Multihoming in an IPv6 world

Tales from a networking course

Olivier Tilmans
olivier.tilmans@uclouvain.be
https://github.com/UCL-INGI/lingi2142

Hands-on experience is a must to understand how networks are operated

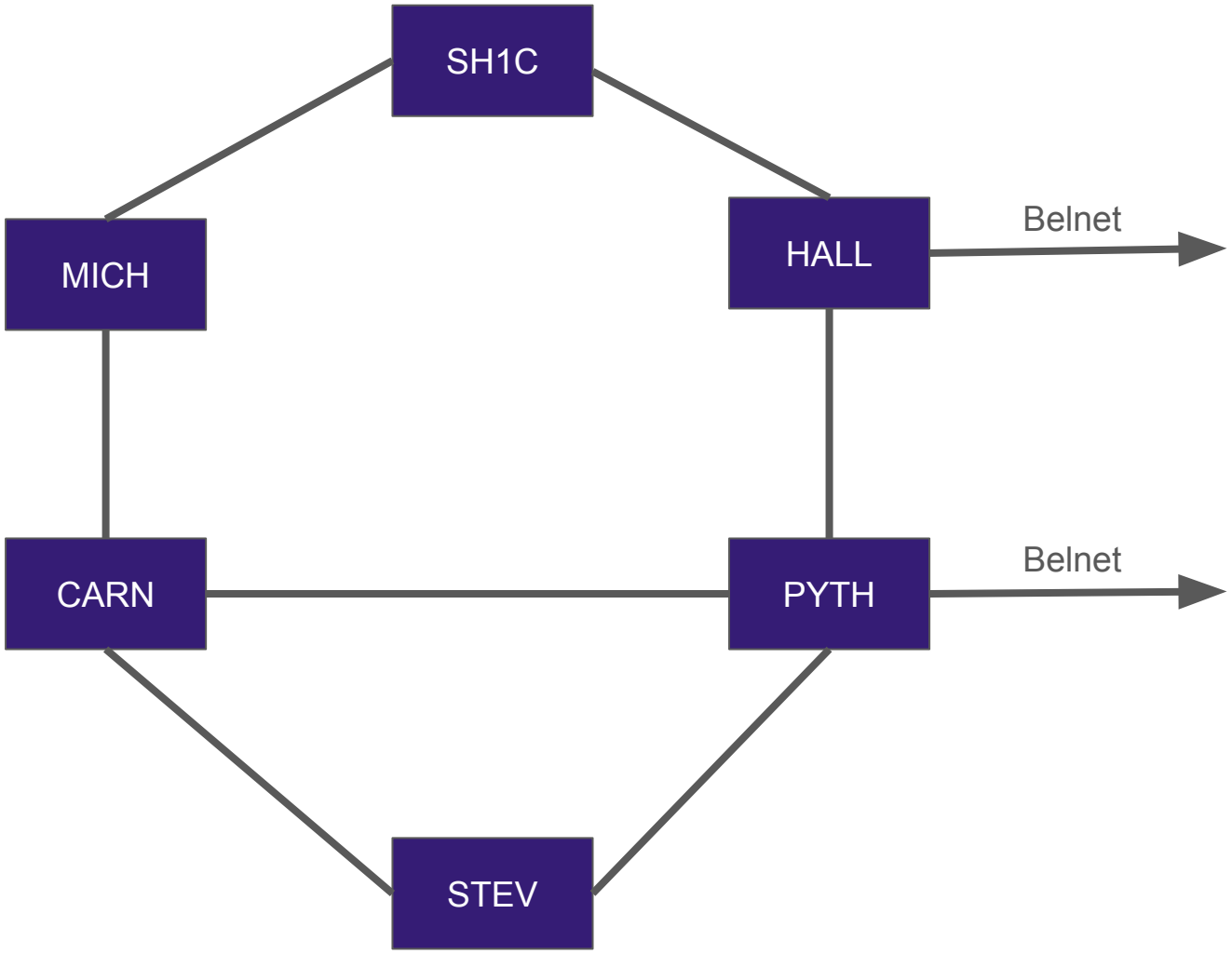
We asked student to configure and operate a campus network

