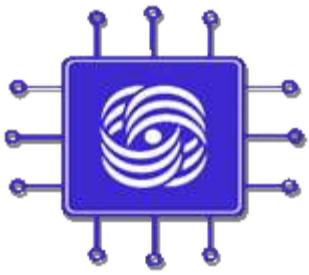
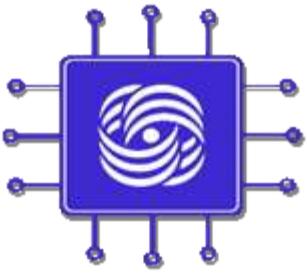


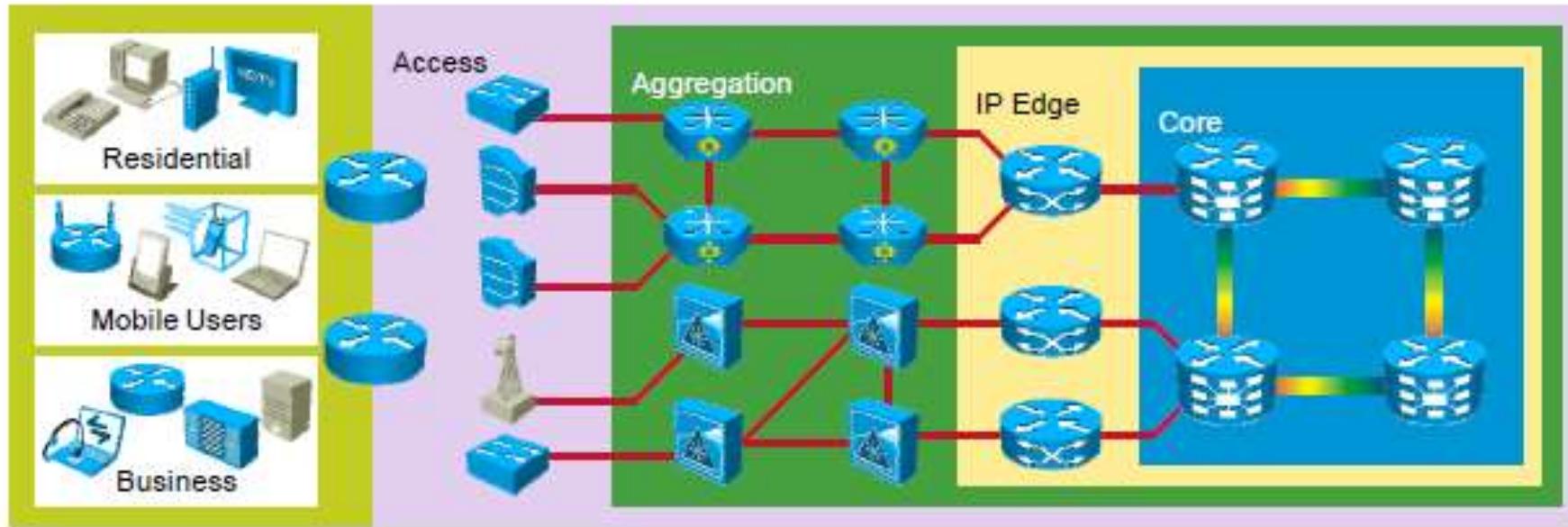
Системы передачи данных: маршрутизация

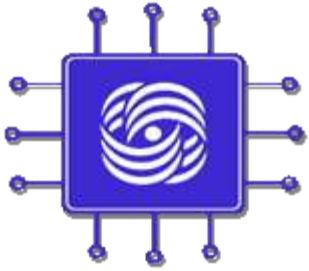


Структура Интернет



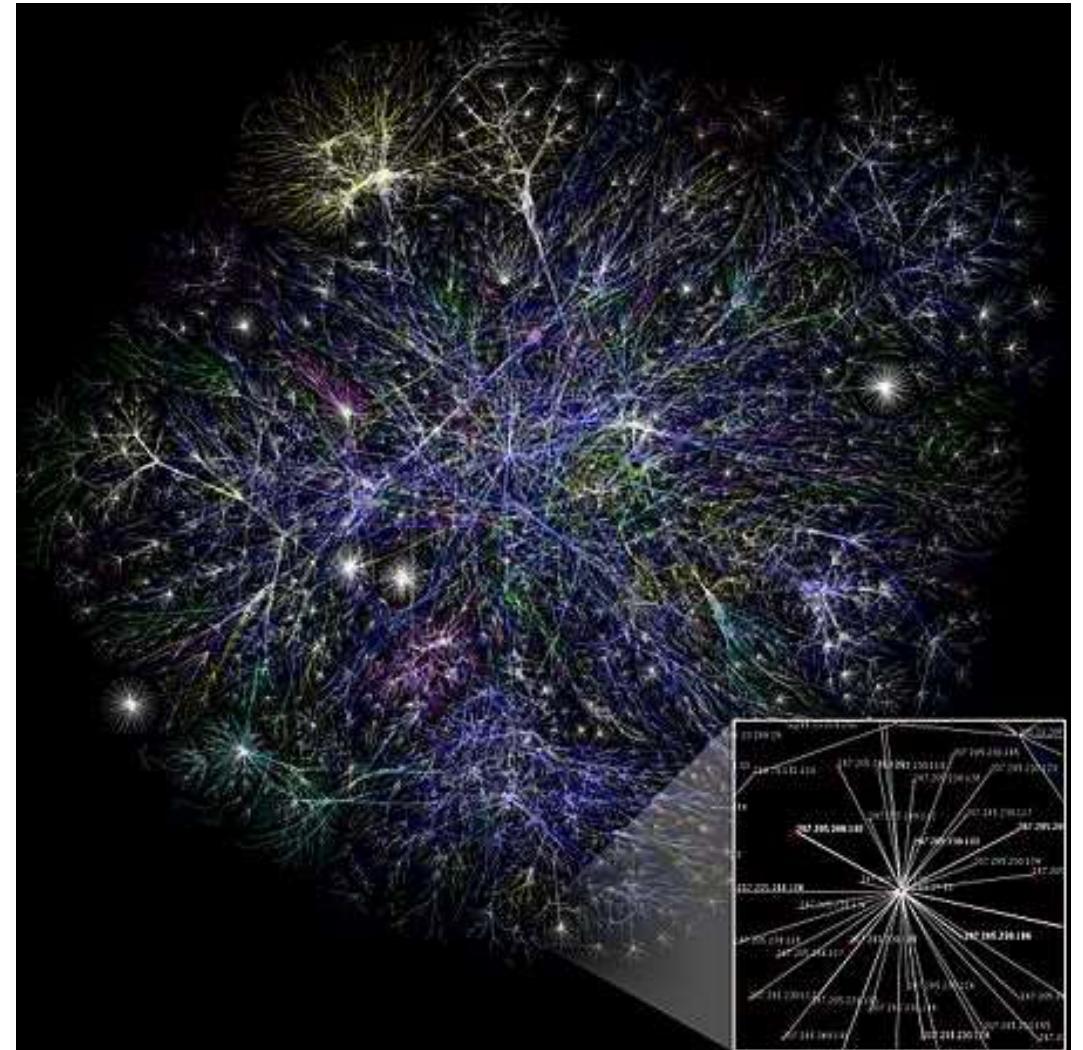
Структура сети

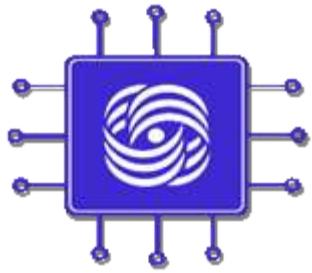




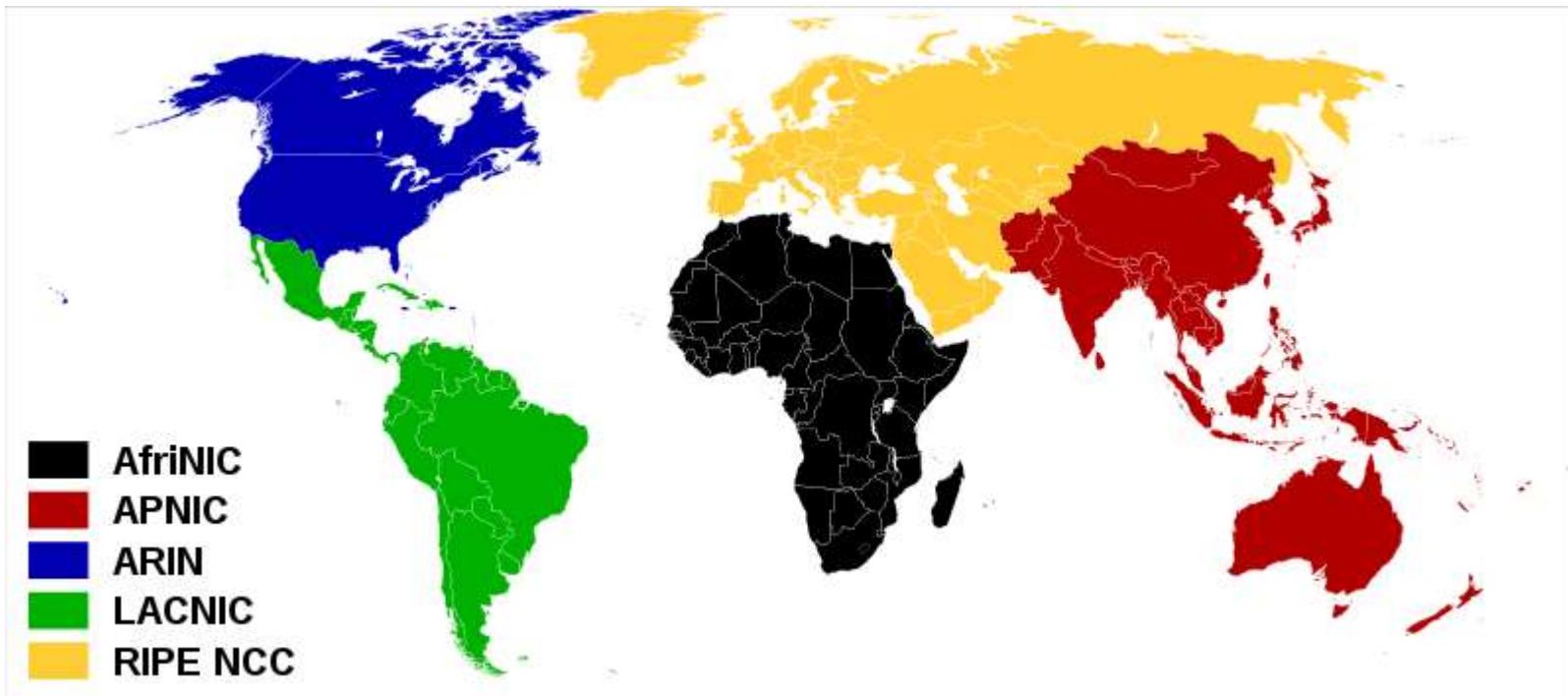
Структура сети

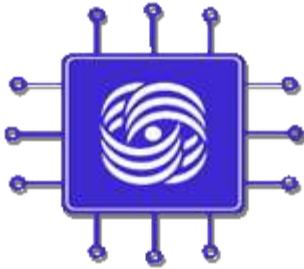
- Частичная карта Интернета, основанная на данных от 15 января 2005 года
- Карта представляет менее чем 30 % сетей класса С, доступных для сбора данных в 2005 году. Цвет линии соответствует её местоположению





RIR и LIR





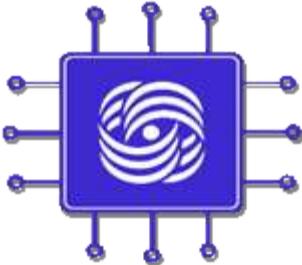
RIR и LIR

- LIR'ом может стать почти любая желающая организация при наличии необходимых документов

При подключении к провайдеру, организация получает диапазон публичных адресов – так называемые PA-адреса (Provider Aggregatable).

Если она не является LIR'ом, то при смене провайдера придётся возвращать и PA-адреса.

У LIR можно приобрести провайдеро-независимый блок адресов (PI) и ASN

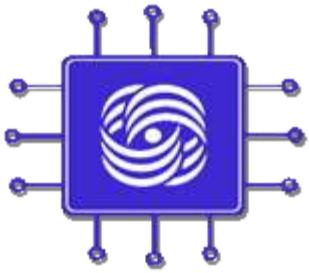


Автономная система

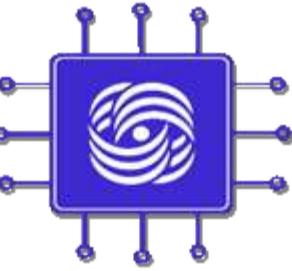
- **Автономная система (autonomous system, AS)** — набор маршрутизаторов, имеющих единые правила маршрутизации, управляемых одной технической администрацией и работающих на одном из протоколов IGP
- До 2007 года были возможны только 16-битные номера AS, то есть всего было доступно 65536, номеров.
- 0-65535 – зарезервированы.
64512-65534 - предназначены для приватных AS
- 64496-64511 – для использования в примерах и документации

Сейчас возможно использование 32-битных номеров AS.

На практике с каждой AS должен быть связан какой-то блок адресов



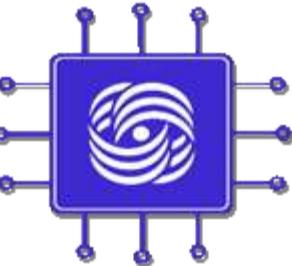
Алгоритмы маршрутизации



Алгоритмы маршрутизации

- Одна из основных задач сетей - транспортировка информации от ЭВМ-отправителя к ЭВМ-получателю.





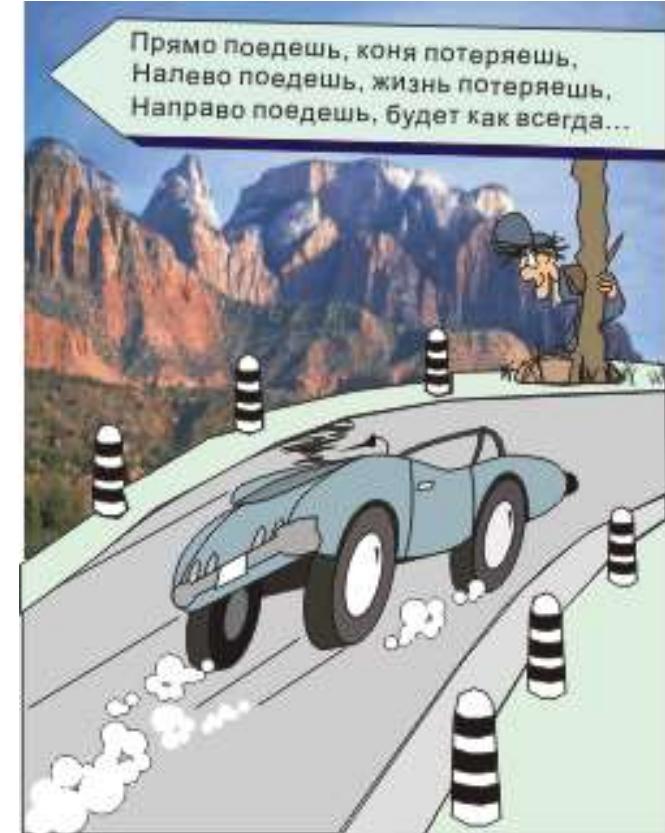
Алгоритмы маршрутизации

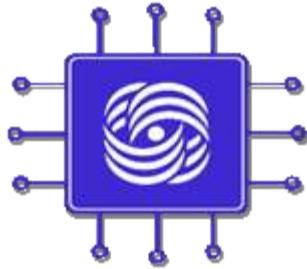
Маршрутизация подразумевает два параллельных процесса:

- подготовка маршрутной таблицы
- переадресация дейтограмм (с помощью этой таблицы).

Формирование маршрутной таблицы производится:

- посредством использования протоколов маршрутизации (динамическая маршрутизация)
- под воздействием инструкций сетевого администратора – (статическая маршрутизация).

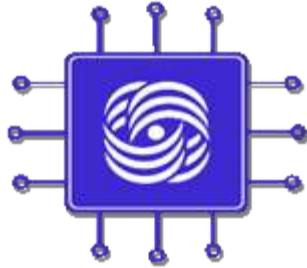




Алгоритмы маршрутизации

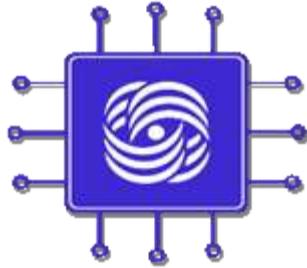
- unicast
- multicast.
- лавина
- маршрутизация от источника
- «горячая картошка».





Метрики маршрутов

- Обычно каждому сегменту, составляющему маршрут, присваивается некоторая величина - оценка этого сегмента.
- Каждый протокол маршрутизации использует свою систему оценки маршрутов.
- Примеры:
 - RIP – на основе подсчета количества маршрутизаторов, которые нужно пройти (hop count)
 - OSPF – на основе пропускной способности каналов

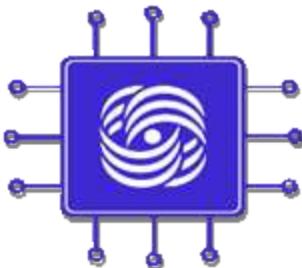


Виды протоколов маршрутизации

- IGP (внутренние по отношению к вашей автономной системе) - ISIS/OSPF/RIP/EIGRP.
- EGP (внешние) - BGP – Border Gateway Protocol.

По используемым типам алгоритмов делятся на:

- DV (Distance Vector)
- LS (Link State).



Протоколы внутренней маршрутизации

Маршрутизация по единым метрикам

RIP (RFC 2453)

(www.rfc.com.ru, <http://www.ietf.org/rfc.html>)

- используют алгоритм по вектору расстояния (алгоритм Б-Ф)
- обновление векторов каждые 30 секунд
- аутентификация при обновлениях не применяется
- изначально был использован в *BSD Unix*
- сегодня применяется редко

OSPF (RFC 2328)

- изменения состояний линии рассылаются лавиной по необходимости
- каждый маршрутизатор использует алгоритм Дейкстры
- изменения аутентифицируются
- АС можно разбивать на области
- Широко используется, сложный аналог *IS-IS (RFC 1142)*, который также широко используется

OSPF Работа протокола

1. Маршрутизаторы обмениваются *hello*-пакетами через все интерфейсы, на которых активирован OSPF. Маршрутизаторы, разделяющие общий канал передачи данных, становятся соседями, когда они приходят к договоренности об определённых параметрах, указанных в их *hello*-пакетах.
2. На следующем этапе работы протокола маршрутизаторы будут пытаться перейти в состояние смежности со своими соседями.
3. Каждый маршрутизатор посылает объявления о состоянии канала маршрутизаторам, с которыми он находится в состоянии смежности.

