

5-1 電子計算機センター 現状と将来構想

5-1-1 過去と現在

分子科学研究所・電子計算機センターは1978年に設立され、今日に至っている。2000年4月より、岡崎国立共同研究機構・計算科学研究センターに改組される。電子計算機センターは日本全国の分子科学研究者に大規模計算を実行する環境を提供する計算機センターとして設立され、22年を経た今日においても所内外の分子科学研究の基盤施設としての重要性は変わらない。実際、「分子研レポート 94」に報告されている通り、外部評価委員、運営委員、所内外の利用者の多くは、本センターが分子科学理論計算分野へ貢献してきた歴史的経緯を高く評価しており、当初の目標を高い水準で達成できていることを認めている。

この22年間にセンターの計算機の性能・記憶容量は大幅に増強され、インターネットの普及により利用形態も大きく様変わりしている。図1はセンターに導入された計算機のCPU能力(理論ピーク性能)を年度別に示したグラフである。分子科学計算に必要な計算

処理の需要は年々増加の一途をたどっている。1995年以降、従来からの計算機借料を汎用計算機借料とスーパーコンピュータ借料に分割した。この分割によって、より性能の高い新型機種を導入できるチャンスと、計算機アーキテクチャの多様化に迅速に適應できる計算機構成となった。2000年3月末から稼動を開始するスーパーコンピュータ(富士通VPP5000とSGI Origin2800)、そして汎用高速演算システム(日本電気製SX-5とIBM製SP2)の総合性能は、1979年1月に初めて導入された日立製作所製M-180(2台)の実に13500倍に至っている。また図2に示すように、性能の大幅な増強に加えて、記憶容量は主記憶メモリで約70000倍、ディスク容量で約1300倍に増加している。この間、計算機アーキテクチャも、単一CPUのスカラ型からベクトル型へ、その後、複数CPUを有するスカラ及びベクトル

