［大学 ICT 推進協議会　2020 年度年次大会論文集より］

東北大学サイバーサイエンスセンター スーパーコンピュータ AOBA の紹介

山下 毅 [1)], 森谷 友映 [1)], 佐々木 大輔 [1)], 齋藤 敦子 [1)]
小野 敏 [1)], 大泉 健治 [1)], 滝沢 寛之 [2)]

1) 東北大学 情報部情報基盤課
2) 東北大学 サイバーサイエンスセンター

yamacta@tohoku.ac.jp

Introduction of Supercomputer “AOBA” at the Cyberscience Center of Tohoku University.

YAMASHITA Takeshi[1)], MORIYA Tomoaki[1)], SASAKI Daisuke[1)], SAITO Atsuko[1)]
ONO Satoshi[1)], OIZUMI Kenji[1)], TAKIZAWA Hiroyuki[2)]

1) Information Infrastructure Division, Information Department, Tohoku Univ.
2) Cyberscience Center, Tohoku Univ.

概要

東北大学サイバーサイエンスセンターは，全国共同利用設備として大規模科学計算システムの整備と，HPCI の資源提供機関としての役割を担っている。本稿では 2020 年 10 月に運用を開始したスーパーコンピュータ AOBA と，ユーザの利用環境および本センターが実施する高速化支援活動について紹介する。

1　スーパーコンピュータ AOBA

東北大学サイバーサイエンスセンター (以下，本センター) では，2020 年 10 月からスーパーコンピュータ AOBA の運用を開始した。スーパーコンピュータ AOBA はサブシステム AOBA-A(SX-Aurora TSUBASA，日本電気株式会社製)，サブシステム AOBA-B(LX 406Rz-2，日本電気株式会社製) の 2 種類の計算機システムと，ストレージシステム (DDN SFA7990XE，DDN 社製)，大判プリンタ，講習会端末およびそれらを接続するネットワーク機器群で構成される。図 1 にシステム構成図を示す。

以下ではスーパーコンピュータ AOBA の 2 つの計算機システムとストレージシステムについて，ハードウェアおよびソフトウェアの特徴と，利用者環境について紹介する。

1.1　サブシステム AOBA-A(スーパーコンピュータ)

■ハードウェア　今回導入した SX-Aurora TSUBASA は, 前スーパーコンピュータシステムの SX-ACE と同じくベクトルアーキテクチャを継承している。アプリケーション演算処理を行うベクトルエンジン (以下, VE) 部と，主に OS 処理を行うベクトルホスト (以下，VH) 部により構成される。PCIe カードに搭載される VE 部はベクトルプロセッサおよび高速メモリから構成され，x86/Linux が動作する VH と PCIe 経由で接続される。

今回本センターが導入した VE(Type 20B) は，理論演算性能 2,456GFLOPS(倍精度) となるマルチコア (8 コア) ベクトルプロセッサを 1 基，主記憶は 48GB を搭載し, 1.53TB/s という高いメモリバンド幅でプロセッサと接続されることで，高い演算性能とメモリ性能の両立を実現している。本センターのサブシステム AOBA-A は，1VH と 8VE が構成単位となる B401-8 モデルを採用し，サブシステム全体では 72 個の VH と 576 個の VE で構成される。VE と VH を合わせたシステム全体の理論演算性能は，1.48PFLOPS(倍精度)，総主記憶容量は 45TB，総メモリバンド幅は 895.68TB/s となる。図 2 に AOBA-A を構成する B401-8 と，それに搭載される VE の外観を図 2 に示す。

■プログラミング言語　SX-ACE と同じく，アプリケーションの実効性能を向上させる高度な自動ベクトル化・自動並列化機能を備えた Fortran/C/C++ コンパイラが利用できる。自動並列化機能および

