

2005年前期 大学院講義科目

広域分散アプリケーション特論

月曜日 3時限

場所: 情報基盤センター3F 多目的講義室

担当 高見 利也 (青柳先生の代理)

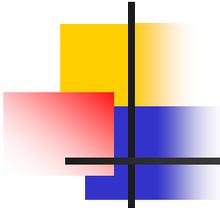
takami@cc.kyushu-u.ac.jp

7月11日(月)

講義の内容, 成績評価方針

サイエンスGrid NAREGI

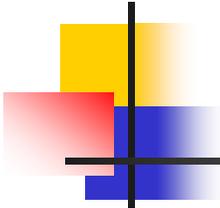
- ・ プロジェクトの概要の復習
- ・ 中間成果報告会での発表内容



講義の内容

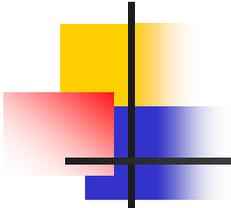
- **グリッドの概要**
 - Gridコンピューティングとは
 - サイエンス分野での利用
 - ビジネス分野での利用
- **計算科学の概論**
 - 主要なシミュレーション手法
- **サイエンスGrid NAREGI**
 - Globus, Unicoreの現状
 - NAREGIミドルウェア概要
 - 連成計算とその類型化
- **UNICORE&Globusのデモ**
- **NAREGI中間成果報告会**

講義資料はWebで公開
server-500.cc.kyushu-u.ac.jp



自己紹介

- 病院地区のコラボステーションIIでNAREGIプロジェクトの九大グループの一員として、たんぱく質の溶液中での最適構造探索システムをグリッド上で構築するための研究を行っている。
- 昨年までは分子科学研究所の計算機センターで、スパコンの導入や運用の仕事をしながら研究を行っていた。
- もともとの専門は物理学で、量子カオスや非線形ダイナミクス。
- 計算機の経歴:
MZ-700, X1 turbo, PC98note, SONY NEWS, 自作PC, ...
- 利用OSの経歴:
SharpBASIC, CP/M, MS-DOS, 386bsd(98), FreeBSD, ...
- プログラミング言語の経歴:
BASIC, Z80, C/C++, Perl, ...



今日の内容

サイエンスGrid NAREGI の復習

- グリッド研究開発(国立情報学研究所)
NAREGIミドルウェア
- ナノサイエンス実証研究(分子科学研究所)
応用研究の内容

NAREGI中間成果報告会

[7月11日(月) 9:30～15:00 国立情報学研究所]

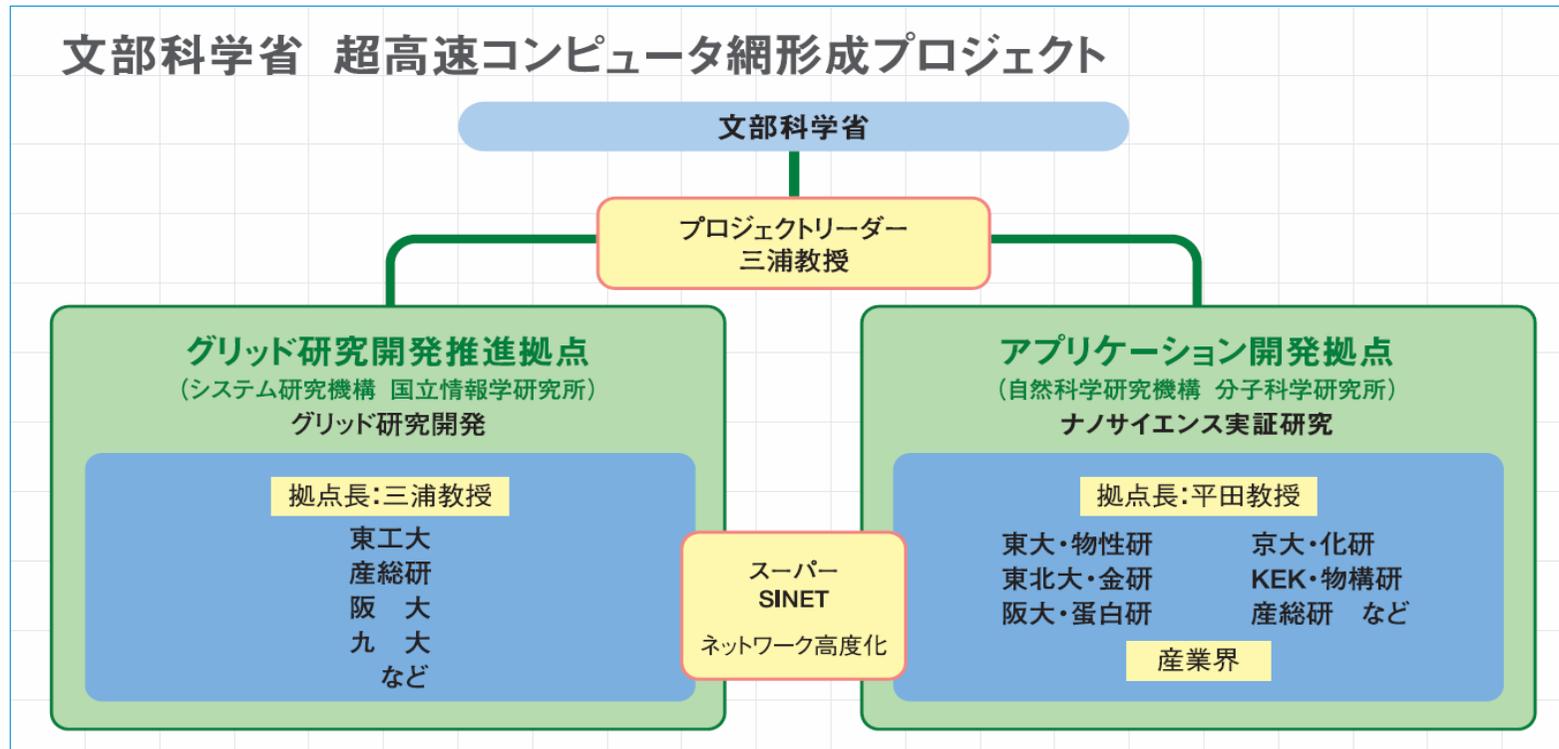
- 全体の発表内容
- グリッド研究開発グループの成果
- ナノアプリケーション実証実験



超高速コンピュータ網形成プロジェクト NAREGIの復習

- プロジェクトの全体像
- 参加組織の構成図
- 各拠点の研究内容

NAREGIプロジェクトの全体像

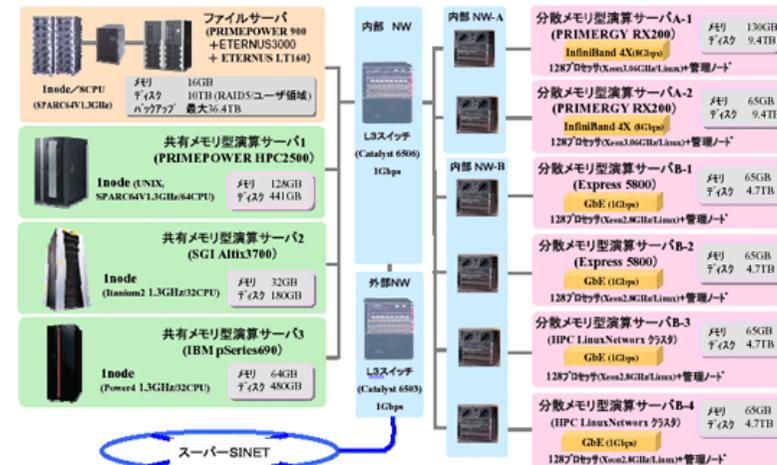


この広域分散アプリケーションの講義では、主に、左半分のグリッド研究開発を中心に解説してきた。
今日の講義では、残りの半分であるアプリケーション開発についても少しだけ覗いてみることにする。

