PCクラスタワークショップin 柏2025

ハイブリッドクラウドとマルチクラウド対応
次世代AI HPCプラットフォームのコスト最適化プロビジョニング

2025年6月27日

コアマイクロシステムズ株式会社
ストレージサイエンスグループ　大川　剛

目次

コアマイクロシステムズの会社概要

会社概要：コアマイクロシステムズ株式会社

社名	コアマイクロシステムズ株式会社
	Core Micro Systems, Inc.
Website	https://www.cmsinc.co.jp
本社	〒174-0041
	東京都板橋区舟渡1-12-11 ヘリオスIIビル3F
戸田事業所	〒335-0034
	埼玉県戸田市笹目7-16-1 リバーサイド共栄
代表者	高橋晶三
設立	1992年(平成4年) 10月16日
事業内容	業界ソリューション事業
	ストレージサイエンス事業
	HPCサイエンス事業
	Cloudサイエンス事業
	システムプロダクツ事業

納入実績	1.産業計測マーケット向け　100PB以上/年間
	2.学術スパコンマーケット向け 10PB～80PB/年間
	3.その他ビッグデータマーケット 3PB以上/年間
	①エンタープライズ
	②メディアエンターテインメント
取得認証	・QMS: 品質マネジメントシステム
	ISO 9001：2015
	・EMS: 環境マネジメントシステム
	ISO 14001：2015
	・ISMS: 情報セキュリティマネジメントシステム
	ISO/IEC 27001: 2013
	・BCMS: 事業継続マネジメントシステム
	ISO 22301：2019 /JIS Q 22301：2020
	・PMS: 個人情報保護マネジメントシステム
	JIS Q 15001: 2017

コアマイクロシステムズ株式会社の5事業について

業界ソリューション事業
プロフェッショナル業界特有の特性に適する
バーティカル/プロフェッショナルマーケット向けの先進的なソリューションの先駆的展開

ストレージサイエンス事業

用途アプリケーション要求における価値特性(機能/性能/価値)極大化を可能にする特性ストレージ(プラットフォーム/HCIを含む)

HPC サイエンス事業

次世代HPCアーキテクチャであるコンポーザブルHCI化による価値特性(機能/性能/価値)極大化と柔軟な拡張機動性を可能にするハイブリッドクラウド化HPCプラットフォームの展開

Cloud サイエンス事業

パブリッククラウドの機動的利用合理性とエッジクラウド化のトレンド合理性に対応するビッグデータ対応ハイブリッドクラウドサービス展開

システムプロダクツ事業
製品展開/ソリューション展開/サービス展開のための自社ハードウェア/ソフトウェア開発と
協業パートナー技術/製品ベースの応用開発、ならびにBTO/CTO/OEM展開

現行のHPCエンジニアリングセンタのご紹介

技術の中核となる大規模環境
戸田事業所エンジニアリングセンタ

コアマイクロシステムズはAI/HPCおよびストレージを始めとした
あらゆる要件に対応するために他では類をみない大型の検証環境
「エンジニアリングセンタ」を整備し、最新鋭・最先端の機器および設備を導入して、
基礎開発研究、検証、評価およびベンチマーク、各種動作確認を行っています。
全23基の19インチラック群を全て自社の検証環境として使用し、
お客様要求に応じて柔軟にシステムを構成して
多種多様な用途や要求に応じた検証に対応します。


ラックマウント状態のサーバ機器群


高密度HPC CPUノード


高密度HPC GPUノード群


液浸冷却ラック 24U


エンタープライズテープライブラリ


液浸マイクロデータセンター


水冷式リアドア型冷却ユニット


水冷式冷却ユニット用 大型チラー

最新・最先端の機器で構成したエンジニアリング設備を整備

■最新のAI/HPC計算機を多数CPU/GPUノードで構成
■超高速/大規模に対応する最新のAI/HPCストレージ設備
■複数種類の液浸冷却装置や水冷式Active型リアドア式熱交換器、
DLC設備(CDU～CDM)など最先端の高効率冷却環境の整備

エンジニアリングセンタ検証環境用途

- HPC／Storage基礎開発検証環境
- HPC／Storage基礎ベンチマーク検証環境
- 大型HPC案件対応システムレベルベンチマーク検証環境
- BIOS／ファームウエア／OS／ファイルシステム ミドルウェアのアップデート検証環境
- システムレベル保守運用検証環境
- 案件固有保守運用検証環境
- カスタムプロジェクト開発対応システム検証環境

環境区分

- 高密度HPC CPUノード群
- 高密度HPC GPUノード群
- 高密度AI／HPC CPU／GPU／メモリ／NVMe コンポーザブルノード群
- 超高速モジュラースケールNVMe SSD Arrayストレージ群
- 高速大規模ラックスケールHybrid Arrayストレージ群
- 高速大規模ラックスケールAll HDD Arrayストレージ群
- 大規模ラックスケールオブジェクトストレージ群
- 大規模ラックスケールテープライブラリストレージ群
- 案件保守運用検証ラック群
- カスタムプロジェクト開発ラック群
- HPC IBインターコネクトスイッチ群
- HPC 100GbEネットワークスイッチ群

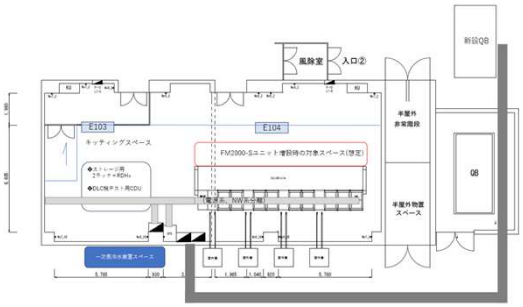
環境デバイス規模

- CPU(Intel/AMD)構成　100ノード以上／10000コア以上
- GPU(NVIDIA/AMD)構成　30ノード以上／50個以上
- NVMe SSD(Gen4/Gen5)構成　500個以上
- SAS SSD(12G/24G)構成　100個以上
- SAS HDD(12G)構成　3000本以上
- SATA HDD(6G)構成　1000本以上
- LTOライブラリ構成　1PB以上
- PCIe/Gen4ファブリックスイッチ構成　16基
- IB/HDR/NDRインタコネクトスイッチ構成　48基
- Rockport 300G スイッチレスダイレクトインタコネクト構成　1式
- 100/200/400GbEネットワークスイッチ構成　8基

環境設備規模

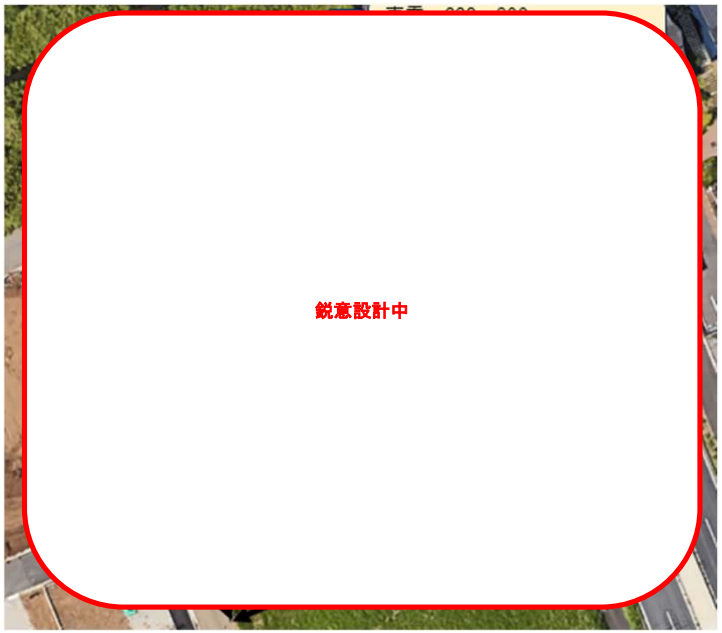
- フロア面積 約572m2
- 19インチ 41Uラック　23基
- 水冷式リアドア型冷却ユニット　3台
- 液浸冷却ラック(24U)　1台
- 水冷式冷却ユニット用 大型水槽付DCインバータチラー　3基
- 電源設備 450KVA

新つくば事業所2拠点誕生予定



環境設備規模

■19インチラック　42U　＊8ラック
■電源設備　1750kVA(最大拡張時)



2025年夏　新つくば事業所　誕生予定
茨城県つくば市に新エンジニアリングセンタを準備中
AI/HPCの開発・検証・評価環境がさらに充実

つくば市東光台

2026年春　新つくば事業所　誕生予定
茨城県つくば市に新生産拠点を準備中
AI/HPCの生産・構築・運用体制さらに充実

つくば市学園西

コアマイクロシステムズのAI HPCビジネス展開について

次世代AI HPCワークロードに求められる革新性

現状の課題例	
インフラ調達の柔軟性	購入サイクルが長く、需要変動に対応しにくい
運用負荷	ソフトウェア環境構築・依存解決に時間がかかる
GPU共有の難しさ	資源の取り合い。公平性・効率性の確保が難しい
クラウド連携の難易度	コスト見積・認証設定・データ転送が煩雑

AI HPC市場における潮流	
ハイブリッド／マルチクラウド化の加速	高負荷時だけクラウドを弾力的に利用（バースト）
クラウドネイティブ技術の導入	Kubernetes, Slurm + API連携など
ノーコード/ローコードの自動化	技術知識がなくてもパイプライン構築が可能
コンプライアンスとガバナンス	予算管理、データ主権への配慮

- ハイブリッドクラウド・マルチクラウドを統合的に使える柔軟性
- ユーザー自身が容易にワークフローを構築・共有できる簡便性
- ガバナンスやコスト制御を前提とした設計
- セキュアかつ再現性の高い実行環境

コアマイクロシステムズとParallel Works戦略的パートナー契約の締結

コアマイクロシステムズ社内の
AI HPC開発評価環境の新規構築

コアマイクロシステムズのハイブリッド・マルチクラウド対応
次世代AI HPCプラットフォームの展開

ハイブリッド・マルチクラウド・コンピューティング・リソースのためのACTIVATEコントロール・プレーンの
プロバイダーである企業のParallel Works社と戦略的パートナー契約を締結

コアマイクロシステムズの革新的なAI HPCワークロードビジネスの展開

1.マイクロデータセンタ構造のAI HPCアプライアンスベースのハイブリッド/マルチクラウド
AI HPCプラットフォームビジネスの展開

2.モジューラデータセンタ構造のAI HPCアプライアンスベースのエッジ/グリッド/ハイブリッド/
マルチクラウドAI HPCプラットフォームビジネスの展開

3.アカデミック及びデータセンタ向けサービスの展開

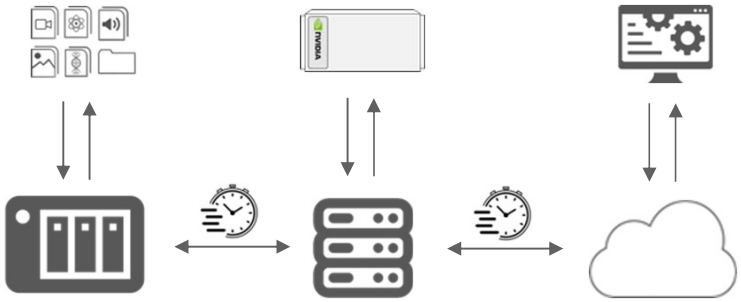
4.ハイパースケーラークラウド、国内クラウド、学術クラウドとのパートナー展開の推進

次世代AI HPCプラットフォームを革新するコントロールプレーンご紹介

次世代HPC・AI ワークロードに求められる課題

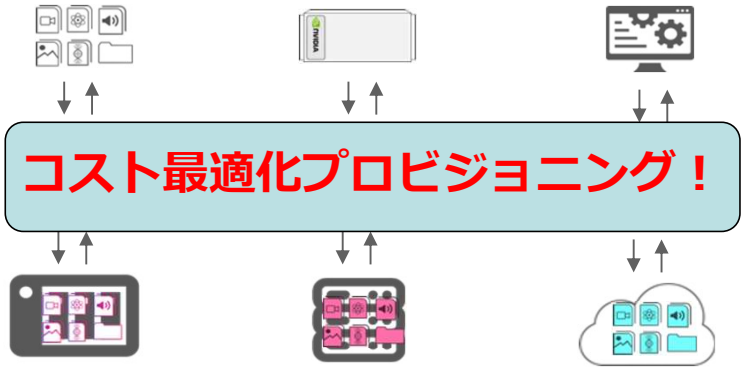
これまで

- ワークフロー作成と管理が複雑
- クラウド／オンプレのリソース統合が困難
- スケーリングと自動化に課題
- データの可視性や追跡性が乏しい



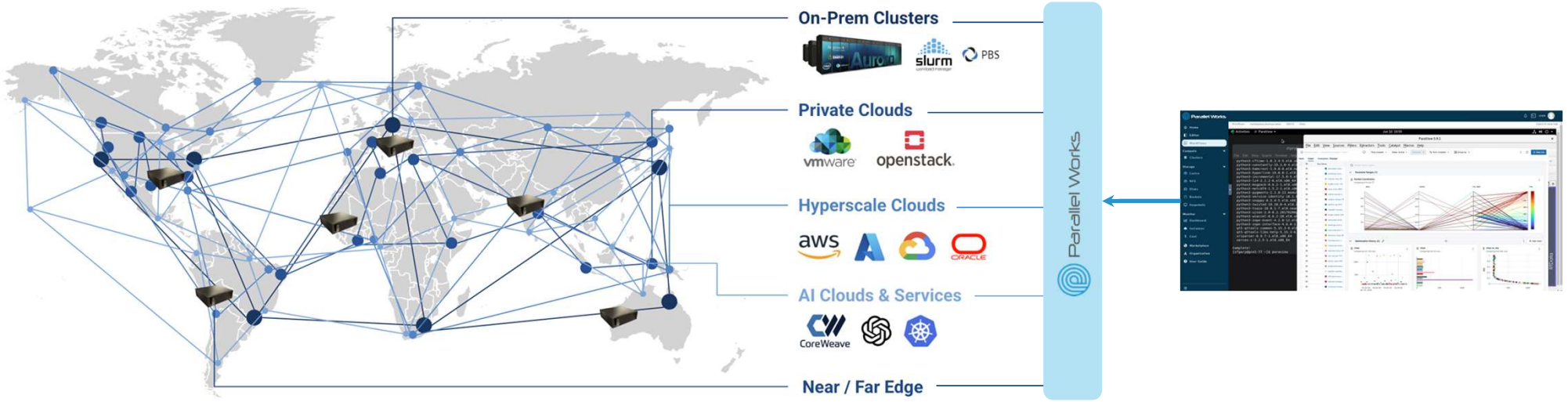
現在～これから

- HPCワークフローをノーコード/ローコードで
- ハイブリッド/マルチクラウド/オンプレ一元管理
- 高度なスケジューラ/ジョブ管理/再現性実行
- チーム開発・共有が容易なコラボレーション



