

BGP - teoria i praktyka

wprowadzenie

Spis Treści

SPIS TREŚCI	2
SPIS ILUSTRACJI	4
BORDER GATEWAY PROTOCOL (BGP)	5
TROCĘ HISTORII.....	5
BGP v4.....	6
JAK TO DZIAŁA – CZYLI WIEDZA TAJEMNA	7
ATRYBUT AS_PATH.....	8
ATRYBUT ORIGIN.....	9
BGP W SIECI TCP/IP	9
REDYSTRYBUCJA TRAS	10
ATRYBUTY BGP	11
LOCAL PREFERENCE.....	12
MED.....	13
WYBÓR NAJLEPSZEJ TRASY	14
PEER GROUP	16
COMMUNITIES	16
FILTROWANIE	18
FILTER LIST.....	18
DISTRIBUTE LIST.....	18
PREFIX LIST.....	19
COMMUNITY LIST.....	20
ROUTE MAPS	20
PREPENDOWANIE AS-PATH	21
ROUTE REFLECTORS – ZBYT DUŻO INFORMACJI	21
KONTROLA DZIAŁANIA	24
PRZYKŁADOWE KONFIGURACJE	27
SCENARIUSZ 1.....	27
<i>Schemat 1</i>	27
<i>Konfiguracja routerów</i>	27
SCENARIUSZ 2.....	28
<i>Konfiguracja routerów</i>	28
SCENARIUSZ 3	29
.....	30
<i>Konfiguracja routerów</i>	30
SCENARIUSZ 4	31
QUAGGA	33
TO TYLKO WSTĘP DO BGP...	34
SŁOWNIK	35
BIBLIOGRAFIA	36

KSIĄZKI.....	36
INTERNET.....	36

Spis ilustracji

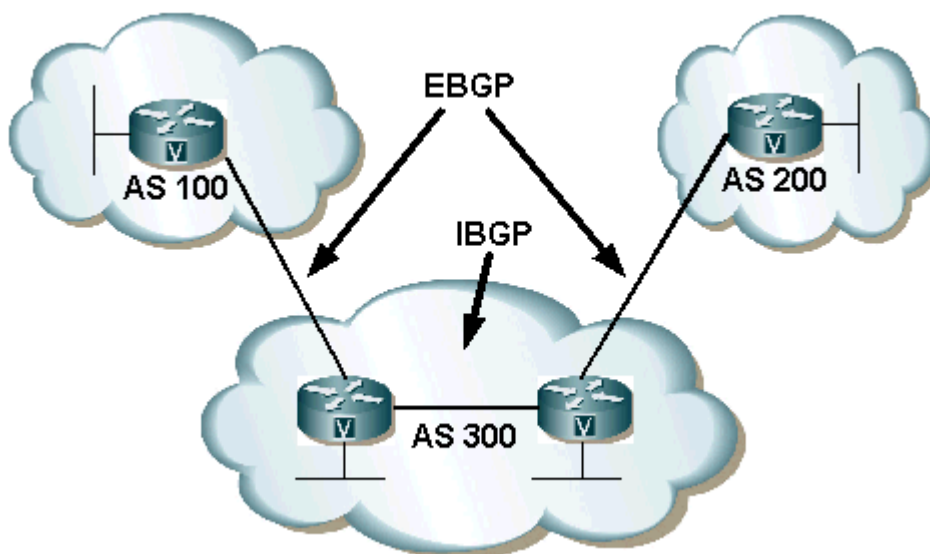
RYSUNEK 1 IBGP VS EBGP.....	5
RYSUNEK 2 BGP POMIĘDZY TRZEMA ODRĘBNYMI SIECIAMI.....	7
RYSUNEK 3 WYKORZYSTANIE LOCAL_PREFERENCE DO WSKAZANIA WYCHODZĄCEJ TRASY.....	13
RYSUNEK 4 ZA POMOCĄ MED MOŻEMY OKREŚLAĆ PREFEROWANE WEJŚCIE DO NASZEGO AS.....	14
RYSUNEK 5 LICZBA POŁĄCZEŃ W SIECI Z BGP ROŚNIE LAWINOWO Z KAŻDYM NOWYM ROUTEREM.....	22
RYSUNEK 6 UŻYCIE ROUTE REFLECTORÓW ZNACZNIE ZMNIEJSZA ILOŚĆ KONICZNYCH DO ZESTAWIENIA SESJI BGP.....	23

Border gateway protocol (BGP)

Tekst ten ma na celu przedstawienie podstawowych cech protokołu BGP i możliwości jego wykorzystania w pracy administratora sieci.

Wprowadzenie

Border Gateway Protocol czyli BGP, jest obecnie podstawowym protokołem za pomocą którego Internet wymienia pomiędzy poszczególnymi sieciami informację o ich dostępności. Bazuje na protokole EGP opisanym w RFC 904, wykorzystywanym w backbonie sieci NSFNET¹. BGP pozwala w znacznym stopniu administratorom sieci na spokojniejszy sen, zapewniając dynamiczną aktualizację informacji o routingu odpowiadającej bieżącemu stanowi sieci.



Rysunek 1 iBGP a eBGP.

Działanie BGP można podzielić na procesy w obrębie jednego ASN oraz różnych ASN, odpowiednio: iBGP i eBGP. iBGP to proces BGP działający w obrębie jednego systemu autonomicznego. eBGP pozwala na połączenie routerów znajdujących się w osobnych AS. To właśnie takie połączenia wymieniają informacje pomiędzy przylegającymi do siebie sieciami i dzięki eBGP każda z tych sieci wie jak dotrzeć do dowolnego zakresu adresów rozgłaszanych w internecie. Routery wymieniające informację za pomocą BGP nie muszą być bezpośrednio połączone, wystarczy, że będzie istniało między nimi połączenie logiczne.

W przypadku kiedy potrzebny jest nam protokół dynamicznego routingu wewnątrz sieci korzysta się najczęściej z bardziej odpowiednich rozwiązań takich jak OSPF, EIGRP, RIP czy mniej popularny IS-IS.

BGP w wersji 1 został opisany w RFC1105, wersja druga protokołu w RFC1163 i kolejno trzecia w RFC 1267. Wersja 4, która jest obecnym standardem używanym w Internecie opisana jest w RFC 1771 z roku 1995. Obecnie prowadzone są intensywne prace nad S-BGP czyli Secure BGP. BGP z wersji pierwszej do czwartej zyskiwało

¹ NSFNET – National Science Foundation Network, we wczesnych latach 90tych, backbone Internetu. RFC 1092, RFC 1093

znacznie na funkcjonalności. Zwiększył się rozmiar pojedynczej wiadomości BGP po wersji 1, wprowadzone zostały atrybuty za pomocą których można w BGP przenosić dodatkowe informacje. Atrybuty te były stopniowo rozbudowywane z wersji na wersję aby doprowadzić do obecnie wykorzystywanej wersji czwartej.

