I 10 motivi principali per scegliere HPE GreenLake per Large Language Models

HPE GreenLake

La maggior parte delle aziende non ha mai potuto permettersi la potenza del supercomputing, per una serie di ragioni che variano dal notevole esborso di capitale iniziale ai lunghi tempi di attesa, fino alla necessità di assumere personale operativo specializzato, solo per citare qualche esempio. Ma oggi le cose sono cambiate.

HPE GreenLake per Large Language Models offre gli stessi livelli di velocità, controllo e governance dei supercomputer on-premise, con l'agilità e la semplicità di utilizzo del cloud. Questa offerta cloud native consente ai clienti di sfruttare attraverso il browser la potenza aggiuntiva del supercomputing per eseguire simulazioni on demand su vasta scala con modelli AI di grandi dimensioni. Grazie a questa offerta di supercomputing as-a-service, le imprese possono sfruttare l'enorme potenza dei supercomputer, eliminando tutti i tradizionali ostacoli all'accesso.

I modelli complessi e gli enormi set di dati necessari per la modellazione, la simulazione e l'intelligenza artificiale (AI) stanno spingendo le aziende a ricorrere a una capacità di elaborazione che, solo fino a pochi anni fa, era riservata alle strutture di supercomputing. Ad esempio IDC prevede che, a breve, le aziende smetteranno di includere i laboratori HPC (High Performance Computing) in una categoria di spesa a parte,[1] mentre fra il 2016 e il 2022 i requisiti di elaborazione per i modelli AI su vasta scala sono raddoppiati ogni 10,7 mesi.[2]

Per dominare il proprio settore, le imprese hanno bisogno di accedere alla potenza del supercomputing per effettuare simulazioni su vasta scala e utilizzare modelli di AI di grandi dimensioni.

È arrivato il momento di vivere la nuova normalità del futuro, scegliendo l'accesso cloud native all'enorme potenza del supercomputing con HPE GreenLake per LLM. Accelera il time to market dei nuovi prodotti e le iniziative AI più cruciali, con tutta la velocità del supercomputing, rivolgendoti allo stesso partner di fiducia che ti fornisce già le tue risorse IT on-premise.

Leggi di seguito i 10 motivi principali per cui HPE GreenLake per LLMs consente di potenziare on demand le tue capacità on-premise:

1 Risparmio di tempo

Grazie agli ambienti di supercomputing preconfigurati e prontamente disponibili, gli utenti possono risparmiare moltissimo tempo sull'installazione e la configurazione dell'infrastruttura di elaborazione. Possono iniziare a eseguire subito i propri LLM senza attendere i tempi normalmente richiesti per la consegna, la distribuzione e il provisioning dell'infrastruttura.

2 Costi contenuti

HPE GreenLake per LLMs evita l'investimento di capitale iniziale necessario per acquistare l'infrastruttura e assumere personale operativo specializzato, riducendo sia CapEx che OpEx. Gli utenti pagano solamente le risorse di supercomputing gestite di cui hanno bisogno, tramite un modello di pagamento a consumo* che può essere molto più economicamente vantaggioso, rispetto alla creazione e alla gestione di una struttura di supercomputing dedicata, con relativo team dedicato.

3 Elasticità

Il progetto offre tutta la scalabilità necessaria per aumentare e ridurre le risorse di supercomputing, senza alcun limite di architettura. Gli utenti possono adattare agevolmente la potenza di elaborazione e la capacità di storage in base alle esigenze di business.

4 Accessibilità

Gli utenti possono accedere alle risorse di supercomputing ovunque sia disponibile una connessione Internet. In questo modo, i membri dei team possono collaborare anche a distanza, oltre che accedere da remoto a potenti risorse di elaborazione, senza necessariamente trovarsi fisicamente vicino all'infrastruttura.

* Può essere soggetto a limiti minimi o potrebbe essere applicabile la capacità di riserva

[1] Worldwide High-Performance Computing Server Forecast, 2023-2027: Enterprise Will Overtake HPC Labs, IDC, aprile 2023

[2] COMPUTE TRENDS ACROSS THREE ERAS OF MACHINE LEARNING, marzo 2022

5 Flessibilità

HPE GreenLake per LLMs offre una vasta gamma di opzioni di elaborazione, che consentono agli utenti di selezionare la configurazione e le specifiche maggiormente in linea con le loro esigenze. Possono scegliere le istanze di supercomputing basate sul Transformer Engine della GPU o della CPU più adatto ai requisiti delle loro applicazioni.

6 Supporto fornito da esperti

Il progetto assicura accesso a tutta l'esperienza tecnica necessaria per ottimizzare le risorse di supercomputing. Gli utenti possono richiedere l'assistenza degli esperti di progettazione delle prestazioni, per determinare la configurazione ideale dei sistemi e ottimizzare il software in modo da completare i processi nel minor tempo possibile.

7 Ottimizzazione delle risorse

HPE GreenLake per LLMs si avvale di sofisticati ambienti software, come HPE Machine Learning Development Environment e HPE Cray Programming Environment, che consentono di accelerare i processi di modellazione e simulazione, oltre che di sviluppare e addestrare modelli di AI di grandi dimensioni su una determinata infrastruttura.

8 Tecnologia all'avanguardia

HPE GreenLake per LLMs si basa su tecnologie di supercomputing exascale per offrire agli utenti la possibilità di accedere alle ultime novità in materia di hardware, software e architettura. Le aziende possono sfruttare le migliori tecnologie senza preoccuparsi di tenere il passo con gli upgrade e l'evoluzione rapida del mercato.

9 Sovranità dei dati

HPE GreenLake per LLMs non addebita alcun costo di uscita dei dati. Grazie alla possibilità di evitare le spese aggiuntive imprevedibili associate al trasferimento dei dati all'esterno del cloud di supercomputing, le aziende sono libere di spostare i dati come desiderano, senza temere sanzioni.

10 Sostenibilità ambientale

Il servizio di supercomputing-as-a-service HPE viene erogato da strutture di colocation sostenibili, che vengono alimentate per almeno il 99,5%[3] da fonti rinnovabili, costituite soprattutto da energia idroelettrica, e riutilizzano il calore disperso dai sistemi. Ultimamente l'impatto ambientale del supercomputing ha suscitato molte polemiche e, proprio per questo, è importante fornire il supercomputing-as-a-service con modalità rispettose dell'ambiente.

Se, per motivi di tempo o di business, hai l'esigenza cruciale di eseguire un LLM per un'iniziativa AI strategica, oppure una simulazione su vasta scala che richiede l'uso prolungato di centinaia (o anche migliaia) di GPU o CPU di fascia alta per un periodo di tempo prolungato, **HPE GreenLake per LLMs** è la scelta giusta per te.

Se l'infrastruttura on-premise attuale non dispone dello spazio necessario per rispondere a questa esigenza, non rimandare oltre e scegli HPE GreenLake per LLMs, che assicura accesso cloud native a tutta la potenza del supercomputing.

[3] QScale Campus Q1, giugno 2023

Ulteriori informazioni alla pagina

HPE.com/GreenLake/llm

Visita **HPE GreenLake**

Chatta ora (commerciale)

Hewlett Packard Enterprise

a50008571ITE