

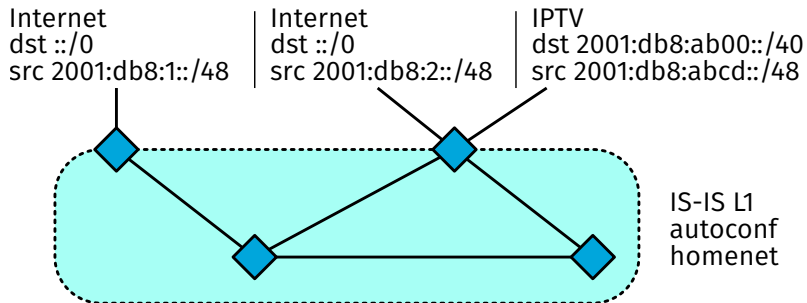
IS-IS Dst/Src Routing & Critical SubTLVs

draft-lamarter-isis-reachability-critical-subtlvs-00
draft-baker-ipv6-isis-dst-src-routing-02

Fred Baker · fred@cisco.com
David Lamarter · david@netdef.org

IETF 91, Honolulu, Nov. 2014

Context



- ▶ ensuring correct exit taken based on packet source address
- ▶ BCP 38 ingress filtering at ISP not to be touched
- ▶ MPTCP needs multiple paths visible at the host
- ▶ application beyond homenet expected (e.g. SMB)

NOT-Context

This is **not** MT / Policy routing with a topology per source prefix.

The expectation is $> 95\%$ dst routes representing prefixes inside the domain, plus $< 5\%$ dst-src routes representing specific external services.

All routes go into **one FIB** which behaves as described in draft-lamparter-rtgwg-dst-src-routing

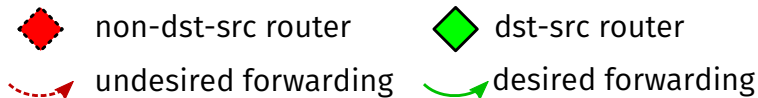
Dst/Src routing is easy

```
+-----+
| TLV 236/237, dst ::/0, etc.
| +-----+
| | TLV X, src 2001:db8:1::/48
| +-----+
+-----+
```

- ▶ this is draft-baker-ipv6-isis-dst-src-routing-02
- ▶ sticking the source prefix in is straightforward
 - ▶ (previous versions forgot to mention MT, that's fixed)
- ▶ implemented, tested, demo'd @ IETF 90
- ▶ open source: <https://git-us.netdef.org/projects/OSR/repos/openwrt-isis-hnet>
- ▶ no changes at all to IS-IS "core"

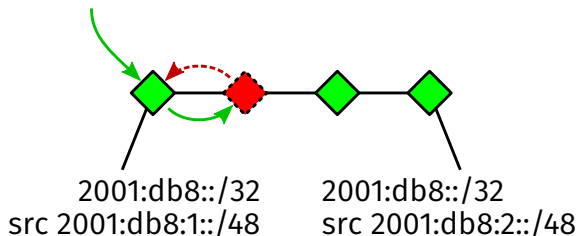
Interop is hard

- ▶ naively mixing dst-src routers and non-dst-src routers leads to persistent loops
- ▶ might be out of scope for homenet
- ▶ certainly in scope for SMB & Campus applications



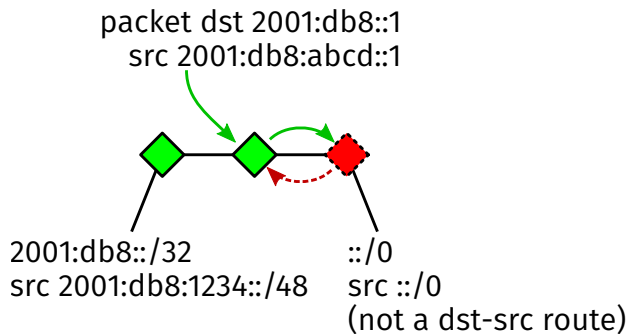
Loop case #1

packet dst 2001:db8::1
src 2001:db8:2::1



- ▶ a dst-src route **MUST NOT** traverse via a non-dst-src nexthop

Loop case #2



- ▶ dst-src routes MUST be ignored by non-dst-src routers

What do we need

To fix these scenarios, we need:

- ▶ to fix loop #1: capability indication
- ▶ to fix loop #2: hiding reachabilities from “old” systems

The goal is correctness, which can imply blackholing traffic (e.g. if there are two islands of dst-src.) Multiple approaches are possible:

- ▶ “try hard”:
 - calculate a separate topology of only dst-src routers
- ▶ “fail quick”:
 - blackhole when dst-only routers on shortest path
- ▶ “fail completely”:
 - don't become adjacent with dst-only routers

Solution attempts

Critical Sub-TLV: “greenfield” / “nice” solution

- ▶ creates new IP Reachability TLVs
- ▶ has 2 kinds of Sub-TLVs:
 - ▶ plain optional Sub-TLVs (share a codepoint namespace)
 - ▶ mandatory/critical Sub-TLVs (entire reachability MUST be ignored if Sub-TLV is not understood)
- ▶ implies separate topology/MT based on capabilities
(in this case: separate topology containing only dst-src routers)

Draft includes a simple capability advertisement mechanism
(too simple?)

Solution attempts

```
+-----+-----+
| TLV = 239 | length=... |
+-----+-----+
| MTID = #0 |
+-----+-----+
| metric |
+-----+-----+
| flags |
+-----+-----+
| dst = ::/0 |
+-----+-----+
| critical:
| +-----+-----+
| | TLV X, src 2001:db8:1::/48
| +-----+-----+
| optional:
| +-----+-----+
| | TLV 1, admin tag 0xdeadbeef
| +-----+-----+
+-----+-----+
```

Alternative solutions

(a) register MTID for dst-src usage

- ▶ probably include capability mechanism to prevent dst-src/non-dst-src mixture in this particular topology (to prevent operator self-foot-shooting)
- ▶ results in 2 MTIDs installed in 1 FIB

-or-

(b) create dst+src main TLV type

- ▶ crude way to hide dst-src routes from non-dst-src routers, still needs capability indication

Further path?

- ▶ need to clear topic and proceed
- ▶ code is running (w/o interop)