

ターゲット・アプリケーションによる 性能評価について

平成18年8月
理化学研究所

ターゲット・アプリケーション選定経緯

ターゲット・アプリケーション

2010年頃にペタスケールのスーパーコンピュータを必要としている
アプリケーション・ソフトウェア。

- 1/24 開発戦略委員会(第1回)
アプリケーションの立場からアーキテクチャの検討、次世代スーパーコンピュータの利用について検討、提言をするための部会「アプリケーション検討部会」を設置。検討部会委員は、各分野の第一人者に依頼した。
- 1/25 アプリケーション検討部会(第1回)
検討方針、スケジュールを確認。
ターゲット・アプリケーション候補の提案を依頼
- 3/1 アプリケーション検討部会(第2回)
提案された候補アプリケーションの分野、優先順位付け方法を確認
- 3/30 アプリケーション検討部会(第3回)
提案アプリケーションの評価結果の取りまとめ
ターゲット・アプリケーション候補の優先順位の決定
- 4/10 開発戦略委員会(第2回)
ターゲット・アプリケーション候補の確定(候補21種 + 次候補4種)

アプリケーション検討部会委員 (平成18年4月現在)

部会長	平尾 公彦	東京大学大学院工学系研究科応用化学専攻 教授
副部会長	中村 春木	大阪大学 蛋白質研究所 教授

宇川 彰	筑波大学 計算科学研究センター センター長 数理物質科学研究科 教授
梅谷 浩之	自動車工業会 ESCAR WG 主査
加藤 千幸	東京大学生産技術研究所 計算科学技術連携研究センター センター長 教授
川本 要次	三菱重工業株式会社 技術本部 高砂研究所 次長
北浦 和夫	産業技術総合研究所 計算科学研究部門 総括研究員
木寺 詔紀	横浜市立大学大学院国際総合科学研究科 教授
阪口 秀	海洋研究開発機構 地球内部変動研究センター 地殻ダイナミクス研究グループ グループリーダー
住 明正	東京大学 気候システム研究センター 教授
泰地 真弘人	理化学研究所 ゲノム科学総合研究センター 高速分子シミュレーション研究チーム チームリーダー
高田 彰二	神戸大学理学部 助教授
寺倉 清之	北海道大学 創成科学共同研究機構 特任教授
土井 正男	東京大学大学院 工学系研究科 教授
中村 振一郎	株式会社三菱化学科学技術研究センター 基盤技術研究所 計算科学技術室長

西島 和三	持田製薬株式会社 医薬開発本部 開発企画推進部 主事
姫野 龍太郎	理化学研究所次世代スーパーコンピュータ開発実施本部 開発グループ グループディレクター
平田 文男	自然科学研究機構分子科学研究所 教授
平山 俊雄	日本原子力研究開発機構 システム計算科学センター 上級研究主席・次長
樋渡 保秋	金沢大学 名誉教授
藤井 孝蔵	宇宙航空研究開発機構情報・計算工学センター センター長
古村 孝志	東京大学地震研究所 助教授
前川 禎通	東北大学金属材料研究所 教授
牧野内昭武	理化学研究所 ものづくり情報技術統合化研究プログラム プログラムディレクター
松本 洋一郎	東京大学大学院 工学系研究科 機械工学専攻 教授
観山 正見	自然科学研究機構国立天文台 教授・台長
山形 俊男	東京大学大学院理学系研究科 地球惑星科学専攻 教授

オブザーバ

渡辺 貞	文部科学省次世代スーパーコンピュータ整備推進本部 副本部長・プロジェクトリーダー
高田 俊和	文部科学省次世代スーパーコンピュータ整備推進本部 リーダー補佐

ターゲット・アプリケーション一覧

(21種 + 次候補4種)

【ライフ分野】

	名称	BMT*略称
候補	巨大タンパク質系の第一原理分子動力学計算	ProteinDF
候補	タンパク質立体構造の予測	SimFold
候補	血流解析シミュレーション	FrontFlow/ blood
候補	オーダーメイド医療実現のための統計的有意差の検証	multilocas
候補	薬物動態の解析・予測	
候補	遺伝子発現実験データからの遺伝子ネットワークの推定	coevolv
次候補	タンパク質 薬物ドッキング計算	myPresto

【ナノ分野】

	名称	BMT*略称
候補	分子動力学計算	ParaMD
候補	FMO分子軌道法計算	GAMESS
候補	疎視化分子動力学計算	OCTA
候補	実空間第一原理分子動力学計算	RSDFT
候補	平面波展開第一原理分子動力学計算	PHASE
候補	溶液中の電子状態の統計力学的解析	RISM
次候補	低次元強相関電子系の数値繰り込み群による解析	