OSPF Работа протокола

4. Каждый маршрутизатор, получивший объявление от смежного маршрутизатора, записывает передаваемую в нём информацию в базу данных состояния каналов маршрутизатора и рассыпает копию объявления всем другим смежным с ним маршрутизаторам.
5. Рассыпая объявления внутри одной OSPF-зоны, все маршрутизаторы строят идентичную базу данных состояния каналов маршрутизатора.
6. Вычисление по алгоритму Дейкстры: каждый маршрутизатор независимо запускает алгоритм Дейкстры найкратчайшего пути
7. Каждый маршрутизатор строит таблицу маршрутизации из своего дерева кратчайших путей.

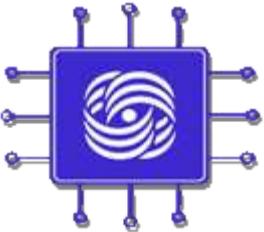
Каждый маршрутизатор находит соединяющее дерево с минимальной стоимостью до каждого другого маршрутизатора

OSPF Типы маршрутизаторов

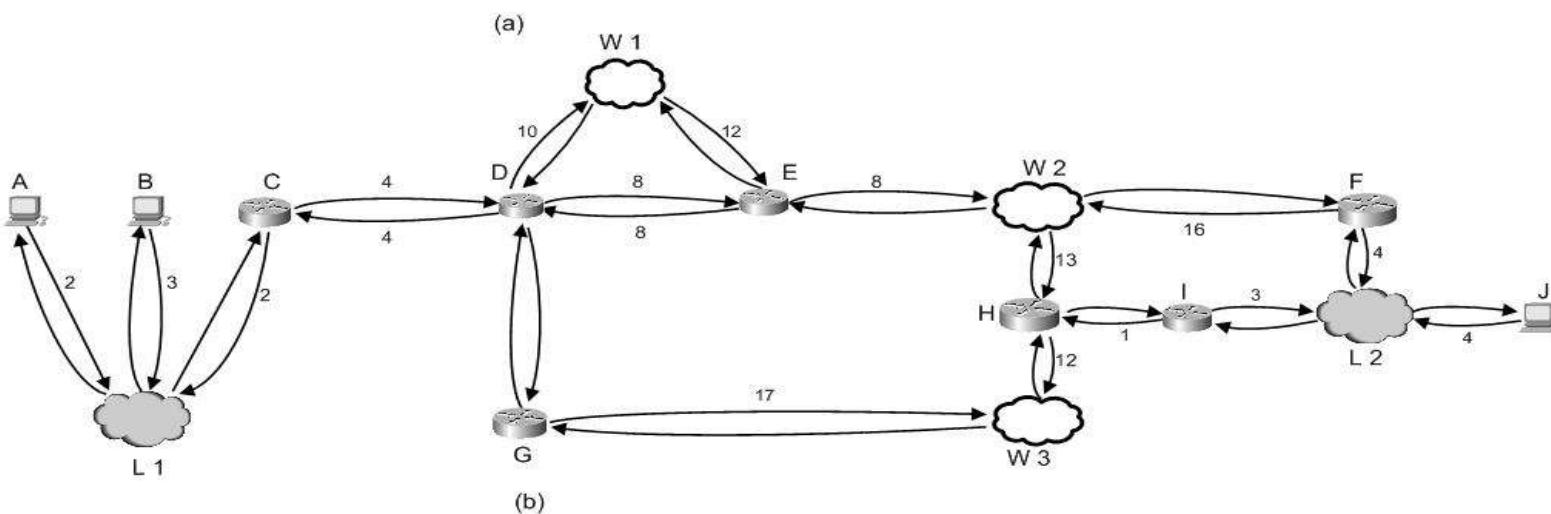
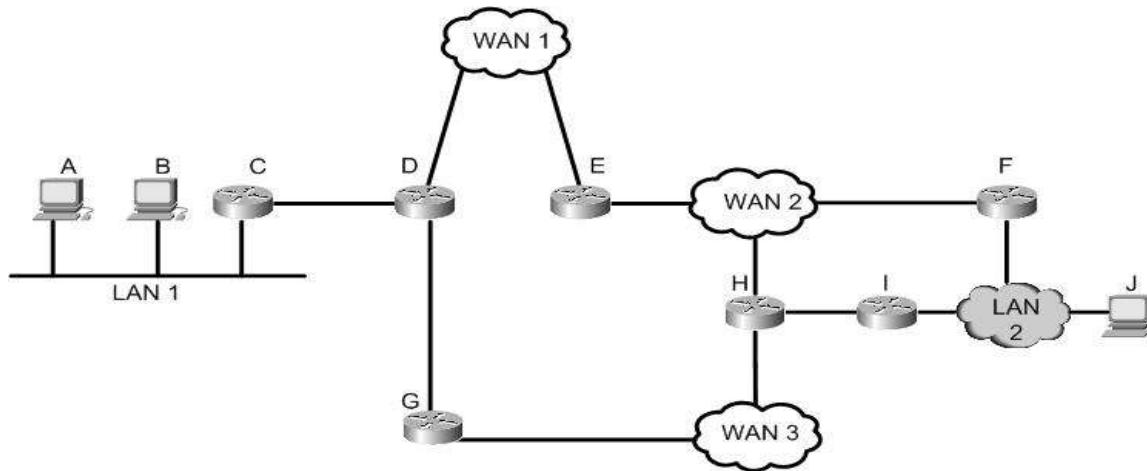
- Внутренний маршрутизатор (*internal router*) – маршрутизатор, все интерфейсы которого принадлежат одной зоне. У таких маршрутизаторов только одна база данных состояния каналов.
- Пограничный маршрутизатор (*area border router, ABR*) – соединяет одну или больше зон с магистральной зоной и выполняет функции шлюза для межзонального трафика.
- Магистральный маршрутизатор (*backbone router*) – маршрутизатор у которого всегда хотя бы один интерфейс принадлежит магистральной зоне.

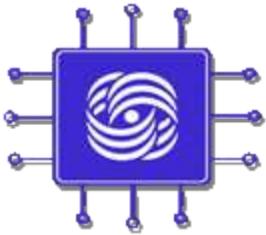
OSPF Типы маршрутизаторов

- Пограничный маршрутизатор автономной системы (*AS boundary router, ASBR*) — это маршрутизатор, один из портов которого находится в домене OSPF протокола, а другой в домене любого из внутренних шлюзовых протоколов (например *RIP* или *EIGRP*).

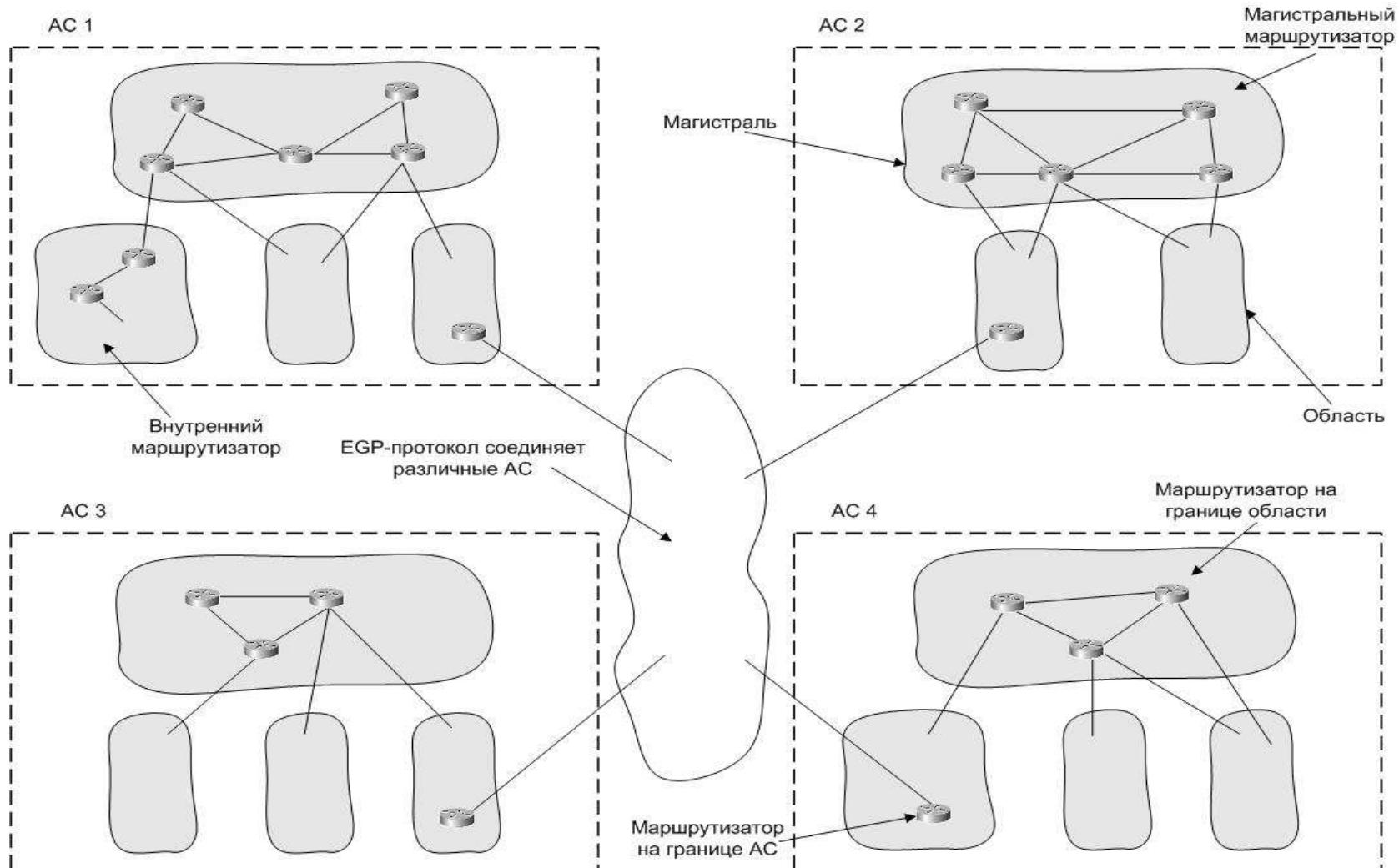


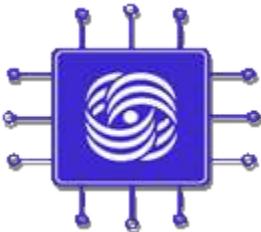
Автономная система изнутри





AS, магистрали и области OSPF

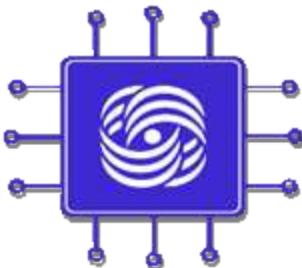




Управляющие сообщения OSPF протокола

- В OSPF маршрутизаторы обмениваются данными не с соседними маршрутизаторами, а со *смежными маршрутизаторами*.
- Каждый маршрутизатор имеет дублера (зеркало).