並列型の混成システムとなった。また複数CPUにおける主記憶メモリへのアクセス機構も分散型、共有型、分散共有

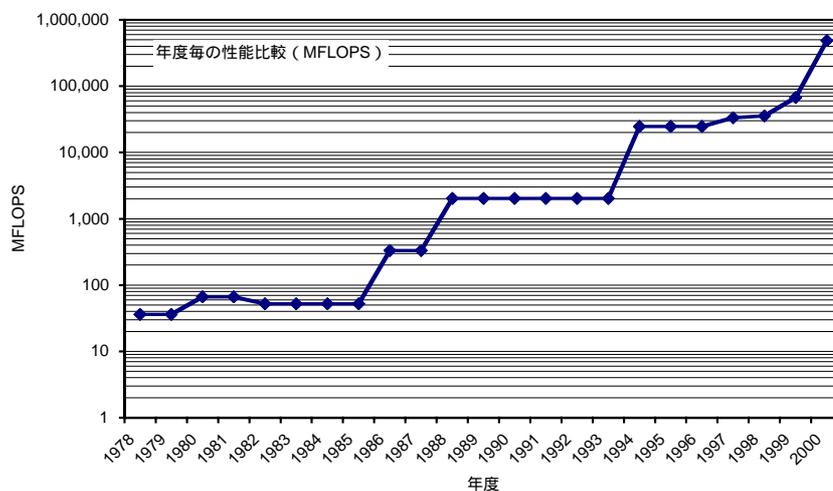


図1 電子計算機センターにおけるCPU能力の増強

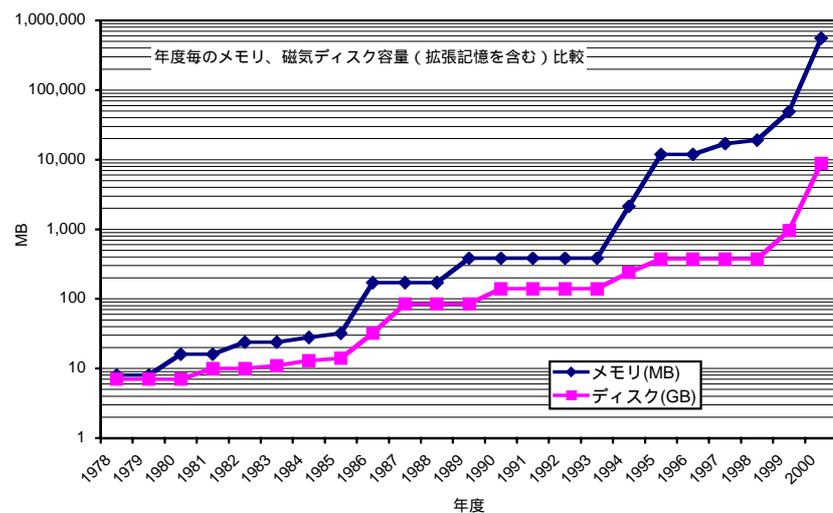


図2 電子計算機センターにおける記憶容量の増強

型等へと高度化及び複雑化している。

分子科学計算の計算機利用分野においても多様化が進んだ。設立以来、初めてベクトル型スーパーコンピュータが導入された1986年以前までは、全CPU資源の60～70%程度が電子状態計算分野に利用されてきた。1986年、日立製作所製 S-810 ベクトル型計算機の導入に伴い、分子動力学、量子反応動力学などの動力学分野での計算機利用が増え、その後ベクトル型計算機としては、同 S-820、日本電気製 SX-3、SX-5 へと更新され、今日においてもこの傾向は続いている。今後の計算機利用においては、ベクトル並列計算方式や分散共有メモリ機能を最大限に活用できるような、新たなソフトウェアが開発され、例えば電子状態計算と動力学計算を融合したアルゴリズム等、媒質との相互作用をも精密に取りこんだ電子構造計算や生体高分子などへの新分野での応用が期待される。

現在の計算機システムは、スーパーコンピュータ NEC SX-3 (3CPU)、並列計算機 IBM 製 SP2 (48CPU) が稼動しており、昨年度の汎用高速演算システムの更新により日本電気製 HSP (2CPU) が同 SX-5 (4CPU, 32GB) に更新された。また CPU 能力を増強するために施設運営費によって日本電気製 HPC (2CPU)、日立製作所製 SR2201 (16CPU) が導入されている。前述の通り、まもなく SX-3 は富士通 VPP5000 (30CPU) と SGI Origin2800 (256CPU) から成るシステムに更新される。過去6年間に導入され利用されてきた計算機が、実際にジョブ処理のために稼動してきた時間を年度毎の1CPU当たりの月平均ジョブ処理時間として図3に示した。本センターでは、年間を通じて定期保守

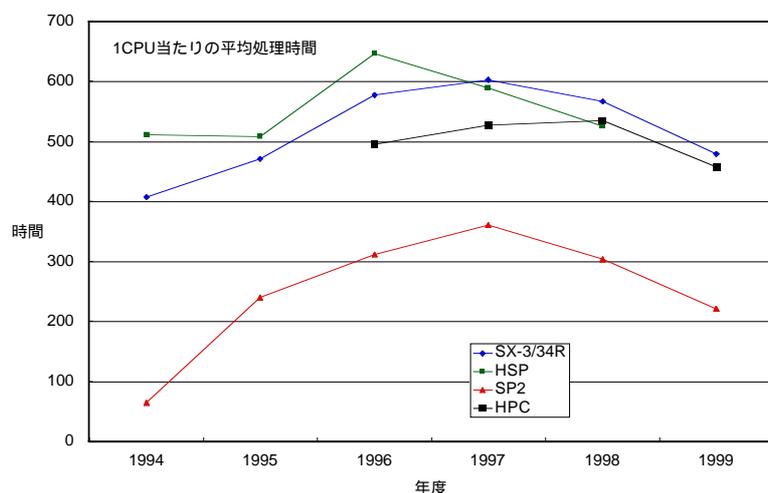


図3 過去6年間のCPU利用状況

及び障害対応時間以外1日24時間体制の運用を行っており、1カ月の通電時間は約700時間であることから500時間(稼働率70%)を越える運用は、利用者から見た場合には、待ちジョブが並ぶ高負荷状態である。48CPUから構成されるSP2は約半数のCPUが並列ジョブクラスに割り当てられているため1CPU当たりで平均すると、他の逐次処理を主体としたベクトル計算機に比べ、ジョブ処理時間は高くない。

