

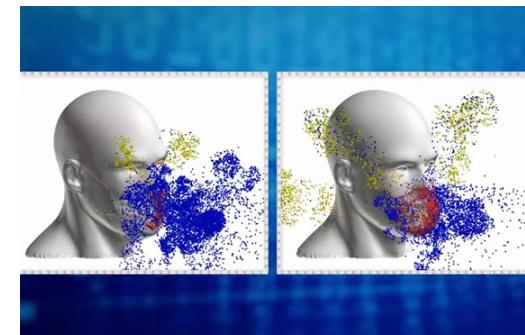
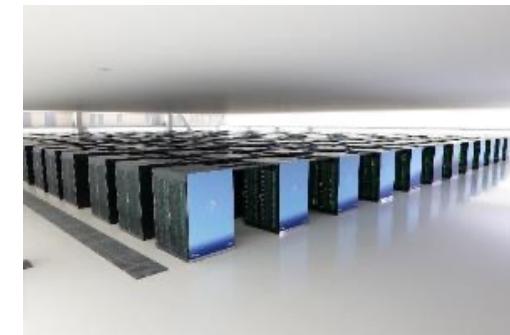
科学技術計算分科会 2022年度会合
「富岳」NEXTへの挑戦 ~現在から未来へ~

「次世代計算基盤に係る調査研究」
理化学研究所チーム システムソフトGの取組
January 20, 2023

理化学研究所 高性能ビッグデータ研究チーム
佐藤 賢斗

フラグシップマシン（富岳・ポスト富岳）開発のこれまで

- 富岳(ポスト京)が検討され始めたのは2010年
 - 2010年から2013年にかけて「戦略的高性能計算システム開発に関するワークショップ（SDHPC）」が開催され、本コミュニティの多くのボランティアの方々のご協力で「今後のHPCI技術開発に関する報告書」がまとめられました。
- 2014年に "FLAGSHIP2020" プロジェクトがスタート
- 2019年3月から富士通ITプロダクツ（石川県かほく市）にて製造を開始。2019年12月3日から理化学研究所 計算科学研究センター（神戸市）への搬入、設置・調整を始め、新型コロナウイルスの渦中、2020年5月13日に筐体の搬入を終了（大型10ントラック 72台分）
- 2019年半ごろ、SDHPCの後継のNGACI (Next-Generation Advanced Computing Infrastructure) 発足
- ISC20, SC20、「富岳」が2期連続で4部門のベンチマークで世界一位、SC20ではゴードン・ベル賞ファイナリストに2件選出。
- 2021年3月9日、幅広い学術・産業分野に計算機資源を提供するため、富岳が共用を開始した
 - 共有開始前の調整中に、COVID-19対策のためにいくつかの計算機資源が提供されていた
- 2021年4月 「次世代先端的計算基盤に関する白書（White Paper on Next-Generation Advanced Computing Infrastructure）」を公開
- ISC21において、「富岳」が3期連続で4部門のベンチマークで世界一位
- SC21において、「富岳」が4期連続で4部門のベンチマークで世界一位、SC21ではゴードン・ベル賞 COVID-19 研究特別賞受賞（COVID-19 の飛沫・エアロゾル拡散モデル）
- 2022年8月に “次世代計算基盤に係る調査研究事業” がスタート
- 2022年11月、5期連続で2部門（HPCG, Graph500）のベンチマークで世界一位。Top500およびHPL-AIでは世界二位。SC22ではゴードン・ベル賞受賞（プラズマミュレーション WarpX）
- 2028-30年ごろにポスト富岳の共用開始が期待。



目的

今後の高性能計算機の持続的な発展を考えるにあたり、AIやビッグデータ技術とのさらなる融合、Society5.0といった新しい応用分野への展開など、さらなる発展も期待される一方で、ムーアの法則の終焉など多くの技術的課題が待ち受けていることも事実である。本活動(NGACI)は、将来の高性能計算環境として、また共用計算機資源としてどのような技術的課題があり、どのような研究開発が必要なのか、コミュニティーとしてどのような活動をしていくべきなのなどに関して、オープンに意見交換をしつつそれをWhite Paperとしてまとめることで本分野の発展に寄与することを目的としています。