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner



Протокол OSPF

Формат заголовка пакета OSPF

Все пять типов пакетов OSPF инкапсулируются непосредственно в IP-пакет. Номер протокола OSPF в IP-заголовке — 89.

У всех пакетов OSPF одинаковый заголовок.

Поля заголовка пакета OSPF:

Версия протокола (Version number) — версия протокола OSPF. Текущая версия для IPv4 — 2.

Тип пакета (Type) — указывает какой тип пакета OSPF передается:

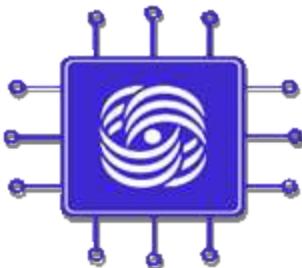
- 1 — Hello
- 2 — Database Description
- 3 — Link State Request
- 4 — Link State Update
- 5 — Link State Acknowledgment

Длина пакета (Packet length) — длина пакета OSPF в байтах. Длина включает в себя и заголовок.

Идентификатор маршрутизатора (Router ID) — определяет какой маршрутизатор отправил пакет.

Идентификатор зоны (Area ID) — определяет в какой зоне сгенерирован пакет.

Контрольная сумма (Checksum) — используется для проверки целостности пакета OSPF, для обнаружения ошибок при передаче.



Протокол OSPF

Формат заголовка пакета OSPF

Тип аутентификации (Authentication type) — тип аутентификации, который используется между маршрутизаторами:

- 0 — аутентификация не используется,
- 1 — аутентификация открытым текстом,
- 2 — MD5-аутентификация.

Данные аутентификации (Authentication) — используется при аутентификации маршрутизаторов.

Поле **Данные** отличается для различных типов пакетов OSPF:

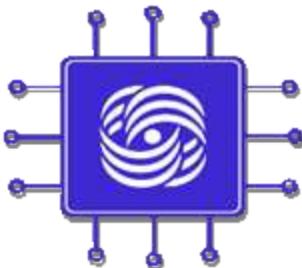
Hello — список известных соседей

DBD — содержит суммарную информацию базы данных состояний каналов, которая включает в себя все известные идентификаторы маршрутизаторов и их последние номера последовательностей (sequence number) и другую информацию.

LSR — содержит тип необходимого LSU и идентификатор маршрутизатора, у которого есть этот LSU.

LSU — содержит полные записи объявления о состоянии канала. Несколько LSA могут передаваться в одном пакете обновлений.

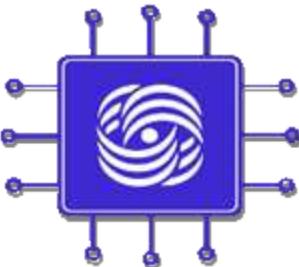
LSAck — поле пустое



Протокол OSPF

Формат заголовка пакета Hello

- С помощью Hello-пакета маршрутизатор обнаруживает своих соседей;
- Он передает параметры о которых маршрутизаторы должны договориться прежде чем они станут соседями;
- Hello-пакеты выполняют роль keepalive-пакетов между соседями;
- Отвечает за установление двухсторонних коммуникаций между соседними маршрутизаторами (двуихсторонняя коммуникация установлена тогда, когда маршрутизатор увидит себя в списке соседей hello-пакета полученного от соседнего маршрутизатора);
- Он выбирает DR и BDR в широковещательных и нешироковещательных сетях со множественным доступом.



Протокол OSPF

Формат заголовка пакета Hello

Кроме стандартного заголовка пакета OSPF, в hello-пакете содержится:

Network mask — сетевая маска назначенная на интерфейсе с которого отправляется hello-пакет;

HelloInterval — частота отправки hello-пакетов;

Options — дополнительные опции, которые может поддерживать маршрутизатор. Это поле есть в пакетах hello, DBD, LSA. Согласно [RFC 2328](#) могут быть установлены такие 5 битов:

E-bit — указывает каким образом распространяются AS-external-LSA

MC-bit — указывает соответствует ли IP multicast датаграмма спецификациям [RFC 1584](#);

N/P-bit — указывает как обрабатываются Type-7 LSA ([RFC 1587](#));

EA-bit — указывает желание маршрутизатора получать и отправлять External-Attributes-LSA;

DC-bit — указывает как маршрутизатор работает с demand circuits ([RFC 1793](#));

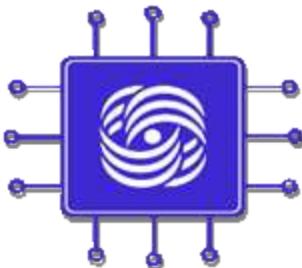
Rtr Pri — приоритет маршрутизатора. Используется для выбора DR и BDR. Если приоритет установлен в 0, то маршрутизатор не может стать DR или BDR;

RouterDeadInterval — интервал времени по истечению которого сосед будет считаться "мертвым";

Designated Router — IP-адрес DR для сети в которую отправлен hello-пакет, с точки зрения отправляющего маршрутизатора. Поле установлено в значение 0.0.0.0, если в сети нет DR;

Backup Designated Router — IP-адрес BDR для сети в которую отправлен hello-пакет, с точки зрения отправляющего маршрутизатора. Поле установлено в значение 0.0.0.0, если в сети нет BDR;

Neighbor — идентификаторы всех маршрутизаторов от который маршрутизатор-отправитель получал hello-пакеты в течение последнего интервала RouterDeadInterval.



OSPF Типы объявлений о состоянии канала (LSA)

Type 1 LSA – Router LSA – объявление о состоянии каналов маршрутизатора.

Type 2 LSA – Network LSA – объявление о состоянии каналов сети. Распространяется DR в сетях со множественным доступом.

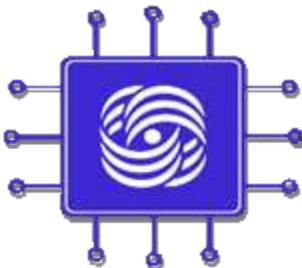
Type 3 LSA – Network Summary LSA – суммарное объявление о состоянии каналов сети. Объявление распространяется пограничными маршрутизаторами.

Type 4 LSA – ASBR Summary LSA – суммарное объявление о состоянии каналов пограничного маршрутизатора автономной системы.

Type 5 LSA – AS External LSA – объявление о состоянии внешних каналов автономной системы.

Type 6 LSA – Multicast OSPF LSA – специализированный LSA, который используют мультикаст OSPF приложения.

Type 7 LSA – AS External LSA for NSSA – объявления о состоянии внешних каналов автономной системы в NSSA зоне.



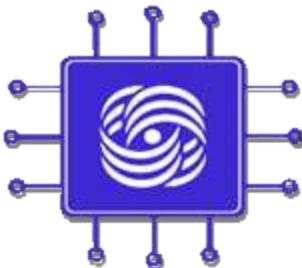
OSPF Типы объявлений о состоянии канала (LSA)

Type 8 LSA – Link LSA – анонсирует *link-local* адрес и префикс(ы) маршрутизатора всем маршрутизаторам разделяющим канал (*link*).

Отправляется только если на канале присутствует более чем один маршрутизатор. Распространяются только в пределах канала (*link*).

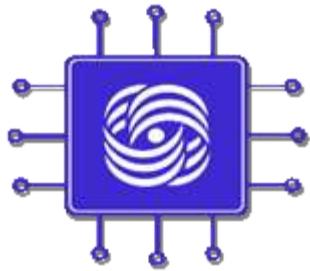
Type 9 LSA – Intra-Area-Prefix LSA ставит в соответствие: список префиксов IPv6 и маршрутизатор, указывая на *Router LSA*, список префиксов IPv6 и транзитную сеть, указывая на *Network LSA*.

Распространяются только в пределах одной зоны.



OSPF Типы зон

- Магистральная зона (*backbone area*)
- Стандартная зона (*standard area*)
- Тупиковая зона (*stub area*)
- *Totally stubby area*
- *Not-so-stubby area (NSSA)*



Border Gateway Protocol (BGP-4) Основные термины

Внутренний протокол маршрутизации (interior gateway protocol)

Внешний протокол маршрутизации (exterior gateway protocol)

Автономная система (autonomous system, AS)

Транзитная автономная система (transit AS)

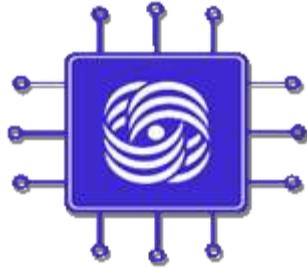
Путь (path)

Атрибуты пути (path attributes, PA)

BGP speaker

Соседи (neighbor, peer)

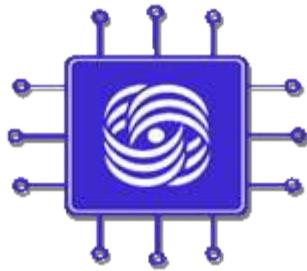
**Информация сетевого уровня о доступности сети
(Network Layer Reachability Information, NLRI)**



Автономные системы - AS

AC – это система IP-сетей и маршрутизаторов, управляемых одним или несколькими операторами, имеющими единую политику маршрутизации.

