1. srečanje članov SIX

1. srečanje članov SIX

- IP-naslovi (naslovna shema)
- preštevilčenje
- nekaj iz prakse

photo: http://www.pivo-lasko.si/

Naslavljanje

- en samo IP-podomrežje (PI za IPv4, poseben prostor za IXP-je za IPv6)
  - naslov članа je dodeljen glede na lokacijo

91.220.194.n/24
n = n_1 = 2..99 na TPL
n = n_1 + 100 = 102..199 na IJS
n = 1, 101 route-reflektorja
n = 254 - testni looking-glass usmerjevalnik

2001:7f8:46:0:L:N::<AS>/64
L = 0 na IJS
L = 1 na TPL
N = 0 za en sam usmerjevalnik, sicer N = 1, 2, ...
AS = članov AS (decimalno)
AS = 51988 za RR

PI for IPv4, IPv6 Assignment for an Internet Exchange Point (IXP) for IPv6 (see RIPE-460 and RIPE-451/policy).

Novi naslovi

AS	član	as-name	zap.št.	IPv4 TPL		IPv4 IJS	
				stari naslov	novi naslov	stari naslov	novi naslov
AS2107	Arnes	ARNES-NET	2	193.2.141.1	91.220.194.2	193.2.141.33	91.220.194.102
AS3212	Telemach (Triera)	TRIERA	3	-	91.220.194.3	193.2.141.48	91.220.194.103
AS5435	Velcom	VELCOM-AS	4	193.2.141.4	91.220.194.4	193.2.141.36	91.220.194.104
AS5603	Telekom Slovenije	SIOL-NET	5	193.2.141.6	91.220.194.5	193.2.141.38	91.220.194.105
AS6764	Perftech	PERFTECH-SLOVENIA-AS	6	-	91.220.194.6	193.2.141.37	91.220.194.106
AS8591	AMIS	AMIS	7	-	91.220.194.7	193.2.141.40	91.220.194.107
AS9119	Softnet	SOFTNET-AS	8	-	91.220.194.8	193.2.141.41	91.220.194.108
AS12644	Telemach	TELEMACH	9	-	91.220.194.9	193.2.141.42	91.220.194.109
AS12778	NETSI.NET	NETSI	10	193.2.141.3	91.220.194.10	193.2.141.35	91.220.194.110
AS16016	Tušmobil	VOLJATEL-AS	11	193.2.141.15	91.220.194.11	193.2.141.47	91.220.194.111
AS21283	Simobil	SIMOBIL-AS	12	193.2.141.20	91.220.194.12	193.2.141.52	91.220.194.112
AS33929	Telemach (Lj-kabel)	Masicom-AS	13	-	91.220.194.13	193.2.141.49	91.220.194.113
AS34779	T-2	T-2-AS	14	193.2.141.19	91.220.194.14	193.2.141.51	91.220.194.114
AS39765	IT TEL	ITTEL-AS	15	-	91.220.194.15	193.2.141.39	91.220.194.115
AS39912	i3B(ASCUS)	i3B-AS	16	-	91.220.194.16	193.2.141.__	91.220.194.116
AS41427	Datacenter	marc-net_AS	17	-	91.220.194.17	193.2.141.45	91.220.194.117
AS41828	Tušhosting	TUSMOBIL	18	-	91.220.194.18	193.2.141.43	91.220.194.118
AS43061	Stelkom	SI-STELKOM	19	193.2.141.12	91.220.194.19	193.2.141.44	91.220.194.119
AS196841	KRS Networks	KRS-NET-AS	20	-	91.220.194.20	193.2.141.53	91.220.194.120
AS44549	MEGA M	MEGA M	21	-	91.220.194.21	193.2.141.54	91.220.194.121
AS51988	route-server na lokaciji TPL			-	91.220.194.1	-	-
AS51988	route-server na lokaciji IJS			-	-		91.220.194.101

Prešteviličenje

- **predpriprava**
  - pripravimo konfiguracijo novih peering-ov
    naj bodo izklopljeni (neighbor <x> shutdown)
  - dopolnimo prefiks-liste, policy direktive/route-map-e, filtre, nadzorni sistemi...
  - po potrebi uredimo routing do 91.220.194.0/24
- **akcija**
  - 15.9.2011 ob 4.00 podremo vse peering-e (neighbor <x> shutdown)
  - zamenjamo naslov
  - preverimo povezljivost z 91.220.194.254
    (ping 91.220.194.254 source <interface>)
  - dvignemo vse nove peering-e (no neighbor <x> shutdown)
  - preverjamo stanje BGP in pravilnost usmerjanja prometa
- **čiščenje** (blokada BGP na starih naslovih)

