

第7回けいはんな「エジソンの会」

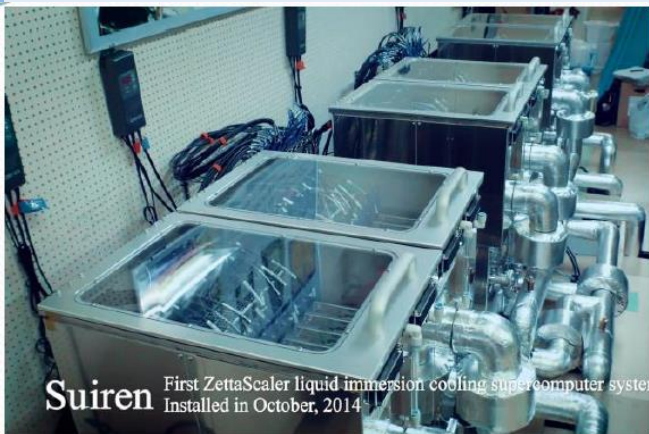
「次世代スパコンと人工知能エンジン による、AI駆動科学の時代に向けて」

2017年 2月14日

齊藤 元章

(株式会社PEZY Computing/株式会社ExaScaler/UltraMemory株式会社
株式会社Deep Insights/株式会社Infinite Curation)

現在稼働する6台の独自スパコンシステム



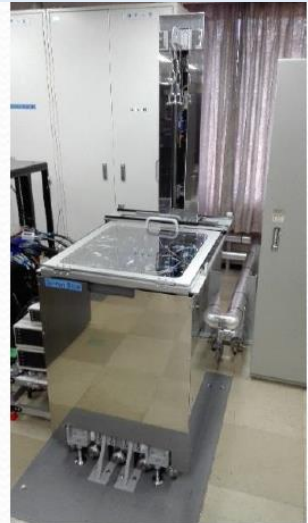
Suiren First ZettaScaler liquid immersion cooling supercomputer system
Installed in October, 2014

Suiren (睡蓮) ZettaScaler-1.5
2014.10 Install 2016.5 Upgrade
(32node to 48node)



Shoubu Biggest System of ZettaScaler system with over 1 PetaFLOPS
Installed in June, 2015

Shoubu (菖蒲) ZettaScaler-1.6
2015.6 Install 2016.5 Upgrade



Suiren Blue (青睡蓮)
ZettaScaler-1.5
2015.5 Install
2016.5 upgrade

Ajisai (紫陽花)
ZettaScaler-1.6
2015.10 Install 2016.5 Upgrade



Satsuki Configured with 20 bricks instead of normal 16 computation bricks
First TOP500 Supercomputer operated in a PRIVATE OFFICE

Satsuki (皐月) ZettaScaler-1.6
2016.5 Install



Sakura First TOP500 Supercomputer operated in a small BUSINESS OFFICE
Used for the development of PLEZY3 power efficient processors

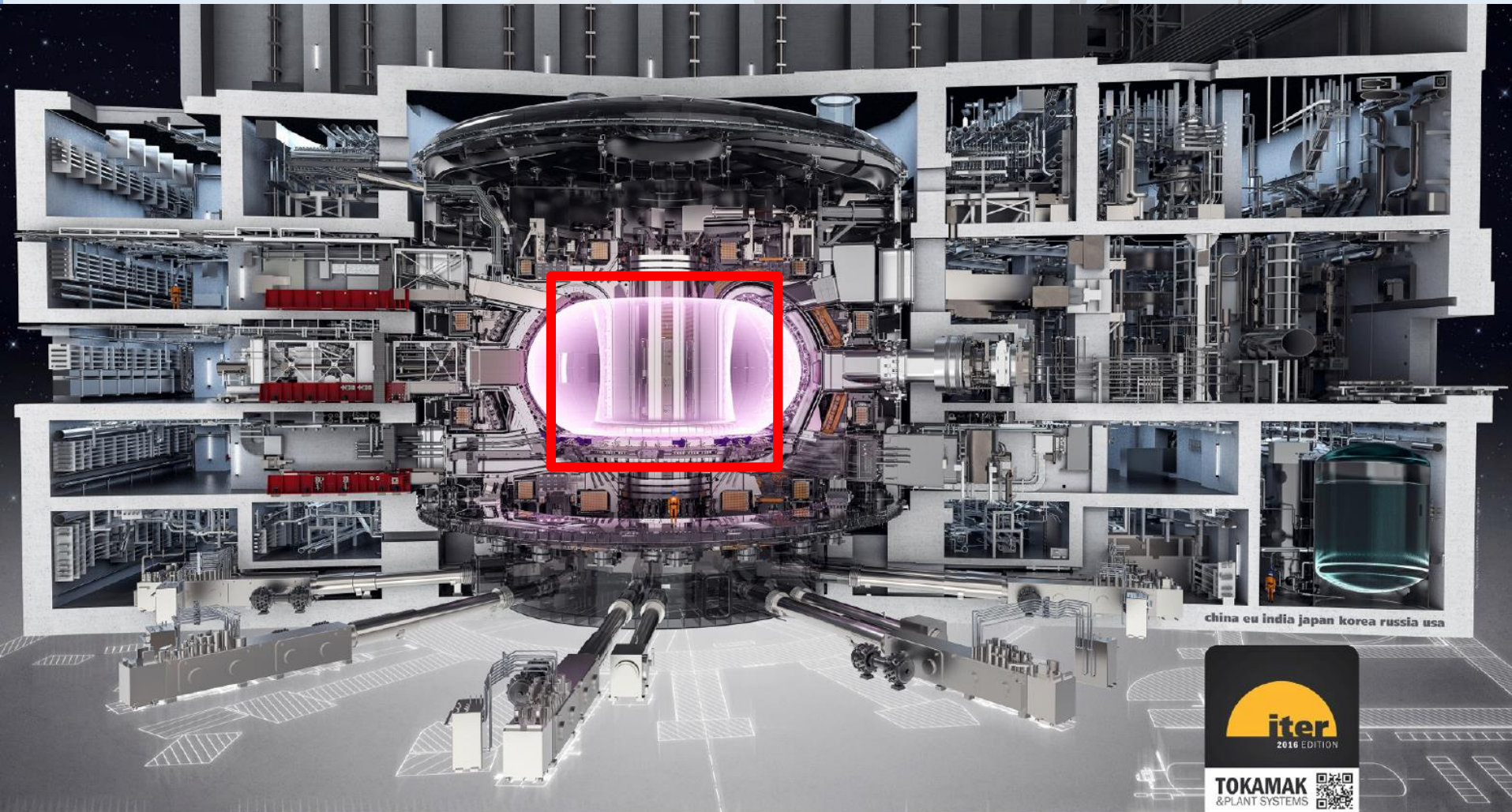
Sakura (さくら)
ZettaScaler-1.6
2016.5 Install

