

mdx（データ活用社会創成プラットフォーム）の現状 と次世代学術情報基盤への一考察

田浦健次郎

東京大学 情報基盤センター長

大学院情報理工学系研究科 教授

JHPCN共同利用共同研究拠点 拠点長

tau@eidos.ic.i.u-tokyo.ac.jp



概要

- 背景：データ活用のための基盤の必要性
- mdx の紹介・現状説明
- 情報基盤センター群・JHPCNにとってのmdx
- 次世代情報基盤の一角としての仮想化基盤

背景

- Society 5.0
- 科学技術基本計画
 - 新たな研究システムの構築（オープンサイエンスとデータ駆動型研究等の推進）
- 大学研究力強化委員会
 - コアファシリティやオープンラボなどの共用システム、共同利用・共同研究拠点の活用、etc.
- 研究データ基盤（Research Data Cloud）
- 次世代計算基盤

学術情報基盤整備をとりまく状況

1. 分野ごとに情報基盤の大規模化・高度化対応するのが困難になっている

- スパコン利用が進んでいたいくつかの研究分野（e.g., 気象、地震、物性材料、ゲノム医科学、etc.）で従来独自のスパコンを調達してきたが必要な計算規模拡大、最新技術の高度化・多様化にそれぞれの分野で外部利用まで含めた対応をするのが難しくなっている

2. 裾野の広い分野に求められる情報基盤が提供できているとはいえない

- データ・機械学習（AI）活用による革新はこれまでの**学術情報基盤（スパコン中心）**利用分野以上に**広い分野**でおきる可能性がある
 - 従来情報技術活用から縁遠かった分野（歴史学、社会科学（経済）、教育学など）でもデータの蓄積、解析などが重要になっている
- そのための情報基盤が必要であるが、**これまでの学術情報基盤がそれに応えられているとは言えない**
 - 一部は利用制度の問題（改善しつつある）
 - 一部は計算機環境・技術の問題（究極的には重視する利用目的の問題）

参考：

- Ian Foster, Daniel Lopresti, Bill Gropp, Mark D. Hill, and Katie Schuman. *A National Discovery Cloud: Preparing the US for Global Competitiveness in the New Era of 21st Century Digital Transformation*
- “... DOE and NSF supercomputers provide access to powerful simulation capabilities, but *with access limited to small communities*. With a few notable exceptions, AI-ready datasets for research use are lacking. *Commercial clouds are accessible to anyone with a credit card, but there is little of the coordination needed to create nationally useful discovery cloud services. ...*”

3. ソフトウェアの進化速度に追隨できる基盤が提供できていない

- オープンソース中心のソフトウェアの進化は目覚ましく、**全ユーザが与えられた単一の環境を使うだけという（現在のスパコンの）モデルでは利用可能範囲、生産性が著しく損なわれる**
 - たった今は必要な機械学習フレームワーク+Jupyterのような環境を即座に用意できるか、というような問題
 - だがより本質的なのは将来の変化へ追従できること

4. 商用クラウド利用が解決策??

- **コスト**（次スライド）
 - 個々の研究者にとっても、国全体としても
- 海外クラウドベンダのシェアの大きさ
 - お金もデータも海外へという図式になる恐れ
- **性能**
 - 通信性能、資源割り当てなど商用クラウドで制御困難な要素
 - 高性能計算とのconvergeは商用クラウドでは困難
- **サポート・技術の蓄積・コミュニティ創成**
 - 計算資源があればいいわけではない

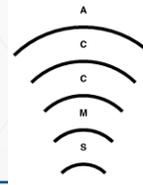
参考:あるクラウドベンダーのpricing

- (a) GPU: NVIDIA A100 8 GPU
 - $\approx 4000\text{円/h} \approx 10\text{万/day} \approx 300\text{万/月} \approx 3600\text{万/年}$
- (b) CPU: Ice Lake 64 vCPU
 - $\approx 300\text{円/h} \approx 7000\text{円/day} \approx 20\text{万/月} \approx 250\text{万/年}$
- 調達した場合（以下はmdxの場合だがスパコンも同様）
 - $\approx (a) \times 40 + (b) \times 368 + \text{ストレージ}$
 - コスト：ハードウェア**30億** + 保守・運用・光熱費（6年）**15億程度**
 - 上記クラウドベンダーのpricingでは**140億**（ストレージ含まず）

