

Standaard database-structuur om over kostenbeheersing te bevragen

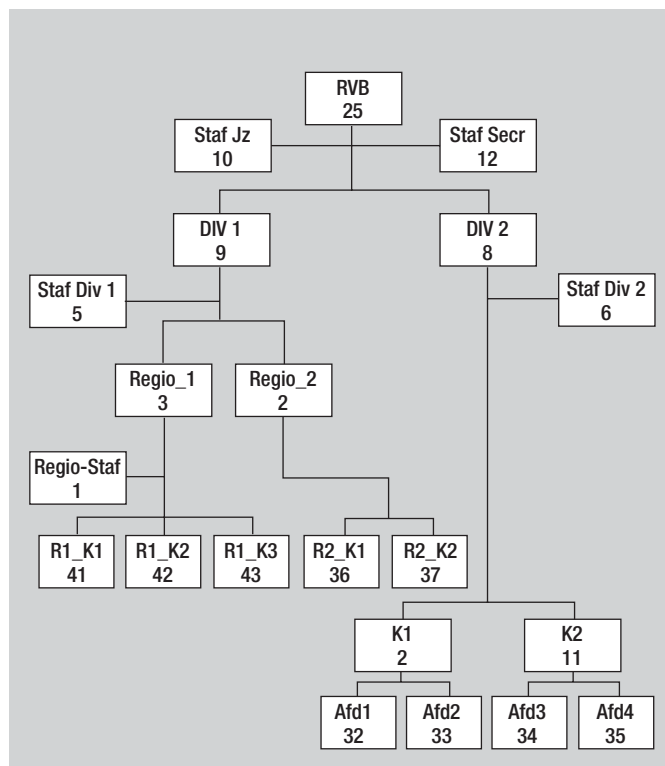
Langs de paden van het organisatorische bos

Paul Hulst en Fons Rooijers

Een eindgebruiker moet zijn vragen op eenvoudige wijze kunnen vertalen naar een database-structuur. Dit artikel behandelt een praktijk-case over kostenbeheersing binnen het UWV, waarbij gekozen is voor een sterstructuur.

Het UWV (Uitvoering Werknemers Verzekeringen) is een organisatie met een ingewikkelde organisatiestructuur, bestaande uit een Raad van Bestuur met stafafdelingen en meerdere divisies met stafafdelingen en regiokantoren. De indeling van de afzonderlijke divisies verschilt ook weer van elkaar (de structuur is te karakteriseren als een ongebalanceerde boom). Zo'n structuur kan getekend worden als een boom die begint met de Raad van Bestuur. Elk punt in de structuur is verantwoordelijk voor alle punten eronder.

De kosten die door de verschillende UWV-onderdelen worden gemaakt worden geregistreerd op die punten in de boomstructuur.



Afbeelding 1: Organisatiestructuur.

Elk punt in de boomstructuur wordt een kostenplaats genoemd. De vragen over die kosten gaan of over de kosten die in een specifiek punt zijn geboekt of over de kosten waarvoor het punt verantwoordelijk is, dus de kosten in het punt zelf plus alle kosten van de punten onder dat punt in de boomstructuur.

Zie als voorbeeld de organisatiestructuur in afbeelding 1. De vraag (nummer een) naar de kosten van de Raad van Bestuur (RVB) zelf is natuurlijk eenvoudig te beantwoorden: RVB = 25. De (tweede)vraag naar *alle* kosten van divisie 1 (DIV 1) is te beantwoorden door de kosten van alle punten *onder* kostenplaats DIV 1 bij elkaar te tellen (inclusief de kosten van DIV 1 zelf). Het antwoord luidt: DIV 1 + Staf DIV 1 + Regio_1 + Regio_2 + Regio Staf 1 + Regio 1 Kantoor 1 + Regio 1 Kantoor 2 + Regio 1 Kantoor 3 + Regio 2 Kantoor 1 + Regio 2 Kantoor 2 = 219. De vraag (nummer 3) naar de kosten van de uitvoerende afdelingen van divisie 1 is al weer lastiger te beantwoorden; hiervoor dienen de kosten van de stafafdelingen buiten beschouwing gelaten te worden.

