



HPCI 整備計画調査研究事業 運用システム（計算機）整備計画調査研究

令和 7 年度 成果報告書

代表機関: 理化学研究所

令和 8 年 3 月 31 日

概要

本成果報告書は、文部科学省の HPCI 整備計画調査研究事業による委託業務として、国立研究開発法人理化学研究所が代表機関として実施した令和 7 年度「運用システム（計算機）整備計画調査研究（以下、FS3.0 と称する）」の成果を取りまとめたものです。FS3.0 は、スーパーコンピュータ「富岳」の後継となる「富岳 NEXT」およびその先の将来を見据え、フラッグシップシステム単体にとどまらず、全国の多様な計算機資源から構成される革新的ハイパフォーマンス・コンピューティング・インフラ（HPCI）全体として、科学技術計算の成果を最大化するための運用システム（計算機）整備計画の策定を目的としている。近年、生成 AI に代表される人工知能技術の急速な進展により、計算科学研究は数値シミュレーション中心の利用形態から、AI for Science を含むデータ駆動型・自動化型の研究スタイルへと移行しつつある。これに伴い、次世代 HPCI には、計算性能の向上に加え、計算機アーキテクチャの多様化、柔軟なシステムソフトウェア基盤、ならびに次世代アプリケーションを支える実行環境が求められている。FS3.0 では、計算機アーキテクチャ、システムソフトウェア、次世代アプリケーション開発の三つの技術領域を対象とし、協調設計（Co-Design）の観点から体系的な調査を行う。また、「富岳 NEXT」開発プロジェクトや HAIRDESC 等との連携を通じて、実現可能性を重視した将来の HPCI 整備方針の整理を進めている。最新の情報については、プロジェクトの Web ページ（<https://www.r-ccs.riken.jp/research/feasibility-study/fs3/>）を参照されたい。

Abstract

This report summarizes the outcomes of the FY2025 “Next-Generation Computing Infrastructure” (commonly referred to as FS3.0), which was conducted by RIKEN as the lead institution under a commissioned project of the Ministry of Education, Culture, Sports, Science and Technology for the “Feasibility Study on the future HPCI” initiative. FS3.0 aims to formulate a development plan for operational systems (computing infrastructure) that maximizes scientific computing outcomes across the entire High Performance Computing Infrastructure (HPCI). Looking ahead to the successor of the supercomputer Fugaku—namely “FugakuNEXT”—and beyond, the study considers not only a standalone flagship system but also an integrated infrastructure composed of diverse computing resources distributed nationwide. In recent years, rapid advances in artificial intelligence technologies, particularly generative AI, have driven a shift in computational science from a simulation-centric paradigm toward data-driven and automated research approaches, including AI for Science. In response, next-generation HPCI systems are required not only to deliver enhanced computational performance, but also to support diversified computing architectures, flexible system software platforms, and execution environments capable of sustaining next-generation applications. FS3.0 conducts a systematic investigation across three key technical domains—computer architecture, system software, and next-generation application development—from a co-design perspective. Furthermore, through collaboration with the “FugakuNEXT” development project and related initiatives such as HAIRDESC, the study advances the formulation of future HPCI development policies with a strong emphasis on practical feasibility. For the latest information, please refer to the project website (<https://www.r-ccs.riken.jp/en/research/feasibility-study/fs3/index.html>).

目次

第 1 章	調査研究の概要	2
1.1	調査研究の背景と目的	2
1.2	調査研究の体制	3
1.3	最終年度に向けた達成目標	5
第 2 章	アーキテクチャ調査研究グループ	6
2.1	調査研究の概要および方針	6
2.2	アクセラレータ、メモリ技術、半導体技術に関する調査	6
2.3	Intel における将来システムの動向調査	17
2.4	NVIDIA における将来システムの動向調査	23
2.5	HPE における将来システムの動向調査	26
2.6	富士通における将来システムの動向調査	30
2.7	NEC における将来システムの動向調査	33
第 3 章	システムソフトウェア調査研究グループ	37
3.1	調査研究概要および方針	37
3.2	HPC プログラミング・通信環境サブグループ	38
3.3	先端演算加速装置プログラミング環境サブグループ	47
3.4	数値計算ライブラリサブグループ	56
3.5	ストレージ・ファイルシステムサブグループ	61
第 4 章	アプリケーション調査研究グループ	68
4.1	調査研究の概要および方針	68
4.2	Japan Scientist AI Jam Session を踏まえた HPCI 整備計画に関する調査	70
4.3	次世代アプリケーションニーズ調査研究サブグループ	75
4.4	次世代アプリケーション開発手法調査研究サブグループ	86

第 1 章

調査研究の概要

1.1 調査研究の背景と目的

近年の計算科学分野においては、AI 技術の急速な進展、とりわけ生成 AI の台頭により、科学技術研究における計算資源の利用スタイルが大きく変化しつつある。従来は数値シミュレーションを中心とした解析が主流であったが、今後は AI による自動化やデータ駆動型の解析と統合された「AI for Science」への移行が加速すると見込まれている。

このような変化を踏まえ、HPCI においては、計算機アーキテクチャや実行環境における多様性と柔軟性、さらにはアプリケーション側からの要件を考慮したシステム設計が求められている。特に、次世代のフラッグシップシステムである「富岳 NEXT」およびその後継となる HPCI 全体のシステム群では、こうした要求に応えるための総合的な検討が不可欠である。

FS3.0 は、このような背景のもと、次世代 HPCI システムの整備計画を策定するために実施されるものであり、科学技術計算における成果の最大化を目指す「運用システム（計算機）整備計画調査研究」として位置付けられている。本調査研究では、単一の計算機システムに閉じることなく、フラッグシップシステムと全国の多様な計算資源が連携した運用体制を想定し、その実現に向けて必要な要素技術の調査および体系的な計画立案を行う。

調査の主な対象は以下の 3 点である。

- **計算機アーキテクチャ**：次世代の計算基盤が備えるべき構成要素を検討するため、最新の半導体技術、パッケージング技術、チップレットや積層構造、光通信技術などを対象に、調査と技術動向の分析を行う。
- **システムソフトウェア**：次世代の HPCI を構成するシステム群が一体的に運用されるために必要なシステムソフトウェア基盤について、プログラミング環境、通信ライブラリ、数値計算ライブラリ、ファイルシステム等の技術的要件を整理し動向調査を行う。
- **次世代アプリケーション**：将来的に科学技術研究を牽引する次世代 HPC・AI アプリケーションについて、開発支援手法、自動化、ベンチマーク整備、教育展開まで含めたニーズ分析を行い、システムに求められる実行基盤要件を明らかにする。

また、これらを独立に検討するのではなく、三者の相互作用に注目した協調設計（Co-Design）の観点を重視し、実用的かつ持続可能な HPCI 整備方針の確立を目指す。

FS3.0 は、理化学研究所を代表機関とし、大学・研究機関・産業界の幅広い参画による体制で実施されてお

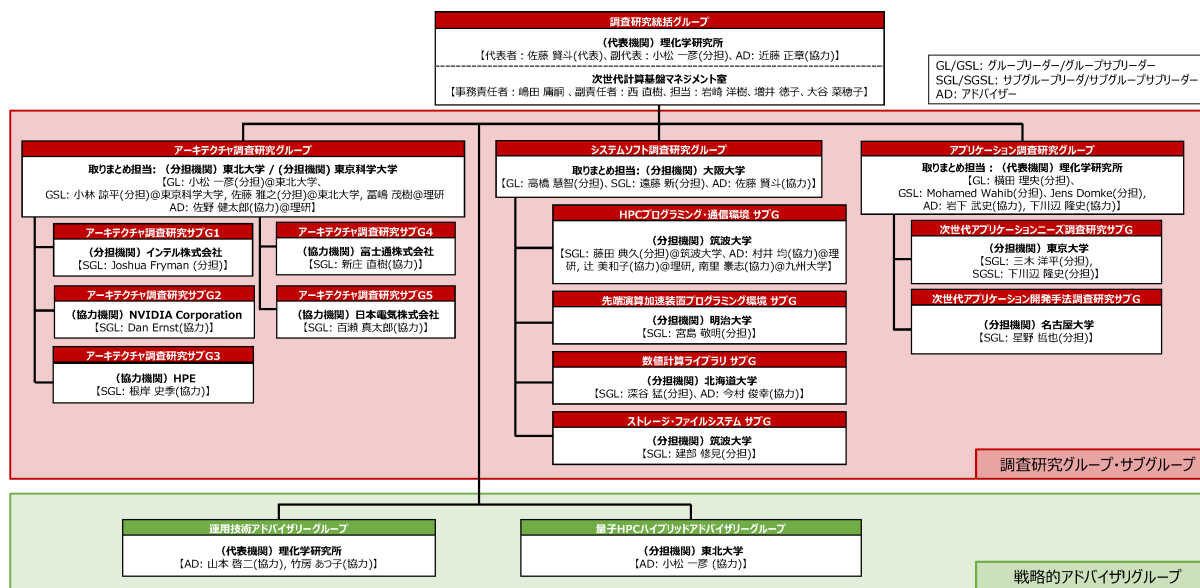


図 1.2.1 FS3.0 における実施体制図

り、主要ベンダーや次世代 HPC・AI 研究開発支援センター（HAIRDESC） [1] との連携や「富岳 NEXT」開発プロジェクトとの整合性の確保を図りながら、HPCI 全体の将来像の構築に貢献することを目的としている。

1.2 調査研究の体制

FS3.0 は、理化学研究所を代表機関とし、大学・企業を含む多様な機関が参画する構成により実施されている。調査研究全体を統括する「調査研究統括グループ」のもとに、以下の 3 つの調査研究グループが設けられている。

- アーキテクチャ調査研究グループ
- システムソフトウェア調査研究グループ
- アプリケーション調査研究グループ

さらに、各グループの下には複数のサブグループが組織されており、詳細な技術領域に対応する形で調査研究が分担されている。各グループ・サブグループには、グループリーダー（GL）、サブグループリーダー（SGL）、およびアドバイザー（AD）が配置されており、技術的な助言と調整が行われる体制となっている。

1.2.1 代表機関および参加機関の体制

代表機関である理化学研究所は、FS3.0 の全体統括と実行責任を担うとともに、「富岳 NEXT」開発とも連携しながら調査の整合性を確保する。調査研究全体の運営を支援するために、代表機関内には PMO（Project Management Office）が設けられており、報告書作成、研究進行管理、他機関調整などの事務的・技術的支援を行っている。これにより、実施体制の透明性と実行力が担保され、円滑な調査研究の推進が可能となってい

- HPCI-CFSP: 継続的な調査研究プラットフォームの構築 (HPCI Continuous FS Platform)
 - → (1) (2) 等のFS3.0の調査研究の成果・ノウハウを継続的に継承させるための仕組みづくり
 - (1): NDA情報を含む技術情報を含め HPCI-RB をどのように管理・継承するか検討
 - (2): 長期的な HPCI-CB の運用方針の策定

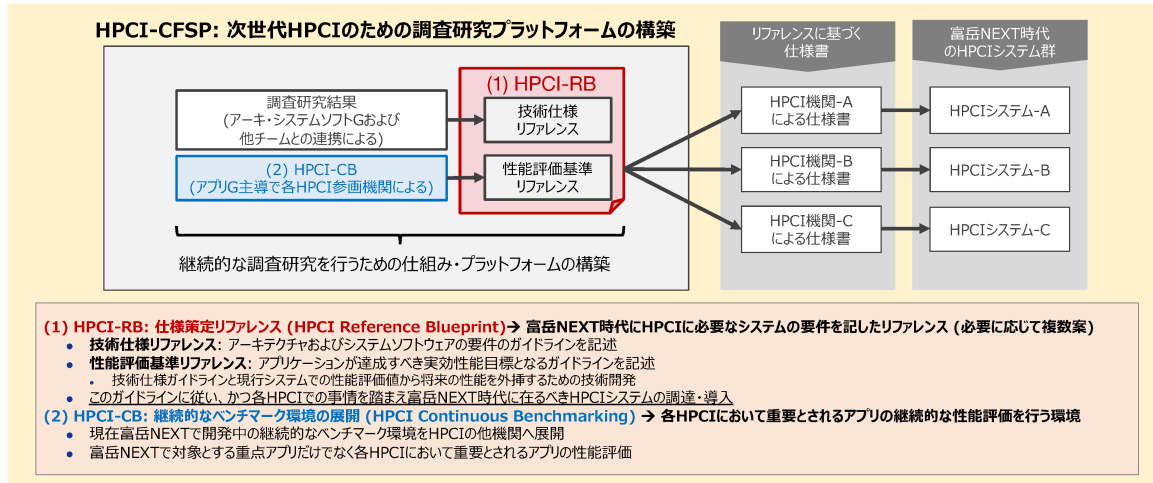


図 1.3.1 次世代 HPCI を支える持続的な調査研究プラットフォームの構築

る。分担機関としては、東北大学、東京科学大学、大阪大学、筑波大学、北海道大学、明治大学、東京大学、名古屋大学などの大学に加え、インテル、NVIDIA、富士通、日本電気 (NEC)、HPE などの産業界からも協力を得ており、広範な視点からの調査研究が可能となっている。また、FS2.0を担当したメンバーの多くがADとして継続的に参画し、技術的知見と実施体制の継承を支援している。調査グループ間では定期的なミーティングを通じて情報交換が行われ、調査結果の相互フィードバックや、全体構想の整合性維持を図る。

1.2.2 連携体制

FS3.0には、富岳NEXTの開発の中心的なメンバーが参画している。具体的には、アーキテクチャ設計、システムソフト設計、運用方針策定に関わるリーダーらがFS3.0に直接参画しており、富岳NEXTプロジェクトと本調査の方向性が一貫したものとなるよう配慮する。この連携により、調査研究の成果が富岳NEXTおよびその後継となるHPCI整備計画に直接的に反映される体制が整っており、実現可能性を踏まえた計画立案が可能となっている。また、HPCI整備計画における他の調査研究（運用体制調査研究、セキュリティ・運用技術実証研究、量子HPCハイブリッド運用環境調査研究など）との連携を促進するために戦略的アドバイザーグループを設けることで、他チームとの整合性を保ちつつ、必要な情報共有と意見交換を可能としている。さらに、HAIRDESCとの連携を通じて、ユーザ側のニーズや実行環境への要求を吸い上げ、それをシステム要件に反映する仕組みも確立する。

1.3 最終年度に向けた達成目標

FS3.0は、「富岳NEXT」時代におけるHPCI全体の計算基盤像を明確化し、それを支える運用システムの整備方針を策定することを目的とする。ここで重要なのは、特定のHPCI運用組織の事情を無視した一方的な技術仕様の提示ではなく、富岳NEXTの実情を踏まえつつ、各組織が採用可能な複数の方向性を提示することである。提案の採否および具体的な導入判断は最終的に各HPCI運用組織に委ねられるが、その際に参照可能な技術的指針を提供することが本調査研究の本質的な目標である。

また、ハードウェア構成に関しては、富岳NEXTのアーキテクチャ（例：Monaka-X [2]、次世代NVIDIA GPU）の普及促進しつつ、一方で特定ベンダーや製品に依存しない選択肢の提示を重視する。国内開発CPUについては、アプリケーション側の要件整理を富士通と連携して進め、後継CPU開発への貢献を目指す。

これらを踏まえた上で、FS3.0は「HPCI-CFSP: 次世代HPCIを支える調査研究プラットフォームの構築」を全体目標とする。HPCI-CFSPは、調査研究で得られた成果・ノウハウを継続的に活用するための基盤を構築する。非公開情報を含む技術資料やHPCI-RBで策定したリファレンスの管理・継承方法を整備し、またHPCI-CBの長期運用方針を策定することで、調査成果を将来世代に引き継ぐ仕組みとする。

1. **HPCI-RB（仕様策定リファレンス）** 富岳NEXT時代に必要とされるシステム要件を複数案として整理し、アーキテクチャやシステムソフトウェアに関する技術仕様リファレンス、アプリケーションの実効性能を示す評価基準リファレンスを策定する。現行システムでの性能値と組み合わせ、将来の外挿評価を可能とする技術も併せて開発することで、各HPCI組織に応じた調達・導入指針を提供する。
2. **HPCI-CB（継続的ベンチマーク環境）** 富岳NEXTで開発中の継続的ベンチマーク基盤を他のHPCI機関へ展開することで、重点アプリケーションに限定されず、各組織にとって重要なアプリケーションの性能評価を可能とする。これにより、ハードウェアやソフトウェアの実効性能を透明かつ信頼性の高い指標で継続的に測定・評価できる環境を構築し、調達判断を支援する。

HPCI-CFSPによりHPCIの調達・運用プロセスの効率化と全体最適化を促進し、富岳NEXT時代にふさわしい次世代HPCIの実現可能な全体像を提示することを目指している。本調査研究の成果は、我が国における学術研究および産業応用の基盤を支えるHPCIの持続的発展に貢献し、国際的な計算科学基盤競争における優位性の確保にも寄与することが期待される。

参考文献

- [1] hairdesc.jp. *HAIRDESC*. <https://hairdesc.jp>. Accessed: 2026-02-13. 2026.
- [2] Fujitsu Limited. *Fujitsu Technical Document (2025)*. <https://global.fujitsu/-/media/Project/Fujitsu/Fujitsu-HQ/ir/documents/tech/2025/12/02-03a-ja.pdf>. Accessed: 2026-02-13. 2025.

第2章

アーキテクチャ調査研究グループ

2.1 調査研究の概要および方針

2025年度は、アーキテクチャ調査研究グループにおける体制の整備を行い、半導体製造技術、アクセラレータ技術、メモリ技術のシステムアーキテクチャの前提となる技術動向を調査する体制を整備した。半導体製造技術についてはデバイス動向・技術に専門性を持つ佐藤 雅之（東北大学）、アクセラレータ技術についてはFPGA・GPU 応用に高い専門性を持つ小林 諒平（東京科学大学）、メモリ技術については国内外ベンダーとの関係を有し、メモリ技術に関する専門性を持つ富嶋 茂樹（理化学研究所）をサブグループリーダーとしてそれぞれ擁立し、実施体制を整備した。

また、サブグループであるインテル株式会社、NVIDIA Corporation、HPE、富士通株式会社、日本電気株式会社からの製品ロードマップ、要素技術・技術的方向についての情報を収集できる体制を整えた。

2.2 アクセラレータ、メモリ技術、半導体技術に関する調査

2.2.1 アクセラレータ技術に関する動向

今年度は、次々世代スーパーコンピュータにおけるアクセラレータ活用方針を検討するため、現行市場の調査と技術トレンドの整理を実施した。TOP500 2025年11月版では、El Capitan (AMD Instinct MI300A)、Frontier (AMD Instinct MI250X)、Aurora (Intel Data Center GPU Max)、JUPITER Booster (NVIDIA GH200) が上位を占めており、最先端システムの設計がすでにアクセラレータ中心に移行していることが確認できる。また、HPL-MxP においても El Capitan、Aurora、Frontier、JUPITER Booster が上位に並んでおり、mixed precision を活用した AI-HPC 融合が実機設計の前提になりつつある。加えて、Green500 2025年11月版では GH200 系システムが上位を占めており、性能のみならず性能/電力がシステム評価の主要軸であることも明瞭である。

本調査の第一の成果は、アクセラレータ市場を単なる製品リストとしてではなく、役割の異なるレイヤとして整理し直した点にある。具体的には、(i) 主計算を担う GPU/APU 層、(ii) pod / system 単位の拡張設計を牽引するクラウド専用 ASIC 層、(iii) 主流 GPU とは異なる設計原理を持つ独自アーキテクチャ層、(iv) 通信・I/O・ストレージをホストから切り離すインフラ・オフロード層、(v) 2030年代を見据えた先行計算レイヤ、の五層として再整理した。第二の成果は、この市場整理を、FS3.0 やその先のシステム設計論に直接つながる「主系統」「探索系統」「インフラ系統」という三本立ての研究開発方針へ変換した点である。

2.2.1.1 アクセラレータが中核となる背景

アクセラレータが将来機の中核になる理由は、大きく二つに整理できる。第一は、性能/電力がシステム設計の最上位制約になったことである。Green500 上位システムが示すように、今後の大規模計算基盤では、ピーク性能だけでなく、電力効率、冷却可能性、給電設計まで含めた全体最適が必要になる。第二は、HPC と AI の収束である。現在の最先端システムは、FP64 の数値計算専用機と低精度 AI 専用機に分化するのではなく、混合精度を活用しながら AI とシミュレーションを同一基盤上で運用する方向へ進んでいる。したがって、次々世代機の論点は、アクセラレータを導入するか否かではなく、どの種類のアクセラレータを、どの階層で、どのように組み合わせるかへ移っている。

2.2.1.2 アクセラレータ市場の五層整理

■**第1層：GPU/APU層** 第1層は、HPC と大規模 AI の基盤となる主流アクセラレータ群である。ここには、NVIDIA H100/H200/GH200、AMD Instinct MI250X/MI300A/MI355X、Intel Data Center GPU Max などが含まれる。この層の特徴は、すでにエクサ級の実機で採用実績があること、およびソフトウェア・エコシステムと運用知見が比較的蓄積していることである。したがって、将来機における主システムの基準線は、当面この層に置くのが妥当である。

■**第2層：クラウド専用 ASIC層** 第2層は、クラウド事業者が自社インフラ向けに設計した専用 ASIC である。Google TPU v6e / TPU7x、AWS Trainium2、Microsoft Maia 200 が代表例であり、いずれもチップ単体ではなく、Pod、UltraServer、あるいはクラスタ単位での拡張性を前面に出している。これらは伝統的な FP64 HPC の直接代替というより、大規模学習・推論、代理モデル、データ同化、生成 AI ワークロードの側から将来システム設計に影響を与えていると位置づけるべきである。

■**第3層：独自アーキテクチャ層** 第3層は、主流 GPU とは異なる設計原理に基づくアクセラレータ群である。国内では PFN の MN-Core 系が代表例であり、MN-Core 2 は AI training and high-performance computing processor として位置づけられ、MN-Core L1000 は generative AI inference に最適化したプロセッサとして開発が進められている。また、Tenstorrent は Blackhole と Galaxy を展開しており、RISC-V ベースの制御系と Ethernet-based mesh、さらに fully open source software stack を組み合わせることで、従来の GPU クラスタとは異なる設計方向を示している。加えて、Cerebras WSE-3、Groq LPU、SambaNova SN40L のように、wafer-scale、SRAM 主記憶 + static scheduling、dataflow + 三層メモリなど、特定の学習・推論形態に深く最適化したシステムも存在する。この層は、主流アーキテクチャの全面置換候補としてではなく、将来差別化の源泉として継続評価すべき対象である。

■**第4層：インフラ・オフロード層** 第4層は、通信、I/O、ストレージ、セキュリティの処理を主計算資源から切り離すオフロード層である。NVIDIA BlueField DPU は networking, storage, cybersecurity の line-rate offload を掲げ、Intel IPU は infrastructure task offload, virtual storage, infrastructure isolation を訴求している。MangoBoost も MB-DPU 向け SDK において、RNIC, NVMe-over-Fabrics Initiator/Target, TOE を supported product とし、programmable infrastructure offload and management を明示している。AI と HPC が同一基盤上で混在する将来機では、データ移動と I/O がボトルネック化しやすく、この層を主計算系と同格に扱う必要がある。

■**第5層：先行計算レイヤ** 第5層は、2030年代を見据えた先行計算レイヤであり、現時点では主システムを全面置換するものではなく、テストベッドとして段階的に統合すべき対象である。本稿では、この層を、光・アナログ計算、ニューロモーフィック、量子-HPC連携の三つに分けて捉える。光・アナログ計算では、LightSolverのLPUが、光をcomputationとmemoryの両方に使うphysics-based computingとして、最適化、固有値問題、PDE等の高速化を狙っている。ニューロモーフィックでは、Intel Loihi 2とHala Pointが、sparse event-driven computationとintegrated memory/computeによる高効率計算の研究基盤として提示されている。量子-HPC連携では、RIKENがReimeiおよびIBM Quantum System Twoと古典系を連携させるQuantum-HPC Hybrid Platformを整備しており、IBMとRIKENは2026年にFugakuと量子プロセッサをclosed-loopで連携させたquantum-centric supercomputingの実証を公表している。

2.2.1.3 共通技術トレンド

■**メモリ中心化とノード内統合** 現在のアクセラレータ競争で最も顕著なのは、設計の重心が「演算器中心」から「メモリ中心」へ移っている点である。H200は141GB HBM3eと4.8TB/sの帯域を備え、MI355Xは288GB HBM3eと8TB/s、MI300AはCPUとGPUが128GBのUnified HBM3を共有する構成を採る。クラウド側でも、TPU v6eは32GB HBMと1600GB/s、TPU7xは192GiB HBMと7380GB/s、Trainium2 UltraServerは6TBの共有アクセラレータメモリと185TB/sの総帯域を前面に出している。このことは、シリコン単体の演算性能よりも、どれだけ大容量・高帯域のメモリを演算器近傍へ置き、どれだけ転送オーバーヘッドを削減できるかが支配的な差別化要因になっていることを示している。

ノード内統合の面でも、同様の傾向が見られる。MI300AはCPUとGPUが同一メモリ空間を共有することでデータ転送の往復を抑制し、GH200は900GB/sのNVLink-C2CによってCPU-GPU間の強結合を実現している。独自アーキテクチャ層においても、PFNのL1000は3D-stacked DRAMによりmemory wallを別方向から崩そうとしており、SambaNova SN40Lはdataflow architectureと三層メモリを組み合わせている。したがって、将来機的设计空間は、デバイス単体性能ではなく、メモリ階層とコンパイラ/ランタイムを含むノード設計全体として評価する必要がある。

■**チップからラックへ、そしてモジュールへ** 第二の共通トレンドは、評価単位がchipからnode、rack、pod、modular systemへ移っている点である。JUPITERはBooster ModuleとCluster Moduleから成る動的モジュラー構成を採り、Booster側にNVIDIA技術、Cluster側に高メモリ帯域のSiPearl Rhea1を配置している。また、AWS Trainium2はchip単体ではなくTrn2インスタンスおよびUltraServer単位で性能を提示し、Microsoft Maia 200も標準Ethernetベースのscale-up networkを前提に6,144アクセラレータ規模までの拡張を想定している。Tenstorrent Galaxyも32基のWormholeをEthernet-based meshで束ねたサーバとして提示されている。すなわち、今後の評価単位は部品単体ではなく、system designとしてのnode / rack / podである。

■**インターコネクト、オフロード、設備の一体化** 第三の共通トレンドは、インターコネクト、オフロード、冷却・給電がアーキテクチャそのものになっていることである。UALink 1.0は200G/laneで最大1,024アクセラレータ規模のpodを想定したopen standardのscale-up interconnectを定義し、UEC 1.0はNIC、switch、optics、cablesを含むEthernetスタック全体をAI/HPC向けに最適化している。CXL 3.2でもCXL Memory Deviceの監視・管理・機能強化が進み、異種メモリ資源の共有と管理の標準化が前進している。

これらの接続技術は、DPU/IPUなどのオフロード層と密接に結びついている。BlueField、Intel IPU、

MangoBoost が共通して示しているのは、通信、ストレージ、セキュリティ、テレメトリをホスト CPU から切り離し、主計算資源を計算そのものへ集中させる方向である。さらに、ラックスケール化は冷却・給電を独立の制約条件ではなく設計変数へ押し上げている。たとえば GB300 NVL72 は fully liquid-cooled の rack-scale design として提示されており、将来機では設備設計と計算機設計を切り離すことができない。

