kathara lab

bgp: multi-homed-stub with frr

Version	1.0
Author(s)	G. Di Battista, M. Patrignani, M. Pizzonia, F. Ricci, M. Rimondini
E-mail	contact@kathara.org
Web	http://www.kathara.org/
Description	configuration of a multi-homed stub network with backup – kathara version of a netkit lab

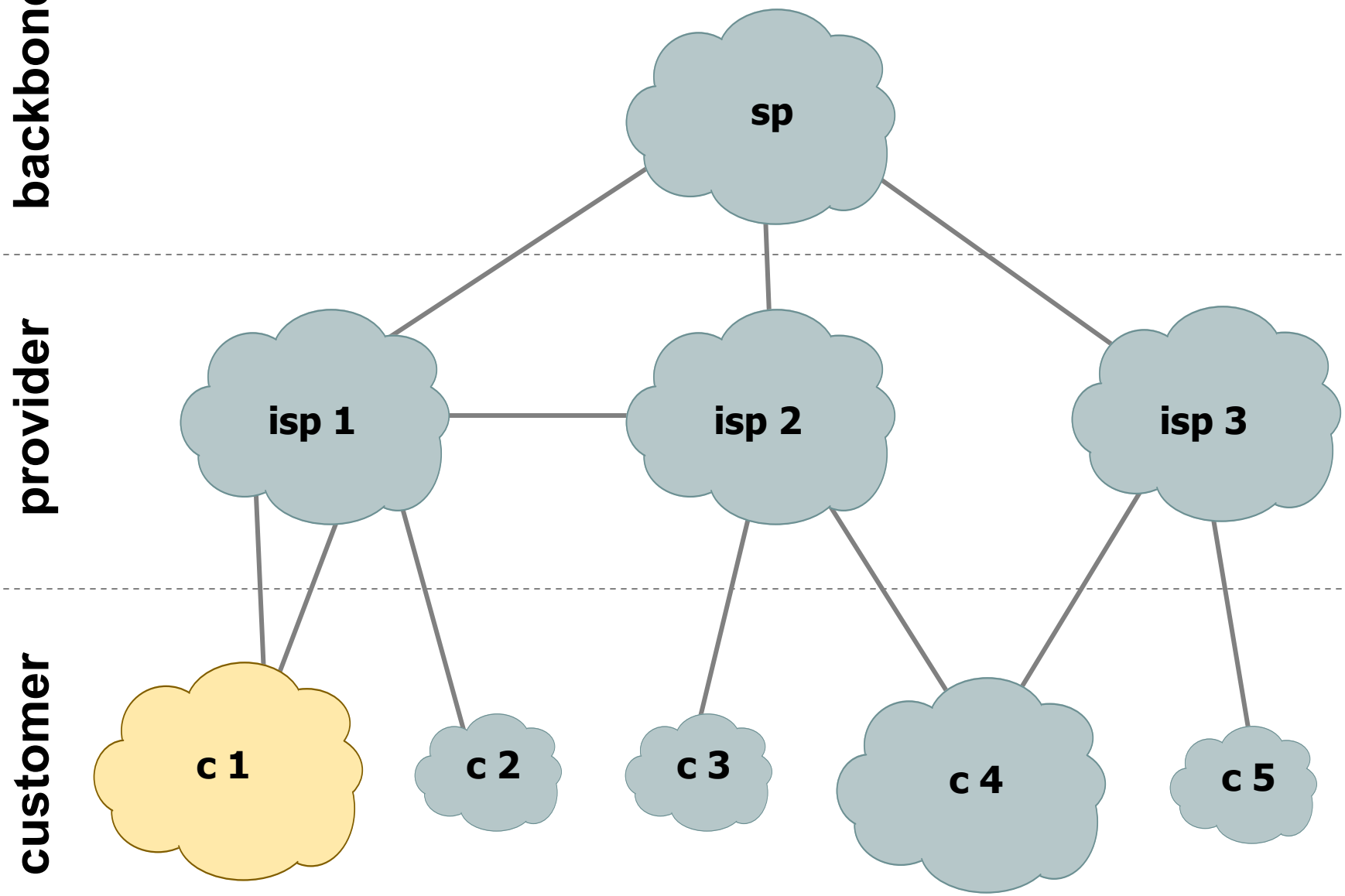
copyright notice

- All the pages/slides in this presentation, including but not limited to, images, photos, animations, videos, sounds, music, and text (hereby referred to as "material") are protected by copyright.
- This material, with the exception of some multimedia elements licensed by other organizations, is property of the authors and/or organizations appearing in the first slide.
- This material, or its parts, can be reproduced and used for didactical purposes within universities and schools, provided that this happens for non-profit purposes.
- Information contained in this material cannot be used within network design projects or other products of any kind.
- Any other use is prohibited, unless explicitly authorized by the authors on the basis of an explicit agreement.
- The authors assume no responsibility about this material and provide this material "as is", with no implicit or explicit warranty about the correctness and completeness of its contents, which may be subject to changes.
- This copyright notice must always be redistributed together with the material, or its portions.

