                        Softwire Problem Statement

Status of This Memo

Copyright Notice

Abstract

   This document captures the problem statement for the Softwires
   Working Group, which is developing standards for the discovery,
   control, and encapsulation methods for connecting IPv4 networks
   across IPv6-only networks as well as IPv6 networks across IPv4-only
   networks. The standards will encourage multiple, inter-operable
   vendor implementations by identifying, and extending where necessary,
   existing standard protocols to resolve a selected set of "IPv4/IPv6"
   and "IPv6/IPv4" transition problems. This document describes the
   specific problems ("Hubs and Spokes" and "Mesh") that will be solved
   by the standards developed by the Softwires Working Group. Some
   requirements (and non-requirements) are also identified to better
   describe the specific problem scope.

Table of Contents

1. Introduction

   The Softwires Working Group is specifying the standardization of
   discovery, control, and encapsulation methods for connecting IPv4
   networks across IPv6-only networks and IPv6 networks across IPv4-only
   networks in a way that will encourage multiple, inter-operable vendor
   implementations. This document describes the specific problems
   ("Hubs and Spokes" and "Mesh") that will be solved by the standards
   developed by the Softwires Working Group. Some requirements (and
   non-requirements) are also identified to better describe the specific
   problem scope. A few generic assumptions are listed up front:

   o Local Area Networks will often support both protocol families in
      order to accommodate both IPv4-only and IPv6-only applications, in
      addition to dual-stack applications. Global reachability requires
      the establishment of softwire connectivity to transit across
      portions of the network that do not support both address families.
      Wide area networks that support one or both address families may
      be separated by transit networks that do not support all address
      families. Softwire connectivity is necessary to establish global
      reachability of both address families.

   o Softwires are to be used in IP-based networks to forward both
      unicast and multicast traffic.

   o Softwires are assumed to be long-lived in nature.

   o Although Softwires are long-lived, the setup time of a softwire is
      expected to be a very small fraction of the total time required
      for the startup of the Customer Premise Equipment (CPE)/Address
      Family Border Router (AFBR).

   o The nodes that actually initiate softwires should support dual-
      stack (IPv4 and IPv6) functionality.

   o The goal of this effort is to reuse or extend existing technology.
      The 'time-to-market' requirement for solutions to the stated
      problems is very strict and existing, deployed technology must be
      very strongly considered in our solution selection.

   The solution to the stated problem should address the following
   points:

   o Relation of the softwire protocols to other host mechanisms in the
      same layer of the network stack. Examples of mechanisms to
      consider are tunneling mechanisms, VPNs (Virtual Private
      Networks), mobility, multihoming (SHIM6 (Level 3 Shim for
      IPv6)),...

   o Operational brittleness introduced by softwire, e.g., potential
      single point of failure or difficulties to deploy redundant
      systems.

   o Effects of softwires on the transport layer. Issue like packet
      losses, congestion control, and handling of QoS (Quality of
      Service) reservation and usage of on-path protocols such as RSVP
      (Resource Reservation Protocol).

   The history of IPv4 and IPv6 transition strategies at the IETF is
   very long and complex. Here we list out some steps we have taken to
   further the effort and it has lead to the creation of this document
   and a few 'working rules' for us to accomplish our work:

   o At the IETF 63 "LightWeight Reachability softWires" (LRW) BOF
      meeting, attendees from several operators requested a very tight
      timeframe for the delivery of a solution, based on time-to-market
      considerations. This problem statement is narrowly scoped to
      accommodate near-term deployment.

   o At the Paris Softwires interim meeting in October, 2005,
      participants divided the overall problem space into two separate
      "sub-problems" to solve based on network topology. These two
      problems are referred to as "Hubs and Spokes" (described in
      Section 2) and "Mesh" (described in Section 3).

   As stated, there are two scenarios that emerged when discussing the
   traversal of networks composed of differing address families. The
   scenarios are quite common in today's network deployments. The
   primary difference between "Spokes and Hubs" and "Mesh" is how many
   connections and associated routes are managed by each IPv4 or IPv6
   "island". "Hubs and Spokes" is characterized with one connection and
   associated static default route, and "Mesh" is characterized by
   multiple connections and routing prefixes. In general, the two can
   be categorized as host or LAN connectivity and network (or VPN)
   connectivity problems. Looking at the history of multi-address
   family networking, the clear delineation of the two scenarios was
   never clearly illustrated but they are now the network norm, and both
   must be solved. Later, during the solution phase of the Work Group
   (WG), these problems will be treated as related, but separate,
   problem spaces. Similar protocols and mechanisms will be used when
   possible, but different protocols and mechanisms may be selected when
   necessary to meet the requirements of each given problem space.

1.1. Terminology

   Address Family (AF) - IPv4 or IPv6. Presently defined values for
   this field are specified in

   http://www.iana.org/assignments/address-family-numbers.

   Address Family Border Router (AFBR) – The router that interconnects
   two networks that use different address families.

   Customer Premise Equipment (CPE) – Under the scope of this document,
   this refers to terminal and associated equipment and inside wiring
   located at a subscriber's premises and connected with a carrier's
   communication channel(s) at the demarcation point ("demarc"). The
   demarc is a point established in a building or complex to separate
   customer equipment from telephone, cable, or other service provider
   equipment. CPE can be a host or router, depending on the specific
   characteristics of the access network. The demarc point for IPv4 may
   or may not be the same as the demarc point for IPv6, thus there can
   be one CPE box acting for IPv4 and IPv6 or two separate ones, one for
   IPv4 and one for IPv6.

