Протокол BGP

проф. д-р инж. Венета Алексиева

Основни моменти

- Автономна система
- BGP същност
- BGP съобщения
- BGP състояния
- BGP принцип на действие
- Конфигуриране

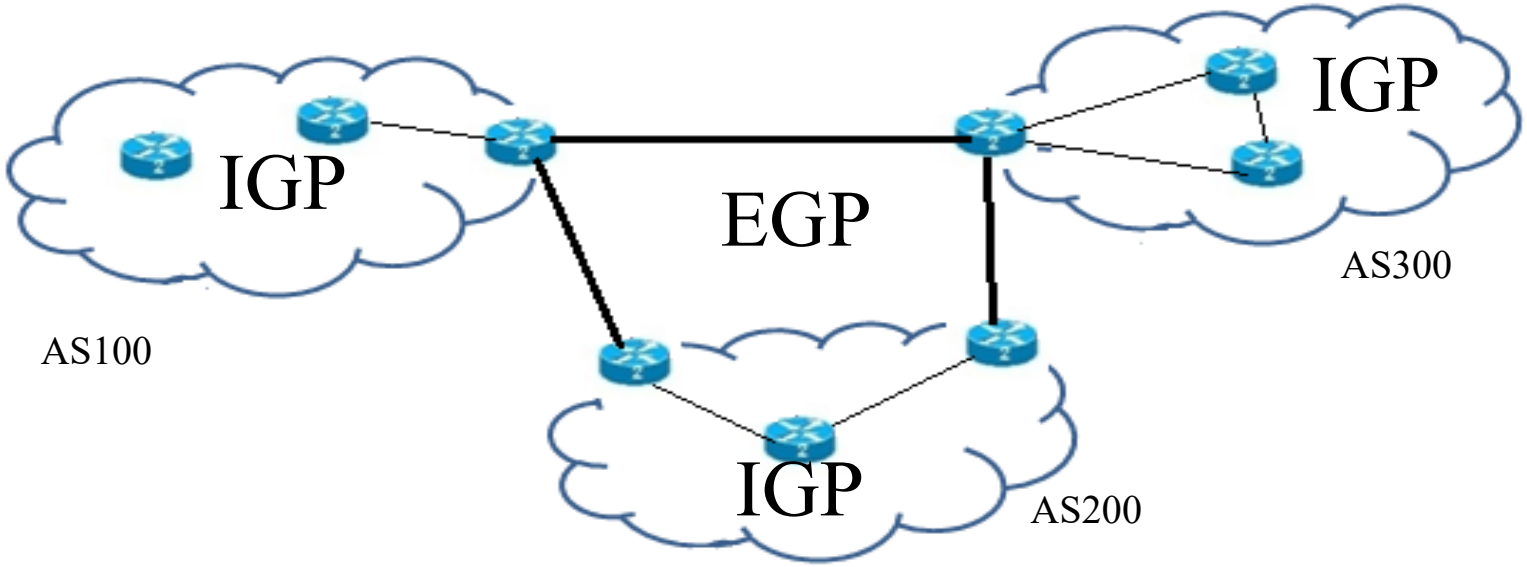
Автономни системи

AS - съвкупност от мрежи под общо администриране

EGP- между AS

IGP- вътре в AS

Номер на AS



BGP- същност

- Border Gateway Protocol (BGP) е EGP.
- BGP е маршрутен векторен протокол (Path Vector Protocol).
- Маршрутната информация съдържа последователност от номера на AS, указваща пътя до дестинацията, съвместно с атрибути на маршрута до нея.
- BGP осъществява маршрутизиране без цикли между AS.
- Работи с VLSM и CIDR.
- Работи с TCP по порт 179.
- Не налага никакви ограничения върху използваната мрежова топология, стига маршрутизирането в AS да е реализирано от IGP (RIP,OSPF…).
- Праща incremental updates, а не цели updates като EGP.

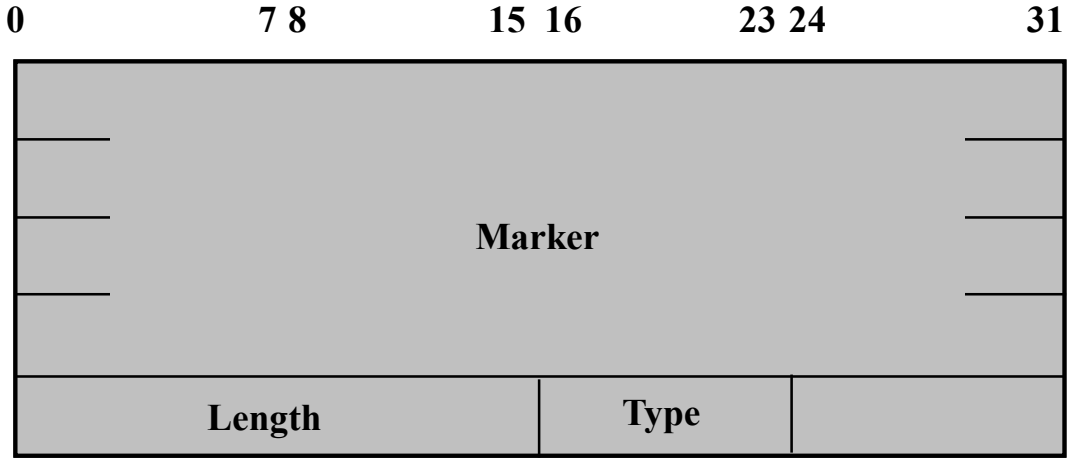
BGP е подходящ когато...

- Трябва да се балансира натоварването при наличие на множество изходни точки, свързващи към един ISP.
- Трябва да се управлява начина за препращане на трафика по няколко пътища от различни ISP.
- Трябва да се избере път по конкретни критерии.
- Ако инфраструктурата на конфигурираната мрежа се използва като транзитна област за трафик на други AS.

История

- RFC 1105 представя за пръв път BGP.
- През 1989 BGP става Интернет стандарт.
- RFC1771 представя характеристиките на BGP-4.
- BGP е един от най-използваните съвременни външни маршрутизиращи протоколи в Интернет.
- BGP строи граф от автономни системи на базата на информация, обменяна между BGP маршрутизаторите.

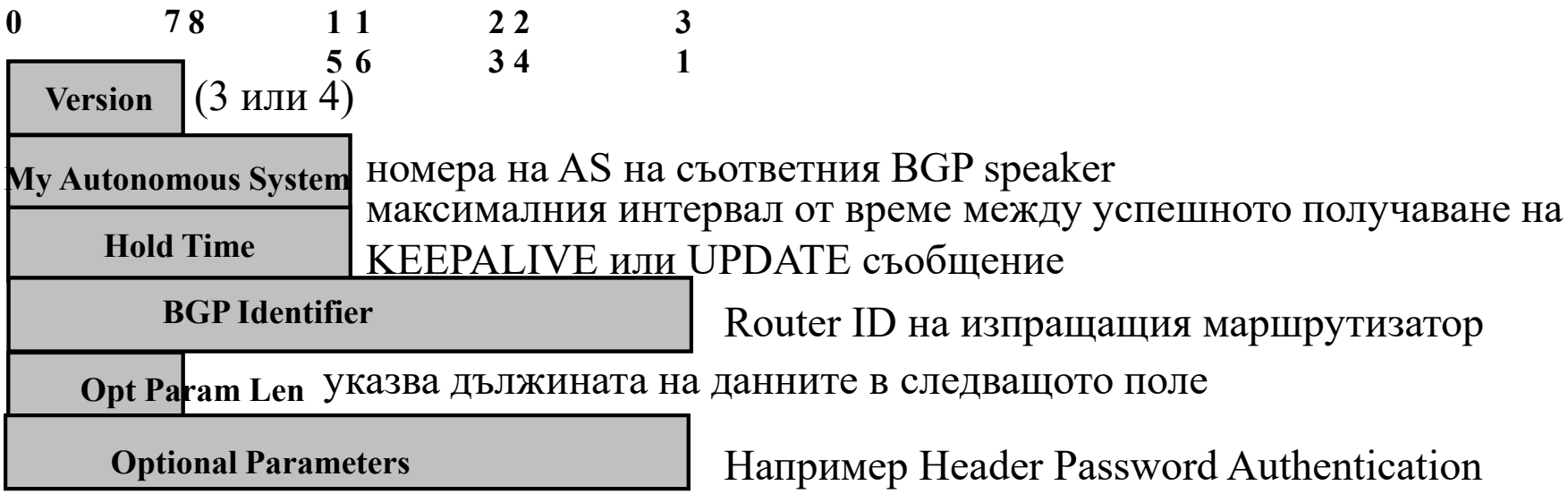
BGP съобщения



- В един TCP сегмент - няколко BGP съобщения
- **Marker** –автентикация на получаваните BGP пакети и за установяване отсъствието на синхронизация между BGP съседи:
  - съдържа само 1-ци - ако типът на съобщението е OPEN, или няма автентикация;
  - съдържа информация, получена на базата на използвания механизъм за автентикация.
- **Length** – указва дължината на съобщението (19 – 4096 байта);
- **Type** – идентифицира типа на BGP съобщението - 5 типа съобщения – **OPEN, UPDATE, KEEPALIVE, NOTIFICATION, ROUTE_REFRESH**

