

Wide Area File Services maken verdere consolidatie mogelijk

WAN op LAN-snelheid

In een poging grip te krijgen op de stijgende kosten van de gedistribueerde IT-infrastructuur, zijn veel IT-organisaties bezig met consolidatie van gedistribueerde servers en opslagsystemen. De IT-wereld gaat terug naar een gecentraliseerde IT-omgeving. Helaas houdt consolidatie vaak op bij nevenvestigingen en bijkantoren, omdat Wide Area Networks nu eenmaal de prestaties van een LAN niet kunnen evenaren. Bram Dons bespreekt het fenomeen Wide Area File Services, dat hierin verandering brengt.

Bram Dons

Van oudsher zijn voor de implementatie van de IT-infrastructuur op nevenvestigingen veel investeringen nodig. Om lokaal op de remote locatie de prestaties van het LAN te evenaren is nu nog een complete IT-infrastructuur op locatie vereist. Het hoeft geen betoog dat het onderhoud van de lokale file- en print-servers, additionele tape- of diskgebaseerde back-upsoftware en -hardware en het extra systeembeheer voor ieder IT-budget een flinke kostenpost is. De reden voor deze dure duplicaatinfrastructuur is dat remote gebruikers dezelfde toegangstijden tot en prestaties van opslagsystemen en werkingstijden van applicaties eisen als ze binnen hun eigen LAN-omgeving mogen verwachten. Tot voor kort maakten de beperkingen van het WAN het onmogelijk om remote gebruikers prestaties te bieden - bijvoorbeeld met betrekking tot de traditionele Microsoft Officeachtige bestanden die zijn gehuisvest binnen het datacenter - die vergelijkbaar zijn met het LAN.

Wide Area File Services

Vandaag de dag zijn er echter nieuwe technologieën ontwikkeld die het mogelijk maken om remote gebruikers met LAN vergelijkbare toegangstijden te bieden bij het openen of sluiten van bestanden via het WAN. Een voorbeeld zijn *Wide Area File Services* (WAFS). Er zijn enkele firma's die sinds een paar jaar de daarvoor benodigde zogenaamde WAFS-apparaten op de markt brengen. Een daarvan is Tacit Networks, met het product Tacit Shared Server Appliance (zie figuur 2). Om een beter begrip te

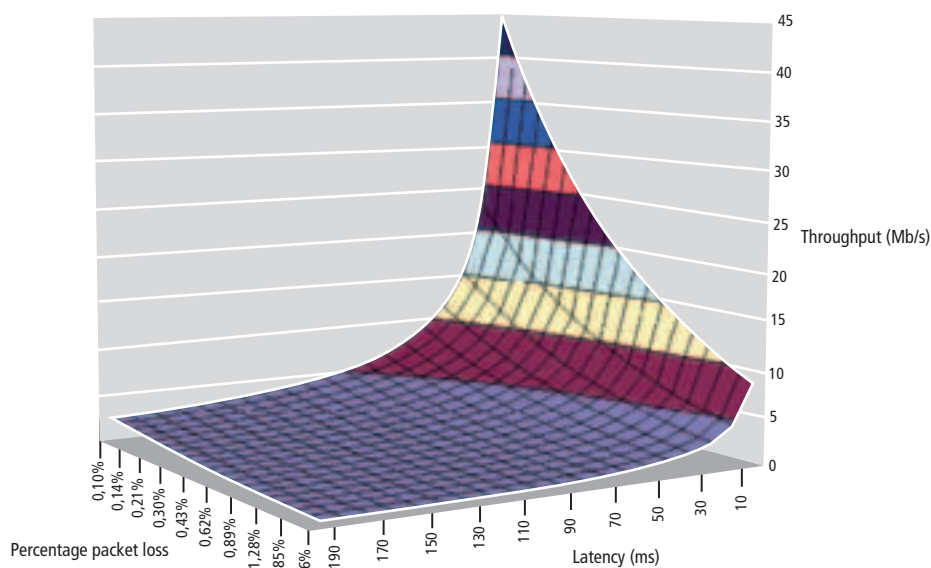
krijgen van de wijze waarop dergelijke WAFS-producten verbeteringen aanbrengen aan de prestaties van WAN's, komen eerst de voornaamste oorzaken van de huidige slecht presterende WAN's aan de orde.