それぞれの計算機は導入以降、徐々に利用率が増加し2・3年後にピークを迎えている。

SX-3、HSPについては利用率が80%以上にも達し飽和状態を経験している。導入当初は個々の計算機に適したソフトウェアへの移植作業を行うなどジョブ処理以外での利用も含まれ、ソフトウェアの完成度が高まり、利用者が新しい計算機に慣れるに従い、より大規模な計算処理を要求することがグラフから覗える。逐次処理コードからの並列化が比較的困難な分散メモリ型計算機(SP2)では、ソフトウェアの移植(並列化)が行われたとみられる時期(1997年度)での利用率の上昇は著しい。導入から2・3年後に利用のピークに達した計算機は、さらに大規模な計算処理を多数要求されるが、計算機資源に限りがあるため常時待ちジョブが並ぶ状態が長期にわたると利用率は徐々に低下してゆく。スーパーコンピュータに比べてワークステーション等の小型計算機の開発期間が著しく短いため、最新鋭の計算機が利用者の研究室に導入されることも、グラフが右肩下がりとなる要因の一つと考えられる。従ってセンターでは、利用のピークを迎えた時期から2・3年以内に次期の更新によるCPU能力の増強を行い、研究室では実行が困難な(または不可能な)計算処理の要求に迅速に対応してゆくことが必要である。

1997年度、文部省に概算要求をしていたスーパーコンピュータ借料の増額によるCPU能力の増強は認められなかった。また導入一時経費もつかないという厳しい結果となったため、外部委員を含む「スーパーコンピュータ検討小委員会」の結論に従い、1999年1月に機種を更新することを断念し、スーパーコンピュータの更新を1年延期することとした。これを補うべく汎用システムの機種更新を1年早め、1999年3月には最新鋭のベクトル並列計算機SX-5が導入された。このように、汎用高速演算システムとスーパーコンピュータシステムは、それぞれ中規模ベクトル演算と大規模スカラ並列演算及びベクトル並列演算のように相補的なシステム構成としての位置付けだけでなく、計算機の更新時期を柔軟に調整する観点からも2システムから成る計算機構成は重要である。

