- **東京大学情報基盤センターのスーパーコンピュータ群の概要**
  - システム紹介
  - **利用事例**
- スーパーコンピュータ（スパコン）を使うための様々な制度の紹介
  - 通常利用（一般・トライアル）
  - お試し利用，講習会
  - HPCI
  - JHPCN
  - 若手・女性，AI for HPC
  - HPCチャレンジ，教育利用
  - 企業利用（一般・トライアル）

(計算＋データ＋学習)融合によるエクサスケール時代の革新的シミュレーション手法(1/2)

http://nkl.cc.u-tokyo.ac.jp/h3-Open-BDEC/

- エクサスケール(富岳$^{+}$クラス)のスパコンによる科学的発見の持続的促進のため，計算科学にデータ科学，機械学習のアイディアを導入した(計算＋データ＋学習(S+D+L))融合による革新的シミュレーション手法を提案
  - (計算+データ+学習)融合によるエクサスケール時代の革新的シミュレーション手法（科研費基盤S，代表：中島研吾(東大情基セ)，2019年度～2023年度)

Wisteria
BDEC-01

Hierarchical, Hybrid, Heterogeneous
h3-Open-BDEC
Big Data & Extreme Computing

（計算＋データ＋学習）融合によるエクサスケール時代の革新的シミュレーション手法（2/2）

http://nkl.cc.u-tokyo.ac.jp/h3-Open-BDEC/

- **革新的ソフトウェア基盤「h3-Open-BDEC」の開発**：東大BDECシステム（Wisteria/BDEC-01），「富岳」等を「S+D+L」融合プラットフォームと位置づけ，スパコンの能力を最大限引き出し，最小の計算量・消費電力での計算実行を実現するために，下記2項目を中心に研究
  - 変動精度演算・精度保証・自動チューニングによる新計算原理に基づく革新的数値解法
  - 階層型データ駆動アプローチ（hDDA：Hierarchical Data Driven Approach）等に基づく革新的機械学習手法
  - **H**ierarchical, **H**ybrid, **H**eterogeneous ⇒ h3

Wisteria BDEC-01

Hierarchical, Hybrid, Heterogeneous h3-Open-BDEC Big Data & Extreme Computing

JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE · FOUNDED 1932 ·

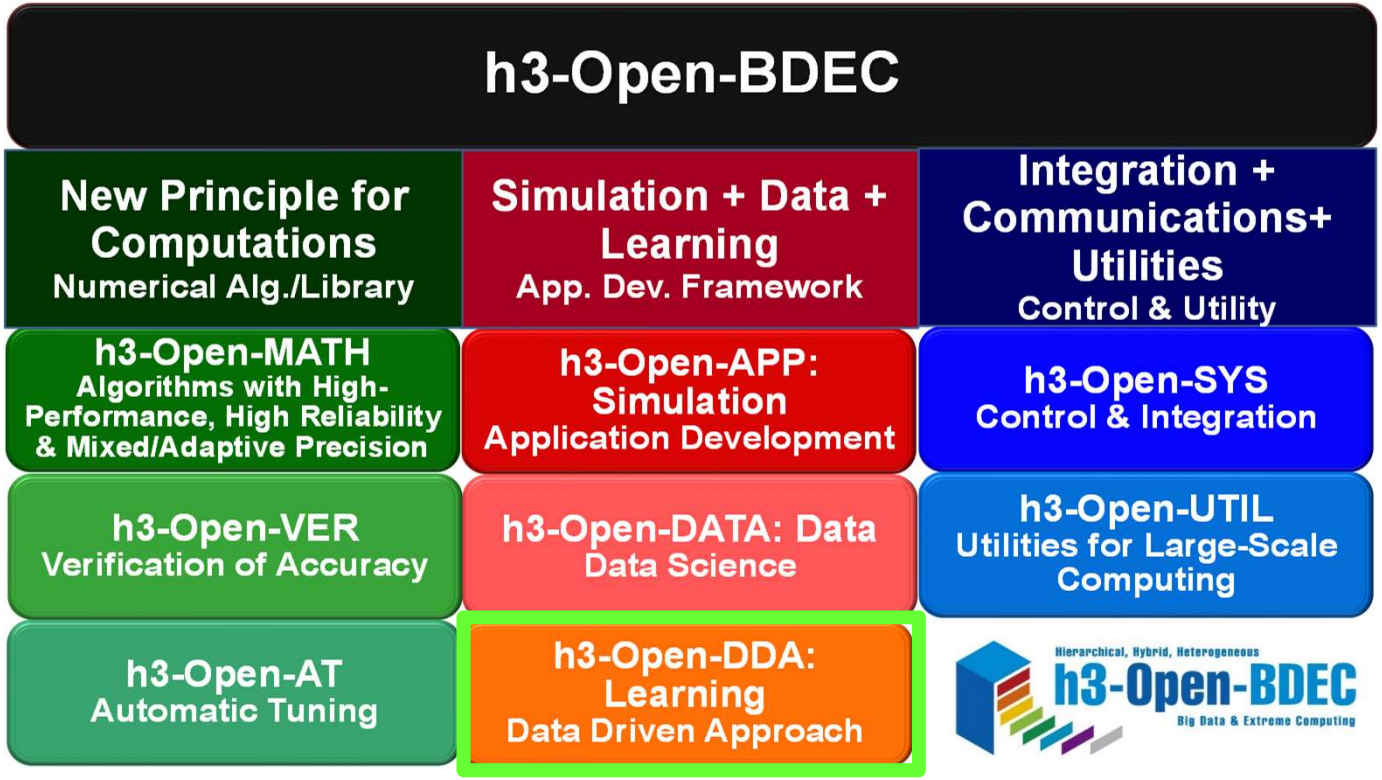
h3-Open-BDEC

「計算＋データ＋学習」融合を実現する革新的ソフトウェア基盤
科研費基盤（2019年度～2023年度，代表：中島研吾，
北大，東大，名大，東工大，九大各センター協力）

① 変動精度演算・精度保証・自動チューニングによる新計算原理に基づく革新的数値解法

② 階層型データ駆動アプローチ（hDDA：Hierarchical Data Driven Approach）等に基づく革新的機械学習手法

- ✓ Hierarchical, Hybrid, Heterogeneous ⇒ h3

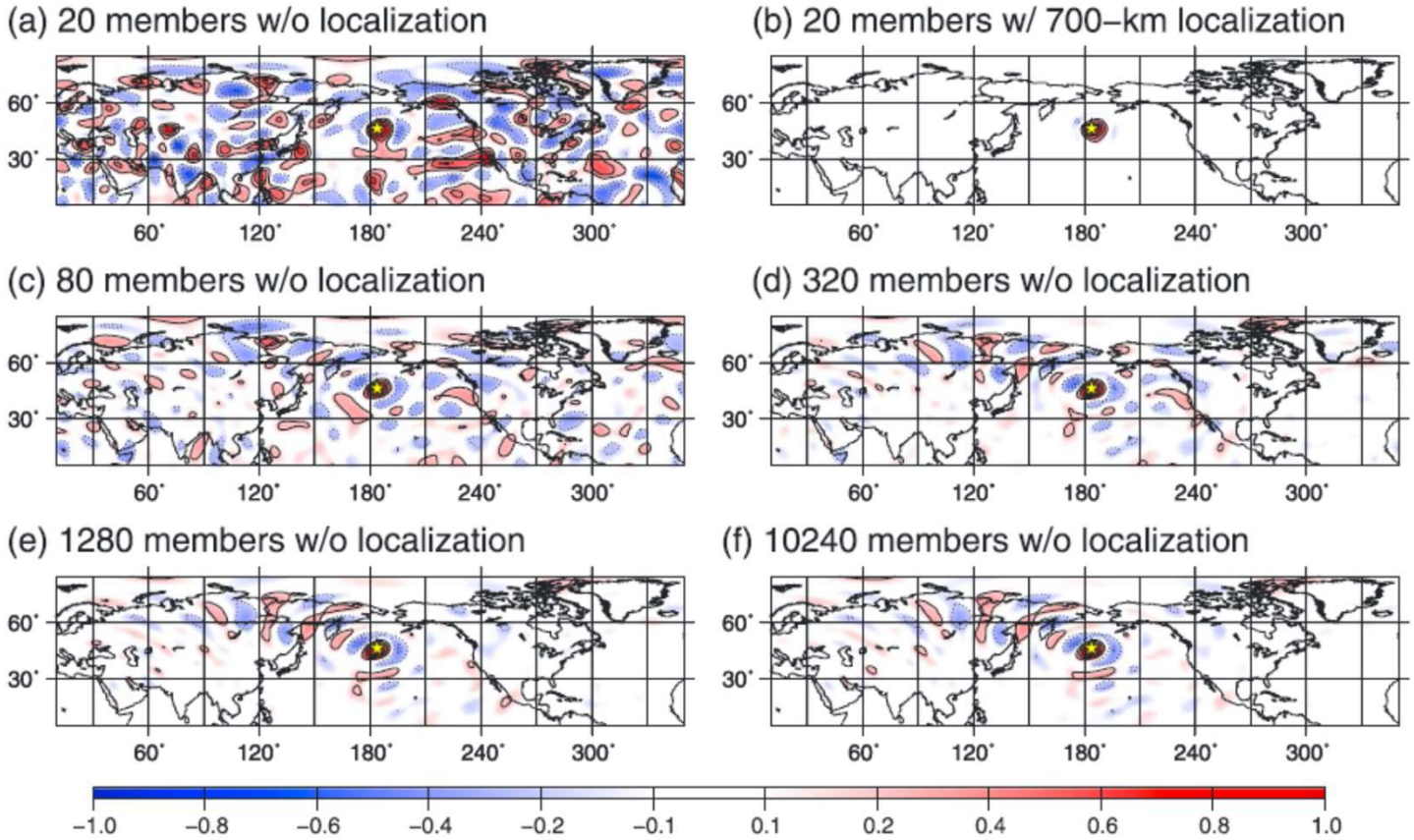


計算科学シミュレーション

- 非線形：多数のパラメータスタディ必要
  - ✓ ケース数削減が非常に重要

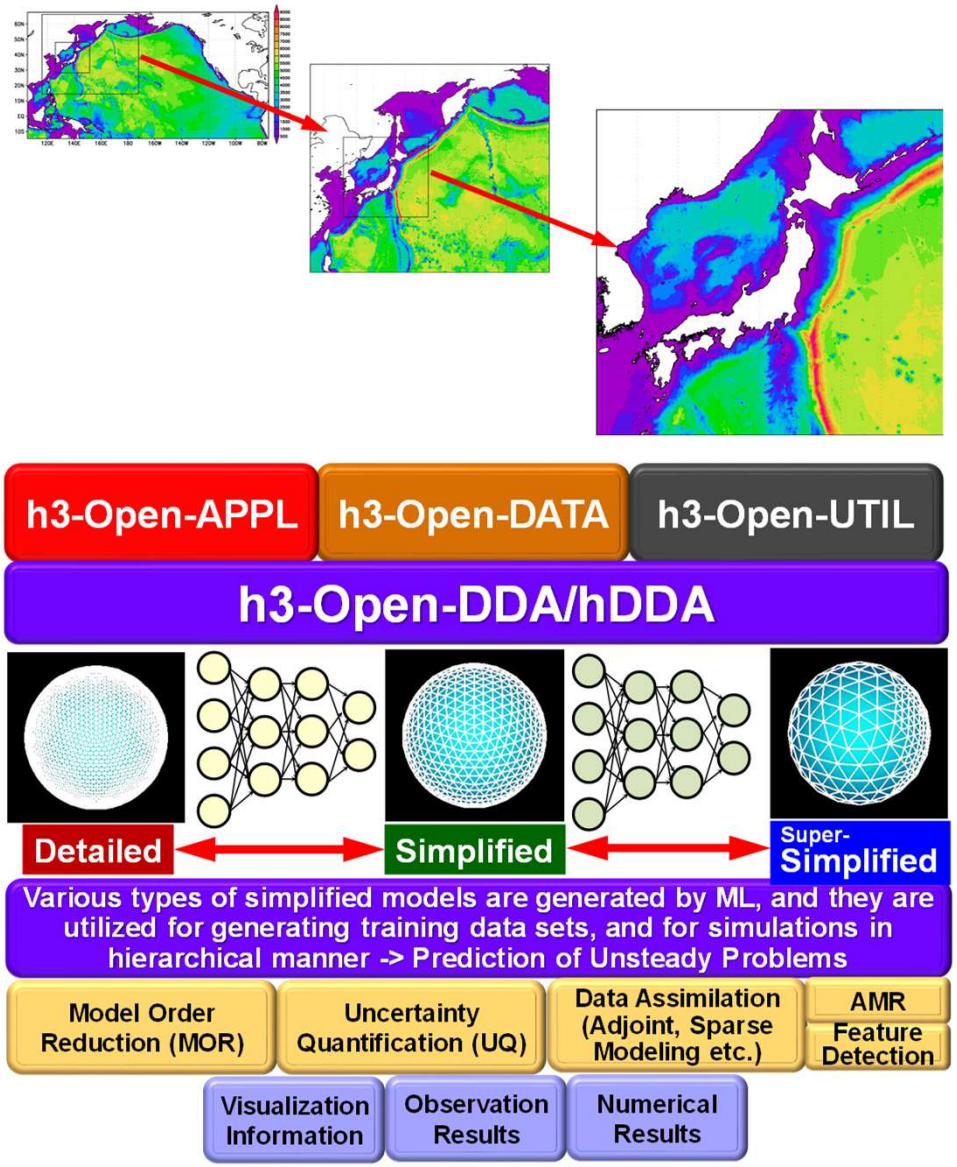
[Miyoshi et al. 2014]

- 天気予報：データ同化
  - ✓ 中期予報（半月程度）：50-100程度のアンサンブル実行，精度良い解のためには1,000程度必要（時間，計算資源の制限）
  - ✓ 機械学習等でパラメータを正確に予測できれば，50-100（あるいはそれ以下）で十分かも知れない

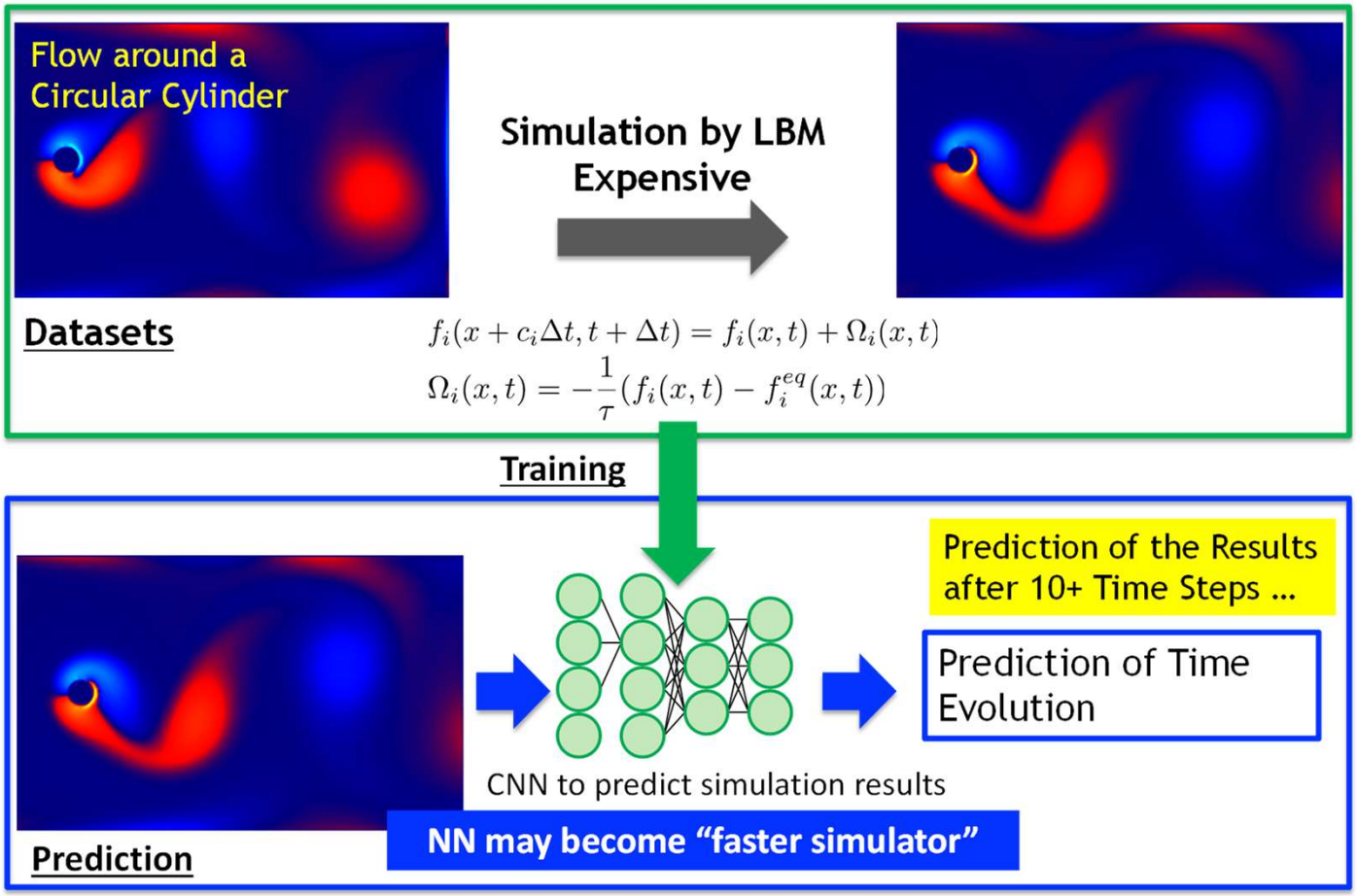


階層型データ駆動アプローチ：hDDA

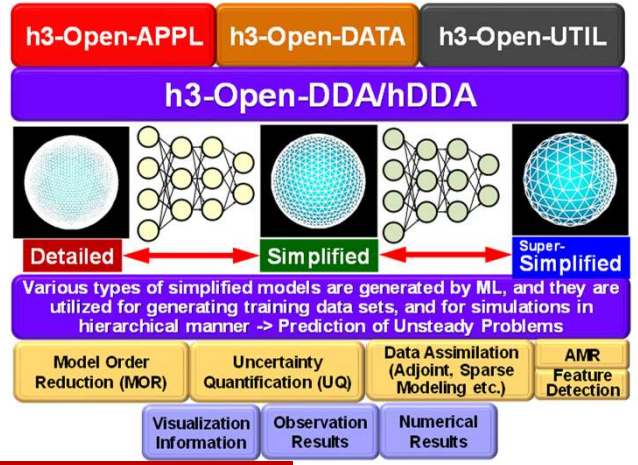
- シミュレーションに機械学習を適用して異なるパラメータでの解を予測するデータ駆動アプローチ（DDA, Data Driven Approach）では，計算を繰り返して教師データを生成する必要がある。
- 階層型DDA（*h*DDA）は，特徴検知，MOR（Model Order Reduction），UQ（Uncertainty Quantification），スパースモデリング，適応格子等の諸機能を駆使して，計算量（メッシュ数，粒子数）を削減した簡易モデルを，機械学習により自動生成，教師データ生成用モデルとして利用する



