17. Маршрутен протокол OSPF

Общи положения

Open Shortest Path First (**OSPF**) е динамичен протокол за маршрутизация.

Той е протокол със следене на състоянието на връзката (link-state routing protocol)

Попада в групата на протоколите за вътрешна маршрутизация (interior gateway protocols - IGP), т.е в рамките на една автономна система – autonomous system (AS).

OSPF Version 2 за IPv4 е дефиниран в RFC 2328 (1998).

OSPF е най-широко приложимия IGP в големите корпоративни мрежи; IS-IS, е по-приложим в големи мрежи на доставчици на мрежови услуги.

Optimized Link State Routing Protocol (**OLSR**) е оптимизиран за мобилни и др. безжични ad-hoc мрежи.

Характеристики на OSPF

OSPF изгражда маршрутни таблици по destination IP адресите в IP пакетите.

Поддържа VLSM, респ. CIDR.

OSPF разпознава промени в топологията като отпадане на връзки много бързо и се конвергира, за секунди, в нова топология, без зацикляния (loop-free).

За целта всеки OSPF рутер събира информация състоянието на връзките (link-state information), за да изгради цялостната топология на мрежата, състояща се от области (*areas).*

Характеристики на OSPF

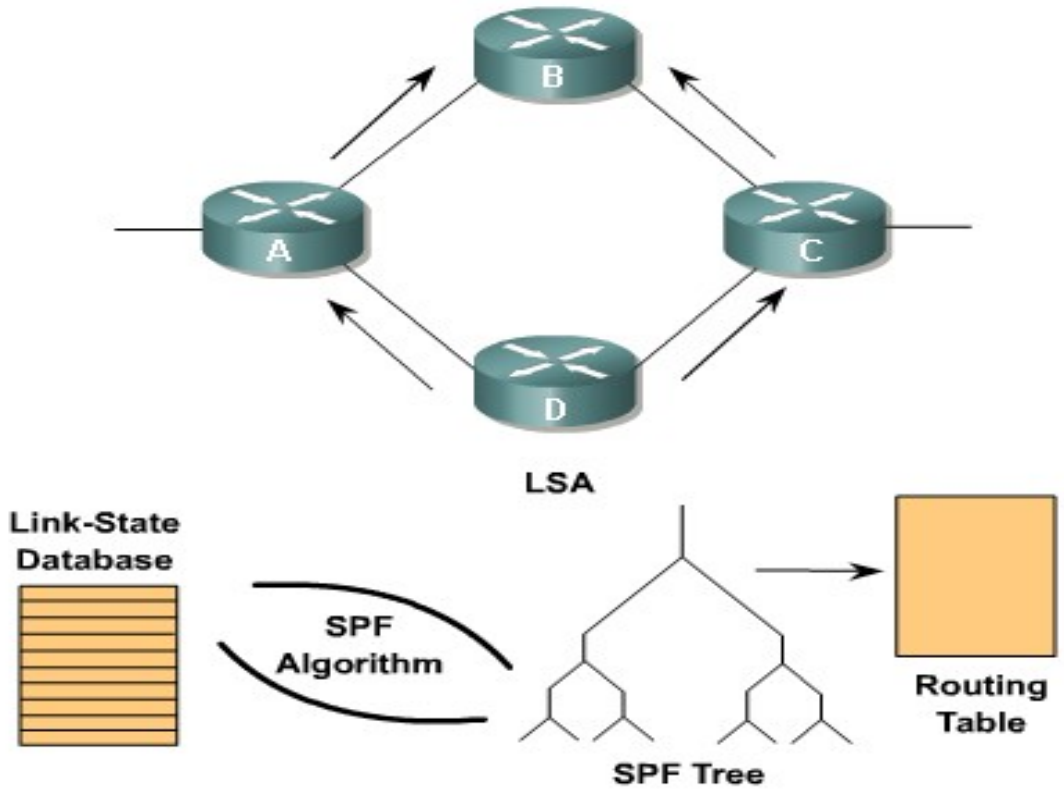
Топологията - граф, възлите на който са рутерите, а клоните – комуникационните линии между рутерите.

На клоните се присвояват стойности, които са обратно пропорционални на скоростта на линиите. Т.е по-бързата връзка е с по-малка стойност.

От тази пълна топология се изчислява “дървото на най-късия път” (shortest path tree), по най-ниската стойност, за всеки маршрут по алгоритъма на (Dijkstra's algorithm).

Информацията за състоянието на връзките се поддържа във всеки рутер под формата база от данни засъстояние на връзките (link-state database - LSDB): мрежовата топология, пълния граф на мрежата.

OSPF. Принцип на работа.



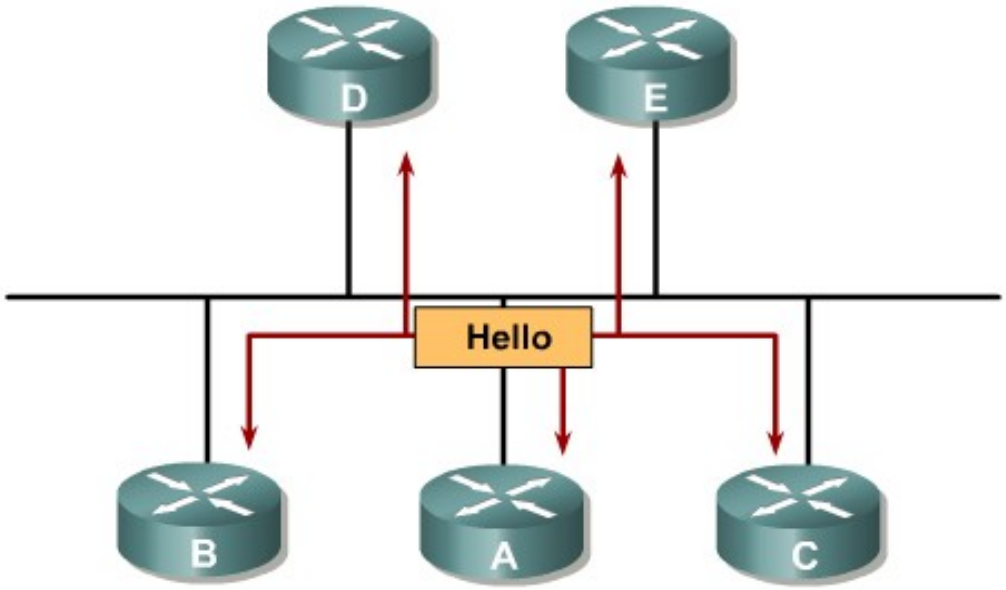
OSPF съседство

OSPF първо установява съседство (adjacency) между директно свързани рутери с цел обмен на маршрутна информация.

Hello протоколът отговаря за установяване и поддържане на съседство.

Hello пакети се изпращат периодически през всички активни интерфейси.

Съседство в LAN



Съседство между рутерите

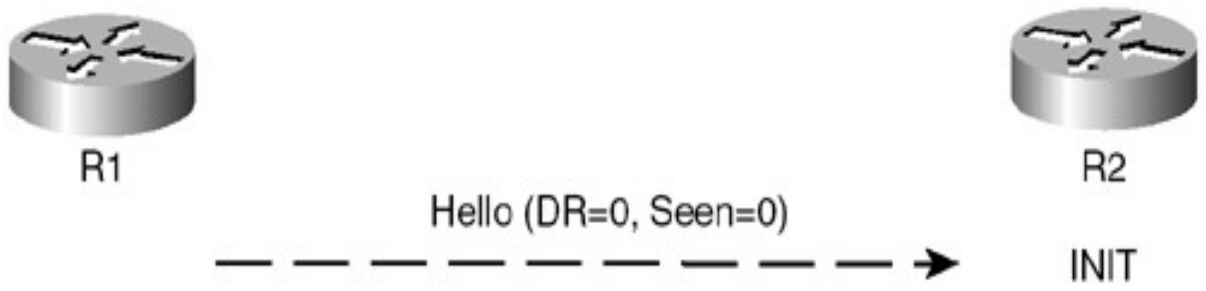
Рутерите в Ethernet локална мрежа избират *designated router* (DR) и резервен, *backup designated router* (BDR), който действа като концентратор с цел намаляване на трафика между рутерите.

OSPF изпращат "hello" пакети и обновления за състоянието на връзките (link state updates) и по unicast, и по multicast.

Състояния на OSPF. Down.

DOWN. От съседа все още не се получава информация.

Състояния на OSPF. Init.

