

# PC クラスタにおけるベンチマークと並列ツールの紹介

廣安 知之，三木 光範，大向 一輝，吉田 純一

同志社大学 工学部 知識工学科  
知的システムデザイン研究室

一般の PC 用のベンチマークは多数存在する．また、並列計算機用のベンチマークとしては linpack や Nas Para のベンチマークが有名である．しかしながら計算機の特性全てはそれぞれのベンチマークで把握できるものではなくいくつかの特徴がある．そこで、本発表ではいくつかのこれらのベンチマークを紹介する．特に流体計算結果に比較的近い値を出すと言われている姫野ベンチや同志社大学で開発した遺伝的アルゴリズムを利用したベンチマークの使用結果を報告する．また、近年、Linux の普及に伴い、Linux 上で作動する free で open な並列処理ツールの数が増えてきている．本発表の後半ではその中からいくつかのツールについて報告する．

## 1 はじめに

安価で高速な並列コンピュータとして、PC クラスタが注目されている．通常の利用でも十分にその高速性は実感できるが、実際にどの程度のパフォーマンスが得られているのかはわかりにくい．また「安価で高速」というものの汎用のスーパーコンピュータとの性能の比較や、コストパフォーマンスの比較も興味のあるところであろう．そこで、本発表では、PC クラスタの性能を厳密に評価するために、いくつかのベンチマークプログラムを紹介する．また、クラスタの利用をより便利にするようなフリーの並列ツールについても紹介する．

## 2 今回使用したクラスタ

本発表で紹介するプログラムは Linux ベースの PC クラスタにおいてテストを行った．我々のクラスタは以下のような仕様になっている．

- ・ OS :  
Debian/GNU Linux(kernel2.2.12)
- ・ 構成台数 :  
1+8 台 (うち 2 台はメンテナンス中)
- ・ CPU :  
Pentium III 500MHz × 2
- ・ DRAM :  
128MB ( Master のみ 256MB )
- ・ ネットワーク構成 :  
100BASE + Swiching HUB
- ・ 並列ライブラリ :  
MPI, PVM3

## 3 ベンチマークプログラム

### 3.1 ベンチマークとは

ベンチマーク ( benchmark ) という語には、一般的に「基準となるもの」という意味がある．しかしながら、コンピュータ用語としてのベンチマークとは、コンピュータのハードウェアやソフトウェアの処理速度を計測する試験問題、およびそれによる性能評価という意味を持つ．つまり、計算機システムの性能を計測すること自体を指す言葉として使われている．コンピュータは、プログラムによってさまざまな挙動を示すうえ、その用途もさまざまあるため、比較の指標を一意に決めることは困難である．そこで、各用途に応じた性能評価の指標を決定し、その尺度にもとづいて性能を評価するということが重要になる．その結果として、さまざまな用途に応じた性能を評価するためのベンチマークが必要になるのである．

ベンチマークの対象となる機能を分類すると次のようになる．

- ・ 演算性能ベンチマーク
- ・ I/O 性能ベンチマーク
- ・ グラフィックス関連性能ベンチマーク
- ・ ネットワーク関連性能ベンチマーク
- ・ データベース性能ベンチマーク
- ・ アプリケーション指向ベンチマーク
- ・ 並列/大規模計算ベンチマーク

本発表ではこの中でも特に PC クラスタに関係が深いと考えられる、並列計算ベンチマークとネットワーク性能ベンチマークについていくつかのプログラムを紹介する．並列計算ベンチマークとして、Nas Parallel Benchmark , ScaLAPACK , 姫野

ベンチ, GA Benchmark を紹介する. また, ネットワーク性能ベンチマークとしては netperf を紹介する.

一般に, PC クラスタは安価で高速な並列・分散計算環境と認識されているが, 実際にどの程度の性能があるのかはわかりにくい. 以下に紹介するベンチマークプログラムをもちいることで, 汎用のスーパーコンピュータとの性能比較も可能となる.

### 3.2 Nas Parallel Benchmark

NAS Parallel Benchmarks (NPB) は, NASA Ames Research Center で開発された, 並列コンピュータのためのベンチマークである. NPB は, 5 つの Parallel Kernel Benchmarks と 3 つの Parallel CFD (Computational Fluid Dynamics) Application Benchmarks から構成されている. それぞれの問題に対し, 問題サイズや反復回数が異なる 5 つのクラス<sup>1</sup>が定義されており, それぞれの環境に応じたベンチマークを行うことができる. NPB は, 並列コンピュータの実効性能を知るうえで, 権威あるベンチマークの 1 つであるといえる.

また, NASA Ames Research Center より定期的に WWW などで行われているレポート (NAS Parallel Benchmark Results) では, プログラム実行に要した経過時間 (elapsed time) と, Y-MP/1 または C90/1 との性能比などについて報告がなされており, 大規模計算機と PC クラスタの性能の比較も可能である.

