# BGP hacks

## Trucos, ideas y consejos

Fernando García Fernández
IP Architect

bics
a belgacom company

# Preguntar cuando querais

bics
a belgacom company

# Tipos de director de TI

- "Selecciona los carrier que dan mas calidad y contrata el ancho de banda que necesites"
  - Animal mitológico
- "Hemos elegido estos dos operadores porque son los mas baratos. Balancea el tráfico para usarlos al máximo"
  - Gallinacea común

# Teoría

- BGP es un protocolo vector-distancia
  - Selecciona ruta en base a reglas definidas
- Longitud del path AS elemento decisorio
- Otros elementos, tie-break
- La operación debería ser:
  - Si BGP usa mas el carrier A, amplia el carrier A

# Práctica

- BGP es un protoclo político
- La operación es:
  - Si BGP el path A es el más elegido y A is más caro…
  - Retoca BGP para que use B

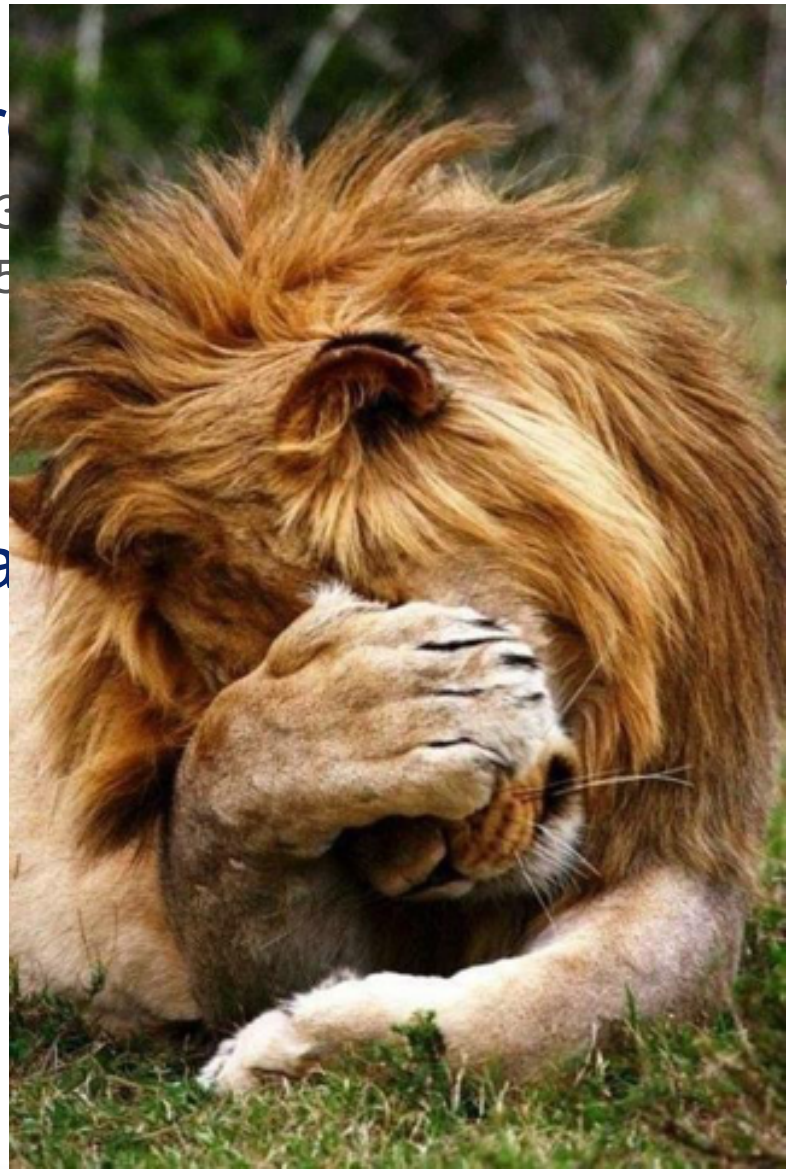
# Hacks tipicos de BGP

- AS path pr

```
 *>i 1.187.3                                    (65000) 174
9583 55644 5                                    5644 55644
55644 55644                                     55644 55644
55644 55644                                     55644 55644
55644 55644
```

- Desagrega

```
*>i 164.128.36
*>i 164.128.36
*>i 164.128.36
*>i 164.128.36
*>i 164.128.36
*>i 164.128.36
*>i 164.128.36
```





