

Mediatorを用いた RISM-FMO連成シミュレーション

国立大学法人 九州大学 青柳 睦
(株)日立製作所 基礎研究所 何 希倫
(株)日立超LSIシステムズ 久保昭一, 福田和幸

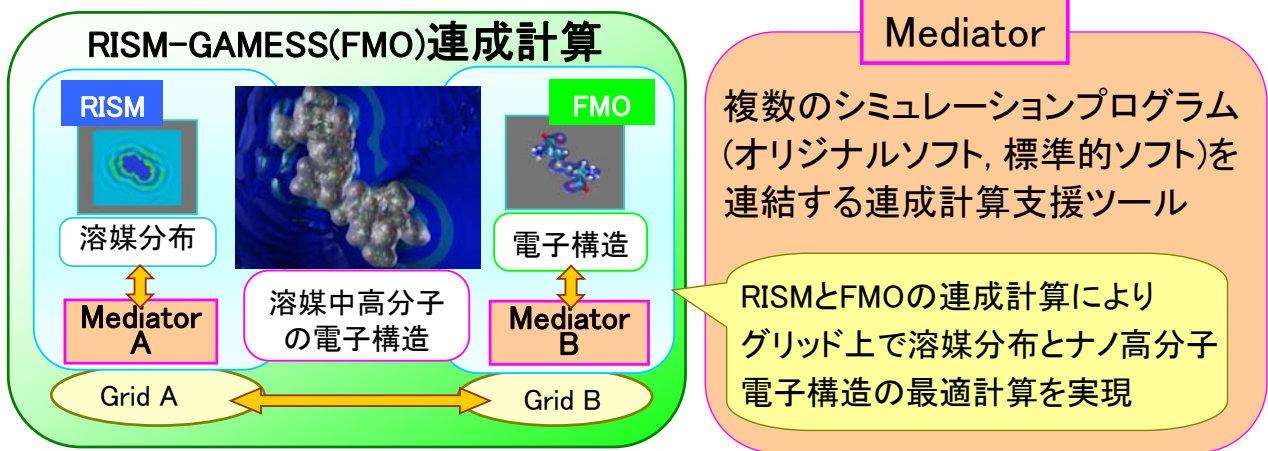
National Research Grid Initiative

目次

1. 序 連成計算プログラムの開発課題
2. Mediatorの基本機能とシステム構成
3. RISM-FMO連成シミュレーション
4. 大規模連成計算の負荷分散
5. まとめ

National Research Grid Initiative

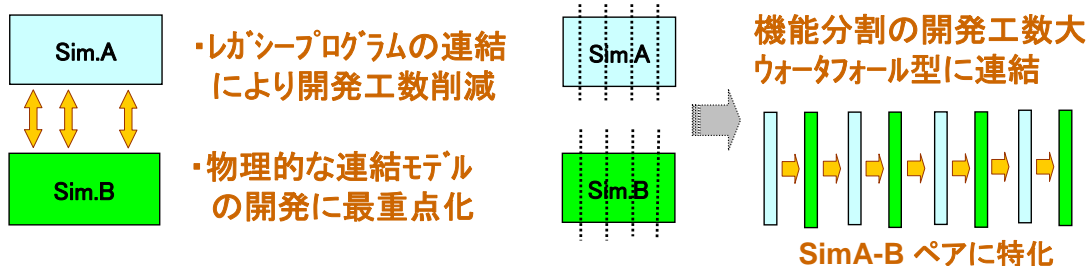
空間・時間スケールの異なる物理現象が複雑に絡み合った
マルチスケール・マルチフィジックス問題の解析ニーズ拡大



RISM	Reference Interaction Site Model	無限系における溶媒分布解析プログラム 分子科学研究所開発
FMO	Fragment Molecular Orbital method	ナノ高分子の電子構造解析プログラム 産業技術総合研究所開発, openソース GAMESS導入

連成計算プログラムの開発課題

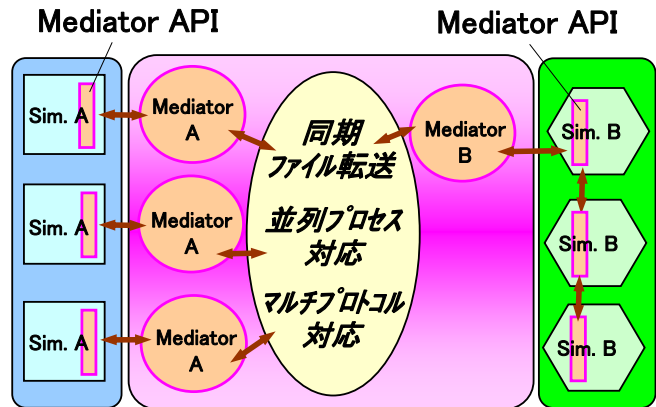
- ❖ レガシーなシミュレーションプログラムに対して
最小限の変更で反復的なデータ交換を実現



