Integrating Networks

Michael Lyngbøl <lyngbol@wheel.dk>, Nina Bargisen <nihb@tdc.dk>
TDC
NANOG41
Oct 14-16 2007

Abstract

- In 2004 TDC AS3292 acquired Song Networks AS3246. It was decided relatively fast that the two networks were to be integrated in order to have a common network with products available throughout the area spanned by the two original networks.
- The main product to be supported was RFC4364 VPNs (L3 MPLS VPN) with access connections throughout the area of the two original networks.

Straightforward Integration

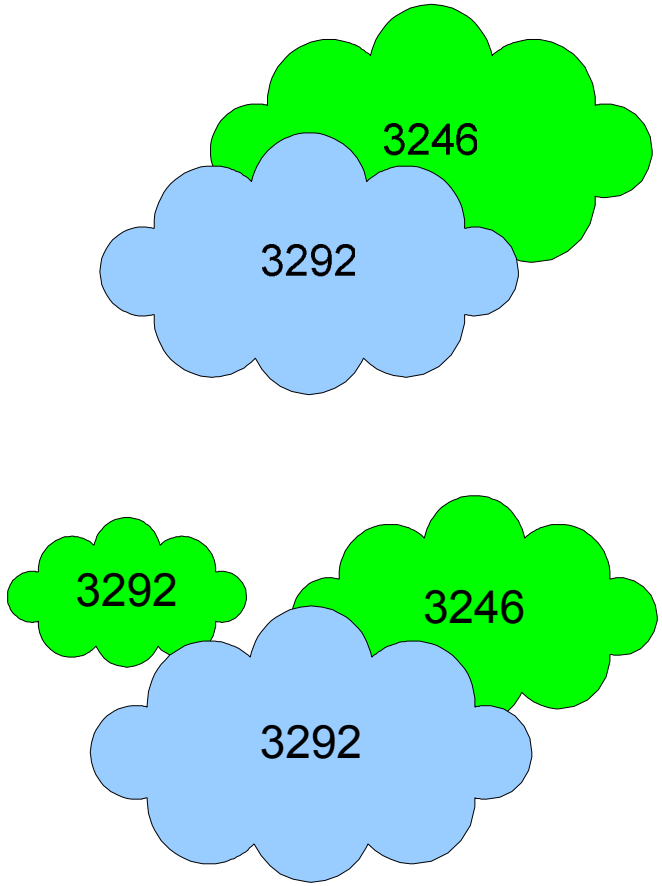
- Both networks were using IS-IS as IGP, BGP for VPN and internet routes.
- So:
- Merge the IS-IS domains
- Change the AS number on PE and Ps
- Align configuration

- DONE!

The Facts

- The acquired network consisted of 4 geographically areas – each operated and designed by local groups of engineers. A fact that was not changed by the acquisition
- Each area had a different set of issues to be taken care off. A few examples:
  - two parallel topologies spanning the same area
  - Old equipment needing replacement
- This meant that integration of the whole acquired network in one big operation was not possible