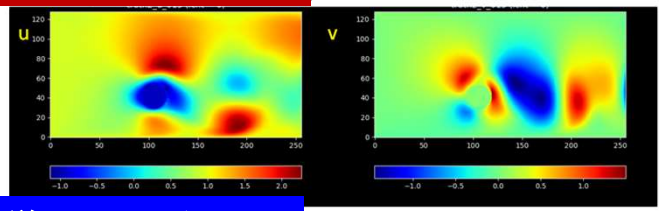
機械学習による非定常数値流体力学シミュレーション高速化（計算＋データ＋学習（S+D+L））融合の実例



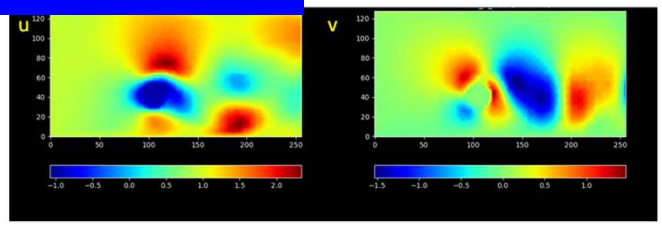
〔Shimokawabe et al. APCOM 2019〕



シミュレーション結果



機械学習による予測



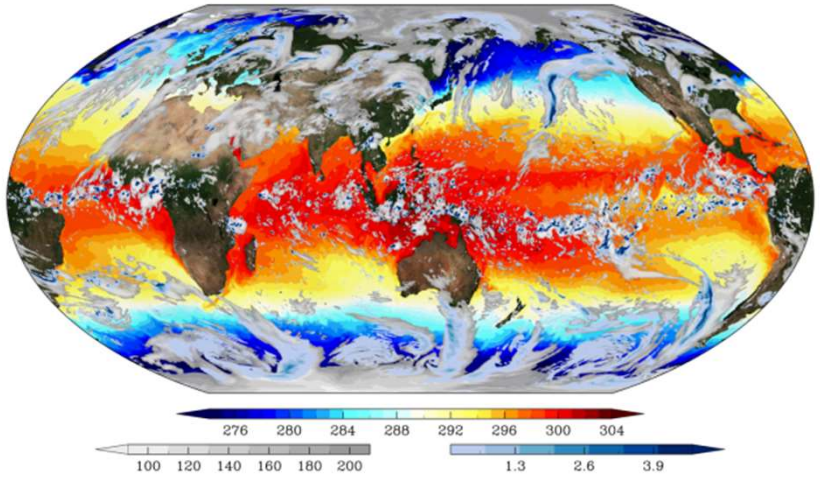
期待される成果と意義

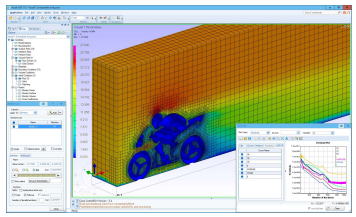
- 計算科学の専門家のみで(S+D+L)融合を容易に実現
  - 機械学習の専門家のサポートを必要としない
- ソースコード，マニュアル類も含めて一般に公開，様々なエクサスケールシステムでの普及を目指す
  - ポスト富岳も含めたポストムーア時代への展開
- h3-Open-BDEC利用による(S+D+L)融合シミュレーションにより従来手法と同等の正確さを保ちつつ，大幅な計算量・消費電力削減を目指す(10分の1が目標)。

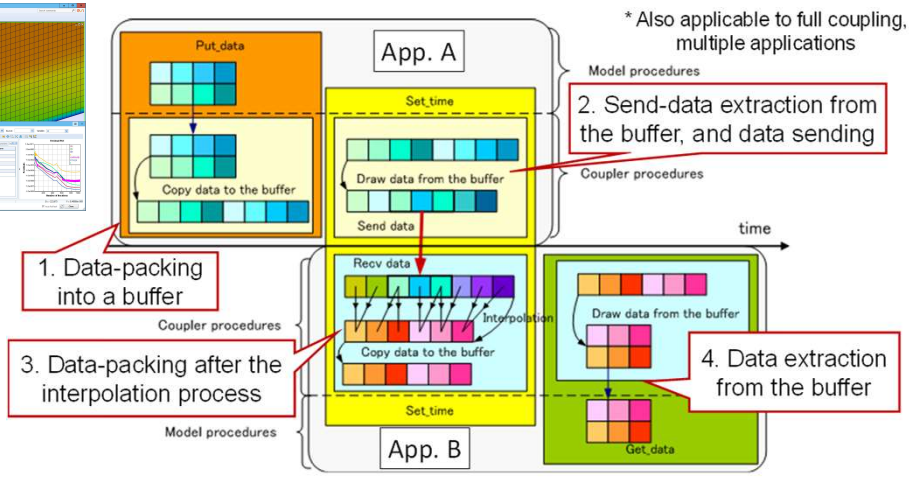
- シミュレーション高度化：パラメータスタディのケース数を削減できる
- リアルタイム災害シミュレーション等への適用

Wisteria/BDEC-01上における h3-Open-BDECを使用した(S+D+L)融合

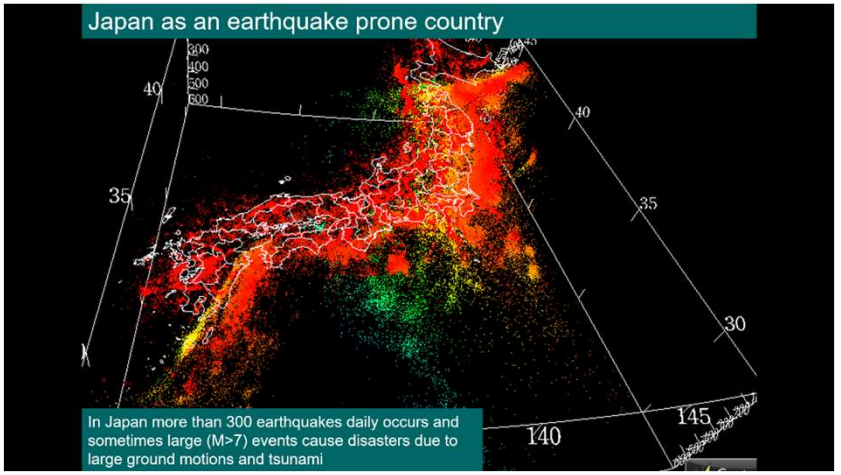
- シミュレーションとデータ同化の融合
  - 典型的・伝統的な(S+D+L)融合
- 気候・気象のための大気海洋連成シミュレーション
  - 東大大気海洋研，理研，国立環境研他
- **リアルタイム同化＋三次元強震動シミュレーション**
  - **東大地震研**
- リアルタイム災害シミュレーション
  - 洪水，津波
- 既存シミュレーションコードの(S+D+L)融合による高度化
  - OpenFOAM



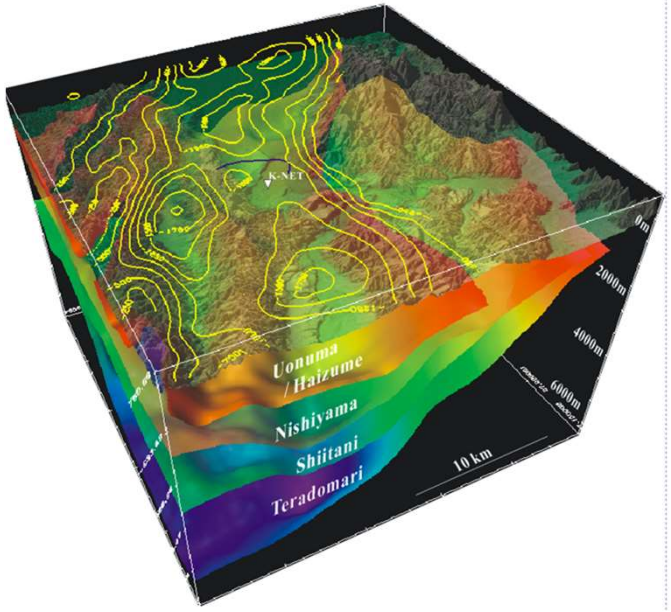




地震シミュレーション：不確実性（Uncertainty）と隣り合わせ



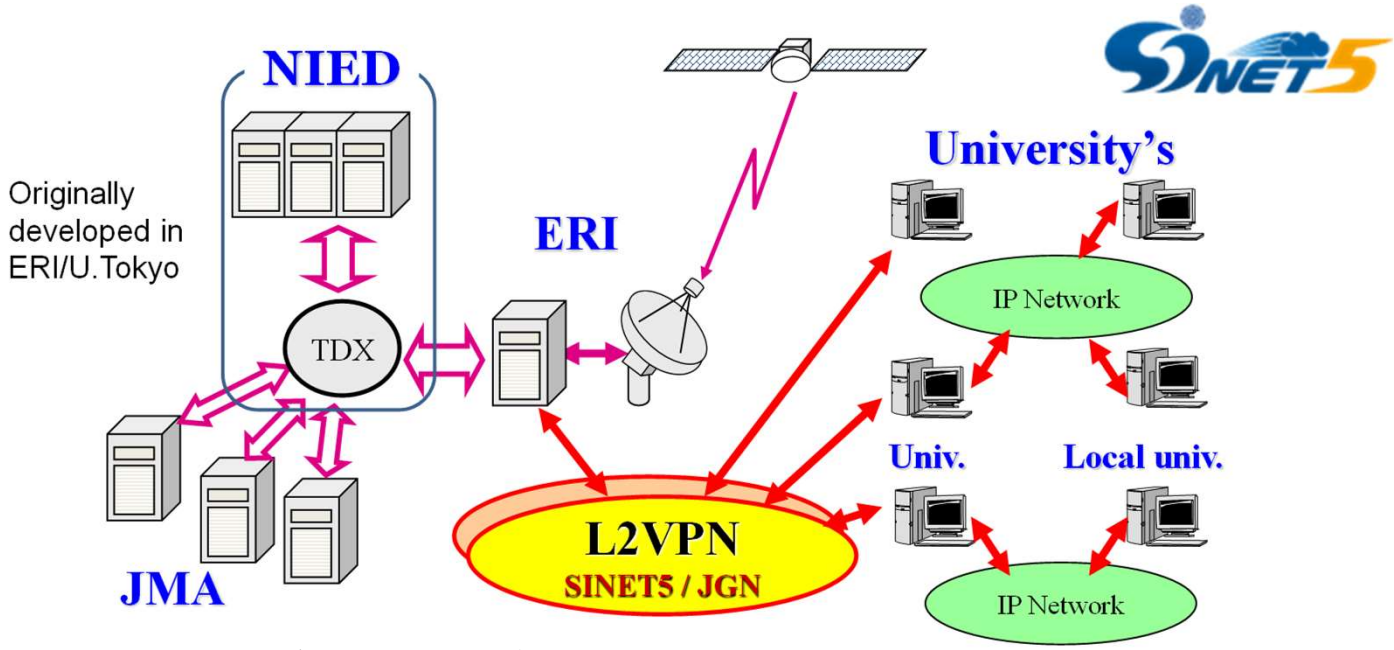
- 地震シミュレーション（強震動シミュレーション）
  - 応力蓄積過程⇒動的破壊⇒地震波動伝播（強震動）
- 地下構造
  - 不均質，不確定
- **シミュレーション・観測融合が不可欠**
- 伝統的なシミュレーション
  - いわゆるフォワードモデリング
  - 「メカニズムの理解」の域を出ない
- **データ同化・リアルタイム観測と融合した手法の開発が必要**
  - **シミュレーション：予測＋観測・データ同化：補正**



[c/o Prof. T. Furumura,ERI/U.Tokyo]

全国地震観測データ流通ネットワーク「JDXnet」

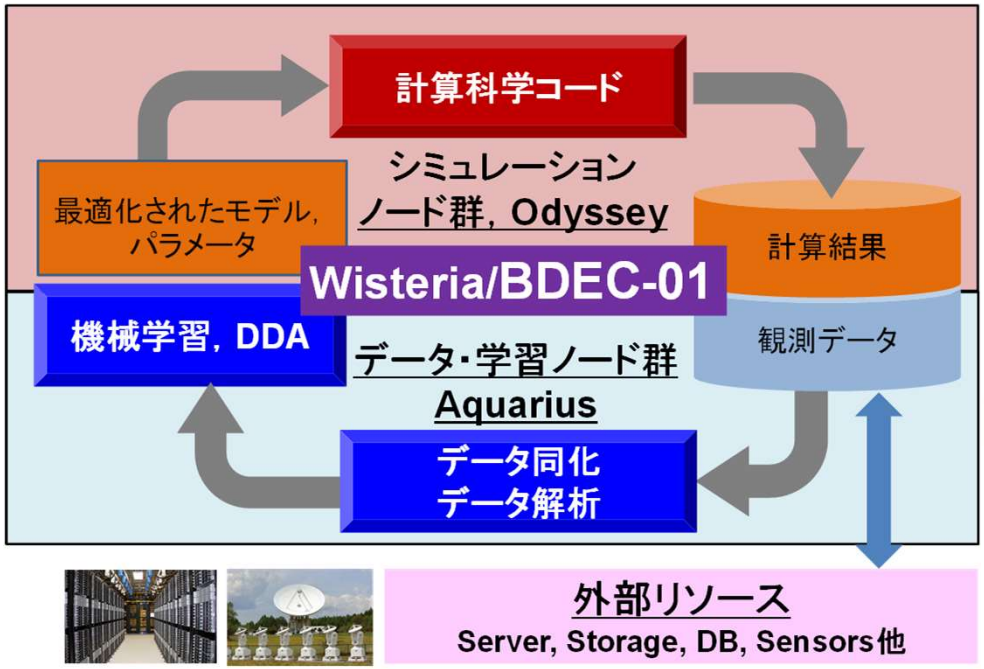
- 国内地震観測点の観測データ（約2,000点，100Hz，3方向）をSINET経由でリアルタイムに取得可能
  - 気象庁，東大地震研，防災科技研，各大学
  - 1日のデータ量：100GB級

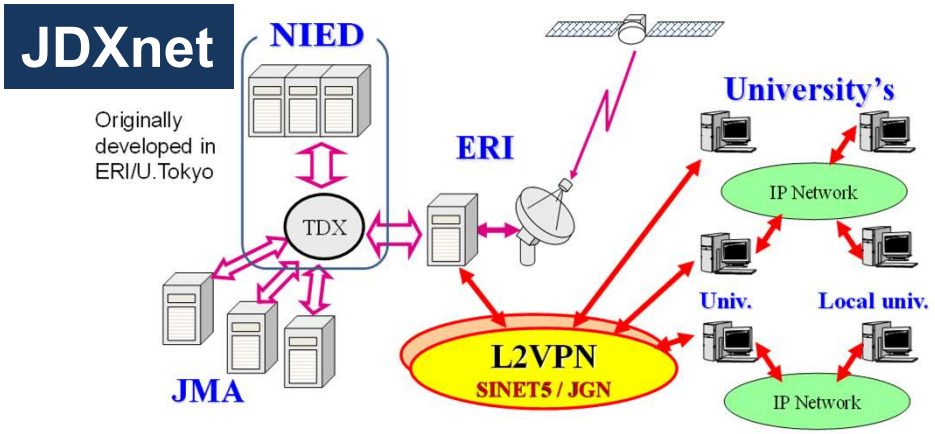


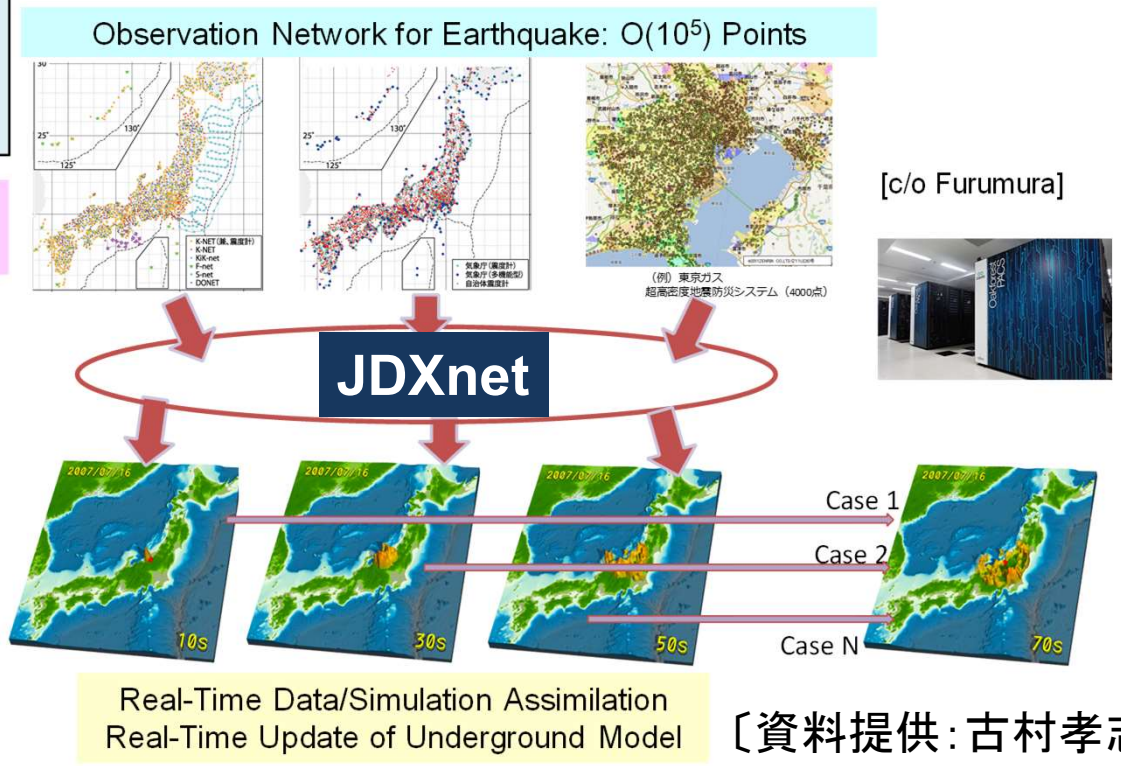
〔資料提供：鶴岡弘准教授（東大・地震研）〕

三次元地震シミュレーション＋リアルタイムデータ同化/観測

JDXnetの観測データを利用したリアルタイムデータ同化/観測







[c/o Furumura]

〔資料提供：古村孝志教授（東大・地震研）〕

最初は「A:同化＋シミュレーション」⇒「B:シミュレーション予測」へ移行

Residual

Assim. Comp. Obs. Comp.

$$\boldsymbol{x}_n^a = \boldsymbol{x}_n^f + \boldsymbol{W}(\boldsymbol{y}_n - \boldsymbol{H}\boldsymbol{x}_n^f)$$

Comp. Assim.