WAN latency en CIFS/NFS

Bij de toepassing van geografische gedistribueerde enterprise-omgevingen die met remote locaties in een corporate netwerk zijn geïntegreerd, worden IT-beheerders geconfronteerd met talrijke nieuwe uitdagingen. De ervaring heeft bijvoorbeeld geleerd dat oorspronkelijk voor het LAN ontwikkelde technologieën binnen een WAN slecht functioneren en presteren. Tot voor kort dacht men dit op te kunnen lossen door meer bandbreedte aan de bestaande WAN-verbindingen toe te voegen, maar dit is zelden een oplossing gebleken voor het 'prestatieprobleem'.

Bij een slechte WAN-verbinding heeft het netwerkverkeer te lijden onder de zogenaamde *packet loss*. De combinatie van netwerklatency (vertraging) en packet loss doen de prestaties nog sneller afnemen: bij enkele tientallen milliseconden (ms) latency en een packet loss van tienenden van procenten zal de *throughput* (in Mb/s) heel snel afnemen (zie figuur 1).

Communicatie tussen remote locaties en datacenters zal altijd last hebben van netwerklatency. Deze kan, afhankelijk van de afstand, enkele tientallen of honderden milliseconden bedragen. De gedachte dat het delen van bandbreedte



Figuur 1 Gevolgen latency en packet loss voor netwerkthroughput

met andere gebruikers of de toepassing van een VPN tot een verbetering zou kunnen leiden, is een misvatting. Sterker nog, daarbij wordt zelfs meer latency geïntroduceerd vanwege verkeerscongestie, de extra benodigde router hops en de invoering van netwerkcarrierpolicy's. Toevoeging van *lanes* aan een netwerkverbinding (te vergelijken met extra paden of stroken op een snelweg) verbetert op geen enkele wijze de latency en dus de doorvoersnelheid van het netwerk. In sommige gevallen kan de op deze manier geïntroduceerde latency zelfs meer bedragen dan de netwerklatency die wordt veroorzaakt door de geografische afstand.

De traditionele file sharingprotocollen als CIFS en NFS zijn zeer gevoelig voor latency, omdat ze van talrijke round trips (zending en bevestiging van ontvangst ofwel een foutmelding) tussen client en server gebruikmaken en daarmee het negatieve effect van netwerklatency nog een keer versterken. Elke round trip introduceert een vertraging, doordat NFS en CIFS wachten op antwoord voordat ze een volgende zending versturen. Hoewel deze vertraging per individuele round

trip betrekkelijk klein kan zijn (lokaal twintig tot dertig ms en meer dan honderdms tussen continenten), is het aantal messages dat nodig is voor het versturen van zelfs al een klein bestand enorm. Bijvoorbeeld: het openen van een Wordbestand van 1,5 Mb via een WAN resulteert in meer dan vierduizend messages. En het opslaan van een wat groter Wordbestand van 10 Mb via een 768 Kbps WAN met 80 ms latency neemt meer dan 12 minuten in beslag, terwijl het via het LAN slechts 14 seconden duurt.

WAN latency en e-mail

Een andere applicatie die in een gedistribueerde omgeving te lijden heeft van netwerklatency is e-mail. Het messagingplatform Microsoft Exchange vormt voor de meeste ondernemingen een kritiek onderdeel van de communicatie-infrastructuur. De prestaties van Exchange in een gedistribueerde netwerk omgeving zijn echter onvoldoende; dit geldt trouwens voor alle op het MAPI-protocol gebaseerde e-mailsystemen. De inefficiënte communicatie tussen de Outlook clients en Exchange Server en de redundante overdracht van data tussen ver-

Prestaties Exchange over een WAN

De volgende factoren zijn de oorzaak van een slecht presterend Exchange in een WAN-omgeving:

- redundante aflevering van data naar gebruikers in *branch offices*;
- meerdere sends/receives van dezelfde message tussen Outlook client en server;
- een communicatieprotocol dat niet in staat is om grote data-transfers (zogenaamde *bulk transfers*) efficiënt via WAN's te versturen.

