3 Nederlandse deelname in EuroHPC

Nederland is founding member van de Joint Understaking.

Het ministerie van EZK heeft voor de jaren '21/’22 eenmalig middelen beschikbaar gesteld om aan de matching van R&I projecten tegemoet te komen. Dit geld is vrijgemaakt in reactie op een brandbrief van Nederlandse topwetenschappers. In deze brief stellen zij niet aan de hoge matchingseisen van 50% te kunnen voldoen, waardoor zij achter dreigen te raken op hun collega-onderzoekers uit andere Europese landen, waar de overheid veelal wel een tegemoetkoming voor de matching heeft. Voor de jaren ‘23 en verder zijn er nog geen toezeggingen over een eventuele tegemoetkoming.

Voor investering in EuroHPC-systemen zijn door de Rijksoverheid tot op heden geen middelen vrijgemaakt. Nederland heeft daarmee geen aandeel in één van de pre-exascale systemen die reeds in gebruik zijn genomen.
In het Nederlands voorstel voor de Recovery & Resillience Facility is een post van 10 Mio opgenomen voor mogelijke deelname in één van exascale systemen die zijn voorzien. Een dergelijke investering zou echter pas in ‘24 / ‘25 resulteren in beschikbare capaciteiten.

Binnen EuroCC is SURF de penvoerder van het Nederlandse competentie centrum. De matching hiervoor wordt bekostigd uit de structurele middelen voor kennis en expertise van SURF.

In het European Processor Initiative (EPI) is Nederland via SURF betrokken bij benchmarking van applicaties die voor Nederlandse onderzoekers van belang zijn.

6.4 Invloed van EuroHPC op het landschap

EuroHPC vormt de spil in de Europese onderzoeksprioriteiten die gebaat zijn bij grootschalige rekenkracht en dataverwerking, zoals de digital twins in biomedisch onderzoek, modellering van de aarde en materiaalwetenschappen[6]. Als de onderzoeksdomeinen hun infrastructuur gaan organiseren rondom deze (pre-)exascalesystemen, zullen deze uitgroeien tot zwaartepunten. Voor een goede aansluiting van de Nederlandse pendanten van deze prioriteiten, is het van eminent belang dat Nederlandse faciliteiten onderdeel uitmaken van het ecosysteem dat rondom de EuroHPC-faciliteiten wordt opgebouwd. Dit ecosysteem is een mix van technische voorzieningen en competenties van onderzoekers en ondersteuners.

Nederland is door deelname in EPI, ComBioMed, EuroCC en PRACE op dit moment alleen zijdelings betrokken bij de vorming van dit ecosysteem, maar heeft door het gebrek aan eigen investering geen stem in de ontwikkeling van de EuroHPC-faciliteiten zelf. Daarnaast is bij consortiavorming voor de ontwikkeling van het gefedereerd systeem van Europese HPC-faciliteiten, deelname in één van de (pre-)exascale voorzieningen randvoorwaardelijk. Nederland heeft daarmee nauwelijks invloed op de ontwikkeling van het Europees HPC-landschap. Opschaling van toepassingen en internationale samenwerking komen hiermee in het gedrang.

[6] https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-powered-high-performance-computing-brochure

7 Het Nederlands HPC landschap

Het High Performance Computing ecosysteem volgt een getrapt model van lokaal (tier-2), nationaal (tier-1) en internationaal (tier-0) toegankelijke faciliteiten met een steeds grotere schaal. Naadloze aansluiting tussen deze faciliteiten is cruciaal voor de opschaling van applicaties uit de wetenschapsgebieden.
Naast High Performance Computing, van oudsher gericht op modelberekeningen en simulaties die complexe parallellisatie vragen, bestaat er ook High Troughput Computing gericht op grootschalige verwerking van data, en Cloudcomputing gericht op flexibele en elastische rekenkracht. De grens tussen deze soorten rekenkracht begint te vervagen. In dit document wordt met name gericht op de HPC.

7.1 Nationale rekenfaciliteiten

De nationale supercomputer Snellius (sinds 18 oktober 2021 in productie) vormt de schakel waarmee onderzoekers hun rekenvragen van institutionele voorzieningen kunnen opschalen, of voorbereiden op gebruik van internationale voorzieningen.