2.2.1.4 今年度におけるアクセラレータ技術動向調査のまとめ

今年度の進捗として、以下の三点を報告できる。第一に、現行主流は依然として GPU/APU 系であり、次々世代機の主計算基盤としてはこの層が基準線であることを、TOP500、HPL-MxP、Green500 と主要ベンダーの仕様を照合することで確認した。第二に、MN-Core、Tenstorrent、MangoBoost、LightSolver、Loihi 2、量子-HPC 連携まで視野に入れると、将来システムの差別化は演算器単体よりも、メモリ、接続、offload、workflow integration に現れることを整理できた。第三に、したがって評価はチップ別の性能比較だけでは不十分であり、代表 workload、設備条件、ソフトウェア可搬性、運用性を含むシステム全体で行う必要がある、という設計上の示唆を得た。

2.2.1.5 来年度に向けたアクセラレータ技術動向調査研究の方向性

以上を踏まえると、来年度の研究開発方針は、主系統、探索系統、インフラ系統の三本立てで構成するのが妥当である。主系統では、市場実証済みの GPU/APU を基準に、代表 workload に対する性能・電力・冷却・可搬性の基準線を確立する。探索系統では、MN-Core、Tenstorrent、クラウド専用 ASIC に加え、LightSolver、ニューロモーフィック、量子-HPC 連携を small-scale testbed 上で継続評価し、どの workload に効果があるかを module 単位で見極める。インフラ系統では、DPU/RNIC、NVMe-oF、TOE、CXL、UALink/UEC などを含む data movement testbed を整備し、主計算資源からどこまで通信・I/O・ストレージを外出できるかを評価する。

この方針の要点は、新しいチップを列挙すること自体ではなく、異種アクセラレータを system design に落とし込む方法論を確立することにある。言い換えれば、主系統で実装可能性を担保し、探索系統で将来差別化を模索し、インフラ系統で実効性能を底上げするという三つの軸を並行に進めることが、次々世代スーパーコンピュータ研究開発の基本戦略になる。

2.2.2 メモリ技術に関する動向

2.2.2.1 標準 DRAM の動向

ここでは、主にオンチップおよびオフチップのメモリ技術の動向調査について述べる。2025 年度には、今後の標準 DRAM の方向性を決める大きな発表が相次いだ。

- HBM4 の JEDEC 標準化完了
- LPDDR6 の JEDEC 標準化完了
- DDR6 の方向性を JEDEC でほぼ合意

この JEDEC での標準化が完了した事により、DRAM ベンダー各社は、実製品製造、テストに注力し、いち早くサンプル出しを目指す動向である。上記の3つのメイン製品の最新動向として特筆されるのは、前世代 (HBM3、LPDDR5、DDR5) と、全く互換性がないという点である。つまり、コンピュータ産業界は、コンピュータボードから新設計し、用意周到に準備をして、実機での検証をしないと、「DDR6 使用ではチューニ

ングが不完全で不良が起こるから、今は前世代メモリの DDR5 を用いて製品化する」といったリスク回避的な対応となる。

HBM4 では、IO 数をこれまで HBM 機種 の 2 倍にあたる 2048 ピンに増やす大きな方向転換を行った。その IO ピン数の増加に伴い、uBump Ball ピッチを縮小したものの、HBM の PKG サイズは 11mmx13mm と前機種より大きくなった。そして、12 層 DRAM ダイ積層が主流となる予想であり、大容量版では 16 層積層も計画されている。また、各社の方針に依存するが、Base Logic ダイが DRAM プロセスからロジックプロセスに変更し、IO 部分の消費電力低減と高性能化を狙うと噂されている。HBM4 の最初のコンセプトとして、IO 数を増やして、IO 周波数が高くなるのを抑える事から、JEDEC の議論はスタートしたが、業界からの高いメモリ BW の要求に対して、既に、IO 周波数（データレート）は高い側にシフトし始めている。2048 ピンという物理配置の新しいチャレンジに加えて 10Gbps を超える高精度なシグナリングと、2 重のチャレンジに、実装後のシステム上での仕上がり具合や実機出荷の時期が懸念されている。2025 年 6 月には、KAIST のラボから HBM ロードマップ (HBM Roadmap Ver 1.7 Workshop、KAIST TERALAB、June 14th 2025) が発表されたが、各技術の深掘りはなく、単に希望的観測に基づく線形的な性能向上予測が示されているだけであった。昨年、HBM4 のスペックが確定した後、JEDEC では既に HBM5 のスペック議論が始まっているとの噂で、今後の方向性を示す情報公開が待ち遠しい。

LPDDR デバイスでは、パッケージやボードに直接はんだ付けする形態の性質上、SoC ベンダーからの強い要望を受け、システム的に ECC 用のビットを用意しなければいけなかった (in-bound, out-bound)。LPDDR6 では、従来のシステム的な ECC 方式から、チップ自体に、その ECC ビットをデバイスレベルで取り込んだ 12b x 2 サブチャネル = 24b のインターフェイスに変更した。従来チップと異なり、システムから見える実メモリ容量と実メモリ BW は、これらの ECC ビット分を差し引く必要が出て来る。チップ構成としては、従来のエッジ IO のダイを積層し、ワイヤボンディングで製造する方式を継承しているが、メモリモジュールとしては、薄型でも大容量を実現してきた Compression Attached Memory Module (CAMM) の次世代版 CAMM2 に移行する予定である。非公式な情報交換の場では、1 チップ内にワイヤボンディングで 16 DRAM ダイ積層する技術は、既に量産レベルまで達しており、十分な信頼性があるとの情報であった。

DDR6 も、LPDDR6 の動向を伺いながら議論していたためか、LPDDR6 に同調するように、24 ビット IO のインターフェイスに変更すると報告されているが、正式な報告は JEDEC からまだない。DRAM ダイのレイ構成やパッド配置まで、LP ダイに倣った変更の必要性は、まだ不明である。また、メモリモジュールの形態も、既存の DIMM から CAMM2 に移行の可能性が大であり、メモリデバイスのみならず、メモリモジュールの動向調査の必要性が伺える。

公開情報として検索できる各機種のまとめを以下の表に示す。

2026 年 2 月に開催された半導体のオリンピックと呼ばれる International Solid-State Circuits Conference (ISSCC) でも、世界で初めて HBM4 の発表が Samsung からあり、注目を集めていた。上述したように、これまで噂されていた Base Die が Logic プロセスでの製造された物と明言されていた。この Logic プロセスへの変更により、消費電力の低減と 13Gbps の高速周波数でもマージンある IO 動作が実現できたとの事であった。また、昔から噂されていた Cu Hybrid Bonding (CHB) は HBM4 でもまだ未使用で、従来通りの uBump ボンディングであるとの事も判明した。フォーラムセッションでも、HBM4 の電源配布ネットワークの言及が、同じ Samsung からされており、従来の HBM3 の 2 倍の TSV の数を DRAM コアダイに配置し、電源の安定供給を目指しているとの事であった。そして、CHB は更なる電源供給安定化への重要な将来技術の一つという位置づけではあるが、HBM4 には使用されていないとの事であった。メモリセッションでの標準 DRAM の発表としては、HBM 1 件、LPDDR6 2 件、GDDR7 1 件、Vertical Cell Transistor (VCT) DRAM 1 件の

	HBM4/4e	DDR6	LPDDR6
IO Count	2048	24	24
Data Rate	8 ~ 13 Gbps	8.8 ~ 17.6 Gbps	10.6 ~ 14.4 Gbps
BW@PKG	~ 3TB/s	~ 187 GB/s	~ 153 GB/s
BW@Module	-	~ 187 GB/s	~ 153 GB/s
Package Module	12/16-hi w/ TSV	CAMM2	SO-CAMM2
DRAM Die Capacity	3 ~ 4 GB	6 ~ 8 GB	~ 4 GB
DRAM Stack@PKG	-	2-hi (RDL) / 16-hi (WB)	Up to 16-hi (WB)
Capacity@PKG	36 ~ 64 GB	16 GB / 128 GB	~ 64 GB
Capacity@Module	-	256 GB / 2048 GB	~ 256 GB

合計 5 件であった。メモリセッション最後の論文（評価点数が低くボーダー擦れ擦れ）であった VCT の内容は、縦型 Floating Body のセルトランジスタで 4F2 の最小セルサイズを実現し、DRAM セルウェハと周辺回路のウェハを CHB でボンディングすると言う YMTC（中国）の NAND フラッシュのような製法で試作したテストデバイスであった。VCT は数十年前前からあるコンセプトであるが、まだ、実現されていない技術であり、そこに、以前 NAND Flash で発表された CHB を用いてセルウェハと周辺ウェハをボンディングするというコンセプトを初めて DRAM に導入したという、攻めた発表ではあったが、既存の技術の融合的な論文件数狙いの発表と感じた。メモリ容量も 16Gb レベルで、リフレッシュ時間など DRAM にとっては重要な情報は記載されておらず、まだまだ、研究開発段階で、製造コストも踏まえた量産には程遠いという感が伺える。

2.2.2.2 イメージングメモリの動向

VLSI Symposium 2025、IEDM 2025 からの論文調査から、以下のような項目の論文件数が多くなってきている傾向が見えてきた。

- IGZO などでは有名な酸化物系の半導体トランジスタとメモリ素子への応用
- Hf 系材料を使った強磁性を使ったメモリ素子
- Resistive 素子にセレクタを使った Selector Only Memory (SOM)

IGZO などでは有名な酸化物系の半導体トランジスタの発表件数が多数見受けられる。まだまだ、IGZO 系の材料の改善（組成比、結晶軸の変更など）や In や他の材料を用いた材料系の発表が大半であるが、DRAM と同じ 1T1C 型のメモリセルは少数で、2T、3T、2T+nC などの Gain Cell と呼ばれる色々なセル構造を用いた発表が多数あった。Gain Cell の場合、基本的には、SRAM の置き換えや L4 などの遅いキャッシュの置き換えを狙っていると予想されるが、通常の DRAM と同じセル・リフレッシュも必要であるため、このリフレッシュ動作による通常の READ/WRITE 動作を邪魔する事が実用上の欠点の一つとなる可能性がある。また、セルの素子数が通常の DRAM より多くなることでセル面積が大きくなるため、面積辺りのメモリ容量という観点からでも、通常 DRAM を置き換える候補にはなり得ないと予想される。

Hf 酸化物に強磁性が発見されてから、この Hf 系材料での強磁性メモリの研究が活発になり、論文数も多い。メモリ素子としては、一般的な 1T1C-FeRAM から、nvSRAM や nvDRAM など、更に、1Fe-FET、1T-nC という様々なセル構造の発表が見受けられる。ただ、まだ、「材料の研究開発」、「テストデバイスの試作」というフェーズであり、量産を踏まえた実機種の発表までには至っていない。また、書き込み動作時に、通常

の先端 CMOS ロジックでは使わない-2.0V~+2.0V という大きな負電圧を含んだ広い電圧範囲のため、先端 CMOS プロセスとの親和性の問題があり、大規模 LSI としての開発には至っていないのが現状である。

2015 年の 3D XPoint (3DXP) 発表と 2017 年以降の Intel Optane 製品化を契機に注目を集めた、DRAM とフラッシュのギャップを埋める Storage-Class Memory の中で、Intel と Micron が提案、開発し、量産化まで行った 3DXP (4F2 セルサイズの X ポイントセル構造で、セクタと呼ばれるダイオード機能の材料を用いてセルトランジスタを使用しない) 関連の発表が、再び件数を伸ばしている。最近では、このようなセクタ機能を中心とするメモリ構造が Selector-Only Memory (SOM) として議論されている。これらの論文は、基本的に、色々な素子材料の議論が中心であり、大メモリ容量の LSI チップとしての発表ではない。面白い事に、Intel より先にこの 3DXP 事業を中止した Micron から、再び論文が出ている点が興味深い。経営のトップ層が変わり、方針が変わったためであろうか？この Micron の論文はメモリデバイスの俯瞰 SEM 写真も載っているが、逆に材料組成比などは何も記載されていない。

イメージングメモリ分野では、材料開発を中心に、メモリ素子開発の論文は多いが、素子レベルや小容量テストシリコンの試作レベルで、量産に近いレベルでの具体的な製品機種の発表はない。ISSCC では、この手の発表が 0 であることがその傾向を顕著に表している。また、これらのデバイスは不揮発であるため、スタンバイ状態からの早期復帰、スタンバイ時の電力低減によりモバイルやエッジ用のデバイスやアプリケーションには効果を発揮するが、通常の READ、特に WRITE 時にはより電力を多く消費する点は、スパコンなどの常時動作しているシステムには、逆に不向きと考えられる。

2.2.2.3 CIM/PIM の動向

ISSCC 2026 では、CIM/PIM の 1 セッションが構成されていた。発表論文 9 件中、SRAM 4 件、ReRAM 1 件、Charge Trap(CT) Transistor 1 件、FeNAND 1 件、Gain Cell 1 件、3D DRAM Stack 1 件であった。数年前と比べると、CIM/PIM ブームがやや落ち着いてきた印象がある。合計 9 件の論文が発表されたが、基本的には、全てアカデミアからの発表で、企業が製品を目的とした発表は無かった。TSMC や Intel の名前も見受けられるが、事業部門ではなく、中央研究部門からの発表であることは指摘すべき点であろう。中国の Fudan University からの発表で、ロジックチップの上に DRAM ダイを 2 層積層する発表があった。これは、コロナ以前に、中国の UniIC が発表していたロジックウェハ上に DRAM (台湾の 25nmDRAM プロセス) を直接ボンディングした発表の後継論文であろう。確かに、UniIC の名前も著者の中に見受けられる。ボンディング技術は 3.5 μ m ピッチの Cu Hybrid Bonding(CHB) とアグレッシブであるが、発表者は W2W か W2C どちらかさえも知らない状態であった。

メモリ IP はファウンダリから提供されているためか、メモリの中身も分からないまま、試作して、発表しているケースが見受けられており、大学でも、新規メモリ IP を使って試作まで可能な環境であるが、メモリ中身や特性を知らずとも論文を出せるという、あるまじき状態に陥っていることを危惧する。ISSCC は回路学会なので仕方のないことであるが、最終製品のシステムや SW Stack まで議論することなく、性能評価していることを不気味に感じる。また、得てして、メモリ容量は、L1 キャッシュ並の小さな容量で議論されている点も不思議である。CIM/PIM の本来の問題であった「大容量メモリから大容量データを転送することが大きな消費電力を必要とする」という大前提を完璧に忘れていると言わざるを得ない。残念ながら、これが世の中にあふれている CIM/PIM の論文の傾向である。

2.2.3 半導体技術に関する動向

半導体技術に関する動向調査においては、大きく分けて3項目の技術に対し調査を行った。2.2.3.1節では半導体プロセス微細化技術について調査結果を示す。2.2.3.2節はチップレットアーキテクチャ・接続技術について調査結果を示す。2.2.3.3節は光電融合技術に関する調査結果を示す。2.2.3.4節は今後の課題である。

2.2.3.1 半導体プロセス微細化技術に関する調査

プロセスノードの微細化は、スーパーコンピュータの性能向上を支える最も根幹的な基盤技術の一つである。トランジスタ寸法の縮小により単位面積あたりに集積できる素子数（演算密度）が増大し、1チップあたりの計算性能は世代ごとに飛躍的な向上を遂げてきた。また、微細化は単に性能を高めるだけでなく、動作電圧の低減や配線の短縮を通じて消費電力あたりの性能、すなわち電力効率（FLOPS/W）の改善にも大きく寄与する。この電力効率の向上は、膨大な電力を必要とする次世代のゼタスケール計算の実現に向けて不可欠な要素である。さらに、トランジスタの高密度化はオンチップキャッシュの大容量化を可能にし、メモリアクセス遅延の低減を通じて実効性能の向上にも直結する。こうした背景のもと、最先端プロセスを実現できる半導体ファウンドリの技術力は、国家レベルの競争力や産業基盤にも大きな影響を与えている。以下に、世界の最先端半導体製造技術を有する主要ファウンドリ各社の動向について整理する。

TSMC（台湾）は、最先端ロジック半導体において引き続き業界をリードしているファウンドリであり、2nm世代プロセス「N2」の量産を2025年Q4に開始する計画である。このN2では、同社として初めてGAA（Gate-All-Around）トランジスタ構造を採用し、従来のFinFETからの大きなアーキテクチャ転換を実現する点が注目されている。すでにAppleやNVIDIAといった主要顧客が生産能力を確保しているとされており、先端ノードにおける需要の高さとTSMCの供給力の重要性が改めて浮き彫りになっている。

Samsung（韓国）は、TSMCに対抗する形でGAA技術をいち早く導入した点が特徴であり、2022年にはすでにGAAベースの量産を開始している。次世代ノードとしては、2nmプロセス「SF2」を2025年Q4に量産開始する計画であり、性能・電力効率の両面で競争力を高めることを狙っている。また、TSMCの先端プロセスにおける需要逼迫を背景に、一部の顧客企業がSamsungへの製造委託を検討する動きも見られる。さらにSamsungは、その先の1.4nmプロセス「SF1.4」を2029年に量産開始するロードマップを掲げており、中長期的にも先端ノード競争に積極的に関与していく姿勢を示している。

Intel（アメリカ）は、IDM（垂直統合型メーカー）としての強みを活かしつつ、ファウンドリ事業の強化も進めている。2025年4月には1.8nm世代相当となる「18A」プロセスの生産開始を予定しており、この世代ではRibbonFETと呼ばれるGAAトランジスタ構造に加え、PowerViaと呼ばれるバックサイド電源供給技術を導入する点が大きな特徴である。これにより、配線遅延の低減や電力効率の改善といったメリットが期待されている。現時点では主に自社CPUへの適用が中心と見られているが、将来的には外部顧客への提供を通じてファウンドリ市場での存在感を高めることも視野に入れていると考えられる。

Rapidus（日本）は、日本発の先端ロジック半導体量産を目指して設立された新興ファウンドリであり、IBMとの技術提携を軸に開発を進めている。2025年7月には2nmプロセスの試作に成功したとされており、2027年後半には量産開始を予定している。さらに、その先の1.4nm世代についても2029～2030年頃の実現を目標に掲げており、長期的な技術開発ロードマップを描いている。顧客候補としては富士通やTenstorrentなどが挙げられており、日本国内外の先端半導体需要を取り込むことで、新たなエコシステムの構築を目指している。

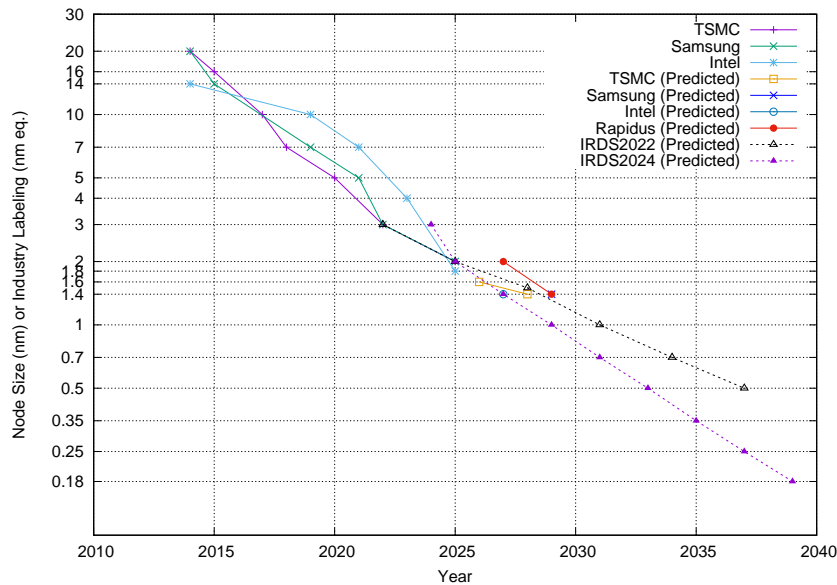


図 2.2.1 半導体微細化技術の年次進行（予測値を含む）

図 2.2.1 は、これまで述べてきた各先端ファウンドリにおける半導体微細化技術の推移を、年次ベースで整理したものである。ここで示す年は研究開発段階ではなく、実際に量産化が達成された時点に基づき、各社の実用技術としての到達度を比較することを目的としている。このように量産年に着目することで、技術的な先行性だけでなく、産業としての実行力や供給能力も含めた現実的な競争状況を把握することが可能となる。

これまでの実績を見ると、2015 年頃を境に Intel は微細化競争において TSMC や Samsung に対して後れを取る状況が続いてきた。一方で、TSMC および Samsung は 2022 年頃までは比較的順調にプロセスノードの縮小を進めていたが、GAA（Gate-All-Around）トランジスタへの移行期に入った 2022 年以降は、その進展がやや鈍化していることが読み取れる。これは、トランジスタ構造の抜本的な変革に伴う技術的難易度の上昇や製造プロセスの複雑化が影響していると考えられる。他方で Intel は、2025 年に 1.8nm 世代の量産開始を迎えることで急速にキャッチアップし、再び先端ノード競争における存在感を強めている。

将来の各社のロードマップについては、各社ともおおむね 2029 年頃までの見通しを公表している。この範囲では、Intel は 1.8nm 世代までの急速な進展の後、その成長速度が徐々に落ち着き、TSMC や Samsung と同程度のペースに収束する傾向が示されている。また、1.4nm 世代の量産開始時期については、Intel が 2027 年中、TSMC が 2028 年中、Samsung および Rapidus が 2029 年中と予測されており、各社の技術開発競争が引き続き数年単位で拮抗している状況が見て取れる。

2030 年以降の動向については各社から具体的なアナウンスはなく、国際半導体技術ロードマップである IRDS の予測が重要な指標となる。2022 年時点のロードマップでは、2031 年頃に 1.0nm 世代へ到達するとされており、本調査でターゲットとする 2030 年代の HPCI 基盤は、この 1.0nm クラスの微細化技術を前提とした半導体によって構築される可能性が高いと考えられる。さらに、0.7nm は 2034 年、0.5nm は 2037 年とされており、これらは 2030 年代半ば以降に位置付けられる技術であり、本調査のターゲットとして想定される HPCI 基盤への実装は時間的制約の観点から限定的になると見込まれる。

一方で、2024 年時点の IRDS では、この見通しに変化が見られ、微細化の進行が加速している点が興味深

い。具体的には、2031年に0.7nm、2033年に0.5nmへ到達するとの予測が示されており、従来の見積もりよりも数年単位で前倒しされている。このシナリオに基づけば、2030年代のHPCI基盤は0.7nmクラスの半導体を前提とする構成となり、導入時期が後ろ倒しとなるシステムについては0.5nmクラスの採用も現実的な選択肢として浮上する可能性がある。

もっとも、これらのIRDS2022およびIRDS2024の予測と、2029年頃までに公表されている各ファウンドリの実際のロードマップを比較すると、現時点ではIRDS2022の想定に近い進捗が示されているように見受けられる。よって、技術的な到達時期は加速シナリオよりもやや保守的に推移している可能性が高い。このため、本調査で想定する時期のHPCI基盤としては、1nm前後のプロセス技術を軸とするシナリオが最も現実的であり、0.7nmから0.5nm世代の技術については、将来的な発展を見据えたセカンドシナリオとして位置付けるのが妥当であると考えられる。

2.2.3.2 チップレットアーキテクチャ・接続技術に関する調査

チップレットアーキテクチャおよびその接続技術は、半導体の微細化が物理的・経済的な限界に近づく中で、性能とコストを両立するための重要なアプローチとして注目されている。従来のモノリシック（単一ダイ）構成では、ダイサイズの増大に伴う歩留まり低下や製造コストの上昇が大きな課題であった。これに対し、機能ごとにチップを分割するチップレット構成を採用することで、小面積ダイの組み合わせによる歩留まり向上が可能となる。また、CPUコア、IO、メモリ制御などを異なるプロセスノードで製造し、それらを統合することで、性能とコストの最適なバランスを実現できる。こうした設計柔軟性により、単なる微細化に依存しない形でシステム全体の性能向上を図ることが可能となっている。

具体的なチップレット構成の事例としては、各社の高性能プロセッサにおいてすでに広く採用が進んでいる。AMDのEPYC Turinでは、最大192コアの計算ダイ（3nmプロセス）を、IOダイ（4nmプロセス）と接続する構成が採られており、異種プロセスの組み合わせによる高性能化とコスト最適化を実現している。Intelも同様にチップレット化を進めており、Sapphire Rapidsでは4分割のチップレット構成、後継のEmerald Rapidsでは2分割構成を採用することで、世代ごとに設計の最適化を図っている。またAppleのUltraシリーズでは、2つのMAXクラスのCPUチップを接続することで、単一ダイでは困難な大規模SoCを実現しており、コンシューマ向け分野においてもチップレット的アプローチが広がっていることが分かる。

これらのチップレットを統合するためには、高帯域かつ低遅延の接続技術が不可欠であり、各社が独自技術を開発・実用化している。IntelのEMIB（Embedded Multi-die Interconnect Bridge）は、シリコンブリッジを用いた2.5D接続技術であり、第2世代ではマイクロバンプピッチが55μmから45μmへと微細化され、より高密度な接続を実現している。TSMCのInFO-LSI（Integrated Fan-Out Local Silicon Interconnect）は、Intel EMIBと同様にローカルなシリコンインターコネクタを再配線層に内蔵することで、巨大なシリコンインターポーザを不要にしつつ高帯域接続を可能とする技術であり、AppleのUltraシリーズに採用されている。さらに、AMDはInfinity Fabricと呼ばれる専用的高速インターコネクタを用いてチップレット間を接続し、大規模プロセッサにおいても高いスケーラビリティと帯域を確保している。

加えて、チップレット接続のエコシステム拡大に向けた標準化の動きとして、UCIe（Universal Chiplet Interconnect Express）が注目されている。UCIeには約100社が参加しており、異なるベンダ間でのチップレット相互接続を可能とする共通仕様の策定が進められている。これにより、従来はベンダ固有であった接続技術がオープン化され、より柔軟なシステム設計やサプライチェーンの多様化が期待される。商用展開は2026年から2028年頃にかけて本格化すると見込まれており、今後の半導体設計における重要な基盤技術の一つになると考えられる。

さらに、チップレット統合技術の発展形として注目されているのが、Rapidus が取り組むガラスインターポーザである。Rapidus は、600mm × 600mm という大面積の四角形ガラス基板を用いた大型再配線インターポーザのプロトタイプ実証を進めており、従来のシリコン（Si）インターポーザとは異なる新しい実装アプローチを提示している。この技術は、量産性、面積効率、配線密度といった観点において従来技術と一線を画す可能性を持っており、今後の大規模チップレット統合基盤として期待されている。

ここでいうインターポーザとは、CPU や GPU、HBM メモリといった複数の半導体チップを高密度に接続するための中間基板であり、いわばチップ間通信を担う基盤である。このインターポーザ上には再配線層（Redistribution Layer, RDL）が形成され、微細な配線によってチップ間を接続することで、高帯域かつ低遅延のデータ通信を実現する。特に HPC 分野では、チップレット化に伴いチップ間通信の重要性が増しており、インターポーザ技術はシステム性能を左右する重要な要素となっている。

Rapidus のガラスインターポーザが注目される理由の一つは、基材としてガラスを用いる点にある。従来のシリコンインターポーザでは、単結晶シリコンという高価な材料を使用するため、大面積化に伴いコストが急激に増大するという課題があった。一方でガラスは比較的低コストで大面積基板の製造が可能であり、製造コストの低減に寄与する。また、600mm 角といった大面積かつ角形の基板を採用することで、ウェハ（円形）に比べて面積利用効率が高く、1 枚の基板からより多くのインターポーザを切り出すことが可能となる。このように、材料と形状の両面からコストおよび生産効率の改善が期待されている。

