

フリースケールネットワーク方式の予備評価 仮想アドレス使用量の予測

Preliminary Evaluation of the Free-Scale Network Architecture

片山 忠和¹ 廣津 登志夫² 福田 健介³ 明石 修⁴ 菅原 俊治⁵ 村上 健一郎⁶

概要

再帰的なネットワークの拡張が可能な Free Scale Network(FSN) 方式では、新しい形式の FSN 実アドレスを IPv4 形式の仮想アドレスに変換して使用する。アドレス変換は Realm Gateway で処理されるが、この Realm Gateway を実現するためには、どの程度の大きさの記憶領域を用意すればよいかといったアドレス変換に関する資源情報の予測が必要である。本研究では、ネットワーク内外にわたる通信を解析することで、ホスト毎の通信傾向を分析し、変換に必要な資源の見積もりを行った。これによって Realm Gateway を実現するための重要な基礎データが得られた。

1 はじめに

現在、インターネットの発展と共に利用者が増加し、全世界でインターネットに接続する需要が高まっている。新しいホストがインターネットを経由した通信を行うとき、ホストに未使用のグローバル IP アドレスを割り当てる必要がある。しかし、インターネットが拡大するにしたがい、現在使われている未配分の IPv4 アドレスの在庫が少なくなり、新たに確保することが困難になっている。IPv4 アドレスの在庫が枯渇すると、増加し続けるインターネット利用者の需要に対応できなくなったり、インターネット上のサービスの新設や拡大ができなくなる。

Free Scale Network(以下 FSN と略す) 方式 [1] はこの問題に対応し、従来の IPv4 と

の互換性を維持したままアドレス空間の拡張を実現する通信方式である。FSN 方式では、Realm と呼ばれるネットワークの構成単位の内部では従来のネットワークプロトコルである IPv4 を使用する。Realm 間は Realm Gateway(以下 RG と略す) によって接続され、ネットワークは階層的に拡張される。ホストは階層的に接続された Realm アドレスにホスト自身の IPv4 アドレスを接続したグローバルかつユニークな FSN 実アドレスを持つ。一方、異なる Realm に所属しているホスト間で通信を行う際には、RG が FSN 実アドレスを IPv4 形式のアドレス空間の一部にマッピングするため、各ホストは従来の IPv4 プロトコルを用いて通信することができる。RG は Realm 外のホストの FSN 実アドレスとそれを Realm 内で参照する仮想アドレスとの対を変換情報として保持し、Realm 内のホスト毎に独立した変換テーブルでこれを管理している。そして、頻繁に使用する変換情報は Translation Look-ahead Buffer(以下 TLB と略す) と呼

¹豊橋技術科学大学 情報工学専攻

²豊橋技術科学大学

³国立情報学研究所

⁴NTT 未来ねっと研究所

⁵早稲田大学

⁶法政大学

ばれる高速な記憶領域に保持されるという設計になっている。

このような FSN 方式を実現するための RG を実装しようとするとき、アドレス変換機構に関する各種パラメータがわからないと実現可能であるかどうかすらわからない。例えば、TLB 上に変換情報が収まらなくなった場合、長期間使用されていない変換情報は低速な記憶領域に置き換えられ、再度参照されると検索効率が低下するが、アクティブな通信に対する影響を抑えるためには、どの程度の TLB サイズが必要になるかといった情報が必要である。

そこで本研究では、Realm 内の各ホストが通信する Realm 外ホスト数の時間経過による変化、継続中である通信のパケット送信間隔の傾向、Realm 内の各ホストが通信する Realm 外ホスト数の合計数 (同一 Realm 外ホストの重複を含む) の時間経過による変化を調査し、アドレス変換の為に必要となる資源量の予測を行った。

