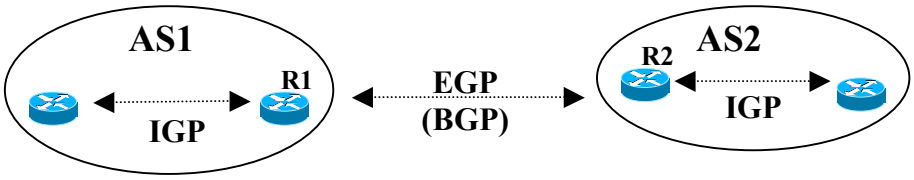
# Konfigurowanie protokołu BGP w systemie Linux

## 1. Wprowadzenie

**Wymagania wstępne: wykonanie ćwiczeń „Zaawansowana adresacja IP” oraz „Dynamiczny wybór trasy w ruterach Cisco”, znajomość pakietu Zebra.**

Internet tworzą połączone ze sobą sieci IP. Centralne zarządzanie siecią komputerową o globalnym rozmiarze jest technicznie niemożliwe, a ponadto niewskazane. Dlatego na najwyższym poziomie sieć Internet podzielona jest na niezależnie zarządzane, numerowane obszary, z których każdy obejmuje wiele sieci IP. Obszary te nazywa się *systemami autonomicznymi* (ang. Autonomous System, w skrócie AS).

Propagacja informacji o trasach odbywa się na dwa sposoby. Wewnątrz pojedynczego systemu autonomicznego jego trasy ogłasza protokół typu wewnętrznego, IGP (ang. Interior Gateway Protocol), na przykład RIP czy OSPF, zaś wymianą informacji o trasach pomiędzy systemami autonomicznymi zajmuje się protokół typu zewnętrznego, EGP (ang. Exterior Gateway Protocol). W chwili obecnej w sieci Internet stosowany jest tylko jeden protokół zewnętrzny, BGP (ang. Border Gateway Protocol), będący przedmiotem niniejszego ćwiczenia. Jego aktualna definicja (BGPv4) zawarta jest w dokumentach RFC1771 i RFC1772.




**Rysunek 1. Zakres stosowania protokołów IGP i EGP**

Rutery łączące różne systemy autonomiczne nazywa się ruterami brzegowymi (ang. border router) lub ruterami BGP. Na rysunku 1 role ruterów brzegowych pełnią R1 oraz R2.

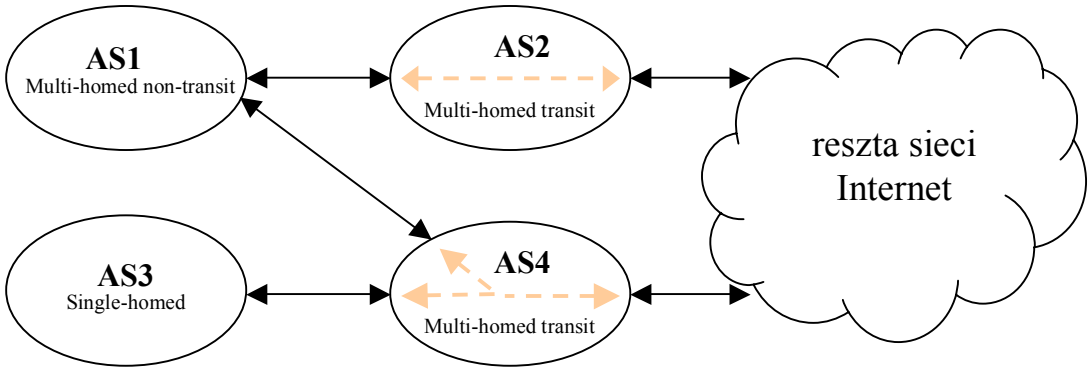
### 1.1 Systemy autonomiczne

Za administrowanie systemem autonomicznym odpowiada jedna instytucja; może nią być dostawca usług internetowych (ISP), duża uczelnia czy inna jednostka. Dla przykładu, w chwili obecnej w Polsce istnieje około 60 systemów autonomicznych.

Ze względu na ich wzajemne związki, wyróżnia się następujące rodzaje systemów autonomicznych:

- single-homed – systemy, które z resztą sieci Internet łączy tylko jeden inny system autonomiczny;
- multi-homed – systemy, które z resztą sieci Internet łączą co najmniej dwa inne systemy autonomiczne. Wśród nich istnieją dwie podgrupy: systemy, które przenoszą ruch sieciowy pomiędzy innymi systemami (multi-homed transit) i takie, które na to nie pozwalają (multi-homed non-transit).

