Network Working Group   X. Li
Internet-Draft   CERNET
Expires: September 1, 2006   A. Durand
                                                             Comcast
                                                             D. Ward
                                                  Cisco Systems, Inc
                                                    S. Dawkins, Ed.
                                                            Futurewei
                                                   February 28, 2006

                      Softwire Problem Statement
                draft-ietf-softwire-problem-statement-01.txt

Status of this Memo

Copyright Notice

Abstract

   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. At the highest level, the Softwires Working Group
   is tasked to identify, and extend where necessary, standard protocols
   to support 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 as part of a solution phase
   following the completion of this document, within a relatively tight
   "time-to-market" as requested by operators at IETF 63. Some
   individual requirements (and non-requirements) are also identified in
   this document at times in order to better describe the specific scope
   for a given problem definition.

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 is describing the scenarios that the
   Working Group is going to focus on leading toward defining solutions.
   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.
      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/MPLS based networks to forward both
      unicast and multicast traffic.

   o Softwires are assumed to be non-ephemeral 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 startup of the Customer Premise Equipment (CPE)/Address Family
      Boundary 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 extending 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 history of IPv4 and IPv6 transition strategies at the IETF is a
   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 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 3) and "Mesh" (described in Section 4).

   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 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 RFC 1700 (see the Address Family Numbers
section).

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 familly.

Softwire (SW) - A "tunnel" that is created on the basis of a control
protocol setup between softwire endpoints with 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 the additional information about the type of the 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 establised a relatively small
number of well connected hubs and then serve smaller airports from
those hubs. There are four variant cases of the Hubs and Spokes
problem which are shown in Reference Diagram 1.

Reference Diagram 1

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.


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

```
                         +-------+  +------------+  +--------+
                         |       |  | 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 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.

```
                         +-------+  +------------+  +--------+
                         |       |  | 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 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.