BGP v4

BGP jak każdy protokół rutowania, w dużym uproszczeniu, odpowiada za wymianę informacji o zakresach adresów, które dostępne są poprzez określone rutery. Zakresy adresów są określone jako prefiksy (ang. prefix), a ścieżki opisujące kolejne systemy autonomiczne (AS) doprowadzające do wybranego prefiksu jako trasy (ang. path). BGP nie przenosi informacji o konkretnych urządzeniach, zamiast tego posługuje się terminem systemów autonomicznych. Taki system najczęściej tworzą logicznie wydzielone organizacje, firmy, dostawcy internetu. To jak w ramach AS (ang. Autonomous System) przekazywana jest informacja o ścieżkach, zależy już jedynie od administratorów danej sieci.

Każdy system autonomiczny aby brać udział w wymianie informacji przez BGP musi posiadać własny numer, który przydzielany jest przez IANA/RIPE². Jest to 32 bitowa liczba - Autonomous System Number (ASN), która używana jest następnie przez BGP przy wymianie informacji. Numery te są unikatowe, aby nie doszło do zapętlenia się routingu. Proces przyznawania numerów ASN jest opisany w dokumencie RIPE-279, RIPE-278 dostępnych na stronach <http://www.ripe.net/>. Podobnie jak w przypadku numerów IP także tutaj można wykorzystać prywatne numery AS. Numery te są często wykorzystywane w dużych sieciach opartych o BGP, przykłady takiego wykorzystania zostaną przedstawione w kolejnych częściach. Prywatne ASN, są to numery z zakresu: 64512-65535.

Przykładowe numery ASN:

GTS	-	8246
TPNET	-	5617
TAMI	-	29620

W bazach whois RIPE i pozostałych agend IANA przechowywane są informacje dotyczące numerów AS i jednostek, którym zostały przypisane. Aby dowiedzieć się więcej o którymś z numerów ASN można wykorzystać program whois:

```
#whois as8246

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

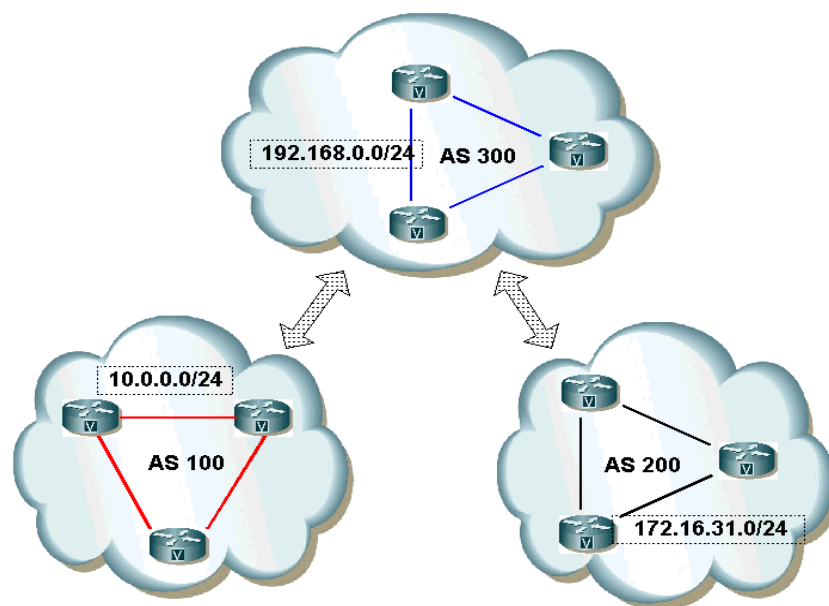
Posiadając własne adresy IP (Provider Independent), niezależne od naszego dostawcy Internetu adresy IP przyznane przez RIR (ang. Regional Internet Registry), w

2 IANA - Internet Assigned Numbers Authority, <http://www.iana.org/>. Pod patronatem ISOC zajmuje się sprawowaniem nadzoru nad rozwojem Internetu. RIPE - Réseaux IP Européens, <http://www.ripe.net/> jest jednostką wydzieloną przez IANA na teren Europy.

przypadku Europy RIPE, możemy zacząć zabawę w konfigurowanie BGP (oczywiście nie posiadając PI, można także korzystać z BGP choćby w sytuacji kiedy mam dwa łącza od jednego operatora). Okazuje się często, że w podstawowych konfiguracjach jest to bardzo proste i nie potrzeba do tego wiedzy tajemnej jak często usiłują nam to wmówić niektórzy administratorzy;) Oczywiście w przypadkach bardziej niż prostych wiedza tajemna może się okazać niezbędna, chociażby do analizowania problemów.

Jak to działa – czyli wiedza tajemna.

Celem BGP jest przenoszenie informacji o dostępnych prefiksach poprzez poszczególne ASy. Jest to informacja nie tylko o trasach dostępnych, ale także o zmianach zachodzących w sieci BGP. Routery wykorzystujące nazywana są neighborami, zestawiają połączenia – peeringi pomiędzy sobą. W obrębie danego AS (iBGP) każdy router musi być połączony z każdym (da się to obejść za pomocą route reflectorów). Routery nie muszą być ze sobą bezpośrednio połączone. W przypadku kiedy routery odległe są od siebie o kilka hopów korzysta się z opcji ebgp-multi-hop.



Rysunek 2 BGP pomiędzy trzema odrębnymi sieciami

Router należący do AS100 informuje pozostałe AS o dostępnej przez niego trasie, czyli 10.0.0.0/24. Podobnie pozostałe ASN informują sąsiadów jakie prefiksy są u nich i przez nich dostępne. AS300 poza tym, że rozgłasza informację o swoim własnym prefiksie 192.168.0.0/24 informuje także sąsiadów jakie innego prefiksy są przez niego dostępne. AS300 informuje AS200, o dostępnej przez niego ścieżce do AS100 i odwrotnie, AS100 jest informowany o AS200. W rzeczywistości wygląda to nieco bardziej skomplikowanie ze względu na ilość AS, oraz strukturę i ich rozmieszczenie. Jeśli jest to router tranzytowy prowadzący do Internetu, tras które otrzymuje może być bardzo dużo (obecnie około 130 000). Jeśli jest to router, który informuje tylko o własnych trasach/zakresach adresów ip, może być ich kilka, a nawet tylko jedna (oczywiście możemy otrzymać pełen zakres tras, ale nie zawsze ma to sens).

Poniższy przykład pokazuje ścieżki otrzymane od ASN8246 - GTS/IPARTNERS. Jak widać poniżej są to tylko trzy trasy.

