Network Working Group
Internet-Draft
Expires: June 10, 2006

X. Li
CERNET
A. Durand
Comcast
S. Miyakawa
NTT Communications
J. Palet
Consulintel
F. Parent
Hexago
D. Ward
Cisco Systems

December 7, 2005

Softwire Problem Statement
draft-ietf-softwire-problem-statement-00.txt

Status of this Memo

Copyright Notice

Abstract

This document defines problem statements for the Softwire Working Group to solve. At the highest level, the softwire WG is tasked to identify, and extend where necessary, standard protocols to support a selected set of IPv4 in IPv6 and IPv6 in IPv4 transition problems. This document describes the distinct problems that will be solved as part of a solution phase following the completion of this document. 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. Requirements Notation

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 [RFC2119].

2. Introduction

The Softwires Working Group is specifying the standardization of discovery, control and encapsulation methods for connecting IPv4 networks across IPv6 networks, IPv6 networks across IPv4 networks in a way that will encourage multiple, inter-operable vendor implementations.

An important aspect of the problem to keep in mind is that softwires are to be used in IP based networks to forward both unicast and multicast trafic. They are also assumed to be non-ephemeral in nature thus, they are peristent or long-lived. Last, the setup time of a softwire is expected to be a very small fraction of the total setup time of the CPE/Address Family Boundry Router (AFBR)

At the Paris softwire 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 "Hub and Spoke" (Described in Section 4) and "Mesh" (Described in Section 5). The primary difference between these two problems are how many connections and associated routes are managed by each IPv4 or IPv6 island. Hub and Spoke is characterized with one connection and associated static default route, and Mesh is characterized by multiple connections and routing prefixes. During the solution phase of the WG, these problems will be treated as related, but separable, problem spaces. Similar protocols and mechanisms will be used when necessary, but may vary when necessary to optimize for the requirements of the given problem space.

2.1. Terminology

Address Family - IPv4 or IPv6

AFBR - Address Family Boundry Router (aka PE)

CPE - Customer Premisis equipment (Host, small router, or "modem")

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 brought up and down on demand from any side of the softwire), but may be very long-lived.

The node hosting the end of the softwire within the customer network is called the softwire initiator.

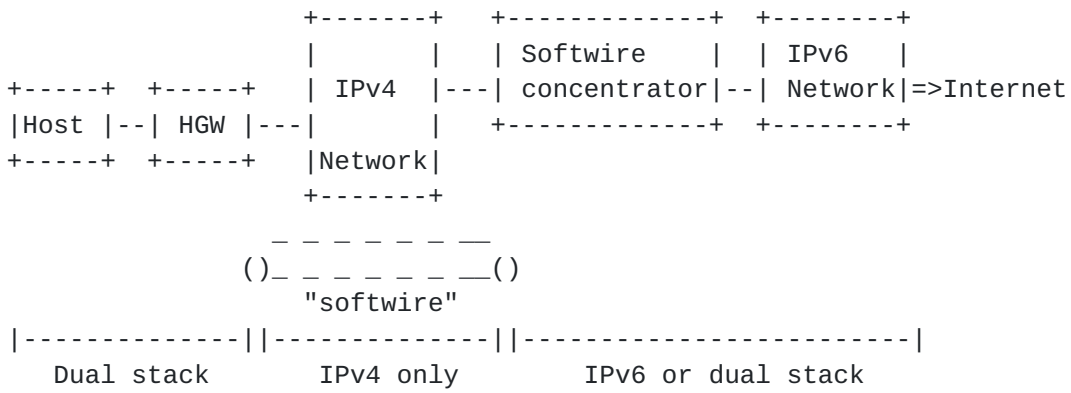
The node hosting the end of the softwire within the ISP network is called the softwire concentrator.

