Mitigation of Breaking Connections

(a.k.a. OLSRd v1 Multi-Gateway & BRDP)

About Me

- Ferry Huberts
- Self-Employed Open Source Entrepreneur

- Lead Committer for OLSRd v1

- Committer in several other projects
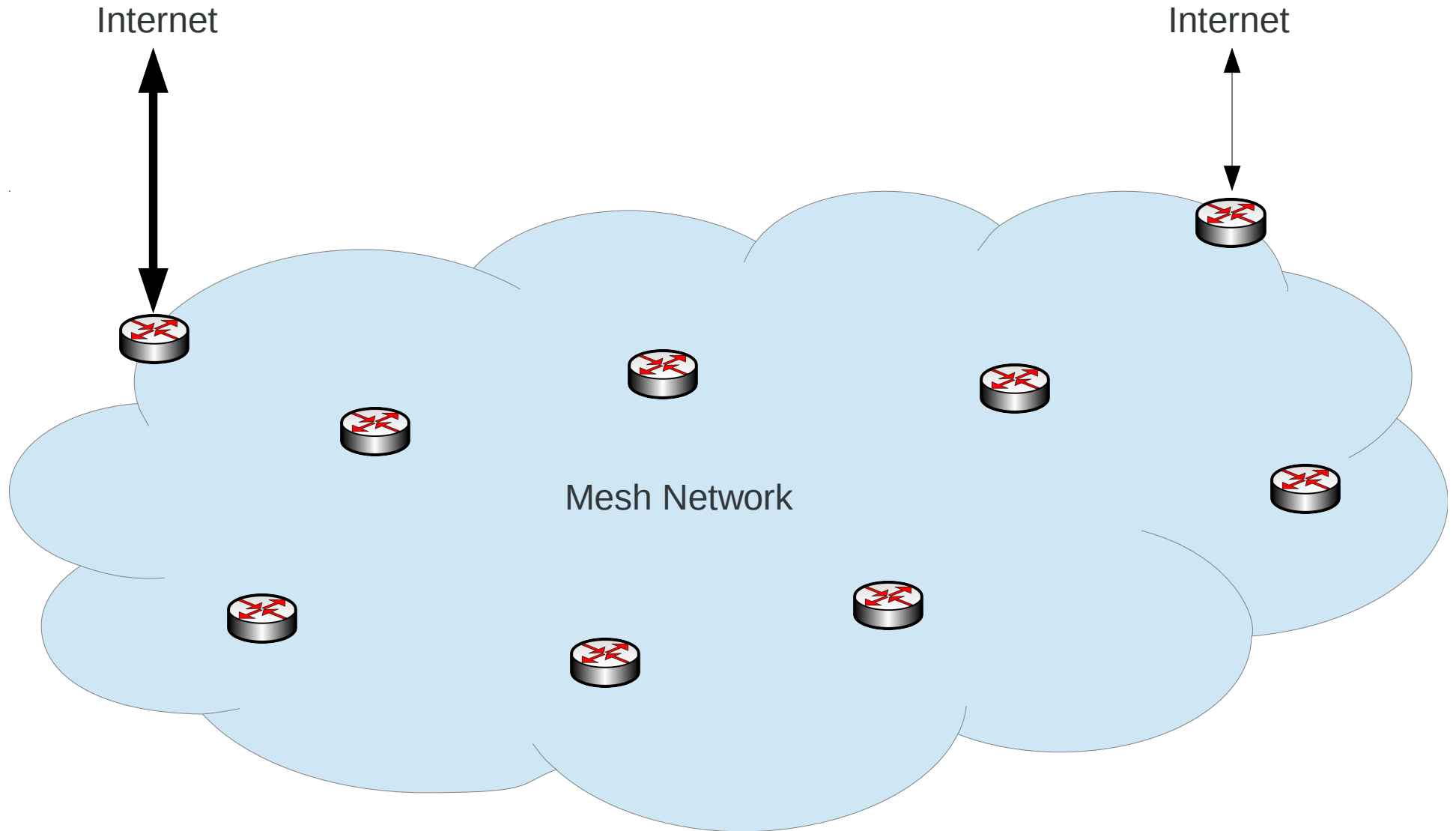 - Mainly OSGi projects

Smart Gateway

Dafuq?
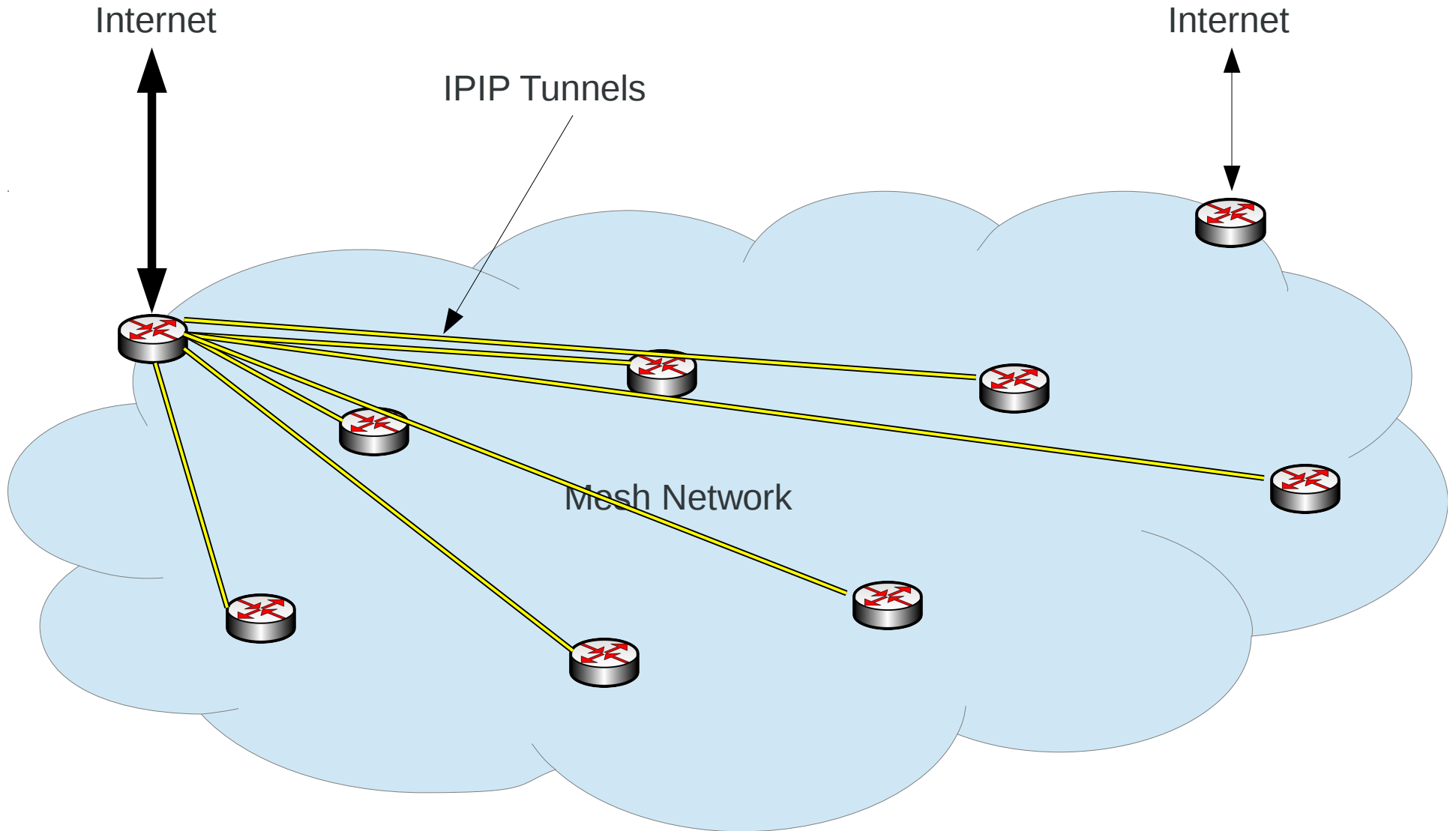
- Original Behaviour:
  - Choose fattest exit link
  - Only select when chosen 6 times in a row
  - Only select **once**
  - Even if it's a uni-directional link
  - Even if it's (or has become) unreachable
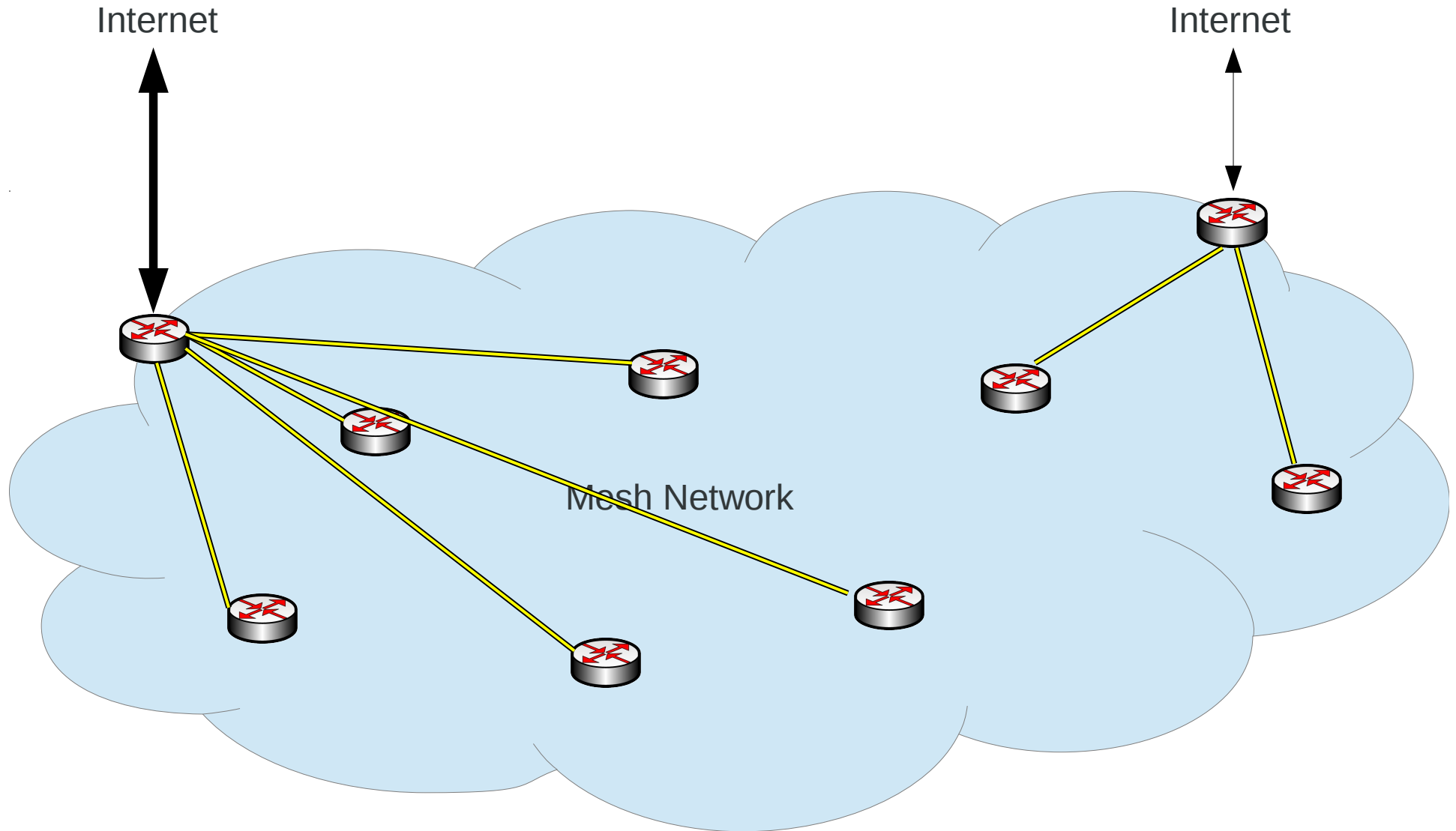
- No so smart...

