

# mdx : データ科学、データ駆動科学 のための情報基盤

田浦健次郎

JHPCN共同利用共同研究拠点 拠点長  
東京大学情報基盤センター長



# 自己紹介

- 東京大学
- 本務 情報理工学系研究科
- 兼務 情報基盤センター長（2018.4～）
  - JHPCN（共同利用共同研究拠点）統括拠点長
  - HPCI（コンソ、計画推進委員会、第2階層まとめ役、etc.）
  - mdx仕様策定委員長
- 研究分野：システムソフトウェア（とくに並列処理、並列プログラミング処理系、データ処理系）
- 好きなソフト：SSH, Emacs, make, sqlite3

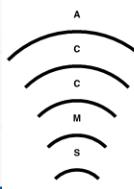
# mdxとは

- 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
- もうすぐ… デビューします



**NII** 大学共同利用機関法人 情報・システム研究機構  
国立情報学研究所  
National Institute of Informatics

 産総研



# mdx (名前) について

-  mdx は何の頭文字? dx は DX?
-  mdx と決めた時には DX なんて言葉は流行ってませんでした
- 今のところいくつかの言葉の量子重ね合わせ状態になっています
  -  mdx
    - = Massive Data Transformation (platform)
    - + Massive Data eXchange (platform)
    - + なんか Digital Transformation (platform)
    - + … (良いアイデア募集) …

# mdxのハイレベルゴール

- **共同研究 5.0**

- ほぼ計算科学  $\Rightarrow$  {計算科学+データ科学}

- **大学マシン 5.0**

- ほぼシミュレーション、DL  $\Rightarrow$  データ利活用、データプラットフォーム構築、Society 5.0実現へ向けたサービス構築、基盤ソフト研究に使える基盤

- 簡単に、迅速に使える基盤

- **大学とベンダの協働 5.0**

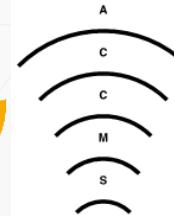
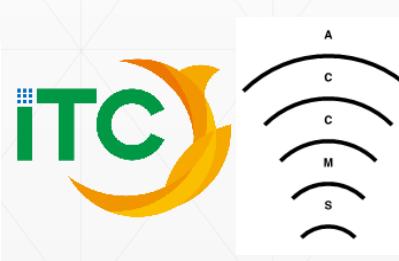
- 大学マシン整備のための大学とベンダの新しい協働の形

# 共同研究 5.0

# 8大学の共同利用・共同研究拠点 (JHPCN) について



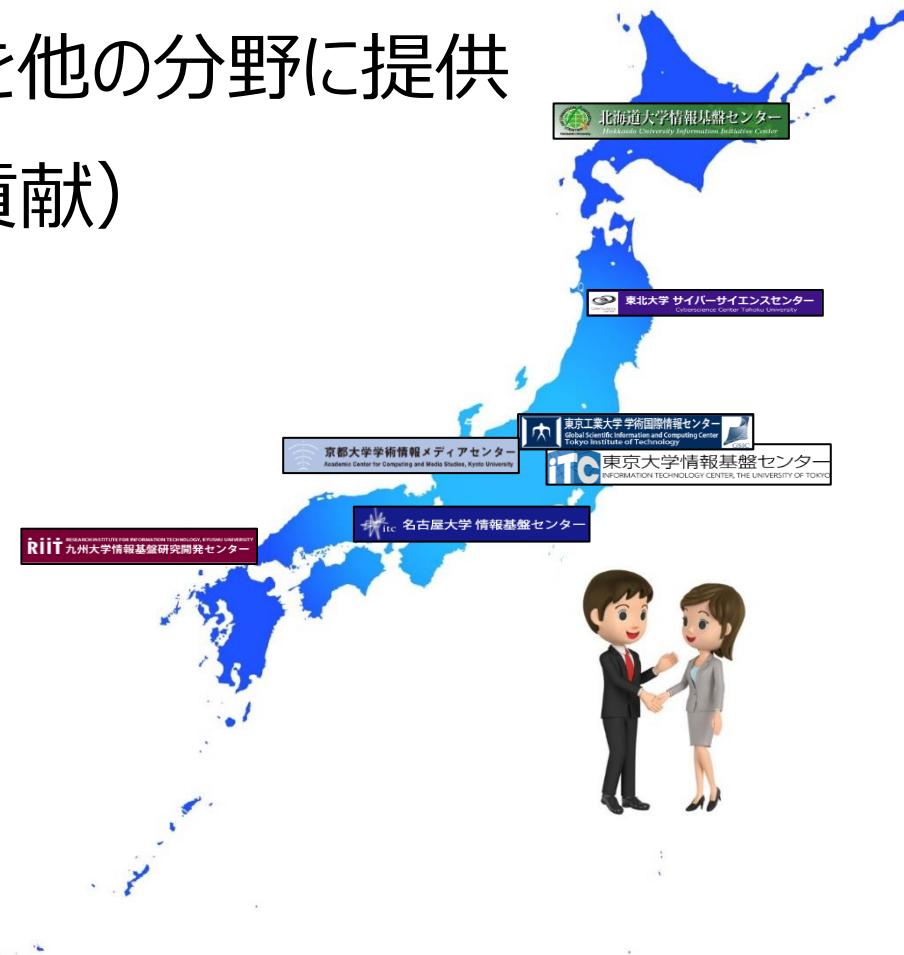
- 8大学の基盤センターが運営する共同研究のための組織
  - 北大、東北大、東大、東工大、名大、京大、阪大、九大
- 全国から共同研究を募集
  - 採択された課題に計算設備利用時間（主にスパコン利用時間）を割り当てる
  - 現在シミュレーション中心の計算科学分野が多く集まる



# 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 × ○○学 (学際貢献)
- コミュニティ形成

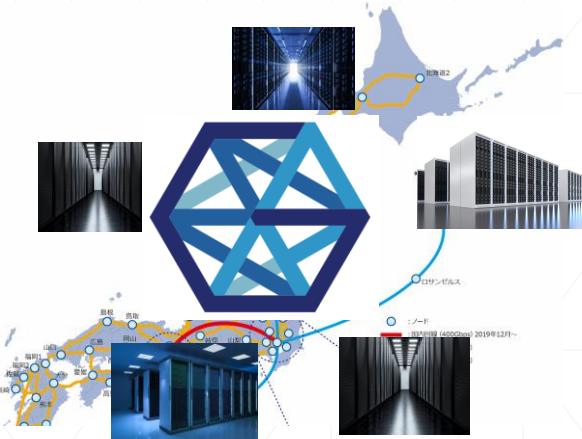
mdxはこの使命をこれまでのシミュレーション中心・計算科学から、**分野も基盤もセクターも**「データ駆動科学・データ活用分野」に広げるための第一歩



# JHPCNとmdx

- JHPCNはこれまで計算科学中心の学際的共同研究で実績がある
- データ科学分野の共同研究を歓迎し、そのための受け入れ態勢(拠点のexpertise)も強化したい

## 活用分野



### 共通ディシプリン：計算科学

原理のシミュレーション中心  
核物理、材料・物性、生体分子、  
燃焼、気象、地震、宇宙、etc.

### 情報基盤

高性能計算機、ストレージ (スパコン) 中心

### 情報系研究分野

数値計算アルゴリズム、  
HPC中心

### 共通ディシプリン：データ科学

データ解析、モデリング中心  
医療、経済、空間、モビリティ、環境、歴史 (アーカイブ等) 、ソーシャルデータ、etc.

