MIKROTIK USERS MEETING
UIO / 19

OSPF & BGP / Ventajas, Desventajas y Casos de Estudio

Andres Ocampo D.

ACERCA DE MI

- **Trayectoria**
  - + 8 años de experiencia en Service Providers / Networking
  - Diseño e implementación de soluciones multivendor a escala mundial
  - OSPF / BGP / MPLS / EIGRP / STP / IS-IS / VPNs (L2TP, IPSec, OVPN)
- **Certificaciones**
  - Mikrotik: MTCNA / MTCTCE / MTCRE / MTCINE
  - Consultor Certificado para Mikrotik desde 2013
  - Otras Certificaciones: CCNA Security / CCNA R&S / CCNP R&S/ CCIE R&S (en proceso)
- **Presentaciones anteriores**
  - MUM / Quito - 2017 / Seguridad informática hacia RouterOS
  - MUM / Ciudad de México - 2018 / BGP Troubleshooting
  - MUM / San Jose - 2018 / OSPF Troubleshooting
  - MUM / Tegucigalpa - 2018 / Como influenciar rutas utilizando BGP & Prefix-list en RouterOS

ACERCA DE LA EMPRESA

Fibramax

- Brindamos servicio de internet a nivel nacional con infraestructura autónoma.
- Cobertura en las principales ciudades:
  - Quito / Cuenca / Guayaquil / Manta / Ibarra / Ambato / Latacunga y creciendo.
- Mayor información de la empresa en el Stand

OSPF

- Principios básicos de operación:
  - Mantiene toda la topología de la red dentro de la base de datos en forma de LSAs.
  - Sumarización únicamente en ABRs o ASBRs (filtrado de summary-LSAs no soporta RouterOS)
  - Timers para convergencia por default; paquete hello enviado cada 10secs / dead-interval de 40 secs.
  - Posibilidad de implementar MPLS Traffic Engineering
- Algoritmo de elección de ruta:
  - Primero escoge entre intra-area, inter-area, external type1, external type2
  - Segundo selecciona en función de los costos de las interfaces
  - La decisión de enrolamiento se basa automáticamente en lo que el router haya calculado a través del SPF y el algoritmo de Djistra.

OSPF / VENTAJAS

- Agilidad en la convergencia de rutas
- Permite realizar balanceo de carga a través de ECMP
- Protocolo confiable si está bien diseñado
  - Manejo de Multi-área (aprox 50 routers máximo por area)
  - Diseño de costos (máximo 65535 por interfaz)
  - Diseño de ASBR (redistribuciones) y ABR (inter-area routers)

OSPF / DESVENTAJAS

- Diseño requiere que todas las areas externas tengan siempre conexión con area de backbone (ID: 0.0.0.0)

- BUGs reportados en RouterOS:

- Virtual-links (generados para conectar un área 2 a través de un área X al area 0, genera problemas al manejar diversos LSAs)

- Elección de un área Stub, la inyección de la ruta por defecto en ciertas versiones de RouterOS se desactiva por error del software

- Autenticación MD5 puede generar problemas en la sincronización de elección en Master/Slave en la parte de sincronización de neighbors en status ex-start de OSPF.

BGP

- Principios básicos de operación
  - Maneja una conexión entre routers a través del protocolo TCP y puerto 179.
  - Existen más de 10 atributos posibles de selección para elegir una ruta.
  - Protocolo aceptado a nivel mundial para intercambio de prefijos

BGP / VENTAJAS

- Protocolo muy escalable en el que valida muchas posibles métricas antes de elegir un camino.
    - Atributos comunes: weight, local-preference, as-path, MED, comunidades
- Permite manejar balanceo de carga por proveedores únicamente al modificar prefijos o atributos desde la red interna.
- Permite la creación de diferentes address-families; ipv4, ipv6, vpn4, l2vpn, l2vpn-cisco para crear VPNs capa3 o capa2 utilizando MPLS / VPLS con BGP.
- Dependiendo de la topología se puede crear una versión de BGP con RouteReflectors o Confederaciones para evitar una full-mesh de sesiones BGP

BGP / DESVENTAJAS

- Tiempo de convergencia para BGP por defecto es 180 segundos, para un protocolo de red interna manejar iBGP como IGP generaría lentitud en la convergencia.

- BUGs reportados para RouterOSv6:
  - RouterOS v6 aún maneja BGP como un proceso single-core, el cual se siente más cuando el router maneja full-routing-table, a menos que se hayan instalado en CHR.
  - RouterOS v6 no soporta comunidades extendidas de 32bits.