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

             _ _ _ _ _ _ _ _ _ _ _ _
           ()_ _ _ _ _ _ _ _ _ _ _ _()                        IPv6 SPH
                    "softwire"
    |-------||----------------------||----------------------|
       Dual          IPv4 only           IPv6 or dual stack
       stack
```
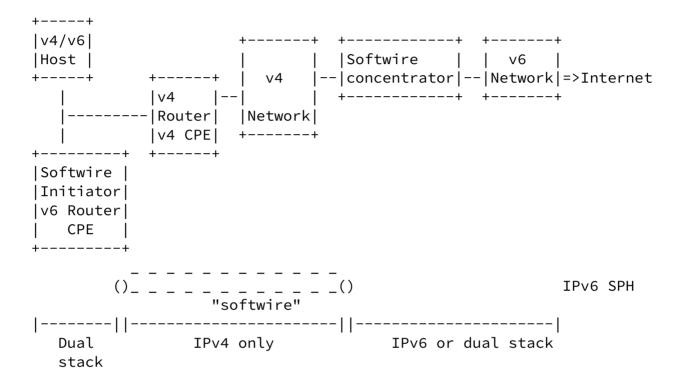
The converse cases exist, replacing IPv4 by IPv6 and vice versa
in the above figures.

Figure 1

2.1. Description

In this scenario, carriers (or large enterprise networks acting as
carriers for their internal networks) have an infrastructure which in
at least one device on any given path (core, access, home gateway)
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 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, 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").

2.2. Non-upgradable CPE router

Residential and small-office CPE equipment may be limited to support

only one address familly. 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 Case 2 and Case 3).

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 (Case 2) or by a dedicated piece of hardware acting
as the "IPv6 router" (Case 3). 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 softwires solution must support NAT
traversal.

If the NAT is not in the home gateway, but in carrier equipment
located at the other end of the access link (typically in an carrier
POP), support for NAT traversal is still required.

Establishing a softwire through NAT or PAT must work by default.
However, there is no requirement for explicitly "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.
Also, if the softwire has to be established from a node behind a CPE
router, that node IP address can also be dynamically assigned. In
cases where static IP addresses are unavailable, dynamic addresses

are a problem for some Internet accessible services. Solutions like
external dynamic DNS and dynamic NAT port forwarding have been
deployed, but it would be simpler if, in IPv6 networks, a static
prefix was delegated to the customer, even in the case of single node
network. That prefix would allow for the registration of stable
addresses in the DNS and also enough room to use either RFC3041
privacy extension or cryptographically generated addresses (CGA)
[RFC3972]. The softwire protocol does not need to define a new
method for prefix delegation however DHCPv6 prefix delegation must be
able to run over a softwire. Note also that the IP addresses of the
softwire link itself do not need to be stable, as, even in the single
PC being attached behind it, a /64 prefix will be delegated.

Link local addresses allocated at both ends of the tunnel are enough
for packets forwarding, but for management purpose like traceroute,
global addresses can be allocated using existing protocols such as

Neighbor Discovery or DHCPv6.

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 simply any kind of node (host, appliance, sensor, etc.).

The softwire initiator does not have to be always the same node
and/or always have been delegated the same IP address. In
particular, in the nomadic case (e.g. a user opening up his laptop in
various Wi-Fi hot-spots), 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 original carrier.

Nomadicity should be supported.

IPv4 provider can also periodically change the IPv4 address allocated
to the gateway. The softwire initiator has to discover in a
reasonable period of time that the tunnel is down and restart tunnel
establishment. 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. An carrier may deploy several softwire concentrators
(for example one per POP) for scalability reasons. A softwire
concentrator is in practice a dual-stack router connected to the
dual-stack core of the carrier or directly to the upstream providers.
Softwire concentrators are not nomadic and have stable IP addresses.
It may be the case that one of the address families is not natively
supported, even if this is not optimal, in the softwire concentrator,
but instead by means of tunnels to the upstreams (or other networks).

Softwire concentrator functionality will be based on existing
standards for tunneling, prefixes and addresses allocation,
management. The working group must define Best Current Practices for
Softwires 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 network (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 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: [I-D.durand-naptr-service-discovery], [I-D.ietf-v6ops-ipsec-
tunnels], and [I-D.palet-v6ops-solution-tun-auto-disc].

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). DNS redirection and/or local anycast
addresses among other choices, coupled with the (to-be-determined)

softwire concentrator discovery solution will enable sharing the load
among 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

Existing multicast solutions can be used over the softwire.
Typically, such solution would be either proxy Multicast Listener
Discovery or Internet Group Membership Protocol and Protocol-
Independent Multicast.

2.11. Security

2.11.1. Authentication, Authorization and Accounting

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, an
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) when desired. This
additional protection must be separable from the tunneling aspect of
the softwire mechanism itself. For IPsec, default profiles must be
defined. [draft-ietf-v6ops-ipsec-tunnels] 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. Other encapsulations, like IPv6/IPv6 or IPv4/IPv4, are nice to have but not on the critical path. There is no intention to place limits on additional encapsulations beyond those explicitly mentioned in this specification.

3. Mesh Problem

The "Mesh" problem is named in reference to typical routing problems in which there are more than one paths to a destination and a routing

protocol is needed to select the best path. It is also extremely
similar to the problems that the L3VPN Working Group is tackling in
which reduced, private and/or overlapping virtual routing and
forwarding tables are announced. The key difference is that the
reachability that must be announced across the transit core will
include more than VPN address family routes.

Reference Diagram 2

