



Serverprestaties in balans met de energieconsumptie

# CIO's

## plukken vruchten van innovaties

Naarmate onze datacenters groter worden en de eisen van webscale- en cloudproviders de markt overnemen, worden servers meer en meer slechts een component van een grotere machine. Hoewel componenten waardevol kunnen zijn, zijn zij niet langer het hele systeem en als zodanig kan hun waarde niet los gezien worden van het datacenter waarin zij gehuisvest zijn. De efficiëntie van het component wordt daarmee belangrijker dan het bezitten van 'coole features'. Wat zijn de actuele ontwikkelingen die moeten zorgen voor een lager energieverbruik van servers?

Vierentwintig uur per dag, het hele jaar door, zijn servers bezig steeds grotere hoeveelheden data te verwerken en uit te leveren via een wereldwijd netwerk van miljoenen devices die deze data consumeren en opslaan voor het nageslacht. Een verscheidenheid aan industrieën heeft fors geïnvesteerd in de benodigde middelen om het bereik, de rijkdom en de snelheid van deze digitale informatiestroom te verbeteren, maar de snelle groei van het energieverbruik door deze weelde aan diensten roept steeds meer vragen op. De uitdaging is om de computing-prestaties te verhogen zonder dat daarbij het energieverbruik explosief groeit. De operationele productiviteit van computersystemen moet kortom verbeteren terwijl de energieprestaties van IT-apparatuur omhoog gaat. Hoewel dit nog niet als een noodzaak wordt onderkend door iedere CIO, lijdt het weinig twijfel dat organisaties die met een duurzame datacenter-operatie een gedegen strategie voor energie-efficiëntie omarmen, toekomstige risico's voor hun bedrijfsvoering zullen minimaliseren.

### Ontwikkeling energie verbruik

Deze focus op energie-efficiëntie heeft

al een duidelijke impact gehad op de ontwikkeling van de IT-hardware, zoals recent onderzoek het datacenter-eindegebruikersplatform Datacenter Pulse ook aangeeft. In dit onderzoek werd gekeken naar de afname van het energieverbruik van servers over de afgelopen tien jaar. Hierbij werden drie typen serversystemen onder de loep genomen:

- 1U rackmount servers (of zo geheten 'pizzadoos-servers');
- bladeservers en
- density-optimized servers, die door marktonderzoeksbureau IDC worden omschreven als 'minimalistische servers', die qua 'form factor' lijken op bladeservers maar waarin alle overbodige componenten zijn weggelaten. Deze density-optimized servers bieden doorgaans vier server-nodes in een 2U-rackchassis – en soms een dozijn nodes in een 4U-chassis – en beschikken over processoren, geheugen, enkele schijven, een paar netwerkpoorten en verder niets.