「次世代スパコン」とは

- 日本では2022年に「京」より100倍高速な「ポスト京」が開発される予定 (ExaFLOPS: エクサフロップス)
- 「100年掛かる計算が、1年で出来る」ことの他にも、「100倍複雑な計算ができる」ことが重要 (ネズミの脳が100倍高速になっても、知性レベルが上がる訳ではない)
- 熱核融合炉の全体をスパコン上に構築して、詳細なモデリングと、シミュレーションを繰り返せる
- 遺伝子解析に引き続き、タンパク質の解析が可能になり、代謝レベル解析も行え、生命科学が大きく進化
- 人工知能搭載の軍事用コンピュータウィルスの開発で、大量破壊兵器が無効化され、その保有がリスクに

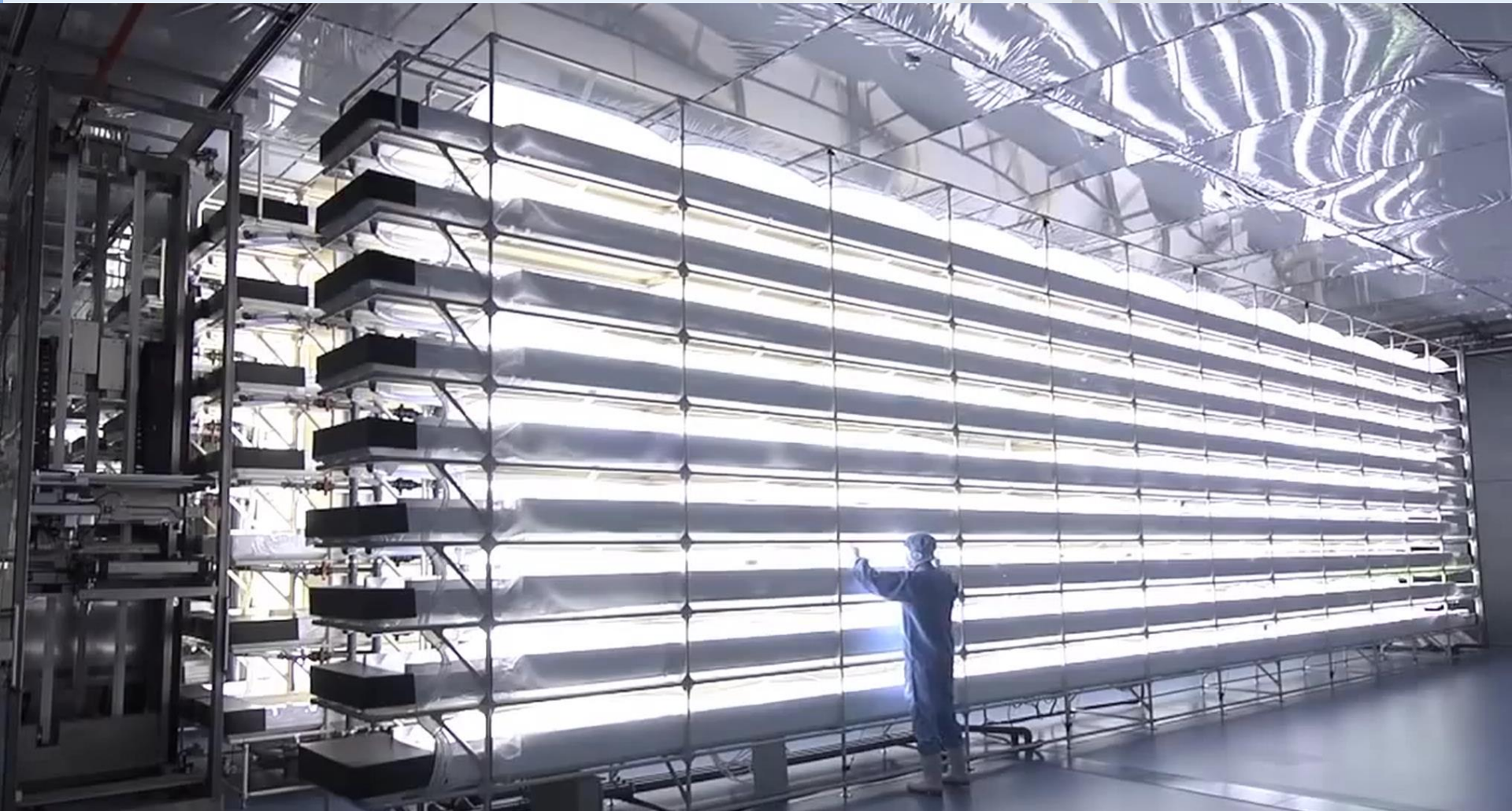
「次世代スパコン」とは

- ・熱核融合炉の全体をスパコン上に構築して、詳細なモデリングと、シミュレーションを繰り返せる



次世代スパコンと次世代植物工場

・LED光源の電気代と、冷却装置の電気代で、8割の植物工場は赤字経営のため普及が進まない



理化学研究所の「Shoubu (菖蒲)」(2015年)