```
                    +----------+              +----------+
                    |differing |              |differing |
                    |AF as core|              |AF as core|
                    |access    |              |access    |
                    |island    |              |island    |
                    +----------+              +----------+
                         |                         |
                         |                         |
                    Dual-Stack                Dual-Stack
                      "AFBR"                    "AFBR"
                         |                         |
                         |                         |
                    +---------------------------+
                    |                         |
                    |                         |
      +-------+     |                         |     +-------+
      |same AF|     |      Single AF only     |     |same AF|
      |as core|-------|       transit core      |-------|as core|
      |access |     |                         |     |access |
      |network|     |                         |     |network|
      +-------+     |                         |     +-------+
                    |                         |
                    |                         |
                    +---------------------------+
                      |   /                \   |
                      |  /                  \  |
                    Dual-Stack                Dual-Stack
                      "AFBR"                    "AFBR"
                      | |                        |
                      | |                        |
                    +--------+                +--------+
                    |dual    |                |dual    |
                    |stack   |                |stack   |
                    |access  |                |access  |
                    |island  |                |island  |
                    +--------+                +--------+
```
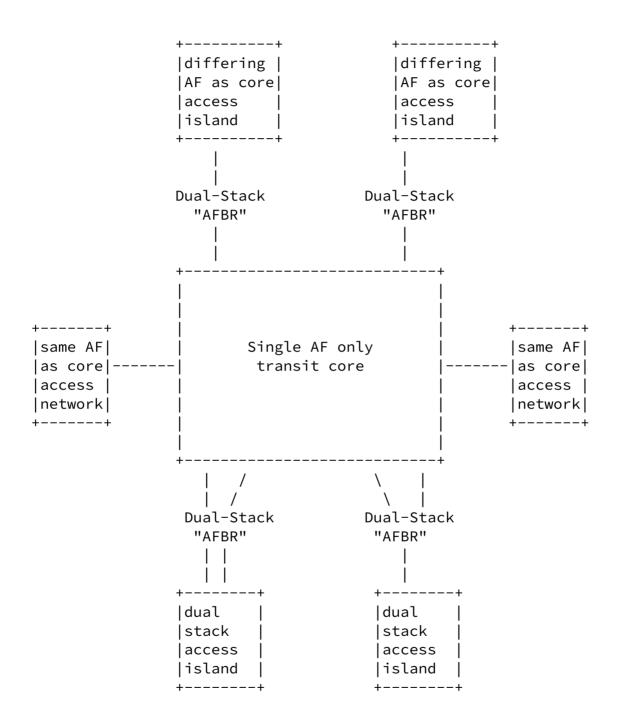
Figure 2

3.1. Mesh Description

In this problem, carriers (or large enterprise networks acting as
carrier for their internal resources) may be required to establish
connectivity to 'islands' of networks of one address family type

across a transit core of a differing address family type. For an

example, See Reference Diagram 2. To provide reachability across the
transit core, dual-stack devices are installed that act as "Address
Family Boundary Routers." These AFBRs can be performing peering
across autonomous systems or, performing as Provider Edge routers
(PE) in VPN parlance within an autonomous system. With respect to
deployment considerations, the islands do not have to be upgraded at
the time of deploying the transit core and interwork as if there was
no awareness of the AFBR.

The AFBRs are the only devices in the carrier's network that must be
able to perform dual-stack operations and setup and encapsulate
softwires in a mesh to the other islands. They then pass
reachability information as appropriate according to policy. They
may be multiply connected to the transit network and thus, have to be
able to exchange appropriate information and make a routing selection
choice as to the best exit point. Note that this creates multipoint-
to-point reachability using a point-to-point logical overlay of
softwire connectivity.

It should also be noted that the mesh problem can be considered as a
derivative of L3VPN, where the core provides transit in one address
family and the islands are connected via L3VPN of another address
family. This analogy only holds true if the islands can to be
represented as VPNs. In general, the diagrams frequently used for
L3VPNs is very similar. The key point is that the reachability
information that is to be exchanged must not be limited to VPNs or
any single AF or SAF or combination of AF/SAF. The solution must be
generic enough to carry any AF or SAF.

In the future a tunnel concentrator may be a different device than
the AFBR that is announcing reacability. In that future phase, the
AFBR may need to announce a third party tunnel concentrator.

3.2. Scaling

In the mesh problem, the number of AFBRs is on the order of the
number of islands though it should be clear that a single AFBR could
handle many islands if the islands have distinct routing and
forwarding tables. A primary issue in the Mesh problem is that the
size of the routing tables exchanged between the islands is of the

order of the 'full Internet' (with respect to the island's native
Address Family) plus any VPNs. These tables plus the routing tables
associated with the transit core (and VPNs of the same AF as the
transit core) must be stored on the AFBRs. The number of peering
points of an AFBR will be on the order of the number of Autonomous
System Border Routers (ASBRs), which are assumed to be multiply
peered to the transit core (multi-homed) for reliability. An island