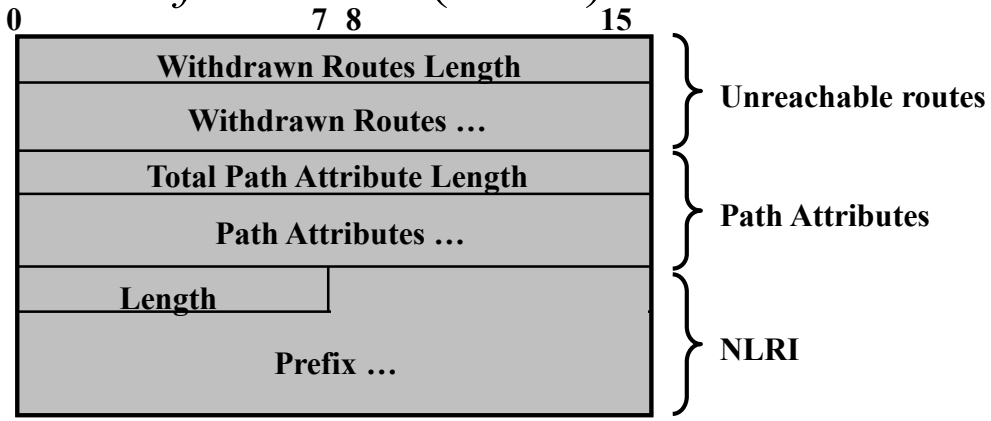
BGP съобщения- OPEN

- 1-то съобщение след установяване на TCP сесия
- За конфигурационна информация и договаряне на общите параметри за сесията между съседите



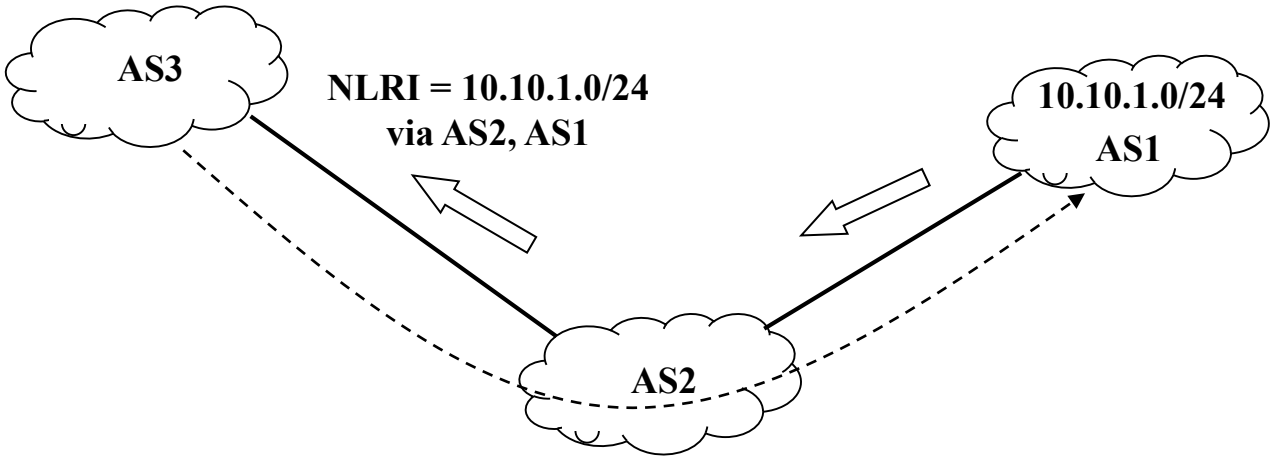
BGP съобщения- UPDATE

- Изпращат се само след установяване на сесия между BGP съседи.
- Съобщението съдържа 3 основни блока:
  - недостижими пътища (*Unreachable Routes*),
  - атрибути на пътя (Например атрибута AS_PATH, който представлява последователност от номера на АС, които трябва да се преминат по маршрута.)
  - информация за достижими пътища-*Network Layer Reachability Information* (NLRI)



Пример

- AS3 получава UPDATE съобщение, указващо, че мрежата 10.10.1.0/24 е достижима през две стъпки – първо AS2 и после AS1.
- Така AS3 може да препрати трафика до мрежата 10.10.1.0/24 транзитно през AS2 до дестинацията в AS1.

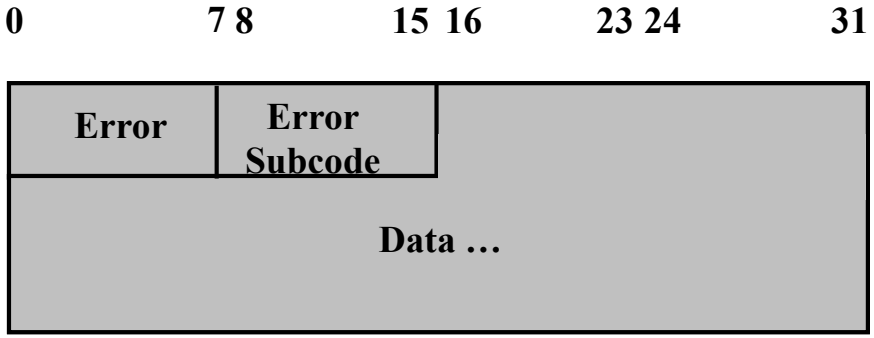


BGP съобщения- KEEPALIVE

- KEEPALIVE съобщенията съдържат единствено BGP хедър и не включват данни.
- Обменят се периодично за установяване наличието на всеки от маршрутизаторите.
- Те се изпращат с честота, осигуряваща Hold таймера да не изтече.
  - Препоръчва се честотата да е 1/3 от стойността на Hold Timer.

BGP съобщения- NOTIFICATION

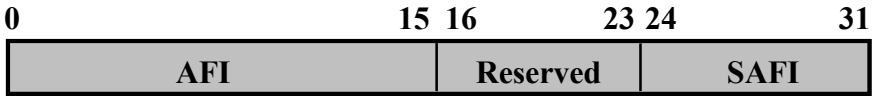
- Изпраща се при откриване на грешка от BGP.
- Следва преустановяване на сесията със съседа и затваряне на TCP връзката.
- Условието за грешка се изпраща към съседа за откриване и анализ на проблема от мрежовия администратор.