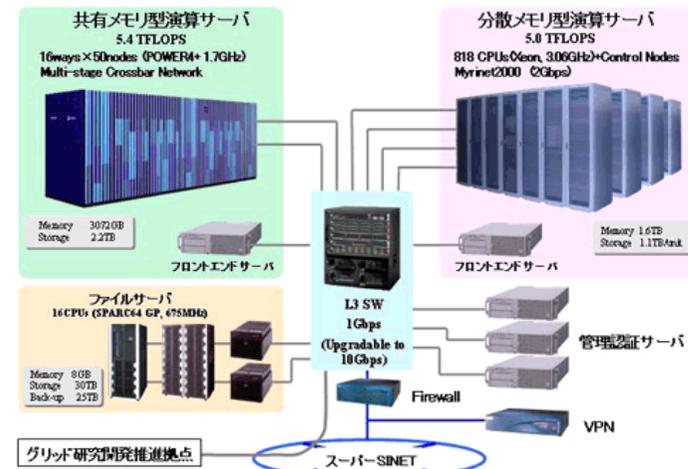
NAREGIの計算機リソース

- グリッド研究開発拠点(国立情報学研究所)とアプリケーション研究開発拠点(分子科学研究所)で動いているマシンは、それぞれ5Tflopsと10Tflopsのピーク性能を持つ。
- どちらの拠点でも、異なるアーキテクチャのマシンをテストできるようにPCクラスと大規模SMPマシンが導入されている。

グリッド研究開発推進拠点 (5Tflops, 700GB)



アプリケーション研究開発拠点(10Tflops, 5TB)



National Research Grid Initiative



グリッド研究開発の概要

基本層 (WP1 東工大 松岡教授)