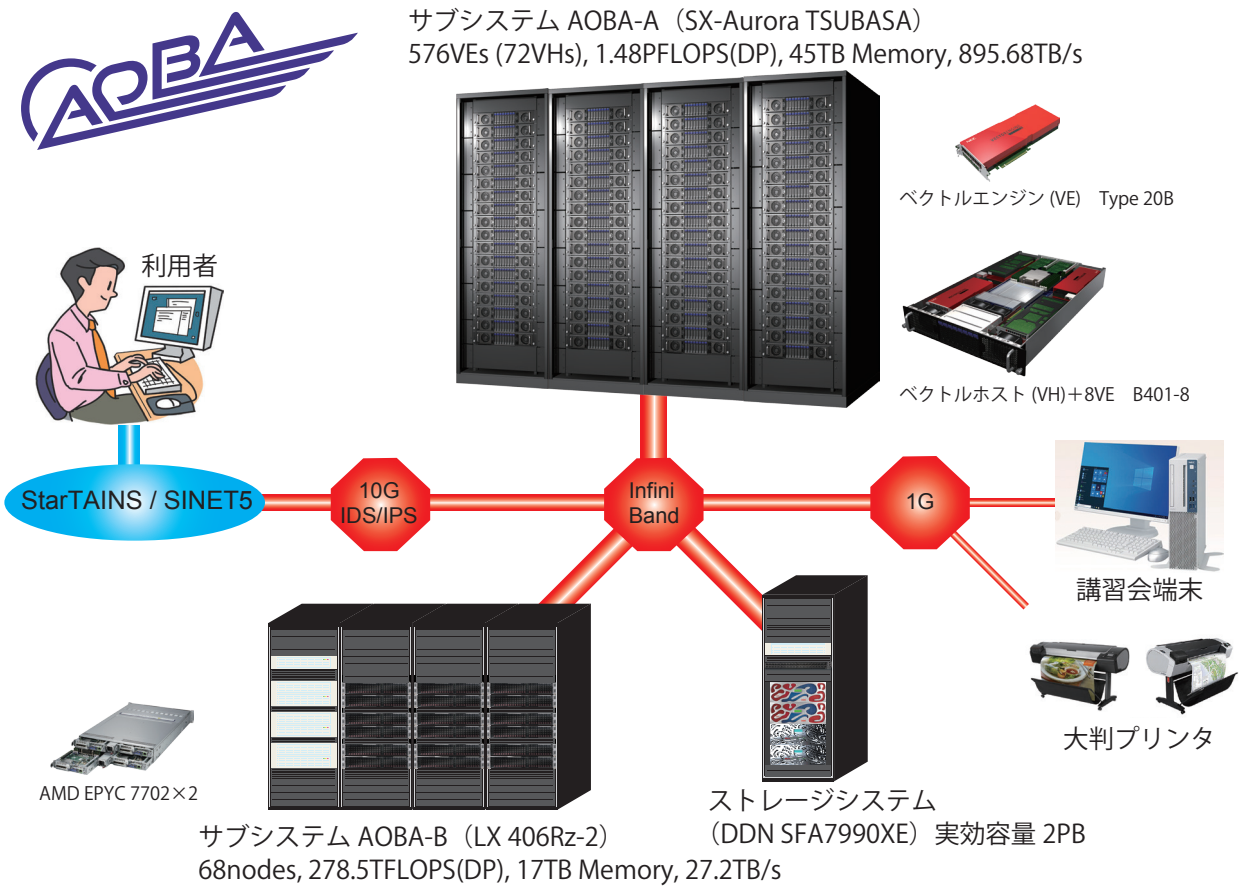

図 1　スーパーコンピュータ AOBA の構成

図 2　サブシステム AOBA-A(VE と B401-8)

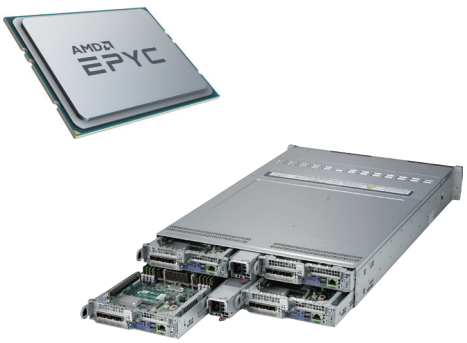

図 3　サブシステム AOBA-B(プロセッサと 4 ノードシャーシ)