Prešteviličenje - če gre kaj narobe

- dosegljivost 91.220.194.254 (LG)
- testni peering z LG
  za vsak primer lahko ta test naredite že danes :-)
  - brez ključa MD5 oz. z enakim ključem, kot ga imate v seji z obstoječim looking-glass usmerjevalnikom
  - LG oglašuje 153.5.250.0/24
  - LG se lahko peer-a tudi na starih naslovih 193.2.141.x

```
lljtpl1#show bgp ipv4 unicast neighbors 91.220.194.254 routes

   Network          Next Hop            Metric LocPrf Weight Path
*> 153.5.250.0/24   91.220.194.254                  170      0 51988 i

Total number of prefixes 1
```

Prešteviličenje - če gre kaj narobe

- klic v sili
- http://www.arnes.si/infrastruktura/six-sticisce-omrezij/status-prestevilcenja.html

- Telefon: 01-4798 911
  E-mail: hw-podpora@arnes.si
  Voda gori - dežurni telefon:
  041-793060

Nekaj iz prakse

- nastavitve vmesnikov
  - na stikalu SIX
  - na usmerjevalniku člana
- dve lokaciji
  - priporočila za lokaliziranje prometa s pomočjo oznak "BGP community"
- zakaj je pomemben "next-hop self"?
- zakaj naj izklopimo preusmeritve ("ICMP redirect")?

Vmesnik na dostopovnem stikalu

- Cisco 4900M
  - strogi “access” način
  - 1 MAC na vmesniku
  - izklop v primeru kršitve, ponoven vklop po 10 minutah

```
interface GigabitEthernet2/24
 switchport access vlan <N>
 switchport mode access
 switchport nonegotiate
 switchport port-security [maximum 2]
 load-interval 30
 storm-control broadcast level 1.00
 storm-control action shutdown
 spanning-tree portfast
 spanning-tree bpduguard enable
 service-policy input COUNTER_IPv4_IPv6
 service-policy output LIMIT-QUEUE-200
!
!
errdisable recovery cause bpduguard
errdisable recovery cause psecure-violation
errdisable recovery cause storm-control
errdisable recovery interval 600
```

```
class-map match-any IPv4_traffic
  match protocol ip
class-map match-any IPv6_traffic
  match protocol ipv6
!
policy-map COUNTER_IPv4_IPv6
 class IPv4_traffic
    police cir 32000
      conform-action transmit
      exceed-action transmit
      violate-action transmit
 class IPv6_traffic
    police cir 32000
      conform-action transmit
      exceed-action transmit
      violate-action transmit
!
policy-map LIMIT-QUEUE-200
 class class-default
    queue-limit 200
!
```

Vmesniki med obema stikaloma

- EtherChannel 2 x 10 Gb/s z LACP
- maksimalni MTU = 9198 Bytov na ethernetu
  - zagotavljamo 9100 Bytov za IPv4/6

```
interface TenGigabitEthernet1/1
 switchport access vlan <N>
 switchport mode access
 switchport nonegotiate
 mtu 9198
 load-interval 30
 channel-protocol lacp
 channel-group 48 mode active
!
interface TenGigabitEthernet1/2
 switchport access vlan <N>
 switchport mode access
 switchport nonegotiate
 mtu 9198
 load-interval 30
 channel-protocol lacp
 channel-group 48 mode active
!
```

```
interface Port-channel48
 switchport
 switchport access vlan <N>
 switchport mode access
 switchport nonegotiate
 mtu 9198
 bandwidth 10000000
!
port-channel load-balance src-dst-ip
```

