



TOHOKU
UNIVERSITY



Cyberscience
Center

東北大AOBAの 運用開始と将来展望

東北大学サイバーサイエンスセンター
スーパーコンピューティング研究部 教授

滝沢 寛之

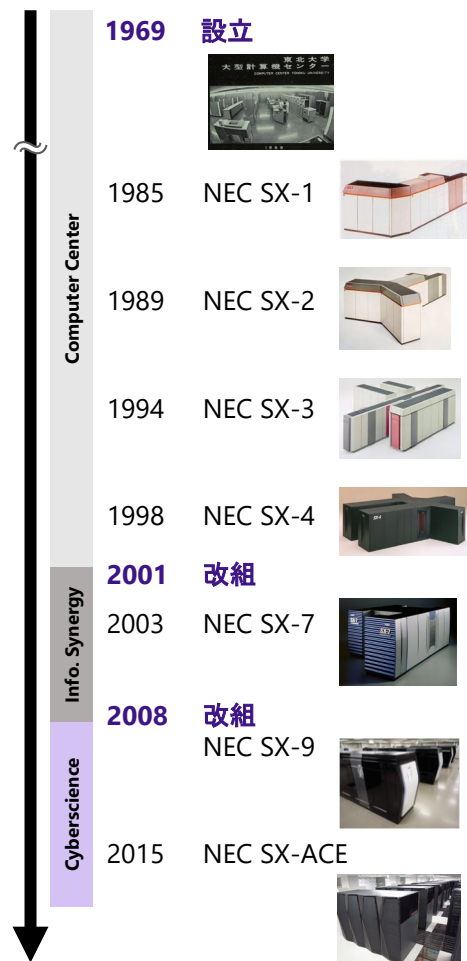
<takizawa@tohoku.ac.jp>

AOBAの運用開始

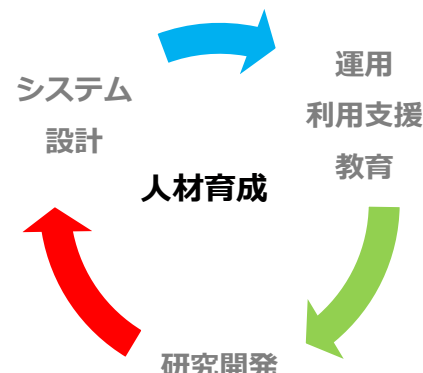
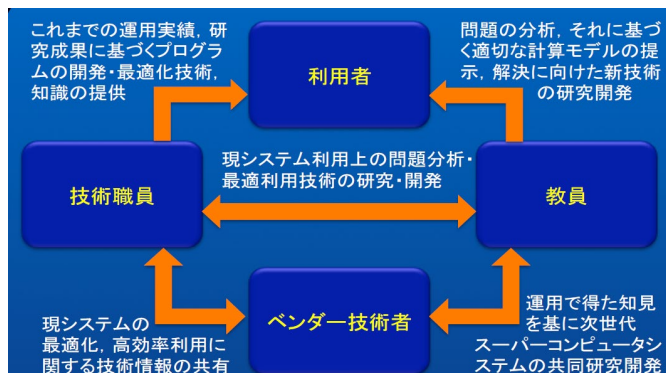
歴史と使命

■ 沿革

■ 使命



- 学術利用のために**最先端の計算環境**を提供
 - NEC SX-ACEやスーパーコンピュータAOBAを運用
 - 利用登録者数 **1,577名** (2021年1月末時点) 2020年10月稼働
- **利用者支援**
 - 性能測定, 性能解析, 高速化支援
 - 高性能計算に関するセミナーや講習会
- **研究開発型**の情報基盤センター
 - 利用者や企業との共同研究
 - 高生産性かつ高性能な計算環境を実現するための次世代高性能計算システムやそのアプリケーションの創出

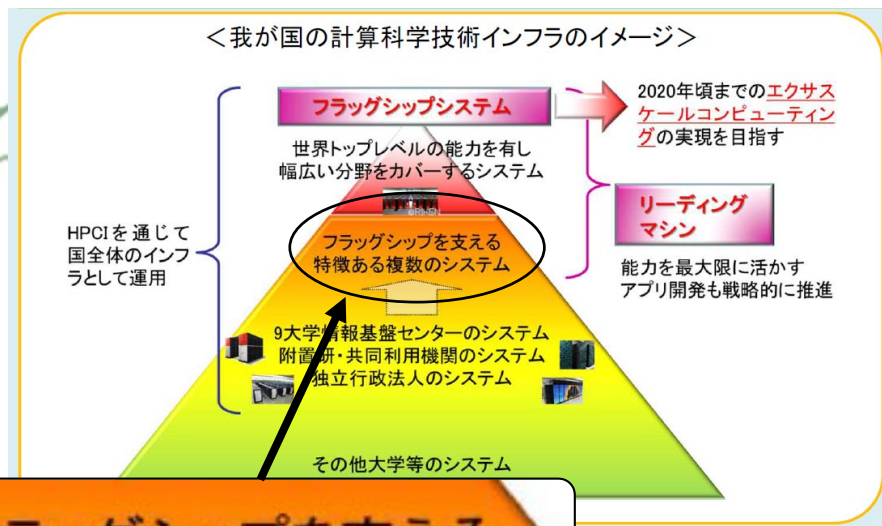


HPCIやJHPCNの拠点として

HPCIシステムを構成する計算機資源を提供するシステム構成機関 (令和元年度)

今後のHPCI計画推進の在り方について (平成26年3月14日)

https://www.mext.go.jp/b_menu/shingi/chousa/shinkou/028/gaiyou/1348991.htm



フラッグシップを支える
特徴ある複数のシステム

サイバーサイエンスセンターが目指す役割

フラッグシップ(富岳)を支える特徴のあるシステムの提供

→ **メモリバンド幅重視**の数値シミュレーション主眼のシステム設計

新システムAOBAの始動

■ 愛称：スーパーコンピュータ **AOBA (あおば)**

- 公募期間：7月1日～31日
- 応募総数：346件
- 選定理由：
 - 新スパコンの愛称としての相応しさ・親しみやすさ
 - 36件の応募があり**他を圧倒する高い支持率** (10.4%)



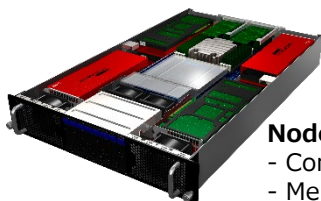
10月1日 運用開始

システム構成

サブシステム AOBA-A



NEC SX-Aurora TSUBASA B401-8 x 72

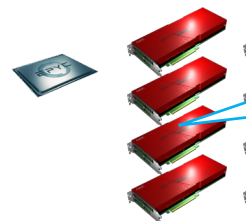


Node Performance

- Comp. : **20.7 TF**
- Mem size : 640 GB
- Mem BW : **12.4 TB/s**

第2世代VE

AMD EPYC 7402P x 1
NEC Vector Engine Type **20B** x 8



Vector Engine Core x 8
HBM2E Module x 6



サブシステム AOBA-B



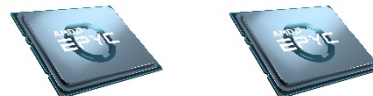
NEC LX 406Rz-2 x 68



Node Performance

- Comp. : **4.1 TF**
- Mem size : 256 GB
- Mem BW : **0.41 TB/s**

AMD EPYC 7702P x 2

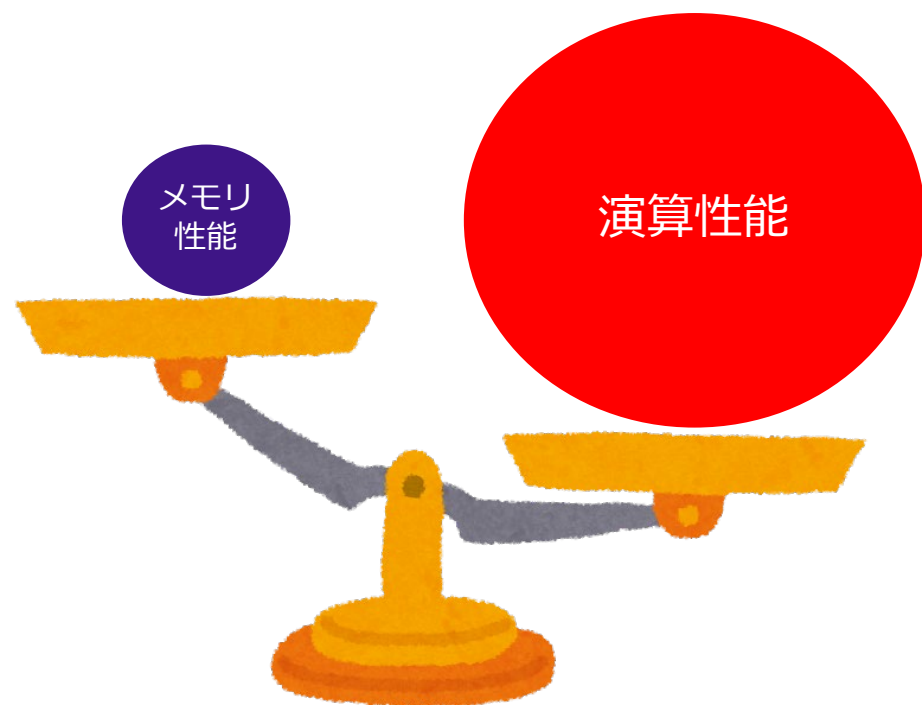


システム全体性能: **1.78 Pflop/s**, **924 TB/s**

真に役立つスパコンを目指して

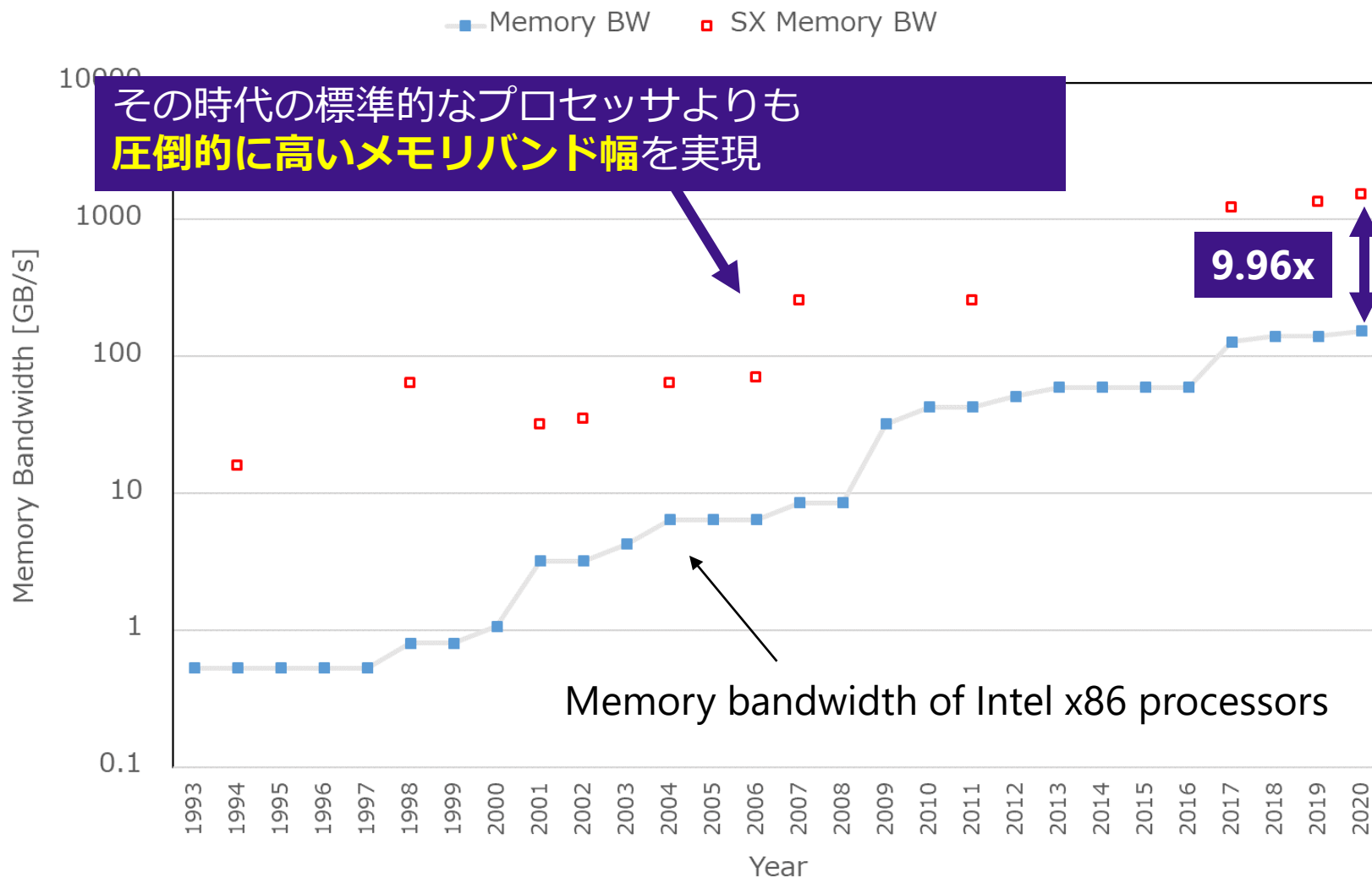
普通のスパコン

バランスのとれたスパコン



↑ 我々はこちらを目指している

SXシリーズの特長



Xeon (DDR4-3200*6) vs SX-Aurora2 (HBM2E*6)