The plan in short

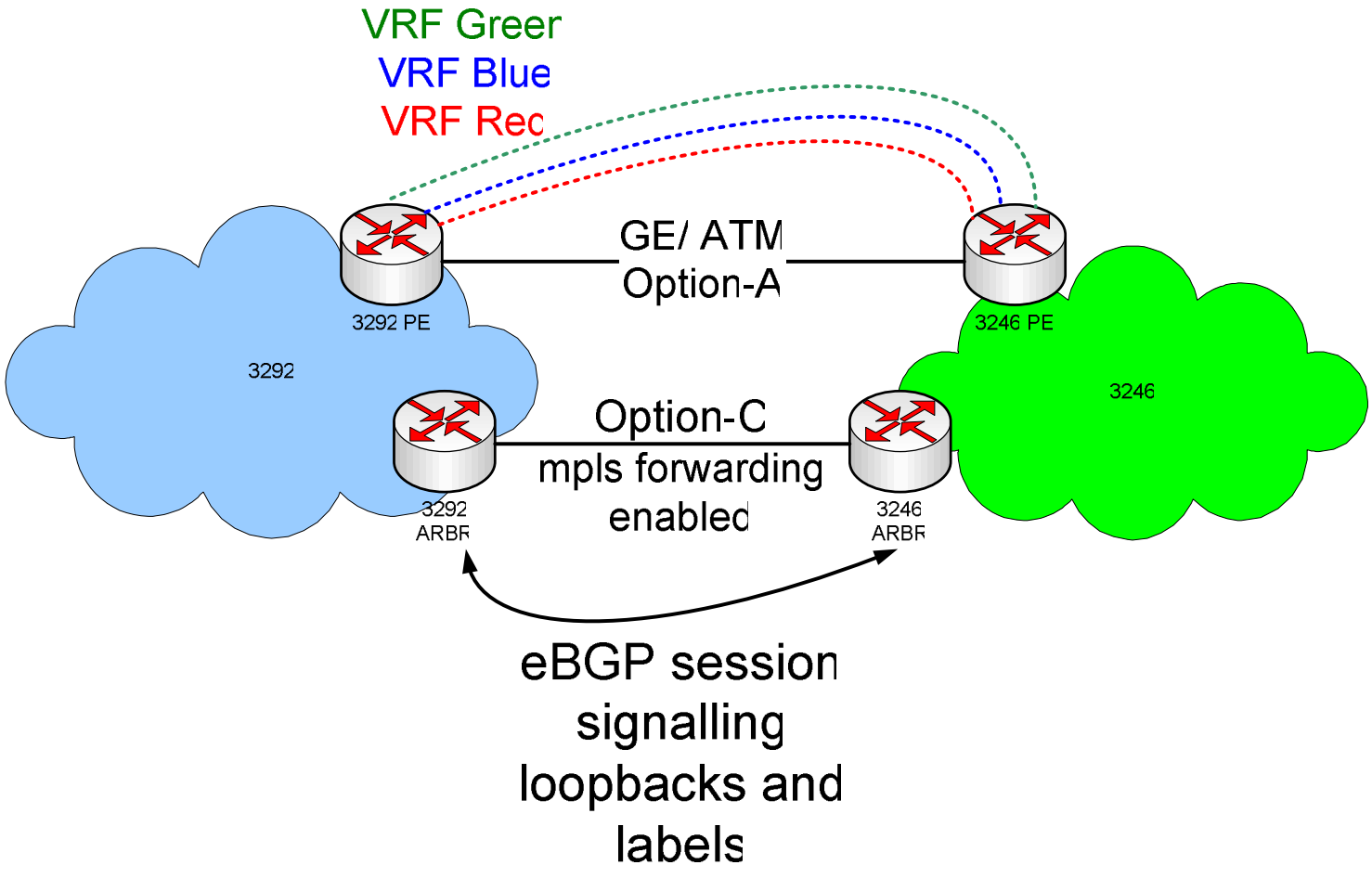


- Split area to be integrated from the rest of the acquired network
- Integrate IS-IS and BGP
- When done start on the next area

RFC4364 jump start

- Use RFC4364 Option-A to connect VRFs in 3292 and 3246 back to back.
  Don’t do this in a big scale, you’ll feel the pain later on when doing Option-C or full IGP integration.
- Use RFC4364 Option-C to connect MPLS networks.
  - eBGP + send-label on border routers gives you MPLS LSP from end-to-end.
  - eBGP between VPNv4 RRs to exchange VPNv4 routes.

RFC4364 jump start



RFC4364 Option-C

- signal loopback adresses and labels for these using eBGP between the two AS’s
- redistribute loopbacks from BGP into ISIS
- enable MPLS forwarding on links between networks
- then LSP’s between PE’s in the two network are set up and can forward traffic.

RFC4364 Option-C

- ISIS-> BGP (IOS)

- router bgp 3292
- address-family ipv4
- redistribute isis level-2 route-map OPTC-ISIS-BGP
- neighbor 1.2.3.2 route-map OPTC-BGP-IN in
- neighbor 1.2.3.2 route-map OPTC-BGP-OUT out
- neighbor 1.2.3.2 send-label explicit-null

- Take care not to redistribute own IS-IS loopbacks into iBGP
-

RFC4364 Option-C

- route-map OPTC-ISIS-BGP permit 10
- match ip address prefix-list BB-LOOPBACK
- set origin igp
- !
- route-map OPTC-BGP-IN permit 20
- match ip address prefix-list PEER-BB-LOOPBACK
- match mpls-label
- set metric 0
- set community no-advertise
- !
- route-map OPTC-BGP-OUT permit 10
- match ip address prefix-list BB-LOOPBACK
- set metric 0
- set mpls-label
- set community no-advertise
- !
- ip prefix-list BB-LOOPBACK permit 10.0.0.0/24 ge 32
- ip prefix-list PEER-LOOPBACK permit 192.168.0.0/24 ge 32

