

# Sieci Komputerowe

## Wykład 13 Dynamiczne protokoły rutowania

# System Autonomiczny

- AS (ang. Autonomous System) to sieć lub zbiór sieci zarządzanych przez jedną organizację (np. UW) i realizujący spójną politykę routingu
- AS jest identyfikowany za pomocą numeru 16 bitowego ASN (ang. Autonomous System Number)
  - 1-64511 – numery publiczne przydzielane przez RIR
  - 64512-65534 – prywatne, nieprzydzielane
  - 0 oraz 65535 zarezerwowane
- Numery AS są przydzielane przez RIR (dla Europy RIPE: <http://www.ripe.net>.)

# Rutowanie dynamiczne

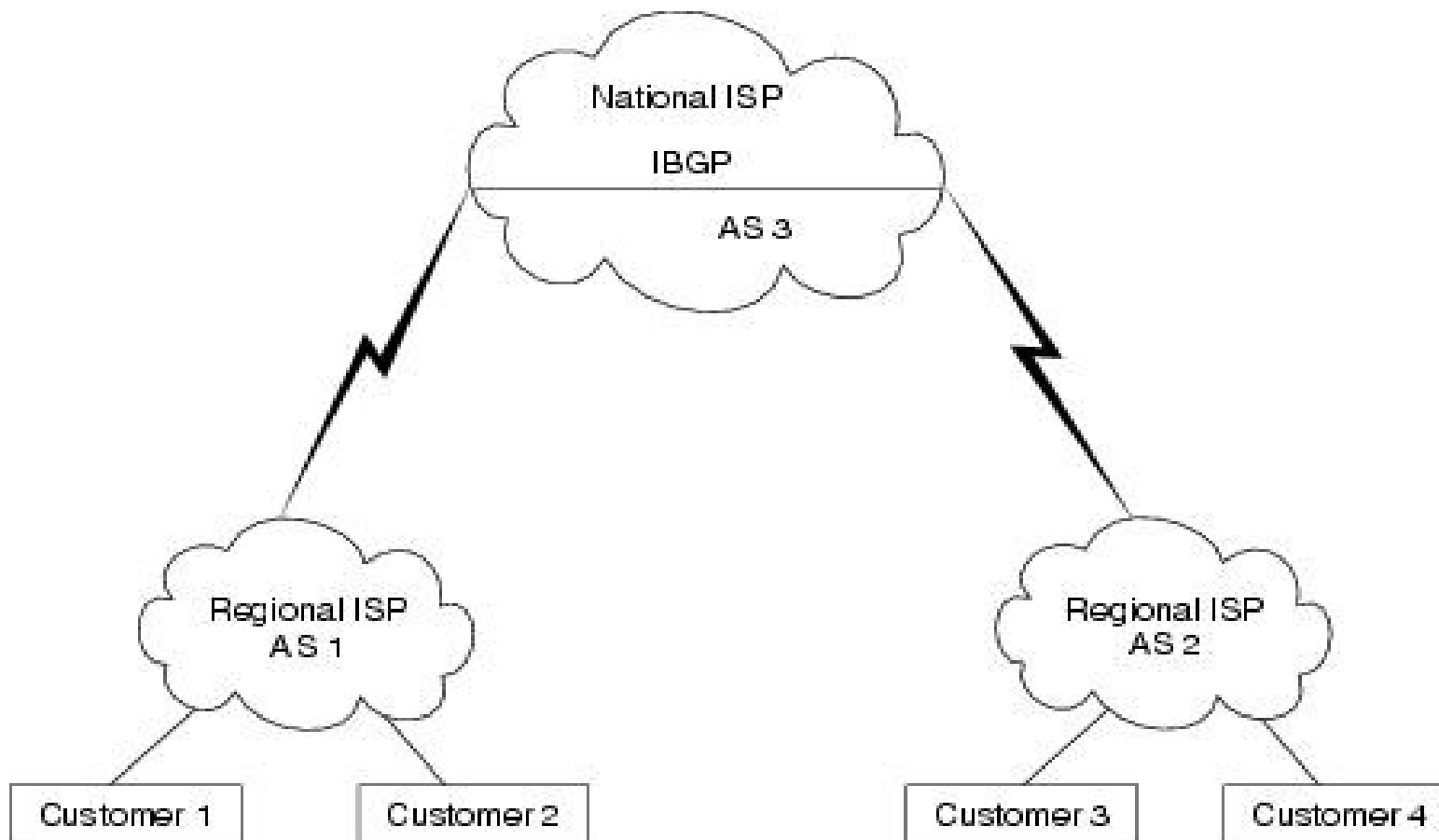
- Zachodzi wtedy, gdy routery informują się wzajemnie o dostępnych sieciach
  - komunikacja następuje za pomocą protokołów routowania
- Rutowanie dynamiczne nie zmienia obsługi procesu routowania w warstwie IP przez jądro systemu operacyjnego
  - Na routerze jest uruchamiany proces, który wprowadza wpisy do tablicy FIB automatycznie

