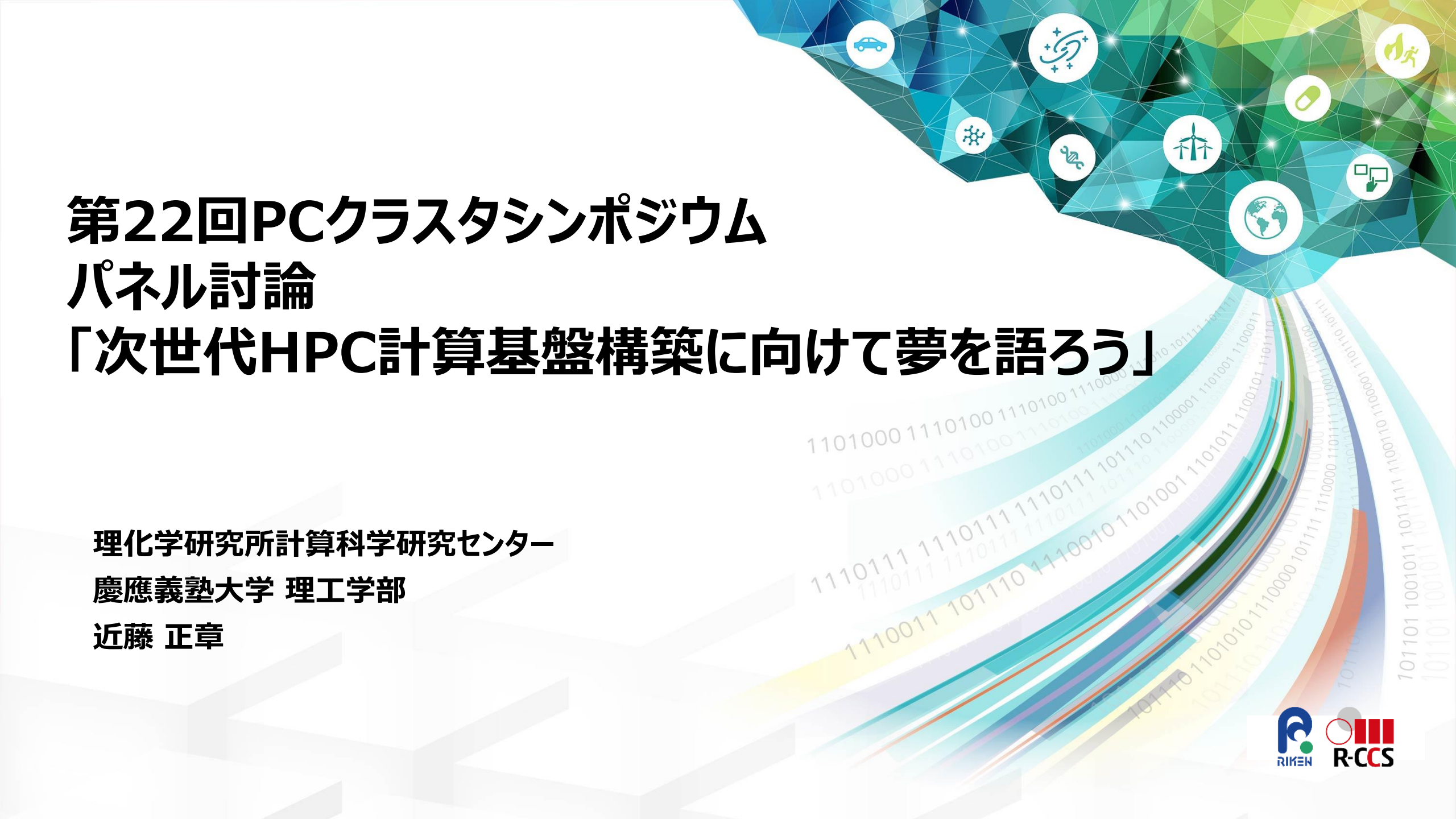




第22回PCクラスタシンポジウム パネル討論 「次世代HPC計算基盤構築に向けて夢を語ろう」

理化学研究所計算科学研究センター
慶應義塾大学 理工学部
近藤 正章

次世代計算基盤に係る調査研究

令和4年度予算額

4.3億円（新規）

背景

- ◆ データ駆動型科学が重要視される中で、シミュレーションやAI 等が連携した研究の重要性がより一層高まっている。さらに、世界的にも研究活動のデジタルトランスフォーメーション（研究DX）の必要性が高まっている。
- ◆ スーパーコンピュータのみならず、データセンターからエッジコンピューティング、それらを繋ぐネットワーク等、様々な形態の社会情報基盤がますます重要となっており、また、これらの基幹技術を自国で保有することは経済安全保障の観点からも重要である。
- ◆ これらの情勢を踏まえると、ポスト「富岳」時代の次世代計算基盤を、国として戦略的に整備することは必要不可欠である。

