

- <https://cs-forum.github.io/hpci-aplfs/roadmap-2014/>
- 創薬・医療、総合防災、エネルギー、社会経済科学、生命科学、物質科学、ものづくり、基礎物理などのテーマで、5-10年後に(富岳で計算パワーが100倍程度になったら)どのようなブレークスルーがあり、社会貢献をし得るかをコミュニティが執筆

### 計算科学ロードマップ 概要

～大規模並列計算によるイノベーションの  
目指す社会貢献・科学的成果～

平成 26 年 5 月

将来の HPCI システムのあり方の調査研究  
「アプリケーション分野」

### 今後の計算科学が貢献しうる社会的課題

#### 創薬・医療

#### 画期的創薬・医療技術の創出

#### 従来の研究

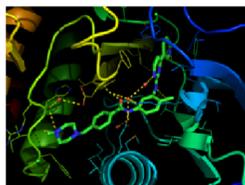
- 小規模なデータ処理
- 個別分野において固有のスケールが進展
- 単純な脳回路等のシミュレーションモデル

#### 今後の科学計算からのアプローチ

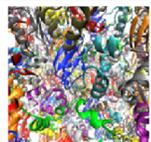
- DNAシーケンサーから得られる大規模データによる遺伝子ネットワーク解析
- 細胞環境下での創薬
- 幅広い時空間にまたがる階層でのモデルの連成
- モデルの大規模化・高精度化
- 詳細な脳神経回路シミュレーションとデータ同化

#### 社会への貢献

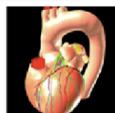
- 個人の遺伝情報に基づき患者個人に最適な治療法を提供するテーラーメイド医療の実現
- 新薬開発の短期化、低コスト化
- 負担が小さい治療の実現による患者の生活の質の向上、早期社会復帰による社会の活性化、医療費の低減



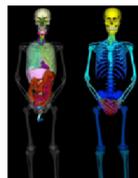
タンパク質と薬の結合



細胞環境下での創薬



臓器の精密シミュレーション(京)



全身スケールシミュレーション

今後のスーパーコンピュータがもたらす莫大な計算能力が、神経系や細胞の詳細なシミュレーション、幅広い時空間にまたがるシミュレーション、そしてそれらのリアルタイムに近いデータ同化<sup>4</sup>などさまざまな面で生命分野の発展に大きく寄与することは間違いなく、ひいては画期的創薬・医療技術創出の重要な科学基盤となり得るものである。

創薬・医療分野において今後、必要となる計算機性能を下表に示す。

| 課題   | 要求性能 (PFLOPS) | 要求メモリ (PB/a) | メモリ量/ケース (PB) | ストレージ量/ケース (PB) | 計算時間/ケース (hour) | ケース数    | 総演算量 (EFLOP) | 概要と計算手法                                     | 問題規模   | 備考  |
|--|---------------|--------------|---------------|-----------------|-----------------|---------|--------------|---|--|---|
| 個人ゲノム解析  | 0.0054        | 0.0001       | 1.8           | 0.1             | 0.7             | 200000  | 2700         | シーケンスマッチング                                  | がんゲノム解析200,000人分のマッピングおよび変異判定  | 1人分の解析を1ケースとした。入力データを分割することで、細かい単位での実行、誤差をまたいだ実行も可能。整数演算中心のため「総演算量」はInstruction数とした。総演算小数は演算量は45.864EFLOPとなる。 |
| 通信ネットワーク解析   | 25            | 89           | 0.08          | 0.016           | 0.34            | 26000   | 780000       | ベイジアンネットワークおよびL1正則化                         | 4万転写物×26,000データセット×280万アレイ   |   |
| 創薬などMD・自由エネルギー計算                                     | 1000          | 400          | 0.0001        |                 | 0.0012          | 1000000 | 4300000      | 全原子分子動力学シミュレーション                            | ケース数:10万化合物×10種の蛋白質(10万原子程度)   | B/F=0.4。数百から数千ケース同時に行うことを想定している。実行時に必要なメモリ量、各ケースの実際の実行時間は、表の数値の数倍となる。メモリ量/ケースは100ノード実行時を想定。                   |
| 細胞環境・ウイルス  | 490           | 49           | 0.2           | 1.2             | 48              | 10      | 850000       | 全原子/粗粒化分子動力学シミュレーション                        | ~1億粒子  | B/F=0.1   |
| 細胞内信号伝達経路シミュレーション                                    | 42            | 100          | 10            | 10              | 240             | 100     | 3600000      | 分子種細胞シミュレーション (格子法)                         | 1000 から 10,000 細胞で構成される細胞集団  | 格子法。整数数の演算性能を要求。ケース数は最低10回、100回程度が望ましいため100回とした。  |
| 高精度創薬  | 0.83          | 0.14         | 1             | 0.001           | 1               | 100     | 300          | 薬品とタンパク質間相互作用の量子化学計算                        | 水和条件下、500種タンパク質+リガンド   | ファイル/OIは終了時に1TBを1秒で書き出すことを想定し、1TB/sが必要とした。  |
| バイオデバイス設計  | 1.1           | 0.19         | 1             | 0.001           | 1               | 100     | 400          | 200-500種程度のタンパク質の分光計算                       | 電子軌道数10万程度   | ファイル/OIは終了時に1TBを1秒で書き出すことを想定し、1TB/sが必要とした。  |
| 血流シミュレーション   | 400           | 64           | 1             | 1               | 170             | 10      | 2500000      | 差分法、準解法(構造・流体・生化学連成シミュレーション)                | 100mm長×100mm径、0.1mm格子、流速10 <sup>-2</sup> m/s、解像度1us、10秒  |   |
| 超音波シミュレーション  | 380           | 400          | 54            | 64              | 240             | 10      | 3300000      | 差分法、準解法(音波・熱シミュレーション)                       | 400mm <sup>3</sup> の計算領域を数細胞とマイクログラフセル千分音響場を捉えるため、225兆点の格子と時間ステップ数として1459200ステップが必要である。また、1格子点あたり演算数1000程度となる。 |   |
| 脳神経系シミュレーション(ヒト全脳解剖モデル)                              | ※             | 6.9          | 7.8           | 56              | 3600            | 0.28    | 100          | 単一コンパートメントHPCシノプス可塑性+通信                     | 1000億ニューロン、ニューロンあたり1万シノプス10 <sup>5</sup> step  | ネットワークのボトルネックはレイテンシー  |
| 脳神経系シミュレーション-昆虫全脳神経モザイク-神経回路マシナリー推定・生理実験とシミュレーションの連成 | ※             | 71           | 60            | 0.2             | ※               | 28      | 20           | マルチコンパートメントHPC(局所クランクニコルソン)シノプス通信 進化的アルゴリズム | 1000億ニューロン 10 <sup>6</sup> 億格子 100世代   | 100MB/s程度の外部との通信も想定   |