3. Hubs and Spokes Problem

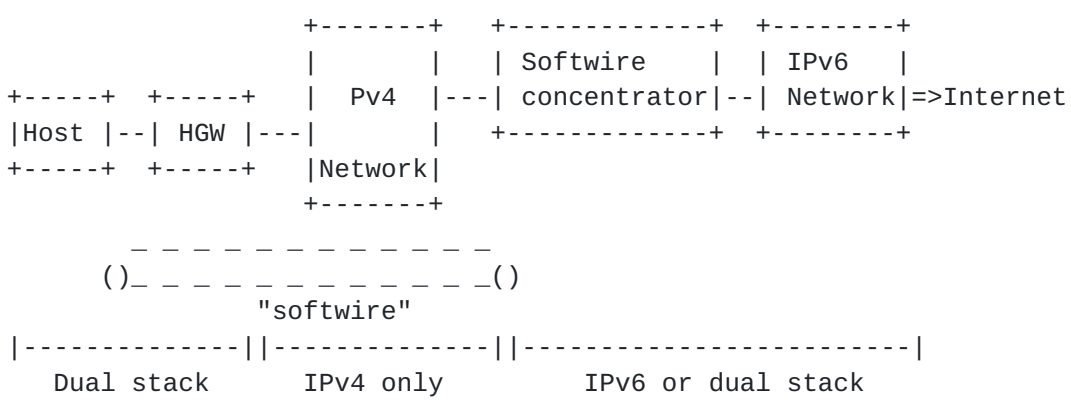
The "Hubs and Spokes" problem is named in reference to the airline industry where major companies have establised a relatively small number of well connected hubs and then deserve smaller airports from those hubs. There are three cases refer to Hubs and Spokes problem which are shown in Diagram 1.

Reference Diagram 1

Case 1: IPv6 connectivity across an IPv4 access network. Softwire initiator is the home gateway.



Case 2: IPv6 connectivity across an IPv4 access network. Softwire initiator is a host.



Case 3: IPv6 connectivity across an IPv4 access network. Softwire initiator is a device acting as an IPv6 router inside the home network.

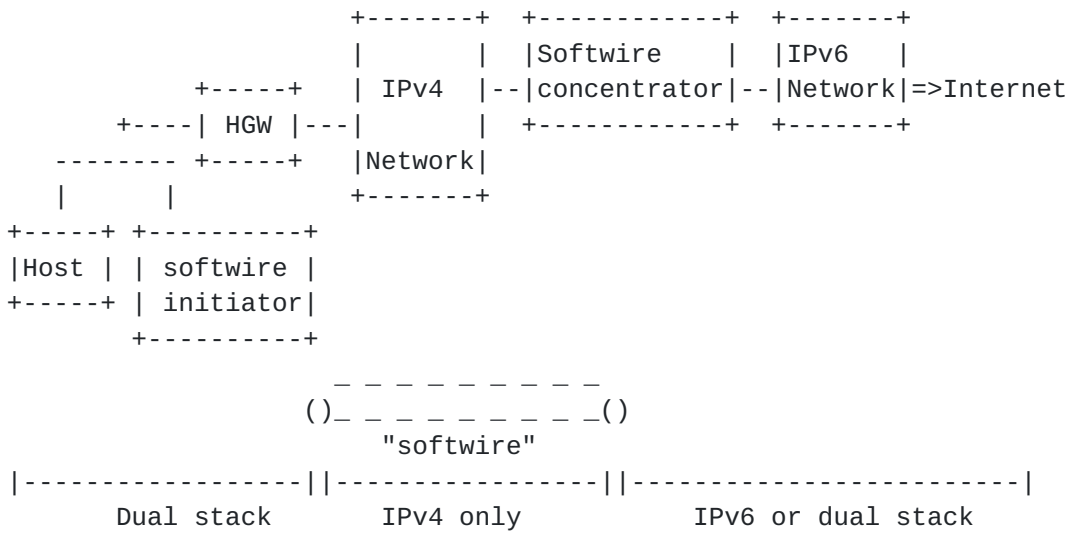


Figure 1

3.1. Description

In this problem, ISPs (or large enterprise networks acting as ISP for their internal resources) establish a dual stack core (either natively or by running tunnels, potentially managed by softwires in a "Mesh" problem) and a number of dual stack Points of Presence (POP) where they connect their customers. However, one or two things may happen:

a) the networks between the CPE router and the POP supports only one address family.

b) the CPE router cannot be easily upgraded to support both address families.

