

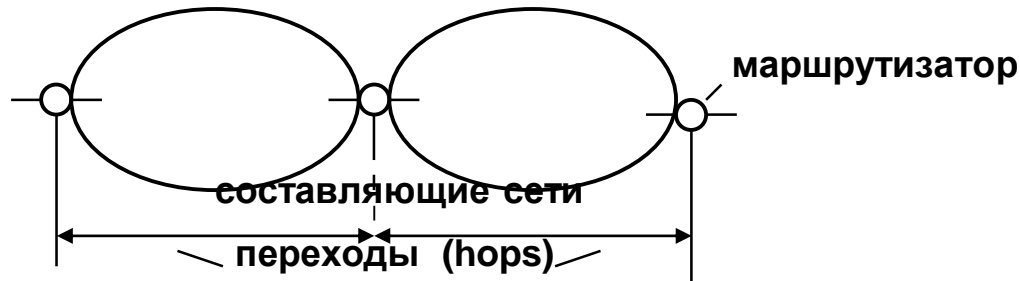
4. Принципы маршрутизации в сети Internet

4.1 Базовые установки

Понятие маршрутизации

Выбор маршрута пересылки IP-дейтаграммы между оконечными системами объединенной сети, через промежуточные составляющие сети. На пути вдоль маршрута дейтаграмма пересылается его узлами – оконечными системами или маршрутизаторами. Участок трассы между смежными узлами называется «переходом, термин на англ. - hop».

Независимо от типа и структуры составляющей сети, на которой базируется переход, она принимается за отрезок среды передачи данных, логически не участвующий в процессе маршрутизации.

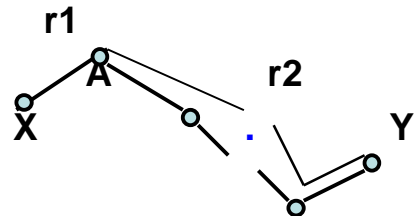


Оптимальные маршруты (маршруты минимальной стоимости)

Решение о маршруте принимается на основании критерия минимальной стоимости. Стоимость маршрута или его отрезка измеряется определенным показателем или комбинацией показателей.

Типичный показатель – число переходов в трассе. Их минимизация одновременно уменьшает задержку и освобождает сетевые ресурсы, т.е. увеличивает пропускную способность сети.

Принцип оптимальности



Если узел A находится на оптимальном маршруте от X к Y, то оптимальный маршрут от A к Y является его частью.

Доказательство: если бы существовал более оптимальный маршрут от A к Y, то добавив его к r1, получили бы более оптимальный маршрут от X к Y.

Следствие 1. Независимо от топологии сети, можно считать, что узел является корнем дерева оптимальных маршрутов, все ветви и части ветвей которого, направленные от корня, являются оптимальными маршрутами.

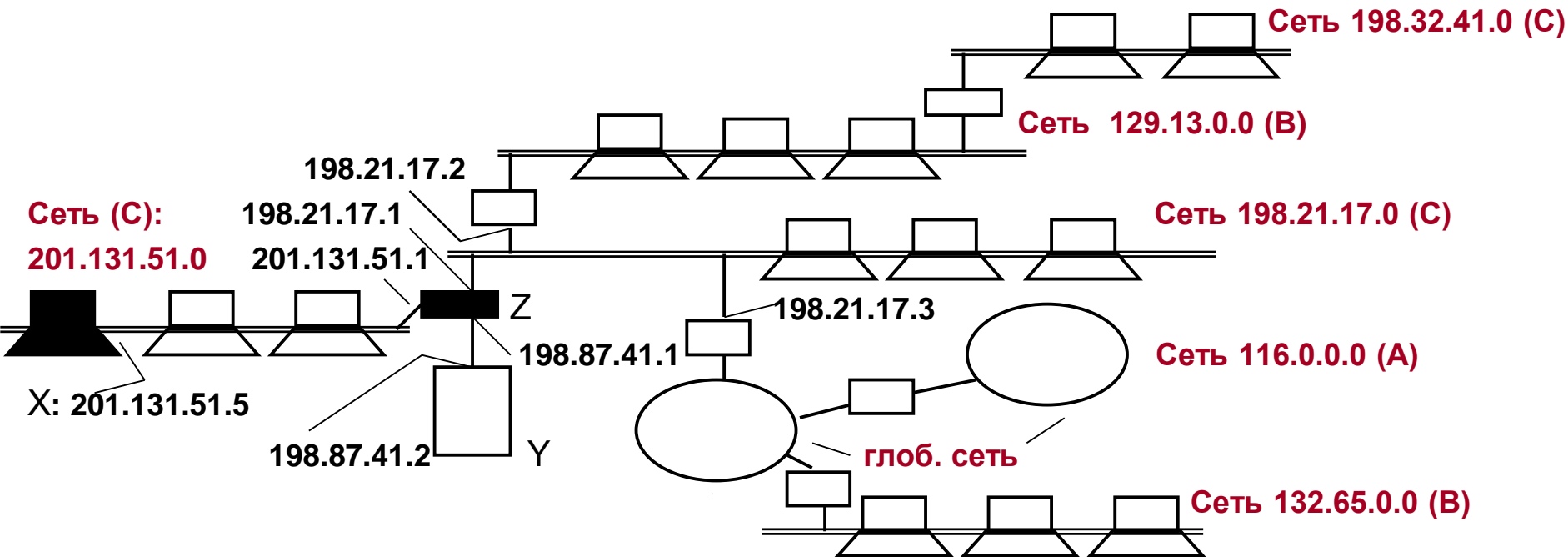
Следствие 2. Если алгоритм маршрутизации предполагает использование таблиц оптимальных маршрутов, то в каждом узле достаточно хранить только ее правый столбец (часть маршрута до смежного узла).