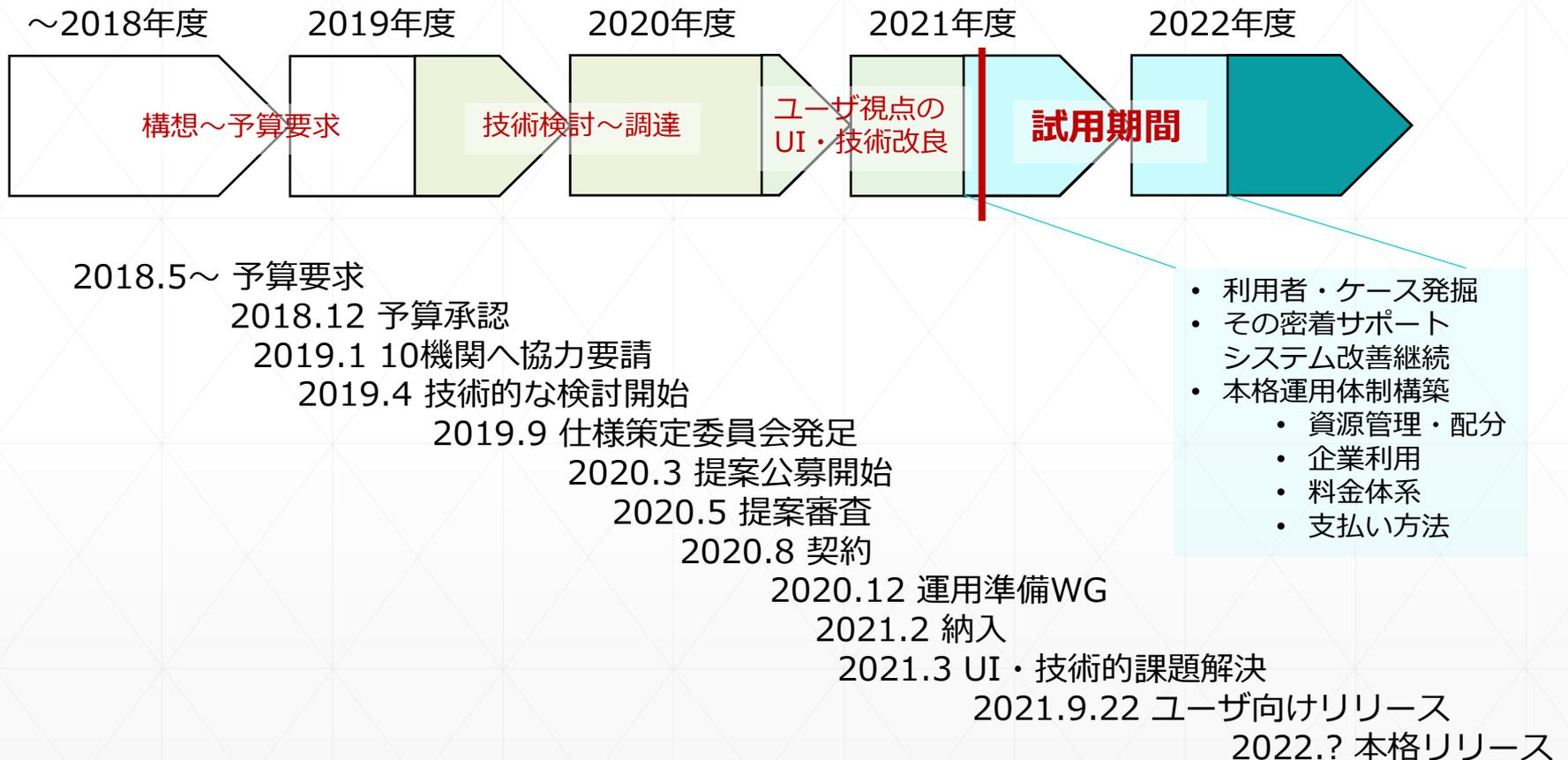
mdx の発足経緯・現状など

mdx

- 9大学2研究所が共同運営し、全国共同利用に供する、データ科学・データ駆動科学・データ活用応用にフォーカスした高性能仮想化環境 <https://mdx.jp/>
- @ 東京大学柏IIキャンパス



経緯

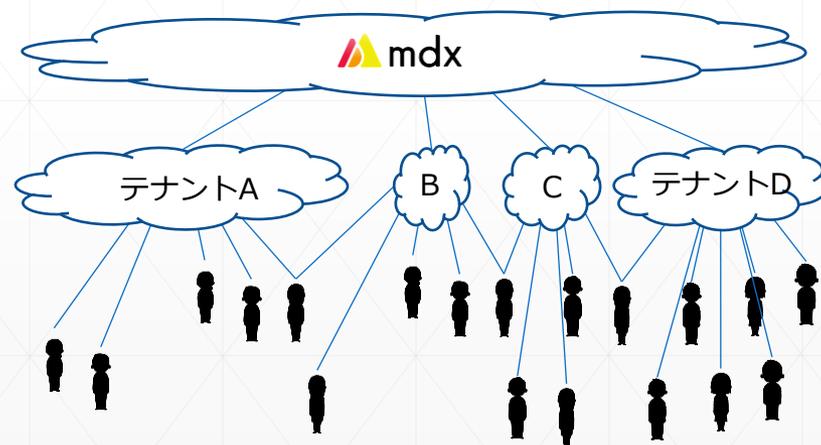
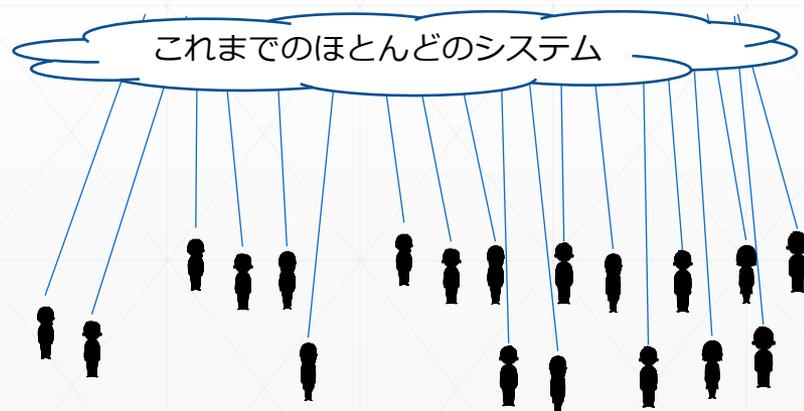


mdx（使用感紹介）

- 申請ポータルから利用申請（学認でサインイン）
- 申請が受理されたらユーザポータルから「仮想マシン」を生成
- 自分（たち）だけの環境が使える ⇒
 - 自前で調達したサーバと同様、自由に環境を構築可能
 - 他のユーザ環境と隔離されている（高セキュリティ）
- ハードウェア的にはそれなりの規模のものを共有 ⇒
 - 簡単・迅速に仮想マシンを追加（環境の拡張）可能
- 「Infrastructure as a Service」型のクラウド環境
- 学認でサインインし、ほかのサービス（e.g., Gakunin RDM）とシングルサインオンで連携

