



**Dell offre un portafoglio più ampio di soluzioni, servizi e opzioni di consumo per supportare i clienti nell'implementazione dei loro progetti di intelligenza artificiale**



**Prestazioni fino al 37% migliori in metà dello spazio su rack**

per l'elaborazione del linguaggio naturale (GPT-J) per dati MLPerf



**I Dell Validated Design includono più architetture di riferimento incentrate sull'AI rispetto alle soluzioni convalidate Supermicro**

## Trovare il percorso per il successo dell'AI con il portafoglio AI Dell

### Confronto tra il portafoglio AI Dell e le offerte analoghe di Supermicro

Per implementare l'intelligenza artificiale (AI) e trarne vantaggi, è necessario considerare tutti gli aspetti di una soluzione. È fondamentale trovare un partner in grado di offrire supporto end-to-end, che comprenda pianificazione, preparazione dei dati, selezione dell'hardware, progettazione dei modelli di AI, test proof-of-concept, architetture di riferimento e supporto end-to-end. In questo riepilogo, evidenziamo i vantaggi in termini di prestazioni e spazio dei server nel portafoglio AI Dell™ rispetto a quelli di Supermicro, basandoci sui risultati dei test di benchmark di MLPerf® disponibili pubblicamente. MLPerf® testa le prestazioni relative sia all'addestramento che all'inferenza su diversi modelli di AI. I dati contenuti in questo riepilogo provengono dai risultati di MLPerf® v3.1 Inference Datacenter pubblicati sul sito web MLCommons® a partire da novembre 2023.<sup>1</sup> Abbiamo anche confrontato le informazioni disponibili pubblicamente sul range di prodotti e soluzioni che ciascun fornitore offre per supportare le implementazioni di AI.

### Server a quattro GPU: i server Dell PowerEdge offrono prestazioni MLPerf migliori rispetto ai server Supermicro

Con un massimo di quattro GPU NVIDIA H100 SXM, Dell PowerEdge XE9640 2U offre la metà della potenza di elaborazione della GPU di PowerEdge XE9680 in un terzo dello spazio.<sup>2</sup> Il server a quattro GPU con GPU NVIDIA HGX H100 di Supermicro negli invii MLPerf® è SYS-421GU-TNXR, che è un server 4U. Nei risultati pubblicati del modello di AI gptj-99.9, PowerEdge XE9640 ha superato le prestazioni del server Supermicro nei test offline, ottenendo fino a 1,37 volte il punteggio (vedi Figura 1).

Abbiamo inoltre confrontato i risultati del server Supermicro SYS-421GU-TNXR con Dell PowerEdge XE8640, un server 4U a quattro GPU che supporta anche GPU NVIDIA H100 HGX. Dell PowerEdge XE8640 è dotato dei più recenti processori scalabili Intel Xeon di quarta generazione e fino a 4 TB di memoria per gestire dataset di grandi dimensioni e calcoli complessi, comuni nell'AI e nell'analisi dei dati.<sup>3</sup> Con il suo fattore di forma 4U, PowerEdge XE8640 è simile a Supermicro SYS-421GU-TNXR sia in termini di densità che di funzionalità GPU. Tuttavia, come abbiamo visto con PowerEdge XE9640, Dell PowerEdge XE8640 ha ottenuto punteggi gptj-99 migliori nei test offline rispetto al server Supermicro (vedi Figura 2).

### Risultati di MLPerf® normalizzati: Dell PowerEdge XE9640 con H100 SXM5 rispetto a Supermicro SYS-421GU-TNXR con H100 SXM5 (L'opzione più grande è migliore)

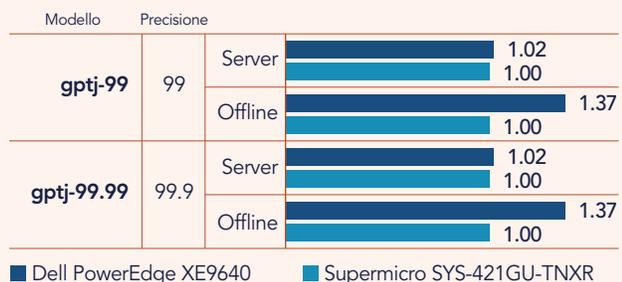


Figura 1. Risultati pubblicati di MLPerf® relativi a Dell PowerEdge XE9640 e Supermicro SYS-421GU-TNXR al 29/11/2023. Entrambi i sistemi utilizzano il fattore di forma SXM della GPU NVIDIA H100. Fonte: Principled Technologies con i dati di MLCommons®.\*

