Internet Engineering Task Force
INTERNET DRAFT

Y. Li
Nortel Networks
W. T. Teo
National University
of Singapore
17 November 1998

IP Private Address Identification (PAID)
draft-yliteo-mobileip-paid-01.txt

Status of This Memo

Abstract

This memo describes a hierarchical IP addressing scheme that provides end-to-end host connectivity across routing realms with overlapping address space. PAID agents are routers that connect an internal private network to the global public Internet. The PAID agents' public IP addresses augment the locally significant private addresses to form globally unique binary IP addresses for private hosts. This extends the IPv4 address space and allows Internet hosts to use private instead of public IP addresses for global Internet communication. The proposal does not need any changes to the current routing infrastruture but requires an extension to the end hosts' network socket descriptors and the domain name system.

1. Introduction

This document describes a proposal to extend the IPv4 address space using a hierarchical IP addressing scheme.

1.1 Private Networks

In an internal private network, host machines are usually assigned IP address from the private IP address space [1]. These addresses typically have no topological significance outside the network i.e. external hosts cannot communicate with the hosts within the private network.

Each of the private networks (and the public Internet) constitute an independent routing realm. IP addresses are unique and valid only within each routing realm, thus routing realms may have overlapping address space. The standard IP routing mechanism cannot deliver datagrams across the different routing realms.

1.2 Datagram Delivery

Using Generic Routing Encapsulation (GRE) tunneling between the end hosts, the current IPv4 routing mechanism already supports the delivery of datagrams across different routing realms. Therefore, using GRE [2], there is no need to modify the routing infrastructure to support PAID operation. Since datagram delivery is not a problem, a solution needs only be found to address the end hosts in different routing realms.

1.3 Address Collisions

Address collision is an implication of using private IP addresses. The problem of address space collisions are further aggravated by merging privately addressed networks. Therefore to maintain the end-to-end semantics of IP addresses, the current IP addressing scheme needs to be augmented.

This document will describe one addressing scheme that can provide a globally unique identification to any private host without modifying the existing network deployment or require the network address renumbering of any hosts and still retain compatibility with the current IPv4 addressing scheme.

1.4 Hierachical IP Addressing

Within a routing realm, the current single-tier IP addressing scheme fulfills the end-to-end host connectivity requirements, thus for the rest of this document, we are only concerned with network communication across connected routing realms.

A two-tier address hierarchy is used to identify end hosts in private networks connected to the public Internet.

At the top level, all private hosts can be identified by the private network which they belong to. Since the private network must be connected to the public Internet for global communication, there must be a router that is connected to both the public Internet and the private network. This router will have a public IP address that is valid in the public Internet. The private network can therefore be globally addressed using this public IP address. Some private networks may have more than one connection to the public Internet. In this case, any of the public IP addresses can be used to identify the private network, provided the chosen address remains constant for a given network stream connection.

At the bottom level, all the end hosts can then be addressed by their private IP addresses. Therefore, private hosts can be globally addressed using a <public, private> address pair.

1.5 Design Goals

The PAID proposal attempts to prevent possible side effects to higher layer network protocols due to the introduction of binary IP addressing. Although the end hosts' network socket descriptors are associated with binary IP addresses, the GRE tunnelled payload retains the original IPv4 header. As in the traditional IPv4 addressing scheme, this encapsulated IP header's destination and source IP addresses are the end hosts' assigned IP addresses.

The design limits PAID impact on existing IP protocols and unless it is mentioned otherwise, the IP protocols should not need any modification for PAID compatibility.

The end-to-end semantics of network addresses are retained with the adoption of binary addresses, ensuring support for any end-to-end network layer security scheme.

1.6 Applicability

Since it is unrealistic to expect all Internet hosts to immediately support PAID, hosts using private IP address will enjoy less services than if they have full traditional IPv4 connectivity.

If the private hosts only require access to foreign servers but do not provide services to foreign clients, they can employ IP Network Address Translators [4] to access end hosts that do not support PAID. Thereafter, the Internet connectivity of the private hosts can expand with PAID deployment without sacrificing the latters' access to the Internet resources.

Currently, PAID is more suitable for complete IP network connectivity between multiple cooperating private networks. Enterprizes with their own Intranets can adopt PAID to merge with related organizations' private networks and form extranets. Consequently, all these extranets will also have full network connectivity with one anther.

