

2001 年 8 月 3 日

サイエンティフィック・システム研究会

科学技術計算分科会 第 1 回会合資料

スーパーコンピューティングを支える大規模データマネジメント

日本原子力研究所 関西研究所 光量子科学研究センター

光量子シミュレーション研究グループ

(兼) 計算科学技術推進センター

ITBL 利用推進室

情報システム管理課

上島 豊

1. はじめに

Chirped Pulse Amplification (CPA) [1]という最新のレーザ技術により、パルス長が 10 フェムト秒 (f seconds= 10^{-15} s) 程度でパワーが 100 テラワット (T watto= 10^{12} W) を越える極短パルス超高強度レーザが出現し、従来、地球上では得られなかったような極限状態が実験室環境で実現できるようになった。このレーザの短パルス性と高強度性について以下いくつかの比較を行う。まず、パルス長が 10fs という時間は、光が $3\mu\text{m}$ しか進めない程短い時間であり、室温での固体中の原子運動 (振動) 周期より短く、原子の運動を止めるフラッシュのような働きができる。また、レーザパワー 100TW という大きさは、地球上で使用されている総エネルギーの 10 倍、別の言い方をすると東京都の全地表面 (2183km^2) に降り注ぐ太陽光エネルギーに匹敵するものである。このレーザを集光して得られる光の強度は、実に太陽光の 10^{20} 倍にも達する。光の電磁界や圧力等は、普段ほとんど意識されないが、これほど強い光になるとそれらも桁違いで、もはや地球上で得られるどんなものよりもはるかに大きな値になる。例えば、レーザの磁場の強度は 0.1TGauss 以上となり、磁気閉じ込め核融合の磁石の 100 万倍以上の強度であり、遠く宇宙の神秘のパルサー等のものにも匹敵する。圧力にしても、地球の中心圧力の 1000 倍以上、つまり太陽中心圧力にも達する非常に強大なものにな

る。

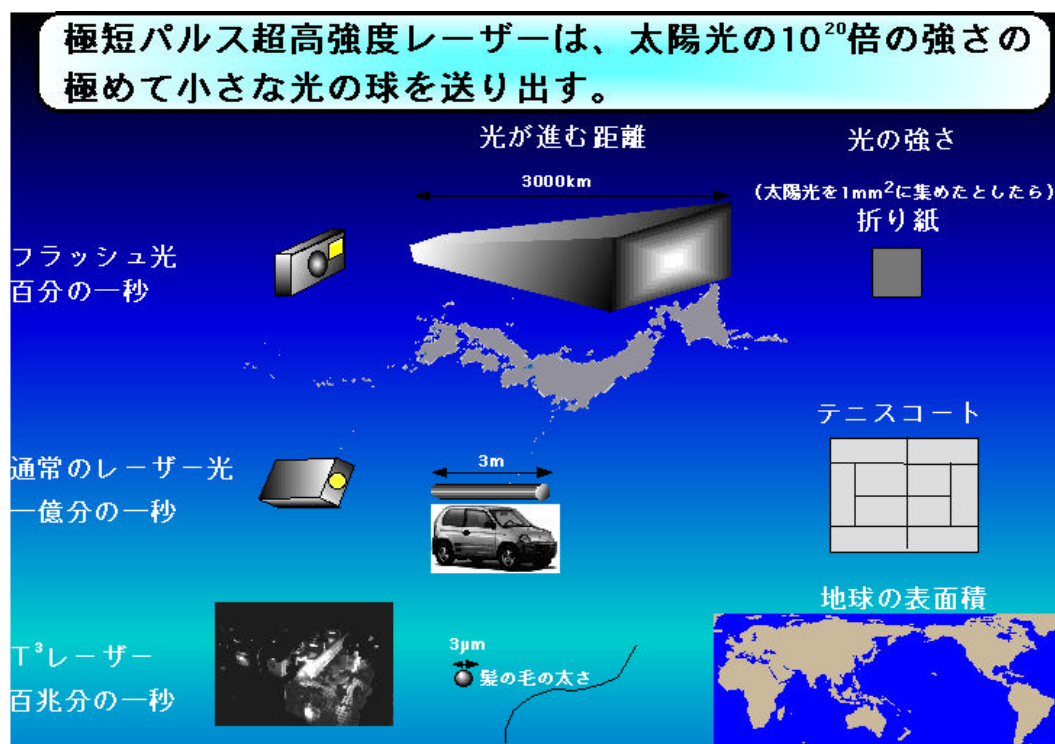


図1 極短パルス超高強度レーザーの説明図

日本原子力研究所関西研究所光量子科学研究センターでは、上記で述べたような極短パルス超高強度レーザーを物質に照射することによって、高密度集積半導体や新薬開発等の産業に大きな影響を与えるX, γ線レーザー、癌治療の旗手として高密度高エネルギーイオン加速器の開発を目標としている。さらに、照射自体によって生じる極限状況を制御し詳しく解析することで天文学、高エネルギー科学等の基礎科学の発展にも大きく寄与することが期待されている。