White Paper on Next-Generation Advanced Computing Infrastructure
(次世代先端的計算基盤に関する白書)

第1.0.0版

2020年11月

Screenshot

NGACI WP (v1.0) 執筆者

全体取りまとめ：近藤正明（慶應大学/理化学研究所）

アーキテクチャWG リーダ：三輪忍(電気通信大学), 佐野健太郎(理化学研究所) , 谷本輝夫(九州大学)	安島雄一郎(富士通), 井口寧(北陸先端大学), 井上弘士(九州大学), 江川隆輔(東京電機大学), 岡本有司(Spin Memory), 小野貴継(九州大学), 鯉渕道紘(国立情報学研究所), 児玉祐悦(理化学研究所), 小林諒平(筑波大学), 小松一彦(東北大学), 佐藤雅之(東北大学), 塩見準(京都大学), 田邊昇(東京大学), 中里直人(会津大学), 吉川隆英(富士通研究所), 福本尚人(富士通研究所), 星宗王(NEC), 三好健文(わさらぼ), 宮島敬明(理化学研究所)
システムソフトウェアWG リーダ： 佐藤賢斗(理化学研究所) , 佐藤幸紀(豊橋技術科学大学)	合田憲人(国立情報学研究所), 遠藤敏夫(東京工業大学), 小柴篤史(理化学研究所), 小松一彦(東北大学), 坂本龍一(東京大学), 高野了成(産業技術総合研究所), 滝沢寛之(東北大学), 辻美和子(理化学研究所), ゲローフィバリ(理化学研究所), 中島耕太(富士通研究所), 深井貴明(理化学研究所), 山本啓二(理化学研究所), 和田康孝(明星大学)
アプリケーション・ライブラリ・アルゴリズムWG 深沢圭一郎(京都大学), 今村俊幸(理化学研究所), 中島研吾(東京大学/理化学研究所)	岩下武史(北海道大学) , 小野謙二(九州大学), 笠置明彦(富士通研究所), 片桐孝洋(名古屋大学), 白幡晃一(富士通研究所), 住元真司(富士通研究所), 高橋大介(筑波大学) , 寺尾剛史(理化学研究所), 長坂侑亮(富士通研究所), 棕木大地(理化学研究所), 村上弘(東京都立大学)
システム運用WG 塙敏博(東京大学) , 野村哲弘(東京工業大学)	大島聰史(名古屋大学), 實本英之(理化学研究所), 庄司文由(理化学研究所), 滝澤真一朗(産業技術総合研究所), 竹房あつ子(国立情報学研究所), 藤原一毅(国立情報学研究所), 三浦信一(理化学研究所)

テーマ	サブテーマ
基盤ソフトウェア	OS、仮想化、コンテナ
大規模並列 / 高性能計算	汎用演算機：CPU、アクセラレータ
	次世代・専用ハードウェア：量子コンピュータ、ニューロモーフィック、データフロー型 etc.
	メモリ：メモリ階層、データ移動
	インタコンエクト：データ転送
	ストレージ：ファイルシステム、I/Oツール、データ管理
プログラミング環境	プログラミングモデル
	コンパイラ
	デバッグ、テスト、再現性
	データ学習&推論フレームワーク（TensorFlow, PyTorchなど）
	データ処理&解析フレームワーク（Hadoop, Sparkなど）
性能解析ツール	プロファイラ、トレーサー、性能解析方法論
	性能モデル、シミュレーション
利用高度化ツール	省電力化システム：DVFSなどアプリの消費電力最適化
	耐障害性システム：高信頼通信ライブラリ、冗長符号化、チェックポイントリストアなど
	ディスクアグリゲーション技術
資源管理	資源モニタリング、ジョブスケジューラ
	リアルタイム処理
外部連携	分散ストレージ、広域ネットワーク、クラウド連携

第1章 はじめに

第2章 スーパーコンピュータの技術動向

- 2.1 ハードウェア技術の動向
- 2.2 システムアーキテクチャ(システムハードウェア)の技術動向
- 2.3 システムソフトウェアの技術動向**
- 2.4 数値計算ライブラリ/ミドルウェア/アルゴリズムの技術動向
- 2.5 運用に関する技術動向

