社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2006－HPC－107 (36)
2006／8／2

科学技術計算用超並列クラスタPACS-CSの実装と基本性能評価

朴 泰祐[†] 佐藤三久[†] 高橋大介[†]
宇川彰[?] 深川正一[††] 藤田不二男[††]
清水正明[††] 住元真司[†††] 久門耕一[†††]

筑波大学計算科学研究センターで開発された超並列クラスタPACS-CS (Parallel Array Computer System for Computational Sciences) は，広範囲な計算科学アプリケーションを対象とした新しい発想の超並列クラスタである．実アプリケーションにおける実効性能を高めるべく，各計算ノードのメモリバンド幅及びネットワークバンド幅を重視した設計になっており，さらに超並列ネットワークとしてGigabit Ethernetをトランク結合した3次元ハイパクロスバ網を独自に開発・実装している．

本稿では同システムの基本設計と実装について延べ，また，初期性能評価としてLinpack性能と，ネットワークの基本性能を評価する．Linpackベンチマークにおいては，2560ノードで10.35TFLOPSを記録した．また，一定以上のメッセージ長に関し，MPIレベルでのネットワークバンド幅が理論ピーク性能の89%が達成され，ソフトウェアによる3次元ルーティングにおいても長メッセージの場合のオーバヘッドは低いこと，さらに3次元ハイパクロスバ網が高いbisectionバンド幅を提供できることを示した．

Implementation and Performance Evaluation of the Large Scale Cluster PACS-CS for Scientific Computation

Taisuke Boku,[†] Mitsuhisa Sato,[†] Daisuke Takahashi,[†]
Akira Ukawa,[?] Masakazu Fukagawa,[††] Fujio Fujita,[††]
Masaaki Shimizu,[††] Shinji Sumimoto[†††] and Kouichi Kumon [†††]

At the Center for Computational Sciences, University of Tsukuba, we have developed a massively parallel cluster named PACS-CS (Parallel Array Computer System for Computational Sciences) for wide fields of computational sciences. To achieve high sustained performance on real applications, each node is designed to be aware of the bandwidth both on memory and network. The massively parallel interconnection network based on 3-D Hyper Crossbar with trunked Gigabit Ethernet is also originally developed.

In this paper, we describe the basic design and implementation of the machine as a large scale PC cluster. We also describe the basic system performance including Linpack benchmark and network performance. The system recorded 10.35 TFLOPS on Linpack. For reasonablly large size of message, the interconnection bandwidth achieves 89% of the peak at MPI level, and even for 3-D routed message, it is confirmed the routing overhead is acceptable and the system provides a large bisection bandwidth.

1. はじめに

筑波大学計算科学研究センター[1]では，今後数年間の主力計算リソースとして，2005年度より超並列クラスタPACS-CSの開発を進めてきた[2),3)]．PACS-CS (Parallel Array Computer System for Computational Sciences) は，筑波大学を中心として開発研究が進められてきたPACS及びPAXシリーズ計算機の第7代目に当たる超並列計算機で，計算科学の諸問題の解決をターゲットに開発された超並列PCクラスタである．

PACS-CSは一般の計算センターに導入されている多くのHPC向けクラスタと異なり，実行されるアプリケーションの特性に傾注し，ピーク性能だけでなく実効性能において高い対価格性能比を実現すべく，マザーボードの独自開発，Gigabit Ethernetのソフトウェアトランク技術に基づく3次元ハイパクロスバ網の開発，数千ノード規模のジョブ実行を支援するソフ

† 筑波大学
Universit of Tsukuba
†† 日立製作所
Hitachi Ltd.
††† 富士通研究所
Fujitsu Laboratories Ltd.

トウェア環境の提供等を行っている．同マシンは2006年7月1日より2560ノードによる稼動を開始した．

本稿ではPACS-CSの実装と稼動状況，及び初期性能評価としてLinpackベンチマークと基本的なネットワーク性能とその特性について述べる．

2. PACS-CSの開発

2.1 設計コンセプト

一般の計算センターで運用されるHPCクラスタでは，不特定多数のユーザ・不特定のアプリケーションを対象とするため，なるべく一般的かつ「無難な」システム構成を取る．これに対し，PACS-CSは以下のようなコンセプトの上で設計されている．