ベクトルエンジン (Type 20B)

ベクトルプロセッサ



高速

2020年6月現在
NEC調べ

世界トップクラスの
コア性能

307GFlops(DP)
614GFlops(SP)



高速

2020年6月現在
NEC調べ

ダントツの
データアクセス性能

1.53TB/s



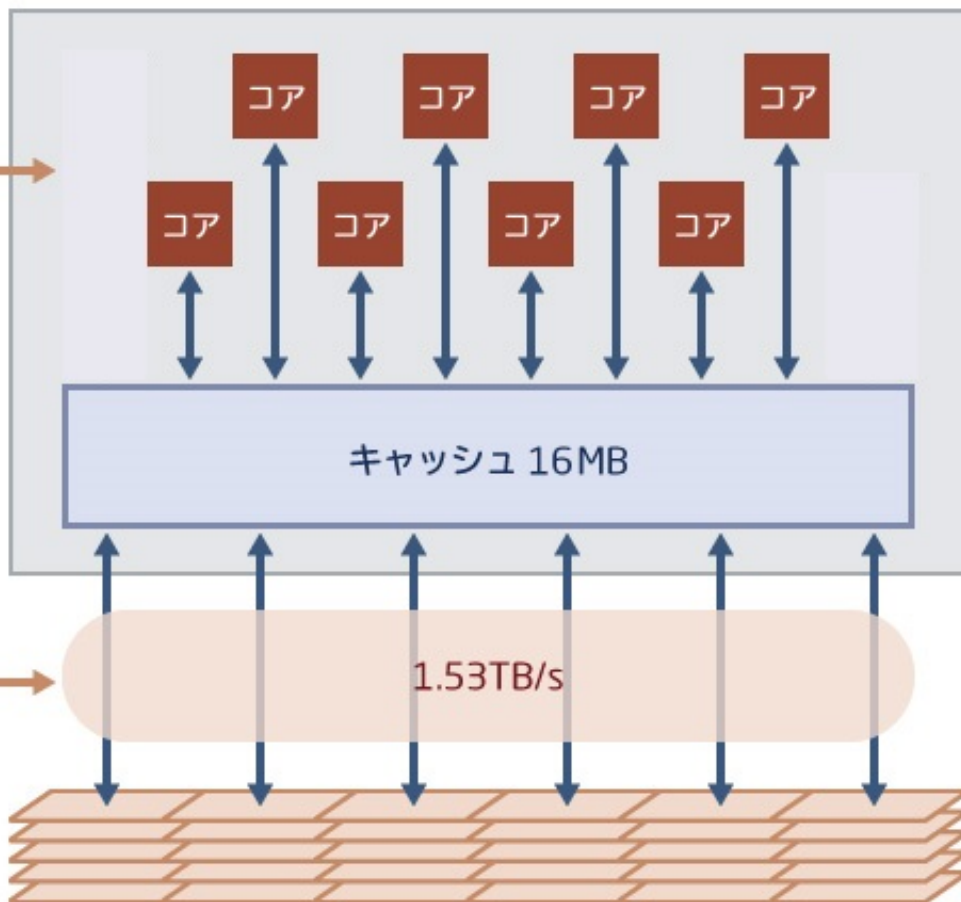
世界初

テクノロジー

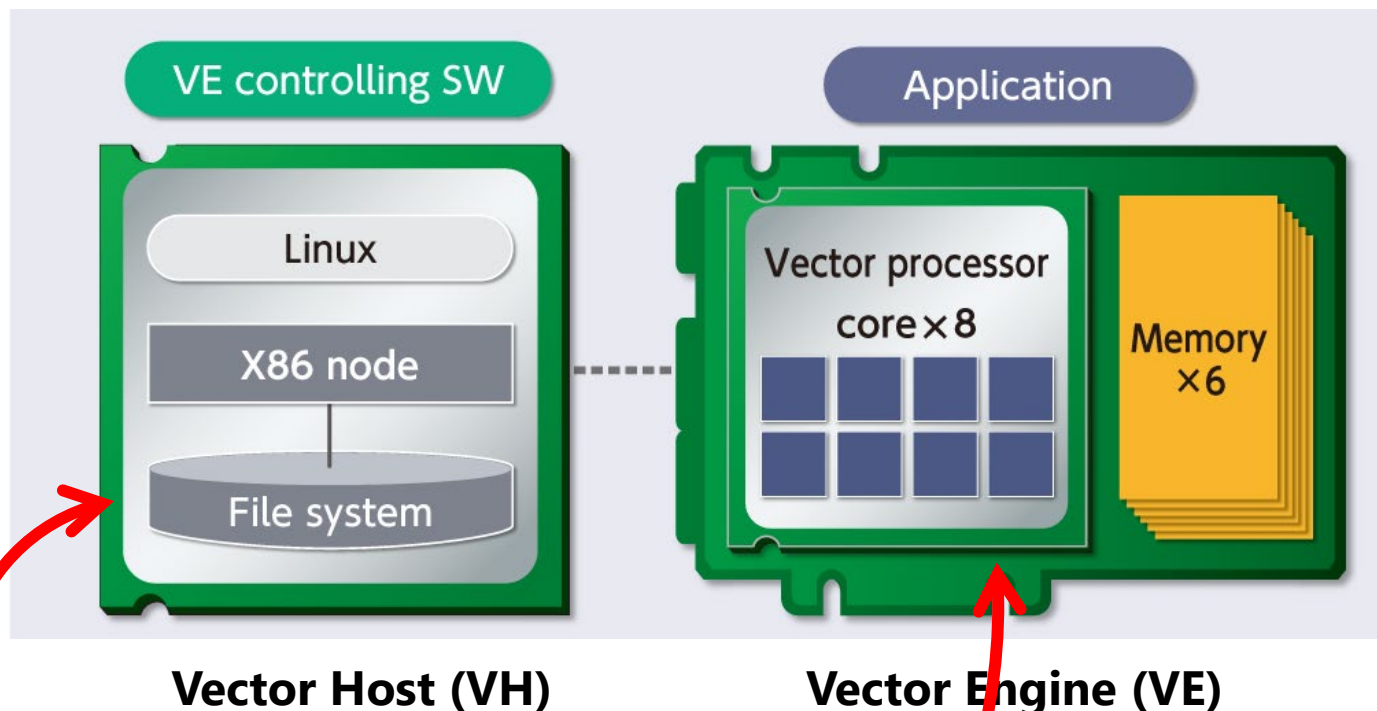
世界初HBM2 x6実装



世界初となるCPUと6個の
3次元積層メモリHBM2搭載技術
をTSMC社と共同開発



標準ソフト + 専用ハード



```
$ gcc matmul.c -o matmul
$ ./matmul

$ ncc matmul.c -o matmul
$ ./matmul
```

AOBA運用開始記念サイト

<https://www.cc.tohoku.ac.jp/aobadebut.html>



東北大学サイバーサイエンスセンターは
令和2年10月1日(木)から **新スーパーコンピュータ「AOBA」**の運用を開始しました

ご紹介

- 東北大の新スパコンAOBA始動！最先端科学から防災減災まで、現代社会を支えるスパコン
- 新スーパーコンピュータサービス開始のお知らせ
- バーチャルスパコン見学



ご挨拶

本日ここに、東北大学サイバーサイエンスセンターの新スーパーコンピュータシステムの更新にあたり、挨拶を申し上げます。

東北大学は、1907年の建学以来、113年にわたり「研究第一主義」の伝統、「門戸開放」の理念、「実学尊重」の精神のもと、多くの指導的人材を輩出し、さらに世界的に卓越した研究成果によって人類の知の地平を拡大して、未来社会へ向けた改革・イノベーションを先導してきました。また、2017年6月には、世界の有力大学と伍していくことを使命とする「指定国立大学法人」の最初の3校に指定されました。

現在、およそ17,800名の学生と6,300名の教職員を擁し、「杜の都」と呼ばれるこの仙台の地に、四季の気配を色濃く映す四つのキャンパスに10の学部と15の大学院研究科、3つの専門職大学院、6の附置研究所に加え、教育研究に携わる機構・センターと図書館・病院などを有しています。



東北大学理事・副学長
青木 孝文

東北大学サイバーサイエンスセンター長
菅沼 拓夫

文部科学省研究振興局参事官（情報担当）付 計算科学技術推進室長
宅間 裕子

日本電気株式会社 代表取締役 執行役員社長 兼 CEO
新野 隆

Director, The High Performance Computing Center (HLRS) in Stuttgart
Michael M. Resch

Director, Research Computing Center, Lomonosov Moscow State University
Vladimir Voevodin

東京理科大学工学部情報工学科 教授 (JAXA 名誉教授)
藤井 孝蔵

NUG (NEC User Group) 会長、大阪大学サイバーメディアセンター 准教授
伊達 進

掲載順 (敬称略)

利用者支援・共同研究

■ 平成9年よりセンター教職員とベンダーの協働による高速化支援体制を構築

→ 利用者支援を重視

- ・ 技術職員1名あたり2～3件の高速化支援を常時担当

これまでの高速化支援実績

年度	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
件数	2	8	8	9	10	7	18	20	8	16	10	15	8	8	13	6	11	9	6	9	7	7
ベクトル化性能向上比	1.9	46.7	4.5	2.5	1.6	2.2	6.7	2.9	1.3	2.9	33	9.4	381	47	16.2	19.7	16.7	10.3	32	19.2	5.9	1.7
並列化性能向上比	11.1	18.4	31.7	8.6	4.9	2.8	18.6	4.5	5.3	8.1	1.9	5.1	3.6	48	17.2	15.3	12.9	8	5	3	1.2	3.6

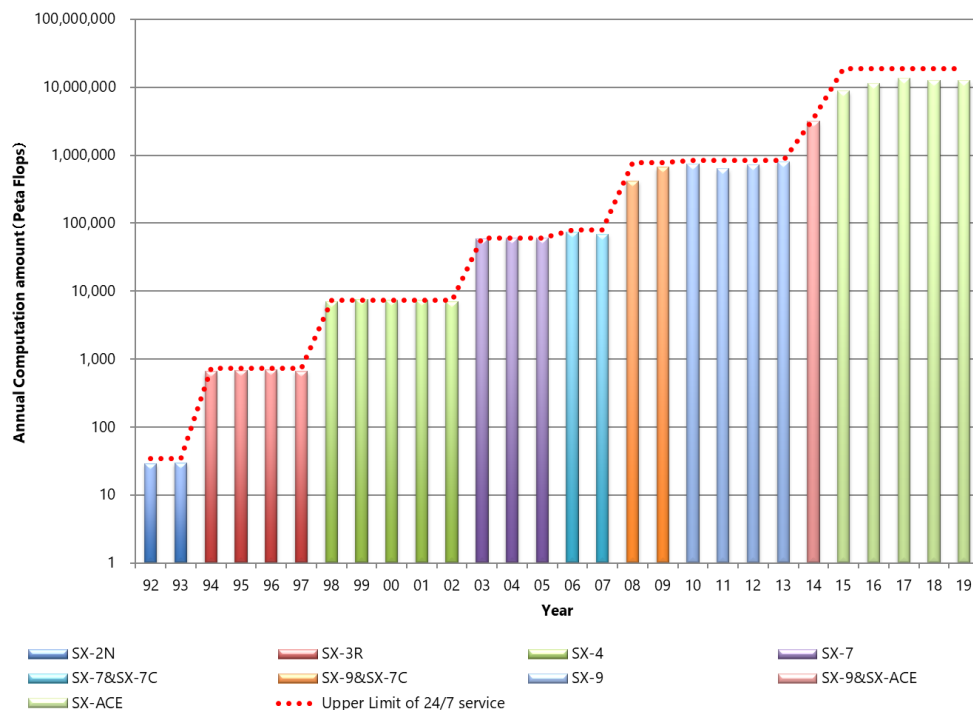
- ・ システム規模は小さいけれども**利用者満足度でナンバー1**を目指している情報基盤センターです！