Equipment cost, operational cost, complexity of running a dual-stack network, reluctance to touch CPE, etc. are all reasons brought forward when asked why the invervening network cannot be dual-stack throughout.

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

When connecting IPv6 islands through IPv4 networks, it is assumed that one or more IPv4 NAT/PATs MAY exist on the intervening IPv4 network. At this point in time, neither IPv6 NAT nor IPv6 PAT has

been defined, so no special consideration will be made for those cases.

There is no requirement to be able to "autodetect" NAT or PAT presence during softwire setup.

3.3. Non upgradable CPE router

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 PC in the home running some ad-hoc software or by a dedicated piece of hardware acting as the "IPv6 router". Such a device is fairly simple in design and only requires one physical network interface.

3.4. Static Prefix Delegation

An important characteristic of this problem in IPv4 networks is that the ISP-facing CPE IP address is typically dynamically assigned. Also, if the softwire has to be establish 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 netwroks, 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). 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.

Similarly, in the case of an IPv4 softwire, the address could be provided by means of DHCP.

3.5. Softwire Initiator

In the Hub and Spoke 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 a simple dual stack host or a local dual stack router. As noticed earlier, this can be the CPE access router, another dedicated CPE router behind the CPE access router or simply a host.

The softwire initiator does not have to be always the same node

and/or always have the same IP address. In particular, in the nomadic case (e.g. a user opening up his laptop in various wifi hot-spots), the softwire initiator could potentially obtain an IP address of one address family outside its original ISP network and still want to obtain the other address family addresses from its original ISP.

3.6. Softwire Concentrators

On the ISP side, softwires are termintated on a softwire contentrator. An ISP may deploy several concentrators (for example one per POP) for scaling reasons. A concentrator is in practice a dual stack router connected to the dual stack core ISP infrastructure. Softwire concentrators are not nomadic and have fixed IP addresses.

3.7. Softwire Concentrator Discovery

When the initiator of the softwire is a CPE, the IP address or DNS hostname of the softwire concentrator must be known. The simplest way for this to be known by the CPE is for it to be configured by the user, or by the provider of the CPE in advance. Alternatively, an automated discovery phase may be run in order to return the IP address(s), or hostname(s) of the concentrator. The details of this discovery problem are outside the scope of this document.

3.8. Scaling

In a hub and spoke model, an ISP MUST scale the solution to millions of softwire inititators by adding more hubs (i.e. softwire concentrator).

3.9. Routing

As customers networks are typically attached via a single link to their ISP, a default or static route is the only thing that is needed for both address families.

3.10. Multicast

The "classic" multicast solutions can be used over the softwire. Typically, such solution would be either proxy MLD/IGMP and PIM.

NOTE: need to add a reference to "classic" multicast.

3.11. Security

3.11.1. Authentication, Authorization and Accounting

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

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

3.11.2. 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.

3.12. Operations and Management (OAM)

As it is assume that the softwire may have to go accross NAT or PAT, a keepalive mechanism MUST be define. Such a mechanism is also useful for dead peer detection. However it may consume unnecessary bandwidth, so turning it on or off MUST be an administrative option.

Other OAM needed features include:

- Usage accounting

- End-point failure detection (must be encapsulated w/in the tunnel in the transmitting direction

- Path failure detection)

3.13. Encapsulations

IPv6/IPv4, IPv6/UDP/IPV4 and IPv4/IPv6 are on the critical path for softwires. Other encapsulations, like IPv6/IPv6 or IPv4/IPv4, are nice to have but not on the critical path.