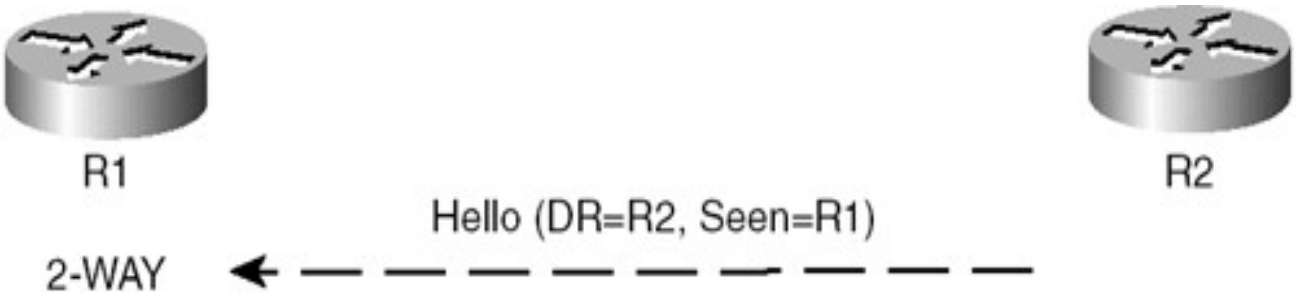


Init е еднопосочен (one-way) Hello.

R1 изпраща Hello пакет.

R2 обявява one-way, защото не вижда своя router ID в Hello пакета.

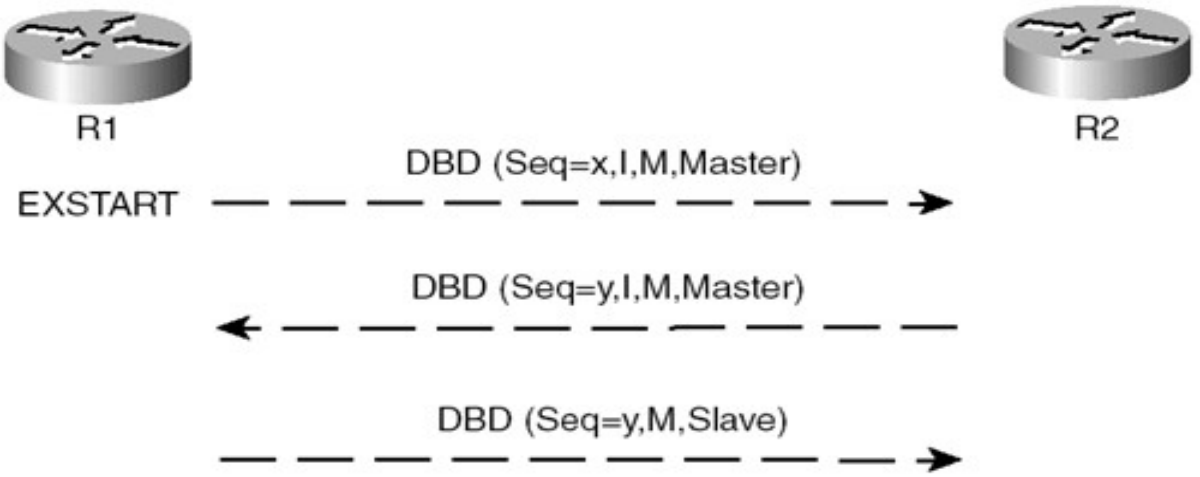
Състояния на OSPF. 2-Way.



2-way състояние при установяване на двупосочни комуникации – начало на OSPF съседство. Избор на DR и BDR.

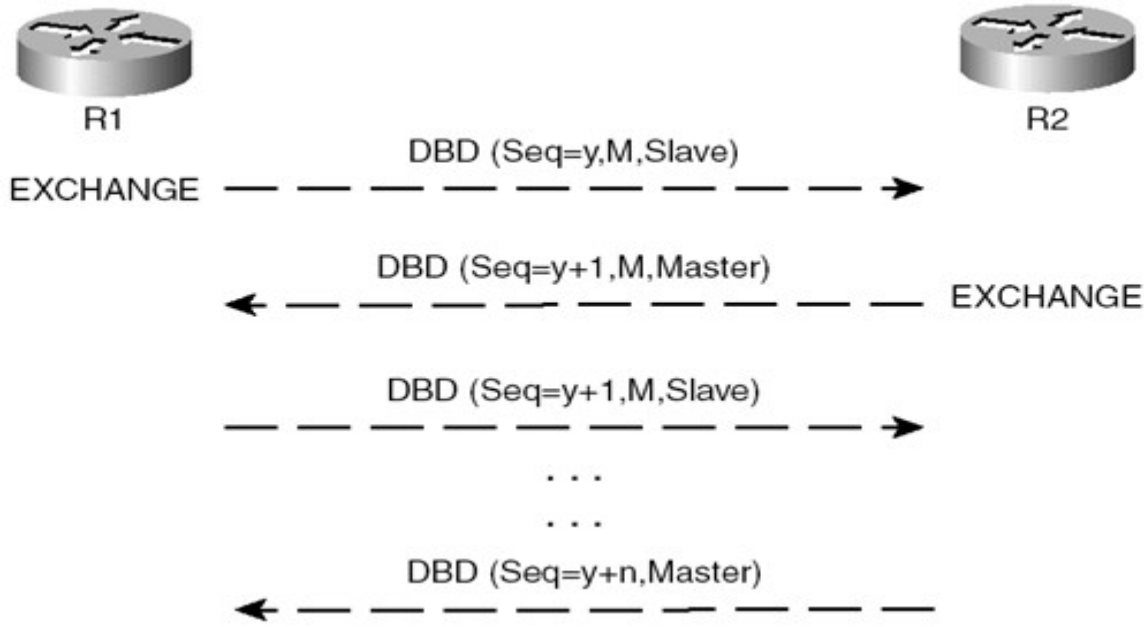
router ID на R2 е по-голям, избран е за DR. (Designated Router или Master)

Състояния на OSPF. Exstart.



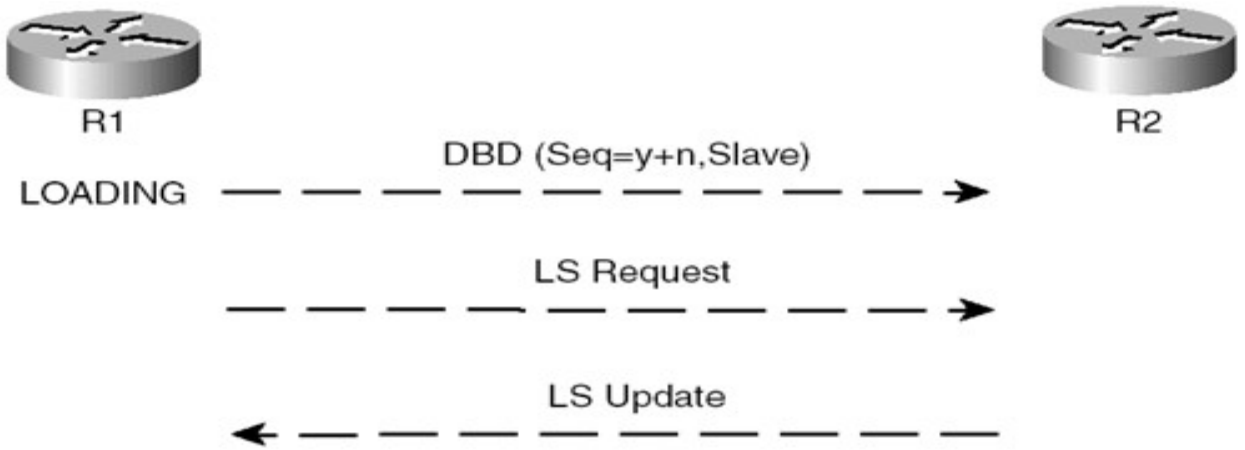
Exstart инициализира процеса на синхронизация на топол. Б.Д. (DBD). Избират се Master и slave (или DR и BDR), Seq0 (началния) за DBD обмена.

Състояния на OSPF. Exchange.



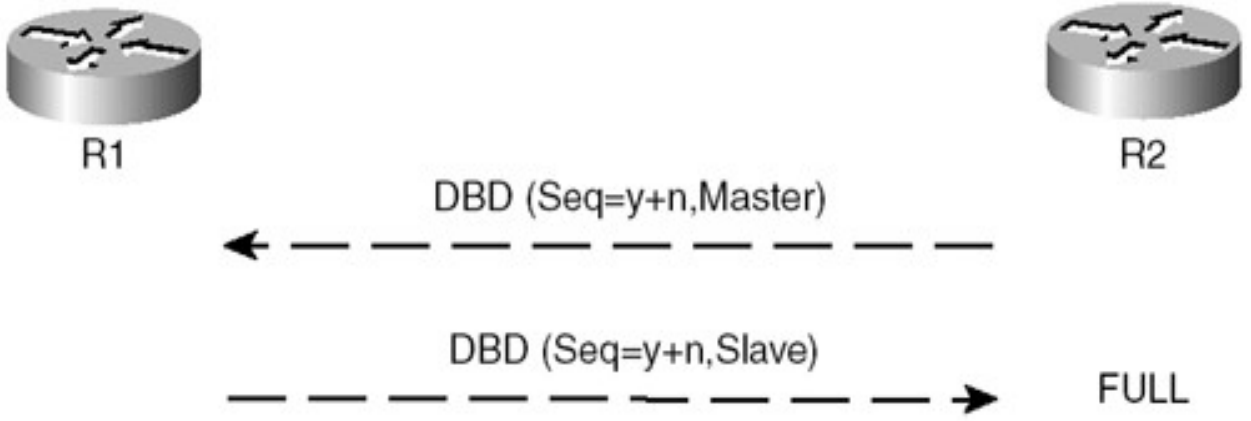
Exchange. Рутерът описва цялата си link-state database чрез DBD пакети. Всеки DBD Seq се потвърждава.