Powyższe definicje ilustruje rysunek 2.




**Rysunek 2. Rodzaje systemów autonomicznych**

Ponadto, podobnie jak w przypadku sieci IP, istnieją publiczne (rejestrowane) i prywatne (nierejestrowane) systemy autonomiczne. Publiczny system otrzymuje numer z zakresu od 1 do 64511, a dla systemów prywatnych zarezerwowano numery od 64512 do największego, 65535.

## 1.2 Protokół BGP

Protokół BGP pozwala systemom autonomicznym wymieniać między sobą informacje o sieciach IP. Ściślej rzecz ujmując, protokół w wersji 4 służy do przesyłania prefiksów CIDR. W przypadku idealnym ruter brzegowy mógłby ogłaszać tylko jeden prefiks, będący adresem uogólnionym wszystkich sieci IP jego systemu autonomicznego.

BGP, jak każdy protokół wyboru trasy, wybiera najlepszą trasę. Jednak, w przeciwieństwie do protokołów wewnętrznych, nie stosuje przy tym metryki technicznej, lecz kilka parametrów administracyjnych, tzw. atrybutów (atrybuty zostały omówione w kolejnym podpunkcie). Ponadto, jednym z podstawowych elementów modelu BGP jest wykorzystanie podejścia nazywanego po angielsku „policy routing". Zakłada ono manipulowanie wysyłanymi (odbieranymi) ogłoszeniami o trasach po to, by akceptować (odrzucać) ruch sieciowy, którego źródłem (celem) jest konkretna sieć IP. Aby dana sieć w systemie autonomicznym A nie otrzymywała pakietów od systemu B, ruter brzegowy systemu A nie ogłasza jej systemowi B (ustala reguły dla ogłoszeń wysyłanych), i w drugą stronę – aby nie wysyłać pakietów do danej sieci przez system B, ruter systemu A nie przyjmuje ogłoszeń o niej od ruterów systemu B (ustala reguły dla ogłoszeń odbieranych).

Z technicznego punktu widzenia protokół BGP należy do kategorii odległość-kierunek (ang. distance-vector). W warstwie transportowej wykorzystuje protokół TCP wraz z portem 179. Po nawiązaniu połączenia TCP dwa rutery utrzymują ze sobą sesję BGP (ang. BGP peering relation). Istnieją dwa rodzaje sesji. Rutery należące do dwóch różnych systemów autonomicznych nawiązują ze sobą sesję zewnętrzną, EBGP (ang. external BGP). Wyróżnia się następujące stany sesji BGP: IDLE (ruter rozpoczyna nawiązywanie sesji), CONNECT, ACTIVE, OPEN SENT, OPEN CONFIRM i ESTABLISHED (sesja nawiązana).

W systemie tranzytowym istnieje wiele punktów styku z siecią Internet, dlatego jego rutery brzegowe muszą współdzielić informacje o trasach do innych systemów. W tym celu rutery brzegowe wewnątrz jednego systemu również nawiązują ze sobą sesje BGP; są to tzw. sesje wewnętrzne, IBGP (ang. internal BGP). Aby te współdzielone informacje były spójne, specyfikacja protokołu dodatkowo wymaga, by każdy ruter brzegowy danego systemu autonomicznego miał sesję IBGP z każdym innym ruterem brzegowym tego systemu. Wspomniane tu szczegóły to właściwie jedyne różnice między sesjami EBGP i IBGP. Rodzaj sesji jest w trakcie jej nawiązywania automatycznie wykrywany przez rutery. Zaraz po

ustanowieniu sesji rutery wymieniają ze sobą całe tablice tras, a późniejsze wymiany mają już charakter przyrostowy.

## 1.3 Atrybuty BGP

Jak już wspomniano, protokół BGP nie korzysta przy wyborze najlepszej trasy z metryk technicznych, lecz stosuje parametry administracyjne, nazywane atrybutami. Atrybuty są związane z konkretną siecią IP i przesyłane razem z informacją o niej. Wyróżnia się następujące kategorie atrybutów BGP:

- well-known mandatory – atrybuty, które muszą być rozpoznawane przez wszystkie implementacje protokołu (well-known) oraz muszą towarzyszyć każdej ogłaszanej trasie (mandatory);
- well-known discretionary – atrybuty, które są rozpoznawane przez wszystkie implementacje protokołu, ale nie muszą być przesyłane razem z ogłaszaną trasą (discretionary);
- optional transitive – atrybuty, które nie muszą być rozpoznawane przez implementację (optional), ale – jeśli towarzyszą ogłaszanej trasie - są wraz z nią przekazywane do innych ruterów;
- optional non-transitive – atrybuty, które ani nie muszą być rozpoznawane przez implementację, ani też przekazane do innych ruterów (non-transitive).