### 3.3 ScaLAPACK

テネシー大学の Jack Dongarra によって開発された, 線形代数問題の LU 分解を行うベンチマーク LINPACK の並列版である. LINPACK は,  $N \times N$  行列について, 単精度および倍精度で  $y(i) = y(i) + a * x(i)$  を演算する. 演算の特性上, ベクトル化による効果が顕著に表れるベンチマークである. また, LINPACK は世界のスーパーコンピュータランキングである「TOP500 Supercomputer」のベンチマークに採用されている.

ScaLAPACK では, 行列を分割し, それぞれを 1 つのプロセスが担当する. プロセス間通信には MPI もしくは PVM を用い, これによって PC クラスタでも実行・計測が可能になっている.

<sup>1</sup> S(ample), W(orkstation), A, B, C の 5 つ. C が一番難易度が高い

<sup>2</sup> Round Trip Time

### 3.4 姫野ベンチ (Himeno Benchmark)

姫野ベンチマークは理化学研究所の姫野龍太郎氏が非圧縮流体解析コードの性能評価のために考案したベンチマークでポアソン方程式解法をヤコビの反復法で解く場合に主要なループの処理速度を計るものである.

コードは非常に短く簡単にコンパイル・実行できるので, 即座に実測速度を求めることが特徴である. Windows や Machintosh でも利用可能である.

### 3.5 GA Benchmark(GAPPA bench)

GA Benchmark は我々の研究室で開発中の PDC 汎用 GA プログラムを応用したベンチマークである. 遺伝的アルゴリズム (GA) は生物の進化を応用した, 確率的最適化手法である. なかでも, 分散遺伝的アルゴリズムでは, 母集団を複数のサブ母集団に分割して GA を実行するため並列化に適したモデルであるといえる.

GA ベンチマークでは分散遺伝的アルゴリズムを用いた関数の最適化に要する時間を計測し, 並列計算性能を評価する.

### 3.6 netperf

netperf ベンチマークはネットワーク性能に対するベンチマークプログラムである. 一般にネットワークベンチマークは

- ・実ネットワークの性能評価
- ・ネットワーク・スタックの性能評価

等に利用される. また, ネットワークの性能は次の 2 点で決定される.

- ・スプールネット:  
単位時間における処理量を示す指標
- ・遅延:  
レイテンシ (待ち時間) と RTT<sup>2</sup> (往復時間)

netperf はネットワークに接続された 2 ノード間における TCP および UDP を利用した通信のスプールネットおよび, ネットワーク遅延を計測するベンチマークである.

## 4 姫野ベンチマーク実行結果

我々のクラスタにおいて、姫野ベンチマークを実行した結果を挙げる。図1は使用したプロセッサ数と性能の関係を示したものである。1CPUの場合に約65MFLOPS,7CPUの場合に約340MFLOPSと、完全に比例する値は出ないが、グラフに見るとおり、プロセッサ数に応じて順調に性能が上がっていることがわかる。この傾向は、32CPU,64CPUといった並列計算機においても同様である。

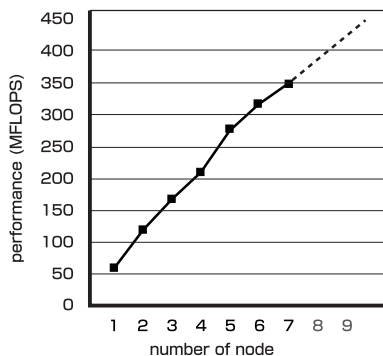


図 1: 姫野ベンチの結果

また、表2は姫野ベンチマークのHPに掲載されていた、他の並列計算機の結果との比較(一部)である。

図 2: 他のマシンとの性能比較

Manufacturer / MACHINE (CPU)	PROC.	MFLOPS
1 FUJITSU VPP700E	16	7986.84
2 FUJITSU VPP5000	1	4137.30
3 HITACHI SR2201	64	3586.95
4 HITACHI SR8000	1	3565.98
5 SGI Origin2000(R12000)	16	3399.00
6 SGI CRAY-T3E (Alpha21164)	64	3275.73
7 SGI Reality Monster(R10000)	16	2503.70
8 NEC SX5/S1	1	1722.00
9 IBM RS/6000SP(PowerPC604e)	64	836.56
10 SUN Ultra Enterprise10000	16	575.49
11 mikilab LinuxCluster(Pentium III)	7	340.97
12 HPC AlphaUP(Alpha21264)	1	228.00

## 5 並列ツールの紹介

### 5.1 Beowulf Watch

クラスタでは多数のノードの状況を把握するのは難しい。すべてのノードの状況を一覧表示してくれるプログラムがあれば便利である。

Beowulf Watch( bWatch )はクラスタの各ノードのメモリやプロセスの状態を表示するための簡

単な Tcl/tkプログラムである。このプログラムは、rsh を利用しているため、ルート以外のいかなるユーザーでも実行可能である。インストールや設定は簡単なのに対し、非常に有効なソフトウェアである。