OpenMP による共有メモリ並列実行と，システム構成に最適化された MPI ライブラリによる，分散メモリ並列実行が可能である。また科学技術計算ライブラリとして，VE に最適化された数学ライブラリのコレクション NEC Numeric Library Collection(NLC)が利用できる。

SX-ACE で動作していたプログラムをサブシステム AOBA-A に移植する場合には，そのプログラムを SX-Aurora TSUBASA 用のコンパイラでコンパイルし直す必要がある。なお，SX-Aurora TSUBASA 用コンパイラでは GNU コンパイラ互換性が強化され，指示行やコンパイルオプションが SX-ACE 用のものから変更されている。このため、移植時にはそれらの差異に注意が必要である。

■アプリケーション　サブシステム AOBA-A では，VE 向けに移植された商用アプリケーションの VASP や，オープンソースソフトウェア (OSS) の Quantum Espresso を利用できる。また，今後も VE 向けに移植されたアプリケーションを拡充する予定である。な

お，VASP の利用には利用者が契約したライセンスの提示が必要である。

1.2　サブシステム AOBA-B(並列コンピュータ)

■ハードウェア　今回導入した LX 406Rz-2 は，1 ノードに AMD EPYC プロセッサ 7702(64 コア) を 2 基と 256GB の主記憶装置を搭載し，合計 68 ノードで構成される。OpenMP，MPI を利用したノード内の並列処理は 128 並列まで可能で，ノードあたりの理論演算性能は 4.096TFLOPS(倍精度) である。サブシステム全体の理論演算性能は，278.5TFLOPS(倍精度)，総主記憶容量は 17TB，総メモリバンド幅は 27.2TB/s となる。サブシステム AOBA-B を構成する LX 406Rz-2 の 4 ノードシャーシと，それに搭載される AMD EYPC プロセッサの外観を図 3 に示す。

サブシステム AOBA-B は，ベクトル演算に不向きなプログラムや，商用アプリケーションや OSS の高速な実行を目的として導入された。

■プログラミング言語　Fortran/C/C++ コンパイラとして，AMD Optimizing C/C++ Compiler(AOCC)，GNU Compiler Collection(GCC) および，Intel Compiler(MKL，Intel MPI 含む) が利用できる。AOCC と GCC は OpenMPI ライブラリによる分散メモリ並列プログラムをコンパイル可能である。科学技術計算ライブラリとして，EPYC プロセッサに最適化された AMD Optimizing CPU Libraries (AOCL) が利用できる。Intel Compiler は旧システムからのソースコード移行用として，ライセンス数限定で利用できる。

■アプリケーション　商用アプリケーションとして Gaussian16 および VASP と，東北大学内利用者向けに MATLAB および Mathematica が利用できる。OSS として OpenFOAM および Quantum Espresso がインストールされている。なお，VASP の利用には利用者が契約したライセンスの提示が必要である。

1.3　ストレージシステム

ユーザのホーム領域として，高速アクセスかつ高密度ストレージである DDN SFA7990XE(DDN 社製) を導入した。図 4 にストレージシステムを示す。上図がストレージのコントローラー部で，下図がスピンドルの格納部である。SFA7990XE 上に ScaTeFS(日本電気株式会社製) の IO サーバを構築し，高速アクセス性能と IO サーバの耐障害性を確保した。ホーム領域は RAID6 で構成され，実効容量は 2PB である。

2　利用者環境

2.1　ログイン認証方式

図 5 に利用者向けサーバを示す。今回のシステムでは利用者の利便性とセキュリティの向上を考慮し，ログインサーバとフロントエンドサーバの 2 段構成としている。ログインサーバは外部ネットワークに公開され，緊急に対応が必要なセキュリティインシデントに迅速に対処可能としている。フロントエンドサーバはログインサーバからのみアクセス可能とし，利用者はフロントエンドサーバ上でソースコードのコンパイルやリクエストの投入を行う。

