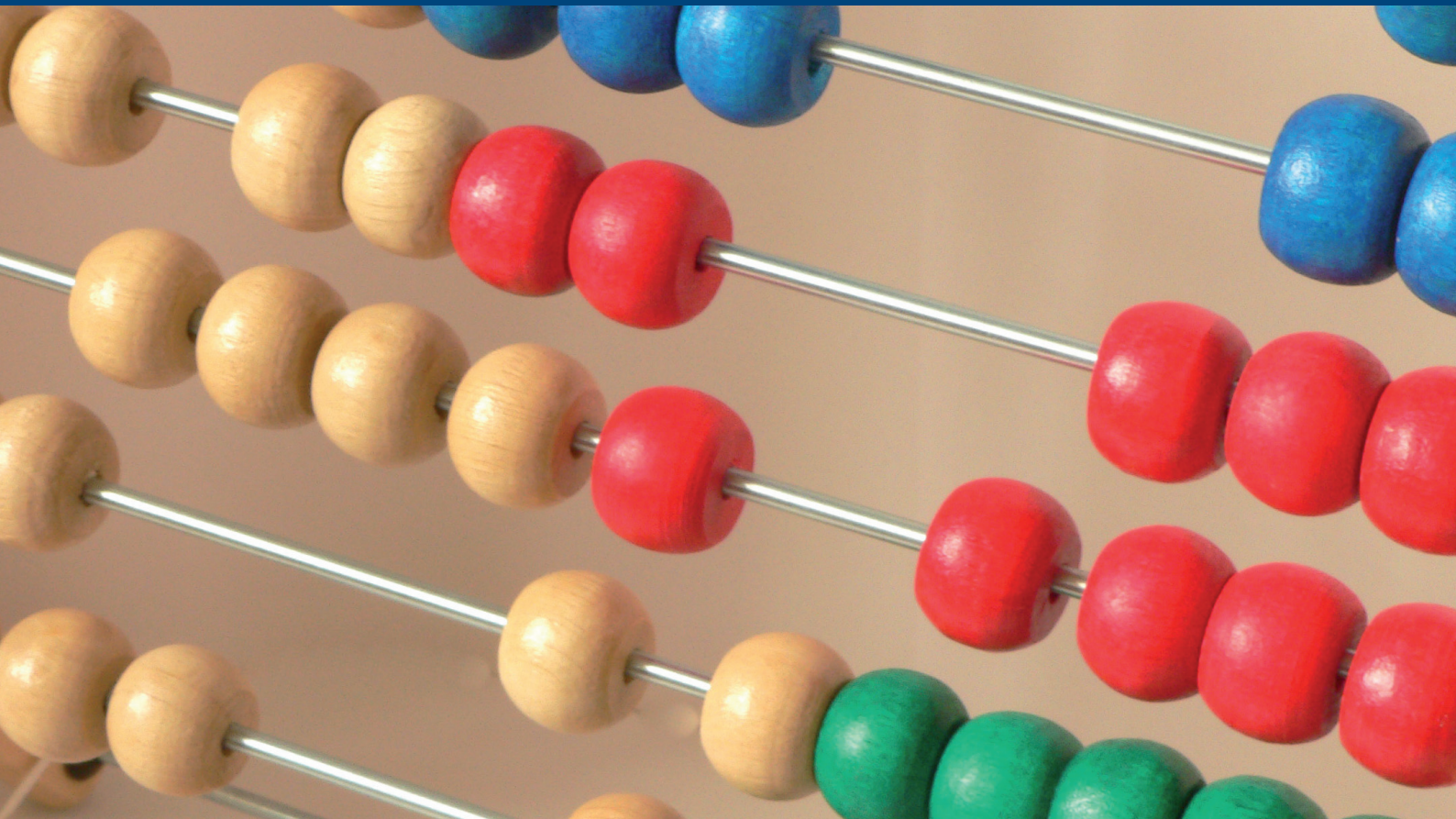


Scriptie



Begroten van datamigratietrajecten bij vervanging van informatiesystemen

Auteur: Edwin Paulissen

Studentnummer: 838191288

Datum: 4 september 2009

Begroten van datamigratietrajecten bij vervanging van informatiesystemen

Afstudeerscriptie

Edwin Paulissen
838191288

4 september 2009

Open Universiteit Nederland,
faculteiten Managementwetenschappen en Informatica
Masteropleiding Business Process Management and IT

Afstudeercommissie:

Voorzitter/begeleider: Prof.dr.ir. A.J. Udink ten Cate, Open Universiteit Nederland

Tweede beoordelaar: Dr. E. Roubtsova, Open Universiteit Nederland

Bedrijfsbegeleider: Ing. T.M.J.C. van Dort, Mediaan/abs

Inhoudsopgave

1.1	Voorwoord	7
1.2	Samenvatting	9
1.3	Summary	11
2	Inleiding	13
3	Datamigratie	17
3.1	Op zoek naar een procesbeschrijving van datamigratie	17
3.2	Datamigratieproces	20
3.3	Datamigratiefasen	23
3.4	Scope bepaling	23
3.5	Migratiestrategie	24
3.6	Interfaces	25
3.7	Infrastructuur	25
3.8	Conclusie	26
4	Mediaan Migratie Methode	27
4.1	Methode	27
4.2	Aspecten van datamigratie	28
4.3	Conclusie	29
5	Functiepuntanalyse	31
5.1	Ontstaan	31
5.2	Afbakening en toepassing	31
5.3	Introductie	32
5.4	Telrichtlijnen	34
5.5	Calculeren van kosten	36
5.6	Kengetallen	38
5.7	Recente ontwikkelingen	38
5.8	Conclusie	39
6	Het begroten van datamigratietrajecten	41
6.1	Quickscan van de praktijk	41
6.2	Calculeren omvang datamigratietrajecten	43
6.3	WBS relateerde uren	46
6.4	Datamigratiefunctiepunten	47
6.5	Productiviteitsfactoren	49
6.6	Telecom case	51
6.7	Conclusie	56
7	Empirisch onderzoek	57
7.1	Migratieaanpak	57
7.2	Migratietool keuze	57
7.3	Uitvoering	58
7.4	Analyse	59
7.5	Conclusie	61

8	Conclusies en aanbevelingen	63
8.1	Onderzoeksvragen	64
8.2	Algemene conclusie	65
8.3	Aanbeveling voor vervolgonderzoek	66
9	Literatuur	67
	Bijlage I Vragenlijst om een datamigratiebegrotingsmethode te ontwikkelen	71
	Bijlage II Mediaan Migratie Methode	73
	Vooronderzoek	73
	Definitie	74
	Ontwerp	76
	Realisatie	77
	Test	78
	Pre-Productie	78
	Productie	79
	Changemanagement	80
	Projectmanagement	80

1.1 Voorwoord

De afgelopen jaren ben ik diverse malen in aanraking gekomen met functiepuntanalyse en heb daarbij de overtuiging gekregen dat de omvang van informatiesystemen uitstekend met functiepuntanalyse gecalculeerd kan worden. Deze methode wordt veel toegepast om ontwikkel- en beheertrajecten op basis van het aantal functiepunten voor een fixed price aan te bieden. Datamigratietrajecten worden veelal op basis van nacalculatie uitgevoerd, maar er is een steeds grotere behoefte om ook bij datamigratietrajecten een fixed price aanbieding te doen. Datamigratietrajecten is een van de speerpunten van Mediaan/abs, mijn werkgever.

Het was niet eenvoudig het onderzoek uit te voeren binnen de ongeveer zes maanden die hiervoor staan. Het onderzoek was helaas niet te combineren met mijn dagelijkse werkzaamheden, waardoor ik aangewezen was op de avonden, het weekend en vele verlofdagen. In totaal heb ik bijna twee jaar aan mijn afstuderen gewerkt en is meer dan de dubbele tijd besteed die voor afstuderen geraamd is.

Een dergelijk onderzoek doe je nooit helemaal alleen. Diverse mensen hebben mij op verschillende manieren geholpen bij de uitvoering van mijn onderzoek. Daarvoor wil ik iedereen ontzettend bedanken. In het bijzonder gaat mijn dank uit naar Alexander Udink ten Cate (Open Universiteit Nederland), Kees Aarts (Open Universiteit Nederland), Anda Counotte-Potman (Open Universiteit Nederland), Ella Roubtsova (Open Universiteit Nederland) en Theo van Dort (Mediaan/abs).

Verder wil ik de medewerkers van Mediaan/abs bedanken die aan mijn onderzoek hebben meegewerkt door middel van het invullen van vragenlijsten en de vele uren dat we over het onderzoek gediscussieerd hebben.

Rest mij nog mijn vrouw te bedanken. Zonder haar ondersteuning en hulp was me dit nooit gelukt.

Edwin Paulissen
Maastricht, mei 2009

1.2 Samenvatting

Het einde van de levenscyclus van een informatiesysteem gaat in veel gevallen gepaard met datamigratie naar de opvolger van het oude informatiesysteem. Het uitvoeren van datamigraties is een complexe aangelegenheid. In de literatuur is weinig beschreven over de hierbij uit te voeren stappen. De eigenaar van het te migreren informatiesysteem heeft meestal geen ervaring met deze werkzaamheden en daarom worden experts benaderd om de datamigratie uit te voeren. De belangrijkste vragen die de experts moeten beantwoorden voor de opdracht gegund wordt, betreft de aanpak van het datamigratietraject en de kosten die daarmee gepaard gaan. In dit verslag wordt een onderzoek gepresenteerd over de manier waarop datamigraties aangepakt en begroot kunnen worden.

Een datamigratietraject is onderdeel van een migratietraject. In het migratietraject wordt het nieuwe informatiesysteem ontwikkeld of ingericht en worden ook de organisatie en de processen ingericht om te werken met het nieuwe informatiesysteem. Het migratietraject zorgt onder andere voor scholing van medewerkers, afsluiten van contracten met leveranciers en uitfasen van het oude informatiesysteem. In dit verslag is het onderwerp datamigratie. In een datamigratietraject is de scope het overzetten van de data van het oude informatiesysteem naar het nieuwe informatiesysteem.

Een literatuurstudie is uitgevoerd om in standaardmethodieken een methode te vinden voor het uitvoeren van datamigratietrajecten. Gedurende het literatuuronderzoek is contact gelegd met een van de grondleggers van ASL en BiSL. In de standaardmethodieken zijn géén aanknopingspunten gevonden voor een projectmatige aanpak van een datamigratietraject. Deze uitkomst ligt niet in lijn met de aanvankelijke verwachtingen.

Daarna is een quickscan gedaan onder een aantal datamigratie experts en is een projectmatige aanpak gevonden waarmee een datamigratietraject uitgevoerd kan worden. Deze best practice is door een werkgroep bij Mediaan/abs ontwikkeld. De methode bevat een aantal fasen die in volgorde worden doorlopen. Meestal wordt een aantal fasen iteratief doorlopen. De fasen zijn verder onderverdeeld in een aantal onderling samenhangende activiteiten en voor elke activiteit is compact beschreven wat in die activiteit gebeurt.

Doordat een datamigratiemethode gekozen is kunnen datamigraties op dezelfde manier uitgevoerd worden, waardoor ervaringen uit vorige trajecten gebruikt kunnen worden in calculaties van komende trajecten.

Alleen een methode om datamigratietrajecten uit te voeren is niet voldoende om een begroting te calculeren. Daar zijn cijfers over omvang en productiviteit voor nodig. Er is literatuuronderzoek gedaan naar functiepuntanalyse en productiviteitscijfers.

Calculeren van de omvang wordt voor delen van een datamigratietraject gedaan met behulp van functiepuntanalyse volgens de NESMA telrichtlijnen (NESMA, 2004). Door het uitvoeren van een globale of detail functiepuntanalyse kan de omvang in functiepunten bepaald worden. Indien de organisatie over productiviteitscijfers op basis van functiepunten beschikt kan voor het uitvoeren van datamigratietrajecten de omvang van de te besteden uren bepaald worden.

Tijdens het onderzoek is van de activiteiten van de datamigratiemethode beschreven of

de omvang bepaald kan worden met functiepunten. Helaas kunnen niet alle onderdelen van een datamigratietraject via functiepunten begroot worden, denk aan strategiebepaling, uitvoeren van de datamigratie en projectmanagement. Deze activiteiten worden via een work breakdown structure begroot.

Een van de moeilijkst te begroten activiteiten in een datamigratietraject is opschoning. Opschonen van data is het verhogen van de kwaliteit van de gegevens door fouten te herstellen die historisch in de data zijn geslopen. Het bepalen van de kwaliteit van de te migreren gegevens en inschatten hoeveel tijd nodig is om de kwaliteit te verhogen is op dit moment niet via een breed geaccepteerde methode uit te voeren. De metingen die nodig zijn voor opschoning worden niet meegenomen in de kostprijsbepaling voor datamigratietrajecten in dit onderzoek.

Voor het calculeren van de omvang van datamigratietrajecten in metingen is een formule opgesteld. Middels een case studie is onderzocht of het aantal tabellen en complexiteit een grote impact heeft op de kostprijs per functiepunt.

Cijfermateriaal van productiviteit voor het uitvoeren van datamigratietrajecten is niet voorhanden. Dit is een onderwerp dat vervolgstudie verdient.

Dit onderzoek vormt een zinvolle bijdrage aan de toegepaste wetenschap, door een op best practices gebaseerde methode te presenteren, om datamigratietrajecten uit te voeren. Daarnaast zijn formules beschreven om de omvang van een datamigratietraject te bepalen. Door middel van een case studie en proof of principle worden de methode en formules getoetst. Functiepuntanalyse kan toegepast worden voor een deel van het datamigratietraject om de omvang in metingen te bepalen.

2 Inleiding

In de praktijk worden bestaande informatiesystemen regelmatig vervangen door nieuwe informatiesystemen, omdat ze bijvoorbeeld aan het einde van hun levenscyclus zijn, een fusie van de organisatie heeft plaatsgevonden of omdat een maatwerkpakket vervangen wordt door een standaardpakket. Bij dit vervangen is het van belang dat bedrijfsgegevens uit de oude systemen worden overgezet naar de nieuwe systemen, zodat de bedrijfsvoering ongehinderd door kan gaan. De gegevens uit een oud systeem kunnen meestal niet eenvoudig worden overgebracht naar een nieuw systeem. Om gegevens van het ene systeem naar het andere systeem over te zetten is een datamigratietraject nodig. In het datamigratietraject worden de gegevens van het bronsysteem (oude informatiesysteem) overgezet naar het doelsysteem (nieuwe informatiesysteem).

Een datamigratietraject is onderdeel van een groter meer omvattend migratietraject. In het migratietraject wordt het nieuwe informatiesysteem ontwikkeld of ingericht en worden de organisatie en de processen ingericht om te werken met het nieuwe informatiesysteem. Het migratietraject omvat onder andere de scholing van de medewerkers, het afsluiten van contracten met leveranciers en het uitfaseren van het bronsysteem. In dit verslag is het onderwerp datamigratie. Een datamigratietraject richt zich op het overzetten van de data van het bronsysteem naar het doelsysteem.

Een datamigratietraject heeft een éénmalig karakter. Datamigratietrajecten zijn aan de orde bij invoering van een nieuw softwarepakket of informatiesysteem, bij samenvoeging of splitsing van informatiesystemen of bij grootschalige aanpassingen aan informatiesystemen.

De trend in de ICT wereld is om datamigratietrajecten uit te besteden, omdat datamigratie een éénmalige activiteit is, die specifieke expertise nodig heeft. Dit is niet zonder problemen. Meer dan 80% van de datamigratietrajecten falen volledig of schieten over hun geplande budgetten en deadlines heen (Standish Group, 1999). Om uit te besteden is het voor de eigenaar van het te migreren informatiesysteem van groot belang dat de kosten en doorlooptijd vooraf geschat kunnen worden en beheersbaar zijn. Beheersbaar in de vorm van volledigheid, juistheid en tijdigheid. Beheersbaarheid is nodig om goede afspraken te maken met de leverancier die het datamigratietraject gaat uitvoeren.

Voor het ontwerpen, ontwikkelen, testen en beheren van informatiesystemen zijn geaccepteerde methoden beschikbaar om projectomvang, doorlooptijd en kosten te berekenen. Voorbeelden hiervoor zijn lines of codes en functiepuntanalyse. Voor datamigratietrajecten bestaat er nog geen algemeen geaccepteerde methode om de omvang in uren te calculeren.

Doel van het onderzoek is te bepalen of functiepuntanalyse (FPA) te gebruiken is bij het vooraf bepalen van kosten, omvang en complexiteit van datamigratietrajecten. De probleemstelling is gericht op het transparant maken van de omvang en complexiteit van datamigratietrajecten, zodat de benodigde inspanning en doorlooptijd om een datamigratie volledig en correct uit te voeren vooraf gecalculeerd kan worden.

De probleemstelling wordt gesplitst in twee centrale vragen:

- Kan FPA gebruikt worden om de omvang van datamigraties te bepalen?
- Hoe kunnen omvang, kosten en doorlooptijd begroot worden van onderdelen die van een datamigratietraject die niet met FPA begroot kunnen worden?

Om te bepalen voor welk deel van het datamigratietraject FPA toe te passen is, zullen de fasen en activiteiten beschreven moet worden die in een datamigratietraject aanwezig zijn.

De invloed van de aanpak van de datamigratie op de kosten dient in kaart gebracht te worden. Hoe wordt de complexiteit en de productiviteit in een datamigratietraject gemeten?

Dit leidt tot de volgende deelvragen:

- Uit welke onderdelen bestaat een datamigratietraject?
- Wat is de invloed van de aanpak van de datamigratie op de kosten?
- Welke complexiteitsfactoren zitten in een datamigratietraject?
- Welke productiviteitsfactoren beïnvloeden een datamigratietraject?

Om de centrale vraag en de deelvragen te beantwoorden is de aanpak gevolgd zoals hieronder beschreven.

Om inzicht te krijgen in de materie van datamigratietrajecten, functiepuntanalyse, complexiteitsfactoren en productiviteitsfactoren is een literatuuronderzoek uitgevoerd.

Het literatuuronderzoek over datamigratie had als doelstelling om in de bekende ontwikkel- en beheersmethodieken een datamigratiemethode te identificeren. Daarnaast is literatuuronderzoek gedaan naar aspecten die tijdens datamigratietrajecten aan bod komen.

Er is uiteindelijk een datamigratieprocesmodel gebruikt waarin alle activiteiten die uitgevoerd moeten worden binnen een datamigratietraject beschreven zijn. Dit model is gebaseerd op best practices die opgedaan zijn tijdens het uitvoeren van datamigratietrajecten binnen het bedrijf waar ondergetekende werkzaam is.

Het literatuuronderzoek naar functiepuntanalyse had tot doel te bepalen of FPA gebruikt wordt om datamigratietrajecten te begroten. Hierover zijn in de geraadpleegde literatuur geen aanknopingspunten gevonden. Daarna is uitgewerkt hoe FPA toegepast en gebruikt kan worden om datamigratietrajecten te begroten. Vervolgens zijn de activiteiten van een datamigratietraject beschreven die niet begroot kunnen worden met FPA.

Als laatste deel van het literatuuronderzoek is onderzoek gedaan naar productiviteitsfactoren, die het aantal uren beïnvloeden in datamigratietrajecten. Naar complexiteit binnen de te ontwerpen, ontwikkelen en testen programmatuur is geen onderzoek gedaan. De reden hiervoor is dat complexiteit een vast onderdeel van functiepunten is dat de omvang vergroot.

Naast de literatuurstudie zijn interviews gehouden met acht datamigratie experts. Deze experts hebben zelf een op basis van best practices meerdere datamigratietrajecten doorlopen. Doel van de interviews was opsporen van knelpunten die in datamigratietrajecten voorkomen. Daarnaast het inventariseren van de manieren die gebruikt worden

om kostprijs en doorlooptijd te bepalen. De interviews hebben geleid tot meer inzicht in de problematiek.

Daarna zijn de activiteiten van het datamigratietraject onderzocht die met FPA begroot kunnen worden. Voor de activiteiten die niet middels FPA begroot kunnen worden is een methode gekozen waarmee wel begroot kan worden. De verwachting is dat tijdens een datamigratietraject niet altijd voldoende gegevens aanwezig zijn om een detailplanning te maken. Daarom zijn kengetallen opgesteld in functiepunten voor datamigratietrajecten op basis van een globale en een detailtelling. Voor een offertezoek van een telecombedrijf is via deze methode het aantal functiepunten bepaald.

Tenslotte is een proof of principle uitgevoerd door over een datamigratietraject, uitgevoerd bij een internationaal bedrijf, gegevens te verzamelen en te analyseren. In dit datamigratietraject zijn meerdere migratietools gebruikt en is maatwerkprogrammatuur ontwikkeld om de datamigratie uit te voeren. Deze analyse heeft onder andere aangetoond dat veel uren in een datamigratietraject besteed worden aan projectmanagement.

3 Datamigratie

In dit hoofdstuk wordt nader ingegaan op het fenomeen datamigraties bij het vervangen van het ene informatiesysteem door het andere. Er zijn veel verschillende typen migraties, zoals migreren van de versie van een database of tool, hardware, pakket, database, van een maatwerkpakket overstappen naar een standaard pakket, enz. In dit verslag is het onderwerp datamigratie van bedrijfsgegevens.

Een datamigratietraject is onderdeel van een migratietraject. In een datamigratietraject is de scope het overzetten van de data van het bronsysteem naar het doelsysteem. Tijdens een dergelijke verandering wordt binnen de organisatie gekeken naar producten, diensten, distributiekanaal, bedrijfsmodellen, procesmodellen, mensen, de organisatie en ICT. Dit geheel van veranderingen wordt meestal in termen van verandermanagement opgepakt. Dit is van belang omdat migratie méér is dan alleen datamigratie. In het gehele migratietraject wordt het nieuwe informatiesysteem ontwikkeld of ingericht en de organisatie en de processen ingericht om te werken met het nieuwe informatiesysteem. Het migratietraject zorgt onder andere voor de scholing van medewerkers, het afsluiten van contracten met leveranciers en het uifasieren van het bronsysteem.

Begonnen is met opstellen van een nieuwe omschrijving van de begrippen migratie en datamigratie. Hiervoor is eerst gezocht binnen algemeen geaccepteerde methoden van beheer en ontwikkeling van informatiesystemen. Aangezien in de bestudeerde literatuur geen bruikbare definitie voor datamigratie is gevonden, zijn de volgende (voorlopige) werkdefinities gehanteerd in het onderzoek.

Werkdefinitie migratie: migratie bestaat uit het proces en de gehanteerde technieken en hulpmiddelen om bestaande software systemen te vervangen door (nieuwe) systemen.

Werkdefinitie datamigratie: datamigratie is het proces om gegevens over te zetten van één of meerdere bronsystemen naar één of meerdere doelsystemen inclusief transformatie teneinde de gegevens geschikt te maken voor gebruik in de nieuwe systemen.

Dit onderzoek richt zich op het aspect datamigratie. Hieronder verstaan we het overzetten van gegevens vanuit één of meerdere informatiesystemen naar één of meerdere (nieuwe) informatiesystemen met als uitgangspunt dat het overzetten ook eenmalig is. De gegevens die gemigreerd dienen te worden, noemen we brongegevens. Brongegevens zijn meestal opgeslagen in de databases van de bronsystemen, maar het komt ook voor dat deze gegevens opgeslagen zijn in Excel-documenten, papieren dossiers en op diverse andere media en gegevensdragers.

3.1 Op zoek naar een procesbeschrijving van datamigratie

Tijdens het literatuuronderzoek is allereerst onderzocht of in de algemeen bekende methoden voor informatiesysteem ontwikkeling en beheer een procesbeschrijving bestaat om een datamigratie uit te voeren. De methoden die verder onderzocht worden zijn ASL, BiSL, ITIL, CMM, QSM en het vakgebied software engineering in het algemeen. Voor deze methoden is gekozen, omdat te verwachten valt dat deze methoden iets over datamigratie beschrijven.

3.1.1 *Application Services Library en Business Information Services Library*

In eerste instantie is onderzocht wat beschreven is binnen Applicatie Service Library (ASL) en Business Information Service Library (BiSL) over datamigraties (ASL, 2008). ASL en BiSL zijn ontwikkeld door PinkRoccade voor outsourcing van informatiesystemen, waarbij BiSL meer business en ASL meer IT georiënteerd is. Eerst is onderzoek gedaan naar ASL en daarna naar BiSL.

De reden om onderzoek te doen naar ASL is dat ASL ingaat op het beheren, onderhouden en vernieuwen van applicaties en databases. De laatste fase in de levenscyclus van een informatiesysteem is vaak een datamigratietraject om de bedrijfskritische gegevens over te brengen naar het nieuwe informatiesysteem.

ASL is een raamwerk waarin de processen van applicatiemanagement met elkaar in relatie worden gebracht. ASL heeft drie niveaus waarop het zich richt, en wel het richtinggevende, sturende en uitvoerende niveau.

Het richtinggevende niveau richt zich op de invalshoeken service en applicatie. We duiden die respectievelijk aan als Organization Cycle Management en Applications Cycle Management. Organization Cycle Management richt zich op processen voor de ontwikkeling van een toekomstvisie op de ICT-serviceorganisatie en de vertaling van die visie naar beleid ter vernieuwing van de ICT-serviceorganisatie. Applications Cycle Management richt zich op de processen die zorgen voor de vormgeving van een lange termijnstrategie voor de verschillende applicaties en het geheel van de informatievoorziening van een organisatie, in relatie tot het lange termijnbeleid van de organisatie.

Het sturende niveau bestaat uit de overkoepelende managementprocessen. Deze processen verzorgen de gezamenlijke aansturing van de uitvoerende processen voor services en applicaties. Zowel het richtinggevende als het uitvoerende niveau voeden de managementprocessen. Daarmee zijn de toekomstgerichtheid en de dagelijkse realiteit sterk en expliciet verankerd in deze processen.

