

環境分野研究のための データ基盤・解析プラットフォーム構築

八代 尚

国立研究開発法人国立環境研究所

地球システム領域 衛星観測センター 主任研究員

環境情報部 研究情報室 室長代行（兼）

Gfarmシンポジウム2025, 秋葉原, 2025/9/19

自己紹介

- 八代 尚 (Hisashi Yashiro)

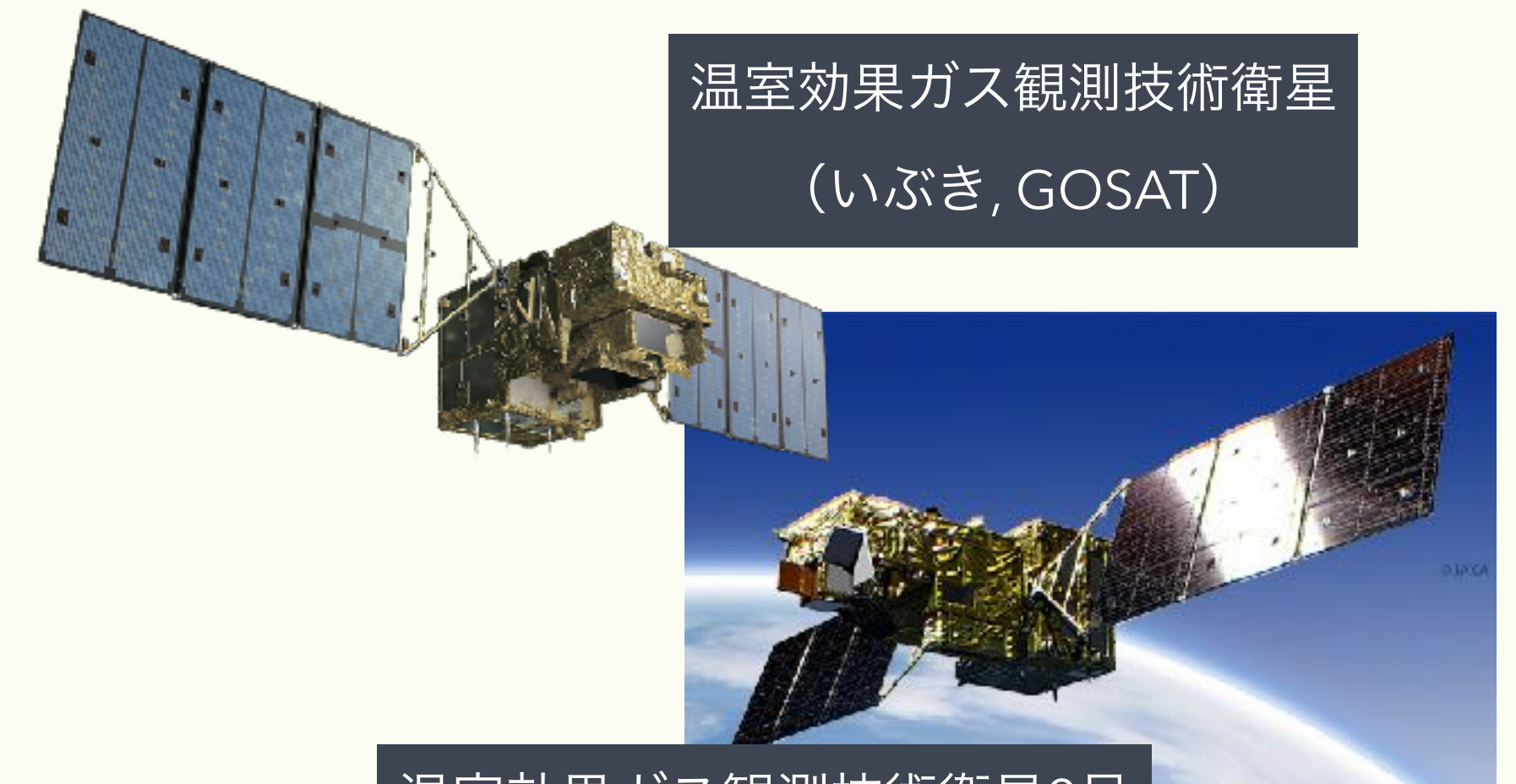
- 1978年生：青森県弘前市
- 専門：大気化学、高性能計算、大気科学
- 海洋研究開発機構→理研R-CCS(AICS)→国立環境研究所
- 次世代気候モデル研究開発を経て、温室効果ガス観測衛星「GOSAT」プロジェクトが現在の本務



スバルバル島ニーオルスン基地(北緯79度)