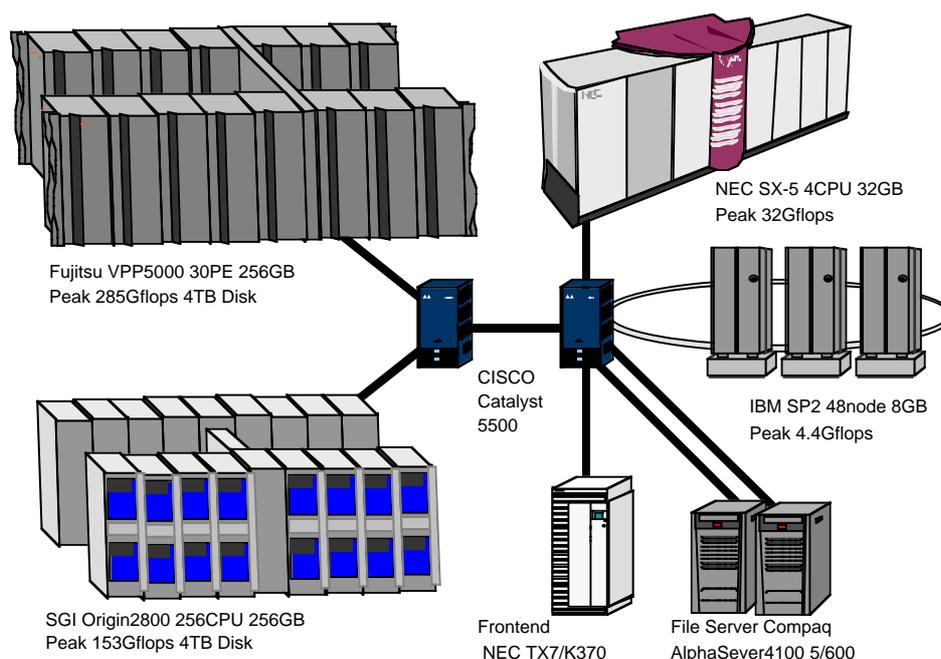


図4 2000年4月以降の計算機構成

2000年4月以降の計算機構成を図4に示す。今回導入が決まったスーパーコンピュータシステムを図の左側に、昨年度に更新を終えた汎用高速演算システムを右側に示す。新スーパーコンピュータシステムは、富士通製 VPP5000 と SGI 製 Origin2800 から構成される。VPP5000 は1CPU当たりの最高演算性能が 9.5 Gflops のベクトル演算装置 30 台から構成され、各CPUに8~16GBの主記憶装置を持つベクトル並列計算機である。一方、SGI Origin2800 は1CPU当たりの最高演算性能が 0.59 Gflops のスカラ演算装置256CPUから構成され、CPU当たり1GBの主記憶をそれぞれのCPUから共有メモリとしてアクセスが可能な分散共有方式の超並列計算機である。VPP5000では高速なベクトル演算能力を活かした大型ジョブの逐次演算処理はもちろん、例えば8台以上のベクトル演算装置を使った大規模なベクトル並列演算が可能となる。Origin2800 は Non Uniform Memory Access (NUMA) 方式と呼ばれる論理的な共有メモリ機構を有する。NUMAは主記憶装置が各CPUに分散して配置されているためCPUから主記憶へのアクセス速度が非等価ではあるが、利用者プログラムから大容量のメモリを容易に利用することが出来るため、大規模な並列ジョブの実行が可能となる。昨年度、導入された SX-5 は1CPU当たり 8Gflops の最高演算能力を持つ共有メモリ型ベクトル計算機であり、SP2は分散メモリ型スカラ並列計算機である。これらの計算機の特徴を活かしつつ、利用者ジョブの効率的な実行環境を構築することがこれからの課題である。

5-1-2 岡崎共同研究機構・計算科学研究センターと分子科学研究所・計算機室

2000年4月には、分子科学研究所電子計算機センターは岡崎国立共同研究機構・計算科学研究センターに改組転換される予定である。スタッフは、教授1、助教授1、助手2から構成されることになっている。しかし、定員増があったわけではなく、分子研から助教授1と助手1が、基礎生物学研究所（基生研）と生理学研究所（生理研）から各助手1の定員が計算科学研究センターに振り替えられることになっている。

改組以前（現在）の岡崎機構内で電子計算機・情報ネットワーク関連のスタッフとしては、分子科学研究所電子計算機センターには、理論系教授が兼任しているセンター長、助教授1、助手2、技官5（うち一人は所内情報ネットワーク専任）があり、また基生研には、培養育成研究施設・電子計算機室に助手1、生理研には、脳機能計測センター・生体情報処理室に助手1がいる。さらに基生研と生理研には計算機・ネットワーク関連の仕事をしている技官がそれぞれ若干名いる。これに加えて数年前より岡崎機構全体の情報ネットワークの企画・管理・運営などを総括する助手が1名採用されて、岡崎機構「情報ネットワーク管理室」に配属されている（組織上はこの助手の所属は分子研計算機センターに属している）。

4月の改組転換以降も、これまで分子科学研究所電子計算機センターが共同利用研究機関の業務としてきた利用者サービスには変更はない。現行の分子科学研究所電子計算機センター規則には、「第1条 岡崎国立共同研究機構分子科学研究所電子計算機センター（以下「センター」という）は、センターの大型電子計算機システムを分子科学の大型計算等のために分子科学研究所内外の研究者の利用に供するとともに、これに必要な研究開発を行い、かつ、岡崎国立共同研究機構に置かれる研究所の研究に関する計算を処理することを目的とする。」となっている。すなわち、これまでも基生研・生理研の研究に関連する計算を処理することにはなっている。そのために、運営委員には各研究所から教授あるいは助教授が参加している。2000年度の計算機利用申請は、従来通りこの3月に開かれる「運営委員会」で審議されるが、今後、改組転換に伴い新たな申請を募集する場合は、計算科学研究センター運営委員会で新たに議論する必要がある。

分子研の電子計算機センターでは、助手1と技官5が「岡機構化」されないが、彼らの働きなしには新計算科学研究センターは全く動きが取れない。したがって、分子研内の措置として「計算機室」をあらたに設置して、この計算機室と計算科学研究センターとの密接な共同作業で、新研究センターを運営していく必要がある。しかしながら、運営諸経費等、日常業務の上ですぐさま問題になることが生じよう。また、従来、分子研所内利用者が負担していた使用料の取り扱いも問題になろう。

