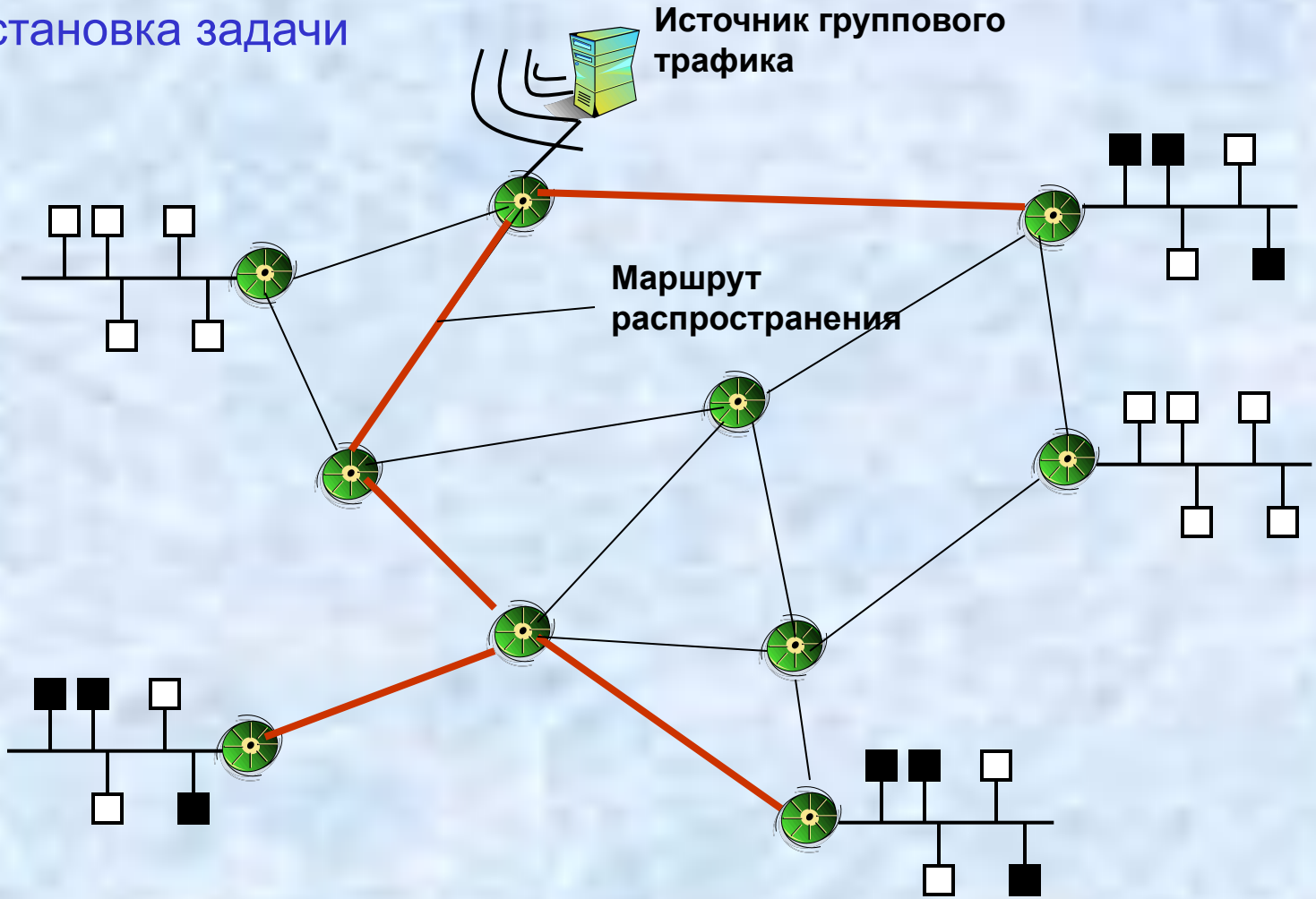
Групповое вещание - Multicasting

Постановка задачи



□ - узел, не принадлежащий группе

■ - узел, принадлежащий группе

Принципы группового вещания

(Steve Deering в 1988 г.)

4Источник посылает пакеты в любое время по групповому адресу (класс D - 224.0.0.0 и выше) по протоколу UDP

4Открытые группы

- Источник не должен знать состав группы
- Источник не обязательно принадлежит к группе
- Группа может быть образована узлами разных подсетей