Primer: konfiguracija vmesnika na usmerjevalniku člana

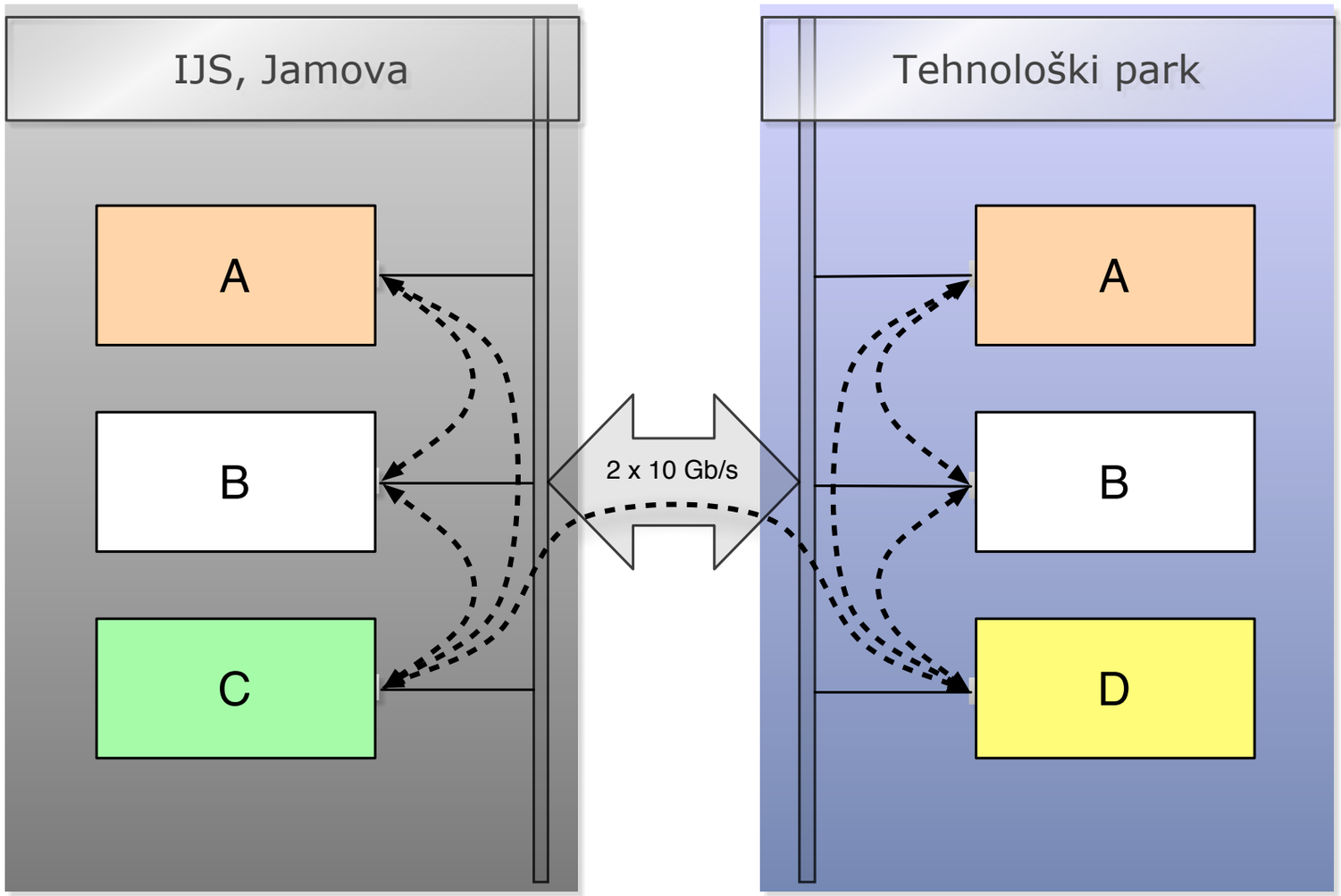
- izklopite vse, razen IP-ja in ARP-a
  - no ICMP redirects
  - brez nestandardnih protokolov, kot npr. CDP
  - brez odvečnih broadcast-ov
  - izklopite IPv6 RA
  - ! pozor: “ICMP unreachable”-sporočila so zelo pomembna za “PMTU discovery”

```
! primer za Cisco IOS
!
interface TenGigabitEthernet3/3
 ip address x.y.z.w 255.255.255.0
 ip access-group IxIncoming in
 ip access-group IxOutgoing out
 no ip redirects
 no ip proxy-arp
 ipv6 address 2001:.../64
 ipv6 enable
 ipv6 traffic-filter IxIncoming6 in
 ipv6 traffic-filter IxOutgoing6 out
 ipv6 nd reachable-time 300000
 ipv6 nd ra suppress
 no ipv6 redirects
 storm-control broadcast level 1.00
 no cdp enable
!
```

zanimivo branje:
http://www.ams-ix.net/config-guide/
http://www.vix.at/vix_bcps.html?&L=1

