

Energie Footprint Manual

1	INLEIDING	2
2	HET PROCES.....	3
2.1	De stappen in het proces.....	3
2.1.1	De Hoofdtransactie(s).....	3
2.1.2	De Scope.....	3
2.1.3	De Deployment View.....	3
2.1.4	De Meetperiode	4
2.1.5	Het Meetplan.....	4
2.1.6	Het Energiemodel	4
2.1.7	Utilisatiedata	4
2.1.8	Verbruiksdata	5
2.1.9	Algemene richtlijnen	Error! Bookmark not defined.
3	APPENDIX.....	6
3.1	Keuze van de Hoofd Business Transactie(s)	6
3.2	Regels voor bepaling van de scope.....	6
3.2.1	Buiten scope	6
3.2.2	Binnen scope.....	6
3.3	Meetstrategieën.....	6
3.3.1	SAN's.....	6
3.3.2	Virtualisatie	7
3.3.3	Mobiele toestellen.....	7
3.3.4	Internet verkeer.....	7
3.4	Parameters	7
3.4.1	Referenties	8

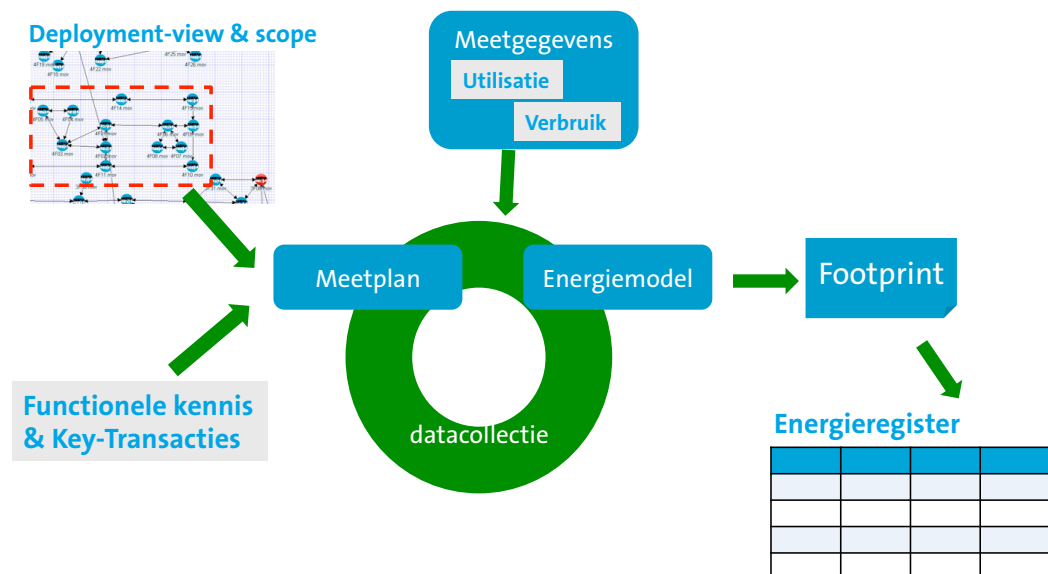
Versie	0.1
Datum	14 Mei 2013
Auteurs	Kay Grosskop Joost Visser Jeroen Arnoldus

1 Inleiding

Dit Document is bevat instructies voor het vaststellen van een energieprofiel van een applicatie of service.

De doelstellingen zijn

- Een gedetailleerde *overzicht* te geven over wat de uitvoering van een dergelijke meting inhoud.
- De lezer in staat te stellen op grond van dit document *zelfstandig* een energieprofiel vast te stellen
- De instructies moeten leiden tot een *herhaalbaar* resultaat



Figuur1: Onderdelen van een voetprint bepaling

2 Het Proces

2.1 De stappen in het proces

De basiselementen van een footprint bepaling zijn weergegeven in Figuur 1.

De bepaling van een energieprofiel bestaat uit de volgende stappen:

1. Het bepalen van een hoofd transactie
2. Het bepalen van de juiste scope
3. Het maken van een meetplan
4. Het maken van een energiemodel
5. Het uitvoeren van het meetplan, de data collectie
6. Het definitieve vaststellen van de resultaten

De doorlooptijd van een scan ligt typisch bij 4-8 weken, maar is in grote mate afhankelijk ervan hoe makkelijk het is de metingen van de productiesystemen te vergaren.

We gaan deze stappen en de daarvoor benodigde artefacten en hun samenhang nu kort toelichten. In de appendix zijn aanvullende richtlijnen te vinden voor het uitvoeren van de stappen.