- **アプリケーションの特性を限定し，大規模利用を想定したクラスタ構築：**実空間モデリングに基づく近接通信中心のスケーラブルなアプリケーションで数千台規模の計算を行う．
- **CPU当たりのメモリバンド幅を重視：**ピーク性能と実効性能の差を生みやすい共有メモリ構成を避け，キャッシュ利用率が必ずしも高くないアプリケーションに対してもある程度の性能を保証．
- **CPU当たりのネットワークバンド幅を重視：**高い通信バンド幅を要求するアプリケーションに対し，CPU当たりのネットワークリンクバンド幅を確保．
- **実空間モデリングに適したネットワークトポロジ：**3次元直交空間における近接通信を，全ノードが一斉に行っても耐え得る高bisectionバンド幅を提供．
- **コモディティネットワークのトランク技術の利用：**3次元直交ネットワークの実現のために，高バンド幅の単一リンクではなく比較的低いバンド幅のリンクを多数トランクし，多次元配線することにより実効バンド幅を増強．

2.2 PACS-CSの設計と基本仕様

PACS-CSの設計については[2),3)]に詳しい．ここでは以降の実装及び性能評価に関連する概略のみを記す．

2.2.1 ノード設計

PACS-CSはコモディティ製品を利用した高性能PCクラスタであり，CPUにはIntel IA-32アーキテクチャを採用している．しかしながら，従来のHPC向けクラスタのように1つのノード上に複数のCPUを搭載しない．XeonのようなIntel系プロセッサでは複数CPUによってチップセット経由でメモリを共有することにより，CPU当たりの実効メモリバンド幅が制限される．我々が主要ターゲットアプリケーションと考えている素粒子物理学のLattice Quantum Chromo-Dynamicsや物性第一原理計算であるReal Space Density Function Theory等の大規模シミュレーションでは，キャッシュに依存した処理が期待できず，有効メモリバンド幅が実効性能に直接的なインパクトを与える．

また，Opteronのような AMD系プロセッサではメモリコントローラをCPU内に内蔵することにより，CPU数にスケーラブルなメモリバンド幅を確保するNUMA型共有メモリが実現可能である．しかし，Linuxを含む現在のOS及びコンパイラ環境は，実効性能を上げるためのデータ局所化を完全にはサポートしていない．加えて，ネットワーク上でのデータ転送時に特定のCPUに負荷がかかり，かつDMAを介したメモリへのデータ転送時に上述のNUMA性により負荷バランスが崩れる可能性もある．以上の考察より，PACS-CSではノードを単一CPUでのみ構成し，メモリバンド幅・ネットワークバンド幅の両者をCPUに占有させる設計となっている．

2.2.2 ネットワーク設計

一般的なHPC向けPCクラスタではSANの導入により，ネットワークリンクに高い性能を求める．しかしながら，システムの大規模化に伴い，ツリー型ネットワークではノード数Nに対し一般的に$N\log N$オーダー以上のスイッチ用ハードウェアが必要となり，さらにツリーのルート付近でのバンド幅ボトルネックは避けられない．我々は数千ノードのシステム全体を使い切るまでの超並列ジョブの実行を想定しており，特に多次元直交空間における隣接通信を行った場合，ツリー型ネットワークではある次元方向での通信の全トラフィックがツリーのルート付近を一斉に通過する場合が想定される．これらを考慮し，CP-PACS[5)]で用いられた3次元ハイパクロスバ網（以下，3D-HXBと略）[6)]をPACS-CSに導入した．

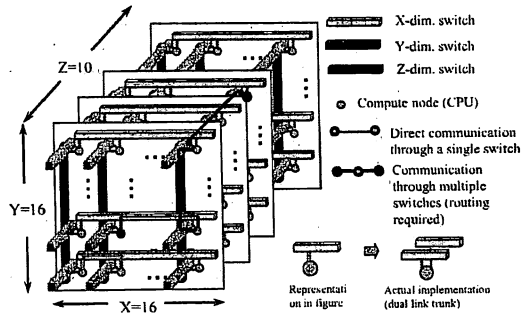

図1 PACS-CSのネットワーク構成

図1にPACS-CSのノード結合ネットワークのブロックダイアグラムを示す．各次元のサイズをX, Y及びZとする（ここでは$X=16, Y=16, Z=10$）．3次元空間に直交配置された各ノードは3本のリンクを持ち，それぞれがx, y及びz次元のスイッチに結合される．x次元方向のスイッチは1つのスイッチがX個のポートを持ち，これが$Y \times Z$個だけある．y, z次元についても同様で，ポート数Yのy次元方向