Smart Gateway

- Fixed
  - Ignore gateways with uni-directional links
  - Ignore unreachable gateways
  - Many more small fixes
- Modified
  - Choose **nearest** fattest exit link
  - Only select when chosen <configurable> times in a row
  - Allow continuous selection
- Added
  - Optionally use a threshold when choosing
  - Optionally use link weighing

- A bit smarter...

Smart Gateway

- Multi-Gateway
    - Allow <configurable> gateways in parallel
    - Continuously select a new (best) gateway
    - Do not take down existing tunnels
    - Use policy routing for connection pinning

Smart Gateway

Internet

Newly selected
gateway tunnels

Previously selected
gateway tunnels

Internet

Mesh Network
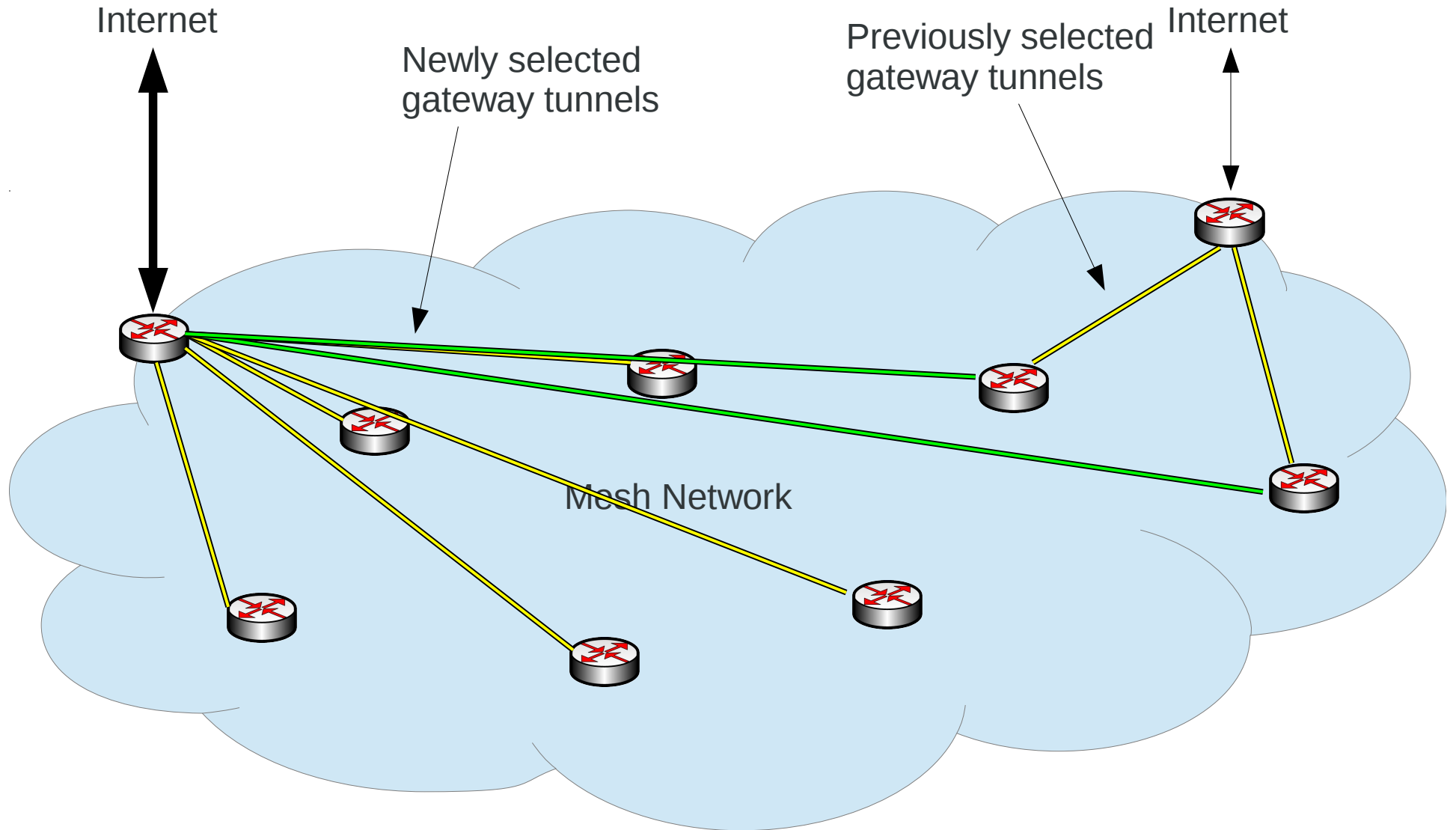
Smart Gateway

- Much smarter :-)

Pelagic

- Issue:
  - Tunnels not preemptively taken down (yet)

- Summary:
  - **Mitigates** the 'breaking connections' problem
  - Does **not** solve it

So how do we **solve**
the 'breaking connections' problem?

BRDP

Border Router Discovery Protocol

BRDP

- ## IETF HomeNet RFC

  – https://datatracker.ietf.org/doc/draft-boot-homenet-brdp/

  BRDP enables multi-homing for small to medium sites, including Homenets, using Provider Aggregatable **IPv6** addresses.

