

# 仮想 IP アドレスを用いたプライベートネットワーク内のノードへの透過的アクセス

山本 剛之† 建部 修見†

## 1. はじめに

一般家庭における PC の普及率はここ十数年で飛躍的に増加している。そしてこれらの PC の基本性能もまた急速に進化している。しかし一般ユーザが PC を使う主な目的はホームページ閲覧や文章作成といった、比較的マシンに対する負荷が小さいものである。

そこでこれらの未使用の資源を有効利用することを考える。ブロードバンド接続の普及により、これらの資源にはインターネットを通じてアクセス可能である。既に実用化されているものとして、CPU に関しては BOINC といったボランティアコンピューティング、ストレージに関しては Gnutella などの P2P ファイル共有アプリケーションが代表的である。

一般家庭では通常外部ネットワークからの不正アクセスを防止するため、NAT などを利用し、家庭内ではプライベートアドレスのネットワークを構築している。NAT 内の PC は外部ネットワークからの接続要求を受けられない。NAT の中から外部へ接続要求を出すことはできるが、このように通信を開始できるのは一方からのみとなってしまう。サーバのようにクライアントからの接続を受けるノードとして使用したいときにこの制約はアプリケーションの設計に大きな影響を及ぼす。

本研究では既存の任意のネットワークでサーバを無変更で実行するための汎用的な機構を検討する。具体的には NAT などによるプライベートネットワーク内の PC をサーバとして外部に公開するためのものである。プライベートネットワークと外部ネットワーク間の通信制約を UPnP などのネットワーク設定なしにソフトウェアのみで解決する。クライアントはライブラリの connect() 部分を修正しプライベートネットワークのサーバに接続させるが、サーバのコードに変更は一切加えない。

## 2. 設計・実装

### 2.1 基本指針

元のアプリケーションのサーバを変更せずプライ

ベートネットワーク内で (かつ外部からの接続を受けられるように) 使うため、PSA (Private Server Agent) と呼ばれるデーモンプロセスを導入しサーバと連係動作させる。プライベートネットワークの内外で接続を確立するには内から接続を開始するしかないがサーバにこのような機能はない。そこでプライベートネットワークの中から外に向けて接続を開始する動作を PSA に担当させる。PSA は外部へ公開したいサーバが存在する各プライベートネットワーク内に設置する。

PSA の接続先は、PSA とクライアントの双方から接続を受けられるような環境に設置された中継ノードである。中継ノードはプライベートネットワークの内外の接続を仲介するサーバである。中継ノードは PSA との接続を常に維持する。PSA はクライアントが通信したいノードに直接接続できるため、クライアントは中継ノードと PSA を経由して目的のノードへアクセスできる。但しクライアントは最初に中継ノードに接続を行うよう動作に変更を加える必要がある。

### 2.2 仮想 IP アドレスの導入

外部ネットワークからはプライベートネットワーク内のノードを特定する ID が存在しない。そこで仮想 IP アドレスを導入し、プライベートネットワーク内の各ノードに割り当てる。

これにはまず、PSA は中継ノード接続後に同ネットワーク内の各ノードのプライベート IP アドレスを送信する。中継ノードは PSA の ID (PSA のディスクリプタ番号等) と受信したプライベート IP アドレスの組から一意に定まるような仮想 IP アドレスをクラス E のアドレス帯から選択し、各ノードに割り振る。

### 2.3 クライアントのライブラリ修正

クライアントが前節で導入した仮想 IP アドレスに対して接続することを可能とするために、クライアントの通信ライブラリの connect() を以下のような機能を持つように置換する。

接続先の IP アドレスがクラス E のものであるかの判定を行い、もしそうでなければ直接その IP アドレスに接続を行う。クラス E のものであれば、中継ノードに対して接続を行い、中継ノードが持つ PSA との接続を利用してプライベートネットワーク内のノードと通信を行う。ライブラリの修正であるため、ユーザ

† 筑波大学大学院 システム情報工学研究科

は仮想 IP アドレスを指定するだけで透過的にプライベートネットワーク内へアクセスできる。

また、プライベートネットワークのノードと通信するにはクライアントのネットワーク状態に即した接続形態をとるようになっている。仮にクライアントがグローバル環境に設置されていれば、PSA からクライアントへ接続を行わせる。こうすることでクライアントは PSA をリレーサーバとして目的のノードと通信できる。もしクライアントもプライベートネットワーク内であれば、クライアントは中継ノードと PSA のふたつをリレーサーバとして目的のノードと通信する。