常に高い利用率

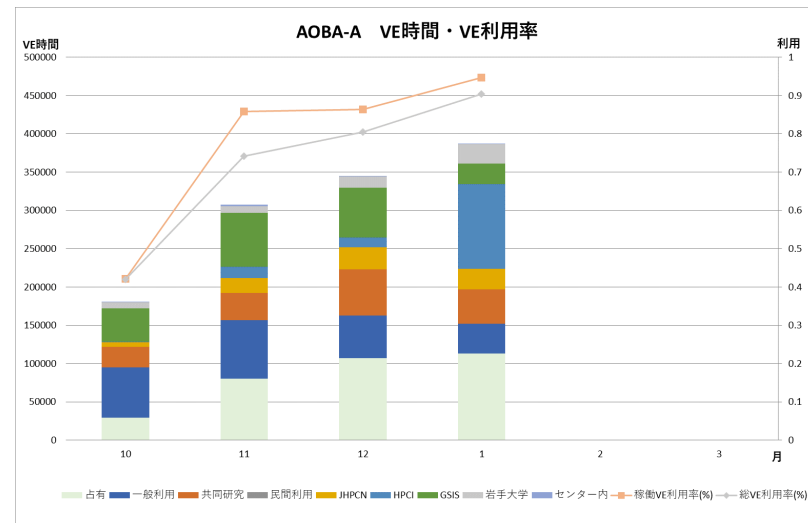
特徴のあるシステムの安定運用

+

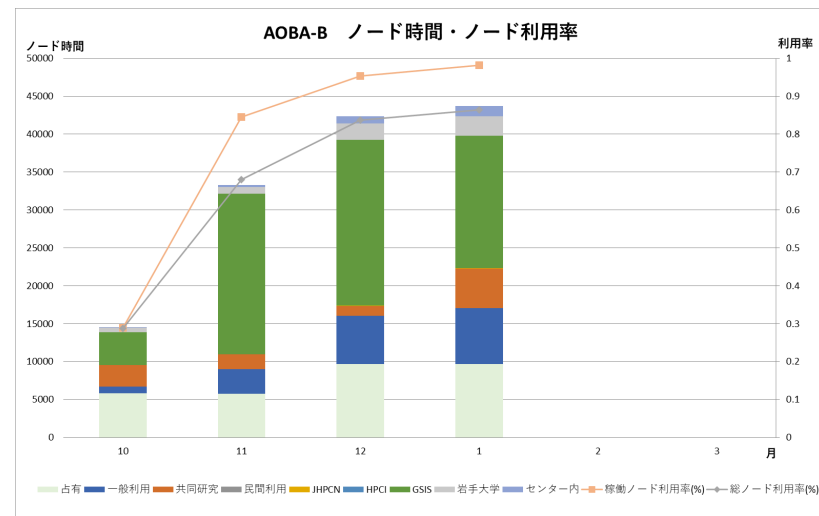
手厚い (+ 熱い) 利用者支援



歴史的に、新スパコン導入直後から提供可能上限に近い利用



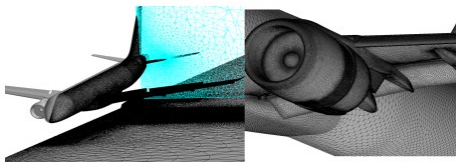
SX-ACEからAOBA-A (SX-Aurora TSUBASA)への移行は順調



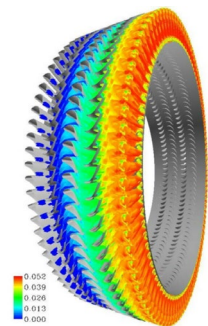
AOBA-B (x86)も運用開始直後から高い利用率

多様な科学技術分野に貢献

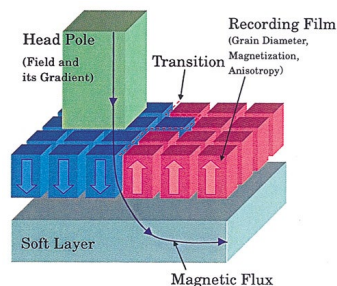
Next-Generation SuperSonic Transport



Turbine Design



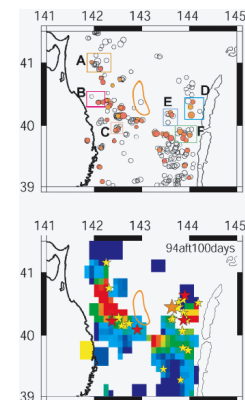
Perpendicular Magnetic Recording Medium Design



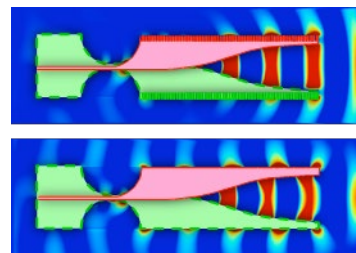
Nano Material Design



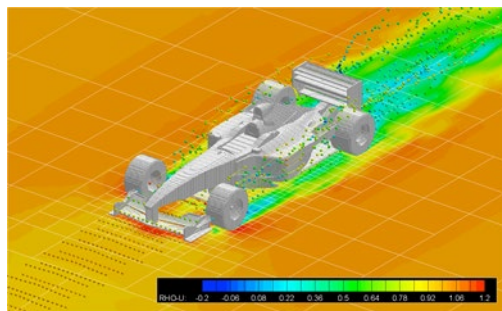
Earthquake Analysis



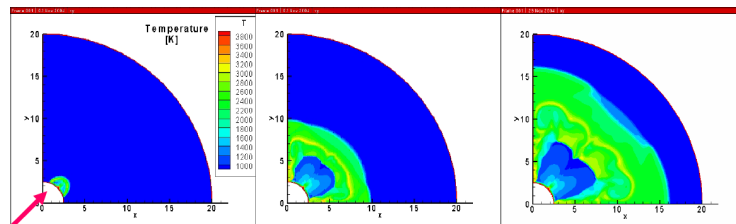
Antenna Analysis



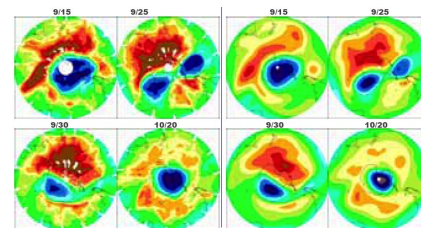
Next-Generation CFD Analysis



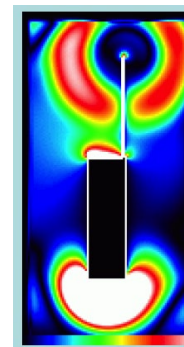
Combustion Flow Simulation



Ozone-hole Analysis



MRJ



リアルタイム津波浸水・被害推定システムの構築

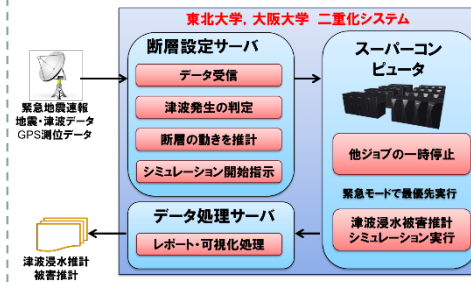
概要

南海トラフ地震を想定し、地震発生直後に津波浸水被害推計シミュレーションをリアルタイムで行い、被害推定情報を提供できる世界初のシステムの研究開発を行った。東北大学および大阪大学のスーパーコンピュータを利用することで、20分以内に津波浸水による被害を推定する。現在では、相模トラフ地震や日本海溝および千島海溝を震源とする地震にも対応できるように拡張されている。また、より多くの市町に実装、普及展開を図り、来る巨大地震津波災害への地域社会のレジリエンス※1向上を目指している。本システムは、内閣府の総合防災情報システム※2の一部として平成30年4月から本格運用されており、地震大国と呼ばれる我が国において、安全安心な生活を支える基盤技術として実活用されている。

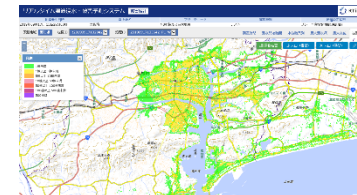
具体的な成果・効果

- 運用中のスーパーコンピュータを地震発生時には津波浸水被害推計シミュレーションに利用し、被害推計の可視化を行い、20分以内に配信できる**世界初**のシステム
- 東北大学、大阪大学、日本電気株式会社、国際航業株式会社、株式会社エイツー、株式会社RTi-castという**産学連携**による共同開発を実施
- 学問的波及効果に加えて、内閣府の総合防災情報システムの一部として採用されるなど、**安全安心を支える社会基盤**として大きな社会的波及効果
- 南海トラフ地震に加えて、対応できる震源の範囲拡大のために継続的にシステムを拡張しており、**社会的要請や期待が大い**ことは明白
- 本研究に関連して、平成29年度「情報化促進貢献個人等表彰」文部科学大臣賞、平成30年度 科学技術分野の文部科学大臣表彰科学技術賞、平成31年 第1回オープンイノベーション大賞総務大臣賞等を受賞

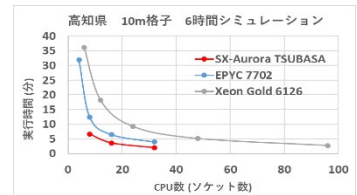
イメージ図



左図はシステムの流れをあらわしている。地震データの受信から被害推計の配信まで自動化を行い、24時間365日の運用を行っている。地震に関するデータは、気象庁の地震情報と国土地理院のGEONETのデータを利用している。



上図は本システムの被害推計結果の一例である。津波による浸水範囲と浸水深から建屋の被害棟数を推計したものである。



上図は津波浸水被害推計シミュレーションの実行時間である。JHPCNにおける共同研究によって複数のシステムで高速化を実現した。

用語解説

※1 被害を最小化して素早く回復すること

※2 総合防災情報システムは、災害発生時に政府等が被災状況を早期に把握し、迅速・的確な意思決定を支援するため、防災関係機関間で防災情報を地理空間情報として共有するシステム

熱中症リスク評価シミュレータの研究開発

概要

熱中症は、体内における熱バランスの崩れや脱水症状が要因となり発症する。本研究では、温帯生育者および熱帯生育者モデルに対して様々な暑熱環境条件における体温変化や発汗量等の解析を行い、これまで蓄積してきた解析結果の高精度化や、熱帯生育者に対するデータベースの拡充を行い、出身地の違いによる熱中症リスクの違いについて考慮することを可能とした。その成果は、日本気象協会の『熱中症セルフチェック』（※1）として健康な暮らしを支える実運用サービスに活かされている。また、短期暑熱順化（※2）の程度が熱中症リスク増大の因子となることや、当日のみならず過去数日間の暑熱気象条件が熱中症に影響することが報告されていることから、それらの要因も考慮した熱中症搬送人員数予測式を提案し、その有用性を示した。

具体的な成果・効果

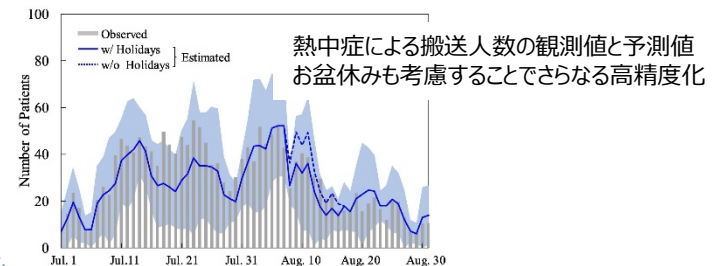
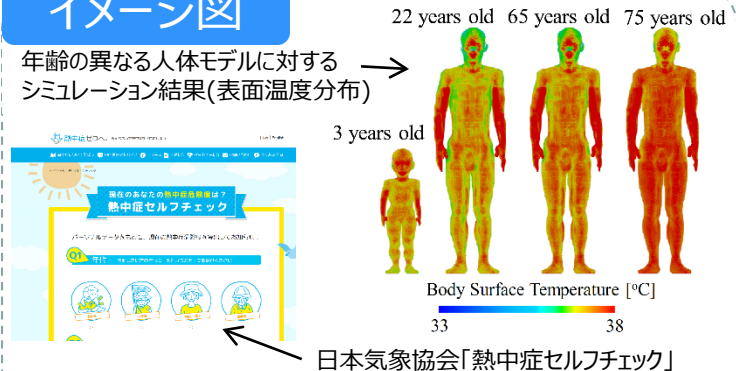
- 年齢や性別など個人特性を考慮したシミュレーション技術を実装・高速化し、3時間後の熱中症リスクを10分で評価する**熱中症リスク評価シミュレータ**を開発した
- 気象予報データと経験から得られた数式を融合させたデータを組み込むことにより、現実的な条件での熱中症リスク評価を可能にした
- 気象データと本技術を用いて構築されたデータベース利用して、**日本気象協会は『熱中症セルフチェック』**を公開し、平成29年4月から令和2年9月にかけて1,033,226件ものアクセスを記録した
- 短期暑熱順化や過去数日間の気象条件も考慮した**熱中症搬送人員数予測**を確立した
- 研究成果は、Environment International (Impact Factor:7.577)をはじめとする一流国際論文誌で発表され、学術的にも高く評価されている
- 研究代表者の平田晃正教授(名古屋工業大学)らは、令和2年度文部科学大臣表彰(科学技術賞 理解増進部門)、第75回電気学術振興賞(進歩賞) 等を受賞した

用語解説

※1 <https://www.netsuzero.jp/selfcheck>

※2 暑さに体が慣れることを暑熱順化といい、特に数日から数週間で慣れることを短期暑熱順化と呼ぶ

イメージ図



熱プラズマ乱流によるナノ粒子群創製プロセスシミュレーション

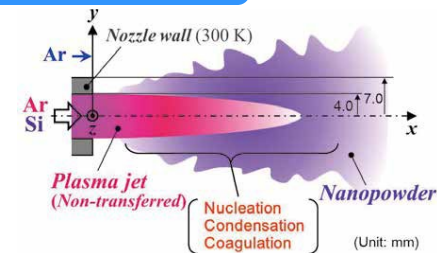
概要

粒径や組成が制御されたナノ粒子（※1）の大量生産が求められているが、その創製は容易ではない。近年、熱プラズマ流（※2）がナノ粒子を効率的に量産できるツールになるとして注目を集めている。しかし10,000 Kを超える熱プラズマについて、実験によって得られる情報には限りがあり、その効率的な制御のために必要な知見はいまだ少ない。また、ナノ粒子の形成現象はナノ秒〜ミリ秒の時間的にマルチスケールなプロセスであり、その集団的形成のメカニズムや輸送現象を直接計測することも困難である。そこでナノ粒子化する集団形成と輸送プロセスをシミュレートし、熱プラズマジェットの外縁近傍におけるナノ粒子群の集団的生成および輸送メカニズムを明らかにした。