【物理・天文分野】

	名称	BMT*略称
候補	天体の起源を探る超大規模重力多体シミュレーション	重力多体 /SPH
候補	格子QCDシミュレーションによる素粒子・原子核研究	格子QCD

【地球科学分野】

	名称	BMT*略称
候補	地震波伝播・強振動シミュレーションモデル	FD
候補	全球雲解像大気大循環モデル	NICAM
候補	超高解像度海洋大循環モデル	COCO
次候補	全球雲渦解像結合モデル	

【工学分野】

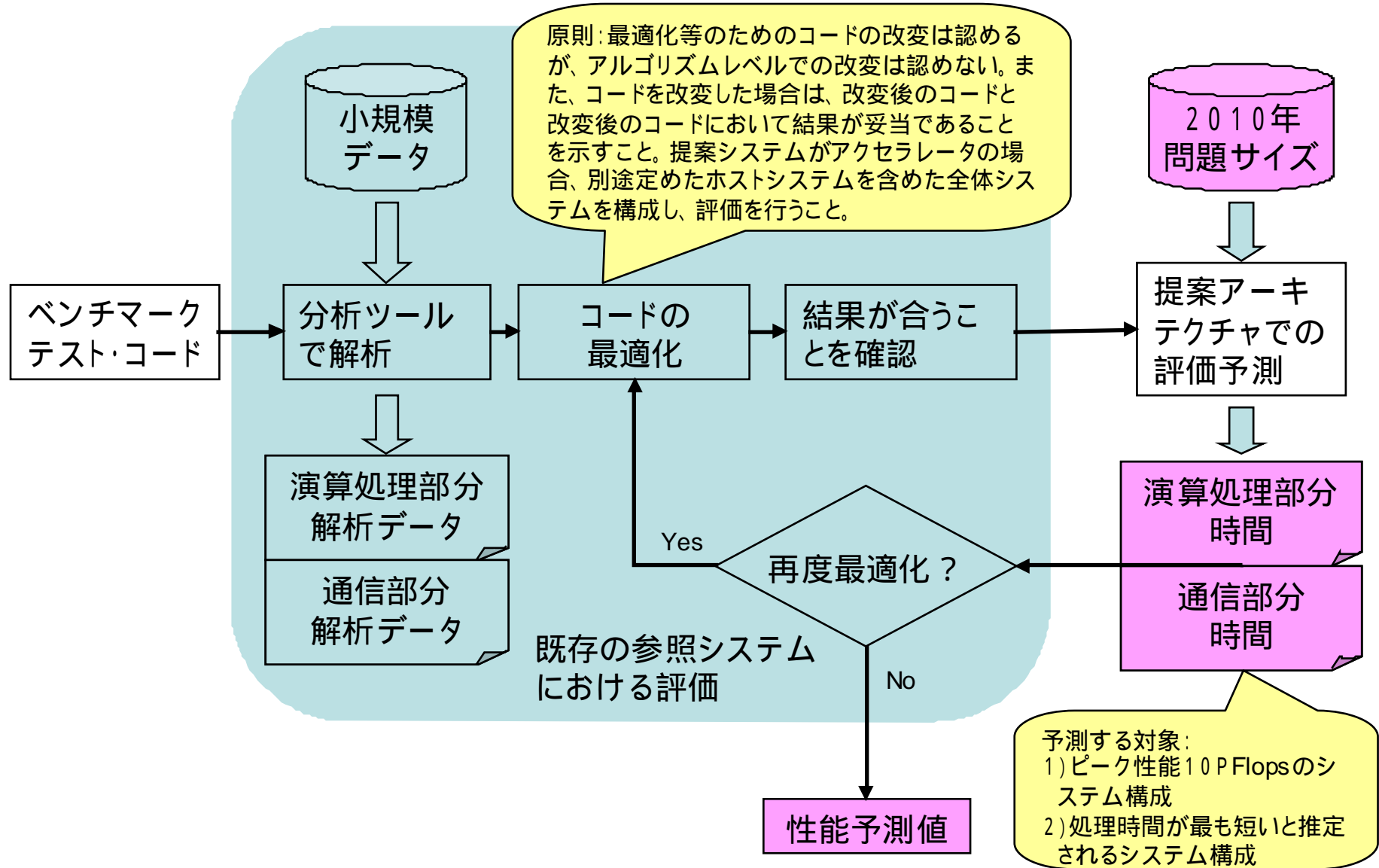
	名称	BMT*略称
候補	有限要素法による構造計算	FrontSTR
候補	有限差分法によるキャビテーション流れの非定常計算	Cavitation
候補	航空機解析における圧縮性流体計算	圧縮性流体
候補	Large Eddy Simulation (LES)に基づく非定常流体解析	FrontFlow/ blue
次候補	ボクセル手法による流体構造連成解析	

