

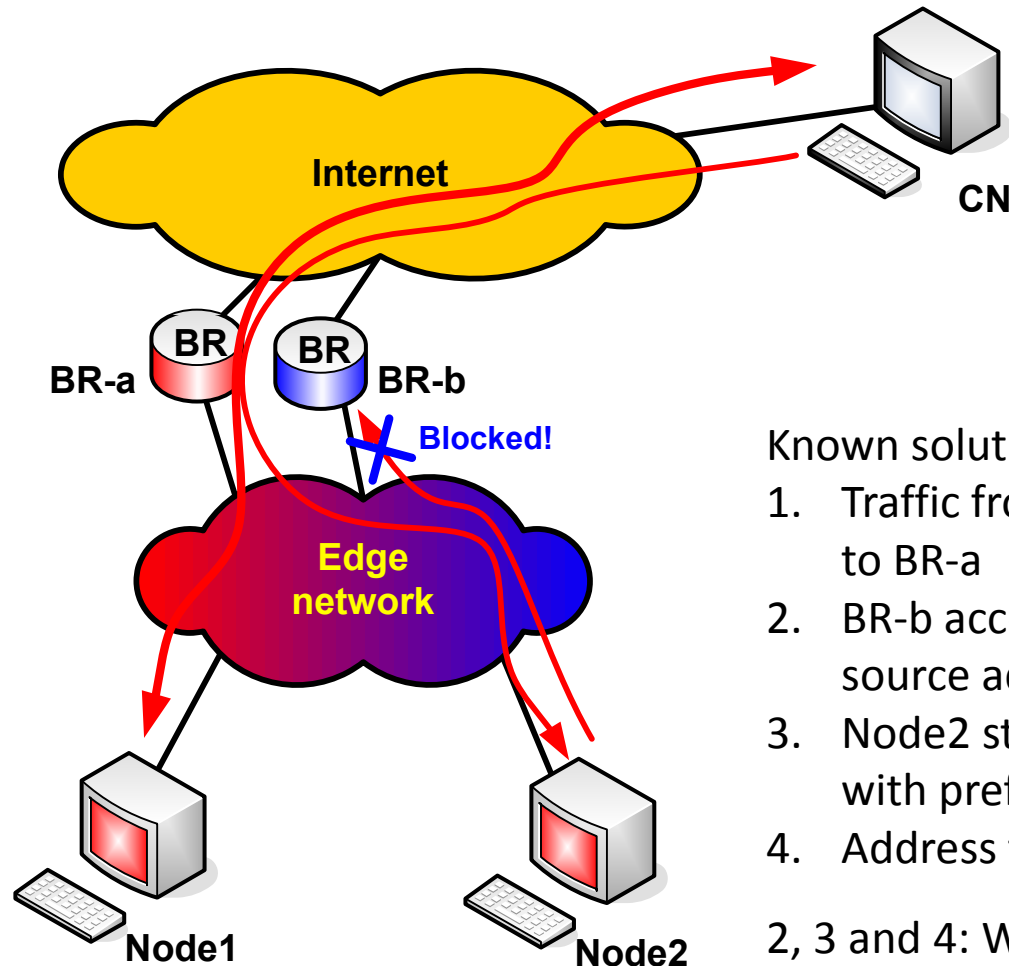
Border Router Discovery Protocol (BRDP) Based Routing

Exit routing for multi-homed networks

`draft-boot-autoconf-brdp-01.txt`
`draft-boot-brdp-based-routing-00.txt`

Teco Boot / 21 November 2008
IRTF Routing Research Group (RRG)

