資料2－1

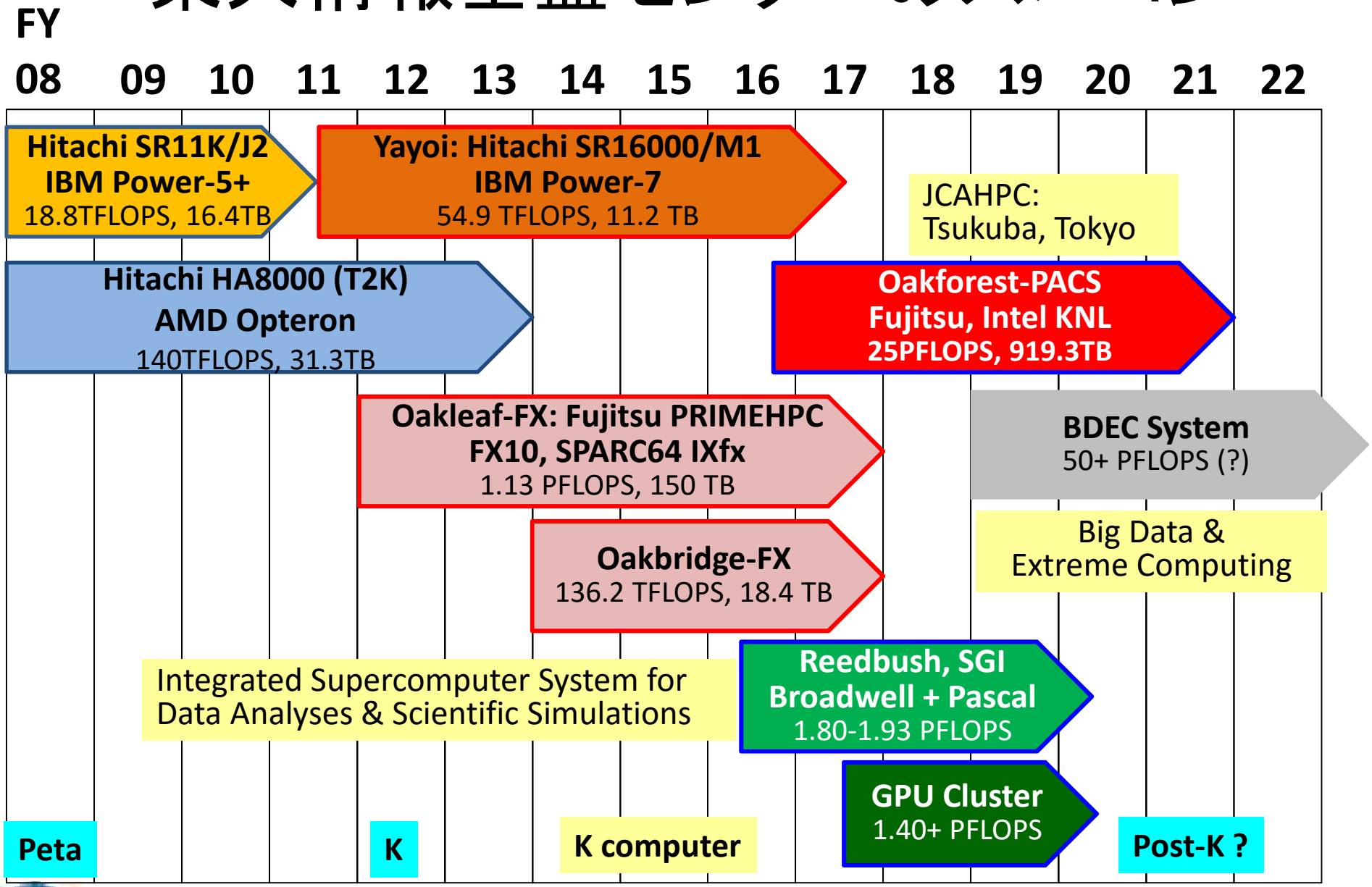
東京大学情報基盤センター
INFORMATION TECHNOLOGY CENTER, THE UNIVERSITY OF TOKYO