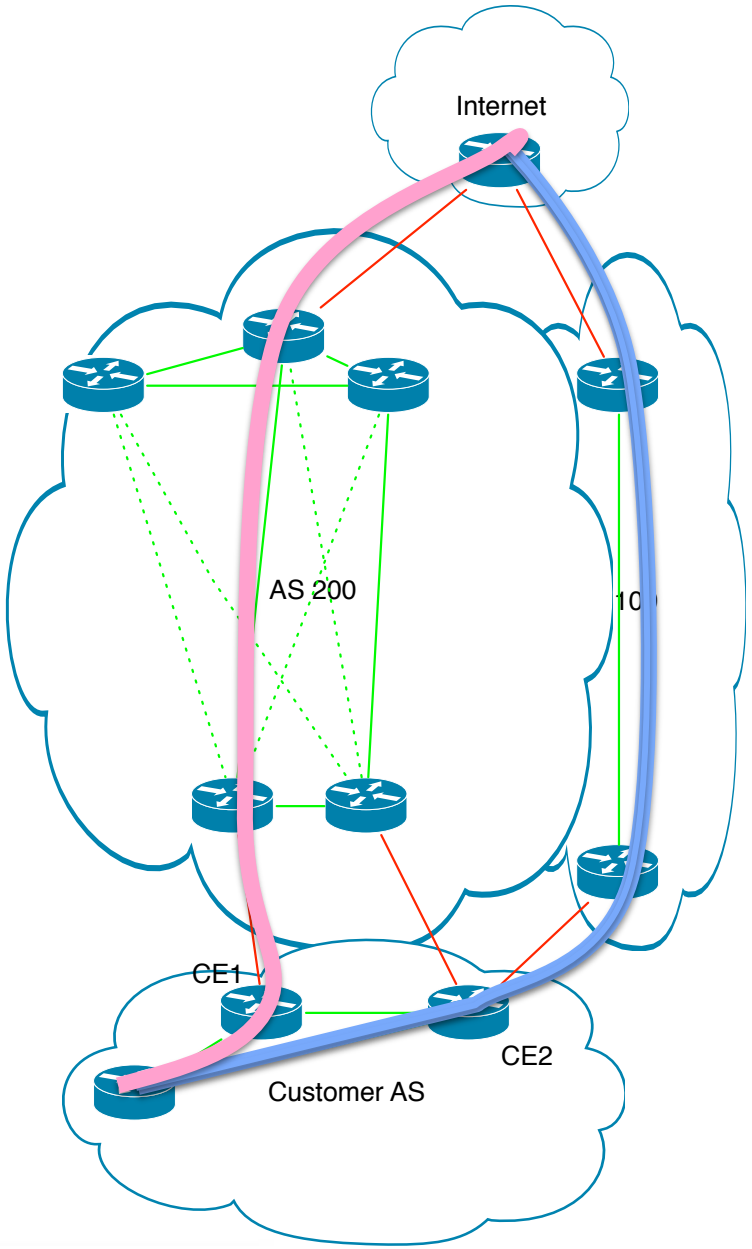
bics
a belgacom company

# PROBLEMAS…

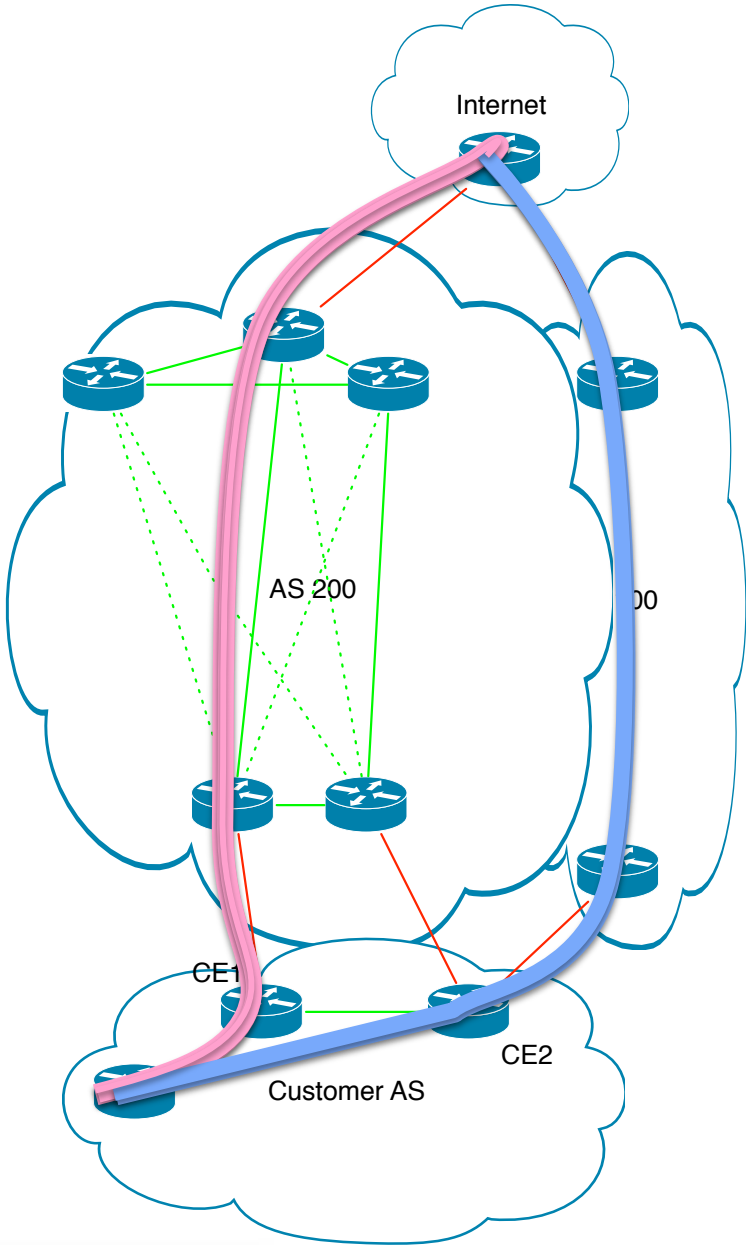


bics
a belgacom company

# AS Prepend: ON/OFF



# Deaggregation: ON/OFF



# Problema añadido

- Para la empresa nacional tipica
- Tiene usuarios
  - Reciben más de lo que envían
  - Tráfico de salida no es problema
  - Tráfico de entrada SI es problema


Let’s hack!!!

bics
a belgacom company

# Herramientas

- La tabla de selección BGP eslarga
  - Espacio a la imaginación…

**Interesantes**
Paths con NEXT_HOP inaccesible
Local Preference superior
Generada localmente con network o agregación
AS Path más corto
Origen IGP<EGP<incomplete
MED inferior
Prefiere salida eBGP sobre iBGP
Mejor métrica IGP al next-hop

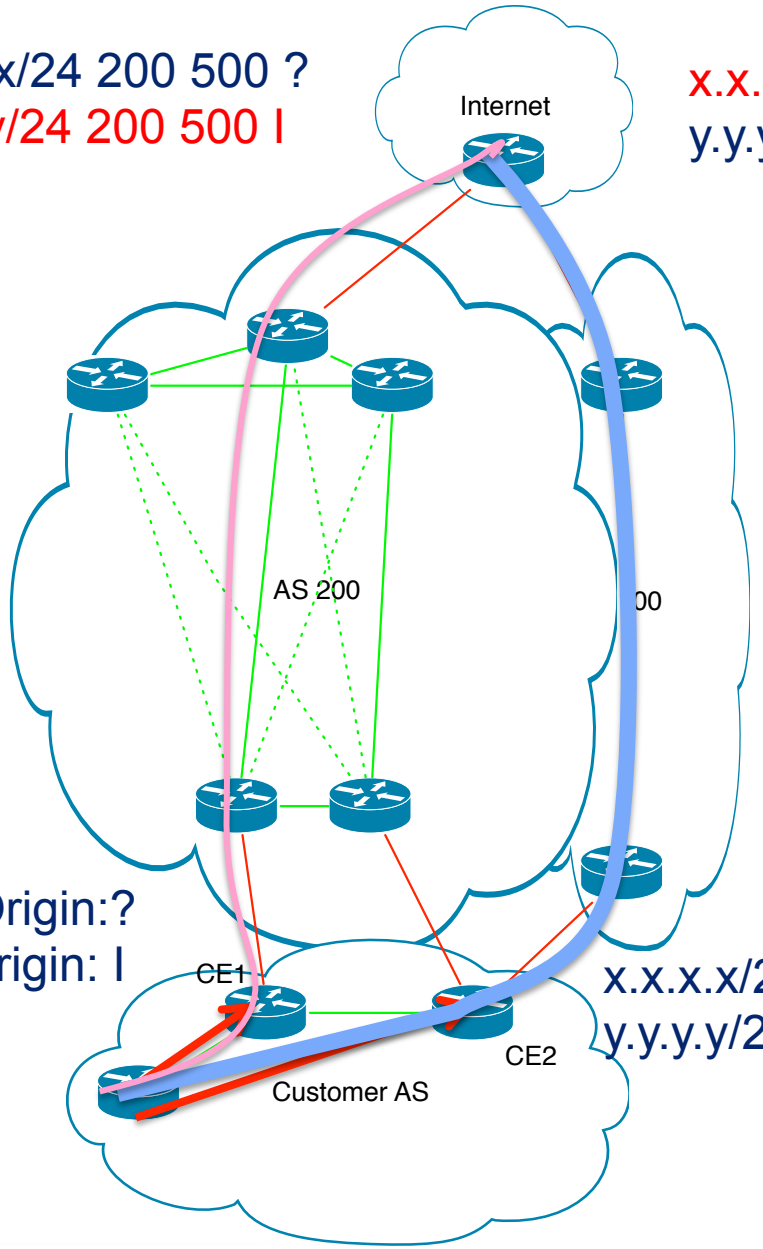
Esta es anterior!!!
Generador de path aleatorios