•**Error** –тип на грешката;
•**Error Subcode** –детайли за грешката;
•**Data** –същността на грешката

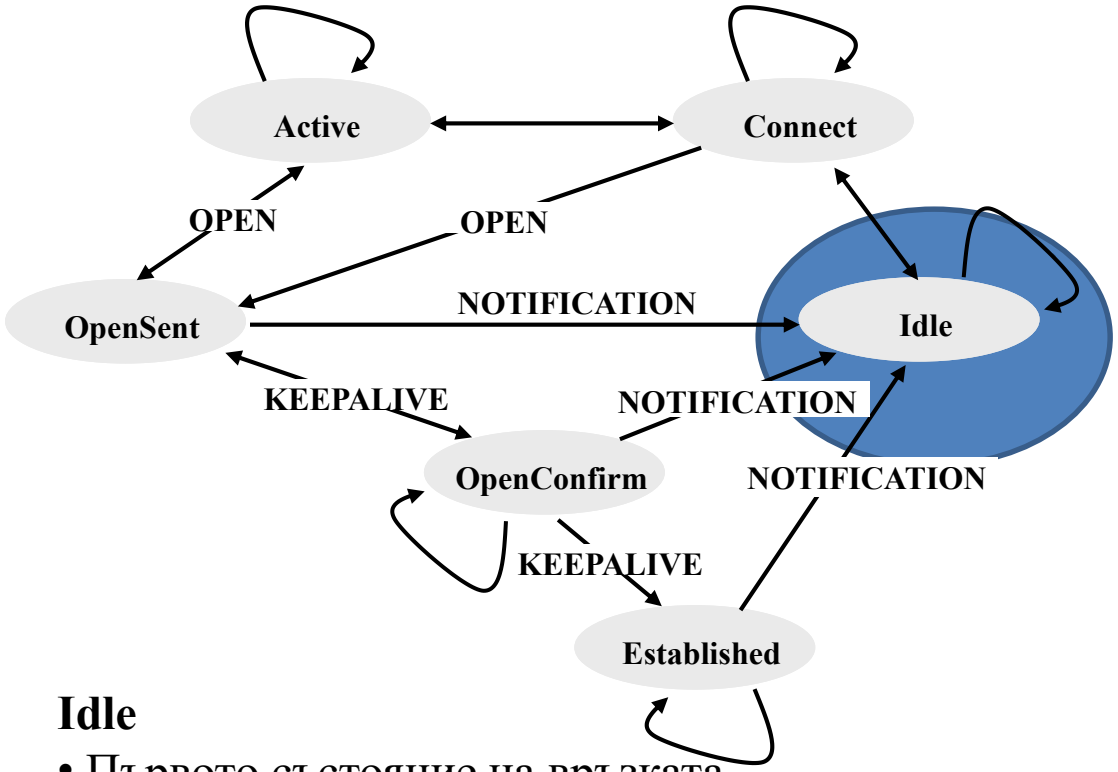
BGP съобщения- ROUTE_REFRESH

- Това съобщение е добавено като допълнителна възможност на BGP-4 в RFC2918.
- За заявяване на цялостно повторно предаване на маршрутната информация от BGP съсед без затваряне на текущата BGP сесия и последващото й създаване.



AFI – Address Family Identifier, указващ IPv4 или IPv6 протокол
Reserved – само нули
SAFI – Subsequent Address Family Identifier, задава unicast или multicast тип на използваните адреси.

И как работи...



Idle

• Първото състояние на връзката.

•Ако настъпи грешка, BGP се връща в **Idle** състояние.

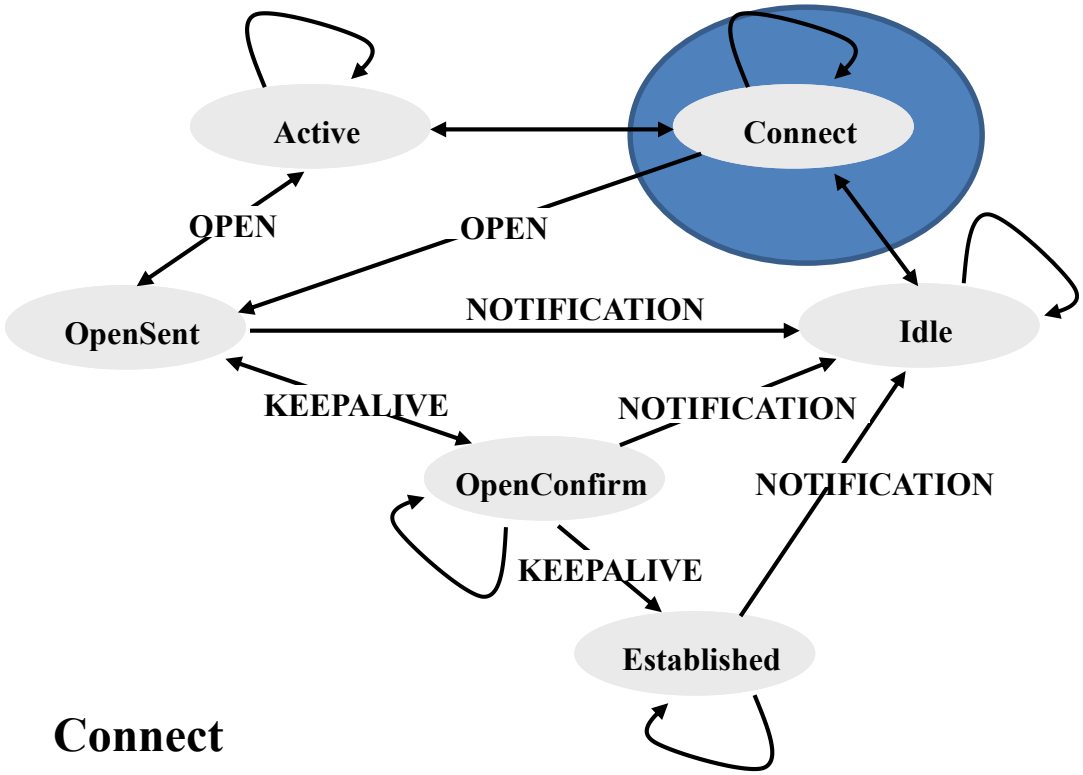
•Инициира се от оператора или BGP системата.

•Инициализира необходимите ресурси,

•Нулира ConnectRetry таймера,

•Инициира TCP връзка и започва очакване на такава, която може да се инициира от отдалечен съсед.

И как работи...



Connect

•BGP очаква установяването на TCP връзка.

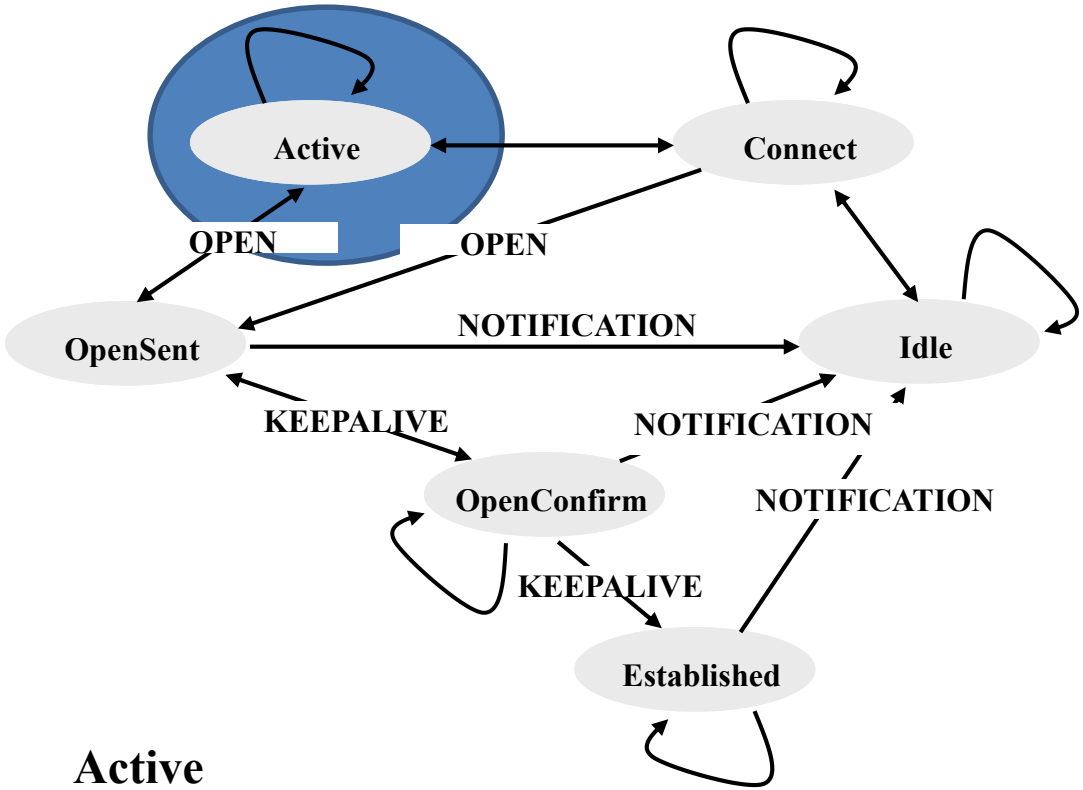
•Ако е успешно, се преминава в състояние **OpenSent** (изпратено е OPEN съобщение).

•Ако е неуспешно, се преминава в състояние **Active**.