- ❖ 同時に実行中の複数シミュレーション間で
同期的なデータ転送は不可欠
- ❖ 広域グリッド上の疎結合計算とローカルサイト上の
密結合計算を統一的な連結方式で実施したい
- ❖ 物理モデル・計算手法の異なるシミュレーション間で
物理量を保存したデータ変換

Mediatorの基本機能

- ❖ データ交換を支援するMediator API関数
レガシープログラムにAPI関数を追加し簡便にデータ交換を実現
- ❖ 同期ファイル転送
同時実行中の複数シミュレーションプロセス間で
同期的にファイル転送を実施するためにポーリング機構を導入
- ❖ 並列プロセス対応
MPI 並列プロセス間の
ファイル転送
- ❖ マルチプロトコル対応
同一APIを用いたGridFTP/
scp/rcp/NFSのファイル転送
- ❖ 意味(セマンティック)変換
異なる物理モデル・計算手法
の物理データ変換を支援

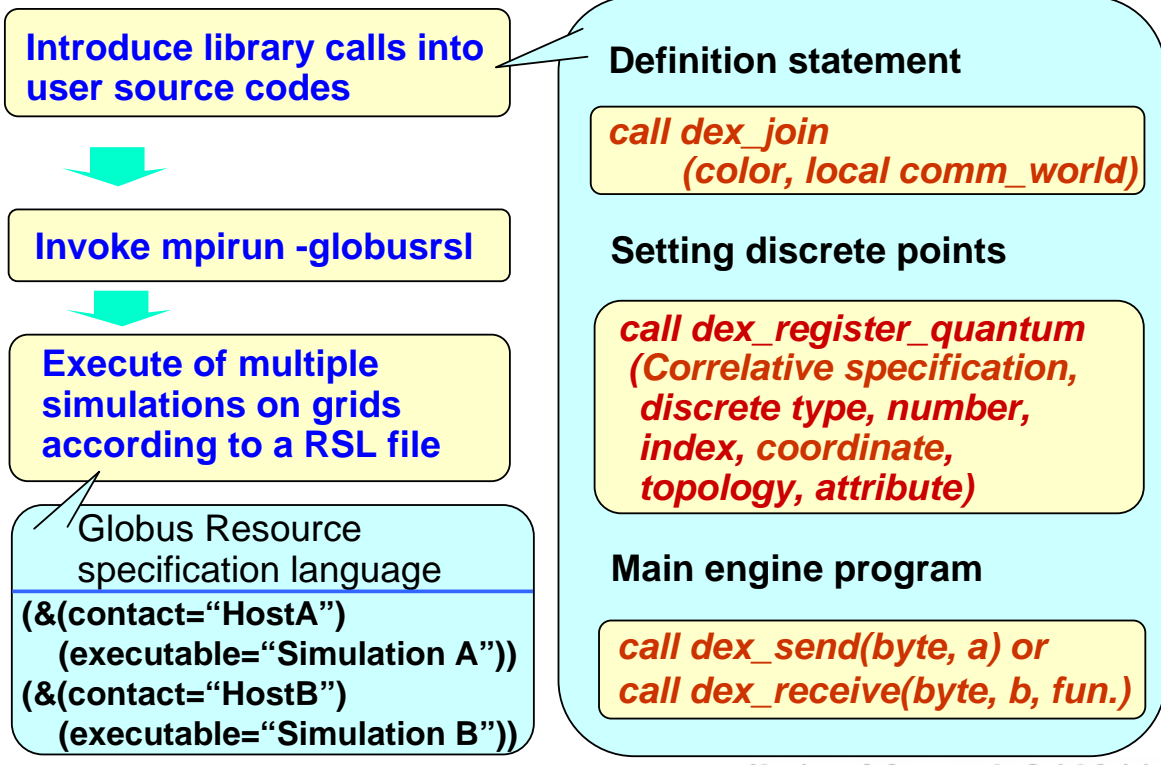


システム構成の階層構造

- ❖ Mediatorは同期制御・データ転送・並列プロセス管理を行う
メインユニットとプログラムに追加するAPIライブラリを構成
- ❖ Mediatorを介したデータ交換として、SBC (Storage-based
communication)のGridFTPなどによる同期ファイル転送と
GridMPIによるデータ通信をNAREGIミドルウェア上で実現