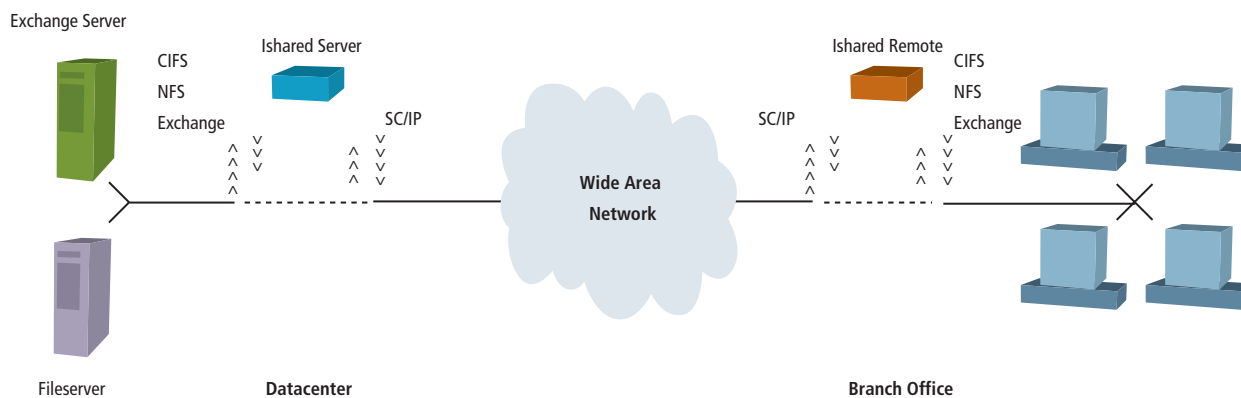
Outlook en Exchange (net zoals trouwens de andere Microsoftapplicaties) verzenden data alleen in kleine datablocks van 8 tot 16 Kb. Na verzending van ieder block moet een bevestiging worden ontvangen voordat het volgende block kan worden verstuurd. Elk van deze round trips introduceert een vertraging. De gebruiker kan dit al merken als de vertraging 30 ms of meer bedraagt.

schillende vestigingen stelt IT-beslissers voor een lastige keuze, namelijk tussen een slecht presterend e-mailsysteem op de decentrale vestigingen – ondanks een overgedimensioneerde WAN-verbinding – en de implementatie van Exchange Servers op elke vestiging. Het zal duidelijk zijn dat dit laatste tot hoge implementatie- en beheerkosten leidt.

Voor deze prestatieproblemen zijn in het verleden al verschillende oplossingen bedacht, waaronder: Outlook Web Access (OWA) en WAN optimization (bandbreedtecompressie). Helaas introduceerden deze technologieën vaak nieuwe problemen, zoals hogere kosten, verminderde functionaliteit en compa-



wide area networks



Figuur 2 Tacit WAFS-architectuur

tibiliteits- en beveiligingsproblemen. Ze boden dus geen afdoende oplossing.

Tacit WAFS-toepassing

De firma Tacit is een pionier op het gebied van file sharing via het WAN. Het bedrijf heeft een product ontwikkeld dat de gevolgen van netwerklatency bij het gebruik van MS Officeapplicaties via het WAN kan minimaliseren. Daartoe heeft Tacit het proprietary *storage caching over IP-protocol (SC/IP)* ontwikkeld, dat is gebaseerd op een aantal technologieën. Die technologieën zijn in het verleden al wel afzonderlijk toegepast om het WAN-latencyprobleem aan te pakken, maar dat loste niet echt het probleem op. Totdat Tacit de gebundelde toepassing van de volgende technologieën introduceerde: file-aware differencing, streaming en compression; temporary file optimization; data aggregation en clustered I/O.

Een verandering aan een MS Officebestand heeft normaliter tot gevolg dat het hele bestand naar het datacenter wordt teruggestuurd. Met file-aware differencing worden, zodra een bestand op een remote office applicatie van Tacit in cache is opgeslagen, alleen de daaropvolgende veranderingen aan het bestand gecompriëerd en via WAN in data streaming mode naar het datacenter teruggestuurd. Het resultaat is een aanzienlijke reductie van te ver-

zenden data en dus tijdswinst voor de gebruiker.