# Protokoły rutowania

- IGP (Interior Gateway Protocol)
  - używany wewnątrz AS
    - OSPF (Open Shortest Path First) – obecnie szeroko stosowany
    - RIP (stary protokół...)
- EGP (Exterior Gateway Protocol)
  - zwykle używany do komunikacji ruterów z różnych AS
    - BGPv4 (Border Gateway Protocol) RFC 4271
    - TCP, port 179
  - Internet, to BGP :)

# IBGP i EBGP

- External i Interior BGP
  - sesje połączeń między routerami w ramach tego samego AS i różnych AS



# Sesja BGP

- BGP wymaga utrzymywania połączenia, routery z różnych AS powinny być połączone bezpośrednio  
tzw. sesja między peerami - peering
  - komunikat BGP typu open
    - wersja protokołu
    - lokalny ASN
    - tzw. hold-time
    - identyfikator BGP – adres IP interfejsu

# Sesja BGP

- komunikat update
  - dzięki niemu routery wysyłają informacje o dostępności podsieci (tzw. prefiksów) wraz z atrybutami
  - komunikat zawiera:
    - długość pola opisującego prefiksy niedostępne
    - niedostępne prefiksy
    - rozmiar pola opisującego atrybuty ścieżek
    - ścieżki (sekwencje numerów AS) wraz z atrybutami
    - dostępne prefiksy
- komunikat notification
  - informacja dla peera o błędzie
    - zawiera kod błędu, oznacza koniec sesji BGP
- komunikat keep-alive
  - podtrzymanie sesji, częstotliwość > hold-time

# Wymiana Informacji między AS

- Aby sieci w dwóch różnych AS mogły się komunikować:
  - Pierwszy AS musi rozgłosić swoje prefiksy do drugiego
  - Drugi musi zaakceptować rozgłaszane prefiksy
  - Drugi AS musi rozgłosić swoje prefiksy do pierwszego
  - Pierwszy musi zaakceptować rozgłaszane prefiksy
    - rozgłaszany prefiks to np. 193.0.96.0/24

