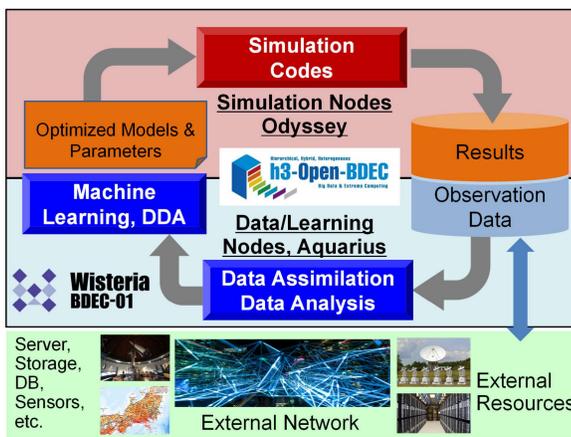


# ヘテロジニアス環境における「計算・データ・学習」融合による新しい計算科学

代表: 中島研吾(東京大学情報基盤センター), 副代表: 古村孝志(東京大学地震研究所)

## 目的・組織

- スーパーコンピューティングの役割は計算科学シミュレーションから、データ科学、機械学習とその範囲を広げている。本研究では「計算(シミュレーション)・データ(Data)・学習(Learning) (S+D+L)」融合による新しい計算科学の開拓と普及を目指し、関連するソフトウェア、数値アルゴリズム、アプリケーションの研究開発、研究成果普及・啓蒙のための環境整備、及び検証を、Wisteria/BDEC-01(東大)、不老(名大)等のヘテロジニアスな構成を有するスパコン、データ活用社会創成プラットフォーム「mdx」により実施する。
- 2021・2022年度は地震シミュレーションを中心とした研究開発を実施してきたが、本年度から大気シミュレーションを加え、材料科学も含めた新たな適用分野への展開も併せて検討する。
- また、ドイツ(Jülich Supercomputing Centre (JSC), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)), フランス(French Atomic Energy Commission (CEA)), クロアチア(Rudjer Boskovic Institute (RBI))の各研究機関(参加者国籍は11ヶ国)も参加し、ソフトウェア、アプリケーションの研究開発を共同で実施、「S+D+L」融合を実現するエコシステム構築へ向けた議論も実施する。

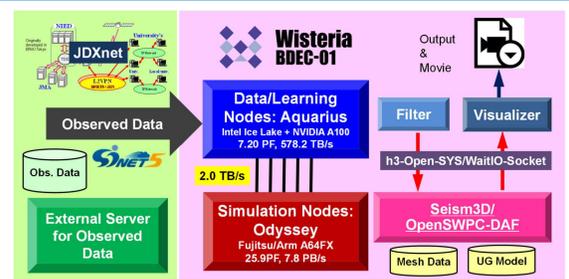


Wisteria/BDEC-01及びh3-Open-BDEC利用による「計算・データ・学習」融合



## 地震シミュレーション

**(a)シミュレーション+リアルタイムデータ同化:** JDXnet地震観測データによるリアルタイムデータ同化と三次元強震動シミュレーション融合については、2021年度にWisteria/BDEC-01上で実装済み、2022年度はシミュレーション部分高速化を実施した。防災・減災教育・啓蒙のためのWebベース実行環境のプロトタイプを整備し、Wisteria/BDEC-01, mdxを利用して、一般市民がスパコン上で「シミュレーション+リアルタイムデータ同化」実行、様々な計算結果を閲覧可能である。2023年度は、フィルタリング、データ同化・シミュレーション、可視化処理部の高速化、複数シナリオ実行機能(2022年度開発中)の最適化、実計算への適用による閲覧用データベース整備の他、mdxによるJDXnet観測データのリアルタイム処理についても検討する。



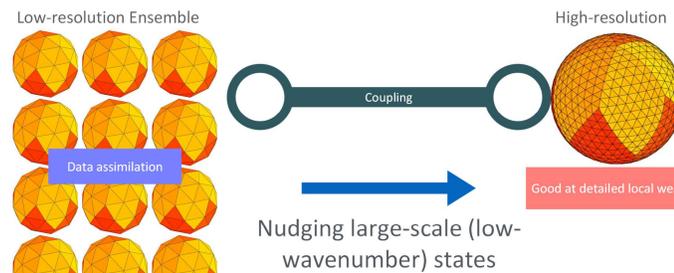
「リアルタイムデータ同化+三次元強震動シミュレーション」融合フレームワーク

**(b)地下構造推定:** (a)で整備した、「リアルタイム観測データ処理+シミュレーション」の枠組みは、大地震発生時の緊急時対応の他、通常時は小規模な地震の観測データの使用により、三次元地下構造の推定、改良に適用できる。2021・2022年度は、大規模機械学習、レプリカ交換モンテカルロ法(REMC) [Kano, Nagao他 2017]について、ツールの整備、Forward解析コード並列化、高速化等を実施した。2023年度以降は、首都圏地震観測網「MeSO-net」等による、より稠密な観測データをを念頭において、大規模機械学習、REMC法の2つのアプローチに基づき、数km四方の比較的ローカルな地盤の地下構造推定を中心として実施する。既存研究では、水平方向に一樣な成層構造等の比較的単純な構造を仮定していたが、より複雑な地下構造の推定が可能となる。2023年度はこれまで開発したツール群、コード群による予備的検討、2024年度はローカルな領域への適用、2025年度は(a)で対象としているような、数百km規模の広域モデル整備へ向けた検討を実施する。