次世代計算基盤検討部会 中間まとめ（令和3年8月）

◆ 次世代計算基盤検討の留意事項

技術動向や周辺状況が急速に進化・変化

ムーアの法則の終焉等、関連技術が転換期にある、性能の向上に伴い要求される電力量も増大

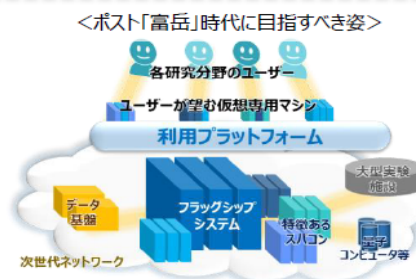
⇒ 半導体やネットワーク等国内外の周辺技術動向や利用側のニーズの調査、要素技術の研究開発等必要な調査研究を行い、多角的な検討が必要。



◆ 次世代計算基盤の在り方

次期「フラッグシップシステム」及び国内の主要な計算基盤、データ基盤、ネットワークが一体的に運用され、総体として持続的に機能する基盤

⇒ 調査研究（FS）を通じ、技術的課題や制約要因を抽出しつつ、実現可能なシステム等の選択肢を提案



次世代計算基盤に係る調査研究

◆ 具体的には以下の取組を実施。

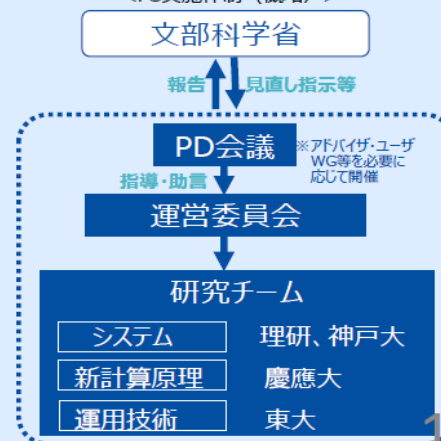
- ・要素技術の研究開発（併せて、我が国として独自に開発・維持すべき技術を特定）
- ・評価指標の検討（例：演算性能、電力性能比、I/O性能、コスト、運用可能性、生産性（アプリケーションのしやすさ）、商用展開・技術展開、カーボンニュートラルへの対応 等）
- ・技術的課題や制約要因の抽出 等

◆ 実施期間：令和4年度～令和5年度 ※令和6年度以降の取組は、調査研究の進捗を踏まえ検討

令和5年度の取組：システム候補の性能評価、アプリケーションのコードデザイン、新たな計算原理を適用すべき領域・分野の検討、多様な計算基盤の一体的運用、これらにおいて必要な要素技術の研究開発 等

