Analyse LEAP Track 1 "Powermanagement"

Gemaakt door Certios/WCooliT
Voor LEAP
in opdracht van de Rijksdienst voor Ondernemend Nederland

AUTEURS

Dirk Harryvan Certios B.V.
Marco Verzijl WCooliT B.V.
Max Amzarakov WCooliT B.V.

VOORWOORD

"Dit rapport dragen wij op aan Mees Lodder. De ideeën over de Server Idle Coëfficiënt die Mees Lodder samen met Dirk Harryvan ontwikkelde, vormden de basis van LEAP Track 1 'Powermanagement'. Mees is in december 2019 overleden."

SAMENVATTING

De afgelopen tien jaar hebben ICT-energie-efficiëntieprogramma's in Nederland geleid tot grote stappen voorwaarts. Dankzij deze aandacht voor energie-efficiëntie is de PUE van datacenters in Nederland stukken verbeterd en zijn de meeste datacenters nu zeer efficiënt als faciliteit.

Tegelijkertijd zijn fabrikanten van ICT-hardware doorgegaan met het ontwikkelen van servers en opslag- en netwerkapparatuur, waardoor de prestaties per kWh enorm zijn toegenomen. Deze vooruitgang wordt ook wel omschreven als de wet van Moore of de wet van Koomey. Daarnaast zijn met name servers op grote schaal voorzien van functies waardoor het energieverbruik van de machines zich aanpast aan de werkbelasting.

De gecombineerde vooruitgang van faciliteits- en ICT-efficiëntie hebben echter niet geleid tot een afname van het energieverbruik voor ICT-diensten in Nederland. Algemeen wordt aangenomen dat de vraag naar energie in de ICT-sector nog altijd enigszins toeneemt, mede als gevolg van een enorme toename in het gebruik van ICT-diensten.

Het "Lower Energy Acceleration Program" (LEAP) draait om verbetering van de energie-efficiëntie van ICT-diensten. LEAP wordt aangestuurd door een kernteam, in samenwerking met partijen uit de hele waardeketen van datacenters.

De eerste track van LEAP is gebaseerd op de constatering dat het totale elektriciteitsverbruik van datacenters nagenoeg constant is. Dit vormt een sterke tegenstelling met de bekende schommelingen in de vraag naar ICT-diensten. Daarom is er een onderzoek gestart om de reden voor dit stabiele energieverbruik te achterhalen en te proberen het totale energieverbruik van datacenters te verlagen door het energieverbruik sterker te koppelen aan de werkbelasting. Het LEAP-kernteam heeft gebruikgemaakt van een groep bedrijven die bereid waren deel te nemen aan pilots waarbij zij gegevens over hun huidige ICT-omgeving moesten aanleveren en bepaalde instellingen moesten wijzigen, zodat het effect op het energieverbruik van de betreffende servers kon worden gemeten.

In dit document worden de resultaten beschreven van de analyse van track 1 "Powermanagement" van LEAP. Ook zijn in dit eindrapport meer verzamelde metingen opgenomen om de betrouwbaarheid van de getrokken conclusies te vergroten. De analyse leidt tot de volgende waarnemingen:

- De meeste respondenten hebben hun servers in een dynamische energiemodus staan. Bij deze modi is het energieverbruik van de betreffende servers afhankelijk van de werkbelasting.
- Alle respondenten passen standaard enige vorm van de instelling “hoge prestaties” toe.
- Veel respondenten passen op het niveau van BIOS en OS tegenstrijdige instellingen toe.

- Als men het powermanagement instelt op modi die meer energie besparen, leidt dit op druk bezette serverknooppunten tot ongeveer 10% energiebesparing. Tijdens het testen van deze energiebesparingsmodi zijn er geen nadelige gevolgen voor de prestaties gemeld.
- De omschakeling van statische naar dynamische instellingen voor hoge prestaties leidt niet automatisch tot energiebesparingen voor één enkele server, maar kan wel voor besparingen zorgen voor een heel cluster machines.
- Zelfs de best bezette servers besteden meer dan een derde van hun energieverbruik aan "idle cycles", terwijl dit bij de slechtst bezette servers bijna 99% is.

In vervolggesprekken met de respondenten over de redenen en belemmeringen die ervoor zorgden dat zij geen energiebesparingsmodi gebruikten, kwamen zeer consistente antwoorden naar voren:

- Ondanks het vierde punt hierboven zijn er nog altijd grote zorgen over prestatieverliezen bij het gebruik van energiebesparing, zelfs nu niets in deze pilot erop wees dat daar daadwerkelijk sprake van zou zijn.

Met enig voorbehoud wegens de kleine schaal van dit onderzoek kunnen op basis van de resultaten van de pilot de volgende conclusies worden getrokken:

- Gebruik van de energiebesparingsmodus draagt sterk bij aan de doelen van LEAP (d.w.z. verbetering van de energie-efficiëntie van ICT in datacenters).
- De mogelijkheden om energie te besparen door middel van virtualisatie blijven enorm. Door naar hogere niveaus te streven kan zowel meer energie als meer geld worden bespaard dan de LEAP coalitie momenteel voor ogen heeft.
- Uit de verzamelde gegevens blijkt dat er verder onderzoek nodig is. Dat er nog steeds zorgen zijn over de gevolgen voor de prestaties, betekent dat er meer inzicht in en onderzoek naar powermanagement nodig is, ook met betrekking tot het effect op de prestaties van de applicaties. Ook is er een uitgebreidere statistische analyse van het energieverbruik van powermanagementfuncties en de gemiddelde CPU-belasting nodig om goed onderbouwde conclusies te kunnen trekken over het algemene gebruik van powermanagementfuncties en de potentiële energiebesparing.
- Er is dringend behoefte aan duidelijke en bij voorkeur eenduidige begeleiding en instructies van software- en hardwareleveranciers ten aanzien van de beste manier om powermanagement-instellingen toe te passen. Daarbij moeten de mogelijke besparingen worden benadrukt en moet worden uitgelegd wanneer het standaard powermanagement strakker of juist minder strak kan worden gemaakt.

Belangrijk is dat het gebruik van "powermanagement" en "virtualisatie" maatregelen vormen in het kader van de "Informatieplicht" voor datacenters die deel uitmaken van het "Activiteitenbesluit".

INHOUDSOPGAVE

1 INLEIDING

1.1 LEAP

De Amsterdam Economic Board, NLdigital, Green IT Amsterdam, de Rijksdienst voor Ondernemend Nederland en de Omgevingsdienst NZKG hebben het "Low Energy Acceleration Program" (LEAP) opgezet. Deze coalitie werkt samen met bedrijven uit de keten van datacenters, kennisinstellingen en de overheid, met ondersteuning van de Dutch Datacenter Association (DDA), om energiebesparingen te realiseren voor de ICT in datacenters, met als doel om de transitie naar een duurzame digitale economie te versnellen.

Het **doel** van LEAP is om inspirerende perspectieven te bieden voor de invoering van (nieuwe) technologieën en de ontwikkelingen te versnellen die voor energiebesparingen kunnen zorgen voor ICT binnen datacenters. Dit doen we om de toekomstbestendige groei van de sector een positieve impuls te geven.

Om de doelen van LEAP te bevorderen heeft de Rijksdienst voor Ondernemend Nederland (https://www.rvo.nl) opdracht gegeven tot de pilot "LEAP Track 1 'Powermanagement'".

LEAP track 1 is gericht op het realiseren van energiebesparingen met bestaande technologie, zoals powermanagement, waarbij energie-efficiënte instellingen van servers worden gebruikt zonder dat dit ten koste gaat van de prestaties. Ook virtualisatie (optimalisatie van de capaciteit van de servers wat energieverbruik betreft) en gebruik van een tool voor het meten van de doelen, om zo het energieverbruik en de gevolgen voor de prestaties structureel te kunnen monitoren en analyseren, kunnen tot de oplossingen behoren. Het **streven** is om samen te werken met leiders en voortrekkers in de waardeketen van datacenters en zo tegen het eind van 2022 energiebesparingen van 20-40% te realiseren.

LEAP is een coalitie van (momenteel) 20 partijen die de doelen van LEAP ondersteunen om energiebesparingen te realiseren voor de ICT binnen datacenters. Dit zijn partijen uit de waardeketen van datacenters:

- datacenters: Interxion, Iron Mountain;
- organisaties met veel dataverkeer en klanten van datacenters: Booking.com, Deloitte, gemeente Almere, gemeente Amsterdam, KPN, NEP The Netherlands, OD NZKG, Rabobank, Royal Schiphol Group, SURFsara, VU Amsterdam;
- (hardware)leveranciers: Dell Technologies, Hewlett Packard Enterprise, IBM, VMware en Red Hat;
- overheid: de gemeentes Amsterdam en Almere, ministerie van Economische Zaken en Klimaat (EZK)/ de Rijksdienst voor Ondernemend Nederland (RVO), OD NZKG;
- branche- en netwerkorganisaties als NLdigital, Green IT Amsterdam, DDA en de Amsterdam Economic Board.

Aan het begin van de pilot zijn er twee hypotheses geformuleerd:

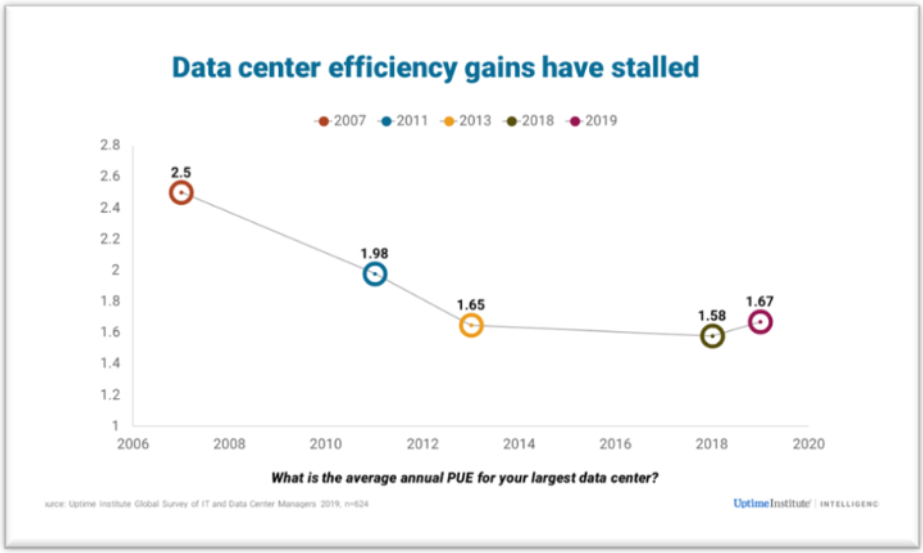
1. **Er is geen direct (lineair) verband tussen de IT-werkbelasting op een server en het energieverbruik van deze server.** Hoewel deze hypothese op de uitkomsten van verschillende praktijkgevallen is gebaseerd, is het belangrijk deze op structurelere wijze te onderzoeken. Met het oog daarop is er een nieuwe variabele geïntroduceerd die wordt gemeten: de Server Idle Coëfficiënt (SIC).

2. **Het gebruik van powermanagementfuncties op servers biedt mogelijkheden voor energiebesparing, zonder dat dit merkbaar ten koste gaat van de prestaties of beschikbaarheid van de server.**

Dit rapport bevat een analyse van de gegevens die in zowel fase 1 als fase 2 van de analyse van track 1 "Powermanagement" van de LEAP-pilot zijn verzameld. Er worden waarnemingen beschreven en conclusies getrokken op basis van die waarnemingen. Aan het eind van dit document worden aanbevelingen geformuleerd op basis van deze conclusies.

1.2 DATACENTERS EN ENERGIE

Het is algemeen aanvaard dat datacenters tot de grote energiegebruikers behoren in de huidige economie. De schattingen verschillen, maar een veelgenoemd cijfer is 2% van de nationale elektriciteitsproductie. Vanwege dit gebruik is de efficiëntie van datacenters al jaren aan aandachtspunt en is er, door verbeteringen die exploitanten van datacenters de afgelopen decennia hebben doorgevoerd, een situatie ontstaan waarbij verdere verbeteringen van de infrastructuur van faciliteiten waarschijnlijk niet meer tot significante energiebesparingen leiden voor deze datacenters. Uit de mondiale cijfers van het Uptime Institute blijkt bijvoorbeeld dat de verbeteringen van faciliteiten zijn gestagneerd (figuur 1):



Figuur 1: wereldwijde gemiddelde PUE in de loop der jaren

Power Usage Effectiveness (PUE) is een ratio die weergeeft hoe efficiënt een computerdatacenter energie gebruikt. Met andere woorden, hoeveel energie de computerapparatuur gebruikt (in tegenstelling tot koeling en andere overhead). Formule 1:

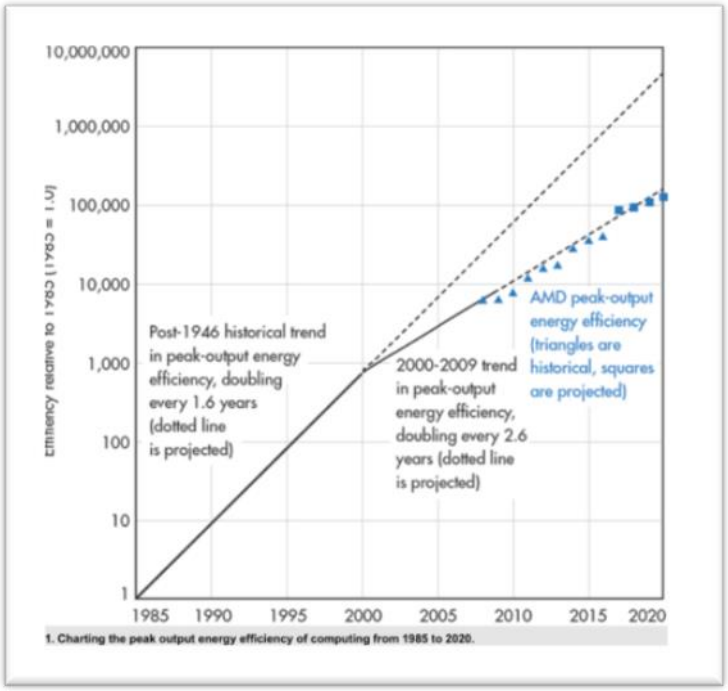
Totaal energieverbruik datacenter = PUE x energieverbruik ICT (1)

Formule 1 laat zien hoe belangrijk energiebesparingen binnen ICT zijn, dat de PUE als multiplier fungeert en dat iedere bespaarde kWh in het energieverbruik van ICT tot een grotere besparing in het totale energieverbruik van een datacenter leidt.

In Amsterdam hebben datacenters een verplichte doelstelling van een PUE van ten minste 1,3, maar de daadwerkelijke PUE kan op minimaal 1,15 liggen (deze cijfers zijn "by design", terwijl in de grafiek de

gemeten PUE wordt weergegeven) (bron: Ruimtelijke Strategie Datacenters – Routekaart 2030 voor de groei van datacenters in Nederland). De Nederlandse datacenters doen het al beter dan de mondiale gemiddelden, en de kans is klein dat verdere verbeteringen nog tot een grote verbetering zullen leiden.