Het antwoord: DIV 1 + Regio_1 + Regio_2 + Regio 1 Kantoor 1 + Regio 1 Kantoor 2 + Regio 1 Kantoor 3 + Regio 2 Kantoor 1 + Regio 2 Kantoor 2 = 213.

In het kader van de beheersing van deze kosten worden ze in een database-structuur opgenomen die ingericht is op het bevragen en niet op het registreren er van. Die structuur moet enerzijds eenvoudig te begrijpen zijn door de gebruiker en anderzijds de mogelijkheid bieden om alle mogelijke vragen te beantwoorden. De gekozen structuur waarmee dit soort vragen door een eindgebruiker op een eenvoudige wijze gesteld en beantwoordt kunnen worden, en hoe de betreffende tabellen gevuld worden vanuit het aangeleverde hiërarchische bestand, zullen verder worden uitgewerkt.

Model van de informatiebehoefte

De bovenstaande informatiebehoefte is in een model om te zetten.

Telgrootheden.

- Eigen kosten;
- Verantwoordelijke kosten.

Selectie- of groepeeringsgrootheden.

Het punt in de organisatiestructuur waarin de gebruiker geïnteres-

seerd is, is aan de hand van de volgende kenmerken te typeren:

- de code;
- de naam;
- of de typering van de kostenplaats.

(De eerste twee vragen zijn voorbeelden van selecties aan de hand van deze kenmerken.)

En elk punt van de organisatiestructuur waarvoor dat geselecteerde punt verantwoordelijk is, is te selecteren aan de hand van:

- de code;
- de naam;
- of de typering van de *afhankelijke* kostenplaats.

De derde vraag is hier een voorbeeld van: er wordt gevraagd naar de kosten van alle afhankelijke kosten die niet van het type 'Stafafdeling' zijn.

Gevolgdewerkwijze

Gegeven de eis dat een eindgebruiker zijn vragen op eenvoudige wijze moet kunnen vertalen naar de structuur, is gekozen voor een sterstructuur. De ervaring binnen UWV is dat zo'n structuur snel uit te leggen is aan de gebruiker en ook goed begrepen wordt. Eén van de mogelijke oplossingen was het in ieder punt vastleggen van niet alleen de eigen kosten, maar ook de kosten van de totale boom eronder, bijvoorbeeld voor DIV 1: eigen kosten = 9; totale kosten 219. Het grote nadeel van deze manier is dat het daarmee niet mogelijk om alle soorten vragen eenvoudig te beantwoorden. Neem bijvoorbeeld vraag 3; deze is alleen te beantwoorden door van de totale kosten de kosten van de twee stafafdelingen af te trekken.

Eén van de oplossingen was het vastleggen van de kosten van de totale boom eronder

Het is natuurlijk ook niet mogelijk om gewoon even alle mogelijkheden uit te rekenen. Dat zou enerzijds te veel ruimte vragen en anderzijds veel te lastig zijn om al die combinaties van een begrijpelijke naam te voorzien.

Kenmerkend voor dit soort vragen over kosten is dat ze gebaseerd zijn op een verzameling punten (kostenplaatsen) die een onderlinge relatie hebben (ouder/kind in 1 of meer stappen) én aan één of meer eisen moeten voldoen.

De eerste vraag is simpelweg te herformuleren als de eigen kosten van het punt 'Raad van Bestuur'. Vraag nummer 2 is te beantwoorden met de verzameling punten die als (groot)ouder de kostenplaats DIV 1 hebben. De derde vraag wordt beantwoord met de verzameling van alle punten die een ouder/kind-relatie hebben met als ouder DIV 1 én als typering ongelijk aan 'staf' hebben. Op deze benadering is de gekozen oplossing gebaseerd. Deze oplossing betreft een sterstructuur waarin een feitentabel is

gedefinieerd (de kostentabel), een dimensietabel (de kostenplaatsentabel) en een relatietabel waarin de relaties tussen de kostenplaatsen worden vastgelegd.