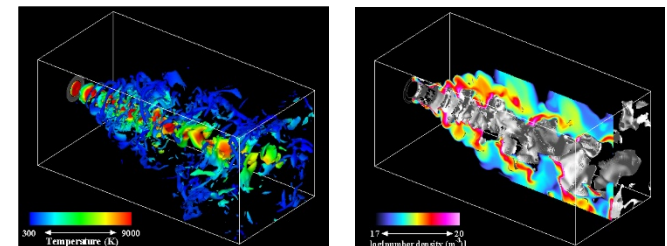
具体的な成果・効果

- 熱プラズマ内のみならず遠方にも複雑な渦構造が誘発されることは20年以上も前から実験により示唆されてきたが、それを**世界で初めて**数値シミュレーションにより再現することに成功した
- 高温の希ガス分子や水分子が解離・電離することで高活性となった熱プラズマジェット中に含まれる金属蒸気が移流・拡散しながら核生成・凝縮・凝集を経てナノ粒子化する集団形成と輸送プロセスをシミュレートし、活性種とイオンから成る熱プラズマジェットの**外縁近傍における金属ナノ粒子群の集団的生成および輸送メカニズムを明らかにした**
- 東北大学のSX-ACEを利用して、条件ごとに計算時間が約59時間かかるため、サイバーサイエンスセンターもコードの高速化を支援した
- 研究成果は、Plasma Chemistry and Plasma Processing (Impact Factor: 2.178)をはじめとする一流国際論文誌で発表され、学術的にも高く評価されている
- 茂田正哉准教授(大阪大学)は、日本学術振興会 プラズマ材料科学賞 奨励部門賞等を数多くの賞を受賞した

イメージ図



熱プラズマジェットによるナノ粒子群創製



熱プラズマジェット内外に誘発された渦構造(左図)および成長・輸送過程にあるナノ粒子群の中心断面分布(右図)

用語解説

※1 直径がナノメートルスケールの超微粒子であり、環境・医療・工学といった様々な分野への応用が期待されている

※2 高温により、原子核の周りの電子が原子から離れた(電離)状態の気流のこと(大気圧程度の圧力を有する)

アカウント登録は**無料**です！ 詳しくはウェブで！

<https://www.ss.cc.tohoku.ac.jp>

サブシステムAOBA-A (SX-Aurora TSUBASA)

利用形態	負担額および利用可能 VE 時間
共有 (無料)	利用 VE 数 1 (実行数, 実行時間の制限あり) 無料
共有 (従量)	課金対象時間 = (利用 VE 数 ÷ 8 を切り上げた数) × 経過時間 (秒) ← ≡ ノード時間 課金対象時間の合計 1 時間につき 125 円 課金対象時間は半期毎 (4~9 月および 10~3 月) に合計し, 1 時間未満を切上げて負担金を請求する。
共有 (定額)	負担額 10 万円 につき 課金対象時間の合計 800時間
占有	利用 VE数8 , 利用期間 3ヶ月 につき 270,000円

サブシステムAOBA-B (LX 406-Rz2)

利用形態	負担額および利用可能ノード時間
共有 (従量)	課金対象時間 = 利用ノード数 × 経過時間 (秒) 課金対象時間の合計 1 時間につき 22 円 課金対象時間は半期毎 (4~9 月および 10~3 月) に合計し, 1 時間未満を切上げて負担金を請求する。
共有 (定額)	負担額 10 万円 につき 課金対象時間の合計 4,600時間
占有	利用 ノード数1 , 利用期間 3ヶ月 につき 47,000円

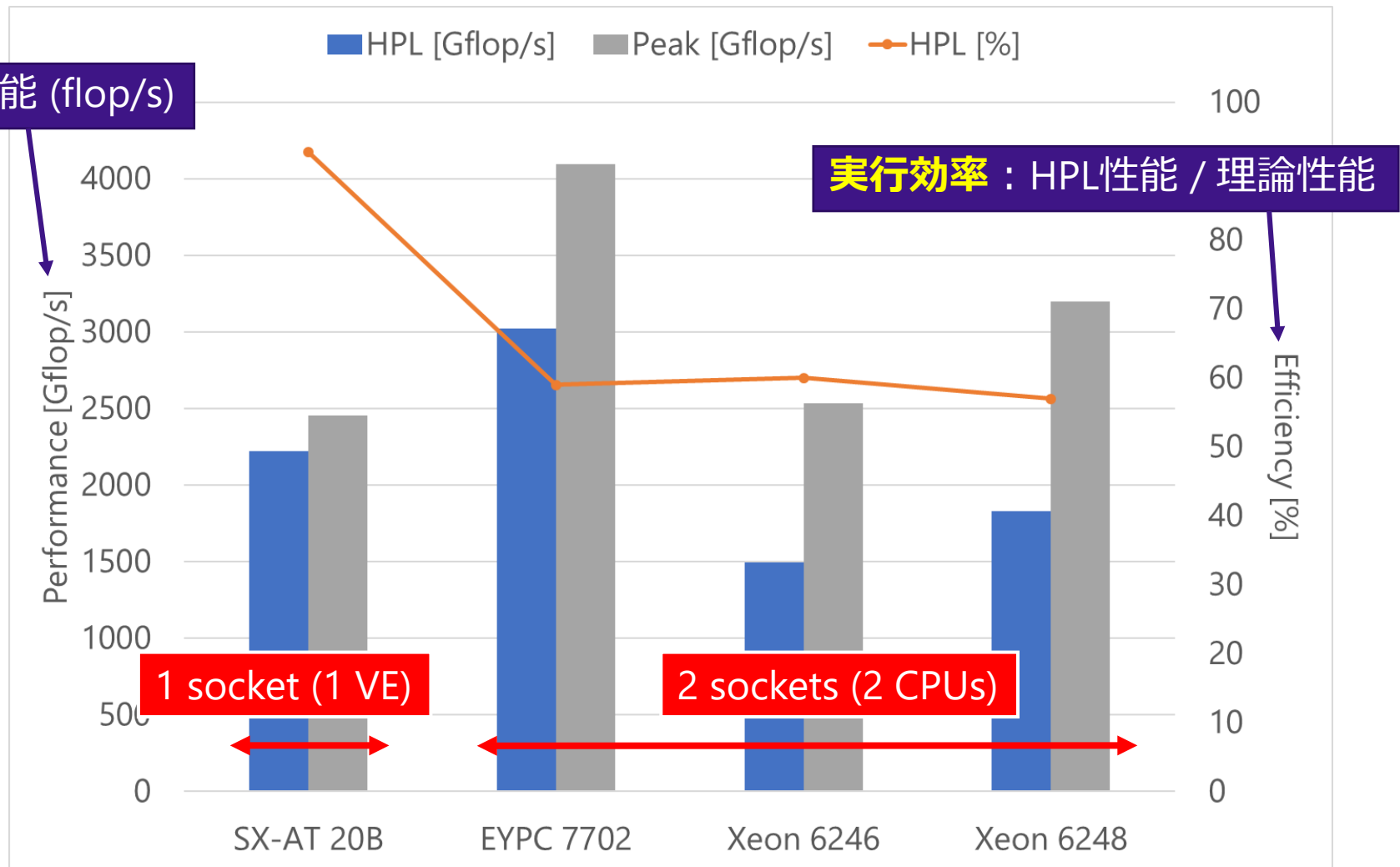
評価環境

		SX-ACE	SX-AT 10B	SX-AT 10AE	SX-AT 20B	EPYC 7702	Xeon 6246	Xeon 6248
Core	Clock Freq. [GHz]	1	1.4	1.58	1.6	2	3.3	2.5
	Peak Perf. [Gflop/s]	64	268	304	307	32	105	80
Socket	No. of Cores	4	8	8	8	64	12	20
	Peak Perf. [Gflop/s]	256	2150	2433	2457	2048	1267	1600
	Mem BW [GB/s]	256	1228	1352	1536	204	140	140
Node	No. of Sockets	1	1	1	1	2	2	2
	Peak Perf. [Gflop/s]	256	2150	2433	2457	4096	2534	3200
	System B/F	1	0.56	0.56	0.62	0.1	0.11	0.09



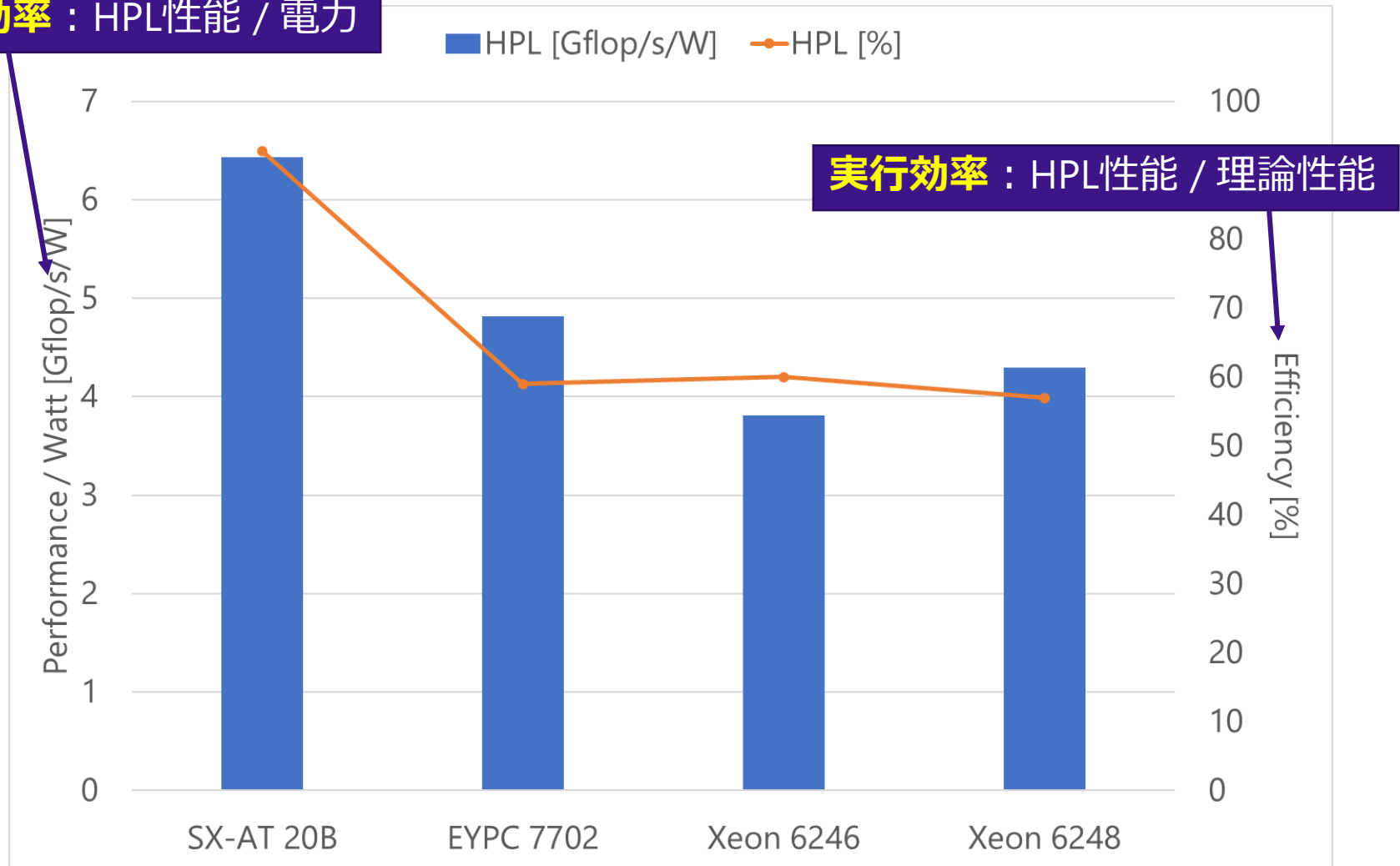
HPL 性能

演算性能 (flop/s)



