

# 分散・大規模データ 可視化の動向

日本電気株式会社  
吉田 有宏

# 本講演のアウトライン

- はじめに ~ 可視化の置かれている状況と課題
- 分散・大規模データ可視化方式の概要
  - 典型的な解決策
  - 画像送信によるリモート可視化
- 世界の動向
  - DOE(米国エネルギー省)関連
  - グリッド関連
- 弊社取り組みのご紹介
  - NEC RVSLIB
  - NAREGI グリッド可視化システム
- まとめ

# はじめに

## 可視化の置かれている状況

### ■ 計算機シミュレーションにおける可視化の重要性

- 計算結果は人間による解釈が事実上不可能な膨大な数値の羅列.
- シミュレーションは計算結果が人間に解釈されてはじめてその目的が達成される. 可視化により初めて研究者は目に見える形で計算結果を観察, 分析できる.

### ■ 計算結果データの大規模化と分散化

- 計算結果データの規模はますます増大.  $\infty$  計算機の高性能化
- 計算サーバの数が増大し計算結果データは遠隔のそれらに散在.  $\infty$  グリッドの普及  
一般的な可視化ソフトでは簡単かつ効率的な可視化がますます困難に.

## 解決すべき課題

### ■ 分散配置された大規模データの円滑な可視化

- 遠隔の複数計算サーバ上に分散されて保存された巨大な計算結果データを, 複雑な操作や煩わしい待ち時間なしに, 研究者のPCより遠隔から可視化できるシステムの実現.

# 分散・大規模データ 可視化方式の概要

# 典型的な解決策 ~1/3~

## ■ 可視化専用サーバ+オンサイト可視化

- HW処理が可能な高性能可視化専用サーバを導入, XやVNC(端的に言えばリモートデスクトップ)を利用してオンサイトの可視化アプリを遠隔操作. … 力技の方法
- 長所:
  - 原則としてストレスなくスムーズに可視化が可能.
- 短所:
  - 専用サーバへの大規模データの転送が大きなボトルネック.
  - 専用サーバのHDDに大規模データが入りきらない可能性.
  - メモリ容量の制約により大規模データの可視化処理が行えない可能性.
  - 可視化操作/表示での通信遅延.
  - ユーザ数増大に追従しがたい. (排他利用, 共有利用のいずれでも)
  - 設備投資額大. 増設, 増強も気軽にはできない.
  - リモートデスクトップのプロトコルはセキュリティ上サイトで許可されにくい. (ただし, 後述の画像転送による方法との組み合わせで回避可)

# 典型的な解決策 ~2/3~

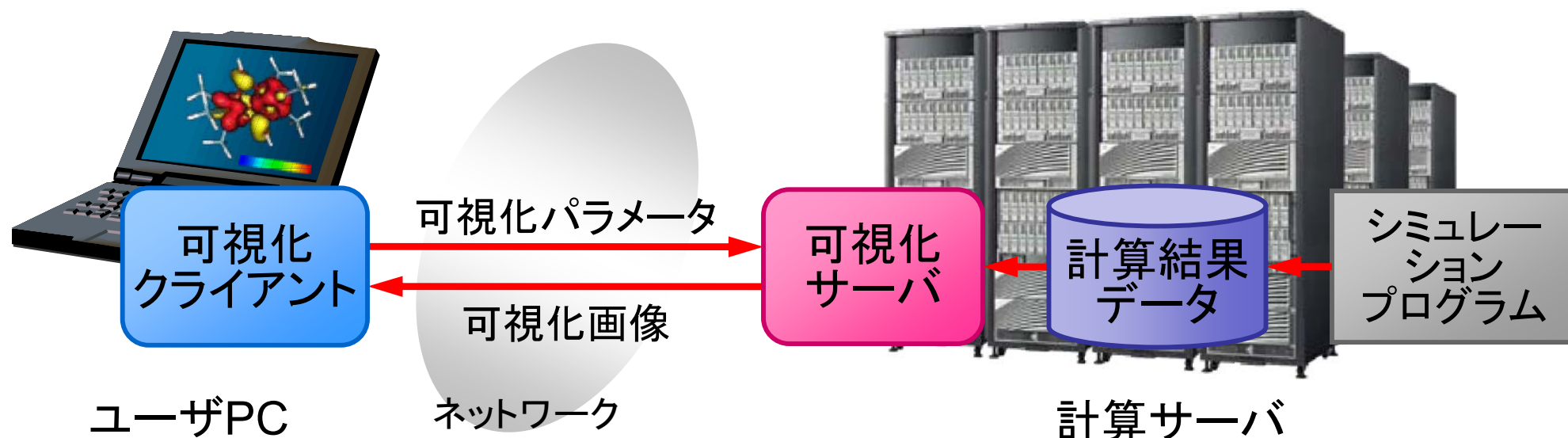
- データリダクション＋ローカル可視化
  - 高精細なデータを低解像度化してデータ量を削減, または高解像度のまま興味領域のみを切り出し, これをPCにダウンロードしてローカルで可視化.
  - 長所:
    - ストレスなくスムーズに可視化が可能.
    - 可視化操作/表示に通信遅延なし.
    - 既存の可視化アプリを利用できる.
  - 短所:
    - PCの性能に応じて, 解像度や領域を妥協しなければならない.
    - データ量を削減したとはいえ, ダウンロードがボトルネック.  
(PCで扱える最大に近いサイズへのデータリダクションが通常のため)

# 典型的な解決策 ~3/3~

- 画像送信によるリモート可視化 … 今回, この方式を詳しく
  - 可視化対象データに直接アクセス可能な計算サーバにて可視化処理を行い, 画像をPCのクライアントに送信.
  - 長所:
    - 可視化対象データ規模に依らずハイエンドPCは不要.  
(高速CPU, 高性能GPU, 大容量メモリ, 大容量HDDを要さない)
    - 可視化対象データ規模に依らずネットワーク負荷が軽い.  
(通信するのは制御パラメータと画像だけのため)
    - データロードのボトルネックが最小化.
    - 低解像度データにリダクション不要.  
(ネットワーク性能, 処理性能, 容量の制限のために妥協しなくてよい)
  - 他の利点は後述
  - 短所:
    - 可視化操作/表示での通信遅延.
- ⌘ 以上の3方式は排他的なものではなく, 合わせ技も可能.

