
BGP hacks
Traffic engineering with BGP
Fernando García Fernández
IP Architect
bics
a belgacom company

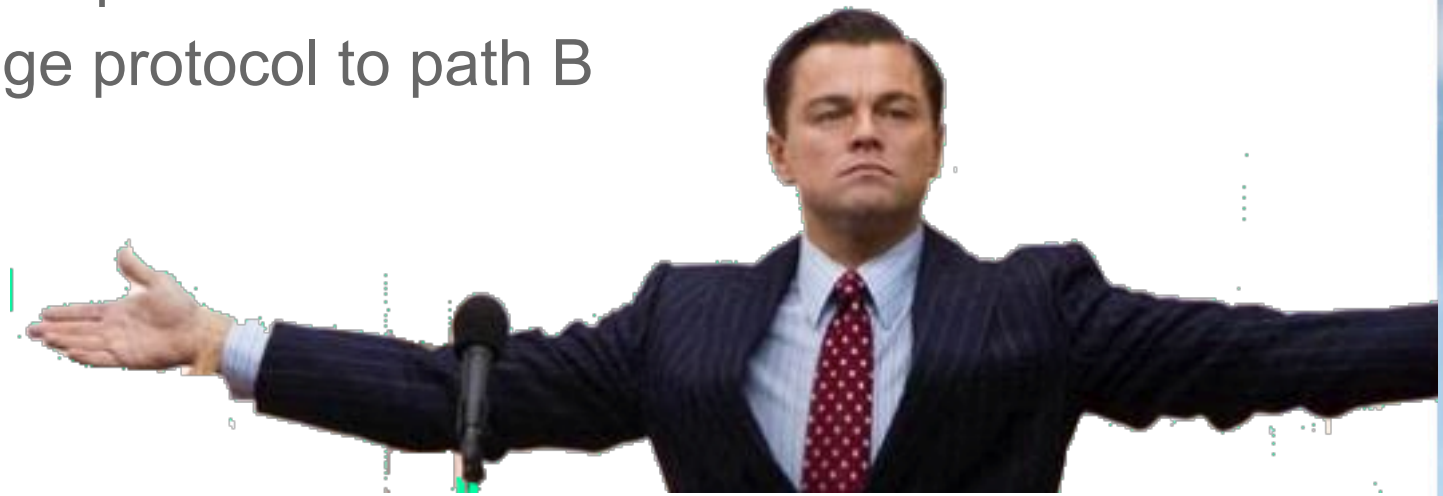
bics
a belgacom company

Theory

- BGP is vector-distance protocol
  - Select best path based on defined rules
- AS path length should be the decision point
- Other elements tie-breaks
- Operation should be:
  - if BGP select mostly path A, upgrade path A

Practice

- BGP is a political protocol
- Operation is:
  - (management) if BGP select path A and path A is more expensive…
  - change protocol to path B



bics
a belgacom company

Typical BGP hacks

- AS path prepend

```
 *>i 1.187.32.0/20    164.128.32.11 (65000) 174
9583 55644 55644 55644 55644 55644 55644 55644
55644 55644 55644 55644 55644 55644 55644 55644
55644 55644 55644 55644 55644 55644 55644 55644
55644 55644 55644 45271 i
```

- Prefix deagregation

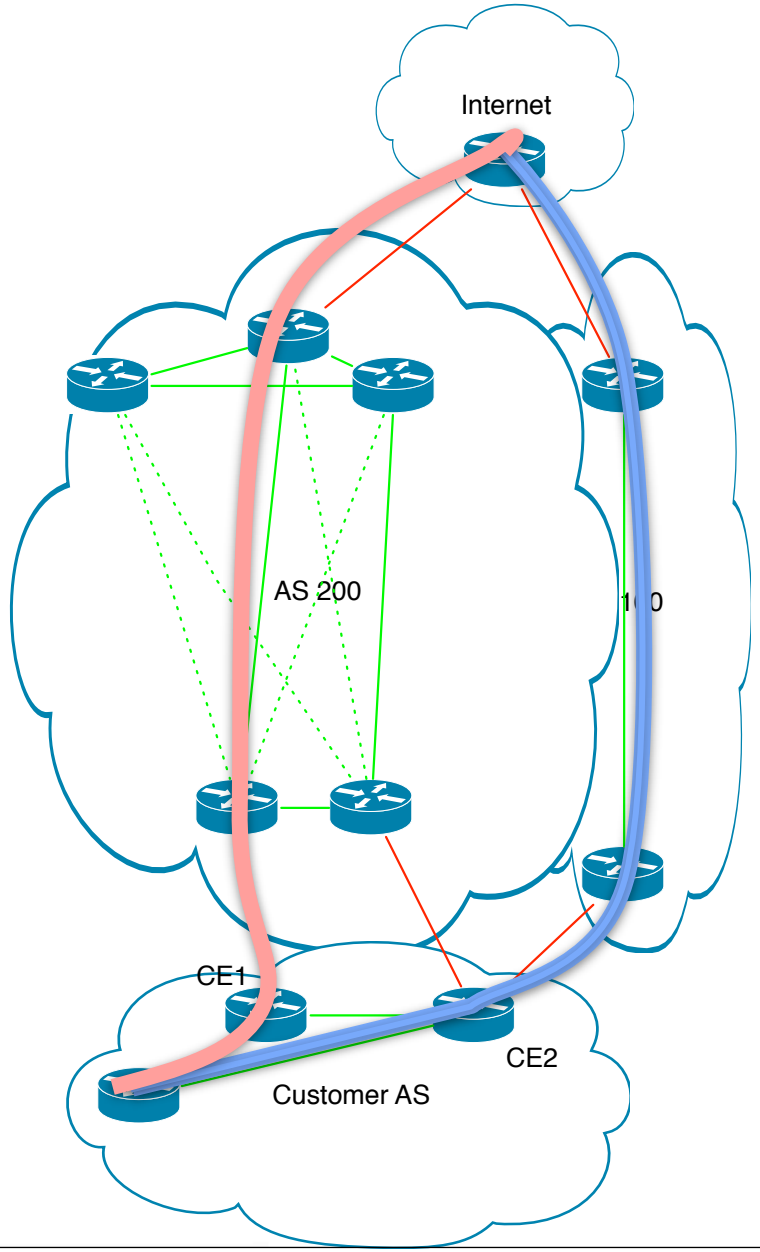
```
*>i 164.128.36.32/32
*>i 164.128.36.34/32
*>i 164.128.36.36/32
*>i 164.128.36.37/32
*>i 164.128.36.48/32
*>i 164.128.36.50/32
*>i 164.128.36.54/32
```