計算科学研究センターへの改組は 統合バイオセンターが3研究所の共同事業として岡崎機構に作られることに伴って進められることになった。他に生理学研究所・動物実験施設が動物実験研究センターに、基礎生物学研究所・アイソトープ実験施設がアイソトープ実験センターに「岡機構化」される。MIT press は、新しく「*Computational Molecular Biology Series*」を発刊するにあたり“Recent developments in Molecular Biology are not evolutionary but revolutionary. Computer Science, Statistics and Mathematics are the driving forces transforming molecular biology from an informational science to a computational science. Computational Molecular Biology is a new discipline, bringing together computational, experimental, and technological methods, that is energizing and dramatically accelerating the discovery of new technologies and tools for molecular biology.”と述べている。「分子生物学を情報科学から計算科学へ転換する」というのは少々言い過ぎとしても、生物学における計算機利用が情報科学的側面からより広い計算科学的利用へと発展しつつあることは事実である。

岡崎国立共同研究機構・計算科学研究センターは、計算分子科学を中心に据えて、計算生物科学をも包含し、世界的に特異な展開を実現しなければならない。そのためには、もちろん優れた研究者の結集が第一の条件であり、同時に、さら

にいつそう高性能な計算機システムの導入も不可欠である。また、国際協力の推進が新センターの発展のために特に重要である。

ネットワーク管理室との関係

岡崎機構・「情報ネットワーク管理室」は引き続き、岡崎機構内の情報ネットワークの維持・管理・企画に責任を持たなければならない。この「管理室」を通じて、数年間3研究所と管理局のネットワーク関係スタッフは共同作業を積み重ねてきた。この作業の中で、各研究所・管理局で責任を持つ部分と、岡崎機構「管理室」が責任を持つ部分とが整理されてきた。各研究所・管理局内の情報ネットワークがそれぞれの責任で運営されている現状の体制を維持し続けなければならない。もちろん、計算科学研究センターは情報ネットワーク管理室と緊密な連携を持ち続けなければならない。

QCLDBの事業化

センターのスタッフが過去に積極的に参加し、量子化学者のみならず広く化学・物理学研究者から高い評価を受けているQCLDB (Quantum Chemistry Literature Data Base) の開発に対し今後も予算的にはもちろんのこと、センターの業務として支援する体制をとり続ける必要がある。現在、米国のいくつかの国立研究機関では、WWW (World Wide Web) を通じてその機関が作成したデータベースを全世界に公開している。このような形の全世界の学会に対する寄与は、特に生物学や素粒子・原子核の分野などではその研究機関の一つの「業績」として高く評価されている。

我が国ではデータベース作成による世界の学会への寄与は、QCLDBを除いて皆無であると言われているが、平成9年夏からはQCLDBも、WWWを通じて登録制の公開を試験的に開始した。この公開に対する世界中の研究者からの反響は著しく高い。この公開をハード・ソフトの両面で長期的に安定運用をするためには、正式に分子科学研究所の事業の一つとする必要がある。1) データベース作成・管理・運用のためのハードウェア整備、2) データベース作成のための謝金と事務費の確保のために早急に予算的措置をとる必要があり、平成12年度概算要求を行う予定である。もちろんQCLDBを開発・作成しているQCDB (Quantum Chemistry Data Base) 研究会と緊密に連携をとってこの事業は進めなければならない。

5-1-3 将来構想検討会議

2000年1月25日に「分子科学研究所電子計算機センター将来構想検討会議」を開催した。所外からは、相田美砂子(広島大)教授、大峰巖(名古屋大)教授、片岡洋右(法政大)教授、柏木浩(九州工大)教授、斎藤晋(東工大)教授、斎藤稔(弘前大)教授、榊茂好(熊本大)教授、橋本健朗(都立大)助教授、中島徹(東大)助手、中野雅由(大阪大)助手、中村恒夫(京大)大学院生、三浦伸一(東工大)助手、山西正人(東大)大学院生の13名が参加した。研究所内からは茅所長を始め、平田文男教授、中村宏樹教授、小杉信博教授、岡本祐幸助教授、谷村吉隆助教授、神坂英幸(総研大)大学院生、西川武志博士研究員および、電子計算機センターの岩田末廣センター長、青柳睦助教授、南部伸孝助手、高見利也助手と技術スタッフ(西本、水谷、南野、手島)が討論に参加した。4時間以上に渉って様々な問題が議論された。その中で主に議論された問題点を以下に整理する。

a) 電子計算機センターはどのような利用者への計算機サービスに力点をおくべきか

センターの設立当初から原子分子の電子構造の理論計算分野を中心に、センターでなければ実行が困難な計算機の利用が盛んに行われ、国内で唯一の分子科学計算センターとしての位置付を確立した。近年では、計算分野も拡大さ