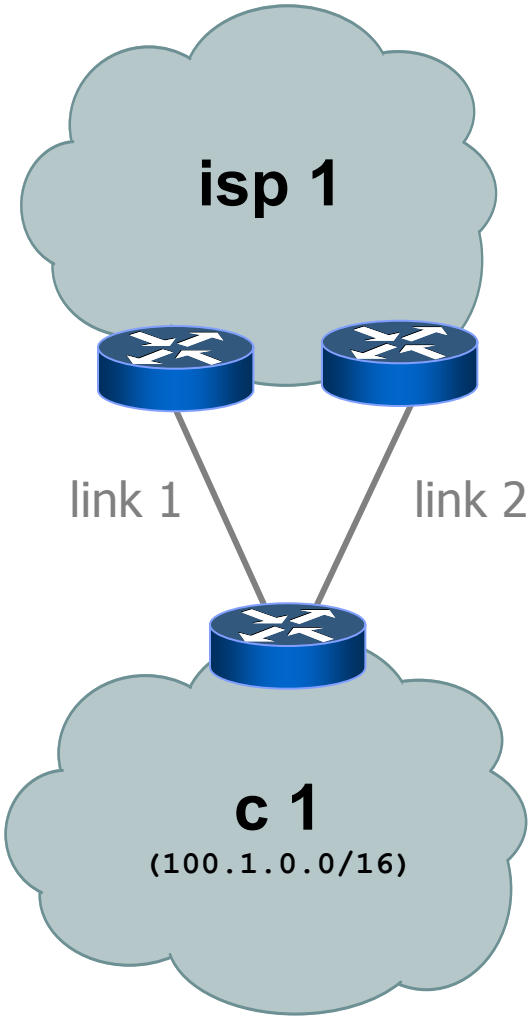
preconditions

- for this lab we assume you have chosen “kathara/frr” as the default image of your Kathará installation
  - execute “kathara settings”
    - select “choose default image”
    - select “kathara/frr”
    - exit from the settings procedure

multi-homed stub network

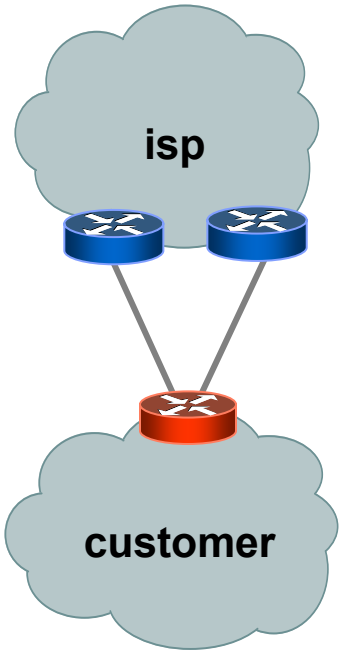


multi-homed stub network

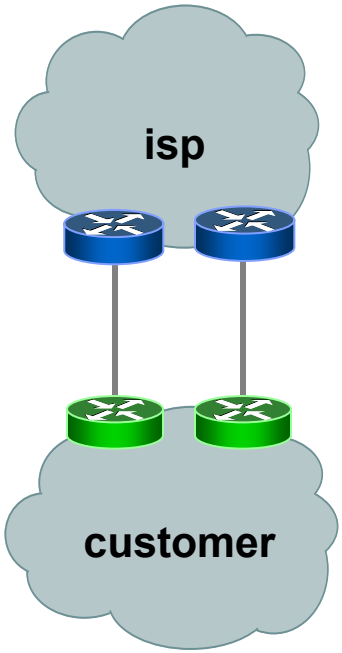


- two links to the same isp
- generally two routers of the customer as are involved

