サイバーメディアセンターの
大規模計算機システムの現状と課題

大阪大学サイバーメディアセンター
応用情報システム研究部門　伊達　進

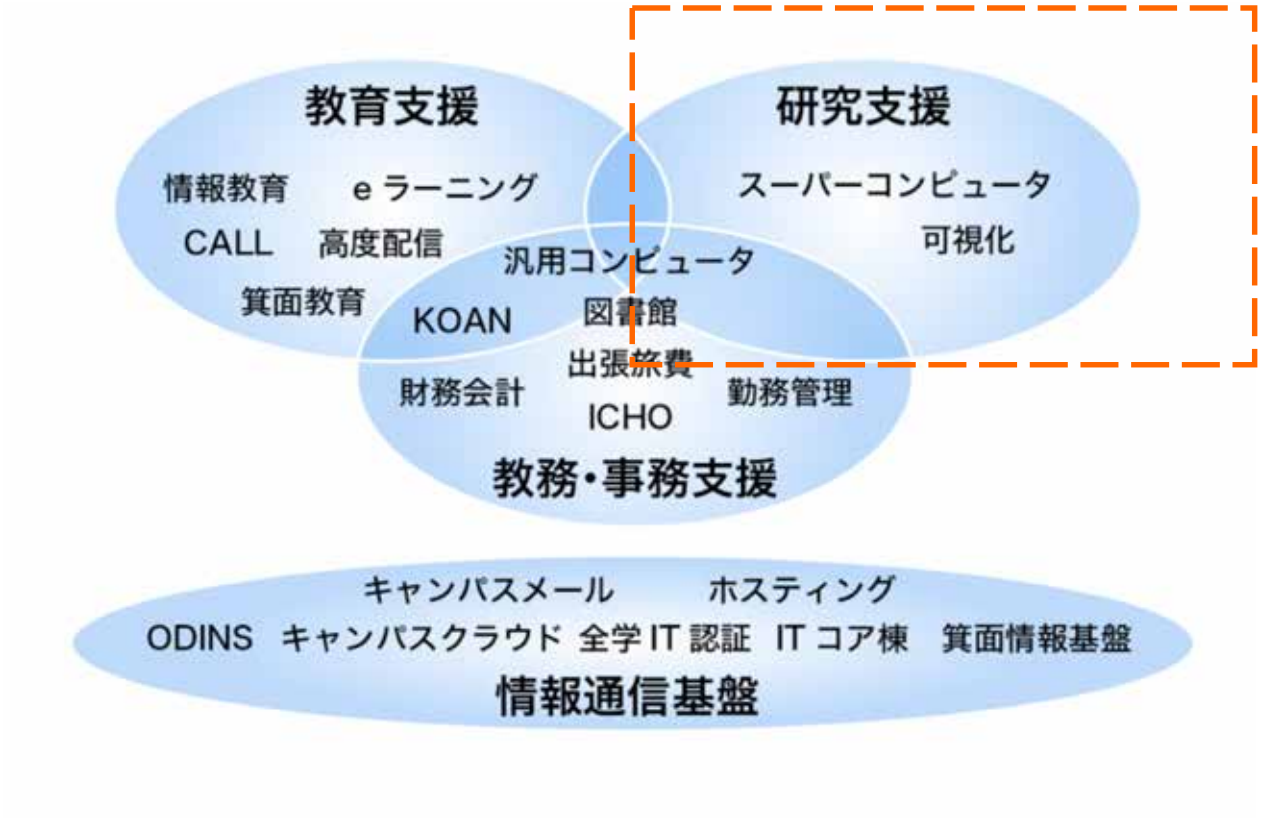
大阪大学サイバーメディアセンターのミッション

吹田CMC 本館

ITコア棟

- 大阪大学の研究・教育を支える情報基盤の整備・運用を担うとともに、大規模計算、情報通信、および、ICT技術を活用した教育に関する最先端の研究開発を推進。
- 学内だけでなく学外の教育・研究組織や産業界と密接に連携したセンターとして機能することが求められた全国共同利用施設でもあり、その一環として、全国の大学の研究者が学術研究・教育に伴う計算及び情報処理を行うことができるよう、種々の高性能な大規模計算機システムを提供。

大規模計算機システムの位置づけ



http://www.cmc.osaka-u.ac.jp/

Cybermedia Centerの大規模計算環境 概要

ベクトル型Supercomputer

- SX-ACE system

スカラ型Supercomputer

- PC cluster system for large-scale visualization (VCC)
- Osaka university Cybermedia CenTer Over-Peta-scale Universal Supercomputer (OCTOPUS)

Large-scale Visualization System

- 24-screen Flat Stereo Visualization System
- 15-screen Cylindrical Stereo Visualization System