東京大学
THE UNIVERSITY OF TOKYO

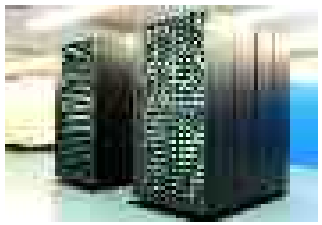
計算科学・データ科学融合へ向けた東大情報基盤センターの取り組み

東京大学　情報基盤センター
中村 宏

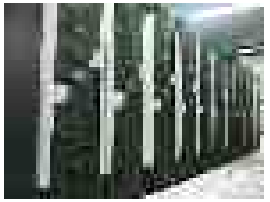
東大情報基盤センターのスパコン



運用中のシステム



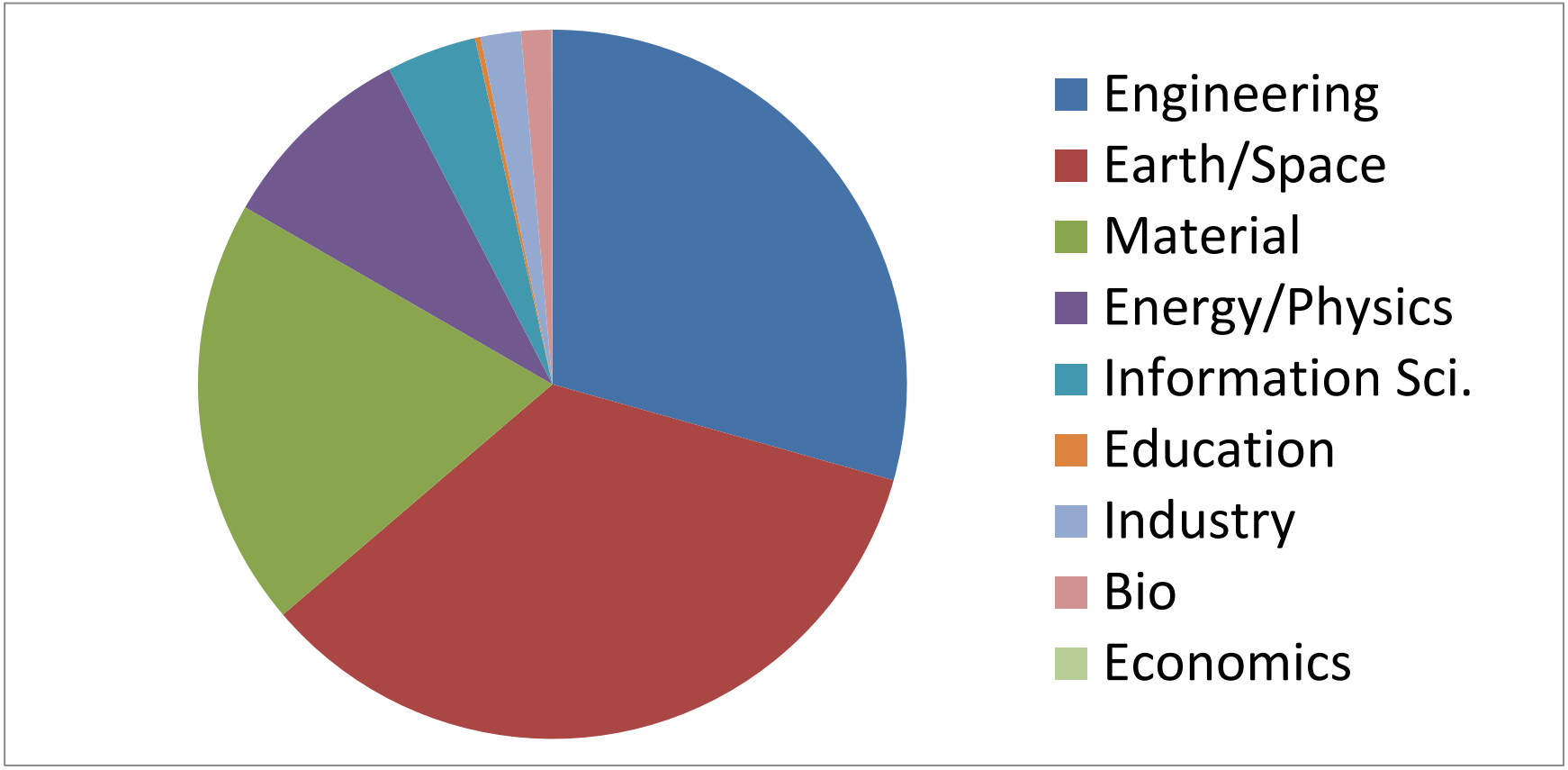

- Yayoi (Hitachi SR16000, IBM Power7)
  - 54.9 TF, Nov. 2011 – Oct. 2017
- Oakleaf-FX (Fujitsu PRIMEHPC FX10)
  - 1.135 PF,「京」商用版, Apr.2012 – Mar.2018
- Oakbridge-FX (Fujitsu PRIMEHPC FX10)
  - 136.2 TF, for long-time use (up to 168 hr), Apr.2014 – Mar.2018
- Reedbush (SGI, Intel BDW + NVIDIA P100 (Pascal))
  - データ解析・シミュレーション融合スーパーコンピュータシステム
  - 1.93 PF, Jul.2016-Jun.2020
  - 東大初のGPU搭載システム
    - 2017年10月に追加システム（4GPU×64ノード以上）を導入



- Oakforest-PACS (OFP) (Fujitsu, Intel Xeon Phi (KNL))
  - JCAHPC (筑波大・東大)共同運用
  - 25 PF, #6 in 48th TOP 500 (Nov.2016) (#1 in Japan)