- IPv6 only
- Open-source tools
- Multihoming using IPv6 PA prefixes

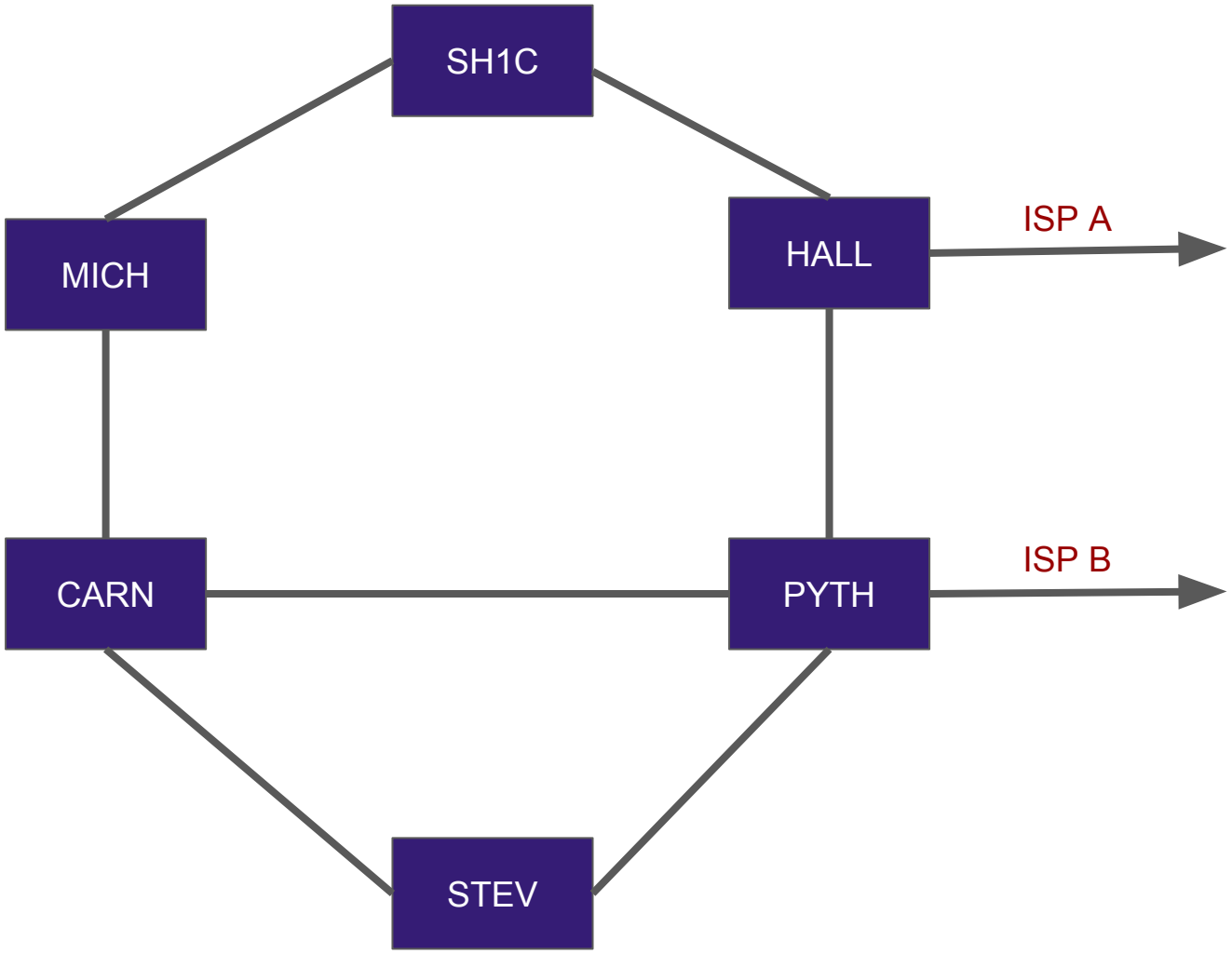
We asked student to configure and operate a campus network

- Addressing plan
- Routing
- Security
- Services
- ...

