

共通基盤ライブラリ・環境 SCALE および SCALE-LES について

西澤 誠也, Team SCALE

理化学研究所・計算科学研究機構



- ◆ **SCALE** (Scalable Computing for Advanced Library and Environment)
 - ◆ 地球・惑星における気象・気候モデルのための共通基盤ライブラリ・環境
- ◆ **SCALE-LES**
 - ◆ 広領域・高解像度実験のための **LES** モデル

SCALE LIBRARY

背景

- ◆ 複数グループによる、(独立した) モデル開発
 - ◆ 多様性/独立性が確保
 - ◆ 比較による信頼性の評価
 - ◆ 非効率/有限リソースを有効に利用出来ていない
 - ◆ モデル開発ができる人材が不足している
 - ◆ (評価システムにも問題有り)
 - ◆ モデル間に統一性がなく、比較が容易でない
 - ◆ ファイルフォーマット/変数名等が異なる
 - ◆ モデル構成が異なる

- ◆ 関係コンポーネントの精緻化・多様化
 - ◆ 物理過程はどんどん増加、複雑化
 - ◆ 組み合わせによっても結果が異なる
 - ◆ それぞれの専門家が、それぞれのモデル用にコーディング

- ◆ 計算機の高度化
 - ◆ アーキテクチャの複雑化
 - ◆ メモリ/IO多階層構造
 - ◆ ノード構成・インターコネクットの複雑化
 - ◆ ヘテロジニアス構成
 - ◆ Byte/Flop値の低下
- ◆ 特定計算機への最適化
 - ◆ コードが読みにくい
 - ◆ バージョンスプリット問題

計算機の専門家との協力が必須に

そこで

共通基盤ライブラリの開発

- ◆ 可能な部分を共通化
 - ◆ 人的・時間的リソースの有効活用
 - ◆ API の共通化
 - ◆ それぞれが開発したコンポーネントは、お互い利用可能
 - ◆ 「口」をそろえることで、独自コンポーネントも共存可能
 - ◆ モデル構造が似ることによって、相互可読性の向上
 - ◆ ドキュメント・知見
 - ◆ IO の共通化
 - ◆ 結果比較を容易に
 - ◆ オープンソース
 - ◆ どのモデル・グループでも利用可能
 - ◆ (特定)計算機最適化コードの共同利用
- ◆ 計算機科学の専門家と協調 (Team SCALE)

Team SCALE

- ◆ Transdisciplinary Research Team in RIKEN AICS



**Computational
Climate Science
Research Team**

TL : H. Tomita



**System Software
Research Team**

TL : Y. Ishikawa



**Programming
Environment
Research Team**

TL : M. Sato



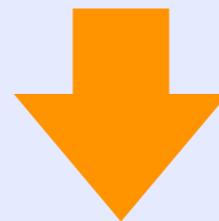
**HPC Programing
Framework
Research Team**

