proxy cache serverを用いたfirewall

一條 博
（株）ケンウッド 研究開発本部

http により情報の入手が簡単になったが、ネットワークへのアクセスの増加により、firewall にトラフィックが集中し情報の入手が遅くなる事がある．そのためキャッシュの機能を持った proxy server をリレー接続して firewall を構築し、情報入手の効率を改善を行った．本稿ではその firewall の構成と proxy cache server によるキャッシュの効果について述べる．

Firewall with relayed proxy cache server

Hiroshi Ichijo
KENWOOD corp R&D division

Nowadays, it becomes ease to collect information from "http", but information delay was be a problem due to much increasing access to network and traffic rush to firewall. Building up firewall with relayed proxy cache server improved efficiency of collecting information. We will mention this construction of firewall and cache effect.

1 はじめに

コンピュータの普及に伴い、その上で各種のアプリケーションの研究が行われた結果、まず実際に ftp の様なプロトコルが使用される様になった．つぎにネットワークを用いたＥメールで情報交換が行われ、だれにでもネットワークを使用する事ができる様になってきた．さらに http を用いた情報の交換が行われる様になると，その取り扱いが容易で有用な事が理解され、またネットワーク接続ができる PC が安価に入手できる様になった事もあり、メールの機能と共に一般化されてきた．企業等においてもネットワークの利用により、内部の情報の流通を円滑にし外部の情報の入手も容易なため、インントラネットとして社内情報の運営に活用されてきている．

ところで企業内で使用されているネットワークを、インターネットに接続するには，内部のネットワークのセキュリティを保つ配慮が必要になり、そのために firewall が使用される．しかし接続には以下の様な問題点が生じる．

- firewall でネットワークの帯域が最も狭くなり、アクセス速度が制限される．
- 内部からのインターネットの利用が増えた結果、インターネットへの接続速度にかかわらずアクセスに制限を受ける事がある．また経費の制約から必ずしも高速な接続を行う事はできず、情報の入手に制限が生じてしまう．
- イントラネットの内部は組織変更等により常に流動的であり、そのトラフィック状況の変化は大きいが、リアルタイム性が必要な http を効率よく利用できる必要がある．

これらの問題点は相反する面を持っているが、大量の情報を早くインターネットから入手する必要があ

り、その実現のためにproxy cache serverをリレー接続したfirewallを構築した.

2 企業におけるネットワーク

最近の様にネットワークの利用が一般化され、企業は外部への情報提供のためにホームページを設け情報入手の目的でインターネット接続を行い、また企業内での情報にも同様な方法を用いる様になってきている.ところが本来ネットワークは“どこにでも接続でき、希望する情報を手に入れる事ができる”という性格を持っているが、企業においては教育機関などとは異なり、外部だけではなく内部にすら公開できない情報もあり、積極的にネットワークを利用するには色々な障害を持っている.この問題を解決するには、まずネットワーク接続にセキュリティを設けることが必要とされ、その実現にfirewallが設けられる.

しかしネットワークの使用をはじめたきっかけや、実際に必要とされるネットワークサービスのプライオリティは部門により異なる事が多いが、一般的に企業で使用されているネットワークサービスは、メールとhttp ftpの利用が多く、これらのサービスで提供される情報の性格は以下の様に異なっており、その為にセキュリティの設けかたも異なってくる.

smtp

[1] 情報は任意のタイミングで送られてくる.
[2] 情報はリアルタイム性が要求されない.
[3] 情報は一般に私的なものでセキュリティが必要とされる.
[4] 情報は双方向にやりとりされる.

http ftp

[1] 要求した情報はなるべく早く獲得する必要がある.
[2] 欲しい時に早く多くの情報を入手する必要がある.
[3] 情報は一般的に公開されている.
[4] 情報は一般に一方向性である.

この様にsmtpとhttpやftpはほとんど逆の性格を持っているが、いずれの場合でもインターネットに接続しているため、不要な——特に悪意のある——ネットワークへの侵入の禁止は企業内の情報の漏洩を防ぐために必要とされる.

3 firewall

firewallを使用する事により、外部のネットワークからの侵入を防ぐ事ができる.一般のネットワークはTCP/IPによる接続を用い、以下の様な処置を行い“防火壁”とする事ができる.

- 通過するアプリケーションプロトコルの制御.
- 通過するユーザの認証.
- 通過するプロトコルのソースアドレス、ディスティネーションアドレスの制御.

実際にこの処置を行うために、インターネットへの接続経路を一つに制限し、要求されるセキュリティや必要なアプリケーションに応じて様々な構成が採られる.