データ共有基盤、データ解析環境、データプラットフォーム構築、実時間データ収集、モバイルネットワーク

機械学習、NLP, 画像処理、大規模データ処理、ネットワーク、クラウド、etc.

# そのために必要なこと

- データ科学・活用は常に分野やセクタをまたがる横断的な活動
- 1-1の共同では済まないこともしばしば

このデータでこんなことがわかるはずだがプログラミングできる学生が…

このデータ、価値を生み  
そうだけど具体的には?

必要なのは  
人をつなぐ仕組み

とある分野研究者A

年〇PBのデータを蓄積・  
バックアップするストレージの運用とかどうする?

とある分野研究者B

ある企業事業部  
リズムはできただけど問  
題設定これでいいのかな?

実データはどこ? こ  
のフィールドの意味  
は?

とある情報研究者

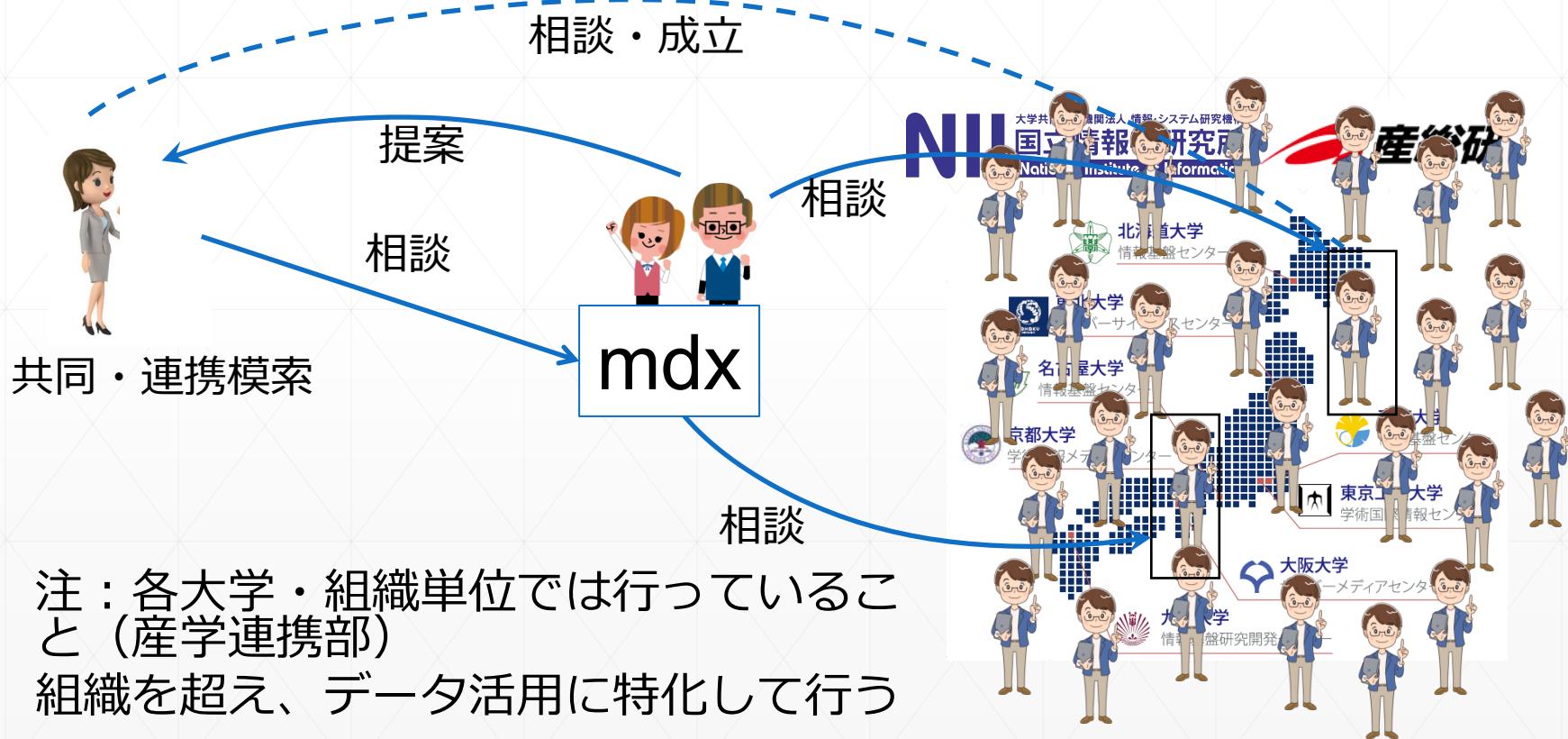
# JHPCNの共同研究の成立過程



- マッチングは応募前になされている
- 似た分野・知った研究者同士の共同向け

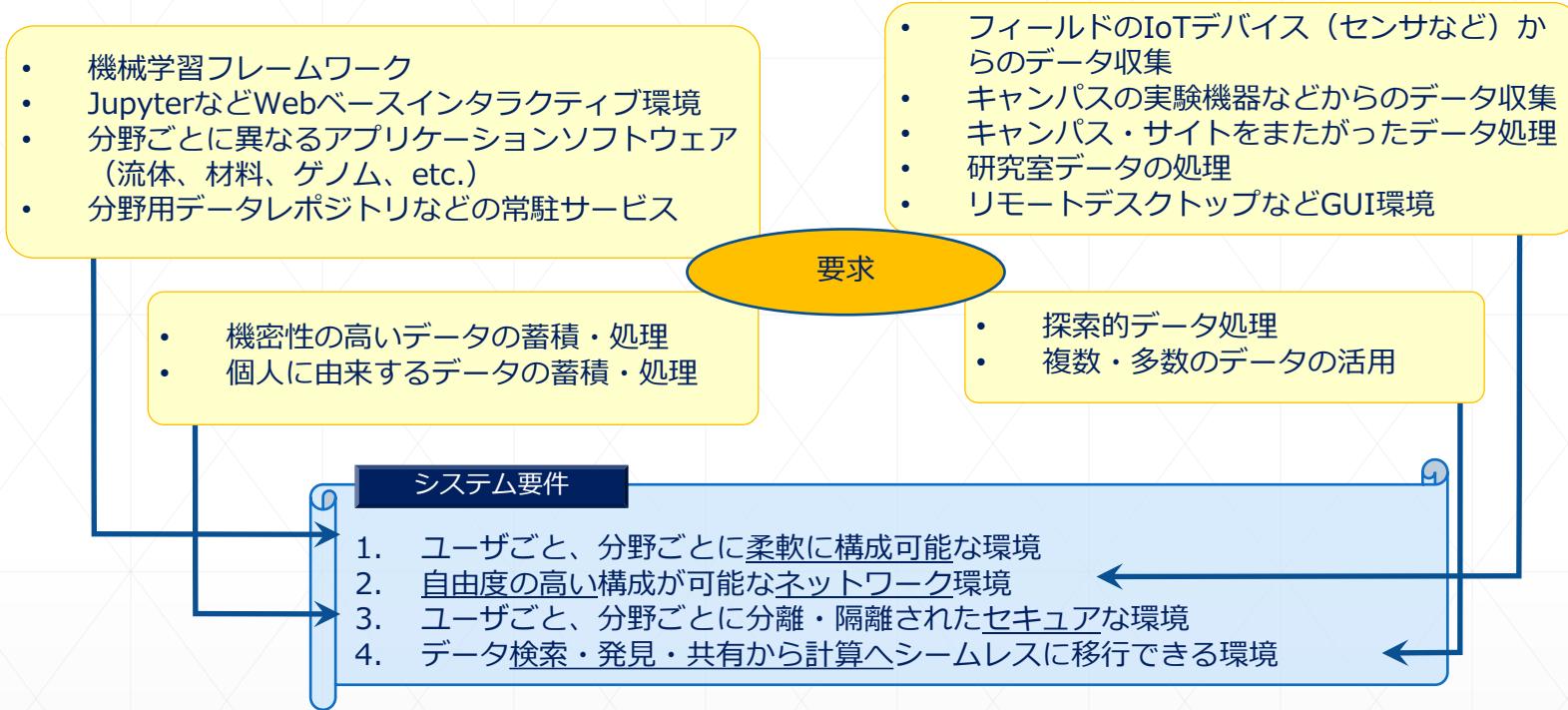
## 今後目指すマッチング

- チーム成立前のマッチングの相談を受付