SX-ACE

Type: Vector

OS: Super-UX

# of nodes: 1536　(3クラスタ)

# of cores: 6144

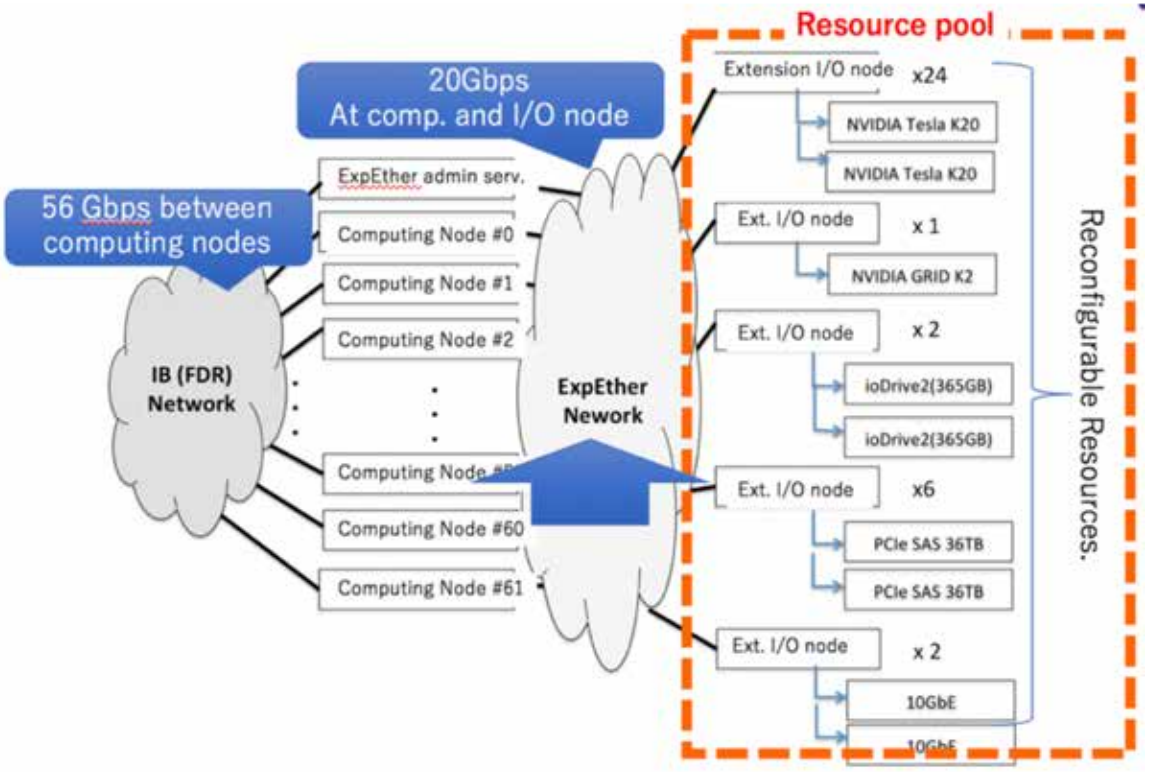
Total memory: 96TB

Peak performance: 423 TFlops

- コアのマルチコア型ベクトルCPU、64GBの主記憶容量を搭載したノード　1536台から構成される”クラスタ化”されたベクトル型スーパーコンピュータ

VCC(大規模可視化対応PCクラスタ)

- システム仮想化技術ExpEtherを応用した再構成可能クラスタシステム



VCC

動的再構成可能型
スーパーコンピュータ

VCC 100.13 TFlops

CPUノード: 66ノード

プロセッサ	Intel Xeon E5-2670v2 (Ivy Bridge / 2.5 GHz 10コア) 2基
主記憶容量	64 GB
インターコネクト	InfiniBand FDR (56 Gbps)

増設ノード: 3ノード

プロセッサ	Intel Xeon E5-2690v4 (Broadwell / 2.5 GHz 14コア) 2基
主記憶容量	64 GB
インターコネクト	InfiniBand FDR (56 Gbps)

再構成可能資源

アクセラレータ	NVIDIA Tesla K20 59基
フラッシュドライブ	ioDrive2 (365 GB) 4個
ストレージ	PCIe SAS (36 TB) 9個