1.7 PAID Requirements

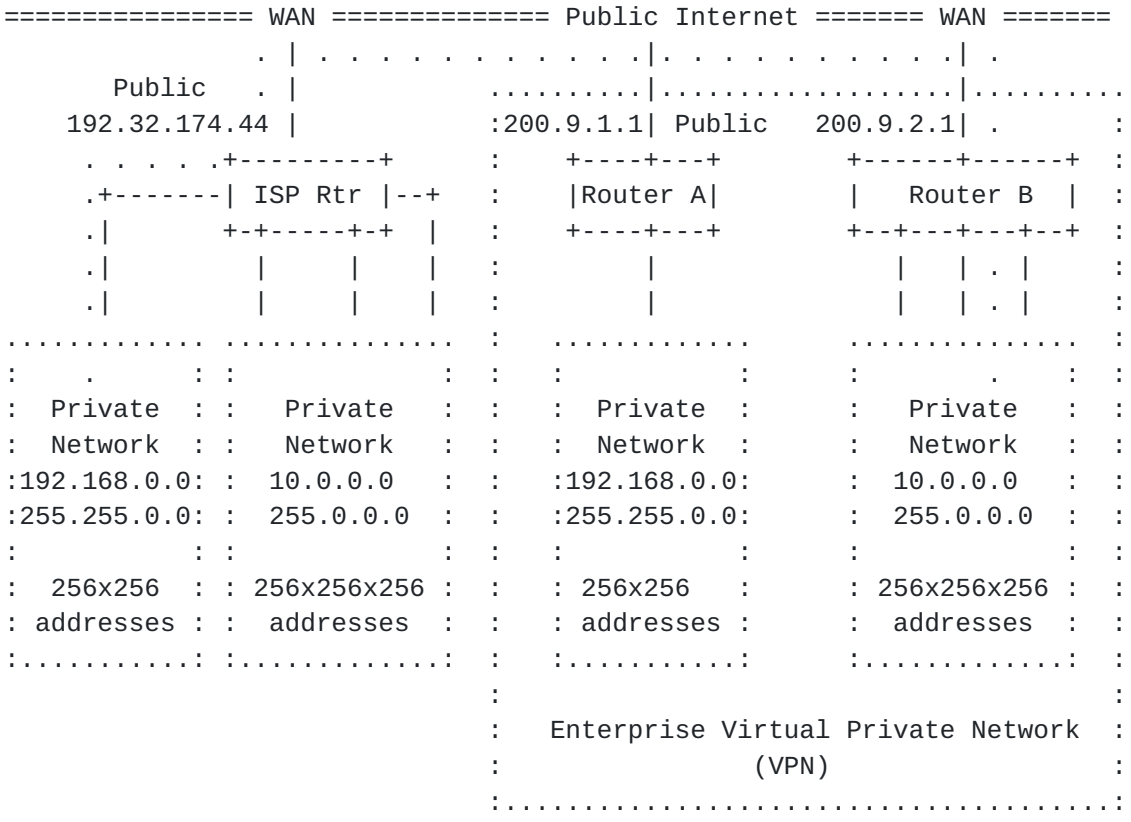


Figure 1 Hosts in different Routing Realms communicate using PAID

The diagram above illustrates a typical network deployment over the Internet. Private Address Identification (PAID) enables hosts to communicate across routing realms.

For example, a private host 10.10.10.10 in the ISP network can communicate with a private host 10.20.20.20 in the enterprise VPN by using the binary IP addresses <192.32.174.44, 10.10.10.10> and <200.9.2.1, 10.20.20.20> respectively. To enable this functionality, these two private hosts must support the binary addressing scheme and GRE encapsulation and decapsulation. Additionally, the ISP router and the Enterprise router B must support GRE tunneling. All other routers can continue to run conventional routing protocols.

2. Terminology and Definitions

2.1 Private Node

A node where all its interface addresses are not reachable from the global public Internet. These addresses are typically from the private address space as specified in RFC 1918 [1].

2.2 Public Node

A node which has at least one public interface address. The public address is routable by the global public Internet as contrast to a private node's address.

