OTUS

Сетевой инженер

STP

otus.ru

Проверить, идет ли запись

Меня хорошо видно && слышно?

Поставьте, пожалуйста “+”, если все хорошо
“-”, если есть проблемы

Тема вебинара

3. Избыточность локальных сетей. Spanning Tree Protocol.

Антон Рожков

Core Network Engineer

anton.a.rozhkov@gmail.com
Telegram: @cskthere

Правила вебинара

Активно участвуем

Обсуждения и офлайн вопросы - в Telegram

Задаем вопрос в чат или голосом

Вопросы вижу в чате,
могу ответить не сразу

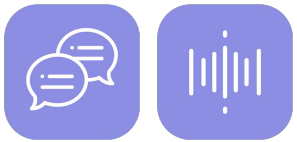
Маршрут вебинара

- Назначение протокола STP
- Принципы работы протокола STP
- Формат BPDU
- Эволюция STP

Назначение протокола STP

Почему L2 петли опасны

Какой заголовок **не является** заголовком фрейма Ethernet?



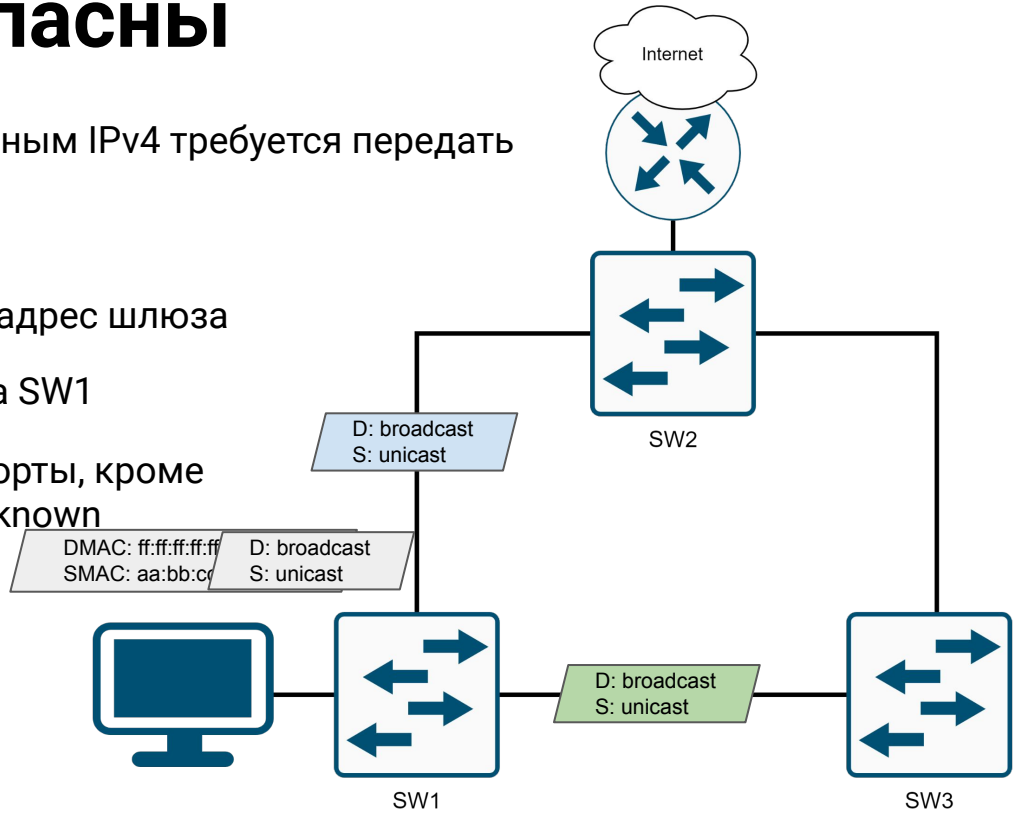
1. Destination MAC address
2. EtherType
3. Source MAC address
4. Time-to-live (TTL)
5. Priority code point

Нет механизма отслеживания времени жизни фреймов в сети ->
фреймы будут передаваться вечно, если для этого будет возможность ->
широковещательный (broadcast) шторм в сети

Почему L2 петли опасны

Что происходит, когда компьютеру с настроенным IPv4 требуется передать данные в Интернет?

