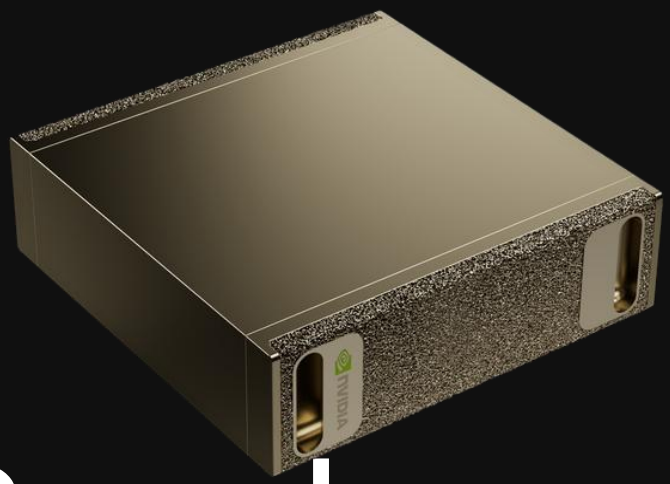
即納体制完備

日本初導入



NVIDIA DGX™ Spark

NVIDIA DGX™ Sparkは、AIの学習・推論処理を強力にサポートする、
小型AIスーパーコンピューターです。
Blackwell世代GPUアーキテクチャと高度なソフトウェアスタックを搭載し
AI開発の高速化と効率化を実現します。

手に取りやすい価格で
ハイエンドなコンピュータパワーを。



販売価格

オープンプライス

NVIDIA GB10 Superchip

NVIDIA Grace Blackwell アーキテクチャを
使用した、FP4 精度で最大1000 AI TOPS の
AIパフォーマンスを体験してください。

128 GB の コヒーレント統合システム メモリ

最大 2,000 億個のパラメーターを持つ AI モデル
を使用してAI 開発とテストのワークロードを、
大容量の統合システムメモリを搭載した
デスクトップで実行できます。

4つの特長

NVIDIA ConnectX ネットワーキング

高性能なNVIDIA Connect-X ネットワーキングで
2つのDGX Spark システムを接続し、
最大4,050 億個のパラメーターを持つ
AI モデルを活用できます。

NVIDIA AI ソフトウェアスタック

ツール、フレームワーク、ライブラリ、
トレーニング済みモデルを含む、生成 AI ワークロー
ド向けのフルスタックソリューションを
活用できます。

DGX Sparkが解決する「3つの壁」と技術的優位性

AI開発現場が直面する制約とDGX Sparkによるブレークスルー

AI開発現場が直面する3つの制約

GPU環境の制約
- 専用ハード導入がボトルネック -

クラウドのコスト
- 利用拡大に伴うコスト増と性能制約 -

セキュリティ懸念
- 機密データ利用時のリスクと規制遵守 -



DGX Sparkが研究・開発現場のこれらの課題を解決します。

すぐに始めてスモールスタート

開発・検証 / 低コスト / GPU環境

複雑な設定なしですぐに開発・検証を始められ、スモールスタートが可能。個人や研究室単位でも導入できる専用GPU環境で、効率的で再現性の高いAI開発を実現します。

オンプレミスでコスト削減と最適化

低コスト化 / リアルタイム性 / ワークロード最適化

ローカル処理により、クラウド依存のコスト増大や遅延を回避。最適構成でワークロード処理

機密データを出さずに安全に利用可能

オンプレミス / 法規制遵守 / セキュリティ強化

オンプレミス完結で、データ主権やセキュリティリスクに対応し、法規制も遵守可能

想定されるDGX Sparkの主なユースケース

AI基盤活用編（生成AI）

生成AIローカル開発

ローカル環境で大規模モデルを活用

- 大規模モデル推論（～200B）
- ファインチューニング（～70B）
- GPT-3相当と同等環境

LLM / Comfy UI / ローカル開発

ハルシネーション対策RAG

信頼性の高い生成AI活用

- 社内文書を安全に活用
- 新情報防止と信頼性確保
- 大規模マニュアル統合

検索拡張生成 / 信頼性向上 / 情報精度

業種別応用編（AIエージェント）

製造・検査AI

品質向上と効率化を同時に実現

- 外観検査・画像AIによる不良品検出
- 外観検査・VLMで製品チェック
- 予知保全・故障パターン予測を検出

画像認識 / 異常検知 / 品質管理

医療研究・ヘルスケア

安全なローカルデータ活用で診断精度と研究再現性を向上

- 診断支援・医用画像で異常検出
- 患者情報要約・診療DB連携
- 医療研究・臨床NLP解析

画像認識 / データ分析 / ローカル運用

DGX Sparkがもたらす業種共通の価値

- データセキュリティ
- 開発高速化
- 拡張性・将来性
- コスト削減
- RAG×安全データで高信頼な生成AI



さまざまなユースケースに展開可能