- スーパー・スケジューラ
- 分散情報サービス
- グリッドVM

上位層

WP3 国情研 宇佐見教授

- PSE
- ワークフローツール
- グリッド可視化

ネットワーク&認証層

WP5 大阪大学 下條教授

- PKI
- 認証連携 VO対応

Programming層

WP2 産総研 関口博士

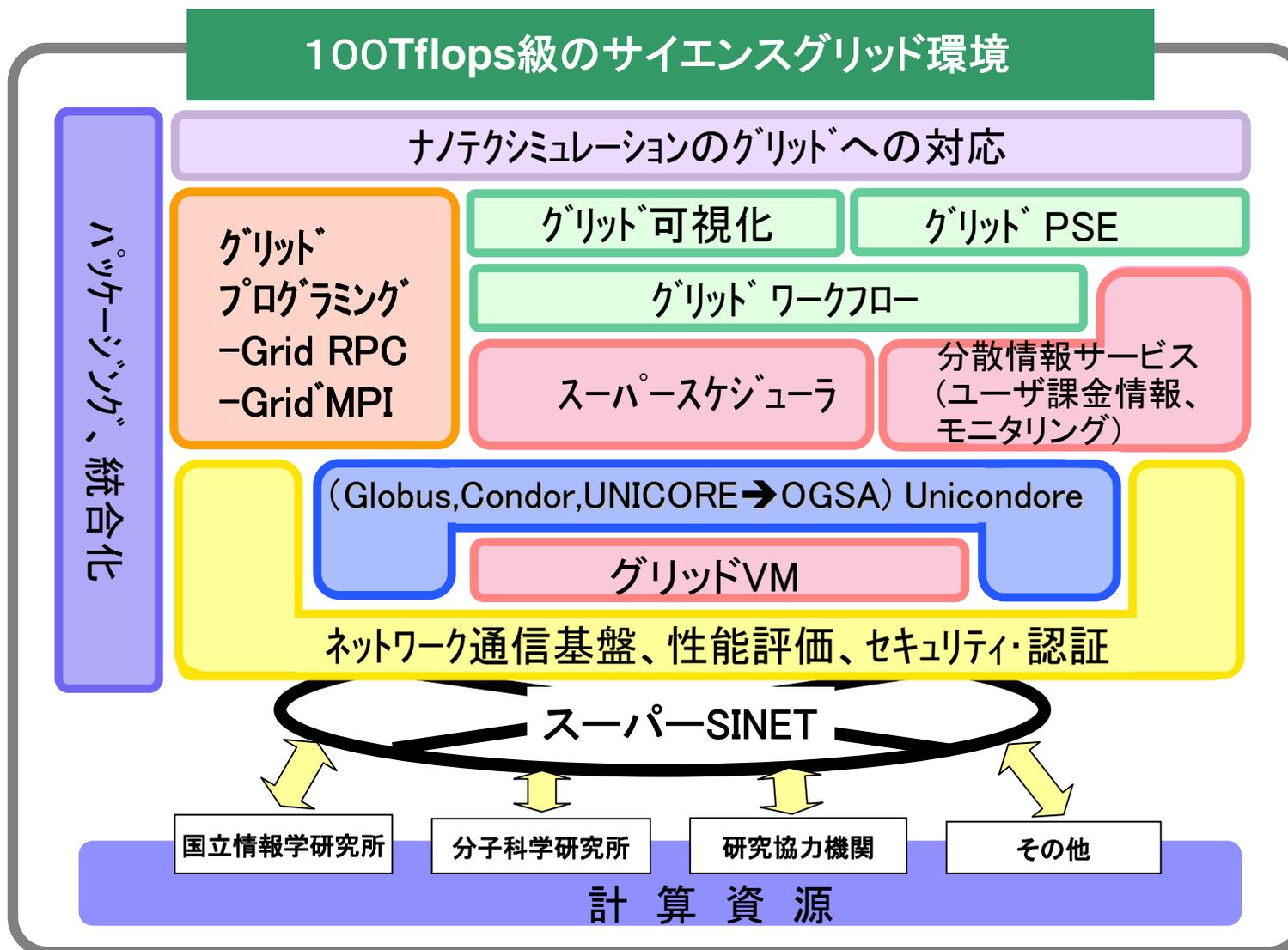
- GridMPI
- GridRPC

Application層

WP6 九州大学 青柳教授

- Mediator
- 連成計算実証試験

NAREGIソフトウェア階層

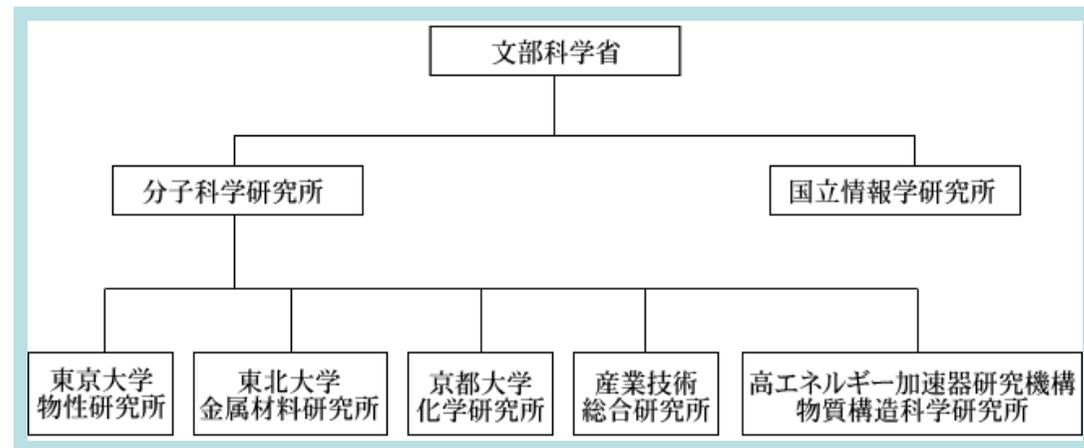




ナノサイエンス実証研究

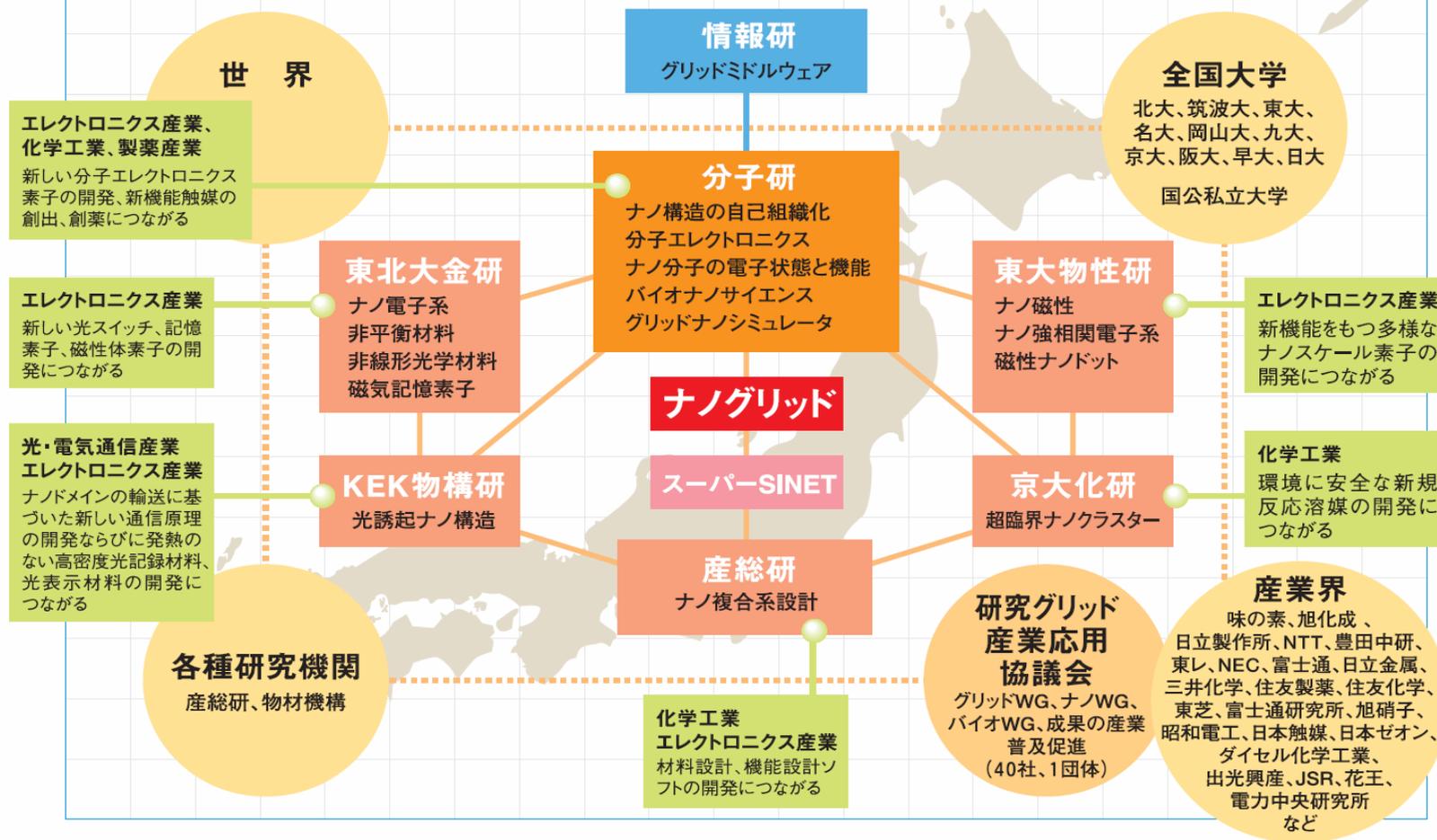
NAREGIのもう一方の柱は、**ナノサイエンス**である。**分子科学研究所**(愛知県岡崎市)では、**ナノサイエンス実証研究拠点**として、主に**ナノアプリケーション**の研究開発が行われている。
<http://nanogrid.ims.ac.jp/nanogrid>