4か月で開発したExaScaler-1.4による5台構成の2 PetaFLOPS級のスパコンを、「Shoubu (菖蒲)」として、理化学研究所和光の情報基盤センターに設置して頂く



RIKEN : Colors 知の旅人たち 全章 [Read More](#)

[研究論文 \(STAP細胞\) に関する取組み、情報等について \(2015年4月9日\)](#)

4つの方針：研究不正の調査・科学的検証の実施・研究論文の取下げ・再発防止の取組みについて掲載しております。

[一般の方](#) >

[研究者・学生の方](#) >

[企業の方](#) >

[報道関係の方](#) >

[理研関係者](#) >

プレスリリース



トピックス



2015年6月25日

[重力によって移動方向が変わらないオーキシンを発見](#)



理研らの国際共同研究グループは、植物ホルモン「オーキシン」の一種であるフェニル酢酸 (PAA) が、重力によって移動方向が変わらないユニークな特徴を持つことを発見しました。 [続きを見る...](#)

2015年6月25日

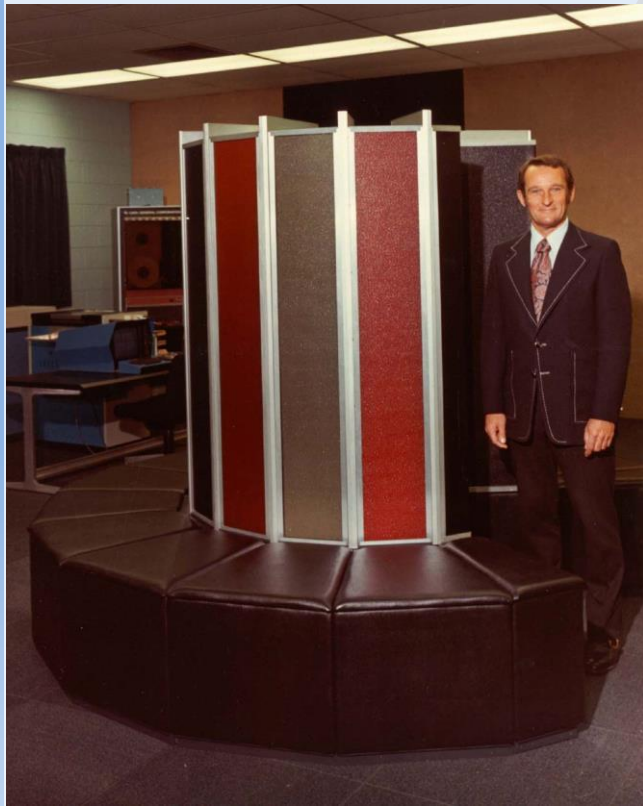
[ExaScaler及びPEZY Computingが、理化学研究所と共同研究契約を締結し理研情報基盤センターに2 PetaFLOPS級の液浸冷却スーパーコンピュータ「Shoubu \(菖蒲\)」を設置](#)



株式会社ExaScalerと株式会社PEZY Computingは理研と共同研究契約を締結し、理研情報基盤センターにExaScaler-1.xの2PetaFLOPS級の液浸冷却スーパーコンピュータ「Shoubu (菖蒲)」を設置します。 [続きを見る...](#)

40年前のスパコンと、現在のスパコン

- 1976年に開発された「Cray-1」と、京速計算機「京」



規模が巨大になり過ぎ、性能向上の速度が大幅に低下中

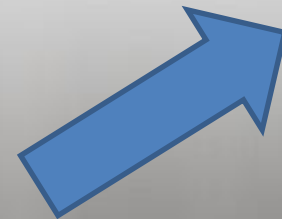


2020年、5nm世代の液浸冷却スパコン

- 第5世代となる超メニーコアプロセッサ「PEZY-SC4」
(16,000コア, 50TFLOPS, 5nmプロセス, 25TB/s帯域, 64bit CPU内蔵)
- 消費電力効率: 100 GFLOPS/W (1Wで1秒間に1千億回)
- タワーサーバ性能: 100 PetaFLOPS (1台で「京」の10倍)
(冷却系を含めた体積効率・性能密度は「京」の1万倍以上)

タワーラック
864台の

京 × 10
K computer



創業法人構成

プロセッサ開発



創業:2010年1月
社員数:26名

株式会社PEZY Computing
(ペジーコンピューティング)

- 独自メニーコア・プロセッサ開発
- 同汎用PCIeボード開発
- 同独自システムボード開発
- 同アプリケーション開発
- 半導体2.5次元実装技術開発
- ウェハ極薄化応用技術開発

積層メモリ開発



創業:2013年11月
社員数:50名

UltraMemory株式会社
(ウルトラメモリ)

- 超広帯域独自DRAM開発
- DRAM積層技術開発
- 磁界結合メモリIF開発
- ウェハ極薄化応用技術開発
- 広帯域、高速DRAM開発
- 最先端汎用DRAM受託開発

液浸冷却開発



創業:2014年4月
社員数:12名

株式会社ExaScaler
(エクサスケaler)

- 液浸冷却技術開発
- HPC液浸システム開発
- 液浸スパコンシステム開発
- 液浸冷却水槽販売
- 液浸冷却システム販売
- 液浸冷却用ボード類販売

株式会社Deep Insights
(ディープインサイト)