can also have multiple AFBRs for reliability as well. Both the

island or the transit core may contain route reflectors or
hierarchical routing with impunity.

An AFBR should be able to pass route filters of data or routing
tables it does not wish to receive. Peering AFBRs must adhere to the
route or route table filters and not send reachability information.
Other attributes that can be sent from one AFBR to the other may
include "no export" or similar mechanisms to prevent subsequent
reannouncements of reachability information. The scaling of the
information to be exchanged is on the order of similar data exchanged
for L3VPNs.

3.3. Persistence, Discovery and Setup Time

Discovery of the AFBRs and softwire encapsulation could be
accomplished by the routing protocol during capability advertisement.
An alternative is that the endpoints could be passed in new data
formats or attributes, within a routing protocol.

The duration of the softwire for inter-island reachability is
considered to be as long as the duration of the peering session.
Thus, dynamicity is very low. The setup time should be on the order
of the same duration to setup L3VPNs.

3.4. Address Family/SAF Reachability

It has been reported that the softwires to connect the islands will
need to be able to perform IPv4/IPv6, IPv6/IPv4 and be able to
exchange multicast and VPN routing tables. The islands will need to
be able to perform multicast routing and if the transit core does not
provide native multicast services, the "classic" multicast solutions
can be used over the softwires. If native multicast services are
enabled, further work may need to be accomplished to optimize the

multicast forwarding path, receiver transmission load or receiver
load.

3.5. Softwire Encapsulation

In the strictest sense, the softwire encapsulation has to be dual
stack. There is no requirement that only one encapsulation technique
must be used. It could be possible to have more than one available
at each AFBR. The AFBR must be able to prioritize which
encapsulation technique it will use if there is more than one
available.

The encapsulations used to traverse the transit core must be enabled
to handle a choice of methods. Common choices that should create a
minimal set would include: L2TPv3, IP in IP, MPLS, IPsec, GRE. The

choice of encapsulation must not be subject to either an island or
peer-wise limitation. Different AF/SAF combinations must be able to
be encapsulated differently according to the requirements of the
network deployment. For example, IPv4 unicast may be encapsulated in
MPLS while IPv4 VPNs may be encapsulated in IPsec or L2TPv3. This
flexibility should not cause multiple peering sessions although it is
not precluded that this may be the desired network deployment. There
must be a scheme in which preferencing the encapsulation to be used
is exchanged between peers. Also, once the softwire encapsulation is
established a minimal amount of information must be passed with
reachability information to connect the AF/SAF reachability to
softwire. The linking of reachability information should not be
passed on a per route basis.

3.6. Security

In contrast with the hubs and spokes problem, routers are advertising
route for relatively large network islands, not individual users, so
fine-grained authentication is not necessary. However the solution
should support security of the softwire mechanism itself or the
softwire data plane or both.

In the softwire initialization mechanism, the softwire solution must
support authentication, but an carrier may decide to turn it off in
some circumstances. This means that if a routing protocol is used to
advertise the softwire encapsulation, it must also support

authentication.

   In the data plane, the softwire solution must support IPsec and an
   IPsec profile will must be defined. (see recommendations in
   [I-D.bellovin-useipsec]).

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

3.7. Operations and Management

   There have been no reports of NATs between the AFBRs (in the transit
   core) so a NAT detection solution is not needed.

   Other OAM needed features include:

   - Usage accounting

   - End-point failure detection (must be encapsulated within the tunnel
   in the transmitting direction)

   - Path failure detection

   Upon failure of a softwire, all reachability information must be
   withdrawn or a backup path used immediately.

3.8. Address Family Encapsulations

   IPv6/IPv4, IPv4/IPv6 and overlapping address space as defined in the
   L3VPN working group (including overlapping RFC-1918 private address
   space) are on the critical path for "Mesh" softwires. Other
   encapsulations, like IPv6/IPv6, IPv4/IPv4 or IP-only LAN Service
   (IPLS) as defined in the L2VPN working group, are nice to have but
   not on the critical path.There is no intention to place limits on
   additional encapsulations beyond those explicitly mentioned in this