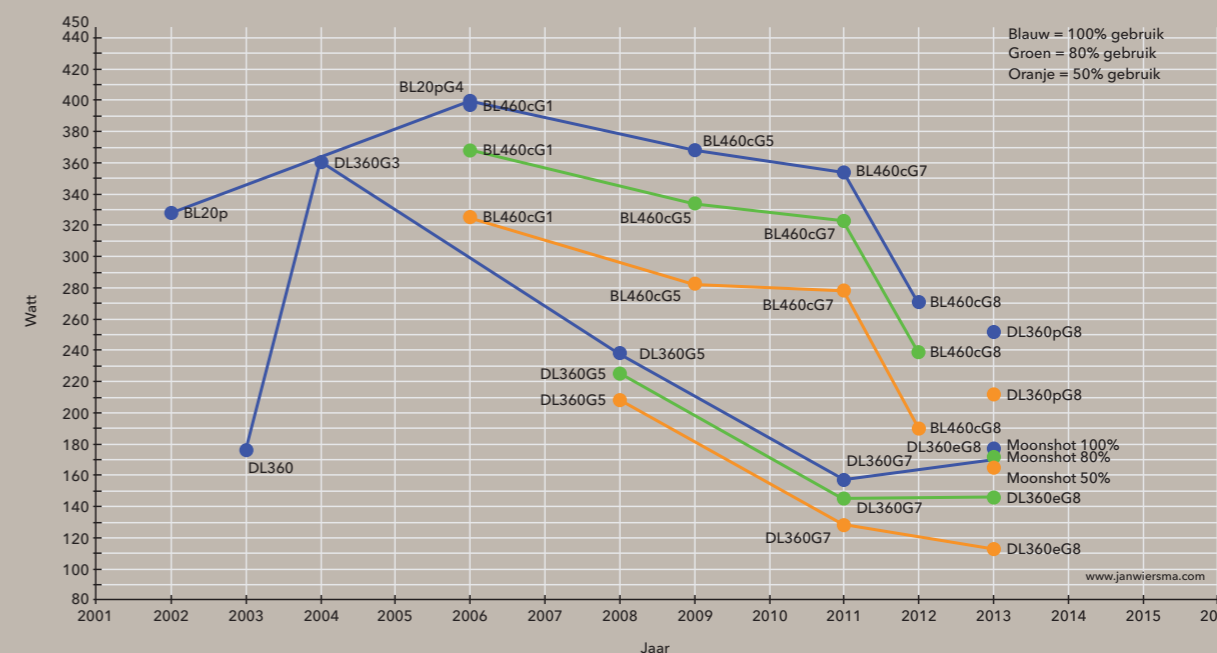
Deze drie typen systemen, in dit voorbeeld HP-systemen, werden tijdens het onderzoek van Datacenter Pulse uiteen gezet in een grafiek die de ontwikkeling van het energieverbruik per 1 hoogte-eenheid (IU) in een rack weergeeft (zie figuur 1). De grafiek

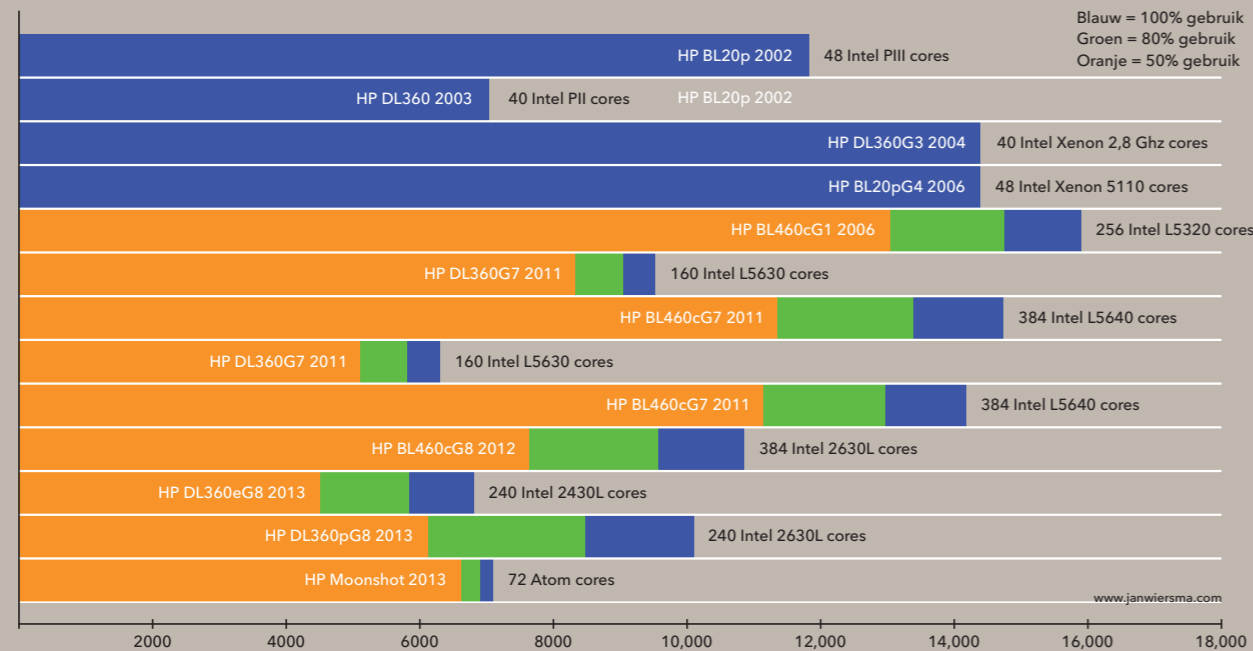
toont rond 2004-2006 een piek in het energieverbruik per IU; de focus lag in deze periode duidelijk nog niet op energie-efficiëntie. Daarna ging het energieverbruik per IU omlaag. Deze afname is voornamelijk toe te kennen aan power-optimized CPU's van Intel en AMD en geheugenmodules. Ook de introductie van Solid State Disks (SSD's) leverde een zeer grote bijdrage aan de energiereductie. Deze daling is uiteraard goed nieuws, maar daarbij is er nu weer een lichte stijging te zien.