ハイブリッドクラウドとマルチクラウド対応型次世代AI HPCプラットフォーム統合

Parallel Works社ACTIVATEによるハイブリッド・マルチクラウドコンピューティングの革新的なコントロールプレーン

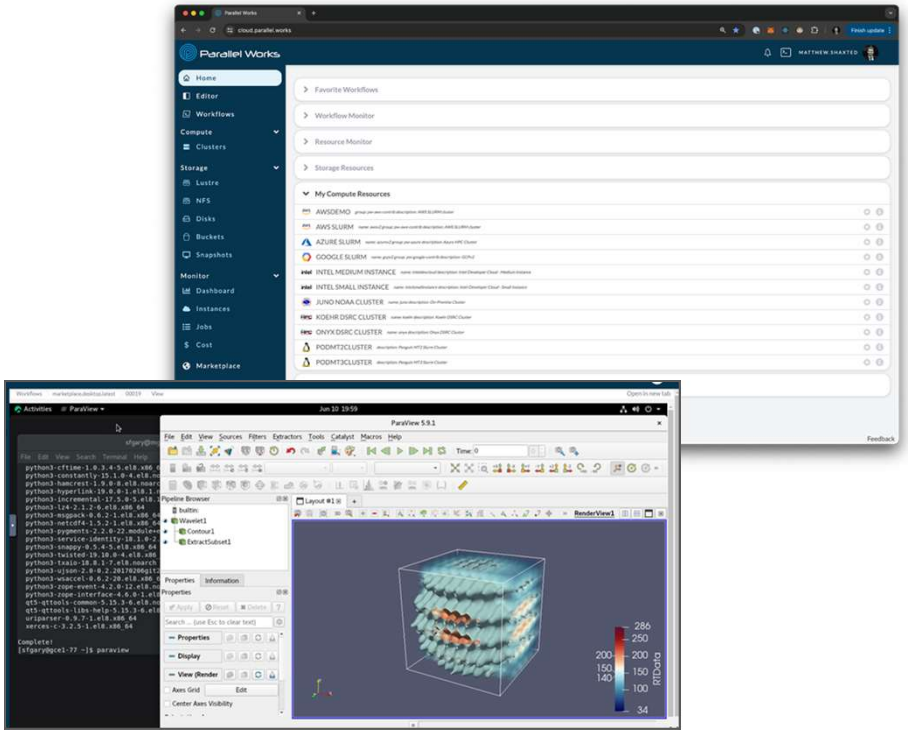


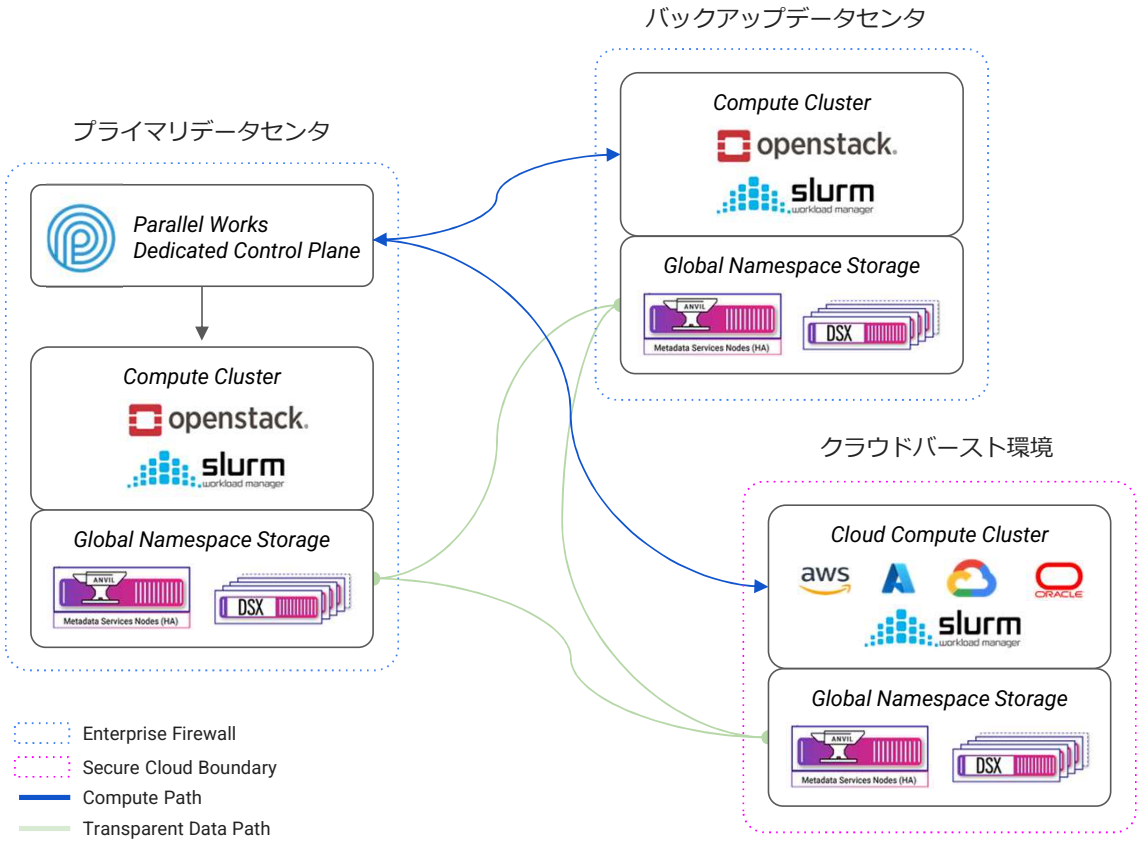
- 多様で遠く離れたコンピューティング・インフラを単一の合理化されたインターフェイスに統合し、規模に応じた制御、コラボレーション、コスト・ガバナンスを可能にします。計算負荷の高いアプリケーションやAIモデルに依存するチームが、シミュレーション、アナリティクス、AI向けに、オンプレミス、クラウド、ハイブリッドのハイパフォーマンス・コンピューティング・リソースを容易かつ一貫してプロビジョニング、管理、共有できるようにします。

Parallel Works

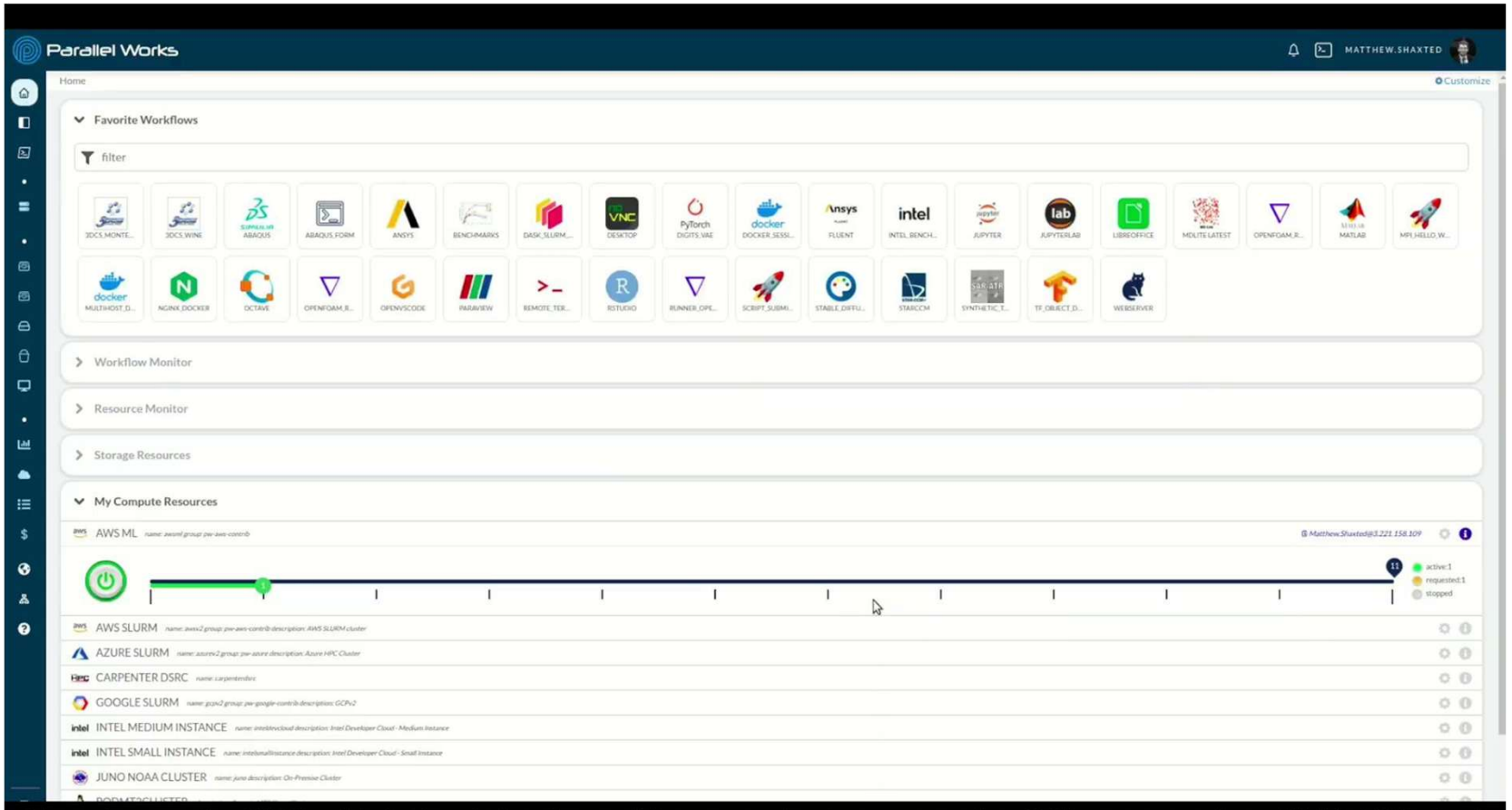
革新的なハイブリッド・マルチクラウドコンピューティングインフラオーケストレーション

Parallel Works社ACTIVATEによるハイブリッド・マルチクラウドコンピューティングの革新的なコントロールプレーン





複雑なHPC AIスタックをクラウドやオンプレ上に容易にデプロイ

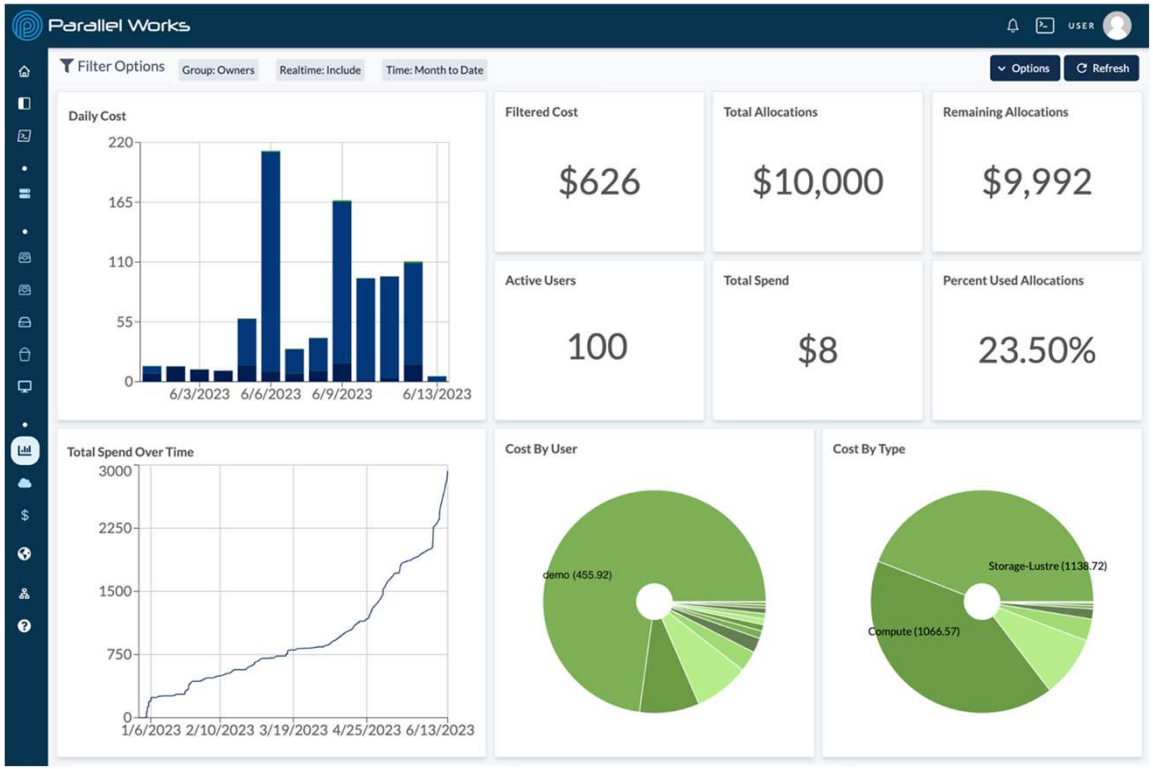


Execute complex ML/GenAI stacks via turnkey webforms on cloud or on-premise GPU resources (e.g., Stable Diffusion).