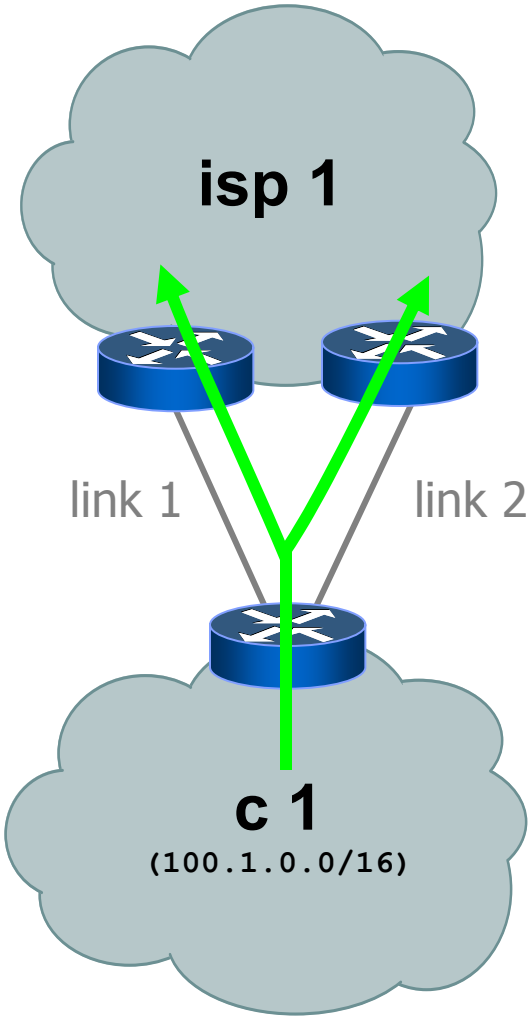
single point of failure



augmented redundancy

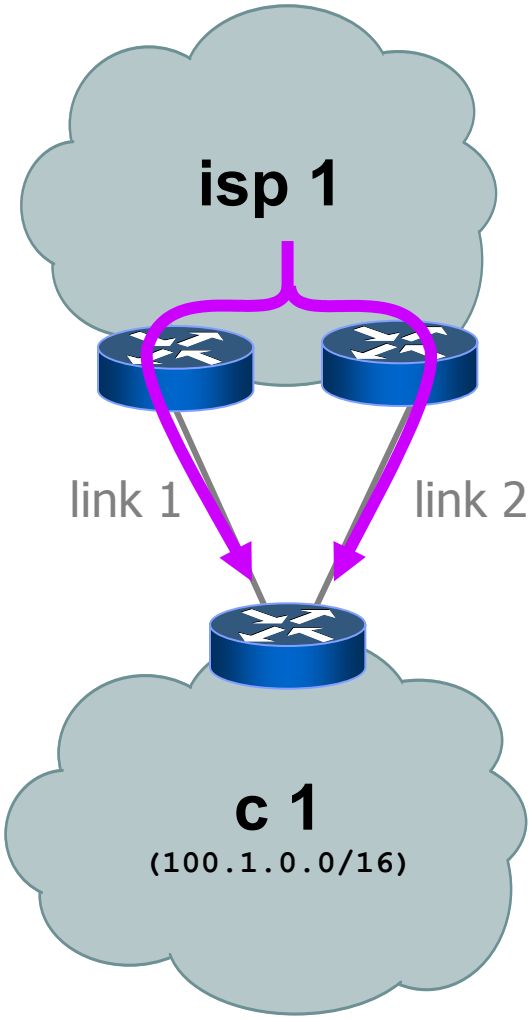


degrees of freedom



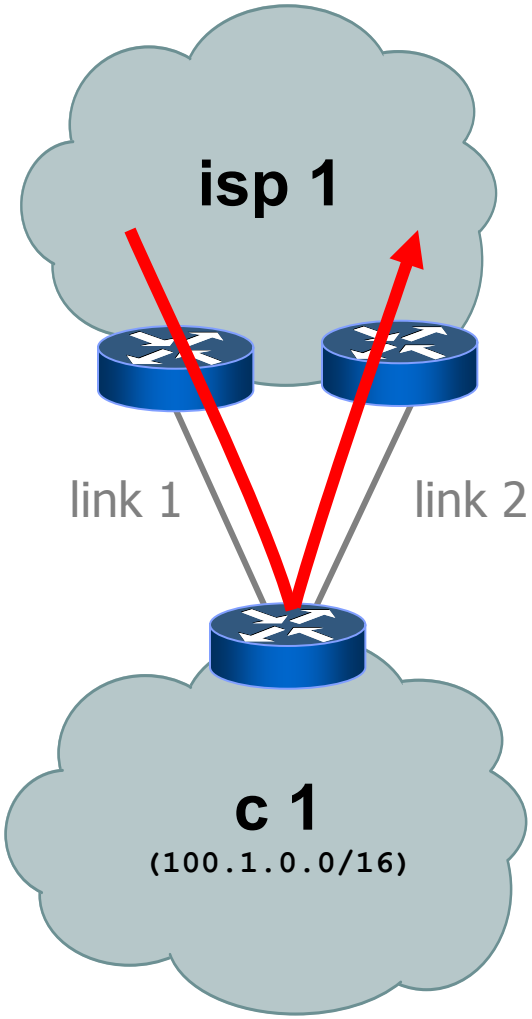
- an outbound packet may be sent through one of the two links in order to reach the internet

degrees of freedom



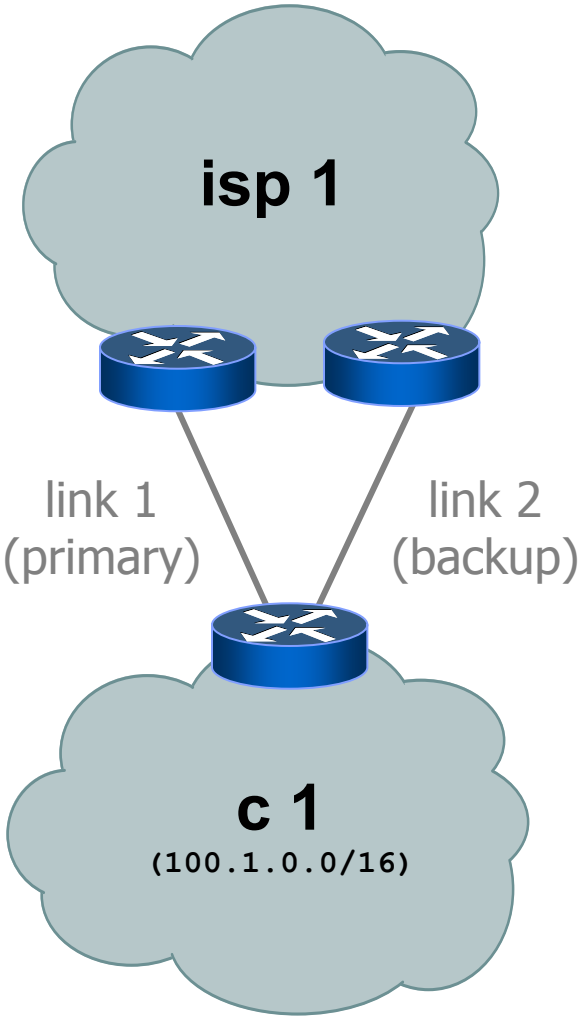
- an outbound packet may be sent through one of the two links in order to reach the internet
- an inbound packet may use any of the two links in order to reach the network