れ、また研究室に比較的安価で高性能のワークステーションが導入され、センターの利用のされ方及びジョブの内容にも変化がみられる。そこで、今後電子計算機センターはどのような利用者への計算機サービスに力点を置いて運用を行うべきか、について委員の方々に議論して頂いた。

- ・ 設立当初、運用が上手くいっていたのは、利用目的がかなり限定されており、ユーザもその利用分野の方が大多数であったからであろう。現在は利用形態、利用分野ともに非常に複雑に、広範になっている。全部をサポートしようとするのは無理ではないか。この様な討論会を数回開いて、方向性を明確に決定して行くのはどうか。
- ・ 大学の大型計算機センターでは、以前は課金係数が高いこと、計算機資源の制限（ディスク、メモリ制限）もあり、大規模ジョブの実行が困難であった。現在では課金も安くなってきたので大型ユーザが研究費を出して大型計算機センターを使う傾向が見られる。
- ・ ユーザがどういう計算機の使い方をしているのかを把握しないと、環境整備はできない。外部ユーザがどの程度の環境を持っているかも把握する必要がある。それをふまえて、棲み分け等を考える必要がある。
- ・ 安いマシンの性能が急速に上がっている。コストパフォーマンスで研究室の計算機環境と比較したらセンターの存在意義は低下してしまうのではないかと。最初の頃のセンターでは新しい方法論を開発していた。現在は計算機アーキテクチャや、ソフトウェアが複雑化し、また応用プログラムも肥大化しているので、本格的な開発は難しいが、センターは新しい方法論の確立をめざす場所でもあって欲しい。
- ・ センターは利用目的と規模において3種類のユーザを持っている。1つは理論研究分野のトップレベルの研究を行っている（大規模計算も行う）研究グループ。1つはある程度のワークステーションを持っている理論研究グループ。もう1つは大規模な計算を行う実験のグループ。それぞれ無視できないのではないかと。
- ・ 所内（理論研究系）が使いやすい計算環境をめざすのも1つの方向。
- ・ 計算機が1台あれば研究が遂行できるグループは、各大学の大型計算機センターで吸収できるのではないだろうか。昔ほど計算環境が無いという緊迫感はない。
- ・ 計算機運用の面で特徴がなければ、大規模計算ユーザはセンターから離れていく。独自性、方向性を明確に出さないとユーザがついてこなくなるので、計算機更新時の借料の増額等、施設設備の拡充も困難となる。
- ・ 議論の観点が、古い分子科学ユーザに特化しすぎているようだ。常に新しいユーザ、新しい分野のことも考えることが必要である。特定分野のために計算機資源を用意するだけでなく分野の拡大も検討する必要があるのではないかと。
- ・ 新規利用者や比較的小規模な研究グループと大規模計算グループのどちらかに決めずに、両方の計算機利用環境を考慮した運用が必要。リプレース直後のリソースが余っている時期だからこそ、大規模計算向けの大型プロジェクトに計算資源を利用する、という試みをして良い。ただしスケジュールが難しい。

b) 新システムをどのように運用するか

4月より新スーパーコンピュータが運用を開始するに先立ち、VPP5000(VPP)30ノードとSGI Origin2800 256 CPUを

どのように運用するか、並列ジョブのジョブクラス等について議論した。

- ・ 並列計算ジョブに利用できるプロセッサ数は多い方が良い。VPPなら8プロセッサ以上。将来はもっと多数の並列が必要であろう。多数プロセッサによる効率的な応用ソフトウェアの開発には、それだけの環境が必要。センターとしては多数の環境を用意していただけるとありがたい。
- ・ VPP ならば16以上、SGI ならば64以上のクラスを作っていたきたい。
- ・ 名古屋大学大型計算機センターでは32CPU及び16CPUの並列ジョブが常に走っている。大規模計算利用者にはメモリ、ディスクが足りない。ユーザ持ち込みのディスクを設置できるようなフレキシブルな運営が必要。
- ・ 計算資源を細切れにするのは、大規模計算を行う環境として適さない。VPPを8プロセッサ程度で分割運用するのでは意味がない。
- ・ 京大では40PEまで普通に使える。SGIも多数を一度に用いて初めて意味がある。

またセンター側から「少数の研究グループに最大リソースを提供する様な運用をしてもよいかどうかを判断しかねている。思い切った運用についてどう思うか？」との問いかけに対し以下の意見を頂いた。