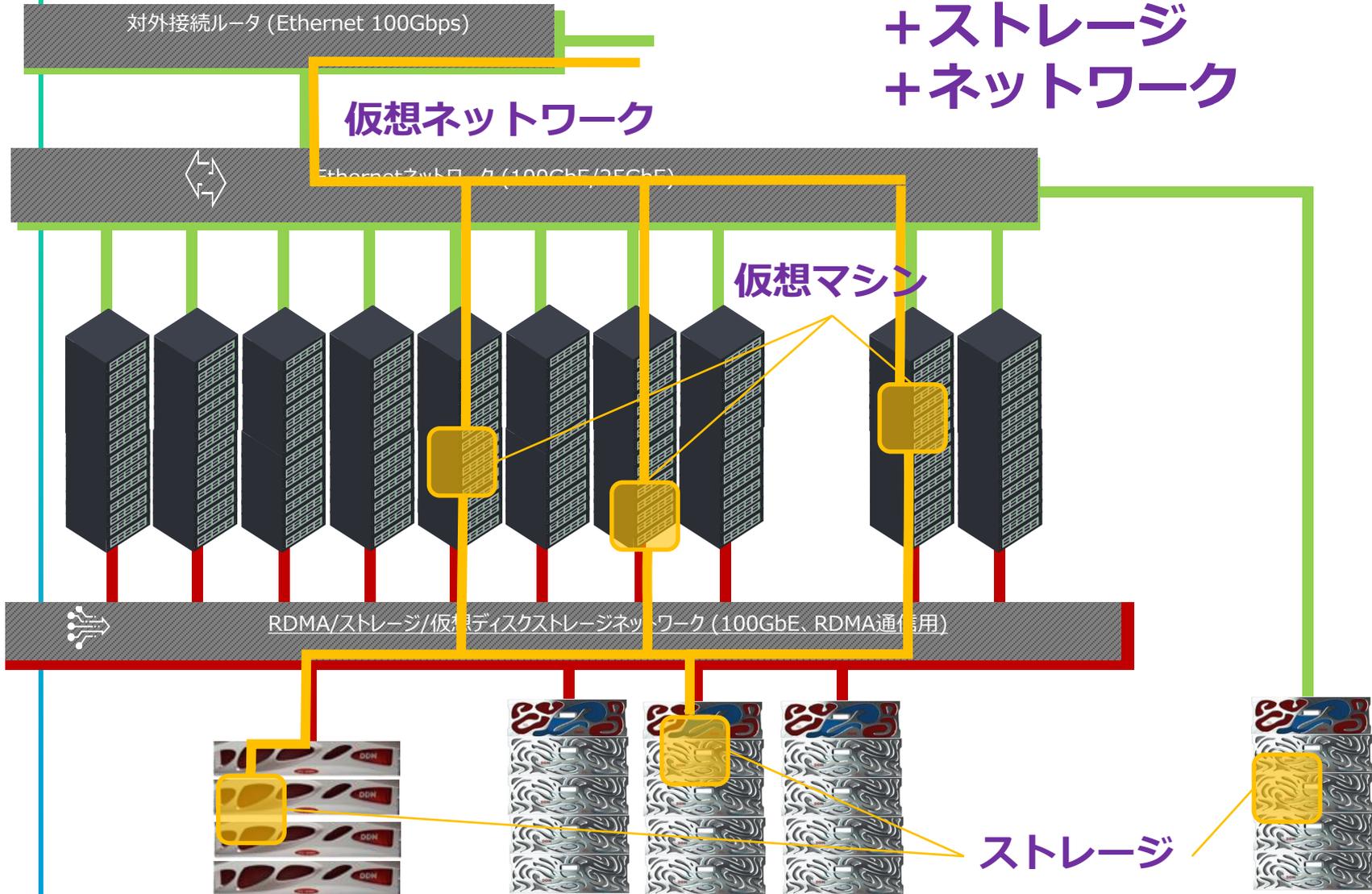
いわゆるスパコンとmdxの違い

- これまで (≈ いわゆるスーパーコンピュータ)
 - 全員が、管理者が設定する**単一の環境**を使う
 - 柔軟性がなく、目的、したがってユーザの分野が限定
- mdx : **仮想化された高性能環境**
 - 各分野、グループごとに合わせた環境を構築可能

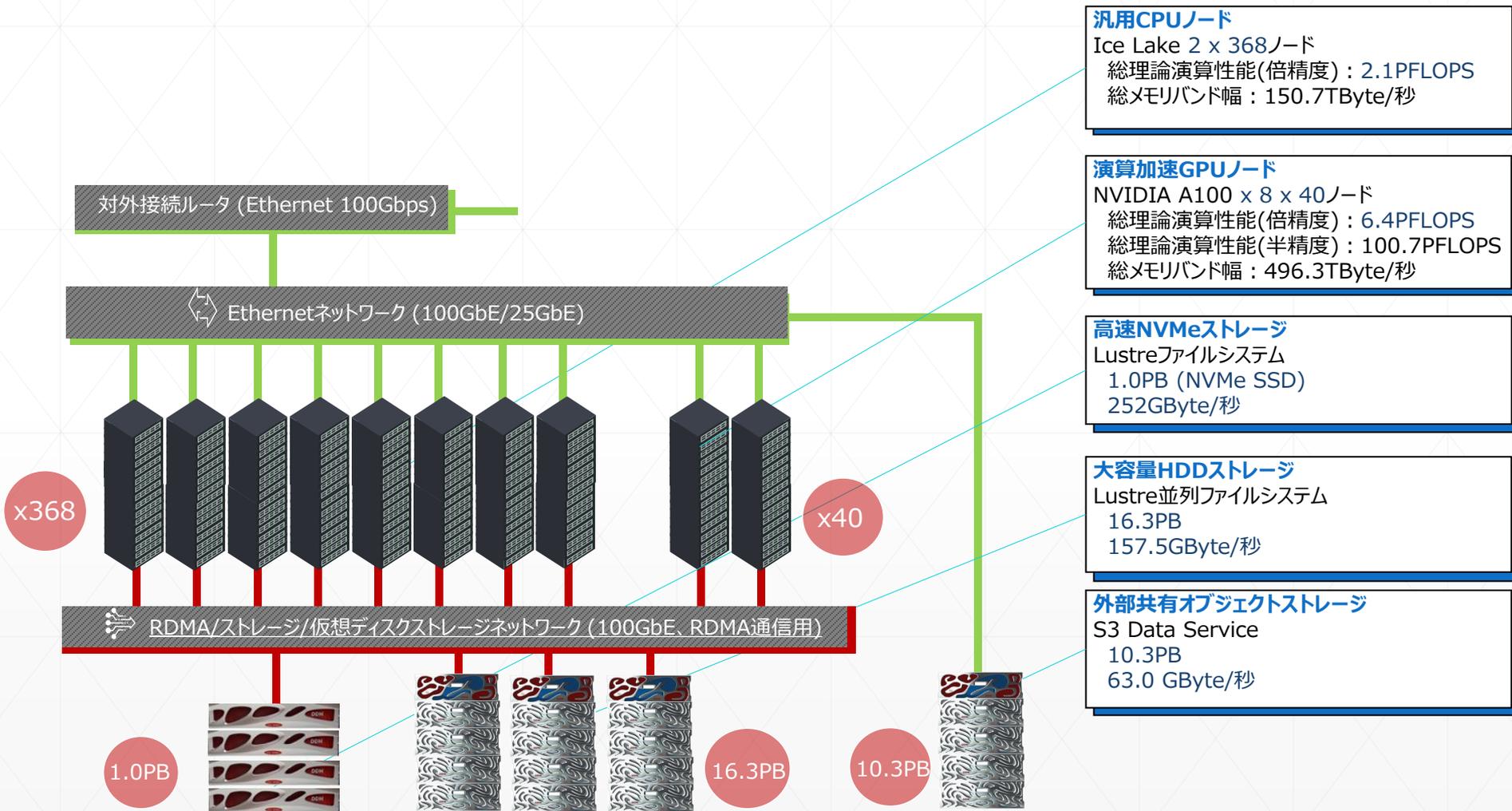


mdx利用モデル (図解)