degrees of freedom



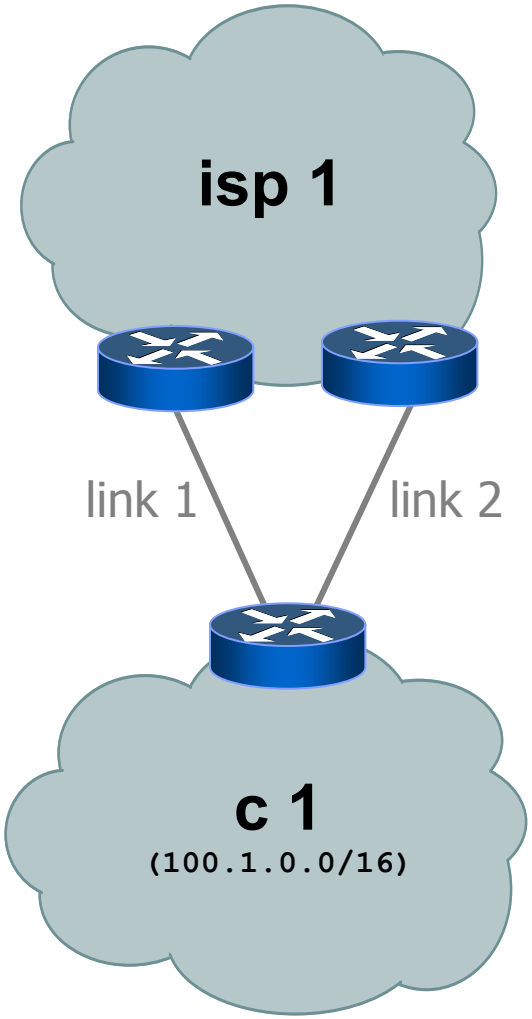
- an outbound packet may be sent through one of the two links in order to reach the internet
- an inbound packet may use any of the two links in order to reach the network
- an internet packet may traverse link 1 and link 2 (or vice versa)

desired policy: backup



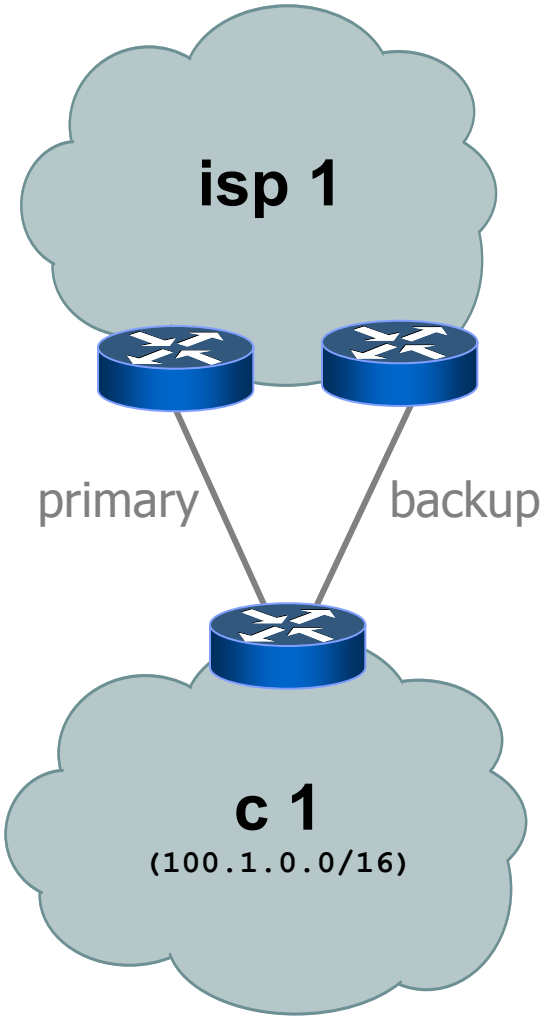
- rule out transit flows
- inbound traffic:
  - use link 1
  - use link 2 when link 1 is unavailable
- outbound traffic:
  - use link 1
  - use link 2 when link 1 is unavailable