JAXA三陸大気球観測所（岩手県大船渡市）

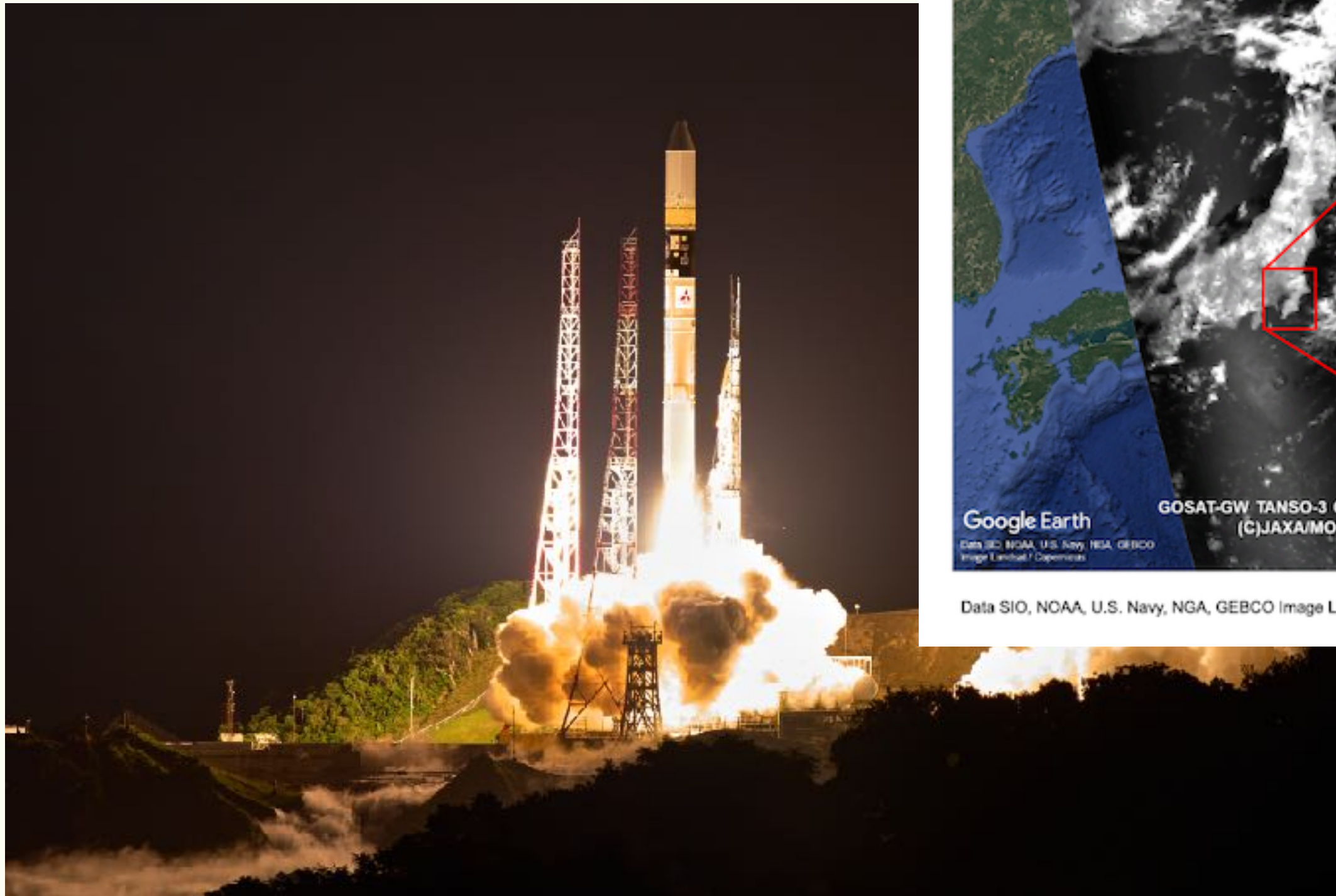


温室効果ガス観測技術衛星
(いぶき, GOSAT)

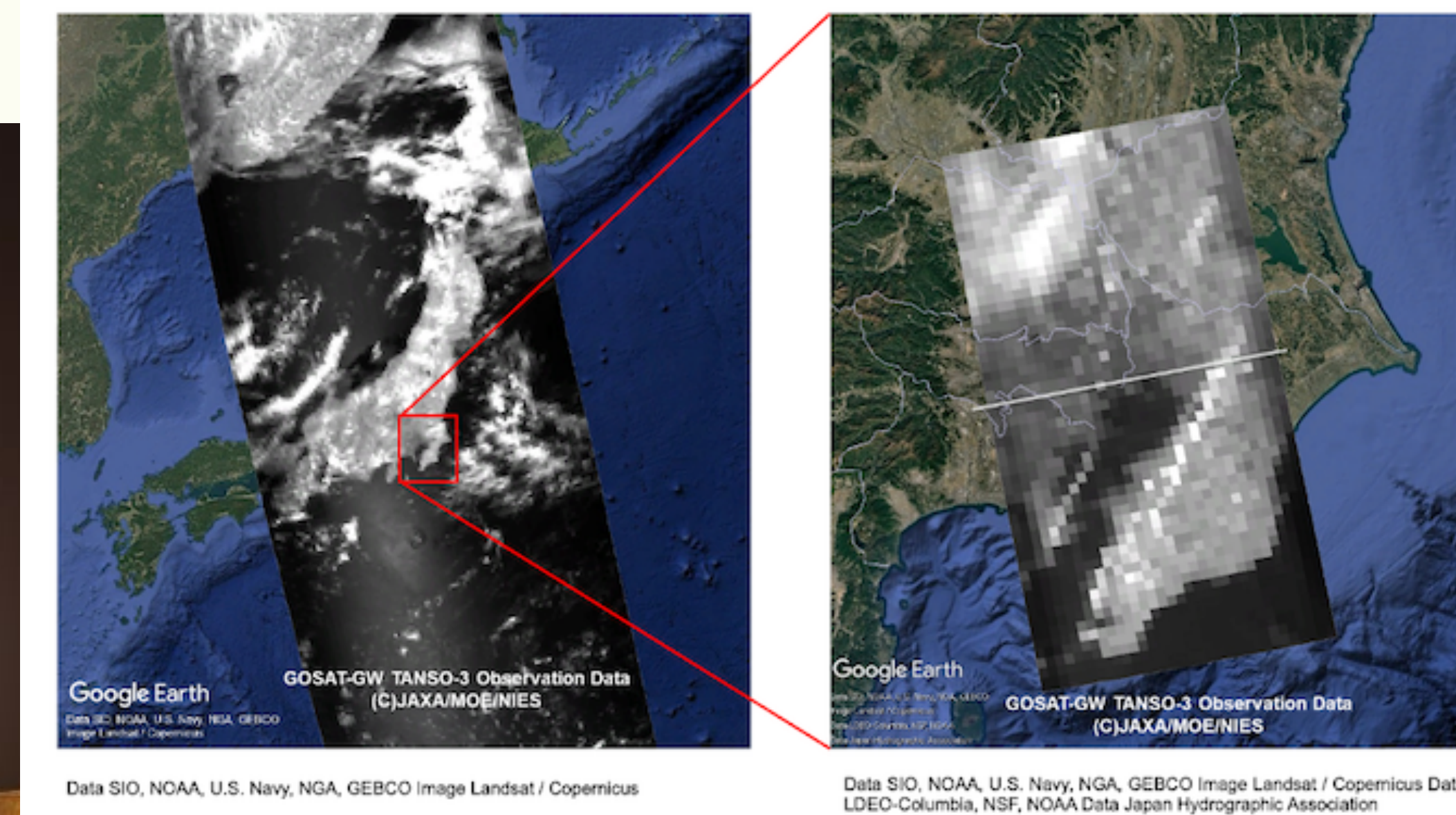
温室効果ガス観測技術衛星2号
(いぶき2号, GOSAT-2)

温室効果ガス・水循環観測技術衛星「いぶきGW」(GOSAT-GW)