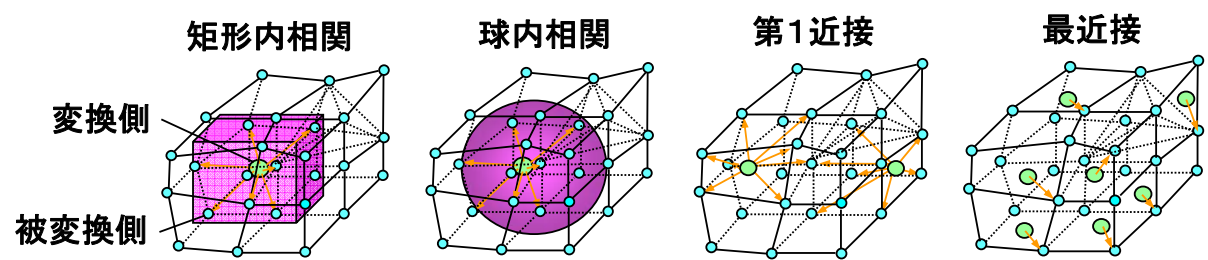
Simulation A		Mediator (Main unit)	Simulation B	
Mediator API			Mediator API	
SBC*1			GridMPI	
GridFTP	rcp/scp	NFS		
NAREGI Middleware				
Globus				
AIX・Linux				
Massive Parallel Computer・PC Clusters				

*1 Storage-based Communication



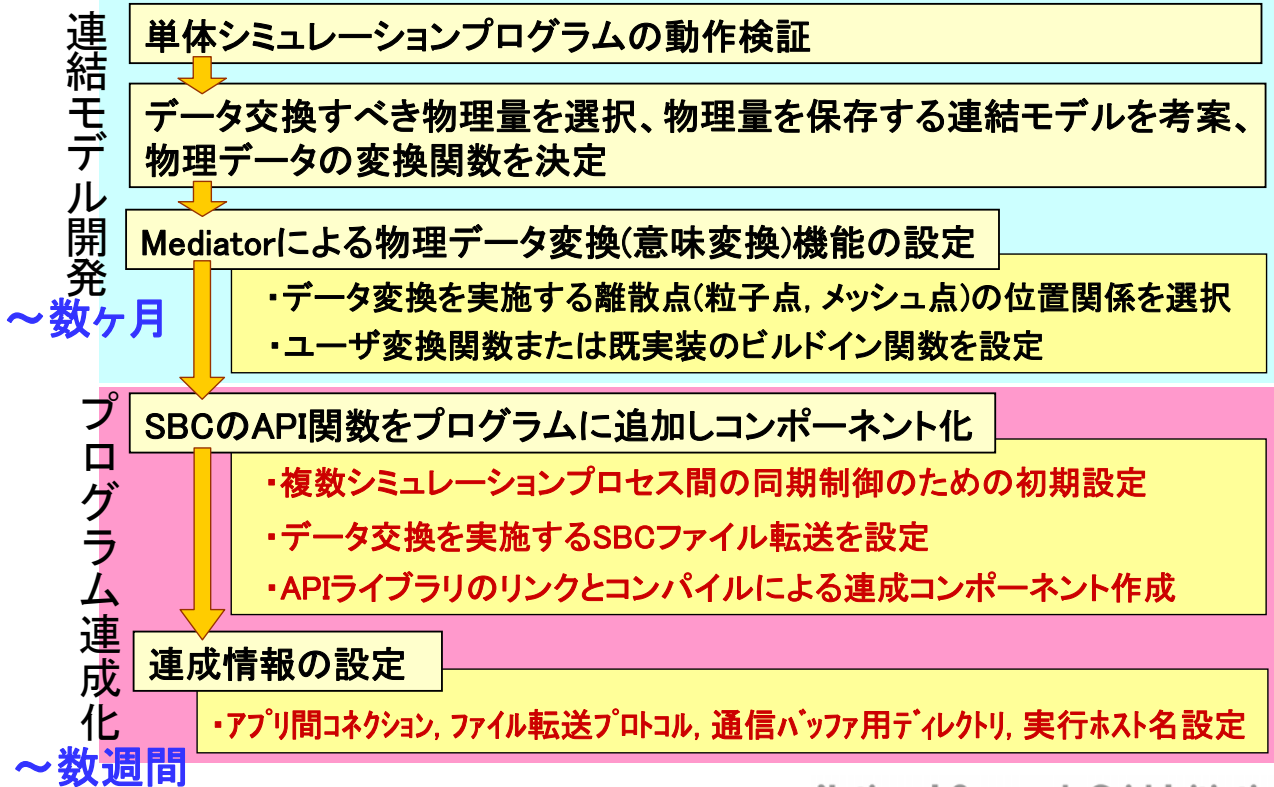
高度セマンティック変換

- 物理量を代表する離散点の位置関係を定義、物理データの設定プロトコルを元に物理保存則を満たす形式でデータ変換
 - 物理モデル: 粒子数, 運動量, エネルギー, 力場等の保存則
誘電率, 電荷量, 散乱レート等(量子論と古典論ギャップ)
 - 数値計算モデル: 離散手法, 座標系変換, 単位系, etc.



