#### 大学の目指す情報基盤の現状とあり方

田浦 健次朗 東京大学 情報基盤センター

次世代計算基盤検討WG 2020.11.30

#### 大学で重要な情報基盤

- 高性能計算機
- 学術情報
- ネットワーク
  - キャンパス内、キャンパス間、遠隔地
- 教育用ICT
  - LMS, オンライン授業, …
- セキュリティ
- データ基盤
- 事務用ICT, 在宅勤務
- 必要な技術や知識の多様化・高度化
- 基盤の大規模化
- 信頼度への要求(基盤への依存度)の増加
- 継続性への要求増加
- 基盤の学内共通化、全国共通化が鉄則の方向性

## 目次

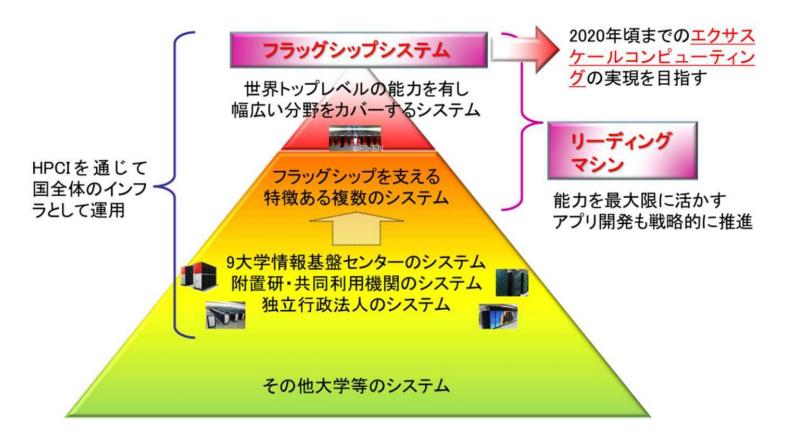
- 情報基盤センター群の高性能計算機
  - -情報基盤センター群とHPCI
  - 情報基盤センター群と共同利用共同研究拠点JHPCN
- ・データ基盤mdx
- それらを踏まえ
  - HPCIの在り方に関する意見
  - 学術情報基盤の高度化に必要な要素に関する意見

#### HPCIと情報基盤センター群

- 次世代計算基盤=HPCIということでは決してないが 少なくとも予算的にHPCI計画がその大きな部分を占 めることは間違いない
- したがって次世代計算基盤と大学とのかかわりを考えるうえで、現在のHPCIと大学とのかかわりを整理する必要がある