Het uitvoerende niveau kent twee processen. Dit zijn beheer en onderhoud/vernieuwing van applicaties. Beheer van applicaties bevat de processen die zorgen voor de optimale inzet van de huidige applicaties ter ondersteuning van het bedrijfsproces met een minimum aan middelen en verstoring in de operatie. Onderhoud/vernieuwing van applicaties bevat de processen die applicaties aanpassen aan nieuwe wensen en eisen als gevolg van veranderingen in de organisatie en haar omgeving. In de gegevensmodellen, de programmatuur en de documentatie worden de noodzakelijke bijstellingen aangebracht.

BiSL is een procesmodel voor professionalisering van functioneel beheer en informatie-management. In het BiSL procesmodel worden de ICT en bedrijfsprocessen naast elkaar in kaart gebracht. Hierdoor ontstaat inzicht in alle processen en de relaties tussen de processen. Het biedt aanknopingspunten voor verbetering van de processen, onder meer via best practices en het verschaft een uniforme terminologie.

BiSL wordt gebruikt aan de bedrijfskant om aan de IT leverancier een verzoek tot levering van IT diensten te doen. ASL wordt door de IT leverancier gebruikt om de gevraagde ICT diensten te leveren.

Concluderend kan gesteld worden dat ASL en BiSL geen beschrijving of aanzet bevatten van een datamigratiemethode.

Tijdens het onderzoek of ASL of BiSL een methode bevatten voor datamigratietrajecten is contact opgenomen met een van de grondleggers van ASL, R. van der Pols. Deze heeft per e-mail (Pols, 2008) de boven beschreven conclusie over de toepasbaarheid van ASL en BiSL bevestigd.

3.1.2 Information Technology Infrastructure Library

Information Technology Infrastructure Library (ITIL) is een referentiekader voor het inrichten van ICT beheerprocessen binnen een organisatie. ITIL wordt gebruikt voor technisch beheer en kan worden beschouwd als de beschrijving van een op exploitatie van IT systemen toegespitst kwaliteitssysteem. De methode is in Engeland door de Central Computer and Telecommunications Agency ontwikkeld. ITIL is een procesgeoriënteerde methode en beschrijft de processen configuratiebeheer, helpdesk, probleembeheer, wijzigingsbeheer en programmatuurbeheer en distributie (<http://www.itil-officialsite.com>). Binnen ITIL is niets beschreven over datamigratie.

3.1.3 Capability Maturity Model Integration

Het Capability Maturity Model Integration (CMMI) is onderzocht, omdat CMMI voor allerlei processen beschrijft waar softwareontwikkelorganisaties aan moeten voldoen om een bepaalde mate van volwassenheid te hebben. Indien CMMI dit beschrijft voor datamigraties kan dit als houvast dienen voor het onderzoek. CMMI is een op best practices gebaseerde procesverbetermethode en bevat op allerlei ICT-gebieden modellen en een aanpak om op dat gebied projecten uit te voeren met meetinstrumenten om de volwassenheid van de methode te meten (<http://www.sei.cmu.edu/cmmi/>). CMMI is ontwikkeld door het Software Engineering Institute van de Carnegie Mellon University in Pittsburgh. CMMI beschrijft volwassenheidsmodellen in de Software Engineering, Systems Engineering, Integrated Product and Process Development en Supplier Sourcing. CMMI-SW heeft 22 Key Process Areas. Deze gebieden zijn weer onderverdeeld in een aantal activiteiten en doelen. Om te voldoen aan een bepaalde mate van volwassenheid moeten deze activiteiten en doelen geïmplementeerd zijn in de organisatie. Het niveau van implementatie bepaalt de volwassenheid en daarmee het CMMI niveau. Ook CMMI biedt geen geformaliseerde aanpak betreffende het aspect datamigratie.

3.1.4 Software engineering

Het vakgebied software engineering is te groot om helemaal in kaart te brengen daarom is in het literatuuronderzoek een toonaangevend boek in de software engineering (Pressman, 2001) bestudeerd. In dit boek is geen methode om datamigratietrajecten uit te voeren gevonden. Vanwege de beschikbare tijd voor het afstuderen is er voor gekozen om niet meer dan één boek te bestuderen.

3.1.5 Quantitative Software Management

Als laatste is gekeken of QSM behulpzaam kan zijn bij voorspellen van kwaliteit of gebruikt kan worden om de doorlooptijd te bepalen bij datamigratieprojecten. Middels QSM (Putnam, Myers, 1992) kan de kwaliteit van een project worden voorspeld.

Binnen QSM is geen methodiek specifiek voor datamigratietrajecten gevonden. Indien functiepunten bepaald worden van een datamigratietraject dan kan QSM behulpzaam zijn om voorspellingen te doen over doorlooptijd en kwaliteit op basis van functiepunten.

3.1.6 Conclusie

Tijdens de literatuurstudie van algemeen bekende methodieken zijn géén aanknopingspunten gevonden voor de aanpak van een datamigratietraject. Omdat migratie en de daarbij horende datamigratie gezien kunnen worden als een onderdeel van de lifecycle van een applicatie ligt deze uitkomst niet in de lijn met de aanvankelijke verwachtingen.

Voor het onderzoek is het van belang de activiteiten te onderkennen die uitgevoerd worden tijdens een datamigratietraject. Door het werken met een gestructureerde datamigratiemethode is het makkelijker gegevens te vergelijken over verschillende projecten en ontstaat een coherent begrippenkader. In de volgende paragrafen wordt een aantal onderwerpen besproken die tijdens de literatuurstudie naar datamigratie naar boven zijn gekomen.

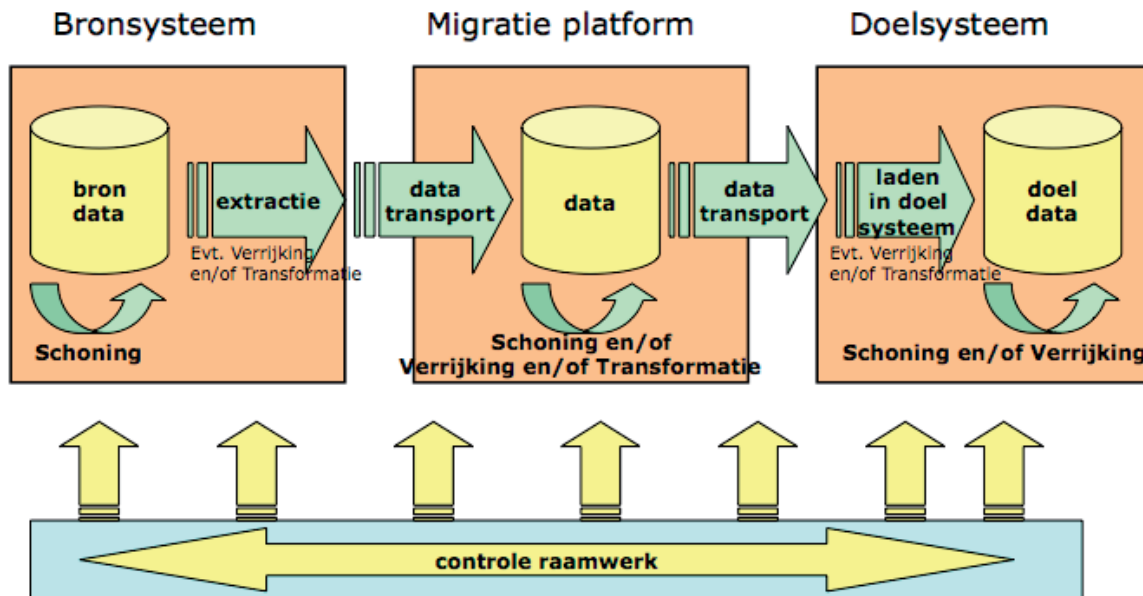
3.2 Datamigratieproces

Het datamigratieproces wordt meestal aangepakt volgens de Extract Transform en Load (ETL) methode (Rahm en Hai Do, 2008). In deze methode wordt eerst bepaald welke data uit het bronsysteem worden gehaald en daarna hoe deze data uit het bronsysteem worden gehaald. Dit proces wordt extractie genoemd. Deze data worden vervolgens opgeslagen in een hulpdatabase. Een ander woord voor hulpdatabase is migratieplatform. Daarna wordt bepaald welke bewerkingen op data nodig zijn. Het veranderen van de data, zodat de data voldoen aan het formaat van het doelsysteem, wordt transformatie genoemd. Vervolgens wordt de transformatie in de hulpdatabase doorgevoerd en aldaar getest. Uiteindelijk worden de getransformeerde data uit de hulpdatabase geladen in het doelsysteem.

Er zijn drie begrippen waaraan de te migreren data onderhevig zijn.

- Transformatie impliceert het omzetten van gegevens in een andere waarde en/of het omzetten van gegevens in een andere structuur (notatievorm).
- Schoning betreft het proces voor het detecteren en corrigeren of verwijderen van incorrecte of overbodige gegevens (fouten, vervuiling, inconsistenties, historie) teneinde de kwaliteit van de gegevens te verbeteren.
- Verrijking betreft het toevoegen van ontbrekende gegevens die nodig zijn voor een correcte werking van het nieuwe systeem.

In figuur 1 is in het datamigratietraject in een processchema weergegeven.



Figuur 1. ETL proces (Rahm en Hai Do, 2008)

Transformatie

Transformatie is het omzetten van gegevens naar het juiste formaat om ingelezen te kunnen worden in het doelsysteem. Allerlei conflicten in de structuur en inhoud van gegevens wordt tijdens transformatie gecorrigeerd. Dit gebeurt in een migratieplatform waar op basis van de ingelezen brontabellen de data getransformeerd worden om te kunnen verwerken in het doelsysteem. In dit migratieplatform komen gegevens uit meerdere tabellen bij elkaar, zodat voor een logische gegevensverzameling alle gegevens voorradig zijn om de data klaar te maken voor datatransport.

Verrijking

Dataverrijking is het toevoegen van gegevens tijdens de datamigratie, zodat het doelsysteem gevuld kan worden. Een voorbeeld van verrijking is dat de provincie in het oude systeem niet bijgehouden wordt en dat tijdens het migratieproces de provincie bepaald wordt, zodat dit in het doelsysteem ingelezen kan worden.

De reden om dataverrijking te doen is meestal dat het doelsysteem gegevens nodig heeft die niet opgeslagen zijn in het bronsysteem. Vaak kan uit andere gegevens de ontbrekende data worden bepaald, maar soms is handmatige actie nodig om gegevens uit dossiers, die niet in de bestaande database zitten, toe te voegen.

Meestal wordt dit toevoegen van gegevens in een hulpdatabase gedaan als het bronsysteem ophoudt te bestaan. Indien het bronsysteem nog gebruikt wordt na de datamigratie dan is het raadzaam om in het bronsysteem de verrijking door te voeren. Dit heeft als voordeel dat het transformeren makkelijker is en dat bij mutaties in het bronsysteem ook de verrijkte gegevens indien nodig aangepast kunnen worden.

Schoning

Gegevens die niet voldoen aan de kwaliteitseisen van data in het doelsysteem worden als uitval betiteld. Deze uitval dient minimaal te zijn. Indien de datakwaliteit niet van voldoende niveau is vindt schoning plaats. Schoning is het aanpassen van gegevens die niet binnen de domeinwaarde liggen te herstellen, of gegevens die niet compleet zijn

aan te vullen, of gegevens die niet in de juiste vorm opgemaakt zijn in de juiste vorm te zetten (Rahm en Hai Do, 2008). Ontdubbelen van gegevens, het oplossen van tegenstrijdigheden in gegevens is ook een onderdeel van schonen. Aandachtspunten zijn domeincontrole en overbodige historie (data die niet meer relevant zijn voor de bedrijfsvoering en wet-/en regelgeving en referentiële integriteit).

Het komt regelmatig voor dat tijdens de testfase een groter percentage uitval optreedt dan aanvankelijk gedacht. Gebruikers schatten de kwaliteit van hun gegevens vaak hoger in dan de kwaliteit werkelijkheid is. Dit veroorzaakt een iteratief proces waarin de ontwerp-, realisatie- en testfase een aantal maal wordt doorlopen om tot een acceptabel uitvalpercentage te komen. Een hoger dan geraamd uitvalpercentage is een indicatie dat in het bepalen van datakwaliteit bepaalde aspecten onderbelicht zijn gebleken. Dit leidt tot problemen die gedurende volgende fases van het datamigratietraject van significante betekenis kunnen zijn en daarmee bepalend zijn voor een verder succesvol verloop.

Controle raamwerk

Een belangrijk onderdeel van een datamigratietraject is de rapportage over de data-migratie, aan het standaard ETL mechanisme voegen we een controle raamwerk toe om te voorzien in rapportages. Indien tijdens het datamigratietraject verschillen in bron- en doelsysteem optreden dan dienen die verschillen verklaard te worden, of - beter nog - van tevoren voorspeld te worden.

Het controle raamwerk is de complete verzameling van alle rapportages en wordt gebruikt om verantwoording af te leggen en inzicht te geven aan de organisatie (en de accountants) in juistheid, volledigheid en traceerbaarheid van de datamigratie.

Met volledigheid wordt bedoeld dat alle gegevens die gemigreerd moeten worden ook gemigreerd zijn. Dit is onderdeel van de verantwoording naar accountants en naar de beheerder van de gegevens dat tijdens de datamigratie alles meegenomen is. Hiervoor worden controleoverzichten en controletellingen ontwikkeld en uitgevoerd tijdens het datamigratietraject. Deze controles worden uitgevoerd om te valideren dat er geen gegevens verloren of verminkt raken tijdens de datamigratie. Voorbeelden zijn het aantal rijen in een tabel voor en na de datamigratie, het bedrag aan uitstaande orders moet voor en na de datamigratie gelijk zijn, enz. Ook geldt dat de betekenis van de data niet veranderd mag zijn door transformatie van gegevens. Hiermee wordt bijvoorbeeld bedoeld dat indien in het nieuwe systeem minder statussen bestaan dan in het oude systeem door transformatie geen data een onjuiste status mogen hebben.

Juistheid geeft aan dat de datamigratie inhoudelijk juist is uitgevoerd. De waarde van het veld in het oude systeem komt overeen met de waarde van het veld in het nieuwe systeem. In bijna elke datamigratie moeten gegevens worden bepaald via transformatieregels. De uitkomst van deze formules is ook onderdeel van de controles op juistheid. Het bepalen van juistheid van de datamigratie is ook onderdeel van de verantwoording naar de accountants, management en eindgebruikers. Voorbeelden hiervan zijn dat een order na datamigratie dezelfde status heeft of dat de tarieven niet veranderd zijn. Referentiële integriteit is ook onderdeel van juistheid. Hiermee wordt bedoeld dat dezelfde adresgegevens gekoppeld zijn aan dezelfde klant.

Traceerbaarheid is het achterhalen waar gegevens vandaan komen en hoe gegevens gebruikt zijn. Dit betekent dat voor elk veld dat in het doelsysteem gevuld is bekend moet zijn op basis van welke gegevens uit het bronsysteem deze waarde bepaald is. In

het controle raamwerk wordt voor een aantal velden de herkomst van gegevens gerapporteerd.

3.3 Datamigratiefasen

Elk traject wordt opgesplitst in fasen die achtereenvolgens of parallel worden doorlopen. Voor een datamigratietraject zal dit niet anders zijn. In de meeste trajecten wordt begonnen met vooronderzoek en definitie, in deze fase wordt de scope vastgesteld en plan van aanpak opgesteld. Daarna worden diverse maatwerkproducten ontwikkeld of een standaardtool geselecteerd (Carreira en Galhardas, 2004) en ingericht om een datamigratietraject uit te voeren. Hier worden de fasen ontwerp, realisatie en test doorlopen. Indien alle tests klaar zijn kunnen proef datamigraties uitgevoerd worden en daarna de daadwerkelijke datamigratie. Het komt regelmatig voor dat bepaalde fasen iteratief worden doorlopen, denk aan meerdere testruns. Over het gehele datamigratietraject vindt projectmanagement en changemanagement plaats om te sturen op resources, deadlines, het nemen van beslissingen en omgaan met wijzigingen tijdens het datamigratietraject, het managen van risico's.

3.4 Scope bepaling

Wat bij een datamigratietraject opgepakt moet worden, zijn alle activiteiten die binnen de scope van het datamigratietraject vallen. De scope is het vullen van het doelsysteem op basis van het bronsysteem. Het bepalen van de scope gebeurt zo vroeg mogelijk in het datamigratietraject door het in kaart brengen van de wijze waarop het doelsysteem gevuld moet worden met behulp van de gegevens van het huidige bronsystemen. Het vullen van het doelsysteem is leidend in de scope bepaling in een datamigratietraject. Dit betekent dat alleen gegevens die nodig zijn voor het vullen van het doelsysteem meegenomen worden in de datamigratie. Hierbij kan bijvoorbeeld gedacht worden om alleen gegevens van klanten die het laatste jaar een bestelling geplaatst hebben mee te nemen in het klantenbestand.

Nadat bepaald is welke doelgegevensverzamelingen gevuld moeten worden, zal per te vullen doelgegevensverzameling een keuze worden gemaakt tussen geautomatiseerd of handmatig migreren. Deze keuze hangt af van foutgevoeligheid, hoeveelheid te migreren gegevens, kosten en opbrengsten en moeilijkheidsgraad van een geautomatiseerde datamigratie. Tevens wordt bepaald welke systemen en processen op basis van onderlinge samenhang onderdeel van de datamigratie zijn. Per systeem zal nog een uitsplitsing gemaakt worden of het hele of een deel van het informatiesysteem wordt gemigreerd. Vaak wordt niet het hele systeem gemigreerd, maar worden alleen gegevens van openstaande processen of bepaalde boekjaren of alleen klanten die nog actief zijn gemigreerd. Mogelijk bevat het nieuwe systeem andere referentiegegevens. Daarnaast wordt beslist of historische gegevens gemigreerd worden (Kelly en Nelms, 2003). Voor historische gegevens en persoonsgegevens geldt wet- en regelgeving, bijvoorbeeld de privacywetgeving. Daarnaast is het vanwege fiscale wetgeving vaak een verplichting om van een aantal kalenderjaren financiële gegevens te bewaren.

Historische gegevens zijn gegevens die op dit moment niet meer actueel zijn. Voorbeelden zijn afgesloten boekjaren, klanten of leveranciers waar geen zaken meer mee gedaan worden, enz.

Referentiegegevens zijn voorgedefinieerde waarden (uit een tabel) voor een bepaald gegevenselement. Voorbeelden zijn bijvoorbeeld artikelen, tarieven, pakketten, provincies, categorieën, enz.

3.5 Migratiestrategie

In de definitiefase wordt de migratiestrategie bepaald, d.w.z. in hoeveel opleveringen wordt de datamigratie doorgevoerd. De opties zijn alles in één keer, gefaseerd migreren of parallel.

Een migratie waarin alle gegevens in één keer overgezet worden wordt ook wel een big bang migratie genoemd. Dit betekent dat er een migratiemoment (een datum) gekozen wordt waarin de gehele migratie uitgevoerd wordt. In de meeste gevallen wordt op een vrijdag aan het einde van de werkdag het bronsysteem afgesloten en de migratie en datamigratie uitgevoerd. Op maandagochtend is dan alleen het oude systeem nog beschikbaar.

Gefaseerd migreren betekent dat onderdelen van het bronsysteem geleidelijk in gebruik genomen worden en dat de migratie over een langere periode uitgesmeerd wordt. Tijdens deze geleidelijke ingebruikname van het doelsysteem wordt ook nog in delen van de bronsystemen gewerkt.

Naast gefaseerde en bigbang migratie bestaat de mogelijkheid om parallel te werken in bron- en doelsysteem. De transacties die in het doelsysteem uitgevoerd worden zullen voor een bepaalde periode ook in het bronsysteem uitgevoerd worden. Men werkt in een bepaalde periode met twee systemen naast elkaar. Dit geeft de mogelijkheid om te gegevens en status van beide systemen te vergelijken. Dit geeft inzicht of het doelsysteem zich op een zelfde wijze gedraagt als het bronsysteem.

In figuur 2 zijn voor elke strategie de voor- en nadelen weergegeven. Dit figuur is tot stand gekomen door kennis opgebouwd uit diverse migratietrajecten.

<p>Big bang</p> <p>Voordelen</p> <ul style="list-style-type: none"> + beperkte belasting uitvoeringsorganisatie + populatie in één systeem + vereenvoudige interface strategie 	<p>Nadelen</p> <ul style="list-style-type: none"> - geen bijstellingsmomenten (proces, systeem, organisatie) - referentie / ervaringskader ontbreekt - hoog afbreukrisico / businessimpact
<p>Gefaseerd</p> <p>Voordelen</p> <ul style="list-style-type: none"> + leercurve, ervaring opdoen met nieuwe systeem, klein beginnen (afbreukrisico) + bijstellingsmomenten (proces, systeem, organisatie) 	<p>Nadelen</p> <ul style="list-style-type: none"> - complexiteit: twee systemen naast elkaar - langere instandhouding oude systeem (kostenaspect) - dubbele interfaces
<p>Parallel (“oud naast nieuw”)</p> <p>Voordelen</p> <ul style="list-style-type: none"> + geen conversie (organisatie belasting / kosten) + minimale consequenties bestaande populatie 	<p>Nadelen</p> <ul style="list-style-type: none"> - twee systemen naast elkaar / verhoogde complexiteit in gebruik - langdurige instandhouding oude systeem - dubbelen interfaces

Figuur 2. Praktijkvoorbeeld van de voor- en nadelen van datamigratiestrategieën

In principe wordt de migratiestrategie bepaald door de business impact op de bedrijfsvoering voor de huidige en toekomstige situatie te vergelijken. Van belang is of de organisatie voldoende adaptief vermogen heeft om twee systemen te kunnen onderhouden. Indien dit het geval is dan behoort schaduw draaien en gefaseerde overgang tot de mogelijke scenario's. Het synchronisatiemoment bepaalt wanneer de datamigratie eenduidig vast te stellen is, wanneer het bronsysteem afgesloten wordt of wanneer tijdelijke systemen en processen gebruikt worden.

3.6 Interfaces

Indien het te migreren systeem koppelingen heeft naar andere systemen dan moet ontworpen worden hoe die interfaces met andere systemen operationeel blijven tijdens en na het datamigratietraject. Mogelijk zijn tijdens de datamigratie gegevens nodig van een ander informatiesysteem en dient er een interface gelegd te worden naar dat systeem om gegevens op te halen.

Er zijn drie typen interfaces te onderkennen:

- Interfaces die gerelateerd zijn aan het bronsysteem
- Interfaces die gerelateerd zijn aan het doelsysteem
- Tijdelijke interfaces die de brug slaan tussen het bronsysteem, het migratieplatform en het doelsysteem.

Het bronsysteem kent interfaces van en naar het aan het bronsysteem gerelateerde systemen. Door het feit dat op een bepaald moment de datamigratie wordt ingezet, impliceert dit mogelijke aanpassingen op deze interfaces. Implicaties kunnen ontstaan door het feit dat data niet, onvolledig en/of veranderd worden aangeleverd. Zowel op procesniveau als op systeemniveau dienen hier expliciete afspraken over gemaakt te worden.

Het doelsysteem kent interfaces naar de aan het doelsysteem gerelateerde systemen. Betreffende interfaces dienen aan onderzoek onderhevig te zijn voor wat betreft de impact van de datamigratie. Als voorbeeld te noemen: als door de datamigratie een groter volume aan data wordt geleverd, zal onderzocht dienen te worden of de interfaces vanuit systeemcapaciteit dit volume aankunnen.

Een voorbeeld van tijdelijke interfaces vanuit architectuur perspectief in een datamigratietraject is het migratieplatform in het ETL proces. Een platform waar vanuit het bronsysteem data worden ontvangen, verrijkt en opgeschoond en vervolgens al dan wel of niet geautomatiseerd worden doorgegeven aan het doelsysteem. Deze interfaces, en daarmee de aanleveringen, dienen beschreven te worden en het tijdelijke beheer dient te worden ingericht.

3.7 Infrastructuur

Als sprake is van een (deels) geautomatiseerde datamigratie, maakt de noodzakelijke infrastructuur (hardware, systeemsoftware) een significant onderdeel uit van het datamigratietraject.

Indien sprake is van een geautomatiseerde datamigratie, vraagt dit om een tijdelijk migratieplatform. Dit zal ingericht moeten worden en er dienen afspraken gemaakt te worden over het tijdelijke beheer van dit migratieplatform. Hier speelt het aspect performance en capaciteitsmanagement een belangrijke rol. Het tijdelijke platform moet namelijk over voldoende capaciteit beschikken om de te migreren data te kunnen verwerken en tijdelijk op te kunnen slaan.

Een belangrijk onderdeel van de infrastructuur is het opbouwen en onderhouden van de testomgevingen. Dit valt uit te splitsen in datamigratie en migratie testomgevingen. De datamigratie testomgeving bevat een productie kopie van alle te migreren doelsystemen en alleen de interfaces die nodig zijn in de datamigratie, plus het migratieplatform en een kopie van de doelsystemen waarnaar toe gemigreerd wordt. In de migratie testomgeving, vaak de gebruikers acceptatie testomgeving genoemd, is de nieuwe systeemarchitectuur klaar gezet zoals het nieuwe systeem met interfaces na de migratie uitziet. Plus een kopie productie omgeving zoals het systeem er nu uitziet om verschillen in data op te sporen.

Indien een gefaseerde aanpak van de migratie of datamigratie gekozen is als strategie dan zal voor elke fase een testomgeving voor migratie en datamigratie aanwezig moeten zijn.

In de testomgeving voor datamigratie wordt getest of de data van het oude naar het nieuwe systeem gemigreerd kunnen worden. In de testomgeving voor migratie wordt getest of met het nieuwe systeem en interfaces de bedrijfsvoering uitgevoerd kan worden.