I E T F®

Let's bind existing stuff together

GLUE

Perimeter detection

Ingress filtering

Neighbor discovery

Address Configuration SLAAC, DHCP

Routing

Prefix delegation

Address selection

Multipath transport
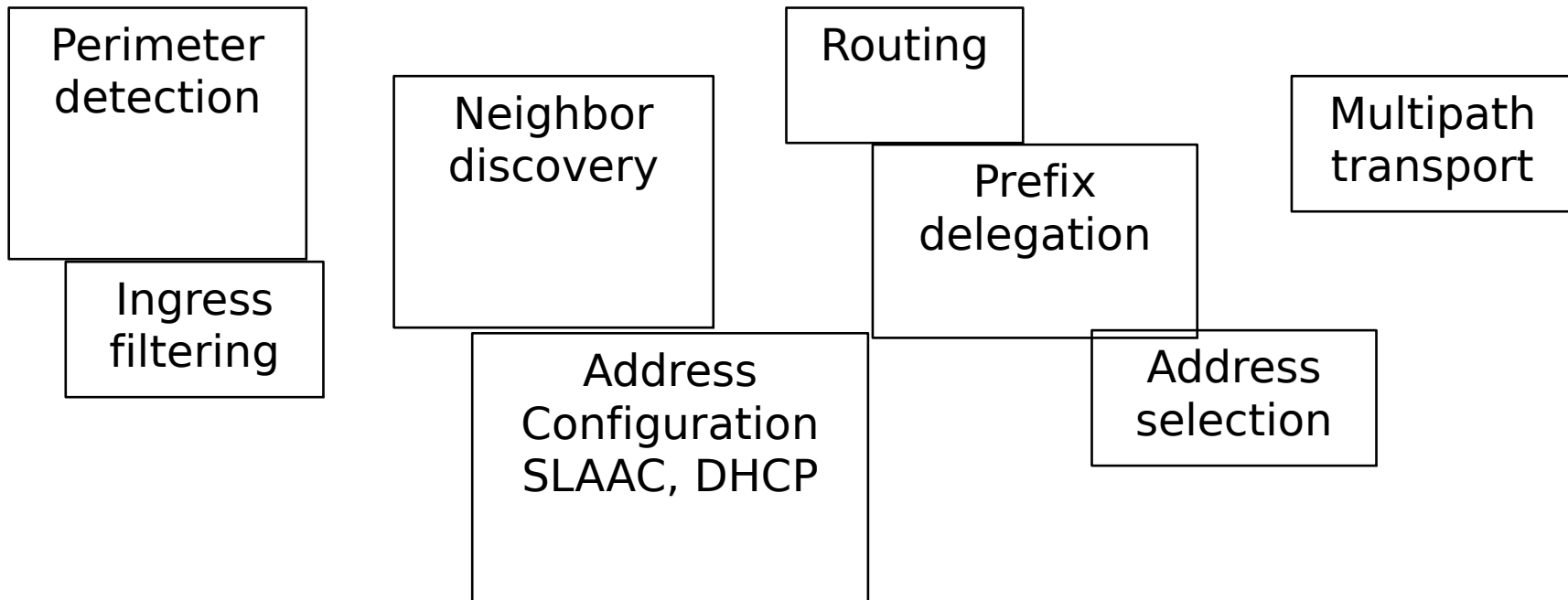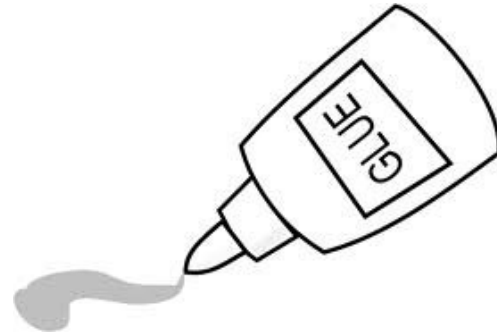
With BRDP, Border Router address and ISP delegated prefix is disseminated in the edge network

Border Routers start sending Border Router Information Options (BRIO), for each prefix. Routers cache and forward the BRIOs downstream.
Hosts may use the BRIO; Neighbor Discovery (ND) is used.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Length     | Prefix Length |D|  reserved  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Sequence Number           |    Hopcount    |  reserved   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Uniform Path Metric                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       reserved                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                              |
+              ( Border Router prefix )                        +
|                                                              |
+              Border Router Address                           +
|                                                              |
+         ( remainder of Border Router Address )               +
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
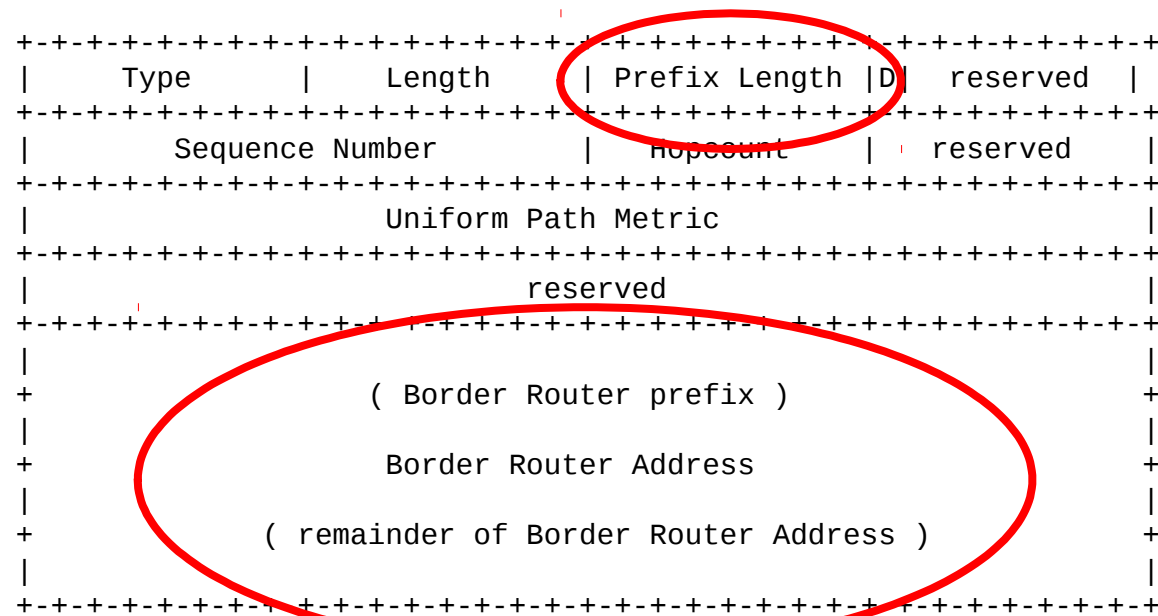
Border Router acts as DHCP server when D-flag is set

Nodes can get IP configuration options for each prefix. Routers can get (/64) prefixes for each subnet with DHCP-PD.
Maybe similar approach for name resolving, with N-flag, for BR as DNS server. Or use DNS server address provided by DHCP.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type      |      Length      |  Prefix Length |D| reserved  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Sequence Number       |   Hopcount   |  reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Uniform Path Metric                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        reserved                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                          |
+               ( Border Router prefix )                   +
|                                                          |
+                Border Router Address                     +
|                                                          |
+        ( remainder of Border Router Address )            +
|                                                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Other fields: for BRIO forwarding and hints for costs to Internet DFZ

BRIOs are distributed downstream, like distance vector.
Uniform Path Metric field can be used for address selection.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type         |       Length      | Prefix Length |D|  reserverd   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Sequence Number          |    Hopcount      |   reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Uniform Path Metric                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          reserved                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                                |
+              ( Border Router prefix )                          +
|                                                                |
+              Border Router Address                             +
|                                                                |
+         ( remainder of Border Router Address )                 +
|                                                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
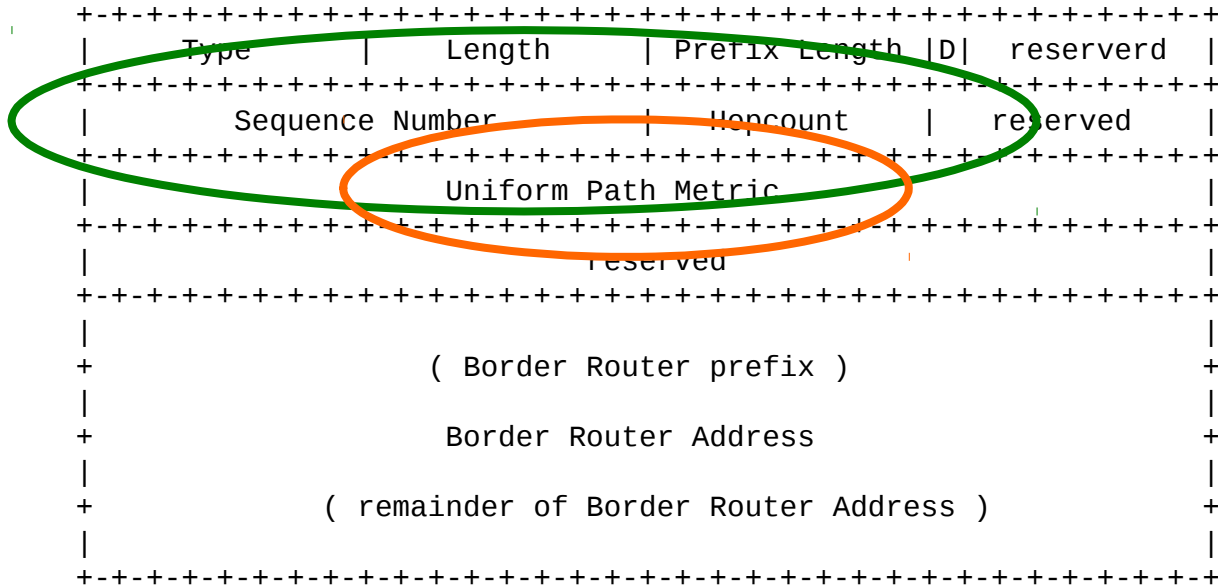
BRDP based routing

Forwarding heuristic slightly updated (related to RFC 1812):
  0) Removal of ::/0 route
  1) Forward based on destination address (as normal)
  2a) If no next-hop found, forward based on source address, with
              lookup in BRIO-cache
  2b) If BR found, forward to next-hop for that BR (rule 1)

This eliminates ingress filter problems, packets are directed to topologically correct Border Router. It also provides distributed ingress filtering.

Now, hosts and routers cooperate for multi-path support.