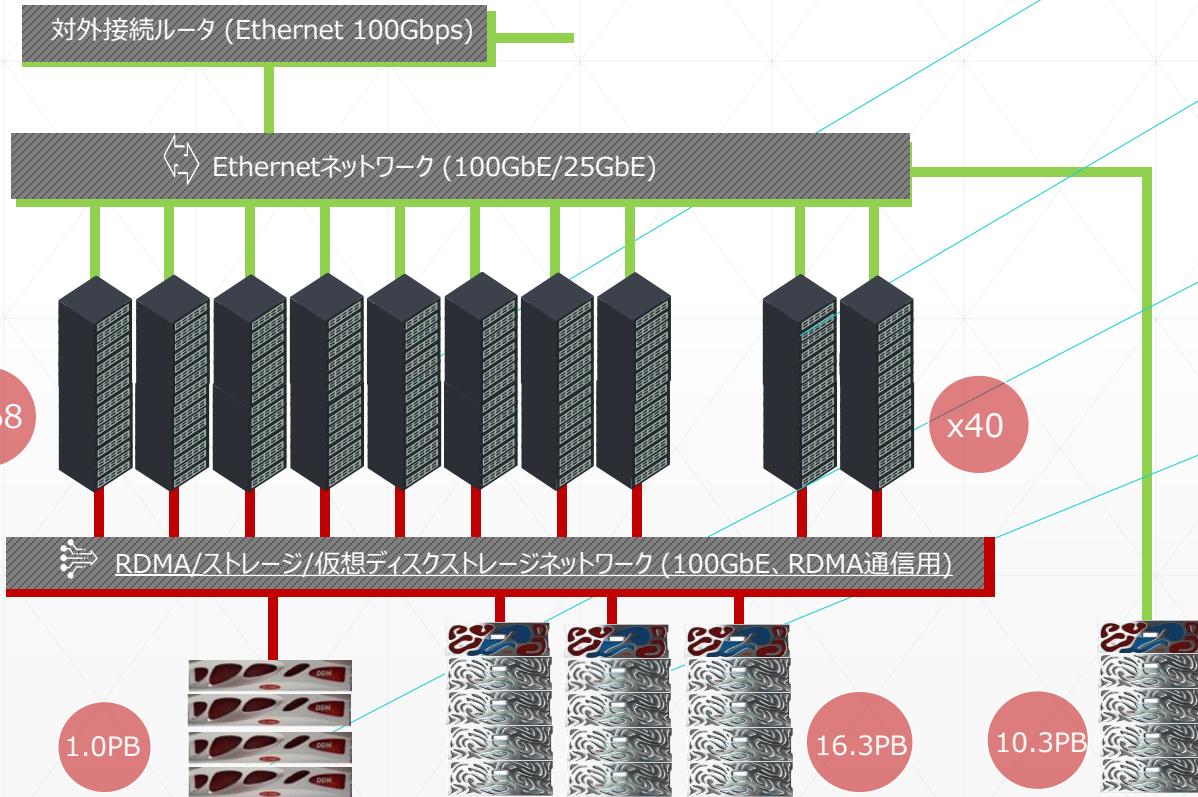
# 大学マシン 5.0

# データ科学のインフラに求められる要件



- ≈ 仮想化されたマルチテナント（クラウド）環境
  - データを蓄積・共有できるだけでは足りない
    - 高性能計算機と一体化の必要（特にML・AI利用）
  - これまでの高性能計算機だけでも足りない
    - GPUとMLフレームワークがあれば済むわけではない

# mdxハードウェアスペック



**汎用CPUノード**  
Ice Lake 2 x 368ノード  
総理論演算性能(倍精度) : 2.1PFLOPS  
総メモリバンド幅 : 150.7TByte/秒

**演算加速GPUノード**  
NVIDIA A100 x 8 x 40ノード  
総理論演算性能(倍精度) : 6.4PFLOPS  
総理論演算性能(半精度) : 100.7PFLOPS  
総メモリバンド幅 : 496.3TByte/秒

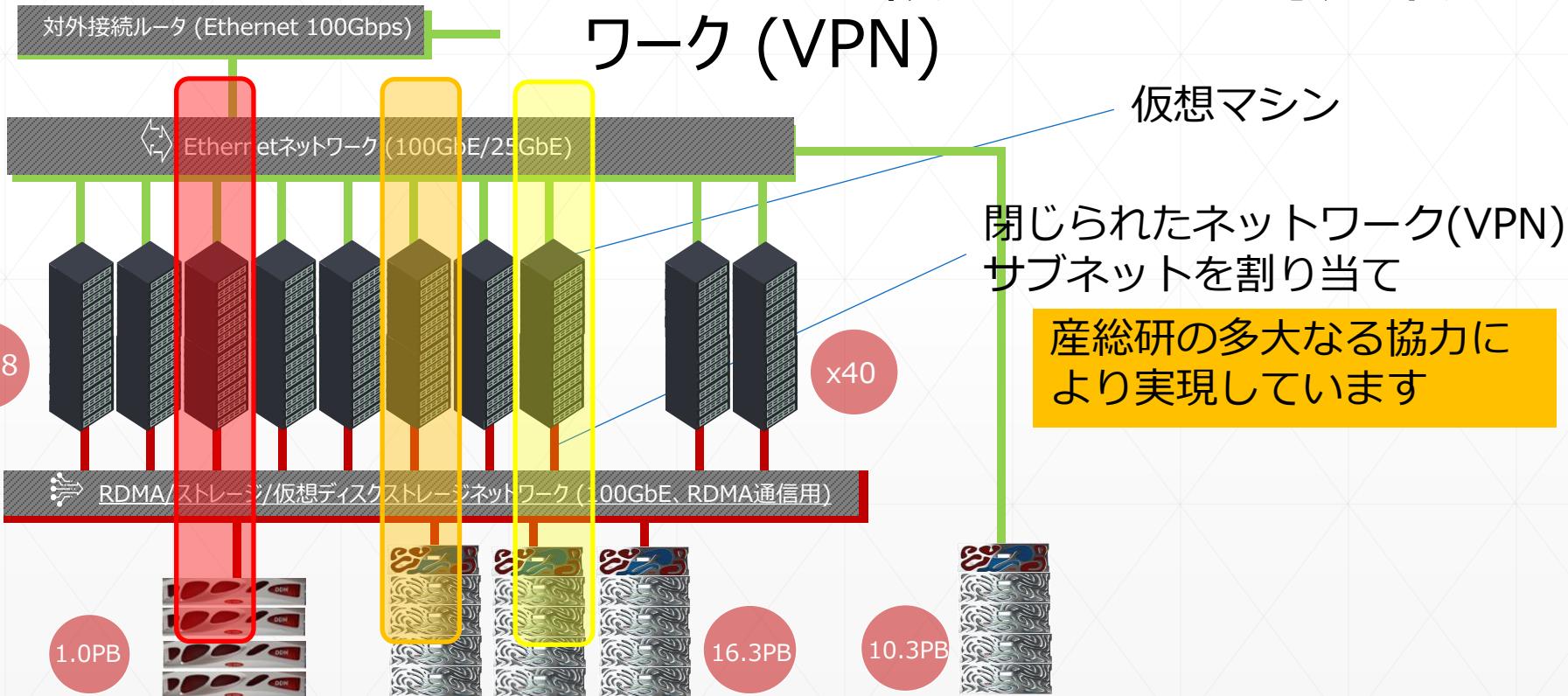
**高速NVMeストレージ**  
Lustreファイルシステム  
1.0PB (NVMe SSD)  
252GByte/秒

