

NECのAI研究用スーパーコンピュータ におけるOSS活用事例

日本電気株式会社

グローバルイノベーション戦略統括部

深層学習基盤開発グループ ディレクター

北野 貴稔

北野 貴稔

KITANO TAKATOSHI



- インド国民13億人が利用する世界最大の大規模生体認証システム（アドハー）を実現、政府・空港等を中心に生体認証PFを世界40か国へと展開
- 機械学習の研究テーマリーダーを経て、現在は国内企業で最大規模となる「NECのAIスーパーコンピュータ」の企画・設計・開発・運用を統括
- 担当の技術レイヤは全て。NECのAIスパコンの開発責任者として、HWの選定・調達から計算機基盤・SW基盤等のアーキ設計・開発・運用の全てを担当

目次

1. NECのAI研究用スパコンの概要
2. NECのAI研究用スパコンでのOSS活用事例
3. まとめ

NECのAI研究用スーパーコンピュータの概要

NECの国内最大級のAI研究用スーパーコンピュータ

928基のA100 GPUを搭載した国内最大級のAIスーパーコンピュータを
2023年3月に稼働開始、NECの研究者数百名が利用しており利用率は100%

現在の利用状況

利用する
ユーザー

利用する
研究者数

AIスパコンの
利用率

NECの研究者 数百名 100%

AIスパコン(A100が928基)

開始時期

GPUの演算能力

国内企業
No.1

2023年3月 100 ⇒ 580 (PFLOPS)

NECの生成AI研究を大幅強化！



生成AI研究を支えるNECのAI研究用スパコンのスペック

2020年から企画、2022年3月に部分稼働・2023年3月に全面稼働
多数の試行錯誤を可能にする圧倒的な計算能力がAIの競争力の源泉

コンピューティング

928基のNVIDIAのA100 GPU を搭載 (580PFLOPS)

ネットワーク

NVIDIAの超高速・低遅延ネットワーク (200GbE)

ストレージ

DDNの大規模・高速ストレージ (16PB)

NECの全研究所での生成AI研究を強化、社会価値創造を加速

価値創造プロセス

NECの先進AI群

顔・虹彩認証で世界No.1、機械学習の難関学会採択数で世界10位

画像認識
映像認識

言語・意味理解
データ分析
予測予兆検知

最適計画
制御

社会課題



見える化



分析



対処

社会価値

ファウンデーションモデル

全AI研究を加速

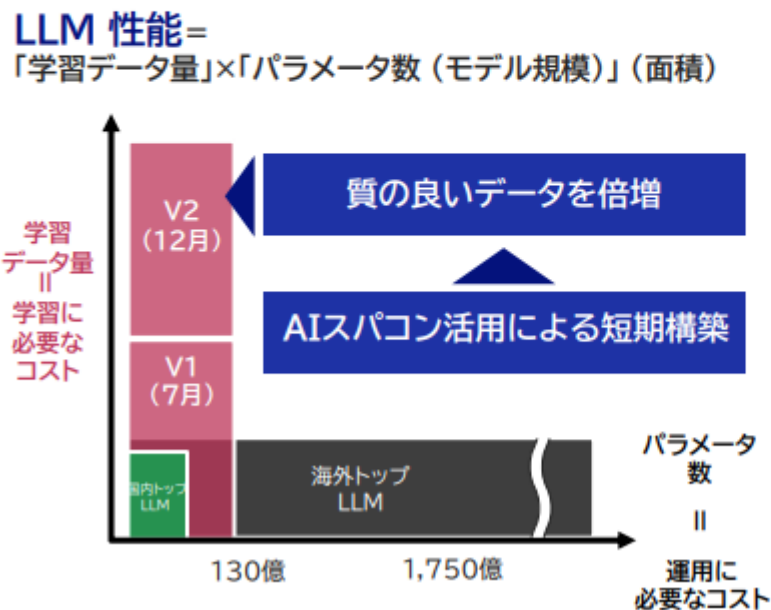
国内企業で最大規模となる
NECのAIスーパーコンピュータ

NECの大規模言語モデルcotomiの短期開発に成功し市場展開

928基のA100 GPUを用いて、高い性能・コンパクトさを両立した日本語特化の大規模言語モデルの短期開発に成功し、2023年7月から事業展開中

- ・ 13B LLMの強化 : 質の良いデータを整備することで、軽量ながら世界トップクラスの日本語能力を堅持
- ・ 長文プロンプト対応 : 「書籍まるごと」の処理も可能な長文処理能力(他社比 最大150倍)