bics
a belgacom company

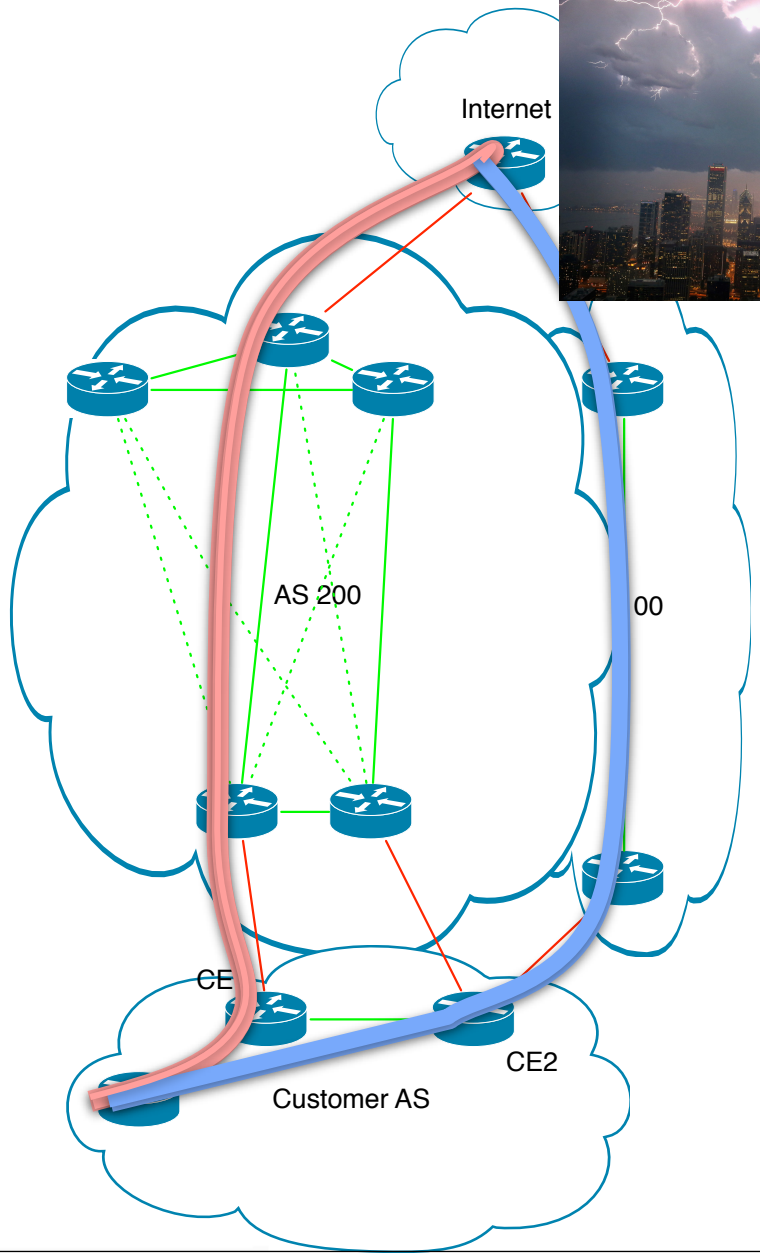
PROBLEMS…

bics
a belgacom company

AS Prepend: ON/OFF



Deaggregation: ON/OFF



bics
a belgacom company

Let’s hack!!!

bics
a belgacom company

Tools

- BGP selection path decision flow is large
  - Gives space to imagination…

Interesting
Paths for which the NEXT_HOP is inaccessible
Higher Local Preference
Locally originated via a network or aggregate
Shorter AS Path
Origin IGP<EGP<incomplete
Lower MED
Prefer exit eBGP over iBGP
Best IGP metric to next hop

This is previous!!!