Състояния на OSPF. Loading.



Loading. LS request се изпращат да заявят по-нови състояния на връзките (LSA - link state advertisement), които не са били получени по време на exchange.

Състояния на OSPF. Full.



Full. Между OSPF съседите е обменена пълната информация.

R1 и R2 са си обменили цялата база от данни и с във Full.

Характеристики на OSPF

Идентични копия на LSDB периодически се обновяват чрез “наводняване” на всички рутери в дадена OSPF област (area).

По принцип area 0 е опорната - "backbone" област на OSPF мрежата, а OSPF областите с други номера обслужват други части от корпоративната мрежа.

Всяка допълнителна OSPF area трябва да има директна или виртуална връзка към backbone OSPF area.

По строга дефиниция областите в OSPF се номерират като IP адресите. Опорната, например е 0.0.0.0. Но това правило не се спазва стриктно. Cisco IOS, например, допуска едноцифрени номерации.

Маршрутизацията между областите се осъществява през опорната.

Multicast адреси

Multicast адресите 224.0.0.5 (всички SPF/link state рутери, AllSPFRouters) и 224.0.0.6 (всички Designated рутери, AllDRouters) са резервирани за OSPF (RFC 2328).

OSPF не се базира на TCP (BGP) или UDP (RIP).

Директно използва IP (IP protocol 89).

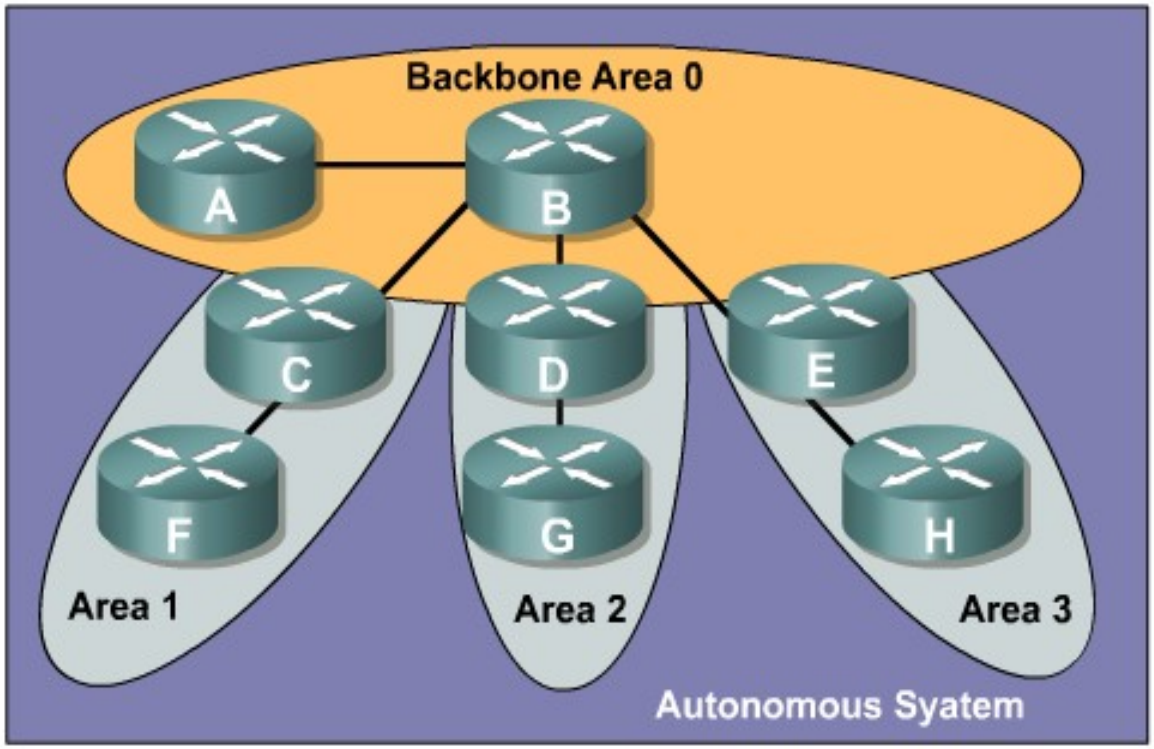
OSPF има собствени механизми за откриване и корекция на грешките.

Сигурност на операциите

Има възможност за защита на операциите в OSPF чрез размяна на пароли между рутерите.

В явен текст или Message Digest 5 (MD5) за автентикация на съседни рутери, преди да са образували съседство и са започнали да приемат link-state advertisements (LSA).

Области в OSPF



Области

OSPF мрежата е разделена на области (*areas*), които се означават с 32-битови идентификатори.

Тези идентификатори се записват подобно на IPv4 адреси. Но те не са IP адреси и спокойно могат да се дублират с истинските.

Повечето реализации на OSPF позволяват да се пише и в друг формат.

Например, area 1 в повечето случаи ще бъде възприета правилно като 0.0.0.1, но за някои ще бъде 1.0.0.0. Затова се препоръчва да се ползва десетично-точковата нотация.

Области

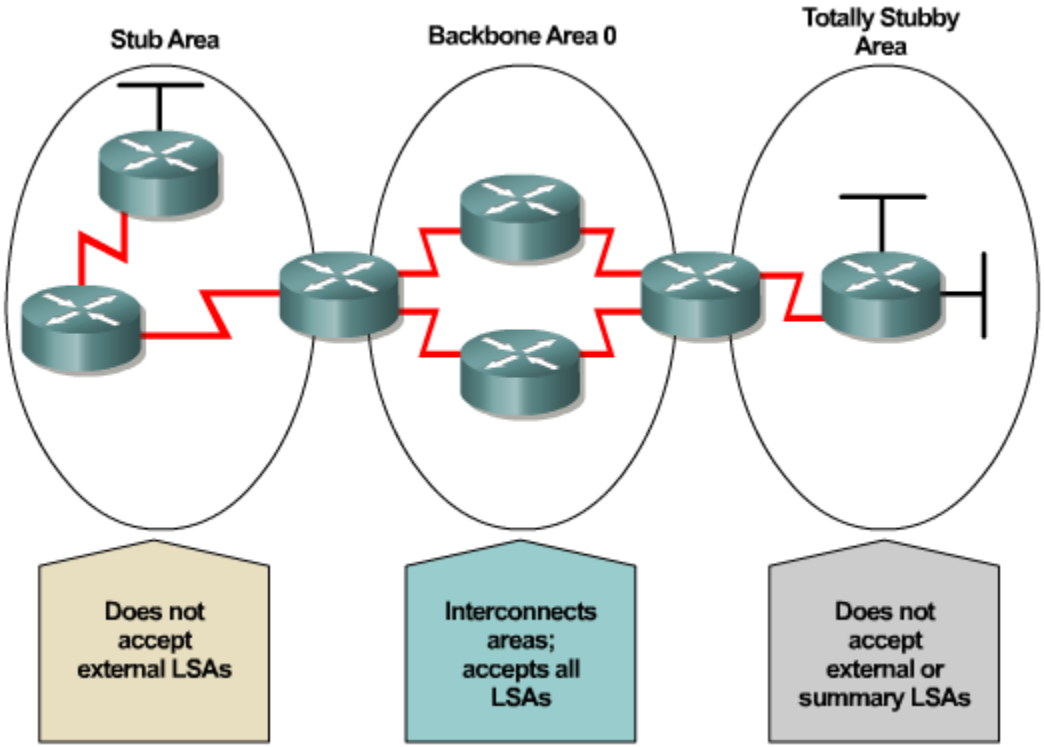
Областите са логически групи от хостове и мрежи, включително и техните рутери, които имат интерфейси към тези мрежи.

Всяка област поддържа собствена link state database, чиято информация може да бъде представена в обобщен вид (summary) пред останалата част от мрежата.

Така топологията на областта ще е неизвестна извън нея.

И ще се намали служебния трафик за маршрутизация в рамките на OSPF мрежата.

Типове области



virtual link

Всички OSPF области трябва да бъдат свързани към опорната.

Но освен по физически, тази връзка може да се осъществи и по виртуален канал.

Например, област 0.0.0.1 има физически канал до 0.0.0.0. Но 0.0.0.2 няма директна връзка с опорната, но има с 0.0.0.1.

Област 0.0.0.2 по виртуален канал през *транзитната област* 0.0.0.1 може да достигне опорната.

Транзитната област не трябва да бъде stub по никакъв начин.

Backbone area

Опорната област (backbone area - *area 0.0.0.0*) е ядрото на OSPF мрежата.