**大容量HDDストレージ**  
Lustre並列ファイルシステム  
16.3PB  
157.5GByte/秒

**外部共有オブジェクトストレージ**  
S3 Data Service  
10.3PB  
63.0 GByte/秒

# 仮想化環境

- 仮想化(VMware)により利用グループごとに占有環境(テナント)を提供
- テナント = 仮想マシン + 専用ネットワーク (VPN)



# 仮想化 ⇒ 使いやすさ・セキュリティ向上

## 1. 仮想化による柔軟性

- ホストを（管理者権限で）自由に設定可能

## 2. VPN (VLAN) による柔軟性

- プロジェクトごとにネットワークの延伸が可能

## 3. 仮想化・VPNによるセキュリティ（分離）

- 他のグループの環境のオペミスや脆弱性が他に影響しない

## 4. 環境の互換性・可搬性を高められる

- VMテンプレート (cf. MateriApps)

スパコン…なんか不便…イメージを過去のものに…

# デモビデオ

- 利用申請 (プロジェクト作成)
- 仮想マシンの作成～起動
- ネットワーク設定～SSH
- デスクトップ環境、Jupyter環境

# 複数環境と環境構築フレームワーク

- 複数の設定済み仮想マシンを提供
  - Ubuntu, CentOS, …
  - 空の仮想マシン (isoから好きなOSをインストール)
- 環境構築を自動化する仕組みを提供
  - Ansible利用
  - 各種用途用のAnsibleのレシピ (Playbook) を提供、コミュニティで開発、共有
- 東大情報基盤センター 空閑先生、中村遼先生

# 使いやすさ: 様々な利用深度を想定

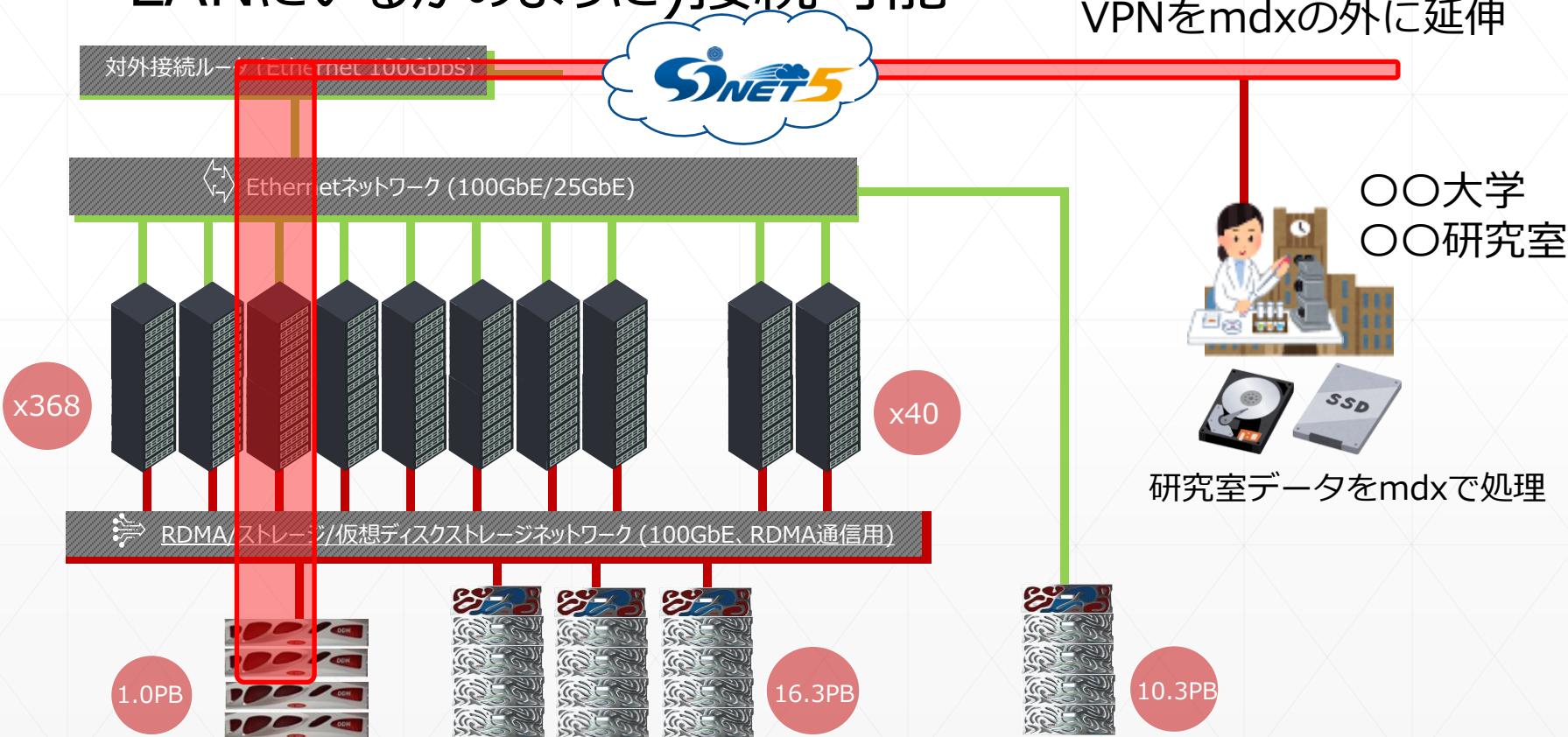
- 対話的利用 (≈ Kaggle, Colab)
  - 規定の環境 (Jupyterhubなど) で直ちに利用開始
  - Gakunin RDMで共有したデータを処理する環境をBinderHubで起動
  - NII藤原先生、東大情報基盤センター松島先生
- 独自環境 (≈ AWS IaaS)
  - 独自にカスタマイズした環境 (Jupyter, 特定MLフレームワーク、etc.)
  - その他、パブリックなIaaS同様、自由に環境を構築可能
  - そのための仮想マシンのテンプレート、そのレシピを公開 (コミュニティで開発・共有)
- 中規模クラスタ環境
  - 多数のノードからなる並列・分散処理環境 (ビッグデータ、中規模AI・ML、データ収集・蓄積・サービス環境)
  - そのためのレシピも共有