**Pseudo aleatorio**
Ruta más antigua
router-id inferior
cluster list inferior
neighbor IP inferior

# 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 ?

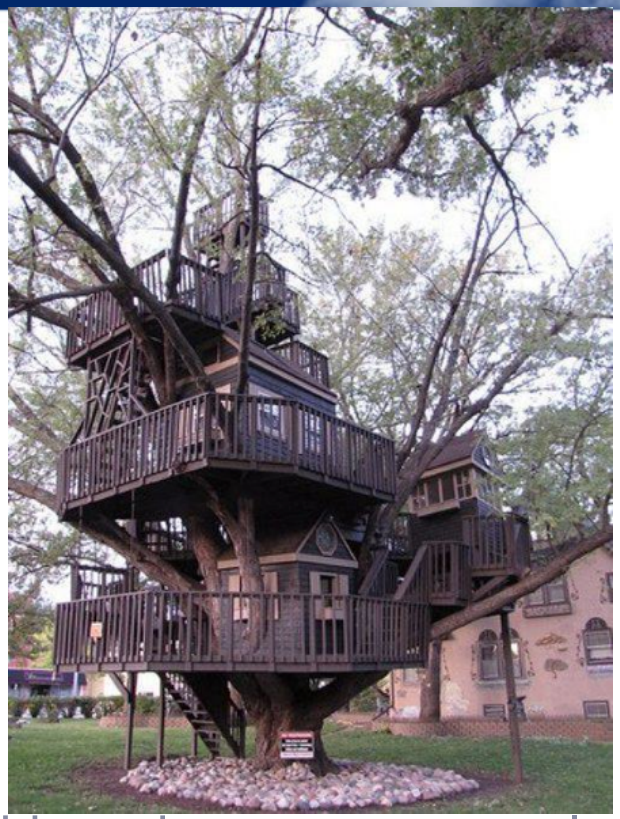



x.x.x.x/24 Origin:?
y.y.y.y/24 Origin: I

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



bics
a belgacom company



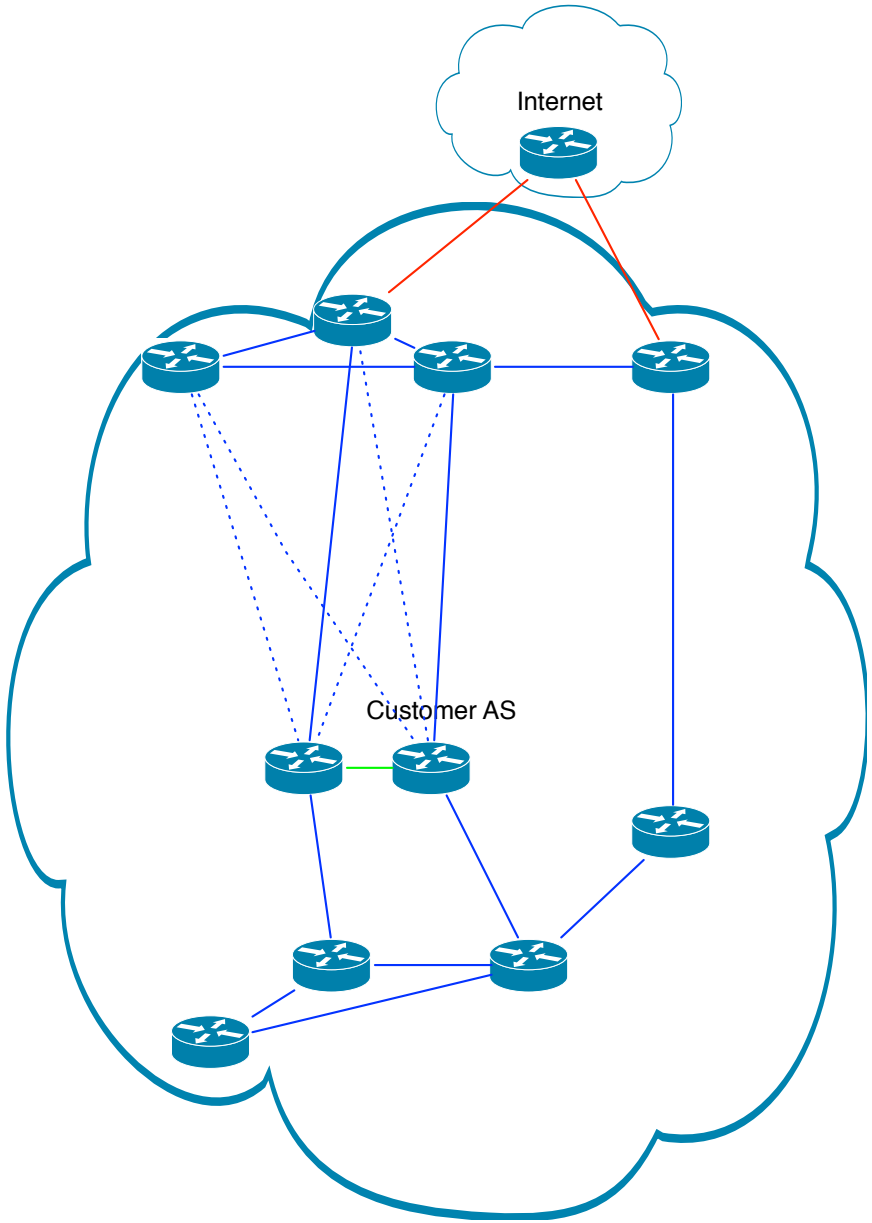
Disclaimer: sólo ideas, hay que procesarlas

# HACKING CENTRALIZADO



bics
a belgacom company

# Una red grande no es fácil de hackear



# Hacking centralizado

Un Anillo para gobernarlos a todos. Un Anillo para encontrarlos,

un Anillo para atraerlos a todos y atarlos en las tinieblas.

# Hacking centralizado

Un router para gobernarlos a todos. Un router para encontrarlos,

un router para atraerlos a todos y enviarlos por nuestro path favorito.

