

Open Ethernet-oplossingen van Dell Technologies voor generatieve AI

De nieuwe grenzen in IT-infrastructuur verkennen

'AI-fabric (backend switching voor GPU-naar-GPU-connectiviteit) zal naar verwachting groeien van \$ 1,2 miljard (2022) naar \$ 15,2 miljard (2027). Dit is een samengestelde jaarlijkse groei van 65% gedurende een periode van vijf jaar.

Ethernet zal naar verwachting een omzetaandeel van 32% en een poortverkoop van 37% behalen voor AI Fabrics (2027)¹

Dell'Oro Research¹

GPU-prestaties zijn sterk afhankelijk van netwerkprestaties. Omdat veel AI-workloads worden uitgevoerd in grote serverclusters waarvoor constante communicatie tussen rekenknooppunten en storage is vereist, zijn er robuuste netwerken nodig om knelpunten te voorkomen. Als de netwerkprestaties onvoldoende zijn voor de workload, worden GPU's inactief en nemen trainings- en inferentietijden toe, waardoor dataverwerking en time-to-insight worden vertraagd.

Inleiding: de netwerkvereisten van generatieve AI

Naarmate generatieve AI-oplossingen (GenAI) zich blijven ontwikkelen en de grenzen van gegevensverwerking en rekenbehoeften verleggen, worden IT-infrastructuren uitgedaagd om manieren te vinden om aan de immense vereisten van deze omgevingen te voldoen. Voor deze modellen, met name grote taalmodellen (LLM's), is niet alleen meer infrastructuur vereist, maar zijn ook zorgvuldig ontworpen systemen nodig om de enorme connectiviteitsbehoeften in GPU-clusters te beheren. Traditionele netwerkoplossingen vormen al snel knelpunten, wat de levensvatbaarheid en het succes van GenAI-initiatieven bedreigt. Voor AI-fabrics zijn een lage latentie, prestaties zonder verliezen en een maximale bandbreedte vereist.

Enorme vereisten voor de verwerking van data en applicaties leiden tot hogere vereisten voor zowel front- als back-end fabrics.

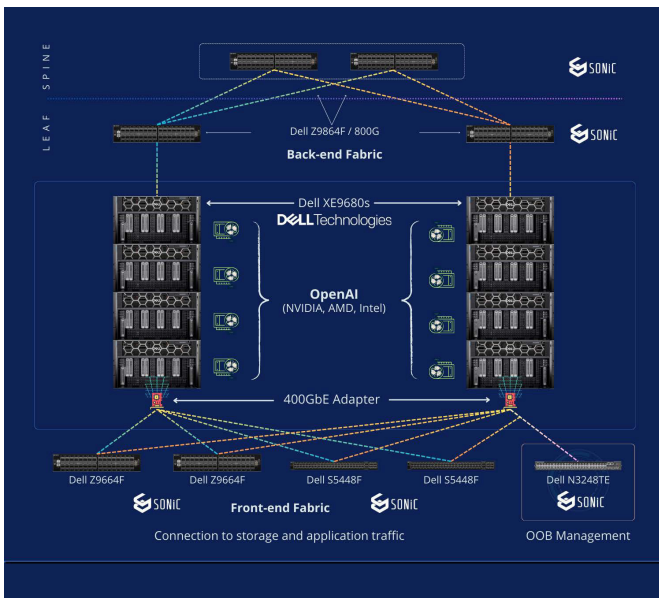
Uitdagingen en behoeften bij de implementatie van een GenAI-infrastructuur

Het implementeren van GenAI-technologieën brengt een reeks uitdagingen met zich mee, van de technische complexiteit die bij nieuwe architecturen komt kijken tot het tekort aan bekwaame professionals die dergelijke implementaties kunnen beheren. Oplossingen die afhankelijk zijn van bedrijfseigen technologieën, zoals InfiniBand, betekenen een extra laag aan complexiteit, beperken de beschikbaarheid van resources en bemoeilijken de integratie met bestaande bewakings- of indelingsplatforms. Bovendien vormen de hoge kosten, lange evaluatietijd en gebondenheid aan bepaalde leveranciers die bij bedrijfseigen oplossingen horen voor aanzienlijke barrières, vooral in een tijdperk van onzekerheid in de leveringsketen. Deze uitdagingen onderstrepen de dringende behoefte aan open, flexibele en robuuste GenAI-infrastructuur oplossingen die aan de unieke vereisten van GenAI-workloads kunnen voldoen.

