日本コンピュータシステム販売店協会 ビデオセミナー

HITACHI
Inspire the Next

生成AIを業務に活用するための日立の取り組み

2025年1月
株式会社 日立製作所
クラウドサービスプラットフォームビジネスユニット
フロントエンゲージメント推進本部　本部長
橋本 進太郎

Contents

1. 日立の取り組みとお客さまの期待
2. 業務適用に向けた業務特化型LLM構築の取り組みのご紹介
3. 新サービスのご紹介

© Hitachi, Ltd. 2025. All rights reserved.

1. 日立の取り組みとお客さまの期待

© Hitachi, Ltd. 2025. All rights reserved.

1-1　日立の生成AIの取り組み

HITACHI
Inspire the Next

オフィスワーカーやフロントラインワーカーの人財不足などの社会課題解決に向けて、One Hitachiで推進

パワーグリッド、原子力、鉄道など

グリーンエナジー＆
モビリティセクター

金融、公共、電力、交通、通信など

デジタルシステム＆
サービスセクター

ファクトリー、テック、ビル、ホームなど

コネクティブ
インダストリーズセクター

お客さま向けサービスの提供

ナレッジの提供

日立社内の改革

日立全体での
ナレッジの蓄積

© Hitachi, Ltd. 2025. All rights reserved.

1-2　日立全体でのナレッジ蓄積のための取り組み

HITACHI
Inspire the Next

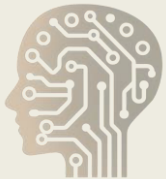
幅広い事業領域・業務で、社内検証を推進中

© Hitachi, Ltd. 2025. All rights reserved.

1-3 生成AIに対する期待の高まり

HITACHI
Inspire the Next

	これまで	これから
生成AIへの期待	生産性の向上 業務の効率化	人財不足の解消 技能継承の実現 競争力の強化
必要な生成AIモデル	汎用知識を広く学習した汎用型生成AIモデル ↓ 汎用LLM利用	事業領域、業務にあわせてカスタマイズが必要 ・自社知識(形式知・暗黙知)の継承 ・自社データ利用で競争力向上 ↓ 業務特化型LLM構築



© Hitachi, Ltd. 2025. All rights reserved.

2. 業務適用に向けた業務特化型LLM構築の取り組みのご紹介

© Hitachi, Ltd. 2025. All rights reserved.

2-1　本格的な業務適用で直面する壁

HITACHI
Inspire the Next

回答精度

・業務レベルの回答が得られない
・熟練者の知見のインプット方法が不明
・学習方法やデータの作り方がわからない

学習環境の整備

・膨大なコンピュートリソースへの投資が困難
・特有な環境の構築・運用ノウハウがない

© Hitachi, Ltd. 2025. All rights reserved.

2-2 本格的な業務適用で直面する壁 - 回答精度

HITACHI
Inspire the Next

回答精度

- 業務レベルの回答が得られない
- 熟練者の知見のインプット方法が不明
- 学習方法やデータの作り方がわからない

学習環境の整備

- 膨大なコンピュートリソースへの投資が困難
- 特有な環境の構築・運用ノウハウがない

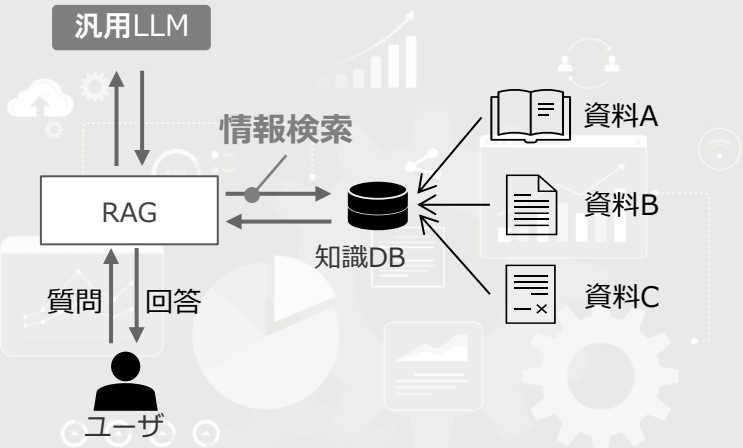
© Hitachi, Ltd. 2025. All rights reserved.

2-2-1　回答精度の壁　RAGの課題

HITACHI
Inspire the Next

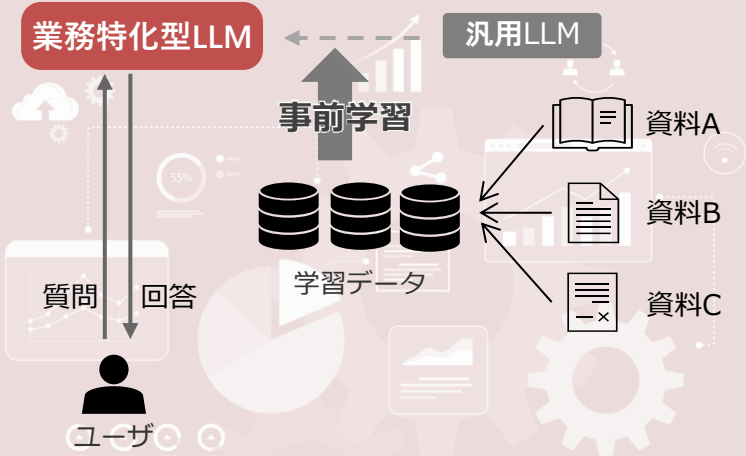
RAGを利用した業務適用が進みつつあるが、複数資料の参照や専門用語が必要な業務では回答精度が不安定

RAGの課題



- 検索でヒットした内容のみ参照するため正確な回答になっていない
- 専門用語、社内用語だと回答精度が悪化

業務特化型LLMによる解決



- **複数ドキュメントを学習するため、回答が正確**
- **専門用語、社内用語にも対応可能**

© Hitachi, Ltd. 2025. All rights reserved.

2-2-2　JP1認定コンサルタントへの挑戦（１）

HITACHI
Inspire the Next

JP1認定コンサルタントとは・・・

JP1
日立のIT運用管理ソフトウェア。
オンプレミス・マルチクラウド環境を対象とした複雑なシステムを統合的に管理

JP1認定コンサルタント
JP1各製品カテゴリーごとに最適なコンサルテーションができるエンジニアを認定

■資格保有者数 **7.1**万人
（JP1認定資格全体）　（2024年3月現在）

■**合格率 24%**の難関資格
（2024年7月現在）

試験内容	認定基準	試験時間
試験問題 30問	70%以上 正解	60分

JP1
CERTIFIED
Consultant

複合的な知識が必要で、単体のドキュメント学習では習得が困難

➡ **複合的な知識を学ばせた業務特化型LLMにより合格をめざす！**

© Hitachi, Ltd. 2025. All rights reserved.