ところでfirewallはネットワークのセキュリティーを守るために使用されるが、ネットワークの接続を制限する為にインターネットワークからの情報入手を妨げてしまう事もあり、特に一般に要望の多いhttp接続に大きな負荷をかけることもある.しかも構成によってはネットワークユーザに負担をかけてしまう事もあり、さらに最近の様に充分にネットワークの知識を持たないユーザが多い場合は、運営面で様々な問題が発生し内部からセキュリティを破壊する事態も発生しえる.そのために特に企業においてはfirewallにはセキュリティを高めるだけではなく、ユーザはその存在を感じる事なくインターネットへのアクセスができ、ネットワークの速度が低下してはならない等の性格が求められる.

3.1 proxy cache server

httpやftpによるネットワークへの要求がfirewallを越えるためにproxy serverがよく使用されるが、このサーバーは内部のネットワークからの要求を代行する機能を持っている.最近はproxy serverにキャッシュの機能を持たせ、さらにリレー接続を行う事でネットワークの有効利用を目指す試みが行わ

れている[1]．そこではproxy serverの持つ“代理”機能を用い、クライアントのみでなく他のproxy serverからの要求をまとめて目的のサーバーや他のproxy serverへ伝えている．ここで使用されるproxy serverはキャッシュの機能も持っているが、この機能はハードディスク等でもアクセス速度の改善に使用され，インターネットへの応用も同様な理由から行われている．

ところでfirewallはセキュリティを守るためにインターネットへの接続ではもっともせまくなる部分であり、内部のネットワークからのアクセス速度の向上のネックの一つとなっている．その様なシステムのインターネット接続にproxy serverのリレー構造を使用しアクセス速度の改善ができないかと考えた．

3.2 proxy serverの評価

実際にキャッシュの機能を持ったproxy serverをfirewallへ組み込む前に、具体的にキャッシュの機能の効果を調査するため、図1の様な“暗室ネットワーク”と名づけたproxy cache serverでクライアントとサーバーの間を中継したトラフィックが少ないネットワークを構築し、webstone[2]を使用し条件を変えてサーバーの動作の比較を行った．webstoneはhttp serverを評価するプログラムであるが、proxyを使用した状態での測定も行う事ができる．

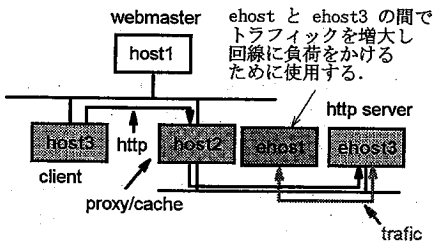


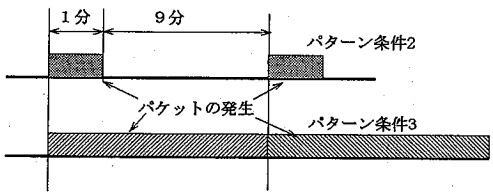
図1: 評価用ネットワークの構成

測定は、全くproxy cache serverを使用しない場合と、使用した場合にわけ、キャッシュ用のメモリーの大きさを変え、図1のehost3とhost2を含むネットワークにehostを使用して以下の様な条件のトラフィックを与え行った．トラフィックは、netperf[3]を用いて図1におけるehostとehost3の間でIPパケットの転送を行い作成した．

条件1 ：
サーバーとクライアントの間に不要なトラフィックが無い状態．

条件2 ：
バースト状のトラフィックがある状態を想定し、ehost3とehost間に16Kbytesの大きさのパケットを１０分毎に一分間ながす．

条件3 ：
非常にネットワークが混み合っている状態を想定し、ehost3とehostの間に16Kbytesの大きさのパケットを連続してながす．



トラフィックの条件

図2: トラフィックのタイミング

図２の様な２つのパターンで16Kbytesの大きさのパケットをトラフィックとして与え、各ホストのハードウエアーと試験に使用したデータの大きさは以下の様になっている．

CPU ： Pentium 200Mhz

OS ： FreeBSD-2.2.2[4]

ネットワーク ： 10 BASE-T

データの大きさ ： 1,2,5,10,20,50,200,500kbytes

cacheのスワップの大きさ ： 100Mbytes

クライアント数 ： ２から３２まで２ずつ増加．

表１に測定で使用したキャッシュメモリーの大きさと、トラフィクの与え方の具体的な条件を示す．

表１：測定の条件

キャッシュ	測定条件	グラフ番号
なし	条件１	NO-1
	条件２	NO-2
	条件３	NO-3
１６Ｍ	条件１	NO-4
	条件２	NO-4
	条件３	NO-4
３２Ｍ	条件１	NO-5
	条件２	NO-5
	条件３	NO-5
６４Ｍ	条件１	NO-6
	条件２	NO-6
	条件３	NO-6