L3 topology of UCL



L3 topology of UCL with a twist



Multihoming without NAT is challenging

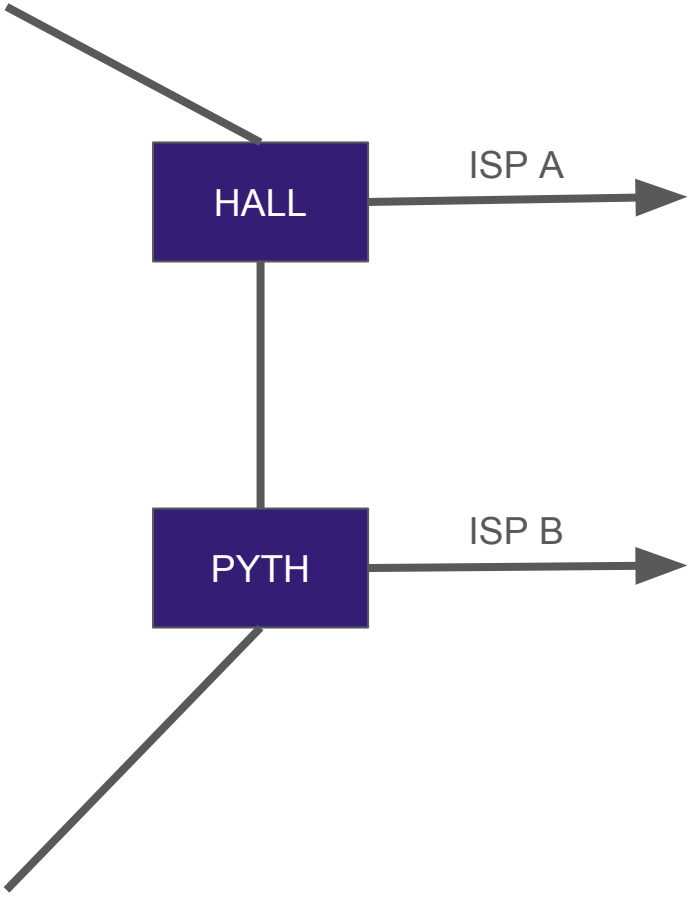
- BCP 38, BCP 38, BCP 38
- Internal routing policies
- Withstanding failures

Multihoming without NAT is challenging

- BCP 38, BCP 38, BCP 38
- Internal routing policies
- Withstanding failures

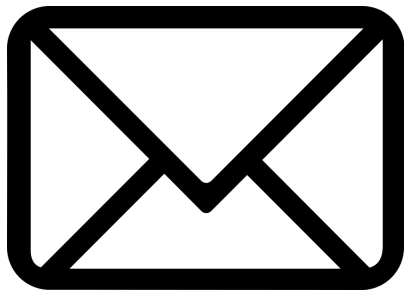
Considering the following setup

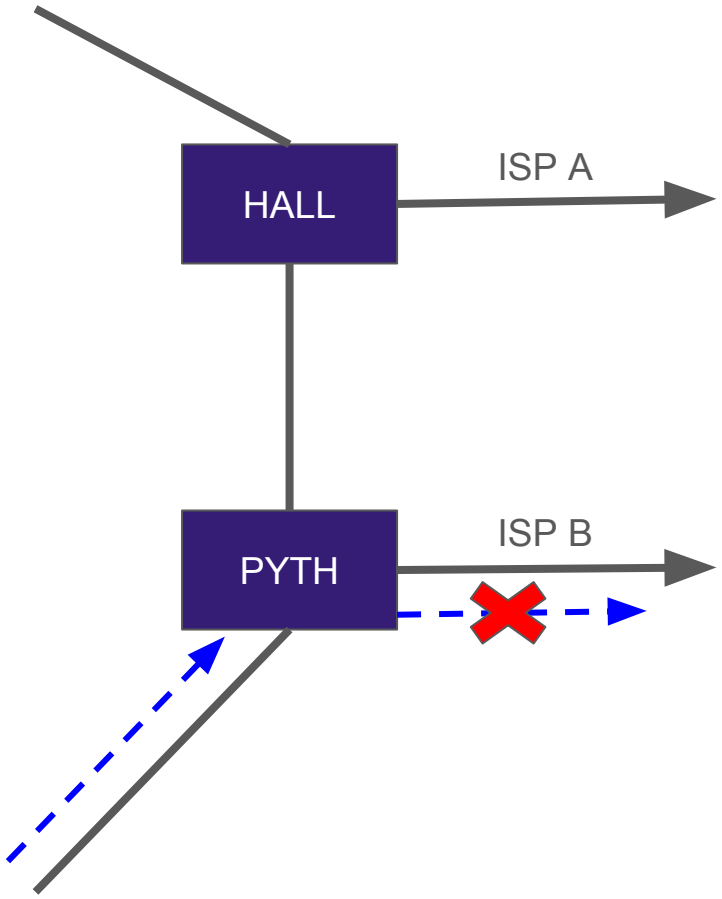
- ISP A assigns 2001:db:a::/48
- ISP B assigns 2001:db:a::/48
- Both advertize ::/0 over BGP