* BMT: Bench Mark Test

ベンチマークテスト・コードによる 性能評価の経緯

- 4 / 6 ~ 4 / 21 次世代スーパーコンピュータ概念構築に関する共同研究機関を募集
- 16組織17提案のうち、具体的なアーキテクチャ提案のあった7組織(東大、筑波大、九州大、国立天文台、富士通、日立、NEC)及び海洋研究開発機構と共同研究を開始
 - 理研
 - 次世代スーパーコンピュータのアーキテクチャ検討
 - ベンチマークテスト・コードの開発
 - 相手機関
 - 提案アーキテクチャでのベンチマークテスト・コードの性能予測
- 4/10 ~ ターゲット・アプリケーションから、ベンチマークテスト・コードの作成
完成したベンチマークテスト・コードを用いて、共同研究機関において性能評価を実施
- 6月末までに、4種の汎用システムアーキテクチャ案と2種のアクセラレータ案に対して、12本のベンチマークテストコードによる性能評価を行った

性能推定方法の概要



性能推定結果とその評価

- 4つのベンチマークテストにおいて、複数のアーキテクチャ案で1ペタフロップス超の達成が予測された
- 多くのベンチマークテストで、アーキテクチャ毎の十分な性能予測結果が得られなかった
 - 数万プロセッサまでスケラブルなコードが少なかった
 - 通信処理の変更はアルゴリズムレベルの変更まで考慮する必要がある
 - 問題サイズが非常に小さい場合がある
 - 演算処理の最適化の不十分な点があった
 - アーキテクチャによる最適化はできる限り行った
 - 実行時の最適化作業の質に差があった
- ベンチマークテストコードの本数を絞り、十分な時間をかけ、詳細な評価を行う

ベンチマークテストコード一覧表(1/5)

ライフ分野						
プログラム	言語	プログラムの概要	主要計算手法	略称	分割方法 (次元数など)	2010年規模での並 列度
ProteinDF:大規模タンパク質の全電子計算	C++	密度汎関数法をベースとして,タンパク質のような金属とペプチド鎖をもつ分子の計算で電子相関を取り込んだ精密な計算が可能.ab initio MD法とDFT-CI法を取り込んでリアルなタンパク質反応シミュレーションを行う.ただし,ab initio MD法とDFT-CI法は開発中.	密度汎関数法,分子動力学	ProteinDF	原子数または電子数	4728 (原子数)
アミノ酸配列情報から擬似的なフォールディングシミュレーションによってタンパク質立体構造を予測する. SimFold	Fortran77	タンパク質の相互作用エネルギー関数と構造生物学の知見を利用した独自の経験的ポテンシャル関数を利用して拡張アンサンブル法を利用して構造探索する.	拡張アンサンブル法	SimFold	逐次	200万
FrontFlow/Blood	Fortran90	基本的にFortFlow/redがコアで1次元解析を行うモジュールを追加して,微小循環系での非ニュートン性,末梢血管の圧力コントロールをフィードバックする.	CG法	FrontFlow/Blood	領域分割 (3次元)	5000万-1億 (要素数)
ゲノム情報を基に表現型(疾患や薬物反応性)に関連する座位の探索	C	ゲノム情報を基に表現型(疾患や薬物反応性)に関連する座位を探索する方法として,連鎖解析と連鎖不平衡を用いた解析がある.連鎖解析の計算は最尤法を用いたものであり,極めて項数の多い多項式.連鎖不平衡を用いた疾患遺伝子や薬物反応性遺伝子の探索では,患者群と健常者群などの二群において,統計的にどのくらい有意差(二群の差が偶然に起こる確率)があるかを判断基準とする.	最尤法	multilocas	サイクリック分割	数10万
実験データからの遺伝子ネットワークの推定	C++	S-systemと呼ばれる微分方程式モデルを用い,このモデルのパラメータを生化学実験によって計測された遺伝子発現データに合うように調整.次々に生成されるモデルパラメータに基づいて微分方程式を解くことで得られた遺伝子発現の予測量と,実際に計測された遺伝子発現量との誤差を遺伝的アルゴリズムで最適化.	常微分方程式,スプライン補間,遺伝的アルゴリズム	coevol	Master-Slave パラメータ分割	5406 (現状)
タンパク質-薬物ドッキング	Fortran90	タンパク質と低分子化合物のスクリーニング.700万化合物を網羅的に標的タンパク質へドッキングさせ,最安定とみなされるタンパク質-化合物複合体構造の予測と結合自由エネルギーの予測を行う.標的タンパク質に化合物を10-100万回ランダムに配置して,複合体構造を作成.これらからエネルギーの最適化計算を最急勾配法などで行う.	最急勾配法	myPresto	逐次	10万~100万

