東京大学
THE UNIVERSITY OF TOKYO

Oakbridge-CX
利用説明会

東京大学情報基盤センター
スーパーコンピューティング研究部門
https://www.cc.u-tokyo.ac.jp/
問合先：uketsuke@cc.u-tokyo.ac.jp

- 東大情報基盤センターについて
- サービス概要
- スーパーコンピュータシステム概要
- 運用
- 質疑

- 東大情報基盤センターについて
- サービス概要
- スーパーコンピュータシステム概要
- 運用
- 質疑

東京大学情報基盤センター

- 沿革
  - 東京大学大型計算機センター（1965-1999）
  - 東京大学情報基盤センター（1999-）
  - 日本最古の学術大型計算機センター
- 学際大規模情報基盤共同利用・共同研究拠点の中核
  - ８大学の基盤センター群からなるネットワーク型の中核拠点
  - 大規模情報基盤（最先端スパコン、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
    - 分野：計算科学（素粒子物理、宇宙物理、物性、生命科学、地球環境）計算工学（構造物解析、材料・・）、バイオ、経済
  - シミュレーションベース　＋　データサイエンス
- 日本のHPCI（High-Performance Computing Infra)の中核

革新的ハイパフォーマンス・コンピューティングインフラ (HPCI)

文部科学省委託事業

http://www.hpci-office.jp/

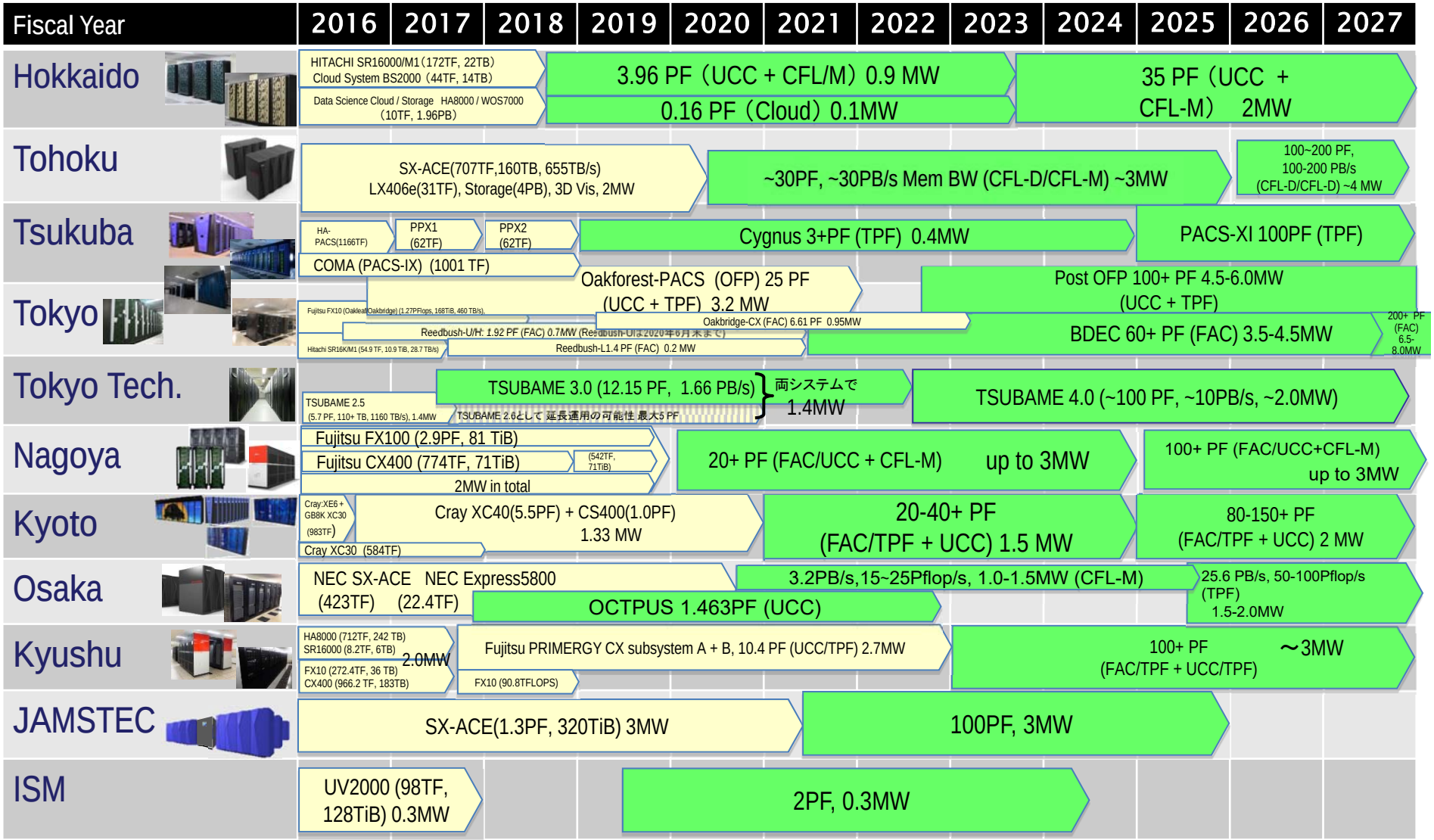
- 使命：我が国における
  - 計算資源（スパコン，大規模ストレージ（東西拠点））
  - 計算科学推進（HPCI戦略プログラム⇒ポスト京重点課題）
- HPCIコンソーシアム（2012～）
  - HPCI計算資源運用
  - 産官学
  - 資源提供者・利用者によるコミュニティ
  - 2012年度発足

情報基盤センター群以外の会員リスト
一般社団法人日本流体力学会
財団法人計算科学振興財団
特定非営利活動法人バイオグリッドセンター関西
自然科学研究機構核融合科学研究所
スーパーコンピューティング技術産業応用協議会
神戸大学
東京大学物性研究所計算物質科学研究センター計算物質科学イニシアティブ（分野2「新物質・エネルギー創成」）
東京大学生産技術研究所（分野4「次世代ものづくり」）
計算基礎科学連携拠点（分野5「物質と宇宙の起源と構造」）
名古屋大学　太陽地球環境研究所
独立行政法人宇宙航空研究開発機構宇宙科学研究所
独立行政法人海洋研究開発機構
一般社団法人日本計算工学会
計算生命科学ネットワーク

国立研究開発法人理化学研究所計算科学研究機構
高エネルギー加速器研究機構 共通基盤研究施設・計算科学センター
情報・システム研究機構　国立情報学研究所
一般財団法人高度情報科学技術研究機構
筑波大学　計算科学研究センター
大阪大学　核物理研究センター
国立研究開発法人産業技術総合研究所 情報技術研究部門
東京大学　物性研究所
東北大学　金属材料研究所
情報・システム研究機構　統計数理研究所
自然科学研究機構分子科学研究所　計算科学研究センター
独立行政法人宇宙航空研究開発機構　情報計算工学センター

HPCI第2階層システム 運用＆整備計画 (2018年11月時点)

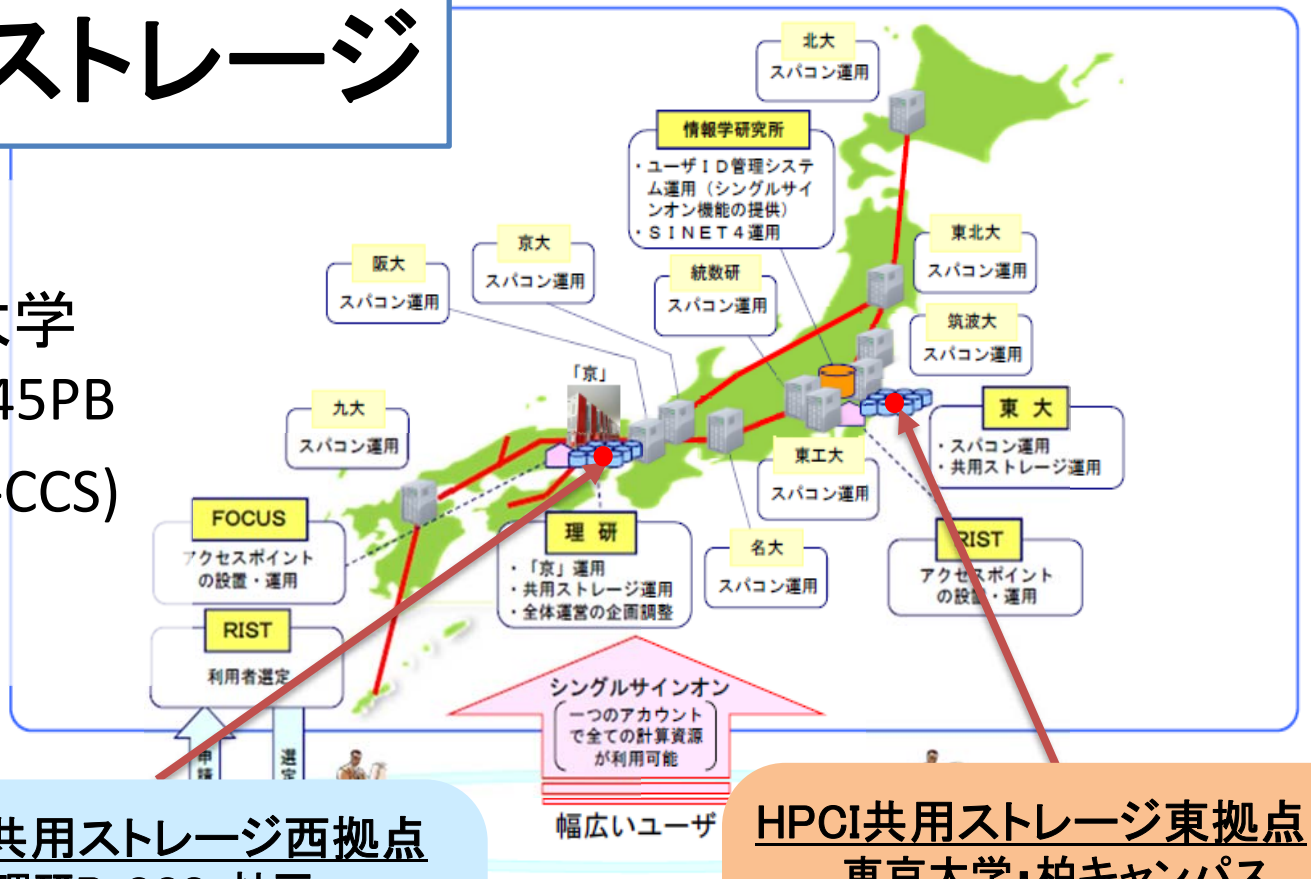
東大のみ2019年6月末時点



電力は最大供給量（空調システム含む）

HPCI共用ストレージ

- 文科省委託費
- 東拠点（東京大学 柏キャンパス）45PB
- 西拠点（理研R-CCS) 45PB



ストレージ

HPCI共用ストレージ西拠点

理研R-CCS・神戸

・データストレージ（総容量 42 PB）
・メタデータサーバ 2 台

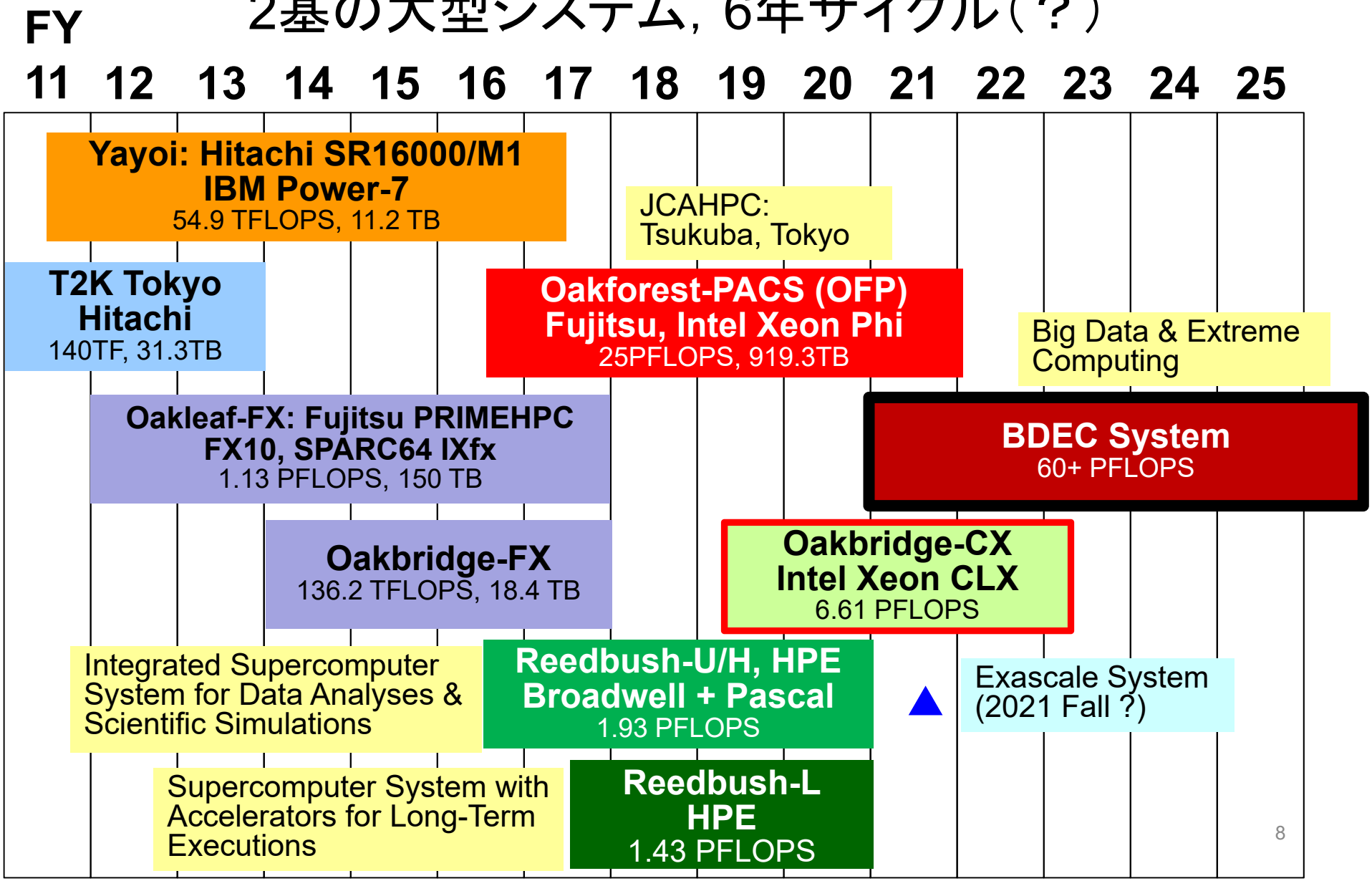
HPCI共用ストレージ東拠点

東京大学・柏キャンパス

・データストレージ（総容量 45 PB）
・メタデータサーバ 2 台
・大容量メモリサーバ、GPUサーバ等

東大センターのスパコン

2基の大型システム，6年サイクル（？）



3システム：利用者2,000+，学外50+%

- 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
  - 3.36 PF, 2016年7月～ 2021年3月末（予定）
  - 東大ITC初のGPUシステム (2017年3月より), DDN IME (Burst Buffer)

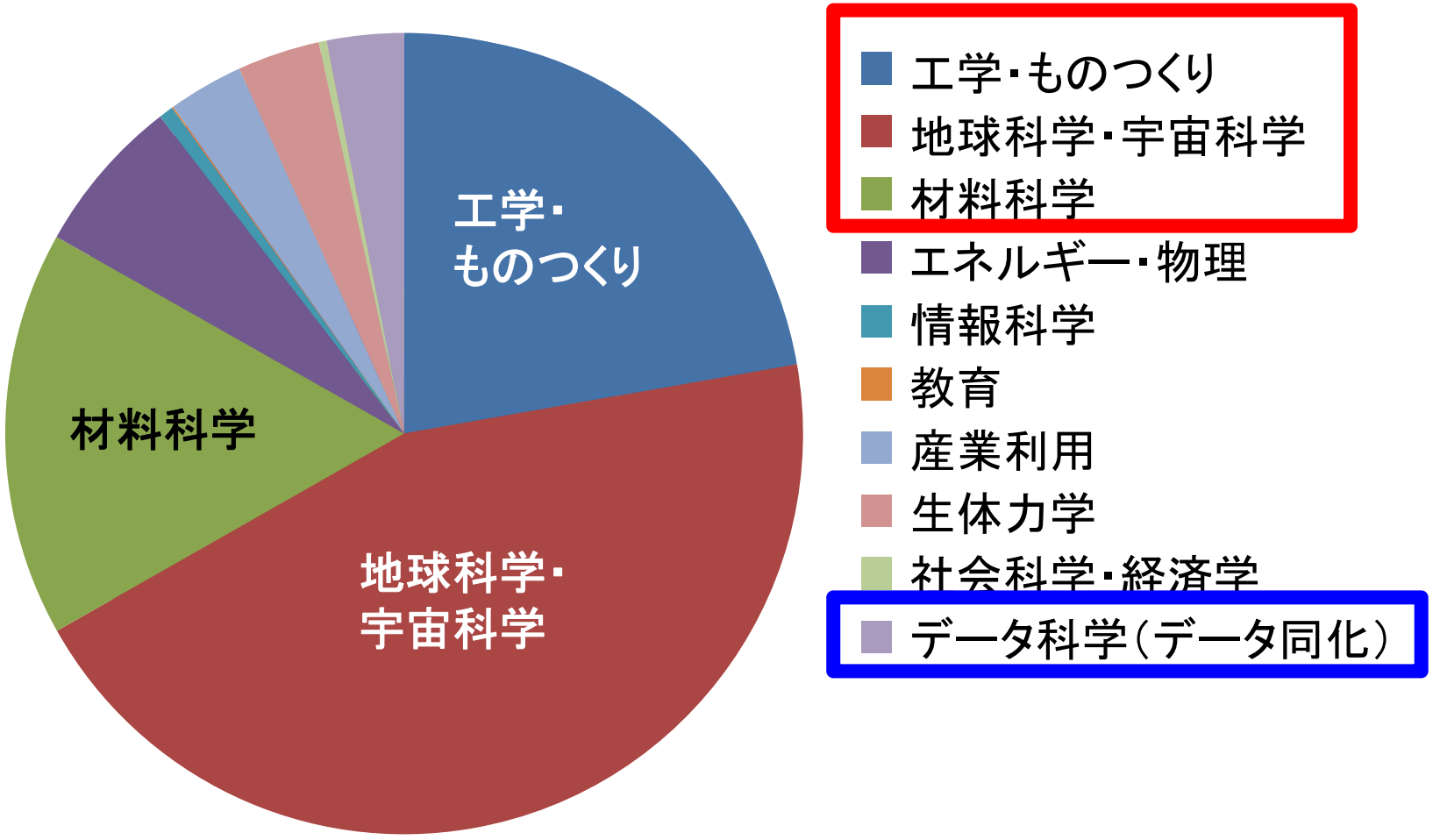
- Oakforest-PACS (OFP) (富士通, Intel Xeon Phi (KNL))
  - JCAHPC (筑波大CCS＆東大ITC)
  - 25 PF, TOP 500で6位 (2016年11月) (日本1位)（初登場時）
  - Omni-Path アーキテクチャ, DDN IME (Burst Buffer)

- Oakbridge-CX (富士通, Intel Xeon Platinum 8280)
  - 大規模超並列スーパーコンピュータシステム
  - 6.61 PF, 2019年7月 ～ 2023年6月
  - 全1,368ノードの内128ノードにSSDを搭載

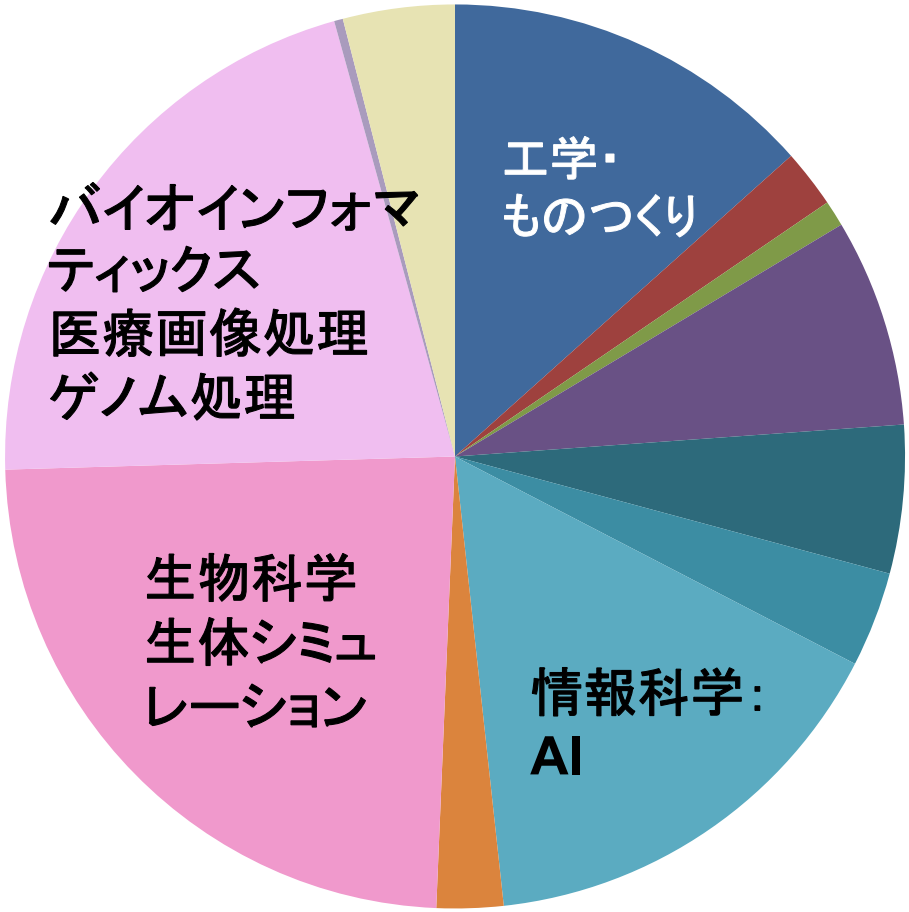
GFLOPS（ピーク性能換算）あたり負担金（～W）

System	JPY/GFLOPS
Reedbush-U (HPE) (Intel BDW)	61.9
Reedbush-H (HPE) (Intel BDW+NVIDIA P100x2/node)	15.9
Reedbush-L (HPE) (Intel BDW+NVIDIA P100x4/node)	13.4
Oakforest-PACS (Fujitsu) (Intel Xeon Phi/Knights Landing)	16.5
Oakbridge-CX (Fujitsu) (Intel Cascade Lake (CLX))	20.7