一方で、このガラスインターポーザの実用化に向けては重要な技術課題も残されている。その代表例が Through Glass Via（TGV）であり、これはガラス基板に微細な貫通孔を形成し、その内部に電極を通す技術である。高密度な外部 I/O や電源供給を実現するためには、この TGV を用いた高密度配線が不可欠であり、これが実現できなければ従来のシリコンインターポーザを置き換えることは難しいと考えられる。逆に、この課題が解決されれば、大規模チップレット統合基盤としての競争力が飛躍的に高まり、「メモリ以外は国産」といった形でのスーパーコンピュータ用半導体の実現にも道が開かれる可能性がある。

2.2.3.3 光電融合技術に関する調査

光電融合技術は、従来の電気配線によるデータ伝送を光通信へと置き換えることで、次世代計算基盤における性能と効率を飛躍的に向上させることが期待されている。電気信号に比べて光信号は高帯域かつ低損失で伝送可能であり、帯域幅の大幅な増加、レイテンシの低減、さらには消費電力の削減といった複数の利点を同時に実現できる。このような特性は、大規模並列処理を前提とする HPC や AI ワークロードにおいて特に重要であり、ノード間・チップ間通信のボトルネックを解消することで、プログラム全体の実効性能とエネルギー効率の向上に直結する技術として位置付けられている。

この分野における中核技術の一つが CPO（Co-Packaged Optics）であり、光トランシーバをスイッチやプロセッサと同一パッケージ内に統合することで、従来の外付け光モジュールに比べて信号損失と消費電力を大幅に削減することが可能となる。Broadcom は、CPO 技術を用いたイーサネットスイッチ「Tomahawk 6-Davisson」（102.4Tbps）を世界で初めて量産化し、超大容量ネットワークの実現に向けた先駆的な取り組みを示している。NVIDIA も、TSMC の COUPE 技術を活用し、Quantum-X InfiniBand や Spectrum-X といった製品においてパッケージ内光変換の統合を進めている。さらに Intel は、Optical Chip Interconnect と呼ばれるプラットフォームを通じて光インターコネクットの垂直統合を推進しており、2028 年頃の実用化を目指している。また、Ayar Labs は光 I/O チップレットを用いた世界初の光チップレット接続技術を開発しており、チップレットアーキテクチャと光通信の融合という新たな方向性を提示している。

一方、データセンターネットワークにおいても光通信の高度化が急速に進展している。光トランシーバ市場

では、Innolight、Coherent、Eoptolink といった主要ベンダが大規模供給を担っており、AI データセンターの需要拡大を背景に技術革新が加速している。通信規格としては、800G Ethernet が 2024 年頃から本格展開され、すでに AI 用途における標準インフラとなりつつある。さらに、次世代の 1.6T Ethernet も 2025 年頃から導入が始まる見込みであり、OSFP-XD といった新しいフォームファクタのもと、ハイパースケール企業を中心に積極的な採用が進んでいる。

2030 年代に向けては、光電融合技術はさらに進化し、チップ内部やチップレット間における光インターコネクットの量産化が進むと予測される。これにより、電気配線の限界を超えた高帯域・低消費電力な通信が実現され、光電融合プロセッサの実用化が現実味を帯びてくる。また、ネットワーク側においても 100Tbps 級の光スイッチング技術や、10Tbps を超える Ethernet 規格の標準化が進むと見込まれ、計算ノードからシステム全体に至るまで、光を中心とした新たな計算・通信基盤が形成される可能性が高い。

2.2.3.4 半導体技術に関する動向調査に関する今後の課題

今後の課題として、まずプロセスノード微細化技術に関しては、2032 年頃の HPCI システムに採用されることが期待されるプロセス技術の動向を継続的に調査する必要がある。IRDS や各社のロードマップは毎年更新されるため、その追跡調査が必要である。また、単にロードマップを追うだけでなく、各世代における性能向上、消費電力、製造コストといった定量指標の推移を整理・分析し、それらが実際のシステム設計に与える影響を評価することが重要である。特に、トランジスタ構造の変化や製造難易度の上昇がコストや歩留まりに与える影響を踏まえ、どのプロセス世代が実用的な選択肢となるかを見極める必要がある。

次に、チップレットアーキテクチャおよび接続技術については、各社が採用する設計思想や実装技術の詳細な動向を把握し、比較評価を行うことが求められる。加えて、2032 年頃の HPCI システムにおいて主流となる可能性のあるチップレット技術、ならびにそれを支えるインターコネクタ技術についても調査を進める必要がある。特に、標準化動向 (UCIe など) と各社独自技術の関係性や、実装密度・帯域・消費電力のトレードオフを踏まえた設計指針の確立が重要な検討課題となる。

さらに、光電融合技術に関しては、CPO (Co-Packaged Optics) や光 Ethernet といった要素技術について、より詳細な技術調査を行う必要がある。加えて、2032 年を見据えた電気・光ハイブリッドシステムにおいて、どの階層 (チップ内、チップ間、ノード間) で光技術を適用するのが最適かという棲み分けの検討も重要となる。通信距離、帯域要求、消費電力、実装コストといった観点から、電気配線と光配線の最適な役割分担を明確にすることが求められる。

以上の各項目について、今後も継続的な情報アップデートを行うとともに、実現可能性や技術的困難度の観点からの詳細な分析を進める予定である。これにより、2030 年代前半の HPCI 基盤に向けた現実的かつ競争力のある技術選択とシステム設計指針の確立を目指す。

2.3 Intel における将来システムの動向調査

Intel は 2030~2032 年にかけての技術ロードマップにおいて、HPC および AI における高度な計算需要に対応するため、ヘテロジニアス計算アーキテクチャ、先端パッケージング技術、ならびにソフトウェアフレームワークの統合を戦略的な重点領域として明確に位置付けている。本ロードマップは、Intel が掲げる「IDM 2.0」戦略、ならびにプロセスリーダーシップ、先端パッケージングの革新、ソフトウェアファーストの設計理念へのコミットメントに基づいて構築されたものである。

ロードマップにおいては、オングストローム世代製造プロセス (Intel 18A) への移行、Foveros 技術による

先進 3D パッケージング、CXL (Compute Express Link) によるメモリファブリックの拡張、高帯域インタコネクトを実現するシリコンフォトニクス統合、多様な計算アーキテクチャを横断する統一プログラミングモデルとしての oneAPI の進化が重要な要素となっている。さらに、エクストリームスケール・コンピューティングが直面する課題に取り組むための革新的なメモリアーキテクチャの実現を目指し、ソフトバンク株式会社の完全子会社である SAIMEMORY 株式会社と協働する ZAM (Z-Angle Memory) 構想も不可欠な要素として位置付けられている。

2025 年秋には、Intel と NVIDIA 社が AI インフラおよびパーソナルコンピューティング製品の構築に向けて協業し、NVIDIA 社が Intel の普通株式に 50 億ドルを投資することを発表した。

Intel の CEO であるリップ・ブー・タンは次のように述べている。

「Intel の x86 アーキテクチャは数十年にわたり現代コンピューティングの基盤となってきた。我々は将来のワークロードを支えるため、製品ポートフォリオ全体で革新を進めている。Intel のデータセンターおよびクライアント計算プラットフォームに加え、プロセス技術、製造、先端パッケージング能力は、NVIDIA 社が有する AI およびアクセラレーテッドコンピューティング分野でのリーダーシップを補完し、業界に新たなブレークスルーをもたらすものである。今回の投資を通じてジェンセン氏および NVIDIA チームが我々に寄せた信頼に感謝するとともに、今後、顧客のための革新と事業成長に向けて共に取り組んでいくことを楽しみにしている。」

2.3.1 調査研究の概要

1) ハイブリッドノードアーキテクチャ：Xeon と GPU の統合ロードマップ

• Xeon ロードマップ (公開情報)

Intel は「Granite Rapids」世代 (2024~2025 年) までの Xeon プロセッサを進化させるべく、2030 年に向けた開発を継続している。コア数の増加、AI アクセラレーション能力の強化、DDR5 および高帯域 HBM の統合によるメモリ帯域幅の向上が重点要素となっている。

• GPU 統合戦略

Intel は Arc GPU アーキテクチャおよびデータセンター向け GPU ロードマップにおいて、CPU と GPU の一体化を一層強化する方針である。Xeon プロセッサと GPU 計算ユニット間でメモリを統一するアーキテクチャやキャッシュ階層を共有する設計に関する研究を推進しており、本技術は 2020 年代後半の導入を目指している。

• ヘテロジニアス計算ビジョン

Intel は 2030 年以降に向けて、CPU、GPU、メモリ、I/O 資源をワークロードに応じて動的に再構成可能とする「分離型 (disaggregated)」計算アーキテクチャを中核に据えている。本アプローチは、Intel が展開する Compute Express Link (CXL) 技術を活用し、多様な HPC および AI 用途に最適化した柔軟かつスケール可能な計算ノードを実現するものである。

• 性能目標

Intel は、プロセス技術の進展、アーキテクチャ改良、ヘテロジニアス計算の統合を組み合わせることにより、2030 年までに HPC ワークロードで 5~10 倍の性能向上を達成することを目標としている。これらの性能向上は、エネルギー効率指標を維持または向上させながら実現するものである。

2) 先端パッケージング技術：HPC 向け EMIB および Foveros

• EMIB 技術の進化

Intel は、チップレット間を単一パッケージ内で高帯域に接続する Embedded Multi-die Interconnect Bridge (EMIB) のロードマップを策定しており、2020 年代後半に 4 Tbps を超える帯域を実現し得る能力を備え、CPU、GPU、メモリコントローラ、専用アクセラレータなど多様な計算要素をシームレスに統合可能とする技術の確立を目指している。

• Foveros 3D パッケージング

Intel は、Foveros 技術のロードマップとして「Foveros Omni」および「Foveros Direct」を位置付けており、それぞれ異なる統合要件に対応する設計となっている。シリコン貫通ビア (TSV) のピッチを 25 ミクロン以下にスケールさせることで、能動シリコンダイを高密度に 3D 統合することを可能とする技術であり、HPC 用途においては、メモリコントローラと計算要素を最適な近接配置に置くことで、レイテンシおよび電力効率の改善を図ることができる。

• チップレットエコシステム戦略

Intel は、Universal Chiplet Interconnect Express (UCIe) 標準の推進を通じて、オープンなチップレットエコシステムの構築を進めている。本アプローチにより、サードパーティ製チップレットや専用アクセラレータを Intel パッケージ内に統合することが可能となり、用途に応じて最適化された柔軟なコンピューティングソリューションの実現につながる。

• 製造プロセス統合

Intel は、先端パッケージング能力を活用し、単一パッケージ内で複数のプロセスノードを併用する設計を進めている。これにより、各チップレットをその機能に最適化したプロセスで製造することが可能となる。高性能ロジックには先端プロセスを適用し、一方、I/O およびアナログ機能には実績のあるコスト効率の高いプロセス技術を活用する方針である。

3) メモリ拡張および CXL 技術の展望

• CXL ロードマップおよび採用状況

Intel は、Compute Express Link (CXL) 技術について、CXL 3.0 以降を見据え、複数ノード間でメモリをプール化し、クラスタ全体にまたがる共有メモリファブリックを構築可能にすることを目指している。また、メモリベンダとの協業を通じて、ノードあたりテラバイト級の容量に対応する CXL 接続メモリモジュールの開発を進めている。

• メモリプーリングおよびディスクアグリゲーション

Intel は、CXL を活用したメモリディスクアグリゲーションにより、アプリケーションの要件に応じてメモリ資源を複数の計算ノード間で動的に割り当てる環境の実現を目指している。本アプローチにより、システム全体の資源利用効率が向上し、各ノードの物理メモリ境界を超えてメモリ資源へアクセスできる環境を実現する。

• 永続メモリ統合

Intel は、CXL インタフェースを用いた永続メモリ技術の統合に取り組んでいる。Optane は一部市場で提供が終了しているものの、Intel は CXL プロトコルを通じてアクセス可能なストレージクラスメモリソリューションの開発を継続している。

- **性能特性**

Intel は、CXL を介したリモートメモリアクセスにおいて 200 ナノ秒未満のレイテンシを目指しており、CXL インタフェース 1 本あたり 512GB/s を超える総帯域を 2020 年代後半に実現することを想定している。これらの性能は、多くの HPC 用途において透過的なメモリプーリングを可能とするものである。

4) ノード間通信に向けたシリコンフォトニクス

- **シリコンフォトニクス技術ロードマップ**

Intel は、シリコンフォトニクス技術に対して大規模な投資を進めており、2020 年代後半の導入を見据えたコパッケージドオプティクス (CPO) ソリューションを掲げている。光インタフェース 1 本あたり 1.6 Tbps を超えるデータレートと、電気式インタコネクタ比で 10 倍のエネルギー効率改善を達成する計画である。

- **統合戦略**

Intel は、フォトニックトランシーバをプロセッサパッケージに直接統合することで、電気-光変換に伴うボトルネックを排除する方針であり、フォトニックインタフェースを統合したプロトタイプシステムをすでに実証し、光部品サプライヤとの協業も発表している。

- **波長分割多重 (WDM) 技術**

Intel は、光ファイバ 1 本あたり 32 波長以上に対応する高度な波長分割多重 (WDM) 技術の開発を進めている。本技術により、帯域密度が大幅に向上し、大規模システムにおけるケーブル構成の複雑性を低減することが可能となる。

- **ネットワークアーキテクチャの革新**

Intel は、HPC 特有の通信パターンに最適化された光スイッチングファブリックおよび再構成可能な光ネットワークに関する研究を進めている。これらのアーキテクチャは、従来のメッセージパッシングインタフェースだけでなく、新たな分散計算モデルにも対応し、最小限のレイテンシオーバーヘッドで通信を実現するものである。

5) oneAPI の進化およびクロスアーキテクチャプログラミング

- **oneAPI ロードマップ**

Intel は、CPU、GPU、FPGA、アクセラレータアーキテクチャを横断する統一プログラミングモデルとして、oneAPI を継続的に進化させる方針である。コンパイラ技術の強化、対応言語の拡充、HPC および AI 分野で広く利用されるフレームワークとのさらなる統合を 2030 年まで進める計画である。

- **コンパイラおよびランタイムの高度化**

Intel は、ヘテロジニアスアーキテクチャ向けに最適化されたコードを自動生成する AI 駆動型コンパイラ最適化技術の研究を進めている。これらのコンパイラは、アプリケーション特性を解析し、プロセッサの割当てやメモリ配置についてインテリジェントに判断する仕組みを備えている。

- **性能ポータビリティ**

Intel は、oneAPI の重点領域として性能ポータビリティを掲げており、アプリケーションが異なるハードウェア世代や構成においても一貫した性能を発揮できるようにする方針である。この取組には、基盤となるハードウェア能力に応じてアルゴリズムやデータ構造を自動的に適応させる仕組みの高度化が含まれている。

- **エコシステム統合**

Intel は、主要な HPC および AI ソフトウェアフレームワークとの連携を進めており、MPI 実装、OpenMP、TensorFlow、PyTorch との統合を発表している。これらの統合により、既存アプリケーションが最小限のコード修正で先端ハードウェアの能力を活用できる環境を実現する。

6) ZAM (Z-Angle Memory) 構想および SAIMMEMORY 社との協働

- **ZAM プログラム概要**

Intel は、SAIMMEMORY 社との戦略的協働を通じて、Z-Angle Memory (ZAM) プログラムを掲げている。SAIMMEMORY 社は東京を拠点とし、画期的な積層型 DRAM アーキテクチャの開発を進めている企業である。この協働は、現行メモリ標準では対応が困難となりつつある AI および高性能計算 (HPC) アプリケーションの需要に応えるために設計された、メモリ技術の重要な進展を示すものである。

- **SAIMMEMORY 社の技術アーキテクチャ**

本協働では、SAIMMEMORY 社が開発する革新的なスタック型 DRAM アーキテクチャに焦点を当てており、現行の高帯域メモリ (HBM) 標準を上回る性能を目指して設計されている。本技術はメモリ容量を大幅に拡大するとともに、電力消費を大幅に削減し、AI システムを極限性能までスケールする際に生じるパッケージング上の重要なボトルネックを解消する能力を備えている。

- **Intel の役割と貢献**

Intel は ZAM プログラムにおいて、テクノロジー、イノベーション、および標準化の協働者としての役割を担うとともに、SAIMMEMORY 社が中核テクノロジー開発、イノベーション主導、商用化マネジメントを担当する構成となっている。Intel の Advanced Memory Technology (AMT) R&D プログラムおよび Next Generation DRAM Bonding (NGDB) イニシアチブで培った基盤的取組を踏まえ、同プログラムに貢献していく予定である。

- **Advanced Memory Technology の基盤**

ZAM プログラムは、米国エネルギー省が支援し、サンディア国立研究所、ローレンスリバモア国立研究所、ロスアラモス国立研究所により運営される Advanced Memory Technology (AMT) R&D プログラムの研究基盤を活用している。Intel は AMT による初期支援の下、ZAM テクノロジーの基盤となる積層 DRAM コンセプトに関する重要な実証および性能検証を行い、同分野の発展に寄与している。

- **次世代 DRAM ボンディングのイノベーション**

Intel は Next Generation DRAM Bonding (NGDB) イニシアチブにおいて、DRAM 性能を大幅に向上させつつ、消費電力を削減し、メモリコストを最適化する画期的なメモリアーキテクチャと革新的なアセンブリ手法を実証している。Intel Government Technologies の Intel フェローおよび CTO である Dr. Joshua Fryman は次のように述べている。「標準的なメモリアーキテクチャでは AI の要求に応えられない。そのため NGDB は、次の 10 年を加速するための全く新しいアプローチを定義したのである」。

- **開発スケジュールおよび商業化**

ZAM の研究開発は開始されており、2027 年度中にはプロトタイプを作製し、2029 年度中の実用化を目指す計画である。このタイムラインは Intel の広範なテクノロジーロードマップと整合しており、次世代 AI および HPC システムで不可欠となるメモリ帯域幅および容量の要件に対応するものである。

- **国際テクノロジーパートナーシップ**

Intel は、AMT 研究から ZAM 協働体制への移行を通じて、米国と日本の間における信頼性の高いテクノロジーパートナーシップを強化しており、国立研究所における研究から世界的な商用展開への移行を加速している。Intel は、このパートナーシップモデルを通じ、世界的な計算課題に取り組む上で国際協働を重視する姿勢を示している。

• 技術的性能目標

具体的な性能指標は現在策定中であるものの、ZAM テクノロジーは現行 HBM 規格と比較して、メモリ密度、帯域幅、エネルギー効率の大幅な向上を目標としている。積層型 DRAM アーキテクチャは、将来のアプリケーションで要求される極端な性能レベルに AI システムをスケールアップする際の制約となる重要なボトルネックを緩和するよう設計されている。

7) 新興テクノロジーと研究方向性

• 先進アーキテクチャ研究

Intel は、データ移動を最小化するために、計算処理をデータ格納位置の近傍に配置するメモリ中心型コンピューティングアーキテクチャに関する研究を進めている。本アプローチは、データ移動に伴うエネルギーコストを低減する上で有効であり、大規模科学計算や AI アプリケーションにとって特に重要である。ZAM プログラムは、こうした研究方向性を構成する重要な要素となっている。

• 量子コンピューティング統合

Intel は、量子コンピューティングに関する研究プログラムにも取り組んでおり、量子と古典計算を組み合わせたハイブリッドコンピューティングシステムを研究方針の一つとして位置付けている。具体的な統合タイムラインは検討段階にあるものの、量子プロセッサ試作機を実証するとともに、特定の計算カーネルを対象とした潜在的な統合経路についても積極的に検討を進めている。

• ニューロモーフィックコンピューティング

Intel は、Loihi ニューロモーフィックプロセッサに関する研究も推進しており、特定の AI 推論および学習アプリケーション向けに、ヘテロジニアスコンピューティングシステムへの統合可能性を視野に入れている。これらのプロセッサは、従来型コンピューティング要素を補完する超低電力 AI 能力を提供するものである。

• プロセステクノロジーのリーダーシップ

Intel の「4 年間で 5 ノード」ロードマップには、Intel 18A プロセステクノロジーが含まれており、2030 年に向けた継続的なスケールアップが位置付けられている。Intel のプロセスリーダーシップは、ZAM およびその他の先進コンピューティングアーキテクチャに必要なトランジスタ密度と性能特性を実現するものである。

8) 実装タイムラインおよび市場展開

• 短期マイルストーン (2025-2027 年)

ハイエンドプロセッサにおける先進的パッケージング技術の初期導入、エンタープライズシステムにおける CXL 2.0/3.0 の採用、データセンターアプリケーションにおける初期シリコンフォトニクス統合、ZAM プログラムの開始 (2026 年第 2 四半期)、および ZAM プロトタイプ開発 (2027 年) が目標として位置付けられている。

• 中期展開 (2028-2030 年)

ロードマップ上には、ハイブリッド CPU-GPU アーキテクチャの本格展開、CXL メモリプーリングの成熟した実装、HPC インタコネクト向け量産シリコンフォトニクスソリューション、そして 2030 年に向けた ZAM テクノロジーの商用化が位置付けられている。

・長期ビジョン（2030-2032 年）

完全に統合されたヘテロジニアスコンピューティングシステムが打ち出されており、メモリおよびコンピュータ資源の透過的な割当、成熟した光インタコネクトファブリック、さまざまな計算パラダイムを支援する包括的な oneAPI ソフトウェアエコシステム、そしてエクストリームスケール・コンピューティング要件に対応する ZAM 対応量産システムが、その重要な要素として位置付けられている。

9) 結論

Intel の 2030～2032 年に向けたテクノロジーロードマップでは、高性能計算および人工知能分野における計算課題に包括的に対応するアプローチが示されており、中でも、SAIMEMORY 社との画期的な ZAM 協働体制は、メモリアーキテクチャに革新的進展をもたらす取り組みとして位置付けられている。先進製造プロセス、革新的パッケージングテクノロジー、メモリファブリックの進展、光インタコネクト、およびブレイクスルーメモリテクノロジーの統合により、人類が直面する最も高度な計算課題に対応可能な次世代科学計算システムの基盤が形成されると見込まれる。

CXL や UCIe などのオープン標準、oneAPI に代表される統一プログラミングモデル、日本との ZAM 協働を含む国際パートナーシップ、そしてエコシステム協働への注力は、世界の研究機関や高性能計算センターが求める多様な要件を支援する Intel のテクノロジーポートフォリオを強化するものである。ZAM プログラムがもたらす変革的メモリ能力を核とした本ロードマップは、将来の科学研究、気候モデリング、材料科学、人工知能アプリケーション、さらにはエクストリームスケール・コンピューティング能力を必要とする新たな計算パラダイムにおいて想定される計算需要に整合している。

ZAM プログラムでの協働体制は、Intel が国際的テクノロジーパートナーシップに取り組む姿勢を体現するものであり、国立研究所における基盤研究が、戦略的産業パートナーシップを通じて商用展開へと円滑に移行しうることが明確にするものである。本取り組みは、最終的に世界の科学技術コミュニティ全体の利益に資するものである。

2.3.2 次年度の調査研究の計画

令和 8 年度（上半期）：CPU+GPU/AI 混載ノード構成含む多様なアーキテクチャ選択肢やロードマップについて調査検討を継続し、加えてチップレット構成技術およびメモリ関連技術の初期の調査検討を実施する。

令和 8 年度（下半期）：チップレット構成技術およびメモリ関連技術の初期の調査検討を継続し、加えてシリコンフォトニクス技術の初期の調査検討を実施する。プログラミング環境の運用事例の調査を実施する。

2.4 NVIDIA における将来システムの動向調査

2033 年までに、科学および産業の発見への道筋は、密接に結び付いた三つの柱、すなわち大規模高性能コンピューティング（HPC）シミュレーション、人工知能（AI）、および量子コンピューティング（QC）に基づくものになる。これらはそれぞれ独立した分野としてではなく、単一の「ディスカバリーエンジン」として連携して機能する。それにより、スーパーコンピュータは受動的なシミュレーションツールから、設計空間を探

索し、新しい物理法則、材料、生体構造をリアルタイムで生成する能動的な協働パートナーへと変化する。

この文脈において、NVIDIA のロードマップは、従来のシリコンスケーリングの限界を超えることを目的としている。これは、データセンター全体を計算単位として扱う「System on a Rack」という思想と、アプリケーション科学者とのハードウェアとソフトウェアの共同設計を通じて実現される。これによりアーキテクチャは、進化し続けるモデルやワークフローに迅速に適応することが可能になる。本概要では、技術の進歩、ソフトウェアの発展、そして日本の主権技術と協業によって推進される、シリコン制約によるスケーリングからデータセンターレベルのスケーリングへの移行について説明する。

2.4.1 調査研究の概要

2.4.1.1 システムアーキテクチャ

Blackwell (2024 年)、Rubin (2026 年)、Feynman (2028 年) に続き、2030 年から 2033 年の期間には、3D 統合技術と、GPU、CPU、ネットワーク、ソフトウェア、電源、冷却、および施設を含む共同設計に基づく「AI ファクトリー」モデルによって支えられた新しい GPU 世代が登場する見込みである。CPU と GPU を統合した「スーパーチップ」は、3D ロジックスタッキングを用いた異種タイルベースの「スーパーノード」へと進化する可能性がある。ここでは、将来の CPU 命令セットアーキテクチャ、GPU アクセラレータ、そして新しいタイプのアクセラレータが垂直統合され、最小遅延かつ極めて高帯域なコヒーレントメモリファブリックを共有することになりなる。

2.4.1.2 メモリアーキテクチャの進化

現在のメモリ性能向上のペースを維持するためには、将来の HBM 世代は GPU あたり数百 TB/s 規模の総帯域幅とテラバイト級の容量を提供する必要がある。しかし、現在の HBM 技術では、ソケットあたりの電力消費を桁違いに増加させることなしにこれらの目標を達成することは不可能である。このギャップに対処するため、複数の技術が検討されている。一つのアプローチは、データ集約型の前処理をオフロードするためのロジックを統合したカスタム HBM である。別の選択肢としては、ウェハ接合によってロジック層とメモリ層を垂直方向に統合する新しい 3D 積層メモリの概念がある。また、Vertical-Gate セルにより、より高密度のスタックが可能になる。さらに、DRAM、SRAM、およびロジックを組み合わせた異種 3D スタックによって、HPC および AI のワークロードに最適化された構成が実現される。

2.4.1.3 先端パッケージング材料

現在、パッケージ接続は主に有機基板技術に依存している。この技術は、電気的特性に優れているが、HBM のような高密度インターフェースを支えるための微細なピッチには対応できない。100mm × 100mm を超えるパッケージでは、有機材料の代替としてガラス基板が検討されている。これにより Through-Glass Via (TGV) が実現され、より細かなピッチと大幅に高い信号密度が可能になる。また、シリコンまたはガラス基板に刻まれたマイクロ流体冷却チャンネルは、計算タイルあたり数 kW に達する熱負荷を管理するための有望な手法である。

2.4.1.4 ネットワークとフォトニクス

2030 年までに、Co-Packaged Optics (CPO) は標準的な技術になり、これにより優れた電力効率が実現される。フォトニクスをネイティブに統合した CPO を備えた Quantum-X InfiniBand および Spectrum-X

Ethernet は、Tb/s 級ポートを低い電力消費で実現し、100 万 GPU 規模のネットワークファブリックを可能になる。NVLink ドメインの拡張により、ラックは長文コンテキスト推論や HPC-AI パイプラインのための計算ユニットとして扱われるようになる。

2.4.1.5 ラックおよび施設インフラ

ラックはメガワット級の電力密度へと移行することを想定している。800V DC 配電を採用することで銅の使用量を 90 %削減し、電力変換損失を低減する。現地でのエネルギー生成（太陽光、風力、原子力）とバッテリー蓄電、AI による監視（リアルタイムの電力および温度追跡、予測的故障解析、施設のロボット化）、および排熱回収により、運用効率を維持しながら環境への影響を最小化することが可能になる。