RECURSIVIDAD

- Recursividad se conoce como la técnica de utilizar la misma tabla de enrutamiento para definir el next-hop de un gateway que no es alcanzable localmente, sino a través de otra ruta

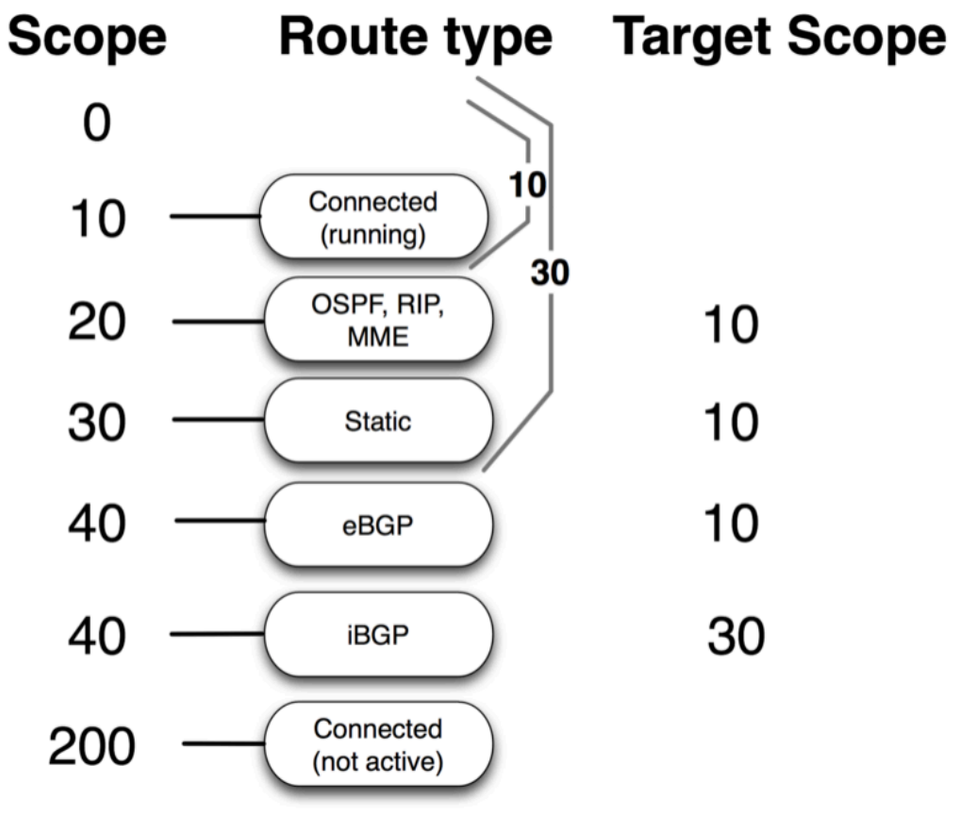
- Scope / Target-Scope son atributos que permiten a RouterOS escoger una ruta como recursiva para otra, es decir:

  - 10.1.1.1 es alcanzable por el next-hop 1.1.1.2 (conectado directamente)

  - Ruta B es alcanzable por el next-hop 10.1.1.1 (no conectado directamente)

  - Si los parámetros de los scope/target scope están bien configurados entonces RouterOS generará que la Ruta B es alcanzable por el next-hop 10.1.1.1 el cual es recursivo a través del next-hop 1.1.1.2.