  - Omni-Path Architecture, DDN IME (Burst Buffer)

CPU時間に基づく利用分野
FX10 in FY.2015 (2015.4~2016.3E)

- 地球科学(大気海洋，地震)，もの作り，物性科学



Oakleaf-FX + Oakbridge-FX

Oakforest-PACS

- 最先端共同HPC 基盤施設　JCAHPC
  - 2013年に東京大学と筑波大学の間で連携推進の協定
  - この協定のもと、2大学の2つのセンターが共同で設置した最先端の大規模高性能計算基盤を構築・運営する組織
    - 東京大学情報基盤センター
    - 筑波大学計算科学研究センター
- Oakforest-PACS:
  最先端共同HPC 基盤施設により調達・運用
  - 予算は2センターの合算
  - 2016/12/1 全系での運用開始
  - 東京大学柏キャンパスの
    東京大学情報基盤センター内に設置

Oakforest-PACSの外観

Oakforest-PACSの特徴

- 計算ノード8208個からなる超並列クラスタ
  - ピーク性能：3TFLOPS×8,208ノード＝ 25 PFLOPS
  - メニーコア 型：Intel Xeon Phi (Knights Landing)
    - 1ノードあたり68コア
    - メモリ（MCDRAM（高速，16GB）＋DDR4（96GB））
- ファイルI/O
  - 並列ファイルシステム:
    Lustre 26PB、総バンド幅500GB/s
  - 高速ファイルキャッシュシステム：
    容量940TB、総バンド幅1560GB/s
    1TB/sec(1000GB/s)を超える実効性能
- TOP500で国内1位（世界6位）