## 使いやすさ: 柔軟性・自由度

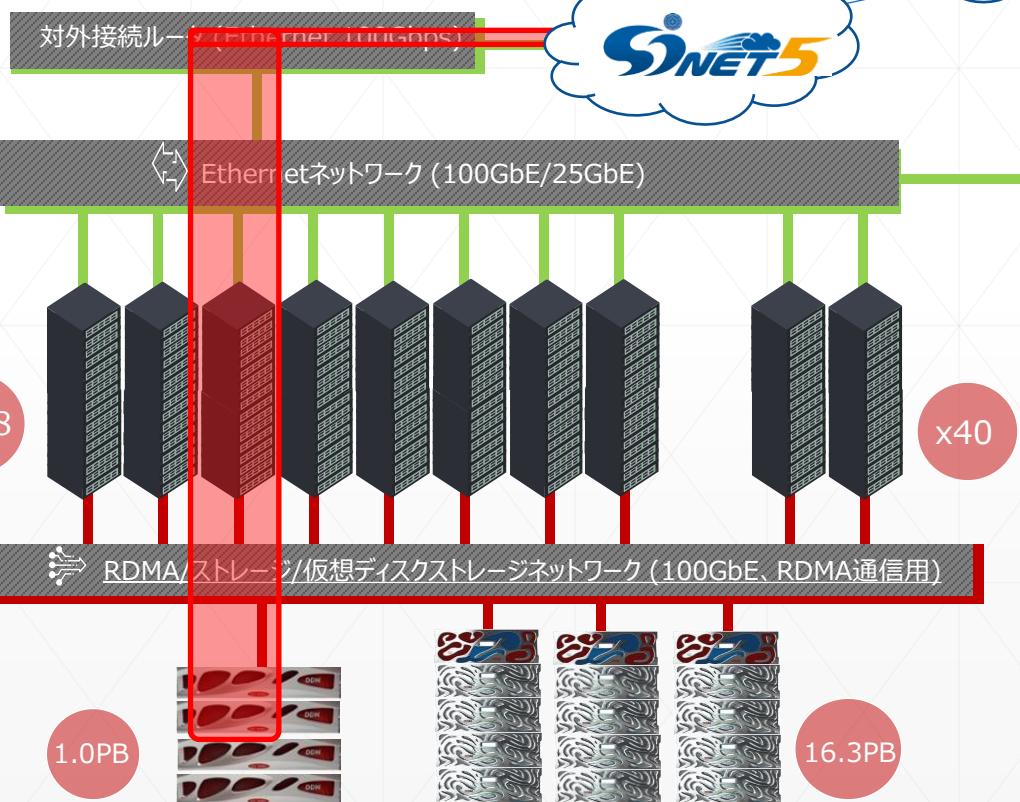
- クラウド風の使い方(バッチスケジューラではなく)
- ≈ 仮想マシンを長時間、常時稼働可能
  - = 申請した分の仮想マシンは一定時間以内に起動すること(実効的な常時稼働)を保証 ⇒ 起動保証VM
- + 「資源が空いている間のみ稼働できる」仮想マシン  
⇒ スポットVM
  - 計算資源が必要、実時間性は不要なジョブ(≈ バッチジョブ)
  - 連携マシン(ABCI, Wisteria/BDEC01)での実行も可能
- 外部通信
  - 各テナントが外部(インターネット)との通信を制御可能

## SINETとの連携

- SINET L2VPNにより研究室, 他大学のマシンと(同一LANにいるかのように)接続可能



# モバイルSINETとの連携



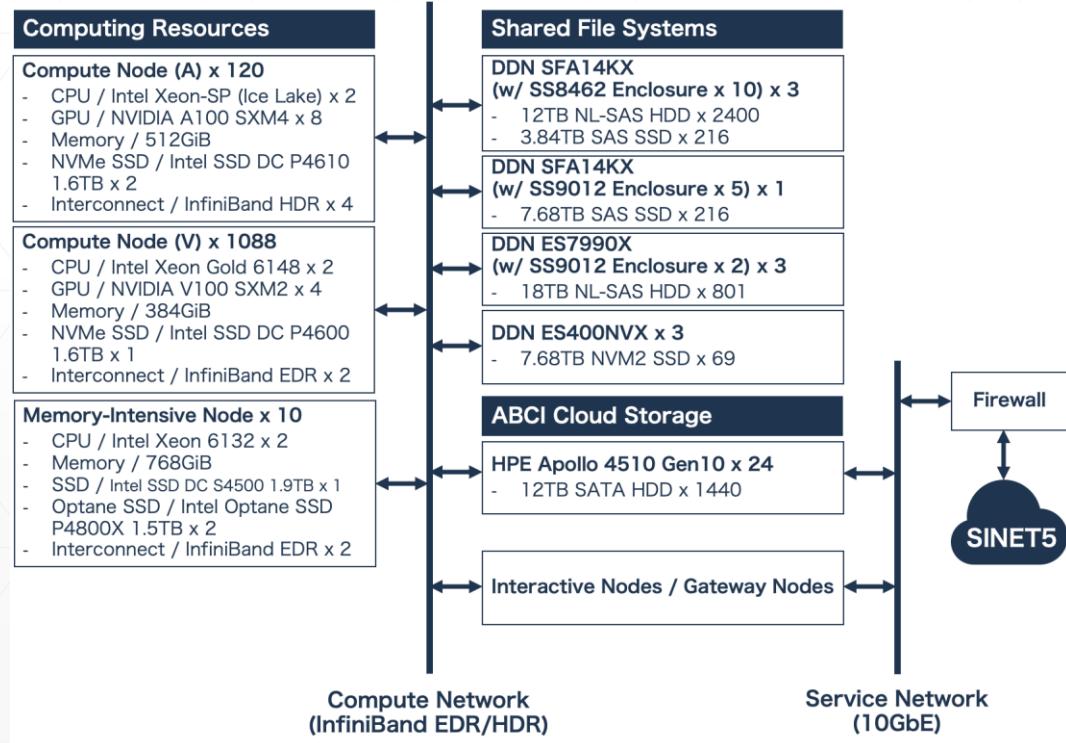
- モバイルSINET ≈ L2VPNをモバイルネットワークに延伸
- 実験機器、野外や実験室のセンサなどからデータを安全に収集可能

# 産総研ABCIとの連携

- mdxと同一キャンパス
  - 共有オブジェクトストレージを通じた連携を予定
- NVIDIA A100 x 8 x 120
- NVIDIA V100 x 4 x 1088



[https://abci.ai/en/about\\_abci/computing\\_resource.html](https://abci.ai/en/about_abci/computing_resource.html)



# 東大 Wisteria/BDEC01 との連携

- mdxと同一建物
- シミュレーションノード部 (Odyssey)
  - Fujitsu A64FX x 7680ノード (25.9 PFLOPS)
  - ファイルシステム経由で連携
- データ・学習ノード (Aquarius)
  - NVIDIA A100 x 8 x 45ノード (7.2 PFLOPS)
  - ファイルシステム経由で連携
  - GPUを占有利用し, そこにmdxテナントのネットワーク(VPN)を延伸することも検討中

# つなげるを前提とした設計

-  **GakuNin** 国内多くの大学・研究機関ですぐにサインアップ可能
-  **GakuNin RDM** と連携
  - 様々なストレージ, データ・コードレポジトリとの連携機能
-  **mdx**
  - Gakunin RDMの外部ストレージとして連携
    - まずはS3共有オブジェクトストレージ
    - (次) 内部大容量ストレージ(Unixファイルシステム)との連携
    - (進行中) BinderHubを用いたストレージ-計算機の連携