TL : N. Maruyama

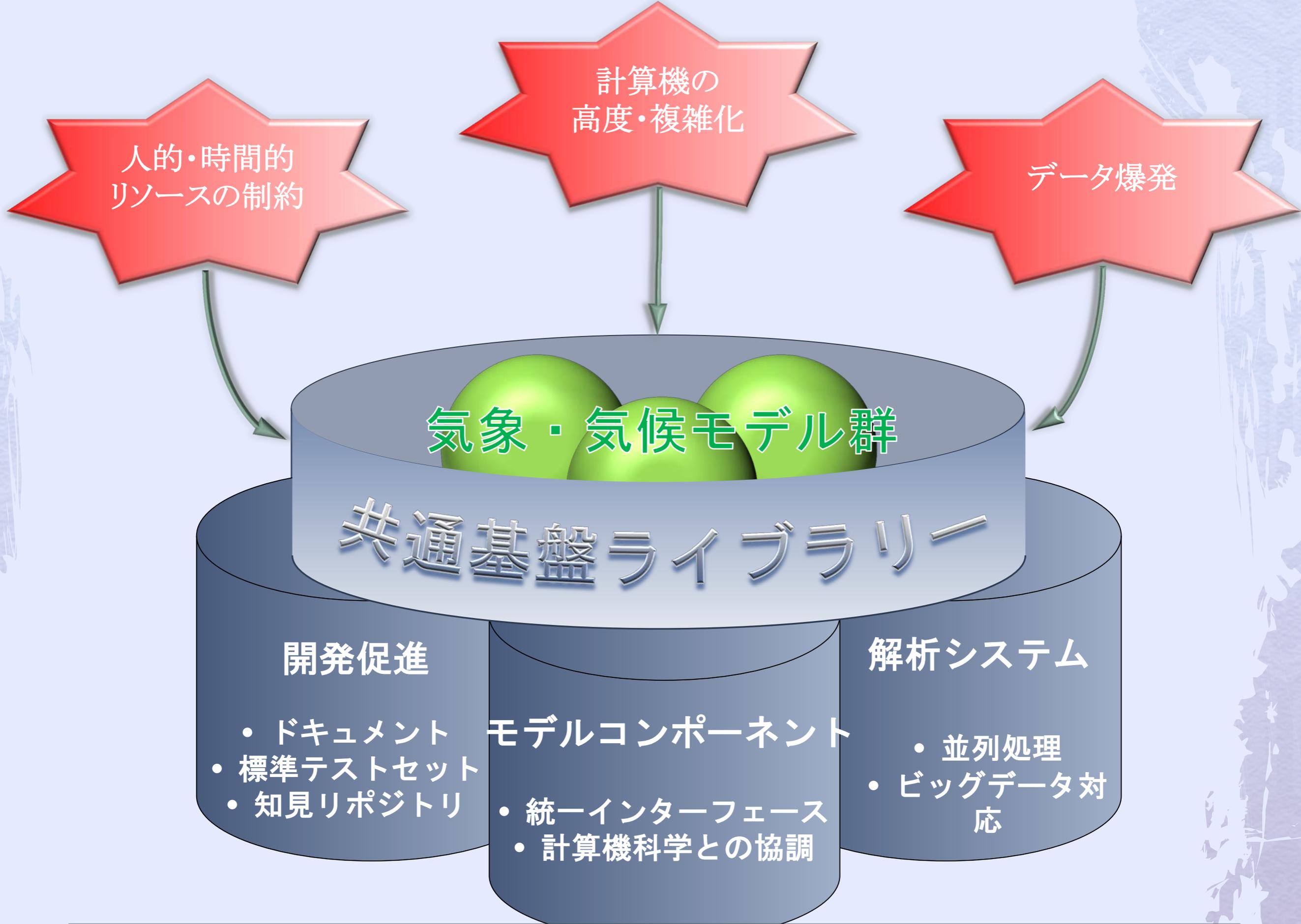


**HPC Usability
Research Team**

TL : T. Maeda



Co-design of application, library, and environment
for weather/climate research



人的・時間的
リソースの制約

計算機の
高度・複雑化

データ爆発

気象・気候モデル群

共通基盤ライブラリー

開発促進

- ドキュメント
- 標準テストセット
- 知見リポジトリ

モデルコンポーネント

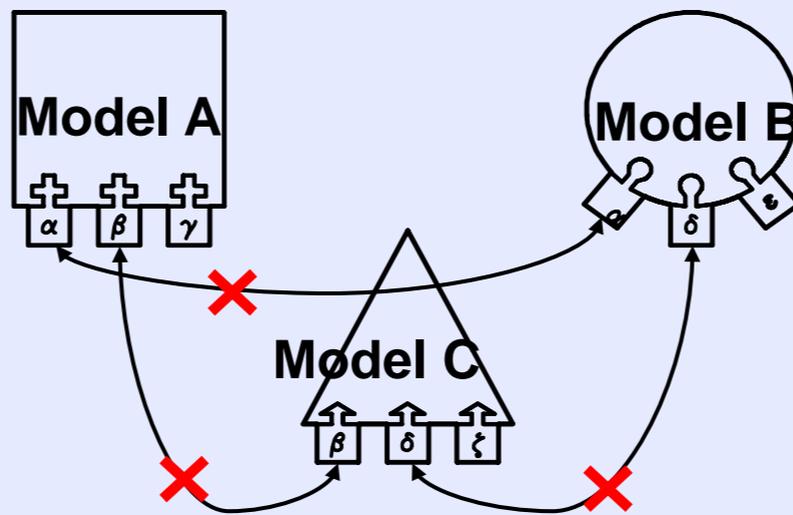
- 統一インターフェース
- 計算機科学との協調

解析システム

- 並列処理
- ビッグデータ対応

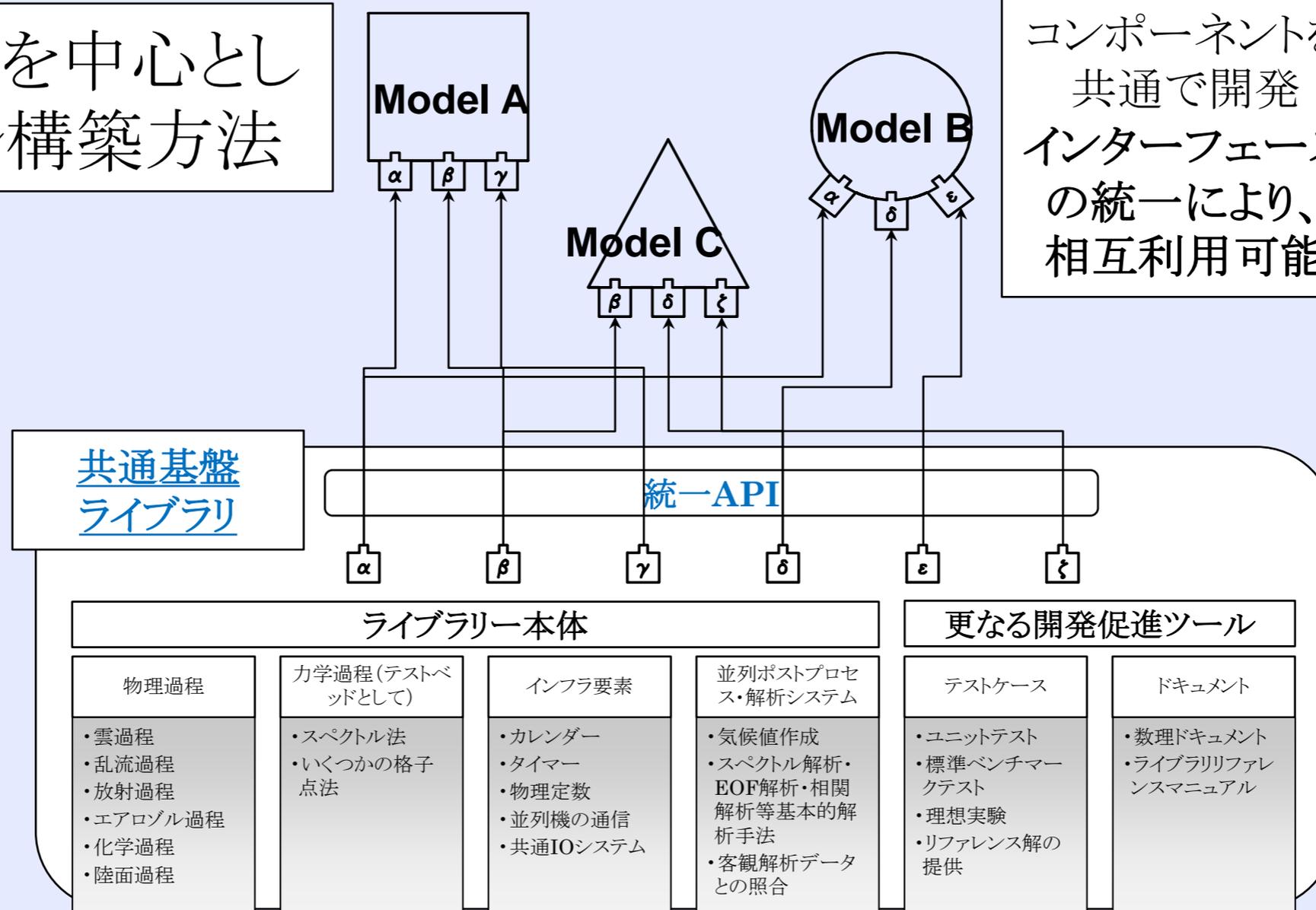
気象・気候研究成果創出へ加速！

従来型モデル構築方法



同等なコンポーネントも、各モデルでそれぞれに開発
 インターフェースが非互換なため、相互利用は困難

共通基盤を中心としたモデル構築方法



コンポーネントを共通で開発
 インターフェースの統一により、相互利用可能

- ◆ 日本国内のモデル開発グループとの協力
 - ◆ ワークショップの開催 (2011秋, 今回)
 - ◆ 勉強会の開催 (2012 春)
 - ◆ 25年度科研費へ応募

全体統括(研究代表者)

- ・富田浩文(理化学研究所・計算科学研究機構)

アドバイス要請

全体マネージャー
進捗管理

ポテンシャル研究協力・アドバイザー