W chwili obecnej istnieje kilkanaście atrybutów BGP, a do najważniejszych z nich należą:

- ORIGIN (well-known mandatory) – określa pochodzenie informacji o danej trasie. Trzy wartości atrybutu są następujące: IGP (informacja pochodzi z tego systemu autonomicznego, od protokołu IGP), EGP (informacja pochodzi z innego systemu autonomicznego, od protokołu EGP), INCOMPLETE (inne źródło informacji, np. trasa statyczna redystrybuowana do protokołu).
- AS_PATH (well-known mandatory) – zbiór lub (najczęściej) sekwencja numerów systemów autonomicznych, przez które wiedzie trasa do danej sieci. Gdy ruter brzegowy (źródłowy lub kolejne) ogłasza sieć IP, wówczas umieszcza numer własnego systemu autonomicznego na początku tej sekwencji. Atrybut AS_PATH wykorzystywany jest do wykrywania pętli w trasach – jeśli ruter odczyta w sekwencji numer własnego systemu autonomicznego, stwierdza wystąpienie pętli i odrzuca ogłoszenie o trasie.
- NEXT_HOP (well-known mandatory) – określa adres IP rutera brzegowego z innego systemu autonomicznego, który powinien być użyty jako brama na trasie do sieci IP w tym systemie. Domyślnie, rutery brzegowe sąsiedniego systemu autonomicznego nie zmieniają wartości tego atrybutu. Dla przykładu, gdyby ruter R1 z rysunku 1 ogłosił pewną sieć systemu AS1 ruterowi R2 w systemie AS2, wówczas R2 i wszystkie inne rutery brzegowe w systemie AS2 jako wartość atrybutu NEXT_HOP odczytają adres IP rutera R1. Oznacza to, że **aby ustalić trasę do sieci w innym systemie autonomicznym, należy dwa razy odczytać informacje o trasach** – wynikiem pierwszego odczytu, z tablicy tras protokołu BGP, będzie wartość atrybutu NEXT_HOP, a dopiero drugi odczyt (ang. **recursive table lookup**), z podstawowej tablicy tras, pozwala znaleźć trasę do sieci, w której znajduje się ruter określany wartością tego atrybutu.
- LOCAL_PREF (well-known, discretionary) – atrybut ten przypisywany jest ruterowi. Jego wartość przesyła się tylko w obrębie jednego systemu autonomicznego. Atrybut służy ruterom brzegowym tego systemu do wybrania spośród siebie rutera pełniącego funkcję bramy do sieci IP w innym systemie autonomicznym. Wartością atrybutu jest liczba naturalna, a preferencja jest rosnąca. Innymi słowy, ruter z największą wartością LOCAL_PREF, związaną z daną siecią IP, zostanie bramą dla pozostałych ruterów BGP w jego systemie autonomicznym.

Ponieważ protokół BGP jest bardzo złożony, umiejętność jego wykorzystania wymaga od administratora sieci bogatego doświadczenia. Studenci zainteresowani pracą u dostawcy usług internetowych na pewno skorzystają z lektury pozycji wymienionych w punkcie 5.

### 1.4 Przydatne polecenia konfiguracyjne

Poniżej wymienione zostały podstawowe komendy konfiguracyjne dla protokołu BGP:
**router bgp** <nr AS> - uruchama protokół BGP i przenosi w tryb konfiguracji BGP;
**show ip bgp** – wyświetla trasy protokołu BGP wraz z wartościami niektórych atrybutów;
**show ip bgp neighbor** – wyświetla informacje o sąsiadach, czyli ruterach, z którymi dany ruter nawiązał sesje BGP;
**clear ip bgp *** - restartuje sesję BGP, powodując zastosowanie wszystkich wcześniej zdefiniowanych reguł;
**neighbor** <adres IP sąsiada> **remote-as** <nr AS sąsiada> - powoduje nawiązanie sesji BGP z sąsiadem w innym (EBGP) lub tym samym (IBGP) systemie autonomicznym;
**network** <adres IP sieci> - określa sieć IP, jaką ruter powinien ogłaszać. Chodzi tu również o sieci IP, w których ruter sam się nie znajduje, a o których dowiedział się za pośrednictwem protokołów IGP;
**access-list** – polecenie tworzy listę dostępu (ACL, ang. Access Control List), która w kontekście BGP pomocna jest przy tworzeniu reguł dla ogłoszeń protokołu (policy routing). Reguły wykorzystujące listy ACL najczęściej usuwają informacje o wybranych sieciach - określanych właśnie przez listę - z odbieranych lub wysyłanych ogłoszeń. Przykłady użycia list ACL zawarto w dodatku;
**neighbor...distribute-list** <nr listy ACL> [**in**|**out**] – polecenie tworzy regułę dla odbieranych (in) lub wysyłanych (out) przez protokół ogłoszeń;
**route-map** <nazwa> [**permit**|**deny**] <nr sekw.> – służy do manipulowania atrybutami BGP oraz do tworzenia reguł dla ogłoszeń;
**neighbor...route-map** – powoduje zastosowanie wcześniej zdefiniowanych atrybutów BGP (lub reguł) dla konkretnego sąsiada.