alternatives to using bgp



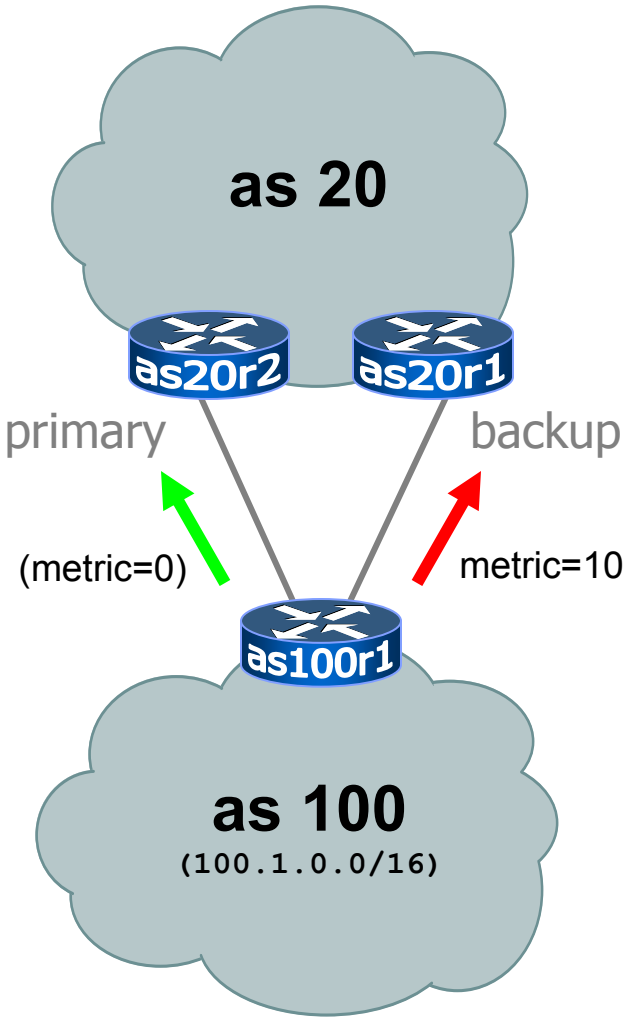
- using an igp (is-is, ospf, rip,…)
  - packets use link 1 or link 2 depending on the shortest path to customer c 1
  - there is no way to rule out transit packets when link 1 and link 2 are on the minimum path between a source and a destination
- using static routes
  - both the routers of the isp and the network have to be coherently configured by hand
  - there is no way to manage an automatic backup mechanism

using bgp



- announce /16 aggregate on each link
  - primary link makes standard announcements
  - backup link increases metric on outbound announcements, and reduces local-pref on inbound announcements
- when one link fails, the announcement of the /16 aggregate via the other link ensures continued connectivity

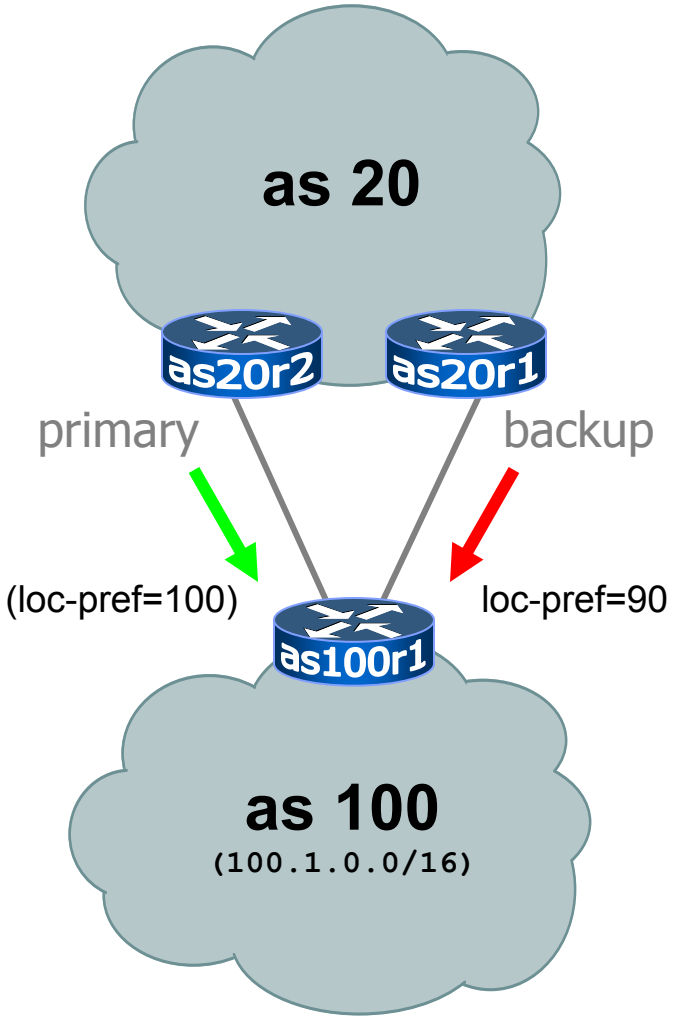
setting metric



- the value of the "multi-exit-discriminator" attribute is called "metric"
- upon receiving the same announcement with two different meds, the provider will (hopefully) adopt the one with the smaller one
- the metric is set on outgoing announcements and manages inbound traffic flows
- metrics are comparable only among announcements coming from the same neighboring as