- ・理化学研究所・計算科学研究機構・複合系気候学研究チームメンバー
- ・同機構・計算機科学系研究チーム(Team SCALE)

気象庁・気象研究所(統一API策定)

- ・室井ちあし(気象庁・予報部数値予報課)
- ・原旅人(気象庁・予報部数値予報課)
- ・平原洋一(気象庁・予報部数値予報課)
- ・金原貴史(気象庁・予報部数値予報課)

MIROC開発チーム(統一API策定、ライブラリー作成)

- ・大越智幸司(海洋研究開発機構・地球環境変動領域):
- ・渡部雅浩(東京大学大気海洋研究所):
- ・建部 洋晶(海洋研究開発機構・地球環境変動領域)
- ・渡辺真吾(海洋研究開発機構・地球環境変動領域):

GAIA(全大気統合モデル)チーム(統一API策定、ライブラリー作成)

- ・三好 勉信(九州大学理学研究院):
- ・陣英克(情報通信研究機構):

NICAM開発チーム(統一API策定、ライブラリー作成)

- ・八代尚(理化学研究所・計算科学研究機構):

CReSS開発チーム(統一API策定、ライブラリー作成)

- ・坪木和久(名古屋大学・地球水循環研究センター):

AFES開発チーム(統一API策定、ライブラリー作成)

- ・吉田聡(海洋研究開発機構・地球シミュレータセンター)

電脳モデル開発チーム(統一API策定、ライブラリー作成)

- ・西澤誠也(理化学研究所・計算科学研究機構):
- ・高橋 芳幸(神戸大学・惑星科学研究センター):

MSSG開発チーム(統一API策定、ライブラリー作成)

- ・馬場雄也(海洋研究開発機構・地球シミュレータセンター)

カップラーメン開発チーム(統一API策定、ライブラリー作成)