RFC4364 Option-C

- BGP -> ISIS (IOS):
- router isis
- redistribute maximum-prefix 1000 95 withdraw
- redistribute bgp 3292 route-map OPTC-BGP-ISIS metric-type external
- !
- route-map OPTC-BGP-ISIS permit 10
- match ip address prefix-list PEER-LOOPBACK
- !
- ip prefix-list PEER-LOOPBACK permit 192.168.0.0/24 ge 32

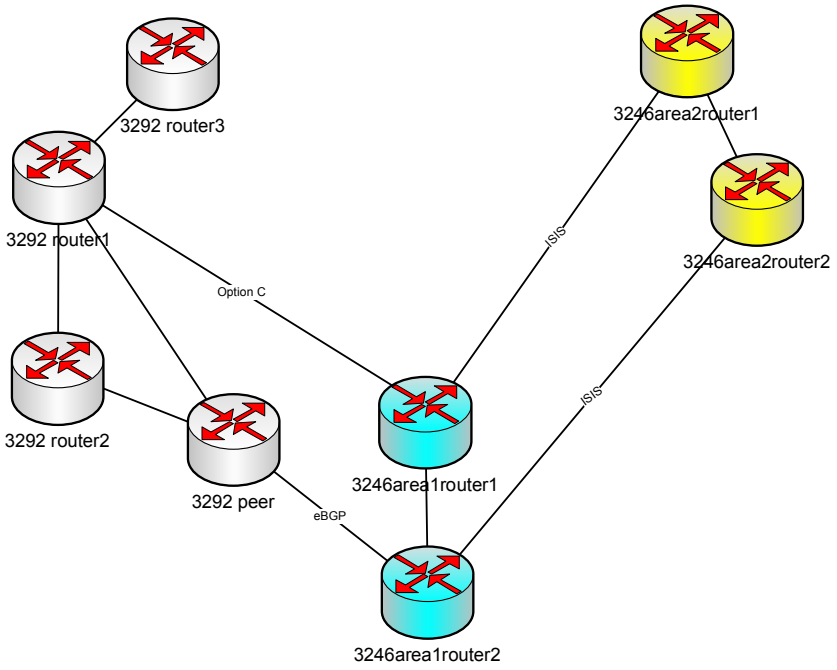
RFC4364 Option-C

- router#sh ip bgp 192.168.0.1/32
- BGP routing table entry for 192.168.0.1/32, version 295665
- Paths: (1 available, best #1, not advertised to any peer)
- Not advertised to any peer
- 3246
- 1.2.3.2 from 1.2.3.2 (192.168.0.100)
- Origin incomplete, localpref 100, valid, external, best
- Community: no-advertise
- **mpls labels in/out 2641/576**

RFC2547 Opt-C (BGP signaled MPLS IGP/next-hop labels)

- router#sh ip cef 192.168.0.1
- 192.168.0.1/32, version 137524, epoch 0, cached adjacency 1.2.3.2
- 0 packets, 0 bytes
- tag information set, all rewrites owned
- local tag: 2641
- fast tag rewrite with Gi2/1, 1.2.3.2, **tags imposed {576}**
- via 1.2.3.2, 0 dependencies, recursive
- next hop 1.2.3.2, GigabitEthernet2/1 via 1.2.3.2 /32 (Default)
- valid cached adjacency
- tag rewrite with Gi2/1, 1.2.3.2, **tags imposed {576}**