De aanpak van Dell Technologies voor GenAI-netwerken

Als antwoord op deze uitdagingen heeft Dell Technologies baanbrekend werk verricht op het gebied van uitgebreide en open Ethernet-oplossingen om te voldoen aan de ingewikkelde vereisten van de GenAI-infrastructuur. Dell Technologies heeft een uitgebreide ervaring in AI-, modellerings- en HPC-omgevingen (HighPerformance Computing) en biedt een reeks oplossingen die voldoet aan zowel front-end als back-end vereisten. Dell Technologies biedt u de essentiële componenten voor een succesvolle GenAI-implementatie, van modulaire rekensystemen die zijn geoptimaliseerd voor versnelling, zoals Dell PowerEdge XE servers, tot AI-storageoplossingen zoals PowerScale. Centraal in deze aanpak staat de implementatie van Ethernet-fabrics van de volgende generatie, mogelijk gemaakt door geavanceerd netwerksilicium. Met de **Dell PowerSwitch Z9864-ON die 800 GbE** aan niet-blokkerende netwerkprestaties levert die essentieel zijn voor GenAI-applicaties, kunnen klanten AI-clusters met een lage latentie implementeren en een hoge doorvoer realiseren met behulp van high-bandwidth switching en nieuwe functies die zijn te vinden in **Dells Enterprise SONiC Distribution by Dell Technologies**, zoals Advanced Routing, RoCEv2, Enhanced Hashing en Priority Flow Control, om de fabric-prestaties te verbeteren en beter controle te houden op drukte.

¹Overzicht van Dell'Oro: Geavanceerd onderzoeksrapport over AI-netwerken voor AI-workloads



Voorbeeld van fabric-architectuur op basis van GenAI

Innovaties in GenAI Networking door Dell Technologies

Dell Technologies is toonaangevend op het gebied van GenAI-netwerkinnovatie en biedt oplossingen die voldoen aan de vereisten van GenAI-omgevingen, van edge tot core tot in de cloud. Doordat Dell Technologies zich richt op open en uitbreidbare oplossingen en gebruikmaakt van silicium en opensourcesoftware, zorgt Dell Technologies voor maximale prestaties en flexibiliteit.

Het gebruik van commercieel verkrijgbare opensource netwerkbesturingssystemen zoals SONiC, in combinatie met de actieve deelname aan en bijdrage van Dell Technologies aan het **Ultra Ethernet Consortium (UEC)**, onderstreept de toewijding van Dell Technologies aan open standaarden en gezamenlijke ontwikkeling op het gebied van Ethernet. Deze inspanningen zijn erop gericht ervoor te zorgen dat Ethernet een cruciale rol blijft spelen in de ondersteuning van de volgende generatie AI-omgevingen.

De Dell PowerSwitch Z serie, waarbij gebruik wordt gemaakt van ultramodern silicium, vormt de ruggengraat voor een krachtig, schaalbaar netwerk dat duizenden knooppunten kan ondersteunen, waardoor de connectiviteitsuitdagingen worden aangepakt die ontegenzeggelijk aan GenAI-applicaties zijn verbonden.

De implementatie van GenAI versnellen met Dell Technologies

De opkomst van GenAI heeft een groot aantal uitdagingen voor IT-infrastructuren met zich meegebracht, waardoor er een nieuwe benadering van netwerken is vereist die zowel innovatief als flexibel is. Dell Technologies beantwoordt deze vraag met open Ethernet-oplossingen die niet alleen voldoen aan de onmiddellijke behoeften van GenAI-implementaties, maar ook de basis leggen voor toekomstige ontwikkelingen.

Om het giswerk uit AI-hardwareoplossingen te halen, biedt Dell in het lab gevalideerde referentiearchitecturen die zijn geoptimaliseerd voor AI-workloads. Deze gevalideerde ontwerpen bestaan uit architectuurconcepten, volledige oplossingsoverzichten, prestaties en andere labvalidaties die de mogelijkheden van de oplossing bewijzen voor de workload waarvoor deze is ontworpen. Ga van 'mogelijk voor AI' naar 'bewezen met AI' met gevalideerde oplossingen waarmee u gemakkelijker en sneller diepgaande inzichten verwerft.

Door te kiezen voor Dell Technologies nemen organisaties een partner in de armen met de expertise, een end-to-end stack met oplossingen en de toewijding om het succes van hun GenAI-initiatieven te garanderen. Met Dell Technologies zijn bedrijven goed toegerust om de complexiteit van GenAI-architecturen te doorzien, zodat hun projecten niet alleen levensvatbaar zijn, maar zich ook in een goede positie bevinden om succesvol te zijn.

Versnel de implementatie en time-to-value voor uw GenAI-omgevingen, verminder risico's en uw operationele complexiteit met Dell Technologies. We nodigen u uit om te ontdekken hoe een open, flexibele en duurzame netwerkoplossing uw GenAI-initiatieven kan transformeren, waardoor uw bedrijf een nieuw tijdperk van innovatie en efficiëntie ingaat.



Meer informatie
over Dell
Networking



Neem contact op
met een Dell
Technologies expert



Lees het
analistenoverzicht
van ESG



Lees het
analistenoverzicht
van IDC