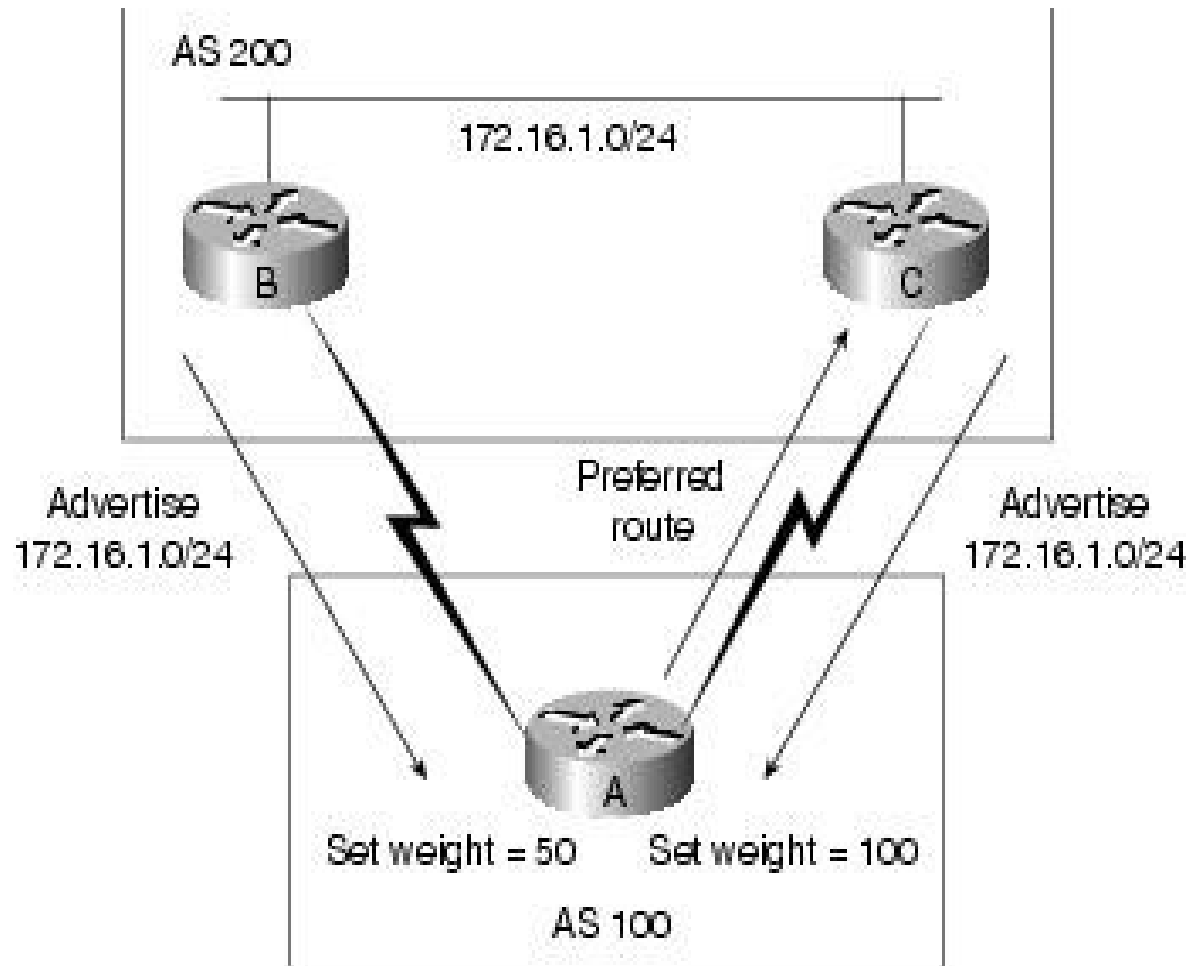


# BGP - atrybuty

- Atrybuty opisują charakterystykę podsieci (prefiksu) otrzymanego lub rozgłaszanego za pośrednictwem BGP.
  - Pozwalają na wybór optymalnej trasy, kontrolę rozgłaszanych informacji i mogą być określane przez administratora, tak aby realizować określoną politykę routowania

# Waga

- Atrybut lokalny dla rutera, nierozgłaszany nawet w ramach tego samego AS
  - wybierana jest trasa o większej wadze

