

Wat is Cloud computing?

L. Ferreira Pires¹

Dit artikel introduceert *cloud computing*, bespreekt de relatie tussen cloudcomputing en *service-oriented architecture*, identificeert verschillende vormen en eigenschappen van cloudcomputing, en bespreekt de voor- en nadelen van deze technologie.

Cloudcomputing is momenteel een zeer populair onderwerp. Een zoekopdracht met de woorden ‘cloudcomputing’ levert meer dan 20 miljard hits, en dit aantal wordt alleen maar groter als wij zoeken naar andere verwante termen zoals ‘software-as-a-service’. Veel mensen praten momenteel over cloudcomputing dus het is belangrijk om stil te staan bij deze technologie en wat zij kan betekenen voor individuen en bedrijven.

1. Oorsprong en definitie

De term cloudcomputing heeft een merkwaardige oorsprong. Voor mensen die veel betekenis achter deze term zoeken, is het soms teleurstellend te constateren dat de *cloud* correspondeert met het symbool (icoon) dat vaak gebruikt wordt om een netwerk of het internet aan te duiden (een wolk, dus ‘cloud’ in het Engels). In essentie correspondeert de cloud dus met het internet, zodat de term cloudcomputing geïnterpreteerd kan worden als ‘computergebruik dat ergens op het internet plaatsvindt’. Deze interpretatie is natuurlijk te vaag voor een definitie, en om meer duidelijkheid te scheppen definieert het Amerikaanse National Institute of Standards and Technology (NIST) cloudcomputing als volgt:²

„A model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.“

Volgens deze definitie, gaat cloudcomputing verder dan het vage begrip van ‘computergebruik ergens in een netwerk’. De definitie geeft ook de bedoeling aan van deze technologie, namelijk het geven van (flexibele) toegang tot computerhardware en -software aan gebruikers. Veel bedrijven zijn momenteel bezig met cloudcomputing, zowel met de infrastructuur als met de technische en organisatorische ondersteuning van systemen die *in de cloud* draaien. Doordat forse investeringen nodig zijn voor een betrouwbare en schaalbare infrastructuur, zijn er momenteel slechts een paar cloudcomputing-infrastructuurleveranciers op grote schaal actief. Voorbeelden zijn Google, Amazon en Microsoft. Andere bedrijven, zoals sourceforge.com, bieden cloudcomputingoplossingen binnen een applicatiedomein (in dit geval bedrijfsapplicaties, met nadruk op ondersteuning van het verkoopproces), terwijl er ook partijen zijn die zich concentreren op consultancy in cloudcomputingoplossingen. Een bedrijf als IBM, bijvoorbeeld,

¹ L. Ferreira Pires is universitair hoofddocent aan de Universiteit Twente (Faculteit Elektrotechniek, Wiskunde en Informatica).

² Peter Mell and Tim Grance, *The NIST Definition of Cloud Computing*, Version 15, 10-7-09, [http://csrc.nist.gov/groups/SNS/cloud computing/cloud def-v15.doc](http://csrc.nist.gov/groups/SNS/cloud%20computing/cloud%20def-v15.doc).

profileert zich in de hele linie als clouddeskundige, van het plannen van cloudcomputingoplossingen tot het bouwen en aanbieden van deze oplossingen aan klanten, in dit geval bedrijven.

2. Service-oriented architecture en cloudcomputing

Als cloudcomputing een oplossing is, wat was dan ook maar weer het probleem? Wij horen heel vaak dat IT (informatietechnologie) als een *showstopper* wordt gezien in grote organisaties, in plaats van wat deze technologie eigenlijk moet zijn, namelijk een katalysator voor betere bedrijfsvoering. In sommige bedrijven zijn IT-afdelingen sloom en zij vertragen de introductie van innovatie. IT-investeringen zijn aan de ene kant gedreven door mode (de nieuwste snuffjes, verzonnen door de snelle jongens), maar aan de andere kant soms zeer conservatief (denk aan het nog aantal draaiende COBOL-applicaties). Deze twee factoren dragen meestal bij aan de complexiteit en ondoorgroendelijkheid van de IT-infrastructuur. Deze problemen worden niet in een klap opgelost met cloudcomputing. Echter, met de opmars van de zogenaamde ‘service-oriented architecture’ (SOA),³ kunnen computerfuncties worden georganiseerd in termen van diensten die meer orde kunnen brengen in de IT-infrastructuur. Met cloudcomputing kunnen deze diensten worden gedetacheerd naar andere bedrijven (in deze special cloudcomputing-serviceproviders (CcSPs), cloudproviders, of cloudleveranciers genoemd), die verantwoordelijk worden voor het onderhouden van de hardware- en softwaresystemen waarop deze diensten draaien. Hierdoor kunnen bedrijven zich concentreren op hun *core business* en komen er middelen vrij voor investeringen.⁴ Dat betekent dat de combinatie SOA/cloudcomputing grote voordelen kan bieden, en dat is de reden waarom deze twee begrippen vaak samen worden beschouwd.