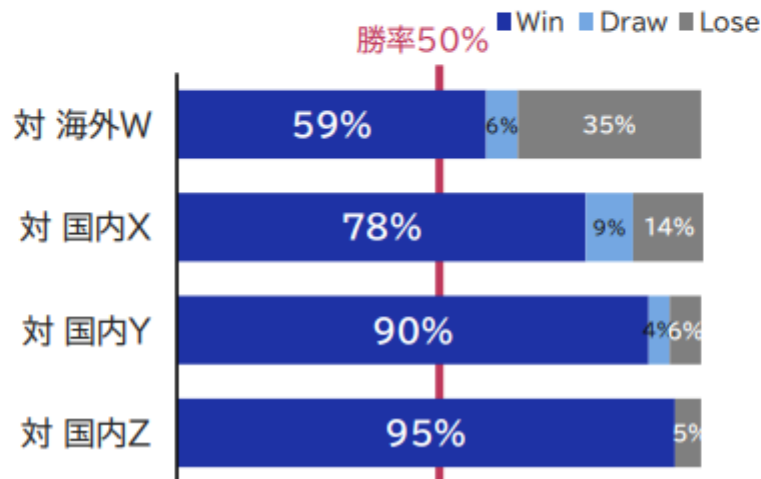
継続的なLLM強化



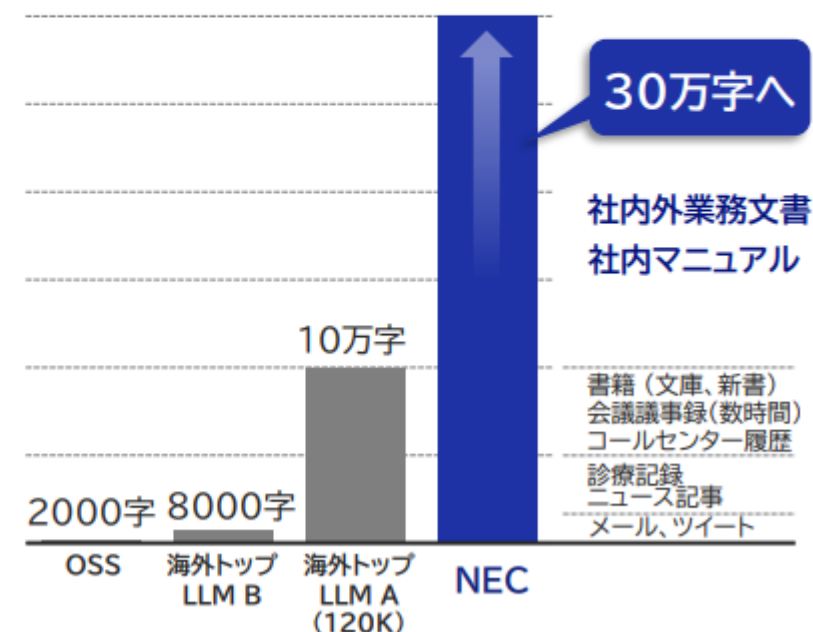
世界トップクラスの日本語処理能力

全ての他社LLMを上回る

日本語対話能力の比較評価 (Rakuda)



他社比 最大150倍の長文処理能力



2023年12月 当社調べ

映像・画像×LLMによるマルチモーダルAIの応用も展開中

映像×LLM

R&D

映像認識技術とLLMによる実世界の把握/文章化

実世界の動画を与えられると、それを詳細に説明する文章を自動で作成。NECの強みは、膨大な数の映像認識AI(映像群⇒文字化)。さらにそれらの文字群を、LLMにより意味のある情報に変換し文章化



2023年12月5日発表

画像・位置情報×LLM

R&D

災害対策ソリューション

NECの類似性判定技術と位置推定技術により、続々と提供される膨大な画像から被災状況と位置を番地レベルで特定。LLMとの融合により、言葉での被災規模や状況の整理を実現し、初動を迅速化



2023年8月25日発表

NECのAI研究用スパコンでのOSS活用事例

何故OSSを全面採用し、どのように活用しているのか？

NECのAIスパコンでは何故OSSを全面採用しているのか？

NECのAIスパコンでは商用SWではなく、全面的にOSSを採用し独自開発。
先進のAI研究を可能にし続けると共に、優れたエンジニアリング文化を組織に宿す

AIの進化に合わせて、研究開発インフラを発展させ続ける

- ✓ これは出来ないというAI研究を作らない。常に先進AIの研究を実現可能にする
- ✓ プラットフォームは利用者のためのもの、利用者価値の最大化を最重要視する

下から上の全ての技術レイヤで高度な技術開発力を内部で保有する

- ✓ どんな問題があっても最後は自力で解決が可能にする
- ✓ 優れたプロダクトを作るには、効率と作りこみの両立が必要

OSSの優れた文化を組織に宿し、優れたエンジニアリング文化を築く

- ✓ OSSには優れたソフトウェアの思想が宿っている。ソフトウェア開発に魂を込めるための文化を組織に宿すため。その文化と思いがプロダクトの質を変える。

NECのAI研究用スーパーコンピュータでのOSS活用

開発言語は基盤・運用開発全てでGoを採用し、高いエンジニアリング効率を実現
Kubernetesを中核として拡張・コンテナを独自開発し、先進のAI研究開発を支える

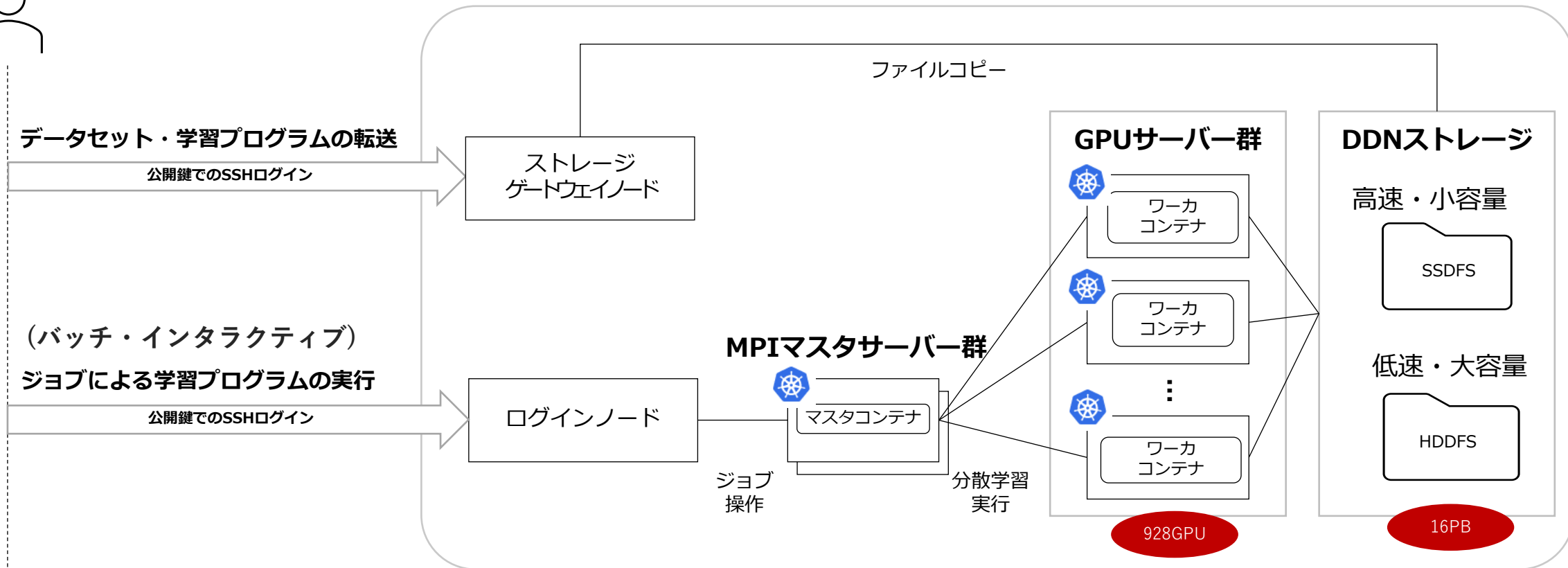
分類	利用している必要なOSS/Software	説明
開発言語	Go	開発・運用で全面的にGoを採用
コンテナ基盤	Kubernetes, Docker	K8sの拡張をGoで開発 ジョブスケジューラをGoで開発
モニタリング	Prometheus, Grafana, Victoria Metrics	ExporterをGoで開発
ログ管理	FluentBit, ElasticSearch, Kibana	FluentBitの拡張をGoで開発
セキュリティー	Falco, eBPF	GoでFalcoの拡張を開発
プロビジョニング	Ansible, Autoinstall	ノードの自動構築
通信	OpenMPI, NCCL	
深層学習	PyTorch, TensorFlow, JAX, Megatron DeepSpeed, Optuna, TensorBoard	一部をAIスパコン向けに拡張し利用
ユーザー管理	OpenLDAP	ユーザー管理、認証など
OS	Ubuntu	全GPU/CPUサーバーのOS