3.8 Conclusie

In de onderzochte standaardmethodieken ASL, BiSL, ITIL, CMMi, Software Engineering en QSM is in de bestudeerde literatuur géén aandacht voor datamigratie gevonden. Dit is onverwacht aangezien datamigratie als onderdeel van de lifecycle van een applicatie gezien kan worden. Wel zijn in de literatuur aspecten van datamigratie gevonden en deze aspecten moeten terug te vinden zijn in een datamigratiemethode.

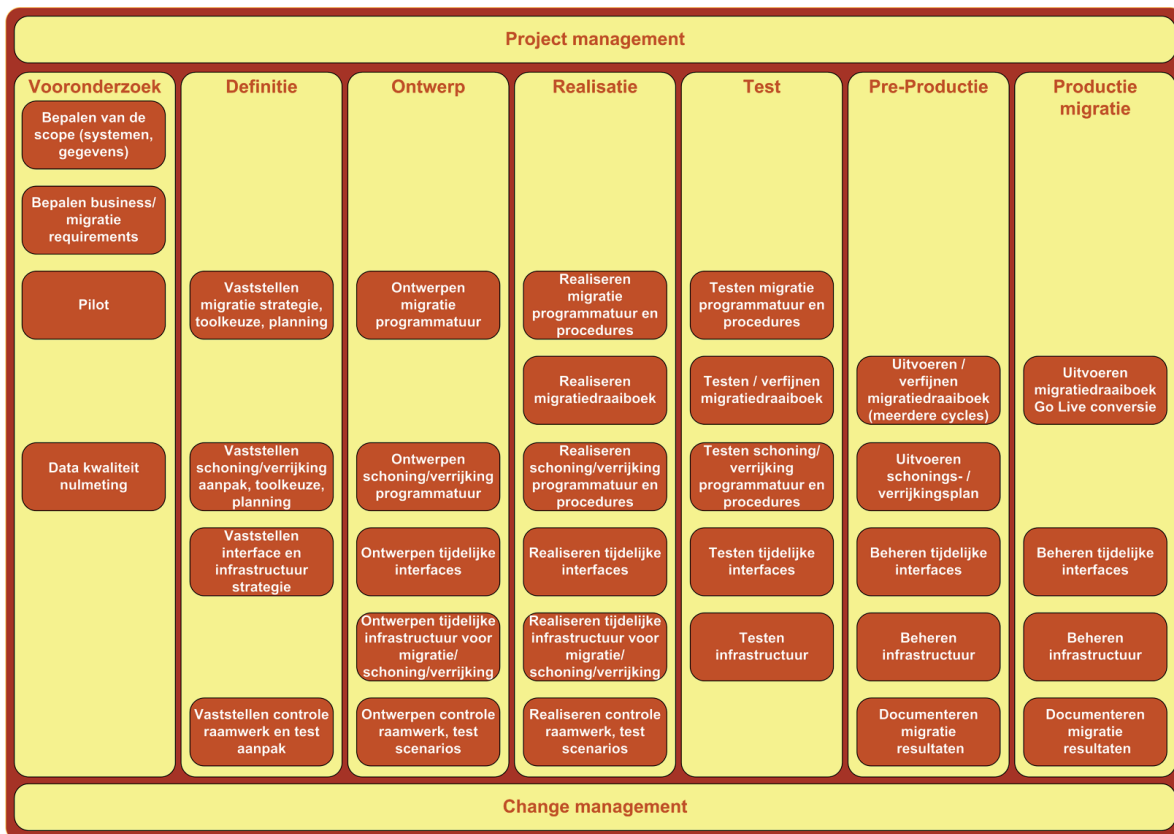
Binnen Mediaan/abs is in een werkgroep een aanpak voor datamigratietrajecten ontwikkeld op basis van best practices vanuit de praktijk. Mediaan/abs is een automatiseringsbedrijf in Heerlen dat gedegen ervaring heeft in het uitvoeren van datamigratietrajecten. In het verdere onderzoek wordt deze door Mediaan ontwikkelde aanpak gebruikt om datamigratietrajecten in kaders te plaatsen. In het volgende hoofdstuk wordt deze methode kort beschreven.

4 Mediaan Migratie Methode

Mediaan/abs hanteert voor het definiëren en uitvoeren van datamigratietrajecten de Mediaan Migratie Methode (MMM). Deze methode is ontstaan door de kennis en ervaring opgedaan uit een groot aantal datamigratietrajecten samen te brengen in een activiteitenmodel (figuur 3). Een werkgroep bestaande uit een zestal Mediaan data-migratie experts heeft deze methode ontwikkeld en getoetst in meerdere projecten. De naamgeving van de activiteiten is tijdens het onderzoek definitief gemaakt.

4.1 Methode

De methode is gebaseerd op het watervalmodel, omdat in elke fase meer details van de datamigratie bekend worden. Gedurende het gehele datamigratietraject wordt aan project en changemanagement gedaan om het traject beheersbaar te maken.



Figuur 3. Mediaan Migratie Methode

De fasen worden van links naar rechts opvolgend doorlopen. Tijdens de fase vooronderzoek wordt de startsituatie in kaart gebracht van de te migreren data, worden de eisen van de business vastgelegd, wordt een pilot gedaan om te valideren of de gekozen aanpak uitvoerbaar is en de kwaliteit van de te migreren data bepaald. Pas als het vooronderzoek is afgerond zal aan de fase definitie worden begonnen. Het volledig afsluiten van een fase is de preconditionie voor het opstarten van de volgende fase. In de definitiefase wordt geïnventariseerd wat nodig is om het datamigratietraject uit te voeren en worden afspraken tussen de betrokken partijen vastgelegd. Nadat vastgesteld is wat ontworpen moet worden begint de ontwerpfasen. In deze fase

worden ontworpen gemaakt voor migratie en schoningsprogrammatuur, indien nodig voor tijdelijke interfaces, tijdelijke infrastructuur en het controle raamwerk. In de volgende twee fasen wordt programmatuur ontwikkeld en getest, infrastructuur en interfaces aangelegd en worden rapportagemiddelen gerealiseerd en gecontroleerd. Als al deze activiteiten uitgevoerd zijn kan de fase pre-productie beginnen, dit is het proefdraaien van de datamigratie. Een proefmigratie kan leiden tot bijstelling van het ontwerp. Meestal worden vier tot zes proefmigraties uitgevoerd voordat gestart wordt met de productiemigratie. Binnen de fasen worden activiteiten iteratief doorlopen, omdat tijdens de proefmigraties allerlei bevindingen naar boven komen die leiden tot aanpassing van specificaties, opnieuw bouwen, testen en opnieuw draaien van een proefmigratie. Het einddoel van het hele datamigratietraject is de productie datamigratie. In deze fase worden de daadwerkelijke data van bron- naar doelsysteem verplaatst en rapportages opgeleverd ter verantwoording.

In het model zijn diverse aspecten van datamigratie te onderkennen. Deze aspecten bevatten logische vervolgactiviteiten in elke fase van het MMM. Deze opsplitsing maakt het mogelijk om met verschillende teams aan een datamigratietraject te werken. In figuur 4 is het model op basis van die aspecten weergegeven.

Project management							
	Vooronderzoek	Definitie	Ontwerp	Realisatie	Test	Pre-Productie	Productie migratie
	Bepalen van de scope (systemen, gegevens) Bepalen business/migratie requirements						
Tools	Pilot	Vaststellen migratie strategie, toolkeuze, planning	Ontwerpen migratie programmatuur	Realiseren migratie programmatuur en procedures	Testen migratie programmatuur en procedures		
Procedures				Realiseren migratiedraaiboek	Testen / verfijnen migratiedraaiboek	Uitvoeren / verfijnen migratiedraaiboek (meerdere cycles)	Uitvoeren migratiedraaiboek Go Live conversie
Opschonen	Data kwaliteit nulmeting	Vaststellen schoning/verrijking aanpak, toolkeuze, planning	Ontwerpen schoning/verrijking programmatuur	Realiseren schoning/verrijking programmatuur en procedures	Testen schoning/verrijking programmatuur en procedures	Uitvoeren schonings- / verrijgingsplan	
Tijdelijke interfaces		Vaststellen interface en infrastructuur strategie	Ontwerpen tijdelijke interfaces	Realiseren tijdelijke interfaces	Testen tijdelijke interfaces	Beheren tijdelijke interfaces	Beheren tijdelijke interfaces
Infrastructuur			Ontwerpen tijdelijke infrastructuur voor migratie/ schoning/verrijking	Realiseren tijdelijke infrastructuur voor migratie/ schoning/verrijking	Testen infrastructuur	Beheren infrastructuur	Beheren infrastructuur
Rapportage		Vaststellen controle raamwerk en test aanpak	Ontwerpen controle raamwerk, test scenarios	Realiseren controle raamwerk, test scenarios		Documenteren migratie resultaten	Documenteren migratie resultaten
Change management							

Figuur 4. Mediaan Migratie Methode op basis van aspecten

4.2 Aspecten van datamigratie

Project en changemanagement

In elk project is aansturing, management, nodig. Deze activiteit is nodig over alle fasen heen en gedurende het hele traject. Uit praktijkervaring is gebleken dat er veel veranderingen in het datamigratietrajecten ontstaan, doordat tijdens datamigratie processen en applicaties voor de nieuwe situatie nog volop in beweging zijn. Hierdoor zullen regelmatig wijzigingen voorkomen in scope, requirements en migratieregels. Om deze verande-

ringen te managen wordt changemanagement ingericht. Activiteiten die in de MMM uitgevoerd in projectmanagement zijn onder andere opstellen projectplan inclusief resourceplanning en risicoanalyse en opstellen van een voortgangsrapportage. In de activiteit changemanagement worden veranderingen in uitgangspunten, scope en requirements inzichtelijk gemaakt, gerapporteerd en afgehandeld.

Tools en procedures

Deze activiteiten hebben tot doel tools te selecteren of te ontwikkelen waarmee de datamigratie uitgevoerd gaat worden. Deze tools moeten ontwikkeld worden indien het nieuwbouw betreft en ingericht worden indien het bestaande tools betreft. In de ontwerpfase wordt bepaald hoe ingericht of ontwikkeld moet worden. Deze tools hebben tot doel het proces van analyseren van data te ondersteunen, teneinde de data geschikt te maken voor extractie, transformatie en laden van data plus valideren of data juist aankomen in het doelsysteem. In de realisatiefase wordt hetgeen wat ontworpen is gerealiseerd en in de testfase wordt ontwerp tegen realisatie getoetst en daarna getoetst of aan de requirements voldaan is.

Opschonen

Op basis van de specificaties, hoe het doelsysteem wordt gevuld, wordt via een nulmeting een schatting gemaakt van de vervuiling en daarmee de kwaliteit van de bron-systemen. Voor het opschonen wordt een plan van aanpak en planning gemaakt. In het opschoontraject bestaat ook weer de keuze om zelf tools te ontwerpen, realiseren en testen of gebruik te maken van standaardpakketten.

Interfaces en Infrastructuur

In deze fases wordt een interfaces en infrastructuur plan opgesteld. De interfaces tussen verschillende systemen en deelsystemen worden in kaart gebracht. Interfaces zijn nodig om toegang te krijgen tot gegevens voor datamigratie. Tijdens gefaseerde datamigraties komt het voor dat er nog extra toegang nodig is om tijdelijk op een andere manier extra gegevens te verwerken of raadplegen. Tevens worden testomgevingen ingericht om de datamigratie te testen en wordt aangesloten bij de testomgevingen voor datamigratie.

Rapportages

Over de gehele looptijd van het project zullen rapportages opgesteld worden. Deze rapportages zijn de rapportages over de datamigratie. In deze rapportages wordt inzicht gegeven over juistheid, volledigheid en traceerbaarheid van de datamigratie. Daarnaast zullen ook rapportages over de kwaliteit van geleverde producten en diensten door middel van kwaliteitsreviews opgeleverd worden.

4.3 Conclusie

De bruikbaarheid van het gepresenteerde model is aangetoond, doordat het in ongeveer dertig datamigratietrajecten met succes gebruikt is. Daarnaast is het model gebruikt in misschien wel het grootste datamigratietraject dat ooit in Nederland uitgevoerd zal worden. Namelijk het samen gaan van alle uitkeringsinstanties in het UVW als gevolg van de inwerkingtreding van de wet SUWI.

Aan dit model wordt het onderzoek opgehangen en in de volgende hoofdstukken zal dit model als uitgangspunt dienen. Nu een datamigratiemodel gekozen is kan binnen dit

model bepaald worden binnen welke activiteiten functiepuntanalyse mogelijk is. In bijlage II wordt dit model verder beschreven. FPA wordt niet gebruikt binnen de Mediaan Migratie Methode. De toevoeging van FPA is een bijdrage van dit onderzoek.

5 Functiepuntanalyse

Om de omvang van informatiesystemen te bepalen wordt in de praktijk functiepuntanalyse gebruikt. Functiepuntanalyse (FPA) is een algemeen geaccepteerde methode die zich op de functionele aspecten van informatiesystemen toespitst (Pressman, 2001). FPA gaat uit van (functionele) transacties en (logische) datafiles die relevant zijn voor de eindgebruiker. De meeteenheid bestaat uit functiepunten en wordt gebruikt om van een informatiesysteem de omvang te bepalen. Als de omvang van een informatiesysteem bekend is, kan op basis van productiviteitsgegevens voor het realiseren van een functiepunt de kostprijs bepaald worden. Het is een eenvoudig toepasbare, objectieve methode waarvoor weinig kennis van ontwerp en bouw van informatiesystemen nodig is (NESMA, 2004). In dit hoofdstuk wordt het ontstaan van FPA, afbakening en toepassing, een korte introductie, calculeren van kosten, een aantal kengetallen en recente ontwikkeling binnen FPA weergegeven. In komende hoofdstukken zal bekeken worden of het mogelijk is de omvang van datamigratietrajecten met behulp van FPA te bepalen.

5.1 Ontstaan

FPA is door Albrecht (1979) geïntroduceerd op basis van productiviteitsonderzoek van een groot aantal projecten (Abran en Robillard, 1996). De oorspronkelijke methode bestond uit vier functietypes, een verzameling van functionele gewichten en tien applicatieparameters voor een maximale correctiefactor van ongeveer 25%. In 1983 heeft Albrecht een tweede versie uitgebracht waarin hij het aantal functietypes heeft uitgebreid naar vijf, met per functietype drie gewichten. De vijf functietypes zijn, opvraagfunctie, invoerfunctie, uitvoerfunctie, interne logische gegevensverzameling en koppel gegevensverzameling. De gewichten zijn eenvoudig, gemiddeld en moeilijk. Het aantal applicatieparameters is uitgebreid naar veertien en de correctiefactor bedraagt maximaal 35%. De correctiefactoren zijn datacommunicatie, gedistribueerde gegevensverwerking, prestatiedoeleinden, transactievolume, online data-entry, gebruikersdoelmatigheid, online update, complexe interne verwerking, produceren van herbruikbare code, installatiegemak, bedieningsgemak, meerdere locaties/organisaties en flexibiliteit.

Toen FPA een geaccepteerde standaard aan het worden was, is de International Function Point User Group 'IFPUG' opgericht. In Nederland is vijf jaar later (1989) de 'NESMA' (toentertijd 'NEFPUG' genaamd) opgericht. In 1993 is de definitie van de functiepunt als functionele meeteenheid binnen ISO opgenomen. In 2004 zijn de richtlijnen NESMA 2.2 geaccepteerd door de ISO (ISO standaard ISO/IEC 24570) (NESMA, 2004).

5.2 Afbakening en toepassing

Zodra de functionele gebruikerswensen van het informatiesysteem op hoofdlijnen bekend zijn, kan FPA worden toegepast. FPA wordt vooral bij administratieve informatiesystemen toegepast.

De reden hiervan is dat FPA uitgaat van het perspectief van de gebruiker. Bij technische systemen speelt het grootste deel van de toepassing zich buiten het gezichtsveld van de gebruiker af. Bij technische informatiesystemen, zoals besturingssystemen wordt veelal

COSMIC toegepast (Vogelezang, 2005; Abran et al, 2001) hiervoor is FPA niet of minder geschikt. Bij COSMIC worden databewegingen gemeten. COSMIC heeft in tegenstelling tot FPA geen maximum aan omvang per te meten object, elke databeweging telt.

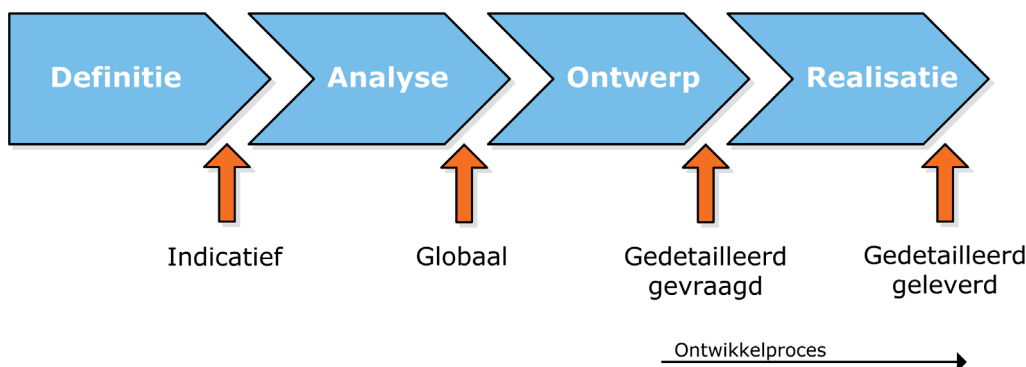
FPA wordt toegepast om allerhande typen projecten te budgetteren (projectomvang) en wordt toegepast op zowel nieuwbouw-, onderhoud- als testprojecten (NESMA, 2004). Het voordeel van FPA is dat het de mogelijkheid biedt om concreet en objectief te kunnen praten over de grootte van een te ontwikkelen/ onderhouden of te testen informatiesysteem. De toepassing van functiepunten biedt hulp bij scope afbakening van informatiesystemen en projecten. Daarnaast maakt FPA het mogelijk een beter beeld te krijgen van kostenschattting, budgettering, projectcontrole, communicatie over projecten en het managen van verwachtingen. Tevens maakt FPA het meten van de productiviteit, het meten van de kwaliteit van een informatiesysteem, het definiëren van een teststrategie in overeenstemming met het beoogde kwaliteitsniveau, het verbeteren van de kwaliteit van het systeemontwikkelingsproces en benchmarking mogelijk. Kortom projecten worden meetbaar waardoor onzekerheid gereduceerd wordt.

Tijdens de definitiefase, de informatieanalysefase en de functionele ontwerpfase wordt FPA toegepast om omvang te bepalen van het te realiseren informatiesysteem. Met FPA kan de benodigde inspanning voor elke fase in de systeemontwikkelingslevenscyclus worden geschat. Tijdens de operationele fase kan met FPA de operationele kosten van informatiesystemen worden geschat. Voor de operationele fase heet de meeteenheid: 'runtime functiepunten per jaar'. Op het moment dat een applicatie geïmplementeerd wordt kan met FPA worden bepaald hoeveel functionaliteit daadwerkelijk is geleverd. Met FPA kan ook worden aangegeven hoeveel functionaliteit getest is of getest dient te worden. De meeteenheid die tijdens testen gebruikt wordt is testfunctiepunten (Pol et al, 2002).

5.3 Introductie

Binnen een ontwikkelingsproces kunnen drie typen functiepuntentellingen worden onderscheiden: indicatieve functiepuntentelling, globale functiepuntentelling en gedetailleerde functiepuntentelling.

Voor elk type functiepuntentelling is een bepaalde mate van informatie nodig om de telling uit te voeren. In onderstaand schema (Figuur 5) worden de fasen van het ontwikkelingsproces getoond en welk type functiepuntentelling op basis van de dan beschikbare informatie uitgevoerd kan worden.



Figuur 5. Functiepuntmethode per fase van het ontwikkelingsproces

De verschillende typen functiepuntentellingen hebben elk een andere nauwkeurigheid (NESMA, 2004). Alle functiepunttelmethoden met hun marge van afwijking en benodigde specificaties is in Tabel 1 verzameld. Indicatieve functiepuntentelling is het minst nauwkeurig (marge 30-50%). Deze telling bepaalt alleen het aantal datafuncties en berekent de functiepuntentelling als volgt indien uitgegaan wordt van een conceptueel datamodel:

$$[\text{Indicatieve grootte}] = 35 \times \text{aantal \{interne logische gegevensverzameling entiteiten uit conceptueel model\}} + 15 \times \text{aantal \{koppeling gegevensverzameling entiteiten uit conceptueel model\}}.$$

Indien uitgegaan wordt van een datamodel in de derde normaal vorm dan geldt:

$$[\text{Indicatieve grootte}] = 25 \times \text{aantal \{interne logische gegevensverzameling entiteiten uit genormaliseerd model\}} + 15 \times \text{aantal \{koppeling gegevensverzameling entiteiten uit genormaliseerd model\}}.$$

Dit is de snelste methode, omdat alleen dataspecificaties noodzakelijk zijn. Deze methode is ook wel bekend als de "Hollandse schoolmethode".

De globale functiepuntentelling is nauwkeuriger dan de indicatieve methode (marge 15-25%). Deze telling bepaalt alle typen functies en hanteert voor de complexiteit een standaardwaarde, gemiddeld voor gebruikerstransacties en eenvoudig voor logische gegevensverzamelingen.

Deze methode is minder arbeidsintensief dan een gedetailleerde functiepuntentelling. Om een globale functiepuntentelling te doen zijn globale specificaties nodig waarin de gebruikersfuncties te onderkennen zijn. Bij een globale functiepuntentelling wordt voor complexiteit gemiddeld genomen.

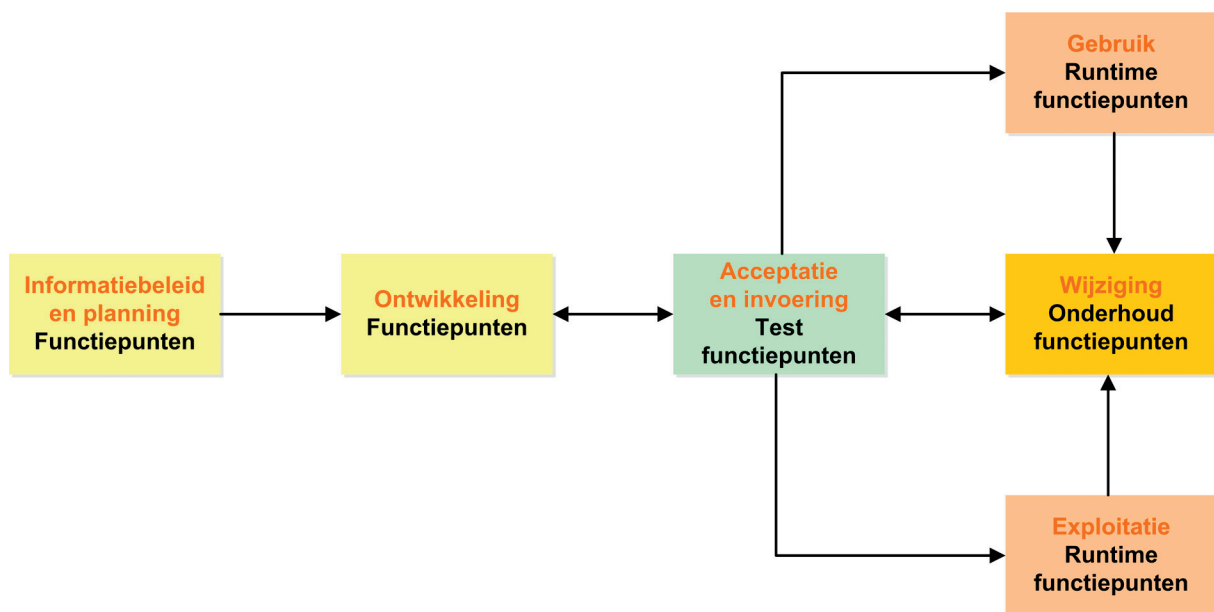
De gedetailleerde functiepuntentelling is het meest nauwkeurig (marge 0-12%). Zij bepaalt alle typen functies, waardeert de complexiteit van elke functie en berekent de functiepuntentelling. Om deze telling te doen zijn gedetailleerde specificaties nodig waarin naast inzicht in de gebruikersfuncties ook complexiteit bepaald kan worden. De inspanning is ongeveer acht uur op duizend uur ontwikkeling (Symons, 1988). Tijdens het ontwikkelingsproces vinden vaak functionele wijzigingen plaats. Daarom wordt, nadat de ontwikkelingsfase voorbij is, nogmaals een functiepuntentelling gedaan om in kaart te brengen hoeveel functiepunten meer of minder opgeleverd is dan bij het ontwerp bepaald is.

Functiepuntenmethode	Marge van afwijking	Benodigde specificaties
Indicatie functiepuntentelling	30-50%	Datamodel
Globale functiepuntentelling	15-25%	Datamodel + globaal ontwerp van de gebruikersfunctie
Detail functiepuntentelling	0-12%	Entiteiten model plus attributen + detail ontwerp op veld niveau

Tabel 1. De functiepunttelmethode in tabelvorm

Iedere fase in de applicatielevenscyclus van FPA kent zijn eigen meeteenheid. De ontwikkelingsfase gebruikt functiepunten, de onderhoudsfase gebruikt onderhoudsfunctiepunten, de testfase gebruikt testfunctiepunten en de exploitatiefase gebruikt runtimefunctiepunten. Elke meeteenheid heeft zijn eigen metrieke. Productiviteit en aantal fouten wordt gebruikt bij onderhoudsfunctiepunten, functiepunten en testfunctiepunten. Operationele kosten is een metriek voor runtimefunctiepunten.