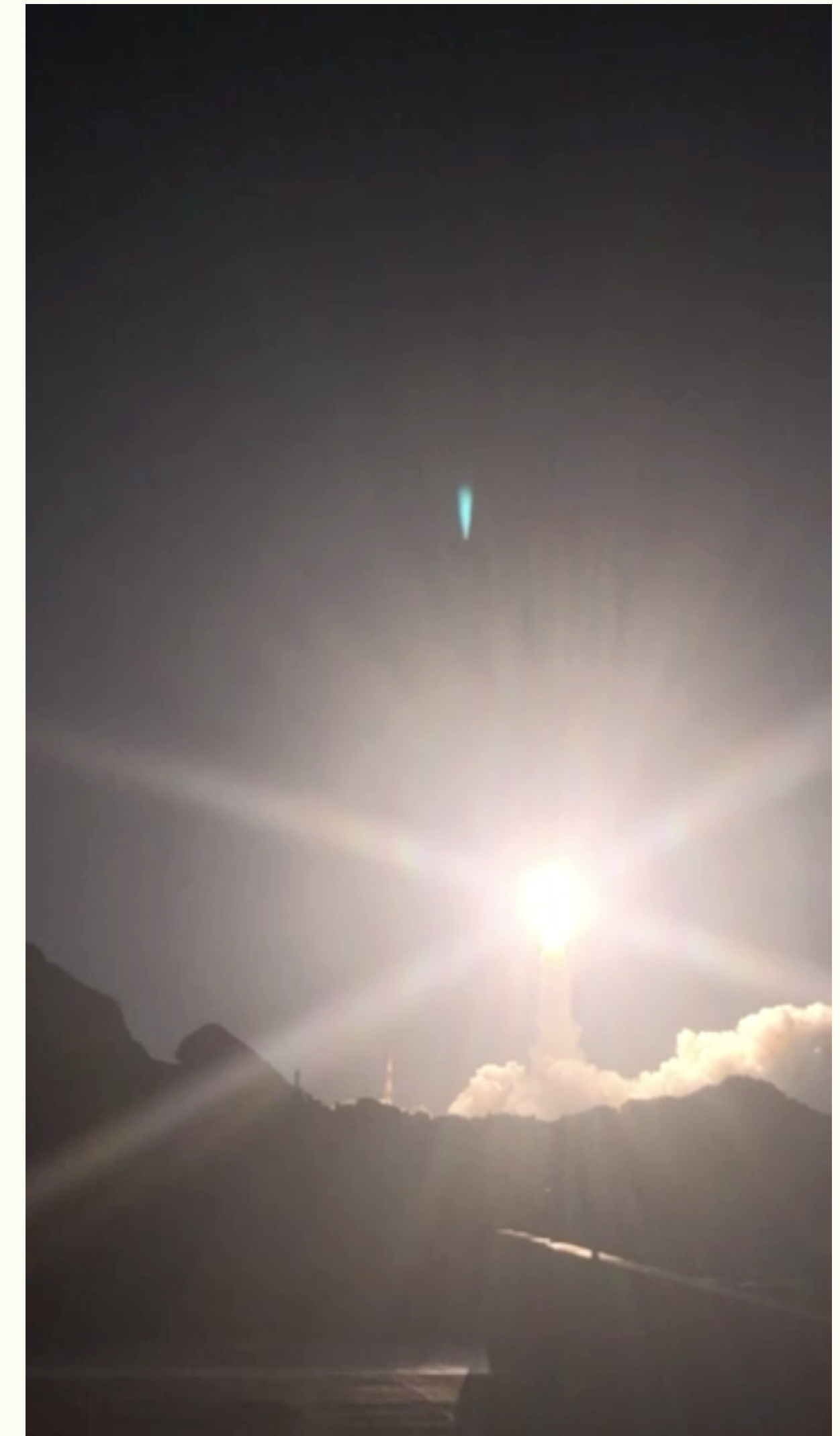
- 2025年6月29日1時33分（日本標準時）
JAXA種子島宇宙センターからH-IIAロケット50号機（最終号機）にて打ち上げ
- 10月まで衛星初期チェック期間、7月14-20日に初観測を実施。



国立環境研究所 成田氏撮影



国環研プレスリリースより



GOSAT-GW衛星データ処理のための地上システム

- 基幹システム：G3DPS

- JAXAから（観測データ）、気象庁から（気象場）、mdxから（温室効果ガス・エアロゾル予測結果）、それぞれデータを日々受領
- 利用者へのプロダクト提供（2026年4月頃開始予定）
- HCI (Nutanix) + Object Storage (Scality RING)：プライベートクラウド

- HPCシステム：GOCF

- センサデータから温室効果ガス濃度を算出する計算を担当
- 200+400TFLOPS, CPU-Only (Intel S.R.), Infiniband HDR
- 1PB SSD Lustre

写真：奥がG3DPS、手前がGOCF
筑波大計算科学研究センター別棟に設置



National Institute for Environmental Studies (NIES)

1974年に国立公害研究所として設立



Supercomputer system in NIES



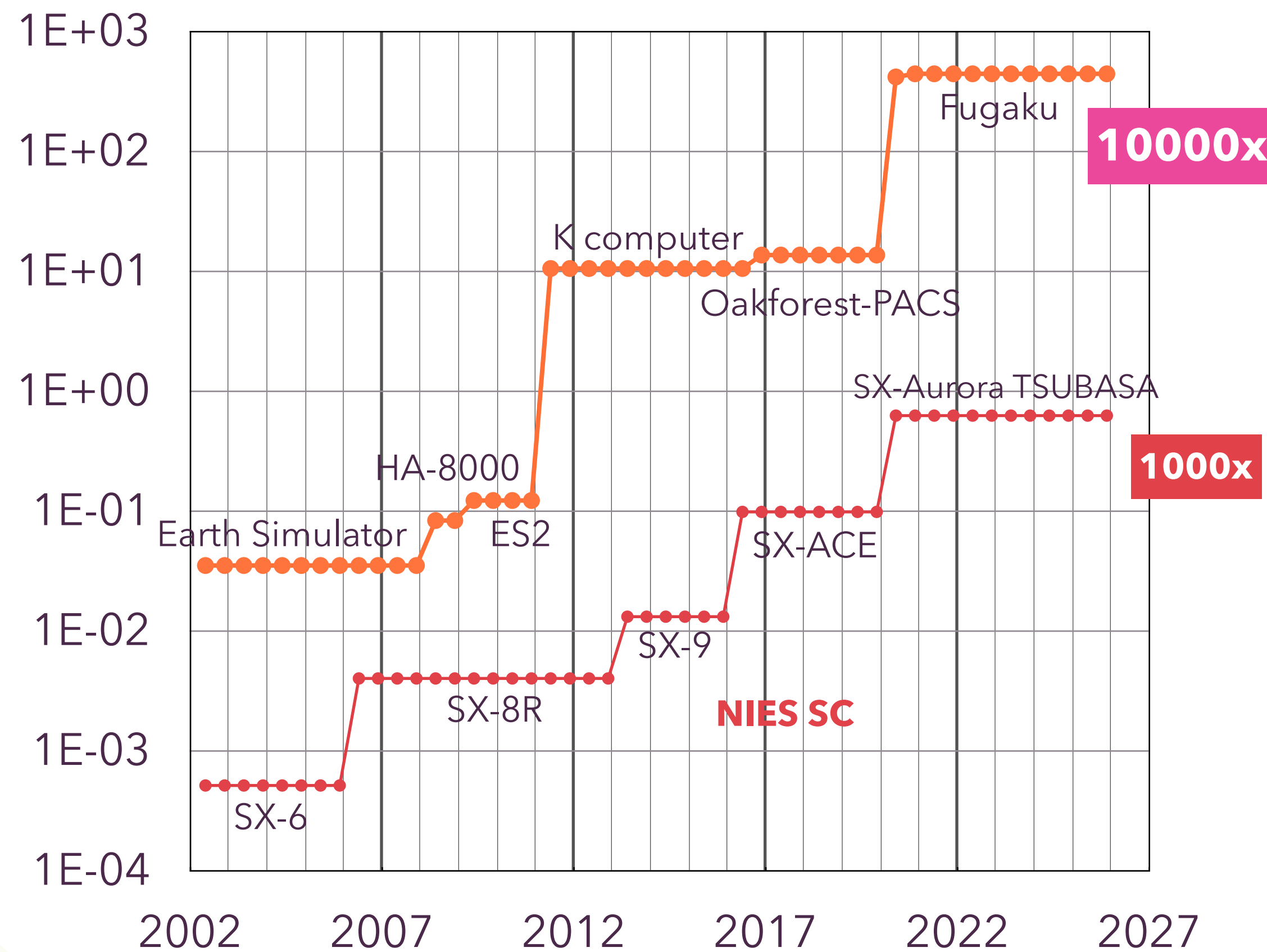
NEC SX Aurora Tsubasa A511-64

	スーパー コンピュータ の機種	型式	導入年度	処理速度	メモリ 転送速度	計算 ノード数	稼働期間
第1号機	NEC SX-3	type 14	1991	5.5GFLOPS*	12.8GB/s	1	1992.2~ 1997.2
第2号機	NEC SX-4	type 32	1996	64GFLOPS	512GB/s	1	1997.2~ 2002.2
第3号機	NEC SX-6	type 64M8	2001	512GFLOPS	2.0TB/s	8	2002.2~ 2007.3
第4号機	NEC SX-8R	type 128M16	2006	4.1TFLOPS	8.2TB/s	16	2007.3~ 2013.6
第5号機	NEC SX-9	type A(ECO)	2013	13.1TFLOPS	16.4TB/s	8	2013.6~ 2015.3
第6号機	NEC SX-ACE		2015	98.3TFLOPS	98.3TB/s	384	2015.6~ 2019.11
第7号機	NEC SX-Aurora Tsubasa	type A511-64	2019	622.8TFLOPS	345.6TB/s	256(VE)	2020.3~ 2026.2 (予定)