テナント
≈ 仮想マシン
+ ストレージ
+ ネットワーク

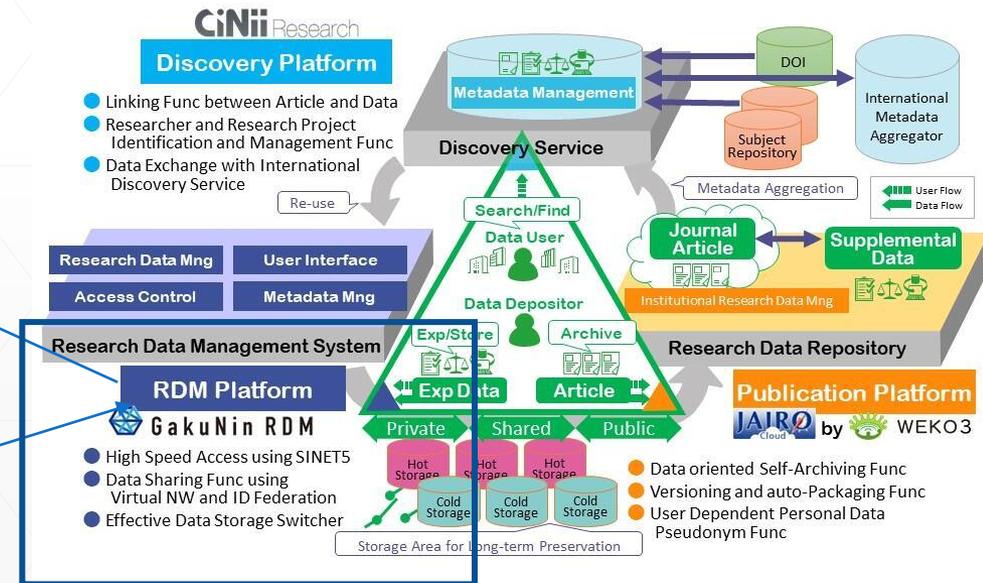
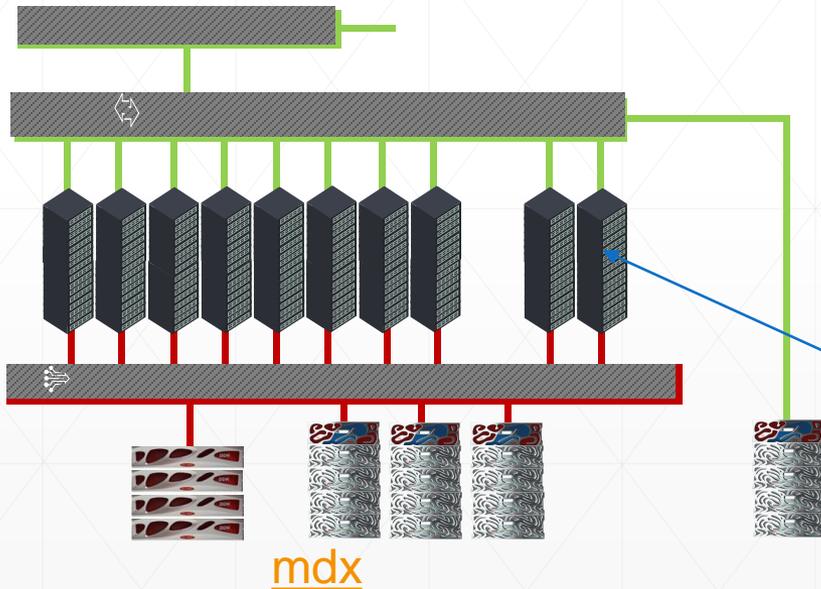


mdxハードウェアスペック



Gakunin RDMとの連携

- mdxの（オブジェクト）ストレージをGakunin RDMのストレージとして利用可能
- 研究データの管理（RDM）と利活用を一体化

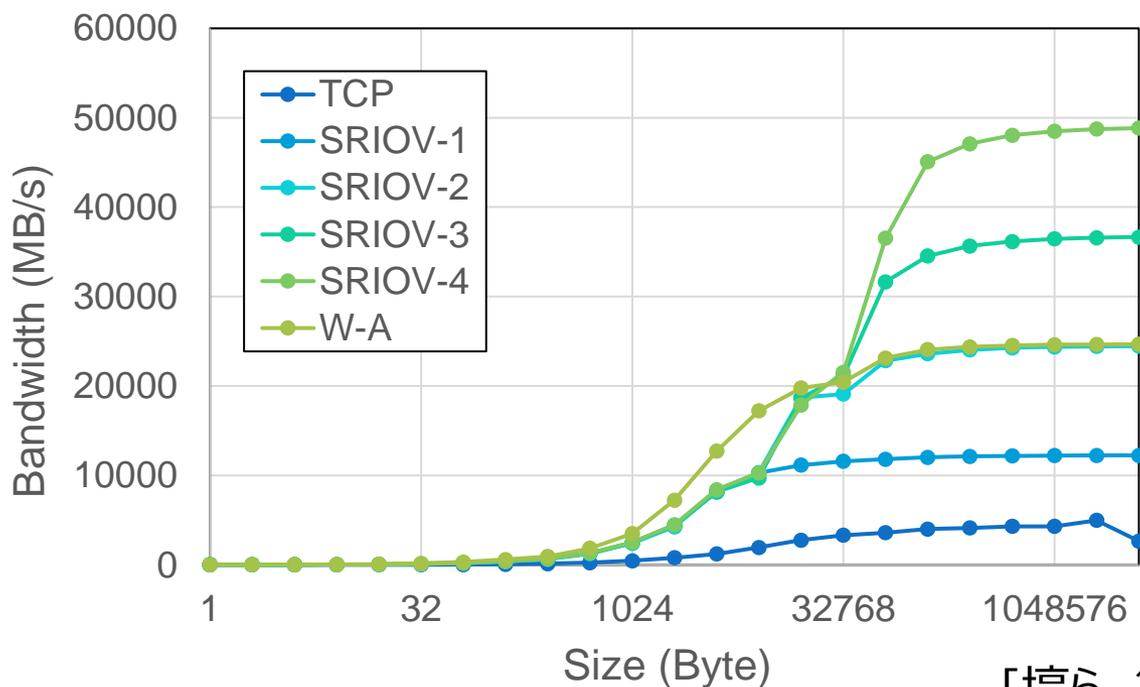


(進行中) ユーザごとに合わせた環境構築をサポートするための取り組み

- 環境自動構築 (Ansible playbook)
 - Special thanks to: 中村遼先生@東大、空閑洋平先生@東大、杉木章義先生@北大
- mdx REST API
 - 大江和一先生@NII, 竹房あつ子先生@NII, 工藤知宏先生@東大
- 高性能コンテナ (Kubernetes) 環境
 - 杉木章義先生@北大
- mdx内Jupyter (PaaS) 環境
 - 華井雅俊先生@東大