分野の主要なスパコン上のシミュレーションのアプリケーションプログラムに関し、用いられるアルゴリズム、要求されるシステムのパラメータ(計算性能、メモリ帯域など)、主要なアルゴリズムなどを、フォーマットに従って記載

※ムーアの法則：「半導体の集積密度は18か月から24か月で倍増する」という経験則

## ● 2030へ向けた変遷 (Post-Mooreの入り口)

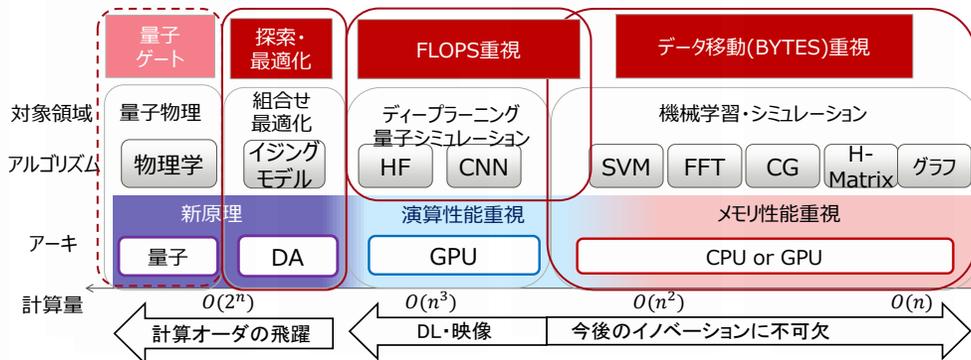
- ・ ムーア則の終焉による、演算性能の進化の終焉
- ・ 新デバイス・パッケージングによるデータ移動コストの革新的削減
- ・ アルゴリズムや利用法の進化による、計算量オーダーの削減(+データ移動の相対的な要求の増加)
- ・ BD/AI/第一原理の融合による、本誌的なデータ中心の流れ

計算の分類(領域)