令和4年度の取組：技術や利用分野の動向調査、評価項目・手法の検討 等

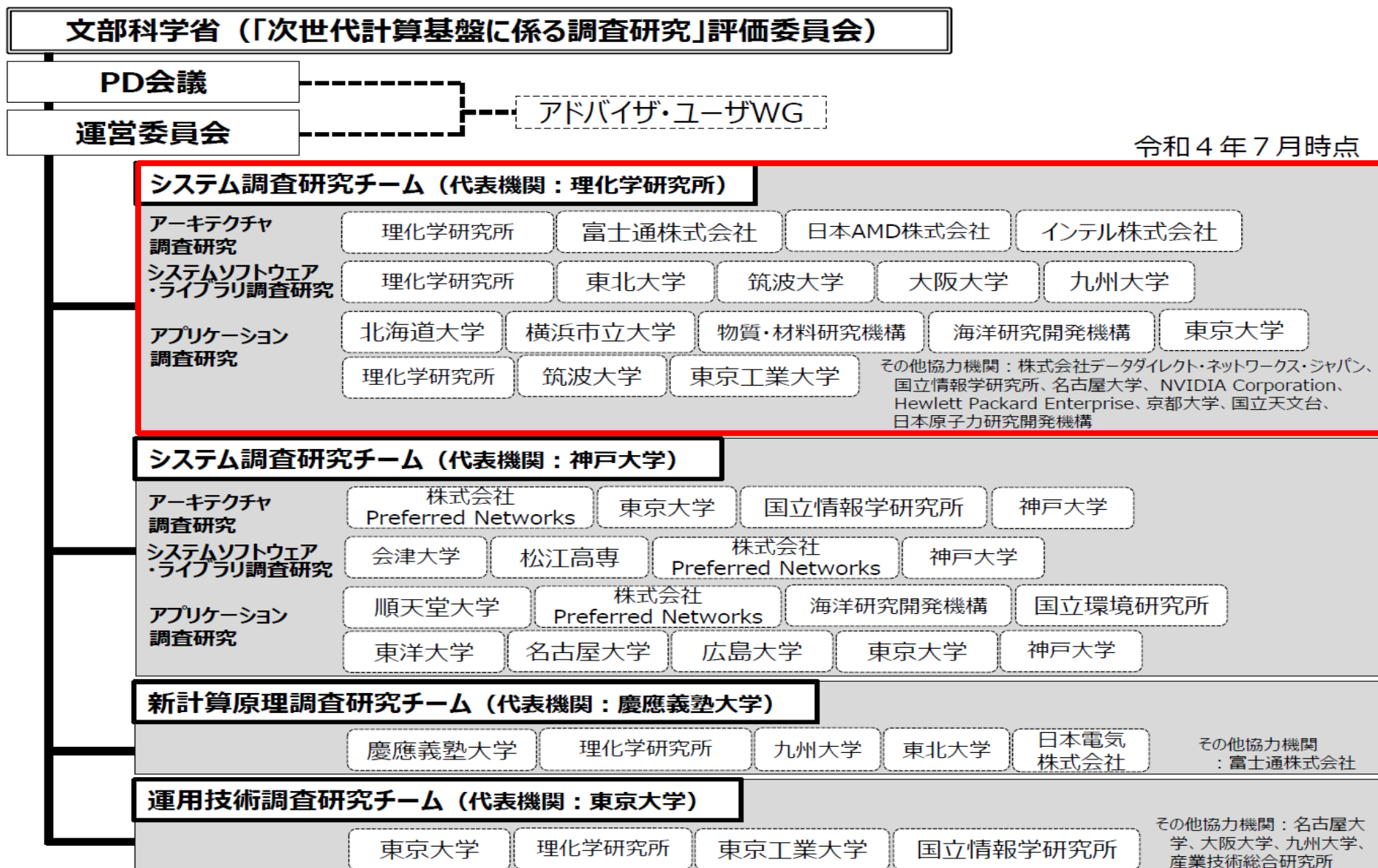
<FS実施体制（概略）>



出典：文部科学省

次世代計算基盤に係る調査研究（FS）の全体推進体制

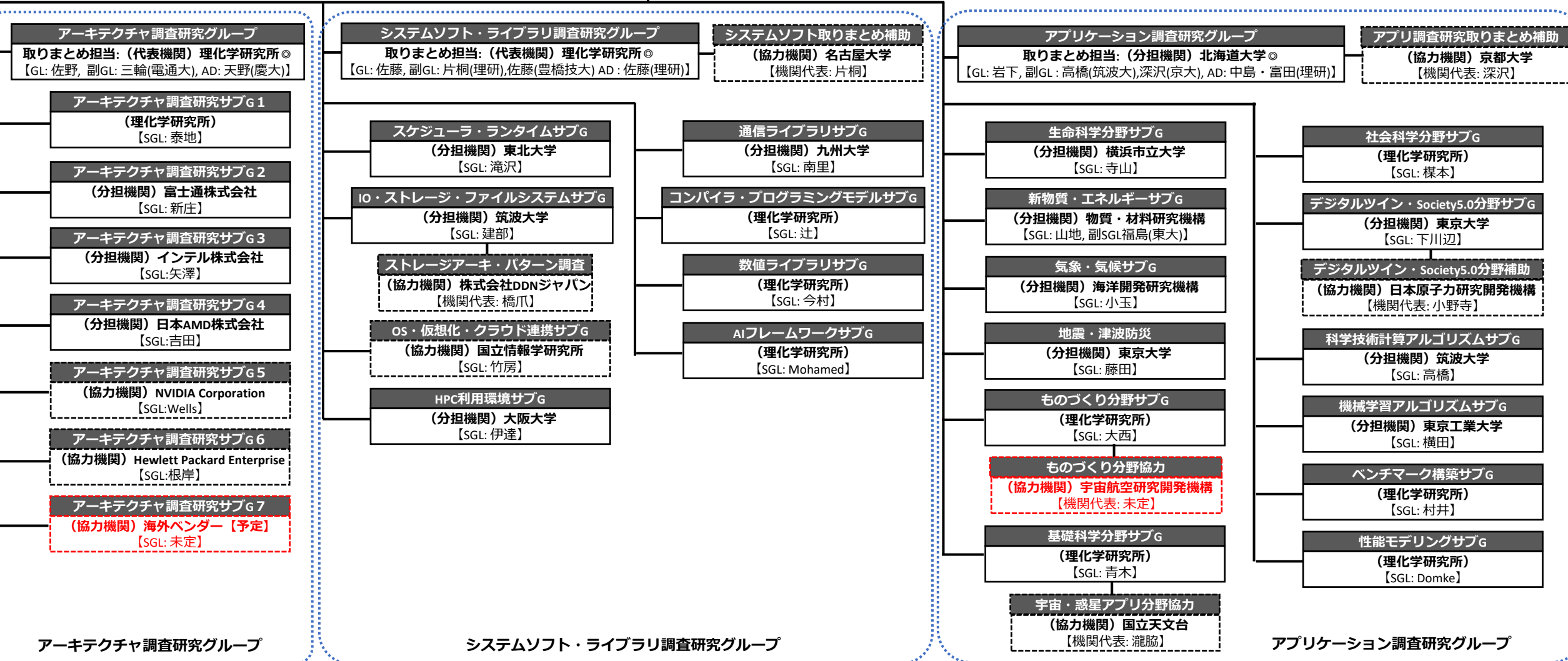
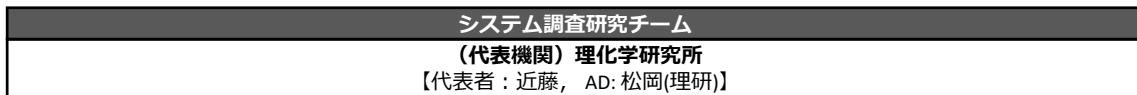
「次世代計算基盤に係る調査研究」実施体制