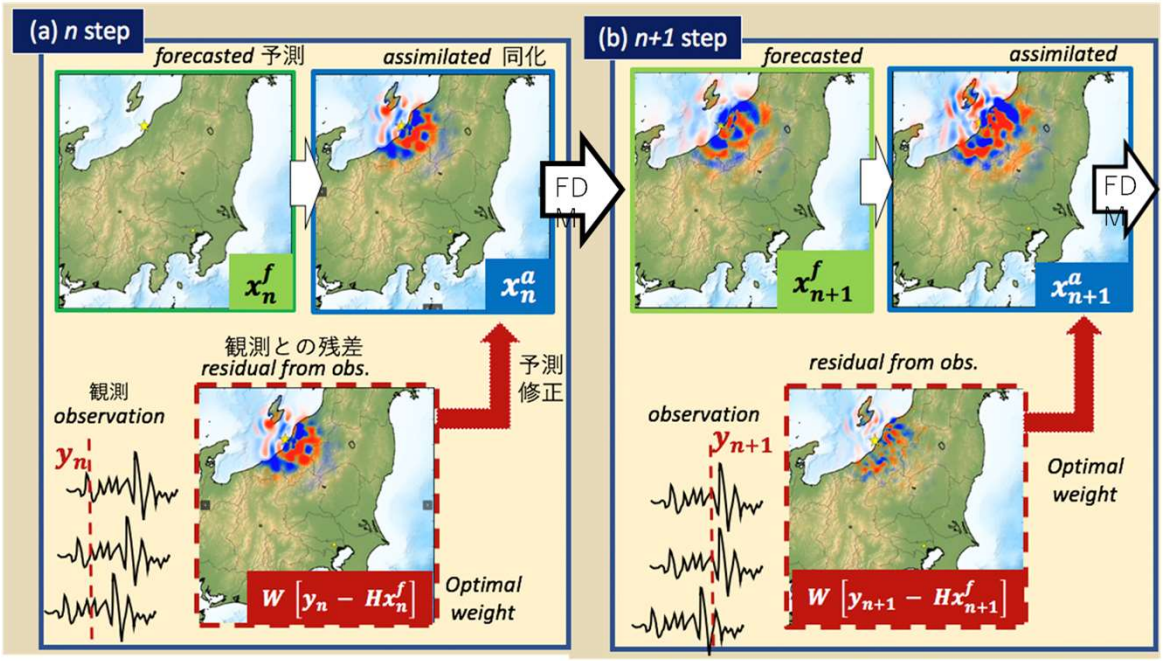
$$\boldsymbol{x}_{n+1}^f = \boldsymbol{F}\boldsymbol{x}_n^a$$

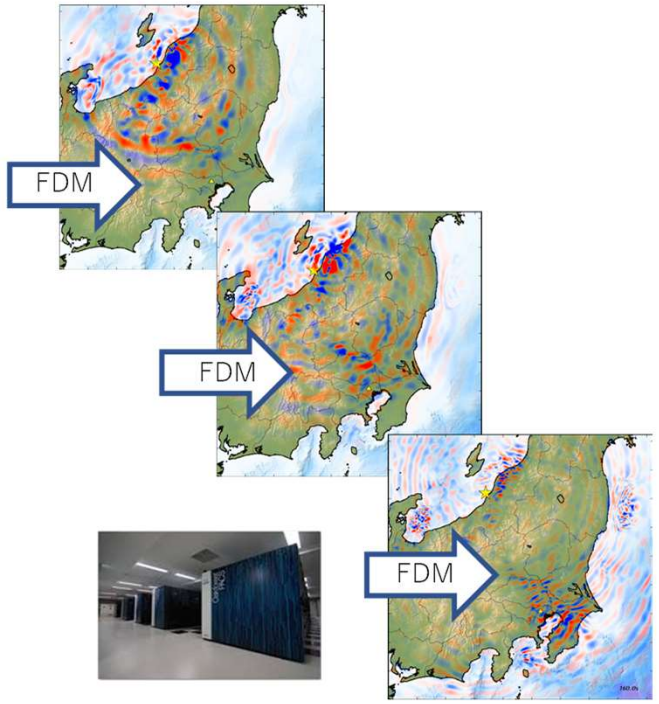
n: Time Step

$\boldsymbol{W}$: **Weighting Matrix**

$\boldsymbol{F}$: Wave Propagation simulation

(A) Assimilation+Simulation → **(B) Pure Simulation/Forecast**

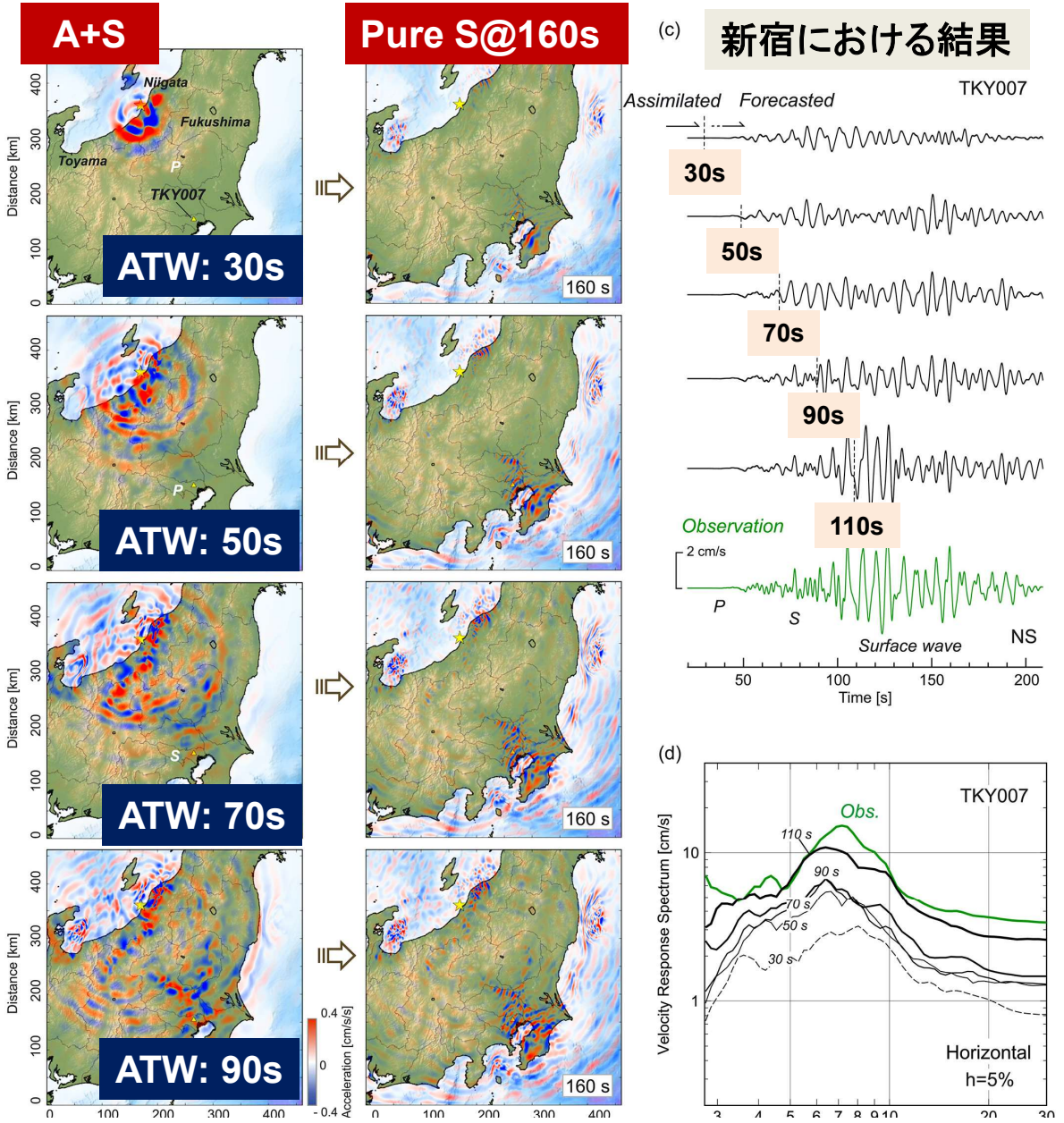




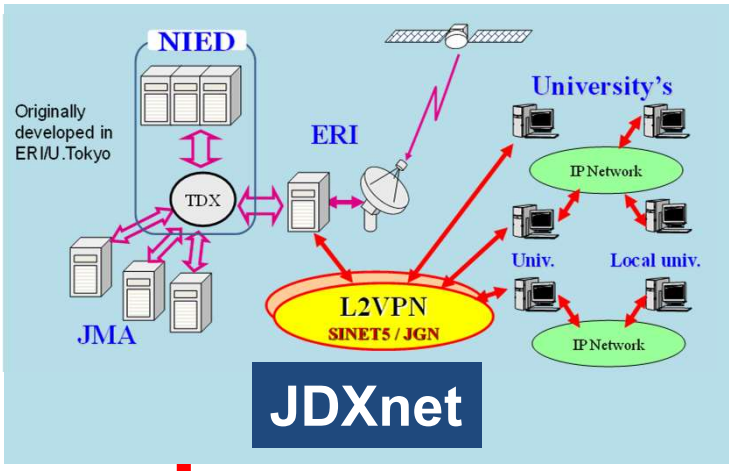
[c/o Prof. T. Furumura, ERI/U.Tokyo]

新潟県中越沖地震（2007年，Mw6.6）

- 同化時間（ATW）を長くとるほど正確な予測が可能
- （A+S）から（Pure S）への移行のタイミングが重要
  - (A+S): データ同化＋シミュレーション
  - (Pure S): シミュレーションのみ



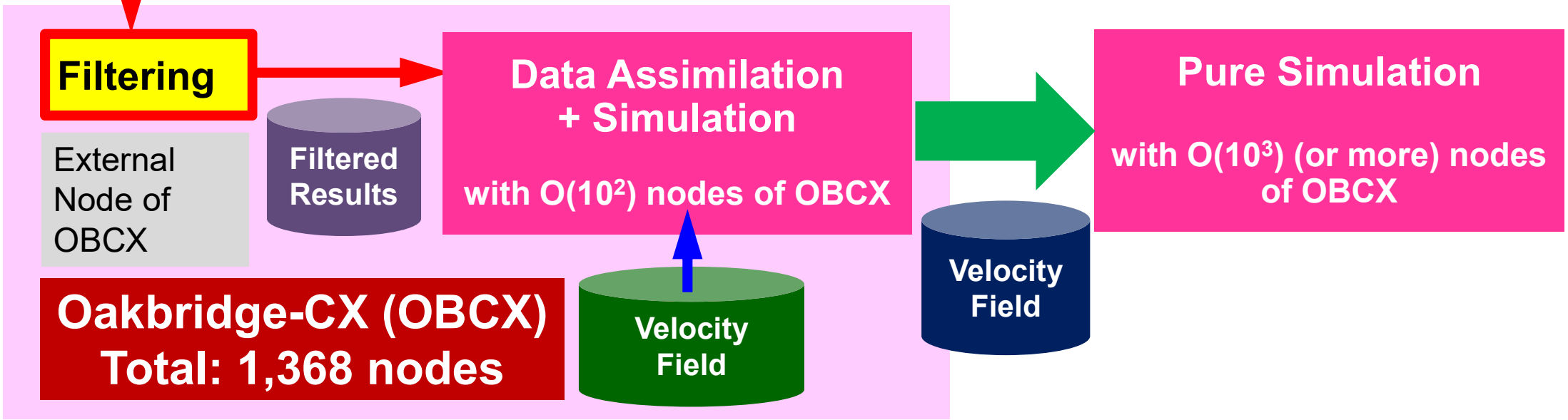
[c/o Prof. T. Furumura, ERI/U.Tokyo]



三次元地震シミュレーション＋リアルタイムデータ同化/観測

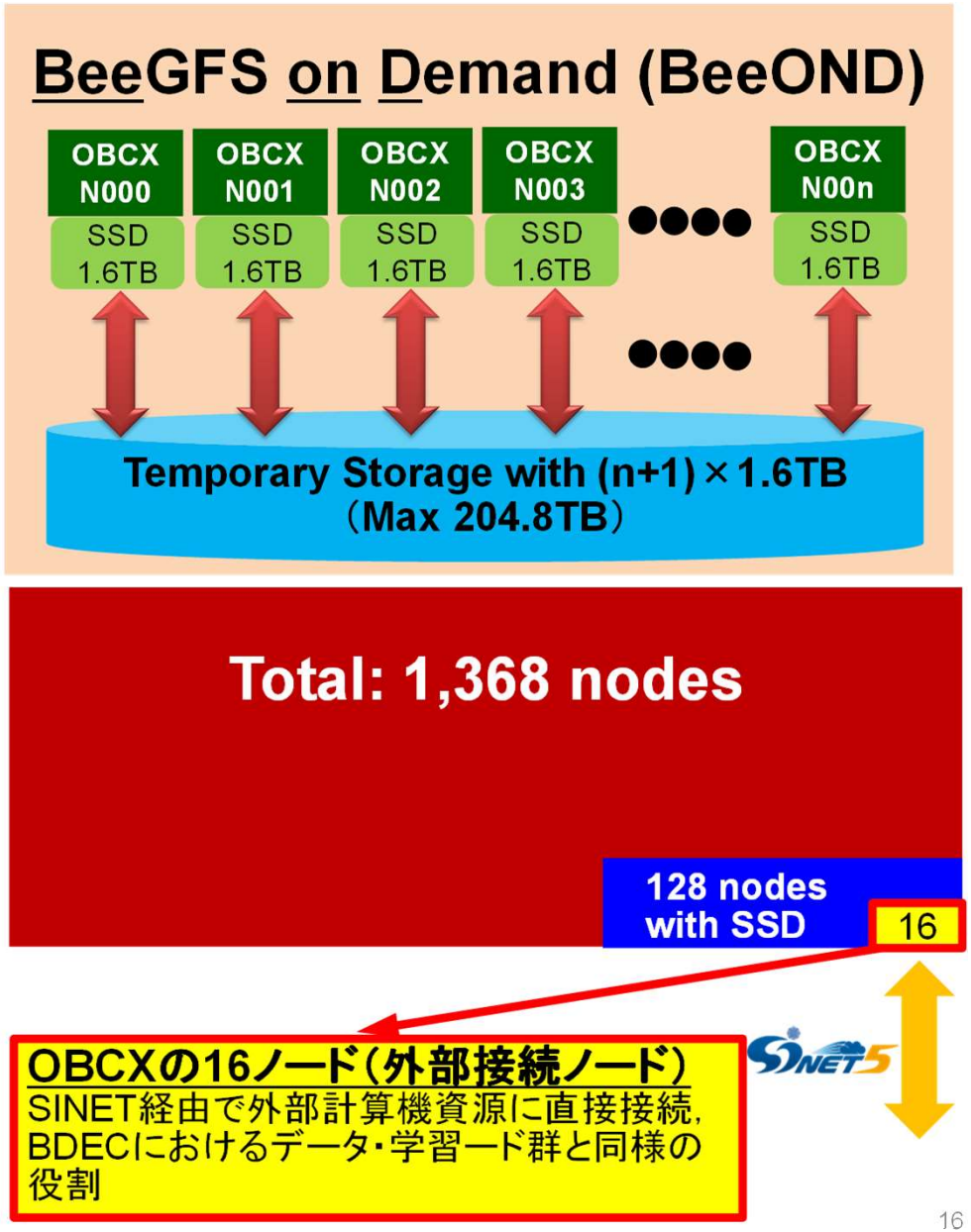
JDXnetの観測データを利用したリアルタイムデータ同化/観測

2019･2020年度はOBCXの外部接続ノードを使って，JDXnetの観測データ取得



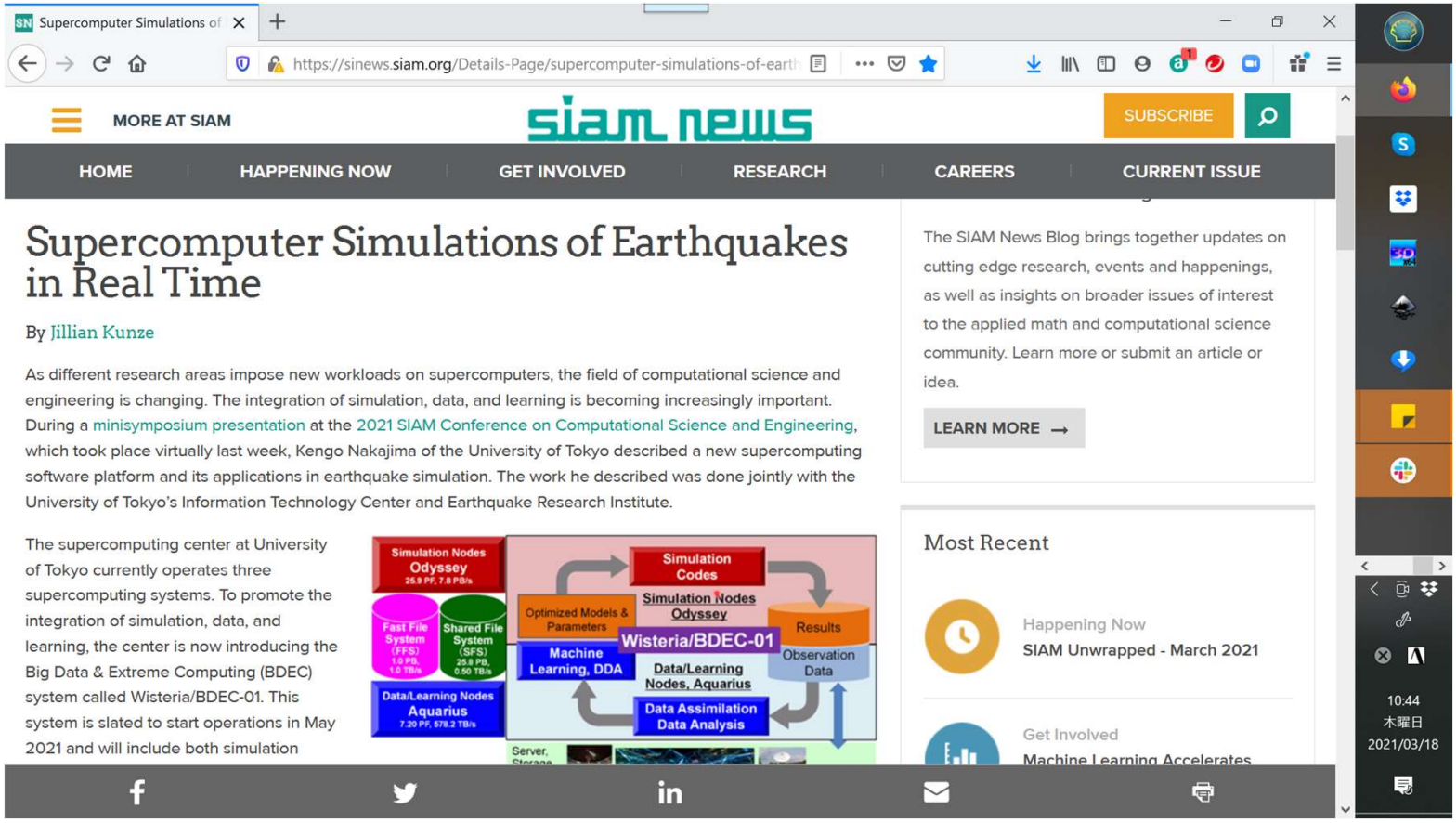
Oakbridge-CX（OBCX）：BDECに向けた実験システム

- 全1,368ノードのうち128ノードにSSD（Solid State Drive）搭載
  - Intel SSD＋BeeGFS
    - 容量：1.6 TB/node
    - 読み書き性能： 3.20/1.32 GB/s/node
    - BeeOND（BeeGFS-on-Demand）によって合計200+TB（128×1.6）の高速ファイルシステムとして使用可能
  - データ科学アプリケーション
    - ソフトウェア類も充実
  - ステージング，チェックポイント
  - 128ノードのうち16ノードはSINET経由で外部リソース（サーバー，ストレージ，センサーネットワーク）に直接接続⇒外部接続ノード