# Dos problemas

- Trafico de salida (upstream)
  - Fácil de gestionar
- Tráfico de entrada (downstream)
  - Más complicado

# Dos escenarios

- Upstream el mismo en ambos casos
- Downstream depende de la configuración
  - Un circuito en cada router
  - El mismo router con varios circuitos

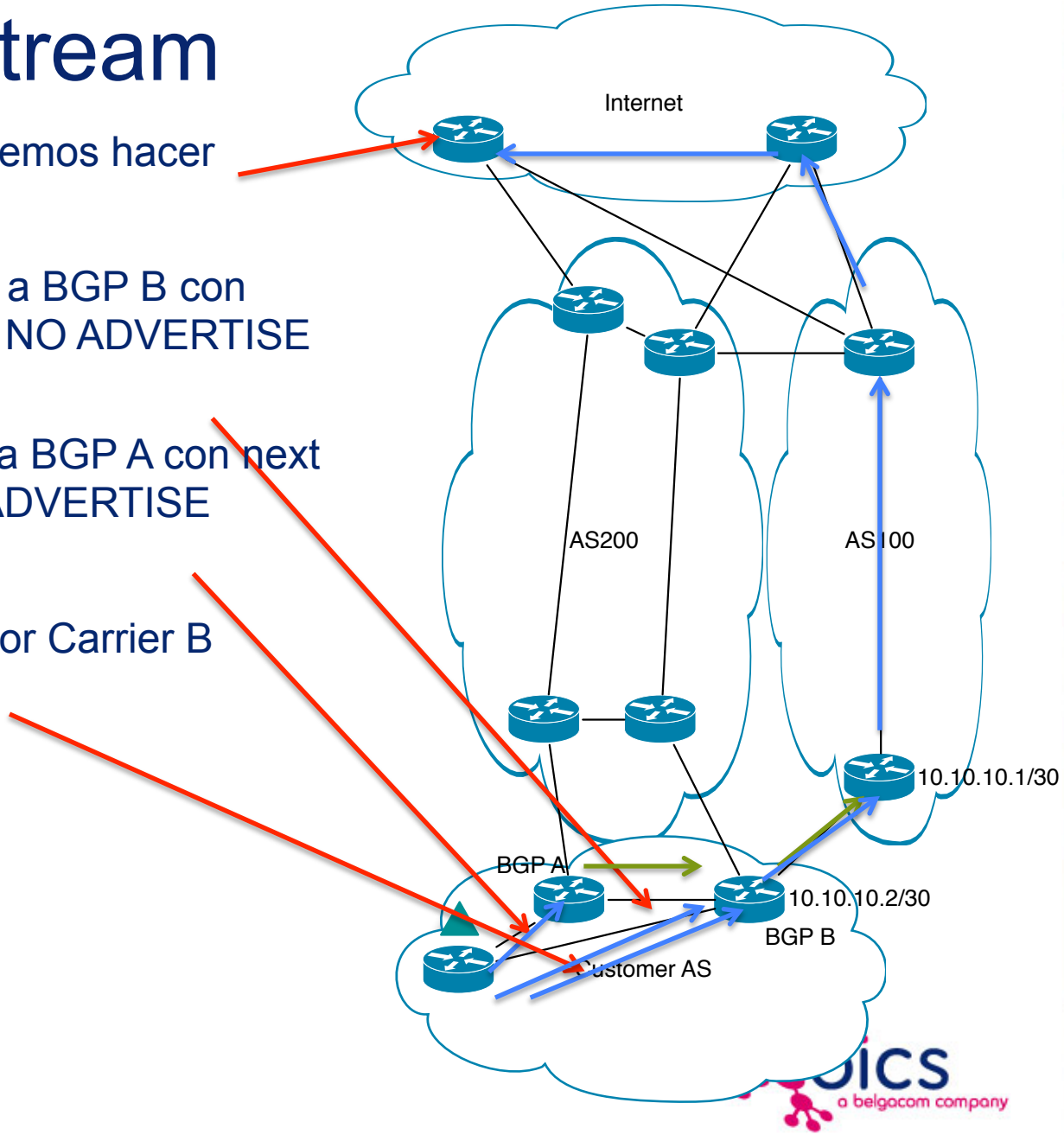
# Tráfico upstream

Prefijo para el que queremos hacer ingenieria

Anunciamos el  destino a BGP B con next-hop 10.10.10.1/30 NO ADVERTISE

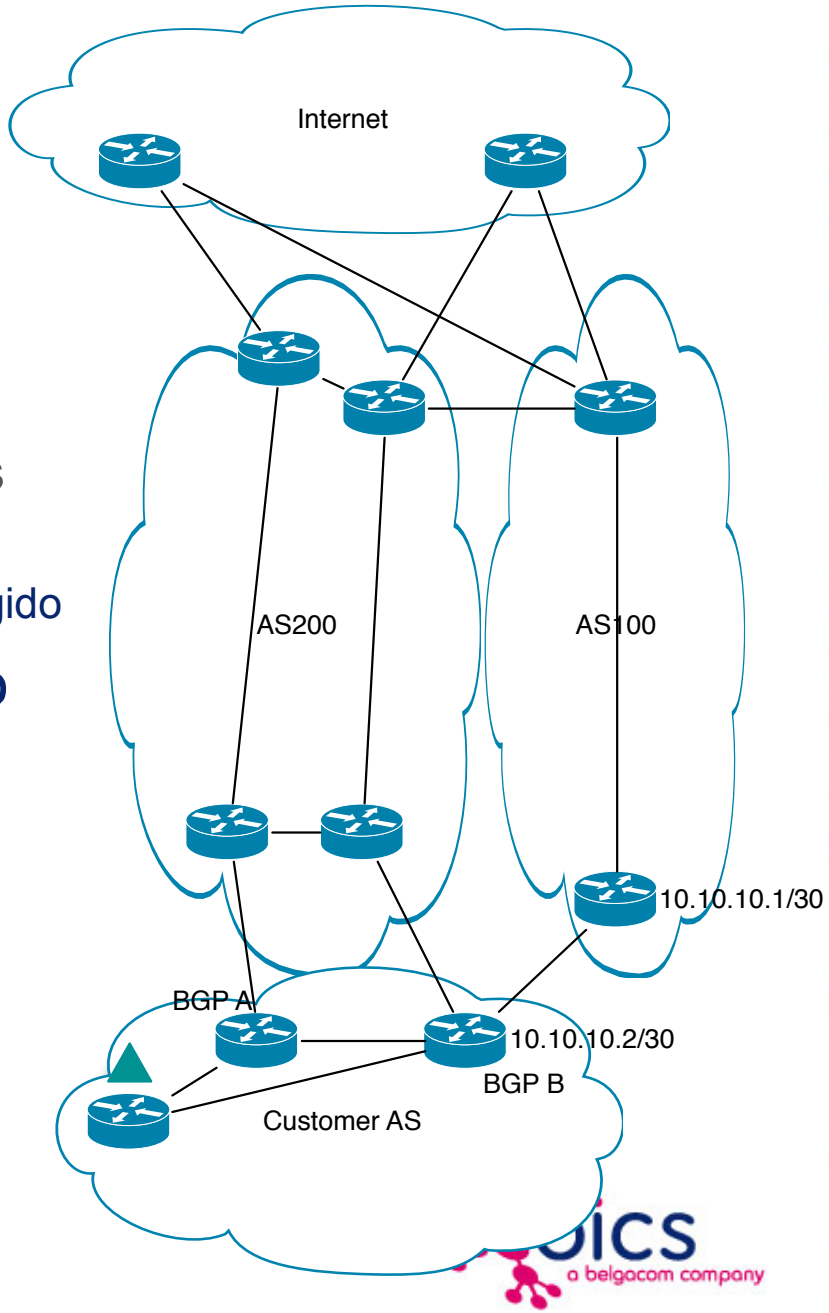
Anunciamos el destino a BGP A con next hop router BGP B NO ADVERTISE