Wanneer verbeteringen van de PUE niet meer effectief zijn, maar het totale energieverbruik nog sterk moet worden verlaagd, zal de vooruitgang moeten komen van verbeteringen in de ICT-apparatuur waarmee deze datacenters gevuld zijn. Dergelijke verbeteringen van de computerefficiëntie zijn al decennia lang zo sterk verweven met de ontwikkeling van apparatuur dat er zelfs een "wet" voor is opgesteld: de wet van Koomey (bron: post van Koomey, https://www.koomey.com/post/153838038643).



Figuur 2: de wet van Koomey, verbeteringen van de computerefficiëntie

De in figuur 2 weergegeven verbeteringen zijn echter representatief voor de maximale prestaties van een server. In de praktijk worden servers zelden in die mate belast en draaien zij veel dichter bij "idle" (inactief). De LEAP-pilot levert gegevens van verschillende belastingen op die gedurende een week zijn verzameld en dus gebaseerd zijn op een realistischer belasting dan die van Koomey.

1.3 DOEL VAN DE METINGEN

In LEAP track 1 hebben we geprobeerd te bepalen hoeveel besparingen powermanagementfuncties die beschikbaar zijn in moderne ICT-servers zouden kunnen opleveren. Om de hypothese te kunnen analyseren zijn tijdens de pilots de volgende zaken gemeten:

- het elektriciteitsverbruik van servers;
- de CPU-benutting van deze servers op hetzelfde moment.

Het doel van de pilot was om deze metingen uit te voeren met twee verschillende instellingen voor powermanagement: eerst de uitgangssituatie – een volle week aan metingen met de huidige instellingen – en vervolgens een tweede week met strenger powermanagement.

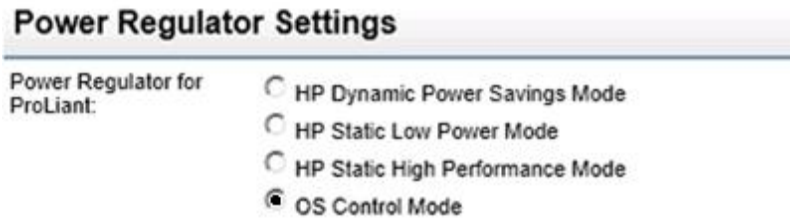
Dergelijke metingen zijn van belang om te kunnen onderzoeken of powermanagement kan worden gebruikt om energie te besparen zonder dat dit merkbaar ten koste gaat van de prestaties. Een secundair doel was het testen van een nieuwe metriek – de Server Idle Coëfficiënt – en de gevoeligheid daarvan voor de gevolgen van powermanagement.

1.4 METHODE VAN DE METINGEN

Voor deze pilot is er een metingsprotocol verspreid onder de deelnemers. Dit protocol bestond uit een aantal eenvoudige stappen voor het vastleggen van een uitgangssituatie en een aanvullende meting om het eventuele effect van wijzigingen in het powermanagement te bepalen.

De eerste stap was het vastleggen van de huidige instellingen voor powermanagement op het niveau van zowel BIOS als OS. Hoewel de interfacing van powermanagement gestandaardiseerd is, bieden verschillende fabrikanten verschillende opties voor de BIOS-instellingen en gebruiken zij verschillende benamingen.

Voorbeelden van de opties binnen de BIOS van een HP ProLiant-server:



en van Dell Poweredge-servers, waarbij DBPM staat voor "Demand-Based Power Management" (powermanagement op basis van vraag).

Statische MAX prestaties	DBPM uitgeschakeld (BIOS stelt P-State op MAX) Geheugenfrequentie = Maximale prestaties, Ventilatoralgoritme = Prestaties
Door OS bestuurd	Door OS bestuurd DBPM ingeschakeld (BIOS stelt alle mogelijke P-states open voor OS) Geheugenfrequentie = Maximale prestaties, Ventilatoralgoritme = Energieverbruik
Actieve energiecontrole	DBPM DellSystem ingeschakeld (BIOS stelt niet alle P-states open voor OS) Geheugenfrequentie = Maximale prestaties, Ventilatoralgoritme = Energieverbruik
Aangepast	**Beheer van energieverbruik en prestaties CPU**: Maximale prestaties | Minimaal energieverbruik | DBPM OS | DBPM systeem | **Beheer energieverbruik geheugen en prestaties:** Maximale prestaties |1333 Mhz | 1067 Mhz | 800 Mhz | Minimaal energieverbruik ventilatoralgoritme | Energieverbruik

Duidelijk is dat de meeste systemen drie verschillende soorten instellingen gebruiken:

- statisch, waarbij er (vrijwel) geen verband is tussen de werkbelasting van de server en het energieverbruik;
- dynamisch, waarbij de CPU-states door de hardware worden bestuurd;
- door het OS bestuurd, waarbij de CPU-states worden bestuurd door het hoofd-OS, de laag die de hardware-abstracties uitvoert.

Binnen deze besturingssystemen bestaan soortgelijke opties, bijvoorbeeld bij VMware ESX:

De OS-instelling treedt pas in werking wanneer de juiste BIOS-instelling wordt toegepast.
Zoals zal blijken, vormen de benaming van en de aanvullende opmerkingen bij de powermanagement-instellingen een belangrijke bepalende factor voor de keuzes die systeembeheerders maken.
Alleen al door de naam lijkt “hoge prestaties” een logische keuze. Uit de pilot blijkt dat deze instelling ook het meest wordt gebruikt. Een gedetailleerd onderzoek naar de daadwerkelijke processen achter deze instellingen toont aan dat de gebalanceerde modus in veel gevallen prestatievoordelen oplevert en dat er zelfs in de modus "laag energieverbruik" geen verslechtering van de prestaties is waargenomen.

Nadat de huidige instellingen zijn vastgelegd, moeten er een week lang metingen worden uitgevoerd. Hoewel kortere meetperioden ook resultaten zouden opleveren, wordt de voorkeur gegeven aan een volledige week, zodat er variaties in de werkbelasting kunnen worden geregistreerd die het gevolg zijn van aan de openingstijden van bedrijven gerelateerde werkpatronen.

De meting bestaat uit twee datapunten die ten minste elk kwartier worden verzameld.

1. Totaal elektriciteitsverbruik [Watt]
2. CPU-belasting [%]

Het elektriciteitsverbruik kan worden verkregen via de systeembeheerconsole. Alle moderne servers leveren deze informatie aan de systeembeheerder.

De CPU-belasting wordt verkregen van het hoofdbesturingssysteem of uit de beheersoftware. De CPU-belasting wordt uitgedrukt als percentage van de beschikbare CPU-capaciteit. De CPU-belasting wordt gemeten over een bepaalde tijdsinterval, waarbij het percentage weergeeft in hoeveel van de totale beschikbare klokcycli instructies zijn verwerkt. Er moet een geschikte interval worden gekozen om snelle variaties in te perken. Een voortschrijdend gemiddelde van 20 seconden wordt voor veel tools voor prestatietoezicht aangeraden, maar de daadwerkelijke interval wordt aan de systeembeheerder overgelaten.

Een klein stukje van zo'n meting ziet er als volgt uit:

Tijd	CPU-%	Stroom [W]
28-05-2020 12:16	24,16	364
28-05-2020 12:31	28,2	359

28-05-2020 12:46	53,57	408
28-05-2020 13:01	24,54	351
28-05-2020 13:16	24,43	356
28-05-2020 13:31	28,85	372
28-05-2020 13:46	35,7	377
28-05-2020 14:01	45,36	392
28-05-2020 14:16	29,22	367

1.5 Server Idle Coëfficiënt

Er is een nieuwe metriek ontwikkeld, die de "Server Idle Coëfficiënt" (SIC) wordt genoemd. Het beginpunt voor de ontwikkeling van deze metriek is een voortdurende zoektocht naar een objectieve manier om de efficiëntie van ICT te meten. Efficiëntiemetrieken worden gedefinieerd als de hoeveelheid energie die nodig is per werkeenheid. Het energieverbruik van ICT-systemen is gemakkelijk te meten, maar over de definitie van een werkeenheid is men het nooit eens kunnen worden.

De nieuwe metriek is gebaseerd op het concept dat men het niet eens kan worden over een "werkeenheid", maar dat het tegenovergestelde, een eenheid van inactiviteit ("idleness") algemeen aanvaard is. Onder "idle" (inactief) wordt een periode zonder CPU-belasting verstaan. Om de idle-coëfficiënt te bepalen meten we het totale energieverbruik van een server en bepalen we hoeveel energie er in de inactieve toestand is gebruikt. De elektriciteit die de server in inactieve toestand nodig heeft, wordt gemeten of op een andere manier vastgesteld. In de berekening wordt dit elektriciteitsverbruik genoteerd als "P_{idle}".

In de LEAP-pilot meten we elektriciteit "P" (Power) en CPU-belasting "CPU-%". De SIC kunnen we als volgt berekenen:

$$SIC = [E_{totaal} / E_{totaal} - E_{idle}] \quad (2)$$

Hierbij varieert de SIC van 1 tot oneindig (zoals ook het geval is bij de welbekende PUE).

Een andere mogelijke weergave van de SIC is:

$$SIC\text{-}\% = [E_{idle} / E_{totaal}] \quad (3)$$

In dat geval varieert de SIC van 0 tot 100% en staat deze voor het percentage van de energie dat wordt gebruikt in de inactieve toestand.

Er is nog een derde weergave voorgesteld:

$$SIC_{score} = 10 * (1 - (E_{idle} / E_{totaal})) \quad (4)$$

In alle bovenstaande berekeningen wordt de energie die in een periode (n) in de inactieve toestand wordt gebruikt, als volgt berekend:

$$E_{idle}(n) = [100\% - CPU\text{-}\%(n)]\ P_{idle} * \text{lengte interval (n)} \quad (5)$$

Totale energie inactieve toestand: $E_{idle} = Som\ [E_{idle}(n)]$ (6)

Totale energie: $E_{totaal} = Som\ [P(n) * \text{lengte interval (n)}$ (7)

1.6 P_{IDLE} BEPALEN

Het bepalen van het elektriciteitsverbruik van de server in inactieve toestand (P_{idle}) is essentieel om de SIC te kunnen bepalen (zie formule 5), maar het bepalen van P_{idle} is niet triviaal.
In de ideale situatie is er sprake van een volledig geïnstalleerde server, inclusief de virtualisatielagen en met het OS geïnstalleerd, maar zonder actieve gebruikersprogramma's.
Deze situatie wordt in benchmarksituaties gecreëerd bij het bepalen van de SPECpower-benchmark.
Het totale elektriciteitsverbruik wordt geregistreerd wanneer het systeem is ingeschakeld, maar er geen programma's actief zijn, wat P_{idle} oplevert.

Deze ideale situatie kan niet worden gebruikt wanneer men probeert P_{idle} te bepalen bij actieve servers. Deze machines kunnen niet worden geïsoleerd en de gebruikersprogramma's kunnen niet worden gestopt om de elektriciteit in inactieve toestand te meten.

Er zijn verschillende andere opties voor het bepalen van het elektriciteitsverbruik in inactieve toestand:

1) Bij een server met een statische energie-instelling is het elektriciteitsverbruik in actieve en inactieve toestand gelijk. In dit geval worden de formules voor de berekening van de Server Idle Coëfficiënt eenvoudiger en is de SIC gelijk aan het gemiddelde CPU-percentage in inactieve toestand.
2) Bij een server met een dynamische energie-instelling waarbij de CPU-belasting in een bepaalde periode minder dan 1% is, kan het gemiddelde elektriciteitsverbruik in deze periode als een goede benadering van P_{idle} worden beschouwd.
3) Bij een server met een dynamische energie-instelling die nooit volledig inactief is, levert de lineaire extrapolatie van de curve van energieverbruik versus CPU-belasting richting 0% belasting een acceptabele waarde voor P_{idle} op.
4) Wanneer er enkel statistieken beschikbaar zijn voor het energieverbruik van de server, zoals het geval kan zijn wanneer er geen of beperkte toegang tot de CPU-statistieken wordt verleend, wordt het gemiddelde van de periode met het laagste geregistreerde elektriciteitsverbruik geacht P_{idle} op te leveren.

Bij het analyseren van de resultaten van de LEAP-pilot zijn al deze methoden toegepast.

1.7 SITUATIES WAARIN POWERMANAGEMENT NIET AAN TE RADEN IS?

Powermanagement is een verzamelnaam voor verschillende technologieën, dus deze vraag heeft betrekking op de gewenste instelling.

In een beperkt aantal gevallen heeft de instelling "hoge prestaties" de voorkeur boven de gebalanceerde of energiebesparingsmodus. Bij instellingen voor hoge prestaties gaan CPU-kernen niet naar hogere C-states. Dit betekent dat alle CPU-kernen altijd actief zijn. Dat is wenselijk wanneer men zeer consistente en snelle reactietijden wil. Let op: het gaat hier niet om de totale rekenkracht van de server, maar om de reactiesnelheid bij een opdracht, zelfs al is de CPU-belasting van de server zelf laag.
Dergelijke situaties doen zich voor bij High-Performance Computing (HPC), waarbij bijvoorbeeld RAM-geheugen van meerdere servers wordt gecombineerd via speciale netwerken, en in de financiële wereld, waar AI op de beurs wordt verhandeld en een milliseconde vertraging al te veel kan zijn. In deze gevallen biedt de instelling voor hoge prestaties de gewenste functionaliteit.

2 ACPI

In een computer biedt de "Advanced Configuration and Power Interface" (ACPI) een open standaard die besturingssystemen kunnen gebruiken om hardwareonderdelen van computers te herkennen en te configureren, powermanagement uit te voeren door (bijvoorbeeld) ongebruikte onderdelen in de slaapstand te zetten, en de statusmonitoring uit te voeren. Bij ACPI, dat in **december 1996** voor het eerst werd geïntroduceerd, wordt het powermanagement naar het besturingssysteem verplaatst, in tegenstelling tot het eerdere BIOS-centrische systeem dat gebruikmaakte van platformspecifieke firmware om het powermanagement- en configuratiebeleid te bepalen. De specificatie staat centraal bij het op het besturingssysteem gerichte configuratie- en powermanagementsysteem (OSPM), een ACPI-implementatie waarbij de verantwoordelijkheden op het gebied van apparaatbeheer worden weggehaald bij legacy firmware-interfaces via een UI.

Intel, Microsoft en Toshiba hebben de standaard oorspronkelijk ontwikkeld, terwijl later ook HP, Huawei en Phoenix eraan meegewerkt hebben. In oktober 2013 is de ACPI Special Interest Group (ACPI SIG), de oorspronkelijke ontwikkelaar van de ACPI-standaard, ermee akkoord gegaan alle middelen over te dragen aan het UEFI Forum, waar alle toekomstige ontwikkelingen zullen plaatsvinden. Het UEFI Forum heeft eind januari 2019 de nieuwste versie van de standaard uitgebracht: "Revision 6.3".

Kort gezegd kan elke werkende server tegenwoordig zijn elektriciteitsverbruik in bepaalde mate aanpassen aan zijn ICT-werkbelasting. De beheersing van het dynamische spectrum ligt echter bij de hardware zelf (via BIOS-instellingen) of bij het besturingssysteem (OS) dat op de hardware draait. De term OS wordt hier in brede zin gebruikt: VMware ESX of Microsoft Hyper V valt hier net zo goed onder als Windows, Linux of Unix OS.