BCP 38 mandates some kind of source routing

src: 2001:db8:a::1
dst: 2001:4860:4860::8888

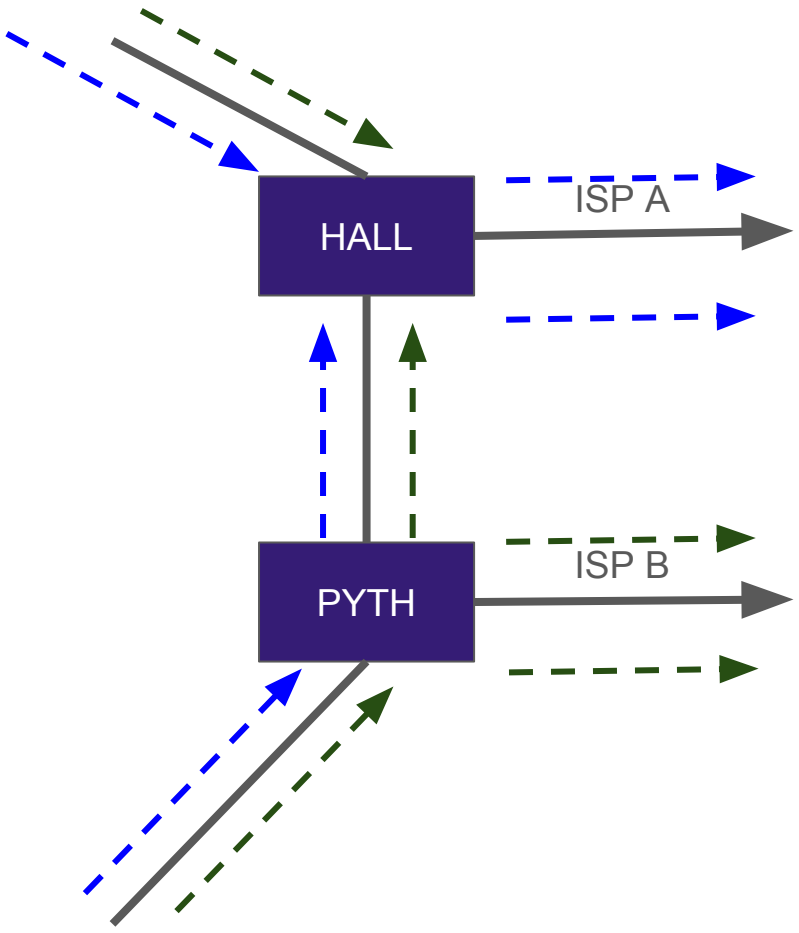




BCP 38 mandates some kind of source routing

src: 2001:db8:a::/48
dst: ‘outgoing traffic’

src: 2001:db8:b::/48
dst: ‘outgoing traffic’



Deploying PBR at the edges ensures no BCP 38 drop will occur

Assuming we are connected to ISP A

```
ip rule add from 2001:db8:b::/48
              to 2001:db8:b::/48 pref 1 table main

ip rule add from 2001:db8:b::/48
              to 2001:db8:a::/48 pref 1 table main

ip route add ::/0 via <other edge router> table 2

ip rule add from 2001:db8:b::/48 pref 10 table 2
```

Internal traffic is forwarded using the main FIB

Assuming we are connected to ISP A

```
ip rule add from 2001:db8:b::/48
              to 2001:db8:b::/48 pref 1 table main

ip rule add from 2001:db8:b::/48
              to 2001:db8:a::/48 pref 1 table main
```

```
ip route add ::/0 via <other edge router> table 2

ip rule add from 2001:db8:b::/48 pref 10 table 2
```