2-2-3　JP1認定コンサルタントへの挑戦（２）

HITACHI
Inspire the Next

オープンLLM をファインチューニングし、JP1の最難関資格である『認定コンサルタント試験』合格

■ LLM別 試験結果

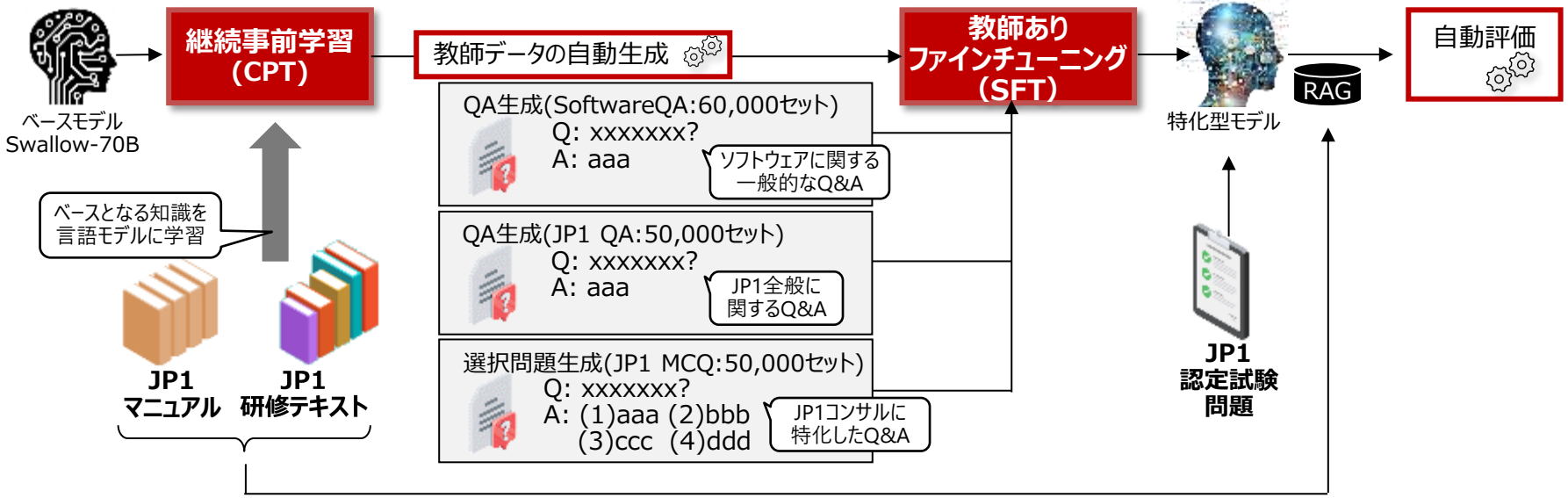
試験種別	JP1認定コンサルタント試験	
LLM種別	RAG無	RAG有
Swallow-70B FT(JP1学習)済	58	70
Swallow-70B(プレーン)	43	63
GPT-3.5 Turbo	31	53
GPT-4	47	65

JP1認定コンサルタント: JP1各カテゴリー製品について、最適なコンサルテーションができるテクニカルスキルを修得したエンジニアを、カテゴリーごとに認定
FT: Fine-tuning　ファインチューニング　LLMを特定の用途に対して最適化させる手法

© Hitachi, Ltd. 2025. All rights reserved.

2-2-4　JP1認定コンサルタントへの挑戦（３）

JP1マニュアル、JP1認定試験受験者向けの研修テキスト、基礎となるテクニカル情報などを複数のファインチューニングのモデルで段階的に学習



※CPT: Continuous pretraining
SFT: Supervised Fine Tuning
前ページの評価結果はJP1 MCQを学習した場合

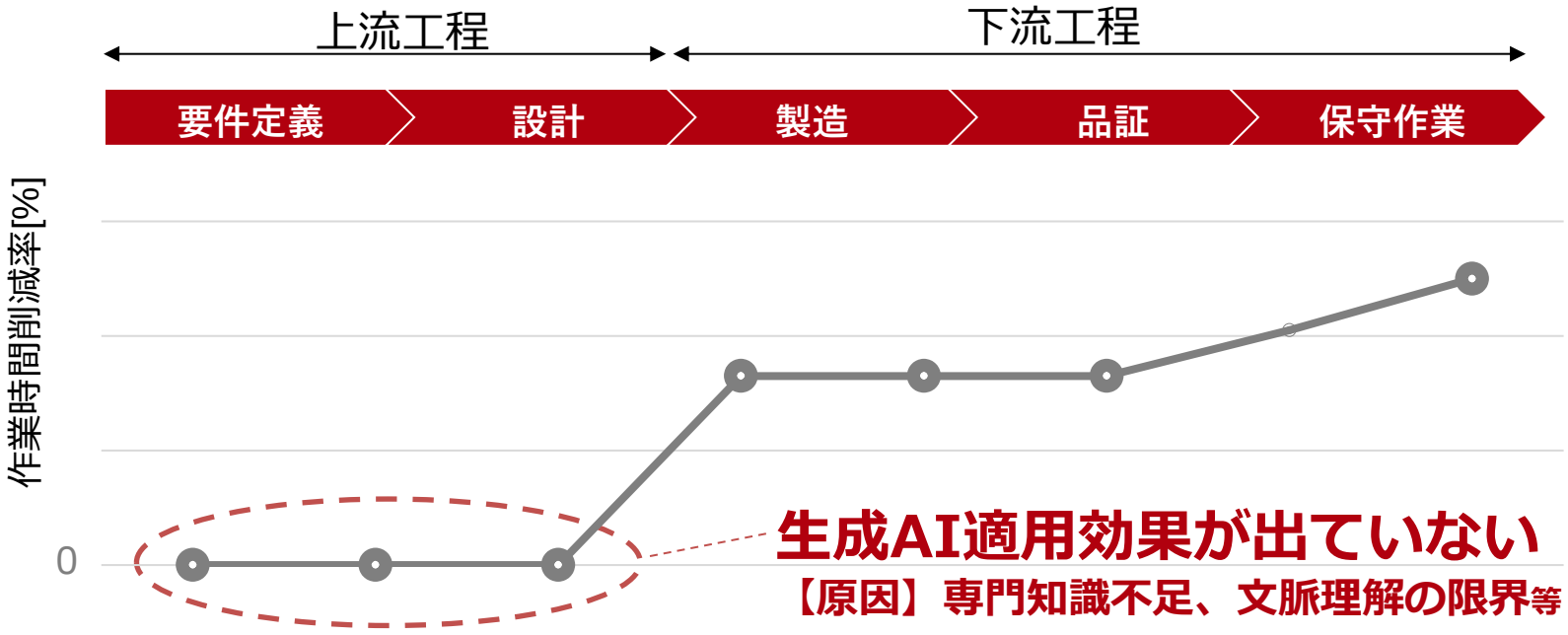
迅速化のため、日立で独自に開発した特化型LLM構築・評価パイプラインを活用

© Hitachi, Ltd. 2025. All rights reserved.

