

MPLS Japan 2025

「AIとAIネットワーク」セッション

IOWN APNを活用した 分散データセンタ検証事例



2025年10月30日

NTTドコモビジネス株式会社
イノベーションセンター IOWN推進室

張 暁晶

自己紹介



NTTドコモビジネス株式会社
エバンジェリスト

張 曉晶

ちょう ぎょうしょう
Xiaojing ZHANG

仕事

2007年～

NTT研究所にてソフトウェア工学や **ソフトウェアテスト効率化** の研究開発

2016年～

NTTドコモビジネスにて **クラウドインフラの設計・構築・運用** の技術検証

- GPU / FPGA / OpenStack / Kubernetes
- AWS, AWS Outposts/ Azure, Azure Local / Google Cloud, GDCE / CCoE

2023年～

IOWNのコンピューティング領域 の技術評価・検証

- APN (All-Photonics Network) の特性把握やユースケース検証
- IOWN光コンピューティング, DCI (Data Centric Infrastructure) の評価・検証 etc.

趣味

読書、映画鑑賞、子育て

社名変更について

最先端テクノロジー・イノベーションを通じ、企業と地域が持続的に成長できる
自律分散型社会を支える「産業・地域DXのプラットフォーマー」へ



NTTコミュニケーションズ株式会社



NTTドコモビジネス株式会社

- 01 IOWN APNの特徴とユースケース
- 02 AIインフラと分散データセンタ
- 03 最新検証事例：GPU over APN
分散学習 / 分散推論 / 拠点間データ転送

01 IOWN APNの特徴とユースケース

「IOWN®」は、NTT株式会社の商標又は登録商標です

IOWNに取り組む背景

現状

生成AI等のICTの進展による
けた違いのデータ量増加など、
電力消費の急増

世界のデータセンタの年間消費電力
(2018年→2050年)
約2,600倍※

既存の情報通信システムでは
伝送能力と処理能力の双方に限界に
+ 半導体競争の過熱

2032年に向けて

持続可能な社会を実現するため
技術革新による新たな価値創造

環境負荷が大幅に下がる
ネットワークやコンピューティング
インフラの高度化の実現にむけて、

ネットワークから端末まであらゆる場所
に**光電融合**デバイスなどの
光（フォトニクス）関連技術を活用した
次世代のインフラへ

IOWN®

※出典：国立研究開発法人科学技術振興機構低炭素社会戦略センター

「IOWN®」は、NTT株式会社の商標又は登録商標です

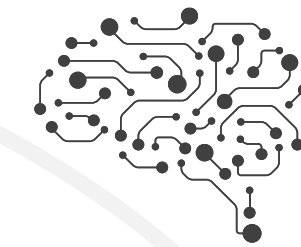
IOWN (Innovative Optical and Wireless Network) とは

つながる。驚きを。幸せを。

スマートな社会の実現に向けた光（フォトンクス）関連技術を活用した次世代情報通信基盤

デジタルツイン
コンピューティング(DTC)
Digital Twin Computing

実世界とデジタル世界の掛け合わせによる未来予測等を実現

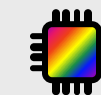


コグニティブ・
ファウンデーション (CF)
マルチオーケストレーター

あらゆるものをつなぎ、
その制御を実現

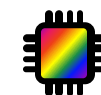
IOWN光コンピューティング

光電融合技術を活用した次世代コンピューティング
Data Centric Infrastructure (DCI)



オールフォトンクス・ネットワーク(APN) 

ネットワークから端末まで、すべてにフォトンクス（光）ベースの技術を導入



光電融合技術
(光プロセッサ)



APNの特徴

つながろう。驚きを。幸せを。

 **docomo Business**

通信ネットワークのすべての区間で光波長を占有することで「大容量」「低遅延」「低消費電力」を実現

IOWN APN

All Photonics Network

低遅延

エンドエンド遅延※1

大容量高品質

伝送容量※2

低消費電力

電力効率※3

**最終目標
(2032年)**

1/200

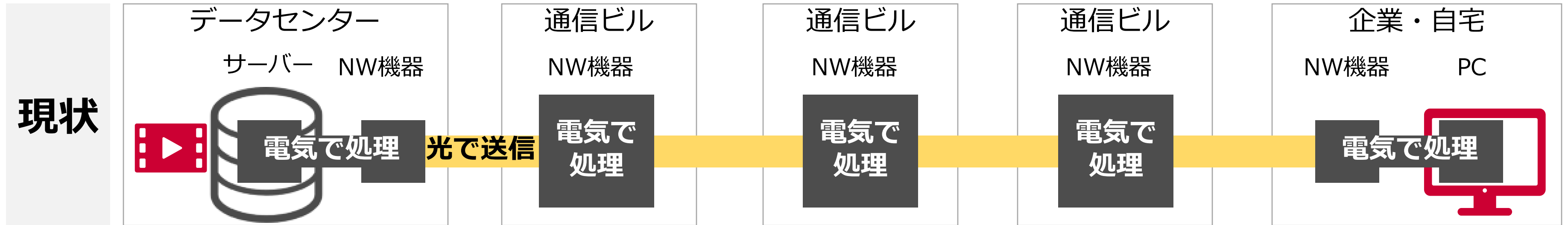
125倍

100倍

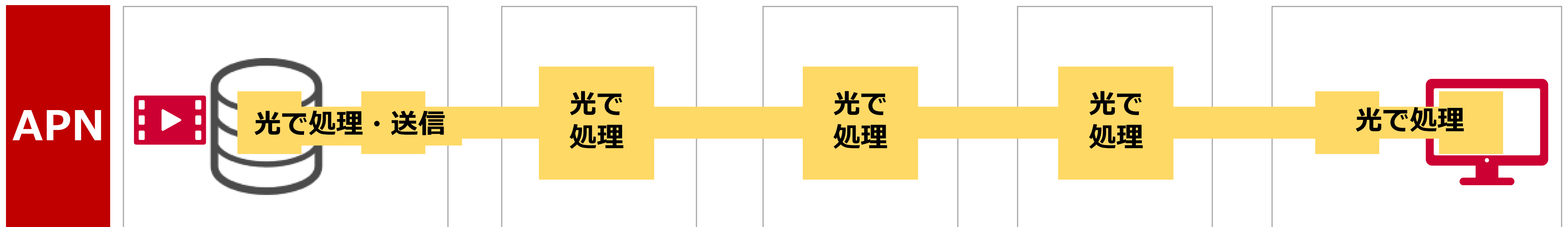
※1 同一県内で圧縮処理が不要となる映像トラフィックでのエンドエンドの遅延の目標値 ※2 光ファイバー1本あたりの通信容量の目標値 ※3 フォトニクス技術適用部分の電力効率の目標値

APNとこれまでの光ネットワークとの違い

電気で処理していたデータを光で処理することで、**大容量・低遅延・低消費電力**を実現
これまでは：データを電気→光→電気・・・と変換、多くのエネルギーを使用



将来的には：サーバからPCまでデータを光で処理、無駄なエネルギーを低減



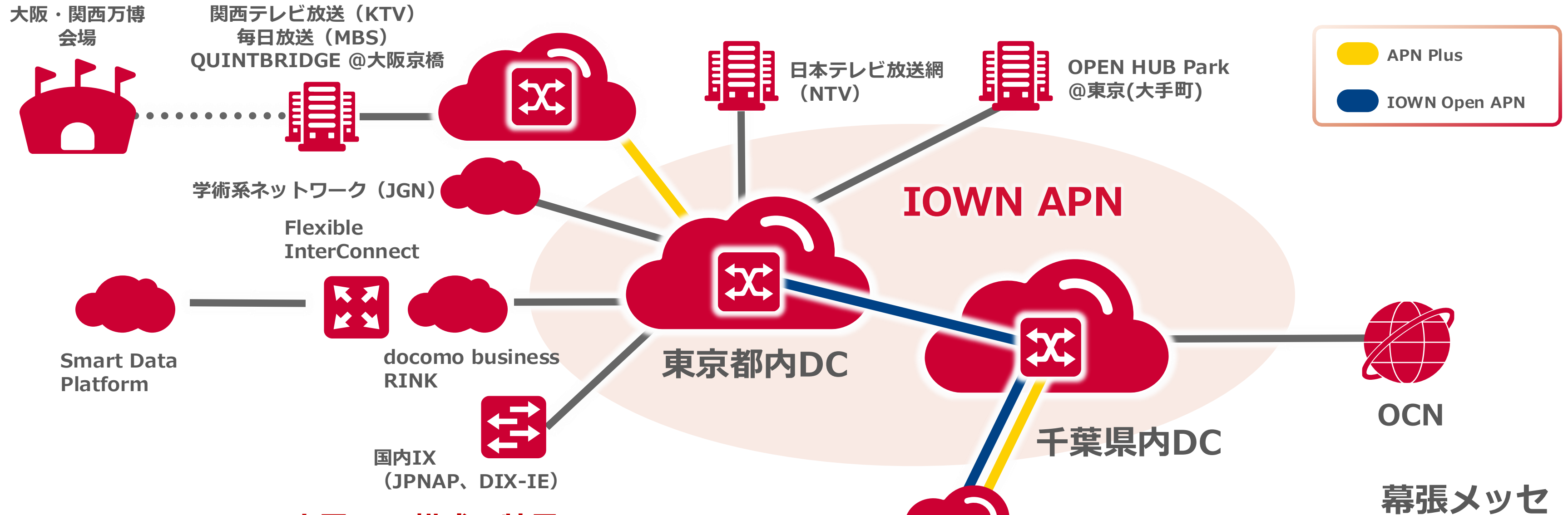
光の波長パスでユーザ拠点間を接続

Interop Tokyo 2025 ShowNetでの取り組み

のべ来場者数13万6000人超

つながろう。驚きを。幸せを。

NTT docomo Business

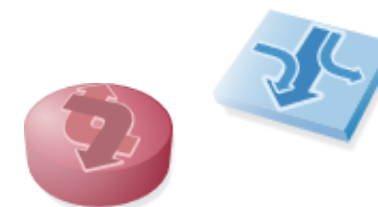


Interop Tokyo 2025出展APN構成の特長

- ・ 光波長多重化による大容量伝送
- ・ All-Photonics Network技術による低遅延・低ジッタ接続
- ・ マルチベンダな装置による波長を収容可能な構成
- ・ ShowNetへ対外接続を提供
- ・ 遠隔拠点との連携デモなどの実アプリケーション



SHOWNET is your network.



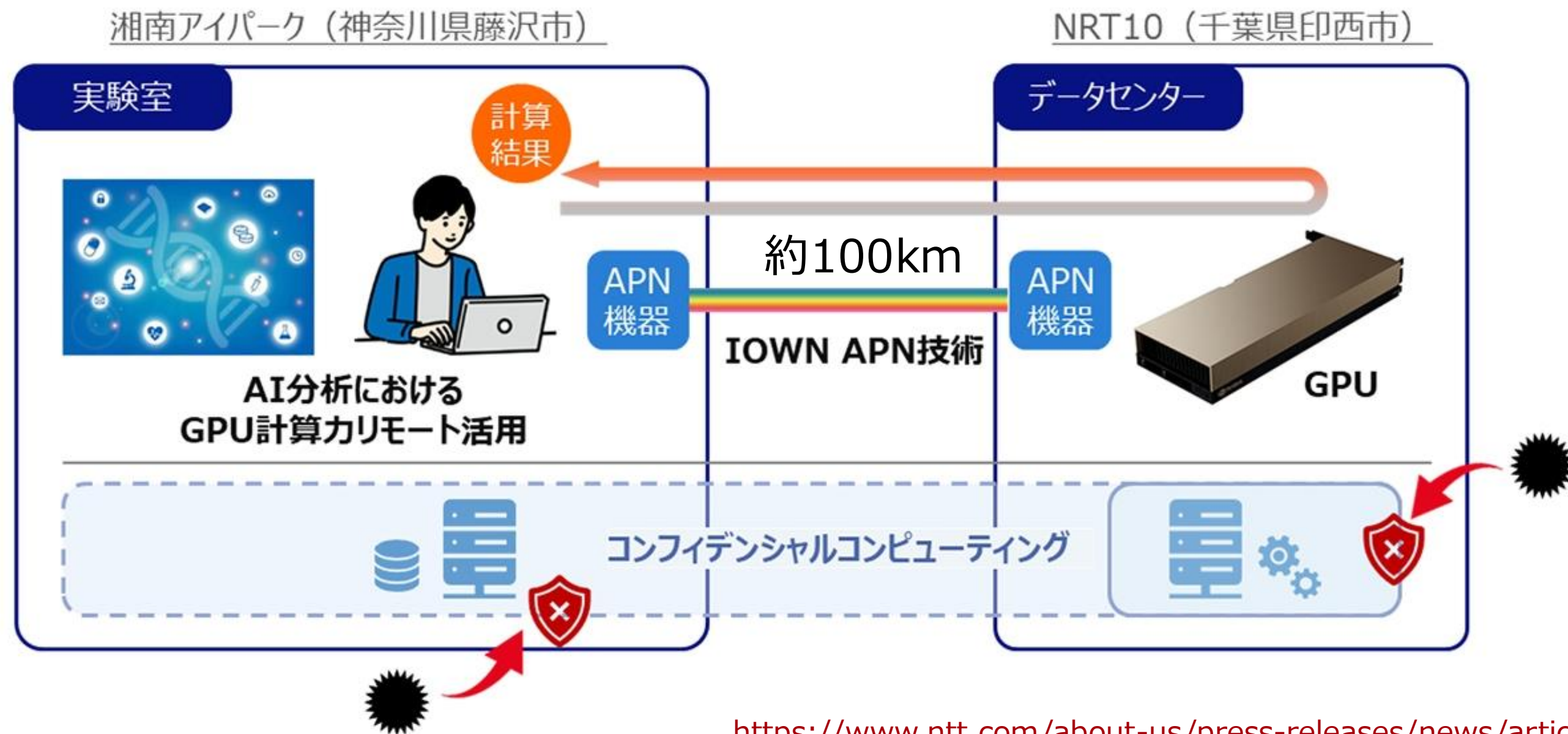
MOC



製薬・創薬研究拠点：湘南アイパーク x IOWN

製薬会社の研究開発拠点と、AIの計算基盤となるデータセンターをNTTの次世代通信基盤「IOWN」でつなぎ、データを高速かつ安全にやりとり

■製薬・創薬業界が求める**セキュリティ要件**を満たせるかの
コンフィデンシャルコンピューティングの実行可能性検証



<https://www.ntt.com/about-us/press-releases/news/article/2025/0217.html>

スマートシティ：千歳市の自動運転バス実証実験

つなごう。驚きを。幸せを。

NTT docomo Business

千歳市
自動運転バス
実証実験



千歳市
City of Chitose

＜期間＞
令和6年11月18日(月)から
27日(水)まで
※土日運行なし

実証実験の目的

全国的に少子高齢化、人口減少が進む中で、公共交通分野においては、運転手の高齢化、人手不足が深刻化しています。本市においても、運転手不足などから、路線バスにおいて減便されており、今後、運転手の年齢構成から、運転手不足はさらに加速するものと推測しています。
そこで、本市においては、バスの運転を自動化することにより、運転に係る人数を削減し、持続可能な地域公共交通を確保するため、自動運転バスの実証実験を実施しています。

実施体制

実施主体： 千歳市
City of Chitose
参画企業：
 A-Drive  AISAN  北海道中央バス  千歳相互観光バス株式会社
 docomo business  NTT Communications  TOKAI RIKAI  SOMPOリスクマネジメント
 NDS 日本データサービス株式会社

IOWNや5Gワイドなどを
活用した路線バス自動運転
実証を千歳市で実施

公共交通機関のさらなる
利便性向上・運転手不足
対応に向けて、

IOWNなどの
最先端デジタルテクノロジー
を活用

その先のさらに住みよい
まちづくりに貢献

遠隔監視システム

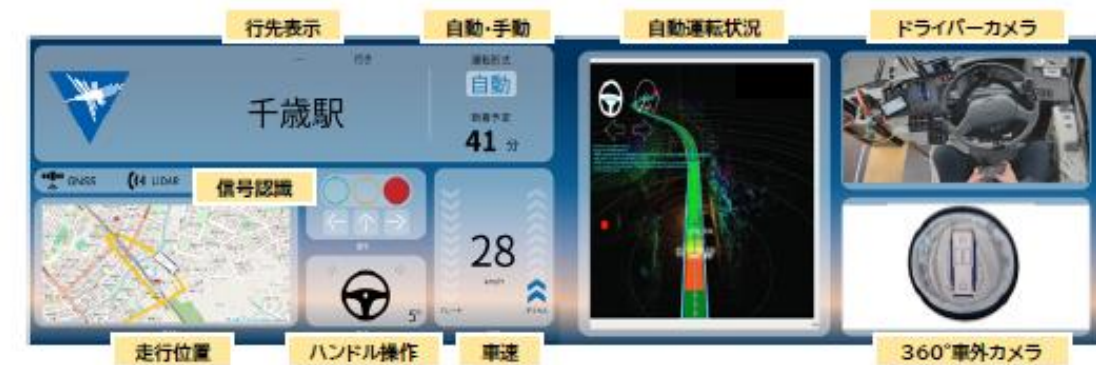
- 自動運転バスの車室内外で撮影されたカメラ映像をモバイル回線(5G、LTE等)を介し遠隔地に伝送します。その際に課題となる回線速度による映像伝送の遅延やデータ圧縮による画像劣化に対し、カメラ映像を画像処理技術により統合化することで、高画質で遅延のない伝送が可能です。



実証期間の10:00～16:00、千歳駅東口待合室、
千歳市役所 市民ロビーに設置

デジタルサイネージ

- 乗客に対し、自動運転制御状況を分かりやすく知らせるため、センシング結果やドライバーの手元を撮影したカメラ映像などを表示するディスプレイを車室内に搭載し、自動運転バスと連携する車両情報(アクセル/ブレーキ、舵角等)を表示します。



ニュース 2024年11月15日:路線バスの運転手不足に対応するIOWNや5Gワイドなどを活用した路線バス自動運転実証を千歳市で実施 | NTTドコモビジネス 企業情報

大阪万博 x IOWN

世界初、リアルタイムに空間が伝送される体験

1970年 吹田 万博記念公園
(電気通信館 跡地)



IOWN APN



3D空間・触覚

2025年 夢洲 万博会場
(NTTパビリオン ZONE2)