Basics:
3.2.2.1. A: Single ISP, Single CER, Internal routers

Service Provider Router

Customer Internet Connection

IPv6 Customer Edge Router

Network A

Network B (E)

H1

H2

Link F

H3

H4

Network E (B)

IPv6 Interior Router

Network C

Network D

H5

H6

H7

H8
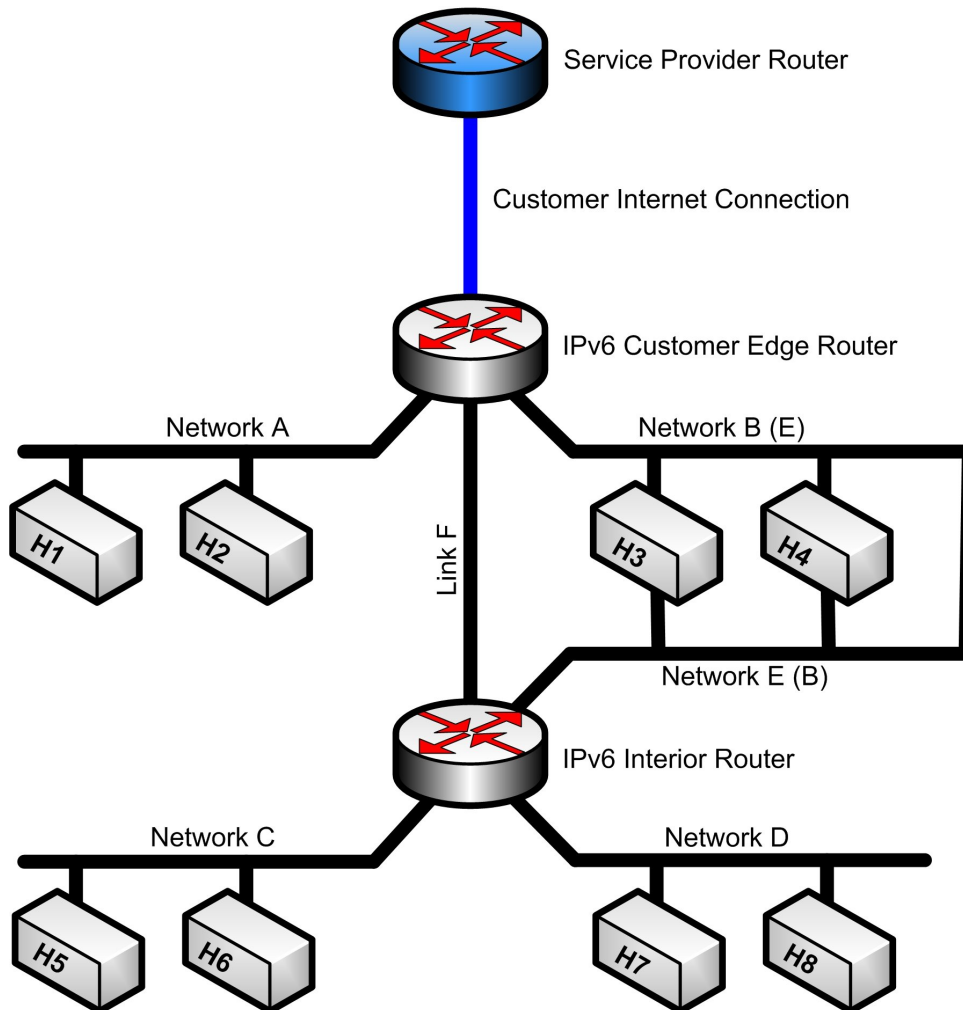
BRDP operation explained using Network topology model A (*Home Networking Architecture for IPv6*)

Assume Customer Edge Router has discovered Customer Internet Connection, and has assigned a prefix.

1) Customer Edge Router configures subnets on its interfaces
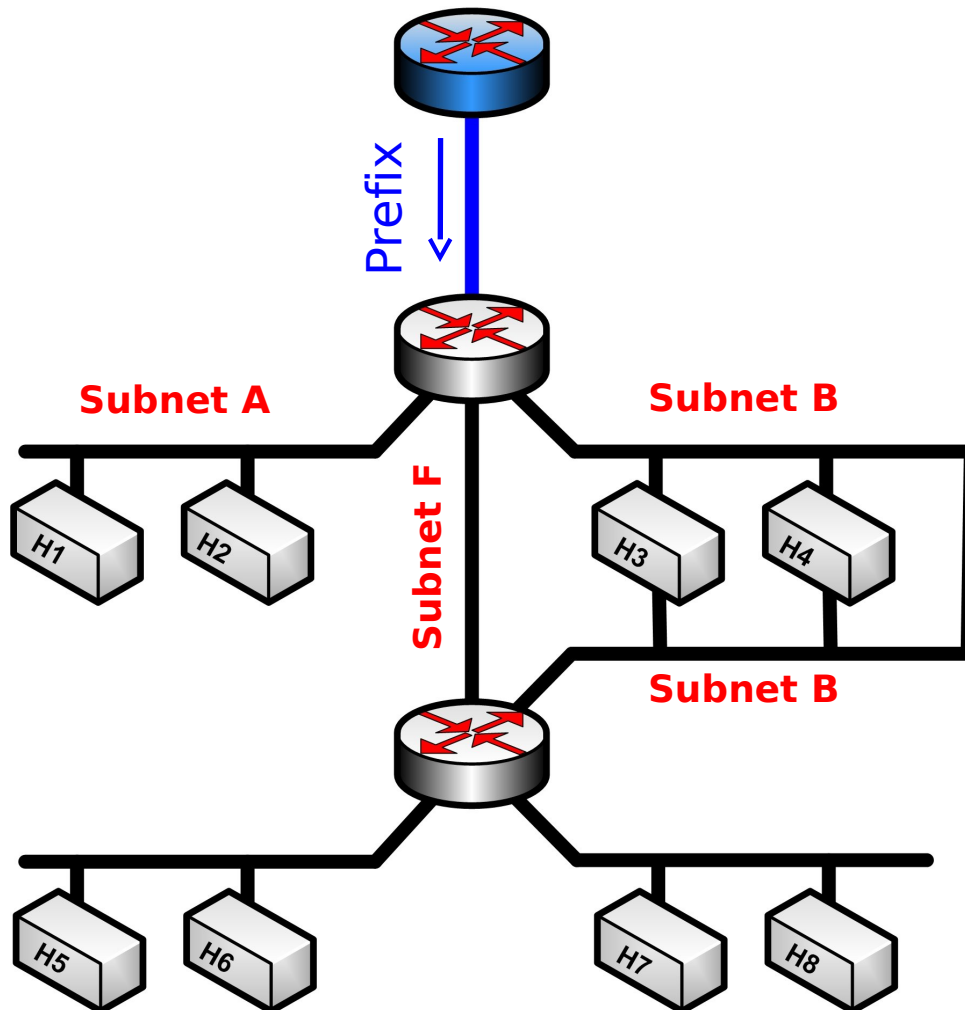
Customer Edge Router has three Homenet interfaces, it configures subnets on these and send out ND RA PIO.

Nodes on subnets A, B and F can get IP addresses and use the Customer Internet Connection.

Nothing special here…

2) Customer Edge Router starts sending BRIO
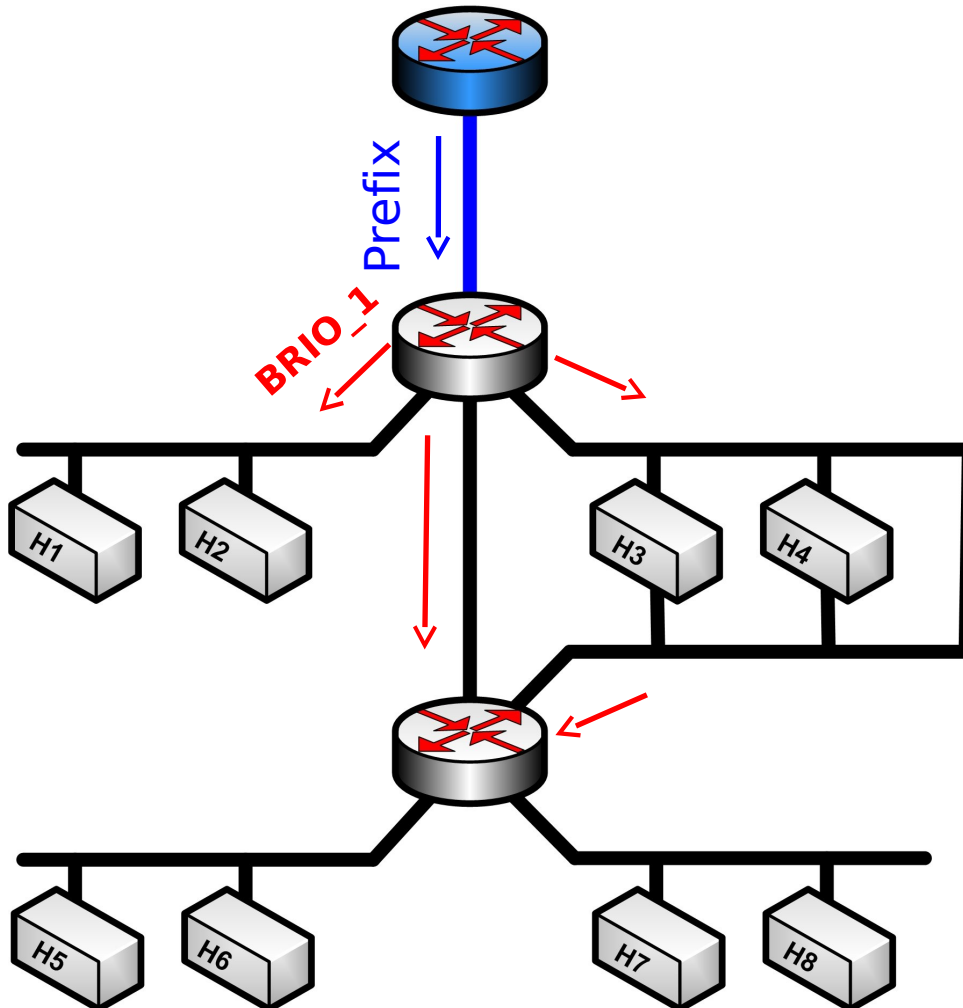
Customer Edge Router sends out its assigned prefix with ND RA BRIO.

