



BIG DATA

EEN ETHERNET FABRIC IS DE ESSENTIËLE FUNDERING VOOR BIG DATA

Tot voor kort was de belangrijkste uitdaging voor de meeste organisaties hoe meer transacties sneller te verwerken dan wel op te slaan. Vandaag ligt de nadruk veel meer op het snel kunnen leveren van meer informatie vanuit scale-out cloud computing clusters (zoals documenten, medische afbeeldingen, films, gene sequences, datastreams, tweets) naar systemen, pc's, mobiele devices en woonkamers. De uitdaging is om een betere manier te vinden voor het analyseren, te gelde maken, en het creëren van nieuwe waarde vanuit al deze informatie. Dit is het tijdperk van Big Data.

door: Sander Bakker

Big Data betreft een nieuwe generatie van technologieën en architecturen waarmee organisaties op een economische manier waarde kunnen onttrekken aan zeer grote hoeveelheden data door middel van het op hoge snelheid vastleggen, bekijken en analyseren van deze data. Big Data vereist een verandering in Compute, Storage en Networking zodat organisaties de toevoer, het management en gedistribueerde processing workloads aankunnen die vereist zijn om de grote hoeveelheden data economisch te analyseren. Organisaties die het beste in staat zijn om realtime beslissingen te nemen met behulp van Big Data zullen floreren, terwijl andere die deze verschuiving niet

doormaken zullen falen, zeker in markten waarin een hoge mate van verandering plaats vindt en agressieve consolidaties. Het netwerk is de fundering die nodig is voor de succesvolle implementatie van Big Data en om grote hoeveelheden data vanuit verschillende bronnen te gebruiken. Automatisering, schaalbaarheid en een hoge performance zijn essentieel voor Big Data implementaties.

Datacenternetwerk

Het volume en de verscheidenheid aan data die gegenereerd wordt groeien zo snel dat CIO's zich realiseren dat zij snellere netwerken van een hogere kwaliteit nodig hebben om niet alleen deze groei voor te blijven maar ook om deze data te kunnen

gebruiken voor betere inzichten en zakelijk voordeel. Het netwerk moet de doelstellingen van Big Data ondersteunen. De beste aanpak hiervoor is een platter netwerk met een fabric architectuur om de gedistribueerde data workloads aan te kunnen.

Automatisering

Data wordt niet langer eenvoudigweg opgeslagen en vergeten; data wordt veel meer proactief gebruikt om beslissingen te nemen. Big Data analytics zorgt voor precieze en tijdige inzichten die zich zullen vertalen in zakelijk voordeel. Deze verschuiving zal niet alleen leiden tot een verandering in de manier hoe data opgeslagen en beheerd wordt, maar zal ook zorgen voor een complete verandering in analysetools, samenwerkingsplatformen en applicaties zoals ERP, CRM en financiële systemen.

De intelligentie van het netwerk wordt hierbij steeds belangrijker. De mogelijkheid om automatisch het netwerk te configureren voor een veranderende netwerkbelasting en falende links, maakt het mogelijk om sneller in te spelen op veranderende omstandigheden. Het resultaat is een enorme vermindering van de beheerlast van operationele teams, en het vermindert daarmee de kans op fouten.

High performance

De complexiteit van het op hoge snelheid vastleggen, bekijken en analyseren vereist dat er verder gekeken wordt dan poort-snelheden en veel meer naar de switch-architectuur. Om er zeker van te zijn dat het netwerk de hoge performance levert dient het netwerk als één geheel getest te worden en niet alleen de individuele switches.

Performance is noodzakelijk voor data-inname, of dit nu bulk, micro-batching of streaming is. Veel Big Data projecten proberen data binnen te halen vanuit transactionele systemen. Dit zijn veelal realtime omzet genereerende systemen en Big Data analyse heeft slechts een beperkte tijd, wanneer de business gesloten is, om dit te doen.

Het proces van het onttrekken, transformeren en reloaden van data dient snel te gebeuren. Als Big Data het transactionele systeem en zijn activiteiten hindert, dan zal het niet worden getolereerd in de organisatie. Netwerk doorvoer is dus essentieel voor een succesvolle implementatie van Big Data in transactionele omgevingen.

Schaalbaarheid

De meeste organisaties starten met een kleine pilot en brengen daarna wat ze geleerd hebben mee naar de bedrijfskritische omgeving. Zij hebben dus behoefte aan een architectuur die van klein naar zeer groot kan schalen. Het is dan ook duidelijk dat zij het juiste netwerk vanaf het begin dienen te kiezen. Investeren in

een initieel kleine Fabric zal het netwerk in staat stellen om gemakkelijk te groeien met de CPU- en storagegroei. Deze modulaire aanpak brengt vereenvoudiging naar klanten en reduceert operationele kosten.

Ethernet Fabric

Brocade's netwerk oplossing voor Big Data is de Ethernet Fabric. Het doel van een Ethernet Fabric is allereerst te voorzien in applicatiebeschikbaarheid en resiliency. Daarnaast biedt een Ethernet Fabric een manier om de complexiteit in het netwerk te verminderen en sneller te kunnen inspelen op veranderende omstandigheden.

- Gedistribueerde intelligentie met automatische port-profielmigratie, zodat elke switch alles weet van elke andere switch om zo naadloos Big Data configuraties aan te passen in het datacenter.
- Schaalbaarheid door middel van scale-out in plaats van scale-up.
- Alle switches vormen één logisch chassis, die er voor netwerkbeheerders eruit ziet als één grote switch met een virtueel management plane. Elke fysieke switch in de Fabric wordt gemanaged alsof het een portmodule is in een chassis. Dit zorgt voor schaalbaarheid zonder handmatige configuratie.

Het datacenternetwerk evolueert naar een applicatiegedreven netwerk

Een Ethernet fabric is zeer geschikt voor Big data:

- Een Ethernet Fabric biedt een multipath, non blocking architectuur met een lage latency zonder de noodzaak van complex en inefficiënte protocollen zoals Spanning Tree.
- Door de onchip buffering levert het meer dan dubbel de throughput als enig concurrent. Essentieel voor velen-naar-één communicatie zoals bij computing frameworks als Hadoop.
- Millisecond failover waardoor communicatie niet verbreekt na een uitval van een pad of link.

- Wanneer andere switches connecteren aan de fabric worden ze automatisch onderdeel van de Fabric en zijn er geen handmatige werkzaamheden nodig voor het opzetten van interswitch links. Je plukt de kabel in de ene switch en de andere kant in de andere switch en je hebt de bandbreedte verdubbeld.

Evolutie

Het datacenternetwerk zit middenin de evolutie van een fixed, datacentrische client/server-topologie naar een applicatiegedreven dynamisch netwerk, dat beter geschikt is voor de eisen van Big Data en de noodzaak van nonstop networking. Door het invoegen van grotere flexibiliteit in het netwerk kunnen datacenters beter inspelen op veranderende omstandigheden. Nu organisaties meer en meer kijken naar de mogelijkheden van realtime analytics moeten ze er zeker van zijn dat het netwerk de enorme hoeveelheden data snel en efficiënt kan verplaatsen. Netwerk performance is essentieel voor elk onderdeel van de oplossing, inclusief met hoge snelheid vastleggen, bekijken en analyseren. Een Ethernet Fabric is de technologie die kan voorzien in de netwerkkintelligentie, ultra low latency en hoge performance die nodig zijn om tegemoet te komen aan de eisen van Big Data. «