Multi-homed edge network with traffic blocked by ingress filter

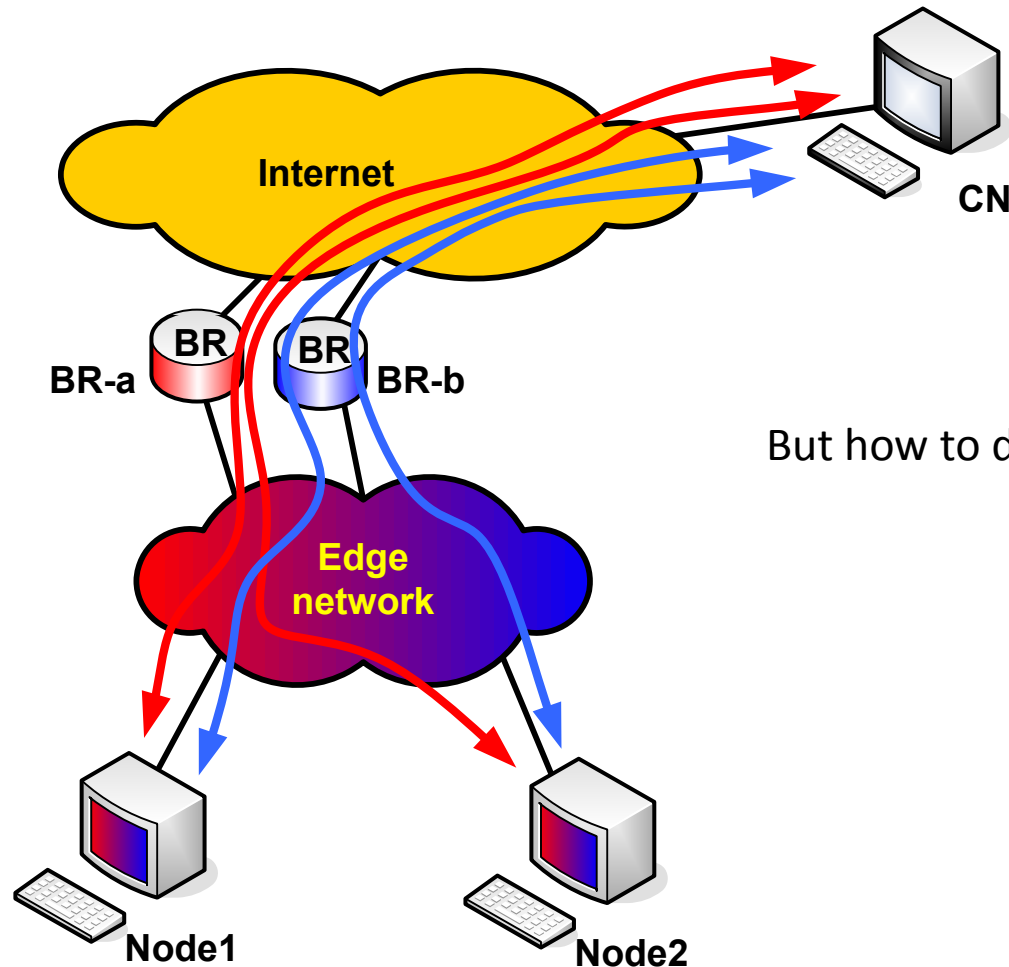


Known solutions:

1. Traffic from Node2 is forwarded to BR-a
2. BR-b accepts BR-a owned source addresses ???
3. Node2 starts new connection with prefix from BR-b ???
4. Address translation ???

2, 3 and 4: What are the issues ?

Goal: support for multi-homed edge networks with multi-homed nodes



But how to do this?

Problems with default gateway routing in a multi-homed network

- Next hop selection is based on destination address
- Multiple DGWs (default gateways):
 - Source has no influence on what DGW is used for sent traffic
 - Three options (at least)
 - Single DGW is used
 - Packet load balancing
 - Flow load balancing
 - Seen from source perspective: used BR (Border Router) is *guesswork*
- BRs may have ingress filters (is a SHOULD)
- Result:
 - In a multi-homed network, outgoing traffic could be blocked on BR when Source Address does not correspond with delegated prefix of used BR**

See also:

RFC5220 section 2.1.2. “Ingress Filtering Problem”

RFC3704 section 4.3. “Send Traffic Using a Provider Prefix Only to That Provider”

draft-ietf-shim6-ingress-filtering section 3. “The problem: Ingress filtering incompatibility”

Problems with default gateway routing in a multi-homed network

- Next hop selection is based on destination address
- Multiple DGWs (default gateways):
 - Source has no influence on what DGW is used for sent traffic
 - Three options (at least)
 - Single DGW is used
 - Packet load balancing
 - Flow load balancing
 - Seen from source perspective: used BR (Border Router) is *guesswork*
- BRs may have ingress filters (is a SHOULD)
- Result:

In a multi-homed network, outgoing traffic could be blocked on BR when Source Address does not correspond with delegated prefix of used BR

See also:

RFC5220 section 2.1.2. "Ingress Filtering Problem"

RFC3704 section 4.3. "Send Traffic Using a Provider Prefix Only to That Provider"

draft-ietf-shim6-ingress-filtering section 3. "The problem: Ingress filtering incompatibility"

SO, LETS FIX THIS!

Current behavior:

RFC1812; Section 5.2.1.2 Unicast:

- (5) The forwarder determines the next hop IP address for the packet, usually by looking up the packet's destination in the router's routing table. This procedure is described in more detail in Section [5.2.4]. This procedure also decides which network interface should be used to send the packet.

RFC1812; Section 5.2.4.3 Next Hop Address:

- (5) Default Route: This is a route to all networks for which there are no explicit routes. It is by definition the route whose prefix length is zero

BRDP Based Routing

- With BRDP Based Routing, the default route is removed.
- New heuristic for finding a next hop, only used when first FIB lookup didn't find a nexthop (draft-boot-brdp-based-routing) :

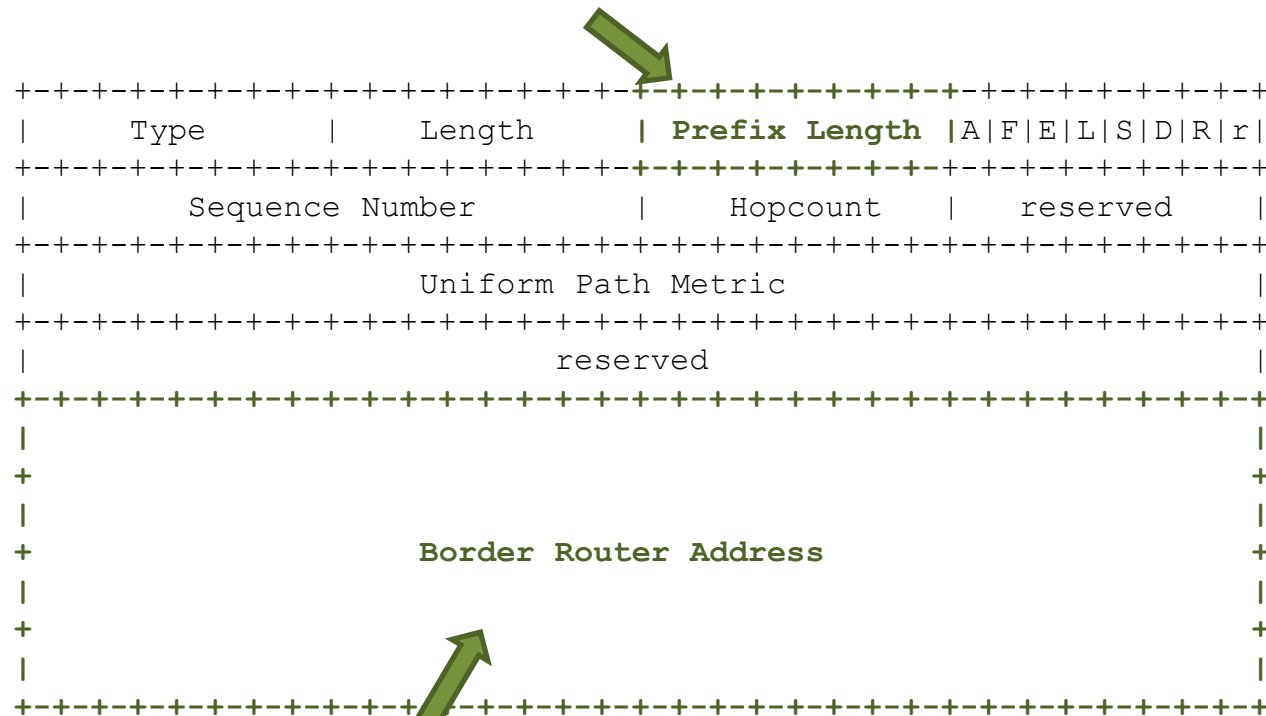
(6) BRDP Route: This is a route to all networks for which there are no explicit routes, and a default route is not used.