ほぼリアルタイムのクラウド費用管理が可能

Parallel Works ACTIVATE provides:

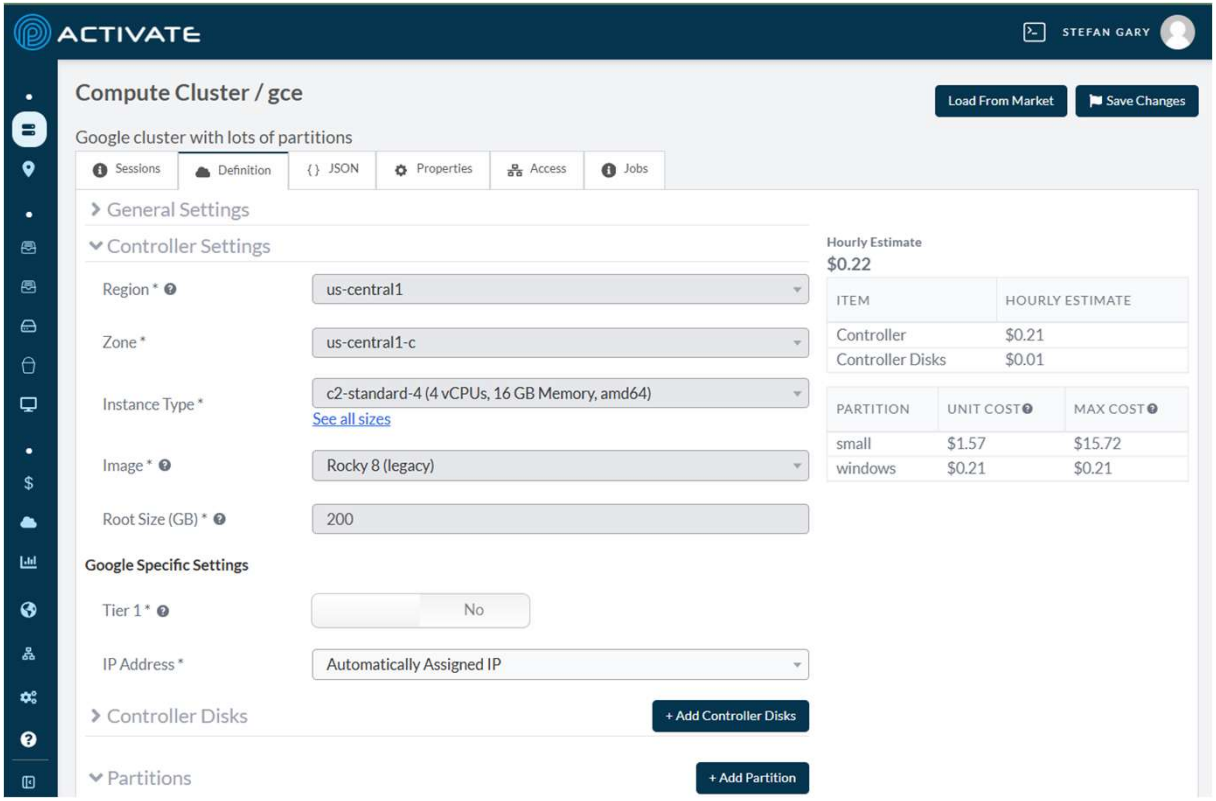
- ほぼリアルタイムのクラウド費用見積り CSPからの一定時間後の請求書と同期するFinOpsツール（3分毎）
- カスタマイズ可能な閾値の強制アクションを備えた管理ポータルを介した、グループベースおよび個別の予算割り当て



複雑なHPC AIスタックをクラウドやオンプレ上に容易にデプロイ

Parallel Works ACTIVATE offers

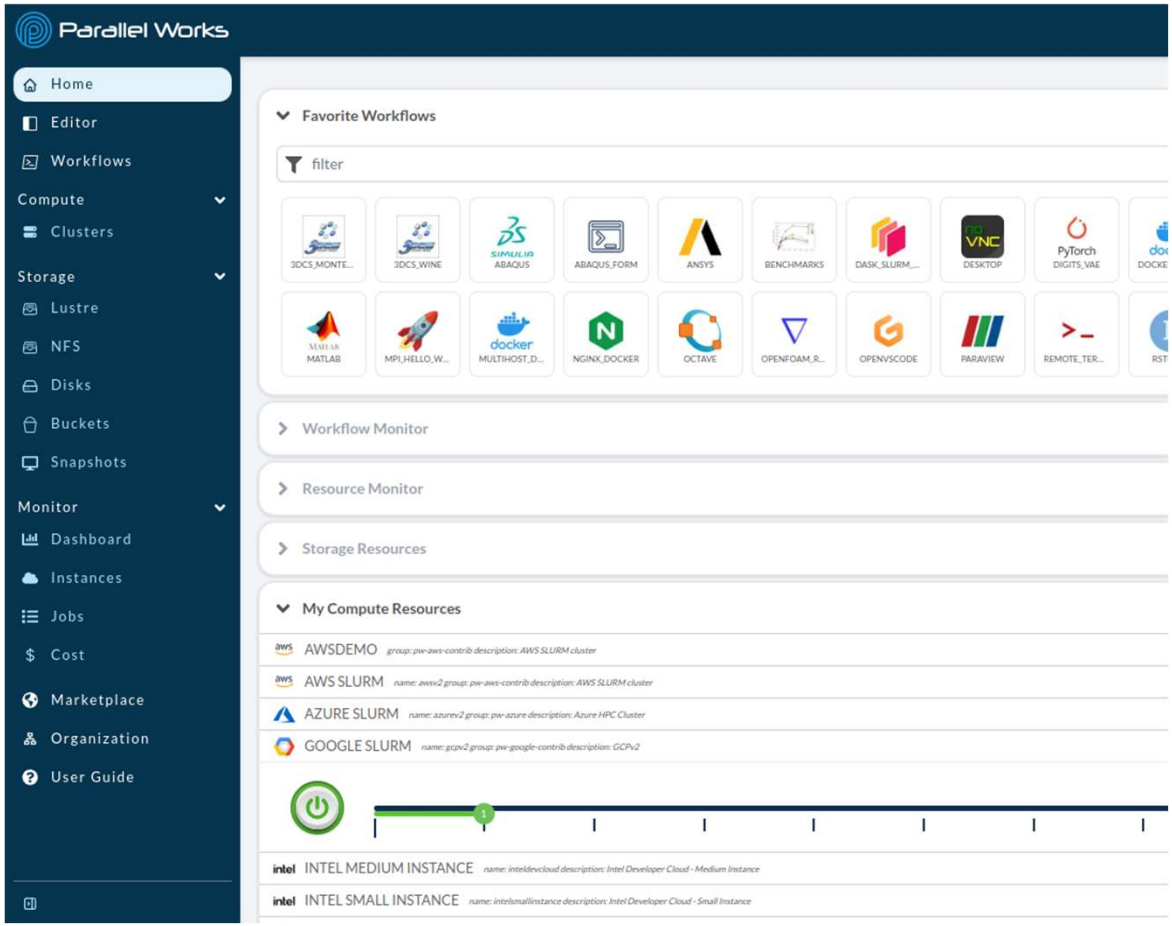
- コードとしてのインフラストラクチャーは、コンピュートとストレージの要素を作成する
- コードとしての環境は、ワークフローと対話型セッションを展開する



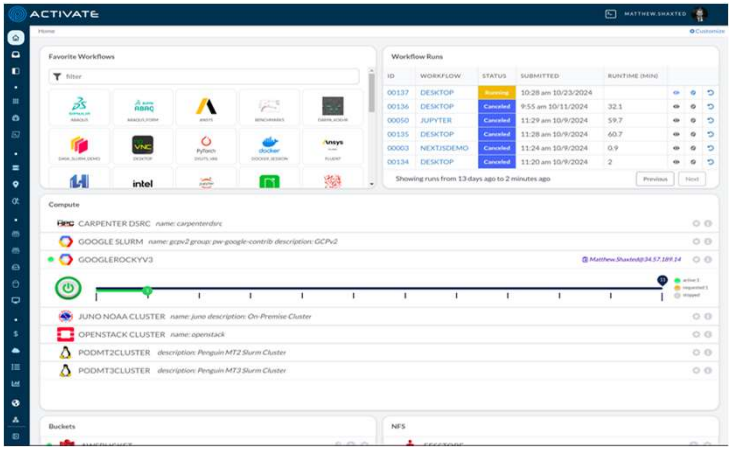
複雑なHPC AIスタックをクラウドやオンプレ上に容易にデプロイ

Parallel Works ACTIVATE offers

- 単一のコントロールプレーンから様々なロケーション（ハイブリッドとマルチクラウド）にまたがる多様なリソースタイプを容易に管理するためのスケジューラーレイヤーでの抽象化
- 複数のクラウドとリソースタイプへの単一のクレデンシャルアクセス



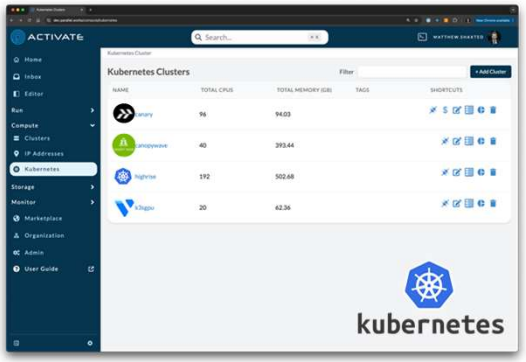
複雑なHPC AIスタックをクラウドやオンプレ上に容易にデプロイ



ユーザーアクセスとセキュリティの管理	短期トークンとロールベースの権限を提供し、新しいユーザーがコンソールに直接アクセスしなくてもクラウド リソースにアクセスできるようにします。
予算の配分と管理	3 分単位の精度でコストを追跡するほぼリアルタイムの FinOps システムを使用して、正確で信頼性の高い予算執行による個人およびグループベースの予算配分を可能にします。
クラウドインフラストラクチャのプロビジョニング	Infrastructure as Code は、コンピューティング、ストレージ、ネットワーク リソースの展開を自動化し、実務者に馴染みのある外観と操作性を備えた調整されたクラウド HPC インスタンスのセットアップとスケーリングを簡素化します。
自動化と一貫性	Environment as Code を使用してワークフローと環境を標準化し、チーム間で一貫性のある繰り返し可能な展開を保証して、イノベーションとコラボレーションを促進します。
コンプライアンスとセキュリティ	独自のセキュリティ ニーズに合わせてカスタマイズ可能な展開など、さまざまなセキュリティオプションをサポートしています。

ACTIVATE AIの導入：規模に応じたAIリソース導入の運用化

ACTIVATEコントロールプレーンには**3**つの新機能が追加され、**AI**とコンピュート環境全体で複雑さを増している組織向けに設計されている：



Kubernetes サポート

Amazon SageMaker
Machine Learning Workspace
OpenAI

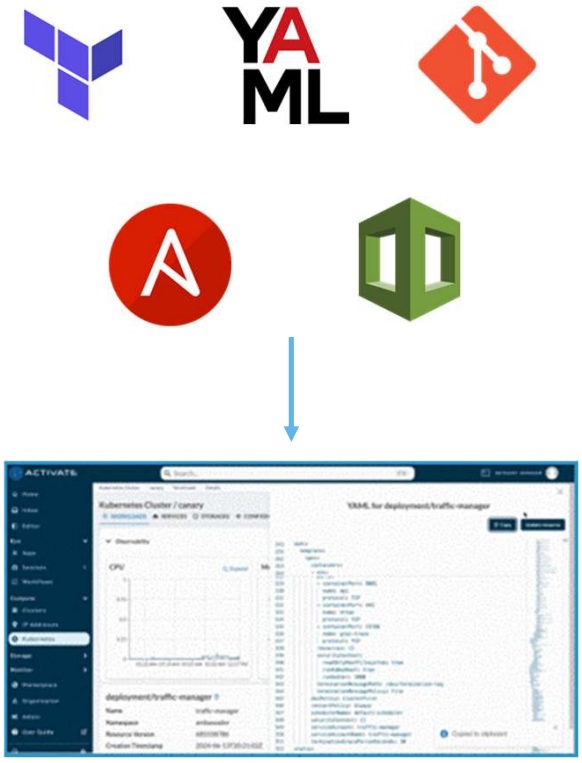
AI リソースの統合

VULTR
CANOPY WAVE
VOLTAGE PARK

NeoCloud サポート

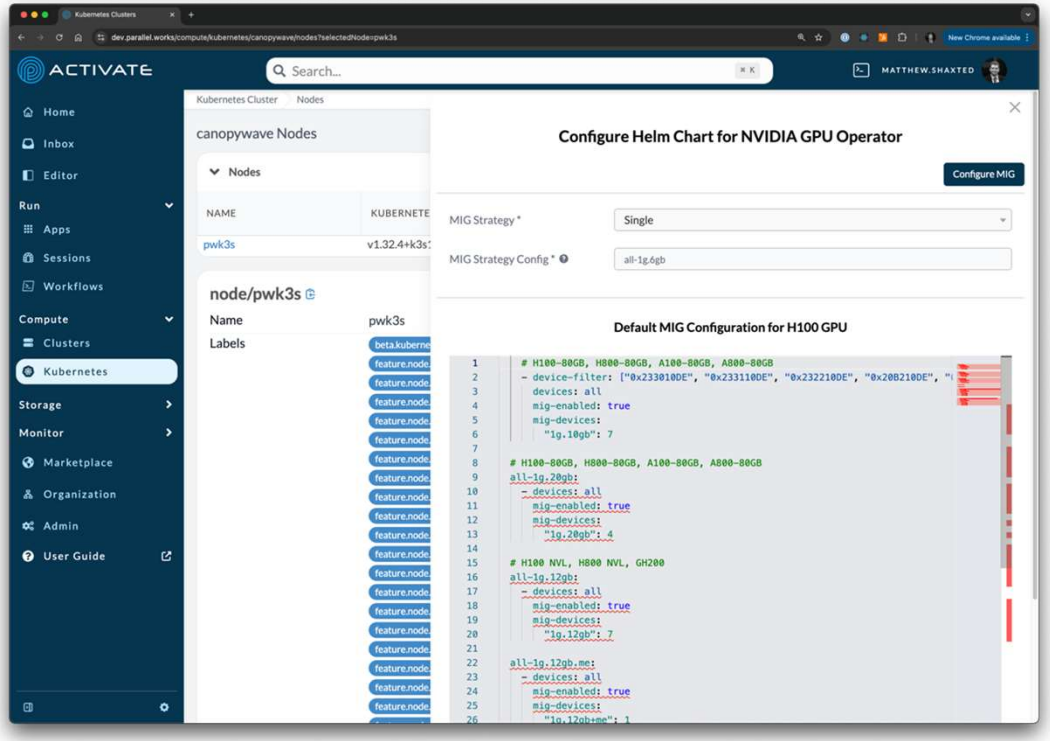
インフラチームとエンド向けのKubernetesサポートを簡素化する

- Kubernetes は、現代のエンタープライズ、AI、ML ワークロードを実行するために不可欠になりつつありますが、使用するには専門知識が必要になることがよくあります。
- 従来、ユーザーは YAML 構成を記述し、それを手動でサーバーに適用し、複雑な CLI ツールを使用してデプロイメントを管理する必要がありました。
- ACTIVATEによりわずか数回のクリックで、組織はリモートデスクトップ、HPC、仮想化ワークロードの実行に既に使用している環境にKubernetesクラスターを接続できます。