3. Vormen van cloudcomputing

De NIST-definitie van cloudcomputing praat over ‘computing resources’, maar wat voor bronnen bedoelen ze dan? In eerste instantie kunnen wij deze bronnen verdelen in software en hardware. Softwarebronnen zijn computerprogramma’s (applicaties) die ergens in het netwerk draaien. Typische voorbeelden zijn webapplicaties, zoals Gmail, Google Docs, en de CRM-applicaties aangeboden door salesforce.com. Deze vorm van cloudcomputing bestaat al lang en is algemeen bekend. Hardware resources zijn iets moeilijker te begrijpen voor mensen zonder een technische achtergrond. Hier worden zogenaamde virtuele servers aan cloudklanten (bijvoorbeeld bedrijven) aangeboden, waarop deze cloudklanten hun software kunnen draaien. Virtuele servers zijn logische systemen die in fysieke systemen (echte computers) draaien. De cloudleverancier zorgt dat meerdere virtuele servers in een fysiek systeem draaien, met de gewenste capaciteit. De cloudklant is veelal eigenaar en beheerder van zijn eigen virtuele servers.

Amazon heeft bijvoorbeeld de Amazon S3 storage resource, waarop cloudklanten data kunnen opslaan, en de Amazon EC2, waarmee cloudklanten virtuele servers kunnen creëren. In dit geval

³ Deze afkorting heeft vervelende associaties in het Nederlands, maar is een gangbare en veel gebruikte afkorting in de ICT-literatuur.

⁴ M. Naghshineh. *Cloud Computing: Incremental or Transformational?*, Keynote presentation, ECOWS 2009.

ontwikkelt de cloudklant zelf de software, maar installeert deze software op de virtuele server van de cloudleverancier (Amazon in dit voorbeeld), die dan verantwoordelijk is voor het onderhoud en draaiend houden van deze server. Deze virtuele servers draaien in de fysieke servers van Amazon met gebruik van zogenaamde virtualisatietechnologie (technologie waarmee virtuele servers gecreëerd kunnen worden). Een cloudprovider als Amazon zorgt dus dat voldoende capaciteit in hun eigen computers ontstaat voor bestaande virtuele servers en de servers die aangevraagd zullen worden.

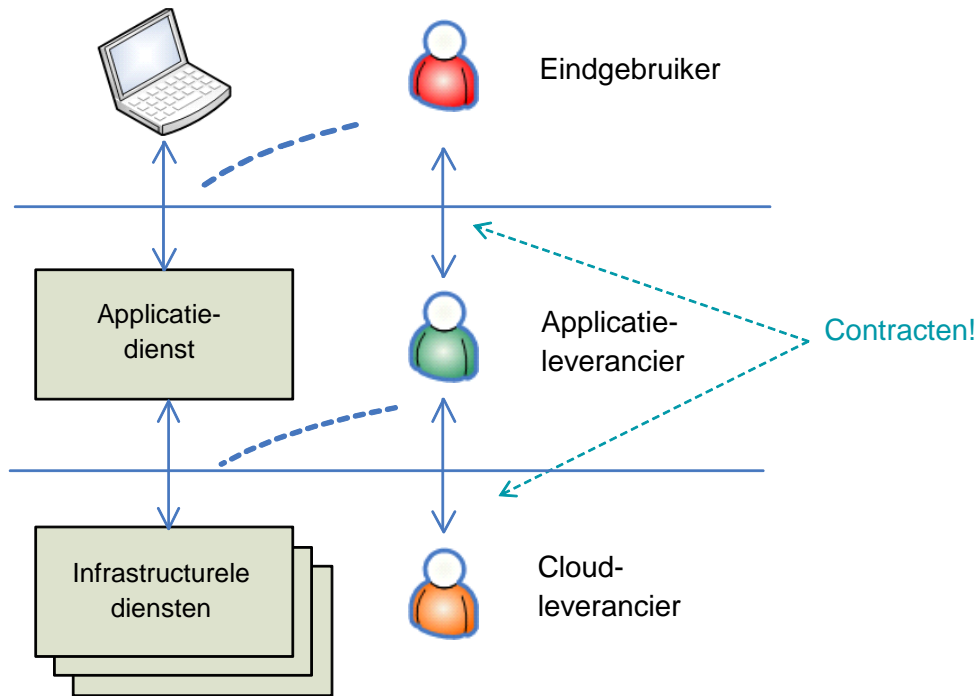
Gebaseerd op de bronnen die via clouddiensten worden aangeboden, kunnen wij verschillende vormen van cloudcomputing onderscheiden,⁵ waarvan de meest populaire zijn:

- software-as-a-service (SaaS): cloudleverancier biedt toegang tot een software pakket of applicatiedienst, vaak via een website. Typische voorbeelden zijn Gmail, Google Docs, salesforce.com;
- platform-as-a-service (PaaS): cloudleverancier geeft toegang tot een ontwikkelomgeving, waarin cloudklanten hun eigen applicatiediensten kunnen ontwikkelen en aan eindgebruikers aanbieden. Typische voorbeelden zijn Google Apps en Force.com;
- infrastructure-as-a-service (IaaS): cloudleverancier geeft toegang tot virtuele systemen, waarop cloudklanten hun eigen systemen (applicatiediensten) kunnen installeren en aan eindgebruikers aanbieden. Een typisch voorbeeld is Amazon Elastic Compute Cloud 2 (EC2).

In de eerder genoemde NIST-definitie worden deze vormen van cloudcomputing aangeduid als *servicemodellen*. In deze vormen van cloudcomputing herkennen we drie typen partijen, namelijk cloudleveranciers, applicatieleveranciers en eindgebruikers. De rechten en plichten tussen deze partijen moeten worden vastgesteld via contracten zoals de zogenaamde *service-level-agreements* (SLAs) die in de cloudapplicatie, het ontwikkelplatform of de infrastructuur nageleefd moeten worden.

Vormen van cloudcomputing kunnen ook onderscheiden worden op basis van hun toegankelijkheid: (i) een *publieke cloud* is beschikbaar en toegankelijk voor alle gebruikers en bedrijven. Betaling in deze vorm van cloudcomputing geschiedt *per gebruik*, d.w.z. gebruikers betalen alleen voor de resources die ze werkelijk hebben gebruikt; (ii) een *privécloud* wordt beheerd en gebruikt door een enkel bedrijf; (iii) een *gemeenschappelijke cloud* wordt beschikbaar gemaakt voor één of meerdere organisaties, en ondersteunt een gemeenschap met gezamenlijke belangen, en (iv) een *hybride cloud* bestaat uit een combinatie van de drie bovengenoemde vormen. In de NIST-definitie worden deze vormen van cloudcomputing als *ontplooiings* ('deployment') *modellen* aangeduid.

⁵ Zie D.S. Linthicum, *Cloud Computing and SOA Convergence in Your Enterprise: A Step-by-Step Guide*, p. 38.



4. Eigenschappen en technologieën

Om cloudcomputingtechnologieën nog beter te karakteriseren, heeft men in de literatuur de volgende eigenschappen aan deze technologieën toegekend:

- diensten op verzoek ('on demand'): computerbronnen worden geleverd op het moment dat de gebruiker deze bronnen vraagt;
- netwerktoegang ('ubiquitous network access'): computerbronnen zijn overal in het netwerk toegankelijk;
- bronnen gedeeld onafhankelijk van locatie ('multitenant model'): meerdere gebruikers delen de bronnen onafhankelijk van waar ze zich ook in het netwerk bevinden;
- snelle elasticiteit: snelle toename of verlaging van bronnen (of hun capaciteit) afhankelijk van de vraag van de gebruikers;
- betalen per gebruik ('pay-per-use'): gebruikers betalen alleen voor de bronnen die ze werkelijk hebben gebruikt.