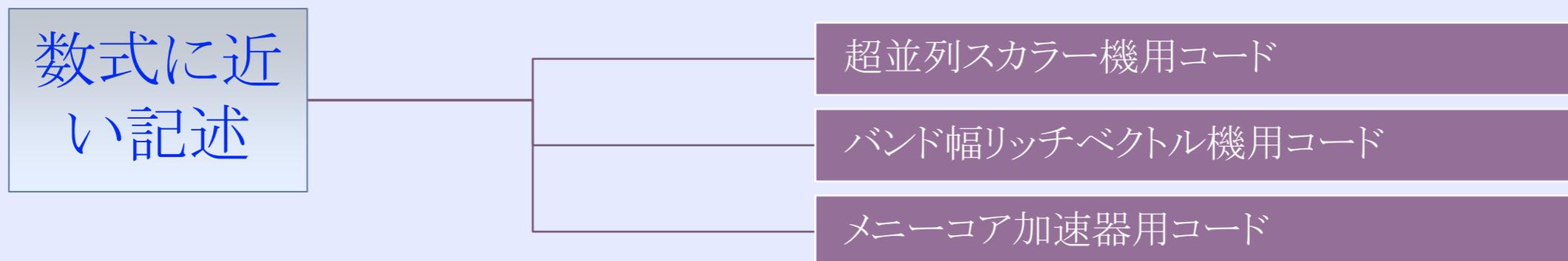
- ・荒川隆(高度情報科学技術研究機構):
- ・井上

解析システム開発チーム(並列解析ツール作成)

- ・堀之内武(北海道大学・地球環境科学研究所)
- ・西澤誠也(理化学研究所・計算科学研究機構)

課題

- ◆ グループ間でコンセンサスをとる
 - ◆ API (subroutine の入力と出力)
 - ◆ IO
- ◆ 汎用 vs. 最適化
 - ◆ ループ、データ構造 etc. の抽象化
 - ◆ CPP マクロ
 - ◆ DSL (Domain Specific Language)



SCALE-LES

- ◆ 気象・気候モデルは、超並列計算機でどのくらい
の計算ができる可能性があるのか
 - ◆ 性能限界
 - ◆ ボトルネック
 - ◆ 最適アルゴリズム

「京」コンピューターに最適化した気象・気候モデルを作成し調査 → SCALE-LES

設計思想

- ◆ 「京」コンピューターにおける性能限界調査
 - ◆ より簡単に, より速く, 精度は必要かつ最低限
 - ◆ FLOPSではなく、実験の実行時間を重視
 - ◆ 高領域実験がターゲット (mesoscale LES)
 - ◆ Full-explicit scheme : 通信は隣接のみ
- ◆ 知見を活かし、ライブラリー構築へ
- ◆ コンパイラやシステムに提言

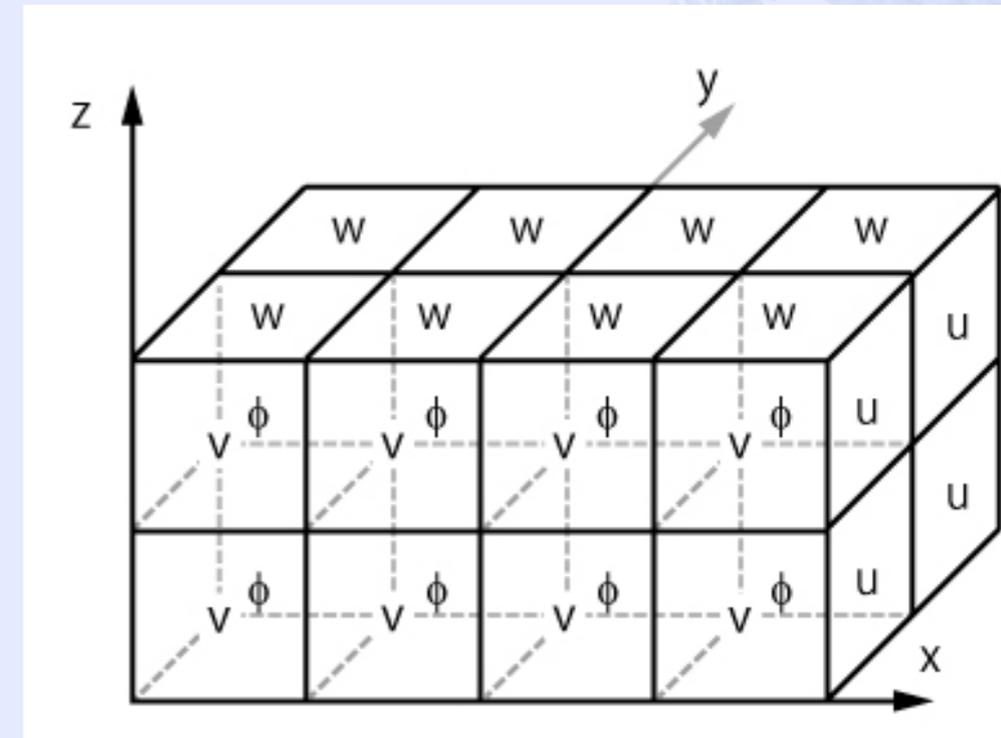
手法

◆ 力学過程:

- ◆ 完全圧縮: 3D Full Explicit
- ◆ 格子系: カーテシアン格子
- ◆ 時間積分ス: 3次RK
- ◆ 空間差分: 4次中央差分
- ◆ その他トレーサー (水蒸気など): 非負保証(FCT)

◆ 物理過程

- ◆ SGS乱流: 成層を考慮したSmagorinskyモデル
- ◆ 雲微物理: ダブルモーメントスキーム (Seiki 2011)、ビン法 (Suzukiモデル)
- ◆ 放射: MSTRNX (Sekiguchi & Nakajima 2008)



開発環境

- ◆ VCS: git + gitlite + gitLab
- ◆ 知見: Track
- ◆ reference ドキュメント: doxygen
- ◆ CI: (gitLab-CI or Jenkins)

