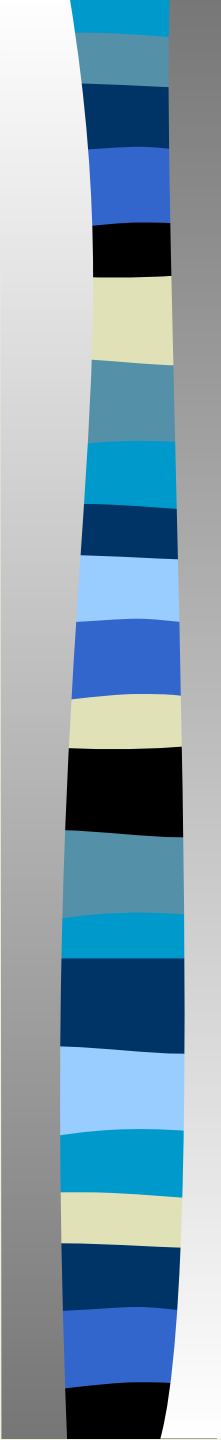


Протоколы маршрутизации

- Внутренние и внешние протоколы маршрутизации
- Протоколы RIP, OSPF, EGP, BGP
- Фильтрация трафика и объявлений маршрутов

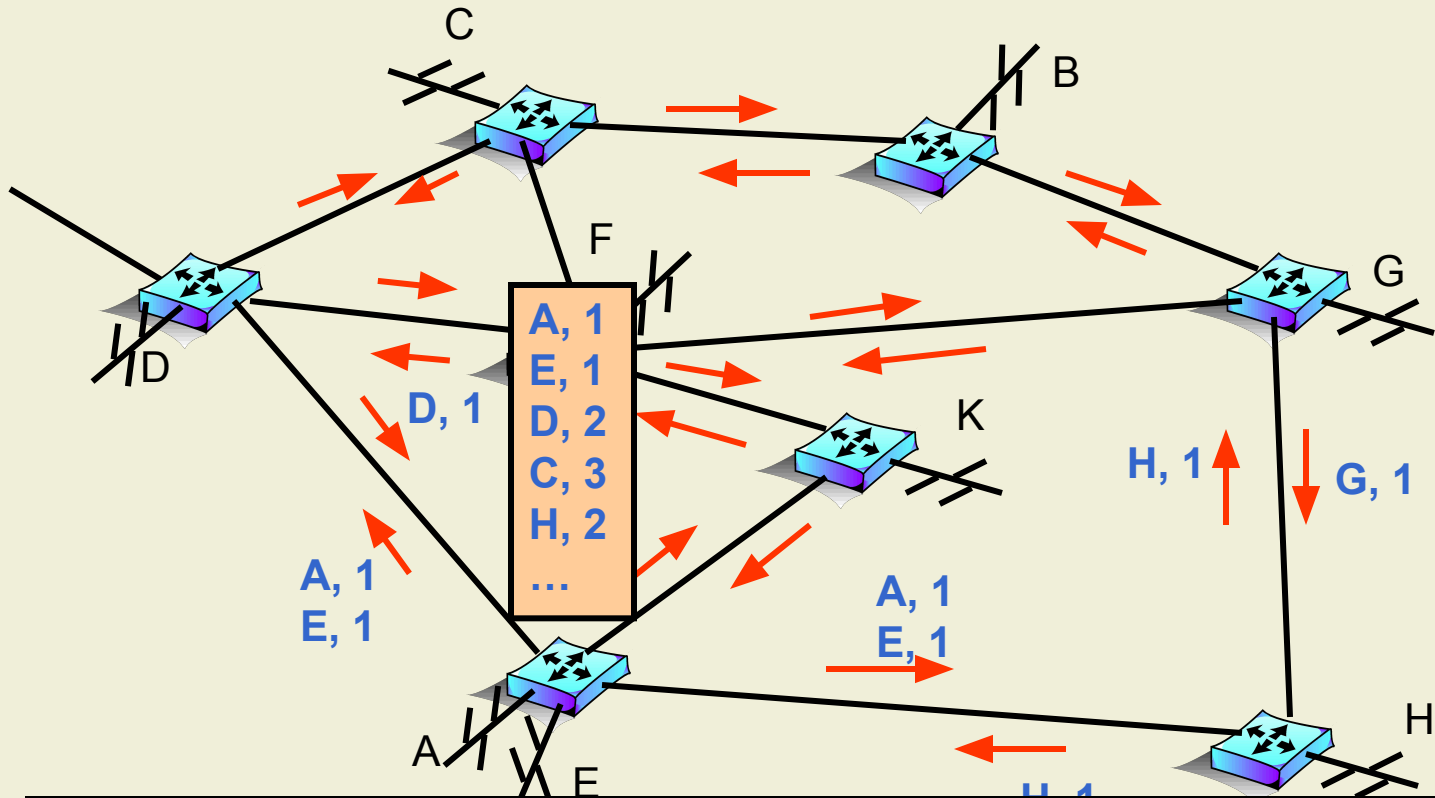


Протоколы обмена маршрутной информацией стека TCP/IP

- ♦ дистанционно-векторные алгоритмы
(*Distance Vector Algorithms, DVA*)
- ♦ алгоритмы состояния связей
(*Link State Algorithms, LSA*)

Дистанционно-векторные алгоритмы

Distance-Vector Algorithms, DVA



- 2.
3. Процесс обмена векторами расстояний периодически повторяется – отслеживается динамика изменения топологии сети

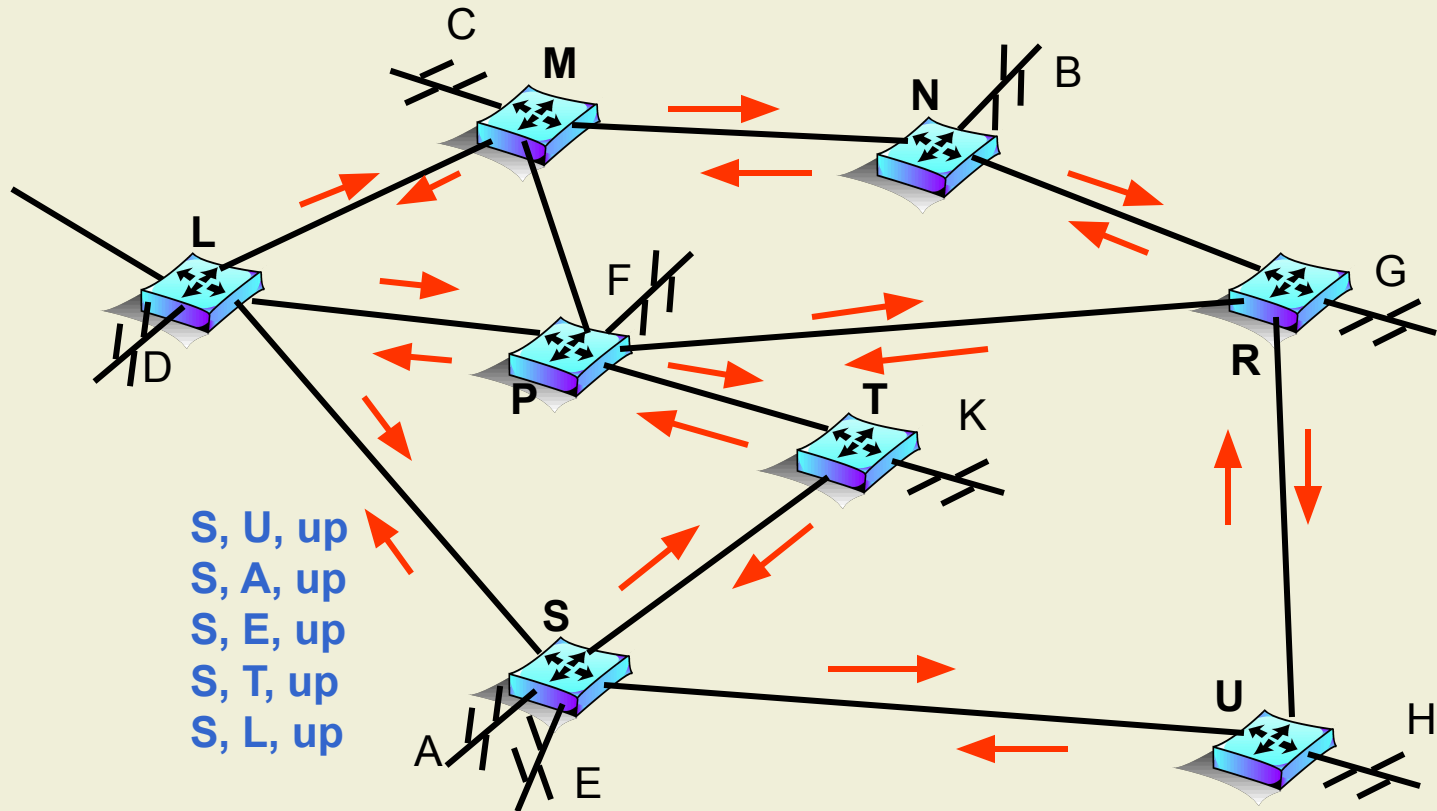


Алгоритм дистанционно-векторного типа

- ◆ Хорошо работают только в небольших сетях
- ◆ Засоряют линии связи интенсивным широковещательным трафиком
- ◆ Изменения конфигурации могут обрабатываться по этому алгоритму не всегда корректно – информация из «вторых» рук может уже устареть и «обмануть»
- ◆ Наиболее распространенным протоколом, основанным на дистанционно-векторном алгоритме, является протокол RIP

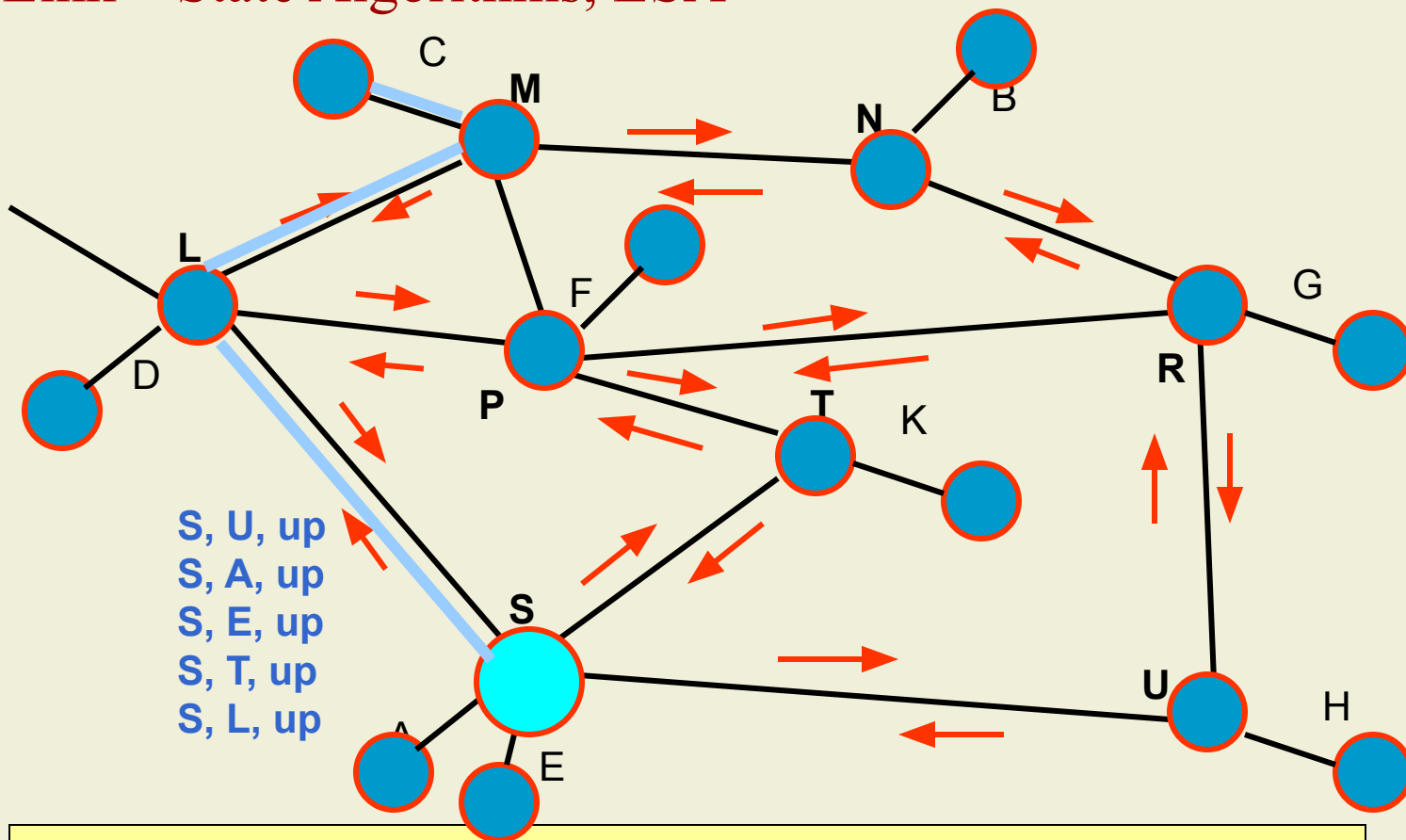
Алгоритмы состояния связей

Link – State Algorithms, LSA



1. Маршрутизаторы обмениваются информацией о топологии сети для построения точного графа. Вершины – маршрутизаторы и сети (A, B, C, ..., S, T, U). Передается информация о связях между вершинами графа.

Алгоритмы состояния связей Link – State Algorithms, LSA



- 2.
3. Каждый маршрутизатор на основе графа строит свою таблицу маршрутизации:
 - находит от себя кратчайший маршрут до каждой сети
 - запоминает из этого маршрута только следующий хоп



Алгоритм состояния связей

- ◆ Периоды нестабильной работы сети при изменении топологии существенно короче, чем у алгоритмов DVA
- ◆ Вычисления кратчайшего маршрута существенно сложнее – нагрузка на маршрутизатор растет
- ◆ Служебный трафик – меньше
- ◆ Open Shortest Path First – протокол типа LSA в стеке IP

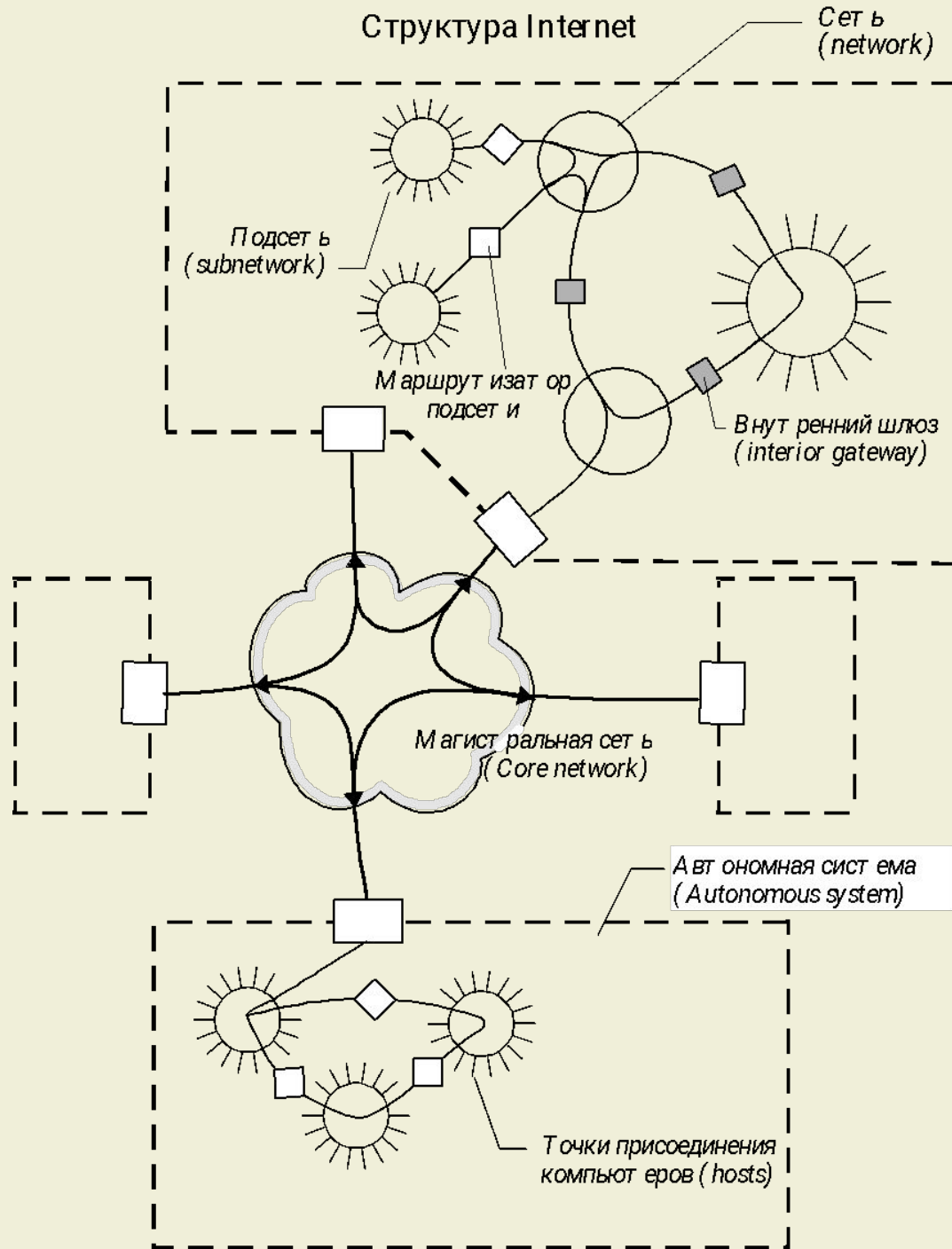


Внутренние и внешние протоколы маршрутизации

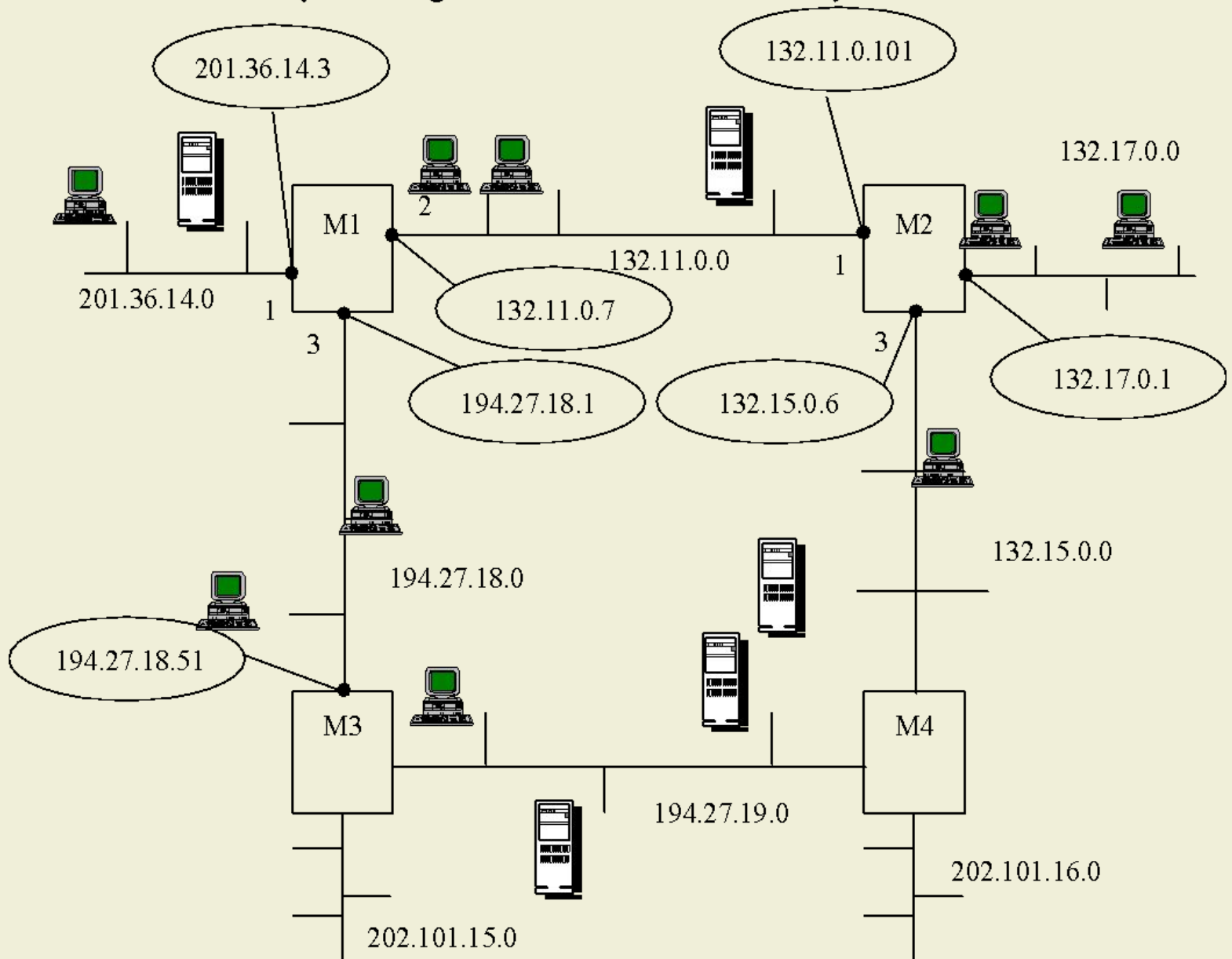
Структура и терминология сет и Internet:

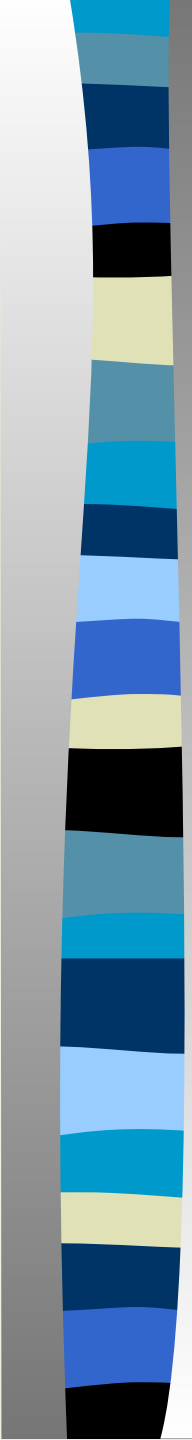
- ◆ магистральная сеть (core backbone network)
- ◆ автономные системы (autonomous systems)
- ◆ протоколы внутренних шлюзов (interior gateway protocol, **IGP**)
- ◆ протоколы внешних шлюзов (exterior gateway protocol, **EGP**)
- ◆ таблица достижимости (reachability table)

Структура Internet



Протокол маршрутизации RIP (Routing Information Protocol)





Начальное состояние таблицы маршрутизации в M1:

| Номер сети | Адрес следующего маршрутизатора | Порт | Расстояние |
|-------------|---------------------------------|------|------------|
| 201.36.14.0 | 201.36.14.3 | 1 | 1 |
| 132.11.0.0 | 132.11.0.7 | 2 | 1 |
| 194.27.18.0 | 194.27.18.1 | 3 | 1 |

Начальное состояние таблицы маршрутизации в M2:

| Номер сети | Адрес следующего маршрутизатора | Порт | Расстояние |
|------------|---------------------------------|------|------------|
| 132.11.0.0 | 132.11.0.101 | 1 | 1 |
| 132.17.0.0 | 132.17.0.1 | 1 | 1 |
| 132.15.0.0 | 132.15.0.6 | 1 | 1 |

RIP-объявления делаются каждым маршрутизатором каждые 30 секунд (Тобъявл.)

Протокол маршрутизации RIP (продолжение)

Таблица маршрутизации в M1 после одного шага:

| Номер сети | Адрес следующего маршрутизатора | Порт | Расстояние |
|--------------|---------------------------------|------|------------|
| 201.36.14.0 | 201.36.14.3 | 1 | 1 |
| 132.11.0.0 | 132.11.0.7 | 2 | 1 |
| 194.27.18.0 | 194.27.18.1 | 3 | 1 |
| 132.17.0.0 | 132.11.0.101 | 2 | 2 |
| 132.15.0.0 | 132.11.0.101 | 2 | 2 |
| 194.27.19.0 | 194.27.18.51 | 3 | 2 |
| 202.101.15.0 | 194.27.18.51 | 3 | 2 |

Правила построения таблицы маршрутизации по RIP

- В таблице для каждой сети остается только одна запись с минимальным расстоянием
- Каждая запись имеет срок жизни
 $6 \times \text{Тобъявл.} = 180 \text{ сек} = 3 \text{ мин}$
- Максимальное количество хопов в поле расстояния - 15
- 16-
признак недостижимой сети
- Объявления о имеющейся в таблице сети, но с худшей метрикой учитываются только от того маршрутизатора, на основании объявления которого была сделана запись

Протокол маршрутизации RIP (продолжение 2)

Таблица маршрутизации в M1
агов: после двух
ш

| Номер сети | Адрес следующего маршрутизатора | Порт | Расстояние |
|--------------|---------------------------------|------|------------|
| 201.36.14.0 | 201.36.14.3 | 1 | 1 |
| 132.11.0.0 | 132.11.0.7 | 2 | 1 |
| 194.27.18.0 | 194.27.18.1 | 3 | 1 |
| 132.17.0.0 | 132.11.0.101 | 2 | 2 |
| 132.15.0.0 | 132.11.0.101 | 2 | 2 |
| 132.15.0.0 | 194.27.18.51 | 3 | 3 |
| 194.27.19.0 | 194.27.18.51 | 3 | 2 |
| 194.27.19.0 | 132.11.0.101 | 2 | 3 |
| 202.101.15.0 | 194.27.18.51 | 3 | 2 |
| 202.101.16.0 | 132.11.0.101 | 2 | 3 |
| 202.101.16.0 | 194.27.18.51 | 3 | 3 |

Формат сообщения RIP

- Протокол RIP передает данные о номерах сетей в пакетах UDP (порт 520).
- В одном пакете RIP может передаваться до 25 номеров сетей.

Пакет RIP имеет следующий формат:

| | | | |
|--------------------------------|--------|-------|----|
| 0 | 8 | 16 | 31 |
| Команда | Версия | 0...0 | |
| Тип адресной информации (IP=2) | | 0...0 | |
| IP адрес | | | |
| 0...0 | | | |
| Метрика | | | |

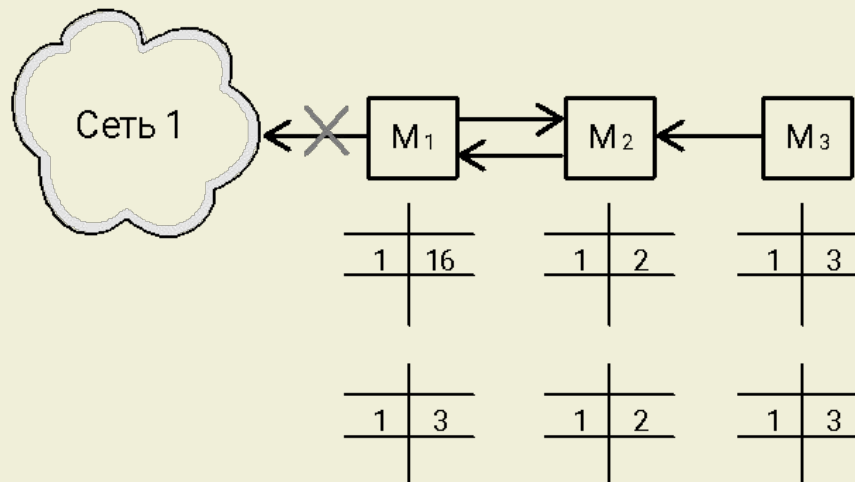
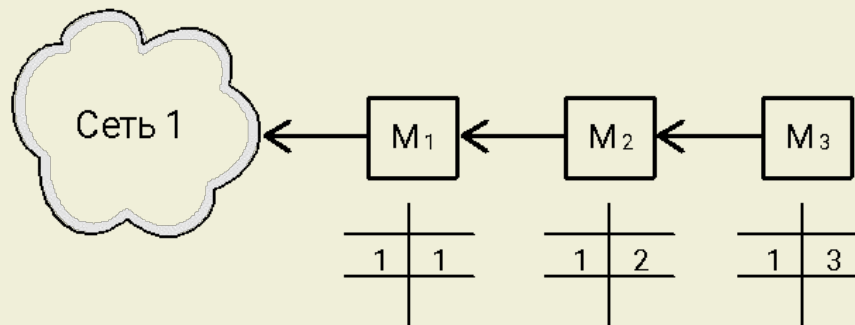
Протокол RIP использует две команды:

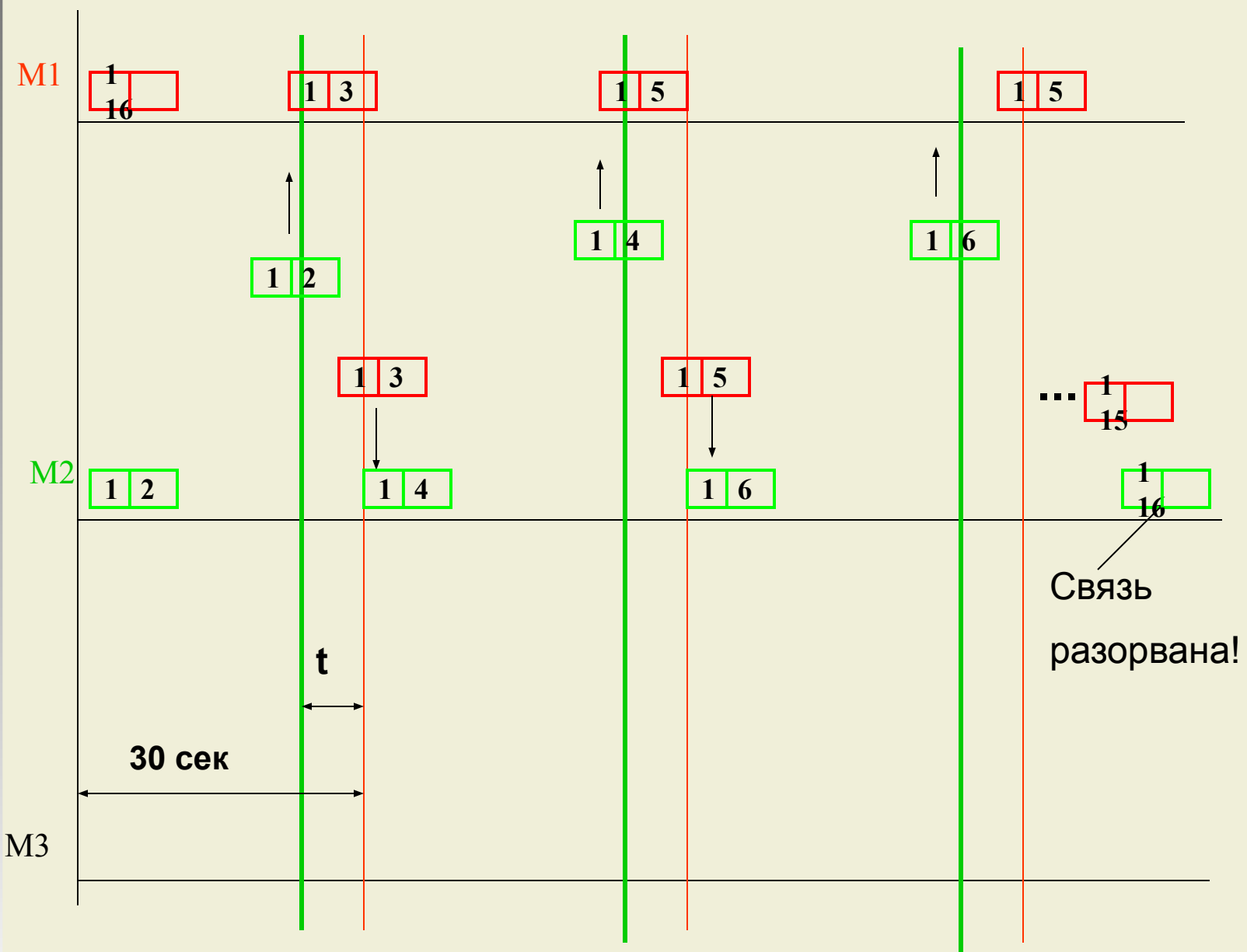
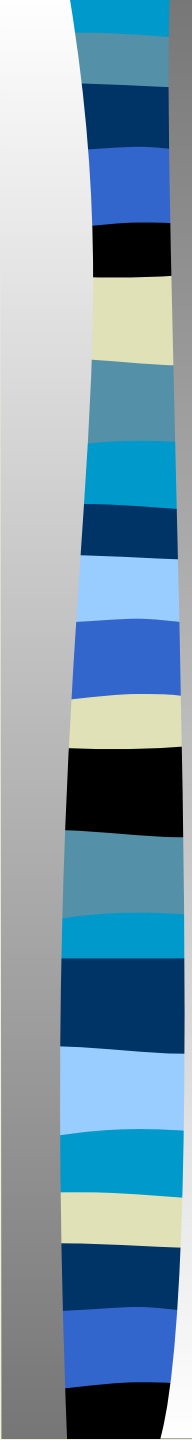
- 1 - запрос о посылке всей или части таблицы маршрутизации И-
- 2 - сообщение, содержащее всю или часть таблицы маршрутизации р-

Нестабильность работы RIP-маршрутизаторов при отказе связей

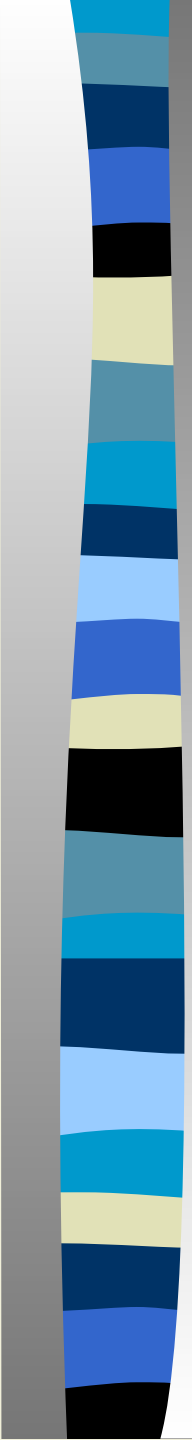
Нестабильность работы - заикливание пакетов, а также отправка пакетов по нерациональным маршрутам

Заикливание пакетов в маршрутных петлях:





Время существования петли = $(16/2-1) \times 30 = 210 \text{ с} = 3,5 \text{ мин}$