リアルタイムに空間を
センシング

高速・低遅延
伝送

忠実に空間として
再現

お問い合わせの多い業種・業界

2030年をターゲットに
中長期ビジョン／デジタル戦略／インフラ戦略／共創
新たな収益獲得、ビジネスモデルなどを視野に

スマートシティ

- ・スマートサービスを支える基盤
- ・デジタルツイン
- ・エンターテインメント

金融

- ・取引の公平性の担保（確定遅延）
- ・遠隔拠点間の大量データ取引
- ・事業継続（BCP）への対応

製造

- ・FA機器の遠隔リアルタイム制御
- ・工場拠点間ネットワーク

放送・メディア

- ・リモートプロダクション
- ・高精細映像配信
- ・没入型メディア

IT事業者

- ・次世代ネットワーク
- ・データセンター間接続
- ・次世代コンピューティング

電子・電機

- ・デバイス×光電融合技術
- ・光半導体のエコシステム

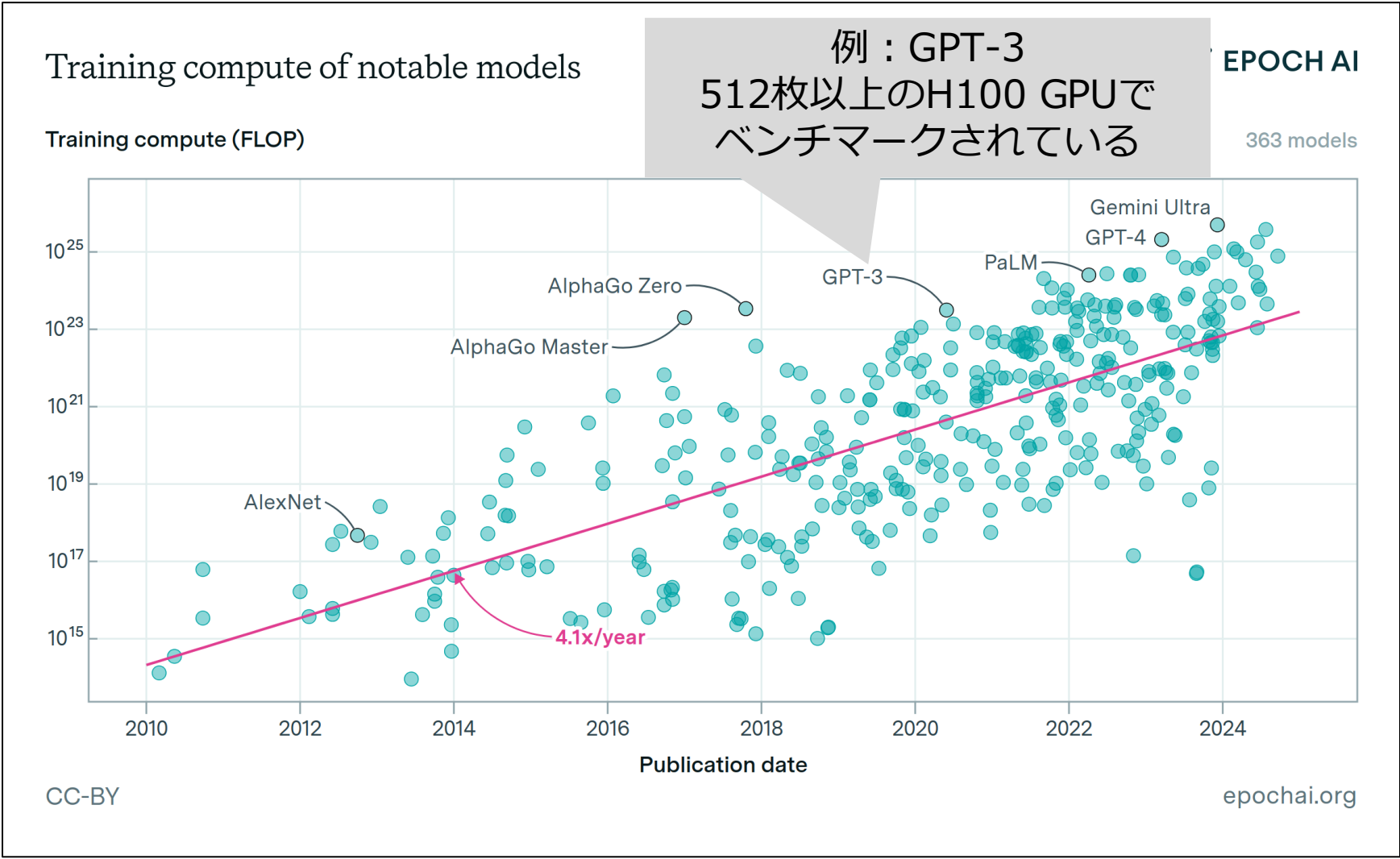
今後の活用イメージ（将来像）

02 AIインフラと分散データセンタ

AI向けGPUインフラ需要増のトレンド

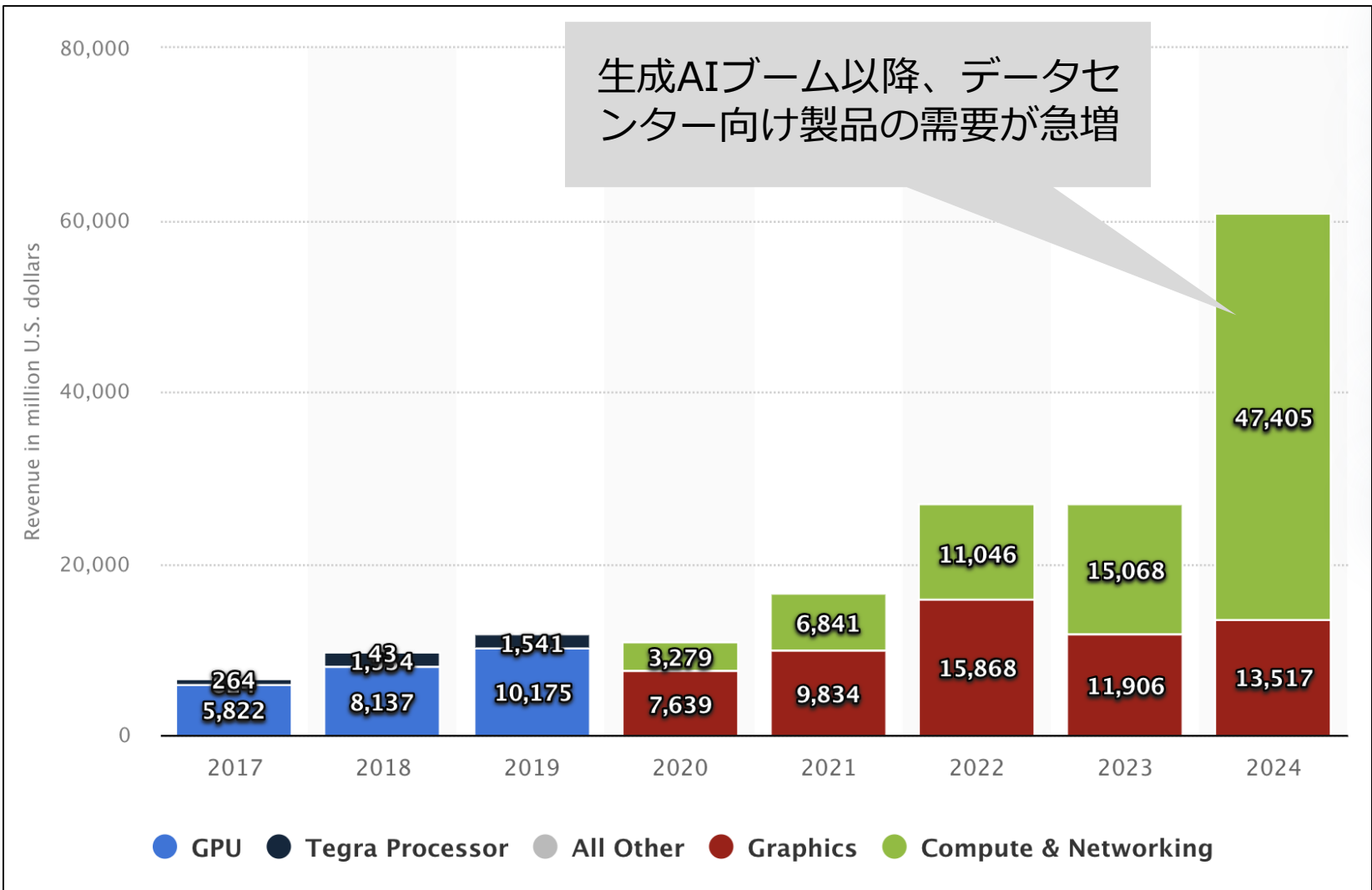
生成AI、データ利活用、画像処理などの分野で、**多くの計算資源が必要**となっている
また、1台のGPUサーバーでは搭載できるGPUの数に限りがあるため、**複数台のGPUサーバーを並べて同時に使う**シーンがますます増えている

AIモデル学習にかかる計算量の推移



<https://epochai.org/data/notable-ai-models>より引用・注釈追加

GPU業界最大手NVIDIA社の売上高の推移



<https://www.statista.com/statistics/988034/nvidia-revenue-by-segment/>より引用・注釈追加

AI向けGPUインフラの特徴

学習・推論等のAIワークロードのためのICTインフラの特徴

1. 高い計算能力と並列処理性能

GPUや専用アクセラレータを多用し、膨大な演算を高速に処理する

2. 大量の電力消費と発熱

特に学習時は消費電力・排熱が大きく、インフラ側の電源・冷却設計が重要

3. 高速インターコネクトと大容量ストレージ

巨大なデータセットやモデルの保存・転送のため、ストレージやネットワーク帯域も高性能が求められる

4. 柔軟なスケーラビリティと運用管理

ワークロードの変動や新技術への対応のため、インフラの拡張性・自動運用機能が重視される

インフラ制約の観点では

ラックあたり 電力 密度	最新AIラックは250-900kW/ラックに達する（2026-27年予測※1） →従来DCの電力供給能力を超過
冷却 能力	通常7-20kW/ラックの空冷ユニットでは120-140kWの熱除去が不可 →液冷対応不可時は負荷分散が必要
床荷重 制限	液冷ラックの荷重が、従来DCの標準床荷重を逼迫または超過する可能性

→「やむをえず」分散へ

※1. <https://www.tweaktown.com/news/101799/ai-servers-of-the-future-rack-density-1000kw-with-nvidias-next-gen-rubin-ultra-gpus/index.html>

事業継続・災害復旧の観点では

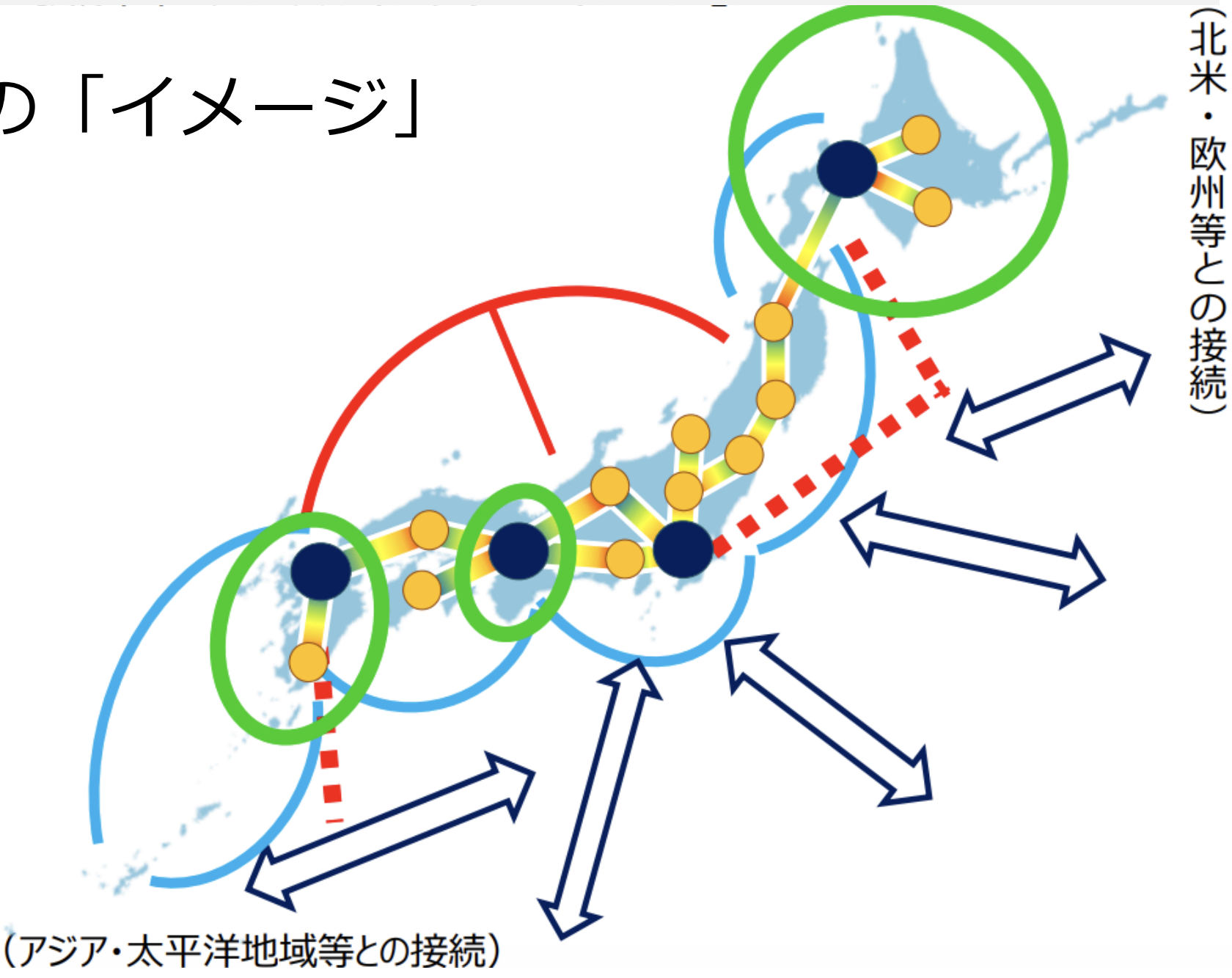
AIワークロードの 継続性 確保	拠点障害時に学習や推論を即時再開し、サービス停止や学習中断リスクを最小化したい
巨大なAIデータセットやモデルの バックアップ	複数拠点にデータやモデルを分散配置し、災害時のデータ消失リスクを低減したい
データ主権 要件や法規制への柔軟な対応	企業ごとや地域ごとにデータやモデルを管理し、企業/国の規制やガバナンスに適合させたい
リソース有効活用と分散処理による精度向上	型落ちや零細なリソースも含めて活用することでAIサービスの効率性を向上させたい

→ 「あえて」 分散へ

AIインフラにかかわる総務省見解

AI・半導体・オール光ネットワーク・量子コンピューター等の技術を軸に、AI革命に続く大きなパラダイムシフトが数年以内に起こり得ることを念頭に置いた柔軟性の確保が重要。

2030年代の我が国のデジタルインフラの「イメージ」



分散データセンター x IOWN

背景/課題

首都圏へのデータセンター一極集中

AIによるGPUニーズへの対応

Why IOWN

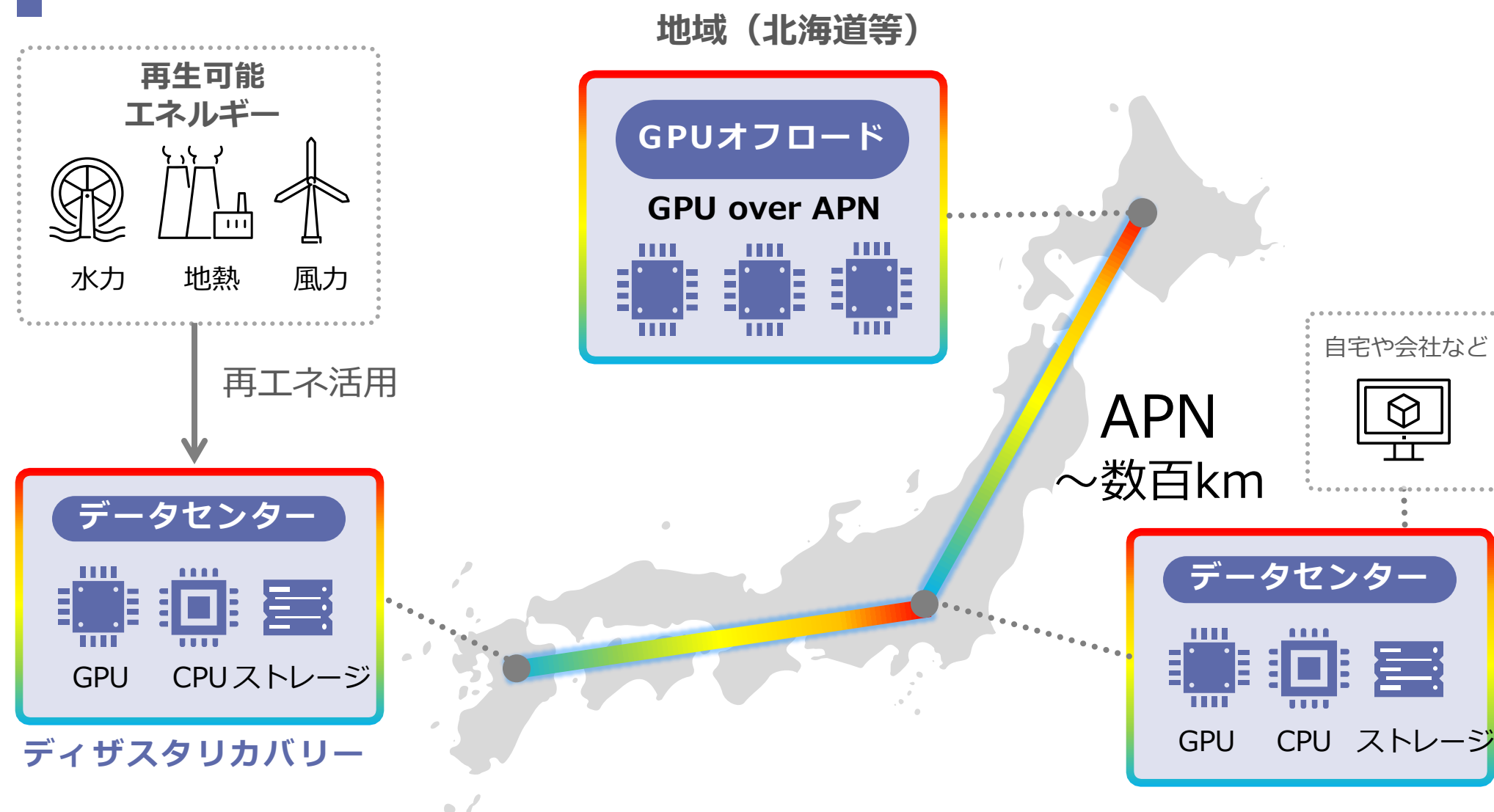
分散されたデータセンター間を
IOWN APNで接続

レジリエンスの向上

環境に配慮した電力使用が可能に

柔軟な拡張と分散DCの一体運用

構成イメージ



03

最新検証事例：GPU over APN

分散学習 / 分散推論 / 拠点間データ転送

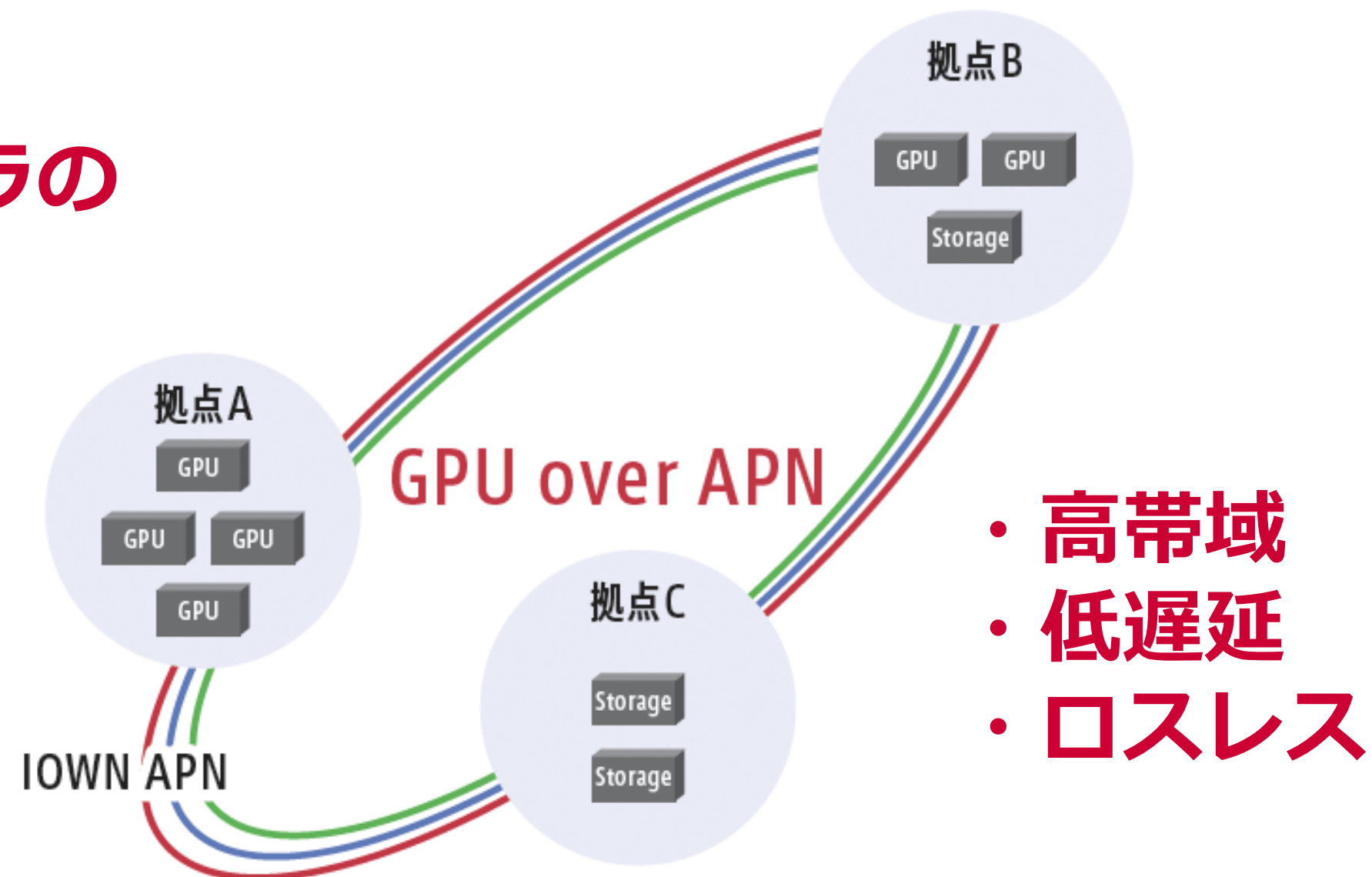
GPU over APN コンセプト

コンセプト/
訴求ポイント

計算資源やデータの最適な分散配置を考慮した柔軟なGPUクラウドを実現

- 処理量の変動に応じてオンデマンドにGPUリソースを確保
- 利用者の拠点から移動できない機密度の高いデータの取り扱いが可能
- 1つのデータセンターの床面積や電力供給能力に制限されない

