INTERNET-DRAFT   Joe Touch
draft-touch-l3vn-arch-00.txt   Lars Eggert
                                                         Yu-Shun Wang
                                                              USC/ISI
                                                            Jun. 2002
                                                   Expires: Dec. 2002

An Architecture for Layer 3 Virtual Networks

Status of this Memo

   This document is an Internet-Draft and is in full conformance with
   all provisions of Section 10 of RFC2026 except for the right to
   produce derivative works.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF), its areas, and its working groups. Note that other
   groups may also distribute working documents as Internet-Drafts.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet- Drafts as reference
   material or to cite them other than as "work in progress."

   The list of current Internet-Drafts can be accessed at
   http://www.ietf.org/ietf/1id-abstracts.txt

   To view the list Internet-Draft Shadow Directories, see
   http://www.ietf.org/shadow.html

Abstract

   This document describes an architecture for layer 3 (IP) virtual
   networks. Virtual networks consist of virtual hosts and virtual
   routers connected by virtual links (tunnels) just like a real
   network. The focus of this draft is to extend the current Internet
   architecture to support virtual networks.

1. Introduction

   This document describes an architecture for layer 3 (IP) virtual

networks. Layer 3 virtual networks use only layer 3 protocols to
provide layer 3 connectivity within each virtual network. Virtual
networks consist of virtual hosts and virtual routers connected by
virtual links (tunnels). The architecture is based on the current
Internet architecture with some extensions and techniques required to
support virtual networks. The components in this architecture are
also examined against the current Internet standards.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in RFC 2119.

1.1 Concepts of Virtual Networks

The concepts and purposes of virtual networks in a physical network
are the same as virtual memory [1] in computer systems:

  - Provide an abstract view of their real counterparts to
    applications.
  - Allow concurrent sharing of the physical resources among
    multiple applications.
  - Ensure isolation among different virtual entities.

Virtual networks are constructed by linking nodes with tunnels, which
encapsulates packets inside virtual networks with additional headers.
One feature unique to virtual networks is the capability to bypass
nodes not in the virtual networks. This property led to the early
adoption of virtual networks to incrementally deploy new protocols
such as multicast [2] and IPv6 [3]. It also enabled the recent
commercial deployment of virtual private networks (VPN) [4], and the
proliferation of peer-to-peer networks [5].

1.2 Virtual Private Networks

Virtual Private Network (VPNs) are an important subset of virtual
networks. By tunneling the traffic, virtual networks provide
isolation among traffic belong to different virtual networks, also
between virtual networks and the underlying Internet. But tunneling
alone does not provide security. Other security mechanisms and
protocols, such as encryption, authentication, access control, and
policy management must be used to secure the nodes and traffic in
virtual networks.

This document does not discuss security measures and policy
management for virtual private networks. While they will certainly
influence how a virtual network is provisioned and managed, they

really should not affect the architecture of the virtual network. A
good virtual Internet architecture should be able to use existing
Internet mechanisms to achieve security. The modifications should not
exceed the changes made to the current Internet to accommodate
security.

1.3 Types of Virtual Networks - Open, Close, & Router Cloud

As with the real network, a "close" virtual network confines the
traffic to within the virtual network. An "open" virtual network
allows communications between nodes inside the virtual network with
nodes outside, either in another virtual network or in the underlying
Internet. Using virtual networks as router clouds is the third
category between "open" and "close". In this case, the virtual
network really acts like a transit network for packets from the
outside.

Normally, virtual networks are designed to be closed for isolation or
security purposes. The use of tunnel encapsulation also prevents
direct communication between the virtual network and the underlying
physical network. For packets passing through the virtual network as
a transit, it is sufficient to encapsulate them with the extra header
to direct these packets to their exit points in the virtual network.
For packet exchange between a node inside a virtual network and a
node in the real network, the boundary nodes need to perform NAT-like
operations to transform packets crossing the boundary.

[2](#). Architecture of Virtual Networks

This section describes the architecture and components of layer 3
virtual networks.

2.1 Virtual Networks

The architecture of the layer 3 virtual networks is very simple, and
is the same as the real network. The architecture is defined below:

A virtual network consists of virtual hosts and virtual routers
connected by virtual links (tunnels) in an arbitrary topology.
Virtual hosts and virtual routers are RFC-conformant hosts and
routers inside a virtual network. While the Internet runs (mostly)
over layer 2 links, a layer 3 virtual network runs over layer 3
tunnels, uses only layer 3 protocols to provide layer 3
connectivity.