2-2-5　金融システム開発での挑戦（１）

HITACHI
Inspire the Next

社内実証の結果、汎用LLMだけでは上流工程において精度に課題があり、業務特化型LLMの開発に挑戦

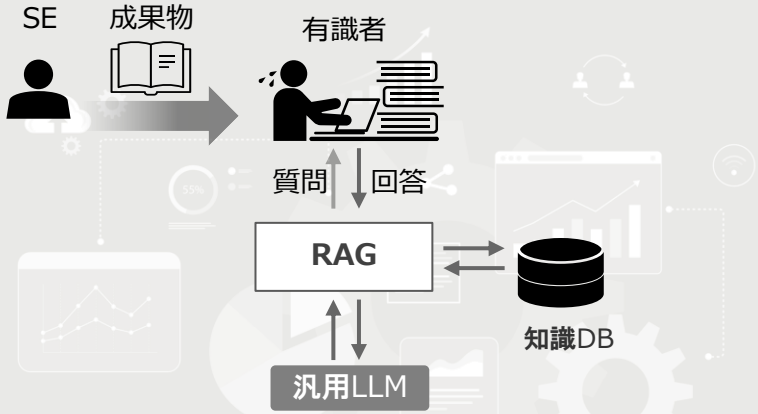


© Hitachi, Ltd. 2025. All rights reserved.

HITACHI
Inspire the Next

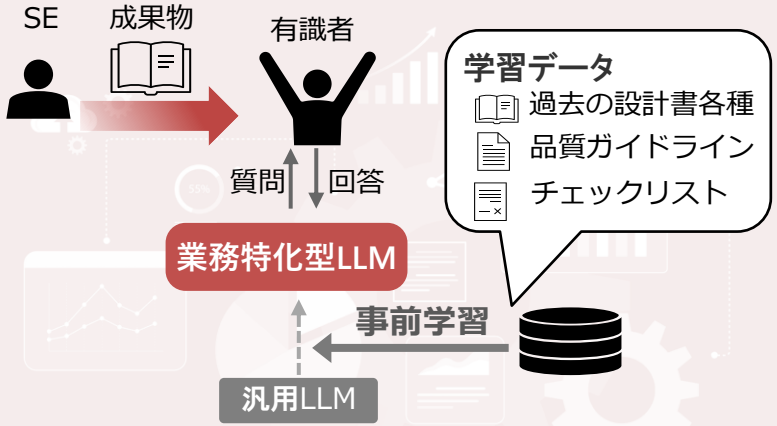
金融系システム開発の設計書レビュー向けに、日立の品質管理ノウハウを学習した業務特化型LLMの構築を推進中

汎用LLMの課題



- 汎用LLMが、品質ガイドラインやチェック項目などを理解していないため、有識者のレビュー負担軽減せず

業務特化型LLMによる解決

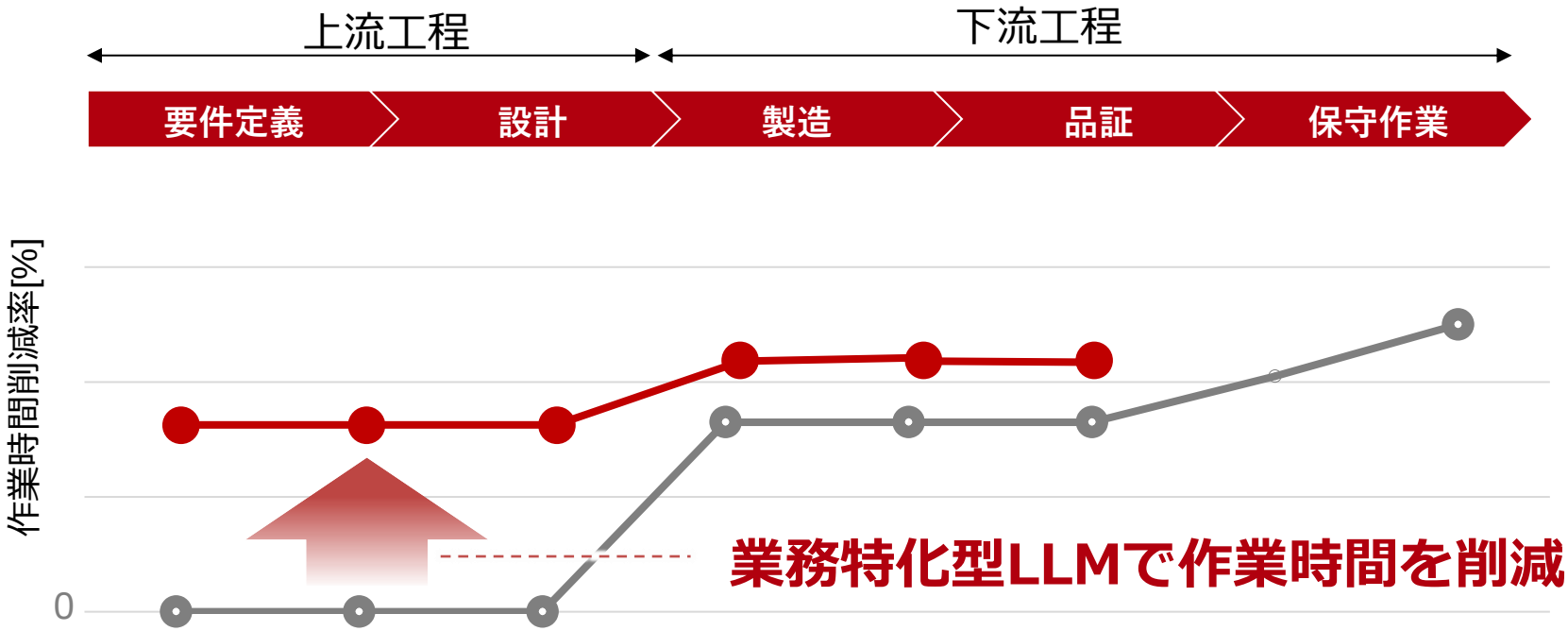


- **品質ガイドライン、チェック項目のレビュー効率化、作業時間削減**
- **指摘漏れなどのヒューマンエラー抑制**

© Hitachi, Ltd. 2025. All rights reserved.

HITACHI
Inspire the Next

業務特化型LLMにより上流工程でも作業効率向上



© Hitachi, Ltd. 2025. All rights reserved.