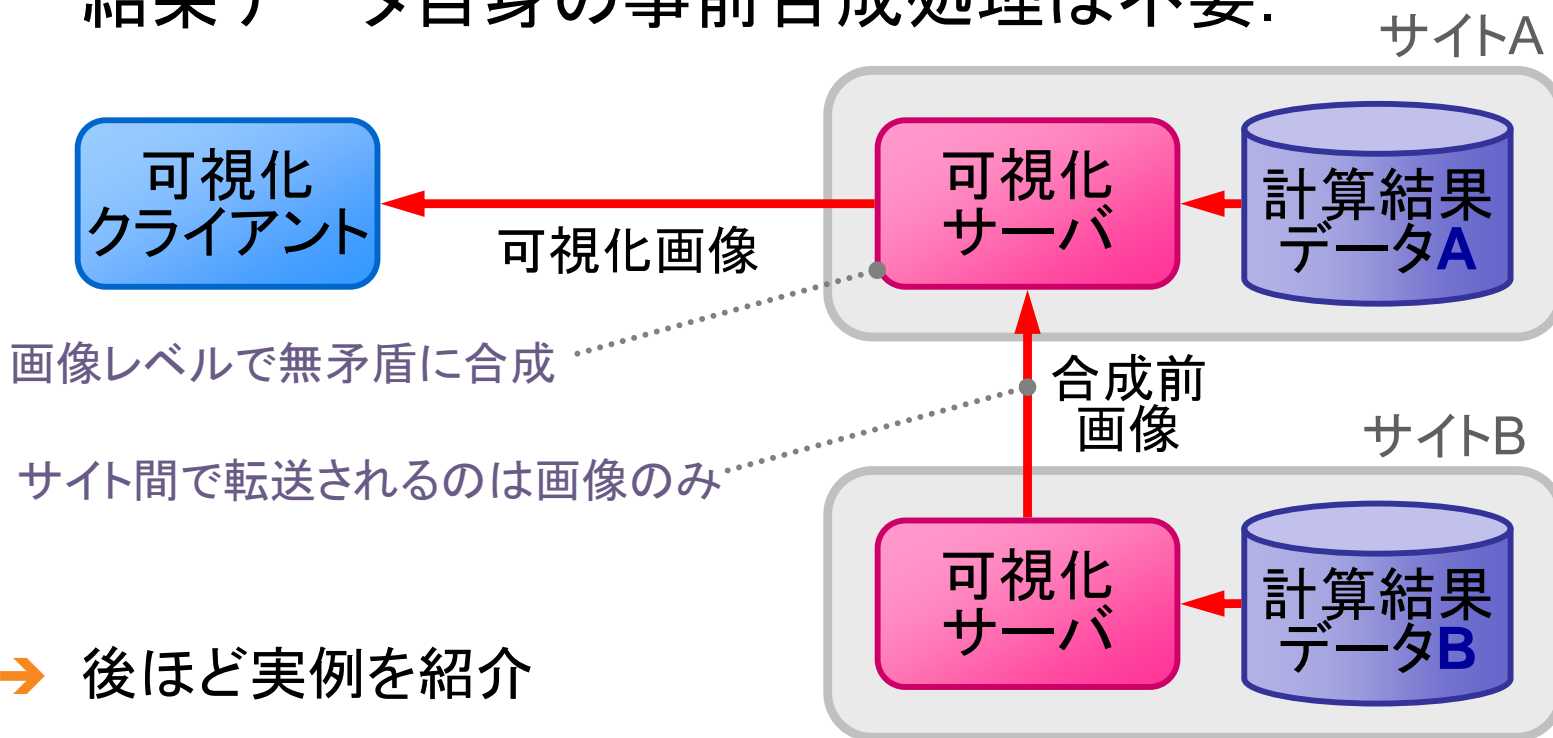
# 画像送信によるリモート可視化 ~1/2~

- 従来のローカル可視化との違い
  - 計算結果データをユーザのPC上ではなく、リモートの計算サーバ上で可視化.
  - ユーザはPC上で起動したクライアントGUIによって、計算サーバ上の可視化を制御でき、そこで生成された画像を見ることができる.



# 画像送信によるリモート可視化 ~2/2~

- 分散データに対する利点 … 大規模データに対しては前述
  - マルチサイト連成/並列計算の場合でも大規模な計算結果データのサイト間転送は不要.
  - マルチサイト連成/並列計算の場合でも大規模な計算結果データ自身の事前合成処理は不要.



→ 後ほど実例を紹介

# 世界の動向

# DOE (米国エネルギー省) 関連 ~1/2~

## ■ ASC (Advanced Simulation & Computing) ... IBASCI

### □ ParaView

- オープンソースの可視化ソフト.
- ローカル可視化とリモート可視化が可能. MPIによる並列可視化も可能. 特に非構造格子に強い.
- Kitware社, Los Alamos国立研究所により2000年に開発開始, 2002年に初の一般リリース. Sandia国立研究所, CSimSoft社が開発に加わり, UI強化の改良版を2007年に一般リリース. メンテが比較的よい.

<http://www.paraview.org/>

### □ VisIt

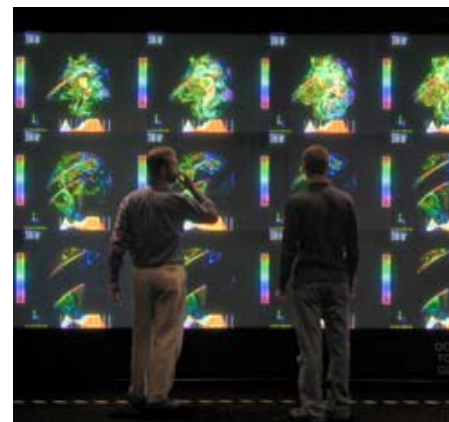
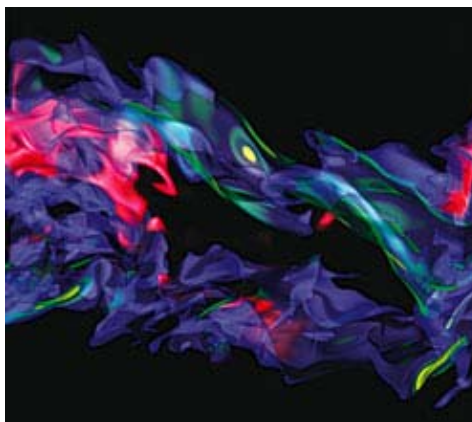
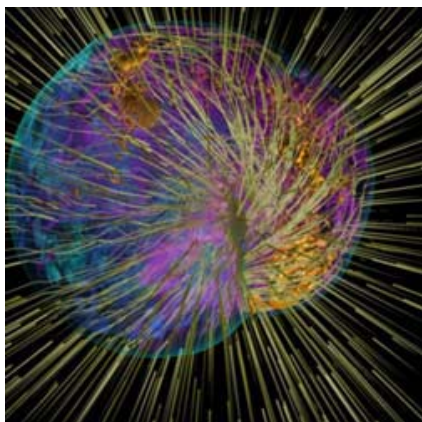
- オープンソースの可視化ソフト.
- UIや使用手順など細部は異なるが, 主たる特徴はParaViewと同様.
- Lawrence Livermore国立研究所により2000年に開発開始, 2002年に一般リリース, その後も定期的に最新版をリリース.