出典：文部科学省

研究体制：ALL Japan+メジャー国内外ベンダーによる実施体制

GL: グループリーダー
AD: アドバイザー
SGL: サブグループリーダー
※赤字は開始後に追加予定機関



取組概要

次世代計算基盤には、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の役割が求められる。そこで、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**として、**広範な計算手法・シミュレーション技法や大規模データを駆使しつつ、それらが密に連携しながら全体のワークフロー実行が可能**な汎用性の高い計算基盤の実現を目指し、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、アプリケーションとのコデザインを通じた調査研究を行う。

特に、システム設計の基本理念として演算精度も考慮しながら必要な計算性能を確保し、**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。

ALL Japan体制のもと、実効的な性能を向上させる次世代計算基盤のシステム構成や要素技術の調査検討、要素技術の開発を、アーキテクチャ・システムソフトウェアとアプリケーションとのコデザインを通じて実施する。



調査内容

アーキテクチャ調査研究

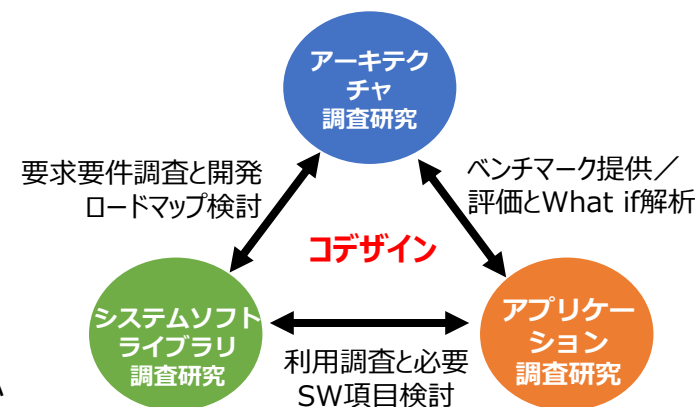
- 半導体技術・パッケージング技術動向を踏まえつつ、**システム全体やその構成要素について考え得る技術的可能性や総合性能を調査**（新規要素技術として三次元積層メモリ技術、強スケーリング・計算インテンシブ向けアクセラレータ、チップ間直接光通信技術などを意識）
- アプリケーション調査研究G提供の**ベンチマークセットの性能解析に基づき将来システムの性能を予測**、また次世代アプリ開発へとフィードバック

システムソフトウェア・ライブラリ調査研究

- 従来のソフトウェアに加え、データ活用促進、機械学習技術と第一原理シミュレーションや大規模リアルタイムデータ処理との高度な融合、高セキュリティの担保などを主要検討項目と据え、**国内で開発すべきソフトを優先度も含めて明らかにしつつ今後の開発ロードマップを策定**

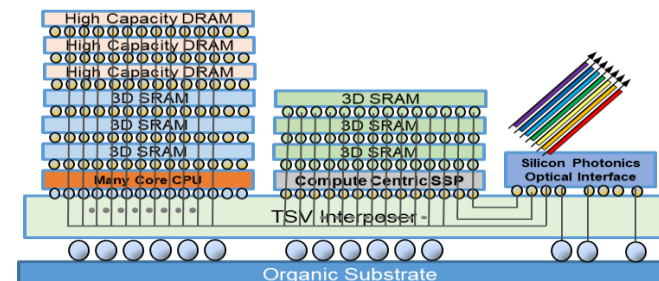
アプリケーション調査研究

- **複数アーキテクチャを統一的に評価するための広範なベンチマークセットを構築**し、それを利用したアーキテクチャ評価結果を踏まえてアルゴリズムやパラメータの改善を検討しつつベンチマークセットを更新しながら、**性能モデルも利用した探索的な“What if”の視点で解析**を行う
- 次世代計算基盤に向けてどのようなアルゴリズムのクラスが大幅な進化が見込まれるか、の指標も抽出して今後のアプリ発展につなげる