この分野では、分子研を中心にして多くの研究機関をまとめる形で**グリッドリソース**や**ミドルウェア**を利用した研究が進められている。

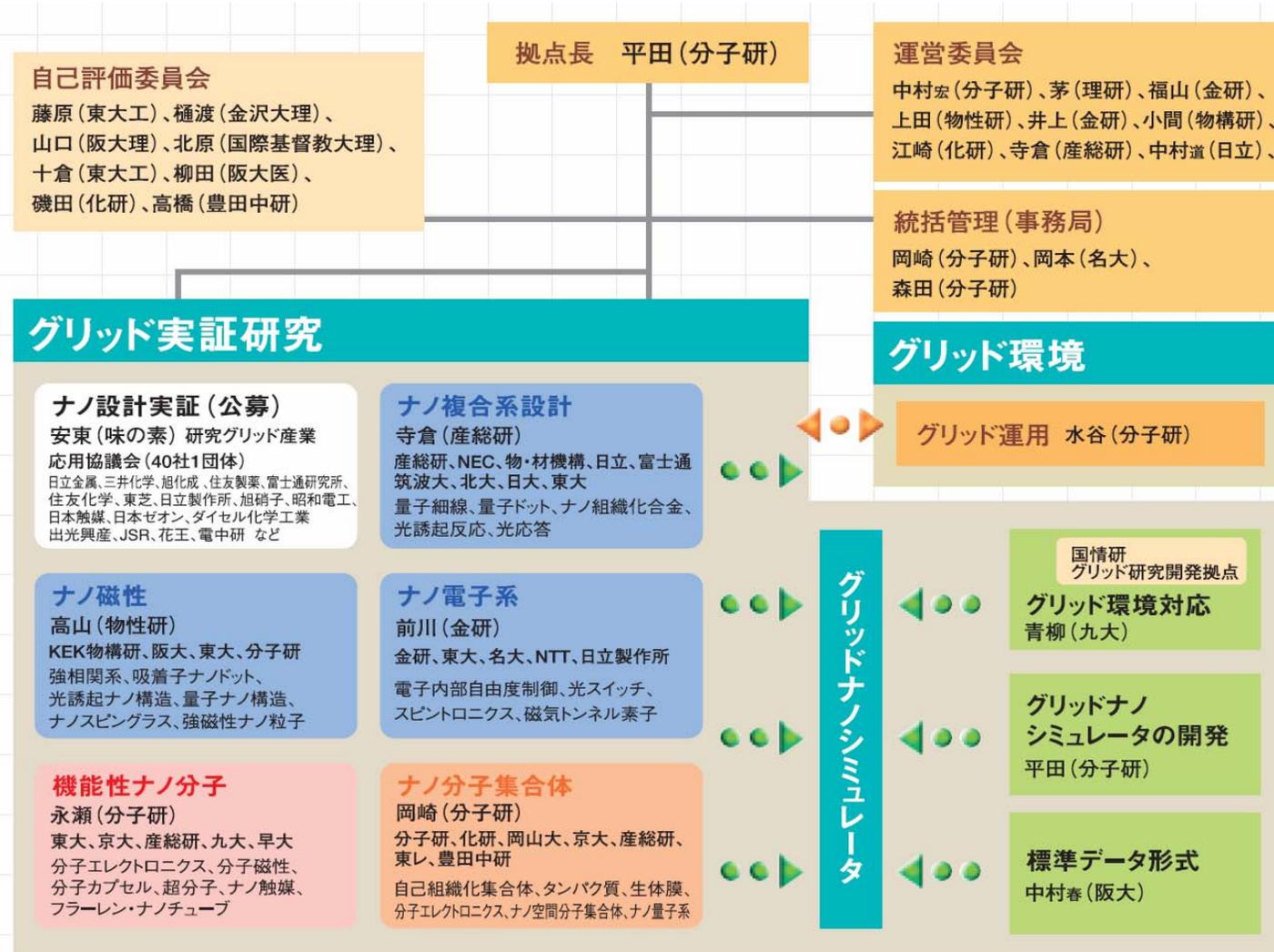


ナノサイエンス実証研究の組織相関

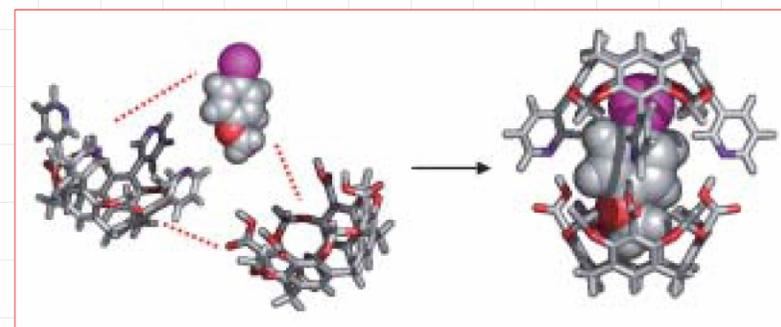
研究体制と期待される成果



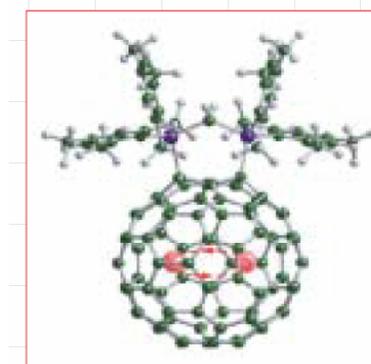
ナノサイエンス実証研究拠点



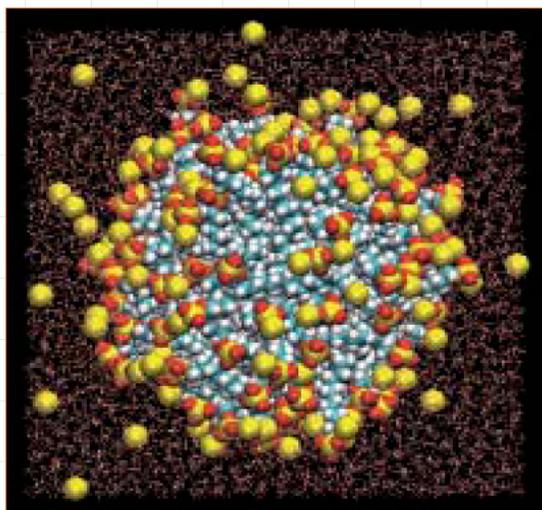
- 機能性ナノ分子(右)
 - 量子化学計算に基づいて新しい機能を持つ分子を設計する。
- ナノ分子集合体(下)
 - たんぱく質や生体膜などの機能発現の原理を、分子動力学や統計力学計算を利用して明らかにする。



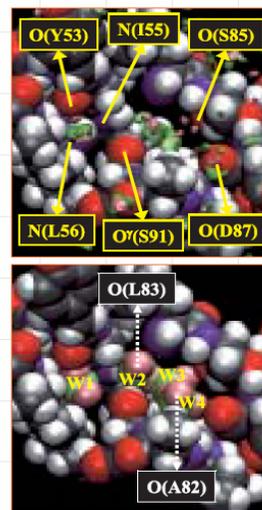
自己集合とカプセル化



化学修飾による回転運動の制御

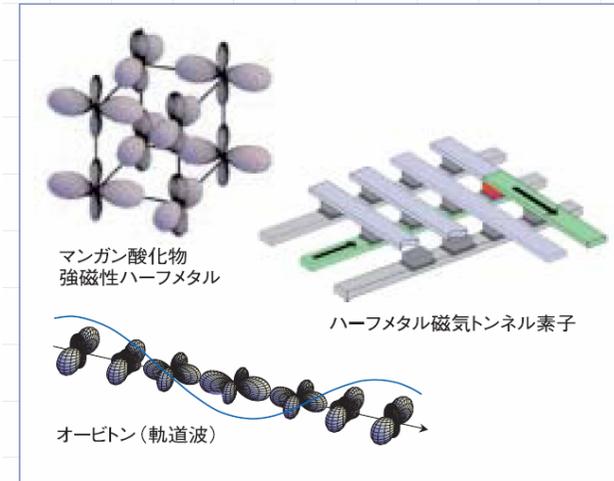


水中のドデシル硫酸ナトリウム (SDS) 球状ミセル

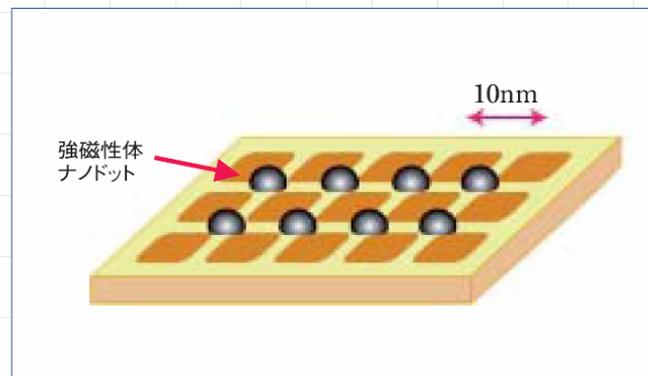


3D-RISMで検出されたタンパク質内部の水分子

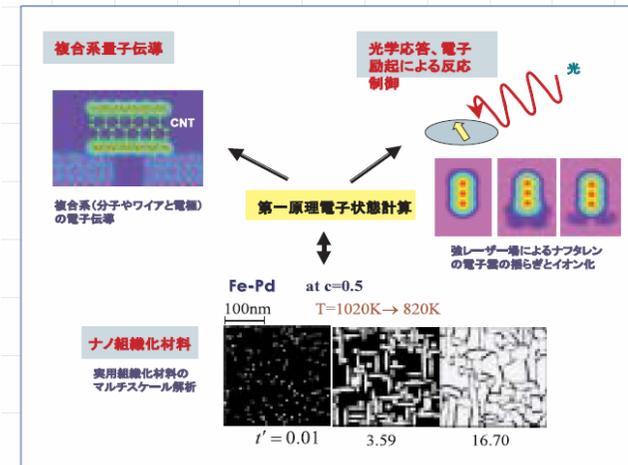
- ナノ電子系(右上)
 - 固体電子論に基づく大規模計算で、ナノ電子スイッチなどへの応用を目指す。
- ナノ磁性(下)
 - 量子モンテカルロ計算などにより、高密度磁気メモリーなどの基礎を研究する。
- ナノ複合系設計(右下)
 - ナノ/マクロ複合系のマルチスケール解析を行う。



マンガン酸化物 強磁性ハーフメタル



自己組織化による磁性ナノドットの配列制御



ナノ複合系設計



NAREGI中間成果報告会

[7月11日(月) 9:30～15:00 国立情報学研究所]

の内容紹介

- ・グリッド研究開発拠点のデモ内容
- ・ナノサイエンス拠点のデモ内容

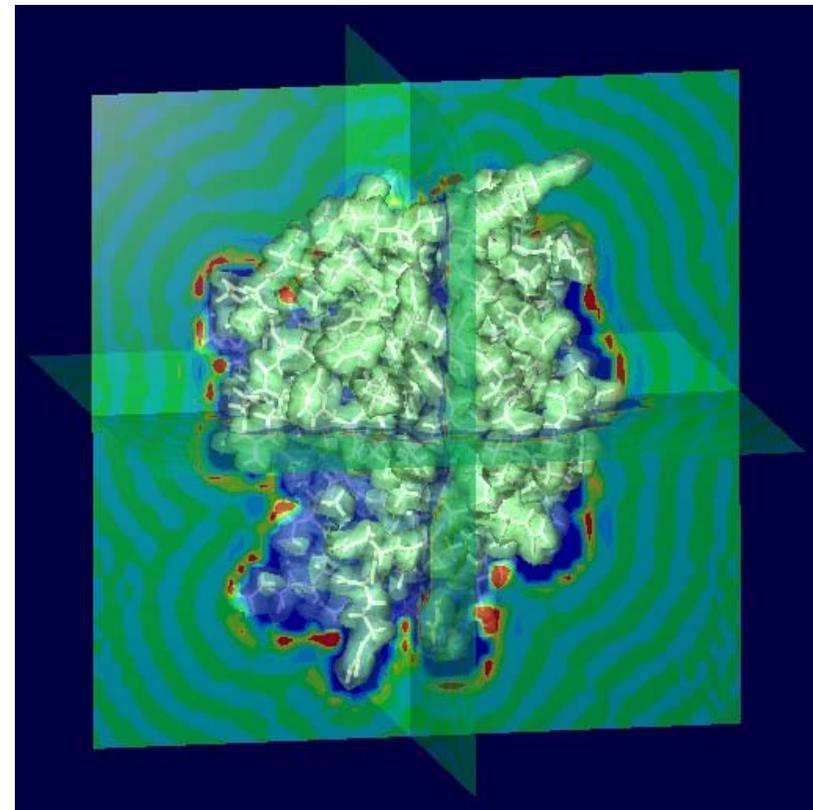
NAREGIプロジェクト中間成果報告会

- 全部で5年のプロジェクトのうち、今年がちょうど真ん中の年になるため、中間評価が行われる。(中間評価は、大きなプロジェクトの場合に行われ、今後の方向性と予算を決定する重要な作業)
- 右のようなプログラムで行われている(本日)。
- 前半がグリッド研究開発分野で、後半がナノサイエンスの分野。