AI向けGPUインフラの
分散ニーズに対して
選択肢を提供したい



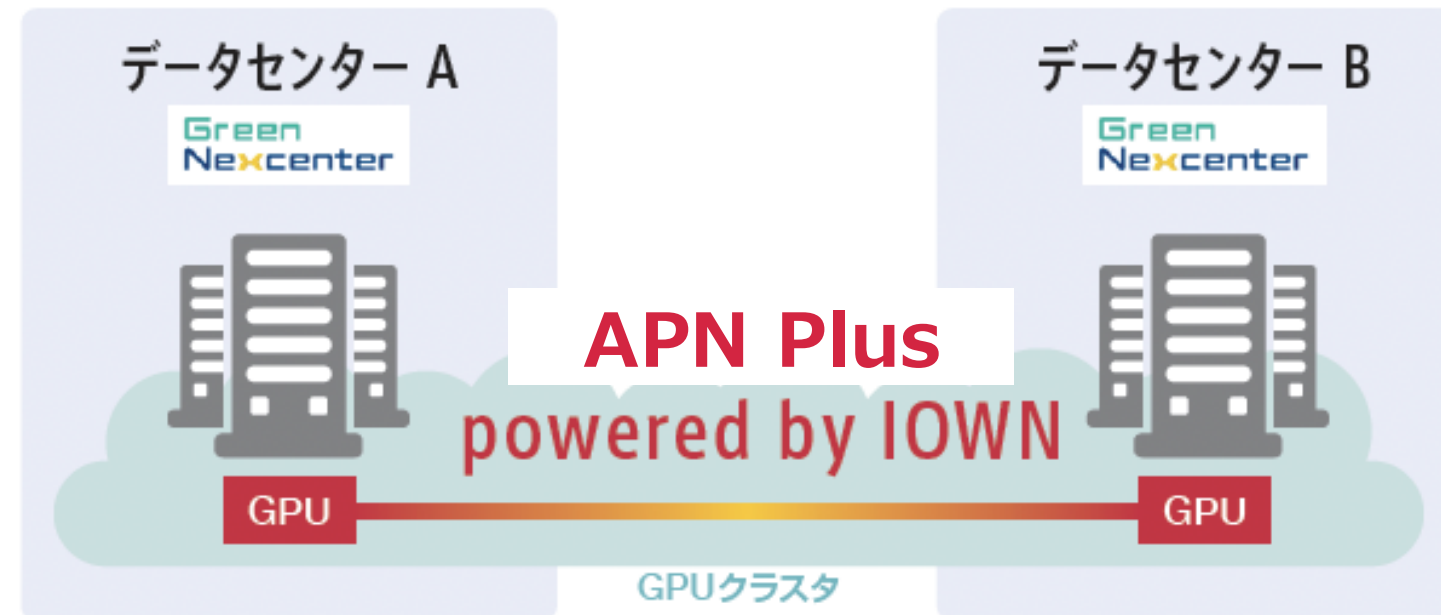
GPU over APN 構成

以下大きく2つの構成方法を想定

複数のデータセンターで
1つのGPUクラスタを構成し、
クラスタ内高速通信を前提と
して計算を実施

クラスタ
拡張

拠点分散

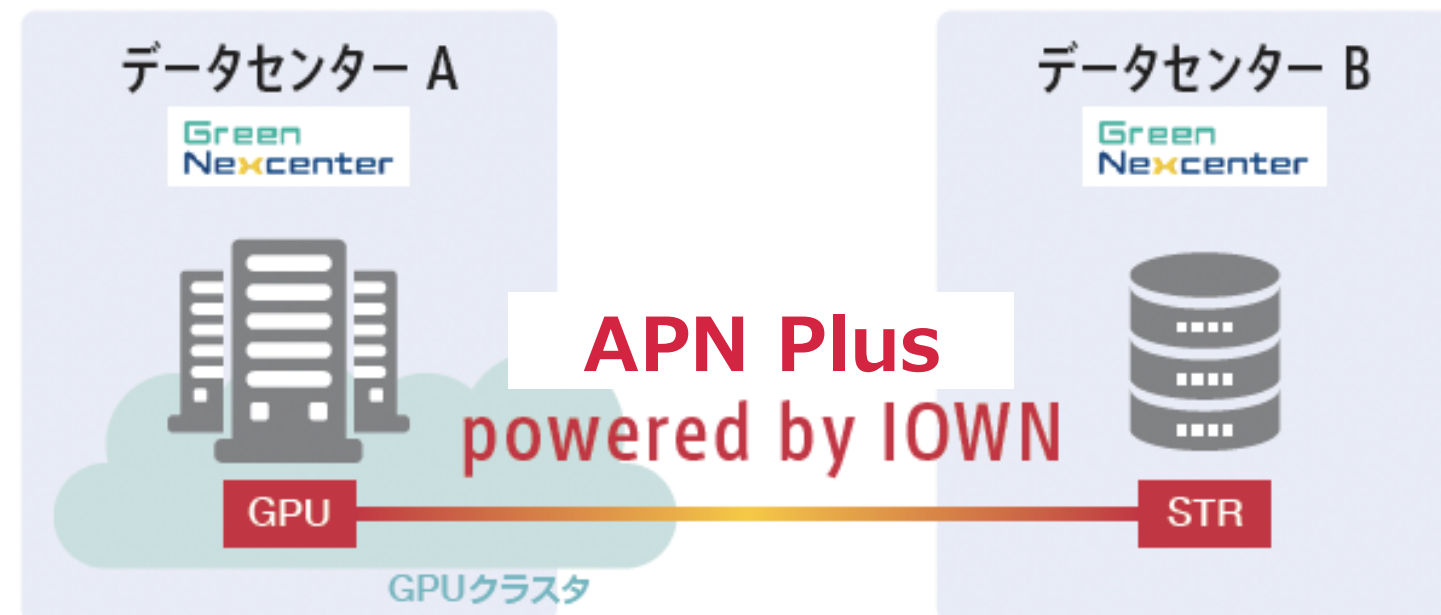


GPUサーバ同士を
引き離してみる

他拠点に置いたストレージ
から高速にデータの読み書き
を実施

LLM学習

大規模
データ

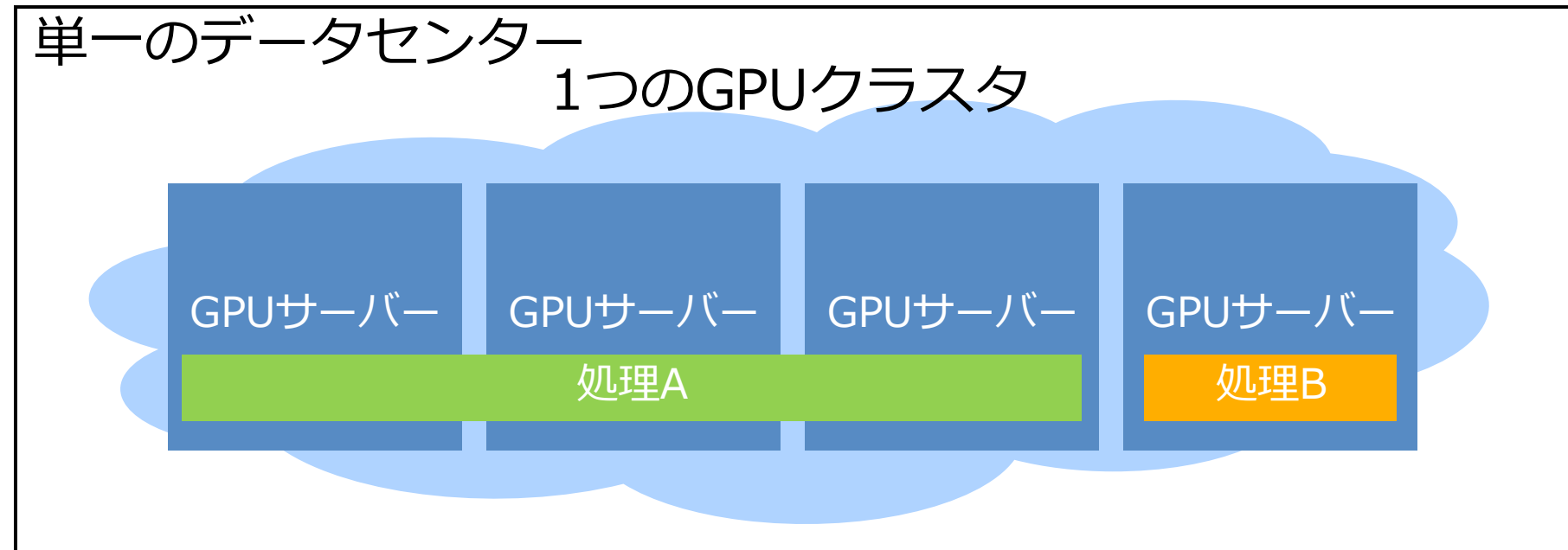


GPUサーバと
ストレージを
引き離してみる

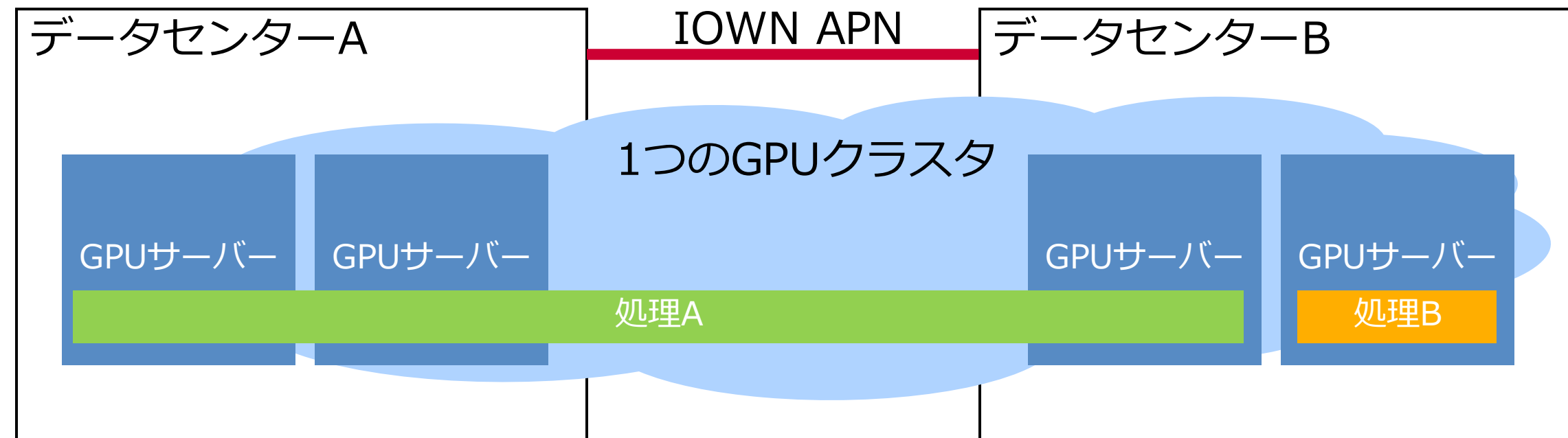
GPUクラスタを分散するとは

本実証におけるGPUクラスタの分散配置とは、
「1つの」GPUクラスタを複数のデータセンターにまたがって配置することを指す

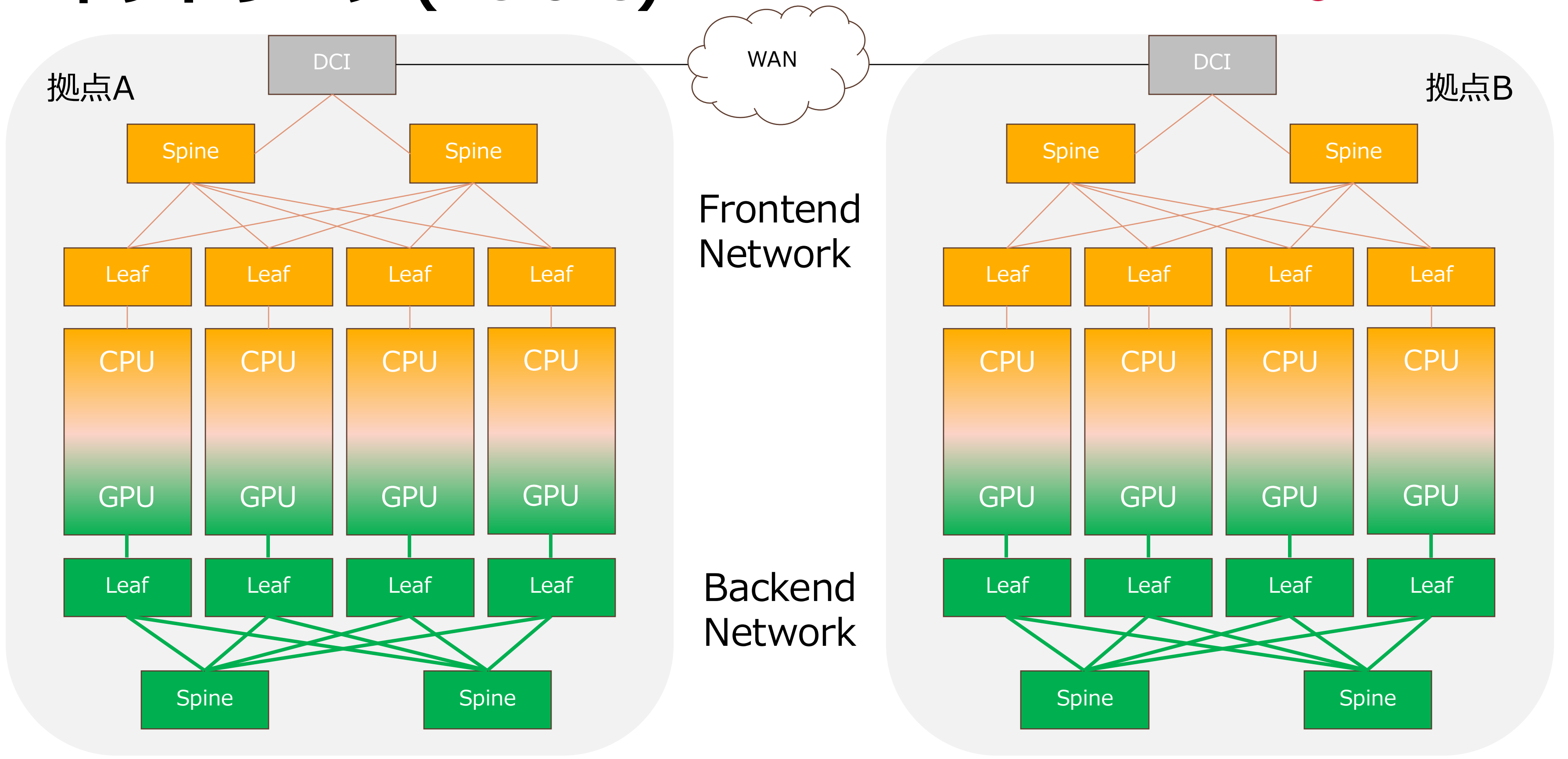
従来



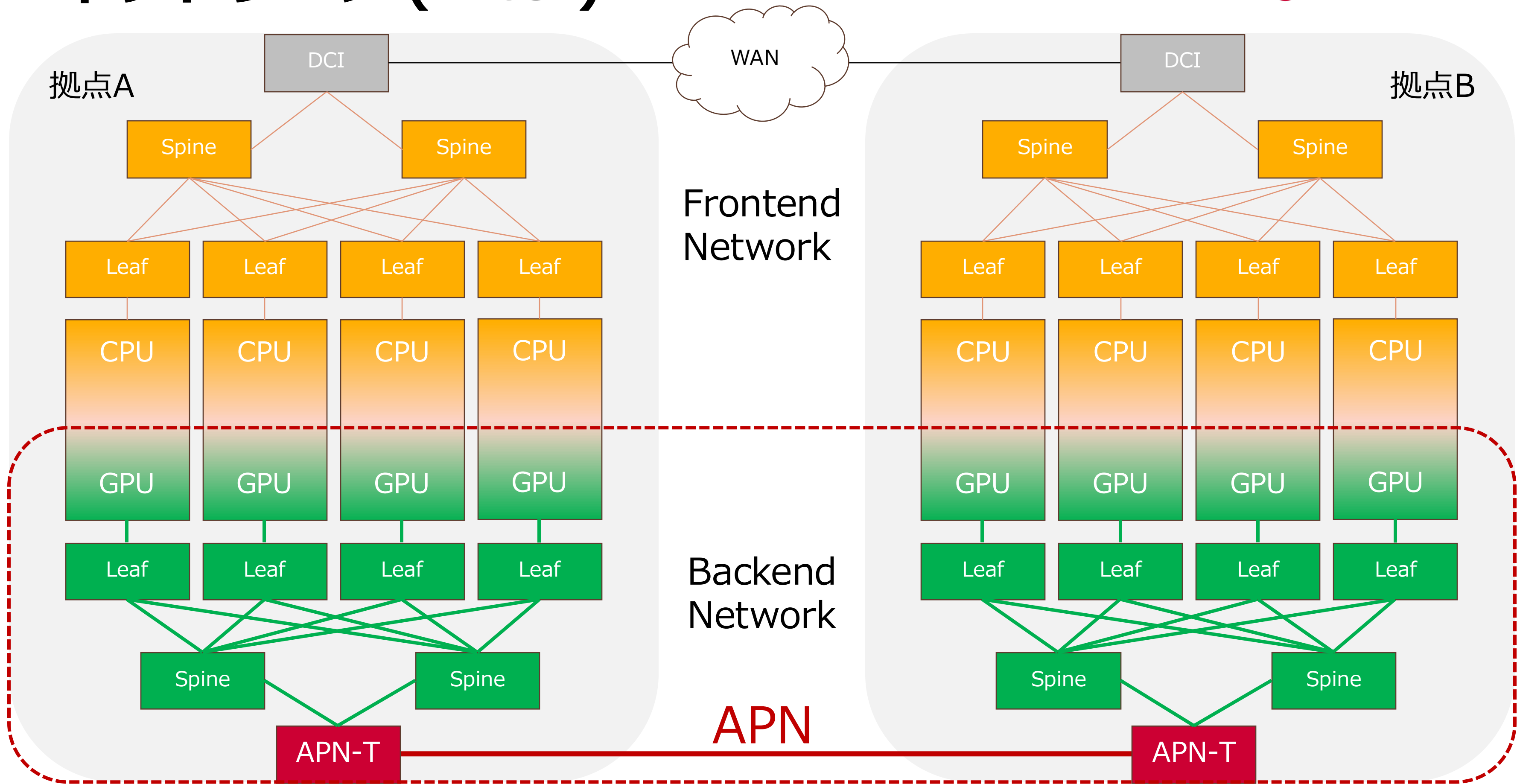
本実証の構想



AIネットワーク (Before)



AIネットワーク (After)



GPU over APN ユースケース

A

AI学習
Training

B

AI推論
Inference

C

拠点間
データ転送
Data Transfer

...

GPU over APN ユースケース

A

AI学習
Training

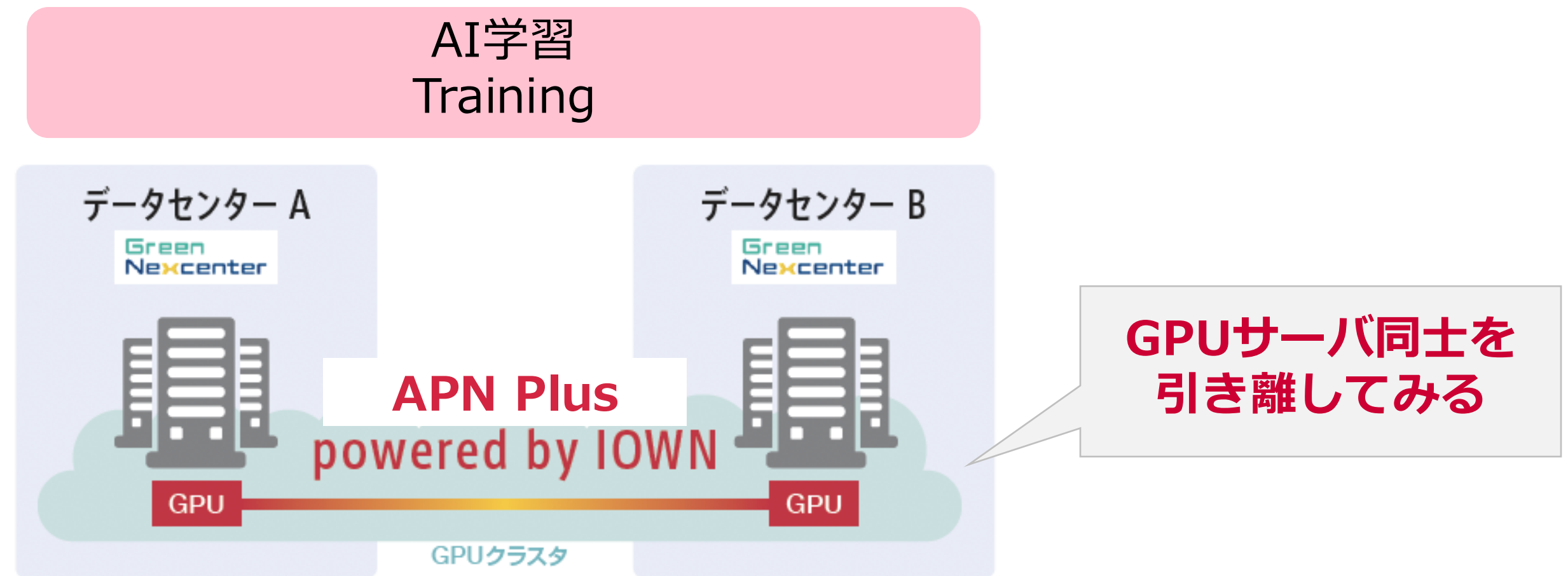
B

AI推論
Inference

C

拠点間
データ転送
Data Transfer

...