2.1.1 De Hoofdtransactie(s)

Een systeem wordt gebouwd om taken te verrichten, de zogenaamde transacties. Meestal kan voor een systeem een enkele hoofdtransactie beschreven worden; bijvoorbeeld voor een core banking systeem het verwerken van een overboeking en voor een e-mailserver het versturen van een bericht. De andere transacties, zoals rapportages, staan in dienst van de hoofdtransactie.

Aan begin van de scan moet dus de hoofd transactie(s) worden vastgesteld. Deze zal vervolgens als eenheid dienen om de utilisatie van het systeem te formuleren en het energieprofiel wordt in termen van deze transactie(s) uitgedrukt. Deze wordt ook wel 'unit of work' genoemd.

2.1.2 De Scope

Het basisidee achter de footprint bepaling is dat elke computing resource die nodig is voor het draaien van de applicatie moet worden meegerekend. En het heeft een groot effect als delen van de infrastructuur worden weggelaten. Er zijn hierbij twee aspecten

- Inclusie: Er moet worden opgelet dat ook werkelijk alle relevante resources in kaart worden gebracht en meegeteld.
- Exclusie: sommige resources, zoals bijvoorbeeld een testserver, worden niet meegeteld.

De output van de scope bepaling is een lijst met resources. Dit zijn normaalgesproken servers, maar ook stukken netwerk of bijvoorbeeld een beeldscherm of SAN.

2.1.3 De Deployment View

De deployment view is nodig om zowel de scope beter te bepalen als ook het executiepad van de hoofdtransactie te begrijpen. Dit laatste is bijvoorbeeld nodig om te begrijpen hoeveel verkeer tussen componenten over het netwerk gaat.

2.1.4 De Meetperiode

Om de footprint van een systeem goed te bepalen is het belangrijk dat de metingen *representatief* zijn voor de hele werkingstijd van een systeem. Hoe lang de meetperiode daarvoor moet zijn hangt af van het systeem. Sommige systemen hebben een wekelijks wederkerend patroon en nauwelijks schommelingen gedurende het jaar. In dat geval is het voldoende om de footprint op meetgegevens over een week te baseren. Soms is er een piek in een bepaalde maand. In dat geval zal de meetperiode een jaar zijn. De langste periodes voor de scan is één jaar.

Er moet bij kortere periodes echter wel op de een of ander manier kunnen worden aangetoond dat de utilisatie over het jaar gelijk blijft.

2.1.5 Het Meetplan

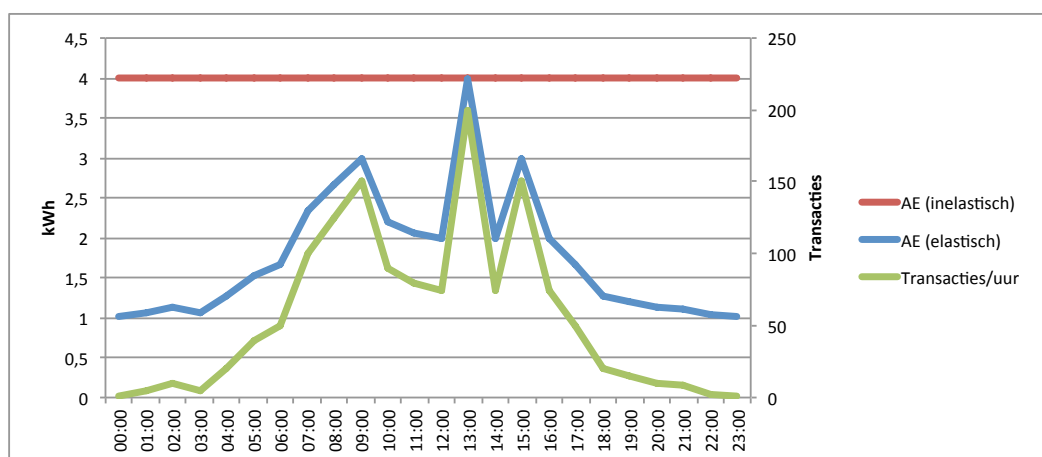
Het meetplan is in principe een lijst van data die moet worden verzameld voor een succesvolle footprint bepaling. Op het meetplan staan

- Wat voor type meting het is (utilisatie-, verbruik- of attributie-data)
- Voor welke resources de metingen zijn (bijv. welke machines)
- Wie een bepaalde soort data moet verzamelen.
- Over welke periode en met welke fijnmazigheid.

2.1.6 Het Energiemodel

Het energiemodel is het rekenmodel waarmee middels de meetdata de footprint wordt uitgerekend. In principe kan dit in vorm van een eenvoudig spreadsheet zijn.