Meer inzicht in de verschillende type functiepunten wordt bereikt als de fasen van de levenscyclus van een informatiesysteem worden bekeken. Hiervoor gebruiken we het alom bekende toestandenmodel van Looijen (1995). Dit model wordt gebruikt om de fasen van de levenscyclus van een applicatie te beschrijven. Tijdens de fasen informatiebeleid, informatieplanning en ontwikkeling worden functiepunten gebruikt om het informatiesysteem tijdens het ontwikkeltraject te begroten. Tijdens de acceptatie en testfase kan door middel van testfunctiepunten de omvang en kwaliteit van testen worden bepaald. Daarna gaat het informatiesysteem in beheer. Tijdens beheer worden wijziging als onderhoudsfunctiepunten geteld. Gebruik en exploitatie kan door middel van runtime functiepunten worden geteld (Stegink, 1993). In figuur 6 is een aangepast toestandenmodel weergegeven met bij elke toestand de functiepuntmethode die in die toestand gebruikt wordt om omvang te meten.

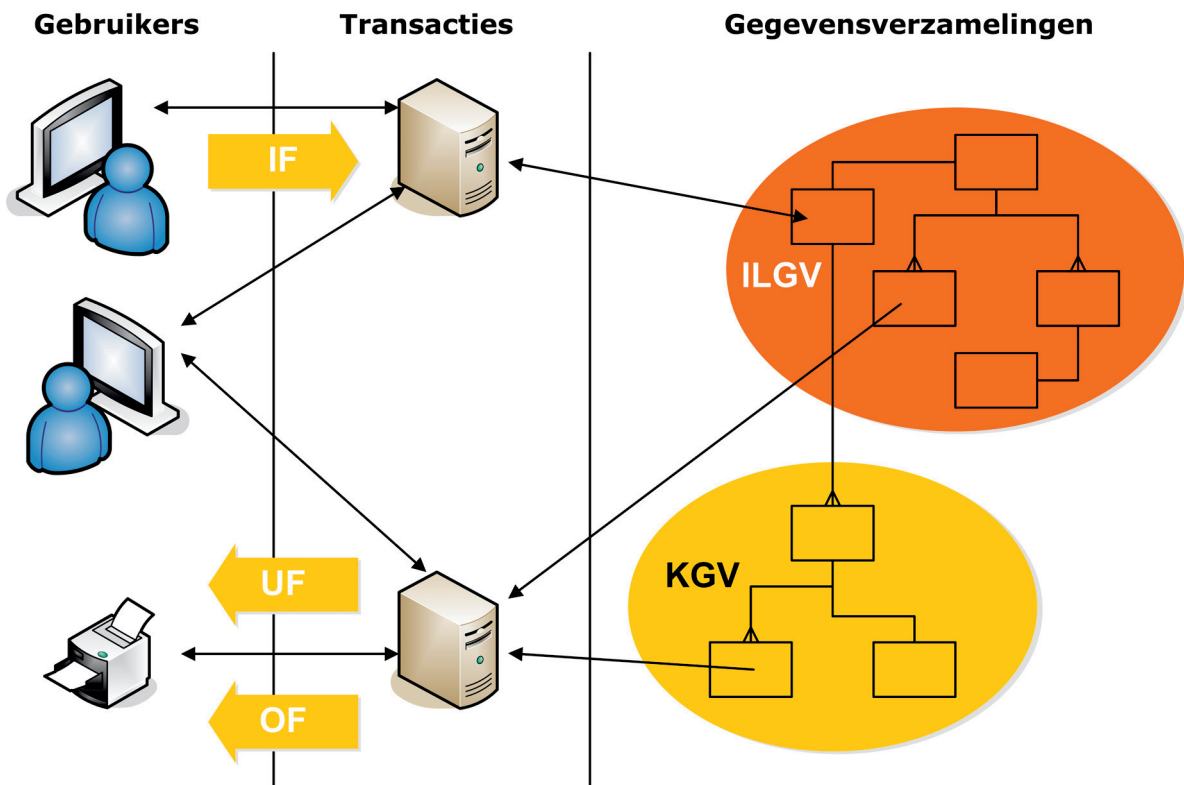


Figuur 6. Toestandenmodel Looijen en bijbehorende typen functiepunten

5.4 Telrichtlijnen

Functiepunten worden bepaald op basis van gegevensverzamelingen en gebruikerstransacties. Van deze gegevensverzamelingen wordt bepaald of ze voldoen aan de NESMA telrichtlijnen (NESMA, 2004) voor interne logische gegevensverzamelingen (ILGV's: systeem-eigen gegevensverzamelingen) of koppelingsgegevensverzamelingen (KGV's: systeemvreemde gegevensverzamelingen). Als een logische gegevensverzameling als ILGV of als KGV is aangemerkt, wordt op basis van de telrichtlijnen de complexiteit bepaald: laag, gemiddeld of hoog. De complexiteit beïnvloedt bij een detailtelling het aantal

functiepunten. Dezelfde methodiek wordt ook toegepast voor de gebruikerstransacties. De gebruikers voeren transacties uit om gegevens uit de gegevensverzamelingen te onderhouden of om gegevens uit de gegevensverzamelingen te tonen. Drie gebruikerstransacties worden onderkend invoerfuncties (IF), de uitvoerfuncties (UF) en de opvraagfuncties (OF) onderkend. Vanuit elke gebruikerstransactie wordt de complexiteit bepaald. In figuur 7 wordt schematisch weergegeven hoe gebruikers middels transacties de gegevensverzamelingen benaderen en welke functies hiervoor gebruikt worden.



Figuur 7. FPA gebruikersfuncties (Vogelezang, 2005)

Nadat de complexiteit van een gebruikersfunctie is bepaald volgens de NESMA telrichtlijnen kan het aantal functiepunten aan de gebruikersfunctie worden toegekend. Dit moet gebeuren volgens de waardering als weergegeven in tabel 2.

	ILGV	KGV	IF	UF	OF
Eenvoudig	7	5	3	4	3
Gemiddeld	10	7	4	5	4
Moeilijk	15	10	6	7	6

Tabel 2. Functiepuntentabel voor het waarderen van gebruikersfuncties

Als alle gebruikersfuncties van een te tellen informatiesysteem via deze methode bepaald zijn dan is de omvang van het informatiesysteem bekend.

5.5 Calculeren van kosten

Hoe de kosten bepaald kunnen worden van softwareontwikkeling is de laatste jaren al veelvuldig beschreven in de literatuur (Jorgensen en Shepperd, 2007). Daarnaast zijn diverse productiviteitsonderzoeken over softwareontwikkeling gedaan (Premaj et al, 2005). De algemene conclusies uit deze onderzoeken zijn dat de beste calculatie van kosten ontstaat als met meerdere methodes hetzelfde project begroot wordt en op basis daarvan een gemiddelde kostprijs wordt bepaald. Hoe verder het project gevorderd is (en het detailniveau van de specificaties toeneemt) hoe nauwkeuriger de schatting wordt.

Uit de literatuur blijkt dat FPA een goede methode is om te begroten (Premaj en Sheppard, 2005). Een andere meeteenheid zoals Lines of code (loc) wordt meestal niet gebruikt, omdat de gebruikte programmeertaal het aantal lines of code beïnvloedt en vergelijken moeilijk maakt (Liebchen en Shepperd, 2006). Daarnaast zijn loc pas beschikbaar op het einde van het project, waardoor loc voor het onderzoek niet relevant is aangezien in het onderzoek vooraf een calculatie gemaakt moet worden. De functionaliteit van een informatiesysteem is vooraf bekend waardoor FPA geschikt is om de omvang te bepalen, zodat op basis van omvang een calculatie van kosten gemaakt kan worden. MARK II is een variant van FPA dat complexiteit in zich heeft, maar vereist historische gegevens (Tran-Cao et al, 2002). Die historische gegevens zijn vooraf vaak niet beschikbaar. Voor datamigratietrajecten zijn die er vrijwel zeker nooit. Daarnaast bestaan er nog andere methoden die in dit verslag niet aan de orde komen (Nashir, 2006). Begroten met functiepunten gebeurt door een kostprijs per functiepunt te bepalen op basis van productiviteitsfactoren.

Op basis van productiviteit wordt een kostprijs per functiepunt bepaald. Productiviteit meten in softwareontwikkeling is een complexe aangelegenheid, omdat er veel variabelen zijn die de productiviteit kunnen beïnvloeden. Hierdoor is het lastig productiviteitscijfers te gebruiken die uit metingen komen.

Een aantal factoren is van invloed op productiviteit. De eerste factor is de grootte van het team, hoe groter het team wordt hoe lager de productiviteit wordt. Een tweede factor is ervaring van het team. Een derde factor zijn de tools waarmee wordt gewerkt. Een vierde factor is de manier (methode) van werken. Een vijfde factor is de branche waarin wordt gewerkt. Een zesde factor is de soort applicatie die ontworpen/gemaakt/getest/gemigreerd wordt (Angelis et al, 2001).

Welke productiviteit gebruikt moet worden om een offerte te maken, heeft te maken met kengetallen. Deze kengetallen zijn de productiviteitsfactoren verzameld uit diverse projecten. Projecten die het bedrijf zelf heeft uitgevoerd, leveren de meest bruikbare cijfers op, omdat dan alle parameters bekend zijn en de validiteit van gegevens hoog is (Grimstad en Jorgensen, 2006). Voor de meeste bedrijven is het een probleem om voldoende cijfers voorhanden te hebben van een soortgelijk project als het project dat wordt begroot (Jeffery et al, 2001). Een alternatief is om een database met projectgegevens te kopen om daar kengetallen uit te bepalen.

De International Software Benchmarking Standards Group (ISBSG) (www.isbsg.org) is een non profitorganisatie dat zich tot doel heeft gesteld het managen van IT-middelen te verbeteren.

Dit doet de ISBSG door projectgegevens over software-ontwikkeltrajecten in databases

te verzamelen. Deze projectgegevens worden deels in functiepunten uitgedrukt en bijna alle functiepuntgebruikersgroepen over de hele wereld zijn aangesloten bij de ISBSG. Op dit moment zitten meer dan 3000 projecten in deze database. De gegevens zijn verdeeld over zes datagroepen. Deze zijn projectattributen, bestede tijd, projectomvang in functiepunten, kwaliteitsgegevens, projectresultaat en projectkosten (Angelis et al, 2001). Uit deze databases kunnen kengetallen worden gehaald. Een van die kengetallen is productiviteit. De productiviteit per functiepunt kan worden gebruikt als basis om te rekenen met functiepunten. Een aantal onderzoekers heeft de ISBSG database verder onderzocht. Een conclusie is dat elk bedrijf een andere productiviteit heeft (Mendes en Lokan, 2005); (Mendes en Lokan, 2006). Zolang er geen productiviteitscijfers zijn, kan de ISBSG database worden gebruikt om productiviteitscijfers te bepalen. Uit onderzoek is gebleken dat gegevens gebruiken van andere bedrijven niet altijd tot het gewenste resultaat leidt (Mendes en Lokan, 2006). De ISBSG database wordt veel gebruikt en diverse onderzoekers valideren de kwaliteit van de ISBSG database (Lokan, 2005) (Lokan et al, 2001).

Om FPA te kunnen gebruiken om een kostprijs in manuren te bepalen moet een bedrijf haar eigen productiviteitsratio's op basis van functiepunten kennen. Productiviteitsratio is het gemiddelde aantal uren dat nodig is om één functiepunt te realiseren. Dit is gebaseerd op eigen ervaringen met vergelijkbare projecten. Vergelijk dit met de branchegemiddelden om het gemiddelde van het eigen bedrijf af te zetten tegen branchegenoten (vergelijkingsmateriaal: ISBSG).

Om een projectbudget te bepalen, dienen vier stappen doorlopen te worden.

1. De eerste stap is het vaststellen van de functionele omvang in functiepunten van het te ontwikkelen informatiesysteem.
2. De tweede stap is nagaan wat de standaard productiviteitsfactor per functiepunt is van het soort project met de gekozen tools binnen het bedrijf dat het project gaat uitvoeren.
3. De derde stap is het bepalen van de verwachte productiviteitsratio van het project. Hierbij spelen allerlei omstandigheden een rol die de standaard productiviteitsratio kunnen beïnvloeden zoals; teamgrootte, gewenste doorlooptijd, kennis en ervaring van de teamleden.
4. De laatste en vierde stap is het feitelijke budgetteren van het project.

Verder is van belang vast te leggen welke activiteiten wel en welke niet worden meegenomen in het projectbudget, duidelijke scopeafbakening. Het is van belang om aandacht te schenken aan de berekening van activiteiten die buiten het berekende projectbudget vallen. Probeer deze kosten zo goed mogelijk in te schatten. Deze kosten worden naast een vaste prijs per functiepunt doorberekend.

Het werken met een vaste prijs per functiepunt heeft als uitdaging het aantal functiepunten dat begroot is te bewaken indien er een vast aantal functiepunten begroot is. Tijdens het project kan autonome groei optreden. Autonome groei is detaillering van specificaties dat meestal leidt tot groei van het aantal functiepunten. Let daarnaast op scopeverschuiving. Scopeverschuiving is veranderende/nieuwe eisen, die meestal leiden tot groei van het aantal functiepunten. Deze extra functiepunten is meerwerk.

Om te bepalen of het aantal uren per functiepunt juist is kan getoetst worden door een schatting te maken van de kosten in mensuren met andere begrotingsmethoden. Als de

omvang in uren bekend is kan een kostprijs in euro's bepaald worden, deze berekening valt buiten de scope van dit onderzoek en wordt verder niet uitgewerkt.

5.6 Kengetallen

Tijdens beheer, test en nieuwbouwtrajecten zijn de uit te voeren activiteiten vaak het hetzelfde en als in een soortgelijke omgeving gewerkt wordt dan zijn de gegevens over die projecten te gebruiken om kengetallen te ontwikkelen.

Een aantal kengetallen, die bepaald kunnen worden met behulp van FPA, zijn onder andere:

- Productiviteit te meten in uren per type functiepunt voor een bepaalde ontwikkelomgeving en type / generatie programmeertaal / type werk (ontwerp, ontwikkeling, onderhoud, testen);
- Kwaliteit van het systeemontwikkelingsproces te meten in aantal fouten per tijdsperiode testen per functiepunt. Fouten zijn de verschillen in hetgeen geprogrammeerd is en hetgeen functioneel gevraagd is;
- Kwaliteit van het informatiesysteem te meten in aantal fouten per tijdsperiode per functiepunt. Fouten die optreden na in productie van het systeem wordt hier bedoeld met fouten;
- Beheerkosten van het informatiesysteem te meten in kosten per runtimefunctiepunt rekening houdende met CPU-tijd, netwerkactiviteit, enz.

De richtlijnen van NESMA zijn richtlijnen om de omvang in functiepunten te bepalen, deze richtlijnen zijn geen wetten. Hierdoor kan een kleine afwijking in aantal functiepunten op treden als meerdere mensen hetzelfde systeem tellen. Het resultaat is een indicatie en geen absolute waarde dit betekent dat een systeem van 1000 functiepunten na een functiepunttelling kan uitkomen op 900 of 1100 functiepunten, maar niet 500 of 1500 functiepunten. FPA werkt met een complexiteitsfactor die het aantal functiepunten beperkt beïnvloedt. Complexiteit heeft grenzen, moeilijker dan moeilijk kan niet en eenvoudiger dan eenvoudig kan niet. Ervaringscijfers zijn gebaseerd op statistiek, de wet van de grote aantallen. Rekenen met een klein aantal functiepunten (< 100 functiepunten) is niet erg betrouwbaar. De betrouwbaarheid neemt toe als het aantal functiepunten toeneemt, verwacht geen 'fout'vrije projectschattingen (Kitchenham, 1997). Alleen als alle regels van FPA gevolgd zijn, is de telling betrouwbaar (Boetticher en Lokhandwala, 2007).

5.7 Recente ontwikkelingen

Wereldwijd gezien is Functional Size Management (FSM) bezig met een opmars. FSM is het geheel van alle methoden om omvang van een informatiesysteem te bepalen. De verwachtingen zijn dat FSM doorgroeit in specifieke gebieden van systemdevelopment, maar ook in andere gebieden van de ICT (bijv. gebruik van FSM in de operatie met RFPA en iteratieve ontwikkelingen (bijv. use case points)) en dé basis voor benchmarking wordt. Daarnaast vindt integrale toepassing van FSM plaats bij projectmanagement, IT-governance en systeemontwikkelmethoden. Gebruik van functiepunten als meeteenheid begint toe te nemen. Dit wordt gestimuleerd doordat bedrijven transparantie nodig hebben. Bijvoorbeeld in outsourcetrajecten, waar afspraken worden gemaakt

over onderhoud van informatiesystemen op basis van een vaste prijs per functiepunt. Cosmic FFP is groeiende en FISMA FPA is onlangs gecertificeerd door ISO. In 2005 heeft the IEEE een groot Metrics symposium gehouden over FPA (Premraj en Sheppard, 2005) (Santillo et al, 2005) (Liebschen en Sheppard, 2005) Mustafa, Growthaman en Khan (Mustafa et al, 2005).

In Nederland is de NESMA actief om het gebruik van FPA te stimuleren. Op dit moment staan testfunctiepunten en alles wat met testen in combinatie met functiepunten te maken heeft in de belangstelling. Op de NESMA-najaarsconferentie in 2007 was testfunctiepunten het thema. In 2008 maakt NESMA zich sterk om certificering van FPA via Exin door te zetten.

In de markt is Sogetti op dit moment bezig om productiviteitsfactoren voor testen te bepalen om zo een vaste prijs per testfunctiepunt af te spreken met klanten.

5.8 Conclusie

FPA is een algemeen geaccepteerde methode om omvang te bepalen voor ontwerp, ontwikkel, test en beheertrajecten. Functiepunten vormen een objectieve meeteenheid, vergelijkbaar met tonnen in de transportwereld of m² in de bouw (Pressman, 2001). Een meeteenheid maakt het mogelijk om te vergelijken. Met FPA kan van alles gemeten en vergeleken worden, voorbeelden zijn kosten van ontwerp of ontwikkeling of testen of onderhoud, productiviteit van teams of tools of bedrijven, efficiency van informatiesystemen en platvormen, kwaliteit van informatiesystemen en serviceorganisaties, aantal fouten, enz. FPA maakt het mogelijk om te budgetteren bij het ontwerpen, ontwikkelen, testen en exploiteren van informatiesystemen. FPA is een eenvoudig te leren administratieve activiteit, die erkend wordt door nationale en internationale organisaties. Wereldwijd zijn er diverse gebruikersorganisaties die FPA ondersteunen en uitdragen, zodat het een wereldwijde gebruikte methode is. FPA is toepasbaar gedurende de gehele levenscyclus van een informatiesysteem. Binnen FPA is niet beschreven hoe om te gaan met datamigratietrajecten.

In dit hoofdstuk is FPA beschreven en in het vorige hoofdstuk is datamigratie beschreven. In het volgende hoofdstuk wordt een aanpak beschreven hoe op basis van FPA datamigratietrajecten begroot kunnen worden.

6 Het begroten van datamigratietrajecten

Binnen de in hoofdstuk 3 onderzochte, bestaande methoden is er geen gevonden die gebruikt zou kunnen worden om datamigratietrajecten te begroten. Ook zijn er geen generieke factoren gevonden die gebruikt kunnen worden om de omvang in uren te bepalen. In hoofdstuk 4 is een datamigratiemethode beschreven voor een projectmatige aanpak, maar hiermee is het maken van een begroting nog niet mogelijk. In hoofdstuk 5 is beschreven wat FPA is en hoe het toegepast wordt. In dit hoofdstuk gaan we in op de vraag of FPA gebruikt kan worden om datamigratietrajecten te calculeren en op welke manier dit zou kunnen. Hiervoor gebruiken we het Mediaan Migratie Model uit hoofdstuk 4 als vertrekpunt. De deelvragen over complexiteit en productiviteit zijn nog niet beantwoord; daar gaat dit hoofdstuk ook verder op in.

6.1 Quickscan van de praktijk

Aangezien Mediaan/abs ongeveer 30 datamigratietrajecten met goed resultaat heeft uitgevoerd; is besloten om met een klein aantal experts een quickscan uit te voeren. Door middel van gestructureerde interviews worden zo praktijkervaringen bij het onderzoek betrokken. De interviews zijn uitgevoerd door het invullen van een vragenlijst aangevuld met een vraaggesprek. In bijlage I is deze vragenlijst opgenomen. Het doel van de quickscan is inzicht te krijgen hoe de migratie experts datamigratietrajecten hebben uitgevoerd en wat hun ervaringen daarbij zijn. Door analyseren van de gebruikte best practices verwachten we bruikbare aanknopingspunten te vinden voor het begroten van datamigratietrajecten.

In totaal zijn acht experts geïnterviewd. Deze waren bij verschillende typen datamigratietrajecten betrokken. Ook de branche en omvang van project waren verschillend. Hieronder volgen de belangrijkste bevindingen.

- a) In bijna alle projecten die uitgevoerd zijn is data geschoond, omdat de kwaliteit van de data niet voldoende was om zonder meer te migreren. De meeste opschooninspanning wordt geleverd door de gebruikersorganisatie van de te migreren data. Hoeveel inspanning het schonen kost kan op dit moment niet vooraf gecalculeerd worden. Dientengevolge worden de schoningskosten buiten de projectbegroting gehouden.
- b) Het datamigratietraject is in bijna alle gevallen een gedeeltelijke datamigratie (actieve klanten, laatste vijf jaar, enz.). Dat niet alle gegevens uit het bronstelsel meegenomen worden verhoogt de complexiteit van het datamigratietraject en de verantwoordingsplicht naar de opdrachtgever.
- c) De opdrachtgever wil in alle gevallen bewijsvoering ontvangen (in de vorm van rapportages) dat de datamigratie juist en volledig is uitgevoerd. Deze bewijsvoering vindt plaats in de vorm van rapportages en aansluitingsoverzichten. De aansluitingsoverzichten zijn nodig om verschillen te verklaren die ontstaan tijdens de migratie. Een voorbeeld hiervan is dat een bedrijf in het oude systeem alle onderdelen in het magazijn per stuk een waarde toekent en in het nieuwe systeem alleen halffabrikaten een waarde toegekend krijgen. De waarde van producten in het magazijn voor en na datamigratie is dan niet exact hetzelfde.

- d) Een datamigratietraject kent meestal meerdere iteratieslagen (3-6) om de uitval van data binnen de gestelde grenzen te krijgen. Deze iteratieslagen omvatten ook tests of de datamigratie binnen de gestelde tijd uitgevoerd kan worden en kunnen gebruikt worden ter controle van de draaiboeken die tijdens datamigratie gevolgd worden.
- e) De opdrachtgever wil minimale business impact, minimale kosten, minimale uitval van data, zekerheid over juistheid en volledigheid van data en zekerheid dat de business na datamigratie zonder problemen uitgevoerd kan worden.
- f) In een aantal datamigratietrajecten is de Migration Workbench (Koorneef, 2008) gebruikt. Dit is een tool om datamigratietrajecten uit te voeren. Het inzetten van een goed tool bij datamigratietrajecten verhoogt de productiviteit bij datamigratie. Vaak is een bijkomend voordeel dat aanpassingen in specificaties doorgevoerd kunnen worden in het migratietool wat automatisch leidt tot aangepaste code. De leverancier van het tool heeft meestal gezorgd dat specificaties goed beheersbaar zijn en dat het tool allerlei rapportageopties bevat.
- g) Binnen alle projecten is via een work breakdown structure (Brykczynski, 2006) een projectmatige aanpak afgedwongen. In een work breakdown structure zijn alle uit te voeren werkzaamheden vooraf uitgewerkt in activiteiten. Voor al deze activiteiten is uitgeschreven wie de activiteit wanneer uitvoert en welke tijd hiervoor gepland staat.
- h) Door wijzigingen in scope van en onvoldoende kennis over de bronsystemen verandert migratieprogrammatuur vaak. In de meeste projecten wijzigt de helft van de programmatuur meer dan één keer in de doorlooptijd van het datamigratietraject.
- i) In alle projecten is het gehele datamigratietraject op basis van nacalculatie uitgevoerd. In geen enkel project is vooraf een vaste prijs gecalculeerd en afgegeven aan de klant.
- j) In één datamigratietraject zijn handmatig lopende dossiers in het nieuwe systeem ingebracht. Dit is overigens het enige project waar handmatige acties onderdeel waren van het project.

Discussie quickscan

Tijdens het houden van de quickscan is geen kant en klare best practice gevonden om de omvang van datamigratietrajecten vooraf te calculeren. Dit was niet de insteek of de verwachting. Wel werd duidelijk dat slechts een gedeelte van een datamigratietraject met functiepunten begroot kan worden.

Op basis van deze acht migratietrajecten wordt geschat dat 30% van het totale migratie projectbudget opgaat aan datamigratietrajecten. Hieruit blijkt dat begroten van datamigratietrajecten van wezenlijk belang is.

De verwachting is dat door gebruik te maken van de Mediaan Migratie Methode en projectmanagement de volgende risico's ondervangen kunnen worden:

- kwaliteit van te migreren gegevens van het bronsysteem
- ontbreken kennis hoe gegevens in het bronsysteem vastgelegd zijn
- wijzigingen in scope van het datamigratietraject
- de bewegelijkheid van de specificaties van het doelsysteem
- ontbreken van datamodel van het doelsysteem

- beschikbaarheid van materiedeskundigen van doelsysteem
- opschooncapaciteit van de gebruikersorganisatie

Een goede datamigratiemethode en veel aandacht voor projectmanagement is van groot belang bij datamigratietrajecten, omdat er in het algemeen veel capaciteit van de business nodig is om het project te doen slagen en veel kennis verzameld moet worden.