AI学習に関する実証

【0: ベース実験】 秋葉原・三鷹の **2拠点** におけるAIモデル学習のスループット性能を測定

2024年10月ニュースリリース <https://www.ntt.com/about-us/press-releases/news/article/2024/1007.html>

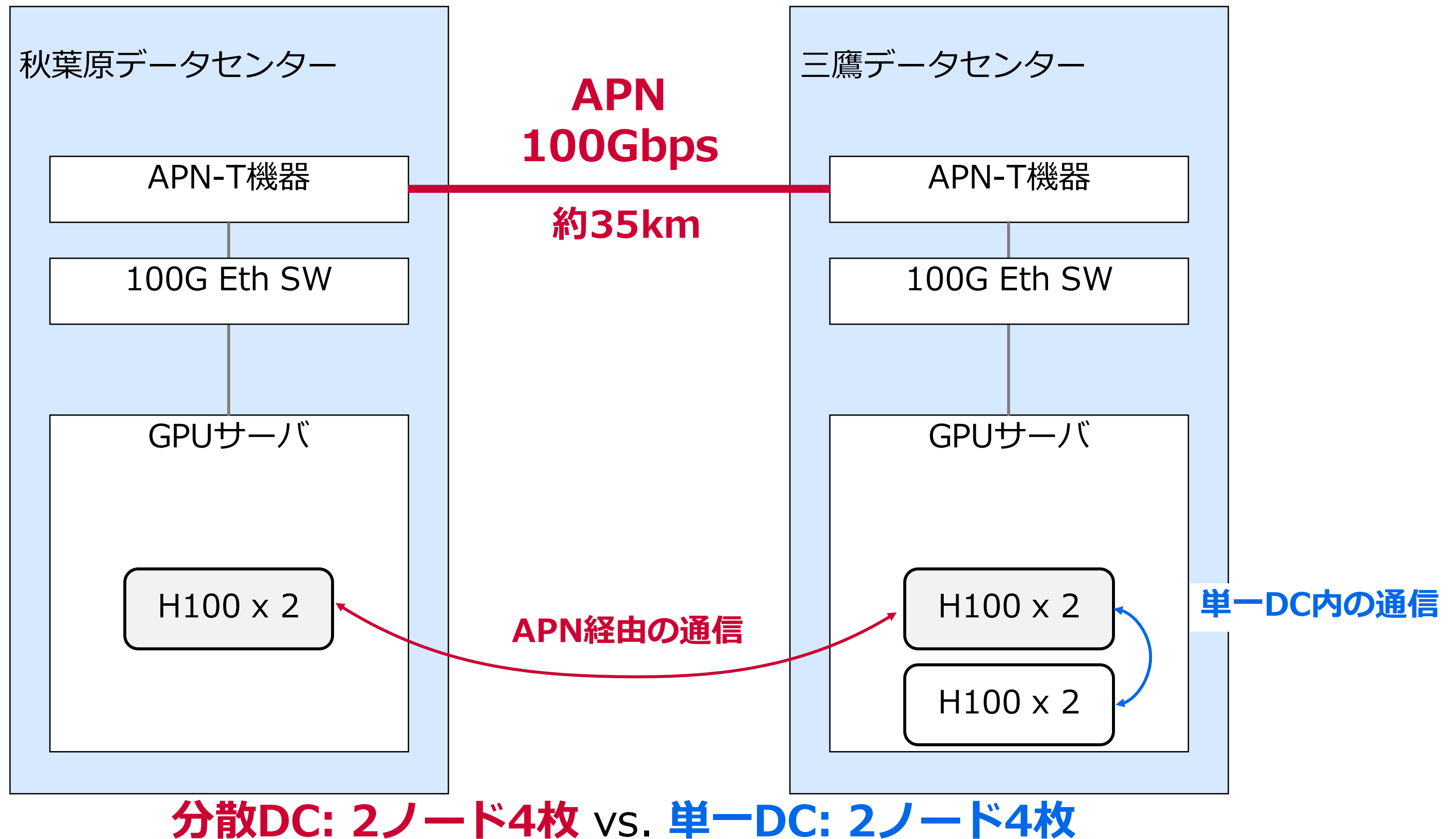
【1: 多拠点対応】 秋葉原・三鷹・川崎の **3拠点** におけるAIモデル学習のスループット性能を測定

2025年3月ニュースリリース <https://www.ntt.com/about-us/press-releases/news/article/2025/0319.html>

【2: 遠距離対応】 3000kmの **超遠距離** を模擬した2拠点間でAIモデル学習のスループット性能を測定

2025年6月ニュースリリース <https://www.ntt.com/about-us/press-releases/news/article/2025/0610.html>

0: 2拠点で分散学習 実証環境

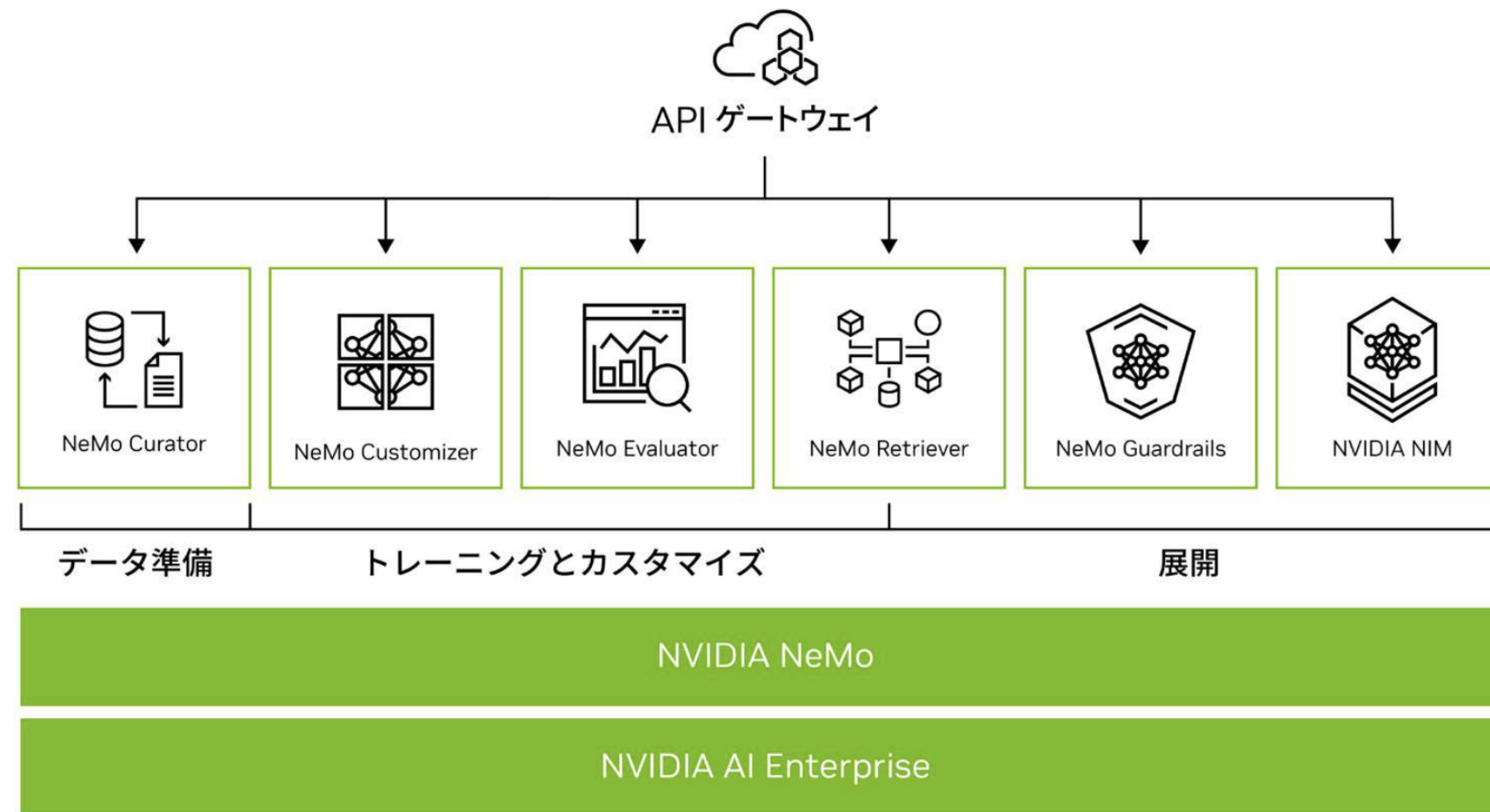


0: 2拠点で分散学習 実証環境

複数GPUサーバーを用いた分散学習に対応したNVIDIA NeMo™ を活用して、実証環境を構築・利用した

NVIDIA NeMo™ :

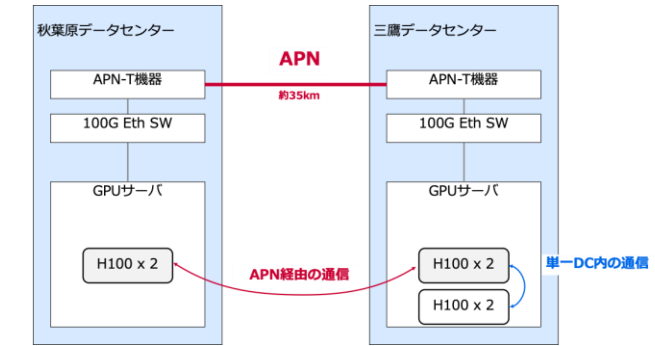
- あらゆる環境でカスタム生成AIを開発できるエンドツーエンドのプラットフォーム
- 正確なデータキュレーション、最先端のカスタマイズ、検索拡張生成（Retrieval Augmented Generation : RAG)、高速化されたパフォーマンスを備えたエンタープライズ対応モデルを提供



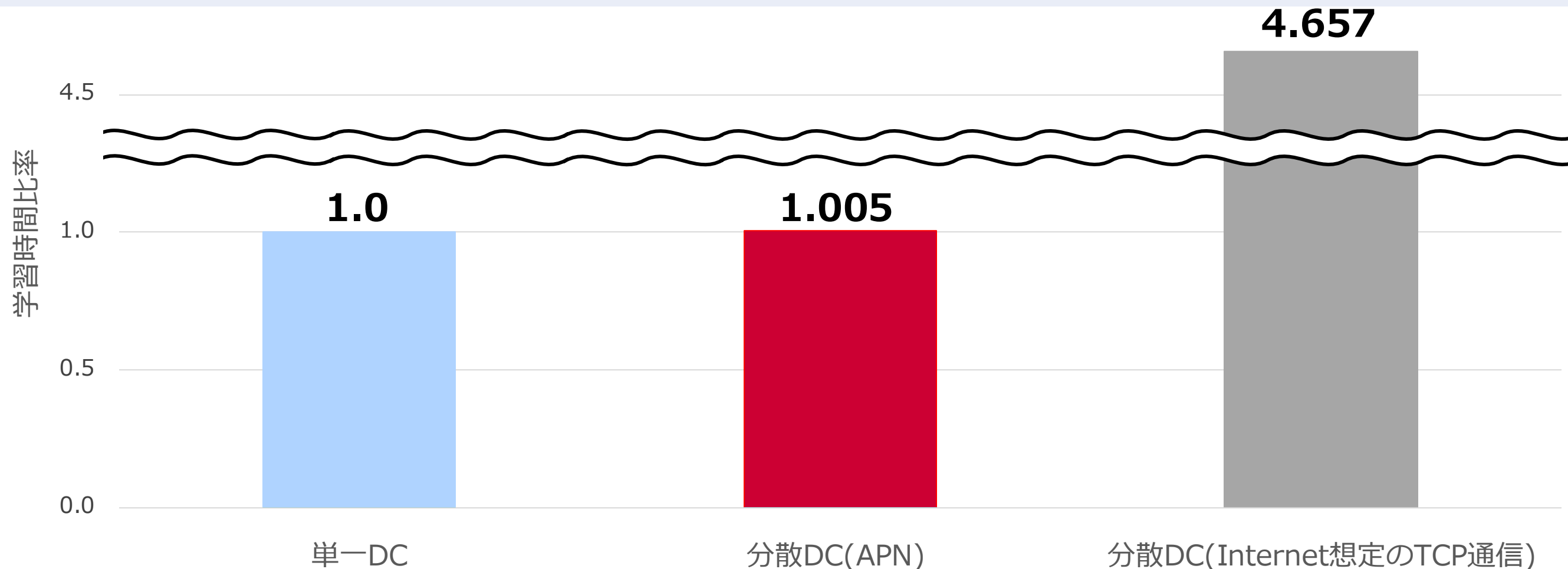
<https://resources.nvidia.com/ja-jp-nemo/nemo-jp> より引用

0: 2拠点で分散学習 実証結果

- **LLM(tsuzumi 7B) の事前学習** を、NVIDIA H100 Tensor コア GPU x2基 x2ノードで実施
- 学習データは各GPUサーバのローカルのディスクに格納
- 処理完了の所要時間（1ステップあたりの時間 x 所定ステップ数）を計算
- 単一DC、分散DC(APN経由※⁰)、分散DC(Internet想定 of TCP通信※¹)、の三者の性能を比較



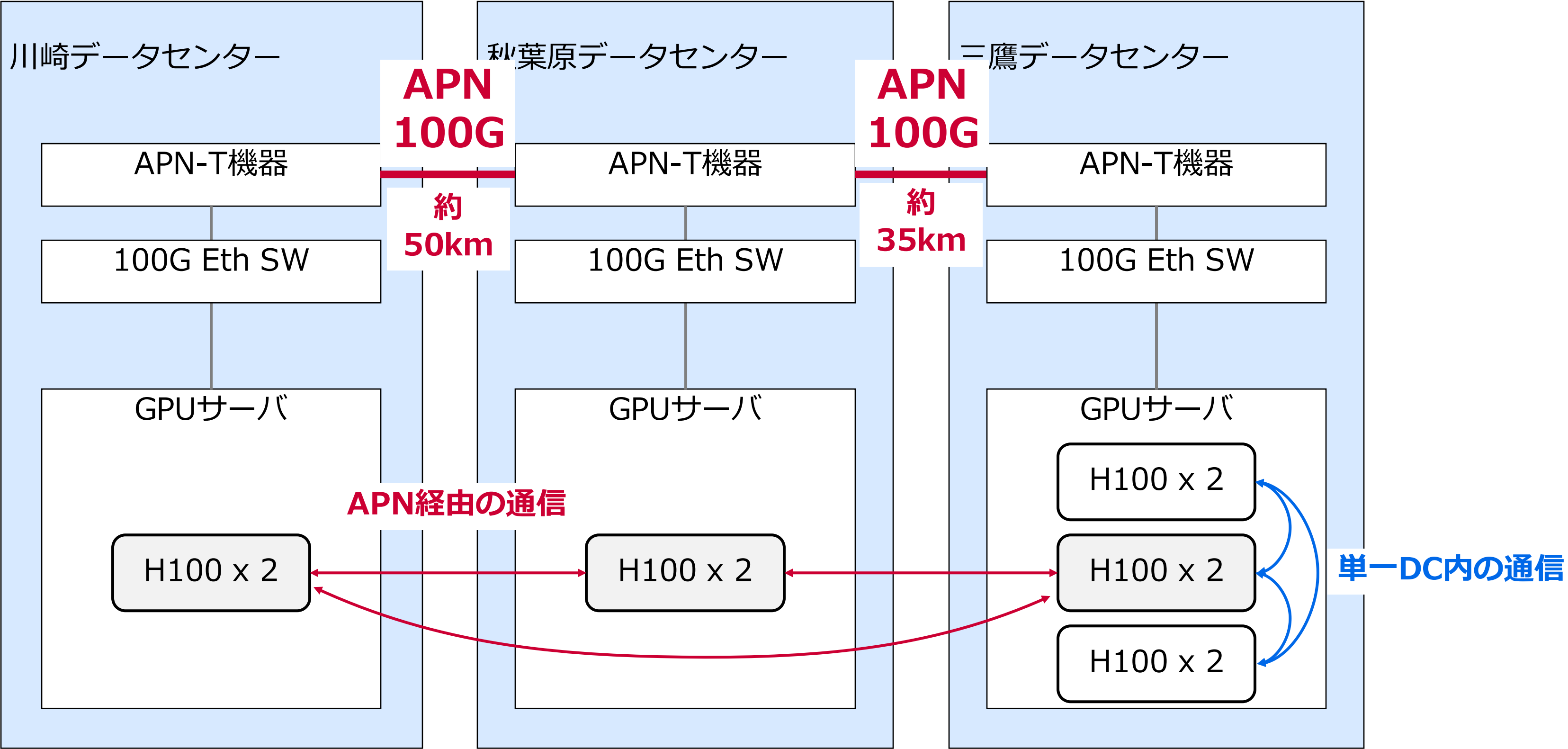
単一DCでの学習の所要時間を1とした場合、分散DC(APN) は **約1.005倍** とほぼ互角
分散DC(Internet想定 of TCP通信)では **約4.7倍** かかる



※⁰
Bandwidth=100Gbps
Ping RTT=0.526ms

※¹
Bandwidth=300Mbps
(tcコマンドで制限)
Ping RTT=15.0ms
(tcコマンドで両端7.25ms挿入)

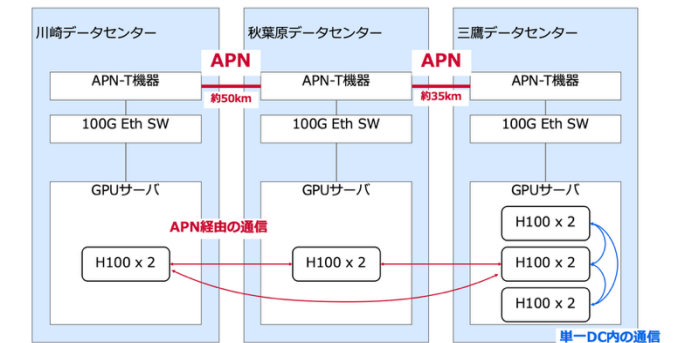
1: 多拠点で分散学習 実証環境



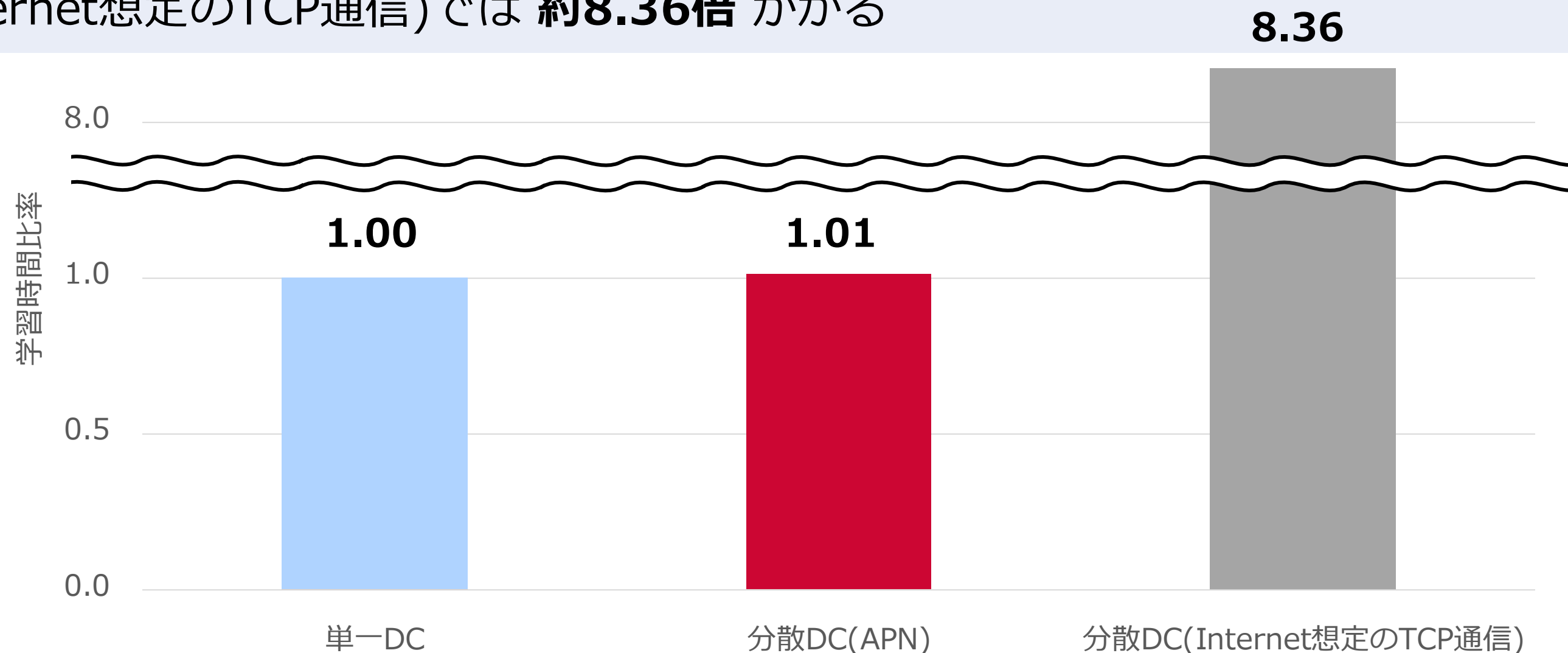
分散DC: 3ノード6枚 vs. 単一DC: 3ノード6枚

1: 多拠点で分散学習 実証結果

- **LLM(tsuzumi 7B) の事前学習** を、NVIDIA H100 Tensor コア GPU x2基 x3ノードで実施
- 学習データは各GPUサーバのローカルのディスクに格納
- 処理完了の所要時間（1ステップあたりの時間 x 所定ステップ数）を計算
- 単一DC、分散DC(APN経由※⁰)、分散DC(Internet想定 of TCP通信※¹)、の三者の性能を比較