これらの資源を計算ニーズに応じて各ノードに割り付けることができます。

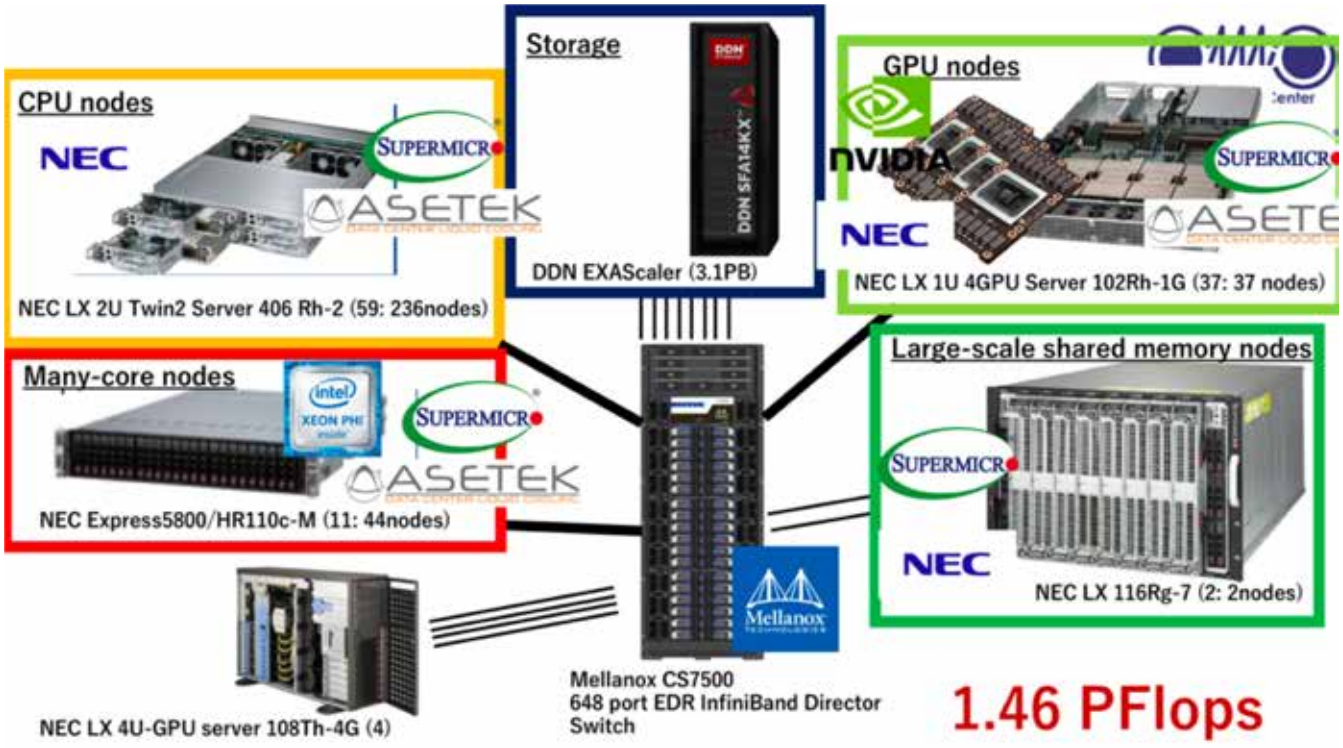
大容量ストレージ

ファイルシステム	NEC ScaTeFS
容量	2 PB

SX-ACEと共通です。

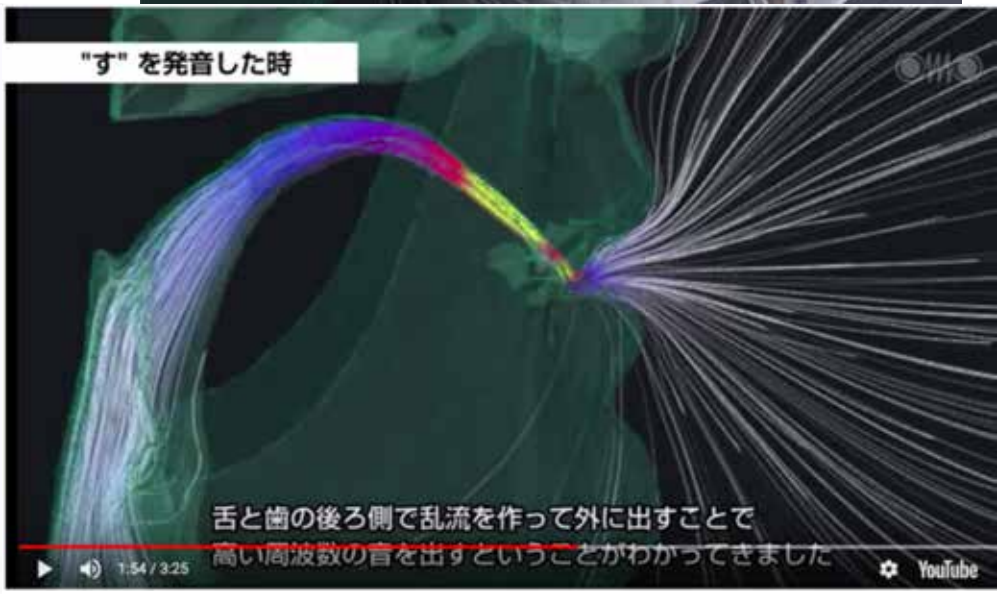
OCTOPUS

- 汎用CPUノード群、GPUノード群、Xeon Phiノード群、大容量主記憶搭載ノード群、大容量ストレージから構成されるハイブリッド型スーパーコンピュータ



Application example from High-performance Scientific News

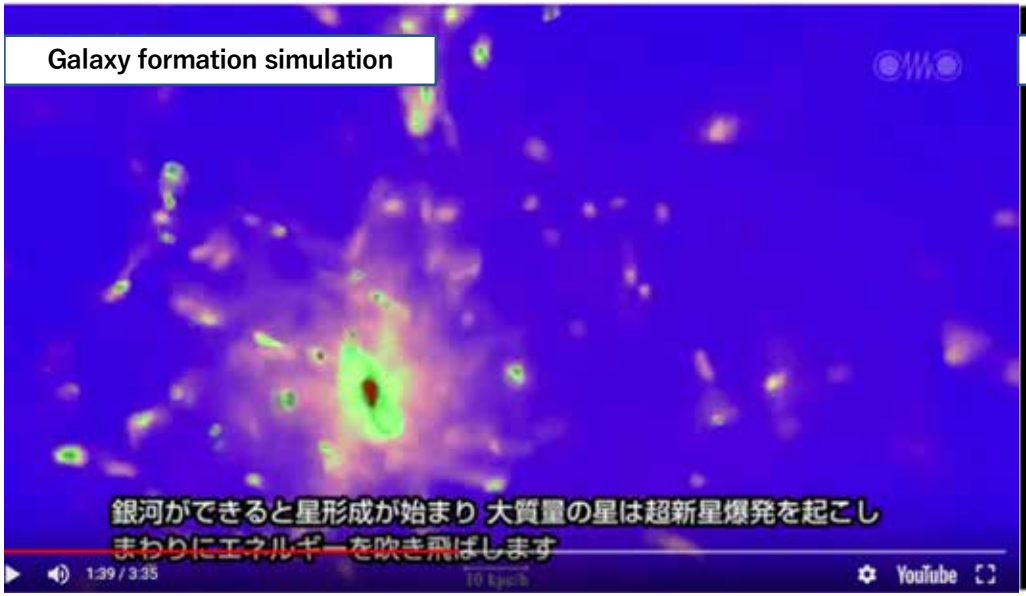
Air flow analysis
on speech production
Dr. Kazunori Nozaki
Osaka University Dental Hospital,
Assistant Professor

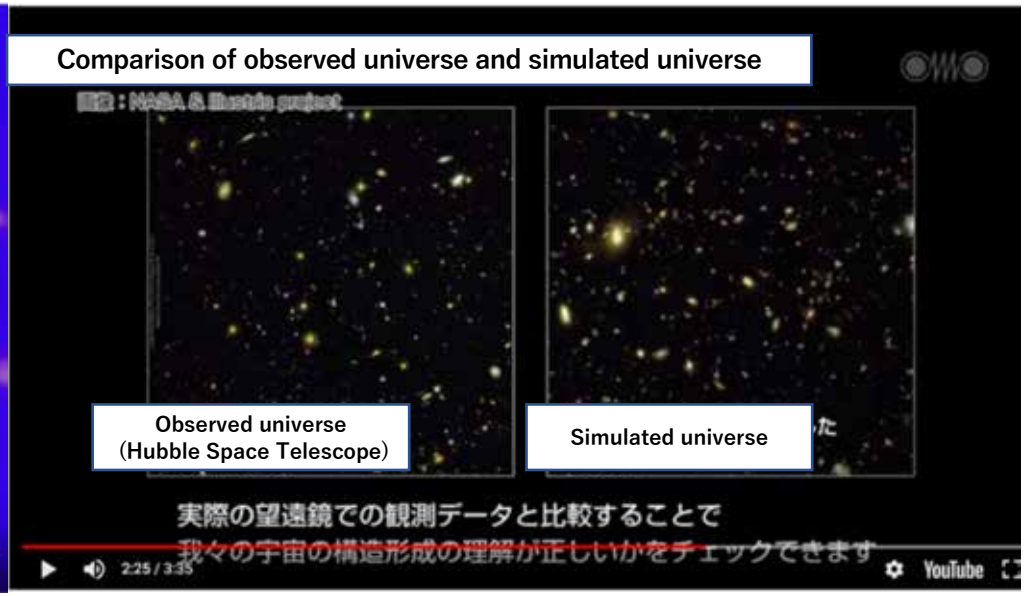


HPSC vol.1 http://www.hpc.cmc.osaka-u.ac.jp/en/hpsc-news/vol1/

Application example from High-performance Scientific News

Cosmology with Numerical Simulations
Prof. Kentaro Nagamine
Graduate School of Science, Osaka Univ.

