



Maarten Spijker, HydroLogic

Frans van Kruiningen, Hoogheemraadschap van Rijnland

Arnold Lobbrecht, HydroLogic en UNESCO-IHE

# High Performance Computing voor waterbeheer

**High Performance Computing (HPC) is nog geen gangbare term in het Nederlandse waterbeheer. Ten onrechte, want deze moderne techniek voor zeer efficiënte simulatie biedt grote mogelijkheden. In dit artikel richten we ons op het snel en adequaat uitvoeren van simulatieberekeningen van het landelijke en stedelijke watersysteem. De toepassingen van HPC hebben betrekking op complexe en tijdsintensieve berekeningen met rekenclusters en snelle computers. Met HPC kunnen we een aantal actuele knelpunten in het strategische en operationele waterbeheer oplossen die thans nog samenhangen met een lange rekentijd. Twee voorbeelden daarvan zijn: gedetailleerde multidimensionale overstromingssimulatie en real-time berekening van de kans op hoogwater. Dit artikel geeft de resultaten weer van onderzoek dat is uitgevoerd op basis van rekenintensieve overstromingssimulaties van de gevolgen van dijkdoorbraken bij het Hoogheemraadschap van Rijnland.**

Modellen spelen een steeds belangrijkere rol bij het beantwoorden van de huidige watervraagstukken. Tot in het recente verleden konden we nog analyses uitvoeren op basis van historische metingen in het watersysteem of op basis van eenvoudige rekenregels. Momenteel is dat bijna niet meer mogelijk door het tempo waarin het watersysteem verandert, evenals de veranderende belasting op dat systeem (klimaat) en de steeds strenger wordende eisen die we aan dat systeem stellen. Het detailniveau waarmee kan, en vaak ook moet worden gemodelleerd, is enorm toegenomen<sup>1)</sup>. Dit leidt tot grote en complexe simulatiemodellen waarbij de tijd, voor het uitvoeren van een berekening ter beantwoording van watervraagstukken, een knelpunt wordt.

In het waterbeheer wordt veel aandacht besteed aan het verkrijgen van correcte invoergegevens en de realisatie van hoogwaardige modellen. Bij een overstromingsberekening worden de beste rekenresultaten bereikt door te werken met een gridgrootte die aansluit bij die van het gecorrigeerde hoogtemodel<sup>2)</sup>, veelal met grids van vijf bij vijf meter. Dat vergt bij grootschalige overstromingssimulaties, bijvoorbeeld als gevolg van een dijkdoorbraak in de Haarlemmermeer, een rekentijd van vele dagen. Dat is in de praktijk veel te lang voor zowel een strategische als een operationele simulatie tijdens een echte dijkdoorbraak. In de praktijk wordt daarom de rekentijd beperkt door te werken met een grover grid, wat een

sterke vereenvoudiging van de werkelijkheid inhoudt.

## Gangbare toepassingen HPC

De hoge verwachtingen van HPC voor het waterbeheer komen voort uit de vele succesvolle toepassingen ervan in andere sectoren. HPC-systemen zijn wijdverbreid in bijvoorbeeld de astronomie, de civiele techniek, de medische wereld en de filmindustrie. Ook bij het voorspellen van tsunami's wordt gebruik gemaakt van een HPC-systeem om in korte tijd een nauwkeurige voorspelling te doen<sup>3)</sup>. HPC-technieken zijn bruikbaar in situaties waar in korte tijd vele, complexe en rekenintensieve simulaties moeten worden uitgevoerd. Hiervoor zijn geen dure super-

