

BGP – podstawy działania, polityka w sieciach TCP/IP.

O czym ta mowa...

- routing w sieciach TCP/IP (forwarding/routing statyczny/dynamiczny, link state, distance vector)
- BGP zasady funkcjonowanie, pojęcie AS, prefix, as_path
- więcej o BGP:
 - atrybuty
 - filtrowanie
 - route mapy
 - community
- przykładowa konfiguracja
- Quagga

Pojęcia pierwotne...;)

FORWARDING

ROUTING

PROTOKOŁY ROUTUJĄCE

DYNAMICZNE/STATYCZNE

LINK-STATE/DISTANCE-VECTOR(PATH-VECTOR)

BGP \$ ^^^H... 4

BGP = **B**order **G**ateway **P**rotocol

EGP – Exterior Routing Protocol

Policy Based – bardzo polityko-lubny

EGP dzisiejszego internetu

podobno skomplikowany (coś w tym jest..)

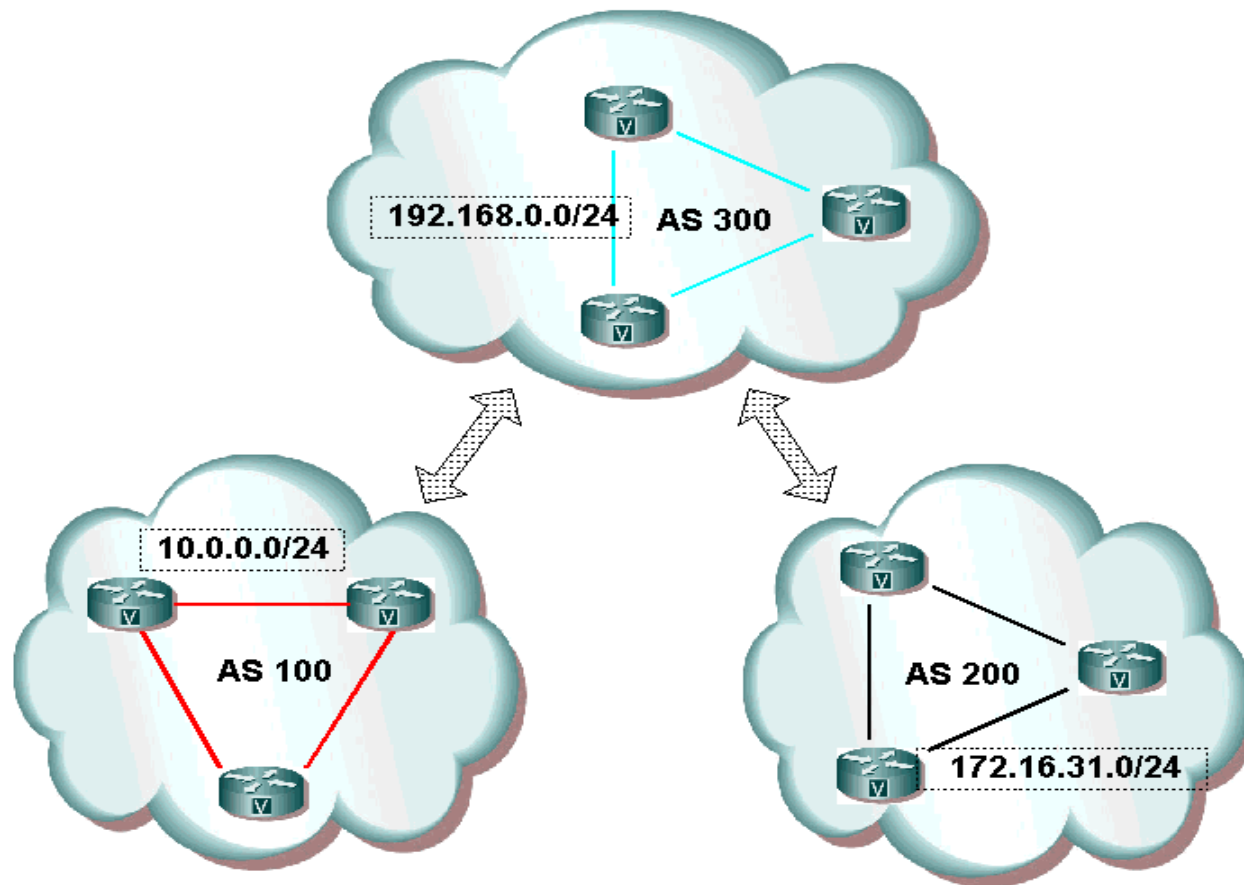
Protokół typu Distance-Vector (choć w zasadzie Path-Vector)

Pojęcie AS – to fundament BGP:)