   Home gateway – Existing piece of equipment that connects the home
   network to the provider network. Usually act as CPE for one or both
   address families.

   Softwire (SW) – A "tunnel" that is created on the basis of a control
   protocol setup between softwire endpoints with a shared point-to-
   point or multipoint-to-point state. Softwires are generally dynamic
   in nature (they may be initiated and terminated on demand), but may
   be very long-lived.

   Softwire Concentrator (SC) – The node terminating the softwire in the
   service provider network.

   Softwire Initiator (SI) – The node initiating the softwire within the
   customer network.

   Softwire Transport Header AF (STH AF) – the address family of the
   outermost IP header of a softwire.

   Softwire Payload Header AF (SPH AF) – the address family of the IP
   headers being carried within a softwire. Note that additional
   "levels" of IP headers may be present if (for example) a tunnel is
   carried over a softwire – the key attribute of SPH AF is that it is
   directly encapsulated within the softwire and the softwire endpoint
   will base forwarding decisions on the SPH AF when a packet is exiting
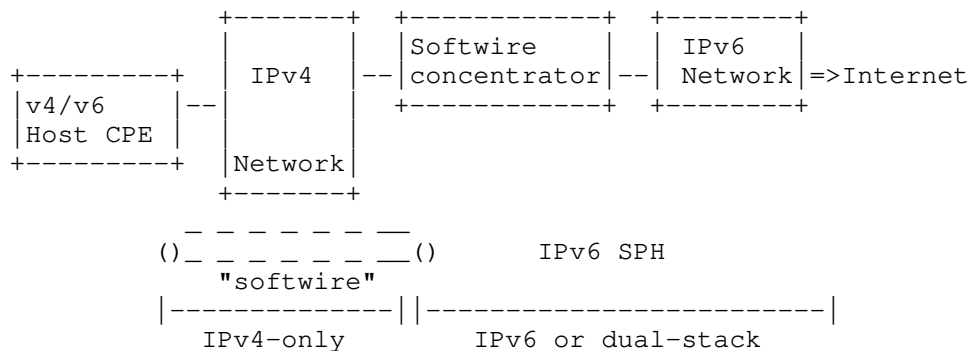   the softwire.

   Subsequent Address Family (SAF) – Additional information about the
   type of Network Layer Reachability Information (e.g., unicast or
   multicast).

2. Hubs and Spokes Problem

   The "Hubs and Spokes" problem is named in reference to the airline
   industry where major companies have established a relatively small
   number of well connected hubs and then serve smaller airports from
   those hubs.

   Manually configured tunnels (as described in [RFC4213]) can be a
   sufficient transition mechanism in some situations. However, cases
   where Network Address Translation (NAT) traversal is a concern (see
   Section 2.3), or dynamic IP address configuration is required,
   another solution is necessary.

   There are four variant cases of the "Hubs and Spokes" problem which
   are shown in the following figures.

```
                          +-------+  +------------+  +--------+
                          |       |  |Softwire    |  | IPv6   |
                          | IPv4  |--|concentrator|--| Network|=>Internet
          +---------+  |       |  +------------+  +--------+
          |v4/v6    |--|       |
          |Host CPE |  |       |
          +---------+  |Network|
                          +-------+
                    _ _ _ _ _ __
                ()_ _ _ _ _ __()        IPv6 SPH
                    "softwire"
                |--------------||-------------------------|
                    IPv4-only       IPv6 or dual-stack
```

   Case 1: IPv6 connectivity across an IPv4-only access network (STH).
   Softwire initiator is the host CPE (directly connected to a modem),
   which is dual-stack. There is no other gateway device. The IPv4
   traffic should not traverse the softwire.