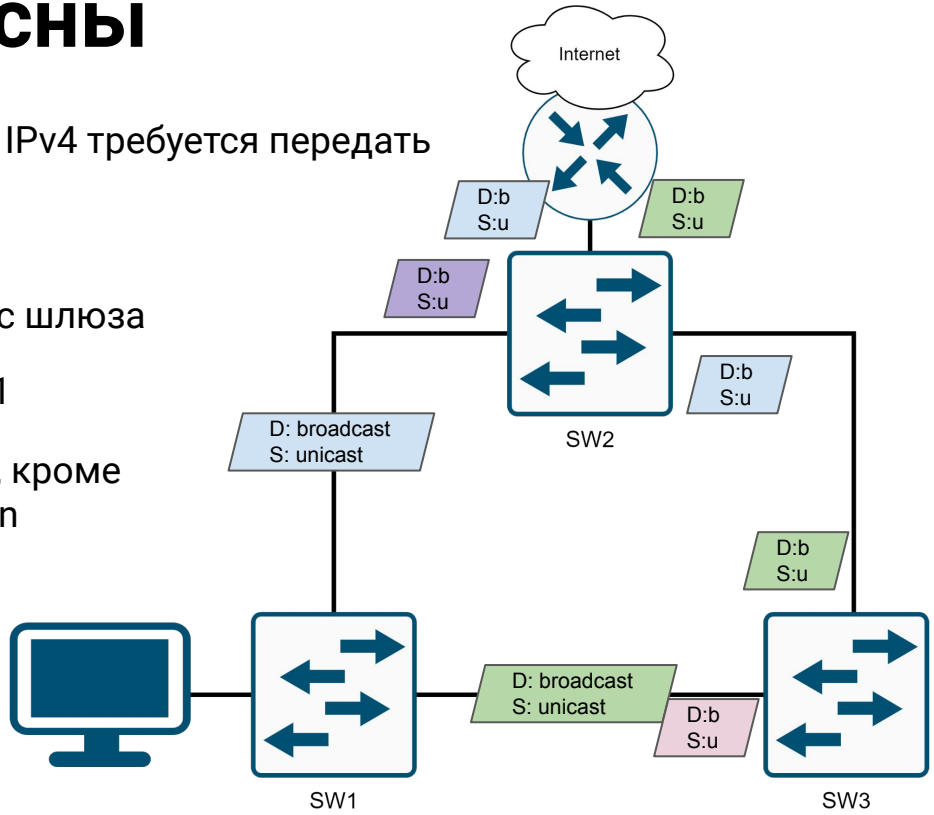
- Отсылка ARP request, чтобы найти MAC адрес шлюза
- Получение фрейма на порт коммутатора SW1
- Широковещательная передача на все порты, кроме того, с которого был получен фрейм (unknown unicast)



Почему L2 петли опасны

Что происходит, когда компьютеру с настроенным IPv4 требуется передать данные в Интернет?

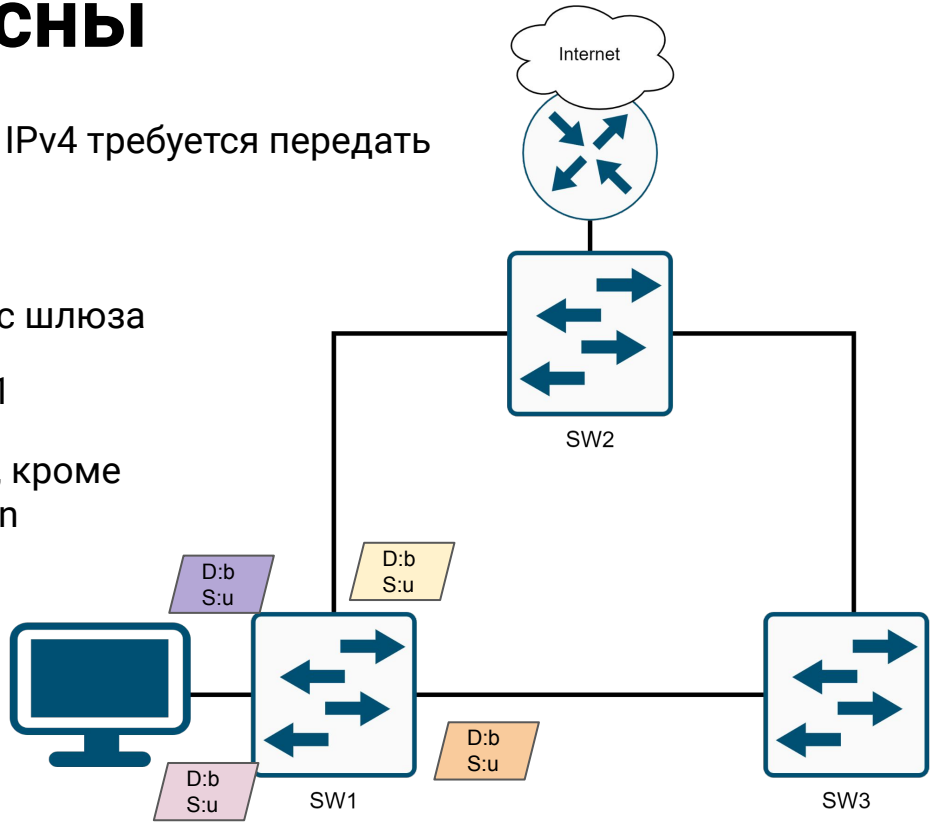
- Отсылка ARP request, чтобы найти MAC адрес шлюза
- Получение фрейма на порт коммутатора SW1
- Широковещательная передача на все порты, кроме того, с которого был получен фрейм (unknown unicast)
- Широковещательная рассылка другими коммутаторами x2