#### HPCIにおける大学「いわゆる第2階層」

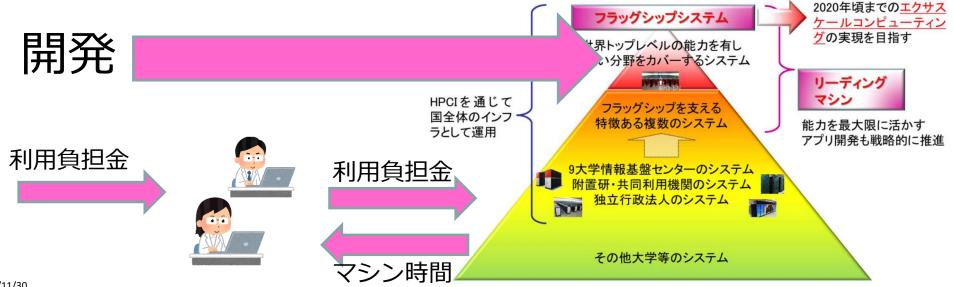
#### く我が国の計算科学技術インフラのイメージ>



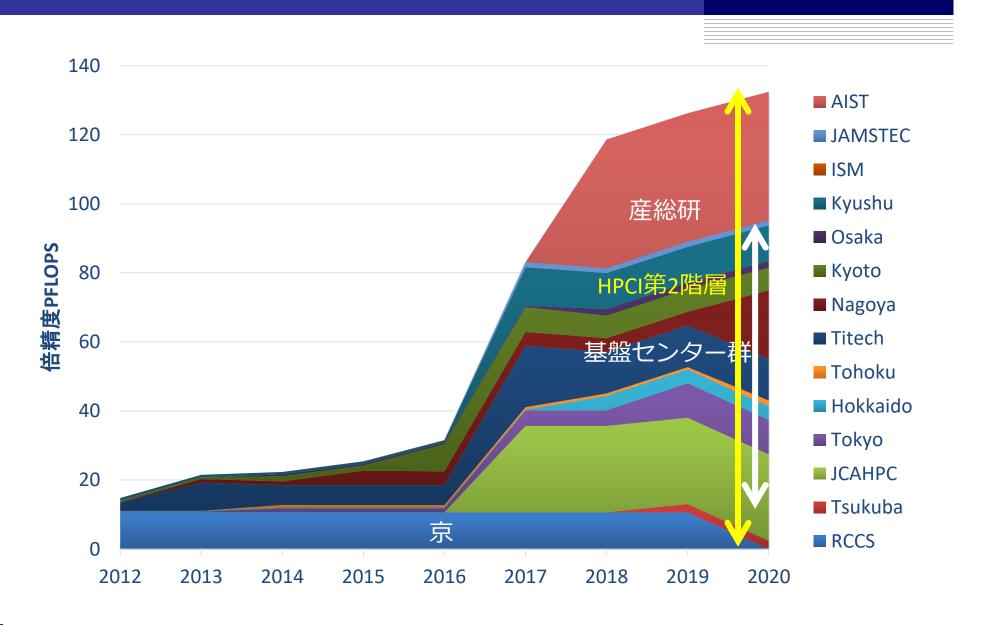
- 資料2-2 今後のHPCIシステムの構築とその利用に関する基本的な考え方について(報告書) (mext.go.jp)
- 第二階層計算資源を提供している各情報基盤セン ター等が、HPCI 計画の推進に対して、より主体的・ 積極的に参画できる体制を構築していくことが、長 期的に見て、我が国の HPCI 計画が強い求心力を維 持しながら推進されるために必須であると考えられ る。そのためには、第二階層計算資源を提供してい る大学情報基盤センター等として HPCI に参画する ことのメリットをより明確に享受できるような HPCI の推進体制を構築していく必要がある

# HPCIの仕組み(I)

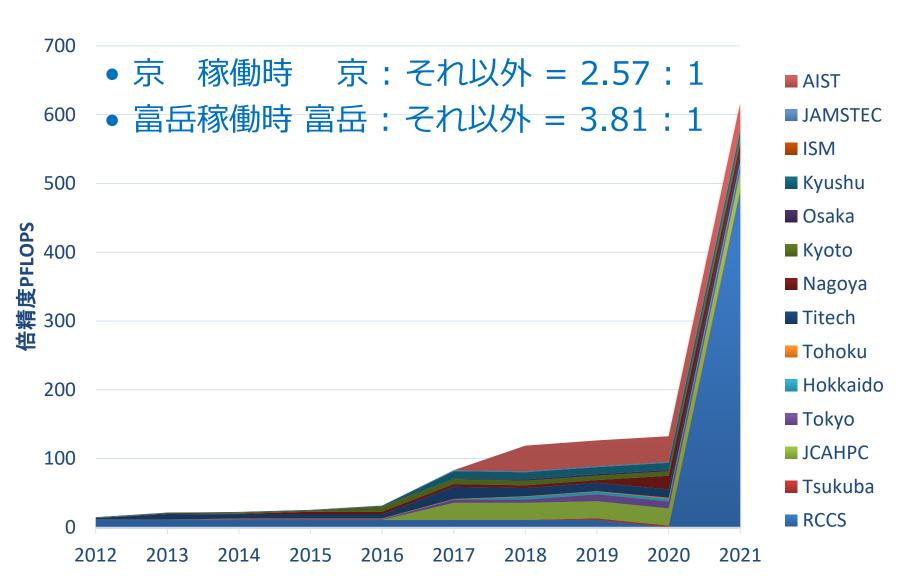
- 第2階層は(HPCIとは無関係に)大学が整備している計算機をHPCIに提供している
- HPCI課題応募者による課題が採択されたらその応募 者の計算機利用負担金が大学に来る



# 倍精度ピークFLOPS (富岳以前)



# 富岳を合わせると(倍精度PFLOPS)



# HPCIの仕組み(II)

- ◆大学は計算機を提供しているだけでなく、多くの人が「連携」のための人的貢献を行っている
  - -連携サービス委員会
  - -連携サービス運営作業部会
  - 共用ストレージ
  - 認証基盤システム
  - 利用者支援システム

## 本当の使命は学際貢献

- 基盤センター群は
  - -連続的に増強する多彩な計算資源(CPU, GPU, ベクトル)で、 日本の計算資源の大きな部分を支えている
  - 京の「端境期」を支えたが容量としてはそれ以前から
- 当然ながら大切なのは量ではなく、利用のされ方、研究・教育・社会への貢献
- 基盤センター群は
  - 学際大規模情報基盤共同利用共同研究拠点(JHPCN)を運営