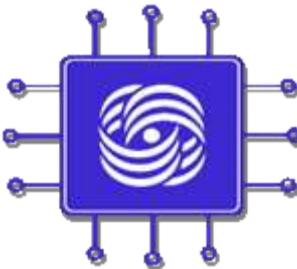
Выдачей номеров AC занимаются RIR (Regional Internet Registry) или LIR (Local Internet Registry).



Междоменная маршрутизация

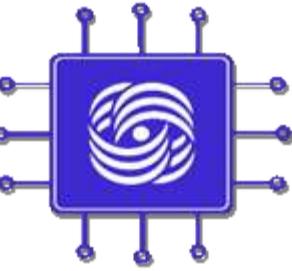
Border Gateway Protocol (BGP-4)

- BGP использует «вектор пути»
- Каждый BGP маршрутизатор рассыпает список путей (путь - список AS)
 - AS_PATH
 - К сети 171.64/16 можно пройти по маршруту {AS7,AS52,AS13}
- Наличие цикла в маршруте определяется локально и такие маршруты игнорируются
- Из множества доступных маршрутов выбирается тот, которые наиболее всего соответствует политике AC
- Если маршрутизатор/линии вышли из строя, то маршрут изымается из списка

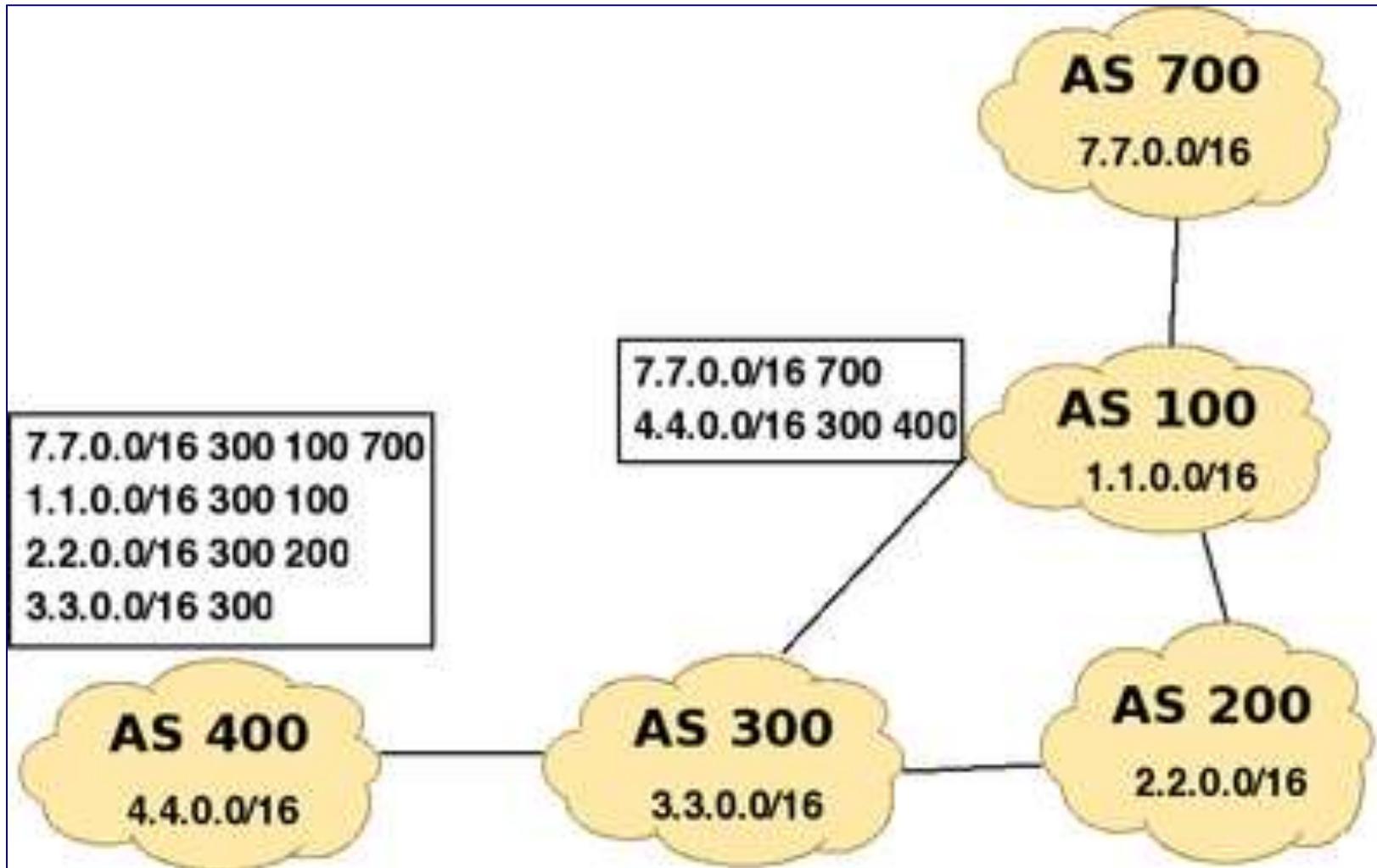


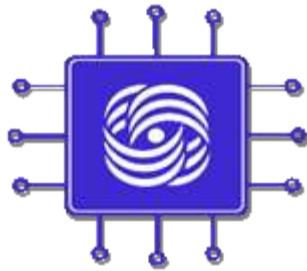
BGP сообщения

- *Open* Установка BGP сессии (TCP port 179)
- *Keep Alive* Проверка работоспособности через регулярные интервалы
- *Notification* Закрытие peering сессии
- *Update* Объявление нового или изъятие ранее объявленного маршрута
- ***BGP объявление = префикс + атрибуты маршрута***
- *Path attributes*
 - Origin кто предложил этот маршрут
 - список AC (path), предпочтения, правило определения выхода
 - следующий скачок (IP для BGP шлюза следующей АС),
 - multi_exit_disc - для выбора одного из нескольких объявляющих
 - Local pref префиксы локальных АС
 - Aggregate - объединение нескольких маршрутов с общим префиксом.