2.3 PAID Agent

A PAID agent is a public node. The PAID agent's public address provide private nodes with a globally unique identification of their routing realm.

A PAID agent is connected to both the public Internet and a private network. In Figure 1, the ISP router is a PAID agent for the private host 10.10.10.10 in the ISP network, and the router B is a PAID agent for the private host 10.20.20.20 in the enterprise VPN.

From the standpoint of routing and security, the PAID agent should be chosen from domain border routers or backbone routers.

2.4 PAID Address

This document uses a binary IP addressing scheme to identify nodes in private routing realms. The PAID address is used to identify a private node globally. The PAID address is a <Agent, Node> IP address pair.

PAID agents connect private routing realms to the public Internet. The Agent component of a PAID address is a PAID agent's public IP address. The Node component of a PAID address is the end host's IP address.

All private routing realms are connected to a common routing realm, the public Internet. The Agent address is optional for public node in this common routing realm. However, private nodes must have an Agent address for communication across routing realms. The Agent address must be reachable from the Node address using the existing routing mechanism available in a private routing realm.

The PAID addresses represent the communication end points for traffic across different routing realms and are only relevant to the end hosts. The PAID agents may also process the PAID addresses in order to handle ICMP messages from within the GRE tunnel. All other non PAID aware nodes will identify both the PAID Agents and end hosts by their Agent and Node IP addresses respectively.

The diagram below illustrates all the PAID entities during communication between two private nodes across the public Internet.

```
  Node "A" <----> Agent "B" <----> Agent "C" <------> Node "D"
10.10.10.10     192.32.174.44      200.9.2.1        10.20.20.20
```

Node "A" PAID address is <B,A> / <192.32.174.44,10.10.10.10>
Node "C" PAID address is <C,D> / <200.9.2.1,10.20.20.20>

3. Datagram Delivery

PAID uses GRE tunneling [3] for datagram delivery across different routing realm.

GRE encapsulation is a means to alter the normal IP routing for datagrams, by delivering them to intermediate destinations that would otherwise not be selected by the (network part of the) IP Destination Address field in the original IP header.

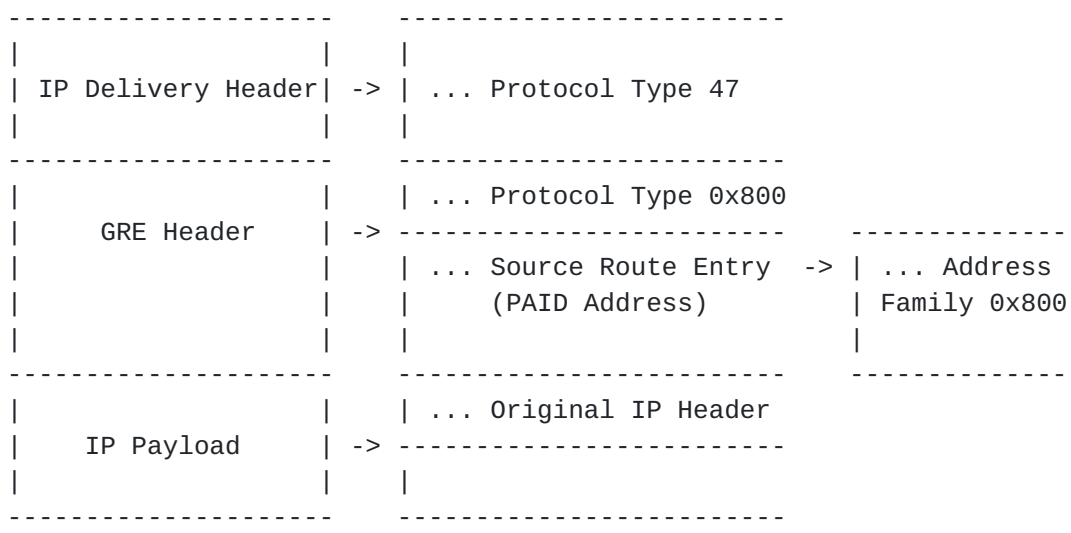
The process of encapsulation and decapsulation of a datagram is frequently referred to as "tunneling" the datagram, and the encapsulator and decapsulator are then considered to be the "endpoints" of the tunnel; the encapsulator node is referred to as the "entry point" of the tunnel, and the decapsulator node is referred to as the "exit point" of the tunnel.