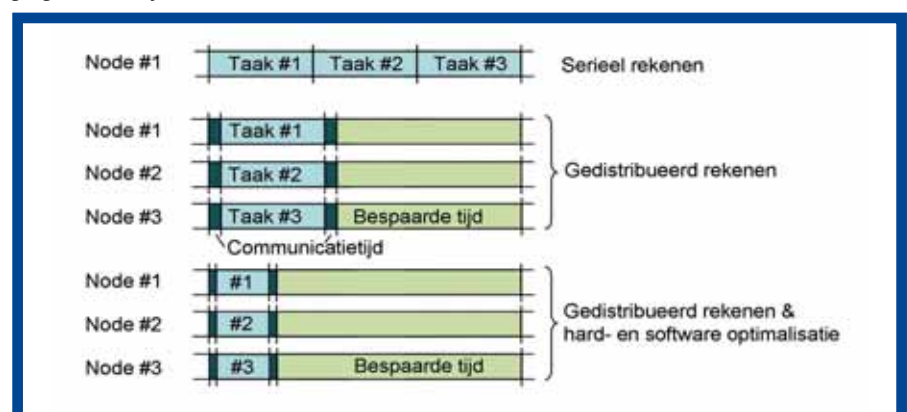
computers meer nodig. Men kan volstaan met standaard kantoornetwerken<sup>4)</sup>. Hiermee krijgt ook de waterbeheerder de mogelijkheid om de vaak te grote rekentijd van complexe simulatiemodellen te verminderen.

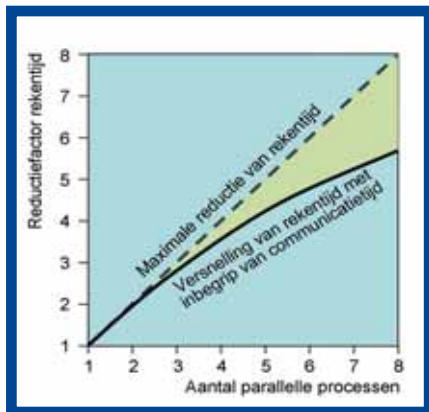
## Waterbeheer

De mogelijkheden van HPC voor het waterbeheer zijn geanalyseerd, waarbij twee aanpakken zijn verkend: gedistribueerd rekenen oftewel het gelijktijdig uitvoeren van een groot aantal berekeningen én de optimalisatie van hard- en software oftewel het in zo kort mogelijke tijd uitvoeren van één berekening.

In afbeelding 1 is aangegeven waartoe HPC-technieken kunnen leiden: het

**Afb. 1: Voorbeeld van de besparing op rekentijd met behulp van HPC-technieken, in vergelijking met de gangbare werkwijze van het serieel rekenen<sup>5)</sup>.**





**Afb. 2: Conceptuele reductie van de rekentijd als functie van het aantal parallelle rekenprocessen.**

aanzienlijk verkorten van de totaal benodigde rekentijd bij individuele en grote aantallen berekeningen. De gangbare aanpak in het waterbeheer is het serieel rekenen, waarbij berekeningen achtereenvolgend worden uitgevoerd. In het geval van een groot aantal berekeningen leidt dit tot een enorme rekentijd en daarmee tot de wens om het aantal uit te voeren berekeningen te beperken. Door de berekeningen over verschillende computers te verdelen, neemt de rekentijd aanmerkelijk af.

Wel is extra communicatietijd nodig om de rekenresultaten van de verschillende computers op een centrale locatie te verzamelen. Hierbij moet worden opgemerkt dat de mate waarin de rekentijd kan worden beperkt ook afhangt van de mate waarin het rekenproces efficiënt over meer subprocessen kan worden verdeeld. Dit is in afbeelding 2 conceptueel aangegeven.

Met de HPC-techniek, hier aangeduid met de optimalisatie van hard- en software, wordt de rekentijd van een individuele berekening verkort. Dit is vooral in operationele situaties zeer relevant, maar biedt ook een toegevoegde waarde bij het uitvoeren van grote aantallen berekeningen. In afbeelding 1 is te zien dat het combineren van beide HPC-technieken leidt tot een aanmerkelijke reductie van de rekentijd.