RECURSIVIDAD

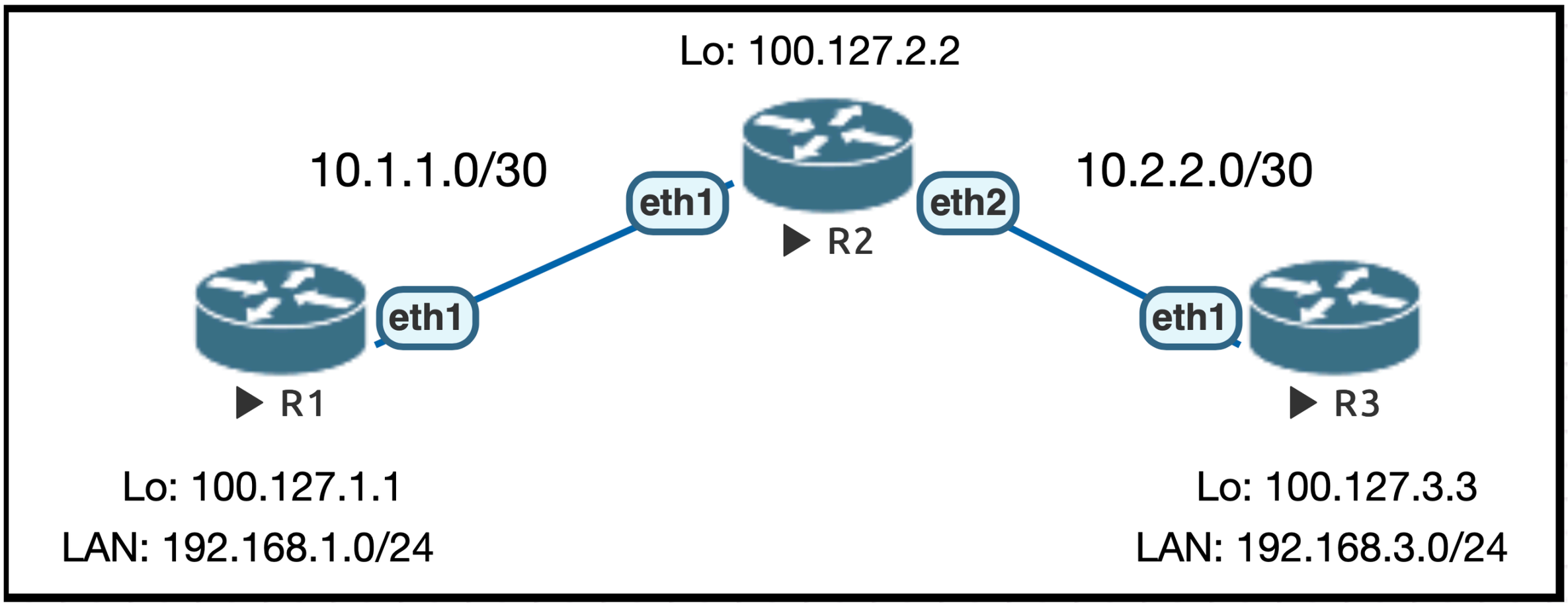


Funciona solo en la tabla de enrutamiento *main*

- La regla: target-scope de la ruta recursiva revisa si el scope de la ruta actual es menor o igual

RECURSIVIDAD

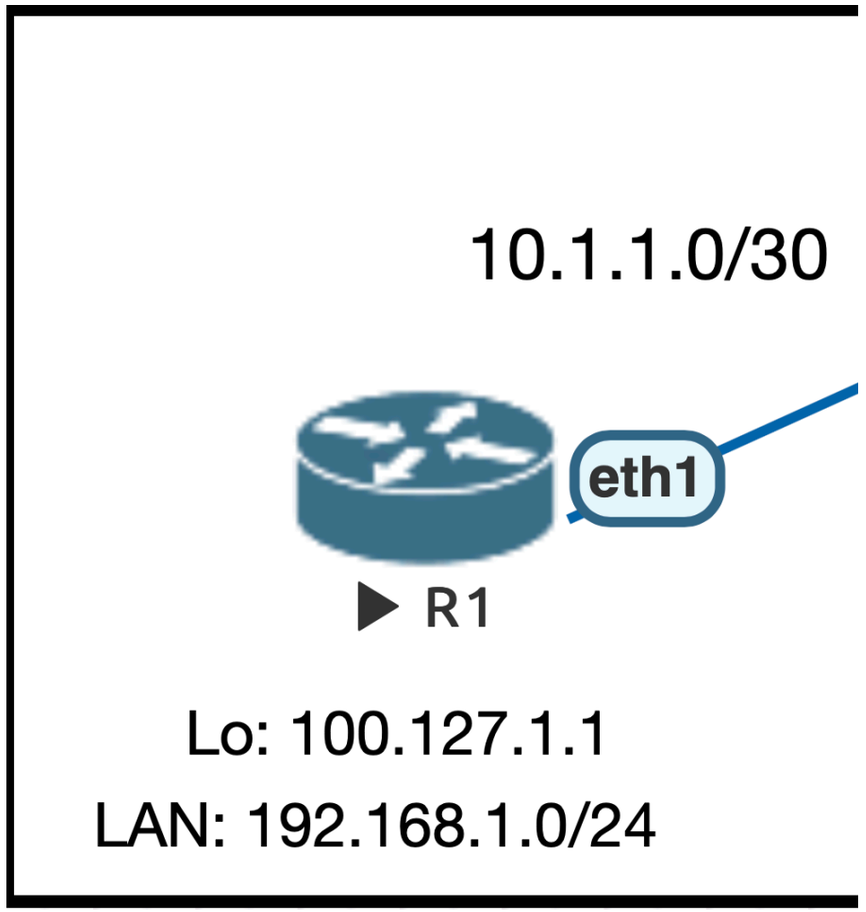
- Ejemplo; rutas estáticas



RECURSIVIDAD

- Ejemplo; LAN a través de Loopback

```
[admin@R1] > ip route print detail
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0 ADC  dst-address=10.1.1.0/30 pref-src=10.1.1.1 gateway=ether1
        gateway-status=ether1 reachable distance=0 scope=10

 1 A S  dst-address=10.2.2.0/30 gateway=10.1.1.2
        gateway-status=10.1.1.2 reachable via  ether1 distance=1 scope=30
        target-scope=10

 2 ADC  dst-address=100.127.1.1/32 pref-src=100.127.1.1 gateway=Lo0
        gateway-status=Lo0 reachable distance=0 scope=10

 3 A S  dst-address=100.127.2.2/32 gateway=10.1.1.2
        gateway-status=10.1.1.2 reachable via  ether1 distance=1 scope=30
        target-scope=10

 4 A S  dst-address=100.127.3.3/32 gateway=10.1.1.2
        gateway-status=10.1.1.2 reachable via  ether1 distance=1 scope=30
        target-scope=10

 5 ADC  dst-address=192.168.1.0/24 pref-src=192.168.1.1 gateway=LAN
        gateway-status=LAN reachable distance=0 scope=10
```