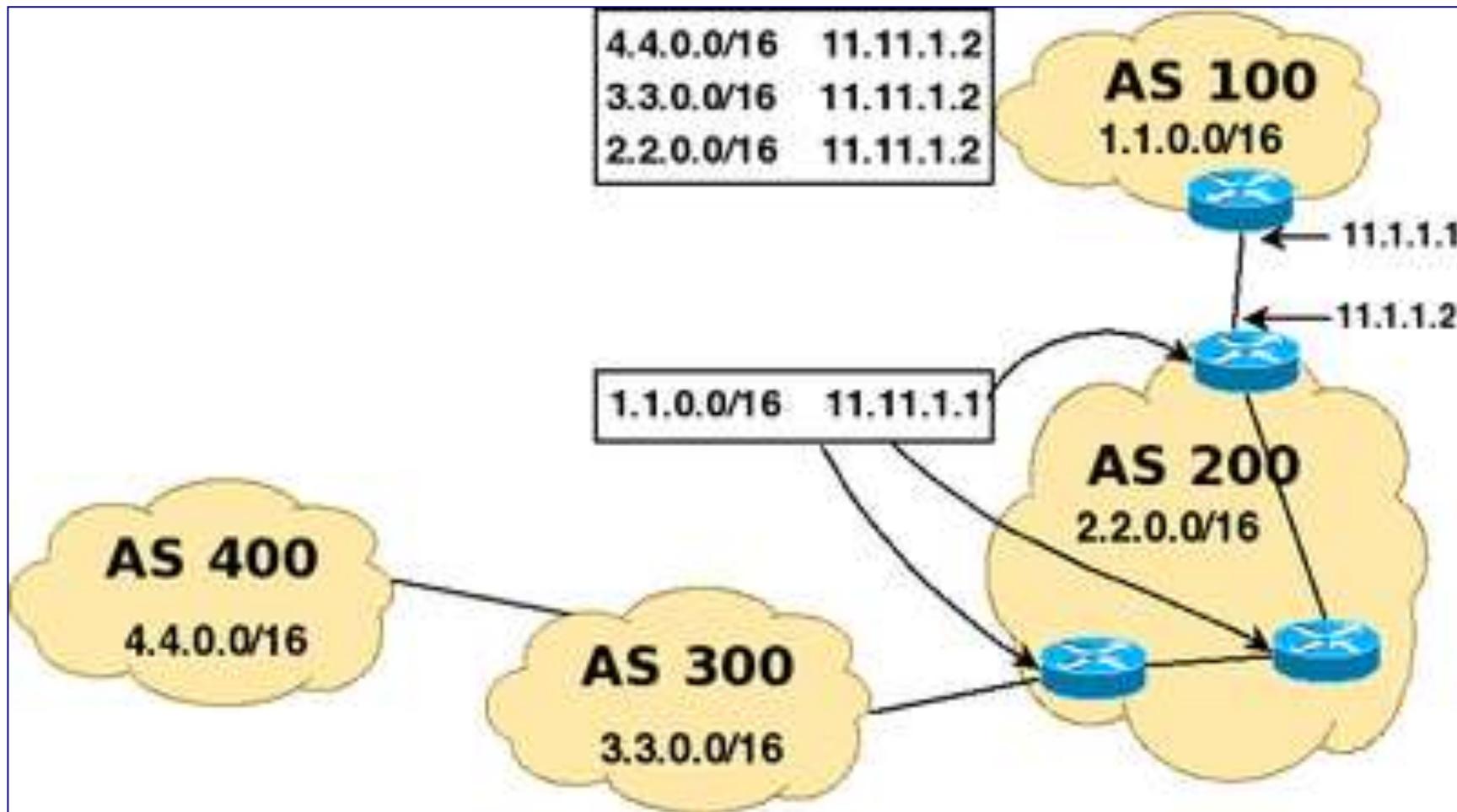


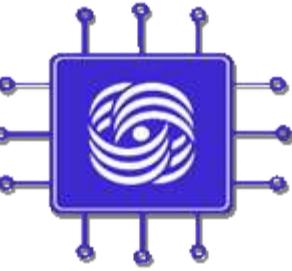
Список АС



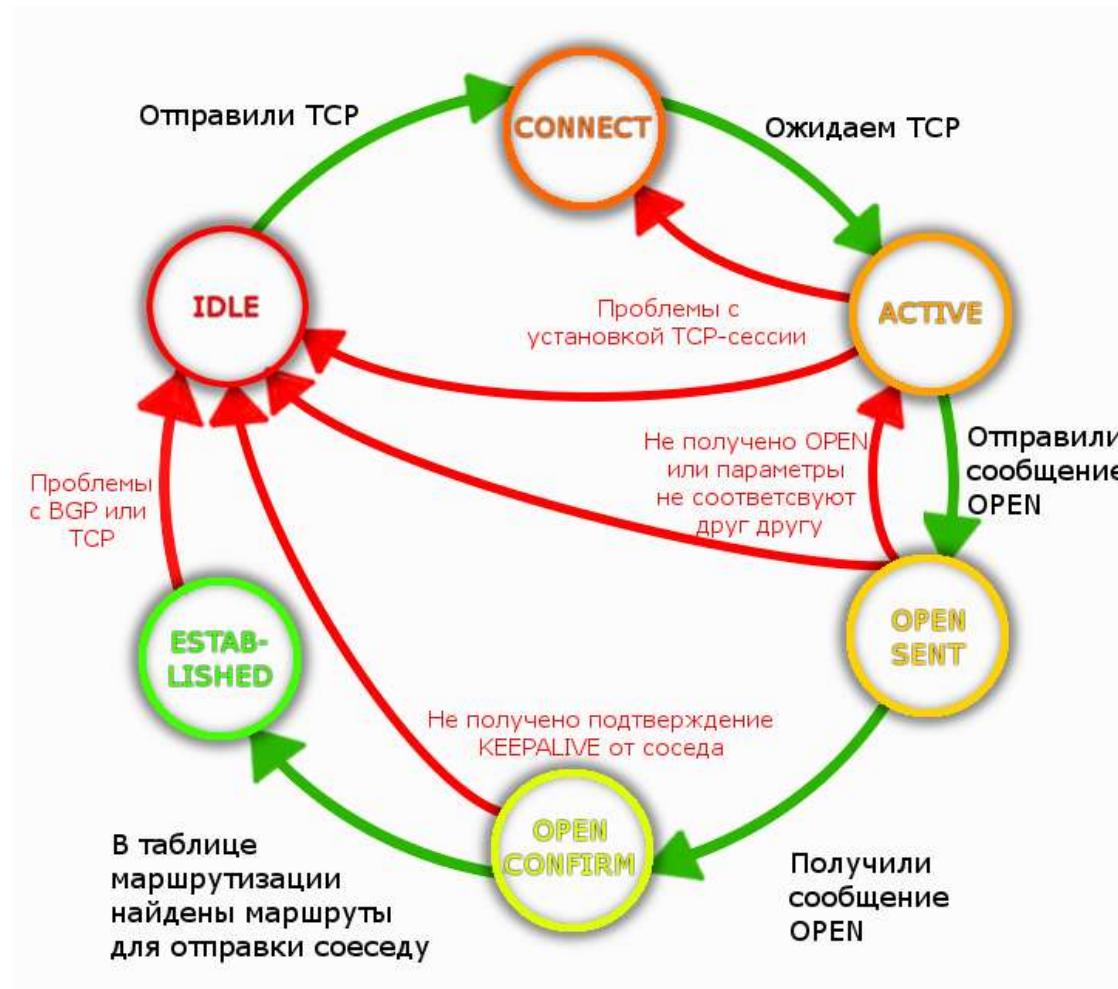


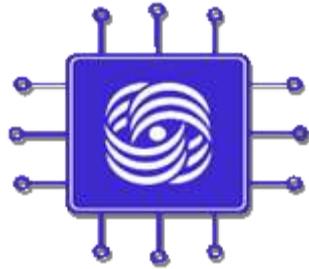
Следующий скачок



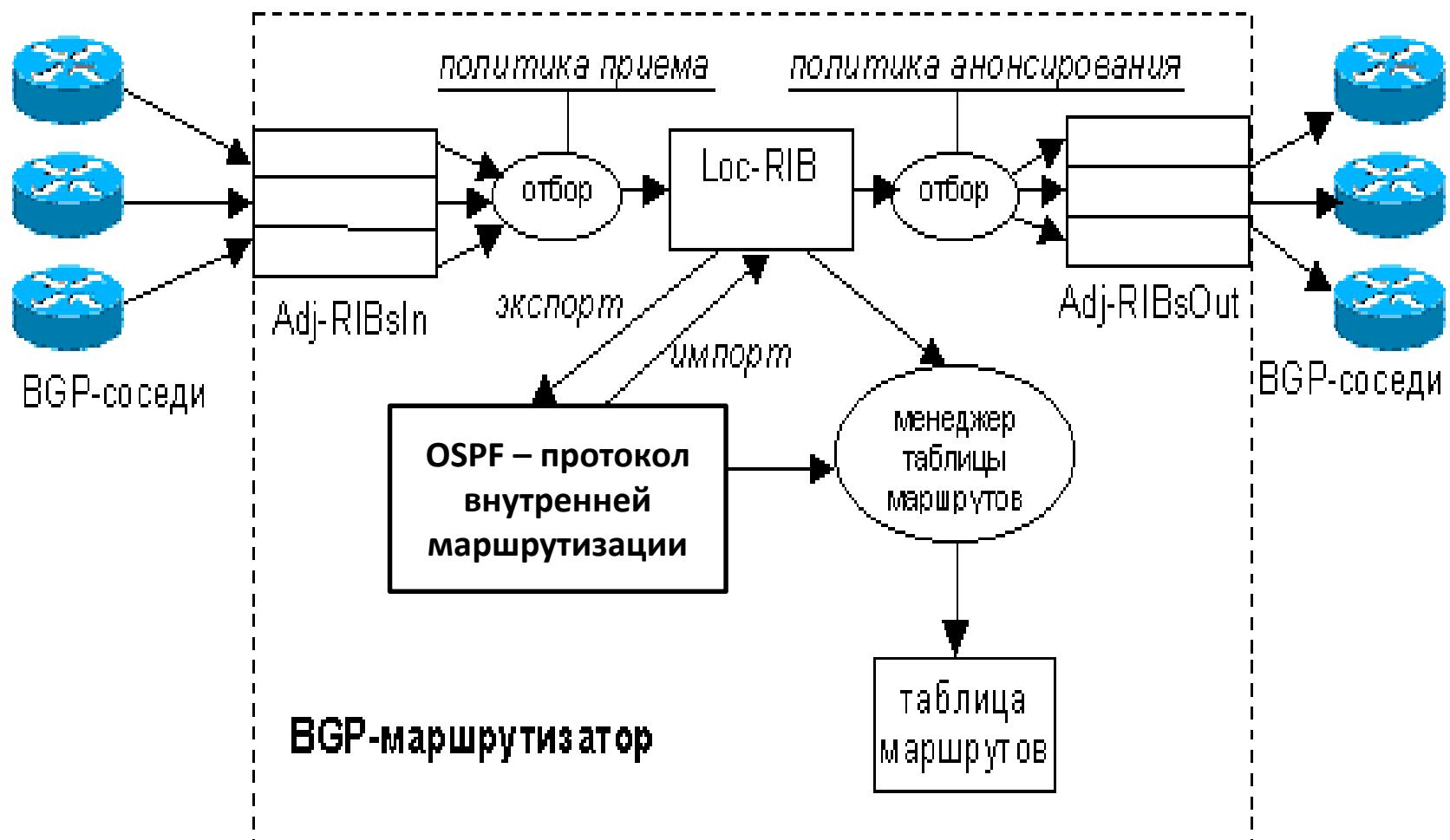


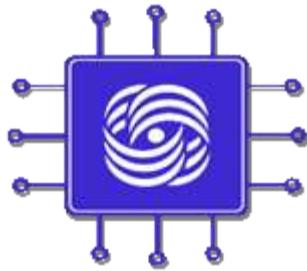
Состояния





Процедура выбора маршрута

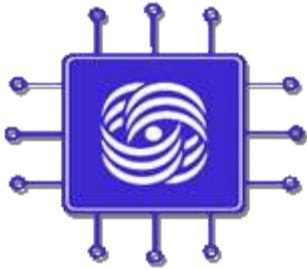




Border Gateway Protocol (BGP-4) Выбор пути

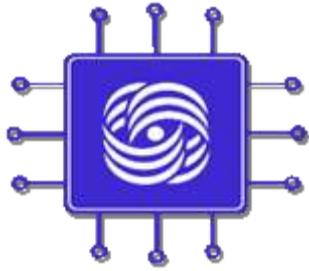
Если не настроены политики выбора пути, выбор происходит следующим образом :

1. Максимальное значение weight (локально для маршрутизатора).
2. Максимальное значение local preference (для всей AS).
3. Предпочесть локальный маршрут маршрутизатора (next hop = 0.0.0.0).
4. Кратчайший путь через автономные системы. (самый короткий AS_PATH)
5. Минимальное значение origin code (IGP < EGP < incomplete).
6. Минимальное значение MED (распространяется между автономными системами).
7. Путь eBGP лучше чем путь iBGP.
8. Выбрать путь через ближайшего IGP-соседа.
9. Выбрать самый старый маршрут для eBGP-пути.
10. Выбрать путь через соседа с наименьшим BGP router ID.
11. Выбрать путь через соседа с наименьшим IP-адресом.