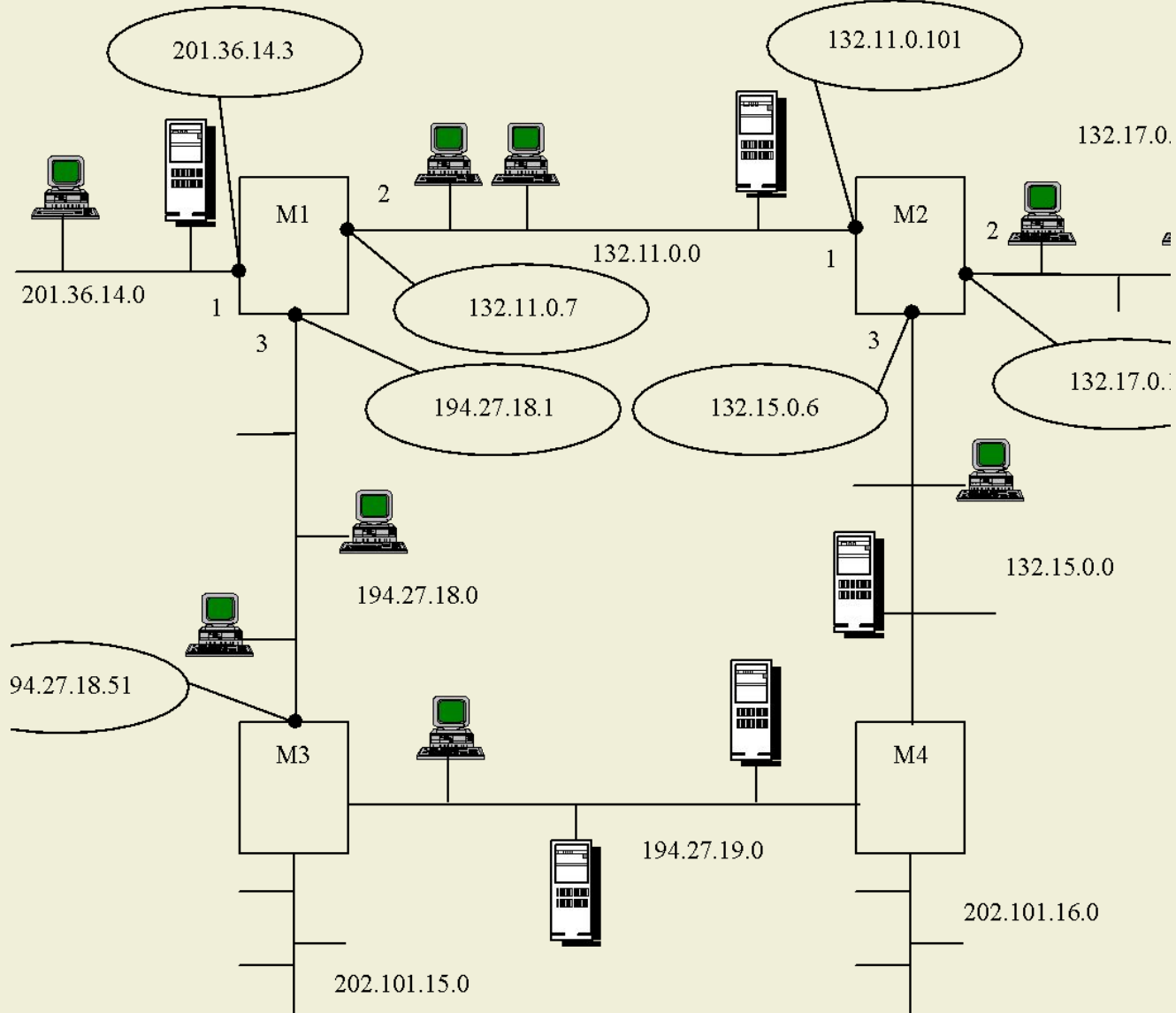


Методы борьбы с ложными маршрутами в протоколе RIP

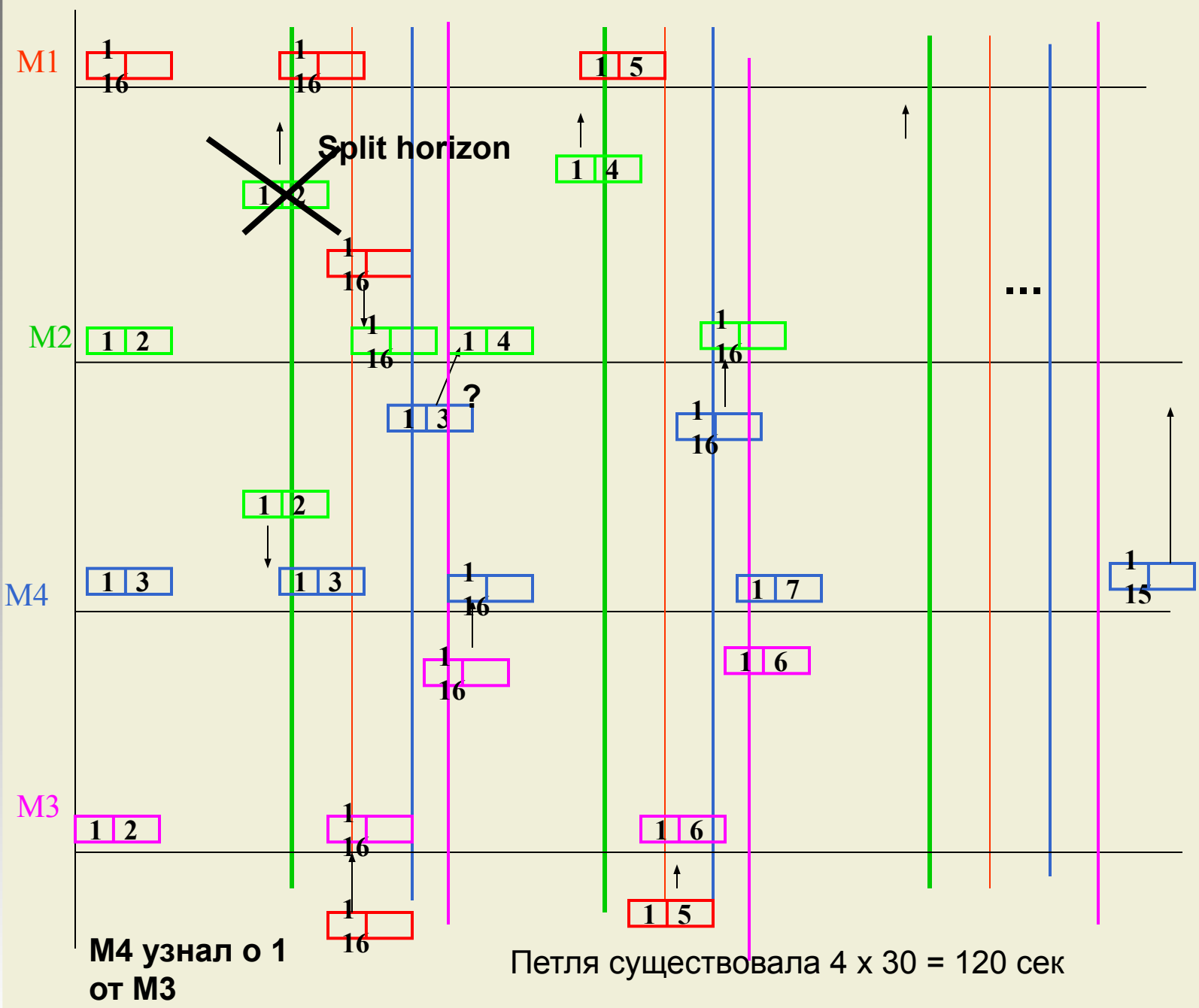
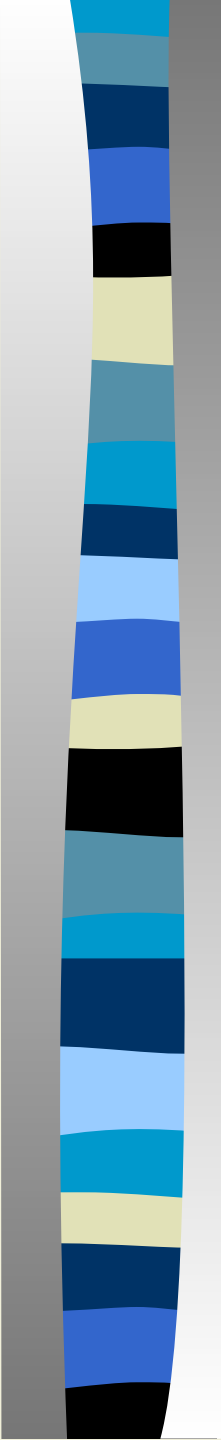
1) Split horizon - " расщепление горизонта"

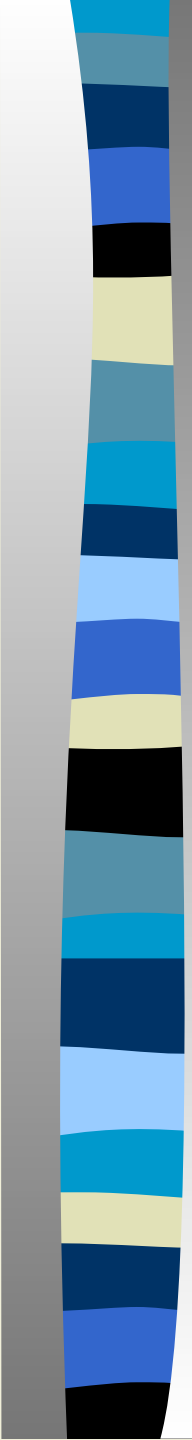
**Маршрутизатор никогда не возвращает данные о
некоторой сети тому маршрутизатору, от которого узнал о
этой сети**

- ◆ Защищает от зацикливания пакетов в петлях, образованных соседними маршрутизаторами
- ◆ Не защищает от зацикливания пакетов в маршрутных петлях, образованных 3-мя и более маршрутизаторами



При отказе связи M1 с сетью 201.36.14.0 информация о достижимости этой сети вернется в M1
 через цепочку M2 - M4 – M3.





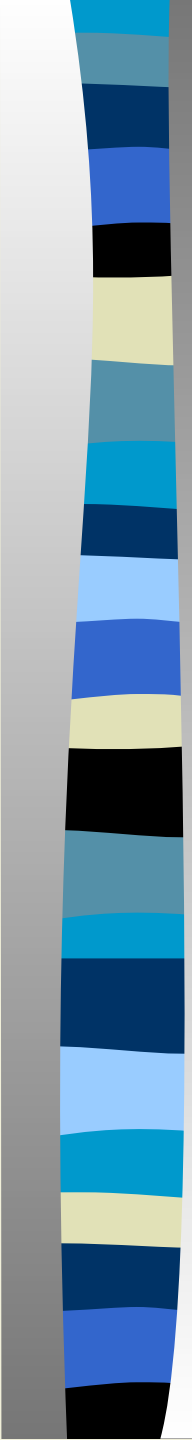
Методы борьбы с ложными маршрутами в протоколе RIP

2) Triggered updates - триггерные обновления

При изменении состояния связи маршрутизатор немедленно делает объявление, не ожидая периода в 30 сек

Небольшая задержка (1-5 сек) – возможность решения проблемы на нижних уровнях

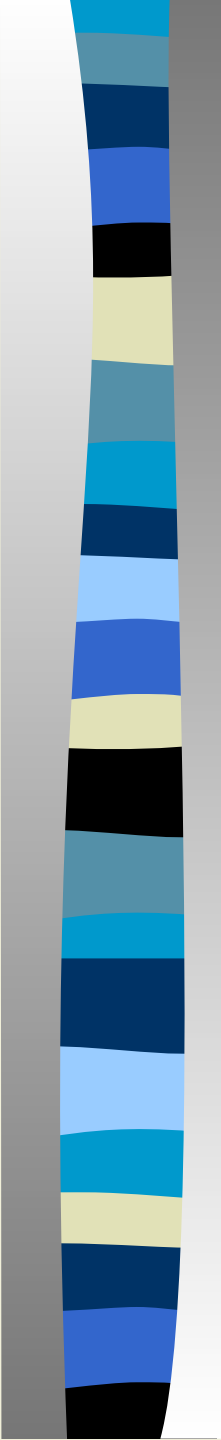
- ♦ В многих случаях предотвращает использование устаревшей информации из вторых рук - быстро заменяет ее новой во всех маршрутизаторах сети
- ♦ Возможны сбои - когда регулярное объявление с устаревшей информацией опережает триггерное объявление



3) Hold down - " замораживанием изменений"

Вводится тайм-аут на принятие новых данных о сети, которая только что стала недоступной.

- ◆ Тайм-аут предотвращает принятие устаревших сведений о некотором маршруте от тех маршрутизаторов, которые находятся на некотором расстоянии от отказавшей связи и передают устаревшие сведения о ее работоспособности
- ◆ Хорошо сочетается с триггерными объявлениями



Протокол RIP v. 2

Новые свойства:

- ◆ поддержка масок подсетей
- ◆ аутентификация маршрутизаторов

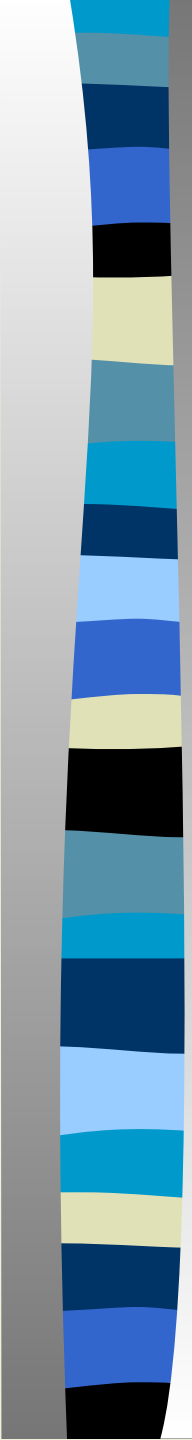
Формат сообщений RIP v.2:

| | | | |
|---------------------------------------|---------------|-----------------------|----|
| 0 | 8 | 16 | 31 |
| Команда | Версия | 0 ...0 | |
| Тип адресной информации (IP=2) | | Метка маршрута | |
| IP | | | |
| Маска подсети | | | |
| Следующий хоп | | | |
| Метрика | | | |

Если "Тип адресной информации" = 0xFFFF, то в первой записи вместо маршрутной информации передается 16-байтовый пароль для аутентификации

"

Метка маршрута" - для номера автономной системы

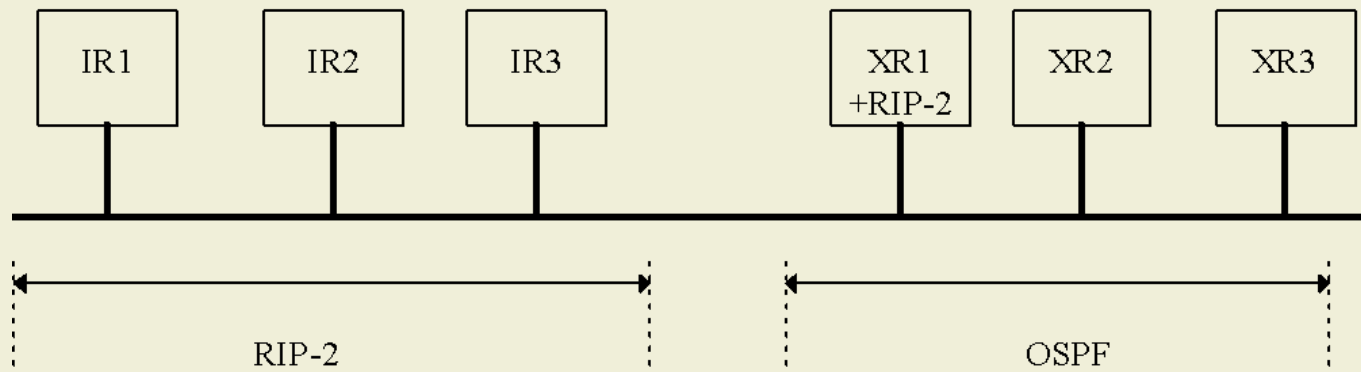


Правила взаимодействия маршрутизаторов RIP-2 с маршрутизаторами RIP-1

- информация, внутренняя для некоторой сети никогда не должна передаваться в другую сеть
- не должны передаваться данные о суперсетях, то есть сетях, имеющих маску более короткую, чем ее "естественная" длина

Поле "Следующий хоп" - связь RIP-2 с OSPF

- ◆ Поле "Следующий хоп" используется для указания адреса маршрутизатора, которому нужно передавать пакеты для сети, адрес которой передается в поле IP адрес.
- ◆ Позволяет задать рациональный маршрут через сеть с другим протоколом маршрутизации, например, OSPF



Протокол IGRP (Interior Gateway Routing Protocol) компании Cisco Systems

- Протокол IGRP - дистанционно-векторный протокол
- Использует приемы split horizon, triggered updates и hold down
- Использует дополнительные по сравнению с RIP механизмы:
 - Использование более тонкой метрики: $(K1/B_e + K2 \cdot D) / R$,
где B_e - пропускная способность канала,
задержка, вносимая каналом,
коэффициент готовности канала
 - Использование алгоритма Route Poisoning для борьбы с устаревшей информацией, получаемой из вторых рук через дельту
- Маршруты, метрика которых увеличивается при очередном обновлении более чем на 10%, отбрасываются

Использование этого алгоритма позволяет отказаться от "замораживания изменений", а это ускоряет установление в сети новой топологии после обрыва связи или отказа маршрутизатора

Протокол маршрутизации OSPF

Каждый маршрутизатор имеет исчерпывающую
информацию о топологии сети -
база данных топологической информации

Все маршрутизаторы имеют идентификаторы

Сети идентифицируются IP-адресами

Этапы работы протокола

- 1) построение топологической базы отдельного маршрутизатора:
- 2) построение общего графа путем обмена топологическими базами с соседями (подобно RIP)
- 3) построение таблиц маршрутизации
- 4) поддержание топологической базы



1

Этап конфигурирования топологической базы отдельного маршрутизатора

- ♦ Маршрутизатор - присоединенная соседняя сеть
(*ручное конфигурирование*)
- ♦ Маршрутизатор - соседний маршрутизатор, непосредственно присоединенный
(*ручное конфигурирование*)
- ♦ Маршрутизатор - соседний маршрутизатор, присоединенный через соседнюю сеть
(*для широкополосных сетей - автоматический опрос, для остальных - ручное конфигурирование*)

Для каждой связи дополнительно определяется 3 типа метрики



4 этап - поддержание топологической базы

- ♦ проверка состояния связей - обмен hello (10 сек)
- ♦ объявления о изменении состояния связи
- ♦ для надежности изредка обмен полными топологическими базами – раз в 30 мин

Дополнительные свойства:

- аутентификация
- разбиение на OSPF-области
- внешние объявления

Пример сети с маршрутизаторами OSPF

Связи типа

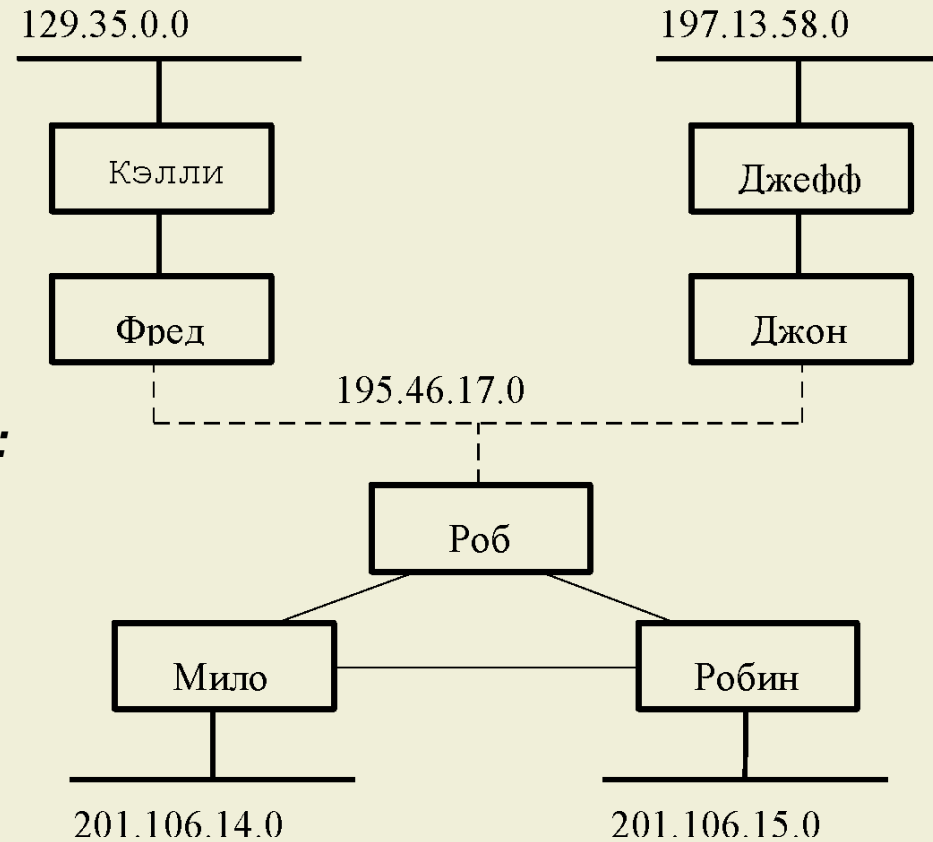
"маршрутизатор - маршрутизатор":

- Мило - Робин
- Мило - Роб
- Роб - Робин
- Кэлли - Фред
- Джеф - Джон

Связи типа

"маршрутизатор - сеть":

- Мило - 201.106.14.0
- Робин - 201.106.15.0
- Фред - 195.46.17.0
- Джон - 195.46.17.0
- Роб - 195.46.17.0
- Кэлли - 129.35.0.0
- Джеф - 197.13.58.0.