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels



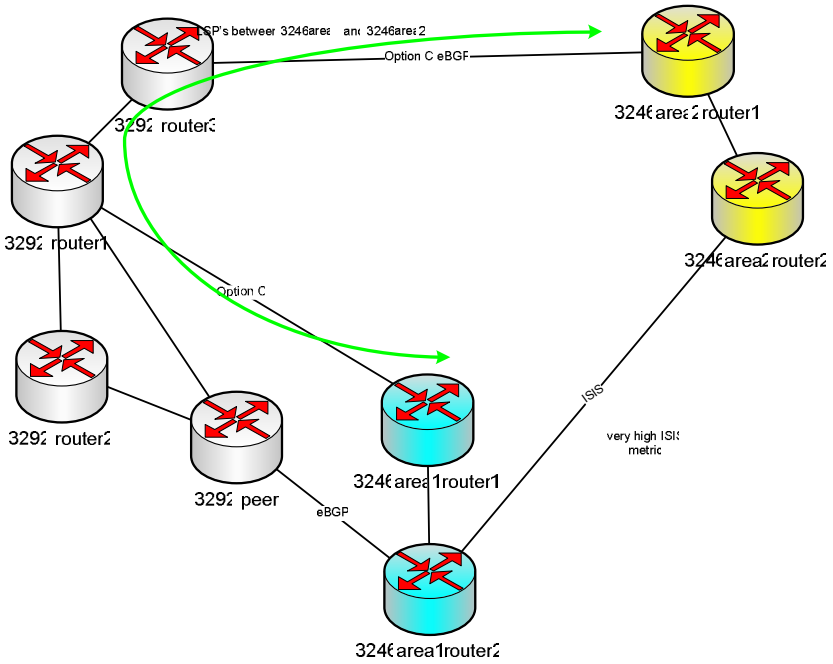
- 3246area1 and 3246area2 are connected with ISIS links
- 3246area1 is connected to 3292 with a dedicated Option C connection and an IP transit connection.

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels

- Split 3246 into two IGP domains in order to integrate one of the domains
  - Necessary to flush unwanted IS-IS LSPs (not to flooded into the other network)
  - IS-IS authentication issues as authentication not possible on all LSPs in 3246

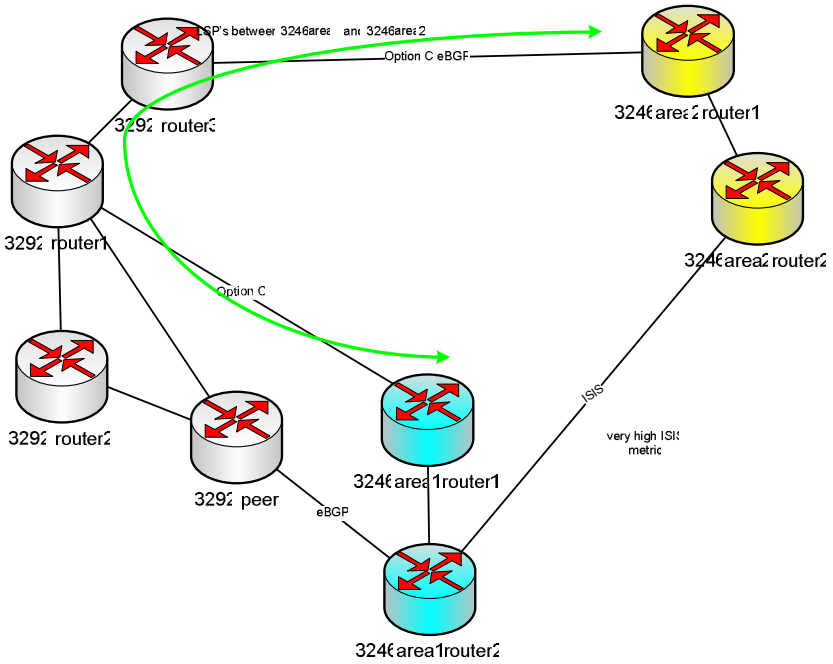
- To minimize time for LSPs to be flushed, configure low LSP lifetime in the network to be integrated (we used default lifetime)

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels



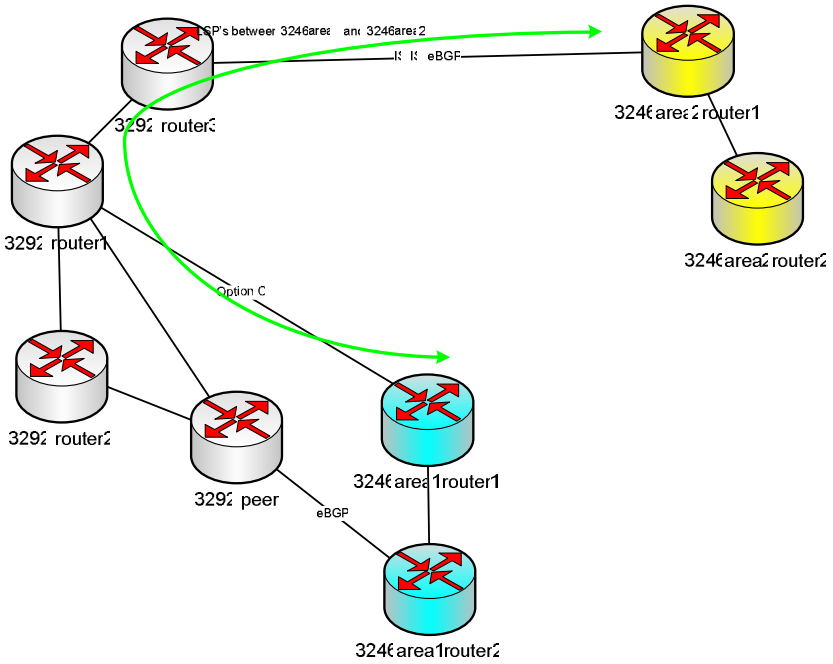
- the link between 3292 and 3246area2 is configured with two vlans
- vlan 1 is Option C vlan 2 is IP transit
- 3246area2 loopbacks and loopback labels are signaled on the vlan1 connection from 3246area2 to 3292
- 3246area1 loopbacks and loopback labels are signaled on the vlan1 connection from 3292 to 3246area2