   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](#). IANA Considerations

   There are no IANA actions requested in this specification.

. Changes from -00

    1.   Individual-draft authors moved to Authors section, and added an
         acknowledgements section.

    2.   Detailed mailing list comments from Alain Baudot (2005/12/20).

    3.   Detailed mailing list comments from Pekka Savola (2005/12/22).

    4.   Detailed mailing list comments from Laurent Toutain
         (2005/12/26).

    5.   Detailed mailing list comments from Francis Dupont (editorial)
         (2005/12/29).

    6.   Detailed mailing list comments from Francis Dupont (non-
         editorial) (2005/12/29).

    7.   Detailed mailing list comments from Tom Pusateri (2005/12/29).

    8.   Detailed mailing list comments from Tom Alain Durant
         (2005/12/30).

    9.   Changed all occurances of "HGW" to "CPE" and added definitio

    10. Removed all occurances of "TEP" (which seemed to be a synonym
         for concentrator anyway).

    11. Changed all occurances of "ISP" to be "operator".

    12. Removed all [RFC 2119](#) language from the specification (since it's
         a problem statement).

    13. Further linguistic clarificatons and edits (2006/01 and 02)

    14. Remove Compare and Contrast section after discussion w/ authors
         (2006/02/19)

7. Acknowledgements

7.1. Authors

   The principal authors for this document are: Xing Li, Alain Durand,
   Shin Miyakawa, Jordi Palet Martinez, Florent Parent, and David Ward.

7.2. 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.

   The authors would also like to acknowledge the participants in the
   Softwires interim meeting in Paris, France (October 11-12, 2005)
   (minutes are at
   http://bgp.nu/~dward/softwires/InterimMeetingMinutes.htm).

   The authors would also like to express a special acknowledgement and
   thanks to Mark Townsley. Without his leadership, persistence,
   editing skills and thorough suggestions for the document; we would
   have not have been successful.

   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 RFC3053. The earlier "V6 Tunnel Configuration" BOF problem
   statement [I-D.palet-v6tc-goals-tunneling] includes a reasonable
   pointer to prior work.

8. References

8.1. Normative References

[RFC1700] Postel, "Key words for use in RFCs to Indicate Requirement
           Levels", BCP 14, RFC 1700, October 1994.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
           Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2663] Srisuresh, P. and M. Holdrege, "IP Network Address
           Translator (NAT) Terminology and Considerations",
           RFC 2663, August 1999.

[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.

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

8.2. Informative References

[I-D.bellovin-useipsec]
           S, "Guidelines for Mandating the Use of IPsec,
           draft-bellovin-useipsec-04", September 2005.

[I-D.durand-naptr-service-discovery]
           A, ""Service Discovery using NAPTR records in DNS",
           draft-durand-naptr-service-discovery-00", October 2004.

[I-D.ietf-v6ops-ipsec-tunnels]

              P, ""Using IPsec to Secure IPv6-in-IPv4 Tunnels",
              draft-ietf-v6ops-ipsec-tunnels-01", August 2005.

   [I-D.palet-v6ops-solution-tun-auto-disc]
              J, ""IPv6 Tunnel End-point Automatic Discovery Mechanism",
              draft-palet-v6ops-solution-tun-auto-disc-01",
              October 2004.

   [I-D.palet-v6ops-tun-auto-disc]
              J and M, ""Analysis of IPv6 Tunnel End-point Discovery
              Mechanisms", draft-palet-v6ops-tun-auto-disc-03",
              January 2005.

   [I-D.palet-v6tc-goals-tunneling]
              J, ""Goals for Tunneling Configuration",
              draft-palet-v6tc-goals-tunneling-00", February 2005.

Authors' Addresses

    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

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

Email: dward@cisco.com

Spencer Dawkins (editor)
Futurewei Technologies
1700 Alma Drive, Suite 100
Plano, TX   75075
US

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

Intellectual Property Statement

   The IETF takes no position regarding the validity or scope of any
   Intellectual Property Rights or other rights that might be claimed to
   pertain to the implementation or use of the technology described in
   this document or the extent to which any license under such rights
   might or might not be available; nor does it represent that it has
   made any independent effort to identify any such rights. Information
   on the procedures with respect to rights in RFC documents can be
   found in BCP 78 and BCP 79.

   Copies of IPR disclosures made to the IETF Secretariat and any
   assurances of licenses to be made available, or the result of an
   attempt made to obtain a general license or permission for the use of

Acknowledgment