本報告では、上記に描画されたような極短パルス超高強度レーザーと物質との様々な相互作用を研究テーマとして取り扱う光量子科学研究と、そこで使われる大規模シミュレーションの現状と将来について論じる。

極短パルス高強度レーザー相互作用は、複雑な現象が絡み合うためシミュレーションによる総合評価が必要である。

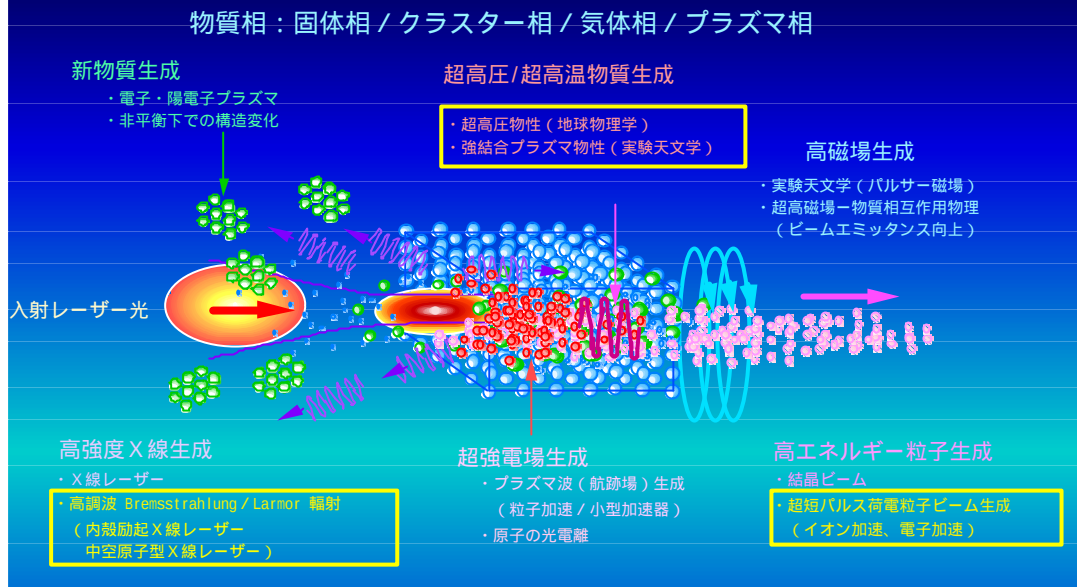


図2 光量子科学における光量子・物質相互作用の様々な物理現象の模式図

2. 光量子シミュレーション

2.1 光量子・物質相互作用

前章で述べたように、極短パルス超高強度レーザを使えば、従来、地球上では得られなかったような極限状態が実験室環境で実現できる。以下、極短パルス超高強度レーザを物質に照射するとどのようなことが起こるか簡単に説明する。

- 1) レーザが物質に到達すると、物質の表面で原子に束縛された電子が強いレーザ電場によって結合を振りほどかれ、自由電子となり強く振動する。
- 2) レーザが最大強度となる頃には、電子を剥ぎ取られた固体表面は、自身が帯電し大きな反発力を持つことになり、粒子群 (クラスター)、原子、イオンなどが様々な形で表面から飛び出す。また、レーザは、固体表面でのプラズマ振動 (電子の集団的振動) と強く相関し合い、より短波長光を含む光に変換 (変調) される。レーザ反射の前面は、レーザの光圧により生じた超高压状態にさらされることになる。反射された変調光は、自らがひきはがしたクラスターやガスプラズマと相互作用しながら反射されていく。また固体内部を伝播していく変調光は、固体内部の原子を電離させたり、電子密度 (電場) の波である航跡場を形成したりしながら進んでいく。
- 3) レーザにより前方に加速される非常に高エネルギーの電子は、固体内部の原子を電離させなが

ら、レーザを追い越し、固体を貫いて進んでいく。そしてこの加速電子に引きずられて、イオンが前方へ加速され高密度高エネルギーで指向性の高いイオンビームとなる。この時、電子やイオンの自己電流により強い磁場が発生し、自身をより収斂させる働き（ピンチ効果）を示す。