スパコンと比肩する通信性能

- 100Gbps RDMA over Converged Ethernet
- GPUノードには 4つのNIC (合計 400Gbps)



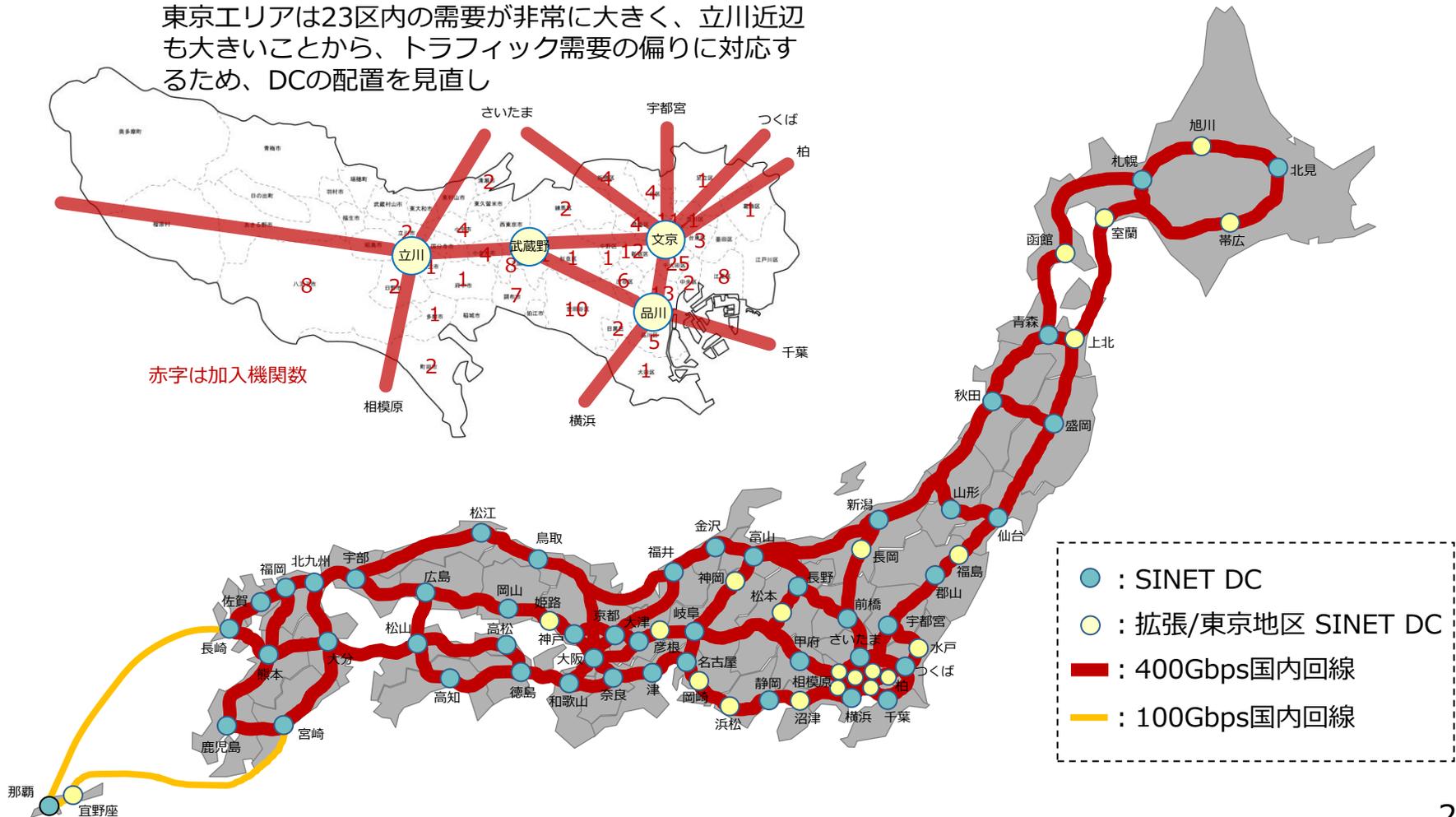
- 仮想環境であるが通信性能も高く、中規模までの並列計算にもスケール可能と期待

[埴ら、第183回HPC研究会]

- 国立情報学研究所が運用する、研究機関を接続する学術ネットワーク
- 国内ほぼ全県を400Gbps*で接続

* 沖縄は技術的な制約により当面100Gbpsベース

東京エリアは23区内の需要が非常に大きく、立川近辺も大きいことから、トラフィック需要の偏りに対応するため、DCの配置を見直し



SINET VPN

図提供：国立情報学研究所

- インターネットに加え閉域網が利用可能で、共考共創の高度なサービスも活用可能

大型実験施設等
(全国各地、海外各地)

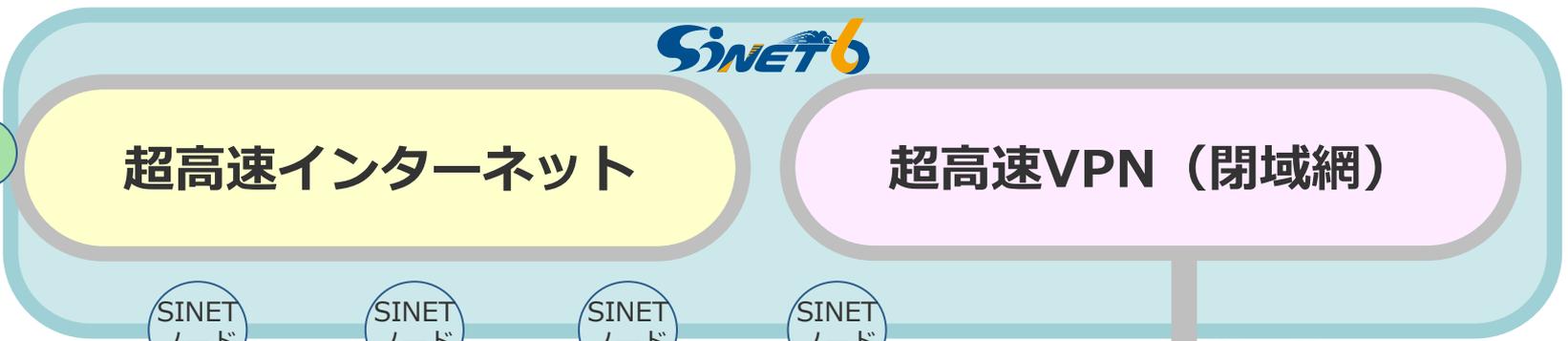
図は例

スパコン
(HPCI 13拠点)