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels



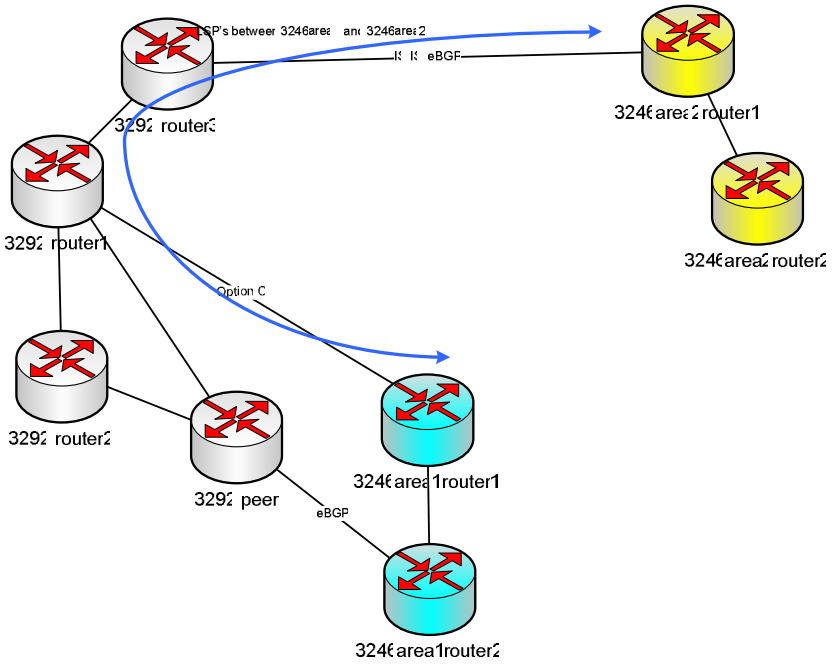
- The Option C connection between 3292 and 3246area1 is changed to only signal 3246area1 loopbacks and loopback labels from 3246area1 to 3292
- 3246area2 loopbacks and loopback labels are signaled from 3292 to 3246area1
- Raising the metric on the ISIS link between 3246area1 and 3246area2 causes LSP's between the two areas to run via 3292.
- ISIS link between the two areas can now be shut

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels



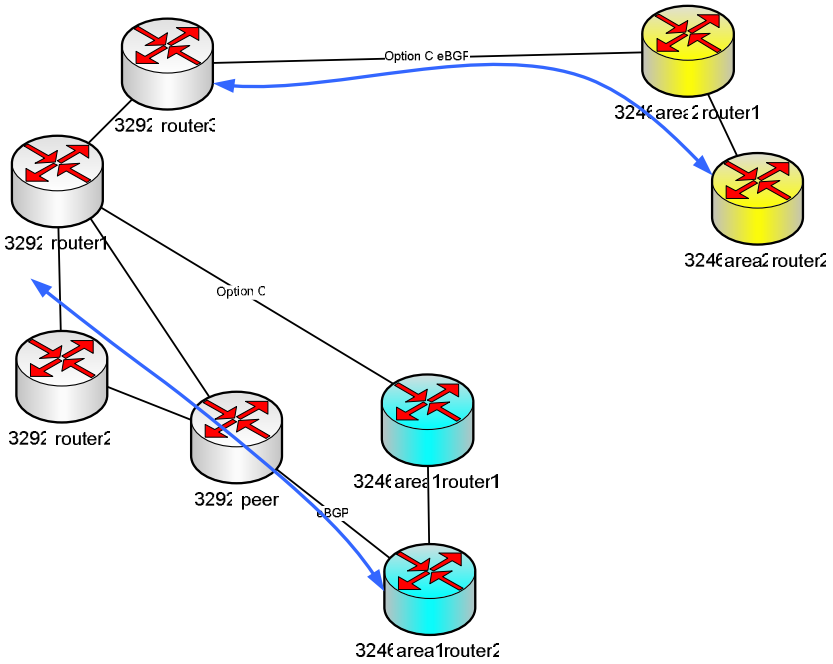
- When IS-IS between 3246area1 and 3246area2 is shut and the IS-IS database in 3246area2 is purged, then shut the option C connection between 3292 and 3246area2 and configure IS-IS on the link
- The IP transit session on the link remain unchanged
- DONE!