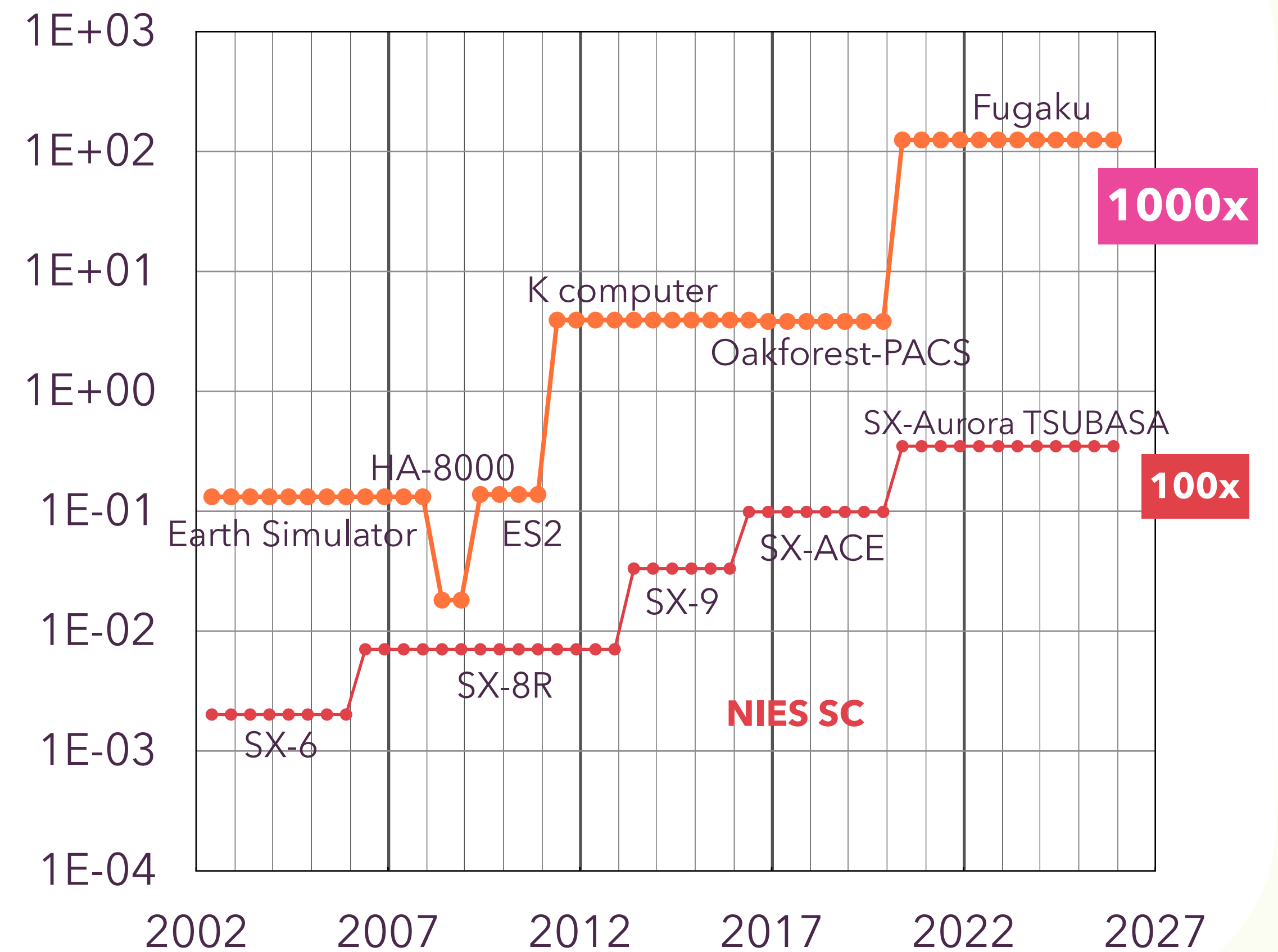
※コンピュータの処理速度をあらわす単位の一つで、1秒間に実行できる浮動小数点数演算の回数。

Performance trends of NIES supercomputer

LINPACK Performance [PFLOPS]



Total Memory Throughput [PB/s]



7th generation of the NIES supercomputer (2019-2025)

623TFLOPS

ベクトル処理用計算機 NEC SX-Aurora TSUBASA

ホスト名 : scv01~scv32
OS : CentOS 7

- 総理論演算性能(VE)
622.8TFLOPS
- 総主記憶容量(VE)
12TiB
- 総メモリバンド幅 (VE)
345.6TB/s



256 VE

86TFLOPS

スカラー処理用計算機 HPE Apollo 2000

ホスト名 : scs01~scs28
OS : Red Hat Enterprise Linux 7

- 総理論演算性能
86.0TFLOPS
- CPUコア数
1120コア
- 総主記憶容量
5.25TiB

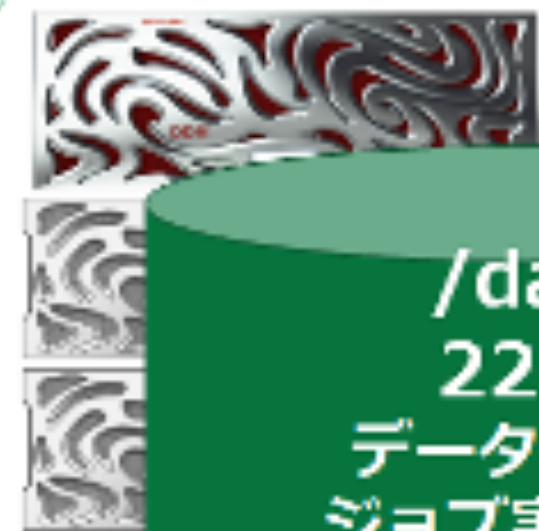


7式 (28ノード)

所内NW

高速ネットワーク
(InfiniBand EDR)

コンピュータシステム接続用
ネットワーク (Ethernet)



/data
22PB
データ保管兼
ジョブ実行領域



/home
62TB
ホーム領域

大容量ファイルシステム

22PByte

フロントエンドサーバ HPE ProLiant DL360 Gen10



ベクトル用 (2ノード)