性能

- ◆ 単体性能
 - ◆ ノード内の演算実行性能
- ◆ 並列性能
 - ◆ スケーリング (weak scale)

単体性能

八代尚

- ✳ $\Delta xyz=5m$, $\Delta t(\text{dyn})=0.008\text{sec}$, $\Delta t(\text{phy})=0.8\text{sec}$, $(k,i,j)=(1256,32,32)$
- ✳ 2.4sec分の積分
: かなりGrid数を増やして、高負荷にした設定

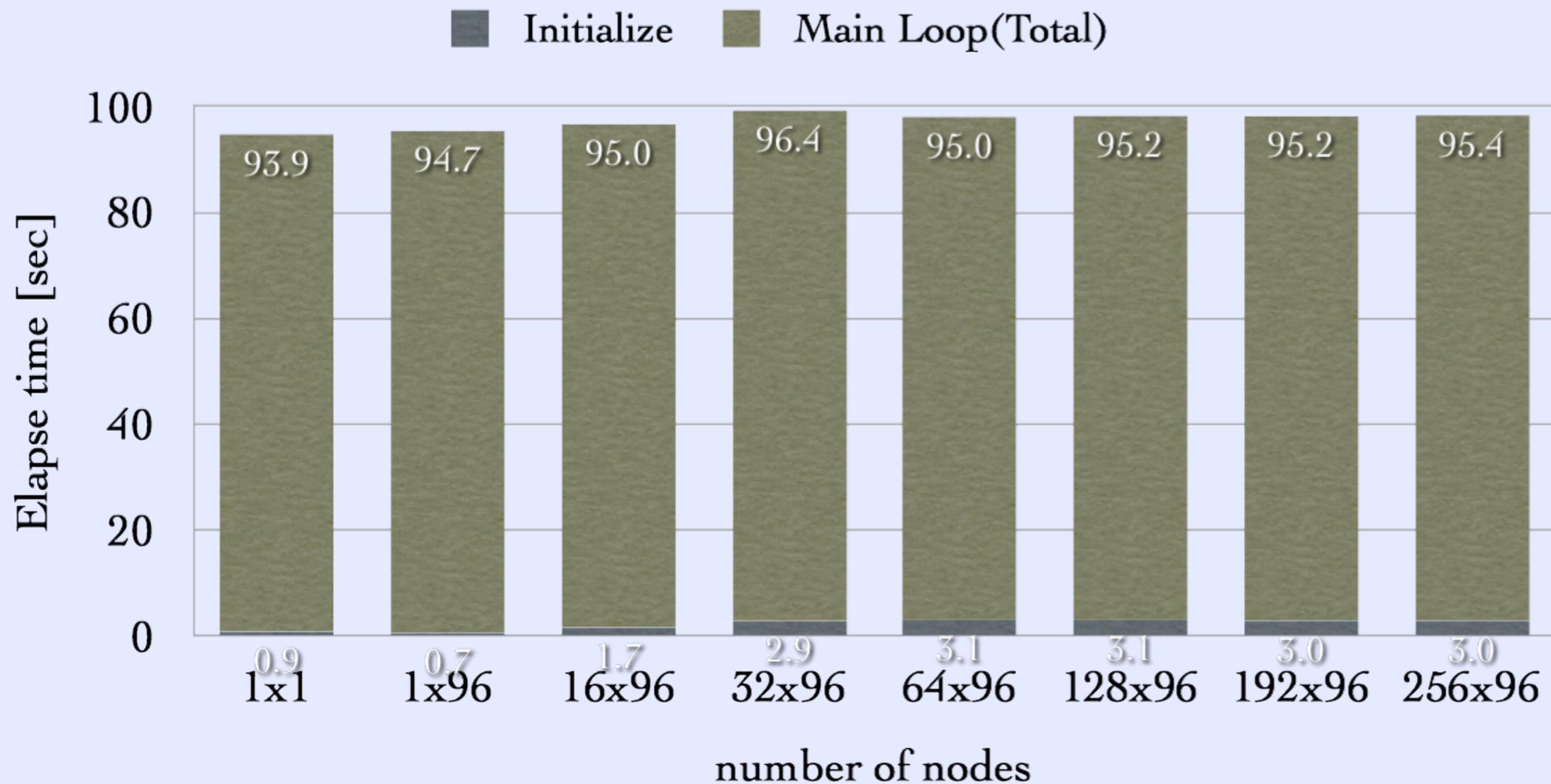
[sec]	CTL	RDMA	w/ kessler	w/ NDW6 orig	w/ NDW6 tuning
Main loop	110.6	102.5	112.7	124.9	116.3
Dynamics	109.5	101.5	108.7	108.7	108.7
TB	0.0	0.0	0.8	1.0	0.9
Micro-	0.0	0.0	1.6	13.7	5.4

力学部分だけだと、大体10%から12, 3%。

スケーリング結果

八代尚

- ✧ $\Delta xyz=5m$, $\Delta t(\text{dyn})=0.008\text{sec}$, $\Delta t(\text{phy})=0.8\text{sec}$, $(k,i,j)=(1256,32,32)$
- ✧ 2.4sec分の積分
: かなりGrid数を増やして、高負荷にした設定