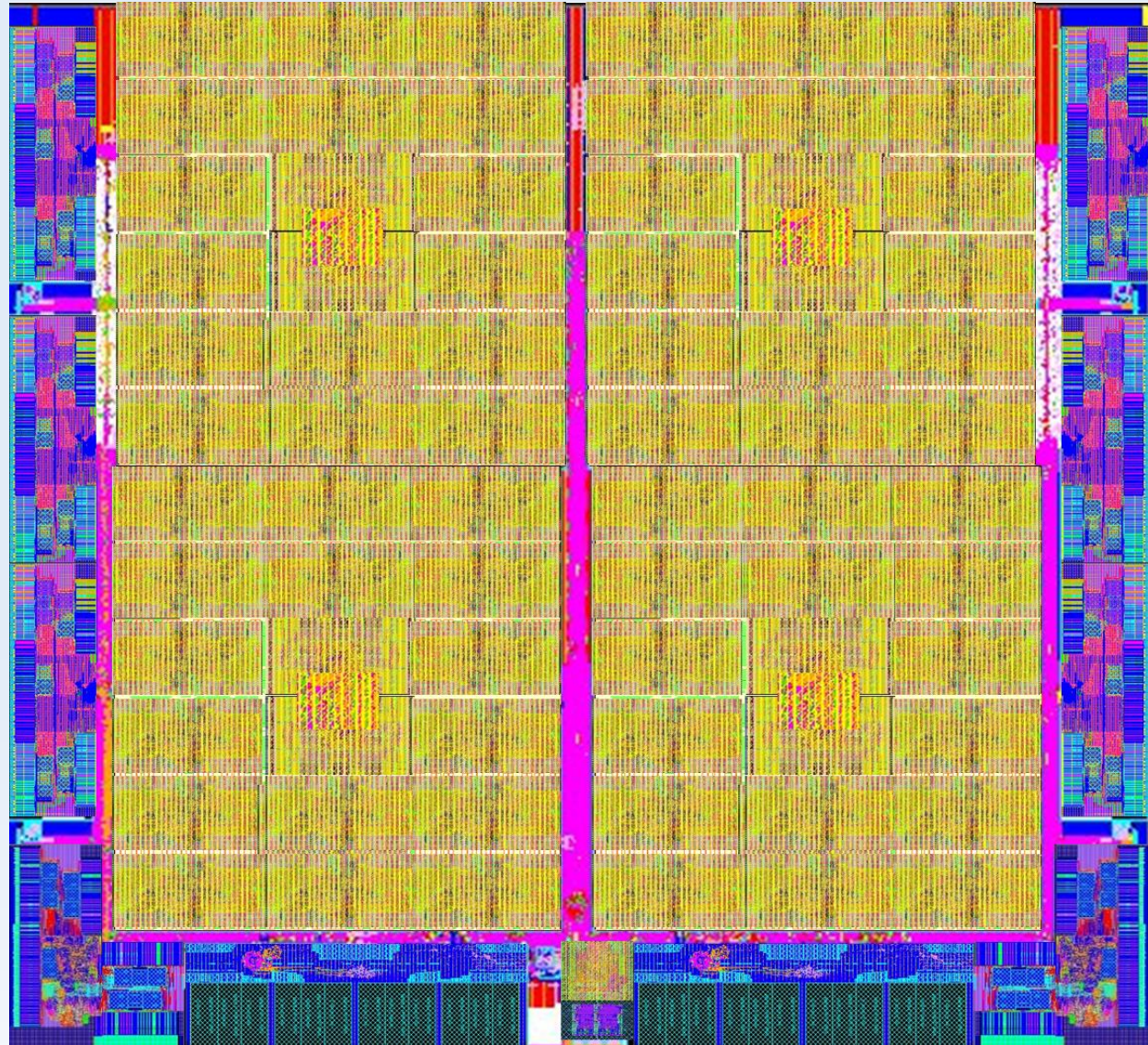
- 高効率人工知能ハード
- 超広帯域大容量メモリ混載
- 精度可変密行列演算, etc.