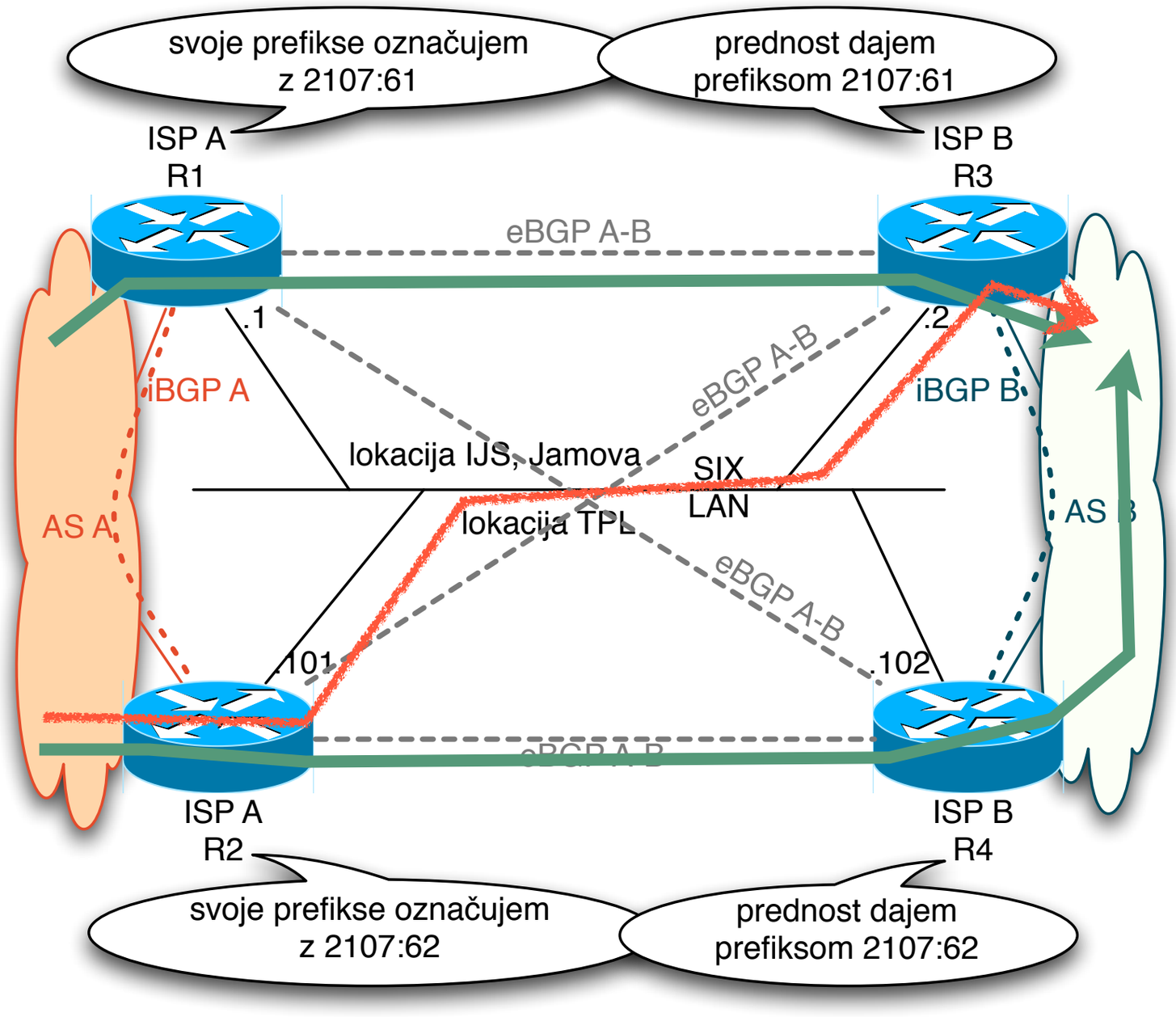
Dve lokaciji

- skrb za učinkovito in gospodarno usmerjanje prometa
  - lokalizirajmo promet
  - minimizirajmo promet med lokacijama



Promet med članoma na dveh lokacijah: prilagoditev metrike

- prefikse označimo glede na lokacijo, kjer jih oglašujemo
- prilagoditev metrike



✔ priporočljiva pot
✗ izognite se tej poti

Član na lokaciji IJS daje prednost prefiksom, ki so označeni za to lokacijo (2107:61).
Priporočamo prilagajanje metrike in ne spreminjanje “local-preference”, saj ima slednja visoko prednost v kriterijih za izbor aktivne poti v BGP-ju, celo višjo od dolžine AS-poti. Zato je prilagajanje metrike bolj primerno/enostavno.