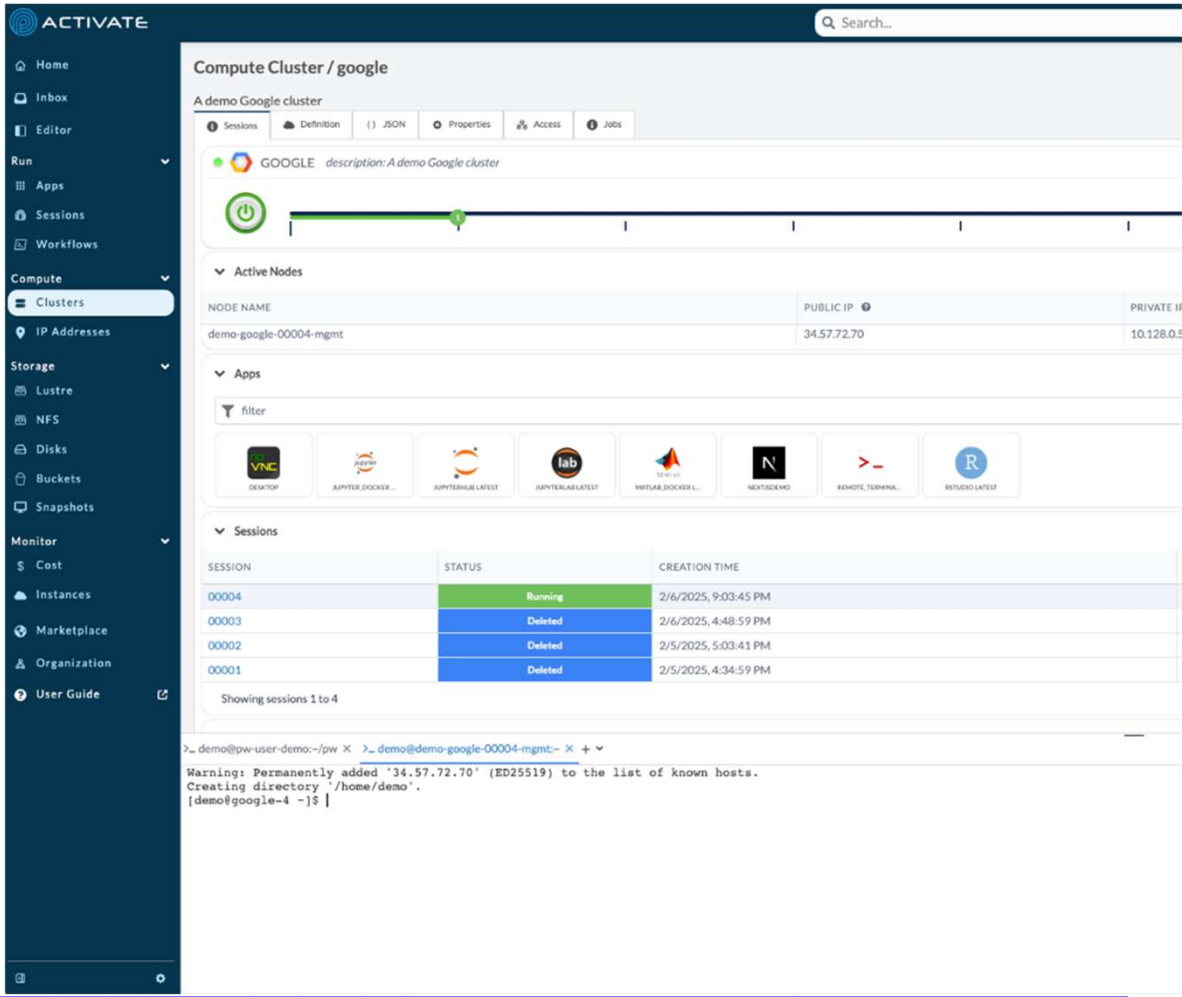
GPUの動的分割:MIG または Juice GPU-over-IP サポート

- マルチインスタンス GPU (MIG) は、GPU リソースを分割するための安全なハードウェア レベルのパーティションを提供します。*Compatible GPUs listed in NVIDIA MIG guide*.
- MIG 部分をクラスター ノードに動的に適用し、ユーザー グループに割り当てることができるようになりました。
- MIG 非対応 GPU の場合、ワークフローに対してシームレスな Juice GPU-over-IP サポートも可能です。



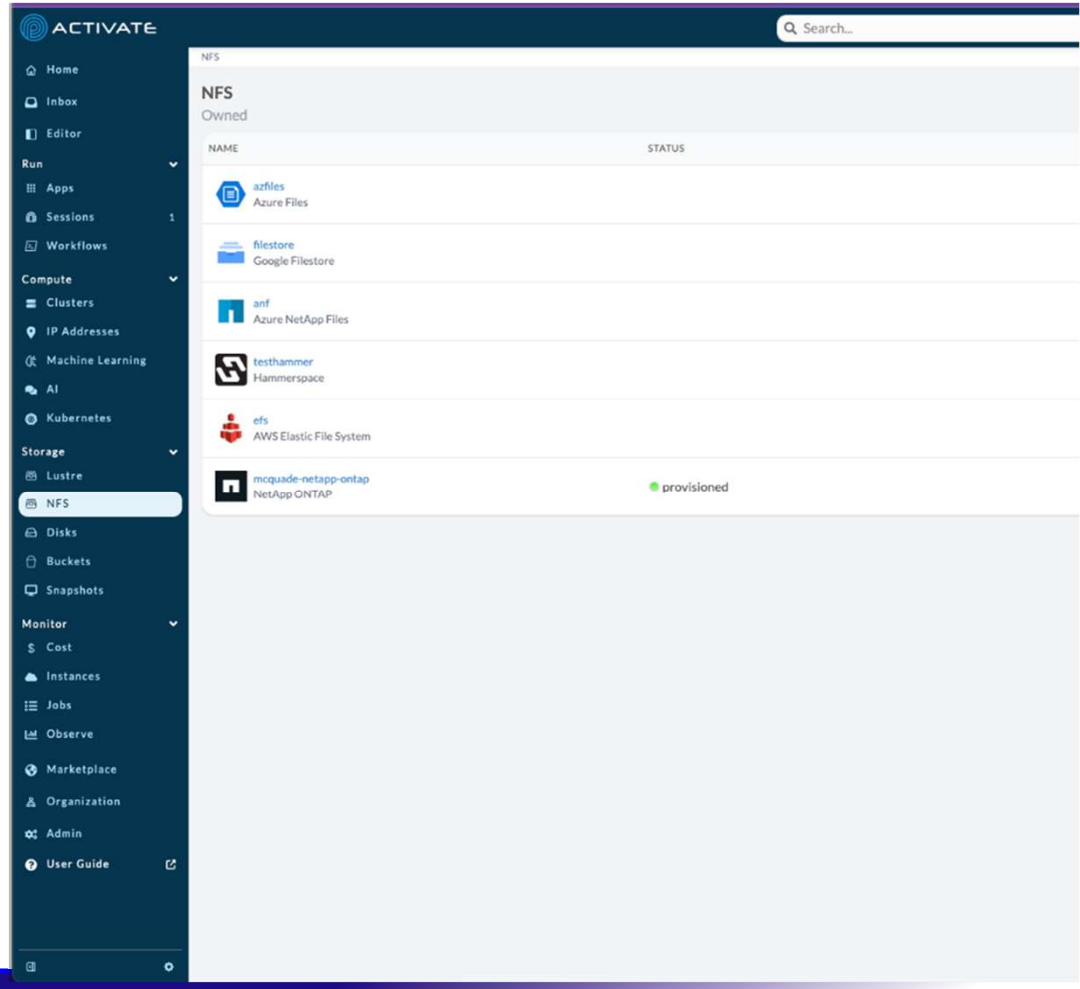
ACTIVATE OpenStack：ワンクリック展開

- ワンクリック展開バッチ HPC スケジューラ (Slurm、PBS、LSF)、直接 VM、または Kubernetes クラスターで構成されたクラスターを即座に起動します。
- 柔軟な共有プロビジョニングされたクラスターをチームまたは組織全体と簡単に共有できます。