default value: 0

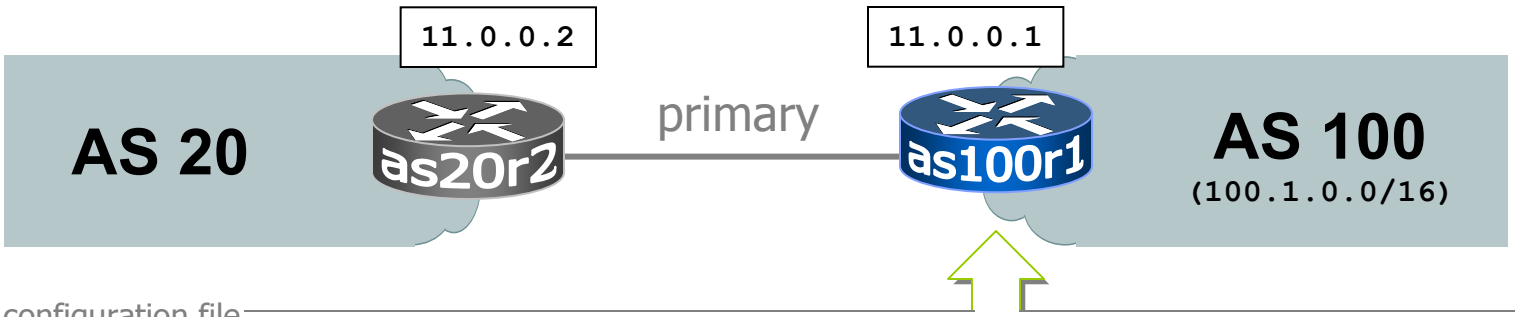
setting local-preference



- the customer assigns a lower local-preference to the announcement coming from the backup peer
- the local-preference attribute is checked before as-path length in the route selection process
- local-preference applies to incoming announcements and manages outbound traffic flows

default value: 100

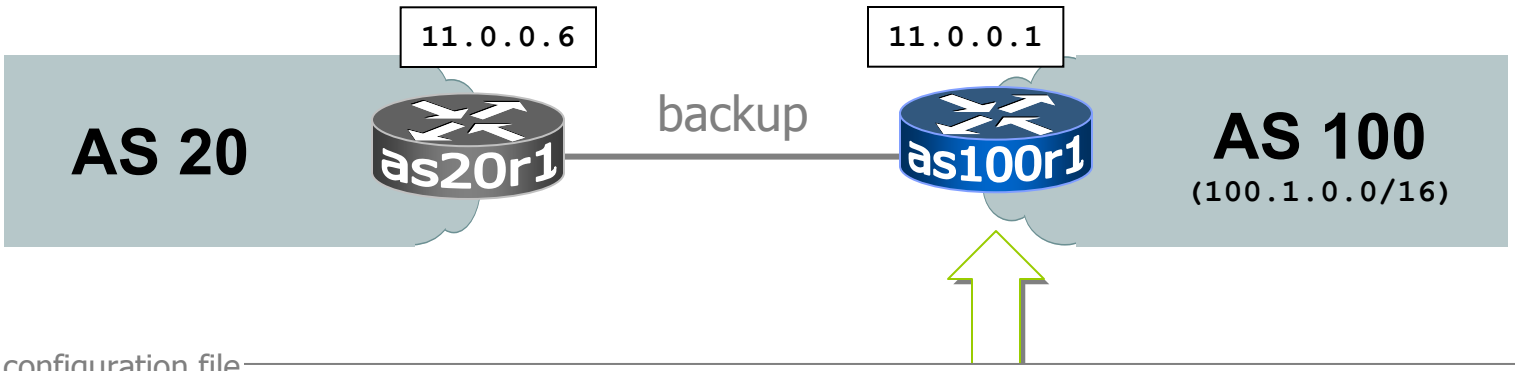
router as100r1 configuration



frr configuration file

```
router bgp 100
!
neighbor 11.0.0.2 remote-as 20
neighbor 11.0.0.2 description as20r2 (primary)
neighbor 11.0.0.6 remote-as 20
neighbor 11.0.0.6 description as20r1 (backup)
!
network 100.1.0.0/16
!
… next slide
```