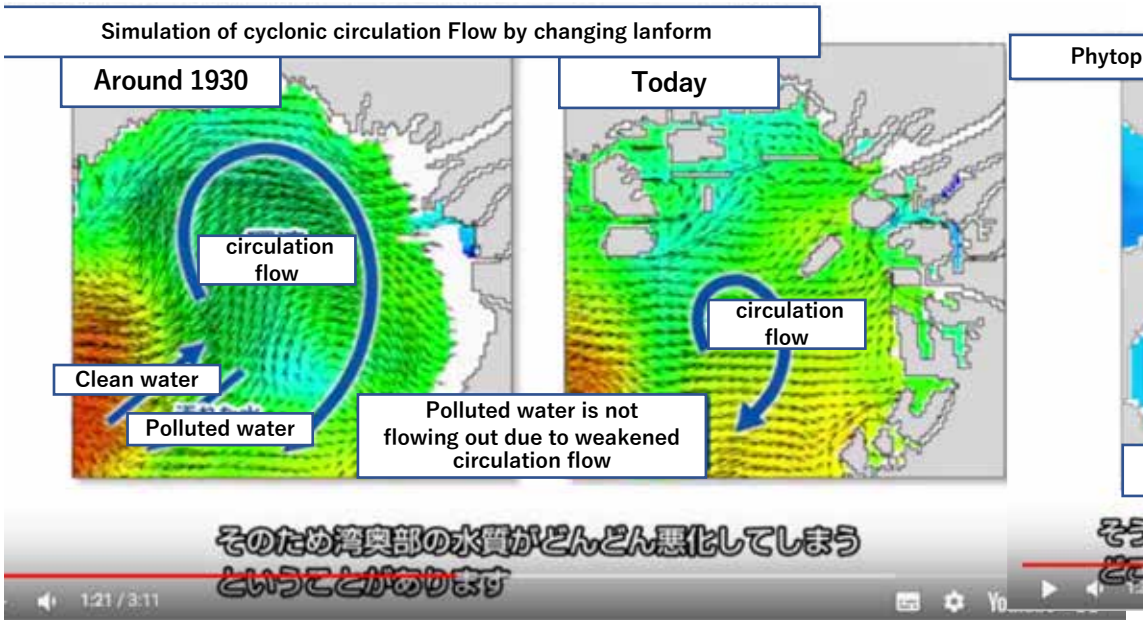


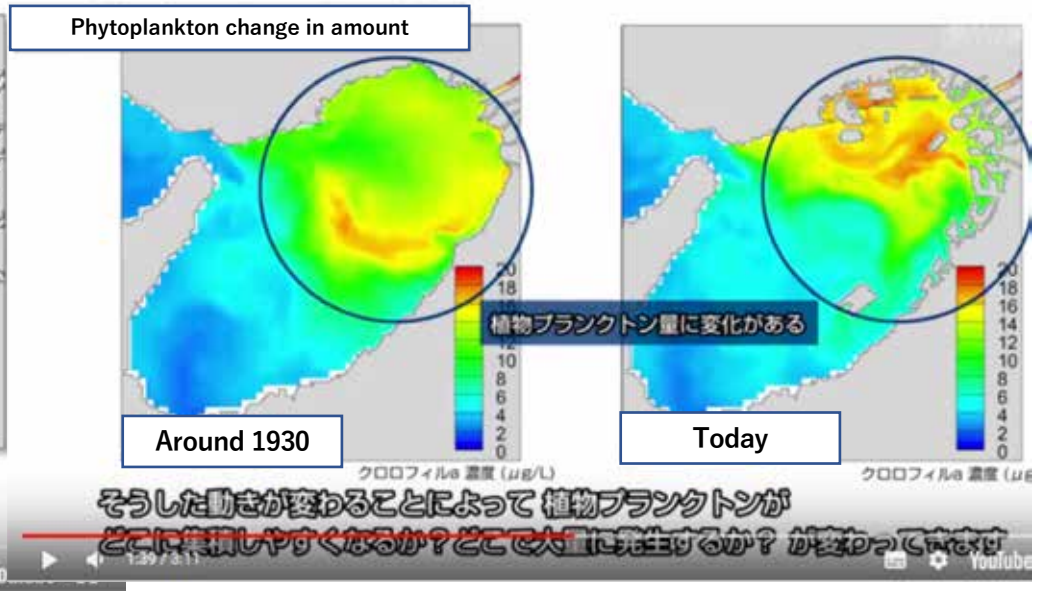


HPSC vol.2 http://www.hpc.cmc.osaka-u.ac.jp/en/hpsc-news/vol02/

Application example from High-performance Scientific News

Numerical simulation of flow and water quality in Osaka Bay
Assit. Prof. Yusuke Nakatani
Graduate School of Engineering, Osaka Univ.





HPSC vol.3 http://www.hpc.cmc.osaka-u.ac.jp/en/hpsc-news/vol03/

大規模計算機システムの利用

1. 一般(学術)利用

- 学術機関を対象とし、利用負担金による利用

2. 産業利用 (成果公開型/成果非公開型)

- 企業を対象とし、利用負担金による利用

3. 公募利用

- 女性・若手支援枠
- 大規模枠
- 特設枠

4. HPCIでの利用

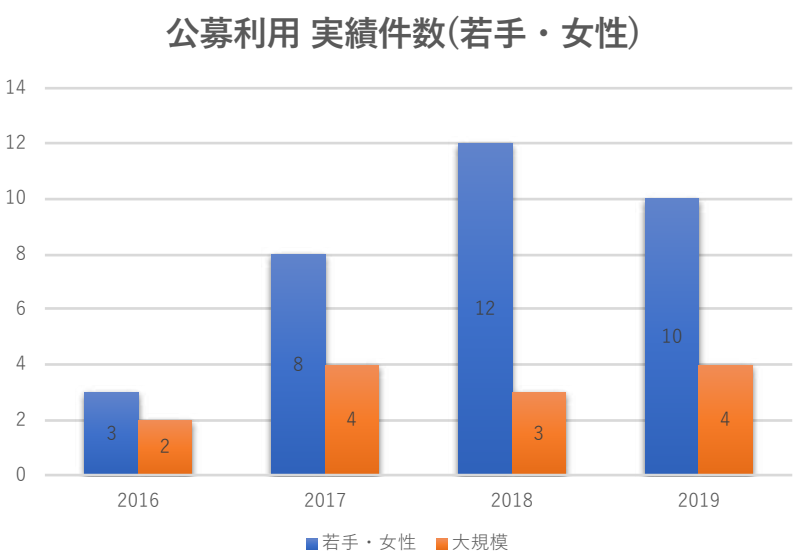
High Performance Computing Infrastructure (https://www.hpci-office.jp/)
の枠組みによる利用