El tráfico a B.B.B.B irá por Carrier B



# Tráfico upstream

- IP del destino
  - Anunciado por el ANILLO al border router
    - Next-hop IP del P2P del carrier.
  - Anunciado por el ANILLO a los otros border routers
    - Next-hop IP del border router elegido
- Tráfico ira por carrier elegido
- Medidas de seguridad en ANILLO
  - Por si se cae el path elegido
  - IP SLA
  - Rutas recibidas desde BGP B



# Downstream
# Un circuito por cada router

bics
a belgacom company

# Requisitos

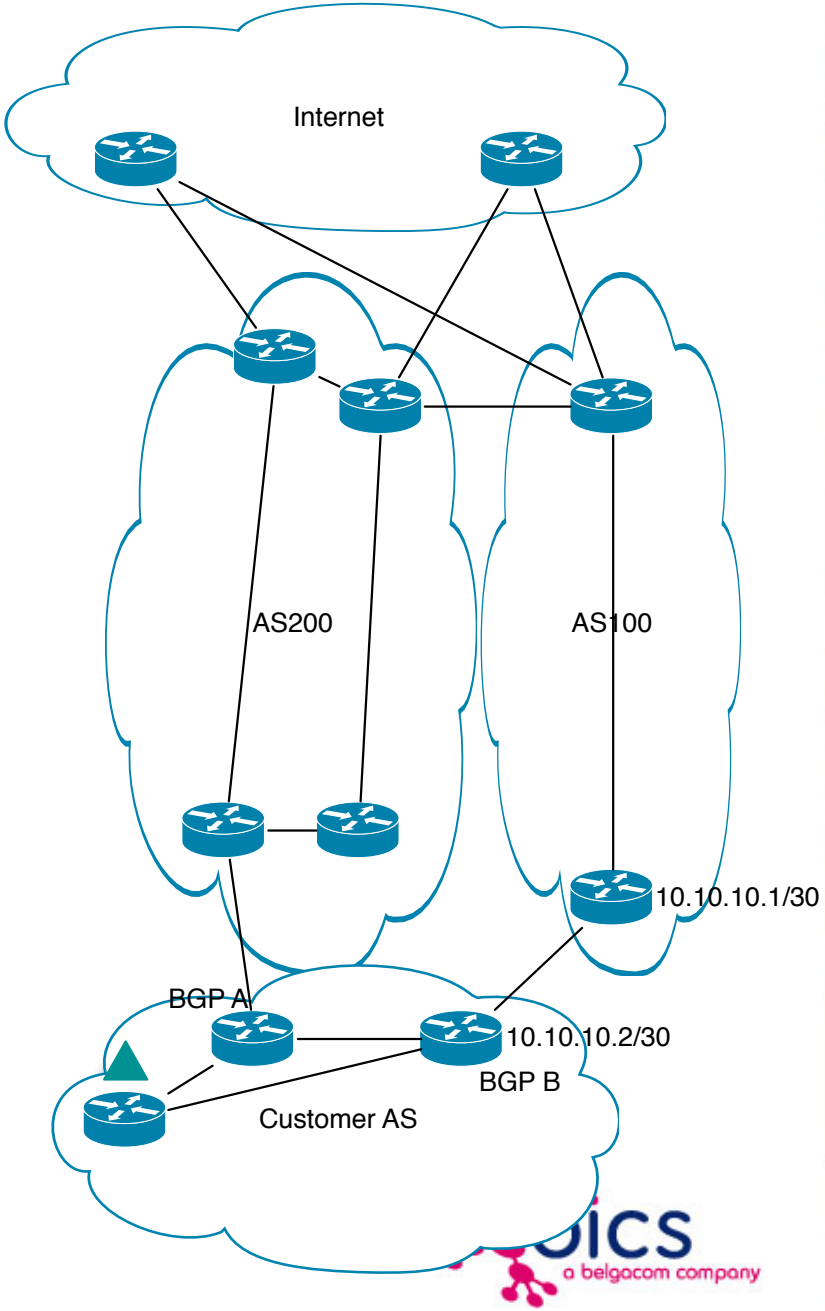
- Variante del método manual
- Prefijos desagregados
  - En la base de datos de RIPE
  - No inferiores a /24
- Te van a mirar mal



bics
a belgacom company

# Tráfico downstream

- Path preferido
  - ANILLO se lo anuncia al router elegido
  - NO inferior /24.
- No lo anuncia a los demás
  - O cambiamos el Origin (AS Path length más complicado)
- IP SLA/Track/Beacons
  - Cambiar el prefijo a otro path



# Tráfico downstream

Prefijo A.A.A.A/24
del cliente

Anuncia el prefijo sólo a
BGP B

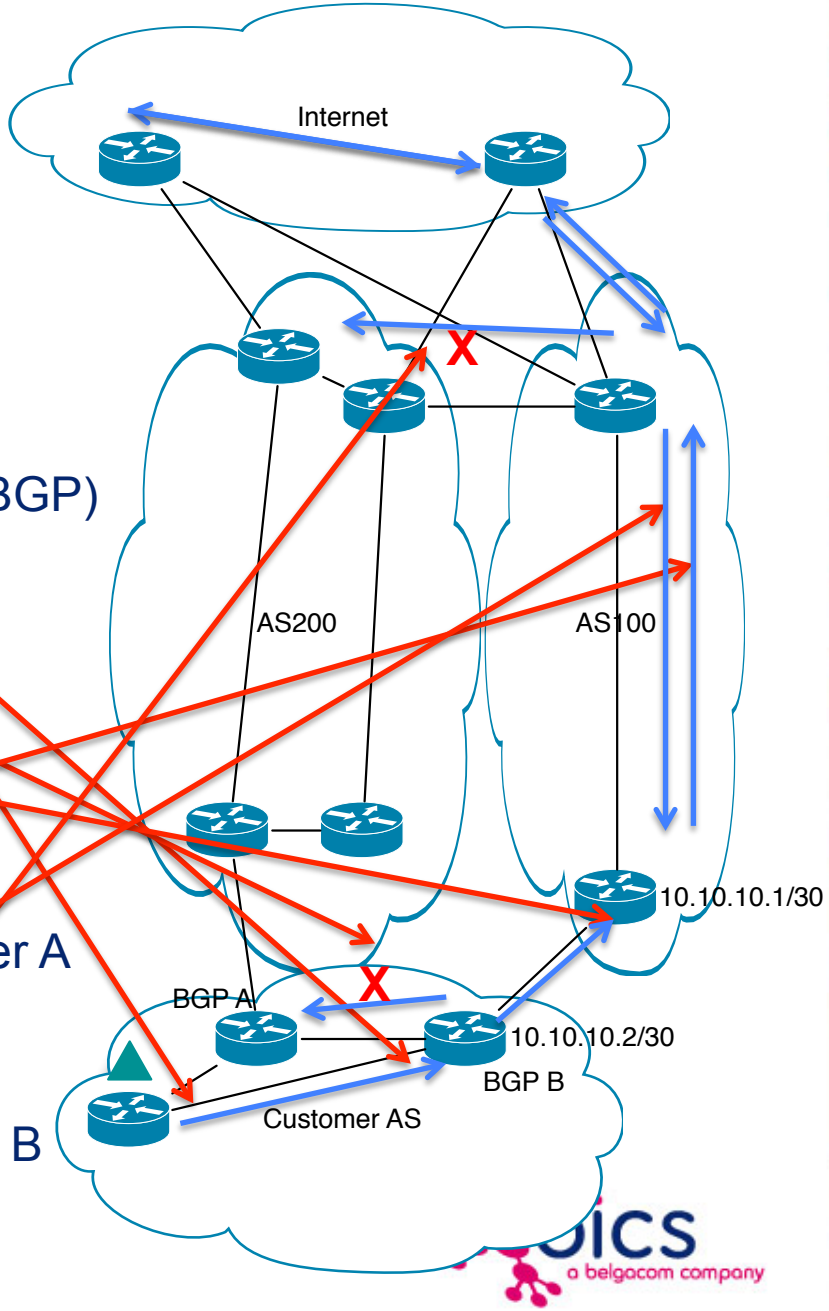
Router BGP B no lo anuncia a BGP A es iBGP)