NECのAIスパコンのSW基盤の全体アーキテクチャの概要

Kubernetesを用いて最新の深層学習FWをジョブスケジューラで高効率で利用可能
先進HW・先端SWを密結合した設計で、高効率・利便性の高い分散学習システムに

AI研究者

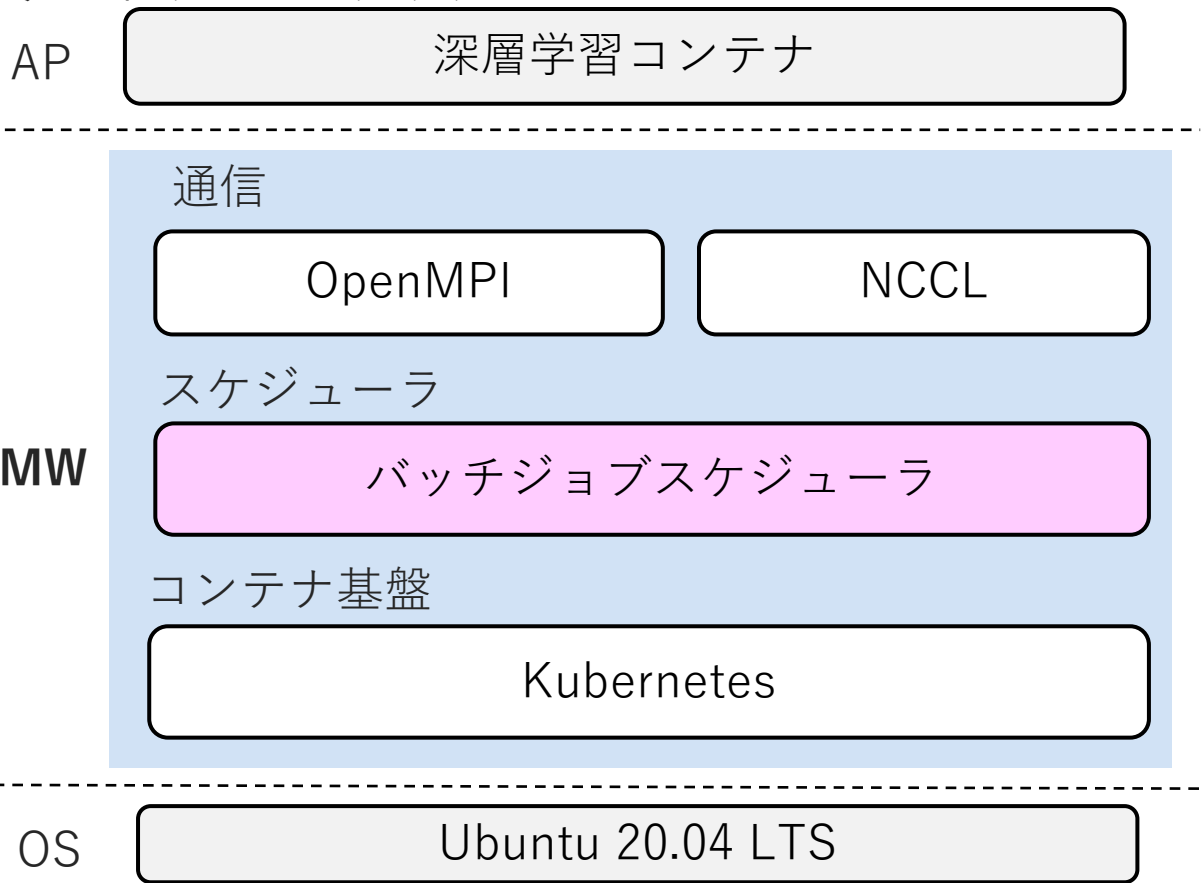
NECのAIスーパーコンピュータの構成概要



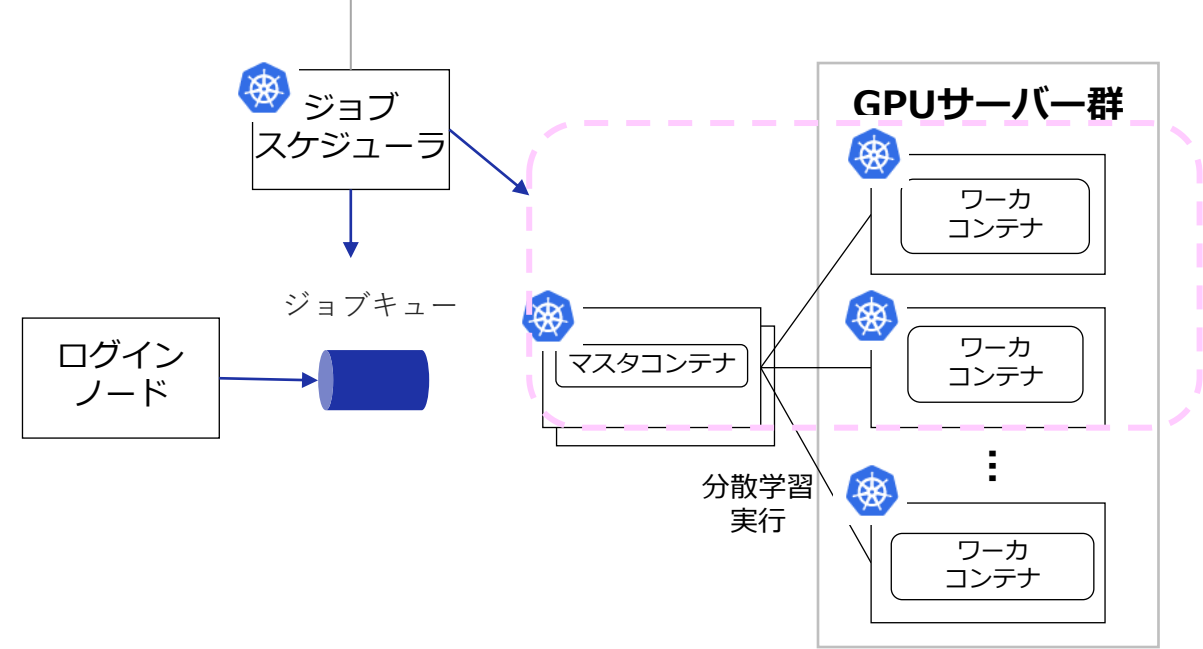
多数GPUサーバーでの利用を支えるSW基盤（ミドルウェア層）