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels



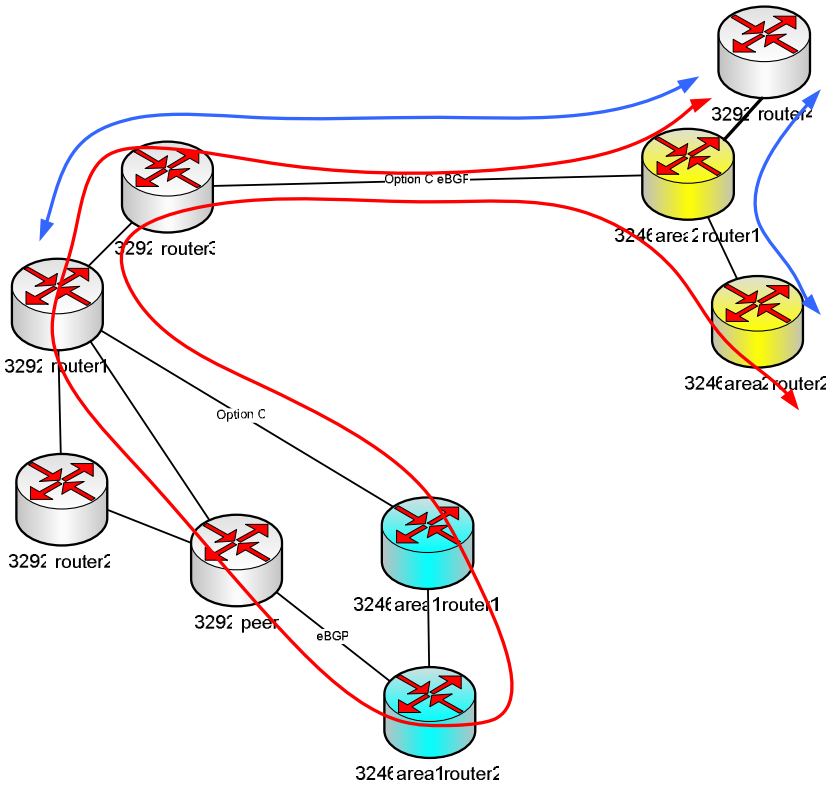
- Internet traffic flow between customer in area1 and customers in area2 and 3246 peers connected in area2
- Internet traffic flow between customer in area2 and customers in area1 and 3246 peers connected in area1

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels



- Internet traffic flow between customers in area1 and customers in 3292 and rest of Internet
- Internet traffic flow between customer in area2 and customers in 3292 and rest of Internet

Splitting 3246 in two using BGP signaled MPLS IGP/next-hop labels



- Internet traffic flow between customer on a 3292 PE router connected in area2 and customers in 3292 and rest of Internet
- Leaks of more specific 3246area2 prefixes to 3292 PE connected in 3246area2 are necessary in order to avoid traffic to trombone via 3246area1

IS-IS – checklist

- Low LSP lifetime in the domain to be integrated is nice when waiting for IS-IS database to purge
- Authentication (or lack of) must be consistent in the two domains to be integrated

IS-IS size

- 3292 is one flat L2 only network. Before integration ~ 1300 LSP (incl. pseudo nodes) in total
- 3246 is one flat L2 only network. ~ 600 LSP before integration.
- With the integration of 2/3 of 3246 we now have a L2 only network of ~ 1600 LSPs
- We’ve seen no need for a split into L1L2.
- No IS-IS issues what so ever.