2-2-8 OT領域における業務特化型LLMへの取り組み

HITACHI
Inspire the Next

例：OT領域におけるベテラン（エキスパート）ノウハウ継承

© Hitachi, Ltd. 2025. All rights reserved.

2-3 本格的な業務適用で直面する壁 – 学習環境の整備

HITACHI
Inspire the Next

回答精度

- 業務レベルの回答が得られない
- 熟練者の知見のインプット方法が不明
- 学習方法やデータの作り方がわからない

学習環境の整備

- 膨大なコンピュートリソースへの投資が困難
- 特有な環境の構築・運用ノウハウがない

© Hitachi, Ltd. 2025. All rights reserved.

2-3-1 学習環境整備の壁の解決

HITACHI
Inspire the Next

整備の壁 ・高性能GPU導入（コスト面、供給面）、導入後の膨大な電力消費
・生成AI特有の環境を整備・運用する人財の確保

解決ポイント1
日立のデータセンタに生成AI基盤を構築、構築・運用技術を蓄積

- ✓高性能NVIDIA GPUサーバを大規模導入
- ✓高速NW(InfiniBand 400Gbps等)で接続
- ✓円滑なデータ学習活用を実現するストレージ環境

解決ポイント2
膨大な発生熱・消費電力対応のための空調/電源設備を完備

- ✓本環境の安定稼働を実現する専用冷却設備、空調を完備
- ✓膨大な電力消費に対応する電源設備、配線を完備

2027年度までにカーボンニュートラル化をめざす横浜データセンターを活用

© Hitachi, Ltd. 2025. All rights reserved.

HITACHI
Inspire the Next

３. 新サービスのご紹介

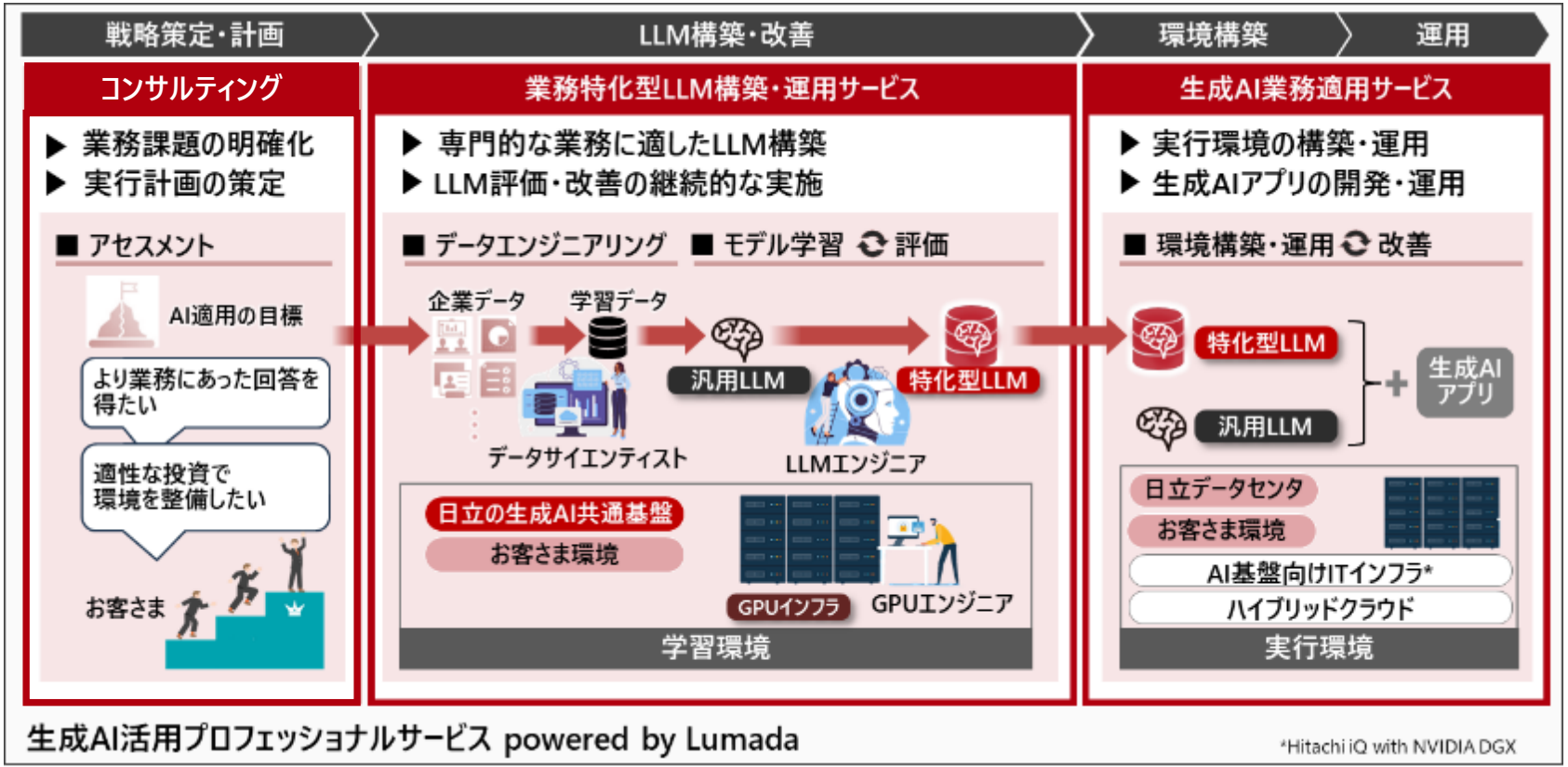
- 業務特化型LLM構築・運用サービス
- 生成AI業務適用サービス
- AIインフラソリューション　Hitachi iQ

© Hitachi, Ltd. 2025. All rights reserved.

3-1　培ったノウハウ・技術でお客さまの生成AI活用を支援

HITACHI
Inspire the Next

試行段階から本格的な業務適用フェーズへ加速する2サービスを2024年10月より提供開始



© Hitachi, Ltd. 2025. All rights reserved.