Маски переменной длины в OSPF

- ◆ Протокол OSPF передает длину маски вместе с номером сети
- ◆ Поддерживаются маски переменной длины
- ◆ IP-подсети могут перекрываться - одна сеть является подмножеством адресов другой
- ◆ Если адрес принадлежит нескольким подсетям в таблице маршрутизации, то маршрутизатор использует наиболее специфический маршрут

Метрика и оптимизация транзитных локальных сетей

В OSPF три категории сетей :

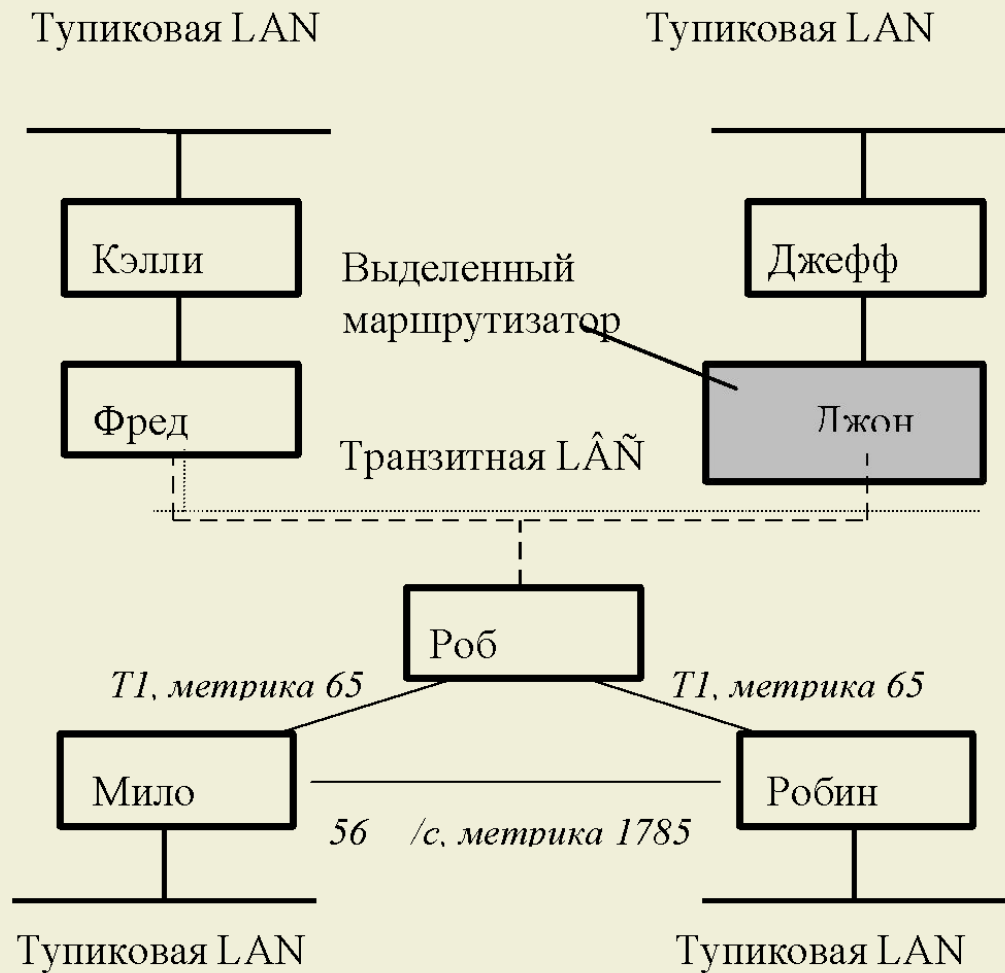
- 1) "хост-сеть» - подсеть из одного адреса
- 2) "тупиковая сеть» - подсеть, подключенная только к одному маршрутизатору
- 3) "транзитная сеть» - подсеть, подключенная к более чем одному маршрутизатору.

◆ Синхронизация маршрутизаторов "каждый с каждым" на транзитной сети:

передача базы $N*(N-1)$ раз

◆ Синхронизация с "выделенным" маршрутизатором:

передача базы $2*N$ раз





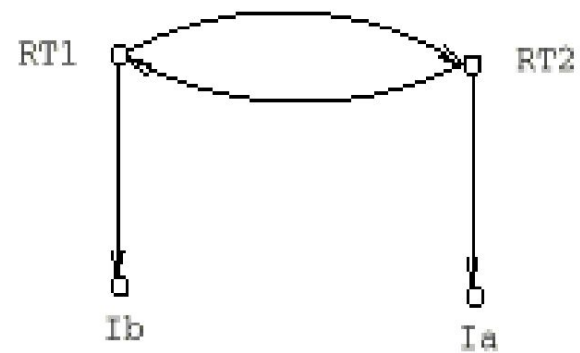
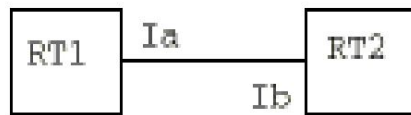
Топологическая база OSPF

Три типа вершин:

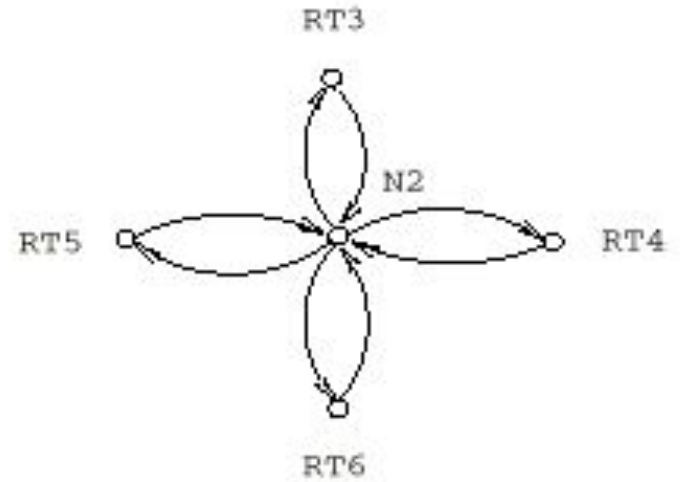
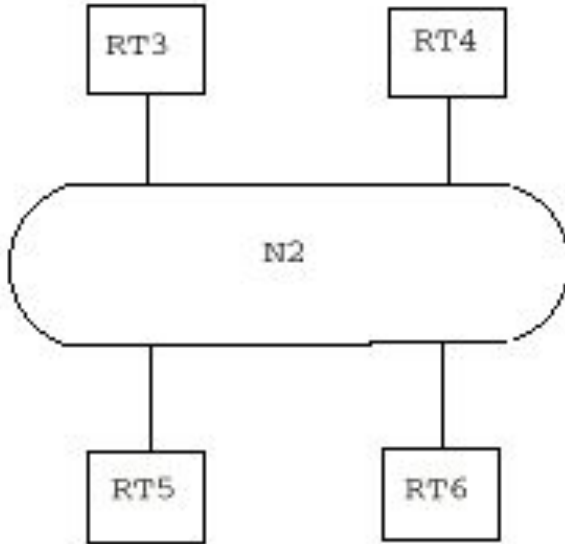
- Маршрутизатор (транзитная)
- Сеть (транзитная)
- Тупиковая (stub) сеть (не транзитная)

Преобразование физических связей:

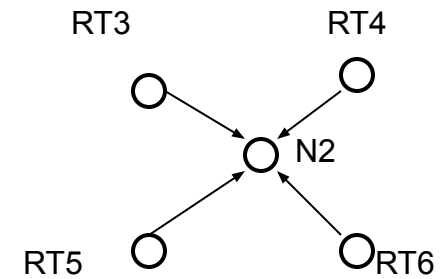
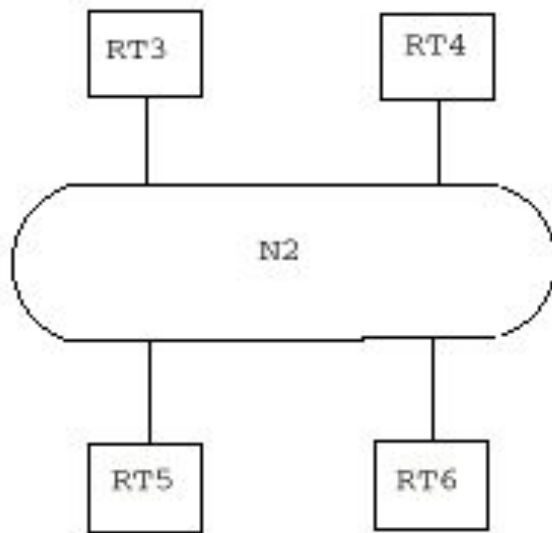
А) физическая связь "точка – точка"

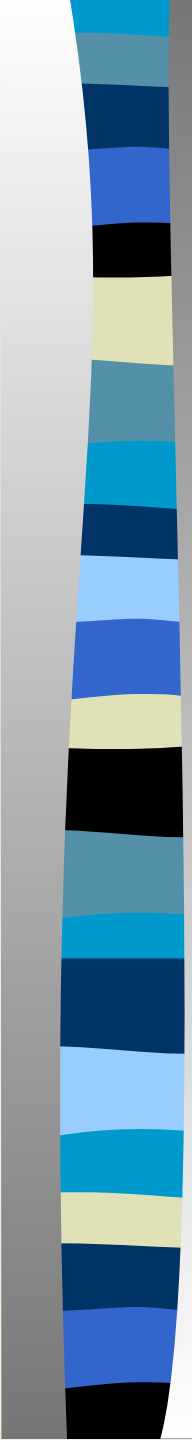


В) Связи через сеть с множественным доступом



С) Связь с тупиковой сетью с множественным доступом

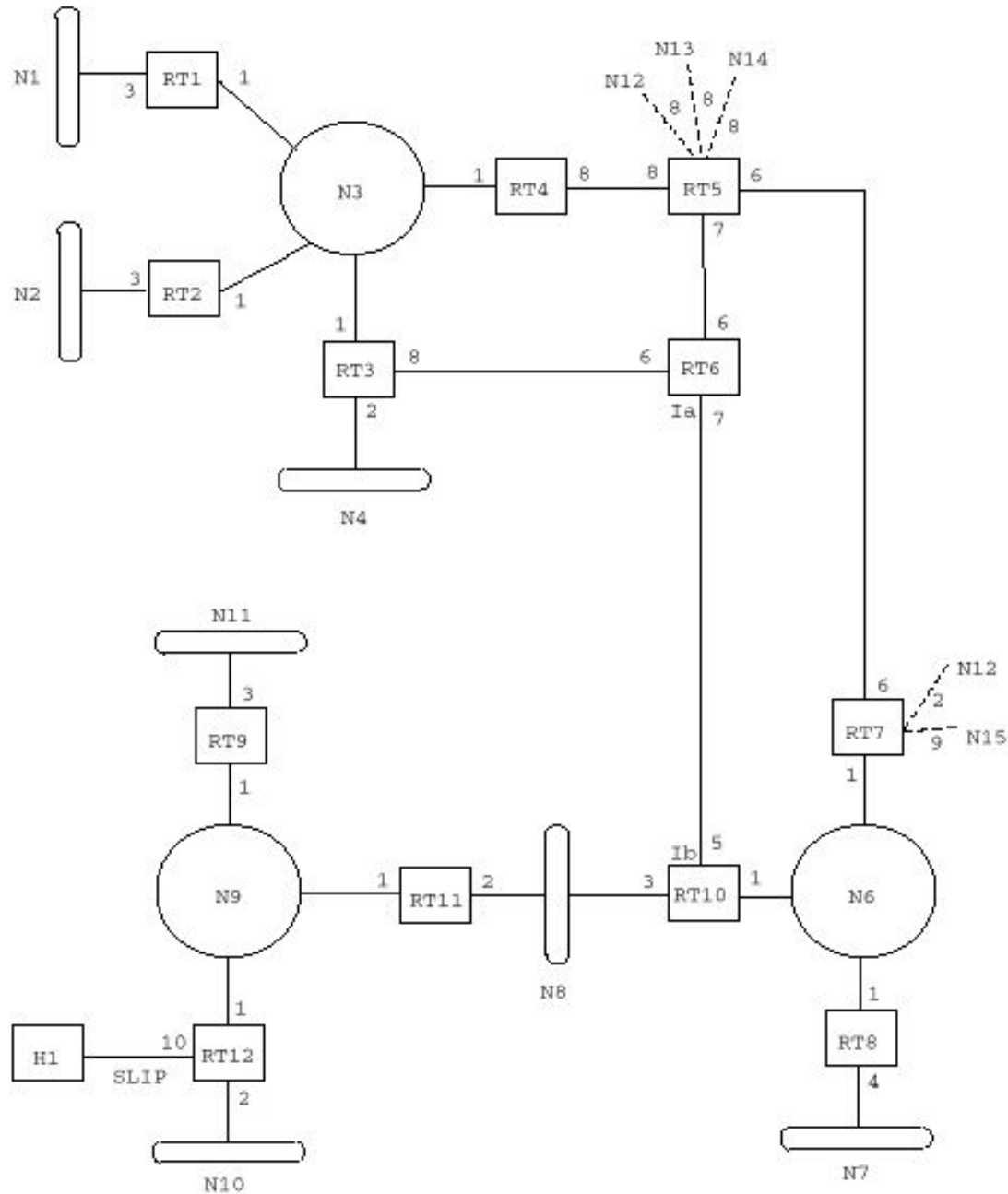




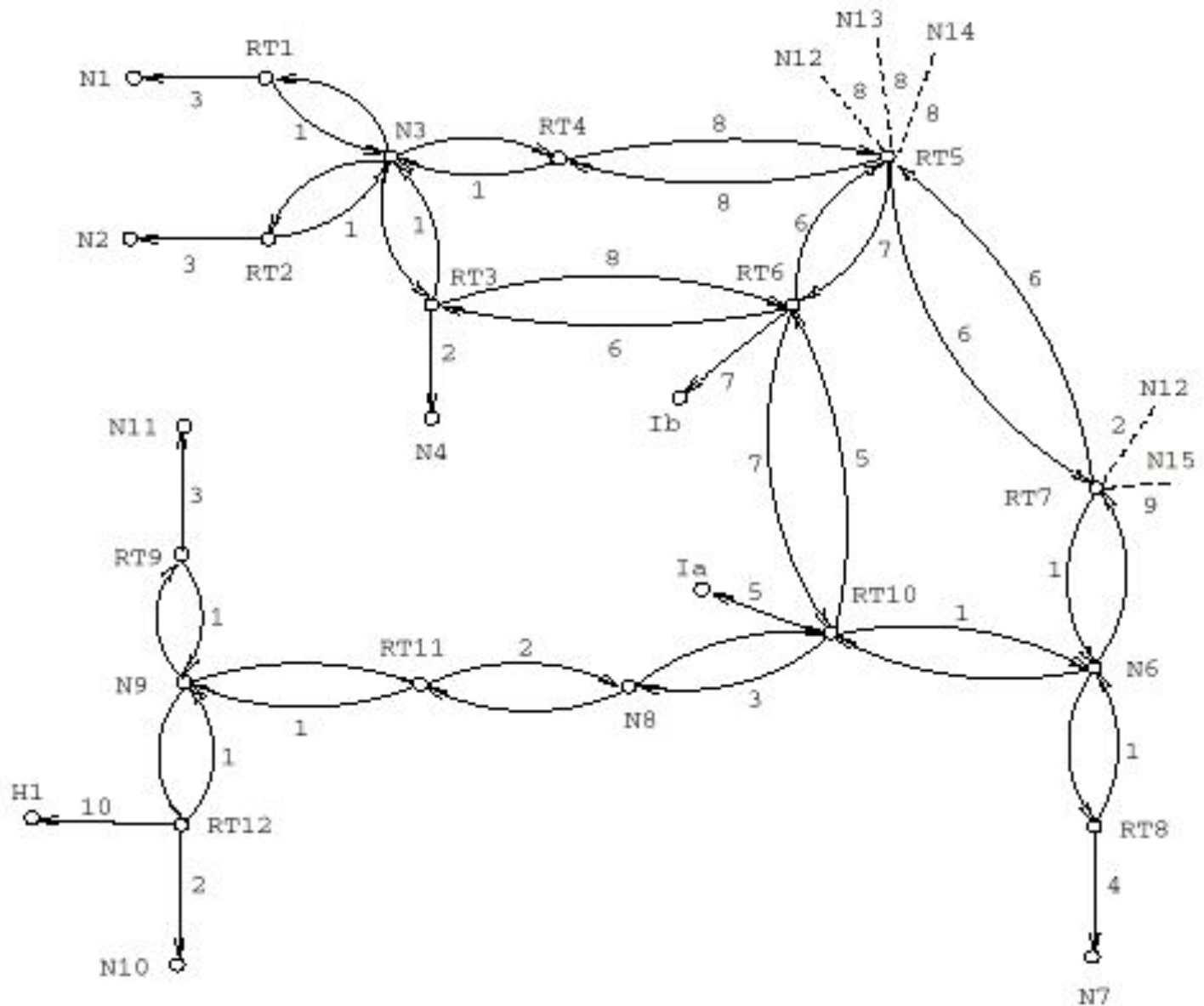
Пример построения топологической базы и таблицы маршрутизации

- A) Исходная сеть**
- B) Результирующий граф**
- C) Объявления о связях маршрутизатора**
- D) Объявления о связях сети**
- E) OSP-дерево маршрутизатора RT6**
- F) Локальные маршруты в таблице RT6**
- G)
Локальные маршруты в таблице RT6**

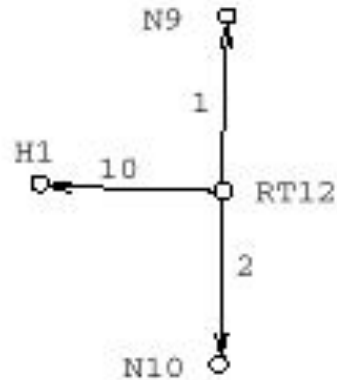
A) Исходная сеть



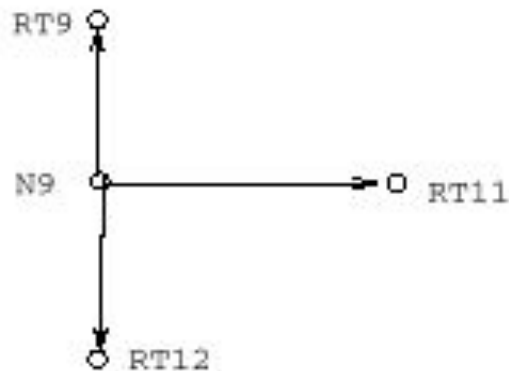
В) Результирующий граф



С) Объявления о связях маршрутизатора

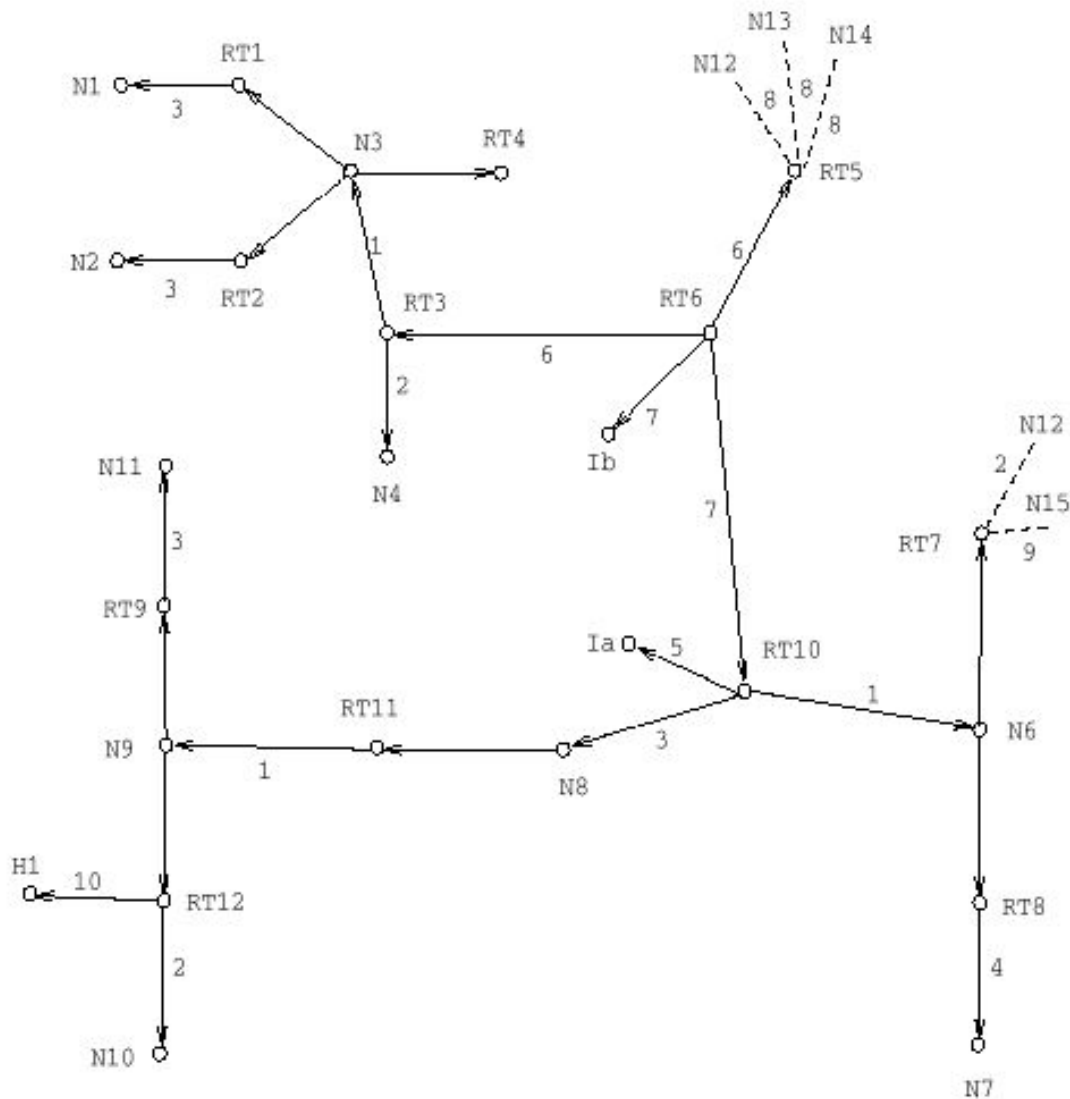


Д) Объявления о связях сети



Метрика в направлении от сети до маршрутизатора всегда равна 0

Е) OSP-дерево маршрутизатора RT6



Ф) Локальные маршруты в таблице RT6

| <i>Destination</i> | <i>Next Hop</i> | <i>Distance</i> |
|--------------------|-----------------|-----------------|
| N1 | RT3 | 10 |
| N2 | RT3 | 10 |
| N3 | RT3 | 7 |
| N4 | RT3 | 8 |
| Ib | * | 7 |
| Ia | RT10 | 12 |
| N6 | RT10 | 8 |
| N7 | RT10 | 12 |
| N8 | RT10 | 10 |
| N9 | RT10 | 11 |
| N10 | RT10 | 13 |
| N11 | RT10 | 14 |
| H1 | RT10 | 21 |
| RT5 | RT5 | 6 |
| RT7 | RT10 | 8 |