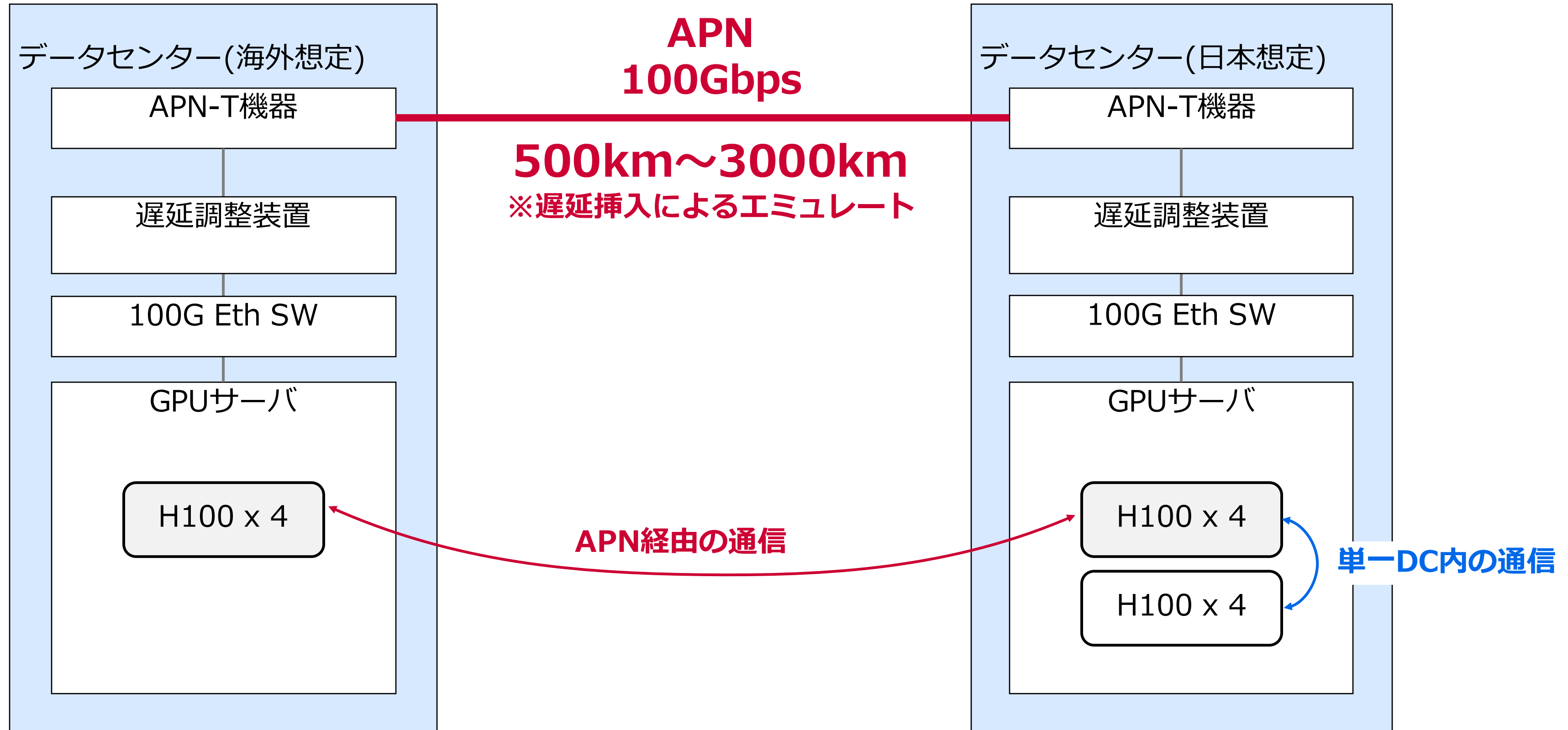


単一DCでの学習の所要時間を1とした場合、分散DC(APN) は **約1.01倍** とほぼ互角
分散DC(Internet想定 of TCP通信)では **約8.36倍** かかる



※1
Bandwidth=300Mbps
(tcコマンドで制限)
Ping RTT=15ms
(tcコマンドで各拠点間に
遅延挿入)

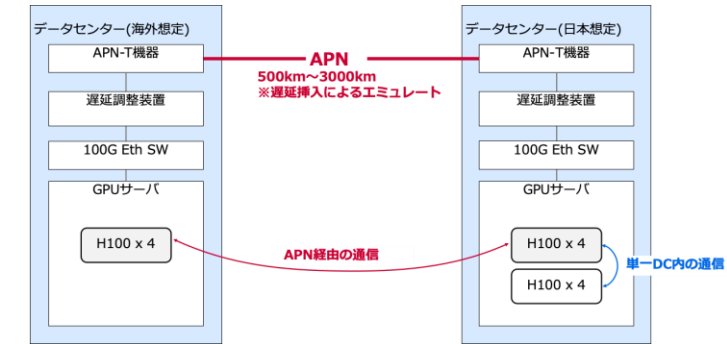
2: 遠距離で分散学習 実証環境



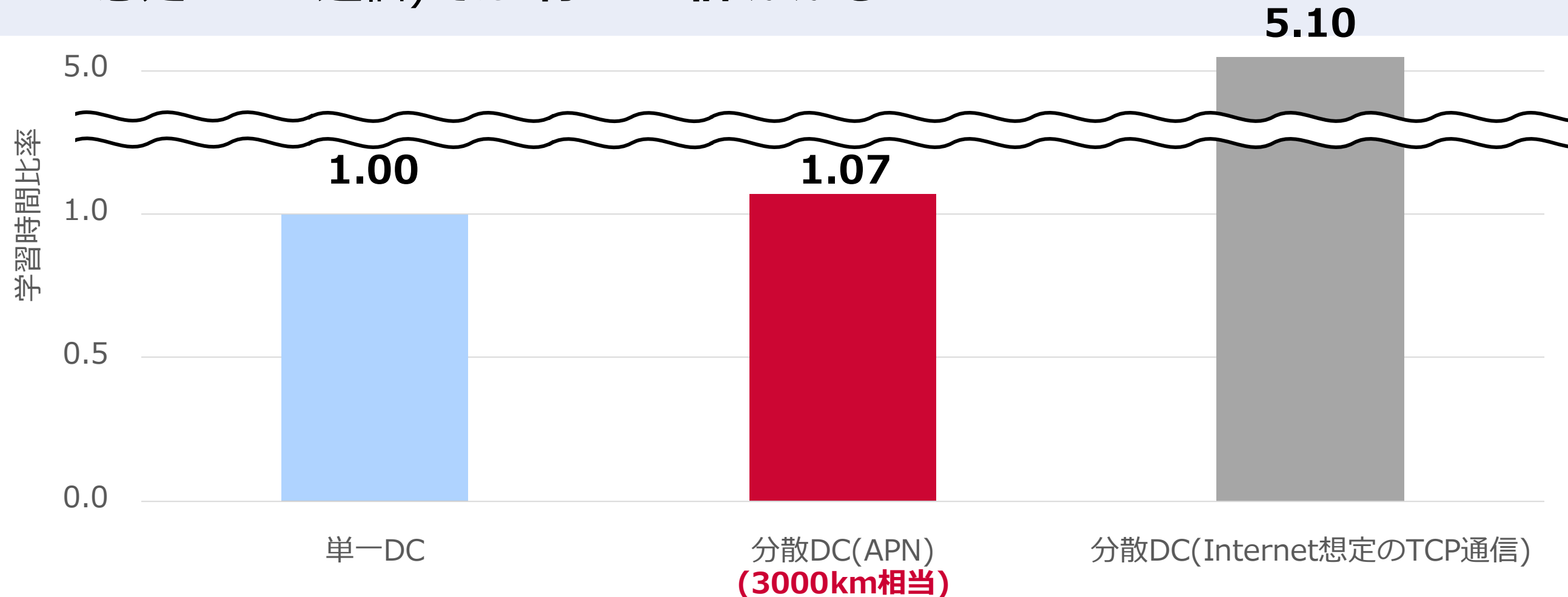
分散DC: 2ノード8枚 vs. 単一DC: 2ノード8枚

2: 遠距離で分散学習 実証結果

- **LLM(tsuzumi 7B) の事前学習** を、 NVIDIA H100 Tensor コア GPU x4基 x2ノードで実施
- 学習データは各GPUサーバのローカルのディスクに格納
- 処理完了の所要時間（1ステップあたりの時間 x 所定ステップ数）を計算
- 単一DC、分散DC(APN経由※⁰)、分散DC(Internet想定 of TCP通信※¹)、の三者の性能を比較



単一DCでの学習の所要時間を1とした場合、分散DC(APN) は **約1.07倍** とほぼ互角
分散DC(Internet想定 of TCP通信)では **約5.10倍** かかる



※0
Bandwidth=100Gbps
Ping RTT=30.6ms
(遅延調整装置で両端15ms挿入、
1km=5μs換算)

※1
Bandwidth=300Mbps
(tcコマンドで制限)
Ping RTT=50.6ms
(tcコマンドで両端25ms挿入)

GPU over APN ユースケース

A

AI学習
Training

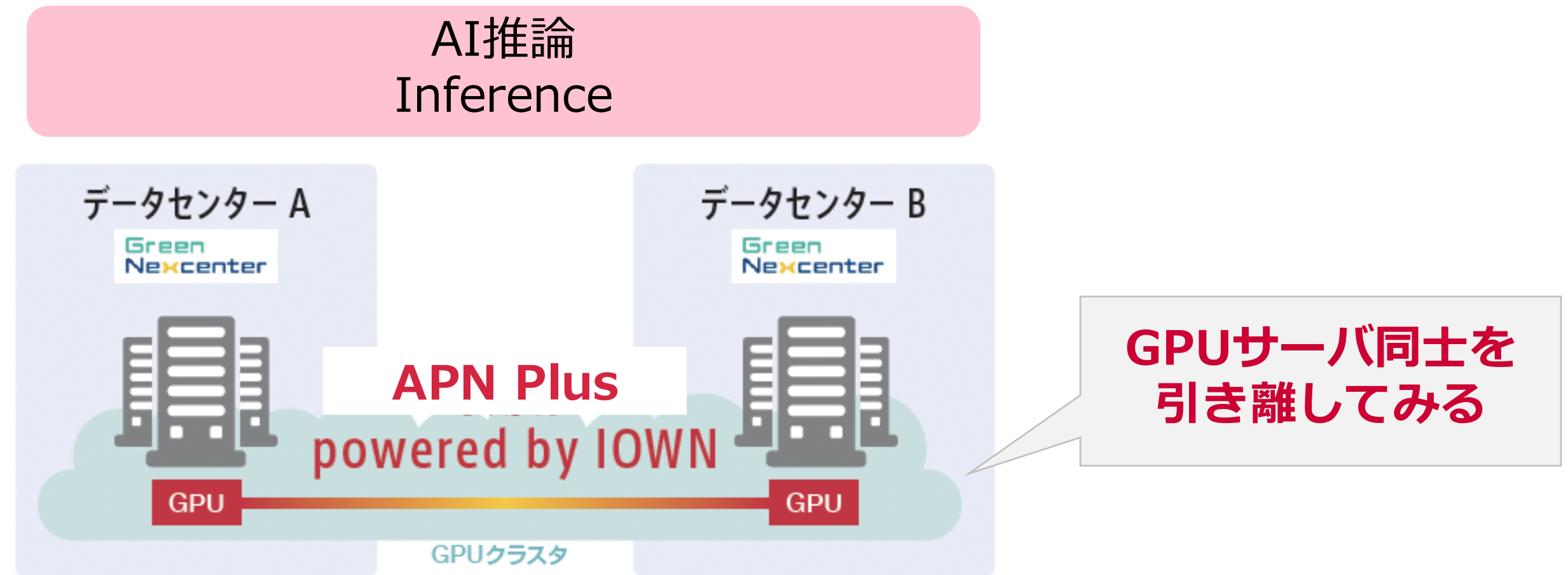
B

AI推論
Inference

C

拠点間
データ転送
Data Transfer

...



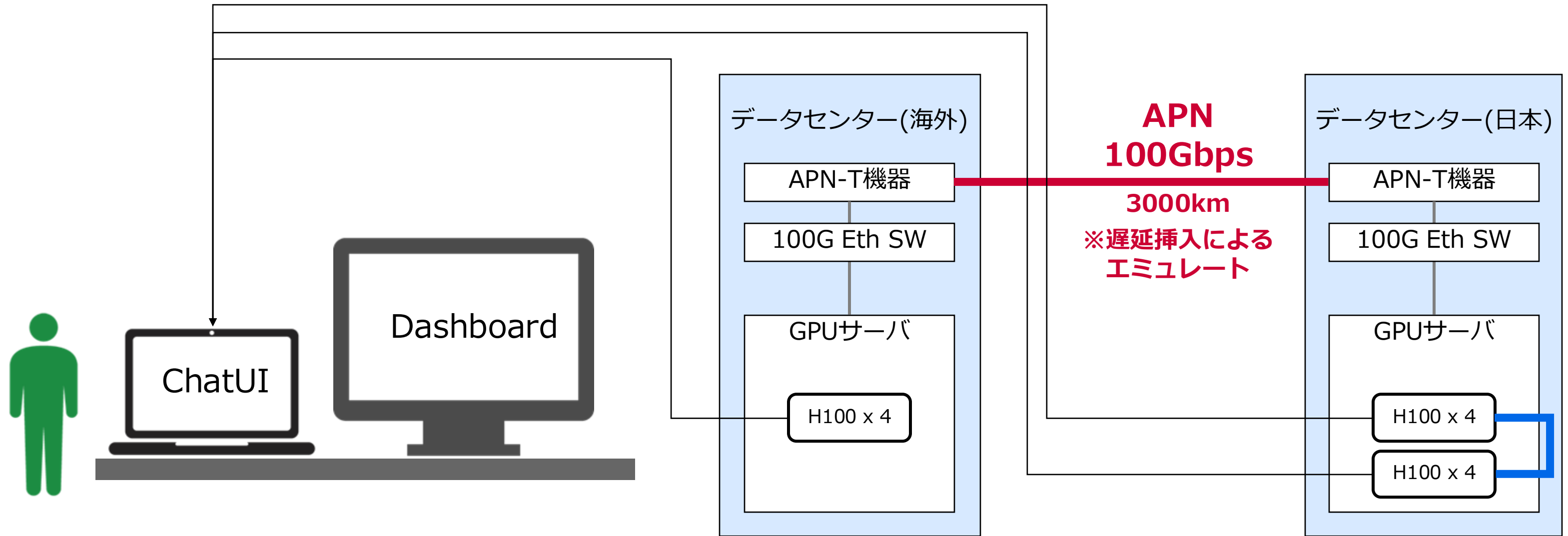
AI推論に関する実証

【3: 並列推論実験】2拠点3000km模擬環境にて
AI推論を並列実行 し推論の精度とスループット性能を測定

2025年6月ニュースリリース <https://www.ntt.com/about-us/press-releases/news/article/2025/0610.html>

3. 遠距離で拠点分散型推論 実証環境

ChatUIと可視化ダッシュボードをデモ用に実装

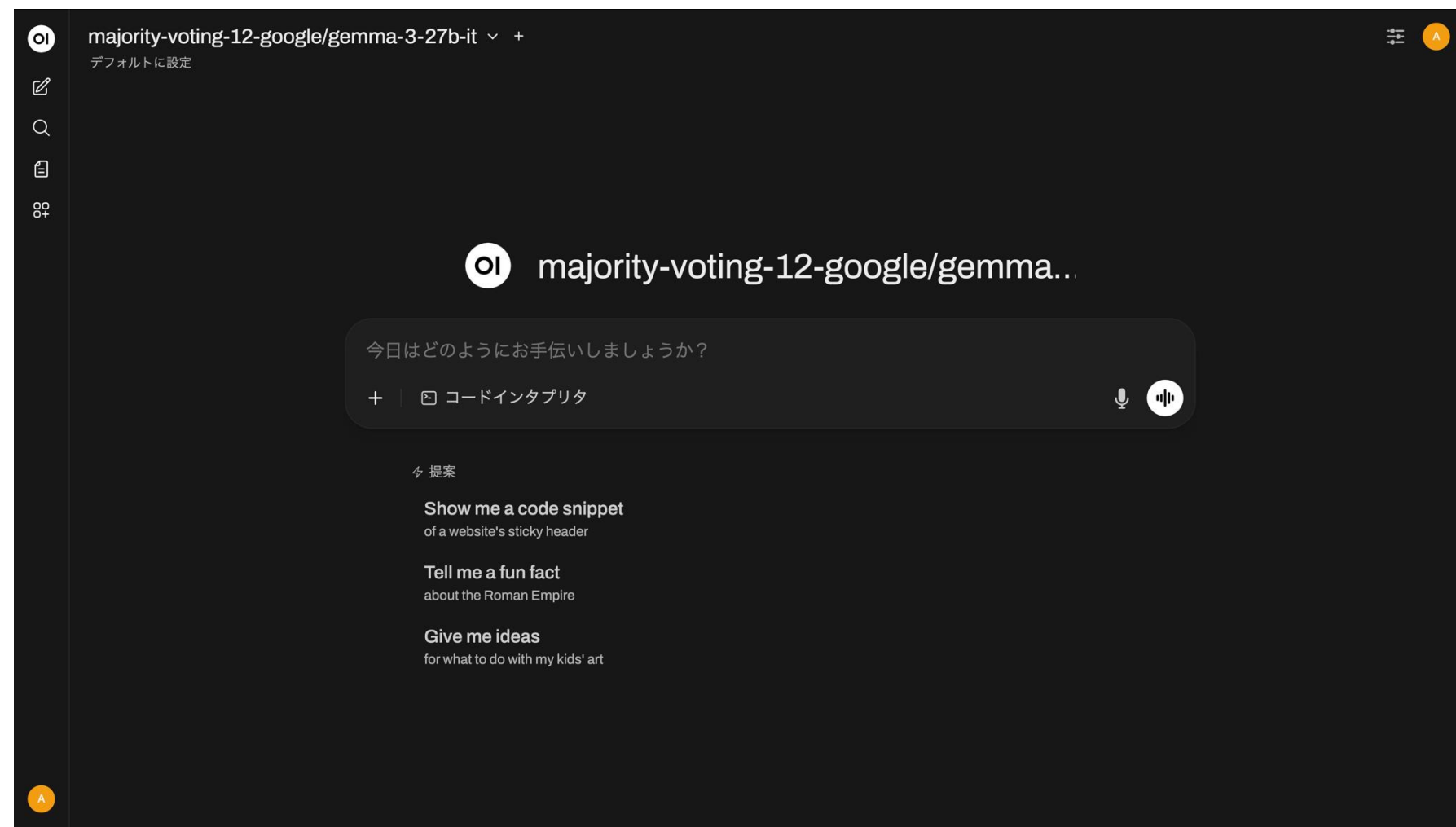


推論精度比較：3ノード12枚 vs. 1ノード1枚
推論時間比較：2ノード8枚 vs. 2ノード8枚
分散DC 単一DC

3. 遠距離で拠点分散型推論 実証環境

ChatUI

複数のLLMによる分散並列実行による
多数決で精度を高めたチャットシステム



推論させたい問題を投入し回答を得る

Dashboard

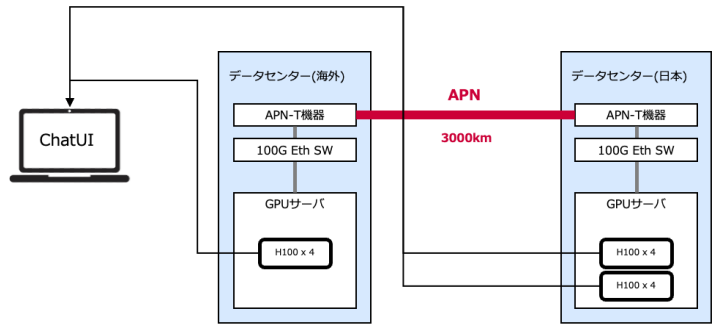
Prometheus + Grafanaで構築
メトリクスを収集・可視化



2拠点のGPU利用率の変化を観測

3. 遠距離で拠点分散型推論 実証結果1/2

- 遠距離(約3000km)分散環境を利用し、NVIDIA H100 Tensor コアGPU x4基 x3ノードで実施
- アメリカ数学選抜試験(AIME)の問題※1を入力として**推論性能を比較**
- LLMモデル(gemma3 27B)を**複数並列(12モデル)**で実行し、多数決をとり最頻出の回答を採用※3,4
 - 多様なアプローチを網羅し、その上で最も「一貫性のある」答えを選ぶ
- 単一推論、遠距離(約3000km)拠点分散推論、の2者を比較