NII藤原先生  
東大松島先生

公開基盤(JAIRO), 検索基盤  
(CiNII)との連携も協議中

# Gakunin RDMを用いたデータ共有とmdx

## 1. 「プロジェクト」を作成

## 2. プロジェクトにストレージを接続

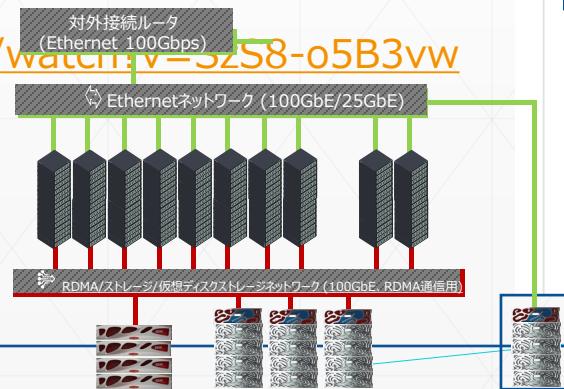
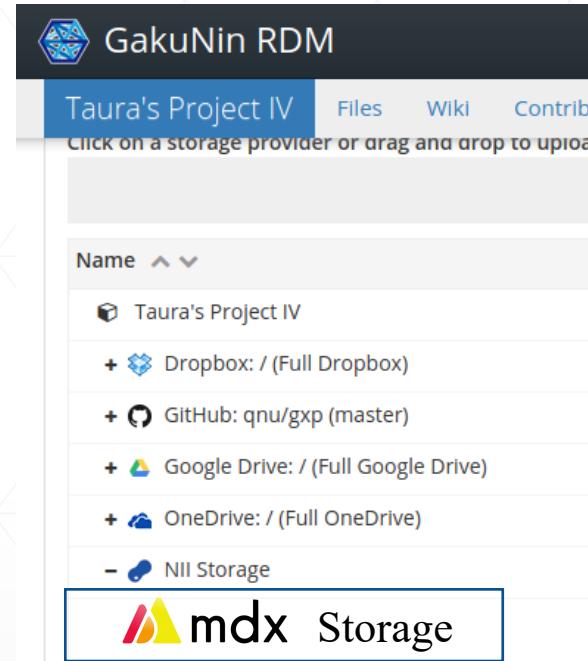
- Google Drive, Microsoft OneDrive, Dropbox, Githubなどを接続可能
- Nextcloud, S3互換ストレージなど、オンラインストレージの接続も可能
- NIIデフォルトストレージを提供

## 3. プロジェクトに共同研究者を招待

- 参考: 北大、名古屋大

<https://www.youtube.com/watch?v=3zS8-o5B3vw>

mdxがGakunin RDMと接続するストレージと、それを処理する計算基盤を提供

GakuNin RDM

Taura's Project IV

CLICK on a storage provider or drag and drop to upload

Name

- Taura's Project IV
- + Dropbox: / (Full Dropbox)
- + GitHub: qnu/gxp (master)
- + Google Drive: / (Full Google Drive)
- + OneDrive: / (Full OneDrive)
- NII Storage

mdx Storage

# 将来像 (mdxの先) NIIと情報基盤センター群のマシン間連携



- 将来的には多数の大学の共同利用マシンがGakunin RDMと接続
  - 大規模データの共有～直接処理までを迅速にできる環境



GakuNin RDM

Taura's Project IV Files Wiki Contrib

CLICK ON a storage provider or drag and drop to upload

Name ▾ ▾

- Taura's Project IV
- + Dropbox: / (Full Dropbox)
- + GitHub: qnu/gxp (master)
- + Google Drive: / (Full Google Drive)
- + OneDrive: / (Full OneDrive)
- NII Storage

mdx Storage

- 大学 □□
- △△大学 ●●
- 大学 ◆◆
- ▲▲大学 ☆☆
- ...

# 将来像 (mdxの先) NIIと情報基盤センター群のマシン間連携

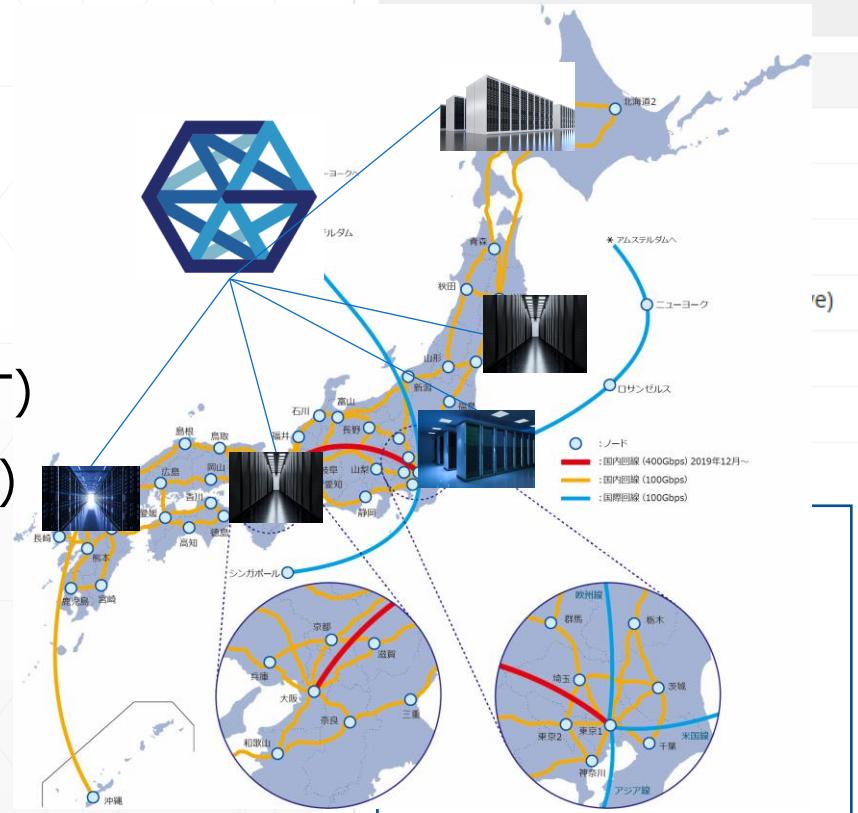
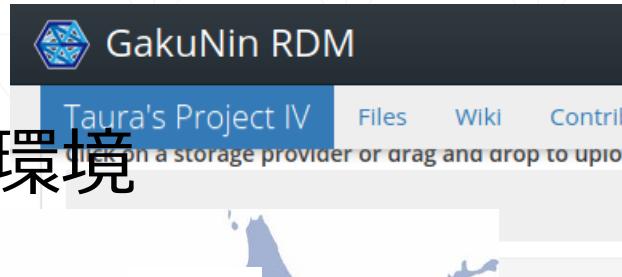


- 大学をまたがった大規模データ処理を  
**SINET 6**経由で高速に、透明にできる環境

- 要開発・研究

- 以下の蓄積

- HPCIシングルサインオン (NII)
- 拠点間高速データ転送**HpFP** (NICT)
- 分散ファイルシステム**Gfarm** (筑波大)



# mdx想定活用シナリオ

## 1. フィールドのデバイスからのデータ収集・処理

- 環境センサー, モバイル機器
- ⇐ 常時稼働, モバイルSINET, 外部通信

## 2. 実験機器からのデータ収集, 実験機器作動

- 電子顕微鏡, 材料合成, 望遠鏡, シーケンサ, etc.
- ⇐ 常時稼働, L2VPN, モバイルSINET, 外部通信

## 3. 研究室環境の拡張

- 研究室ローカルデータの処理, 計算機拡張
- ⇐ L2VPN

# mdx想定活用シナリオ

## 4. 少数研究グループ間でデータ+計算機を共有した連携

- Gakunin RDMやクラウドストレージで論理的な共有(ダウンロード可能にする)は可能だが、大量データの共有にはストレージと一体の計算資源が必要
  - ⇛ Gakunin RDM連携