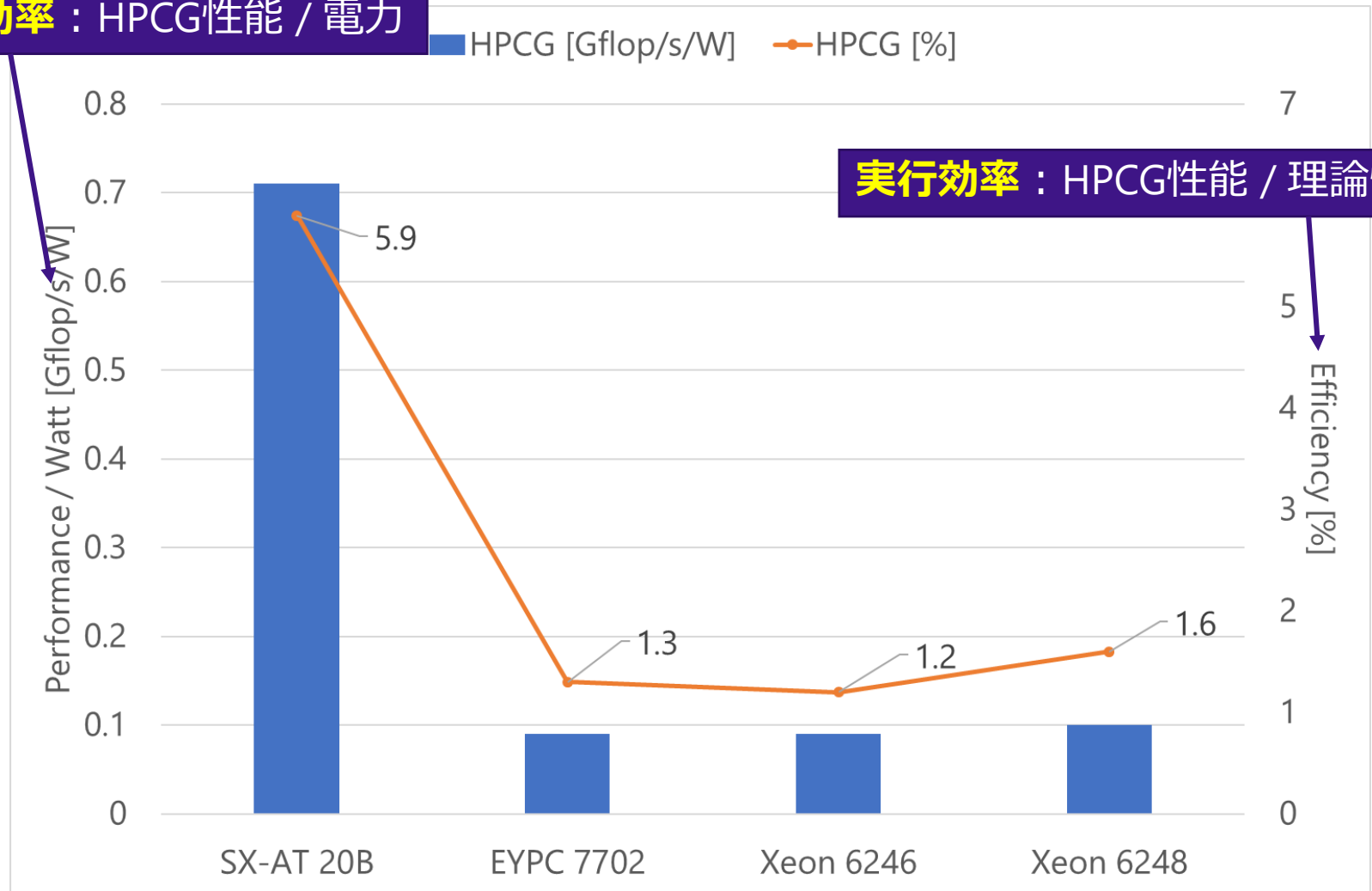
HPL 電力効率

電力効率 : HPL性能 / 電力



HPCG 電力効率

電力効率 : HPCG性能 / 電力



HPCG & HPL on AOBA

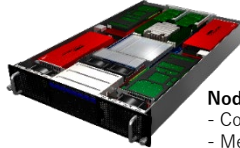
Egawa et al.@PMBS20



Subsystem AOBA-A



NEC SX-Aurora TSUBASA B401-8 x 72



Node Performance

- Comp. : 20.7 TF
- Mem size : 640 GB
- Mem BW : 12.4 TB/s

【SX-Aurora】

72VH(576VE) Peak 1415.6TF

HPL : 1140.1TF (Efficiency 80.5%)

HPCG : 75.3TF (Efficiency 5.32%)

Subsystem AOBA-B



NEC LX 406Rz-2 x 68



Node Performance

- Comp. : 4.1 TF
- Mem size : 256 GB
- Mem BW : 0.41 TB/s

【AMD Rome 7702】

68node Peak 278.5TF

HPL : 161.7TF (Efficiency 58.08%)

HPCG : 2.55TF (Efficiency 0.92%)

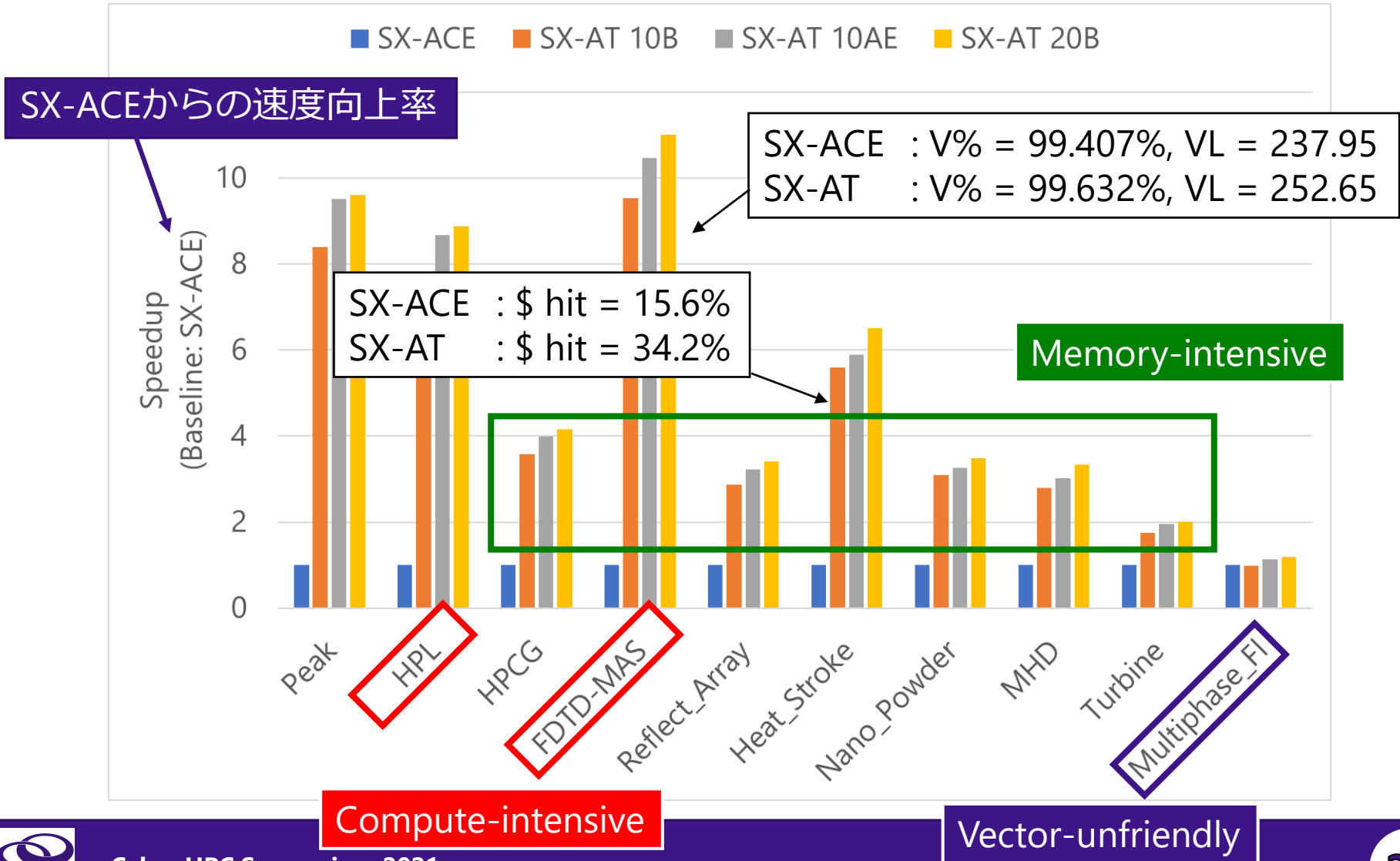
一般的なx86サーバと比較して圧倒的に高い効率を達成

実アプリによる評価

	Vectorization [%]	Ave Vec Len	No. of Processes	No. of VEs	B/F
HPL	99.3	254.4	256	32	0.31
HPCG	99.0	237.5	Compute-intensive		11.08
FDTD-MAS	99.6	252.7	1	1	0.29
Reflect_Array	98.8	216.1	8	1	3.74
Heat_Stroke	99.3	214.2	160	20	1.87
Nano_Powder	98.9	255.1	32	32	2.40
MHD	99.6	249.6	Memory-intensive		1.33
Turbine	92.0	175.5	159	20	2.52
Multiphase_FI	29.1	18.0	1	1	12.88

Vector-unfriendly

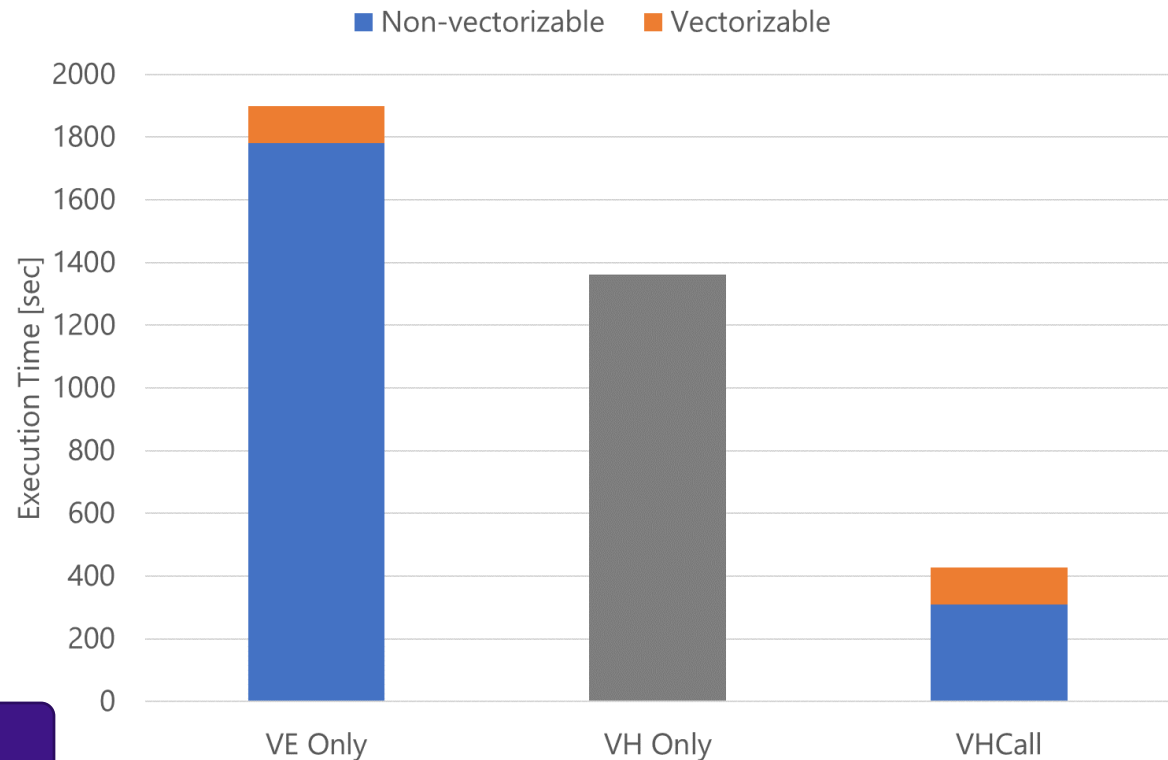
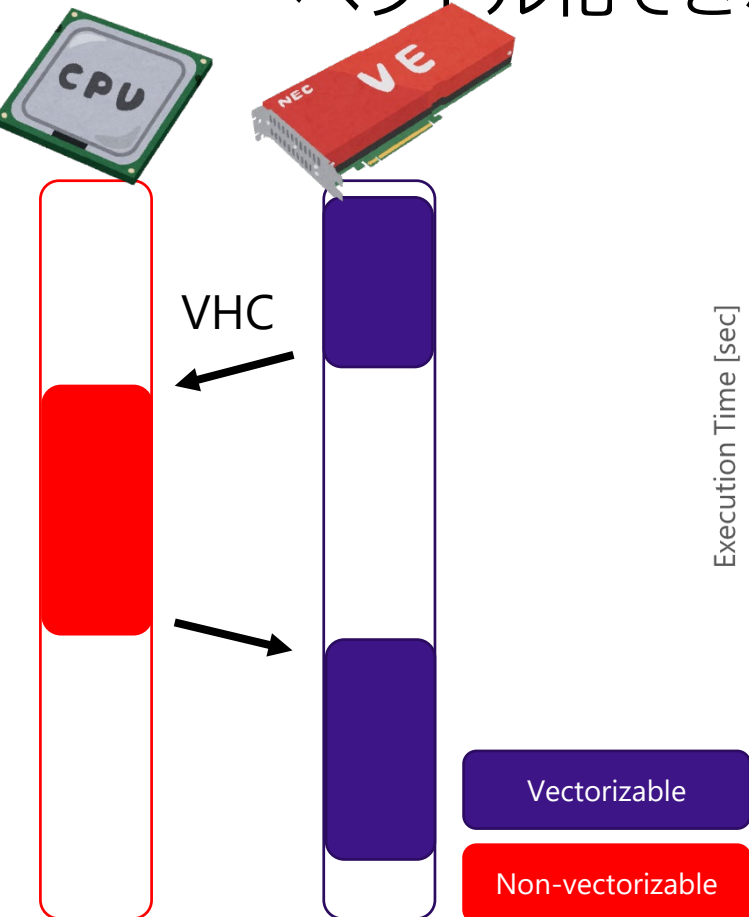
性能評価結果



VHとVEの連携

■ Vector Host Call (VHC)