4Динамические группы

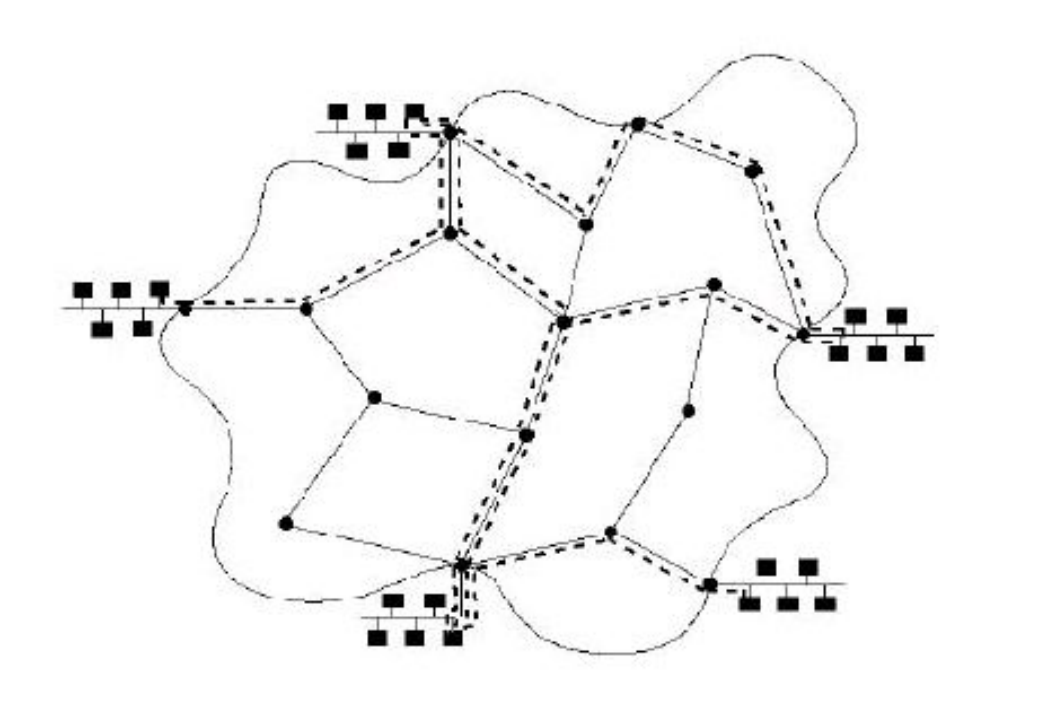
- Узлы могут присоединяться или покидать группы без регистрации или переговоров с централизованным элементом управления группового вещания

Магистраль MBONE

Появилась в 1992 году

Объединяла 20 сетей с помощью туннелирования группового трафика – *mrouted*

В конечной точке туннеля пакет распространялся широковещательно



Формат группового адреса (multicast)

1	1	1	0	Идентификатор группы (ID) – 28 бит

Адреса класса D

224.0.0.0 - 224.0.0.255	***Well known*** адреса
224.0.0.255 -238.255.255.255	групповые адреса, использующиеся в Internet
238.255.255.255 - 239.255.255.255	групповые адреса, использующиеся в автономных локальных сетях

Постоянные группы *(permanent host groups)* имеют зарезервированные (*well known)* идентификаторы:

- 224.0.0.1 – все узлы данной сети
- 224.0.0.2 - группа, в которую входят все маршрутизаторы в сети,
- 224.0.1.7 - группа получателей аудио-новостей.
- 224.0.1.9 – все маршрутизаторы RIPv2
- … - определены в RFC1700 – Assigned Numbers

Область действия группового адреса (scope) ограничивается TTL

Например, TTL=3 - адрес может пересечь 2 маршрутизатора

Групповой адрес определяет не узел, а сервис – аналог порта UDP/TCP

Отображение

групповых адресов IP на групповые адреса Ethernet

От **00:00:5e:00:00:00** до **00:00:5e:ff:ff:ff**

блок MAC-адресов, закрепленных за IANA.