1. De KOSTENPLAATSEN-tabel (dimensietabel)

De KOSTENPLAATSEN-tabel bevat voor iedere kostenplaats de code, een typering en andere kenmerken van de kostenplaats, zoals bijvoorbeeld een naam. In deze tabel legt men ook vast wat de direct hogere kostenplaats in de boomstructuur is. Dit gegeven wordt later gebruikt voor het bepalen van de relaties.

Kolomnaam	Datatype	Omschrijving
Kostenplaatsnummer	String(15)	Logische sleutel, de code van de kostenplaats.
Kostenplaats_naam	String(30)	De naam van de kostenplaats.
Type_kostenplaats	String(15)	Codering van het type kostenplaats, mogelijke waarden: DIV = divisie, DIV-ST = staf van de divisie, REGIO = regiokantoor, REGIO-ST = staf van de regio.
Kostenplaats_ouder	String(15)	De code van de direct hogere kostenplaats waartoe de betreffende kostenplaats behoort.

Als voorbeeld:

KOSTENPLAATS-NUMMER	KOSTENPLAATS-NAAM	TYPE-KOSTENPLAATS	KOSTENPLAATS-OUDER
DIV 1	Divisie 1	DIV	RVB
STAF DIV 1	Stafafdeling Divisie 1	DIV-ST	DIV 1
REGIO 2	Regio 2 van Divisie 1	DIV_UITV	DIV 1
R2-K2	Kantoor 2 van Regio 2	KANT	REGIO 2
RVB	Raad van Bestuur	Overig	n.v.t.

2. De KOSTEN-tabel (feitentabel)

In de KOSTEN-tabel worden de geboekte kosten per maand vastgelegd bij de direct verantwoordelijke kostenplaats.

Kolomnaam	Datatype	Omschrijving
Kostenplaats	String(15)	De kostenplaatscode waar de kosten geboekt zijn.
Maand	String(6)	De maand waarin de kosten geboekt zijn.
Bedrag	Number(13,2)	Het geboekte bedrag.

Bijvoorbeeld:

Kostenplaats	Maand	Bedrag
DIV I	200312	9,00
STAF DIV I	200312	5,00
R2-K2	200312	37,00

3. De RELATIES-tabel

In de RELATIES-tabel wordt alle ouder-kind relaties vastgelegd die bestaan tussen de kostenplaatsen. Tevens wordt elke kostenplaats met een verwijzing naar zichzelf opgenomen.

Kolomnaam	Datatype	Omschrijving
Kostenplaats_ouder	String(15)	De kostenplaats waartoe de kindkostenplaats gerekend moet worden.
Kostenplaats_kind	String(15)	De kostenplaatscode waar de kosten geboekt zijn.
Afstand	Number(11)	De afstand tussen de twee kostenplaatsen.

Bijvoorbeeld:

KOSTENPLAATS_OUDER	KOSTENPLAATS_KIND	AFSTAND
DIV I	DIV I	0
DIV I	STAF DIV I	1
DIV I	R2-K2	2
REGIO 2	R2-K2	1

In deze tabel wordt dus van elk punt in de boom (ouder) vastgelegd welke punten (kinderen) de totale kosten van het betreffende punt bepalen. Anders gezegd: in deze tabel geeft men voor elke kostenplaats (kind) aan, aan welke hogere kostenplaats (ouder) de kosten moeten worden toegerekend. In de tabel wordt elke subboom van de oorspronkele boomstructuur vastgelegd. De bovenbeschreven tabellen zijn op een wijze aan elkaar gekoppeld zoals in afbeelding 2 te zien is. De feitentabel KOSTEN bevat voor elke kostenplaats de kosten per maand. Die tabel wordt niet direct aan de KOSTENPLAATSEN-tabel gekoppeld maar indirect en wel via de tabel RELATIES, waarbij de koppeling wordt gelegd met het veld KOSTENPLAATS_KIND. De KOSTENPLAATSEN-tabel waarin de eigenschappen van de individuele kostenplaatsen (bijvoorbeeld naam) zijn opgenomen, wordt twee keer aan de RELATIES-tabel gekoppeld. Eén keer voor de ouder (met alias KOSTENPLAATSEN_OUDER) en de tweede keer voor het kind (met alias KOSTENPLAATSEN_KIND).