第3章 アプリケーションの要求性能分析

- 3.1 次世代システムの予測性能とアプリケーションの要求性能
- 3.2 要求性能に対するアプリケーション分析
- 3.3 アプリケーションの要求性能との比較

第4章 次世代(2028年頃)システムの候補

- 4.1 汎用システム型
- 4.2 専用システム混載型および新たなる可能性

第5章 次世代型運用への要求

- 5.1 新しい利用形態とシナリオ
- 5.2 設備・管理
- 5.3 ユーザ利用・課金モデル

第6章 技術課題と研究開発ロードマップ

- 6.1 デバイス・アーキテクチャ
- 6.2 システムソフトウェア**
- 6.3 数値計算ライブラリ・アルゴリズム

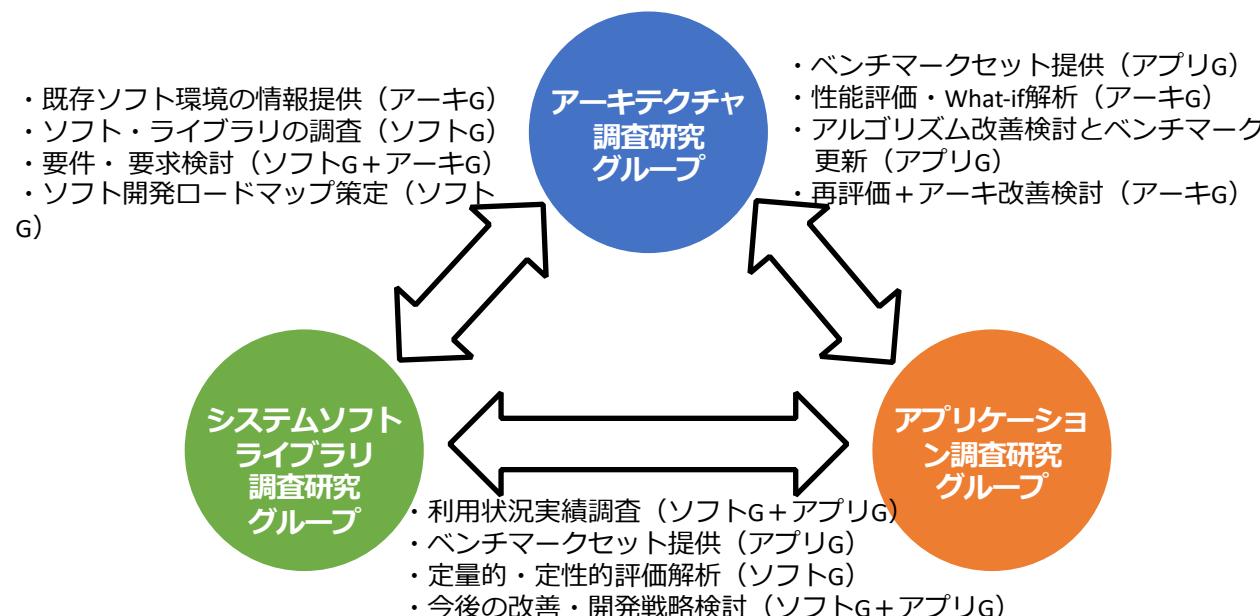
第7章 おわりに

46ページ / 163ページ

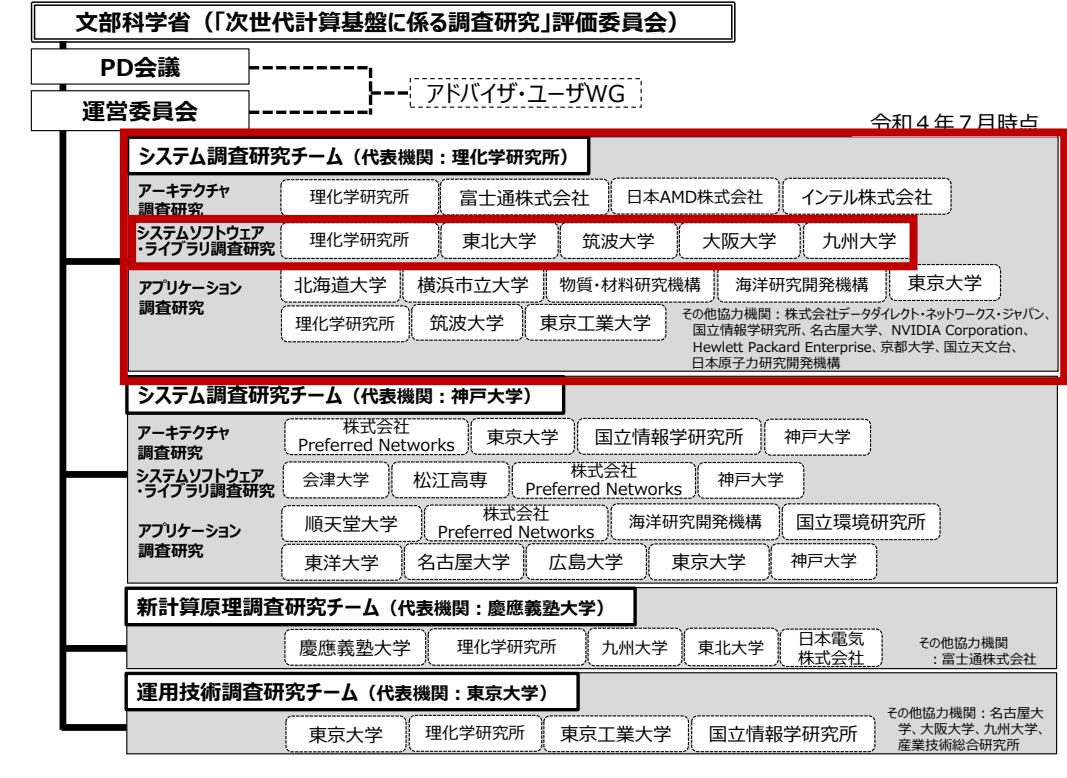
本事業全体の趣旨