<https://wci.llnl.gov/codes/visit/>

### □ 他にいくつか様々な可視化プロジェクトあり

# DOE (米国エネルギー省) 関連 ~2/2~

- SciDAC (Scientific Discovery through Advanced Computing)
  - SciDAC-2
    - SciDACプロジェクト(2001~2005)の後続として発足(2006~).  
可視化に限らず広く結果分析手法の研究開発がミッション.  
<http://www.scidac.gov/>
  - Ultra-Scale Visualization Institute
    - SciDAC傘下に設立された大規模可視化に関する専門組織(2006~).  
方法論の研究, 教育促進, ライブラリ開発への投資などがミッション.  
(DirectorのKwan-Liu Maは精力的に学会開催. SC併催Workshop等.)  
<http://vis.cs.ucdavis.edu/Ultravis/>



# グリッド関連 ~1/2~

## ■ UK e-Science

### □ RealityGrid

- 英国EPSRC(工学物理研究会議)からファンドされたグリッドプロジェクトの一つ(2001~). London大学, Edinburgh大学, Manchester大学等々が参加.
- シミュレーションの動的なコントロールとリアルタイム可視化(計算途中結果のモニタリング)を可能とするステアリングライブラリを開発. その他いくつかのミドルウェアも開発している. – <http://www.realitygrid.org/>

### □ e-Viz

- 英国EPSRCからファンドされた可視化グリッドプロジェクトの一つ(2003~2007). Manchester大学, Leeds大学等々が参加.
- ユーザ指定の要件に基づき, グリッド環境上で利用可能な各種可視化サーバや可視化アプリの中から適したものを, ブローカリングによりシステムが選択するのが主な特徴. 実験的試みで終了. – <http://www.eviz.org/>

### □ 他にも関連可視化プロジェクトあり

- RAVE (Resource Aware Visualization Environment) @Cardiff大学等.

# グリッド関連 ~2/2~

## ■ AUSTRIAN GRID

### □ gsh/glogin

- GSI通信を用いてグリッド環境上でインタラクティブな遠隔操作を可能とするミドルウェア。シミュレーションのステアリングやインタラクティブリモート可視化、ストリーミング等が可能。命名はrsh, rloginのアナロジー。
- Johannes Kepler大学にて研究開発。詳細情報は学会発表等でしか行われていない。

## ■ OGF

### □ Grid Visualization and Steering

- グリッド可視化やステアリングについて、APPS-RG (Applications Developers and Users RG) ないしSAGA-RG (Simple API for Grid Apps RG) の一環で継続的にWorkshopを開催。
- OGF22にてRISGE-RG (Remote Instrumentation Services in Grid Environment RG) との協業や、CG (Community Group) としての立ち上げについて議論したもよう。… 参加していないので結論は不明

# 弊社取り組みのご紹介

# RVSLIB *Real-time Visual Simulation Library*

## ■ 沿革

- 1991年 研究着手
- 1992年 プロジェクト本格始動
- 1997年 R1.0リリース リアルタイム/リモート可視化の先駆け
- 2007年 R5.1リリース

## ■ 特徴

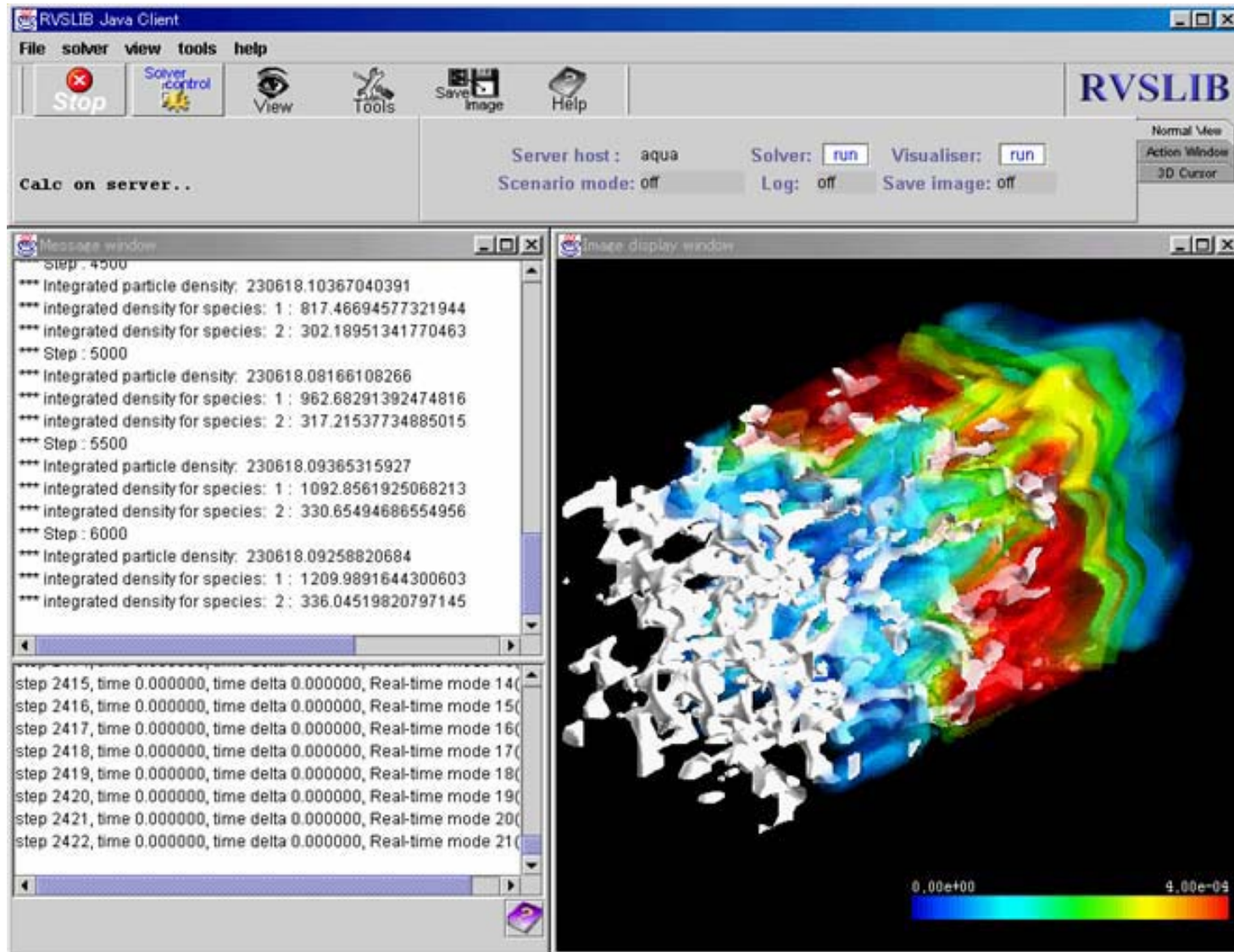
- リアルタイム可視化(シミュレーションの過程をモニタ可能)
- ステアリング(シミュレーションのパラメータを動的に操作可能)
- サーバはPCクラスタ, SXシリーズに対応
- 分散メモリ並列(MPI), ベクトル化による高速な描画
- 流体, 気象, 衝突解析等, 様々な分野に適用可能