創業:2016年6月
社員数:~15名

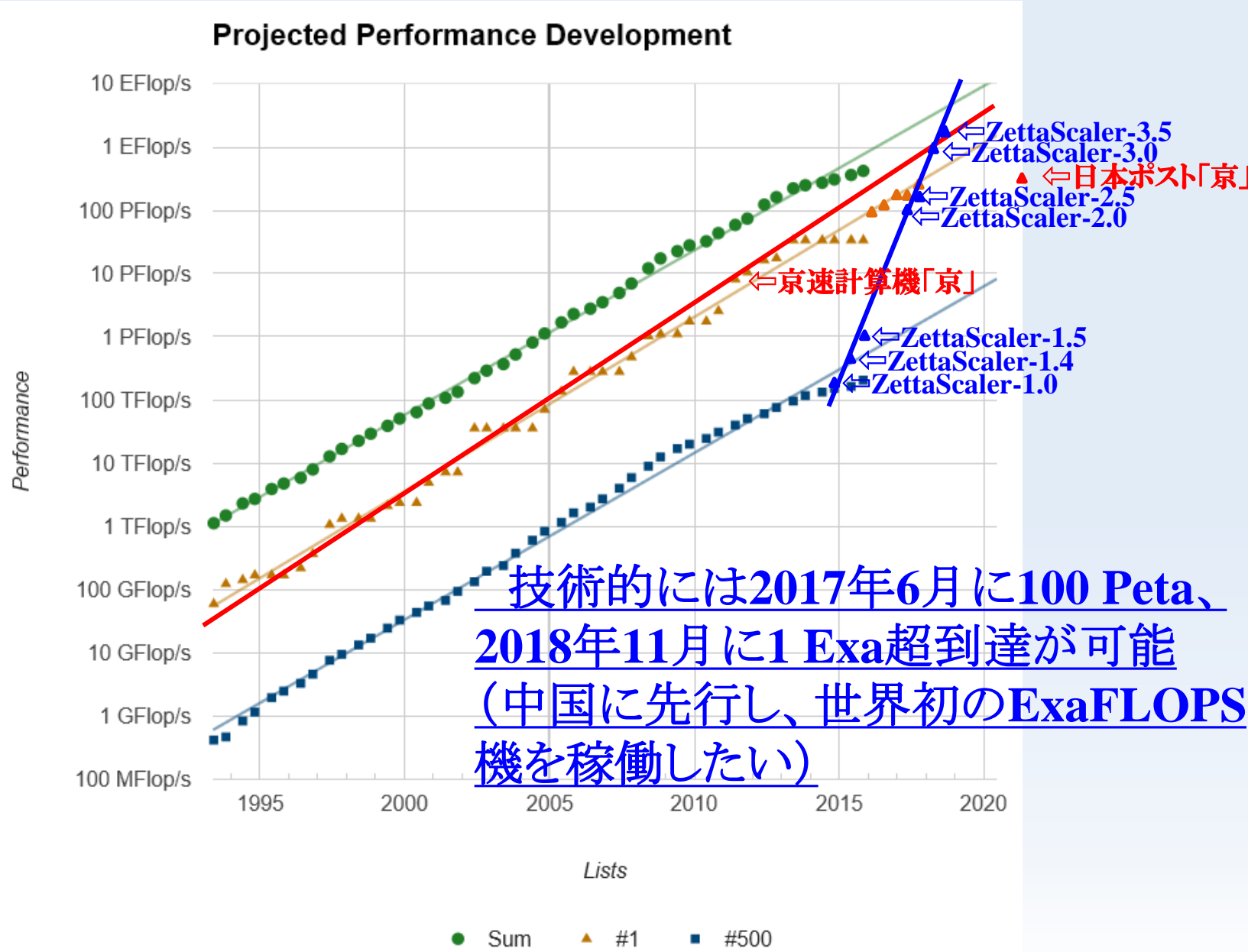
新人工知能エンジン開発

2014年に開発した1,024コア「PEZYZ-SC」



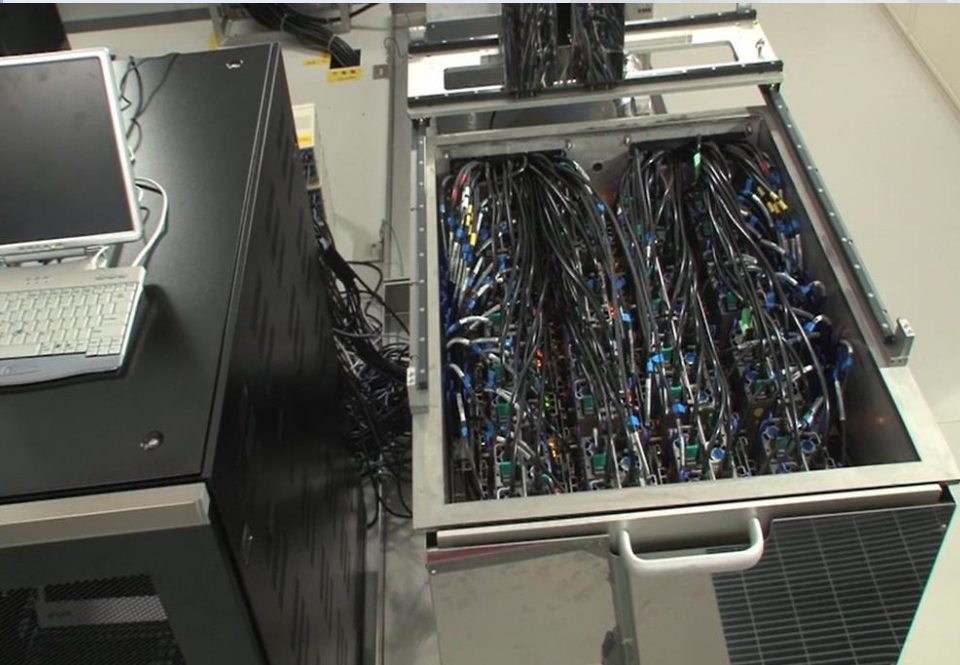
僅か2cm角の半導体に、世界最多となる1,024個の独立演算コアを集積

小型液浸スパコンで性能向上を再加速



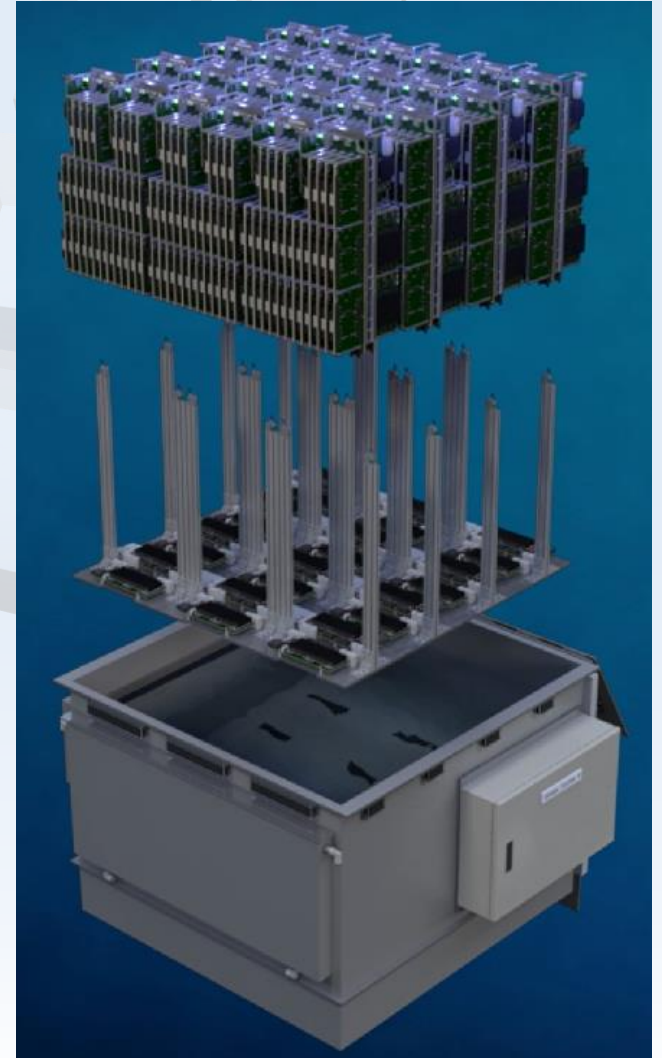
次世代スパコンの小型化・高性能化

- 設置規模(床面積・体積)は1/1000に
- 10 PetaFLOPSの京速計算機「京」が、僅か1台の液浸槽に
- 以下のスパコンと室外冷却装置による構成で
 - 2016年 6月: 0.25 PetaFLOPS (ZettaScaler-1.6)
 - 2017年 6月: 1.50 PetaFLOPS (ZettaScaler-2.0)
 - 2018年11月: 8.00 PetaFLOPS (ZettaScaler-3.0)
 - 2020年11月: 20.00 PetaFLOPS (ZettaScaler-4.0)



現在試験中の試作機「ZettaScaler-1.8」

1 PetaFLOPS/m³を超える性能密度 (世界初)



大規模スーパーコンピュータ開発へ

JSTトップ > プレス一覧 > 科学技術振興機構報 第1240号

平成29年1月20日

科学技術振興機構報 第1240号

東京都千代田区四番町5番地3
科学技術振興機構(JST)

産学共同実用化開発事業(NexTEP)未来創造ベンチャータイプ 平成28年度緊急募集における新規課題の決定について

JST(理事長 濱口 道成)は、産学共同実用化開発事業(NexTEP)未来創造ベンチャータイプ平成28年度緊急募集における新規課題2件を決定いたしました(別紙)。

NexTEPは、大学などの研究成果に基づくシーズを用いた、開発リスクが高く規模の大きい開発を支援し実用化を目指す技術移転支援事業です。

NexTEPでは、開発を行う企業などに開発費(原則、総額1億円以上50億円以下)を支出し、開発が成功した場合は支出した開発費の全額の返済を求め、開発不成功の場合は開発費支出の10%の返済を求めます。JSTが開発のリスクを負担することで、大学などの研究成果に基づくシーズの開発を促進します。