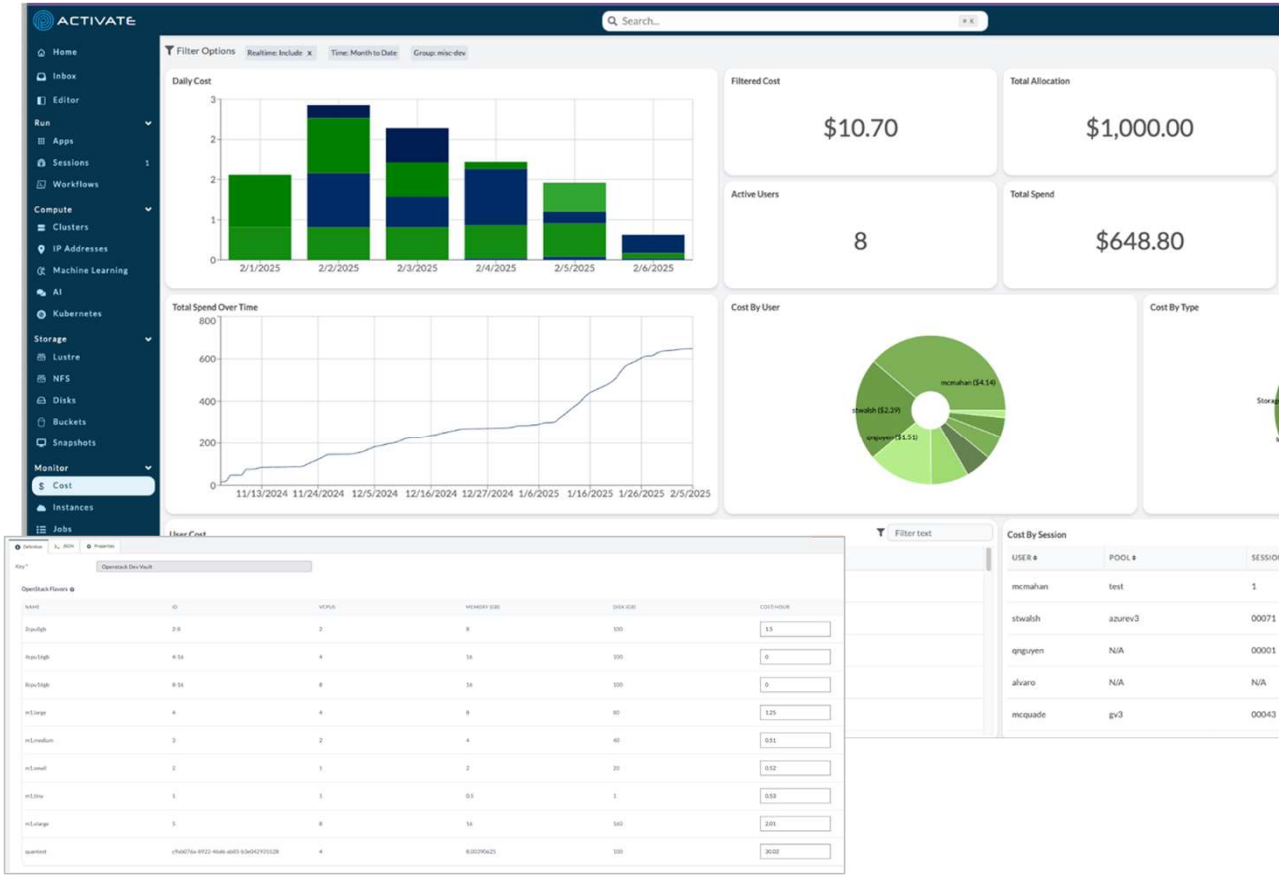
ACTIVATE OpenStack：さまざまなストレージプロビジョニング

- さまざまなストレージ タイプをシームレスにプロビジョニングおよびマウントし、データの永続性とパフォーマンスを確保します。



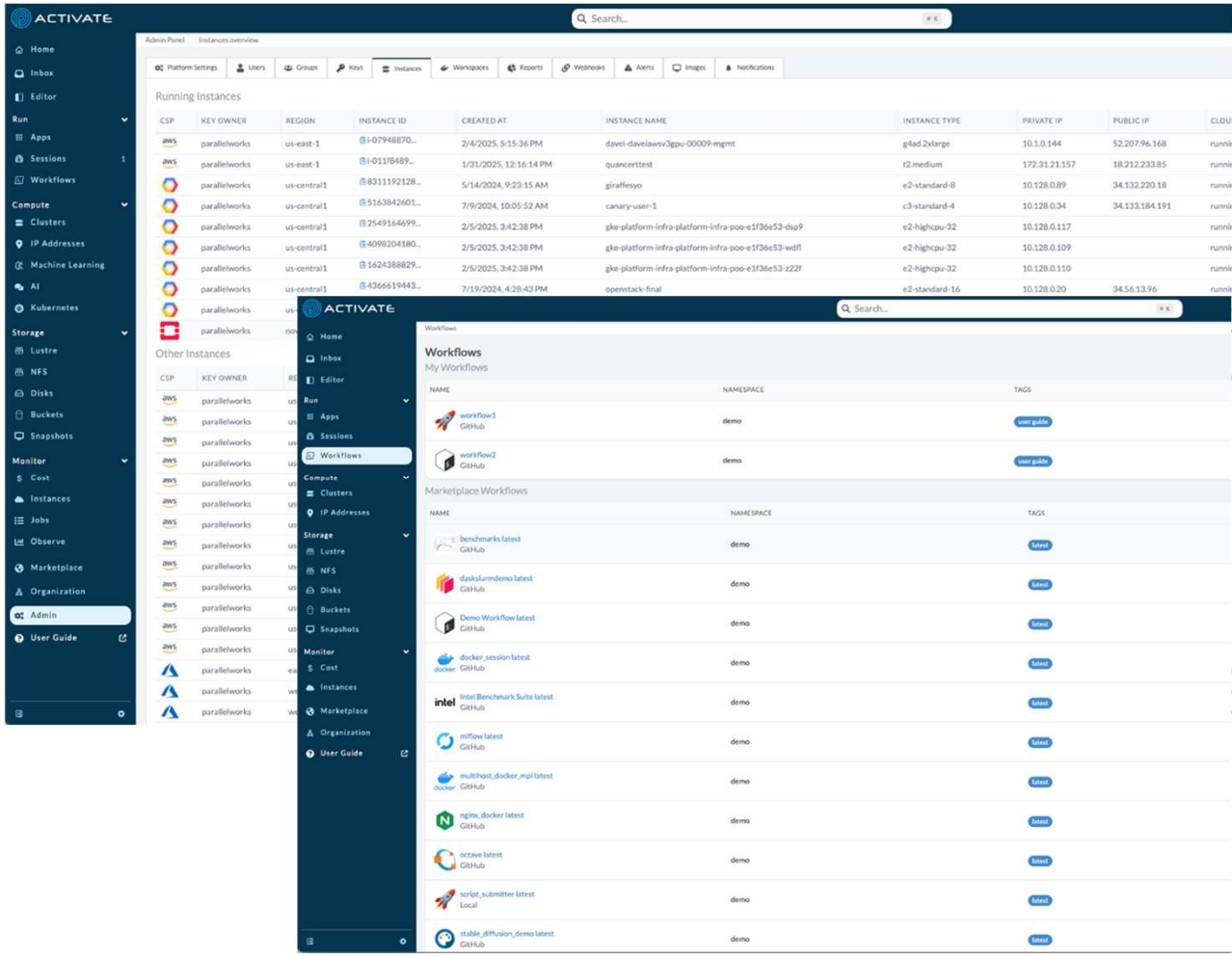
ACTIVATE OpenStack :動的なコスト管理と使用状況の追跡

- さまざまなストレージ タイプをシームレスにプロビジョニングおよびマウントし、データの永続性とパフォーマンスを確保します。



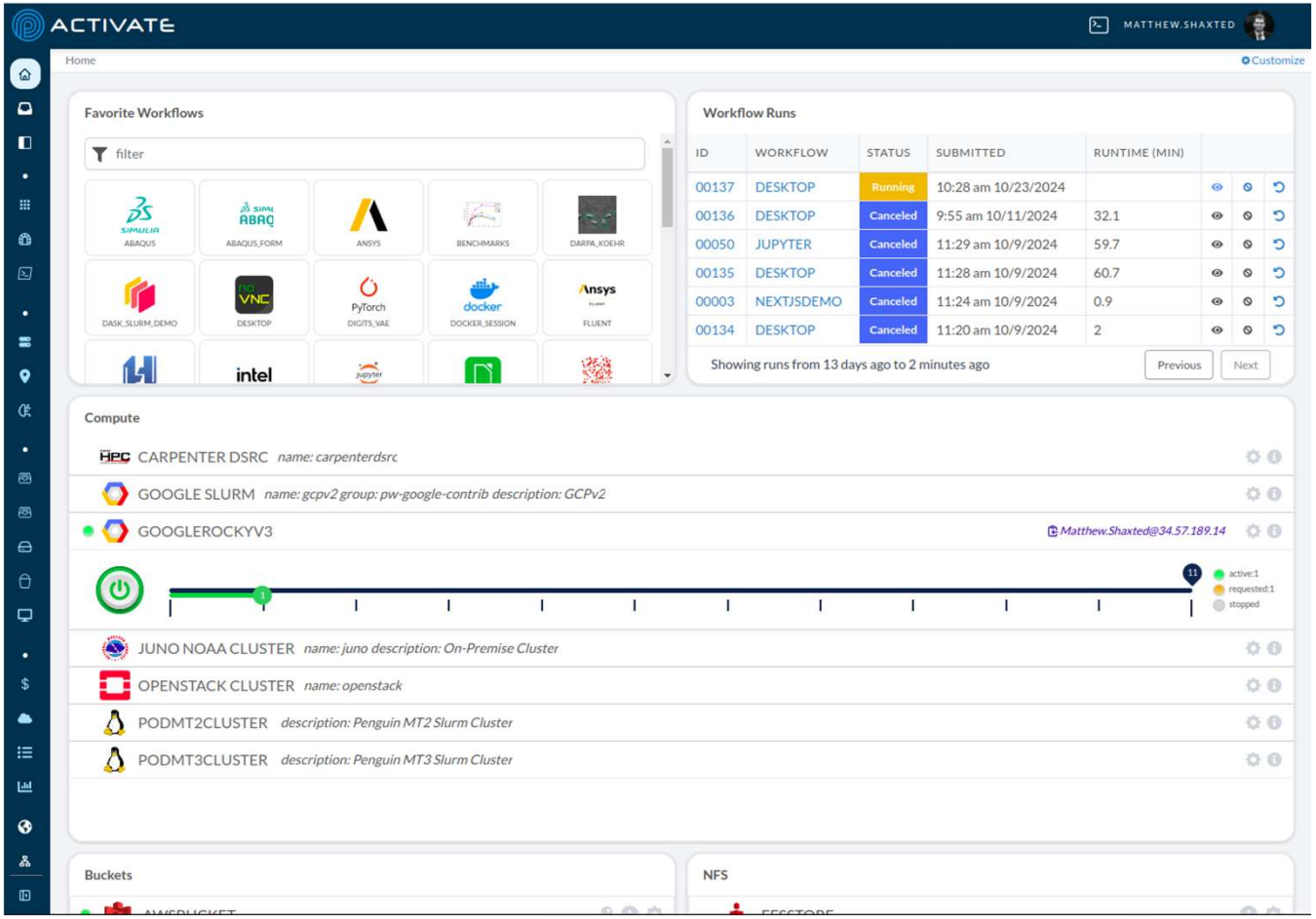
ACTIVATE OpenStack :革新のエンドユーザエクスペリエンス

- 統合アクセスポータル:
  すべてのユーザーが単一の直感的なインターフェースを使用して OpenStack ポータルに直接アクセスする必要がなくなります。
- インタラクティブなワークロード:
  ワンクリックで Jupyterノートブック、リモート デスクトップ、RStudio セッションなどを起動し、迅速な実験と生産性を高める環境を構築します。



まとめ：

コアマイクロシステムズのAI HPCプラットフォームにParallel Works ACTIVATEを組み合わせることによって、次世代AI HPCワークロードのハイブリッド/マルチクラウド環境をコスト最適化プロビジョニングさせることが容易に達成できます。