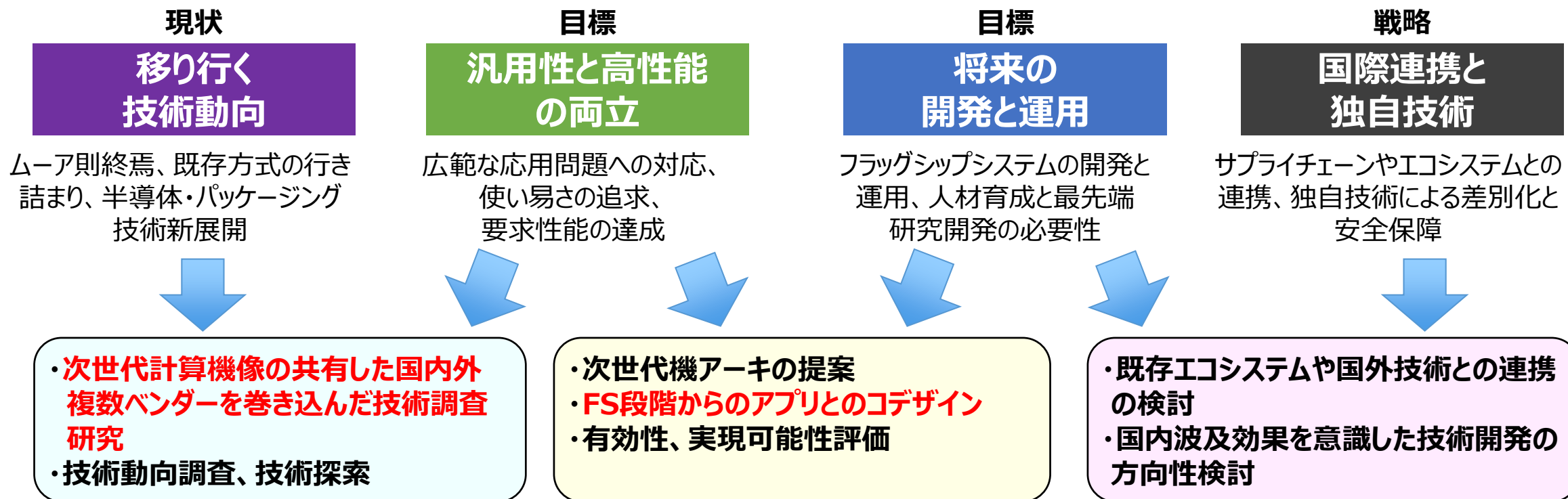


スケジュール

	2022 Q3	2022 Q4	2023 Q1	2023 Q2	2023 Q3	2023 Q4	2024 Q1
アーキテクチャ							
システムソフトウェア							
アプリケーション							
	アーキテクチャ	テクノロジー・アーキ技術の調査・検討	ベンチマーキングによる性能解析/予測	アーキ再検討			
		既存ツールや利用動向調査	新規開発ソフト項目検討・定量的評価	将来の要件整理			
		アプリ調査とベンチマーク設計	ベンチマーク評価に基づく性能分析	アルゴリズム最適化検討			



プロセッシングエレメントの検討要素技術例



調査検討技術項目（案）

・要素技術

- ・ CMOS技術
- ・ パッケージング技術
(3次元積層、チップレット)
- ・ シリコンフォトニクス

・ ノードアーキテクチャ

- ・ 汎用プロセッサ（CPU）
- ・ アクセラレータ
- ・ DRAM（外部メモリ）
- ・ SRAM（キャッシュメモリ等）

・ システムアーキテクチャ

- ・ 相互接続網
- ・ I/O、ストレージ
- ・ 信頼性、可用性、保守性
- ・ 電力効率

- システムソフトに関する技術動向調査と研究開発ロードマップの策定

- システムソフト8分野（下図参照）に渡る網羅的な調査研究

- 加えて分野の横断的に必要な技術（セキュリティ、自動チューニング等）やプラットフォーム化に関する調査

- フラグシップマシンや第二階層システムにおける利用状況、既存のソフトウェアエコシステム、開発工数等を考慮し、国内でどこまでを開発すべきか、もしくは国際協調が必要かを判断し、システムソフトウェアの優先度を決定