2.4.1.6 ソフトウェアとワークロードの進化

AI for Science (AI4S) および AI for Industry (AI4I) の進展により、ソフトウェアの役割は大きく変化している。従来は静的なライブラリを用いて計算を実行することが中心であったが、今後は数十万 GPU 規模のポストエクサスケール並列計算を統合的に管理する自律型のシステムスケールオーケストレーターへと進化していく。こうしたソフトウェアの進化により、HPC シミュレーション、AI、量子コンピューティング (QC) の計算資源を統合的に活用することが可能となる。このような環境を支えるため、次世代の NVIDIA ソフトウェアスタックは、従来の HPC 開発で用いられてきた Fortran、C、C++ などの言語と、Python や PyTorch といった現代の AI フレームワークを橋渡しする役割を担うものである。これにより、従来のシミュレーションコードと AI 技術を同一の計算環境の中で連携させることが可能となる。さらに、システム全体の運用においてもソフトウェアの役割は拡大する。例えば、自己修復型ファブリックでは AI エージェントがハードウェア障害を事前に予測し、ネットワークトラフィックを自動的に再ルーティングする仕組みが導入される。また、ソフトウェア定義の自律的電力管理は「パワーガバナナー」として機能し、厳しいエネルギー制約の下でもワークロードの進行を維持しながら科学的成果を最大化するようにハードウェアを自動制御するものである。こうしたソフトウェア基盤の進化に伴い、計算ワークロードの形態そのものも変化すると考えられる。2033 年頃には、従来の HPC におけるバッチ処理型のジョブは、継続的に更新されるデジタルツイン型のワークロードへと移行していくと見込まれる。その中で、Physics-Informed Neural Networks (PINNs) や Neural Operators は、従来の数値シミュレーションで用いられてきた反復型ソルバーを補完する役割を担うものである。さらに、ユーザーの利用形態も変化する。生成的設計やエージェント型ワークロードでは、ユーザーは設計の制約条件のみを指定すればよく、システムが自動的に多数の候補を生成し、それらを設計、シミュレーション、検証まで一体的に実行する。これにより、数百万規模の設計候補を数分以内に評価することが可能となる。このような統合的な計算環境の一例として、NVIDIA の Omniverse for Science がある。これは、物理システムのモデルと AI による「ワールドモデル」を同一のソフトウェア基盤上で扱うことができる統合環境を提供するものである。

2.4.1.7 日本の技術との融合

技術、研究、産業の各分野においてリーダーシップを有する日本は、将来のスーパーコンピューティングソリューションの方向性を定義していく上で不可欠なパートナーである。富岳 NEXT における理化学研究所、富士通、NVIDIA によるコデザインは、単一のシステム開発にとどまらず、今後のスーパーコンピューティング、ひいてはより広い産業へのアプローチを共同開発によって拡張するモデルを示すものである。

また、日本は先進パッケージング、次世代メモリ、材料技術の重要な拠点でもある。日本のガラスメーカー

は、チップレットベースの HPC-AI パッケージに最適化されたパネルレベルのガラスコアおよびガラスインターポーザの開発を推進している。日本のシリコンフォトリソグラフィエコシステムは、材料、デバイス設計、高精度製造の分野で強力な能力を有しており、CPO や高速インターコネクタの実現を支える基盤となっている。さらに、基板エコシステム (ABF、FC-BGA) および新たな HBM 製造投資は、日本を将来の HBM および高速 DRAM の重要な拠点として位置付けている。加えて、日本企業は将来のプラットフォーム要件に適合する高性能サーマルインターフェース材料 (TIM) や冷却ソリューションの開発を進めている。また、日本では新しい 3D 統合メモリ概念や異種積層技術に関する研究も積極的に進められており、これらの技術は今後 10 年の間に成熟していくことが期待されている。

さらに、CUDA ベースのライブラリ、性能ツール、ならびに気候、地球科学、製造分野向けのドメインフレームワークについても、NVIDIA と日本の学術機関および産業界が連携して開発を進めることにより、トレーニングプログラム、共通ベンチマーク、オープンソースの取り組みを通じたエコシステムの強化が可能となる。これにより、AI for Science および AI for Industry 分野における日本のリーダーシップを示す、より強固なエコシステムの構築が期待される。

また、モジュール化されたエネルギー効率の高い施設設計、再生可能エネルギーの統合、AI を活用した施設監視などの分野においても共同の取り組みを進めることで、日本は持続可能な HPC-AI インフラの分野においてリーダー的な役割を果たす可能性がある。

2.4.2 次年度の調査研究の計画

次年度においては、本年度整理した内容について、調査検討を継続する。

2.5 HPE における将来システムの動向調査

HPE はポスト・エクサスケール時代における顧客ニーズに応えるため、Cray システムアーキテクチャをさらに前進させる。

2.5.1 調査研究の概要

HPE は、Top500 リストにおける上位 3 つのエクサスケールシステムのアーキテクチャである Cray EX Supercomputing の後継製品となる HPE Cray Supercomputing GX(Cray GX) を発表した。2030-2032 年において、HPE は Cray GX によって示された方向性をさらに発展させることを目標としている。

Cray GX は AI と HPC が融合する時代におけるスーパーコンピューティングのブレークスルーを可能にするために特別に設計されたシステムである。Cray GX は Open Infrastructure の原則に基づいたシステムであり、OCP を中心としたコミュニティの重要な成果を取り込みながら、Multi-kW 級のシステムに到達する最新の HPC/AI プロセッサの電力・冷却課題に対応する。また、最新の Slingshot Ethernet Interconnect に加え、デバイスからノード、ラック、さらにはシステム全体へと広がる階層型システム管理によって、HPE ならではの価値を実現している。主要要素としては、OCP 高密度電力供給ソリューションと 100% 直接液体冷却を組み合わせることで、システムを構成するあらゆる要素が、きめ細かな電力および冷却制御のもとで、エネルギー効率が最も高い領域で動作するよう保証している。さらに、Cray GX はテナント間処理のセキュリティを確保するため、Multi-Tenancy の運用をサポートする。システム管理は重要なコンポーネントをリアルタイムで監視・管理し、最大稼働時間を確保する。

HPC と AI 駆動のワークロードおよびワークフローの組み合わせは、気候変動といった地球規模の課題やスマートシティを含む社会的目標に対処するために世界が必要とする科学的発見のペースを維持する上で不可欠となる。HPE はこれらの課題に対処するために、インテリジェントな HPC および AI システムの活用が極めて重要だと考えており、その影響を加速するために複数の技術に投資している。本章で取り上げる分野では、今後の協業により影響力を加速させることが可能である。HPE は社内で開発した技術と業界の主要テクノロジーパートナーとともに開発した技術の連携により、将来のシステムアーキテクチャのニーズに応えることに注力している。これには AMD、NVIDIA、Intel といった長年のパートナーに加え、Cerebras や NextSilicon のような新興企業、さらに量子分野の新たなパートナーが含まれる。HPE は Hewlett Packard Lab では大規模な量子プロセッサを最も生産的に活用する道を切り拓くことを目的として、古典コンピューティングと量子コンピューティングの組み合わせに焦点を当てた共同研究プログラムを展開している。HPE は将来のワークロードとワークフローのニーズに関する重要な洞察を提供してくれる顧客およびエンドユーザーと継続的に連携し、将来のシステムアーキテクチャの形成に役立てている。

HPE が注力する主要な技術は以下である。

- HPE Slingshot Ethernet High Performance Interconnect : 計算ノード間のインターコネクト基盤
- DAOS : 先進的な HPC および AI ワークフローが要求する高性能データアクセスを満たす
- システム管理 : 大規模システム導入における効率的な運用管理を実現
- ユーザーサービス : HPC や AI のユースケースに対し、最も生産的なワークフロー環境を提供
- 先進的なシステム製造および構築サービス : 大規模な最先端システムの提供を可能とし、迅速に本稼働への移行を保証

HPE は Juniper Networks を買収し、ネットワーク主導のイノベーションを顧客課題の解決に結びつける企業戦略を進めている。ネットワーク主導のビジネス・フォーカスの一環として、Slingshot Ethernet を大幅に進化させ、先進的なプロセッサノード同士を結合するバランスの取れたシステムアーキテクチャを構築するための能力を提供すべく、重要な投資を行っている。これには、プロセッサとチップレットレベルでの統合の可能性や、シリコンフォトニクスによるノード間通信の活用も含まれている。また、Slingshot テクノロジーの恩恵を受けうる日本のプロセッサ開発者との協業に前向きである。HPE はエクサスケールシステムにおける Slingshot の成功から得られた知見を活かし、Ultra Ethernet Consortium(UEC, <https://ultraethernet.org/>) の一員として High Performance Ethernet に関する共通のオープンスタンダードを確立することを目的として活動している。UEC は、プロトコル、物理層、リンク層、トランスポート、ソフトウェア、ストレージコンプライアンスといった個々の重点分野ごとにワーキンググループに分かれている。各分野の詳細は UEC のウェブサイトで開催されており、HPE は、理化学研究所および理化学研究所と協力するその他のテクノロジーパートナーが UEC への参加を検討することを歓迎している。UEC の取り組みは HPC および AI ワークロードの要件に主眼を置いており、理化学研究所が想定するターゲットアプリケーション要件とも高い整合性を持つと見込まれる。この活動を通じて、次世代計算機基盤プログラムにおけるテクノロジー採用のリスクが低減され、計算ノードを構成するネットワークアダプタとスイッチング基盤間のネットワーク分離において、日本の技術プロバイダとの相互運用可能なコンポーネントおよび技術共有の機会が生まれる。さらに、1,600Gbps NIC 世代における Ultra Ethernet Transport(UET) の完全実装に向けて急速に前進している。米国のエクサスケールシステムの性能を下支えしてきた Slingshot の Connectionless RDMA 機能が、HPE の実装基盤となる。2030-2032 年において、コネクション型プロトコルに注力する他ベンダーに対し、HPE は明確なリードを確保することになる。

将来のワークロードの新たなデータ要件に対応する取り組みの一環として、HPE は DAOS(Distributed Architecture for Object Storage) オープンソース・ソフトウェアプロジェクト (<https://daos.io/>) に投資し、このプロジェクトに貢献する開発者チームを拡大している。HPE は DAOS が今後の HPC、AI 連携型 HPC、および AI アプリケーションとワークフローに必要なストレージ性能を提供する鍵になると考えている。DAOS の性能ポテンシャルは IO500 (<https://io500.org/>) の最新の結果にも表れており、これは潜在的な将来のソフトウェアおよびハードウェアの探索や実装が含まれる。HPE は昨年 11 月に DAOS をベースとした初めての製品である K3000 を発表し、この基盤の上に今後 10 年で更なる強化を行う。HPE は既存の DAOS 製品 (例えば Argonne 国立研究所の Aurora システム - IO500 ランキング 1 位) を引き続きサポートし、さらに Oak Ridge 国立研究所の Discover システムに K3000 ストレージシステムに採用されるという新たな支持を得ている。HPE としては、次世代コンピューティング基盤のニーズに応えるべく、次世代 DAOS に求められる要件について議論を進めていきたい。なお、DAOS は AI・データ解析向けの低レイテンシ IO を担うオブジェクトベースの高速データレイヤとして機能し、HPE Cray Supercomputing Storage Systems(通称 ClusterStor) は大容量データを管理する Lustre ベースの並列ファイルストレージとして機能する。HPE が想定する将来のストレージアーキテクチャにおいて、DAOS と ClusterStor は、相補的に共存する関係と位置付けられている。

AI はいかに生産性を高め、科学の経済的インパクトを加速させるかをコミュニティに再考させている。汎用モデルは研究者の開発体験を変革し、問題と要件の精緻化に集中しながら共同でコードを生成することを可能にし、特化型 AI モデルはシミュレーションのアンサンブルを操舵し、科学ワークフローのスループットを高めるために使用されている。AI モデルの学習と推論にシミュレーションを密接に結合させる高度なワークフローが、モデリングとシミュレーションのあり方を再定義する。最後に、AI エージェントを含む AI 連携型 HPC ワークフローに実験施設やラボ設備を組み合わせたワークフローは、自動化と科学的発見における新たなパラダイムになり得る。HPE のソフトウェアエコシステムは急激に進化するこの状況を支援し、ユーザーワークフローからシステム管理にまたがるユースケースをサポートするために拡張を続ける。

HPE はオープンソースコミュニティへのコミットメントを通じ、DragonHPC(<https://github.com/DragonHPC/dragon>) や SmartSim(<https://github.com/CrayLabs/SmartSim>) のような技術を開発してきた。これらは科学者やエンジニアがリーダーシップクラスのスケーリングとパフォーマンスを保ちつつ、HPC、AI、およびデータ処理の洗練されたワークフローを容易に開発可能にする。HPE はより広範囲なコミュニティと連携し、HPE のエコシステムを活用して RHAPSODY(<https://arxiv.org/pdf/2512.20795>) などのツールを拡張し、AI の取り組みで恩恵を得られる多様な科学ユースケースを支援している。

AI 連携型・エージェント型ワークフローのために、HPC システムは膨大な電力を必要とする。次世代情報基盤において許容可能な電力予算の範囲内で性能目標に到達するために、Holistic Power Management(全体最適のための電力管理) は必要不可欠である。Cray GX で導入される手法を基盤として、HPE の Holistic Power Management 研究をソフトウェアエコシステムに組み込むことで、全システムレベルでの電力効率向上を図る。電力管理能力の強化に向け、HPE は次世代情報基盤のプロセッサ提供ベンダーと協力し、この取り組みを選定される新コプロセッサと密接に結合することが可能である。

AI Factory のコンセプトはユーザー体験を再定義し、従来は分離していたクラウドネイティブなデータ・機械学習と HPC の橋渡しとなる可能性がある。AI エージェントやエージェント型ワークロードを可能にする環境は、AI Factory において HPC 環境と密接に統合される必要がある。現在は Kubernetes ベースのワークフローを HPC ワークロードと統合する機能 (多くの場合は Slurm などの従来のワークロードマネージャー内で実現される) が必要とされており、新たなオーケストレーション手法や環境カプセル化の形へと進化する

想定される。この進化がどのような形をとろうとも、HPE のソフトウェアエコシステムはそれに歩調を合わせ、将来のハードウェア・プラットフォーム上でこれらの Capability や体験を実現できるよう発展していく。

HPE は従来の HPC システム管理環境と、HPE Performance Cluster Manager(HPCM) や HPE Cray System Management といった現行の HPE 製品のクラウドライクな運用環境の組み合わせを基に、さらに発展させることを目指す。HPE は新たなオープンソース・プロジェクトの Open CHAMI(<https://openchami.org/>) の創設メンバーであり、日本のテクノロジーパートナーとの協業に向けても前向きである。また、HPE はスーパーコンピューターシステムの計画と構築の双方を支援するデジタルツインの取り組み (例：<https://exadigit.github.io/>) をさらに発展させようとしている。HPE はエージェント型 AI が複雑なシステムの運用管理の自動化を提供し、リソース利用の最大化に寄与できると考える。

HPE は Slingshot と Processing SOC のシステム・シミュレーションを実施しており、期待されるノードおよびインターコネクティブ性能の目標が達成可能であるかの調査を共同で行うことが可能である。HPE は Sandia 国立研究所の Open-Source System Simulation Toolkit(SST) シミュレータ (<https://www.sandia.gov/ccr/project/structural-simulation-toolkit-sst>) の活用を推奨する。HPE は既に Slingshot インターコネクティブで接続される先進的なプロセッサノードに基づく設計の現行のシミュレーションに SST を使用している。HPE は SST 上で動作する自社デバイスの機能的かつ準サイクル制度のモデルを保有している。これらを用いて、顧客のトラフィック・パターンでネットワーク性能を評価することができる。このため、このシミュレーション環境を用いて次世代計算機基盤向けに選定されるベンチマークの主要要素を評価することが可能である。HPE は UEC Industry Consortium を通じて、業界のパートナーや大規模顧客 (クラウドサービスプロバイダなど) と連携し、リーダーシップクラスのスーパーコンピューティングおよび AI システムのニーズに応える High Performance Ethernet を中心としたエコシステムの構築に取り組んでいる。目標は世界発のエクサスケールシステム (米国 ORNL の Frontier) で示された Slingshot Ethernet の成功を基盤に、標準的な Ethernet の機能を提供することである。

HPE は 2.5.1 に示すような潜在的パートナーとの協業の可能性を検討しており、Slingshot エコシステムの低レベルなシリコン IP からチップ、シャーシ、ラック機能に至るまでの様々なレベルにおいて、協業に対して前向きである。HPE は次世代計算機基盤の中核となるプロセッシングテクノロジー (SOC) の提案に取り組む潜在的パートナーと密接に連携し、Slingshot Ethernet をノードプロセッサと緊密に連携する最善の方法を検討可能である。

2019 年以来、HPE は一貫して「Pervasive Security」の採用を積極的に提唱してきた。Pervasive Security にはハードウェア、ソフトウェア、データのサプライチェーン、ASIC の Root of Trust、およびそれらに関連する測定、また、人と機械エンティティがシステムとやり取りをする際の高度な認証・認可機能を結集する総合的アプローチが必要である。

最新の脅威モデルでは、以下の 3 つの主要領域に対処することが求められる：

- Hardware Root of Trust 機能：BMC から実行環境に至るまで、コンピューティング機能に依らない低レベルの Root of Trust の確立をサポートする。これには Secure Boot、および Measured State Model で動作するシステムの効率的な運用を可能にする管理ソフトウェアも含まれる。
- Software and Data Supply Chain Fidelity：システムにインストールされるソフトウェアが意図通りであることを確実にするため、Comprehensive Cryptographic Provenance を提供するよう強化する必要がある
- Identity Management and Authorization：リーダーシップクラスのシステムを効率よく活用する

Slingshot Collaboration Opportunities

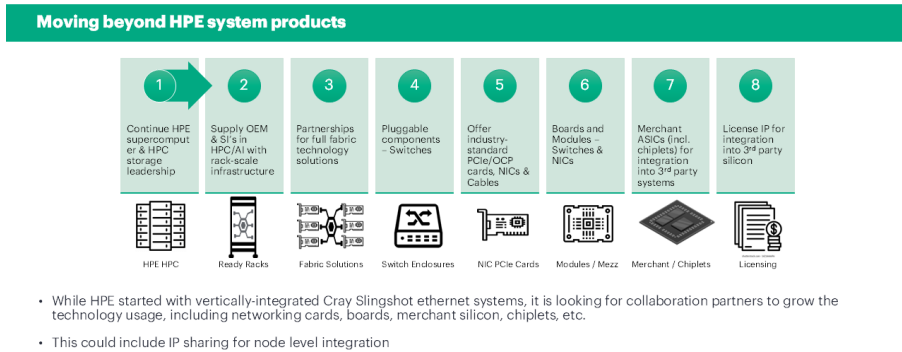


図 2.5.1 HPE の Slingshot Ethernet の協業可能性

には、広く受け入れられた Identity Management と Authorization のオープン標準が必要であり、Attribute-Based Control(ABAC) 機能が求められる

ここ数年で、HPC/AI 業界は科学的発見を加速するテクノロジーの最速の進歩の恩恵を受けている。2030-2032 年にどのようなテクノロジーが提供されるかを正確に予測するのは難しいものの、HPE は業界で最も効率的な計算機基盤を構築するための独自の能力をもたらし、賛同者とのパートナーシップを結ぶことにコミットしている。これはハードウェア・ソフトウェア技術レベルでのコラボレーション、オープン業界団体を通じた取り組み、およびオープンソースソフトウェアのイニシアティブが含まれる。

2.6 富士通における将来システムの動向調査

富士通は「富岳 NEXT」時代に向けた将来世代プロセッサとして MONAKA-X NPU および MONAKA-XX と呼ぶ 2 つのプロセッサ開発を検討している [3]。MONAKA-X NPU は MONAKA-X Core Die と AI 性能を強化する Accelerator Die(s) を Package(Socket) で統合し、70-200B 規模の小中規模 AI 推論を低消費電力、高スループットで実行できるプロセッサである。Accelerator Die(s) には富士通独自の HW 技術 Reconfigurable(再構成可能エンジン)Engine と量子化処理加速機構を搭載し、行列積計算や活性化関数を高速化する。MONAKA-XX は MONAKA-X よりも進んだプロセスノードを採用する。アーキテクチャは AI エージェントや量子連携に向けて最適化される予定であり、CPO(Co-Packaged Optics) の搭載も検討されている。本調査研究では MONAKA-XX に向けたアーキテクチャ技術と、MONAKA-X NPU に向けたシステムソフトウェア技術の動向を調査した。

2.6.1 調査研究の概要

まずアーキテクチャの調査研究では、2030 年以降を見据えた先端半導体製造技術、チップレットおよび三次元実装技術、メモリ技術、高帯域通信技術（電気および光インターコネクト）、ならびに CPU と加速器から構成されるノード・システムアーキテクチャの動向を体系的に整理した。これらの技術潮流を俯瞰し、「富岳

「NEXT」時代に求められる HPC + AI 統合基盤の構成要素を明確化するとともに、将来世代プロセッサへの展開を見据えた設計方向について整理した。

2030 年以降の計算基盤においては、トランジスタ微細化の進展と並行して、パッケージング、メモリ階層、高帯域通信を含むシステム統合設計が性能向上の主要因になると予測されている。国際半導体技術ロードマップ (IRDS) では、2nm 世代以降において配線遅延、電源品質、消費電力密度およびデータ移動コストが主要な制約となることが示されている [4, 5, 6]。特に IRDS 2024 では、システムレベルでの電力効率およびパッケージ統合技術の重要性が強調されている。このため、微細化の継続に加え、設計・実装技術の高度化が不可欠である。

デバイス構造の将来動向としては、Å 世代に向けたトランジスタ構造の高度化が検討されている。Complementary FET (CFET) は nFET と pFET の垂直積層によりセル面積効率を向上させる構造として報告されている [7, 8]。CFET は従来の GAA から大幅に構造が変化するため、Å 世代でも MONAKA-XX での利用が予想される A10 や A7 には間に合わない可能性がある。また、裏面電源供給構造 (Backside Power Delivery Network, BSPDN) は、電源配線と信号配線の分離により IR ドロップ低減および配線輻輳緩和を可能とする技術として報告されている [9]。BSPDN は特定のプロセス世代から全面的に導入されるのではなく、従来の電源供給構造のプロセスと並列に提供されると予想される。BSPDN は製造技術の複雑さから低い歩留りが予想されるので、コストに対する利益を勘案して利用する必要がある。

パッケージングおよび三次元実装技術では、チップレット設計とヘテロジニアス統合が主流化している。ハイブリッドボンディングを用いた高密度三次元積層は、ロジックとキャッシュあるいはメモリの近接配置を可能とし、レイテンシ削減および帯域向上に資する技術として報告されている。チップレット間接続の標準化としては UCIe 仕様が策定されている [10]。

メモリ技術に関しては、HBM を中心とする三次元積層型 DRAM が高帯域化を牽引している。AI 学習および HPC 数値計算ではメモリ帯域が性能律速要因となる傾向が強く、オンパッケージ高帯域メモリおよび三次元キャッシュ層の統合が重要である。

高帯域通信技術については、ノード内では高密度電気インターコネクットの高度化が継続する一方、ノード間およびラック間通信ではシリコンフォトニクスを活用した光インターコネクットの研究開発が進展している。

以上の技術動向を踏まえ、「富岳 NEXT」時代に向けた将来世代プロセッサ (MONAKA-XX) については、先端ロジック技術の活用に加え、三次元実装およびチップレットを前提としたモジュラー構成、三次元キャッシュ統合やオンパッケージ高帯域メモリの活用を含むメモリ階層最適化を主要な設計要素として位置づけている。また、光インターコネクット技術の段階的導入を視野に入れた I/O 構成についても検討を進めている。これらの取り組みにより、HPC および AI を統合的に処理可能な異種統合型アーキテクチャの構築を志向している。

次にシステムソフトウェアの調査研究では、MONAKA-X NPU を対象に最新動向を調査した。近年のシステムソフトウェアは、OS やファイルシステム、ツールチェーンから成る共通基盤と、その上で動作する HPC や AI、データセンターといった各領域向けのソフトウェア群で構成される。この構成は、CPU と NPU のような異種の計算ユニットを混載したシステム上で、HPC、AI、データセンターの各ワークロードを統合的に実行するための基盤を形成する。

共通基盤として、OS には RHEL や Ubuntu などの Linux ディストリビューションが、オブジェクトストレージや分散ファイルシステムには Ceph や Lustre が、ツールチェーンには GCC や LLVM などの標準コンパイラや、OpenBLAS などの数値計算ライブラリが広く利用され、システムの安定稼働と効率的な開発環境

	AI	Data Center	HPC
FW・ライブラリ	AI FW/Lib : ONNX Runtime, OpenVINO(OpenTensorRT), llama.cpp, TensorFlow, PyTorch	Data Analysis : Apache Spark, Trino, Apache Kafka, Apache Hadoop	並列実行環境 : OpenMPI, UCX, FCCL , UCC with xCCL
	ツールチェーン: GCC, LLVM, LLVM(SYCL) , CUDA Toolkit, OneDNN, fjDNN	UXL(OneDNN 以外), Arm Performance Libraries, Arm Compute Library, Arm Optimized Routines	OneTBB, OpenBLAS, FFTW3, OpenMP, NVIDIA HPC SDK
サービス・アプリ開発	推論サービング : Ollama, vLLM, Triton	CCA : CoCo, Trustee, Veraison	性能分析ツール : perf, Linaro Forge
	RAG : Milvus, TEI	監視 : prometheus, Grafana/Loki	
	ML/LLMOps: MinIO, LangChain, MLflow	サービス開発 : pypiserver, Quetz, PostgreSQL, Jupyter Notebook	ジョブ/システム運用管理 : Slurm, Altair PBS Professional, Singularity, Foreman,
運用管理	コンテナ運用基盤 : Kubernetes, Harbor, Kubernetes ISV distro.	仮想化基盤 : OpenStack, KubeVirt, QEMU/KVM, OpenStack ISV distro.	Ansible, ISV cluster mng. software
OS・ファイルシステム	オブジェクトストレージ : Ceph, Ceph ISV distro.		
	分散ファイルシステム : HDFS, Lustre, DDN EXAScaler, IBM Storage Scale		
	ディストリビューション : RHEL, SLES, Ubuntu, Rocky Linux		

注：2025年12月時点のソフトウェアであり今後変更される可能性あり。太字はNPU向けソフトウェア。

を提供する。

HPC 領域では、大規模な計算資源を効率的に管理するため、Slurm などのジョブスケジューラ、Foreman や Ansible などのシステム管理ソフトが用いられる。また、アプリケーションの並列化には Open MPI といった並列実行環境が中心的な役割を担い、OpenBLAS や FFTW3 のような高度な数値計算ライブラリも用いられる。

AI 領域では、ONNX Runtime のようなフレームワークが、GPU や CPU の性能を最大限に引き出すため、CUDA や oneDNN といったライブラリを呼び出し推論処理が実行される。また、Ollama に代表される推論サービングや RAG の構成要素である Milvus などのベクターデータベース、LangChain や MLflow などの ML/LLMOps を支援するソフトウェア群の活用が進んでおり、それらの管理のために Kubernetes などのコンテナ運用基盤の導入が進んでいる。

データセンター領域では、OpenStack などの仮想化基盤でサーバリソースの効率利用を図るとともに、大規模データを効率的に処理・分析するための Apache Spark などのデータ分析ソフトウェア、データ保護を目的として CoCo に代表される Confidential Computing ソフトウェアの導入が進み、多様なワークロードを支えている。