Beide aanpakken zijn in een pilotstudie onderzocht door gebruik te maken van het rekeninstrumentarium van het calamiteiteninformatiesysteem (CIS), dat HydroLogic voor Rijnland ontwikkelde. Het systeem bestaat uit een via GIS toegankelijke overstromingsbibliotheek van vooraf gemaakte simulatieberekeningen van dijkdoorbraak over het gehele beheergebied van Rijnland. In totaal zijn 400 verschillende dijkdoorbraken berekend en beschikbaar via het CIS. De gedachte daarbij is dat het uitvoeren van een nauwkeurige real-time berekening tijdens de dijkdoorbraak teveel tijd in beslag neemt en dat in die situatie beter kan worden gewerkt met een tweetraps strategie: de actuele situatie analyseren met de best vergelijkbare voorberekende situatie (keuze doorbraaklocatie uit 400) en tegelijkertijd een real-time berekening starten (dijkdoorbraak op de feitelijke locatie, met de werkelijke



**De Ringvaart van de Haarlemmermeer, begrensd door een regionale waterkering. In de pilot zijn onder andere dijkdoorbraken van deze waterkering gesignaleerd.**

bresdimensies en hydrologische randvoorwaarden). Deze laatste berekening moet dan wel vooruit lopen op het werkelijke proces, om nog een in de operationele praktijk bruikbaar resultaat op te kunnen leveren.

### Het principe van gedistribueerd rekenen

Voor het vullen van de overstromingsbibliotheek van het CIS Rijnland zijn 400 tweedimensionale overstromingsberekeningen uitgevoerd (2D-simulaties). Dit vergde een totale rekentijd van circa twee maanden. In de bibliotheek is rekening gehouden met de actuele situatie in het watersysteem. Er is geen rekening gehouden met toekomstige veranderingen, zoals veranderingen in de gegevens van maaiveldhoogte (als gevolg van bodemdaling of nauwkeuriger hoogtemeting) en ook niet met ingrepen in de ruimtelijke ordening die nieuwe doorgangen of barrières vormen in de overstroming (nieuwe wegen, spoorlijnen, viaducten, woonwijken).

Bij een rekentijd van twee maanden is het onpraktisch of zelfs onmogelijk om het CIS bij iedere wijziging in het watersysteem te actualiseren met nieuwe berekeningsresultaten. Deze lange rekenperiode vormt eenvoudigweg een te grote drempel. Daardoor lopen de gegevens in het systeem achter bij de werkelijkheid, wat ongewenst is gezien het belang van de verstrekte informatie.

Bij de geschetste situatie met een rekentijd van twee maanden is gewerkt met een rekgid van 25 x 25 meter. De wens van het hoogheemraadschap is om deze gridgrootte te verkleinen naar vijf bij vijf meter. Dat zou overeenkomen met een verlenging van de rekentijd met een factor 10 à 20. Daarmee zou de totale rekentijd meer dan een jaar bedragen, wat geheel ontoelaatbaar is.

In de pilot is de systematiek van gedistribueerd rekenen beproefd. Hierbij is gebruik

gemaakt van een rekencluster bestaande uit tien parallel geschakelde computers, waarvan er één een coördinerende rol kreeg. Deze zorgt ervoor dat de rekeninspanning slim over de beschikbare computers wordt verdeeld. De totale rekendoorlooptijd kon met dit HPC-principe met een factor 8 worden verkort (van twee maanden naar één week).

Dit resultaat stemt overeen met een vergelijkbaar onderzoek dat recentelijk is uitgevoerd door Barreto van UNESCO-IHE<sup>4</sup>. In dat onderzoek is met HPC-technieken een stedelijk watersysteem met het simulatiepakket MOUSE geoptimaliseerd op basis van een multicriteriabepaling. HPC leidde met vier parallelle computers tot een reductie van de rekentijd met meer dan een factor 3. Hiermee kregen in zijn voorbeeld de beleidsmakers en belanghebbenden de mogelijkheid tot het simuleren van veel meer varianten dan voorheen.

### Watertoepassingen van gedistribueerd rekenen

Een mogelijke toepassing van HPC bij Rijnland is de inzet van een groot deel van de aanwezige computers die bij het hoogheemraadschap zijn opgesteld. Deze worden gemiddeld nog geen 20 procent van de tijd gebruikt (denk aan nachten en weekeinden). Door deze computers via het kantoor netwerk parallel te laten rekenen, ontstaat dezelfde situatie als met de computers in een cluster. Een voorzichtige schatting geeft aan dat meer dan 100 computers kunnen worden gebruikt die voor 80 procent van de tijd beschikbaar zijn voor het uitvoeren van een breed scala van berekeningen, waaronder overstromingsberekeningen voor het CIS. Hierdoor kan de geconstateerde drempel voor het actualiseren van de gegevens in het CIS worden weggenomen. Alle 400 berekeningen kunnen dan in één nacht worden uitgevoerd.