Всички други области са свързани към нея.

Маршрутизацията между областите се осъществява чрез гранични рутери, имащи интерфейси и в опорната, и в някоя от другите области.

Тези рутери се наричат Area Border Routers (ABRs).

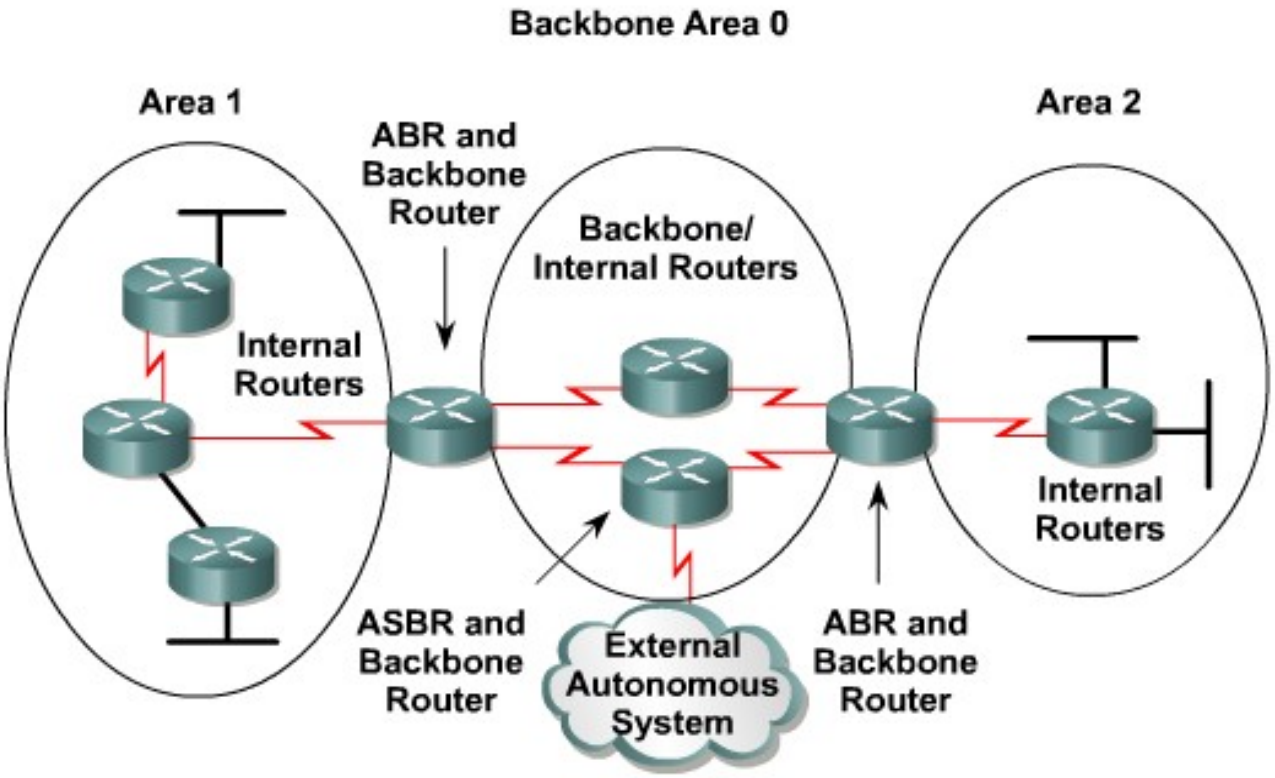
Backbone area

От area 0.0.0.0 се реализира и връзката на OSPF мрежата с външния свят с помощта на рутер, който има външен интерфейс - Autonomous System Border Router (ASBR).

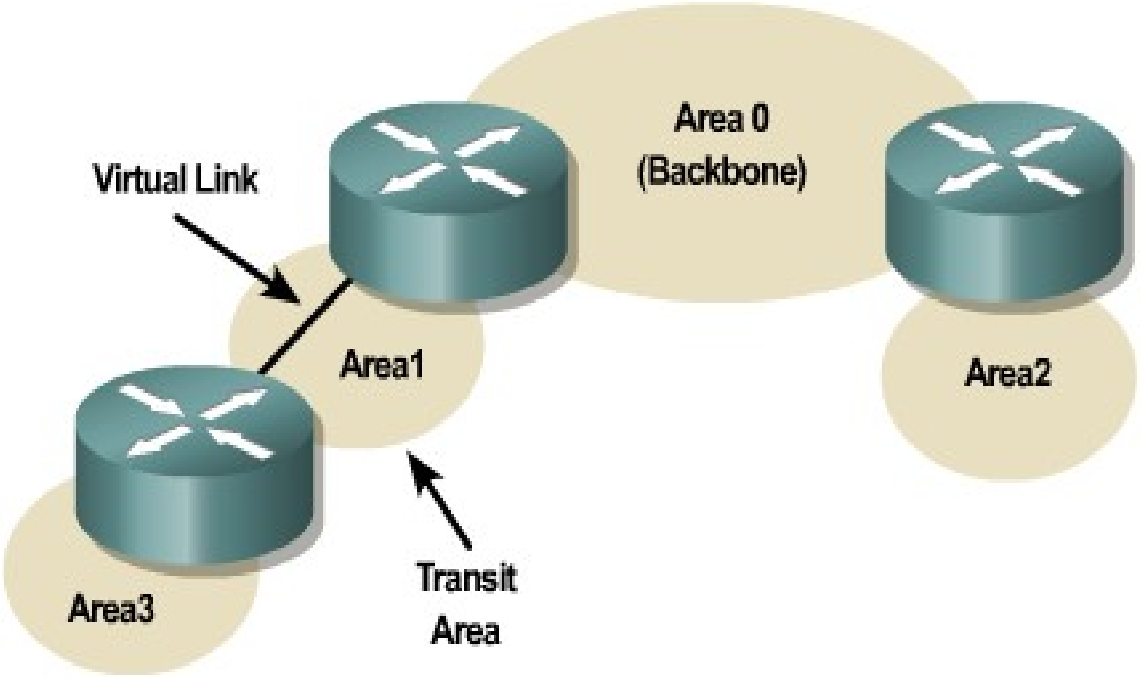
По наследство на OSPF мрежата с всичките й области се казва “автономна система”.

Това няма нищо общо с видимите през Internet автономни системи, съгласно RFC1996, които могат да съдържат множество OSPF, RIP и др. мрежи.

Междуобластни рутери



virtual link



Stub area

stub area е област, която не получава външни за OSPF мрежата маршрути с изключение на маршрута по подразбиране за навън (default route)

Но областта получава маршрути от другите области (*inter-area*).

Дефинирането на такава област е полезно, ако целият излаз към Internet минава през рутери (ASBRs) в Area 0.0.0.0, и трябва да се избере най-добтия път до друга област.

Totally stubby area

totally stubby area (TSA) е дефиниция, въведена от Cisco (IOS)

Подобна е на stub area, но освен външни маршрути не позволява и *summary* маршрути, т.е *inter-area* (IA), междуобластните маршрути не се обобщават в TSA.

Единственият начин трафикът да се рутира извън такава област е чрез default route.

TSA - в периферията на мрежата - рутери от по-нисък клас.

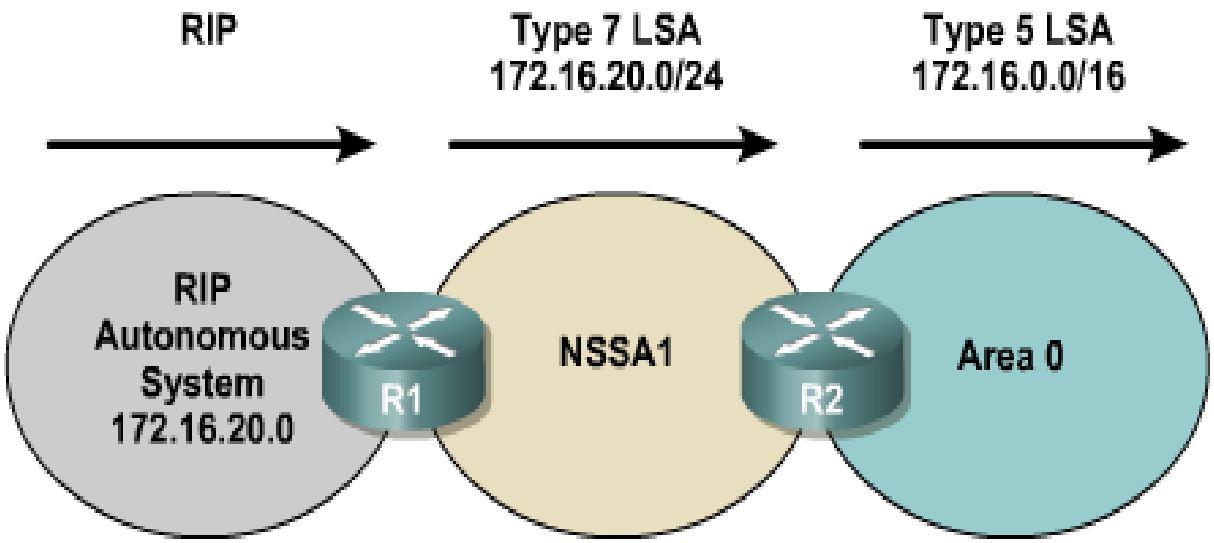
TSA може да има повече от един ABR. Трафикът извън областта се праща през ABR с най-ниска метрика (“най-близкия").