以上の動向を踏まえ、将来世代プロセッサ MONAKA-X NPU のシステムソフトウェアとしては、CPU と NPU を協調動作させて AI 性能を最大限に引き出すソフトウェアの検討を進めている。その一つとして NPU

に最適化された富士通独自ライブラリ fjDNN(仮称) を検討している。ONNX Runtime などの上位フレームワークにより CPU と NPU の違いを吸収し、NPU での処理に適したワークロードを fjDNN に振り分けることで、Ollama のような推論サービングツールがハードウェアを意識することなく高速な LLM 推論を実行できるシステムソフトウェアの構築を志向している。

また、富士通は 2030 年に 10,000 超の量子ビットの実現を目標に量子コンピュータの開発を進めると同時に、MONAKA シリーズを搭載する HPC との連携にも取り組んでいる。そして、これら量子・HPC のハイブリッド計算基盤を、AI プラットフォーム「Kozuchi」を通じて統合し、金融や材料など様々なアプリケーションで利用可能にすることを目指している。

この実現に向けて、量子コンピュータを管理するための基盤ソフト OQTOPUS(Open Quantum Toolchain for OPERATORS and USERS) や、量子アルゴリズムやアプリケーションを研究開発するためのパッケージソフト QARP(Fujitsu Quantum Application Research Package)、計算の種類や特性に応じて、CPU、GPU、量子コンピュータの中から最適な計算資源を自動選択するミドルウェア技術 CWB(Computing Workload Broker) などの研究開発に取り組んでいる。

2.6.2 次年度の調査研究の計画

次年度においては、本年度整理した 2030 年以降のアーキテクチャ技術動向について、技術成熟度、適用時期、システム影響の観点から調査を継続する。CPU と加速器からなるノード・システムアーキテクチャを支えるシステムソフトウェアについては、開発の方向性について調査を継続する。

2.7 NEC における将来システムの動向調査

NEC は 1983 年の SX-2 以来、40 年以上にわたりベクトル型スーパーコンピュータの開発・販売を行ってきた。最新機種 SX-Aurora TSUBASA は HPCI (High Performance Computing Infrastructure) として日本国内では海洋研究開発機構、東北大学、大阪大学などの大型計算機センターで現在採用されており、海外ではドイツ気象庁やチェコ気象庁などの気象予報システム、フランス CRIANN 大学の大型計算機センターで現在採用されているなど、日本国内に留まらずグローバル市場においても HPCI としての販売・運用実績がある。

SX ベクトルアーキテクチャはハードウェアの開発だけではなく、ハードウェア性能を引き出すためのコンパイラなどのソフトウェアについても長年の技術の蓄積があり、NEC は現在、ベクトルアーキテクチャ型の次世代機の検討を進めている。本節では、2029 年頃の製品リリースを目指して現在研究・開発を進めている Next Vector プロジェクトと、それに付随する技術的な検討状況について概要を述べる。

2.7.1 調査研究の概要

NEC は 2030 年以降のアーキテクチャについての調査研究として、高実効性能と高電力効率を両立するアーキテクチャの検討を行っている。プロセッサやアクセラレータのメモリ帯域と演算性能のバランスである Bytes/Flop (B/F) はメモリ帯域律速型のアプリケーションを高速に実行するうえで性能の決定要因の一つとなることが知られている。以前から、プロセッサやアクセラレータの B/F の低下のトレンドとそれに伴うメモリ帯域律速型アプリケーションの性能が十分に得られないことが HPC における大きな問題となってい

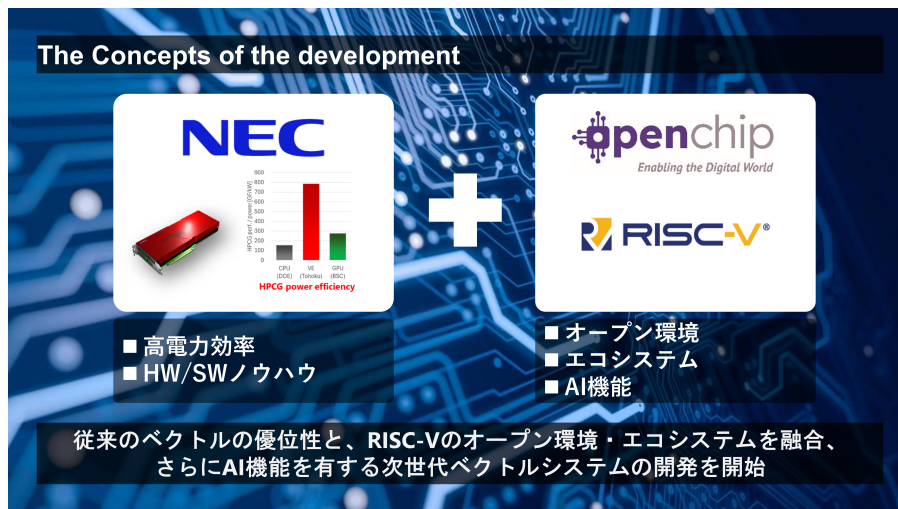


図 2.7.1 Next Vector Project の開発コンセプト

る。また、プロセスルールのスケーリングが鈍化し、今後も大幅な改善が見込めない中で、プロセッサの消費電力が爆発的に増加しており、HPC システムの消費電力が増大している問題がある。HPC による計算科学の進歩を今後も継続的に、かつ持続的に進めていくために、高実効性能と高電力効率を両立するプロセッサやアクセラレータの開発が求められている。

NEC は Next Vector プロジェクトへの参画を通して次世代のベクトルプロセッサの研究・開発を行っている。Next Vector プロジェクトはスペインの Openchip & Software Technologies 社が中心となって進めているプロセッサ開発プロジェクトであり、RISC-V アーキテクチャをベースに高性能スカラコア部、ベクトル演算部、及び AI 処理向けマトリックスエクステンション部を主とする構成を予定しており、2030 年頃の製品リリースを目指し開発を進めている。Next Vector プロジェクトにおいて NEC はベクトル演算部の開発協力を行っている。

Openchip と NEC は、HPC および AI 分野における次世代計算アーキテクチャの開発を目的として、2024 年より協業を開始した。Openchip は RISC - V ベースの高性能アクセラレータの開発を行う欧州企業であり、欧州の半導体戦略の中核を担う存在として位置づけられている。一方、NEC は長年にわたりベクトル型スーパーコンピュータの開発を通じて高性能計算技術を蓄積してきた実績がある。両社の協業は、AI 需要の急拡大やスーパーコンピュータの多様化に対応するため、オープンで柔軟性の高い RISC - V アーキテクチャと、高実行性能や高電力効率を特徴とするベクトル計算技術を組み合わせた新たな計算基盤の必要性を背景として始まったものである。これにより 2024 年に、Openchip、NEC、及び Barcelona Supercomputing Center (BSC) の三者が、RISC - V ベースのスーパーコンピュータ向けアクセラレータの共同研究が開始された [11]。

両社の協業は、AI 需要の急拡大やスーパーコンピュータの多様化に対応するため、図 2.7.1 に示す通りオープンで柔軟性の高い RISC - V アーキテクチャと、高実行性能や高電力効率を特徴とするベクトル計算技術を組み合わせた新たな計算基盤の必要性を背景として始まったものである。2024 年には、Openchip、NEC、Barcelona Supercomputing Center (BSC) の三者が、RISC - V ベースのスーパーコンピュータ向けアクセラレータの共同研究が開始された。

2025 年 11 月には、Openchip と NEC が次世代 Vector Processing Unit (VPU) の共同開発フェーズへ

移行したことが正式に発表された [12]。両社は Openchip の RISC - V ベースのハードウェアおよびソフトウェアスタックを活用しつつ、NEC のベクトルアーキテクチャを発展させた次世代ベクトルプロセッサの開発を進めている状況である。本協業で開発される技術は、RISC - V の持つ拡張性とオープン性を活かしながら、ベクトル計算の高い実効性能・高い電力効率を組み合わせる点に特徴がある。これにより、AI 推論や学習、科学技術計算など多様な用途に対応可能な柔軟かつ高性能なアクセラレータの実現が期待されている。また、Openchip は欧州の Next Generation EU の支援を受けており、今回の協業は日本および欧州のデジタル主権戦略の観点からも重要な位置づけを持つものと言える。

2.7.2 次年度の調査研究の計画

前年度に引き続き、NEC が HPCI システム開発において重視している実アプリケーションの実効性能、電力効率、使い勝手などの観点から調査・検討を行う。アーキテクチャグループの協力機関と議論・連携を行い、調査検討の分担、及び令和 9 年度までの調査研究範囲、計画について決定する。調査研究では特に NEC の長年のベクトルスーパーコンピュータの開発実績、及び国内外において HPC システムを提供してきたシステムベンダーとしての観点からの情報を整理する。さらに、将来の HPCI のシステムのあるべき姿、及びこれを実現するためのプロセッサ、メモリ、半導体技術、実装技術、冷却技術などのトレンド、及び課題について NEC 独自の視点からの調査研究を実施し、最終的に製造性やコスト、運用性なども鑑みた報告をまとめる。

参考文献

- [3] Vivek Mahajan. 量子コンピュータと HPC のハイブリッド計算基盤のビジョン. 2025.
- [4] International Roadmap for Devices and Systems (IRDS). *IRDS 2022 Edition*. Tech. rep. IEEE, 2022.
- [5] International Roadmap for Devices and Systems (IRDS). *IRDS 2023 Update*. Tech. rep. IEEE, 2023.
- [6] International Roadmap for Devices and Systems (IRDS). *IRDS 2024 Edition*. Tech. rep. IEEE, 2024.
- [7] J. Ryckaert et al. “The Complementary FET (CFET) for CMOS scaling beyond N3”. In: *IEEE Symposium on VLSI Technology*. 2018.
- [8] P. Weckx et al. “Complementary FET (CFET) for Sub-2nm Technology Nodes”. In: *Proc. IEEE International Electron Devices Meeting (IEDM)*. 2020.
- [9] B. Chava et al. “Backside Power Delivery as a Scaling Knob for Future Systems”. In: *Design-Process-Technology Co-optimization for Manufacturability XIII*. 2019.
- [10] UCIE Consortium. *Universal Chiplet Interconnect Express (UCIE) Specification 1.0*. Tech. rep. UCIE Consortium, 2022.
- [11] OpenChip, NEC, and BSC Explore Collaboration to Develop Next-Generation Supercomputers Based on RISC-V. <https://openchip.com/openchip-nec-and-bsc-explore-collaboration-to-develop-next-generation-supercomputers-based-on-risc-v/>. Accessed: 2026-03-17.

- [12] Openchip. *Openchip and NEC advance their collaboration on next-generation vector accelerator*. Press Release. Nov. 2025. URL: <https://openchip.com/openchip-and-nec-advance-their-collaboration-on-next-generation-vector-accelerator/>.

第3章

システムソフトウェア調査研究グループ

3.1 調査研究概要および方針

システムソフトウェア調査研究グループでは、次世代の HPCI 構成システムが備えるべきシステムソフトウェア基盤について、主に以下の2つの観点から技術的要件の整理と動向調査を行う。

第1に、「富岳 NEXT」をはじめとする次世代 HPCI を構成する多様なシステム群を一体的に運用するためのソフトウェア基盤の検討である。HPCI 構成システム間でソフトウェアやデータを共有するためには性能可搬性・保守性が欠かせない [13]。また、これまでの HPCI 構成システムで培われてきたソフトウェア資産をさらに継承・発展させるために、既存ソフトウェアスタックとの親和性を考慮するとともに、国際的な技術動向・コミュニティ動向との整合性も踏まえた検討を行う。

第2に、AI for Science をはじめとする新たなアプリケーション領域への対応である。近年、生成 AI の台頭やデータ駆動型研究の進展により、従来の数値シミュレーションに加えて AI、データ科学、およびシミュレーションを融合した研究手法が急速に広がりつつある [14]。こうした新しい利用形態を支えるためのプログラミング環境やライブラリ等のソフトウェア基盤についても調査・検討を行う。

本目的を達成するため、システムソフトウェア調査研究グループでは以下の4つのサブグループを設置し、プログラミング環境、通信ライブラリ、先端演算加速装置向けプログラミング環境、数値計算ライブラリ、ストレージ・ファイルシステムに関する調査研究を実施する。

- HPC プログラミング・通信環境サブグループ（サブグループリーダー：藤田 典久、分担機関：筑波大学）
- 先端演算加速装置プログラミング環境サブグループ（サブグループリーダー：宮島 敬明、分担機関：明治大学）
- 数値計算ライブラリサブグループ（サブグループリーダー：深谷 猛、分担機関：北海道大学）
- ストレージ・ファイルシステムサブグループ（サブグループリーダー：建部 修見、分担機関：筑波大学）

2025年度は、各サブグループの役割分担を整理するとともに、計算機アーキテクチャ・次世代アプリケーションとの相互作用に注目した Co-Design に向けて、他グループとの連携・実施体制を確立し、調査範囲とするシステムソフトウェアの明確化と調査戦略の策定を行った。また、FS2.0におけるシステムソフトウェア・ライブラリ調査研究グループの成果を確認・整理し、本調査研究において取り組むべき課題項目の抽出と具体的な調査項目の策定を行った。

3.2 HPC プログラミング・通信環境サブグループ

3.2.1 調査研究の目的

分担機関との役割分担及び次世代 HPC・AI 開発支援拠点形成事業との連携体制を確認・整備し、円滑な連携体制を確立する。また、将来動向の把握に向けた基礎調査として、現行の HPCI システムで利用されているプログラミング環境及び通信基盤について調査を実施し、求められるプログラミング・通信環境を明らかにする。

3.2.2 調査研究の結果

3.2.2.1 概要

分担機関との役割分担および、次世代 HPC・AI 開発支援拠点との連携体制を確認・整備し、円滑な連携体制を確立した。将来動向の把握に向けた基礎調査として、現行の HPCI システムで利用されているプログラミング環境について調査を実施した。本年度において具体的には、NVIDIA GPU と AMD GPU を対象として、ネイティブ開発環境である CUDA および HIP の機能比較と、指示文を用いてプログラミングを行う手法としてメジャーな OpenMP と OpenACC について機能比較や両 GPU における対応状況について調査を行った。

3.2.2.2 今年度の調査方針

HPCI における演算加速装置の利用状況を踏まえると、NVIDIA GPU は広く用いられている一方で、AMD GPU については運用及び利用の経験が相対的に少ない。このため、本年度は AMD GPU の開発環境を重点的な調査対象とした。調査にあたっては、QST Plasma Simulator 「双星」及び筑波大学で 3 月より運用開始したスーパーコンピュータ「Sirius (PACS12.0)」を利用し、実機に基づく検証を進めた。

本サブグループでは、従来の HPC プログラミングで用いられるプログラミング手法及び通信環境の調査を担当する。特定のハードウェアに限定されたプログラミング環境、AI に関する開発・通信環境、又は先端演算加速装置に関する内容については、他のサブグループが担当する、もしくは他のサブグループと協力して調査を進める体制とする。

想定する主たる利用場面としては、既に NVIDIA 環境で動作している CUDA 又は OpenACC ベースのプログラムを AMD 環境へ移植する場合を対象とした。これは、現実の HPC アプリケーションにおいて既存資産の継続活用が重要であり、移植可能性の把握が実務上の意義を持つためである。なお、本年度は Intel GPU を対象外とした。これは、HPCI 機関において現有リソースがなく、現時点で導入予定も確認されていないためである。また、調査対象のプログラミング環境としては、ネイティブ言語である CUDA (NVIDIA) 及び HIP (AMD)、並びに指示文ベースのプログラミング環境である OpenACC (NVIDIA) 及び OpenMP (AMD、NVIDIA) を対象とした。

3.2.2.3 MI300A の性能評価

前述のとおり、AMD GPU に関する知見は現時点で十分とはいえない。そこで、AMD MI300A を搭載したスーパーコンピュータ「Sirius (PACS 12.0)」を用いて、基礎性能評価を実施した。MI300A APU は、Zen 4 世代の CPU コア、CDNA 3 世代の GPU ユニット、HBM3 メモリを 1 つのパッケージに統合した製品で

表 3.2.1 MI300A の主要諸元

項目	性能
GPU FP32 Vector ピーク性能	122.6 TFLOPS
GPU FP32 Matrix ピーク性能	122.6 TFLOPS
GPU FP64 Vector ピーク性能	61.3 TFLOPS
GPU FP64 Matrix ピーク性能	122.6 TFLOPS
メモリ容量	128 GB
メモリピーク帯域	5.3 TB/s

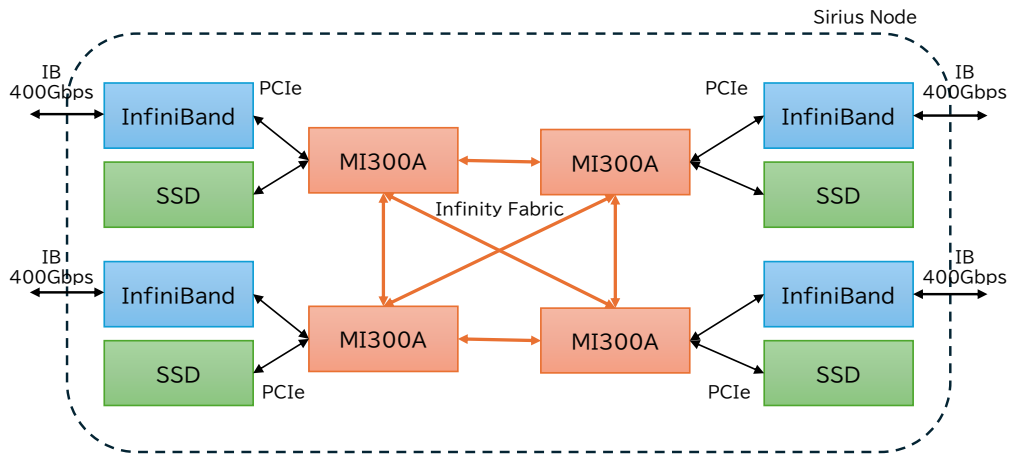


図 3.2.1 Sirius (PACS 12.0) のノード構成

あり、その諸元を表 3.2.1 に示す。

表 3.2.1 に示すように、GPU のピーク性能（行列積）は単精度で 122.6 TFLOPS、倍精度でも 122.6 TFLOPS であり、メモリ容量は 128 GB、ピーク帯域は 5.3 TB/s である。MI300A は CPU と GPU が同一メモリを共有するアーキテクチャであり、物理的にも論理的にも CPU-GPU 間でメモリが共有されている点の特徴である。また、メモリは HBM3 のみをサポートしており、DDR メモリを追加して大容量構成とすることはできないという制約がある。従来のアーキテクチャよりも総メモリ容量が少なくなるため、この点は、アプリケーションのメモリ使用特性を考えるうえで重要である。

Sirius のノード構成を図 3.2.1 に示す。Sirius では、AMD MI300A、InfiniBand NDR (400 Gbps)、3.84 TB PCIe Gen5 NVMe SSD がそれぞれ 4 台搭載されている。さらに、すべての APU が AMD Infinity Fabric で相互接続され、Non Uniform Memory Architecture (NUMA) を構成している。このため、計算機に含まれる 4 つの APU に属するすべての CPU と GPU が、1 つの共有メモリドメインを形成する構造となっている。

性能評価には、BLAS ライブラリを用いた単精度及び倍精度の行列積性能評価と、BabelStream を用いたメモリバンド幅測定を用いる。行列積では正方行列を用い、CPU 側には AMD Optimizing CPU Libraries (AOCL)、GPU 側には hipBLAS を利用した。これにより、MI300A における CPU 単独時と GPU 活用時

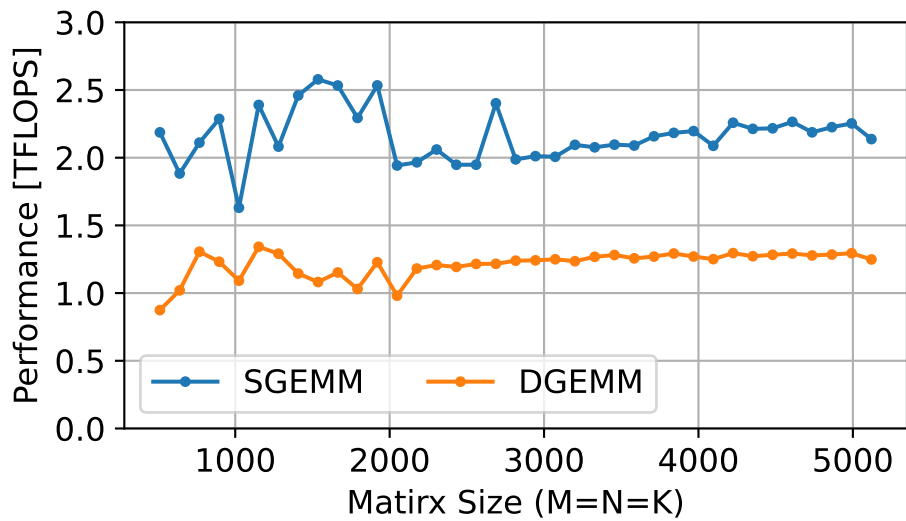


図 3.2.2 MI300A における CPU 側 GEMM 性能評価結果

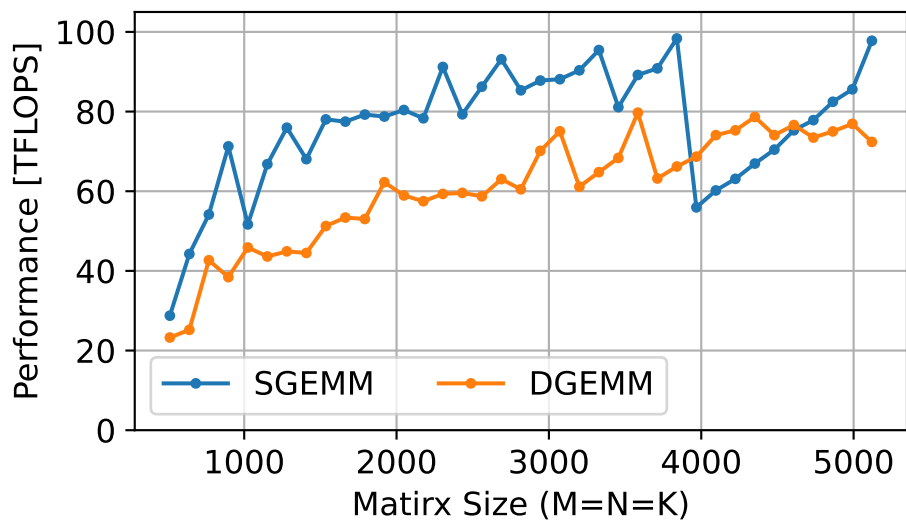


図 3.2.3 MI300A における GPU 側 GEMM 性能評価結果

の基本性能差を定量的に把握できる。

行列積性能の評価結果を図 3.2.2、図 3.2.3 及び表 3.2.2 に示す。GPU を用いた場合、SGEMM で 98.35 TFLOPS、DGEMM で 79.72 TFLOPS の性能が得られた。一方、CPU のみでは SGEMM で 2.58 TFLOPS、DGEMM で 1.34 TFLOPS にとどまり、演算性能を十分に引き出すには GPU 活用が不可欠であることが明らかとなった。なお、SGEMM では $N = 4000$ 前後で性能の落ち込みが見られるが、再現性があるため、実装上の問題である可能性が高いと考えられる。

メモリバンド幅の評価結果を図 3.2.4 及び表 3.2.3 に示す。GPU では Copy で 3780.9 GiB/s、Mul で

表 3.2.2 MI300A における GEMM 性能 [TFLOPS]

	単精度行列積 (SGEMM)	倍精度行列積 (DGEMM)
CPU	2.58 ($N = 1536$)	1.34 ($N = 1152$)
GPU	98.35 ($N = 3840$)	79.72 ($N = 3584$)

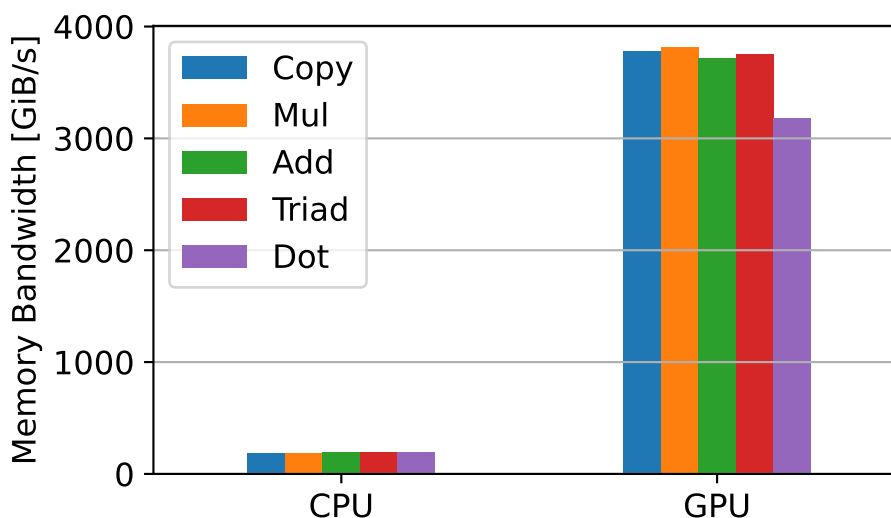


図 3.2.4 MI300A における BabelStream 性能評価結果

表 3.2.3 MI300A における BabelStream 測定結果 [GiB/s]

	Copy	Mul	Add	Triad	Dot
CPU	185.9	185.9	195.6	195.3	196.1
GPU	3780.9	3815.2	3718.1	3754.4	3177.5

3815.2 GiB/s、Add で 3718.1 GiB/s、Triad で 3754.4 GiB/s、Dot で 3177.5 GiB/s という高い性能を確認した。これに対し、CPU ではおおむね 185–196 GiB/s 程度にとどまっており、メモリバンド幅についても GPU を活用することが MI300A の性能を引き出す鍵であることが分かる。

以上より、MI300A は CPU 単独では十分な演算性能及びメモリバンド幅を發揮できないが、GPU を活用することで高い基本性能を示すことが確認できた。総じて、NVIDIA GPU と同等の基本性能が得られることを確認しており、AMD GPU 環境を HPC 用途で活用できる可能性を示す結果であるといえる。

3.2.2.4 ネイティブ開発環境に関する調査

NVIDIA GPU は CUDA、AMD GPU は HIP がネイティブ開発環境に該当する。どちらも、C++ 言語をベースに、GPU 向け拡張機能を追加した言語系である。CUDA と HIP は類似した文法を有しており、一方を学習していれば他方も比較的扱いやすいと考えられる。このため、既存の CUDA 資産を HIP へ移行する際

表 3.2.4 CUDA と HIP における主要ライブラリ対応

CUDA	HIP	用途
cuBLAS	hipBLAS	線形代数、行列積
cuFFT	hipFFT	高速フーリエ変換
cuSPARSE	hipSPARSE	疎行列演算
cuRAND	hipRAND	乱数生成
Thrust	rocThrust	GPU 向け C++ テンプレートライブラリ

表 3.2.5 GPU 間通信環境の対応

CUDA	HIP	用途
MPI	MPI	高性能計算向け汎用通信仕様 (CPU 間通信も含む)
NCCL	RCCL	GPU 間通信に最適化されたベンダ提供の通信ライブラリ

の学習障壁は必ずしも高くない。ただし、GPU アーキテクチャに強く依存した機能については互換性の確保が難しい。たとえば、NVIDIA GPU における warp 分割実行や非同期メモリ転送のような機能は、単純な置換では AMD GPU に対応しにくく、個別の検討が必要である。