KubernetesをSW基盤の中核に採用。カスタムの独自ジョブスケジューラを開発、最新の深層学習FWをコンテナを利用可能することが出来、高効率かつ利便性・柔軟性の高い分散学習環境を実現

ソフトウェアスタック



ラック配置・NWトポロジに基づいてGPUサーバー間のNW距離が最小となるようスケジュール



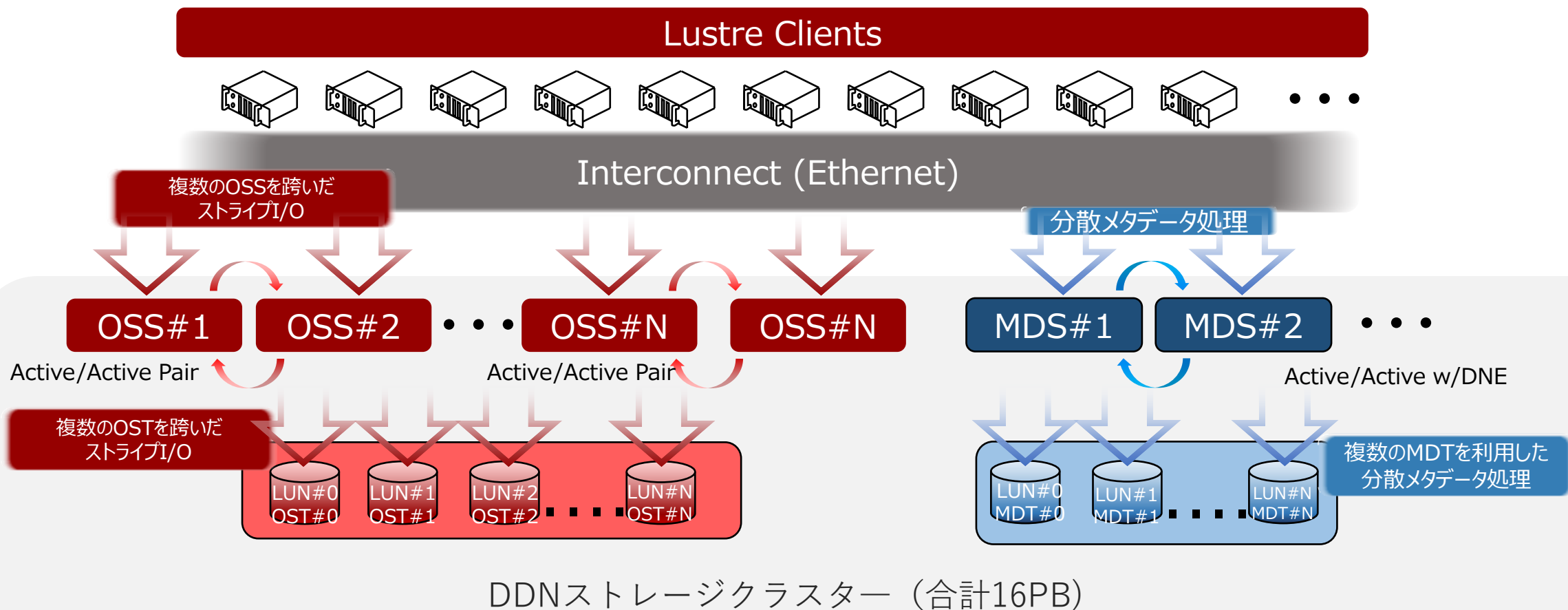
最新の生成AI研究を支えるSW基盤（AP層）

深層学習FWは、ドメイン毎にFW/Libraryを含んだものをコンテナイメージとして開発・提供
Horovod/Optuna等を用いて多数GPUノードでの分散学習・ハイパラ探索が可能