本調査研究は、ポスト「富岳」時代の次世代計算基盤の具体的な性能・機能等について、サイエンス・産業・社会のニーズを明確化し、それを実現可能なシステム等の選択肢を提案するものである。その際、我が国として独自に開発・維持するべき技術を特定しつつ、システム（アーキテクチャ、システムソフトウェア・ライブラリ、アプリケーション）、新計算原理、運用技術を対象とした調査研究及び要素技術の研究開発等を実施。（“評価項目”も参照）

システム調査研究チーム（代表機関：理化学研究所）



「次世代計算基盤に係る調査研究」実施体制



評価項目：演算性能、電力性能比、I/O性能、**開発コスト**、**生産性**（アプリ開発のしやすさ）、商用展開・技術展開、カーボンニュートラル

次世代計算基盤に係るシステム検討ワーキンググループ

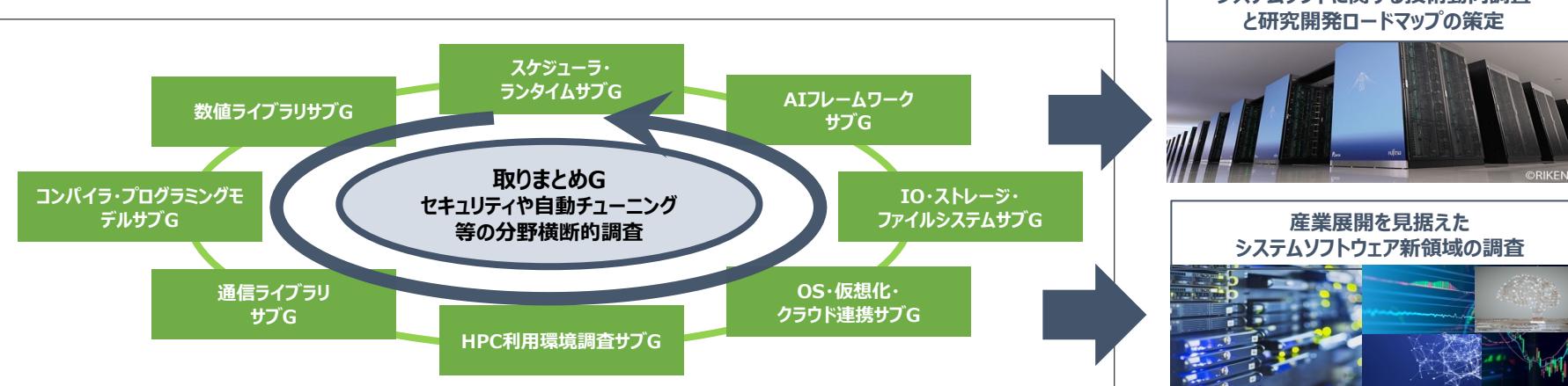
システム調査研究/システムソフト・ライブラリ調査研究Gの概要（システムソフトG、SysSoftG）