製品HP <http://www.nec.co.jp/APSOFT/SX/rvslib>



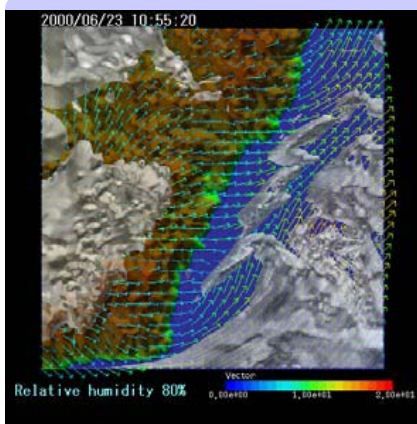
~ NEC RVSLIB ~

# ユーザインタフェース



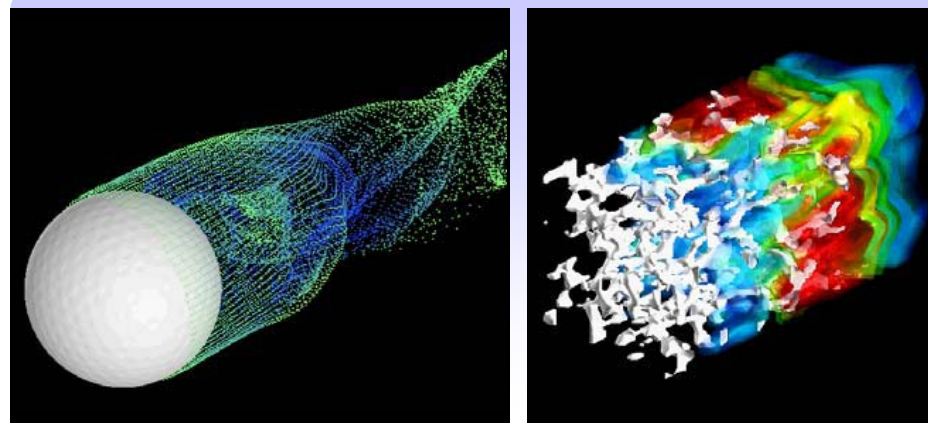
# 可視化適用事例

## ■ 気象



オーストラリア気象庁様

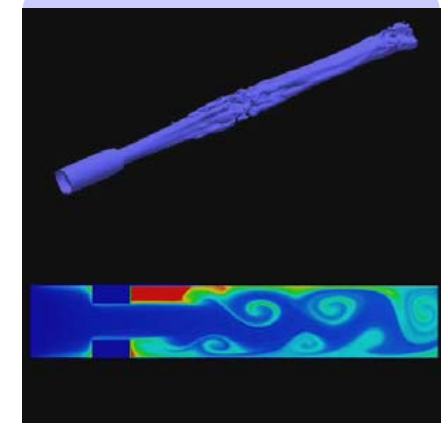
## ■ 流体



東北大学様

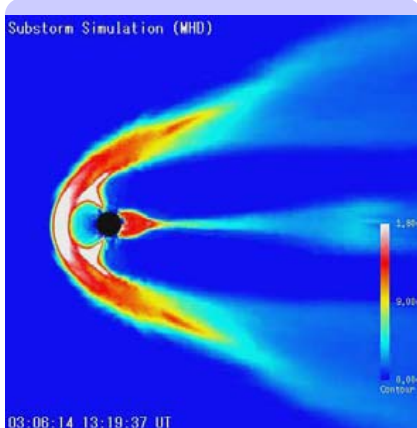
NEC

## ■ 医療



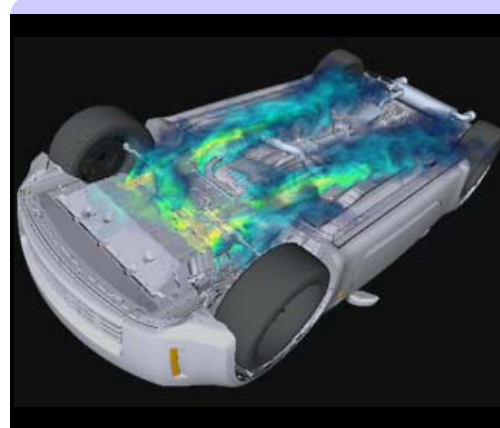
シェフィールド大学様

## ■ 宇宙



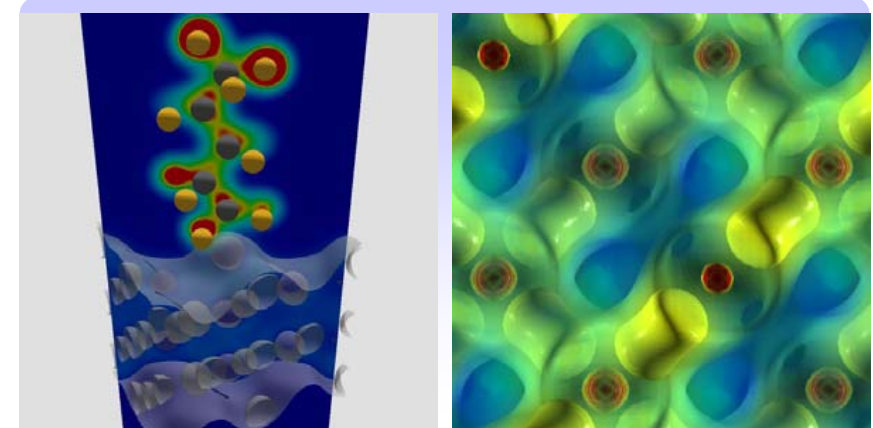
情報通信研究機構様

## ■ 製造



日産自動車様

## ■ ナノサイエンス



同志社大学様, けいはんな様

NEC

# グリッド可視化システム: GVS

## ■ 沿革

- 文科省NAREGIプロジェクト\*(2003～)にてNECが開発
- 2006年 β 版公開
- 2008年 V1.0公開予定

### \*NAREGIプロジェクト

正式名称「超高速コンピュータ網形成プロジェクト」。  
科学技術計算向けグリッド環境の実現、普及を目的。

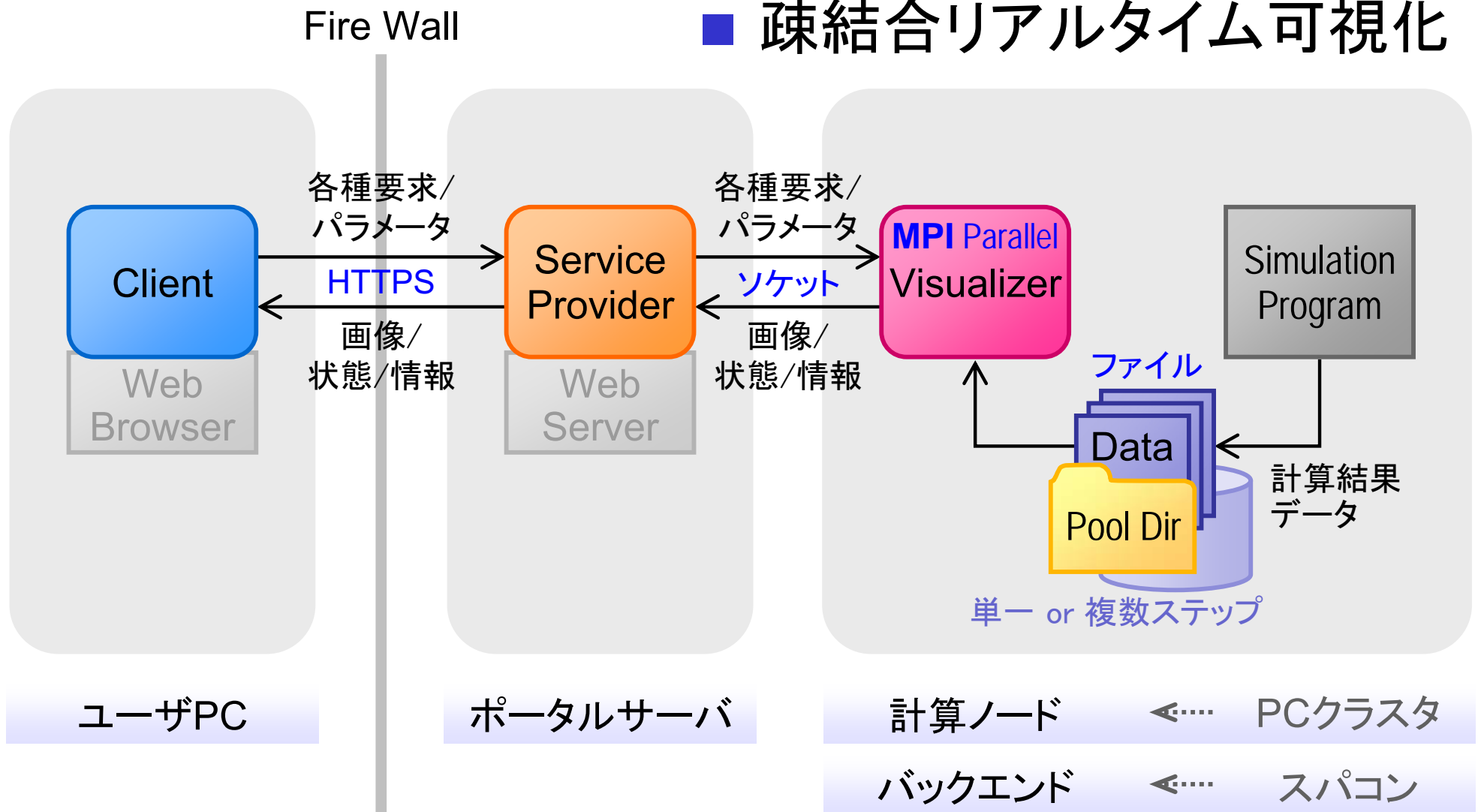
## ■ 特徴

- オープンソース
- 分散/連成シミュレーションにも対応
- ライブラリ組込み不要のリアルタイム可視化も可能
- HTTPストリーミングによるインタラクティブ性向上
- ナノ分野向けを手始めに機能を開発