- 物理モデルや計算モデルを融合するための抽象階層を設定
 効率的な連結を支援する高度なセマンティック変換技術を開発
 - 異なる機能のプログラムを融合し, 高い付加価値を創生
 - 異なる計算精度を融合し高精度性と高速性の両立を追究

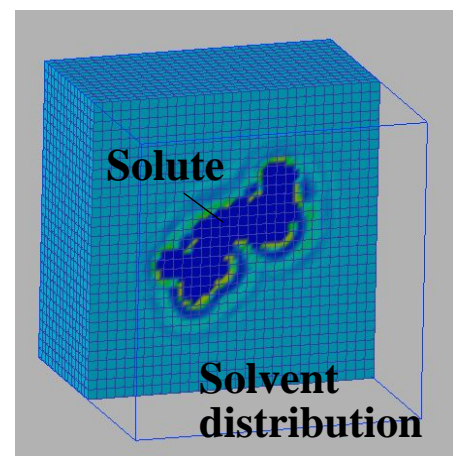
SKIP 連成計算プログラムの構築手順



National Research Grid Initiative

シミュレーション・モデル(RISM*)

- RISM (Reference Interaction Site Model)法を用いて溶媒分子と溶質分子の間の分布関数を解析。
- Closureモデル: PLHNC
van der Waalsモデル: Amber
溶質: Lysozyme 129アミノ基
1960原子
解析領域: 54 \AA^3
メッシュ構造: 128^3 等間隔メッシュ
収束条件: 残差 $< 1 \times 10^{-7}$



■ High density

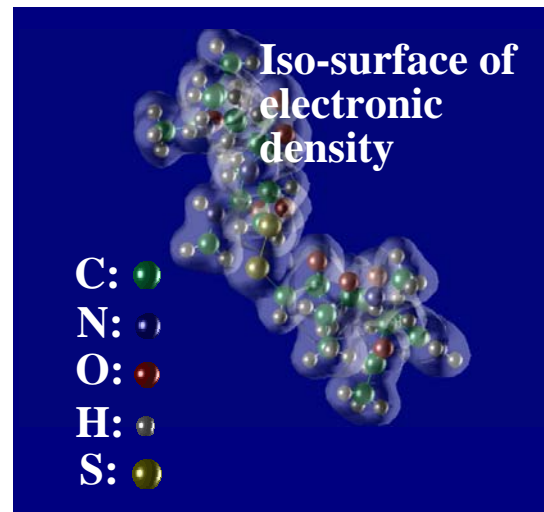
■ Low density

*オリジナルコードは 分子科学研究所にて開発。

National Research Grid Initiative

- FMO (Fragment Molecular Orbital method) を用いて溶質の電子エネルギー及び構造を計算。溶質分子はフラグメントに分割され、フラグメント間は静電相互作用を考慮。

- 溶質: Lysozyme 129アミノ基
電子状態計算: DFT
基底関数: 6-31G*, 16973 基底数
溶質: 1960 原子
収束条件: 残差 $< 1 \times 10^{-5}$