The GRE encapsulation provides a Source Route Entry (SRE) in the tunnel header. Using a SRE with an Address Family indicating an IP source route (and the Strict Source Route flag cleared), the intermediate destinations can be specified.

In the case of PAID, the SRE's IP address list will include the Agent and Node addresses. The SRE list also represents the PAID addresses of the end hosts. The end points of the GRE tunnel must therefore be the communicating end hosts.

3.1 Overall PAID packet

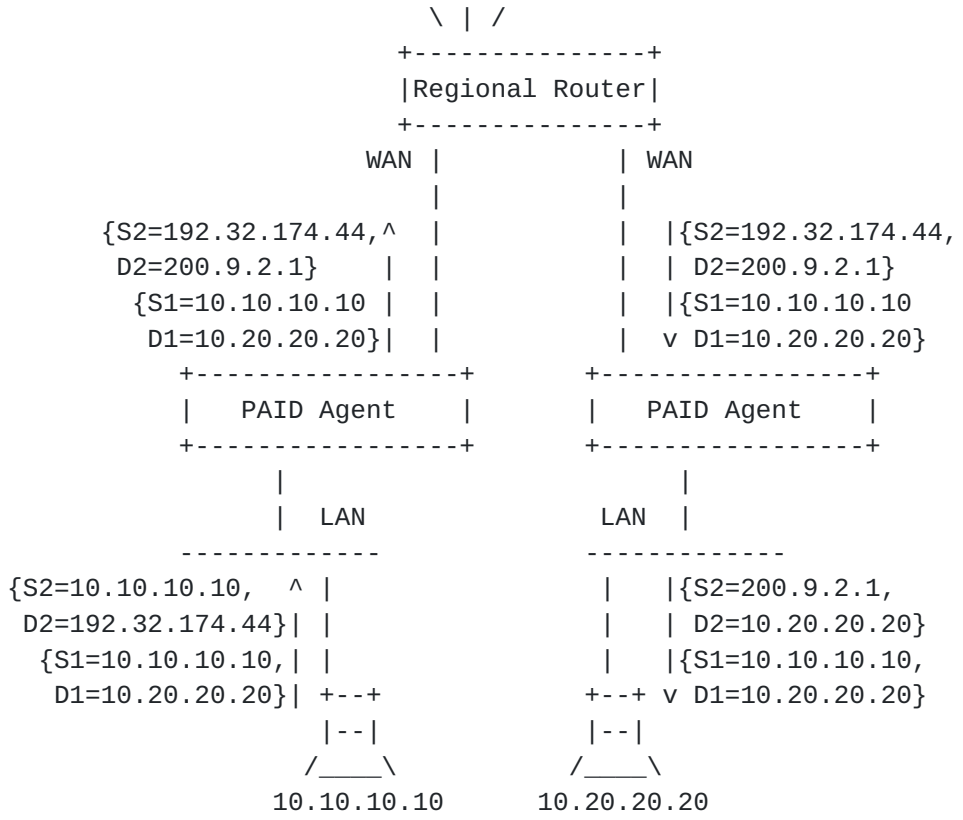
The diagram below provides an overview of the entire PAID packet.



3.2 GRE SRE Processing

PAID agents should process the GRE's SRE as specified in RFC 1702 [2]. The diagram illustrates the processing of a PAID packet in the example in [Section 1.7].

S1 and D1 represent the original IP header's source and destination addresses respectively. S2 and D2 represent the IP delivery header's source and destination addresses respectively.