- ベクトル化できない部分を**VE**から**VH**にオフロード

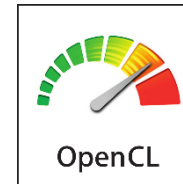
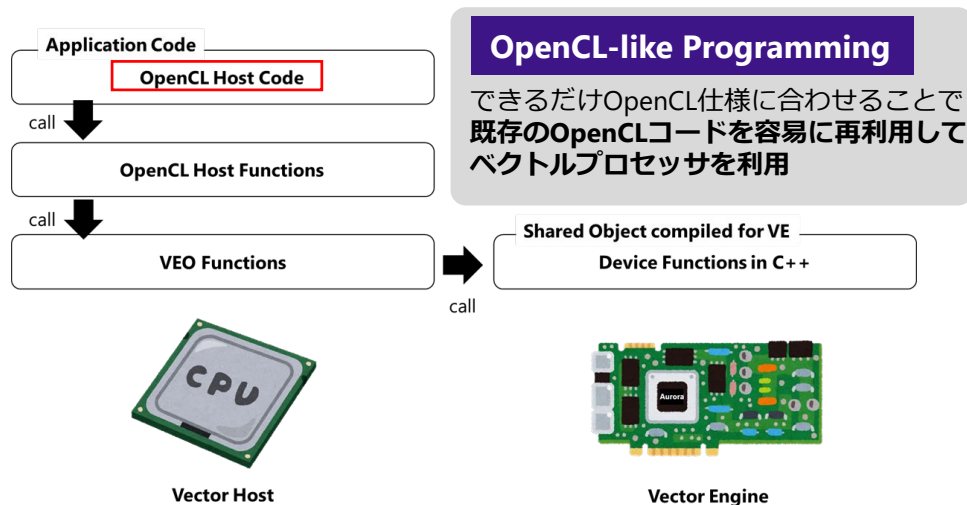


適材適所でプロセッサを利用した効果

研究活動：成果例 1

・次世代HPCプログラミング環境の創成

標準規格の議論に参加予定



現在はSYCLへ移行中 (neoSYCL)

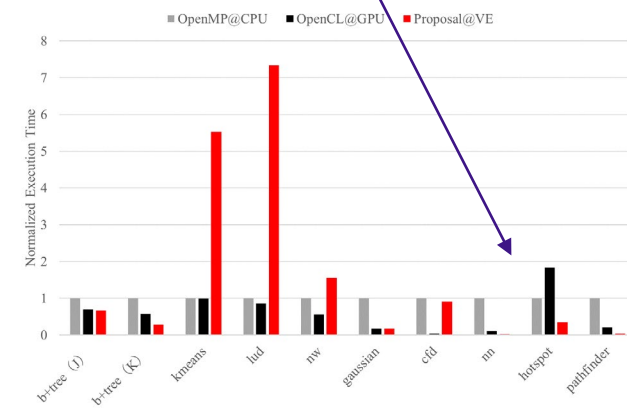
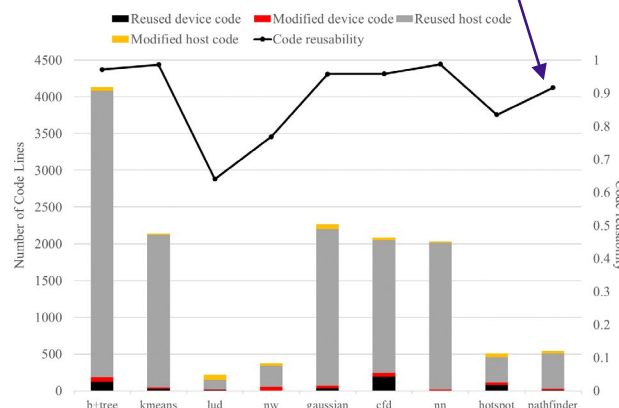
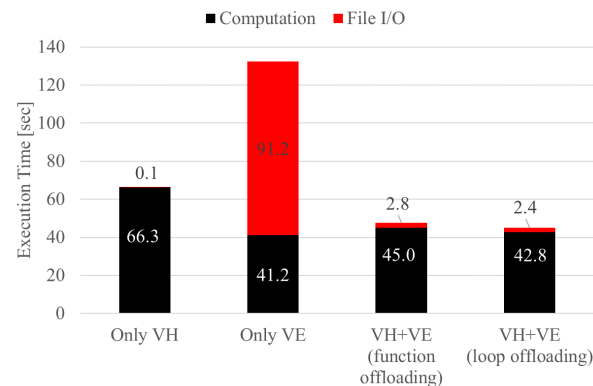


コードの複雑化を抑えつつオフロードを実装可能 (BFSの例)

	NLOC	AvgCCN	AvgToken
オフロード無し	116	4.5	196.2
従来(VEO)	225	7.4	302
neoSYCL	133	4.5	248

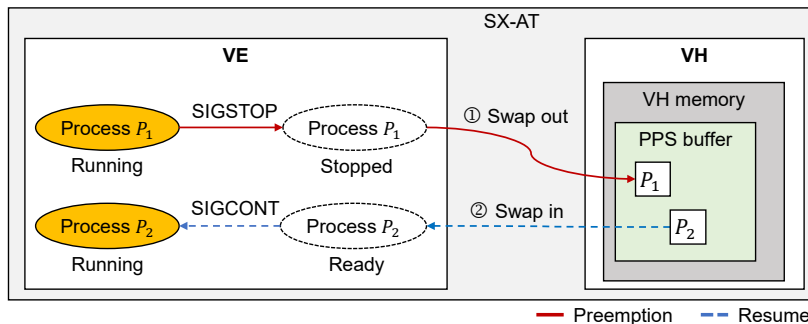
2種類のプロセッサを適材適所で利用

既存のコードの大半を利用しつつ ベクトル向き処理で高性能を実現

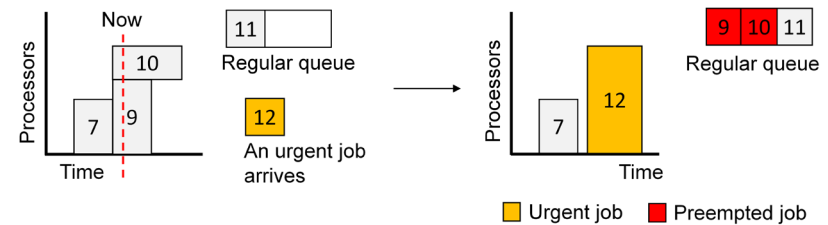


研究活動：成果例 2

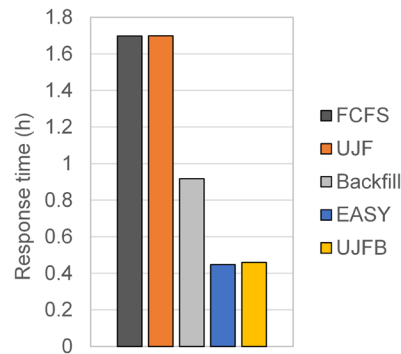
- スパコンを中核とするSociety5.x実装に向けて
- 運用システムを緊急時にも活用するための基盤技術
 - ヘテロ構成スパコンでの**緊急ジョブの割り込み実行**



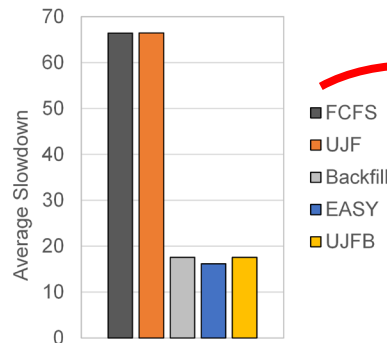
VE上のジョブをVHへ**部分的に退避**してから
緊急ジョブを実行 (**Partial Process Swapping**)



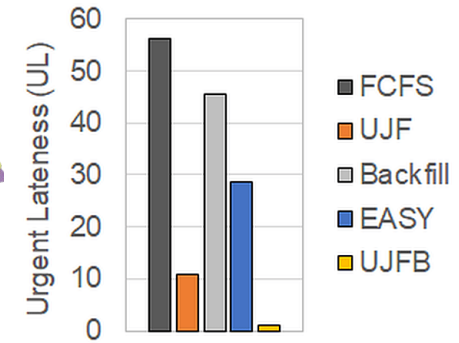
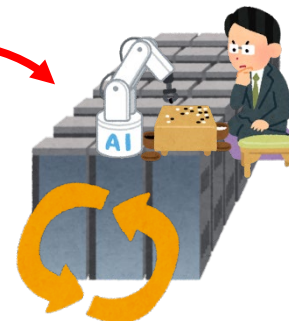
- 通常ジョブの性能への影響を抑える**緊急ジョブスケジューリング**



通常ジョブの実行時間



通常ジョブの速度低下率

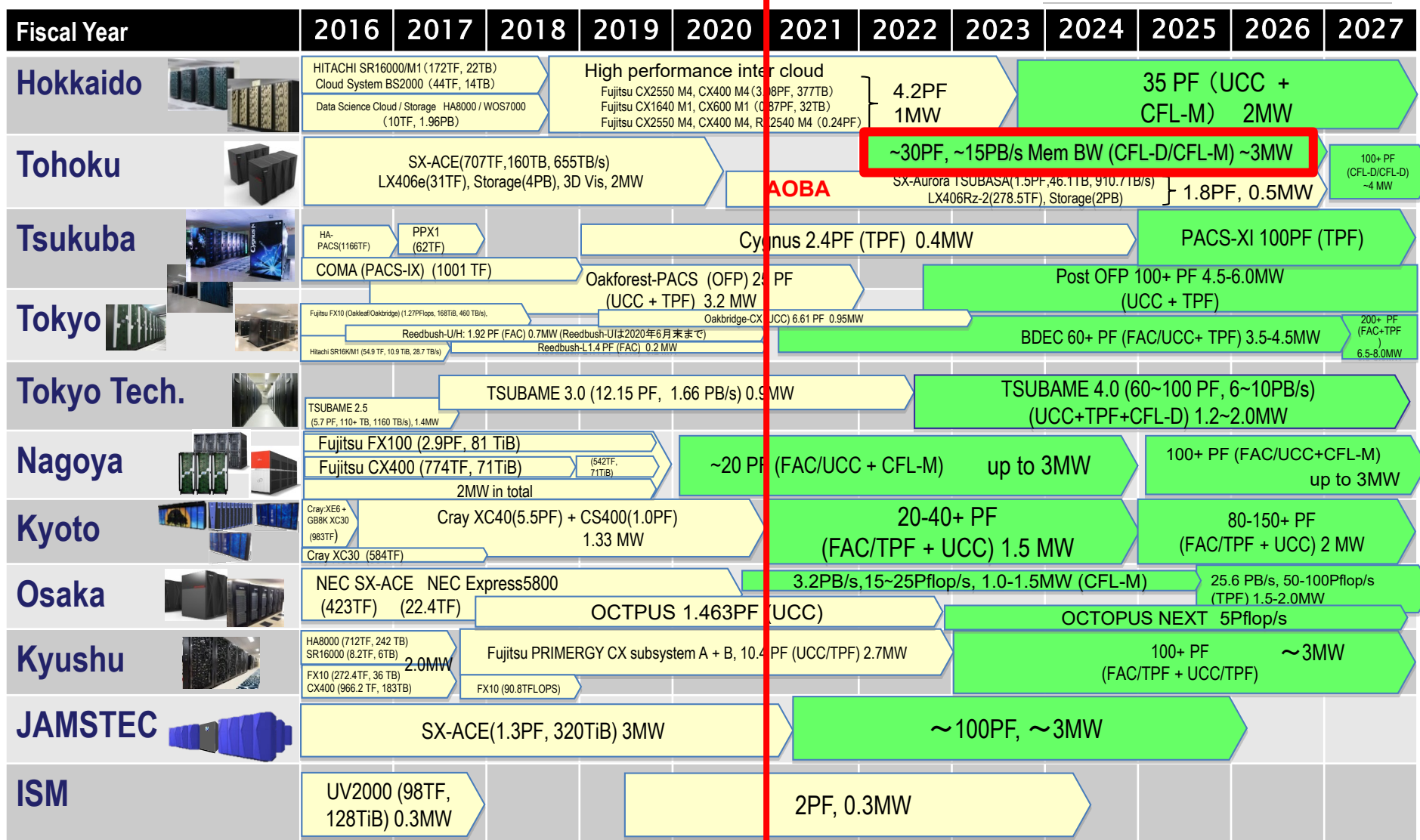


緊急ジョブの開始遅延



将来展望

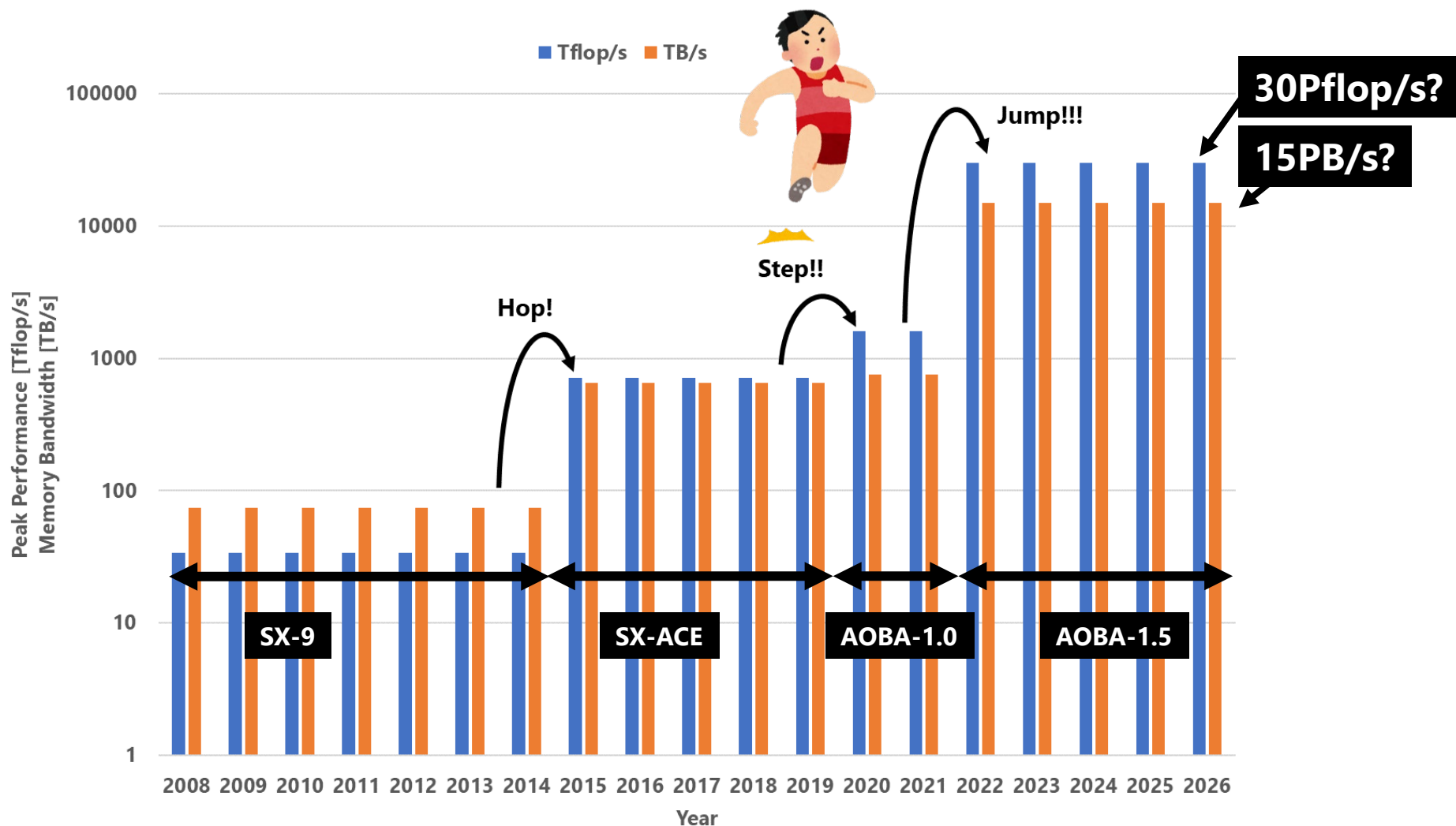
HPCI Tier 2 Systems Roadmap (As of June 2020)