Naast het beschreven calamiteiteninformatiesysteem kunnen we ook aan vele andere

toepassingen van het principe van *veel rekenen* denken in het waterbeheer:

- ensemble hoogwaterverwachtingen door het met een operationeel beslissings-ondersteunend systeem doorrekenen van 50 parallele weerscenario's van het Europese weerinstituut ECMWF<sup>(5),(6),(7)</sup>;
- real-time optimalisatie van de beschikbare bergings- en transportcapaciteit van landelijke en stedelijke watersystemen, ter voorkoming van overstroming en overstort uit rioolstelsels<sup>(4),(8)</sup>;
- nauwkeuriger stochastisch rekenen met meer variabelen en fijnere discretisaties, bijvoorbeeld voor het toetsen van watersystemen aan de geldende normen of voor het vaststellen van de optimale norm;
- het vaker actualiseren van toetsingen van watersystemen conform de WB21-systematiek<sup>(9)</sup>;
- het optimaliseren van KRW-maatregelenpakketten via het parallel doorrekenen van lange perioden met hydrodynamische waterkwaliteitsmodellen. Voor de KRW zijn lange berekeningen vereist om de effecten op waterkwaliteit en ecologie te bepalen. Deze lengte (tien tot 30 jaar) werkt beperkend in het praktische aantal door te rekenen maatregelen;
- het efficiënt en automatisch kalibreren van modellen, wat een zeer rekenintensief iteratieproces<sup>(10)</sup> is. Op basis van ervaring met het nauwkeurig simuleren van het gedrag van landelijke en stedelijke watersystemen kunnen we stellen dat de nauwkeurigheid van het model sterk samenhangt met de mate van precisie waarmee dit iteratieve proces is doorlopen.

## Optimalisatie hard- en software

De optimalisatie van hard- en software heeft tot doel om een enkele berekening in zo kort mogelijke tijd uit te voeren. Dat is vooral van belang voor operationele toepassingen, waarbij de rekentijd cruciaal is voor een adequate reactie. Als voorbeeld nemen we het actuele knelpunt van het real-time uitvoeren van overstromingsberekeningen met het calamiteiteninformatiesysteem van Rijnland. Hierbij moet de actuele situatie worden nagebootst om de precieze effecten van een dijkdoorbraak beschikbaar te hebben, geruime tijd voordat deze in werkelijkheid optreden. Alleen in dat geval kan de waterbeheerder de ingelanden op tijd waarschuwen en het juiste protocol inzetten voor het beperken van extreme schade en het voorkomen van slachtoffers. Nu is de rekentijd van een enkele overstromingsberekening nog te groot om deze berekening met de juiste nauwkeurigheid

real-time te kunnen uitvoeren. De voorspelende waarde van het huidige berekeningsresultaat is daarmee beperkt en het doorrekenen van het effect van operationele maatregelen in een calamiteuze situatie is niet goed mogelijk.

In grote lijn geldt in de huidige situatie bij een overstromingsberekening in het regionale systeem, met regionale waterkeringen, voor de rekentijd op een modern werkstation (medio 2008) het volgende<sup>(11)</sup>:

- dijkdoorbraakberekening met grid van vijf bij vijf meter: tijdsverloop in modelsimulatie komt overeen met het werkelijke tijdsverloop (circa een dag per gesimuleerde dag),
- dijkdoorbraakberekening met grid van 25 bij 25 meter: tijdsverloop in modelsimulatie circa 20 maal sneller dan het werkelijke tijdsverloop (circa een uur per gesimuleerde dag).

Het bovenstaande noodzaakte het hoogheemraadschap om te kiezen voor berekeningen met het minder nauwkeurige grid van 25 bij 25 meter. Voor het behalen van werkbare rekentijden bij een dijkdoorbraak van een primaire kering is opschaling naar een nog grover grid nodig, in verband met het grotere geïnundeerde areaal.