研究分野別利用CPU時間割合（Oakleaf-FX+Oakbridge-FX）（「京」商用版） 2017年度

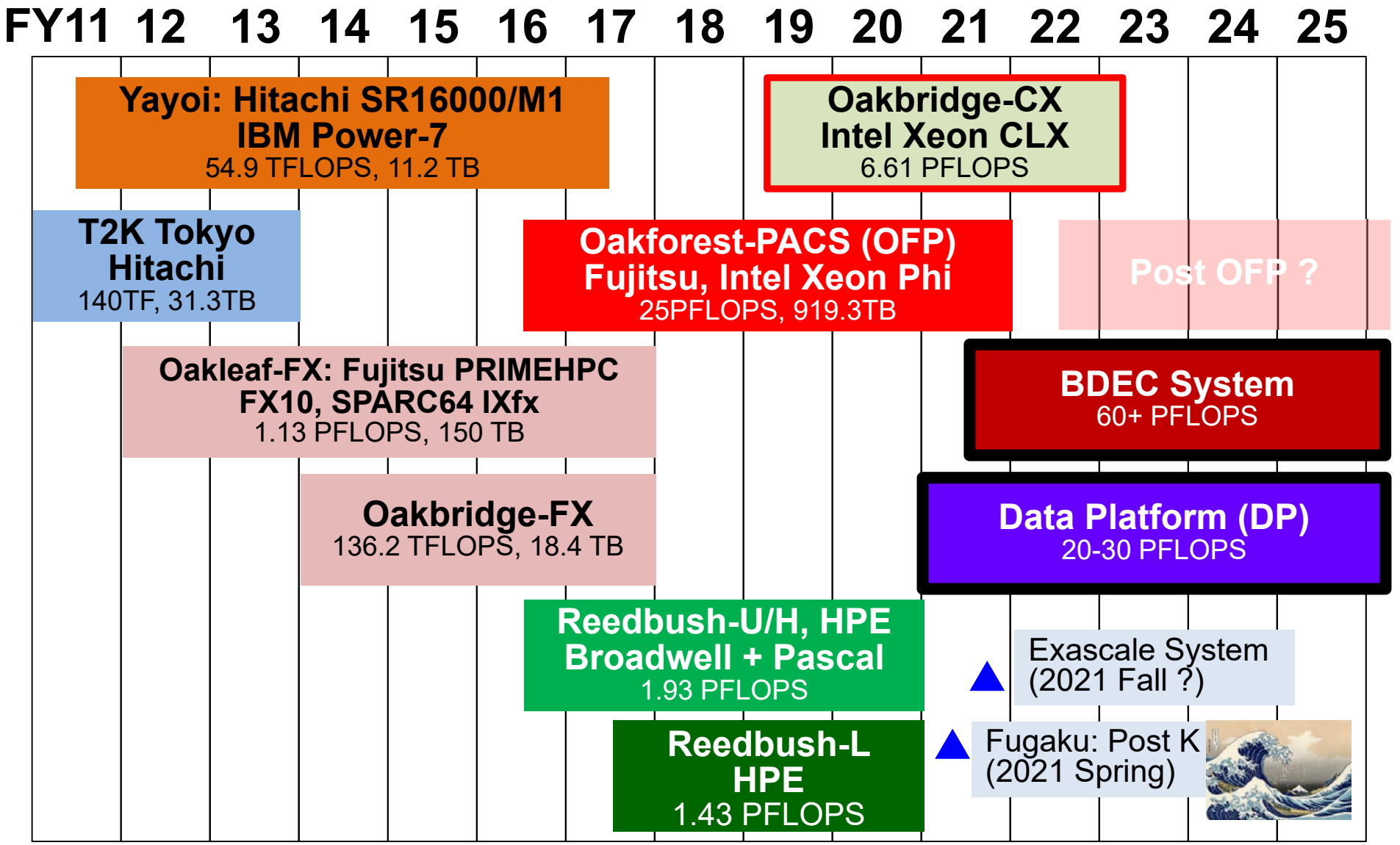


研究分野別利用CPU時間割合（Reedbush-H，2GPU/ノード）2018年度



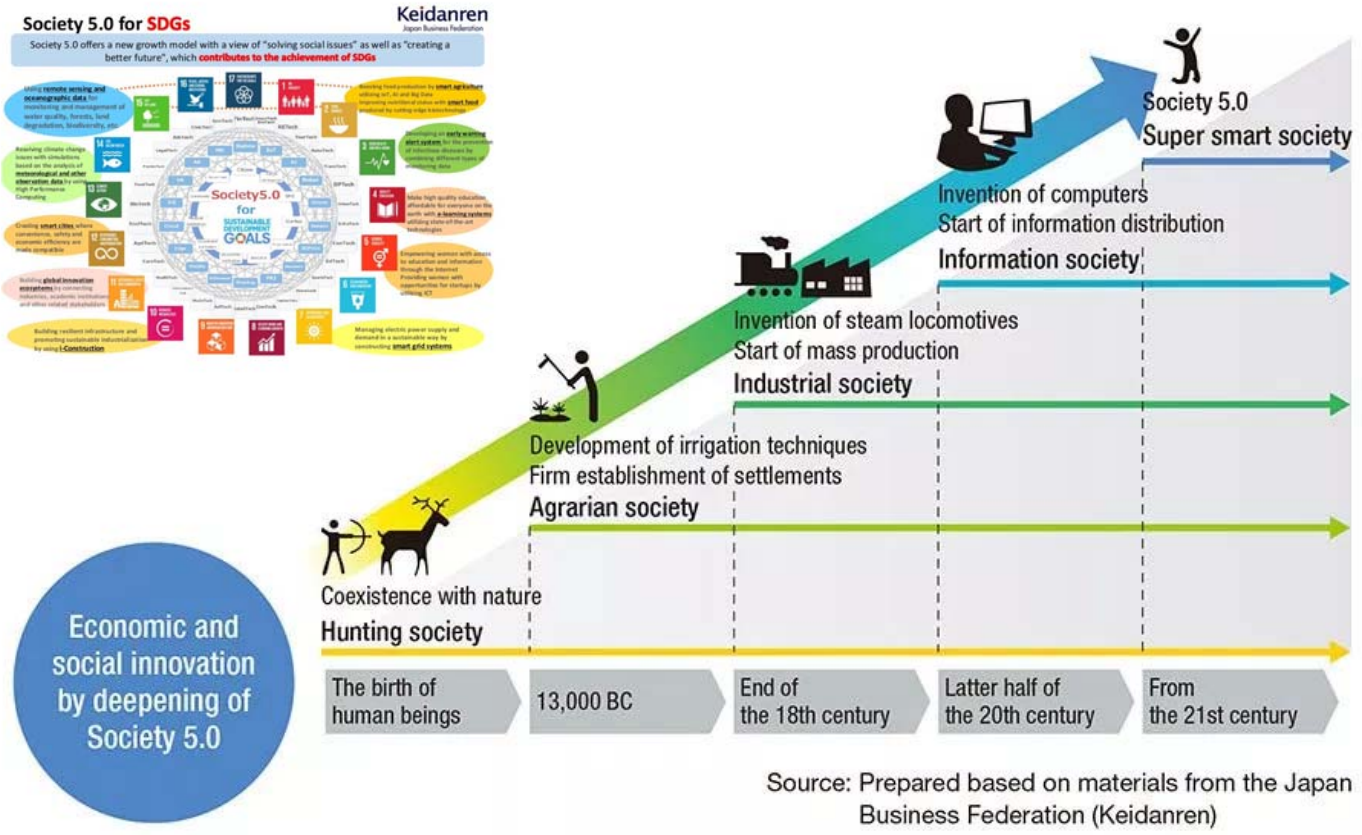
- 工学・ものつくり
- 地球科学・宇宙科学
- 材料科学
- エネルギー・物理
- 情報科学：システム
- 情報科学：アルゴリズム
- 情報科学：AI
- 教育
- 産業利用
- 生物科学
- バイオインフォマティックス
- 社会科学・経済学
- データ科学（データ同化）

東大センターのスパコン



Society 5.0：日本が提唱する未来社会のコンセプト

- デジタル革新・イノベーション（IoT，AI，ビッグデータ等）により知識集約型・超スマートな社会への変革を目指す



5.0: Super Smart

4.0：情報

3.0：工業

2.0：農耕

1.0：狩猟

Source: Prepared based on materials from the Japan Business Federation (Keidanren)

新タイプの利用者

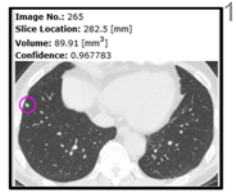
- 計算科学・工学中心
- データ科学，機械学習，AI
  - ゲノム解析
  - 医用画像処理

Society 5.0実現へ向けた新手法

- （シミュレーション：S＋データ：D＋学習：L）融合

BDEC (Big Data & Extreme Computing)

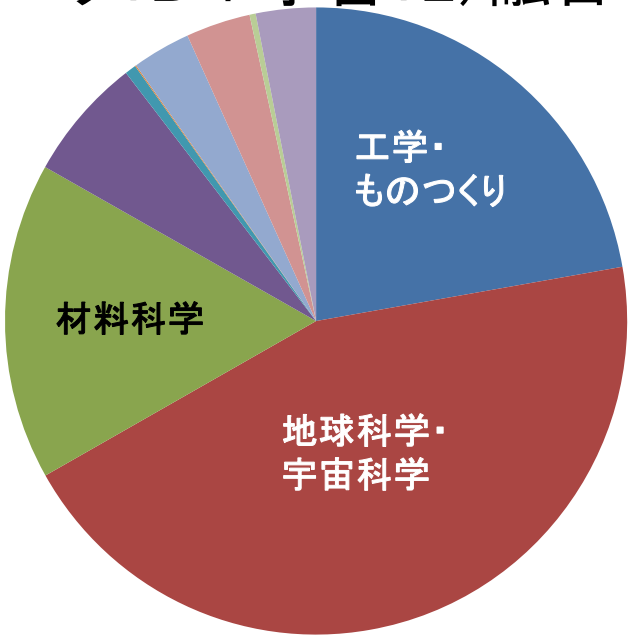
- 60＋PFピーク性能，2021年10月以降運用開始
- Society 5.0へ向けた（S+D+L）プラットフォーム
- Reedbush・Oakbridge-CXはBDECのプロトタイプ



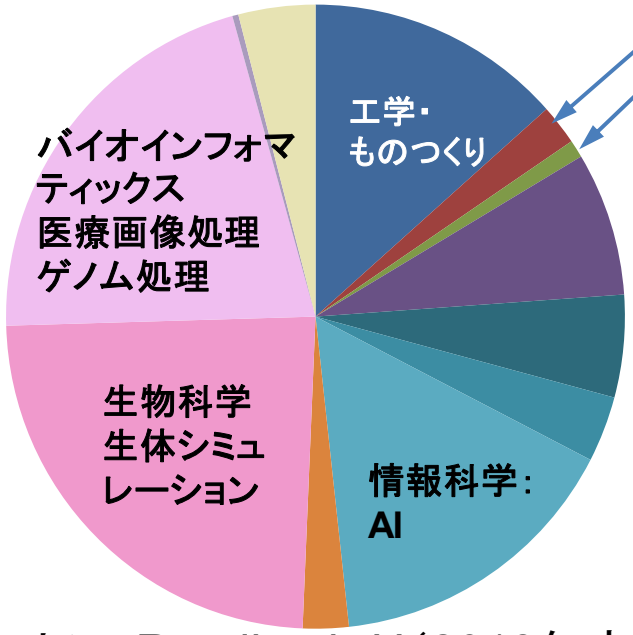
Lung Nodule by Chest CT
(○:detected lesion)

Lung Nodule by Head CT
(○:detected lesion)

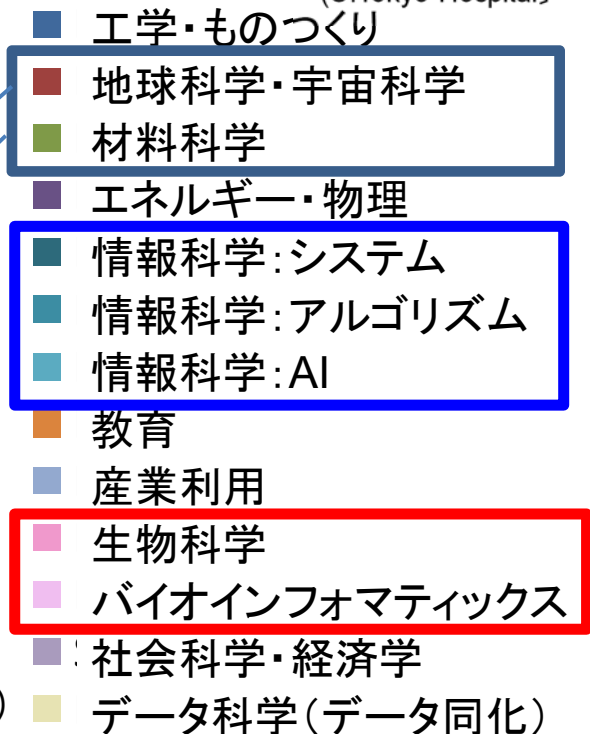
〔c/o Dr. Y. Nomura (U.Tokyo Hospital〕



Oakleaf/Oakbridge-FX（2017年度）
「京」商用版

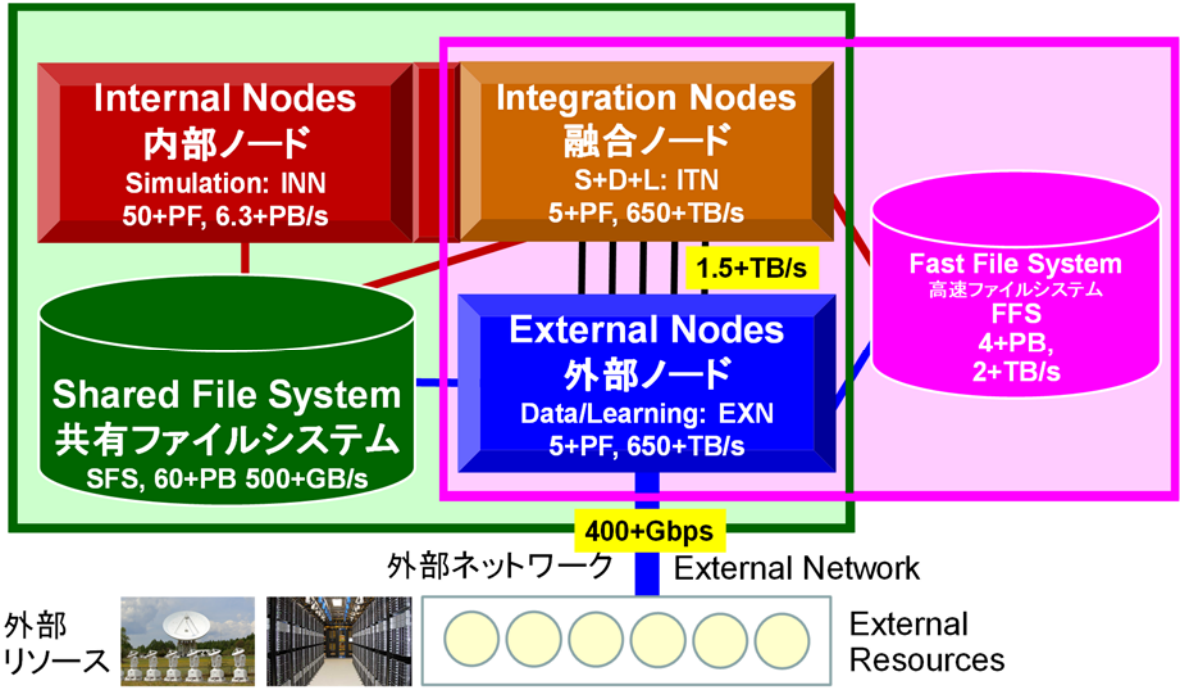


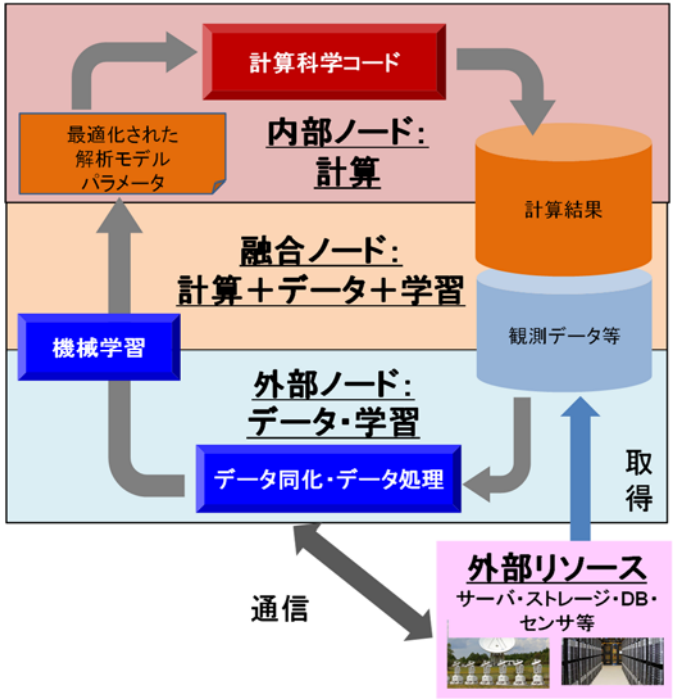
Reedbush-H（2018年度）
Intel BDW + NVIDIA P100



BDEC System (1/4): 2021年10月以降

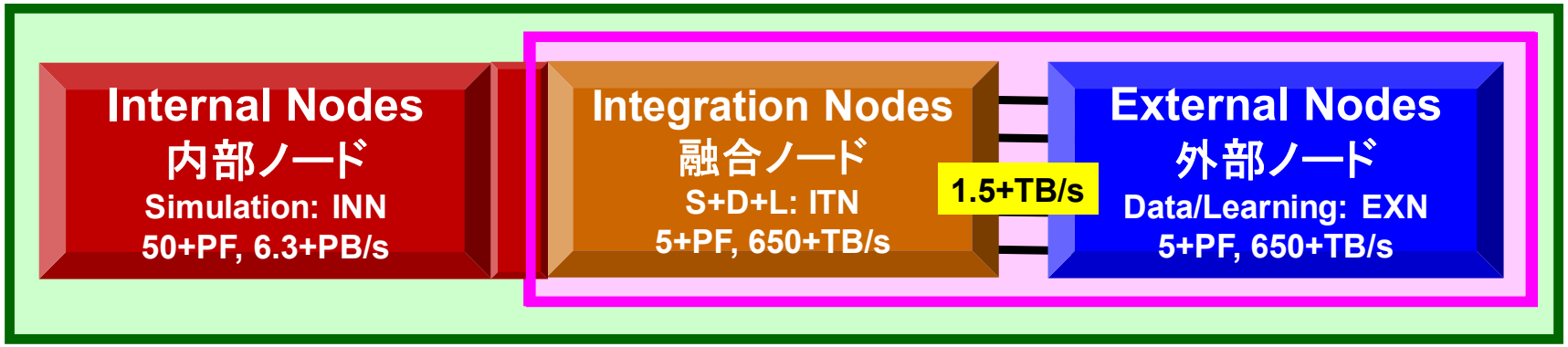
- 60+ PF, 3.5-4.5 MW（冷却込み）, ~360 m^2
  - 外部データ（External Nodes, EXN）：データ取得・処理，学習
  - 内部ノード（Internal Nodes, INN）：計算，50+PF, 6.3+PB/s
  - 融合ノード（Integration Node，ITN）：（S+D+L）融合
    - EXNのアーキテクチャは（INN+ITN）と異なっても良い
  - 共有ファイルシステム (SFS) + 高速ファイルシステム (FFS)





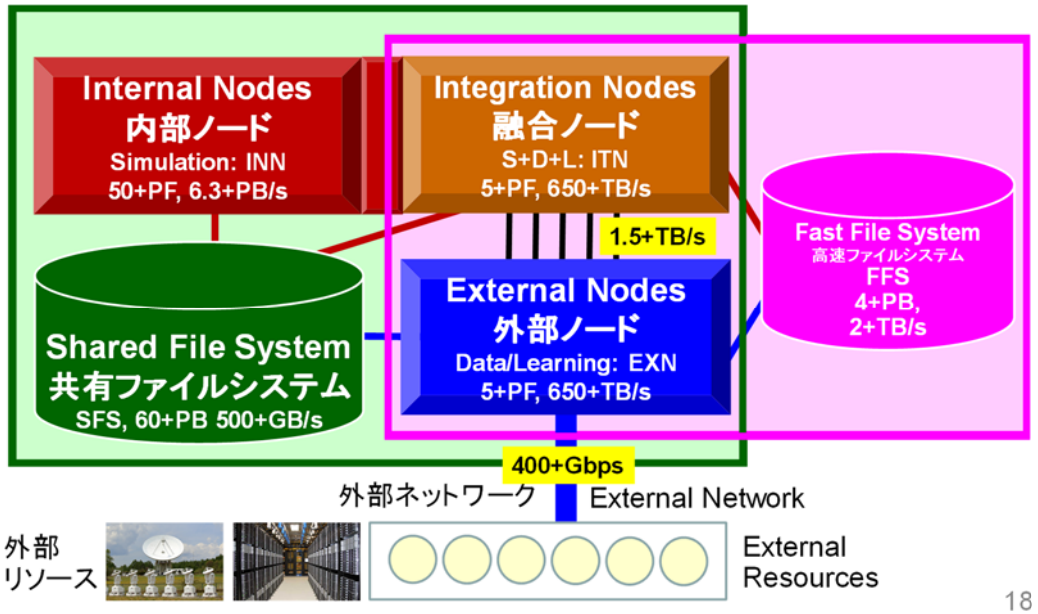
BDEC System (2/4): 合計60+PF

- 内部ノード：Internal Nodes (INN)
  - 50+PF, 6.3+PB/s
  - 従来のスパコンと同様の役割，シミュレーション・計算
- 融合ノード：Integration Nodes (ITN) ： (S+D+L)融合
  - 5+PF, 650+TB/s
  - 内部ノード（INN）と同じアーキテクチャ，一体
- 外部ノード：External Nodes (EXN)
  - 5+PF, 650+TB/s
  - 外部リソースに直接接続：400+Gbps
  - （ITN+INN）とは異なったアーキテクチャでも可