ホスト名 : scfrv01, scfrv02



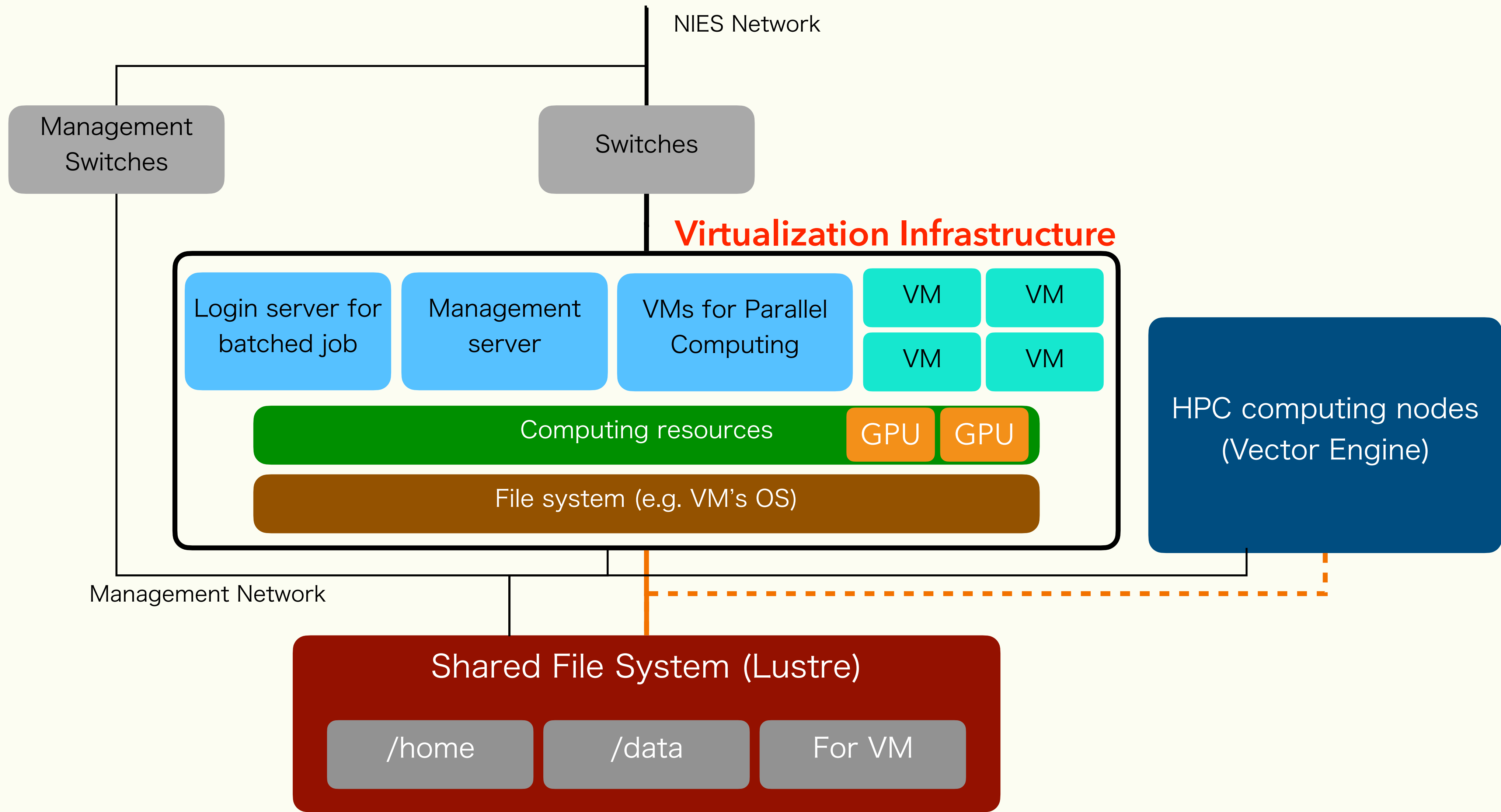
スカラー用 (2ノード)

ホスト名 : scfrs01, scfrs02

OS : Red Hat Enterprise Linux 7
ノードあたりのCPUコア数 : 40
ノードあたりのメモリ容量 : 192GiB

~40 active users (vector)
~50 active users (scalar)

8th generation of the NIES supercomputer (2026-)



「環境研究共創拠点」構想（NIES Environmental Research Hub）

（仮称）環境研究共創拠点NIES Environmental Research Hubの将来構想



環境データ先端研究プラットフォーム

産・官・学の環境研究コミュニティをデータで繋ぐ

ビッグデータ・AIを含む、先端的な研究手法を支える大規模数値計算・解析基盤を運用・提供

環境情報システム基盤

貴重な環境研究データをインフラ改革により守り支える

大規模化する環境データを効率的に収集し、セキュリティに配慮して適切に保管・管理

環境情報流通データベース

産・官・学/一般ユーザへ環境データを届ける

環境データを産官学それぞれに使いやすい形で提供し、国内外での環境データリポジトリと連携して環境データ活用型社会を実現



【図の出典】環境省（2024） 新たな「環境研究・環境技術開発の推進戦略」の概要 p.12

「共創拠点」 計算基盤・データ基盤の狙うところ

- 計算機環境の全体最適化

- ハウジング領域に乱立する個人サーバを削減し、トータルの保守管理コストを削減
- セキュリティガバナンスのさらなる徹底（サポート切れOS・ライブラリ対応、脆弱性対応、etc.）
- 多様な計算ニーズへの対応：GPU、大容量メモリ、並列計算、Pythonライブラリパッケージの個別構築
- 似たような計算ニーズへの一括対応：プライベートリポジトリ運用、R・Jupyter環境のテンプレート公開