Почему L2 петли опасны

Что происходит, когда компьютеру с настроенным IPv4 требуется передать данные в Интернет?

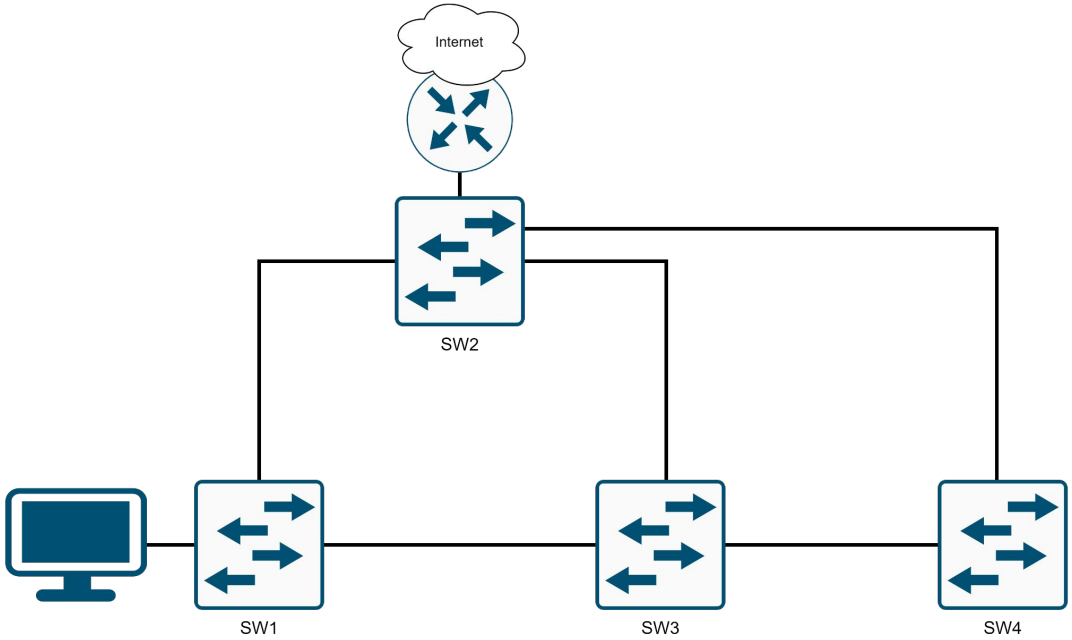
- Отсылка ARP request, чтобы найти MAC адрес шлюза
- Получение фрейма на порт коммутатора SW1
- Широковещательная передача на все порты, кроме того, с которого был получен фрейм (unknown unicast)
- Широковещательная рассылка другими коммутаторами x2
  x3
  x4
  x5



Почему L2 петли опасны

При усложнении сети количество фреймов, передающихся по сети с петлей, увеличивается.