Cloudcomputing wordt mogelijk gemaakt door technologieën als (i) virtualisatie op verschillende niveaus (processoren en geheugen, maar ook servers en applicaties), (ii) geavanceerde beheersystemen voor deze virtuele elementen, vaak in de vorm van webapplicaties, (iii) takenautomatisering gebaseerd op beleid en rollen, en (iv) monitoring infrastructures.

Vanuit een puur technologisch standpunt is cloudcomputing niet echt nieuw. Het is gebaseerd op het oude principe van *time sharing* (gezamenlijk gebruik van middelen door meerdere partijen op dezelfde tijd), dat veelvuldig is toegepast in de begintijd van de computer toen verschillende eindgebruikers gebruik moesten maken van grote, dure, schaarse mainframes. In cloudcomputingsystemen worden vaak virtualisatietechnieken gebruikt die al lang in omloop zijn. Hoewel de onderliggende technologieën al lang bestonden, hebben de recente groei in snelheid en capaciteit van communicatiesystemen en computerhardware cloudcomputing

bruikbaar en rendabel gemaakt. In het bijzonder ontwikkelingen in netwerktechnologie hebben een belangrijk aandeel gehad in de intrede van cloudcomputing, doordat computernetwerken wereldwijd verspreid, sneller en goedkoper werden.

Cloudcomputing wordt vaak geassocieerd met de term *commodity computing*, een vorm van computergebruik die vergelijkbaar is met (het consumptiegoed) elektriciteit. Om deze vergelijking verder te trekken, in het begin van de industriële revolutie had elke fabriek zelf een generator om elektriciteit op te wekken, maar later werd duidelijk dat het goedkoper was om een elektriciteitsnetwerk te bouwen, waarvan niet alleen bedrijven maar ook individuen konden profiteren. De term *utility computing* wordt ook vaak gebruikt in de context van cloudcomputing, waarin computergebruik wordt vergeleken met telefoon, elektriciteit, water of gas, en ook 'uit de muur' komt. Ook dit idee is niet nieuw. Het stamt uit de jaren 60,⁶ maar de computer- en communicatietechnologie die nodig is om dit te realiseren is alleen sinds kort aanwezig.

5. Voor- en nadelen

Cloudcomputing heeft een aantal voordelen. Bedrijven kunnen de last van het kopen, installeren, onderhouden en beheren van IT-systemen vermijden door hun systemen *in de cloud* te laten draaien. Het gebruik van energie en ruimte is efficiënter als het computergebruik wordt toegekend en aangeleverd door een cloudleverancier. Het laatste wordt onder meer duidelijk uit het feit dat gemiddeld maar 10 tot 15% van de ICT-middelen door een bedrijf wordt gebruikt,⁷ maar dat meer ingekocht moet worden voor pieksituaties. En daar kan nu de cloudprovider in voorzien. Daarom is cloudcomputing ook vaak gezien als een belangrijk element van *greencomputing*, als een oplossing voor de toename van energiegebruik veroorzaakt door de wereldwijde groei in het gebruik van ICT-systemen. Verder maakt cloudcomputing het voor bedrijven mogelijk om de vraag naar computerbronnen (servers, geheugen, enzovoorts) naar boven of naar beneden aan te passen, afhankelijk van de vraag van eindgebruikers. Bijvoorbeeld, een webwinkel kan de capaciteit van zijn webapplicaties laten afhangen van het verwachte aantal bezoekers, naar boven tijdens de kerstperiode en tijdens grote acties, en naar beneden tijdens de vakantieperiode. Bedrijven kunnen hun klanten sneller bedienen via applicaties die in een cloudcomputingomgeving draaien, doordat de installatie van software op virtuele systemen kan gebeuren zonder dat nieuwe hardware gekocht en geconfigureerd moet worden, wat soms een paar maanden in beslag neemt. Bovendien kunnen alle beheeracties, zoals veranderingen in servers en software versies, veranderingen in proces- en geheugencapaciteit, enzovoorts, op elke moment van de dag, zeven dagen per week, gebeuren.