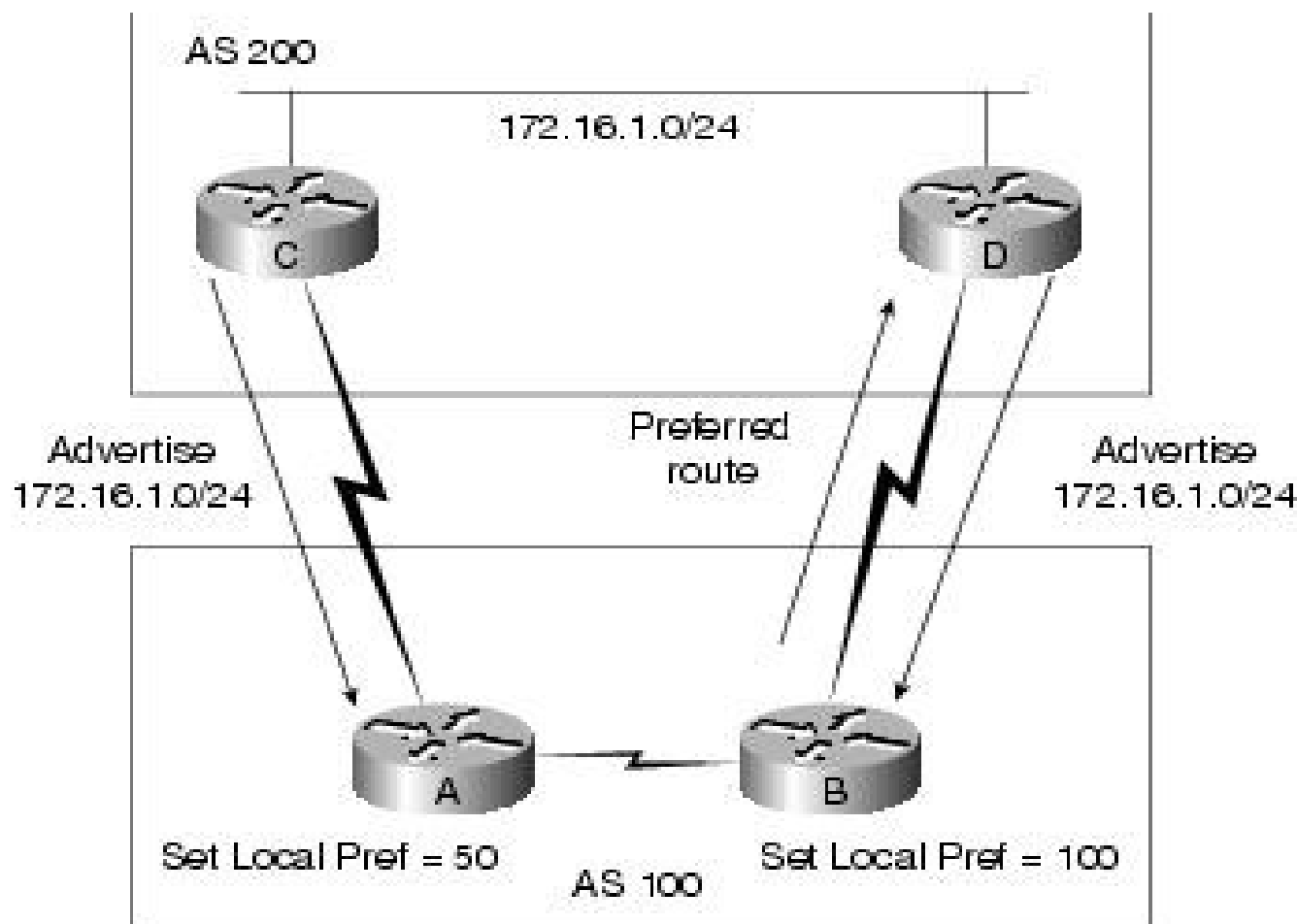


# Local Preference

- Podobnie jak waga, służy do określenia „punktu wyjścia” dla ruchu wychodzącego poza AS (jeśli jest więcej niż 1 peering)
- Wartość local preference jest rozgłaszana w ramach tego samego AS

# Local Preference c.d.

- Preferowana jest wyższa wartość więc ruch do 172.16.1.0/24 opuści AS 100 przez ruter B