The nexthop IP address is found by means of a Border Router Information Cache (BRIO-Cache) lookup based on the source address and, if a matching BRIO-Cache entry is found, a subsequent FIB lookup based on the selected Border Router address.

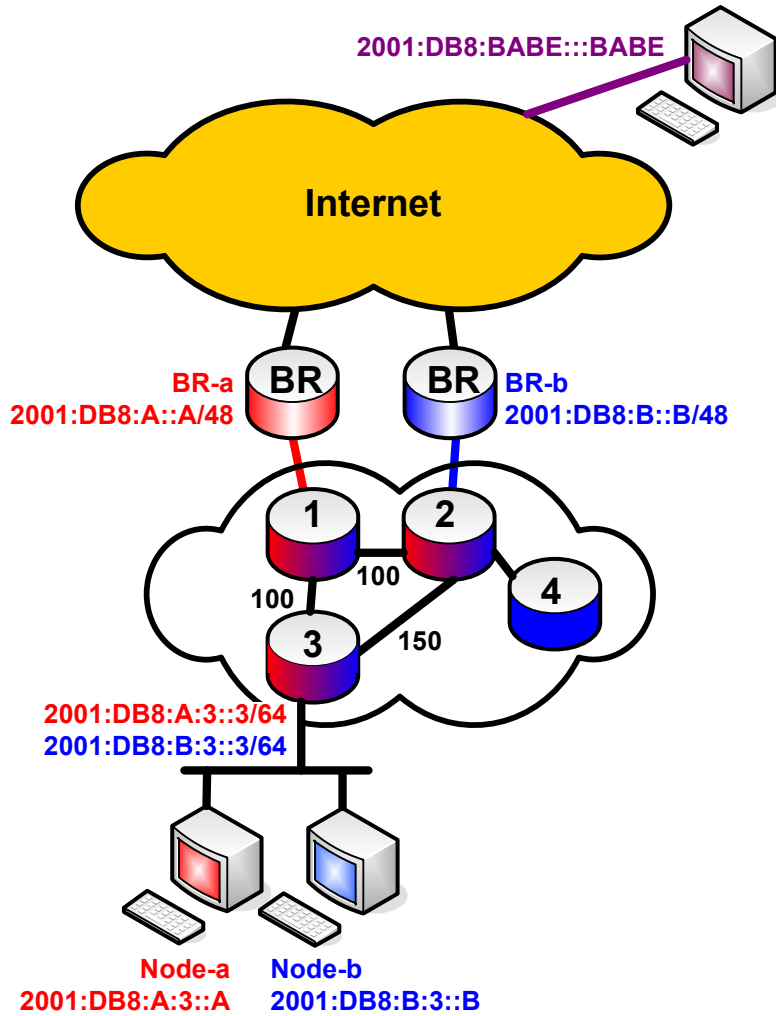
BRIO-Cache:

- BRIO-Cache:
 - Table, populated with received or initiated Border Router Information Options (BRIOs)
- Information is used for generating globally unique addresses in a MANET (BRDP is designed for Autoconf)
- A BRIO provides ingress filter information for the Border Router
- BRIOs are flooded hop-by-hop using ND RA
 - Adjustments on metric, hopcount