2.2 Virtual Hosts

  Virtual hosts, like their real counterparts, are either packet
  sources or sinks, but never packet transits inside a virtual network.
  Virtual hosts MUST conform to the Internet Host Requirements [6][7].
  A virtual host can be located inside a real multihomed host or a
  router in the underlying real network as discussed below.

2.2.1 Virtual Hosts in Multihomed Hosts

  When a virtual host is inside a real host, the host becomes
  multihomed because it has one host (home) for the real network and at
  least another host (home) in the virtual network.

  RFC 1122 specified two models for multihomed hosts: Strong End System
  (ES) Model and Weak ES Model. The Strong ES model states that the
  destination address of an incoming packet MUST match the address of
  the physical interface through which it is received, and an outgoing
  packet must be sent through the interface that corresponds to the
  source address of the packet. The models are updated to include
  virtual interfaces associated with tunnels in a virtual network in
  addition to physical interfaces.

  A multihomed host SHOULD follow the Strong ES model instead of the
  Weak ES Model to ensure the isolation of the traffic in a virtual
  network from those of the underlying network and other virtual
  networks.

  The multihomed host architecture also supports a host participating
  in more than one virtual network. There might be administrative
  reasons not to do so, but the architecture supports it.

  Another requirement regarding multihomed hosts is the capability to
  "route" outbound packets among the "homes". This also implies that a
  multihomed host MUST be able to forward packets like a router. It
  doesn't mean a multihomed host needs to run routing protocols like a
  real router, but it must conform to the Router Requirements [8][9]
  regarding packet forwarding operations.

2.2.2 Virtual Hosts in Real Routers

  As described in Section 1.3, when a virtual network is open or acts
  as a transit network, routers could act as virtual hosts inside a
  virtual network. There are three possible communication patterns
  depending on the entities involved and the type of virtual networks:

    1. Virtual networks as transit clouds:

In this case, the edge nodes of the virtual network will
encapsulate the incoming packets from the real network with
additional headers to send them through the virtual network to
their corresponding exit points. Note that the real source and
destination are hidden from the nodes in the virtual network.
These packets appear to originate from one such edge node and
destined to another edge node. Though those edge nodes are
routers in the real network, they act as packet sources and
sinks inside the virtual networks. As far as the virtual
network is concerned, they really are just virtual hosts.

2. Communication between virtual network and physical network:

This requires the edge nodes between the virtual network and
the physical network to perform a NAT-like [10] operation on the
packets crossing the boundary to "elevate" or "transform" the
packets into the virtual network space. But to the virtual
network, those packets effectively originated from and destined
to those edge nodes, which makes them just like hosts in the
virtual network.

3. Inter virtual network communication

Strictly speaking, the edge nodes in this case are not virtual
hosts, but should be virtual routers. They perform the same
functionality as the border routers of the Internet AS's,
exchanging routing information among different virtual networks,
and forwarding packets across the boundary between different
virtual networks.

Note that this is only possible if the address spaces of the
connected virtual networks do not overlap. Otherwise, the
edge nodes will need to perform NAT-like translation on
packets crossing the boundary, and this will make the edge
nodes again behave like virtual hosts in the virtual network.

In all the cases, additional mechanisms and protocols are required to
handle the dissemination of address and routing information between
virtual networks and the physical networks, and the encapsulation and
decapsulation of the packets crossing the boundary of virtual
networks and the real network. The former can be done using one of
the exterior gateway routing protocols, the latter could be
implemented as a variation of NAT [10]. But neither should be
consider as part of the virtual network architecture because as far
as the virtual network is concerned, those mechanisms and protocols,
while non-trivial, are nothing more than "applications" using the
virtual network (generating and receiving packets).

2.3 Virtual Routers

   Virtual routers are routers inside virtual networks. The fundamental
   concepts of virtual routers described in this draft are very similar
   to those described in [11]. Virtual routers work exactly the same way
   in virtual networks as physical routers do in real networks, which
   means they must also conform to the Internet router requirements
   [8][9] except some constraints regarding the use of virtual links
   instead of physical ones, and containment and isolation of the
   routing information among virtual networks and between virtual
   networks and the underlying physical networks.

   The main functionality of virtual routers is to forward packets
   inside a virtual network. The dissemination of reachability
   information and the construction of the routing table can be achieved
   either by static manual configurations or by running dynamic routing
   protocols inside the virtual network. Like the real network, the
   exchange of routing information in the case of dynamic routing
   protocols must be limited within the boundary of a virtual network.
   Any exchange crossing the boundary of virtual networks and real
   networks must be treated with extreme care and can only be done when
   the address spaces of different virtual networks do not overlap.