また、CUDA と HIP は、100% 互換ではないものの、HPC で頻用されるライブラリについては双方に対応する実装が提供されている。対応例を表 3.2.4 に示す。これにより、線形代数、高速フーリエ変換、疎行列演算、乱数生成など、主要な機能については両環境で一定の移植性が確保されている。

GPU 間通信については、表 3.2.5 に示すように、Message Passing Interface (MPI) を用いればベンダー非依存に通信を記述できる。また、GPU 間通信に特化したベンダが提供する最適化通信ライブラリとして、NVIDIA GPU 向けには NVIDIA Collective Communications Library (NCCL)、AMD GPU 向けには ROCm Communication Collectives Library (RCCL) が提供されている。したがって、通信環境についても両ベンダーで一定の選択肢が用意されているといえる。

CUDA から HIP への自動変換ツールとしては、AMD から `hipify-perl` 及び `hipify-clang` が提供されている。`hipify-perl` は文字列置換を基本原理とするため、マクロが絡む変換や名前空間の扱いには限界があるものの、コンパイル可能な入力である必要がなく、ライブラリをインストールせずとも変換可能であり、さらにコードスニペットのような部分的な入力にも適用できる利点がある。一方、`hipify-clang` は LLVM Clang のフロントエンドを用いるため変換精度は高いが、入力ソースコードがコンパイル可能でなければならず、外部ライブラリを含めた環境整備や CUDA のインストールが必要となる。このため、入力コードの性質に応じて適切なツールを選択する必要がある。

図 3.2.5 及び図 3.2.6 に、`hipify-perl` を用いたコード変換の例に対応する形で、CUDA コードと HIP コードの対応関係を示す。いずれのツールを用いる場合でも、出力結果の検証は必須である。また、`hipify` ツールはコードを単純に変換するものであって、AMD GPU 向けの最適化まで自動で行うものではない点に留意する必要がある。

さらに、ボルツマン方程式における衝突項を求めるベンチマークコードに対して、実際に `hipify` ツールを適用した。このコードでは、計算の主たる部分が FFT による畳み込み計算であり、小さい 3 次元 Batched

図 3.2.5 CUDA コードの例 (hipify-perl 変換前)

```

cudaError_t err;

// allocate memory on a_dev with size bytes
err = cudaMalloc((void**)&a_dev, size);
if (err != cudaSuccess) {
    // handle error
}

// copy from host a to device a_dev
err = cudaMemcpy(a_dev, a, size, cudaMemcpyHostToDevice);
if (err != cudaSuccess) {
    // handle error
}

// launch kernel with N blocks and M threads
kernel<<<N, M>>>(...);

// check launch error
err = cudaGetLastError();
if (err != cudaSuccess) {
    // handle error
}

// synchronize
err = cudaDeviceSynchronize();
if (err != cudaSuccess) {
    // handle error
}

```

FFT を大量に実行する必要がある。本ベンチマークは、もともと CUDA で開発されたコードであるが、hipify-perl を用いることで AMD GPU 対応を実施した。本コードにおいては、ツールによる自動変換のみで AMD GPU 対応が可能であることがわかった。GPU カーネルコードに加えて CUDA API 呼び出し及び FFT ライブラリ呼び出しも変換される。一方で、ビルド方法、たとえば Makefile 等は hipify では変換できず、手動で対応する必要があった。

性能測定結果を図 3.2.7 に示す。ここでは、計算全体、FFT のみ、及び FFT 以外のカーネル部分の時間を比較し、MI300A と GH200 の性能差を評価した。GH200 については、最先端共同 HPC 基盤施設 (JCAHPC) が運用するスーパーコンピュータ「Miyabi-G」を利用した。その結果、全体としては GH200 の方が 7.5% 高速であったが、本計算では FFT の影響が大きく、MI300A では FFT 部分が遅いために全体性能でも不利となった。一方、FFT 以外のカーネル計算に限れば MI300A の方が高速であり、実アプリケーションに近いコードにおいても MI300A が高い性能を示す可能性が確認された。

この結果は、hipify ツールが実際のコードで有効に動作することを示すとともに、AMD GPU への移植可能性を実証するものである。ただし、比較は 1 つのベンチマークコードに限られており、一般化にはさらなる

図 3.2.6 変換された HIP コードの例 (hipify-perl 変換後)

```
hipError_t err;

// allocate memory on a_dev with size bytes
err = hipMalloc((void**)&a_dev, size);
if (err != hipSuccess) {
    // handle error
}

// copy from host a to device a_dev
err = hipMemcpy(a_dev, a, size, hipMemcpyHostToDevice);
if (err != hipSuccess) {
    // handle error
}

// launch kernel with N blocks and M threads
kernel<<<N, M>>>(...);

// check launch error
err = hipGetLastError();
if (err != hipSuccess) {
    // handle error
}

// synchronize
err = hipDeviceSynchronize();
if (err != hipSuccess) {
    // handle error
}
```

検証が必要である。考察として、本事例で用いた FFT サイズは 32^3 であり、AMD GPU におけるスレッドの制御単位である wavefront (NVIDIA GPU における warp に相当) のサイズが 64 であるため、AMD GPU が性能面で不利であると考えられる。実際、 64^3 の場合には FFT 性能差はほとんどなく、MI300A は約 2% 遅い程度であった。もっとも、実計算では 32^3 サイズの方が主要であることを、ベンチマーク開発者である宇宙物理分野の専門家から確認しており、この点は実運用上無視できない。

3.2.2.5 指示文を用いたプログラミング

指示文 (ディレクティブ / Directive) を用いた GPU プログラミング手法としては、OpenACC と OpenMP Target Offload がある。以降、CPU のみを対象とした従来の OpenMP と区別するため、GPU プログラミングに用いる OpenMP を OpenMP Target Offload と呼ぶ。表 3.2.6 に示すように、AMD 及び NVIDIA の GPU 向け開発環境は、指示文を用いたプログラミングをサポートしている。

OpenMP Target Offload は、NVIDIA GPU と AMD GPU の両方を対象にできる点が大きな利点である。そのため、ベンダーをまたいだ移植性を重視する場合には有力な選択肢である。ただし、言語仕様が共

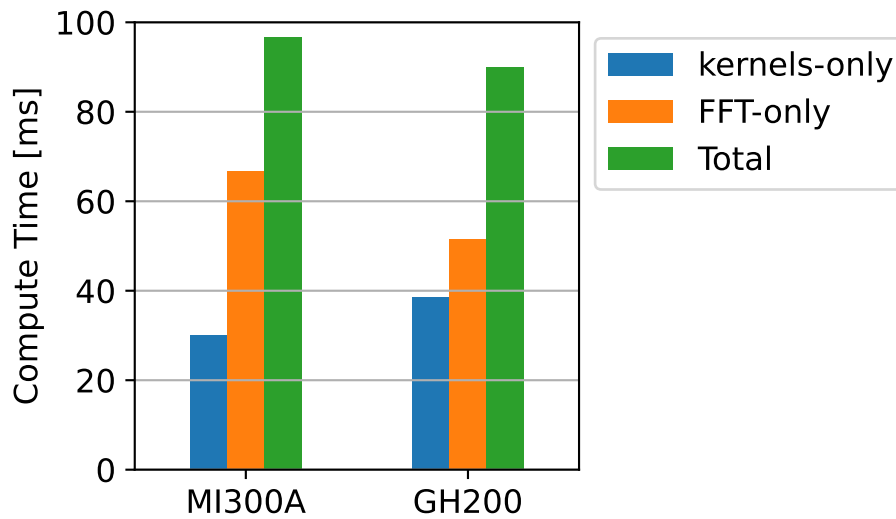


図 3.2.7 衝突項ベンチマークコードにおける MI300A と GH200 の性能比較

表 3.2.6 主要 GPU 開発環境における Directive 対応状況

実装	バージョン	OpenACC	OpenMP Target Offload
NVIDIA HPC SDK	26.1	OpenACC 2.7 の大部分に対応	OpenMP 5.0 のサブセット対応
ROCm	7.2.0	非対応	OpenMP 4.5 対応 5.0 / 5.1 / 5.2 はサブセット対応

通であっても性能特性が同じとは限らない。これは、CPU プログラミングにおいて異なるコンパイラごとに「速い書き方」が異なるのと同様であり、GPU でも詳細な性能調査が必要である。また、NVIDIA GPU では OpenACC の方が性能上有利となる傾向が見られる一方で、CPU 並列化との親和性は OpenMP Target Offload の方が高いと考えられる。

一方、OpenACC は対応コンパイラが NVIDIA HPC SDK に事実上限られており、現状では NVIDIA GPU 向けの仕様として理解するのが妥当である。AMD GPU のコンパイラ基盤は LLVM に基づいているため、将来的に Clang が OpenACC をサポートすれば利用可能となる可能性はあるが、直近での実装可能性は高くないと見込まれる。また、gcc でも利用可能ではあるものの、現時点では実用レベルには達していないと判断した。

3.2.2.6 次世代 HPC・AI 開発支援拠点連携

次世代 HPC・AI 開発支援拠点（HAIRDESC）との連携方法について確認した。サブグループリーダーである藤田が、HAIRDESC 中核機関の一つである筑波大学の運営委員を兼任しているため、組織間でダイレクトに連携可能であることを確認した。このことは、今後の調査研究において、アプリケーション側の要請や実運用上の課題を迅速に把握するうえで有効である。

3.2.3 次年度の調査研究の計画

3.2.3.1 開発環境に関する調査研究の計画

今年度は Fortran を調査対象外としたが、来年度はこれを重要な調査対象として扱う予定である。NVIDIA 側では、PGI 社のコンパイラに由来するコンパイラ (nvfortran) がよく知られており、利用実績も多く、比較的安定した開発基盤が整備されている。nvfortran コンパイラを用いることで、OpenACC 及び OpenMP を用いた GPU プログラミングが可能であり、既存の Fortran 資産を活用しやすいと見込まれる。

これに対して、AMD 環境における Fortran コンパイラは、LLVM Flang ベースであるため現時点では未知数の部分が多い。OpenMP を用いた GPU プログラミングは可能であるものの、まだ簡単な例でしか確認できておらず、安定性や完成度については十分に把握できていない。したがって、来年度は Fortran 資産を有する既存 HPC アプリケーションの将来展開を見据え、AMD 環境における実用可能性を重点的に検証する必要がある。

また、来年度は従来手法にとられない新たなアプローチについても探索する。具体的には、Kokkos、Julia、Mojo に関する調査を予定している。Kokkos は、ベンダニュートラルな GPU プログラミング環境として、主として米国において一定のシェアを有するが、C++ ライブラリであるため C++ の利用が避けられず、C++ に習熟していない開発者にとって導入障壁が高い。テンプレートの利用が必要なことや、テンプレートに関係する長大で複雑なエラーメッセージなど、実用上の障壁も多い。抽象度は CUDA や HIP と同程度であり、ベースとなる言語が C++ で同じことから、ベンダーニュートラルな CUDA のように位置付けることはできると考えられる。

C++ の利用障壁の問題を解決するために、Julia や Mojo のように、異なる言語として開発されている開発環境についても調査すべきであると考えている。もっとも、これらの言語の利用シェアは必ずしも明確ではなく、実利用の広がりや継続性を慎重に見極める必要がある。Julia については、一定の利用例が存在し、日本国内でも活用事例が見られることから、候補としての妥当性を有すると考えられる。

3.2.3.2 通信環境調査に関する計画

高性能 GPU クラスタは、システム全体の計算機を相互接続する Scale-out ネットワークと、限られた台数の GPU を直接広帯域に接続する Scale-up ネットワークの双方によって構成される。「富岳 NEXT」においても、2種類のネットワークを併用することが明らかにされており、今後の大規模 GPU システムではこのような階層的通信構成が重要になると考えられる。

このため、次年度は Scale-out ネットワークと Scale-up ネットワークが混在する通信環境に関する調査を実施する予定である。Scale-up ネットワークは接続台数が限られる代わりに高い通信性能を持つが、それだけでは利用可能な GPU 数に限界がある。そのため、大規模並列計算を実現するには、両者のネットワークを効率よく使い分け、あるいは協調的に利用する方法を明らかにすることが必要である。

3.2.3.3 次世代 HPC・AI 開発支援拠点連携に関する計画

次年度は、HAIRDESC との連携を通じて、アプリケーション側開発者のニーズ及び計算の性質に関する調査を進める予定である。具体的には、利用言語として C、C++、Fortran 等、通信方式として MPI や PGAS 系 (OpenSHMEM、GloblArrays など)、さらに BLAS、LAPACK、FFT、乱数等のライブラリ、及び NetCDF 等の I/O 環境について実態を把握する。これにより、実アプリケーションに即した開発環境及び

通信環境の要件を整理し、優先的に検討すべき技術課題を抽出する。

特に HAIRDESC では、GPU 化が未対応、もしくは十分ではないアプリケーションを扱うことが多いと見込まれる。このため、単に GPU で動作可能かどうかだけでなく、GPU 化の容易さや GPU 上で十分な性能を得やすいかどうかといった観点も重視して調査を進める必要がある。こうした検討を通じて、将来の HPCI システムにおいて実効性の高い開発支援の方向性を明らかにする予定である。

3.3 先端演算加速装置プログラミング環境サブグループ

3.3.1 調査研究の目的

AI 技術の急速な進歩に伴い、計算性能の需要と消費電力が急激に増加している。これらの問題を解決すべく、AI 向けに設計された先端的演算加速装置（AI アクセラレータ）が数多く登場している。近年の先端的演算加速装置は、密行列積や Transformer モデルを高速かつ低消費電力で実行することに注力しているため、そのアーキテクチャはそれらに特化したものになっている。これらの処理を行なう範囲では、既存の CPU や GPU よりも高い計算性能や電量性能比を示すものも多い。しかし、プログラミング環境は AI 向けプログラミング環境である PyTorch や ONNX が中心で、プログラミングパラダイムもこれまでの計算機とは異なる点も散見される。そのため、これらの先端演算加速装置の使用時に必要な考え方を明確化する必要がある。

このような背景を踏まえ、本サブグループでは、プログラミング環境および性能特性理解の面から、将来の技術動向の把握と先端演算加速装置の科学技術計算への適用可能性を検討する。プログラミング環境については、各装置の現状把握と今後の動向などを調査・比較する。また、性能特性の理解には具体的なプロキシアプリなどではなく、HPC 分野で頻出の計算カーネルを念頭に置く。具体的には、Barkley Dwarfs で示される、Dense Linear Algebra, Sparse Linear Algebra, FFT, N-Body, Structured Grids, Unstructured Grids, Monte Carlo, Graph である。

2025 年度の目的は、各種先端演算加速装置の動向調査、主たる調査対象の選定、調査対象の選定と各社との連携、関連サブグループとの役割分担の明確化である。

3.3.2 調査研究の結果

本節では、各種先端演算加速装置の動向調査、選定した主たる調査対象の先端演算加速装置、関連サブグループとの役割分担について述べる。

3.3.2.1 各種先端演算加速装置の動向調査

各種先端演算加速装置の広範な調査と動向について述べる。表 3.3.1 に報告書執筆時点（2026 年 02 月）で利用可能な各種先端演算加速装置とその特徴を列挙する。発表時期と利用可能となる時期が異なるため、各装置の発表時期は必ずしも同じではないが、執筆時点で利用可能であることに重点を置いた。なお、“-” は非搭載であることを示す。

メモリについては、これまでの計算機と比べて大幅な性能向上が見られる。ワークロードや環境に依存するが、推論のスループットはメモリバンド幅とメモリ容量に大きく影響されるため、これまでの計算機よりもバンド幅が高きく、容量も大きい。HBM 系のメモリデバイスを利用して、バンド幅が 10 TB/s 前後でメモリ容量が 100 GB 超を提供するものもあり、科学技術計算も恩恵があると考えられる。また、Groq や Cerebras の装置では SRAM をメインメモリとし、HBM 系よりも 2 桁以上大きなバンド幅を達成している。その反面、

メモリ容量は HBM 系よりも 1 ほど桁小さく、このアーキテクチャを採用する限りこの傾向は大きくは変わらないと考えられる。

計算性能については、低精度化が著しい。推論は単位時間あたりのトークン数など応答遅延が主たる指標であり、パラメータ数と数値精度を下げて計算量とメモリ転送量を減らすことで応答遅延を短くしている。この状況に対応するために、装置側は一定の回路面積の中で 32bit 以上の高精度な演算器を排除して、16bit 以下の低精度な演算器を搭載する方針を採用している。また、ほぼすべての装置がベクトル演算器だけでなく、行列（テンソル）演算を行なう演算器を搭載している。報告書執筆時点では、16bit 以下の行列演算の性能を高める傾向に進んでおり、8bit や 4bit の演算器も搭載されつつある。

計算機アーキテクチャについては、キャッシュや共有メモリの排除が見受けられる。深層学習モデルはデータフローで表現できるため、計算とデータ移動をグラフとして設計・最適化でき、データの再利用や共有の必要性が一般的な計算アルゴリズムと異なる。キャッシュや演算器間でのデータ共有の機能を搭載しないことで、演算器密度やメモリ性能を向上させる方針が採用されている。これらの機能は非常に複雑であり、回路面積や消費電力が単なる演算器よりも大きい。

プログラミング環境については PyTorch が中心であり、汎用プログラミング環境が提供されているのは、Cerebras Systems 社、Tenstorrent 社、Preferred Networks 社、PEZY Computing 社、NEXTSILICON 社の 5 つだけであることがわかった。どのベンダーの装置でも C/C++ と同等の水準の言語が利用できるが、計算機アーキテクチャが既存の CPU や GPU と異なるため、プログラミングモデルやパラダイムが大きく異なるものもある。

その他、ベンダーに栄枯盛衰はあるが創業から 10 年程度経っている例もあることや、計算機アーキテクチャを学術論文として公開している例は少なくないこともわかった。

これらの動向調査を踏まえて以下の選定要件を設定し、主たる調査対象を選定した。その結果、Cerebras Systems 社、Tenstorrent 社、Preferred Networks 社の装置を主たる調査対象とした。

1. 科学技術計算への適用可能性検討の高さ（計算機アーキテクチャ）：単精度浮動小数点演算器（IEEE 754-2008 binary32, FP32）をハードウェアマクロとして搭載していること
2. 科学技術計算への適用可能性検討の高さ（プログラミング環境）：C/C++ 言語などの汎用プログラミング環境が提供されていること
3. 技術的信用性、透明性の高さ：計算機アーキテクチャや計算性能を学術論文として公開していること

3.3.3 Cerebras Systems 社 Cerebras CS-2/3

CS-2 は WSE-2 をプロセッサとして搭載し、それを収容・冷却・電源供給するための専用システムである。x86 サーバに 100 Gbps イーサネットケーブル 12 本で接続され、ネットワークアタッチドアクセラレータとして動作する。ここでは WSE-2 のアーキテクチャについて説明し、その汎用コンピューティング向け開発環境である Cerebras Hardware SDK について述べる。

3.3.3.1 計算機アーキテクチャ：Wafer-Scale Engine

WSE-2 は、全体として 215mm × 215mm のウェハサイズのチップ上に、約 85 万個の計算コアを 2 次元メッシュで接続した構造となっている。各 PE は完全に非同期で動作し、メモリの一貫性も保証されないため、実質的には分散メモリシステムとして扱われる。動作周波数は 850 MHz で、科学技術計算を行う際の

ベンダー	カテゴリ	創業時期	アーキテクチャ分類	名称 (Core / System)
Intel	汎用計算機 専業		VLIW / TU ¹	Gaudi3
AMD	汎用計算機 専業		SIMT / TU / 共有メモリ	CDNA3 / Instinct MI300X
NVIDIA	汎用計算機 専業		SIMT / TU / 共有メモリ	Blackwell / B200
Groq (NVIDIA)	AI 計算機 専業	2016~2026	SIMD / dataflow	LPU / NVIDIA Groq 3 LPU
Microsoft	クラウド事業者		TU /	- / Maia 200
AWS	クラウド事業者		TU /	NeuronCore-v4 / Trainium3
Google	クラウド事業者		Systolic Array	- / TPU v7
Meta	クラウド事業者		MIMD / TU / 分散メモリ	MTIA 300
Cerebras	AI 計算機 専業	2015~	MIMD / Spatial / 分散メモリ	WSE-3 / CS-3
PFN	AI 統合	2014~	Wide-SIMD / 分散メモリ	MN-Core 2 / -
Tenstorrent	AI 計算機 専業	2016~	MIMD / TU / Spatial / 分散メモリ	Tensix Core / Blackhole
SambaNova	AI 計算機 専業	2017~	Dataflow / CGRA	RDU / SN40L
NEXTSILICON	高性能計算機 専業	2017~	Dataflow / Spatial	- / Maverick
PEZY Computing	高性能計算機 専業	2010~	MIMD	PEZY-SC4 / -
ベンダー	FP64 性能 [TFlops/s]	FP32 性能 [TFlops/s]	FP16/TF16 性能 [TFlops/s]	メモリバンド幅 [TB/s]
Intel	-	14	1,835	3.7
AMD	81.7	163	1,307	5.3
NVIDIA	40	80	5,000	8.0
Groq (NVIDIA)	-	-	???	150.0
Microsoft	-	???	1,268	7.0
AWS	-	183	661	4.9
Google	-	-	2,307	7.4
Meta	-	-	600	6.1
Cerebras	-	7,200	14,400	21,000.0
PFN	12	49	393	0.5
Tenstorrent	-	1.68	372	0.5
SambaNova	-	???	638	1.6
NEXTSILICON	???	???	???	6.4
PEZY Computing	24.6	49	589	3.2
ベンダー	メインメモリ容量	汎用プログラミング環境	対応分野	
Intel	128 GB HBM2e	-	推論	
AMD	192 GB HBM3	○	学習、推論、HPC	
NVIDIA	186 GB HBM3e	○	学習、推論、HPC	
Groq (NVIDIA)	0.5 GB SRAM	-	推論	
Microsoft	216 GB HBM3e	-	推論	
AWS	144 GB HBM3e	-	推論	
Google	192 GB HBM2e	-	推論	
Meta	216 GB HBM3e?	-	推論	
Cerebras	44 GB SRAM	○	推論、(HPC)	
PFN	16 GB GDDR6X	○	推論、HPC	
Tenstorrent	32 GB GDDR6	○	推論	
SambaNova	64 GB HBM3, 1.5TB DDR5	-	推論	
NEXTSILICON	192 GB HBM3e	-	高性能計算	
PEZY Computing	96 GB HBM3	○	ゲノム解析	

表 3.3.1 各種先端演算加速装置のカタログ値

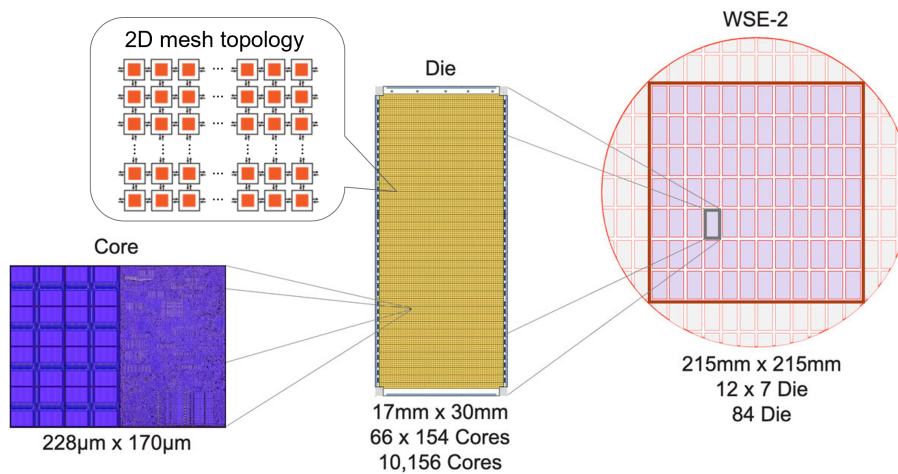


図 3.3.1 WSE-2 の物理的構造

CS-2 システム全体の最大消費電力は約 16.5kW である。図 3.3.1 に示す様に、チップ上の物理的な PE の総数は 853,104 (= 10,156/ダイ × 84 ダイ) であるが、ユーザーが使用できる総数は 745,500 (= x 方向：750 × y 方向：994) である。各ダイは切り出されず、レティクル・スティッチング (Reticle Stitching) と呼ばれる特殊な方法で物理的に接続されており、全体として 1 つの巨大なチップとして動作する。WSE-2 は物理的に約 40GB のメモリを搭載するが、ユーザーが使用できるのは 35.78 GB (= 48 KB × 745,500 PEs) である。なお、WSE-3 でユーザーが使用可能な値は、893,064 個 (x 方向：762 個、y 方向：1,176 個) であり、総メモリ容量は 42.9 GB、メモリバンド幅の理論値は 17.1 PB/s である。各 PE は完全に独立して動作し、メモリ一貫性は保証されないため、科学計算を行なう際には分散メモリシステムとして扱われる。このため、WSE-2/3 は分散メモリアーキテクチャとして利用可能で、通信制御を含めた明示的な分散並列プログラミングが必要となる。本調査研究で利用するアルゴンヌ国立研究所に設置された WSE-2/3 の動作周波数は、750 MHz もしくは 850 MHz である。

WSE-2 の PE アーキテクチャを図 3.3.2 に示す。WSE-3 の PE アーキテクチャの詳細は公開されていないため、WSE-2 の PE アーキテクチャを示す。PE は演算を行う Compute Element (CE) と、PE 間でデータを交換するルータで構成される。熱管理のためサーマルスロットリングが発生する可能性がある。サーマルスロットリングは動作周波数を下げるのではなく、NOP 命令を動的に挿入することで実現されている。そのため、実効性能はワークロードによって変動するが、パフォーマンスカウンターなどは提供されていないため、実際にどの程度の NOP 命令が挿入されたかはわからない。

PE の CE は、48 KB の PE 専用ローカルメモリと 7 段のインオーダー実行パイプラインで構成される。両者は 2 つの 64bit 読み出しポートと 1 つの書き込みポートで接続されている。ローカルメモリは 48 KB のスクラッチパッドメモリであり、キャッシュの機能は持たない。CE の内部は 4 つの主要機能ブロックで構成される。データパス部には 4 つの 128bit の SIMD 演算ユニットがあり、4-way 32bit、8-way 16bit または 16-way 8bit SIMD 演算ユニットとして動作するが、全ての演算がこの SIMD 幅で実行されるわけではない。CE には行列演算ユニットは存在しないが、単一命令で配列要素に対してアドレス計算を省略して最大 4 次元のテンソルを効率的に処理できる。また、Fused Multiply-Add (FMA) 演算が行えるが、後述の DSD を利用して明示的に SIMD 演算ユニットを利用する必要がある。