- ・ 4月以降の計算機リソースは確実に増加する。3月の運営委員会時点での割り振りでリソースが多少余るのではないだろうか。その余剰分で思い切った運用するのは可能。運用当初、テスト的にやるには好機である。
- ・ 京大でVPPを使っているが、ベクトルの効かないものはワークステーションかまたはパソコン並の性能である。大きなメモリを使いたいただければ、VPPを使わせるよりもパソコンクラスタを用意すべき。
- ・ 土日が大規模計算のみに占有されるようなのはきつい。月に1度くらい試行されるのならば問題ないのでは。年に4、5回が限度ではないか。
- ・ できるだけ大きく使いたい。SGIならば128までは並列で使えるようにしてもらいたい。VPPも並列で大きくとれるとうれしい。常に監視して適切なルールを適応して運用してほしい。
- ・ 少数の研究グループに最大リソースを提供する様な運用を行う場合は、審査の仕方、審査委員の選定も2重化していく必要がある。従来の審査方法とは違った方法が必要である。

c) バイオサイエンス分野の研究課題について

2000年4月には、分子科学研究所・電子計算機センターは岡崎共同研究機構・計算科学研究センターに改組される予定である。また機構の共通センターとして統合バイオサイエンスセンターが設置される予定であり、バイオサイエンス分野の計算機利用について議論した。

- ・ 分子研がバイオサイエンス含め生体系などの複雑系に進出していくのは基本的に良いことだと思う。機構化によって、分子研が主導を担う形でバイオ分野での計算機利用が促進されることが理想だと思っている。当面の移行にあたっては、センターのリソース不足等、危惧すべき点は少ないと思う。
- ・ バイオロジの分野でもゲノムという情報処理的な用途だけでなく、次のレベルに発展しようとしてきている様だ。現時点のリソースでは、そこまでは難しいので、当面は分子科学分野に限った利用に限定せざるを得ない。近い将来、次のステップではバイオ関連の計算機利用を念頭においたシステムに変わっていくことになるだろう。

- ・ 分子科学はこのままではいけない。分子科学がどういう分野に成長していくかが重要で、バイオ分野に成長していくことは重要である。それに連れて、計算機利用分野も変わってゆく。
- ・ 生物系であれ、生体系であれ、生命系であれ蛋白質1つを扱っている限りは生命を扱っているとは思えない、と言うのは言い過ぎであろうか。
- ・ 「生体分野の信頼のおける計算方法を開発する」ことを目標にしてはいかかが？今のバイオロジーの研究者の利用しているソフトウェアをセンターに準備してサービスを行うだけでは研究開発とは言えない。
- ・ ゲノム解析ではマルチプルアライメントをする人、構造予測する人、NMRの予測をする人などがいて、情報処理を含め色々な計算機利用を行っているようだ。このような情報处理的な利用として、バイオ関係のサービスを始めると、単にサービスをするだけのセンターになるのではないか？

d) 新しい組織及び、国際化について議論した。

- ・ 電子計算機センターにソフトウェア開発部門を作ったらどうだろうか。アメリカだと、センターには10人くらいのドクターがいてソフトウェア開発をやっている。そういう環境でなければ応用ソフトウェアが生まれてこない。オブジェクト指向のプログラミングは並列化に適している。そういうものが普及しないのはセンター内に開発部門が無いからでは。
- ・ 大学のセンターには人は結構いるけど目標が定まらない。センターが目標を掲げることが必要。

現在のセンターの運用方針では、外国人でも国内の研究グループに所属していれば利用できるが、国際研究集会などで、外国人主導の利用の要請がある。すなわち、国内に研究グループを持たないが、インターネットを通じて国外からセンターの計算機資源を利用するといった要請である。この点について委員から以下の意見を頂いた。

- ・ 現在の計算リソースでは無理だが、将来的には外国人の客員部門なども作ってはどうか。計算機センターは計算機のおもり役ではなく、サイエンスの牽引役として進展すべきである。
- ・ 方法論の国際シンポジウムを開く、というのも国際化の選択肢にあげられると思う。

今回の議論を参考にして、3月の電子計算機センター運営委員会でも議論を重ね、4月からの運用を考えていく必要がある。今後もこのような会議を開いて、広い世代の研究者の意見を採り入れ、センターの方向性を考える機会を増やし運用に反映させていかなければならない。