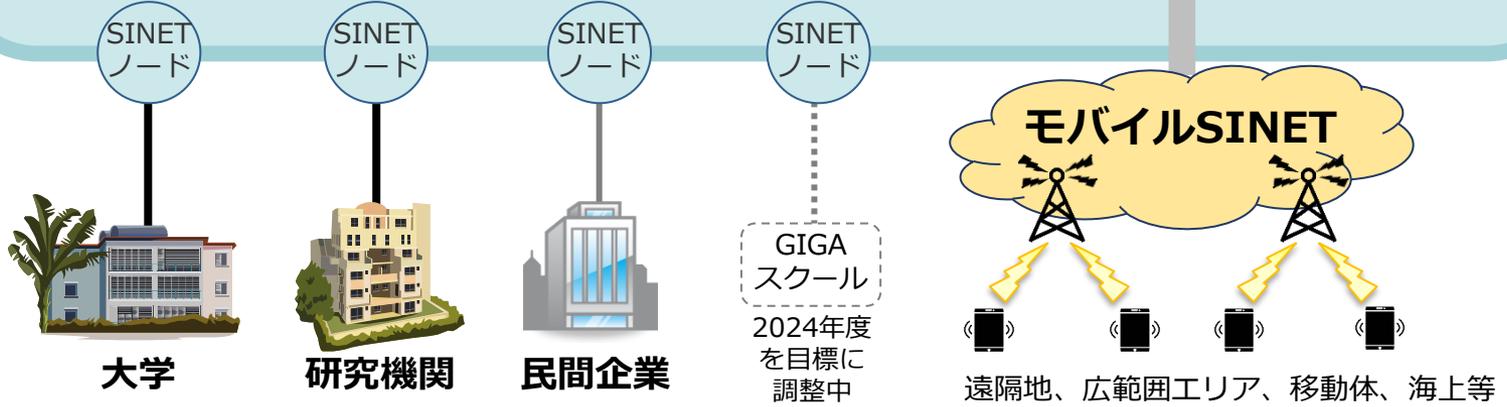
mdx

**研究データ
基盤**

直結クラウド
(34社45拠点)

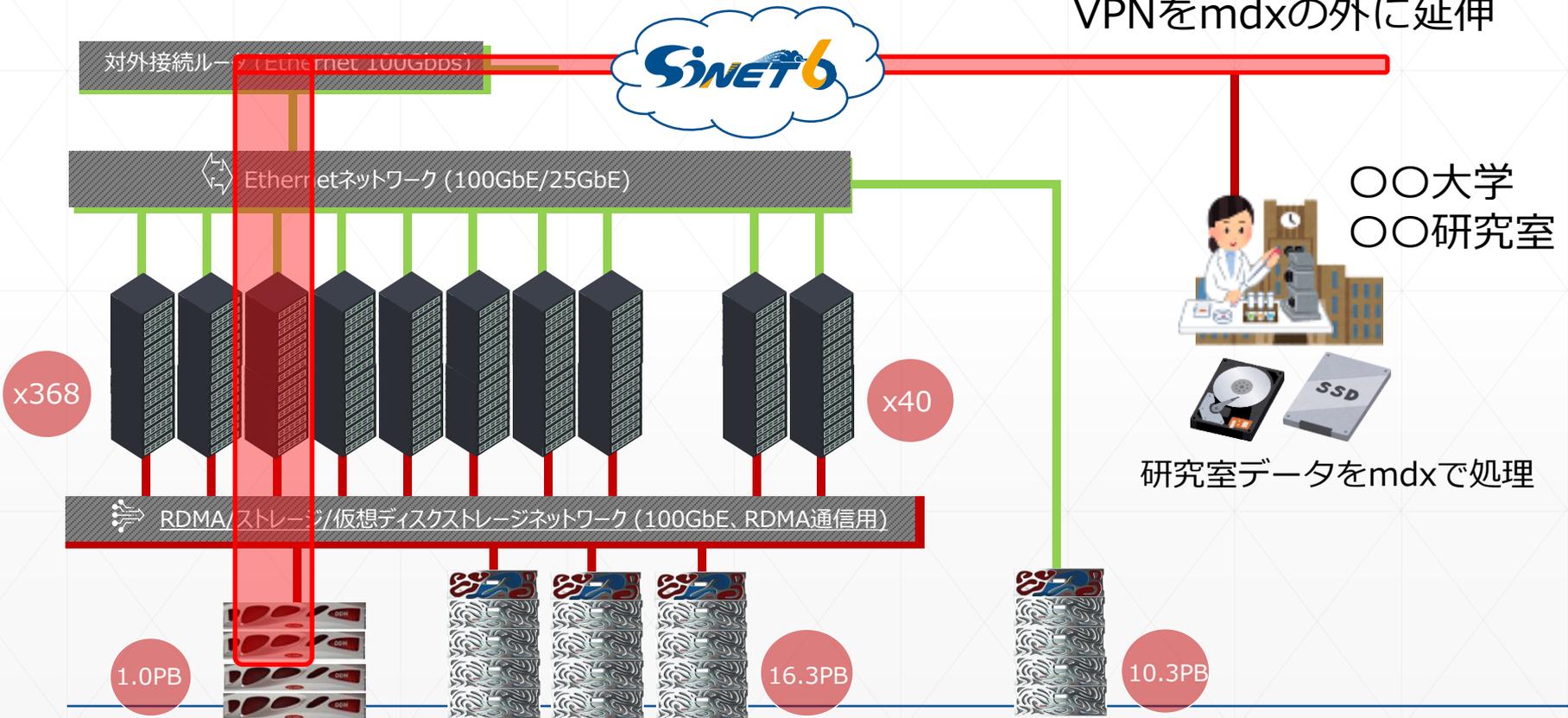


990 機関
(2022年3月末現在)