G) Внешние маршруты в таблице RT6

| <i>Destination</i> | <i>Next Hop</i> | <i>Distance</i> |
|--------------------|-----------------|-----------------|
| N12 | RT10 | 10 |
| N13 | RT5 | 14 |
| N14 | RT5 | 14 |
| N15 | RT10 | 17 |

Обработка внешних маршрутов в OSPF

- Внешние маршруты (полученные маршрутизатором по протоколам EGP или BGP) распространяются по AS без изменений
- Каждый внешний маршрут содержит IP-адрес назначения и IP-адрес маршрутизатора AS, через который следует передавать пакеты к сети назначения (Forwarding Address)
- Существует 2 типа внешней метрики: 1-й и 2-й:
 - Метрика типа 1 эквивалентна метрике, используемой OSPF внутри данной AS
 - Метрика типа 2 несоизмерима с внутренней метрикой OSPF — ее значения всегда больше любых значений внутренней метрики
- При использовании метрики типа 1 оптимальный маршрут к внешней сети находится путем суммирования внутренней и внешней метрик
- При использовании метрики типа 2 оптимальный маршрут к внешней сети находится путем использования только внешней метрики



Пример:

- Если внешняя метрика к сети N12 — 1 , то цена маршрута к сети N12 через маршрутизатор RT5 будет равна $6+8=14$, а через маршрутизатор RT7 $\rightarrow 1+7+2=10$. Выбирается маршрут через маршрутизатор RT7
- Если внешняя метрика к сети N12 — 2 , то цена маршрута к сети N12 через маршрутизатор RT5 будет равна 8 , а через маршрутизатор RT7 $\rightarrow 2$. Выбирается маршрут через маршрутизатор RT7

Разделение своей автономной системы на области

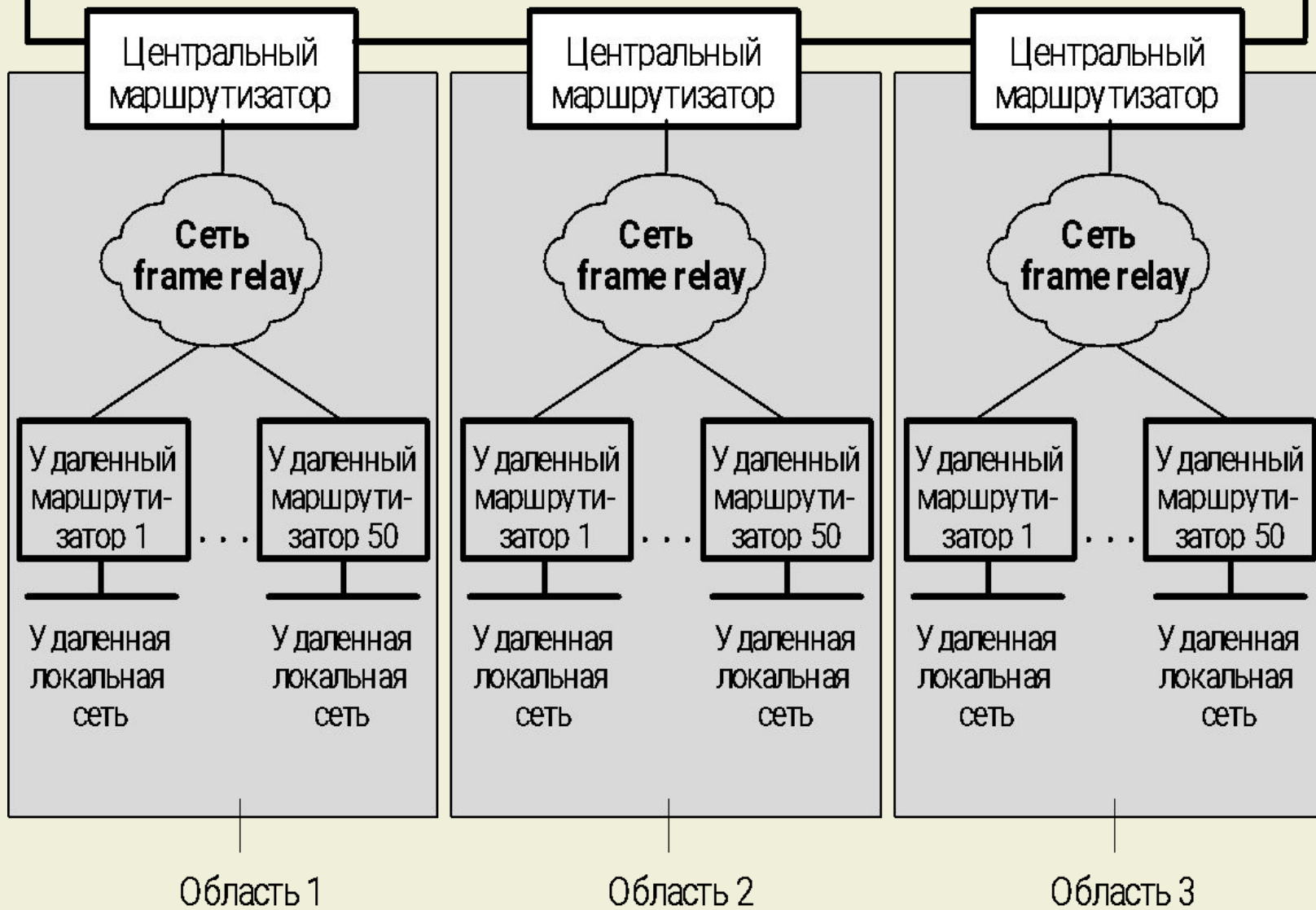
- ♦ Цель - сократить размерность графа связей сети до приемлемых величин
- ♦ Граф строится только в пределах области
- ♦ Между областями передаются вектора расстояний
- ♦ Области соединены через общую область - магистраль (нет маршрутных петель между областями)

Пограничный маршрутизатор области (ABR, area border router)
- это маршрутизатор с интерфейсами в двух или более областях,
одна из которых является магистральной (backbone area)

ABR распространяет обобщенные объявления (summary links advertisements) в формате Net – NextHop (адрес пограничного маршрутизатора) :

- ♦ объявления об отдельном маршруте
- ♦ обобщения нескольких маршрутов в наименее специфический адрес
- ♦ маршрут по умолчанию

Магистральная область (backbone)



Разделение AS на области (Areas)

Автономная система разбита на 4 области:

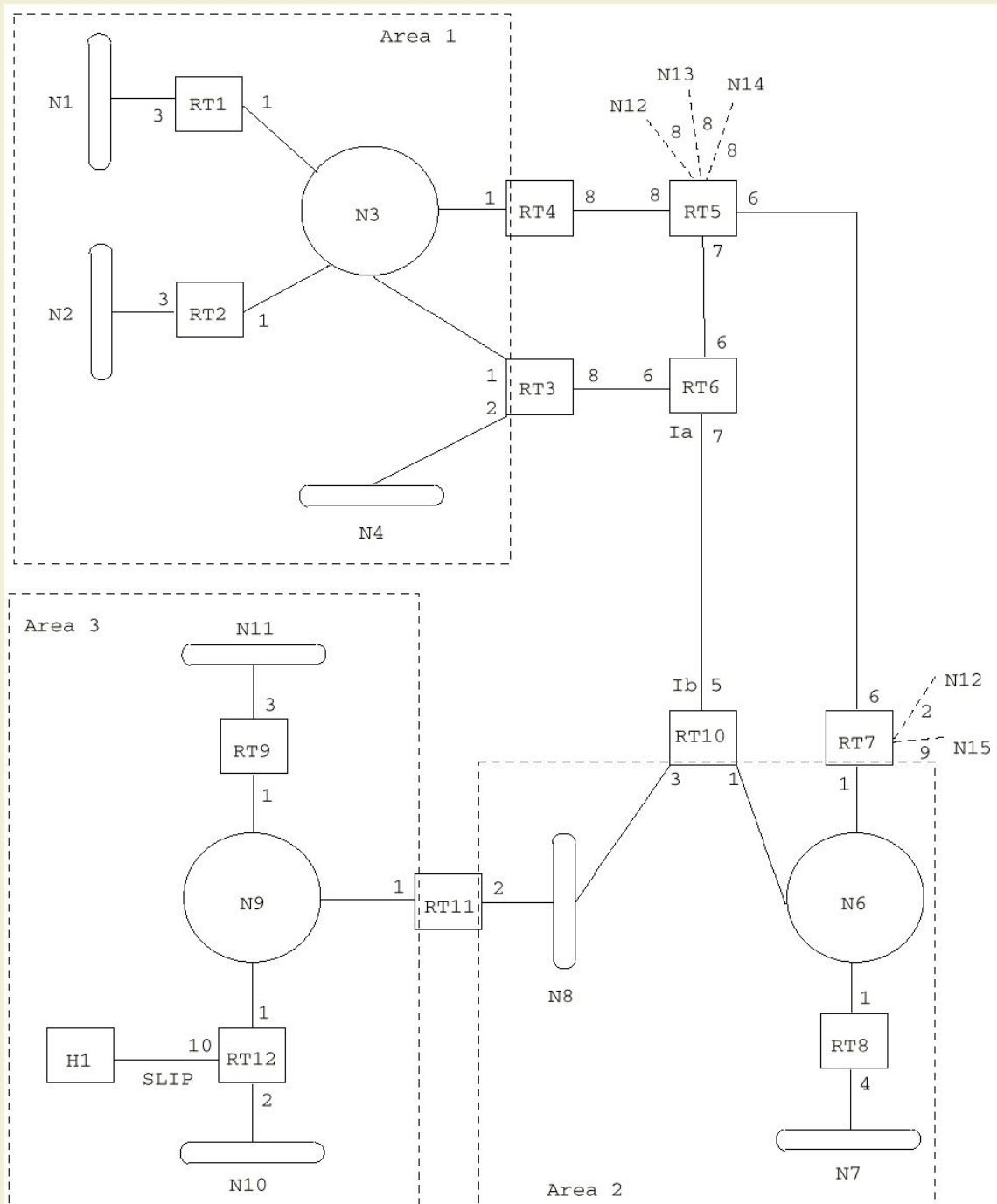
Area 1: RT1, RT2, RT3, RT4, N1, N2, N3, N4

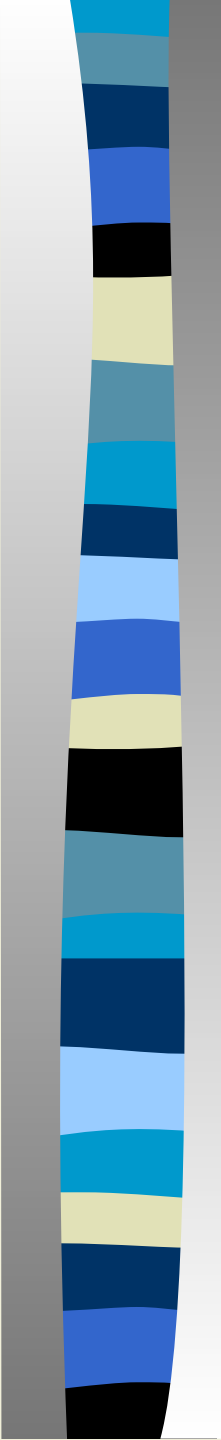
Area 2: RT10, RT7, RT8, RT11, N6, N7, N8

Area 3: RT11, RT9, RT12, N9, N10, N11, H1

Backbone (магистраль): RT3, RT4, RT5, RT6, RT7, RT10, RT11, Ia, Ibs

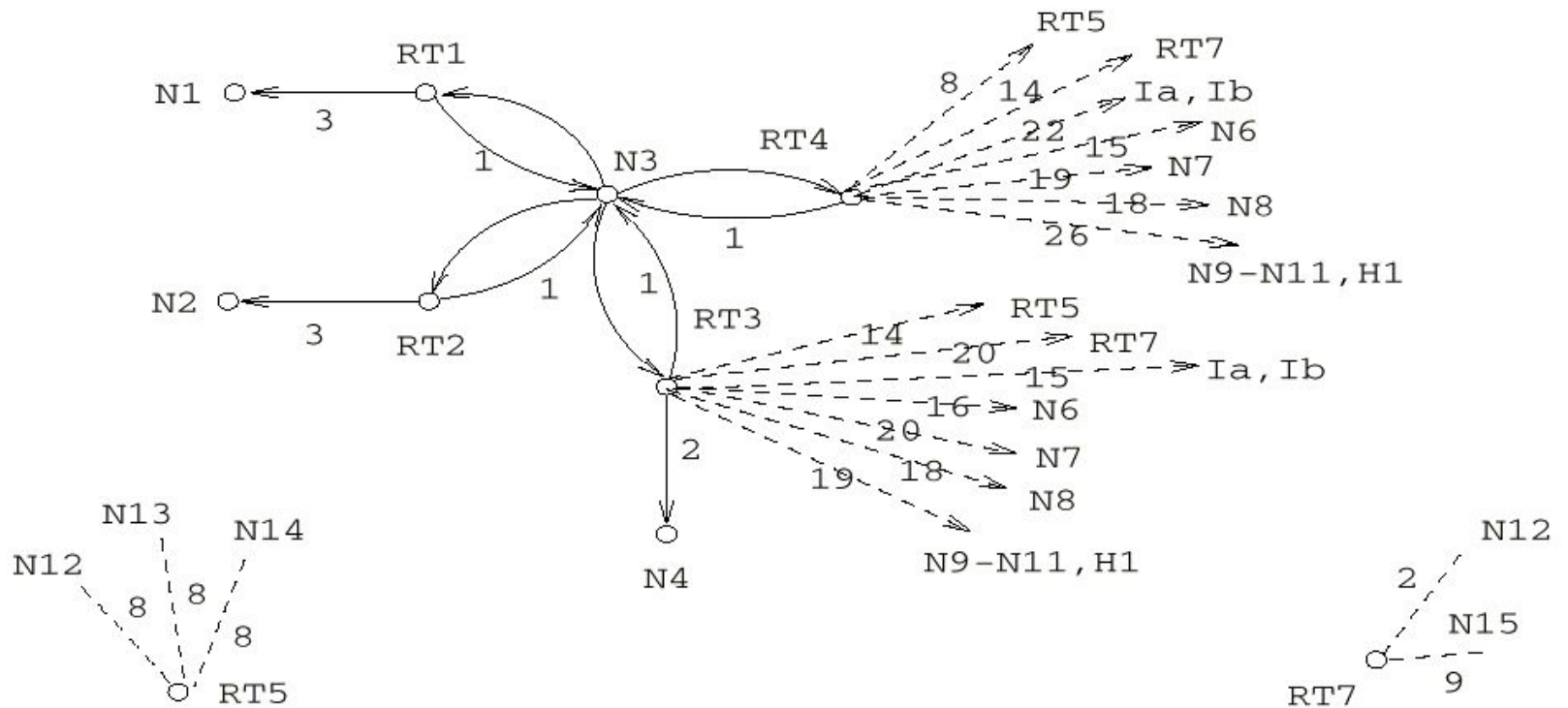
RTx — пограничные маршрутизаторы области (Area Border Router, ABR)



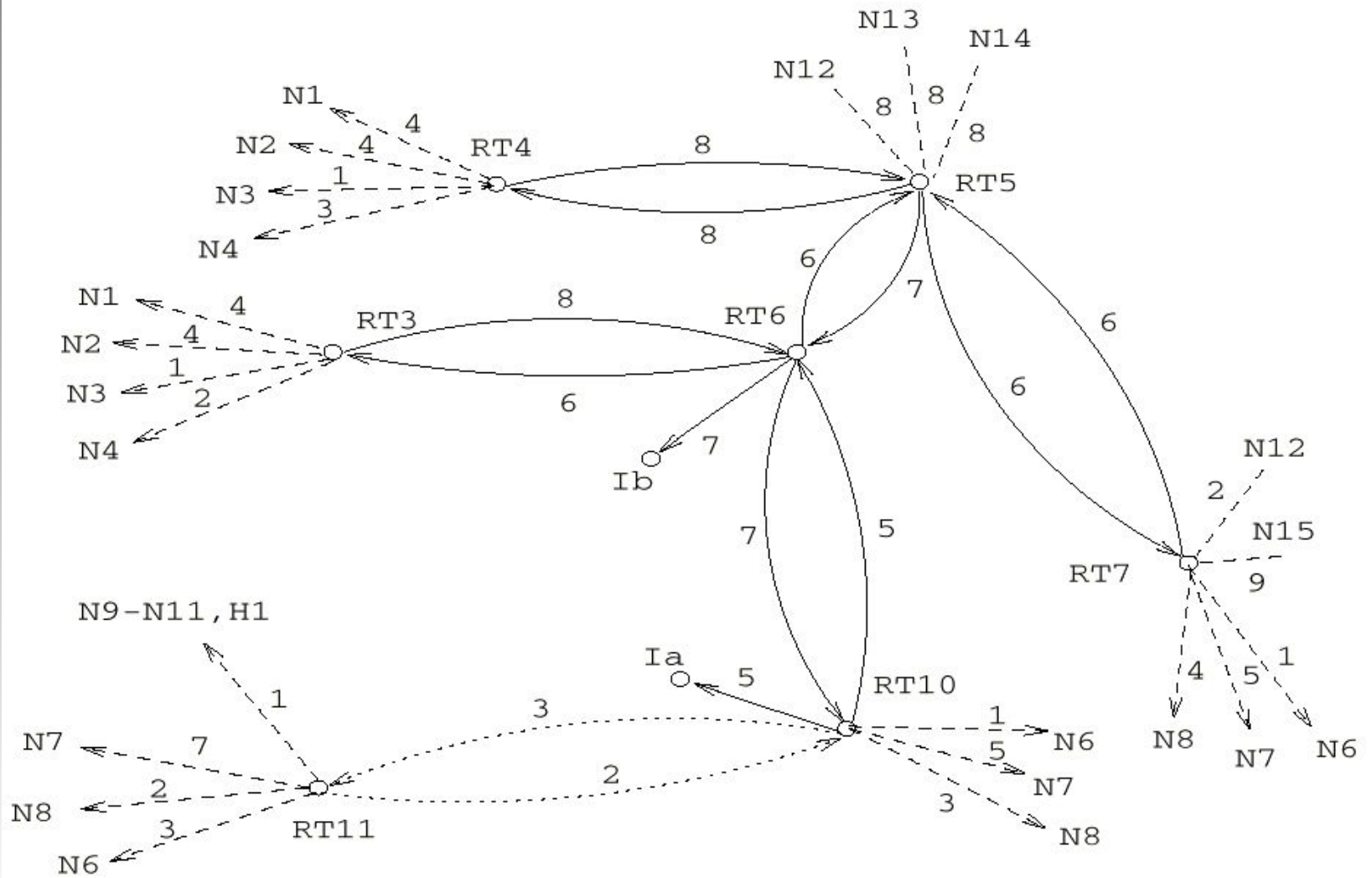
- 
- Области обмениваются объявлениями только через магистраль
 - Пограничные маршрутизаторы области распространяют обобщенные (summary) объявления о маршрутах к сетям в других областях
 - Магистраль может быть образована как физическими, так и виртуальными связями