### Meer cores in een rack

De populariteit van power-optimized CPU's is ook te verklaren door de steeds grotere honger naar consolidatie van IT-systemen in het rack en daarbij betere uitnutting van de beschikbare datacenter ruimte. Zowel IDC als Gartner ziet groei in de markt in de segmenten van bladeservers (nu 16,3 procent van de totale markt) en de density-optimized servers (66,4 procent omzetgroei, jaar over jaar). "Beide soorten modulaire form factors presteren beter dan de totale servermarkt, wat aangeeft dat klanten specialisatie in het serverontwerp steeds meer op prijs stellen", aldus Jed Scaramella, bij IDC research manager voor Enterprise Servers. "IDC neemt een toenemende interesse waar voor

Figuur 1. Ontwikkeling energieverbruik per IU in een rack





Figuur 2. Stroomverbruik van een 40U-rack, gevuld met de vermelde systemen.

## GELIJKSTROOM NAAR HET RACK

Bij uitrol van serversystemen op grote schaal, wordt er ook gekeken naar de manier waarop energie getransporteerd en omgezet wordt. Daar waar deze in een traditionele serveromgeving enkele malen van voltage verlaagd wordt (10.000V – 400V – 230V) en soms enkele malen van gelijkstroom (DC) naar wisselstroom (AC) wordt omgezet, is er winst te behalen door complete serverracks direct te voorzien van gelijkstroom.

geconvergeerde systemen die blade-servers gebruiken als bouwstenen. Enterprise IT-organisaties zien geconvergeerde systemen als een methode om het beheer te vereenvoudigen en sneller waarde te kunnen toevoegen." Kijkend naar deze ontwikkelingen, dan ziet men op een 40U rack-vulling al snel 240 x86 CPU kernen (cores) voor rackmount servers, 384 x86 cores voor bladeservers en zelfs 720 Atom-cores voor Density Optimized servers (zie figuur 2). Indrukwekkende aantallen die gepaard gaan met even indrukwekkende energie-afnamen van minimaal 6 tot 10kW per rack.

## Hogere temperaturen

Hoe klein de form factor ook wordt, CPU's blijven een grote bron van energieverbruik en daarmee een grote warmtebron binnen servers. Mainboard, geheugenmodules (RAM), hard-disk en systeemventilatoren volgen op de voet. Hier valt winst te boeken door de toevoertemperatuur te verhogen. Om alle componenten te voorzien van voldoende koeling, schrijven de meeste leveranciers in hun garantievoorwaarden een toevoertemperatuur voor die ligt tussen 10 en 35 °C. De aanname is dat bij hogere temperaturen er meer uitval van componenten optreedt. Inmiddels hebben diverse testen van bijvoorbeeld Intel en Microsoft in hun eigen datacenters laten zien dat de meeste servers hogere toevoertemperaturen gemakkelijk aan kunnen zonder extra uitval te creëren. Dell is hier op ingesprongen met het Fresh-Air-initiatief, waarbij de toevoertemperatuur met behoud van garantie kan worden ingesteld op 45 °C bij een relatieve luchtvochtigheid van 90 procent. Dit betekent dat men in het Nederlandse klimaat helemaal niet meer hoeft te koelen, aangezien daar nooit een buitentemperatuur van 45 °C gehaald wordt.