Het buiten de begroting houden van de activiteit “opschonen van data” is een werkwijze die overgenomen wordt in dit onderzoek. Aangezien de meeste tijd van schoning voor rekening komt van de gebruikersorganisatie is alleen het kwaliteitsaspect van gemigreerde data een onderdeel van het datamigratietraject.

In de projecten is te weinig cijfermatig vastgelegd, waardoor na afloop van het traject weinig inzicht is waaraan uren besteed zijn. Als gevolg hiervan is de omvang in functiepunten niet achteraf te calculeren. Hierdoor is het onmogelijk om op basis van reeds uitgevoerde projecten rekenregels op te stellen om omvang te bepalen.

Gebruik maken van een datamigratieworkbench is productiever dan zelf programmatuur ontwikkelen. Projectplanning wordt in dit geval gedaan op basis van praktijkervaring van de bewuste expert.

6.2 Calculeren omvang datamigratietrajecten

Dat FPA gebruikt kan worden om de omvang van datamigratie te begroten, moet nog aangetoond worden. Wel is duidelijk dat niet de hele omvang van een informatiesysteem in functiepunten relevant is bij datamigratie. Het gaat immers vooral om het aantal logische gegevensverzamelingen in het bron- en doelsysteem. Waar in nieuwbouwtrajecten vooral schermen en rapporten geteld worden in functiepunten wordt bij datamigratie gekeken naar rapporten, extractie-, transformatie- en laadfuncties. Hiermee is het gebruik van functiepunten bij datamigratietrajecten wezenlijk anders dan bij nieuwbouw.

Voor activiteiten die niet uitgedrukt worden in functiepunten gebruiken we de work breakdown structure (WBS). WBS is de opdeling van het uit te voeren werk in componenten van samenhangende activiteiten. Deze componenten kunnen worden opgedeeld in deelcomponenten. Deelcomponenten worden opgedeeld in concrete activiteiten met daaraan gekoppeld de te besteden tijd en wie de activiteit gaat uitvoeren. Het aantal componenten, deelcomponenten en activiteiten is afhankelijk van de complexiteit en omvang van de uit te voeren werkzaamheden.

Het schonen van data is niet met functiepunten te begroten, omdat er geen functionaliteit vanuit de gebruiker wordt gevraagd. Meestal zijn experts van het bronstelsel en het doelsysteem nodig om te bepalen welke data er precies geschoond moet worden. Op welke manier data wordt geschoond is verschillend per migratietraject. In sommige gevallen kan voor een groot aantal gegevens geautomatiseerd worden geschoond, maar heel vaak is veel gebruikersinteractie en handmatig werk nodig om de data geschoond te krijgen.

Het is verstandig om schonen van data niet mee te nemen in de calculatie van kosten

van een datamigratietraject. Voor schonen van data dient een apart schoningstraject opgezet te worden met een eigen calculatie van kosten in euro's en doorlooptijd in uren.

Het schoningstraject kan begroot worden middels een WBS, maar het aantal te besteden uren is niet met enige mate van zekerheid vooraf vast te stellen. Vooraf kan bepaald worden waaraan data moeten voldoen, maar hoeveel data niet voldoen aan de eisen wordt pas later duidelijk. Vervuilde gegevens bepalen in grote mate de complexiteit, kosten en doorlooptijd van het schoningstraject. De hoeveelheid data die geschoond moet worden en de wijze van schoning zal per schoningstraject anders zijn. Mogelijk is er een correlatie tussen de omvang in functiepunten en de tijd besteed aan schoning te vinden. Het is raadzaam om de uren van schoning op nacalculatiebasis te calculeren.

Voor het bepalen van de omvang van het datamigratietraject wordt bij de Mediaan Migratie Methode (MMM) een onderscheid gemaakt tussen activiteiten die via WBS of via FPA begroot kunnen worden (figuur 9). In formule:

$$T = W + F-S \quad (1)$$

Hierin is:

- T {MMM} de verzameling die alle activiteiten uit de MMM omvat
- W {WBS} de verzameling van activiteiten uit het MMM die met WBS begroot kunnen worden
- F {FPA} de verzameling van activiteiten uit het MMM die met FPA begroot kunnen worden
- S {WBS} de verzameling van activiteiten uit het MMM die met schonen te maken hebben en met WBS begroot zouden kunnen worden.

Voor het bepalen van de omvang in uren geldt:

$$Q = \Sigma A + P \times FP \quad (2)$$

Hierin is:

- Q [mensuur] de omvang van het datamigratietraject.
- ΣA [mensuur] de grootte bestaande uit een verzameling van datamigratieactiviteiten die terug te vinden zijn in de MMM in uren die niet functiepunten is uit te drukken (denk aan bepalen scope, bepalen requirements, bepalen datakwaliteit, pilotproject, definitiefase, beheerfase en opschonen/verrijken van data, projectmanagement, rapportage). Dit komt overeen met deelverzameling W-S uit formule (1).
- P [mensuur per functiepunt] de datamigratieproductiviteit per functiepunt.
- FP [functiepunt] is het aantal functiepunten van een aantal activiteiten die in de MMM in functiepunten zijn uit te drukken. Dit zijn het logische gegevens model van bron- en doelsysteem, de gebruikersfuncties die nodig zijn om de datamigratie uit te voeren (bedrijfsregels en datamigratieprogrammatuur) plus een aantal functies voor rapportage en beheer. De functiepunten worden bepaald volgens de NESMA telrichtlijnen. Dit komt uiteen met de verzameling F uit formule (1).

Functiepunten kunnen bepaald worden voor de activiteiten ontwerpen, realiseren, testen en beheren. Van de processen van ETL, zoals beschreven in hoofdstuk 3, is het aantal functiepunten te bepalen. De overige activiteiten van een datamigratietraject zijn niet gerelateerd aan functiepunten.

Projectmanagement, zijnde teamleiding, van een ontwikkeltraject wordt vaak op een bepaald percentage van de ontwikkeltijd gezet. De ontwikkeltijd is dan: Omvang in functiepunten gedeeld door de productiviteit in functiepunten. Dus als projectmanagement een percentage vermenigvuldigt met de ontwikkeltijd is, dan is projectmanagement een percentage van de gemiddelde productiviteit per functiepunt.

In formule:

$$\text{PMT} = \text{pct} * \text{Q} \quad (3)$$

Waarbij

PMT [mens uur] de omvang van projectmanagement is.

Pct [percentage] een opslag is.

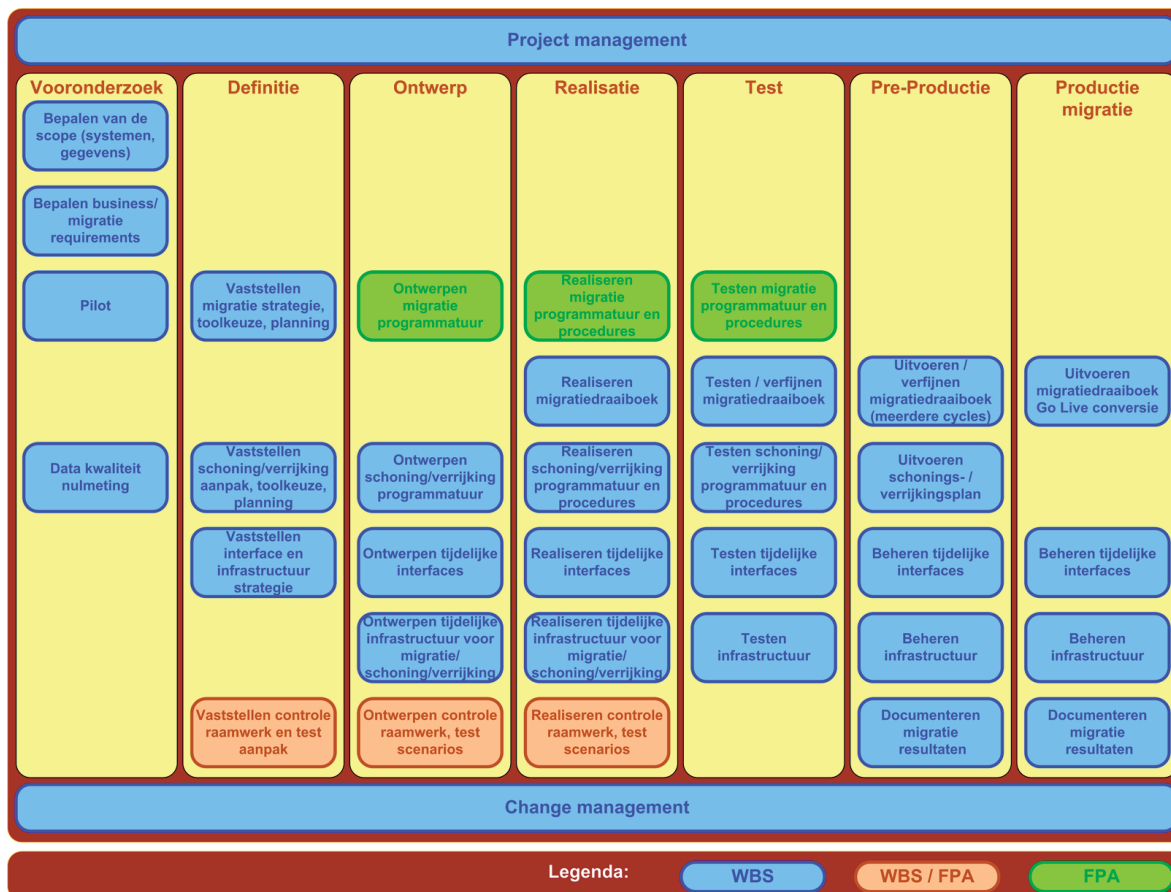
Q formule (2).

Een andere benadering is dat projectmanagement, zijnde teamleiding, afhankelijk is van de doorlooptijd, omdat men bijvoorbeeld zegt één dag projectmanagement per week te besteden. Het aantal weken kan bepaald worden op basis functiepunten gedeeld door productiviteit (FP / P). Dus ook weer FP gerelateerd.

In de theorie van de Nesma is projectmanagement niet afhankelijk van de omvang. Hierbij wordt uitgegaan van Prince2, dan is projectmanagement geen teamleiding. Hierbij bepaalt niet de omvang, maar klant, omgeving, deelprojecten, product, enz. de te besteden tijd. De werkzaamheden van de projectmanagement bestaan in dit geval uit plannen, meten, rapporteren over de voortgang. Of over 1000 fp of 20000 fp rapporteerd wordt maakt in het werk van een projectmanager heeft weinig uit. Hij krijgt zijn cijfers immers van anderen. Die anderen doen het werk.

Changemanagement is in feite ook een projectmanagement functie. Waarin de wijzigingen binnen het project met stuurattributen wordt bijgehouden. Deze gevens zijn de basis van rapporteren, plannen en bewaken van wijzigingen. Deze activiteit staat los van functiepunten.

In figuur 8 is op basis van de Mediaan Migratie Methode aangegeven bij welke activiteit met FPA of WBS gecalculereerd wordt.



Figuur 8. Mediaan Migratie Methode, uitsplitsen activiteiten naar WBS en/of FPA

6.3 WBS gerelateerde uren

De ΣA uren worden via een WBS bepaald door alle activiteiten in kaart te brengen en vervolgens te voorzien van een gepland aantal te besteden uren. Indien voor elk project dezelfde aanpak toegepast wordt (Mediaan Migratie Methode) zullen de uit te voeren activiteiten over het algemeen voor elk datamigratietraject vergelijkbaar zijn.

Waarschijnlijk zal het aantal keer dat een bepaalde activiteit uitgevoerd dient te worden per project verschillen. Het ene project heeft meer afstemming nodig om bijv. de migratiestrategie te bepalen dan het andere project. Hoeveel proefmigraties gedaan worden zal per project verschillen, enz.

Door van elk datamigratietraject de schatting vooraf te vergelijken met het daadwerkelijk bestede aantal uren, kan de betrouwbaarheid van de WBS vergroot worden. Vastleggen welke handelingen uitgevoerd zijn in een activiteit en gegevens over de activiteit zijn nodig bij het analyseren van de planning in een vervolgttraject. Bijvoorbeeld bij het uitvoeren van een kwaliteitsmeting is informatie over aantal tabellen, platform en uitgangsdokumentatie van belang om voor volgende datamigratietrajecten een goede planning in uren te maken voor die activiteit.

Alle uren in de WBS opgeteld is de factor ΣA in formule (2).

6.4 Datamigratiefunctiepunten

Om voor datamigratie de functiepunten te kunnen tellen zal de NESMA werkwijze en richtlijnen voor FPA worden gehanteerd. In het verdere onderzoek introduceren we de term “datamigratiefunctiepunten” om aan te geven dat het hier functiepunten betreft die gebruikt worden in een datamigratietraject.

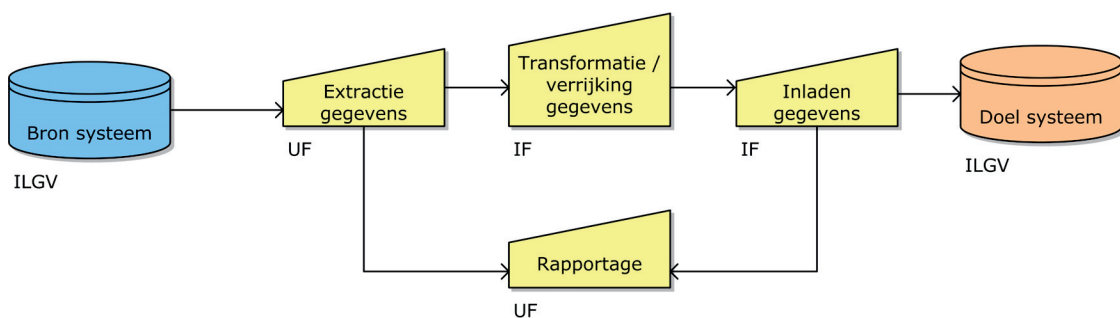
Binnen datamigratietrajecten is het nodig om via een aantal standaard mechanismen snel een FPA uit te voeren. Net zoals bij nieuwbouwtrajecten een snelle methode bestaat om een globale telling uit te voeren zonder het hele systeem vooraf in detail bekend te hebben.

In de komende paragrafen wordt een methode uitgewerkt om een globale telling met FPA bij datamigratietrajecten uit te voeren. Uitgangspunt is de manier waarop volgens de gebruiker het doelsysteem moet worden gevuld. Tweede uitgangspunt is dat binnen een organisatie de logische gegevensverzamelingen van bron- en doelsysteem sterk op elkaar lijken: beide systemen bevatten ongeveer dezelfde logische gegevensverzamelingen. Het doelsysteem kan niet gevuld worden middels een datamigratie als daar geen gegevens voor beschikbaar zijn. Hieruit kan geconcludeerd worden dat voor ieder te vullen gegevensverzameling uit het doelsysteem één of meerdere bronsysteemgegevensverzamelingen aanwezig zijn. Hierdoor ontstaat een functiepuntenmigratiestructuur.

Tijdens de offertefase is wel het bronsysteem bekend en in veel gevallen een concept-datamodel van het doelsysteem. Gekozen is om tijdens dit onderzoek uit te gaan van het aantal te vullen doelsysteemtabellen. Op basis van dit uitgangspunt zal per doelsysteemgegevensverzameling, één of meerdere extractie-, transformatie- en laadfuncties nodig zijn, naast één of meer rapportagefuncties. Verder is er een gegevensverzameling waar de brongegevens uitgehaald worden.

6.4.1 Globale datamigratiefunctiepunten telling

In de meeste datamigratietrajecten zal tijdens de start van het datamigratietraject op basis van de beschikbare documentatie alleen een globale functiepunten telling uitgevoerd kunnen worden. In figuur 9 is in schemavorm de benodigde FPA componenten weergegeven die nodig zijn om één doelgegevensverzameling te vullen. Binnen migratiefunctiepunten wordt op basis van de te vullen doelsysteemgegevensverzameling ETL functies geteld, plus een rapportage en bronsysteemgegevensverzameling. ETL is beschreven in hoofdstuk 3.2. Voor deze opzet is gekozen, omdat vanuit gebruikersoogpunt er een doelsysteem aanwezig moet zijn waar gegevens uitgehaald moeten worden, aanpassing van deze gegevens nodig is, die daarna in een bepaalde structuur ingelezen worden in het doelsysteem.



Figuur 9. Globaal model datamigratietraject per logische gegevensverzameling

Het schema op pag. 43 uitgedrukt in functiepunten (NESMA, 2004) voor een globale functiepunten telling is:

- (A) Bronsysteem = 1 ILGV = 7 FP
 - (B) Extractie gegevens uit bronsysteem = 1 UF = 5 FP
 - (C) Transformatie = 1 IF = 4 FP
 - (D) Rapportage = 1 UF = 5 FP
 - (E) Inladen gegevens = 1 IF = 4 FP
 - (F) Doelsysteem = 1 ILGV = 7 FP
- Bij een globale telling per doelsysteem ILGV is de omvang 32 FP

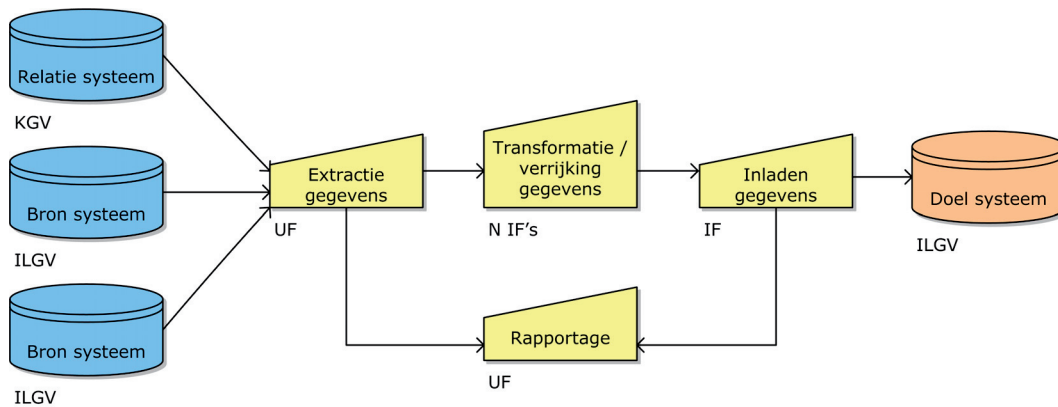
Waarin:

- (A) de functiepunten van het bronsysteem zijn voor het valideren van data uit het bronsysteem, zodat kwaliteit en uitval bepaald kan worden.
- (B) de functiepunten van het extractieprogramma zijn om data uit het bronsysteem te halen en klaar te zetten voor transformatie.
- (C) transformatie is het transformeren van brongegevens, zodat het ingelezen kan worden in het doelsysteem. Indien nodig worden data verrijkt om te voldoen aan eisen van het doelsysteem. Uitval en data die voldoen aan de eisen van het doelsysteem, zijn het resultaat van de functiepunten die gebruikt worden tijdens transformatie.
- (D) uitgegaan wordt van een rapport waarin gerapporteerd wordt over het hele data-migratietraject dat afgelegd wordt om de doelsysteem ILGV te vullen.
- (E) de functionaliteit betreffende het klaarzetten van de gegevens om het doelsysteem te vullen, dit kan een tabel zijn maar ook een inlaadprocedure.
- (F) de functiepunten de ILGV van het doelsysteem en het daadwerkelijk vullen van gegevens de functiepunten zijn die hier geteld worden.

Om een globale functiepuntentelling voor een datamigratietraject uit te voeren wordt het aantal doel logische gegevensverzamelingen vermenigvuldigd met 32 functiepunten. Dus bij 100 doeltabellen komen we hier op 3200 functiepunten.

6.4.2 Detail datamigratiefunctiepuntentelling

Indien er gedetailleerde documentatie beschikbaar is of de informatie nodig voor een detailtelling verkregen wordt in de vooronderzoekfase, kan een detail functiepuntentelling uitgevoerd worden. Het verschil tussen een globale en detailtelling is dat in de detailtelling de complexiteit wordt meegenomen (zoals beschreven in H5.4). Dit levert een preciezer telling op, bijv. dat er meer gegevensverzamelingen betrokken zijn of meer of minder rapporten nodig zijn. Daarnaast wordt ook het aantal attributen en de complexiteit van functies meegenomen in de functiepuntanalyse. Per doeltabel wordt bepaald wat voor die doeltabel het aantal benodigde gebruikersfuncties en logische gegevens is wat nodig is om de omvang in FP te bepalen voor die doeltabel. Een willekeurig voorbeeld van een detailplanning om een doelsysteem logische gegevens verzameling te vullen is in figuur 10 weergegeven.



Figuur 10. Detail model datamigratietraject per logische gegevensverzameling

Bij bovenstaand model gaan we uit van 2 ILGV's, een extra koppeltabel uit het relatiesysteem en twee opschoonfuncties en een rapport. Een detail functiepuntentelling kan diverse verschillende gebruikersfuncties bevatten, voor het onderzoek is het niet relevant voor elke mogelijkheid een figuur op te stellen.

Deze detailtelling komt uit op:

- A) 1 relatietabel uit een ander systeem = 1 KGV = 5 FP
 - B) 2 interne logische gegevensverzameling uit het bronsysteem = 2 ILGV = 2 x 7 FP = 14 FP
 - C) 1 extractiefunctie = 1 UF = 5 FP
 - D) 2 transformatiefuncties = 2 IF x 4 FP = 8 FP
 - E) 1 rapport = 1 UF = 5 FP
 - F) 1 inlaadfunctie = 1 IF = 4 FP
 - G) 1 interne logische gegevensverzameling van het doelsysteem = 1 ILGV = 7 FP
- Deze berekening komt uit op 48 FP

Op basis van het ETL model aangevuld met rapportages kan voor elke doelgegevensverzameling in kaart gebracht worden welke gebruikersfuncties nodig zijn om de doeltabel te vullen.

6.5 Productiviteitsfactoren

Productiviteitsfactoren beïnvloeden de prijs per functiepunt, zoals aangegeven in formule (2). In deze paragraaf wordt een aantal factoren genoemd die de prijs per functiepunt beïnvloeden. Op basis van een case, uitgewerkt in de volgende paragraaf, wordt het risico getoond om met een vaste prijs te werken als er geen productiviteitsfactoren bekend zijn.

De methode, tools, teamsamenstelling en omgeving waarin de datamigratie wordt gedaan bepalen mede de productiviteit. Op dit moment zijn geen cijfers voorhanden welke productiviteit gehaald wordt in datamigratietrajecten. Daardoor kunnen er (nog) geen kengetallen ontwikkeld worden m.b.t. productiviteit bij datamigratietrajecten.

Om een betrouwbare formule voor productiviteit te krijgen moeten gegevens worden verzameld over datamigratietrajecten. Per project moet het aantal functiepunten worden geteld en moeten de uren van de FPA gerelateerde activiteiten worden vastgelegd. De

vastgelegde uren moeten uitgesplitst kunnen worden per activiteit van het datamigratie-traject.

Daarnaast zullen kenmerken vastgelegd moeten worden waardoor datamigratietrajecten met elkaar vergeleken kunnen worden. De kenmerken die we onderkennen om trajecten onder te verdelen zijn: gebruikte methode, tools, teamsamenstelling en omgevingsfactoren. De productiviteit met tool A kan significant anders zijn dan tool B. Kenmerken van een datamigratietraject zijn vaak ook weer productiviteitsfactoren, in vervolgonderzoek zal dit verder uitgewerkt moeten worden.

Hieronder worden de datamigratieproductiviteitsfactoren die op dit moment onderkend zijn beschreven. Dit zijn performance en aantal rijen (1), verdeling automatisch en handmatig (2), privacy (3), tools (4), scenario (5), platform (6), resources (7), beschikbare kennis (8) en doorlooptijd (9).

- 1) Performance en aantal rijen: hiermee wordt de doorlooptijd bedoeld waarmee een deel van de datamigratie wordt uitgevoerd. Voor deze factor wordt aangegeven of de performance de doorlooptijd beïnvloed heeft en zo ja met hoeveel uur. De verwachting is dat een datamigratie van gegevensverzamelingen met veel rijen meer inspanning kost dan een datamigratie met weinig rijen. De te ontwikkelen tool is waarschijnlijk hetzelfde ongeacht het aantal rijen. Het aantal unieke voorkomens van verschillende soorten uitval neemt waarschijnlijk ook toe met het aantal rijen. Waardoor schoning meer inspanning kost bij een groter aantal rijen. De looptijd van een datamigratiecyclus, de tijd nodig om de data van bron- naar doelsysteem over te hevelen, zal met een groter aantal rijen langer duren dan bij een klein aantal rijen. Dit heeft consequenties voor de doorlooptijd van het project en wachttijd voordat nieuwe testresultaten bekend zijn. Indien de doorlooptijd langer duurt dan de beschikbare tijd op het moment van naar productie gaan, zal aan de performance gewerkt moeten worden. Technische oplossingen zijn mogelijk nodig om de doorlooptijd te versnellen.
- 2) Automatische versus handmatige datamigratie: niet alle gegevens worden automatisch overgezet. Als een beoordeling nodig is per geval dan zal dat deel van de datamigratie handmatig plaatsvinden. Indien er handmatig en geautomatiseerd gemigreerd wordt, zal inzichtelijk moeten worden gemaakt welk deel van de functiepunten handmatig en welk deel geautomatiseerd gemigreerd wordt. Handmatige datamigratie is over het algemeen arbeidsintensiever, maar bij een datamigratie van kleine omvang en grote complexiteit kan een handmatige datamigratie voordeliger zijn. Voor het handmatige en automatische deel zal het aantal bestede uren beschikbaar moeten zijn.
- 3) Privacy: dit zijn gegevens waar voorzichtig mee omgegaan moeten worden in verband met privacy. Ook geheime gegevens horen hierbij. Mogelijk moeten gegevens worden versleuteld waardoor lastig te bepalen is of de gegevens voldoen. Bovendien kost het tijd om de gegevens te versleutelen.
- 4) Tools: hier gaat het om met welke tools de datamigratiesoftware ontwikkeld is. Zijn er datamigratietools ingezet tijdens het datamigratietraject of zijn maatwerktools ontwikkeld. Welk deel van de datamigratie is met maatwerk en standaard datamigratietools. Inzetten van meerdere tools kost inspanning en mogelijk extra inter-

faces om de tools op elkaar aan te sluiten.