## 5. 分野用環境の構築

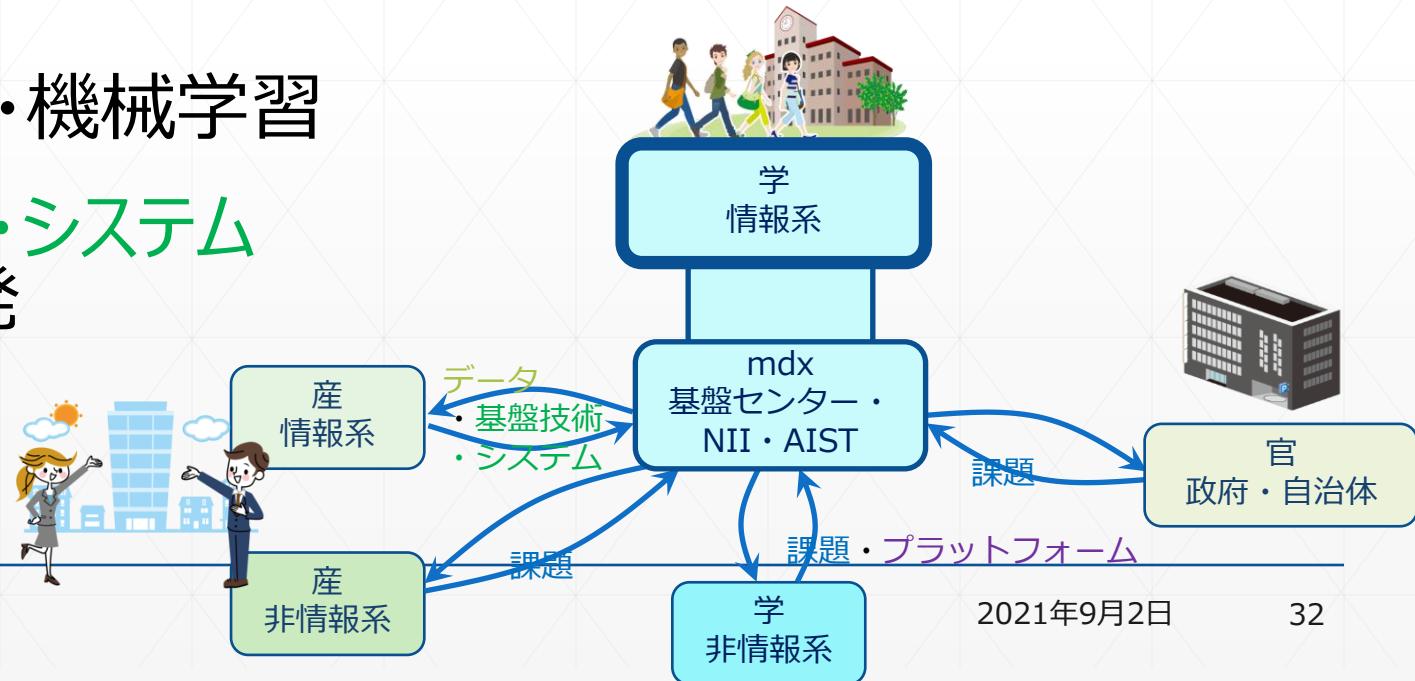
- 分野で必要とされるアプリをそろえた仮想マシンのテンプレート

## 6. 分野データプラットフォームの構築

- アプリ + データをそろえたPaaS環境

# 多様な利用・連携を歓迎・追及

- データ活用課題持ち込み（分野の、地方の、企業の、…）
- 分野データの整備・プラットフォーム構築
- データ提供・サービス（収益）化模索（⇒ 商用クラウドへ展開）
- 高性能AI・機械学習
- 基盤技術・システム研究・開発
- …



# マテリアル先端リサーチインフラ(ハブ): 東大浅野キャンパスに集約

## mdx: データ活用社会創成 プラットフォーム

- ・ 高性能計算機(\*)  
　スパコン性能: 25 PFLOPS
- ・ 大容量データストレージ(\*)  
　26 PB
- ・ 高速ネットワーク  
　SINET 5

(\*スーパーコンピューター:  
Oakforest-PACS)