スイッチが $X \times Z$ 個，及びポート数 Z の z 次元方向スイッチが $X \times Y$ 個だけ，それぞれ存在する．もし転送元と転送先のノードのアドレスが x 次元方向でのみ異なる場合は，x 次元スイッチを一度経由するだけで転送先ノードに到達できる（y 次元，z 次元方向の 1 次元だけが異なる場合も同様）．もし 2 つあるいは全ての次元のアドレスが異なる場合，x 次元→ y 次元→ z 次元のスイッチを各々一度ずつ経由することにより，必ず転送先ノードに到達できる．

このように 3D-HXB 網では多次元方向への通信を複数のリンクに分散させるため，複数のネットワークインタフェース (NIC) を複数の PCI バスに分散させ，ノード上での PCI バスボトルネックを解消できる．それと同時に，1 つのメッセージは 3 次元どれかの方向のリンクから送出されるため，メッセージ当たりに利用可能なバンド幅は相対的に小さくなってしまう．しかし，我々の主要アプリケーションでは多次元実物理空間上での多次元格子にマップ可能なデータ構造を主眼としており，多次元方向へのデータの同時送出がアルゴリズム上可能である．よって，3 次元方向のリンクを全て同時に利用した「多次元同時転送」を行うことにより，バンド幅を最大で 3 倍まで増強することが可能である．PACS-CS では 1 次元方向の論理リンクを 2 本の Gigabit Ethernet（以下，GbE と省略）を束ねて用いることにより，バンド幅を 2 倍に増強している．これを 3 次元同時転送に適用することにより，最大で GbE リンクの 6 倍のバンド幅を得ている．

2.2.3 ネットワークドライバ

PACS-CS の OS としては通常の HPC 向け PC クラスタと同様，Linux を用いる．ただし，上述の 3D-HXB 網を構成するために，GbE リンクのトランキングと多次元 HXB のためのルーティング機能を持つ特殊なネットワークドライバソフトが必要となる．我々はこれを SCore[8] 上で以前から提供されている PM/Ethernet[7] をベースに，新たに PM/Ethernet-HXB として開発した[9),10)]．同ドライバでは数千ノードからなる超並列構成の SCore 環境を想定し，次元当たり複数本の GbE NIC のトランキングと，多次元 HXB 上でのルーティング機能を提供する．さらに今後，ユーザレベルから直接 PM ドライバを簡単に利用できるようにするために，拡張された PM ドライバ API の開発を進めている[11)]．本 API により，ユーザメモリ空間における連続領域あるいは非連続領域上のデータ転送を各ノード当たり 1 回のコピーで転送することが可能で，かつ多次元同時転送もスムーズにプログラム可能になり，MPI レベルよりもさらに高速なデータ通信を可能とする予定である．

2.2.4 PACS-CS の諸元

以上の基本設計に基づいて構築された，PACS-CS の諸元を表 1 に示す．

表 1 PACS-CS の諸元

ノード台数	2560 (16 × 16 × 10)
理論ピーク性能	14.3 Tflops
ノード構成	単一 CPU/ノード
CPU	Intel Low Voltage Xeon EM64T 2.8 GHz, 1MB cache
メモリ容量	2GB/ノード
メモリバンド幅	6.4GB/s (1Gflops 当り 1.14GB/s)
並列処理ネットワーク	3 次元ハイパクロスバ
リンクバンド幅	単方向 250MB/s/次元 単方向 750MB/s (3 次元同時)
bisection バンド幅	640 GB/s
ローカルディスク容量	320 GB/ノード (RAID-1 により実効 160GB)
ファイルサーバディスク容量	10TB (RAID-5)
オペレーティングシステム	Linux (FedoraCore3)
システム管理ソフトウェア	SCore
プログラミング言語	Fortran90, C, C++, MPI
システム規模	総ラック数：59 総消費電力：545kW

3. PACS-CS の実装

3.1 シャーシ実装

PACS-CS の開発に当たっては，ハードウェア及び全体システムの実装を日立製作所，PM/Ethernet-HXB の実装を富士通がそれぞれ担当し，筑波大学を中心とした三組織での産学連携体制で行った．