De talrijke door MS Office aangebrachte tijdelijke bestanden verhogen ook de toegangs- en opslagtijd van Officebestanden via het WAN. De Tacittechnologie herkent deze tijdelijke bestanden en beperkt de hoeveelheid via het WAN te versturen data aanzienlijk door deze 'intelligent' af te handelen. Tenslotte werken Officeapplicaties onafhankelijk van elkaar wanneer ze de informatie weer op het datacenter opslaan. Met behulp van data aggregation en geclusterde I/O-technologieën worden de extra round trips via het WAN voorkomen door data van de verschillende applicaties te bundelen, zelfs voordat gebruikers in nevenvestigingen hun bestandsveranderingen naar de centrale opslagplaats terugschrijven.

Het back-upprobleem

Behalve de op CIFS en NFS gebaseerde applicaties is remote back-up een ander terrein waarop de WAFS-technologie kan worden ingezet om een probleem op het gebied van continuïteit op te lossen. Veel ondernemingen hebben aandacht besteed aan de beveiliging van het datacenter, maar er worden ook grote hoeveelheden kritische data gecreëerd en opgeslagen buiten het datacenter, op de remote kantoorlocaties. Ondernemingen proberen de risico's hiervan te onder- vangen door replicatie van remote data

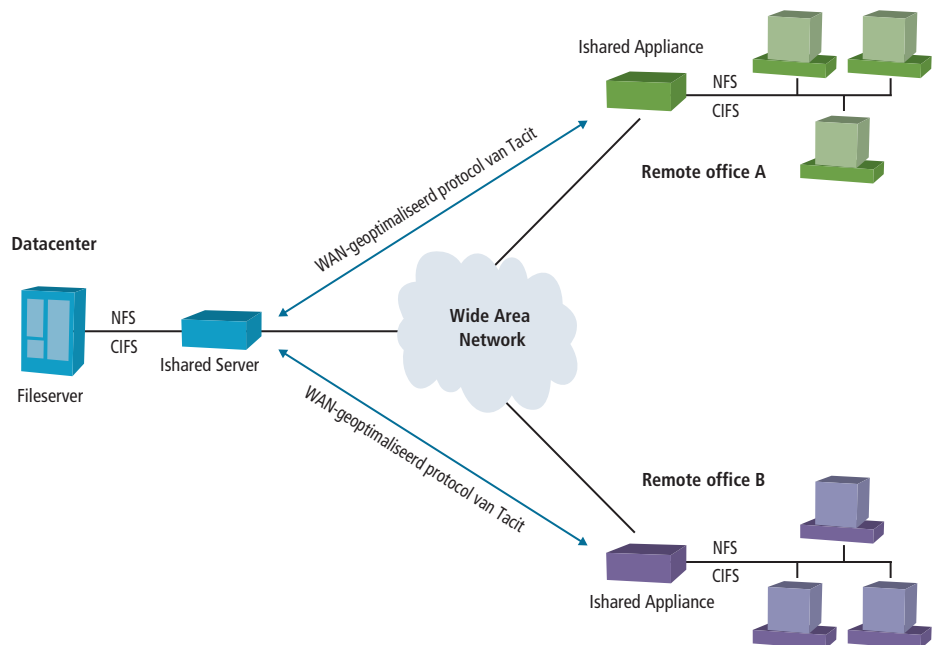
naar het datacenter en de toepassing van back-uphardware, maar de procedurele richtlijnen zijn niet altijd perfect. Zo laten deze maar al te vaak de beslissing tot implementatie aan het kantoor zelf over, wat resulteert in een ongelijke dataprotectie binnen de organisatie. De huidige back-uptechnologieën en praktijktoepassingen proberen zo goed mogelijk de data op de remote offices te beschermen, maar hebben te lijden onder slechte prestaties van het WAN, schaalbaarheid en beperkingen met betrekking tot het beheer. Ook ontbreken de schaalvoordelen bij inkoop van het back-upproduct en de bijbehorende training.