Router BGP B lo anuncia a
Carrier B

El anuncio se propaga a través
de carrier B

Para evitar reenrutado del trafico por carrier A
podemos hacer prepend de carrier A

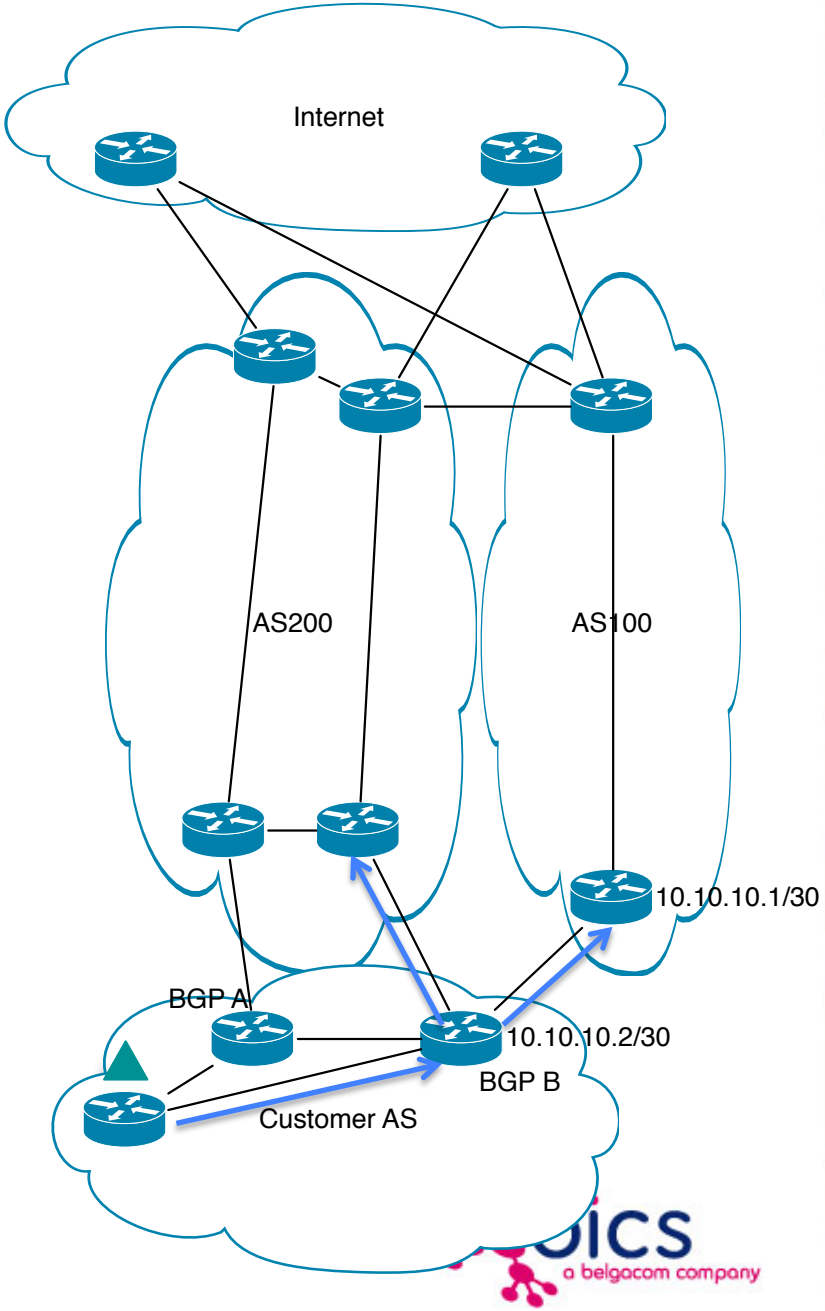
El tráfico a A.A.A.A siempre irá por Carrier B



# Downstream
# Mismo router para varios proveedores

# Problema

- Enviamos el anuncio al router
- No discrimina entre vecinos
  - El mismo anuncio por todos



# Solución

- Configurar una Community en router frontera
  - Community CUS:1
    - En vecino 1: quita community y anuncia
    - En vecino 2: descarta anuncio
  - Community CUS:2
    - En vecino 1: descarta anuncio
    - En vecino 2: quita community y anuncia
  - No community
    - Anuncia

# Solución soft

- Configurar una Community en router frontera
  - Community CUS:1
    - En vecino 1: quita community y anuncia con origen I
    - En vecino 2: quita community y anuncia con origen ?
  - Community CUS:2
    - En vecino 1: quita community y anuncia con origen ?
    - En vecino 2: quita community y anuncia con origen I
  - No community
    - Anuncia