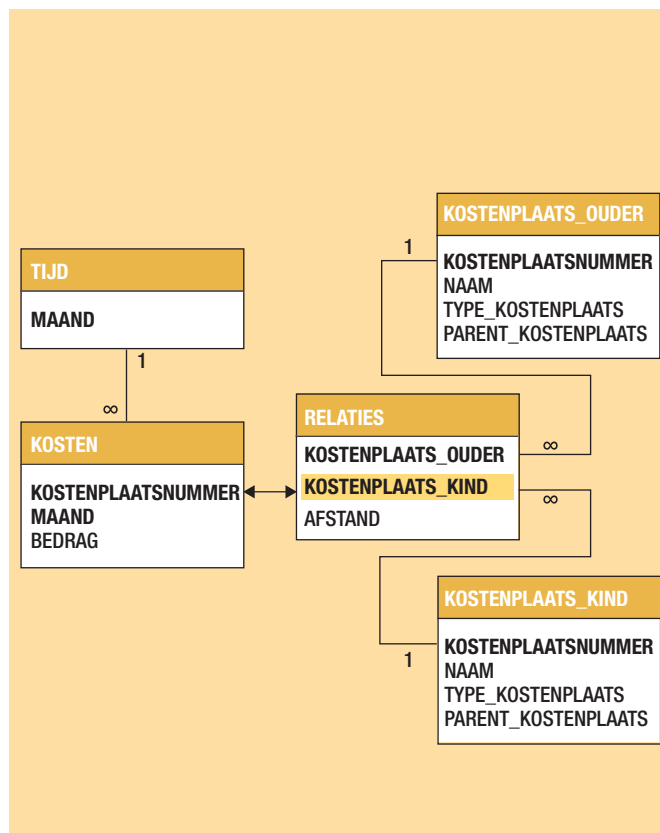
Via de tabel RELATIES kan men voor elke kostenplaats de 'ouders' vaststellen, waartoe de geboekte kosten gerekend moeten worden. Omgekeerd is voor elke ouder-kostenplaats (in de tabel KOSTENPLAATSEN_OUDER) vast te stellen welke kosten er zijn gemaakt door de kinderen van deze kostenplaats (via de tabel RELATIES). Met behulp van deze structuur kunnen de drie vragen uit het begin van het artikel beantwoord worden.

Vraag 1. Geef de kosten van de Raad van Bestuur zelf.

```

Select sum(bedrag)
From KOSTEN,
      RELATIES
      KOSTENPLAATS_OUDER
Where KOSTEN.KOSTENPLAATSNUMMER =
      RELATIES.KOSTENPLAATS_KIND
AND RELATIES.KOSTENPLAATS_OUDER =
      KOSTENPLAATS_OUDER.KOSTENPLAATSNUMMER
AND KOSTENPLAATS_OUDER.NAAM = 'Raad van Bestuur'
AND RELATIES.AFSTAND = 0
    
```

Het opnemen van de eis dat de afstand 0 moet zijn, zorgt ervoor dat alleen de relatie RVB - RVB wordt gevonden en daarmee dat alleen de eigen kosten van de kostenplaats 'Raad van Bestuur' worden gesommeerd.



Afbeelding 2: De tabellen en hun relaties.

Vraag 2. Geef alle kosten van de Divisie 1.

```
Select sum(bedrag)
From KOSTEN,
      RELATIES,
      KOSTENPLAATS_OUDER
Where KOSTEN.KOSTENPLAATSNUMMER =
      RELATIES.KOSTENPLAATS_KIND
AND RELATIES.KOSTENPLAATS_OUDER =
      KOSTENPLAATS_OUDER.KOSTENPLAATSNUMMER
AND KOSTENPLAATS_OUDER.NAAM = 'Divisie 1'
```

In deze selectie wordt *geen* beperking gelegd aan de afstand van het kind ten opzichte van de ouder. Daarmee worden alle kosten waarvoor ouder 'Divisie 1' verantwoordelijk is meegenomen.