取組概要

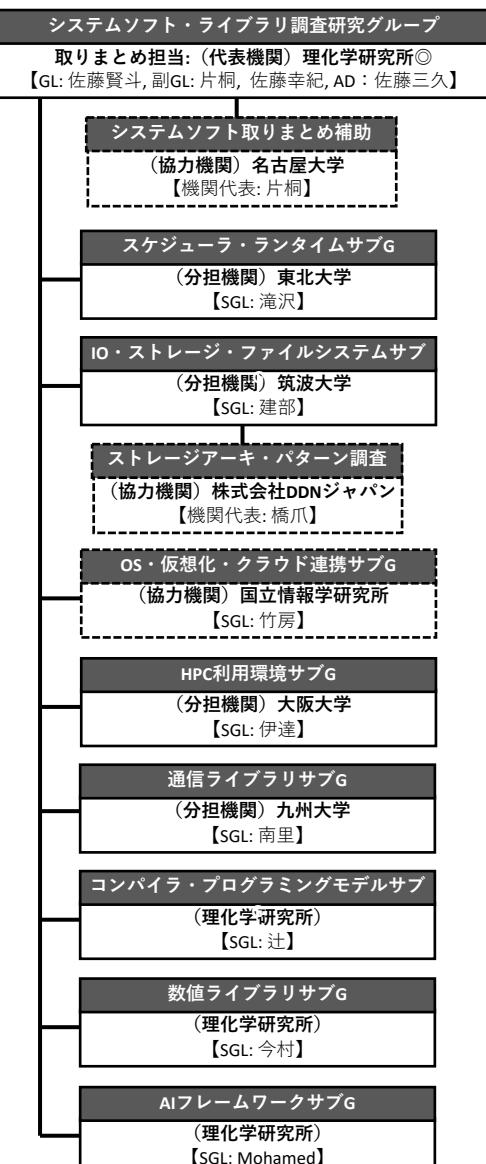
システムソフトGに関しては、これまでのフラグシップや第二階層システムのソフトウェアの開発工数や利用状況、そして既に構築されているエコシステムとの親和性も踏まえ、ソフトウェア資産として何をベースに、国内でどこまでを開発すべきか、もしくは国際協調が必要なソフトウェアを優先度をつけて明らかしつつ、今後のロードマップを策定することが必要である。このため、次世代計算基盤を産業界を含む幅広い応用分野での活用を促すためにも、従来のシステムソフトウェアだけでなくデータ利活用の促進、機械学習技術と第一原理シミュレーション、さらには大規模リアルタイムデータ処理の高度な融合、従来の共用HPCシステムとは次元の異なる高セキュリティの担保、などを主要検討項目として見据えた今後の日本の次世代計算基盤として行うべきソフトウェア開発について調査研究を行う。

調査内容

次世代計算基盤に係る調査研究は、候補アーキテクチャを取り巻くソフトウェアエコシステムの状況やその後の商業展開を見据えた調査が不可欠であり、網羅的なシステムソフトウェアの調査研究が必要となる。このためシステムソフト・ライブラリ調査研究グループでは、①スケジューラ・ランタイム、②IO・ストレージ・ファイルシステム、③OS・仮想化・クラウド連携、④HPC利用環境調査、⑤通信ライブラリ、⑥コンパイラ・プログラミング、⑦数値ライブラリ、⑧AIフレームワークのシステムソフトウェアの技術動向の調査を行うとともに、利用実態やAPIと連携し（可能であれば）優先順位をつけ、それに基づき、国内で開発・移植・チューニングすべきものと国際協調によりソフトウェア開発を行うべきかの検討（を可能にするための調査）を行う。評価項目には、演算性能、電力、性能比、I/O性能、開発コスト、生産性（アプリ開発のしやすさ）、商用展開・技術展開、カーボンニュートラルが挙げられる。