   The main difference between a virtual router and a physical router is
   the links they operate on. For virtual routers, links in a virtual
   networks are actually point-to-point tunnels in the underlying
   network. The dynamic routing protocols chosen for virtual networks
   must be able to exchange routing information using the tunnels. This
   would ensure the isolation and containment of the routing information
   among different virtual networks and also the real network. For
   routing protocols with special requirements regarding the properties
   of the links, e.g., broadcast vs. point-to-point media, specific link
   layer protocols, etc., the tunneling mechanisms might limit the
   choice of routing protocols used in a virtual network.

 2.4 Virtual Links

   Virtual links are links among different (virtual) nodes in a virtual
   network. Virtual links are often implemented as point-to-point
   tunnels encapsulating packets with extra header(s). This is different
   from physical links with broadcast media. The implication is that
   some protocols, like multicast, might not work over virtual links
   because they assume certain link properties.

 2.5 End-to-End vs. Router-Cloud/Transit Virtual Networks

End-to-end virtual networks are virtual networks that extend to the
hosts, while router-cloud or transit-net virtual networks stop at the
edge routers. As described in Section 2.2.2, the router-cloud model
is actually a special case of the end-to-end virtual network in which
the virtual hosts sit inside real routers and the packets from the
real hosts will be converted when entering (encapsulation) and
leaving (decapsulation) the virtual network. Within the virtual
network, the source and destination addresses of the real hosts are
hidden from the virtual routers. As far as the virtual network is
concerned, edge routers are acting like sources and sinks of packets
inside the virtual network.

An alternative view of the router cloud model is to include the real
hosts connected from outside the router cloud as part of the virtual
network and try to management them altogether. These often result in
added complexity to the deployment and management of virtual networks
when what is really needed are for the edge nodes (of the virtual
network) to exchange reachability information for those hosts using
(but not inside) the virtual networks.

2.6 Layer "X" VPNs (where X > 1)

The following is a discussion of implementing virtual networks in
different layers of the network protocol stack.

2.6.1 Link Layer (L2) Virtual Networks

Different types of link-layer (L2) virtual networks [12] were
proposed to provide layer-2 connectivity, (or forward packets based
on layer-2 information and incoming links,) across the Internet.
While this can certainly be done, there are potential issues
regarding this approach:

1. Most layer-2 (LAN) protocols have latency bounds. Using LAN
   protocols over wide area MAY breaks those specifications.
   There MUST be mechanisms to detect when this will occur and
   disable the L2 protocols, rather than allow inconsistencies
   to exist.

2. Most LAN protocols assumes broadcast media while most virtual
   networks use point-to-point tunnels. Some implementations of
   L2 virtual networks use software copy to emulate broadcast.
   The problem is that atomicity and ordering of broadcast
   packets MUST be maintained in such implementations. Otherwise
   the resulting virtual network is not broadcast-capable, which
   MUST inhibit L2 protocols that rely on such capability.

3. There are no hierarchical structures in the address space of most
   layer 2 protocols, and often no routing infrastructure and no
   mechanisms to disseminate address lookup information in wide area
   for the same reason. This means routing (reachability) and address
   lookup within such virtual networks require additional protocols
   and mechanisms.

To overcome the issues caused by using layer-2 protocols in an
environment they are not designed for, network engineers inevitably
have to reinvent many new mechanisms to circumcise those problems and
make the resulting frameworks unnecessarily complex.

Another issue is security. Most link layer protocols have no security
mechanisms, makes it necessary to use additional security mechanisms
in higher layers of the protocol stack. While this is not a
shortcoming of layer-2 virtual networks, it is an important factor
when security is required, e.g., virtual private networks.

2.6.2 Network Layer (L3) Virtual Networks

Network-layer virtual networks, or layer-3 virtual networks, assume
network-layer connectivity, and use only network-layer protocols to
provide layer-3 virtual networks. An example is virtual networks
constructed by using IP-IP tunnels.

Constructing virtual networks over layer 3 protocols have the
following advantages over other layers:

1. Layer 3 (IP) has global naming and addressing structure.
   Many virtual networks in layers above L3 actually map their
   node IDs to L3 addresses.

2. Layer 3 has routing protocols in the architecture. This
   means a layer 3 virtual networks can just use the same
   routing protocols within themselves.

As will be discussed in the next section, many virtual network
architectures in other layers often reinvent these functionalities of
layer 3.

2.6.3 Layer 4 (and above) Virtual Networks

The problem with implementing virtual networks on layer 4 to layer 7
is that they don't have naming/addressing and routing functionality.
An L4 (or L7) virtual network protocol often ends up reinventing
these functionalities, usually by either routing on L7 names or
providing a separate tag space so L7 names map to L3 addresses.