•Ако таймерът ConnectRetry изтече, остава се в същото състояние **Connect**, таймерът се нулира и отново се инициира TCP връзка.

•В случай на настъпване на друго събитие се преминава в състояние **Idle**.

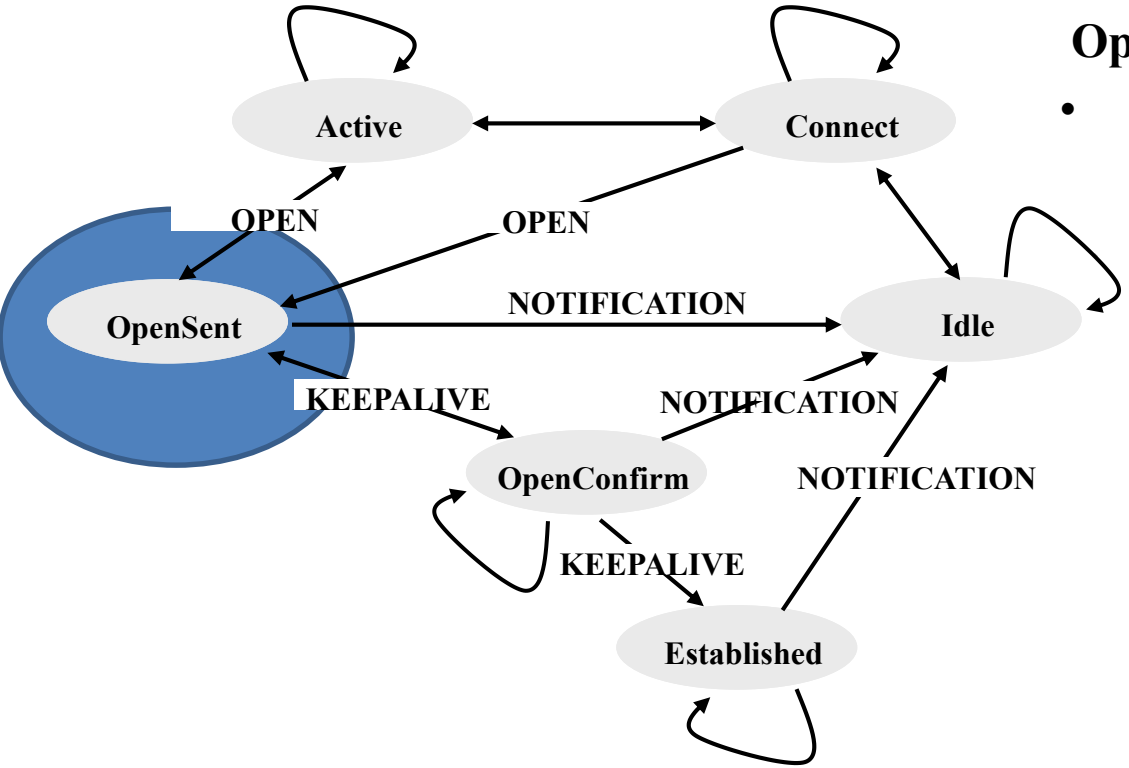
И как работи...



Active

- BGP се опитва да започне нова TCP връзка със съсед.
- Същевременно се очаква връзка, която може да се инициира от отдалечен маршрутизатор.
- Ако състоянията преминават между **Connect** и **Active**, това е индикация за проблем в TCP комуникацията - некоректна IP конфигурация, прекалено много TCP препредавания поради лошо качество на връзката...

И как работи...

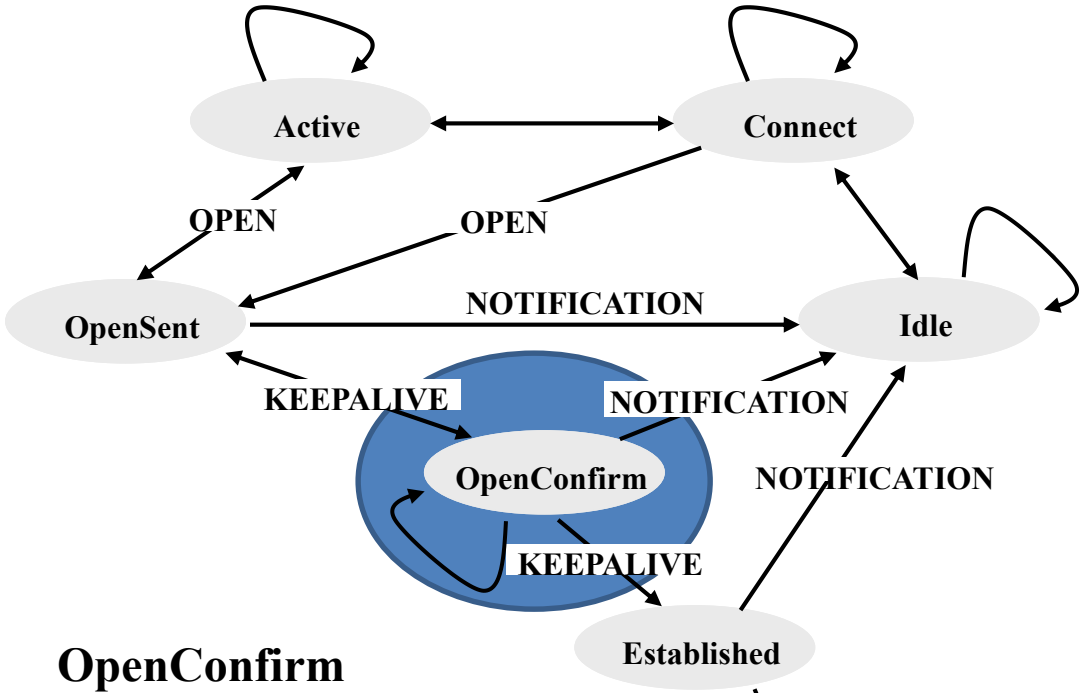


OpenSent

- Полученото съобщение OPEN се проверява за грешки (некоректен номер на версия или AS)
  - При грешки- NOTIFICATION съобщение и се преминава обратно в състояние **Idle**.
  - без грешки - KEEPALIVE съобщения и нулира KeepAlive таймера. В този момент се извършва договарянето на стойноста на Hold таймера, като се избира по-малката от предложените. Ако тази стойност е 0, двата таймера (Hold и KeepAlive) не се нулират.

- По номера на AS на своя съсед се определя дали той е вътрешен(*iBGP*) или външен (eBGP).
- Ако TCP връзката се разпадне, преминава се обратно към **Active** състояние.
- При наличие на други грешки (изтичане на Hold таймера), BGP изпраща NOTIFICATION съобщения със съответен код и се връща в състояние **Idle**.

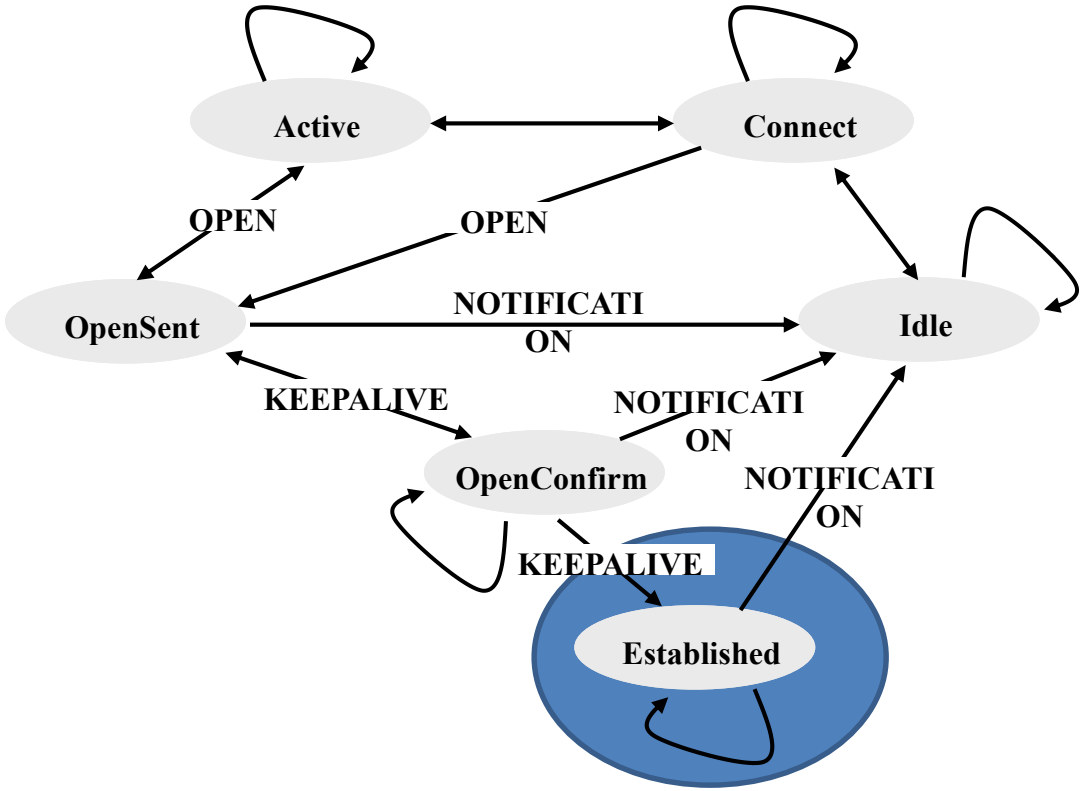
И как работи...