以上の様な条件のもとでキャッシュの使用によりサーバーからのデータ速度にどの様な効果が得られるかを調べた．図３の測定結果は、クライアントのプロセス数を変えた時のサーバーからデータの供給される速度を示しているが、これより以下の様な事が判断できる．

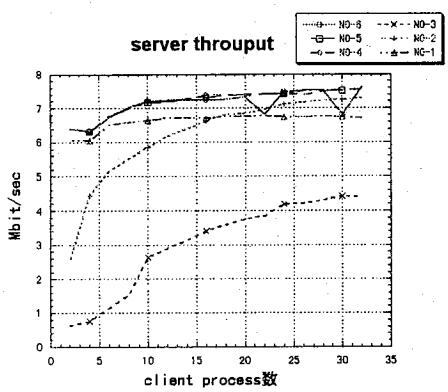


図 3: サーバースループットの変化

（１）　トラフィックの無い理想的環境ではキャッシュの効果はない．

（２）　連続したトラフィックが存在する場合は、データ転送速度は低下する．

（３）　バースト状のトラフィックが存在する場合は、データの量がトラフィックより少ない場合大きな影響をうけ、同じくらいの量の場合はあまり依存しない．

（４）　proxy cache server を使用した場合は、トラフィックの影響をうけない．

クライアントとサーバー間にトラフィックが存在するとデータ転送速度が低下するが、特に連続したトラフィックでは顕著である．しかし proxy cache server を使用することでその影響が少なくなる．

4　proxy server を用いた firewall

評価の結果より、proxy cache server を使用する事でトラフィックの影響が少なくなる事が確かめられた．

ところで一般に firewall では proxy cache server をインターネット側にのみ設ける事が多く、主にインターネットへの要求を減らす事を目的としているが、必ずしも充分な効果が得られておらず、また firewall の利用面についての改善にもなりえていない．そこで firewall のチョークポイントにあるホストを２つの proxy cache server ではさみ、以下の様な方針で firewall を構成する事とした．

［１］　扱うアプリケーションは smtp http ftp に制限し、ユーザがコンピュータを扱う知識が充分にない事を前提とし、http と ftp をなるべく簡単に firewall を越える事ができる様にする．

［２］　firewall のチョークポイントを通過するプロトコルをなるべく早く処理される様にし、もっとも帯域の狭い部分におけるパケットの滞留を減らし全体の処理を早くする．

［３］　キャッシュの効果を改善して firewall の内側のヒット率を高し、インターネット側へのアクセス数を減らす．

以上の様な条件を前提にし、firewall は図４に示す様な proxy cache server によるリレー構造を持った３階層の構成とした．以下に各階層の詳細を説明をする．

［一階層目］

ローカルなネットワークのクライアントから auto proxy 機能により選択される proxy server であり．

二階層目のproxyへプロトコルをリレーし、ローカルなhttp serverへの要求は直接行われる.

[二階層目]

インターネットへの接続のルートを設ける目的と、他のサブドメインとのhttpの処理を行う目的を持つproxy serverである.

[三階層目]

インターネットへの接続を行うproxyである.

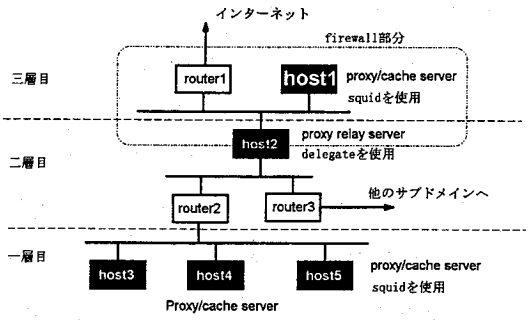


図 4: firewall周辺の構造

各階層にあるproxy serverは表2の様なアクセス制御を行いfirewallとして機能させている.

表2 : 各ホストのアクセス制御条件

host1	host2 からのみ proxy cache server の機能を使用でき、インターネット上の任意のホストへの接続を許可.
host2	host1 へのみ接続を許可し、イントラネット内のホストからのみ接続使用できる.
host3 4 5	イントラネット内のホストから接続し、host2 へ接続を行う.

4.1 構成

各ホストのハードウエアーとソフトウエアーの構成は次の様になっているが、proxy cache serverであるhost1 3 4 5は動作を改善するために、キャッシュ専用のディスクを設けさらにカーネルの改造を行っている.

表3 : 各ホストの構成

host	OS名 CPU	Cache Disk Memory	Proxy
host1	FreeBSD2.2 K6 300Mhz	4Gbytes 128M	squid-2.1
host2	IRIX 6.2 Indy R5000	- -	delegate
host3	FreeBSD2.2 K6 200Mhz	4Gbytes 64M	squid-2.1
host4	FreeBSD2.2 K6 200Mhz	4Gbytes 64M	squid-2.1
host5	FreeBSD2.2 K6 200Mhz	4Gbytes 64M	squid-2.1

4.2 autoproxy

autoproxy[5] はJava scriptによるプログラムをブラウザに供給し、その記述に従い自動的にproxy serverを選択する機能である. クライアントはこの機能を利用して、イントラネット側のproxy cache serverに接続するために以下の様な条件でネットワーク内の3つのproxy serverを選択する様にしている.