IS-IS experience
platforms and software

- Cisco 7200/7300/7600/GSR/CRS-1
  IOS 12.3/12.2SB/12.2SRA/12.0S/3.4.2
- Juniper M10i/M160/M320
  JUNOS 7.5/8.2E

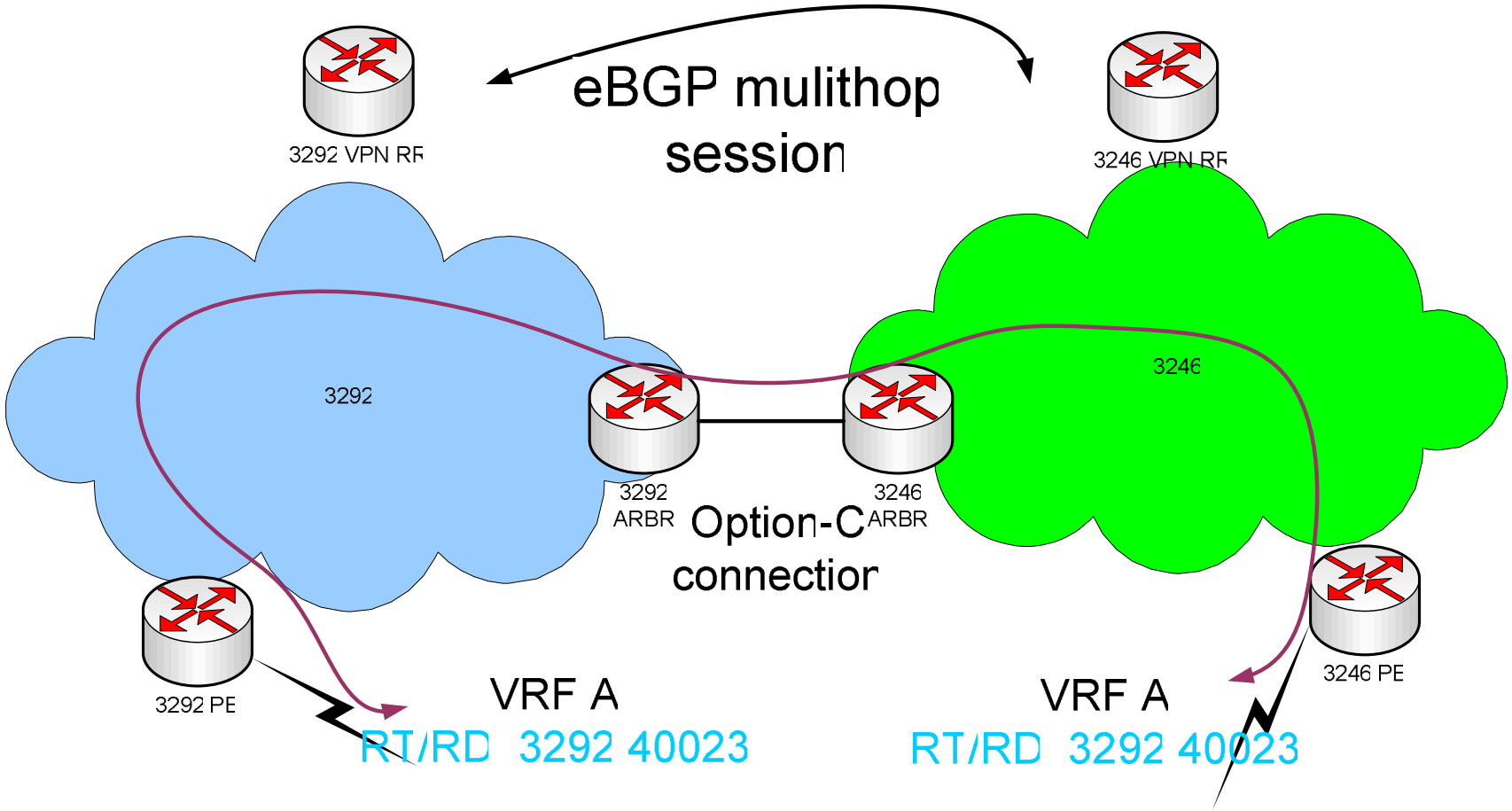
RFC4364 Opt-A vs. Opt-C

- Opt-A is the easy way to interconnect two VPNs but
  - Configuration is necessary on a per customer per VPN per interface basis
  - Capacity is not easily added
  - QoS mapping is troublesome (no end-to-end QoS)
  - PE interconnect router scalability is an issue as PE needs to carry all interconnected VRFs VPNv4 tables