OpenConfirm

- В това състояние BGP очаква KEEPALIVE съобщение. Ако се получи такова, се преминава в състояние **Established**, като договарянето със съседа е завършило. При получаването се рестартира Hold таймера.
- Ако се получи NOTIFICATION съобщение, преминава се обратно в **Idle** състояние. При разпадане на TCP връзката или ако администраторът спре BGP процеса, състоянието се връща в **Idle**. В случаи на всяко друго събитие, BGP изпраща NOTIFICATION съобщение с код за грешка *Finite State Machine Error* и отново преминава в **Idle** състояние.

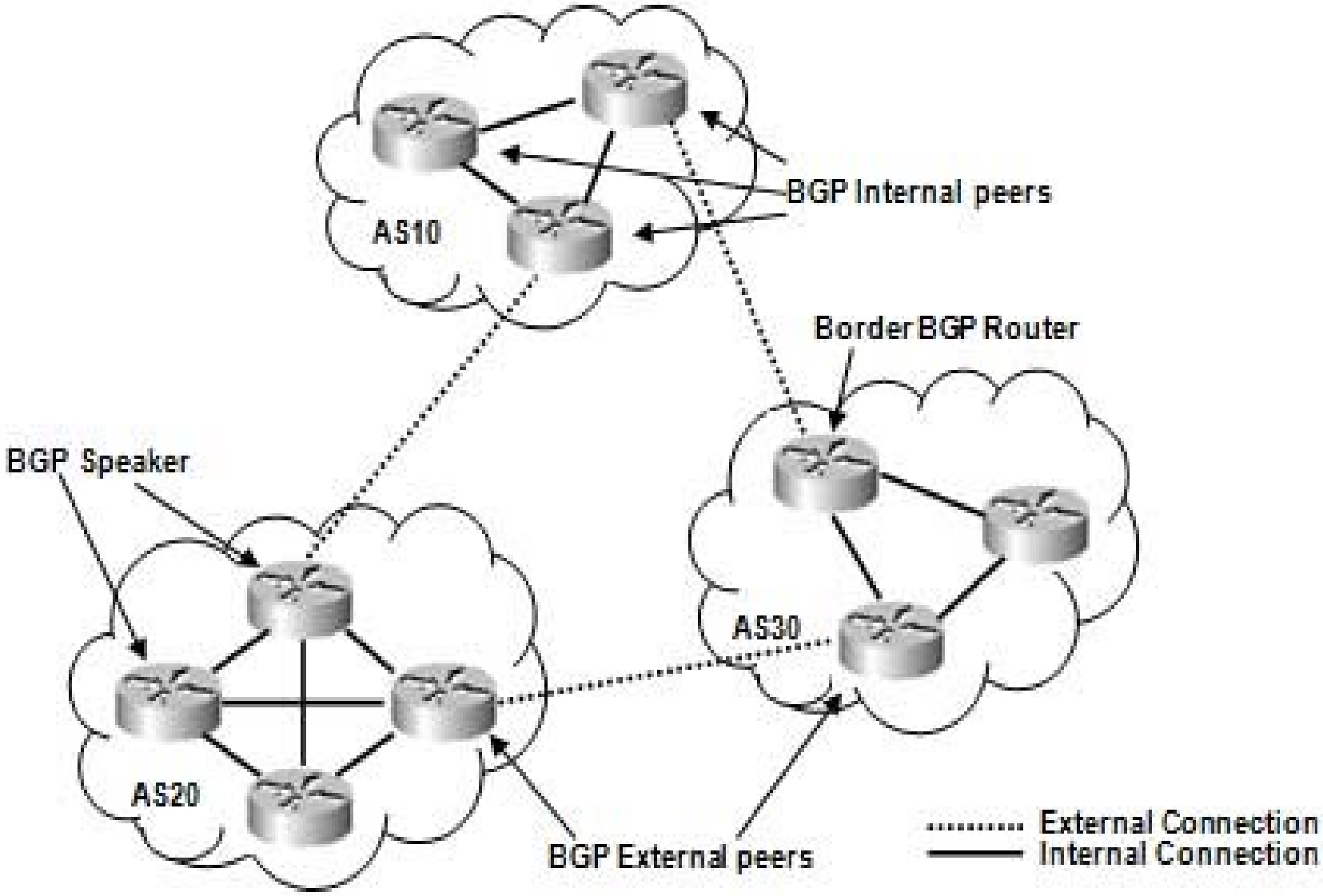
И как работи...



Established

- Последното състояние на процеса на установяване на сесия със съсед.
- В случай на получаване на NOTIFICATION съобщение, състоянието преминава обратно в **Idle**.

BGP топология



Таймери

- **ConnectRetry** - само когато BGP се опитва да осъществи TCP връзка със съсед и определя колко често е започвала TCP връзка;
- **Hold** – Времето, изтекло след получаването на последно UPDATE или KEEPALIVE съобщение, след което съседът се счита за отпаднал от мрежата;
- **KeepAlive** – За генериране на KEEPALIVE съобщения;
- **MinASOriginationInterval** –за регулиране на честотата на изпращане на съобщения за настъпили вътрешни промени в AS;
- **MinRouteAdvertisementInterval** – за регулиране на честотата на изпращане на съобщения за промяна на даден маршрут.

Маршрутите

- Routing Information Bases (RIB):
  - **Adj-RIB-In –** пътища, получени от други BGP speakers
  - **Loc-RIB-** пътища, използвани от локалните BGP speakers.
- Forward Information Base (FIB):
  - следващия хоп на маршрутите
  - **Adj-RIB-Out -** пътища, които трябва да се изпратят към други BGP маршрутизатори

Определяне на маршрута

- Първата фаза определя степента на предпочитание на маршрута:
  - ако маршрутът е от локалната AS - на базата на атрибута LOCAL_PREF или от предварително конфигурираната политика;
  - ако маршрутът не е от локалната AS - само от конфигурираната политика.