2 従来のアドレス拡張方式

IPv4 と互換性を持つネットワークの拡張方式がいくつか提案されている。ここでは IPv4+4[2]、IPNL[3] について説明する。

IPv4+4 では、各ホストはゲートウェイの IPv4 アドレスとホスト自身の IPv4 アドレスを接続したアドレスを持つ。ホスト自身が所属するプライベート空間内ではホストの IPv4 アドレスを使用し、グローバル空間ではゲートウェイのグローバル IPv4 アドレスを使用する。グローバル空間の通信では、あらかじめ指定された経路によってパケット転送が行われており、経路の対称性のために経路の切り替えができない。

IPNL では、各ホストはインターネットに直接接続している Middle Router の Middle Realm IP Address、自身が所属する Realm

の Realm Number、ホスト自身の IPv4 アドレスの 3 つを接続したアドレスを持ち、識別するアドレスが経路によって変化する。障害時に別経路を選択して対応するとき、ホストのアドレスが変わることになり、継続中の接続が切断されてしまう可能性がある。

3 Free Scale Network 方式

FSN 方式では、従来の IPv4 で通信しているネットワークを Realm と呼び、Realm 同士を RG で相互接続させる。これを、階層的に何層にも繰り返すことで再帰的なネットワークの拡張を可能にする。

FSN 方式には以下のような特徴がある。

- 再帰的なネットワークの拡張
- 従来の IPv4 との互換性
- 下層の Realm がショートカットを持つ
- 多重仮想アドレス空間方式

FSN 方式では、FQDN(Fully Qualified Domain Name) に対する DNS のクエリをトリガーとして、FSN 実アドレスを IPv4 形式の仮想アドレス空間にマッピングすることで、従来のプロトコルで通信することを可能としている。このアドレス変換においては、多重アドレス空間方式と呼ばれる手法を用いており、Realm 内のホスト毎に仮想アドレス空間が用意される。パケットを転送する経路は、RG が持つ Realm の経路情報が書かれている Realm 経路テーブルから最適な経路が選択される。FSN 方式では、追加するネットワーク機器は RG だけであり、従来の機器やホストを変更する必要がない。これらの特徴から、FSN 方式によって展開と運用の費用が安価で、互換性と透過性のあるネットワークの拡張が可能になる。図 1 に FSN 方式の概要を示す。以降、個々の技術について説明する。