Outgoing traffic using the source address from ISP B must use the other egress

Assuming we are connected to ISP A

```
ip rule add from 2001:db8:b::/48
              to 2001:db8:b::/48 pref 1 table main

ip rule add from 2001:db8:b::/48
              to 2001:db8:a::/48 pref 1 table main

ip route add ::/0 via <other edge router> table 2

ip rule add from 2001:db8:b::/48 pref 10 table 2
```

Encapsulate the traffic forwarded to the alternate egress

Assuming we are connected to ISP A

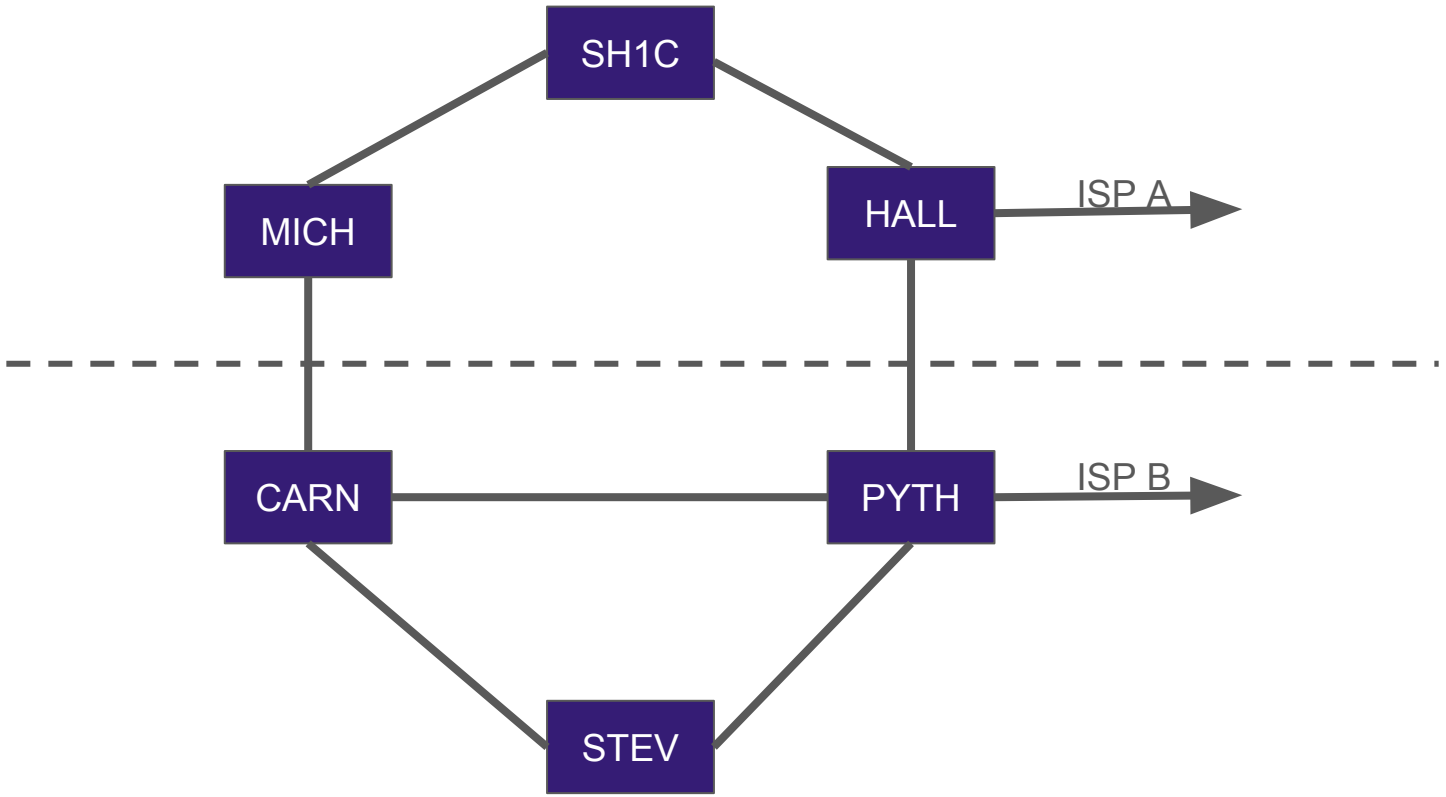
```
ip rule add from 2001:db8:b::/48
              to 2001:db8:b::/48 pref 1 table main

ip rule add from 2001:db8:b::/48
              to 2001:db8:a::/48 pref 1 table main

ip route add ::/0 via <other edge router> table 2

ip rule add from 2001:db8:b::/48 pref 10 table 2
```