3.3 Source Node

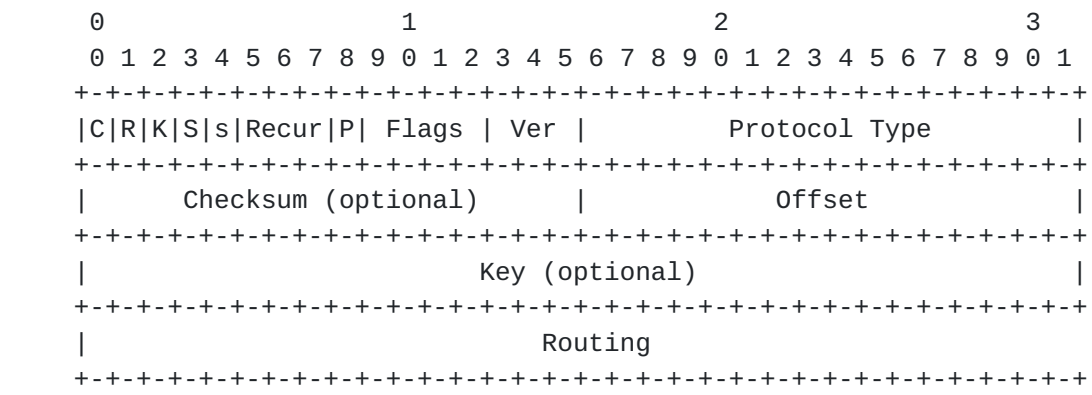
When a private node generates a packet destined to another routing realm, it must perform GRE encapsulation for the packet.

IP Delivery Header

The TTL, TOS and IP options of the original IP header must be copied into the same fields in the IP delivery header.

GRE Header

The GRE Checksum and Key values are optional for PAID packets. The GRE Sequence Number is not required for PAID packets.



Checksum Present (bit 0)

If the Checksum Present bit is set to 1, then the Checksum field contains valid information.

BOTH the Checksum and Offset fields are present in the GRE packet.

Routing Present (bit 1)

The Routing Present bit is set to 1.

Key Present (bit 2)

If the Key Present bit is set to 1, then it indicates that the Key field is present in the GRE header.

PAID Present (bit 8)

The PAID Present bit is set to 1, to indicate PAID addresses is used.

The Sequence Number Present (bit 3), Strict Source Route (bit 4), Recursion Control (bits 5-7) and Version Number (bits 13-15) bits

are set to 0.

Protocol Type (2 octets)

The Protocol Type field is 0x800.

Offset (2 octets)

The offset field indicates the octet offset from the start of the Routing field to the first octet of the active Source Route Entry to be examined. The default value is 0.

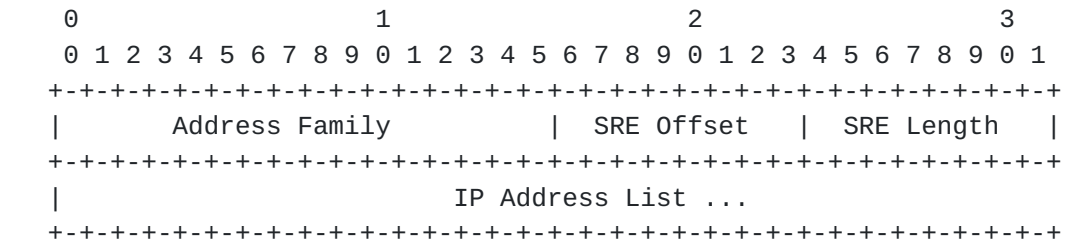
Checksum (2 octets)

The Checksum field contains the IP (one's complement) checksum of the GRE header and the original IP payload.

Key (4 octets)

The Key field contains a four octet number which is inserted by the source node. It may be used by the destination node to authenticate the source of the packet. The techniques for determining authenticity are outside of the scope of this document.

The Routing field is a list of Source Route Entries (SREs). Each SRE has the form:



The routing field is terminated with a "NULL" SRE containing an address family of type 0x0000 and a length of 0.

Address Family (2 octets)

The Address Family field is set to 0x800.

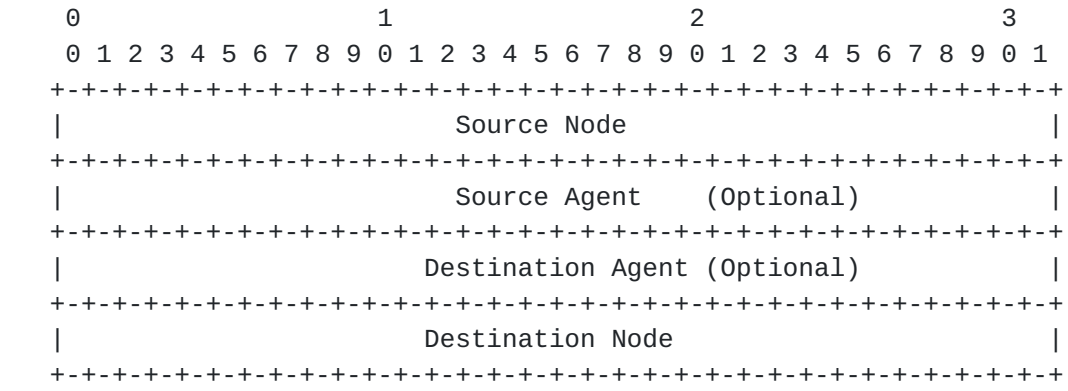
SRE Offset (1 octet)