Random path generator

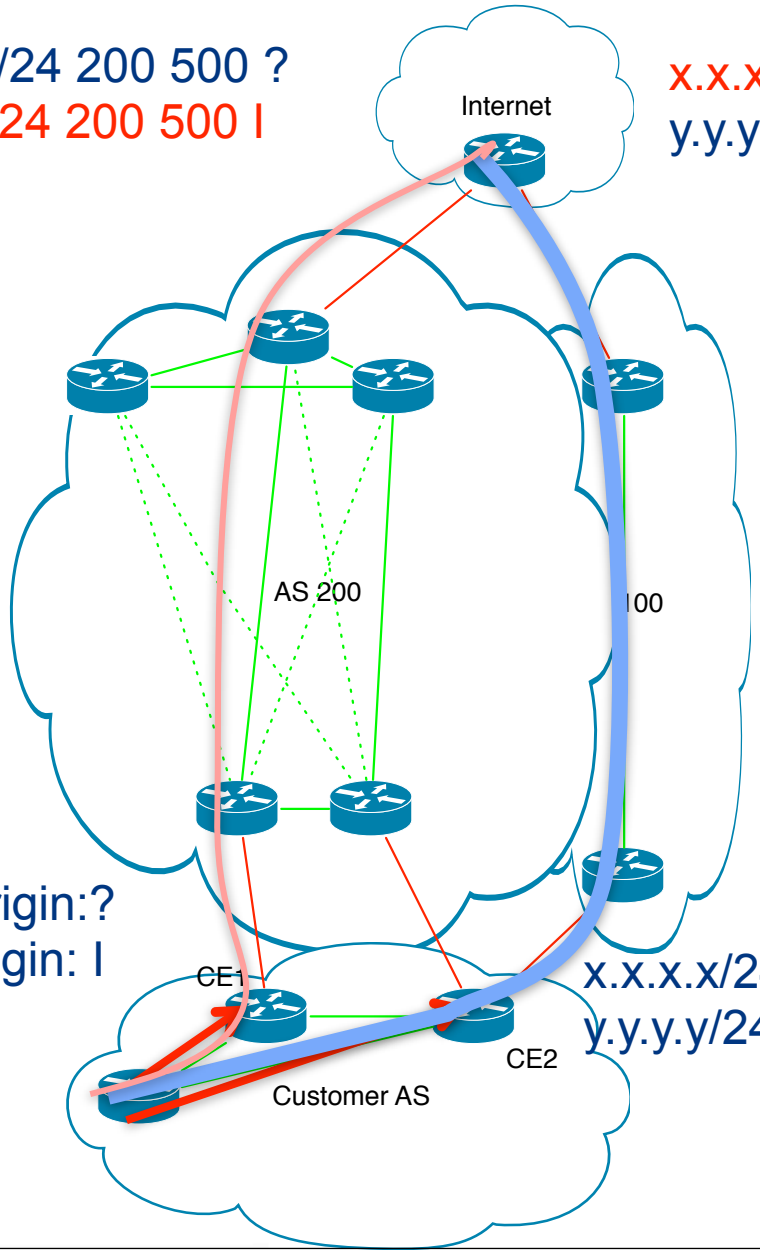
Pseudo random
Older route
Lower router-id
Lower cluster list
Lower neighbor IP

bics
a belgacom company

AS Prepend: ON/OFF

x.x.x.x/24 200 500 ?
y.y.y.y/24 200 500 I

x.x.x.x/24 200 500 I
y.y.y.y/24 200 500 ?



x.x.x.x/24 Origin:?
y.y.y.y/24 Origin: I

x.x.x.x/24 Origin:I
y.y.y.y/24 Origin: ?