今回の募集は、平成28年10月12日(水)から10月25日(火)まで行い、2件の応募がありました。募集締め切り後、外部専門家の協力のもと、評価委員会にて課題の独創性(新規性)および優位性、目標設定の妥当性、イノベーション創出の可能性、提案内容の実行可能性、事業化の可能性、開発に伴うリスクなどの観点から審査(事前評価)し、その結果をもとに課題を決定しました。

ホームページURL: <http://www.jst.go.jp/jitsuyoka/index.html>

<添付資料>

別紙: [産学共同実用化開発事業\(NexTEP\)未来創造ベンチャータイプ平成28年度緊急募集における新規課題一覧](#)

参考: [産学共同実用化開発事業\(NexTEP\)未来創造ベンチャータイプについて](#)

<お問い合わせ先>

科学技術振興機構 産学共同開発部

大竹 利也(オオタケトシヤ)、関谷 亮英(セキヤ アキヒデ)

〒102-0076 東京都千代田区五番町7 K's五番町

Tel: 03-6380-8140 Fax: 03-5214-0017

E-mail: jitsuyoka@jst.go.jp

別紙

産学共同実用化開発事業(NexTEP)未来創造ベンチャータイプ 平成28年度緊急募集における新規課題一覧

| 課題名 | 新技術の代表研究者 | 開発実施企業 |
|--------------------------------------|-------------------------|---------------|
| 磁界結合DRAM・インタフェースを用いた大規模省電力スーパーコンピュータ | 慶應義塾大学 理工学部 教授 黒田 忠広 | 株式会社ExaScaler |
| 草本バイオマスの高効率・高品質ガス化装置 | 長崎総合科学大学 名誉教授 坂井 正康 | バイオマスエナジー株式会社 |

人工知能による革命は、まだ序章

ディープラーニングによる機械学習は、「抽象的概念」を既に獲得して、猛烈な勢いで進化中



5年以内に生ずる「新」産業革命

- 人工知能が処理可能であるのは、人間が手間暇を掛けてお膳立て(前処理)をした後の情報が殆どで、前処理には膨大な時間と作業を要する … という大きな誤解(人間の驕り)
- 実は、そうした面倒で複雑な前処理こそ、本来は人工知能が得意とするところ。しかし、現実には十分な処理能力が無いために、人間が行っているだけである。この部分が解決すると、現時点の人間の知的労働・生産の大部分が置換される
- 「旧」産業革命においては、人間や家畜の肉体労働・生産が置換されたのみならず、人間や家畜が行えなかった規模の工事や事業や大型生産設備が登場し、人類が月面にまで到達した。
- 今回は、現在の人間の知的労働・生産の置換を遥かに上回る規模の極めて大きな衝撃に備える必要がある