BGP - ИТОГ

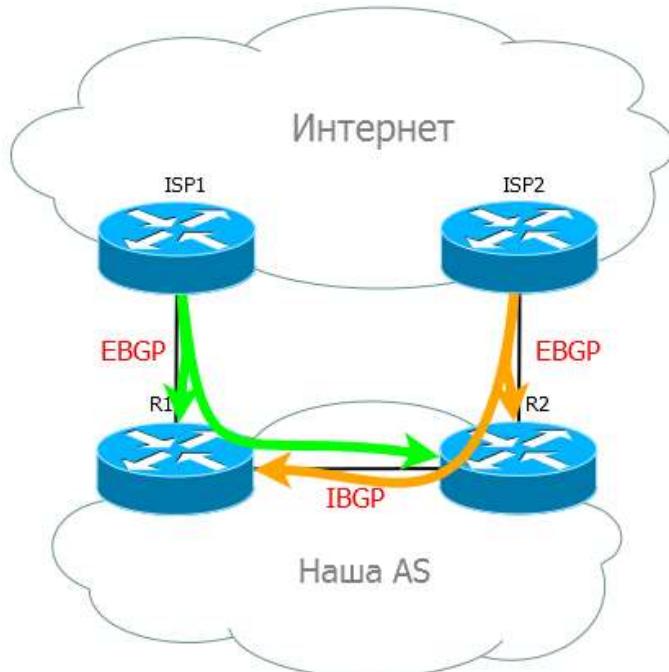
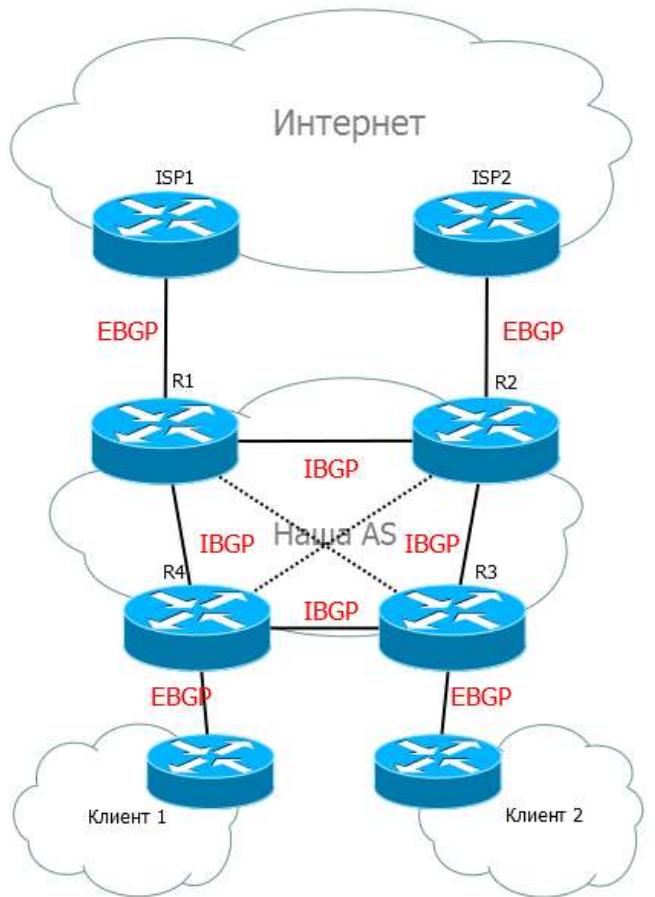
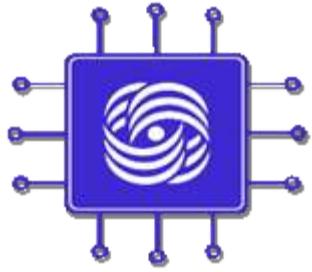
- Все АС для взаимодействия в Интернете должны использовать BGP-4
- BGP-4 использует алгоритм маршрутизации по вектору пути, которые легко распознает циклы
- BGP-4 имеет сложный интерфейс, позволяющий каждой АС устанавливать свою локальную политику маршрутизации
- Каждая АС устанавливает свою политику для построения маршрутов, безопасности и локальных особенностей



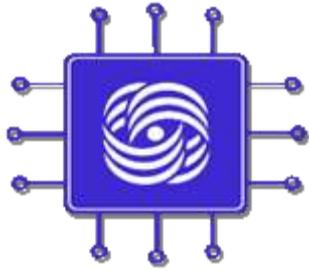
Что такое IBGP

- Internal BGP (IBGP) - это тот же самый BGP, но работающий **внутри AS**.
- **Резервирование.** Когда есть несколько подключений к провайдерам и не нужно замыкать всё на одном граничном маршрутизаторе.
- **Подключение клиентов по BGP.** Если стоит задача подключить клиента по BGP, при этом у вас больше, чем один маршрутизатор.

Что такое IBGP



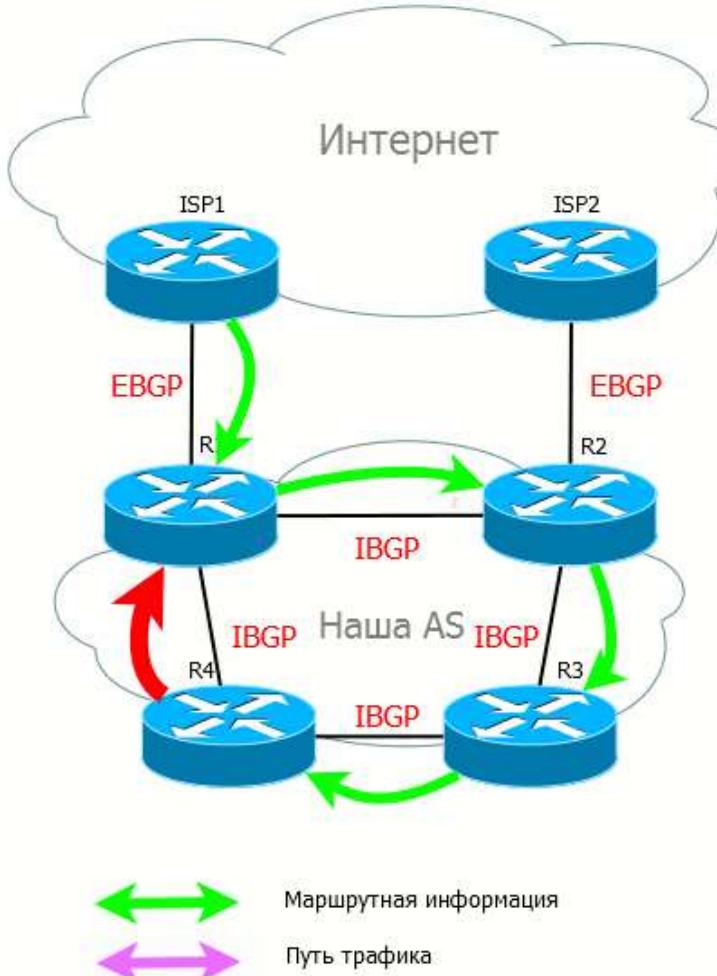
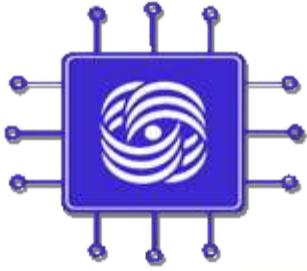
↔ Маршрутная информация



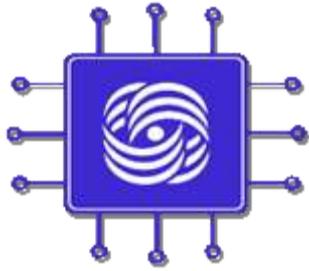
Что такое iBGP

- Если iBGP-маршрутизаторы работают в нетранзитной AS, то соединение между ними должно быть full mesh.
- Это следствие принципов работы протокола — если маршрутизатор, находящийся на границе AS, получил обновление, то он передает его всем соседям; соседи, которые находятся внутри автономной системы, больше это обновление не распространяют, так как считают, что все соседи внутри AS уже его получили.

Что такое IBGP



- В случае полносвязной топологии и правила Split Horizon, исключается ситуация, когда в качестве приоритетного может выбраться даже новый маршрут от R4, это, неэффективно, т.к. маршруты будут изучены неверно, может образоваться петля и трафик не попадёт к точке назначения.,
- R1 получив анонс от ISP1, передаёт его сразу всем своим соседям: R2, R3, R4. А те в свою очередь эти анонсы сохраняют, но передают только EBGP-партнёрам, но не IBGP, именно потому, что получены от IBGP-партнёра. То есть все BGP-маршрутизаторы имеют актуальную информацию и исключены петли



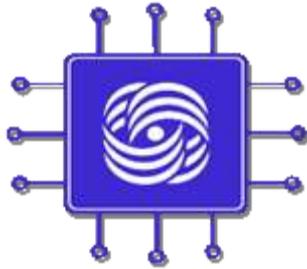
Различия IBGP и EBGP

1) В EBGP справляются с петлями с помощью AS-Path. Если в списке уже был номер собственной AS, то такой маршрут отбрасывается.