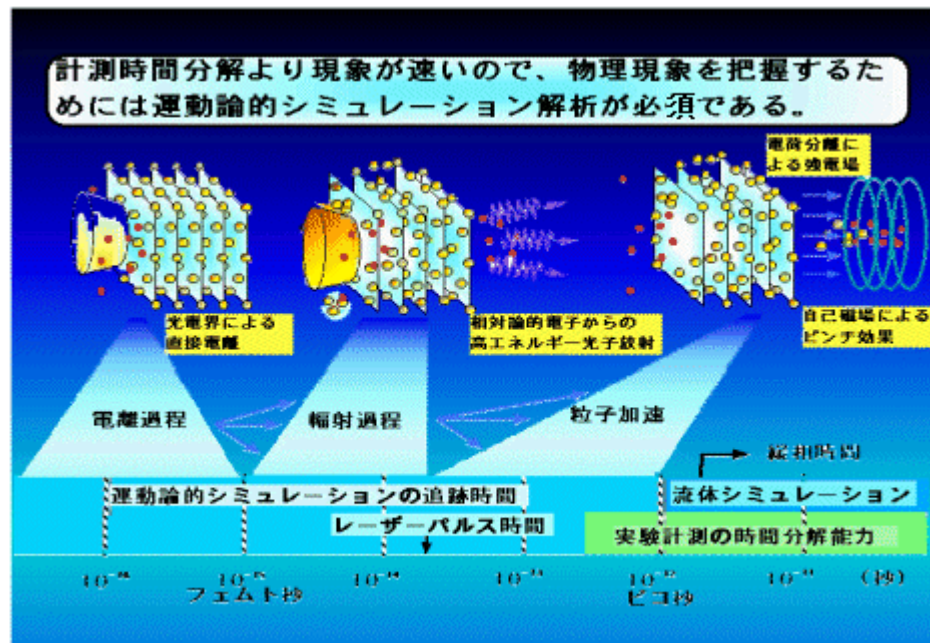


図3 極短パルス超高強度レーザ照射時の物理過程の時間発展と時間分解能

このような過程は、全て $10 \sim 100\text{fs}$ という非常に短い時間の現象であり、実験による計測が現在の技術では不可能な領域である。従って、光量子科学研究を進めていく上で第3の科学と呼ばれているシミュレーションは理論・実験と同様に不可欠な要素である。

2.2 シミュレーションモデル

極短パルス超高強度レーザを物質に照射した場合、 10fs という瞬間的な時間に多数の複雑な過程が絡み合って現象が時間発展していくことになる。従って、単純な理論ですべてを理解することは困難であり、シミュレーションを使った現象理解が不可欠となる。高強度レーザと物質の相互作用において主要なシミュレーション技法は、大きく分けて流体法（熱力学量の時空間発展: 連続体としての温度、密度、流速等）と運動論法（分布関数の時空間発展: 粒子群としての速度・空間分布関数）の2種類がある。流体法は、比較的現象が穏やかな、つまり、準熱平衡的な状況にしか適応できなく、また、レーザの伝播過程が扱えないという大きな欠点をもっている。しかし、時間空間の

大きさを自由にとれる利点を持っている。注目すべき量が長時間後の振る舞いであり、熱力学量やそれに強く関連する量の場合は、非熱的な中間過程を様々な近似を用いて適用範囲の拡張を図ることができる。光量子科学研究においても、プラズマの熱力学量の振る舞いが長時間にわたって必要な X 線レーザーの発振過程等は、流体法にレーザー吸収過程と原子過程を近似的に結合させることで評価を進めている。しかし、元来、非熱力学的である粒子加速過程やレーザーの伝播過程そのもの（粒子の分布関数）を評価する必要がある問題では、流体法は無力であるため、運動論シミュレーションを行う必要がある。

運動論シミュレーションは、ボルツマン方程式を直接差分していく方法が通例であるが、対象がイオンや電子の荷電粒子群（プラズマ）となる場合、粒子・空間格子法（PIC: Particle In Cell）法[2]という変則的な運動論シミュレーションがよく使われる。PIC 法の特徴は、ボルツマン方程式のように分布関数をあらかじめ座標・運動量空間格子に分割して直接解かず、粒子運動を運動方程式に従って解いてゆき、その結果として分布関数の時間発展が求まるところである。ボルツマン方程式の方法は、座標・運動量空間で動的に局在化するような場合、局在化と無関係に座標・運動量空間格子に分割しておく必要があり、非常に計算効率が悪い。しかし、PIC 法では、運動量空間を格子点分割しないため、分布関数が座標・運動量空間で局在化した場合でも、局在化に対して効率的に計算できる。

