

## CP-PACS プロジェクト終了にあたって思うこと

吉江 友照

筑波大学物理学系／計算物理学研究センター(当時)

筑波大学数理物質科学研究科／計算科学研究センター(現在)

私は計算素粒子物理学の研究者の一人として、CP-PACS プロジェクトに参加させて頂いた。CP-PACS の記念冊子への寄稿をしたためようと、十余年にわたる CP-PACS プロジェクトを振り返ると、実に色々な形でこのプロジェクトに関わってきたものだと、感慨深いものがある。CP-PACS の開発初期段階では、計算素粒子物理学のアプリケーションである QCD シミュレーションプログラムを分析しベンチマークプログラムを纏めた。設計段階では、計算機工学者やメーカーである日立製作所の提示するアーキテクチャに基づいて演算・通信・ディスク 10 の各側面での机上性能評価を行い、CP-PACS のアーキテクチャの決定に寄与した。CP-PACS 完成後は、プログラムの開発チューニングに携わり、QCD シミュレーションのマイルストーンとも呼ぶべき『クエンチ近似のハドロンスペクトラム計算』を自ら実行した。また、CP-PACS 稼働期間中は一ユーザーとしてのみならず、計算機システム運用委員の一人として、主として素粒子物理ユーザーと計算機システム及びメーカーとの橋渡役も行なわせて頂いた。実働部隊の一人として、プロジェクトの多くの側面を勉強させて頂き、微力ながらも各側面でプロジェクトに寄与できたのは、望外の幸せであった。

私にとっての CP-PACS プロジェクトのハイライトは、上述の『クエンチ近似のハドロンスペクトラム計算』である。その素粒子物理学に於ける意義や成果の詳細は、本冊子の『CP-PACS の成果、計算素粒子物理学』の章に纏められているので、ここでは、スペクトラム計算と計算機としての CP-PACS との関わりについて、既に種々の報告書にまとめた事柄も含むが、本記念冊子の本文で詳細に述べる事が出来なかった内容を、雑記風にまとめておきたい。

スペクトラム計算の主要なサブルーチンである QCDMULT や QCDMR

の演算の特徴は、ひとことで言えば、『ベクトル計算機向き』である。つまり、ロードストアの割合が加算乗算演算のそれに匹敵する。(コーディング手法にも依存するが、QCDの場合、概ね、加算:乗算:メモリアクセスは、1:1:1である。)キャッシュに頼らず高いメモリスループットを確保し、メモリアクセスレイテンシを隠蔽する為に、CP-PACSでは、バンクメモリとスライドウィンドウによる疑似ベクトル機構が導入された。CP-PACSプロジェクトでは新たなベクトルプロセッサを開発するのは不可能であったので、現在でも最善の選択であったと思っている。但し、「疑似ベクトル計算機はベクトル計算機に比べて、高い効率を実現するのにユーザーの努力をより必要とする計算機である」というのが、正直な感想である。ベクトル計算機では、フォートランレベルでのコーディングの工夫とチューニングで、これらのサブルーチンで80%程度の効率を実現するのはそれ程困難ではないが、CP-PACSでは30%の効率を実現するのがやっとであった。結局、我々は、QCDMULTやQCDCMRのアセンブラコードを開発し、それを実際の計算に用いることになったわけである。この場合、60%以上の効率を実現できた。参考までに、フォートラン版とアセンブラ版のQCDMULTやQCDCMRの性能の時間方向格子サイズ依存性を図1に再録した。アセンブラ化による2倍以上の高速化は、CP-PACSで多くの素粒子物理の成果を得る事におおいに寄与したと思っている。なお、QCDMULTのアセンブラコードは、最初のバージョンを日立製作所に作製して頂いた。この場を借りて、改めて深謝したい。QCDCMRは私が作製した。実は、私が実アプリケーションで使われるアセンブラコードを書いたのはこれが唯一であり、貴重な体験をしたと思っている。

並列計算機としてのCP-PACSの特徴の一つは、ブロックストライド転送をサポートしたリモートDMA機構であるが、これもQCDの通信パターンを分析し、それを高速に実行する為に開発された方式である。その過程で私が一番興味深かったのは、「メモリコピーが意外に重い」という点である。ブロックストライド転送はユーザーレベルのgather/satterのコストを削減し、リモートDMA機構はユーザーシステム間のメモリコピーを削減する。この様な工夫によって、QCDの実アプリケーションでピーク性能の50%の

性能を達成できた(時間方向の格子サイズが 40 以上の場合). 今後も幾つかの並列計算機の開発プロジェクトに参加する機会もあるかと思うが, その際には, CP-PACS での経験を生かして, ユーザーの立場から, 「メモリコピーをギリギリまで削減した通信」の提案をしていきたいと思っている.