При передаче маршрута внутри Автономной Системы AS-Path не меняется. В IBGP используется **полносвязная топология — Full Mesh**.

При этом маршрут, полученный от IBGP-соседа не анонсируется другим IBGP-соседям.

Это позволяет на всех маршрутизаторах иметь все маршруты и при этом не допустить петель.



Различия IBGP и EBGP

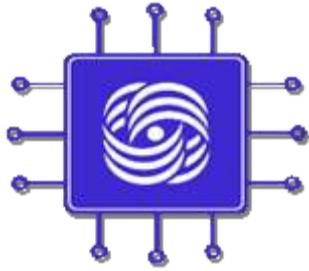
2) адрес Next Hop. В External BGP маршрутизатор при отправке анонса своему EBGP-соседу сначала меняет адрес Next-Hop на свой, а потом уже отсылает.

Если маршрутизатор передаёт анонс IBGP-соседу, то адрес Next-Hop не меняется.

Понятие Next-Hop отличается от того, которое используется в IGP. В IBGP он сообщает о точке выхода из локальной AS.

Важно, чтобы у получателя такого анонса был маршрут до Next-Hop. Если его не будет, то маршрут будет помещён в таблицу BGP, но его не будет в таблице маршрутизации.

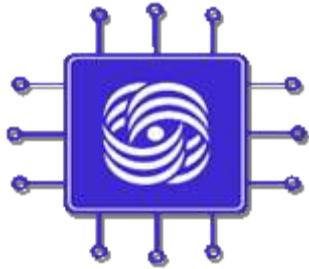
Такой процесс называется рекурсивной маршрутизацией.



Различия IBGP и EBGP

3) Если в EBGP обычно подразумевается прямое подключение двух соседей друг к другу, то в Internal BGP соседи могут быть подключены через несколько промежуточных устройств.

- в EBGP также можно настраивать соседей, которые находятся через несколько промежуточных устройств друг от друга, но для IBGP это работает по умолчанию.
- позволяет устанавливать IBGP-партнёрство между Loopback-адресами.
- в EBGP это не используется



Взаимодействие с OSPF

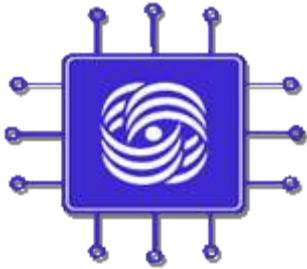
Как правило, IBGP настраивается с использованием существующего на сети IGP. IGP обеспечивает:

- связность всех маршрутизаторов между собой по IP,
- быструю реакцию на изменения в топологии
- перенос маршрутной информации о внутренних сетях.

1. создадим Loopback-интерфейсы :

Эти настройки будут использованы для определения Router ID и для OSPF и для BGP.

2. Настройка внутренней маршрутизации OSPF



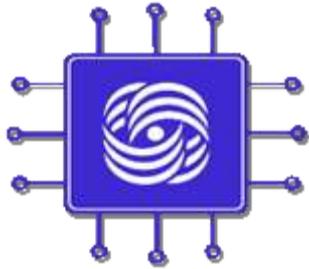
Взаимодействие с OSPF

Наша задача, чтобы все знали обо всех подключенных подсетях, адресах Loopback-интерфейсов и, естественно, о наших белых (публичные/глобальные/внешние) адресах.

Пример конфигурации OSPF (*Cisco*):

```
router ospf 1
network 1.1.1.1 0.0.0.0 area 0
network 10.0.0.0 0.255.255.255 area 0
network 100.0.0.0 0.0.1.255 area 0
```

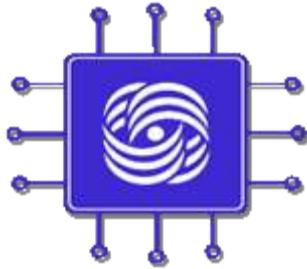
После этого появляется связность со всеми Loopback-адресами.



Настраиваем BGP

На каждом узле нужно настроить всех соседей вручную:

```
router bgp 64500
network 100.0.0.0 mask 255.255.254.0
neighbor 2.2.2.2 remote-as 64500
neighbor 2.2.2.2 update-source Loopback0
```



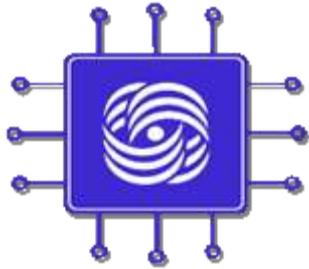
Настраиваем BGP

Адреса Next-Hop полученных маршрутов не меняются.

Для решения этой проблемы можно:

- 1) Настроить статические маршруты до этих адресов.
- 2) Добавить эти интерфейсы (в сторону провайдеров) в домен IGP-маршрутизации. Внешние сети не рекомендуется добавлять в IGP.
- 3) Менять адрес Next-Hop при передаче IBGP-соседям.

Добавляем в BGP ещё такую команду: `neighbor x.x.x.x Next-Hop-self`. Для каждого соседа, на каждом узле

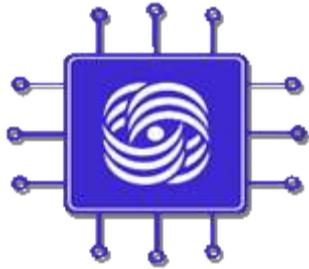


Может ли BGP обойтись без IGP

Вопросы:

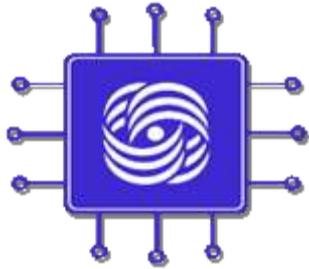
- Почему нельзя информацию про публичные адреса передавать по IBGP?
- Почему нельзя обойтись вообще одним BGP, без OSPF или IS-IS, например.

Самое главное препятствие – Full Mesh. Придётся устанавливать соседство со всеми маршрутизаторами вручную. Даже не смотря на наличие Route Reflector и скриптов.



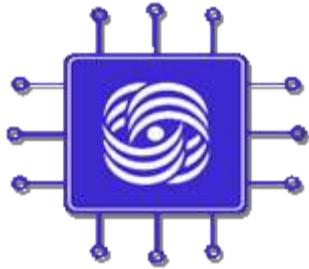
Может ли BGP обойтись без IGP

- медленная реакция и Дистанционно-Векторный подход к распространению маршрутной информации.
- отсутствие возможности автоматического изучения соседей. Что ведёт к ручной их конфигурации.



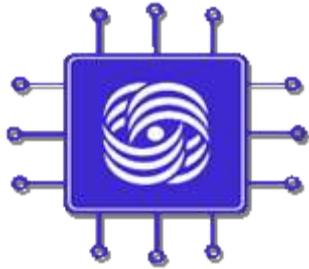
Может ли BGP обойтись без IGP

- *IGP и IBGP работают совместно, каждый выполняет свою работу.*
- *IGP обеспечивает внутреннюю IP-связность, реакцию на изменения в сети, оповещение всех узлов. Он же знает о публичных адресах нашей AS.*
- *IBGP занимается обработкой внешних маршрутов в нашей AS и их транзитом к клиентам и обратно. Обычно он ничего не знает о структуре внутренней сети.*



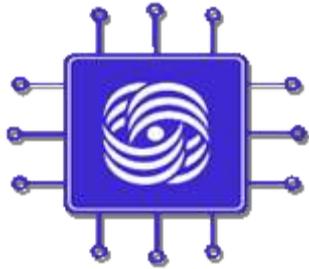
Дополнительные возможности IBGP

- *Route Reflector*
 - *Route Reflector* – это специальный IBGP-маршрутизатор, который выполняет функцию «отражения» маршрутов – ему присыпает маршрут один сосед, а он рассыпает его всем другим. То есть на IBGP-маршрутизаторах нужно настроить сессию только с одним соседом – с *Route Reflector*'ом. Прямая аналогия DR OSPF.



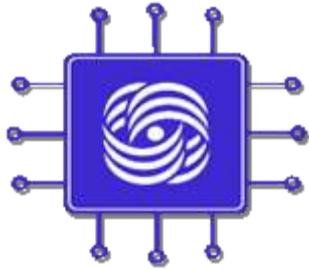
Дополнительные возможности IBGP

- Конфедерации
 - Другой способ решения проблемы *Full-Mesh* – это конфедерации или иначе их называют *sub-AS*, под-*AS*. Каждая конфедерация ведёт себя как *AS* – внутри полная связность, снаружи – *IBGP* работает по принципу *EBGP* (с некоторыми оговорками), граничные маршрутизаторы конфедераций, ведут себя как *EBGP*-соседи, должны быть подключены напрямую.



Дополнительные возможности IBGP

- Возможности *Community*
 - С помощью атрибута *Community* можно из своей *AS* управлять поведением маршрутизаторов другой *AS*. Управление своими анонсами в чужой *AS* с помощью *community* поддерживается подавляющим большинством вендоров.
Community – это опциональный передаваемый атрибут (*Optional Transitive*) размером 4 байта. Он представляет из себя запись вида *AA:NN*, где *AA* – двухбайтовый номер *AS*, *NN* – номер коммьюнити (например, *64500:333*).



Дополнительные возможности IBGP

- Возможности *Community*
 - Существует 4 так называемых *Well-Known community*:
Internet – Нет никаких ограничений – передаётся всем.
No-export – Нельзя экспорттировать маршрут в другие AS. При этом за пределы конфедерации передавать их можно.
No-export-subconfed (называется также *Local AS*) – Как *No-export*, только добавляется ограничение и по конфедерациям.
No-advertise – Не передавать этот маршрут никому – только сосед будет знать о нём.