

国立環境研究所 研究用計算基盤概要

国立研究開発法人国立環境研究所
環境情報部

1. 研究用計算基盤構成

1.1 ハードウェア構成

国立研究開発法人国立環境研究所（以下、NIESという。）は、2026（令和8）年3月より本研究用計算基盤ERFa（NIES Environmental Research computing Facility）の運用を開始する。本研究用計算基盤は、ベクトル計算機（NEC製SX-Aurora TSUBASA C403-8）、仮想化基盤（HPE製 ProLiant DL360 Gen11、ProLiant DL380a Gen11）、大容量ファイルシステム（DDN製 ES400NVX2）を中心に構成し、高速ネットワーク接続にはInfiniBandとEthernet（200GbE）を採用している。また、2031（令和13）年度までの運用を予定している。

NIESの大規模計算機システムとしては、1991年に導入されたNEC製ベクトル演算型計算機 SX-3から数えて7代目となる。

図1：主な構成機器



1.2 全体構成

2025年11月まで運用していた旧システムは、ベクトル計算機とスカラ計算機で構成されていた。本研究用計算基盤でもベクトル計算機、スカラ計算機については、利用方法は基本的に変わらない。また、本研究用計算基盤では新たにプロジェクトVM（Virtual Machine：仮想マシン）を利用できるようになった。

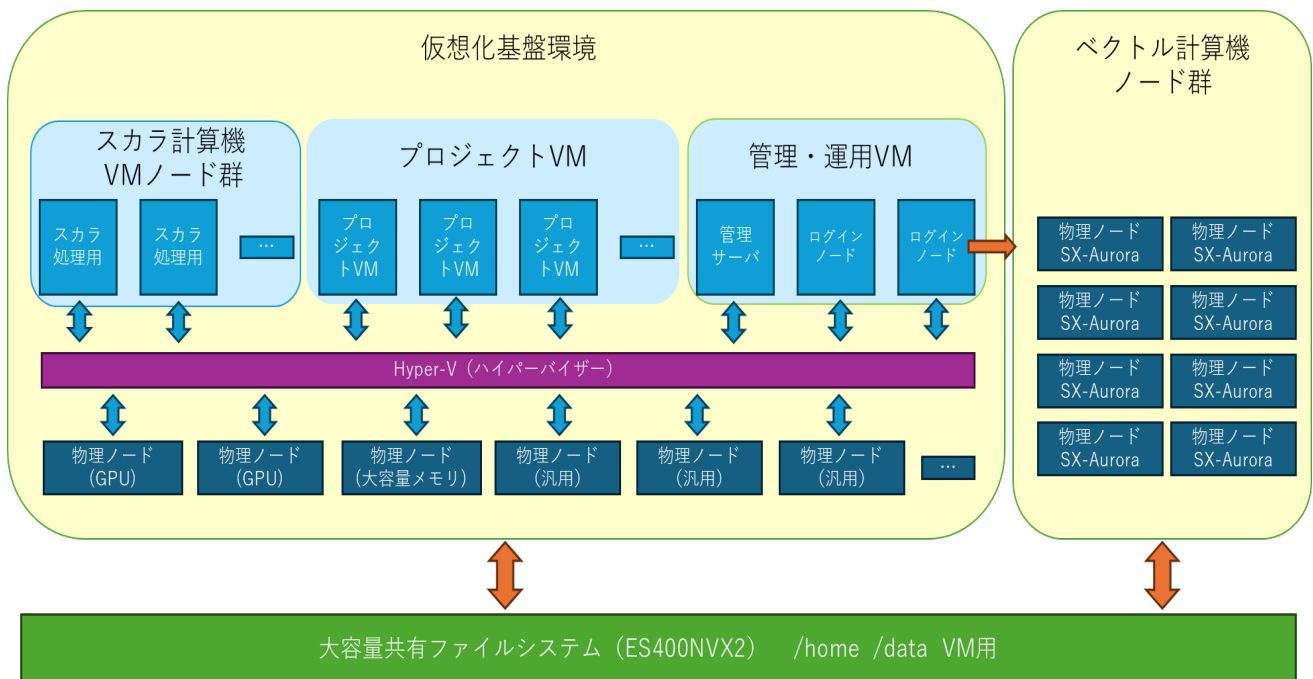
仮想化基盤環境では、ハードウェアをそのまま使用するのではなく、複数の物理コンピュータをとりまとめて仮想化基盤とし、その上でソフトウェア的に作られたVMを稼働させる。ハードウェアの一部が故障しても自動的に正常なノードにVMを移動させて動作を継続する機能が提供される。また、プロジェクトVMの管理者は、割り当てられたリソースの範囲内で、VMの作成・削除・起動・停止などを行うことができる。