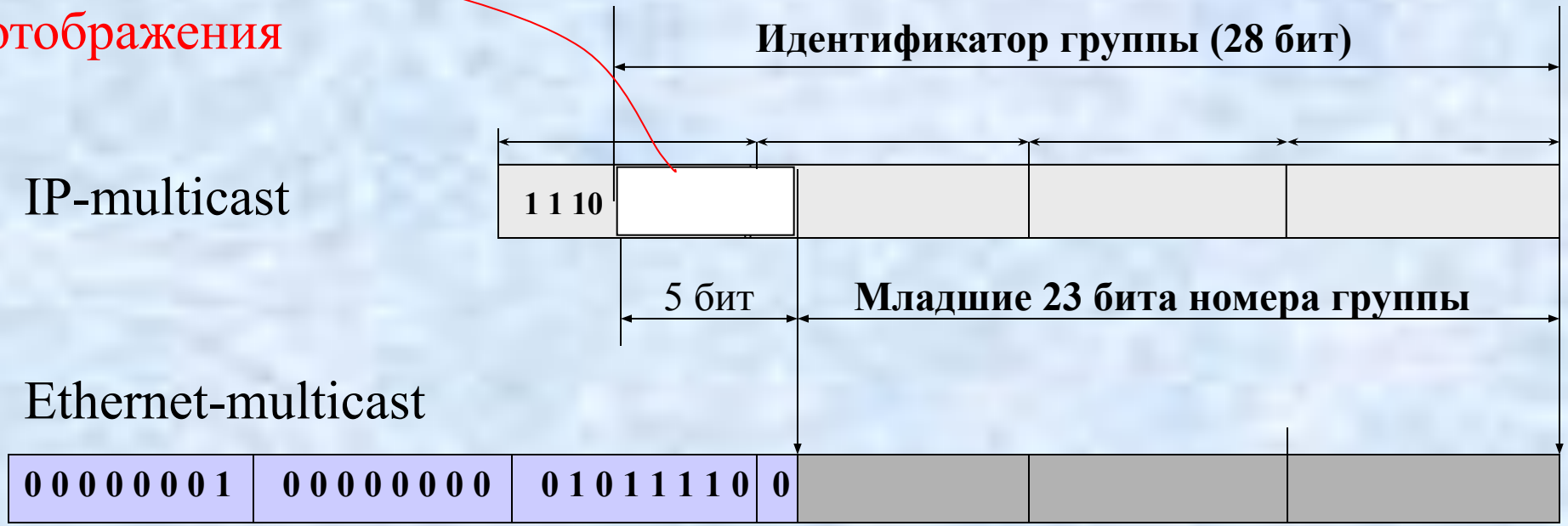
От **00:00:5e:00:00:00** до **00:00:5e:7f:ff:ff** -

половина из них отведена для групповой адресации.

Поскольку **01**- признак группового MAC- адреса, имеем следующий диапазон для отображения групповых IP-адресов:

от **01:00:5e:00:00:00** до **01:00:5e:7f:ff:ff**

Неоднозначность отображения



01 5e

224.128.64.32 (11100000 10000000 01000000 00100000) и
224.0.64.32 (11100000 00000000 01000000 00100000)

отображаются на один и тот же MAC-адрес **(Какой?)**

01:00:5e:00:40:20

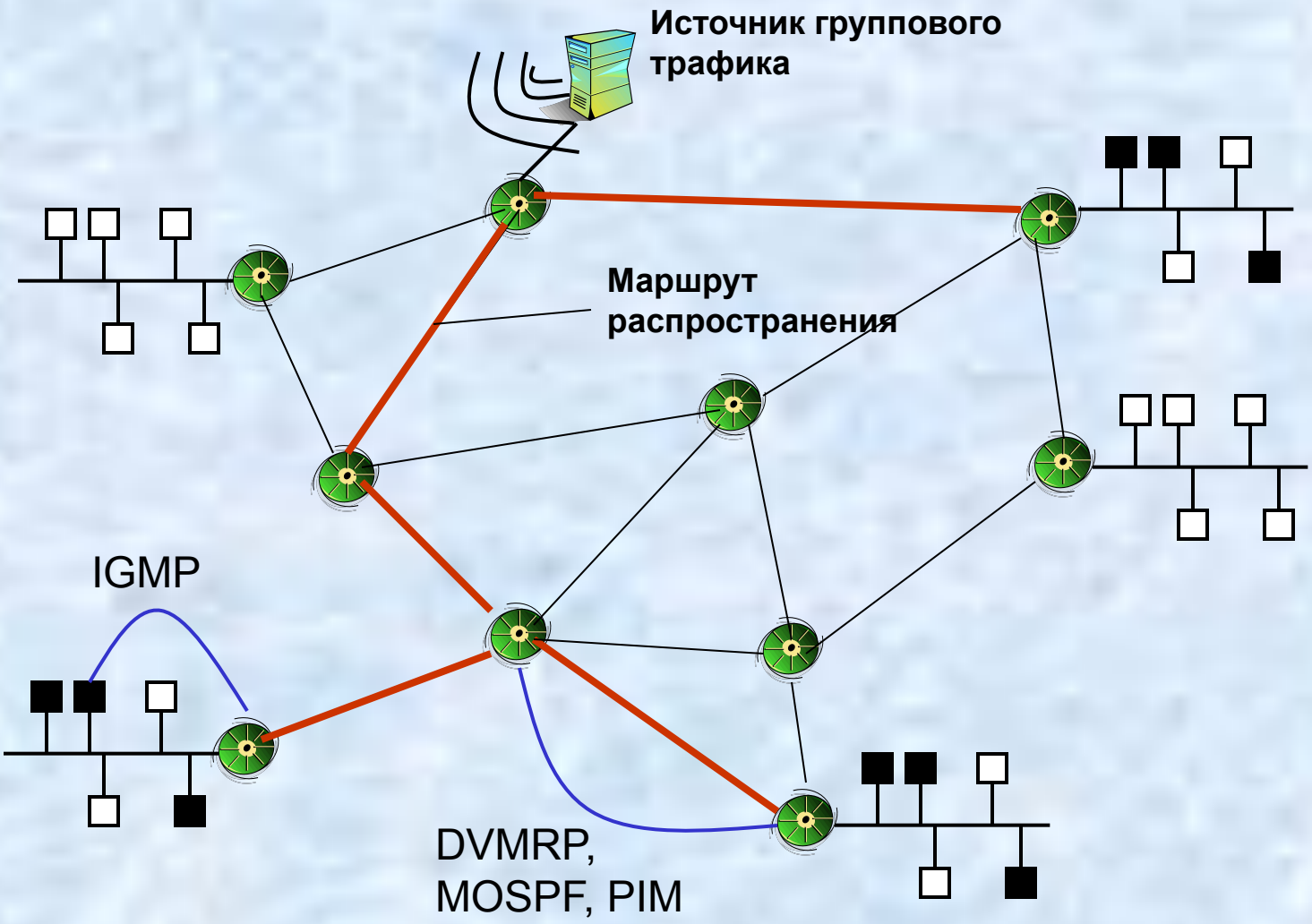
Решение проблемы неоднозначности:

1. Сетевой интерфейс может принять не предназначенный ему кадр

2. Модуль IP проверяет адрес, указанный в заголовке IP, сравнивая его с собственным групповым адресом

3. Даже при такой избыточности групповая адресация намного эффективнее широковещания.

Протоколы группового вещания



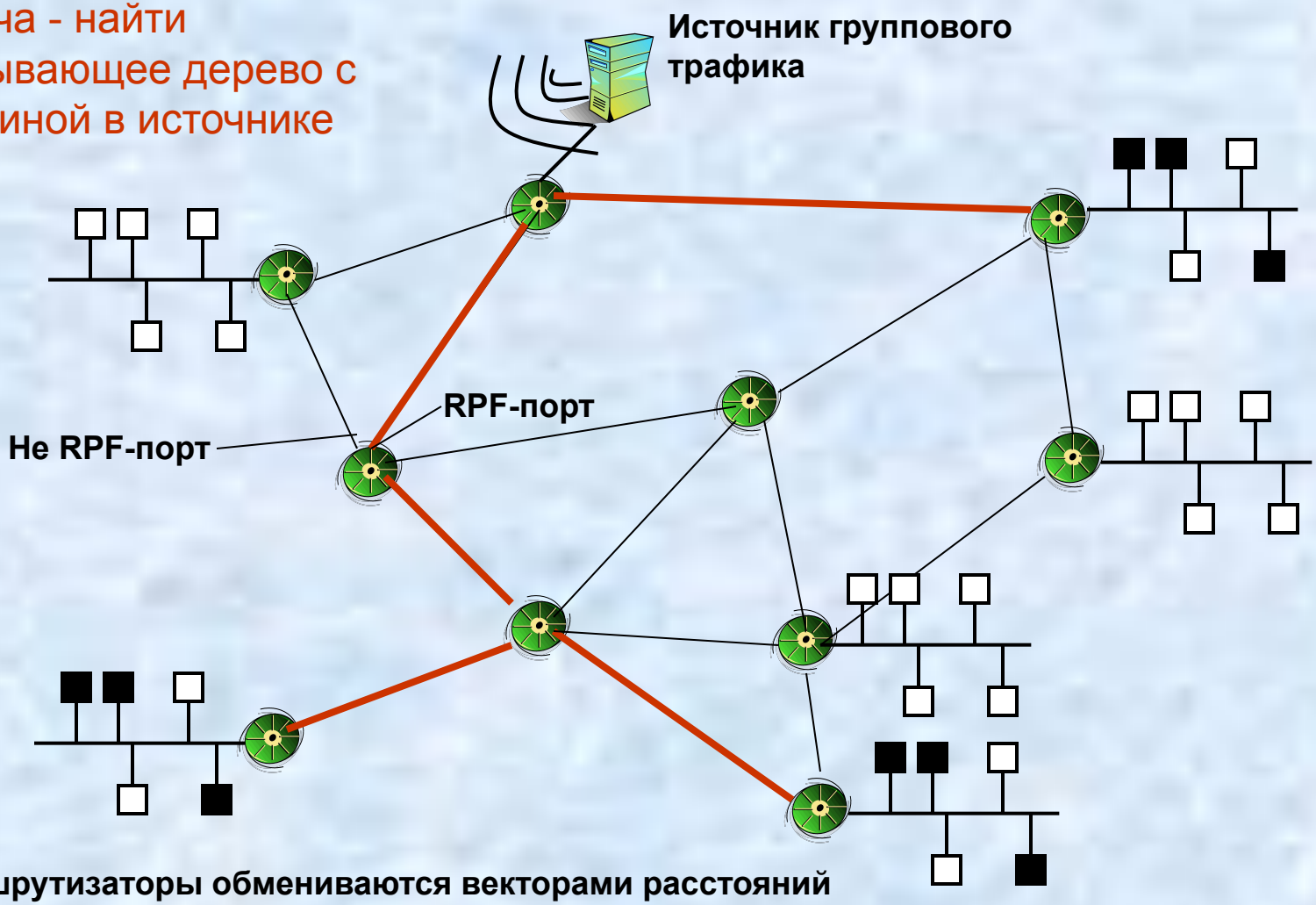
Internet Group Management Protocol (IGMP)