2022の現状

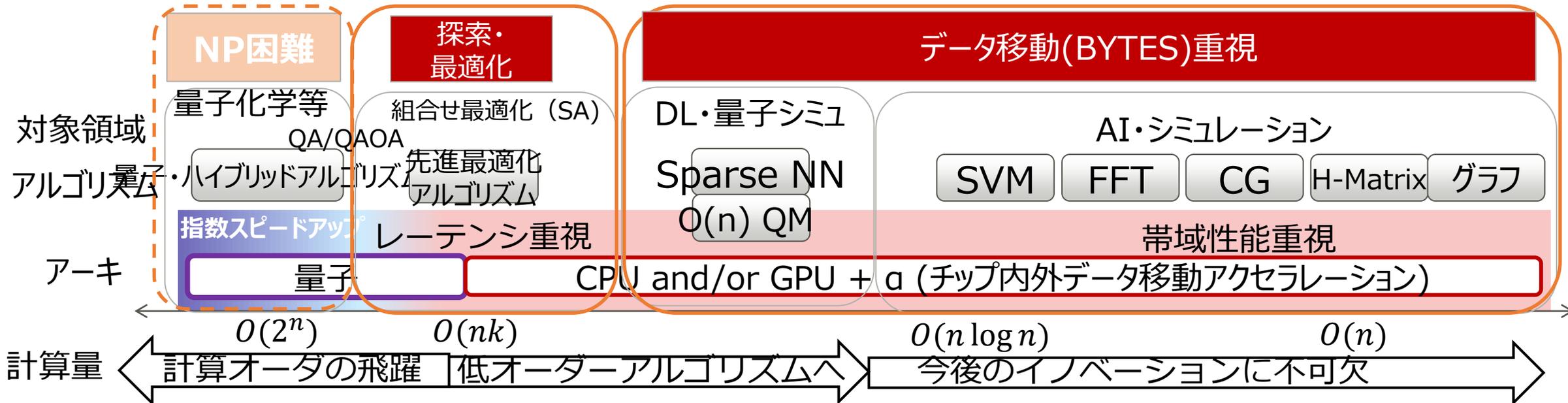
FUJITSU

- CPUでのデータ処理におけるノイマンボトルネックの解消が課題に
- ヘテロ・特化型(=汎用CPU+特化)は使いこなしが困難



### 量子の未来(Hybrid)

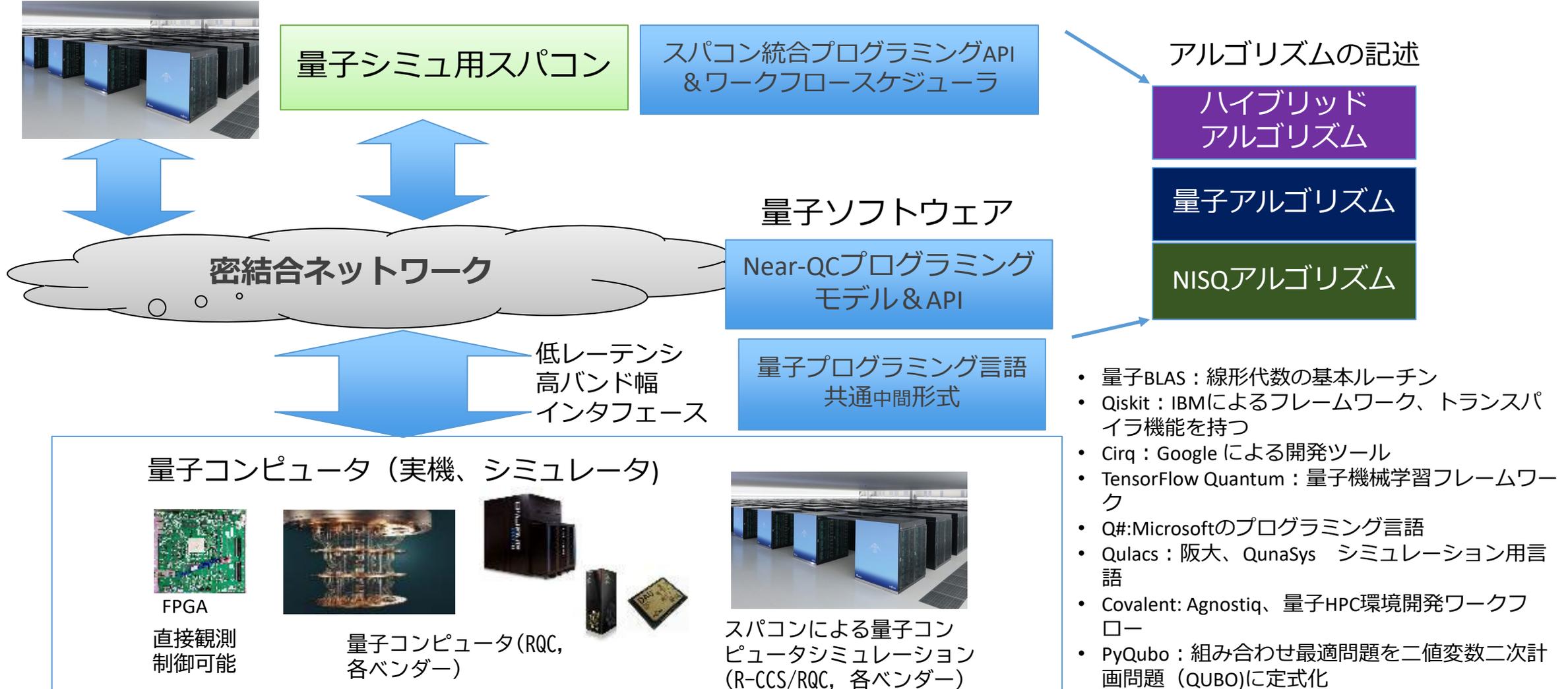
### 古典の未来(Fugaku NEXT, FugakuNEXT^2)