計算ノード

並列ファイル
システム

高速ファイル
キャッシュ
システム

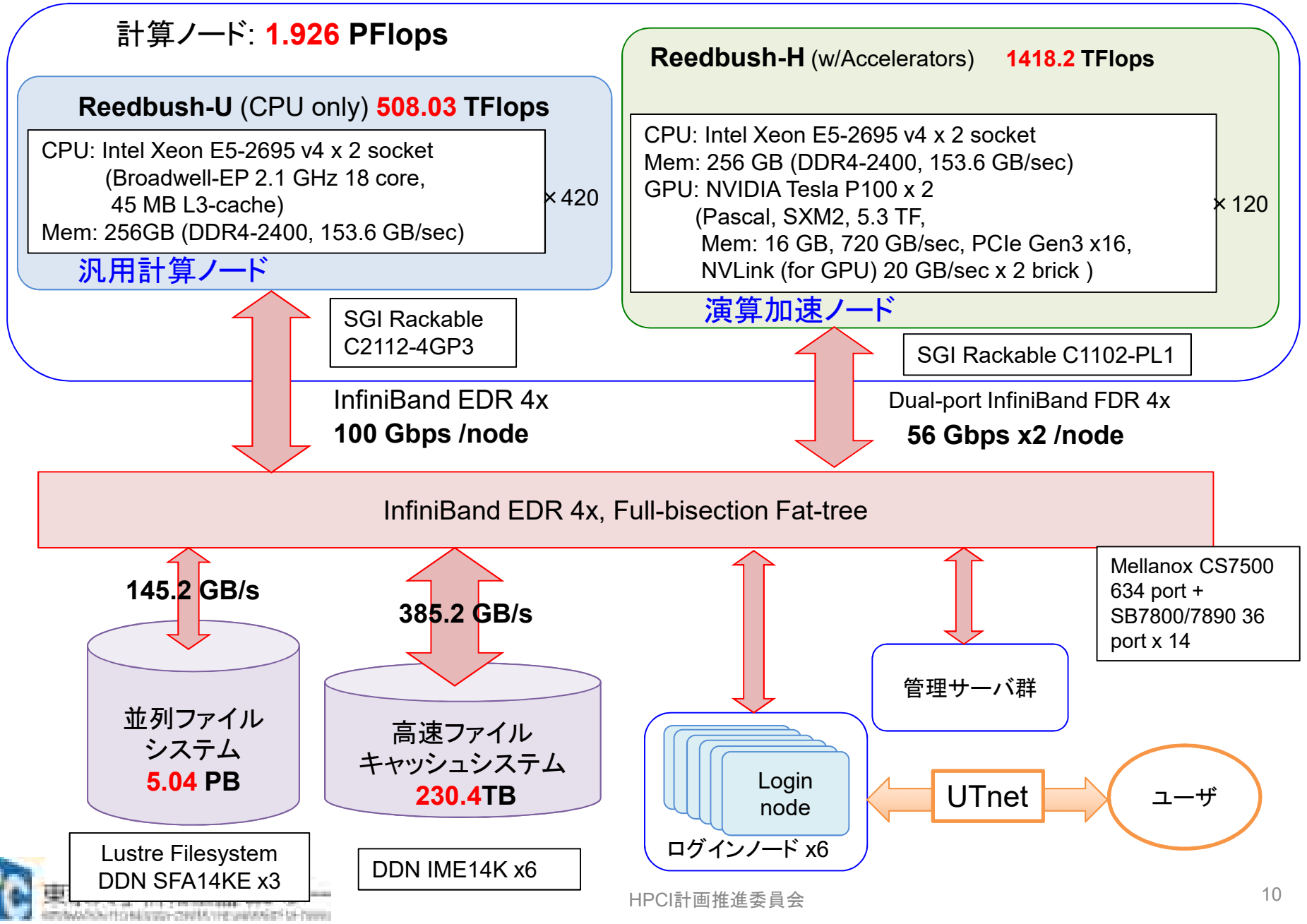
東大情報基盤センターの中長期的展望

- 新しい利用分野の開拓
  - これまでは計算科学分野が中心
    - 「温室効果ガス観測技術衛星「いぶき」（GOSAT）」観測データ処理に適用した実績はある：T2K東大，ノード固定
  - データ科学・深層学習・人工知能
    - ゲノムデータ解析，医用画像処理等での利用は既に開始
- 計算科学・データ科学を融合した新研究手法の創成と活用
  - 当センターの大口ユーザー（大気海洋科学，地震学，ものづくり，物性科学）は，データ同化，境界条件生成，シミュレーション結果評価に観測・実験データを利用する
  - 様々なデータを効率的に活用する計算科学手法開発，環境整備
    - Data Driven Approach
    - 観測データ等を活用したリアルタイムシミュレーション・緊急時対応システム
- BDECシステム（Big Data & Extreme Computing）
  - 2019年4月頃を想定：Reedbushはそのプロトタイプ

Reedbushシステム外観
（データ解析・シミュレーション融合スパコン）

本郷地区（浅野キャンパス）に設置

Reedbushの構成



Reedbushのソフトウェア

- ライブラリ／フレームワーク
  - OpenCV：コンピュータ・ヴィジョン・ライブラリ
  - Theano：Python数値計算ライブラリ
  - ROOT：ビッグデータ向けのライブラリ
  - TensorFlow ：
    Google開発の機械学習向けライブラリ
  - Chainer：Neural Network向けフレームワーク
  - NVIDIA Deep Learning SDK

Reedbushの特徴

- 計算ノード：理論性能 1.926PFLOPS
  - 汎用計算ノード：CPUのみ（Intel Xeon Broadwell-EP)
  - 演算加速ノード：GPUを搭載（NVIDIA Tesla P100）
  - インターコネクト: 100G bps / ノード（Full Bisection Fat-Tree)
- ファイルI/O
  - 並列ファイルシステム (Lustre) : 5.04 PB, 145.2 GB/sec
  - 高速ファイルキャッシュシステム: SSD: 230 TB, 385.2 GB/sec

	演算処理能力	並列ファイル システム	高速ファイル キャッシュシステム
Oakforest-PACS	25PFLOPS	26PB (500GB/s)	940TB (1560GB/s)
Reedbush	1.9PFLOPS	5PB (145GB/s)	230TB (385GB/s)