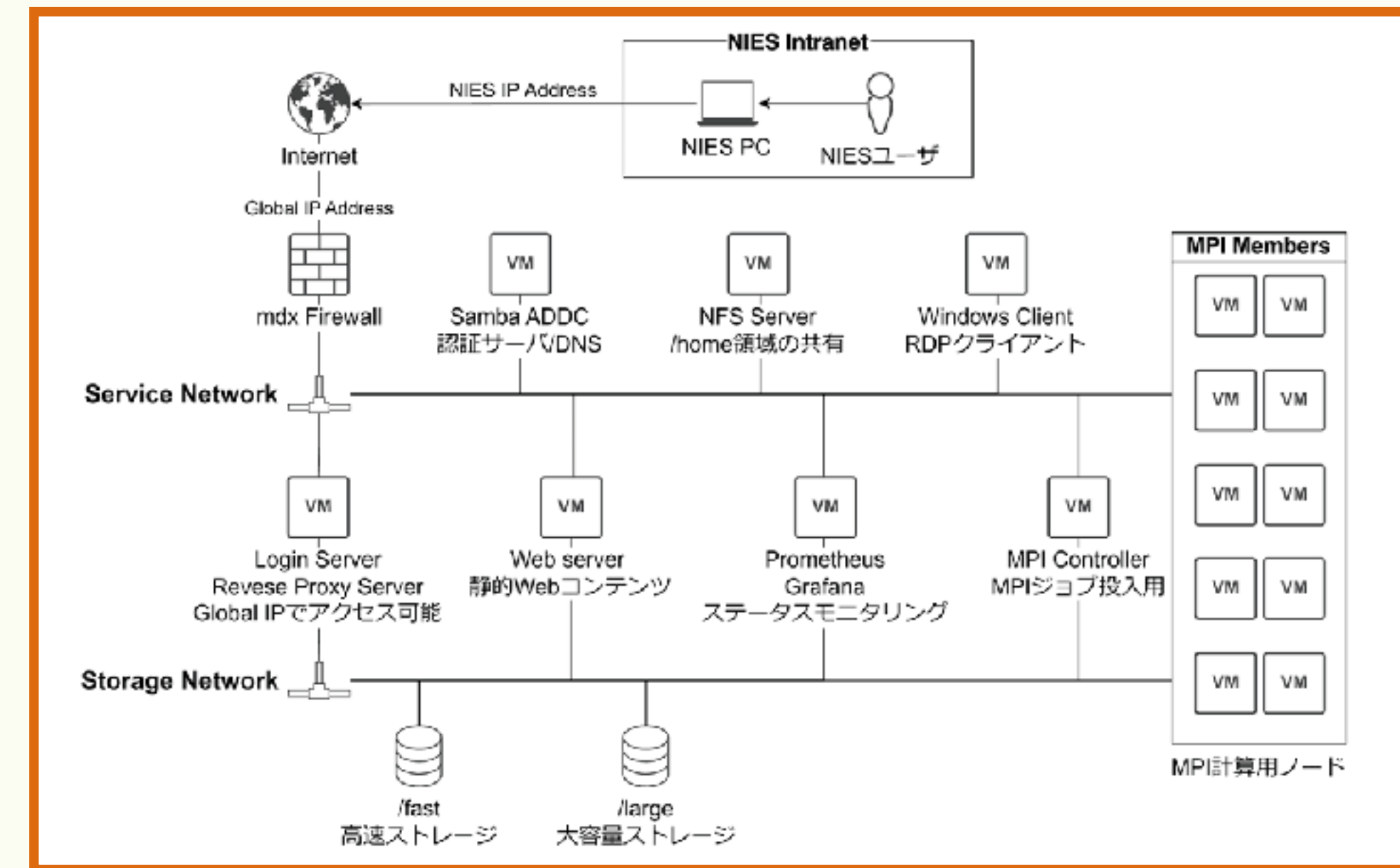
- 環境データの集約と利活用促進

- 各研究課題に仮想マシンを提供し、解析プラットフォームとして利用してもらう
- 「その場解析」を促進するための環境づくり
- 研究データ蓄積、管理、共有、公開を1カ所で
- 適切な公開範囲を設定した環境データのプロジェクト間共有
- 他機関のデータ基盤と所全体として連携

mdxを用いたフィージビリティスタディ

- 仮想化基盤の多様な利用形態を検証

- 5年間での運用の中で、「共用・並列・バッチ処理」と「個別カスタマイズ・定常処理」での計算リソース利用率を調整しながら運用したい
- 2023年よりmdxを利用し、仮想化基盤上に構築した並列HPC計算ノード群の性能評価、運用ノウハウを蓄積
- NICAM-LETKFを用いた大気環境解析・予測システムを日々実行（温室効果ガス・エアロゾル予測）
- mdx利用事例4として紹介されている
(<https://mdx.jp/use-case/case4/>)
- 物理サーバのみで構築した並列計算環境と比較して、10%程度の性能低下に留まる



国環研が発信する様々な環境データ・ツール



「共創拠点」 データ基盤の狙うところ（データの公開）

トップページ > 研究紹介 > データベース/ツール

Read this page in English

データベース/ツール

▼地球環境 ▼健康・化学物質 ▼大気・水環境 ▼自然・生物 ▼ごみ・リサイクル ▼その他

地球環境

- 地球環境データベース**
地球温暖化をはじめとする地球環境問題に関する自然科学・社会科学分野の成果を公開。
リアルタイム観測データ（地上モニタリングステーションで観測している観測データ）解析に有用な（オンライン）ツール（METEX：トランジェクトリ解析、供）。
- 温室効果ガスインベントリ・オフィス（GIO）-日本の温室効果ガス排出量データ**
日本の1990年度以降の温室効果ガス排出量データ
- 温室効果ガス排出シナリオに係るデータベース**
世界中の数百種類以上の異なる温室効果ガスの排出予測を整理したデータベース
- 化石燃料燃焼による二酸化炭素排出量の地球1kmデータ(ODIAC)**
化石燃料の燃焼による人為起源二酸化炭素排出量の地球分布を1x1kmの空間解像度で示したデータを提供
- 温暖化影響モデル**
温暖化に対する高山は、気候変動に関するCJでも指摘される必要性は世界的に共す。そこで、日本選で、積雪・融雪時期のため、定点カメラモニタリングを行っ

健康・化学物質

- 化学物質データベース Webdis Plus**
化学物質安全情報提供システムや消費者公衆資料等に記載されている化学物質のリスク関連情報や分析法などを様々な方法で検索できるシステム（旧Webdis-Plusと旧EivMethocを統合してリニューアル）
- 環境リスク統合アセスメントプログラム MuSEM**
オランダの国立公衆衛生・環境保護研究所（RIVM）の化学物質評価グループが開発したUSEs（uniform system for the evaluation of substances）を基にして開発した環境リスクの統合アセスメントプログラム
- 環境多媒体モデル G-CIEMS**
環境中に放出された化学物質の一大・水・土壌・底質など媒体間での輸送や分配と点間移動を同時に推定するモデル
- 環境展望台-環境GIS**
全国や地域のさまざまな環境データや法規制の状況などについて、システム（GIS）を用いて提供できるシステム
- 環境展望台-大気汚染常時監視データ**
全国の大気汚染常時監視測定局の年間値、月間値、時間測定結果のデータをダウンロード提供
- つくば大気質モニタリングデータ**
国立環境研究所内大気モニター群で観測された大気成分の常時モニタリングデータ
- 固定発生源NOx、SOx、PM排出係数データベース（EF-JASS）-大気汚染物質に関するインベントリ分析のための基礎情報**
わが国の固定発生源に関する窒素酸化物（NOx）、硫黄酸化物（SOx）および粒子状物質（PM）の排出係数を取組したデータベース
- 東アジア域の黄砂・大気**

自然・生物

- 侵入生物データベース**
日本国内で外来生物・生態系への悪影響が懸念されている外来種について、生態学的情報を体系的に整理して提供するデータベース
- ため池の生物多様性**
兵庫県南部ため池群の研究で蓄積された生物多様性情報を地理情報システム（GIS）を用いて提供するシステム
- 野生動物遺伝資源データベース**
国立環境研究所の環境試料タイムカプセル棟および京都大学野生動物研究センターで保存している野生動物の細胞・組織・DNAのリストをデータベース化した検索システム
- 微生物系統保存施設-保存株データベース**
微生物系統保存施設で維持されている保存株の基本的な株情報、立派情報、細胞のイメージデータ等
- 鳥インフルエンザ侵入リスクマップ**
海外から鳥インフルエンザウイルスが侵入するリスクの高い国内地域を予測するマップ（行政機関対応）
- 絶滅危惧種サンプルデータベース：藻類**
微生物系統保存施設で保有している絶滅危惧藻類の保存株を検索できるシステム
- 日本長期生態学研究ネットワークWe-tacatサービス**
長期生態学研究の推進を目的とした、生態系観測データの検索サービス