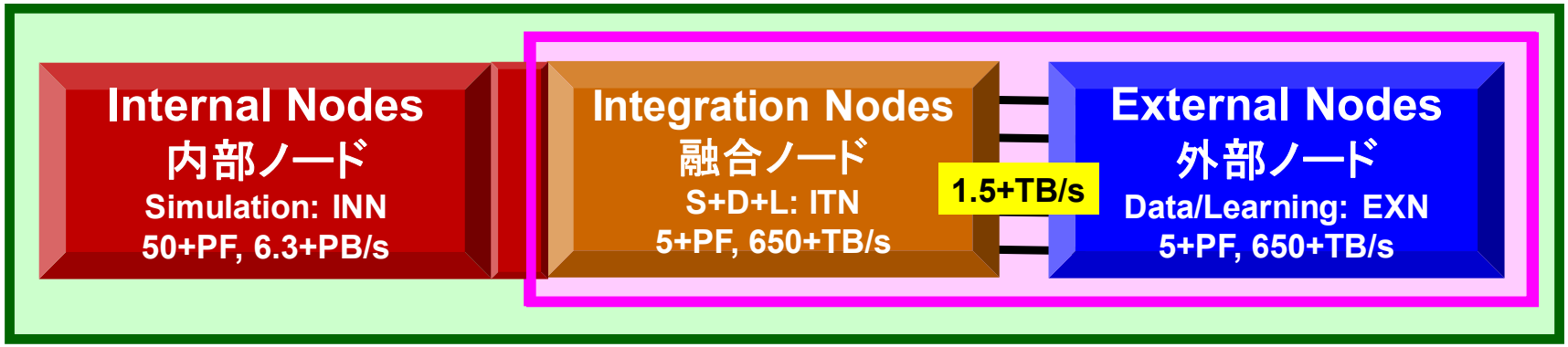
BDEC System (3/4)

- INN+ITN: 15+TB/s (バイセクションバンド幅)
  - 単一のスパコンシステム
  - INNとITNをまたいで単一のMPIジョブを実行できる：必須要件
- (INN+ITN)-EXN: 1.5+TB/s (バイセクションバンド幅)
  - (INN+ITN+EXN)をまたいで単一ジョブを実行できる：必須要件
  - 単一MPIジョブであることは必須ではないが加点
- 高速ファイルシステム：Fast File System: FFS: 4+PB, 2+TB/s
  - EXN・ITNにより共有
  - SSD可能
- 共有ファイルシステム：Shared File System: 60+PB, 500+GB/s
  - EXN・ITN・INNにより共有
  - 現存・将来の他システムからも利用可能



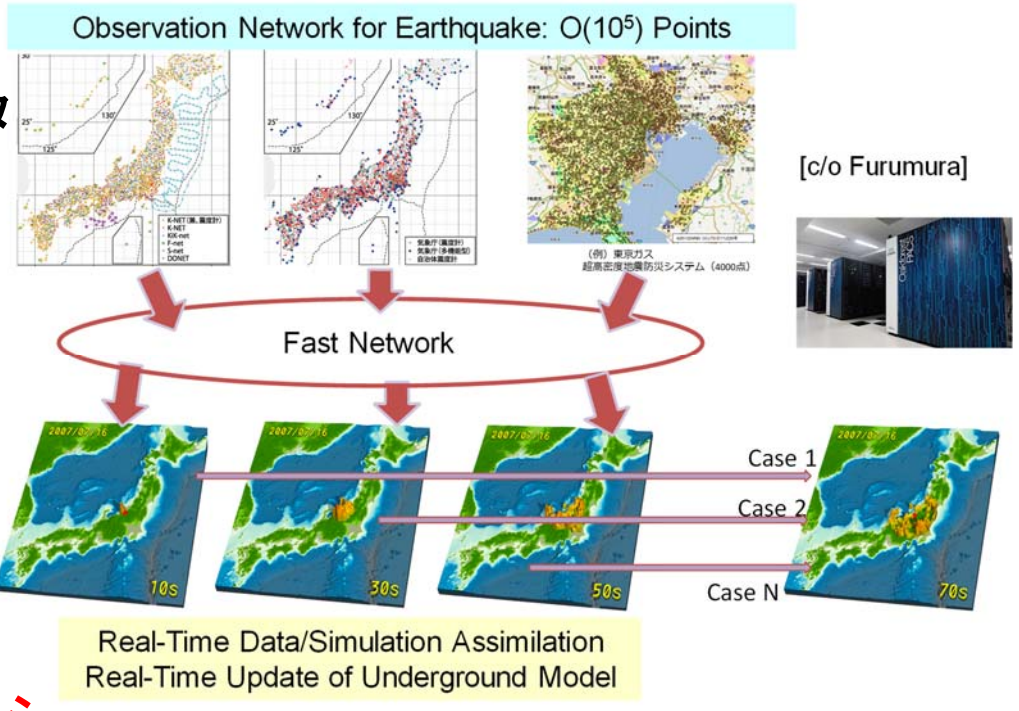
BDEC System (4/4): 想定アプリ

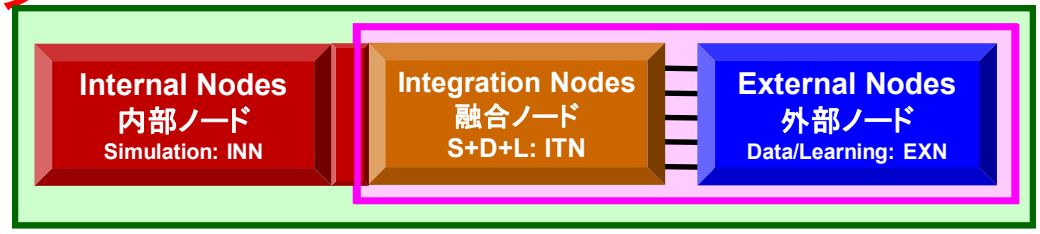
- 大規模シミュレーション＋データ同化
- 大気海洋連成シミュレーション＋データ同化
  - 東大・大気海洋研，理研
- 地震シミュレーション＋データ同化
  - 東大・地震研（次頁）
- リアルタイム災害シミュレーション
  - 洪水，津波，地震



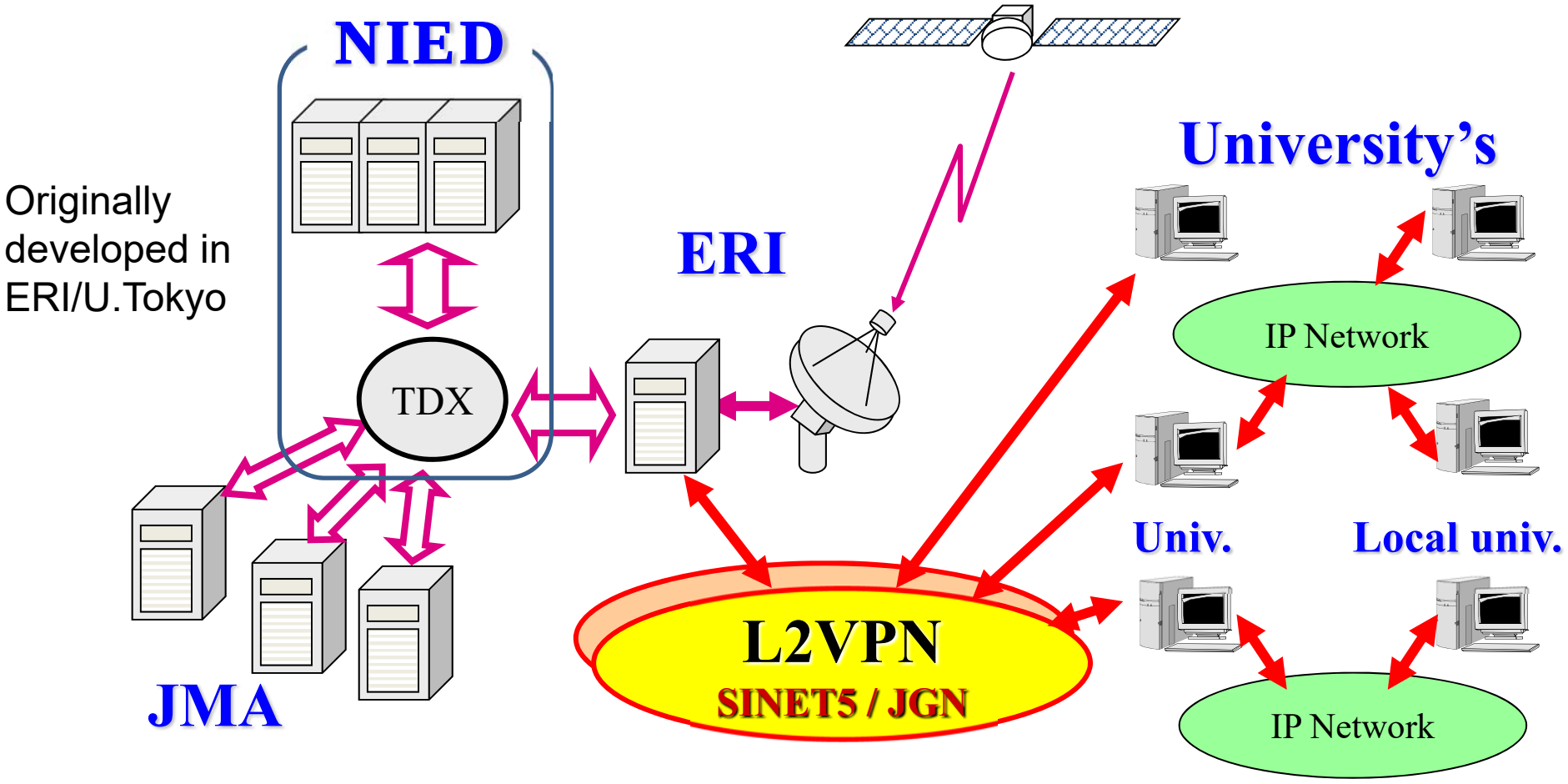
リアルタイム（地震Sim.＋データ同化）

- 全国の地震観測網（気象庁・防災科技研・地震研）は2,000点規模（100Hz，3方向）⇒100GB/Day，JDXnetによってSINET経由でリアルタイムに入手できる
- 外部ノード
  - JDXnetリアルタイムデータ
- 融合ノード
  - リアルタイムデータ同化
  - 全体制御
- 内部ノード
  - 大規模シミュレーション
- 通常時
  - 地下モデル更新（データ＋シミュレーション）：
  - パラメータスタディ：機械学習

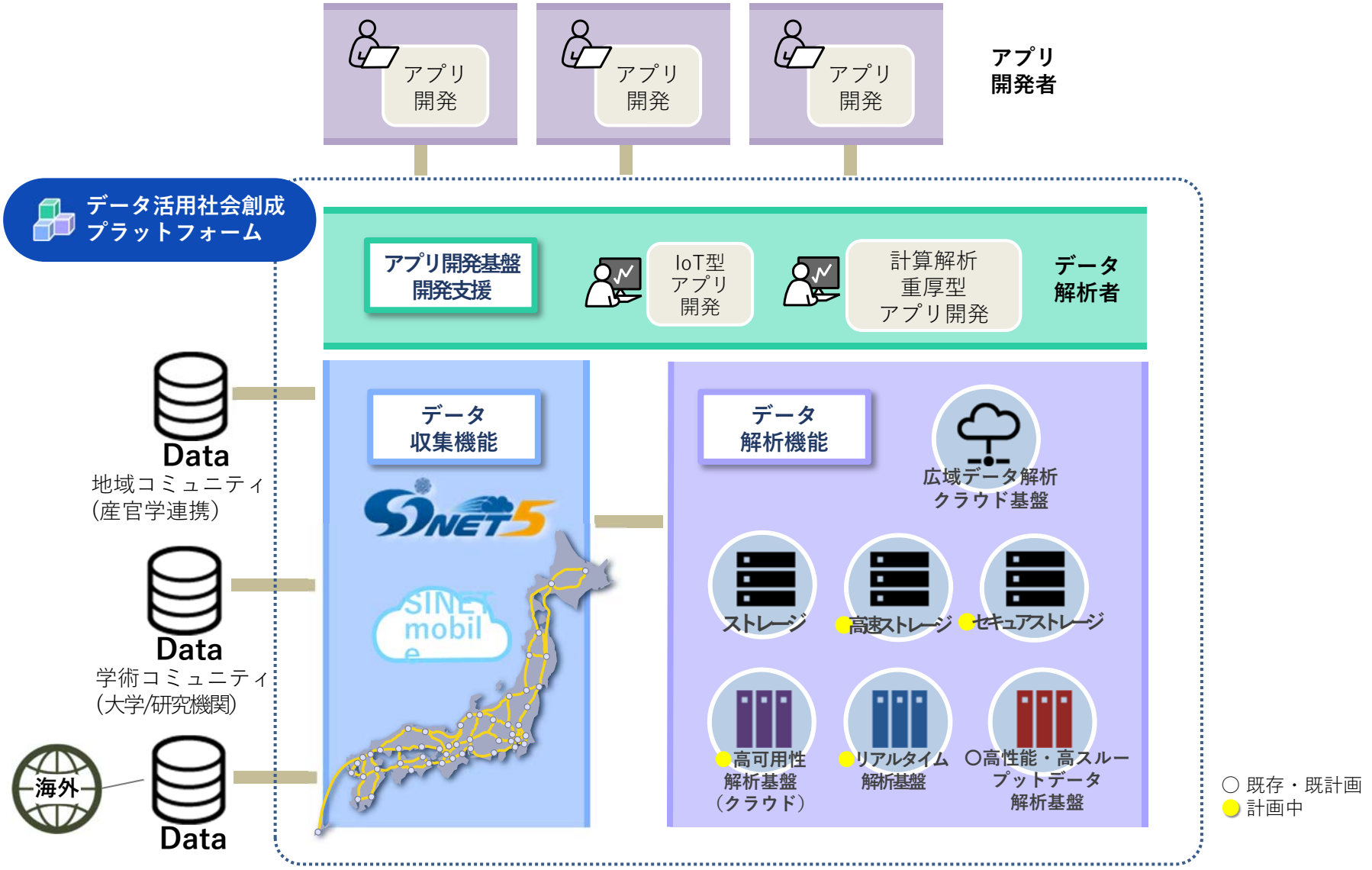




JDXnetによりSINET経由で国内地震観測データのリアルタイム取得が可能

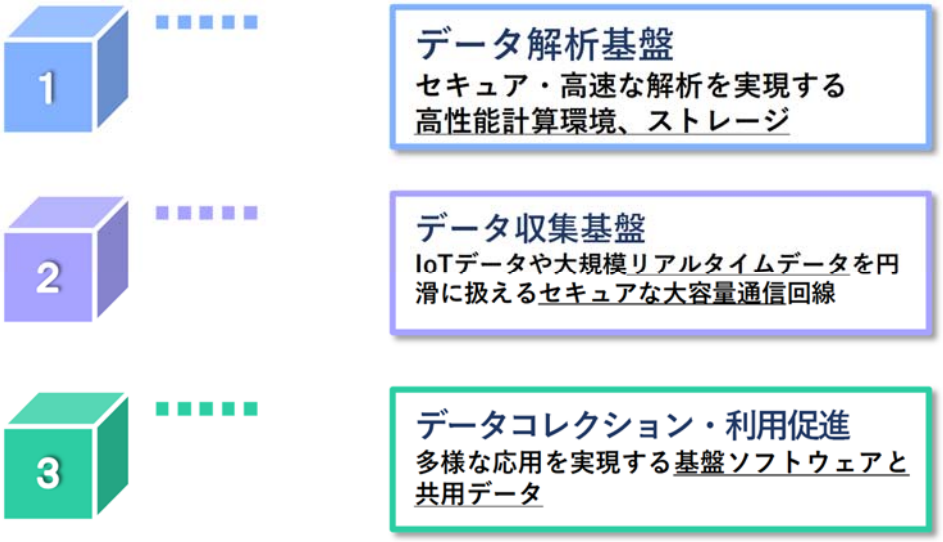


データ活用社会創成プラットフォーム（データプラットフォーム）



データプラットフォーム：Data Platform DP

- Society 5.0を目指したデータ利活用プラットフォーム
- 2019年度概算要求：2020年度末に導入
  - 国立大学＋国立研究所
    - 8大学（北大，東北大，東大，東工大，名古屋大，京大，阪大，九大）
    - 国立情報学研究所（NII）
    - 産業技術総合研究所（AIST）

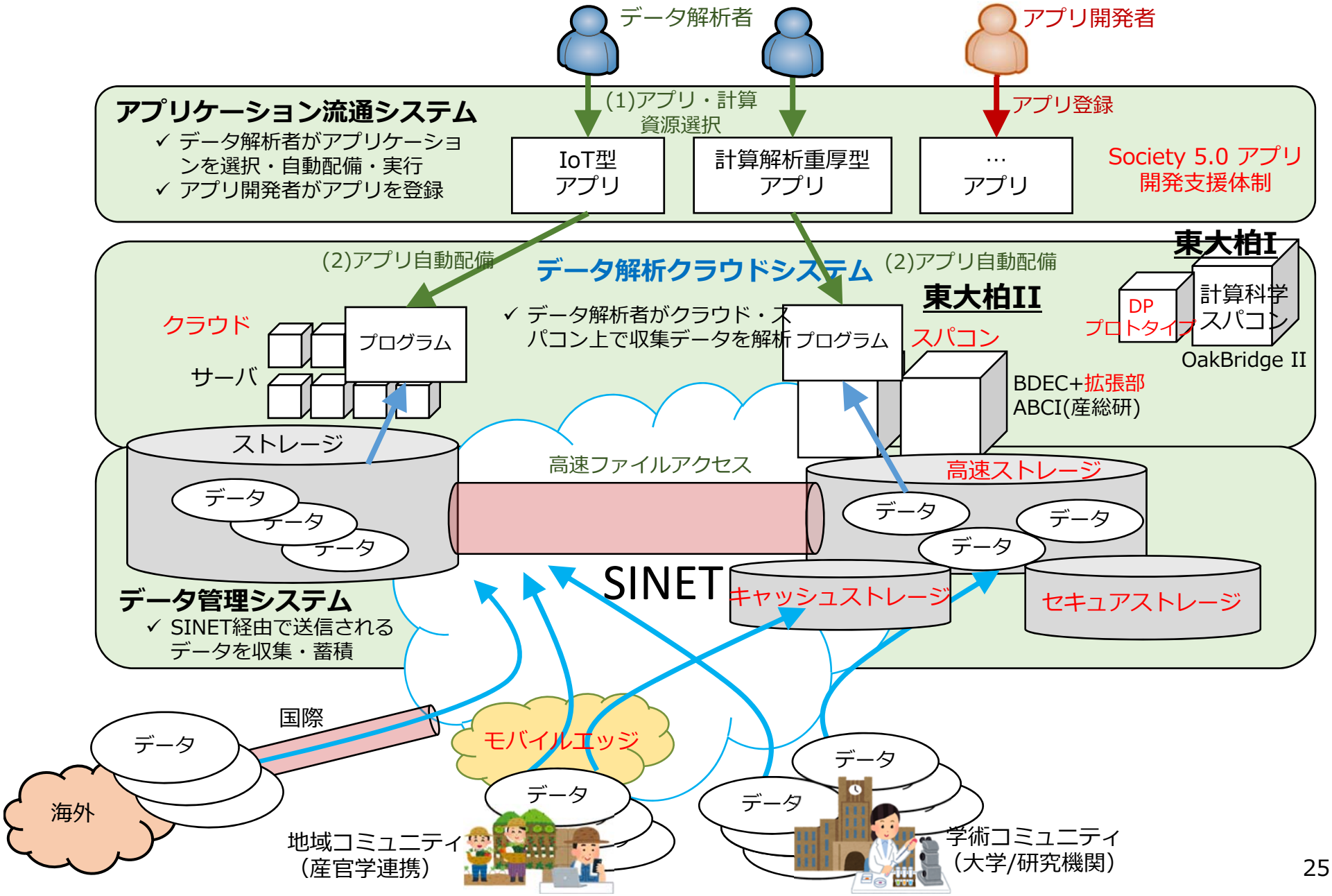


データプラットフォームとは

- 「知」の抽出により、Society 5.0における「ビッグデータ等の新たな技術をあらゆる産業や社会生活に取り入れてイノベーションを創出」するための「場」
  - アイデアはあっても実際にデータを収集・集積・解析する仕組みを作ることが困難な（地方）大学や産学連携体に「場」を提供、「知」の抽出を容易に行えるようにする
  - 短時間でのプロトタイピングを可能にし、機を逃さないシステム構築を支援する
- データプラットフォームの３本柱

1. SINETを活かしたリアルタイム収集・集積・解析環境の動的な構築
   - 遠隔地のセンサーやストレージ、データプラットフォームの計算資源、ストレージをつないで、リアルタイムに入力から出力を得られるアプリケーションごとの収集・集積・解析環境（スライス）を、使いたいときに即時に構築する
   - SINETモバイル基盤によりセンサー等のデータを安定してセキュアにつなぐ
2. 高性能計算環境によるデータ科学と計算科学の融合
   - データ科学、計算科学の手法を融合し、さらに国内最高の計算環境を用いて他に無い高精度の予測を行えるようにする
3. 開発支援とコミュニティーの形成
   - データ科学、計算科学、活用可能データなどの知見を基にコンサルティング・開発支援を行う
   - データ保持者、データ利用者等のコミュニティーを形成し、新たなデータ活用につなげる。