Ook andere serverleveranciers zoals SuperMicro leveren nu dit soort serversystemen. Dit alles overigens zonder overmatige energieconsumptie van de serverventilatoren, die door betere sensoren en intelligente BIOS-software ook al lang niet meer excessief hard draaien. Mocht al deze luchtkoeling in de toekomst dan toch tekortschieten, dan blijft vloeistofkoeling een redelijk alternatief. Ondanks het feit dat vloeistofkoeling natuurkundig gezien het meest efficiënt is, blijft dit tot op heden een niche. Dit heeft met name te maken met de praktische toepasbaarheid ervan, bijvoorbeeld bij het verwisselen van defecte onderdelen.

## Voortstuwende initiatieven

De genoemde innovaties zijn op het moment vooral geschikt voor de grote cloudproviders die honderdduizenden servers onder hun hoede hebben. Veel van de eerder genoemde innovaties worden dan ook gedreven door deze bedrijven, die verantwoordelijk zijn voor de geprognosticeerde aankoop van 1,5 miljoen serversystemen in 2013. Hun eigen innovaties en marktfragen leiden op het moment tot een explosie aan nieuwe hardwareoplossingen van diverse

bestaande en nieuwe leveranciers. Een van de initiatieven die energie-efficiëntie in serverontwikkeling voortstuwt, is het project Open Compute van Facebook. Bij de groei van de Facebook-omgeving liep men al snel tegen schaalbaarheidsproblemen aan. Hierbij werd duidelijk dat met de wijzigende functionaliteit in hun applicatielandschap de hardware niet altijd even efficiënt werd benut. Het bleek erg lastig om met de hedendaagse servers dit gat tussen hardware en software goed te dichten. Ook liep men tegen een stijgende energieconsumptie aan die verhoudingsgewijs een steeds groter deel van de OPEX ging uitmaken. Om beide problemen te adresseren werden enkele nieuwe hardware-oplossingen ontwikkeld, die vervolgens onder de open-sourcegedachte gedeeld werden via het Open Compute-project. Dit gaf vervolgens hardwareleveranciers over de hele wereld de mogelijkheid om hun deel van de engineering op te pakken volgens deze Open Compute-serverconcepten. Na één jaar had Facebook vervolgens een grotere keuze uit verschillende leveranciers, die beter voldeden aan hun wensen. Op het vlak van energie-efficiëntie zijn deze servers ook geschikt voor een bredere temperatuurrange (-5 °C tot +45 °C) waardoor men minder hoeft te koelen. Door optimale positionering van temperatuursensoren is het samenspel met de serverventilatoren geoptimaliseerd zodat deze bij hogere temperaturen niet onnodig hard draaien en daarmee minder energie afnemen. Ook is er door het verwijderen van behuizings-elementen de weerstand binnen de server verlaagd, wat een positief effect heeft op de energie afnamen van de ventilatoren. De door Open Compute en Facebook ontwikkelde serverconcepten sluiten aan bij een nieuw form-factorsegment in de serverindustrie, die door IDC als de eerder genoemde Density Optimized is bestempeld. Hiermee is een belangrijke impuls gegeven aan een nieuw marktsegment dat beter aansluit bij de klantvraag.