AIチップ開発が死活問題に

世界各国で、AI・ディープラーニングの開発が進み、米・英・イスラエル・韓ではAIプロセッサ(チップ)の開発が多数進められるが、日本ではソフト開発が中心である

ソフト開発は、新しいアルゴリズムや革新的な演算モデルが発明されると、一瞬で全く価値が消失する可能性が高い

ハードウェア、特にチップ開発が不可欠であることが明白ながら、日本で最先端半導体開発を行えるグループは2-3社しかなくなってしまった

Deep Insights社が日本のAI開発の重責を担う

AIチップ開発が死活問題に

| 社名 | Deep Insights | 富士通 | Preferred Networks | 東芝 |
|-----------|--------------------------|-------------------------------------|--------------------|--------------------------------|
| 完成/出荷時 | 2018年中ごろ | 2018年度中 | 2019年頃 | (研究開発段階) |
| 1チップの目標性能 | 現行GPUの1000倍以上の演算性能を目指す | 他社比10倍の演算性能を目指す | 1ベタ回/秒の演算が可能 | 17ワットの消費電力で48.5テラ回/秒の演算を実証 |
| 演算精度 | 単精度(32ビット)~1ビット演算まで対応 | 単精度(32ビット)をサポート。他の精度への対応は未定 | (未公開) | 1ビット |
| 演算回路数 | 4ビット精度の演算コア100万個を集積 | 単精度(32ビット)の演算コア換算で2万個以上を集積 | (未公開) | 65nmプロセス、1.9mm×1.9mmに32768個を集積 |
| その他の特徴 | 帯域100テラバイト/秒の専用メモリーも開発する | 学習用途を狙うが、推論も可能。Caffeを含む主要なフレームワーク対応 | (未公開) | エッジデバイス向けを想定 |

国内のAIチップ開発は僅か数社が行うのみ
その中でも、最先端の半導体プロセスで開発を行えるのは、富士通とDeep Insights社のみ

AIチップ開発のスタートアップ企業であるGraphcore社(英国)が、Samsung社とBosch社から3,600億円もの出資を受け入れて開発を加速しているなど、世界では大型開発が目白押しであるところ、日本ははるかに遅れている現状に甘んじており、抜本的な対策が必要なことが明らかである

AIエンジンとスパコンの開発方向は真逆

次世代スパコン:

少なくとも倍精度 (DP: 64bit) 浮動小数点演算が必要
今後、4倍精度 (128bit)、8倍精度 (256bit)、16倍精度 (512bit) といった多倍長演算が必要になることに

新AIエンジン:

単精度 (SP: 32bit) 浮動小数点演算が基本
半精度 (HP: 16bit)、1/4精度 (8bit) で良い場合もあり、
1/8精度 (4bit)、1/16精度 (2bit)、更にはビット演算でも
十分な事例も報告が出て来ている
一方で、メモリ帯域とメモリ容量への要望は際限がない

1チップで100万コア、100TB/s、DRAM一体、100Wの
積層型半導体エンジンが必要

⇒ 1年半での製品化を予定

今後の「医療革新」に関する予測

- 1) AI・スパコンが、人間の医師の診断能力を上回る
- 2) AI・スパコンが、既存薬の新しい効能を発見する
- 3) AI・スパコンが、新しい診断法を確立する
- 4) AI・スパコンが、新しい治療法・治療薬を開発する
- 5) AI・スパコンが、新しい病気を発見する
- 6) AI・スパコンが、病気の概念を再定義する
- 7) AI・スパコンが、生殖・成長・老化・進化の神秘を解明する
- 8) AI・スパコンが、医療と生命科学をゼロから再構築する
- 9) AI・スパコンが、生命体のリデザインを進める

最強の科学技術基盤の出現

1,000倍高速な
人工知能エンジン
(仮説の立案)

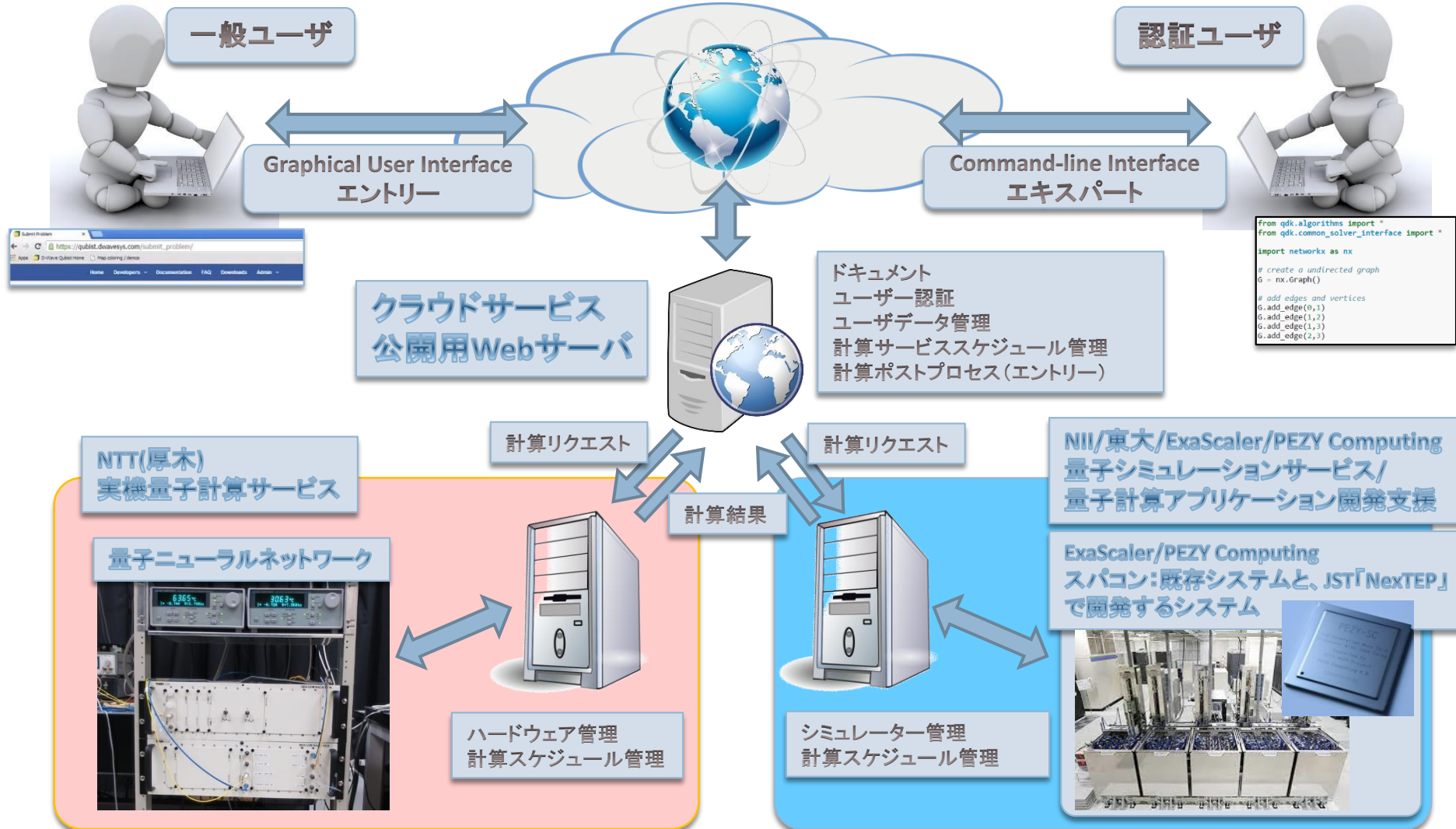
1,000倍高速な
次世代スパコン
(仮説の検証)