# 理研TRIPプロジェクトにおける古典・量子ハイブリッド計算環境 (理研 R-CCS, RQC w/iTHEMS, AIP、他センター、FS新計算原理etc.)

古典スパコン環境

量子スーパーコンピューティングに  
必要なソフトウェア階層を試作・検討



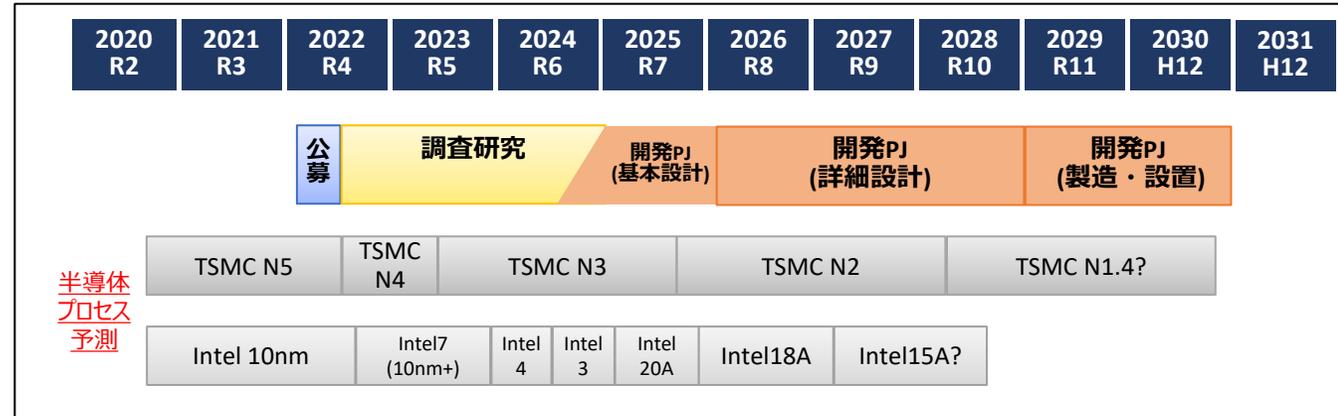
- 量子BLAS : 線形代数の基本ルーチン
- Qiskit : IBMによるフレームワーク、トランスパイラ機能を持つ
- Cirq : Google による開発ツール
- TensorFlow Quantum : 量子機械学習フレームワーク
- Q#: Microsoftのプログラミング言語
- Qulacs : 阪大、QunaSys シミュレーション用言語
- Covalent: Agnostiq、量子HPC環境開発ワークフロー
- PyQubo : 組み合わせ最適問題を二値変数二次計画問題 (QUBO)に定式化

- **更なるAI-5.0・デジタルツインの高度化・拡大・普及を目指す基盤および技術的礎として**
  - →第一原理シミュレーション・大規模かつ「デジタルツイン」向けの柔軟なAI・リアルタイムを含む大規模データ同化をシームレスに行えるスパコンは？
  - 将来の量子計算との融合も見据える（理研TRIPなど）
- **我が国も関与する最先端のデバイス・パッケージングを前提**
  - 我が国の半導体戦略との整合→他省庁や米国(米国エネルギー省や産業界)との連携
- **ムーア側の減速に伴う、アーキテクチャ・アルゴリズム・ソフトウェアの変化**
  - デジタルツインの種々のシナリオ・ワークフローとの新たなコ・デザイン
- **世界のITプラットフォーム・エコシステムへの技術的訴求**
  - 世界レベルでの主要プレイヤーとのパートナーシップ
  - 「ベンチマーク一位」「AI(のみ)でリード」などの一点豪華主義傀儡目標は立てない

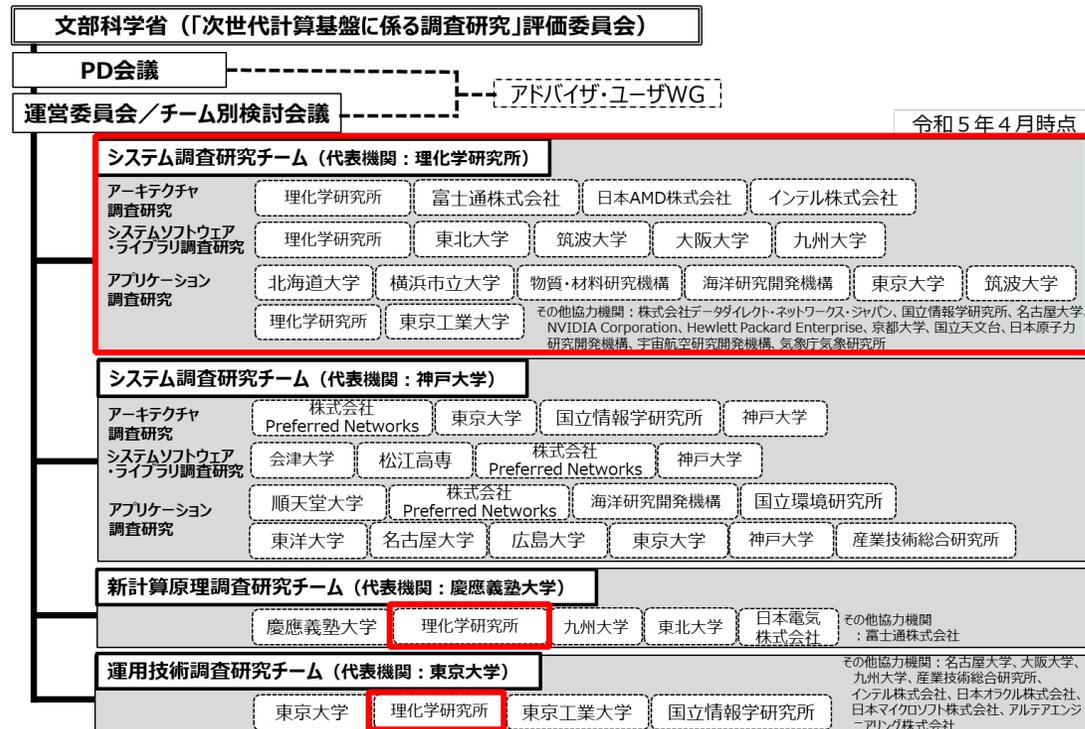
## ● 文部科学省「次世代計算基盤に係る調査研究」事業

- 実施期間：令和4～6年度
- 令和5年度予算額：10.2億円  
(令和4年度予算額：4.3億円)
- ポスト「富岳」時代の次世代計算基盤の戦略的整備に向けた調査研究→**実現可能なシステムの選択肢の提示**
  - 要素技術の研究開発（日本で開発する独自技術の特定）
  - 技術課題や制約要因の抽出