項目	内容	説明
深層学習FW	<ul style="list-style-type: none">PyTorch2TensorFlow2TransformersJAXMegatron DeepSpeed	最新の深層学習フレームワークを、ドメイン毎に典型コンテナイメージを作成し、2-3か月に1回程度の頻度で更新し展開
分散学習	<ul style="list-style-type: none">HorovodDDP・FSDP	分散学習は、HorovodやPyTorch（DDP・FSDP）等をサポート
ハイパラ探索	<ul style="list-style-type: none">Optuna	OptunaにHorovod/torch distributed向けの拡張を行い利用

NECのAIスパコンのストレージシステムのアーキテクチャ

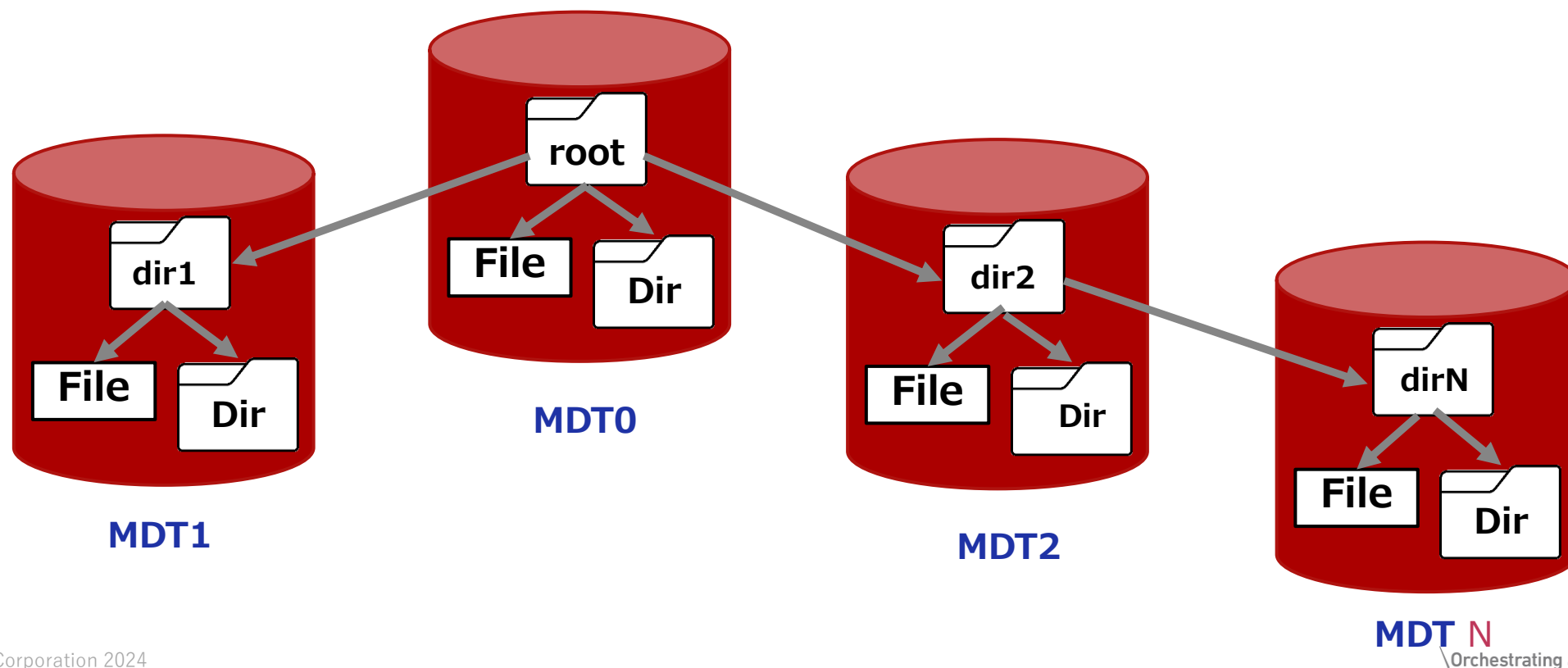
MDSがActive/Activeでメタデータ処理、OSSがActive/Activeでデータ管理し、高可用性・高信頼性を実現
メタデータの性能はDNEを用いて複数のMDSを使いスケール、データのスループットはOSSをNペア用意しスケールアウト



Lustre DNEを用いたメタデータ性能のスケールアウト

ユーザーが属するグループを、複数のMDT数に合わせて自動配置する機能をNEC側で開発
DNEにより複数のMDSを活用してMetadata性能をスケールアウト、大規模FSの性能向上

