

De Polis Papers: Size Matters

Niet veel mensen weten het nog, maar sinds 2006 wordt voor iedereen die in Nederland werkt, een uitkering, pensioen of lijfrente heeft, het inkomen en een heleboel daarmee samenhangende gegevens geregistreerd. Iedere maand doen dik 500.000 bedrijven en instellingen loonaangifte in de vorm van ruim 70 gigabytes aan XML bij de Belastingdienst. De Belastingdienst stuurt die deze ruim 20 miljoen inkomstenverhoudingen voor 14 miljoen mensen door naar UWV, dat deze gegevens aanvult met gegevens over bedrijven en personen (o.a. uit de GBA) en opslaat in één enorme database, de Polisadministratie. De Polisadministratie database is fysiek groot – nu ruim één terabyte – maar moet tegelijk ook een enorme stroom inkomende mutaties verwerken en een explosief groeiende stroom gegevensleveringen verzorgen. Inclusief inkomende en uitgaande stromen praten we te zijner tijd over een fors aantal terabytes per jaar. De Polisadministratie staat bij UWV omdat het de bedoeling is dat UWV met een druk op de knop, op basis van de inkomensgegevens, een uitkering kan berekenen; maar ondertussen verandert de Polisadministratie het Nederlandse informatielandschap ingrijpend. Het CBS hoeft geen loonenquêtes meer te houden. De belastingdienst kan zonder rompslomp allerlei toeslagen berekenen. Pensioenfondsen kunnen premies en pensioenrechten vaststellen. Zorgverzekeraars en deurwaarders kunnen gericht loonbeslag leggen. En uzelf kunt straks kijken of uw inkomensgegevens netjes zijn geregistreerd en hoe hoog uw uitkering wordt bij werkloosheid of verminderde arbeidsgeschiktheid.

Een multi-terabyte terabyte database waarin maandelijks honderden gigabytes instromen en terabytes uitstromen vereist een *state of the art* systeem. UWV had zo'n systeem laten bouwen, maar *state of the art* bleek niet genoeg te zijn. Behalve onvoldoende performance was het struikelblok dat een groot deel van de opgespaarde loonaangifteberichten niet kon worden verwerkt omdat ze niet aan de specificaties voldeden. Voorjaar 2007 liep het daarom mis: de loonaangifteketen bleek niet te werken. Belastingdienst en UWV stelden een Integrale Probleemanalyse op en structurele oplossingen voor de toekomst. Drie fasen werden geschetst; fasen waarin de loonaangifteketen werkend, stabiel en robuust gemaakt moet worden. De situatie werd goeddeels gered door een in drie maanden gebouwde noodvoorziening in de vorm van een berichtendatabase waarin alle loonaangiften zonder slimigheden waren opgestapeld en waar vanuit de het CBS, de Belastingdienst en een aantal andere belanghebbenden konden worden bediend. De eerste versie van die nieuwe Polisadministratie is gebouwd tussen oktober 2007 en juni 2008 en is sinds maart 2008 operationeel. In die korte periode is door een klein ontwikkelteam een systeem neergezet dat noodgedwongen een aantal voorzieningen bevat die min of meer nieuw zijn in de wereld van administratieve systemen. In deze technische artikelen willen we een aantal van die innovaties voor het voetlicht brengen en zodoende de *state of the art* in databaseland een stukje vooruit helpen. Veel van de grote logische en technische problemen rond de Polisadministratie komen op kleinere schaal elders ook voor. Waar dat het geval is hopen we met deze serie tenminste te kunnen bijdragen aan het vermijden van kostbare valkuilen.