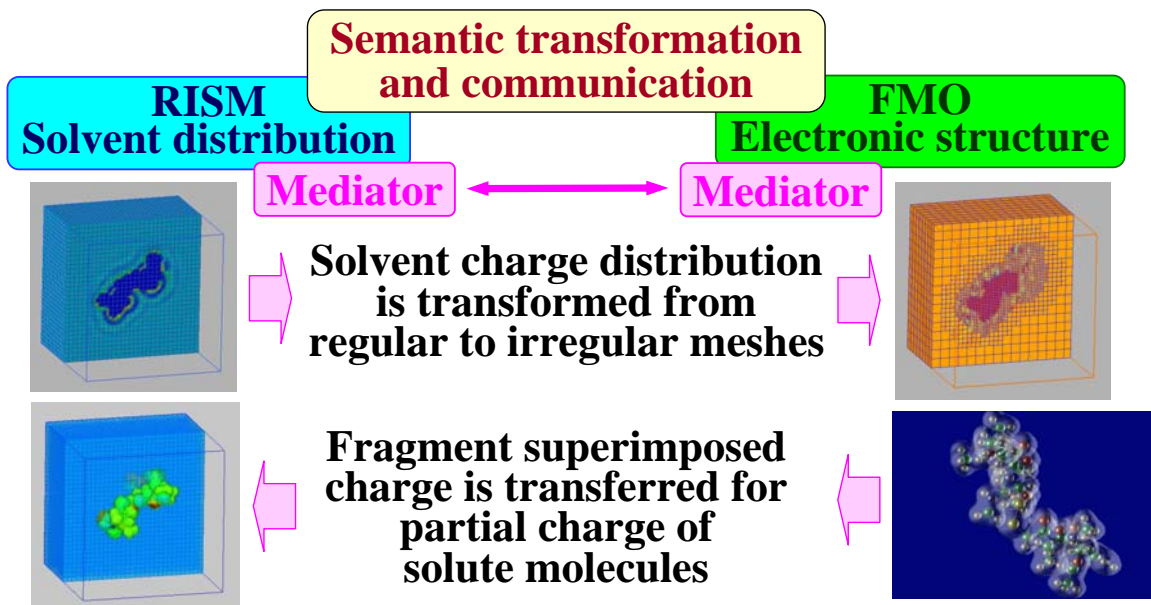


*オリジナルコードは産業技術総合研究所にて開発

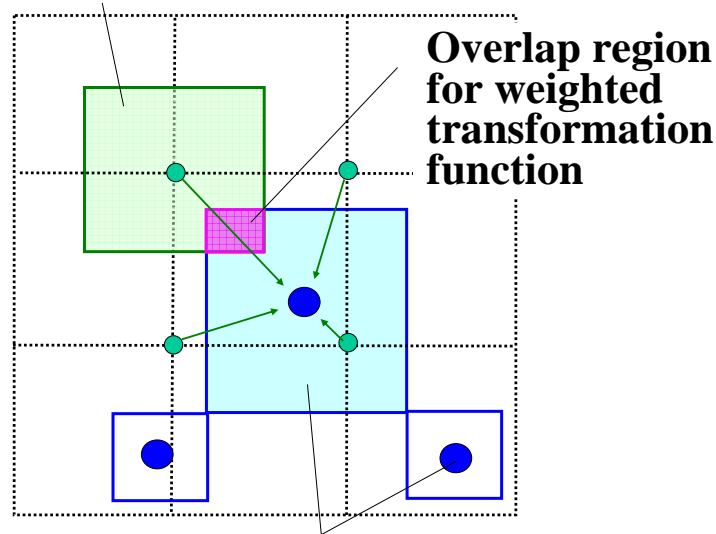
National Research Grid Initiative

RISM-FMO coupled simulation

Electronic structure of Nano-scale molecules in solvent is calculated self-consistently by exchanging solvent charge distribution and partial charge of solute molecules.



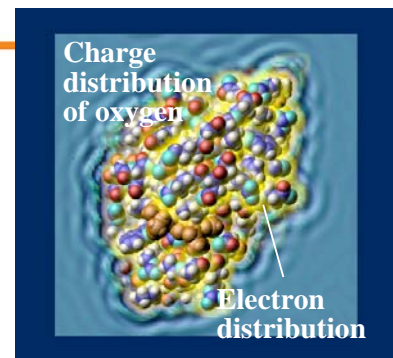
Correlated uniform mesh points with shared region



Scattered particle points with rectangle region

RISM-FMO連成計算の対象と条件

溶媒中Lysozyme分子(129アミノ基,1960 原子) に対して、下記のシミュレーション条件で RISM-FMO連成計算を実施する。



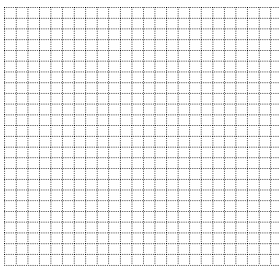
	シミュレーション条件	使用リソース
RISM	解析領域: 54 Å ³ メッシュ構造: 128 ³ 等間隔メッシュ 収束条件: 残差 < 1 × 10 ⁻⁷	CPU (SX-7)数=1 メモリ2GB
FMO	電子状態計算: DFT 基底関数: 6-31G*, 16973 基底数 非等間隔メッシュ数: 184万 収束条件: 残差 < 1 × 10 ⁻⁵	CPU(PCクラス)数 =並列プロセス数 メモリ2GB

RISM-FMO連成における高演算負荷部

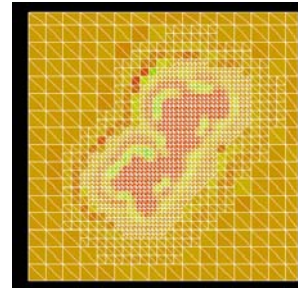
溶媒電荷-フラグメント(溶質)間の1電子積分計算

$$F_{\mu\nu}^I \leftarrow \sum_{\gamma} \rho q_{\gamma} h_{\gamma}(\mathbf{r}_g) \langle \mu | \frac{1}{|\mathbf{r}' - \mathbf{r}_g|} | \nu \rangle \quad (\mu, \nu \in I)$$