HITACHI
Inspire the Next

NVIDIAとの戦略的パートナーシップによるAI環境向け製品を提供

NVIDIA

- AIソフトウェアなどの高い技術
- 先進的なGPUおよび生成AI向けツール

戦略的パートナーシップにより
相互の革新的技術を融合

日立

- OT分野でのドメインナレッジ
- 次世代サーバー・ストレージ

Hitachi iQ

NVIDIAのGPUを搭載したサーバーと日立の次世代ストレージを組み合わせたAI向けインフラを提供

NVIDIA GPU 搭載サーバー × Hitachi Vantara 次世代ストレージ

- ’24年3月　社外発表
- ’24年7月　国内販売開始
- ’24年10月　ラインナップ強化

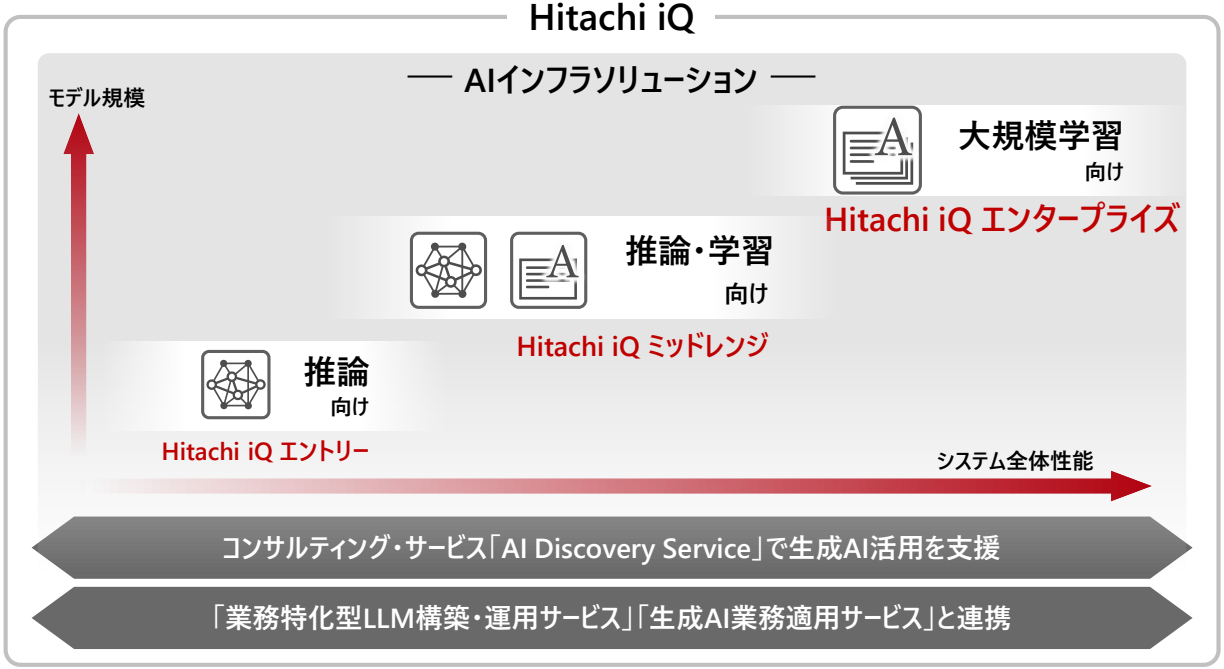
GPU: Graphics Processing Unit

© Hitachi Vantara, Ltd. 2024. All rights reserved.

3-2-1 AIインフラソリューション Hitachi Iq ラインアップ

小規模LLMの推論環境やPoC用途などに適したコンパクトなエントリーモデルから、
大規模学習やクラウドサービス基盤などに適用できるエンタープライズモデルまでラインアップ

Hitachi iQのラインアップ



各モデルの概要

Hitachi iQ エンタープライズ

NVIDIA GPU搭載サーバーと
日立高速ストレージによる
BasePOD構成

Hitachi iQ ミッドレンジ/エントリー

NVIDIA GPU搭載サーバー

ミッドレンジ

… GPU最大4枚
搭載可能

エントリー

… GPU最大3枚
搭載可能

LLM: Large Language Model

© Hitachi Vantara, Ltd. 2025. All rights reserved.

まとめ

01 業務での生成AI活用には
業務特化型LLMが必要だが、
回答精度、人財、環境の壁がある

02 日立は、お客さま用途に特化した
業務特化型LLM構築、AIインフラ含む
環境整備をend to end支援

03 適用領域を継続して拡大し、
現場の障害対応や生産工程の高度化など
フロントラインワーカーの業務へも
順次拡大予定

© Hitachi, Ltd. 2025. All rights reserved.

HITACHI
Inspire the Next

END

生成AIを業務に活用するための日立の取り組み

© Hitachi, Ltd. 2025. All rights reserved.

付録：AIインフラプラットフォームのご紹介
～ラインアップ～

© Hitachi, Ltd. 2025. All rights reserved.

付録1. Hitachi iQの構成

- NVIDIA DGX BasePOD™認定と日立グループ内での利用で培ったナレッジをベースにお客様にご提供
- スケールアウト可能な分散ファイルシステムが学習データを高速に渡すことでGPUサーバの稼働率を向上
- オブジェクトストレージとの組み合わせにより大量の学習データもコスト効率良く運用が可能

InfiniBandスイッチ

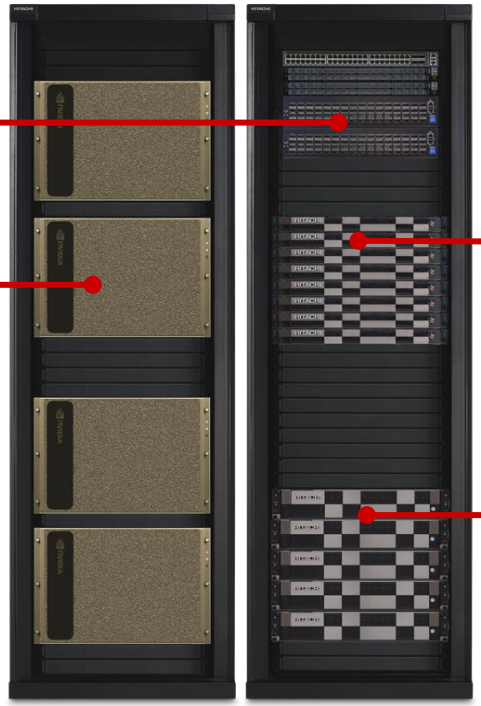
NVIDIA Quantum-2™
高帯域(200Gb/400Gb)なInfiniBandに対応