Het is belangrijk om op te merken dat, om te zorgen dat deze controlemechanismen tegelijk aan de systeembeheerders worden gepresenteerd, de juiste modus voor door het OS bestuurde operaties een BIOS-instelling van "door het OS bestuurd" zou zijn, gevolgd door de juiste instelling binnen het OS. De kans is groot – al is dit niet bevestigd – dat iedere andere BIOS-instelling de OS-instellingen zal opheffen, maar er is meer onderzoek nodig om het effect van tegenstrijdige instellingen in het BIOS en OS aan te tonen.

Powermanagement bestaat uit verschillende stappen:

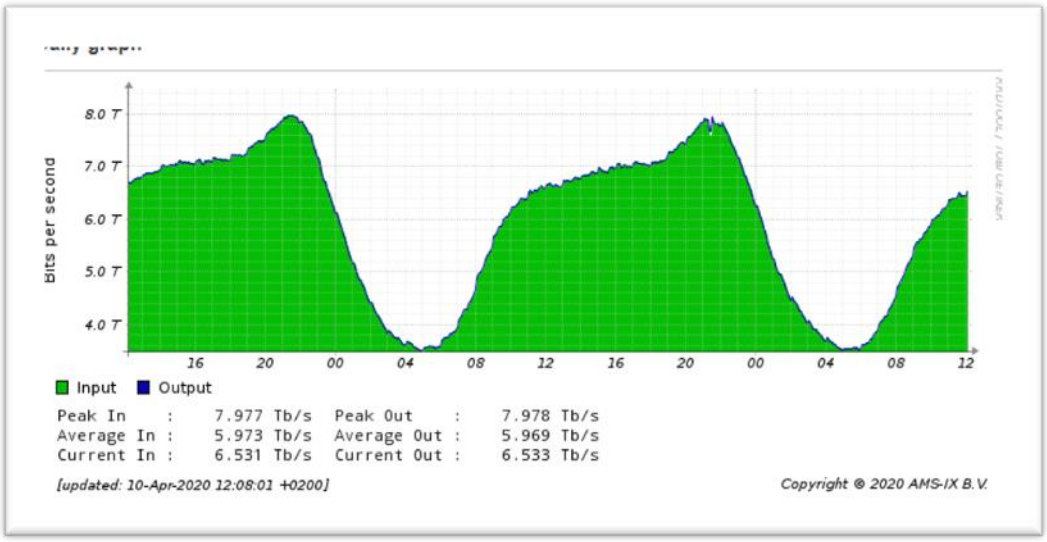
- HP = hoge prestaties. Dit houdt in dat er weinig energie wordt bespaard wanneer de CPU van de server inactief is. In veel gevallen worden er alsnog aanpassingen gemaakt in de kloksnelheid. Deze aanpassingen vallen onder de zogeheten ACPI P-states en vinden plaats als een CPU niet inactief, maar onderbelast is.
- Extra powermanagementstappen: veel servers hebben meerdere instellingen voor powermanagement. Deze kunnen per merk en per type server verschillen. Zij dienen om steeds meer energie te besparen naarmate er minder CPU-capaciteit nodig is en/of een of meerdere kernen worden uitgeschakeld (dieper). Voorbeelden van CPU-status:
    - C0 = actief;
    - C1 = minst agressieve vorm van terugschakeling bij inactieve toestand. Opstarttijd van een uitgeschakelde kern is ongeveer 0,5 microseconden;
    - C6 = zwaarste C-state: CPU heeft helemaal geen stroom. Opstarttijd van C6 is ongeveer 40 microseconden.

Als de CPU op 3,3 GHz "loopt", duurt opstarten vanuit C-states naar C0 1.650-13.200 klokcycli.

Om deze mogelijke vertragingen in perspectief te plaatsen moeten we ons allereerst realiseren dat een CPU een behoorlijke tijd niet gebruikt moet zijn om deze in een C6-state te kunnen zetten. ACPI C-states zijn uitsluitend op inactieve CPU's van toepassing. De vermelde opstartvertraging doet zich slechts eenmaal voor, namelijk wanneer een inactieve CPU aan de pool van actieve CPU's moet worden toegevoegd.

Daarnaast moeten we kijken naar verschillende andere vertragingen die zich zo nu en dan kunnen voordoen tijdens een computerverwerking. De reactietijd van een harde schijf ligt rond de 10 ms, maar ook netwerkverkeer kan voor milliseconden vertraging zorgen. Zelfs zonder behandelingsvertragingen (zenden/ontvangen) duurt een heen- en terugverplaatsing over 100 m glasvezel al 1 microseconde. We kunnen wel stellen dat het voor een eindgebruiker onmogelijk zou zijn 40 microseconden extra vertraging in de reactietijd van een applicatie op te merken.

De werking van ACPI-states (powermanagement) lijkt met name nuttig wanneer we het werkbelastingsprofiel in aanmerking nemen, zoals gepubliceerd door de Amsterdam Internet Exchange (AMS-IX).



Figuur 3: dagelijks verkeer AMS-IX

Deze grafieken tonen het netwerkverkeer over twee perioden van 24 uur en het enorme verschil in internetverkeer in de loop van een dag. We moeten aannemen dat deze grote variaties in het netwerkverkeer gepaard gaan met vergelijkbare variaties in de CPU-belasting van de server. Door gebruik te maken van de powermanagementstanden kunnen servers hun energieverbruik verlagen wanneer de werkbelasting daalt.

In de LEAP-pilot komen twee soorten serverbelasting voor: applicaties die van machine naar machine werken ("machine-to-machine") en applicaties die van machine naar eindgebruiker werken ("machine-to-end-user").

Over het algemeen is er in de CPU-belasting van servers waarop applicaties draaien die van machine naar eindgebruiker werken, een terugkerend patroon zichtbaar van hoge belasting wanneer gebruikers actief zijn en lage belasting wanneer zij niet actief zijn. In commerciële omgevingen vallen deze actieve perioden vaak samen met kantoortijden. In de onderstaande gegevens zijn de meest voor de hand liggende applicaties die van machine naar eindgebruiker werken applicaties voor een "virtuele desktop", zoals Citrix (zie figuur 10). Gebruikers melden zich 's ochtends aan, werken en melden zich 's avonds

weer af. Servers waarop uitsluitend dit soort applicaties draait, zijn in feite 120 van de 168 uur per week inactief en zouden veel voordeel hebben van de hoogst mogelijke powermanagement-instellingen.

Applicaties die van machine naar machine werken hebben geen rechtstreeks verband met activiteit van eindgebruikers. In de pilot spelen servers waarop monitoringapplicaties draaien een grote rol. Deze applicaties monitoren de gezondheid van netwerken, systemen en applicaties via regelmatige peilingen. Dergelijke applicaties hebben geen inactieve perioden, waardoor de werkbelasting over het algemeen zeer constant is. Doordat de werkbelasting zo voorspelbaar is, kan deze optimaal over de beschikbare hardware worden verdeeld, wat leidt tot een hoge gemiddelde CPU-belasting met zeer weinig variatie (zie figuur 9). Bij dit soort systemen wisselen CPU's gewoonlijk niet tussen inactieve en actieve toestand (vanwege de constante werkbelasting), maar zij kunnen alsnog voordeel hebben van powermanagement, aangezien niet alle rekenkracht waarover deze machines beschikken nodig is voor de uitvoering van hun taken.

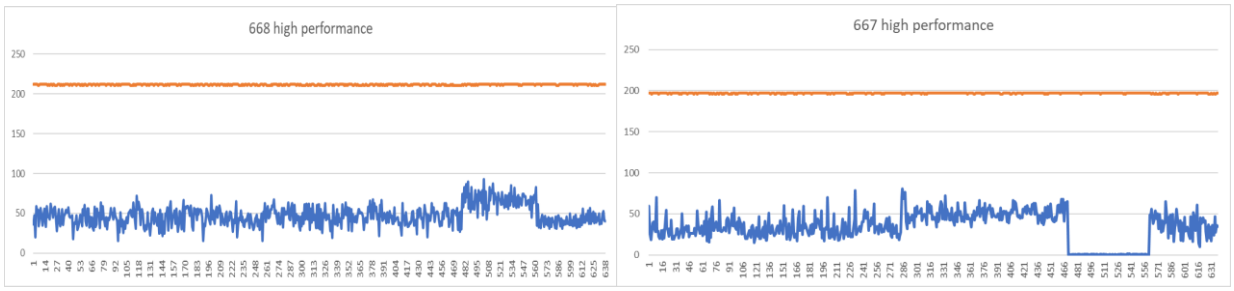
3 GEGEVENSANALYSE

In totaal hebben tijdens de startvergadering op 12 december 2019 dertien bedrijven aangegeven bereid te zijn gegevens aan te leveren over de uitgangssituatie en de powermanagement-instellingen te wijzigen. Na verlenging van de meetperiode hadden in september 2020 in totaal negen partijen gegevens aangeleverd. Alle bijdragen zijn geanonimiseerd. Verschillende door de eigenaars van de systemen geselecteerde serversystemen zijn gemonitord, en het CPU- en energieverbruik van deze systemen is geregistreerd. In de onderstaande gedeelten worden specifieke resultaten van de gemonitorde apparaten uiteengezet en geanalyseerd. Deze specifieke datasets zijn gekozen op basis van een in de betreffende dataset waargenomen situatie, instelling of effect. In hoofdstuk 5 worden vervolgens conclusies getrokken op basis van de gegevens.

3.1 STATISCHE HOGE PRESTATIES

Een van de datasets die tijdens de meetperiode zijn verkregen was afkomstig van een bedrijf dat consistent de instelling "**statische hoge prestaties**" gebruikt voor zijn HP Blade-infrastructuur.
Het effect van een statische instelling op het niveau van de hardware wordt in figuur 4 weergegeven.
Er zijn elk kwartier metingen gedaan. In elke grafiek wordt voor een week aan gegevens van één server weergegeven.

Type server	HP BL460C Gen8 (2016)
Powermanagement	
Hardware (BIOS)	**STATISCHE hoge prestaties**
Besturingssysteem	Hoge prestaties
Type CPU	Nr.
Intel(R) Xeon(R) CPU E5-2680 0 @ 2.70GHz	1
Intel(R) Xeon(R) CPU E5-2680 0 @ 2.70GHz	2
Besturingssysteem	VMware



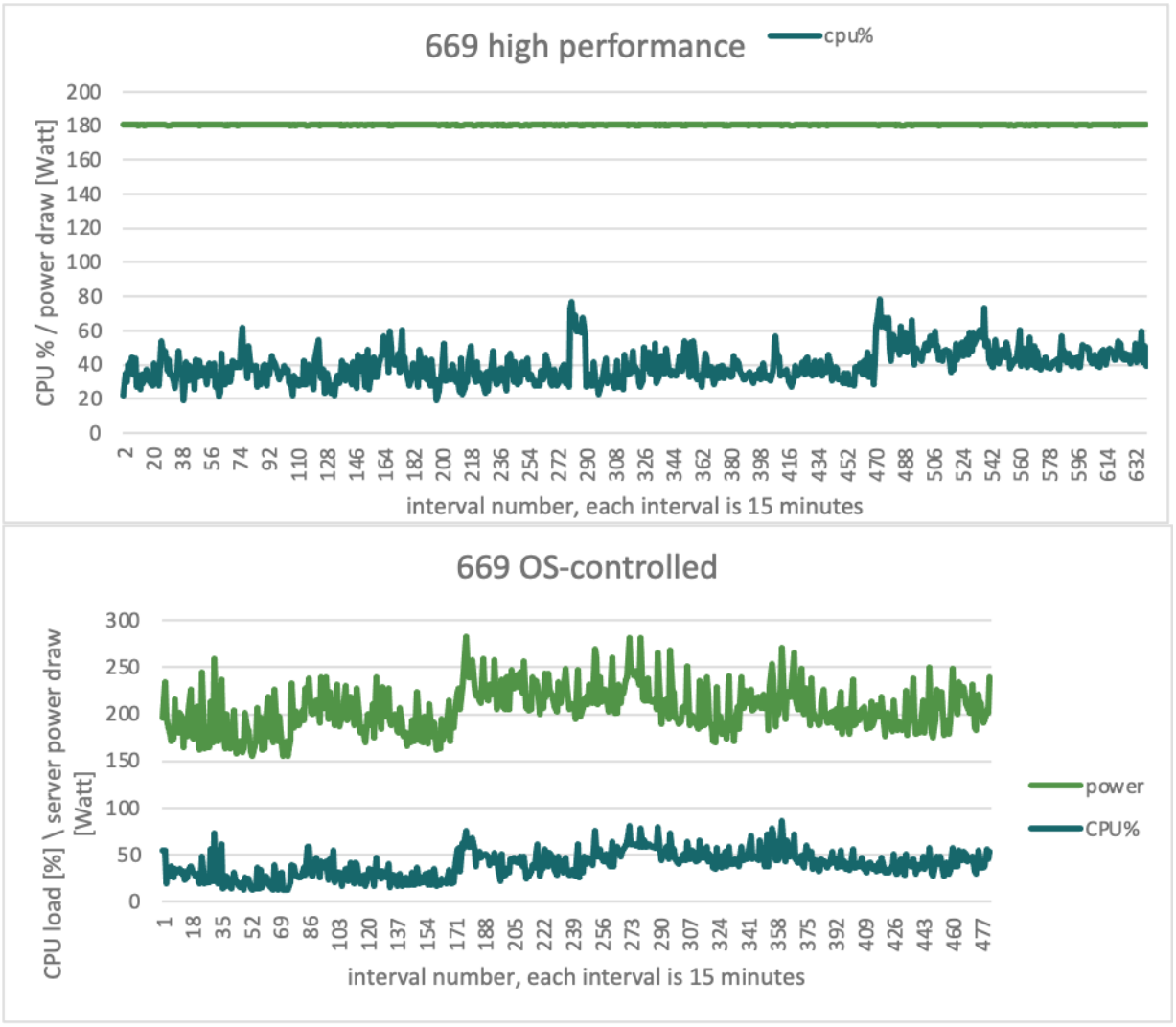
Figuur 4: twee servers in de modus statische hoge prestaties. In deze grafieken heeft de verticale as twee betekenissen. De blauwe lijn geeft de CPU-belasting weer als percentage, van 0% (inactief) tot 100% (volledig belast). De oranje lijn geeft het elektriciteitsverbruik P weer in Watt. Op de horizontale as wordt enkel het nummer van de meetinterval weergegeven.

Zoals u kunt zien is er geen variatie in het energieverbruik gemeld. De grafieken zijn specifiek geselecteerd om de omvang van het effect weer te geven. De juistheid van de CPU-belasting van nul voor server 667 gedurende een periode van 24 uur is gecontroleerd en bevestigd: het VMware-cluster heeft de werkbelasting van dit knooppunt in de betreffende periode correct vastgesteld (zie figuur 4, grafiek over server 668).

Het is zeer interessant om te zien wat er gebeurt als de instellingen worden gewijzigd naar een door het OS bestuurde modus. Hier zijn twee voorlopige waarnemingen van belang:

1. Om de besturing van het powermanagement te verplaatsen van de hardware naar het OS moest de server opnieuw worden opgestart. Dit is voor deze specifieke configuratie aangehaald als een van de redenen dat de wijziging niet op andere platforms is doorgevoerd.
2. De keuze voor deze servers was ook gebaseerd op de noodzakelijke herstart, in combinatie met het gebrek aan kennis over eventuele gevolgen voor de prestaties van de applicatie. Het geobserveerde systeem is bedoeld voor intern gebruik door de afdeling systeembeheer. Er draaien applicaties op die van machine naar machine werken, wat verklaart waarom er geen consistente variatie te zien is in de dagelijkse werkbelasting.