3. 大規模シミュレーション

前章では、光量子科学研究において問題となる物理対象と PIC 法を基盤とする運動論シミュレーションのモデルについて述べた。本章では、実際に具体的な問題を解く場合のソフトウェア・ハードウェアを含めたいくつかの条件についてさらに詳しく論じておく。まず、PIC 法のコードを使って気体中の相対論的レーザーの伝播過程や電子・イオン加速過程を評価することを考える。レーザー（光）というものは、その波の特性上、縦横厚さすべての方向において自身の波長（極短パルスレーザーでは、約 1 ミクロン）の数倍程度までしか集中させることができない。それゆえ、シミュレーションでは最低でも 10 ミクロン立方程度の大きさが必要である。また、PIC 法では、光の波の伝播を正確に計算するにはレーザーの 1 波長を 10 格子程度、さらに 1 格子中に電子・イオンはそれぞれ 10 個程度で表現しなければならない。したがって、格子数で 1000^3 、粒子数で 200 億、言い換えるとプログラムのメモリで 1 TByte 以上必要である。

このような大規模シミュレーションを実行するためには、従来の高速な中央演算処理装置（CPU）を

1 つだけ持つ形のスーパーコンピュータを越えたハードウェア=超並列計算機と、極めて困難とされている超並列計算用ソフトウェア開発が不可欠となる。日本原子力研究所関西研究所光量子科学研究センターでは、5 年前に当時世界でも有数の計算能力を誇るスカラ超並列計算機"Intel Paragon XP/S75 MP"を大阪府寝屋川市に導入し、昨年度その後継として京都府木津町の光量子科学研究センターに"Compaq ES40/227 SC シリーズ"を導入した。このシステムは、最大 720CPU (1.20TFLOPS , 1.44TByte) の超大規模計算ができるように運用されている。



図4 前システム"Intel Paragon XP/S75 MP"と現システム"Compaq ES40/227 SC シリーズ"

現在までに、前計算機システムでは計算不可能であった 2 次元大規模計算や 3 次元計算が行われ、いくつかの新しい結果が生まれつつある。演算速度に関しても、前システムの 10 倍から 20 倍の性能が調達時のベンチマークやその後のユーザからの結果から確かめられている。

ところで、大規模シミュレーションについて上記の議論では、単純にメモリの大きさと大規模化が比例していてソフトウェア的な工夫がないかのように思われがちであるので、現在までの技術の変遷とこれから将来に達成されるべき技術について以下で簡単に触れておく。

まず、5 年前は、ほとんどのシミュレーションコードは、単一 CPU のベクトル計算機向けにプログラミングされていた。無限とも思われるメモリ容量と超高速頭脳を持った計算機は、それまでの長い間メモリ確保の格闘やスカラ型の悩みであるキャッシュミスヒットの悪夢から物理屋を解放してくれた。というより、これ以前では、ほとんどの物理屋は大規模計算など考えもしなかった。ベクトル型計算機の最大の恩恵は、物理屋の数学的思考をそのまま（物理の基本法則のほとんどが単純な配列演算に帰するので）プログラムに反映しても、けっこう高速なプログラムになっていたことである。また、メモリ容量が大きくなっていったので、単純に大規模・多次元計算をしていても新しい発見が

次々見つかった時代でもある。今振り返ればこの時代は、物理屋から見ればシミュレーションに関してひとときの平和な時代であったのかも知れない。もちろん、現在、ベクトル計算機の単一のメモリやCPUはその当時からさらに大容量かつ高速にはなっているが、単価ベースでは並列計算機の伸び率から比べると雲泥の差である。

現在では、どちらにせよ超大規模計算を実施するためには、ベクトル並列型計算機かスカラ超並列計算機のどちらか、すなわち並列計算機に頼るしかない。どちらの計算機を選ぶかは、計算対象とその解法に大きく依存すると思われる。流体法による計算の場合、大規模な計算になればなる程、スカラ特有のキャッシュミスヒットが増大し、実効性能は極端に小さくなってしまっているので、現在の価格差を考えてもベクトル並列型計算機の方が優勢のように思われる。もちろん、PIC 法のプログラム(ベクトル型計算機用)であっても、その事実はかわりない。しかし、前述したように PIC 法では、空間格子の計算ではなく粒子の計算がもっとも負荷が大きいいため、スカラ計算機用に配列の宣言を変更することで、キャッシュミスヒットを効果的に減らすことができる。実際に、ベクトル型 PIC 法をスカラ型 PIC 法に変更することで 60%以上の速度向上が見られた。結果として、単一スカラ型計算機で論理性能の 55%以上の実効性能を引き出すことができることになった。また、光量子科学研究における PIC 法では、より現実的な相互作用(衝突、イオン化など)をモデリングするために Monte-Carlo 法を多用しているため、ベクトル型計算機に対してのスカラ型計算機の優位性はさらに強固なものになっている。これからの傾向についても、流体に見られるような大規模な偏微分方程式の差分が主要な演算コストである問題は、これからもベクトル並列型計算機の方が有利で、PIC や分子動力学のような常微分方程式系は、スカラ超並列計算機の方が有利であると思う。しかし、ベクトル並列計算機は、アメリカをはじめ海外ではもはや開発されていないことを考慮すると、偏微分方程式の解法を如何にしてスカラ計算機に適したアルゴリズムに変更できるかが重要なテーマであると考えられる。