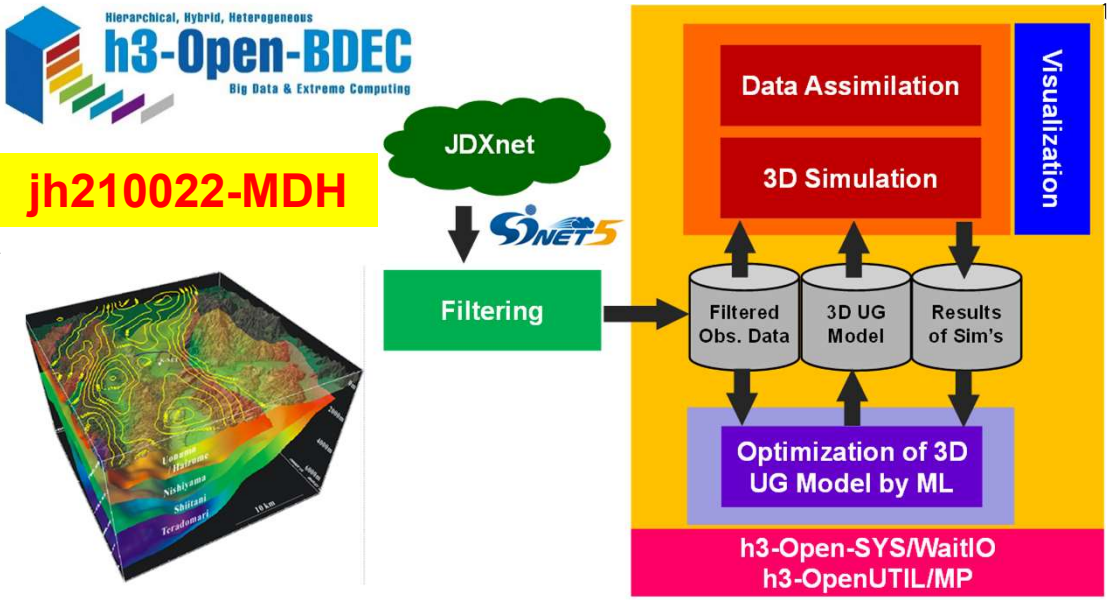
The article on my presentation@CSE21 appears in *SIAM News*

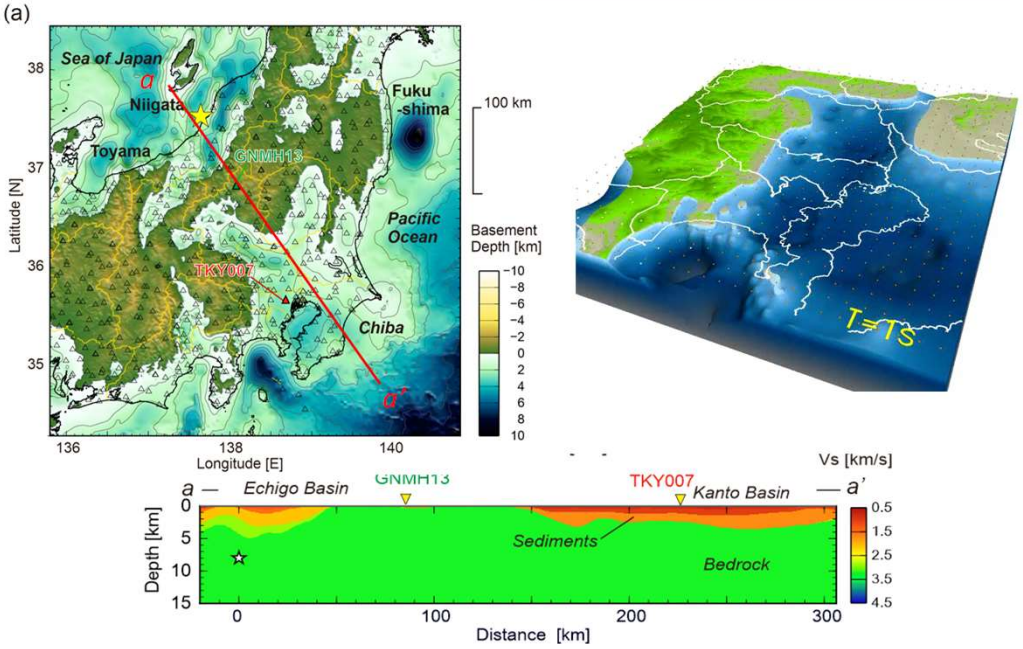
https://sinews.siam.org/Details-Page/supercomputer-simulations-of-earthquakes-in-real-time



リアルタイムデータ同化＋3D強震動Sim. 融合（1/2）

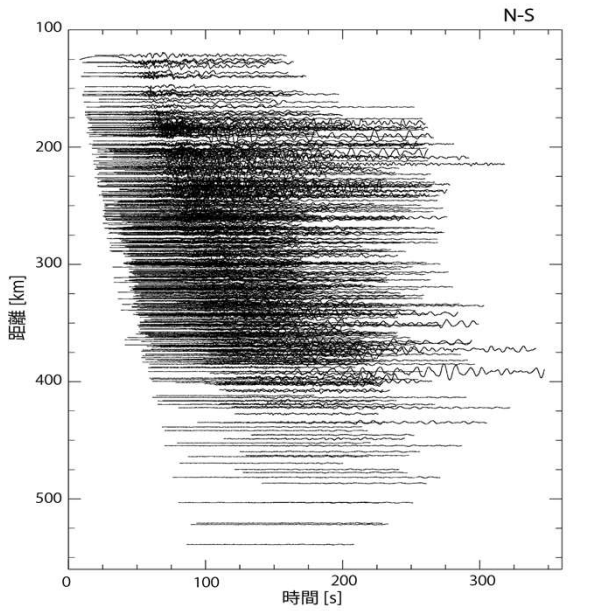
- 大地震時：正確な地震波動伝播予測
  - 地震波観測データを同化して得られる変位分布を初期条件として入力
  - 100秒先の予測を10秒以内に計算
  - 的確な避難計画の策定
- 通常時：地下構造モデル改良
  - 地下構造は複雑，不均質，実はよくわかっておらず，大小の地震時の逆解析等によりモデルを少しずつ改良する，のが現状
  - 観測データ・三次元シミュレーション・データ解析を元に地下構造モデルの改良に使う
    - ✓ 例えばMw=3.0+の地震が起きた場合
    - ✓ 機械学習により精度高いモデル生成
  - ⇒（S+D+L）融合へ

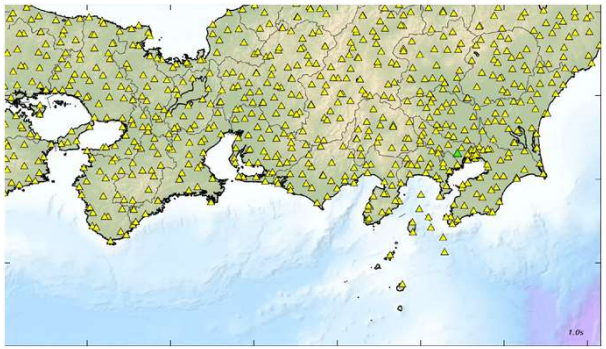




紀伊半島南東沖地震（2004年，Mw 7.4）[c/o Oba & Furumura]

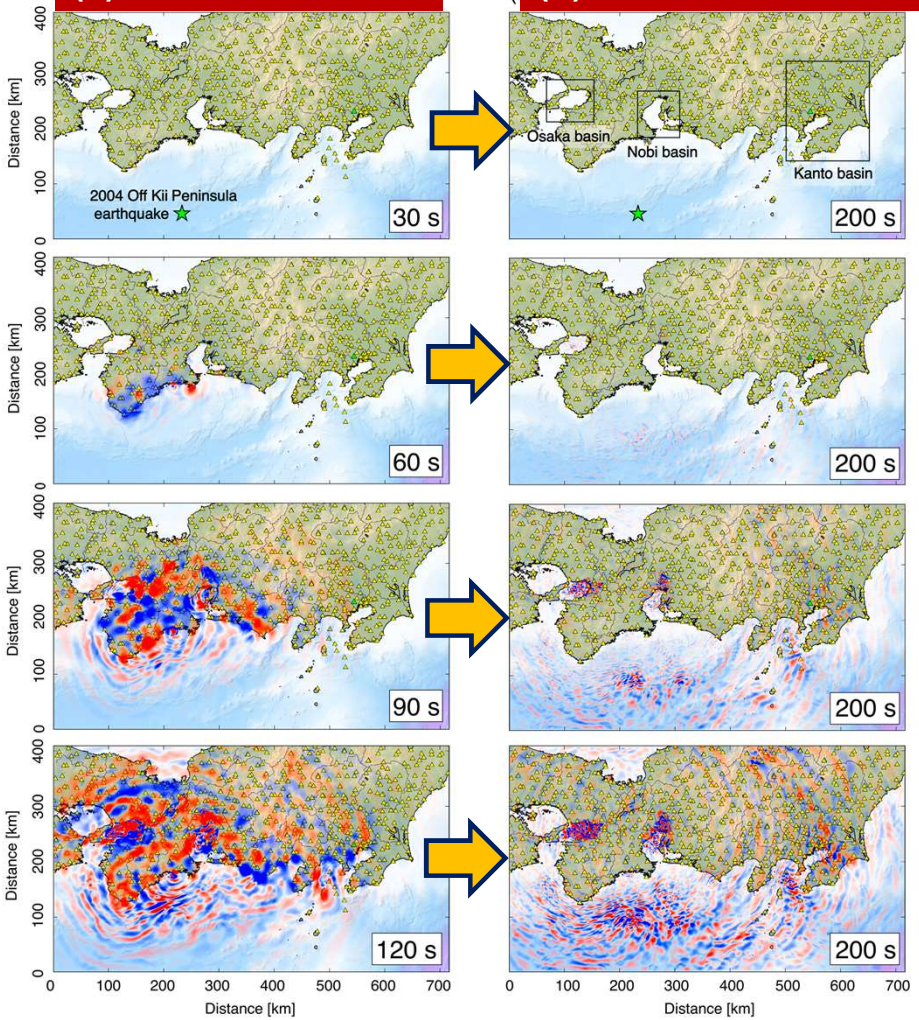
○ **Observation (**K-NET, KiK-net 446 pts**)**



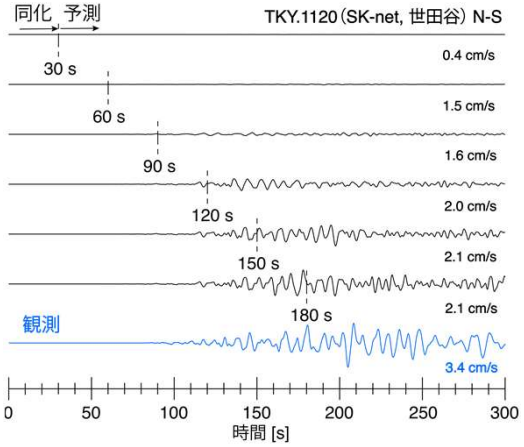


(a) Sim. + Assim.

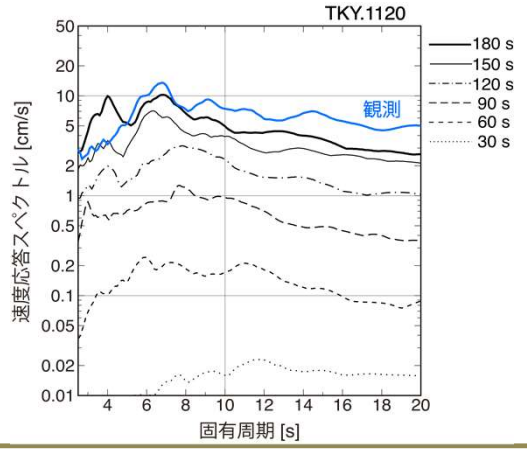
(b) Pure Simulation



Long Wave Propagation in Tokyo



Response Spectrum

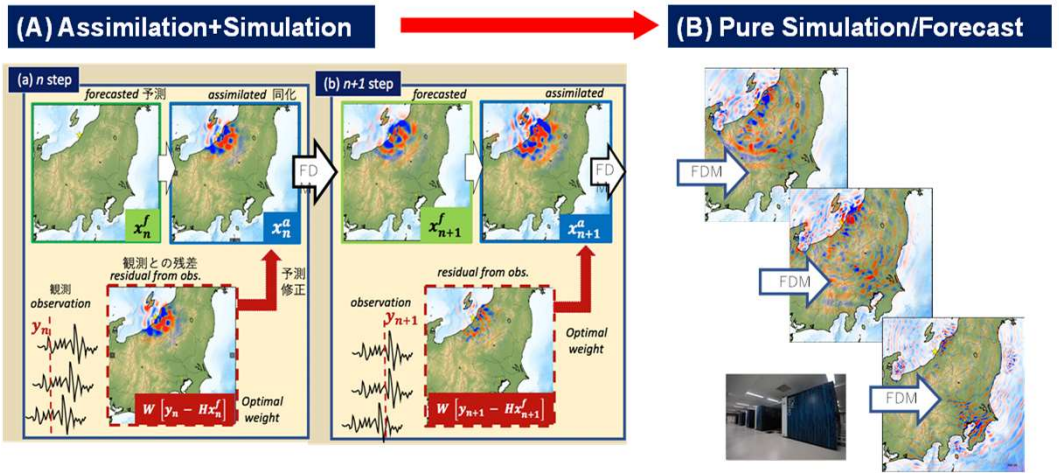


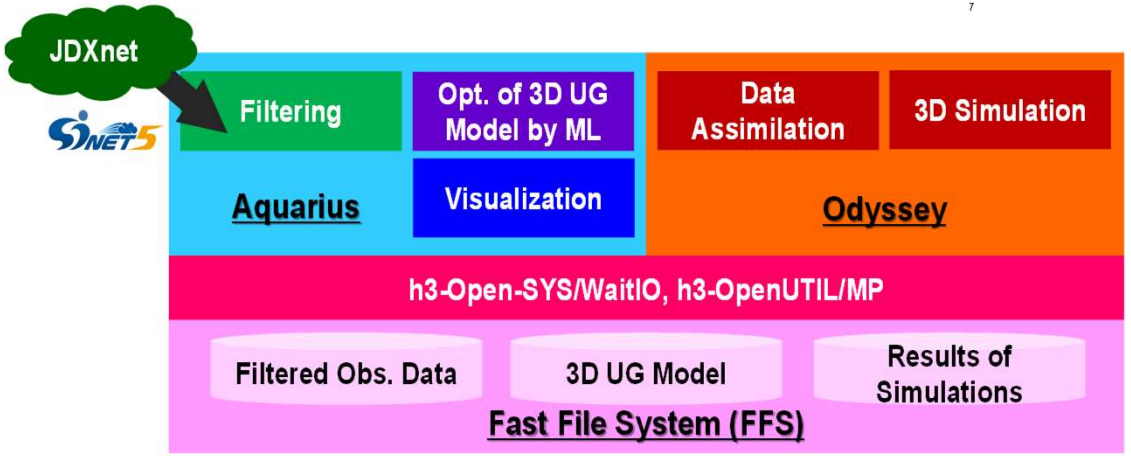
リアルタイムデータ同化＋3D強震動Sim 融合（2/2）

jh210022-MDH

Wisteria BDEC-01

- Wisteria/BDEC-01の利用
- データ同化＋シミュレーション
  - Optimal Interpolation Technique：高速
- 三次元地下構造モデル高度化
  - リアルタイム性はそれほど重要ではない
  - より高度なデータ同化手法（e.g.四次元変分法）を適用
- Odyssey
  - データ同化，シミュレーション
- Aquarius
  - フィルタリング，機械学習，可視化



