

# Soluzioni Open Ethernet Dell Technologies per l'AI generativa

## Le nuove frontiere dell'infrastruttura IT

"Si prevede che la fabric AI (switching di backend per la connettività da GPU a GPU) aumenterà da \$ 1,2 miliardi (2022) a \$ 15,2 miliardi (2027), con un CAGR quinquennale del 65%.

Si prevede che l'Ethernet raggiungerà il 32% di quota di fatturato e il 37% di spedizioni di porte per le fabric AI (2027)"

### Studio Dell'Oro<sup>1</sup>

Le prestazioni della GPU dipendono fortemente dalle prestazioni di rete. Con molti carichi di lavoro AI in esecuzione su cluster di server di grandi dimensioni, che richiedono una comunicazione costante tra i nodi di elaborazione e lo storage, occorrono reti robuste per evitare colli di bottiglia. Se le prestazioni di rete sono insufficienti per il carico di lavoro, le GPU diventano inattive e i tempi di addestramento e inferenza aumentano, rallentando l'elaborazione dei dati e il tempo di analisi degli stessi.

### Introduzione: le esigenze di rete dell'AI generativa

Man mano che le soluzioni di AI generativa (GenAI) continuano a evolversi, allargando i confini delle esigenze computazionali e di elaborazione dei dati, le infrastrutture IT devono trovare il modo di supportare gli immensi requisiti di questi ambienti. Tali modelli, in particolare i modelli linguistici di grandi dimensioni (LLM), richiedono non solo un'infrastruttura più ampia, ma anche sistemi appositamente progettati per gestire le enormi esigenze di connettività dei cluster della GPU. Le soluzioni di rete tradizionali si stanno trasformando in colli di bottiglia, minacciando la fattibilità e il successo delle iniziative GenAI. Le fabric AI richiedono bassa latenza, prestazioni ineccepibili e la massima larghezza di banda. Enormi requisiti di elaborazione dei dati e delle applicazioni determinano un aumento degli stessi per le fabric front-end e back-end.

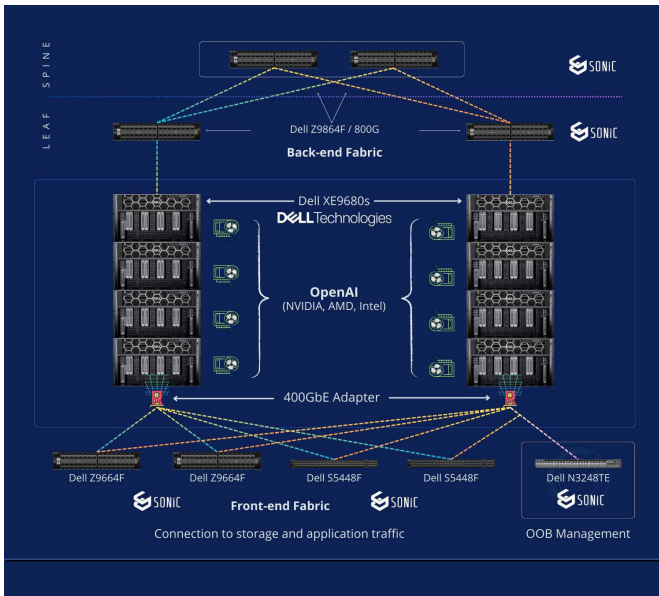
### Sfide ed esigenze dell'implementazione dell'infrastruttura GenAI

L'implementazione di tecnologie GenAI introduce una serie di sfide, dalle complessità tecniche associate alle nuove architetture alla scarsità di professionisti qualificati in grado di gestire tali implementazioni. Le soluzioni che si basano su tecnologie proprietarie, come InfiniBand, aggiungono un altro livello di complessità, limitando la disponibilità delle risorse e complicando l'integrazione con le piattaforme di monitoraggio od orchestration esistenti. Inoltre, i costi elevati, i lunghi tempi di valutazione e la dipendenza dai fornitori associati alle soluzioni proprietarie rappresentano ostacoli significativi, soprattutto in un'epoca di incertezze nella supply chain. Queste sfide evidenziano l'urgente necessità di soluzioni di infrastruttura GenAI aperte, flessibili e robuste in grado di soddisfare le esigenze specifiche dei carichi di lavoro GenAI.