利用者の公開鍵は旧システムで利用していたものを引き続き利用できるので，ローカル PC に保存済みの秘密鍵とパスフレーズによるログインが可能である。

新規利用者は本センターウェブサイト上に提供される，公開暗号鍵ペア作成機能を用いてログインのための秘密鍵を作成する。

また，利用者のローカル PC とストレージシステム間で大規模なデータ転送を行うために，データ転送サーバも同じ鍵ペアを用いてログインと利用が可能である。

なお，GSI-SSH 認証によるログインはログインサーバを介さず，HPCI 利用者用のフロントエンドサーバから利用する。

2.2　プロジェクトコード

本センターではバッチリクエストの処理に NEC Network Queuing System V (以下，NQSV) を採用している。NQSV のジョブアカウント機能によって，ユーザが異なるプロジェクトで計算機資源を利用する際に，リクエスト単位の課金と予算管理を行うことが出来る。新システムでも引き続きこのプロジェクト

図 4　ストレージシステム

図 5　利用者向けサーバ

コードの機能を利用し，1 つの利用者番号で複数の請求先の使い分けを可能としている。プロジェクトコードと請求先の関係を図 6 に示す。

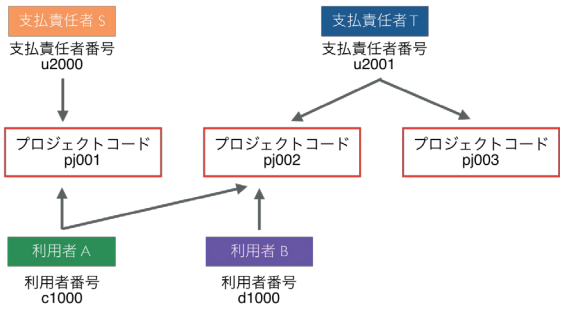

図 6　プロジェクトコード

■複数の請求先の利用　近年では研究費での利用に加え，課題採択形式で利用されるケースが増加している。プロジェクトコードを用いることにより，利用者 A は 1 つの利用者番号 (c1000) から請求先の異なる複数のプロジェクトコード (pj001, pj002) を使い分けることが可能である。バッチリクエスト投入の際に NQSV のジョブアカウント機能を用い，請求先としたいプロジェクトコードを指定することで利用者が複数の請求先を使い分けることが可能となる。また，支払責任者が複数のプロジェクトコード (pj002, pj003) を保有することも可能である。

■課題利用期間とプロジェクトコード　採択課題の利用期間が終了したものについては，該当するプロジェクトコードを無効にすることで利用者はリクエストを投入不可となる。また，利用可能な課題が追加された場合は，利用者番号に対してプロジェクトコードを追加設定することでリクエストの投入が可能となるので，それまで利用していた環境を引き続き利用することが可能である。

3　利用負担金と実行形態

3.1　利用負担金

大規模科学計算システムの利用負担金表を表 1 に示す。この表は大学・学術利用に適用され，民間企業利用は成果公開型の場合で本表記載の金額の 2 倍，成果非公開型の場合で本表記載の金額の 4 倍となる。

課金対象時間は各リクエストの利用 VE 数または利用ノード数と経過時間の積を秒単位で記録し，半年間の請求期毎に合算した後に時間単位に切り上げたものである。この課金対象時間に負担額を乗じた金額が請求金額となる。また，負担金を前払いすることで一定の課金対象時間まで利用することの出来る，定額制の導入も行った。定額制による利用は，年度途中に負担金を追加することによる利用継続も可能である。一定数の VE またはノードを研究グループで占有して利用する，占有利用も引き続き利用可能である。

3.2　サブシステム AOBA-A の実行形態

サブシステム AOBA-A で実行する場合の実行形態を表 2 に示す。今回導入したシステムでは，計算資源の効率的な利用と，リクエストの待ち時間短縮など利用者の利便性を考慮して，VH を共有する実行形態および VH を共有しない実行形態を利用者が選択できるようにした。それぞれの実行形態の例を図 7 に示す。どちらの実行形態も利用者の利便性を考慮し，最大経過時間を既定値 72 時間，最大値 720 時間として長時間のリクエスト実行を可能とした。