```
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

Atrybut as_path

Informacja o prefiksach przekazywana jest między routerami danych AS. Przechodząc przez nie, dodatkowo oznaczana jest numerami ASN sieci przez które przechodzi. Atrybut ten nazywany jest *as_path*. Przekazywanie informacji o ścieżce *as_path* oprócz informacji samej w sobie zapobiega występowaniu pętli, przez wykluczenie sytuacji w której ASN mógłby się powtórzyć *as_path*. Więcej o atrybutach wkrótce.

```
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
```

Atrybut origin

Drugim ważnym atrybutem jest origin. Określa on pochodzenie informacji. Może przyjmować trzy wartości:

- IGP: informacja o trasach pochodzi z polecenia *network* definiującego rozgłaszane informacje, lub informacja pochodzi z redystrybucji przez inny protokół routingu IGP (np. OSPF). Trasy takie są oznaczane w tablicy BGP przez „i”.
- EGP: informacja została otrzymana przez EGP, oznaczana jest ona w tablicy BGP przez „e”
- INCOMPLETE: proces BGP nie jest w stanie zidentyfikować źródła informacji, informacja ta może pochodzić także z redystrybucji tras statycznych. Informacje tego typu oznaczane są w tablicy BGP przez „?”

BGP w sieci TCP/IP

Zanim zagłębimy się w czeluści szczegółów BGP i konfiguracji warto wiedzieć jak BGP działa w warstwie sieciowej modelu OSI/ISO. Do komunikacji używany jest protokół TCP oraz port 179. Dzięki TCP protokół rutowania nie musi się martwić o utrzymywanie

połączenia i sprawdzanie poprawności danych. Niektóre z protokołów jak EIGRP wykorzystują własne protokoły stworzone specjalnie do komunikacji, w tym przypadku RTP, niektóre nie robią tego w ogóle jak RIP czy IGRP i korzystają z protokołów bezpołączeniowych jak UDP .

Powyżej warstwy transportowej BGP używa własnych mechanizmów do zestawiania sesji i wymiany danych. BGP tworzy trwałe połączenia pomiędzy ruterami bezpośrednio komunikującymi się. Używane jest kilka typów komunikatów do komunikowania się ruterów. Potrzebne są one do ustanawiania sesji, podtrzymywania jej, informowania routera sąsiada o zmianach oraz zamykania sesji.

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. Rutery 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.

Podczas zestawiania sesji BGP, wyróżnia się kilka stanów na podstawie których, możemy określić w jakim momencie znajduje się skonfigurowane przez nas połączenie BGP. Początkowo neighbor BGP znajduje się w stanie „**Idle**”, po zmianie konfiguracji routera wywołany jest proces, który próbuje nawiązać połączenie TCP z sąsiadem – stan „**Connect**”. Po wysłaniu pierwszych pakietów kiedy sesja TCP zostanie zestawiona, sesja BGP znajduje się w stanie „**OpenSent**”. Jeśli połączenie TCP z jakichś powodów nie może być ustanowione, sesja przechodzi w stan „**Active**”. Następnie jeśli sesję uda się otworzyć, jeden z ruterów wysyła pakiety dotyczące identyfikacji i sesja BGP przechodzi w stan „**OpenSent**”. Jeśli rozmówca potwierdzi informacje autoryzującą, sesja przechodzi w stan „**OpenConfirm**”, a następnie „**Established**”. Procedura ta bardzo dokładnie została opisana w RFC-1771 w rozdziale 8.

Poniżej widać fragment opisu połączenia BGP w stanie „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
```

Po ustanowieniu połączenia pomiędzy ruterami, wymieniana jest pełna informacja jaką rutery posiadają za pomocą pakietów UPDATE. Może to wygenerować w przypadku wymiany pełnej tablicy routingu całkiem spore obciążenie procesora i znaczny ruch. Następnie przekazywane są pomiędzy speakerami BGP jedynie informacje o zmianach, o

dostępnych nowych ścieżkach oraz o ścieżkach, które stały się niedostępne.

Redystrybucja tras

Ważną sprawą jest to w jaki sposób proces BGP danego routera otrzymuje informacje o dostępnych trasach. Poza otrzymywaniem informacji z innych routerów uczestniczących w procesie wymiany BGP jest kilka sposobów na „wstrzykiwanie” tych informacji do BGP.

```
network numer-sieci [mask maska-sieci]
```

Polecenie to informuje BGP o sieciach które są dostępne przez router. Ponieważ BGP wspiera class-less routing³, czyli może przesyłać informację o maskach dowolnej długości poza numerem sieci możemy określić też maskę sieci. Przy użyciu polecenia network bardzo ważne jest aby router wiedział jak dotrzeć do danej sieci. Jeśli informacji o tym będzie brakować w tablicy routingu, informacja o sieci nie będzie przekazywana do BGP. Sieć ta może być umieszczona w tablicy routingu jako trasa statyczna, trasa przekazana przez protokół IGP (np. OSPF) lub sieć bezpośrednio połączona (ang. directly connected).

Przy trasach statycznych możemy to wymaganie trochę ominąć przez skierowanie danej sieci na interfejs null:

```
router bgp 100
network 10.0.0.0 mask 255.0.0.0

ip route 10.0.0.0 255.0.0.0 null 0
```

Inny sposób na przekazywanie informacji do BGP to redystrybucja informacji z IGP (IGRP, OSPF, RIP, EIGRP, etc.). Należy w tym wypadku zwrócić szczególną uwagę na to jakie informacje są przekazywane z IGP do BGP. Informacje to można filtrować, tak aby BGP nie zostało zalane ścieżkami, które nie powinny się dostać do BGP.

Atrybuty BGP

Wraz z informacją o prefiksach i ich długości przesyłana jest dodatkowa informacja w pakietach UPDATE. Pozwalają one dokonywać procesowi BGP wyboru optymalnej ścieżki. Za ich pomocą administrator może wpływać na działanie procesu BGP i filtrować wybrane ścieżki.

Wcześniej przedstawiony został krótko atrybut *as_path*. Jest to jeden z najważniejszych atrybutów, ale jest ich znacznie więcej. Można je podzielić na takie które każda implementacja BGP musi znać i przekazywać i takie, które może ale nie musi znać oraz przekazywać.

well-known mandatory – są to atrybuty, które muszą występować w pakiecie UPDATE, oraz muszą być rozpoznawane przez wszystkie implementacje BGP. W przypadku braku takiego atrybutu wysyłany jest pakiet NOTIFICATION, który powoduje, że sesja BGP zostaje zerwana.

³ class-less routing - routing w którym protokoły przesyłają informację o masce sieci podczas przesyłania informacji o danej sieci. class-less routing pozwala na wykorzystanie Variable-Length Subnet Mask (VLSM) i supernetting.

- as-path
- next-hop
- origin

well-known discretionary – atrybut, który także musi być rozpoznawany przez wszystkie implementacje BGP, ale niekoniecznie musi się pojawiać w pakietach UPDATE.

- local preference
- atomic aggregate

optional transitive – są to atrybuty, które nie muszą być rozpoznawane przez proces BGP, ale powinny być przenoszone. Proces BGP nie zajmuje się treścią pakietu, przekazuje go dalej.

- aggregator
- communities

optional non-transitive – atrybuty o najmniejszym znaczeniu, nie dość, że nie jest wymagane aby były rozpoznawane przez poszczególne implementacje BGP, dodatkowo nie będą przekazywane dalej do peerów BGP

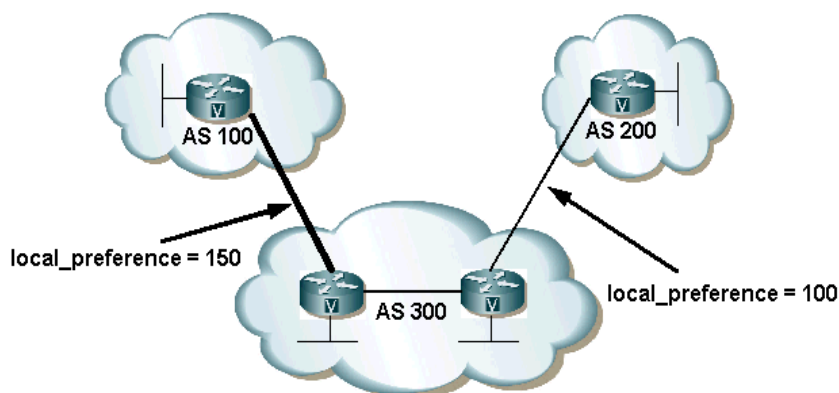
- MED – multi exit discriminator

Przy pomocy tych atrybutów BGP podejmuje decyzje co do wyboru ścieżek. Dzięki temu, że można wpływać na wartość tych atrybutów, możemy zmieniać działanie BGP.

Na szczególną uwagę zasługują *local_preference* i *MED*, zostaną omówione pokrótce poniżej.

Local Preference

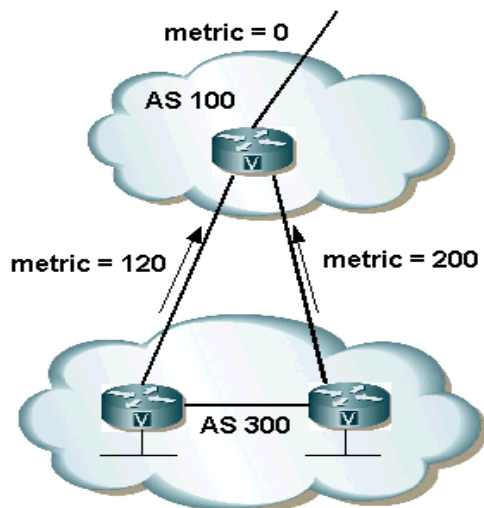
Za pomocą tego atrybutu możemy wskazać BGP, która trasa wychodząca z AS jest preferowana. Ścieżka, która została oznaczona *local_preference* o najwyższej wartości będzie preferowana. Domyślna wartość to 100. Wartość ta może być zmieniana za pomocą route map. W ten sposób możemy kierować ruchem wychodzących z AS. Jak będzie ruch wychodzący rozkładany zależy tylko od naszego "widzimisie", niestety nie ma złotej recepty na to żeby to robić w sposób optymalny.



Rysunek 3 Wykorzystanie local_preference do wskazania wychodzącej trasy

MED

Multi Exit Discriminator - atrybut ten w przeciwieństwie do local preference pozwala na kontrolowanie, którą ścieżką ruch będzie wchodził do danego AS. Zmieniając wartość tego atrybutu możemy wpływać, na wybór ścieżki, która zostanie wybrana przez router w AS do którego mamy co najmniej dwa różne połączenia.



Rysunek 4 Za pomocą MED możemy określać preferowane wejście do naszego AS

Preferowana jest niższa wartość, domyślnie ma wartość 0. MED może być przenoszony jedynie pomiędzy AS sąsiadującymi. Router, który otrzymuje informację o ścieżkach z tego samego AS, porównuje MED. Aby wybrać ścieżkę najlepszą.

Wybór najlepszej trasy

Proces BGP podejmuje decyzję, która ze ścieżek przekazać do tablicy routowania na podstawie wielu argumentów. Oczywiście zakładamy że dany router może mieć i ma wiele ścieżek do tego samego miejsca. Na wiele z tych argumentów na podstawie których podejmowane są decyzje można wpływać za pomocą route-map. Poniżej przedstawiony jest proces wybierania najlepszej trasy.

1. Jeśli NextHop jest niedostępny przechodzimy do pkt 2
2. Wybierana jest ścieżka z najwyższym atrybutem WEIGHT
3. Jeśli WEIGHT są identyczne wybierana jest z najwyższym Local-Preference
4. Jeśli Local-Preference są równe wybierz trasę, która pochodzi z procesu BGP pracującego na tym routerze
5. Wybierz najkrótszy AS-PATH
6. Jeśli wszystkie ścieżki pochodzą z poza routera wybierz tą z najniższym origin type (IGP<EGP<INCOMPLETE)
7. Wybierz ścieżkę z najniższym MED.
8. Preferuj ścieżkę : "zewnątrzną nad „wewnętrzną”
9. Jeśli synchronizacja jest wyłączona, wybierz najbliższą ścieżkę pochodzącą z IGP
10. Wybierz trasę o najniższym adresie IP, który wskazuje na router id

W tym schemacie uwzględniony został dodatkowo atrybut "weight", który jest używany jedynie w routerach Cisco.

Peer Group

Kiedy zarządzamy dużą siecią BGP i mamy do administracji wiele sesji BGP może zdarzyć się, że część z nich może wymagać podobnych ustawień. Za pomocą konfiguracji dotyczącej danej peer grupy możemy wpływać na konfigurację wielu sesji BGP. Każdy z sąsiadów może mieć skonfigurowane własne dodatkowe parametry. W ten sposób można wpływać jedynie na przychodzące informacje.

```
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
```

Obydwa routery, 172.16.11.1 oraz 10.0.0.2 należą do tej samej peer grupy o nazwie wan-link. Do tej peer grupy została przypisana filter lista o numerze 1.

Communities

Communities są szczególnymi atrybutami na które warto zwrócić większą uwagę. Za ich pomocą można oznaczać trasy wymieniane w sesji BGP. Dzięki temu można filtrować otrzymywane informacje czy wpływać na to jak są wymieniane. Communities mogą być definiowane przez administratorów danych AS. Mogą na przykład oznaczać wybrane trasy danego operatora. Poza dowolnie definiowanymi communities istnieją zdefiniowane, powszechnie znane:

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

Po szczegółowy opis odsyłam do bardziej szczegółowych dokumentów.

Bardzo często właściciele danego ASN opisują używane przez nich communities w bazach whois aby ułatwić konfigurację administratorom systemów z nimi połączonych za pomocą BGP:

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

Możemy też sprawdzić czy rzeczywiście communities opisywane przez danego ISP są używane.