- 5) Scenario: hoeveel iteraties zijn doorlopen, is er big bang of gefaseerd gemigreerd, volledige of gedeeltelijke datamigratie. Leg bij gedeeltelijke datamigratie vast welke eisen zijn gesteld voor het wel of niet meenemen van data.
- 6) Platform: beschrijving architectuur en techniek van bron- en doelsysteem. Een SAP omgeving en een Oracle database worden beide anders gevuld, hebben een ander aantal tabellen, zijn anders te benaderen. Het aantal doel- en bronsystemen dat betrokken is in het datamigratietraject beïnvloedt de productiviteit.
- 7) Resources: het is belangrijk te weten hoeveel resources ingezet zijn, hoe ervaren de resources zijn en de projectstructuur. Een groot team of meerdere teams kost extra inspanning in aansturing en er is meer overleg nodig tussen de teams.
- 8) Beschikbare kennis: hoe makkelijk was het om kennis te krijgen van bron- en doelsysteem. De productiviteit wordt erg laag indien het veel moeite kost om kennis te krijgen van de systemen.
- 9) Doorlooptijd: de doorlooptijd van het datamigratietraject kan een indicator zijn dat veel gedaan moest worden in een korte tijd waardoor extra tijd nodig is voor projectmanagement. Een lange doorlooptijd kan erop duiden dat veel tijd nodig was om de geëiste kwaliteit te halen, veel iteraties nodig zijn en extra projectmanagement uren.

Naast bovengenoemde factoren zal per datamigratietraject vastgelegd moeten worden welke knelpunten in het datamigratietrajecten opgetreden zijn en achtergrondinformatie over het datamigratietraject. Met deze informatie kunnen uitschieters in productiviteit verklaard worden en kunnen datamigratietrajecten die vanwege bepaalde knelpunten of ontwikkelingen in het datamigratietrajecten niet representatief zijn weggelaten worden bij het ontwikkelen van kengetallen voor productiviteit.

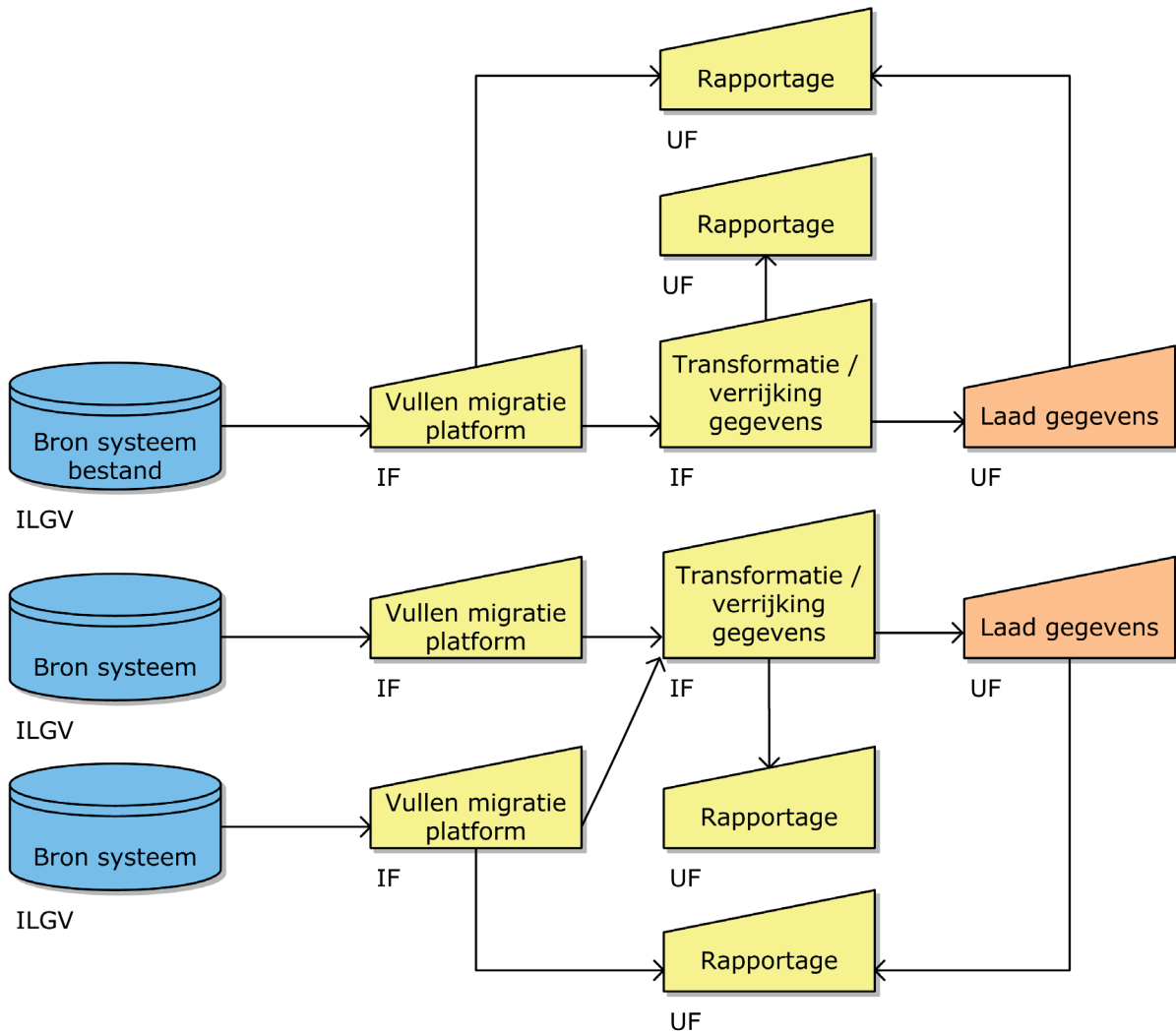
6.6 Telecom case

In deze paragraaf worden drie scenario's bekeken om voor een groot internationaal telecom bedrijf de omvang van een datamigratietraject met functiepuntanalyse uit te voeren. Vooraf was bekend dat het doelsysteem bestond uit 80 tabellen en het bronsysteem uit 250 tabellen. De partij die op dit moment het bronsysteem onderhoudt, levert de geëxtraheerde gegevens aan die nodig zijn voor datamigratie. Daarnaast wil het telecom bedrijf zelf de gegevens laden in het doelsysteem. Binnen het telecom bedrijf worden detailfunctiepunttellingen uitgevoerd als met functiepunten wordt gewerkt. In deze case wordt van alle gebruikersfuncties de complexiteit meegenomen in de bepaling van het aantal functiepunten.

Het uitgangspunt dat 250 brontabellen (scenario 1) de input vormen levert veel functiepunten op waar relatief weinig moeite voor gedaan hoeft te worden. Het is goed mogelijk dat de toeleverancier deze gegevens zelf al groepeerd in bijvoorbeeld 150 (scenario 2) of zelfs 80 (scenario 3) invoerbestanden. De kosten zullen ongeveer gelijk blijven aangezien ongeveer dezelfde gegevens geleverd worden alleen in een ander aantal brongegevensverzamelingen.

6.6.1 Scenario 1

In dit scenario gaan we ervan uit dat vanuit het bronsysteem 250 tabellen geëxtraheerd worden die als bron voor de datamigratie dienen. In alle scenario's gaan we uit dat 30 transformatie-functies nodig zijn. Dit aantal is bepaald door 25-50% van het aantal doeltabellen te nemen, voor 80 doeltabellen is dat 20-40 dat is gemiddeld 30 transformatiefuncties. Dit levert de nodige gebruikersfuncties op, zie figuur 11. Voor scenario 2 en 3 is geen figuur opgesteld.



Figuur 11. Migratie functies telecom case

Op basis van deze gegevens bepalen we de volgende aantallen:

- Logische gegevensverzamelingen = 250 invoertabellen
- Gebruikersfuncties zijn onder te verdelen in
 - 250 invoer functies
 - 30 transformatie functies
 - 80 uitvoer functies
- Controle raamwerk
 - Kwaliteitsrapporten over de inhoudelijke datamigratie
 - 250 invoer controle rapporten
 - 30 transformatie controle rapporten

- Kwantiteitsrapporten over de aantallen tijdens datamigratie
 - 250 invoer controle rapporten
 - 30 transformatie controle rapporten
 - 80 uitvoer controle rapporten

Om dat in functiepunten uit te drukken worden aannames gedaan over de complexiteit van functies dit is ook meegenomen in de uitwerking in functiepunten.

- Logische gegevensverzamelingen = 2125 FP
 - 150 x eenvoudige complexiteit (7 FP) = 1050 FP
 - 75 x gemiddelde complexiteit (10 FP) = 750 FP
 - 25 x moeilijke complexiteit (15 FP) = 325 FP
- Gebruikersfuncties = 1490 FP
 - 250 invoer functies x eenvoudige complexiteit (3 FP) = 750 FP
 - 30 transformatie functies x moeilijke complexiteit (6 FP) = 180 FP
 - 80 uitvoer functies x moeilijke complexiteit (7 FP) = 560 FP
- Controle raamwerk, kwaliteitsrapporten over de inhoudelijke datamigratie (1210 FP)
 - 250 invoer controle rapporten x eenvoudige complexiteit (4 FP) = 1000 FP
 - 30 transformatie controle rapporten x moeilijke complexiteit (7 FP) = 210 FP
- Controle raamwerk, kwantiteitsrapporten over aantallen in de datamigratie (1610 FP)
 - 250 invoer controle rapporten x eenvoudige complexiteit (4 FP) = 1000 FP
 - 30 transformatie controle rapporten x moeilijke complexiteit (7 FP) = 210 FP
 - 80 uitvoer controle rapporten x gemiddelde complexiteit (5 FP) = 400 FP

Totaal aantal functiepunten is 6435 FP.

Tijdens verkennende gesprekken is bepaald dat er een vaste prijs per functiepunt berekend wordt. De datamigratie zal gebeuren met een migratieworkbench. Er zijn zes testruns nodig voor de datamigratie richting productie kan. Aangezien er op dit moment nog te weinig gegevens voorradig zijn om een goede prijs per functiepunt te bepalen is besloten een prijs te bepalen middels de traditionele methodieken. Deze schatting van de kostprijs is voor het uitvoeren van het datamigratietraject zoals beschreven door de klant. Daarnaast worden functiepunten bepaald om zo een prijs per functiepunt op te stellen. De kosten die nodig zijn om het datamigratietraject uit te voeren worden geschat op 460K€ aan WBS activiteiten en 400K€ aan FPA activiteiten. De klant gebruikt de prijs per functiepunt om de leverancier te kiezen en bij aanpassingen in het project de meer en minderwerk procedure te ondersteunen.

Voor 80 bestanden komt dit uit op 6435 FP en 860 K€ op basis van deze gegevens kunnen we een aantal kengetallen bepalen. Cijfers zijn afgerond op gehele getallen.

Kosten per bestand = $860 \text{ K€} / 80 \text{ bestanden} = 10.75 \text{ K€}$

Gemiddeld aantal FP per bestand = $6435 \text{ FP} / 80 \text{ bestanden} = 80 \text{ FP}$

Gemiddeld aantal euro per functiepunt = $860 \text{ K€} / 6435 \text{ FP} = 134 \text{ €}$, bestaand uit $460 \text{ K€} / 6435 \text{ FP} = 72 \text{ €}$ voor WBS en $400 \text{ K€} / 6435 \text{ FP} = 62 \text{ €}$ voor FPA.

6.6.2 Scenario 2

In dit scenario gaan we ervan uit dat het aantal aangeleverde brontabellen 150 is, het aantal te vullen doeltabellen blijft 80.

Op basis van deze gegevens bepalen we de volgende aantallen:

- Logische gegevensverzamelingen = 150 invoer tabellen
- Gebruikersfuncties zijn onder te verdelen in
 - 150 invoer functies
 - 30 transformatie functies
 - 80 uitvoer functies
- Controle raamwerk
 - Kwaliteitsrapporten over de inhoudelijke datamigratie
 - 150 invoer controle rapporten
 - 30 transformatie controle rapporten
 - Kwantiteitsrapporten over de aantallen tijdens datamigratie
 - 150 invoer controle rapporten
 - 30 transformatie controle rapporten
 - 80 uitvoer controle rapporten

Om dat in functiepunten uit te drukken worden aannames gedaan over de complexiteit van functies dit is ook meegenomen in de uitwerking in functiepunten.

- Logische gegevensverzamelingen = 1305 FP
 - 90 x eenvoudige complexiteit (7 FP) = 630 FP
 - 45 x gemiddelde complexiteit (10 FP) = 450 FP
 - 15 x moeilijke complexiteit (15 FP) = 225 FP
- Gebruikersfuncties = 1190 FP
 - 150 invoer functies x eenvoudige complexiteit (3 FP) = 450 FP
 - 30 transformatie functies x moeilijke complexiteit (6 FP) = 180 FP
 - 80 uitvoer functies x moeilijke complexiteit (7 FP) = 560 FP
- Controle raamwerk, kwaliteitsrapporten over de inhoudelijke datamigratie (810 FP)
 - 150 invoer controle rapporten x eenvoudige complexiteit (4 FP) = 600 FP
 - 30 transformatie controle rapporten x moeilijke complexiteit (7 FP) = 210 FP
- Controle raamwerk, kwantiteitsrapporten over aantallen in de datamigratie (1210 FP)
 - 150 invoer controle rapporten x eenvoudige complexiteit (4 FP) = 600 FP
 - 30 transformatie controle rapporten x moeilijke complexiteit (7 FP) = 210 FP
 - 80 uitvoer controle rapporten x gemiddelde complexiteit (5 FP) = 400 FP

Totaal aantal functiepunten is 4515 FP.

Voor 80 bestanden komt dit uit op 4515 FP en 860 K€ op basis van deze gegevens kunnen we een aantal kengetallen bepalen. Cijfers zijn afgerond op gehele getallen.

Kosten per bestand = $860 \text{ K€} / 80 = 10.75 \text{ K€}$

Gemiddeld aantal FP per bestand = $860 \text{ K€} / 4515 \text{ FP} = 56 \text{ FP}$

Gemiddeld aantal euro per functiepunt = $860 \text{ K€} / 4515 \text{ FP} = 191 \text{ €}$, bestaand uit $460 \text{ K€} / 4515 \text{ FP} = 102 \text{ €}$ voor WBS en $400 \text{ K€} / 4515 \text{ FP} = 89 \text{ €}$ voor FPA.

6.6.3 Scenario 3

In dit scenario gaan we ervan uit dat het aantal aangeleverde brontabellen 80 is, het aantal te vullen doeltabellen blijft 80.

Op basis van deze gegevens bepalen we de volgende aantallen:

- Logische gegevensverzamelingen = 80 invoer tabellen
- Gebruikersfuncties zijn onder te verdelen in
 - 80 invoer functies

- 30 transformatie functies
- 80 uitvoer functies
- Controle raamwerk
 - Kwaliteitsrapporten over de inhoudelijke datamigratie
 - 80 invoer controle rapporten
 - 30 transformatie controle rapporten
 - Kwantiteitsrapporten over de aantallen tijdens datamigratie
 - 80 invoer controle rapporten
 - 30 transformatie controle rapporten
 - 80 uitvoer controle rapporten

Om dat in functiepunten uit te drukken worden aannames gedaan over de complexiteit van functies. Dit is ook meegenomen in de uitwerking in functiepunten.

- Logische gegevensverzamelingen = 675 FP
 - 50 x eenvoudige complexiteit (7 FP) = 350 FP
 - 25 x gemiddelde complexiteit (10 FP) = 250 FP
 - 5 x moeilijke complexiteit (15 FP) = 75 FP
- Gebruikersfuncties = 980 FP
 - 80 invoer functies x eenvoudige complexiteit (3 FP) = 240 FP
 - 30 transformatie functies x moeilijke complexiteit (6 FP) = 180 FP
 - 80 uitvoer functies x moeilijke complexiteit (7 FP) = 560 FP
- Controle raamwerk, kwaliteitsrapporten over de inhoudelijke datamigratie (530 FP)
 - 80 invoer controle rapporten x eenvoudige complexiteit (4 FP) = 320 FP
 - 30 transformatie controle rapporten x moeilijke complexiteit (7 FP) = 210 FP
- Controle raamwerk, kwantiteitsrapporten over aantallen in de datamigratie (930 FP)
 - 80 invoer controle rapporten x eenvoudige complexiteit (4 FP) = 320 FP
 - 30 transformatie controle rapporten x moeilijke complexiteit (7 FP) = 210 FP
 - 80 uitvoer controle rapporten x gemiddelde complexiteit (5 FP) = 400 FP

Totaal aantal functiepunten is 3115 FP.

Voor 80 bestanden komt dit uit op 3115 FP en 860 K€ op basis van deze gegevens kunnen we een aantal kengetallen bepalen. Cijfers zijn afgerond op gehele getallen.

Kosten per bestand = $860 \text{ K€} / 80 = 10.75 \text{ K€}$

Gemiddeld aantal FP per bestand = $860 \text{ K€} / 3115 \text{ FP} = 39 \text{ FP}$

Gemiddeld aantal euro per functiepunt = $860 \text{ K€} / 3115 \text{ FP} = 276 \text{ €}$, bestaand uit $460 \text{ K€} / 3115 \text{ FP} = 147 \text{ €}$ voor WBS en $400 \text{ K€} / 3115 \text{ FP} = 129 \text{ €}$ voor FPA

Discussie

Het grote verschil in de drie scenario's is het aantal meegenomen brontabellen. Hieruit kan geconcludeerd worden dat het aantal invoerbestanden het aantal functiepunten beïnvloedt. Om een vaste prijs per functiepunt af te spreken dient het aantal brontabellen en doelbestanden bekend te zijn. In dit voorbeeld is complexiteit van functies meegenomen om een realistischer beeld van de omvang te krijgen. Indien het aantal bron- en doel tabellen kleiner wordt, neemt de invloed van het aantal functiepunten vanwege complexiteit af. Productiviteitsfactoren zoals beschreven in paragraaf 6.5 zijn niet gebruikt om de kosten per functiepunt te bepalen. Doordat de prijs plus aantal doeltabellen niet wijzigt in de case blijven de kosten per bestand gelijk.

6.7 Conclusie

In dit hoofdstuk is de afbeelding gemaakt voor welke activiteiten in het Mediaan Migratie Methode (MMM) FPA bruikbaar is en voor welke niet. Voor de activiteiten waar FPA niet geschikt is kan WBS gebruikt worden.

Er is een vast aantal functiepunten gevonden om een globale functiepunttelling uit te voeren bij datamigratietrajecten. Op basis van ETL en rapportages worden datamigratietrajecten uitgevoerd, hierdoor is een model te maken van de gebruikersfuncties en logische gegevensverzamelingen. Indien in het offertetraject van het datamigratietraject helemaal geen detailinformatie bekend is zal het aantal functiepunten per doelgegevensverzameling gelijk zijn (namelijk 32 FP). Indien op hoofdlijnen bekend is wat moet gebeuren kan een globale functiepuntanalyse doorgevoerd worden. Dan kan geen complexiteit gehangen worden aan de gebruikersfuncties en zijn alle gebruikersfuncties even complex. Dit betekent dat op basis van functiepunten de migratie van het de ene gegevensverzameling even complex en even groot is in omvang binnen hetzelfde informatiesysteem als de andere.

Het aantal invoer bestanden heeft een grote invloed op het aantal functiepunten van het datamigratietraject.

Het verbeteren van de kwaliteit van data door schoning dient buiten de begroting van het datamigratietraject gehouden te worden. De reden hiervoor is dat het vaststellen van de kwaliteit van data vooraf op dit moment nog niet mogelijk is en de meeste inspanning in verbeteren van kwaliteit gedaan wordt door de gebruikersorganisatie. Grote afwijkingen in specificatie vanwege datakwaliteit wordt opgevangen in changemanagement.

Projectmanagement bij datamigratieprojecten is erg belangrijk vanwege het grote aantal mensen uit de gebruikersorganisatie die nodig zijn bij een datamigratietraject en de vaak grote onduidelijkheden bij de start van het datamigratietraject.

Het doel van het onderzoek was om een formule op basis van FPA te ontwikkelen om datamigratietrajecten te begroten. Het is niet gelukt om factoren als aantal rijen per gegevensverzameling, mate van vervuiling, te gebruiken tools, rapportage-eisen generiek te maken en daardoor een factor over FPA heen te leggen zoals wel gebruikelijk is in nieuwbouwprojecten. Deze factoren moeten voorsnog los gezien worden van functiepunten en zullen meegenomen moeten worden in de productiviteit per functiepunt.

Ondersteuning in wat de productiviteit binnen migratietrajecten is, is in dit onderzoek niet gegeven, omdat er te weinig gegevens beschikbaar zijn. Het is helaas niet mogelijk om betrouwbare gegevens achteraf te verzamelen van de uitgevoerde projecten. De redenen hiervoor zijn dat gegevens over aantal functiepunten niet beschikbaar zijn en niet meer te achterhalen zijn. Bestede uren zijn niet toegespitst per activiteit en niet van alle teamleden zijn de uren te achterhalen.

Indien productiviteit gemeten is binnen meerdere datamigratietrajecten, binnen een gelijksoortige omgeving met hetzelfde werkproces dan is functiepunten wel degelijk een zeer bruikbaar hulpmiddel bij begroten van datamigratietrajecten.

7 Empirisch onderzoek

Om de voorgestelde methodiek te toetsen op bruikbaarheid is een datamigratietraject als voorbeeld genomen bij een groot internationaal bedrijf dat producten en diensten levert op het gebied van document- en informatiemanagement.

Het doel van het project is het implementeren van nieuwe processen en systemen, o.a. gebaseerd op SAP, om binnen alle Europese landen waar het bedrijf opereert naar een standaard manier van werken over te gaan.

Het empirisch onderzoek richt zich op het datamigratiedeel binnen het implementatieproject voor de werkmaatschappij Nederland. De start van het datamigratietraject was 1 september 2007, de geplande implementatiedatum was 2 juni 2008, de einddatum van het datamigratietraject stond gepland voor 30 juni 2008 i.v.m. nazorg en afronden rapportages. Uiteindelijk is besloten om de implementatiedatum uit te stellen tot 1 december 2008, o.a. vanwege het niet tijdig kunnen trainen van alle eindgebruikers. Het uitstel heeft niets met datamigratie te maken en is dus niet verwerkt in de project-details t.a.v. bestede en geplande uren.

Het datamigratietraject omvat acht procesgebieden. In totaal zijn het 72 objecten die gemigreerd moeten worden. De te migreren gegevens komen uit vijf bronsystemen en twee doelsystemen moeten gevuld worden.

7.1 Migratieaanpak

Als methode is de Mediaan Migratie Methode gebruikt. Alle stappen van het model zijn een of meerdere keren doorlopen. De planning is opgesteld met behulp van een WBS. Er is geen FPA uitgevoerd tijdens het datamigratietraject.

Als aanpak is gekozen voor semi big bang, wel in één keer alles over, maar in twee stappen. De stamgegevens van leveranciers, medewerkers, klanten, en het productassortiment worden ca. twee weken voor het datamigratieweekend al gemigreerd en vanaf dat moment met interfaces actueel gehouden. De proces- en transactiegegevens worden in het datamigratieweekend gemigreerd. Dit betekent dat de oude systemen niet meer nodig zijn na de tweede datamigratiestap.

7.2 Migratietool keuze

Export van gegevens uit de bestaande systemen wordt deels via interfaces, deels via reeds bestaande datamigratieprogrammatuur en deels via nieuwe datamigratieprogrammatuur geregeld. De interfaces gaan in het nieuwe systeemlandschap het uitwisselen van masterdata regelen. Reeds bestaande datamigratieprogrammatuur kan hergebruikt worden, omdat in 2006 een soortgelijk datamigratietraject is uitgevoerd in de Duitse vestiging. De nieuwe datamigratieprogrammatuur moet opnieuw worden ontworpen, ontwikkeld en getest.

De geëxporteerde stamgegevens worden verwerkt via bestaande middleware. SAP heeft

standaardmodules voor deze stamtabellen. De bedrijfsspecifieke gegevens worden via de Migration Workbench (Koorneef, 2008) verwerkt. Dit is een datamigratietool dat op basis van platte tekstbestanden gegevens importeert in een MS Access database en door het instellen van vele parameters in cobol migratiecode regels genereert. Daarnaast is gebruik gemaakt van de SAP Legacy System Migration Workbench (LSMW).

Het laden van gegevens in doelsystemen gebeurt deels via interfaces die in het nieuwe systeemlandschap het uitwisselen van stamgegevens regelen en deels via reeds bestaande datamigratieprogrammatuur (standaard SAP LSMW tools voor het laden van gegevens). Hierbij waren wel aanpassingen noodzakelijk om de programmatuur in te richten.

Diverse rapportages zijn ontwikkeld om voortgang en de volledigheid en juistheid van gemigreerde gegevens vast te leggen.

7.3 Uitvoering

In de datamigratiefases vooronderzoek en definitie is de scope bepaald, een aanpak bepaald, is een toolselectie gedaan, is het project ingericht in de organisatie en zijn mensen aangetrokken om de datamigratie uit te voeren. De doorlooptijd van deze fases was van 1 september 2007 t/m 30 september 2007 en twee personen hebben fulltime gewerkt aan deze fases van het datamigratietraject. De totale inspanning was 320 uur, onder te verdelen in 160 uur projectmanagement en 160 uur datamigratiespecialist. Deze korte doorlooptijd is bereikt doordat binnen hetzelfde bedrijf een soortgelijke datamigratie al voor de Duitse vestiging gedaan is in 2006. Hierdoor was al veel informatie bekend en zijn de uren niet representatief voor andere datamigratietrajecten.