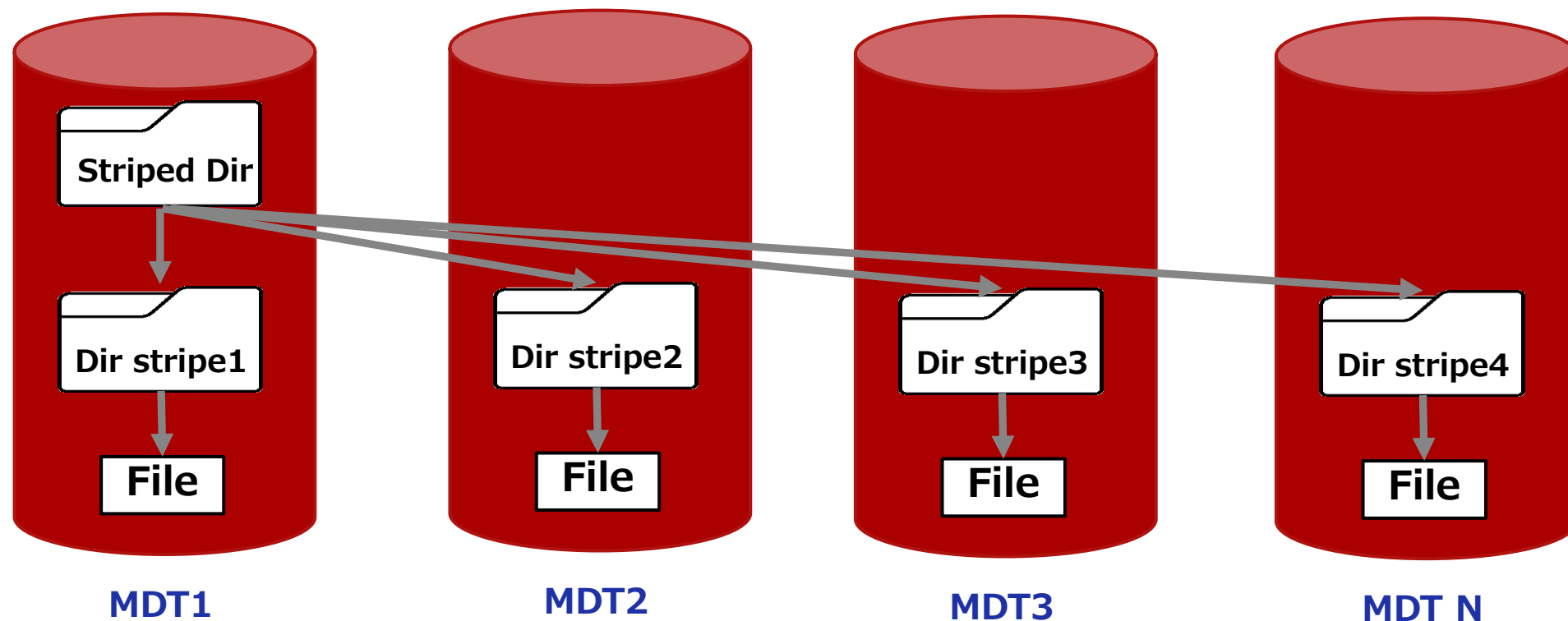
以下はDNEのイメージ例



DNE2 (Striped Directory) でメタデータ性能をスケールアウト

Striped Directoryを使ってまたデータ処理を複数のMDTに分散することでボトルネックを解決可能
Striped Directoryを要望に応じ部分設定する機能開発を行い展開。全面的利用はi-node数の関係で行っていない。

Striped Directoryのイメージ例

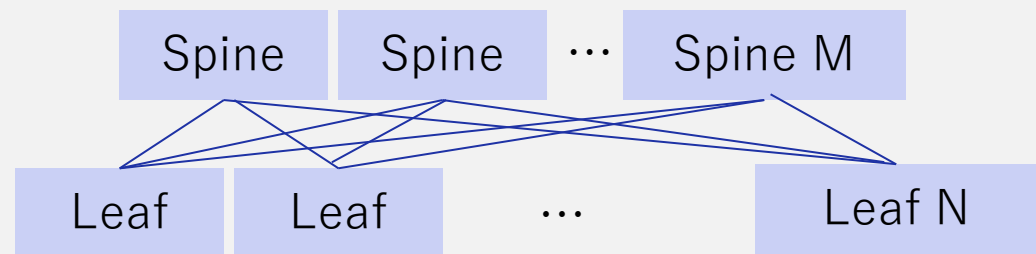


NECのAIスーパーコンピュータのネットワークの構成

全NWはSpine/Leafアーキテクチャで構成、全サーバー間はRoCEv2で低レイテンシ・高バンド幅で通信可能。低レイテンシでのサーバー間通信により、大規模な分散学習を高速に実行可能

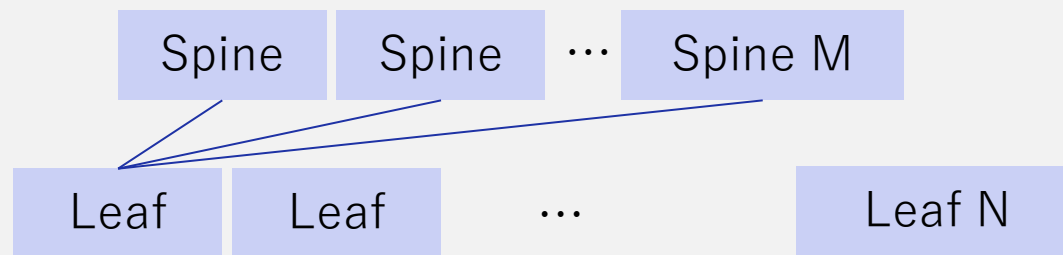
Management Network

NVIDIA Spectrum SN3700 (200GbE)



Compute Network

NVIDIA Spectrum SN3700 (200GbE)