### L'approccio Dell Technologies al networking GenAI

In risposta a queste sfide, Dell Technologies ha aperto la strada a soluzioni Ethernet complete e aperte, progettate per soddisfare le complesse esigenze dell'infrastruttura GenAI. Sfruttando la vasta esperienza in ambienti AI, modellazione e High Performance Computing (HPC), Dell Technologies offre una suite di soluzioni che soddisfano i requisiti di front-end e back-end. Dai sistemi di elaborazione modulari ottimizzati per l'accelerazione, come i server Dell PowerEdge XE, alle soluzioni di storage incentrate sull'AI come PowerScale, Dell Technologies fornisce i componenti essenziali per un deployment di successo della GenAI. Al centro di questo approccio vi è il deployment delle fabric Ethernet di nuova generazione con silicio di rete avanzato. Con **Dell PowerSwitch Z9864-ON** che fornisce **800 GbE** di prestazioni di rete non bloccanti fondamentali per le applicazioni GenAI, i clienti implementano cluster AI con bassa latenza e throughput elevato utilizzando switching a elevata larghezza di banda e le nuove funzionalità disponibili in **Dell Enterprise SONiC Distribution di Dell Technologies** come Advanced Routing, RoCEv2, Enhanced Hashing e Priority Flow Control, per migliorare le prestazioni della fabric e monitorare la congestione.

<sup>1</sup>Brief Dell'Oro: Advanced Research Report on AI Networks for AI Workloads



*Esempio di architettura della fabric GenAI*

## Innovazioni nel networking GenAI di Dell Technologies

Dell Technologies è all'avanguardia nell'innovazione delle reti GenAI, offrendo soluzioni che soddisfano i requisiti degli ambienti GenAI di oggi e di domani, dall'edge al core fino al cloud. Concentrandosi su soluzioni aperte ed estendibili, sfruttando silicio commerciale e software open source, Dell Technologies garantisce massime prestazioni e flessibilità.

L'utilizzo di sistemi operativi di rete open source disponibili in commercio come SONiC, insieme alla partecipazione attiva e al contributo di Dell Technologies all'**Ultra Ethernet Consortium (UEC)**, sottolinea l'impegno verso gli standard aperti e lo sviluppo collaborativo nello spazio Ethernet. Tali iniziative mirano a garantire che l'Ethernet continui a svolgere un ruolo fondamentale nel supportare la prossima generazione di ambienti AI.

Dell PowerSwitch serie Z, che utilizza silicio all'avanguardia, fornisce la dorsale per una rete scalabile a prestazioni elevate in grado di supportare migliaia di nodi, affrontando così le sfide di connettività inerenti alle applicazioni GenAI.

## Accelerare l'implementazione della GenAI con Dell Technologies

L'aumento della GenAI ha introdotto una serie di sfide per le infrastrutture IT, richiedendo un nuovo approccio al networking che risulti allo stesso tempo innovativo e flessibile. Dell Technologies risponde a questa richiesta con soluzioni aperte e basate su Ethernet, che non solo soddisfano le esigenze immediate dei deployment GenAI, ma gettano anche le basi per i progressi futuri.

Per eliminare le incertezze dalle soluzioni hardware AI, Dell offre architetture di riferimento convalidate in laboratorio e ottimizzate per i carichi di lavoro AI. Queste progettazioni convalidate includono concept di architettura, panoramiche complete della soluzione, convalide delle prestazioni e altre di laboratorio, che dimostrano le capacità della soluzione sul carico di lavoro per il quale è stata progettata. Passa dalle possibilità dell'AI a quelle comprovate con soluzioni convalidate che semplificano la fornitura di informazioni rapide e approfondite.

Scegliendo Dell Technologies, le organizzazioni ottengono un partner con l'esperienza, lo stack di soluzioni end-to-end e l'impegno per garantire il successo delle loro iniziative GenAI. Con Dell Technologies, le aziende affrontano le complessità delle architetture GenAI e garantiscono che i loro progetti non solo siano fattibili, ma anche idonei al successo.

Accelera il deployment e il time-to-value per gli ambienti GenAI, riducendo i rischi e la complessità operativa con Dell Technologies. Scopri in che modo una soluzione di rete aperta, flessibile e sostenibile trasforma le iniziative GenAI, proiettando il business in una nuova era di innovazione ed efficienza.



Scopri di più su  
Dell Networking



Contatta un esperto  
Dell Technologies



Leggi il brief  
degli analisti  
ESG



Leggi il brief  
degli analisti  
IDC