Lokalizacija: primer konfiguracije

- Cisco IOS

```
! usmerjevalnik R3 na lokaciji IJS
ip community-list 61 permit 2107:61
!
route-map AnnounceToIX permit 10
 set community 2107:61
!
route-map AcceptFromIX permit 10
 ! this location (IJS)
 match community 61
route-map AcceptFromIX permit 20
 ! other location - worse metric
 set metric +1
!
router bgp <member-AS>
 template peer-policy IX
  route-map AcceptFromIX in
  route-map AnnounceToIX out
  next-hop-self
  send-community
!
 address-family ipv4|6
 neighbor <R1> inherit peer-policy IX
 neighbor <R2> inherit peer-policy IX
!
```

```
! usmerjevalnik R4 na lokaciji TPL
ip community-list 62 permit 2107:62
!
route-map AnnounceToIX permit 10
 set community 2107:62
!
route-map AcceptFromIX permit 10
 ! this location (TPL)
 match community 62
route-map AcceptFromIX permit 20
 ! other location - worse metric
 set metric +1
!
router bgp <member-AS>
 template peer-policy IX
  route-map AcceptFromIX in
  route-map AnnounceToIX out
  next-hop-self
  send-community
!
 address-family ipv4|6
 neighbor <R1> inherit peer-policy IX
 neighbor <R2> inherit peer-policy IX
!
```