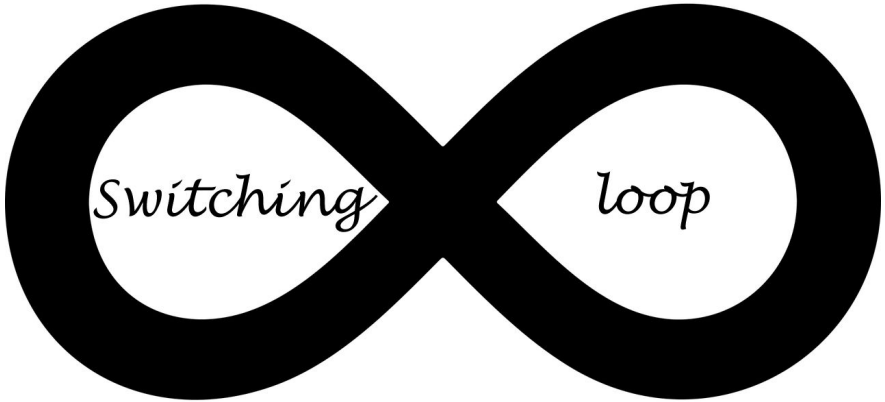
Коммутаторы SW2 и SW3 удваивают количество фреймов.



Почему L2 петли опасны

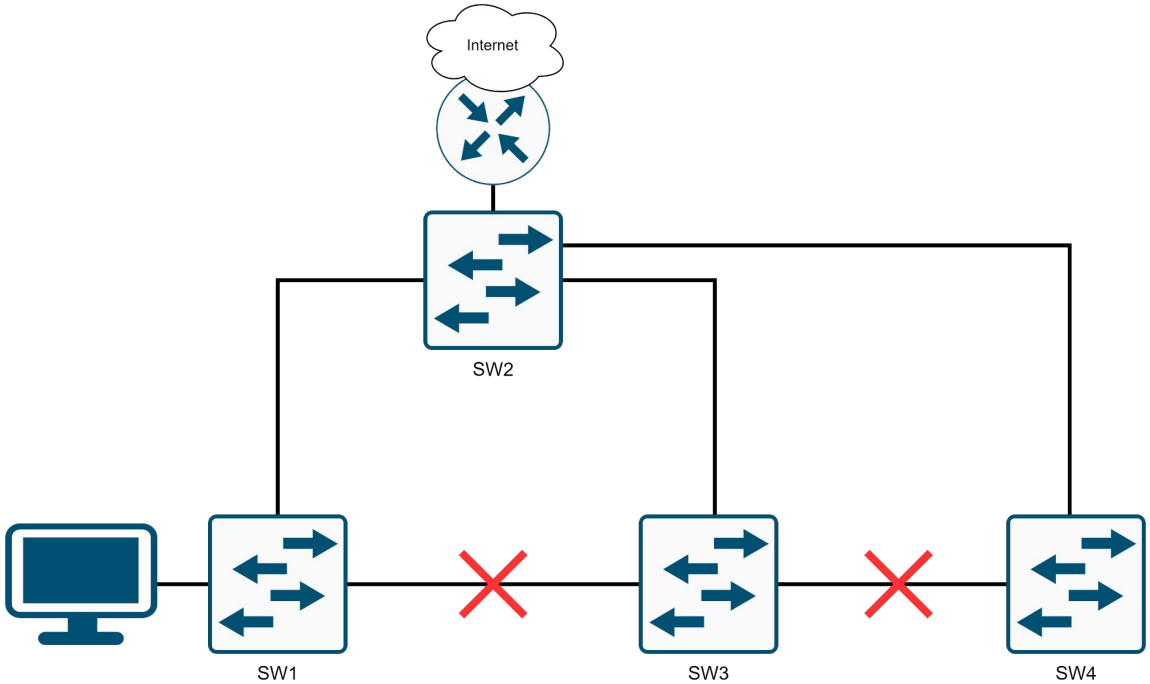
Чем опасен broadcast шторм в сети:

1. Утилизация линков
2. Нагрузка на CPU хостов сети -> вместо полезной нагрузки хост занят обработкой broadcast фреймов
3. Нестабильность таблицы MAC адресов
4. Нагрузка на CPU сетевых устройств (коммутаторы, маршрутизаторы, межсетевые экраны и др.)-> страдают сетевые протоколы