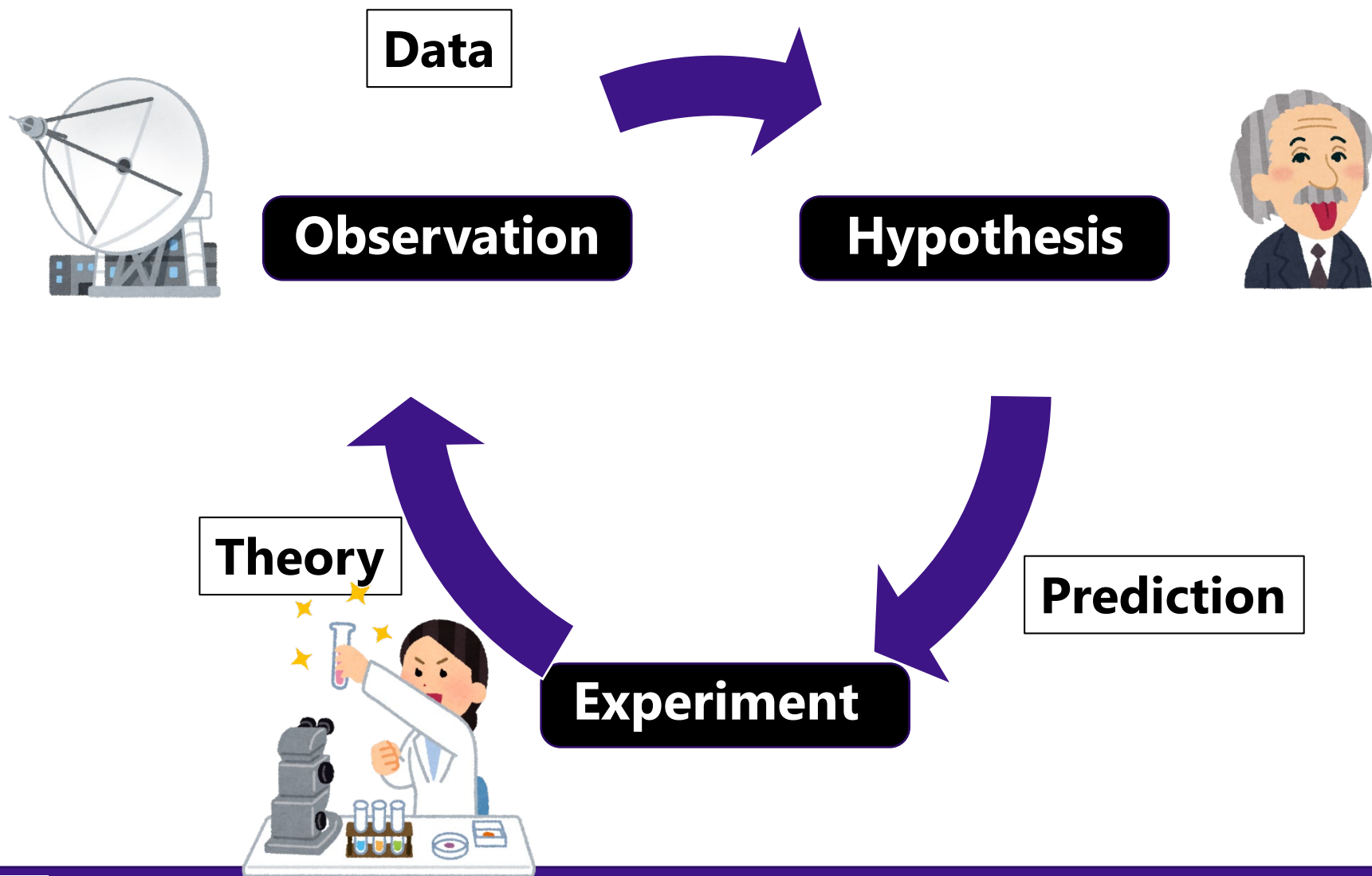
現在

power is the maximum consumption including cooling

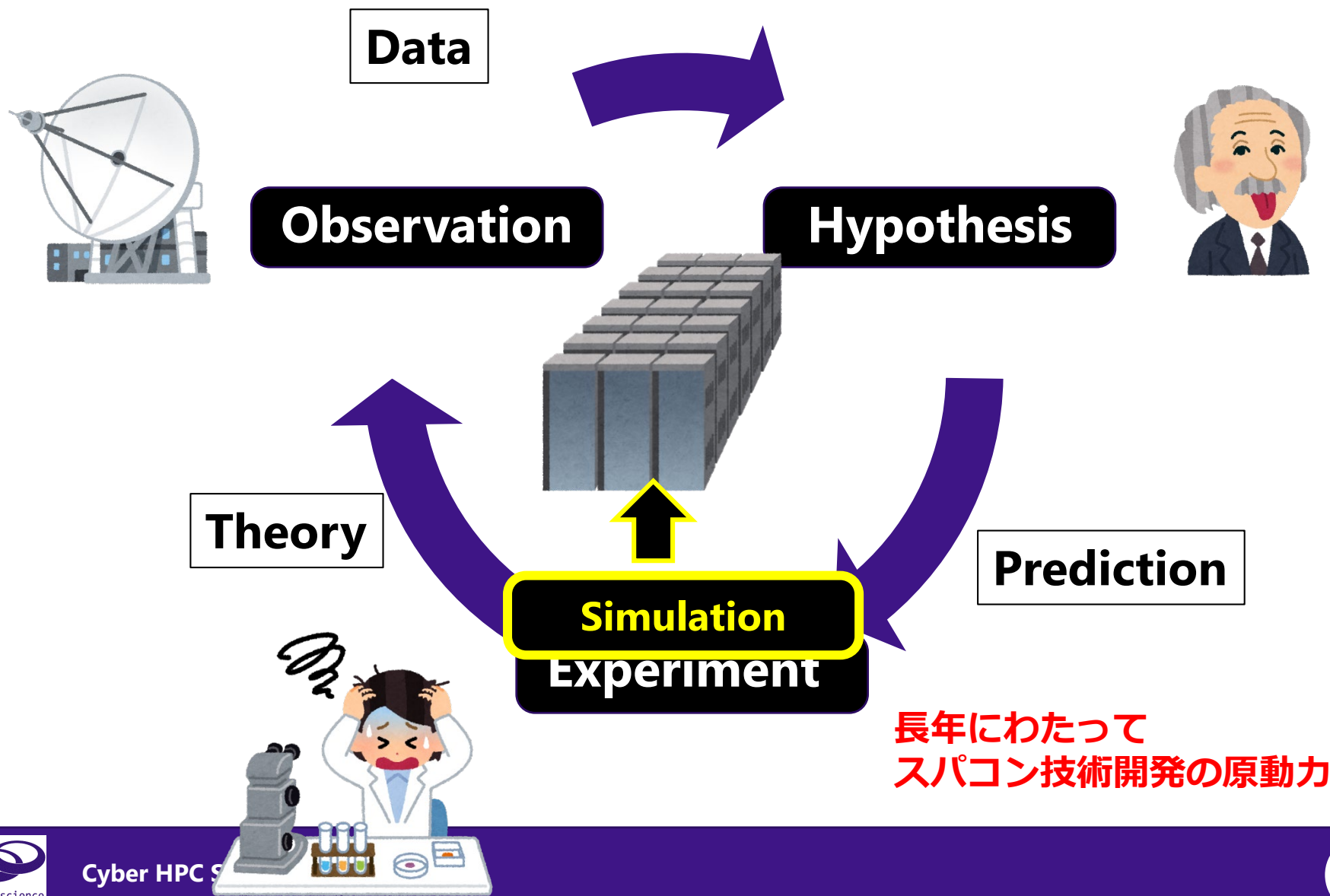
システム更新計画



科学技術発展のサイクル

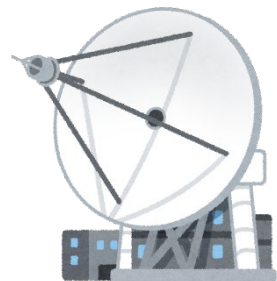


数値シミュレーションの役割



データ爆発の時代

膨大なデータを解析する
技術の重要性



Data

Data Science

Observation

Hypothesis



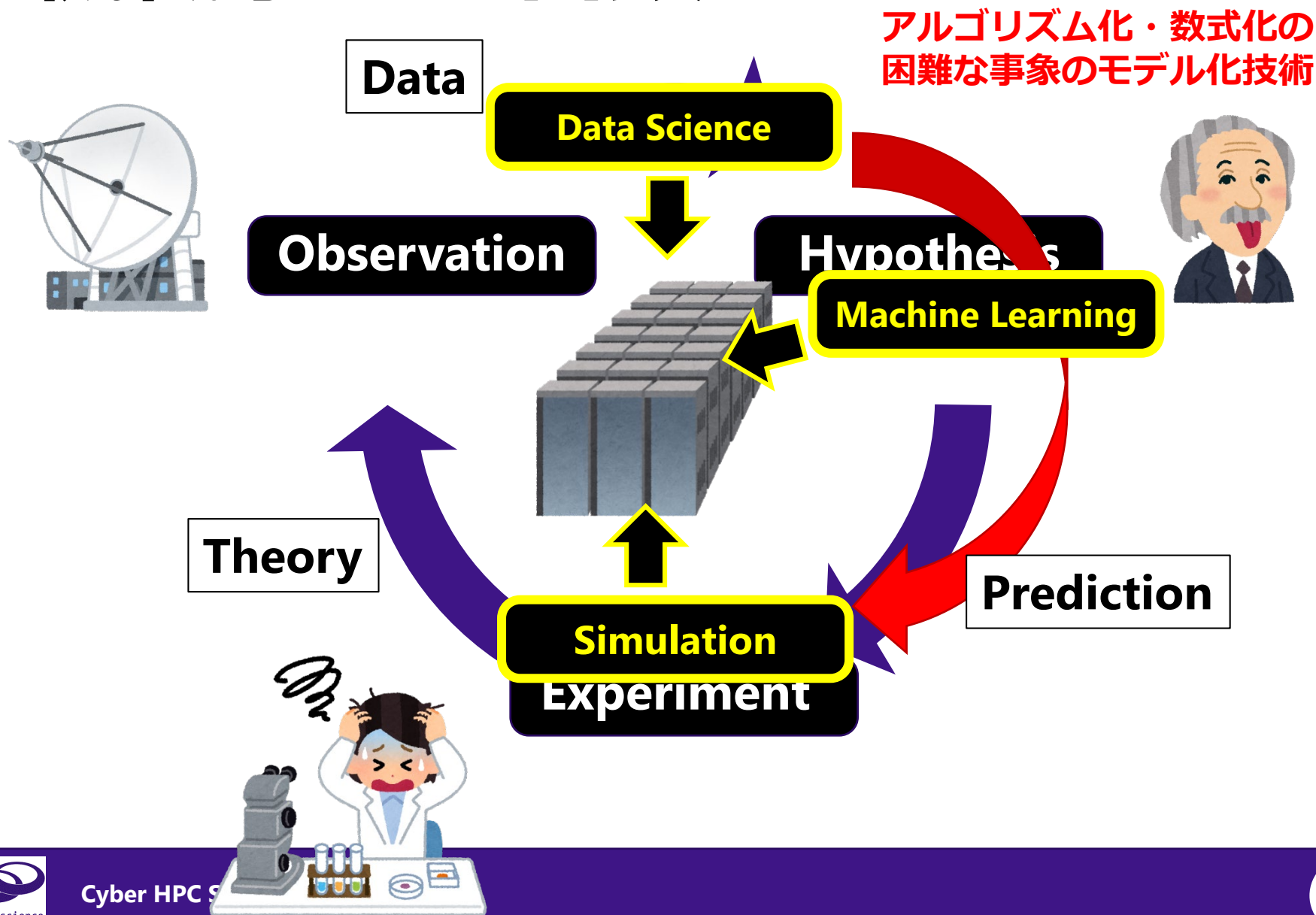
Theory

Prediction

Simulation
Experiment



機械学習の台頭



AIを支えるスパコン技術

- 計算能力の限界で**AIの進歩が頭打ちに？** WIRED (2020/7/11)