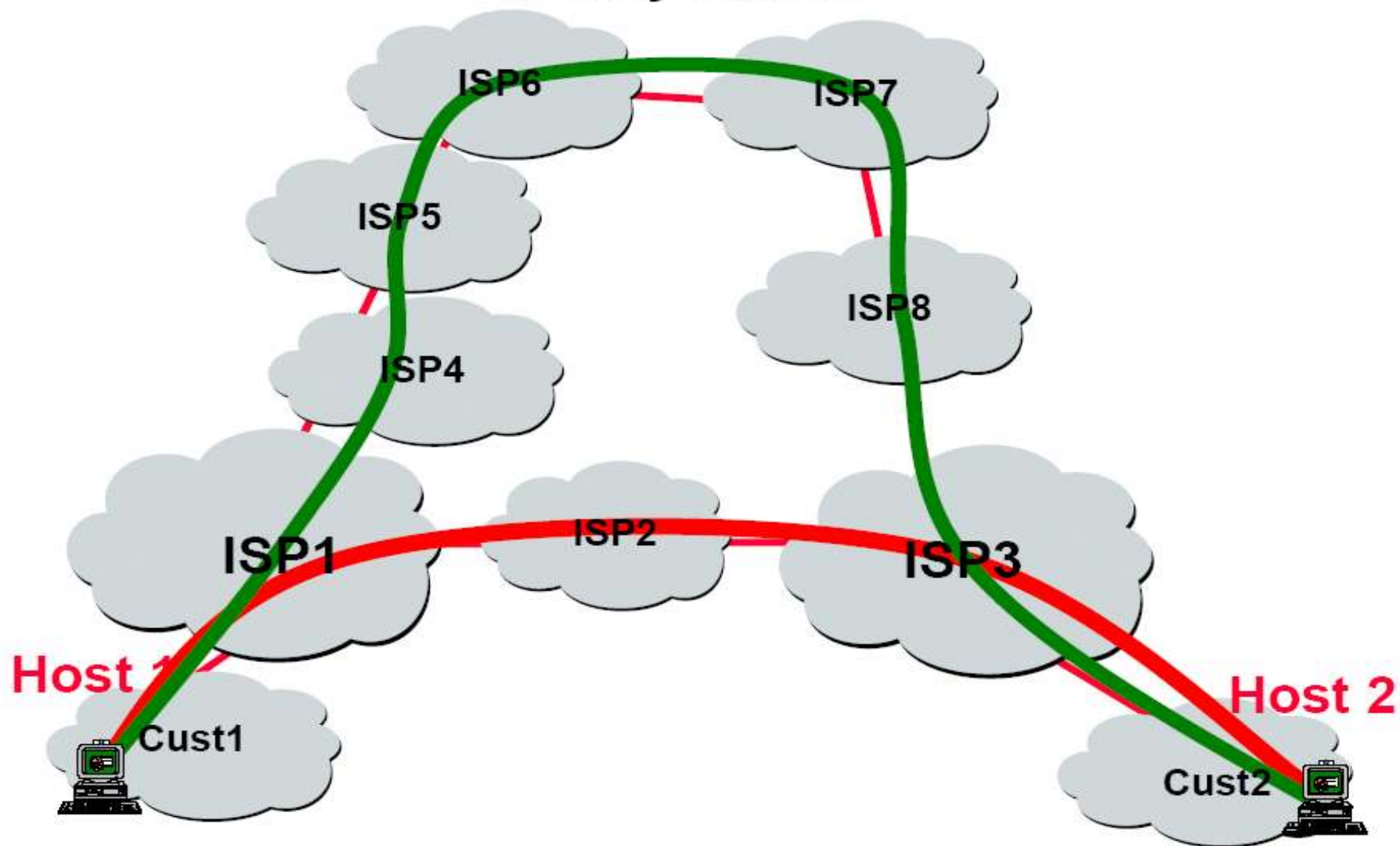
```
#whois as8246
```

```
aut-num:          AS8246
as-name:          INTERNET-TECHNOLOGIES-POLSKA-AS
descr:           Internet Technologies Polska
descr:           GTS Internet Partners
descr:           al. Niepodleglosci 69
descr:           02-626 Warsaw, Poland
...
...
```

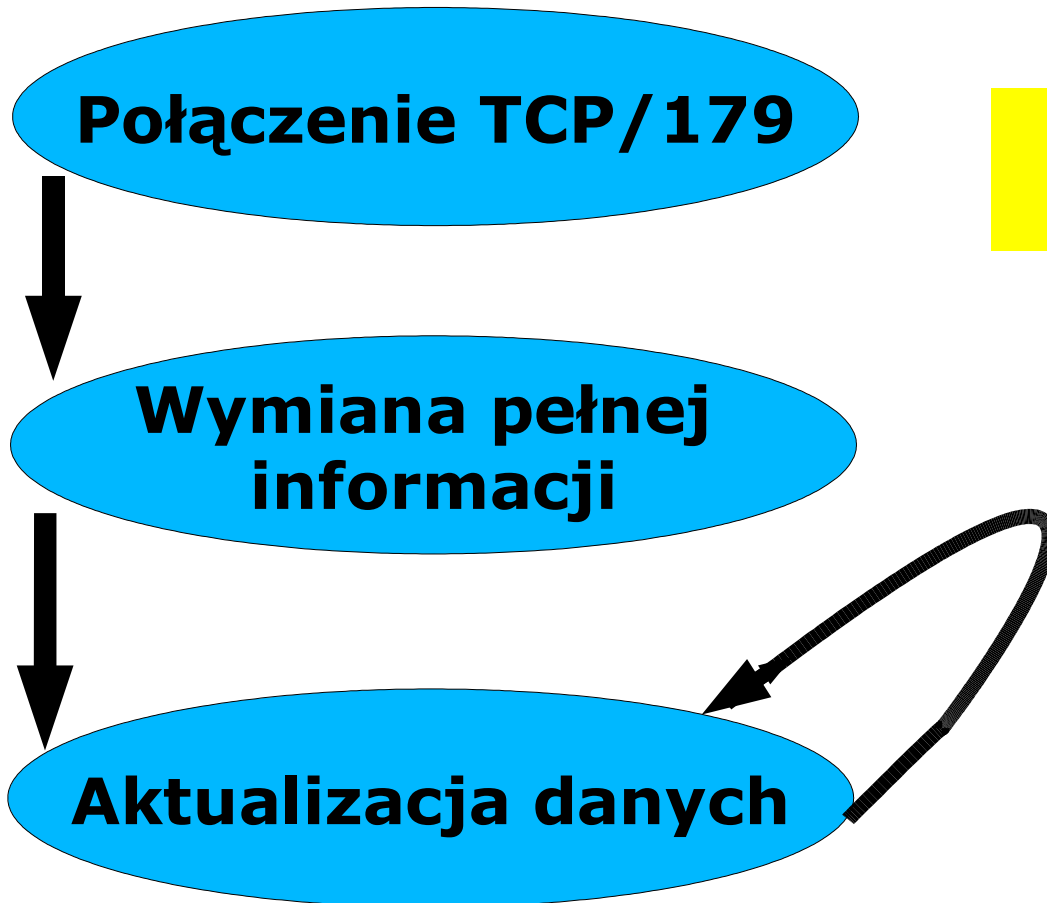
BGP to wymiana informacji pomiędzy AS



Money talks...