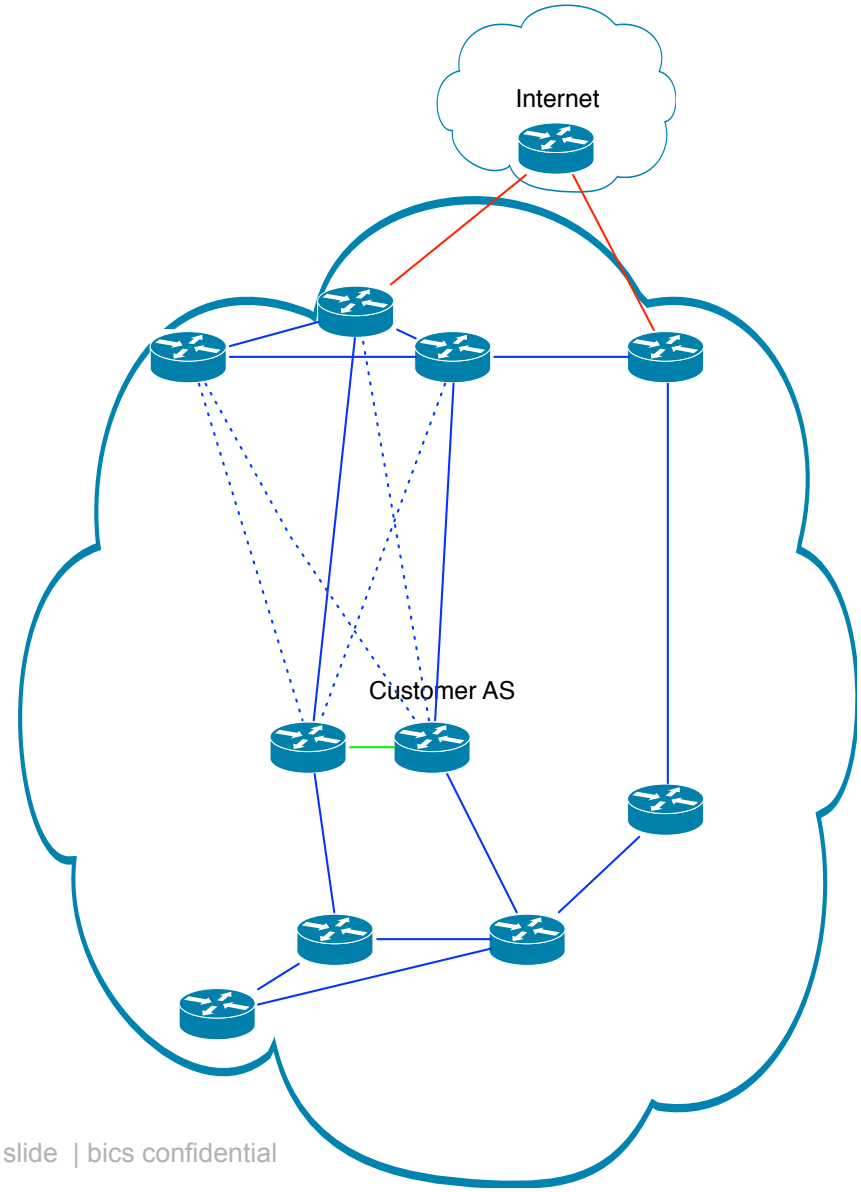
bics
a belgacom company

Disclaimer: ideas for your brain. To be processed

CENTRALIZED HACKING

bics
a belgacom company

Large network not easy to hack



bics
a belgacom company

Centralized hacking

One Ring to rule them all, One Ring to find them,

One Ring to bring them all and in the darkness bind them

bics
a belgacom company

Centralized hacking

One Router to rule them all, One Router to find them,

One Router to bring them all and in the darkness bind the

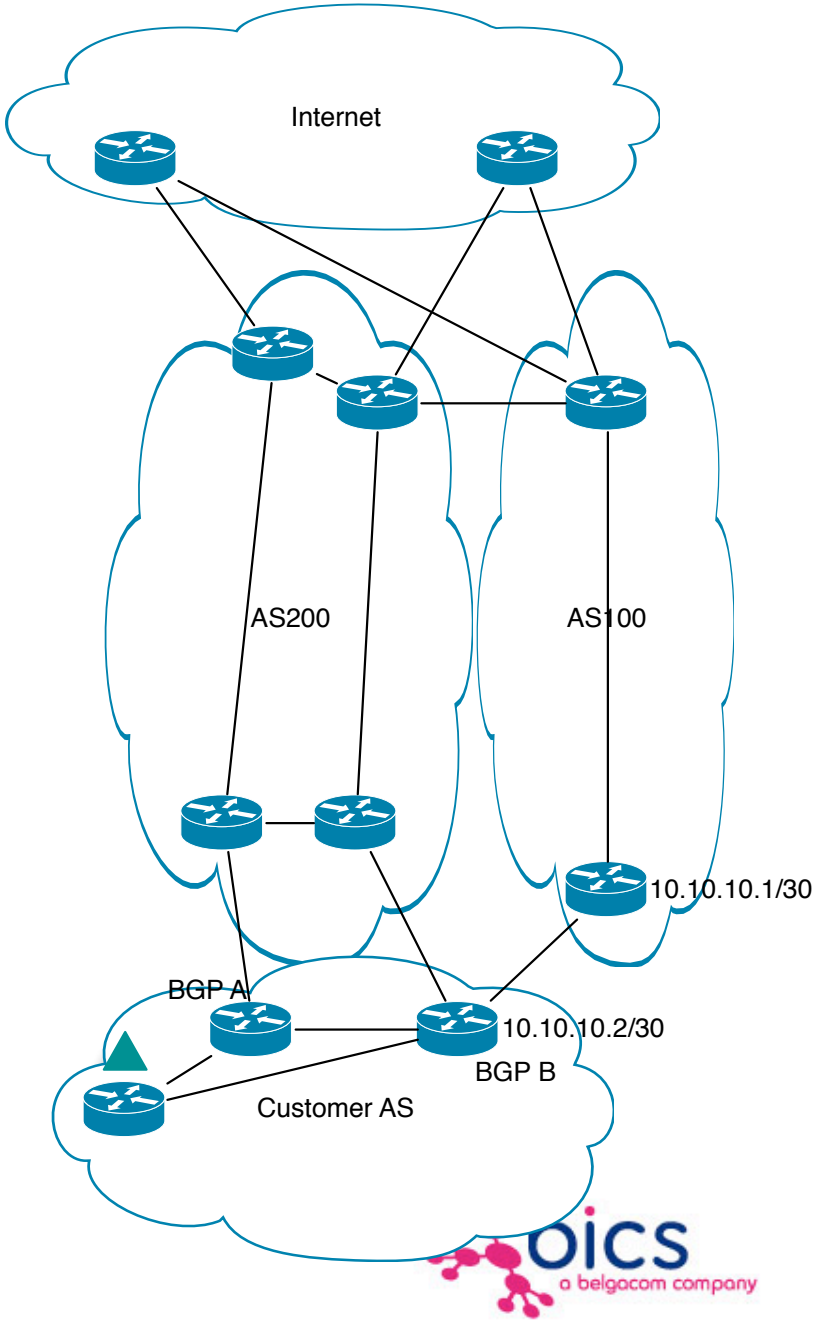
bics
a belgacom company

2 Scenarios

- Upstream the same for all 2
- Downstream based on customer config
  - 1 circuit with each router
  - Same router for several circuits