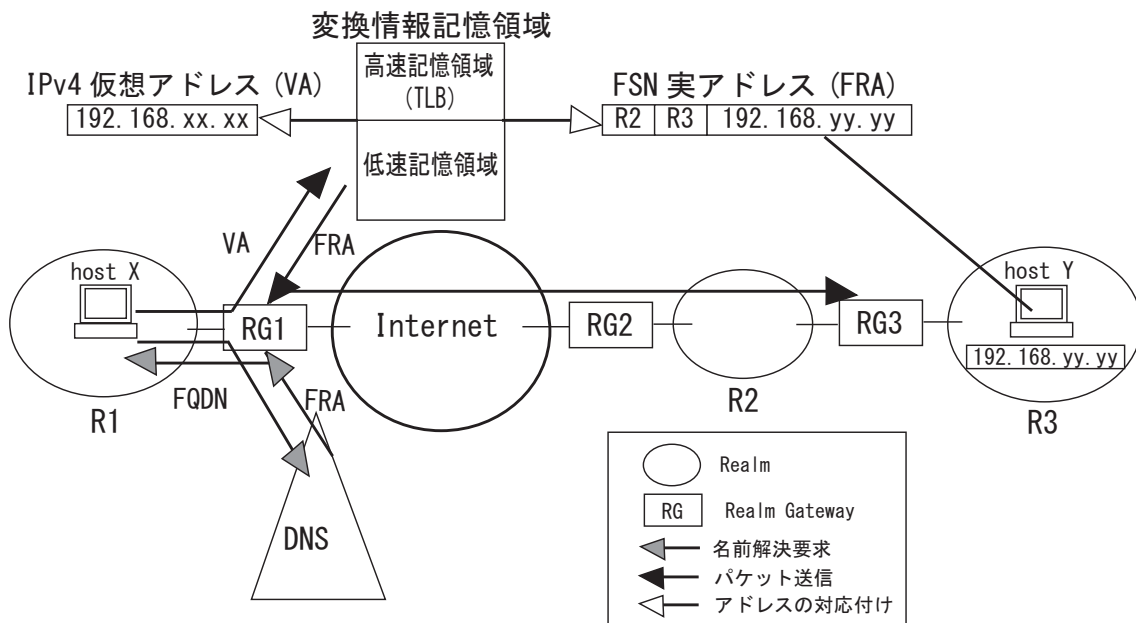


図 1: FSN 方式のアドレス変換

3.1 Realm アドレス

各 Realm は固有の Realm アドレスを持つ。インターネットに直接接続しているトップレベルの Realm の Realm アドレスはグローバル IPv4 アドレスである。図 1 では、Realm R1 と R2 がトップレベルの Realm である。トップレベルより下層の Realm は直属の親 Realm 内でユニークな識別子を持ち、Realm アドレスは階層的に表記される。具体的には、トップレベルより下層の Realm の Realm アドレスはトップレベルから自身までの経路上に存在する Realm の識別子を接続したものに自身の識別子を接続したものとなる。

また、下層の Realm は自身に付けられた Realm アドレスの別名としてグローバル IPv4 アドレスを持つことも可能である。別名のグローバル IPv4 アドレスをトップレベルの経路情報に載せておくことによって、親 Realm に障害が発生しても別経路を選択することが可能となる。

3.2 FSN 実アドレス

Realm 内の各ホストは自身の IPv4 アドレスと所属する Realm の Realm アドレスを接続した FSN 実アドレスを持つ。各層においてユニークな識別子と IPv4 アドレスを階層的に接続することによって、FSN 実アドレスはグローバルかつユニークなものになる。

例えば図 1 では、トップレベルの Realm R2 の Realm アドレスである「R2」とその子 Realm である Realm R3 の識別子「R3」を接続した「R2 R3」が Realm R3 の Realm アドレスとなる。Realm R3 に所属しているホスト Y の FSN 実アドレスは、Realm R3 の Realm アドレスと自身の IPv4 アドレス「192.168.yy.yy」を接続した「R2 R3 192.168.yy.yy」となる。この FSN 実アドレスが FSN 方式においてホスト Y を識別するものとなる。

3.3 FQDN からのアドレス変換

Realm 外のホストへパケットの送信を開始するとき、Realm 内の送信元ホストは宛先ホストの FQDN を使用する。FQDN をアドレスに変換する要求は RG へ送られ、RG は自身の DNS リゾルバの機能を用いて、DNS サーバから当該 FQDN の FSN 実アドレスを取得する。そして、以下のいずれかの操作が実行される。

1. FSN 実アドレスが透過的に Realm 外へパケットを中継すべきアドレス空間にある場合
そのまま送信元ホストに返す
2. それ以外の場合
送信元ホストの変換テーブルから変換情報を調査する
 - (a) 既に FSN 実アドレスに対して仮想アドレスが割り当てられている場合
その仮想アドレスをアドレス解決要求の応答として返す
 - (b) 割り当てられていない場合
未使用の仮想アドレスを割り当てて、それを返す

3.4 FSN 実アドレスと仮想アドレスの変換

FSN 実アドレスは新しい形式のアドレスであり、通常のホストはそのままではこれを扱うことができない。そこで、従来の IPv4 との互換性を実現するために、FSN 実アドレスを IPv4 形式の仮想アドレス空間の一部にマッピングする。この仮想アドレス空間には IPv4 形式のプライベートアドレスを使用する。そして、通常のホストは FSN 実アドレスに対応付けられた仮想アドレスを使用して通信を行う。