## 富岳Next スケジュール予測



## 「次世代計算基盤に係る調査研究」実施体制



## 取組概要

SDGs・Society5.0の実現に向けた課題解決のためのプラットフォームとして、今後の科学に「研究DX」をもたらす**高度なデジタルツイン実現の基盤**となる汎用性の高い次世代計算基盤実現を目指し、あるべきアーキテクチャやシステムソフトウェア・ライブラリ技術について、アプリとのコデザインを通じた調査研究を行う。

**電力制約の下でデータ移動を高度化・効率化する“FLOPS to Byte”指向のシステム構築**を、アーキテクチャ開発からアルゴリズム設計、アプリケーション技術に至るまで実践する。

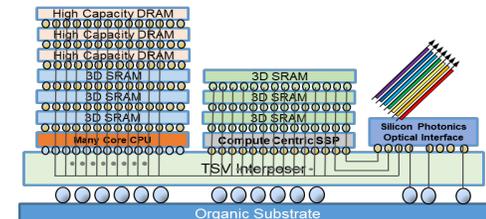
## 調査研究の現状

### アーキテクチャ調査研究

- 複数ベンダーとシステム全体や構成要素の可能性を調査  
(三次元積層メモリ技術、チップ間光接続技術など)
  - アーキ候補の抽出、ベンチマーク評価・解析を実施中
- ### システムソフトウェア・ライブラリ調査研究
- 重用システムソフトウェアの富岳へのポータビリティ評価を計画  
(OneAPI, DAOS, Cuda Quantumなど)

### アプリケーション調査研究

- アーキ評価可能なベンチマークセットを構築、将来アプリの調査  
(将来的なCI/CDベンチマーキング環境構築の準備)