Multihoming without NAT is challenging

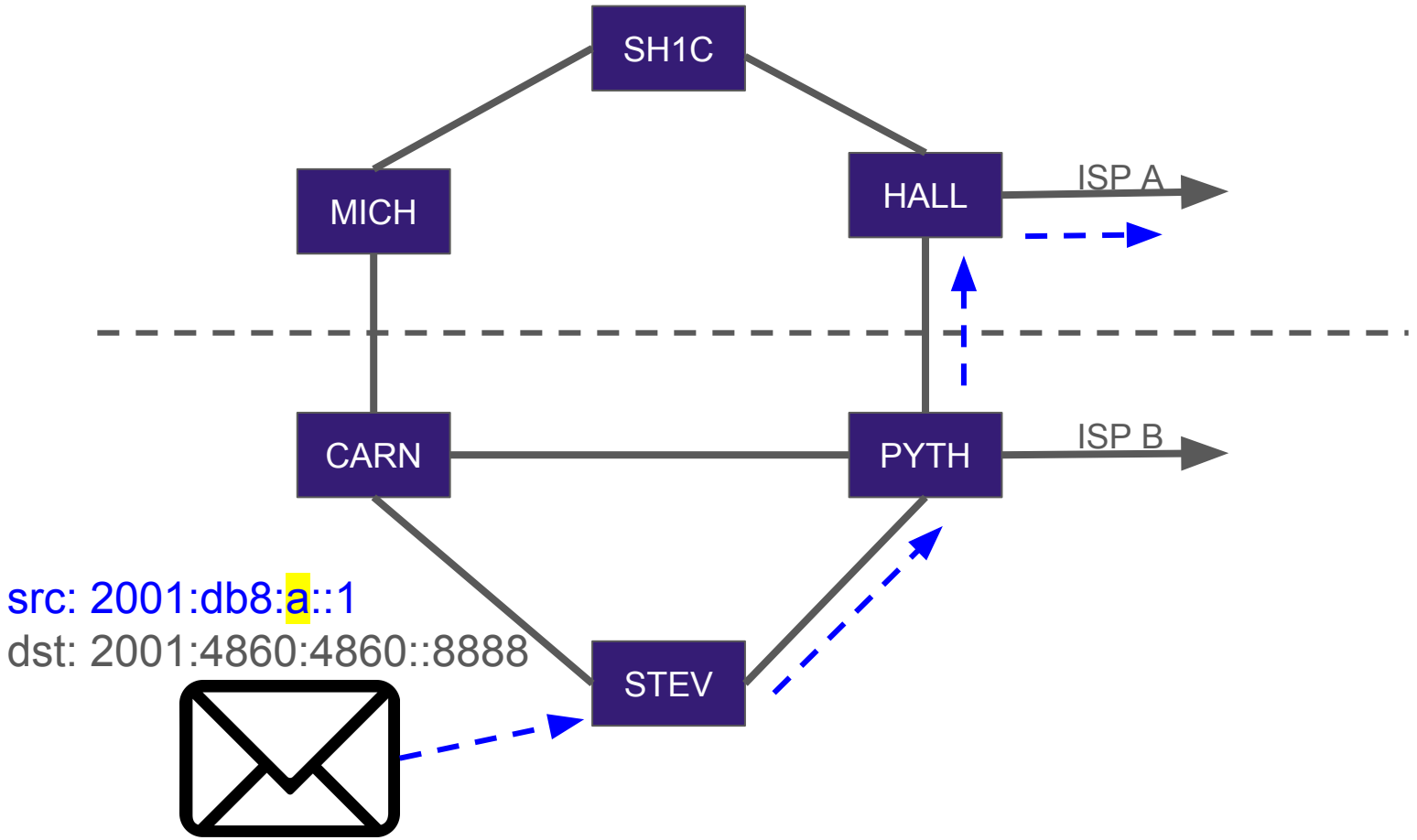
- BCP 38, BCP 38, BCP 38
- Internal routing policies
- Withstanding failures