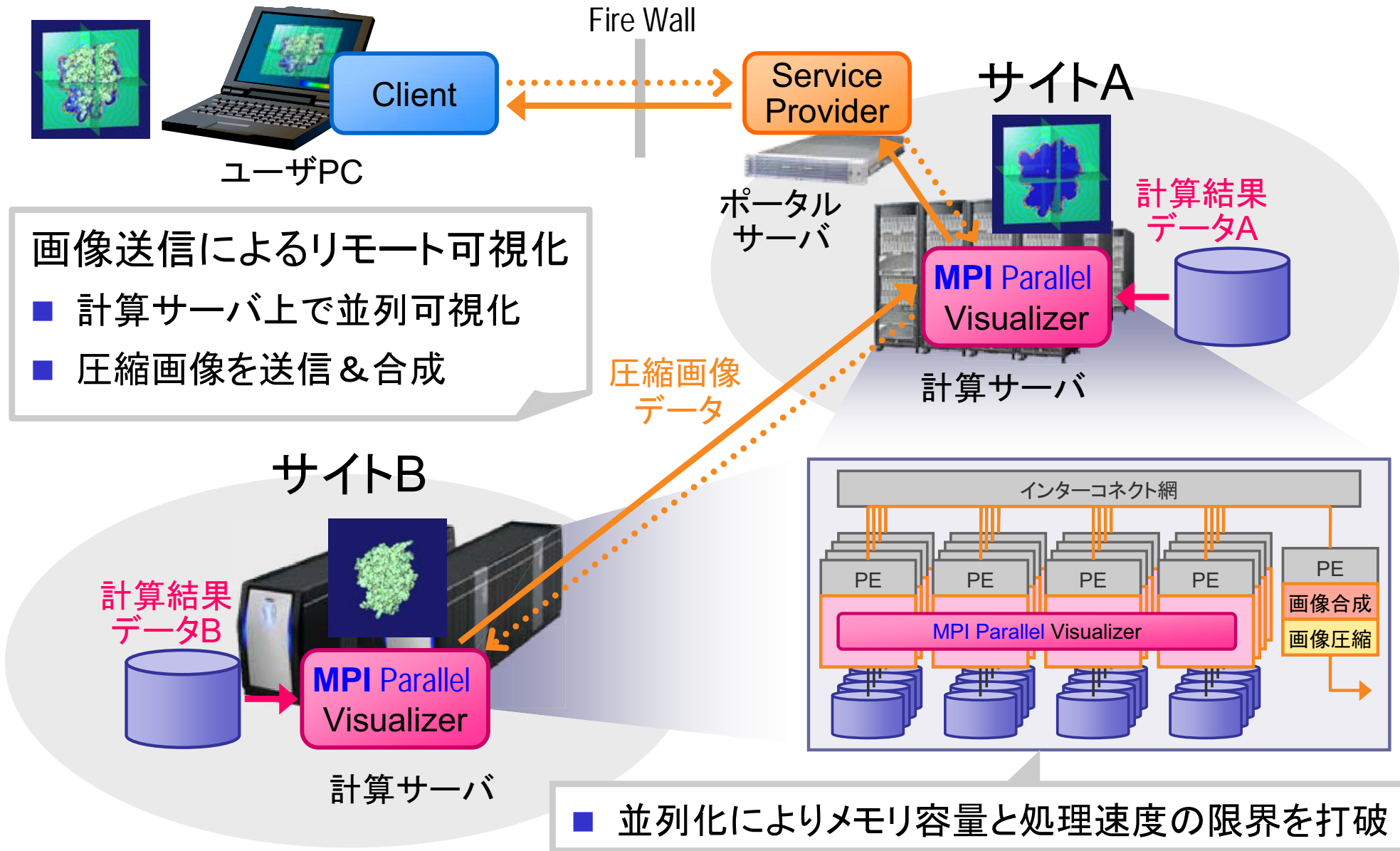
プロジェクトHP <http://www.naregi.org/>

# 基本構成(インタラクティブ可視化)

- ポスト可視化
- 疎結合リアルタイム可視化



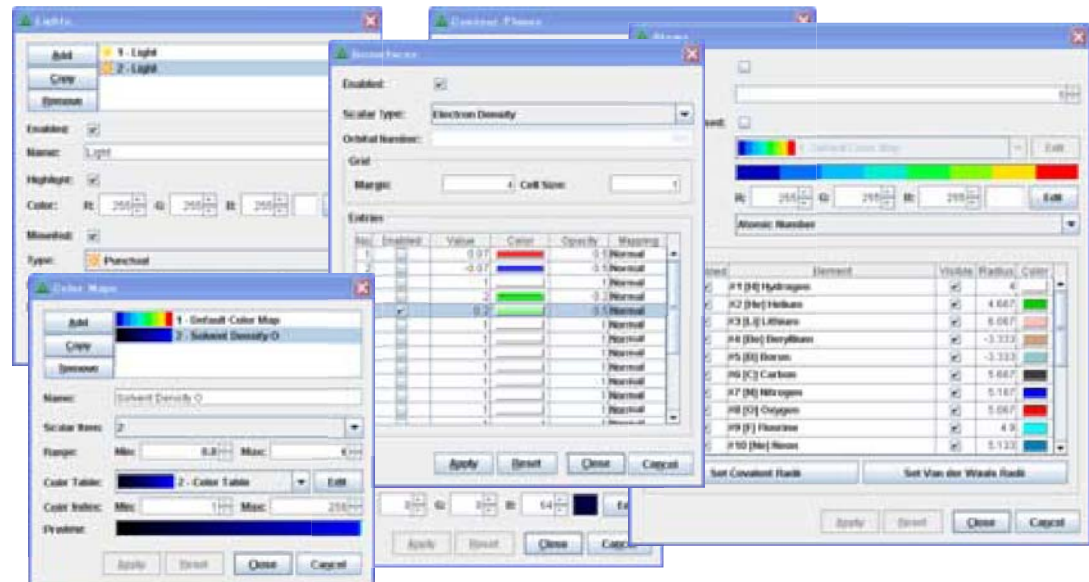
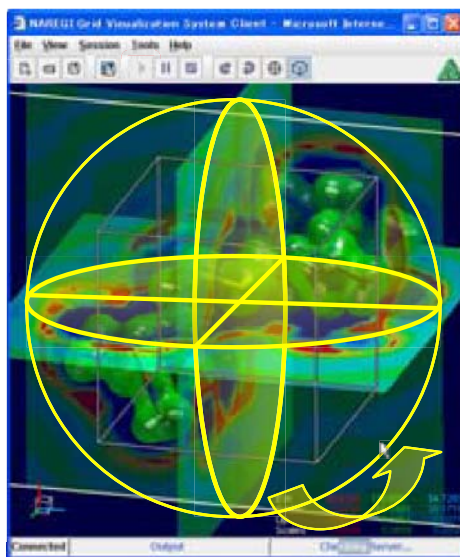
# 大規模分散リモート可視化の構成例