Позволяет маршрутизатору выяснить, есть ли в его подсетях члены групп:

- Маршрутизатор периодически рассылает сообщения Host Membership Query (HMQ) с нулевым адресом группы
- Узлы отвечают сообщениями Host Membership Replay (HMR), указывая групповой адрес, по которому они хотят получать пакеты

Distance-Vector Multicast Routing Protocol DVMRP

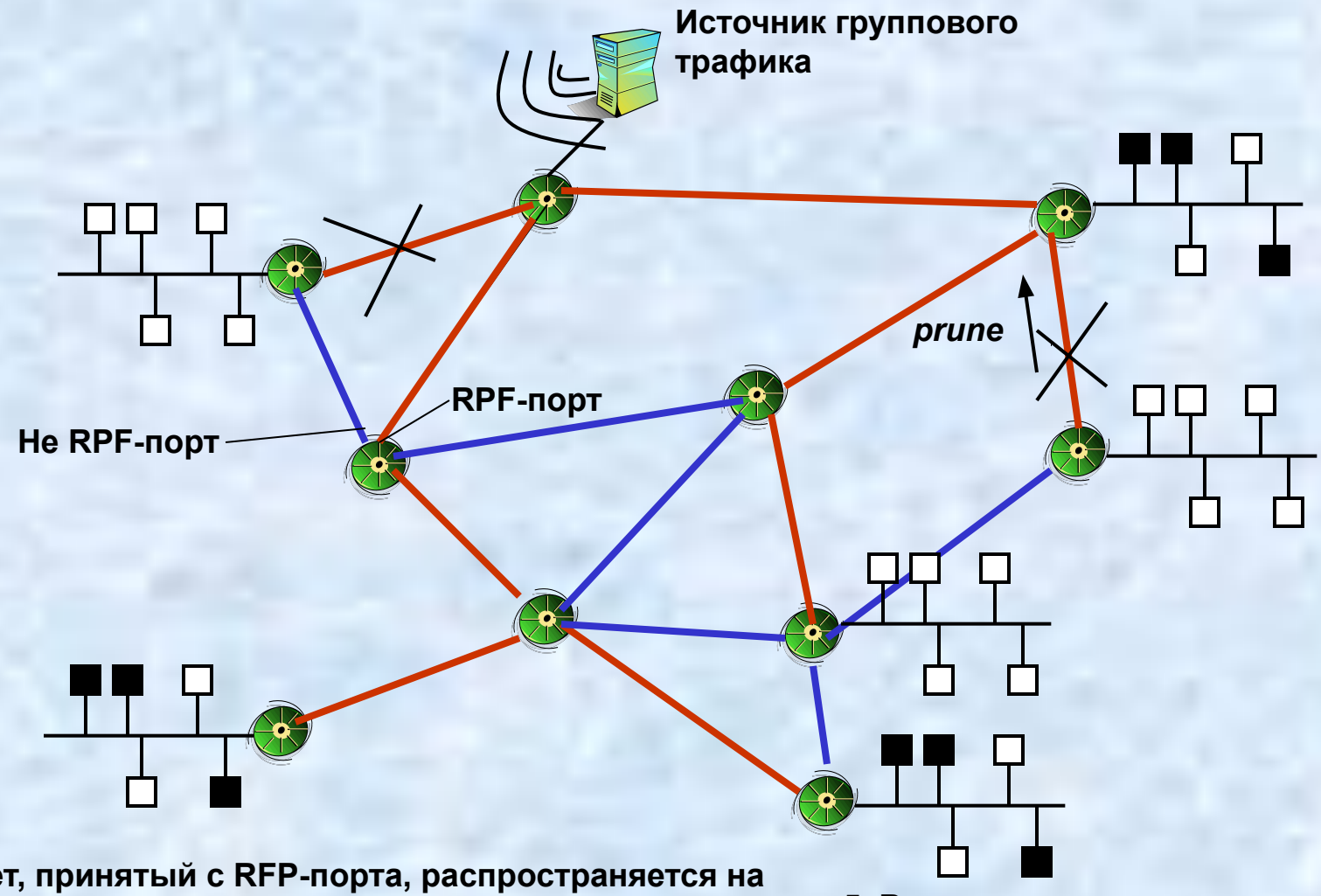
Задача - найти покрывающее дерево с вершиной в источнике



1. Маршрутизаторы обмениваются векторами расстояний до сетей, как и в RIP

2. При приеме пакета с multicast-адресом определяется получен ли он с «корневого» порта - Reverse Path Forwarding Port (RPF) - если нет, то пакет отбрасывается - исключаются петли