Customer Interior Router learns how to get prefixes for its attached interfaces that do not have a subnet already.

Assume D-flag is set and Customer Edge Router acts as DHCP server, including DHCP-PD.

3) Customer Interior Router configures subnets on its interfaces
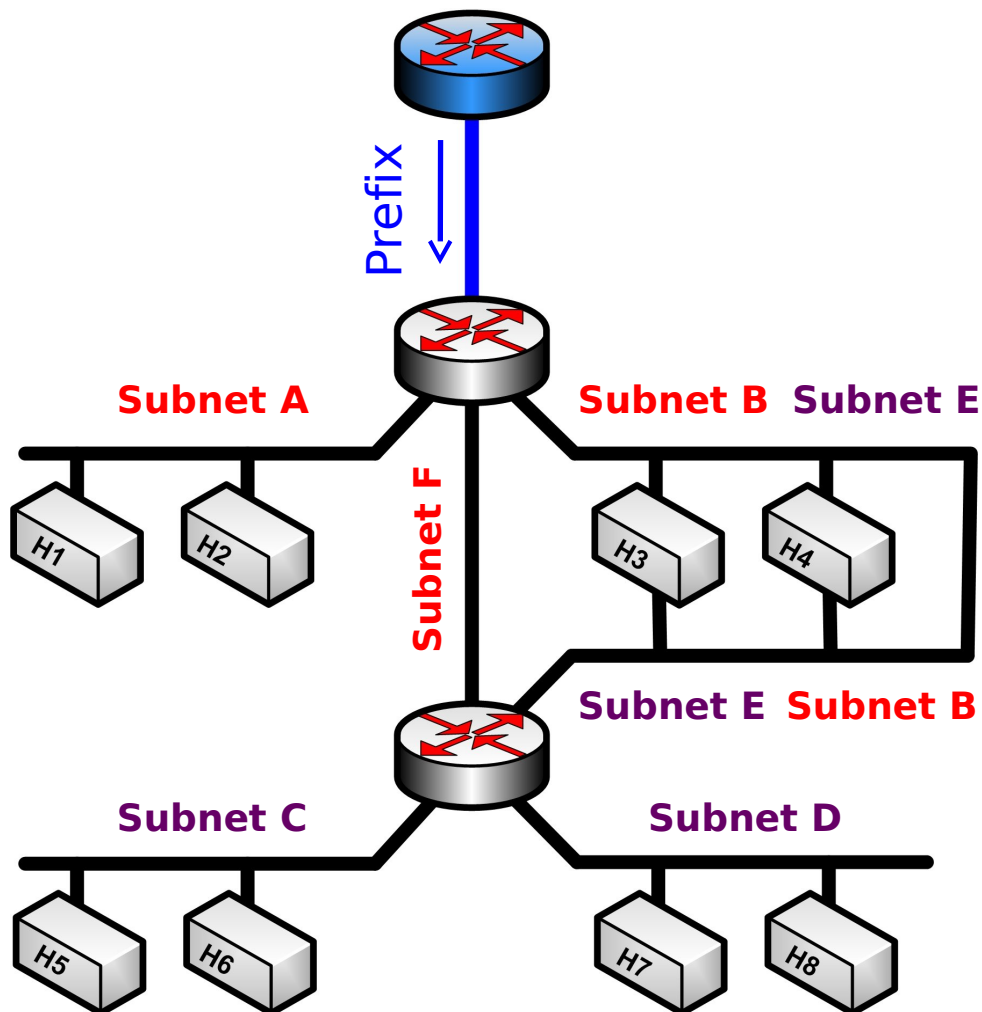
Customer Interior Router has two interfaces, which do not have a subnet configured already. It configures subnets on it and sends out ND RA PIO

Nodes on subnets C and D can get IP addresses and use Customer Internet Connection.

There is no need to configure subnet E on interface with subnet B, because both subnets B and E would be in same BRIO prefix. For the example, subnet E is configured. To be discussed.

4) Customer Interior Router starts sending BRIO
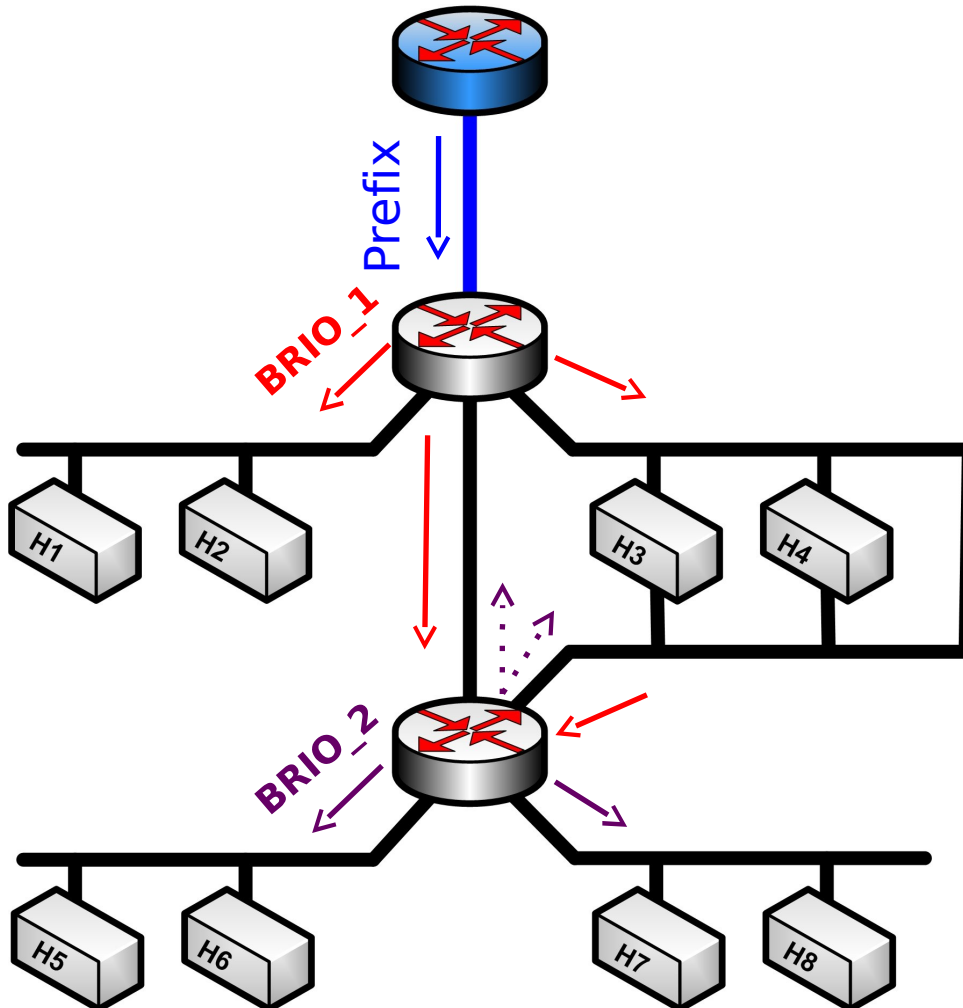
IPv6 Customer Interior Router sends BRIO information with its ND RA.

Hop-count and Uniform Path Metric fields are adjusted.

Option for forwarding back on same interface (needed on MANET interfaces)

5) Interior nodes select IP addresses as source address
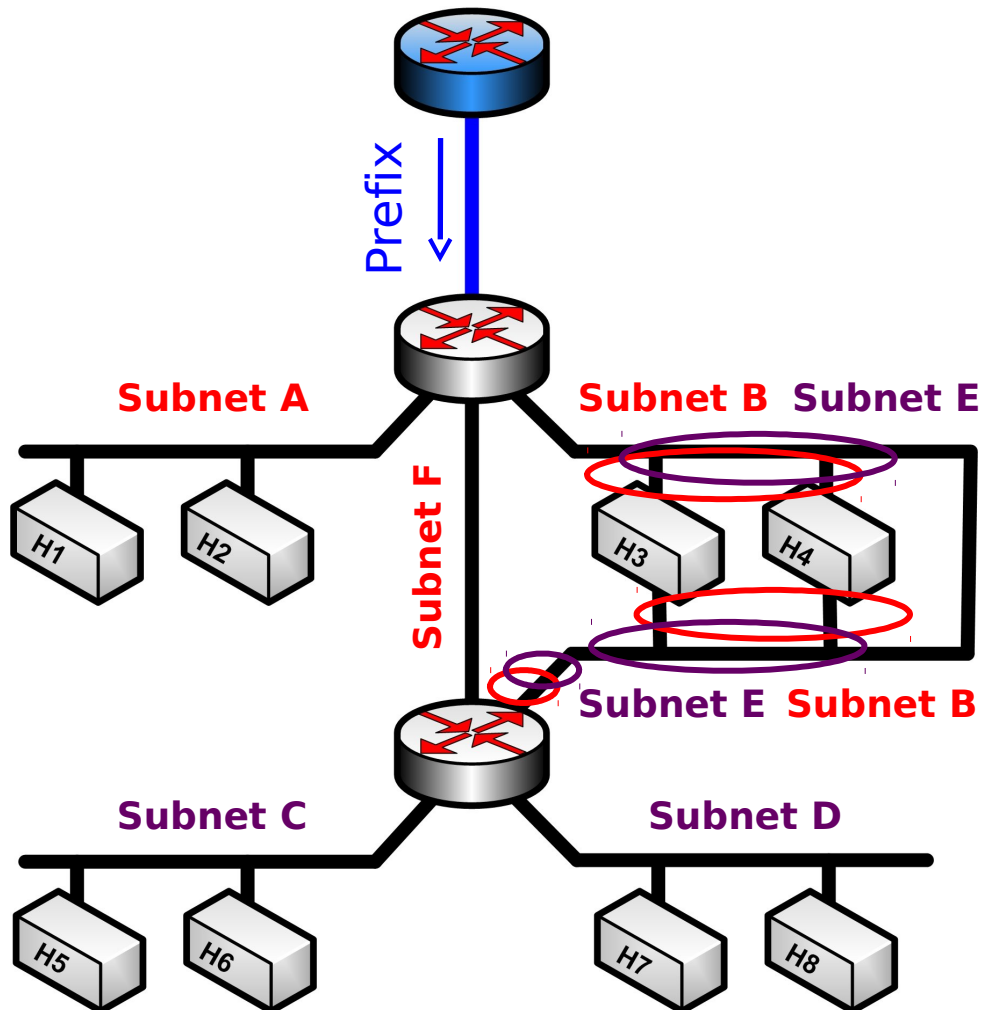
In this example, source selection for globally unique addresses is straightforward, there is only one provider prefix.