The drawbacks are twofold:

1. Reinvented protocols often incompletely recapitulate the
   discovery of errors, and usually not as well-implemented
   as the existing L3 protocols.

2. Similar functionalities at different layers could interfere
   with each other, resulting in mis-routing, re-routing, or
   dead-ends, etc.

[3]. Service

This section lists some of the services provided by virtual networks
to hosts and routers inside the virtual networks, and at the edge of
the virtual networks in the case of virtual networks as transit nets.

3.1 Addressing

Private IP addresses defined in RFC [13] SHOULD be used inside the
virtual network if it is a closed network isolated from the Internet.
Addresses assigned to the components within a virtual network must be
unique throughout the virtual network, and when communicating with
other virtual networks, the scope of uniqueness must be extended to
cover all the connected virtual networks.

While the isolated nature of virtual networks allows network
administrators to re-use real IP addresses in the virtual networks,
this will make it difficult or confusing to manage and differentiate
routing information for the real network versus the virtual ones when
they touch down on the same node(s).

3.2 IP Connectivity

A virtual network must provide IP connectivity to hosts and routers
inside. Packet forwarding must be transparent to the hosts just like
in the Internet.`

3.3 Routing

3.3.1 Within a Virtual Network

Inside a virtual network, the construction of routing tables on
virtual routers can be done by using static routes and manual

configurations, or by using dynamic routing protocols among the
virtual routers to exchange the reachability information.

No matter which approach is used, it is important to ensure that
layer 3 routing must work in a virtual network just as in the base
network. Failure of an existing routing mechanism in the virtual
network means there are problems in the virtual network architecture.

3.3.2 Between Multiple Virtual Networks

When the address spaces of multiple virtual networks do not overlap,
it is possible to run exterior routing protocols (e.g., BGP [14])
among them to exchange routing information just like inter-AS routing
of the current Internet.

3.3.3 Between Virtual Networks and the Internet

Depending on the type and use of virtual networks, the routing
information exchange between a virtual network and the underlying
Internet is either out of scope, or no different from the practice of
the current Internet.

1. Closed virtual networks with private [RFC 1918] addresses:
   This is often the case when the virtual network is isolated from the
   underlying network. The intention is to have a "closed" abstract
   network environment which does not support communication across the
   boundary of virtual networks and the Internet.

2. Open virtual networks with valid Internet addresses:
   While not a normal way of using virtual networks, virtual networks of
   this type are the same as autonomous systems (AS's). Just run BGP or
   other inter-domain routing protocols at the edge of the virtual
networks.

3. Virtual networks as transit clouds:
   As discussed in Section 2.2.2, and 2.5, a virtual network in this case
   becomes a transit domain for the real network. It is necessary to
   exchange reachability information among all the edge nodes of the
   virtual network, and between the virtual network and the Internet.
   Again, this is exactly the same as how routing exchange is done for
   transit AS's in the Internet.

3.4 Security

As discussed in Section 1.2, security mechanisms and protocols
required to implement VPNs, while very complex and non-trivial to
deploy, are orthogonal to the architecture of virtual networks.

These security issues are currently being addressed in several IETF
working groups regarding the current Internet. Once they are
established and standardized, the virtual networks SHOULD be able to
utilize them, just like in the Internet.

 3.5 QoS

   For a virtual network to support QoS, the components of the virtual
   network must be able to do the following:

      1. Virtual routers can assure, enforce, and reserve bandwidth
         constraints; and,
      2. Virtual links (tunnels) can reserve and enforce bandwidth
         requirements.

   Virtual networks can support QoS to the extend that the components
   can, and QoS in virtual nets works ONLY if QoS is continuously
   deployed on all possible underlying network (physical networks,
   Internet, etc.) paths. This is currently not the case now. QoS is
   one of the few properties that virtual networks could just bypass the
   parts of the physical networks that don't support it and still make
   it work.

   QoS in virtual networks also assumes that reservations on the
   underlying network are then served as reservable services to the
   higher layer virtual network. This kind of QoS mechanism does not
   exist, though a 2-level variant was proposed [16].

 3.6 Multicast

   Current deployment of multicast relies heavily on the broadcast
   capability of the LAN protocols in the stub networks. Point-to-point
   tunnels in the virtual networks break this assumption, and makes it
   necessary to explicitly specify those tunnels in the multicast
   routing configurations.