本研究用計算基盤では、このVMを積極的に活用する。

従来のスカラ計算機の代わりとして、汎用CPUが搭載された仮想化基盤上にスカラ計算ノード群のVMを構築する。この汎用CPUが搭載された基盤上のVMは、申請したユーザも使用することができ、プロジェクトVM（個別サーバVM）と呼ぶ。加えて、GPUが搭載された仮想化基盤と大容量メモリが搭載された仮想化基盤が提供されるようになり、これら基盤上に構築されたプロジェクトVMではGPUや大容量メモリを活用した演算も可能になった。

従来、研究室で個別に購入していたサーバの代わりにプロジェクトVMを利用できる。ユーザからは1台のVMが1台の物理サーバのように扱えるため、プロジェクトVMはOSから先の管理責任だけが必要になる「個人サーバ」「グループ用サーバ」のように考えてもらってよい。また、プロジェクトVMごとに独立したサーバ環境となるため、VMの中身は用途に合わせて自由に構築可能で、別途、所内の審査・承認を経て、構築したWebサーバを外部に公開することもできるようになった。

図2：イメージ図



※1：水色のエリアがソフトウェア的に実装するノード（VM）。

※2：GPUの利用に関して、厳密には、VMから仮想化した環境をすり抜けて物理GPUを使用するGPUパススルー方式となる。

1.3 ベクトル計算機

ベクトル計算機としては旧システムとほぼ同等の性能であり、使い方も従来どおりログインサーバにログインしてバッチジョブを投入する方法になる。

ベクトル計算機はベクトル処理を行うVE（ベクトルエンジン）と、VEの制御やOS機能を担うVH（ベクトルホスト）で構成され、地球環境シミュレーションなどの大規模計算に必要なベクトル演算能力と大容量メモリを有する。

本研究用計算基盤では、1つのVHは8つのVEを、1つのVEは16のコアを、研究用計算基盤全体で128のVEを持つ。（図3）ハードウェア構成として1つのVH単位で1台のノードとなる。

演算ピーク性能628TFLOPS、総メモリ容量12.0TiBである。（表1）

図3：ベクトル計算機 SX-Aurora TSUBASA C403-8 の VEとノード



表1：ベクトル計算機の性能と前システムの比較

		旧システム SX-Aurora TSUBASA/A511-64	本システム SX-Aurora TSUBASA/C403-8
プロセッサ (VEあたり)	コア数	8	16
	FLOPS	304GFLOPS (コア) 2,432TFLOPS (VE)	307GFLOPS (コア) 4,915TFLOPS (VE)
ノード (VHあたり)	CPU数	1	1
	VE数	8	8
	FLOPS	19.456TFLOPS	39.28TFLOPS
	メモリサイズ	384GiB	768GiB
システム 合計	ノード数	32	16
	VE数	256	128
	総コア数	2048	2048
	FLOPS	622.8TFLOPS	628.48TFLOPS
	メモリサイズ	12TiB	12TiB
OS		CentOS Linux 7	Rocky Linux 10
ジョブ管理ソフト		NEC NQSV	NEC NQSV

旧システムに比べて、VH数とVE数は半分に減ったが、1つのVEあたりのコア数が2倍となったため、ベクトル計算機としては旧システムとほぼ同等の性能である。

1.4 仮想化基盤

HPEのProLiantシリーズを中心とするノード群で構成し、仮想化基盤環境を構築している。この仮想化基盤環境で、管理用サーバ、ログインノード、スカラ計算機ノード群、プロジェクトVMなどをVM（仮想マシン）として構築して運用する。

表2：仮想化基盤の物理ノード構成

機種・ノード数		CPU	GPU	メモリ	用途
HPE ProLiant DL360 Gen11 ×8ノード 62.2TFLOPS	1ノード 8ノード 合計	128コア 1,024コア	なし なし	256GiB 2,048GiB	ログインノード スカラ計算機ノード群、管理用VM プロジェクトVM (標準的な計算やアプリケーション、汎用)
HPE ProLiant DL380a Gen11 ×5ノード 638.9TFLOPS	1ノード 5ノード 合計	128コア 640コア	NVIDIA H100 NVL 94GB 4枚 20枚	1,024GiB 5,120GiB	プロジェクトVM (GPUを使用した計算)
HPE ProLiant DL360 Gen11 ×1ノード 7.7TFLOPS	1ノード	128コア	なし	2,048GiB (2TB)	プロジェクトVM (大規模データ解析、巨大行列計算など)

※1 1つのVMのスペックが1台の物理ノードのスペックを超えることはできない。例えば、2台の物理ノードが空いているとしても、2台の物理ノードのスペックを合算した1台のVMを構成することはできない。