Топологические базы областей

База области 1:



База магистрала:



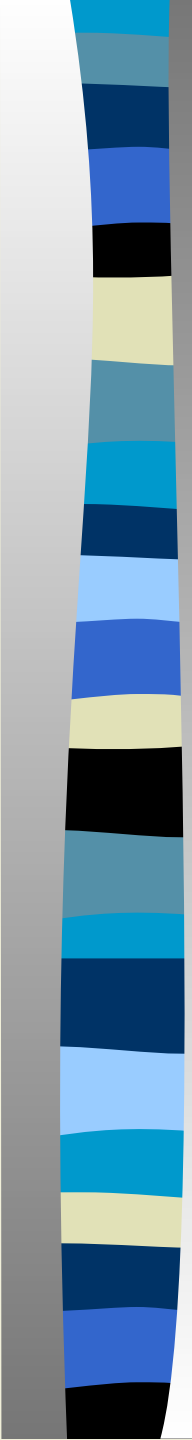
Топологические базы областей

Объявления, передаваемые пограничными маршрутизаторами из области 1 в магистраль:

| <i>Network</i> | <i>RT3 adv.</i> | <i>RT4 adv.</i> |
|----------------|-----------------|-----------------|
| N1 | 4 | 4 |
| N2 | 4 | 4 |
| N3 | 1 | 1 |
| N4 | 2 | 3 |

Расстояния, вычисленные RT3 и RT4 до остальных пограничных маршрутизаторов

| <i>Area border router</i> | <i>dist from RT3</i> | <i>dist from RT4</i> |
|---------------------------|----------------------|----------------------|
| to RT3 | * | 21 |
| to RT4 | 22 | * |
| to RT7 | 20 | 14 |
| to RT10 | 15 | 22 |
| to RT11 | 18 | 25 |
| to Ia | 20 | 27 |
| to Ib | 15 | 22 |
| to RT5 | 14 | 8 |
| to RT7 | 20 | 14 |



Объявления, передаваемые пограничными маршрутизаторами RT3 и RT4 внутренним маршрутизаторам области 1:

| <i>Destination</i> | <i>RT3 adv.</i> | <i>RT4 adv.</i> |
|--------------------|-----------------|-----------------|
| Ia,Ib | 15 | 22 |
| N6 | 16 | 15 |
| N7 | 20 | 19 |
| N8 | 18 | 18 |
| N9-N11,H1 | 19 | 26 |
| RT5 | 14 | 8 |
| RT7 | 20 | 14 |

Сравнение протоколов RIP и OSPF по затратам на широковещательный трафик

Трафик, создаваемый протоколом RIP:

$$F_{RIP} = \left(\frac{\text{число объявляемых маршрутов}}{25} \times 8 \left(\frac{\text{байтов в сообщении}}{\text{битов в байте}} \right) \times 528 \right) \times \text{число копий в единицу времени}$$

Трафик, создаваемый протоколом OSPF:

$$F_{OSPF} = \{ [20 + 24 + 20 + (4 \times \text{число соседей})] \times 8 + \text{число копий HELLO в единицу времени} \times \text{число объявлений} \times \text{средний размер объявления} \times 8, \text{число копий объявлений в единицу времени} \}$$

где

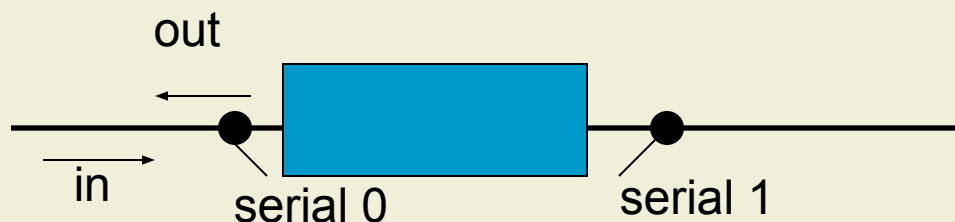
- 20 - размер заголовка IP-пакета,
- 24 - количество заголовков пакета OSPF,
- 20 - размер заголовка сообщения HELLO,
- 10 - количество HELLO на каждого соседа
- 10 - количество объявлений OSPF - каждый 30 минут

Пример: 10 маршрутизаторов, 100 сетей, маршрутизаторы соединены друг с другом через сеть frame relay

$$F_{RIP} = (100 \text{ маршрутов} / 25 \text{ маршрутов в объявлении}) \times 528 \times (10 \text{ копий} / 30 \text{ сек}) \times 8 = 5\,632 \text{ б/}$$

$$F_{OSPF} = \{ [20 + 24 + 20 + (4 \times 10)] \times (10 \text{ копий} / 10 \text{ сек}) + [100 \text{ связей} \times (32 + 24 + 20) \times (10 \text{ копий} / 30 \text{ сек})] \times 60 \text{ сек} \}$$

Фильтрация пакетов с помощью маршрутизаторов



Задача – не принимать (in) или не отправлять (out) пакеты, удовлетворяющие определенным признакам

Язык фильтрации маршрутизаторов Cisco Systems:

1. Стандартный список доступа:

```
access-list access-list-number {deny | permit} {source-address [source-wildcard] | any} [log]
```

```
interface serial 0
```

```
ip access-group 1 in – применение списка доступа
```

Определение списка доступа

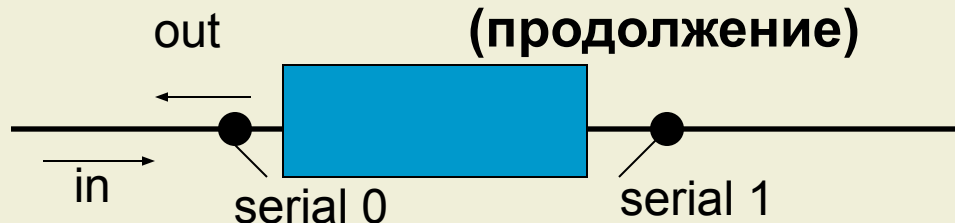
```
access-list 1 deny 192.78.46.0 0.0.255
```

```
access-list 1 deny host 193.24.15.8
```

```
access-list 1 permit any
```

[*access-list 1 deny any*] – неявный оператор в конце каждого списка

Фильтрация пакетов с помощью маршрутизаторов (продолжение)



Задача – не принимать (in) или не отправлять (out) пакеты, удовлетворяющие определенным признакам

Язык фильтрации маршрутизаторов Cisco Systems:

1. Расширенный (extended) список доступа:

```
access-list [list number] [permit|deny]
  [protocol|protocol key word]
  [source address source-wildcard mask] [source port]
  [destination address destination-wildcard mask]
  [destination port] [log options]
```

Запрещает ли список 101 доступ к серверу 192.78.46.8 по TCP, разрешая остальной доступ по IP?

```
access-list 101 permit IP any host 192.78.46.8
access-list 101 deny TCP any host 192.78.46.8
```

Запрет ping хоста 192.78.46.8:

```
access-list 101 deny ICMP any host 192.78.46.8 eq 8
```

Фильтрация объявлений протоколов маршрутизации



Маршрутизаторы Cisco Systems:

1. Подавление отправки объявлений на определенном интерфейсе:

```
passive-interface serial 0
```

2. Управление объявлением определенных маршрутов

```
distribute-list 2 out serial 1
```

```
access-list 2 deny 194.12.34.0 0.0.0.255
```

```
access-list 2 deny 132.7.0.0 0.0.255.255
```

```
access-list 2 permit any
```

3. Управление приемом определенных маршрутов

```
distribute-list 2 in serial 1
```

В ОС Unix – gated поддерживает язык фильтрации

объявлений

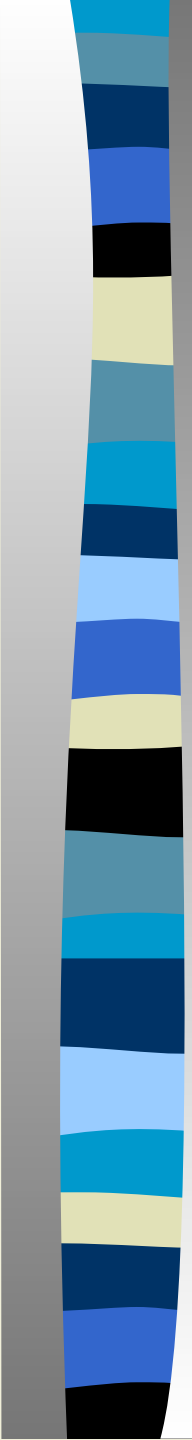
Связь между различными протоколами маршрутизации



```
interface serial 0
  ip address 130.93.1.1 255.255.255.0
!
interface ethernet 0
  ip address 130.94.1.1 255.255.255.0
!
router ospf 9000
  network 130.93.0.0 0.0.255.255 area 0
  redistribute rip metric 1
!
router rip
  network 130.94.0.0 0.0.255.255 area 0
  redistribute ospf 9000 metric 1
```

Протокол EGP

- ◆ Протокол EGP (Exterior Gateway Protocol) - 6-менас маршрутной информацией между автономными системами (AS) протоколом
- ◆ EGP успешно работает только для иерархических связей AS
- ◆ При существовании маршрутных петель - возможно закливание пакетов
- ◆ Каждая автономная система имеет уникальный 16-битный идентификатор
- ◆ Магистраль Internet рассматривается как одна автономная система, которая использует свой собственный внутренний протокол маршрутизации
- ◆ Маски и подсети не поддерживаются



Протокол EGP определяет обмен маршрутной информацией между *непосредственными соседями (direct neighbor)*

Два EGP-маршрутизатора могут стать непосредственными соседями в том случае, если:

- ♦ они связаны через общую сеть (локальную или канал "точка-точка") или
- ♦ они связаны через интерсеть, которая прозрачна для них, то есть для обмена сообщениями маршрутизаторам не нужно знать внутреннюю топологию связей этой интерсети

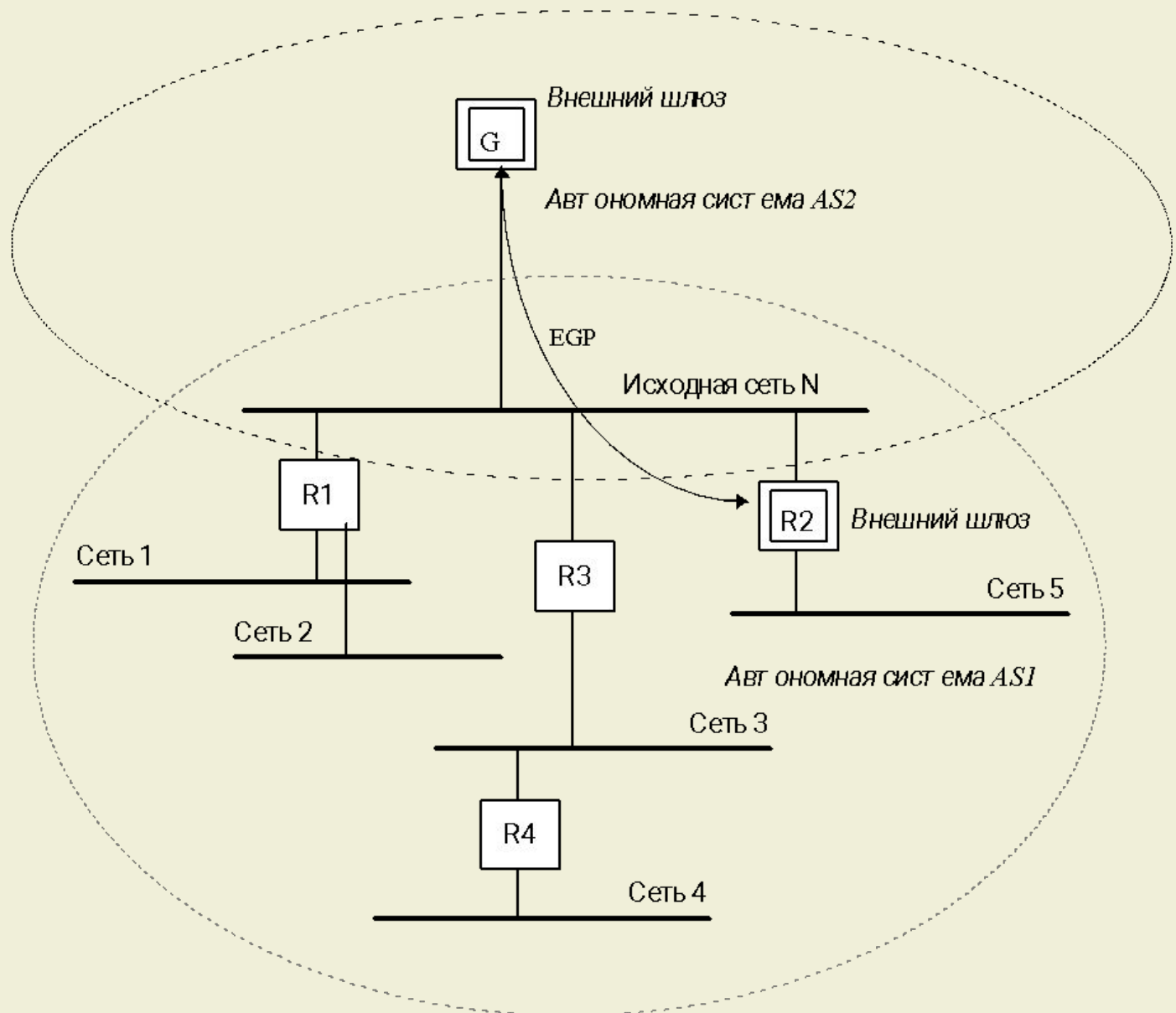


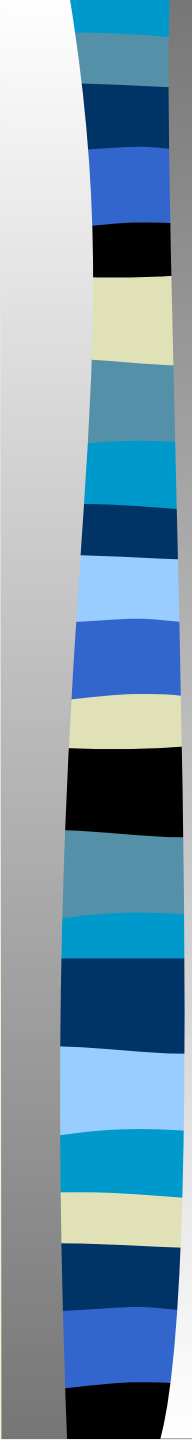
В протоколе EGP определены три основные функции:

- ◆ установление соседских отношений
- ◆ подтверждение достижимости соседа
- ◆ обновление маршрутной информации о достижимости сетей в AS

С помощью протокола достижимости сетей внешний маршрутизатор передает список сетей, для которых данный маршрутизатор является наилучшим входным маршрутизатором относительно исходной общей сети N

Пример двух соседних автономных систем, объединяемых протоколом EGP



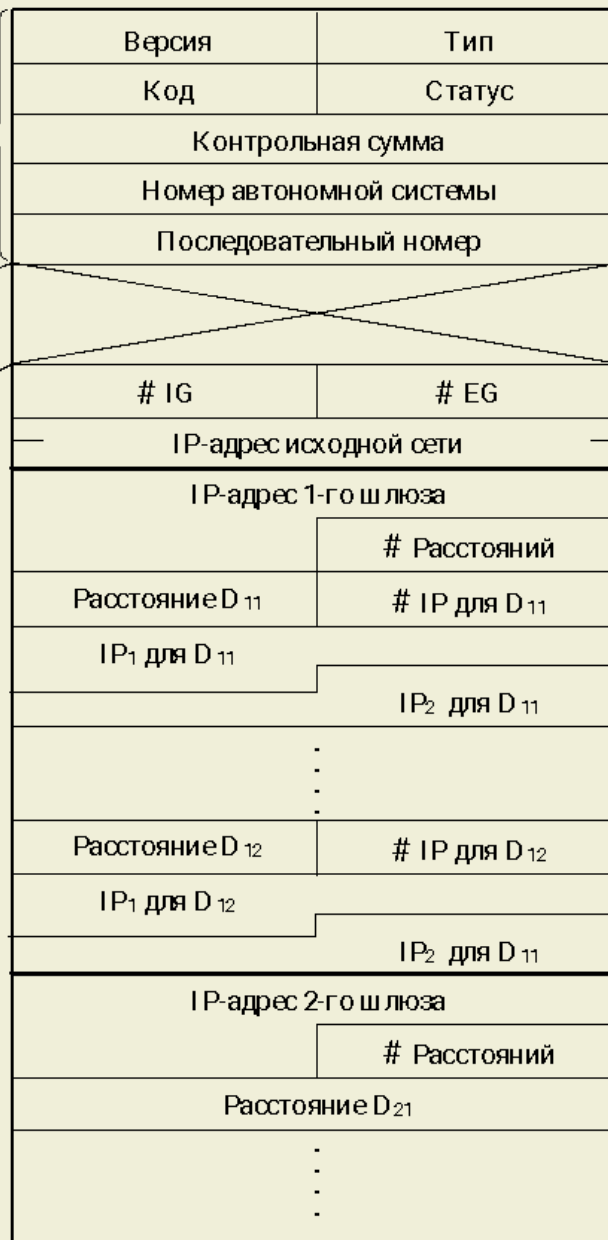
- 
- ◆ Маршрутизаторы R2 и G - внешние "соседи" по EGP
 - ◆ Маршрутизаторы R2, R1 и R3 - внутренние "соседи" по EGP
 - ◆ Маршрутизатор R1 и R3 - "косвенные" соседи G через R2
 - ◆ Маршрутизатор R2 передает G свои данные и данные маршрутизаторов R1 и R3
 - ◆ "
 - ◆ "Соседи" назначаются вручную
 - ◆ Список сетей для объявлений формируется вручную или берется от внутреннего протокола RIP
 - ◆ Объявления о достижимости сетей выполняются на основе схемы "запрос - ответ"