For nodes with addresses in subnet B and E, the BRIO cache provides information that subnet B has a lower cost than subnet E.

Assume hosts send packets with source E to Interior Router, source B to Edge Router. How to deal with this is a MIF subject.

6a) Customer Routers run a routing protocol
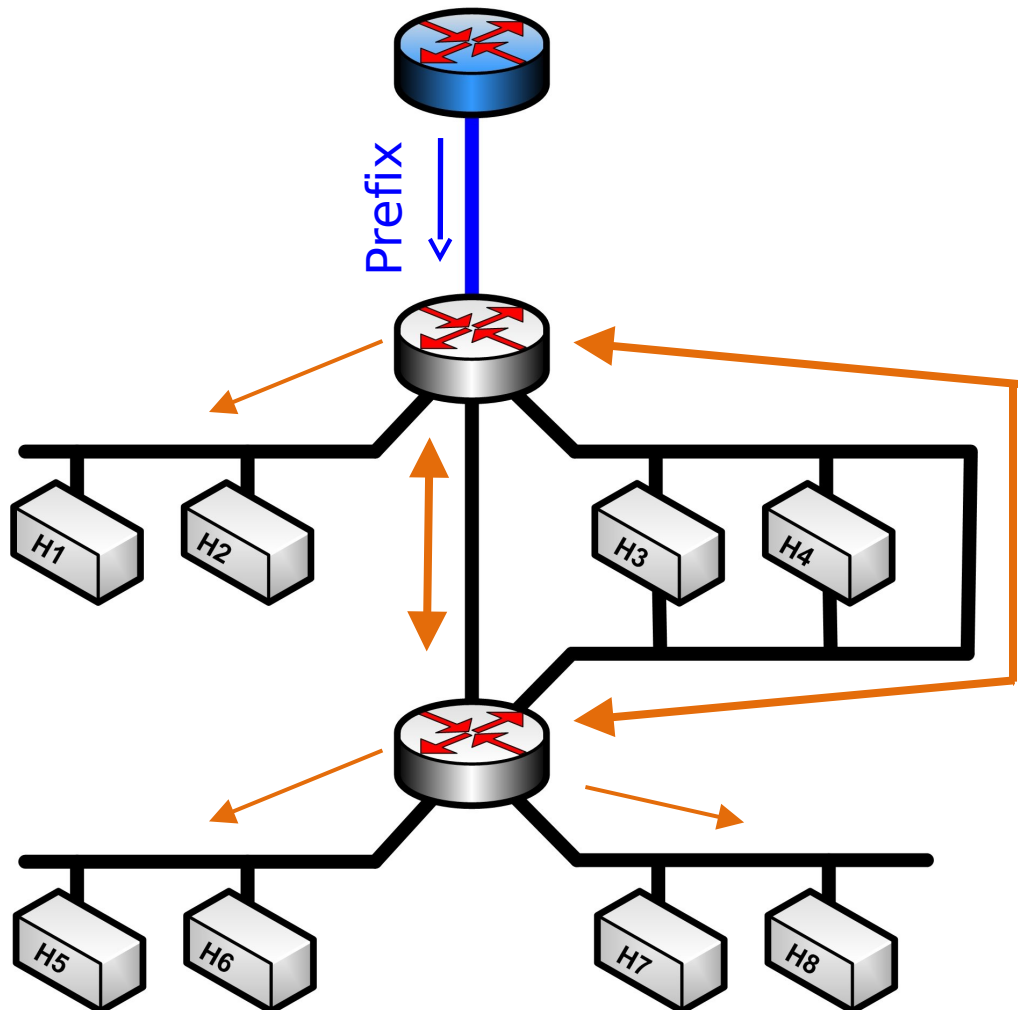
Prefix

The two Customer Routers have adjacency, using two links.

Routing uses link local addresses for adjacency. Reachability to prefixes is exchanged.

BRDP works with any routing protocol.

6b) Forwarding for destinations inside the Homenet
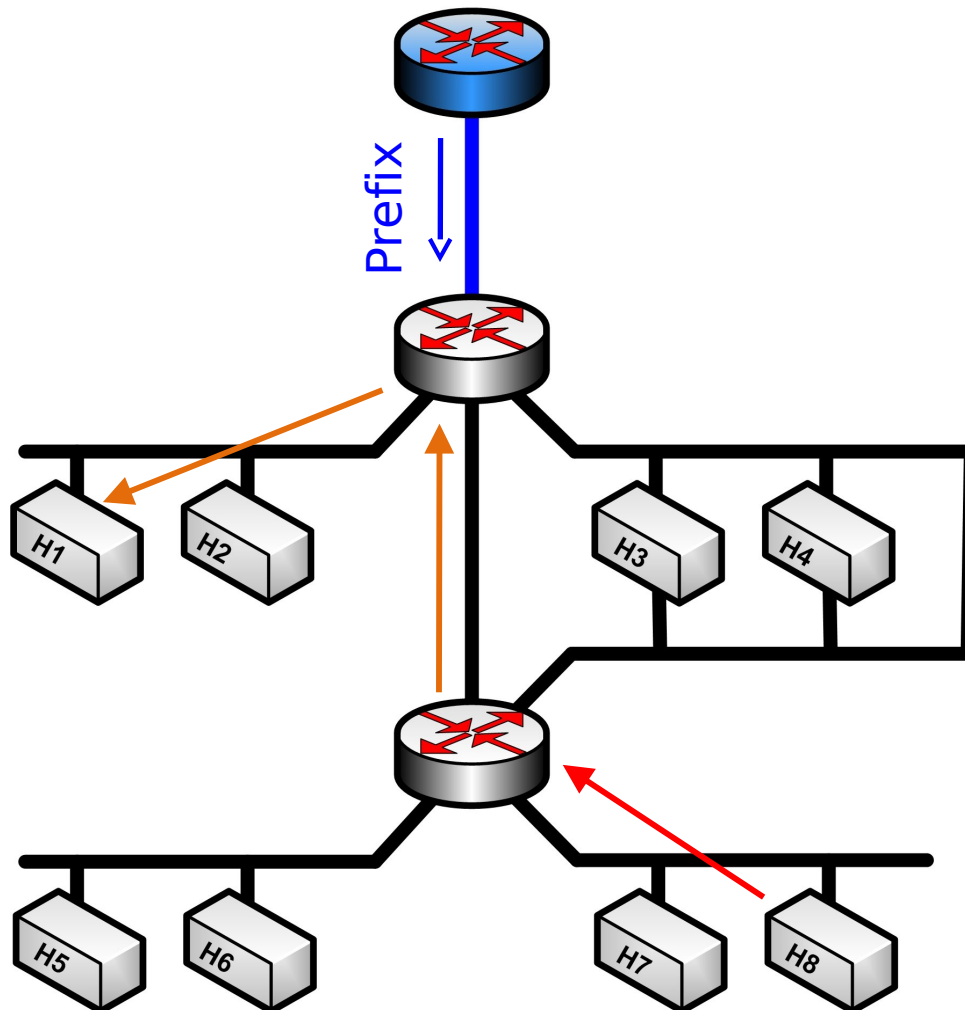
Routing for interior nodes is based on destination address.