De strategie van deze zogenaamde tier-1 voorziening is beschreven in de Lange-termijnstrategie High Performance Computing. De financiering van nationale High Performance Computing faciliteiten vanuit de Rijksoverheid is onderdeel van het Uitvoeringsplan Digitialisering in de wetenschap[7] . Via een convenant tussen universiteiten, UMC's, NWO-instituten en SURF is een aanvullende bijdrage van kennisinstellingen aan reserveringen voor nationale High Performance Computing overeengekomen. Op deze wijze kan over vijf jaar € 30 miljoen aan investeringen gedaan worden om tegemoet te komen aan de toekomstige rekenbehoefte van onderzoekers.

De toegang tot de nationale supercomputer is onderdeel van het financieringsinstrument Rekentijd Nationale Computersystemen van NWO[8]. Via dit instrument wordt op advies van de commissie Wetenschappelijk Gebruik Supercomputer beschikbare rekentijd verdeeld op de nationaal toegankelijke supercomputer (HPC), grootschalige dataverwerkingsfaciliteiten (HTC) en cloud systemen. Naast de rekentijd en dataopslagcapaciteit, kunnen onderzoekers via dit instrument ook expertise aanvragen om hun toepassingen te optimaliseren voor gebruik op grootschalige rekenfaciliteiten.

---

[7] https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2019/10/01/uitvoeringsplan-investeringen-digitale-onderzoeksinfrastructuur/16825243-uitvoeringsplan-investeringen-digitale-onderzoeksinfrastructuur-nwo.pdf
[8] https://www.nwo.nl/onderzoeksprogrammas/rekentijd-nationale-computersystemen

7.2 Lokale rekenfaciliteiten

De rekenfaciliteiten bij kennisinstellingen vormen een brede basis voor de Nederlandse wetenschappelijke rekenbehoefte. Verschillende kennisinstellingen investeren in eigen zogenaamde tier-2 systemen. Voorbeelden daarvan zijn Rijksuniversiteit Groningen, Universiteit Leiden, de Technische Universiteit Delft en Nikhef. De Universiteit van Amsterdam en de Vrije Universiteit werken volgens een model waarin zij investeren in een centraal systeem (LISA) dat wordt beheerd door SURF (voorheen SARA).

In toenemende mate wordt door kennisinstellingen het belang gezien om naast investeringen in faciliteiten, ook te investeren in kennis en expertise die nodig is om een onderzoeker effectief gebruik te laten maken van rekenfaciliteiten. Voorbeelden zijn de Technische Universiteit Eindhoven, Wageningen Universiteit & Research en de Rijksuniversiteit Groningen. Deze expertise is ook cruciaal om onderzoekers voor te bereiden op de stap naar rekensystemen op nationaal niveau.

Nederland kent een langlopende samenwerking tussen Nikhef, Rijksuniversiteit Groningen en SURF voor de grootschalige dataverwerkingsvoorzieningen. In 2020 is er een begin gemaakt om de onderlinge afstemming tussen andere kennisinstellingen te intensiveren, met als doel om kennisuitwisseling, nationale samenwerking en aansluiting tussen nationale en lokale prioriteiten te bevorderen.

7.3 Toegang tot internationale faciliteiten

Nederlandse onderzoekers kunnen via PRACE toegang krijgen tot andere nationale systemen in Europa. De rekentijd wordt op basis van excellentie verdeeld.
Nederlandse onderzoekers zijn historisch gezien succesvol in het meedingen naar rekentijd op de zogenaamde tier-0 faciliteiten in onder meer Duitsland, Frankrijk, Zwitserland, Italië en Spanje. Zij hebben zich daarvoor kunnen voorbereiden op de nationale voorzieningen.

De Nederlandse bijdrage aan PRACE wordt gedaan via SURF in de vorm van een jaarlijkse contributie voor rekentijd op de tier-0 faciliteiten en het beschikbaar stellen van rekentijd op de Nationale Supercomputer voor internationaal gebruik. In het PRACE model vergroot deze deelname de kansen op honorering van voorstellen door Nederlandse onderzoekers[9].

[9] https://prace-ri.eu/wp-content/uploads/Terms_of_Reference_Call24.pdf