データプラットフォームの概要



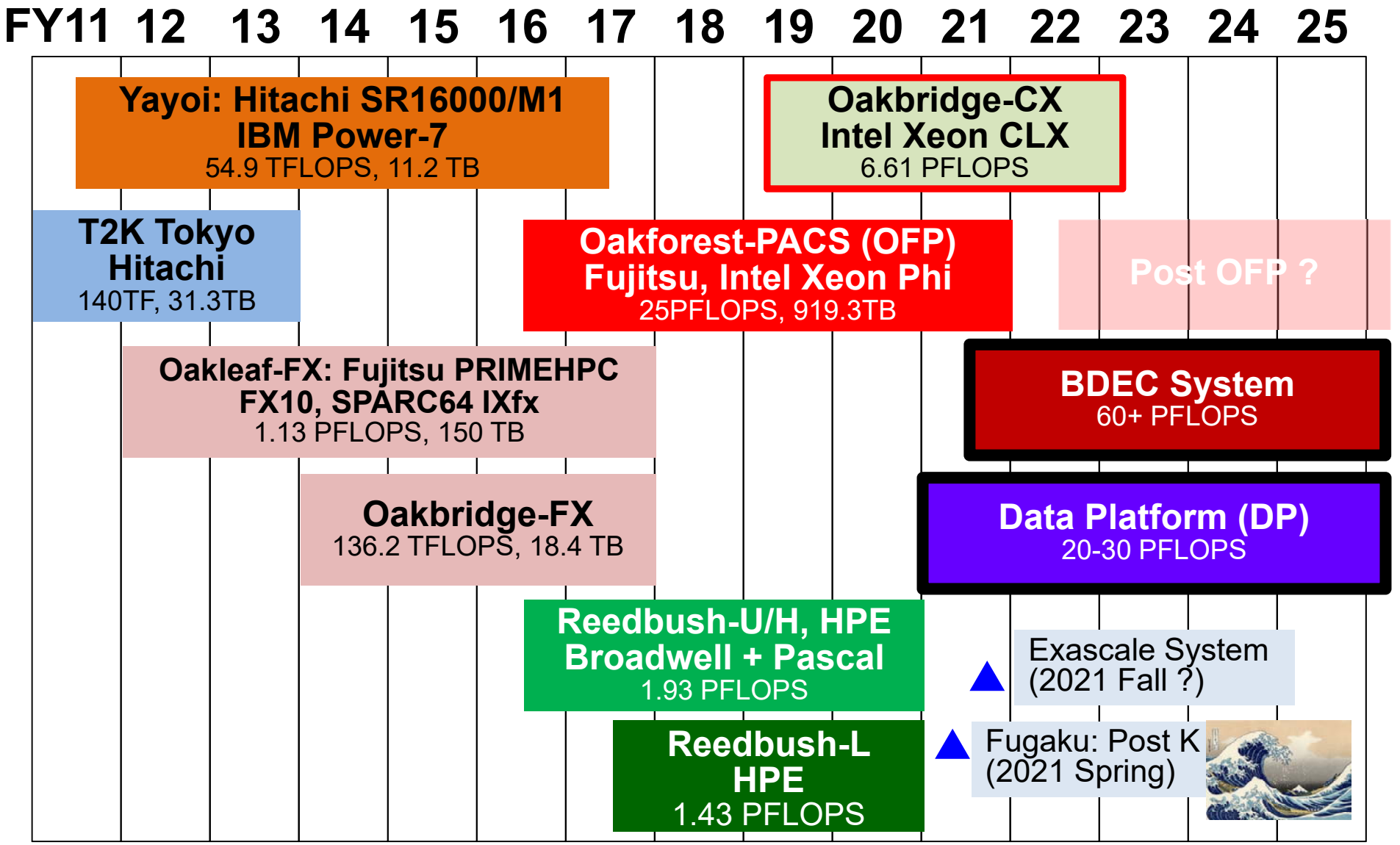
データプラットフォームの機能

- データ収集：モバイルエッジなどからSINETを介してデータを集める
- データ集積：データを集積
  - さまざまなデータ：フィールドから収集するデータ、シミュレーションで生成するデータ
  - 高速ストレージ：大量のデータを集積
  - セキュアストレージ：セキュリティが必要なデータを集積
  - キャッシュストレージ：リアルタイム処理に必要なデータを集積
- データ解析：データ科学、計算科学を融合してデータを解析
  - 国内最高レベルの計算機環境を提供
  - 解析に必要な大量データを供給できるネットワーク環境を提供

建　物

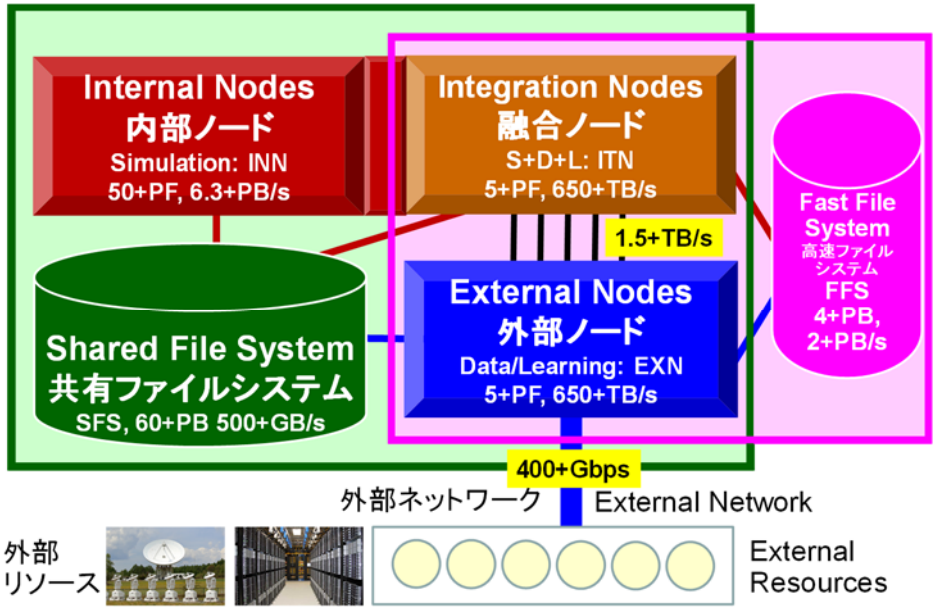
- 浅野キャンパス
  - Reedbush（2021年3月末退役予定）
  - Oakbridge-FX（退役）
- 柏キャンパス（第二総合研究棟）
  - 1F：Oakleaf-FX（退役），Oakbridge-CX（2019年7月運用開始）
  - 2F：Oakforest-PACS
- 柏IIキャンパス（H29補正）
  - 学術高速大容量ネットワーク 拠点整備構想知識集約型社会に向けた基盤整備
    - NIIと東大センター（スパコン・ネットワーク）による拠点形成
  - 2020年に完成予定
  - BDECシステム（2021年10月以降運用開始）
  - Data Platform（2021年4月運用開始）
    - 両者は同じ「部屋」に設置

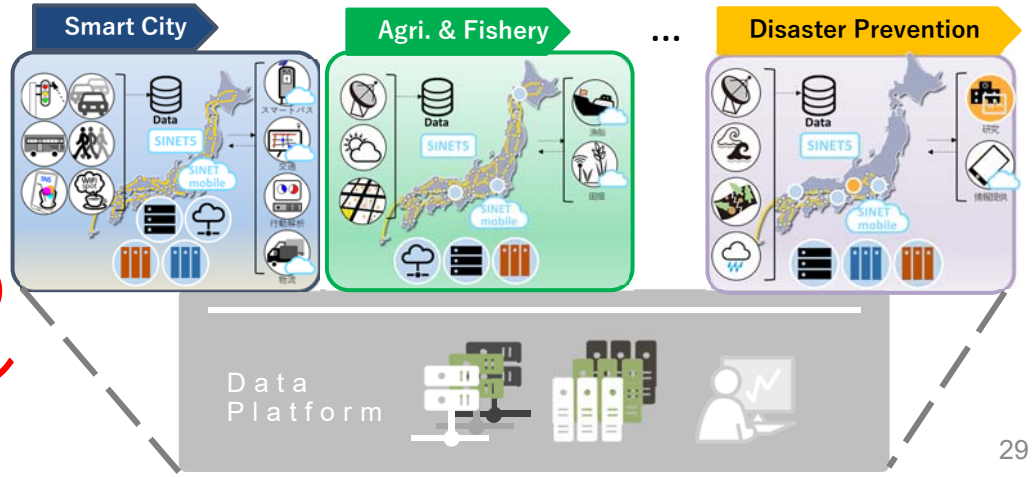
東大センターのスパコン



BDECとデータプラットフォーム（DP）

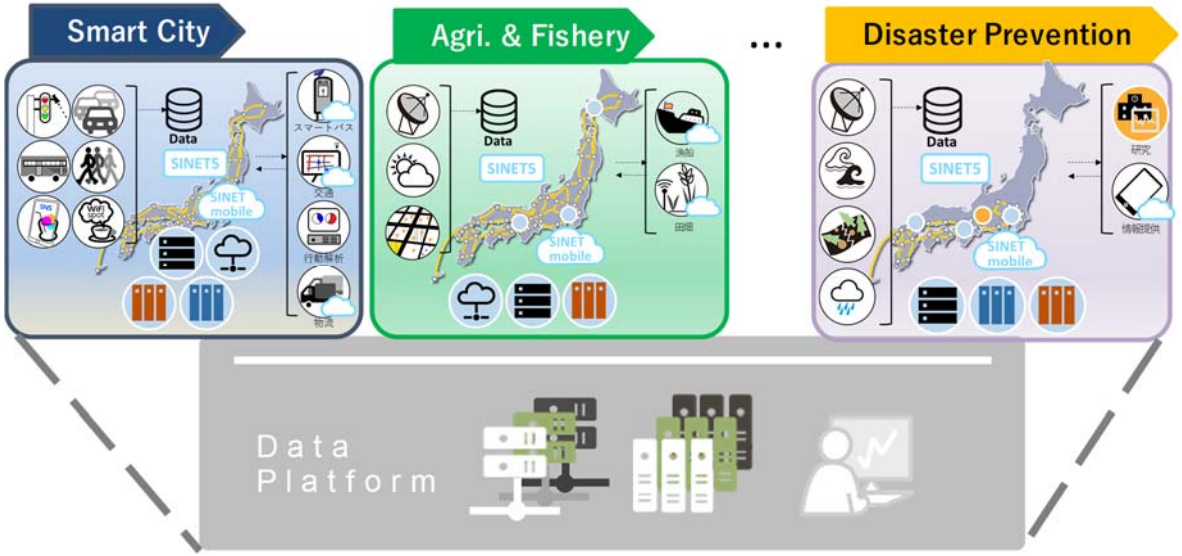
- BDEC
  - データ処理・学習機能をもったスパコン
    - 総ピーク性能60+PF
- データプラットフォーム：DP
  - スパコン＋ストレージ
    - 20-30 PF（倍精度ベース）
    - 30-50%（規模：BDEC比）
  - BDECよりデータ解析・利活用に重点
  - 計算機システムとしてはBDECの外部ノード（EXN）と似ているが，よりフレキシブル，よりsecure



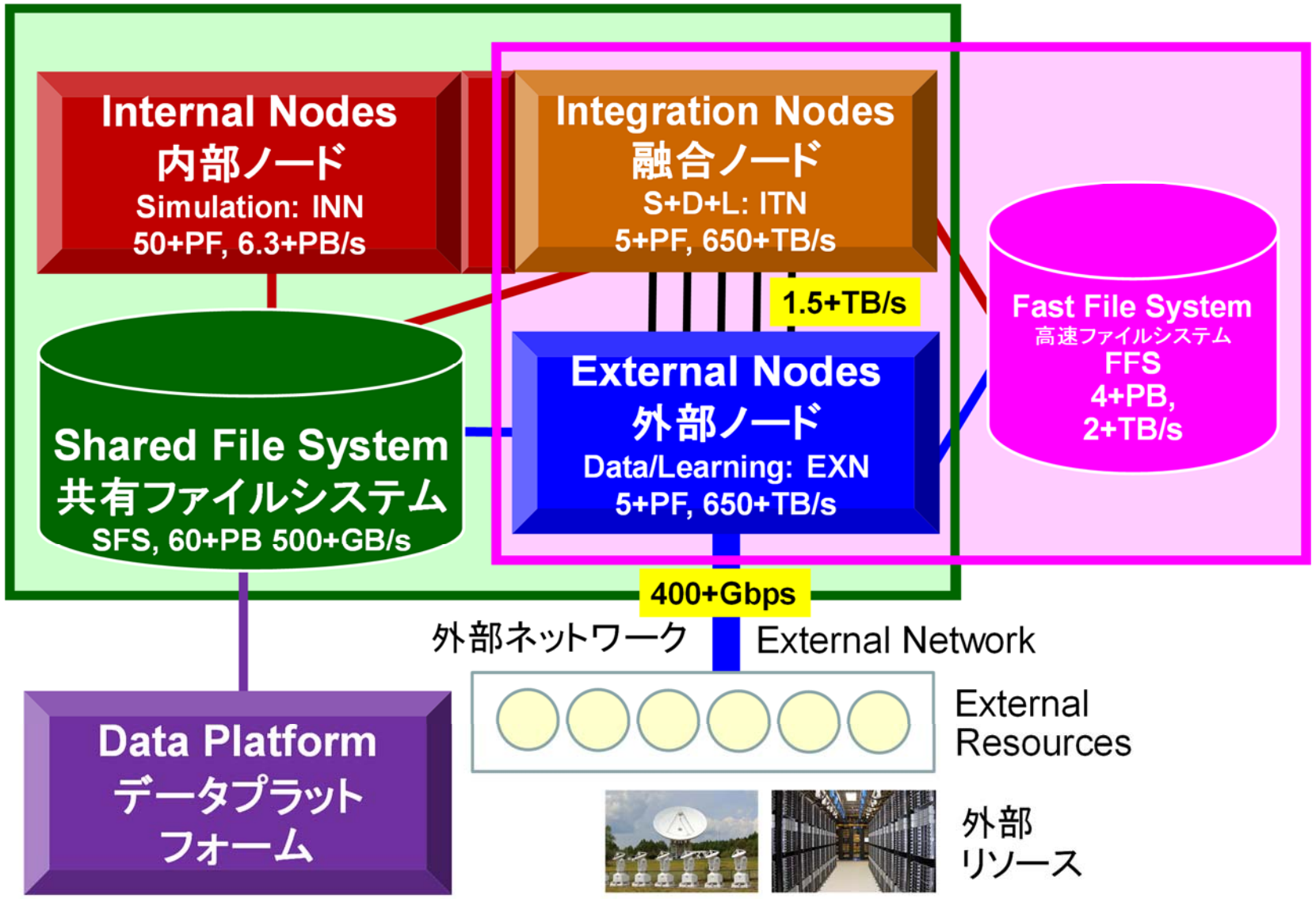


データプラットフォームの特徴

- オンデマンドなConfigurationが可能
- リソースマネージメント & スケジューリング
  - Kubernetes, OpenStack, …
  - Slurm, PBS, …
- 「スライス（slice）」または仮想プラットフォーム
  - Isolated & Securedなリソース群（ストレージ，ネットワーク，計算機他）
  - VLAN, SDN
  - NVMe over Fabrics



BDECの共有ファイルシステム（SFS）はDPからアクセス可能（必須）



- 東大情報基盤センターについて
- **サービス概要**
- スーパーコンピュータシステム概要
- 運用
- 質疑

教育・人材育成(1/2)

- お試しアカウント付き並列プログラミング講習会
  - 現在Oakbridge-CX を利用した講習会を企画中
  - 既存利用者に限定せず，企業の技術者・研究者も受講可能
    - ✓受講者の3分の2以上は企業から受講：裾野拡大に大きな貢献
    - ✓PCクラスタコンソーシアム(実用アプリケーション部会)と共催
  - 1～2日間の講習，1ヶ月有効な「お試しアカウント」
    - ✓MPI基礎，MPI応用(並列有限要素法)，マルチコアプログラミング
    - ✓ライブラリ利用(センター教員開発のライブラリ普及)，OpenFOAM
  - 2019年度からの新企画
    - ✓Altair HyperWorks実行，Altair ultraFluidX入門
- 学部・大学院・高専の講義での利用(学外含む)
  - 提案書ベース，無料，専用のバッチジョブキュー
    (Oakbridge-CX: 8ノード，15分)
  - 年10件程度(うち2~3は学外から)，センター教員の講義

教育・人材育成(2/2)

- 若手・女性支援
  - 40歳以下(女性は年齢制限無し):無料でOakbridge-CX, Oakforest-PACS, Reedbush を使用可能
  - 公募型(年2回，各半年間，連続して2回応募可能⇒1年間無料)
  - 学生を対象とした「インターンシップ」制度，グループ制度
  - https://www.cc.u-tokyo.ac.jp/guide/young/
- 計算科学アライアンス(2014～)
  - 学際的研究の拠点，全学的なHPC教育プログラムの策定
  - H28概算要求(正式に始動)，関連部局
  - Double Degreeを見据えた横断型・学際型プログラム
  - 情報理工学系・理学系・情基セの協力が元になっている
  - http://www.compsci-alliance.jp/

平成28年度概算要求
情報理工学系研究科

計算科学アライアンス：自然・社会科学と情報科学の連携教育研究

背景

- 多数の社会活動・学術分野で計算科学・情報の知見の必要度が増大
  しかし，計算科学，自然科学，情報科学が専門分化して俯瞰力が低下
- 計算科学ソフトウェアが高度化・大規模化する中，並列処理など
  計算効率の原理が変化
- 海外ソフトの利用が多く，中国・ブラジルなど新興国も追い上げ，
  ポスト京アプリ開発が始動するも，国産ソフト開発・競争力強化が必要
- 企業もシミュレーションを重要視しているが，企業内の専門家はわずか

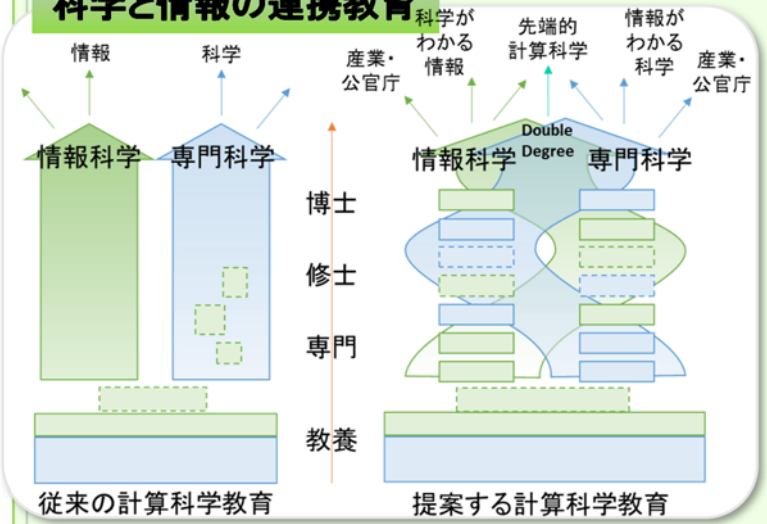
目的

- **自然・社会科学と情報科学を俯瞰した標準教育体系を整備**
- **大規模ソフトウェア，並列計算などの手法を教える**
- **国産シミュレーションソフトの開発力・競争力を強化**
- **企業のシミュレーション能力・活用を向上する人材を育成**

提案：計算科学アライアンス

自然・社会科学と情報科学が連携して
計算科学の体系的教育・人材育成を行う

科学と情報の連携教育



- 柔軟なコース設計：4学期・国際化対応
- 学部から大学院まで，習熟度別・英才教育
- キャリアパス整備：企業へのアピール

英才教育

- 自然・社会科学とソフト・ハード・計算を横断・俯瞰して理解・活用できる人材育成
- ダブルディグリー制度を目指す
  - 例：物理博士・情報修士
  - 入試や学位のありかた検討
- スーパーコンピュータ，ソフトウェア工学などの新しい情報技術を教育

基礎教育

- 自然・モデル・計算の基礎教育
- バイリンガル教材 → 国際的に通用
- 社会人教育対応コース，
  リテラシー教育

国際連携先端研究

- 海外トップ研究者を招聘
- 国際共同研究を探る WS
- 国際的に開かれた計算科学サマースクール，オンライン講座の提供
- 学生の留学・短期派遣支援
- 優秀な留学生・女性獲得活動
- 外国人・女性の教員・研究員採用

組織

大規模部局連携で
本学全体を底上げ

理学系研究科
工学系研究科
情報理工学系研究科
情報基盤センター
新領域創成科学研究科
数理科学研究科

外部連携

生産技術研究所
物性研究所
地震研究所
大気海洋研究所
IPMU

共同利用共同研究拠点
（7大学，筑波大，東工大）
理化学研究所
官庁・自治体・企業等
海外大学・研究所等

資源

部局資源の活用・強みを集中
・再配分されたポストを活用
・基盤センターのスパコン利用

学術的必要性

- ポスト京の計算科学で突出した成果を挙げる
- 国際的な先導的立場を維持，発展させる
- あらゆる学問分野で高性能計算を活用

社会的必要性

- 日本の産業と科学を支える人物の輩出
- 我が国の国際社会における優位性
- 計算科学の力を活用する産学官リーダー

期待される成果

- 自然・社会科学と情報科学を俯瞰した標準体系カリキュラムの整備
- 世界を先導する先端的成果と国際連携拠点の形成
- 計算科学で世界的成果を挙げ日本の科学・産業を先導する人材の育成
- ハード・ソフト・自然科学を横断して先導できる人材の育成
- 本学のあらゆる専門学問領域において計算機活用による高度化を促進
- 計算科学を理解し，強みを活かす企業トップ，学術リーダー，官僚等の輩出
- 企業等においてシミュレーション等を駆使して競争力を発揮する人材の輩出