ごみ・リサイクル

- 日本の一般廃棄物データベース**
1970年代からの一般廃棄物処理実態調査データや広域化情報を収録
- アジア・太平洋地域における都市廃棄物管理に関するデータベース（DAM-SAR）**
アジア・太平洋地域の開発途上国で調査された500指標11.6.1に関連する都市廃棄物データを収録したデータベース
- 製品使用年数データベース LIVES**
日本をはじめとする16の国において調査・推定された1,352の使用年数・寿命分布データを収録
- 災害廃棄物情報プラットフォーム**
災害廃棄物処理に係る情報を一括して収集することができるWebサイト

環境と社会

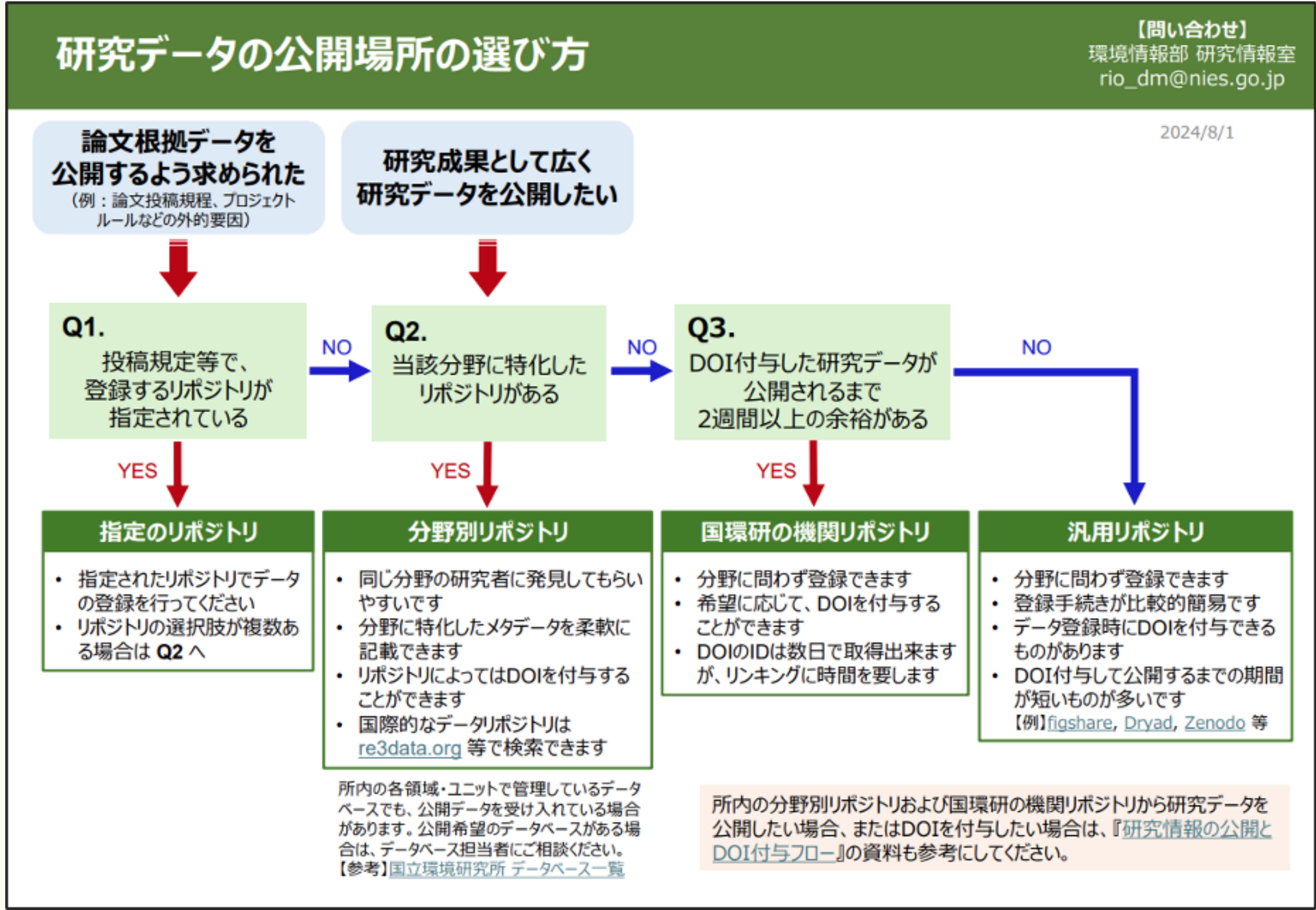
- 持続可能な社会に向けた日本の状況**
SUSBヘッドライン指標に基づいて持続可能性指標指標体系を用いて日本の状況をモニタリングした結果を掲載
- 産業連関表による環境負荷原単位データブック（3EID）**
およそ400に区分されたわが国の経済活動部門別に産業連関分析法の応用により、各部門における環境負荷原単位を算出したデータベース
- 脱炭素型ライフスタイルの選択**
ライフスタイルがもたらす気候変動への影響（カーボンフットプリント）と衣食住に関する産業系ライフスタイルの選択によるGHG削減効果をもとに、57都市・16地方・4大都市圏別に紹介するデータベース
- NIES Global Trade of Materials**
世界各地の金属資源の輸出入量、消費量、ストック量の可視化ツール
- 国が策定する持続可能な開発指標（SDI）のデータベース**
28の国および国際機関等が策定した持続可能な発展に関する指標（SDI）1,848件を管理したデータベース
- 気候研究学園都市の景観変化**
つくば市にわたる70地点における41年間（1980年、1991年、2006年、2021年）の景観変化の記録画像

- 現状はプロジェクト・研究ユニット・グループ・個人単位で作って様々な方法で公開しているデータベース、データセットのリンク集でしかない
- 各研究分野単位で受け入れてくれるデータベースがすべてあるわけでもない

国環研Webページより

Gfarmシンポジウム2025, 秋葉原, 2025/9/19

研究データ公開支援との連携



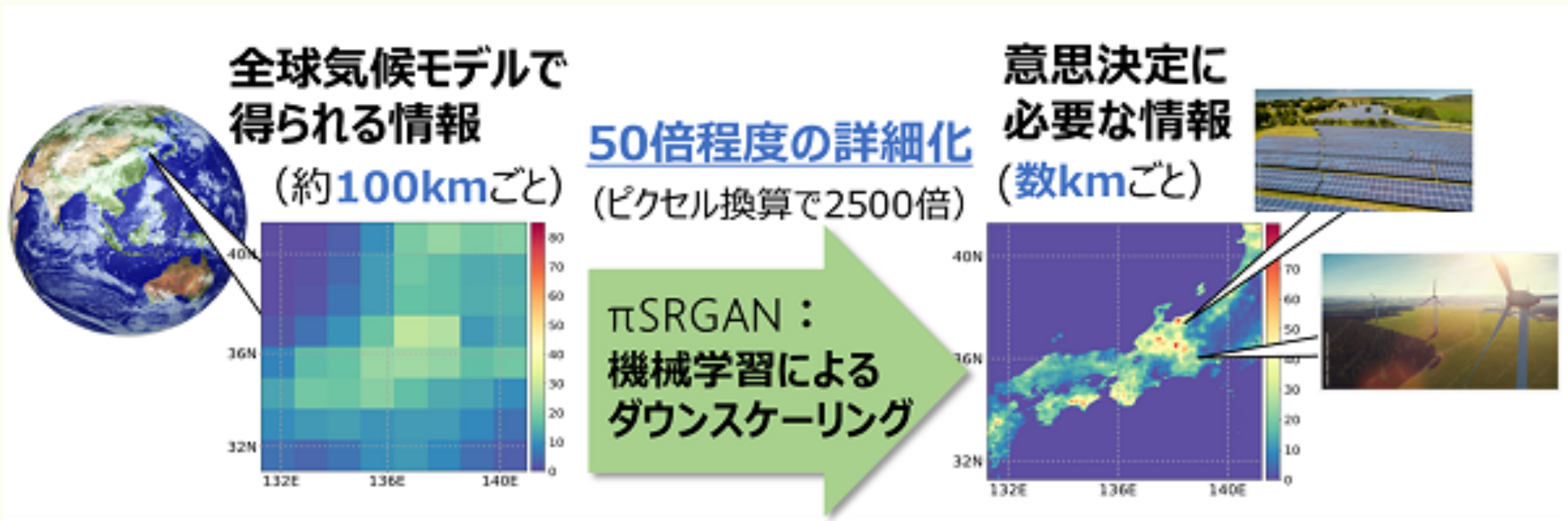
- 外部資金獲得と連動した研究データ管理計画の策定、オープンデータ戦略、適切な研究データ公開など、研究者が対応すべき事項を所として支援
- 集約されたデータ保管場所の提供を起点に、分野横断的な検索、その場解析、他機関データベースとの連携に繋げるのが共創拠点の役割