Powermanagement	
Hardware (BIOS)	**Bestuurd door OS**
Besturingssysteem (VMware)	Gebalanceerde prestaties



Figuur 5: één server geconfigureerd met statische (boven) en dynamische (onder)stroom. In deze grafieken heeft de verticale as twee betekenissen: de CPU-belasting wordt weergegeven als percentage, van 0% (inactief) tot 100% (volledig

belast); het elektriciteitsverbruik P van de server wordt weergegeven in Watt. Op de horizontale as wordt enkel het nummer van de meetinterval weergegeven.

Naast de grafieken kunnen er ook gemiddelden worden berekend voor de meetperiode:

Weekly averages STATIC High performance:
Power **180,9**
CPU% **39,3**

Weekly averages DYNAMIC (OS controlled):
Power **205,2**
CPU% **39,6**

Zoals u kunt zien levert dit onverwachte resultaten op. De statische hoge prestaties leiden tot een lager gemiddeld energieverbruik dan de door het OS bestuurde modus. Hoewel het onmogelijk is te testen wat er in deze situatie daadwerkelijk gebeurt, is de verwachting dat de modus "statische hoge prestaties" wordt bereikt door hoge CPU P-states niet meer toe te staan. Deze standen met hoge frequenties, soms ook wel turbomodus genoemd, kosten veel stroom, maar leveren wel hogere prestaties op. Uit de gegevens wordt zichtbaar dat het energieverbruik bij lagere belasting daalt tot onder de 180 W die wordt geassocieerd met de modus statische hoge prestaties. Bij een CPU-belasting van 40% of meer stijgt het energieverbruik van de CPU echter aanzienlijk en gebruikt het systeem 250 Watt aan energie.

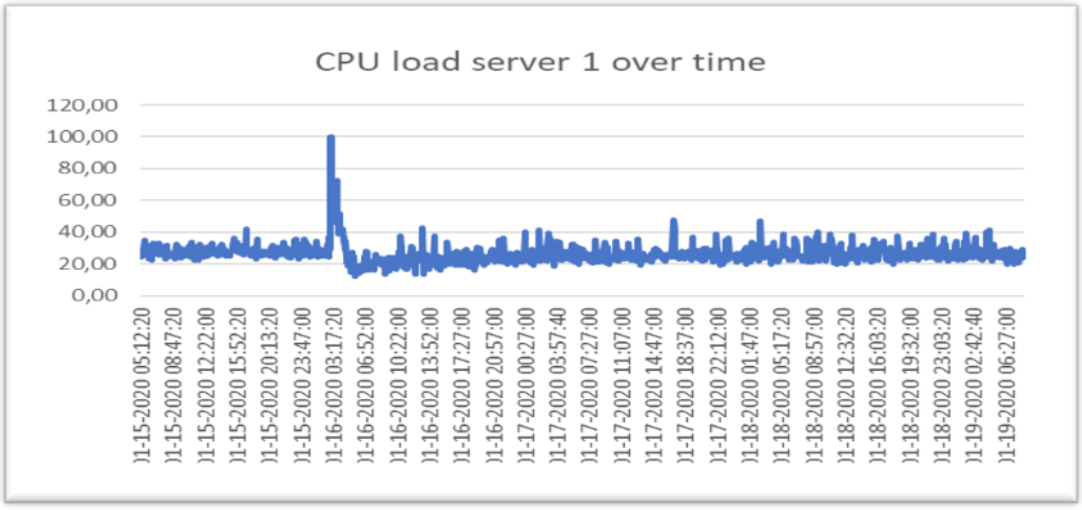
De kans is zeer groot dat de prestaties van de applicatie veel hoger liggen in de door het OS bestuurde periode dan met de statische hoge prestaties, maar er zijn geen gegevens die deze bewering staven. Wel is bevestigd dat het energieverbruik bij inactiviteit in de dynamische modus veel lager is dan in de statische modus. Dat het totale energieverbruik in dit specifieke geval stijgt, kan worden toegeschreven aan de consistent hoge belasting van de applicatie. De specifieke instelling zal waarschijnlijk leiden tot energiebesparingen bij minder zwaar belaste servers.

3.2 DYNAMISCHE PRESTATIES

De instelling die het meest voorkwam onder de gemonitorde servers was een variant van dynamische hoge prestaties. In deze modus reageert de server op af- en toenames van de belasting.

De modus "bestuurd door OS" is met twee verschillende instellingen correct toegepast op een groot VMware-cluster van een van de respondenten:
BIOS: Bestuurd door OS

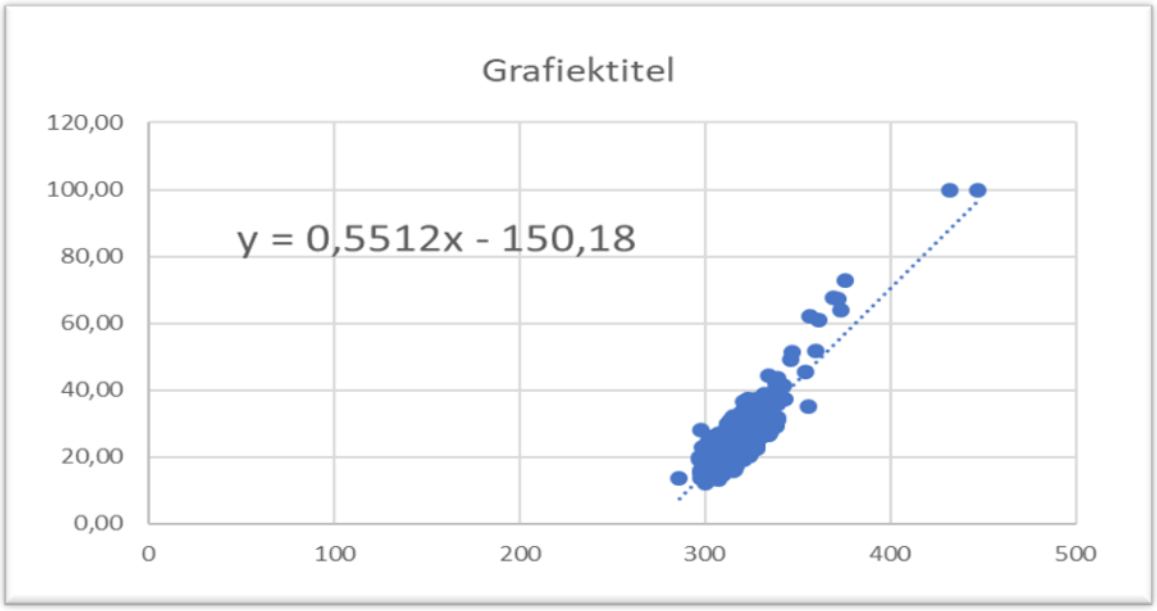
HPE BL460 Gen9
2 x Intel(R) Xeon(R) CPU E5-2697A v4 @ 2.60 GHz
+/- 60 vm's per knooppunt in het cluster van 20 knooppunten
Hypervisor: VMware ESXi 6.5.0 build-13635690
Meting 1: hoge prestaties
08-01-2020 t/m 17-01-2020 10:43
Meting 2: gebalanceerd
17-01-2020 10:43 t/m 20-01-2020



Figuur 6: CPU-belasting server 1 in de loop van de tijd

De server die voor de LEAP-pilot is gebruikt, is enkel bestemd voor intern gebruik binnen het bedrijf. Het betreft een sterk gevirtualiseerde omgeving met veel vm's per fysiek knooppunt. Duidelijk is dat er geen verband is tussen de gemiddelde totale CPU-belasting van de 89 vm's en het moment van de dag. De belasting is in feite constant. De piek (100%) doet zich voor op de 16e om 3 uur 's ochtends, maar het is niet duidelijk waarom dit zo is.

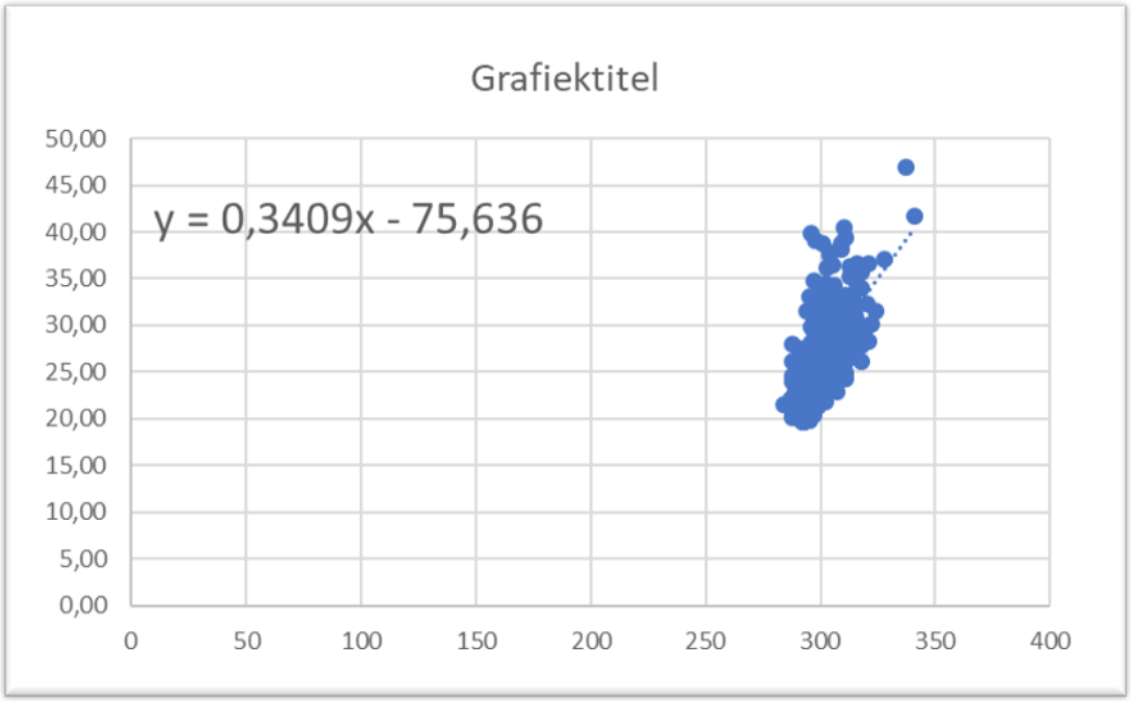
Door een grafiek te maken met daarin het verband tussen de CPU-belasting en het energieverbruik van de server kunnen de metingen handig worden weergegeven. Een duidelijk verband wordt als wenselijk gedrag beschouwd. Zelfs in de modus "hoge prestaties" is dit verband duidelijk zichtbaar, zoals in figuur 7 te zien is.



Figuur 7: CPU-belasting versus energieverbruik, server 1, hoge prestaties. In deze grafiek wordt op de verticale as de CPU-belasting van de server weergegeven en op de horizontale as het gemeten energieverbruik. Ieder punt verwijst naar een meetmoment waarop gegevens over zowel de CPU-belasting als het energieverbruik zijn verzameld (bijvoorbeeld belasting van 20%, 300 Watt).

De stippellijn is de best mogelijke lineaire weergave van de gemeten punten. Als we deze lijn extrapoleren naar een CPU-belasting van nul, levert dat een berekend energieverbruik bij inactiviteit (P_{idle}) van 273 Watt op. De foutmarge is groot, vanwege het grote aantal metingen met een CPU-belasting van 10-40%. Het gebruik van deze extrapolatiemethode wordt in paragraaf 3.5 verder toegelicht.

Van de metingen voor de instelling "energiebesparing" kan een soortgelijke grafiek worden gemaakt:



Figuur 8: server 1, energiebesparingsmodus. In deze grafiek wordt op de verticale as de CPU-belasting van de server weergegeven en op de horizontale as het gemeten energieverbruik. Ieder punt verwijst naar een meetmoment waarop gegevens over zowel de CPU-belasting als het energieverbruik zijn verzameld (bijvoorbeeld belasting van 25%, 300 Watt).

De extrapolatie van deze stippellijn leidt tot een inactief gebruik van de blade van 222 Watt. De foutmarge is ook hier behoorlijk groot, maar de invloed van de energiebesparingsmodus is duidelijk. Zelfs als de belasting voortdurend vrij hoog is, zorgt energie-efficiëntie voor een beduidend lager energieverbruik.

Een simpel gemiddelde van de gegevens uit de twee meetperioden bevestigt dit:

Server gemeten met instelling "hoge prestaties"
Gemiddeld energieverbruik: 321,4 Watt
Gemiddelde CPU: 26,97%

Dezelfde server gemeten met instelling "gebalanceerd"
Gemiddeld energieverbruik: 300,4 Watt
Gemiddelde CPU: 26,7%

Het effect van de wijziging blijft niet beperkt tot één enkel knooppunt uit de cluster, maar is ook zichtbaar in de gemiddelde belasting en het gemiddelde energieverbruik van alle 20 knooppunten in de cluster.

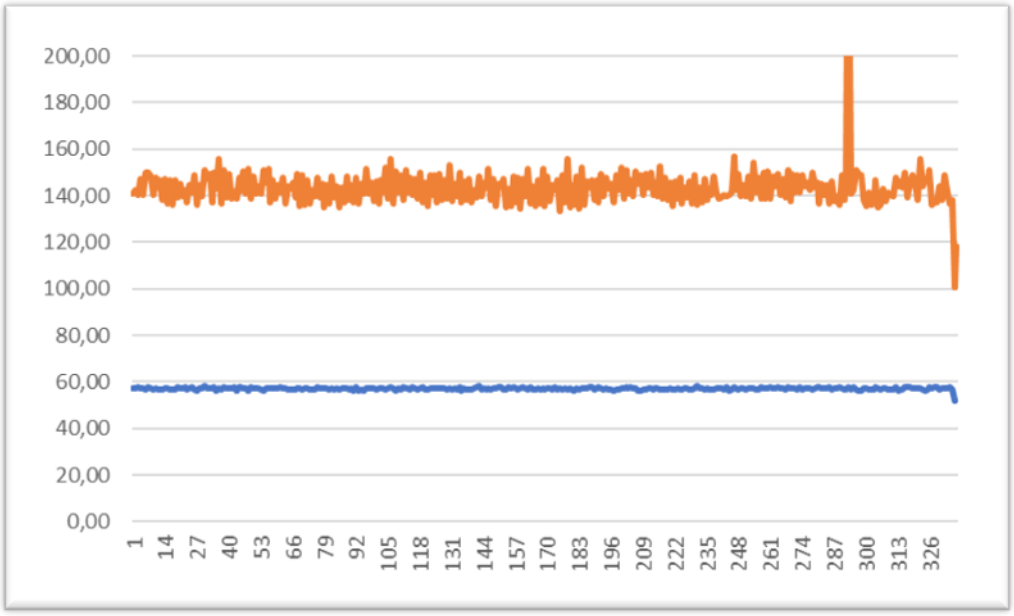
Cluster gemeten met instelling "hoge prestaties"
Gemiddeld energieverbruik: 271 Watt/knooppunt
Gemiddelde CPU: 18%/knooppunt

Dezelfde server gemeten met instelling "gebalanceerd"
Gemiddeld energieverbruik: 252 Watt/knooppunt
Gemiddelde CPU: 17%

Uit de gegevens zijn twee belangrijke dingen op te maken:

1. Zelfs in door het OS bestuurde modus en met de instelling "hoge prestaties" vertoont de server dynamisch gedrag. Dit is een duidelijk verschil ten opzichte van de situatie uit paragraaf 3.1: met de instelling "statische hoge prestaties" is er geen dynamisch gedrag te zien.
2. Zelfs met zeer hoge belasting leidt het inschakelen van de gebalanceerde modus in VMware nog tot een energiebesparing van 7%. Tijdens de periode in de energiebesparingsmodus zijn er geen nadelige gevolgen voor de prestaties gemeld.