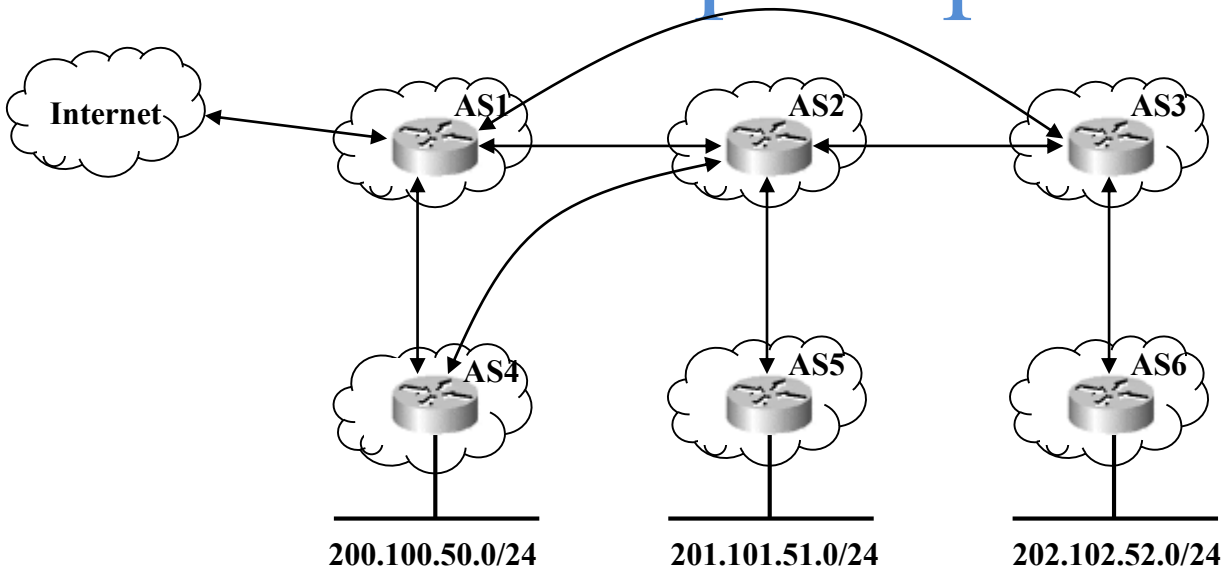
Определяне на маршрута

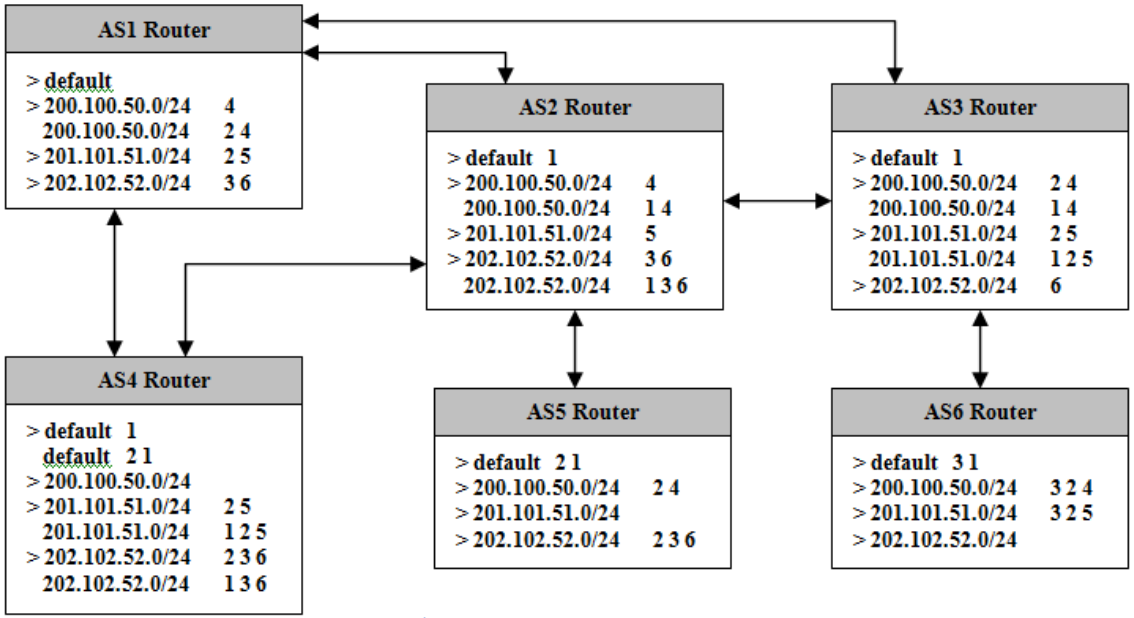
- Втората фаза се състои в определяне на маршрути, подходящи за добавяне в Loc-RIB базата:
  - Избира се маршрут или с най-висока степен на предпочитане или единичен път към дестинацията.
  - Ако се получат няколко пътища към една и съща дестинация се прилагат следните критерии:
    - Избира се маршрут с най-малка стойност на атрибута MULTI_EXIT_DISC;
    - Избира се маршрут с най-малка цена (вътрешна дистанция);
    - Ако има няколко маршрута с еднаква цена, то:
    - Избира се маршрут, изпратен от BGP speaker в съседната АС с най-малък идентификатор;
  - В противен случай, избира се маршрут, изпратен от BGP speaker, чийто ID е най-малък;

Определяне на маршрута

- Третата фаза на процеса определя разпространението на маршрутите, при настъпването на някое от следните условия:
  - При промяна на локалната база Loc-RIB;
  - При промяна на локално генерираните маршрути, които са научени извън протокола BGP;
  - При осъществяване на нова връзка между BGP speakers.

Пример

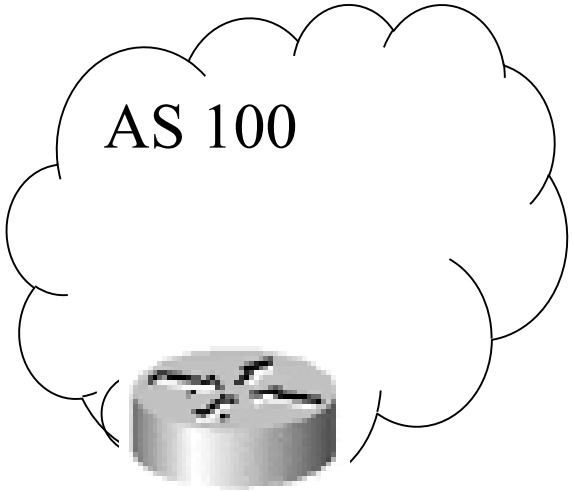

Internet
AS1
AS2
AS3
AS4
AS5
AS6
200.100.50.0/24
201.101.51.0/24
202.102.52.0/24


AS1 Router
> default
> 200.100.50.0/24 4
200.100.50.0/24 2 4
> 201.101.51.0/24 2 5
> 202.102.52.0/24 3 6
AS2 Router
> default 1
> 200.100.50.0/24 4
200.100.50.0/24 1 4
> 201.101.51.0/24 5
> 202.102.52.0/24 3 6
202.102.52.0/24 1 3 6
AS3 Router
> default 1
> 200.100.50.0/24 2 4
200.100.50.0/24 1 4
> 201.101.51.0/24 2 5
201.101.51.0/24 1 2 5
> 202.102.52.0/24 6
AS4 Router
> default 1
default 2 1
> 200.100.50.0/24
> 201.101.51.0/24 2 5
201.101.51.0/24 1 2 5
> 202.102.52.0/24 2 3 6
202.102.52.0/24 1 3 6
AS5 Router
> default 2 1
> 200.100.50.0/24 2 4
> 201.101.51.0/24
> 202.102.52.0/24 2 3 6
AS6 Router
> default 3 1
> 200.100.50.0/24 3 2 4
> 201.101.51.0/24 3 2 5
> 202.102.52.0/24

Конфигуриране на BGP

/usr/local/etc/bgpd.conf

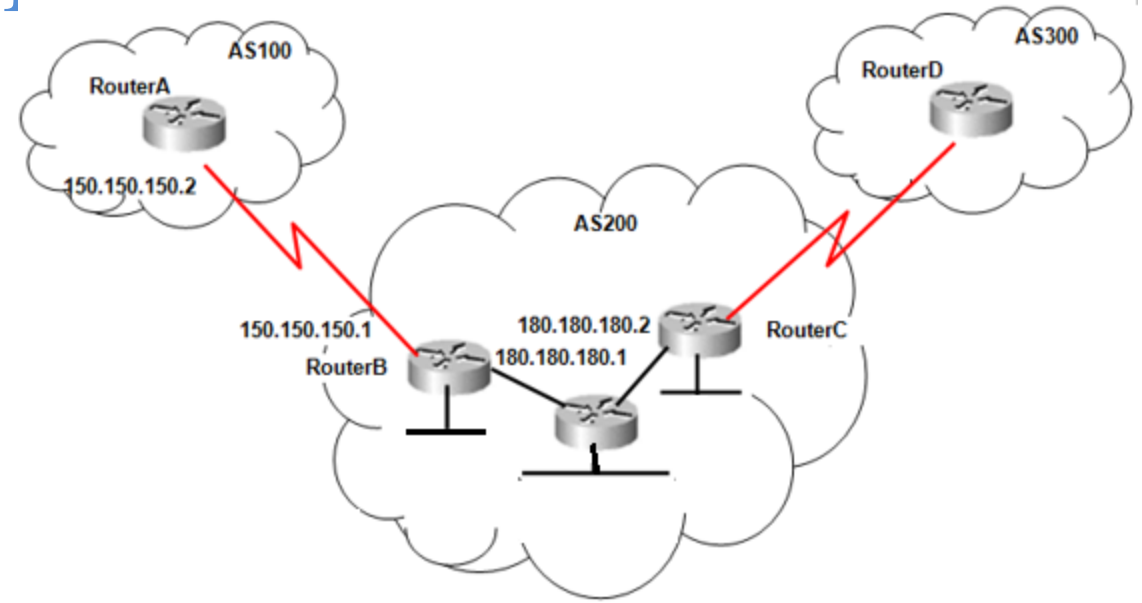
```
router bgp 100
  bgp-router 192.168.0.11
```



По подразбиране маршрутизаторът използва като идентификатор най-големия IP адрес измежду всичките си интерфейси.

