

VAN DER LANS



# Operational BI

De meningen over de term Operational BI zijn verdeeld. De een bedoelt ermee dat BI-applicaties worden ontwikkeld voor diegenen die zich bevinden op de operationele laag van de organisatie, dus dicht bij de bedrijfsprocessen, terwijl voor een ander het betekent dat analisten toegang krijgen tot zeer actuele gegevens, de zogenaamde operationele gegevens. En naast deze twee interpretaties zijn er nog andere in omloop. In deze column gaan we uit van de tweede betekenis.

Als we het over het implementeren van Operational BI hebben, dan fronsen we allemaal ons voorhoofd, want we weten allemaal dat dit niet eenvoudig gaat worden. Ons datawarehouse wordt slechts één keer per dag ververst, onze datamarts om de dag, dus als we die gaan benaderen, krijgen we gegevens te zien die één of twee dagen verouderd zijn. En dat is niet wat we bedoelen met Operational BI. We weten dat we ons in allerlei technische bochten moeten wringen om toch BI-rapporten te creëren die actuele gegevens tonen. We weten dat dit een technisch hoogstandje gaat worden.

We kunnen uiteraard proberen onze BI-applicaties direct op de productiedatabases te draaien. Maar nagenoeg iedereen zal ons voor gek verklaren. Er zijn dan minimaal twee potentiële problemen. Ten eerste, de query's die dan afgevuurd worden zullen de workload op de productiedatabase danig verstoren. We kunnen bijvoorbeeld problemen met locking krijgen, of de query consumeert zoveel resources dat de applicaties van andere gebruikers gewoonweg niet afgehandeld worden. En ten tweede, als we in onze rapporten oudere gegevens willen combineren met operationele gegevens, kunnen we het wel helemaal vergeten, want de productiedatabases bevatten geen historie, alleen het datawarehouse bevat deze. Hierbij doen we wel een aanname, namelijk dat de productiedatabase geen historische gegevens bijhoudt. Maar wat nu als deze dat wel zou doen? Wat nu als er een ontwerp bedacht is waardoor wijzigingen wel degelijk vastgehouden worden? In dat geval valt het tweede probleemgebied weg en blijft alleen die verstoring over.

Kort geleden ben ik in contact gekomen met twee organisaties die inderdaad in productiedatabases historie bijhouden. Beide organisaties behoren niet tot de duizend grootste organisaties ter wereld, maar ze zijn zeker niet klein. Het zijn beide middelgrote, internationaal opererende organisaties, met een productiedatabase van ongeveer 1 Terabyte. Voor beide is door het in Nieuwegein gevestigde Aorta een BI-omgeving opgetuigd die direct de productiedatabase benadert. De ontwerpers van deze productiedatabases hebben deze zodanig ontwikkeld dat histori-

sche gegevens niet verloren gaan. Een van deze twee organisaties houdt al historische gegevens in de productiedatabase bij vanaf 1985. Dus voor de duidelijkheid, deze organisaties hebben geen datawarehouse, geen datamarts, geen ETL, niets. Alle gegevens liggen één keer opgeslagen en wel in de productiedatabases; eigenlijk zoals het hoort, zoals het in de schoolboekjes staat. Wat een eenvoud en wat een visie hadden deze ontwerpers jaren geleden!

Kortom, beide organisaties ervaren het eerstgenoemde probleem niet als een probleem. Blijft het tweede probleem over: performance verstoring. De ene organisatie meldt daar geen problemen mee te hebben. Bij de ander lijkt het alsof een bepaald rapport wat langzaam begint te draaien. De ontwerpers van Aorta hadden bij het begin al bepaald dat ze Oracle BI Server zouden inzetten om de rapportage los te koppelen van de werkelijke gegevensopslag. In feite is wat zij gedaan hebben in lijn met de ideeën van het Data Delivery Platform waarover ik recentelijk gepubliceerd heb. Als ze willen, kunnen ze een datamart, kubus, of datawarehouse toevoegen zonder dat de rapporten aangepast moeten worden. Het is puur een performance versus kosten overweging. Misschien kunnen ze het probleem ook oplossen door een snellere server aan te schaffen. Dit is een mooi voorbeeld waaruit blijkt dat het Data Delivery Platform als architectuur in de praktijk werkt: ontkoppel de opslagstructuur van de rapporten. Maar het mooiste was wel dat toen aan de klant gevraagd werd of ze blij waren met hun moderne Operational BI-omgeving, zij vroegen wat daar precies mee bedoeld werd. Zij zagen het speciale van Operational BI niet in. Het waren 'gewoon' hun rapporten. Beide organisaties hebben per ongeluk Operational BI in gebruik.

Wat we hier zeker van kunnen leren is dat we in onze cursussen over database-ontwerp, en dan bedoel ik het ontwerp van productiedatabases, moeten doceren dat historie altijd bijgehouden moet worden; gooi geen gegevens weg. Misschien zijn die historische gegevens initieel niet van belang, maar ze later toevoegen is een hels karwei. We voorkomen hiermee dat we alsnog complexe datawarehouses moeten ontwikkelen of andere constructies moeten optuigen om de gewenste rapporten te kunnen ontwikkelen.

In de literatuur doet men erg moeilijk over Operational BI. Operational BI hoeft niet moeilijk te zijn, mits je maar de correcte architectuur voor de datawarehouse-omgeving opzet.

**Rick van der Lans** is zelfstandig IT-consultant.