ハードウェアの開発上特に重要だったのは，多数の GbE ポートをオンボード実装した新規開発マザーボードである．PACS-CS では並列処理データ転送用に 1 次元当たり GbE × 2 ポート，3 次元合計で 6 ポートの GbE NIC を必要とする．これらに加え，一般的な UNIX ネットワーク処理（login, NFS 等）を行うため，ツリー構造の GbE ネットワーク（運用系ネットワークと呼ぶ）を構築し，さらに各ノードに対する BIOS レベルのアクセス（リセット，電源 ON/OFF 等の制御）を行うための Serial over LAN 回線として別途 GbE ネットワーク（制御系ネットワークと呼ぶ）を構築している．従って，各ノードには合計 8 ポートの GbE NIC が必要となる．従来の Intel Xeon を用いた HPC 向けクラスタと同等の，19 inch ラックマウントシャーシで 1U サイズの中に 2 ノードを実装するため，8 ポートの GbE NIC はオンボード実装する必要がある．このため，PACS-CS では高密度のブレード型マザーボードを新たに開発した．

図 2 に 1U シャーシ内の様子を示す．図の左側の上下に 1 ノード分のマザーボードが実装されており，中央の電源を共有している．右側には 4 台の HDD が 19 inch 幅にほぼぴったりに装着されている．HDD 部とマザーボードの間の基板はファン，HDD 及び電源の制御ボードである．2 台の独立な PC が電源及び空冷

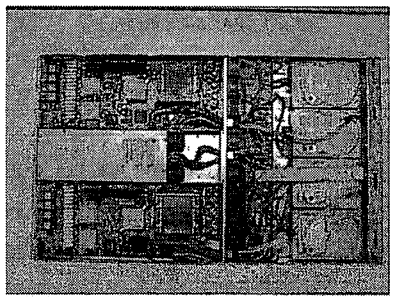

図 2 1U シャーシ内の様子（2 枚のマザーボードと 4 ドライブの HDD）

ファンを共有しているため，ブート時には 1 ノード起動でもファンが回転し，シャットダウン時は両ノードが落ちるまではファンを止めないという動作を行う．

3.2 筐体配置とケーブリング

PACS-CS の筐体は標準の 19 inch ラックである．全システムはノードラック（40 台），3D-HXB ネットワークスイッチラック（15 台），汎用ネットワークスイッチラック（3 台），サーバ用ラック（1 台）の計 59 ラックからなる．1 台のノードラックには 2 ノードを内蔵する 1U シャーシ 32 台（64 ノード分）がマウントされる．3D-HXB ネットワークスイッチラックは XY スイッチラックと Z スイッチラックに分けられる．全ての GbE ネットワークスイッチは，日立電線製 Apresia4348GT で，48 個の GbE ポートを備える．3D-HXB の構成は $16 \times 16 \times 10$ なので，最大で 16 ポートの GbE スイッチがあれば足りる．よって，1 台のスイッチを VLAN によって 3 つの 16 ポートスイッチとして仮想的に分割し，X, Y または Z 次元スイッチに充てる．

3D-HXB 上のシステムは，16×16 の XY 次元平面を構成する合計 256 ノードを 1 つのラックグループとし，これが Z 次元方向に 10 面分あるという構造になっている．1 つのラックグループは 4 台のノードラックと 1 台の XY 次元用スイッチラックからなる．4 台のノードラックに XY 平面が完全に包含されるため，これらの間で閉じる X 次元及び Y 次元の全スイッチ（16 ポート換算で 32 台相当，実際の 48 ポートスイッチで 11 台）を 1 台のスイッチラックに格納している．図 3 に，ノードラックとスイッチラックのケーブリングの様子を示す．

Z 次元スイッチは，全てのラックグループから束ねられた Z 次元方向の GbE リンクを集約する．16 ポート換算で 256 台相当のスイッチを格納するために 5 台のスイッチラックが充てられ，2 つずつのラックグループの中央に位置するよう配置されている．

