Postavit základy pro Enterprise AI je netriviální úkol

Devět z deseti firem zavádí AI, meziroční tempo růstu investic do AI infrastruktury dosahuje 37 %,[1] a firmy využívající AI mají tržní valuaci o 22 % vyšší, než srovnatelná konkurence, která tak nečiní.[2]

Adopce AI ve firmách a veřejném sektoru však často provází řada obtíží, které nasazení a škálování AI projektů brání. Pouze 10% AI projektů se dostane do produkce, přičemž tato cesta může trvat více než 7 měsíců.[3]

Opakujícími se výzvami jsou zajištění vysoké kvality dat, jejich správa a ochrana napříč hybridním prostředím, nedostatek kvalifikovaných specialistů a zvyšující se náklady plynoucí z růstu využívání AI v public cloudu.

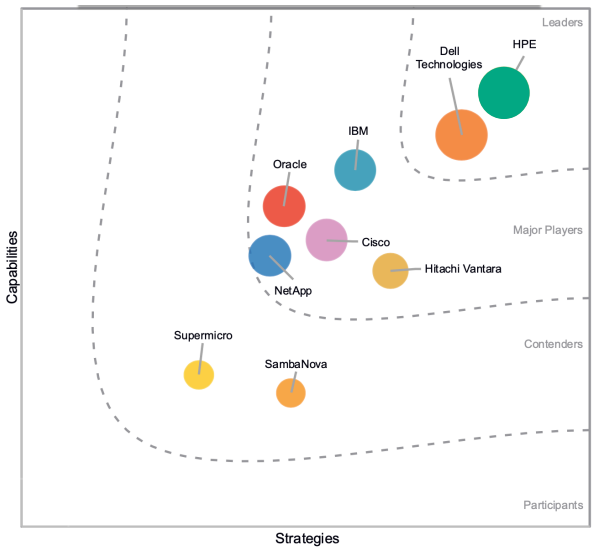
Jaká řešení se nabízejí?

Rozšířeným přístupem je využívání public cloudu, který s sebou však pro AI projekty nese rizika v oblasti bezpečnosti a regulace dat, omezuje kontrolu nad daty i modely, často neumožňuje splnění compliance požadavků a vede k neočekávaným nákladům.

Vybudování a údržba vlastního HW a SW stacku (tzv. DIY přístup) znamená vysoké nároky na interní znalosti a pomalý přechod do produkce – vybudování vlastního produkčního AI stacku může znamenat časovou investici 14-21 měsíců, nehledě na absenci jednotné podpory.[4]

HPE Private Cloud AI je platforma na klíč, poskytující cloudovou provozní zkušenost, připravená k rychlému nasazení v datacentru dle zákazníkova určení

IDC MarketScape Worldwide Private AI Infrastructure Systems, 2025



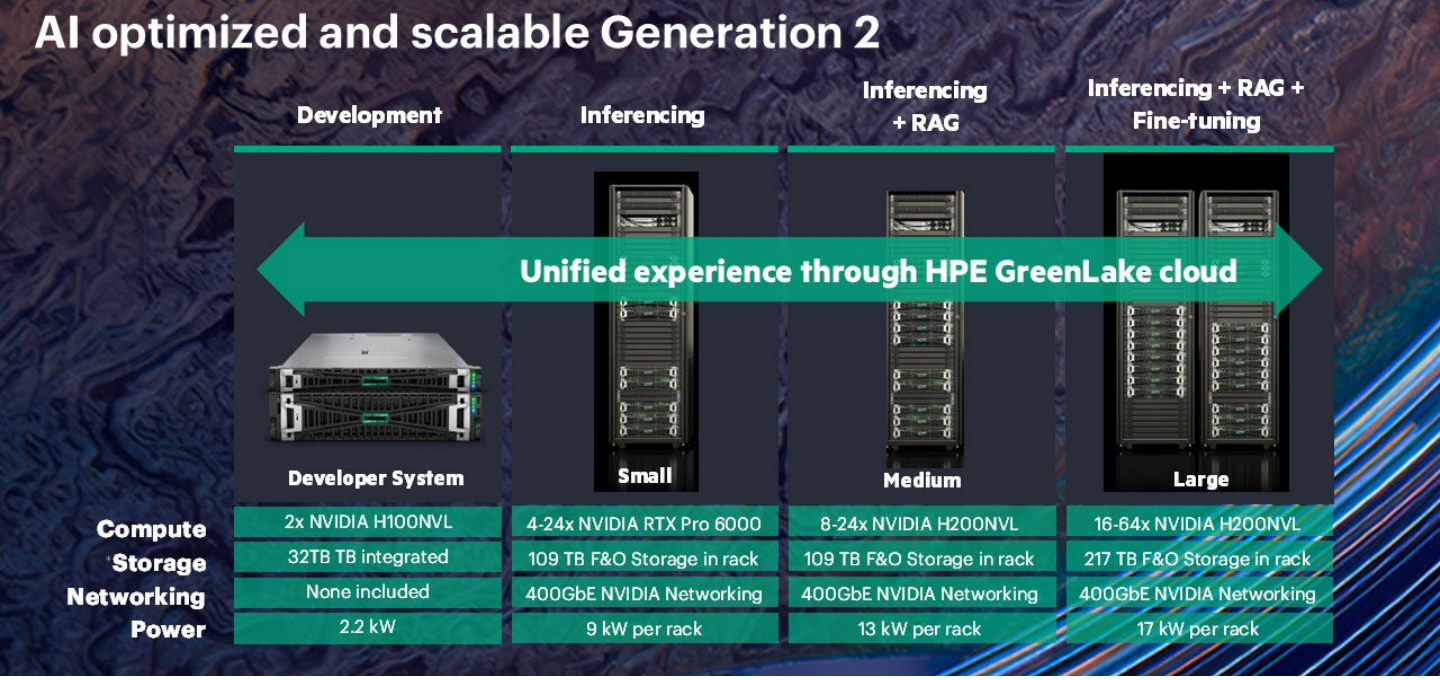
HPE je Leadrem v Private AI Systémech dle IDC