Storage Network

大規模ネットワークの安定運用を支える配線チェックシステム

光ケーブルは数千本も存在、結線ミス・故障等があると動作せず。

デプロイメント・運用の容易化を行う配線チェックツール・ラベリングツールをOSSを用いて開発

配線チェックツール

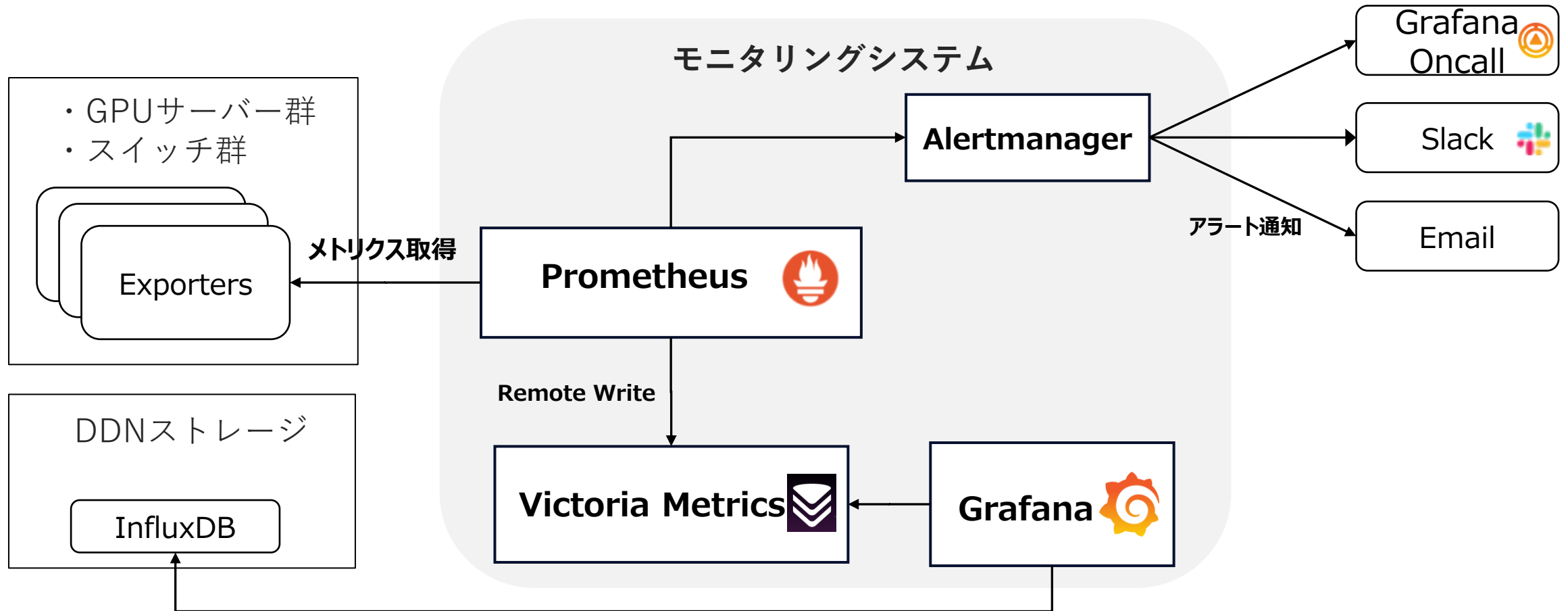
1. ポート管理表から配線情報を自動生成
2. LLDPから実際のスイッチの接続先ポートを認識し実結線情報を生成
3. ポート管理表から生成された配線情報とLLDPからの実結線情報を突き合わせて結線の正しさを自動チェック
4. 配線間違い・ケーブルの故障を自動抽出

ラベリングツール

1. ポート管理表からテプラのラベル情報を自動生成
2. バッチでラベルをプリント

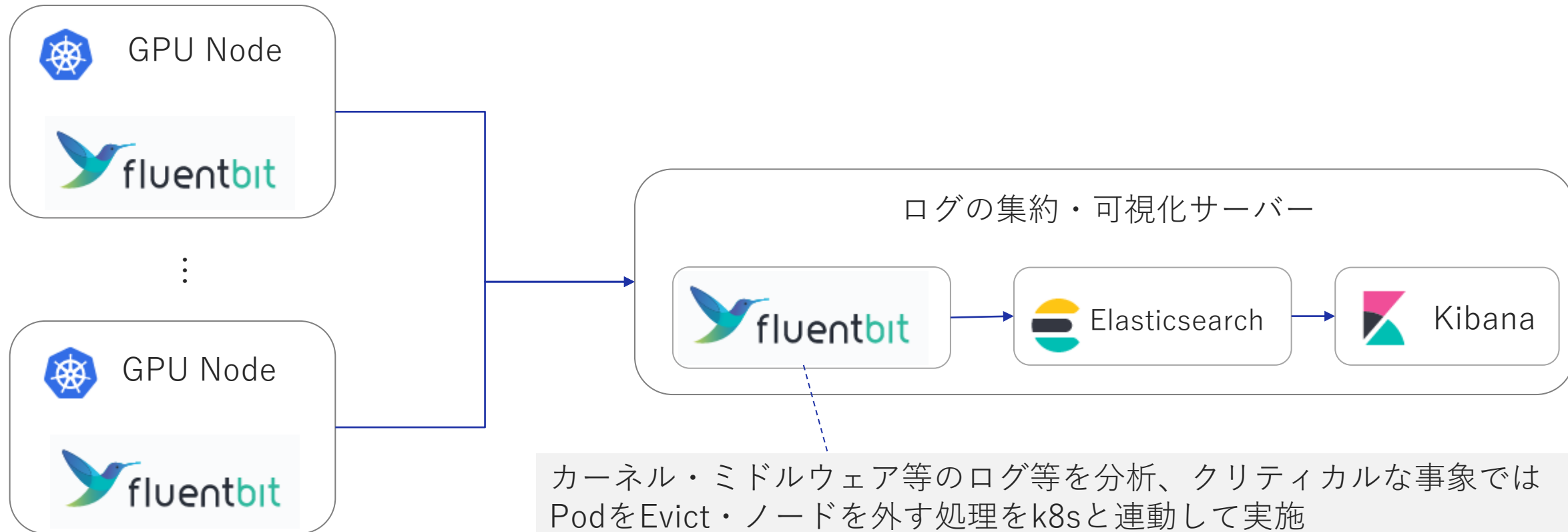
効率的・安定的な分散学習を支えるモニタリングシステム

各種サーバーのリソース、Spine/Leaf等のスイッチ（RoCEv2系等）、GPUサーバー間の通信量、DDNストレージのスループット・メタデータ性能、等をCustomのExporterをGoで開発して全体をモニタリング