SINET VPNとmdx

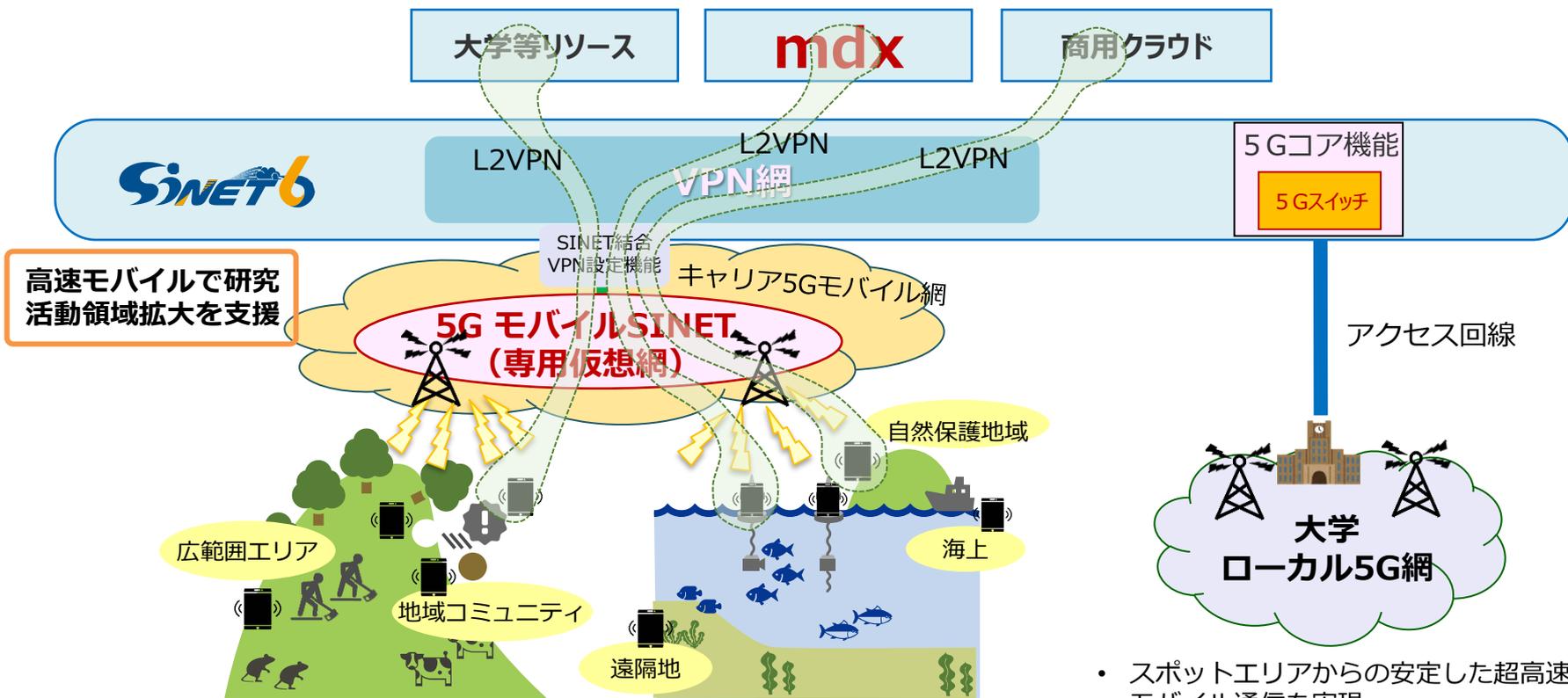
- **SINET VPN**によりmdx内に構築した環境を研究室、他大学のマシンと(同一LANにいるかのように)接続可能
VPNをmdxの外に延伸



mdxとモバイルSINET

図提供：国立情報学研究所

- モバイルSINET：商用モバイル網の中にSINET専用の仮想5G網（4G/3Gも可）を形成してSINET VPN網と接続することで、セキュアな通信環境を実現
- ローカル5G：SINET側に5Gコア機能を実装することで大学のローカル5G網の経済的な導入を支援（まずは小さな規模で実証実験を開始予定）



- 遠隔地、広範囲エリア、移動体、海上等を広くカバー
- セキュアな通信環境を実現

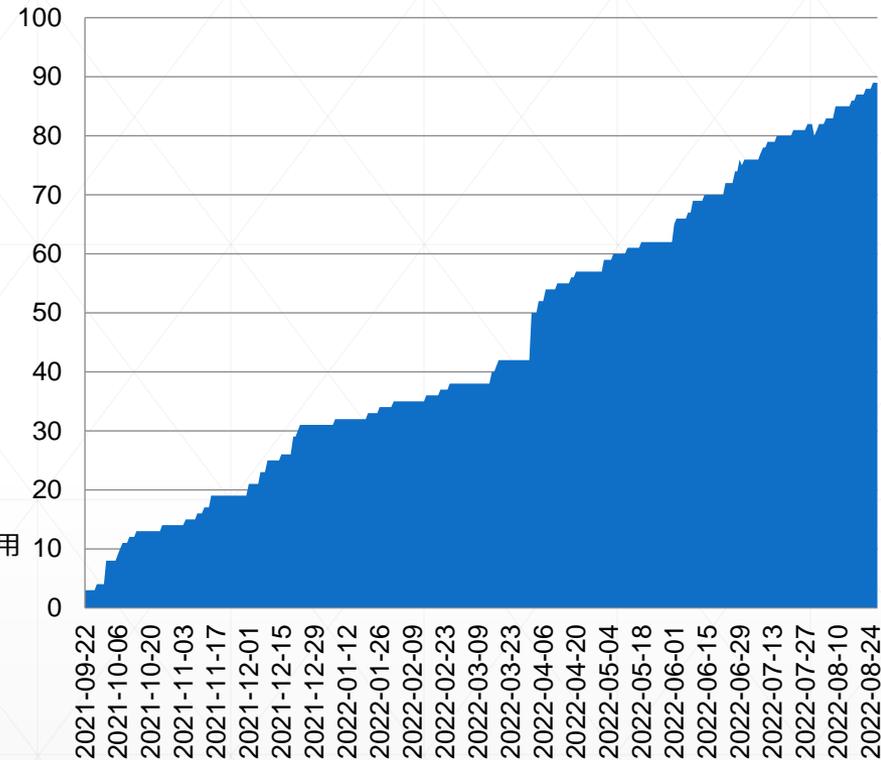
- スポットエリアからの安定した超高速モバイル通信を実現

mdxが目指すもの：まとめ

- 研究データ管理（研究者の日常）と隣接したデータ利活用環境
- 多様な研究分野に環境を提供
- 持続的な環境を提供
- オープンな環境を提供
- オンデマンドに、スケールする環境を提供

現在のユーザグループ（プロジェクト）数

- 財務ビッグデータの可視化と統計モデリング
- 医療・介護領域の人材マッチングに最適化された大規模グラフニューラルネットワーク
- エージェントモデルと統計データを用いた全国規模の疑似人流データの開発
- 大規模な日本語モデル構築・共有のためのプラットフォームの形成
- グラフニューラルネットワークとマルチタスク学習による汎用的物性予測モデルの構築
- ビヨンド・"ゼロカーボン"を目指し地域と技術をつなぐ情報基盤の構築
- 単語間に区切りのない書写言語における係り受け解析エンジンの開発
- 多次元高精細地表情報（MHESD）の地球科学・歴史考古学における高度利活用
- ...