Vraag 3. Geef de kosten van alle uitvoerende afdelingen van de Divisie 1.

```
Select sum(bedrag)
From KOSTEN,
      RELATIES,
      KOSTENPLAATS_OUDER,
      KOSTENPLAATS_KIND
Where KOSTEN.KOSTENPLAATSNUMMER =
      RELATIES.KOSTENPLAATS_KIND
AND RELATIES.KOSTENPLAATS_OUDER =
      KOSTENPLAATS_OUDER.KOSTENPLAATSNUMMER
AND RELATIES.KOSTENPLAATS_KIND =
      KOSTENPLAATS_KIND.KOSTENPLAATSNUMMER
AND KOSTENPLAATS_OUDER.NAAM = 'Divisie 1'
AND KOSTENPLAATS_KIND.TYPE <> 'Staf'
```

Vertaling van de boom naar paren

In de KOSTENPLAATSEN-tabel worden alleen de directe parent/child-relaties tussen de verschillende kostenplaatsen opgenomen. Met het onderstaande, niet recursieve, algoritme wordt vanuit die tabel de RELATIES-tabel gevuld.

Bijvoorbeeld de KOSTENPLAATSEN-tabel:

KOSTENPLAATS- NUMMER	NAAM	TYPE_ KOSTENPLAATS	KOSTENPLAATS_ OUDER
RVB	Raad van bestuur	RVB	n.v.t.
DIV I	Divisie I	DIV	RVB
REGIO I	REGIO I	DIV_UITV	DIV I

Uit deze tabel kan men met behulp van het volgende proces de RELATIES-tabel waarin ALLE paren zijn opgenomen opbouwen. Start met een lege RELATIES-tabel.

Restrictie

Door het opnemen van een restrictie ten aanzien van de afstand om het onderscheid te maken tussen de eigen en de verantwoordelijke kosten, is het niet mogelijk om in één query beide kosten te tonen. Door die restrictie niet op te nemen als een beperking voor de verzameling op te halen kostenplaatsen, maar op te nemen in het te sommeren bedrag, wordt dit wel mogelijk.

Als voorbeeld: Geef de eigen kosten én de verantwoordelijke kosten van de 'Raad van Bestuur'.

```
Select sum(case afstand
            = 0, bedrag
            else 0) eigen kosten,
      sum(bedrag) verantwoordelijke kosten
From KOSTEN,
      RELATIES,
      KOSTENPLAATS_OUDER
Where KOSTEN.KOSTENPLAATSNUMMER =
      RELATIES.KOSTENPLAATS_KIND
AND RELATIES.KOSTENPLAATS_OUDER =
      KOSTENPLAATS_OUDER.KOSTENPLAATSNUMMER
AND KOSTENPLAATS_OUDER.NAAM =
      'Raad van Bestuur'
```

Voer voor elke rij in de KOSTENPLAATSEN-tabel de volgende drie stappen uit. Doe dit in de volgorde waarin de kostenplaatsen in de organisatiestructuur voorkomen, startend met de top. (Dit is eenvoudig op te halen met de volgende SQL-statement: `select * from kostenplaatsen_dim where kostenplaats_ouder = 'n.v.t.'`).

1. Zoek in de RELATIES-tabel alle rijen waarvan de `kostenplaats_kind` gelijk is aan de `kostenplaats_ouder` in de KOSTENPLAATSEN-tabel van de te verwerken rij;
2. Voeg deze rijen toe aan de RELATIES-tabel met `kostenplaats_ouder = kostenplaats_ouder` uit de RELATIES-tabel en `kostenplaats_kind = kostenplaatsnummer` uit de KOSTENPLAATSEN-tabel en `afstand = afstand` uit de RELATIES-tabel + 1;
3. Voeg een nieuwe rij toe aan de RELATIES-tabel met `kostenplaats_ouder = kostenplaatsnummer` van de betreffende rij uit de KOSTENPLAATSEN-tabel en `kostenplaats_kind = kostenplaatsnummer` van de betreffende rij uit de KOSTENPLAATSEN-tabel en `afstand = 0`.