RISM 等間隔メッシュ



FMO アダプティブメッシュ



離散点 変換

$$h_{\gamma}(\mathbf{r}_g)$$

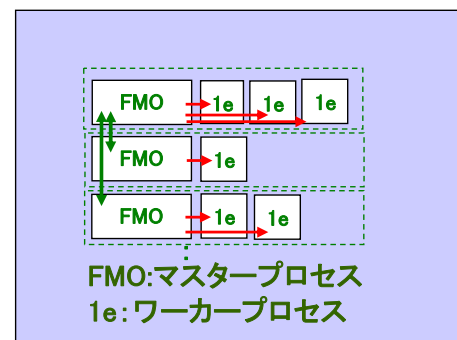
National Research Grid Initiative

負荷分散アルゴリズム

FMOでは、フラグメントを最小単位としてモノマー間SCF計算を複数のマスタープロセスに分配し並列計算する。今回、溶媒電荷-フラグメント(溶質)間に関する1電子積分計算を溶媒原子単位で、マスタープロセスとワーカープロセスに新たに分配する負荷分散機構を新たに開発。

全PE数:	$NPE_{all} = \sum_{i=1}^{N_{fragment}} \{ NPE_{master,i} + NPE_{workr,i} \}$
フラグメント <i>i</i> のマスターPE数:	$NPE_{master,i} = 1$
フラグメント <i>i</i> のワーカーPE数:	$NPE_{workr,i} = T_{hsandt1atom} \times (Nv - Nv_{master,i}) / (T_{th} - T_{others})$

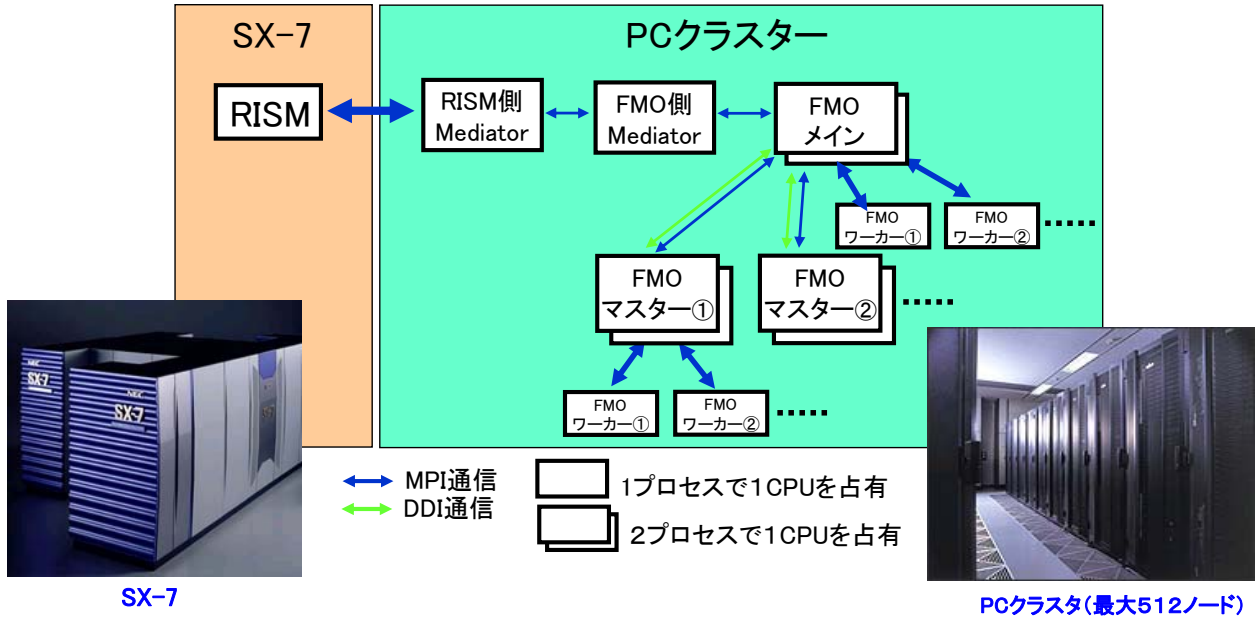
- $N_{fragment}$: フラグメント数
- Nv : 溶媒全原子数
- $Nv_{master,i}$: マスターで1電子積分計算する溶媒原子数
- T_{th} : 反復1回当たりの計算時間上限
=MAX(FMOプロセス計算時間、RISM計算時間)
- $T_{hsandt1atom}$: 溶媒1原子当たりの1電子積分計算時間
- T_{others} : モノマーSCFの波動関数計算時間



National Research Grid Initiative

連成計算のコンフィグレーション @理研RSCC

RISMをSX-7上、FMOの各プロセス(メイン、マスター、ワーカー)をPCクラスター上に配置。また、データ送受信を支援するRISM側のMediatorとFMO側のMediatorをPCクラスター上に配置。