表 1　基本利用負担金【大学・学術利用】

区分	項目	利用形態	負担額及び課金対象時間
演算負担経費	スーパーコンピュータ	共有(無料)	利用 VE 数 1(実行数，経過時間の制限有) 無料
		共有(従量)	課金対象時間 ＝ (利用 VE 数÷ 8 を切り上げた数) ×経過時間 (秒) 課金対象時間 1 時間につき　125 円
		共有(定額)	負担額 10 万円につき課金対象時間 800 時間分使用可能
		占有	利用 VE 数 8　利用期間 3 ヶ月につき　270,000 円
	並列コンピュータ	共有(従量)	課金対象時間＝利用ノード数×経過時間 (秒) 課金対象時間 1 時間につき　22 円
		共有(定額)	負担額 10 万円につき課金対象時間 4,600 時間分使用可能
		占有	利用ノード数 1　利用期間 3 ヶ月につき　47,000 円
ファイル負担経費		共有	5TB まで無料，追加容量 1TB につき年額　3,000 円
		占有	10TB まで無料，追加容量 1TB につき年額　3,000 円
出力負担経費	大判プリンタによるカラープリント		フォト光沢用紙 1 枚につき　600 円 クロス紙 1 枚につき　1,200 円

備考

1. 負担額が無料となるのは専用のキューで実行されたものとし，制限時間を超えた場合は強制終了する。
2. 演算負担経費の課金対象時間については半期毎 (4 月から 9 月及び 10 月から 3 月) に合計し，1 時間未満を切上げて負担金を請求する。
3. 演算負担経費について定額制を選択した場合はスーパーコンピュータ及び並列コンピュータを課金対象時間の範囲内で共用できる。
4. 占有利用期間は年度を超えないものとし，期間中に障害，メンテナンス作業が発生した場合においても，原則利用期間の延長はしない。
5. ファイル負担経費については申請日から当該年度末までの料金とする。運用期間が 1 年に満たない場合は，月割りをもって計算した額とする。

表 2　サブシステム AOBA-A の実行形態

投入キュー名	利用可能 VE 数	最大メモリ	リクエストの実行形態	最大経過時間
sxf	1	48GB	無料の 1VE リクエスト (VH を共用する)	最大値 1 時間
sx	1	48GB	1VE リクエスト (VH を共用する)	既定値 72 時間 最大値 720 時間
sx	2〜256	12TB	8VE 単位で確保 (VH を共用しない)	
sxmix	2〜8	384GB	1VE 単位で確保 (VH を共用する)	
占有利用	契約 VE 数	48GB ×契約 VE 数	VE および VH を占有する	最大値 720 時間

表 3　サブシステム AOBA-B の実行形態

投入キュー名	利用可能ノード数	最大メモリ	リクエストの実行形態	実行時間制限
lx	1〜16	12TB	1 ノード単位で確保 (ノードを共用しない)	既定値 72 時間 最大値 720 時間
占有利用	契約ノード数	256GB ×契約ノード数	ノードを占有する	最大値 720 時間

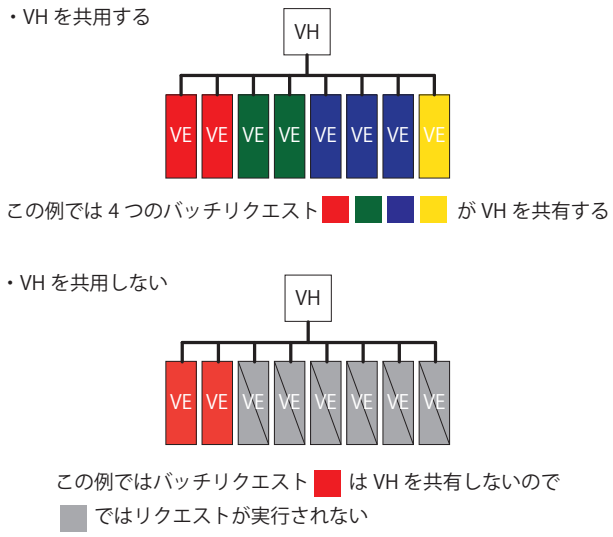

図 7　プロジェクトコード

■VH を共用する　投入したリクエストは他のリクエストと VH を共用して実行される。1VE を利用すると指定したリクエストは，VH を共有して実行される。また例として，2 個の VE を使うと指定したリクエストを sxmix キューに投入した場合，他 6 個の VE で別のリクエストが実行されることがある。利用する VE 数が 2〜8 個の場合，sxmix キューを選択すると実行に必要な VE 数が確保されやすく，リクエスト混雑時にも待ち時間を短縮することが出来る。