○プログラム	
9:30~9:35	開会の挨拶 松川 憲行 文部科学省研究振興局情報課長
9:35~9:40	情報科学技術委員会 主査 ご挨拶 土居 範久 中央大学理工学部情報工学科教授
9:40~14:55	プロジェクト中間成果報告
9:40~10:20 (質疑 20 分含)	○プロジェクト全体、グリッド基盤ソフトウェア開発 三浦 謙一 国立情報学研究所教授
10:20~10:30	○グリッド環境における資源管理 松岡 聡 東京工業大学学術国際情報センター教授
10:30~10:40	○グリッドプログラミング環境 関口 智嗣 産業技術総合研究所グリッド研究センター長
10:40~10:50	(質疑応答)
10:50~11:00	○グリッドアプリケーション環境 宇佐見 仁英 国立情報学研究所教授
11:00~11:10	○グリッド対応通信基盤 下條 真司 大阪大学サイバーメディアセンター教授
11:10~11:20	○ナノシミュレーションのグリッド化 青柳 睦 九州大学情報基盤センター教授
11:20~11:30	(質疑応答)
11:30~12:00 (質疑 10 分含)	○ナノ分野アプリケーション開発・実証 平田 文男 分子科学研究所教授
(12:00~12:50)	昼食・休憩
12:50~13:00	○ナノグリッドシミュレータの開発 岡崎 進 分子科学研究所教授
13:00~13:10	○ナノ分子集合体実証研究 岡崎 進 分子科学研究所教授
13:10~13:20	○機能性ナノ分子実証研究 永瀬 茂 分子科学研究所教授
13:20~13:30	(質疑応答)
13:30~13:40	○ナノ電子系実証研究 前川 禎通 東北大学金属材料研究所教授
13:40~13:50	○ナノ磁性実証研究 高山 一 東京大学物性研究所教授
13:50~14:00	○ナノ複合系設計実証研究 寺倉 清之 産業技術総合研究所研究コーディネータ
14:00~14:10	(質疑応答)
14:10~14:50	○統合デモンストレーション
14:50~14:55	全体質疑
14:55~15:00	閉会の挨拶 星野 利彦 文部科学省研究振興局情報課企画官

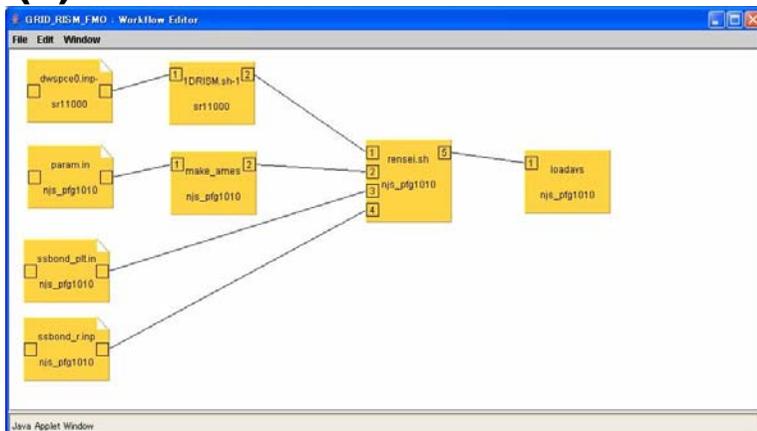
グリッド研究開発拠点のデモ内容(1)

- RISM-FMO連成の大規模計算をNIIと分子研の計算機リソースを利用して行う。
- 計算対象分子: リゾチウム(2021原子)
- 計算機リソース:
 - (RISM) 分子研SR11000 16CPU
 - (FMO) NII PCクラスタ 200CPU
- 並列実行はGridMPIによる
- SuperScheduler, GridVM によりコアロケートし、JSDLで記述した資源要件に従って、NAREGI Workflowから実行
- 計算資源の負荷を情報サービスで監視
- 計算結果をグリッド可視化で表示

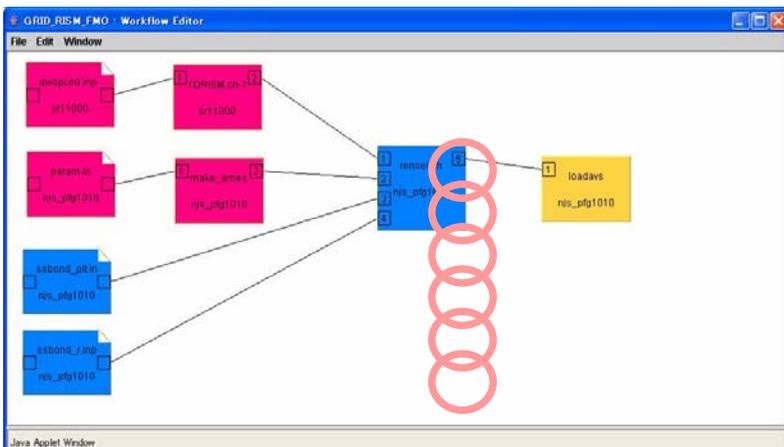


グリッド研究開発拠点のデモ内容(2)