### Risultati di MLPerf® normalizzati: Dell PowerEdge XE8640 con H100 SXM5 rispetto a Supermicro SYS-421GU-TNXR con H100 SXM5 (L'opzione più grande è migliore)

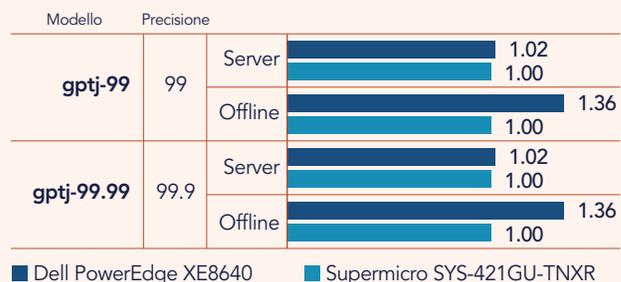


Figura 2. Risultati pubblicati di MLPerf® relativi a Dell PowerEdge XE8640 e Supermicro SYS-421GU-TNXR al 29/11/2023. Entrambi i sistemi utilizzano il fattore di forma SXM della GPU NVIDIA H100. Fonte: Principled Technologies con i dati di MLCommons®.\*

## Oltre l'hardware: uno sguardo al quadro generale

Le organizzazioni che cercano flessibilità nelle offerte di workstation GPU possono trovare una gamma più ampia di opzioni nel portafoglio Dell AI. Il più ampio portafoglio Dell AI comprende molto più di un semplice hardware: offre ai clienti un'unica soluzione di servizi e supporto per ogni fase del deployment. Al contrario, il portafoglio Supermicro AI offre servizi limitati a manuali, risoluzione dei problemi e restituzione dei prodotti. Nella nostra ricerca, non siamo riusciti a trovare servizi di progettazione, implementazione, gestione o formazione nel portafoglio Supermicro AI. L'adozione dell'AI è molto complessa e, grazie ai servizi che offre, Dell è un partner di formazione più promettente per il deployment di questi carichi di lavoro. Inoltre, abbiamo rilevato che i Dell Validated Design includono più architetture di riferimento incentrate sull'AI rispetto alle soluzioni convalidate Supermicro, oltre a offrire indicazioni più approfondite.

Leggi il nostro report completo, che esplora in modo approfondito i vantaggi offerti dal portafoglio Dell AI in termini di opzioni di elaborazione, storage e rete, servizi professionali per l'AI, partnership di terze parti per i carichi di lavoro AI, Dell Validated Design, che eliminano le incertezze dalle soluzioni hardware AI, servizi di gestione e iDRAC.

\*Punteggio MLPerf verificato di v3.1 Inference Closed. Consultato il 5 dicembre 2023 da <https://mlcommons.org/benchmarks/inference-datacenter/>, voci 3.1-0133, 3.1-0066 e 3.1-0067. Il nome e il logo MLPerf sono marchi registrati e non registrati di MLCommons. Association negli Stati Uniti e in altri Paesi. Tutti i diritti riservati. L'uso non autorizzato è severamente vietato. Consultare [www.mlcommons.org](http://www.mlcommons.org) per maggiori informazioni.

1. MLCommons, "MLPerf Inference: Datacenter Benchmark Suite Results", consultato il 7 febbraio 2024, <https://mlcommons.org/en/inference-datacenter-31/>.
2. Dell, "PowerEdge XE9640 Rack Server", consultato il 5 gennaio 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>.
3. Dell, "PowerEdge XE8640 Rack Server", consultato il 5 gennaio 2024, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>.

► La versione originale in inglese di questo riepilogo è disponibile all'indirizzo <https://facts.pt/kRTwky3>

Clicca qui per leggere il report ►



Facts matter.®

Principled Technologies è un marchio registrato di Principled Technologies, Inc. Tutti gli altri nomi di prodotto sono marchi dei rispettivi proprietari. Per ulteriori informazioni, consulta il report.