# ユーザインタフェース

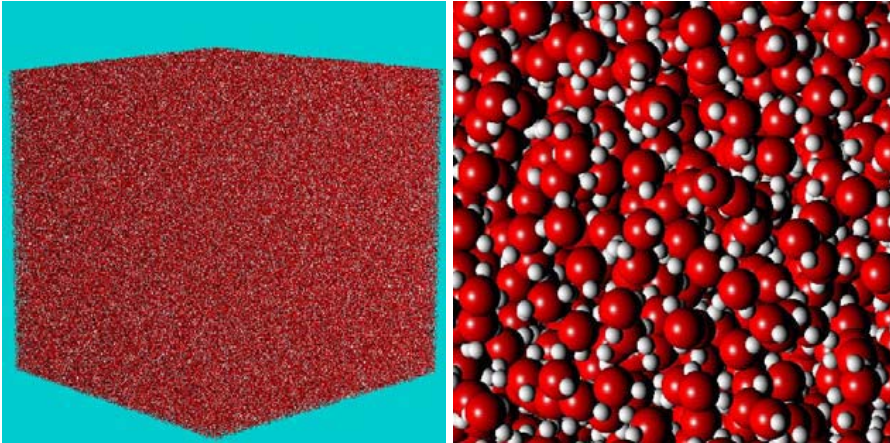
- 可視化セッション開始
  - プロファイルを設定/選択
- 可視化操作
  - マウス操作による回転, 拡大縮小
  - 可視化パラメータの設定
  - 可視化パラメータの保存/復帰