Tunnel upstream

- Prefix to be reached
  - Announced by Remote manager to border router
    - Next-hop IP of carrier P2P.
  - Announced by Remote manager to other border routers
    - Next-hop IP of selected border router
- Traffic will be send to selected router/carrier



Tunnel upstream

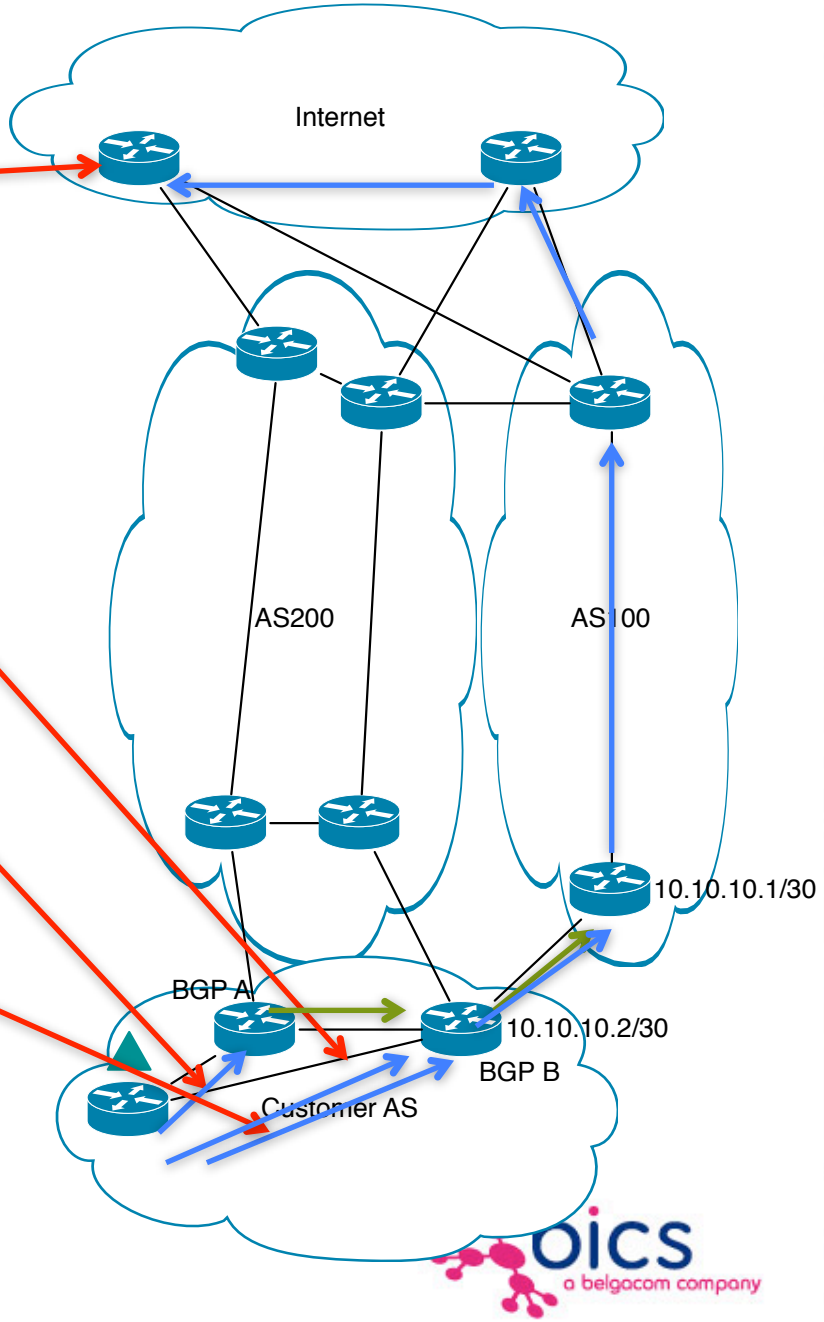
Prefix in Internet to be engineered

Anounces the Prefix to router BGP B only with next hop 10.10.10.1/30 NO ADVERTISE

Anounces the loopback to router BGP A only with next hop router BGP B NO ADVERTISE