Not-so-stubby area

not-so-stubby area (NSSA) също е stub area, но която може да импортира външни за OSPF мрежата маршрути и да ги изпраща към опорната област.

Но тя не е в състояние да получава външни маршрути от опорната или други области.

NSSA


RIP
Type 7 LSA
172.16.20.0/24
Type 5 LSA
172.16.0.0/16
RIP
Autonomous
System
172.16.20.0
R1
NSSA1
R2
Area 0

OSPF метрика

Метриката в OSPF е стойността на пътя - маршрута (*path cost)*.

Тя се определя от скоростта (bandwidth), зададена на интерфейса, водещ към съответния маршрут.

Cisco работи с метрика 10^8/bandwidth.

Тази базова стойност 10^8 по подразбиране може да бъде променена.

Скорости от 1000 Mbps (1 Gbps) вече преобладават, а 10 Gbps бурно навлизат в опорните (backbone) линии.

При 10^8, 100 Mbps ще е със стойност 1, а 10 Mbps - 10 и т.н. Но за по-бързи от 100Mbit/s, cost <1.

Типове OSPF метрика

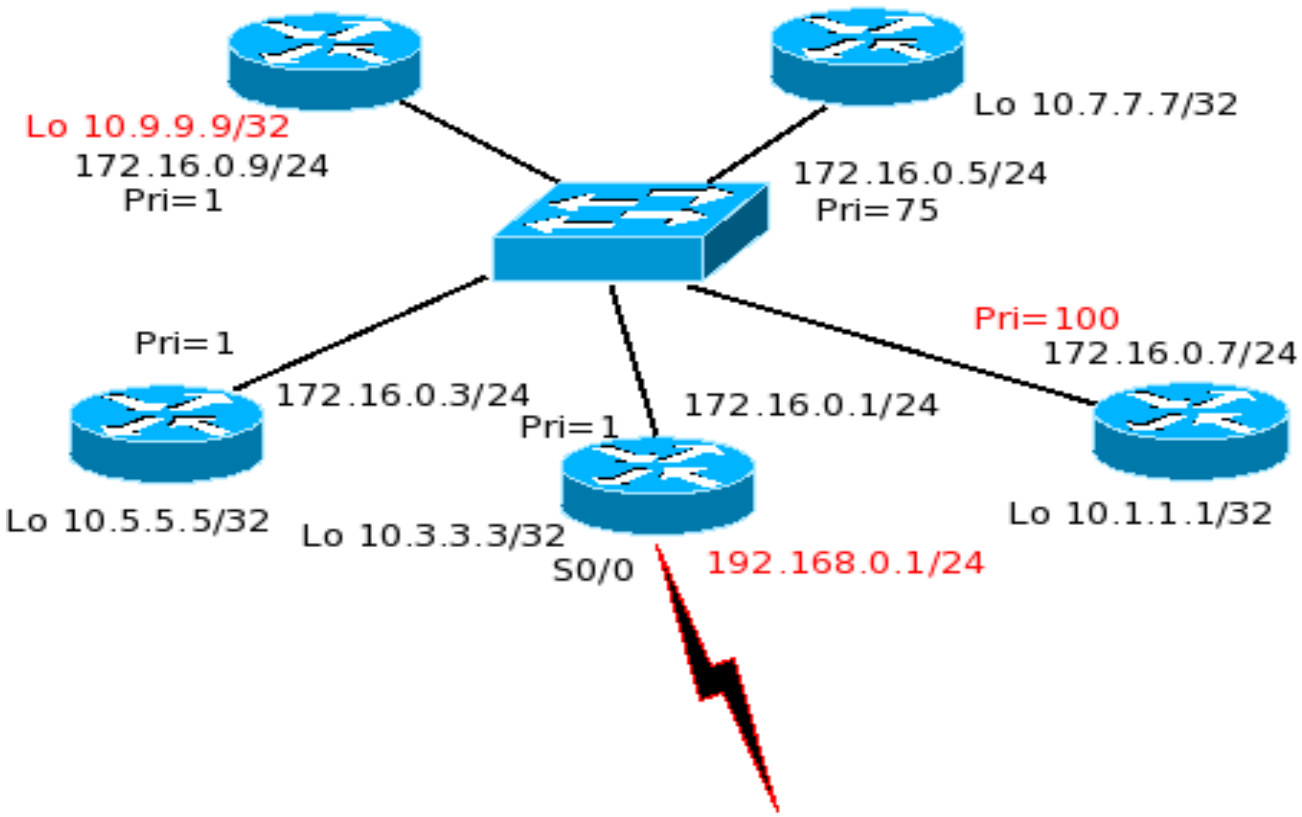
Описаната в предишния слайд метрика е директно приложима само, ако е от един и същи тип.

Имаме четири типа метрика, принадлежащи на 4 типа маршрути.

Редът на предпочитане на един от тези 4 типа е както следва:

1. Intra-area – вътрешни за областта;
2. Inter-area – между областите;
3. External Type 1 – включва стойността на външния маршрут плюс сумата от стойностите на вътрешните пътища до ASBR, който е анонсирал маршрута;
4. External Type 2 - включва единствено стойността на външния маршрут.

Designated router



Designated router

designated router (DR) е интерфейс на рутер избран между съществуващите в сегмент с множествен достъп и общодостъпно предаване (broadcast multiaccess), т.е LAN Ethernet.

DR не е определен тип OSPF рутер. Един от интерфейсите на даден физически рутер може да е DR, а друг - backup designated (BDR), трети да са non-designated.

При първоначално включване на OSPF рутерите в даден сегмент, след установяване на съседство всеки-с-всеки се избират DR и BDR по следните критерии:

Избор на DR и BDR

Рутерът, изпращащ Hello пакети с най-висок приоритет печели “състезанието” за DR, евентуално за BDR.

Ако два и повече рутера са с еднакъв най-висок приоритет, рутерът, изпращащ Hello пакети с най-висок RID (Router ID) “побеждава”.

RID е най-високият логически (loopback) IP адрес на рутера.

Ако такъв не е конфигуриран, избира се най-високия IP адрес, конфигуриран на активен интерфейс.

(напр. 192.168.0.1 е по-висок от 10.1.1.2).

Избор на DR и BDR

Рутер със следващ по-големина приоритет или RID става BDR.

Стойностите на приоритета варират от 0 - 254, колкото е по-голяма стойността, толкова по-големи са шансовете за успех в избора на DR или BDR.

В рутерите на Cisco приоритетът по подразбиране е 1. Т.е, ако администраторът не се намеси (не е конфигурирал и Loopback интерфейси), RID ще се определя по най-високия IP на активните интерфейси.

Желателно е да използваме приоритет и/или Loopback, за да наложим избора на най-мощните рутери за DR и BDR, които да комуникират с всички останали в сегмента.

Избор на DR и BDR

Ако OSPF рутер с по-висок приоритет стане онлайн след избора, той няма да стане DR или BDR, докато не отпадне някой от вече избраните.

Ако избран DR отпадне, BDR става новия DR. Провежда се избор на нов BDR.

Ако новият DR отпадне, а оригиналният DR същевременно се е възстановил, той отново става DR, но текущият BDR си остава такъв.

Designated router

Целта на избора на DR е да се намали служебния трафик в мрежата.

Осигурява се един единствен източник на обновления на маршрутите.

DR поддържа пълна топологична таблица и изпраща обновления (updates) до другите рутери по multicast.

Всички маршрутизатори в областта са в slave/master отношения с DR. Те формират съседства само с DR и BDR.

Designated Router (Con't)

Slave рутер изпраща update към DR и BDR на multicast адрес 224.0.0.6.

DR след това разпраща update до другите рутери в областта на multicast адрес 224.0.0.5.

Така няма да се налага рутерите постоянно да си обменят “update-и” всеки-с-всеки, а ще ги получават от един единствен източник.

Използването на multicasting допълнително намалява натоварването в мрежата.

Външни маршрути. load-sharing.

OSPF осигурява по-добро балансиране на натоварването (load-sharing) на външните връзки от другите IGPs.