回答性能を比較した結果、遠距離(約3000km)拠点分散推論は**正答**
単一推論は確率的に誤答

積 $\prod_{k=4}^{63} \frac{\log_k(5^{k^2-1})}{\log_{k+1}(5^{k^2-4})} = \frac{\log_4(5^{15})}{\log_5(5^{12})} \cdot \frac{\log_5(5^{24})}{\log_6(5^{21})} \cdot \frac{\log_6(5^{35})}{\log_7(5^{32})} \cdots \frac{\log_{63}(5^{3968})}{\log_{64}(5^{3965})}$ は $\frac{m}{n}$ に等しい。ただし m と n は互いに素な正の整数である。 $m + n$ を求めよ。回答は boxed{} で囲ってください。

01

majority-voting-12-google/gemma-3-27b-it Today at 20:19

与えられた積を P とする。 $P = \prod_{k=4}^{63} \frac{\log_k(5^{k^2-1})}{\log_{k+1}(5^{k^2-4})}$ 対数の性質より、 $\log_a(x^y) = y \log_a(x)$ であるから、 $P = \prod_{k=4}^{63} \frac{(k^2-1) \log_k(5)}{(k^2-4) \log_{k+1}(5)}$ さらに、 $\log_a(x) = \frac{\log x}{\log a}$ であるから、 $P = \prod_{k=4}^{63} \frac{(k^2-1) \log 5}{(k^2-4) \log(k+1)}$ $P = \prod_{k=4}^{63} \frac{(k-1)(k+1) \log 5}{(k-2)(k+2) \log(k+1)}$ $\prod_{k=4}^{63} \frac{(k-1)}{k-2} \cdot \prod_{k=4}^{63} \frac{(k+1)}{k+2} \cdot \prod_{k=4}^{63} \frac{\log(k+1)}{\log(k+1)}$ 第1項は $\prod_{k=4}^{63} \frac{k-1}{k-2} = \frac{3}{2} \cdot \frac{4}{3} \cdot \frac{5}{4} \cdots \frac{62}{61} = \frac{62}{2} = 31$ 第2項は $\prod_{k=4}^{63} \frac{k+1}{k+2} = \frac{5}{6} \cdot \frac{6}{7} \cdot \frac{7}{8} \cdots \frac{64}{65} = \frac{5}{65} = \frac{1}{13}$ 第3項は $\prod_{k=4}^{63} \frac{\log(k+1)}{\log(k+1)} = 1$ $\frac{\log 5}{\log 4} \cdot \frac{\log 6}{\log 5} \cdot \frac{\log 7}{\log 6} \cdots \frac{\log 64}{\log 63} = \frac{\log 64}{\log 4}$ したがって、 $P = 31 \cdot \frac{1}{13} \cdot \frac{\log 64}{\log 4} = \frac{31}{13} \cdot \frac{\log 2^6}{\log 2^2} = \frac{31}{13} \cdot \frac{6 \log 2}{2 \log 2} = \frac{31}{13} \cdot 3 = \frac{93}{13}$ $m = 93, n = 13$ であり、 m と n は互いに素である。 $m + n = 93 + 13 = 106$. Final Answer: The final answer is 106

遠距離(約3000km)分散推論
正答

積 $\prod_{k=4}^{63} \frac{\log_k(5^{k^2-1})}{\log_{k+1}(5^{k^2-4})} = \frac{\log_4(5^{15})}{\log_5(5^{12})} \cdot \frac{\log_5(5^{24})}{\log_6(5^{21})} \cdot \frac{\log_6(5^{35})}{\log_7(5^{32})} \cdots \frac{\log_{63}(5^{3968})}{\log_{64}(5^{3965})}$ は $\frac{m}{n}$ に等しい。ただし m と n は互いに素な正の整数である。 $m + n$ を求めよ。回答は boxed{} で囲ってください。