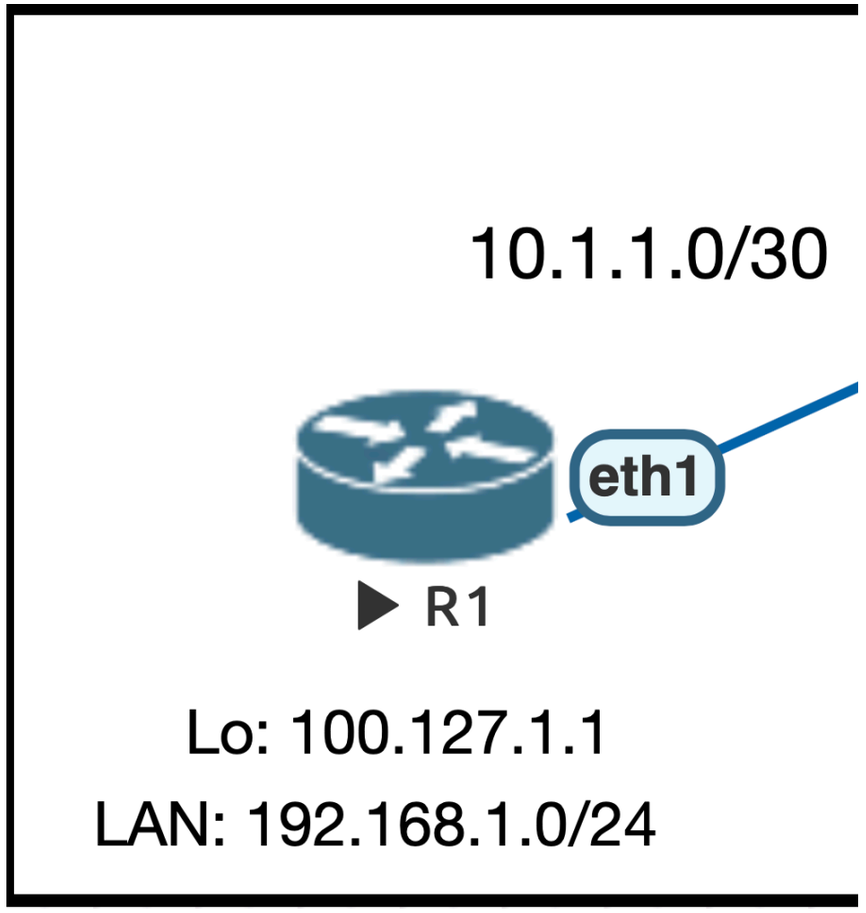
RECURSIVIDAD

- Ejemplo; rutas estáticas

```
[admin@R1] > ip route add dst-address=192.168.3.0/24 gateway=100.127.3.3
```

```
[admin@R1] > ip route print detail where dst-address=192.168.3.0/24
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0   S  dst-address=192.168.3.0/24 gateway=100.127.3.3
        gateway-status=100.127.3.3 unreachable distance=1 scope=30
        target-scope=10
```

```
[admin@R1] > ip route print detail where dst-address=100.127.3.3/32
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0 A S  dst-address=100.127.3.3/32 gateway=10.1.1.2
        gateway-status=10.1.1.2 reachable via  ether1 distance=1 scope=30
        target-scope=10
```