Traffic toward B.B.B.B will always go through Carrier B

Should be a more specific of a general announce to cope with network failure



Downstream
1 circuit with each router

bics
a belgacom company

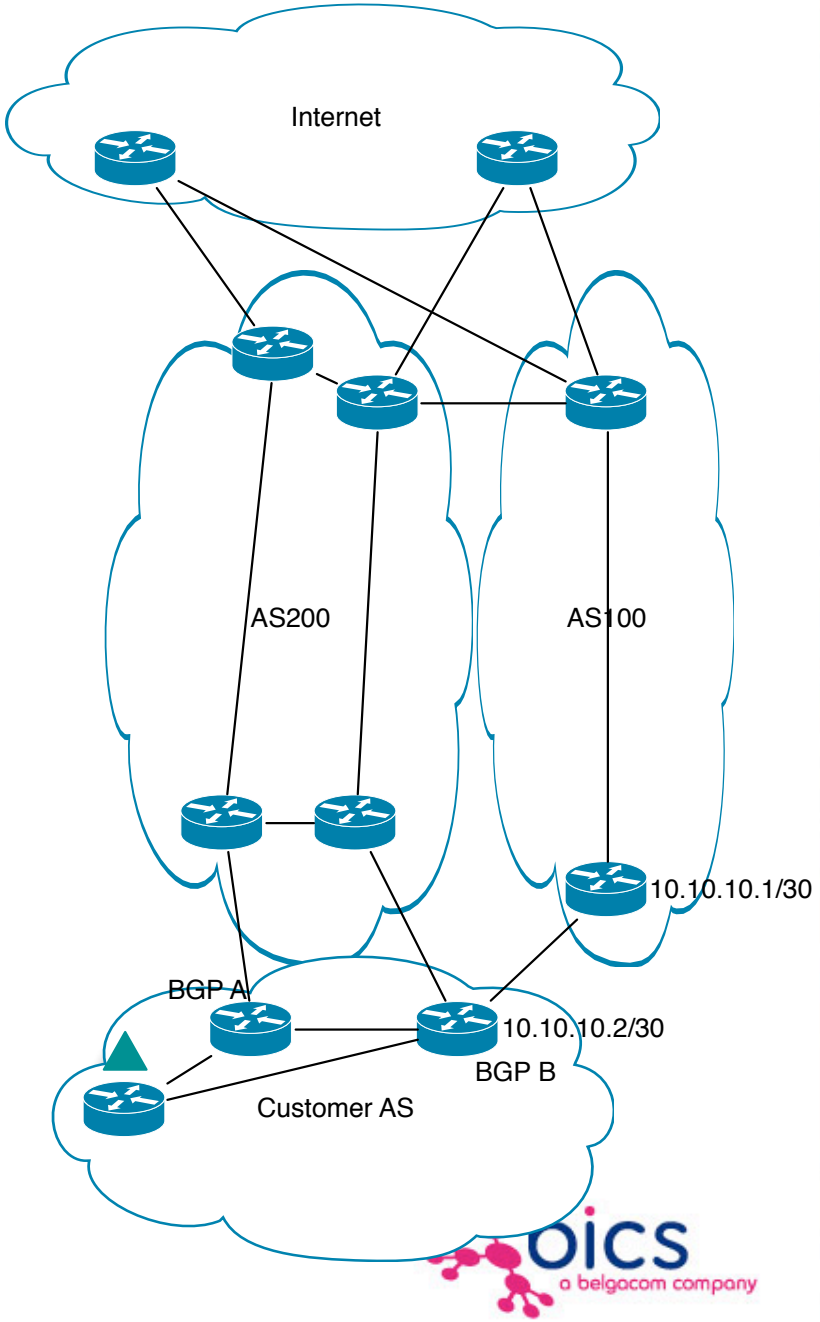
Requirements

- Deaggregated prefixes
  - Added to RIPE database
  - No lower of /24
- You'll get bad reputation

bics
a belgacom company

Tunnel downstream

- Border router prefered
  - Send the prefix for the announced deaggregated prefix
  - NO less than /24.
- Do not announce to other border routers
- Internet will receive the prefix only thru that carrier