Host Name	Num Users	Time	1 min Load	5 min Load	15 min Load	Total Mem	Free Mem	Shared Mem	Buffers	Cache	Total Swap	Free Swap
node1	0 users	4:17	0.00	0.07	0.15	377 Mb	1626 Kb	58 Mb	7 Mb	337 Mb	702 Mb	0 Mb
node2	0 users	4:17	4.35	0.85	1.14	61 Mb	38288 Kb	11 Mb	0 Mb	18 Mb	0 Mb	0 Mb
node3	0 users	4:17	1.08	0.54	0.56	61 Mb	40376 Kb	8 Mb	0 Mb	18 Mb	0 Mb	0 Mb
node4	0 users	4:17	1.02	0.58	0.67	61 Mb	32880 Kb	9 Mb	0 Mb	20 Mb	0 Mb	0 Mb
node5	0 users	4:17	1.08	0.50	0.52	61 Mb	40544 Kb	7 Mb	0 Mb	18 Mb	0 Mb	0 Mb
node6	0 users	4:17	0.99	0.89	0.51	61 Mb	40552 Kb	7 Mb	0 Mb	18 Mb	0 Mb	0 Mb
node8	0 users	4:18	0.99	0.89	0.51	61 Mb	41532 Kb	7 Mb	0 Mb	17 Mb	0 Mb	0 Mb
node9	0 users	4:18	3.25	2.79	2.88	61 Mb	41388 Kb	7 Mb	0 Mb	17 Mb	0 Mb	0 Mb
node10	0 users	4:18	2.07	1.61	1.06	61 Mb	41552 Kb	7 Mb	0 Mb	17 Mb	0 Mb	0 Mb
node11	0 users	4:18	0.99	0.90	0.51	62 Mb	27040 Kb	8 Mb	0 Mb	32 Mb	0 Mb	0 Mb
node12	0 users	4:18	3.37	2.74	2.82	62 Mb	44032 Kb	6 Mb	0 Mb	16 Mb	0 Mb	0 Mb
node13	0 users	4:19	14.77	11.04	6.72	62 Mb	16036 Kb	3 Mb	48 Mb	0 Mb	0 Mb	0 Mb

図 3: Beowulf Watch

### 5.2 XPVM

並列プログラムを作成して実行した場合に、そのプログラムが良いのか悪いのか、悪いとすればどこが悪いのか、どのように実行されているのかなど誰もが知りたいところである。このような要求に応えるのがXPVMである。

XPVMはPVMのためのグラフィカルなコンソールとモニタであり、PVMで書かれたプログラムの実行状況を簡単に確認することができる。

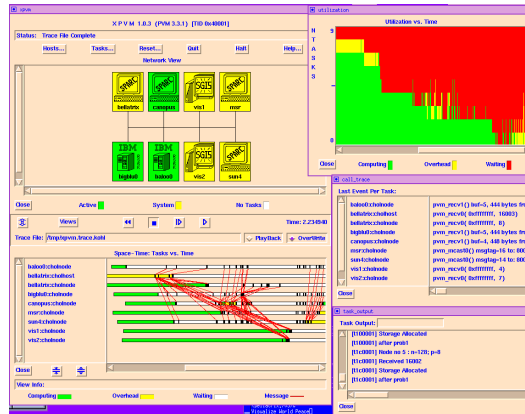


図 4: XPVM

## 6 おわりに

ここで紹介したベンチマークプログラムや並列ツールは次節に示したサイトから入手可能である。また、これらのプログラムのインストール方法などは超並列計算研究会のホームページ( <http://www.is.doshisha.ac.jp/SMPP/> )上で紹介していく予定である。

## 7 関連サイト

本発表で紹介したプログラムの入手先をまとめた．ダウンロードの方法などは各サイトの説明を参照のこと．

### 7.1 ベンチマークプログラム

NAS Parallel Benchmark :

<http://www.nas.nasa.gov/Software/NPB/>

ScaLAPACK Project :

<http://www.netlib.org/scalapack/>

姫野ベンチマーク :

<http://w3cic.riken.go.jp/HPC/HimenoBMT/>

GA Benchmark :

<http://mikilab.doshisha.ac.jp/dia/>

Netperf :

<http://www.netperf.org>

### 7.2 並列ツール

Beowulf Watch : :

<http://www.sci.usq.edu.au/staff/jacek/bWatch/>

XPVM :

<http://www.netlib.org/utk/icl/xpvm/xpvm.html>

### 7.3 その他

Beowulf-underground(並列ツール全般) :

<http://www.beowulf-underground.org/software.html>

Netlib ( 並列関連アーカイブ全般 ) :

<http://www.netlib.org/>

TOP500 Supercomputing Sites :

<http://phase.etl.go.jp/top500/>

超並列計算研究会 :

<http://www.is.doshisha.ac.jp/SMPP/>