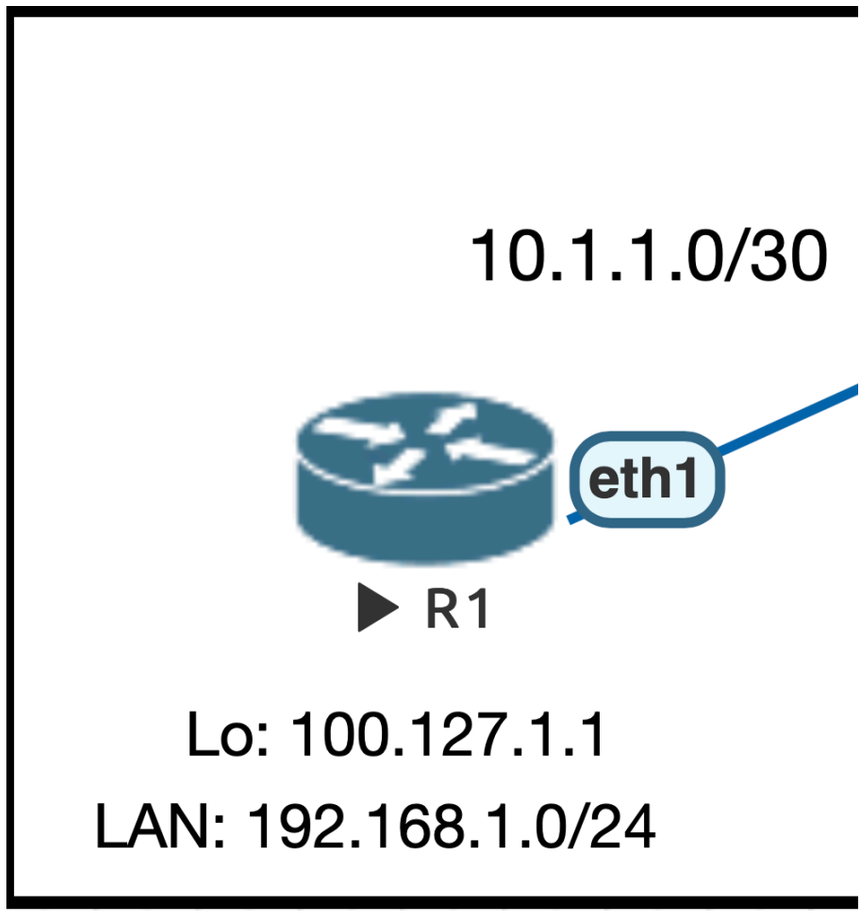


RECURSIVIDAD

- Ejemplo; rutas estáticas

```
[admin@R1] > ip route set [find where dst-address=192.168.3.0/24] target-
scope=30
```

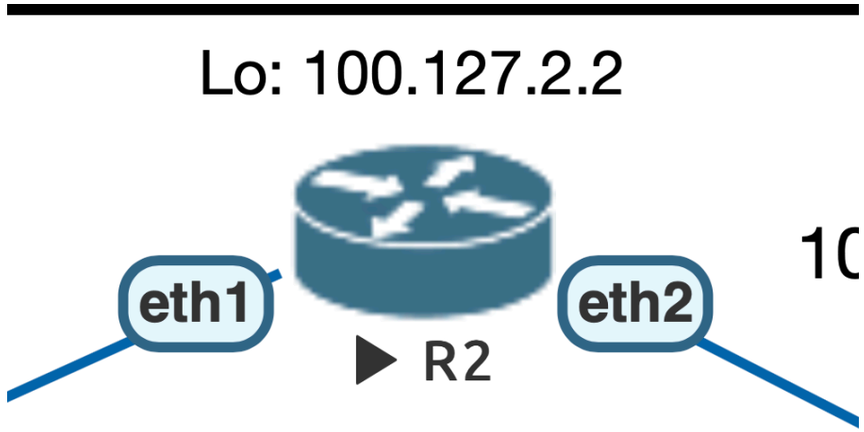
```
[admin@R1] > ip route print detail where dst-address=192.168.3.0/24
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0 A S  dst-address=192.168.3.0/24 gateway=100.127.3.3
        gateway-status=100.127.3.3 recursive via 10.1.1.2 ether1 distance=1
        scope=30 target-scope=30

[admin@R1] > ip route print detail where dst-address=100.127.3.3/32
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0 A S  dst-address=100.127.3.3/32 gateway=10.1.1.2
        gateway-status=10.1.1.2 reachable via  ether1 distance=1 scope=30
        target-scope=10
```