4.2 Основы динамического построения маршрутных таблиц

Определяющая проблема – в процессе маршрутизации максимально снизить вычислительную нагрузку на узлы. Отсюда – принцип намеренного сокрытия информации – маршрутизация на основании минимально необходимых данных:

1. Маршрутизация «только на шаг вперед» - в таблицах пары значений (N,R) , где N – адрес доставки, R – адрес ближайшей точки перехода.
2. Маршрутизатор в ближайшей точке перехода задается своим интерфейсом, который имеет IP-адрес; сам маршрутизатор, как единое устройство, адреса не имеет.
3. Обозначение адреса доставки IP-дейтаграммы через адрес сети (полный адрес узла только в необходимых случаях).
4. Стандартный маршрут следования дейтаграмм (маршрут по умолчанию) – объединение нескольких элементов таблицы в один блок.

4.3 Маршрутные таблицы на основе классовой адресации



Упрощенные маршрутные таблицы

Адрес доставки	Адрес след. маршрутизатора	Выходной интерфейс	Расстояние
X: 201.131.51.0	—	201.131.51.5	0 (подключен)
Специф. (узел Y)			
198.87.41.2	201.131.51.1	201.131.51.5	1
Станд.	201.131.51.1	201.131.51.5	—
Z: 201.131.51.0	—	201.131.51.1	0 (подключен)
198.21.17.0	—	198.21.17.1	0 (подключен)
129.13.0.0	198.21.17.2	198.21.17.1	1
198.32.41.0	198.21.17.2	198.21.17.1	2
Специф. (узел Y)			
198.87.41.2	—	198.87.41.1	0 (подключен)
Станд.	198.21.17.3	198.21.17.1	—

Выбор адреса доставки IP-дейтаграмм, принятых маршрутизатором (классовая адресация)

1. Извлечь из дейтаграммы IP-адрес назначения D и определить адрес доставки по номеру сети N.
2. Если N совпадет с номером одной из сетей, к которым подключен узел, то выполнить прямую доставку дейтаграммы по D (определить аппаратный адрес D в базовой сети, инкапсулировать IP-дейтаграмму в сетевой (аппаратный) формат и отправить по назначению).
3. Если в таблице указан специфический адрес D, то переслать IP-дейтаграмму по заданному для него адресу следующего маршрутизатора.
4. Если в таблице указан адрес доставки с номером сети N, то переслать IP-дейтаграмму по заданному адресу следующего маршрутизатора.
5. Если в таблице указан стандартный маршрут, то переслать IP-дейтаграмму по адресу заданного для него следующего маршрутизатора.

4.4 Транзит IP-дейтаграмм при бесклассовой адресации CIDR

Предпосылки

По замыслу Internet возможность доступа к сети в типичном варианте предоставляется поставщиками услуг (Internet Service Providers, ISPs), которые обеспечивают подключение через их точки доступа (Points of Presence, POPs) и назначение IP-адресов.

Поставщик (провайдер) получает для распределения непрерывный блок классовых IP-адресов от провайдера более высокого уровня.

Основные принципы распределения:

экономия адресного пространства,

уменьшение числа записей в таблице маршрутизации,

структуризация сетей пользователей на подсети для локализации трафика, отдельной диагностики и реализации политики безопасности.

Типичный ресурс провайдеров – блоки адресов классов B и C.

Класс B, как таковой, не способствует экономии и структуризации (2^{16} узлов),

класс C предоставляет ограниченное число (2^8) узлов, требующее дробить сети, в которых число узлов > 256 , и строить таблицы с большим числом записей.

Отсюда необходимость гибкой системы адресации с возможностью выбора длины сетевой части адреса. Использование масок.

Принцип распределения пула адресов на основе бесклассовой адресации

Пул адресов для сети, формируемой на основе **бесклассовой** адресации, выделяется провайдером из базового адресного пространства, описываемого методом **классовой** адресации.