- 産業展開を見据えたシステムソフトウェア新領域の調査

- 学術領域でだけでなく産業界を含む幅広い応用分野での活用を促すための新規開発領域の調査

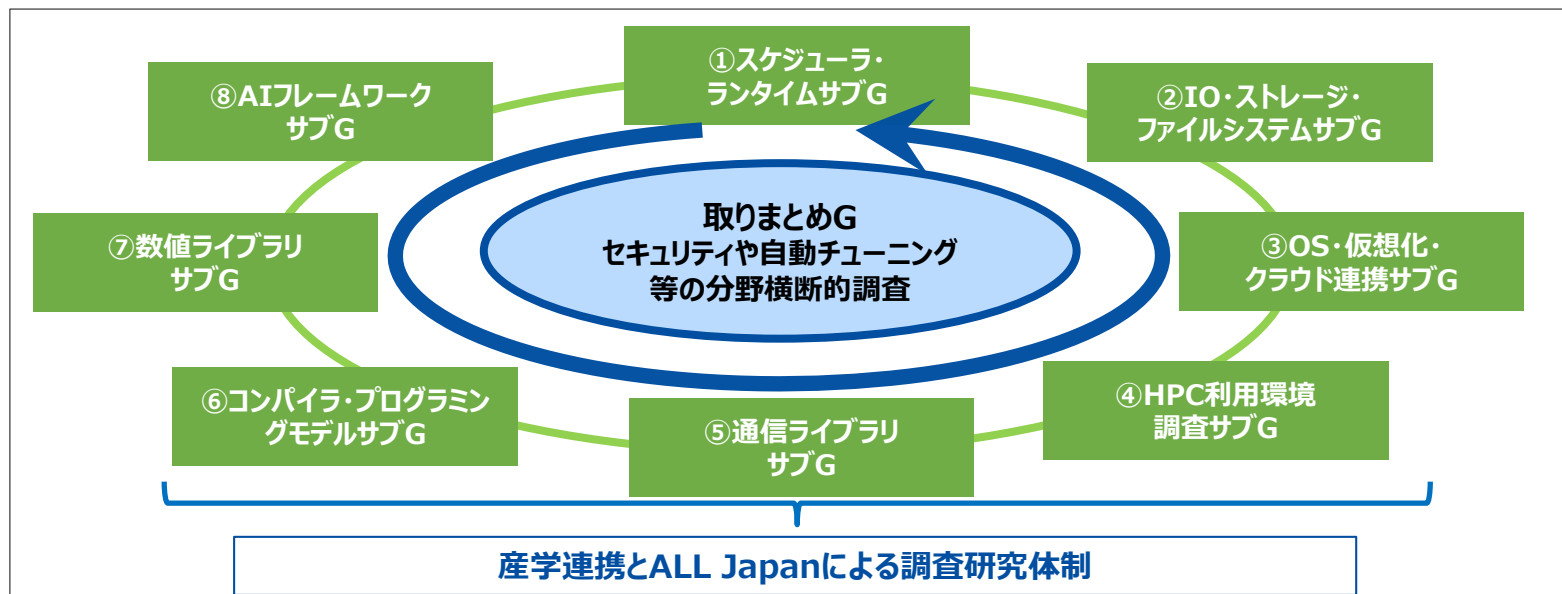
- 民間データセンターやクラウドでも応用可能な技術領域を対象

- 産業利用を目指しプラットフォーム化したデジタルツイン環境の提供、データ利活用の促進、機械学習技術と第一原理シミュレーションの融合、高度な大規模リアルタイムデータ処理、高セキュリティ化など

- 産学連携によるALL Japanの調査研究体制

- 代表機関である理化学研究所が強みを持つ⑥～⑧に関しては自ら調査研究

- ①～⑤に関しては、当該研究分野では各大学の研究者と分担・協力したALL Japan体制（①は東北大学、②は筑波大学およびDDN、③は国立情報学研究所、④は大阪大学、⑤は九州大学）



システムソフトに関する技術動向調査
と研究開発ロードマップの策定



産業展開を見据えた
システムソフトウェア新領域の調査



各アプリ分野サブG



旧戦略5分野アプリ群に加え社会科学やAI・BigDataを含めた幅広いワークロードを網羅的に調査研究

デジタルツインのための裾野の広い計算基盤の構築へ貢献

アプリサブG + ベンチ構築サブG + 性能モデリングサブG

カテゴリ	チーム名	アプリ名	アプリの内容
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)
計算科学	計算科学チーム	SPRINT	SPRINT (Scalable Parallel Runtime)