演算能力は1/10以下だが、ファイルI/O性能は約1/4を維持
データ解析・シミュレーション融合スパコンとしてのReedbush

Reedbushでの試み

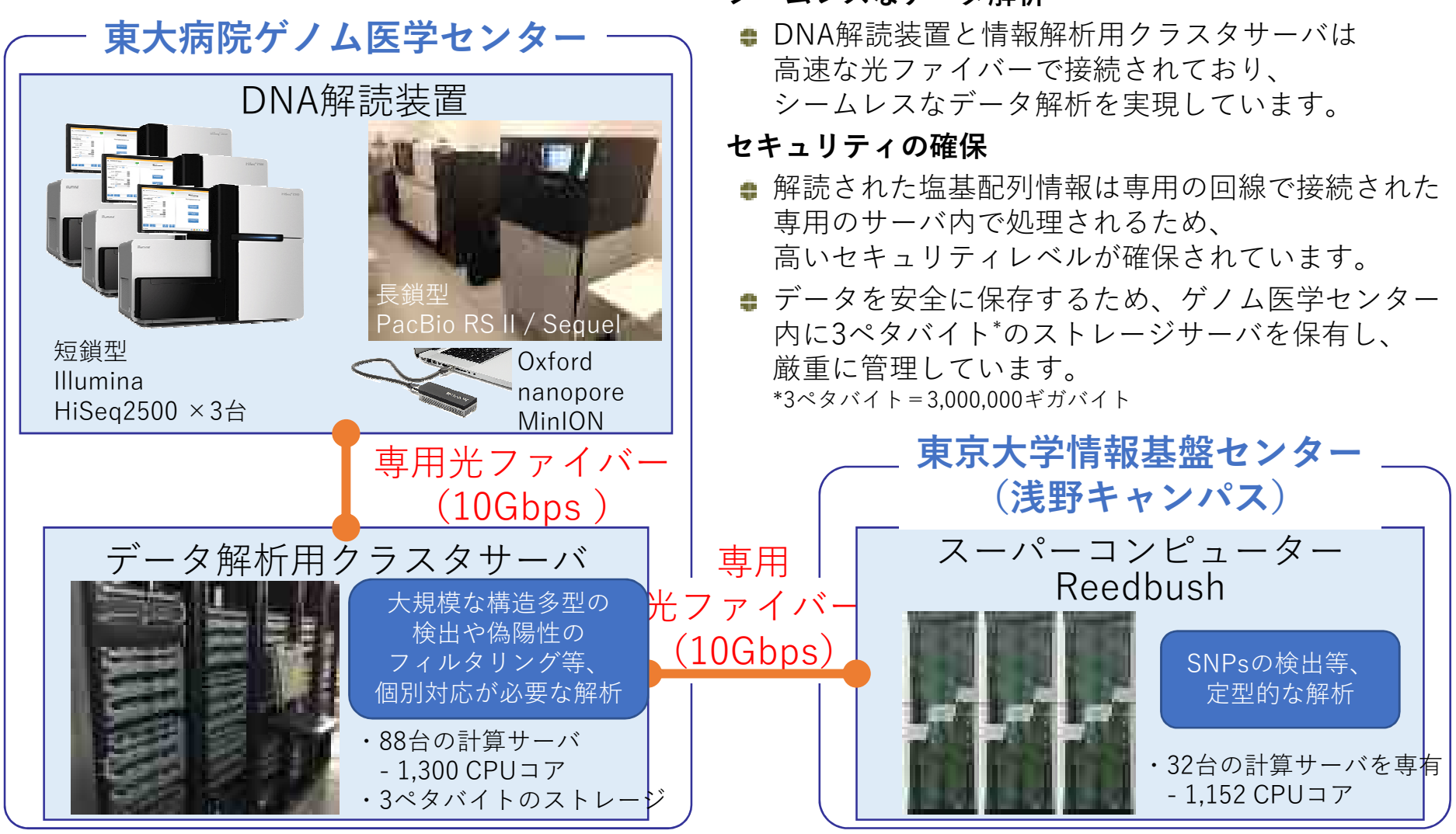
① GPUの導入

- OpenACCの普及：OpenMPと類似したインタフェース
  - CUDAより使い易いが性能悪かった
  - 最近は性能が向上しCUDAと大きな差がなくなった
- データ科学，深層学習（Deep Learning）
  - 新規分野・学際領域の新しいユーザー
    - 旧来のシミュレーションサイエンス → データドリブンサイエンス
  - 東京大学ゲノム医科学研究機構、東京大学医学部附属病院

② ノード固定の導入

- 従来は「ノード×時間」の利用権を購入
- 実行するノードを固定できる利用プログラム
  - データを他利用者から隔離
  - 専用のログインノード＋ストレージをユーザがカスタマイズ可能

東大病院ゲノム医学センターとの連携



シームレスなデータ解析

- DNA解読装置と情報解析用クラスタサーバは高速な光ファイバーで接続されており、シームレスなデータ解析を実現しています。

セキュリティの確保

- 解読された塩基配列情報は専用の回線で接続された専用のサーバ内で処理されるため、高いセキュリティレベルが確保されています。
- データを安全に保存するため、ゲノム医学センター内に3ペタバイト*のストレージサーバを保有し、厳重に管理しています。