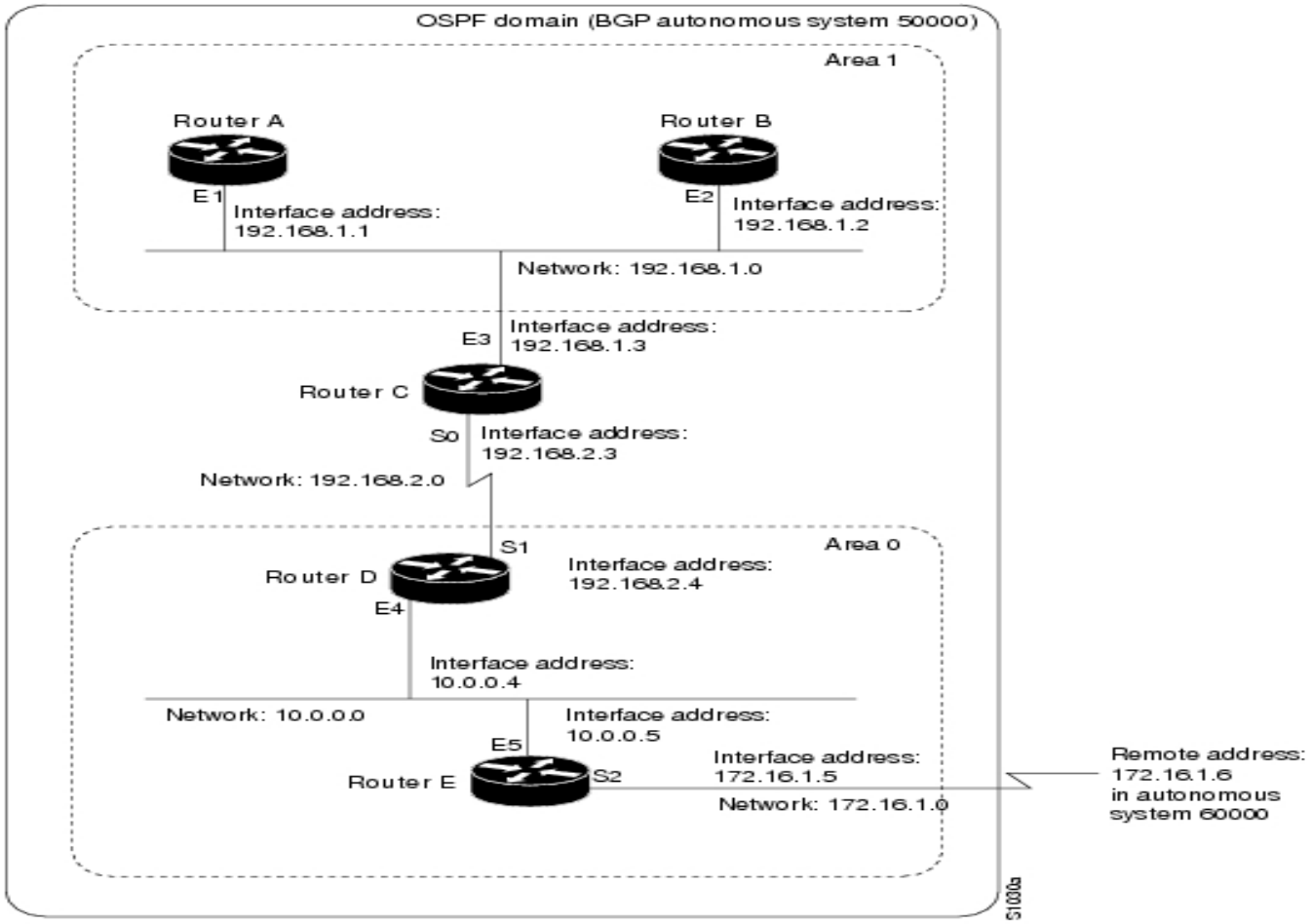
default route до ISP се инжектира в OSPF от множество ASBRs като Type I външен маршрут.

Външната стойност на маршрута е една и съща, но другите маршрутизатори ще изпратят външния си трафик до ASBR, до който имат път с най-малка стойност.

Чрез нагласяване на стойностите натоварването на ASBRs може да се разпредели равномерно.

Ако обаче default route от различни ISP-та се инжектира с различни външни стойности като Type II външен маршрут, маршрутът по подразбиране с по-ниска стойност става основен, а другите ще са резервни.

Конфигуриране на OSPF



Конфигуриране на OSPF

Имате 5 рутера, конфигурирани с OSPF:

- Router A и Router B са вътрешни за Area (област) 1.
- Router C е OSPF area border router (ABR). Area 1 - интерфейс E3, Area 0 - S0.
- Router D е вътрешен за Area 0 (backbone area).
- Router E е OSPF autonomous system boundary router (ASBR). BGP маршрути се вмъкват (редистрибутират - redistribute) в OSPF и съответно се анонсират от OSPF.

(!!!това е само пример на Cisco, но не трябва да се допуска, ще задръсти рутера)

Router A—Internal Router

```
interface ethernet 1
 ip address 192.168.1.1 255.255.255.0

router ospf 1
 network 192.168.1.0 0.0.0.255 area 1
```

Wildcard mask

Забележете!

0.0.0.255 вместо 255.255.255.0

В Cisco IOS реализацията на OSPF се прилага wildcard mask (огледална на subnet mask). Т.е вместо .AND. >> .OR.

В quagga си е с префикс: /m,

router ospf

network 10.0.0.0/8 area 0

Router B—Internal Router

```
interface ethernet 2
 ip address 192.168.1.2 255.255.255.0

router ospf 202
 network 192.168.1.0 0.0.0.255 area 1
```

Router C—ABR

```
interface ethernet 3
 ip address 192.168.1.3 255.255.255.0

interface serial 0
 ip address 192.168.2.3 255.255.255.0

router ospf 999
 network 192.168.1.0 0.0.0.255 area 1
 network 192.168.2.0 0.0.0.255 area 0
```

Router D—Internal Router

```
interface ethernet 4
 ip address 10.0.0.4 255.0.0.0

interface serial 1
 ip address 192.168.2.4 255.255.255.0

router ospf 50
 network 192.168.2.0 0.0.0.255 area 0
 network 10.0.0.0 0.255.255.255 area 0
```

Router E—ASBR

```
interface ethernet 5
 ip address 10.0.0.5 255.0.0.0

interface serial 2
 ip address 172.16.1.5 255.255.255.252

router ospf 109
 network 10.0.0.0 0.255.255.255 area 0
 redistribute bgp 65001 metric 1 metric-type 1

router bgp 65001
 network 192.168.0.0
 network 10.0.0.0
 neighbor 172.16.1.6 remote-as 65000
```

Router E – default route

...
ip route 0.0.0.0 0.0.0.0 172.16.1.6

router ospf 109
 network 10.0.0.0 0.255.255.255 area 0
 default-information originate

OSPFv3

OSPFv3 (RFC 2740) - OSPF за IPv6:

Основните механизми (подобно на RIPng и RIPv2) - flooding, избор на DR, области (area), SPF са запазени.

(Мултикаст адреси: FF02::5 All и FF02::6 DR)

Но, подобно на RIPng и RIPv2:

OSPFv3 не е обратно съвместим с OSPFv2.

OSPF рутер да маршрутизира и IPv4, и IPv6, трябва да поддържа и OSPFv2, и OSPFv3 процес.

Промени в OSPFv3

Заради особеностите на IPv6.

IPv6 – "link" (връзка): "комуникационно средство или среда, по която възлите комуникират на каналния слой (link layer)".