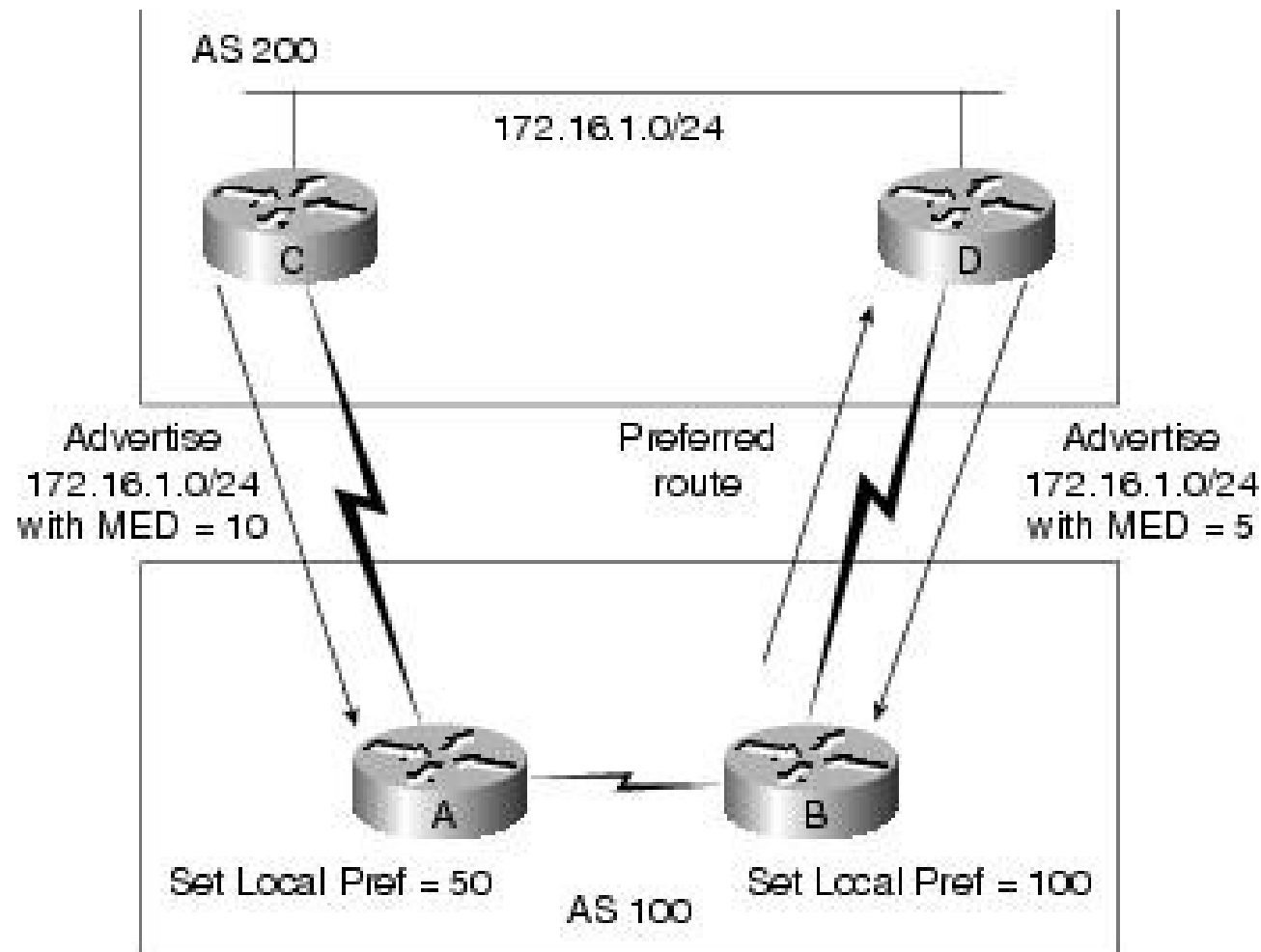


# MED

- Multi-Exit Discriminator (metric attribute)
  - sugestia wyboru trasy do nas dla zewnętrznego AS, do którego rozgłaszamy MED dla danego prefiksu
    - możemy wpływać na ruch przychodzący, o ile zewnętrzny AS nie korzysta z innych – własnych atrybutów dla danego prefiksu wpływających na trasę do naszego AS

# MED c.d.

- Preferowana jest trasa z mniejszą wartością metryki



# Atrybut Origin

- Ten atrybut wskazuje, jak BGP uzyskał informacje o trasie:
  - IGP
    - trasa jest trasą wewnętrzną dla danego AS
  - EGP
    - trasa została uzyskana w wyniku ogłaszania przez zewnętrzny AS
  - Incomplete
    - nie jest znane pochodzenie informacji o trasie