The SRE Offset field indicates the octet offset from the start of the Routing Information field to the first octet of the active entry in Source Route Entry to be examined. The initial default value is 8.

SRE Length (1 octet)

The SRE Length field contains the number of octets in the SRE. If the SRE Length is 0, this indicates this is the last SRE in the Routing field.

The IP Address List indicates the intermediate destinations from the source node to the destination node that the PAID packet has to traverse. The list also represent the PAID addresses of the end hosts.



The source node's PAID address is <Source Node, Source Agent>. The destination node's PAID address is <Destination Node, Destination Agent>.

The first and last IP address list entries MUST be the end nodes' IP addresses.

Either the Source or Destination Agent entry MUST be present. If there is only one Agent IP address entry, the latter may be treated as the Source or Destination Agent in the end nodes' PAID address.

3.4 Destination Node

On receiving the PAID packet, the destination node should record the PAID addresses in the GRE's SRE, before performing GRE decapsulation. The TTL, TOS and IP options of the IP delivery header must be copied into the same fields in the original IP header.

3.5 ICMP messages

The PAID agents should handle ICMP messages from within the GRE tunnel as specified in [5], including the maintenance of tunnel "soft state".

4. PAID DNS

This section describes the enhancements to DNS required to support PAID addressing and the procedure for PAID hosts to perform PAID DNS lookup.

PAID DNS resolution can typically be performed through a two step DNS resolution by the source host. The traditional public DNS server need to be enhanced to provide the public addresses of PAID agents and private DNS servers of a private network.

This document operates on the paradigm that interconnecting routing realms may have overlapping address space but the Fully Qualified Domain Names (FQDN) of hosts that fall within IN-ADDR.ARPA domain must be unique for the PAID DNS service to work.

4.1 Public DNS Server

A name server contains "in-addr.arpa" records and enable reverse "address to name" lookup. To support PAID, two new DNS records type are introduced in the public DNS server - the PX and DX records. The solution is a generalization of the MX record scheme [6] currently used to identify a mail exchange in a domain.

The PX records list the PAID agents' public addresses in a domain and the DX record points to the public address of the domain's private DNS server.

4.2 Private DNS Name Server

A private DNS server contains the resource records (RRs) of the private nodes in its domain. This DNS server should be separated from the general Internet DNS referrals to prevent the non routable private IP addresses from propagating on public networks.

The private DNS server should support secure DNS name service and may sign replies that originate from the external world.

4.3 DNS Resolver

The source hosts' resolvers must support an extension to the iterative mode DNS resolution process [4]. The resolver must retrieve the PX and DX records of the destination domain before querying the destination domain's private DNS server.

For simplicity, the proposal assumes the source node resolver has access to the public Internet. This access may be via network address (port) translators [4] or any other application gateways. Alternatively, the DNS server in a private network may support recursive service to access the private DNS server in the destination

domain, provided it can guarantee the source host is PAID aware. The

DNS server acts as an intermediatery to cross the routing realm boundaries. The rest of this document assumes the DNS server does not support recursive service.

The resolver should typically return the PAID address to the user program.

4.4 PX and DX records

In a definative scheme, it would be neccessary to define the DNS record type and the corresponding format. Currently for easier deployment, the two entries may use the generic "text" record to register the PAID agents and private name server of a domain. This record is desiged for general purpose extensions in the DNS, and its content is a text string.

The PX record will contain three fields :

- A record identifier composed of the two characters "PX".

- A cost indicator, encoded up to 3 numerical digits.