```
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

Filtrowanie

BGP pozwala na kontrolowanie przekazywanej informacji na kilka sposobów. Dane mogą być kontrolowane na podstawie prefiksów, ścieżek AS czy communities. Całość daje bardzo elastyczne narzędzie do kontroli informacji przychodzącej i wychodzącej z BGP.

Filter list

Za pomocą filter-list można zmieniać informację na podstawie atrybutu as-path. Dzięki temu możemy wpływać na wiele tras dotyczących konkretnego dostawcy, czy układu dostawców w ścieżce.

W poniższym przykładzie od sąsiada ASN200 przyjmujemy tylko trasy rozgłaszane przez niego, a wysyłamy do niego jedynie trasy od ASN300 (nie lokalne!, as-path dla lokalnych tras jest pusty).

```
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
```

Lista o numerze 1 pozwala na przyjmowanie tras z AS o numerze ASN 200 i odrzucanie tras w odpowiedzi w których AS-PATH zaczyna się od ASN 300.

Lista 11 określa jakie trasy mogą być wysyłane, w tym przypadku będą to trasy z AS-PATH rozpoczynającym się od ASN 400.

Dzięki wyrażeniom regularnym stosowanym w powyższych access listach możemy w bardzo dokładny sposób filtrować informacje przekazywane. Najlepszym sposobem aby sprawdzić czy dane wyrażenie regularne dopasowuje te ścieżki AS o które nam chodziło jest wykonanie następującego polecenia:

```
show ip bgp regexp <regexp>
```

Distribute list

Dodatkowo możemy filtrować konkretne trasy które są przekazywane do tablicy routingu określając je za pomocą standardowych i rozszerzonych access-list.

Poniższa konfiguracja pozwala na przyjęcie/odrzucenie od routera 172.16.11.254 w sesji iBGP (te same ASN) jedynie tras dopasowanych do access-listy o numerze 15.

```
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
```

Dzięki wykorzystaniu access listy 15, router komunikując się z routerem o adresie 172.16.11.254 otrzyma informację o trasie 172.16.0.0/16, natomiast trasa 172.15.0.0/16 zostanie odfiltrowana i informacja o niej nie dostanie się do procesu BGP.

Prefix list

Za pomocą prefix-list możemy ograniczyć informację o trasach jakie są przekazywane od danego peera do tablicy routingu. Poniżej prosty przykład nie pozwalający aby od sąsiadów przekazywane były default routy.

```
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
```

Prefix listy pozwalają na dużą kontrolę dzięki możliwości określania długości prefixów

```
ip prefix-list lista-20 permit 172.16.0.0/16 ge 17
```

Lista ta pozwala na filtrowanie klas o długości większej od /17 bitów. Jedynie pierwsze /16 musi być dopasowane do nadchodzących update'ów.

```
ip prefix-list lista-20 permit 172.16.0.0/16 le 32
```

Lista poniżej pozwala na filtrowanie całej klasy B, czyli od prefixu /16 do /32.

```
ip prefix-list lista-20 permit 172.16.0.0/16 ge 18 le 31
```

Zakresy dopasowań można łączyć jak powyżej. Dopasowane zostaną wszystkie prefixy o długości od /18 do /31.

Community list

Listy te pozwalają na filtrowanie tras za pomocą communities. Bardzo często wykorzystywane są przez administratorów sieci do oznaczania tras wg bardziej funkcjonalnych i nie technicznych kryteriów. To miejsce gdzie do BGP wkrada się trochę polityki i tzw. warstwy 9tej czyli finansowej;).

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

Lista ta pozwala na przyjmowanie tras, które są oznaczone community o numerze 8246:666 oraz 8246:100. W ten sposób możemy organizować trasy wg bardziej abstrakcyjnych kryteriów. Np. położenia geograficznego, czy funkcjonalnego.

```
ip community-list 1 permit 100:1 100:2
```

Listę community do których będą dopasowywane ścieżki można także podać jako listę, jak powyżej.

Communities mogą być ustawiane w route mapach, które zostaną opisane poniżej.

Route maps

Route mapy są używane przy BGP do kontroli i modyfikacji tablicy routingu oraz wpływają na redystrybucję informacji. Route mapy są zbiorem zdań za pomocą których możemy określić warunki, które muszą być spełnione aby route mapa została użyta, oraz wyrażenia, które zostaną wykonane w określonej sytuacji. Zdania w route mapach mogą być ponumerowane sekwencyjnie. Route mapa ze słowem kluczowym permit powoduje że ścieżki, które będą dopasowane przez match będą wprowadzane do tablicy routingu. Jeśli mają identyczne nazwy stanowią jedną route mapę.

Route mapy są bardzo elastycznym i bardzo przydatnym narzędziem. Poniżej przykład przedstawiający wykorzystanie przy zmianie atrybutu *local-preference*, oraz zapobieganiu wysyłania śmieci do internetu.

```
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-path

W niektórych sytuacjach możemy chcieć mieć wpływa na długość ścieżki AS-PATH. Pozwala to na kontrolowanie w pewnym stopniu wyboru najlepszej ścieżki oraz informacji wysyłanej przez proces BGP. Prependowanie pozwala na wydłużenie ścieżki Asów, co najłatwiej pokazać na przykładzie.

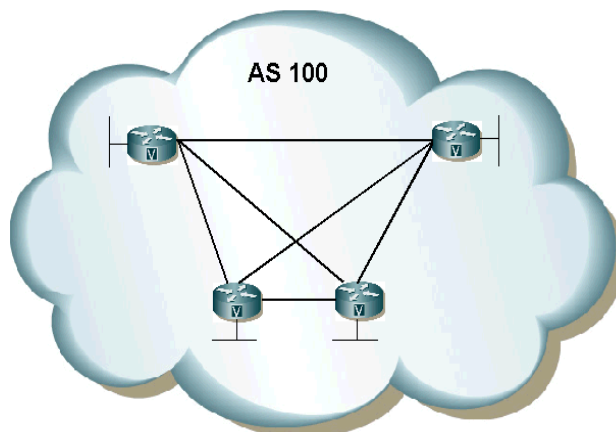
```
router bgp 100
    network 172.16.0.0
    neighbor 10.0.0.1 remote-as 200
    neighbor 10.0.0.1 route-map SETPATH out

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

W ten sposób rozgłaszana przez AS100 sieć będzie niosła informację o ścieżce składającą się z „100 100 100”, zamiast „100”. Ścieżka ta w procesie je wybory będzie mniej preferowana, przez wydłużony atrybut as-path.