De datamigratiefases ontwerp, realisatie en test zijn gestart op 1 oktober 2007 en zijn afgerond op 31 januari 2007. Dit omvat het deel: ontwerpen, bouwen en testen van de datamigratietools, de schoningstools, het datamigratiedraaiboek, tijdelijke infrastructuur en interfaces en het controle raamwerk. Testen gebeurde in drie proefdatamigraties: cycle 1, cycle 2 en cycle 3.

Functionele omvang:

Bron: 161 gegevenstabellen,

Doel: 92 gegevenstabellen,

Totaal 72 objecten die gemigreerd worden, waarvan 8 via interfaces en 64 via workbench

Totale inspanning in deze periode inclusief project en changemanagement was 5292 uur. Deze uren zijn onder te verdelen in 2040 uur project en changemanagement en 3252 uren ontwerp, realisatie en test.

De preproductie omvat een volledige proefmigratie naar een acceptatie testomgeving (cycle 4) en een extra cycle 5 om de laatste issues op te lossen. Totale inspanning in deze periode inclusief project en changemanagement was 3880 uur. Deze uren zijn onder te verdelen in 1488 uren project en changemanagement 1196 uren onderhoud migration workbench en interfaces en 1196 uren in uitvoeren van de daadwerkelijke datamigratie (twee keer uitgevoerd van 1 feb t/m 30 april).

De productiedatamigratie omvat de volledige datamigratie naar de productieomgeving,

de nazorg en het rapporteren van het resultaat t.b.v. audit en accountantsdienst. De totale geplande inspanning in deze periode inclusief project en changemanagement was 2400 uur. Deze uren zijn onder te verdelen in 1080 uren project en changemanagement en 1320 uren in uitvoeren van de datamigratie. (1 mei t/m 30 juni)

7.4 Analyse

Het hele datamigratieproject bevat meer dan duizend pagina's aan specificaties over hoe velden gemigreerd moeten worden. Dit is niet in een aantal dagen volledig te analyseren. Datamodellen van bron- en doelsystemen zijn er niet, omdat het deels SAP systemen zijn en deels beperkt gedocumenteerde maatwerksystemen zijn.

In dit traject is de MMM gebruikt en deze methode is geschikt gebleken voor het uitvoeren van datamigratietrajecten. Alle activiteiten van het model zijn gebruikt en er zijn diverse iteratie stappen uitgevoerd, omdat meerdere testruns in het traject nodig waren. De projectleiders hebben aangegeven deze methode in volgende projecten weer te gebruiken. De planning is in dit datamigratietraject volledig met WBS gedaan, maar de projectleiders geven aan dat ondersteuning door middel van functiepunten een nuttige toevoeging is.

Ook in dit traject gebruikt men de methode van Extractie, Transformatie en Laden (ETL) plus rapporteren over juist en volledig overzetten van data.

Wie hoeveel uren heeft besteed is bekend, maar waaraan de uren precies besteed zijn is niet allemaal bekend. Het totale aantal uren in het project is bijna 12000, ongeveer twee derde daarvan is besteed aan ontwerp, realisatie, test en onderhoud van programmatuur voor het verwerken van gegevens uit het SAP systeem.

De uren in de fasen vooronderzoek en definitie zijn uren die we onder projectmanagement rekenen. De uren in de fasen ontwerp, realisatie en test worden verdeeld naar projectmanagement, rapportage, FPA gerelateerde uren en migration workbench. De meewerkend voorman uren worden bij de FPA gerelateerde uren geteld. De FPA gerelateerde uren zijn uren besteed zijn aan interfaces, extractie, transformatie en laadprocedures en rapportages. De uren van de migration workbench worden apart gehouden, omdat via deze workbench migratiesoftware gemigreerd wordt via een code generator. In de (pre)productie fase zijn diverse testruns gedraaid waaruit bevindingen gekomen zijn die opgelost moesten worden. Deze uren zijn op dezelfde manier verdeeld als de ontwerp, test en realisatie fase. In tabel 3 is de verdeling van fase, activiteit en uren weergegeven.

Migratie fase	Activiteit	Bestede uren	Subtotaal
Vooronderzoek + Definitie	Projectmanagement	160	320
	Vooronderzoek + Definitie	160	
Ontwerp + Realisatie + Test	Projectmanagement	1020	5324
	Meewerkend voorman	680	
	Rapportage	340	
	Interface werk bronsystemen	1660	
	Laden sap	984	
	Migration workbench	608	
PreProductie	Projectmanagement	843	3992
	Meewerkend voorman	298	
	Rapportage	347	
	Interface werk bronsystemen	1176	
	Laden sap	912	
	Migration workbench	416	
Productie	Projectmanagement	792	2448
	Rapportage	288	
	Interface werk bronsystemen	528	
	Laden sap	648	
	Migration workbench	192	
Totaal			12084

Tabel 3. Bestede uren datamigratietraject

In tabel 4 zijn de uren verdeeld in projectmanagement, FPA gerelateerd en migratie workbench gerelateerd. Op basis van deze gegevens kan formule 1 ingevuld worden en hieruit komt dat 75% met behulp van FPA bepaald wordt en 25% met WBS.

Migratie onderwerp	Uren	%	
Projectmanagement	2975	24,6	WBS
FPA gerelateerd	7861	65,1	FPA
Migratie workbench	69	10,3	FPA
Totaal	12084	100,0	WBS/FPA

Tabel 4. Uren per migratieonderwerp

De bestanden die via de migration workbench verwerkt worden zijn wel inzichtelijk en ook de totale uren die besteed zijn door projectmedewerkers. In tabel 5 zijn de acht business proces areas die worden verwerkt weergegeven. Voor elke area is aangegeven hoeveel bestanden er in de workbench verwerkt worden plus de tijd die besteed is aan inrichten en testen om de datamigratie met behulp van migration workbench uit te voeren.

Area	Bestanden	Uren	Uren / bestand
1	23	127	5,5
2	5	52	10,4
3	11	69	6,3
4	8	100	12,5
5	10	185	18,5
6	54	454	8,3
7	6	105	17,5
8	16	156	9,8
Totaal	133	1248	9,38

Tabel 5. Bestede uren migration workbench

Gemiddeld komt dit uit op 9,38 uren per bestand die via de MW uitgevoerd zijn. Hoe de bestanden te verdelen zijn over de 64 gegevens verzamelingen is in dit empirisch onderzoek niet verder onderzocht.

Voor dit datamigratietraject wordt op het einde van het project op basis van de verzamelde cijfers de calculatieformule ingevuld. In de formule 2 is bekend dat Q 12084 uren is. Hiervan zijn 2975 uren besteed in activiteiten die niet via FPA te berekenen zijn. Deze uren zullen via WBS gepland moeten worden. Dit is de ΣA van de formule. Uren besteed aan activiteiten die wel met functiepunten te meten zijn is 9109 uren. Dit is het totaal van $P \times FP$. Het totaal aantal functiepunten wordt middels een "globale datamigratiefunctiepunten telling" bepaald. Aangezien over alle gegevensverzamelingen gewerkt wordt met ETL en wordt gerapporteerd naar de opdrachtgever gaan we uit van 32 FP per gegevensverzameling. De omvang in functiepunten komt uit op 32 functiepunten * 72 gegevensverzamelingen = 2304 functiepunten. Deze 2304 fp zijn te verdelen over 64 gegevensverzamelingen (2048 fp) die via de migration workbench gemigreerd zijn en 8 (256 fp) die via interfaces gemigreerd zijn.

De productiviteit per functiepunt komt uit op bijna 4 uur per functiepunt over het gehele traject. Van de 9109 uren die besteed zijn aan met FPA te calculeren activiteiten is 1248 uren besteed aan werkzaamheden die uitgevoerd konden worden met de migration workbench en migratie via interfaces 7861 uren. Dit betekent dat de productiviteit van de workbench 2048 fp / 1248 uur is en via interfaces 256 fp / 7861 uur. Met de migration workbench is dat ongeveer 0,5 uur per functiepunt en met interfaces is dat 30 uur per functiepunt.

Ongeveer een kwart is besteed aan projectmanagement en deze werkzaamheden kunnen met WBS begroot worden. Een goed WBS is daardoor van cruciaal belang bij datamigratietrajecten.

7.5 Conclusie

Alle in dit datamigratietraject uitgevoerde activiteiten zijn beschreven in de Mediaan Migratie Methode (MMM). Dit empirisch onderzoek onderbouwt dat de MMM alle uit te voeren activiteiten in een datamigratietraject beschrijft en dat het volgen van deze

aanpak aan te bevelen is.

De planning van werkzaamheden is gedaan door een WBS op te stellen en bij te stellen indien nodig. De projectmanager heeft hiermee voldoende tools in handen gehad om te sturen op het project. Werken met een WBS in datamigratietrajecten heeft voordelen. Door FPA toe te voegen kan nog een extra slag gemaakt worden in planning. In de analyse is achteraf op basis van gegevens uit het project een FPA gemaakt. Hieruit blijkt dat de productiviteit hoger ligt dan als een bestaand datamigratietool gebruikt wordt in plaats van zelf migratie tools ontwikkelen.

Gezien het aantal uren dat aan management activiteiten besteed is kan geconcludeerd worden dat projectmanagement en changemanagement in een datamigratietraject van groot belang is.

8 Conclusies en aanbevelingen

Tijdens het literatuuronderzoek is literatuur bestudeerd over datamigratietrajecten door diverse algemeen bekende methoden voor informatiesysteem ontwikkeling en beheer te bestuderen. Met als doel een proces te beschrijven dat gebruikt kan worden om een datamigratie uit te voeren.

Literatuuronderzoek naar een procesbeschrijving van een datamigratietraject in standaard beheersmethoden en software engineering heeft geen procesbeschrijving opgeleverd. In het bestudeerde boek van software engineering is geen informatie gevonden over datamigratie. Het onderwerp is blijkbaar niet tot weinig beschreven in de literatuur, wat het onderwerp uitermate geschikt maakt voor onderzoek.

De in de literatuur gevonden methode extract transform load (ETL) die gebruikt wordt om datawarehouses te vullen is een bruikbare methode voor datamigratie. Vormen van transformatie bij datamigratie zijn verrijking en schoning. Een toevoeging aan standaard ETL is het controle raamwerk, dit is het geheel aan rapportages om aan te tonen dat de datamigratie volledig en juist uitgevoerd is.

De literatuur van functiepuntanalyse is bestudeerd in het onderzoek en er is uitgelegd hoe functiepuntanalyse (FPA) gebruikt wordt. FPA is een methode om de functionele omvang van informatiesystemen vanuit gebruikers perspectief te meten in omvang. FPA wordt vooral gebruikt bij nieuwbouw en onderhoud, tijdens ontwerp, ontwikkeling, test en beheer van administratieve transactieverwerkende systemen. De telmethode van FPA kan uitgevoerd worden op globale en detailspecificaties. Indien geteld wordt met globale specificaties wordt niet gewerkt met complexiteitsfactoren. Door te werken met FPA kunnen ervaringscijfers per datamigratietraject worden verzameld. Met de ervaringscijfers kunnen op basis van functiepunten kengetallen ontwikkeld worden voor productiviteit, kwaliteit en kosten. Hierdoor kunnen datamigratietrajecten met elkaar worden vergeleken.

Aangezien literatuuronderzoek niet opgebracht heeft wat we hoopten is een quickscan van de praktijk uitgevoerd door een aantal experts uit te nodigen om een vragenlijst in te vullen en deel te nemen aan een interview. Uit deze quickscan blijkt dat in elk datamigratietraject:

- a) data geschoond wordt;
- b) een deel van de data niet gemigreerd wordt;
- c) grote behoefte is aan bewijsvoering (rapportages) dat de datamigratie volledig en juist uitgevoerd is;
- d) meerdere testruns nodig zijn (minimaal drie en maximaal zes) om de kwaliteit productieklaar te maken;
- e) de opdrachtgever hoge eisen stelt om business uitval te voorkomen;
- f) inzetten van een migratietool de productiviteit verhoogd;
- g) planning belangrijk is voor een projectmatige aanpak;
- h) veel wijzigingen in scope en specificaties doorgevoerd worden. Changemanagement is daardoor heel erg belangrijk;
- i) calculatie over kosten gedaan wordt op basis van een prijs achteraf;
- j) geautomatiseerde gegevens worden overgezet, in een project zijn dossiers handmatig in het nieuwe systeem ingevoerd.

Door het uitvoeren van een quickscan is een best practice methode gevonden voor het uitvoeren van datamigratietrajecten. Deze best practice methode is de Mediaan Migratie Methode (MMM) en is ontstaan uit de ervaringen van een aantal migratie experts die bij Mediaan/abs werkzaam zijn. De methode bevat een beschrijving van de fasen en activiteiten in een datamigratietraject en kan gebruikt worden bij het uitvoeren van datamigratietrajecten. De methode is in meerdere (grote) datamigratietrajecten met succes toegepast en de resultaten van de uitgevoerde migratietrajecten waren zeer goed.

Een datamigratietraject doorloopt fasen volgens het waterval model. Net als in een traditioneel nieuwbouwproject wordt bij datamigratie begonnen met definitie en ontwerp, daarna bouw, test, pre-productietests en als laatste in productie name.

De eerste stappen zijn het bepalen van de scope van de datamigratie en de migratiestrategie. Bepalen van de scope is bepalen waaraan data moet voldoen om in aanmerking te komen voor migratie. De migratiestrategie bepaalt of de migratie in een keer volledig uitgevoerd wordt of in delen. De capaciteit van de gebruikersorganisatie bepaalt in grote mate de migratiestrategie.

Tijdens het datamigratietraject zijn interfaces nodig om datauitwisseling tussen bron-, doel- en hulpinformatiesystemen mogelijk te maken. Er is ook infrastructuur nodig om testomgevingen op te tuigen om tests te doen en pre-productieruns te draaien.

Niet alle activiteiten in een datamigratietraject zijn te begroten met behulp van functiepunten. De activiteiten die niet via FPA te begroten zijn, worden begroot met een work breakdown structure (WBS). Om te bepalen welke activiteiten van een datamigratietraject gecalculeerd kunnen worden met FPA is het van belang de uit te voeren activiteiten in beeld te hebben. Op basis van de MMM is vastgelegd welke activiteiten met behulp van FPA uitgevoerd kunnen worden en welke met WBS. Schoning wordt buiten de begroting van datamigratietrajecten gehouden.

Gebaseerd op de aanname dat voor elke doeltabel er minimaal één brontabel aanwezig moet zijn die de gegevens bevat om de doeltabel te vullen, is bepaald dat indien alleen het aantal doeltabellen bekend is, het aantal functiepunten bepaald kan worden door doeltabellen te vermenigvuldigen met 32 functiepunten. Dit komt overeen met een globale functiepuntentelling.

Ook zijn twee formules uitgewerkt om omvang van datamigratietrajecten te calculeren en is een case uitgewerkt op basis van een detailfunctiepuntentelling.

In het laatste hoofdstuk is een empirisch onderzoek uitgewerkt, waarin aangetoond wordt dat de MMM in de praktijk met succes gebruikt kan worden. Op basis van een globale functiepuntentelling achteraf blijkt dat gebruik van een datamigratietool een productiviteit biedt die vele malen hoger ligt dan het ontwikkelen van migratietools.

8.1 Onderzoeksvragen

In het onderzoek zijn twee onderzoeksvragen gedefinieerd. Deze onderzoeksvragen zijn:

- Kan FPA gebruikt worden om de omvang van datamigratietraject te bepalen?
- Hoe kunnen omvang, kosten en doorlooptijd begroot worden van onderdelen die van

een datamigratietraject die niet met FPA begroot kunnen worden?

Daarnaast zijn een aantal deelvragen gedefinieerd:

- Uit welke onderdelen bestaat een datamigratietraject?
- Wat is de invloed van de aanpak van de datamigratie op de kosten?
- Welke complexiteitsfactoren zitten in een datamigratietraject?
- Welke productiviteitsfactoren beïnvloeden een datamigratietraject?

Het antwoord op de onderzoeksvragen is dat FPA gebruikt kan worden voor een deel van het datamigratietraject en dat voor de delen die niet met FPA gecalculeerd kunnen worden de work breakdown structure (WBS) een goede aanpak is. De MMM beschrijft het hele datamigratietraject, figuur 8 laat op basis van de MMM zien voor welke onderdelen van een datamigratietraject middels FPA de omvang bepaald worden en voor welke onderdelen van een datamigratietraject niet middels FPA de omvang bepaald worden. WBS is de methode om een planning en schatting van kosten te maken voor de niet functiepunt gebonden activiteiten van een datamigratietraject.

De antwoorden op de deelvragen zijn:

- De activiteiten waaruit een datamigratietraject bestaat is uitgewerkt in de MMM (zie bijlage 2).
- Uit de quickscan is gebleken dat een goede aanpak een aantal risico's vermindert en de kans op slagen van de migratie verhoogt.
- Complexiteit wordt verwerkt in functiepunten volgens de NESMA telrichtlijnen en door inschatting van doorlooptijd in het WBS.
- De productiviteitsfactoren die een datamigratietraject beïnvloeden zijn beschreven in paragraaf 6.5.

8.2 Algemene conclusie

Vroeg starten met een datamigratietraject in een migratietraject is een must. De migratiestrategie en datakwaliteit zijn van grote invloed op het datamigratietraject. Brongegevens zijn nodig om de datakwaliteit te bepalen, deze zijn altijd beschikbaar. Specificaties van doelsysteem zijn pas op een bepaald moment in het migratietraject beschikbaar.

Datamigratie kan een bijdrage leveren aan de rest van het datamigratietraject, door betrouwbare productiegegevens op te leveren voor acceptatietesten. Door te testen met productiekwaliteitgegevens zal de acceptatie van het migratietraject toenemen.

De extract load transform (ETL) methode (figuur 1) is een goede manier om datamigratietrajecten aan te pakken.

Tijdens een datamigratietraject wijzigt meer dan 50% van de code meer dan één keer (Koorneef). Hierdoor is changemanagement een belangrijk onderdeel van een datamigratietraject.

Goed vastleggen van gegevens over bestede uren en geld toegespitst op activiteit is van belang om te meten of de geplande inspanning overeenkomt aan de uitgevoerde inspanning. Dan kan bepaald worden of de calculatiemethode juist is of dat bijstelling nodig

is. Een groot aantal datamigratietrajecten zal volgens de calculatiemethode nog uitgevoerd moeten worden om wetenschappelijk vast te stellen of de methode correct is.

In dit onderzoek is een veelbelovende datamigratiecalculatiemethode beschreven om met functiepunten kostprijs te bepalen.

Deze heeft zijn beperkingen, dit zijn:

- de productiviteit in functiepunten is nog onbekend
- methode moet nog getoetst worden door vele projecten te begroten en de resultaten te gebruiken voor validatie van de ontwikkelde methode

Daarnaast is in dit onderzoek de MMM beschreven die in meerdere datamigratietrajecten zijn waarde heeft bewezen. Door het werken met een gestructureerde datamigratiemethode is het makkelijker gegevens te vergelijken over verschillende projecten en ontstaat een coherent begrippenkader.

WBS is een methode om niet-functiepunt gebonden activiteiten te plannen. Door achteraf de werkelijke kosten in tijd (en geld) te bepalen kan de WBS door ervaring bijgesteld worden en kunnen na een groot aantal projecten kengetallen bepaald worden. Hieruit blijkt het belang van goed documenteren van bestede tijd en kosten per activiteit.

8.3 Aanbeveling voor vervolgonderzoek

Het verzamelen van gegevens over uitgevoerde werkzaamheden maakt het mogelijk om verder onderzoek te doen naar productiviteit en het bepalen van ervaringscijfers en kengetallen. Om dit te bereiken zullen vele datamigratietrajecten uitgevoerd moeten worden met de in dit onderzoek beschreven aanpak. Door in alle datamigratietrajecten de Mediaan Migratie Methode toe te passen zijn gegevens uit verschillende datamigratietrajecten met elkaar te vergelijken. Door de niet aan functiepunten gerelateerde activiteiten via een vast WBS te begroten kan de betrouwbaarheid van het WBS vergroot worden. Daardoor wordt het ook mogelijk om correlatie van te besteden uren tussen activiteiten te bespelen. Voor dat gestart wordt met het verzamelen van gegevens zal vastgelegd moeten worden welke gegevens verzameld moeten worden.

De methode van kostprijs bepalen op basis van functiepunten zal gebruikt moeten worden. Het aantal functiepunten dat nodig is voor een globale telling zal moeten blijken als vele projecten uitgevoerd zijn. Als van de projecten vastgelegd wordt in welke omstandigheden het project uitgevoerd is, dan is het op termijn mogelijk productiviteitscijfers te bepalen.

Voor de NESMA gebruikersfuncties zijn de factoren voor complexiteit bekend. Tijdens vervolgonderzoek kan bepaald worden of voor de transformatiefunctie een complexiteitstabel nodig is.

9 Literatuur

Abran A. en Robillard P.N., 1996, Function Point Analysis: An Empirical Study of Its Measurement Processes, IEEE Transactions on Software Engineering, vol 22. No. 12 blz. 895-910

Abran A., Symons C., Oligny S., 2001, *An overview of COSMIC-FFP field trial results*, ESCOM 2001

Angelis L., Stamelos I. en Morisio M., 2001, Building a Software Cost Estimation Model Based on Categorical Data, Proceedings of the Seventh International Software Metrics Symposium

ASL, <http://www.aslfoundation.org/nl/aslbislfoundation/index.html>, geraadpleegd 20 januari 2008

BISL, <http://www.aslfoundation.org/nl/aslbislfoundation/index.html>, geraadpleegd 20 januari 2008

Boetticher G. D. en Lokhandwala N., 2007, Assessing the Reliability of a Human Estimator, Third International Workshop on Predictor Models in Software Engineering

Bryczynski B., 2006, *SOFTWARE ENGINEERING PROJECT MANAGEMENT*, Wiley-India, blz. 183 - 194

Carreira P. en Galhardas H., 2004, Efficient development of data migration transformations, SIGMOD

CMM, <http://www.sei.cmu.edu/cmimi/>, geraadpleegd 12 februari 2008

Grimstad S. en Jorgensen M., 2006, A framework for the Analysis of Software Cost Estimation Accuracy, ISESE

ISBSG, www.ISBSG.org, geraadpleegd 11 november 2007

Jeffery R., Ruhe M. en Wieczorek I., 2001, Using Public Domain Metrics to Estimate Software Development Effort, Proceedings of the Seventh International Software Metrics Symposium

Jorgensen M. en Shepperd M., 2007, A Systematic Review of Software Development Cost Estimation Studies, IEEE Transactions on Software Engineering, vol 33. No. 1 blz. 33-53

Koorneef R, www.gegevensconversie.nl, geraadpleegd 3 februari 2008

Kelly C. en Nelms C., 2003, Roadmap to checking data migration, Elsevier Ltd

Kitchenham B., 1997, The problem with Function points, IEEE Software

Liebchen G. A. en Shepperd M., 2005, Software Productivity Analysis of a Large Data Set and Issues of Confidentiality and Data Quality, 11th IEEE International Software Metrics Symposium (METRICS 2005)

Lokan C., 2005, What Should You Optimize When Building an Estimation Model?, 11th IEEE International Software Metrics Symposium (METRICS 2005)

Lokan C. en Mendes E., 2006, Cross-company and Single-company Effort Models Using the ISBSG Database: A further Replicated Study, ISESE

Lokan C., Wright T., Hill P. R. en Stringer M., 2001, Organisational Benchmarking Using the ISBSG Data Repository, IEEE Software September/October 2001

Looijen M., 1995, *Beheer van informatie systemen*, ten Hagen & Stam uitgevers, Den Haag

Mendes E., Lokan C., Harrison R. en Triggs C. , 2005, A Replicated Comparison of Cross-company and Within-company Effort Estimation Modules using the ISBSG Database, 11th IEEE International Software Metrics Symposium (METRICS 2005)

Mendes E. en Lokan C., 2006, Cross Company and Single-company Effort Models using the ISBG Database" A further replicated Study, ISESE

Mustafa K., Gowthaman K. en Khan R.A. 2005, Measuring the Function Points for Migration Project: A Case Study, American Journal of Applied Science 2 (8), blz. 1218-1221

Nashir M., 2006, A Survey of Software Estimation Techniques and Project Planning Practices, Proceedings of the Seventh ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing

NESMA Definitives en telrichtlijnen voor de toepassing van functiepuntnalyse; Nesma Functional size measurement method conform ISO/IEC 24570; Handboek van de Nederlandse software metrieke gebruikers associatie, 2004 (versie 2.2)

Pol M., Teunissen R. en Veenendaal E., 2002, *TMAP*, Tutein Nolthenius, 's-Hertogenbosch

Pols van der R., 2008, e-mail ontvangen op 30-01-2008.

Premraj R. en Shepperd M., 2005, An Empirical Analysis of Software Productivity Over Time, 11th IEEE International Software Metrics Symposium (METRICS 2005)

Premraj R., Shepperd M., Kitchenham B. en Forselius P., 2005, An Empirical Analysis of Software Productivity Over Time, 11th IEEE International Software Metrics Symposium (METRICS 2005)

Pressman R.S., 2001, *Software Engineering: A practitioner's approach*, McGraw-Hill Professional, blz. 89-94.