Пул адресов выделяется с использованием адресной маски, которая определяет префикс формируемой сети – ее номер в составе Internet. Таким образом, устанавливается соответствие между сетью и маской. Условно говоря, «каждая сеть, заданная методом бесклассовой адресации, «находится под своей маской»

Правые разряды базового адресного пространства, не находящиеся под маской, определяют номера узлов в формируемой сети. Если номера узлов задаются n разрядами, то первый номер узла в пуле должен быть равен n нулям, а остальные n -разрядные номера - увеличиваться в непрерывной последовательности.

Фактор экономии адресного пространства

Если в составе пула адресов, выделенных для **первой сети, находящейся под данной маской**, есть незанятый блок последовательных адресов, его по тому же принципу можно использовать для **второй сети, находящейся под другой маской**. Маска второй сети образуется из маски первой сети путем ее удлинения справа на число разрядов, оставляющее возможность перекрыть все номера узлов в незанятом блоке. Тем же образом незанятый блок адресов во второй сети можно использовать для третьей сети и так далее.

Выбор сети назначения в маршрутизаторе

Если в Internet введен маршрутизатор, который обеспечивает подключение всех сетей, сформированных таким образом из единого базового адресного пространства, то возможна ситуация, когда поступившая извне IP-дейтаграмма будет иметь адрес назначения, показывающий принадлежность к нескольким сетям с бесклассовыми адресами, находящимися под разными масками.

В этом случае маршрутизатор направляет IP-дейтаграмму в ту из них, которая находится под самой длинной маской.

Пример распределения пула адресов на основе бесклассовой адресации

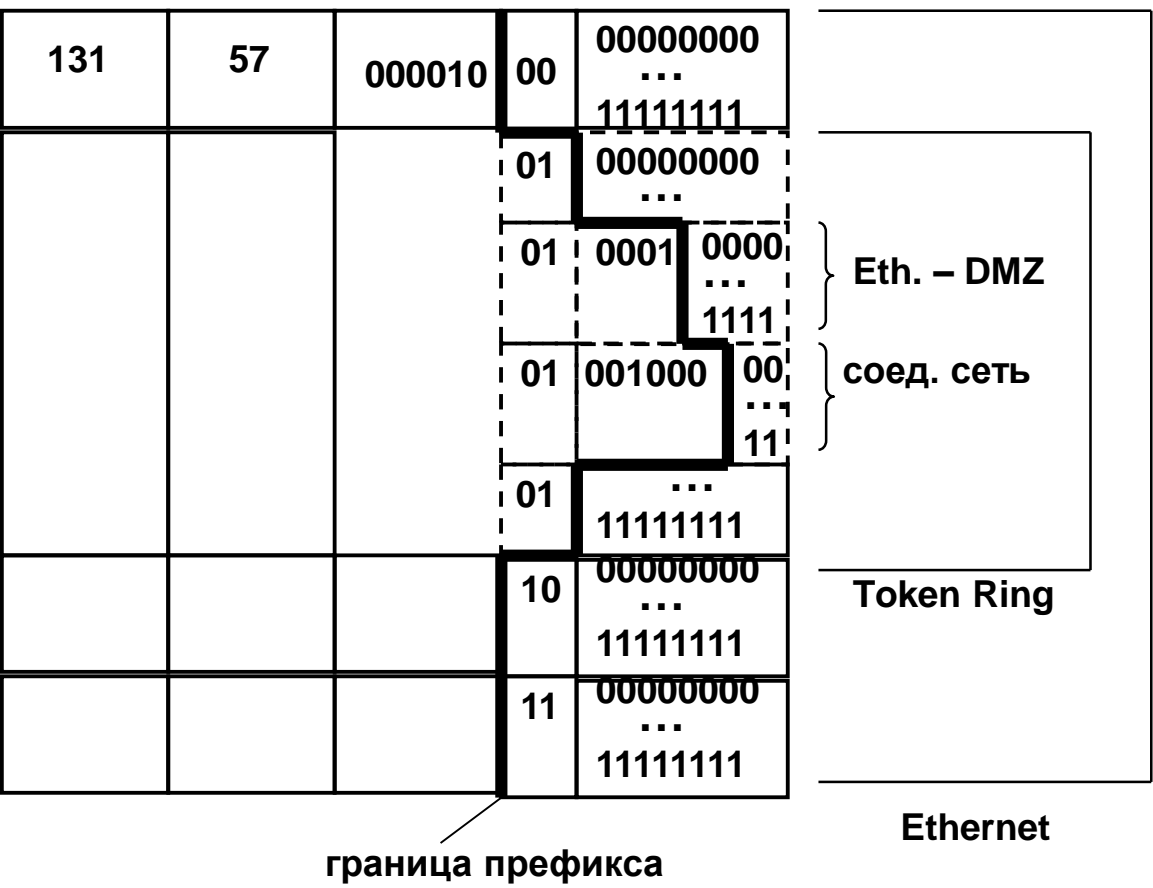
Провайдер располагает адресным пространством класса В, в которое надо ввести пул адресов для сети S, включающей:

сеть Ethernet - 0 – 600 станций (офисная),

сеть Token Ring – 200 станций (офисная)

сеть Ethernet – DMZ – 10 станций (открытый доступ, Demilitarized Zone).