Een aansprekend voorbeeld van de voordelen van cloudcomputing is gerapporteerd in een blog in 2007.⁸ Dit voorbeeld gaat over de Amerikaanse krant *The New York Times* (NYT), die op zoek was naar een oplossing voor het beschikbaar stellen van alle artikelen tussen 1851-1922 in leesbaar (PDF-) formaat via het internet. Deze artikelen waren al gescand en bestonden uit stukken plaatjes die nog aan elkaar geplakt moest worden. Ze zochten een oplossing waarmee alle leesbare documenten van tevoren zouden kunnen worden aangemaakt, zodat de lezer snel het document kon krijgen. Het zelf inrichten van een computercentrum om dat te doen was geen

⁶ Bijvoorbeeld, toespraak van John McCarthy tijdens de 'MIT Centennial' in 1961.

⁷ Uitspraak van A. Selipsky, Amazon VP.

⁸ 'Self-Service, Prorated Supercomputing Fun', *New York Times* 1 november 2007, <http://open.blogs.nytimes.com/2007/11/01/self-service-prorated-super-computing-fun/>.

haalbare oplossing vanwege de te hoge kosten, en deze klus uitbesteden was ook uitgesloten om dezelfde reden. Sommige mensen binnen de *NYT* hebben toen besloten om cloudcomputing als een experiment te gebruiken om dit probleem op te lossen. Uiteindelijk waren ze in staat om 11 miljard artikelen binnen 24 uur te verwerken⁹ met gebruik van 100 virtuele machines in Amazon EC2 voor de verwerking van de plaatjes, en met gebruik van Amazon S3 voor data-opslag. En de hele operatie heeft slechts 890 dollar gekost,¹⁰ wat niet eens genoeg is om één fatsoenlijke desktopcomputer aan te schaffen.

Naast deze potentiële voordelen, kent cloudcomputing ook een aantal potentiële nadelen. Veiligheid blijft een punt van zorg, want met het uitbesteden van de verwerking wordt er in principe ook de mogelijkheid aan de cloudleverancier gegeven om de informatie in te zien en te manipuleren. Er zijn technieken om dat te voorkomen, maar uiteindelijk moet de cloudklant de cloudleverancier vertrouwen. Terug naar de vergelijking met elektriciteit, water en gas van hierboven; wij zijn gewend om de nutsbedrijven te vertrouwen, maar zover zijn we nog niet met cloudcomputing.

De fysieke locatie van de data is ook een mogelijk probleem. Bijvoorbeeld, voor prestatieoptimalisatie is het mogelijk dat een provider van een SaaS-applicatie de informatie van haar klanten kopieert naar een server in de buurt van waar deze klanten zich bevinden. Dit betekent dat kopieën van de informatie overal in het netwerk kunnen bestaan. Deze data moeten worden verwijderd om ruimte te maken voor nieuwe data, en het is mogelijk dat dit na een onbepaalde termijn gebeurt door een proces dat van tijd tot tijd ongebruikte data verwijderd. Deze oplossing is relatief goedkoop, maar heeft als consequentie dat informatie kan worden verspreid binnen de cloudinfrastructuur zonder dat de klant een enkele vorm van controle over deze verspreiding heeft. In dit geval, als gebruikers hun informatie niet op bepaalde plekken willen hebben, bijvoorbeeld vanwege wetgeving, wordt de cloudleverancier genoodzaakt om hun technische infrastructuur aan te passen. In de Amazon S3 dataopslag diensten,¹¹ bijvoorbeeld, is de locatie van de datacenters bekend want Amazon biedt snelle toegang tot de informatie en de gebruiker kan bepalen in welk datacenter (in welke regio) de data worden opgeslagen. Google, aan de andere kant, ondersteunt onder andere wereldreizigers met diensten zoals Gmail, zodat het moeilijk kan worden om te bepalen waar informatie over deze personen in hun netwerk te vinden is. De bepaling van de fysieke locatie van data kan dus afhankelijk zijn van de vorm van cloudcomputing en de toepassing.