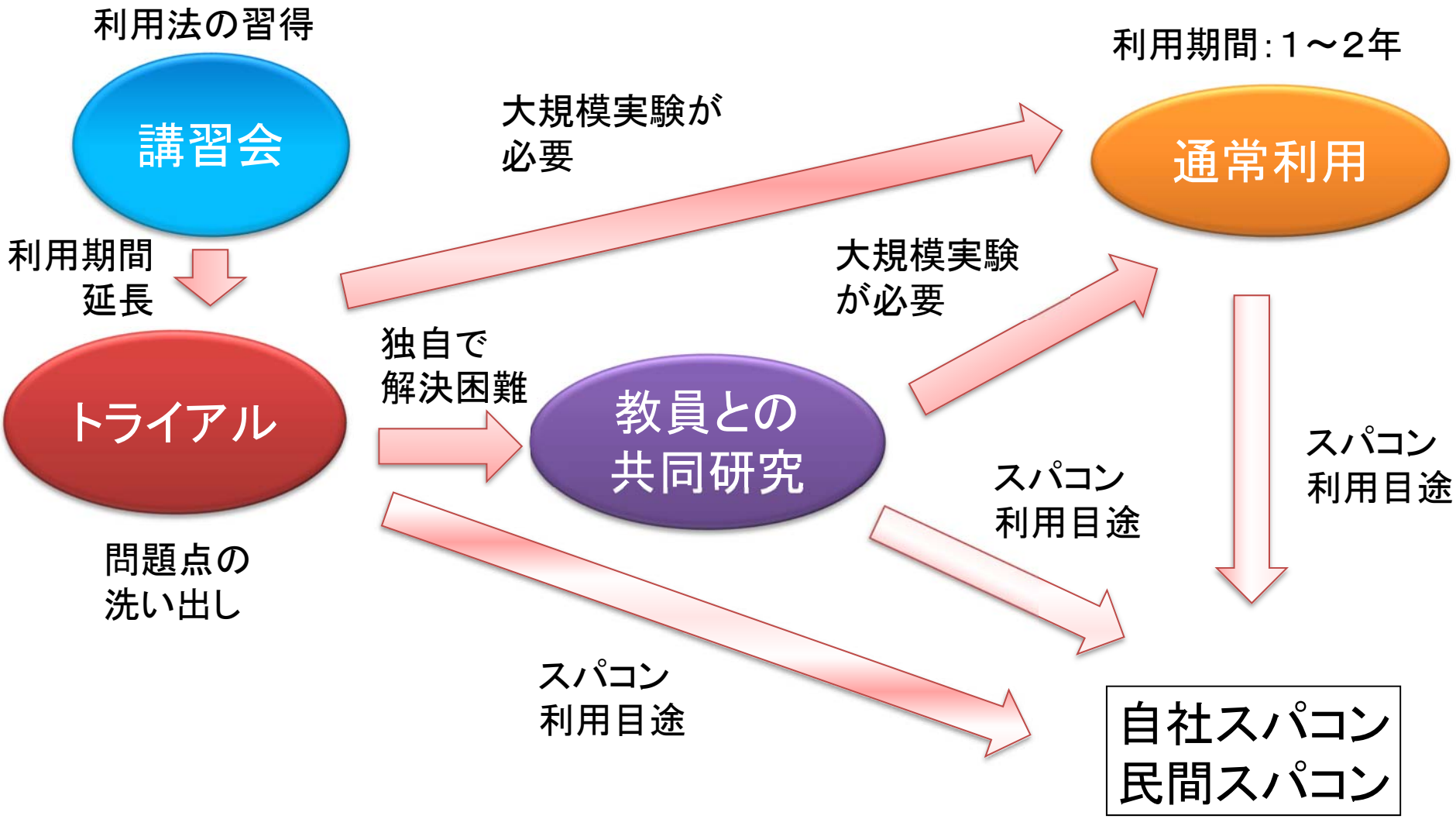
本学・我が国の競争力強化

企業利用

https://www.cc.u-tokyo.ac.jp/guide/company/

- 2008年度より開始
  - 大規模並列計算普及，社会貢献
  - ビジネスへの萌芽的段階での支援，データセンターと競合しない
  - 成果は原則公開，全資源の10%以下
  - アカデミック利用者と異なる審査基準（年2回募集），負担金体系
- 様々な利用体系
  - 通常グループ利用（いわゆる「企業利用」）
    - ✓毎年3－4グループ，基礎的な研究が多い
  - トライアルユース（グループ（有償，無償），個人）
    - ✓お試しアカウント付き講習会を受講するとパーソナルトライアルユース可能
  - 大学等との共同研究（アカデミック料金で利用可能）：2-3件
  - オープンソース，自作コードに限定（ISVアプリ無し）
    - ✓方針再検討も考えている：ご意見，ご要望をお聞かせください

東大スパコン企業利用制度の展開イメージ



大規模HPCチャレンジ

Oakforest-PACS, Reedbushで実施中
Oakbridge-CXでは本年10月以降実施予定

- https://www.cc.u-tokyo.ac.jp/guide/hpc/
- 月1回1日（24時間），1,280ノードを1グループで占有して実行できる，公募制，無料。
- OBCXユーザー以外も応募可能。
- 成果公開を義務づける
  - センター広報誌への寄稿
  - センター主催各種催しでの発表，各種外部発表への情報提供
  - 速報結果の査読付国際会議への投稿等による迅速，国際的な成果公開が望ましい。
- 企業からの申し込みも受け付ける（成果公開を義務づけ）
- 自作プログラム，オープンソースプログラム利用に限定

- 東大情報基盤センターについて
- サービス概要
- **スーパーコンピュータシステム概要**
- 運用
- 質疑

3システム：利用者2,000+，学外50+%

- 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
  - 3.36 PF, 2016年7月～ 2021年3月末（予定）
  - 東大ITC初のGPUシステム (2017年3月より), DDN IME (Burst Buffer)

- Oakforest-PACS (OFP) (富士通, Intel Xeon Phi (KNL))
  - JCAHPC (筑波大CCS＆東大ITC)
  - 25 PF, TOP 500で6位 (2016年11月) (日本1位)（初登場時）
  - Omni-Path アーキテクチャ, DDN IME (Burst Buffer)

- Oakbridge-CX (富士通, Intel Xeon Platinum 8280)
  - 大規模超並列スーパーコンピュータシステム
  - 6.61 PF, 2019年7月 ～ 2023年6月
  - 全1,368ノードの内128ノードにSSDを搭載

Reedbush (1/2)

- システム構成・運用：SGI => HP
- Reedbush-U (CPU only，2016年7月～)
  - Intel Xeon E5-2695v4 (Broadwell-EP, 2.1GHz, 18core) x 2ソケット (1.210 TF), 256 GiB (153.6GB/sec)
  - InfiniBand EDR, Full bisection BW Fat-tree
  - システム全系: 420 ノード, 508.0 TF
- Reedbush-H (with GPU，2017年3月～)
  - CPU・メモリ：Reedbush-U と同様
  - NVIDIA Tesla P100 (Pascal世代 GPU: 5.3TF, 720GB/sec, 16GiB) x 2 / ノード
  - InfiniBand FDR x 2ch, Full bisection BW Fat-tree
  - 120 ノード, 145.2 TF(CPU)+ 1.27 PF(GPU)= 1.42 PF
- Reedbush-L (with GPU: 長時間ジョブ用，2017年10月～)
  - CPU・メモリ：Reedbush-U と同様
  - NVIDIA Tesla P100 (Pascal世代 GPU: 5.3TF, 720GB/sec, 16GiB) x 4 / ノード
  - InfiniBand EDR x 2ch, Full bisection BW Fat-tree (U, Hとは少し遠い)
  - 64 ノード, 76.8 TF(CPU)+ 1.35 PF(GPU)= 1.43 PF

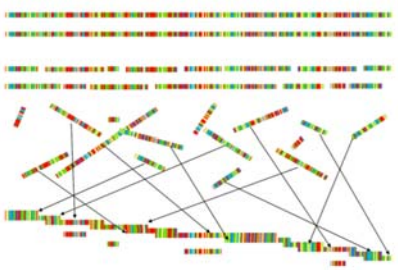
Reedbush (2/2)

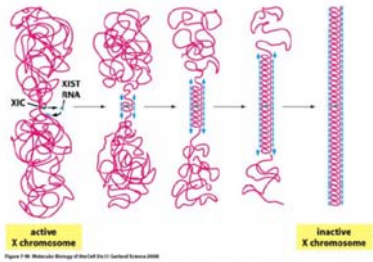
- ストレージ/ファイルシステム
  - 並列ファイルシステム (Lustre)
    - 5.04 PB, 145.2 GB/sec
  - 高速ファイルキャッシュシステム: Burst Buffer (DDN IME (Infinite Memory Engine)) : SSDによるキャッシュ
    - Reedbush-U,H: 230.4 TB, 385.2 GB/sec
    - Reedbush-L: 153.6 TB, 166.4 GB/sec
- 電力, 冷却, 設置面積
  - 空冷, 368 kW (RB-U,H) + 134 kW (RB-L) (冷却除く)
  - < 90 m^2
- データ解析、ディープラーニング向けソフトウェア・ツールキット
  - OpenCV, Theano, Anaconda, ROOT, TensorFlow, Torch, Caffe, Chainer, GEANT4

	Reedbush-U	Reedbush-H	Reedbush-L
	Integrated Supercomputer System for Data Analyses & Scientific Simulations		Supercomputer System with Accelerators for Long-Term Executions
CPU/node	Intel Xeon E5-2695v4 (Broadwell-EP, 2.1GHz, 18core) x 2 sockets (1.210 TF), 256 GiB (153.6GB/sec)		
GPU	-	NVIDIA Tesla P100 (Pascal, 5.3TF, 720GB/sec, 16GiB)	
Infiniband	EDR	FDR × 2ch	EDR × 2ch
Nodes #	**420**	**120**	**64**
GPU #	**-**	**240 (=120 × 2)**	**256 (=64 × 4)**
Peak Performance (TFLOPS)	**509**	**1,417 (145 + 1,272)**	**1,433 (76.8 + 1,358)**
Total Memory Bandwidth (TB/sec)	**64.5**	**191.2 (18.4+172.8)**	**194.2 (9.83+184.3)**
since	2016.07	2017.03	2017.10

GPUの導入

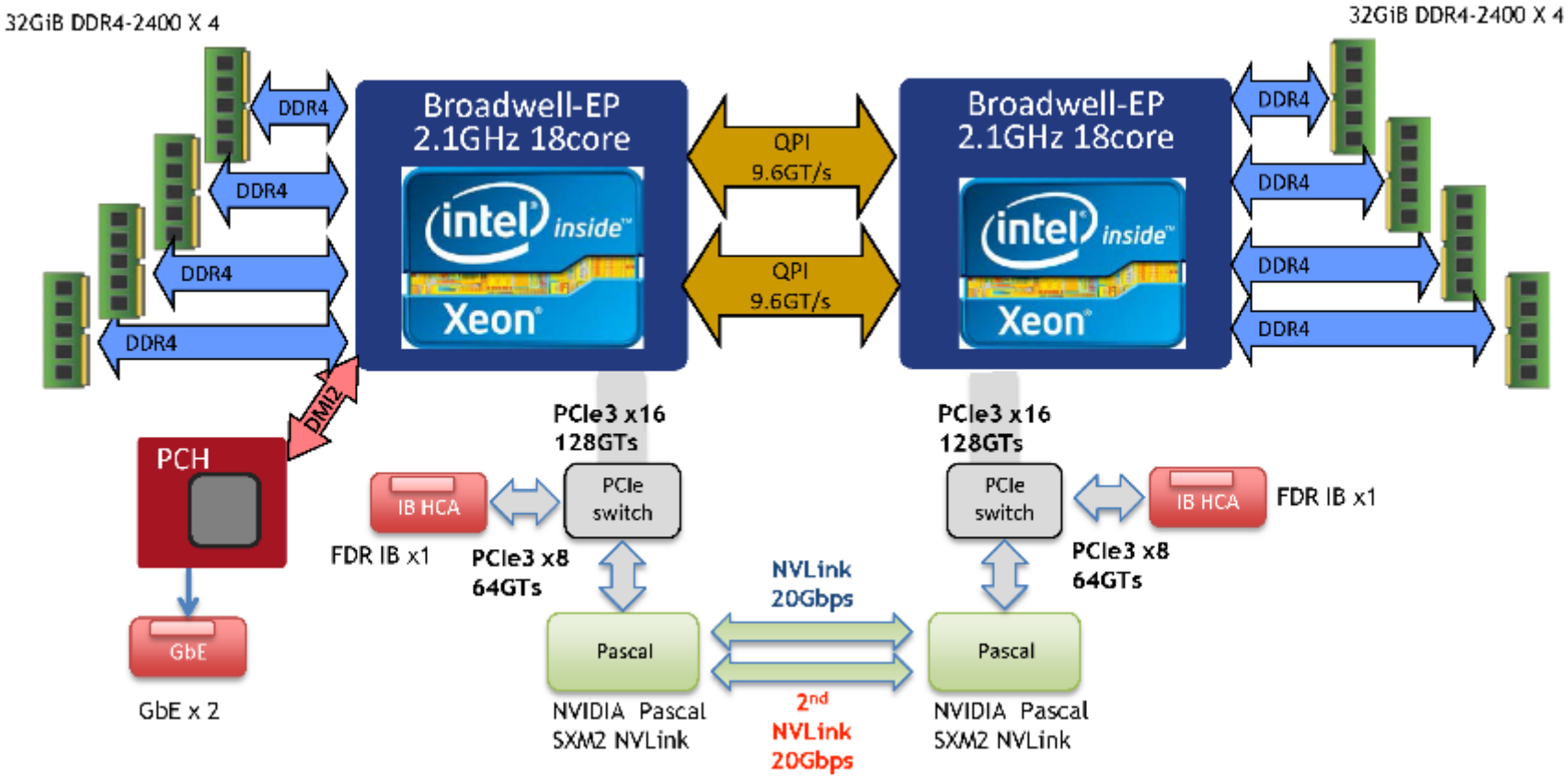
- OpenACC
  - OpenMPと類似したインタフェース：使いやすいが性能悪かった
    ⇨昨今の性能向上，CUDAとそれほど大きな差がなくなった
    - NVIDIA研究者との共同研究
  - OpenACC専門家など、GPUに詳しい人材の情報基盤センターへの加入
- データ科学，深層学習（Deep Learning）
  - 従来の計算科学，計算工学分野とは異なった分野の新規ユーザー開拓が急務：電気代＝負担金
  - 東京大学ゲノム医科学研究機構
  - 東京大学病院
    - 医療画像処理への深層学習適用





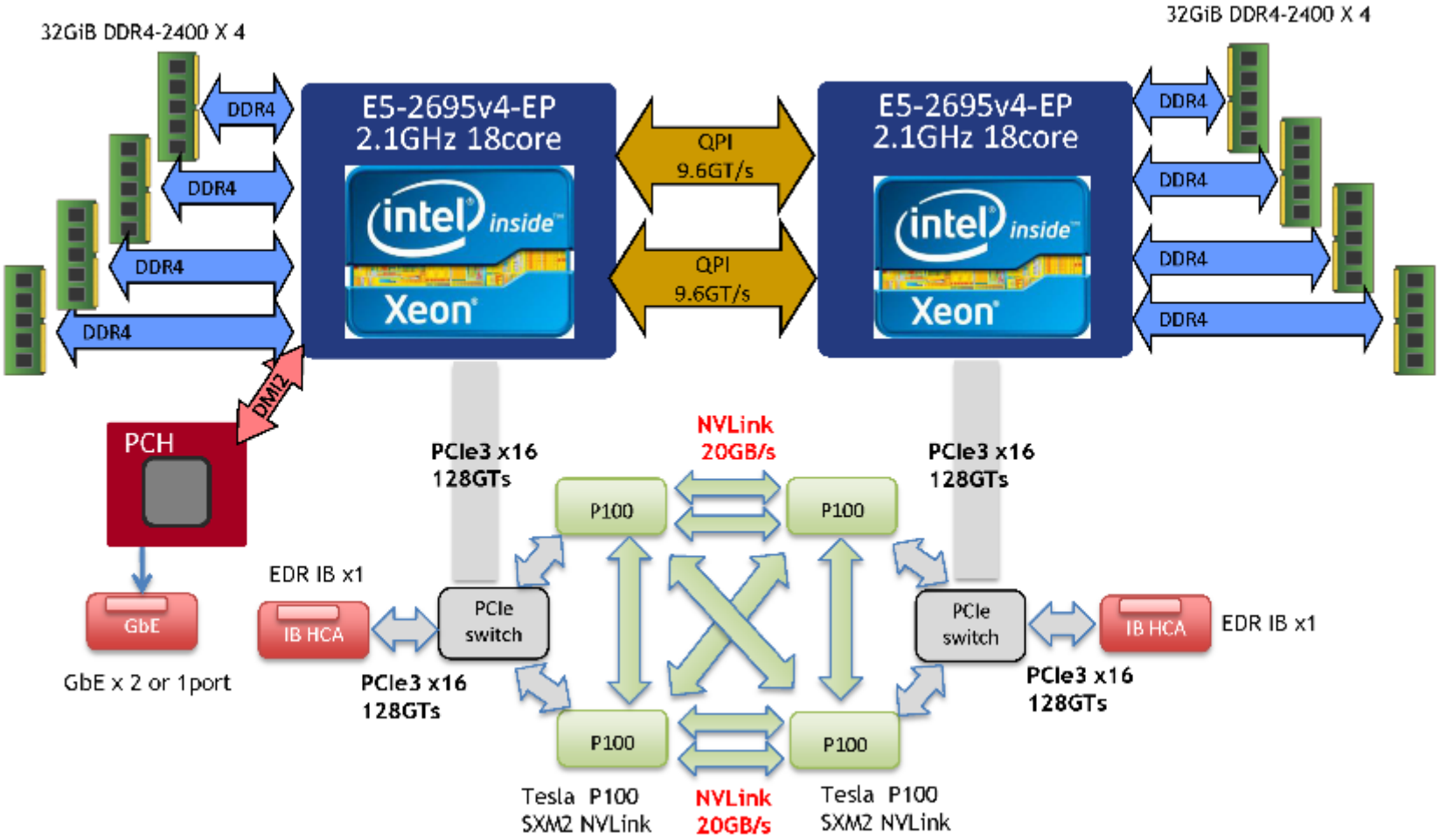
Compute Node of Reedbush-H

Reedbush-L: 各ソケットにPascal 1個=>2個ずつ，FDR=>EDR

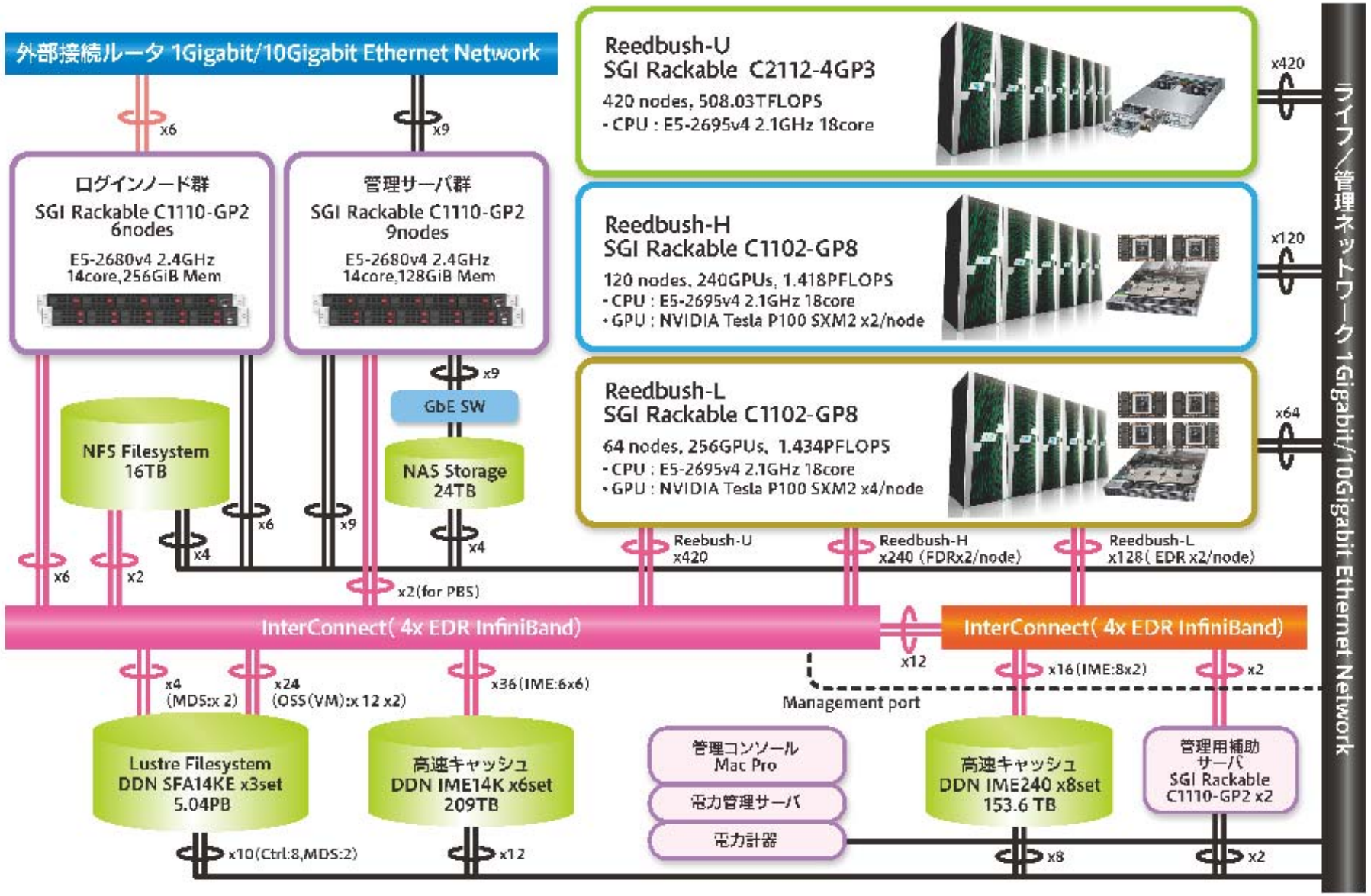


Compute Node of Reedbush-L

Reedbush-L: 各ソケットにPascal 1個=>2個ずつ, FDR=>EDR



Reedbushシステム概要



ソフトウェア構成

項目	Reedbush-U	Reedbush-H, L
OS	Red Hat Enterprise Linux 7	
コンパイラ	GNU コンパイラ Intel コンパイラ (Fortran77/90/95/2003/2008、C、C++)	
		PGI コンパイラ (Fortran77/90/95/2003/2008、C、C++、OpenACC 2.0、CUDA Fortran) NVCC コンパイラ (CUDA C)
メッセージ通信ライブラリ	Intel MPI, SGI MPT, Open MPI, MVAPICH2, Mellanox HPC-X	
		GPUDirect for RDMA: Open MPI, MVAPICH2-GDR
ライブラリ	Intel 社製ライブラリ(MKL): BLAS、LAPACK、ScaLAPACK その他ライブラリ: SuperLU、SuperLU MT、SuperLU DIST、METIS、MT-METIS、ParMETIS、Scotch、PT-Scotch、PETSc、FFTW、GNU Scientific Library、NetCDF、PnetCDF など	
		cuBLAS、cuSPARSE、cuFFT、MAGMA、OpenCV、ITK、Theano、Anaconda、ROOT、TensorFlowなど
アプリケーション	OpenFOAM、ABINT-MP PHASE、FrontFlow、FrontISTR、REVOCAP、ppOpen-HPC など	
デバッガ、プロファイラ	Total View, Intel VTune, Trace Analyzer & Collector	