Jak to działa?



**BGP route =
prefixy + atrybuty**

Ustawianie połączenia

- Idle
- Connect
- OpenSent
- Active
- Confirmed
- Established

```
router-bgpd# show ip bgp neighbors 157.25.1.16
BGP neighbor is 157.25.1.16, remote AS 8246, local AS
29620, external link
Description: "Link do GTS"
BGP version 4, remote router ID 157.25.1.16
BGP state = Established, up for 01w0d15h
```

Typy pakietów BGP

OPEN MESSAGE – pakiet ten jest wymieniany pomiędzy routerami zaraz po zestawieniu sesji TCP. Przekazywane są w nim podstawowe informacje potrzebne do skonfigurowania połączenia.

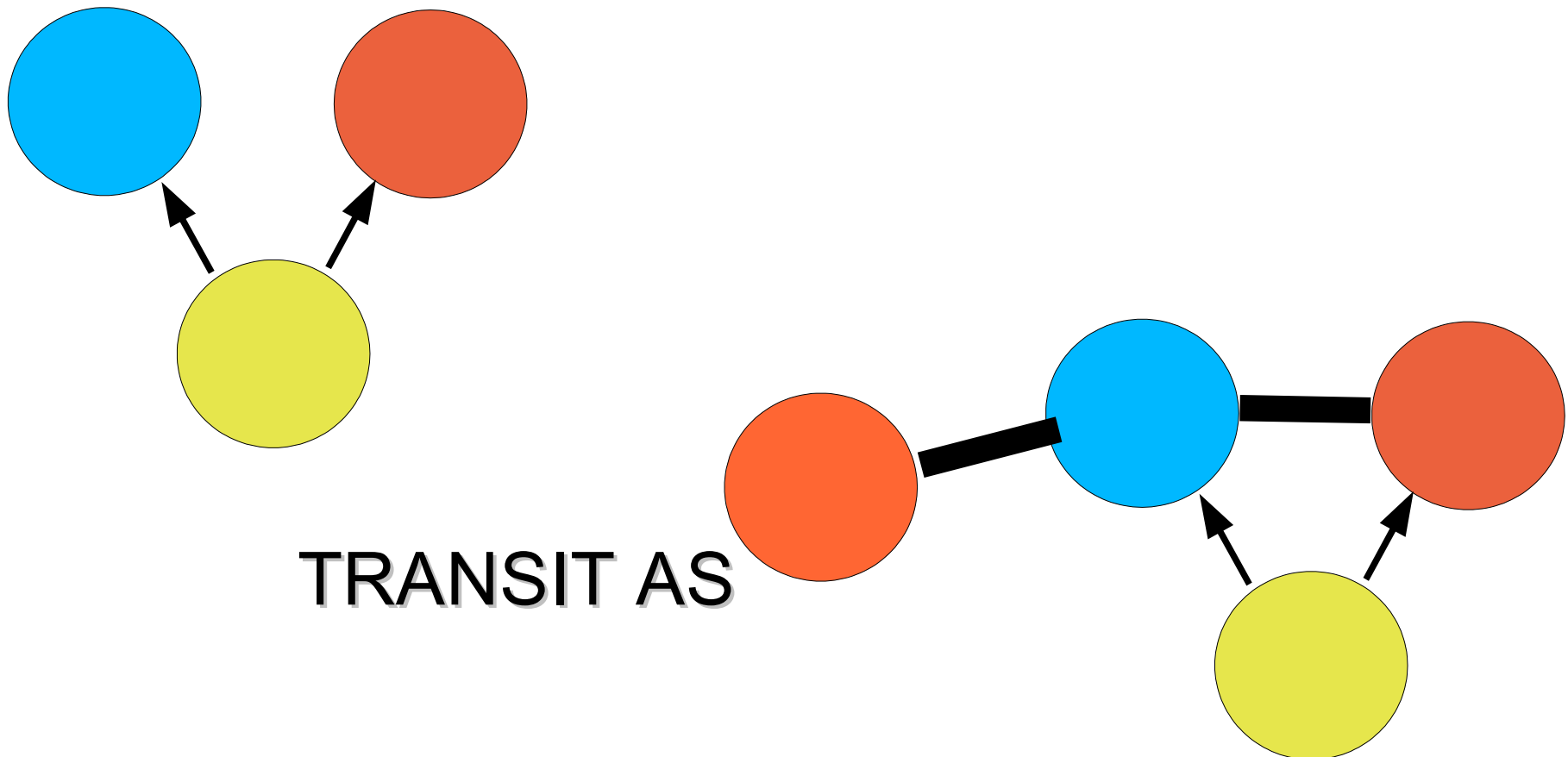
UPDATE MESSAGE – jest to typ pakietu najczęściej wymieniany. W nim znajduje się informacja o trasach dodawanych lub usuwanych oraz związanymi z nimi parametrami.

NOTIFICATION MESSAGE – przesyłany jest w sytuacji wystąpienia jakiegokolwiek błędu. Po wysłaniu tego typu pakietu połączenie BGP jest przerywane.

KEEPALIVE MESSAGE – wysyłane są kiedy sesja BGP jest zestawiona, mają one za zadanie podtrzymanie sesji BGP. Co 60 sekund przesyłany jest pakiet o wielkości 19 bajtów. Routery informują się za pomocą tych pakietów, że połączenie jest wciąż aktywne. W przypadku gdy router otrzyma pakiet UPDATE nie jest konieczne wysyłanie pakietu KEEPALIVE przez dany okres czasu.