RECURSIVIDAD

- Ejemplo; rutas estáticas

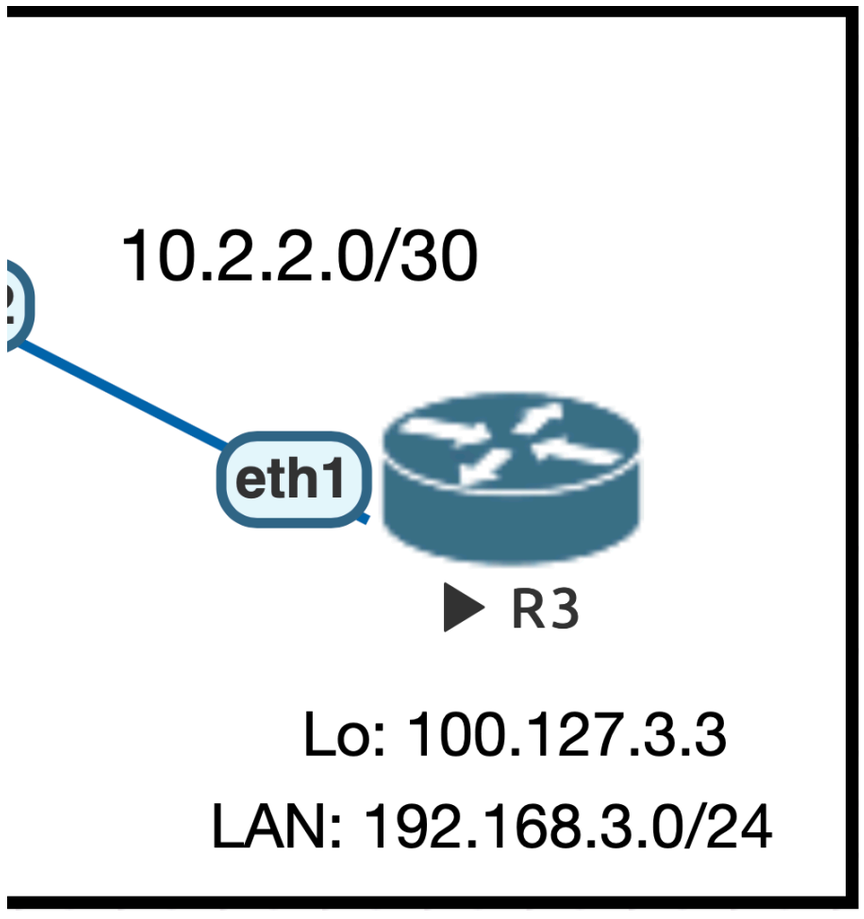
```
[admin@R2] > ip route print detail
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0 ADC  dst-address=10.1.1.0/30 pref-src=10.1.1.2 gateway=ether1
        gateway-status=ether1 reachable distance=0 scope=10

 1 ADC  dst-address=10.2.2.0/30 pref-src=10.2.2.1 gateway=ether2
        gateway-status=ether2 reachable distance=0 scope=10

 2 A S  dst-address=100.127.1.1/32 gateway=10.1.1.1
        gateway-status=10.1.1.1 reachable via  ether1 distance=1 scope=30
        target-scope=10

 3 ADC  dst-address=100.127.2.2/32 pref-src=100.127.2.2 gateway=Lo0
        gateway-status=Lo0 reachable distance=0 scope=10

 4 A S  dst-address=100.127.3.3/32 gateway=10.2.2.2
        gateway-status=10.2.2.2 reachable via  ether2 distance=1 scope=30
        target-scope=10

 5 A S  dst-address=192.168.1.0/24 gateway=100.127.1.1
        gateway-status=100.127.1.1 recursive via 10.1.1.1 ether1 distance=1
        scope=30 target-scope=30

 6 A S  dst-address=192.168.3.0/24 gateway=100.127.3.3
        gateway-status=100.127.3.3 recursive via 10.2.2.2 ether2 distance=1
        scope=30 target-scope=30
```



RECURSIVIDAD

- Ejemplo; rutas estáticas

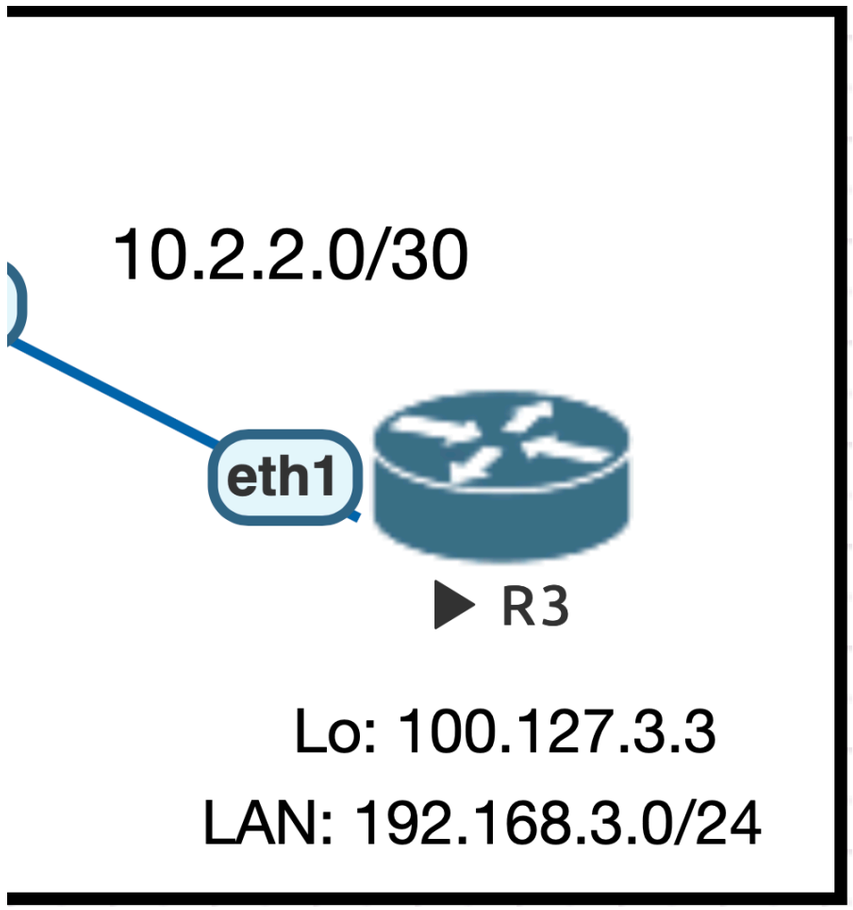
```
[admin@R3] > ip route print detail
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
 0 A S  dst-address=10.1.1.0/30 gateway=10.2.2.1
        gateway-status=10.2.2.1 reachable via  ether1 distance=1 scope=30
        target-scope=10

 1 ADC  dst-address=10.2.2.0/30 pref-src=10.2.2.2 gateway=ether1
        gateway-status=ether1 reachable distance=0 scope=10

 2 A S  dst-address=100.127.1.1/32 gateway=10.2.2.1
        gateway-status=10.2.2.1 reachable via  ether1 distance=1 scope=30
        target-scope=10

 3 A S  dst-address=100.127.2.2/32 gateway=10.2.2.1
        gateway-status=10.2.2.1 reachable via  ether1 distance=1 scope=30
        target-scope=10

 4 ADC  dst-address=100.127.3.3/32 pref-src=100.127.3.3 gateway=Lo0
        gateway-status=Lo0 reachable distance=0 scope=10

 5 A S  dst-address=192.168.1.0/24 gateway=100.127.1.1
        gateway-status=100.127.1.1 recursive via 10.2.2.1 ether1 distance=1
        scope=30 target-scope=30

 6 ADC  dst-address=192.168.3.0/24 pref-src=192.168.3.1 gateway=LAN
        gateway-status=LAN reachable distance=0 scope=10
```