最後に、並列化技術に関していくつかの指摘と今後の展望を述べてこの章を結ぶ。現在のプログラムの並列化は、もっぱら人力に頼っている状況である。確かに、HPF や OpenMP などの自動並列化言語の開発や SMP のような共有メモリ領域に対するスレッドを CPU で自動的にロードバランスさせる方法等でプログラミングの負荷を減らせる努力は続けられている。論理性能が 10 倍になって、実効性能が 3 倍で許容できるような場合(マーケティング的にはここが最大顧客であろう)は、これでよい。しかしながら、これらの方法では前段で述べたような論理性能に匹敵するような性能を達成することが困難と言うより不可能となってしまうのが現状である。「できるだけ大規模に」という大規模シミュレーションの本来の性向を考えると、マニュアルプログラミングで回避できるのに自動化で

それを台無しにするということは、許容できない問題である。したがって、大規模シミュレーションを目指す最先端シミュレーションでは、今後も常にマニュアルプログラミングによる並列化・高速化は必要であると思われる。

このように悲観的な意見ばかり述べると並列化の将来はないように聞こえるかも知れないが、実際にはもう少し楽観的に考えられる兆しもある。実際に、並列プログラミングを行うとマニュアルプログラミングと言えど科学技術用計算の場合は、大抵領域分散なのでたいした種類があるわけではない。現在のプログラムを振り返ってみても、並列=通信と考えるといくつかの種類のI/O、エラー処理によるプログラム停止、領域分割からくるノード境界量の授受、ノード境界を跨ぐ粒子の授受程度である。これらは、極めて汎用性のある問題で流体コードや分子動力学コードでも同様のプログラムを用いることができるはずである。この手の問題は、ベクトルプログラムの時にも存在した。特に、PIC法では、粒子の運動量を空間格子に割り付ける（電流の計算）時に、そのままでは参照関係が問題で、ベクトル化できなかったが、ベクトル化作業領域を設けることで、強制的にベクトル化ができるプログラミング技法が開発され、皆がそれを使っていた。並列計算でもこれと同様程度の問題しかないように思われる。たぶん、並列化計算の障壁はアルゴリズムの問題と言うより、プログラマーの思考回路の問題と思われる。従来、プログラムを設計する時には、頭の中でのしるフローチャートやパッドなどでおよその骨組みを描いてから開発を行っていたはずである。また、デバッグの時も同様であった。しかし、これに対応する並列化の場合の手続きが極めて貧弱なのである。確かに、同期、非同期やノード間の関係を記述するプログラム設計書がないのである。並列プログラムの簡単な設計書ができてその使い方に熟知すれば、ベクトル計算機のとときと同程度の閾値で、並列計算ができるように思われる。蛇足ではあるが、並列化のためのアルゴリズム開発は十分に行われており、並列計算をしている人は既に十分の恩恵を受けていることも付け加えておく。例えば、ポアソン型方程式の解法は、 N^2 オーダの通信の高速フーリエ変換から N オーダの通信の双極モーメント法へ移行し、また、PIC法で使われていたFFTは、完全局所化アルゴリズムを利用することで並列化に適さないアルゴリズムの多くは既に駆逐されている。

4. 大規模データマネージメント

光量子科学研究センターでは、レーザ加速・X線レーザ実験などの実験研究とともに、テラフロップス、テラバイト級の超並列計算機およびギガビット級のネットワーク整備を進め、物理過程解明

のためのシミュレーション研究を強化していく計画である。光量子科学研究では、実験、シミュレーション研究ともに極めて高精度の時間空間分解計測が必要なため、大量のデータが発生する。しかしながら、現在のデータ転送・解析・管理技術では、研究者に大きな負担をかけて創造的研究の遂行に大きな支障をきたしているのが現状である。前章では、大規模シミュレーションについて計算の側面から問題となるところを時代比較を交えて論じた。本章では、大規模シミュレーションによって発生するデータを如何に効率良くマネジメントし、科学者の本分である創造的研究活動につなげるかを考えてみる。