※2 仮想化基盤（汎用CPU/大容量メモリ/GPU）の物理ホストのCPUは Intel Xeon Platinum 8592+（2ソケット）構成。

1.5 大容量ファイルシステム

大容量ファイルシステムはDDN製 ES400NVX2というアプライアンス製品を中心にLustreファイルシステムと呼ぶ分散並列ファイルシステムを構成する。総容量は18PBであり、ユーザに提供する領域としては、主にhome領域、data領域の2つである。

ベクトル計算機ノード群、VMとも、この大容量ファイルシステムをストレージとして使用する。

なお、バックアップについては以下のとおりとする。

表3：大容量ファイルシステムのバックアップ

領域	バックアップ
home領域	1日に1回バックアップし、そのバックアップを7世代保管する。
data領域	バックアップしない

※ data領域に関しては、必要があればユーザがバックアップすること。

2. 研究用計算基盤の運用

2.1 計算機運用事務局

旧システムでは、ベクトル計算機については地球システム領域、スカラ計算機については環境情報部と、窓口・事務局が分かれていた。本研究用計算基盤の運用窓口（申請受付・アカウント発行・グループ管理、ユーザサポート等）は、すべて環境情報部 研究情報室が担当する。

2.2 基本的な使用法

本研究用計算基盤に接続して利用する基本的な使い方は以下のとおりである。

以下の表のいずれの場合も、セキュリティ強化のため、ログインノードへのパスワードによるSSHログインは禁止し、公開鍵認証+ワンタイムパスワードの2要素認証にする。

表4：接続方法等

利用方法	接続方法
ベクトル計算機の利用	ログインノードにSSH接続、ログインノードでコンパイル等の作業を行なって、NQSVMにてバッチジョブを投げる。 計算時のデータは共有ファイルシステムから読み書きする。
スカラ計算機の利用	ログインノードにSSH接続、ログインノードでコンパイル等の作業を行なって、NQSVMにてバッチジョブを投げる。 計算時のデータは共有ファイルシステムから読み書きする。
プロジェクトVMの利用	ログインノードにSSH接続、そこから自分のプロジェクトVMにログインして使用する。 プロジェクトVMのローカルストレージも、共有ファイルシステムも、どちらも読み書きできる。 プロジェクトVM内は、必要なソフトウェアをインストールするなど、自由に構築、使用できる。

2.3 ユーザアカウントについて

利用者アカウントの管理方法、命名規則等についても、昨今の情報セキュリティ、個人情報保護、研究インテグリティに即したルールとする。

- 一人がひとつのアカウントを持ち、共有アカウントは作らない。（アカウントの貸し借りは以前から禁止）
- アカウント名から個人を特定できない通番のアカウント名とする。
- データは利用グループ単位で管理し、退職者のデータ引継ぎはグループ内で対処する。
- 割り当て領域のデータは、割り当てられたグループでしかアクセスできない。（他のグループの領域にはアクセスできない）
- データ共有には別途共有領域の設定ができる。
- 課題（グループ）単位で参加者を管理し、代表者（または受入責任者）がメンバー管理を行い、毎年度メンバーの棚卸を行う。（所外利用者を含む）

2.4 データ管理

- home領域、data領域 は利用グループ（課題）単位でクォータ（割り当て容量制限）管理を行う。
- グループ間のデータ共有が必要な場合は、共有用のグループ（共有領域）を別途設定できる。
- 共有領域のデータを参照しつつ、別グループの計算資源で実行することは可能だが、実行結果ファイルの所有者・グループやアクセス権に留意すること。

2.5 ジョブ管理

- ベクトル計算ノード群およびスカラ計算ノード群のバッチ実行は、ジョブ管理システム（NQSVM）により行う。
- ログインノード上での長時間計算は行わず、原則としてバッチジョブを投入する。

2.6 プロジェクトVMの運用と責任分界

- プロジェクトVMは、利用者がOS上の管理（ソフト導入、設定、セキュリティ対策等）を行う。
- 外部公開を伴う利用（公開VM）は、別途所内の手続き・審査の対象となる。
- VMの外部接続は、所定のネットワーク設計（NAT等）に基づき提供する。また、外部向け通信に関しては、セキュリティ上の視点から、制約やログ取得がある。

3. 利用支援

3.1 利用者向け講習会

研究用計算基盤の利用者向け講習会を開催する。

3.2 プログラム開発支援

科学技術ソフトウェアの開発・利用のための、コンパイラ利用、プログラミング、ジョブ実行スクリプト作成等に関する技術的な支援、情報提供を行う。

3.3 ソフトウェアの移植及び性能向上に関する技術支援

他のコンピュータシステム等で開発されたソフトウェアについて、本研究用計算基盤（ERFa）で実行するための移植作業を支援する。また、ベクトル化・並列化等による計算処理性能向上に関する技術的な支援を行う。

以上