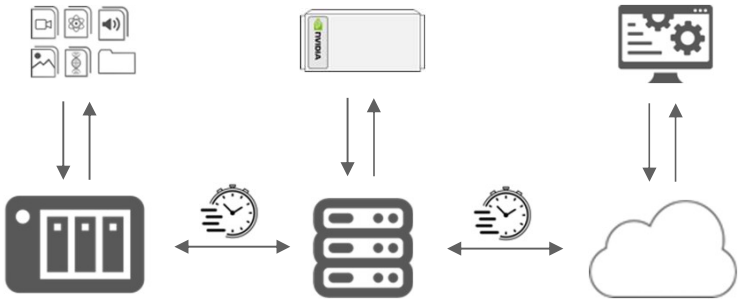


広域AI HPCリソースのオーケストレーションご紹介

広域分散＆次世代HPC・AI ワークロードに求められる課題

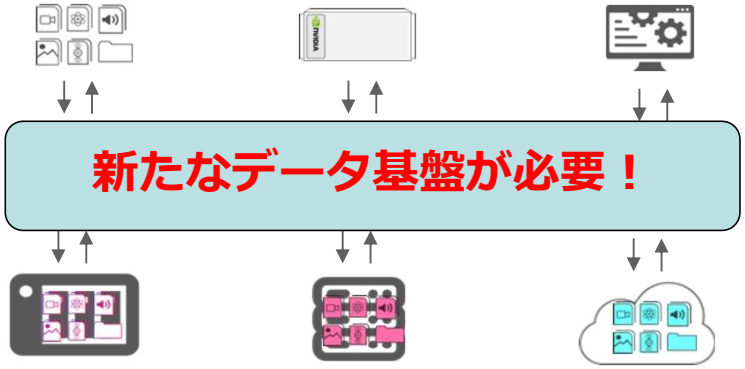
これまで

- サイロ化が進む非構造化データ
- パフォーマンスとリソースの限界
- クラウド利用における俊敏性の欠如
- 手作業によるデータ管理

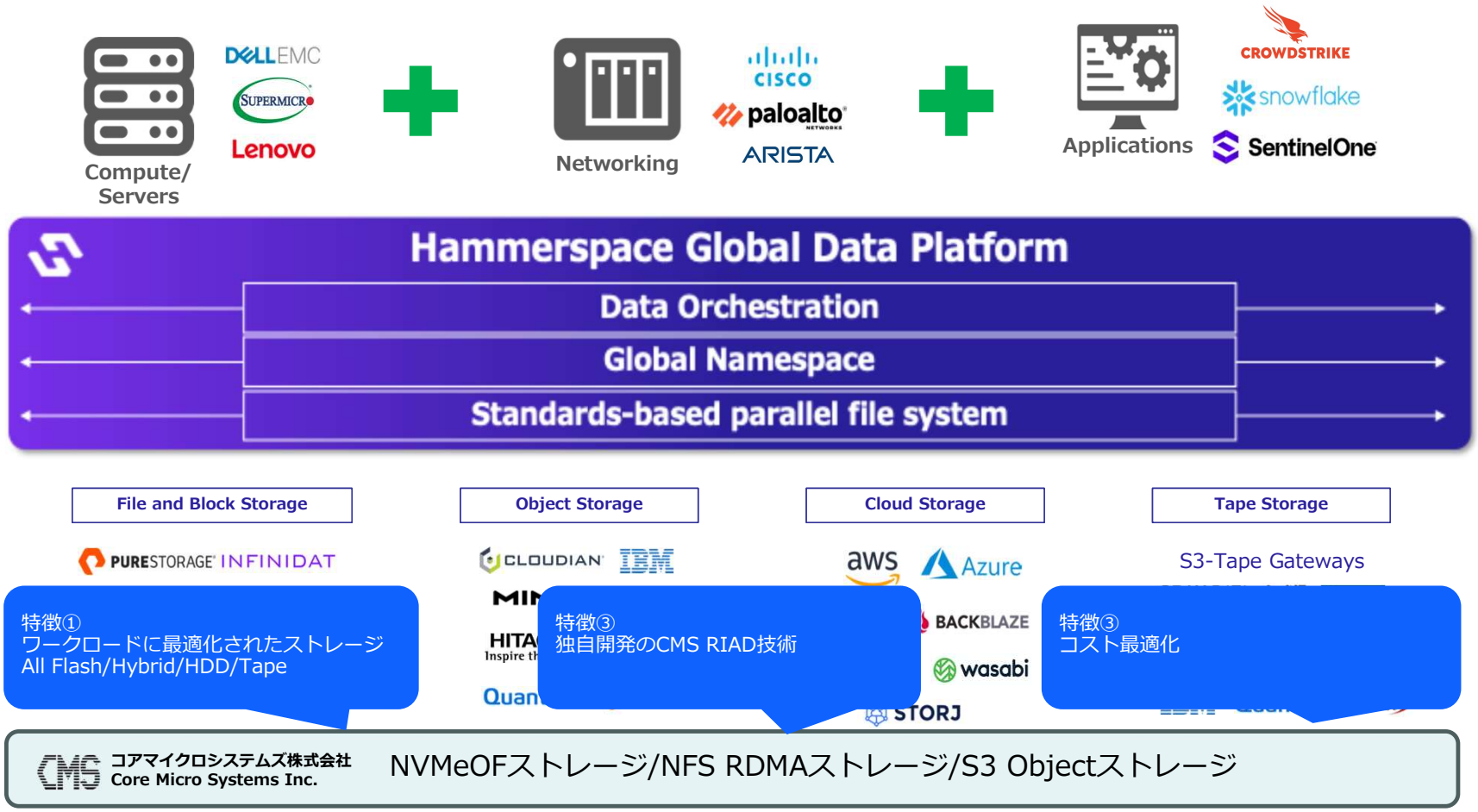


現在～これから

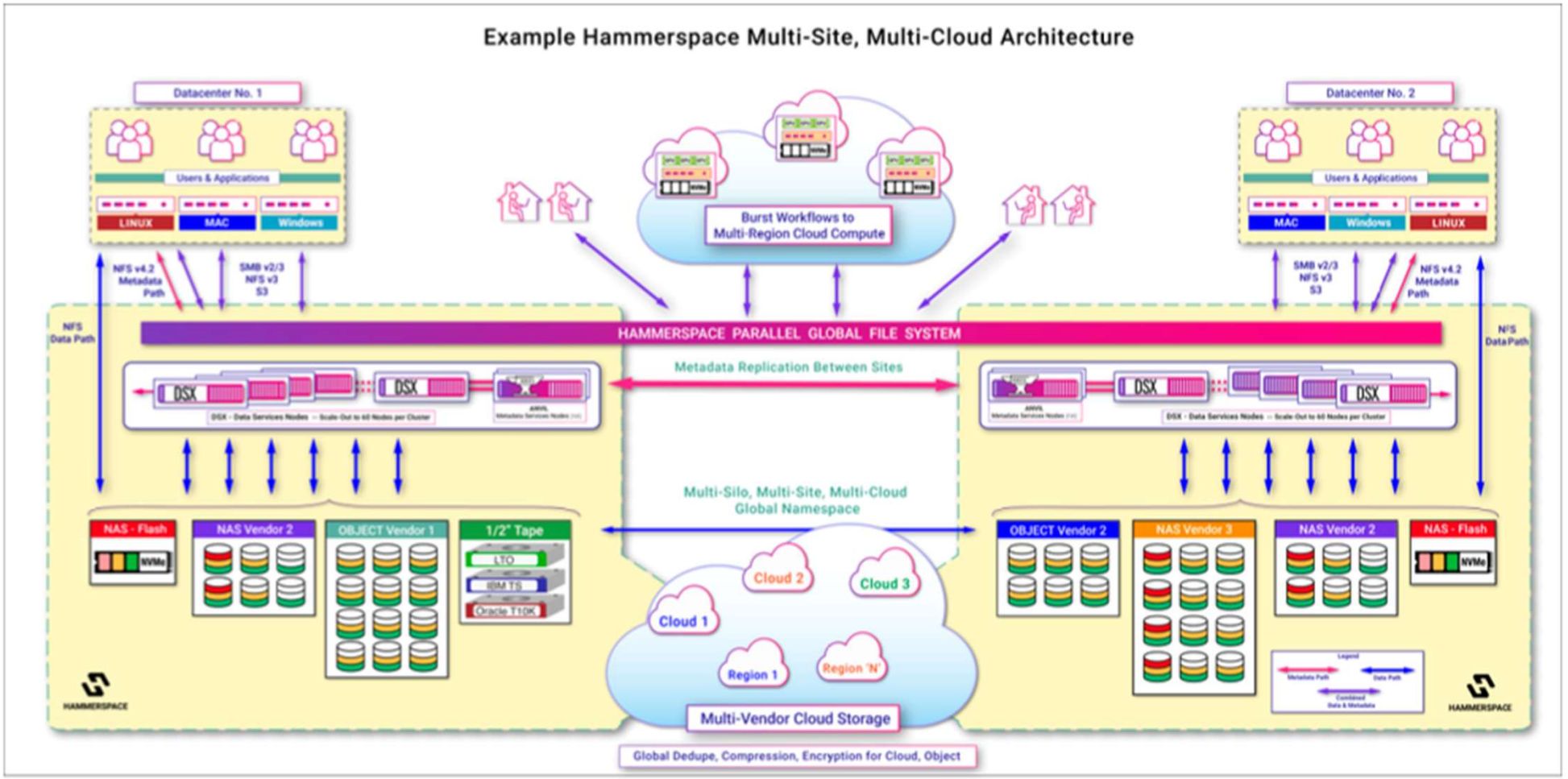
- マルチサイトのデータ利用可能
- CPU/GPUの使用率を最大化
- クラウドへのシームレスな連携
- 自動データオーケストレーション



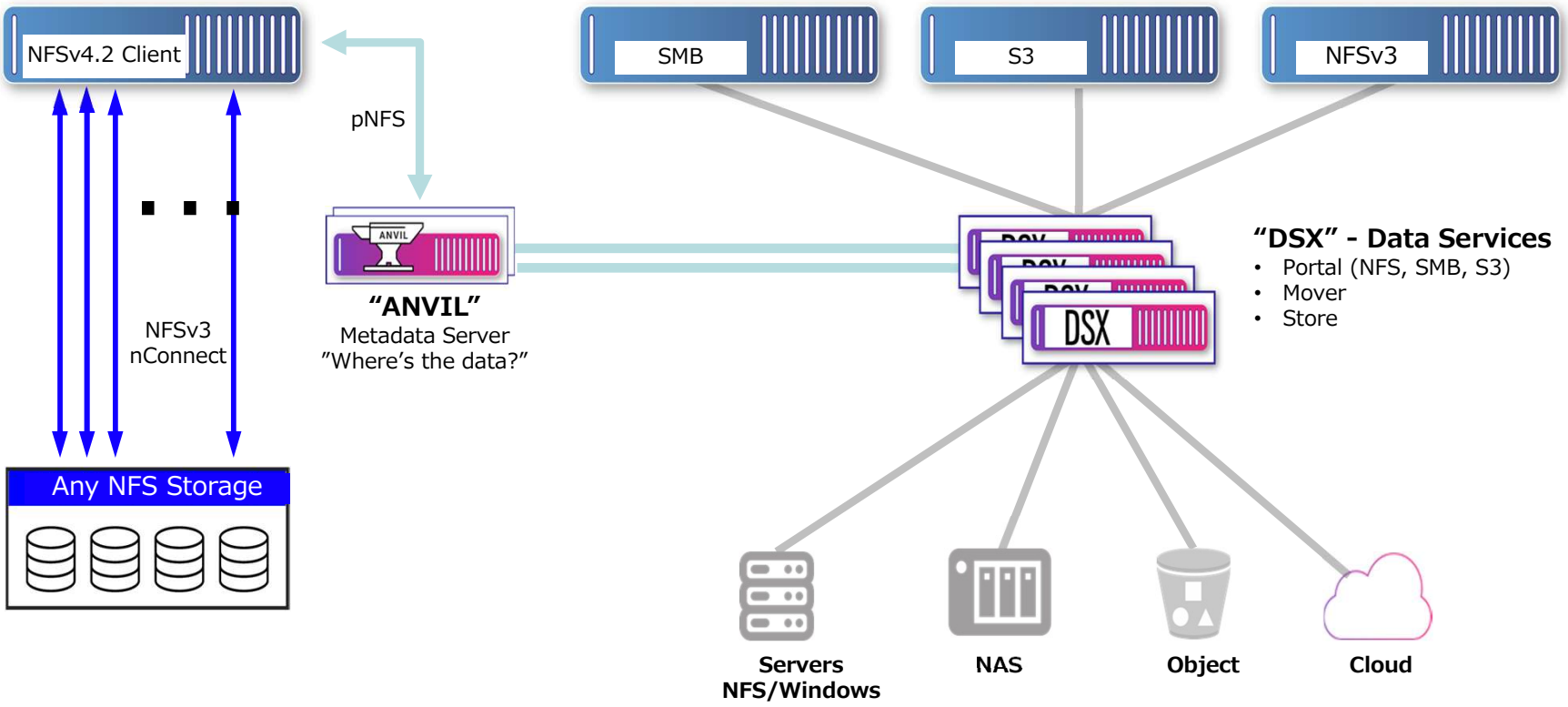
広域AI HPCストレージHammerspace + CMSによる課題解決



Hammerspace アーキテクチャ



Hammerspace キーコンポーネント



Hammerspace キーコンポーネント

Hammerspace は、AnvilとDSXという２つのキーコンポーネントでハイ・パフォーマンスと
シングルネームスペースを実現しています

「Anvil （メタデータサーバー）」

・メタデータの管理（ローカル＆リモート）

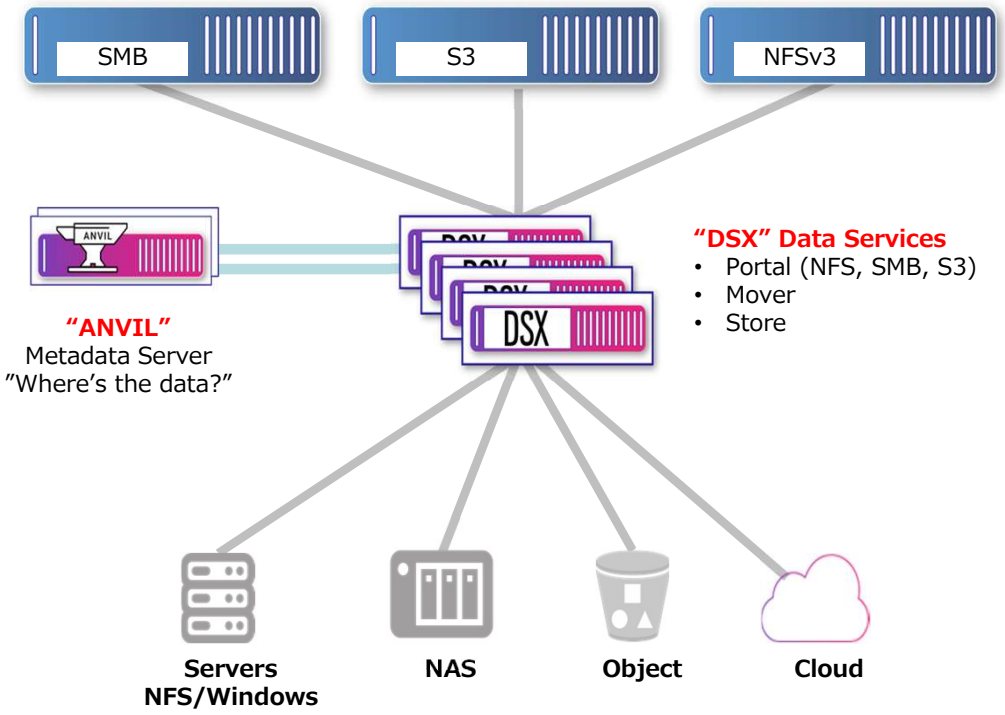
・リモートサイトのAnvilとメタデータを同期

・I/Oエラー発生時でもデータの可用性を最大化

・ファイル、オブジェクト、インスタンスに
関するすべての情報の保存

「DSX （データサービスサーバー）」

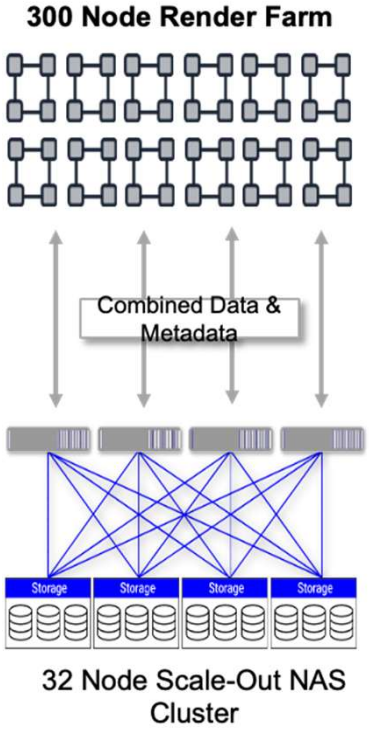
・Portal
ヘテロな環境のファイルとオブジェクトに対するアクセス
・Move
データの移動　・Store
ストレージとしてのデータ管理



ハイパフォーマンスソリューション - Hyperscale NAS

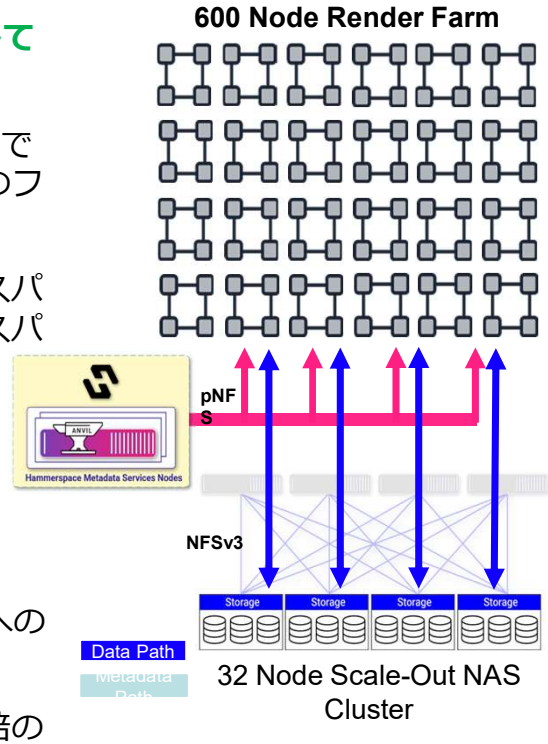
Before Hammerspace:
Scale-Out NAS Bottlenecks

- スケールアウトNASシステム、パフォーマンス要求に対応できず
- コントローラーがボトルネック
- メタデータの処理がシステム・リソースを消費
- レンダリングジョブが定期的にタイムアウトし、停止する。
- 300ノードのレンダーファームが十分に活用されておらず、フル稼働できていない



After Hammerspace :
Hyperscale NAS Doubled Performance

- **NFS v4.2の仕様として追加（RFC8435）**
- Hammarspace 経由でアクセス可能な全てのファイル（２PB)
- メタデータのアクセスパスをデータのアクセスパスと分離
- ノード間の競合やキャッシュの競合がない理想的なアクセスパターン
- 既存のNASシステムへの変更は不要
- レンダーファームが倍の600ノードに!