プロセッシング要素の要素技術例

# 参考資料

## 理化学研究所の国内拠点



## R-CCS研究センター + データセンター



計算機室 50 m x 60 m = 3,000 m<sup>2</sup>  
**電力消費 最大 37 MW**  
 水冷システム

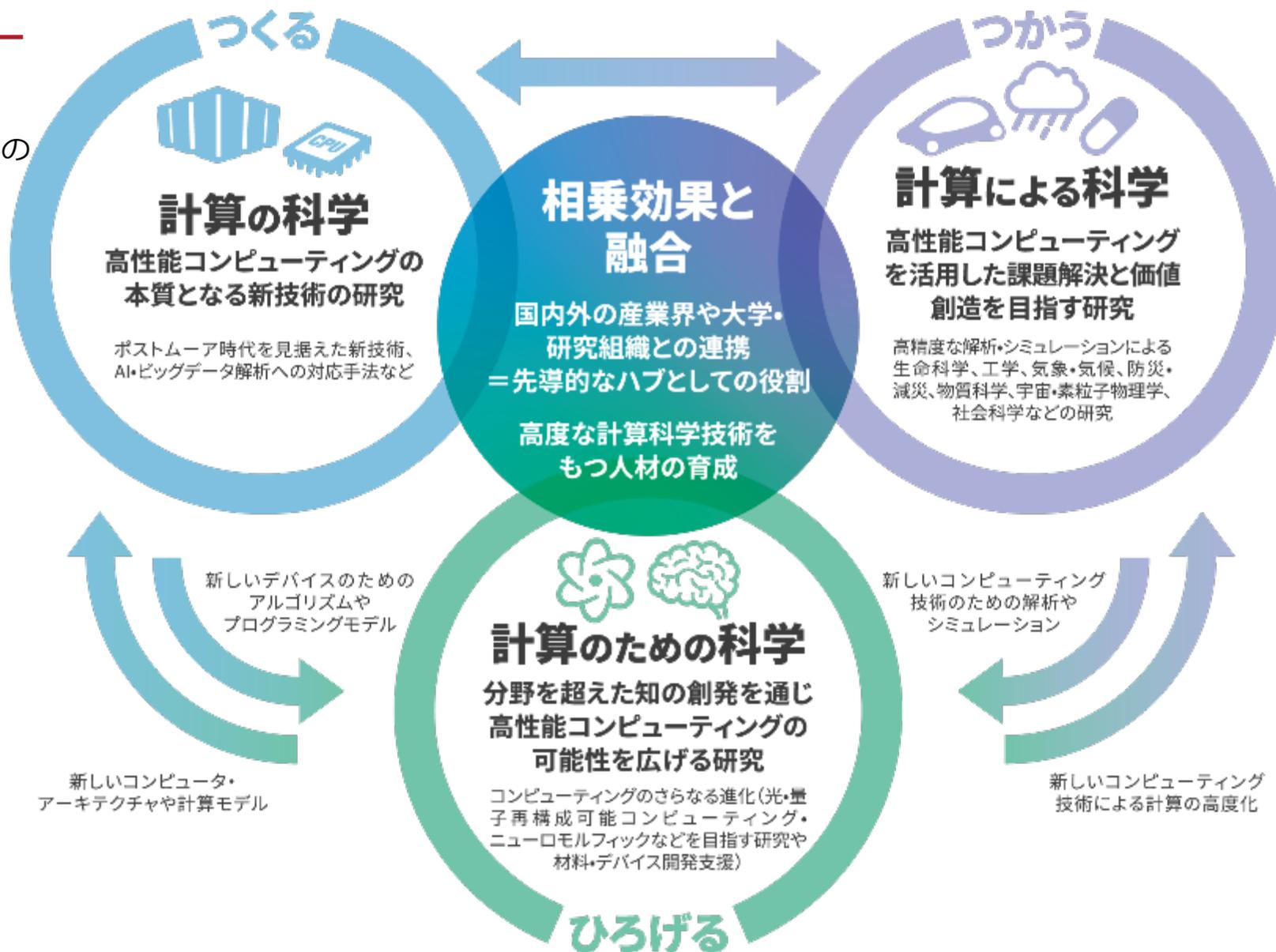
ガスタービンコージェネレーション設備  
 5 MW x 2

# 「計算の 計算による 計算のための科学」

卓越したサイエンスの創出と、Society5.0実現の要となることを目指して

## 理研 計算科学研究センター (R-CCS)

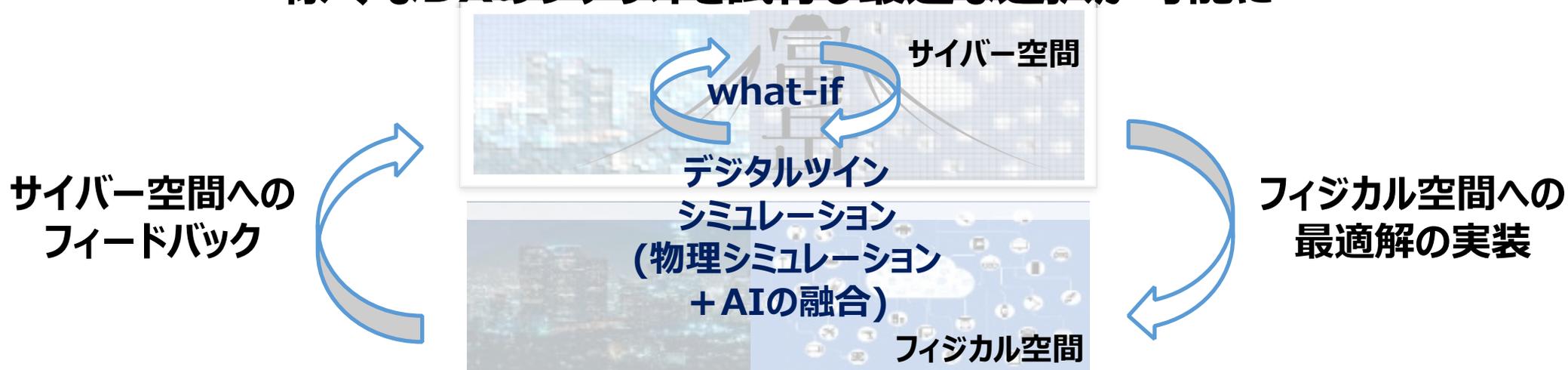
理研の13研究センターの一つ  
であると同時に高性能計算科学の  
トップの国家拠点



## Society5.0におけるDX

- 単に従来ある単純現象(サイエンスの現象からビジネスにおける業務まで)をデジタルで置き換えるだけではない
- 「実験では観測できない」、「通常のビジネスではハイリスクで行えない」、などのシナリオをデジタルツインでwhat-ifで数多く再現し、最適なものを選択可能
- **デジタルツインは、物理シミュレーションによるものとAIによるもののコンバージェンスが大事**