Simple policy

Balance outgoing flows across egresses



Hosts choosing a source address from 2001:db8:a::/48 could violate the policy



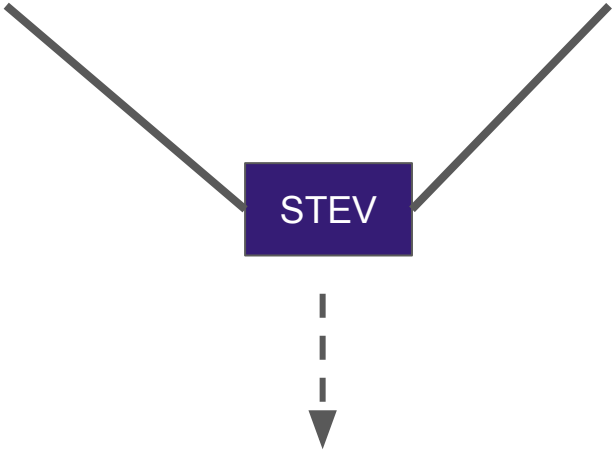
Implementing routing policies requires to influence the source address selection of the end hosts

RFC 6724 §5 specifies IPv6 source address selection rules

- Rule 3: Avoid deprecated addresses
- Rule 5.5: Prefer address in a prefix advertized by the next hop
- Rule 8: Longest prefix match

Going nuclear
Configuring the lifetime of RA PIOs

- We want hosts behind STEV to favor addresses from ISP B
- Can statically define which addresses is used
- Independent from the destination prefix

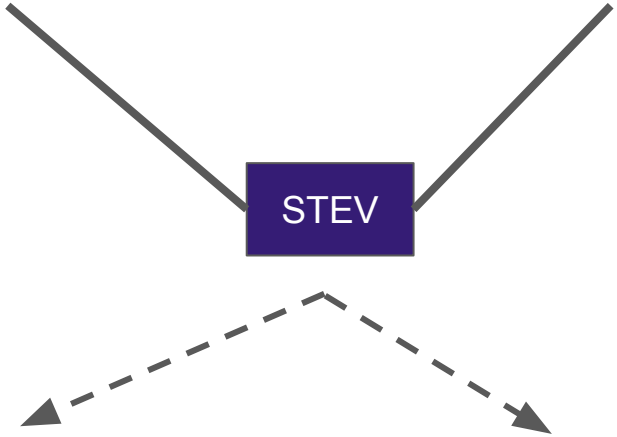


Router Advertisement
PIOs
2001:db8:a:101::/64
preferred_lft=0
2001:db8:b:101::/64
preferred_lft=1234

RIO
::/0

Finer-grained policies

Using extra LL addresses on the router



Router Advertisement
(using LL_STEV_1)
PIO
2001:db8:b:101::/64

RIO
::/0

Router Advertisement
(using LL_STEV_2)
PIO
2001:db8:a:101::/64

RIO
<prefix we want to reach through ISP A>

Finer-grained policies can be achieved by using extra LL addresses on the router

- Rule 3: Avoid deprecated addresses
- Rule 5.5: Prefer address in a prefix advertized by the next hop
- Rule 8: Longest prefix match