<https://www.wired.com/story/prepare-artificial-intelligence-produce-less-wizardry/>

- 深層学習が要求する計算能力は指数的に増加 Thompson et al. (2020)

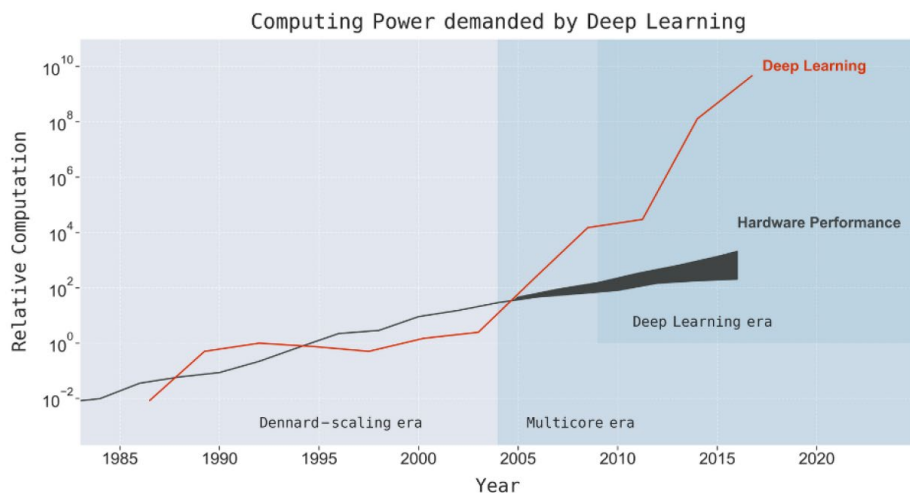
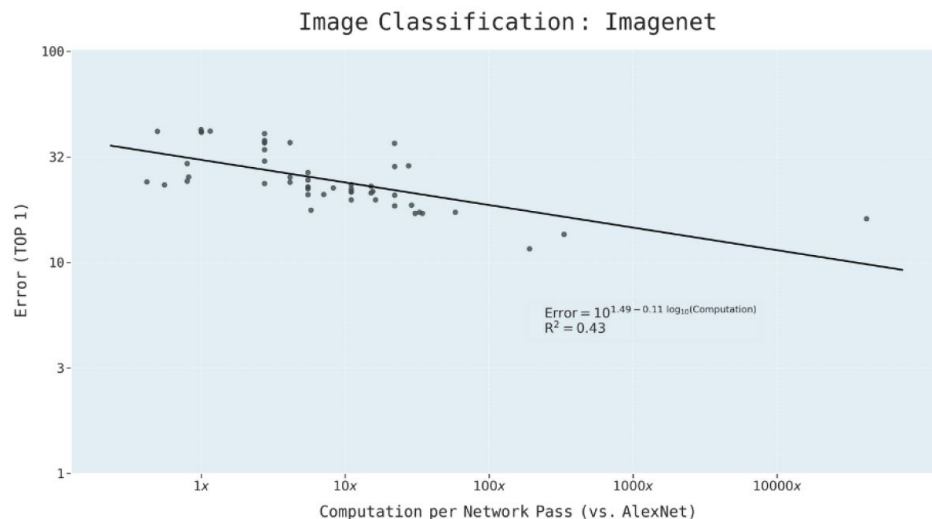


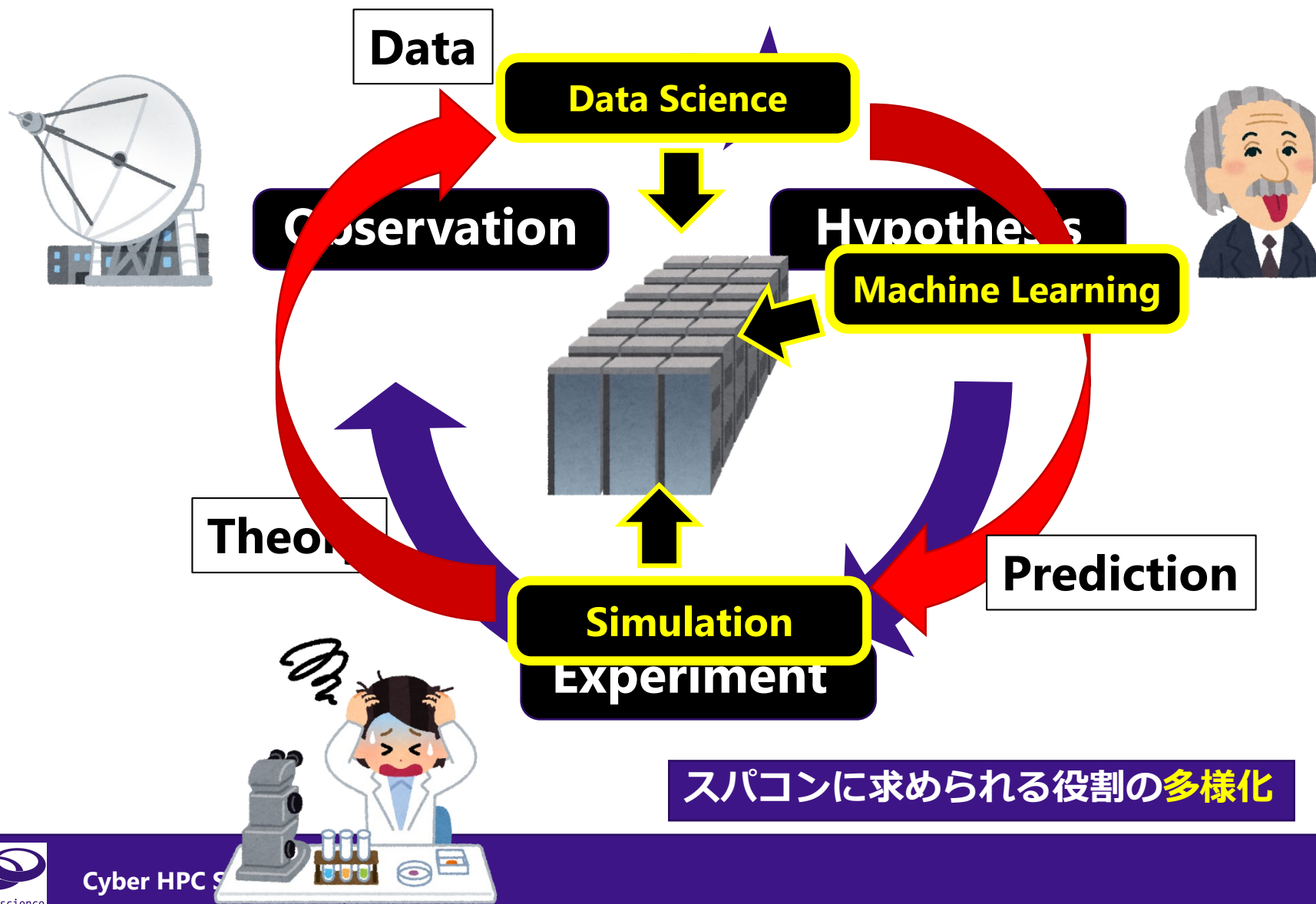
Image Classification: Imagenet

深層学習が求める計算能力の上昇は
プロセッサの進化よりも速い



誤差を減少させるために必要な計算能力は
指数的に増加

スパコンの役割の変化

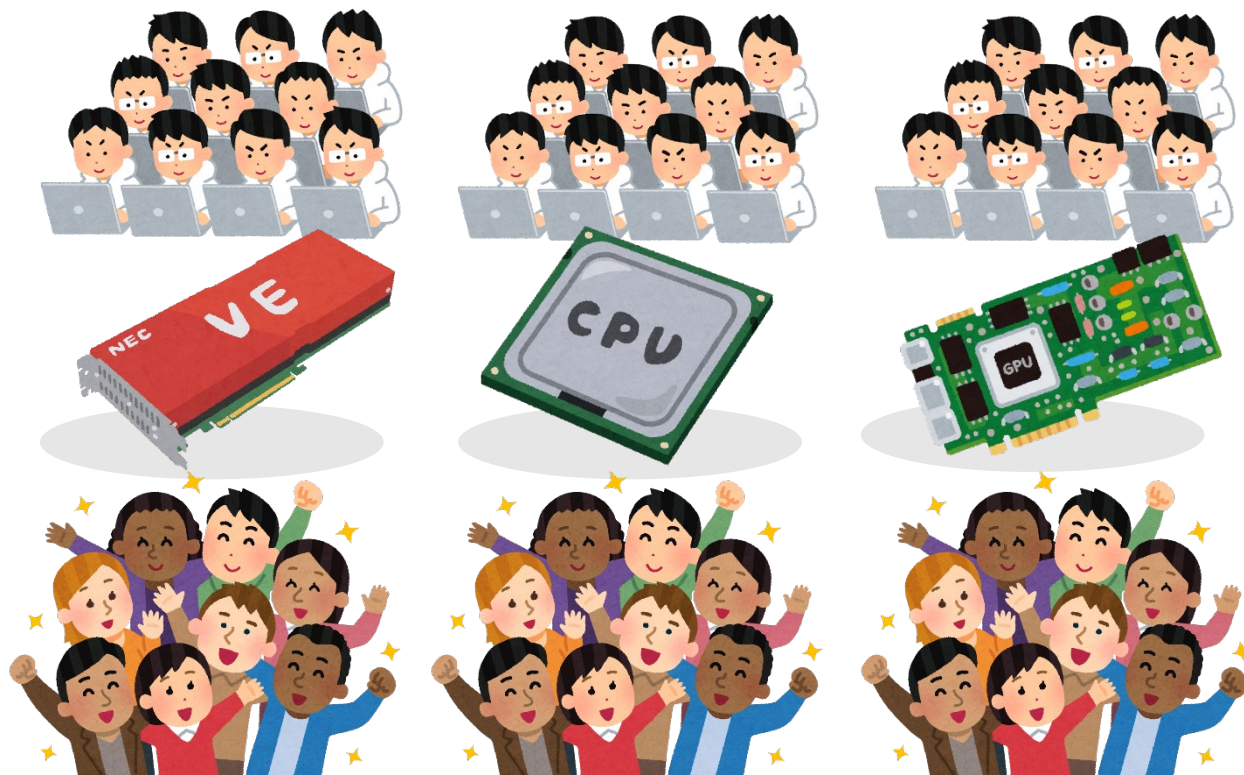




ではどうする？

多様化時代のセンターの姿

それぞれのスパコンに対して十分なサポート体制を背後に構築する必要



利用者支援・共同研究 (再掲)

- 平成9年よりセンター教職員とベンダーの協働による高速化支援体制を構築

→ **利用者支援を重視**

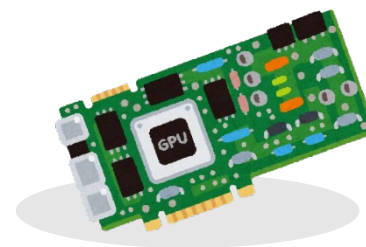
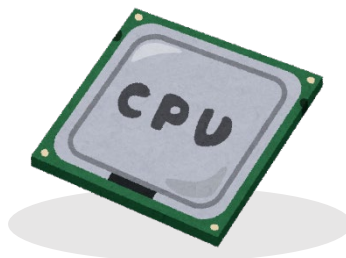
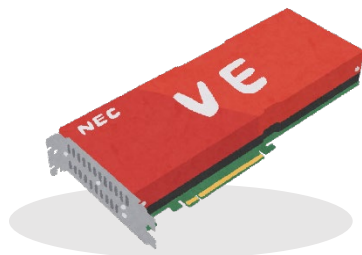
- ・ 技術職員1名あたり2～3件の高速化支援を常時担当

これまでの高速化支援実績

年度	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
件数	2	8	8	9	10	7	18	20	8	16	10	15	8	8	13	6	11	9	6	9	7	7
ベクトル化性能向上比	1.9	46.7	4.5	2.5	1.6	2.2	6.7	2.9	1.3	2.9	33	9.4	381	47	16.2	19.7	16.7	10.3	32	19.2	5.9	1.7
並列化性能向上比	11.1	18.4	31.7	8.6	4.9	2.8	18.6	4.5	5.3	8.1	1.9	5.1	3.6	48	17.2	15.3	12.9	8	5	3	1.2	3.6

- ・ システム規模は小さいけれども**利用者満足度でナンバー1**を目指している情報基盤センターです！

支援できるか？



支援体制も含めたシステム



現在の強みを延ばしつつ新領域の開拓も目指す

まとめ

■ 東北大学サイバーサイエンスセンター

- ・ システム性能ではなく顧客満足度で1位を目指しているセンターです
- ・ スパコンそのものだけではなく支援体制も含めたトータルシステムで考えています

■ 現在はベクトルを前提とした支援体制

- ・ 技術職員＋ベンチャー技術者による**強力なサポート！**

■ 新領域の開拓も**もちろん重要**

↑ **支援体制も構築できること前提で** (努力中)

- ・ 教員の体制は再構築中 → **ガチ公募予定**

ご応募お待ちしております！