3.5 変換テーブルと Translation Look-ahead Buffer

FSN 実アドレスと仮想アドレスとの変換情報は RG 上に用意された変換テーブルが保持する。特定のホストの異常な振る舞いが他に波及することを防ぐために、変換情報の記載された変換テーブルは Realm 内のホスト毎に用意される。

RG は自身を通過するパケットに含まれるアドレスに対応した変換情報を変換テーブルから検索または追加し、パケット内のアドレスの書き換えを行い、書き換えられたアドレスによって転送を行う。RG が変換情報を高速に検索することを可能にするために、高速な記憶領域である TLB を用意しこれに変換情報を保存していく。時間が経過すると TLB 上に変換情報が収まりきらなくなる場合がある。溢れた変換情報を捨ててしまうと継続している通信が切断される恐れがあるので、TLB 上から溢れたものを低速な記憶領域に保存する。これらの様子を図 2 に示す。

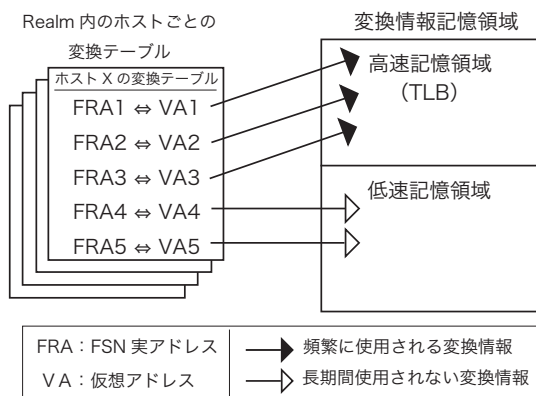


図 2: 保存された変換テーブル上の変換情報

3.6 パケットの送信

パケットの送信手順を図1を用いて説明する。ホストXがホストYにパケットを送信する場合、ホストYのFQDNを使用して通信を開始し、FQDNの名前解決要求をRG1へ送る。RG1は自身のDNSによってFQDNからFSN実アドレス「R2 R3 192.168.yy.yy」に変換し、仮想アドレス「192.168.xx.xx」を対応させ、これをホストXに返す。ホストXはこの仮想アドレスを宛先アドレスとしてパケットを送信する。パケットがRG1を通過するとき、RG1はホストXの変換テーブルから仮想アドレスに対応するFSN実アドレスを取得し、これをパケットに付加されたシムヘッダ部の宛先アドレスとして転送する。

4 評価事項

本研究の目的はRGの設計についての指針を得ることである。RGはRealm内の各ホスト毎に仮想アドレス空間を用意するため、各ホストが通信するRealm外ホスト数に対応できるだけの仮想アドレス空間が必要になる。そこで、Realm内の各ホストが通信するRealm外ホスト数の傾向から、ホストあたりにどの程度の大きさの仮想アドレス空間が必要となるか予測する。

通信途中の変換情報がTLB上から低速な記憶領域に置き換えられたとき、この情報が再度参照されるとTLB上に保存されている場合よりも検索時間が長くなってしまい、通信効率の低下を引き起こす。これを防ぐためには、通信が終了するまで変換情報が出来るだけTLBから追い出されないようなTLBサイズを設定する必要がある。これを実現するために、TLBのサイズがどの程度必要になるか予測する。

最後に、RGの実現可能性を決める最も重要な点は、Realm内外の全ての通信に対する仮想アドレスへの変換エントリを保持できるかである。このためには、多数のホストの通信に対する外部アドレスの出現傾向をみる必要がある。