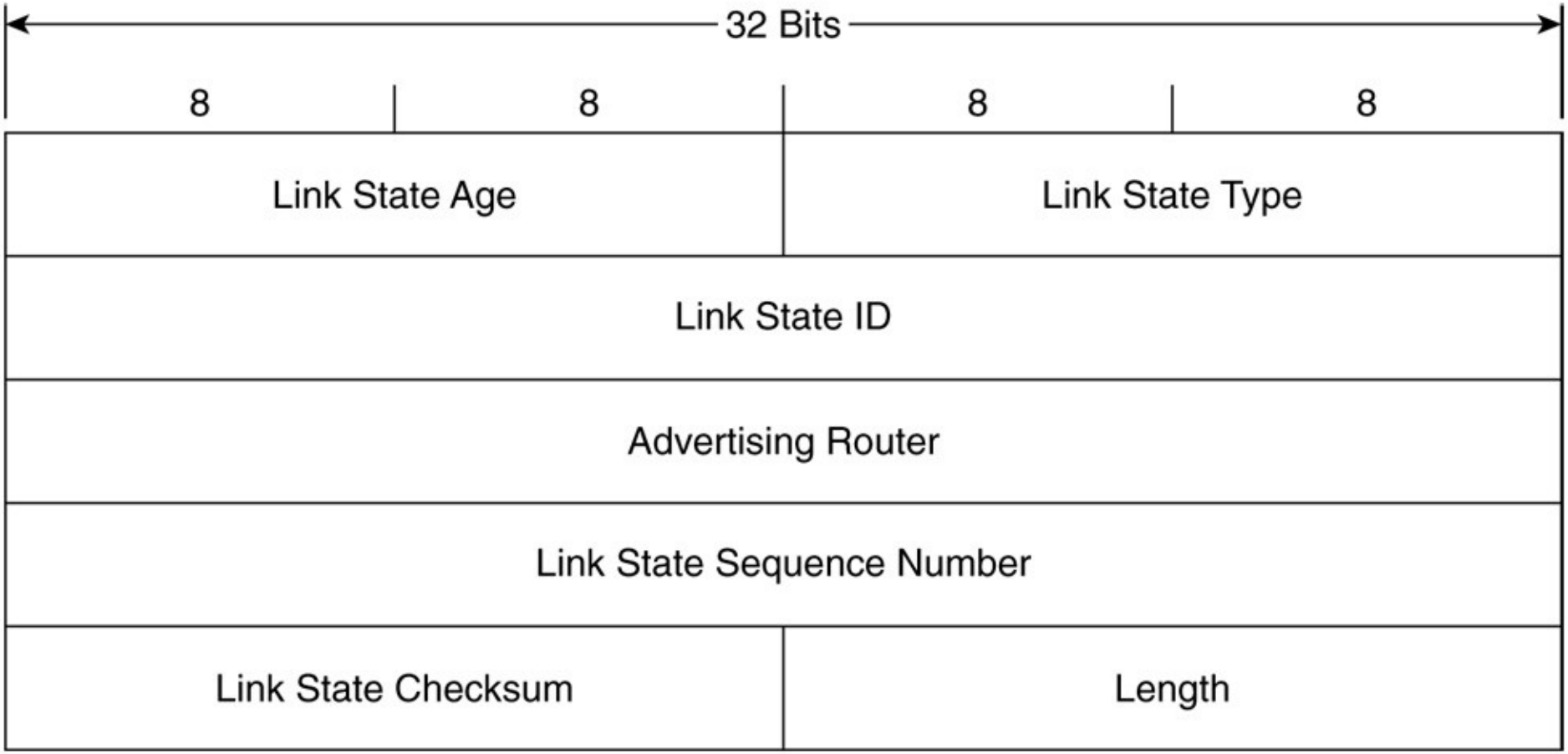
≥ 1 IP подмрежи (IPv6 prefix) на **1 link**, като два възела могат да “говорят” директно по връзката, даже и да не са на една и съща IP подмрежа.

Промени в OSPFv3. Адресна семантика.

Премахване на адресната семантика в OSPF пакетите и LSAs с изключение на LSA payloads в Link State Update Packets.

OSPF Router IDs, Area IDs и LSA Link State IDs остават с IPv4 формат, 32-bits.

Промени в OSPFv3. LSA.



Промени в OSPFv3. Flooding.

Три отделни обсега на разпространение на LSAs (flooding scopes):

- Link-local scope. Link-LSA - използва link-local IPv6 address (FF80::/10).

- Area scope. LSA се разпространяват само и единствено в дадена OSPF област.

- AS scope. LSA се flood-ват из целия маршрутен домейн.

Промени в OSPFv3. Аутентикация.

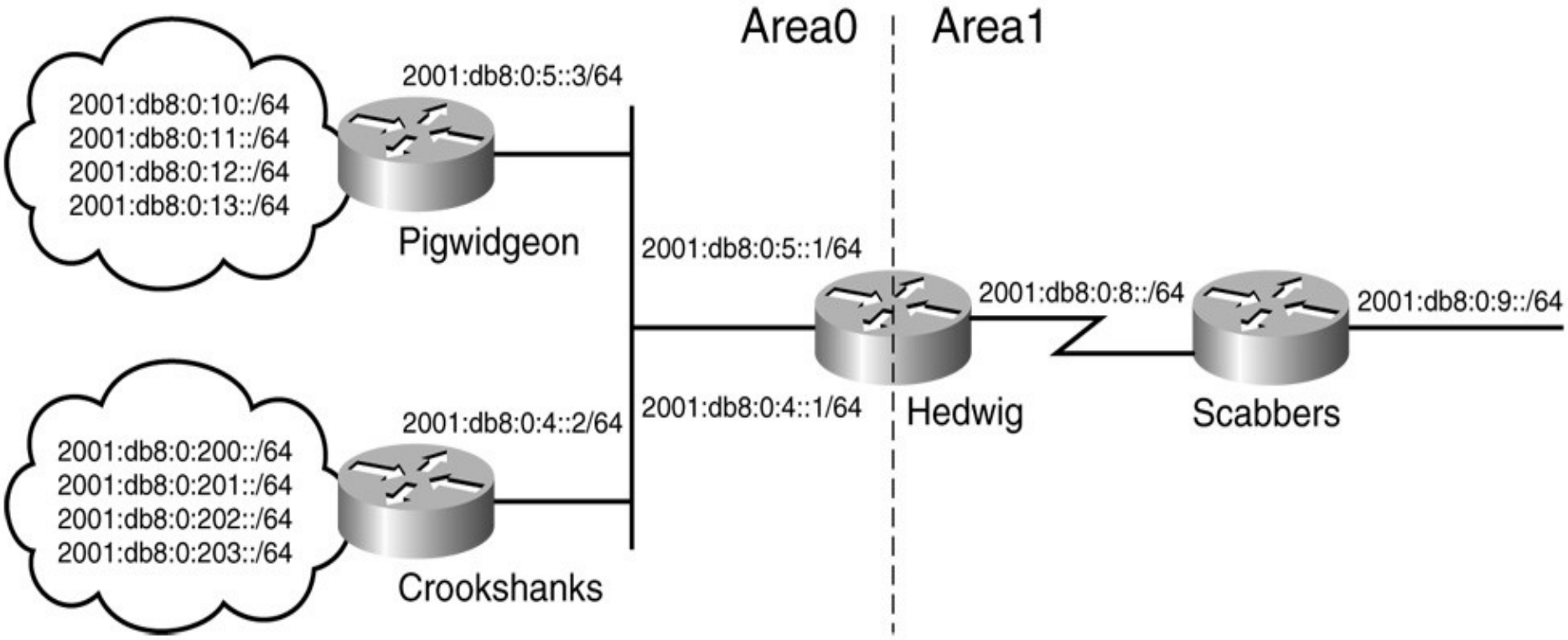
Поддържа **≥ 1** процеса (instances) / връзка.

Напр., Network Access Point (NAP), множество OSPF рутери, свързани към един Ethernet канал, но от различни домейни.

Не формират “adjacency” помежду си.

Аутентикацията между рутерите е премахната. Разчита се на IPv6 Authentication Header и Encapsulating Security Payload.

Конфигуриране на OSPFv3



Конфигуриране на OSPFv3

Hedwig

interface Serial 0/0

ipv6 address 2001:db8:0:8::1/64

ipv6 ospf 1 area 1

Interface fe 0/0

ipv6 address 2001:db8:0:4::1/64

ipv6 address 2001:db8:0:5::1/64

ipv6 ospf 1 area 0

Pigwidgeon

interface fe 0/0

ipv6 address 2001:db8:0:5::3/64

ipv6 ospf 1 area 0

interface Serial 0/0

ipv6 address 2001:db8:0:10::1/64

ipv6 ospf 1 area 0

Конфигуриране на OSPFv3

Hedwig (cont'd)

ipv6 router ospf 1

router-id 192.168.1.1

Pigwidgeon (cont'd)

ipv6 router ospf 1

router-id 192.168.5.1

Show команди

Hedwig#show ipv6 protocol

```
IPv6 Routing Protocol is "connected"
IPv6 Routing Protocol is "static"
IPv6 Routing Protocol is "ospf 1"
 Interfaces (Area 0):
  Ethernet0/0
 Interfaces (Area 1):
  Serial0/0
 Redistribution:
  None
```