Distance-Vector Multicast Routing Protocol DVMRP



3. Пакет, принятый с RFP-порта, распространяется на все порты - затопление сети

4. Тупиковый маршрутизатор, не имеющий в локальной сети членов группы, посылает через RFP-порт сообщение *prune - усечение* дерева

5. Все маршрутизаторы посылают сообщение prune на не RFP-порты

В сети устанавливается покрывающее дерево

Недостатки DVMRP

- Эффективно работает в «плотно» населенных multicast-группами сетях (dense mode)
- Маршрутизатор должен хранить большое количество информации - пропорционально количеству источников - для каждого источника свое дерево!
- В разряженных сетях (sparse mode) протокол DVMRP порождает много избыточного трафика:

  - за счет затопления сети на этапе существования неусеченного дерева

  - за счет служебного трафика DV, требуемого для построения таблицы маршрутизации (протокол работает назависимо от протоколов обычной маршрутизации)

Протокол MOSPF

Multicast Extensions to OSPF, RFC 1584

1. Каждый маршрутизатор выясняет по IGMP членство в группах в непосредственно присоединенных сетях

2. Эта информация рассылается в специальных объявлениях Group Membership - состояние членства

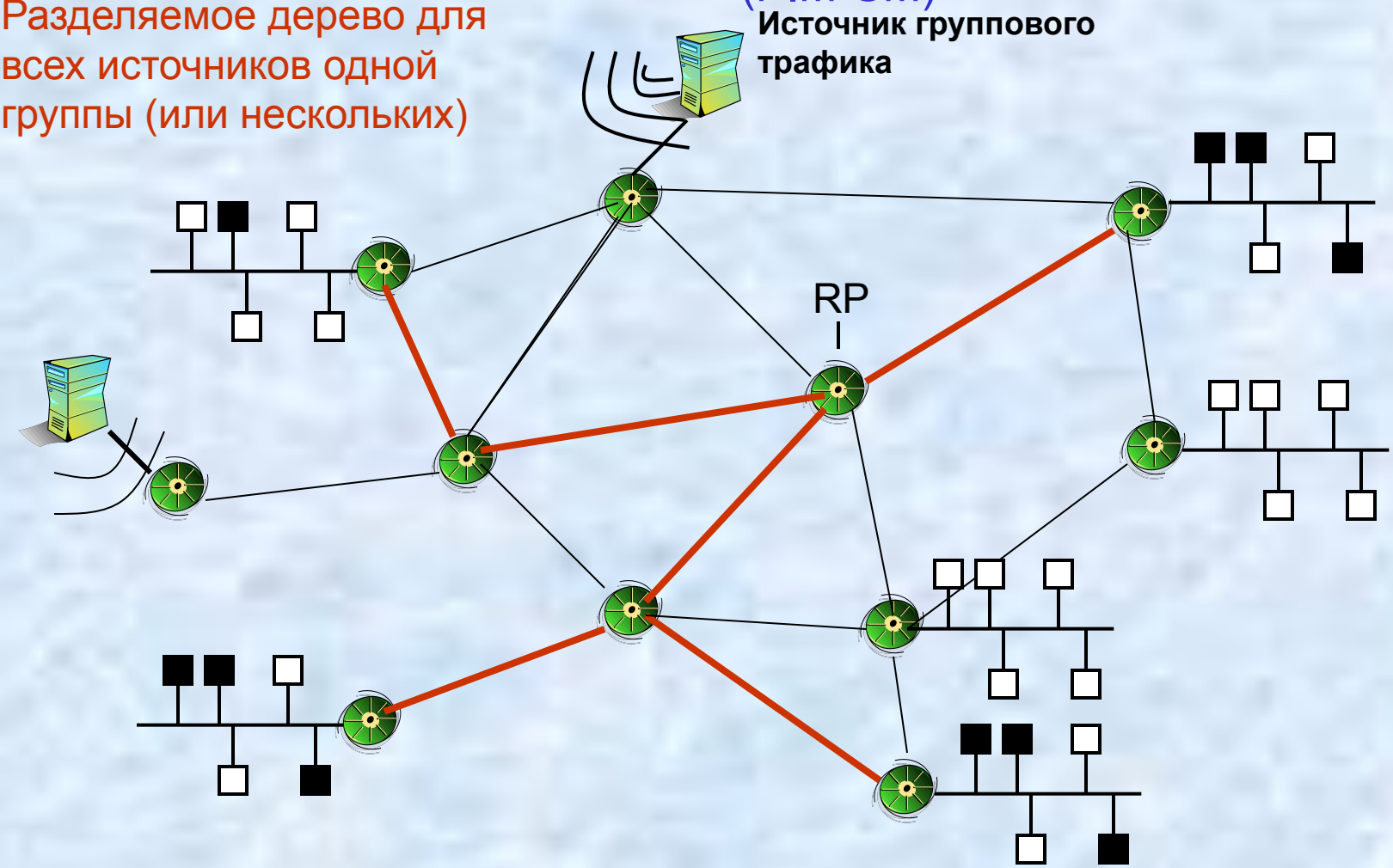
3. Каждый маршрутизатор находит дерево кратчайших маршрутов от себя до членов каждой группы – тот же недостаток, что и у DVMPR: для каждой группы своя таблица маршрутизации