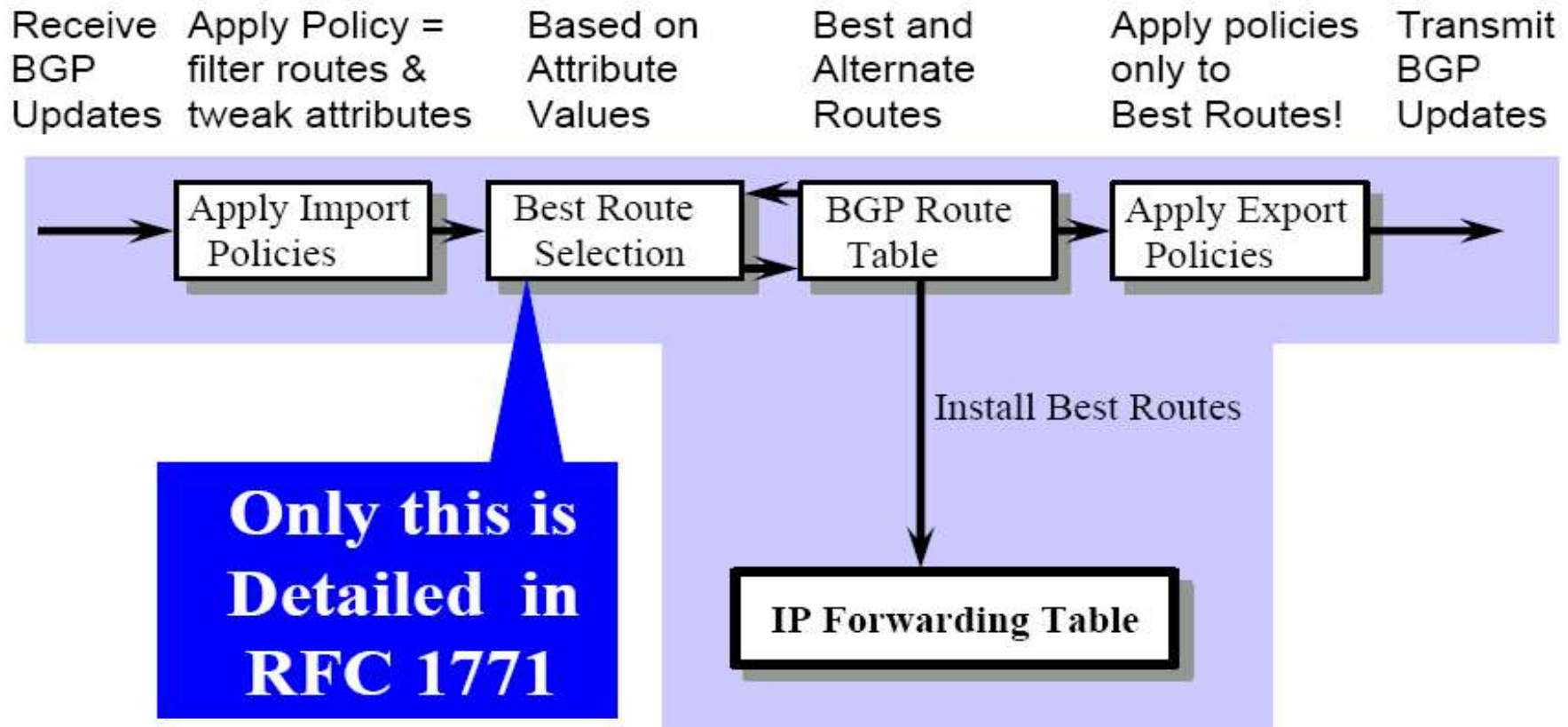
Najczęściej spotykane konfiguracje

MULTIHOMING - 2 i więcej dostawców



TRANSIT AS

BGP Route Processing



Atrybuty BGP

well-known mandatory

- as-path
- next-hop
- origin

well-known discretionary

- local preference
- atomic aggregate

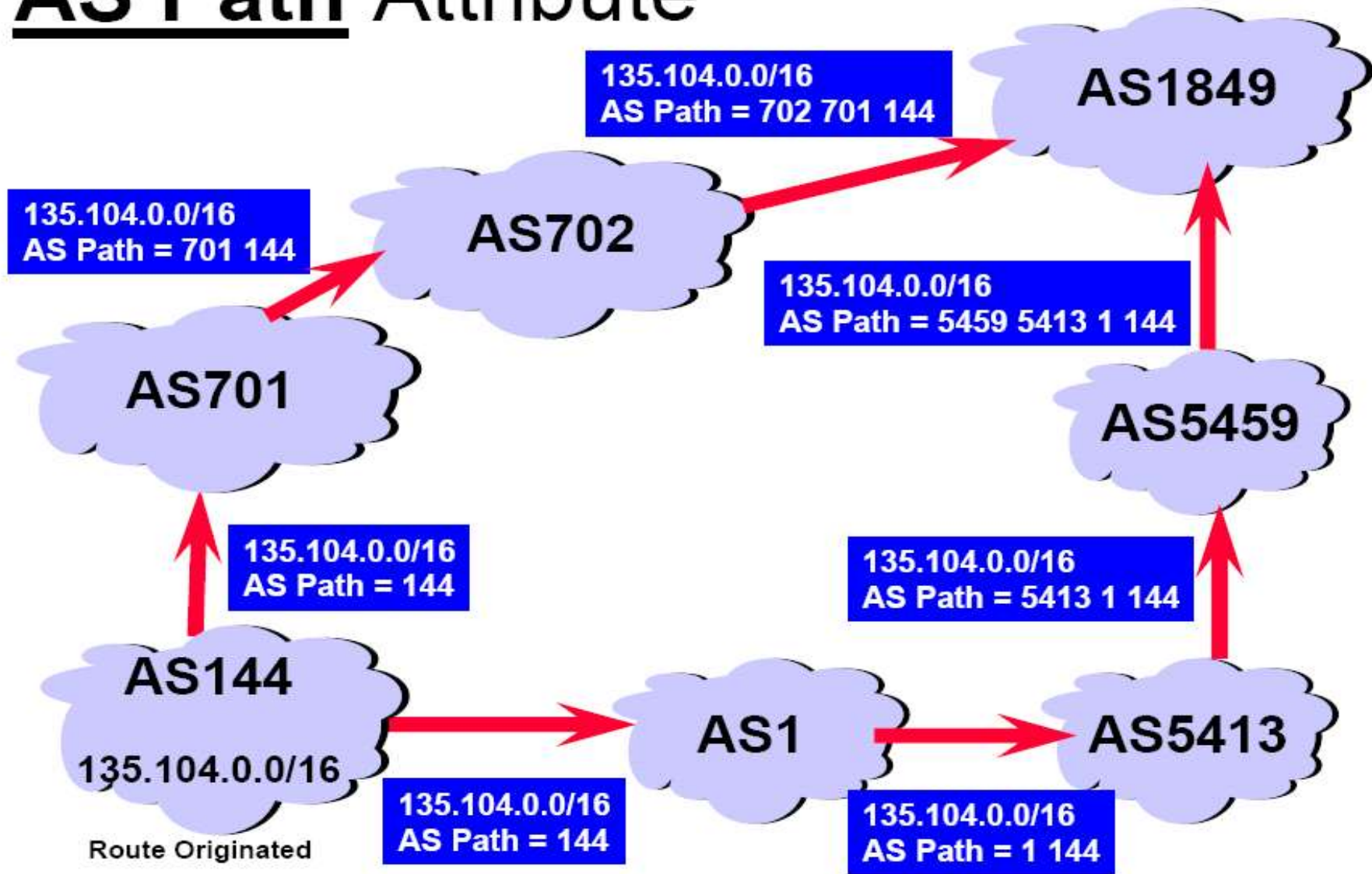
optional transitive

- aggregator
- communities

optional non-transitive

- MED – multi exit discriminator

AS Path Attribute