Als dit proces wordt uitgevoerd, dan ontstaan de volgende (tussen)resultaten. Rij 1 in de kostenplaatsen tabel:

KOSTENPLAATS- NUMMER	NAAM	TYPE_ KOSTENPLAATS	KOSTENPLAATS_ OUDER
RVB	Raad van Bestuur	RVB	n.v.t.
DIV I	Divisie I	DIV	RVB
REGIO I	REGIO I	DIV_UITV	DIV I

Stap 1. Zoek in de RELATIES-tabel alle rijen waarvan de kostenplaats_kind gelijk is aan de kostenplaats_ouder in de KOSTENPLAATSEN-tabel van de te verwerken rij. Omdat de RELATIES-tabel nog leeg is, zullen er geen rijen zijn die voldoen.

Stap 2. Voeg deze rijen toe aan de RELATIES-tabel (kostenplaats_ouder, kostenplaats_kind, afstand) = (RELATIES.kostenplaats_ouder, KOSTENPLAATSEN.kostenplaatsnummer, RELATIES.afstand + 1). Ook nu nog blijft de RELATIES-tabel leeg.

Stap 3. Voeg een nieuwe rij toe aan de RELATIES-tabel met (kostenplaats_ouder, kostenplaats_kind, afstand) = (KOSTENPLAATSEN.kostenplaatsnummer, KOSTENPLAATSEN.kostenplaatsnummer, 0).

De RELATIES-tabel na stap 3 voor de eerste rij:

KOSTENPLAATS OUDER	KOSTENPLAATS KIND	AFSTAND
RVB	RVB	0

Voer hetzelfde proces nu uit voor de tweede rij in de KOSTENPLAATSEN-tabel.

KOSTENPLAATS- NUMMER	NAAM	TYPE_ KOSTENPLAATS	KOSTENPLAATS_ OUDER
RVB	Raad van bestuur	RVB	n.v.t.
DIV I	Divisie I	DIV	RVB
REGIO I	REGIO I	DIV_UITV	DIV I

Stap 1. Zoek in de RELATIES-tabel alle rijen waarvan de kostenplaats_kind gelijk is aan de kostenplaats_ouder in de KOSTENPLAATSEN-tabel van de te verwerken rij. Dit levert 1 rij op (namelijk de rij die in de RELATIES-tabel aanwezig is).

KOSTENPLAATS_OUDER	KOSTENPLAATS_KIND	AFSTAND
RVB	RVB	0

Stap 2. Voeg deze rijen toe aan de RELATIES-tabel (kostenplaats_ouder, kostenplaats_kind, afstand) = (RELATIES.kostenplaats_ouder, KOSTENPLAATSEN.kostenplaatsnummer, RELATIES.afstand + 1). De RELATIES-tabel na stap 2 voor de tweede rij:

KOSTENPLAATS_OUDER	KOSTENPLAATS_KIND	AFSTAND
RVB	RVB	0
RVB	DIV I	0 + 1 = 1

Stap 3. Voeg een nieuwe rij toe aan de RELATIES-tabel met (kostenplaats_ouder, kostenplaats_kind, afstand) = (KOSTENPLAATSEN.kostenplaatsnummer, KOSTENPLAATSEN.kostenplaatsnummer, 0).

RELATIES-tabel na stap 3 voor de tweede rij:

KOSTENPLAATS_OUDER	KOSTENPLAATS_KIND	AFSTAND
RVB	RVB	0
RVB	DIV I	1
DIV I	DIV I	0

Voer ten slotte hetzelfde proces nog eens uit voor de laatste rij in de kostenplaatsen tabel.