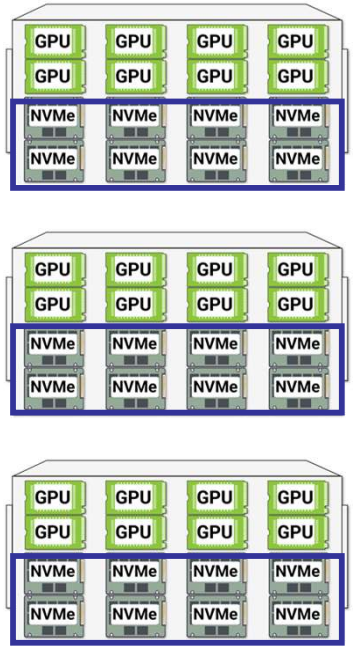
ハイパフォーマンスソリューション - Tier 0

お客様の声

・大規模GPUコンピューティング環境の導入（DGX、HGXなど）
・インフラコストや制約と同様に、Time-to-Valueが重要
・外付けストレージシステムはコスト増、電力消費、時間かかる

GPU Server の ローカルストレージは使われていない

・サイロ化されているため、データに利用されていない
・GPUサーバーは2025年には1-2PBのストレージ搭載が可能になる



16 SSDs per Server

16 x 61.44TB = 983.04TB

16 x 122.88 = 1966.08TB

Tier 0 適用例　・・・　チェックポイント処理の短縮

シナリオ

- 容量100ペタバイト、スループット1TB/秒
- GPUサーバー1,000台、各128 TB（128 PB）
- GPUサーバーのNVMeはすでに導入済み（cost)

インフラコストと消費電力

- インフラのコスト削減 ≈ **$40M**
- 電力の削減 ≈ **Millions of K-watt-hours**

チェックポイント処理の性能比較

- 長いチェックポイント時間＝GPU利用機会の損失
- 外部ストレージ = 600秒/チェックポイント
- Hammerspace のTier 0 = 6秒/チェックポイント

GPU の利用機会のコスト削減

- 年間で節約されたGPU利用時間 ≈ **929K hours**

  →約 **848 GPUs** に相当
- 現在のGPUコスト = $20K-$40K/

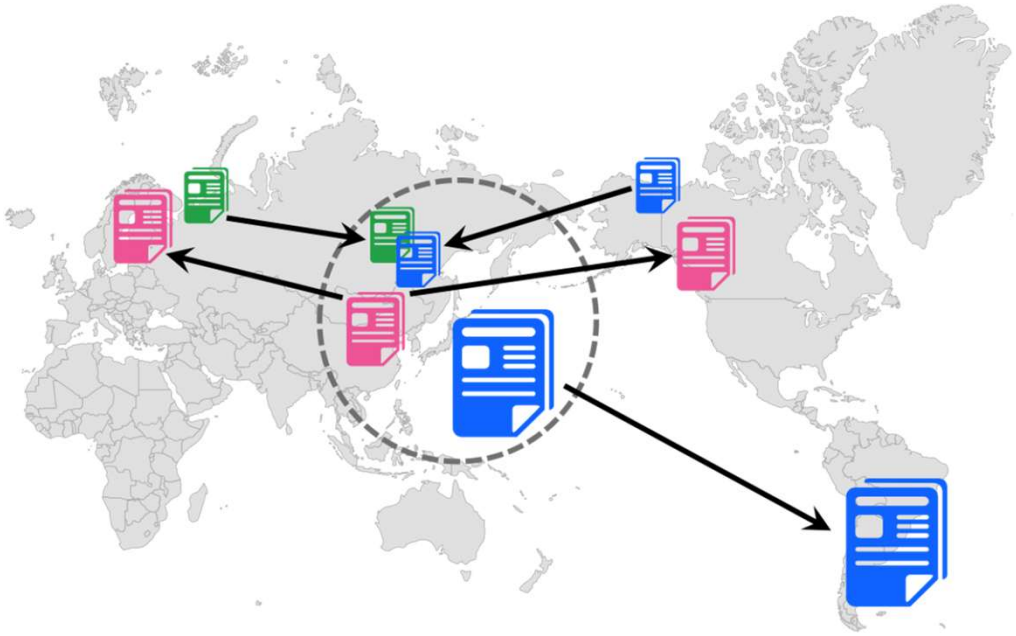
  → ≈ 約**$17M-$33M** GPU コストの削減

注：数字はすべて概算です。費用は、地域、ドライブとネットワークの種類、密度、ワークロードによって異なります。

Hammerspace Global Name space

複数拠点間におけるデータ連携の必要性

- データサイロ化の解消
- データの分配
- 各拠点データの収集
- データの(災害復旧用)コピー



メタ情報の連携とポリシーベースのILM

- ポリシーベースのILMを実現
- サイト間のメタ情報共有による遅延の回避
- キャッシュによる性能および容量の効率化
- ポリシーベースのデータ保護
- 既存資産の有効活用

データの保護 - Snapshot

- HammerspaceはShareレベルのスナップショット
  - 即時またはスケジュール
  - デフォルトおよびカスタムのスケジュールと保持ルール
  - Shareのスナップショットは日時で命名されます
  - Shareあたり4096のスナップショット
- 古いスナップショットファイルを安価なティアに配置する
  - Object-Storageでは、重複排除により、事実上「永遠に増分」が可能
- すべてのサイトのスナップショットの一貫性の保証
  - スナップショットの管理が必要なのは1つのサイトだけで、すべてのサイトがアクセス可能

データセキュリティ – WORM、暗号化、コンプライアンス

- 指定されたI/O操作をブロックまたは拒否する機能（排他制御）
  - block-*は、I/Oがブロックまたは無効化され、アプリケーションは（通常）ブロックが解除されるまで待機します
  - ウイルススキャンが保留中の場合など有効です

block-open
block-create
block-write
block-read

- deny-* は、I/O が失敗し、「アクセスが拒否されました」と返されることを意味します
  - 長期にわたって有効なもの、または「7年間変更不可」などの規制コンプライアンス要件を満たす文書をロックするのに適しています

deny-open
deny-create
deny-read
deny-write
deny-delete

- Files with deny-* or block-*は、階層間およびサイト間を移動可能
  - ストレージプラットフォームの寿命よりも長期にわたる保存要件がある場合、ファイルを移行して保存し、その保存目標を達成することができます

ファイルのバージョニング

- versioningが有効にされたファイルが、少なくとも3分間閉じられた後に書き込まれた場合、以前のバージョンが保存されます
  - 拡張子の前にタイムスタンプとサイトIDがファイル名に挿入されます
  - filename[#M=<YYYYMMDD.HHMMSS.SSSSS #S=n-nn>].ext
- Share、Directory、Fileで設定した目的によって有効化されます
- Recovery options
  - .snapshot/current を参照/移動し、バージョンを表示して見つけることができます
  - GUI – .snapshot/current を参照し、スナップショット復元アクションアイコンをクリックします
  - クライアント – .snapshot/current を参照または移動し、コピーしてタイムスタンプを削除するために名前を変更します

メタデータの保護

- Anvil（メタデータサーバー）はシングルまたはHA (High-Availability)構成
  - AnvilのHAペア構成(Active-Standby)はベストプラクティスです
  - Secondary Anvilのフルタイムのジョブは、プライマリからのメタデータデータベースのレプリケーション

- NFSエクスポート先へのメタデータと設定のバックアップ
  - 「メタデータと構成のバックアップが構成されていません」と警告されます
  - CLIのbackup-createコマンドを使用して構成します（即時またはスケジュール）
  - アプライアンスが破損した場合は、新しいアプライアンスを設置し、バックアップを指定すると、自動的に再構成され、メタデータがロードされます

Hammerspace Global Data Platform – 事例（Meta）

“What Hammerspace does is pure magic.”
-Paul Saab, Principal Engineer Meta

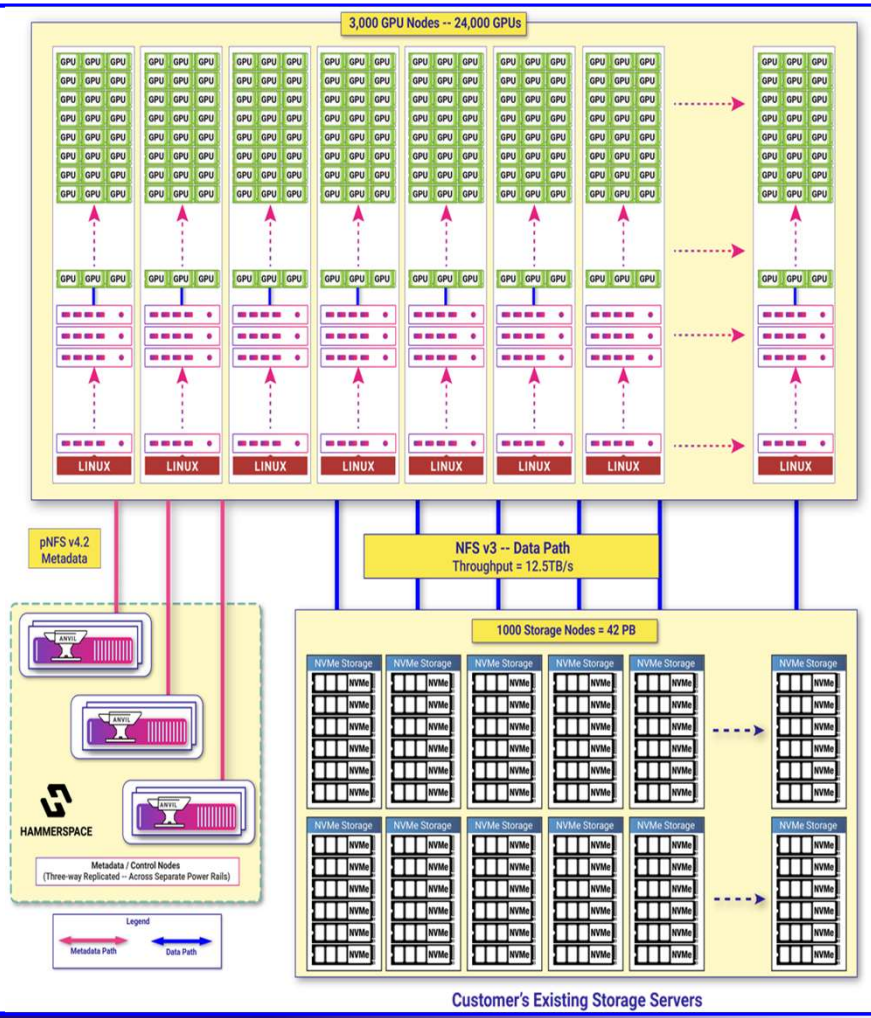
Meta

お客様について

- 世界最大のSNSサービス会社
- LLMやその他のAIモデルのトレーニング
- 膨大なパフォーマンスとスケールの要求
- 主要ストレージ・ベンダーの評価

Hammerspaceのソリューション

- どのベンダーもHammerspaceの足元にも及ばない能力
- 1,000ノード以上のHammerspaceストレージクラスタ
- 24,000個のGPUに給電、まもなく350,000個、そして1,000,000個へ
- 12.5TB/秒（100Tb/秒）の総合性能
- すべてが標準ベースでプラグアンドプレイ
- 既存のOCPストレージ・サーバーの活用



Hammerspace Global Data Platform – 事例（Jellyfish Pictures）

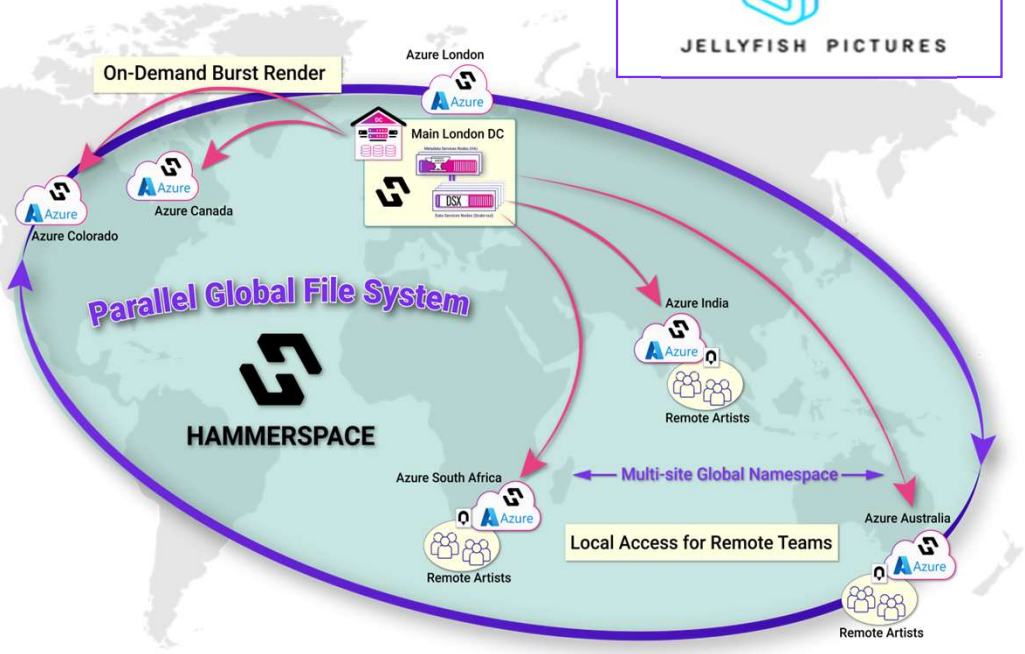
問題：マルチサイト/クラウド - グローバルでのファイル共有

- 世界有数のビジュアルFX・アニメーションスタジオ
- 英国の2カ所に300人のアーティストを擁し、世界各地でリモートワークを展開
- 2019年から仮想化、複数の場所とクラウドに保管
- アーティストの生産性を維持するために大容量ファイルをグローバルに共有することは「現実的な問題」だった

ソリューション：Hammerspaceデータオーケストレーション

- グローバルなデータ移動をオーケストレーション
- オンプレミス環境とAzureクラウド環境の両方
- Jellyfishのプロダクションスケジューラと統合：Autodesk Shotgrid
- 100M以上のファイルにグローバルにアクセスでき、数分で読み書き可能
- カナダと米国の低コストのコンピュート・リソースにファイルを移動することでレンダリング・コストを30万ドル削減。
- アーティストのグローバルな人材プールへのアクセスが可能に

JELLYFISH PICTURES



コアマイクロシステムズの広域AI HPCソリューションご紹介

サイロ化された既存ストレージ群の統合と広域共有の実現

既存環境

サイロ化が進む非構造化データ環境

パフォーマンスとリソースの限界

クラウド利用における俊敏性の欠如

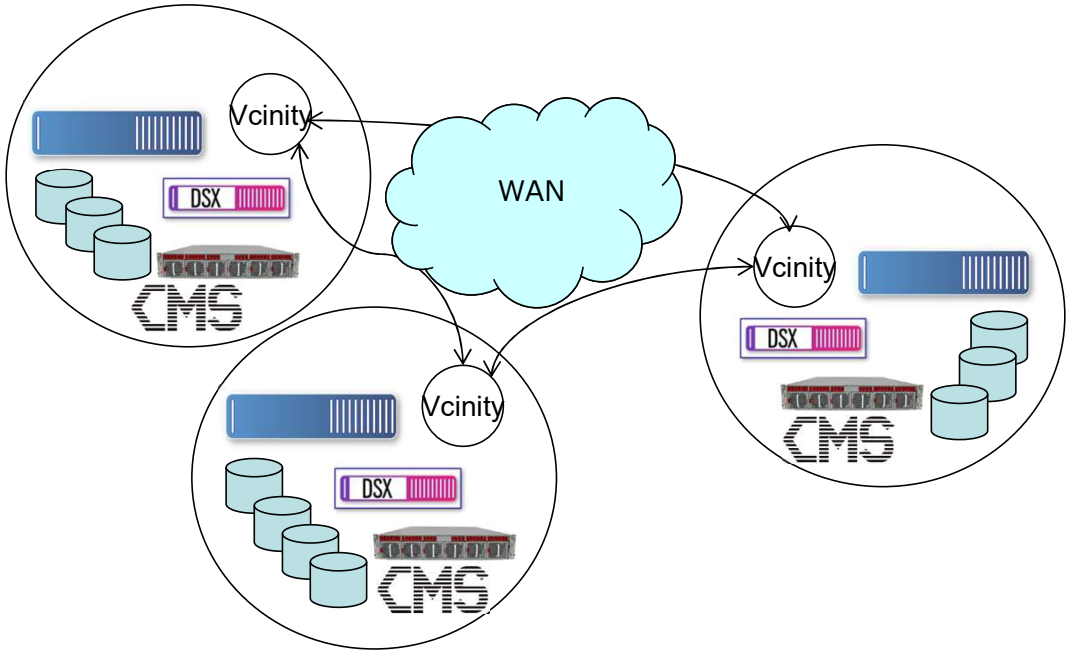
手作業によるデータ管理

ストレージ環境の統合階層化と広域共有化

グローバルシングルネームスペース化(Hammerspace)