ローカルメモリは 48 KB のスクラッチパッドメモリとソフトウェア管理キャッシュで構成される。CE か

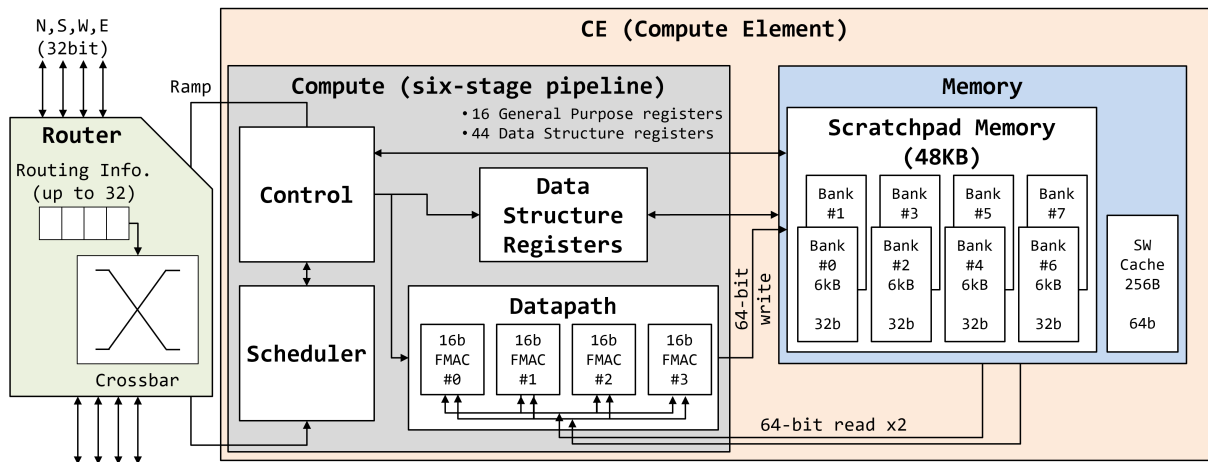


図 3.3.2 WSE-2 の PE の内部構造

らのスクラッチパッドメモリへのアクセスは 1 サイクルで行なわれる。ただしキャッシュ機能は備えておらず、ユーザーが明示的に全てのデータ転送を管理する必要がある。これらのメモリは各 PE ごとに独立しているため、一般的なコンピュータのような共有メモリは存在せず、メモリー貫性は保証されない。スクラッチパッドメモリは 8 つの 32bit シングルポートバンクで構成される。他のコンピュータと同様、メモリ性能を最大化するにはバンク競合を回避する必要がある。バンク競合が発生しなければ、1 サイクルで 2 つの 64bit 読み取りと 1 つの 64bit 書き込みを並列実行できる。ソフトウェア管理キャッシュは 256byte で、頻繁にアクセスされるデータを格納するために使用される。

ルータは、東西南北の 4 方向および自 PE を含めた 5 方向の通信が可能で、32bit 単位のデータを 1 サイクルで送受信できる。また、静的ルーティング方式を採用しており、通信経路はプログラム時に明示的に指定する必要がある。

3.3.3.2 汎用プログラミング環境：Cerebras Hardware SDK

Cerebras Hardware SDK は、Cerebras CS-3 システム上で汎用計算を行うためのソフトウェア開発キットである。これは NVIDIA GPU 向けの NVIDIA HPC SDK や AMD GPU 向けの HIP に相当する、低レイヤー向けのプログラミング環境である。図 3.3.3 と図 3.3.4 に、それぞれ、Cerebras Hardware SDK を用いた開発フローと CSL 言語の例を示す。

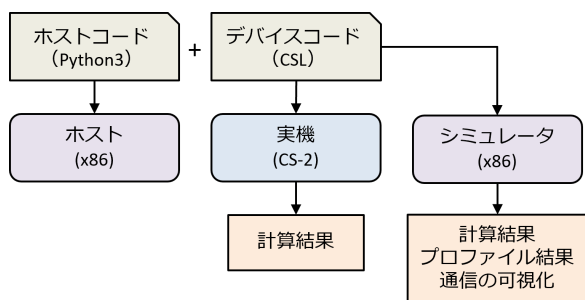


図 3.3.3 開発フロー：検証用シミュレータが存在する

```
fn gemv() void {
  for (@range(u16, N)) |i| {
    @fmacs(y_dsd, y_dsd, A_dsd, x[i]);
    A_dsd = @increment_dsd_offset(A_dsd, 1, f32);
  }
  @fadds(y_dsd, y_dsd, b_dsd);
}
```

図 3.3.4 CSL 言語の DSD を用いた行列ベクトル積の例

SDK には CSL プログラミング言語、数値計算ライブラリ、シミュレータ、サンプルプログラムが含まれる。SDK 内では、CS-3 はデバイス、ホスト CPU クラスはホスト、両者を接続するイーサネット接続はホスト I/O と呼ばれる。デバイス側コードは専用言語 CSL で記述され、WSE-3 上で実行されるタスクと通信を記述する。ホスト側のコードは Python で記述され、イーサネットを用いたホストとデバイス間のデータ転送、デバイス側コードの実行、デバッグなどの機能を提供する。CSL は Data Structure Descriptors (DSD) と呼ばれる特殊な記述子を提供し、テンソルやその他のデータ構造へのアクセスパターンを記述する。DSD と組み込み関数を用いることで、前述の単一命令で最大 4 次元のテンソルを効率的な処理を実現する。

3.3.4 Preferred Networks 社 MN-Core2

MN-Core シリーズは、Preferred Networks (PFN) 社が開発した、深層学習を中心とする大規模行列演算に特化した AI アクセラレータである。演算性能あたりの消費電力を重視した設計を特徴としており、従来の GPU を用いた計算と比較して、電力効率の面で大きな優位性を持つ。第 1 世代である MN-Core を搭載したスーパーコンピュータである MN-3 は、Green500 において複数回 1 位を獲得している。第 2 世代である MN-Core 2 では、第一世代の設計思想を継承しつつ、各種の性能向上に加えて、ソフトウェアスタックの整備が進められている。これにより、AI ワークロードに加えて数値計算を中心とする HPC ワークロードへの対応も進められている。このように MN-Core 2 は、AI アクセラレータとしての高い電力効率を維持しながら、より広い計算分野への適用を視野に入れた装置である。

3.3.4.1 計算機アーキテクチャ

一般的な GPU や多くの高性能プロセッサでは、命令スケジューラ、キャッシュなどの制御ロジックがハードウェアとして実装されている。これにより、既存のコード資産を大きく変更することなく、高い性能を引き出せるよう設計されている。一方で、このような設計では、制御ロジックがチップ上の回路面積の多くを占め、演算器が占める割合は必ずしも高くならない。これに対して MN-Core では、シリコン上におけるハードウェア制御ロジックを極力削減し、演算器のシリコン利用効率を最大化する設計思想を採用している。MN-Core 2 においては、トランジスタ数に占める演算器の割合が約 7% であり、他の汎用プロセッサや演算加速装置と比較しても、非常に高い水準である。このような設計により、チップ面積あたりの演算性能および電力効率の向上が図られている。

MN-Core のもう一つの大きな特徴として、各計算コア (Processing Element, PE) がプログラムカウンタや命令デコーダを持たず、すべての PE が単一の命令列に基づいて完全に同期して動作する点が挙げられる。MN-Core では、ホスト CPU 側で生成された命令列が MN-Core に供給され、全 PE が同一の命令を同時実行する。装置全体が大きな SIMD (Single Instruction Multiple Data) の様に実行されるモデルを採用している。これにより、演算単位ごとに実行進度がばらつくことによって生じる演算負荷の偏りや、それに伴う同期処理のオーバーヘッドを原理的に排除している。この SIMD 型実行モデルは、MN-Core において明確な階層構造として実装されている。最小の演算単位は PE であり、4 つの PE が 1 つの行列演算器 (Matrix Arithmetic Unit, MAU) を共有することで、Matrix Arithmetic Block (MAB) を構成する。MAB は、演算とメモリアクセスを一体として扱う基本単位であり、PE 群が完全に同期した状態で計算を実行する。さらに、16 個の MAB が集約されて Level-1 Broadcast Block (L1B) を形成し、8 個の L1B が Level-2 Broadcast Block (L2B) を構成する。MN-Core (第一世代) では、4 個の L2B が 1 ダイを構成し、4 ダイによって 1 ボードが構成されている。一方 MN-Core 2 では、8 個の L2B によって 1 ボード全体が構成されている。こ

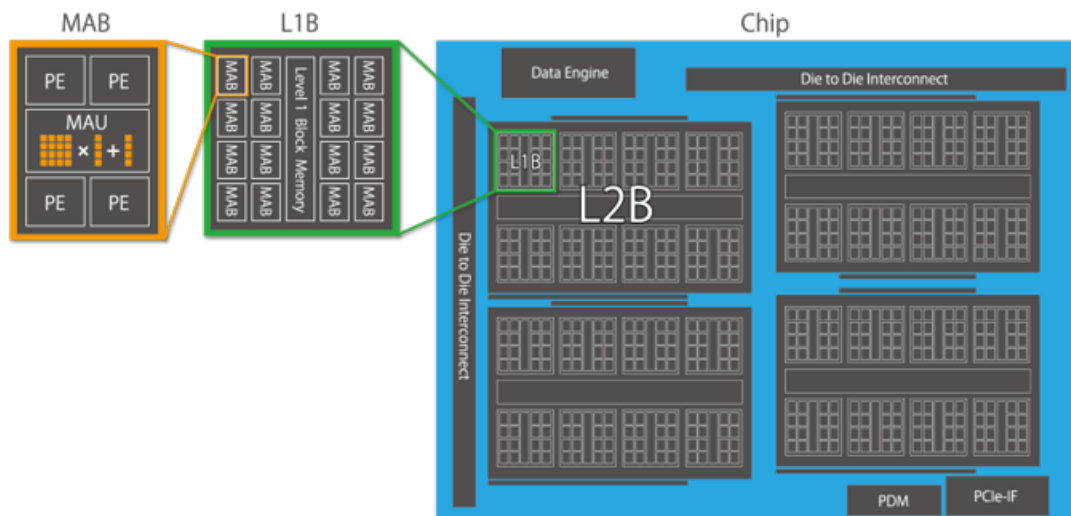


図 3.3.5 MN-Core2 の内部構造 (Preferred Networks 社, “MN-Core 2 Whitepaper”, 2023-11-12 版 より)

のように MN-Core では、PE から MAB、L1B、L2B、そしてボードレベルに至るまで、すべての階層において同期的な SIMD 実行モデルが一貫して採用されている。

一方で、このようなアーキテクチャでは、命令供給やデータ配置、メモリアクセス制御をハードウェア側に任せることができない。そのため、各 PE に対して計算に必要なデータを継続的かつ適切に供給しなければ、演算器が十分に活用されず、性能が大きく低下する可能性がある。すなわち、MN-Core の性能を最大限に引き出すためには、ハードウェア構造を十分に理解したうえで、ソフトウェア側での明示的な最適化を行うことが不可欠である。

ここからは、MN-Core を構成する PE、MAU、MAB、L1B、L2B およびメモリ構成について、その役割とその階層構造を概説する。

Processing Element (PE)

PE は演算に必要なデータの保持および演算処理を担う。データを保持するための汎用レジスタファイル (General Register File, GRF) およびローカルメモリ (Local Memory, LM) を備えており、これらから読み出されたデータが演算器 (Arithmetic and Logic Unit, ALU) に入力される。演算結果は、再び GRF や LM へ書き戻される。また、演算途中の一時データを保持するための T レジスタや、演算結果のフラグを保持するマスクフラグレジスタも備えている。さらに、上位階層である L1B とのデータ転送を行う機能を持ち、これらのデータ移動や演算処理は、PE 命令によって制御される。

Matrix Arithmetic Unit (MAU) と Matrix Arithmetic Block (MAB)

4 つの PE が 1 つの MAU を共有し、これらをまとめた構成単位が MAB である。MAU は、PE から入力データを受け取り、行列演算を中心とした演算処理を行う専用の演算器である。演算結果は再び PE へと返却され、後続の処理に利用される。MAU は行列演算に用いる専用の行列レジスタを備えており、行列ベクトル積演算器において使用される行列データを保持する。これらの行列レジスタは二重化されており、一方の行列レジスタに次の演算用データを書き込みながら、もう一方の行列レジスタを用いて演算を実行できる。

Level-1 Broadcast Block (L1B)

L1B は、16 個の MAB と Level-1 Broadcast Memory (L1BM)、Level-1 Broadcast Reduction Unit か

ら構成される。L1BM と MAB (PE) の間では、L1BM から MAB へ多様な転送モードが利用可能である。データ転送は PE 命令によって制御され、プログラムから明示的に指定される。

Level-2 Broadcast Block (L2B)

L2B は、8 個の L1B と Level-2 Broadcast Memory (L2BM)、Level-2 Broadcast Reduction Unit から構成される。L2BM と L1B の間では、L2BM から L1B へ多様な転送モードが利用可能である。L1B と同様に、データ転送は PE 命令によって制御され、計算パターンに応じた柔軟なデータ供給が可能である。さらに L2B は、ホストインターフェースおよび外付け DRAM との間で多様なデータ転送を行う役割を担っており、MN-Core 全体としての入出力およびデータ供給の中核を形成している。

3.3.4.2 汎用プログラミング環境：MN-Core HPCSDK

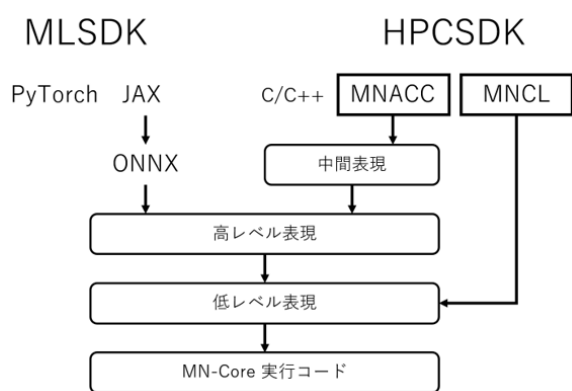


図 3.3.6 MLSDK と HPCSDK のビルドフロー

```
#pragma mnacc config ni_loop(NI) nj_loop(NJ) nk_loop(NK)
#pragma mnacc data first_last
for(n=0 ; n<nn ; n++) {
  gosa = 0.0;

#pragma mnacc parallel for loopcounter(i,j,k) result(wrk2)
  for(i=0 ; i<imax ; i++) { // 変更点1 始め
    for(j=0 ; j<jmax ; j++) {
      for(k=0 ; k<kmax ; k++) {
```

図 3.3.7 MNACC を利用した姫野ベンチマークの抜粋

MN-Core の設計思想は、制御ロジックをソフトウェア側に委ねることで高い電力効率を実現する一方、ユーザがハードウェア構造を深く理解した上で最適なコードを記述することを前提としている。一方で、各 PE へのデータ供給やメモリ階層を考慮しながらアセンブリレベルで最適化を行うことは、一般のユーザにとって容易ではない。このような課題に対処するため、MN-Core 2 ではデータ構造の定義からコード生成に至るまでを自動化する仕組みを備えたソフトウェアスタックが開発されている。

MN-Core 2 向けのソフトウェアスタックは、用途に応じて AI ワークロード向けの MLSDK と、HPC ワークロード向けの HPCSDK の 2 種類に分けられる。図 3.3.6 に両方のビルドフローを示す。MLSDK は、PyTorch で記述された深層学習モデルを入力とし、演算の分割やデータ配置、コード生成を自動的に行うことで、MN-Core 2 上での効率的な実行を可能とする開発環境である。これにより、ユーザは MN-Core のハードウェア構造を意識することなく、深層学習モデルを実行できる。一方、HPC ワークロード向けに提供される HPCSDK には、MNCL および MNACC の 2 種類の開発環境が含まれている。MNCL は、OpenCL に類似したプログラミングモデルを採用した MN-Core 2 向けの API であり、ユーザが計算内容やデータ配置を明示的に制御することで、高い性能を引き出すことが可能である。プログラム記述の自由度が高い分、開発者には MN-Core のアーキテクチャに関する深い理解が求められ、記述難易度や開発負担は依然として大きい。これに対して MNACC は、OpenACC に類似したディレクティブを用いることで、既存の C/C++ コードに比較的少ない修正を加えるだけで、MN-Core 2 上での並列実行を可能とするプログラミングモデルを提供している。MNACC は、細かな最適化をコンパイラ側に委ねることで、性能と実装容易性の両立を目指した開発環境である。図 3.3.7 に、MNACC を用いた姫野ベンチマークの実装を一部抜粋する。OpenACC の一般

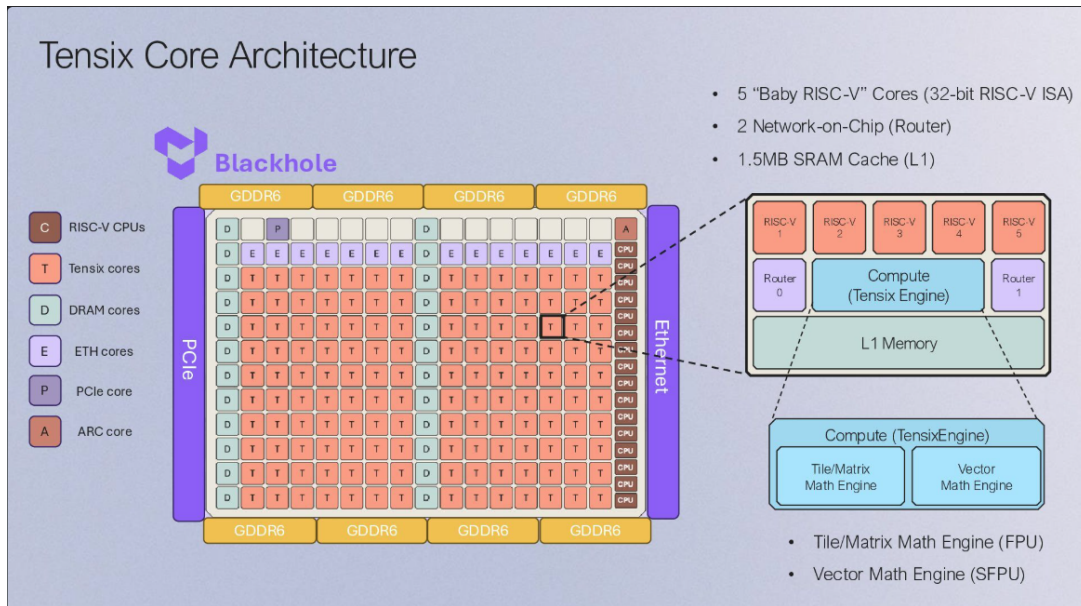


図 3.3.8 Tensix Core のアーキテクチャ (TT-Metalium によるカーネルプログラミング - Tenstorrent Tech Talk #2, Session2 - Tenstorrent Japan より)

的な節と句に加え、`first_last` などの独自の句が追加されている。

3.3.5 Tenstorrent 社 Blackhole

Tenstorrent 社の Blackhole は、ISA に RISC-V を採用し、ハードウェアだけでなくソフトウェアの多くをオープンにした、演算加速装置である。また、HBM よりも安価な DRAM のみを搭載してハードウェアのコスト効率を重視している。Blackhole 世代では TF32 や BF16 といったテンソル計算コアを搭載しており、近年の深層学習ワークロードに適した低～中精度の演算を効率的に処理する。今回選定した 3 機種の中では AI 寄りの設計として位置づけられ、推論・学習ワークロードを想定した最適化が施されている。アーキテクチャは 14×10 個の演算素子 (PE) を 2 次元リングトポロジで接続する、データフローを念頭においた形となっている。計算性能やメモリ性能の実効値の情報が限られ、具体的な数値の測定が必要と考える。図 3.3.8 に Blackhole と Tensix Core のアーキテクチャを、図 3.3.9 にソフトウェアスタックを示す。詳細な調査は来年度の課題とする。

3.3.6 次年度の調査研究の計画

引き続き各種先端演算加速装置の動向調査を行い、状況把握に努める。また、Cerebras Systems 社、Tenstorrent 社、Preferred Networks 社の先端演算加速装置の実効性能やプログラミング環境について調査を行なう。他のサブグループと連携し、収集された要件を基に重点的に考慮すべきアプリケーションを特定し、それらに対応するベンチマークや計算カーネルの候補を抽出する。具体的には、次世代アプリケーションニーズ調査研究サブグループとともに評価すべき項目を特定し、数値計算ライブラリサブグループとともに必要となるライブラリを明確化する。また、STREAM ベンチマークなどのマイクロベンチマーキングを行ない、各装置の性能評価とする。さらに、各装置のスケーラビリティの性能を比較する指標を検討したい。例えば、ダ

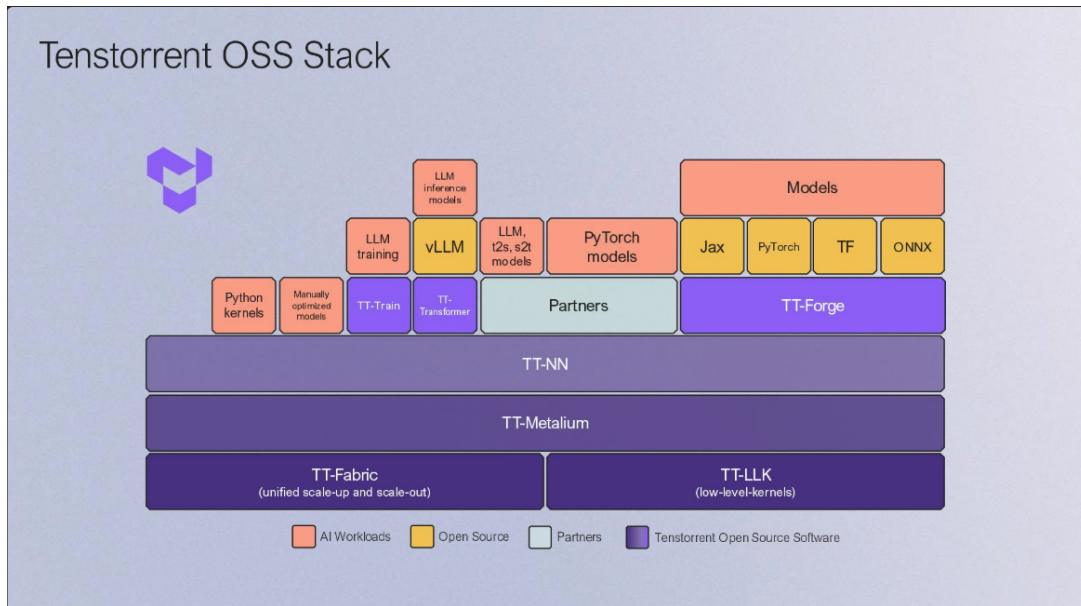


図 3.3.9 Tenstorrent 社のソフトウェアスタック (TT-Metalium によるカーネルプログラミング - Tenstorrent Tech Talk #2, Session2 - Tenstorrent Japan より)

イ (チップ) 間の通信帯域幅をダイ間バンド幅と定義し、ダイ間のデータ移動が全体の性能に直結かどうかを調査する。得られた知見を横展開して調査研究内の各グループへ共有・技術検証を推進する。

3.4 数値計算ライブラリサブグループ

3.4.1 調査研究の目的

数値計算ライブラリは、物理現象の精密なシミュレーションから、近年の爆発的な発展を見せる AI・データ分析のアプリケーションに至るまで、幅広い科学技術計算を支える基盤技術の一つである。近年のハイパフォーマンス・コンピューティング (HPC) 環境は、GPU を筆頭とするヘテロジニアスなアーキテクチャの普及や、AI・データサイエンス分野との融合により、劇的な変化を遂げている。これに伴い、科学技術計算の基盤を支える数値計算ライブラリに対しても、単なる高速化だけでなく、多様な計算精度への対応や、新しいアルゴリズムの実装が強く求められている。さらに、ソフトウェアの保守や維持の観点から、それを支える数値計算ライブラリ自体を、開発者のライフタイムを超えて持続的に維持することも重要な課題となっている。

このような背景を踏まえ、本サブグループは、将来の HPC システム上で実行される多種多様なアプリケーションを支える数値計算ライブラリについて、その技術動向を調査する役割を担っている。具体的には、国内外の最新動向を把握し、アーキテクチャやアプリケーションの関連グループと連携しながら、数値計算ライブラリの開発・整備の方向性を検討することを目的としている。将来の HPC システムのアーキテクチャや主要アプリケーションにおけるニーズ、応用数学分野における新たな数値計算技術などを総合的に踏まえ、今後の開発・整備に向けた指針を示すとともに、持続的な開発体制の確立に向けて課題や方針を整理することを目指している。また、次世代旗艦システム (NFS) や次世代情報インフラ (NIS) の設計に有益な情報を整理し、システム間のギャップを軽減することでアプリケーションの可搬性向上に貢献することも重要な狙いである。

主要な数値計算ライブラリの利用実態や要望を調査し、それらに関する解説や性能ベンチマーク結果を含む資料を作成・提供することで、利用者の利便性向上と運用側のコスト削減を同時に実現することを目指す。

3.4.2 調査研究の結果

本サブグループにおける令和7年度の活動は、将来の数値計算ライブラリの在り方を定義するための基盤構築期間と位置づけ、以下の二段階のプロセスを経て実施する計画を策定した。

まず、年度前半にあたる10月から12月にかけては、数値計算ライブラリに関する詳細な調査戦略の具体化を最優先課題とした。本活動を実効性のあるものとするため、各参加者の専門性を活かした役割分担を再定義するとともに、アーキテクチャグループやアプリケーショングループといった関係グループとの密接な連携体制を明確化する。これと並行して、令和4年度から6年度にかけて実施された先行フェーズ（FS2.0）の実施内容および最終報告を精査する。この精査プロセスを通じて、これまでの調査で明らかになった既存ライブラリの現状と課題を再確認し、次世代システム（NFS/NIS）に向けて重点的に取り組むべき技術要素や調査項目を抽出・整理する作業を行う。

続いて、年度後半の1月から3月にかけては、前半に整理した重点項目をさらに深化・具体化させるフェーズへと移行する。具体的には、国内の数値計算・HPC分野における有識者との多角的な議論を通じて、調査すべき項目の優先順位をさらに明確化する。また、国内にとどまらず、国際会議等への参加を通じて、混合精度計算や乱択アルゴリズム、GPU最適化技術といった関連分野における世界最先端の技術動向を直接的に調査する。さらに、これらの調査活動と並行して、令和8年度から本格的に開始する予定のベンチマーク評価に向けた具体的な準備に着手する。これには、評価対象とする計算機環境の選定やソフトウェアスタックの構築、および予備的な性能評価の実施が含まれており、次年度の円滑な研究遂行を確実なものとするための重要なステップとなる。

3.4.2.1 FS2.0の報告内容の振り返り

本サブグループの活動を推進するにあたり、令和4年度から令和6年度にかけて実施された「FS2.0」におけるシステムソフトウェア・ライブラリ調査研究グループの数値ライブラリ調査研究サブグループによる調査成果を詳細にレビューした。数値ライブラリ調査研究は今村俊幸氏（理化学研究所）をサブグループリーダーとして進められたものであり、既存の数値計算ライブラリの現状把握から実機での性能評価、さらには将来的な課題の抽出まで、極めて広範かつ詳細な調査がなされている。

まず、ライブラリの網羅的な把握という点において、FS2.0では主要な数値計算ルーチンの現状が極めて正確にリストアップされている。具体的には、BLAS、LAPACK、ScaLAPACKといった基本線形計算ライブラリをはじめ、疎行列直接法、高速フーリエ変換（FFT）、疑似乱数生成などが対象となっている。これらのライブラリについて、単精度や倍精度といった対応精度の違いだけでなく、C/C++やFortranといったプログラミング言語への対応状況についても精緻な調査が行われた。これにより、現在のHPC環境においてどのライブラリがどの程度の汎用性を持って利用可能であるかが明確化されている。

性能評価の面では、CPUおよびGPUの両環境において、代表的なルーチンのベンチマークが実施された。CPU環境においては、A64FXやIntel Xeonといった主要なプロセッサ上でのBLAS/LAPACKの性能が測定されており、令和6年度にはGPU環境におけるベンチマークも追加実施されている。これらのデータは、計算機アーキテクチャの違いが数値計算の実行効率に与える影響を客観的に示すものであり、次世代システム（NFS）の開発を念頭に置いた重要ライブラリや機能の選定において、極めて重要な判断材料となっている。

また、ライブラリの「提供側」の視点だけでなく、「利用側」の視点を取り入れている点も FS2.0 の大きな特徴である。HPC 利用者に対するアンケート調査が実施され、実際にどのようなライブラリやデータフォーマット、計算精度が現場で求められているのかという実態が浮き彫りにされた。この調査結果に基づき、世界的に通用する国際 HPC ソフトウェアとしての重要度が高いライブラリの選定が行われている。