MPIよりRDMAを直接つかう！

大野義之

```
do large_step = 1, 3
  ! dynamics
  do small_step = 1, 100
    computation
    call COMM_rdma_vars8( 17 , 5 )
    computation
    call COMM_rdma_vars8( 22 , 5 )
    computation
    call COMM_rdma_vars8( 1 , 5 )
    computation
    do iq = 1, 11
      computation
      call COMM_rdma_vars8( 27 , 3 )
      computation
    enddo
    computation
    call COMM_rdma_vars8( 6 , 11 )
    computation
  enddo
enddo
```

- ◆ Dynamics 300 small step分の通信時間を測定
 - ◆ グリッドのサイズを, (Z,X,Y) = (1256, 32+2*HALO, 32+2*HALO)とし, 1回の通信でHALOを隣接8方向(東西南北+斜め)に通信
 - ◆ 変数は29種類(29個の3次元配列)
 - ◆ 通信回数 4500回
 - ◆ 5配列の送受信 900回
 - ◆ 3配列の送受信 3300回
 - ◆ 11配列の送受信 300回
 - ◆ 今回の測定では, computationは空

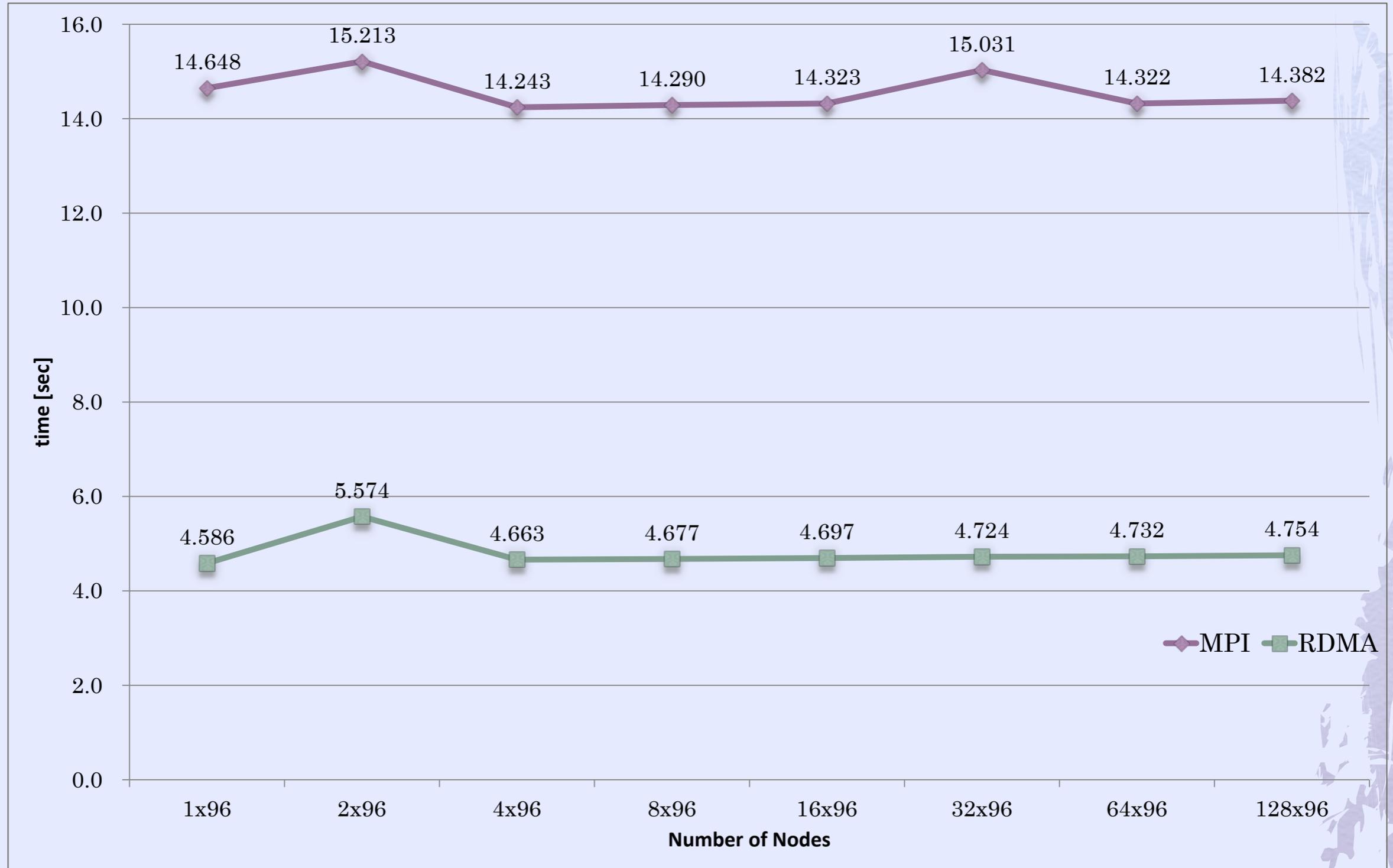
`COMM_rdma_vars8(vid , cnt)`

- 配列には integer の ID が割り振られており, ID を引数に与えて, 通信する配列を指示
- `COMM_rdma_vars8` 関数1回で, `vid` から `vid+cnt-1` までの `cnt`個の配列を転送