                          Figure 1: Case 1

```
                      +-------+  +-------------+  +--------+
                      |       |  |  Softwire   |  |   v6   |
   +-----+  +------+  |  v4   |--| concentrator|--| Network|=>Internet
   |v4/v6|--|v4/v6 |--|       |  +-------------+  +--------+
   |Host |  |Router|  |Network|
   +-----+  |v4/v6 |  |       |
            | CPE  |  +-------+
            +------+

                      _ _ _ _ _ __
                  ()_ _ _ _ _ __()                      IPv6 SPH
                      "softwire"
   |--------------||--------------||-----------------------|
      Dual-stack       IPv4-only        IPv6 or dual-stack
```
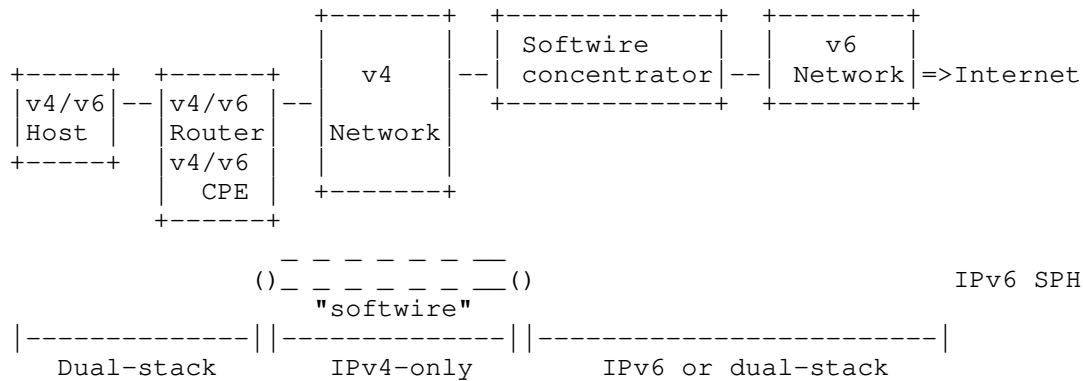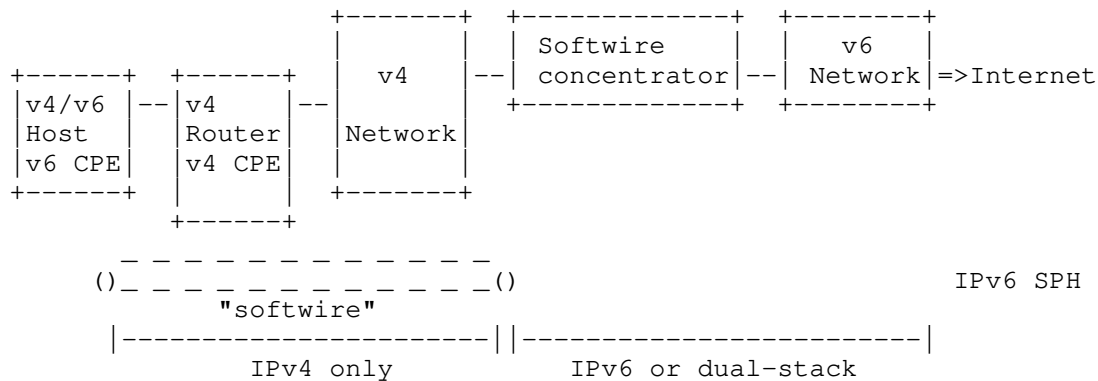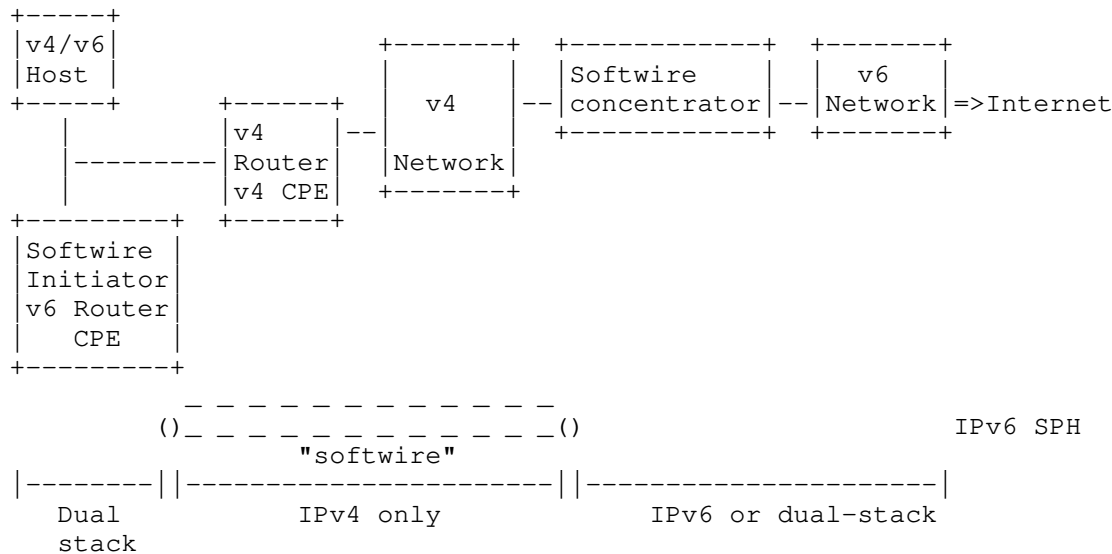
Case 2: IPv6 connectivity across an IPv4-only access network (STH).
Softwire initiator is the router CPE, which is a dual-stack device.
The IPv4 traffic should not traverse the softwire.

                      Figure 2: Case 2

```
                      +-------+  +-------------+  +--------+
                      |       |  |  Softwire   |  |   v6   |
   +------+  +------+ |  v4   |--| concentrator|--| Network|=>Internet
   |v4/v6 |--|v4    |-|       |  +-------------+  +--------+
   |Host  |  |Router|-|Network|
   |v6 CPE|  |v4 CPE| |       |
   +------+  |      | +-------+
             +------+

           _ _ _ _ _ _ _ _ _ _ _
       ()_ _ _ _ _ _ _ _ _ _ _()                        IPv6 SPH
              "softwire"
       |----------------------||-----------------------|
            IPv4 only                IPv6 or dual-stack
```

Case 3: IPv6 connectivity across an IPv4-only access network (STH).
The CPE is IPv4-only. Softwire initiator is a host, which act as an
IPv6 host CPE. The IPv4 traffic should not traverse the softwire.