 3.7 DHCP

 3.8 MPLS

[4]. Issues

4.1 Provisioning

   Other protocols and mechanisms are required to deploy a virtual
   network:

   Required:

      - Resource discovery
      - Link (tunnel) configuration
      - Routing configuration
      - Address allocation
      - Membership management

   Optional:

      - Security policy management and enforcement
      - QoS policy management and enforcement

4.2 Routing

   There are situations where tunneling and security protocols interfere
   with routing protocols within the virtual networks. Please refer to
   the documents regarding different types of tunneling mechanisms,
   security protocols, and routing algorithms to resolve the conflicts
   [17].

4.3 Architectural Components vs. Management Boundary

   There are hosts and routers, servers and clients in the current
   Internet architecture. The only place where management entities are
   used to define network architecture is for inter-domain routing as
   opposed to intra-domain routing. The same principle applies to
   virtual networks. While a virtual network may extend cross several
   administrative domains, same components in different domains still
   perform the same functions. There is no difference architecture-wise
   across management boundaries.

   One example is the current trend [15] in the IETF in defining
   different frameworks for privider-edge PPVPNs (PE-based PPVPNs) and
   customer-edge PPVPNs (CE-based PPVPNs). Unless the provider-edge
   routers behave differently from the customer-edge routers in a
   virtual network, administrative boundary (or equipment placement)
   along does not sufficiently define different architectures.

5. Security Consideration

This draft describes an architecture for virtual networks. Security
measures are required if the resulting virtual networks need to be
secure [17][18].

Other security consideration are not discussed in this draft.

6. Conclusion

This draft describes an architecture for layer 3 (IP) virtual
networks, the components of virtual networks, and service provided by
these virtual networks.

Acknowledgments

The authors would like to thank the members of the X-Bone and
DynaBone projects at USC/ISI for their contributions to the ideas
behind this draft, notably Greg Finn and Amy S. Hughes.

References

[1] Denning, P. J., "Virtual Memory," ACM Computer Surveys, Vol. 2,
No. 3, September 1970, pp. 153-189.

[2] Eriksson, H., "MBone: The Multicast Backbone," Communications of
the ACM, August 1994, pp.54-60.

[3] 6-Bone web pages - http://www.6bone.net

[4] Scott, C., Wolfe, P., Erwin, M., Virtual Private Networks,
O'Reilly & Assoc., Sebastapol, CA, 1998.

[5] Oram, A. (Editor), Peer-to-Peer: Harnessing the Benefits of a
Disruptive Technology, O'Reilly, Sebastapol CA, 2001.

[6] Braden, R. (Editor), "Requirements for Internet Hosts --
Communication Layers", RFC 1122, October 1989.

[7] Braden, R. (Editor), "Requirements for Internet Hosts --
Application and Support", RFC 1123, October 1989.

[8] Baker, F. (Editor), "Requirements for IP Version 4 Routers," RFC 1812, June 1995.

[9] Senie, D., "Changing the Default for Directed Broadcasts in Routers," RFC 2644, August 1999.

[10] Srisuresh, P., Egevang, K., "Traditional IP Network Address Translator (Traditional NAT)," RFC 3022, January 2001.

[11] Knight, P. (Editor) et al., "Network based IP VPN Architecture using Virtual Routers," (work in progress), February 2002.

[12] Kompella, K. et al., "Layer 2 VPNs Over Tunnels," (work in progress), June 2001.

[13] Rekhter, Y., et al., "Address Allocation for Private Internets," RFC 1918, February 1996.

[14] Rekhter, Y., Li, T. (Editors), "A Border Gateway Protocol 4 (BGP-4)," RFC 1771, March 1995.

[15] Callon, R. (Editor) et al., "A Framework for Layer 3 Provider Provisioned Virtual Private Networks," (work in progress), April 2002.

[16] Terzis, A., Krawczyk, J., Wroclawski, J., Zhang, L., "RSVP Operation Over IP Tunnels," RFC 2746, January 2000.

[17] Touch, J., Eggert, L., "Use of IPsec Transport Mode for Dynamic Routing," (work in progress), June 2002.

[18] Kent, S., Atkinson, R., "Security Architecture for the Internet Protocol," RFC 2401, November 1998.

Author Information

  Joe Touch
  Lars Eggert
  Yu-Shun Wang

  Information Sciences Institute
  University of Southern California
  4676 Admiralty Way, Suite 1001
  Marina del Rey, CA 90292-6601
  USA

    Phone: +1 310 448-9151
    Fax: +1 310 448-9300

    URL: http://www.isi.edu/{touch,larse,yushunwa}
    Email: {touch,larse,yushunwa}@isi.edu

Attribution and Disclaimer

The views and conclusions contained herein are those of the authors and
should not be interpreted as necessarily representing the official
policies or endorsements, either expressed or implied, of the Defense
Advanced Research Projects Agency (DARPA), the Air Force Research
Laboratory, or the U.S. Government.