以上のことから、本研究では実際の通信トラフィックを解析して、以下の項目について予測を行った。

- ホストあたりに用意すべき仮想アドレス空間の大きさ
- 通信に悪影響を与えないだけのTLBの大きさ
- Realm Gatewayが保持する変換情報の合計数

5 仮想アドレス使用量の予測

仮想アドレスの使用量を予測するために、豊橋技術科学大学の対外接続セグメントにおける、大学内と大学外間の実際の通信データを用いて、仮想アドレスの出現に関する傾向の調査を行った。ここでは、「学内側のネットワークをRealmの内側」、「学外側のネットワークをRealmの外側」と仮定し、対外接続ルータをRGに対応させた。対外接続ルータを通過する学内から学外へ送信されるパケットについて、アドレス変換が行われると想定して解析した。

今回、Realm内に相当するホストの数1,299台についての、42時間のデータを評価に使用した。また、TCPセッションの総数は3,339,114であった。

5.1 ホストあたりに用意すべき仮想アドレス空間の大きさ

Realm内ホストあたりの通信するRealm外ホスト数が時間経過によってどの程度増

加するか調査した．これは，ある期間内に Realm 内の各ホストによって，使用される仮想アドレスの量に相当し，これらの分布から，ホスト毎に必要な仮想アドレス空間を予測する．

設定した単位時間内に各ホストが使用する Realm 外のホストのアドレス数を集計し，結果をホストあたりの外部アドレス数で順位付けして並び変えを行い，最高順位が1となるように正規化した．集計の単位時間を1時間とした結果を図3に示す．この結果から，非常に大きなアドレスを消費するホストは少数で，99%以下の大部分は数百個程度のアドレスしか使用しないことが分かる．

更に，時間経過による外部アドレス数の増加傾向を調査するために，正規化された変換情報数の順位の0.5，つまり50%の順位にあたるものの数と98%，99%，100%（最大）の順位にあたる数を調査した．この順位の位置は図3に書き込まれた縦線の位置にあたる．集計する単位時間は1分，5分，10分，30分，60分，120分，180分，240分，360分に設定し，各順位にあたる外部アドレス数の最少，平均，最多を求めた結果を図4に示す．大部分の99%以下のホストの外部アドレス使用量は時間が経過しても小さく，99%に位置するホストの Realm 外アドレス数の平均値は，1時間で653，6時間で1,675である．

5.2 通信に悪影響を与えないだけの TLB の大きさ

通信データ内の Realm 内の各ホストが Realm 外のホストへ送信する TCP パケットに対して，一連の TCP セッションが終了するまでのパケット送信間隔の傾向を調査した．ここで，TLB サイズの所要量の見積もりの為には，ある TLB サイズを仮定して

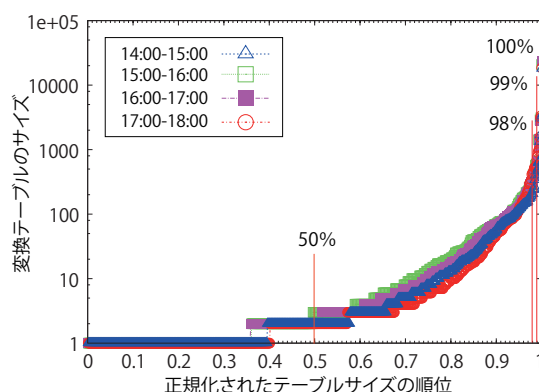


図 3: 各ホストが使用する外部アドレス数

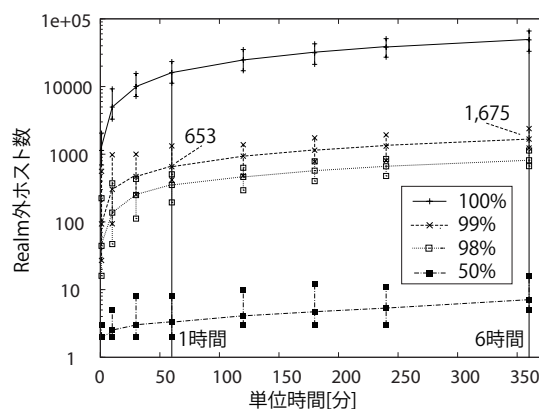


図 4: 外部アドレス数の増加傾向

アドレスの利用傾向で TLB 上から継続中の通信に使用されている変換情報が溢れてしまう TLB ミスがどの程度発生するかをみるという方法が考えられる．しかし，ここでは所要量と TLB ミスの発生に関する傾向を把握することを重視して，一定期間使用されない変換情報はタイムアウトして TLB から追い出されると仮定し，タイムアウトに対するミスの発生と必要となる TLB サイズの両方をみるようにした．