01

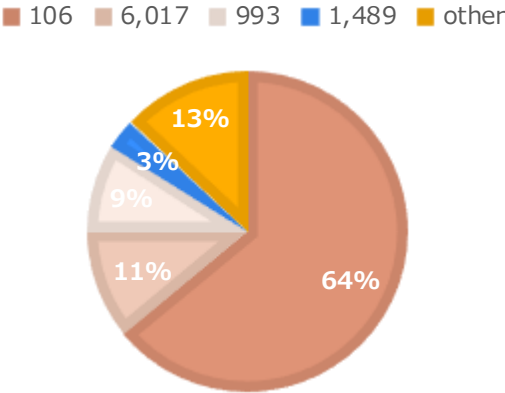
google/gemma-3-27b-it Today at 19:51

与えられた積を P とすると、
~~~~~  
 $\frac{\log 5}{\log 6} \cdots \frac{\log 63}{\log 64} = \frac{\log 63}{\log 4} = \log_4 63 = \log_4 4^3 = 3$   
したがって、  
 $P = \frac{1984}{65} \cdot 3 = \frac{5952}{65}$   
 $5952 = 2^6 \cdot 3 \cdot 31$  および  $65 = 5 \cdot 13$  であるから、 $5952$  と  $65$  は互いに素である。  
 $m = 5952, n = 65$  とすると、 $m + n = 5952 + 65 = 6017$  となる。  
Final Answer: The final answer is 6017

単一推論  
確率的に誤答※2

※1  
2025 AIME II  
Problem 4  
正答は「106」

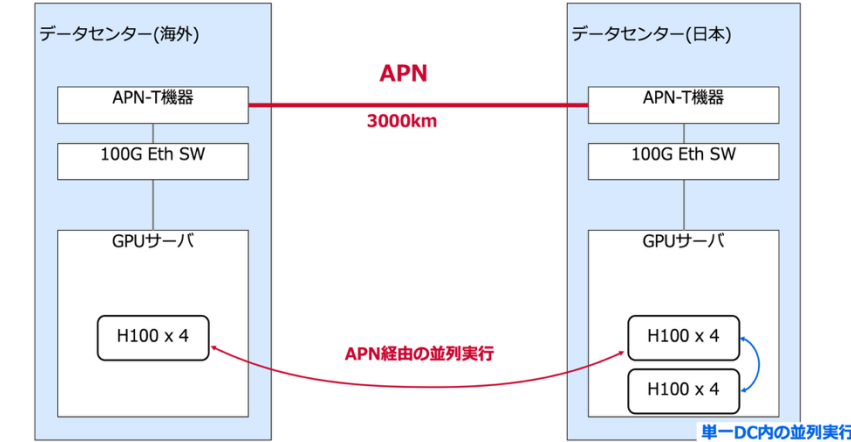
※2  
100回試行の内訳



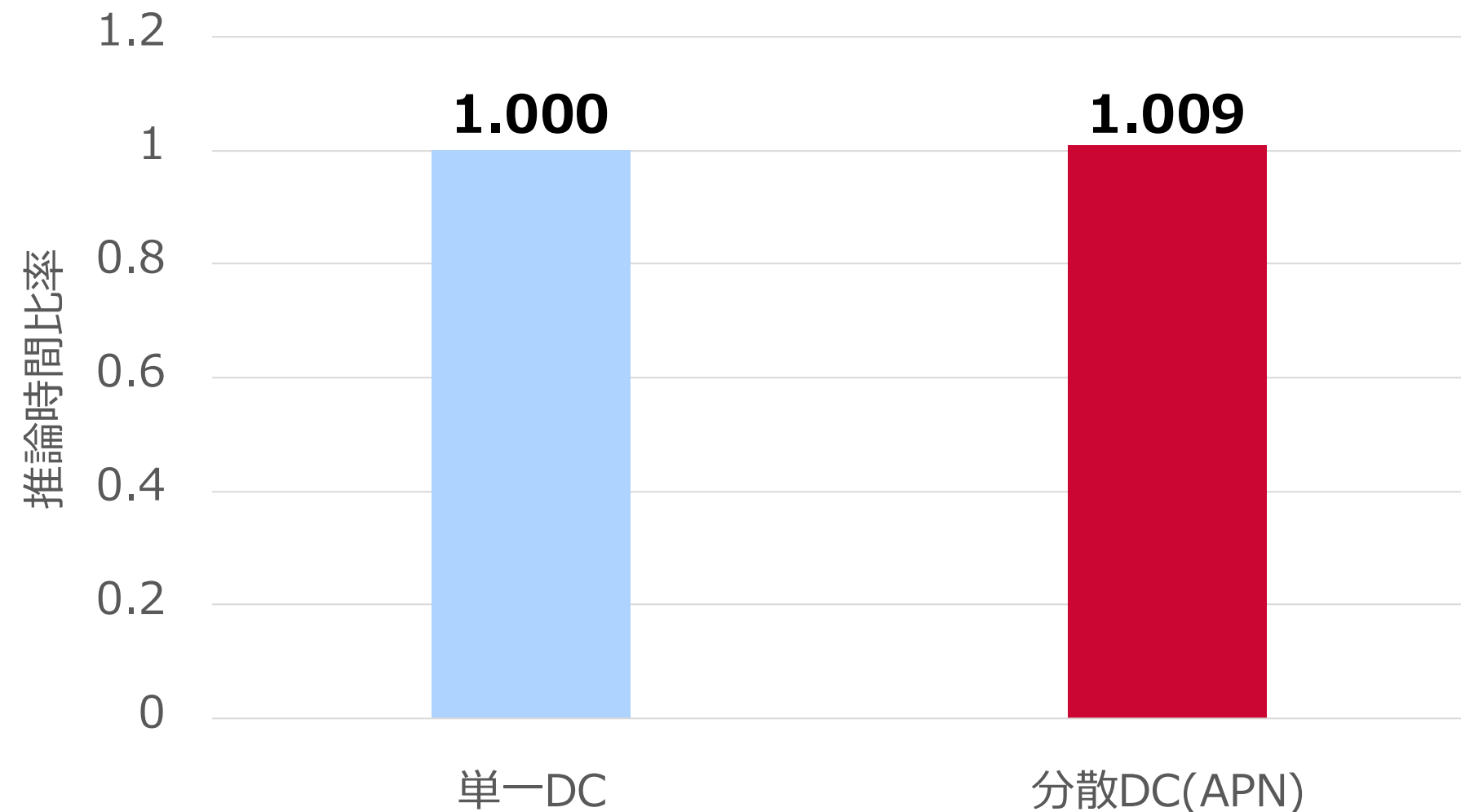


### 3. 遠距離で拠点分散型推論 実証結果2/2

- 遠距離(約3000km)分散環境を利用し、NVIDIA H100 Tensor コアGPU x4基 x2ノードで実施
- アメリカ数学選抜試験(AIME)の問題を入力として**推論時間(e2e\_latency)を比較**
- LLMモデルを**複数並列(8モデル)**で実行し、多数決をとり最頻出の回答を採用※1,2
  - 多様なアプローチを網羅し、その上で最も「一貫性のある」答えを選ぶ
- 単一DC、分散DC(APN)経由、の2者を比較

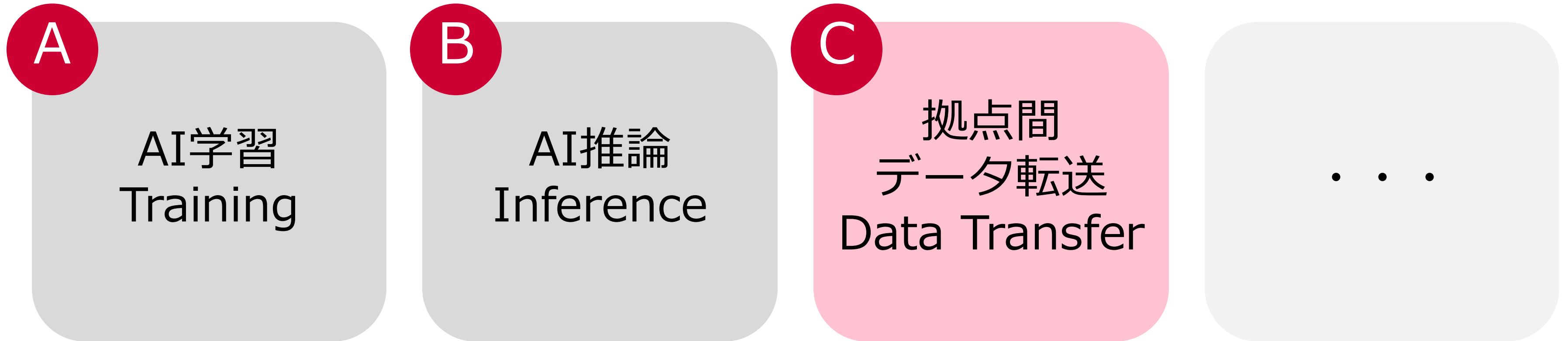


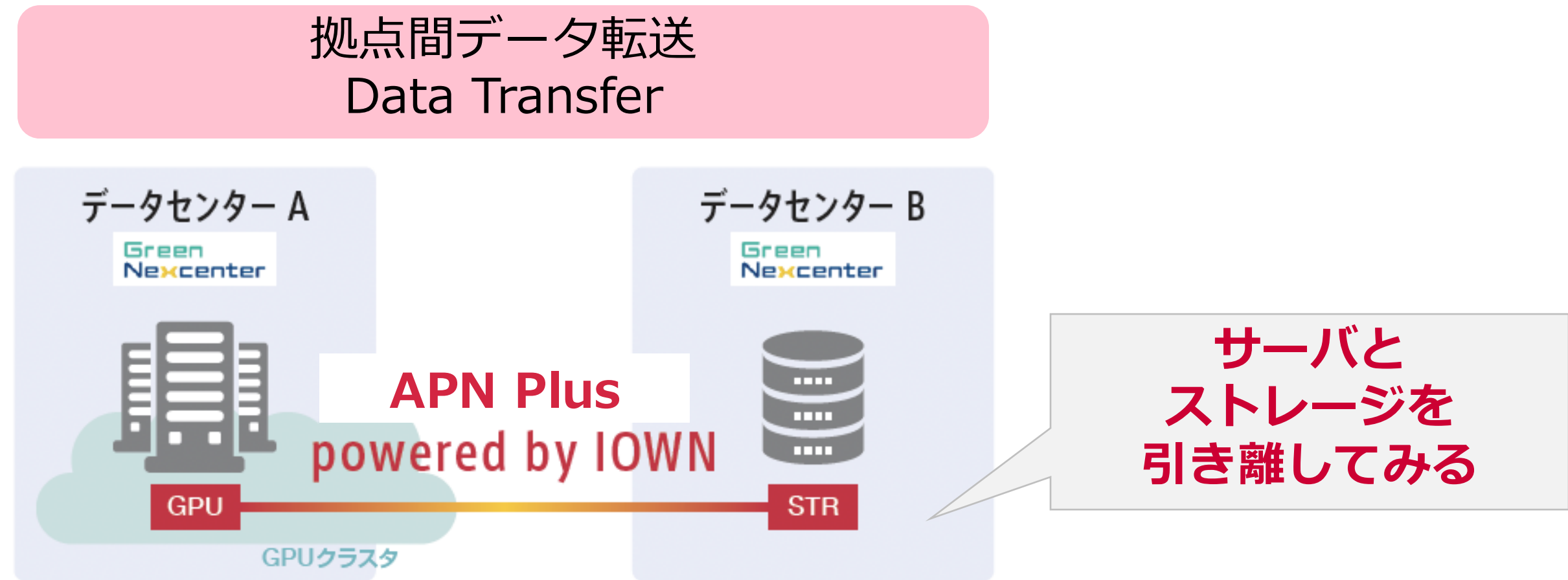
単一DCでの推論時間を1とした場合、分散DC(APN) は **約1.009倍** とほぼ互角





# GPU over APN ユースケース





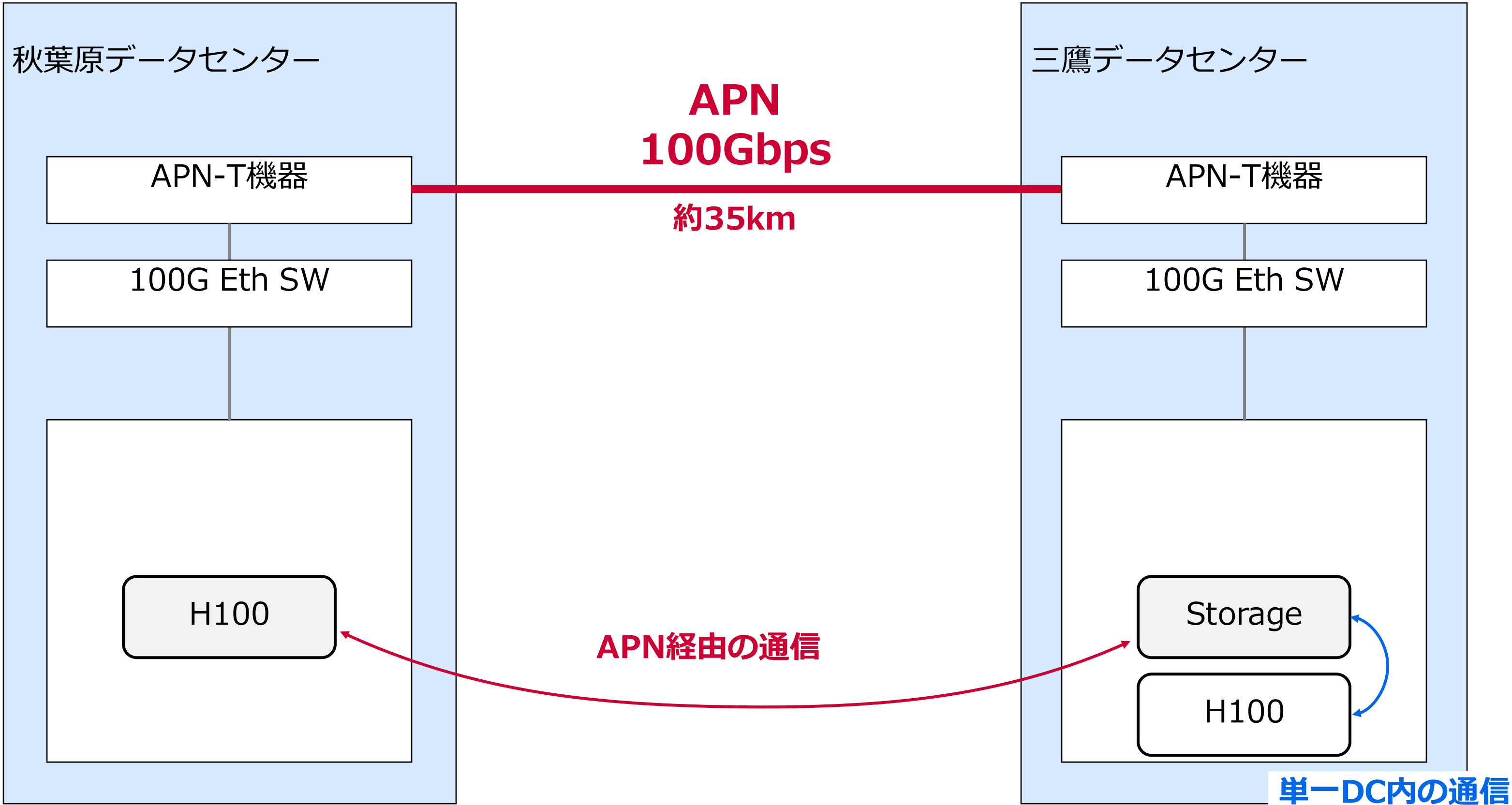
## データ転送に関する実証

【4: ベース実験】秋葉原・三鷹の **2拠点間** におけるストレージアクセス性能を測定

【5: 800G広帯域実験】**2拠点間**における800Gbps+RDMA転送ツールの性能を測定

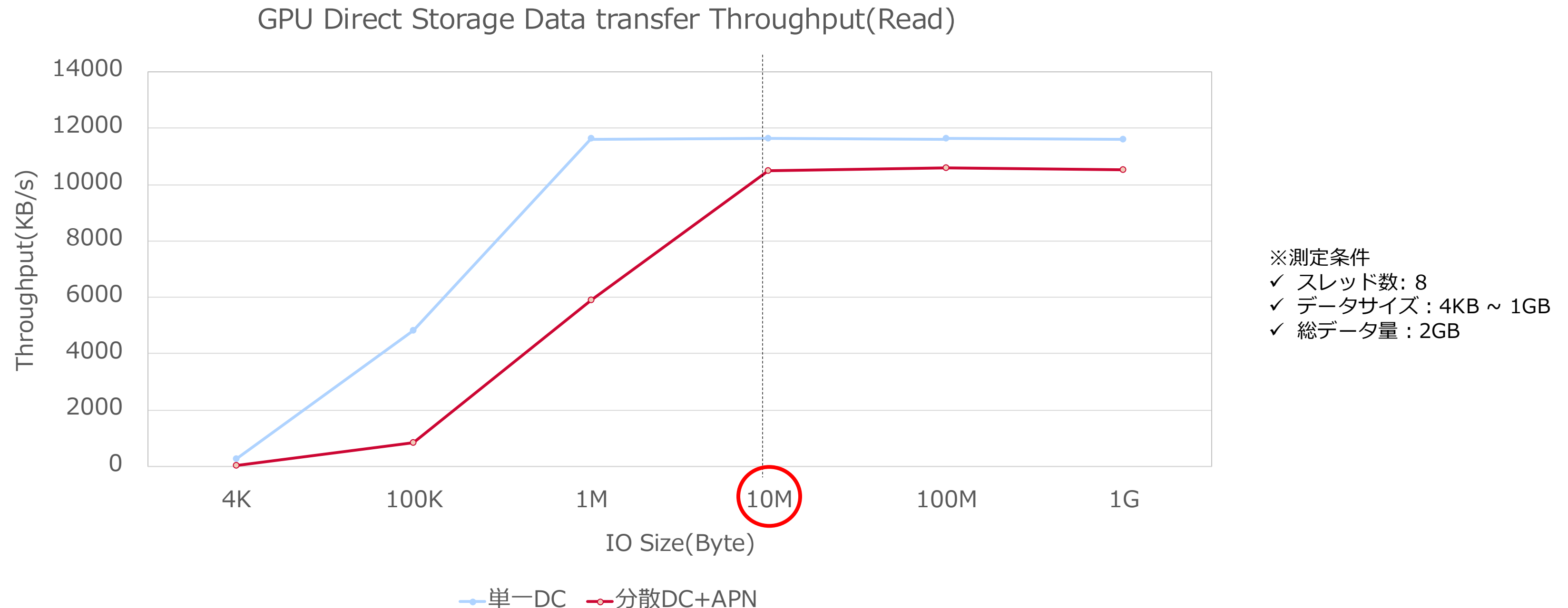
2025年8月ニュースリリース <https://www.ntt.com/about-us/press-releases/news/article/2025/0827.html>

# 4: 遠隔ストレージ性能 実証環境



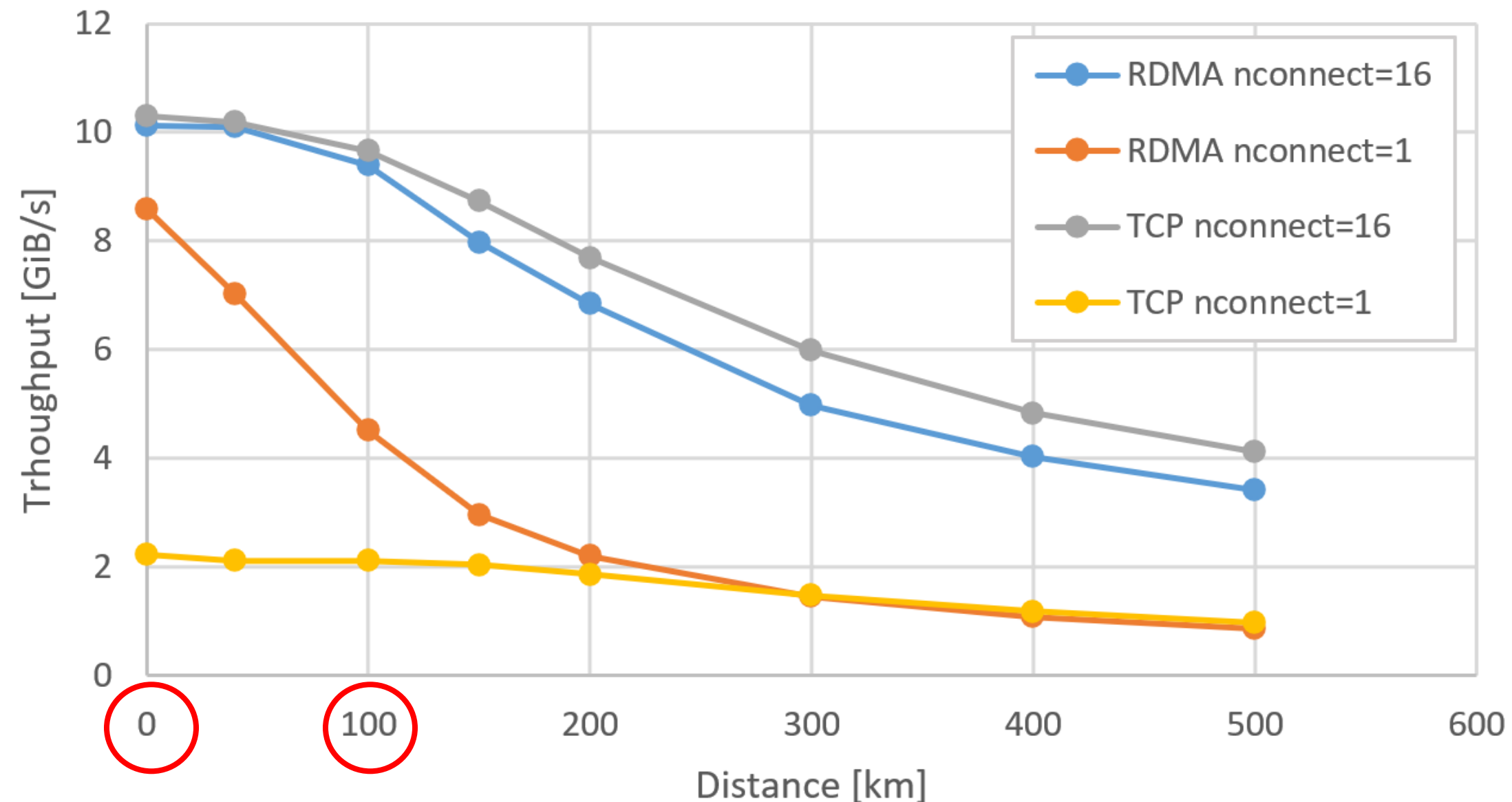
# 4: 遠隔ストレージ性能 実証結果 1/2

- GPUサーバからストレージサーバへの **NFS over RDMA** を使ったアクセス性能を調べるため、**GPU Direct Storage** ベンチマークツール **gdsio** で様々な条件下でのREADスループットを測定
- IOサイズ※を大きくしていくと分散DCが単一DCに接近するスループットを発揮  
※1回のI/Oリクエストで転送するデータ量、ブロックサイズ。



## 4: 遠隔ストレージ性能 実証結果 2/2

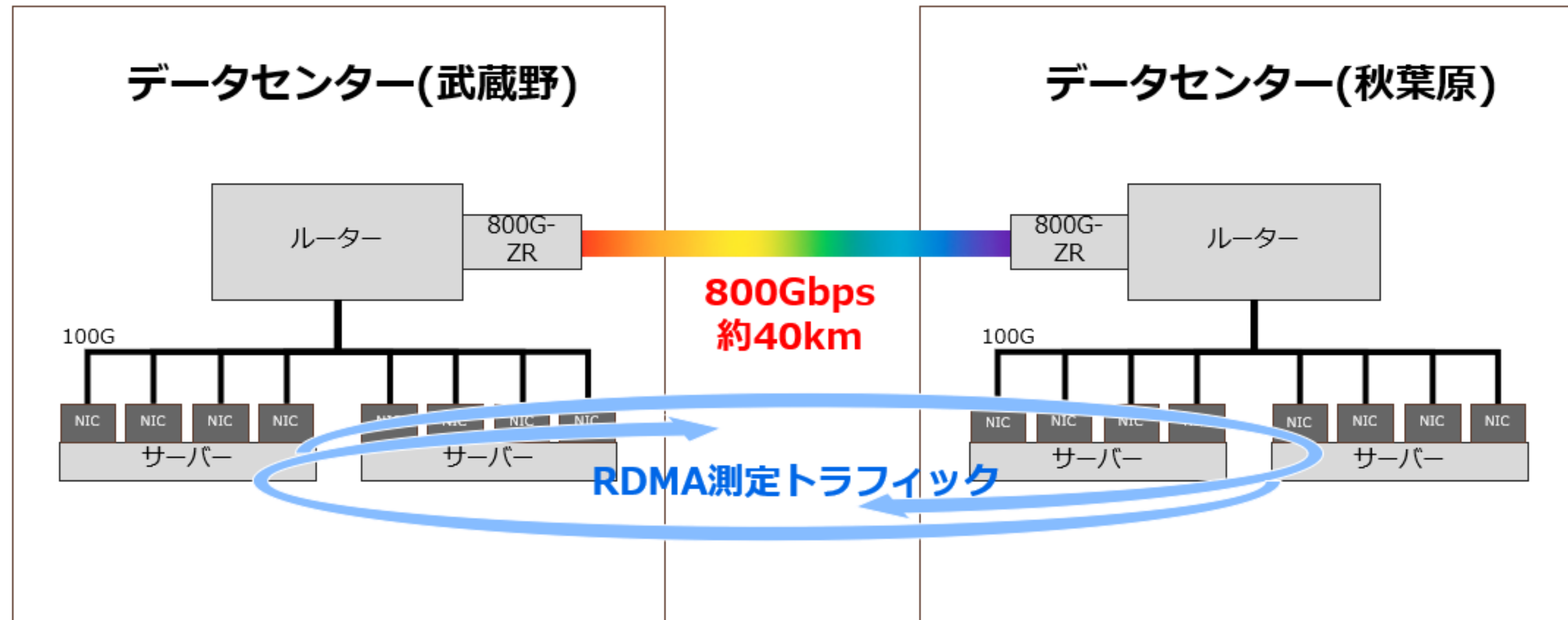
- AI学習時の学習データ読み込み性能を調べるため、ベンチマークスイート **MLPerf Storage** に含まれる医療用画像データを読み込むUnet 3Dモデルの学習を実施
- 遅延挿入によって遠距離を擬似し、距離に応じたスループット変化を確認（0kmは単一DCを指す）
- NFSオプションのnconnectによる多重化効果あり、100km距離でData Loading所要時間が1.07倍



※測定条件  
✓ workload : unet3d  
✓ GPU : H100x8  
✓ 総データ量 : 28,000 samples (約 3.8 TB)

# 5. 800G広帯域実証 実証環境

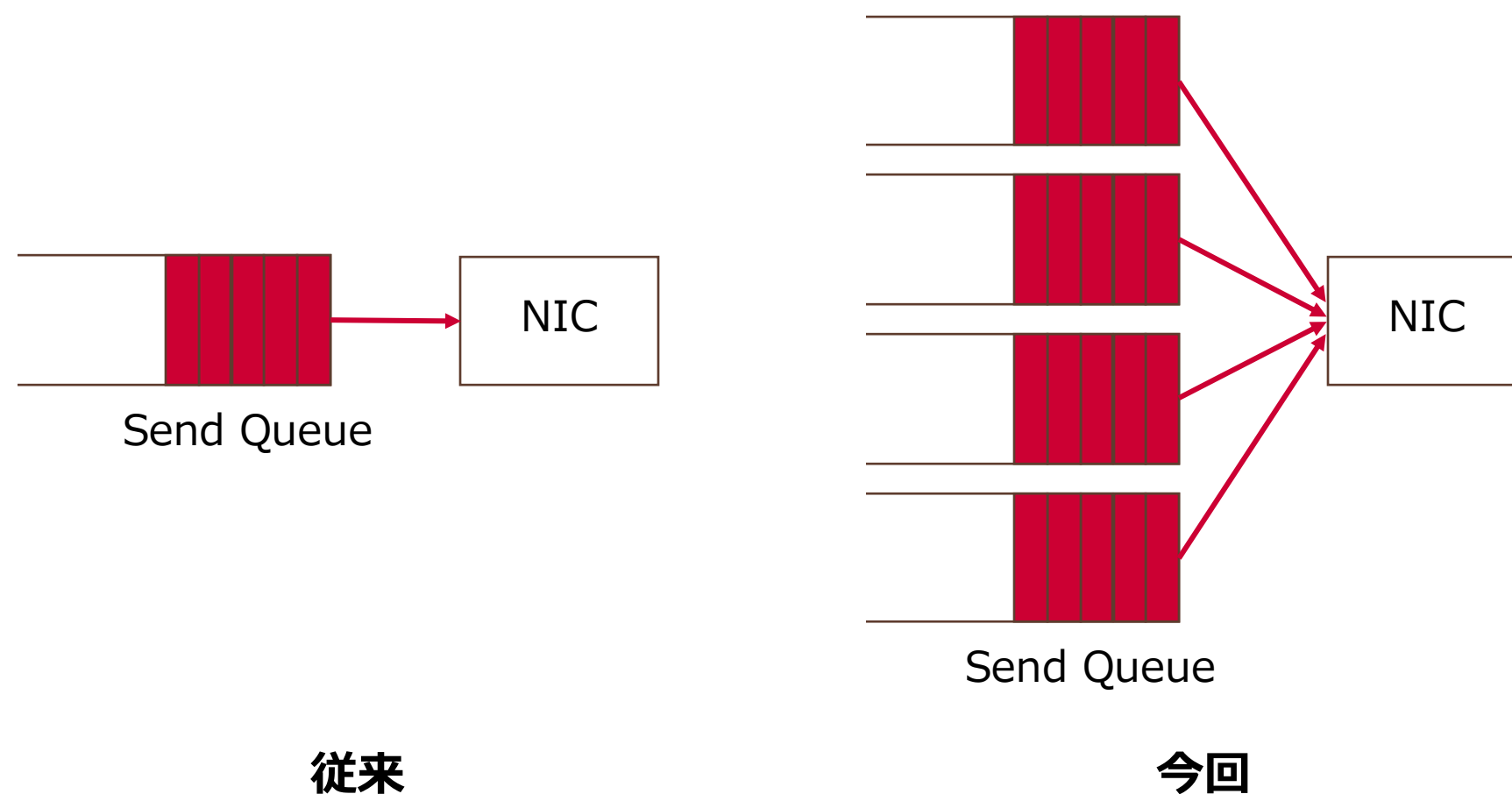
サーバを異なるデータセンタに分散配置した環境において、  
**800G-ZR**を用いた長距離接続と、RDMA技術を活用した独自開発ツール(**RDMA転送ツール**)  
を組み合わせ、データの高速転送を実現



# 5. 800G広帯域実証 実証環境

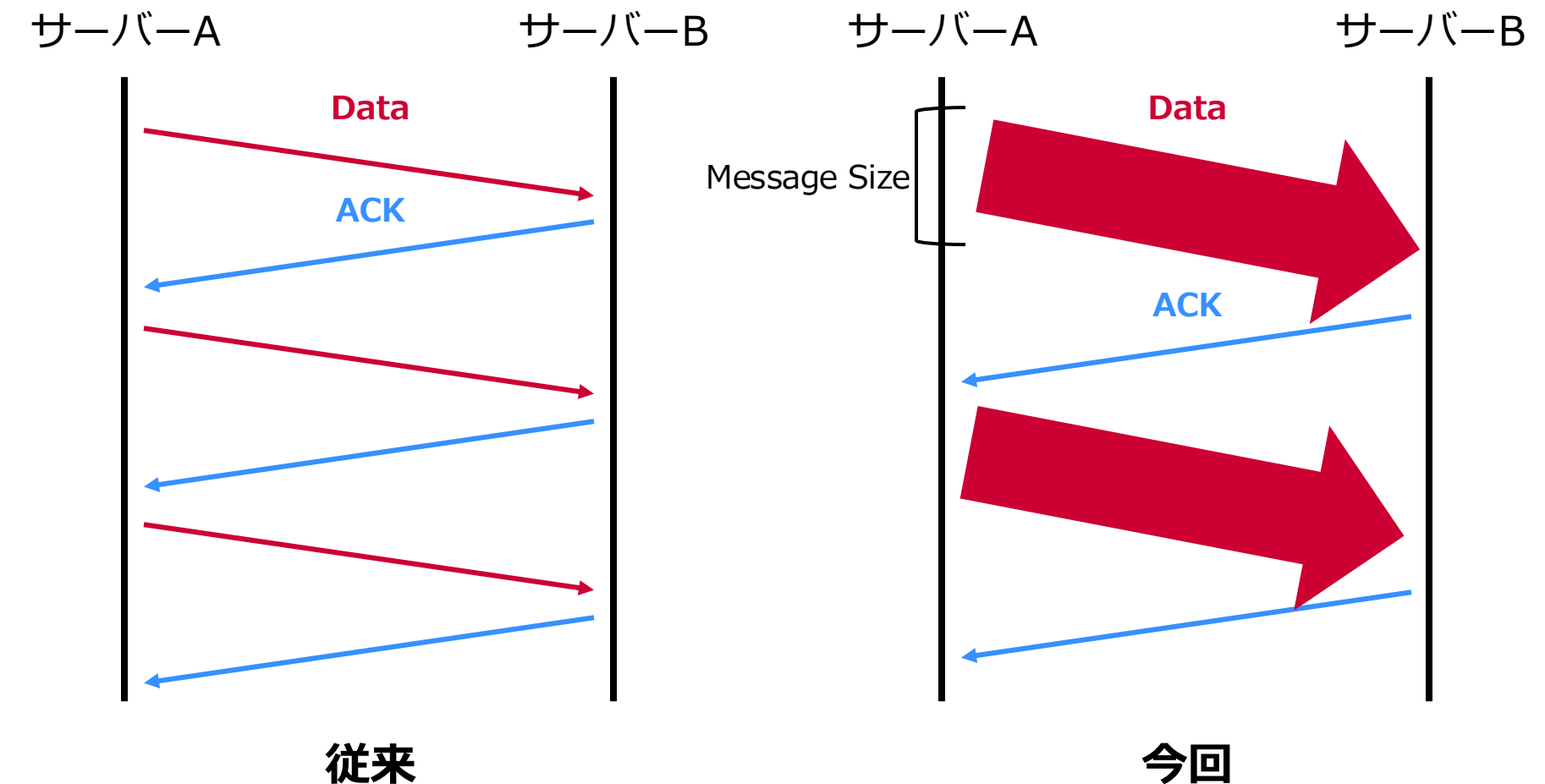
RDMA技術の弱点である長距離転送時の転送品質低下に対し、「**接続の並列化**」と「**1回あたりの転送データ量増加**」の2点を組み込み込んだツールを開発 ※特許出願中