4. Mesh Problem

The "Mesh" problem in named in reference to typical routing problems.

Reference Diagram 2

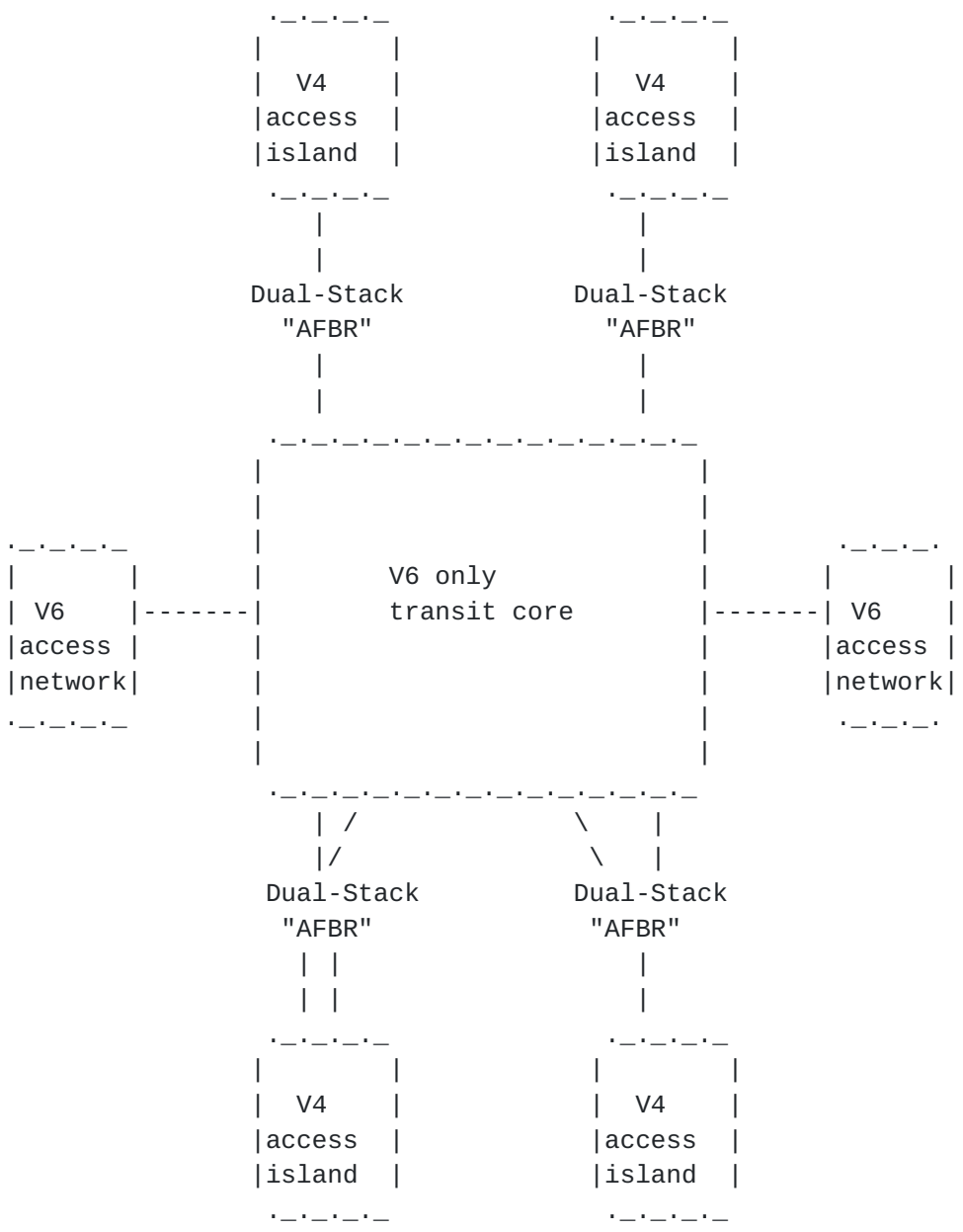


Figure 2

4.1. Mesh Description

In this problem, ISPs (or large enterprise networks acting as ISP for their internal resources) 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 Figure 1. Note that this is just an example and the converse AF problem may exist. 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) within an autonomous system. 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 AFBR's are the only devices in the 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 informations and make a routing selection choice as to the best exit point. Note that this creates a multipoint to point reachability but, in essence a point to point logical overlay of softwire connectivity.