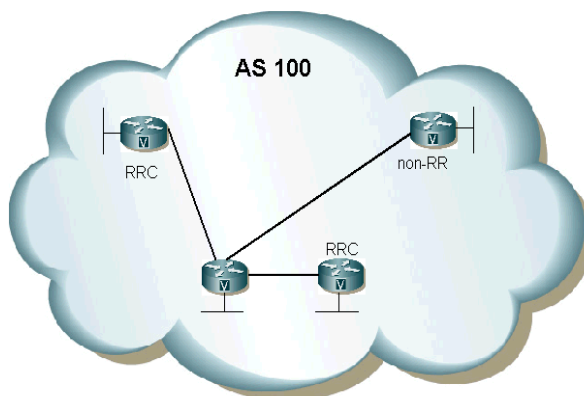
Route reflectors – zbyt dużo informacji...

BGP wymaga aby wszystkie routery w danym AS tworzyły ze sobą bezpośrednie sesje. Poza wysiłkiem administracyjnym jaki można sobie wyobrazić przy większej ilości routerów, powoduje to duży ruch w sieci.



Rysunek 5 Liczba połączeń w sieci z BGP rośnie lawinowo z każdym nowym routerem

Rozwiązaniem tego problemu jest wykorzystanie route reflectorów. Pozwala to na wydzielenie w sieci pewnych logicznych grup routerów pomiędzy którymi wymieniana jest informacja. Część routerów pracuje jak route reflectory, stają się one jakby punktem styku z innymi routerami nazywanymi „route reflector clients”. Dzięki temu routery-klienci komunikują się jedynie z route reflectorami, te zaś jedynie między sobą. Grupa routerów komunikujących się z jednym route reflektorem nazywana jest klustrem. Komunikacja jedynie między klastrami i routerami niekorzystającymi z tych mechanizmów w znacznym stopniu ogranicza ilość koniecznych do zestawienia sesji BGP i wymienianej informacji.



Rysunek 6 Użycie route reflectorów znacznie zmniejsza ilość koniecznych do zestawienia sesji BGP

Kontrola działania

Zakładając, że mamy już skonfigurowane BGP pomiędzy routerami, warto znać kilka podstawowych poleceń, które będą niezbędne przy monitorowaniu jego działania.

Poniżej widać 2 sesje z których krótsza działa od ponad 4 dni. Z obydwu sesji otrzymano ponad 128000 prefiksów, czyli tras do różnych sieci IP.

```
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
```

Na routerze, na którym działa Quagga zostały zestawione dwie sesje. Daemon Quagga zajmujący się wprowadzaniem tras do tablicy routingu jak widać rozdzielił dość nieproporcjonalnie trasy pomiędzy dwu dostawców.

```
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
```

Tu sprawdzamy jaka informacja jest przechowywana w procesie BGP o trasie 217.153.107.0. Jak widać poniżej, pierwsza została wybrana jako najlepsza (best #1). Ma najwyższy local_preference, poza tym krótszą ścieżkę AS – as_path.

```
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)
```

Jak widać w tablicy routingu została umieszczona jedynie trasa najlepsza.

Jeśli chcemy sprawdzić jakie trasy do sieci TPNET otrzymujemy od naszych

peerów wystarczy sprawdzić jakie informacje dostajemy o AS 5617.

```
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
```

Poniżej widać fragment zrzutu z działającego tcpdump nasłuchującego ruchu TCP na porcie 179 czyli BGP. Widać wyraźnie przesyłane pakiety keepalive oraz update. Jeszcze lepszą wizualizację tego co znajduje się w tych pakietach pozwala uzyskać program trafshow.

```
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)
22:40:42.128646 host-ip101-160.crowley.pl.bgp > host-ip62-199.crowley.pl.58266: P
441:754(313) ack 20 win 16118: BGP [|BGP UPDATE] [tos 0xc0] [ttl 1]
22:40:42.128680 host-ip62-199.crowley.pl.58266 > host-ip101-160.crowley.pl.bgp: .
ack 754 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)
```

I na koniec pokazana została pełna informacja o jednym z sąsiadów BGP naszego routera:

```
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
  Last read 00:00:05, hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received (old and new)
    Address family IPv4 Unicast: advertised and received
  Received 1226920 messages, 2 notifications, 0 in queue
  Sent 31874 messages, 10 notifications, 0 in queue
```

```
Route refresh request: received 0, sent 0
Minimum time between advertisement runs is 30 seconds
```

```
For address family: IPv4 Unicast
Inbound soft reconfiguration allowed
Community attribute sent to this neighbor (both)
Inbound path policy configured
Outbound path policy configured
Incoming update prefix filter list is *10
Route map for outgoing advertisements is *localonly
129167 accepted prefixes
```

```
Connections established 14; dropped 13
Last reset 01w0d16h
External BGP neighbor may be up to 15 hops away.
Local host: 217.153.71.52, Local port: 179
Foreign host: 157.25.1.16, Foreign port: 42451
Nexthop: 217.153.71.52
Read thread: on Write thread: off
```

Przykładowe konfiguracje

Najczęściej używane konfiguracje w których wykorzystuje się BGP to:

- transit AS
- dwa i więcej łącz od różnych AS
- dwa i więcej łącz od tego samego dostawcy o różnych przepustowościach i różnym koszcie (koszt w sensie zł:)

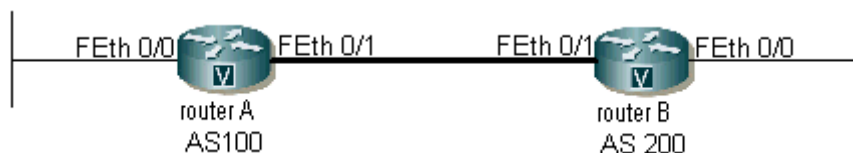
Kiedy nie warto, a nawet nie powinno się wykorzystywać BGP:

- mamy jedno połączenie do sieci
- nie mamy odpowiedniego sprzętu/routera
- nie mamy odpowiedniej obsługi potrafiącej poprawnie skonfigurować router

Scenariusz 1

Przykład ten jest typowo edukacyjny, ma zadanie jedynie przedstawić najprostszą możliwą sytuację. Jedno połączenie z AS100 do AS200. W praktyce nie ma to wielkiego sensu, ponieważ wystarczyłoby w AS100 ustawić routing domyślny do AS200 i odwrotnie.