#### 学際大規模情報基盤共同利用共同研究拠点

● 8大学(北大、東北大、東大、東工大、名大、京大、 阪大、九大)が取り組む「ネットワーク型」拠点

共同研究·萌芽研究課題数 120 100 北海道大学情報基盤センタ・ 40 **JHPCN** 東北大学 サイバーサイエンスセンタ 20 39 H22-1 2010 2015 2016 2018 2019 2020 2011 2012 2013 2014 2017 ■共同研究 ■萌芽研究 **从** 大阪大学 サイバーメディアセンタ・ 名古屋大学 情報基盤セン RIIT 九州大学情報基盤研究開発センタ

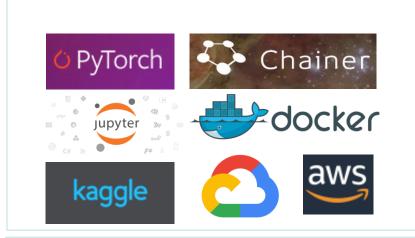
## 目次

- 情報基盤センター群の高性能計算機
  - -情報基盤センター群とHPCI
  - 情報基盤センター群と共同利用共同研究拠点JHPCN
- データ基盤mdx
- 次世代計算基盤へ向けて
  - HPCIの在り方に関する意見
  - 学術情報基盤の高度化に必要な要素に関する意見

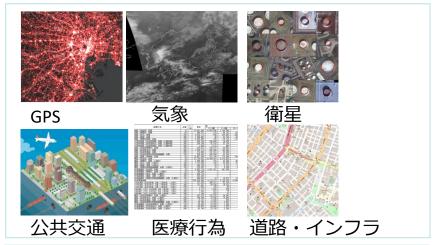
#### mdxとは

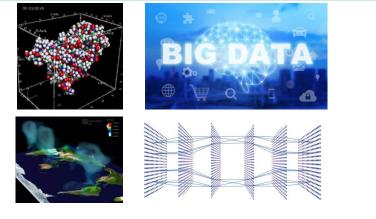


データ科学、データ駆動科学、データ活用、の ために設計された基盤









#### mdxのタイムラインと現在

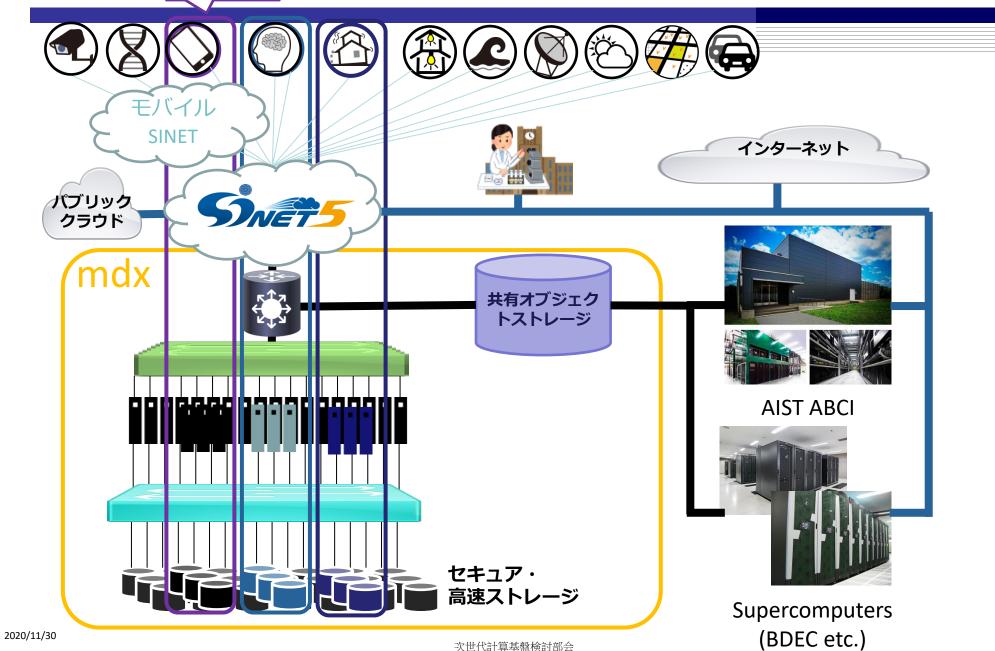
- 2018年度よりNII+東大で構想開始(予算措置以前)
- 予算措置直後からJHPCN全大学へ協力要請、共同で運用を前提に共同で仕様策定
- 2020年度末稼働



