

CWI en Databricks gaan samenwerken

Geplaatst op 25 april 2017 om 20:11 uur



Het Amsterdamse Centrum voor Wiskunde en Informatica en het big data analyse- en data science softwarebedrijf Databricks hebben een nieuwe samenwerking bekendgemaakt. Dat gebeurde op de Hannover Messe, in aanwezigheid van Minister Kamp van Economische Zaken.

Speciaal voor de samenwerking met het CWI, en voor het database engineering-talent ter plekke, opende Databricks eerder dit jaar een R&D-centrum in Amsterdam. Het bedrijf uit San Francisco leidt de ontwikkeling van de open-source software Apache Spark, wereldwijd de meest gebruikte softwaretool om grote hoeveelheden data te analyseren. Databricks biedt Spark aan als service in de cloud: met Spark kunnen bedrijven en organisaties efficiënt en effectief hun data analyseren in een beheerde omgeving.

Databricks financiert onderzoek in de Database Architectures-groep van CWI, de onderzoeksgroep die eerder de bekende databasesystemen VectorWise en MonetDB ontwikkelde. De CWI-onderzoekers werken aan databasetechnieken die voor Databricks interessant kunnen zijn. Het gaat om methoden waarmee gebruikers grote hoeveelheden data kunnen analyseren, waarbij het niet alleen om tabellen hoeft te gaan, maar ook om (sociale) netwerken met groeiende en veranderlijke gegevens.

Verrijking van het data science-ecosysteem

Peter Boncz, senior researcher in de Database Architectures-onderzoeksgroep, coördineert de samenwerking. Boncz: "De komst van Databricks naar Amsterdam is een verrijking van het lokale data science-ecosysteem en het onderstreept de reputatie van het CWI op het gebied van big data-technologie".

Boncz ziet duidelijke kansen voor het fundamentele CWI-onderzoek: "Voor CWI levert de samenwerking de kans op om achter de schermen mee te kijken bij Databricks. We krijgen inzicht in de grote diversiteit aan data-analyse-problemen waarmee gebruikers worstelen, en ontdekken zo de open vragen op het gebied van data-analyse. Daarnaast zijn bij Databricks professoren van Berkeley en Stanford betrokken, dus er ontstaan ook nieuwe wetenschappelijke samenwerkingen."

Sneller en meer schaalbaar

"Databricks kijkt ernaar uit onze R&D-aanwezigheid in Amsterdam te versterken, en baanbrekend werk te verrichten op het gebied van big data-analyse en data science", zegt Ram Sriharsha, interim vestigingsmanager van het nieuwe Amsterdamse R&D-center van Databricks. "De samenwerking met het CWI, en vooral de expertise van het CWI op het gebied van high-performance databases, was doorslaggevend bij onze keuze voor Amsterdam. We hopen dat we dankzij deze samenwerking Spark en Databricks sneller en schaalbaarder kunnen maken."