Lokalizacija: primer konfiguracije

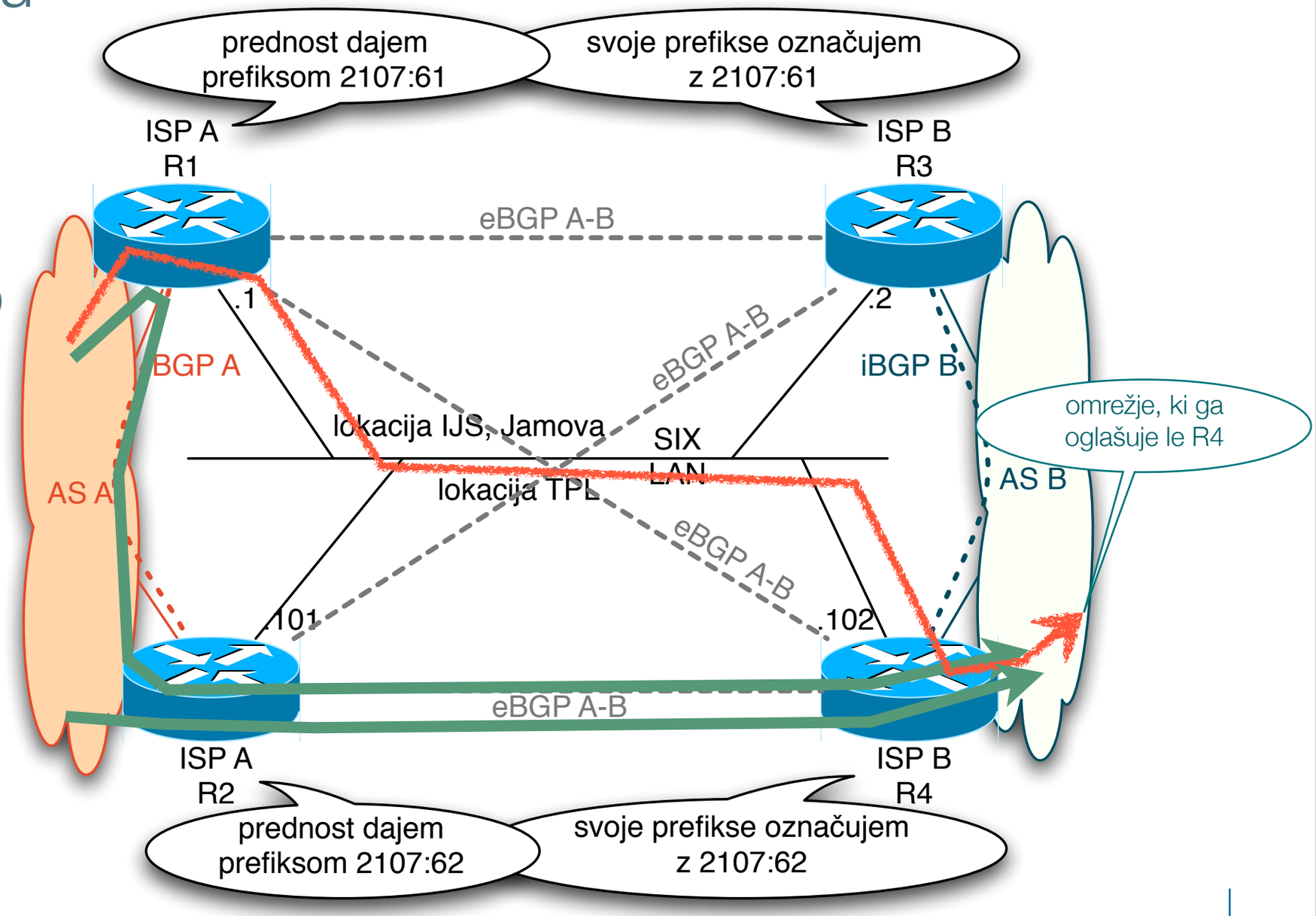
- Juniper JUNOS

```
/* usmerjevalnik na lokaciji IJS */
protocols {
    bgp {
        local-as <member-AS>;
        group Ix {
            type external;
            import [ LocalizeTraffic AcceptFromIx ];
            export AnnounceToIx;
        }
    }
}
policy-options {
    policy-statement AcceptFromIx {
        <member policy at receive>
    }
    policy-statement AnnounceToIx {
        term Localize {
            then {
                community set IxLocationIJS;
                next term;
            }
        }
        <member policy for announcements>
    }
    policy-statement LocalizeTraffic {
        term LocalTraffic {
            from community IxLocationIJS;
            then next policy;
        }
        term OtherTraffic {
            then {
                metric add 1;
            }
        }
    }
    community IxLocationIJS members 2107:61;
}
```

```
/* usmerjevalnik na lokaciji TPL */
protocols {
    bgp {
        local-as <member-AS>;
        group Ix {
            type external;
            import [ LocalizeTraffic AcceptFromIx ];
            export AnnounceToIx;
        }
    }
}
policy-options {
    policy-statement AcceptFromIx {
        <member policy at receive>
    }
    policy-statement AnnounceToIx {
        term Localize {
            then {
                community set IxLocationTPL;
                next term;
            }
        }
        <member policy for announcements>
    }
    policy-statement LocalizeTraffic {
        term LocalTraffic {
            from community IxLocationTPL;
            then next policy;
        }
        term OtherTraffic {
            then {
                metric add 1;
            }
        }
    }
    community IxLocationTPL members 2107:62;
}
```

Še en primer lokalizacije in “next-hop self”

- pomemben je “next-hop self” v iBGP-ju
- sicer bi šel promet z R1 direktno na R4