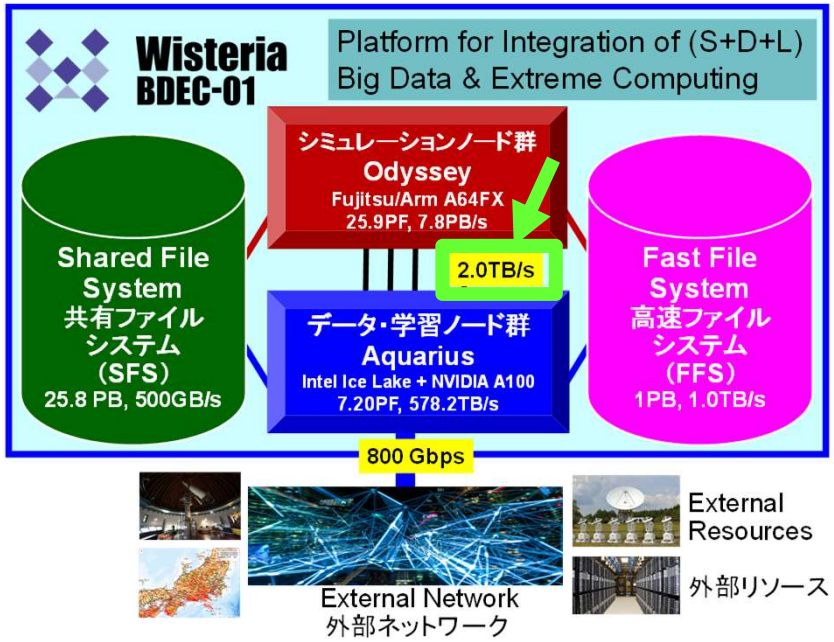


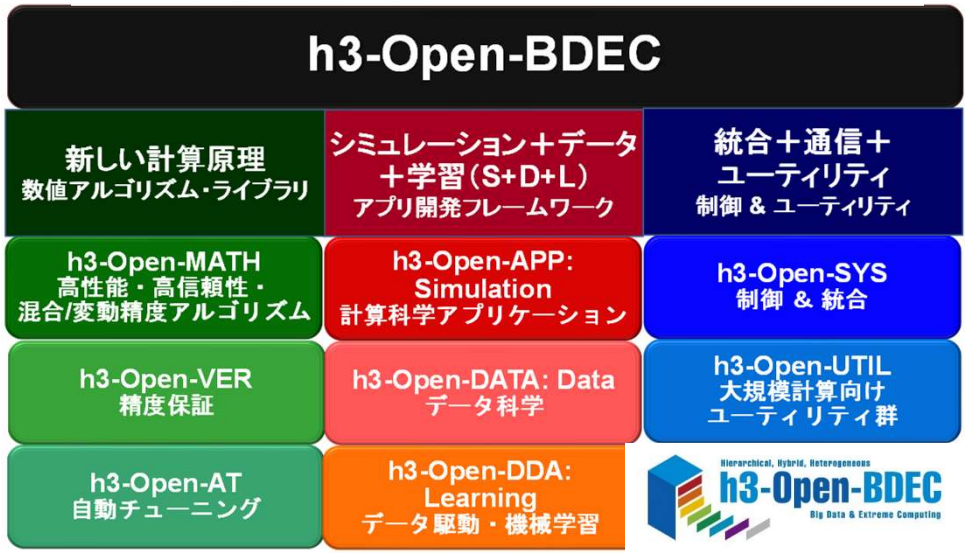
AI for HPC の実現

Wisteria BDEC-01

Hierarchical, Hybrid, Heterogeneous h3-Open-BDEC Big Data & Extreme Computing

- Odyssey-Aquarius連携
  - MPIによる通信は不可
    - O-Aを跨いでMPIプログラムは動かない
  - Odyssey-Aquarius間はInfiniband-EDR（2TB/sec）で結合されている
- ソフトウェア開発
  - O-A間通信：h3-Open-SYS/WaitIO
    - IB-EDR経由
    - 高速ファイルシステム（FFS）経由連携
  - 高機能カプラー：h3-Open-UTIL/MP

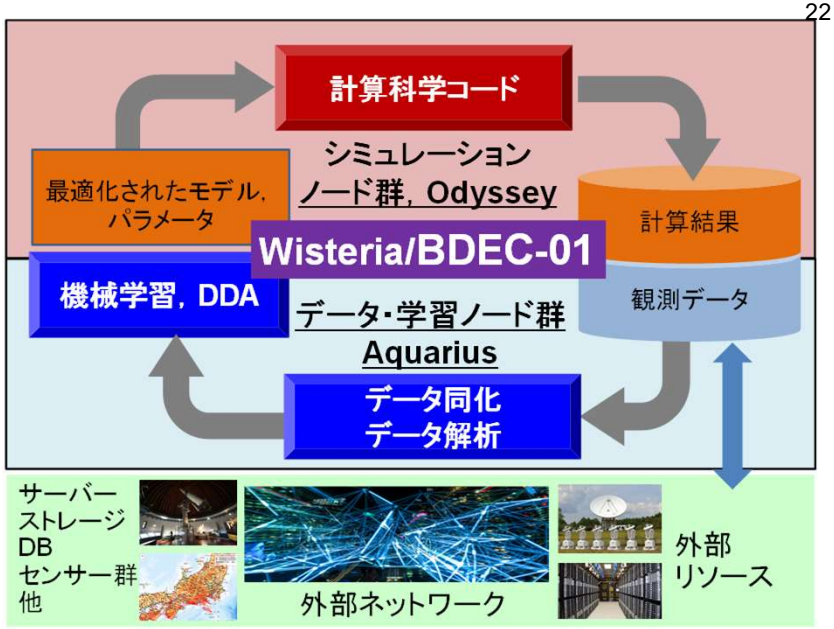


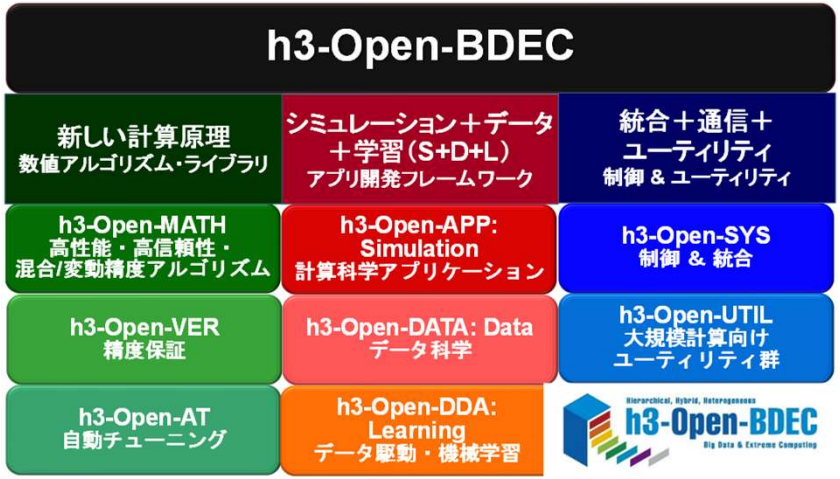


h3-Open-SYS/WaitIO

データ受け渡しライブラリ〔松葉，2020〕

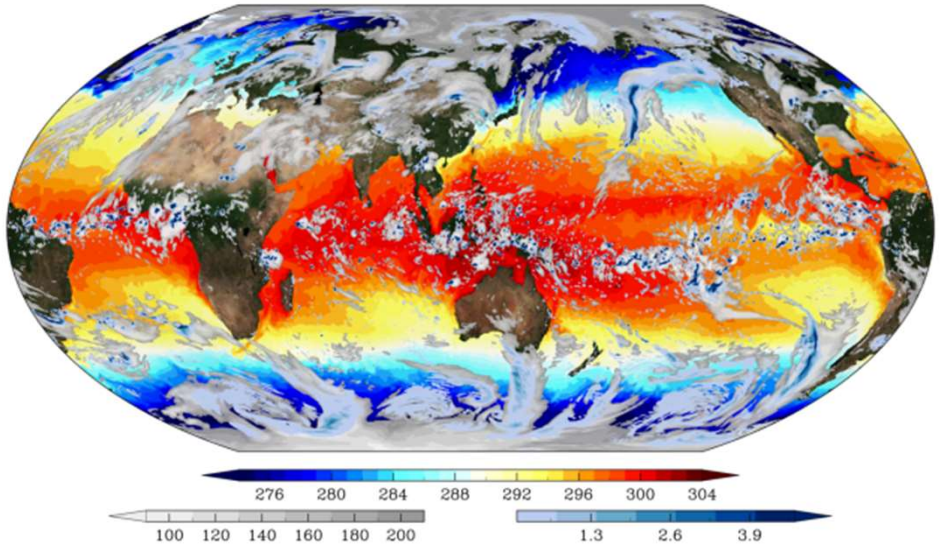
- ヘテロジニアス環境下での異なるコンポーネント間ファイル経由連携ライブラリとして考案
- 機能
  - ✓ Odyssey～Aquarius間連携
    - □ IB-EDR経由通信
    - □ ファイル経由
  - ✓ 外部からのデータ取得（観測データ等）
  - ✓ 読み込み・書き出しの同期
- API：C/C++，Fortranから呼び出し可能
  - ✓ MPIライクなインタフェースを提供
- 多機能カプラー（h3-Open-UTIL/MP）との連携

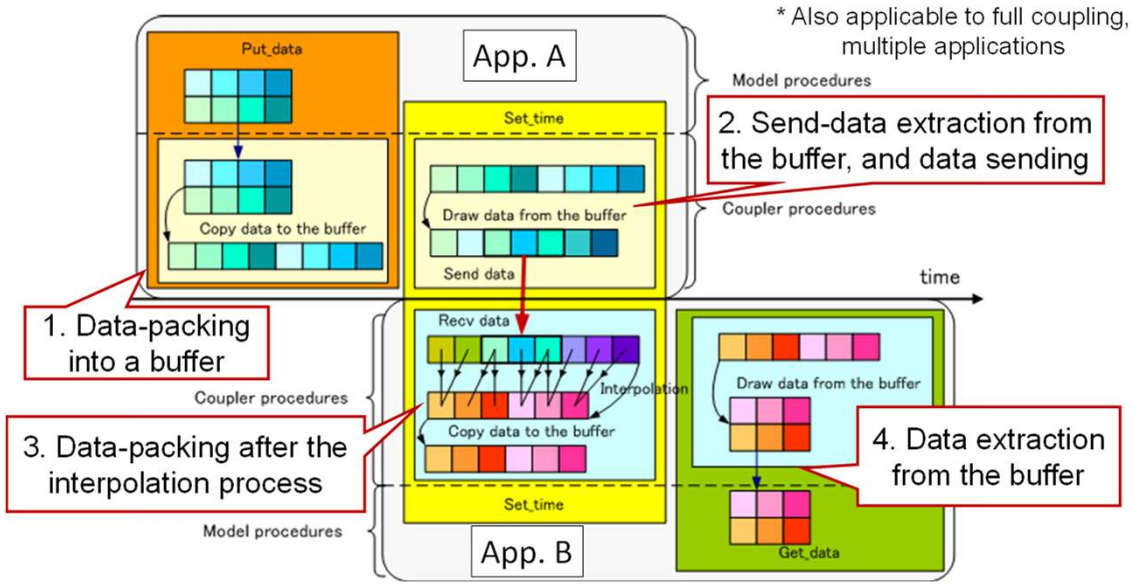




連成シミュレーションのためのカプラー〔荒川，八代〕

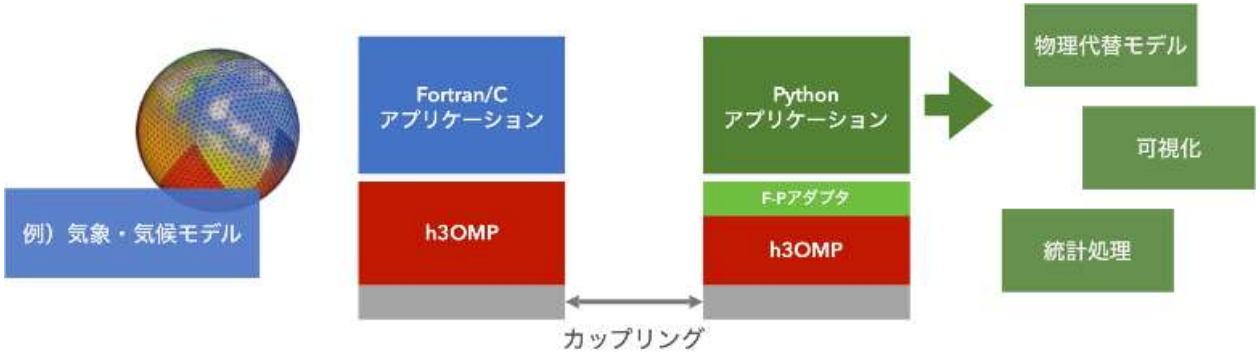
- 従来のカプラー（Coupler）：ppOpen-MATH/MP
  - 複数（通常2つ：大気（NICAM）＋海洋（COCO））のアプリケーションの弱連成（Weak Coupling）をサポート
  - 各アプリケーションは1種類の計算をやる





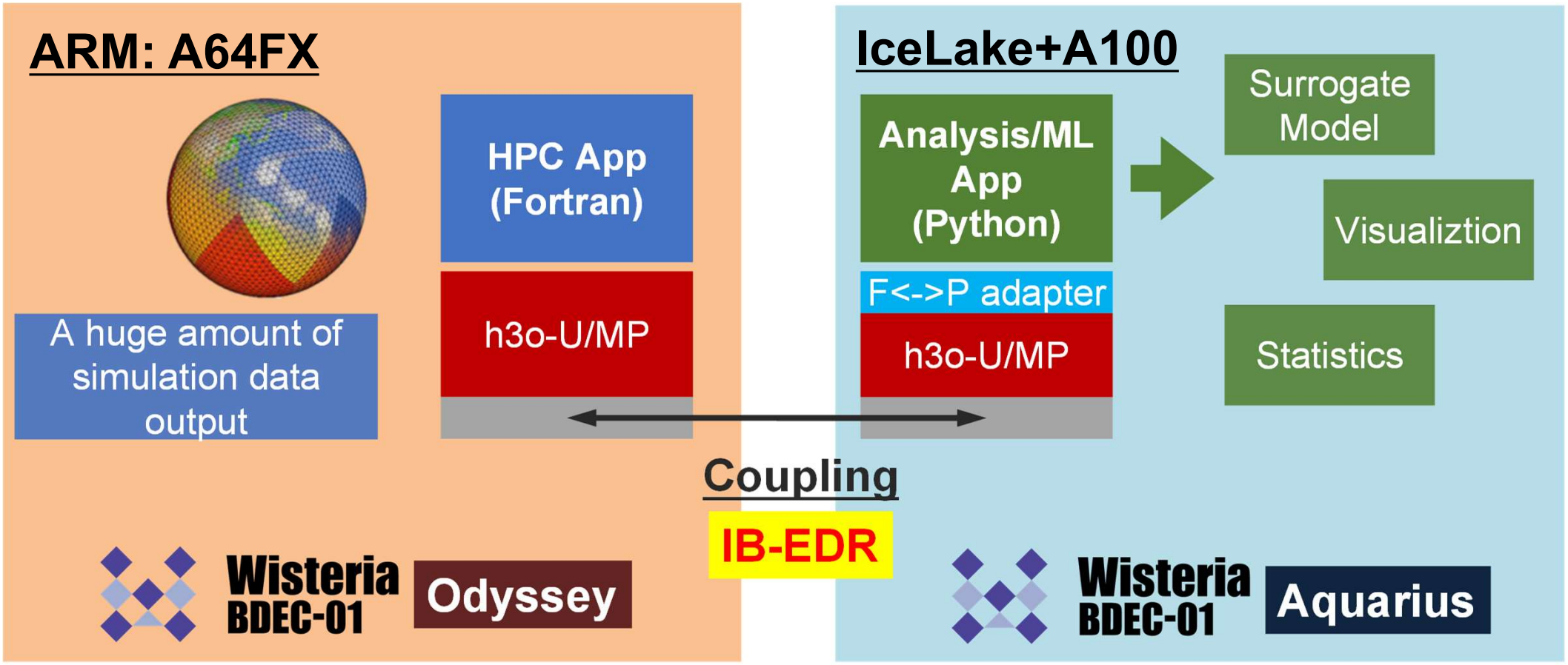
「計算＋データ＋学習」融合を支援する多機能カプラーh3-Open-UTIL/MP

- 異なる物理モデル連成のアンサンブル実行を支援・統合するための機能
  - MPI通信、時刻同期、格子系間マッピング等の管理機能の他、従来のカプラーには無い、複数の弱連成結合シミュレーションのアンサンブル実行、片側のモデルのみをアンサンブル実行する多対1の弱連成結合が可能
  - スパコン上で、全地球大気海洋連成シミュレーションによって動作検証済み
- Fortran/Cコード（物理モデル）とPythonコードの弱連成を実現する機能
  - FortranやCで記述されたプログラム同士の連成計算に限って開発を行ってきたカプラーを、Pythonによって記述されたAI・機械学習、可視化処理系のワークロードからも活用できるよう機能拡充。
- O-A利用：WaitIOとの連携

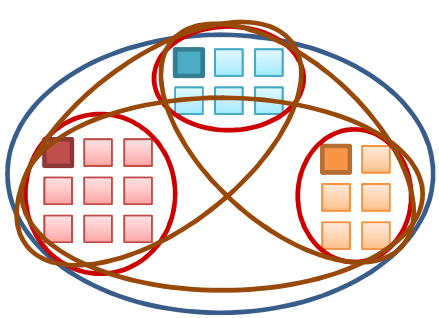


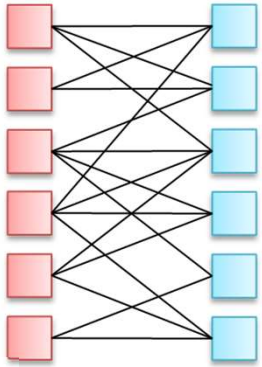
Fortran/CアプリとPythonアプリの連成計算の模式図
〔八代・荒川　2020〕

h3-Open-UTIL/MP (h3o-U/MP) + h3-Open-SYS/WaitIO

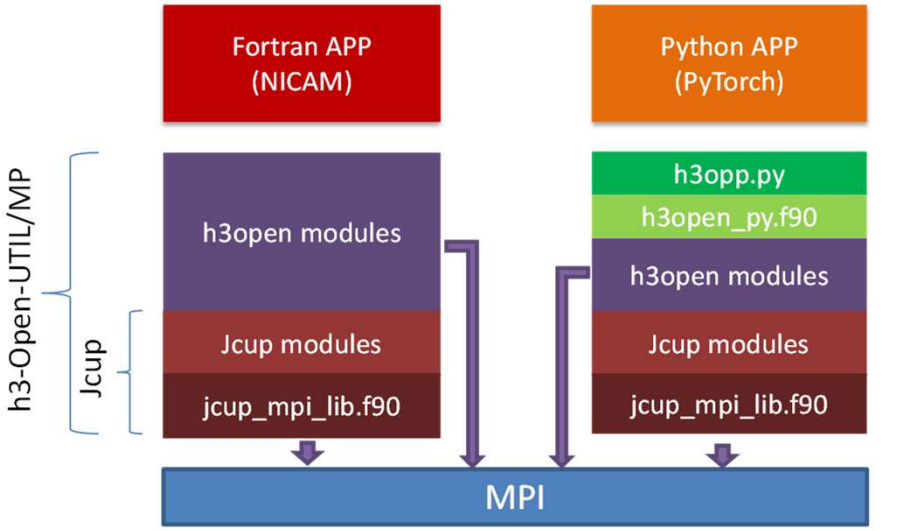


h3-Open-UTIL/MP・h3-Open-SYS/WaitIO連携

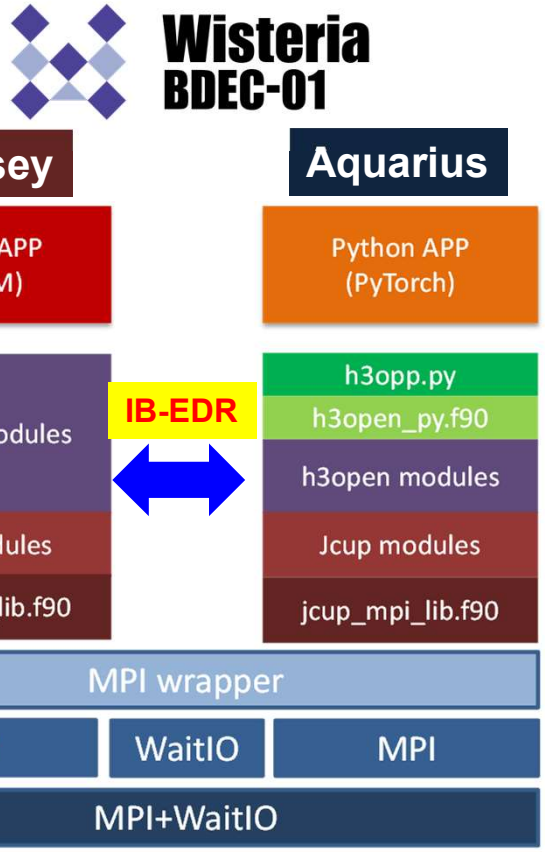




- h3-Open-UTIL/MP
  - （現状）MPIによるコンポーネント間通信：1対1，集団通信
  - Odyssey-Aquarius間はMPIによる通信は不可⇒ h3-Open-SYS/WaitIOによりO-A間通信実現