ここでは，一連の処理のまとめりとして TCP に着目し，Realm 内のホストが Realm 外のホストへ送信している TCP パケットの送信間隔が設定した各タイムアウト時間よりも長い割合を調査した．これは，置き換えによって Realm 内の各ホストの通信と TCP

セッションに影響が及ぶ割合にあたる。この結果を図5に示す。TLB ミスの割合を0.01%に抑えるならば、図5より、タイムアウト時間を30分間と設定することに相当する。すなわち、30分間に使用されるアドレス数を保存可能なTLBを用意することになる。

次に、保存する具体的な数を予測するために、Realm内の各ホストがRealm外のホストと通信するとき使用されるアドレス数を集計し、Realm内の全ホストについての総数が時間経過によってどの程度増加するか調査した。Realm内のホスト毎に変換テーブルが用意されていることから、これらを全て記憶領域上に保存する場合、必要となる記憶領域のサイズはこのRealm内の各ホストの変換テーブルサイズの合計となる。

設定した単位時間内に各ホストが使用したRealm外ホストのアドレス数の総数を調査した。集計する単位時間は1分、5分、10分、30分、60分、120分、180分、240分、360分に設定した。各単位時間内の総数の最少、平均、最多を求めた結果を図6に示す。この結果から、30分間にRealm全体のホストが使用するアドレス総数が平均で331,700個になると予測でき、これらを保存可能なTLBの大きさがTLBミスの割合を0.01%に抑えられる為に必要なものとなる。そして、Realm内に相当するホスト数は1,299台であるため、単純に平均すると1ホストあたり約255個となる。

5.3 Realm Gatewayが保持する変換情報の合計数

5.2節のRealm内の全ホストについての外側ホストのアドレス使用量の総数の調査結果より、ある期間内にRealm内の全ホストによって、使用される変換情報を全てRGの記憶領域上に保存させる場合、必要とな