国環研イントラWebページより

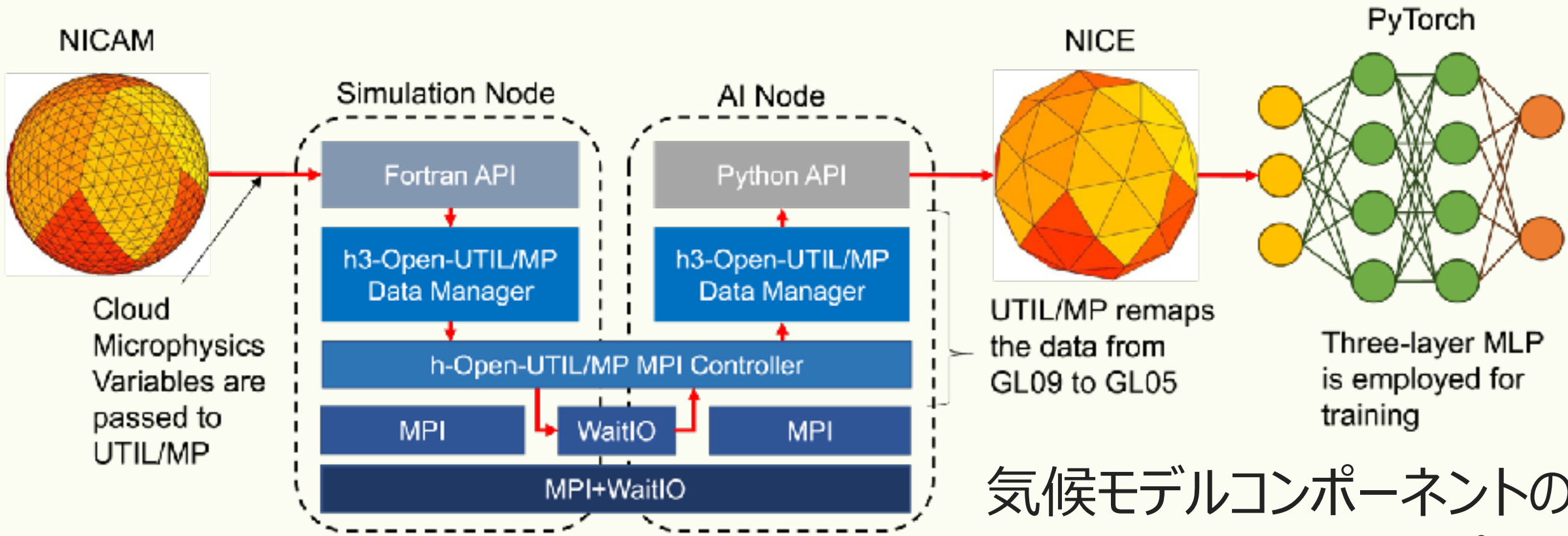
AI for Scienceの推進



セミの大合唱から個別種を識別するAIを開発
2025/8/25 国環研プレスリリース



気候予測データを機械学習により詳細化する技術の開発に成功
2023/4/27 国環研プレスリリース

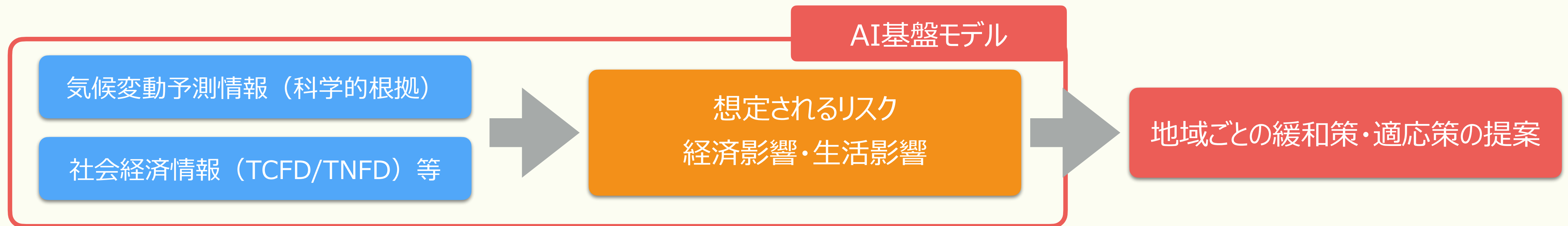


気候モデルコンポーネントのOn-the-fly学習システムの構築
<h3-Open-BDECプロジェクト、中島・荒川・住元・八代>

社会実装のためのAI活用

- 生成AIを用いた環境行動計画策定支援

- 例えば：行政機関・地方自治体・企業等が作成しようとする、気候変動対策の立案を支援
(例：JAMSTEC松岡さんのNEDO研究課題「地域気候サービスのための生成AI基盤モデル開発」)



- 生成AIを用いた伴走型データ利用支援

- マニュアル・フォーマット説明書を読まずに、対話的にデータ検索、抽出、可視化を指示
- 利用するデータのメタデータを説明してもらう
- 統計処理、可視化処理を行なった時のスクリプトを提供し、利用者は再利用

さらに将来的な構想

- オンプレミス脱却は可能か
 - 今回の更新でもクラウド利用は検討した：機動的な計算リソースのスケーリングを考えるならクラウド
 - ストレージのサイズ、データ流量、データ長期保存に関するコストがネック
 - 段階を追うなら、所内オンプレ→外部ハウジングでのプライベートクラウドか
- 環境研なので、再エネ選択&PUEの高いデータセンターを使いたい
 - 営業お待ちしております
- Gfarmの活用について検討するなら
 - 他機関データストレージ・データ基盤との連携：気象データ等は、毎回APIでアクセスしてはネットワーク帯域がもたない可能性があるが、自分たちで全体をコピーして持っておけるだけのストレージの余裕はない
 - 利用頻度の高いデータをキャッシュ的に複製保持できないか



ご清聴ありがとうございました