                      Figure 3: Case 3

```
   +-----+
   |v4/v6|
   |Host |                +-------+ +------------+ +-------+
   +-----+      +------+   |       | |Softwire    | |  v6   |
      |         |v4    |-- |  v4   |--|concentrator|--|Network|=>Internet
      |---------|Router|   |       | +------------+ +-------+
      |         |v4 CPE|   |Network|
   +---------+  +------+   +-------+
   |Softwire |
   |Initiator|
   |v6 Router|
   |   CPE   |
   +---------+
               _ _ _ _ _ _ _ _ _ _ _
          ()_ _ _ _ _ _ _ _ _ _ _ _()              IPv6 SPH
                  "softwire"
   |-------||----------------------||--------------------|
     Dual           IPv4 only           IPv6 or dual-stack
     stack
```

   Case 4: IPv6 connectivity across an IPv4-only access network (STH).
   The routing CPE is IPv4-only. Softwire initiator is a device acting
   as an IPv6 CPE router inside the home network. The IPv4 traffic
   should not traverse the softwire.

                          Figure 4: Case 4

   The converse cases exist, replacing IPv4 by IPv6 and vice versa in
   the above figures.

2.1. Description

   In this scenario, carriers (or large enterprise networks acting as
   carriers for their internal networks) have an infrastructure that in
   at least one device on any given path supports only one address
   family, with customers who wish to support applications bound to an
   address family that cannot be routed end-to-end. The address family
   that must be "crossed" is called the Softwire Transport Header, or
   STH AF (which could be either IPv4 or IPv6).

   In order to support applications bound to another address family (the
   Softwire Payload Header Address Family, or SPH AF), it is necessary
   to establish a virtual dual-stack infrastructure (end-to-end),
   typically by means of automatically-established tunnels (Softwires).
   The traffic that can traverse the network via its native AF must not
   be forced to take the softwire path. Only the traffic that otherwise
   would not be able to be forwarded due to the AF mismatch should be

   forwarded within the softwire. The goal is to avoid overwhelming the
   softwire concentrator (SC).

   A network operator may choose to enable a single address family in
   one or several parts of this infrastructure for policy reasons (i.e.,
   traffic on the network is dominant in one of the address families, a
   single address family is used to lower Operations and Management
   (OAM) cost, etc.) or for technical reasons (i.e., because one or more
   devices are not able to support both address families).

   There are several obstacles that may preclude support for both
   address families:

   a) One or more devices (routers or some other media-specific
   aggregation point device) being used across the infrastructure (core,
   access) that supports only one address family. Typically the reasons
   for this situation include a lack of vendor support for one of the
   address families, the (perceived) cost of upgrading them, the
   (perceived) complexity of running both address families natively,
   operation/management reasons to avoid upgrades (perhaps temporarily),
   or economic reasons (such as a commercially insignificant amount of
   traffic with the non-supported address family).

   b) The home gateway (CPE router or other equipment at the demarc
   point), cannot be easily upgraded to support both address families.
   Typically the reason for this is the lack of vendor support for one
   of the address families, commercial or operational reasons for not
   carrying out the upgrade (i.e., operational changes and/or cost may
   need to be supported by the carrier for all the customers, which can
   turn into millions of units), or customer reluctance to change/
   upgrade CPE router (cost, "not broken, so don't change it"). Note
   that the impracticality of systematic upgrades of the CPE routers is
   also hindering the deployment of 6to4 based solutions [RFC3056] in
   IPv4 networks.

2.2. Non-Upgradable CPE Router

   Residential and small-office CPE equipment may be limited to support
   only one address family. Often, they are owned by a customer or
   carrier who is unwilling or unable to upgrade them to run in dual
   stack mode (as shown in Figure 3 and Figure 4).

   When the CPE router cannot run in dual-stack mode, a softwire will
   have to be established by a node located behind that CPE router.
   This can be accomplished either by a regular host in the home running
   softwire software (Figure 1 or Figure 3) or by a dedicated piece of
   hardware acting as the "IPv6 router" (Figure 4). Such a device is
   fairly simple in design and only requires one physical network

   interface. Again, only the traffic of the mismatched AF will be
   forwarded via the softwire. Traffic that can otherwise be forwarded
   without a softwire should not be encapsulated.

2.3. Network Address Translation (NAT) and Port Address Translation
      (PAT)

   A typical case of non-upgradable CPE router is a pre-existing IPv4/
   NAT home gateway, so the softwire solution must support NAT
   traversal.

   Establishing a Softwire through NAT or PAT must be supported without
   an explicit requirement to "autodetect" NAT or PAT presence during
   softwire setup. Simply enabling NAT traversal could be sufficient to
   meet this requirement.

   Although the tunneling protocol must be able to traverse NATs,
   tunneling protocols may have an optional capability to bypass UDP
   encapsulation if not traversing a NAT.

2.4. Static Prefix Delegation

   An important characteristic of this problem in IPv4 networks is that
   the carrier-facing CPE IP address is typically dynamically assigned.
   (The IP address of the node establishing the softwire behind the CPE
   router can also be dynamically assigned.)

   Solutions like external dynamic DNS and dynamic NAT port forwarding
   have been deployed to deal with ever changing addresses, but it would
   be simpler if, in IPv6 networks, a static prefix was delegated to
   customers. Such a prefix would allow for the registration of stable
   addresses in the DNS and enable the use of solutions like RFC 3041
   [RFC3041] privacy extension or cryptographically generated addresses
   (CGA) [RFC3972].

   The softwire protocol does not need to define a new method for prefix
   delegation; however, the Dynamic Host Configuration Protocol for IPv6
   (DHCPv6) prefix delegation [RFC3633] must be able to run over a
   softwire.