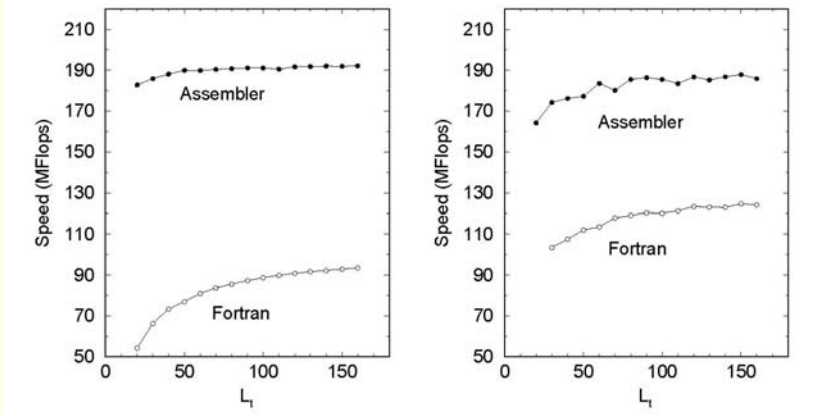


図1. QCDMULT(左)とQCDCMR(右)の実行性能の時間方向格子サイズ  $L_t$  依存性. 格子の時間方向のdoループを疑似ベクトル化し, 空間方向はその繰り返しとした. Even/odd algorithmを用いているので, ベクトル長は  $L_t/2$  である.

計算機の性能の指標の一つはディスク I/O 性能である. 特にクエンチ近似のスペクトラム計算では, 問題規模やパラメータにも依るが, 実行時間の 15-20%がディスク I/O に費やされ, これを如何に削減するかも CP-PACS の課題であった. 実際, CP-PACS の設計段階から QCD プログラムのディスク I/O を分析し, I/O ノードの設計に反映した訳であるが, 私がもっとも興味深かったのは, 「I/O 性能を引き出す為には, ディスクへのシーケンシャルアクセスが必須である」という点である. CP-PACS では 16 個の PU からのディスク I/O が一つの 10U に集中するが, クエンチ計算では, 各 PU からの I/O 命令をシリアライズする事によって, 同一ファイルがディスク上に連続に配置される様工夫した. その他, あらかじめディス

ク上の全てのファイルを消去しておく、処理の軽い C の 1/0 関数を大きなバッファを確保して用いる等の工夫によって、クォーク伝搬関数の出力段階のシーケンシャルな書き込みでは、約 5MB/sec/10U の性能を達成した。これは、ディスク 1/0 の理論ピーク性能の 25% である。クォーク伝搬関数からハドロン伝搬関数を構成するステップは、論理的にランダムな読み込みが発生し、1.3MB/sec/10U の実効速度となった。

以上の様に、クエンチ近似のハドロン質量計算では、演算、通信、ディスク 1/0 の全ての側面で、殆んどギリギリまでチューニングしたプログラムを作製した。よくここまで効率的にシステムを使い果たしたものだ、感慨深い。

並列計算機としての CP-PACS の特徴の一つは、パイパークロスバネットワークである。この方式の利点は多々あるが、QCD 計算に限れば、システム全体を複数のメッシュトラスのパーティションに分割しての運用が可能にした点が、最大の利点であったと思う。実際、CP-PACS 全体を一つのジョブで占有した実アプリケーションは、クエンチ近似の最大サイズの計算で、それは約半年の期間だけであった。運用期間の殆んどすべてで、CP-PACS は計算の粒度に応じて臨機応変に種々のサイズのパーティションに分割して運用された。『CP-PACS の成果、計算素粒子物理学』の項に纏められている様に、CP-PACS では、QCD に限っても実に色々な計算が行われたが、それらを効率良く実行できたのは、「柔軟な分割運転が可能であった」ことに依るところが大きかったのではなかろうか。この点は、なかなか定量的に評価できない事であるが、パイパークロスバという QCD には一見贅沢なネットワークに費用を投じるのは、正しい選択であったと思う。

最後に、CP-PACS の運用に関して感ずることを述べたい。計算物理学(科学)研究センターでは、CP-PACS やそのフロント計算機システムの運用を『計算機システム運用委員会』が担当してきた。特筆したいのは、「運用委員会には、計算機工学者だけでなく、CP-PACS ユーザーである物理サイドの各分野の代表も加わった」点である。システムの問題点や改善すべき事項に最初に気づくのは、多くの場合、実際に計算機を使って

いるユーザーである。その分野のユーザの研究や計算をよく知っている運用委員会のメンバ(ユーザー自身の場合も多々あった)が問題を的確に把握-整理し、計算機工学者やメーカ担当者と協力して問題解決にあたった。ユーザーは我儘なもので、自分達の研究 / 計算の進行がシステムの問題が原因で遅延するのは我慢ならないと感ずるものである。問題の解決を人任せにせず、問題解決の責任の一端をユーザー自身が負っているという方式は、問題解決にかかる時間を短縮するのに、おおいに寄与したのではないかと思う。CP-PACS の成功の一因は、計算機的设计段階から「物理研究者と計算機工学者とメーカが同じ土俵にたって緊密に協力した」ことであるが、緊密な協力関係が10年近い長い運用期間に渡って継続的に維持されたことが、CP-PACS の高い稼働率に、ひいては、CP-PACS による多くの物理学上の成果につながったのだと思う。

2006 年 4 月吉日