5. JHPCNでの利用

学際大規模情報基盤共同利用・共同研究拠点(http://jhpcn-kyoten.itc.u-tokyo.ac.jp/ja/)の
枠組みによる利用

http://www.hpc.cmc.osaka-u.ac.jp/service/intro/

公募利用制度　実績　(since 2016)



(備考) 2019年度追加募集分は未反映

2018年度 実績例

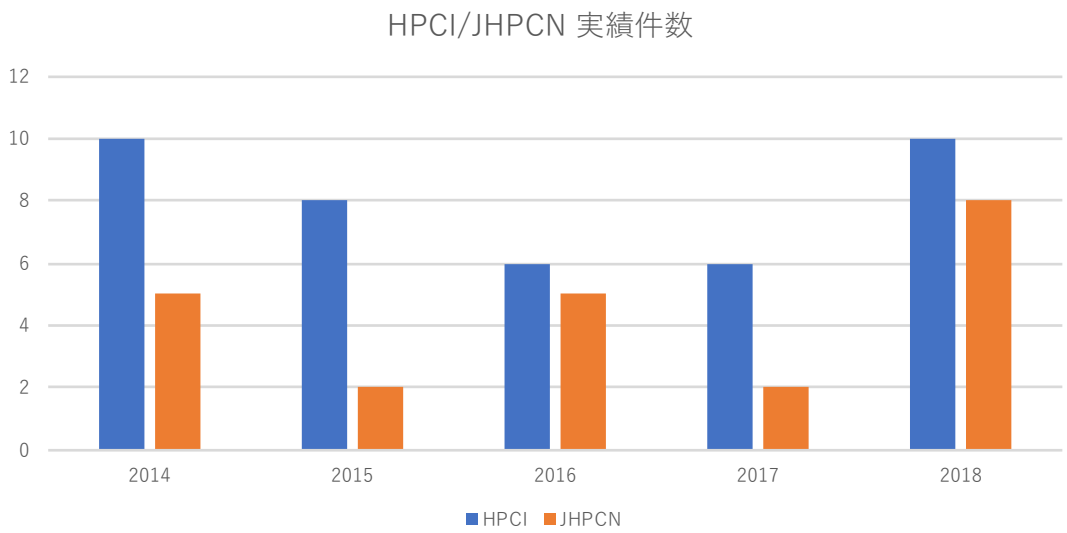
[若手・女性]

- 格子ゲージ理論によるダークマターの研究
- 有限温度・有限密度2カラーQCDの相図と超流動性の解明
- 厳密なZ3対称性を持つ量子色力学による格子計算
- 水素の室温大量貯蔵・輸送を実現する多孔性材料の分子ダイナミクスに基づく解明と先導的デザイン
- キロテスラ級磁場下における超高強度レーザープラズマ相互作用の物理
- 大規模シミュレーションで見る宇宙初期から現在に至る星形成史の変遷
- 格子QCDシミュレーションによる南部-ゴールドストン粒子の質量生成機構の研究
- 日本の全世帯の位置情報付き仮想個票データの統計データからの生成 (追加募集)

[大規模]

- 格子量子色力学を使った高密度物質の研究
- 高強度レーザーによるイオン加速の研究
- 宇宙の大規模構造と銀河形成
- ゴム材料中のナノ粒子構造に対するディープラーニング画像認識モデルの分散学習による高速最適化技術手法の検討 (追加募集)

HPCI/JHPCN 実績



- SX-ACE、OCTOPUS共に堅調に利用されている。

解決したい主な課題

[課題1] 待ち時間の縮小

OCTOPUSはとても安定していて使いやすい。けど、なかなかジョブが実行されない時がある。

年度末は利用負担金を高くするなどの工夫をして、
混雑回避はできないのか？

季節係数の導入

原則：計算機利用に伴う電気代を負担いただく利用負担金制度

(4)OCTOPUSの負担額

(A) 占有

基本負担額	占有ノード数
191,000 円/年	汎用CPUノード群 1 ノード
793,000 円/年	GPUノード群 1 ノード
154,000 円/年	XeonPhiノード群 1 ノード

(B) 共有

コース	基本負担額	OCTOPUSポイント
	10万円	1,000 ポイント
	50万円	5,250 ポイント
	100万円	11,000 ポイント
	300万円	34,500 ポイント
	500万円	60,000 ポイント

(C) ディスク容量追加

基本負担額	提供単位
3,000 円/年	1TB

ノード群	消費係数	季節係数
汎用CPUノード群	0.0520	平成30年度は 通年1で稼動 （季節変動無し） 詳細はこちら
GPUノード群	0.2173	
XeonPhiノード群	0.0418	
大容量主記憶搭載ノード群	0.3703	

- **汎用CPUノードを10ノード並列実行で3時間使用した場合（季節係数：1）**
  10 × 3 × 0.0520 × 1 = 1.560
  →1.56 OCTOPUSポイントが消費されます。

- **GPUノードを10ノード並列実行で3時間使用した場合（季節係数：1）**
  10 × 3 × 0.2173 × 1 = 6.519
  →6.519 OCTOPUSポイントが消費されます。

- **汎用CPUノードを10ノード並列実行で3時間使用した場合（季節係数：0.8）**
  10 × 3 × 0.0520 × 0.8 = 1.248
  →1.248 OCTOPUSポイントが消費されます。

季節係数の導入

Cybermedia Center
Osaka University

季節係数について

季節係数は、前年度の利用率に応じて設定される係数です。
前年度の利用率が低い時期を小さく設定し、通常よりもOCTOPUSポイントの消費量を小さくすることで過疎期の利用率を向上させ、混雑期の待ち時間を緩和させることを目的に導入された係数です。試験運用的な側面もありますため、季節係数の変動がない場合もあります。
0 < 季節係数 ≤ 1 の範囲で設定されるため、消費量が通常よりも割高になることはありません。
例として、季節係数が0.8の時期にジョブを実行すると、そのジョブによって消費するOCTOPUSポイントが本来の消費量の80%に低減されます。
利用者の皆様が計算計画を立てやすいよう、4月の段階で当該年度1年間の季節係数を当ページにて公開いたします。