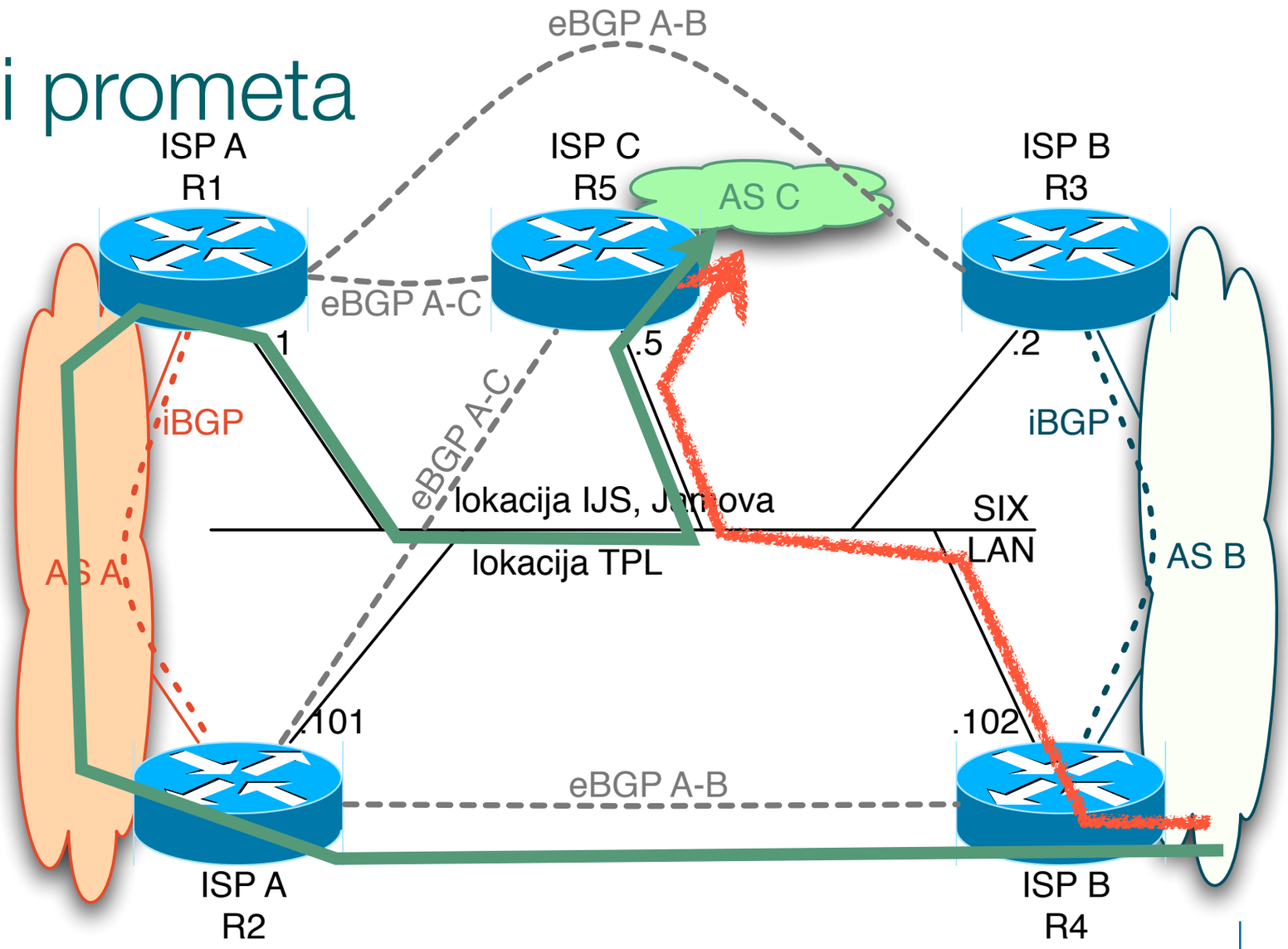


✔ priporočljiva pot
✗ izognite se tej poti

V iBGP-ju pomemben next-hop self (loopback-i) -- npr., nekatere B-jeve prefikse označuje le R4. Promet iz R1 mora iti preko A-jeve interne infrastrukture do R2 in od tam na R4 in ne direktno na R4. Brez next-hop self, bi bil next-hop direktno R4 (.102).

Še enkrat next-hop self in izklop preusmeritev

- A želi prejemati B-jev promet za C in ga posredovati C-ju
- B ne sme poslati prometa direktno C-ju
  - next-hop self tudi v eBGP-ju
  - "no ICMP redirects"



✔ priporočljiva pot
✗ izognite se tej poti

Recimo, da C peer-a z A, ne pa z B.
C oglašuje nek prefiks x A-ju z next-hop = R5 (.2). Predpostavimo, da A ta prefiks x posreduje naprej B-ju. A želi, da pride promet za x k njemu, zato nastavi next-hop nase (na .1 na R1 oz.na .101 na R2). Izklopiti mora tudi ICMP redirect sporočila. S tem prepreči, da bi B-ja (R4) obvestil, da je vrnil paket za x na isti vmesnik, kjer ga je sprejel in, da naj ta promet v bodoče pošilja direktno C-ju na .5 (C sploh nima peering-a z B-jem).

Vprašanja, komentarji, ...