#### mdx 特徵

- 2020年度内(末)稼働@東大柏IIキャンパス
- 仮想化された(複数テナント)環境
  - 設計に際しては仮想化基盤、仮想ネットワークなどこれまでのスパコンにない要素の作りこみが不可欠であった
- 計算ノード
  - 350+ CPUノード, 300+ GPUs
- ストレージ
  - 15+PB (含 NVMe SSD 1PB)ファイルシステム Lustre Multitenant拡張
  - 10PB S3オブジェクトストレージ
- 高性能な仮想化されたネットワーク
  - SR-IOV, PVRDMA
  - SINET VPN, モバイルSINETとの連携



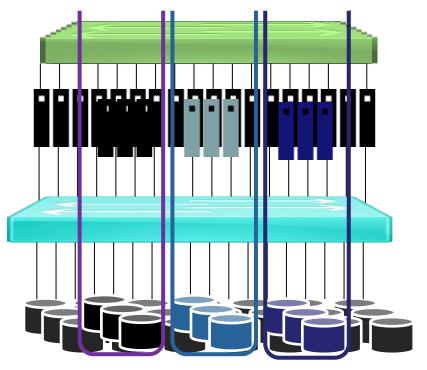


## 仮想プラットフォーム

仮想マシンとVPNを用いて互いに隔離された 「疑似占有環境」

#### 柔軟性

- 各プラットフォームごとに 自由に(管理者権限で)環 境設定可能
- 常時稼働が必要なデータ公 開サービスなどを運用可能
- セキュリティ
  - ひとつの仮想プラット フォームが侵入を受けても 他へ影響しない



#### mdxのハイレベルゴール

- 基盤センターや拠点の分野拡大
  - 黄金公式(HPC+スパコン+計算科学)以外のパターン
  - データ活用、情報学との連携を切望しているのはシミュレーションだけではない
    - ◆ 社会科学(経済)、人文科学(歴史、デジタルアーカイブ)、教育工学、医学、薬学、物性・材料、…
  - すべてのセンターの共通課題?
- 実現のためには基盤センター群とNIIの知恵・ノウハウ結集、エフォートの結合が不可欠、そのための基盤も不可欠

## 目次

- 情報基盤センター群の高性能計算機
  - -情報基盤センター群とHPCI
  - 情報基盤センター群と共同利用共同研究拠点JHPCN
- ・データ基盤mdx
- 次世代計算基盤へ向けて
  - HPCIの在り方に関する意見
  - 学術情報基盤の高度化に必要な要素に関する意見

# 次世代基盤へ向けて(I)

- ●情報学の専門家と様々な分野研究者の学際的な研究 が行われるハブとしての役割が重要
- 今後は両側で分野の拡大が必要
  - 情報: HPC → データ処理、AI
  - 分野:計算科学 → データ駆動科学、データ科学
- 大学はそのために貢献を果たしたいが、HPCIがその ための機会やリソースになっているかは…??? (残念 ながら…)

# 次世代基盤へ向けて(II)

- フラッグシップのような基盤(ハード・ソフト)を 「開発」するならばその興奮を分かち合える仕組みが 必要
  - 多くの若い研究者がお膳立て(FSやWP執筆など)をするが、 プロジェクトスタート後は、お膳立てをした人たちの貢献は ほとんど忘れられている
  - -注:単にFSをfeatureすればいいというわけではない(FSだけして楽しいという、HW・SWの研究者がいるのだろうか)
  - -注:そういう人が開発機関に次から次へと移籍すればいいというわけではない
  - 開発プロジェクト自身をもっとオープンに、組織を超えた貢献ができ、貢献をした個人が報われるような立て付けにする必要があるのではないか

# 次世代基盤へ向けて(III)

- チャレンジングな基盤整備のために…
- フラッグシップ以外でも新規技術の「共同開発」が必要な場面が今後増えてくる
  - 例えばmdxでは仮想化・分離のためにネットワークの詳細設計が必要であり、第一交渉権者決定後に密な議論を行って詳細を決定した
- このようなこと(発注者と受注者の共同開発)はもっと広範に(フラッグシップ以外でも)行えるようになるべき
- ◆制度上の足かせは「既製品しか提案できない」という 入札方式