Example: H8 send a packet to H1, via Interior Router.
Red: hosts' default router list
Orange: forwarding on DA

6c) Forwarding for destinations outside the Homenet
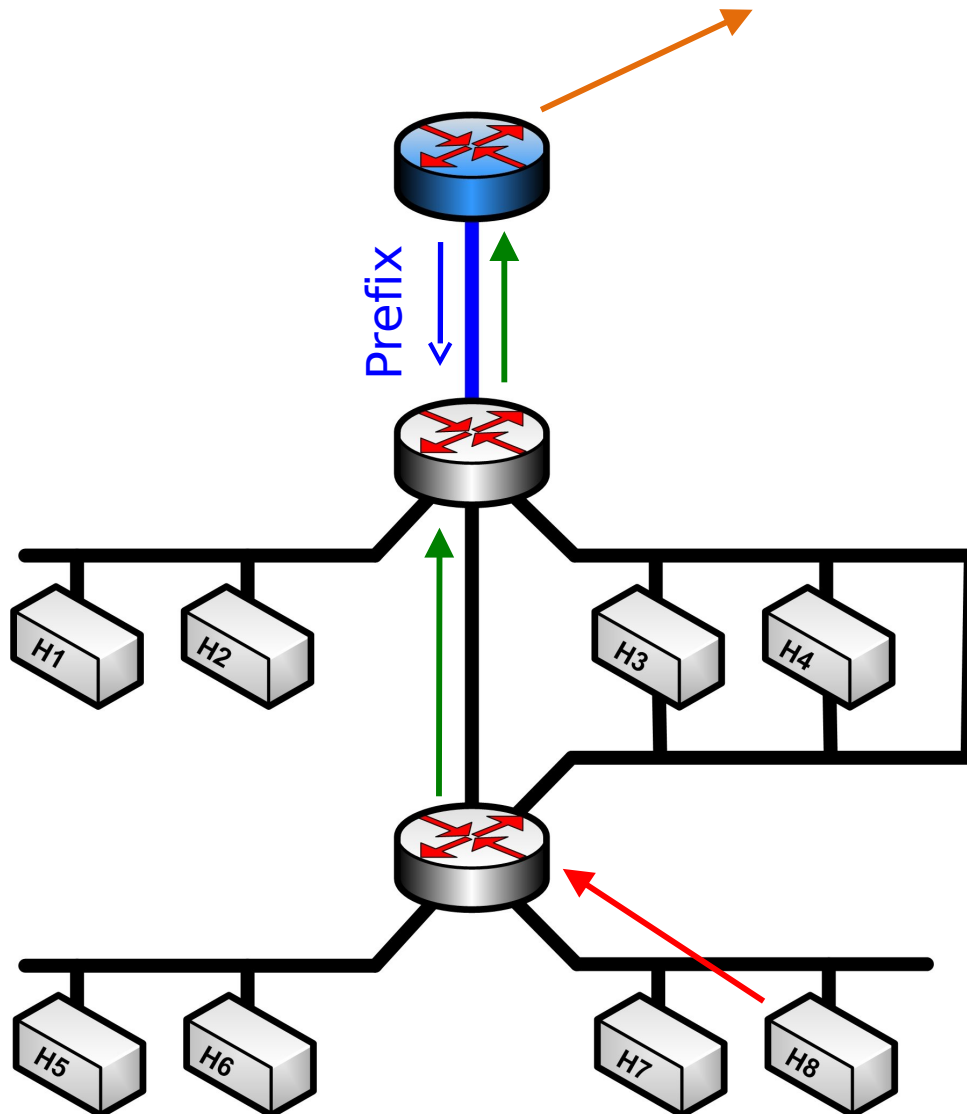
Prefix

For destinations outside the Homenet, a lookup takes place based on source address. This prepares for multi-homing.

Source addresses that matches the prefix are send to the Border Router, the Customer Edge Router.

Example: H8 send a packet to Internet node, via Interior Router.
Red: hosts' default router list
Green: forwarding on SA

# Multi-homing:
3.2.2.2. B: Two ISPs, Two CERs, Shared subnet

Service Provider Router A

Service Provider Router B

Customer Internet Connection A

Customer Internet Connection B

IPv6 Customer Edge Router 1

IPv6 Customer Edge Router 2

Network A

Network B

H1

H2

H3

H4
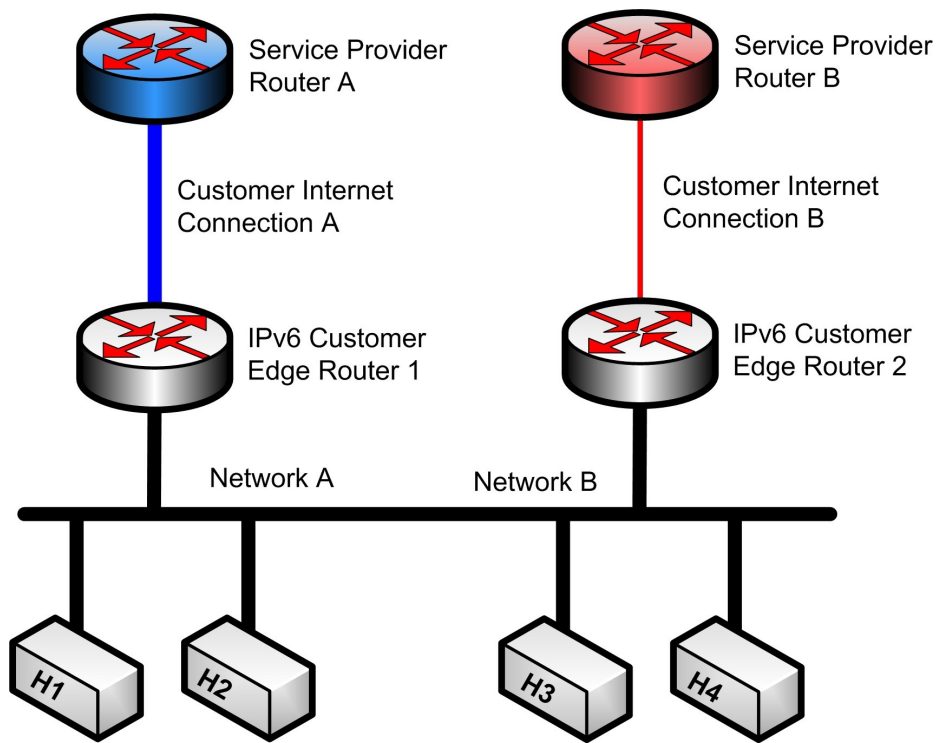
More fun: BRDP multi-homing operation explained using Network topology model B (*Home Networking Architecture for IPv6*)

Again, the Customer Edge Routers has discovered the Customer Internet Connections, and has been assigned prefixes.

Assume providers have ingress filters.

In the example, links to Service providers have different capacity, to point out capabilities of BRDP.
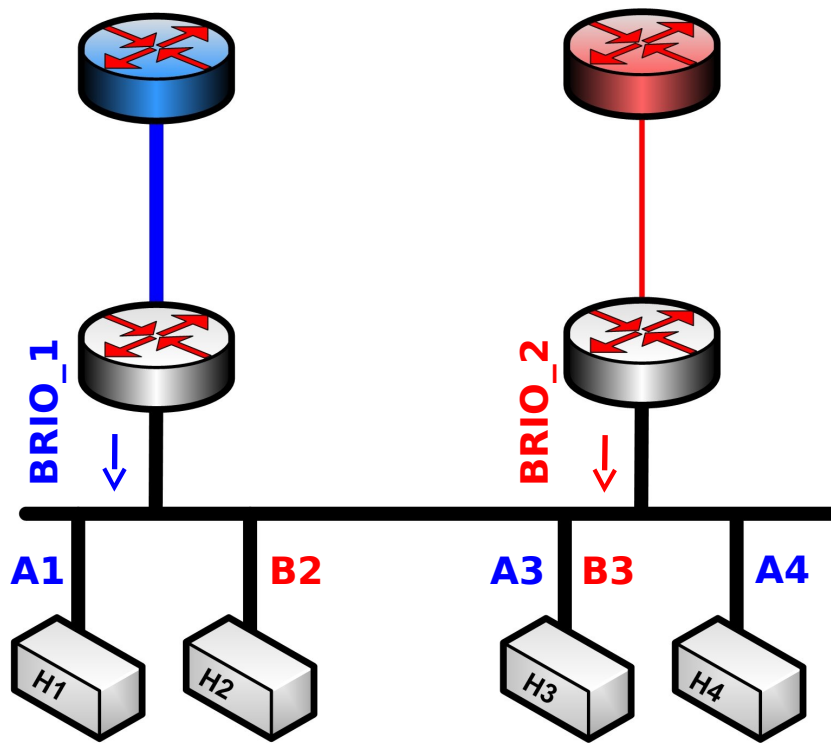
1) BRDP Basics are applied
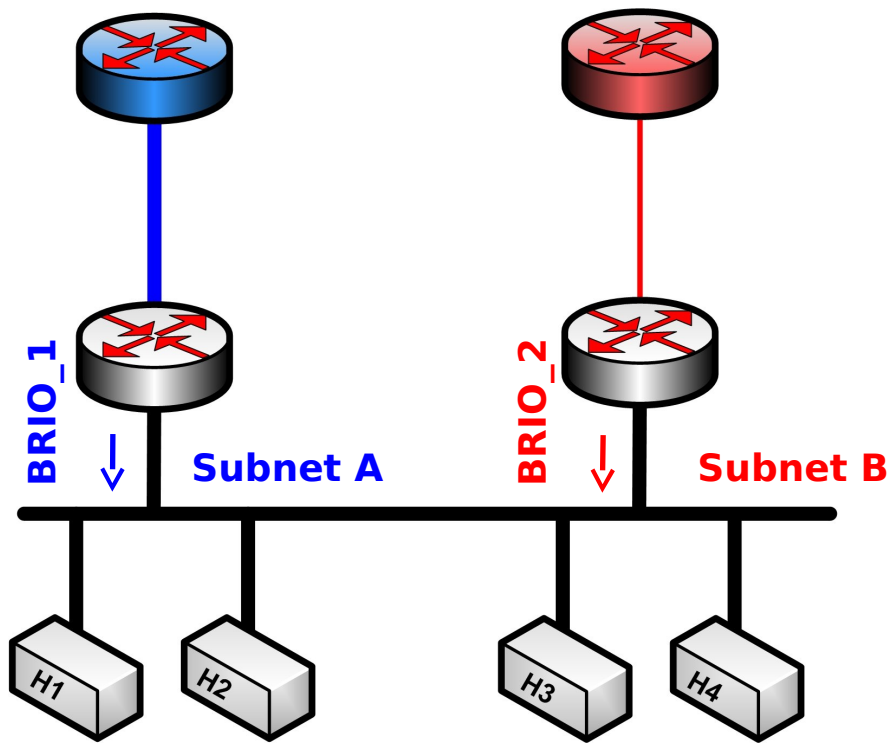
Routers send ND RA with PIO and BRIO.