В адресном пространстве провайдера есть пул на 1024 адреса.



Пример сетевой реализации пула бесклассовых адресов

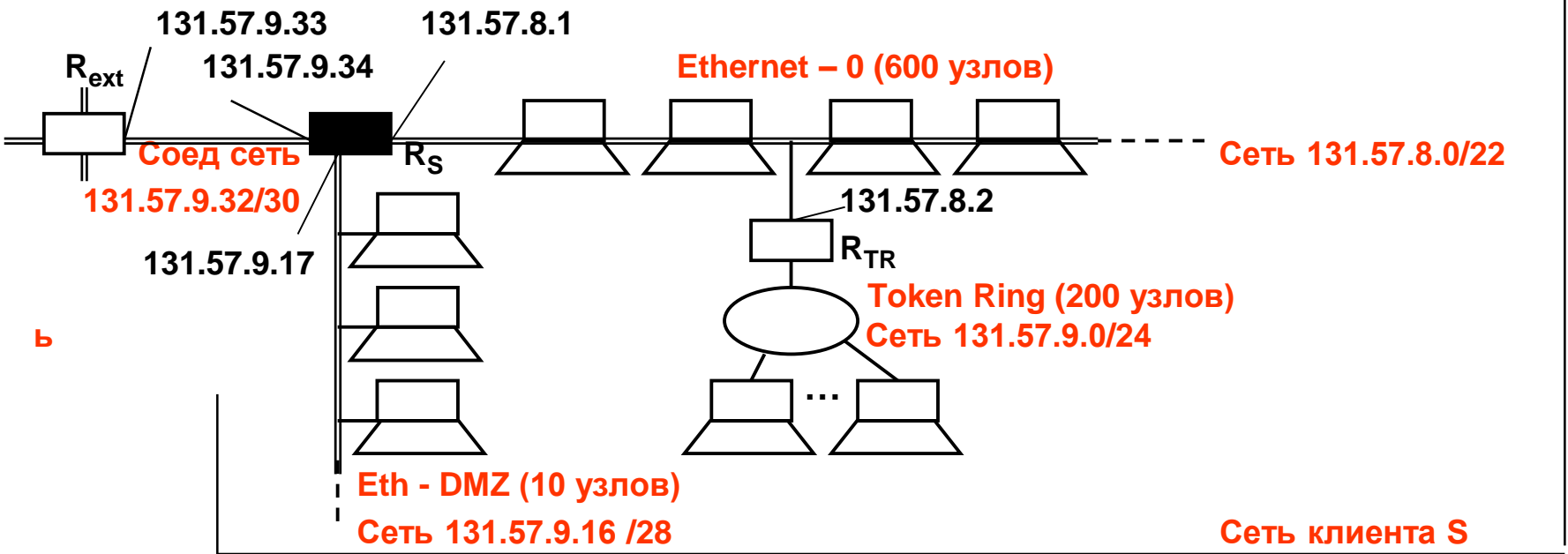


Таблица маршрутизации R_{ext}

Адрес доставки	Маска	Адрес след. маршрутизатора	Вых. интерфейс	Расстояние
131.57.8.0	255.255.252.0.	131.57.9.34	131.57.9.33	1

.....

Таблица маршрутизации R_{1s}

Адрес доставки	Маска	Адрес след. маршрутизатора	Вых. интерфейс	Расстояние
131. 57.8.0 (Eth – 0)	255.255.252.0	—	131.57.8.1	0
131.57.9.0 (TR)	255.255.255.0	131. 57.8.2	131.57.8.1	1
131.57.9.16 (Eth-DMZ)	255.255.255.240	—	131.57.9.17	0
Стандарт (Internet)		131.57.9.33	131.57.9.34	

Выбор адреса доставки IP-дейтаграмм, принятых маршрутизатором (бесклассовая адресация)

Извлечь из дейтаграммы IP-адрес назначения D.

1. Если в таблице указан специфический адрес D то переслать дейтаграмму по адресу следующего маршрутизатора.
2. Просмотр таблицы, при котором с каждой строкой выполняются следующие действия:
 - маска логически умножается на IP-адрес назначения,
 - полученное число сравнивается с адресом доставки,
 - если совпадение, то эта строка отмечается.
3. Просмотр закончен:
 - если не было совпадений, то на 4,
 - если есть одно совпадение, то дейтаграмма отправляется по этому маршруту,
 - если есть несколько совпадений (адрес принадлежит нескольким подсетям, то дейтаграмма отправляется по маршруту, в котором совпадает больше разрядов).
4. Если указан стандартный маршрут, то дейтаграмма отправляется по нему, иначе отбрасывается.

5. Методы и протоколы построения маршрутов пересылки IP-дейтаграмм

5.1 Основные классы

Статические методы

Маршруты строятся в стороне от работы сети. Сформированные (возможно администратором) маршрутные таблицы загружаются в узлы и остаются неизменными до следующей загрузки.