Schemat 1



Konfiguracja routerów

Router A

```
interface fastethernet 0/1
```

```

description "Polaczenie WAN"
ip address 10.1.0.1 netmask 255.255.255.0

interface fastethernet 0/0
description "Polaczenie LAN"
ip address 172.16.0.0 netmask 255.255.255.0

router bgp 100
neighbor 10.2.0.1 remote-as 200
network 172.16.0.0 mask 255.255.255.0

```

Router B

```

interface fastethernet 0/1
description "Polaczenie WAN"
ip address 10.2.0.1 netmask 255.255.255.0

interface fastethernet 0/0
description "Polaczenie LAN"
ip address 192.168.0.0 netmask 255.255.255.0

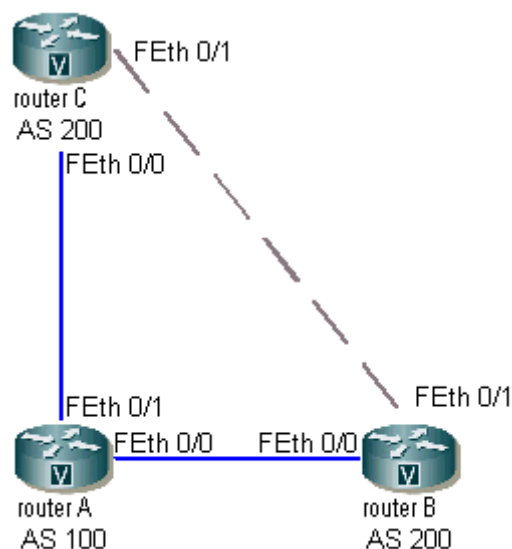
router bgp 200
neighbor 10.1.0.1 remote-as 100
network 192.168.0.0 mask 255.255.255.0

```

Konfiguracja powyższa jest bardzo prosta, w obydwu routerach zostały skonfigurowane interfejsy ethernetowe, przypisane im zostały adresy IP. W konfiguracji BGP podany został jedynie adres sąsiada i określona sieć która będzie ogłaszana.

Scenariusz 2

Router A znajdujący się w AS100 ma dwa połączenia z routerem B i C, znajdującymi się w AS200. Jest to sytuacja kiedy klient ma dwa połączenia z jednym dostawcą. Obydwa połączenia są równoważne.



Konfiguracja routerów

Router A

```

interface fastethernet 0/0
description "Polaczenie do routera B"
ip address 172.16.0.1 netmask 255.255.255.252

```

```
interface fastethernet 0/1
  description "Polaczenie do routera C"
  ip address 10.1.0.1 netmask 255.255.255.252

router bgp 100
  neighbor 10.1.0.2 remote-as 200
  neighbor 172.16.0.2 remote-as 200
```

Router B

```
interface fastethernet 0/0
  description "Polaczenie do routera A"
  ip address 172.16.0.2 netmask 255.255.255.252

interface fastethernet 0/1
  description "Polaczenie do routera C"
  ip address 192.168.0.1 netmask 255.255.255.252

router bgp 200
  neighbor 10.1.0.1 remote-as 100
  neighbor 192.168.0.2 remote-as 200
```

Router C

```
interface fastethernet 0/0
  description "Polaczenie do routera A"
  ip address 10.1.0.2 netmask 255.255.255.252

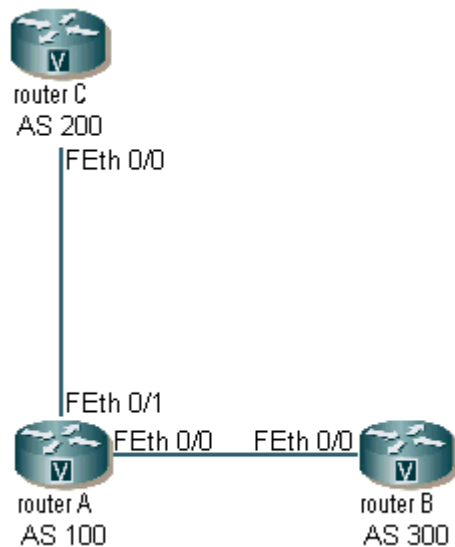
interface fastethernet 0/1
  description "Polaczenie do routera B"
  ip address 192.168.0.2 netmask 255.255.255.252

router bgp 200
  neighbor 10.1.0.1 remote-as 100
  neighbor 192.168.0.2 remote-as 200
```

Jest to jedna z popularniejszych konfiguracji, pozwalających na zapewnienie sobie backupowego połączenia w przypadku korzystania jedynie z usług jednego dostawcy. Konfiguracja w zasadzie nie zawiera nic ponadto co jest niezbędne do zestawienia połączeń. Router A ma zestawione połączenia BGP z routerem B i C. Ponieważ połączenia te są równoważne pozwalamy BGP decydować o tym z którego AS100 ma korzystać. Pomiedzy routerami B i C zestawiona jest sesja iBGP zapewniająca komunikację w AS200.

Scenariusz 3

W scenariuszu tym drugie połączenie zapewnia inny dostawca. W ten sposób uniezależniamy się od większej awarii, która mogła by nam wyeliminować obydwie połączenia do internetu jak w scenariuszu drugim.



Konfiguracja routerów

Router A

```

interface fastethernet 0/0
  description "Polaczenie do routera B - AS300"
  ip address 172.16.0.1 netmask 255.255.255.252

interface fastethernet 0/1
  description "Polaczenie do routera C - AS200"
  ip address 10.1.0.1 netmask 255.255.255.252

router bgp 100
  neighbor 10.1.0.2 remote-as 200
  neighbor 10.1.0.2 route-map localonly out
  neighbor 172.16.0.2 remote-as 300
  neighbor 172.16.0.2 route-map localonly out

ip as-path access-list 20 permit ^$

route-map localonly permit 10
  match as-path 20
  
```

Router B

```

interface fastethernet 0/0
  description "Polaczenie do routera A"
  ip address 172.16.0.2 netmask 255.255.255.252

router bgp 300
  neighbor 10.1.0.1 remote-as 100
  
```

Router C

```

interface fastethernet 0/0
  description "Polaczenie do routera A"
  ip address 10.1.0.2 netmask 255.255.255.252

router bgp 200
  neighbor 10.1.0.1 remote-as 100
  
```

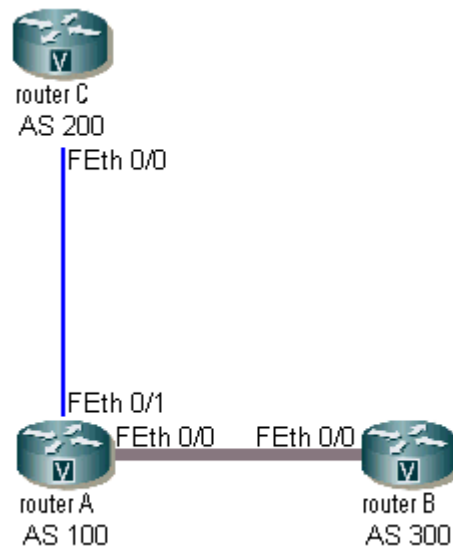
Konfiguracja ta podobna do poprzedniej także zapewnia nam backupowe

połączenie tym razem do innego dostawcy. Z routera A wyprowadzane są dwa połączenia EBGP do AS200 i AS300. Pomiędzy nimi może ale nie musi być dodatkowe połączenie. One zapewniają połączenie z resztą internetu i backup w przypadku awarii jednego z nich. Obydwa są równoważne więc można mówić o nadmiarowości połączeń raczej niż o ich backupie. Konfiguracja jest bardzo podobna do poprzedniej.