まず、大規模シミュレーションの2つのタイプの計算について説明する。一方は定型処理型研究、もう一方は発見探索型研究と以下では呼ぶことにする。定型処理型研究とは、シミュレーションを行う前からあらかじめ評価する量が決まっていってその定量性をもとに研究を進めるものである。また、その特性上シミュレーションを実施する前から、観測（計測）する物理量が明確なため、統計処理や出力変数等に大きな制限を設けることができるため比較的输出データの種類の数は少なくサイズは小さくなる傾向がある。

発見探索型研究とは、定型処理型研究とは対称的にシミュレーションを行う前からあらかじめ評価する量は未定（もちろん物理屋のセンスで適当に当りはつけるがけっこう外れるものである）で、現象の定性的な理解を促進し、新現象の発見・原因探索を目指すものである。また、その特性上シミュレーションを実施する前から、観測（計測）する物理量が明確でないため、統計処理や出力変数等に制限を設けにくく、出力データの種類の数は多く、サイズも大きくなる傾向がある。

シミュレーションを実験・理論にたいして単に手助けだけでなく、第3の科学として成立させるためには、定型処理型研究だけでなく、発見探索型研究をより推進していかなければならない。以下、大規模シミュレーションを使った発見探索型研究を行うために必要な技術・条件について経験をもとに論じる。まず、前章でも述べた大規模シミュレーションからの出力についてまとめる。大規模シミュレーションを行う場合、極限にまでメモリの効率利用を図るため、必然的にプログラム内の使用メモリは、全システムの物理メモリとほぼ同程度になる。また、全システムの物理メモリを使うような大規模シミュレーションは、一般的に現在の計算能力とメモリ容量の比（FLOPS/Byte） ~ 1 では、数日から数週間の時間が必要なためユーザによるリスタートが必要である。さらに、1TB、1TFLOPSのマシンを考えると、最大規模の計算はリスタート時、1TBのデータが出力される。NQSのクラスは精々10時間程度なので、すべてのリスタートファイルを1ヶ月分確保するためには、60TBを超える容量が必要である。100MB/sの速度でテープメディアにセーブするとしても $60000000/100=600000$ 秒=167 時間もかかることになり、再計算によるデータの再収集は現実的で

ない事がわかる。したがって、再計算をしなくてすむようにデータの出力を考える必要がある。例えば、出力データの種類には、以下のような候補が考えられる。

- 1) リスタート用：全システムの物理メモリと同程度
- 2) 動画用：サンプリングや1Byte 型データ等の利用し、1ファイルは小さいが、大量枚数の出力
- 3) 詳細解析用：浮動小数点で出力するため1ファイルが大きく、出力回数は少ない

発見探索型研究のシミュレーションではデータ量が多く、計算時間を圧迫することになるのでできるだけ高速入出力を行う必要がある。そのため、分割ファイルによる並列入出力やマシン固有の並列入出力および出力形式を使うことになり、フォーマットなどの規格化が一般的に困難となる。したがって、計算機毎にフォーマットは異なるものと考えるべきである。現在、大規模シミュレーションで出力される動画用、詳細解析用データは、8時間の NQS のクラスで50 100GByte、100 1000 枚程度にもなる。

大規模シミュレーションでは、上述したような高速計算や高速入出力が重要な要素ではあるが、それだけではもちろん発見探索型研究はおろか定型処理型研究でさえも実質的に研究活動はできない。くり返しになるが、科学者の本分である創造的研究活動と大規模シミュレーションをつなげるためにはまだ多くのハードルがある。まさか、TBを超えるような binary データを見て、科学的発見等できようはずがない(もちろん、そんな大きなデータがASCII であっても、一生かけてもすべてに目を通すことさえできないが...)。つまり、科学的認識ができるようなデータに解析加工する必要がある。「科学的認識ができるデータ」というのは、定義は難しい、というより一般的に定義しても無意味で、各々の分野で考えるべき問題である。しかし、現状で客観的にこのようなことについて科学者間で議論されることが非常に少ないこと自体が1 番の問題である。さらにそれに追い討ちをかけて、発見探索型研究では、データをファイルに保存しつつ、シミュレーションの最中(ファイルベース擬似リアルタイム)に結果のサマリーを監視し、解析をしていく必要がある。実際にこのような自動化技術が伴わない大規模シミュレーションは、非常に手間が多く、創造的研究に費やすことのできる時間が極めて少なくなる。事実、数年前まではPC にデータを取り込み、適当な統計操作等を行い1 枚1 枚手動で画像やグラフを作っていたが、絵を作り上げる時にPC のクリックという非常に非創造的な作業に時間が費やされることになっていた。また、50 100GByte、100 1000 枚程度にも及ぶデータを手動で解析、可視化を行い画像にしていくと、数週間はかかる。また、時間をかけて作られた画像も後日振り返ると、どのような実験、シミュレーションでどのような解析処理をかけたのかに分からなくなる可能性が非常に高かった。また、ノートやファイルのヘッダーにメモを残していても8 時間で100 1000 枚程度のデータが出てくると必要な時に捜し出せないか、もしくは非常に時間