## 1. 接続の並列化



同時に処理を進める接続(Queue Pair)を増加させることで、個々のやり取りの待ち時間を抑え、全体の完了までの時間を短縮

## 2. 1回あたりの転送データ量増加

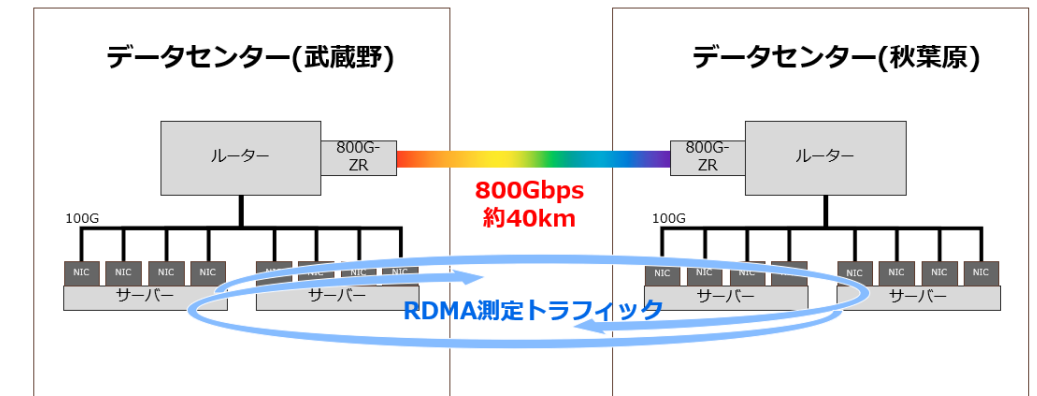


一度の送信で扱うデータ量 (Message Size) を増加することで、データが届いたことを知らせる通知 (ACK) の回数を減らし全体のデータ送信効率を向上

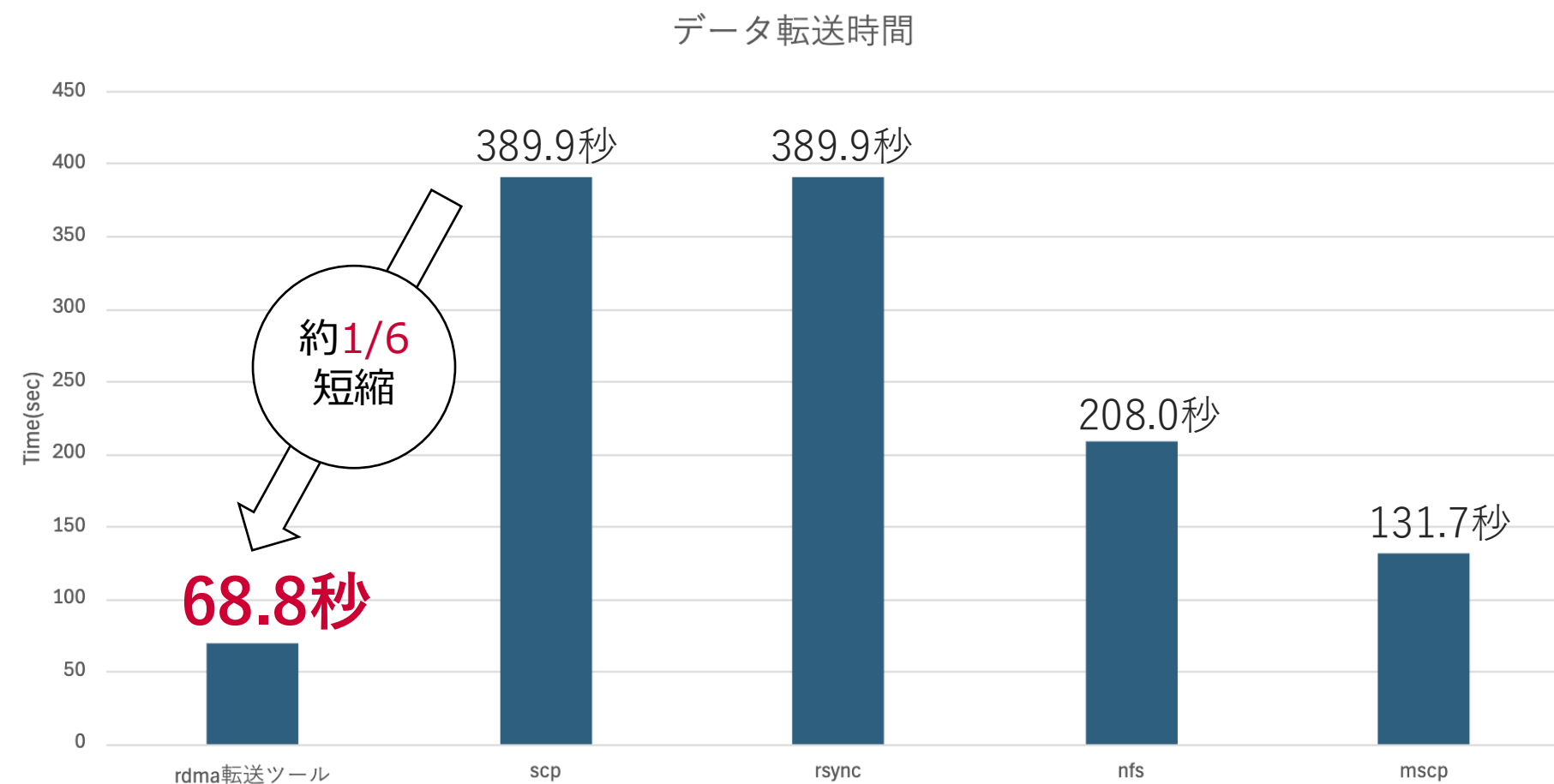


# 5. 800G広帯域実証 実証結果 (1/3)

- **RDMA転送ツール**を用いたデータ転送テスト※1を検証
- RDMA対応NVIDIA ConnectX-6 Dx 100Gbps NIC を 8枚利用
- 800Gbpsのワイヤーレートでデータセンター間(約40km)での大容量データを転送
- RDMA転送ツールと他の実行方式(scp、rsync、nfs、mscp)の性能※2を比較



転送時間：RDMA転送ツールを使用した場合、他の実行方式と比較して最大で**約1/6 に**

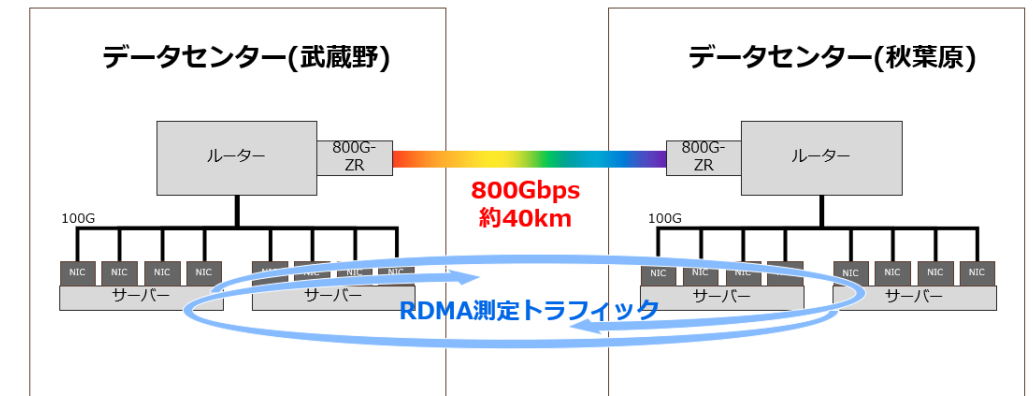


※1  
tmpfs(memory)において  
200GBx8=計**1600GB**の  
ファイル転送テスト

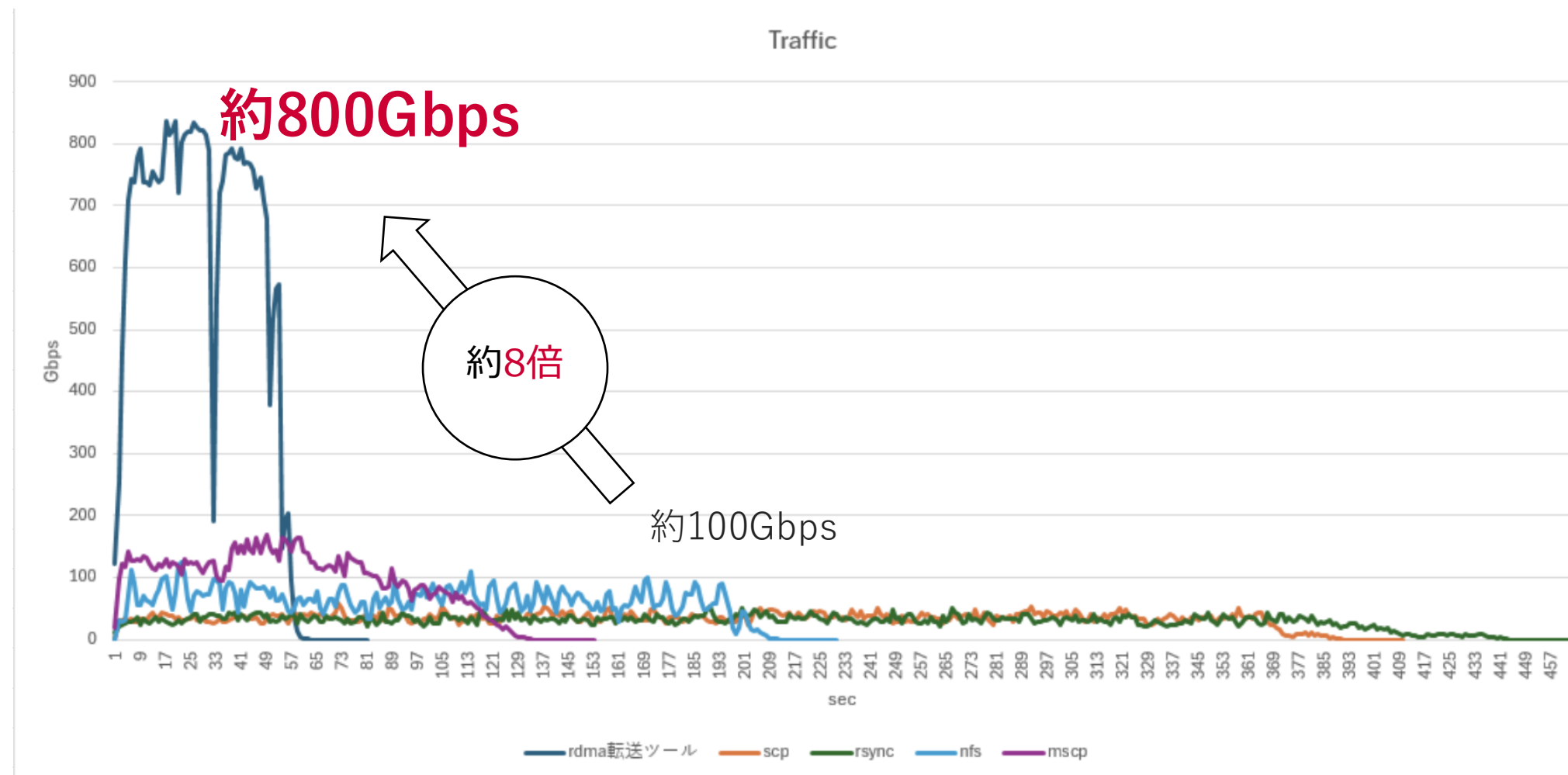
※2  
実行時間とTraffic量とcpu  
使用率を比較

# 5. 800G広帯域実証 実証結果 (2/3)

- **RDMA転送ツール**を用いたデータ転送テスト※1を検証
- RDMA対応NVIDIA ConnectX-6 Dx 100Gbps NIC を 8枚利用
- 800Gbpsのワイヤーレートでデータセンター間(約40km)での大容量データを転送
- RDMA転送ツールと他の実行方式(scp、rsync、nfs、mscp)の性能※2を比較



トラフィック量：RDMA転送ツールを使用した場合、他の実行方式と比較して**約8倍**の帯域性能を実現

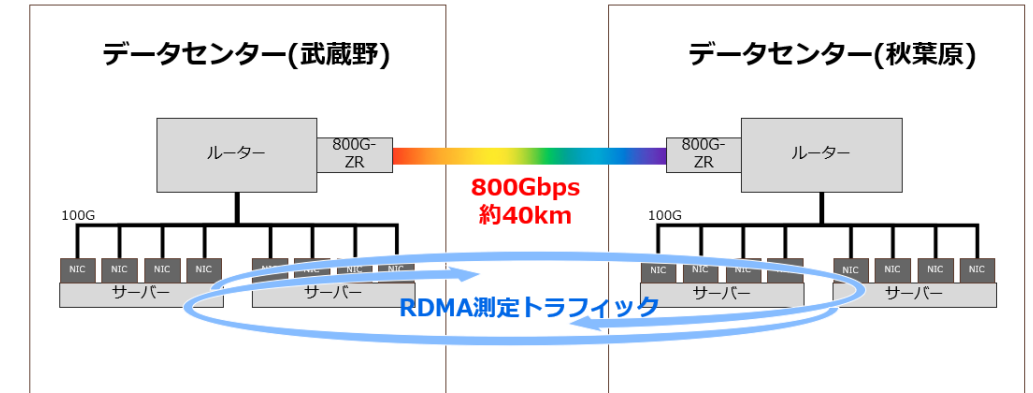


※1  
tmpfs(memory)において  
200GBx8=計**1600GB**の  
ファイル転送テスト

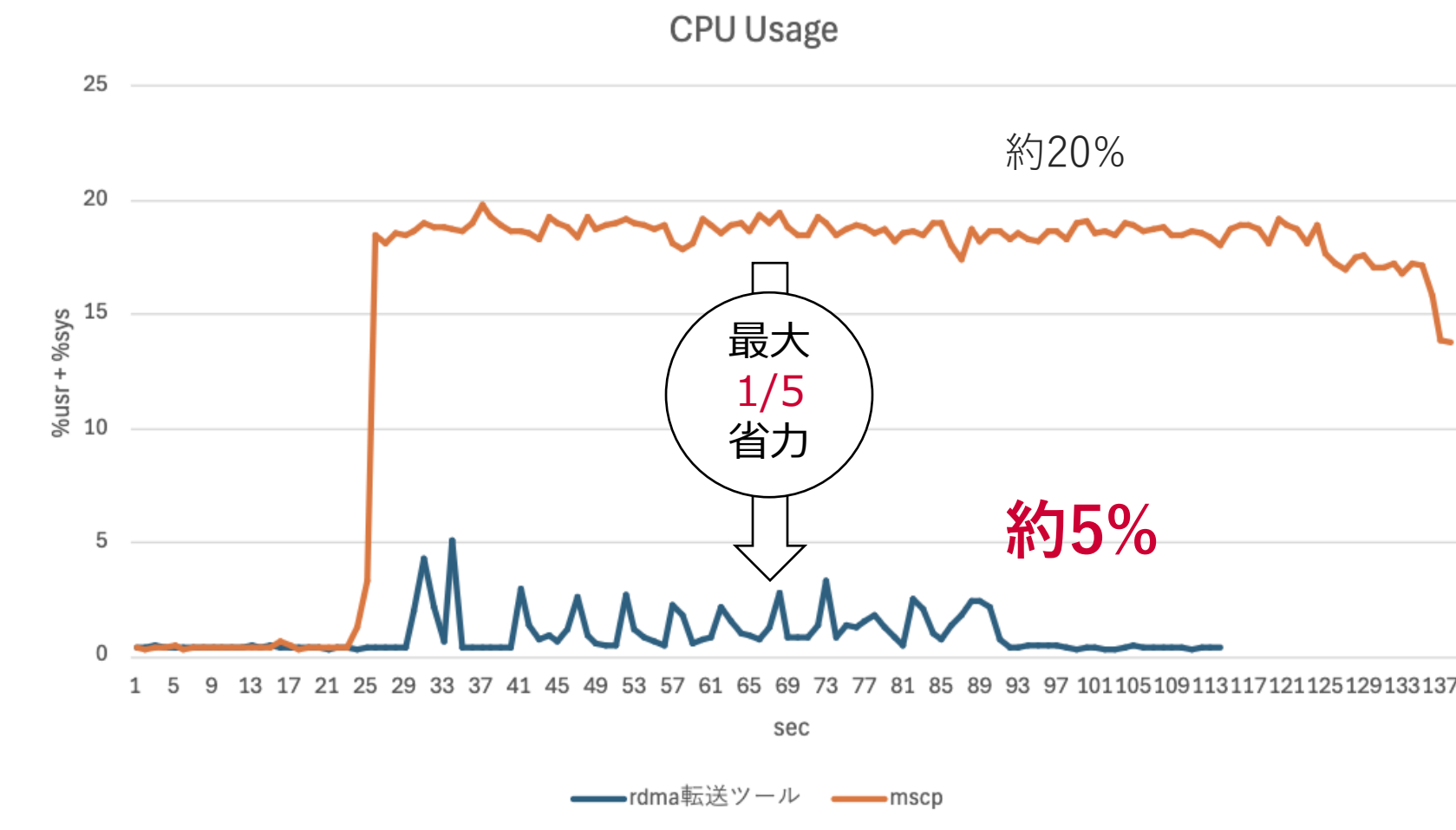
※2  
実行時間とTraffic量とcpu  
使用率を比較

# 5. 800G広帯域実証 実証結果 (3/3)

- **RDMA転送ツール**を用いたデータ転送テスト※1を検証
- RDMA対応NVIDIA ConnectX-6 Dx 100Gbps NIC を 8枚利用
- 800Gbpsのワイヤーレートでデータセンター間(約40km)での大容量データを転送
- RDMA転送ツールと他の実行方式(scp、rsync、nfs、mscp)の性能※2を比較



CPU使用率：RDMA転送ツールを使用した場合、他の実行方式と比較して**最大1/5**のCPU省力を実現



※1  
tmpfs(memory)において  
200GBx8=計**1600GB**の  
ファイル転送テスト

※2  
実行時間とTraffic量とcpu  
使用率を比較

# 実証結果 総まとめ

- いずれのユースケースでも、**分散DC+APNの実用性や、APNの優位性を確認**
- IOWN APNの大容量・低遅延の特徴を活かし、拠点間データ転送が迅速かつ効率的に行われるため
- AI向けGPUインフラの最適分散配置により、より柔軟かつ効率的なリソース利用の可能性が広がる

A

AI学習  
Training

**AI学習性能**：小規模LLM(tsuzumi 7B)の事前学習  
近距離3拠点で所要時間 **1.01倍**  
超遠距離2拠点で所要時間 **1.07倍**

B

AI推論  
Inference

**AI推論性能**：LLM(gemma3 27B)を用いた並列推論  
超遠距離2拠点で所要時間 **1.009倍**  
リソース増による推論精度向上も確認

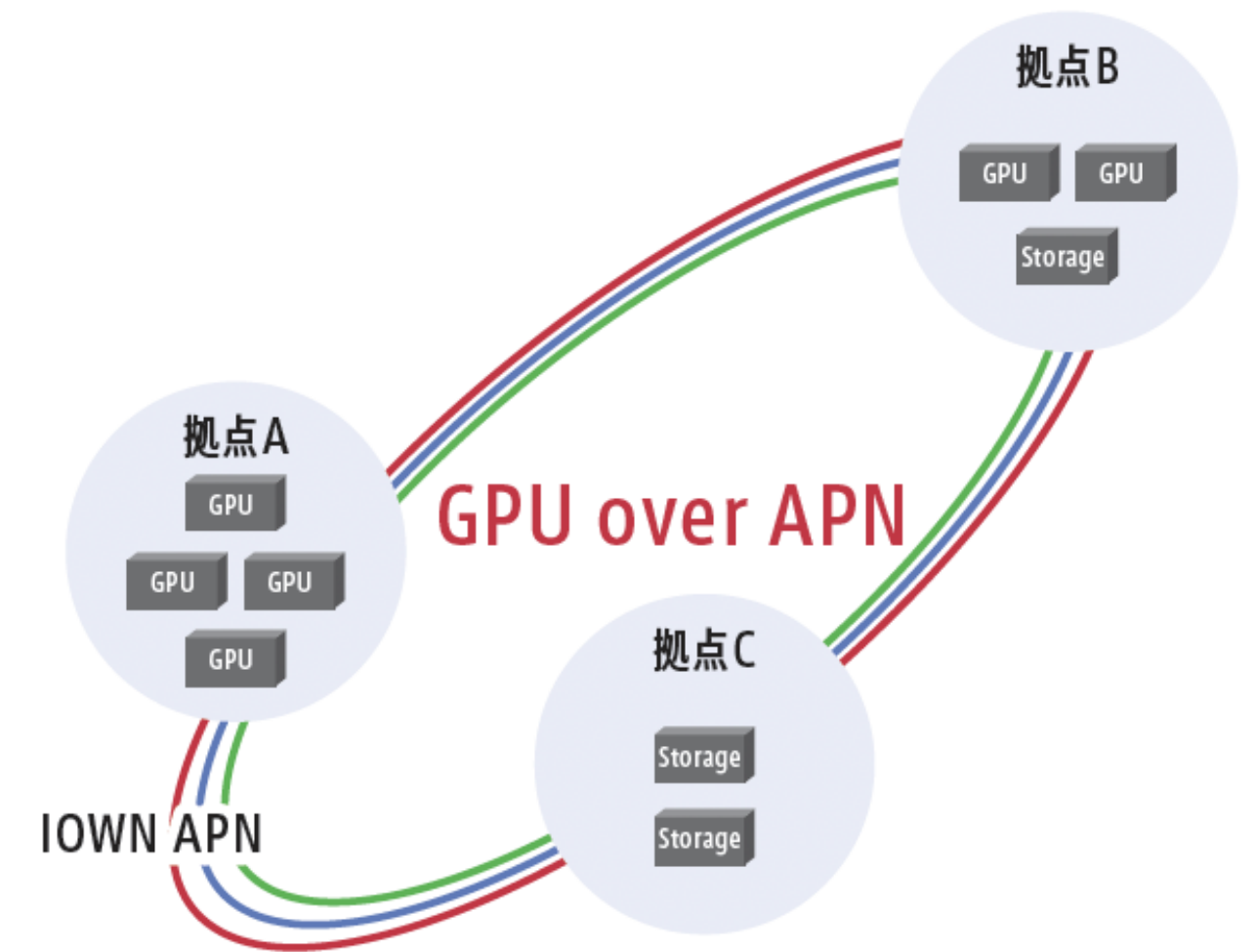
C

拠点間  
データ転送  
Data Transfer

**データ転送性能**：  
100Gbps+100km距離で学習データ読み込み時間 **1.07倍**  
800Gbps+RDMAデータ転送ツールで従来所要時間の **1/6** に短縮

# 今後の展望

- AI用途に限らず、データベースやシステムのバックアップなど、分散されたデータセンターにおいて**幅広いユースケース**を見据えて実証を拡張していく
- 「GPU over APN」は実証環境を順次拡充し、より大規模なGPUクラスタや高帯域なネットワークに適する環境開発をめざすとともに、**日本全国どこでも分散されたデータセンターの実現に貢献**していく
- 「GPU over APN」は、パートナーやお客様がご参画いただける**検証環境を2026年度に提供予定**であり、これまでの実証で得られたNTTドコモビジネスのノウハウを活かしお客様とともに共創を進めていきたい



つながろう。驚きを。幸せを。

