



# Top 10 reasons to choose HPE GreenLake for Large Language Models

**HPE**   
**GreenLake**

Traditionally most enterprises have struggled to adopt supercomputing for a variety of reasons, like large upfront capital outlay, long lead times, and specialized operational skills requirements just to name a few. Not anymore.

HPE GreenLake for Large Language Models combines the speed, control, and governance of on-premises supercomputers with the cloud's agility and ease of use. With this cloud-native offering, customers can tap into additional supercomputing power for large-scale simulations of large AI models on demand — through a browser. This supercomputing as-a-service offering brings the superpower of supercomputing to enterprises — removing all traditional access barriers.

Complex models and larger data sets for modeling, simulation, and artificial intelligence (AI) are pushing enterprises onto a level of computing capability that just a few years ago was reserved for supercomputing sites. For example, IDC is forecasting that enterprises will soon overtake high-performing computing (HPC) labs from an HPC spending perspective,<sup>1</sup> and the computing requirements for large-scale AI models doubled every 10.7 months from 2016 to 2022.<sup>2</sup>

Enterprises need access to supercomputing power for large-scale simulations and large AI models to stay ahead in their industry.

Hop aboard the next normal and get cloud-native access to the superpower of supercomputing with HPE GreenLake for LLM. Accelerate time-to-market for new products and key AI initiatives at supercomputing speed with the partner you already trust with your on-premises IT.



Here are the top 10 reasons to choose HPE GreenLake for LLM to augment your on-premises capabilities on

1

### Time savings

By leveraging preconfigured and readily available supercomputing environments, users can save significant time setting up and configuring their computing infrastructure. They can quickly start running their large language models without the delays associated with infrastructure shipment, deployment, and provisioning.

2

### Cost-effectiveness

HPE GreenLake for LLM helps eliminate the need for up-front investment in infrastructure and operational skills, reducing capital and operational expenses. Users pay for the managed supercomputing resources they need on a pay-as-you-go\* basis, which can be more cost-effective than building and maintaining a dedicated supercomputing facility and team.

3

### Elasticity

The project enables users to scale their supercomputing resources up or down — without architectural limitations. They can easily adjust the computing power and storage capacity based on business demands.

4

### Accessibility

Users can access supercomputing resources from anywhere with an internet connection. This enables collaboration among geographically dispersed teams and facilitates remote access to powerful computing capabilities without the need for physical proximity to the infrastructure.

\* May be subject to minimums or reserve capacity may apply

<sup>1</sup> [Worldwide High-Performance Computing Server Forecast, 2023–2027: Enterprise Will Overtake HPC Labs](#), IDC, April 2023

<sup>2</sup> [COMPUTE TRENDS ACROSS THREE ERAS OF MACHINE LEARNING](#), March 2022

5

**Flexibility**

HPE GreenLake for LLM offers a range of computing options, allowing users to choose the configuration and specifications that best suit their needs. They can select transformer engine GPU-based or CPU-based supercomputing instances that best fit their application requirements.

6

**Expert support**

The project offers access to technical expertise for optimizing supercomputing resources. Users can benefit from the assistance of performance engineering experts who help with optimal system configuration and software optimization to get the job done in the fastest time possible.

7

**Resource optimization**

HPE GreenLake for LLM employs sophisticated software environments like HPE Machine Learning Development Environment and HPE Cray Programming Environment that can speed up modeling and simulation jobs, as well as develop and train large AI models on a given infrastructure.

8

**Cutting-edge technology**

HPE GreenLake for LLM is built on exascale supercomputing technologies helping ensure access to the latest advancements in hardware, software, and architectural innovations. Enterprises can leverage state-of-the-art technology without the burden of keeping up with rapid advancements and upgrades.

9

**Data sovereignty**

HPE GreenLake for LLM does not charge data egress fees. Enterprises can avoid additional and unpredictable charges associated with transferring their data out of the supercomputing cloud enabling penalty-free data mobility.

10

**Environmental sustainability**

The HPE supercomputing-as-a-service offering is delivered out of sustainable colocation facilities that are powered by a minimum of 99.5%<sup>3</sup> renewable energy — mainly hydropower — and capture waste heat for reuse. The environmental impact of supercomputing has been getting a lot of attention recently, so it is important to provide supercomputing-as-a-service in an environmentally sustainable way.

**HPE GreenLake for LLM** is for you if you have a business- and time-critical need to run a large language model for a strategic AI initiative or a large-scale simulation on hundreds (or even thousands) of high-end GPUs or CPUs for an extended time.

If you have insufficient space to implement that demand on your existing on-premises infrastructure, then do not wait any longer and get cloud-native access to the superpower of supercomputing with HPE GreenLake for LLM.

<sup>3</sup> QScale Campus Q1, June 2023

**Learn more at**

[HPE.com/GreenLake/llm](https://HPE.com/GreenLake/llm)

Visit **HPE GreenLake**



**Chat now (sales)**

**Hewlett Packard  
Enterprise**

© Copyright 2023 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

a50008571ENW