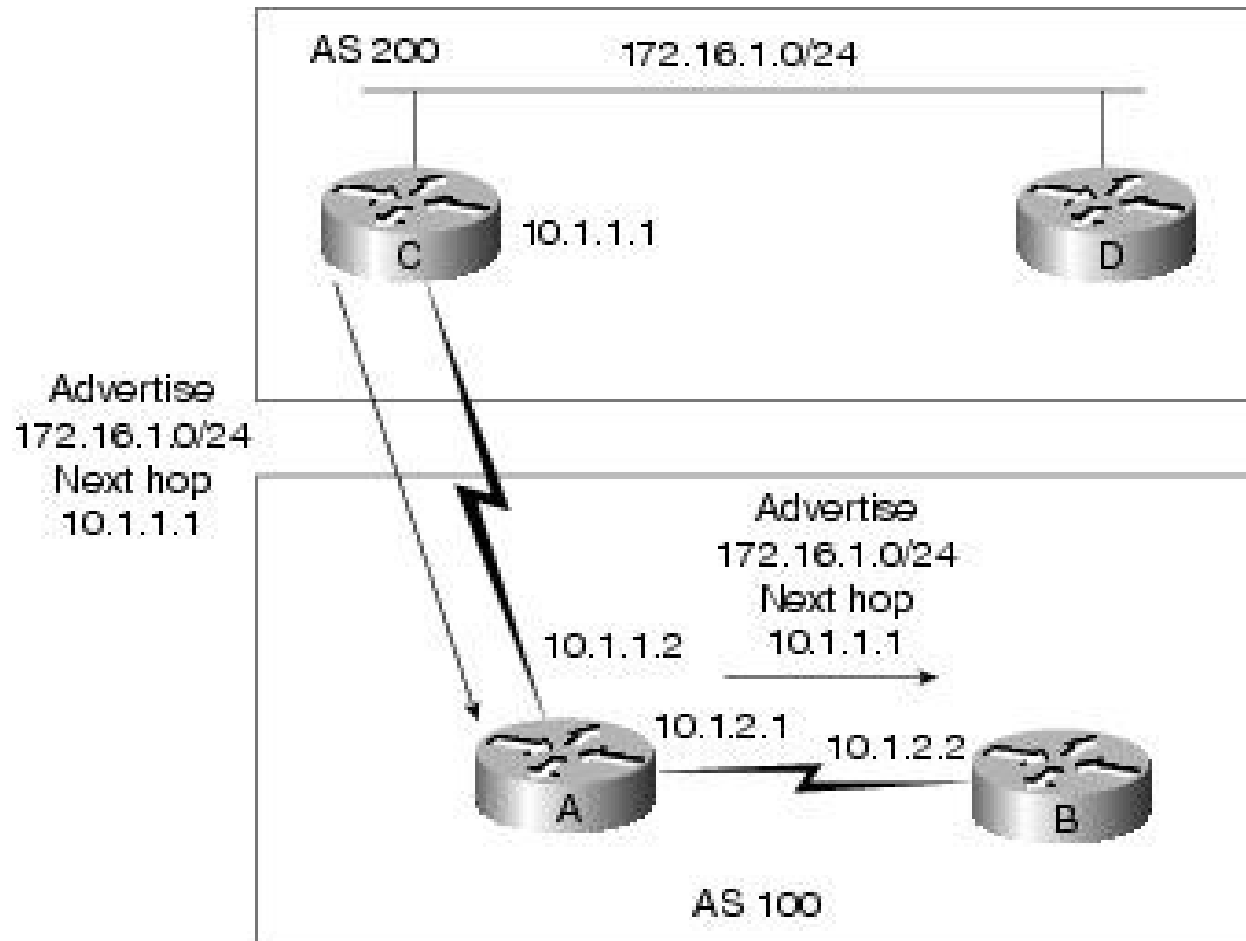
# Atrybut AS\_path

- AS\_path jest to sekwencja numerów AS.
  - Gdy ogłoszenie prefiksu przechodzi przez jakiś AS, ASN jest dodawany i powstaje lista – ścieżka
  - Zapobiega pętlom