Динамические методы

Маршрут строится в процессе работы сети.

Лавинная маршрутизация

Дейтаграмма пересылается отправителем во все соседние узлы кроме того, от которого она принята. Таким образом имеет место многократная доставка.

Проблема: ограничить и уничтожить этот поток.

Решение: значение счетчика переходов, установленного на максимальное число переходов, уменьшается на 1 в каждом промежуточном узле. Есть усовершенствованные методы.

Свойства:

- использование всех возможных маршрутов (живучесть – военные применения);
- по крайней мере одна копия пройдет по маршруту с минимальной стоимостью (можно внести его в таблицу).

Адаптивная маршрутизация

Маршруты устанавливаются в процессе работы сети и соответствуют ее текущему состоянию (возможно введение новых сетевых объектов и прекращение активности действующих).

Проблемы:

- решение более сложно, чем при лавинной технологии, увеличивается объем обработки в узлах;
- информация о состоянии сети вырабатывается в одном месте, но используется в другом – служебные пересылки: чем они чаще, тем точнее выбор – но больше нагрузка на сеть;
- быстрая перестройка состояний может вызвать колебания, при медленной нет адаптации.

Однако:

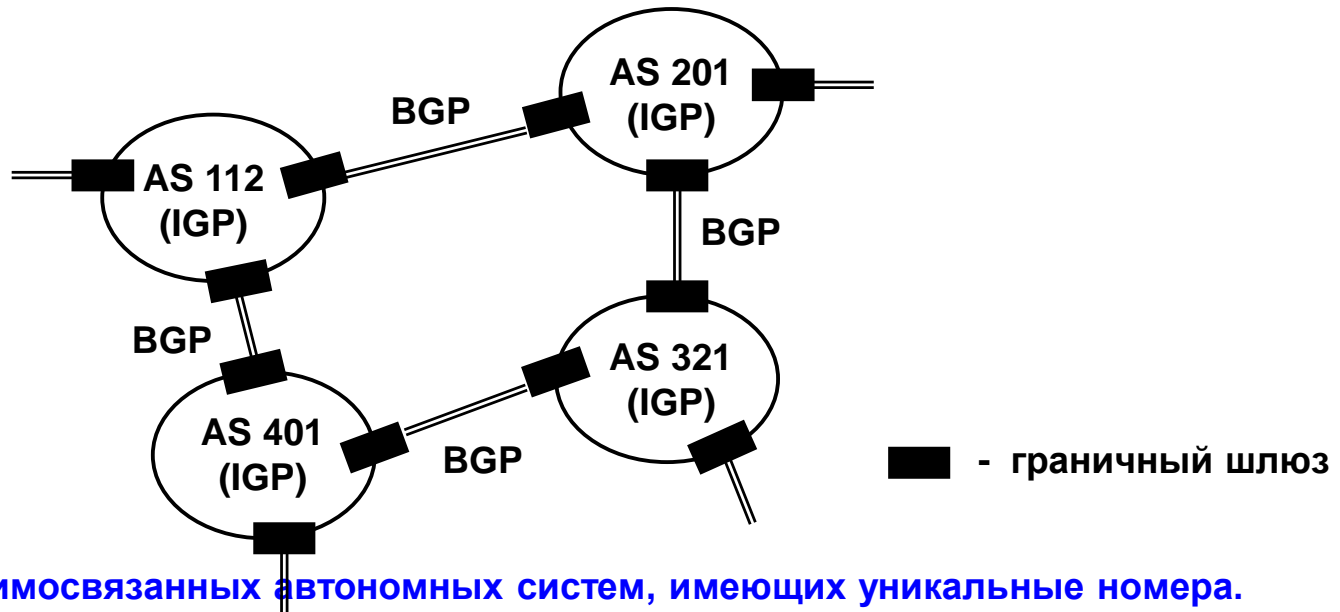
- заметно для пользователя повышается быстродействие сети;
- облегчается управление перегрузкой.

5.2 Концепция автономных систем

Проблема масштабируемости: возможность наращивать количество узлов и протяженность связей в широких пределах без существенного усложнения и снижения показателей сетей.

Решение проблемы: введение автономных систем (autonomous systems, ASs).

Автономная система – совокупность сетей под единым административным управлением, обеспечивающая общую для нее политику маршрутизации. Обычно AS управляется одним поставщиком услуг Internet.



Internet – набор взаимосвязанных автономных систем, имеющих уникальные номера.

Автономные системы соединяются внешними шлюзами и образуют магистраль Internet.

Маршрут в Internet определяется через:

последовательность AS – протоколом граничного шлюза (Border Gateway Protocol, BGP),

последовательность сетей в пределах AS – протоколами внутреннего шлюза (Internal Gateway Protocols, IGP)

5.3 Маршрутизация внутри автономных систем

5.3.1 Протоколы на основе дистанционно-векторного алгоритма (DVA)

5.3.1.1 Построение маршрутов

Общий принцип

В объединенной сети, входящей в состав автономной системы, содержится N сетей.