In de pilotstudie is onderzocht wat bepalend is voor de rekentijd bij een representatieve dijkdoorbraaksimulatie met het CIS Rijnland. Hieruit bleek dat voor het realiseren van een zo kort mogelijke rekentijd drie punten bepalend zijn: de wijze van modelleren, de gebruikte hardware en de toegepaste simulatiesoftware.

Het eerste betreft de wijze waarop de dijkdoorbraak is geschematiseerd en de hierbij gehanteerde modelinstellingen. Optimalisatie van de modelschematisatie leidde tot een verkorting van de rekentijd met factor 2 à 5, met behoud van simulatienauwkeurigheid. In dit voorbeeld is het vooral van belang om in de buurt van de dijkdoorbraak het model zo stabiel mogelijk te laten rekenen door het knopen/takken systeem hierop te verfijnen. Tevens kan rekentijd worden bespaard door de iteratieinstellingen van het instrumentarium af te stemmen op het meest kritische gedeelte van het rekenproces.

Als tweede is onderzocht of extra winst kan worden behaald door inzet van optimale hardware. In de pilotstudie is dit aspect onderzocht voor het veel toegepaste simulatieprogramma Sobek. Uit een beperkt aantal hardwaretests bleek een winst te behalen te

zijn met een factor 1.5 à 2. Bepalend hierin bleek niet de kloksnelheid van de processor te zijn, maar de architectuur ervan en met name de bandbreedte van het geheugen. Het laatste punt in de vergroting van de rekensnelheid is de optimalisatie van de simulatiesoftware zelf. De verwachting is dat een behoorlijke tijdswinst is te behalen door een goede afstemming van de software op de huidige generatie processoren.

De huidige simulatiesoftware dient zodanig te worden geoptimaliseerd dat deze zelf het rekenproces verdeelt, zodat efficiënt gebruik wordt gemaakt van de huidige generatie processoren. De verwachte snelheidswinst is in dezelfde orde van grootte als die welke wordt behaald met het inzetten van een cluster van parallel rekenende computers (zie afbeelding 2). Deze verwachting wordt ondersteund door de resultaten van een door het Danish Hydraulic Institute uitgevoerde test met simulatiesoftware. Hierbij reduceerde toepassing voor een 2D-simulatiemodel op een 8-core processor in vergelijking met een 1-core processor de rekentijd van het onderzochte simulatiemodel met een factor 8<sup>(12)</sup>.

## Conclusies en aanbevelingen

De beschreven experimenten tonen aan dat HPC grote mogelijkheden biedt voor het sterk versnellen van rekenintensieve processen in het waterbeheer (zie de tabel). De gunstige resultaten worden behaald door optimalisatie van de modelschematisatie en een optimale inzet van de hardware en simulatiesoftware. Bij deze laatste versnelling wordt effectief gebruik gemaakt van moderne multi-core processors.

Belangrijke knelpunten in de ontwerp-, analyse- en beheerfasen van watersystemen kunnen met toepassing van HPC worden opgelost, zoals de veel te lange rekentijd die op het moment nog gepaard gaat met het uitvoeren van 2D-overstromingsberekeningen.

Het real time doorrekenen van overstromingen heeft een enorme toegevoegde waarde voor het sneller en effectiever bestrijden van overstromingen en de gevolgen van die overstromingen. Dit zal tot uiting komen in het voorkomen van slachtoffers, het verminderen van de maatschappelijke impact en het beperken van schade. HPC biedt uitkomst bij deze rekenintensieve en beslissingsondersteunende taak.

Anno 2008 beperkt de rekentijd van simulatiemodellen de waterbeheerder nog te veel in de te maken keuzen. Met HPC kunnen deze beperkingen worden opgeheven, waardoor minder vaak vereenvoudigde aannamen hoeven te worden gedaan, complexe maatregelenpakketten beter kunnen worden geanalyseerd en onderbouwd, effectief probabilistisch rekenen binnen handbereik komt én hoogwaterverwachtingen nauwkeuriger kunnen worden gemaakt.