## Hogere consumptie

Nu kan men uiteraard denken dat al deze hardware-ontwikkelingen mooi zijn voor mega-infrastructuren (web-scale) en grote cloudproviders, maar niet van toepassing zijn op de IT-omge-

ving van de gemiddelde CIO. Echter met de trend van het bouwen van eigen private clouds, kan deze hardware-ontwikkeling wel eens heel snel interessant worden. Als men er vanuit gaat dat een private cloud alle voordelen gaat opleveren die men ervan verwacht, zoals sneller uitrol en infrastructuur tegen lagere kosten, dan zou een succesvolle uitrol wel eens kunnen leiden tot meer consumptie binnen de organisatie. Cloudomgevingen zijn bij uitstek gebouwd om dit succes op te vangen, schijnbaar onbeperkt te kunnen opschalen en daarmee de organisatie te helpen om nieuwe innovatieve

## CIO's kunnen energie-efficiëntie als onderdeel van hun serverselectieproces niet langer kunnen negeren.

oplossingen uit te rollen en de business beter te ondersteunen. De benodigde infrastructuur moet echter wel geïnvest worden en voorzien van energie; dit alles tegen de juiste kostenbalans. Dat is exact de uitdaging waar vele grote cloudproviders voor stonden en waar CIO's vandaag de dag de vruchten van kunnen plukken.

## Hulpmiddelen voor inzicht

Voortbouwend op de aanzienlijke vooruitgang in de IT hardware-sector tot op heden, zijn er extra hulpmiddelen nodig om tot een uniforme beoordeling en verbetering van de efficiëntie van IT-apparatuur te komen. Een dergelijk hulpmiddel is een universele maatstaf (of metriek) voor serverefficiëntie die van toepassing is op de meerderheid van de servermarkt. Zo'n metriek voorziet eindgebruikers van een inzicht in de energieprestatie van het systeem en zorgt voor een opstap naar slimmere inkoop van energie-efficiënte servers.

Een voorbeeld is de SPECpower metriek en met name de SERT; SPEC's Server Efficiency Rating Tool. Dit soort metrieke en tools geven gedetailleerd inzicht in de verhouding tussen performance en energie-efficiëntie van een

## SMARTPHONE-CPU'S IN SERVERS

Een markante ontwikkeling van de afgelopen jaren, die moet zorgen dat de serverprestatie beter in balans blijft met de energieconsumptie, is de inzet van CPU's die ontwikkeld waren voor smartphones en tablets. Hierbij werden servers ontwikkeld die gebruikmaken van de ARM CPU of de Atom CPU. Deze laatste wordt door Intel ontwikkeld en is bijvoorbeeld terug te vinden in de recent door HP gelanceerde Moonshot-serverlijn.

AMD is de samenwerking aangegaan met ARM en ontwikkelt een CPU onder de codenaam "Seattle" die in 2014 uit moet komen voor de servermarkt. De rekenkracht van dergelijke processoren zijn uiteraard beperkt ten opzichte van hun grotere broers zoals de Intel Xeon, maar uitstekend inzetbaar voor cloud scale-out omgevingen en het draaien van webservers.

systeem. De SERT, die door de U.S. Environmental Protection Agency (EPA) gebruikt wordt in hun 'Energy Star for Servers'-programma, test diverse workloads op verschillende servercomponenten en geeft inzicht in het energieverbruik op deze onderdelen. Dit geeft eindgebruikers de mogelijkheid om de meest efficiënte server te kiezen bij hun workload-profiel. Dergelijke metrieke kunnen CIO's helpen bij het opstellen van hun RFP voor serveraanschaf en daarmee enige orde in een steeds complexere markt scheppen. CIO's kunnen energie-efficiëntie als onderdeel van hun serverselectieproces niet langer kunnen negeren. Het goede nieuws is dat de grote web-scale- en cloudleveranciers zoals Google, Microsoft, Facebook en Amazon de innovatiedans op servergebied leiden met maximale focus op efficiëntie. Hierbij kan geconcludeerd worden dat servers niet langer de hoeksteen zijn van een IT-strategie.

Onderzoek: [http://datacenterpulse.org/blogs/jan.wiersma/where\\_rack\\_density\\_trend\\_going](http://datacenterpulse.org/blogs/jan.wiersma/where_rack_density_trend_going)

Jan Wiersma, interim-CTO in het dagelijks leven en International Director EMEA voor non-profit platform Datacenter Pulse.