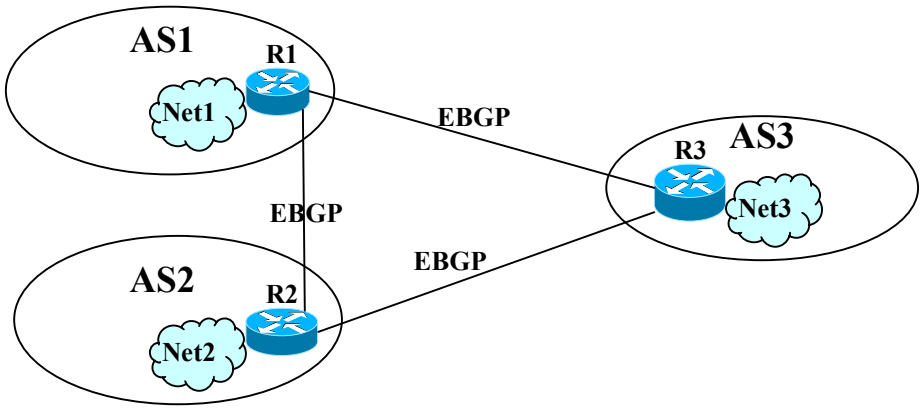
## 2. Organizacja, wymagany sprzęt i oprogramowanie

- zadanie wykonywane jest przez grupę 2- lub 3-osobową;
- sprzęt: 3 komputery PC;
- oprogramowanie: system Linux z zainstalowanym pakietem Zebra.

## 3. Zadania

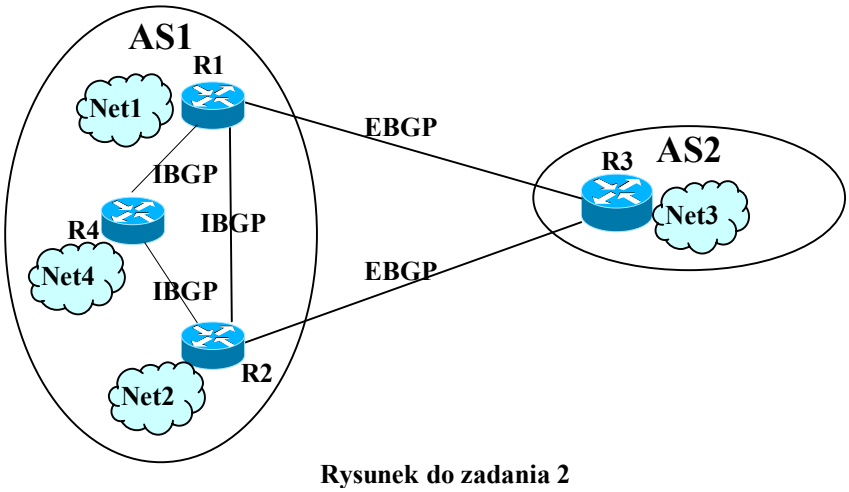
1. Skonfigurować środowisko sieciowe ukazane poniżej. Następnie wykonać zadania:
a) Wyświetlić trasy protokołu BGP. Odczytać wartości atrybutów NEXT_HOP, LOCAL_PREF i AS_PATH oraz sprawdzić, ile protokół pamięta tras do każdej sieci IP.
b) Zasymulować w ruterze R1 awarię interfejsu łączącego R1 z R3 i po odczekaniu jednej minuty (domyślny okres między komunikatami update) sprawdzić, jak wpłynie to na zmianę trasy z R1 do sieci Net3. Następnie wycofać wprowadzone zmiany, przywracając konfigurację pierwotną.

c) Konfigurując tylko ruter R3 spowodować, by trasa z rutera R1 do sieci Net3 wiodła przez system autonomiczny AS2 (policy routing). Wykorzystać polecenia: **`access-list`** (do określenia sieci Net3) oraz **`neighbor...distribute-list`** (do usunięcia sieci Net3 z ogłoszeń wysyłanych do rutera R1). Uwaga – po wprowadzeniu powyższych komend, trzeba odnowić sesje BGP przez wykonanie w ruterze R3 polecenia **`clear ip bgp *`**.