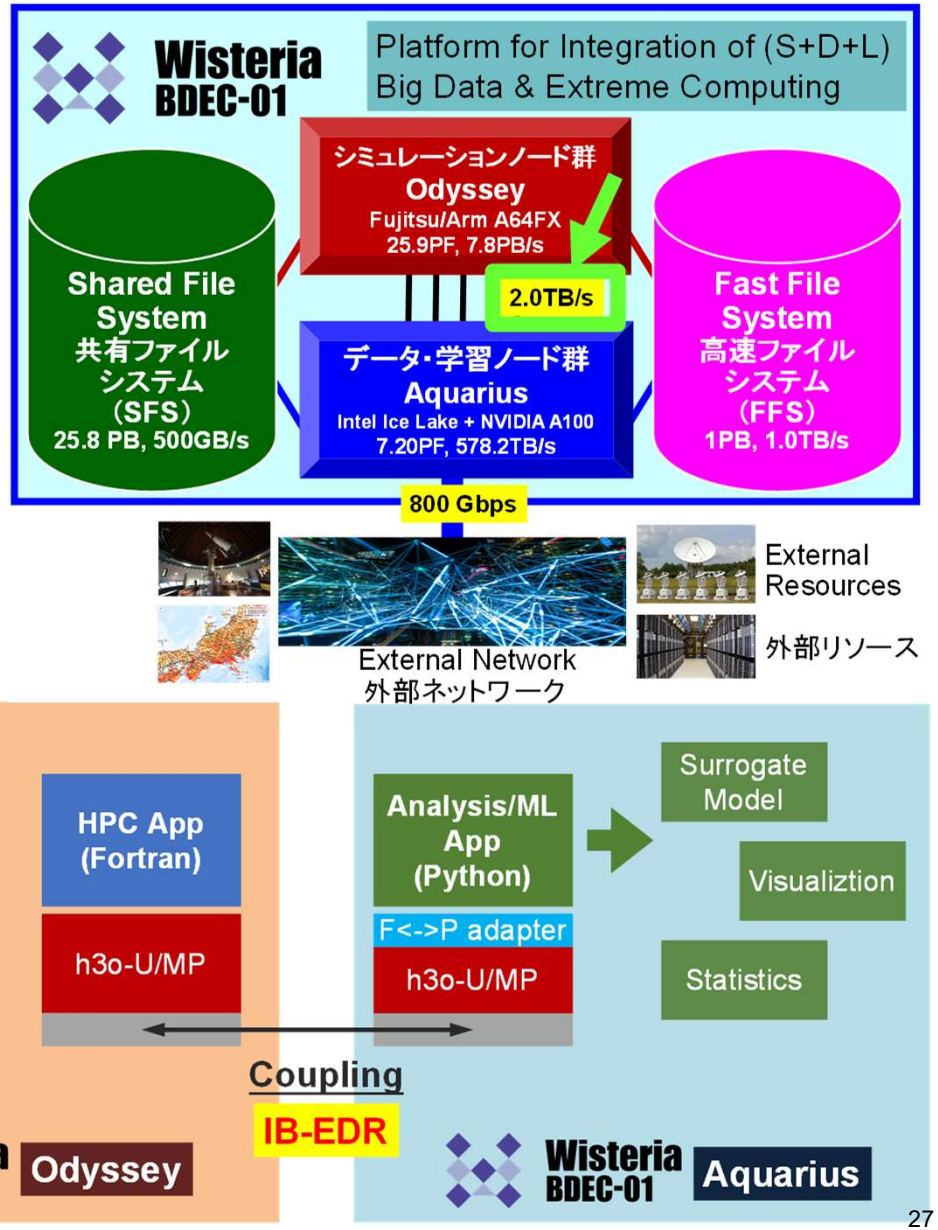


現状：MPI通信可能な環境を前提



整備・公開のスケジュール

- h3-Open-SYS/WaitIO
  - 2021年10月（Odyssey+Aquarius，直接通信）
  - 2022年度（O+A，FFS経由）
- h3-Open-UTIL/MP（HPC+Python）
  - 2021年10月（Oのみ）
- h3-Open-UTIL/MP+h3-Open-SYS/WaitIO
  - 2022年1月～4月（O+A，直接通信）
  - 2022年度（O+A，FFS経由）
- 協力者求む！
  - ユーザーアカウント＋トークン付与
  - 2022年度JHPCN新規応募



Wisteria BDEC-01

Hierarchical, Hybrid, Heterogeneous
h3-Open-BDEC
Big Data & Extreme Computing

参考リンク(ビデオ)

Wisteria
BDEC-01

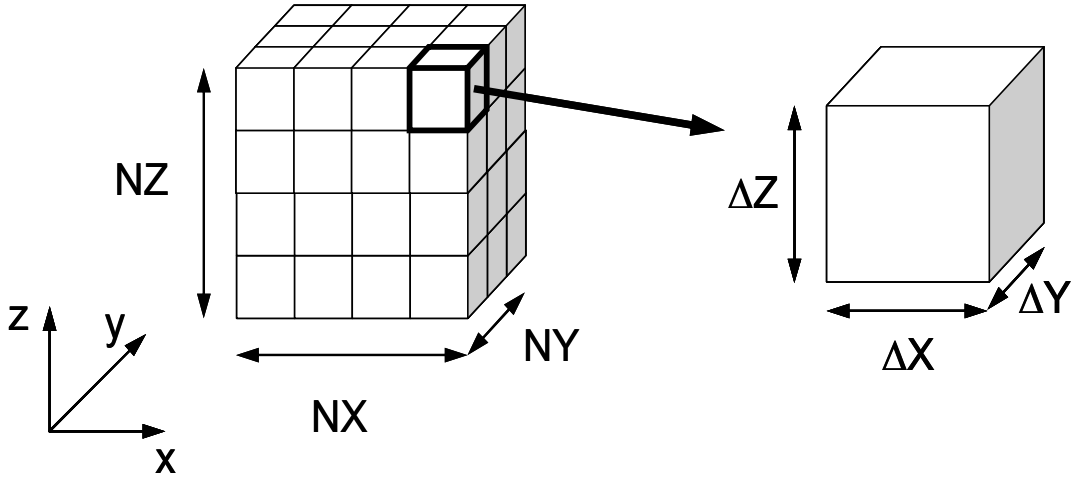
- Wisteria/BDEC-01利用説明会
  - https://www.youtube.com/watch?v=1bbZVO6-UQg
- h3-Open-BDEC:プロジェクトHP(工事中)
  - http://nkl.cc.u-tokyo.ac.jp/h3-Open-BDEC/
- Wisteria/BDEC-01 & h3-Open-BDEC紹介講演(日本語)
  - https://www.youtube.com/watch?v=CsJ_9aGNXCg
  - https://www.pccluster.org/ja/event/pccc20/exhibition/itc-u-tokyo.html
- Wisteria/BDEC-01 & h3-Open-BDEC紹介講演(英語)
  - https://www.youtube.com/watch?v=jX51NF2LniE

予備的計算事例

- **Odyssey**
  - **三次元熱伝導問題（OBCXとの比較）**
- FX700による評価（A64FX搭載）
  - 疎行列演算，半精度実数（FP16）〔中島他 HPC175, 2020〕
  - ステンシル計算〔星野他 HPC178，2021〕
- Aquarius
  - 宇宙物理シミュレーション〔三木，2021〕
    - P100，V100との比較

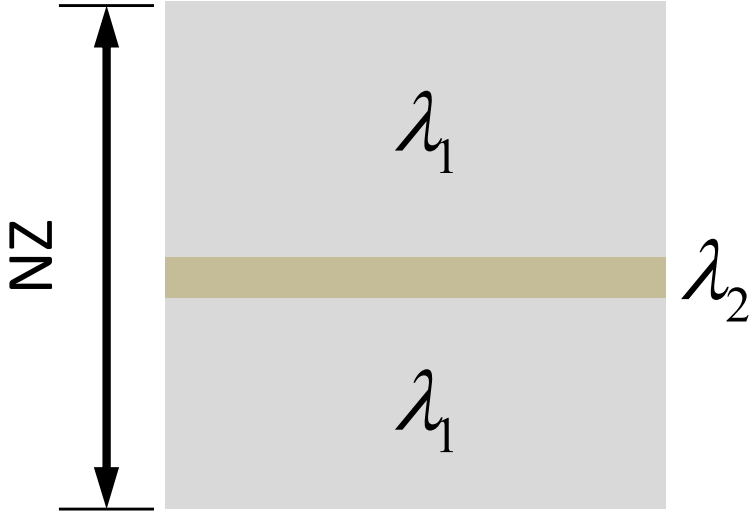
三次元定常熱伝導問題

$$\nabla \cdot \left(\lambda \nabla \phi \right) + f = 0$$



- 7点差分
- 有限体積法
- 本来不均質問題向けだがここでは $\lambda_1 = \lambda_2$ の均質な条件を仮定
- 係数行列：対称正定
- 前処理付き共役勾配法（PCG）
- Fortran 90 + OpenMP

$$\phi = 0 @ z = z_{\max}$$



前処理付き共役勾配法(PCG)

```
Compute r(0)= b-[A]x(0)
for i= 1, 2, …
    solve [M]z(i-1)= r(i-1)
    ρi-1= r(i-1) z(i-1)
    if i=1
      p(1)= z(0)
     else
      βi-1= ρi-1/ρi-2
      p(i)= z(i-1) + βi-1 p(i-1)
    endif
    q(i)= [A]p(i)
    αi  = ρi-1/p(i)q(i)
    x(i)= x(i-1) + αip(i)
    r(i)= r(i-1) - αiq(i)
    check convergence |r|
end
```

実際にやるべき計算は:

$$\{z\} = [M]^{-1}\{r\}$$

「近似逆行列」の計算が必要:

$$[M]^{-1} \approx [A]^{-1}, \quad [M] \approx [A]$$

究極の前処理:本当の逆行列

$$[M]^{-1} = [A]^{-1}, \quad [M] = [A]$$

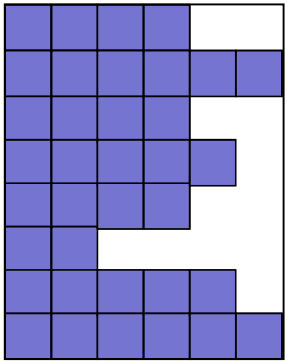
対角スケーリング:簡単＝弱い

$$[M]^{-1} = [D]^{-1}, \quad [M] = [D]$$

疎行列の格納形式

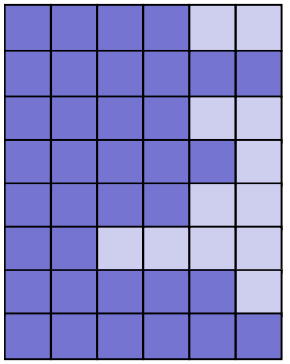
CRS（Compressed Row Storage）

```
do i= 1, N
  W(i,Q)= D(i)*W(i,P)
  do k= indexLU(i-1)+1, indexLU(i)
    W(i,Q)= W(i,Q) + AMAT(k)*W(itemLU(k),P)
  enddo
enddo
```



ELL (ELLPACK/ITPACK)

```
do i= 1, N
  W(i,Q)= D(i)*W(i,P)
  do j= 1, 6
    W(i,Q)= W(i,Q) + AMAT(j,i)*W(itemLU(j,i),P)
  enddo
enddo
```



- CRS: Compressed Row Storage
  - 非ゼロ非対角成分のみ格納⇒メモリ節約できるが，計算効率悪い
- ELL：ELLPACK/ITPACK
  - 非ゼロ非対角成分数固定⇒0のところには0を入れる
  - 記憶容量，計算量は増えるがメモリアクセス性能向上⇒Prefetch

```
[XYZ@wisteria01 run]$ cd ../src-f0
[XYZ@wisteria01 src-f0]$ diff poi_gen.f ../src-f1/poi_gen.f
25,29c25,31
```

src-f0

```
<       PHI    = 0.d0
<       BFORCE= 0.d0
<         D    = 0.d0
<
<       INLU= 0
---
```

src-f1

```
> !$omp parallel do private (icel)
>       do icel= 1, ICELTOT
>         PHI    (icel)= 0.d0
>         BFORCE(icel)= 0.d0
>         D      (icel)= 0.d0
>         INLU   (icel)= 0
>       enddo
```

	Thread #	sec	Speed-up	Parallel Efficiency (%)
src-f0	12	5.27	12.00	100.00
	24	2.78	22.72	94.68
	36	1.95	32.49	90.24
	48	1.60	39.54	82.38
src-f1	12	5.26	12.00	100.00
	24	2.70	23.39	97.45
	36	1.86	33.83	93.96
	48	1.44	43.77	91.18

```
71,72c73,75
```

src-f0

```
<       indexLU= 0
<
<       do icel= 1, ICELTOT
<         indexLU(icel)= INLU(icel)
<       enddo
---
```

src-f1

```
>       indexLU(0)= 0
>
> !$omp parallel do private (icel)
>       do icel= 1, ICELTOT
>         indexLU(icel)= INLU(icel)
>       enddo
```

```
85,86c88,94
```

src-f0

```
<       itemLU= 0
<         AMAT= 0.d0
---
```

src-f1

```
> !$omp parallel do private (icel,k)
>       do icel= 1, ICELTOT
>       do k= indexLU(icel-1)+1, indexLU(icel)
>         itemLU(k)= 0
>         AMAT   (k)= 0.d0
>       enddo
>       enddo
```

CRS：src-f0（First Touch ナシ），src-f1（同アリ）

PCG法計算時間

NX=NY=NZ=128

PCG法計算時間（48スレッド，N=128³, 826反復）

Odyssey：最適化を施さないと性能が出ない，ナイーヴな実装ではむしろOBCXが速い

	OBCX	Odyssey
src0（初期設定）	1.738	1.779
src1（First Touch）	1.217	1.453
src2（+ELL）	1.189	0.760
src3（+omp-parallel削減）	1.180	0.665
src4(+omp-do無し，reduction無し)	1.189	0.657

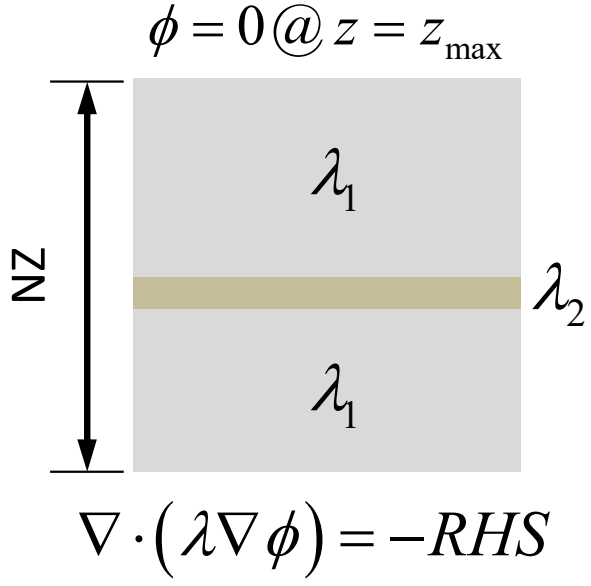
	OBCX	Odyssey
CPU名称	Intel Xeon Platinum 8280 (Cascade Lake, CLX)	Fujitsu A64FX（2.2 GHz）
コア数	56	48
理論演算性能（GFLOPS）	4,838	3,379
主記憶容量（GB）	192	32
メモリ性能（GB/sec）	282	1,024

予備的計算事例

- Odyssey
  - 三次元熱伝導問題（OBCXとの比較）
- **FX700による評価（A64FX搭載）**
  - **疎行列演算，半精度実数（FP16）〔中島他 HPC175, 2020〕**
  - ステンシル計算〔星野他 HPC178，2021〕
- Aquarius
  - 宇宙物理シミュレーション〔三木，2021〕
    - P100，V100との比較

FX700（A64FX搭載）による混合精度演算

- 三次元不均質場熱伝導方程式，FVM，ICCG法
- 倍精度（FP64），単精度（FP32），半精度（FP16）
- 反復法の前処理とそれ以外（SpMV，DAXPY，内積）
- 「それ以外」（SpMV，DAXPY，内積）
  - 倍精度・単精度
- 前処理
  - 倍精度・単精度・半精度
- 半精度前処理
  - ベクトルは単精度にしないと破綻した
  - 係数行列のみ半精度となっている
- λ_1/λ_2 を変化させる，CRS，128^3



〔中島他 SWoPP 2020〕

GFLOPS（ピーク性能）当たり利用負担（円）：電気代 GFLOPS/W（Green 500）

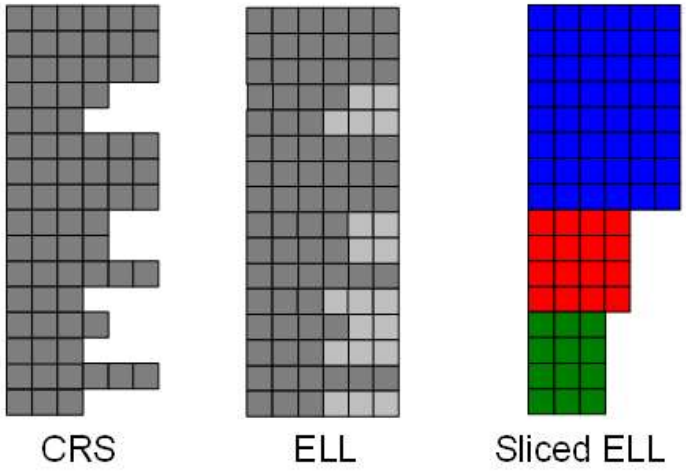
System	JPY/GFLOPS Small is Good	GFLOPS/W Large is Good
Oakleaf-FX/Oakbridge-FX (Fujitsu) (Fujitsu SPARC64 IXfx)	125	**0.866**
Reedbush-U (HPE)　(Intel Xeon Broadwell (BDW))	61.9	**2.310**
Reedbush-H (HPE)　(Intel BDW+NVIDIA P100x2/node)	15.9	**8.575**
Reedbush-L (HPE)　(Intel BDW+NVIDIA P100x4/node)	13.4	**10.167**
Oakforest-PACS (Fujitsu)　(Intel Xeon Phi/KNL)	16.5	**4.986**
Oakbridge-CX (Fujitsu)　(Intel Xeon Cascade Lake)	20.7	**5.076**
Wisteria-Odyssey（Fujitsu/Arm A64FX）	17.8	**15.069**
Wisteria-Aquarius（Intel Xeon Ice Lake + NVIDIA A100x8）	9.00	**24.058**