2019年度の季節係数

	4月 - 6月	7月 - 9月	10月 - 12 月	1月 - 3月
汎用CPUノード群	1.0	1.0	1.0	1.0
GPUノード群	1.0	1.0	1.0	1.0
Xeon Phiノード群	0.5	0.7	1.0	1.0
大容量主記憶搭載ノード群	0.8	0.8	1.0	1.0

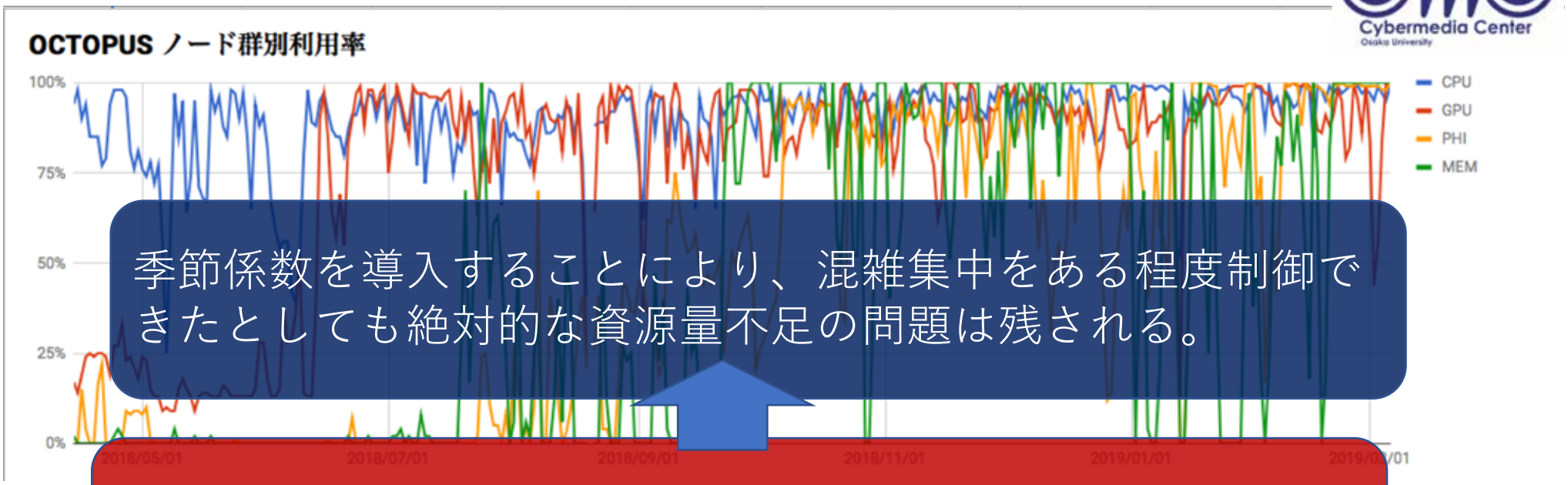
2018年度の季節係数

OCTOPUS稼動初年度に当たる平成30年度は通年"1"での稼動となります。

	4月 - 6月	7月 - 9月	10月 - 12 月	1月 - 3月
汎用CPUノード群	1.0	1.0	1.0	1.0
GPUノード群	1.0	1.0	1.0	1.0
Xeon Phiノード群	1.0	1.0	1.0	1.0
大容量主記憶搭載ノード群	1.0	1.0	1.0	1.0

http://www.hpc.cmc.osaka-u.ac.jp/system/manual/octopus-use/octopus_point/#seasonality

[課題1] 待ち時間の縮小



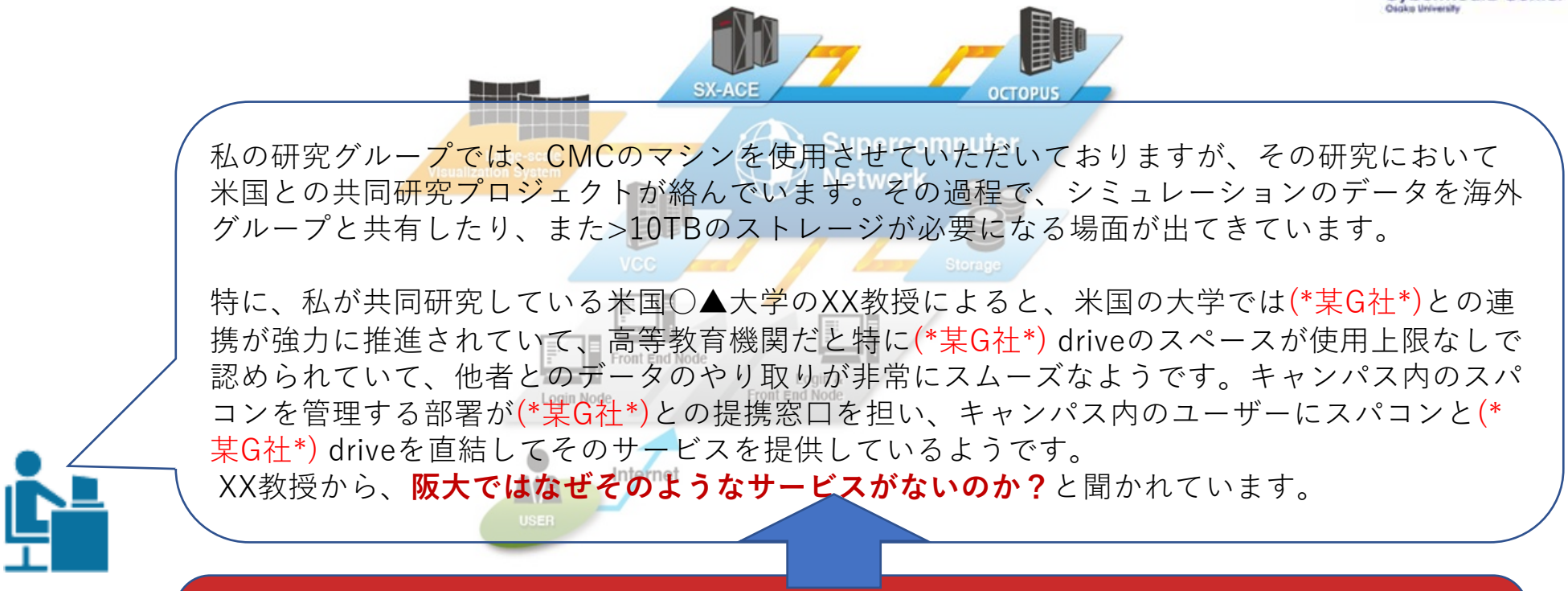
季節係数を導入することにより、混雑集中をある程度制御できたとしても絶対的な資源量不足の問題は残される。

