

PCクラスタの性能評価

-その1 PCクラスタの構築-

日大生産工(学部) 松田 貴行
日大生産工(院) 高橋 大士
日大生産工 角田 和彦

1. はじめに

高度情報化社会で取扱われるデータ量は年々増加する傾向にある。これらの大量なデータを高速に処理するためには並列型コンピュータシステムの開発が必要となってきた。その中でもシステム構築が容易にできるPCクラスタの開発が進められている¹⁾。

本論文では、8台のPCを用いたベオウルフ型クラスタを構築し、その構築の際に必要なNFS (Network File System)²⁾ や NIS (Network Information Service)³⁾ などについても言及する。

2. NFS²⁾

UNIXシステムで利用されるファイル共有システムである。Sun Microsystems社によって開発され、UNIX系OSにおける標準的な分散ファイルシステムとなっている。

NFSを使うと離れた場所にあるコンピュータのファイルを、あたかも自分のコンピュータにあるファイルのように操作することができる。ファイルを他のコンピュータに提供するコンピュータをNFSサーバ、他のコンピュータにあるファイルを利用するコンピュータをNFSクライアントという。UNIX以外のプラットフォームからも利用することができる。

3. NIS³⁾

Sun Microsystems社が開発した、ネットワーク上の複数のUNIXコンピュータ間でユーザー情報を共有するシステムである。

NISを使うことで、サーバからネットワーク上のマシンにユーザー情報(パスワードやユーザーIDなど)を提供でき、ユーザーの

一元管理ができる。NFSを使う時には必要なサービスとなる。

4. PCクラスタの構築

ここでは、実際に構築したネットワーク内のPCクラスタの環境について概説する。まず、構築する際のOSとしては、全て、Fedora Core1(LINUX)を採用し、8台のPCで構成している。構築したPCクラスタの実物写真を図1に、そのシステム構成を図2に示す。



図1 PCクラスタの実物写真

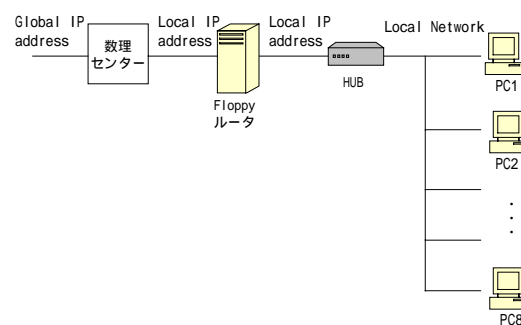


図2 PCクラスタシステム構成図

Performance Evaluation of the PC Cluster -Part1 Construction of the PC Cluster-

Takayuki MATUDA, Daisi TAKAHASI and Kazuhiko KAKUDA

4-1. PCクラスタ

PCクラスタ (clusterの意味はブドウの房) とは、ネットワークに単一で稼動するコンピュータの集まりが接続された状態で、一つの計算資源として使用可能な並列もしくは分散システムである。その際にNFSサーバを設定し、クラスタ内でネットワーク情報を共有するためのデータ管理としてNISの設定も行う。また、MPIの設定も必要となる。

これらのサービスを設定する事によって、PCクラスタとして、並列処理を行う事ができる。なお、PCクラスタの型・構造は、Beowulf型(ベオウルフ)・スクリーンドホスト構造を採用している。

4-2. ベオウルフ型

ベオウルフ型とは特定のシステムを表したのではなく、市販コンピュータ+ネットワーク+TCP/IP通信+フリーソフトウェアで構築されたクラスタシステムのことである。OSはLinux、並列処理プログラミング環境はMPICHやLAM(Local Area Multicomputer)などで構成されている。

4-3. スクリーンドホスト構造

この構造は、1台のルータによりパケットのフィルタリングを行い、内部ネットワークのサーバのみを外部に公開する方法であるが、本研究ではサーバを公開させていない。

4-4. Floppyルータ⁴⁾

ファイアウォール機能が有り、1枚のフロッピーだけで起動ができるルータである。次の特徴と欠点を持っている。

・特徴

(1)メモリ上で動作するためハードディスクは必要ない。

(2)インストールや設定が簡単である。

(3)不意に電源が切れても支障はない。

・欠点

情報 (例えば、DHCPを設定した場合のIPアドレス情報) はメモリ上に保存されるため、ルータの再起動時でそれらの情報は失われる。

5. 使用機器

8台のPCクラスタを構築する際に使用した機器は以下のとおりである。

(1)Floppyルータ

(2)16ポートスイッチングHUB

(3)8CPU切り替え機

PC8台を、ディスプレイ・マウス・キーボードをそれぞれ一つで操作するために

必要である。

(4)8台のPC

PCの性能を表1に示す。

表1 PCの性能

	COMPAQ (6台)	カスタマイズ PC (2台)
OS kernel	Fedora Core 1 2.4.22	Fedora Core 1 2.4.22
CPU	Pentium (600MHz)	Celeron (2.4GHz)
メモリ	256MB	256MB
HDD	6GB	40GB

6. 並列計算ライブラリ

並列計算を行うに当たって、各CPUには一つのプロセスが起動される。並列計算はデータや仕事をCPUの数で分割して計算するわけだから、それぞれのCPUで動いているプロセスに相互関係がある。そこで各プロセス間でデータの交換を行う。あるプロセスのメモリ空間から別のプロセスのメモリ空間へのデータ転送のことをメッセージ・パッシングと言い、PCクラスタではこれを用いて計算を行う。

この通信の仕組みに、通常は、メッセージパッシング・ライブラリと呼ばれる予め用意されたライブラリを使用する。これが並列計算ライブラリである。

メッセージパッシング・ライブラリとしてPVM(Parallel Virtual Machine)とMPI(Message Passing Interface)の二つが代表的であるが、本研究ではMPIを採用した。

7. おわりに

NFSやNISなどの機能を利用し、8台のPCによるクラスタシステムを構築してきた。その際、外部ネットワークとのフィルタリングにはファイアウォール機能などを有しているFloppyルータ(floppyfw)を採用した。

その2では構築したPCクラスタの計算能力を検証している。

「参考文献」

1) <http://www.phys.aoyama.ac.jp/~naoya/ARK/ark000.html>

2) <http://e-words.jp/w/NFS.html>

3) <http://e-words.jp/w/NIS.html>

4) 角田和彦, スクリーンドサブネットワーク内でのPCクラスタの構築, 学術講演会, 2002