がかかることになる。さらに、個人 PC ベースで行われている解析は、解析条件や解析結果の情報の非共有化がおり、複数で議論しながら研究を進める時にコンセンサスが非常にとりにくくなる。このような問題を打開するためには、超並列プログラミング以外の多くの分野の技術が必要であることは重要な事実である。

光子科学研究センターで実際に使われている大規模シミュレーションに関して、現在、どの程度の自動化が達成できているかを以下で紹介する。

- 0) シミュレーションのディレクトリ作成 (超並列計算機とファイルサーバ) やリスタート、再実行等を自動化する。
- 1) 超並列計算機のローカルディスクに作られたファイルを、ファイルができ次第ファイルサーバへ連続的に自動転送する。
- 2) ファイルサーバへファイルが転送されたら、可視化 (現在は、AVS) の実行スクリプトをそのファイルに応じて自動生成する。
- 3) 可視化実行スクリプトを実行する前に可視化用ディレクトリ作成を行い、続いて可視化ソフトウェアを実行し、可視化データを保存する。
- 4) 可視化データを、プリントアウト用、PDF 等の配信用、html 等の WWW 用にデータに自動変換し、自動再編集する。

また、現在開発中で今年度中に開発が終わる機能としては、次のようなものがある。

- 5) データベースサーバとリンクして計算結果データおよび可視化データ、可視化実行スクリプトなど各種ファイルを自動登録し、検索、WWW と連係した閲覧機能を開発する。
- 6) シミュレーションの入力データを作成を支援するシステムを WWW と CGI ベースで開発する。
- 7) シミュレーションの途中経過や JOB の実行状況、さらに上記の自動システムの実行状況を WWW から監視できるシステムを開発する。

さらに将来的に必要な技術としては、

- 8) ディスクスペースを消費し尽くさないようファイル転送やディスク空き容量を監視し、DB と連係して、効率的なデータ退避を行うシステムを開発する予定である。
- 9) 最も効率的な通信路と計算機を自動的に選ぶロードバランシング機能と、どこかの通信路や計算機が不通になった場合のリダンダント機能を追加する予定である。

これらの開発において必要とする技術は次のようなものがある。

高度なシェルスクリプト作成、AVS 等の可視化ソフトウェアの取り扱い、可視化ソフトウェアへの自己作成ツールの追加、データマイニング的な手法による効率的なデータ解析手法、各種ファイルフォ

ーマットへの対応、特にポストスクリプトの編集作業、各種 CGI スクリプトの作成、WWW を使ったインテリジェンスなインタフェース作成、オラクル等の DBMS の取り扱い、SQL を使った DB の作成、それからこれら複数の計算機を統括して制御する機能分散制御システム化技術などがある。

これらのシステム構成で考慮すべきところは次の通りである。プログラムは高速化や超並列化が最優先事項のためできるだけ最低限の制限で動作するようにシステムを構築する必要がある。つまり、外付け的（オブジェクト指向的？）にシステム化を進める必要がある。したがって、すべての制御はプログラムへ特殊なものを組み込まずに外部ファイル等からコントロールできるようにプログラムされるべきである。また、発見的・研究過程ではプログラムの開発が日進月歩であり、このようなプログラム変更に対応できるシステムでなければならない。