   Link local addresses allocated at both ends of the tunnel are enough
   for packet forwarding, but for management purpose like traceroute,
   global addresses can be allocated using existing protocols such as
   stateless address auto-configuration [RFC2462] or DHCPv6 [RFC3315].

   The IP addresses of the softwire link itself do not need to be
   stable, the desire for stability only applies to the delegated
   prefix. Even if there is a single node attached behind a softwire

   link, nothing prevents a softwire concentrator to delegate it a /64
   prefix.

   Similarly, in the case of an IPv4 softwire, the address could be
   provided by means of DHCP [RFC2131]. In the case of an IPv4
   softwire, a mechanism should be available in order to delegate an
   IPv4 prefix [SUBNET].

   Note about 6to4: This is one of the main differences between
   Softwires and 6to4. 6to4 addresses will change every time the CPE
   router gets a new external address, where a DHCPv6 delegated prefix
   through a softwire link could be stable.

2.5. Softwire Initiator

   In the "Hubs and Spokes" problem, softwires are always initiated by
   the customer side. Thus, the node hosting the end of the softwire
   within the customer network is called the softwire initiator. It can
   run on any dual-stack node. As noted earlier, this can be the CPE
   access device, another dedicated CPE router behind the original CPE
   access device, or actually any kind of node (host, appliance, sensor,
   etc.).

   The softwire initiator node can change over time and may or may not
   be delegated the same IP address for the softwire endpoint. In
   particular, softwires should work in the nomadic case (e.g., a user
   opening up his laptop in various Wi-Fi hot-spots), where the softwire
   initiator could potentially obtain an IP address of one address
   family outside its original carrier network and still want to obtain
   the other address family addresses from its carrier.

   If and when the IPv4 provider periodically changes the IPv4 address
   allocated to the gateway, the softwire initiator has to discover in a
   reasonable amount of time that the tunnel is down and restart it.
   This re-establishment should not change the IPv6 prefix and other
   parameters allocated to the site.

2.6. Softwire Concentrator

   On the carrier side, softwires are terminated on a softwire
   concentrator. A softwire concentrator is usually a dual-stack router
   connected to the dual-stack core of the carrier.

   A carrier may deploy several softwire concentrators (for example one
   per POP) for scalability reasons.

   Softwire concentrators are usually not nomadic and have stable IP
   addresses.

It may be the case that one of the address families is not natively
supported on the interface facing the core of the carrier.
Connectivity must then be provided by other tunnels, potentially
using the softwire mesh model.

Softwire concentrator functionality will be based on existing
standards for tunneling, prefixes, and addresses allocation,
management. The working group must define a softwire concentrator
architecture and interaction between these protocols and recommend
profiles. These recommendations must take into account the
distributed nature of the Softwires Concentrator in the provider
network and the impact on core IPv6 networks (for instance: prefix
aggregation).

2.7. Softwire Concentrator Discovery

The softwire initiator must know the DNS name or IP address of the
softwire concentrator. An automated discovery phase may be used to
return the IP address(s) or name(s) of the concentrator.
Alternatively, this information may be configured by the user, or by
the provider of the softwire initiator in advance. The details of
this discovery problem are outside the scope of this document,
however previous work could be taken in consideration. Examples
include: [SERVICE-DIS], [RFC4891], and [TUN-AD].

2.8. Scaling

In a "Hubs and Spokes model", a carrier must be able to scale the
solution to millions of softwire initiators by adding more hubs
(i.e., softwire concentrators).

2.9. Routing

As customer networks are typically attached via a single link to
their carrier, the minimum routing requirement is a default route for
each of the address families.

2.10. Multicast

Softwires must support multicast.

2.11. Security

2.11.1. Authentication, Authorization, and Accounting (AAA)

The softwire protocol must support customer authentication in the
control plane, in order to authorize access to the service, and
provide adequate logging of activity (accounting). However, a

carrier may decide to turn it off in some circumstances, for
instance, when the customer is already authenticated by some other
means, such as closed networks, cellular networks, etc., in order to
avoid unnecessary overhead.

The protocol should offer mutual authentication in scenarios where
the initiator requires identity proof from the concentrator.

The softwire solution, at least for "Hubs and Spokes", must be
integrable with commonly deployed AAA solutions (although extensions
to those AAA solutions may be needed).

2.11.2. Privacy, Integrity, and Replay Protection

The softwire Control and/or Data plane must be able to provide full
payload security (such as IPsec or SSL (Secure Socket Layer)) when
desired. This additional protection must be separable from the
tunneling aspect of the softwire mechanism itself. For IPsec,
default profiles must be defined. [RFC4891] provides guidelines on
this.

2.12. Operations and Management (OAM)

As it is assumed that the softwire may have to go across NAT or PAT,
a keepalive mechanism must be defined. Such a mechanism is also
useful for dead peer detection. However in some circumstances (i.e.,
narrowband access, billing per traffic, etc.) the keepalive mechanism
may consume unnecessary bandwidth, so turning it on or off, and
modifying the periodicity, must be supported administrative options.

Other needed OAM features include:

- Logging

- Usage accounting

- End-point failure detection (the detection mechanism must operate
   within the tunnel)

- Path failure detection (the detection mechanism must operate
   outside the tunnel)

2.13. Encapsulations

IPv6/IPv4, IPv6/UDP/IPv4, and IPv4/IPv6 are on the critical path for
"Hubs and Spokes" softwires. There is no intention to place limits
on additional encapsulations beyond those explicitly mentioned in
this specification.