Tunnel downstream

Prefix A.A.A.A/24
From the customer range

Anounces the prefix to router BGP B only

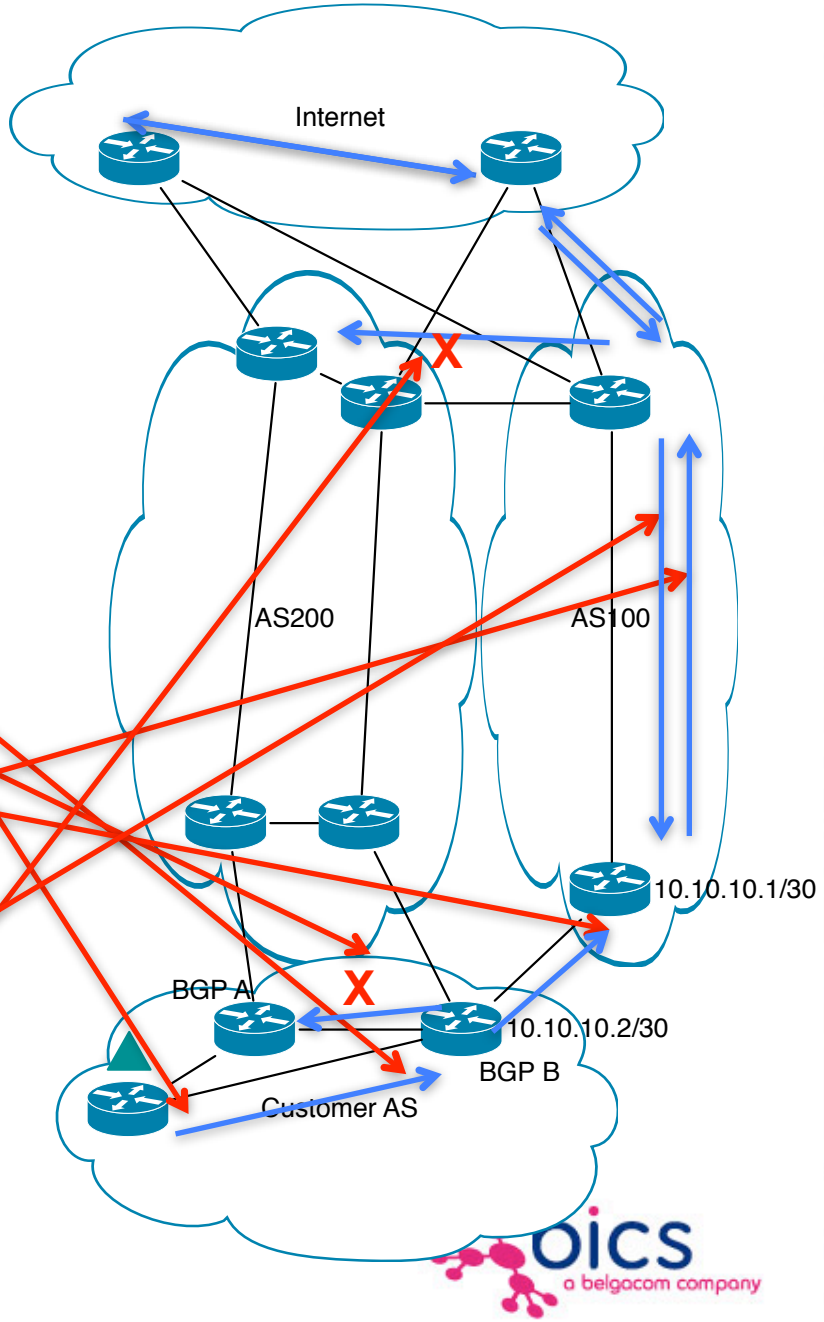
Router BGP B don't announce To router BGP A (it's iBGP)

Router BGP B announces to Carrier B

Internet receives the announce through Carrier B

To avoid rerouting of the traffic, a prepend of the AS of carrier A can be made in the announcement