*3ペタバイト＝3,000,000ギガバイト

新領域創成科学研究科　森下先生・吉村先生よりご提供のものを改変

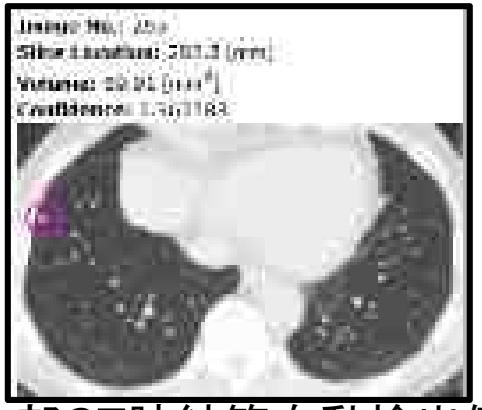
東大病院との連携　～コンピュータ支援検出～

Computer-assisted detection (CAD)

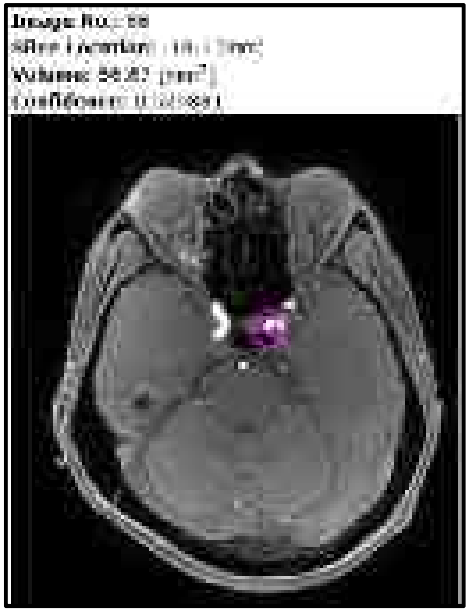
- コンピュータ上で画像解析を行い、自動検出された病変の位置を提示
- 東大病院放射線科及び関連講座における臨床主導型CAD開発
  - 2種類のCAD（胸部CT肺結節検出、頭部MRA脳動脈瘤検出）の開発、臨床使用が実現

  ⇒ 今後対象を増やしていく方針
- 学習と判定（診断）
  - 数千～数万症例（データサイズ：数百GB～数TB）に対応した開発環境の構築が必要

  ⇒ 学習をスパコンReedbushで行う



胸部CT肺結節自動検出例
（○：検出された病変）



頭部MRA肺結節自動検出例
（○：検出された病変）

東大病院　野村行弘先生よりご提供のものを改変

Data Driven Approach

- スーパーコンピュータによる大規模シミュレーション
  - 実アプリケーション＝非線形問題：膨大な計算量（流体，衝突など）
  - 実験・観測データとのValidation，製品設計などのためには膨大なパラメータスタディが必要となり，大量の計算資源が必要
    - 天気予報では既に観測データ同化とシミュレーション（パラメータスタディ）を組み合わせて予測が行われている

Data Driven Approach

1. 機械学習とシミュレーションの融合
   - 詳細シミュレーションモデルの計算結果と機械学習モデルにより，未実施のパラメータによるシミュレーション結果を予測可能な手法
2. 少ないパラメータスタディ数で高精度化が可能
3. 簡略化モデルの構築にも適用可能
   - 詳細計算を予め実施しデータベース，簡易モデルを自動生成
   - リアルタイム観測データ活用シミュレーション，緊急時対策に有効
     - リアルタイム観測データ＋簡易モデルによるシミュレーション