as_path w procesie BGP

```
router-bgpd# show ip bgp 161.12.12.0
BGP routing table entry for 161.12.12.0/22
Paths: (2 available, best #1, table Default-IP-Routing-Table)
```

```
Not advertised to any peer
```

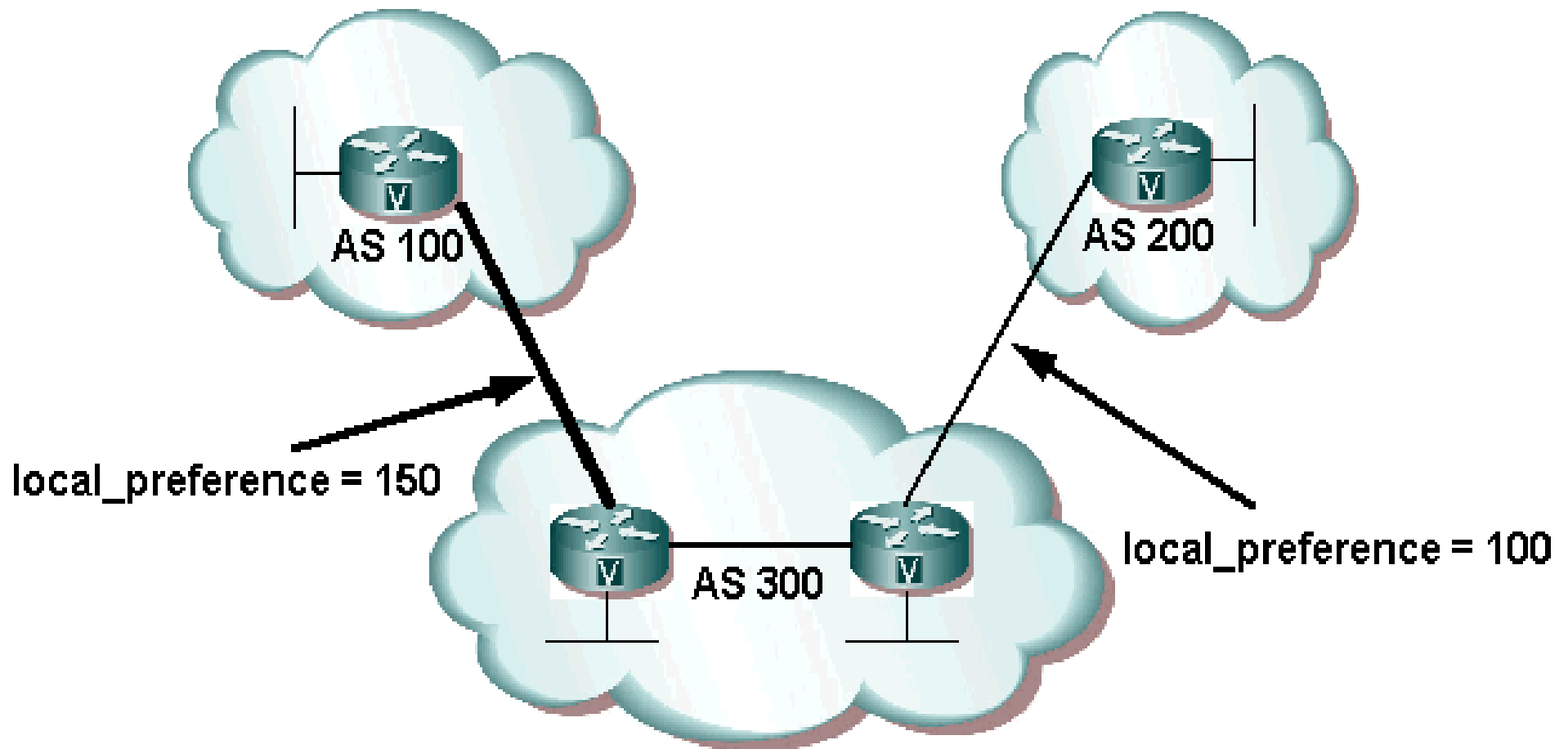
```
12968 3549 3356 <- AS-PATH
```

```
62.111.160.101 from 62.111.160.101 (213.134.144.1)
Origin IGP, localpref 100, valid, external, best
Last update: Tue Dec 30 17:24:48 2003
```