### 3. 性能評価

#### 3.1 接続確立時間

本機構を利用した際のプライベートネットワーク内のノードとの接続確立時間を測定した。実験はローカルネットワーク上で行った。クライアントが PSA からの接続を受けれる環境にある場合は 1.693ms, 受けれない場合は 2,005.571ms となった。後者が約 2,000ms も多くなっているのは、現実装においてクライアントのネットワーク状態を調べるために PSA がクライアントへ 2,000ms の接続試行を行うためである。この時間内に接続が成功しなければクライアントもまた別のプライベートネットワークにいるものとし、クライアントは中継ノードと PSA を通して通信を行う。しかし、最初から中継ノードと PSA によるリレーの方法を取ることを仮定すれば接続試行は削減できる。

#### 3.2 スループット

本機構を Iperf<sup>2)</sup> に適用し、InTrigger<sup>3)</sup> 上でクライアント (hongo クラスタ) とサーバ (imade クラスタ) 間でのスループットを測定した。imade クラスタはヘッドノード以外はすべてプライベートネットワーク内にある。これらのクラスタ間の通信遅延は 13.521ms であった。実験結果を図 1 に示す。

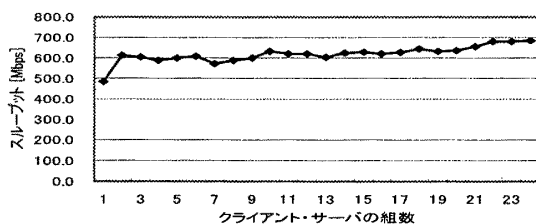


図 1 InTrigger 上でのスループット

InTrigger のネットワークは広域のギガビットであるが、共有回線であり複数のユーザが利用していることを考慮すれば本実験の最大スループットである 685.1Mbps という値は妥当といえる。実際、InTrigger でグローバル IP アドレスを持つノード同士で Iperf を用いてスループットを測ってもギガビットの性能が

フルに出ることはほとんどない。その一例として、予備実験で行った hongo と chiba クラスタ間の 16 並列での測定の結果は平均 510Mbps であった。今回の実験では 24 組のクライアントとサーバを用いたが台数を増やしてもあまり効果がないことがわかる。これはノード性能よりもネットワーク速度がボトルネックになっているためである。今回は、PSA と中継サーバを一組準備すれば十分であったが、より広帯域、高遅延の場合はそれぞれ複数準備することにより対応できる。

#### 3.3 Gfarm

本機構が実用的なアプリケーションでも動作することを示すため、Gfarm ファイルシステム<sup>1)</sup> への適用を試みた。Gfarm においてディスクを提供するノードであるファイルシステムノードをプライベートネットワークに設置し、外部ネットワークからアクセスする。Gfarm 上でファイル転送を行いそのスループットを測定した。使用したルータは BEFSR41C-JPV3 で、事前に測ったこのルータの最大スループットは 68.5Mbps であった。実験の結果、クライアントの状態に関わらず 67.9Mbps の速度でファイル転送が行われた。これはルータ性能の 99% の値であるため、このルータを使用したローカル環境では本機構を用いても性能を出し切れていることがわかる。また、本機構によりプライベートネットワーク内の PC をファイルサーバとして Gfarm ファイルシステムで束ねることに成功した。

### 4. おわりに

本稿では仮想 IP アドレスを用いてプライベートネットワーク内のノードへ透過的にアクセスするための機構について述べた。接続確立時間はクライアントが PSA から接続不可であるときは長い時間がかかってしまっているが、最初から中継ノードと PSA によるリレーを仮定することで短縮できる。また Gfarm ファイルシステムにおいては本機構によりプライベートネットワーク内のファイルシステムノードをファイルサーバとして束ねることに成功した。

謝辞 本研究の一部は、文部科学省科学研究費補助金特定領域研究課題番号 19024009 および基盤研究 (A) 課題番号 17200002 による。

### 参考文献

- 1) 建部 修見, 森田 洋平, 松岡 聡, 関口 智嗣, 曾田 哲之, 「ペタバイトスケールデータインテンシブコンピューティングのための Grid Datafarm アーキテクチャ」, 情報処理学会論文誌: ハイパフォーマンスコンピューティングシステム, 情報処理学会, Vol.43, No.SIG 6 (HPS 5), pp.184-195, (2002).
- 2) Iperf. <http://dast.nlanr.net/Projects/Iperf>.
- 3) InTrigger Platform. <https://www.logos.ic.u-tokyo.ac.jp/intrigger/registration/>