ソフトウェア構成：データ解析向け

- OpenCV
  - コンピューター・ヴィジョン・ライブラリ
- ROOT
  - ビッグデータ向けのライブラリ
- TensorFlow, Keras, Chainer, Horovod, …
  - 機械学習向けライブラリ
- Singularity
  - Docker互換コンテナ

など…要望に応じて追加

3システム：利用者2,000+，学外50+%

- 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
  - 3.36 PF, 2016年7月～ 2021年3月末（予定）
  - 東大ITC初のGPUシステム (2017年3月より), DDN IME (Burst Buffer)

- Oakforest-PACS (OFP) (富士通, Intel Xeon Phi (KNL))
  - JCAHPC (筑波大CCS＆東大ITC)
  - 25 PF, TOP 500で6位 (2016年11月) (日本1位)（初登場時）
  - Omni-Path アーキテクチャ, DDN IME (Burst Buffer)

- Oakbridge-CX (富士通, Intel Xeon Platinum 8280)
  - 大規模超並列スーパーコンピュータシステム
  - 6.61 PF, 2019年7月 ～ 2023年6月
  - 全1,368ノードの内128ノードにSSDを搭載

Oakforest-PACS （OFP）

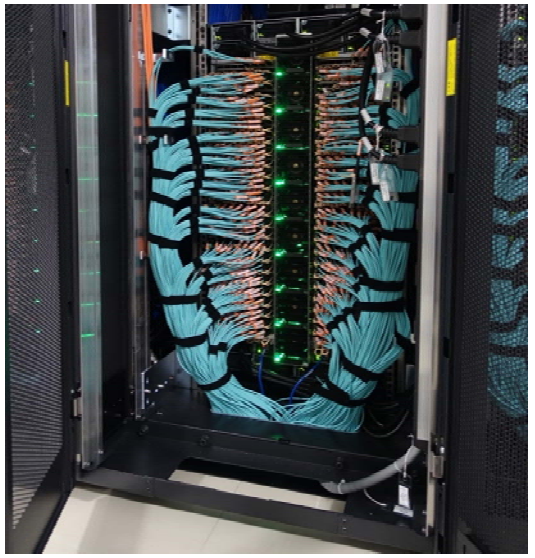
- 2016年12月1日稼働開始
- 8,208 Intel Xeon/Phi (KNL)，ピーク性能25PFLOPS
  - 富士通が構築
- **TOP 500　16位（国内2位），HPCG 9位（国内3位）（2019年6月）**
- **最先端共同HPC 基盤施設(JCAHPC: Joint Center for Advanced High Performance Computing)**
  - 筑波大学計算科学研究センター
  - 東京大学情報基盤センター
    - 東京大学柏キャンパスの東京大学情報基盤センター内に、両機関の教職員が中心となって設計するスーパーコンピュータシステムを設置し，最先端の大規模高性能計算基盤を構築・運営するための組織
  - http://jcahpc.jp

JCAHPC

東京大学 THE UNIVERSITY OF TOKYO　筑波大学 University of Tsukuba

Oakforest-PACSの特徴 (1/2)

- 計算ノード
  - 1ノード 68コア，3TFLOPS×8,208ノード＝25 PFLOPS
  - メモリ（MCDRAM（高速，16GB）＋DDR4（低速，96GB））
- ノード間通信
  - フルバイセクションバンド幅を持つFat-Treeネットワーク
  - 全系運用時のアプリケーション性能に効果，多ジョブ運用
  - Intel Omni-Path Architecture



Oakforest-PACS の仕様

総ピーク演算性能			25 PFLOPS
ノード数			8,208
計算ノード	Product		富士通 PRIMERGY CX600 M1 (2U) + CX1640 M1 x 8node
	プロセッサ		Intel® Xeon Phi™ 7250 （開発コード: Knights Landing） 68 コア、1.4 GHz
	メモリ	高バンド幅	16 GB, MCDRAM, 実効 490 GB/sec
		低バンド幅	96 GB, DDR4-2400, ピーク 115.2 GB/sec
相互結合網	Product		Intel® Omni-Path Architecture
	リンク速度		100 Gbps
	トポロジ		フルバイセクションバンド幅Fat-tree網

Oakforest-PACS の特徴(2 / 2)

- ファイルI/O
  - 並列ファイルシステム:
    Lustre 26PB
  - ファイルキャッシュシステム
    (DDN IME):
    1TB/secを超える実効性能,
    約1PB
    - 計算科学・ビッグデータ解析
      ・機械学習にも貢献
- 消費電力
  - Green 500でも世界6位(
    2016年11月)
  - Linpack: 2.72 MW
    - 4,986 MFLOPS/W(OFP)
    - 830 MFLOPS/W(京)

並列ファイル
システム

ファイルキャッシュ
システム

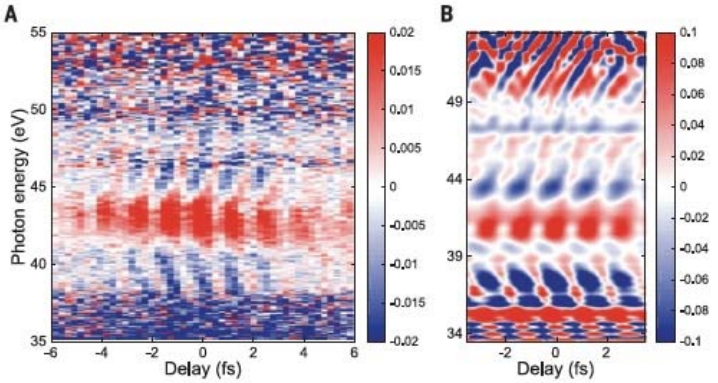
ラック当たり120ノード
の高密度実装

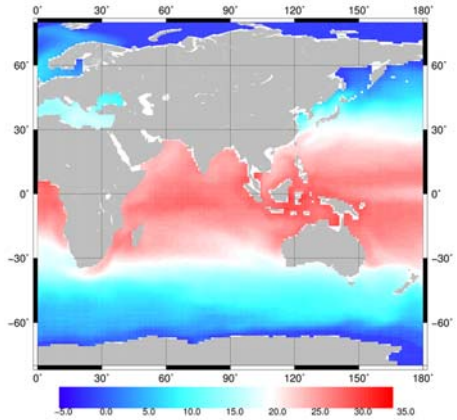
Oakforest-PACS の仕様（続き）

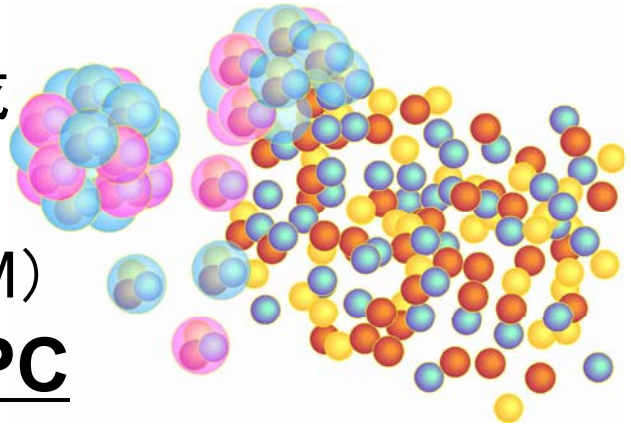
並列ファイルシステム	Type	Lustre File System
	総容量	26.2 PB
	Product	DataDirect Networks SFA14KE
	総バンド幅	500 GB/sec
高速ファイルキャッシュシステム	Type	Burst Buffer, Infinite Memory Engine (by DDN)
	総容量	940 TB (NVMe SSD, パリティを含む)
	Product	DataDirect Networks IME14K
	総バンド幅	1,560 GB/sec
総消費電力		4.2MW（冷却を含む）
総ラック数		102

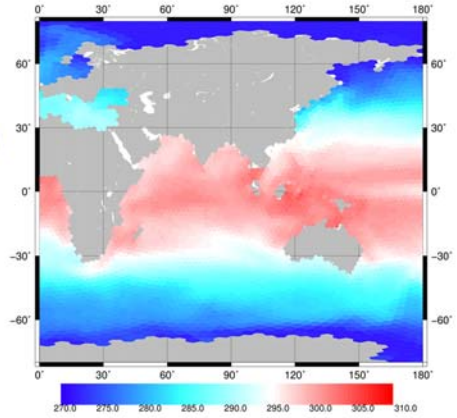
Oakforest-PACS：代表的アプリケーション

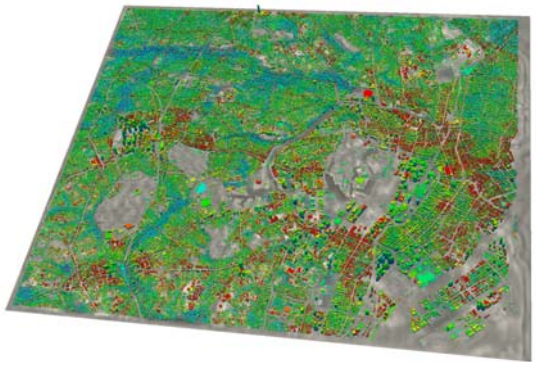
- **SALMON/ARTE**
  - 電子動力学
- **Lattice QCD**
  - 格子量子色力学
- **NICAM-COCO**
  - 全地球大気・海洋連成
- **GHYDRA**
  - 三次元地盤震動（FEM）
- **Seism3D/OpenSWPC**
  - 三次元広域波動伝搬（FDM）

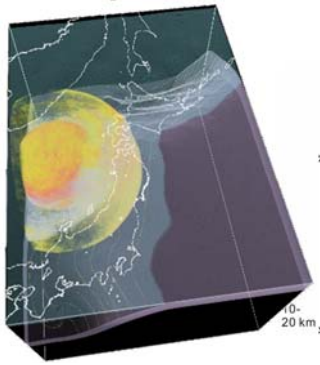


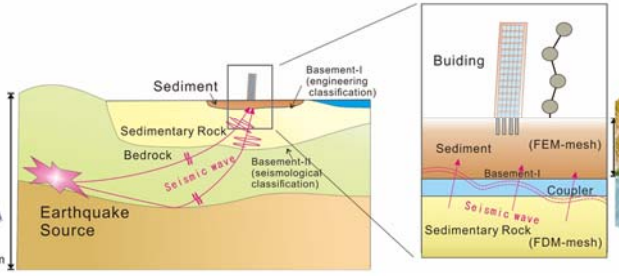


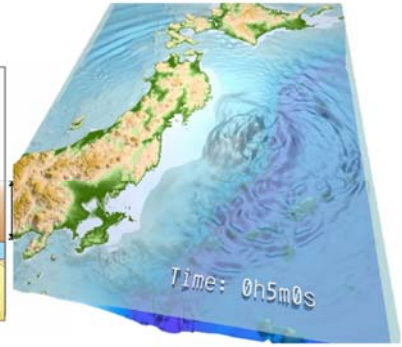












各種ベンチマーク

- TOP 500（Linpack，HPL）
  - 連立一次方程式ソルバー（直接法），計算速度（FLOPS値）
  - 規則的な密行列：連続メモリアクセス
  - 計算性能
- HPCG
  - 連立一次方程式ソルバー（反復法），計算速度（FLOPS値）
  - 有限要素法から得られる疎行列（ゼロが多い）
    - 不連続メモリアクセス
    - 実アプリケーションに近い
  - メモリアクセス性能，通信性能
- Green 500
  - HPL（TOP500）実行時のFLOPS/W値

53rd TOP500 List (June, 2019)

R_{max}: Performance of Linpack (TFLOPS)
R_{peak}: Peak Performance (TFLOPS),
Power: kW

http://www.top500.org/

	Site	Computer/Year Vendor	Cores	R_{max} (TFLOPS)	R_{peak} (TFLOPS)	Power (kW)
1	**Summit, 2018, USA** DOE/SC/Oak Ridge National Laboratory	IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	2,414,592	148,600 (= 148.6 PF)	200,795	**10,096**
2	**Sieera, 2018, USA** DOE/NNSA/LLNL	IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	1,572,480	94,640	125,712	**7,438**
3	**Sunway TaihuLight, 2016, China** National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	10,649,600	93,015	125,436	**15,371**
4	**Tianhe-2A, 2018, China** National Super Computer Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000	4,981,760	61,445	100,679	**18,482**
5	**Frontera, 2019, USA** Texas Advanced Computing Center	Dell C6420, Xeon Platinum 8280 28c 2.7GHz, Mellanox Infiniband HDR	448,448	23,516	38,746	
6	**Piz Daint, 2017, Switzerland** Swiss National Supercomputing Centre (CSCS)	Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100	387,872	21,230	27,154	**2,384**
7	**Trinity, 2017, USA** DOE/NNSA/LANL/SNL	Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect	979,072	20,159	41,461	**7,578**
8	**ABCI (AI Bridging Cloud Infrastructure), 2018, Japan** National Institute of Advanced Industrial Science and Technology (AIST)	PRIMERGY CX2550 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR	391,680	19,880	32,577	**1,649**
9	**SuperMUC-NG, 2018, Germany** Leibniz Rechenzentrum	Lenovo, ThinkSystem SD650, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path	305,856	19,477	26,874	
10	**Lassen, 2019, USA** DOE/NNSA/LLNL	IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta V100, Dual-rail Mellanox EDR Infiniband	288,288	18,200	23,047	
16	**Oakforest-PACS, 2016, Japan** Joint Center for Advanced High Performance Computing	PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path	556,104	13,556	24,913	2,719

HPCG Ranking (June, 2019)

	Computer	Cores	HPL Rmax (Pflop/s)	TOP500 Rank	HPCG (Pflop/s)
1	**Summit**	2,414,592	148,600	1	2.926
2	**Sierra**	1,572,480	94.640	2	1.796
3	**K computer**	705,024	10.510	20	0.603
4	**Trinity**	979,072	20,159	7	0.546
5	**ABCI**	391,680	19,880	8	0.509
6	**Piz Daint**	387,872	21.230	6	0.497
7	**Sunway TaihuLight**	10,649,600	93.015	3	0.481
8	**Nurion (KISTI, Korea)**	570,020	13.929	15	0.391
9	**Oakforest-PACS**	556,104	13.555	16	0.385
10	**Cori (NERSC/LBNL, USA)**	632,400	14.015	14	0.355

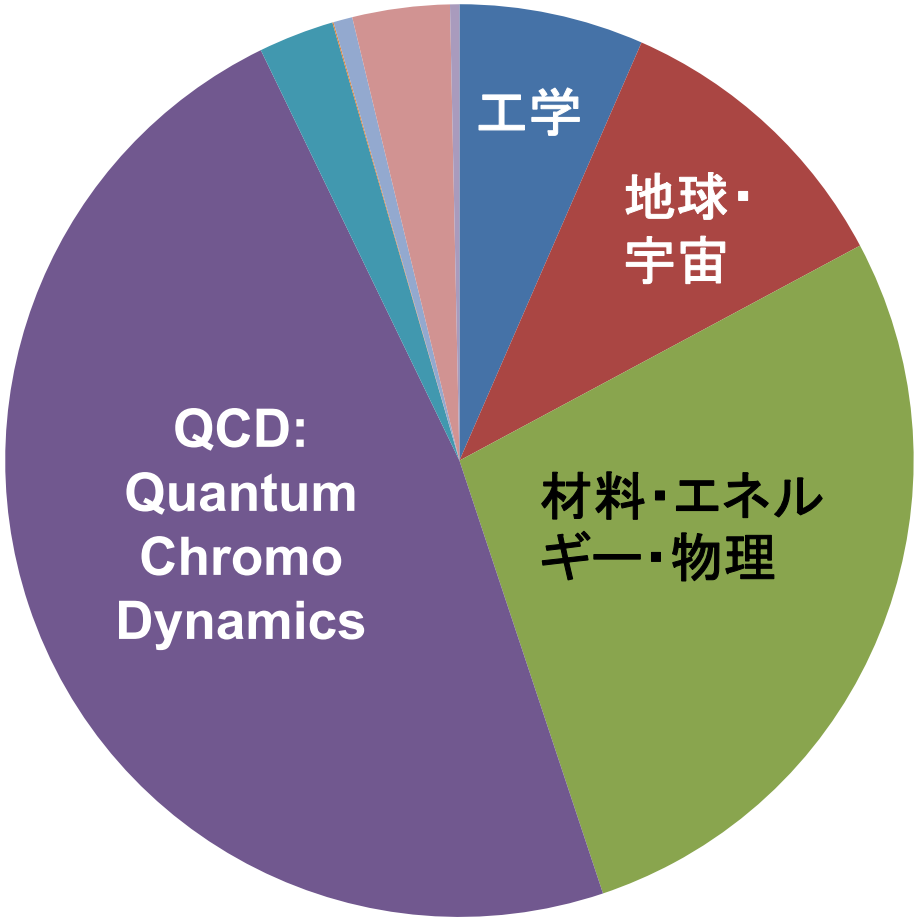
http://www.hpcg-benchmark.org/

Green 500 Ranking (June, 2019)

http://www.top500.org/

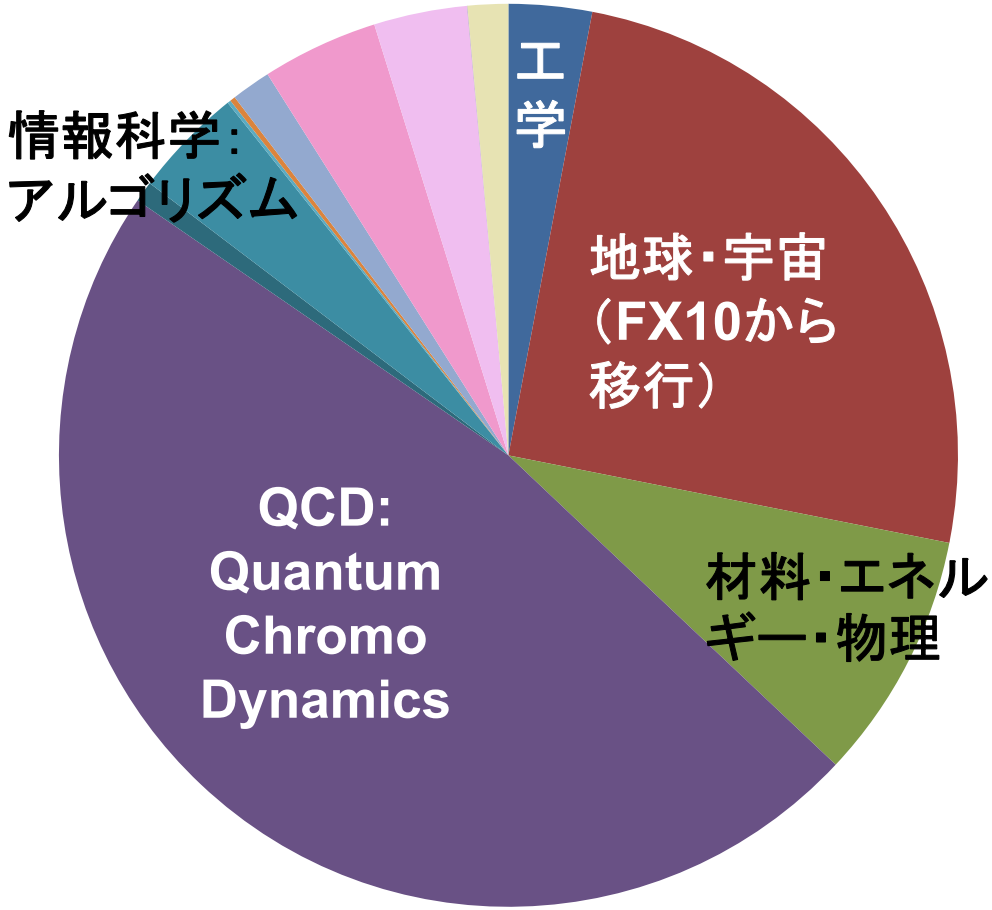
	TOP 500 Rank	System	Cores	HPL Rmax (Pflop/s)	Power (MW)	GFLOPS/W
1	472	Shoubu system B, Japan	953,280	1,063	60	17.604
2	470	DGX SaturnV Volta, USA	22,440	1,070	97	15.113
3	1	Summit, USA	2,414,592	148,600	10,096	14.719
4	8	ABCI, Japan	391,680	19,880	1,649	14.423
5	394	MareNostrum P9 CTE, Spain	18,360	1,145	81	14.131
6	25	TSUBAME 3.0, Japan	135,828	8,125	792	13.704
7	444	PANGEA III, France	291,024	17,860	1,367	13.065
8	2	Sierra, USA	1,572,480	94,640	7,438	12.723
9	43	Advanced Computing System (PreE), China	163,840	4,325	380	11.382
10	23	Taiwania 2, Taiwan	170,352	900	798	11.285
13	June'18	Reedbush-L, U.Tokyo, Japan	16,640	806	79	10.167
19		Reedbush-H, U.Tokyo, Japan	17,760	802	94	8.576