利用者の満足度向上にむけ、限られた予算内で計算資源(on-premise/Cloud)の増強が必要？

OCTOPUSはとても安定していて使いやすい。けど、なかなかジョブが実行されない時がある。

年度末は利用負担金を高くするなどの工夫をして、
混雑回避はできないのか？

[課題2] 海外研究機関・大学とのデータ共有



私の研究グループでは、CMCのマシンを使用させていただいておりますが、その研究において米国との共同研究プロジェクトが絡んでいます。その過程で、シミュレーションのデータを海外グループと共有したり、また>10TBのストレージが必要になる場面が出てきています。

特に、私が共同研究している米国〇▲大学のXX教授によると、米国の大学では(*某G社*)との連携が強力に推進されていて、高等教育機関だと特に(*某G社*) driveのスペースが使用上限なしで認められていて、他者とのデータのやり取りが非常にスムーズなようです。キャンパス内のスパコンを管理する部署が(*某G社*)との提携窓口を担い、キャンパス内のユーザーにスパコンと(*某G社*) driveを直結してそのサービスを提供しているようです。
XX教授から、**阪大ではなぜそのようなサービスがないのか？**と聞かれています。

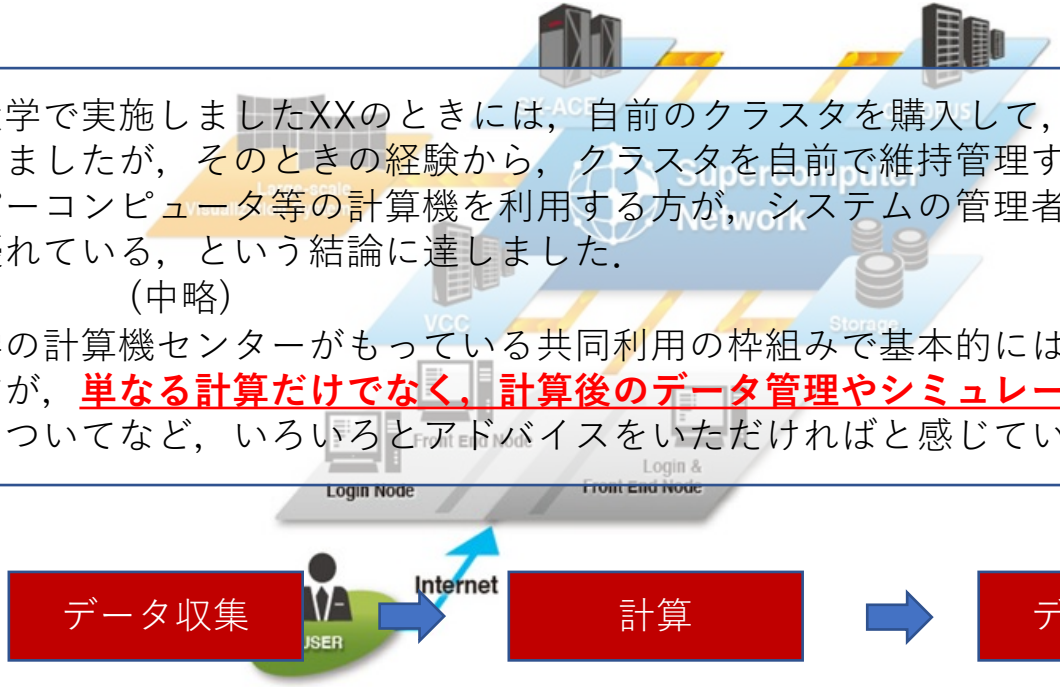
クラウドの(ストレージ)の活用・連携はやはり必要か？

[課題3] 計算前後のデータ利活用を意識した計算環境構築

○▲大学で実施しましたXXのときには，自前のクラスタを購入して，○▲大学にて維持管理をいたしましたが，そのときの経験から，クラスタを自前で維持管理するのではなく，国立大学のスーパーコンピュータ等の計算機を利用する方が，システムの管理者の人件費や組織管理の観点から優れている，という結論に達しました．

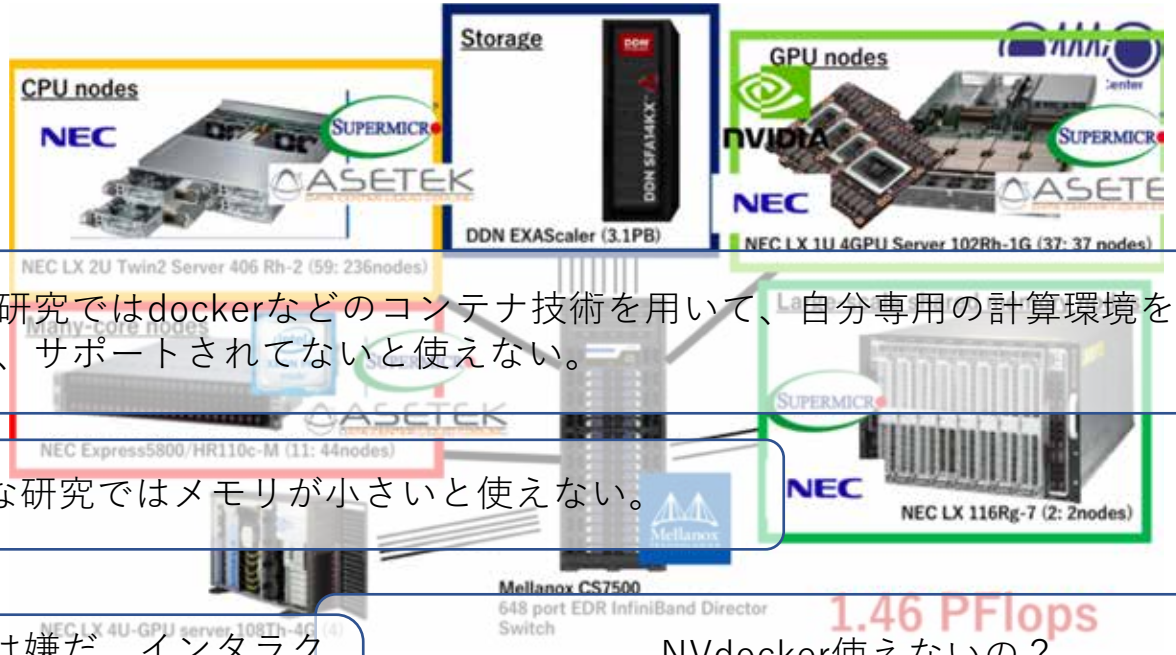
（中略）

各大学の計算機センターがもっている共同利用の枠組みで基本的には実施していきたいと思っていますが，**単なる計算だけでなく，計算後のデータ管理やシミュレーション結果のデータの公開方法**についてなど，いろいろとアドバイスをいただければと感じています．