RECURSIVIDAD

- Ejemplo; rutas estáticas

```
[admin@R3] > ping 192.168.1.1 src-address=192.168.3.1
  SEQ HOST                                     SIZE TTL TIME  STATUS
    0 192.168.1.1                                56  63 3ms
    1 192.168.1.1                                56  63 5ms
    2 192.168.1.1                                56  63 3ms
    3 192.168.1.1                                56  63 3ms
    4 192.168.1.1                                56  63 3ms
    sent=5 received=5 packet-loss=0% min-rtt=3ms avg-rtt=3ms max-rtt=5ms

[admin@R3] > tool traceroute 192.168.1.1 src-address=192.168.3.1
 # ADDRESS                          LOSS SENT    LAST     AVG    BEST   WORST
 1 10.2.2.1                           0%    4   6.8ms     7.5     1.8    19.4
 2 192.168.1.1                        0%    4   3.7ms     3.1     2.6     3.7

[admin@R3] >
```



OSPF (IGP) & BGP (EGP)

- **Nueva Estructura**
  - La manera más util de optimizar una red es hablando OSPF y BGP en conjunto
  - **IGP** (interior gateway protocol) de la red, para RouterOS OSPF (algún soporte futuro de ISIS??)
    - Convergencia entre WANs o enlaces redundantes
    - Tránsito de Loopbacks para establecer sesiones iBGP
    - Balanceo de Carga a través de ECMP
  - **EGP** (exterior gateway protocol) de la red, el único protocolo en esta parte es BGP
    - Tránsito de prefijos y subredes que no sean tránsito
    - Manejo de la ruta por defecto
    - Full mesh no es necesaria si se manejan topologías con route-reflectors o confederaciones

CASOS DE ESTUDIO

- Casos de estudio sobre esta topología:
  - Sesiones BGP hacia proveedores internacionales con AS público, NAT no necesario en el borde.
  - Proveedor entrega un /30 o /29 de públicas para salir al internet, NAT casi obligado en el borde

ASN PUBLICO

- Requisitos para este caso de estudio, una ASN público asignado por LACNIC
  - Para este ejemplo: ASN 1000 (WISP)
- Sesiones eBGP hacia proveedores internacionales (AS100 & AS200)
- Sesiones iBGP hacia la topología interna (route-reflector HA design)

Rutas dentro de Tier1:
8.8.8.8
9.9.9.9

Tier1 (Century Link)
AS: 3356
Tier1
eth2
eth1
4.0.0.0/30
4.0.0.4/30

ISP1 -> AS: 100
Provider1
eth2
eth1
100.0.0.0/30

Provider2
eth2
eth1
200.0.0.0/30
ISP2 -> AS: 200

DOWN desde 8.8.8.8 a 150.0.0.1

UP desde 150.0.0.1 a 8.8.8.8

BGP: AS: 1000
OSPFv2:
area -> backbone

eBGP01
Lo: .100
R1
eth1
.0/30
eth1
LSR01
Lo: .3
eth4
.12/30
eth1
RR1
RR1
Lo: .1

eth2
R2
eth3
.8/30
eth2
R4
LSR02
Lo: .4
eth2
.20/30
eth3
.24/30
eth4
.16/30
eth1
eth2
R3
Customer 01
Lo: .10
Public: 150.0.0.1

eth2
eBGP02
Lo: .200
eth1
R5
.4/30
eth1
eth1
RR2
RR2
Lo: .2
WISP/ISP

WANs: 10.1.0.x/30
Loopbacks: 10.2.0.x/32
Publics: 150.0.0.0/22

Datos a explicar en el ejemplo:
- Configuración de OSPF / punto a punto
- Gateway de BGP sobre OSPF
- Validar la topología de OSPF
- LSR sin BGP
- Convergencia entre Routers
- Configuración de BGP / route-reflectors
- Enrutamiento de BGP (ruta por defecto y prefijos)
- Recursividad entre BGP y OSPF
- Enrutamiento de BGP proveedores
- Asimetrías provocadas