GPUサーバ

NVIDIA DGX™ H100
AI開発に最適化された最新サーバ

NVIDIA Base Command™
AI開発用のプラットフォーム

NVIDIA AI Enterprise™
AI開発を支援するソフトウェア群



分散ファイルシステム

Hitachi Content Software for File
スケールアウト可能な高速ファイルシステム

オブジェクトストレージ

Hitachi Content Platform
大量データのアーカイブに最適なストレージ

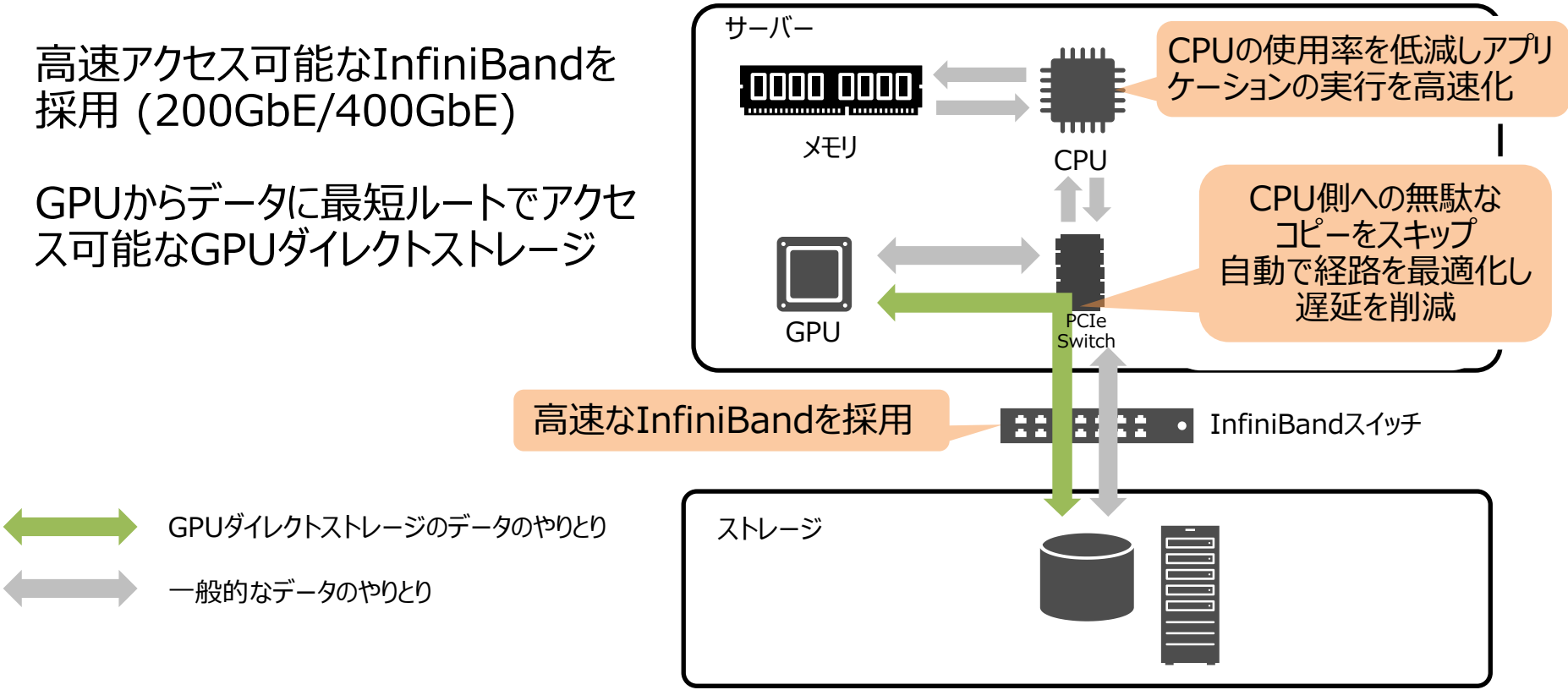
© Hitachi Vantara, Ltd. 2025. All rights reserved.

付録1. Hitachi iQ：大量のデータを高速に転送する仕組み

HITACHI
Inspire the Next

大量データのAI学習で採用されている「InfiniBand＋GPUダイレクトストレージ」の構成を採用

- 高速アクセス可能なInfiniBandを採用 (200GbE/400GbE)
- GPUからデータに最短ルートでアクセス可能なGPUダイレクトストレージ



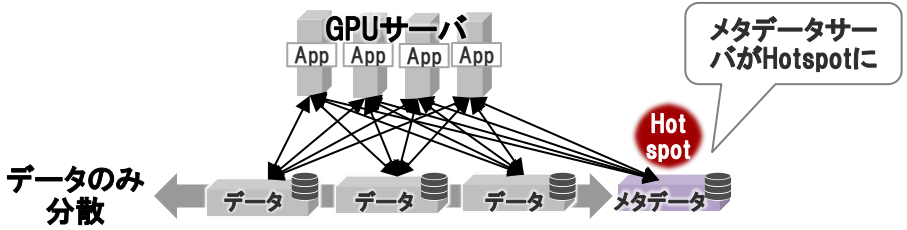
© Hitachi, Ltd. 2025. All rights reserved.