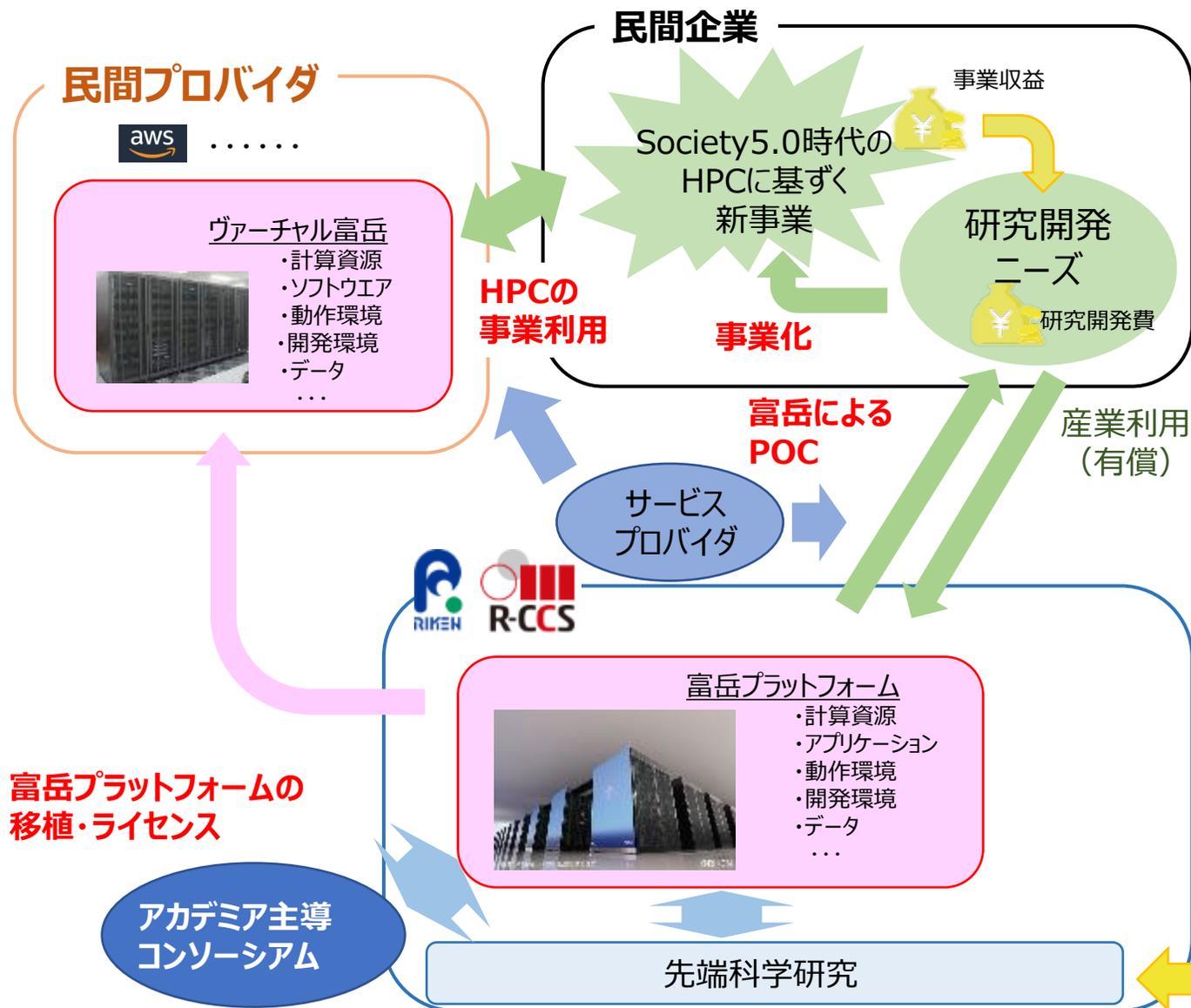
様々なDXのシナリオを試行し最適な選択が可能に



産学官デジタルツイン基盤

# ヴァーチャル富岳による持続的「事業化」環境の提供

～富岳におけるPOCの後、ヴァーチャル富岳で事業化～

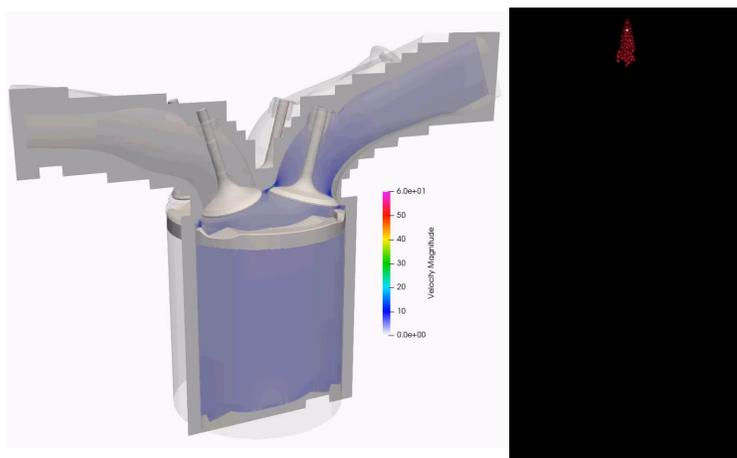


高速かつ大量にシミュレーションモデルを作成し、スパコン性能を最大限に活用できるソフトウェア

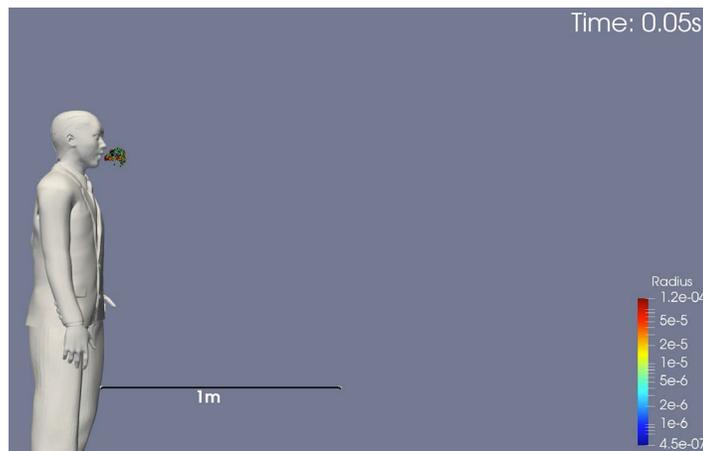
カーボンニュートラル社会と対コロナ飛沫感染シミュレーション

- 2012年より理化学研究所で産学連携で開発。
- スパコン「京」を活用して、自動車、燃焼システム、建築防災分野で多くの実績。
- 2020年初頭、Society5.0時代のものづくりへ向けて富岳でチューニング中に、新型コロナパンデミックが発生。