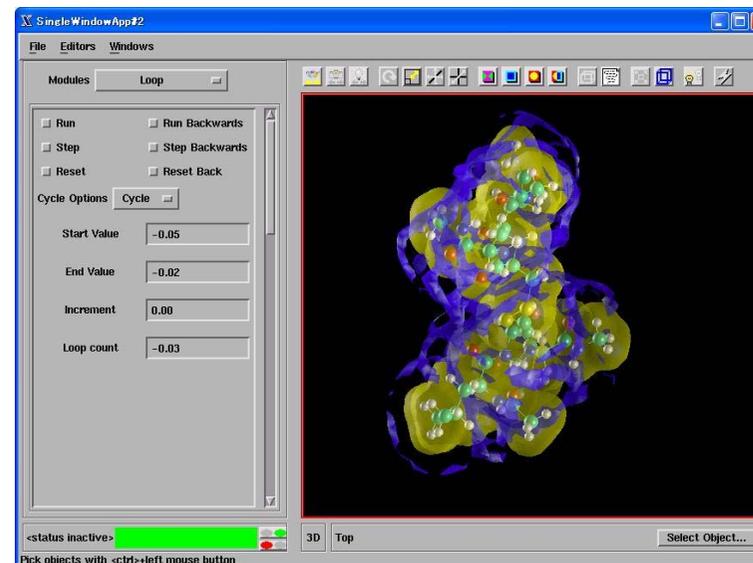
(1) NAREGI Workflowによる記述



(2) 実行



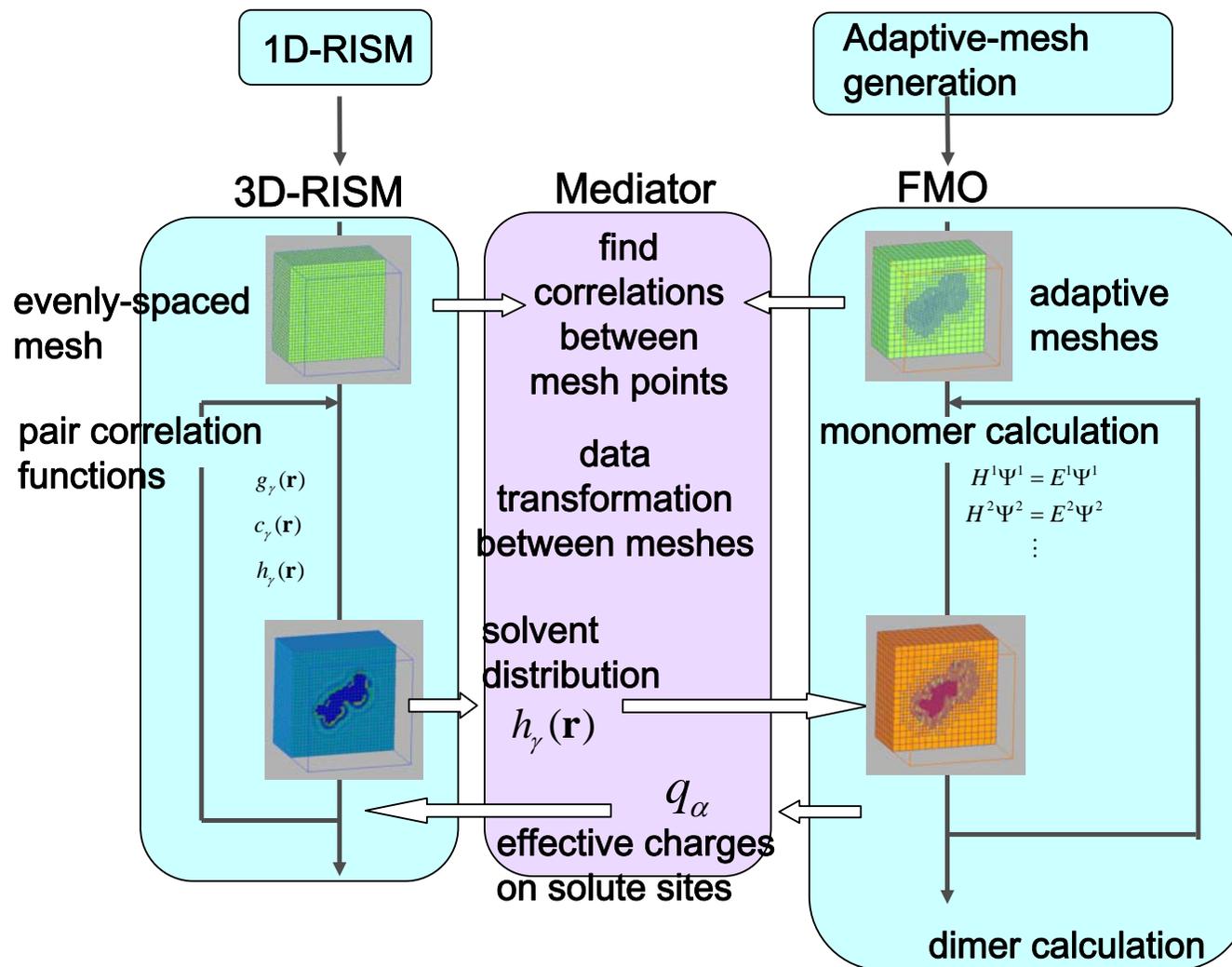
(3) グリッド可視化システムによる表示



連成計算内容の詳細

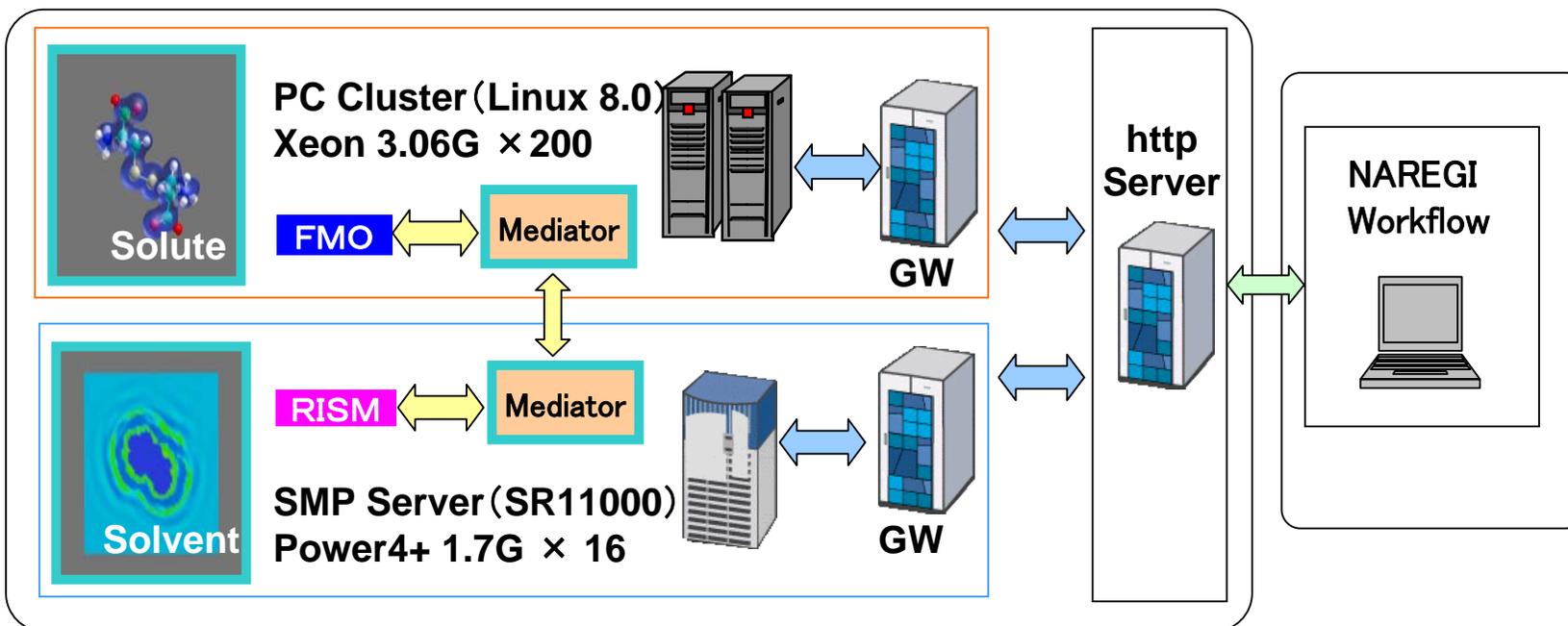
- FMO: 大規模な系の第一原理量子化学計算が可能
分子軌道法、1電子積分、2電子積分、対角化
- RISM: 溶媒-溶質間相互作用の統計力学計算
三次元周期境界箱中のFFT計算、古典力学的相互作用
- データ交換はMediator (九大-日立共同開発)
異なる物理表現間の意味変換
- 今回のデモでは、サイエンスの内容よりも、すべての
グリッドミドルウェアが統合して動くことが重要
- 計算時間:
RISM-FMO連成計算の収束までに約10日
可視化のためのデータ生成(全MO計算)に約30時間

RISM-FMO連成プログラム



デモの実行

NII Resources



IMS Resources



ナノサイエンス実証研究拠点のデモ

- これまでに手に入っている内容を紹介
 - 1000万原子のMDシミュレーション
 - ただひたすら大規模なシミュレーション
 - MD-MO連成シミュレーション
 - 市販プログラムを連成に組み込む
 - レプリカ交換MC-RISMシミュレーション
 - 溶媒中でのエネルギー最小状態の探索

1000万原子のMDシミュレーション

- 世界最大の分子動力学(MD)計算(通常はせいぜい100万)
- 水分子約300万個の運動
- これは、現在の計算機の限界に近いが、実験で使われる水の量とはまだ何桁も開きがある。
- より正確な物理量を予言するには、もっと大きな系で計算したいところ。