Router adjacency is formed.

Nodes have configured addresses, optionally for both Service Providers.

The following slides describe BRDPs advanced address configuration, address selection and routing.

2) BRDP based address configuration

Hosts that configure only one address could pick such at random; a fat pipe (blue) or the narrow band alternative path (red). Legacy host H2 has bad luck.

Hosts can benefit from information on path costs to the Service Providers. BRDP provides such, multiple hops away (not shown in this scenario). This is not supported by RFC 4191. Hosts H1 and H4 opt for the high capacity blue fat pipe, with luck or by using BRDP provided information.

Host H3 has multipath support. It will check both paths.
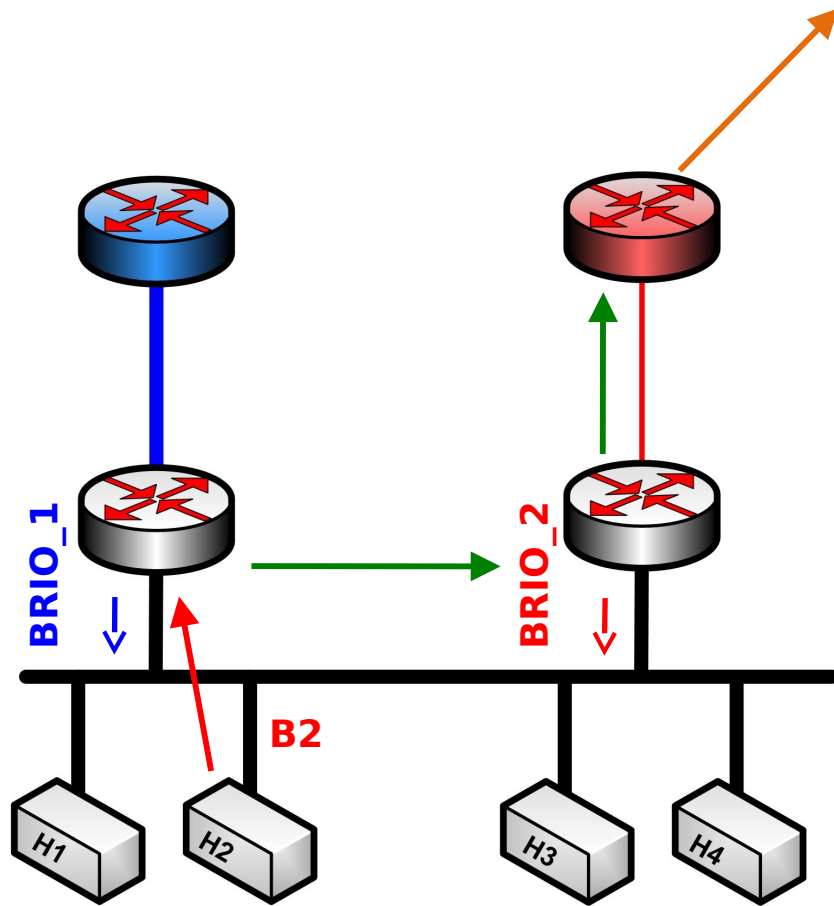
2) BRDP based address selection
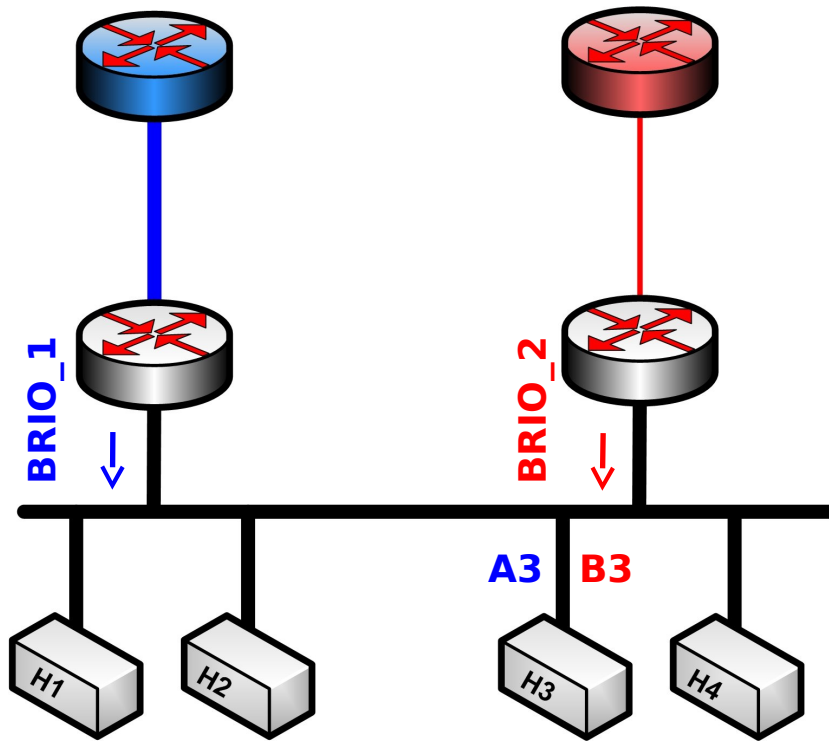
Only host H3 has to select an address out of the candidates, because it has two. Again, it can take randomly. It could use happy eyeballs.

It can do better, using the BRIO provided information. Why not try the route with the best metric first? In this case, this is address A3.

For best performance, it uses both addresses with a multipath transport protocol, e.g. MPTCP.

3a) BRDP based routing, redirecting packets

Host H2 had bad luck with selecting the source address. Murphy is with him, it sends the packet to the less optimal router (red).

With BRDP based routing, Router 1 redirects the packet to Router 2 (green). This because the destination is not within the Homenet and the source belongs to the prefix, advertized by Router 2.

Router 2 forwards the packet to the Service Provider, and the packet is delivered to the final destination.

3b) BRDP based routing, multipath
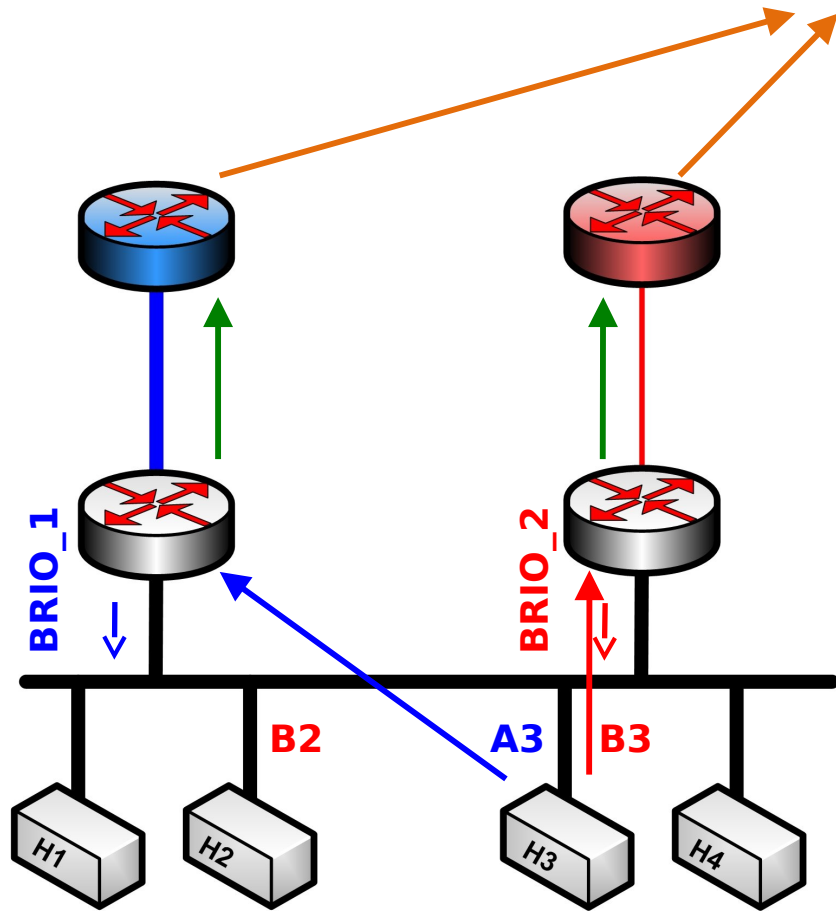
Host H3 has two addresses, one for each Service Provider. It sends the packets the the next hop based on information received with BRIOs.

Forwarding to the correct Customer Edge Router is supported in multi-hop networks also. It does not depend on a routing protocol.

Let's implement BRDP!

Contact

OLSR Development List
<olsr-dev@lists.olsr.org>

Ferry Huberts
<ferry.huberts@pelagic.nl>