MPI版では, 以下に置き換え

```
do i = 1, cnt
  call COMM_vars8( var(:, :, :, vid+i-1) )
enddo
do i = 1, cnt
  call COMM_wait( var(:, :, :, vid+i-1) )
enddo
```

◆ RDMAを使うと, MPIよりも3倍高速 結果



サイエンスターゲット

◆現在の主対象

- ◆層積雲の維持・崩壊メカニズム

- ◆火星境界層

 - ◆collab w/ CPS

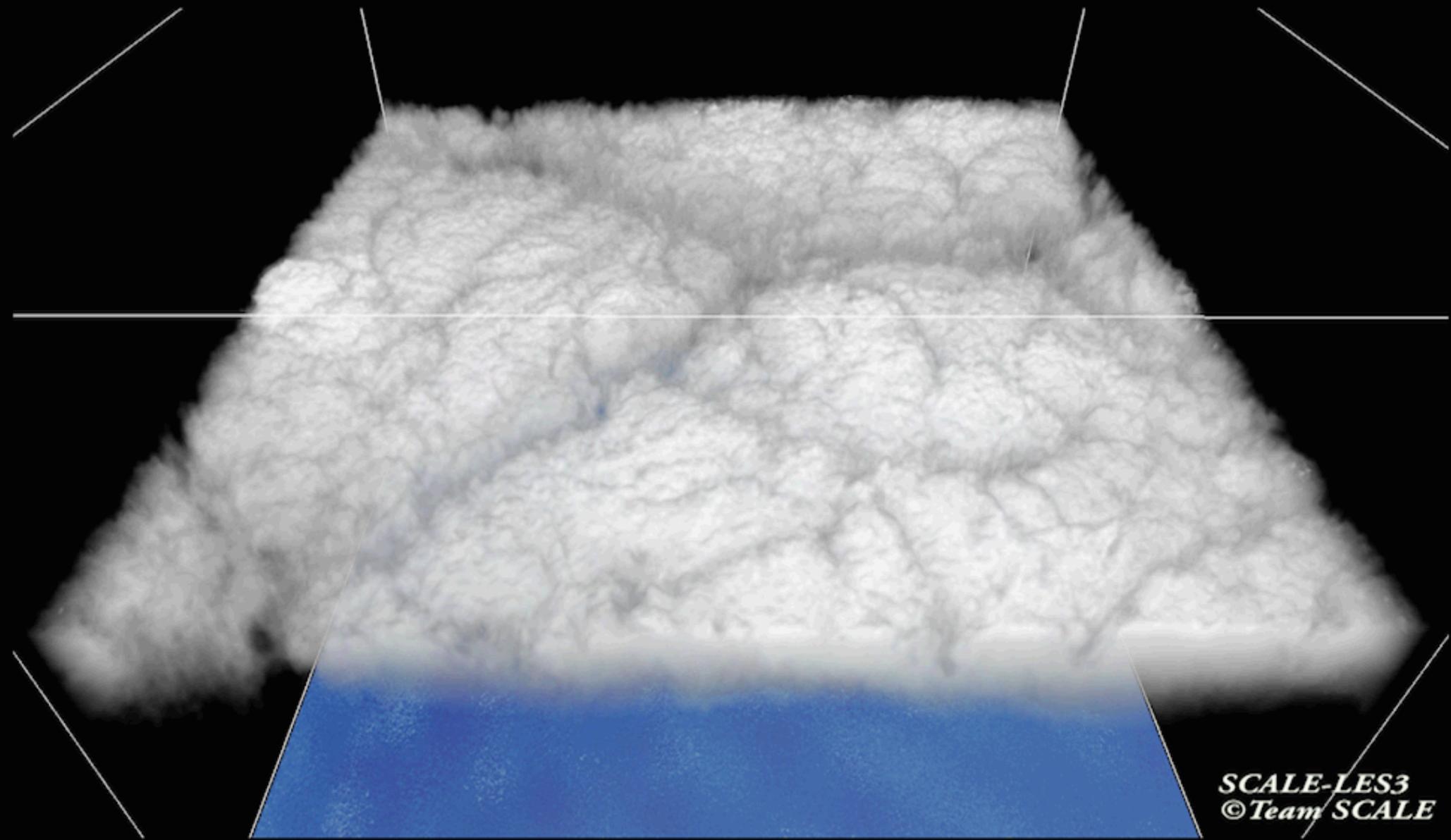
- ◆都市豪雨

 - ◆collab w/ 京都大学, 兵庫県・神戸市

層積雲

- ◆ 層積雲
 - ◆ 亜熱帯・東海岸で発現
被覆率:50 %以上
- ◆ 気候に与えるインパクト大
 - ◆ 特に日傘効果大
- ◆ 層積雲の崩壊
 - ◆ 層積雲から積雲への遷移のメカニズムは分かっていない
- ◆ 時間・空間スケール大きい
 - ◆ 現象を記述する物理過程は非常に細かい