### 情報基盤センター



代表: 幾原雄一  
スライド提供:  
ナノ工学研究センター

- ・ データ管理・構造化
- ・ データ利活用
- ・ データ人材育成

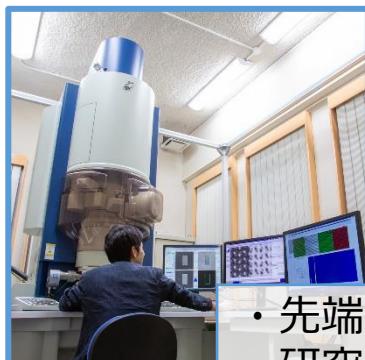
### 情報基盤センター mdx

### ナノ工学研究センター

### 武田先端知 スーパークリーンルーム



### 微細構造解析PF



- ・ 先端ナノ計測
- ・ 研究開発・支援
- ・ データ創出

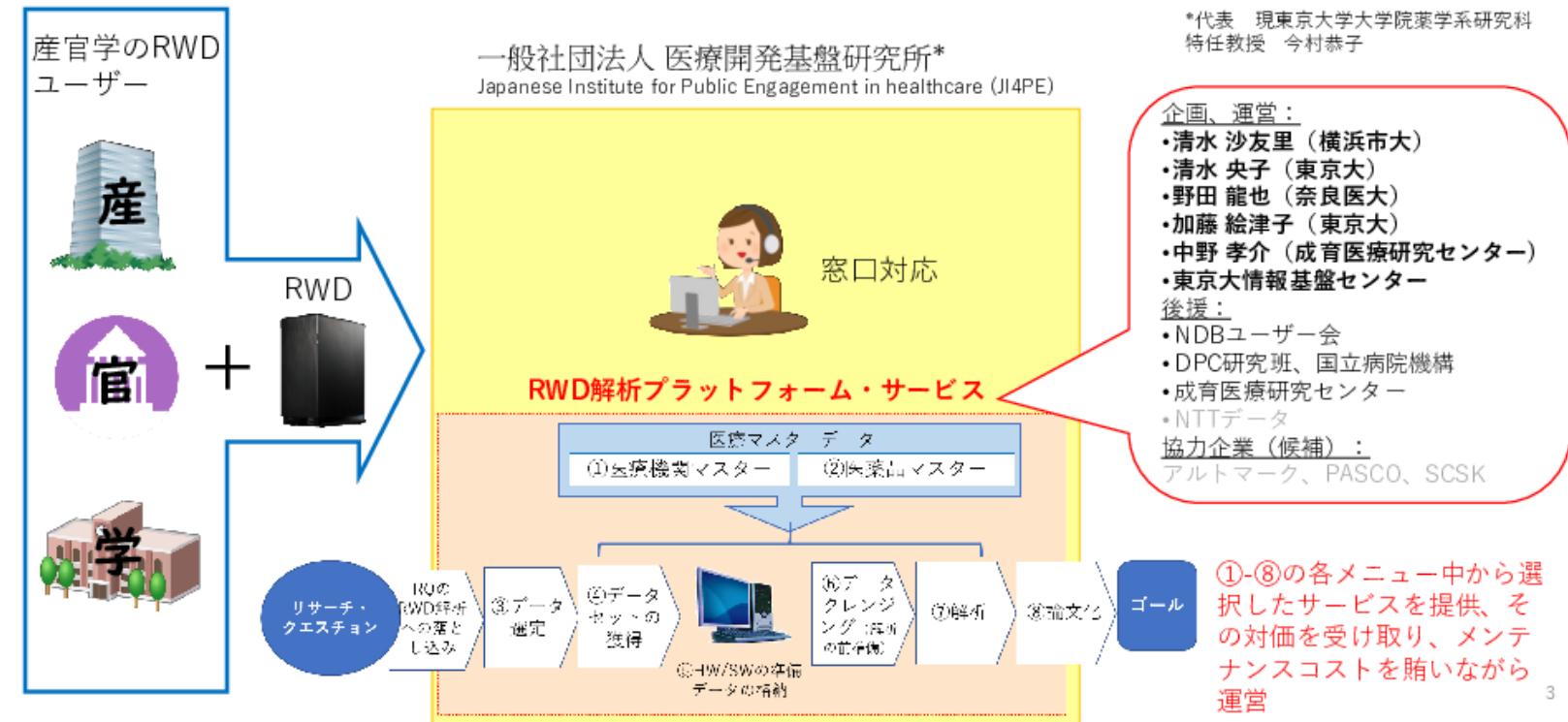
### 微細加工PF



- ・ 先端デバイス加工
- ・ 研究開発・支援
- ・ データ創出

# 医療リアルワールドデータ解析プラットフォーム

産官学のRWDユーザーのもつリサーチクエスチョンに対して、一気通貫した解析に必要なプラットフォームやサービスを提供しようという試み



スライド提供: 東京大学薬学系ITヘルスケアセンター 清水央子先生

JST共創の場(COI-NEXT) 地域気象データと先端科学による戦略的社会共創拠点  
PL: 先端研 中村尚 先生

## 日本域気象再解析: 観測データと数値モデルの融合による大気場再現

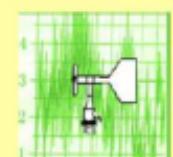
過去の断片的  
各種観測データ



衛星観測データ



高層観測データ



地上観測データ



海洋観測データ

最新の数値予報モデルに取込む「領域再解析」により、①大気状態の全体像を長期にわたり均質に矛盾なく4次元的に再現することで、②時空間的に断片的な観測データに加えて観測不能量も含めた全ての気象要素に関する高時空間分解能データとして社会で広く活用できるよう付加価値を高める。

### 全球大気再解析

最新の全球数値予報モデルと  
データ同化システムによる、  
過去観測データの全球再解析

JRA-55 (2013公開\*)

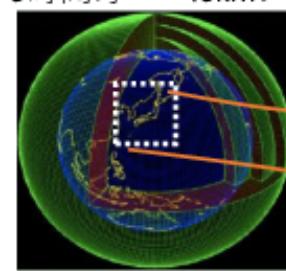
\* 2013以降も継続中

(1958-2012)  
60km メッシュ・6時間毎

JRA-3Q (2021公開予定)

(1947-2021)

40km メッシュ・6時間毎



世界トップクラス

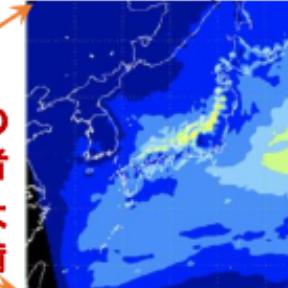
### 日本域気象再解析

最新の領域数値予報モデルと  
データ同化システムによる、  
過去観測データの領域再解析

RRJRA (仮称)

従来にはない過去～現在の  
詳細な4次元地域気象データ

~5km メッシュ (1時間毎?)

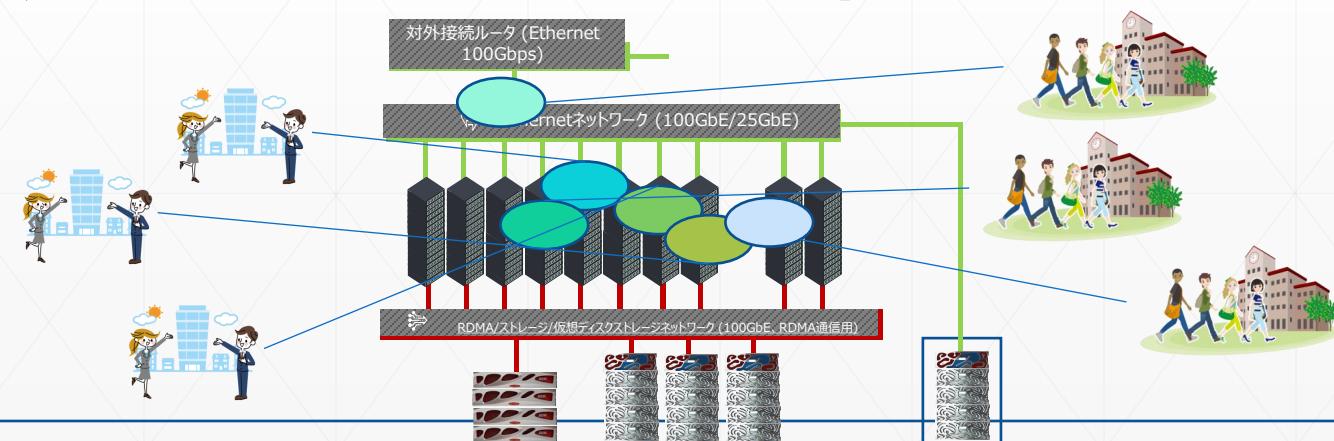


2021年9月2日  
※地表付近は1km メッシュに  
ダウンスケール (15分毎?)

格段に多くの  
潜在的利用者  
だが、日本では  
未整備

# 産業界との連携への期待

- mdxはアカデミアに利用限定するものではない
- 学術・社会への貢献を広くとらえ、それに通じるオープンな（取り組みや知見を共有できる）利用を広く歓迎します
  - そのためにアカデミアが果たすべき役割を果たす「ツール」がmdx
- 最初の敷居を低くし、利用者が交流し、関係する取り組みが出会い、混ざり合っていくことを目指したい



# 大学ベンダ協働 5.0

# mdxは大学とベンダの合作マシン

- これまでの大学の共同利用情報基盤 ≈ スパコン
  - ベンダーに長年の蓄積
- mdx は、それとは大きく異なる環境・異なる運用
  - 仮想化 (VMware) ・プロジェクトごとのネットワーク分離
- 大学にとってもベンダーにとっても初めてのチャレンジ
  - ユーザ環境ごとのVPN設定の自動化
  - ユーザポータル、管理者ポータルの作りこみ

# 大学に必要な意識（と自分が思うもの）

- 大学が「ベンダーの長年の蓄積」だけに頼っては情報基盤はソフトウェアの急激な進化・ユーザの期待値変化に対応できない
- 新しい環境を共に創っていくという意識
  - while (1) {  
ベンダは確実に納品できるものを提案  
大学は確実に納品できることを仕様に書く  
} ループからのbreak
  - 大学はただのお客になってはいけない
    - ベンダとともにリスクをとる
    - ゴールはしっかり据え、やることは臨機応変に



## まとめ

- mdx : 2020年度末もうすぐデビューの高性能仮想化基盤
- 9大学2研究所で離陸
- 使いやすさ、利用を通じた利用者の交流
- ⇒ mdxを通じたセクター、分野を超えたマッチング・共同が行われることを目指す

ご清聴ありがとうございました。

今後のご指導、ご意見、ご利用、ご協力よろしくお願ひします。



## 参考

- データ活用社会創成プラットフォームについて
  - <https://www.itc.u-tokyo.ac.jp/wp-content/uploads/2019/06/45079c24b24f942f050753afec7383c6.pdf>
- データ活用社会創成プラットフォーム
  - <https://www.itc.u-tokyo.ac.jp/wp-content/uploads/2019/02/dd70658fb1895375ac996ec257181671.pdf>

# SINET 5・モバイルSINET

- SINET 5 (有線)
  - 全都道府県にノード
  - ノード間は  $\geq 100\text{Gbps}$
  - 冗長経路
  - 広域VPN (L3/L2)
- モバイルSINET (広域データ収集基盤)
  - SINET VPNに直結できるモバイル環境