# Solución

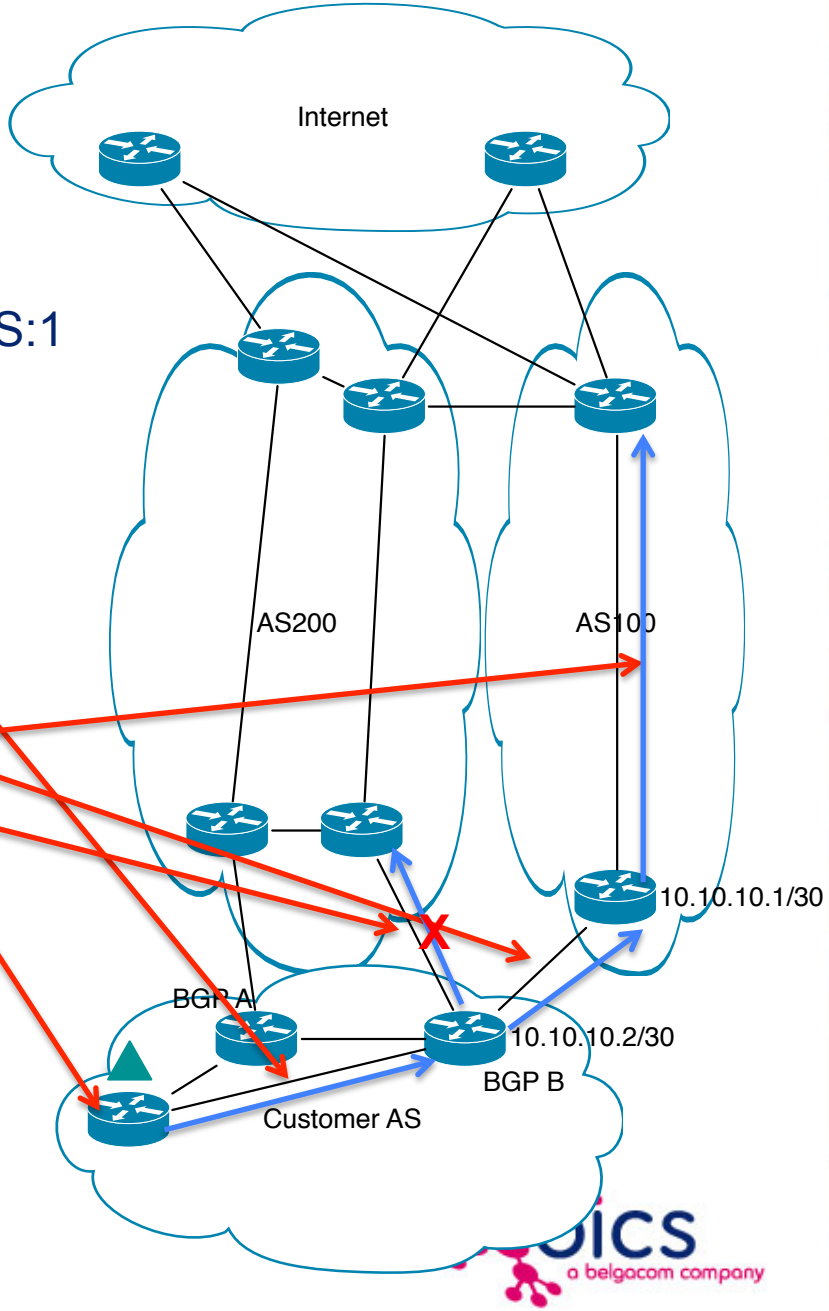
Prefijo desagregado

Anuncia el prefijo a BGP B community CUS:1

Router BGP B lo anuncia a AS100

Router BGP B non anuncia a AS200

El anuncio se propaga por AS100



# Ventajas del hacking centralizado

- Tolerante a fallos
  - Se cae el ANILLO, routing vuelve a BGP estándar
- Gestión más segura
- Routing inteligente
  - Basado en SDN
  - Basado en solución casera:
    - Cisco IP SLA con un ping a la loopback de máquina Unix
    - Máquina Unix shut/no shut basado en netflow, etc.
    - Anuncio de route-map basado en IP SLA

# DEPURACIÓN

# Traceroute

- Era una herramienta útil
- Ahora medicina homeopatica

PROBLEMA

- Tráfico no es simétrico
- No enseña el camino de vuelta

PROBLEMA PEOR

- Los clientes han aprendido a usarlo

# Traceroute.org

- Apartado Looking glass
- Recopilación de servidores looking glass
- No demasiado actualizado
- http://www.traceroute.org/



bics
a belgacom company

# Alternativa a traceroute

- ping con record route
- Puede grabar ida y vuelta
- limitada a 9 saltos
- http://www.traceroute.org/#Route%20Servers

# ping record route

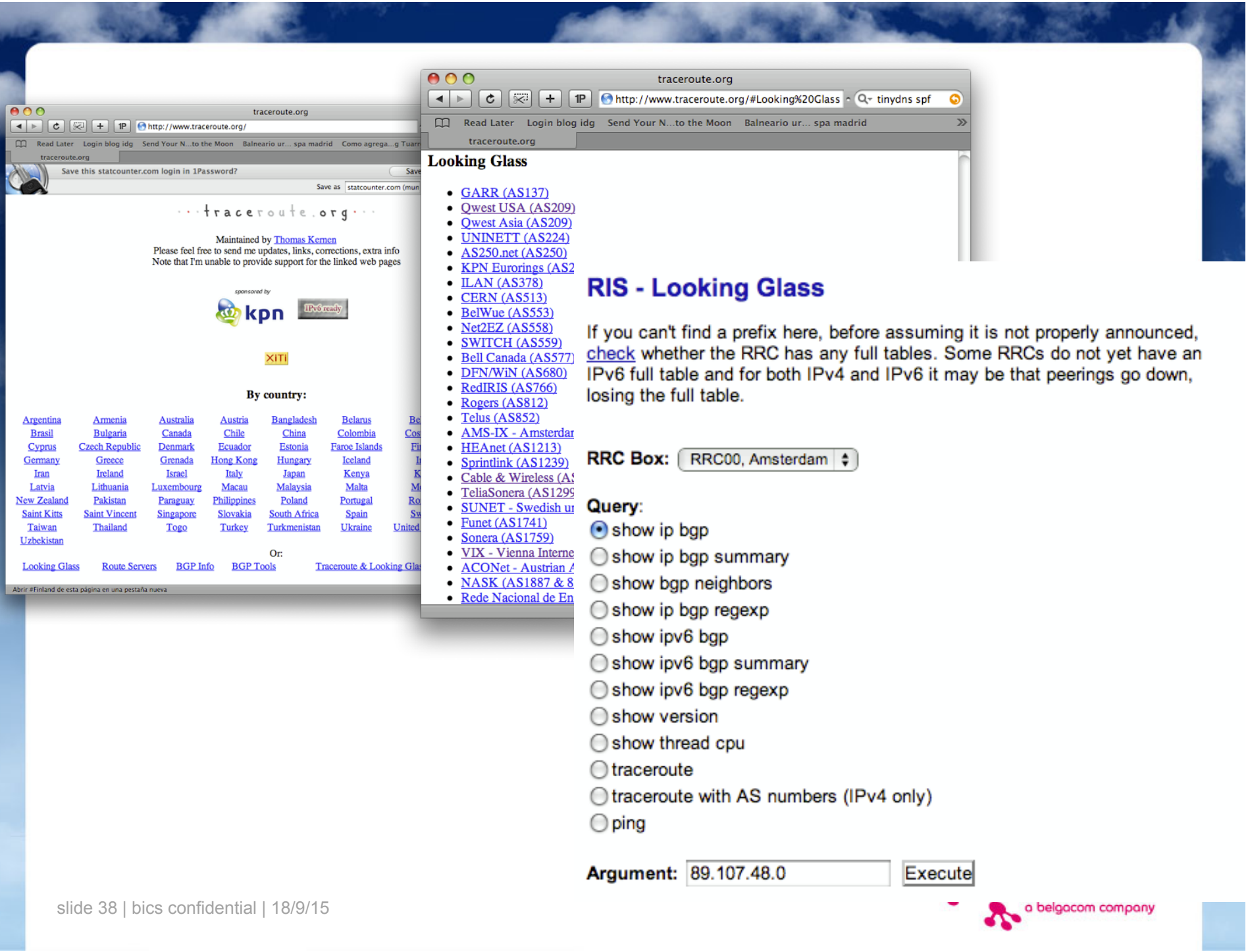
```
route-server>ping
Protocol [ip]:
Target IP address: 89.107.48.1
Repeat count [5]:
Datagram size [100]:
Timeout in seconds [2]:
Extended commands [n]: y
Source address or interface:
Type of service [0]:
Set DF bit in IP header? [no]:
Validate reply data? [no]:
Data pattern [0xABCD]:
Loose, Strict, Record, Timestamp, Verbose[none]: r
Number of hops [ 9 ]:
Loose, Strict, Record, Timestamp, Verbose[RV]:
Sweep range of sizes [n]:
Type escape sequence to abort.
```