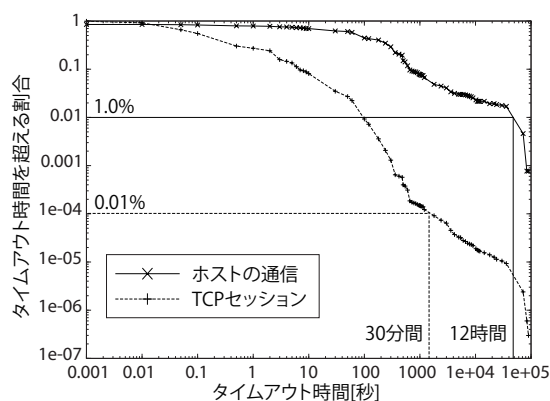


図5: タイムアウト時間による通信への影響

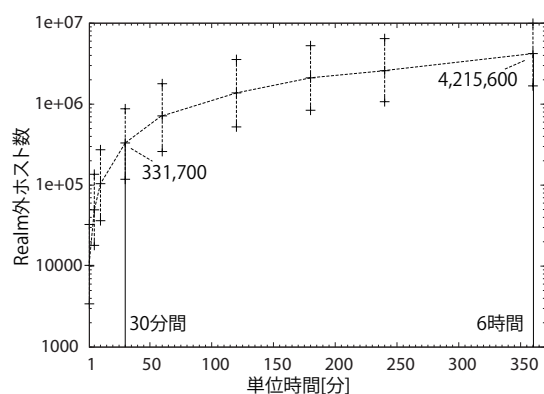


図6: 各ホストが使用する Realm 外のアドレス数の総数

る記憶領域の大きさがどの程度になるか予測した。図6の結果より、1時間以上では増加の割合が小さくなっている。より長い6時間でも増加し続けているため、保存する数の上限はこれ以上であるといえる。図5より、99%のホストが使用するアドレス数を保存したとき、1%のものが超えるタイムアウト時間は約12時間となる。今回は、6時間以上のアドレス数を集計していないため、6時間以上では4時間から6時間の間の増加率のまま増加すると仮定すると、6時間の平均値4,215,600から12時間まで増加した9,090,336が99%のホストに必要な記憶領域の大きさの平均値となる。今回の評価に使用した通信データに含まれる Realm

内に相当するホスト数は 1,299 台で、その 99 % である 1,286 台に対する数であり、単純に平均すると 1 ホストあたり約 7,068 個となる。

6 考察

5.1 節の調査結果から、99 % のホストに対して満足させられる仮想アドレス空間の割り当てで十分としたとき、99 % に位置するホストの外部アドレス数を基準にする。そのとき、平均値は 60 分で 653, 360 分で 1,675 であり、ホストあたりにクラス C アドレス 7 本分の仮想アドレス空間を用意すれば十分なものとなる。このことから、FSN 方式では、Realm 内の各々のホストに対して変換テーブル、すなわち仮想アドレス空間を用意することから、仮想アドレス空間の大きさという観点では FSN は十分に実現可能であると言える。

次に必要となる記憶領域の具体的なサイズを予測する。変換情報のサイズは Realm 外のホストが所属している Realm の階層の深さによって変化する。ここでは、インターネットを 1 層目とし、3 層目に位置する Realm に所属しているホストの FSN 実アドレスに対する仮想アドレスの対を 1 つの変換情報サイズとして考察する。これは、Realm アドレス (4[Bytes]+4[Bytes]) とホストアドレス (4[Bytes]) を接続した FSN 実アドレスに対応する仮想アドレス (4[Bytes]) の組み合わせで、合計サイズは 16[Bytes] となる。Realm 内の 99 % のホストが使用するアドレスを保存するために必要な記憶領域は、1 台あたり約 7,068 と予測した。そのとき、クラス C アドレスを全て使用しているネットワークでは約 28.9[MB] の記憶領域が必要となり、クラス B では約 7.4[GB] となる。更に、TLB ミス率を 0.01 % に抑えるとき、ホストあたり約 255 個の変換情報を保存可能な TLB が必要

であると予測した。必要な TLB サイズを同様に計算すると、クラス C では約 1.0[MB]、クラス B では約 267.4[MB] となる。これらを、現行のサーバマシン等で実現した場合、TLB を主記憶装置とし、それから溢れたものを二次記憶装置に保存すれば十分実現可能である。

7 まとめ

本研究で、実際の組織内外のトラフィックを解析することで FSN 方式で必要となる変換アドレス数の予測を試みた。ここで調査したのは一つのネットワークの特性であるためにまだ初期的な解析結果である。TLB 上に記憶する変換情報数は Realm 内のホスト数に依存すると考えられるので、今後は Realm の大きさに対するスケーラビリティの観点から更なる調査を行いたい。

参考文献

- [1] 村上健一郎, 菅原俊治, 明石修, 福田健介, 廣津登志夫, "フリースケールネットワーク方式", 情報処理学会研究報告, No.104, pp 81-88, January 2007.
- [2] Zoltan Turanyi, Andras Valko, Andrew Campbell, "4+4: an architecture for evolving the Internet address space back toward transparency", ACM SIGCOMM Computer Communication Review, Vol. 33, No 5, pp 43-54, October 2003.
- [3] P.Francis, R.Gummadi, "IPNL:A NAT-Extended Internet Architecture", SIGCOMM' 01, P69-80, August 2001.