Warto zwrócić uwagę na to aby nie rozgłaszać informacji otrzymywanych od AS200 do AS300 i odwrotnie, chyba że chcemy to robić celowo. Tak może być w przypadku gdy jesteśmy punktem tranzytowym. Route mapa o nazwie localonly nie pozwala na wysyłanie niczego co nie pochodzi z lokalnego AS. Wszystkie informacje wysyłane z AS100 posiadają pustą ścieżkę Asów, czyli atrybut as_path można dopasować za pomocą regexpa $^{\$}$.

Scenariusz 4

Kolejnym krokiem w dostosowywaniu połączenia do Internetu może być preferowanie jednego z nich. Kryterium może być przepustowość, cena czy stabilność. Załóżmy, że połączenie z AS300 ma większą przepustowość, tam będziemy wymieniać cały ruch. Połączenie z AS200 jest traktowane jako typowy backup w przypadku zerwania połączenia podstawowego.



Router A

```
interface fastethernet 0/0
  description "Polaczenie do routera B - AS300"
  ip address 172.16.0.1 netmask 255.255.255.252

interface fastethernet 0/1
  description "Polaczenie do routera C - AS200"
  ip address 10.1.0.1 netmask 255.255.255.252

router bgp 100
  neighbor 10.1.0.2 remote-as 200
  neighbor 10.1.0.2 route-map localonly out
  neighbor 172.16.0.2 remote-as 300
  neighbor 172.16.0.2 route-map localonly out
  neighbor 172.16.0.2 route-map as300_best in

ip as-path access-list 20 permit ^$

route-map localonly permit 10
```

```
match as-path 20

route-map as300_best permit 10
set local-preference 150
```

Konfiguracja ta została rozbudowana w stosunku do poprzedniej o jedną route mapę. Powoduje ona, że ścieżki otrzymywane z AS200 otrzymują wyższy niż standardowy, równy 100 atrybut `local-preference` - 150. Powoduje, że w procesie wyboru najlepszych ścieżek, są one preferowane i wprowadzane do tablicy routingu.

Konfiguracje routerów B i C w nie zmieniają się.

Quagga

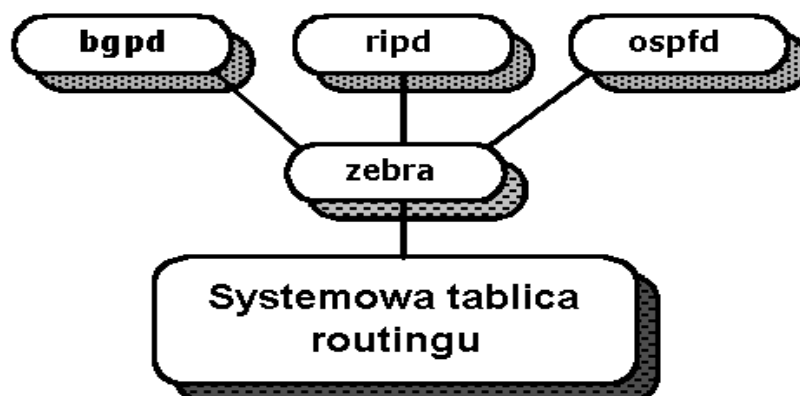
Quagga jest jednym z niewielu rozwiązań OpenSource pozwalających na stosowanie protokołów dynamicznego routingu w Linuxie. Jest to oprogramowanie bazujące na kodzie projektu Zebra (<http://www.zebra.org/>), której rozwój pozostawia ostatnio wiele do życzenia.

Poniższy opis pochodzi z dokumentacji, opisuje on w krótki sposób to czym jest Quagga:

"Quagga jest oprogramowaniem implementującym protokoły dynamicznego routingu dla protokołu TCP/IP takie jak RIPv1, RIPv2, RIPv6, OSPFv2, OSPFv3, BGP-4, i BGP-4+. Quagga posiada także zaimplementowany mechanizm Route Reflectorów i Route Serverów znany z BGP. Poza wsparciem dla protokołów routingu związanymi z Ipv4, Quagga wspiera także protokoły związane z Ipv6."

Co ciekawe i wygodne dla ludzi znających interfejs Cisco jest to, że autor Zebry, a obecnie główny deweloper Quaggi starali się go naśladować. Osoba, która używała interfejsu Cisco nie będzie miała kłopotu z konfigurowaniem Quaggi.

Quagga działa jako zespół demonów komunikujących się wzajemnie ze sobą. `zebra` odpowiada za komunikację z systemową tablicą routingu i komunikację z pozostałymi demonami `bgpd`, `ripd`, `ripngd`, `ospfd`, `ospf6d`. Te obsługują poszczególne protokoły i komunikują się jedynie z `zebrą` lub innymi routerami. Do poszczególnych demonów można się dostać poprzez telnet na port 2601 w przypadku `zebry` i 2605 w przypadku `bgpd`. Warto skorzystać z udogodnienia jakim jest `vttysh`, pozwalający na komunikację z wszystkimi działającymi demonami podczas jednego połączenia.



Oficjalnie wspierane platformy przez Quagga:

- GNU/Linux
- FreeBSD
- NetBSD
- OpenBSD
- Solaris

Quagga stanowi w pełni funkcjonalny odpowiednik routera, który może być wykorzystywany na styku pomiędzy operatorami.

To tylko wstęp do BGP...

Jak wspomniałem we wstępie jest to tylko krótkie wprowadzenie do tego czym jest BGP. Brakuje w nim bardzo wielu rzeczy, ale wtedy z krótkiego opracowania tekst ten zamieniłby się w kopię RFC:). Na co warto zwrócić uwagę przede wszystkim to problem synchronizacji, możliwość agregowania tras, wykorzystanie prywatnych ASN, korzystanie z route reflectorów i konfederacji BGP. Wiele problemów wynika w przypadku heterogenicznych sieci z IBGP. Zainteresowanych odsyłam do książek z bibliografii.

Bibliografia

Książki

1. Internet Routing Architectures, Sam Halabi, Danny McPherson, ISBN: 157870233X
2. Cisco BGP-4 Command & Configuration Handbook, William R. Parkhurst (Author), Ph.D., William R. Parkhurst, ISBN: 158705017X
3. BGP, Iljitsch Van Beijnum, ISBN: 0596002548

Internet

1. <http://www.ietf.org/rfc/rfc1771.txt>
2. <http://www.quagga.net/>
3. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm
4. <http://www.cisco.com/warp/public/459/bgp-toc.html>
5. http://www.cisco.com/en/US/tech/tk365/tk80/tech_tech_notes_list.html
6. <http://www.net-tech.bbn.com/sbgp/sbgp-index.html>
7. http://www.amysegowski.com/bgp_history.html
8. <http://en.wikipedia.org/wiki/BGP>
9. <http://www.ciscopress.com/articles/article.asp?p=169556>