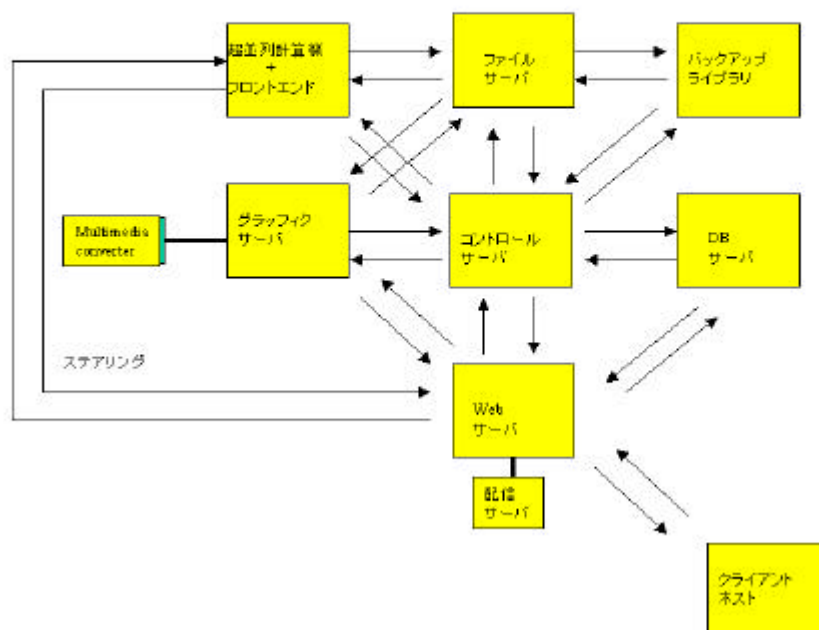


図5 機能分散制御システムの概念図

5. 広域機能分散コンピューティング

発見探索型研究に限らず多くのシミュレーションによる最先端科学研究は、現在、1 人もしくは地理的・知識的に集中している少数の専門研究者により行われている。そのため、現状では、非常

に狭い専門領域の人たちにしか理解できない構造になっている。一方、最近になって本当に興味深い発見は、多くの分野の境界・複合領域から生み出されていることも事実である。そこで、IT 技術を有効に使うことにより地理的知識的に制約されない専門家や異分野の多くの人たちが関わる新しいタイプの研究を模索していく必要がある。従って、前章に述べた機能分散制御システム化技術をさらに広域 (Internet) にまで広げ、広域機能分散制御システムを構築することが急務である。このような OPEN 系での研究では、情報共有とセキュリティー、それに関連して先取権 (著作権) 問題をうまく取り扱わなければならない。つまり、研究者の成果の独占欲からくる情報遮断をできるだけ排除し、かつ、競争意識を駆り立てる構造が必要である。また、お互い持っている情報量が異なることで誤解が広がる可能性があることは、遠隔地における共同研究をしている場合には特に注意が必要な点である。このような技術的課題が全て解決した先には、現在の研究形態とは全く異なり、物理的組織に拘束されず参加意識のある専門家や異分野の研究者ネットワークからなるプロジェクトを常時スクラップアンドビルドする形の仮想研究所的な形態へ進化する可能性がある。このような形態 (研究者が地理的知識的に制約なくつながった環境) の研究では、研究プロジェクト (仮想研究所) 数は、原理的には研究者の総数の階乗の数だけ選択肢が広がることになる。なぜなら、研究者の集合に対してすべての補集合の集合の数は、研究者の総数の階乗の数にまで大きくなり、情報処理の観点からも単なる並列処理を超えたニューラルネット的な質的な変貌が期待できる。まさに、研究活動における IT 革命が起こる可能性を秘めているのである。

6. おわりに

光量子科学研究における大規模シミュレーションの現状と将来展望についていくつかの観点から議論を行った。大規模シミュレーションを行うためには、スカラ超並列計算機に適したプログラミングを心掛けるべきである。特に現在、立ち後れてと思われることは、同期、非同期やノード間の関係を記述するプログラム設計書がないことである。並列プログラムの簡単な設計書ができてその使い方に熟知すれば、ベクトル計算機のときと同程度の閾値で、並列計算ができるように思われる。また、光量子科学研究のような発見探査的側面の強い研究では、スカラ超並列計算機から出力される大規模データのマネージメントも極めて重要であり、このシステム化の出来が物理考察における効率を決定することになる。つまり、どんなに速く計算が終わってもこのシステム化が十分でなければ、実効的に遅い計算をしたことになり、出力されたデータは人目に触れぬまま廃棄される=研究に使われないことを指摘した。最後に、本当の興味深い発見は多くの分野の境界複合領域から

生み出されている事実に注目し、IT 技術を有効に使うことにより地理的知識的に制約されない専門家や異分野の多くの人たちが関わる新しいタイプの研究を模索していく必要があることを指摘した。最後に、本研究を進めるにあたり光量子科学研究センター長をはじめとするセンター各員、計算科学技術推進センターのITBL利用推進室、情報システム管理課関西駐在の運用支援要員各位の協力に感謝いたします。

参考文献

- [1] G. A. Mourou, C. P. J. Barty, and D. Perry, Phys. Today 51, N1, 22 (1998).
- [2] T. Tajima, Computational Plasma Physics With Applications to Fusion and Astropysics, Addison-Wesley, (1989)