В каждом узле i (маршрутизаторе или конечной системе) хранится вектор минимальной стоимости маршрутов:

$$D_i = \begin{pmatrix} d_{i1} \\ d_{i2} \\ \dots \\ d_{iN} \end{pmatrix}, \text{ где } d_{ij} - \text{текущая оценка стоимости маршрута до сети } j.$$

Выбор параметра стоимости среди параметров, определенных протоколом, в общем случае передается на усмотрение администратора.

Каждый узел периодически (в протоколе RIP – через 30 с) посылает сообщение со своим вектором смежным узлам и получает вектор от каждого из них. В начале процесса вектор D_i содержит только элементы, соответствующие смежным узлам, которые находятся в сетях, непосредственно связанных с узлом i .

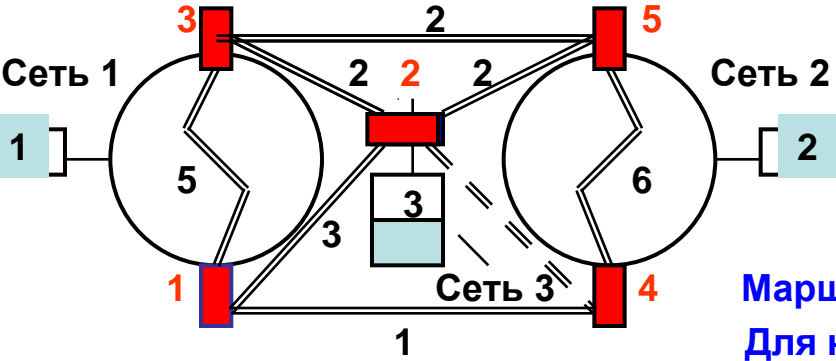
При получении вектора D_k от смежного узла k узел i корректирует свой вектор D_i в следующих случаях:

1. D_k содержит номер сети, которого нет в D_i .
2. Для каждой сети j , которая обозначена в обоих векторах (D_i и D_k) вычисляется

$$d_{ij \text{ нов}} = d_{kj} + l_{ik}, \quad \text{где } l_{ik} - \text{стоимость звена между } i \text{ и } k.$$

Если $d_{ij \text{ нов}} < d_{ij \text{ тек}}$ (текущее значение в D_i), то строка сети j в D_i заменяется. Из нескольких $d_{ij \text{ нов}}$ выбирается наименьшее.

Пример построения маршрута. Сеть



Объединенная сеть включает:

Сеть 1, Сеть 2, Сеть 3.

Каждая сеть соединена с другой через маршрутиз. высокоскоростными каналами.

Маршрутизаторы 1,2,3,4,5. Оконечные системы 1,2,3.

Для каждого канала задается его стоимость.

Соединения каждой сети для надежности и повышения эффективности обменов дублируются.

Маршрутизатор 2 помимо подключения к сети 3 может выполнять обмен между сетями 1 и 2.

Связь маршрутизаторов 4 и 2 временно отключена.

Построение в маршрутизаторах 1,2,3,4,5

Этапы	сеть	D_i	маршрут	следующий	D_i	марш.	след.	D_i	марш.	след.	D_i	марш.	след.	D_i	марш.	след.
0	1	0	Yes	Yes	No	No	No	0	Yes	Yes	No	No	No	No	No	No
	2	N	N	N	N	N	N	N	N	N	0	Y	Y	0	Y	Y
	3	N	N	N	0	Y	Y	N	N	N	N	N	N	N	N	N
1	1	0	Y	Y	2	2-3	3	0	Y	Y	1	4-1	1	2	5-3	3
	2	1	1-4	4	2	2-5	5	2	3-5	5	0	Y	Y	0	Y	Y
	3	3	1-2	2	0	Y	Y	2	3-2	2	N	N	N	2	5-2	2
2	1	0	Y	Y	2	2-3	3	0	Y	Y	1	4-1	1	2	5-3	3
	2	1	1-4	4	2	2-5	5	2	3-5	5	0	Y	Y	0	Y	Y
	3	3	1-2	2	0	Y	Y	3	3-2	2	4	4-1-2	1	2	5-2	2
											8	4-5-2	5			

Построение в оконечных системах

○1: сеть 1 (Y); с2 (след. 1); с3 (сл. 3) ○2: с1 (сл. 4); с2(Y); с3 (сл. 5) ○3: с1 (сл. 2); с2 (сл.2); с3 (Y)

5.3.1.2 Протоколы RIP1 и RIP2 (Routing Information Protocols)

Краткая характеристика

Элементы вектора D_i :

RIP1 – IP-адрес сети, стоимость маршрута

Метрика: число переходов (хопы)

RIP2 – IP-адрес сети, маска, стоимость маршрута

Метрика: хопы с коэф., проп. способн., задержка

Адаптация маршрутизатора к изменениям в сети

Хорошая новость распространяется быстро, плохая – медленно.