De derde respondent die wijzigingen in de powermanagement-instellingen heeft doorgevoerd, heeft uitsluitend gebruikgemaakt van door de hardware bestuurde modi.
De betreffende servers zijn van BIOS-instelling "Efficiëntie – nadruk op prestaties" in week 1 gewijzigd naar BIOS-instelling "Minimaal energieverbruik" in week 2.
Er is bij de gemeten servers echter vrijwel geen sprake van variatie in belasting en, met uitzondering van twijfelachtige waarden, ook vrijwel geen sprake van variatie in energieverbruik.
De constante CPU-belasting die is waargenomen, strookt met de functie van de server. Zoals gezegd is de server uitsluitend voor intern gebruik bestemd en is de draaiende applicatie een applicatie voor monitoring die gegevens van andere servers verzamelt voor de beheerders.



Figuur 9: één server, geconfigureerd als "nadruk op prestaties". In deze grafiek heeft de verticale as twee betekenissen. De blauwe lijn geeft de CPU in inactieve toestand weer als percentage, van 0% (volledig belast) tot 100% (inactief). De oranje lijn geeft het elektriciteitsverbruik P weer in Watt. Op de horizontale as wordt enkel het nummer van de meetinterval weergegeven.

De meest logische manier om het effect van de wijziging van energie-instellingen te meten is door simpelweg gemiddelden van de gegevens te nemen. De volgende variabelen zijn bijgehouden:

Week 1:

Energieverbruik	CPU			
Watt	Systeem	Gebruiker	Inactief	IO Wait
143,46	19,42	21,68	57,03	1,87

Week 2:

Energieverbruik	CPU			
Watt	Systeem	Gebruiker	Inactief	IO Wait
126,04	19,38	21,64	57,11	1,88

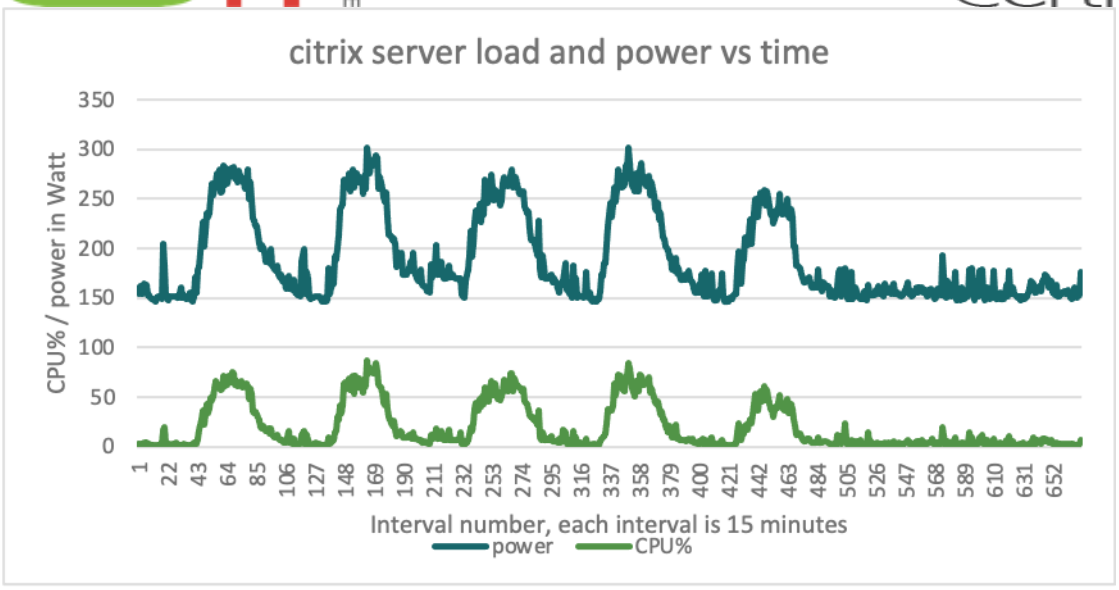
Het was duidelijk dat aanpassing van het powermanagement geen negatieve gevolgen had voor de prestaties en dat powermanagement zelfs bij een vrijwel constante belasting van 40% nog effect heeft. De kans is groot dat enkele CPU-kernen niet door de applicatie gebruikt worden (40% belasting) en dus naar een hogere C-state zijn gegaan.

Wijziging van de energie-instellingen leidt tot een besparing van 19,4 Watt (13%).

3.3 VOORSPELBARE DAGELIJKSE VERSCHILLEN IN BELASTING

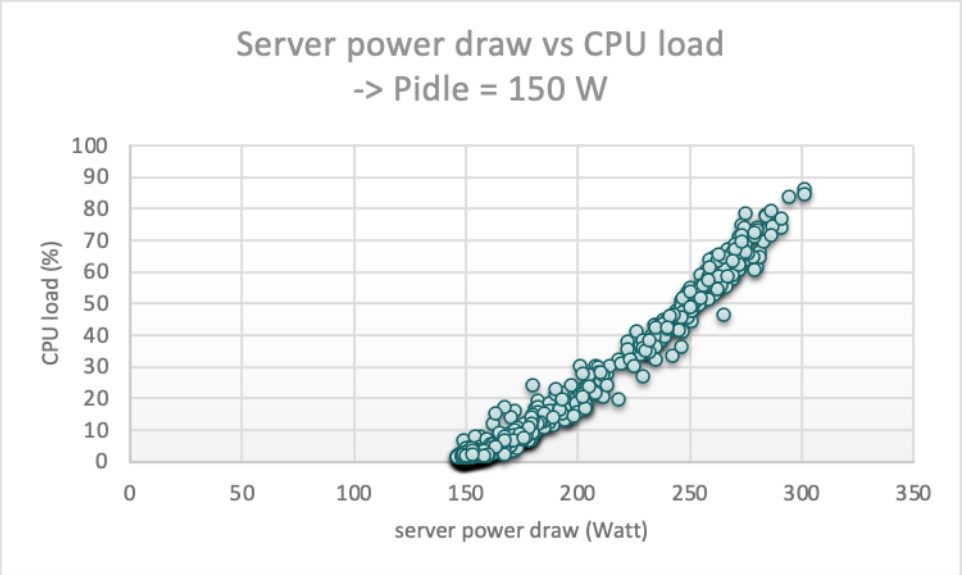
Sommige servers die onderdeel waren van de pilot vertoonden heel duidelijk gedrag dat overeenkomt met de verwachte verschillen tussen dag en nacht, zoals weergegeven in de grafieken over dagelijkse belasting van AMS-IX.

Het duidelijkste voorbeeld van voorspelbare verschillen in belasting is afkomstig van een Citrix VDI-server van een van de deelnemers. De pieken in de CPU-belasting komen precies overeen met de pieken in het energieverbruik. Onverwacht is daarbij dat de server naar verluidt in de modus "statische hoge prestaties" stond. Waarom de server zich dan vrijwel perfect volgens een dynamische modus gedraagt is niet bekend. Het is mogelijk dat de instelling niet correct is geregistreerd of dat de instelling, als deze wel klopt, om een of andere reden niet heeft gewerkt. Een andere mogelijkheid is dat, zoals ook bij andere deelnemers het geval was, de voorheen daadwerkelijk statische modus "hoge prestaties" alsnog dynamisch is. De onderliggende details van alle modi zijn niet geregistreerd en kunnen per server verschillen.



Figuur 10: Citrix-server, geconfigureerd als "hoge prestaties". In deze grafiek heeft de verticale as twee betekenissen. De groene lijn geeft de CPU-belasting weer als percentage, van 0% (inactief) tot 100% (volledig belast). De blauwe lijn geeft het elektriciteitsverbruik P weer in Watt. Op de horizontale as wordt enkel het nummer van de meetinterval weergegeven.

In figuur 10 wordt een volledige week aan metingen weergegeven, waarbij de eerste piek op maandag en de laatste op vrijdag iets lager zijn dan die op de andere werkdagen, met een aanmerkelijk sterkere daling doordat mensen naar huis gaan voor het weekend. Er zijn verschillende punten in de gegevens waarop de CPU-belasting 0 is. Daardoor kunnen we de Server Idle Coëfficiënt berekenen zonder dat we daar lineaire extrapolaties voor nodig hebben.



Figuur 11: energieverbruik versus CPU-belasting

In figuur 10 en 11 worden zowel verwacht als wenselijk gedrag weergegeven voor een gebruikersgerichte dienstverlening. Er zijn duidelijke verschillen in belasting die overeenkomen met werktijden en duidelijke verschillen in energieverbruik die overeenkomen met de verschillen in belasting. Er is geen meting uitgevoerd met een aangepast niveau van powermanagement.

Als we de gegevens voor de meetperioden optellen (zie formule 5-7), levert dat de volgende cijfers op:
Energie in inactieve toestand: 20,2 kWh
Totale energie: 31,9 kWh
Gemiddelde CPU-belasting: 19,9% (CPU inactief 80,1%)

Dit leidt tot:
SIC-% = 63,3%
Of, anders weergegeven: SIC = 2,7

Het is duidelijk dat de belasting voor iedere server die zo sterk aan kantoortijden verbonden is, bijna drie kwart van de tijd zeer laag is (weekenden en nachten). Door het dynamische energiegedrag daalt de SIC van de inactieve CPU van 80% naar 63%.
Toch wordt nog steeds meer dan 60% van de energie gebruikt als de server inactief is. Naar verwachting zou een hoger niveau van powermanagement in dit geval zorgen voor een lager energieverbruik in inactieve toestand, en dus tot een lager totaal energieverbruik en een betere SIC.

Tijdens de besprekingen met de partij die verantwoordelijk is voor de infrastructuur is voorgesteld een extra stap in te bouwen in het powermanagement: S-states. Door systeemstates kan een besturend systeem volledige systemen in een cluster uitschakelen wanneer de belasting onder een bepaalde ondergrens komt. Eventuele resterende werkbelasting kan worden verplaatst naar een knooppuntencluster dat nog actief is. Door deze states te gebruiken kan de in inactieve toestand gebruikte energie met ongeveer 10 kWh per week worden verlaagd. Daardoor daalt uiteraard ook het totale energieverbruik met 10 kWh, waardoor de SIC rond de 2 komt te liggen (SIC-% = 50%).

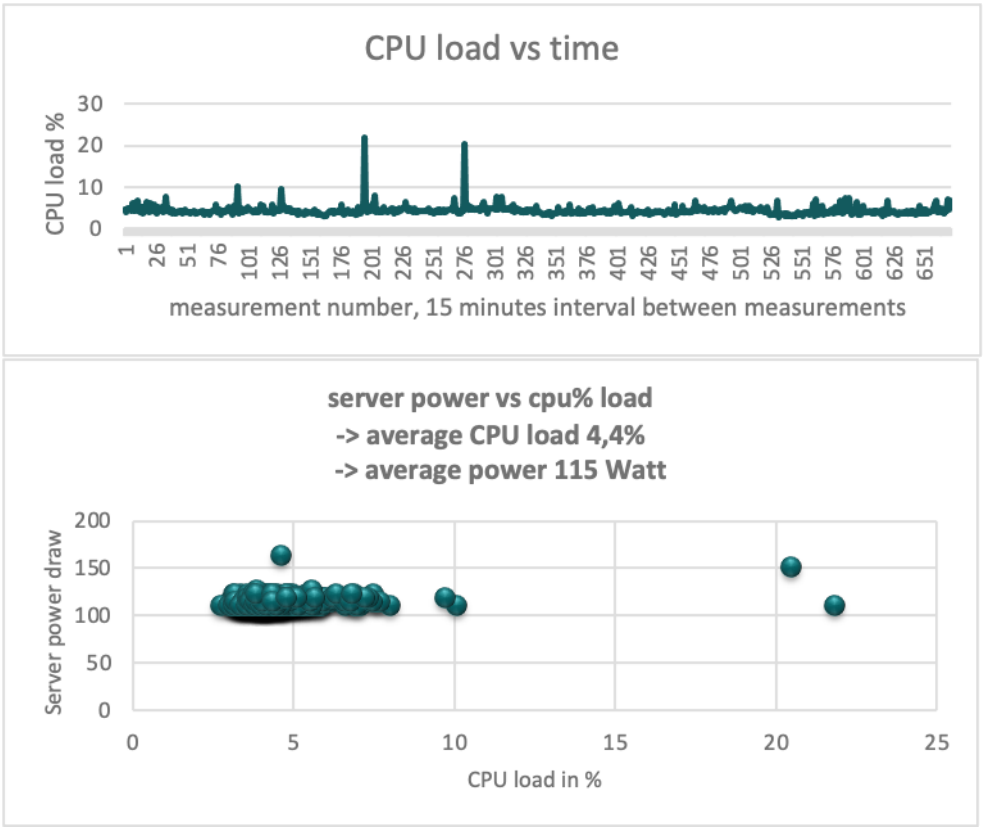
3.4 ONDERBELASTING ONTDEKT IN DE DATASETS

In de aangeleverde gegevens is ook een duidelijk verschil opgemerkt tussen de systemen die geoptimaliseerd zijn voor maximale plaatsing van virtuele machines en de systemen die specifieke doelen hebben of directer betrokken zijn bij de dienstverlening aan eindgebruikers.

Onderbelasting van servers of, anders gezegd, een grotere infrastructuur dan nodig is voor een bepaalde dienst, is nog altijd de duurste vorm van een verkeerd gebruik van middelen. Een dergelijke configuratie zorgt niet alleen voor hoge operationele kosten (onderhoud, licenties en energie), maar is ook kostbaar vanuit het oogpunt van CAPEX, aangezien er veel geld is geïnvesteerd in het datacenter en de servers, terwijl dat niet nodig was.

Voorbeelden van een te grote infrastructuur zijn in veel van de aangeleverde datasets te vinden. De onderstaande grafiek is slechts één voorbeeld van een dergelijke configuratie. Het betreffende systeem is geconfigureerd als "maximale prestaties", wat overeenkomt met een BIOS-instelling van statische hoge prestaties.

Power management:	
(BIOS)*	Max performance
CPU Type	RAM
2 x silver 4110	128 GB



Figuur 12: onderbelaste server, geconfigureerd als "statische hoge prestaties".

In figuur 12 worden twee grafieken weergegeven: CPU-belasting versus tijd (boven) en energieverbruik versus CPU-belasting (onder), waarin het CPU-% wordt geanalyseerd (waarbij 100% staat voor een volledig belast systeem). We zien dat er gedurende de meetperiode slechts driemaal een belasting van meer dan 10% is geregistreerd. De gemiddelde CPU-belasting is 4,4% en het gemiddelde energieverbruik is 115 W.

De powermanagement-instellingen zorgen voor een dynamisch spectrum dat vrijwel gelijk is aan nul. Onder dergelijke omstandigheden is P_{idle} gelijk aan P en is het SIC-% gelijk aan het CPU-% in inactieve toestand.