3. Mesh Problem

3.1. Description

   We use the term "Mesh Problem" to describe the problem of supporting
   a general routed topology in which a backbone network that does not
   support a particular address family can be used as part of the path
   for packets that belong to that address family. For example, the
   path for an IPv4 packet might include a transit network that supports
   only IPv6. There might (or might not) be other paths that the IPv4
   packet could take that do not use the IPv6 transit network; the
   actual path chosen will be determined by the IPv4 routing procedures.

   By saying that the transit network supports only a single address
   family, we mean that the "core" routers of that network do not
   maintain routing information for other address families, and they may
   not even be able to understand the packet headers of other address
   families. We do suppose though that the core will have "edge
   routers" or "border routers", which maintain the routing information
   for both address families, and which can parse the headers of both
   address families. We refer to these as "Address Family Border
   Routers" (AFBRs).

   The following figure shows an AF2-only network connected to AF1-only
   networks, AF2-only networks, and dual stack networks. Note that in
   addition to paths through the AF2-only core, other paths may also
   exist between AF1 networks. The AFBRs that support AF1 would use BGP
   to exchange AF1 routing information between themselves, but such
   information would not be distributed to other core routers. The
   AFBRs would also participate in the exchange of AF2 routing
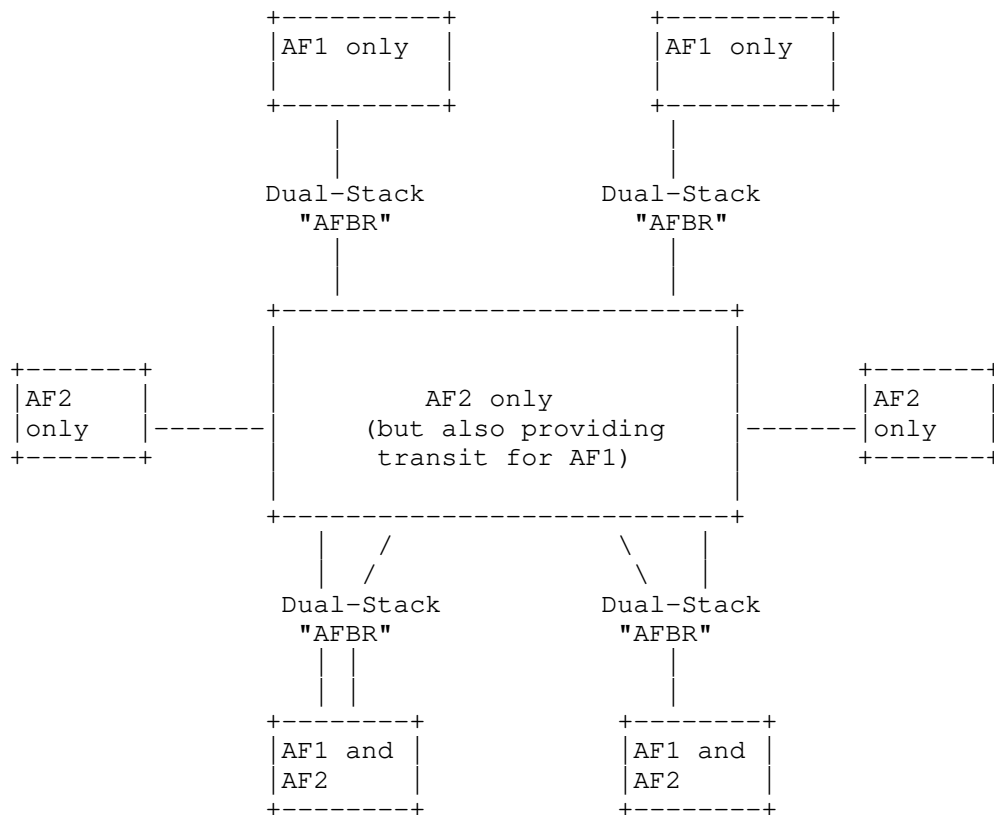   information with the core nodes.

```
            +----------+              +----------+
            |AF1 only  |              |AF1 only  |
            |          |              |          |
            +----------+              +----------+
                 |                         |
                 |                         |
            Dual-Stack                Dual-Stack
              "AFBR"                    "AFBR"
                 |                         |
                 |                         |
         +---------------------------+
         |                           |
 +-------+ |                         | |         +-------+
 |AF2    | |                         | |         |AF2    |
 |only   |-------|     AF2 only      |-------|only   |
 +-------+ |   (but also providing   | |         +-------+
         |     transit for AF1)      |
         |                           |
         +---------------------------+
             |   /            \   |
             |  /              \  |
         Dual-Stack           Dual-Stack
           "AFBR"               "AFBR"
             | |                  |
             | |                  |
         +--------+           +--------+
         |AF1 and |           |AF1 and |
         |AF2     |           |AF2     |
         +--------+           +--------+
```

                       Figure 5: Mesh Topology

   The situation in which a pair of border routers use BGP to exchange
   routing information that is not known to the core routers is
   sometimes known, somewhat misleadingly, as a "BGP-free core". In
   this sort of scenario, the problems to be solved are (a) to make sure
   that the BGP-distributed routing updates for AF1 allow a given AFBR,
   say AFBR1, to see that the path for a given AF1 address prefix is
   through a second AFBR, say AFBR2, and (b) to provide a way in which
   AFBR1 can send AF1 packets through the AF2-only core to AFBR2. Of
   course, sending AF1 packets through an AF2-only core requires the AF1
   packets to be encapsulated and sent through "tunnels"; these tunnels
   are the entities known as "softwires".