■VH を共用しない　投入したリクエストは他のリクエストと VH を共用しないで実行される。利用する VE 数を 2〜7 個と指定をしたリクエストを sx キューに投入した場合は，8 個の VE と 1 個の VH を確保する。他のリクエストと VH を共用しないため他リクエストのストレージへの I/O や VH 間通信の影響を受けにくく，演算時間のバラツキが少なくなる。

3.3　サブシステム AOBA-B の実行形態

サブシステム AOBA-B で実行する場合の実行形態を表 3 に示す。共有利用は lx キューのみであり，利用者は利用するノード数を 1 ノード単位で指定してリクエストを投入する。サブシステム AOBA-B でも利用者の利便性を考慮し，最大経過時間を既定値 72 時間，最大値 720 時間として長時間のリクエスト実行を可能とした。

4　高速化支援活動

本センターでは 1997 年より、ユーザアプリケーションの高精度化、大規模化の支援を目的とした高速化支援活動を，また 1999 年より共同研究制度を実施している。利用者、計算機科学を専門とするセンター教員、技術職員、およびベンダー技術者が連携してアプリケーションの高速化に取り組んでいる。

前スーパーコンピュータシステムの SX-ACE を運用した 5 年間においては，合計で 30 件の高速化支援を行った。単体性能では平均約 16.7 倍の性能向上を，並列性能では約 2.4 倍の性能向上を得ることが出来た。

図 8 に 1999 年から本センターで取り組んでいるセンター独自の共同研究、学際大規模情報基盤共同利用・共同研究拠点 (JHPCN) 課題および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) 課題採択数の推移を示す。本センター独自の共同研究は恒常的に年 10 課題ほど実施されていることに加え、近年では JHPCN、HPCI を介した共同研究数が増加している。これは、センターの共同研究を通してユーザアプリケーションが高度化・大規模化し、JHPCN、HPCI 採択課題へとステップアップしており、我々の継続的な高速化支援活動が一定の成果を上げていると言える。

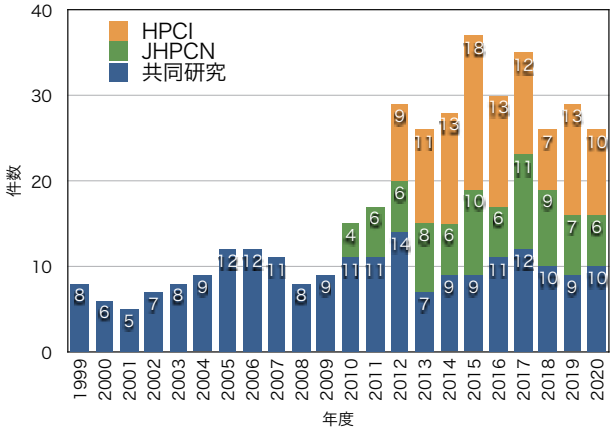

図 8　課題採択件数

5　おわりに

本稿では 2020 年 10 月に運用を開始した，サイバーサイエンスセンターのスーパーコンピュータ AOBA について紹介した。研究室のサーバでは実行できなかったプログラムやアイデアを実現する研究の強力なツールとして，最新鋭のスーパーコンピュータ AOBA をご活用いただければ幸いである。各システムの利用法の詳細，本センターからのお知らせ，問い合わせ，利用相談，高速化の依頼方法などついては本センターのウェブサイト[*1]を参照いただきたい。

[*1] https://www.ss.cc.tohoku.ac.jp/