SIC-%= 95,6% of

SIC = 22,7 in de PUE-weergave

Zoals gezegd komt ernstige en structurele onderbelasting van CPU's veel voor. In dit document worden niet alle verkregen datasets weergegeven. Verschillende systemen uit de dataset hebben een gemiddelde CPU-belasting van minder dan 4% en halen zelfs op piekmomenten zelden de 10%. Deze systemen zouden voordeel hebben van powermanagement-instellingen, maar er kunnen veel energie en CAPEX worden bespaard door deze werkbelasting in een beter benutte (gedeelde) omgeving onder te brengen.

Het spectrum van dynamisch gedrag is in verschillende datasets behoorlijk groot, maar servers halen maar zelden consistent een belasting van 0%, zoals het geval is bij de Citrix-server uit paragraaf 3.3. Om het energieverbruik in inactieve toestand, P_{idle}, te bepalen, wat nodig is om de Server Idle Coëfficiënt te kunnen berekenen, zijn de gegevens over energieverbruik versus CPU-belasting omgevormd tot een lineaire functie. De gegevens bleken verrassend geschikt voor een dergelijke lineaire benadering over een groot aantal cijfers voor CPU-belasting in alle datasets die tijdens dit onderzoek zijn verkregen.

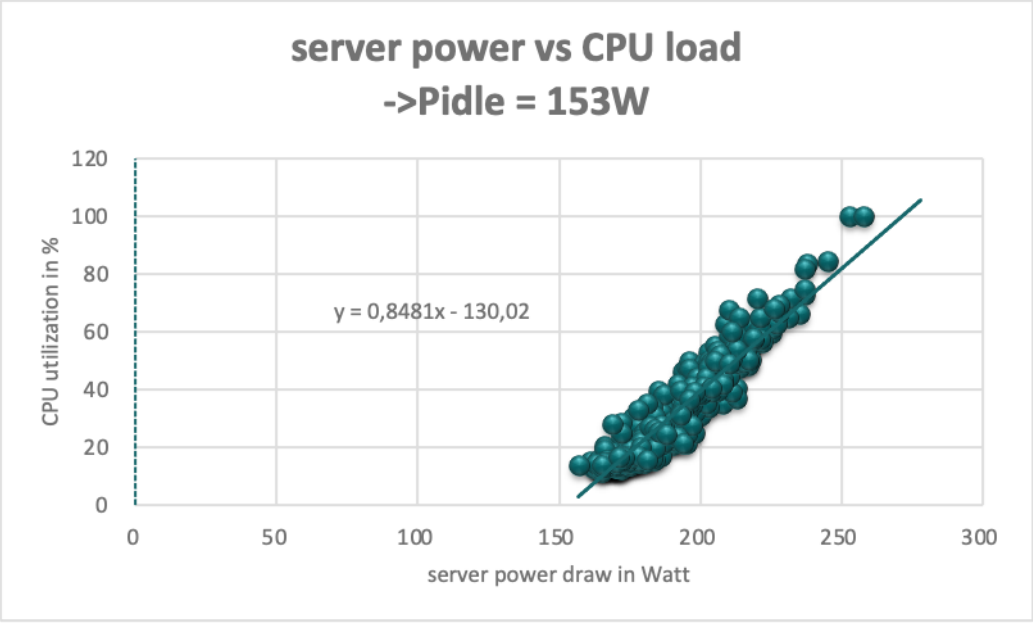
Op basis van figuur 11, de Citrix-server waarvan de belasting varieert van 0% tot 85%, en de gegevens die in verschillende online beschikbare benchmarks zijn gepubliceerd, weten we dat een dergelijke lineaire benadering niet geschikt is voor extreem hoge of lage CPU-belasting (0-10% en 90-100%), maar de gegevens laten zien dat deze weergave het waargenomen gedrag in het tussenliggende belastingsspectrum zeer dicht benadert.

De "trendlijn" die dit oplevert kan worden gebruikt om te extrapoleren naar een belasting van 0%. Deze methode hebben we gebruikt om het energieverbruik in inactieve toestand, P_{idle}, te bepalen op basis van de aangeleverde gegevens. Met behulp van formule 5 kunnen we E_{idle} berekenen, en daarmee vervolgens de SIC/SIC-%/SIC_{score}.

In figuur 13 wordt een dergelijke extrapolatie weergegeven, waarbij de stippellijn staat voor de volgende formule:

$$y = 0{,}8481x - 130{,}02 \quad (8)$$

Daarbij staat y voor de CPU-belasting (%) en x voor het energieverbruik (Watt). Instelling y= 0 levert x = 153 W op, de waarde voor P_{idle}.



Figuur 13: CPU-belasting versus energieverbruik. Server in door het OS bestuurde modus. VMware draait met gebalanceerd profiel.

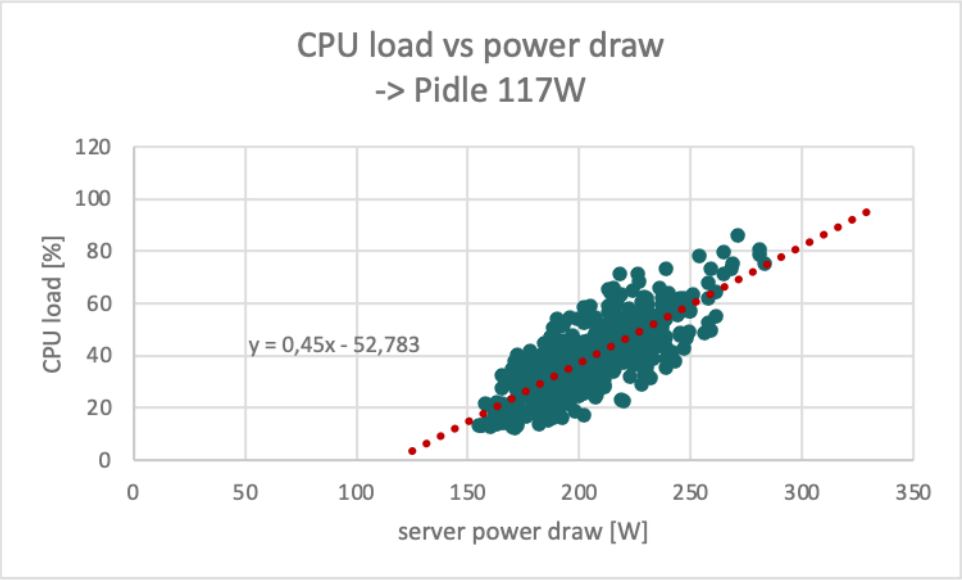
Energieverbruik inactief	Totaal energieverbruik
17,96 kWh	31,25 kWh

SIC-% = 57%
SIC = 2,3

In deze specifieke situatie was er geen meting voor een strengere powermanagement-instelling.

Wanneer we de servers analyseren waarvan de gegevens in de voorgaande paragrafen zijn weergegeven, zien we hoe gevoelig de SIC is voor wijzigingen in de powermanagement-instellingen.

Na overschakeling van statische hoge prestaties naar dynamische gebalanceerde prestaties (figuur 5, paragraaf 3.1) kunnen we de SIC voor server 669 berekenen:



Figuur 14: CPU-belasting versus energieverbruik voor server 669 in gebalanceerde modus, P_{idle} op basis van extrapolatie 117 Watt

Uit de gegevens over server 669 kunnen we de volgende cijfers opmaken:

Totaal energieverbruik in de betreffende periode: 24,5 kWh

Totaal energieverbruik in inactieve toestand: 8,43 kWh

Gemiddelde deel van de tijd dat CPU inactief is: 60,4%

Met behulp van de juiste formules berekenen we het SIC-% en de SIC in gebalanceerde modus als volgt:

SIC-% = 8,43/24,5 = 34,4%

SIC = 1,5

Zoals te zien in figuur 5 geldt, wanneer dat systeem is ingesteld op statische hoge prestaties, dat P_{idle} = P en dus dat:

SIC-% = CPU-% in inactieve toestand = 60,4%

SIC = 2,53

Deze berekeningen lijken erop te wijzen dat de SIC een sterke indicator is van en gevoelig is voor grote wijzigingen in het powermanagement. Ook is uit de berekeningen op te maken dat een wijziging naar een dynamische energiemodus niet altijd hoeft te leiden tot absolute besparingen, maar wel zorgt voor een verschuiving naar een nuttiger gebruik van energie, al moet er nog verder onderzoek worden gedaan naar de gevolgen voor de prestaties van de applicatie.

Bij de in paragraaf 3.2 genoemde systemen is een veel kleinere wijziging in de powermanagement-instellingen doorgevoerd. Deze systemen stonden al in een dynamische modus en zijn gewijzigd naar een strengere instelling waarbij een diepere slaapstand beschikbaar is. Voor server 1 (figuur 7 en 8) leverde dit de volgende cijfers op:

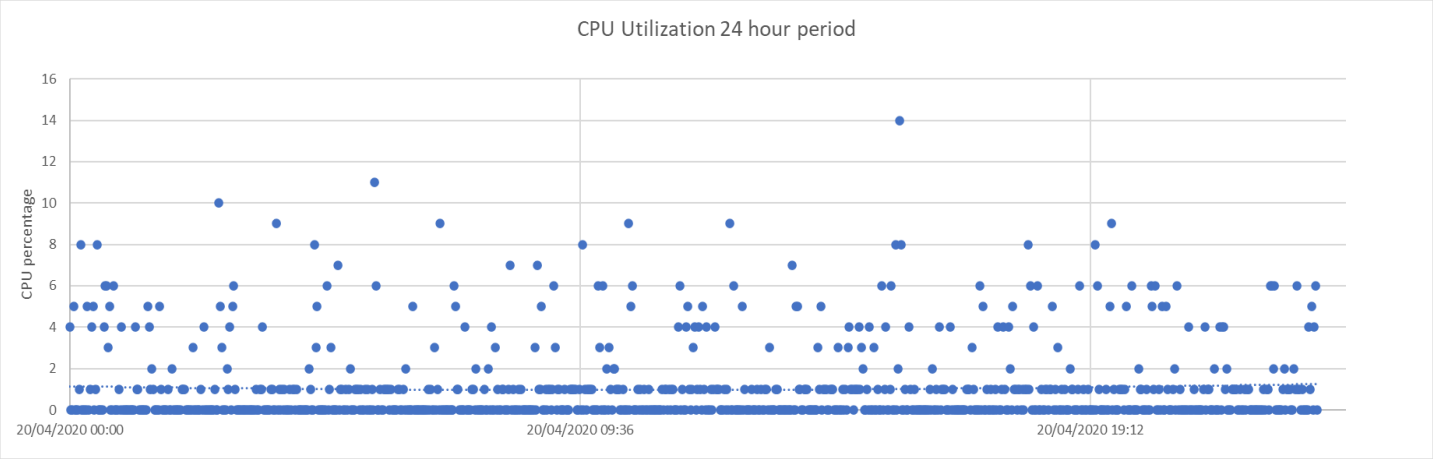
Server gemeten met instelling "dynamische prestaties":
Gemiddeld energieverbruik: 321,4 Watt
Gemiddelde CPU: 26,97%
SIC-% = 64% (SIC = 2,8)

Dezelfde server gemeten met instelling "dynamisch energie-efficiënt":
Gemiddeld energieverbruik: 300,4 Watt
Gemiddelde CPU: 26,7 %
SIC-% = 64% (SIC = 2,8)

In dit specifieke geval blijkt de SIC compleet ongevoelig te zijn voor de wijziging in het powermanagement. Zoals gezegd is er meer onderzoek nodig, maar het lijkt erop dat in deze specifieke modus zowel het maximale energieverbruik (bij belasting van 100%) als het energieverbruik in inactieve toestand lager wordt. De verlaging van het energieverbruik in inactieve toestand kan worden toegeschreven aan het gebruik van diepe C-states, die uitsluitend worden gebruikt bij inactieve CPU-kernen en dus geen invloed hebben op de prestaties. Daarentegen is de verlaging van het maximale energieverbruik wellicht toe te schrijven aan de afschaffing van P-states voor hoge prestaties (turbomodi). Dit zou wel invloed hebben op de maximale prestaties van de server. In het tussenliggende spectrum (tot een CPU-belasting van 80-90%), waar nog voldoende speelruimte is, kan het effect van de afwezigheid van turbomodi op de prestaties van de applicatie zo klein zijn dat het niet kan worden waargenomen. Deze waarnemingen geven, zoals gezegd, duidelijk aan dat er uitgebreider onderzoek nodig is waarin ook de prestaties van de applicatie worden gemeten.

Op basis van de gegevens van de verhoogde powermanagement-instelling in de hardware die in figuur 9 (paragraaf 3.2) is weergegeven, kan geen SIC worden berekend. De variatie in de CPU-belasting is in deze specifieke situatie zo klein dat het niet mogelijk is veranderingen in het energieverbruik waar te nemen. P_{idle} kan dus niet worden bepaald.

Tot slot kan de SIC worden gebruikt om het gedrag van de server te analyseren wanneer er geen CPU-statistieken beschikbaar of bruikbaar zijn. In een dergelijk geval bij een van de deelnemers zijn de CPU-statistieken waarschijnlijk met een te korte interval geregistreerd. Aangezien het onderzochte systeem sterk onderbelast is, vertonen de CPU-statistieken onjuist gedrag, dat voornamelijk uit een belasting van 0% bestond. Interessant is dat het patroon van het energieverbruik van de server veel vlakker is, wat de verwachte belasting van de server weergeeft. De betreffende server wordt gebruikt voor kantoorproductiviteit en is dus 's nachts nagenoeg inactief en, zoals uit de metingen blijkt, overdag zeer licht
belast.



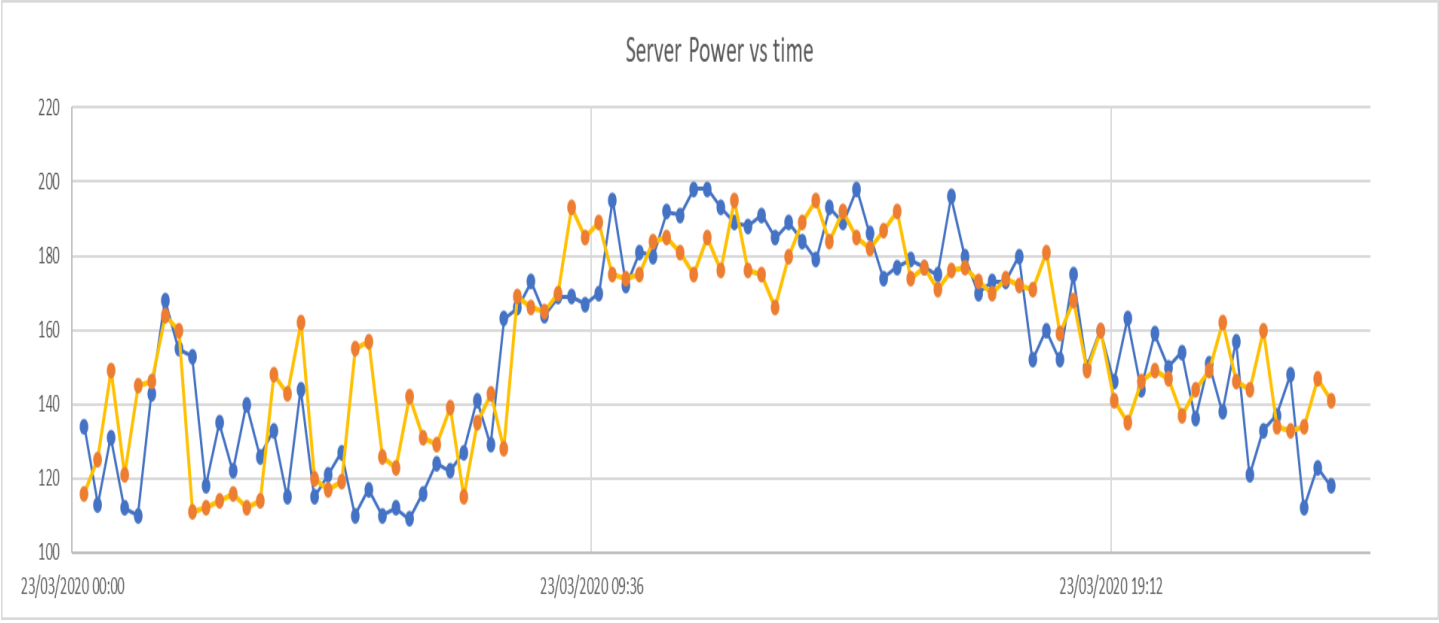
Figuur 15: registratie van CPU-belasting gedurende 24 uur

Het gemiddelde van de metingen is een belasting van 1,0%, met één piek van 14%.