KOSTENPLAATS- NUMMER	NAAM	TYPE_ KOSTENPLAATS	KOSTENPLAATS_ OUDER
RVB	Raad van bestuur	RVB	n.v.t.
DIV I	Divisie I	DIV	RVB
REGIO I	REGIO I	DIV_UITV	DIV I

Stap 1. Zoek in de RELATIES-tabel alle rijen waarvan de kostenplaats_kind gelijk is aan de kostenplaats_ouder in de KOSTENPLAATSEN-tabel van de te verwerken rij.

Dit levert de volgende rijen op uit de RELATIES-tabel:

KOSTENPLAATS_OUDER	KOSTENPLAATS_KIND	AFSTAND
RVB	DIV I	1
DIV I	DIV I	0

Stap 2. Voeg deze rijen toe aan de RELATIES-tabel (kostenplaats_ouder, kostenplaats_kind, afstand) = (RELATIES.kostenplaats_ouder, KOSTENPLAATSEN.kostenplaatsnummer, RELATIES.afstand + 1).

RELATIES-tabel na stap 2 voor de derde rij:

KOSTENPLAATS_OUDER	KOSTENPLAATS_KIND	AFSTAND
RVB	RVB	0
RVB	DIV I	1
DIV I	DIV I	0
RVB	REGIO I	1 + 1 = 2
DIV I	REGIO I	0 + 1 = 1

Stap 3. Voeg een nieuwe rij toe aan de RELATIES-tabel met (kostenplaats_ouder, kostenplaats_kind, afstand) = (KOSTENPLAATSEN.kostenplaatsnummer, KOSTENPLAATSEN.kostenplaatsnummer, 0).

RELATIES-tabel na stap 3 voor de derde rij:

KOSTENPLAATS_OUDER	KOSTENPLAATS_KIND	AFSTAND
RVB	RVB	0
RVB	DIV I	1
DIV I	DIV I	0
RVB	REGIO I	2
DIV I	REGIO I	1
REGIO I	REGIO I	0

Nu is de RELATIES-tabel gevuld met alle bestaande relaties tussen deze drie kostenplaatsen.

De pl-sql die nodig is voor deze bewerking is:

```
Procedure create_paren IS
l_kostenplaats kostenplaatsen_dim%ROWtype;
cursor cur_kost is
  SELECT * FROM Kostenplaatsen_dim kopladim
  START WITH kopladim.kostenplaatsnummer =
    (SELECT kostenplaatsnummer
     from kostenplaatsen_dim
     where kostenplaats_ouder = 'Nvt')
  CONNECT BY kopladim.kostenplaatsnummer =
    PRIOR kopladim.kostenplaats_ouder
  ORDER BY LEVEL;
Begin
FOR l_kostenplaats IN cur_kost
LOOP
  INSERT INTO Relatie_tabel (Kostenplaats_ouder,
    Kostenplaats_kind,
    Afstand)
  Select rel.kostenplaats_ouder,
    kopl.kostenplaatsnummer,
    rel.afstand + 1
```

```
From Relatie_tabel rel,
      Kostenplaatsen_dim kopl
where kopl.kostenplaats_ouder =
      rel.kostenplaats_kind
and   kopl.kostenplaatsnummer =
      l_kostenplaats.kostenplaatsnummer;

INSERT INTO Relatie_tabel (Kostenplaats_ouder,
  Kostenplaats_kind,
  Afstand)
Select l_kostenplaats.kostenplaatsnummer,
      l_kostenplaats.kostenplaatsnummer,
      0
from dual;
Commit;
END LOOP;
End;
```

Conclusie

Een grote organisatie heeft vaak een ingewikkelde organisatie-structuur en die complexiteit maakt het moeilijk om vragen over kosten te beantwoorden. Beschreven is hoe dit soort vraagstukken met behulp van een standaard-structuur beantwoord kunnen worden zonder over kennis van de betreffende organisatiestructuur te hoeven beschikken.

Paul Hulst en Fons Rooijers

Fons Rooijers (fons.rooijers@uwv.nl) werkt bij UWV en werkt aan het datawarehouse-project van UWV.

Paul Hulst (phulst@deloitte.nl) werkt bij Deloitte & Touche en was betrokken bij het datawarehouse-project van UWV.