Er zijn voldoende technische oplossingen voor problemen met data-locatie en -veiligheid, maar de nodige aanpassingen in de infrastructuur brengen extra kosten met zich mee, waardoor de diensten voor de klanten ook duurder kunnen worden. In het algemeen geldt dat hoe hoger de kwaliteitseisen (veiliger, sneller, betrouwbaarder, enzovoorts) hoe duurder deze diensten worden. Eisen aan de diensten die verder gaan dan hun basisfunctionaliteit worden vaak met de term 'quality of service' (QoS) aangeduid.¹²

⁹ Eigenlijk twee keer, want ze vond de eerste keer een fout in het programma.

¹⁰ D. Hilley. *Cloud Computing: A Taxonomy of Platform and Infrastructure-level Offerings*. College of Computing, Georgia Institute of Technology, April 2009.

¹¹ Amazon Simple Storage Service (Amazon S3), <http://aws.amazon.com/s3/>.

¹² De term 'quality-of-service' is afkomstig uit de telecomindustrie, en werd vaak gebruikt voor het aanduiden van kwaliteitsaspecten van communicatiesystemen, zoals vertraging en betrouwbaarheid, maar ook kosten en veiligheid. Met de opmars van SOA wordt deze term steeds vaker gebruikt voor alle soorten diensten, en dus ook voor cloudcomputingdiensten.

Als wij Google opnieuw als voorbeeld nemen, eindgebruikers van diensten als Picasa, Google Docs en Blogger kunnen extra opslagcapaciteit tegen betaling krijgen, maar de dienst wordt nog steeds ‘as is’ aangeboden.¹³ Google committeert zich om deze diensten op de best mogelijke manier te leveren, maar is niet aansprakelijk voor de schade die een gebruiker kan oplopen door deze diensten te gebruiken. Als Google ooit beslist om deze diensten niet te leveren (tijdelijk of definitief) kunnen gebruikers geen aanspraak maken op vergoedingen, en lopen ze ook het risico dat ze hun documenten kwijt raken.

Hoewel kostenbesparing vaak als de belangrijkste reden wordt gegeven om cloudcomputing te gebruiken, wordt de cloudgebruiker geadviseerd om de kosten te controleren voor zijn specifieke situatie voordat een keuze voor cloudcomputing wordt gemaakt. Vooral de kosten van *hosting* en dataverkeer kunnen oplopen, waardoor een cloudcomputingoplossing duurder kan worden dan een in-house implementatie. Een ander nadeel van de huidige cloudcomputingoplossingen is dat op het moment dat een keuze wordt gemaakt voor een cloudleverancier het zeer moeilijk of kostbaar is om naar een andere cloudleverancier te verhuizen, het zogenaamde ‘lock-in’ effect.¹⁴

6. Conclusies

Cloudcomputing heeft grote mogelijkheden, maar een aantal uitdagingen belemmert nog het vertrouwen in deze systemen. Verschillende cloudomgevingen kunnen moeilijk met elkaar communiceren, met nadelen voor de integratie van services die via verschillende cloudleveranciers worden aangeboden. Het is ook vaak moeilijk en duur om een systeem van een cloudleverancier naar een andere te laten verhuizen. De veiligheid van sommige cloudoplossingen laat te wensen over. Er is ook behoefte aan afspraken die vanuit de sociale wereld doorlopen naar de diensten die via cloudomgevingen worden aangeboden. Deze afspraken moeten dan op eenduidige manieren worden ondersteund en nageleefd door de technische infrastructuur, via monitoring en audit. In principe is het alleen mogelijk om zeker te zijn over het naleven van een afspraak als cloudleveranciers inzicht bieden in hun bedrijfsvoering, in het bijzonder de interne organisatie van hun infrastructuur. Aan de andere kant, de interne details van de technische infrastructuur zijn bepalend voor de competitiekracht van de cloudleverancier, en worden vaak beschouwd als bedrijfsgeheim. Ergens in dit spanningsveld moet de klant voldoende handvatten vinden om toch de gewenste invloed op de cloudinfrastructuur te kunnen uitoefenen.

¹³ Zie Google Voorwaarden voor Dienstverlening, <http://www.google.com/accounts/TOS?hl=nl&loc=NL>.

¹⁴ *Lock-in* is beschreven en uitvoerig besproken in C. Shapiro & H.R. Varian, *Information Rules*, Harvard Business School Press, 1999.