# Next-Hop

- Next-hop to adres IP rutera przez który osiągalna jest rozgłaszana trasa

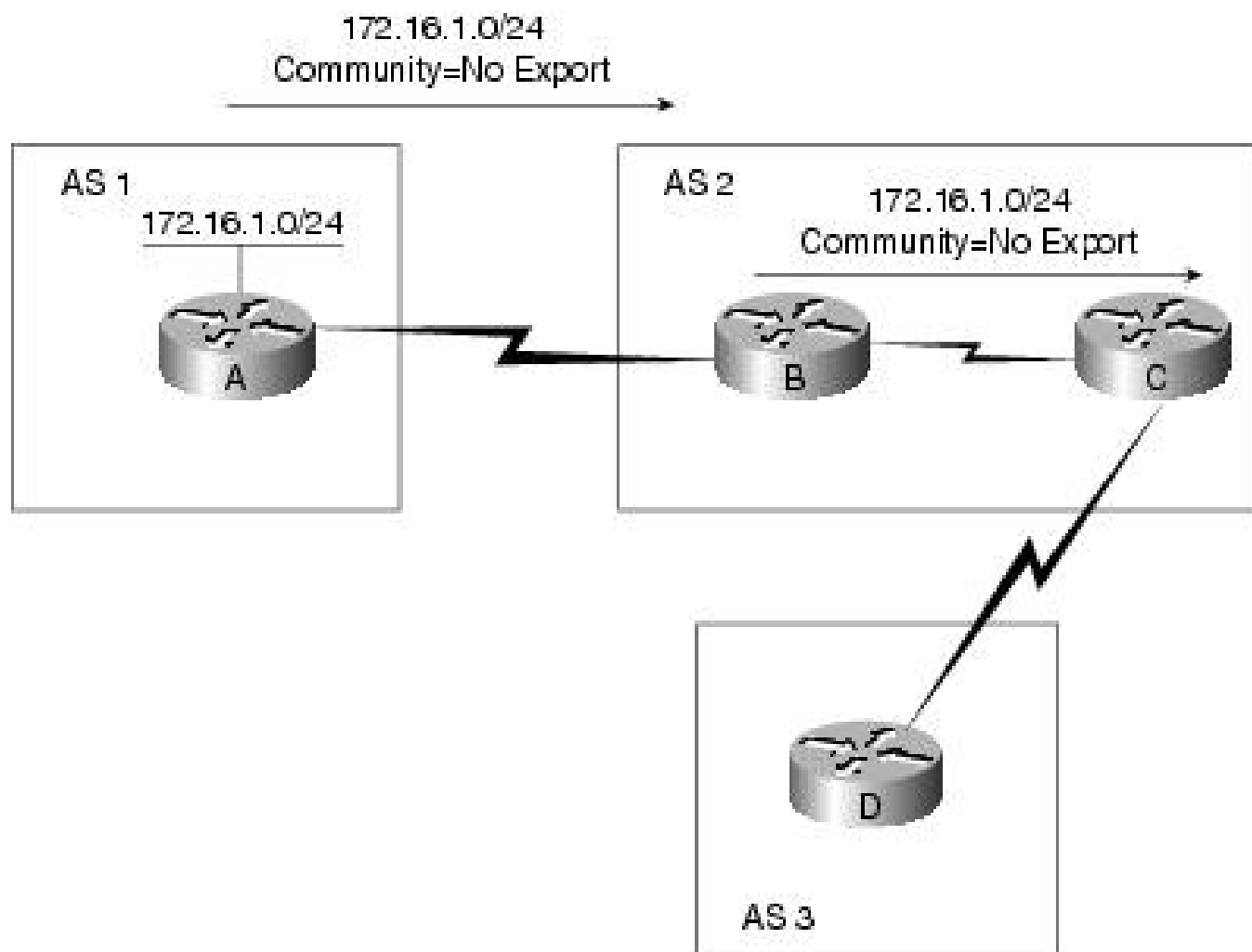


# Community

- Umożliwiają znakowanie tras dla których mogą być podejmowane pewne decyzje, np. zmiana local preference u peera
  - definiujemy znaczenie za pomocą liczby 32 bit
  - ISP publikują listy honorowanych community i na tej podstawie mogą zmieniać politykę rutingu
- Są także predefiniowane communities:
  - no-export – nie ogłaszaj tras do zewnętrznych AS
  - no-advertise – nie ogłaszaj w ogóle, do żadnego peera
  - internet – ogłaszaj wszystkim

# Community c.d.

- Trasa nie będzie wysłana poza AS 2



# Bardzo ważne: filtrowanie

- ip as-path access-list 10 permit .\*
  - .\* - cokolwiek
  - ^\$ - trasy lokalne dla danego AS
  - \_100\$ - trasy stworzone w AS 100
  - ^100\_ - trasy otrzymane z AS 100
  - \_100\_ - trasy przez AS 100
  - \_200\_100\_ - trasy przez AS100, następnie przez AS 200
- ip prefix-list cisco seq 10 permit 0.0.0.0/0 le 19
  - Dopuszcza trasy z maską co najmniej (less or equal) 19 bitów

# Podejmowanie decyzji o wyborze ścieżki

- BGP wybiera tylko jedną trasę, umieszcza ją w tablicy rutowania IP i (ewentualnie) rozgłasza
- Uwzględniane kryteria wyboru trasy w następującej kolejności:
  - Nie uwzględniaj ścieżki, dla której next-hop jest nieosiągalny
  - Preferuj trasę z większą wagą
  - Jeśli wagi są takie same, uwzględnij większą wartość local preference
  - Preferuj trasę z krótszym AS\_path
  - Jeśli wszystkie ścieżki mają taki sam AS\_path, uwzględnij tę z mniejszym atrybutem origin
    - IGP < EGP < incomplete

# Podejmowanie decyzji o wyborze ścieżki c.d.

- Jeśli atrybuty origin są takie same, uwzględnij najmniejszą wartość MED
- Jeśli wartość MED jest taka sama, preferuj ścieżkę zewnętrzną
- Jeśli ścieżki są takie same wybierz ścieżkę przez ruter bliższy w sensie IGP
- W końcu uwzględnij najmniejszą wartość adresu IP, odczytaną z komunikatu BGP

# Kiedy *nie* potrzebujemy BGP

- Jeden provider i jedno łącze
  - wystarczy default route...
- Jeden provider, dwa łącza z inną adresacją
  - wystarczy skonfigurować policy routing