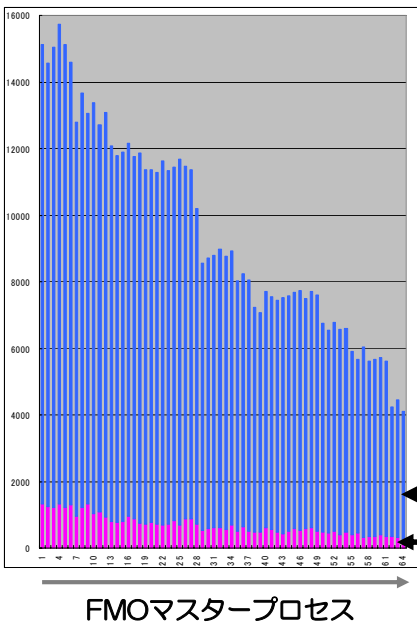
SX-7

PCクラスター(最大512ノード)

連成計算の実測結果(マスタープロセスのみの場合)

SX-7と64台のPCクラスター使用してRISM-FMOの大規模計算を実施し、時間測定を実施。全フラグメント数は125個により、各マスタープロセスには、1~2個のフラグメント計算を分配する

最大15734秒



詳細分析

各マスタープロセスの計算時間を分析するため、フラグメント毎、サブルーチン毎の詳細分析を実施。

サブルーチン毎計算時間(秒)

フラグメントNo.	1	2	3	4	5
アミノ基	LYS001	VAL002	PHE003	GLY004	ARG005
原子数	22	16	20	7	24
基底数	135	123	183	66	191
HSANDT	4103	3870	6871	1412	7041
JANDK	11	11	34	1	32
FMOESP	193	205	375	71	424
WFN	15	22	76	3	65
その他	12	17	29	7	36

HSANDT: 1電子積分ルーチン

溶媒1電子積分計算時間

溶媒1電子積分以外の計算時間

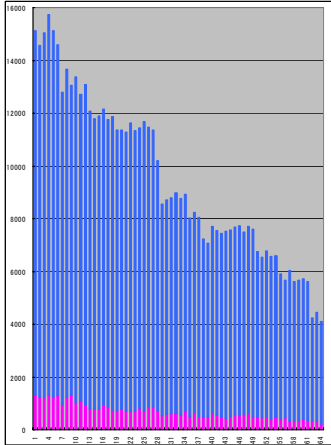
FMO側

FMOマスタープロセス

負荷分散アルゴリズムの計算時間予測

連成計算の実測と分析結果をもとに、負荷分散アルゴリズムを適用しPCクラスタ512台を利用した場合の計算時間と加速化率^{*}を予測。RISMが十分に高速化され1440秒の計算時間とすれば、 $T_{th}=1440$ 秒として高い加速化率を得ることができる。

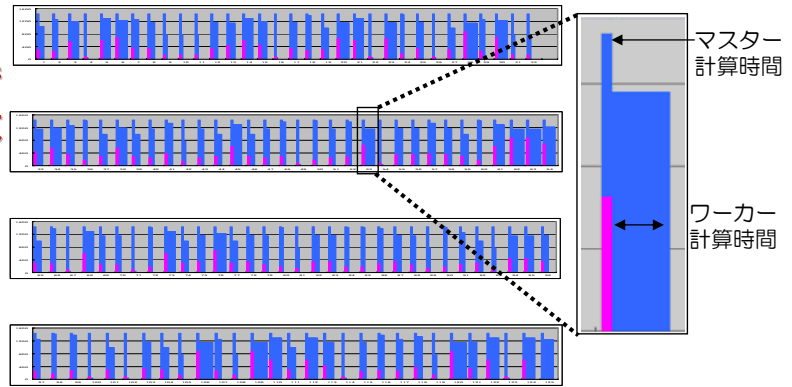
マスターPE数：64
最大15734秒



青：1電子積分計算時間
赤：1電子積分以外の計算時間

マスター PE数：125 } 497並列
ワーカー PE数：372 } 加速化率 295
 $T_{th}=1440$ 秒

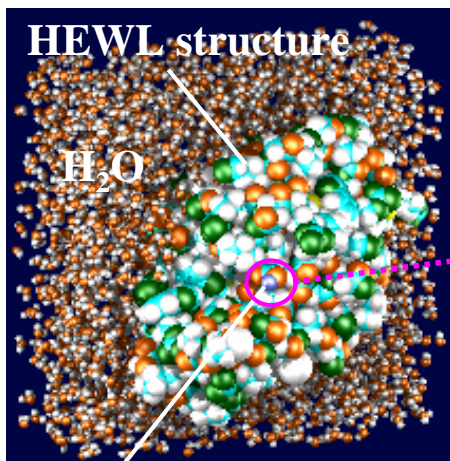
負荷分散



^{*}加速化率=全FMOプロセス計算時間合計/ T_{th}

Solvent interactions of hydrolysis in Lysozyme





A catalytic mechanism of hydrolysis in Lysozyme is Analyzed by RISM-FMO, in which proton transfer leads to bacteriolysis in peptidoglycan of bacterial cell wall.