研究分野別利用CPU時間割合 Oakforest-PACS，2017年度



- 工学・ものつくり
- 地球科学・宇宙科学
- 材料科学・エネルギー・物理
- QCD：素粒子物理関連
- 情報科学
- 教育
- 産業利用
- 生体力学
- 社会科学・経済学
- データ科学（データ同化）

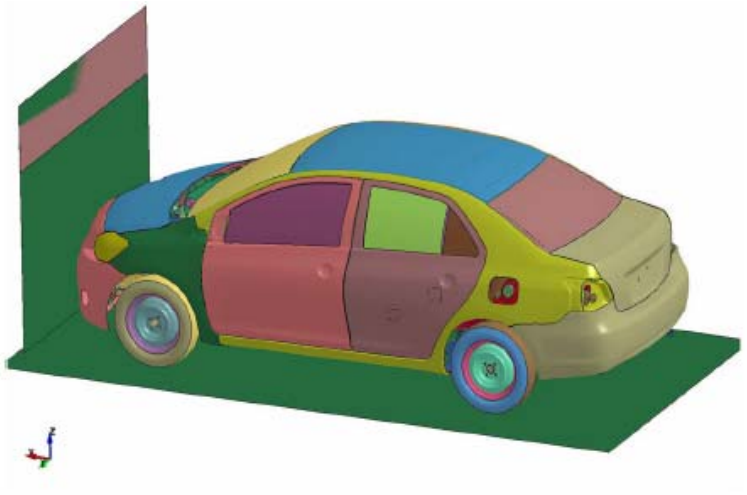
研究分野別利用CPU時間割合
Oakforest-PACS，2018年度
FX10（Oakleaf/Oakbridge-FX）は2018年3月運用停止

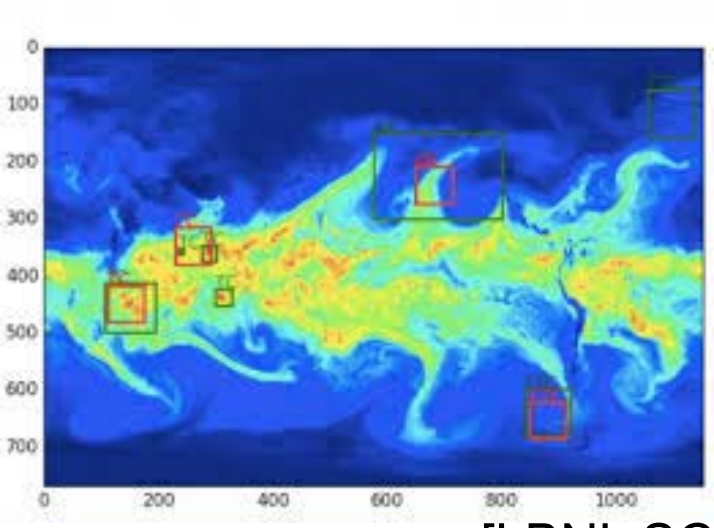


- 工学・ものつくり
- 地球科学・宇宙科学
- 材料科学・エネルギー・物理
- QCD：素粒子物理関連
- 情報科学：システム
- 情報科学：アルゴリズム
- 情報科学：AI
- 教育
- 産業利用
- 生物科学
- バイオインフォティクス
- 社会科学・経済学
- データ科学（データ同化）

OFPにおけるその他のチャレンジ

- 日本自動車工業会（JAMA）
  - 衝撃解析コードLS-DYNAのOFPへの移植を，Intel（米国・フランス），ソフトウェアベンダーと協力で実施
- 超並列環境下での深層学習
  - TensorFlow，Caffe，Chainer等は利用可能





[LBNL SC17]

3システム：利用者2,000+，学外50+%

- 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
  - 3.36 PF, 2016年7月～ 2021年3月末（予定）
  - 東大ITC初のGPUシステム (2017年3月より), DDN IME (Burst Buffer)

- Oakforest-PACS (OFP) (富士通, Intel Xeon Phi (KNL))
  - JCAHPC (筑波大CCS＆東大ITC)
  - 25 PF, TOP 500で6位 (2016年11月) (日本1位)（初登場時）
  - Omni-Path アーキテクチャ, DDN IME (Burst Buffer)

- Oakbridge-CX (富士通, Intel Xeon Platinum 8280)
  - 大規模超並列スーパーコンピュータシステム
  - 6.61 PF, 2019年7月 ～ 2023年6月
  - 全1,368ノードの内128ノードにSSDを搭載

Oakbridge-II（現Oakbridge-CX）

- GPU等のアクセラレータを使わないクラスター
  - Intel Xeon/Skylake or Cascade Lake
  - AMD EPYC, IBM P9, ARM
- 当初計画
  - (5+ PFLOPS), 350+ TB/sec, 250+ TiB
  - Fast Cache: SSD (e.g. Intel Optane): 500+GB/node, 0.5% of memory BW
    - Staging, Check-Pointing, Data Intensive App., External Memory
  - 100+ Gb, Full Bi-Section
  - Storage: 10+ PB, 150+ GB/sec
  - Approx. 1MVA

Oakbridge-CX (OBCX)

計算ノード

Chassis: PRIEMRGY CX400 M1 x342 <4node / Chassis>
Node: PRIMERGY CX2550 M5 x1,240, CX2560 M5 x128

x1,368 node

全体性能

理論演算性能:	6.61PF
主記憶容量:	256.5TiB
メモリバンド幅:	385.1TB/s
ラック数:	21ラック
SSD搭載:	128ノード

ノード単体

理論演算性能: 4.8384 TF
手記憶容量: 192GiB
メモリバンド幅: 281.6GB/s

計算ノード間ネットワーク (Omni-Path Architecture)
通信性能 100Gbps

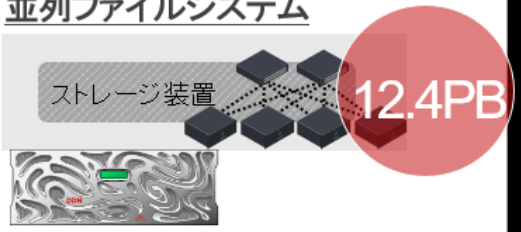
ログインノード

FUJITSU Server
PRIMERGY CX2560 M5 x 10

管理サーバ群

FUJITSU Server
PRIMERGY RX2530 M4 x 15
(ジョブ、運用、認証、Web、
セキュリティログ保存)

並列ファイルシステム

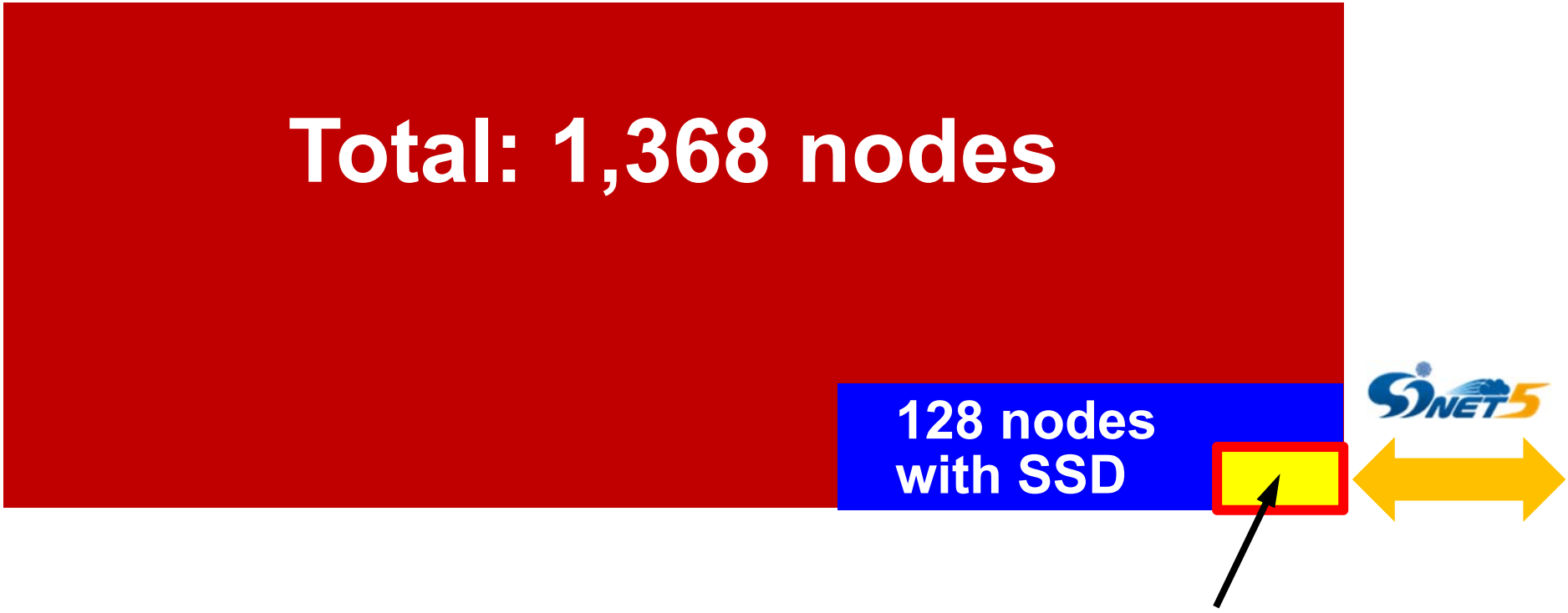


ストレージ装置:DDN ES18KE x2セット
ファイルシステム:DDN ExaScaler
(Lustreベースファイルシステム)

Oakbridge-CX
2019年7月1日運用開始

- Intel Xeon Platinum 8280 (Cascade Lake, CLX), 富士通
  - 1,368 nodes, 6.61 PF peak, 385.1 TB/sec,
  - **4.2+ PF for HPL #45 in 53rd Top500**
  - Fast Cache: SSD's for 128 nodes: Intel SSD, BeeGFS
    - 1.6 TB/node, 3.20/1.32 GB/s/node for R/W
    - Staging, Check-Pointing, Data Intensive Application
    - **128ノードのうち16ノードは外部計算機資源（サーバー，ストレージ，センサーネットワーク等）に直接接続可能（SINET経由）**
- Network: Intel Omni-Path, 100 Gbps, Full Bi-Section
- Storage: DDN EXAScaler (Lustre)
  - 12.4 PB, 193.9 GB/sec
- Power Consumption:
  - 950.5 kVA

Oakbridge-CX
2019年7月1日運用開始



128ノードのうち16ノードは外部計算機資源（サーバー，ストレージ，センサーネットワーク等）に直接接続可能（SINET経由）

全体構成

項 目		仕 様
総理論演算性能		6.61 PFLOPS
総ノード数		1,368=1,240+112+16
総主記憶容量		256.5 TiB
ネットワークトポロジー		Full-bisection Fat Tree
並列ファイルシステム	システム名	Lustreファイル システム
	サーバ(OSS)	DDN ES18K
	サーバ(OSS)数	8
	ストレージ容量	12.4 PB
	ストレージデータ 転送速度	193.9 GB/s

ノードの構成

項 目			仕 様	
製品名			FujitsuPRIMERGY CX2550 M5	Fujitsu PRIMERGY CX2560 M5
ノード数			1240	112+16
CPU	プロセッサ名		Intel® Xeon® Platinum 8280 （開発コード名：CascadeLake）	
	プロセッサ数(コア数)		2 (28+28)	
	周波数		2.7 GHz	
	理論演算性能		4.8384 TFLOPS	
Memory			192 GiB(DDR4)	
インターコネクト			Intel ® Omni-Path ネットワーク (100 Gbps)	
SSD		容量	-	1.6 TB(NVMe)
		読み出し性能		3.20 GB/s
		書き込み性能		1.32 GB/s

ソフトウェア構成

項 目	構 成
OS	Red Hat Enterprise Linux 7, CentOS 7
コンパイラ	GNU コンパイラ Intel コンパイラ(Fortran77/90/95/2003/2008, C, C++)
メッセージ通信ライブラリ	Intel MPI, Open MPI, Intel Omni-Path Fabric Software
ライブラリ	Intel社製ライブラリ(MKL)(BLAS, CBLAS), その他(LAPACK, ScaLAPACK, SuperLU, SuperLU MT, SuperLU DIST, METIS, MT-METIS, ParMETIS, Scotch, PT-Scotch, PETSc, Trillinos, FFTW, GNU Scientific Library, NetCDF, Parallel netCDF, HDF5, Cmake, Anaconda, Xabclib, ppOpen-HPC, ppOpen-AT, MassiveThreads
アプリケーション	Mpijava, OpenFOAM, ABINIT-MP, PHASE, FrontFlow/blue, FrontISTR、REVOCAP-Coupler, REVOCAP-Refiner, OpenMX, xTAPP, AkaiKKR, MODYLAS, ALPS, feram, GROMACS, BLAST, R packages, Bioconductor, BioPerl, BioRuby, BWA, GATK, SAMtools, Quantum ESPRESSO, Xcrypt, Paraview, VisIt, POV-Ray
フリーソフトウェア	Autoconf, automake, bash, bzip2, cvs, emacs, nndutils, gawk, gdb, make, grep, gnuplot, gzip, less, m4, perl, ruby, sed, ubversion, tar, tcsh, tcl, zsh, FUSE, git 等
コンテナ仮想化	singularity (dockerイメージ利用可)

SSD搭載ノード

- Regularキューのみで使用可能
  - 最大 112ノード
  - Intel SSD: DC-P4600 (NVMe IF)
    - 容量: 1.6 TB
    - 性能： Read: 3.2 GB/s, Write: 1.32 GB/s

- 一時ファイルの置き場所として利用

利用方法

1. 各計算ノードのSSDを個別に使用
2. 単一の並列共有ファイルシステムを構成
   （BeeGFS on Demand: BeeOND）
   - ステージングも可能

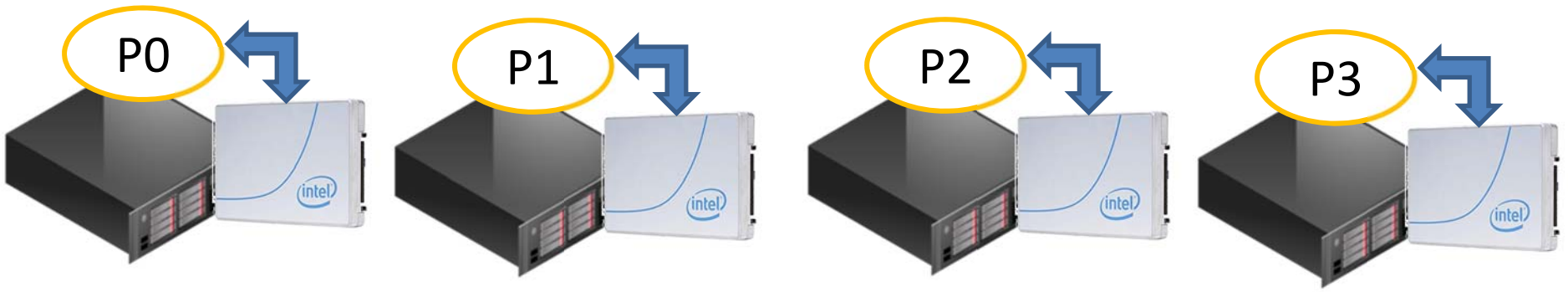
BeeGFS®

SSDノードの指定（個別に使用）

- 環境変数 PJM_SSDで搭載ノード数を指定
- ジョブ中では PJM_SSDDIR 環境変数で割り当てディレクトリ参照

```
$ pjsub -x PJM_SSD=4 run.sh
```

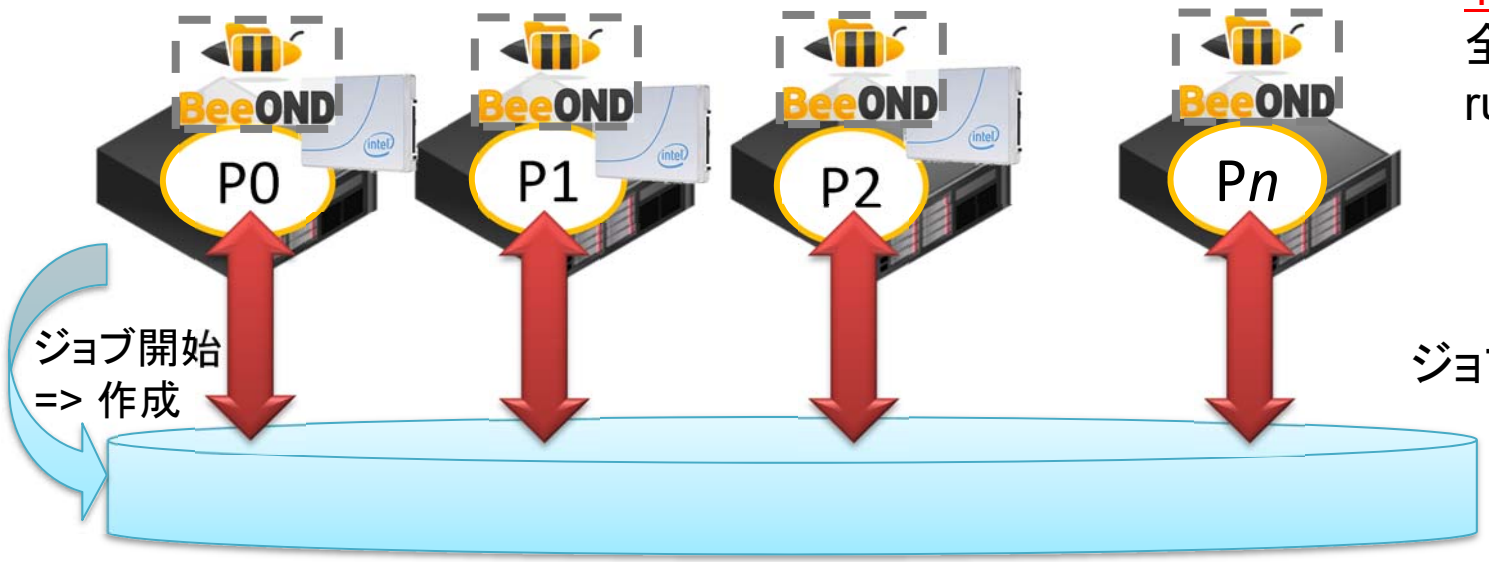
export TMPDIR=$PJM_SSDDIR



SSDノードの指定（BeeOND）

- ジョブ内で単一の並列共有ファイルシステムを動的に構成
- 環境変数 PJM_BEEONDでBeeONDの使用を切り替え
  - 環境変数 PJM_SSDでSSD搭載ノード数を指定、SSD無しノードからも利用できる
  - PJM_BEEONDDIR 環境変数でディレクトリ参照

```
$ pjsub -x PJM_SSD=4 -x PJM_BEEOND=1 run.sh
```