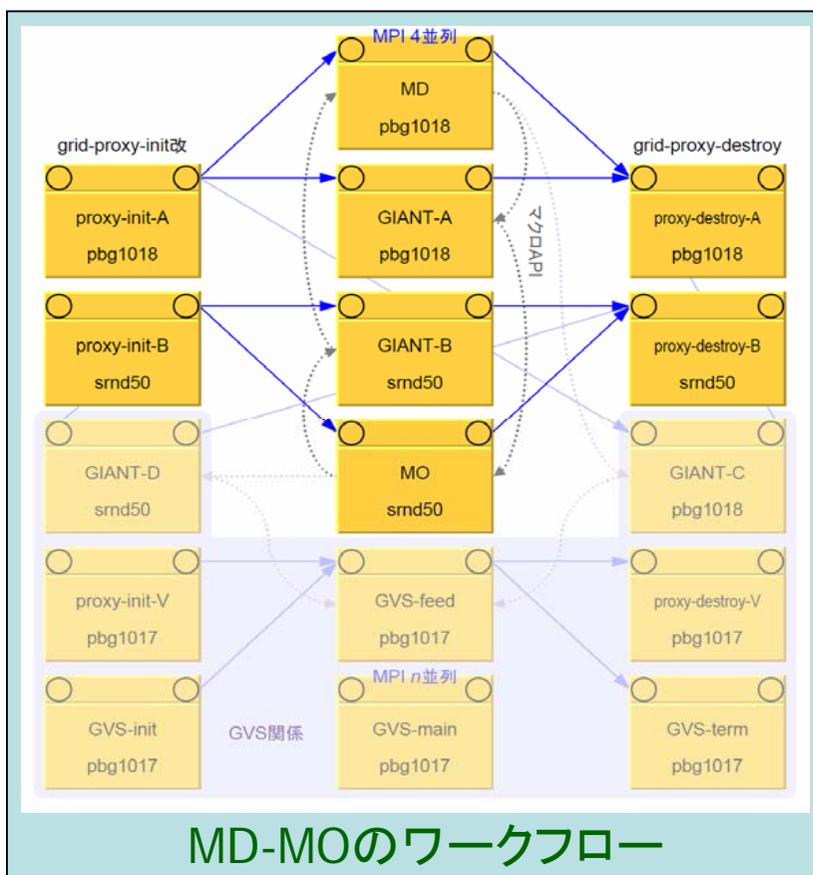


MD-MO連成シミュレーション(1)

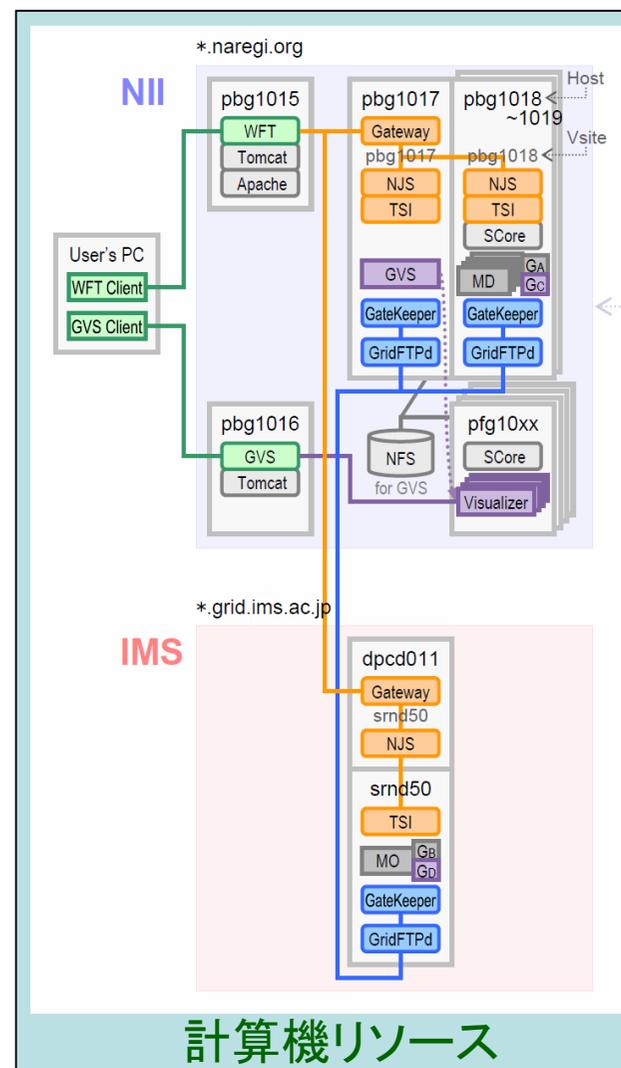
- Gaussianという**市販プログラム**による量子化学計算(分子軌道(MO)法)と、分子動力学(MD)計算との**連成シミュレーション**
 - 注目している分子にはMO法を用いて、量子力学的な力を求める。
 - 周りの溶媒分子に働く力は、通常の古典力学的な力の計算による。
 - 全体の分子の運動は、MD法による。
- **シミュレーションの内容**
 - 化学反応: $\text{ClCH}_3 + \text{NH}_3 \rightarrow \text{Cl}^- + \text{CH}_3\text{NH}_3^+$
- データ交換は分子研が開発している **GIANT** (General Inter-Application Translator)を利用
- **グリッド可視化システム**との連携により結果を表示

MD-MO連成シミュレーション(2)

- 使用したリソース
 MO: 分子研SMPマシン(SR11000)
 MD: NIIクラスター
 可視化: NIIクラスター



MD-MOのワークフロー



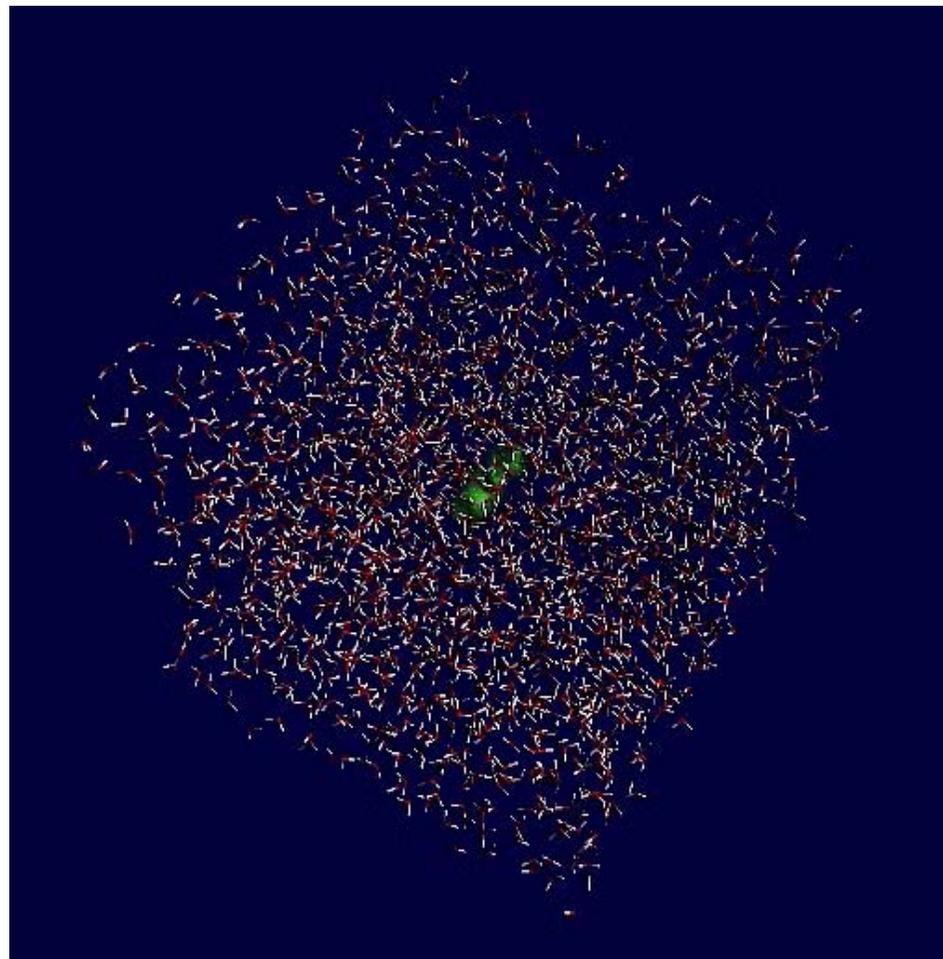
計算機リソース

MD-MO連成シミュレーション(3)