Het basisidee van het energiemodel is, dat het totale energieverbruik van de applicatie wordt opgebouwd uit het energieverbruik van alle componenten ervan. Dit met een bepaalde granulariteit, zodat niet alleen het gemiddelde over de hele meetperiode bekend is, maar zodat ook fluctuaties zichtbaar zijn. Parallel aan verbruiksdata wordt ook utilisatie-data verzameld. Hieronder wordt verstaan hoeveel gebruik op een gegeven moment van het systeem wordt gemaakt. Door correlatie van deze twee soorten data kan een grafiek zoals hieronder worden verkregen



Figuur 2: Energieverbruik en utilisatie van een fictieve applicatie gedurende een dag

2.1.7 Utilisatie-data

Utilisatie-data (of *belasting*) van een systeem wordt gemeten in termen van de hoofdtransactie(s) (of *unit of work*). Dit is typisch te verkrijgen door middel van analyse van

logdata aan de buitenkant van het hele systeem zoals een server access log of expliciete functionele logging.

2.1.8 Verbruiksdata

Onder verbruiksdata wordt hier verstaan de *elektriciteitsconsumptie* van de hardware componenten waar het systeem gebruik van maakt. Deze zal vaak nog moeilijk te verkrijgen zijn. Naast het feit dat het gebruik van energiemeters nog geen standaard bij het systeem beheer spelen er met name twee moeilijkheden

1. **Gedeelte resources:** Hardware wordt gedeeld door meerdere systemen. Bijvoorbeeld als een database server wordt gedeeld.
2. **Niet bereikbare resources:** Hardware is niet toegankelijk voor metingen, bijvoorbeeld als gebruik wordt gemaakt van een remote service of een netwerk infrastructuur van een andere organisatie.

De strategie voor het eerste geval is *attributie*. Hierbij wordt gezocht naar leen verdeelsleutel om de totale energieconsumptie te kunnen toewijzen aan de verschillende applicaties.

In het tweede geval wordt vaak gebruik gemaakt van default waardes of schattingen. Een voorbeeld hiervan is het berekenen van internet energie verbruik door middel van een parameter voor internet verkeer energie intensiteit. Zie de appendix voor meer informatie omtrent deze strategieën.

2.1.9 Data collectie

De data collectie fase is in principe het uitvoeren van het meetplan. Het is echter belangrijk te beseffen, dat het zeker tijdens een eerste meting nodig zal zijn zowel het meetplan als ook het energiemodel aan te passen naar aanleiding van problemen of nieuwe inzichten tijdens de data collectie. Dit is dus een iteratief proces. Pas bij een hermeting kan ervan worden uitgegaan dat meetplan, energiemodel en de beschikbare data goed op elkaar aansluiten.

Over het algemeen zal de data collectie vrij pragmatisch moeten zijn. Als gegevens niet of alleen met grote inspanning te verkrijgen zijn kan worden overgegaan naar schattingen of het gebruik van default waardes. Hier geldt echter altijd de regel, dat deze schattingen 'pessimistisch' moeten zijn. Zo dient een default waarde ervan uit te gaan de het systeem *niet* energie-efficiënt is ingericht.

3 Appendix

3.1 Keuze van de Hoofd Business Transactie(s)

Aan begin van de scan moet de hoofd transactie(s) worden vastgesteld aan hand van de volgende criteria

- Het systeem ontleent zijn bestaansrecht aan deze transactie
- De transactie moet begrijpelijk zijn voor de business. Het is *functionaliteit* die wordt aangeboden worden Het moet niet een low-level technische transactie zijn.
- De transactie is representatief voor andere transacties.
- De transactie is niet te groot. De hoeveelheid transacties die het systeem verwerkt per tijdseenheid moet meetbaar zijn.
- Reduceer het aantal transacties mogelijkst tot één.

3.2 Regels voor bepaling van de scope

De bepaling van welke hardware componenten eigenlijk tot de applicatie gerekend moeten worden heeft een groot invloed op het resultaat. Afhankelijk van de eigen doelen zullen de keuzes hieromtrent afwijken. Voor het maken van een herhaalbare meting en een betere vergelijkbaarheid van het resultaat is het echter belangrijk regels op te stellen zodat op een uniforme manier wordt gemeten.

3.2.1 Buiten scope

- Alleen de productieomgeving wordt in beschouwing genomen. Ontwikkel- test- en acceptatieomgevingen worden genegeerd. Ook leeromgevingen voor gebruikers vallen buiten de scope.
- Periferie hardware, zoals printers, vallen in principe buiten de scope
- De energiebepaling gaat tot het publieke elektriciteitsnetwerk ('het stopcontact'). Inefficiënties in bijvoorbeeld het data-center (PUE) en laadtoestellen voor batterijen vallen binnen scope, maar verlies in het elektriciteitsnetwerk niet.