# Kiedy BGP jest potrzebne

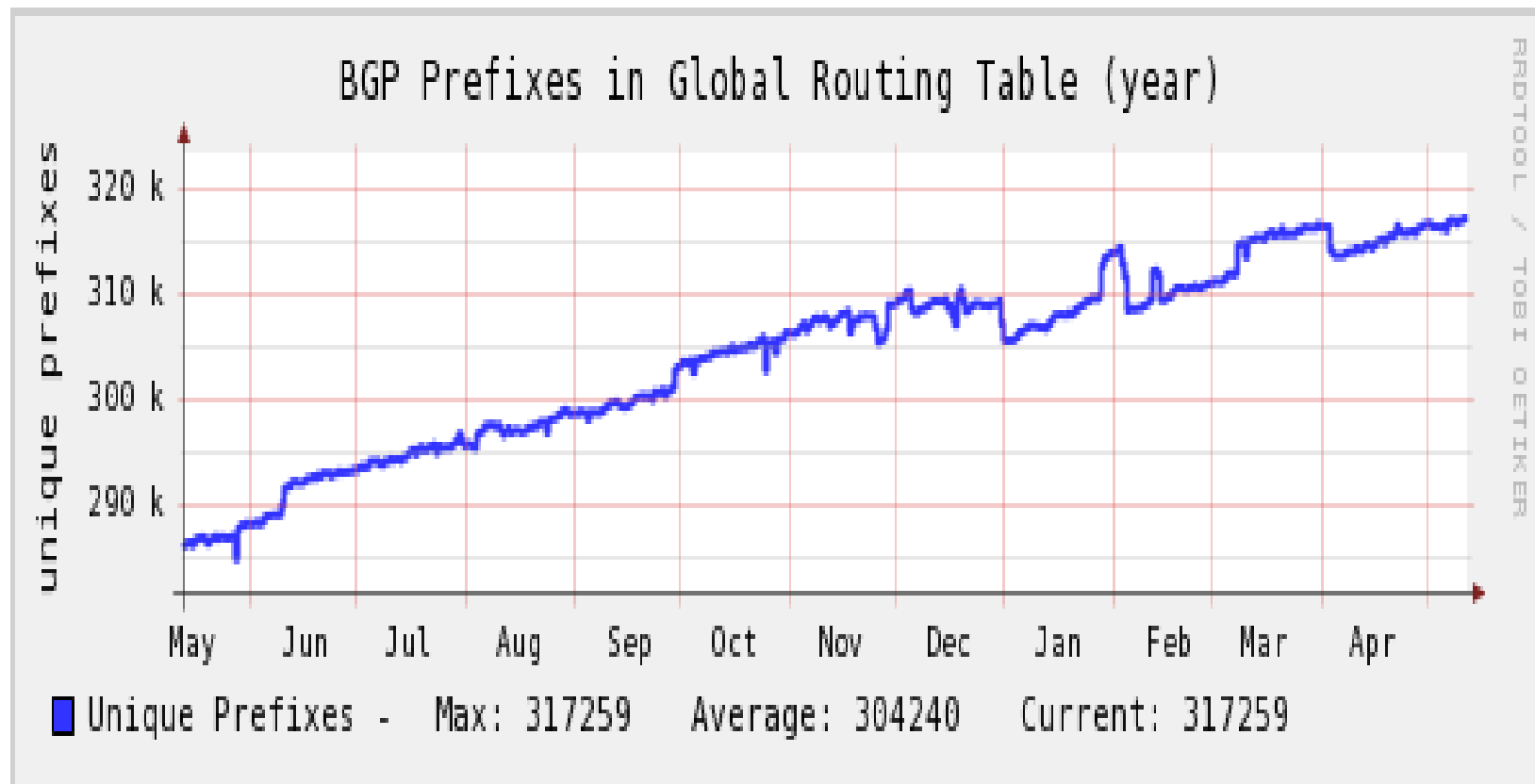
- 2 ISP
  - redundancja
  - możliwość stosowania własnej polityki routingu, rozkładanie obciążenia
- Tranzyt przez nasz AS do innych AS



# Problemy z BGP

- Rozgłaszanie prefiksów nieprzydzielonych lub należących do klas prywatnych: 1/8, 2/8, 10/8
  - stosowanie filtrowania (tzw. bogon filters)
    - spowodowały kiedyś problemy z neostradą – 83/11
    - ip prefix-list bogons seq 5 deny 1.0.0.0/8
  - należy uaktualniać filtrowaną listę prefiksów bogons:
    - <http://www.team-cymru.org/Services/Bogons/>

# Liczba prefixów w tablicach FIB



# BGP pod Linuksem

- Quagga (<http://www.quagga.net>)
  - wywodzi się z projektu zebra
  - implementuje bgp4, ospf, rip

# Monitorowanie

- Za pomocą SNMP:
  - cacti, mrtg (np. obciążenie łącza)
- Na różne sposoby:
  - nagios