Общий формат сообщений протокола EGP

Биты

1 8 9 16

фиксированный заголовок



Тип = 5 = достижимость
 Код = 0 = привет
 = 1 = я-услышал-тебя

HELLO-интервал
 Интервал опроса

Тип = 3 = установление соседских отношений
 Код = 0 = запрос устан.
 = 1 = подтвержд уст.
 = 2 = отказ
 = 3 = запрос разрыва
 = 4 = подтв. разрыва

не используется
 IP-адрес исходной сети

Тип = 2 = запрос данных
 Код = 0 = привет

Тип = 1

Протокол BGP (Border Gateway Protocol) v.4

- ◆ Поддерживает произвольную структуру связей между автономными системами, в том числе и петли
- ◆ Передает маршрут достижимости - список AS, которые нужно пересечь для достижения некоторых сетей - для исключения петель
- ◆ Поддерживает "бесклассовую" маршрутизации, основанную на стратегии CIDR
- ◆ Использует надежный транспортный протокол TCP для обмена сообщениями
- ◆ Таблица маршрутизации пересылается целиком только при инициализации соединения между маршрутизаторами, а затем пересылаются только обновления маршрутов

- Протокол BGP поддерживает "**policy based**" маршрутизацию
- выбор маршрутов для объявления в другие AS производится на основании данных о сетях + набора правил, задаваемых администратором системы
- В одной AS может быть несколько BGP-маршрутизаторов, которые должны поддерживать согласованность своих таблиц с помощью внутреннего протокола маршрутизации (например, OSPF)

Основной единицей информации, которой оперируют маршрутизаторы BGP, является маршрут (route)

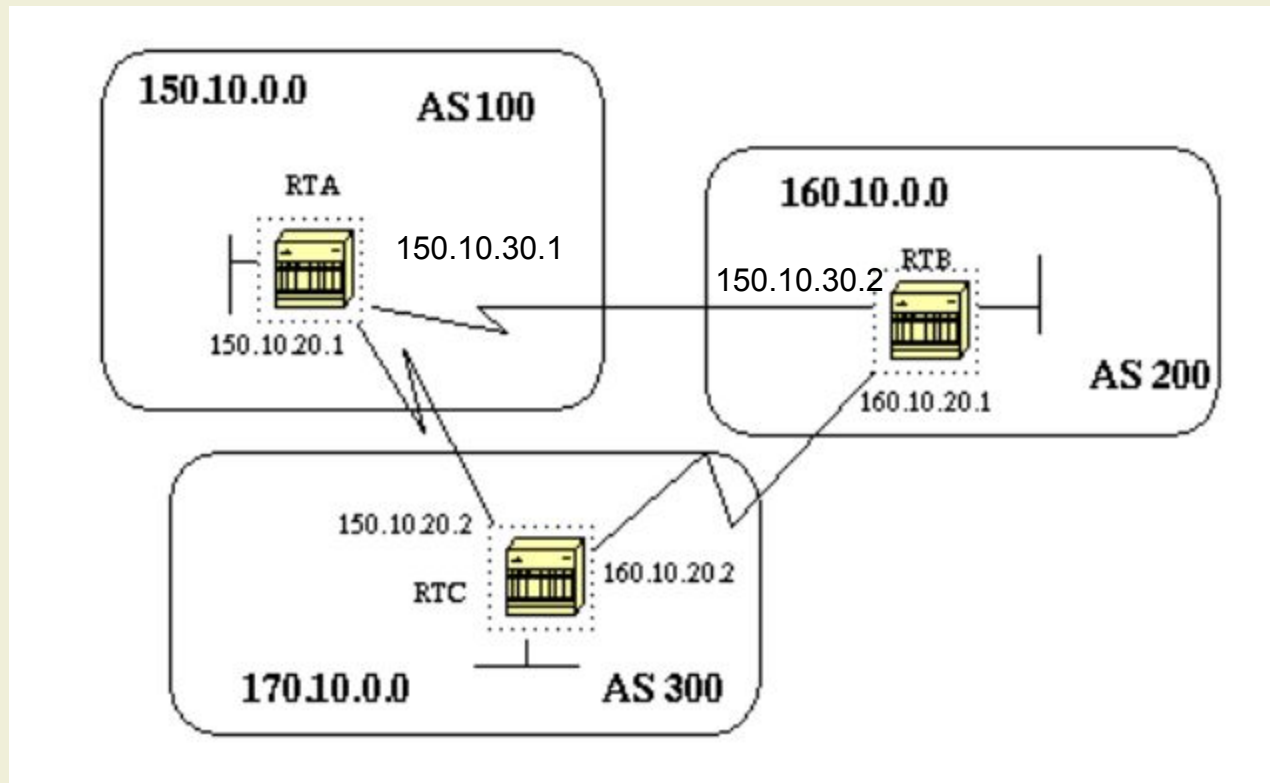
Маршрут

назначения характеризуется IP-адресами нескольких сетей

Путь -

это последовательность промежуточных автономных систем и IP-адреса портового маршрутизатора в автономной системе, к которой принадлежит сеть назначения

Исключение петель в BGP

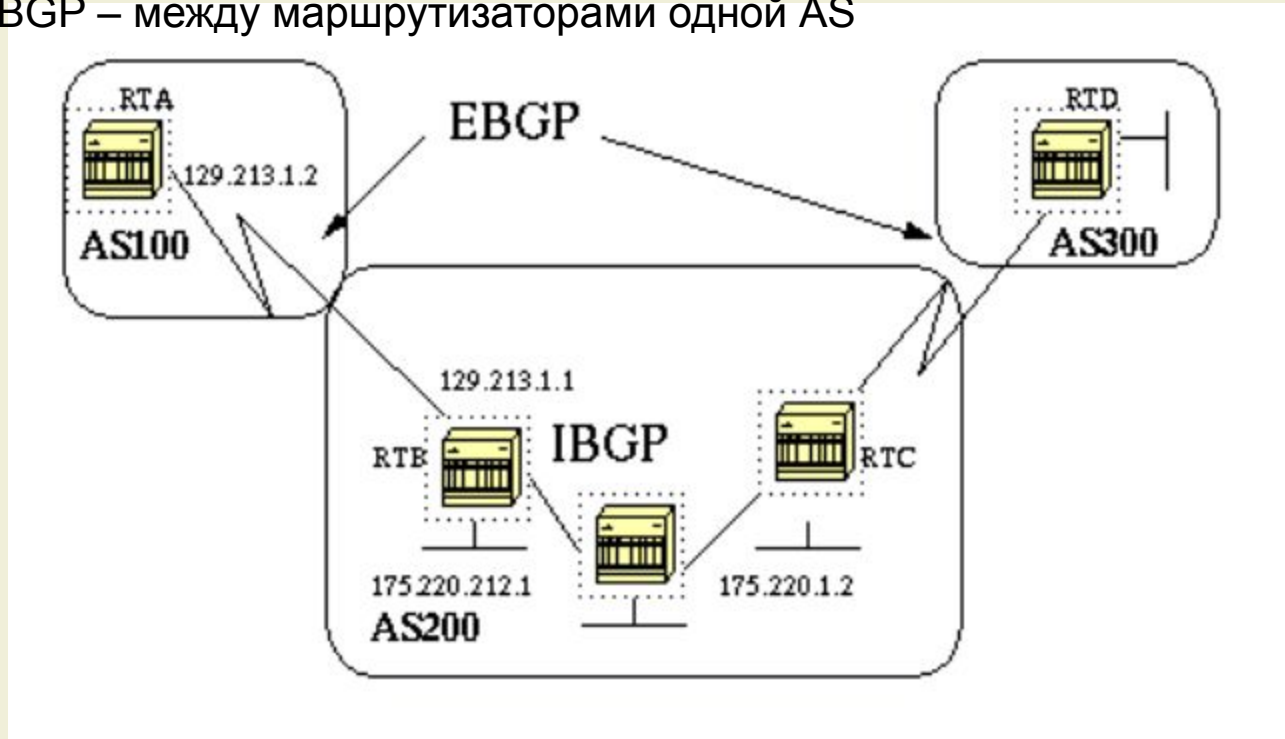


1. Объявление RTA->RTB: AS100; NH=150.10.30.1; Net=150.10.0.0 255.255.0.0
2. Объявление RTB->RTC: AS200, AS100; NH = 160.10.20.1; Net=150.10.0.0 255.255.0.0
3. Объявление RTC>RTA: AS300, AS200, AS100; NH = 150.10.20.2; Net=150.10.0.0 255.255.0.0 – отбраковывается, так как в пути встретился номер своей AS

Внешнее и внутреннее использование BGP

Exterior BGP – между AS

Interior BGP – между маршрутизаторами одной AS



RTA#

```
router bgp 100 ( – номер AS)
```

```
neighbor 129.213.1.1 remote-as 200
```

RTB#

```
router bgp 200
```

```
neighbor 129.213.1.2 remote-as 100
```

```
neighbor 175.220.1.2 remote-as 200
```

RTC#

```
router bgp 200
```

```
neighbor 175.220.212.1 remote-as 200
```


Фильтрация объявлений в BGP

RTC#

```
router bgp 300
```

```
neighbor 2.2.2.2 remote-as 100
```

```
neighbor 2.2.2.2 route-map  
STOPUPDATES out
```

```
route-map STOPUPDATES permit 10
```

```
match ip address 1
```

```
access-list 1 deny 170.10.0.0  
0.0.255.255
```

```
access-list 1 permit 0.0.0.0  
255.255.255.255
```

match as-path

match community

match clns

match interface

match ip address

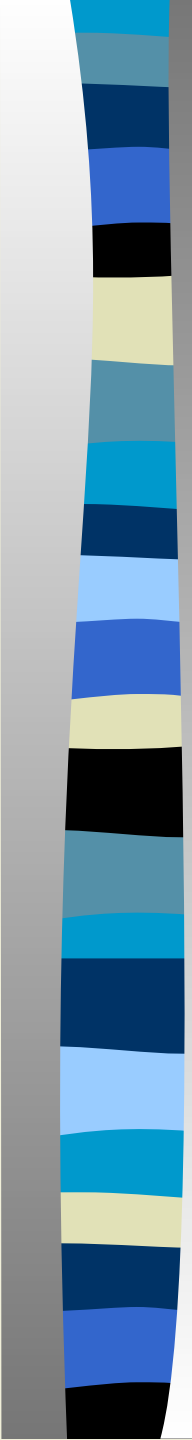
match ip next-hop

match ip route-source

match metric

match route-type

match tag



Организация базы маршрутной информации в BGP

Вся маршрутная информация хранится в базе данных маршрутов - Routing Information Base

RIB

состоит из трех частей:

- ◆ Adj-RIBs-In -
необработанные маршруты, поступающие из других AS и из своей собственной AS в сообщениях UPDATE протокола BGP.
- ◆ Loc-RIB -
локальная маршрутная информация, которую маршрутизатор BGP выбирает в соответствии с его локальной политикой из базы Adj-RIBs-In.
- ◆ Adj-RIBs-Out -
маршруты, выбранные в соответствии с другой BGP маршрутизатором для объявления другим BGP маршрутизаторам



Типы сообщений BGP

Протокол BGP использует четыре типа сообщений:

- OPEN** - открывает логическое соединение между BGP-маршрутизаторами
- UPDATE** - содержит данные о новом маршруте или отзываемых маршрутах
- NOTIFICATION** - предназначены для уведомления об ошибках, обнаруженных в полученных сообщениях UPDATE или OPEN
- KEEPALIVE** - служат для поддержания сессии между соседями в открытом состоянии



Формат сообщения UPDATE

В одном сообщении UPDATE маршрутизатор может объявить один достижимый маршрут или отозвать несколько недостижимых маршрутов:

| |
|---|
| Unfeasible Routes Length (2 octets) |
| Withdrawn Routes (variable) |
| Total Path Attribute Length (2 octets) |
| Path Attributes (variable) |
| Network Layer Reachability Information (variable) |



- ◆ **Withdrawn Routes** -

маршруты описывает отзываеемые недостижимые

Каждый маршрут в поле Withdrawn Routes описывается парой чисел:

- ◇ Длина префикса (1 байт)
- ◇ Префикс (переменной длины)

- ◆ **Network Layer Reachability Information** -

сетей, к которым ведет новый маршрут, описанный полем Path Attributes

Перечень сетей задается парами значений "Длина префикса - Префикс"

- ◆ **Path Attributes** о атрибутах путей между AS:

➤ **AS_PATH** содержит данные

представляет собой последовательность номеров автономных систем, которые должен пересечь маршрут, описываемый сообщением UPDATE. Исходная информация для отбраковки петель

➤ **NEXT_HOP** - маршрутизатора, через который нужно передавать пакеты сетям, которые указаны в поле Network Layer Reachability

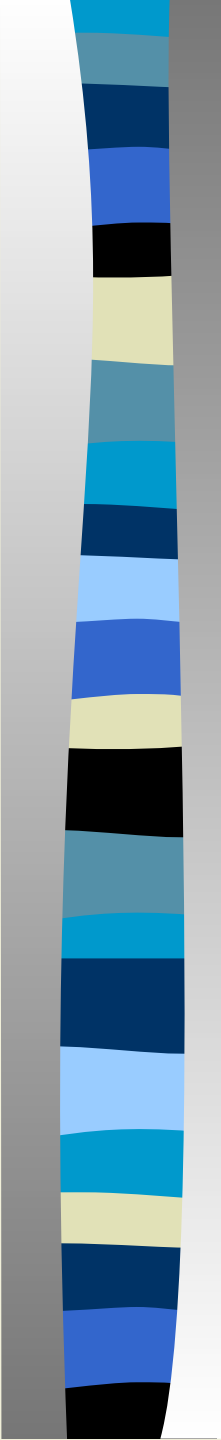


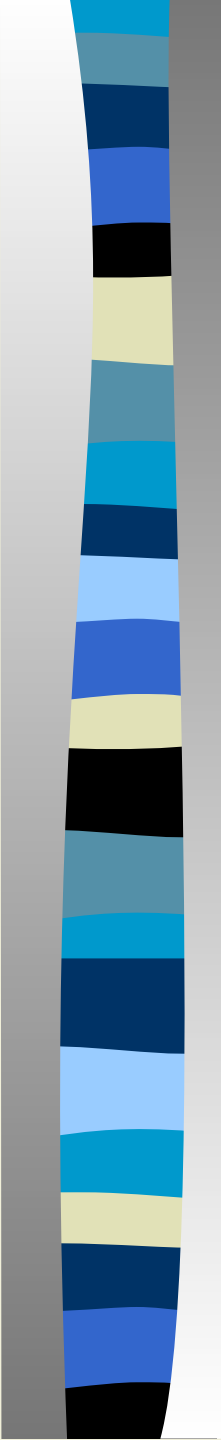
◆ Обработка маршрутов по протоколу BGP

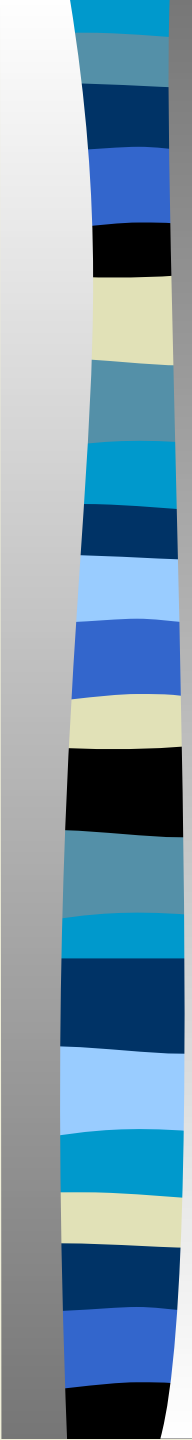
Этап 1. Проверяется, не совпадает ли новый маршрут с одним из имеющихся в базе Adj-RIB-In.

списка достижимых сетей. Для этого совпадение списка достижимых сетей, указанных в поле Network Layer Reachability, с списком сетей каждого маршрута, хранящегося в Adj-RIB-In

- ◆ Если список сетей маршрута совпадает со списком какого-либо маршрута, но путь, указанный в параметре Path Attributes, отличается от пути этого маршрута, то такой маршрут называется **новым** Adj-RIB-In. Он запускается Decision Process, и удаляется старый из базы. Затем не может использоваться так как старый маршрут больше

- 
- ◆ Маршрут называется перекрывающимся (overlapping route), если его список сетей имеет общее подмножество со списком сетей какого-либо другого маршрута
 - ◇ Если новый маршрут является перекрывающимся и более специфическим, то вызывается Decision Process. Например, существует маршрут 194.27.0.0/16, новый - 194.27.192.0/18.
 - ◇ Если новый маршрут имеет параметры пути, содержащиеся в каком-либо имеющемся в Adj-RIB-In маршруте, и является более специфическим, то никакие действия не выполняются

- 
- ◆ Если новый маршрут содержит список сетей, который не совпадает ни с одним списком, уже имеющимся в Adj-RIB-In, Adj-RIB-In, то он должен быть помещен в Decision Process а затем должен быть запущен Decision Process
 - ◆ Если новый маршрут является перекрывающимся и менее специфическим, чем имеющийся, то нужно запустить Decision Process назначения, которые надписываются только менее специфическим маршрутом



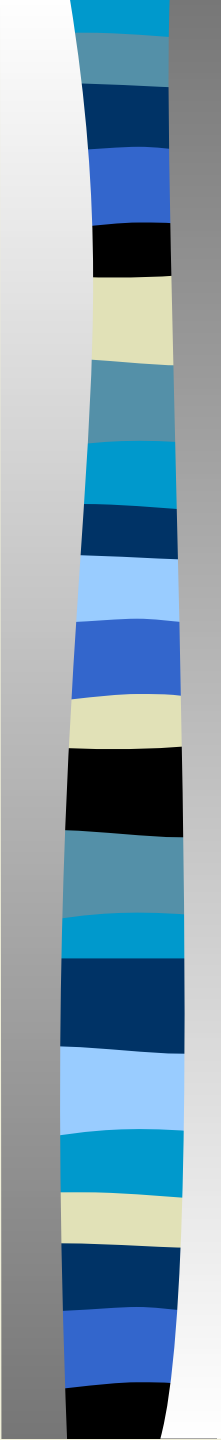
Этап 2. Decision Process выбирает маршруты для последовательного объявления другим BGP-маршрутизаторам путем применения к маршрутам, хранящимся в Adj-RIB-In, Policy Information Base - PIB. правил из локальной базы

Decision Process
предпочтения маршрутов формализуется путем определения функции

Функция предпочтения не должна принимать во внимание:

- ⇒ существование других маршрутов
- ⇒ сущ
- ⇒ параметры других маршрутов
- ⇒ отсутствие других маршрутов
- ⇒ некоторый маршрут должен быть оценен "сам по себе"