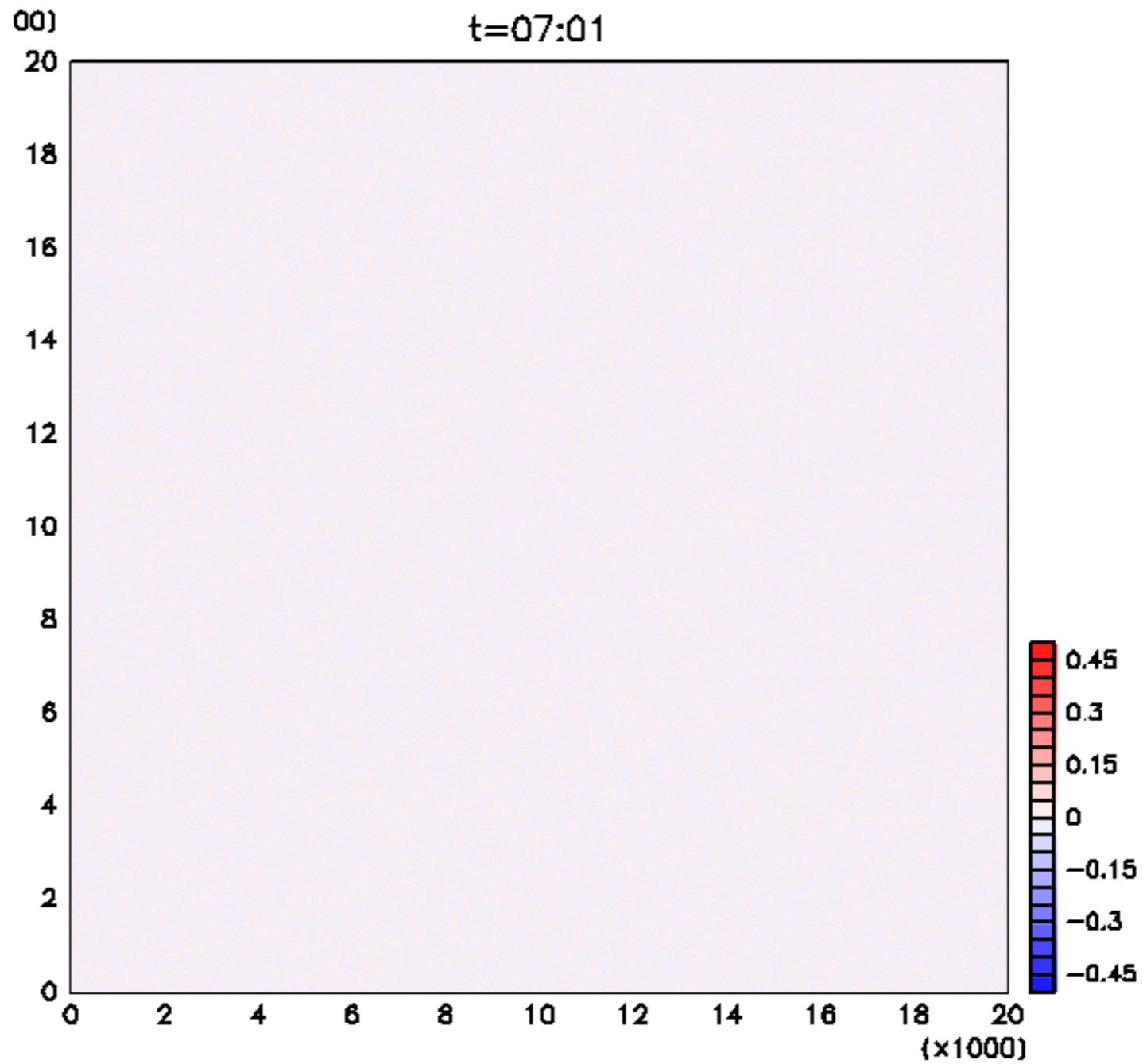


Simulated by Y. Sato (AICS)
Visualized by R. Yoshida (AICS)

火星境界層

- ◆ 火星境界層
 - ◆ 日中は 10km 程度まで発達
 - ◆ ダストデビルがみられる
- ◆ ダストの巻き上げ
 - ◆ ダストストームが、全球規模に発展する場合もある
- ◆ 広領域・高解像度実験が必要
 - ◆ セル構造の表現には広領域が必要
 - ◆ 地表面付近の運動のスケールは非常に小さい

渦度



$\tau=0.0$, $\Delta xyz=25\text{m}$, $L=20\text{km}$, $z=37.5\text{m}$

立ちはだかる未知の領域

◆ 音波! 音波! 音波!

- ◆ 世界は音波に満ち満ちあふれている
- ◆ 問題があると、偽の音波源から音波が発生
 - ◆ モデル上端・下端を節とする mode 1 の音波
 - ◆ large time step を周期とした音波
 - ◆ 力学: small time step, 物理過程: large time step
 - ◆ 波長と近い幅のストレッチング層で、2-grid noise
 - ◆ etc

今ここで、完全圧縮世界の経験を積んで、
知見を蓄積しておく

まとめ

- ◆ 地球・惑星の気象・気候モデル共通基盤ライブラリー構築 (SCALE)
 - ◆ リソースの有効活用
 - ◆ 最適化＋汎用性(可読性)のバランス
- ◆ 広領域・高解像度実験のためのLESモデルの開発 (SCALE-LES)
 - ◆ 単体性能・並列性能は、まずまず
 - ◆ 種々のテストケースで検証
 - ◆ 現在の主なターゲット
 - ◆ 層積雲, 火星境界層, 都市豪雨