システム名	Oakforet-PACS	Oakbridge-CX	Oakleaf-7（FX700）
略称	OFP	OBCX	OL7
CPU名称	Intel Xeon Phi 7250 (Knights Landing, KNL)	Intel Xeon Platinum 8280 (Cascade Lake, CLX)	Fujitsu A64FX（1.8GHz）
コア数/ソケット	68	28	48
ソケット数/ノード	1	2	1
理論演算性能（GFLOPS）/ノード	3,046	4,838	2,765
主記憶容量（GB）/ノード	MCDRAM: 16 DDR4: 96	192	32
メモリ性能（GB/sec）/ノード STREAM Triad	MCDRAM: 490 DDR4: 84.5	202	809
コンパイラ	Intel Parallel Studio 2019		Fujitsu FCC 4.0.0

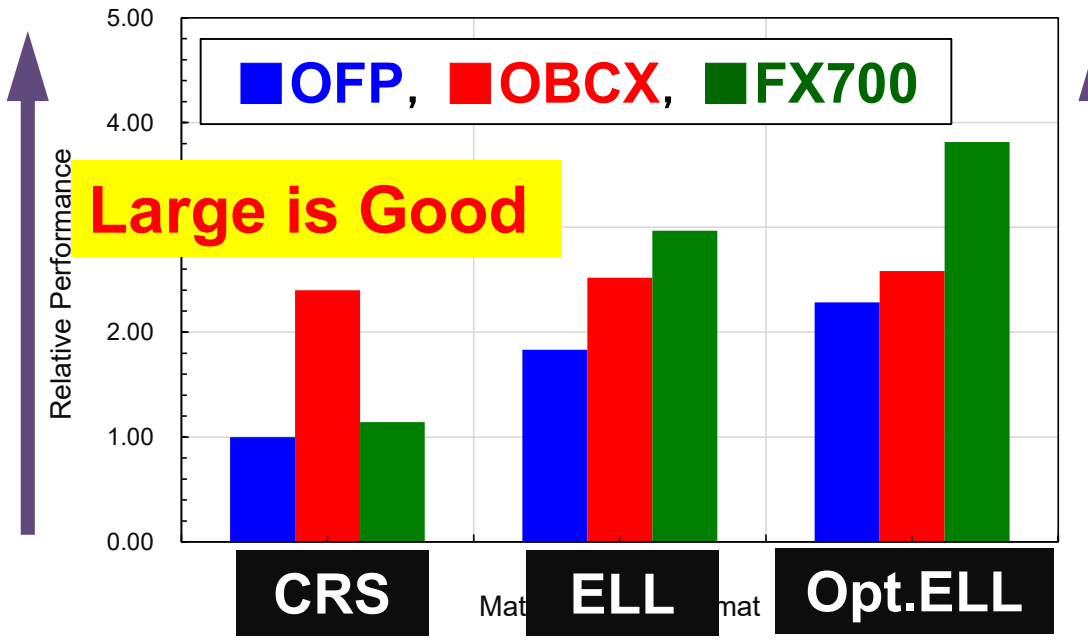
ICCG法計算時間比による性能比（Small is Good，倍精度）

OFP・CRSで無次元化，$\lambda_1 / \lambda_2 = 1$

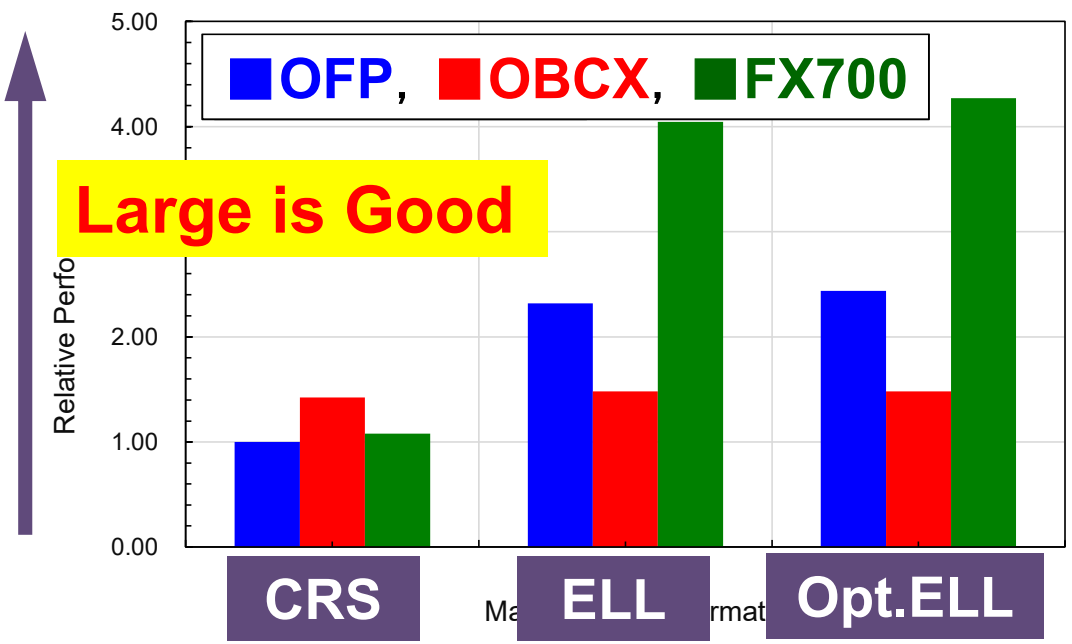
〔中島他 SWoPP 2020〕



Medium：128^3



Large：256^3



P3D ICCG法ソルバー混合精度演算・FX700

	反復法主要部	前処理	前処理部分 ベクトル
D-D	FP64	FP64	FP64
D-S	FP64	FP32	FP32
D-H	FP64	FP16	FP32
S-S	FP32	FP32	FP32
S-H	FP32	FP16	FP32

[KN et al. SWoPP 2020]

P3D ICCG法ソルバー混合精度演算・FX700 前進代入部(CRS)実装

[KN et al. SWoPP 2020]

FP64
FP32
FP16

D-S

```
!$omp parallel do private(ip,i)
      do ip= 1, PEsmpTOT
      do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
        Ws(I,Z)= W(I,R)
      enddo
      enddo

!$omp parallel private(ic,ip,ip1,I,WVALs,k)
      do ic= 1, NCOLORtot
!$omp do
        do ip= 1, PEsmpTOT
          ip1= (ip-1)*NCOLORtot + ic
        do i= SMPindex(ip1-1)+1, SMPindex(ip1)
          WVALs= Ws(i,Z)
          do k= indexL(i-1)+1, indexL(i)
            WVALs= WVALs -  ALs(k) * Ws(itemL(k),Z)
          enddo
          Ws(i,Z)= WVALs * Ws(i,DD)
        enddo
        enddo
      enddo
!$omp end parallel

(Backward Substitution)

!$omp parallel do private(ip,i)
      do ip= 1, PEsmpTOT
      do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
        W(I,Z)= Ws(I,Z)
      enddo
      enddo
```

S-H

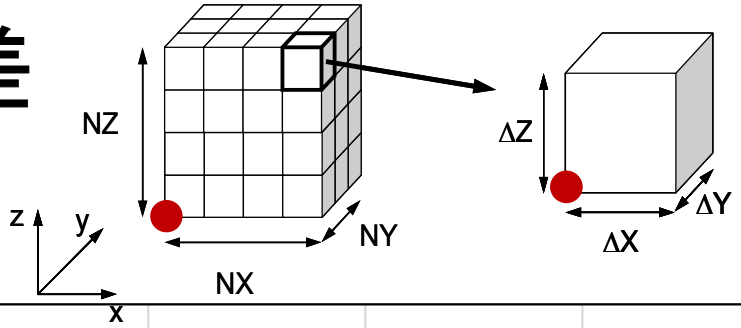
```
!$omp parallel do private(ip,i)
      do ip= 1, PEsmpTOT
      do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
        Ws(i,Z)= Ws(i,R)
      enddo
      enddo

!$omp parallel private(ic,ip,ip1,i,WVALs,k)
      do ic= 1, NCOLORtot
!$omp do
        do ip= 1, PEsmpTOT
          ip1= (ip-1)*NCOLORtot + ic
        do i= SMPindex(ip1-1)+1, SMPindex(ip1)
          WVALs= Ws(i,Z)
          do k= indexL(i-1)+1, indexL(i)
            WVALs= WVALs -  ALh(k) * Ws(itemL(k),Z)
          enddo
          Ws(i,Z)= WVALs * Wh(i,DD)
        enddo
        enddo
      enddo
!$omp end parallel
```

D-H

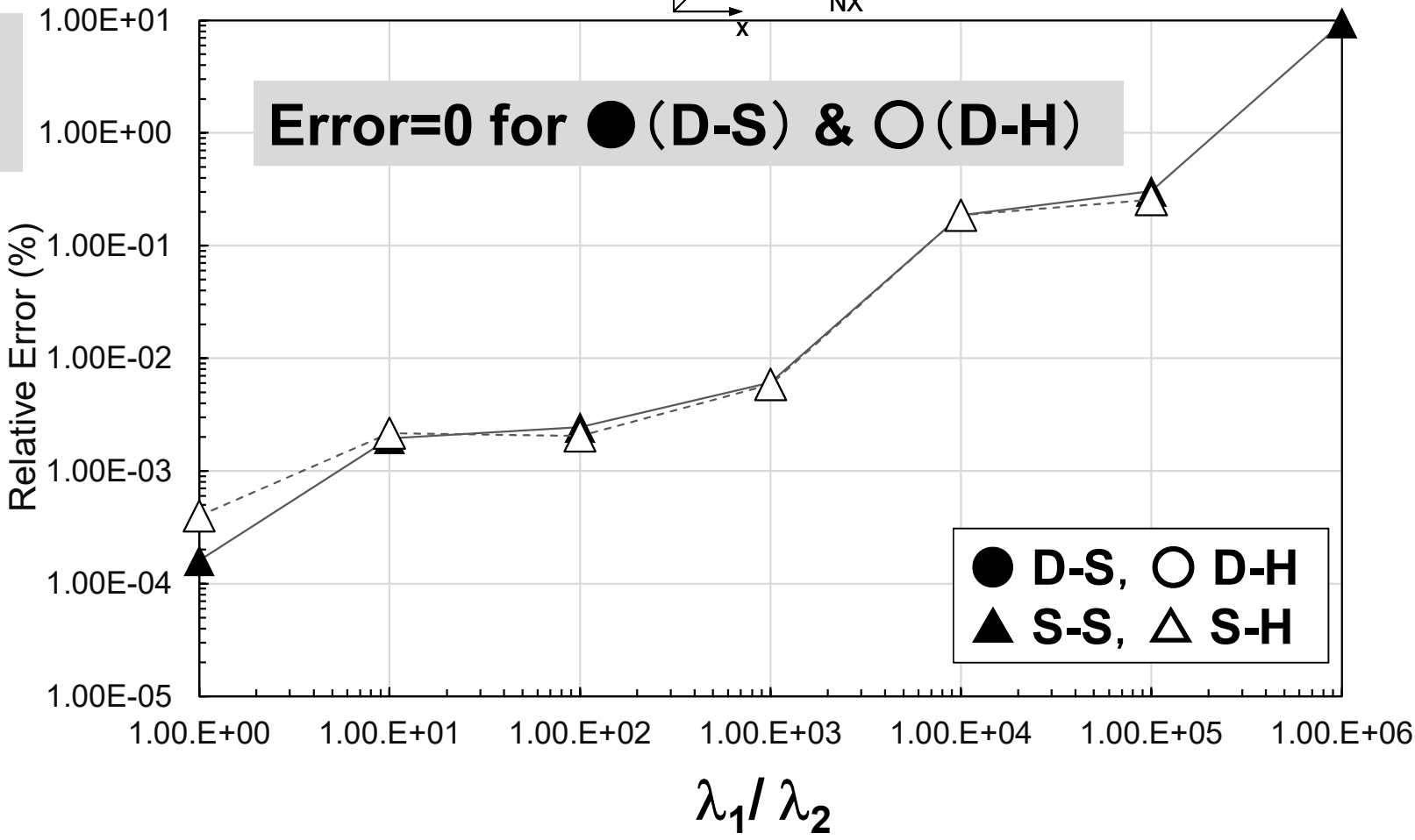
```
!$omp parallel do private(ip,i)
      do ip= 1, PEsmpTOT
      do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
        Ws(i,Z)= W(i,R)
      enddo
      enddo

!$omp parallel private(ic,ip,ip1,i,WVALs,k)
      do ic= 1, NCOLORtot
!$omp do
        do ip= 1, PEsmpTOT
          ip1= (ip-1)*NCOLORtot + ic
        do i= SMPindex(ip1-1)+1, SMPindex(ip1)
          WVALs= Ws(i,Z)
          do k= indexL(i-1)+1, indexL(i)
            WVALs= WVALs -  ALh(k) * Ws(itemL(k),Z)
          enddo
          Ws(i,Z)= WVALs * Wh(i,DD)
        enddo
        enddo
      enddo
!$omp end parallel

(Backward Substitution)

!$omp parallel do private(ip,i)
      do ip= 1, PEsmpTOT
      do i= SMPindex((ip-1)*NCOLORtot)+1, SMPindex(ip*NCOLORtot)
        W(I,Z)= Ws(I,Z)
      enddo
      enddo
```

混合精度演算計算結果：相対誤差

D-H，S-Hは$\lambda_1/\lambda_2 = 10^6$で収束せず



●点におけるD-D
との相対誤差（%）



[KN et al. SWoPP 2020]

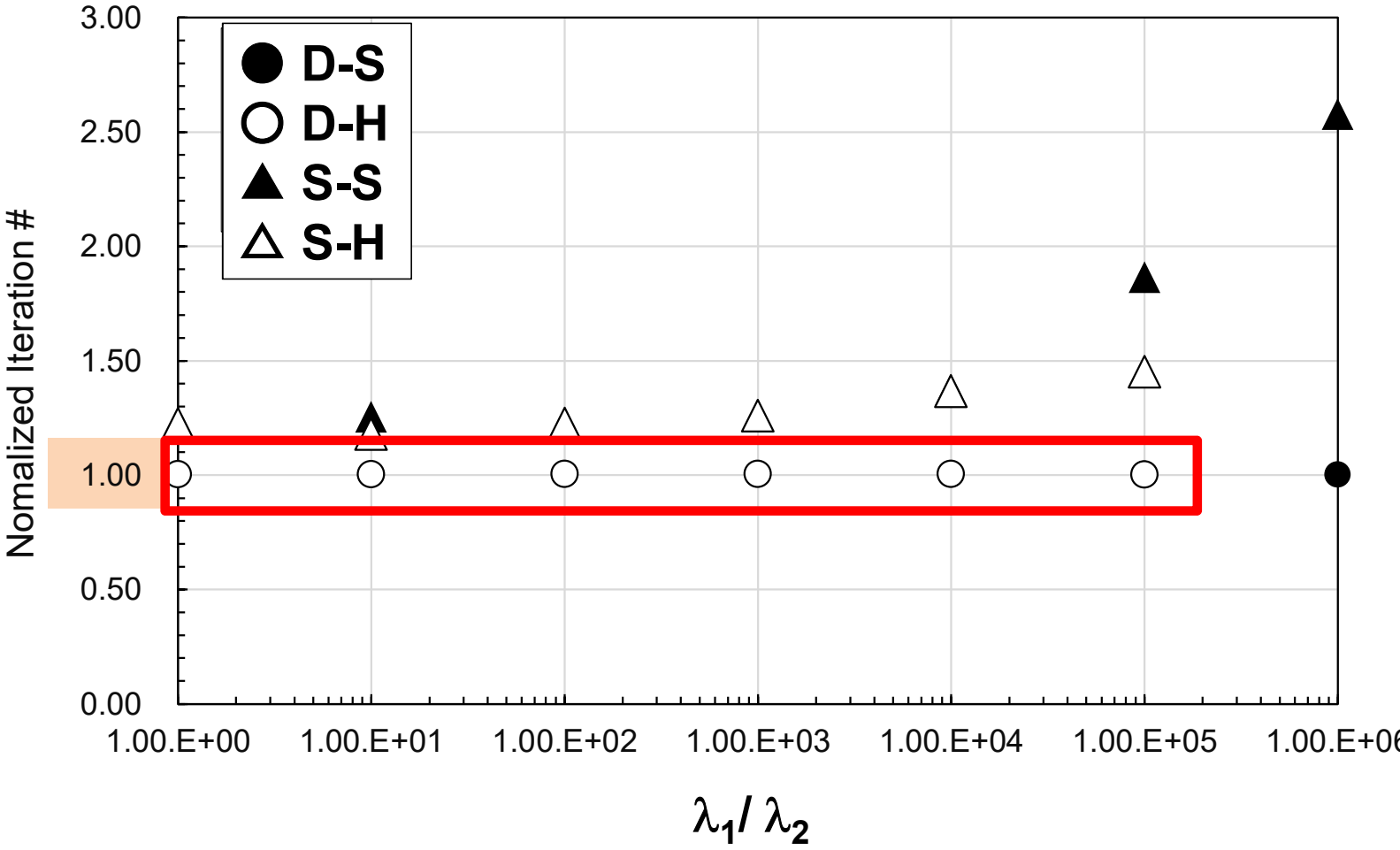
混合精度演算計算結果：反復回数

[KN et al. SWoPP 2020]