**(c)四次元変分法(2nd-order Adjoint法):** 長尾等が提案する「2nd-order Adjoint法」[池田, 伊藤, 長尾他2018]は、予測結果の不確実性評価が可能な新しい四次元変分法であり、地下構造推定への適用を進めている。2022年度は、データ同化部分の並列化、Forward解析部分の時空間並列化(PinST)を実施した。2023年度は、「Forward解析・データ構造・地下構造改良」を繰り返し実行するフレームワークを整備、提案手法のフィージビリティスタディ、最適化を実施。2024年度以降は、上記(b)と同様に進める。

## 大気シミュレーション・その他のシミュレーション

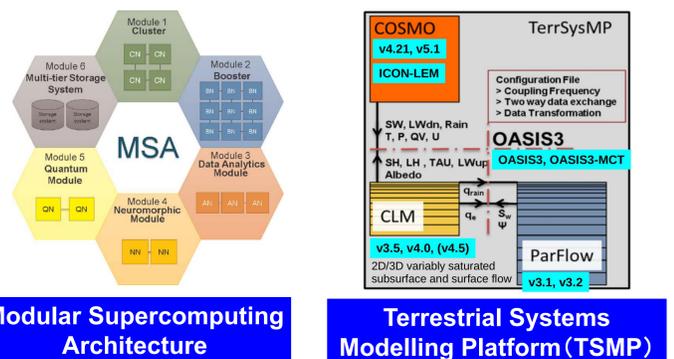
**(a)全球雲解像シミュレーション:** 大気シミュレーションについては、h3-Open-BDECを利用した、「S+D+L」融合の予備的検討が既に進められている。本研究では、全球大気シミュレーションコード「NICAM」を使用した検討を実施する。(a)は、粗いモデルを使用した「Parameterization」のプロセスを機械学習モデルで置き換える試みである。予備的検討では、機械学習による極値解像が不十分であり、2023年度はこの部分の改良を主に実施する。天気予報では、現状の中長期予報では計算資源、計算時間の制限のため、データ同化のためのアンサンブル数は50程度だが、1,000程度実施すれば精度の高い予報が可能である[Miyoshi他2014]。



Ensemble Coupling

**(b)Ensemble Coupling:** アンサンブルシミュレーションに解像度の粗いモデルを適用し、詳細モデルによるシミュレーションと組み合わせる手法である。試算では、詳細アンサンブルシミュレーションモデルと比較して、20分の1程度の計算時間で実行可能である。2023年度は、h3-Open-UTIL/MPを改良したツールを適用した予備計算を実施し、Ensemble Couplingの手法を確立する。

**(c)その他のシミュレーション:** JSC, CEA等の海外協力機関で開発されたコード群をWisteria/BDEC-01-h3-Open-BDEC環境下で実行する。特にJSCで整備された地球科学シミュレーションコード(Terrestrial Systems Modelling Platform (TSMP)), 量子化学コード(Chebyshev Accelerated Subspace Eigensolver (ChASE))は、Modular Supercomputing Architecture (MSA)のアイデアに基づき、JSCで開発されたヘテロジニアスな環境で動作するシミュレーションコードで、様々なワークロードを汎用CPU, GPU, FPGAに割り当てて、連携して計算可能である。



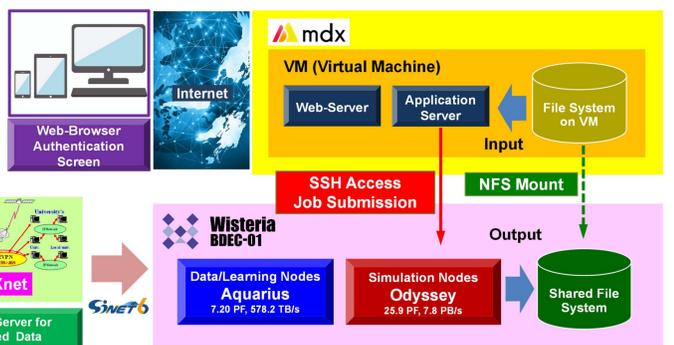
Modular Supercomputing Architecture

Terrestrial Systems Modelling Platform (TSMP)



## ソフトウェア・実行環境開発

**(a)h3-Open-BDEC:** 2022年度は、h3-Open-BDECのうち、多機能カプラーh3-Open-UTIL/MP, 複数アプリケーション間通信ライブラリh3-Open-SYS/WaitIOがWisteria/BDEC-01上で一般ユーザーに利用可能となり、講習会も実施した。WaitIOはTCP-IP経由のWaitIO-Socketの他、2022年度はファイルシステム経由のWaitIO-Fileを開発し、性能評価を実施した(Best Paper Award, PDCAT2022)。2023年度は、各項目に対応した機能の追加、改良、Wisteria/BDEC-01上での性能最適化を実施する。FAUの研究グループと協力して、WaitIOの性能評価モデルを開発し、WaitIO-Socket・WaitIO-Fileの切り替えの自動設定、アルゴリズムの改良に資する。



Webベース実行環境

**(b)Webベース実行環境:** 2022年度は、防災・減災教育・啓蒙のためのWebベース地震シミュレーション実行環境のプロトタイプ(附属資料参照)を整備した。現状はフィルタリング済みの観測データをWisteria/BDEC-01上に置いているため、2023年度は、JDXnetのリアルタイム観測データを直接利用できるように改良を実施する。また、様々なアプリケーションへの拡張、クラスタ用WebポータルOpen OnDemandの適用についても検討する。