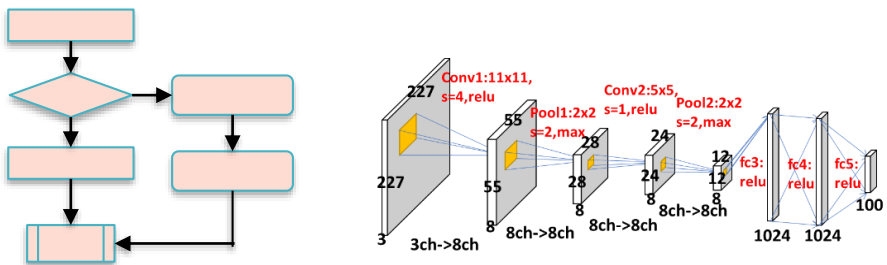
重要アプリカーネルを選出しベンチマーク化
各種アーキ上で簡便に評価可能なツール構築



ベンダー1
ベンダー2
ベンダー3
:

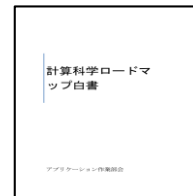
各種カーネルレポジトリ化し
共通ベンチマークセットを各ベンダーに提供+性能解析

科学技術計算／機械学習アルゴリズムサブG



次世代計算基盤のアーキ特徴を踏まえ最新・有用なアルゴリズム・解法（時空間タイリングや通信回避アルゴリズムなど）また機械学習手法とその第一原理計算との連携に関して調査研究

各アプリ分野サブG



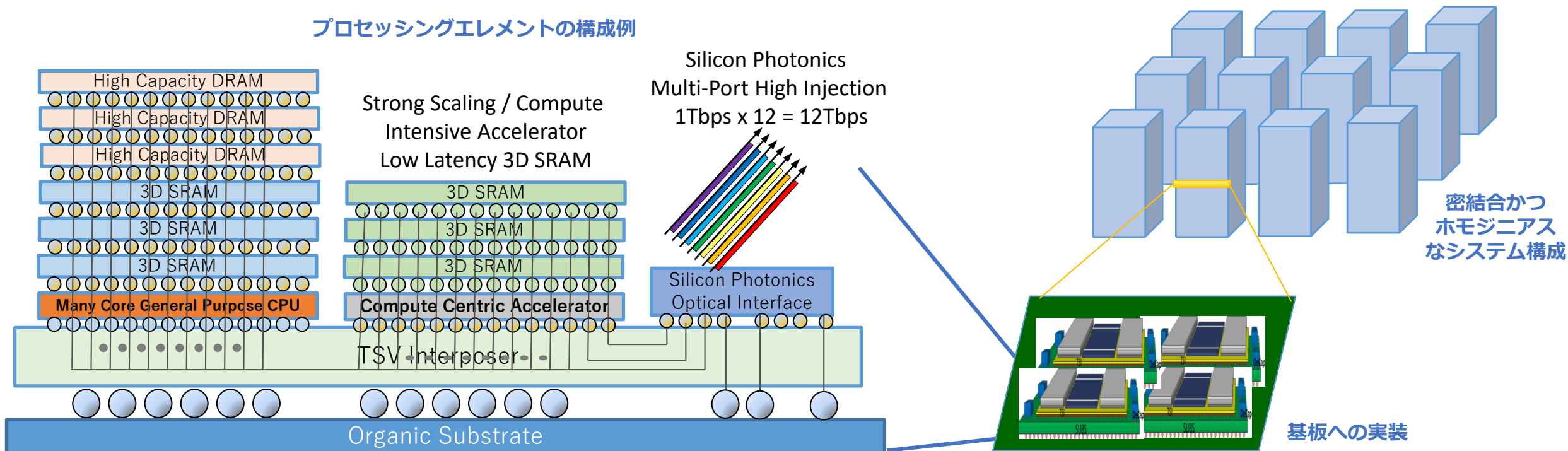
ベンチマーク評価解析に基づき次世代のアプリチャレンジを検討、サイエンスロードマップへ反映

- ✓msの分子シミュにより新規機能タンパク質の設計・細胞内タンパク質の機能解明へ
- ✓高温超伝導体の量子結核や超伝導ギャップの直接シミュレーションが可能に
- ✓将来気候予測の信頼性に対して質的向上が期待
- ✓自然界のクォーク質量より軽い深カイラル領域の本格探索
- ✓ペルソナモデルに基づく社会シミュレーション