Включение узла A

A	B	C	D
1	1	1	
N	N	N	(начало)
1	N	N	(1-й этап)
1	2	N	(2-й этап)
1	2	3	(3-й этап)

Отключение узла A

A	B	C	D
1	1	1	
1	2	3	(начало)
3	2	3	(1-й этап)
3	4	3	(2-й этап)
5	4	5	(3-й этап)
5	6	5	(4-й этап)
7	6	7	(5-й этап)
7	8	7	(6-й этап)

Разрыв связи при увеличении до 16

Решение: алгоритм расколотого горизонта

Узел C, установив через B стоимость маршрута до A ($C-A=2$),

- сообщает C-A узлу D,
- сообщает узлу B, что $C-A = 16$ (закрывает для B доступ к A через себя).

Аналогично D закрывает для C доступ к A через себя.

Если A отключается, то B полностью изолируется от A, } плохая новость распространяется
на след. этапе C изолируется от A, } со скоростью один переход за 30 с.
на след. этапе D изолируется от A.

Тайм-аут наличия сети в векторе D – при отсутствии в векторе в течение 180с сеть отключается.

5.3.1.3 Общая оценка протоколов с алгоритмом DVA

Метод прост в реализации – обмен только с соседями,

но:

большой объем сообщений, передающих вектор D, в котором данные о всех сетях, вычисления пути наименьшей стоимости распределены по сети, отсюда - меньшая достоверность, быстрое изменение маршрутов, медленное распространение - результат м. быть нестабильным, возможное усовершенствование алгоритма (расколотый горизонт) усложняет протокол и при определенных конфигурациях не работает.

5.3.2 Протоколы на основе алгоритмов маршрутизации по состоянию связей (Line State Advertisements, LSAs)

5.3.2.1 Алгоритм Дейкстры

В исходном узле содержится база данных, собранная от всех узлов сети.

По ней определяется кратчайшее расстояние до каждого узла сети.

Пусть:

N – множество узлов сети,

S – исходный узел,

T – множество узлов, уже обработанных алгоритмом,

$w(i,j)$ – стоимость маршрута от I к j, причем

$w(I,i)=0$, $w(I,j)=\infty$, если узлы не соединены непосредственно, $w(I,j) \geq 0$, если соединены.

$L(n)$ – стоимость маршрута минимальной стоимости от S к n.

1. Инициализация:

$T = \{S\}$ - сам узел,
 $L(n) := w(s,n)$ - для $n \neq s$, начальная стоимость маршрута к другому узлу сети.

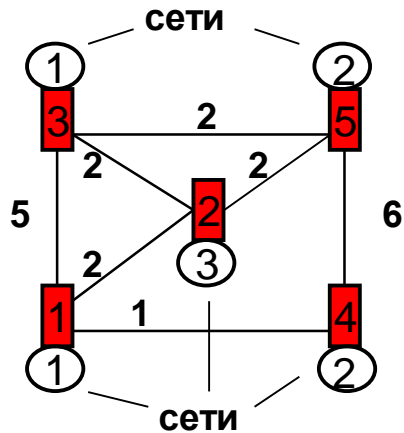
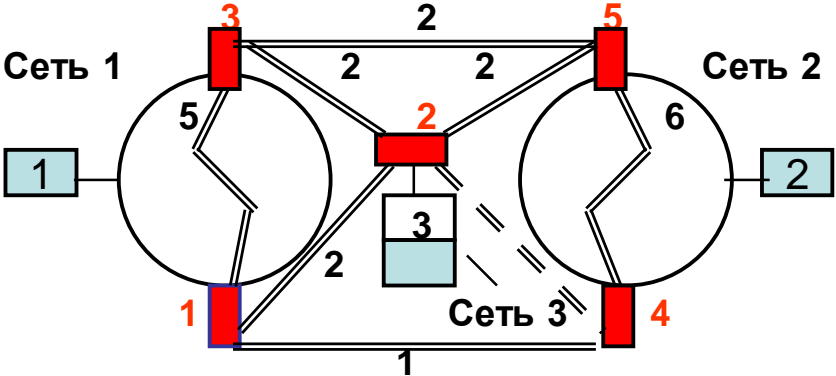
2. Добавление нового узла:

Найти узел x , не входящий в T , который имеет маршрут минимальной стоимости от узла S .
 Добавить узел x в T .

3. Обновление путей:

$L(n) := \min [L(n), L(x) + w(x,n)]$ для всех n , не принадлежащих T .

Этапы 2 и 3 повторяются, пока все узлы не войдут в T .



Маршрутизатор 1:

Итерация	T	L(2)	маршрут	L(3)	маршрут	L(4)	маршрут	L(5)	маршрут
0 (иниц.)	1	2	1-2	5	1-3	1	1-4	∞	-
1	1,4	2	1-2	5	1-3	1	1-4	7	1-4-5
2	1,2,4	2	1-2	4	1-2-3	1	1-4	4	1-2-5
3	1.2.3.4				без изменений				
4	1,2,3,4,5				без изменений				

сеть	следующий
1	Y
2	4
3	2

Применение алгоритма Дейкстры к БД показывает, что маршрутизатор 1, подключенный к сети 1, имеет оптимальный путь до сети 2 через маршрутизатор 4 и оптимальный путь до сети 3 через маршрутизатор 2.