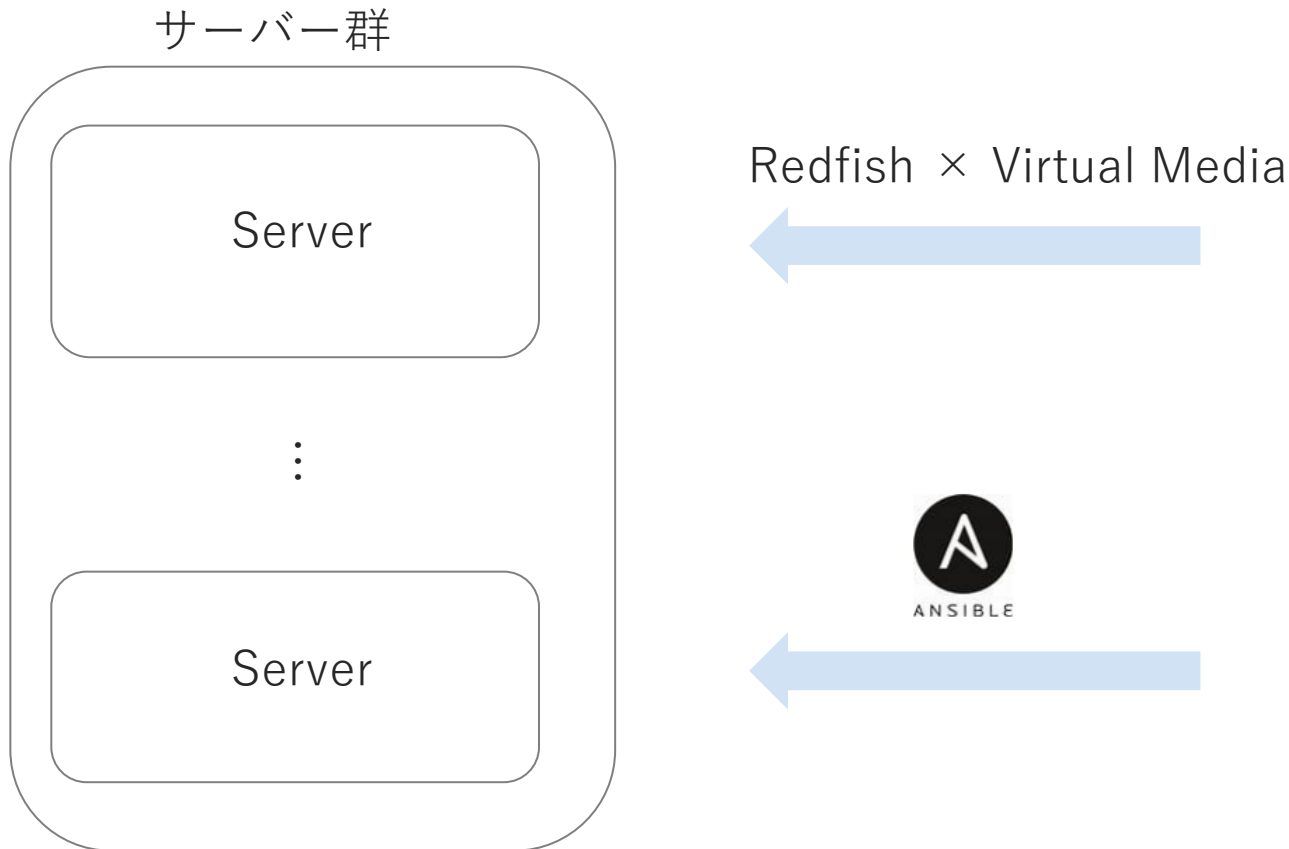
AIスパコンの安定運用を支えるロギングシステム

分散システムでは必ず何かが故障、全ホスト・コンテナのログを収集・分析、即時対応を可能に、EFKのスタックを採用し拡張はGoで開発、異常ノードを即時離脱させ、全体の安定性を確保



多数サーバーの効率的な構築を支えるプロビジョニングシステム

OSのISOイメージを動的に生成、Redfish API と Autoinstall を使い OS を自動インストール
OSインストール後に Ansible で環境構築し、サーバー環境の構築を自動化



① OSの自動インストール

- ✓ **Autoinstall** 入りの OS の ISO イメージを動的生成
- ✓ **Redfish+IMPI** で仮想ディスクとしてマウントして自動インストール
- ✓ DHCPサーバーいらず、シンプル構成

② 環境の自動構築

- ✓ **Ansible+AWX** で自動構築
- ✓ 環境更新が多いため、Goldenイメージ方式ではなく、Ansible側に寄せ自動化

NECのAIスパコンのセキュリティを支えるシステム

KubernetesのAdmission Webhook等の拡張による認証・認可

Falcoによるロギングと危険オペレーション検知による自動停止・アラート

KubernetesのAdmission Webhookを開発して認証・認可

- ✓ 権限の絞り込みなどにより、安全な利用を可能に

Falcoによる操作ログの監視で、危険オペレーションを検知

- ✓ 操作ログを残すことにより後からの監査も可能に
- ✓ 危険操作の強制停止・アラートにより安全性を強化する枠組みを開発
- ✓ Goで拡張を容易に追加可能にすることで、システムの安全性を継続強化

まとめ

まとめ

NECのAI研究スパコンでは、OSSをフル活用しPFエンジニアリングを推進。
優れたOSSの文化を組織に宿し、高度なエンジニアリングで、生成AI研究を加速

- ◆ NECではAIスパコンを用いて、全研究所で生成AI研究の開発を加速している
NECのAI研究者数百名が常時利用しており、利用率は100%である
- ◆ NECのAIスパコンは、先進的で高性能な計算機基盤とKubernetesを中心とした拡張性の高い柔軟なソフトウェア基盤で構成される
- ◆ 全面的にOSSを採用し、下から上まで基盤・運用開発を全てGoに統一して開発することで、高いエンジニアリング効率でAIスパコンを進化させ先進AI研究を実現

NECのAIスパコンに関する社外発表情報

NECのAIスパコンに関する技術的内容や開発の狙いなどの情報を発信しています
ご興味ある方は、是非ご覧になってください。



AIスパコンでNECを世界トップレベルのAIテックカンパニーへ
<https://jpn.nec.com/rd/lablife/forstudents/interviews/aisupercomputer.html>

NEC北野によるAIスパコン開発の狙い、技術的特徴等のインタビュー記事

NEC \Orchestrating a brighter world



社会価値を創造するNECのAIスーパーコンピュータ
<https://jpn.nec.com/rd/aisupercomputer/index.html>

NECのAI研究用スーパーコンピュータの技術スペック等を解説



NVIDIA/DDNのイベントの基調講演（技術説明）
<https://www.youtube.com/watch?v=DmXRIae1Q8k>

NEC北野によるAIスパコンのアーキテクチャの解説

Truly Open, Truly Trusted -This is NEC.

All Share Benefitの世界を実現する

Truly Open

オープンイノベーションを加速させる
オープンなエコシステムをリードする

Truly Trusted

ミッションクリティカルを支える力
技術を正しいことに使う姿勢