この他，PACS-CS では運用系ネットワーク，制御系ネットワークの他に，監視系ネットワークを設けている．これは，システム中に分散配置されている大量

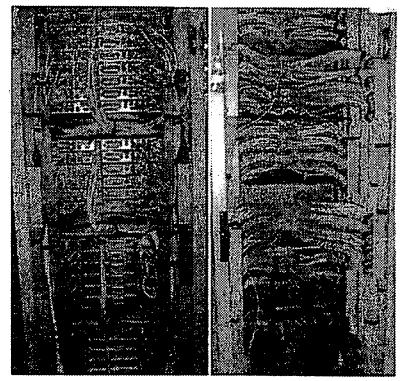

図 3 ノードラック（左）及びスイッチラック（右）の GbE ケーブル配線の様子

の GbE スイッチを SNMP により監視し，どのスイッチのどのポートに異常が発生しても即座にその位置を特定し，停止・交換等の保守作業に対応できるようにするためのものである．3D-HXB 以外のこれらのスイッチ群は，3 台の汎用ネットワークスイッチラックと，XY 次元スイッチラックの隙間に分散配置されている．

全システムは 3 台の SCore 管理ノードと 1 台の統合管理ノードにより制御される．PACS-CS は全システム 2560 ノードを最小 64 ノード，最大 2560 ノードの様々なサイズのパーティションで分割あるいは一括運転することが可能であり，統合管理ノードからサブミットされたジョブが適宜 3 台の SCore 管理ノードに分散され，各々の上で OpenPBS[12] によるジョブキュー管理の下で実行される．この他，容量 10TB の RAID ファイルサーバがあり，これらをまとめて 1 台のサーバラックにマウントしている．図 4 に，全 59 ラックからなる PACS-CS システムの全貌を示す．3D-HXB を構成する 15360 本とそれ以外のネットワーク用の約 5000 本の Category5E Ethernet ケーブルは全て床下で配線されている．

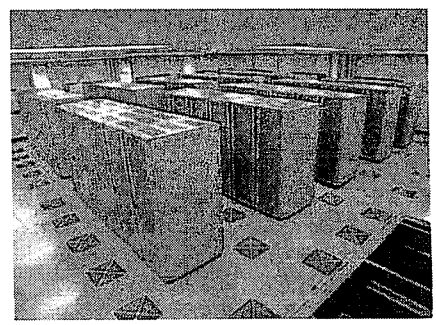

図 4 PACS-CS の全景（2560 ノード，59 ラック）

4. Linpack 性能評価

全2560ノードを用いた性能評価の例として，Linpack ベンチマーク[4] 結果を示す．PACS-CS はバンド幅を意識した実効性能指向のクラスタである．Linpack は一般的にネットワーク及びメモリバンド幅を消費しないベンチマークであるが，高性能計算機の性能指標として世界標準的に用いられているため，これを測定した．測定には HPL (High Performance Linpack)[13] を用い，ソースコードには特に修正を加えていない．PACS-CS 上で利用可能な MPICH-1.2 on PM/Ethernet を，独自の PM/Ethernet-HXB ドライバ上で運用する環境を用いた．ソースコードのコンパイルには gcc-3.4.4 を用いた．

表 2 Linpack ベンチマーク結果

$P \times Q$	N	R_{max} (Tflops)	Efficiency (%)
16 × 160	706560	10.33	72.05
32 × 80	722944	10.35	72.20

表 2 に結果を示す．HPL では N^2 の行列を $P \times Q$ という 2 次元のノード空間に適当なブロックサイズでサイクリックマッピングするが，我々は表に示す 2 種類を試した．16 × 160 のケースは，X 次元を第 1 次元に，YZ 次元を合わせて第 2 次元に割り当てた場合である．これに対し，32 × 80 は Y 次元が 2 × 8 に分割され，$P = 16(X) \times 2(Y)$，$Q = 8(Y) \times 10(Z)$ というように 2 次元の論理ノード平面が構成されている．HPL では対象行列の行方向の入れ替えと列方向のブロードキャストがそれぞれ P 及び Q の各方向で独立に発生する．従って，前者のマッピングでは通信は 1 次元内の転送で閉じるが，後者のマッピングでは XY 次元及び YZ 次元に跨る通信が発生する．

一般的に，通信性能に偏りがない場合，$P = Q$ に近い構成，すなわち 2 次元ノード空間を正方形に近い形に取るのが最も効率的である．しかし，3D-HXB 構造を持つ PACS-CS においては，多次元に跨る転送でのソフトウェアルーティングにより通信性能低下が生じるため，P,Q の最適値は一意に決定できず，演算性能と通信性能とトレードオフが生じることになる．実際，表に示すように両者の差は僅かであるが，今回は多次元ルーティングが生じても演算バランスを取った 32 × 80 の構成が若干高い性能を示した．