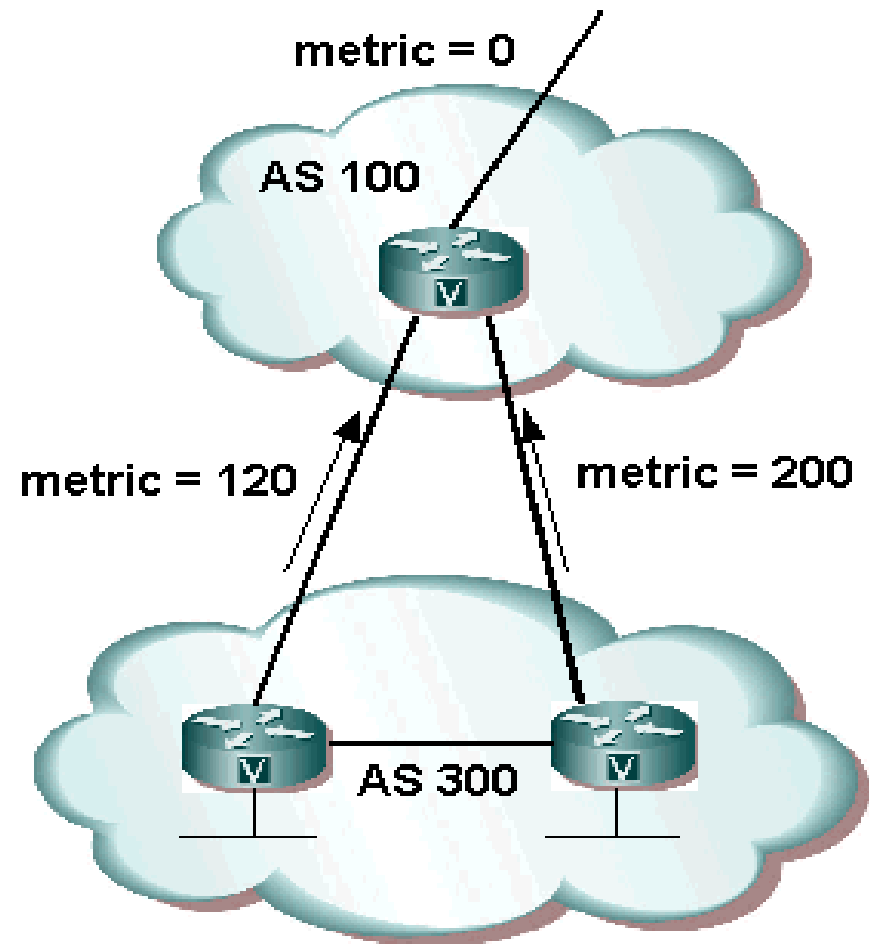
```
8246 5588 1239 3356 <- AS-PATH
```

```
157.25.1.16 from 157.25.1.16 (157.25.1.16)
Origin IGP, localpref 100, valid, external
Community: 5588:1001 5588:3001 8246:667 8246:1080
Last update: Tue Dec 23 07:24:15 2003
```

local_preference



MED



communities

- internet
- no-export
- no-advertise
- local-as

```
vix:~# whois as8246 | grep 8246: | more
remarks:      8246:666 Polish operators (TPNET+NASK+POL34)
remarks:      8246:667 Foreign operators (EBONE, SPRINT)
```

```
router-bgpd# show ip bgp community 8246:666
BGP table version is 0, local router ID is 195.149.118.1
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
*>	32.239.31.0/24	157.25.1.16		0	8246	5617 2686 ?
*>	32.239.135.0/24	157.25.1.16		0	8246	5617 2686 ?
*>	80.48.0.0/13	157.25.1.16		0	8246	5617 i
*>	80.249.0.0/20	157.25.1.16		0	8246	8308 24671 i

Wybór najlepszej ścieżki w BGP

Jeśli NextHop jest niedostępny przechodzimy do pkt 2

Wybierana jest ścieżka z najwyższym atrybutem WEIGHT

Jeśli WEIGHT są identyczne wybierana jest z najwyższym Local-Preference

Jeśli Local-Preference są równe wybierz trasę, która pochodzi z procesu BGP pracującego na tym routerze

Wybierz najkrótszy AS-PATH

Jeśli wszystkie ścieżki pochodzą spoza routera wybierz tę z najniższym origin type (IGP<EGP<INCOMPLETE)

Wybierz ścieżkę z najniższym MED.

Perferuj ścieżkę :”zewnętrzną nad „wewnętrzną”

Jeśli synchronizacja jest wyłączona, wybierz najbliższą ścieżkę pochodzącą z IGP

Wybierz trasę o najniższym adresie IP, który wskazuje na router id

peer-group

```
router bgp 31400
  neighbor wan-link peer-group
  neighbor wan-link filter-list 1 out
  neighbor 172.16.11.1 peer-group wan-link
  neighbor 172.16.11.1 remote-as 100
  neighbor 10.0.0.2 peer-group wan-link
  neighbor 10.0.0.2 remote-as 200
```

Filtrowanie

Filter list

```
router bgp 100
  neighbor 172.16.65.11 filter-list 11 in
  neighbor 172.16.65.11 filter-list 1 out
...
ip as-path access-list 1 permit ^200$
ip as-path access-list 1 deny ^300
ip as-path access-list 11 permit ^400
```

Distribute list

```
access-list 15 deny ip 172.15.0.0 0.0.255.255
access-list 15 permit ip 172.16.0.0 0.0.255.255

router bgp 31400
  neighbor 172.16.11.254 remote-as 31400
  neighbor 172.16.11.254 distribute-list 15 in
```

Filtrowanie cd...

Prefix list