次のシステムでは、計算だけでなく、研究者の研究活動における計算前後のデータ利活用を意識した計算環境整備が必要？
もうクラウドで全部できちゃう？

[課題4] テーラーメイド計算環境



バイオな研究ではdockerなどのコンテナ技術を用いて、自分専用の計算環境をつくるようになっているので、サポートされてないと使えない。

バイオな研究ではメモリが小さいと使えない。

バッチは嫌だ。インタラクティブに使えないの？
Jupyter Notebookは？

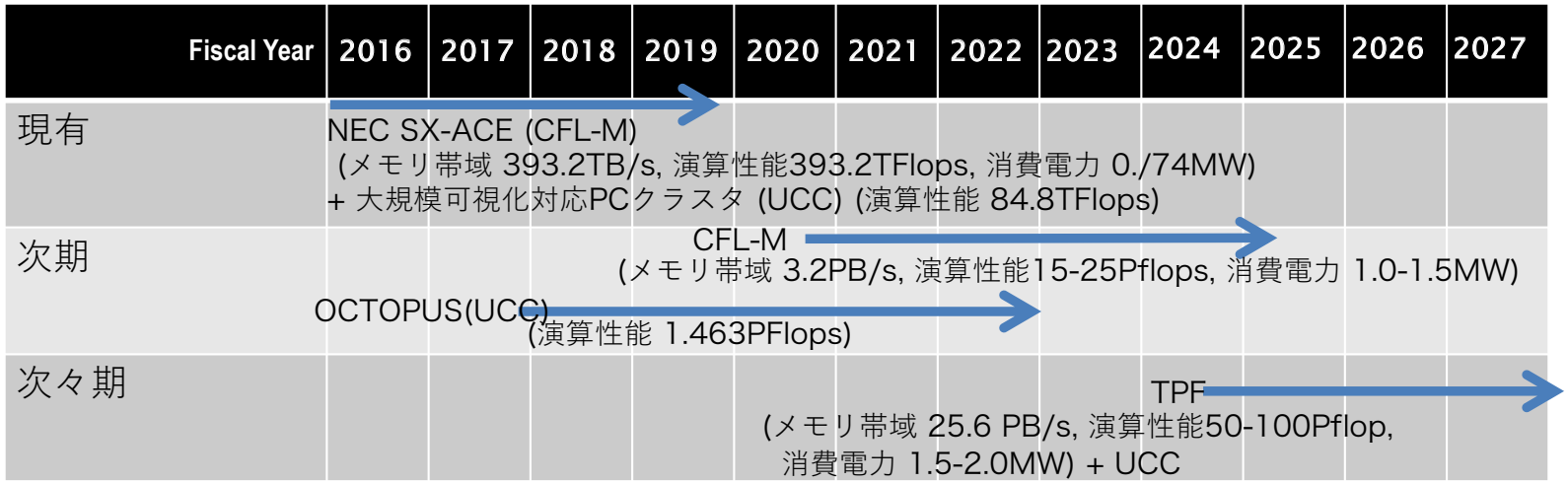
NVdocker使えないの？

従来型のHPCだけでなくHPDAの多様なニーズを収容できる、テーラーメイドな計算環境の構築

大阪大学サイバーメディアセンターの10年計画

システム整備に関する基本方針・戦略

- 高メモリバンド幅を要する計算を主ターゲットとし、更新毎にメモリ帯域を8-10倍に向上
- ベクトル、スカラーをアプリケーションごとに最適化するクラスタ型アーキテクチャの推進
- データ収集から解析結果の大規模可視化までのフロー全体を効率化



産業利用・共同研究の拡大

- 社会貢献、地域貢献の視点から、産業利用支援および共同研究を積極展開
- 新しい計算ニーズの探求

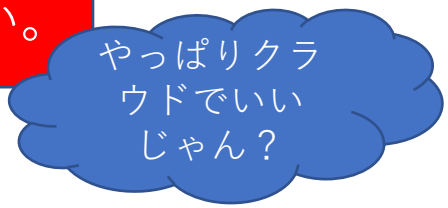
人材育成

- 大規模計算および可視化を駆使してe-Scienceの諸問題に対応できる人材の育成
- ベクトル・スカラ融合に基づく新しいアーキテクチャを先導できる人材の育成

まとめ

- サイバーメディアセンターの大規模計算機システムの現状とともに、クラウドとの連携が解決しうる主たる課題4点について報告した。
  - 課題1: **待ち時間の縮小**
  - 課題2: **海外研究機関・大学とのデータ共有**
  - 課題3: **計算前後のデータ利活用を意識した計算環境構築**
  - 課題4: **テーラーメイド計算環境**

次のスーパーコンピュータシステムにむけて是非解決したい。




Cyber HPC Symposium
大阪大学サイバーメディアセンター サイバーHPCシンポジウム
2015年3月20日(金)
参加費無料 (受付開始 9:00)
大阪大学銀杏会館


CYBER HPC
SYMPOSIUM
2016
大阪大学サイバーメディアセンター サイバーHPCシンポジウム
2016年3月25日(金)
参加費無料
大阪大学


大阪大学 サイバーメディアセンター
サイバーHPCシンポジウム
2017
CYBER HPC SYMPOSIUM
2017年3月16日(木) 参加費無料
大阪大学


CYBER HPC
SYMPOSIUM
CYBERMEDIA CENTER
OSAKA UNIVERSITY
2018
大阪大学 サイバーメディアセンター
サイバーHPCシンポジウム
大阪大学
2018年3月16日(金) 参加費無料