3.2.2 Binnen scope

- Tot de productieomgeving horen wel eventuele failover omgevingen.
- Netwerk en met name internet verkeer hoort binnen de scope
- Client installaties horen binnen de scope. Dit omvat bijvoorbeeld ook beeldschermen van een desktop PC.

3.3 Meetstrategieën

In deze sectie worden een aantal bestaande strategieën voorgesteld voor vaker voorkomende problemen tijdens het bepalen van een energie footprint

3.3.1 SAN's

SAN's worden vaak door meerdere applicaties gedeelt. De totale elektriciteitsconsumptie van de SAN moet dus worden opgedeelt onder de gebruikers. Hierbij moet op het volgende worden gelet.

- Er zijn meerdere indicatoren die als verdeelsleutel kunnen dienen. Afhankelijk van de situatie kan een (of meerdere) van de indicatoren geschikter zijn.
 - De gebruikte opslagcapaciteit

- De exclusief gereserveerde opslagcapaciteit
 - Het aantal IOPs
- Het totale energieverbruik moet worden verdeelt. Het is dus ongeldig om het verbruik door de totale capaciteit te delen en dan alleen het gebruikte deel aan te rekenen. Het ongebruikte gedeelte moet ook naar ratio worden toegekend onder de gebruikers.

3.3.2 Virtualisatie

Voor virtualisatie wordt vaak een fysieke server door meerdere virtuele machines gedeeld die niet voor dezelfde applicatie in gebruik zijn. Directe energiebepaling op VM level is vaak niet mogelijk. De volgende leidraden kunnen worden opgevolgd

- Als alle virtuele machines op een fysieke server worden gebruikt door applicatie-componenten die binnen de scope vallen, dan kan deze in principe behandeld worden als een niet-gevirtualiseerde machine: alle verbruikte energie is toe te rekenen aan de applicatie.
- Als de fysieke resources van de server vast geclaimd zijn (bijv. een CPU per VM en een vaste hoeveelheid geheugen) dan kan het verbruik op fysiek niveau naar een vaste ratio worden opgedeeld.
- Als resource consumptie variabel is, dan kan het gebruik van de meest energie-intensieve resource als verdeelsleutel worden genomen. Bij default is dat de CPU. Er moet dus worden bepaald hoeveel cycli van het totaal aantal zijn gebruikt door de applicatie.

3.3.3 Mobiele toestellen

De consumptie van mobiele toestellen wordt vaak bepaald aan hand van de stroom die aan de batterij wordt onttrokken. Echter moet ook het verlies bij het laden worden meegenomen. Per default wordt hier een factor 2 gehanteerd voor verlies door opladers.

3.3.4 Internet verkeer

Voor verkeer via internet wordt een factor gehanteerd per hoeveelheid data. Dit getal gaat jaarlijks omlaag met 30%. We hanteren een waarde van 0,05 kWh/GB. Het is dus belangrijk om een schatting te maken van al het verkeer voor de applicatie dat via internet gaat. Hieronder valt bijvoorbeeld ook data-replicatie naar een failover locatie.

3.3.5 Externe services

Voor een externe service call hanteren we een default waarde voor het geval de gemiddelde energiekosten van de service niet bekend zijn. De default is het gemiddelde uit het Energieregister.

3.4 Parameters

In de onderstaande tabel vindt u default waardes en andere parameters die in het energiemodel van een applicatie gebruikt kunnen worden.

Parameter	Waarde	Opmerking
Internet traffic energie intensiteit	0,5 kWh/GB	Gebaseerd op research van Koomey et al. [1] en een jaarlijkse correctiefactor
Default mobile charger verlies factor	2	
Gemiddeld PUE voor een Europees data center	2,5	Gebaseerd op research van DigitalReality [2]
Gemiddelde power consumptie van een laptop	40 W	
Gemiddelde power consumptie van een server	200 W	
Gemiddelde power consumptie van een desktop PC		
Default energie voor een externe service call	80 Wh	Gebaseerd op gemiddelde van het energieregister

3.4.1 Referenties

[To be completed]

[1] Weber, C. L., Koomey, J. G. and Matthews, H. S. (. "The energy and climate change implications of different music delivery methods." *Journal of Industrial Ecology* 14.5 (2010): 754-769.

[2] europe campus survey result 'what is driving the EU datacenter market' Jan 2013 Digital reality.