付録1. Hitachi iQ：高パフォーマンスを支える独自のファイルシステム技術

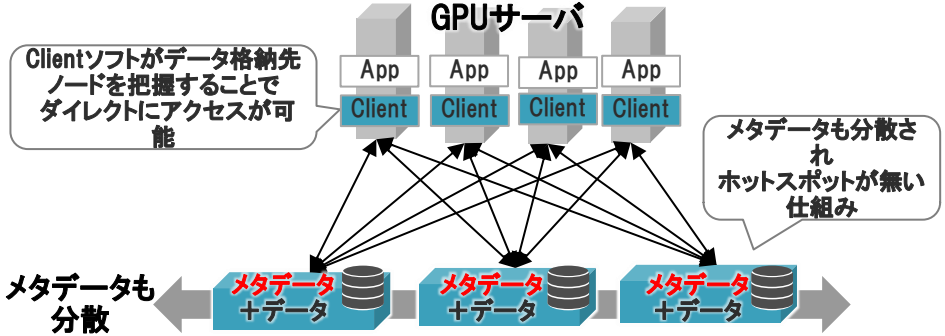
HITACHI
Inspire the Next

ホットスポットがない分散ファイルシステムで性能がリニアに向上

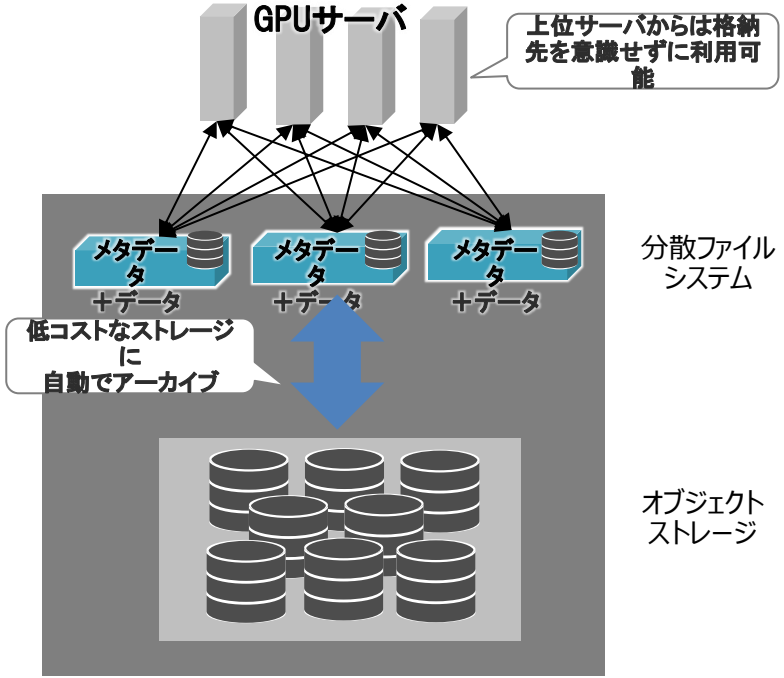
通常の分散ファイルシステム



Hitachi Content Software for File



アーカイブストレージとの組み合わせにより大量の学習データもコスト効率良く運用



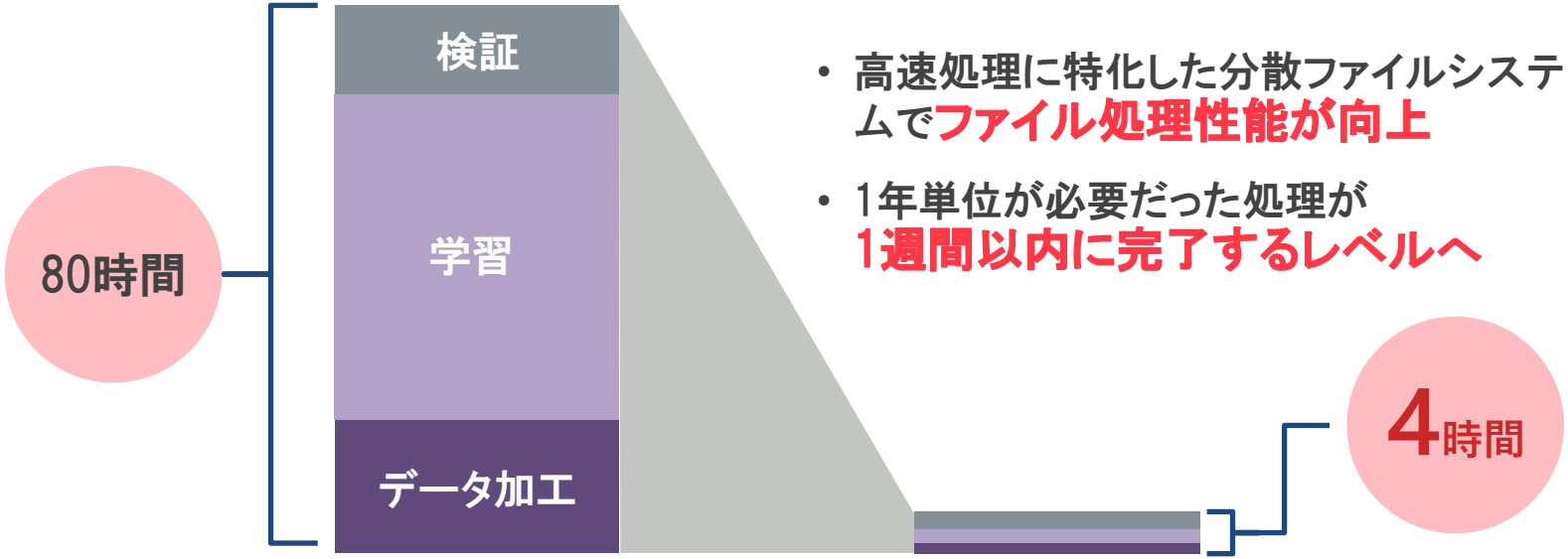
© Hitachi Vantara, Ltd. 2025. All rights reserved.

付録1. Hitachi iQ：適用効果の事例（ディープラーニングへの適用事例）

HITACHI
Inspire the Next

自動車メーカー某社にてディープラーニングの処理が大幅に短縮出来ることが検証され採用

Hitachi iQが提供する分散ファイルシステムの
優れた性能で処理時間が短縮



- 高速処理に特化した分散ファイルシステムで**ファイル処理性能が向上**
- 1年単位が必要だった処理が
1週間以内に完了するレベルへ

自動車メーカー某社でのテスト結果（ディープラーニング）

© Hitachi Vantara, Ltd. 2025. All rights reserved.

HITACHI
Inspire the Next

HA8000V Gen11 AI推論ソリューション向けGPU搭載サーバ

1. エッジでのコンピュータービジョンAI

HA8000V DL320 Gen11
最大2枚のNVIDIA L4 GPU

目的/ワークロード	説明、業種
• コンピュータービジョンとインテリジェントビデオ分析 • 損失の防止 • スマートスペース	• エッジに最適なコンピュータービジョンのリアルタイム推論 • カメラからビデオ映像を分析し、物体/人間/行動を検出して追跡することで、運用/セキュリティ/スマートスペース体験を改善 • 小売業、接客業、製造業に最適

2. ビジュアル生成AI

HA8000V DL380 Gen11
最大3枚のNVIDIA L40 GPU
NVIDIA AI Enterprise

目的/ワークロード	説明、業種
• 3Dアニメーションとレンダリング • ビデオコンテンツの作成 • 画像生成	• キャラクターの動きや、物理現象のシミュレーション、環境エフェクトなど、実物に近い動的な3Dアニメーションを生成 • 現実のデータを模倣した画像やビデオなどの、新しいビジュアルコンテンツを作成 • メディア、エンターテイメント、ヘルスケア、製造業に推奨

3. 自然言語処理AI

HA8000V DL380a*1 Gen11
最大4枚のNVIDIA H100/L40S
NVIDIA AI Enterprise

目的/ワークロード	説明、業種
・自然言語処理(NLP) ・スピーチAI ・不正検出 ・予測メンテナンス	・不正検出やスピーチAIなどのユースケースで高度なNLPが可能 ・製造業、顧客サービス向けに推奨

*1 2024年9月27日販売開始予定

© Hitachi, Ltd. 2025. All rights reserved.

付録2. HA8000V GPU搭載サーバ

GPU商材強化と在庫確保に努めてまいりますので、ご活用をよろしくお願い致します

- **生成AI向けGPU商材強化： DL380a Gen11+H100 NVL (2024/9/27 販売開始)**
- **タイムリーな最新GPU(NVIDIA B100など)サポートとGPU搭載検証機の配備を順次提供予定**

【ワークロード毎の構成例】

社外サイト：https://www.hitachi.co.jp/products/it/ha8000v/products/gpu/index.html

【GPUラインナップ】

凡例:適正ワークロード3段階評価 ◎>○>△

カテゴリー	GPUモデル (アーキテクチャ)	ワークロード			
		Training (学習)	Inference (推論)	Graphics	vGPU (仮想化)
AI	H100 NVL (Hopper)	◎	◎		○
	H100 PCIe (Hopper) *1	◎	◎		○
AI/ Graphics	L40S (Ada Lovelace)	○	○	◎	○
	L40 (Ada Lovelace)	△	△	◎	○
	L4 (Ada Lovelace)		△	○	○
VDI	A16 (Ampere)				○
Edge/ Entry	A2 (Ampere)		△		△

*1 販売一時停止中、個別対応可能

© Hitachi Vantara, Ltd. 2025. All rights reserved.