It should be noted that according to reports the islands do not want to achieve network connectivity via tunneled Layer 2 mechanisms but, as distinct Layer 3 or MPLS routers. This clearly helps scaling and Layer 2 discovery performance issues. It also prevents having to have fully meshed point to point Layer 2 connectivity between the nodes in differing islands as Layer 2 technology choice must be preserved.

It should also be noted that the mesh problem can be treat as a special case of L3VPN, where the core provides transit in one address family and the islands are connected via L3VPN of another address family.

4.2. Scaling

In the mesh problem, the number of AFBRs is on the order of the number of islands though it should be clear that an AFBR could handle many islands if they 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 islands native AF) plus, VPNs. The number of peering points of an AFBR will be on the order of any Autonomous System Border Router (ASBR) which are assumed to be multiply peered to the transit core for reliability. An island can

also have multiple AFBRs for reliability as well. Both the island or the transit core can contain route reflectors or hierarchical routing with impunity.

4.3. Persistence, Discovery and Setup Time

Discovery of the AFBRs and softwire encapsulation can be accomplished by the routing protocol during capability advertisement. Or, the endpoints can be passed in new data formats or attributes, yet to be defined. The duration of the softwire for inter-island reachability is considered to be as long as the peering session. Thus, dynamicity is very low. The setup time should be on the order of the same duration to setup L3VPNs.

4.4. AF/SAF Reachability

It has been reported that the softwires to connect the islands will need to be able to perform IPv4 in IPv6, IPv6 in IPv4 and be able to exchange L3VPN 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 softwire. 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.

4.5. 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.

4.6. Security

In contrast with the hub and spoke problem, routers are advertizing routers for relatively large islands, and never a single user so there is no "user authentication" necessary. However, if running over an untrusted network, control or data plane security may be necessary.

In the control plane, the softwire solution has to support authentication, but an ISP may decide to turn it off in some circumstances.

In the data plane, the softwire solution must support IPsec and an IPsec profile will have to be defined. (see Steve Bellovin

recomendations)

4.7. OAM

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 w/in the tunnel in the transmitting direction

- Path failure detection)

4.8. Encapsulations

IPv6/IPv4,IPv4/IPv6 and overlapping address space as defined in the L3VPN working group are on the critical path for softwires. Other encapsulations, like IPv4/IPv4 or IPLS as defined in the L2VPN working group, are nice to have but not on the critical path.

5. Problems: Contrast & Compare

An important distinction between the "Hub & Spokes" and " Mesh" problems is that the former defines client-initiated tunnels and the "spoke" is a device on the client premises (and may be owned by the client). The latter discusses about provider-initiated tunnels, and the devices participating in the mesh are on the provider premises and owned/managed by the provider.

6. Security Considerations

The softwire protocol (control plane) must support user authentication, and should be configurable to offer mutual authentication. The authentication mechanism has to be compatible with existing AAA protocols. The authentication must be protected using an already defined secure mechanism.

The data plane is the established tunnel which transports IPv6 over an IPv4 network, or IPv4 over an IPv6 network. This tunnel can be protected using IPsec. Guidelines on this are described in [draft-ietf-v6ops-ipsec-tunnels].

As with any tunneling protocol, using this protocol may introduce a security issue by circumventing a site security policy implemented as ingress filtering.

7. References

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

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

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

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.