広域RDMA対応データ共有コラボレーション(Vcinity)

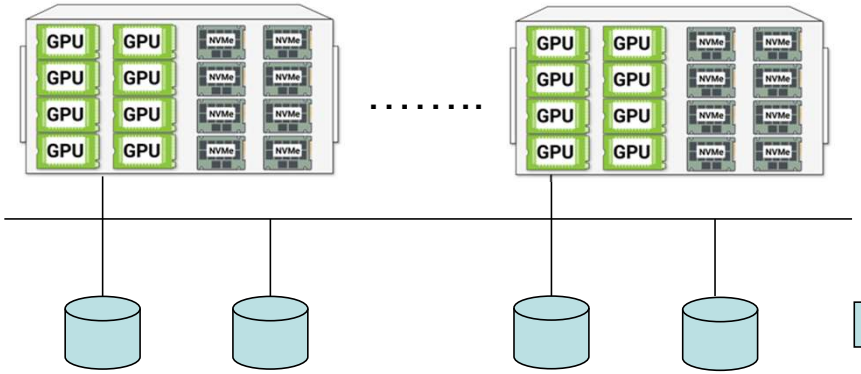
ストレージ階層化による性能及びコストの最適化(CMS)



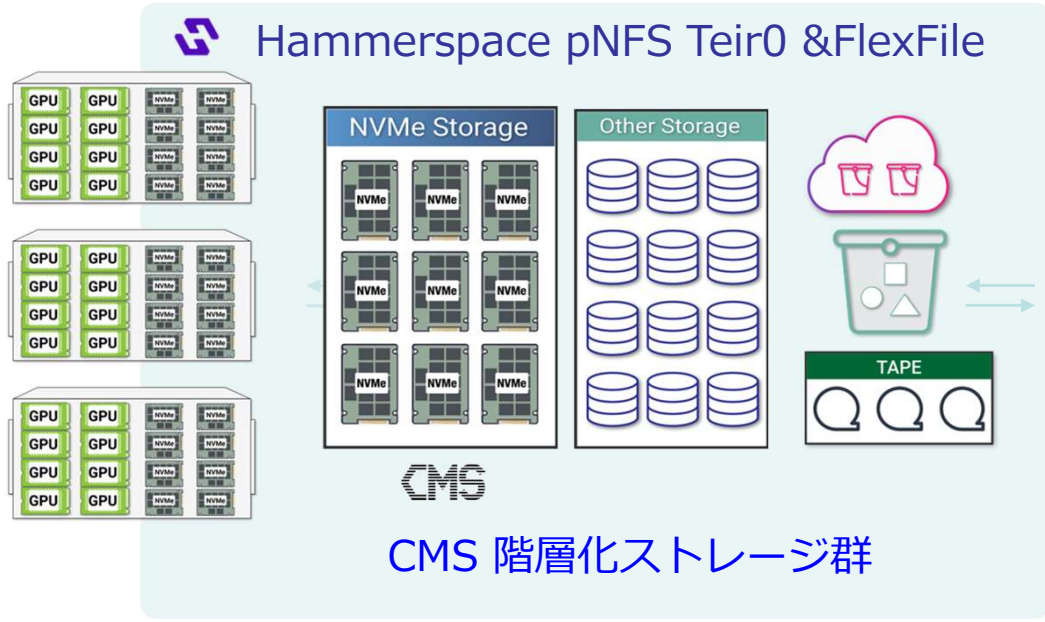
HAMMERSPACE
DATA IN MOTION

CMS

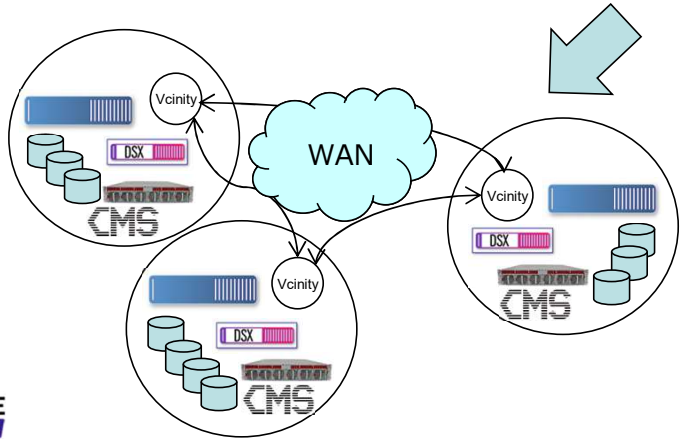
既存のHPC/AIプラットフォーム環境の改善



1.GPUサーバ内の高速NVMe有効利用と外部ストレージの階層化



3.複数拠点の広域データ共有コラボレーション



2.階層化ストレージ群による性能/容量/コストの最適化

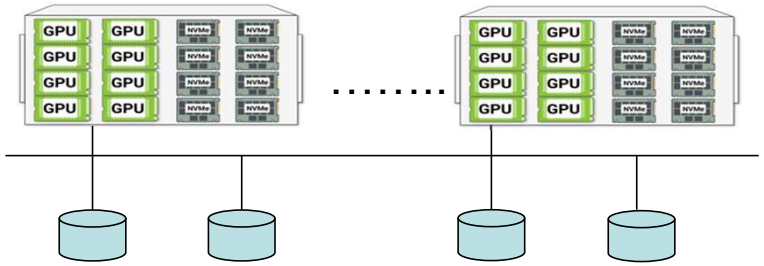
Tier 0	Tier 1	Tier 2	Archive
Microseconds	Milliseconds	Seconds	Minutes

HAMMERSPACE
DATA IN MOTION

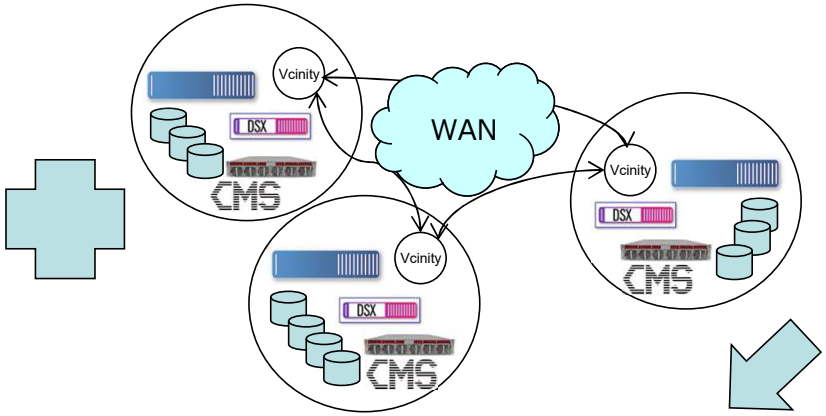
CMS

既存のHPC/AIプラットフォームの新規拠点構築時の効率統合

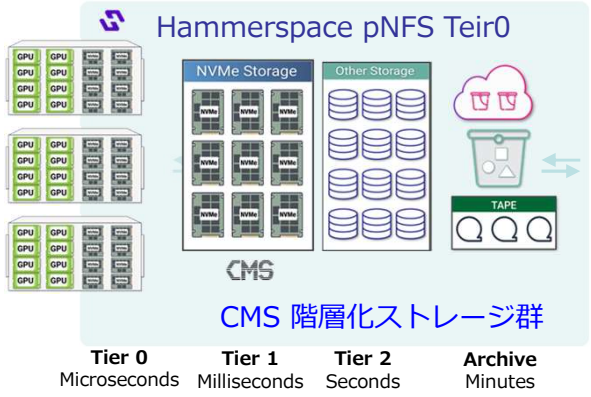
1.既存HPC/AIプラットフォームとの運用改善
Parallel Worksによる、リソース統合の効率化・
プロビジョニング・Job管理の最適化



2.複数拠点ストレージのHammerspace＆Vcinityによる
リアルタイム広域RDMAデータ共有コラボレーション化



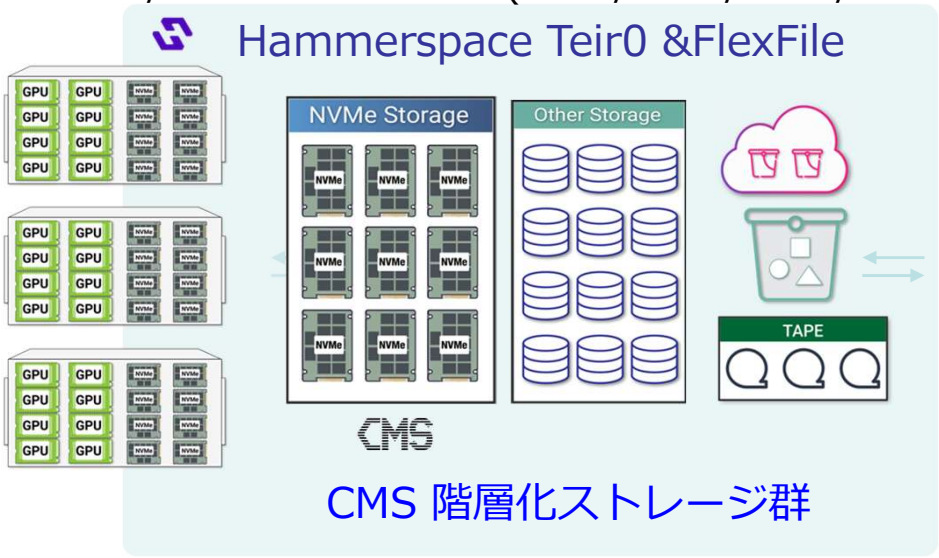
3.CMS階層化ストレージ群による性能とコストの最適化



4.データオーケストレーションと
スムースなデータ移行

新規HPC/AIプラットフォーム環境の高効率な構築

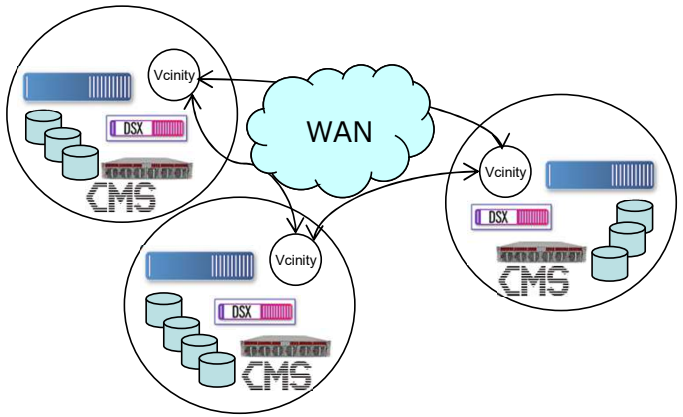
1.新規HPC/AIプラットフォーム(Tier0/Tier1/Tier2/Archive)構築



Tier 0	Tier 1	Tier 2	Archive
Microseconds	Milliseconds	Seconds	Minutes

2.CMS階層化ストレージ群による最適化

3.複数拠点ストレージ群のシングルネームスペース化と広域データ共有コラボレーション化

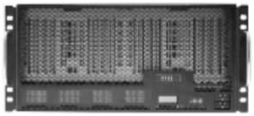


4.Parallel Works & Hammerspaceによる
コンピュートリソース&ストレージリソースの
広域統合オーケストレーション化

HAMMERSPACE
DATA IN MOTION

CMS

CMSストレージソリューション・ポートフォリオ（抜粋）

CMSストレージソリューション・ポートフォリオ（抜粋）				
CMS GpSCALERシリーズ			各種ストレージ（抜粋）	
GpSCALER NVPro	GpSCALER NVMax	GpSCALER Hybrid	CMS Ceph（オブジェクト）	テープ（外部＆長期保管）
超高速 All Flash	大容量All Flash	SSD/HDD ハイブリッド	アーカイブ＆長期保管	長期＆外部保管
				
-2U 32NVMe CiB (Active/Active HA) -Gen 5 PCIe ベース -400G(Infiniband/Ethernet)x16ポート -実行バンド幅 384GB/s -最大物理容量 983TB (30.72TB x 32)	-4U 80NVMe CiB (Active-Active HA) -Gen4 PCIeベース -200G (Infiniband/Ethernet)x16ポート -実効バンド幅 192GB/s -最大物理容量 9.83PB (122.88TB x 80)	**ヘッド部** -4U 80NVMe CiB (Active-Active HA) -Gen4 PCIeベース -200G (Infiniband/Ethernet)x16ポート -実効バンド幅 192GB/s -最大物理容量 9.83PB (122.88TB x 80) **JBOD部** -4U 80NVMe CiB (Active-Active HA) -Gen4 PCIeベース -200G (Infiniband/Ethernet)x16ポート -実効バンド幅 192GB/s -最大物理容量 9.83PB (122.88TB x 80)	-4U 90 SAS/SATA -10Gbps Ethernet x 2 -８ノード/ラック -ラックスケール (詳細についてはお問い合わせ下さい)	IBM社製テープライブラリ装置 （TSシリーズ）など (詳細についてはお問い合わせ下さい)
＊各製品の仕様については予告なく変更になる事があります。				

ご清聴どうもありがとうございました。

コアマイクロシステムズ株式会社

URL : https://www.cmsinc.co.jp
tokawa@cmsinc.co.jp