なお，この結果は 2006 年上半期の TOP500[4] リストに登録されており，PACS-CS は 34 位にランクされた．これは，国内ベンダーによるマシンとしては地球シミュレータ (10 位) に次ぐ第二位の値である．

5. ネットワーク性能評価

GbE トランクによる 3D-HXB という，PACS-CS の特殊なネットワーク構成を評価するため，MPI による一対一の pingpong 転送性能を測定した．GbE が一般的な SAN に比べ大きなレイテンシを持つことから，実アプリケーション上でどの程度のメッセージ長を確保することが実効性能につながるかを示す指標となる．加えて，通信が 2 次元または 3 次元に跨る場合，中継ノードにおけるソフトウェアルーティングのオーバヘッドの影響も調べる．

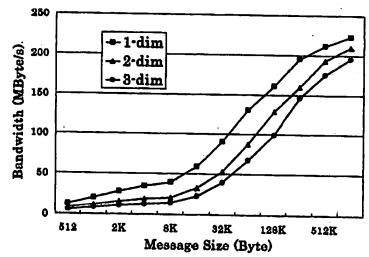

図 5 多次元転送を行った場合のバンド幅

図 5 に，1 次元〜3 次元通信における MPI メッセージ長と実効バンド幅の関係を示す．512 ノード (8×8×8) のパーティションで，ノード (0,0,0) を基準として，ノード (1,0,0) への通信 (“1-dim”)，ノード (1,1,0) への通信 (“2-dim”) 及びノード (1,1,1) への通信 (“3-dim”) の 3 通りを測定した．HXB であるため，転送元と転送先のアドレスが何次元分異なったかが重要であり，同一次元上での距離 (例えば相手が (1,0,0) なのか (2,0,0) なのか) は重要ではない．よって，この 3 通りが本質的な多次元ルーティングによる性能差を示している．なお，先述した「多次元同時転送」と，ここで言う 2〜3 次元転送は意味が異なる．前者は「複数次元方向のリンクに対する同時データ送出」を表しており，それはこの性能評価での 1 次元転送を複数個同時実行することに相当する．この評価で示している多次元ルーティングとは，1 つのメッセージの転送が 1 つの次元方向のみの通信で収まらない場合を指す．

1 次元転送の場合，最大で 223MB/sec の通信性能が得られている (メッセージ長 1MB)．これは，2 本の GbE リンクをトランクした場合の理論ピーク性能である 250MB/sec の 89%に当たる．半性能長は約 60KB となるが，科学技術計算における実アプリケーションでは，粒度の高い並列処理を行った場合，数百 KB 以上のデータを送ることが多く，十分高い効率で通信が可能であると言える．続いて，通信が 2 次元または 3 次元に跨る場合，通信性能が 1 次元の場合に比べて低下しているのがわかる．特にメッセージ長が短い場合，例えば 32KB の場合でそれぞれ 1 次元の場合の 59%及び 44%に留まっている．しかし，128KB メッセージでは 2 次元は 1 次元の場合の 80%，3 次元でも 62%に達しており，1MB メッセージでは 3 次元転送で

もほぼ 200MB/sec に達している．我々は PACS-CS におけるデータ通信は多次元隣接通信が基本であると考え設計を行っている．よって，多数次元に跨るルーティングはランダム的な通信を行うアプリケーションで発生する特殊な場合であると捉え，その場合でもある程度の通信性能，特に長メッセージの通信では 1 次元の場合とほぼ等しい性能が得られていることから十分満足できると考えている．

最後に，3D-HXB による bisection バンド幅について評価する．図 6 は，サイズ 8×8×8 のパーティションにおいて 2×1×1 から 8×8×8 までのあらゆるサイズの直方体のノード空間を対象に，X 次元方向の中央でノードグループを半分に分割して X 次元の低いアドレスのグループと高いアドレスのグループ間で一対一 pingpong 転送を一斉に行った場合の総バンド幅を表している．pingpong 転送であるため，総バンド幅は実際の bisection バンド幅の半分に相当する．なお，ここではバースト転送のような状況での最大バンド幅を求めるため，MPI でのメッセージ長を 1MB に固定している．