Data Driven Approachの適用

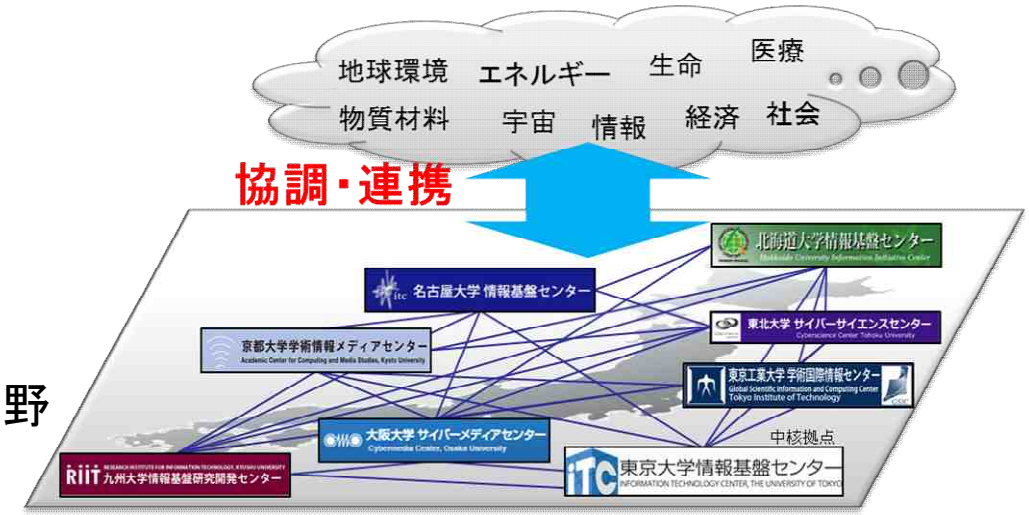
- 地球科学分野
  - 天気予報，長期気候変動予測
  - （衛星観測データ＋詳細シミュレーション）による
    沿海域短期的海況予測 ⇒ 漁業への適用（既に事例あり）
  - 建設・建築・地震防災
    - 長期地殻変動，強震動伝播，避難シミュレーション
    - 三次元国土モデル（国交省 i-Construction）活用
    - インフラの調査・設計・施工・維持・管理・防災
    - 鉄道，道路，物量，損保等の業界で地盤・建物モデル共通化，共有
  - 台風，洪水のリアルタイム警報システム
- 経済・金融関連，交通シミュレーション・渋滞予測
- もの作り：自動車等の最適設計（流体，衝突）

技術的問題点

  - 連続生成される大規模データをスパコンで処理する仕組み
  - セキュリティ（システム，個人データ）

学際大規模情報基盤共同利用・共同研究拠点(JHPCN)の中核として

- 8大学の基盤センター群からなるネットワーク型の拠点
  - 多様な大規模情報基盤と利用技術を活用し、学際研究を発展、産業応用へ展開
- 公募型の学際的共同研究を実施：HPCIシステムの一部も利用可能
- H29年度より、大規模データ・大容量ネットワーク利用課題を奨励
  - 国立情報学研究所の協力によりSINET5が提供する広域・大容量ネットワーク(L2VPNサービスなど)と密に結合可能な資源を用意
  - 広域・大容量ネットワークの利用を前提とした研究を可能に
- 課題数推移 (H26→H27→H28→H29)
  - 超大規模数値計算系応用分野 (29.5→28→33.5→33.5)
  - 超大規模データ処理系応用分野 (0.5→1.5→3.5→3.5)
  - 超大容量ネットワーク技術分野 (1→0.5→0→4.5)
  - 超大規模情報システム関連研究分野 (3→5.5→2→4.5)



8センターの計算資源と人的資源の連携による相乗効果

おわりに

- 東京大学情報基盤センターの取り組み
  - 2016年度に導入した2つのスパコン
    - Oakforest-PACS: 筑波大学と共同調達、国内最高性能
    - Reedbush:データ解析・シミュレーション融合スパコン
  - 最先端スパコンと大容量ストレージの活用により、学際研究を発展、産業応用への展開
    - File I/O 能力の重要性
- 中長期的展望 : 計算科学・データ科学を融合した新研究手法の創成と活用
  - Data Driven Approach
  - 課題
    - 連続生成される大規模データをスパコンで処理する仕組み
    - セキュリティ