Multihoming without NAT is challenging

- BCP 38, BCP 38, BCP 38
- Internal routing policies
- Withstanding failures

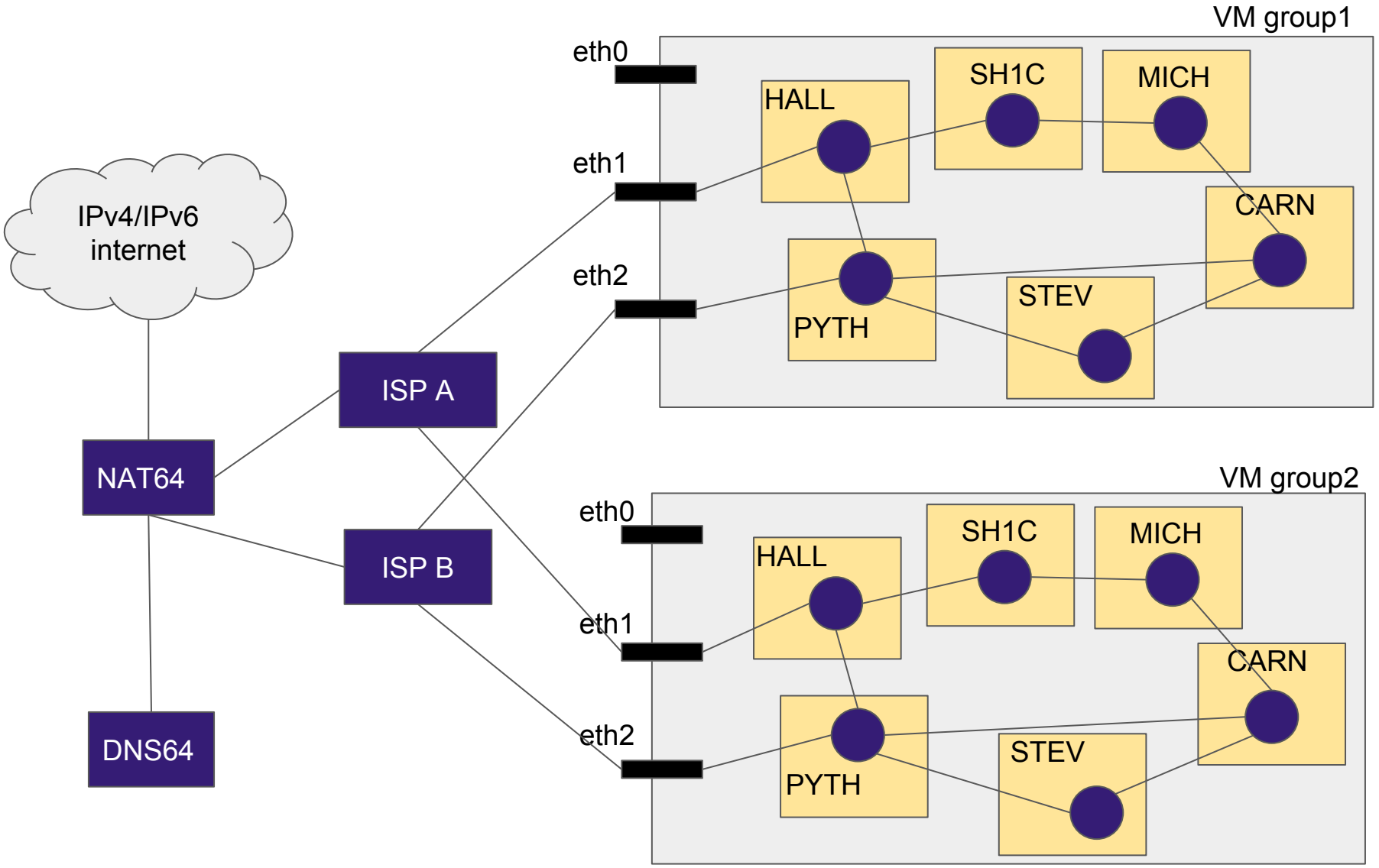
End hosts cannot use source addresses from a failed uplink for egress traffic

- Generate ICMPv6 T1/Code 5 (source address failed ingress/egress policy) at failed egress
- Must bind RA generation to uplink status
- No standardized mechanism
  to inform DHCP/DNS/... servers of network events
- MPTCP at the end hosts enables seamless handovers

You can experiment yourself today at https://github.com/UCL-INGI/lingi2142

- Scripts to build a VM and create an emulated network
- Scripts to interconnect a cluster of multihomed VM
- *<Soon>* configuration/scripts from the students

Hosted VM cluster



Multihoming in an IPv6 world

More details available in draft-bowbakova-rtgwg-enterprise-pa-multihoming-01

Olivier Tilmans
olivier.tilmans@uclouvain.be
https://github.com/UCL-INGI/lingi2142