router as100r1 configuration



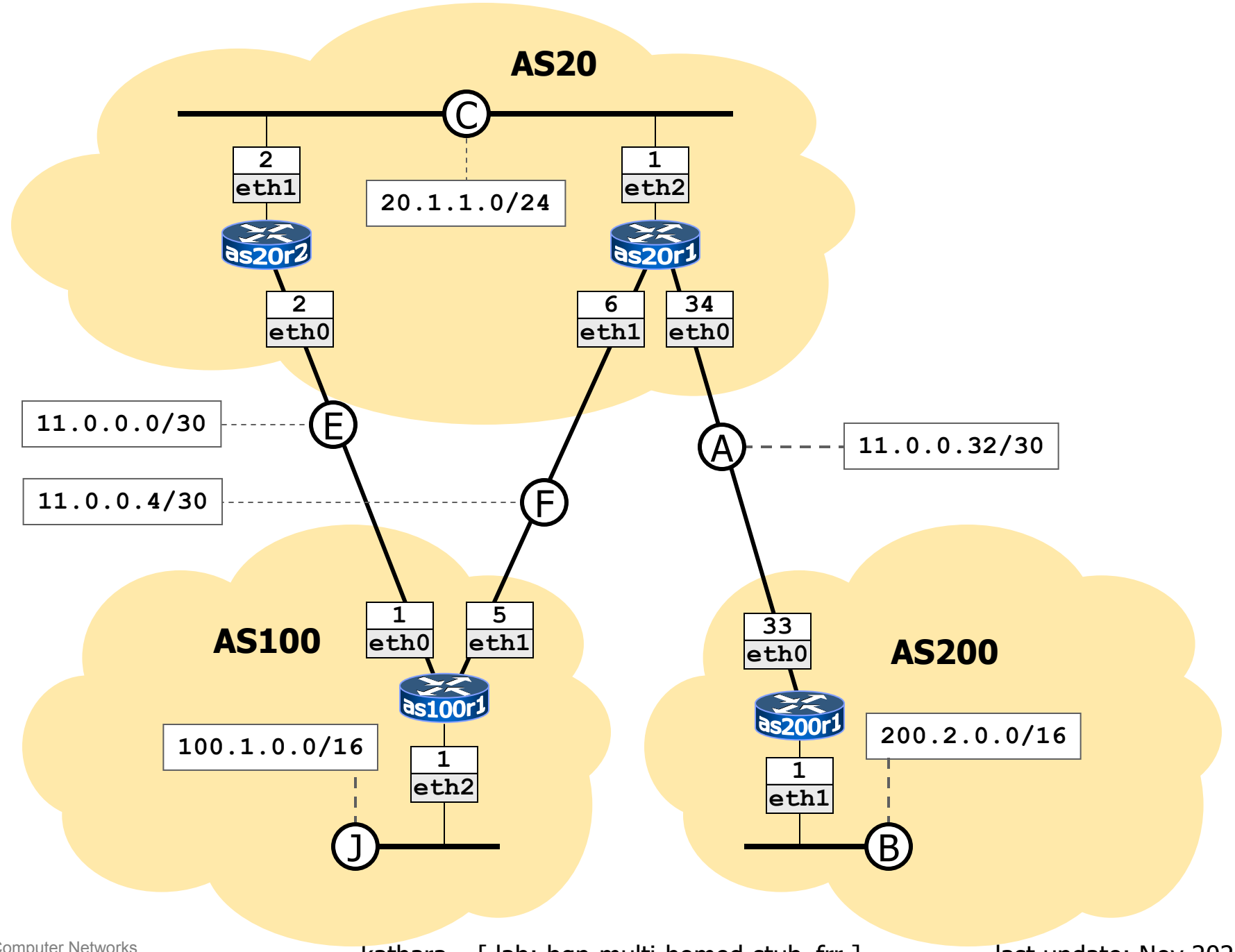
frr configuration file

```
!
neighbor 11.0.0.2 prefix-list mineOutOnly out
neighbor 11.0.0.2 prefix-list defaultIn in
neighbor 11.0.0.6 prefix-list mineOutOnly out
neighbor 11.0.0.6 route-map metricOut out
neighbor 11.0.0.6 prefix-list defaultIn in
neighbor 11.0.0.6 route-map localPrefIn in
!
… next slide
```

router as100r1 configuration

frr configuration file

```
access-list myAggregate permit 100.1.0.0/16
!
ip prefix-list mineOutOnly permit 100.1.0.0/16
ip prefix-list defaultIn permit 0.0.0.0/0
!
route-map metricOut permit 10
match ip address myAggregate
set metric 10
!
route-map localPrefIn permit 10
set local-preference 90
!
```



multi-homed stub

- start the lab

host machine

```
user@localhost:~$ cd kathara-lab_bgp-multi-homed-stub_frr
user@localhost:~/kathara-lab_bgp-multi-homed-stub_frr$ kathara lstart
```

- ping as100r1 from as200r1

as200r1

```
root@as200r1:~# ping 100.1.0.1
PING 100.1.0.1 (100.1.0.1) 56(84) bytes of data.
64 bytes from 100.1.0.1: icmp_seq=1 ttl=62 time=1.39 ms
64 bytes from 100.1.0.1: icmp_seq=2 ttl=62 time=1.88 ms

--- 100.1.0.1 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1022ms
rtt min/avg/max/mdev = 1.398/1.642/1.886/0.244 ms
```

multi-homed stub

- traceroute from as200r1 to as100r1

as200r1

```
root@as200r1:/# traceroute 100.1.0.1
traceroute to 100.1.0.1 (100.1.0.1), 30 hops max, 60 byte packets
 1  11.0.0.34 (11.0.0.34)  6.665 ms  6.222 ms  0.129 ms
 2  20.1.1.2 (20.1.1.2)  0.214 ms  0.179 ms  0.170 ms
 3  100.1.0.1 (100.1.0.1)  0.277 ms  0.208 ms  0.210 ms
root@as200r1:/#
```