# ping record route

```
Reply to request 0 (64 ms).  Received packet has options
 Total option bytes= 40, padded length=40
 Record route:
   Route-Server-gi0-1.belwue.net (129.143.103.78)
   Stuttgart-NWZ-Server-10GE-1-1.belwue.net
  (129.143.103.170)
   Stuttgart-NWZ-1-10GE-0-2-0-1.belwue.net
  (129.143.103.169)
   sgrt-b1.telia.net (62.115.128.80)
   ffm-bb2.telia.net (213.248.64.225)
   prs-bb2.telia.net (213.248.64.254)
   mad-b2.telia.net (80.91.255.171)
   (10.209.0.77)
   95.39.41.177.static.user.ono.com (95.39.41.177)
   <*>
```



bics
a belgacom company



# Show BGP en remoto

```
BGP routing table entry for 89.107.48.0/21
Paths: (14 available, best #7, table Default-IP-Routing-Table)
  Advertised to non peer-group peers:
  195.28.164.125 203.119.0.116
  3549 12956 3352 39780 39780 39780 39780
    208.51.134.248 from 208.51.134.248 (67.17.80.217)
      Origin incomplete, metric 2937, localpref 100, valid, external
      Community: 3549:2293 3549:30840
      Last update: Wed Oct  8 11:46:20 2008

  3333 5511 12479 39780
    193.0.0.56 from 193.0.0.56 (193.0.0.56)
      Origin IGP, localpref 100, valid, external
      Last update: Wed Oct  8 09:12:53 2008

  42109 41965 41877 12389 8928 31479 39780 39780 39780 39780 39780
    91.103.24.1 from 91.103.24.1 (91.103.24.1)
      Origin EGP, localpref 100, valid, external
      Last update: Tue Oct  7 23:16:55 2008

  3.5 1125 1103 3257 31479 39780 39780 39780 39780 39780
    145.125.80.5 from 145.125.80.5 (145.125.80.5)
      Origin IGP, localpref 100, valid, external
      Community: 1103:1000 3257:4000 3257:5034
      Last update: Tue Oct  7 14:23:41 2008
```



bics a belgacom company

# RIPE Stat

- Resumen de datos de un AS
- Gran cantidad de información histórica y estadística

https://stat.ripe.net/asXXXX



bics
a belgacom company

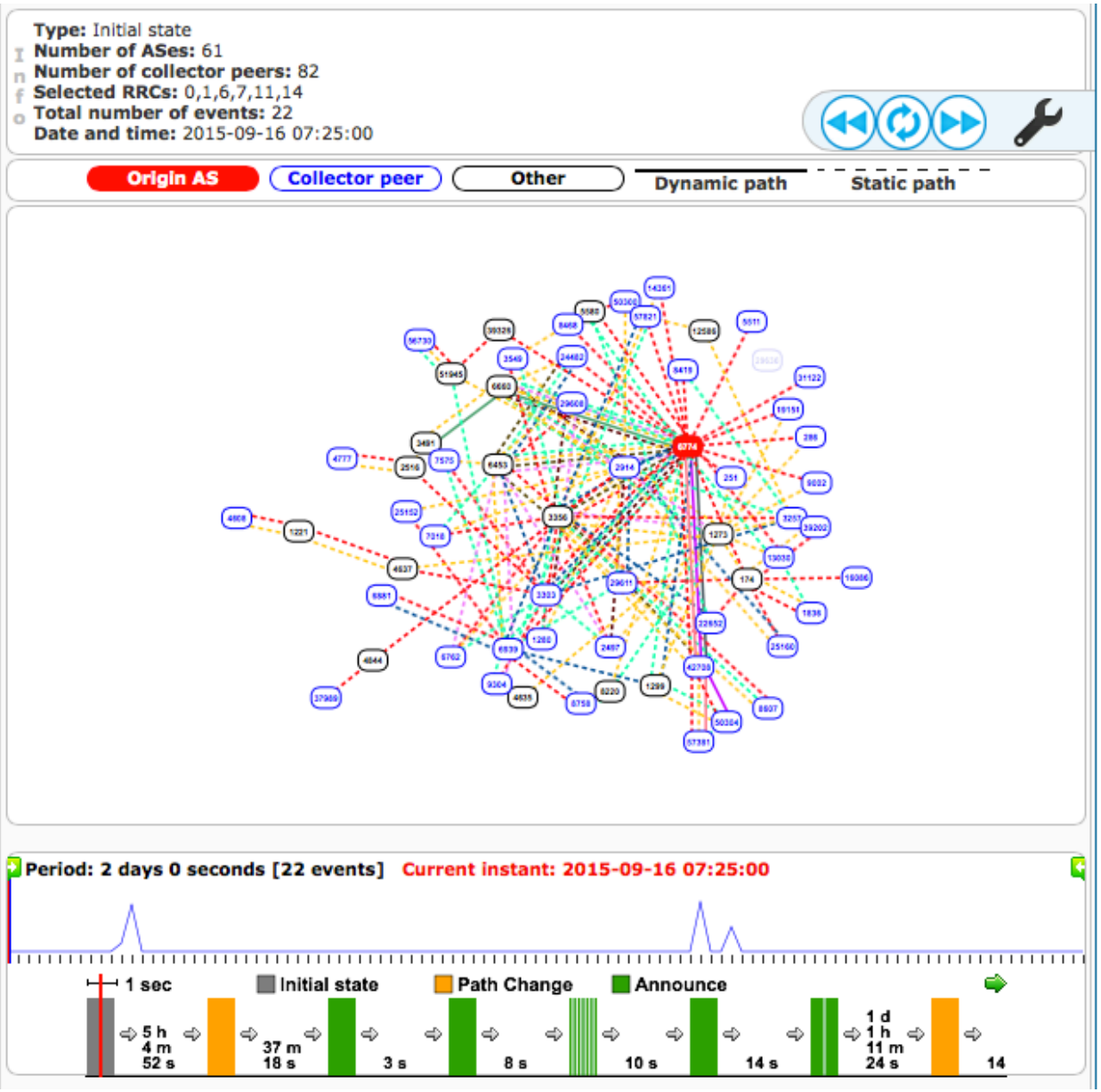
# BGP Play

- Visión gráfica e histórica de un prefijo en la red
- Visión parcial pero bastante útil

http://bgplay.routeviews.org/bgplay/

http://www.ris.ripe.net/bgplay/

# BGP Play






bics
a belgacom company


bics
a belgacom company