Van alle mogelijkheden die er zijn om simulatietijden te beperken, is het ombouwen van simulatiesoftware naar

**Reductiefactoren voor de totale rekentijd ten opzichte van serieel rekenen (huidige praktijk).**

casus	gedistribueerd	optimalisatie hardware en software			maximale reductiefactor
	rekenen	model	hardware	software	
1 CIS-simulatie	1	2-5	1.5-2	≤ 8	80
400 CIS-simulaties op 100 PC's	~ 100*	2-5	1.5-2	≤ 8	8.000
Barreto met 4 PC's <sup>(5)</sup>	3	-	-	-	
DHI Mike-casus <sup>(13)</sup>	-	-	-	~ 8	

\* Bepaald op basis van extrapolatie van een experiment met tien parallele computers.

multi-core software het meest ingrijpend. Evenwel is een deel van de stappen naar de snelst denkbare simulatieberekening relatief snel gezet en binnen handbereik van de waterbeheerder. Specifiek denken we aan gedistribueerd rekenen, modeloptimalisatie en toepassing van de juiste hardware. Met deze stappen kunnen al gauw snelheidswinsten van een factor 10 worden behaald voor de situatie zonder gedistribueerd rekenen, tot orde grootte 1000 voor grote aantallen simulaties met gedistribueerd rekenen (zie de tabel). Met de inzet van deze vormen van HPC kunnen de Nederlandse waterbeheerders zich op korte termijn aanzienlijk beter wapenen tegen de risico's die nu nog samenhangen met extreem lange rekentijden van simulatiesoftware voor hoogwater- en overstromingsberekening.

#### LITERATUUR

- 1) Spijker M. en A. Lobbrecht (2006). Naar een efficiënt integraal waterbeheer met 'het nieuwe modelleren'. H<sub>2</sub>O nr. 14/15. pag. 24-25.
- 2) Zantvoort M., F. van Kruiningen, N. ten Heggeler en M. Spijker (2008). 2D-modelleren waardevol voor regionaal waterbeheer. H<sub>2</sub>O nr. 13, pag. 41-44.
- 3) Hendry A. (2008). Debian Linux cluster beats supercomputer in tsunami warnings - Philippine weather service finds forecasting cheaper with Linux. PCworld.
- 4) Barreto W., Z. Vojinovic, R. Price en D. Solomatine (2008). Multi-tier modelling of urban drainage systems: Multi-objective optimization and parallel computing. In: Proceedings of the 11th International Conference on Urban Drainage, Edinburgh. In druk.
- 5) Lobbrecht A., G. Hiemstra, M. Talsma en Z. Vonk (2003). Neerslaginformatie voor het waterbeheer. H<sub>2</sub>O nr. 23, pag. 22-25.
- 6) Van Andel S. en A. Lobbrecht (2006). Ensemble weather forecasts and operational management of regional water systems. 7th International Conference on Hydroinformatics (ed. by P. Gourbesville, J. Cunge, V. Guinot & S. Liong). Research Publishing Services, pag. 1351-1358.
- 7) Loos S., A. Lobbrecht, H. van Norel en J. Esenkbrink (2008). Nieuwe beslissingsondersteunende systemen voor het operationele waterbeheer. H<sub>2</sub>O nr. 13, pag. 45-48.
- 8) Lobbrecht A. (1997). Dynamic water-system control - design and operation of regional water-resources systems. A.A. Balkema Publishers.
- 9) Reichard H., M. Spijker, G. Tromp en B. Pijpers (2006). Eerste landelijke toetsingskaart NBW-werknormen. H<sub>2</sub>O nr. 16, pag. 6-7.
- 10) Solomatine D. (1999). Random search methods in model calibration and pipe network design. In: Water Industry Systems: modelling and optimization applications (ed by D. Savic & G. Walters). Research Studies Press, pag. 317-332.
- 11) De Graaf J., M. Zantvoort en M. Spijker (2007). CIS Rijnland achtergronddocument modellering. Hoogheemraadschap van Rijnland.
- 12) DHI (2008). Simulation time now significantly reduced for MIKE 3/21. The NetWork - a community newsletter for MIKE users.