注: 画面は旧β版のもの



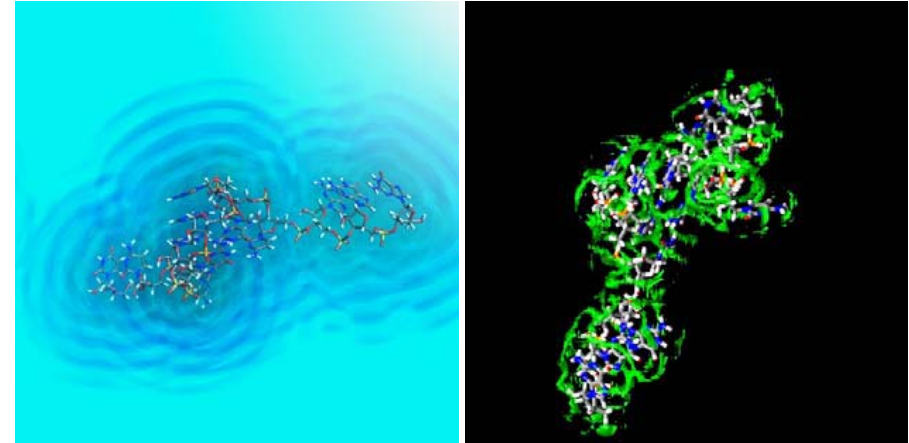
# 大規模/連成可視化事例 ～1/2～

## ■ MD(1000万原子)\*



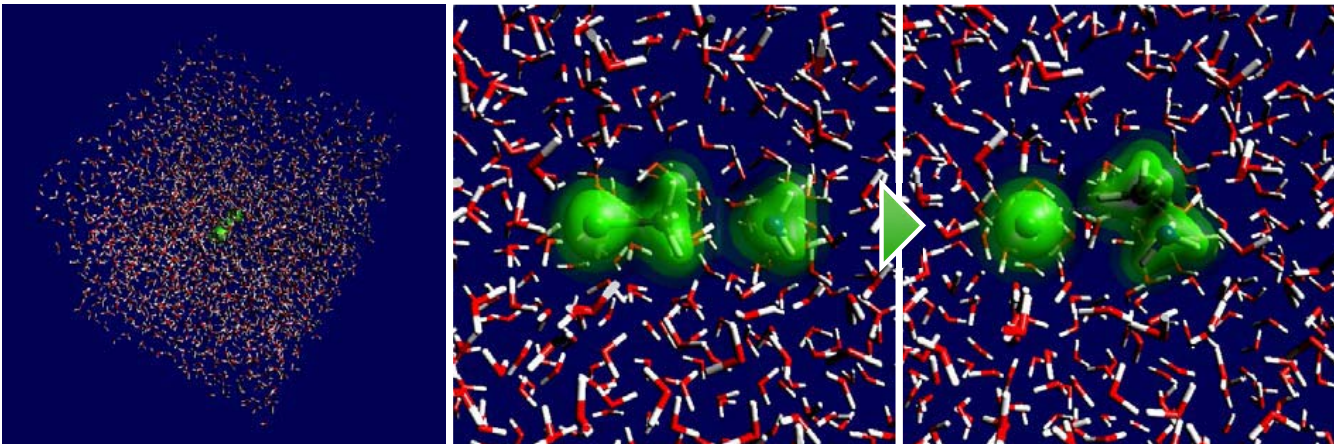
Dynamics of water molecules

## ■ レプリカ交換 MC-3D RISM\*

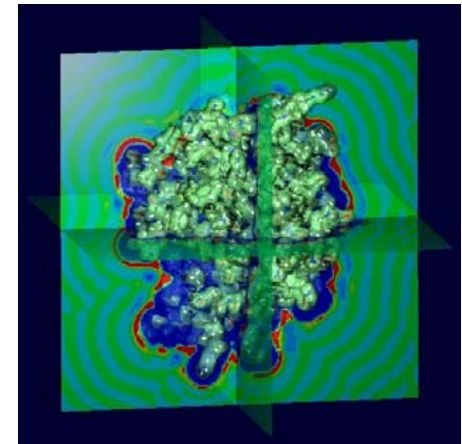


Molecular structure optimization of RNA in water

## ■ MD-DFT\*

 $S_N2$  reaction in water

## ■ RISM-FMO\*\*



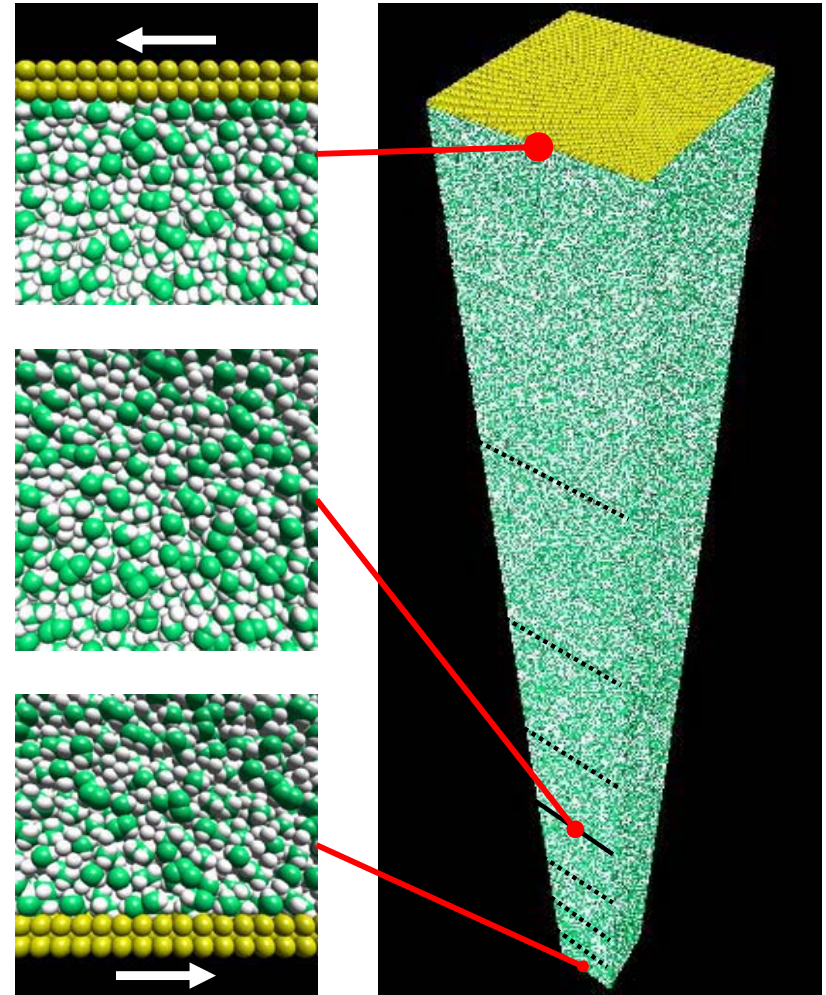
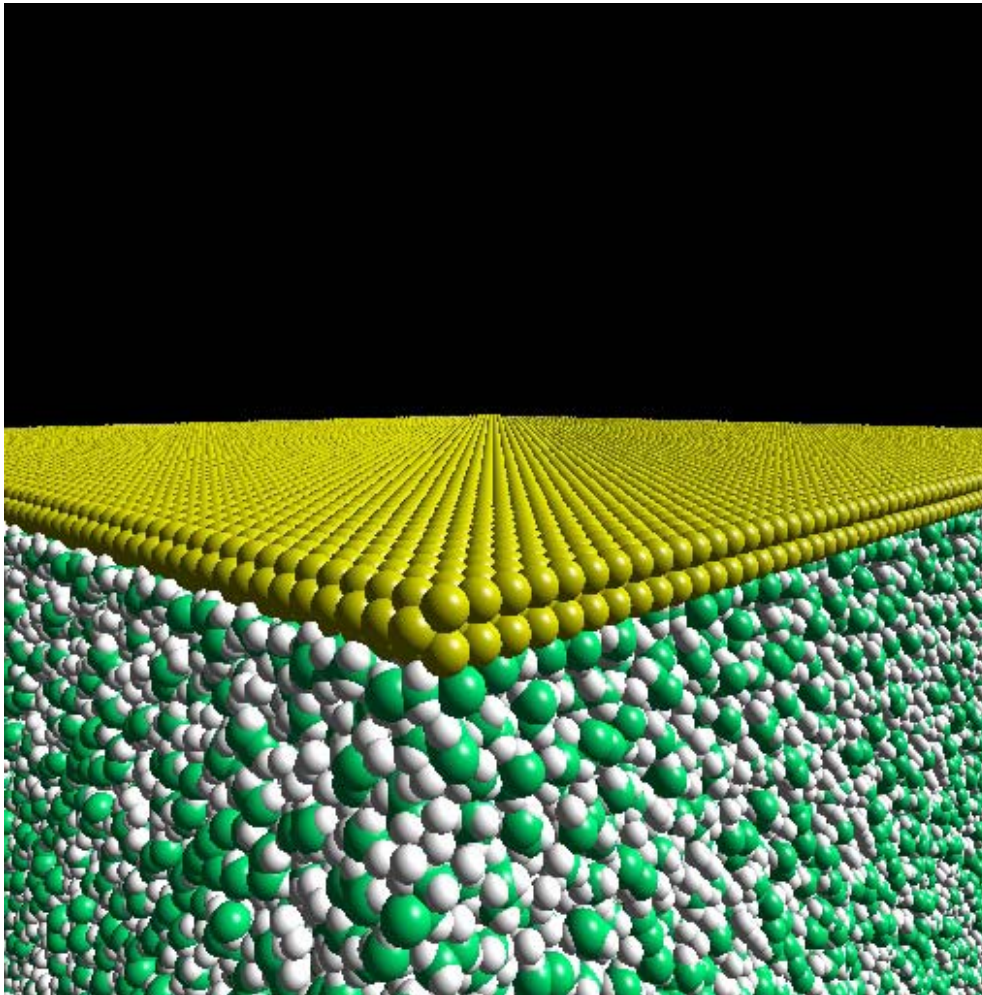
Electron and solvent structure of protein in water

データ提供: \*分子科学研究所様, \*\*九州大学様

# 大規模/連成可視化事例 ～2/2～

## ■ MD (2000万原子)

※ シミュレーション自身は50万原子  
水平方向に仮想セルを追加して可視化



Dynamics of oil molecules between metal solids

データ提供: 豊田中央研究所様

# まとめ

- 可視化を取り巻く状況の変化
  - 計算機の高性能化(特に急速なMPP化) →データ大規模化
  - 計算サーバ保有数増加, グリッドの普及 →データ分散化
- 分散・大規模データ可視化方式
  - 画像送信によるリモート可視化が近年やっと当たり前化
- 世界の動向
  - 欧米のHPCプロジェクトでは可視化を始めから重視
    - 計算後の重要な最終タスクとしての的確に認識(国内と対照的で賢明)
    - HW導入, SW開発の両面で十分な投資が行われ, 幅広く野心的に活動
- 弊社の取り組み
  - RVSLIBとして10年前にどこよりも先駆けて製品化
  - NAREGI等のプロジェクトにて可視化の分野で貢献

Empowered by Innovation

**NEC**