付録2. HA8000V GPU搭載サーバ構成例

HITACHI
Inspire the Next

HA8000V/DL380 Gen11 + GPU NVIDIA H100 NVL 搭載構成 例

HA8000V/DL380 Gen11 + GPU NVIDIA H100 NVL 1枚構成 例

品名	数量	単価[税抜]	合計[税抜]
HA8000V/DL380 Gen11 8SFF構成 第4,第5世代インテルXeonスケーラブル・プロセッサー搭載モデル,翌日以降(9-17時)保守モデル(3年),OSレスモデル,標準搭載LANレス, プロセッサレス,メモリレス,ハードドライブレス,ラックモデル,200V 1.8m C13-C14電源ケーブル(パワーサプライ数分)標準添付 工場設定 RAID 1+スペア (1 Raidグループ)	1	762,000	762,000
DL3XX Gen11 Easy Install Rail 3 Kit	1	23,000	23,000
INT Xeon-G 6526Y CPU(2.8GHz/16コア/32スレッド)	2	728,000	1,456,000
64GiB 2Rx4 PC5-5600B-R Smart Kit	4	796,000	3,184,000
BCM 5719 1Gb 4p BASE-T OCP Adptr	1	85,000	85,000
1800W-2200W FS Ti Ht Plg PS Kit	2	175,000	350,000
MR408i-o Gen11 SPDM Storage Cntlr	1	394,000	394,000
1.92TB SATA MU SFF BC MV SSD	3	422,000	1,266,000
NVIDIA H100 NVL 94GB PCIe Accelerator【増設用】	1	16,890,000	16,890,000
DL380/DL560 Gen11 2U GPU Pwr Cbl Kit【増設用】	1	12,000	12,000
9.5mm SATA DVD-ROMドライブ	1	14,000	14,000
Smartストレージバッテリー 96W 145mm	1	19,000	19,000
DL380 G11 2U x8/x16/x8 Sec Riser Kit	1	37,000	37,000
DL380 Gen11 2U 8SFF x1 TM Kit	1	87,000	87,000
DL380/DL560 G11 High Perf 2U HS Kit	2	32,000	64,000
DL380/DL560 G11 2U High Perf Fan Kit	1	134,000	134,000
DL360 Gen11 Stg Cntrl Enable Cbl Kit	1	6,000	6,000
DL380 Gen11 SFF Univ Media Bay Kit	1	17,000	17,000
DL360 Gen11 CPU1/OCP2 x8 Enable Kit	1	17,000	17,000
iLO Advanced 1サーバーライセンス【工場搭載】	1	65,000	65,000
		ハードウェア合計[税抜]	24,882,000
		ハードウェア合計[税込(10%)]	27,370,200

希望小売価格　　：¥ 24,882,000 (税抜)

© Hitachi Vantara, Ltd. 2025. All rights reserved.

付録3. VSP One製品ラインナップ

VSP One ストレージ： ハードウェア高圧縮による圧縮率向上・パフォーマンス向上
高速IFサポートによるパフォーマンス向上
省電力、省資源化、メディアのドライブ増設によりDC運用コスト低減に寄与

ハイブリッドクラウド ： 災害対策・データ利活用を可能とするリモートコピー機能をサポート
VSP One ストレージは'24/4Q以降に連携可能とする予定です

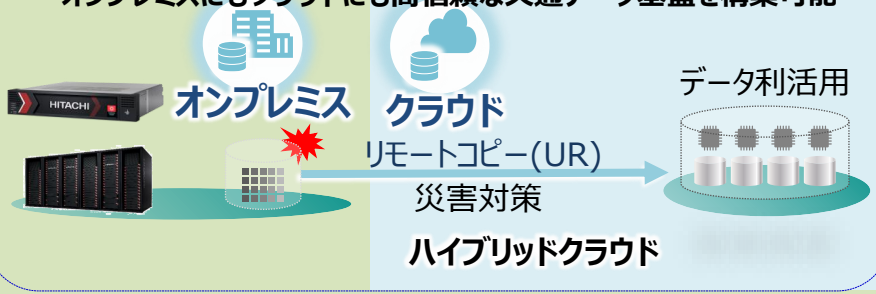
VSP One ストレージ（ミッドレンジ）

New

VSP One 2U Block Appliance
(2Uラックマウント型オールフラッシュアレイ)

VSP One B28　VSP One B26　VSP One B23

オンプレミスにもクラウドにも高信頼な共通データ基盤を構築可能



VSP F/Gシリーズ

VSP G150
VSP G130
VSP G350
VSP F350

VSP Eシリーズ

VSP E390(H)
VSP E790(H)
VSP E590(H)

VSP E1090(H)

ハイエンドストレージ

VSP 5200(H)　VSP 5600(H)

VSP One : Hitachi Virtual Storage Platform One

© Hitachi, Ltd. 2025. All rights reserved.

HITACHI
Inspire the Next

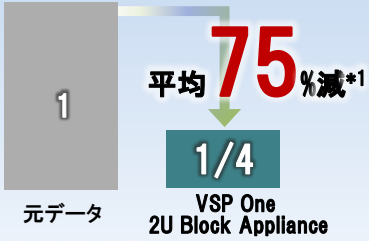
NEW

Hitachi Virtual Storage Platform One 2U Block Appliance

効率性

性能影響を抑えたデータ圧縮で
データ量の増大と分析・活用ニーズに応える

✓ データ削減率



✓ コスト効率(価格)

—20%(SPL比)*2

信頼性

システムの安定稼働を支え
重要データを守る

✓ データ保護機能強化

最大1024世代スナップショット ＋ 変更・削除を許さないプロテクション

✓ ドライブ障害影響低減

データ復旧時間最大 **1/3** 短縮

環境配慮

省電力・省資源化で
サステナブルな社会の実現に貢献

✓ 製品のライフサイクル全体で環境負荷を低減

旧機種比
最大
—66%*3
消費電力量

サーキュラーエコノミーに向けた省資源化

*1 圧縮アクセラレータ搭載モデルVSP One B28,B26での、圧縮・重複排除有効時のデータ削減率。 VSP One B23ではデータ削減率は異なります。対象となるデータの種類により削減率は変動するため、記載の値を保証するものではありません。
*2 VSP G700(100TB)とVSP One B28(100TB)を比較。
*3 VSP E790 (販売開始時仕様) と VSP One B28 の1TBあたり消費電力 (NVMe AFA最大構成時) を比較。
引用:令和3年版「環境・循環型社会・生物多様性白書」(環境省)

© Hitachi, Ltd. 2025. All rights reserved.

Hitachi Social Innovation is
POWERING GOOD