Putnam, Myers, 1992, *Measures for excellence: Reliable software on time, within budget*, yourdon press computing series, Prentice Hall Building

Rahm E. en Hai Do H., 2000, Data Cleaning: Problems and current approaches, *IEEE Techn. Bulletin on Data Engineering*

Santillo L., Conté M. en Meli R. , 2005, Early & Quick Function Point: Sizing More with Less, 11th IEEE International Software Metrics Symposium (METRICS 2005)

Standish Group, 1999, http://www.standishgroup.com/sample_research/chaos1998.pdf

Stegink J., 1993, *Organisatie van servicegericht systeembeheer; Een methode voor het bespreekbaar en beheersbaar maken van service*, Mediaan/abs Heerlen

Symons C. R. , 1988, Function Point Analysis: Difficulties and Improvements, *IEEE Transactions on Software Engineering*, vol 14. No. 1 blz. 2-11

Tran-Cao D., Levesque G. en Abran A., 2002, Measuring Software Functional Size: Towards an Effective Measure of Complexity, *Proceedings of the International Conference on Software Maintenance*

Vogelezang F., 2005, COMIC Full Function Points “The next generation of functional sizing, proceeding SMEF 2005

Bijlage I

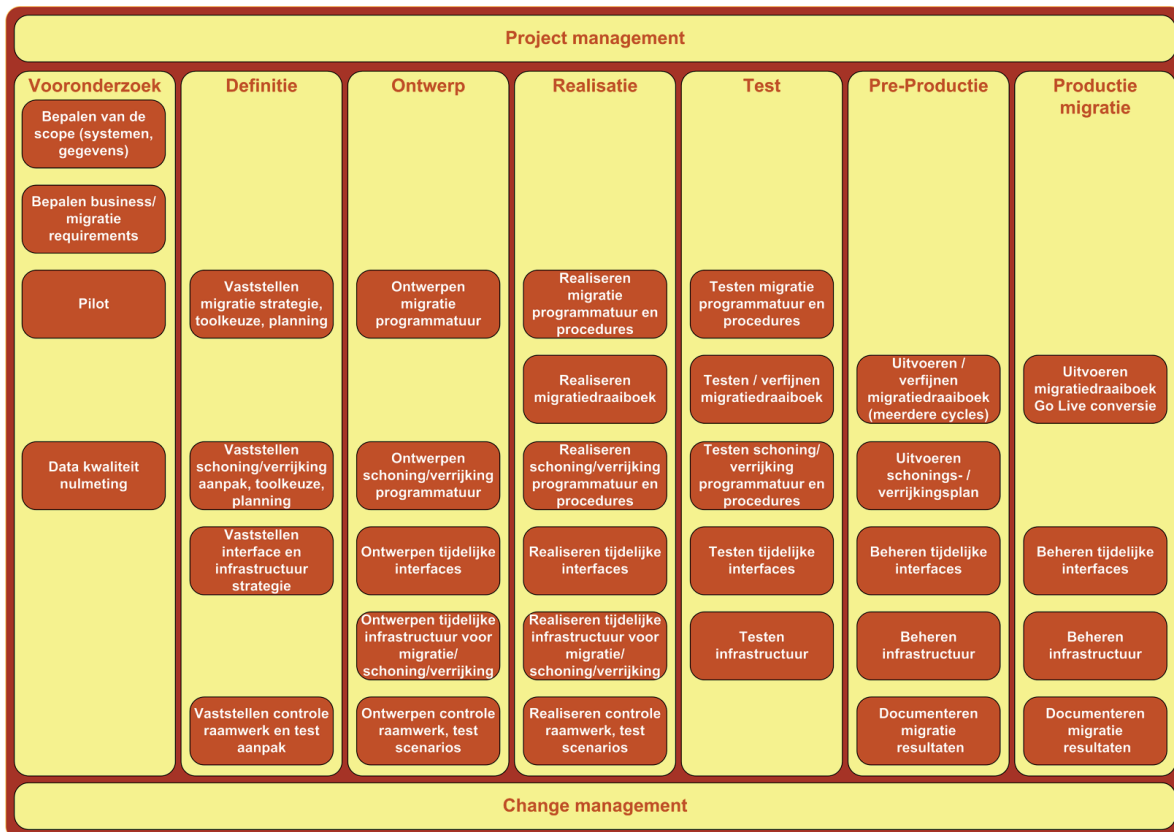
Vragenlijst om een datamigratiebegrotingsmethode te ontwikkelen

- *Vragen over de start van een datamigratietraject*
 - Bij wat voor soort bedrijf is het datamigratietraject uitgevoerd (branche, land, omvang)
 - Hoe is de datamigratie uitgevoerd? (bigbang/ gefaseerd)
 - Is een volledige datamigratie uitgevoerd of een gedeeltelijke? (bepaald land / boekjaar / lopende zaken)
 - Welke onderwerpen van een datamigratietraject zijn uitgevoerd?
 - Hoeveel iteraties zijn uitgevoerd?
- *Is er een projectplan opgesteld voor de datamigratie?*
 - Zo ja, op basis van wat is dat projectplan opgesteld?
 - Was vooraf bekend dat er data geschoond of verrijkt moest worden?
 - Hoe is dat begroot?
 - Hoe is de omvang, doorlooptijd en kosten begroot?
 - Wat was de omvang, doorlooptijd en kosten die vooraf begroot zijn?
 - Hoe is de omvang, doorlooptijd en kosten opgebouwd? (denk ook aan nacalculatie, fixed price, capaciteit)
 - Welke manier van rapportage is afgesproken?
 - Welke randvoorwaarden stel jij om de kosten, omvang en doorlooptijd te kunnen bepalen van het datamigratietraject?
 - Welke eisen stelt de klant meestal voor een datamigratietraject?
 - Zijn de doelstellingen die de klant heeft om de datamigratie uit te voeren van invloed op de kosten?
- *Vragen over de periode tijdens een datamigratietraject*
 - Met welke meeteenheid is de voortgang van het datamigratietraject gemeten?
 - Hoe is bepaald wat de voortgang was en of het traject nog op koers ligt?
 - Welke werkwijzen methodes zijn toegepast in het datamigratietraject?
 - Is gebruik gemaakt van een tool om de datamigratie uit te voeren?
 - Is maatwerk ontwikkeld om de datamigratie uit te voeren?
 - Hoe wordt bepaald of gegevens schoon zijn?
 - Hoe worden de kosten bepaald om gegevens te schonen?
- *Vragen na afloop van het datamigratietraject*
 - Welke valkuilen zaten in het datamigratietraject die kosten / doorlooptijd beïnvloeden?
 - Hoe wordt bepaald of de datamigratie succesvol verlopen is?
 - Is de manier om kosten en doorloop tijd te bepalen juist geweest in het traject?
 - Waar is verbetering mogelijk in bepalen van omvang, kosten en doorlooptijd?
 - Wat heb je geleerd op het gebied van bepalen van kosten, omvang en doorlooptijd in het afgesloten traject?
 - Welke factoren bleken tijdens het project de grootste invloed te hebben om de doorlooptijd, omvang en kosten van het datamigratietraject?
 - Wat was na afloop van het datamigratietraject de omvang, doorlooptijd en kosten?
- *Vragen voor datamigratie experts*

- Welke valkuilen heb je ontdekt door de jaren heen die kosten / doorlooptijd beïnvloeden?
- Is de manier om kosten en doorloop tijd te bepalen juist geweest in de afgelopen trajecten?
- Wordt het steeds makkelijker om een datamigratietraject te begroten?
- Zie je een datamigratietraject als een stuk te ontwikkelen functionaliteit?

Bijlage II Mediaan Migratie Methode

Mediaan/abs hanteert voor het definiëren en uitvoeren van datamigratieprojecten de Mediaan Migratie Methode (MMM). Deze methode is ontstaan door de praktijkervaring van een zestal datamigratie experts in een model te gieten. Het model is getoetst, door het in een aantal datamigratietrajecten toe te passen. De Mediaan Migratie Methode is weergegeven in de volgende figuur:



Figuur 3. Mediaan Migratie Methode

In de volgende paragrafen worden de activiteiten binnen de MMM verder toegelicht.

Vooronderzoek

Bepalen van de scope

- Vaststellen welke bedrijfsgegevens gemigreerd moeten worden. Alleen actuele gegevens of ook historie?
- Onderscheid tussen “stamgegevens” die “stabiel” zijn (bijv. klantadres, productassortiment, prijzen) en transactiegegevens die continu wijzigen (orders, leveringen, voorraad, openstaande posten)
- Uit welke bestaande applicaties zijn deze gegevens afkomstig, in welke applicaties dienen deze gegevens geladen te worden.
- Ook duidelijk vermelden welke bedrijfsgegevens niet gemigreerd worden.
- Hoe om te gaan met historische gegevens, indien die niet gemigreerd worden, hoe

kunnen deze dan na de datamigratie geraadpleegd worden.

Bepalen van de business datamigratie requirements

- De business requirements zijn o.a. input om in een volgende stap de migratie strategie te kunnen bepalen.
- Welke eisen stelt de business aan de datamigratie?
- T.a.v. het uitvoeren van de datamigratieactiviteiten, bijv. hoelang mag de “business” stilvallen tijdens het uitvoeren van datamigraties (applicaties zijn mogelijk gedurende enige uren/dagen niet beschikbaar op het moment dat de datamigratie wordt uitgevoerd)
- Prioriteiten (welke gegevens eerst, bijv. eerst zakelijke klanten, later consumenten)
- Eisen vanuit een audit department (controleren van de resultaten, rapportage)
- Acceptatiecriteria t.a.v. de kwaliteit van de datamigratie (bijv. het berekende “Netto bedrag” in de nieuwe applicatie mag in 99% van de gevallen max. ±0,50 afwijken van het berekende “Netto bedrag” in de oude applicatie en in 100% max. ±1,00)
- Business freeze, periode waarin er geen (substantiële) wijzigingen mogen worden doorgevoerd (bijv. introductie nieuwe producten en diensten, prijswijzigingen)

Pilot

- De pilot is een test waarmee onderzocht wordt of de gekozen scope, aanpak, planning, resources enz. realiseerbaar en uitvoerbaar is.
- Testen van enkele “complexe gevallen” en een “steekproef van reguliere gevallen”.
- Probleem hierbij is vaak het nog niet beschikbaar zijn van de doelapplicaties omdat deze nog in ontwikkeling zijn.

Data kwaliteit nulmeting

- De mate van data vervuiling is van grote invloed op de uit te voeren datamigratie (veel vervuiling leidt tot fouten, noodzakelijk reparatieacties, extra proefmigraties)
- Er dient dan ook te worden vastgesteld (een ruwe indicatie) wat de kwaliteit is van de bedrijfsgegevens die gemigreerd moeten worden.
- Worden er in de reguliere bedrijfsvoering dataschoningsacties uitgevoerd (bijv. actualiseren klantenbestand, verwijderen overbodige gegevens)?
- Hoe vaak worden gegevens gemuteerd
- Zijn er in de nieuwe applicatie(s) nieuwe gegevens nodig die nog niet bestaan in de bestaande applicaties (= >verrijking)
- Zijn er hulpmiddelen om de bestaande gegevens te onderzoeken (queries, dumps, rapportages)?

Definitie

Bepalen migratiestrategie, toolkeuze, planning

- Op basis van de gegevens die in het vooronderzoek zijn verzameld worden meerdere migratiescenario's (gefaseerde migratie, big bang) geanalyseerd en wordt vastgesteld welke scenario het meest geschikt is voor het project.
- Ook wordt vastgesteld welke tools voor het datamigratietraject ingezet zullen worden, zoals:
 - (handmatig met ondersteuning van “office”-hulpmiddelen als Excel, Access),
 - zelf te bouwen programmatuur,
 - tools behorende bij de bestaande of nieuwe applicaties (pakketten hebben

- vaak ook “data import” tools),
 - een specifiek datamigratietool (migration workbench)
- Planning van het datamigratietraject
- Hoe vaak testen (o.a. als gevolg van datakwaliteit)
- Beschikbaarheid van infrastructuur en andere middelen
- Beschikbaarheid van bron en doel applicaties (doel applicatie is vaak nog in ontwikkeling)

Bepalen aanpak voor schoning en verrijking, toolkeuze, planning

- Vaststellen op welke wijze de schoning uitgevoerd zal gaan worden
 - In de bronsystemen?
 - Tijdens datamigratie in een datamigratietool?
 - Handmatig (geval voor geval)?
 - Geautomatiseerd?
- Van belang hierbij is ook: indien de bestaande (oude) applicaties in gebruik blijven na de datamigratie is schoning in de bron meestal de enige optie.
- Welke eisen stellen de nieuwe applicatie(s) aan de aan te leveren gegevens?
 - > Dit bepaalt mede welke schoning er uitgevoerd moet gaan worden
 - Verplichte velden
 - Veldwaarden (domeinen)
 - Schoning binnen bestanden/databases
 - Schoning om inconsistenties tussen bestanden/databases te corrigeren (bijv. “order” waarvoor geen klantgegevens bestaan)
- Verrijking: aanleveren van nieuwe gegevens die noodzakelijk zijn om de nieuwe applicaties correct te laten werken.
 - Vaak kunnen deze nieuwe gegevens (attributen) niet in de oude applicaties vastgelegd worden, en moet er dus een alternatief verzonnen worden zoals hulpbestanden (spreadsheets?) of het invullen van defaults tijdens het laden in de nieuwe applicatie
- Planning van de schoning en verrijkingsactiviteiten (uitvoering meestal door business resources, ondersteuning door IT),
- Doelstellingen kunnen zijn: 100% kwaliteit bij golive, $\geq 80\%$ bij een acceptatietest

Vaststellen strategie t.a.v. infrastructuur en interfaces

- Op welke infrastructuur kunnen de datamigratietools gebouwd/getest worden
- Welke infrastructuur is er nodig voor het uitvoeren van proefmigraties
- Kopie bronsystemen, kopie doelsystemen, kopie datamigratieplatformen
- Welke applicaties moeten er via interfaces gekoppeld worden om een proefmigratie uit te kunnen voeren
- Hoe wordt het (versie) beheer van deze infrastructuur geregeld?

Vaststellen controle raamwerk en test aanpak

- Migraties via de Mediaan Migratie Methode zijn transparant t.a.v. volledigheid en juistheid. Het controle raamwerk is een “dashboard” waarmee de volledigheid van de datamigratie bewaakt kan worden, gebaseerd op aantallen en/of bedragen van de relevante bedrijfsgegevens. Gegevens worden op meerdere momenten tijdens het datamigratietraject (bron, migratieplatform, doel) vastgelegd en verwerkt in het raamwerk zodat een volledig beeld ontstaan over de volledigheid van de verwerkte gegevens.
 - Welke bedrijfsgegevens moeten in het controle raamwerk worden opgenomen?
 - Welke eisen heeft de Audit of Accountsdienst t.a.v. de bewaking van de datamigratie? Deze eisen moeten in het raamwerk worden verwerkt.

- Er moeten (functionele) testscripts worden vastgelegd waarmee na afronding van het datamigratietraject in de doelomgeving gecontroleerd kan worden of de gegevens correct gemigreerd zijn. Voorbeeld kunnen zijn:
 - Op basis van een of meerdere steekproeven bekijken van gegevens via een scherm of rapport en relevante attributen controleren
 - Het uitvoeren van een aantal functionele activiteiten, bijv. verwerken van orders, het aanmaken facturen, waarna de resultaten vergeleken met soortgelijke resultaten uit verwerking in de bronsystemen
 - Het aanmaken van extracties uit de bronsystemen waarna de gegevens (geautomatiseerd) vergeleken worden met de gegevens uit de bronsystemen (bijvoorbeeld: nieuw en oud adressenbestand met elkaar vergelijken en de verschillen tonen)

Ontwerp

Ontwerpen datamigratieprogrammatuur

- Architectuur keuze t.a.v. de plaatsen waar gegevensbewerkingen worden uitgevoerd (bron / migratieplatform / doel), waar wordt gefilterd, waar worden codevertalingen uitgevoerd.
 - Bij voorkeur (praktijkervaring!): volledige tabellen uit bron halen, bewerkingen in migratieplatform (migration workbench), minimale bewerkingen tijdens laden
 - Voordeel van deze aanpak: alle bewerkingen op één plek, eenvoudiger beheer, eenduidig, documentatie, foutanalyse etc.
- Er moeten specificaties worden bepaald van:
 - de gegevensextracties uit de bronapplicaties,
 - de bewerkingen die op de gegevens tijdens de datamigratie moeten worden uitgevoerd (filtering van records, splitsen van bestanden, combineren van bestanden, functionele/kwaliteitscontroles, code vertalingen, samenstellen van bestanden voor de doelapplicaties,
 - het laden van de gegevens in de doelapplicatie(s)
 - het aanleveren van controle tellingen voor het controle raamwerk (op alle relevante plekken: bron, migratieplatform, doel)

Ontwerpen schoning/verrijking programmatuur

- Functionale specificaties van de checks die op de brongegevens kunnen worden uitgevoerd om de kwaliteit te kunnen vaststellen
- Functionele specificaties van de handmatige of automatische bewerkingen op de gegevens in bron of migratieplatform
- Specificatie van de tools / hulpbestanden / bewerkingen om de noodzakelijke dataverrijking uit te voeren

Ontwerpen tijdelijke interfaces

- Specificeren / architectuur keuzes t.a.v. interfaces die voor het datamigratietraject noodzakelijk zijn zowel voor het testen als voor de uiteindelijk productie datamigratie.
- Voorbeeld: koppelingen met applicaties die stamgegevens beheren (bijv. produkt-assortimentsgegevens)

Ontwerpen tijdelijke infrastructuur voor schoning/verrijking en migratie

- Specificeren van de infrastructuur, welke applicaties, welke versies, welke hulpmiddelen

- Tijdens proefmigraties wordt meestal gewerkt met kopieën van alle systemen zodat een onafhankelijke test kan worden uitgevoerd.
 - Bronsystemen
 - Migratieplatformen
 - Doelsystemen
 - Hulpmiddelen
 - Middleware (messaging platformen)
 - Noodzakelijke kopieacties
 - Beveiliging / toegang
- Denk hierbij ook aan de sizing van de platformen voor het kunnen verwerken van grote hoeveelheden gegevens tijdens een korte periode (migratieweekend?)

Ontwerpen controle raamwerk, test scenario's

- Specificeren van het controle raamwerk. In detail aangeven welke tellers, welke gegevens op het dashboard zullen worden weergegeven, uit welke datamigratieprocesstapen telgegevens verzameld moeten worden
- Per gegevensobject een of meerdere testscripts specificeren om de juistheid van de gemigreerde gegevens te kunnen vaststellen

Realisatie

Realiseren datamigratieprogrammatuur en procedures

- Bouwen van de programmatuur d.m.v. maatwerk of het inrichten van de gekozen tools
- Meestal: Eenvoudige extractieprogrammatuur/queries aan de bron, gebruik van standaard laadtools aan de doelkant (bijv. SAP: LSWM, Oracle: xxx).
- Voor de migratieomgeving is dit afhankelijk van de gekozen tooling.
- Migratieprocedures worden gerealiseerd, incl. de noodzakelijke businessactiviteiten (bijv. stoppen met invoeren van orders of een workaround procedure voor het callcenter indien ze tijdens de datamigratie tijdelijk geen toegang hebben tot de applicaties)

Realiseren datamigratiedraaiboek

- Het uitvoeren van een datamigratietraject omvat een veelvoud aan activiteiten die in de juiste volgorde moeten worden uitgevoerd.
- Hiervoor wordt een datamigratiedraaiboek samengesteld waarin alle activiteiten van business, IT, datamigratieteam, infrastructuur enz. in detail worden vastgelegd.
- Alle activiteiten worden gepland incl. doorlooptijd, benodigde resources, afhankelijkheden t.a.v. andere activiteiten.
- In het draaiboek worden ook Go/NoGo beslismomenten opgenomen.

Realiseren schoning/verrijking programmatuur en procedures

- Bouwen van de programmatuur d.m.v. maatwerk of het inrichten van de gekozen tools

Realiseren tijdelijke interfaces

- Bouwen van de interfaces d.m.v. maatwerk of het inrichten van de gekozen tools (middleware, messaging platformen)

Realiseren controle raamwerk, test scenario's

- Realiseren van het controle raamwerk
- Realiseren van evt. tools voor testen

Test

- In de Test-fase worden alle gerealiseerde middelen getest. Hiervoor worden een of twee proefmigraties uitgevoerd. Doel is om alles technisch en functioneel te testen.
- In de eerste testrun wordt vooral gekeken naar de technische werking van de programmatuur, in de tweede testrun (ervan uitgaande dat de techniek dan al grotendeels in orde is) kan er al gekeken worden naar functionele aspecten.

Testen datamigratieprogrammatuur en procedures

- Testen extractieprogrammatuur en procedures
- Testen datamigratietools (workbench of zelfgebouwde tools)
- Testen laadprogrammatuur
- De datamigratieprogrammatuur levert telresultaten (aantallen) t.b.v. het controle framework

Testen datamigratiedraaiboek

- Het testen wordt uitgevoerd op basis van de planning in het datamigratiedraaiboek.
- Hierbij worden bevindingen meteen in het draaiboek verwerkt (nieuwe activiteiten die nog niet in het draaiboek waren opgenomen, aanpassingen van de volgorde, doorlooptijden)

Testen schoning/verrijking programmatuur en procedures

- Ook de schonings- en verrijkingsprogrammatuur en procedures worden getest.

Testen tijdelijke interfaces

- Voor het uitvoeren van de testen dienen de benodigde interfaces operationeel te zijn.

Testen tijdelijke infrastructuur

- Voor het uitvoeren van de testen dient de infrastructuur operationeel te zijn.

Pre-Productie

Er worden één of meerdere proefdatamigraties uitgevoerd teneinde het geheel van procedures, tools, werkinstructies, draaiboek, testscripts enz. te toetsen. Meestal is er een proefmigratie naar een acceptatietestomgeving. Dit heeft een tweeledig doel, namelijk het testen van de datamigratie als zodanig, maar ook het vullen van de acceptatietestomgeving met data voor het kunnen uitvoeren van de gebruikersacceptatietest (nieuwe processen en applicaties).

Uitvoeren/verfijnen datamigratiedraaiboek

Elke proefmigratie wordt het draaiboek uitgevoerd en waarnodig verbeterd.

Uitvoeren schoning / verrijkingsplan

- Uitval van gegevens bij het laden in de doelapplicaties kan een gevolg zijn van niet

correct werkende datamigratieprogrammatuur of ten gevolge van problemen in de data-kwaliteit. Mogelijk moeten er extra schonings/verrijkingsactiviteiten gepland worden.

Beheren tijdelijke interfaces

- Voor het uitvoeren van de testen dienen de benodigde interfaces operationeel te zijn.

Beheren infrastructuur

- Voor het uitvoeren van de testen dient de infrastructuur operationeel te zijn.

Documenteren datamigratieresultaten

- De datamigratieprogrammatuur levert telresultaten (aantallen) t.b.v. het controle raamwerk. Deze worden te samen met de resultaten van de functionele tests (testscripts) vastgelegd in datamigratierapporten.
- Een datamigratierapport bevat voor alle gemigreerde bedrijfsgegevens o.a. de volgende informatie:
 - Volledigheid (op basis van controle raamwerk)
 - Juistheid (op basis van de uitgevoerde testscripts)
 - Eventuele issues in procedures of tools
 - Eventuele nog uit te voeren schoningsactiviteiten

Productie

De productiedatamigratie is de definitieve datamigratie waarin de data wordt gemigreerd van de operationele bronapplicatie(s) naar de productieomgeving van de doelapplicatie(s). Afhankelijk van de gekozen migratiestrategie is er een “big bang” datamigratie (bijv. in één (lang?) weekend alles migreren), of zijn er meerdere productiedatamigraties (per keer wordt een deel van de gegevens overgezet naar productie).

Uitvoeren datamigratiedraaiboek, Go Live

Het draaiboek wordt uitgevoerd. Op de aangegeven momenten dient business verantwoordelijken een Go/NoGo besluit te nemen over het continueren of stopzetten van het datamigratietraject.

Beheren tijdelijke interfaces

In de productieomgeving dienen de interfaces te zijn ingericht.

Beheren infrastructuur

- De infrastructuur van de productieomgeving moet gereed zijn voor het uitvoeren van de datamigratie en het live gaan na afloop van de datamigratie.
- Denk ook aan performance, backups / restore momenten.
- Gebruikersautorisatie.

Documenteren datamigratieresultaten

- Er wordt een gedetailleerd datamigratierapport samengesteld met een verslag van alle uitgevoerde activiteiten, Go/NoGo beslissingen en datamigratieresultaten.
- Het rapport wordt door de opdrachtgever (stuurgroep) goedgekeurd.
- Het rapport dient als bewijs voor een volledige en juiste uitvoering van de datamigratie en de acceptatie van de resultaten door business management. Het rapport wordt beschikbaar gesteld aan audit en accountantsafdelingen.

Changemanagement

- Migratieprojecten hebben in het algemeen een langere doorlooptijd. Het is onvermijdelijk dat er gedurende het traject wijzigingen plaatsvinden t.a.v. de scope, aanpak, specificaties, eisen aan infrastructuur etc.
- Deze wijzigingen moeten op een adequate wijze gemanaged worden, waarbij zowel de noodzakelijke (door het projectmanagement goedgekeurde) wijzigingen gerealiseerd zullen worden, maar ook het lopende datamigratietraject zijn voortgang kan behouden.
- Voor elke wijziging zal daarom een impactanalyse gemaakt worden en zal bekeken worden op welke wijze en welk moment de wijziging gerealiseerd kan worden.

Projectmanagement

- Migraties worden projectmatig uitgevoerd, alle reguliere projectmanagement activiteiten komen hierbij aan bod:
 - Overall planning van het datamigratietraject
 - Afstemming met de implementatieplanning
 - Risicobeheersing
 - Resource management, “knowledge resources” met
 - Kennis van de business
 - Kennis van bronapplicaties (IT+business)
 - Kennis van doelapplicaties (IT+business)
 - Migratiespecialisten
 - Voortgangsrapportage
 - Aan Implementatie management en/of stuurgroep
 - Issue management