- the packets traverse router as20r2 (20.1.1.2) to reach 100.1.0.1
  - the backup link is not used

multi-homed stub

- let us have a look at frr bgp table

as20r1

```
root@as20r1:~# vtysh
.....
as20r1-frr> show ip bgp
BGP table version is 6, local router ID is 20.1.1.1, vrf id 0
Default local pref 100, local AS 20
Status codes:  s suppressed, d damped, h history, * valid, > best, = multipath,
               i internal, r RIB-failure, S Stale, R Removed
Nexthop codes: @NNN nexthop's vrf id, < announce-nh-self
Origin codes:  i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
   0.0.0.0/0        0.0.0.0                  0         32768 i
*>i11.0.0.0/30      20.1.1.2                 0    100      0 i
*> 11.0.0.4/30      0.0.0.0                  0         32768 i
*> 11.0.0.32/30     0.0.0.0                  0         32768 i
* i20.1.1.0/24      20.1.1.2                 0    100      0 i
*>                  0.0.0.0                  0         32768 i
*  100.1.0.0/16     11.0.0.5                10             0 100 i
*>i                 11.0.0.1                 0    100      0 100 i
*> 200.2.0.0/16     11.0.0.33                0             0 200 i
Displayed  7 routes and 9 total paths
```

recursive lookup problem

- some implementations of bgp would not accept a route if the next-hop is not known via igp
- this below is an example from the analogous lab using quagga

as20r1

```
as20r1:~# telnet localhost bgpd
.....
bgpd> show ip bgp
BGP table version is 0, local router ID is 20.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
.....
*>i11.0.0.0/30      20.1.1.2                 0    100      0 i
.....
* i100.1.0.0/16     11.0.0.1                 0    100      0 100 i
*>                  11.0.0.5                10             0 100 i
.....
Total number of prefixes 7
```

not accepted by quagga because next hop is only known via bgp and not via igp

multi-homed stub

- now shut down the primary connection on as100r1

as100r1

```
root@as100r1:~# vtysh

Hello, this is FRRouting (version 7.5.1).
Copyright 1996-2005 Kunihiro Ishiguro, et al.

as100r1-frr# configure terminal
as100r1-frr(config)# router bgp 100
as100r1-frr(config-router)# neighbor 11.0.0.2 shutdown
as100r1-frr(config-router)# quit
as100r1-frr(config)# quit
as100r1-frr# show ip bgp summary

IPv4 Unicast Summary:
BGP router identifier 100.1.0.1, local AS number 100 vrf-id 0
BGP table version 3
RIB entries 2, using 384 bytes of memory
Peers 2, using 43 KiB of memory

Neighbor  V  AS  MsgRcvd  MsgSent TblVer InQ OutQ  Up/Down State/PfxRcd  PfxSnt
11.0.0.2  4  20       43       45      0   0    0 00:00:20 Idle (Admin)       0
11.0.0.6  4  20       44       44      0   0    0 00:39:33             1       1

Total number of neighbors 2
as100r1-frr#
```

multi-homed stub

- check the backup

as100r1

```
root@as100r1:/# route
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
default         11.0.0.6        0.0.0.0         UG    20     0        0 eth1
11.0.0.0        0.0.0.0         255.255.255.252 U     0      0        0 eth0
11.0.0.4        0.0.0.0         255.255.255.252 U     0      0        0 eth1
100.1.0.0       0.0.0.0         255.255.0.0     U     0      0        0 eth2
```

as100r1

```
root@as100r1:/# traceroute 200.2.0.1
traceroute to 200.2.0.1 (200.2.0.1), 30 hops max, 60 byte packets
 1  11.0.0.6 (11.0.0.6)  5.575 ms  5.158 ms  4.913 ms
 2  200.2.0.1 (200.2.0.1)  4.720 ms  4.448 ms  4.207 ms
```

multi-homed stub

- restart the primary connection and check that the primary link is back

as100r1

```
root@as100r1:/# vtysh

Hello, this is FRRouting (version 7.5.1).
Copyright 1996-2005 Kunihiro Ishiguro, et al.

as100r1-frr# configure terminal
as100r1-frr(config)# router bgp 100
as100r1-frr(config-router)# no neighbor 11.0.0.2 shutdown
as100r1-frr(config-router)# quit
as100r1-frr(config)# quit
as100r1-frr# quit
root@as100r1:/# traceroute 200.2.0.1
traceroute to 200.2.0.1 (200.2.0.1), 30 hops max, 60 byte packets
 1  11.0.0.2 (11.0.0.2)  4.466 ms  4.230 ms  4.112 ms
 2  20.1.1.1 (20.1.1.1)  4.014 ms  3.867 ms  3.753 ms
 3  200.2.0.1 (200.2.0.1)  3.641 ms  0.234 ms  0.120 ms
root@as100r1:/#
```