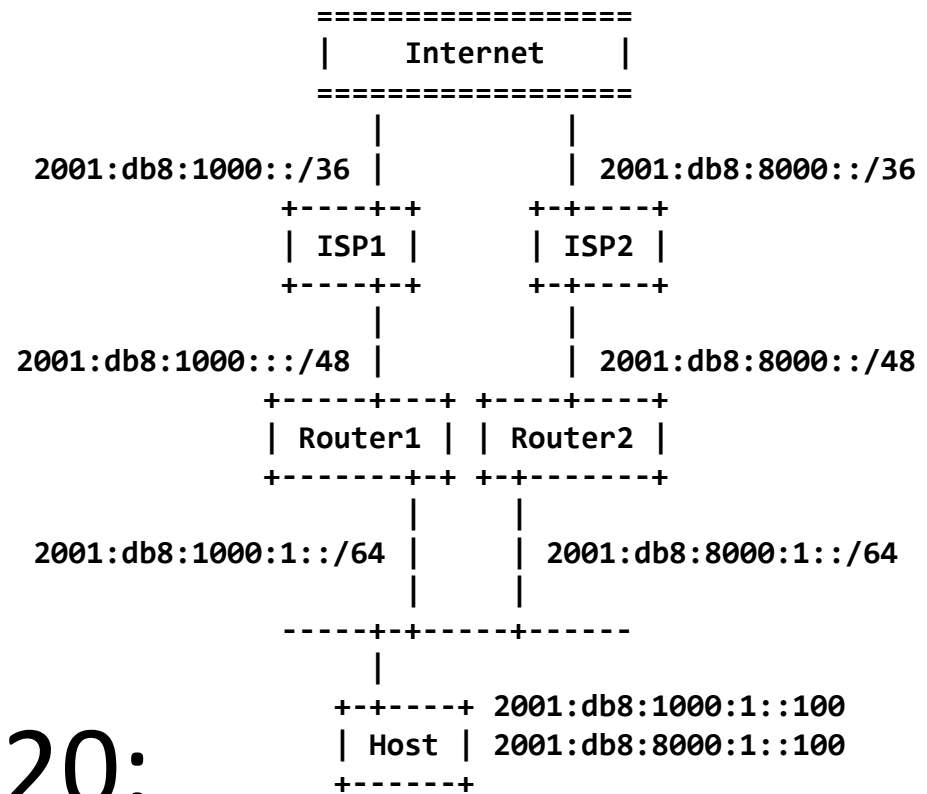
BRIO format



BR address and valid prefix for this BR



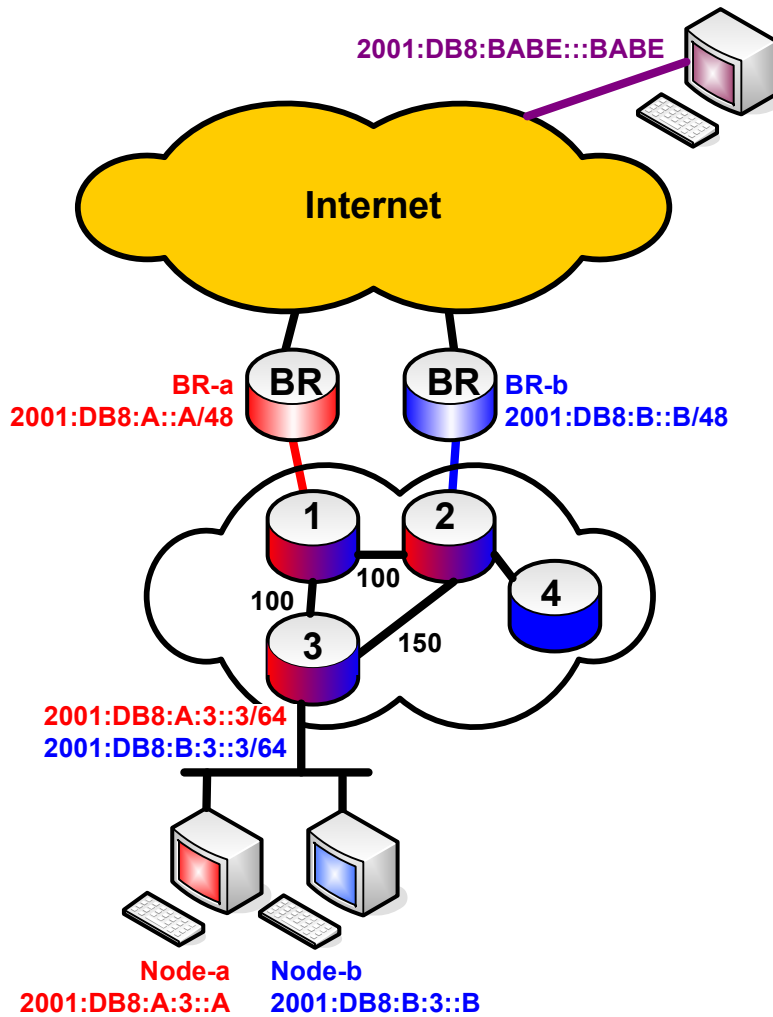
← My example



RFC 5220:

Figure 1

BRDP based routing IETF-73



Example

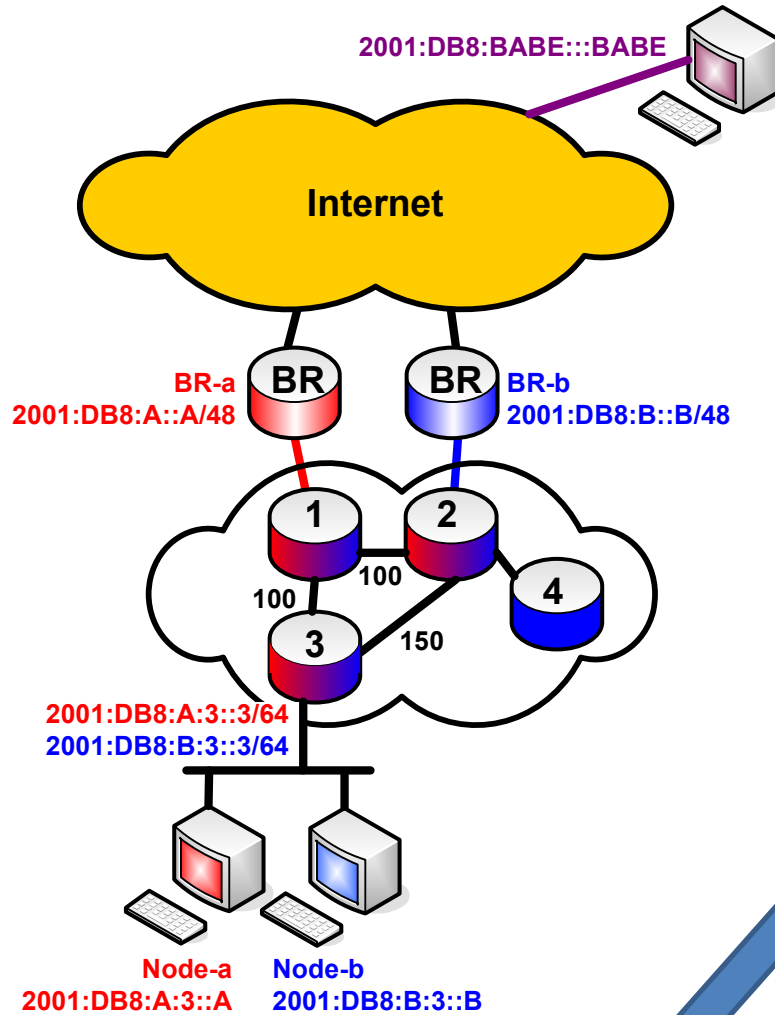
FIB R3:

2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

Example - FIB



All internal routes are found in FIB
 Forwarding based on DA
 No default route !!

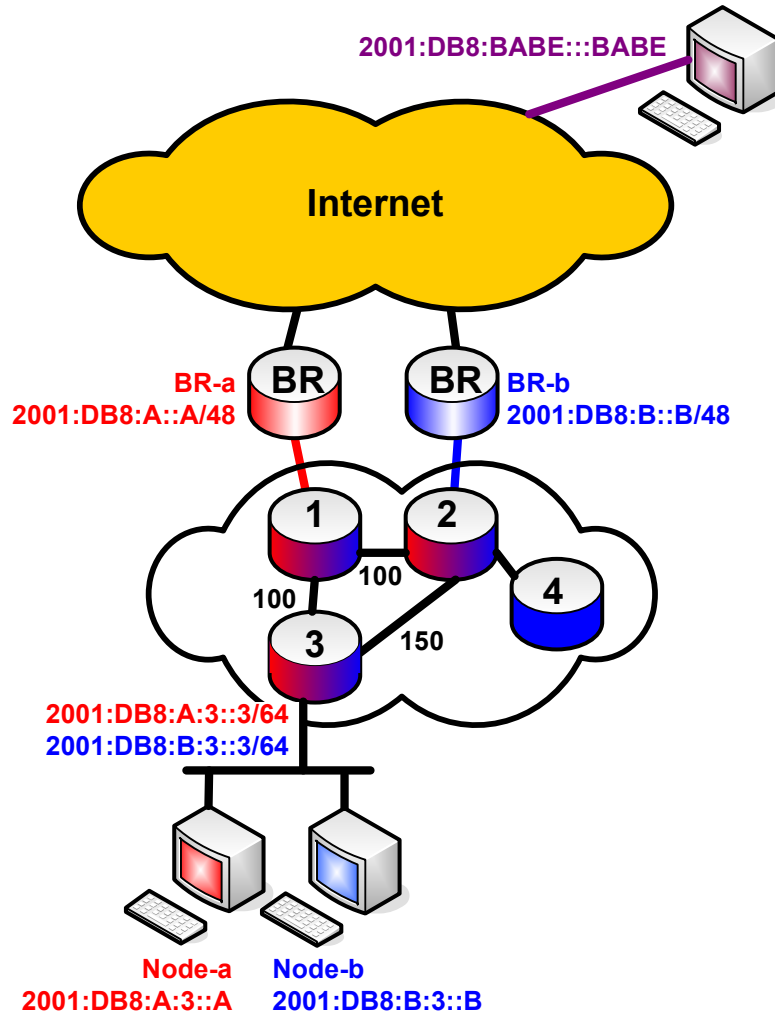
FIB R3:

2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

Example BRIO-Cache



Border Router information,
with sender info and metrics

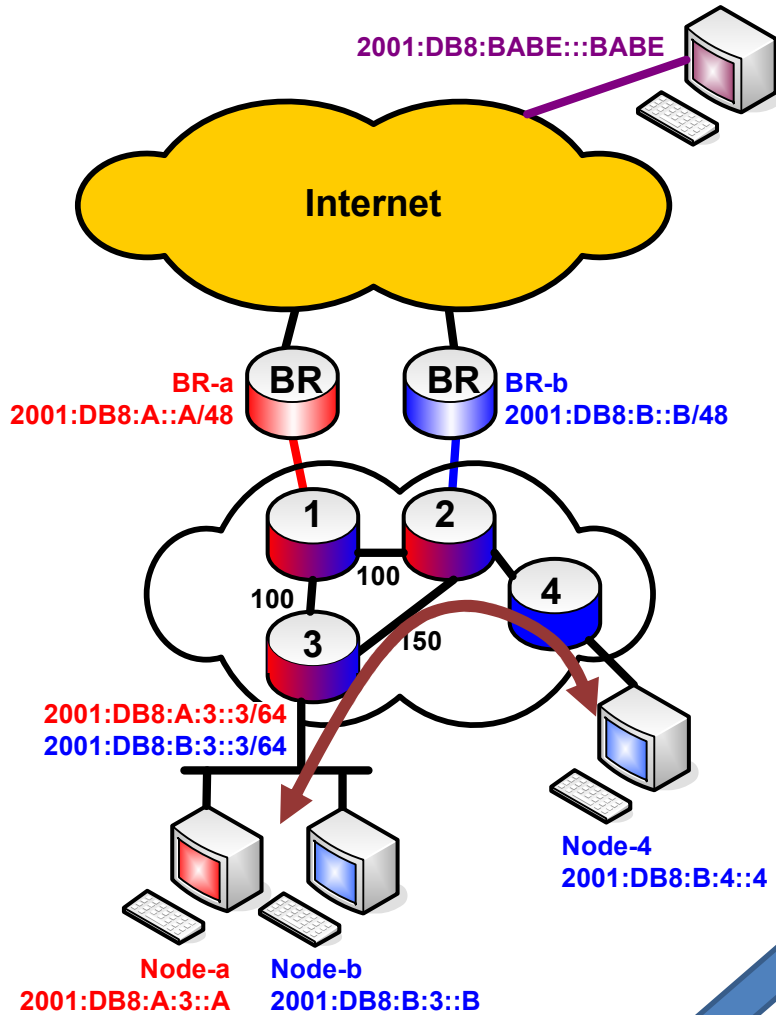


FIB R3:

2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200



Node-a or Node-b sends to Node-4

FIB R3:

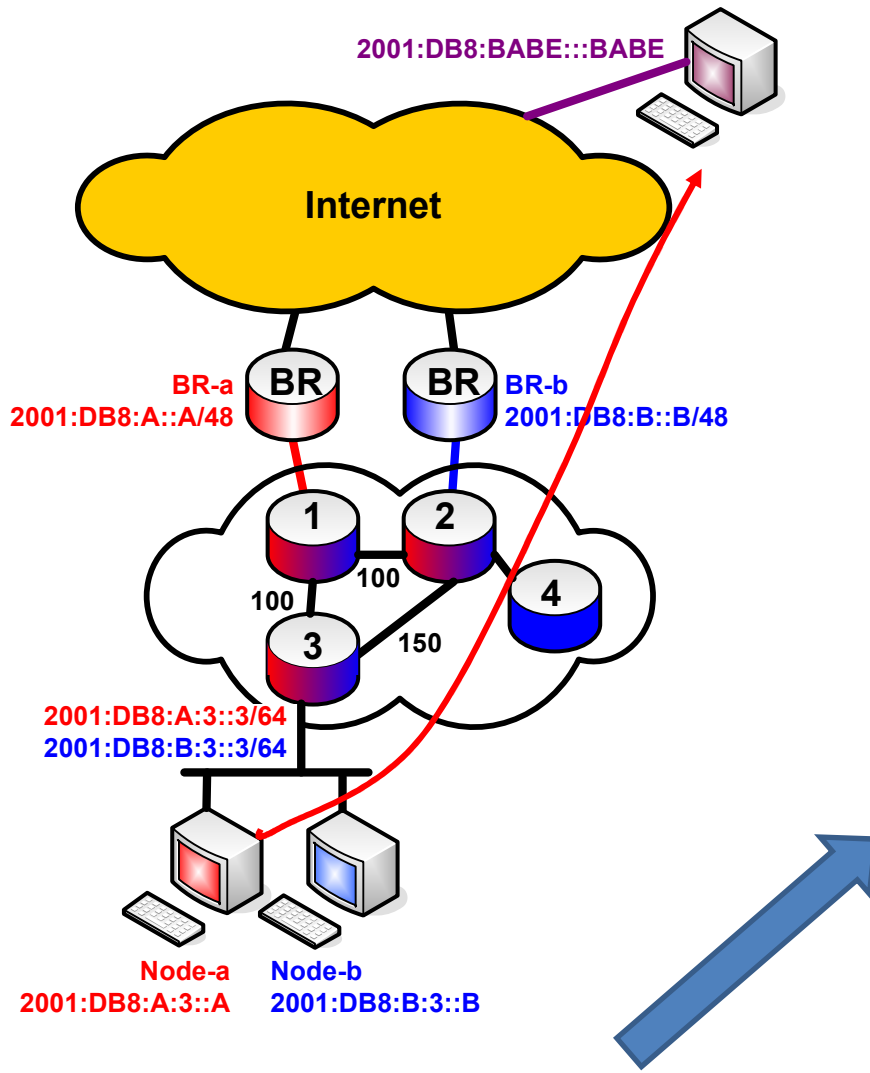
2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

1) 2001:DB8:B:4::4 in FIB?

Yes



Node-a sends to Babe

FIB R3:

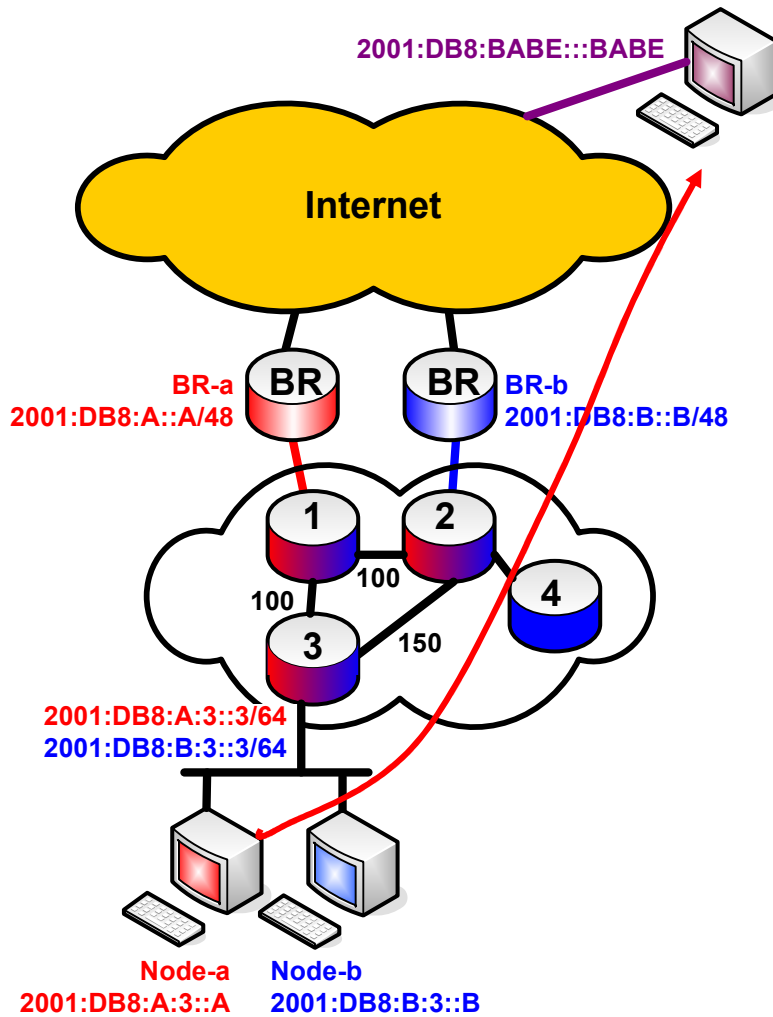
2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

1) 2001:DB8:BABE::BABE in FIB?

NO



Node-a sends to Babe

FIB R3:

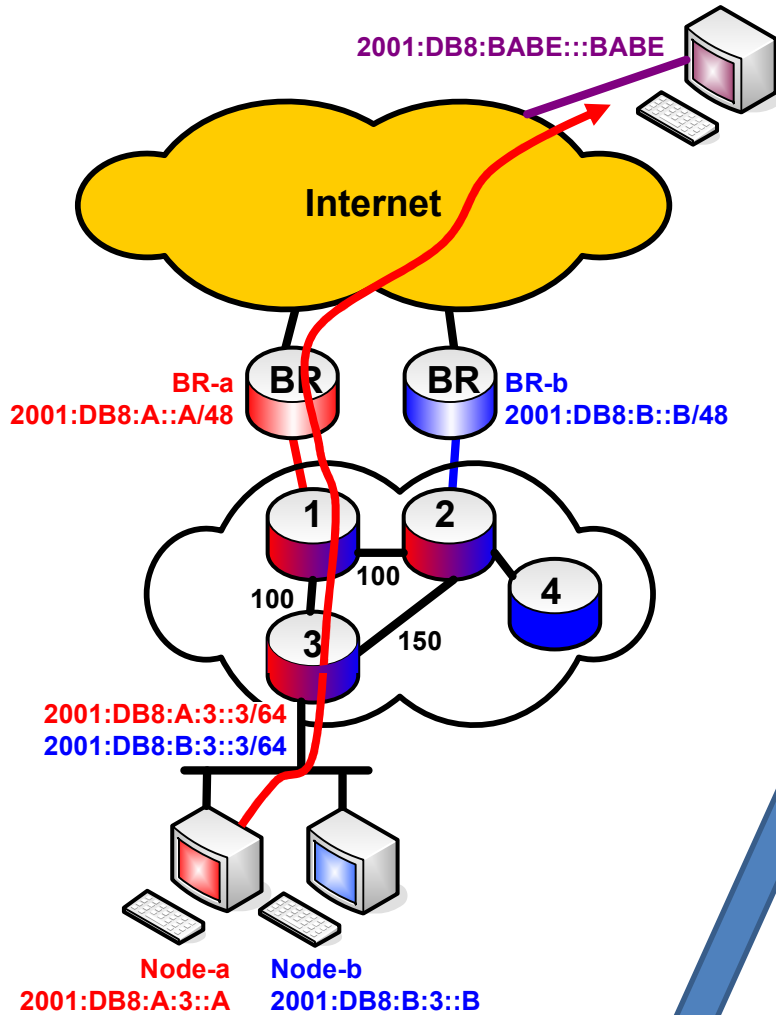
2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

2) 2001:DB8:A:3::A match in BRIO Cache?

YES: 2001:DB8:A::A



Node-a sends to Babe

FIB R3:

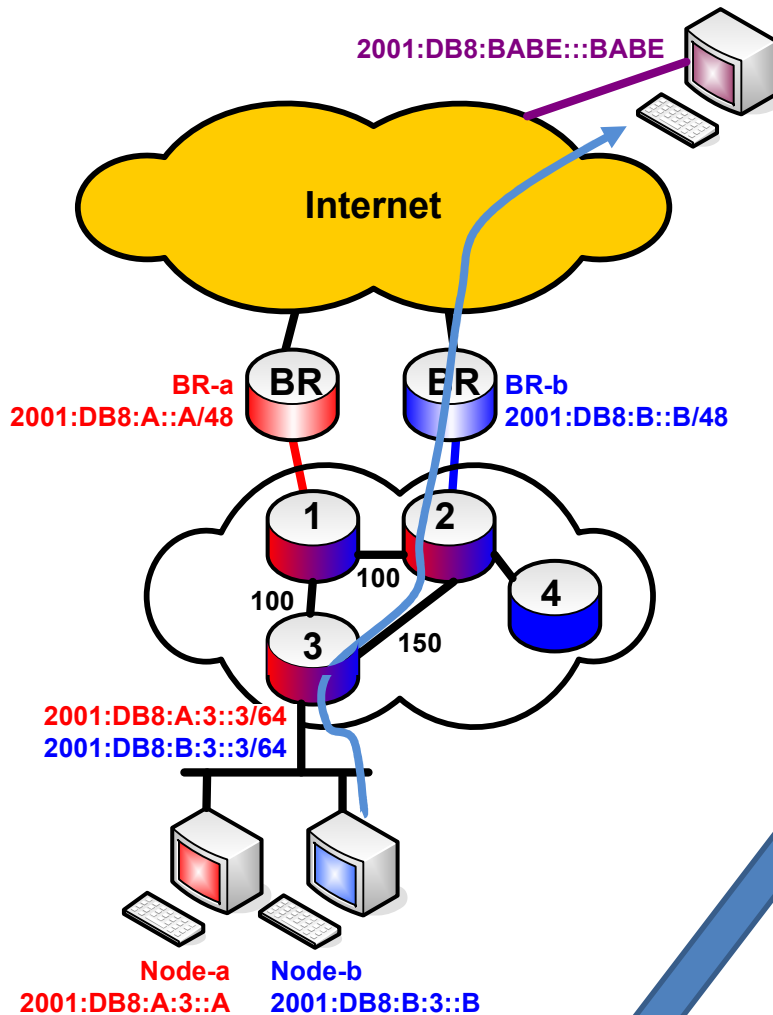
2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

3) 2001:DB8:A::A in FIB?

YES: 2001:DB8:A::/64 -> FE80::1



Node-b sends to Babe

FIB R3:

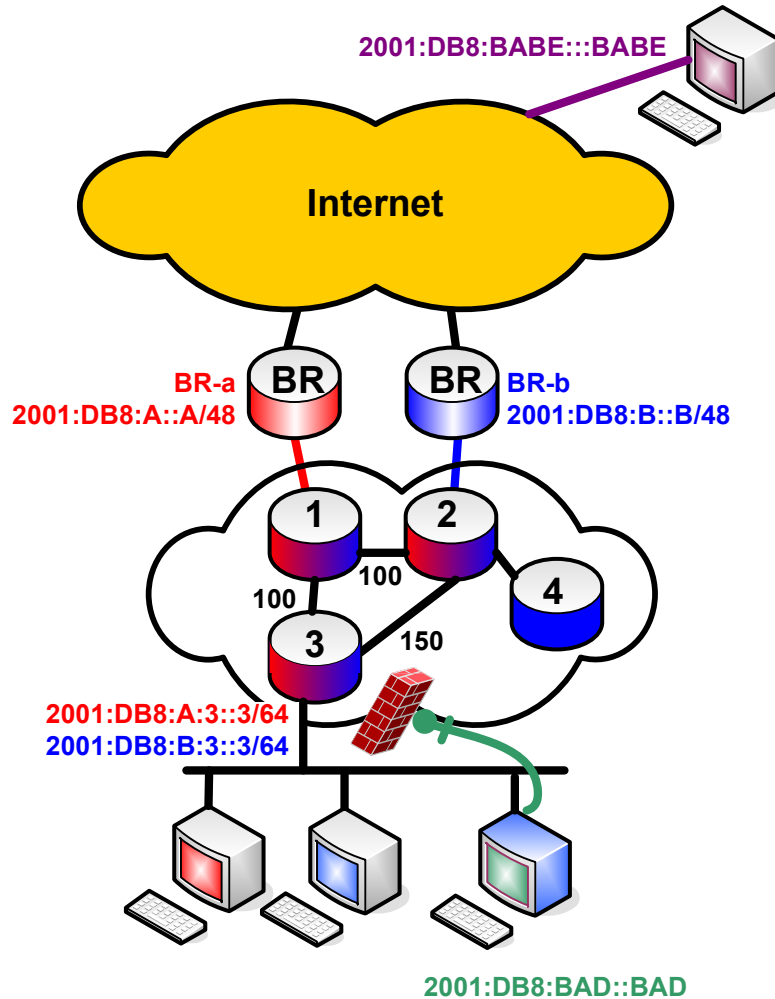
2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

- 1: 2001:DB8:BABE::BABE in FIB: **NO**
- 2: 2001:DB8:B:3::B match in BRIO Cache :
Yes, 2001:DB8:B::B
- 3: 2001:DB8:B::B in FIB: Yes, FE80::2

Bad sends to Babe



1: 2001:DB8:BABE::BABE in FIB: **NO**

2: 2001:DB8:BAD::BAD match in BRIO Cache : **NO**

3: **DROP**

FIB R3:

2001:DB8:A::/64	-> FE80::1	# BR-a
2001:DB8:A:1::/64	-> FE80::1	
2001:DB8:A:2::/64	-> FE80::2	
2001:DB8:A:3::/64	-> local	
2001:DB8:B::/64	-> FE80::2	# BR-b
2001:DB8:B:1::/64	-> FE80::1	
2001:DB8:B:2::/64	-> FE80::2	
2001:DB8:B:3::/64	-> local	
2001:DB8:B:4::/64	-> FE80::2	

BRIO Cache R3:

2001:DB8:A::A/48	<= FE80::1 metric 100
2001:DB8:A::A/48	<= FE80::2 metric 250
2001:DB8:B::B/48	<= FE80::2 metric 150
2001:DB8:B::B/48	<= FE80::1 metric 200

Next steps

- Rework on Internet Drafts
 - Support for IPv4, SA selection, next-hop selection for hosts
 - Get running code
-
- Website: <http://www.inf-net.nl/brdp.html>
 - Mail: teco@inf-net.nl
 - Any help is welcome !
 - Any comment is welcome !

Thanks for your attention !