Ако е необходимо използването на друг идентификатор, той се задава с **bgp-router.**

Дефиниране на съседите




```
RouterA#
  router bgp 100
    neighbor 150.150.150.1

RouterB#
  router bgp 200
    neighbor 150.150.150.2 remote-as 100
    neighbor 180.180.180.2 remote-as 200

RouterC#
  router bgp 200
    neighbor 180.180.180.1 remote-as 200
```

При промяна в конфигурацията

- При промяна на BGP конфигурацията на маршрутизатор е необходимо да се рестартират (*reset*) TCP връзките със съседите, за могат новите параметри да се приемат.

```
router bgp 100
  neighbor 199.199.199.1 shutdown
```

- В този случай маршрутизаторът ще затвори връзката със своя съсед 199.199.199.1 и ще започне нова процедура за установяване на BGP сесия с последно дефинираните параметри.

Анонсиране на мрежи

```
router bgp 100
    network 200.100.50.0
    network 192.168.0.0 255.255.0.0
```

- Първата мрежа е 200.100.50.0/24
- Втората мрежа 192.168.0.0/16 се явява супермрежа (получена от сумаризация на няколко мрежи), което е указано от 16-битовата мрежова маска.

Редистрибутиране

- Редистрибутират се източниците:
  - **Connected**
  - **Static**
  - **RIP**
  - **OSPF**
- Метриката се задава в диапазона 0-16777214:

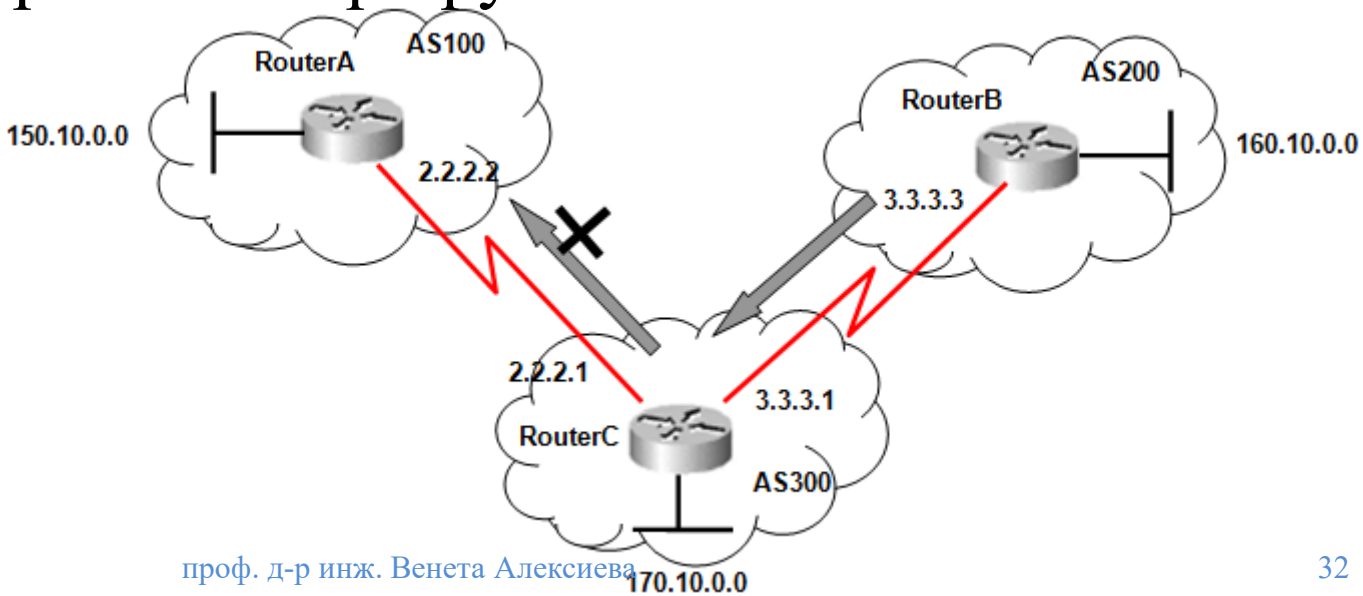
```
router bgp 100
  redistribute rip metric 10
  redistribute ospf metric 16
```

или

```
router bgp 100
    default-metric 1
    redistribute rip
    redistribute ospf
  redistribute bgp metric
```

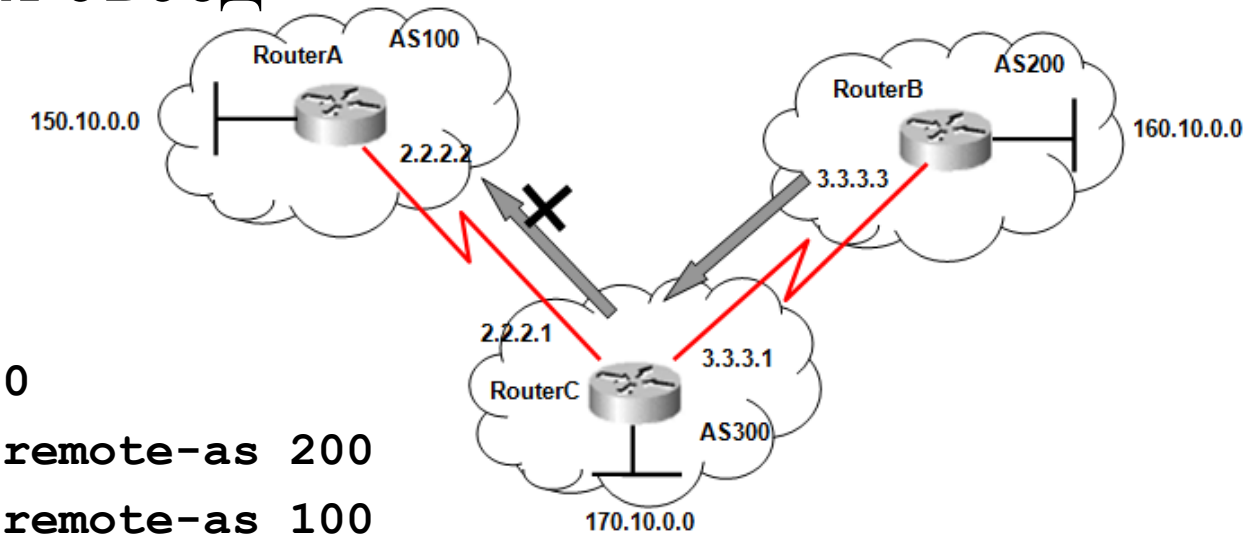
Филтриране на BGP обмен

- Изпращането и получаването на BGP съобщения може да бъде контролирано по няколко начина:
  - базирано на филтриране на мрежи,
  - филтриране на маршрути,
  - чрез групиране на маршрути.