ベンチマークテストコード一覧表(2/5)

ナノ分野						
プログラム	言語	プログラムの概要	主要計算手法	略称	分割方法 (次元数など)	2010年規模での並列度
高並列汎用分子動力学シミュレーションソフト	Fortran90 C	任意の原子・分子集合体に対する高並列汎用分子動力学ソフトで、1000万原子系のような巨大システムや自由エネルギー計算も可能。RESPA法による運動方程式の数値解、NVE、Nose-Hoover chain、Parrinello-Rahman法によるNVT、NPTアンサンブルが生成可能。SHAKE、RATTLE、ROLL法により拘束条件付きの動力学もOK。周期境界での無限遠を見積もるため、Ewald法、PME法。巨大系にセル多極子展開法。	様々な手法を含む分子動力学計算	ParaMD	粒子分割	1000万 (粒子数)
量子化学基盤ライブラリ(GAMESS)	Fortran77 C	ナノ分野の代表的なアプリケーションにおいて、共通に高負荷となる部分・演算を複数抽出し、それに対する超並列アルゴリズムを開発して共通のライブラリを実装する。1.二電子積分計算部、2.数万次元以上の密行列対角化、3.数億次元以上の疎行列対角化、4.三次元FFT、5.密行列連立一次方程式。	二電子積分、密行列・疎行列対角化、三次元FFT、密行列連立一次方程式	GAMESS	2段階の並列化 ・二電子積分部分の分割 ・フラグメント単位の分割	10万程度
ソフトマターの統合シミュレータOCTA	C, C++	高分子系のマイクロ、メソ、マクロを扱う統合シミュレータ。SUSHI: OCTAの動的平均場シミュレータ。高分子の平均場理論に基づいて、高分子や界面活性剤の自己組織、界面構造をシミュレート可能。 Muffin: 連続体シミュレータ。	分子動力学	OCTA	逐次	
実空間第一原理ナノ物質シミュレータ	Fortran90	1)密度汎関数法:全エネルギー・電子構造計算と2)密度汎関数法: 拡張アンサンブル分子動力学法計算。1)には実空間密度汎関数理論を用いる。1),2)は結合されて実空間第一原理ナノ物質シミュレータとなる。1)は長距離相互作用計算において、FFTではなく、実空間Poisson方程式解法が可能。2)はカー・パリネロ型ダイナミックスの分子動力学と拡張アンサンブルによる統計的位相空間の結合法である。	Poisson方程式、カー・パリネロ分子動力学、拡張アンサンブル法	RSDFE	空間分割 (3次元)	10077696
擬ポテンシャルと密度汎関数法によるナノ材料第一原理分子動力学計算プログラム「PHASE」	Fortran90 C	PHASEは局在基底ではなく平面波基底を用いることで、分子から固体まで多くの物質に対して高精度な電子状態計算が可能。PHASEはFFTを多用する。	FFT	PHASE	エネルギーバンドおよび波数空間について分割	17496(バンド数)
RISM/3D- RISM	Fortran90	RISM理論および三次元RISM理論に基づき溶液内のナノ粒子(巨大分子)の構造安定性や溶媒和構造を求める。量子化学計算やモンテカルロ法などと組み合わせることにより、溶液内のタンパク質の電子状態やフォールディング問題に適用可能。本手法は統計学に基づいてその分布関数(統計的重み)を求めるところが特徴。三次元FFTを繰り返し行うため、FFTがポイント	FFT	RISM	三次元FFTの実装依存	現状512

ベンチマークテストコード一覧表(3/5)