Traffic to A.A.A.A will always go through Carrier B

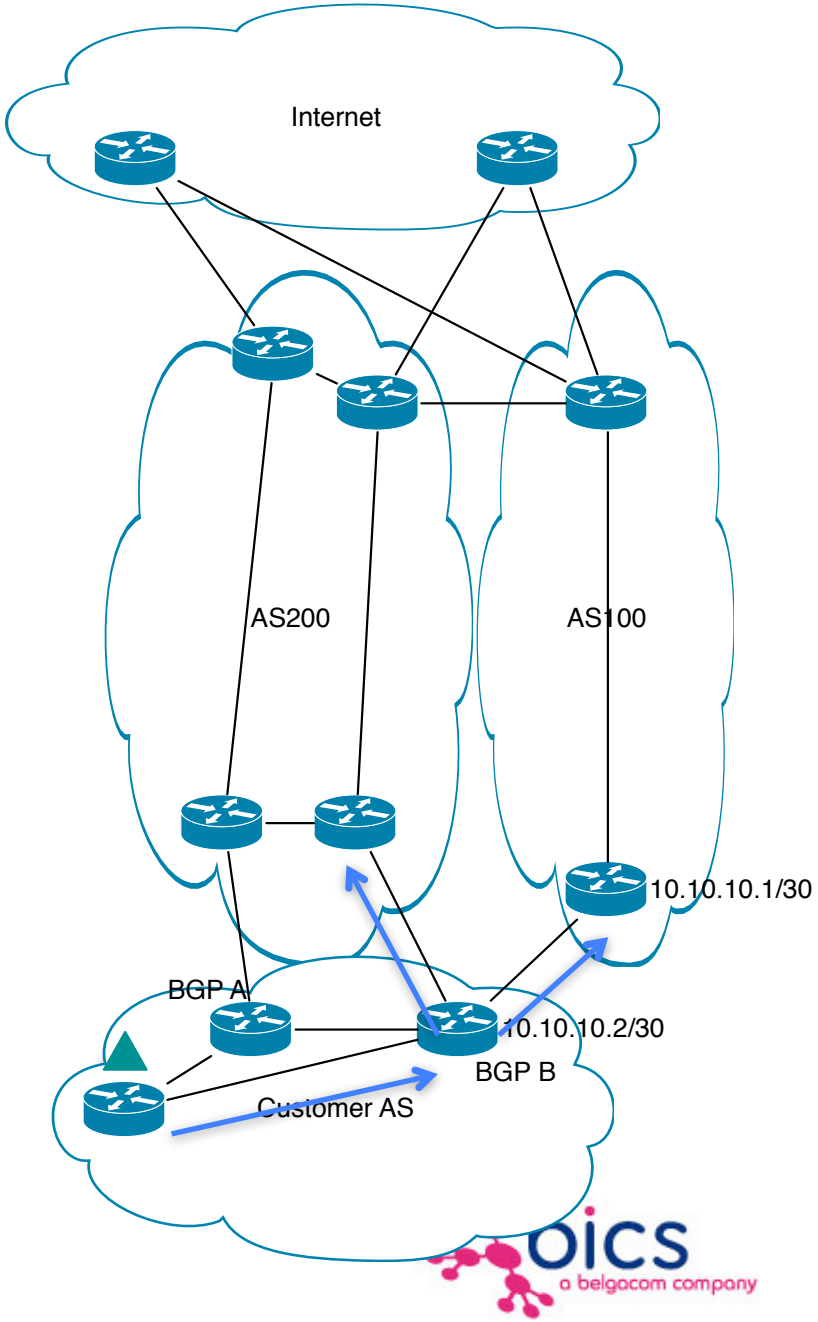


Downstream
Same router for several circuits

bics
a belgacom company

Problem

- We send the announce to the router
- Don’t discriminate between circuits
  - Same announce made to all the circuits



Solution

- Customer must configure a Community in the CE(s)
  - Community CUS:1
    - On circuit 1: Remove community and announce
    - On circuit 2: Drop announce
  - Community CUS:2
    - On circuit 1: Drop announce
    - On circuit 2: Remove community and announce
  - No community
    - Announce

Solution

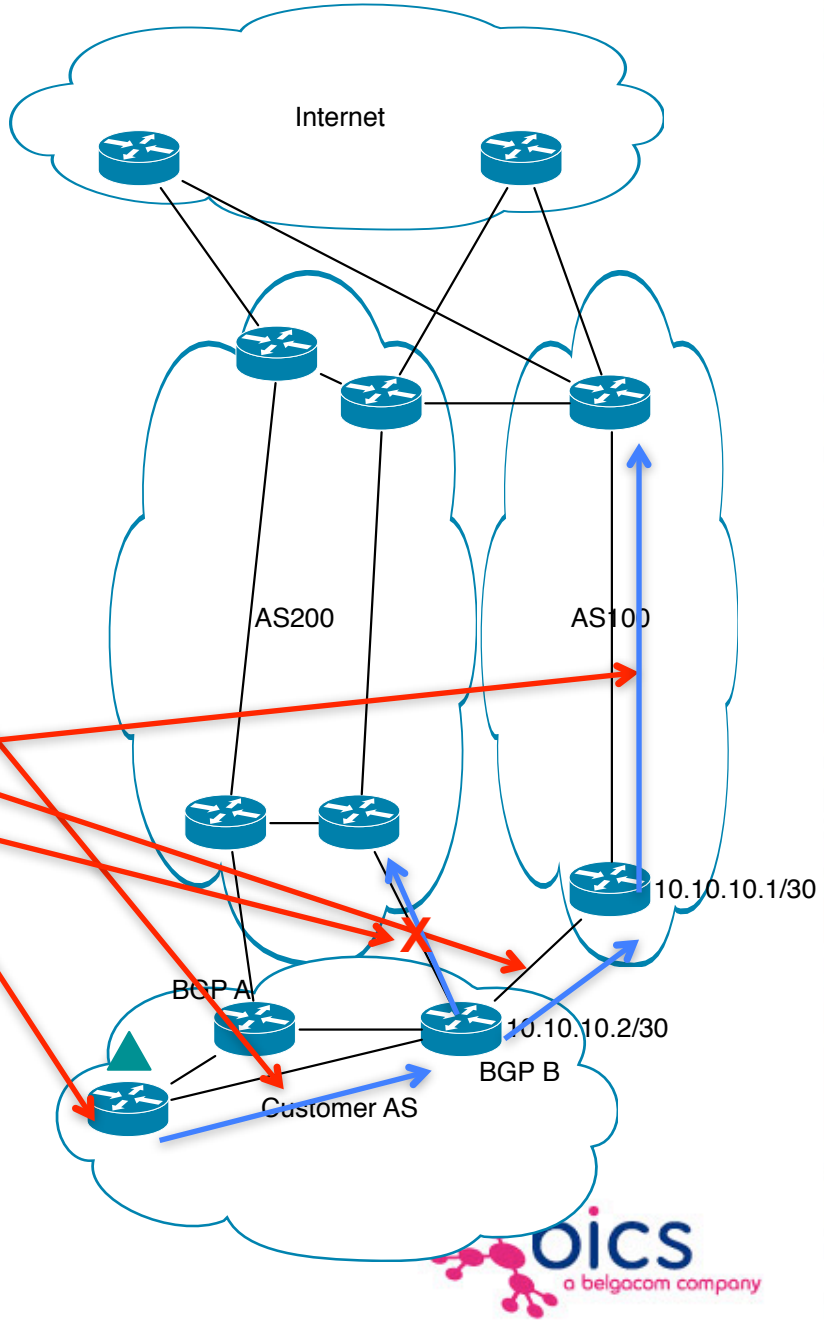
Deaggregated prefix
From the customer range

Anounces the prefix to
router BGP B community CUS:1

Router BGP B announces to
AS100

Router BGP B drops announce
to AS200

Internet receives the announce through
AS100



Advantages of centralized hacking

- Probability of fail = num routers ^2
- More secure management
- Inteligent routing
  - SDN based
  - House-made based:
    - Cisco IP SLA with ping to a loopback in Unix machine
    - Unix machine shut/no shut based on netflow, etc.
    - route-map announce based on IP SLA


bics
a belgacom company