Het betreffende systeem is in vier verschillende toestanden gemeten:

De eerste twee metingen zijn gedaan om te controleren of de nieuwe verwachting klopte dat de instelling voor hoge prestaties niet meer statisch was. Ook dienden zij als een vergelijking tussen door hardware bestuurde en door het OS bestuurde hoge prestaties.

Deze eerste twee metingen worden in figuur 16 weergegeven, allebei met de instelling hoge prestaties. Blauw is door de hardware bestuurd (BIOS in hoge prestaties), oranje is door het OS bestuurd. Dat laatste houdt in dat het systeem opnieuw is opgestart met de BIOS-energie-instelling "door OS bestuurd" en de OS-energie-instelling "hoge prestaties".



Figuur 16: energieverbruik server. Blauwe lijn is door hardware bestuurd, oranje lijn is door OS bestuurd.

Uit figuur 16 is op te maken dat de onderzochte server tussen middernacht en het begin van de werkdag om 7.30 uur nagenoeg inactief is. Op basis van die periode kunnen we schatten dat P_{idle} 110 W is.

Het energieverbruik van dit systeem in de betreffende periode van 24 uur kan worden berekend als het gebied onder de lijn. In beide gevallen komt dit gebied uit op:

E_{totaal} = 3,72 kWh

Aangezien de CPU-belasting van het systeem zeer laag is, schatten we het energieverbruik in inactieve toestand op 110 W x 24 uur.

E_{idle} = 2,64 kWh

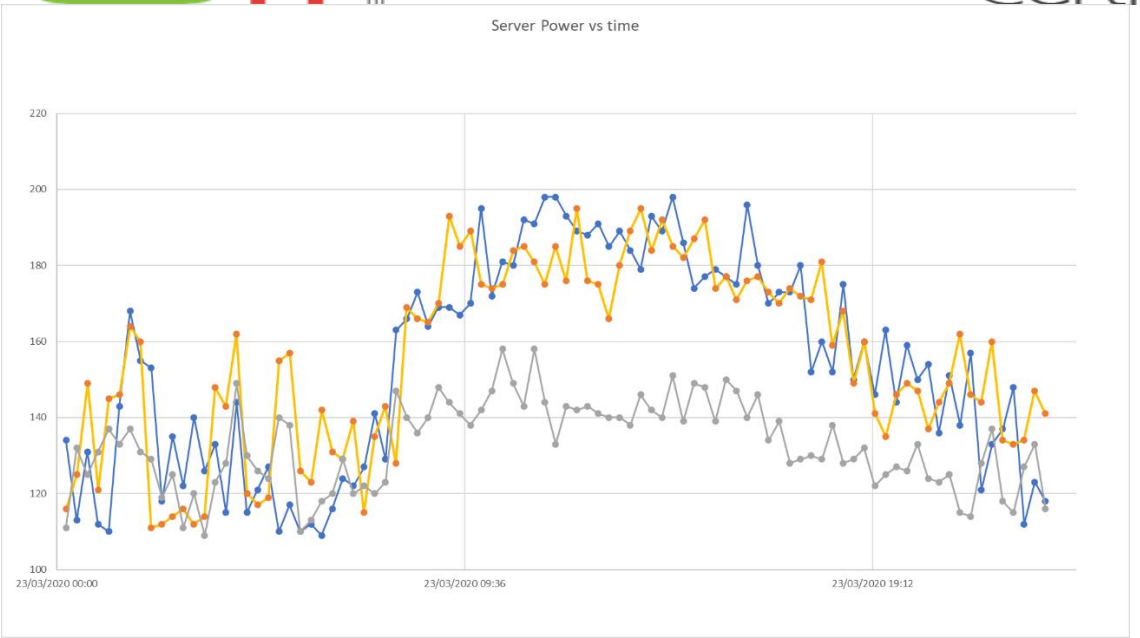
De SIC is dan 71%.

Aangezien de gemiddelde CPU-belasting zeer laag is (99% inactief), kunnen we het positieve effect van het dynamische energiegedrag zien.

Na bespreking van deze resultaten met de betreffende deelnemer is er een wijziging doorgevoerd in de powermanagement-instelling van het OS. De derde meting is uitgevoerd met het powermanagement van het OS op "gebalanceerde prestaties" en de vierde met de OS-instelling op "laag energieverbruik".

Tijdens deze meetperioden zijn er geen veranderingen in de prestaties van de applicatie gemeld. Gezien de lage belasting van het systeem en de eerdere vaststelling dat de hogere CPU-turbomodi ook met de gebalanceerde energie-instelling beschikbaar zijn, viel te verwachten dat de waargenomen prestaties beter zouden zijn met de gebalanceerde instelling, maar er zijn geen onafhankelijke metingen gedaan van de prestaties van de applicatie.

Het effect op het energieverbruik van het systeem is echter zeer groot. De derde meting, met de gebalanceerde instelling, is in de grafiek samengevoegd met de voorgaande metingen, wat de grijze lijn oplevert:



Figuur 17: Metingen voor hoge prestaties en gebalanceerd

Interessant is dat het energieverbruik in inactieve toestand vroeg op de ochtend niet significant veranderd is: dit ligt nog steeds rond de 110 W. Het energieverbruik in actieve toestand is echter sterk verlaagd. Het totale energieverbruik is over een periode van 24 uur met 14% afgenomen.

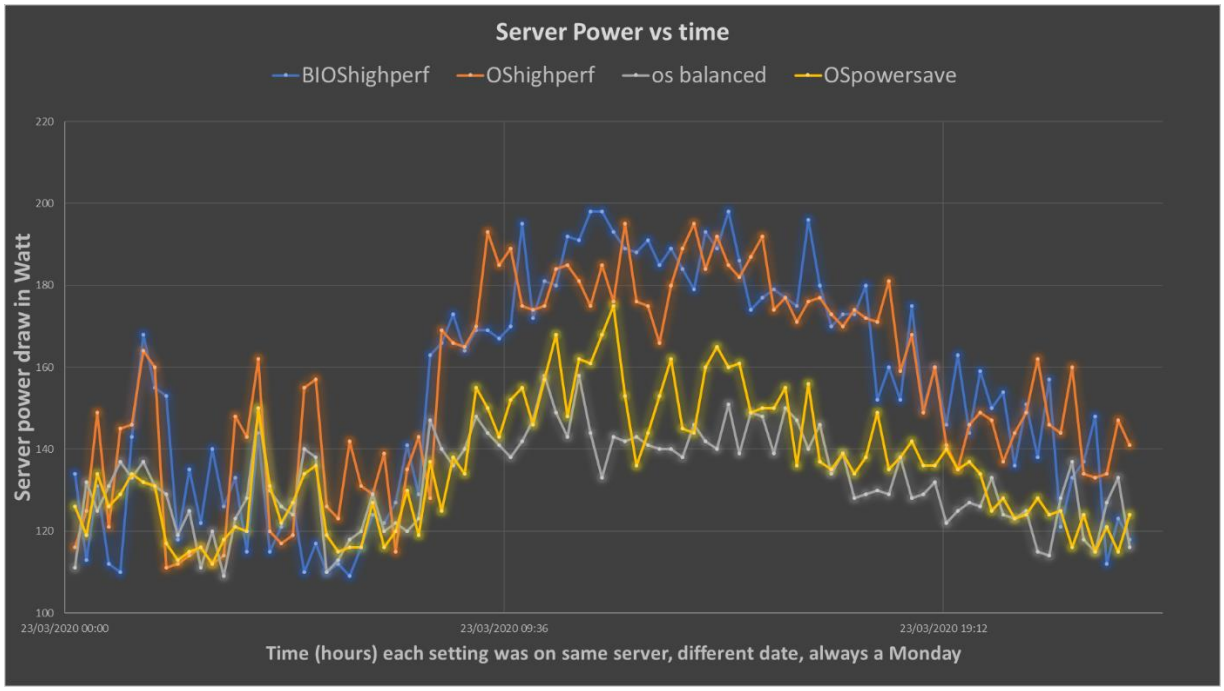
Het gemiddelde energieverbruik levert E_{totaal} = 3,18 kWh op.
Aangezien de CPU-belasting van het systeem niet veranderd is, geldt: E_{idle} = 2,64 kWh

De SIC is dan 83%.

Dat de SIC is gestegen (wat onverwacht is), heeft te maken met de blijkbaar onjuiste aanname dat het energieverbruik overdag volledig toe te schrijven is aan de CPU-belasting.
Het resultaat voor het energieverbruik is zeer positief. Interessant genoeg is het effect van powermanagement bij deze lage CPU-belasting niet te zien aan het energieverbruik in inactieve toestand, maar doordat er sneller naar de inactieve toestand wordt teruggeschakeld. Het effect is dus juist zichtbaar in het energieverbruik in actieve toestand.

Gezien de resultaten die de gebalanceerde modus had opgeleverd, was het interessant om te testen of de server nóg minder energie zou gebruiken wanneer de energie-instelling in het OS op "energiebesparing" werd gezet. De modus "energiebesparing" verlaagt de maximale prestaties van de server door overklokken (turbomodi) te beperken. In dit geval, met een lage belasting, zou het voorkomen van overklokken geen problemen opleveren, want de maximale CPU-prestaties waren nooit nodig. Aan de andere kant zou het effect op het energieverbruik ook klein zijn omdat de hoogste turbomodi zelden ingeschakeld werden. Het interessantste deel van de meting was de inactieve periode (van middernacht tot 7.00 uur). De verwachting was dat de energiebesparingsmodus diepere C-states mogelijk zou maken en dus het energieverbruik van de server in inactieve toestand zou beperken.

Na besprekingen met de betreffende deelnemer is er afgesproken dat er een vierde meting zou worden uitgevoerd op maandag 20 september 2020, waarbij de instelling energiebesparing op de gehele Hyper-V-cluster zou worden toegepast.
De resultaten worden, samen met de eerdere metingen, in figuur 18 weergegeven.



Figuur 18: Metingen inclusief de modus "energiebesparing"

De resultaten van de energiebesparingsmodus waren niet waar we op hoopten. Als we de schommelingen in energieverbruik in aanmerking nemen, verschilt de gele lijn (energiebesparing) niet significant van de grijze (gebalanceerd). Gedurende de periode van middernacht tot 7.00 uur overlappen de lijnen zelfs. Het kleine verschil overdag is waarschijnlijk het gevolg van kleine verschillen in de CPU-belasting.

Ook hier zijn er geen duidelijke gevolgen voor de prestaties van de server gemeld, maar gezien de profielen lijkt het logischer om voor deze server de voorkeur te geven aan de gebalanceerde modus.

Naar aanleiding van de metingen tijdens de LEAP-pilot heeft deze deelnemer het advies opgevolgd en zijn Hyper-V-cluster permanent overgeschakeld naar de gebalanceerde modus.

4 KWALITATIEVE ANALYSE VAN GESPREKKEN

Nadat de gegevens waren geanalyseerd, is er contact opgenomen met alle dertien coalitiepartners die aan de pilots deelnamen, om de resultaten van de analyse te bespreken of, als er geen gegevens waren aangeleverd, te bespreken met welke redenen en belemmeringen dit te maken had. Het doel van deze semi-gestructureerde gesprekken was om te achterhalen waarom er voor bepaalde instellingen was gekozen, of er tijdens de tweede fase van de pilot veranderingen in de prestaties van de applicatie waren waargenomen, en wat er zou moeten gebeuren om ervoor te zorgen dat powermanagement op bredere schaal wordt toegepast.

Tijdens deze gesprekken hebben we gezien dat de deelnemers de energie-efficiëntie van hun ICT graag willen verbeteren. In veel gevallen hebben de pilots het bewustzijn van het energieverbruik in verband met gegevensverwerking binnen deze organisaties vergroot. Toch bleek het aanleveren van de juiste gegevens lastiger dan verwacht. Dit lijkt te maken te hebben met een combinatie van de volgende zaken:

- Gebrek aan de noodzakelijke technische kennis om de juiste pilotgegevens aan te leveren:
  - Kennelijk is er te weinig kennis over de rol van virtualisatie en powermanagement met betrekking tot kosten en energieverbruik. Dit leidt tot grote inefficiënties in de installatie en het beheer van servers.
- Vooroordelen over powermanagement:
  - Alle coalitiepartners noemden de gevolgen van powermanagement voor de prestaties van de applicatie als reden om de instellingen niet te wijzigen. Zelfs de partijen die in week 2 strengere powermanagement-instellingen hebben gebruikt voor hun servers, zijn teruggegaan naar de oorspronkelijke instellingen ondanks dat zij geen verandering in de prestaties van de applicatie hadden waargenomen.
- Gebrek aan prioriteit en beleid:
  - Maar weinig organisaties hebben beleid voor het gebruik van powermanagement, en als dit er wel is, staat er meestal in dat de modus voor hoge prestaties moet worden gebruikt.
  - We hebben weinig reacties ontvangen, ondanks instructies, beschikbare ondersteuning (ook van VMware- en hardwareleveranciers) en herinneringen.
  - Er is nauwelijks gebruikgemaakt van de LEAP-helpdesk: in totaal hebben slechts vier organisaties contact met de helpdesk opgenomen. Ook hebben maar een paar organisaties hardware- en softwareleveranciers om hulp gevraagd.

Hardwareleveranciers ontwikkelen verschillende technische of organisatorische oplossingen om de energie-efficiëntie van hardware te verbeteren. Ook zijn hardware- en softwareleveranciers (leveranciers van besturingssystemen) zich ervan bewust dat er niet veel gebruik wordt gemaakt van hun trainingen en instructies met betrekking tot het gebruik van powermanagement. Deze zijn beschikbaar, maar voor veel organisaties zijn ze moeilijk te vinden en te begrijpen.

5 AFSLUITING

Als onderdeel van LEAP track 1 heeft een aantal bedrijven gegevens aangeleverd die nuttig zijn gebleken voor het analyseren van de potentiële energiebesparingen via powermanagement voor servers. De gegevens zijn echter niet doorslaggevend. In bepaalde gevallen heeft de wijziging naar een dynamische energiemodus tot een hoger gemiddeld energieverbruik geleid in een meetperiode. In andere gevallen zorgen zowel de door het OS bestuurde als de door de hardware bestuurde energiebesparingsmodus voor duidelijke energiebesparingen.