サイエンスへの波及効果例

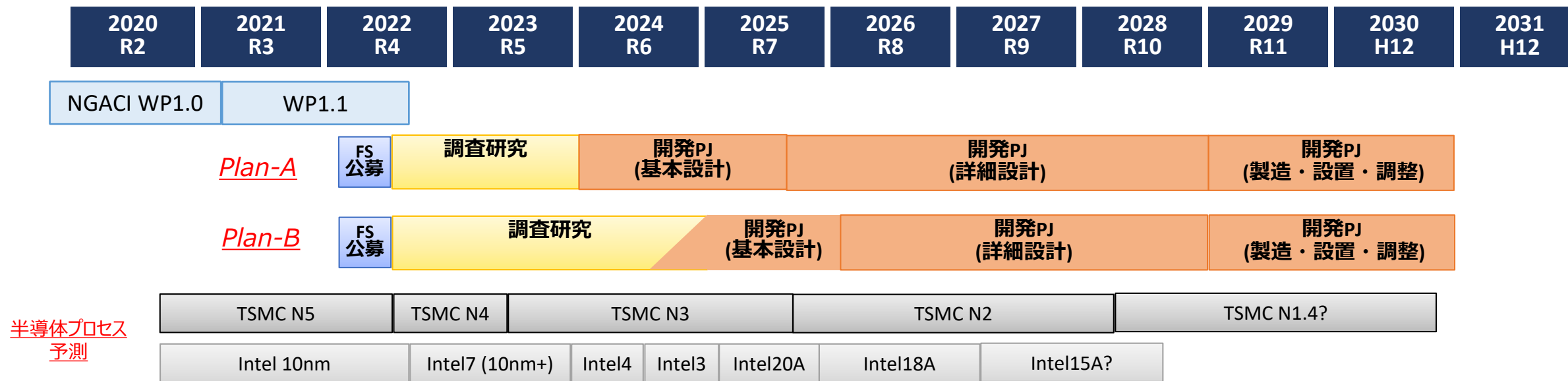
- データ移動効率化指向のアーキテクチャ・アルゴリズムの一例
- 三次元積層メモリ技術を駆使した相対メモリバンド幅の大幅な向上
- シリコンフォトニクスによるリモートメモリアクセスへの高バンド幅確保
- 強スケーリング実行での実行効率の確保：データフロー型の導入などによる低遅延実行
- 混合精度演算の積極的利用とアーキ・ソフトウェアからのサポート

プロセッシングエレメントの構成例

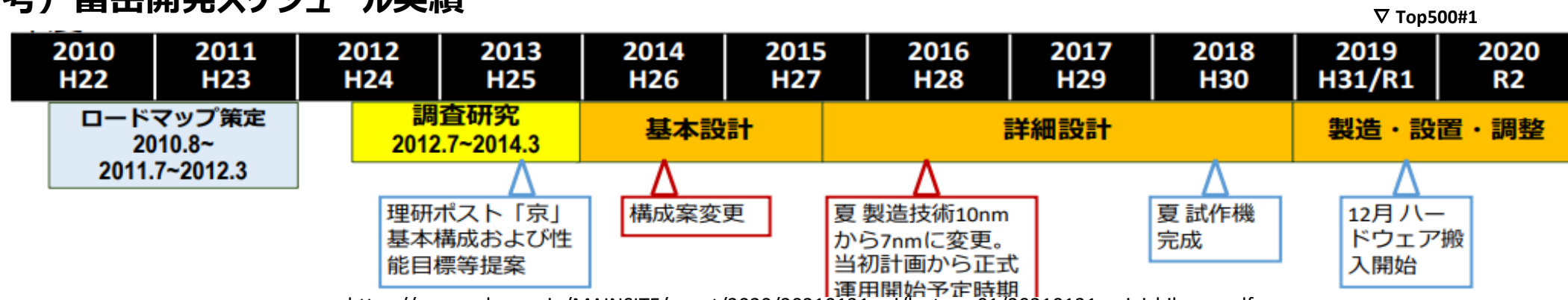


「3次元メモリ技術」＋「光技術」：日本の技術力を結集して破壊的イノベーションへ

富岳Next スケジュール予測



(参考) 富岳開発スケジュール実績



- 「アプリケーションファースト」な計算基盤の構築
 - Top500などの指標に捕らわれない設計目標の設定
 - ソフトウェア・エコシステムを当初から意識したハードウェア・ソフトウェアの開発
 - 利用者の裾野を広げつつ最適化も可能なアプリケーションフレームワークへの投資とそのためのシステムソフト開発
 - 計算システム内、またエッジ・計算システム間でのワークフローを効率的に実行できるシステム構築
 - ターゲットコードに特化し過ぎない「コデザイン2.0」手法の構築
 - 数世代にわたりアップデートされ続ける（≡商用展開で成功する）アーキテクチャの設計
- 破壊的イノベーションをもたらす計算基盤の開発
 - 最先端テクノロジーのアーリーアダプターとして世界をリードする高い実行性能を誇るシステムの実現
 - データを基にした「計算」が価値を生む今後の世界に向けた日本のプレゼンス強化
- 「フラグシップだからできない」ではなく「フラグシップだからこそできる」思い切った機能的・運用的な進化への挑戦
 - 計算基盤と実験系や新計算原理とのより密な連携など
- 日本に夢を与えられる次世代計算基盤開発プロジェクトへ