Филтриране на BGP обмен

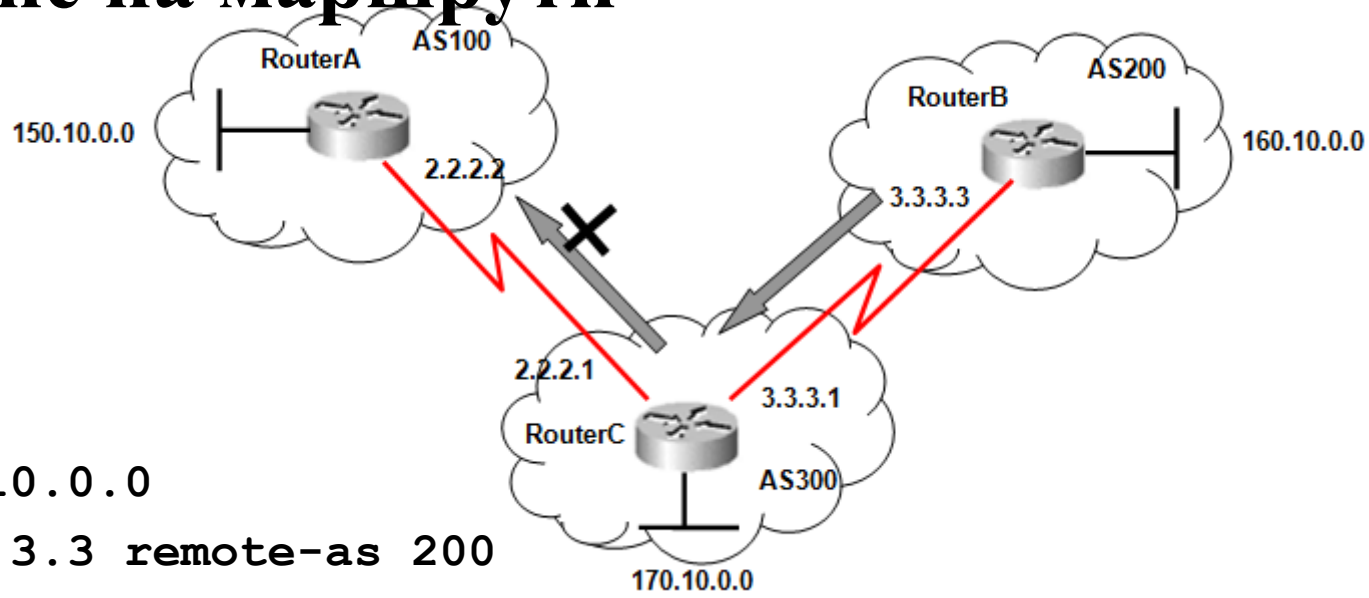
- **Филтриране на мрежи -**чрез дефиниране на списък за достъп (*access-list*) и прилагането му към съответния съсед




```
RouterC#
  router bgp 300
    network 170.10.0.0
    neighbor 3.3.3.3 remote-as 200
    neighbor 2.2.2.2 remote-as 100
    neighbor 2.2.2.2 distribute-list filter out

access-list filter deny 160.10.0.0/16
access-list filter permit any
```

Филтриране на BGP обмен

- **Филтриране на маршрути**




```
RouterC#
  router bgp 300
    network 170.10.0.0
    neighbor 3.3.3.3 remote-as 200
    neighbor 2.2.2.2 remote-as 100
    neighbor 2.2.2.2 filter-list newfilter out

ip as-path access-list newfilter deny ^200$
ip as-path access-list newfilter permit .*
```

Регулярни изрази- символи

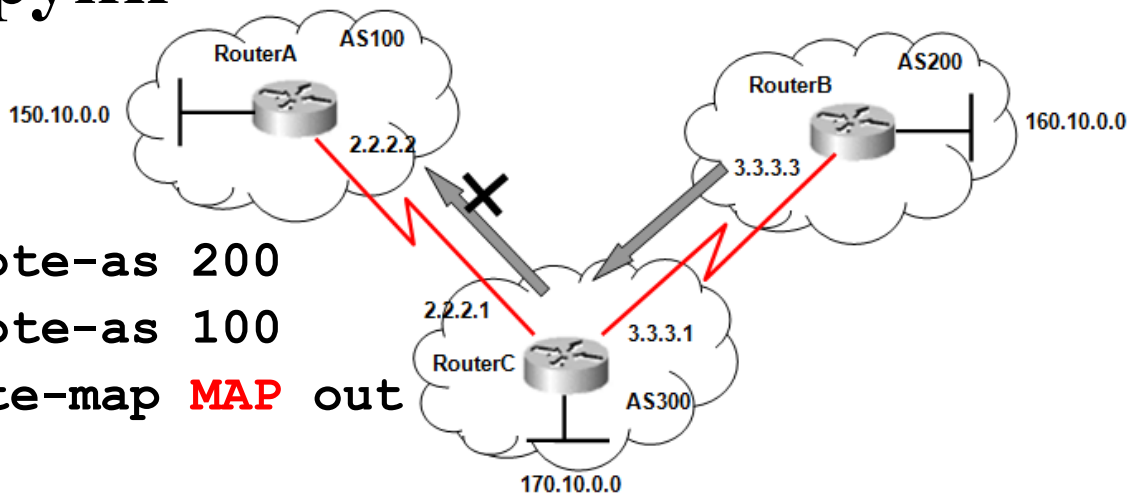
Символ	Значение
.	Съвпада с всеки символ
*****	Съвпада с 0 или повече последователни символа
+	Съвпада с 1 или повече последователни символа
?	Съвпада с 0 или 1 следващ символ
^	Съвпада с началото на входния низ
$	Съвпада с края на входния низ
_	Съвпада със символите ‘,’, ‘{‘, ‘}’, ‘ ‘, началото или края на входния низ
[	Начало на последователност от символи
]	Край на последователност от символи
-	Задава диапазон от символи

Примери

Регулярен израз	Значение
a*	Всяко наличие на символа **a**, включително и без него
a+	Поне едно наличие на символа **a**
ab?a	Съвпада с последователност **aa** или **aba**
_100_	През AC100
^100$	Произход е AC100
^100.*	Идващо от AC100
^$	Произход от същата AC

Филтриране на BGP обмен

- **Филтриране по групи**

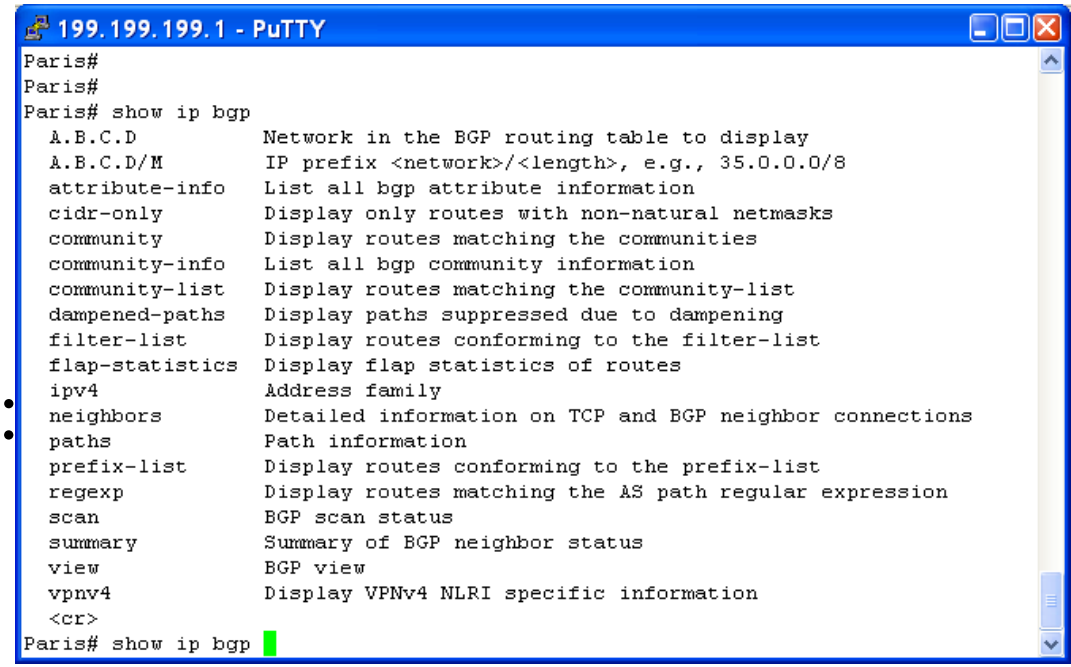



```
RouterC#
  router bgp 300
    network 170.10.0.0
    neighbor 3.3.3.3 remote-as 200
    neighbor 2.2.2.2 remote-as 100
    neighbor 2.2.2.2 route-map MAP out

route-map MAP deny 10
  match ip address denyfilter
route-map MAP permit 20
  match ip address permitfilter

access-list denyfilter deny 160.10.0.0/16
access-list permitfilter permit any
```

Стартиране и наблюдение

```
# zebra -d
# bgpd -d
```

telnet на порт 2605:




```
show ip bgp
show ip bgp neighbors
show ip bgp summary
```

Въпроси ?

Благодаря за вниманието !