HPE Private Cloud AI

HPE Private cloud AI (PCAI) je hotová platforma pro provoz i vývoj AI aplikací, kterou lze rychle uvést do provozu a snadno rozšiřovat. Nabízí pohodlí cloudu, robustní bezpečnost a efektivní řešení klíčových potřeb týmů komerčních a veřejných organizací.

Kompletní integrace SW a HW a garantovaná podpora výrobce výrazně snižuje nároky na interní kapacity potřebné pro implementaci, provoz, správu a rozvoj AI. Díky platformnímu přístupu a široké paletě integrovaných nástrojů zároveň umožňuje rychlé dosažení výsledků. Predikovatelné provozní náklady pak podporují pozitivní návratnost investice.

[1]Worldwide Semiannual Artificial Intelligence Infrastructure Tracker, IDC · [2] Journal of Financial Economics, volume 151, 2024 · [3]Reasons Why Generative AI Pilots Fail To Move Into Production, Forbes, 2024 · [4]The Economic Benefits of HPE Private Cloud AI with NVIDIA AI Computing, Enterpriise Strategy Group, 2025



Co dělá z HPE Private Cloud AI ideální platformu pro Enterprise AI

Okamžitá AI produktivita

Dodání do 4–6 týdnů, on-site instalace během 1 dne. Integrovaný HW a SW stack, doručený plně konfigurovaný, připravený k okamžitému vývoji a provozu AI/ML projektů.

AI Developer začíná pracovat. Nástroje má hned k dispozici, případně AI Admin snadno přidá další.

Cloudová zkušenost

Tričkové velikosti umožňují rychlý start. Systémy Small, Medium a Large nabízí možnost navyšovat výkon podle projektových potřeb v průběhu životního cyklu platformy.

HPE GrenLake Cloud Portál poskytuje pokročilé real-time reporty o stavu služby, data governance, end-to-end nákladech PCAI.

All-in-one update celého PCAI stacku z jednoho místa a v jednom otestovaném aktualizačním balíčku.

Nekompromisní bezpečnost

Zero trust architektura. Zabezpečení nejen na úrovni autentizace uživatelů a rolí. Bezpečnost řešíme na všech úrovních - u datových zdrojů, u jednotlivých workloadů, dokonce i na úrovni API pro jednotlivé AI modely. Airgapped režim pro chod systému bez připojení k internetu umožňuje práci s těmi nejcitlivějšími daty.[5]

AI týmy snadno spolupracují

Pomocí vestavěné funkce Enterprise Multitenancy mohou jednoduše spolupracovat na sdíleném projektu a využívat přidělené zdroje, nebo naopak izolovat svůj projekt od všech ostatních. To vše efektivně a bezpečně.

Jednotný přístup k datům

Vestavěné konektory do existujících datových zdrojů – strukturovaných, nestrukturovaných, datalaků, objektových i souborových úložišť, u vás i v cloudu.

AI aplikace mohou využívat data na stávajících zařízeních, a lze je také zpřístupnit v rychlé PCAI storage.

Architektura a technologie

SW NVIDIA AI Enterprise

Zajistí přístup k potřebným knihovnám, frameworkům a nepřeberné škále základních AI modelů různých velikostí, výkonů a zaměření. Můžete si zprovoznit vlastní RAG pipeline nebo snadno importovat předpřipravený NVIDIA Blueprint. To vše díky integrovanému SW NVIDIA AI Enterprise.

SW HPE AI Essentials

Tvoří základ PCAI. Orchestruje kontejnery, zajišťuje bezpečnost a dohled, integruje moderní AI/ML nástroje jako Apache Airflow, Spark, Kubeflow, MLFlow, inferenční server MLIS pro správu životního cyklu nasazených jazykových modelů a řadu dalších. Navíc - všechny vestavěné SW komponenty mají technickou podporu od HPE.

Specifikace PCAI Developer systému

HPE AI Essentials SW	Zahrnuto
NVIDIA AI Enterprise SW	Zahrnuto
Instalace & Startup	Zahrnuto
Support & Maintenance 24x7, 4h odezva	Poskytuje HPE pro kompletní HW a SW stack, 3/5 let
Control Node	1x HPE ProLiant Gen11
Worker Node	1x HPE ProLiant Gen11
CPU (Worker node)	2x Xeon 32 Core CPU
GPU	2x H100NVL
Úložiště	32TB Internal NVMe SSD File/Object
NIC Speed (AI Network)	200Gb NICs

[5] Funkctionalita Airgapped režimu se vztahuje na PCAI systém Medium.