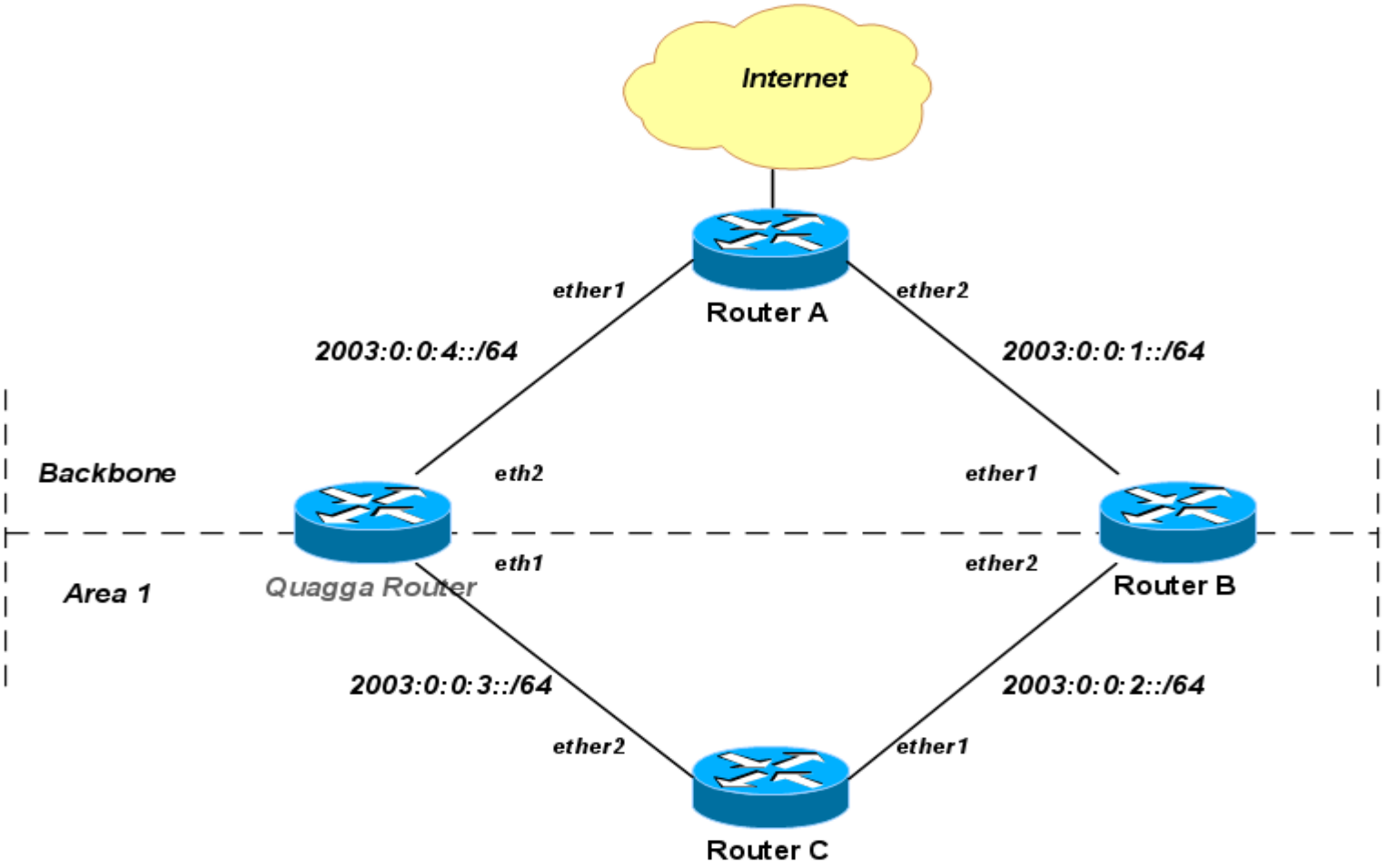
Show команди

Crookshanks#show ipv6 ospf neighbor

(*Hedwig RID: 10.1.1.1; Pigwidgeon RID: 10.1.3.1*)

```
Neighbor ID     Pri   State           Dead
Time   Interface ID  Interface
10.1.1.1          1   FULL/DROTHER
00:00:30   3          Ethernet0/0
10.1.3.1          1   FULL/BDR      00:00:37
 3          Ethernet0/0
```

OSPFv3 c Quagga (**ospf6d**)



Quagga Router (**ospf6d**)

```
debian:~# ip -6 addr add 2003:0:0:3::4/64 dev eth1
debian:~# ip -6 addr add 2003:0:0:4::4/64 dev eth2
```

debian:~# cat /etc/quagga/ospf6d.conf

...

interface eth1
  ipv6 ospf6 cost 10

interface eth2
  ipv6 ospf6 cost 10

router ospf6
  router-id 0.0.0.4
  interface eth1 area 0.0.0.1
  interface eth2 area 0.0.0.0

Optimized Link State Routing Protocol

Optimized Link State Routing (**OLSR** - rfc3626) - за мобилни **ad hoc** мрежи.

Multipoint relays (MPRs). MPRs са избрани възли, които единствено и само те препращат broadcast съобщения по време на flooding процеса.

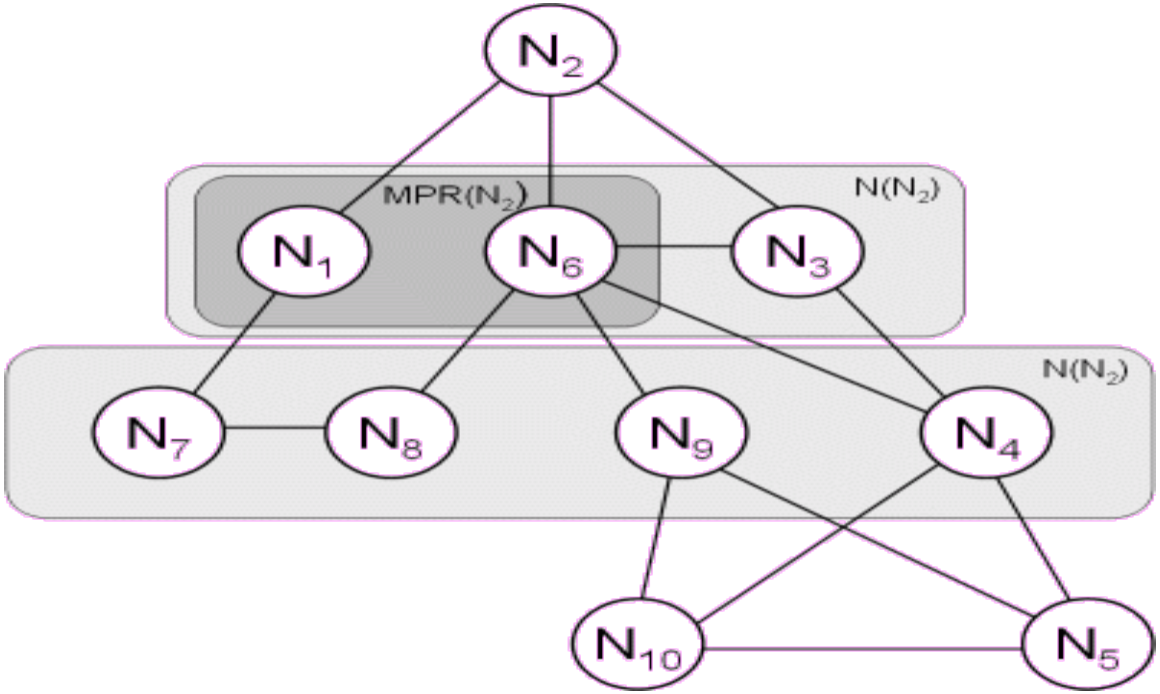
Така значително се редуцира служебния трафик.

MPR анонсира периодически информация, че има достижимост до възлите, които са го избрали като такъв.

Т.е разпространява се частична LS информация.

OLSR създава оптимални маршрути (метрика брой хопове).

OLSR. Избор на MPR.

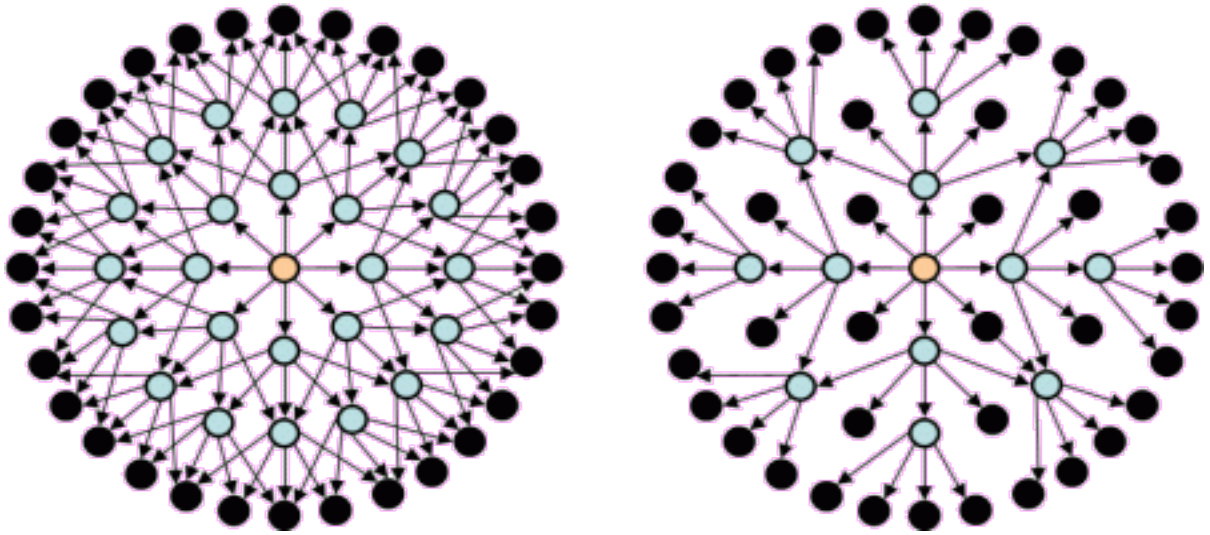


Даден възел (N2) избира своя MPR:

(1) повече 2-hop възли да бъдат обхванати;

(2) да се минимизира броя на MPRs.

OLSR. Сравнение.



Нормален процес на flooding (ляво).
MPR flooding (дясно).