   One of the goals of the mesh problem is to provide a solution that
   does not require changes in any routers other than the AFBRs. This
   would allow a carrier (or large enterprise networks acting as carrier
   for their internal resources) with an AF2-only backbone to provide
   AF1 transit services for its clients, without requiring any changes

whatsoever to the clients' routers, and without requiring any changes
to the core routers. The AFBRs are the only devices that perform
dual-stack operations, and the only devices that encapsulate and/or
decapsulate the AF1 packets in order to send and/or receive them on
softwires.

It may be recognized that this scenario is very similar to the
scenario handled by the Layer 3 Virtual Private Network (L3VPN)
solution described in RFC 4364 [RFC4364]. The AFBRs correspond to
the "Provider Edge Routers" (PE) of RFC 4364. In those L3VPN
scenarios, the PEs exchange routing information in an address family
(e.g., the VPN-IPv4 address family), but they send VPN data packets
through a core which does not have the VPN routing information.
However, the softwire problem is NOT focused on the situation in
which the border routers maintain multiple private and/or overlapping
address spaces. Techniques which are specifically needed to support
multiple address spaces are in the domain of L3VPN, rather than in
the domain of Softwires.

Note that the AFBRs may be multiply connected to the core network,
and also may be multiply connected to the client networks. Further,
the client networks may have "backdoor connections" to each other,
through private networks or through the Internet.

3.2. Scaling

In the mesh problem, the number of AFBRs that a backbone network
supporting only AF2 will need is approximately on the order of the
number of AF1 networks to which it connects. (This is really an
upper limit, since a single AFBR can connect to many such networks).

An AFBR may need to exchange a "full Internet's" worth of routing
information with each network to which it connects. If these
networks are not VPNs, the scaling issues associated with the amount
of routing information are just the usual scaling issues faced by the
border routers of any network which is providing Internet transit
services. (If the AFBRs are also attached to VPNs, the usual L3VPN
scaling issues apply, as discussed in RFC 4364 [RFC4364] and RFC 4365
[RFC4365].) The number of BGP peering connections can be controlled
through the usual methods, e.g., use of route reflectors.

3.3. Persistence, Discovery, and Setup Time

AFBRs may discover each other, and may obtain any necessary
information about each other, as a byproduct of the exchange of
routing information (essentially in the same way that PE routers
discovery each other in L3VPNs). This may require the addition of
new protocol elements or attributes to existing protocols.

The softwires needed to allow packets to be sent from one AFBR to
another should be "always available", i.e., should not require any
extended setup time that would impart an appreciable delay to the
data packets.

3.4. Multicast

If the AF2 core does not provide native multicast services, multicast
between AF1 client networks should still be possible, even though it
may require replication at the AFBRs and unicasting of the replicated
packets through Softwires. If native multicast services are enabled,
it should be possible to use these services to optimize the multicast
flow.

3.5. Softwire Encapsulation

The solution to the mesh problem must not require the use of any one
encapsulation. Rather, it must accommodate the use of a variety of
different encapsulation mechanisms, and a means for choosing the one
to be used in any particular circumstance based on policy.

In particular, the solution to the mesh problem must allow for at
least the following encapsulations to be used: Layer 2 Tunneling
Protocol version 3 (L2TPv3), IP-in-IP, MPLS (LDP-based and RSVP-TE-
based), Generic Routing Encapsulation (GRE), and IPsec. The choice
of encapsulation is to be based on policy, and the policies
themselves may be based on various characteristics of the packets, of
the routes, or of the softwire endpoints themselves.

3.6. Security

In the mesh problem, the routers are not advertising routes for
individual users. So the mesh problem does not require the fine-
grained authentication that is required by the "Hub and Spoke"
problem. There should however be a way to provide various levels of
security for the data packets being transmitted on a softwire. The
softwire solution must support IPsec and an IPsec profile must be
defined (see recommendations in [USEIPSEC]).

Security mechanisms for the control protocols are also required. It
must be possible to protect control data from being modified in
flight by an attacker, and to prevent an attacker from masquerading
as a legitimate control protocol participant.

The verification of the reachability information exchanged and issues
surrounding the security of routing protocols themselves is outside
the scope of the specification.

4. Security Considerations

   Security considerations specific to the "Hubs and Spokes" and "Mesh"
   models appear in those sections of the document.

   As with any tunneling protocol, using this protocol may introduce a
   security issue by circumventing a site security policy implemented as
   ingress filtering, since these filters will only be applied to STH AF
   IP headers.

5. Principal Authors

   These are the principal authors for this document.