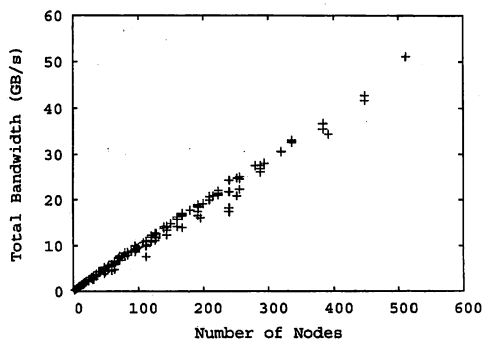

図 6　X 次元方向への一斉転送時の総バンド幅

結果からわかる通り，ノード数の増加に伴い，総バンド幅もほぼリニアに増加している．全 257 通りのノード空間サイズにおいて，一組のノード対間での pingpong 通信バンド幅の平均は 196.5MB/sec であった．また，8 × 8 × 8 = 512 ノードの総バンド幅は 51.1GB/sec であった．この場合の理論ピーク総バンド幅は 64GB/sec であるから，理論ピークの 80%のバンド幅が得られていることになる．PACS-CS における 3D-HXB は全次元について等しく対称に構成されている．従って，この特性は他の全ての次元についても共通するものである．

6. おわりに

本稿では 2006 年 7 月 1 日に稼動開始した超並列クラスタ PACS-CS の実装と初期性能評価について述べた．PACS-CS は従来の HPC 向けクラスタと同等の実装密度を保ちつつ，単一 CPU ノードと GbE トランクによる 3D-HXB により，実空間モデリングによる計算科学において非常に高い実効性能が期待される．Linpack ベンチマークにおいては GbE トランク 2 本（多次元同時転送は用いていない）での通信で，2560 ノードを用いて 10.35TFLOPS（効率 72.2%）を達成した．また基本的な通信性能として，比較的長いメッセージ長において，1 次元当たり理論ピーク性能の 89%までの通信バンド幅が確認され，512 ノードまでのパーティションにおいて，システムサイズにスケールした通信バンド幅が得られることがわかった．

本システムはまだ稼動を開始したばかりであり，ネットワークを中心としたより詳細な性能評価が待たれる．特に，今回時間の関係で省略した，多次元同時転送性能の大規模システムでの評価を行い，また実アプリケーションにおける全体的な性能評価を行っていく予定である．

謝辞　本プロジェクトを進めるに当り、多くの面で協力を頂いた、筑波大学計算科学計算センター，日立製作所，富士通の関係者諸氏に深く感謝する。

参考文献

1) http://www.ccs.tsukuba.ac.jp/
2) 朴 泰祐, 他 "計算科学のための超並列クラスタ PACS-CS の概要", 情処研報「ハイパフォーマンスコンピューティング」, No.2005-HPC-103, pp.133-138, 2005.
3) T. Boku, et. al., "PACS-CS: A large-scale bandwidth-aware PC cluster for scientific computations", Proc. of CCGrid2006, pp.233-240, 2006.
4) http://www.top500.org/
5) T. Boku, et al., "CP-PACS: A massively parallel processor for large scale scientific calculations", Proc. ICS'97, 1997.
6) 朴 泰祐 他, "ハイパクロスバ・ネットワークにおける転送性能向上のための手法とその評価", 情報処理学会論文誌 Vol.36, No.7, pp. 1610-1618, 1995.
7) 住元 真司 他, "複数の Ethernet を束ねる Network Trunking 機構の提案と 1024 プロセッサ PC クラスタ上での性能評価", HPCS2004 論文集, 2002.
8) http://www.pccluster.org/
9) 住元 真司 他, "複数 Gigabit Ethernet を用いた PACS-CS のための高性能通信機構の設計と評価", SACSIS2006 論文集, 2006.
10) S. Sumimoto, et. al., "A Scalable Communication Layer for Multi-Dimensional Hyper Crossbar Network Using Multiple Gigabit Ethernet", Proc. of ICS2006, 2006.
11) 住元 真司, 他 "PACS-CS のための高性能通信ライブラリインターフェイスの設計", 情処研報 2006-HPC-107 (SWoPP 高知 2006), 2006.
12) http://www.openpbs.org/
13) http://www.netlib.org/benchmark/hpl/