**Rysunek do zadania 1**

2. Skonfigurować środowisko sieciowe ukazane poniżej. Następnie wykonać zadania:
a) Odczytać w ruterach R1 i R2 wartość atrybutu NEXT_HOP dla wszystkich tras do sieci Net3. Jaka jest jego wartość dla trasy nie uznanej za najlepszą? Czy dla tej trasy istnieje konieczność rekurencyjnego odczytu informacji o trasach (reverse table lookup; patrz omówienie atrybutu NEXT_HOP)?
Odczytać domyślną wartość atrybutu LOCAL_PREF w ruterach R1, R2 i R3.
b) Wykorzystując atrybut LOCAL_PREF i konfigurując ruter R1 spowodować, by trasa z rutera R3 do sieci Net3 w systemie AS2 wiodła przez ruter R1. Wykorzystać polecenia: **`access-list`** (do zdefiniowania sieci Net3), **`route-map`** (tu podpolecenia **`match ip address`** do powiązania mapy z siecią Net3 oraz **`set`** do określenia nowej wartości atrybutu LOCAL_PREF) oraz **`neighbor...route-map`** (do ogłoszenia ruterom R2 i R3 nowej wartości atrybutu dla sieci Net3). Uwaga – po wprowadzeniu powyższych komend, trzeba odnowić sesje BGP przez wykonanie w ruterze R1 polecenia **`clear ip bgp *`**.




**Rysunek do zadania 2**

## 4. Pytania sprawdzające

1. Jak przebiega proces wyboru najlepszej trasy w protokole BGP?
2. W jaki sposób rutery automatycznie wykrywają rodzaj sesji BGP podczas jej ustanawiania?
3. Jaka jest przeciętna liczba tras w tablicy szkieletowego rutera BGP? Jak szybko ta liczba wzrasta? Wskazówka – przeczytać wybrany raport BGP, opublikowany w sieci Internet.
4. Jakie znaczenie dla wydajności pracy ruterów szkieletowych miało wprowadzenie wersji 4 protokołu BGP?
5. Co oznaczają pojęcia Route-Reflector oraz BGP Confederation?

## 4. Literatura


1. Dokumenty RFC1771 i RFC1772.
2. Książka Alberto Leon-Garcia, Indra Widjaja „Communication Networks – Fundamental Concepts and Key Architectures".
3. Serwis internetowy "BGP – the Border Gateway Protocol, Advanced Internet Routing Resources" http://www.bgp4.as
4. Omówienie atrybutów BGP: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm
5. Konfiguracja ruterów Cisco: serwis internetowy www.cisco.com
6. Dokumentacja pakietu zebra: www.zebra.org

# DODATEK – użycie list ACL

Składnia:
**access-list** <nr listy> **permit**/**deny** <adres IP> <negacja maski>

Negacja maski służy temu samemu celowi, co tradycyjna maska sieci. W wypadku maski zanegowanej, do adresu określonego w definicji listy ACL dopasowane zostaną te bity adresu źródłowego/docelowego pakietu, którym odpowiadają zera w zanegowanej masce.

Przykłady:
**access-list** 1 **permit** 192.168.1.0 0.0.0.255
Akceptacja pakietów pochodzących z sieci IP o adresie 192.168.1.0/24.

**access-list** 1 **deny** 150.254.17.96 0.0.0.31
Odrzucenie pakietów, których źródłem jest sieć IP o adresie 150.254.17.96/27.

W kontekście protokołów wyboru trasy listy dostępu służą do filtrowania nie pakietów, a informacji o trasach zawartych w ogłoszeniach protokołu.

Przykładowo, aby w protokole BGP zrealizować regułę dla ogłoszeń zakładającą istnienie komunikacji z siecią 192.168.1.0 i brak komunikacji z siecią 192.168.2.0, należy zdefiniować następującą listę ACL:
**access-list** 1 **permit** 192.168.1.0 0.0.0.255
**access-list** 1 **deny** 192.168.2.0 0.0.0.255,

a następnie zastosować ją do filtrowania tych sieci z ogłoszeń BGP odbieranych od sąsiadów:
**neighbor** <adres sąsiada BGP> **distribute-list** 1 **in.**