Replicatie naar het datacenter vraagt om een afweging tussen het risico van gegevensverlies door uitgestelde back-ups en de overbelasting van het WAN door het back-upverkeer. Verder moeten beheerders back-upprocessen ontwikkelen en hun servers en opslagsystemen op de juiste manier inzetten, om ervoor te zorgen dat de honderden of soms duizenden remote back-ups volgens een bepaald schema worden uitgevoerd. Toepassing van afzonderlijke back-uphardware op bijkantoren is duur en druist in tegen het streven om de data van de onderneming te consolideren. Wanneer het mogelijk zou zijn om de remote offices realtime online met het

datacenter te verbinden, dan zou dit een oplossing kunnen zijn.

De Tacit Networks Ishared-producten maken realtime lezen en schrijven van gegevens via het WAN mogelijk, met de snelheid, betrouwbaarheid en transparantie van het LAN. Bij toepassing ervan in nevenvestigingen en in het datacenter worden nieuw door de remote-gebruiker gecreëerde data automatisch en realtime binnen het datacenter opgeslagen. De remote-gebruiker ervaart dan dat hij bestanden lokaal opslaat. Met deze toepassing hoeven ondernemingen er zich niet langer druk over te maken of er op remote offices nu wel of niet een back-up wordt bijgehouden. De toepassing van realtime opslag van remote data maakt het plaatsen van speciale fileservers op bijkantoren overbodig, zodat één infrastructuur voor opslag kan worden gecreëerd (zie figuur 3).

De appliances vragen een minimale configuratie en onderhoud die in vergelijking met de traditionele toepassing van remote fileservers en back-upapplicaties, beduidend minder zijn. Kortom, deze methode van opslagconsolidatie bespaart veel aan onderhouds- en installatiekosten, én aan ergernis voor zowel de systeembeheerders als de gebruikers op locatie.



Figuur 3 Opslagconsolidatie via het WAN

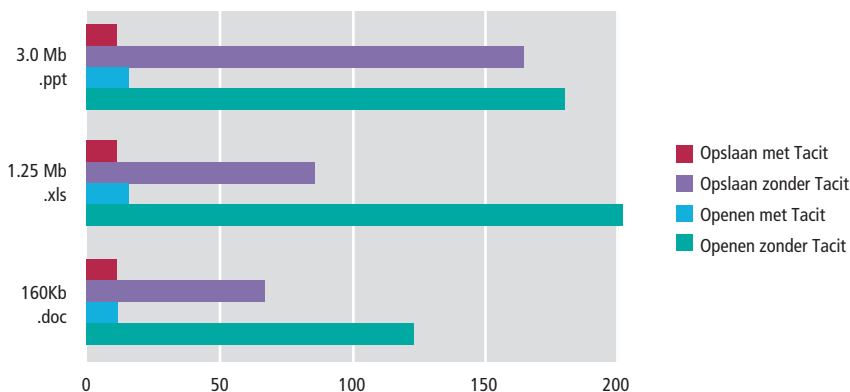
Prestaties

Er zijn diverse onafhankelijke instituten en testlaboratoria die de prestaties van WAFS over een periode van twaalf maanden hebben gemeten in een gesimuleerde WAN-omgeving. Volgens een rapport van de Taneja Group¹ liet de WAFS-toepassing bij een bestandsoperatie over een native WAN een verbetering met factor 78 zien. In figuur 4 zijn de prestaties te zien bij het openen en opslaan van

bestanden via een WAN tussen New York en Hong Kong.

Wanneer we kijken naar de toepassing van WAFS over een periode van drie jaar, dan zien we een aanmerkelijke daling van de hardwarekosten. Ook de investeringen dalen bij de toepassing van WAFS. De keuze om wel of niet een WAFS-product in te zetten, is afhankelijk van de latency van het WAN en de toegepaste applicaties. Waar het break-even point precies ligt, is dus vooraf moeilijk te zeggen. Een toetsingscriterium dat meestal niet in TCO-berekeningen wordt opgenomen is de gebruikerservaring. En die kan wel eens sterk van invloed zijn op de beslissing om wel of niet een WAFS-product te implementeren.

Bram Dons is onafhankelijk IT-analist. E-mail: b.dons@it-trendwatch.nl.



Figuur 4 Gemiddelde tijd (in seconden) voor openen en opslaan MS Officebestanden via een WAN (bron: Tacit)

WWW

www.tacitnetworks.com