ASN PRIVADO

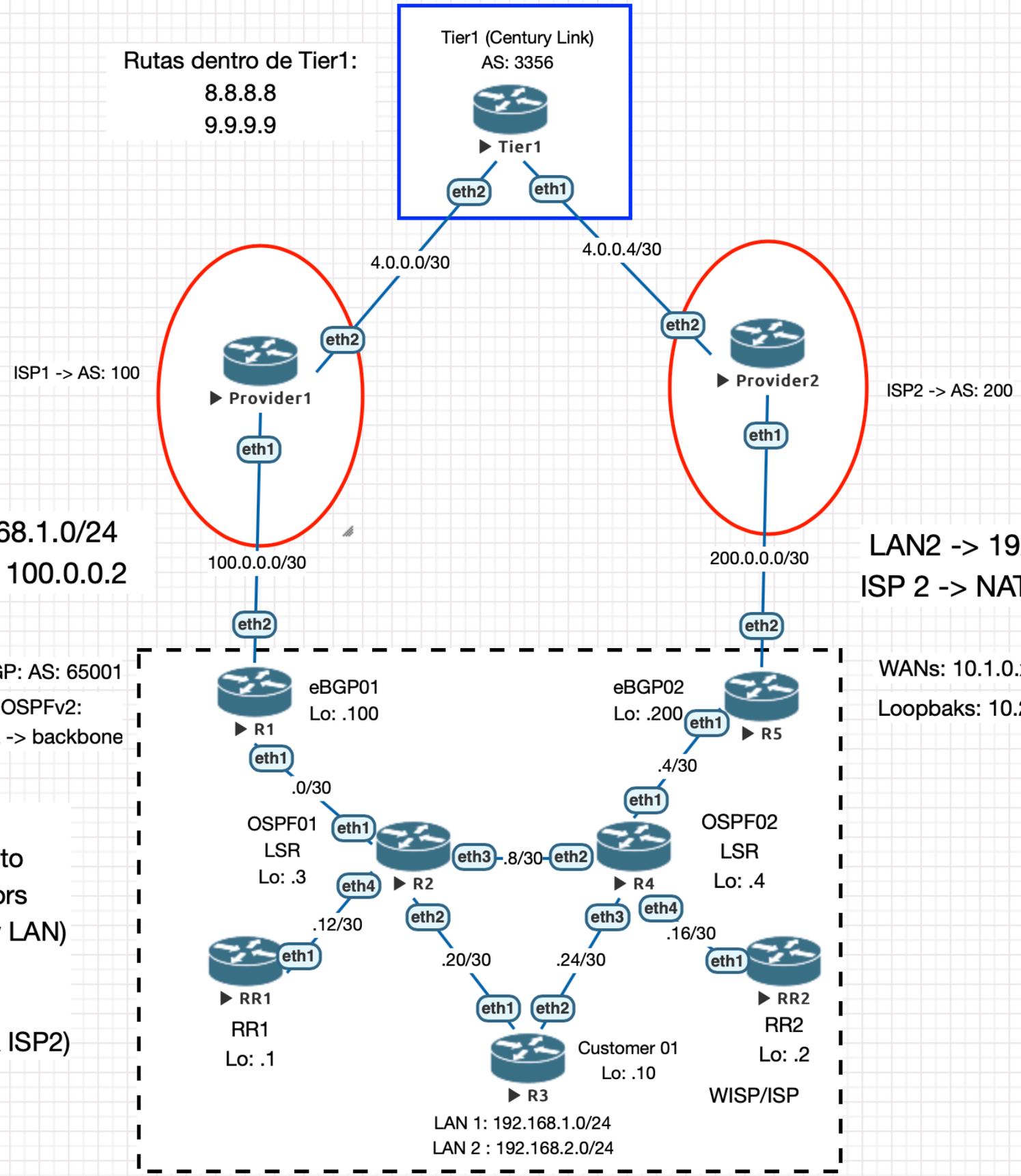
- Requisitos para este caso de estudio:

- Para este ejemplo: ASN 65001 (WISP)

- Sesiones iBGP hacia la topología interna (route-reflector HA design)

- IPs públicas de cada proveedor /30

Rutas dentro de Tier1:
8.8.8.8
9.9.9.9



LAN1 -> 192.168.1.0/24
ISP 1 -> NAT to 100.0.0.2

LAN2 -> 192.168.2.0/24
ISP 2 -> NAT to 200.0.0.2

WANs: 10.1.0.x/30

Loopbaks: 10.2.0.x/32

Datos a explicar en el ejemplo:
- Configuración de OSPF / punto a punto
- Configuración de BGP / route-reflectors
- Enrutamiento de BGP (ruta por defecto y LAN)
- Recursividad entre BGP y OSPF
- NAT en los equipos de borde
- Balanceo de carga LAN1 & LAN2 (ISP1 & ISP2)

MUCHAS GRACIAS!!

Preguntas…