物理・天文分野						
プログラム	言語	プログラムの概要	主要計算手法	略称	分割方法 (次元数など)	2010年規模での並 列度
天の川銀河の現在, 過去未来(第一原理に基づく宇宙論的銀河形成過程の解明)	C	<ul style="list-style-type: none"> ・中身はN-body/Smoothed Particle Hydrodynamics+GRAPE/Treeハイブリッドプログラム ・ダークマター及び星の系での重力相互作用の無衝突系粒子の計算 ・星間ガスを流体近似しSPHで計算 ・重力相互作用はTree法とGRAPE-6 	SPHと重力多体計算	重力多体/SPH	再帰直交分割	32 (32ノード以上はコード修正が必要)
星・惑星の誕生(超大規模多体シミュレーションによる惑星系の構造の起源)	C	GRAPE-6を使った重力多体シミュレーション. 粒子の起動積分は4次エルミート積分. 相互重力計算はGRAPE-6.	重力多体計算		2次元分割	1000以上
Wilson-clover作用を用いた格子QCDシミュレーション	Fortran90	各クォークの効果をハイブリッドモンテカルロ法を利用してグルオン配位を生成する. クォーク行列を係数行列とする連立一次方程式にはBiCGStab法.	CG法(BiCGStab)中の行列ベクトル積	格子QCD	領域分割(3次元)	48x48x48=11万程度
オーバーラップ作用を用いた格子QCDシミュレーション	Fortran90	グルオン配位の生成は上と同じ. オーバーラップ作用の適用に分数多項式近似を利用. クォーク行列を係数行列とする連立一次方程式の解法には現状CG法を利用.	CG法中の行列ベクトル積			

ベンチマークテストコード一覧表(4/5)

地球科学分野						
プログラム	言語	プログラムの概要	主要計算手法	略称	分割方法 (次元数など)	2010年規模での並 列度
地震波伝播・強振動シミュレーション	Fortran77	不均質な地下構造における地震波動の伝播を運動方程式(応力の釣り合いの式), 応力 - 歪みの構成方程式の2種類の方程式を陽的差分法により求める. 差分計算は時間2次, 空間16次精度.	FDM	FD	領域分割 (鉛直1次元)	2560 (各領域のZ方向サイズは偶数である必要)
「全球雲解像モデル」NICAM	Fortran90	全球を数Kmメッシュで覆う大気モデルで地球全体の雲を解像するモデル. 完全圧縮性非静力学方程式系を修正型正20面体格子を使って有限体積法で離散化. Split-explicit(時間分割陽解)法を用いて, 並列化効率も高い. まだ開発中だが, 雲物理過程の高精度化のため雲粒子を粒径方向に分割した「ビン法」を用いることを想定.	FVM, 陽解法	NICAM	領域分割 (水平2次元)	10x2 ²⁴
超高解像度海洋大循環および海洋生態系モデル	Fortran77一部90	従来からの海洋大循環モデルを高解像度化して, 潮汐などの気候モデルを陽に取り入れ, 同時に海洋-大気間の二酸化炭素交換に関する化学課程および海洋植物による二酸化炭素固定と海洋生物の食物連鎖を表現する生物過程を取り込む予定. 三次元直交座標系での偏微分方程式系を有限差分法で離散化する. 方程式系は粘性流体の運動方程式と温度・塩分の移流拡散方程式をベース. モードスプリット法を適用しているが, モードスプリットで必要とされる通信がボトルネック	FDM, モードスプリット法	COCO	領域分割 (水平2次元)	1280x960 =819200 (東西, 南北方向に指定できる各々の分割数は, それぞれのグリッド数の約数とする必要がある)

ベンチマークテストコード一覧表(5/5)

工学分野						
プログラム	言語	プログラムの概要	主要計算手法	略称	分割方法 (次元数など)	2010年規模での並 列度
主要マクロ力学系の大規模 並列構造解析ソフトウェアを 自在に組み合わせた連成解 析	Fortran90 C	連成解析ソフトウェアの構造部分。構造解析機能として、静的 (熱)弾性、動弾性、固有値、動解析、弾塑性、大変形が取り扱 える。	FEM	FrontSTR	領域分割 (3次元)	128 ³ =2097152
有限差分法での流体機械内 部キャビテーション流れの非 定常計算	Fortran77/90	一般座標系での非線形な気泡の体積運動を考慮したキャビ テーションモデルおよび乱流モデルを考慮した方程式群の時間 発展計算	FDM, 界面追跡	Cavitation	領域分割	4 (羽の枚数)
圧縮性流体のシミュレーショ ン	Fortran77/90	一般座標系の移動格子において圧縮性流体の時間発展計算	FDM, ADI	圧縮性CFD	領域分割	数万
LESに基づいた流体解析 コードFrontFlow/blue	Fortran77	LESを利用して非定常流れを精度良く解析する。非定常解析の 飛躍は流れの非定常性に起因する振動、騒音、反応、燃焼と いった現象の予測を設計に反映できる。	FEM	FrontFlow/blue	BMで使用する領 域分割用ツールで は、分割は一方向 のみ	1000