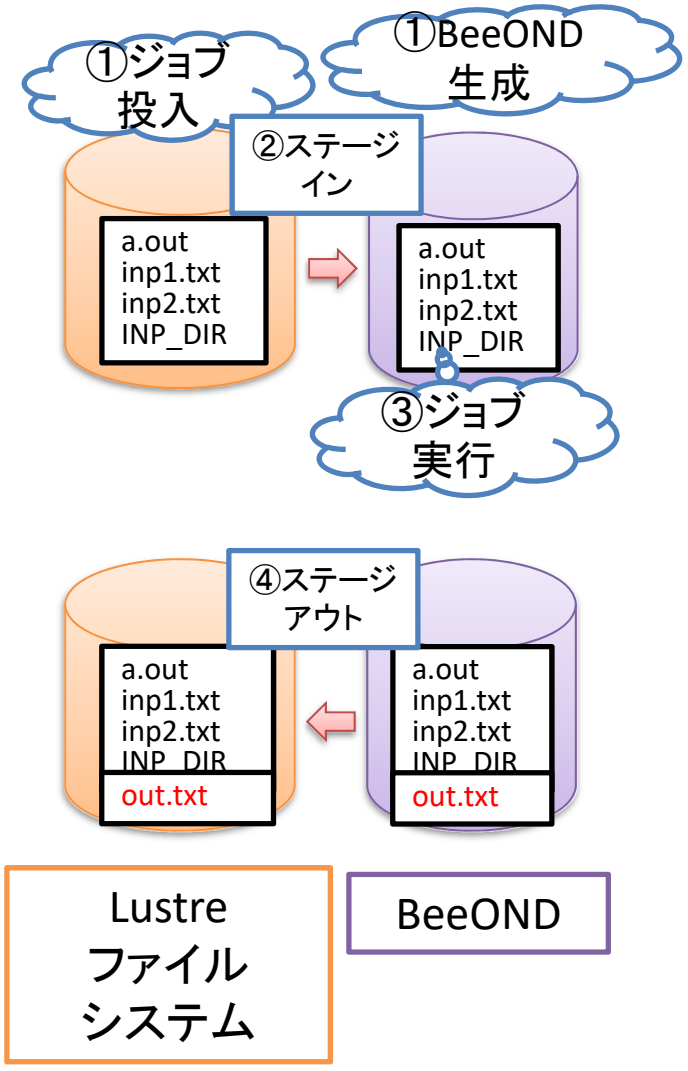


4ノードにSSD搭載
全ノード数は別途
run.shに指定

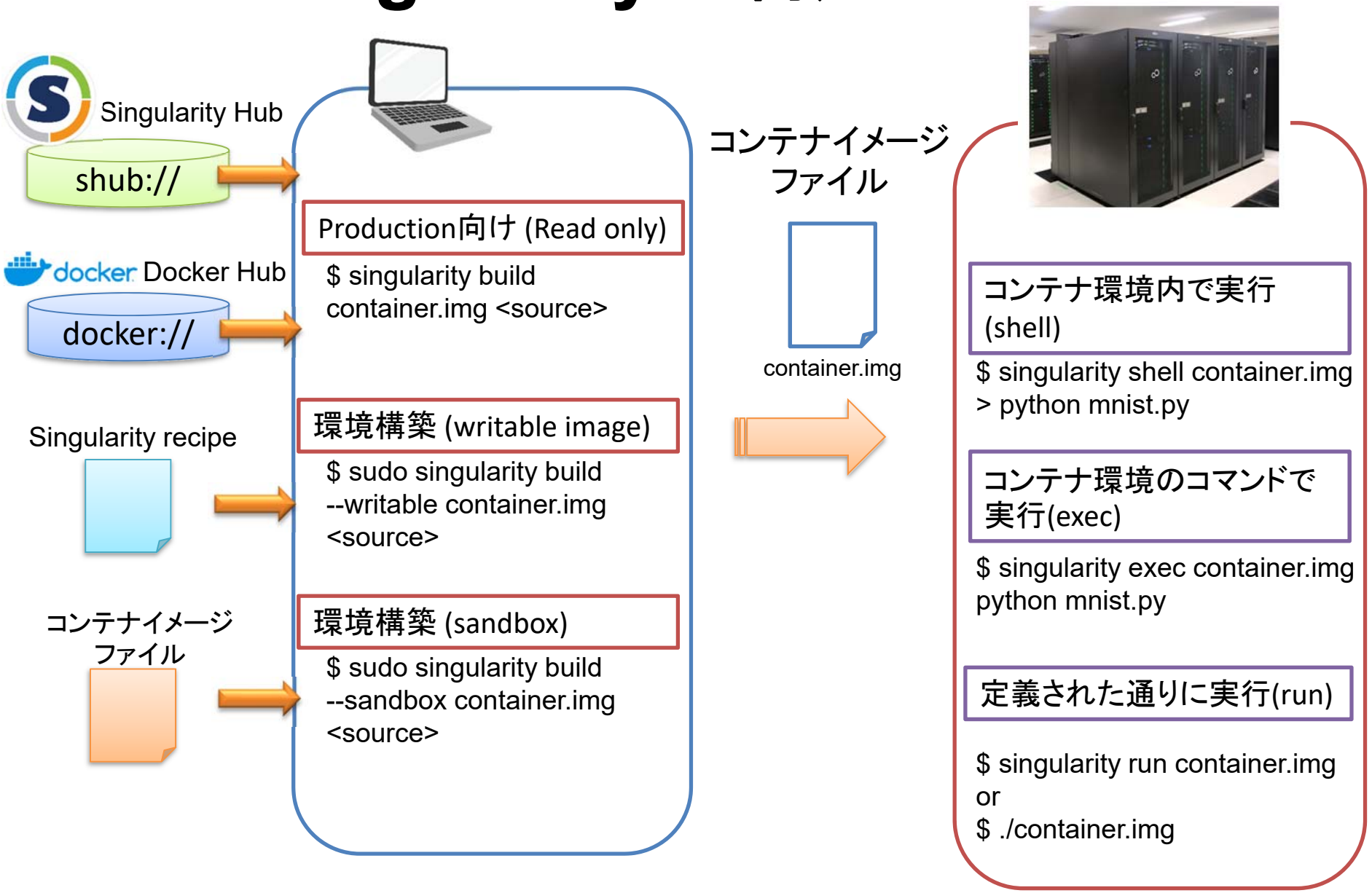
BeeOND+ファイルステージング

ジョブスケジューラと連携、
高速ファイルキャッシュにステージング

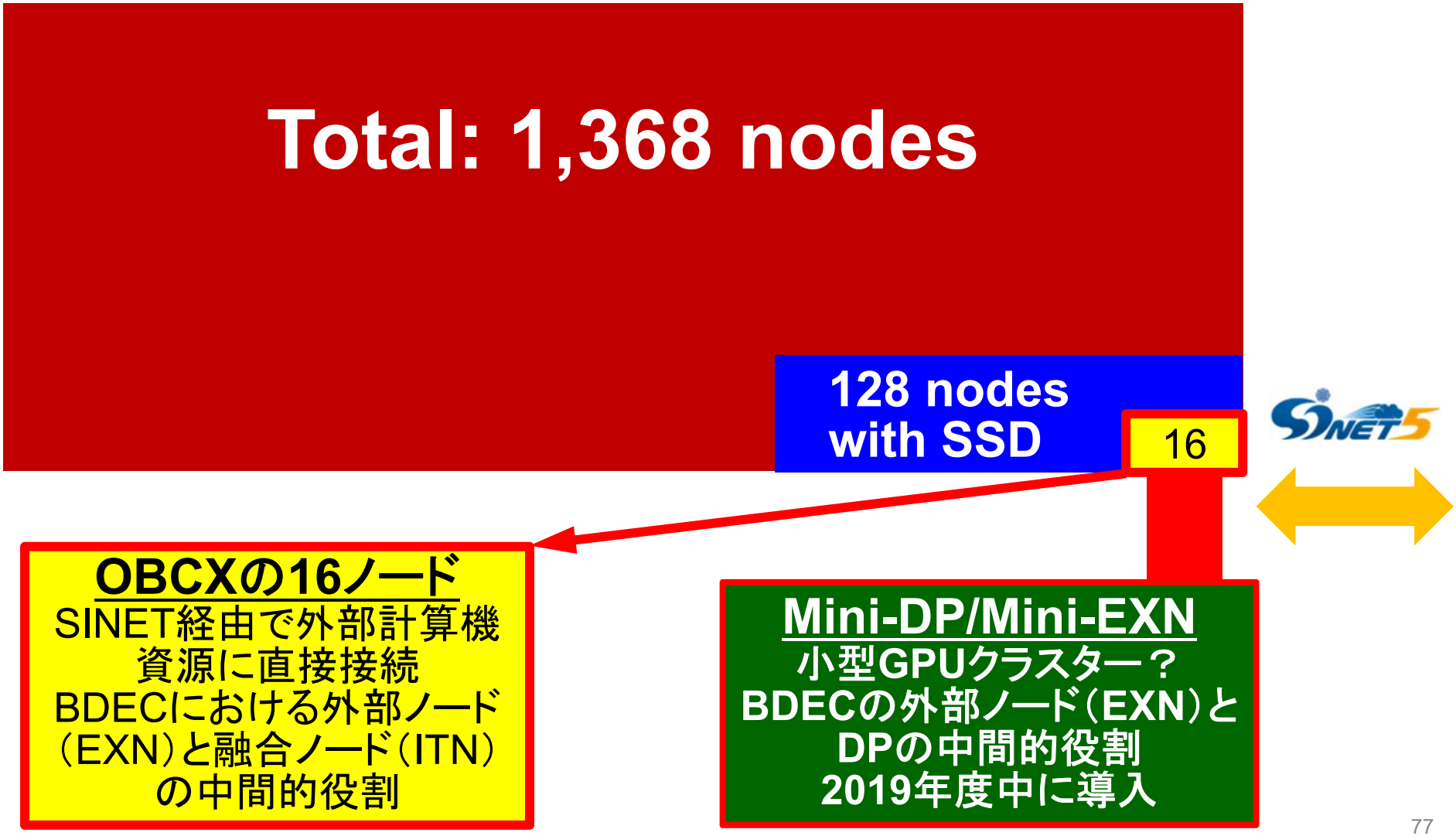
- ステージインリスト、ステージアウトリストをファイルに記述:
  input_file.txt, output_file.txt
  - ⇨ ジョブ投入時 : 並列ファイルシステム
    ファイルキャッシュ システム
  - ⇨ ジョブ終了時: ファイルキャッシュシステム　並列ファイルシステム



Singularityの利用イメージ

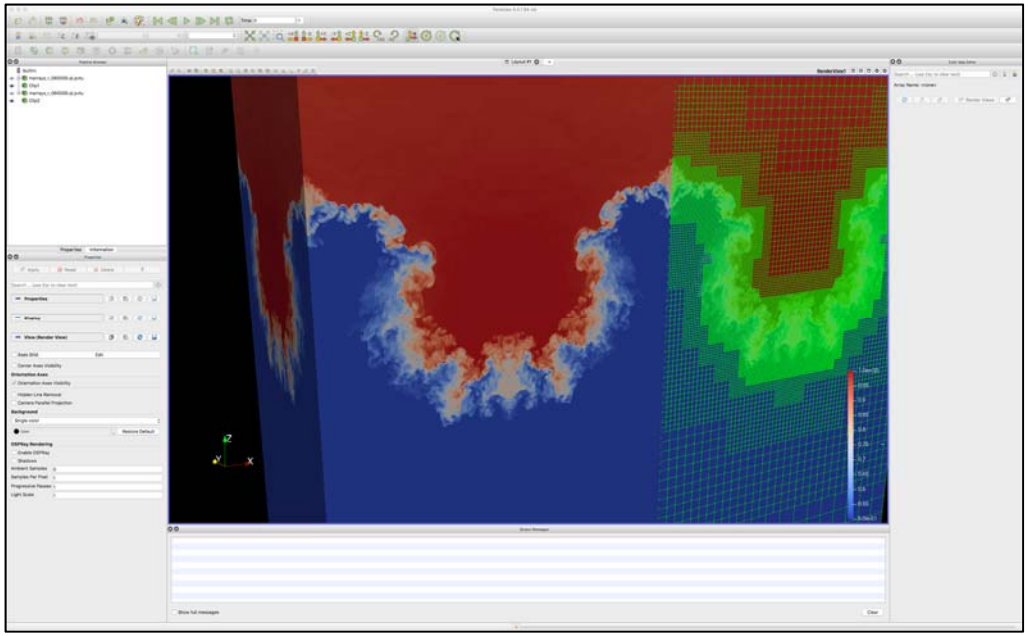


BDEC・DPに向けた実験システム
OBCX+Mini-DP/Mini-EXN



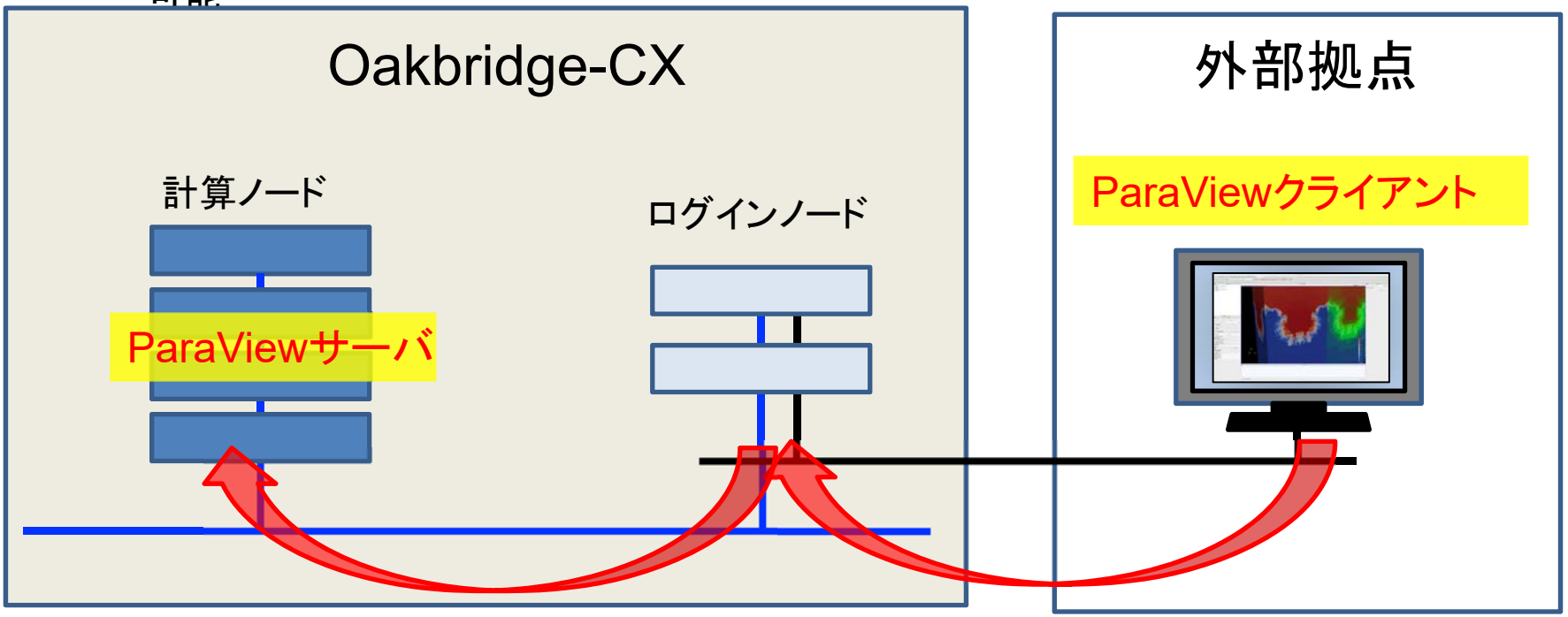
ParaView

- オープンソース、マルチプラットフォームの可視化アプリケーション
- 大容量データの可視化のための分散環境に対応
- クライアント・サーバモデルによるリモート可視化対応
- 様々なデータ形式に対応
- OpenFOAMデータ可視化などに利用される
- https://www.paraview.org



リモートからのParaView利用

- クライアント・サーバモデルによるリモート可視化対応
  - 手元のPCのクライアントからGUI操作によって、Oakbridge-CXのParaViewサーバを制御可能



科学技術計算用アプリケーション

■商用ソフト

・Altair HyperWorks

（総合**CAEプラットフォーム**）

http://altairhyperworks.jp/

HyperWorks, OpenFOAM,
を用いた講習会を実施

■スパコン向けオープンソースコードが充実

・OpenFOAM

(Open source Field Operation And Manipulation)

http://www.openfoam.com/

・ppOpen-HPC

(自動チューニング機構を有するアプリケーション開発・実行環境)

http://ppopenhpc.cc.u-tokyo.ac.jp/ppopenhpc/

・FrontISTR

(大規模並列有限要素解析オープンソースソフト)

http://www.multi.k.u-tokyo.ac.jp/FrontISTR/

・FrontFlow

(乱流燃焼解析ソフトウェア)

http://www.ciss.iis.u-tokyo.ac.jp/dl/

Altair HyperWorksのライセンス

■国内アカデミックユーザ（大学、短大、大学校、高専等に所属の方）

・**ライセンス料：無料**（基盤センター所有のライセンスで利用可能）

・HyperMesh, HyperViewなどクライアントPCで動作するソフトも利用可能

■その他のスパコンユーザ（企業や研究機関に所属の方）

・基盤センター所有のライセンスは利用不可

・**ライセンス個別購入**により利用可能
（ご希望の方はuketsuke@cc.u-tokyo.ac.jpまで問い合わせ下さい）

Oakbridge-CXに関する情報

- 全般
  - https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/
- 利用コース
  - https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/course.php
- ジョブクラス
  - https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/job.php
- 試験運用のお知らせ
  - https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/obcx_test.php
- 利用申込・利用負担金
  - https://www.cc.u-tokyo.ac.jp/guide/application/guideline.php

- 東大情報基盤センターについて
- サービス概要
- スーパーコンピュータシステム概要
- **運用**
- 質疑

Oakbridge-CX利用

- カテゴリー
  - 一般（大学・公共機関），公募（企業（有料），HPCI等（無料））
  - グループ，パーソナル
  - 通常，ノード固定
- 通常利用
  - パーソナル
    - 一般（大学・公共機関）のみ
  - グループ
    - 一般（大学・公共機関
    - 企業（有料・審査あり）
    - その他公募（無料・提案書）
      - 若手・女性支援，HPCI
      - 学際大規模情報基盤共同利用共同研究拠点（JHPCN）

- ノード固定（グループのみ）
  - https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/fixed_node.php
- 若手・女性支援
  - https://www.cc.u-tokyo.ac.jp/guide/young/
- HPCI
  - http://www.hpci-office.jp/
- 学際大規模情報基盤共同利用・共同研究拠点（JHPCN）
  - https://jhpcn-kyoten.itc.u-tokyo.ac.jp/ja/

計算資源配分のポリシー

- 利用するコース (パーソナル・グループコース), 利用申込したノード数に応じて, 計算ノードの利用可能時間である「トークン」を割当てます。
  - 割り当てられたトークン内であれば (一部のコース, サービスを除き) 利用できるノード数制限はなく, 最大利用可能ノード数まで, バッチジョブの実行を可能。
- 1つのシステムに申し込めば他のシステムへトークンを移し, 複数システムを使用可能な場合もある(トークン移行)
  - カテゴリーとしては「通常・一般(大学・公共機関)」の場合(グループ・パーソナル)

トークン（token）

参考：https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/token.php

- トークンは，バッチジョブ実行ごとに消費。（インタラクティブノードを利用したバッチジョブを除く）
  - ノード時間積「経過時間 ×ノード数 × 消費係数」により消費
  - 消費係数は，申込ノード数までは 1.00
    - 申込ノード数を超えた範囲について 2.00
- トークンを使い果たすとジョブ実行不可：1日単位モニター
  - トークン不足の場合，ジョブはsubmitできない
  - 計算資源に余裕がある場合にのみ，トークンを追加可能
- トークンは，利用を許可された有効期間内に全量が利用できることを保証するものではありません。
- 利用を許可された期間のみを有効期間としているため，次年度への繰り越しや返金等は不可。

トークン移行

- 対象システム
  - Oakbridge-CX
  - Oakforest-PACS
  - Reedbush-U/H/L
- 1つのシステムに申し込めば他システムへトークンを移し，複数システム利用できる（換算表は次頁）
  - 通常・一般（大学・公共機関，グループコース，パーソナル）のみ
  - 公募型（企業，若手・女性，HPCI，JHPCN等），ノード固定（Oakbridge-CX, Reedbush）は不可
  - 「通常かつ一般」の場合「以外」はReedbush-U，Reedbush-H，Reedbush-L も別々に申し込む必要がある。

トークン移行

**1つのシステムに申し込めば他システムへトークンを移し，
複数システム利用可能（企業利用，公募利用を除く）**

現在ご利用のシステム	トークン移行先のシステム		
	Reedbush	Oakforest-PACS	Oakbridge-CX
Reedbush （基準ノード数 4ノード）	－	1.5 6	0.75 3
Oakforest-PACS （基準ノード数 8ノード）	0.6 6	－	0.5 4
Oakbridge-CX （基準ノード数 4ノード）	1.3 6	2.0 8	－

- 移行先に追加されるトークン量（ノード時間）＝移行トークン量×上段の係数
- 移行先の消費係数切替点（ノード）＝移行元の申込ノード数÷基準ノード数×下段の係数

トークン移行（Oakbridge-CX）

	トークン移行
パーソナル（一般）（大学・公共機関等）	○
グループ（通常・一般）（大学・公共機関等）	○
グループ（ノード固定・一般）（大学・公共機関等）	×
グループ（通常・公募）（企業・HPCI等）	×
グループ（ノード固定・公募）（企業・HPCI等）	×

Oakbridge-CX：グループコース（ノード固定）

- グループコース（ノード固定）は，研究グループ等で利用するためのコースで，4ノード単位で申込めます。
  - 申込み分のノードを占有できる
  - パーソナルコース，グループコース（通常）とは別のバッチジョブキューを使います。
  - グループコース（通常）と同じ計算機資源を使用することも可能：その分トークンは減る
    - トークンがなくなった時点で占有ノードも使えなくなります（資源に余裕があればトークン追加購入可能）
  - 審査制，全体の10%程度を上限とする
- 設定のカスタマイズが可能となる
  - 個別のログインノード・ストレージ（データ隔離）等柔軟に対応

試験運用について

- 試験運用期間（予定）
  2019年7月1日（月）～ 9月27日（金）
- 利用負担金：無料（試験運用期間のみ）
- ノード固定の利用開始は10月1日以降
- 利用申込は下記より

https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/application.php

※試験運用期間中は、システムの設定変更等のため、予告なく運用の停止、運用仕様の変更を行う場合があります

※試験運用終了時に実行中および待ちジョブは全て削除されます

バッチジョブキュー

- パーソナルユーザ，グループユーザ（一般）で共通
- interactive はトークンを消費しない

代表キュー名	キュー名	最大ノード数	実行制限時間（経過時間）		ノード当たりメモリ量 (GB)
			試験運用期間	正式サービス	
interactive	interactive_n1	1	15 min	30 min	168
	interactive_n8	2 ～ 8	5 min	10 min	168
debug	debug	1 ～ 16	30 min	30 min	168
short	short	1 ～ 8	4 h	8 h	168
regular	small medium large x-large	1 ～ 16 17 ～ 64 65 ～ 128 129 ～ 256	12 h 12 h 12 h 6 h	48 h 48 h 48 h 24 h	168

パーソナルコース負担金

一般（大学・公共機関）のみ

- 研究者が個人単位で使用
- 下記の基本セットを元に最大3口まで申込可能

コース	利用負担金 (年額，税込)	利用可能 ノード数	トークン	ディスク量 /work
パーソナルコース （申込1口あたり）	100,000円/年	最大256ノード	8,640 ノード時間 （1ノード×360日相当） 消費係数： 4 ノードまで 1.00 4 ノード超過 2.00	4TB

※ 申込口数に関わらず、利用可能ノード数（256）までのジョブが実行可能です。

グループコース負担金(年)

- 研究グループ等で利用するコース

	利用負担金 (年額, 税込)	利用可能 ノード数	割当トークン量(年間) 及び消費係数	ディスク量 /work
一般 申込4ノード以上 (4ノード当り)	大学・公共機関: 400,000円 企業: 480,000円	最大256ノード	34,560 ノード時間 (4ノード×360日相当) 消費係数 : 申込ノードまで 1.00 申込ノード超過 2.00	16TB (グループ当り)
ノード固定 申込4ノード以上 (4ノード当り)	大学・公共機関: 600,000円 企業: 720,000円	最大256ノード	34,560 ノード時間 (4ノード×360日相当) 消費係数 : 申込ノードまで 1.00 申込ノード超過 2.00	16TB (グループ当り)

※ 申込ノード数に関わらず、利用可能ノード数（256）までのジョブが実行可能です。

- 東大情報基盤センターについて
- サービス概要
- スーパーコンピュータシステム概要
- 運用
- **質疑**

ご質問･連絡先

- 東京大学情報システム部情報戦略課研究支援チーム
  - 電話（平日09-12，13-17） 03-5841-2717
  - uketsuke(at)cc.u-tokyo.ac.jp