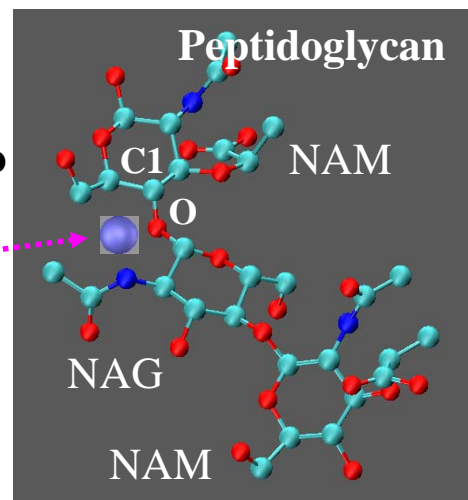


Active site of Glu35

Bacteriolysis by hydrolysis due to H dissociation

H^+

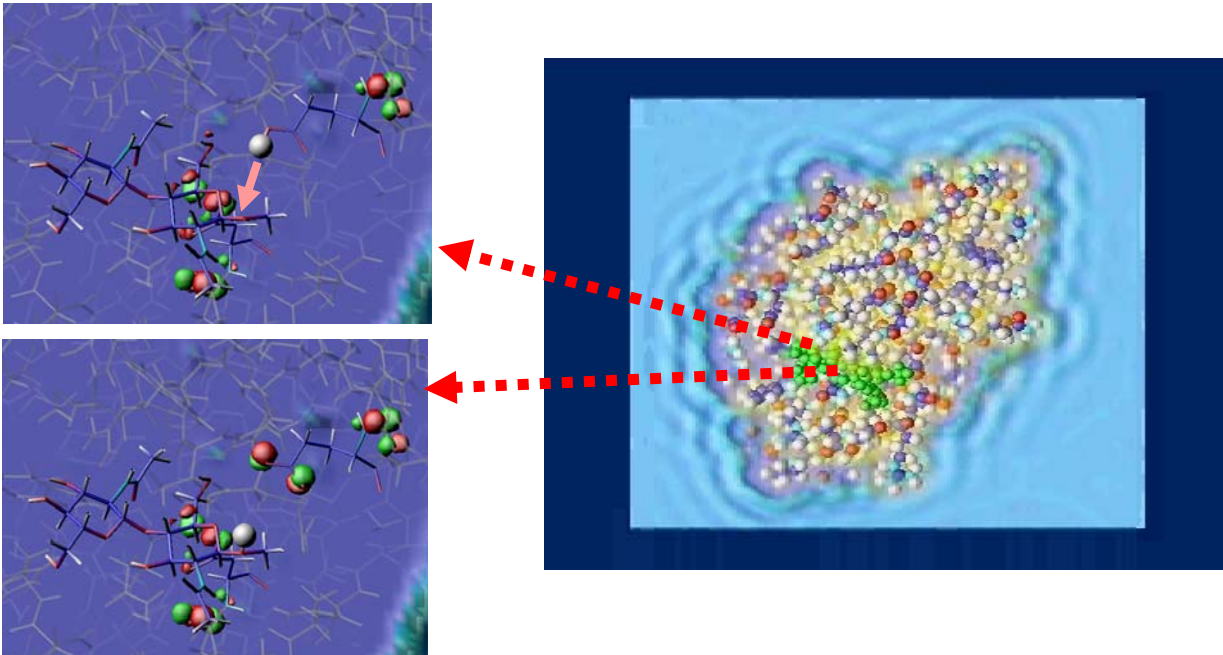
H: 
C: 
N: 
O: 



NAM: N-Acetylmuramin
NAG: N-Acetylglucosamin

Proton transfer in Lysozyme (in solution)

proton transfer from active site of Glu35 to peptidoglycan



RISM-FMO coupled simulations lead to microscopic picture of proton transfer of Lysozyme in solution National Research Grid Initiative

まとめ

連成計算支援ツールMediatorを用いたRISM-FMO連成シミュレーションを開発した。SX-7とPCクラスタ上に、RISMプロセス、FMOプロセス(メイン、マスター、ワーカー)とMediatorを配置し実測した。計算対象は、Lysozyme (129アミノ基, 1960原子, 184万溶媒点)の溶媒中全電子構造計算。

RISM-FMO連成計算の連結部である1電子積分の負荷分散アルゴリズムを考案。PCクラスタ64台の実測と分析結果から、512台の場合を予測。FMO計算は497並列(マスター数125,ワーカー数:372)時、高い加速化率295を予測。

Mediatorによりレガシープログラム間をAPIを用いた連結が可能。
異なる離散化手法間の物理データをセマンティック変換。
ローカルサイト上と広域グリッド上の連成データを統一的に連結。

マルチスケール・マルチフィジックス計算科学における次世代のシミュレーション基盤ツールのひとつを目指し、今後も連成ツールを開発してゆきたい。