- 左の分子:
 ClCH_3
- 右の分子:
 NH_3

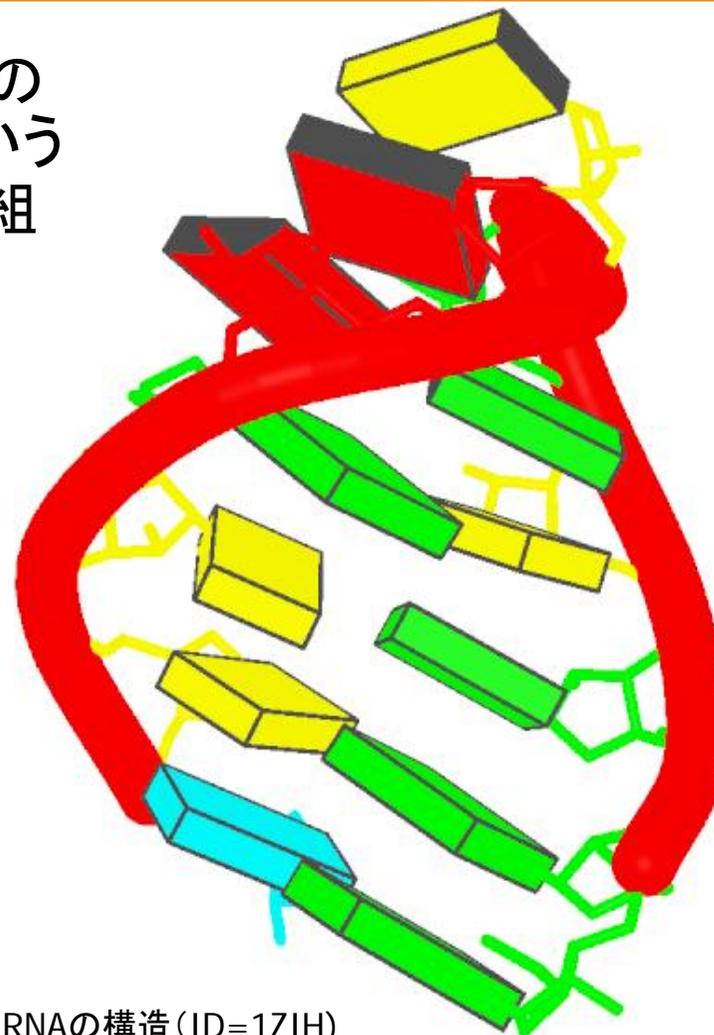


- 結合の切断:
 $\text{Cl} - \text{x} - \text{CH}_3$
- 新たな分子:
 CH_3NH_3^+



レプリカ交換MC-RISM(1)

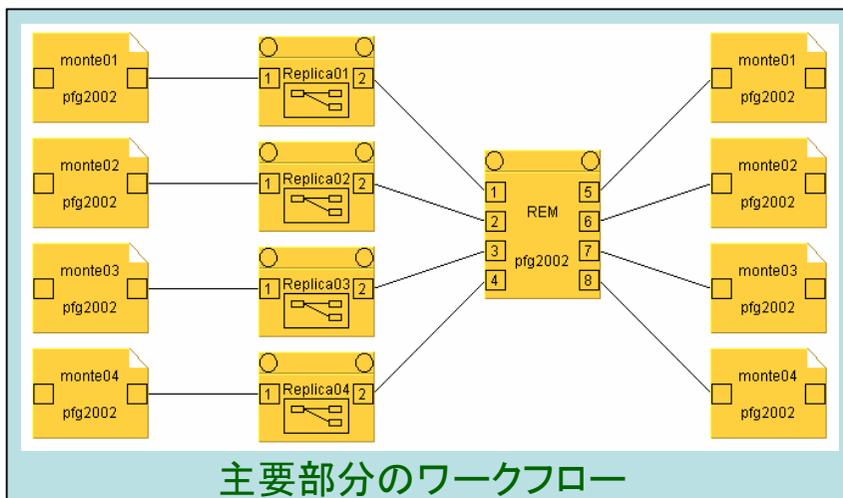
- エネルギー最小構造を求めるためのモンテカルロ(MC)計算をRISMという溶液の統計力学計算プログラムと組み合わせる。
- 計算対象: GCAA RNA Tetraloop
- PDB-ID: 1ZIH
- 塩基配列: G G G C G C A A G C C U
- 原子数: 389



RNAの構造 (ID=1ZIH)

<http://ndbserver.rutgers.edu/>

レプリカ交換MC-RISM(2)

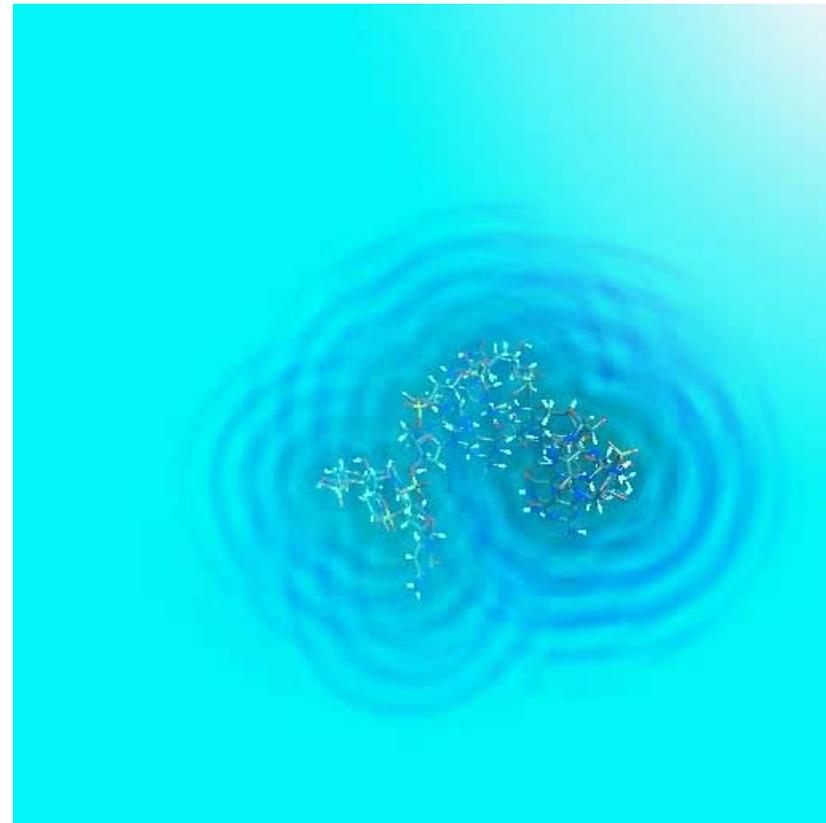


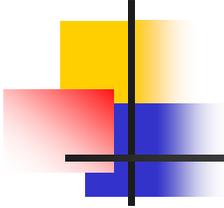
- 計算のワークフロー
 - 繰り返し部分の構造は左の通り
- 計算機リソース
 - NIIとIMSのPC/SMPマシンを連携



レプリカ交換MC-RISM(3)

- 計算方法: レプリカ交換モンテカルロ
 - 同種のレプリカ分子を用意して、それぞれにRISM-モンテカルロ(MC)連成計算を行い、時々レプリカ間で温度などを交換することで、局所最小状態にトラップされることを避ける方法。
 - パラメタ交換以外の時間では、レプリカ間はまったく独立に運動するため、効率のよい並列化が可能。
- 分子のエネルギー計算は古典力学的相互作用による。溶媒部分の記述はRISM (Reference Interaction Site Model)による。
- 多くの原子が含まれるたんぱく質や塩基鎖のMC計算では、いかに局所最小状態にとらわれずに全体の構造をサーチできるかが重要





事務連絡

- 広域分散アプリの講義はこれでおしまいです。
- すでに提出されたレポート、および、出席点により成績評価がなされます。
- 試験はありません。
- お疲れ様でした。