調査研究体制



調査研究項目

- 【調査研究1】システムソフトウェアの動向調査
 - (1-1) 既存および将来システムソフトウェアの動向調査
 - 移植性・生産性、性能について調査
 - (1-2) 大規模システムにおけるシステムソフトウェアの調査（担当：HPC利用調査サブG）
 - 導入調査、利用頻度、その他優先順位の決定のための情報収集
 - 海外のスパコンセンターの動向の調査
 - センターや採用しているアーキテクチャーによる違い 等
- 【調査研究2】国内で開発すべきか否かを検討するための調査
 - (1) 国内では維持しない技術、独自開発しないソフトウェア
 - (a) OSSコミュニティによる開発（例：gcc, LLVM）
 - (b) 海外企業による独自開発（例：SYCL/DPC++）
 - (c) OSSの移植（工数の低い移植：As-is or 軽微な移植）（例：VeloC）
 - (d) OSSの移植（工数の高い移植：チューニング）（例：OneDNN）
 - (e) 国際共同による新規開発
 - (2) 国内で維持する技術、独自開発するソフトウェア
 - (a) 国内独自開発予定（例：富士通コンパイラ、LLIOソフトウェア）
 - (b) 開発&導入未定
- 【調査研究3】類似ソフトウェアに関しては差異の調査
 - 例1) OpenMPI vs. MVAPICH はAPIはほぼ同等
 - 例2) OneMKL vs SSL-II, Numpy v.s. cuNumeric
 - Intel x86 (OneMKL) とFujitsu Arm (SSL-II) と同等の開発環境が提供できるか、できなければ一方が他方に対して何が足りないか

最終報告書のイメージ（プロセッサ版）

評価指標：
開発コスト、生産性(アプリ開発のしやすさ)
商用展開・技術展開

HPC利用状況調査GやアプリGからの要望に基づき、どのようなシステムソフトウェア・ライブラリを開発/移植・チューニングすべきかを決定

候補アーキテクチャ

		アーキ候補1 (x86)	アーキ候補2 Arm9	アーキ候補3 Arm9+X(Chiplet)	アーキ候補4 x86+GPU	アーキ候補5 ...
コンパイラ		gcc	(1-a)			
新規コンパイラ		LLVM/clang				
低位 数値ライブラリ		icc	(1-b)	N/A		
新規コンパイラ		OneMKL	(1-b)			
低位 数値ライブラリ		OpenBLAS				
数値計算ライブラリ		Eigen		(2-a)		
数値計算ライブラリ		SSL-III				
高位 数値計算ライブラリ		...				
高位 数値計算ライブラリ		Numpy	(1-a)	(1-d)		
高位 数値計算ライブラリ		Scipy				
高位 数値計算ライブラリ		cuNumeric				
高位 数値計算ライブラリ		...				

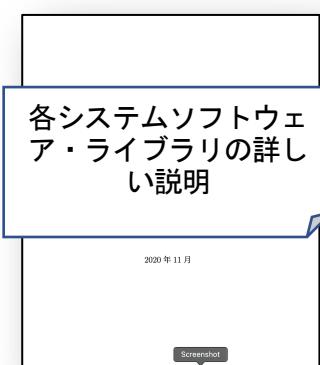
国内で開発すべきか否かを検討するための調査

複数の類似システムソフトがある場合、差異の調査

(主要な) システムソフトに関しては、異なるアーキテクチャでの性能評価

- (1-a) OSSコミュニティによる開発
- (1-b) 海外企業による独自開発
- (1-c) OSS (工数の低い移植 : As-is or 軽微な移植)
- (1-d) OSS (工数の高い移植 : チューニング)
- (1-e) 國際共同開発
- (2-a) 国内独自開発 (重点開発 or 富岳ソフトの移植)
- (2-b) 開発未定

※インターフェクトやストレージに関する調査も同様



今後の調査研究計画