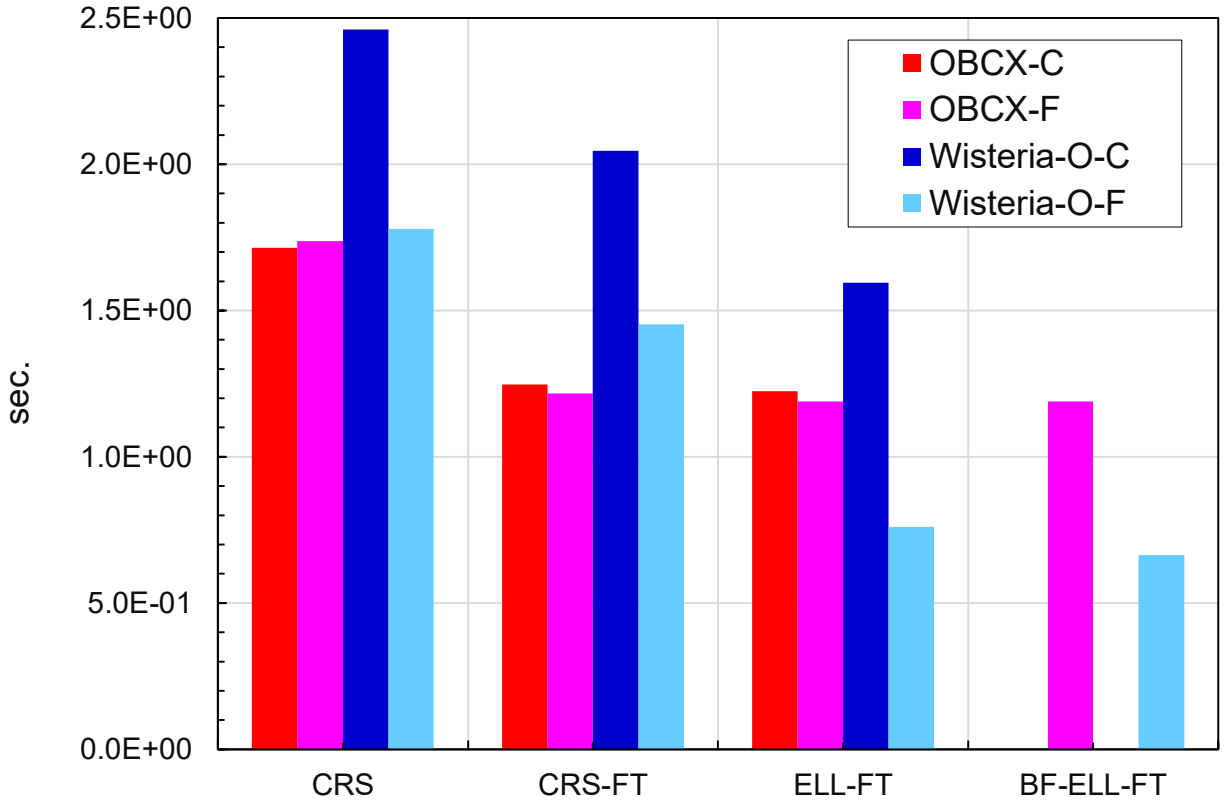
D-H，S-Hは$\lambda_1/\lambda_2 = 10^6$で収束せず

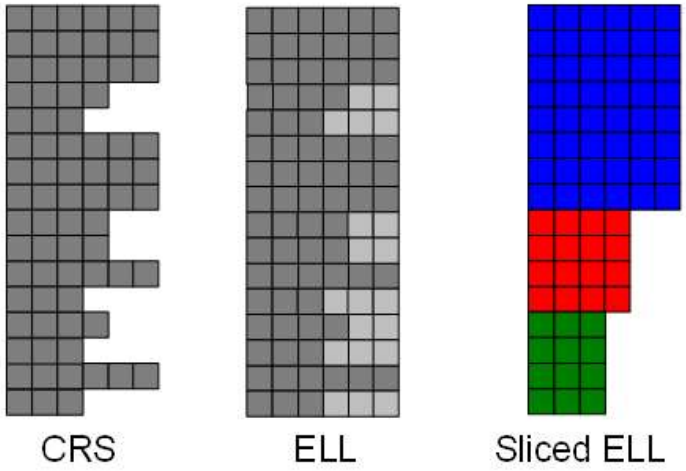
D-Dの反復回数で無次元化
●～○～D-D，▲～△

(D-S，D-H) の結果はD-Dと一致 (if $\lambda_1/\lambda_2 \leqq 10^5$)



7点ステンシルPoisson方程式 PCG法，N=128^3 1ノード48コア，4CMGまたぎ





- CRS
- CRS-FT: First Touch
- ELL-FT
- BF-ELL-FT: Barrier Free
- OBCXではC■，Fortran■とで同様の傾向
- Wisteria-OではC■が極遅
  - -Kfast,openmp

予備的計算事例

- Odyssey
  - 三次元熱伝導問題（OBCXとの比較）
- **FX700による評価（A64FX搭載）**
  - 疎行列演算，半精度実数（FP16）〔中島他 HPC175, 2020〕
  - **ステンシル計算〔星野他 HPC178，2021〕**
- Aquarius
  - 宇宙物理シミュレーション〔三木，2021〕
    - P100，V100との比較

FX700 vs. CLX vs. KNL (7点ステンシル)

- 拡散方程式のカーネル
  - 自身と前後左右上下の7点を参照して値を更新
- ディリクレ境界条件
  - 境界では自身の値を参照
- 倍精度
- サイズ：256*384*384
  - A64FXが48コアのため、48*8 = 384
- Byte/Flop = 16/13
  - 各ステップで読み込んだデータは全てキャッシュに乗ると仮定して性能評価
- 7点ステンシルへよく知られた一連の最適化を適用
  - 基本的には直前までの実装で速かったものに最適化を加算

〔星野他 HPC178，2021〕

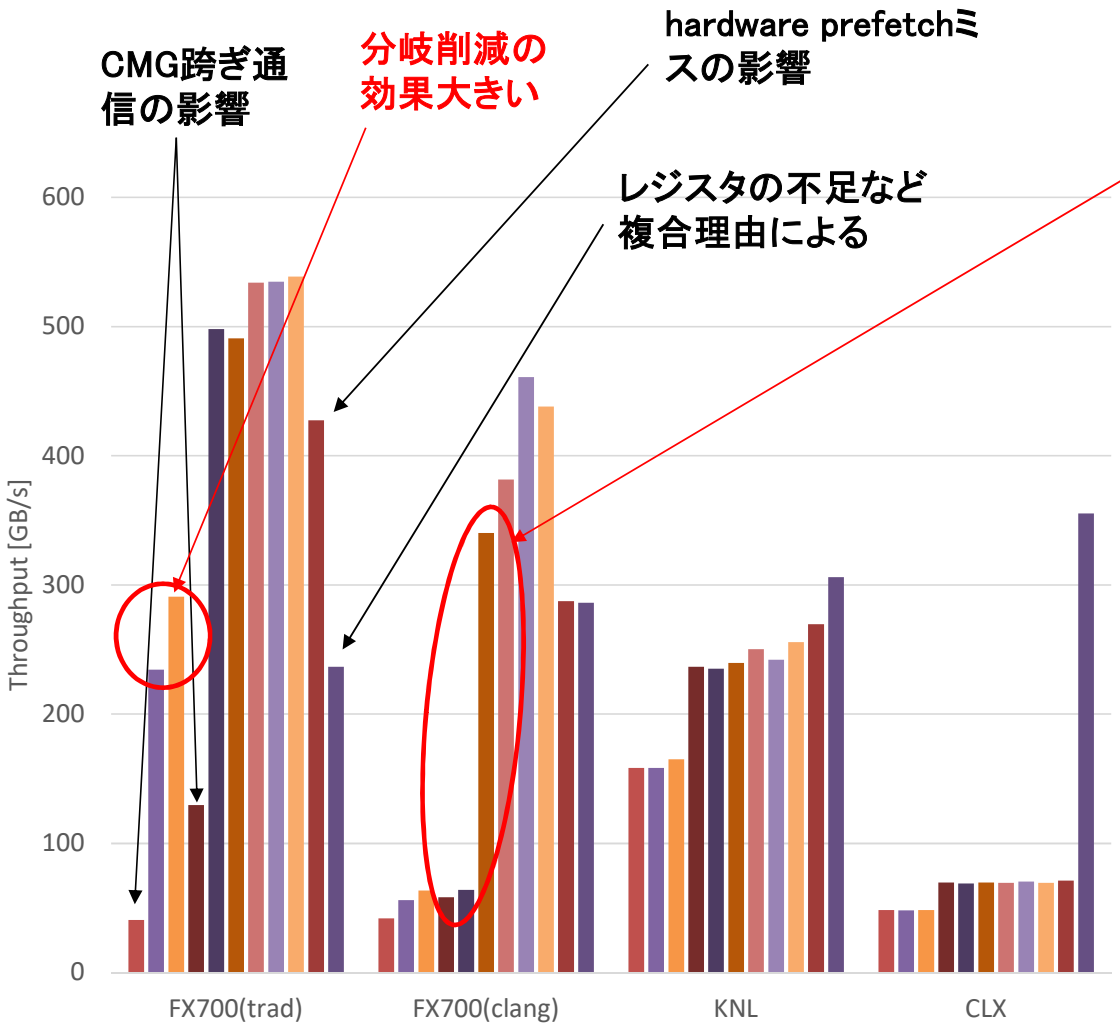
→ 7点ステンシルのプログラム

```
do {
#pragma omp parallel for
   for (int z = 0; z < nz; z++) {
     for (int y = 0; y < ny; y++) {
       for (int x = 0; x < nx; x++) {
         int c =  x + y * nx + z * nx * ny;
         int w = (x == 0)    ? c : c - 1;
         int e = (x == nx-1) ? c : c + 1;
         int n = (y == 0)    ? c : c - nx;
         int s = (y == ny-1) ? c : c + nx;
         int b = (z == 0)    ? c : c - nx * ny;
         int t = (z == nz-1) ? c : c + nx * ny;
         f2_t[c] = cc * f1_t[c]
            + cw * f1_t[w] + ce * f1_t[e]
            + cs * f1_t[s] + cn * f1_t[n]
            + cb * f1_t[b] + ct * f1_t[t];
       }
     }
   }
   double *tmp = f1_t;
   f1_t = f2_t;
   f2_t = tmp;
   time += dt;
} while (time + 0.5*dt < 0.1);
```

↓ 計算環境

略　称	FX	KNL	CLX
名称	Fujitsu A64FX (1.8GHz)	Intel Xeon Phi 7250 (Knights Landing)	Intel Xeon Platinum 8280 (CascadeLake)
コア数	48	68	28
理論演算性能	2,765 GFLOPS	3,046 GFLOPS	2,419 GFLOPS
主記憶容量	32 GB	MCDRAM: 16 GB DDR4: 96 GB	96 GB
STREAM Triad性能	809 GB/sec	MCDRAM: 490 GB/sec DDR4: 80.1 GB/sec	101 GB/sec

FX700 vs. CLX vs. KNL (7点ステンシル)



	最適化内容
BASE	ベースライン実装
1stT	ファーストタッチ
PEEL	loop peeling (branch hoisting) により最内ループでの分岐の削減
Ydim	Yループに #pragma omp for nowait を適用
Y-Zdim	ZループをCMG分割、YループをCMG内スレッド分割
Intrin	Intrinsics を用いた実装
w/oFMA	Fused Multiply Addを敢えて使わず、MULとADDのみを用い、パイプライン処理を効率化
UNR	ループアンローリングによりパイプライン処理の効率化
REG	レジスタブロッキング
TILE	タイリングによるキャッシュ利用の効率化
TB	テンポラルブロッキング

〔星野他 HPC178，2021〕

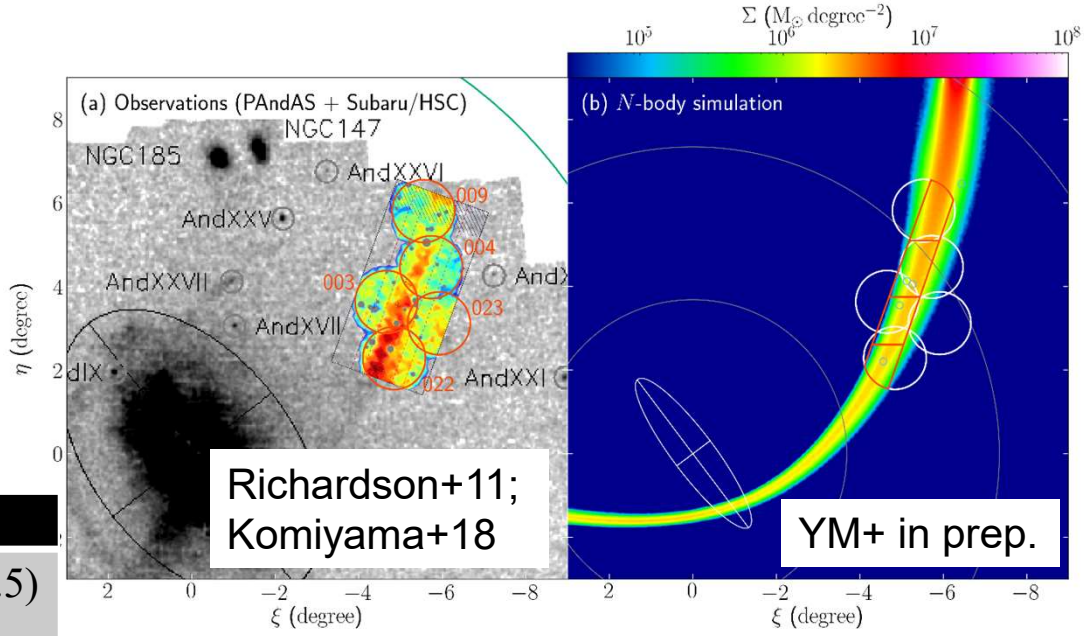
予備的計算事例

- Odyssey
  - 三次元熱伝導問題（OBCXとの比較）
- FX700による評価（A64FX搭載）
  - 疎行列演算，半精度実数（FP16）〔中島他 HPC175, 2020〕
  - ステンシル計算〔星野他 HPC178，2021〕
- **Aquarius**
  - **宇宙物理シミュレーション〔三木，2021〕**
    - **P100，V100との比較**

A100上での性能評価：重力ツリーコード（1/2）

- GPU向けに最適化された重力ツリーコードGOTHIC（実装はCUDA）
  - Miki & Umemura (2017), New Astronomy, 52, 65
  - Miki (2019), International Conference on Parallel Processing (ICPP 2019)
- ほとんどの浮動小数点演算は単精度
- 幅優先探索を採用
- 動的な最適化を実装
- 各世代のGPU向けの詳細な調整
  - A100ではL2キャッシュ制御を追加
  - テンソルコア（TC）は使用していない

GPU名称	P100	V100	A100
理論演算性能（TFLOPS）	DP: 5.3 SP: 10.6	DP: 7.8 SP: 15.7	DP: 9.7 (+TCで 19.5) SP: 19.5
主記憶容量（GB）	16	32	40
メモリ性能（GB/sec）	732	900	1,555

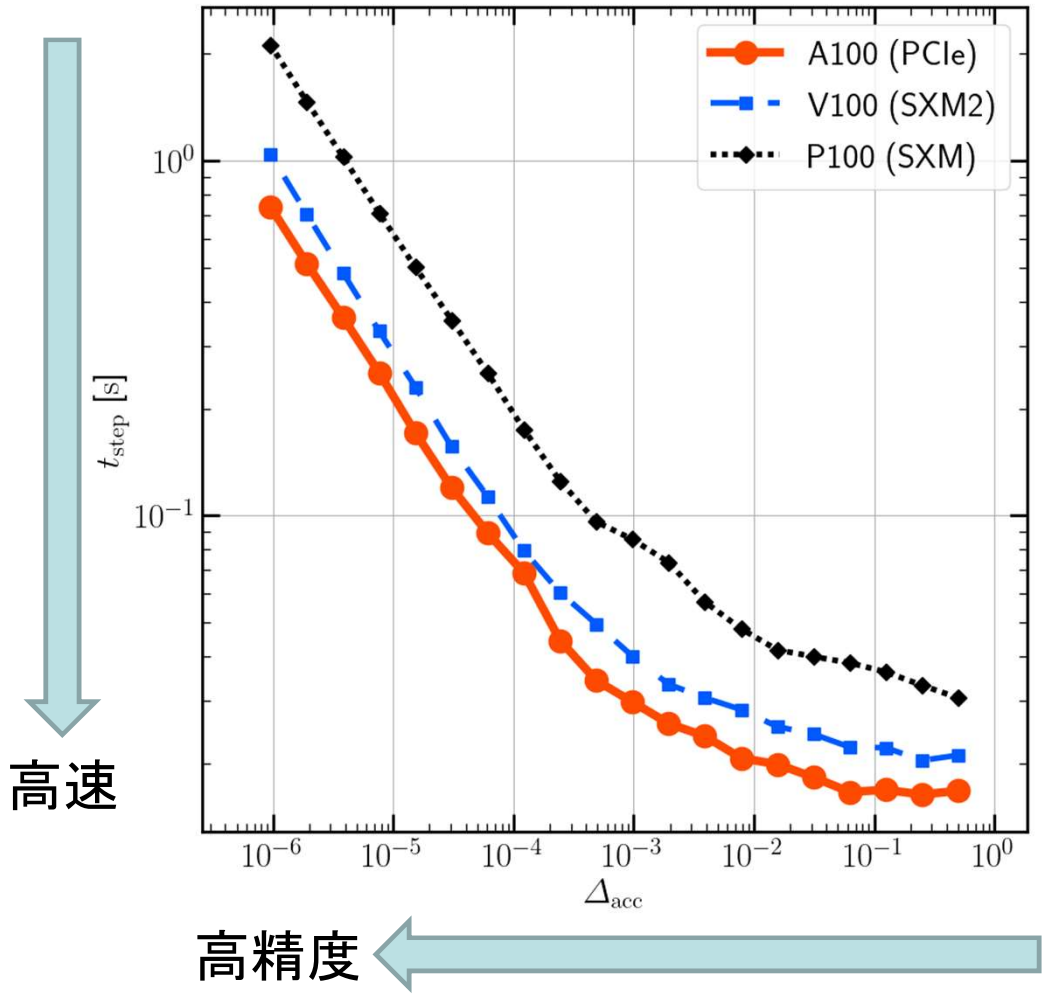


〔三木　2021〕

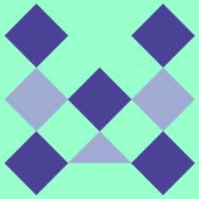
A100上での性能評価：重力ツリーコード（2/2）

- A100（PCIe版）上での測定結果
  - M31 model, N = 8M
  - 4096ステップの計算時間を測定
- 理論ピーク性能比よりも高速！
  - P100（SXM：Reedbush-H/L搭載）よりも2.6倍高速
    （理論ピーク性能比は1.8倍）
  - V100（SXM2）よりも1.3倍高速
    （理論ピーク性能比は1.2倍）
  - SXM4版（Wisteria-A搭載）はPCIe版よりも1割程度高速と期待される

〔三木　2021〕



技術的な特徴など



Wisteria
BDEC-01

- Odyssey
  - SVE (Scalable Vector Extension)
    - Armv8-A命令セットアーキテクチャをスーパーコンピュータ向けに拡張
  - FP16
  - 機械学習・AIワークロードへの適用
- Aquarius
  - HPC・計算科学への適用
  - CPU：Intel Xeon Ice Lake
    - 3rd Generation Intel Xeon Scalable Processors
    - 推論，単独での利用は難しいが
  - GPU：NVIDIA A100 Tensor Core
    - Tensor Core + Tensor Float [TF32]
- Odyssey-Aquarius
  - InfiniBand-EDR