5.3.2.2 Открытый протокол первоочередного поиска кратчайшего пути (Open Shortest Path First, OSPF)

Основные сообщения

HELLO – контроль связей. Передается только соседям, интервал передачи обычно 10с.

Если ответа нет – связь потеряна, передается сообщение об обновлении.

Элементы БД - регулярные сообщения, передаваемые всеми маршрутизаторами в режиме заливки; одна запись БД содержит IP- адрес анонсирующего маршрутизатора и описание его соединения с соседом (маршрутизатором или составляющей сетью).

Таким образом формируется топология объединенной сети в составе AS (все соединения каждого маршрутизатора), которая обрабатывается алгоритмом Дейкстры.

Длительный интервал передачи – 30 мин! – экономит пропускную способность сети.

Сообщение об обновлении состояния соединения – широковещательно рассылается анонсирующим маршрутизатором только при изменении состояния.

Запрос о состоянии соединения – отсылается, если при обновлении БД обнаружены отличия.

Основные свойства протокола OSPF

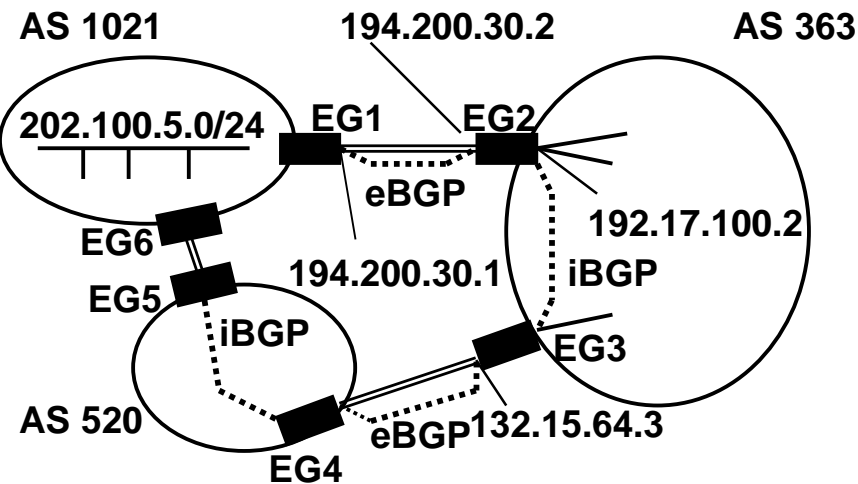
- Меньший (по сравнению с RIP) размер сообщений, только данные о соседях
- При вычислении оптимального пути каждым маршрутизатором используются одни и те же данные, которые не зависят от вычислений в других узлах (что способствует правильности).
- Оперативное распространение информации об изменениях (сообщения и запросы).

Дополнительные возможности

- Использование разных таблиц с учетом типа обслуживания (D,T,R), заданного в дейтаграмме.
- Распределение нагрузки при наличии путей одинаковой стоимости к одному получателю.
- Разбиение объединенной сети на зоны со скрытой топологией, маршрутизация между зонами.
- Классовая и бесклассовая адресация

и другие.

5.4 Маршрутизация между автономными системами. Протокол граничного шлюза (Border Gateway Protocol, BGP)



Связь между маршрутизаторами задается администраторами при конфигурировании путем передачи сообщения OPEN (протокол eBGP): для EG2 посылается адрес 194.200.30.1 для EG1 – адрес 194.200.30.2.
Все сообщения передаются на уровне TCP.

В общей форме новый маршрут задается следующим образом:

BGP Route = AS Path; Next Hop; Network/Mask_length.

Например, появление сети 202.100.5.0/24 в AS1021 передается от EG1 к EG2 сообщением типа UPDATE:

AS 1021; 194.200.30.1; 202.100.5.0/24.

EG1 → EG2: маршрутизатор EG2 запоминает эту новость в таблице:
сеть 202.100.5.0/24
следующий 194.200.30.1
получено от BGP

и посылает информацию о новой сети к маршрутизаторам в AS 363 по внутреннему протоколу OSPF, указывая в качестве анонсирующего адреса 192.17.100.2.

EG2 → EG3: получив от EG2 сообщение UPDATE (протокол iBGP), маршрутизатор EG3 пересылает к EG4 через протокол eBGP сообщение

AS 363, AS1021; 132.15.64.3; 202.100.5.0/24

и далее - с оповещением от EG4 всех маршрутизаторов AS 520 о новой сети и сообщением EG4 – EG5.

Процесс заканчивается на обмене EG5 – EG6, когда EG6 обнаруживает в AS Path номер своей AS (1021).

Принцип BGP позволяет исключить маршрутные петли и отдельно администрировать AS.