- An IP address, encoded as a text string following the "dot" notation.

The three strings will be separated by a single comma. An example of a PX record would thus be:

domain	type	record	value
*.2.9.200.in-addr.arpa	IP	TXT	RX, 10, 200.9.2.1

The DX record is similar to the PX record except there are only two fields :

- A record identifier composed of the two characters "DX".

- An IP address, encoded as a text string following the "dot" notation.

4.5 PAID DNS requirements

The scheme is valid only if the PX, DX records and the private DNS server of the destination domain can be accessed from the public Internet.

A private domain that wants to obtain dynamic connectivity using this scheme will have to replicate its domain name service info so as to provide them through servers accessible from the core of the

Internet.

5. Name Lookup

For Host name to Host Address address query requests :

Assuming the destination node is dst.private.com.

The source host resolver will query the DNS with the name for the destination host and obtain the PX and DX records. The PX record lists the Agent address for the destination node.

The name resolver on the source host node will send the name lookup query (A record) for dst.private.com to the private.com domain's private DNS server listed in the DX record. The Node address for the destination node is returned.

For Host address to Host name queries :

Assuming the destination node is <192.32.174.44,10.10.10.10>

The source host resolver will query the DNS with the host address 192.32.174.44 and obtain the DX record.

The name resolver on the source host will send a inverse name lookup query (PTR record) for "10.10.10.10.IN-ADDR,ARPA." to the private domain's private DNS server listed in the DX record. The host name for the destination node is returned.

5.1 Choosing a PAID agent

Having only one PAID agent will be a single point of failure and a possible bottleneck device.

The PX records should carry associated with each PAID agent a preference identifier. To select a PAID agent, one has to rely on heuristical approaches. The easiest may be to always choose the "preferred PAID agent" - the PX entry with the minimal preference index or alternatively chose one PAID agent randomly within the list for each stream network connection. This will spread the traffic on several routes and introduce better load sharing and more redundancy to the network.

5.2 Performance

The initial DNS exchanges required for loading the record information may induce a response time penalty for the users. Some caching strategy of each private routing realm's PAID agents and private DNS server should be sufficient to alleviate the performance effect.

6. Applications Consideration

Any application protocol that embeds the end nodes' IP addresses in the application payload may need to be PAID aware if these addresses are consequently used to create another separate network connection.

7. Security Considerations

This document proposes a scheme to allow network communication across routing. GRE is a clear text encapsulation mechanism and does not protect sensitive data over the unsecure global Internet. Since PAID retains the end-to-end semantics across routing realms, additional security mechanism such as IPSEC should be used to protect the original IP payload.

Organizations may not want to enable full network connectivity for all private hosts nor allow access from all external hosts. The PAID agents should support host access controls. The firewall rules may be enhanced to deny or allow access based on PAID addresses in the GRE' SRE.

Acknowledgements

Many thanks to Dr. Y. C. Tay at the National University of Singapore for supporting this joint work as well as for his valuable comments.

This work was supported in part by National University of Singapore ARF grant RP960683.

References

[1] Rekhter, Y., Moskowitz, B. Karrenberg, D., G. de Groot, and Lear, E. "Address Allocation for Private Internets", RFC 1918, February 1996

[2] S. Hanks, T. Li, D. Farinacci and P. Traina, "Generic Routing Encapsulation over IPv4 networks", RFC 1702, October 1994

[3] S. Hanks, T. Li, D. Farinacci, P. Traina, "Generic Routing Encapsulation (GRE)", RFC 1701, October 1994

[4] P. Srisuresh, K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", <draft-ietf-nat-traditional-01.txt> - work in progress, November 1998

[5] C. Perkins, "IP Encapsulation within IP", RFC 2003, May 1996

[6] P. Mockapetris, "Domain Name - Concepts and Facilities", RFC 1034, November 1987

Author's Address

Y. Li
Nortel Networks
BL60-304
600 Technology Park Drive
Billerica, MA 01821

Phone: 1-978-916-1130
Fax: 1-978-670-8760
E-mail: yunli@NortelNetworks.COM

W. T. Teo
Department of ISCS
National University of Singapore
Lower Kent Ridge Crescent
SINGAPORE 119260

E-mail: teoweetu@iscs.nus.edu.sg