4. При изменении членства в группе маршрутизатор сразу же делает объявление о новом состоянии членства

Относится к классу Sparse Mode протоколов

Protocol Independent Multicast - Sparse Mode (PIM-SM)

Разделяемое дерево для всех источников одной группы (или нескольких)



Для определенной группы или нескольких групп назначается Rendezvous Point - связной маршрутизатор

Построение разделяемого дерева по PIM-SM

1. С помощью специального протокола между маршрутизаторами распространяется информация о том, по какому уникальному IP-адресу находится RP, обслуживающий данный групповой адрес

2. Каждый маршрутизатор-получатель (у которого есть подсети с членами группы) посылает по адресу RP сообщение *joint* о присоединении к разделяемому дереву

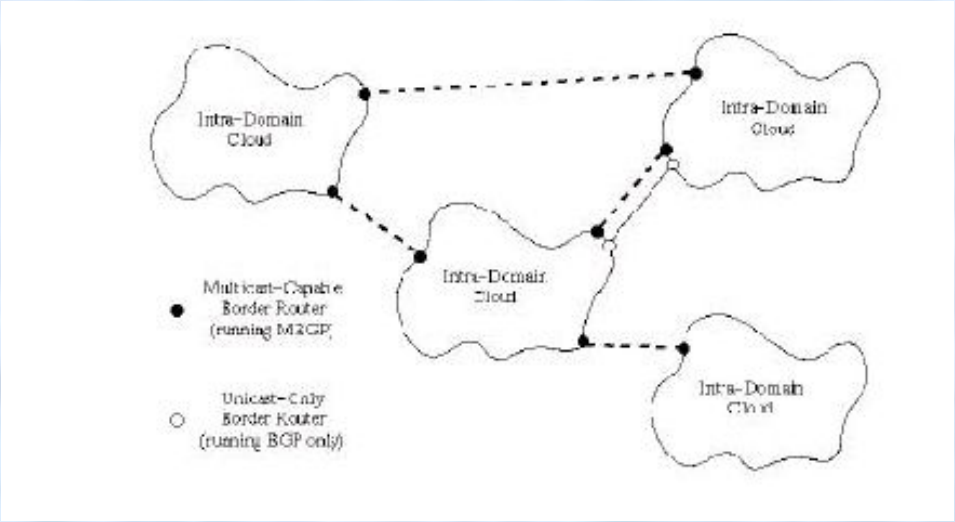
3. Вдоль пути *joint* маршрутизаторы фиксируют состояние продвижения группового адреса

4. Любой источник для группы посылает RP по уникальному адресу первые пакеты, инкапсулированные в сообщение регистрации

5. RP посылает источнику сообщение *joint*, после чего источник начинает посылать пакеты по групповому адресу - его присоединили к дереву