```
router bgp 29620
...
  neighbor 157.25.1.16 prefix-list lista-10 in
...
ip prefix-list lista-10 seq 10 deny 0.0.0.0/0
ip prefix-list lista-10 seq 20 permit any
```

Community list

```
ip community-list 50 permit 8246:666
ip community-list 50 permit 8246:100
ip community-list 1 permit 100:1 100:2
```

ROUTE MAP

```
router bgp 29620
...
  neighbor 157.25.1.16 route-map localonly out
  neighbor 157.25.1.16 route-map gts in
...
ip as-path access-list 30 permit ^8246_
ip as-path access-list 40 permit ^12968$
ip as-path access-list 40 permit ^12968 [0-9]*$
ip as-path access-list 45 permit ^12968_
ip as-path access-list 50 permit ^$

route-map gts permit 10
  match as-path 30
  set local-preference 100

route-map localonly permit 10
  match as-path 50

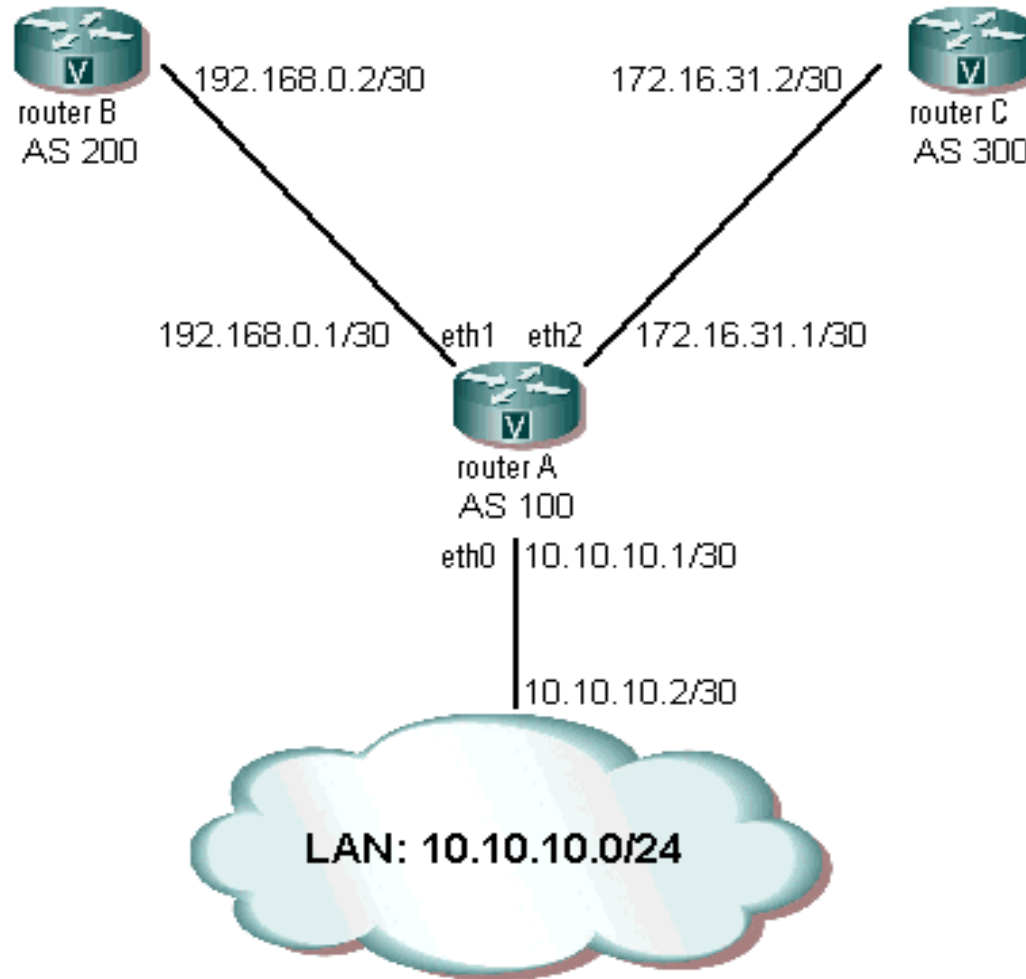
route-map cdp permit 10
  match as-path 40
  set local-preference 200
```

Prependowanie AS

```
router bgp 100
  network 172.16.0.0
  neighbor 10.0.0.1 remote-as 200
  neighbor 10.0.0.1 route-map SETPATH out

route-map SETPATH
  set as-path prepend 100 100
```


Przykładowy schemat w konfiguracji Multihoming



Konfiguracja bgpd (Quagga)

```
9  router bgp 100
10  network 10.10.10.0/24
11  neighbor 192.168.0.1 remote-as 200
12  neighbor 192.168.0.1 description "Link do ISP 1"
13  neighbor 192.168.0.1 prefix-list 10 in
14  neighbor 192.168.0.1 route-map localonly out
15  neighbor 172.16.31.1 remote-as 300
16  neighbor 172.16.31.1 description "Link do ISP 2"
17  neighbor 172.16.31.1 prefix-list 10 in
18  neighbor 172.16.31.1 route-map localonly out
19  !
20  ip prefix-list 10 seq 10 deny 0.0.0.0/0
21  !
22  ip as-path access-list 50 permit ^$
23  !
24  route-map localonly permit 10
25  match as-path 50
```

Konfiguracja zebra (Quagga)