「京」による自動車エンジンシミュレーションと燃料噴霧噴射の様子



「富岳」による飛沫飛散シミュレーション



HPCを活用した自動車次世代CAEコンソーシアム



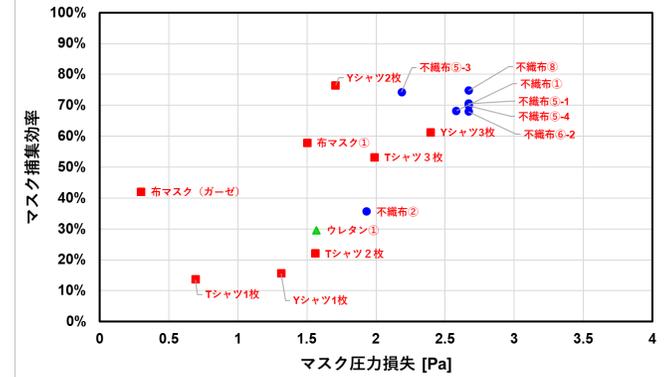
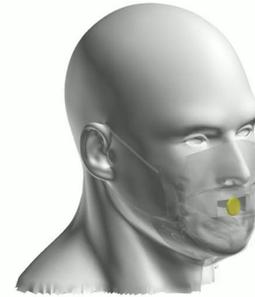
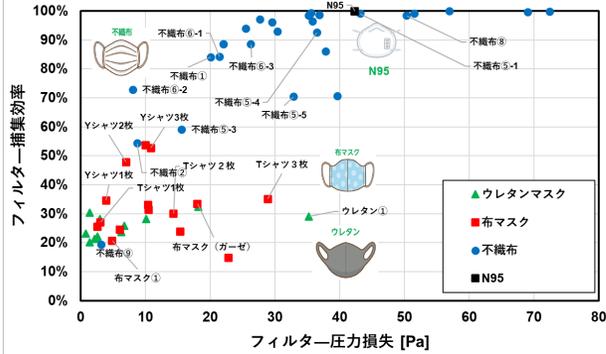
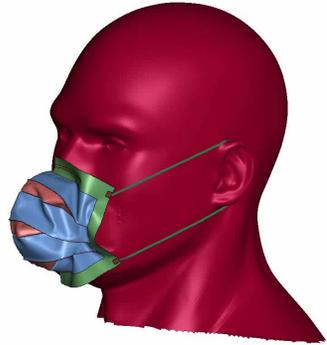
燃焼システム用次世代CAEコンソーシアム



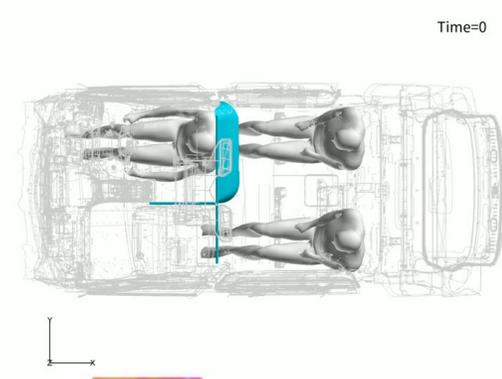
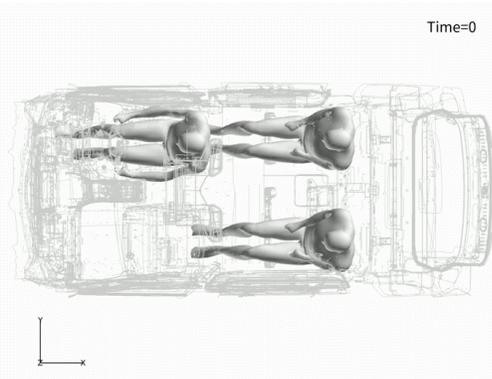
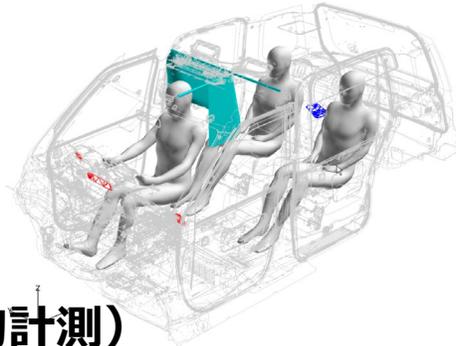
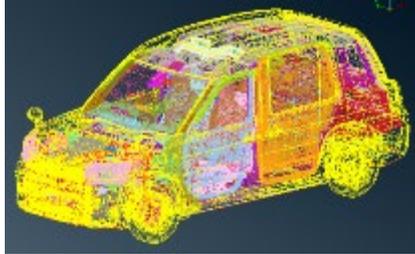
都市・建築CFD コンソーシアム (東京工業大学)



- マスクシミュレーション (シミュレーションによる形状データ自体の作成 + 実験による必要データの計測)



- 公共交通機関での感染リスク評価 (CAD形状データの活用)



- 室内での感染リスク評価 (実物計測)