Развитие междоменного группового вещания

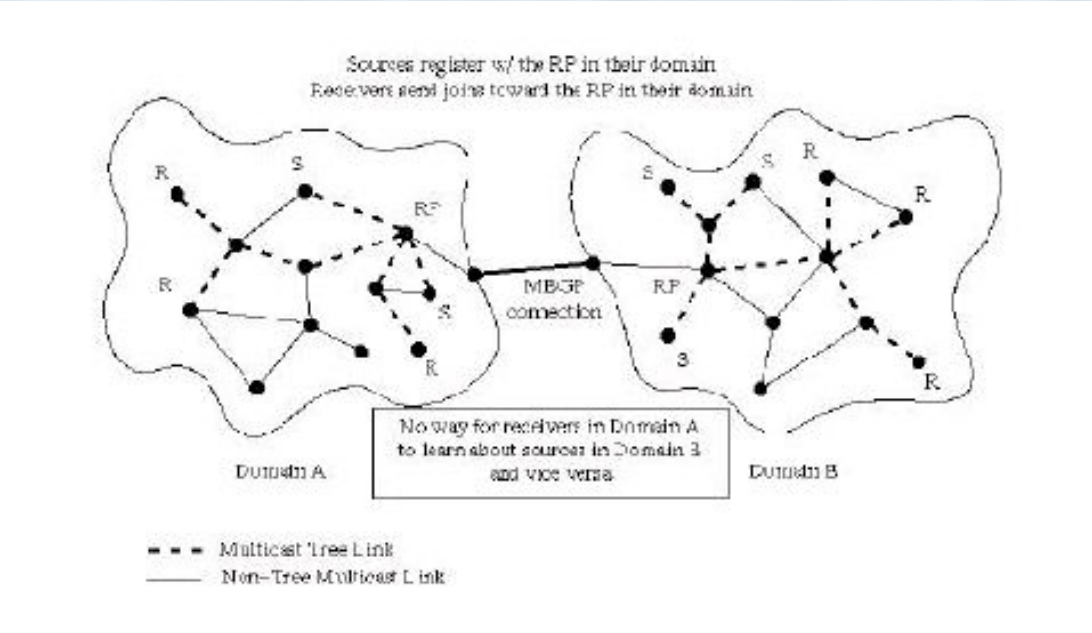
1. MBGP – BGP Multiprotocol Extensions



Сам по себе MBGP недостаточен – он только объявляет о наличии в AS получателей группового адреса – нужен протокол построения дерева вещания

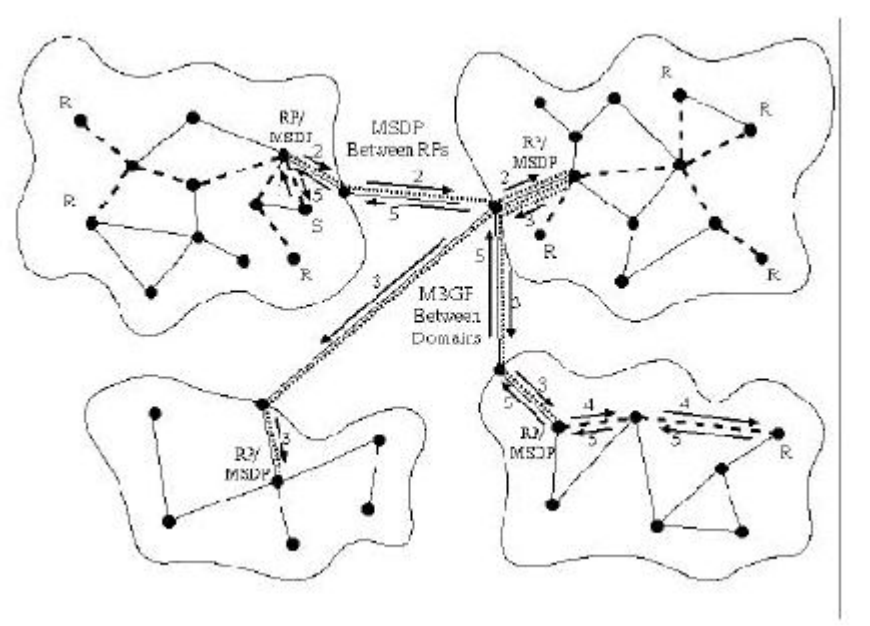
Проблемы MBGP:

1. В какой AS размещать точку встречи?
2. Как уведомить получателей одной AS о существовании источника в другой AS? – MBGP не переносит эту информацию



MBGP узнает о наличии в домене получателей определенной группы по протоколу PIM-SM

Новый дополнительный протокол – Multicast Source Discovery Protocol, MSDP + PIM SP внутри AS



1. Когда новый источник для группы становится активным, он регистрируется в точке RP домена по PIM.

2. Протокол MSDP, работающий в данном домене, обнаруживает возникновение нового активного источника и посылает сообщение «Источник активен» (Source Active, SA) всем узлам MSDP, с которыми у него сконфигурировано одноранговое TCP-соединение.

3. Сообщение SA затем периодически повторяется

4. MSDP-партнеры, которые получают сообщение SA, выполняют одноранговую RPF-проверку – не зациклилось ли SA – на основе данных MBGP

5. Внутри домена MSDP-партнер, который является точкой RP, проверяет, есть ли у него данные о членах группы в домене.

6. Если такие данные есть, то RP посылает сообщение протокола PIM о присоединении по адресу источника, объявленному в сообщении SA

7. Сообщение о присоединении проходит обратный путь до источника, используя междоменные пути, определяемые с помощью протокола MBGP, и фиксируя в маршрутизаторах состояние продвижения для данного группового адреса

. 8. Как только обратный путь продвижения установлен, точки RP начинают продвигать данные, в том числе и между доменами.

Недостатки подхода MBGP/PIM/MSDP

- Громоздкое решение – 3 протокола
- Возникают задержки присоединения, поскольку сообщения SA распространяеются периодически
- Большой объем служебного трафика между доменами – MSDP работает по принципу начального затопления, как DVMPR, только между доменами

Долговременные решения междоменного группового вещания

Потокол BGMP (Border Gateway Multicast Protocol) – более перспективное решение

- Строит дерево с одной RP для нескольких доменов
- Решение о том, какой домен будет содержать точку RP для определенных групп принимается динамически
- Использует схему динамического распределения групповых адресов между доменами MASC (Multicast Address-Set Claim (аналог DHCP) – видна принадлежность получателя AS – основа для взаиморасчетов между операторами