「最強の科学技術基盤」

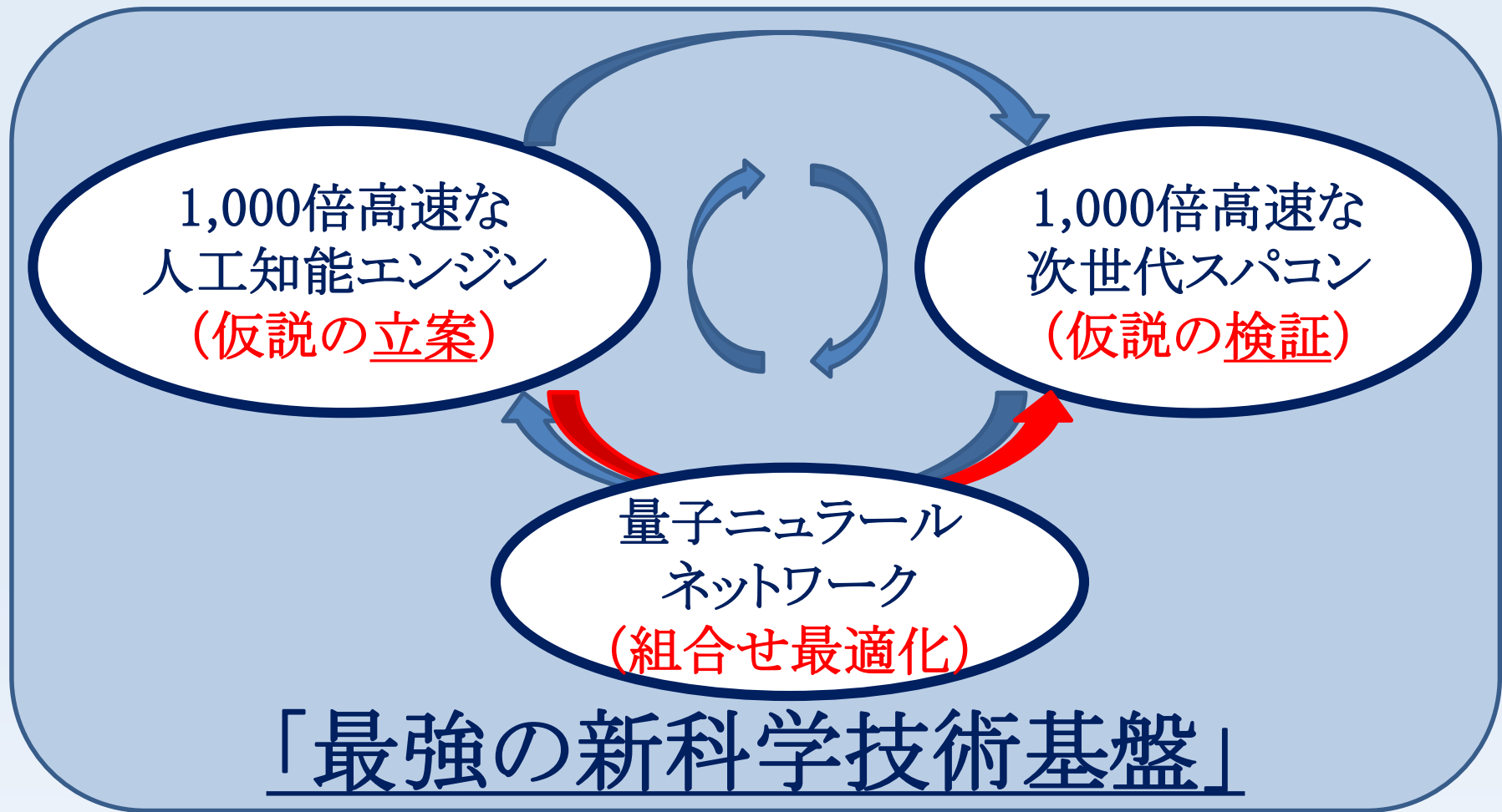
人間には抽出できない複雑で無数の特徴点・特徴量から、更に規則性・法則性が抽出されることで膨大な仮説が立案され、それらが検証されることで、人間に構築できない次元の理論が、多数生まれてくることに

(ノーベル賞級ではなく、「新しいノーベル賞」が幾つも創設されるレベルの理論が次々と産み出される)

量子ニューラルネットワークで共同研究



最強の新科学技術基盤の出現



人間には抽出できない複雑で無数の特徴点・特徴量から、更に規則性・法則性が抽出されることで、膨大な仮説が立案され、それらが高速に検証され、最適化されることで、人間には決して構築できない次元の理論が、多数生まれることに

御清聴、有難う御座いました。