```
1 hostname router-zebra
2 password 8 xxxxxx
3 enable password 8 xxxxxxxx
4 log file /var/log/quagga.log
5 log trap informational
6 log record-priority
7 service advanced-vty
8 service password-encryption
9 !
10 interface eth0
11 description "Siec wewnetrzna"
12 link-detect
13 ip address 10.0.0.1/24
14 !
15 interface eth1
16 description "ISP 1 – AS200"
17 link-detect
18 ip address 192.168.0.2/30
19 !
20 interface eth2
21 description "ISP 2 -AS300"
22 link-detect
23 ip address 172.16.31.2/30
24 !
25 interface lo
26 ip address 127.0.0.1/8
27 !
```

ethereal – czyli naga prawda o BGP

przykładowy pakiet BGP:

<http://www.poznan.linux.org.pl/wyklad/materialy/20040107/zrzutekranu1.png>

UPDATE msg z withdrawn routes:

<http://www.poznan.linux.org.pl/wyklad/materialy/20040107/zrzutekranu2.png>

Pakiet BGP z atrybutami:

<http://www.poznan.linux.org.pl/wyklad/materialy/20040107/zrzutekranu3.png>

NOTIFICATION msg:

<http://www.poznan.linux.org.pl/wyklad/materialy/20040107/zrzutekranu4.png>

Zestawianie sesji BGP TCP/IP

<http://www.poznan.linux.org.pl/wyklad/materialy/20040107/zrzutekranu5.png>

BGP okiem tcpdump

```
gamma:~# tcpdump -i eth2 port 179
```

```
tcpdump: listening on eth2
```

```
22:40:29.850573 host-ip101-160.crowley.pl.bgp > host-ip62-199.crowley.pl.58266: P
367:386(19) ack 1 win 16137: BGP (KEEPALIVE) [tos 0xc0] [ttl 1]
22:40:29.850622 host-ip62-199.crowley.pl.58266 > host-ip101-160.crowley.pl.bgp: . ack
386 win 20904 (DF)
22:40:32.962497 host-ip62-199.crowley.pl.58266 > host-ip101-160.crowley.pl.bgp: P 1:20
(19) ack 386 win 20904: BGP (KEEPALIVE) (DF)
22:40:33.165747 host-ip101-160.crowley.pl.bgp > host-ip62-199.crowley.pl.58266: . ack 20
win 16118 [tos 0xc0] [ttl 1]
22:40:42.119417 host-ip101-160.crowley.pl.bgp > host-ip62-199.crowley.pl.58266: P
386:441(55) ack 20 win 16118: BGP [|BGP UPDATE] [tos 0xc0] [ttl 1]
22:40:42.119461 host-ip62-199.crowley.pl.58266 > host-ip101-160.crowley.pl.bgp: . ack
441 win 20904 (DF)
22:41:12.487287 host-ip101-160.crowley.pl.bgp > host-ip62-199.crowley.pl.58266: P
754:785(31) ack 20 win 16118: BGP (UPDATE: (Withdrawn routes: 8 bytes)) [tos 0xc0] [ttl
1]
22:41:12.487332 host-ip62-199.crowley.pl.58266 > host-ip101-160.crowley.pl.bgp: . ack
785 win 20904 (DF)
22:41:12.498122 host-ip101-160.crowley.pl.bgp > host-ip62-199.crowley.pl.58266: .
785:1321(536) ack 20 win 16118: BGP [|BGP UPDATE] [tos 0xc0] [ttl 1]
22:41:12.498155 host-ip62-199.crowley.pl.58266 > host-ip101-160.crowley.pl.bgp: . ack
1321 win 20904 (DF)
22:41:12.502497 host-ip101-160.crowley.pl.bgp > host-ip62-199.crowley.pl.58266: P
1321:1360(39) ack 20 win 16118: BGP [tos 0xc0] [ttl 1]
22:41:12.502528 host-ip62-199.crowley.pl.58266 > host-ip101-160.crowley.pl.bgp: . ack
1360 win 20904 (DF)
```

Informacja o wskazanym prefixie

```
router-bgpd# show ip bgp 217.153.107.0
BGP routing table entry for 217.153.0.0/16
Paths: (2 available, best #1, table Default-IP-Routing-Table)
  Not advertised to any peer
  8246
    157.25.1.16 from 157.25.1.16 (157.25.1.16)
      Origin IGP, localpref 250, valid, external, best
      Last update: Mon Mar 29 05:26:28 2004

  12968 8246
    62.111.160.101 from 62.111.160.101 (213.134.144.1)
      Origin IGP, localpref 100, valid, external
      Last update: Sun Mar 28 21:52:08 2004

router-zebra# show ip route 217.153.107.0
Routing entry for 217.153.0.0/16
  Known via "bgp", distance 20, metric 0, best
  Last update 2d16h25m ago
  * 157.25.1.16 (recursive via 217.153.71.33)
```

Informacje o AS w procesie BGP

```
router-bgpd# show ip bgp regexp 5617$
BGP table version is 0, local router ID is 195.149.118.1
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
*> 63.167.185.0/24	62.111.160.101				0 12968 3549 1239
25617 i					
*> 80.48.0.0/13	157.25.1.16		150	0	8246 5617 i
*	62.111.160.101			0	12968 24748 5617 i
*> 83.0.0.0/11	157.25.1.16		150	0	8246 5617 i
*	62.111.160.101			0	12968 24748 5617 i

Informacja o wybranym AS

```
router-bgpd# show ip bgp regexp ^8246$
BGP table version is 0, local router ID is 195.149.118.1
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
*>	195.94.192.0/19	157.25.1.16			0	8246 i
*>	217.8.160.0/19	157.25.1.16			0	8246 i
*>	217.153.0.0/16	157.25.1.16			0	8246 i

```
Total number of prefixes 3
```


Podsumowanie informacji o zestawionych sesjach BGP

```
router-bgpd# show ip bgp summary
```

```
BGP router identifier 195.149.118.1, local AS number 29620
```

```
43688 BGP AS-PATH entries
```

```
322 BGP community entries
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down
State/PfxRcd								
62.111.160.101	4	12968	6575310	31966	0	0	0	4d17h27m
128107								
157.25.1.16	4	8246	1226765	31877	0	0	0	01w0d15h
129167								

```
Total number of neighbors 2
```

```
gamma:~# ip route | grep 217.153.71.33 | wc -l  
98072
```

```
gamma:~# ip route | grep 62.111.199.61 | wc -l  
31419
```

```
gamma:~# ip route | wc -l  
129492
```

PYTANIA?