

主 論 文 要 旨

報 告 番 号	甲 乙 第	号	氏 名	鈴 木 紀 章
主論文題名				
超並列計算機 JUMP-1 の実装と評価				
内容の要旨				
<p>キャッシュコヒーレントな分散共有メモリをもつマルチプロセッサシステム CC-NUMA (Cache Coherent Non-Uniform Memory Access model) は、中、大規模な並列計算機の代表的な構成方法である。CC-NUMA は、バス結合型の小規模並列計算機に比べて多数のプロセッサを接続でき、またこれらの小規模並列計算機で開発されたプログラムを容易に移植できる利点もある。このため、SGI の Origin、Sequent の NUMA-Q など、商用化が進んでいる。</p> <p>文部省重点領域研究の一環として1994年から開発が行われたJUMP-1は、数千プロセッサを越す超並列計算機上に効率の良い分散共有記憶や同期、メッセージ転送を実現するためのテストベッドである。JUMP-1では従来のものよりはるかに多数のプロセッサ上でキャッシュコヒーレントな分散共有記憶を実現するため様々な画期的な手法が提案されている。JUMP-1は数千プロセッサ規模への拡張を可能とするために、結合網 RDT (Recursive Diagonal Torus) で相互接続されたクラスタ構造を持つ。RDTは二次元トーラスのFat-tree状の階層構造を持つ結合網で、メッセージマルチキャストと応答収集機能を持つルータチップを用いて実現されている。分散共有メモリ管理には、専用プロセッサであるMBP (Memory Based Processor)-lightを用いている。MBP-lightは、シンプルなコアプロセッサ、さらに、各クラスタは並列I/Oサブシステムを構成する高速シリアルリンクSTAFF-Linkにより、ディスクおよび画像データ用Frame Bufferと接続する。また、各クラスタに対してメンテナンス用ホストPCを設け、システムのブートアップやモニタを行う。</p> <p>JUMP-1は、2000年3月に16クラスタ64プロセッサのハードウェアが完成したため、以後実機を用いて様々な評価を行っていくこととなった。本研究では、完成した実機を用いて、RDTネットワークの応答パケットの自動生成、自動収集機構に関する評価、MBP-lightの命令セットアーキテクチャの評価等を行った。本論文では、これらを含めてJUMP-1プロジェクトで得られた知見についてまとめる。</p> <p>RDTネットワークのバンド幅は、最大パケット長である15フリットの場合で100MBytes/secとなる事が分かった。応答パケットの自動生成機構はソフトウェア処理に比べて1.8倍程度、自動収集機構は2.4倍程度高速である事が分かった。JUMP-1のメモリ管理プロセッサMBP-lightは、Buffer-Register Architectureをとることにより、Homeクラスタで5.64%、Remoteクラスタで6.27%の性能向上を達成している。特殊命令では、ハッシュ値を求める命令のみが有効に働いており、これにより2.80%の性能向上を達成している。しかし、頻繁に利用されている命令は共通RISC命令であり、別のアーキテクチャをとる事により、より高性能なものを実装できる可能性がある事が分かった。また、JUMP-1のI/Oシステムは、クラスタと同数の入出力ユニットを用いた場合、readでクラスタあたり2.1Mbps、writeでは4.5Mbpsの実行転送速度を得られ、4クラスタ4入出力ユニットまで直線的に性能が向上する事が分かった。JUMP-1のDSM (Distributed Shared Memory) では、Homeクラスタに有効なデータがある場合のリモートリードで469サイクル(50MHz動作の場合9.38 μ sec)、Home以外のクラスタが有効なデータを保持している場合で850サイクル(50MHz動作の場合17.00 μ sec)となっている。行列積プログラムの実行では、4プロセッサのシステムで3.72倍、8プロセッサのシステムで6.61倍の台数効果が得られる事が分かった。</p>				