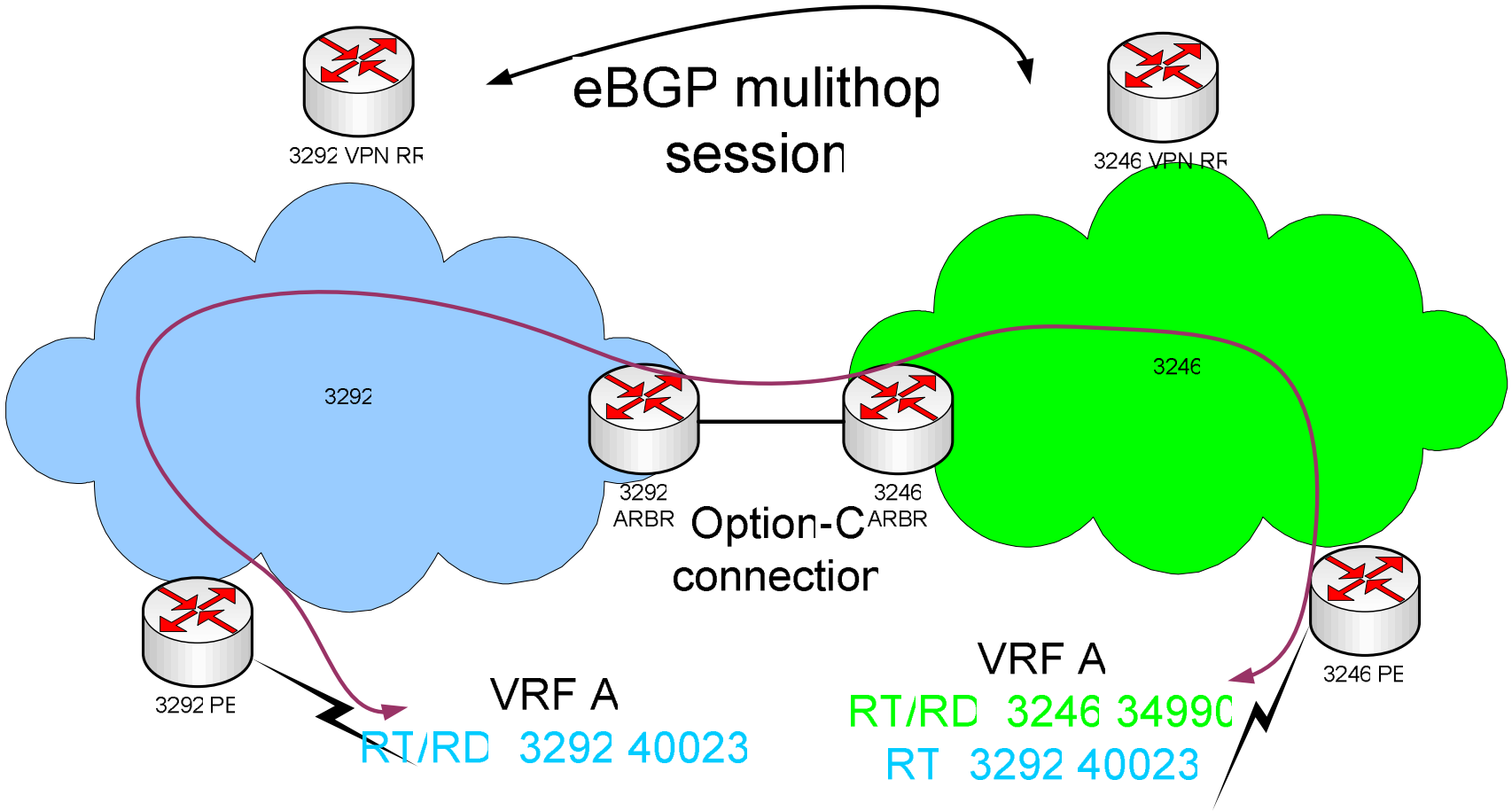
RFC4364 Opt-A vs. Opt-C

- Opt-C is a better way to interconnect networks as:
  - there is no need for customer configurations on the interconnects between the networks
  - VRFs can be transparently configured on PEs in both networks
  - end-to-end QoS works
  - EoMPLS/VPLS will work by default
  - capacity can be added easily

Route reflector design for RFC4364 Option-C VPNs



Route reflector design for RFC4364 Option-C VPNs



Status for the integration

- Two domains are integrated
- No major issues happened during the integrations
- Last domain will be integrated this fall – preparations are on-going

Q & A

- nihb@tdc.dk
- lyngbol@wheel.dk