プロジェクト名はJHPCN（後述）でのmdx利用課題（公開）

情報基盤センター群・JHPCN拠点に つてのmdx

共同利用・共同研究拠点の枠組み

- https://www.mext.go.jp/a_menu/kyoten/
 - 「個々の大学の枠を越えて大型の研究設備や大量の資料・データ等を全国の研究者が共同で利用したり、共同研究を行う」仕組み
- 現在100拠点が認定されている（一部は国際共同利用・共同研究拠点）
 - 物理系：
 - 宇宙線研（東大）、理論物理学研究拠点（京大）、…
 - 情報系：
 - JHPCN（8大学）、計算科学研究センター（筑波大学）、…
 - …

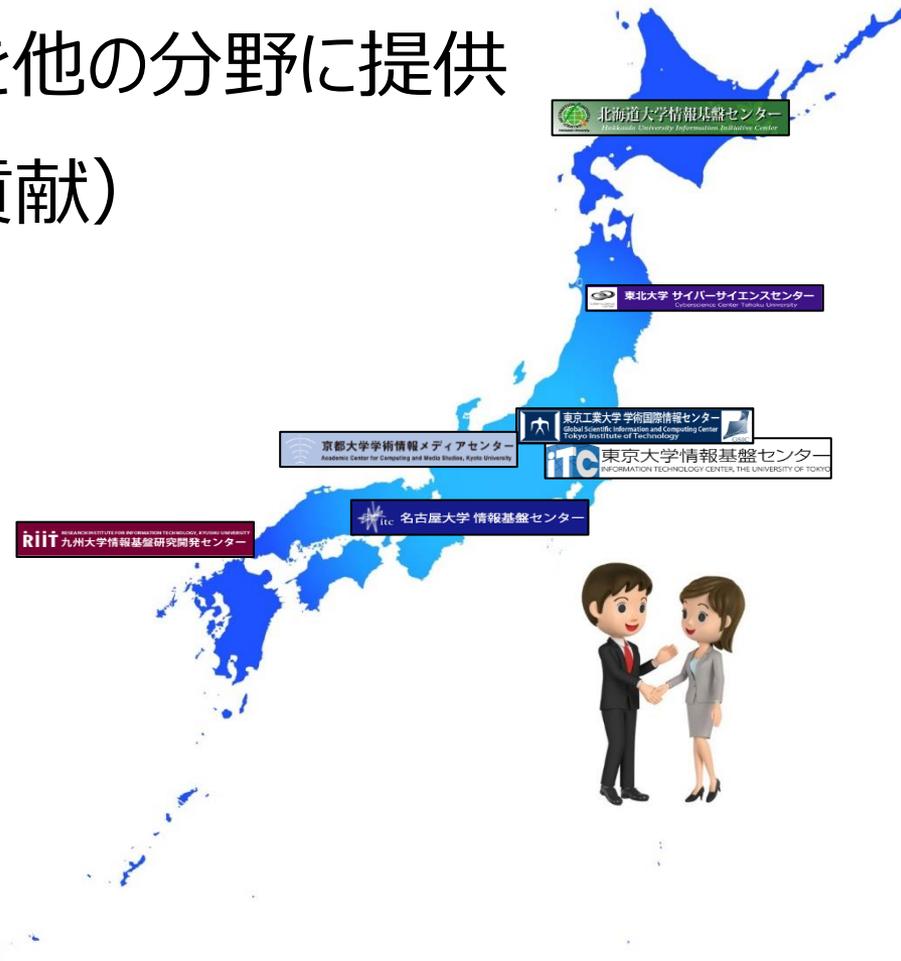
JHPCN（正式名：学際大規模情報基盤 共同利用共同研究拠点）

- 8大学の基盤センターが運営する共同研究のための組織
 - 北大、東北大、東大、東工大、名大、京大、阪大、九大
- 全国から共同研究を募集
 - 現在高性能計算シミュレーション中心の分野が多く集まる
 - 採択された課題に計算機利用時間割り当て
- 2022年度より「データ科学・利活用分野」募集を開始
- <https://jhpcn-kyoten.itc.u-tokyo.ac.jp/>

基盤センターとJHPCN拠点の使命

- 情報の専門家と情報基盤を他の分野に提供
- 情報学 x ○○学（学際貢献）
- コミュニティ形成

mdxはこの使命をこれまでのシミュレーション中心・計算科学から、**分野も基盤もセクターも**「データ駆動科学・データ活用分野」に広げるための第一歩



次世代計算基盤に関する一考察

本質的な目標

- 規模の拡大が可能（大規模な環境を共有）
 - 「1人で30」持つよりも「1000人で30000」持つ方がよい
- 多様な使い方をサポート（柔軟な環境構築）
 - 自分で調達すれば簡単に動かせるものが動かないことがない
 - ソフトウェア資産の継承・発展
- 我が国の技術蓄積
 - プロセッサ、ネットワーク、コンパイラ
 - 柔軟な環境構築（仮想化技術）との両立
 - 柔軟な環境構築のための高度な運用技術
- 経済安全保障

My Two Cents

	フラッグシップ	仮想化環境 (e.g., mdx)
科学技術計算に最適化されたプロセッサ	○ (A64fx)	
超低遅延ネットワーク	○ (Tofu)	
超大規模ワークロード (capability computing)	○ (Tofu)	
日本独自のプロセッサ・ネットワーク技術	○ (A64fx, Tofu)	
資源利用率	○ (バッチスケジューラ)	
環境構築の自由度		○ (ホスト仮想化)
インターネットとの接続性		○ (ネットワーク仮想化)
資源利用の自由度		○ (常時稼働ホストなど)
ソフトウェア資産		○

- 長期検討課題：両立 (converge) が可能か？
 - 我が国の競争力維持のための技術開発・産学共創課題となりうるか
- 目下の課題：次期計算基盤でそれらをどう両立させるか