      Xing Li
      CERNET
      Room 225 Main Building, Tsinghua University
      Beijing 100084
      China

      Phone: +86 10 62785983
      Fax: +86 10 62785933
      Email: xing@cernet.edu.cn

      Alain Durand
      Comcast
      1500 Market st
      Philadelphia
      PA 19102
      USA

      Email: Alain_Durand@cable.comcast.com

      Shin Miyakawa
      NTT Communications
      3-20-2 TOC 21F, Nishi-shinjuku, Shinjuku
      Tokyo
      Japan

      Phone: +81-3-6800-3262
      Fax: +81-3-5365-2990
      Email: miyakawa@nttv6.jp

      Jordi Palet Martinez
      Consulintel
      San Jose Artesano, 1
      Alcobendas – Madrid
      E-28108 – Spain

      Phone: +34 91 151 81 99
      Fax: +34 91 151 81 98
      Email: jordi.palet@consulintel.es

      Florent Parent
      Hexago
      2875 boul. Laurier, suite 300
      Sainte-Foy, QC G1V 2M2
      Canada

      Phone: +1 418 266 5533
      Email: Florent.Parent@hexago.com

      David Ward
      Cisco Systems
      170 W. Tasman Dr.
      San Jose, CA 95134
      USA

      Phone: +1-408-526-4000
      Email: dward@cisco.com

      Eric C. Rosen
      Cisco Systems
      1414 Massachusetts Avenue
      Boxborough, MA, 01716
      USA

      Email: erosen@cisco.com

6. Contributors

   The authors would like to acknowledge the following contributors who
   provided helpful inputs on earlier versions of this document: Alain
   Baudot, Hui Deng, Francis Dupont, Rob Evans, Ed Koehler Jr, Erik
   Nordmark, Soohong Daniel Park, Tom Pusateri, Pekka Savola, Bruno
   Stevant, Laurent Totain, Bill Storer, Maria (Alice) Dos Santos, Yong
   Cui, Chris Metz, Simon Barber, Skip Booth, Scott Wainner, and Carl
   Williams.

   Tunnel-based transition mechanisms have been under discussion in the
   IETF for more than a decade. Initial work related to softwire can be
   found in RFC 3053 [RFC3053]. The earlier "V6 Tunnel Configuration"
   BOF problem statement [GOALS-TUN] a reasonable pointer to prior work.

7. References

7.1. Normative References

   [RFC3041]   Narten, T. and R. Draves, "Privacy Extensions for
                   Stateless Address Autoconfiguration in IPv6",
                   RFC 3041, January 2001.

   [RFC3053]   Durand, A., Fasano, P., Guardini, I., and D. Lento,
                   "IPv6 Tunnel Broker", RFC 3053, January 2001.

   [RFC3056]   Carpenter, B. and K. Moore, "Connection of IPv6
                   Domains via IPv4 Clouds", RFC 3056, February 2001.

   [RFC3972]   Aura, T., "Cryptographically Generated Addresses
                   (CGA)", RFC 3972, March 2005.

   [RFC4213]   Nordmark, E. and R. Gilligan, "Basic Transition
                   Mechanisms for IPv6 Hosts and Routers", RFC 4213,
                   October 2005.

7.2. Informative References

   [GOALS-TUN]   Palet, J., "Goals for Tunneling Configuration", Work
                   in Progress, February 2005.

   [RFC2131]   Droms, R., "Dynamic Host Configuration Protocol", RFC
                   2131, March 1997.

   [RFC2462]   Thomson, S. and T. Narten, "IPv6 Stateless Address
                   Autoconfiguration", RFC 2462, December 1998.

   [RFC3315]   Droms, R., Ed., Bound, J., Volz, B., Lemon, T.,
                   Perkins, C., and M. Carney, "Dynamic Host
                   Configuration Protocol for IPv6 (DHCPv6)", RFC 3315,
                   July 2003.

   [RFC3633]   Troan, O. and R. Droms, "IPv6 Prefix Options for
                   Dynamic Host Configuration Protocol (DHCP) version 6",
                   RFC 3633, December 2003.

   [RFC4364]   Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private
                   Networks (VPNs)", RFC 4364, February 2006.

   [RFC4365]   Rosen, E., "Applicability Statement for BGP/MPLS IP
                   Virtual Private Networks (VPNs)", RFC 4365,
                   February 2006.

   [RFC4891]   Graveman, R., Parthasarathy, M., Savola, P., and H.
                   Tschofenig, "Using IPsec to Secure IPv6-in-IPv4
                   Tunnels", RFC 4891, May 2007.

   [SERVICE-DIS]   Durand, A., "Service Discovery using NAPTR records in
                   DNS", Work in Progress, October 2004.

   [SUBNET]   Johnson, R., "Subnet Allocation Option", Work in
                   Progress, June 2007.

   [TUN-AD]   Palet, J. and M, "Analysis of IPv6 Tunnel End-point
                   Discovery Mechanisms", Work in Progress, January 2005.

   [USEIPSEC]   Bellovin, S., "Guidelines for Mandating the Use of
                   IPsec", Work in Progress, February 2007.

Authors' Addresses

   Xing Li (editor)
   CERNET
   Room 225 Main Building, Tsinghua University
   Beijing,   100084
   China

   Phone: +86 10 62785983
   Fax: +86 10 62785933
   EMail: xing@cernet.edu.cn

   Spencer Dawkins (editor)
   Huawei Technologies (USA)
   1700 Alma Drive, Suite 100

   Plano, TX   75075
   US

   Phone: +1 972 509 0309
   Fax: +1 469 229 5397
   EMail: spencer@mcsr-labs.org

   David Ward (editor)
   Cisco Systems
   170 W. Tasman Dr.
   San Jose, CA   95134
   US

   Phone: 1-408-526-4000
   EMail: dward@cisco.com

   Alain Durand (editor)
   Comcast
   1500 Market St
   Philadelphia, PA   19102
   US

   EMail: alain_durand@cable.comcast.com

Full Copyright Statement

Intellectual Property

Acknowledgement