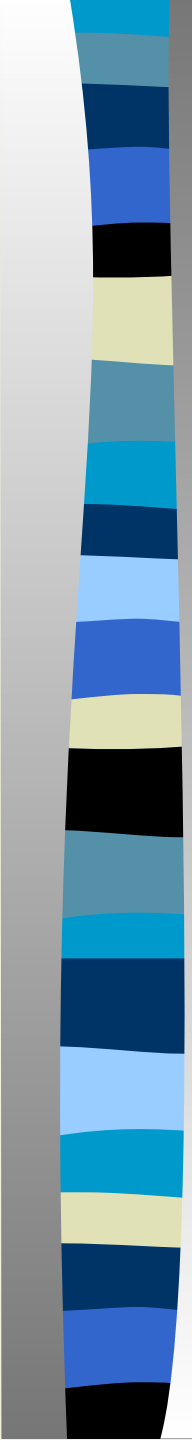
После оценки всех маршрутов из совпадающих маршрутов выбирается тот, у которого степень предпочтения выше



Decision Process

ответственен за:

- ♦ выбор маршрутов для объявлений BGP-маршрутизаторам, находящимся в своей автономной системе
- ♦ выбор маршрутов для объявлений BGP-маршрутизаторам, находящимся в соседних автономных системах
- ♦ агрегирование маршрутной информации за счет поглощения более специфических маршрутов менее специфическими в том случае, если пути у этих маршрутов совпадают



Decision Process

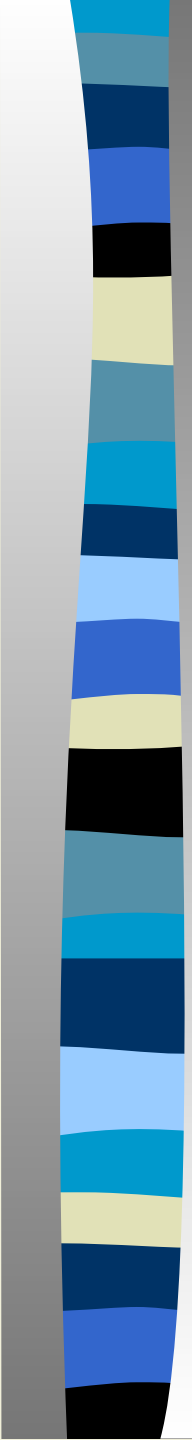
состоит из трех фаз

Фаза 1:

- ♦ вычисление степени предпочтения для каждого маршрута, полученного от соседней автономной системы
- ♦ объявление BGP-маршрутизаторам из своей автономной системы всех маршрутов, которые имеют высшую степень предпочтения
- ♦ Стандарт BGP не определяет вида этих правил выбора маршрута, за исключением одного - если в маршруте уже имеется номер своей автономной системы, то это значит, что маршрут содержит петлю и поэтому он должен быть отбракован
- ♦ Любые другие правила администратор AS может сконфигурировать в том виде, в каком их позволяет конфигурировать разработчик BGP-маршрутизатора

Примеры правил:

- ♦ запрещение использования некоторых AS как транзитных в путях до сетей назначения
- ♦ предпочтение одним AS перед другими для транзита

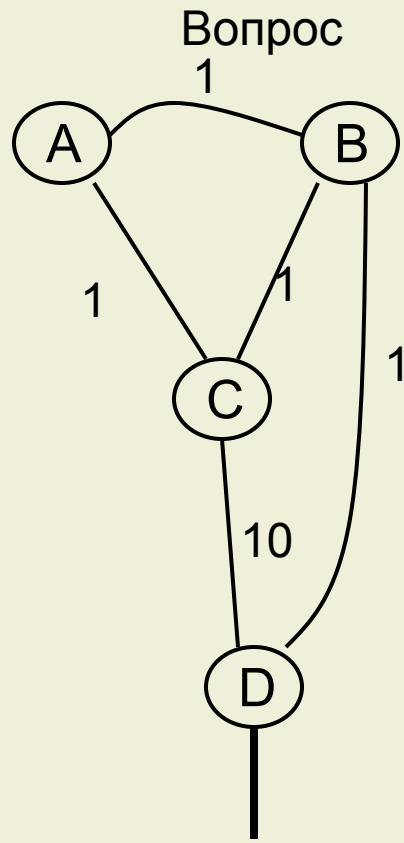


Фаза 2

- ◆ Вызывается по завершению фазы 1
- ◆ Ответственна за выбор наилучшего маршрута из всех, имеющихся для каждого отличного от других набора сетей назначения, и установки этого маршрута в локальную базу Loc-RIB. В Loc RIB включаются также и сети своей AS
- ◆ При установке маршрута BGP-маршрутизатор ДОЛЖЕН определить непосредственный следующий маршрутизатор к адресу, приведенному в поле NEXT_HOP, с помощью внутреннего протокола маршрутизации, то есть протокола IGP, используемого в данной автономной системе

Фаза 3

- ◆ Вызывается после завершения работы фазы 2
- ◆ Ответственна за распространение маршрутов из Loc-RIB BGP-маршрутизаторам из соседних AS в соответствии с правилами, содержащимися в базе RIB
- ◆ Правила могут запрещать:
 - ⇒ объявление некоторых сетей из своей или других областей
 - ⇒ объявления своих сетей для некоторых AS
 - ⇒ объявления сетей всех других AS для некоторой AS
- ◆ Фаза 3 выполняет также агрегирование маршрутов
- ◆ При передаче маршрута, описывающего сети назначения, находящиеся в другой автономной системе, BGP-маршрутизатор должен модифицировать значения полей:
 - ⇒ Path Attributes -
 - ⇒ добавить свой номер AS
 - ⇒ NEXT_HOP - указать маршрутизатор, который соединен с общей исходной сетью с той AS, куда идет объявление

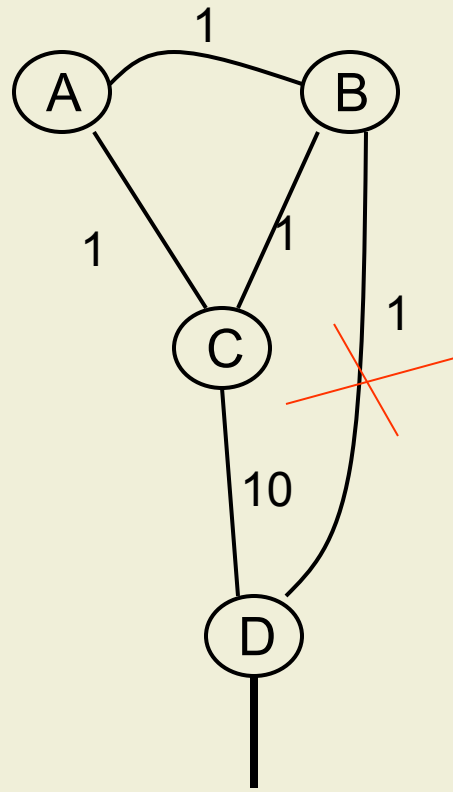


Имеется сеть:

Сеть
назначения

Записи о пути к сети назначения

- D: direct, metric 1
- B: route via D, metric 2
- C: route via B, metric 3
- A: route via B, metric 3



Сеть
назначения

Связь B-D отказывает
Маршрутизаторы не поддерживают Split Horizon

Что происходит в сети при наихудшей синхронизации объявлений?

Записи о пути к сети назначения

| | | | |
|----|--------|--|--------|
| D: | dir, 1 | | dir, 1 |
| B: | 16 | | |
| C: | B, 3 | | |
| A: | B, 3 | | |

Конфигурирование маршрутизаторов при работе через нешироковещательные (NBMA) сети

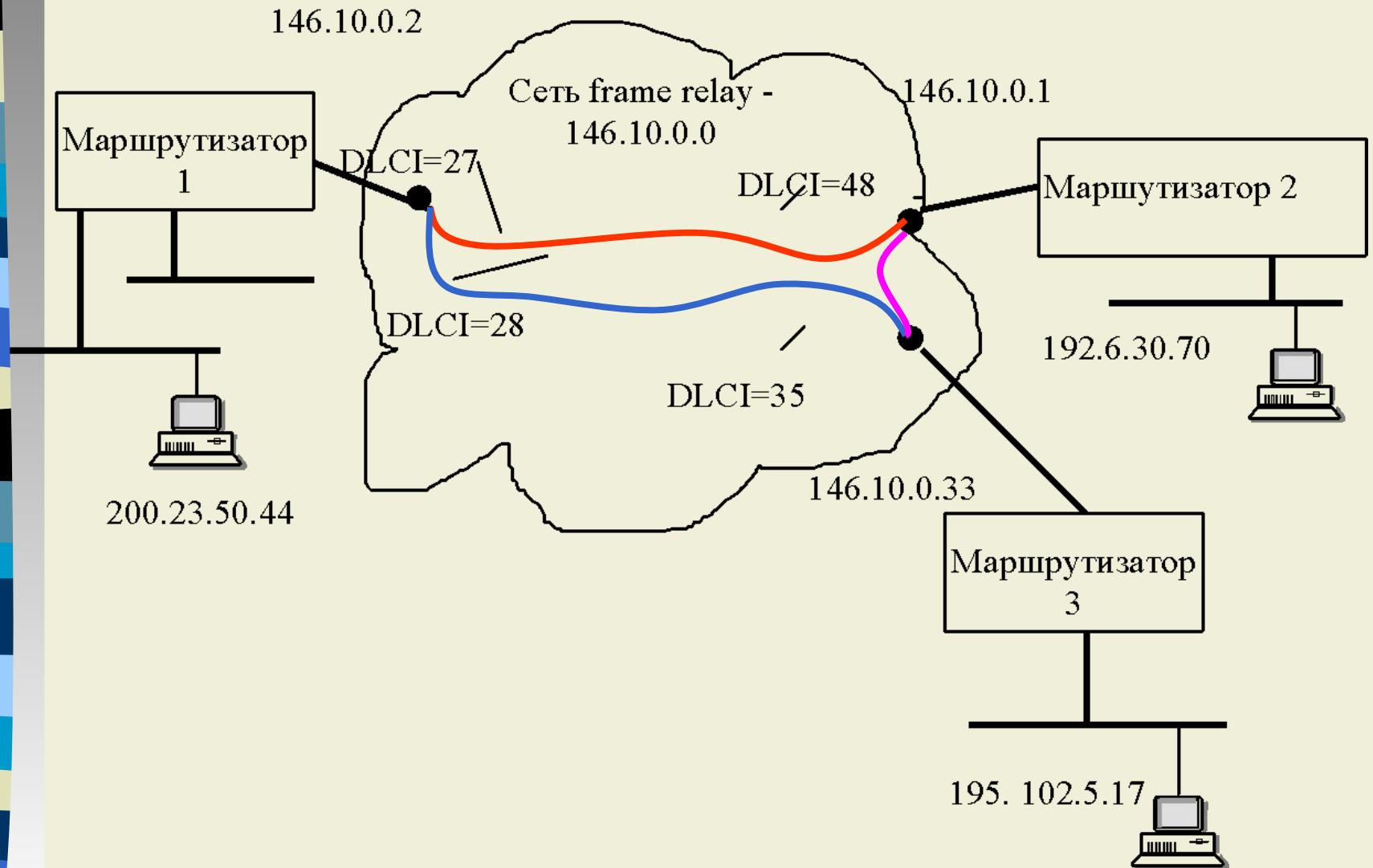


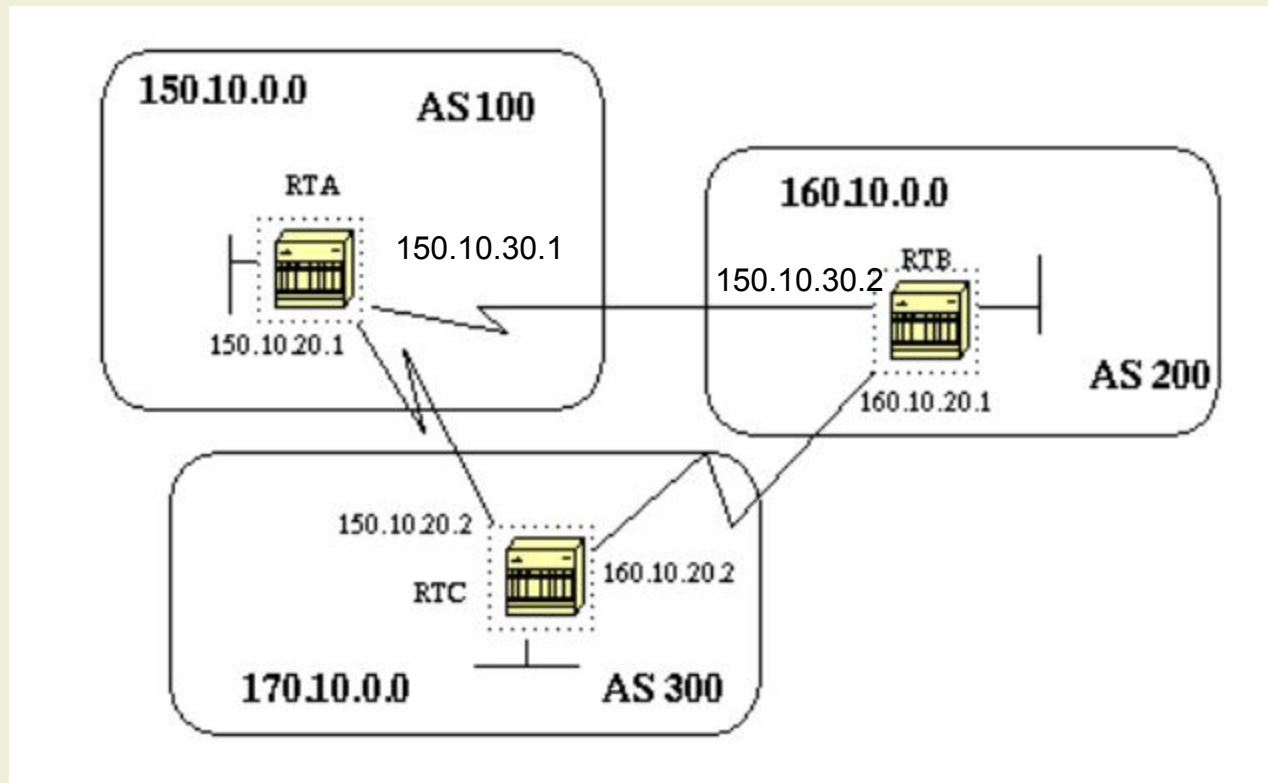
Таблица маршрутизации M1:

| Сеть | Маска | Следующий маршрутизатор | Интерфейс | Метрика |
|-------------|---------------|-------------------------|------------|---------|
| 146.10.0.0 | 255.255.0.0 | 146.10.0.2 | 146.10.0.2 | 1 |
| 200.23.50.0 | 255.255.255.0 | 146.10.0.2 | 146.10.0.2 | 1 |
| 192.6.30.0 | 255.255.255.0 | 146.10.0.1 | 146.10.0.2 | 2 |
| 195.102.5.0 | 255.255.255.0 | 146.10.0.33 | 146.10.0.2 | 2 |

ARP-таблица маршрутизатора M1:

| | |
|-------------|----|
| 146.10.0.1 | 27 |
| 146.10.0.33 | 28 |

Исключение петель в BGP

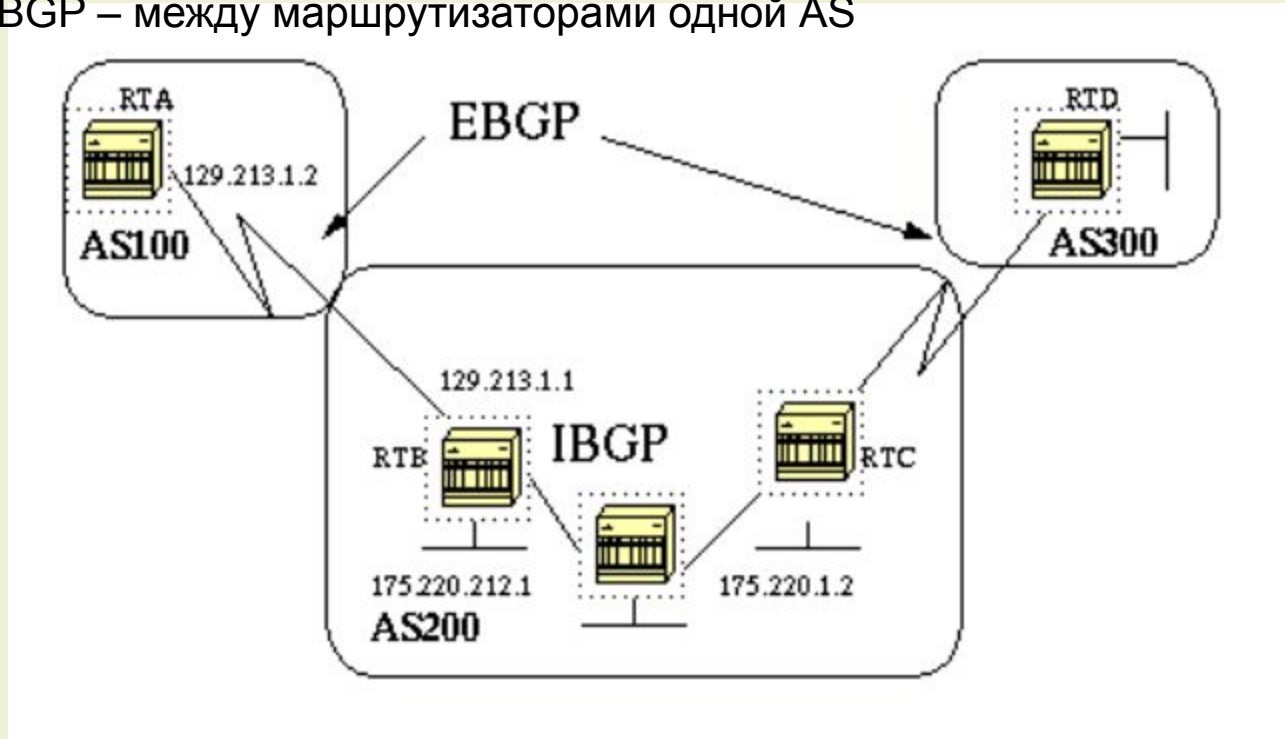


1. Объявление RTA->RTB: AS100; NH=150.10.30.1; Net=150.10.0.0 255.255.0.0
2. Объявление RTB->RTC: AS200, AS100; NH = 160.10.20.1; Net=150.10.0.0 255.255.0.0
3. Объявление RTC>RTA: AS300, AS200, AS100; NH = 150.10.20.2; Net=150.10.0.0 255.255.0.0 – отбраковывается, так как в пути встретился номер своей AS

Внешнее и внутреннее использование BGP

Exterior BGP – между AS

Interior BGP – между маршрутизаторами одной AS



RTA#

```
router bgp 100 ( – номер AS)
```

```
neighbor 129.213.1.1 remote-as 200
```

RTB#

```
router bgp 200
```

```
neighbor 129.213.1.2 remote-as 100
```

```
neighbor 175.220.1.2 remote-as 200
```

RTC#

```
router bgp 200
```

```
neighbor 175.220.212.1 remote-as 200
```

Фильтрация объявлений в BGP

RTC#

```
router bgp 300
```

```
neighbor 2.2.2.2 remote-as 100
```

```
neighbor 2.2.2.2 route-map  
STOPUPDATES out
```

```
route-map STOPUPDATES permit 10
```

```
match ip address 1
```

```
access-list 1 deny 170.10.0.0  
0.0.255.255
```

```
access-list 1 permit 0.0.0.0  
255.255.255.255
```

match as-path

match community

match clns

match interface

match ip address

match ip next-hop

match ip route-source

match metric

match route-type

match tag