特筆すべきは、これらの技術的調査の総括として提示された「数学ソフトウェアの開発・維持に関する課題」への提言である。FS2.0 の報告書では、数学ソフトウェアの寿命が開発者個人の活動期間（ライフタイム）と密接にリンクしているという、ソフトウェア・サステナビリティの本質的な問題が指摘された。開発から 10 年、20 年と経過した後に、コミュニティベースで真にメンテナンス可能な体制を維持できるかという問いは、単なる技術論を超えた、科学技術基盤の継続性に関わる深刻な課題である。次世代計算基盤の構築に際しては、コミュニティに積極的に働きかけるなど、外部を意識した戦略的な活動が必要不可欠であるという認識は、本サブグループ（FS3.0）の活動においても確固たる指針となるものである。

以上のレビューを踏まえ、FS2.0 で蓄積された膨大な知見を継承しつつ、FS3.0 ではさらにその先を見据えた調査を重点的に行う。具体的には、これから数値計算ライブラリに組み込まれることが期待される最新のアルゴリズム動向の把握、および欧米を中心とした海外における開発・保守体制の現状調査、そして次世代情報インフラ（NIS）におけるライブラリの在り方の再定義である。FS2.0 が築いた強固な基礎の上に、新たな技術潮流と国際的な運用体制の視点を加えることで、より実効性の高い提言を目指していく。

3.4.2.2 最新技術動向調査：SIAM PP26 を中心に

2026 年 3 月 3 日から 6 日にかけてドイツ・ベルリンで開催された「SIAM Conference on Parallel Processing for Scientific Computing (PP26)」は、次世代の並列数値計算アルゴリズムとその実装技術を議論する国際的に最も重要な会議の一つである。本サブグループでは、現地でのセッション参加および公式プログラムの精査を通じて、将来の HPC システム（NFS/NIS）における数値計算ライブラリの在り方を占う上で極めて重要な知見を得た。以下に、主要な技術項目ごとの詳細な調査結果を報告する。なお、SIAM PP26 のプログラムおよびアブストラクトは https://meetings.siam.org/program.cfm?CONF_CODE=PP26 から確認できる。

■混合精度計算（Mixed Precision）の深化と多様化

本会議において最も大きな潮流となっていたのは、計算精度を柔軟に制御することで演算性能とメモリ効率を最大化する混合精度計算の研究である。単なる「単精度と倍精度の併用」という段階を超え、アルゴリズムの構造そのものに多段階の精度を組み込む試みが目立った。

- 行列積エミュレーション（Ozaki スキーム）

Tensor コアに代表される高性能な低精度行列積ユニットを用いた高精度な行列積のエミュレーション手法として注目されている Ozaki スキームに関連した発表が散見され、その注目度や有効性を再確認した。4 つの講演が全て Ozaki スキームに関連するセッション（MS38）がその代表例である。一方で、行列積の形で演算が可能となる数値計算アルゴリズムが限られている点が実用上の大きな問題の一つであり、Ozaki スキームのような行列積エミュレーション手法の研究開発と並行して、それを活用するための数値計算アルゴリズム自体の本質的な改良を行うことも重要である。

- HODLR 行列および階層型行列への適用

HODLR（Hierarchical Off-Diagonal Low-Rank）行列は、密行列の非対角ブロックを低ランク近似することで計算コストを削減する手法である。Xiaobo Liu 氏（MS10）らの発表では、この HODLR 構造に低精度演算を導入する「混合精度 HODLR」が議論された。後退誤差解析に基づき、行列ベクトル

積や LU 分解において、近似精度が元々低いブロック（低ランク部分）に対しては倍精度を用いるのは冗長であり、低精度演算に置き換えても全体の精度に影響を与えないことが理論的に示された。これはメモリ帯域がボトルネックとなる将来のアーキテクチャにおいて、有効なデータ削減指針となる。

- 行列方程式（Sylvester 方程式）への展開

Sylvester 方程式の解法である Bartels-Stewart アルゴリズムに対し、最もコストのかかる Schur 分解を低精度で行い、後半の代入計算において反復改良（Iterative Refinement）を適用することで高精度な解を得る手法が提案された（MS27）。従来の反復改良とは異なり、「係数行列そのものに低精度に起因する摂動が含まれる」特殊な状況下での収束条件が誤差解析を用いて議論されており、3つ以上の精度（Mixed Precision）を用いたアルゴリズムへの発展も示唆されている。

- GPU 最適化と固有値問題

Chris Baker 氏（MS62）らは、AI 向け GPU の演算性能を引き出すため、固有値問題に対する混合精度 Jacobi 法の数値実験結果を報告した。Double-Single および Double-Single-Half の組み合わせが試されており、ハードウェアの特性（Tensor Core 等）をいかに数値計算カーネルに適合させるかが重要なテーマとなっている。

- テンソル分解への応用

分散並列環境におけるテンソルトレイン（TT）分解に対し、QR 分解の過程で低精度演算を活用する検討が進められている（Eda Oktay 氏, MS62）。これは通信コストの削減に直結するため、大規模並列環境での実効性能向上が期待される。

■疎行列直接解法におけるモダンな実装技術

疎行列直接解法は、シミュレーションの基盤として依然として重要であり、最新のハードウェア（マルチ GPU）への適合が焦点となっている。

- マルチ GPU 対応とランタイム連携

SuperLU 等の代表的なライブラリにおいて、マルチ GPU 化や通信プロトコルの改良が進んでいる。特に、タスク並列ランタイムとの連携や、one-sided communication（片方向通信）を利用した計算と通信のオーバーラップが性能向上の鍵となっている。

- ベンダー提供ソルバの動向

NVIDIA が提供する「cuDSS (CUDA Direct Sparse Solver)」の動向も注目を集めた。特に、直接解法において大きなコストを占めるシンボリック分解の高速化をいかに行うか、という点について活発な議論がなされた。一方で、分散環境におけるピボット選択（Partial Pivoting）の困難さや、数値的な安定性の確保は依然としてオープンな課題として残されている。

■乱択アルゴリズム（Randomized Numerical Linear Algebra）

乱択アルゴリズムは、行列の「スケッチング（次元圧縮）」により計算量を劇的に削減する手法として期待されているが、本会議での発表件数は期待されたほど多くはなかった。

- スケッチング手法とカーネルの課題

Gaussian Sketch や Count Sketch、SRHT (Subsampled Randomized Hadamard Transform) といった多様なスケッチング手法の特性比較や、疎行列への適用（Sparse JLT）が議論された。しかし、Randomized BLAS/LAPACK のような汎用的なライブラリとして体系化されるには至っていない。

その要因として、従来の BLAS 等の行列積カーネルとは異なる、ランダムサンプリングやハッシュ関数ベースの新しいカーネルの実装が必要となり、既存のソフトウェアスタックとの親和性が低い点が指摘された。

- 特定分野での活用事例

材料科学分野における DFT（密度汎関数理論）計算向けの乱択固有値ソルバ（Rand-JD 等）など、特定のアプリケーションに特化した形での導入が進んでいる事例（MS54）が報告された。

■実装言語とエコシステムの変容

ライブラリの開発・利用環境において、Python や Julia といった高水準言語のプレゼンスが急速に高まっている。

- Julia 言語による HPC 開発

MS70 では、Julia を用いた JACC (Julia Actively Copied Kernels) や QCD (格子量子色力学) アプリケーションへの適用が紹介された。これは、ライブラリの可搬性と開発効率の向上において、C/C++ や Fortran 以外の選択肢が現実的になってきていることを示している。

- 精度自動チューニングツール

CADNA や PROMISE といった、演算精度を自動的に探索・最適化するソフトウェアツールの開発も進展しており、複雑な混合精度アルゴリズムを人間が手動で実装・検証する負担を軽減するアプローチが示された。

以上の調査に基づき、今後の数値計算ライブラリ整備において考慮すべき点は以下の通りであると考えられる。

- 混合精度の標準化:

各種行列方程式や固有値問題において混合精度アルゴリズムが個別に研究されているが、これらを共通の API で利用できるライブラリの整備が急務である。

- ハードウェアとアルゴリズムの乖離:

GPU の AI 機能 (Tensor Core 等) を活用するためには、従来の数学的定式化を「行列積」主体に再構成する必要があり、ライブラリ側での吸収が求められる。

- サステナビリティの再確認:

FS2.0 でも指摘された通り、数学ソフトウェアの開発体制は脆弱であり、欧米 (ExaNLA 等) の組織的な保守体制を参考に、国内でも持続的なメンテナンス体制を構築する必要がある。

本年度の調査により、最新のアルゴリズム動向と、それをライブラリとしてパッケージ化する際の障壁 (実装言語やカーネルの不一致等) が明確になった。これらは次年度以降に実施するベンチマーク評価の対象選定において、重要な技術的指針となる。

3.4.3 次年度の調査研究の計画

令和 8 年度は、上述した令和 7 年度の調査結果を踏まえた上で、引き続き、関連分野の国内会議や国際会議、有識者との議論を通じて、数値計算技術に関連する分野の技術動向調査を継続する。特に、欧州および米国における数値計算ライブラリの開発・整備状況や体制に関する情報収集を行う。また、並行して、CPU 環境および GPU 環境において、混合精度型アルゴリズム等のベンチマークを実施する。加えて、アプリケー

ション分野における数値計算ライブラリへのニーズ等の調査を行うことも視野に入れている。より具体的な計画は以下の通りである。

3.4.3.1 技術動向調査

数値計算技術の進展は極めて速く、特に混合精度演算や乱択アルゴリズムの実用化に向けた研究は世界的に加速している。そのため、令和 8 年度においても引き続き、関連分野の国内会議（HPC 研究会、SWoPP 等）や国際会議（SC、ISC、SIAM 関連会議等）、および国内外の有識者との直接的な議論を通じて、数値計算技術に関連する最新の技術動向調査を継続する。特に、欧州および米国における数値計算ライブラリの開発・整備状況や、その背後にある維持・管理体制に関する情報収集を重点的に行う。具体的には、米国の BLAS/LAPACK プロジェクトや欧州の ExaNLA プロジェクト等の動向を精査し、開発者のライフタイムに依存しない持続可能なソフトウェア・エコシステムの在り方について、日本国内への適用可能性を検討する。

3.4.3.2 ベンチマーク

まず、疎行列を係数とする連立一次方程式向けの混合精度型アルゴリズムの性能評価を行う。具体的には、最新の CPU 環境および GPU 環境の両面において、低精度演算ユニットを積極的に活用した手法の優位性を検証する。前処理や一部の重い演算を低精度（FP32/FP16 等）で行い、反復改良法や Ozaki スキーム等のエミュレーション技術を組み合わせることで、最終的な解の精度を倍精度相当に保ちつつ、実行時間や消費電力をどの程度削減できるかを調査する。特に、行列の条件数や物理問題の特性が収束性に与える影響を詳細に解析し、実用的なアプリケーションにおける適用限界を明らかにする。

次に、乱択アルゴリズム（Randomized Algorithms）の性能検証を実施する。次元圧縮やランダムサンプリングに基づくこれらの手法が、大規模なデータセットを取り扱う線形計算において、従来の決定論的なアルゴリズムと比較してどの程度の計算量削減に寄与するかを評価する。この際、計算時間の短縮だけでなく、近似解の誤差範囲が実用上の許容範囲内に収まるかを、CPU および GPU の各アーキテクチャ上で多角的に検証する。

加えて、これらの最新アルゴリズムを次世代情報インフラ（NIS）の広範なユーザーが活用することを想定し、Python や Julia といった高水準言語から利用する場合の実効性能についても調査を行う。具体的には、数値計算カーネルを、モダンな言語環境から呼び出す際のデータ型変換やランタイムに伴うオーバーヘッドを測定する。これにより、専門的な低レベルプログラミングの知識を持たないユーザーであっても、最新の数値計算技術の恩恵を最大限に享受できるような、最適なライブラリの提供形態およびバインディングの実装指針を議論する。

3.5 ストレージ・ファイルシステムサブグループ

3.5.1 調査研究の目的

HPCI 共用ストレージ、HPCI 各構成機関のストレージシステムについて今後の要求の調査および動向調査を実施する。

共用ストレージを利用していますか？
18 responses

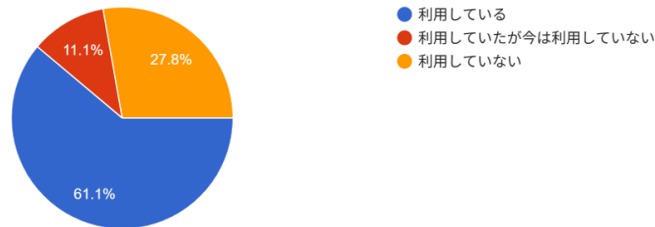


図 3.5.1 HPCI 共用ストレージを利用していますか？

HPCI共用ストレージの利用用途を教えてください。
11 responses

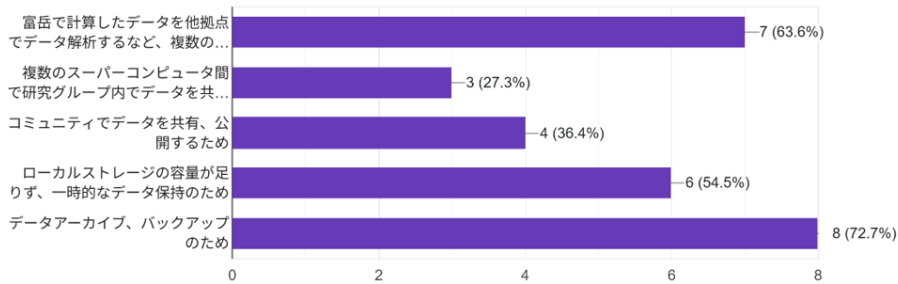


図 3.5.2 HPCI 共用ストレージの利用用途

3.5.2 調査研究の結果

2025 年度の調査研究においては、HPCI 共用ストレージについての利用者アンケートを実施し、現状の問題点の把握および要求を調査した。また HPCI 各構成機関のストレージシステムの調査として現在多くの構成機関で利用している Lustre ファイルシステムと、欧米において利用が増えてきている VAST について動向調査を実施した。

3.5.2.1 HPCI 共用ストレージの利用調査

HPCI 共用ストレージの利用者に対し、利用アンケート (<https://forms.gle/KbnMw1dfNm9wgtJZA>) を実施した。アンケートの内容は、HPCI 共用ストレージの利用用途、利点、何か問題が生じている場合はその問題点、要望・改善点などについてである。2026 年 3 月 8 日までに集まった 18 件のアンケートについて以下にまとめる。回答数の絶対数はまだ多くはないため、引き続きアンケートへの回答を依頼する必要がある。

図 3.5.1 は HPCI 共用ストレージを利用しているかどうかのアンケート結果であり、61% が利用していると回答している。

図 3.5.2 は HPCI 共用ストレージの利用用途についての結果である。以下が主要用途である。

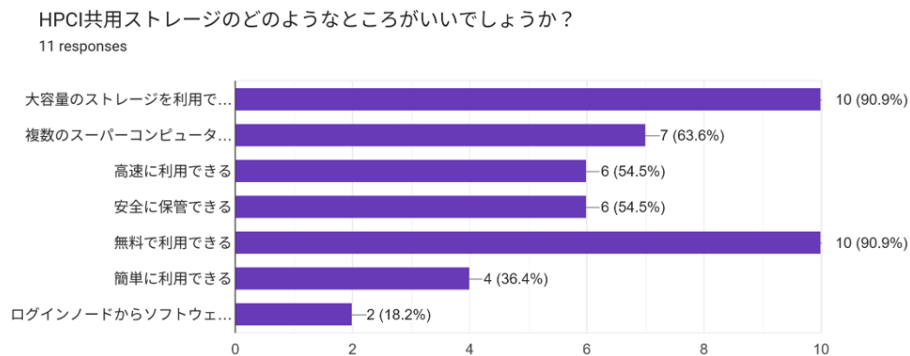


図 3.5.3 HPCI 共用ストレージのどのようなところがいいでしょうか？

- データアーカイブ、バックアップのため
- 富岳で計算したデータを他拠点でデータ解析するなど、複数のスーパーコンピュータ間でデータを共有するため
- ローカルストレージの容量が足りず、一時的なデータ保持のため

また

- コミュニティでデータを共有、公開するため
- 複数のスーパーコンピュータ間で研究グループ内でデータを共有するため

についても重要な用途であることが分かる。

図 3.5.3 は HPCI 共用ストレージのいいところを問い合わせた結果である。

- 大容量のストレージを利用できる
- 無料で利用できる

が多く、次いで

- 複数のスーパーコンピュータ間でデータを共有できる
- 高速に利用できる
- 安全に保管できる

が評価されている。利用者は大容量のストレージを求めており、また HPCI 共用ストレージが高速、安全に利用できるといった利用者のニーズを満たしていることが分かる。

図 3.5.4 は HPCI 共用ストレージを利用していない、あるいは利用していたが今は利用していない利用者について、その理由を尋ねたものである。大容量のストレージが不要、複数のスーパーコンピュータ間でのデータ共有が不要、というコメントがある一方で、使い方がよくわからない、HPCI 共用ストレージをよく知らないというコメントもあり、利用者説明会等による広報活動が必要であると思われる。

また、要望その他として以下の意見があった。

1. 非常に有用なサービスをありがとうございます。とても使いやすいです。

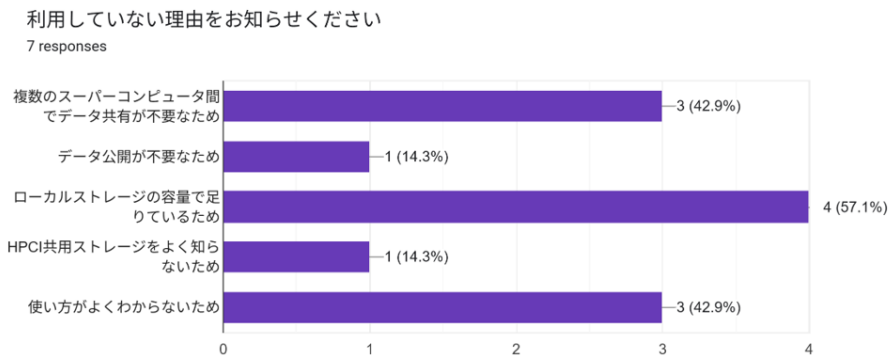


図 3.5.4 HPCI 共用ストレージを利用していない理由

2. データ量は増加の一方であり、HPCI 共用ストレージは非常に重要。今後も強化してほしい。
3. このまま安定なサービスを提供してほしい。
4. 必要になった時に申込できるようにしてほしい。
5. HPCI の課題採択とは別に利用したい。不採択時にデータ退避を求められる危険性がある。
6. Web アプリの利用方法を詳しく知りたい。
7. 利用マニュアルをより分かりやすく。便利な使い方の周知を。

1、2 は非常にポジティブなコメントであり、HPCI 共用ストレージの有用性、必要性が強く示されている。3 については、HPCI 共用ストレージの運用がうまく機能していることを示している。4、5 のコメントについては、HPCI 共用ストレージの運用側も認識しており、HPCI 共用ストレージ（共有型）利用研究課題について随時募集をしているが、利用者に十分周知されていないことが分かる。6、7 については、これまでも利用者説明会等を開催しているが、今後も継続して開催するとともに、便利な使い方などについて周知することが必要と考えられる。

3.5.2.2 Lustre ファイルシステムの動向調査

Lustre ファイルシステムは 25 年以上の実績を持つ POSIX 準拠並列ファイルシステムである。HPC ワークロードにおける長所としては主に以下の点があげられる。

- 世界のスーパーコンピュータの TOP100 システムのうち 60% が Lustre を利用している。
- POSIX の厳密な一貫性モデルを備えており、広範囲のアプリケーションで動作実績がある。
- 長い運用実績があり、多くの利用者、未知のアプリケーション、複雑な運用形態においても安定稼働できる。
- 数万ノード規模から 10 万ノード超の大規模環境での運用実績があり、高いスケーラビリティを有している。
- 構築から運用開始まで、特別な事前テストやチューニングを必要とせず、タイムロスがない。
- 複数サーバでメタデータ処理を行うことができ、高いメタデータ性能およびスループット性能を有している。
- オープンソースソフトウェアであり、マルチベンダによる活発な開発とエコシステムが構築されて

いる。

- 幅広いハードウェアで動作し、高いハードウェア互換性を有している。

一方、HPC ワークロードの短所としては主に以下の点があげられる。

- 機能追加や拡張は継続的に行われているが、長年の設計を前提としているため、複雑性が増し、実装・適用に時間を要する場合がある。
- Lustre 自体には統合管理機能が存在しないため、3rd party ツール、またはベンダー提供の管理ソフトウェアが必要である。
- Lustre はシンプルな Failover/Failback モデルを採用しており、フェイルオーバー時には I/O 待ち時間が発生する可能性がある。
- 厳密な一貫性が不要なアプリケーションでは、一貫性保持のためのオーバーヘッドが性能制約となる。
- ハードウェアの進化に対し、長年の設計を前提としているためメタデータ性能の効率的な活用が難しいケースがある。
- I/O アラインメントを前提としないワークロードでは性能が出にくい。

また、近年利用が多くなってきている AI ワークロードについては以下の長所がある。

- 多くの大規模 AI システムにおける運用・稼働実績がある。
- 高い書き込み性能を活かしたチェックポイント処理を可能である。
- KV Cache オフロードのインテグレーションの実績がある。

一方、短所としては以下の点があげられる。

- クラウドネイティブ環境との統合性に制約がある。
- メタデータのタグ付けおよび検索機能の不足している。

これらについては、今後の開発が必要な点と考えられる。

今後の Lustre の開発ロードマップにおいては以下が含まれる。

- Lustre クライアントの Linux カーネルアップストリームへのマージ
- 障害時のダウンタイムをなくすための可用性の向上
 - Raft による MGT の冗長化実装
 - メタデータの複製
 - ファイルデータの Erasure Coding
- 性能向上
 - メタデータ Writeback Cache
 - メタデータ性能の効率化・最適化

Linux カーネルのアップストリームへのマージにより、より利用が広がることが想定される。可用性の問題についてもダウンタイムをなくすための機能開発が予定されている。また、さらなる性能向上についての開発も計画されている。

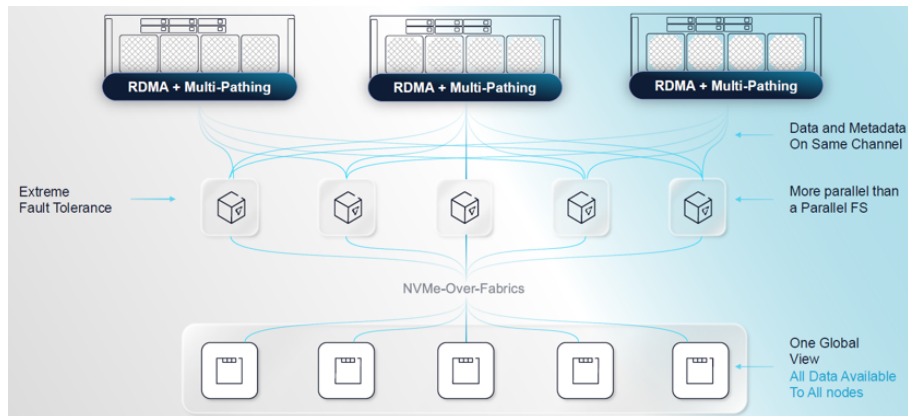


図 3.5.5 VAST のアーキテクチャ (VAST の Kyle Lamb のスライドから引用)

3.5.2.3 VAST について

VAST Data は 2016 年に設立された米国の AI データプラットフォーム企業で、機械学習やディープラーニングなど AI 時代のデータ集約型ワークロードに特化したストレージを提供している。図 3.5.5 に示すような、サーバ (C ノード) とデータ (D ノード) を分離し、C ノードが全ての D ノードを共有する DASE (Disaggregated Shared-Everything) アーキテクチャを採用している。C ノードと D ノードを独立して拡張できるため無駄のないスケールアウトが可能である。

C ノードと D ノードの間は NVMe-oF で接続され、またクライアントからはマルチパス NFS や RDMA を用いてアクセス性能を向上させている。C ノードはステートレスであり、またデータは Erasure Coding で冗長化されるため耐障害性が高い。

NFS, SMB, S3 などの複数のプロトコルでアクセス可能であり、全てを KV ストア (Element Store) 上で提供している。また、単なるストレージに留まらず、AI 開発全体を支える AI OS としてのプラットフォームを展開している。構造化データと非構造化データのサービスをネイティブに提供し、ベクトルデータベースなども提供している。サーバレス関数の実行や外部サービスのトリガ、エージェントシステムの実行、メタデータタグの付加と通知もサポートしている。

HPC ワークロードにおける長所としては主に以下の点があげられる。

- 米国のスーパーコンピュータセンタにおいていくつかの大型の導入計画がある
 - NERSC Doudna において VAST と IBM を導入予定 (2025 年 7 月発表)
 - TACC Horizon (300 PFlops FP64, 80 EFlops FP4) において 400 PB、10 TB/s の VAST を導入予定 (2026 年 4 月運用開始予定)
 - ALCF (Argonne Leadership Computing Facility) にも導入予定
- Quality of Service 機能があり性能が保証される
- 障害時の停止時間がない、性能レジリエンスが高い

米国における複数のスーパーコンピュータセンタにおける導入計画が発表されており、期待の高さを感じられる。Quality of Service 機能があり、他のジョブの影響を大きく受けることなくアクセスが可能となっている。障害発生時においてもフェイルオーバーに頼ることなくアクセスを継続でき、性能レジリエンスも高いのが特徴

である。

一方、短所としては以下の点があげられる。

- IO性能の世界ランキングである IO500 において高い性能を出しているシステムがない。
- プロプライエタリソフトウェア
- オールフラッシュ構成前提で HDD 構成をとれない。

2025年11月のIO500ランキングにおいてVASTは35位（総合スコアは47、バンド幅スコアは19、メタデータ性能スコアは119）であるが、一方Lustreは3位（総合スコアは839、バンド幅スコアは438、メタデータ性能スコアは1,604）である。IO500ランキングはHPCワークロードにおける簡単なケースと複雑なケースのバンド幅とメタデータ性能を基に算出されるものであり、このスコアはHPCワークロードにおいて重要と考えられている。プロプライエタリソフトウェアであるため、コミュニティやマルチベンダによる開発ができず完全にVAST Dataに頼ることになる。オールフラッシュ構成前提のため容量に対するコストが高くなるが、データ圧縮や重複排除により物理的なデータ容量以上の容量を格納可能となっている。

AIワークロードについては、特に以下のような長所がある。

- AIパイプライン統合
- ベクトルデータベースのネイティブサポート
- 障害に強い

ストレージとしての機能を越えたAIに必要な機能の開発が行われている。

3.5.3 次年度の調査研究の計画

HPCI共用ストレージについてはアンケートをより広範囲で実施しより多くの利用者の意見、コメントを集約し、要求をまとめ、今後の整備、機能開発における提案をまとめる。HPCI各構成機関のストレージシステムについて、Lustreファイルシステム、VASTについて引き続き動向調査を実施するとともに、その他のストレージシステムについても動向調査を実施する。

参考文献

- [13] H. Carter Edwards, Christian R. Trott, and Daniel Sunderland. “Kokkos: Enabling manycore performance portability through polymorphic memory access patterns”. In: *Journal of Parallel and Distributed Computing* 74.12 (2014), pp. 3202–3216.
- [14] Jack Dongarra et al. “The Convergence of HPC and AI: A New Paradigm for Scientific Discovery”. In: *Communications of the ACM* 64.1 (2021), pp. 56–65.