[1] 接続する目的のURLを示すキャラクターを単純加算し、proxy serverの数で除算した剰余により使用するproxy serverを指定する.

[2] 割り当てられたserverが動作を停止していた場合は、他のホストへ割り当てを変更する.

一般にホームページから得られるドキュメントは複数のファイルで構成されている事が多く、そこにアクセスした場合は同時に複数のファイルを処理する必要があるが、[1]の機能により処理をhost 3 4 5に分散してキャッシュの動作を改善を行い、また[2]の効果により全くインターネット接続ができなくなる可能性を減らし、さらに管理者はサーバーのメンテナンスが容易になる結果も得られる.

4.3 proxy cache serverの効果

現在はインターネットとの接続が64Kbpsと遅い速度であり、また内部にある他のサブドメインとの接続も考慮する必要がある. そのため複数種の

proxy server を使用してキャッシュとプロトコルのスイッチ機能を分離しておき、スイッチの部分にパケットの滞留がおこらない様に配慮した．具体的には squid[6] と delegate[7] でプロトコルのリレーを行い、squid でキャッシュを行い delegate で http のルートの切り替えをおこなっている．

ところでこの firewall は日常業務に利用しているため、暗室ネットワークで行った様な方法で評価する事はできない．そこで実際に各 proxy cache server の一週間単位のコネクション数についてキャッシュヒット率の経過を測定した．その結果は図 5 に示す様に、各 proxy cache server のヒット率は host1 で約 25%から 28%、host3 host4 host5 では 45%から 52%という値が得られた．その結果この firewall の内側でのヒット率が高くなり、インターネットへのアクセス数を少なくし、かつ host2 を通過する http プロトコルの量を減らし、外部へのアクセス量を減らす効果が得られた．

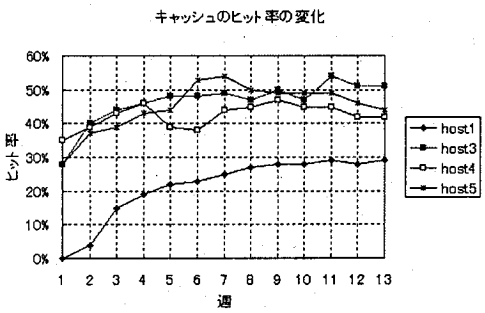


図 5: キャッシュのヒット率の変化

5 終わりに

autoproxy 機能によりキャッシュへの接続を分散しディスクの待ち時間を短縮してキャッシュの動作を改善した．これにより内と外の proxy server のコネクション時間を短縮し host2 への負担を減らす事ができた．さらにユーザへ firewall の存在を感じさせない環境とする事ができている．また proxy cache server をリレー接続した結果、50%程度のキャッシュへのヒット率が得られている．この様にかなり良好な結果が得られていると考えられるが、キャッシュのヒット率は一種の確率であり、ユーザの嗜好が似通っている場合でも良好なヒット率が得られる事もあるが、測定の結果からキャッシュの学習効果を伺うこともできている．また内部のネットワークからの多くの要求が host2 を越える事無く、インターネットへのアクセス量を低減する結果が得られている．

参考文献

[1] http://www.apan.net/ ,Asia-Pacific Advanced Network Consortium(APAN)

[2] webstone http://www.sgi.com/Products/

[3] netperf http://www.cup.hp.com/netperf/

[4] FreeBSD http://www.FreeBSD.org/

[5] http://home.netscape.com/eng/mozilla/2.0/relnotes/demo/proxy-live.html,Navigator Proxy Auto-Config File Format,March 1996

[6] squid http://squid.nlanr.net/v2/

[7] delegate http://wall.etl.go.jp/delegate/

[8] 暗室ネットワークの構築とその効果 ： 一條 島田，１９９８年３月，情報処理学会ソフトウエアー工学研究会

[9] http://naragw.sharp.co.jp/sps/,Super Proxy Script,August 9,1996

[10] http://cache.jp.apan.net/index-j.html

[11] The measured access characteristics of World-Wide-Web client proxy caches. Bradley M. Duska, David Marwood and Michael J. Feeley. In Proceedings of the USENIX Symposium on Internet Technologies and Systems, Monterey, California, December 8-11 1997.

[12] firewall 内の proxy/cache server による リレー ： 一條 山口，１９９９年４月２３日，Japan Web Cache Workshop