

COM 01 Databeheer van BSIK-KvR-projecten

Projectleider	dr. Eric Boom		
Instituut	Dutch Space B.V.		
Email	e.boom@dutchspace.nl		
Consortium	Dutch Space B.V.		
Project website			
Startdatum	voorjaar 2006	Einddatum	voorjaar 2007 (fase 1) voorjaar 2008 (fase 2)

Context / maatschappelijk probleem

Binnen Klimaat voor Ruimte worden in veel projecten grote hoeveelheden data geproduceerd. Deze data zijn niet alleen van belang voor deze projecten zelf, maar zijn vooral waardevol als ze als (data)product ook voor alle potentiële gebruikers (andere Klimaat voor Ruimte projecten, maar ook partijen daarbuiten) van deze data beschikbaar zijn. De KvR-projecten zelf hebben in veel gevallen niet de expertise en/of middelen beschikbaar om de data voldoende laagdrempelig toegankelijk te maken. De prioriteit van de projecten richt zich typisch op het verwerven van de data en deze te vertalen in kennis, en minder op het voor anderen toegankelijk maken hiervan. Een vergelijkbaar probleem speelt op het gebied van (software)gereedschappen om de data te bewerken (zoals modellen, conversie van formaten, visualisatie) – ook hier kan uitwisseling en hergebruik van bestaande middelen een waardevolle bijdrage leveren aan samenwerking tussen de projecten binnen (en buiten) het BSIK KvR-programma.

Wat is al bekend, wat niet?

Sectoren als ruimtevaart, genomics, astrofysica of kernfysica hebben vaak te maken met enorme datastromen en moeten voor de verwerking daarvan over zeer zware reken capaciteit en geëigende gereedschappen kunnen beschikken. Om daar effectief mee om te kunnen gaan, zou men het liefst van elkaars kennis en reken capaciteit gebruik kunnen maken, maar in de praktijk stuit dat op allerlei technische en organisatorische problemen. Wereldwijd wordt daarom aan zogenaamde Grid-technologie en Service-Oriented Architectures (SOA) gewerkt, waarmee zonder hindernissen computersystemen over de hele wereld op transparante en gebruikersvriendelijke wijze aan elkaar kunnen worden gekoppeld – en dit zonder dat er ingewikkelde en noodgedwongen uitwisselingen van algoritmen, modellen en/of broncodes moeten plaatsvin-

den. Binnen Klimaat voor Ruimte wordt gebruik gemaakt van een aantal zeer data-intensieve deelprogramma's waar deze technologie erg belangrijk voor kan zijn. Het is echter afhankelijk van de specifieke karakteristieken van de projecten (zowel als producent als consument van data) in welke vorm bovengenoemde moderne informatietechnologie het beste kan worden ingezet.

Wat wordt nu onderzocht?

Het concept van een 'virtual data centre' vormt de kern. Doel is een centrale toegangspoort te creëren voor projectleiders en externe gebruikers tot (geconsolideerde) dataproducten uit geselecteerde projecten binnen adaptatie, mitigatie en klimaatscenario's. Er wordt een infrastructuur ontwikkeld waarbij alle her en der verspreide databronnen en computerplatforms tot een soort samenwerkingsverband aan elkaar worden gekoppeld. Daardoor wordt het mogelijk data uit te wisselen en gebruik te maken van elkaars kennis (en indien gewenst, van elkaars computersystemen en gereedschappen). Tegelijkertijd blijft men als project toch zelf controle over de eigen, specifieke algoritmen en dataverzamelingen houden. De eerste fase start met inventarisatie van gebruikerseisen. Daarna begint de ontwikkeling en de bouw van het prototype voor de centrale portal waarbinnen partijen kunnen communiceren.

Wat is het resultaat?

Het idee is om binnen Klimaat voor Ruimte het virtual data centre op te tuigen als een portal. Via deze portal kan kennis tussen de deelnemers aan Klimaat voor Ruimte worden uitgewisseld, maar kan ook gezamenlijk gebruik worden gemaakt van kennis en kunde. Kennisuitwisseling dus, en dat op een zeer gedistribueerde basis; intensief en over de volle breedte van het programma. Het COM01-project zal zich hierbij vooral concentreren op integratie van de bij de verschillende partners aanwezige ICT infrastructuur om zo op efficiënte en effectieve wijze de ge-



wenste integratie en communicatie van kennis, informatie en diensten tot stand te brengen. De verwachting is dat het systeem eind 2007 operationeel zal zijn. Intussen blijkt er uit de hoek van olie-exploratie en genomics belangstelling te bestaan voor de (Grid-)technologie zoals die mede binnen BSIK wordt ontwikkeld.