Voordat we onze waarnemingen beschrijven, moet er echter een belangrijke opmerking worden gemaakt. Het aantal fysieke servers dat momenteel in Nederland actief is, ligt rond de 1 miljoen. In dit onderzoek is naar 60 daarvan gekeken. Bovendien zijn deze servers niet willekeurig geselecteerd: zij zijn door de LEAP-coalitiepartners geselecteerd op basis van toegankelijkheid en/of hun specifieke niet-cruciale functies in de gehele IT-infrastructuur.

Het is daarom niet juist of realistisch om gemiddelden te bepalen op een hoger niveau dan dat van de afzonderlijke machines. De gemiddelde CPU-belasting die in dit document wordt gebruikt, geeft uitsluitend het gemiddelde van de betreffende machine weer, en er kunnen geen conclusies worden getrokken over de algemene stand van zaken van de ICT, de energie-instellingen en de potentiële besparingen.

5.1 WAARNEMINGEN

Niet alle gegevens die in de pilot zijn verkregen, zijn in het vorige hoofdstuk besproken. Wel zijn alle gegevens geanalyseerd en gebruikt als achtergrond voor de waarnemingen en conclusies.
In de onderstaande tabel worden de verkregen gegevens samengevat, indien van toepassing onder vermelding van de bijbehorende paragraaf.

IT-environment			OS		CPU%	Power [W]	Delta compared with measurement A	Comment
server type	workload	Bios-setting	layer	OS-setting	%	(avg)		
machine to machine	mixed	Static High Performance	Vmware	High Performance	39,3%	181		
machine to machine	mixed	OS-controlled	Vmware	Balanced performance	39,6%	205	13%	see paragraph 3.1
End user	mixed, average over all 20 nodes is shown	OS-controlled	Vmware	High Performance	18,0%	321		
End user	mixed, average over all 20 nodes is shown	OS-controlled		Balanced	17,0%	300	-7%	
machine to machine	monitoring	favor performance			41,0%	143		
machine to machine	monitoring	low power	Linux	throughput	41,0%	126	-13%	
machine to machine	monitoring	OS Control	Linux	power save	43,0%	160		
machine to machine	monitoring	OS control	Linux		43,0%	169	5%	there is not enough data to explain the discrepancy
enduser	VDI	high performance		n.a.	20,0%	190		a variation of machines were measured with varying settings for power management, the chosen example reflects dynamic high performance. Loading of this machine varried between 0 and 90% with a office hour pattern.
end user	frontend averaged 7 servers	Max performance		n.a.	5,4%	130		two servers show 1% average CPU load, severe underload
end user	backend averaged 4 servers	Max performance		n.a.	15,5%	580		average CPU load on one server 38%, other 3 below 10%
End user	3 servers on-site	Dynamic High Performance	Hyper-V	High performance	unknown, <<10%	154		
End user	3 servers on-site	OS controlled	Hyper-V	High performance	unknown, <<10%	155	0%	
End user	3 servers on-site	OS controlled	Hyper-V	Balanced performance	unknown, <<10%	132	-14,30%	savings when idle are 0%, office hour saving 23% saving
End user	3 servers on-site	OS controlled	Hyper-V	Powersave	unknown, <<10%	136	-12%	difference with balanced mode is negligable
enduser	storage	high performance	CentOS	n.a.	1,0%	72		storage server for HPC support (sporadic use)
enduser	storage	high performance	CentOS	n.a.	5,5%	228		storage server for HPC support (sporadic use)
enduser	mixed	high performance	CentOS	n.a.	9,0%	311		low utilization, storage and test vm's
enduser	HPC	high performance	CentOS	n.a.	79,0%	330		extremely high utilization consistent with HPC
enduser	HPC	high performance	CentOS	n.a.	79,0%	406		extremely high utilization consistent with HPC
enduser	mixed, averaged over 5 nodes	high performance	hyperV	n.a.	4,6%	346		extreme low utilization, customer is fixed on high performance due to advice from software vendor.
enduser	mixed averaged 5 nodes	high performance	Vmware	n.a.	40,0%	357		well utilized Vmware cluster, possibly profit from different power setting like case 2 cluster

Tabel 1: samenvatting van de gegevens

Uit de verkregen gegevens kan een aantal zaken worden opgemaakt die nuttige inzichten opleveren in de bruikbaarheid van powermanagement voor servers en in de besparingen die in specifieke situaties kunnen worden gerealiseerd, zowel door toepassing van powermanagement als door virtualisatie in combinatie met consolidatie van de werkbelasting, zoals in eerdere rapporten is beschreven:

- De meerderheid van de respondenten geeft aan dat hun servers dynamisch energiegedrag vertonen. Bij deze modi is het energieverbruik van de betreffende servers afhankelijk van de werkbelasting.
- Alle respondenten passen standaard enige vorm van de instelling "hoge prestaties" toe. Dit is naar aanleiding van de pilots niet veranderd.
  De genoemde redenen dat de instelling "hoge prestaties" wordt gebruikt, zijn vaak semantisch. De naam van de instelling doet vermoeden dat dit de meest logische optie is of dat deze instelling door de hardware- of softwareleverancier wordt geadviseerd.
  Interessant genoeg is dit advies jaren geleden gegeven, soms als reactie op een IT-incident, soms a priori. Aangezien dit advies nooit is ingetrokken, zijn de instellingen overgenomen door de huidige generatie servers.
- Veel respondenten passen tegenstrijdige instellingen toe op het niveau van BIOS en OS. Uit de gesprekken met de respondenten kan worden opgemaakt dat de belangrijkste reden hiervoor een gebrek aan kennis met betrekking tot de instellingen voor powermanagement is.
- Als men het powermanagement instelt op modi die meer energie besparen, leidt dit zelfs op druk bezette serverknooppunten tot ongeveer 10% energiebesparing. Tijdens het testen van deze energiebesparingsmodi zijn er geen nadelige gevolgen voor de prestaties gemeld.
- De omschakeling van statische naar dynamische instellingen voor hoge prestaties leidt niet automatisch tot energiebesparingen.
- Zelfs de best bezette servers besteden meer dan een derde van hun energieverbruik aan "idle cycles", terwijl dit bij de slechtst bezette servers bijna 99% is.

Eerdere schattingen van de mogelijke energiebesparingen door verbeterd powermanagement lagen hoger dan het waargenomen gemiddelde van 10%, namelijk rond de 20-40%. In specifieke gevallen, zoals de situatie die in paragraaf 3.4 is beschreven, blijkt dat er grotere besparingen kunnen worden gerealiseerd. In het specifieke geval van het lage gebruik daalde het gemiddelde energieverbruik tijdens de werkzaamheden overdag met 23%, terwijl het energieverbruik 's nachts nauwelijks veranderde. Daaruit blijkt dat besparingen samenhangen met de werkbelasting. Een ander element waardoor de besparingen lager waren dan vooraf verwacht, is het feit dat de modus "hoge prestaties" niet meer statisch is. Bijna alle servers vertoonden dynamisch energiegedrag, zelfs wanneer de instellingen op "hoge prestaties" stonden. Dit dynamische gedrag levert op zichzelf al aanzienlijke besparingen op. De vastgestelde besparingen van 10% door de omschakeling naar een gebalanceerde modus kunnen als "extra" worden gezien.

Daarnaast is duidelijk (als bij-effect van de pilot) dat virtualisatie een enorme invloed heeft op energie-efficiëntie en voor servers nog altijd meer mogelijkheden biedt dan powermanagement. Dit geldt met name voor servers die (zwaar) onderbenut worden. Virtualisatie en consolidatie van tien of meer van deze servers is gemakkelijk te doen en leidt tot een energiebesparing die niet rond de 10%, maar een orde van grootte hoger ligt.

De sterk benutte servers uit ons onderzoek ondersteunden tot 90 vm's per fysiek knooppunt. Als er overal sprake zou zijn van een hoge mate van consolidatie van de werkbelasting, zou dit leiden tot een zeer sterke daling van het aantal actieve fysieke servers en dus tot een enorme vermindering van het energieverbruik, laat staan financiële investeringen.

Uit de gegevens kunnen we ook opmaken dat een constante hoge CPU-belasting (tot 85%) niet tot problematisch gedrag leidt. Verschillende systemen hebben voortdurend een hoge belasting en de applicaties lijken op die machines zonder problemen te draaien.

Tot slot hebben we ontdekt dat de Server Idle Coëfficiënt (SIC) een nuttige maatstaf is voor de energieverspilling van computers. De SIC bleek niet erg gevoelig te zijn voor kleine wijzigingen in het powermanagement, maar is desondanks een veelbelovend instrument om te bepalen of een server effectief wordt gebruikt, via een hoge CPU-belasting en, zeer belangrijk, goede dynamische spectra van deze servers. Hoe groter het dynamische spectrum, des te beter sluit het energieverbruik aan bij de belasting van de server. De best benutte servers hadden een SIC van 1,5 tot 3, maar bij de onderbelaste servers lag de SIC tussen de 20 en 100. De SIC in percentages noteren leverde weinig reactie op, dus het is effectiever om deze op dezelfde manier te noteren als de PUE, namelijk als een getal tussen de 1 en oneindig.

5.2 CONCLUSIES

Met de eerder genoemde voorbehouden in het achterhoofd kunnen op basis van de resultaten van de pilot de volgende conclusies worden getrokken:

In vervolggesprekken met de respondenten over de redenen en belemmeringen die ervoor zorgden dat zij geen energiebesparingsmodi gebruikten, kwamen zeer consistente antwoorden naar voren: er zijn nog altijd grote zorgen over prestatieverliezen bij het gebruik van energiebesparing, zelfs als niets erop wijst dat hier sprake van is.

In verschillende gevallen geven klanten aan van hun softwareleveranciers of systeembeheerders te horen te krijgen dat zij andere instellingen dan "hoge prestaties" moeten vermijden.

Consistent gebruik van de gebalanceerde of energiebesparingsmodus zou sterk bijdragen aan de doelen van LEAP (d.w.z. verbetering van de energie-efficiëntie van ICT in datacenters). Aangezien geen van de deelnemende partijen de instelling "energiebesparing" en slechts een kleine minderheid een gebalanceerde instelling gebruikte voor zijn servers, kunnen we aannemen dat deze instellingen in het algemeen weinig worden gebruikt. Het gebruik van de gebalanceerde instelling leverde energiebesparingen van bijna 10% op bij twee groepen sterk belaste servers. Er kan redelijkerwijs worden aangenomen dat deze 10% een goede schatting is voor de energiebesparingen voor alle servers wereldwijd.

Naar verwachting kunnen er grote energiebesparingen worden gerealiseerd door consolidatie van virtuele omgevingen. Door naar hogere niveaus te streven kan zowel meer energie als meer geld worden bespaard dan de LEAP-coalitie momenteel voor ogen heeft. Zoals aangetoond door één deelnemer, die tot wel 90 virtuele machines per fysiek knooppunt gebruikt, met een gemiddelde van 60 vm's per knooppunt voor het hele cluster, kunnen er in een stabiele productieomgeving zeer hoge niveaus van consolidatie van de werkbelasting/server worden gerealiseerd. Als er soortgelijke niveaus

van gevirtualiseerde consolidatie worden toegepast, zouden er voor de werkbelasting van de meeste andere deelnemers veel minder servers nodig zijn.

Uit de verzamelde gegevens blijkt dat er verder onderzoek nodig is. Dat er nog steeds zorgen zijn over de gevolgen voor de prestaties, betekent dat er meer inzicht in en onderzoek naar powermanagement nodig is, ook met betrekking tot het effect op de prestaties van de applicaties. Ook is er een uitgebreidere statistische analyse van het energieverbruik van powermanagementfuncties en de gemiddelde CPU-belasting nodig om goed onderbouwde conclusies te kunnen trekken over het algemene gebruik van powermanagementfuncties en de potentiële energiebesparing. Verrassend genoeg heeft slechts de helft van de organisaties die dit hadden toegezegd, ook daadwerkelijk gegevens aangeleverd, en van de partijen die gegevens hebben aangeleverd, heeft de helft de powermanagement-instellingen niet gewijzigd. Wanneer zij hiernaar werden gevraagd, bleek dat veel organisaties niet volledig aan de pilot hebben deelgenomen vanwege gebrek aan kennis en angst voor de gevolgen, die op zichzelf weer voortkomt uit een gebrek aan kennis.

Er is dringend behoefte aan duidelijke en bij voorkeur eenduidige begeleiding en instructies van software- en hardwareleveranciers ten aanzien van de beste manier om powermanagement-instellingen toe te passen. Daarbij moeten de mogelijke besparingen worden benadrukt en moet worden uitgelegd wanneer het standaard powermanagement strakker of juist minder strak kan worden gemaakt.

Belangrijk is dat het gebruik van "powermanagement" en "virtualisatie" maatregelen vormen in het kader van de "Informatieplicht" voor datacenters die deel uitmaken van het "Activiteitenbesluit". Zoals vermeld bij de waarnemingen gebruiken de meeste partijen een dynamische energie-instelling. Een dergelijke instelling kan worden gezien als een vorm van powermanagement.

5.3 AANBEVELINGEN

Track 1 van LEAP ‘ heeft enorm veel informatie opgeleverd, zowel over de bruikbaarheid van powermanagement als over de menselijke factor die het gebruik van deze functie bepaalt.

Uit de combinatie van de ruwe gegevens en de gesprekken met de deelnemende partijen komen enkele aanbevelingen voor volgende stappen voort:

1) Begeleiding door leveranciers van hardware en besturingssystemen.
2) Meer technisch onderzoek naar de prestaties van applicaties bij verschillende energie-instellingen.
3) Meer statistisch onderzoek naar het huidige gebruik van powermanagement en belastingspercentages.
4) Openheid van (grote) datacenters over het daadwerkelijke energieverbruik van de faciliteit in de loop van de tijd.

De behoefte aan begeleiding door leveranciers is bij de conclusies besproken. Ter ondersteuning daarvan moet de behoefte aan technisch onderzoek in kaart worden gebracht. Wanneer er gegevens beschikbaar worden over het gedrag van veelvoorkomende applicaties bij verschillende energie-instellingen in belaste toestand, is de kans groot dat de ongegronde angst voor een merkbare verslechtering van de prestaties van de applicatie zal verdwijnen.

Punt 3 en 4 houden ook verband met elkaar. Er is momenteel te veel onzekerheid over het energieverbruik van datacenters. De informatie is wel beschikbaar, maar wordt niet gedeeld, waardoor

er geen conclusies worden getrokken op basis van het totale energieverbruik en informatie over tijdgebonden energieverbruik. Zonder deze informatie en een gedetailleerdere analyse van het daadwerkelijke gebruik en de configuratie van ICT-apparatuur kunnen er geen nauwkeurige schattingen worden gemaakt van het werkelijke potentieel van powermanagement en virtualisatie dat met de bestaande middelen niet wordt benut.

Het zou nuttig zijn om een statistisch relevant onderzoek te verrichten naar de daadwerkelijke powermanagement-instellingen en virtualisatie die worden gebruikt bij servers in Nederland. Vanuit een dergelijk onderzoek kunnen gerichtere stappen worden bepaald om de potentiële energiebesparingen ten volle te benutten.