Как избавиться от петли

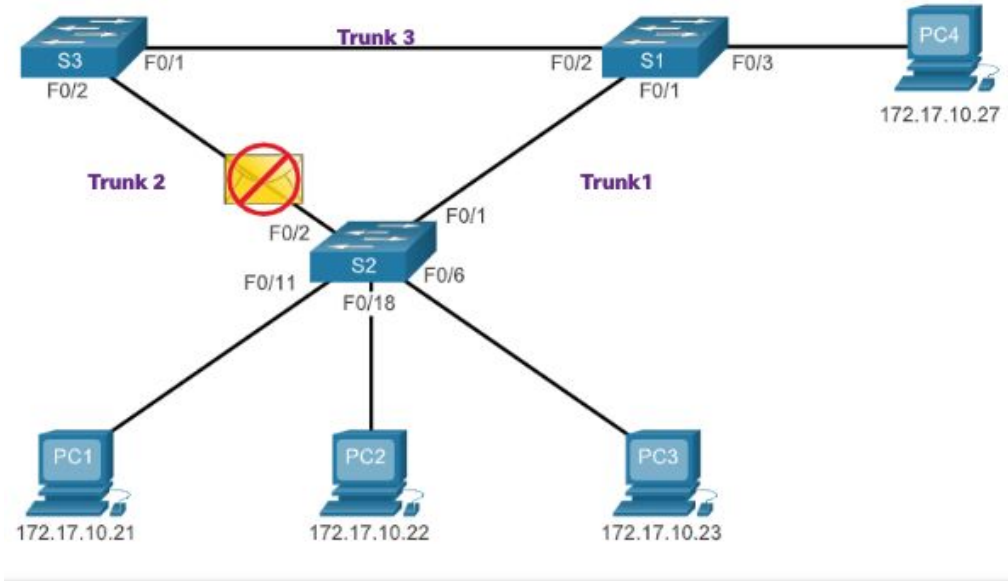
Выключить избыточные линки ¯_(ツ)_/¯



Назначение протокола STP

Spanning Tree Protocol

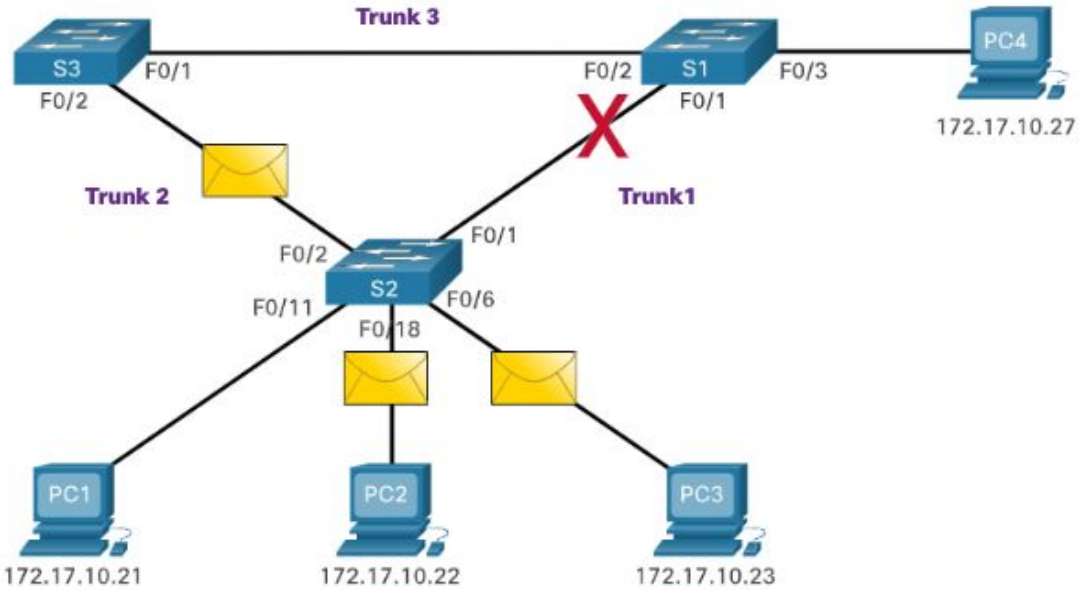
- Протокол связующего дерева (STP) - это сетевой протокол предотвращения петель, который обеспечивает избыточность при создании топологии уровня 2 без петель.
- STP логически блокирует физические петли в сети уровня 2, предотвращая бесконечное хождение кадров в сети.



Назначение протокола STP

Пересчёт STP

- STP компенсирует сбои в сети путем перерасчета и открытия ранее заблокированных портов.



Назначение протокола STP

Проблемы с избыточными каналами коммутатора

- Резервирование путей обеспечивает необходимую доступность множества сетевых сервисов, устраняя вероятность перебоев в работе всех сетевых служб в случае отказа в отдельной точке. При наличии нескольких путей между двумя устройствами и отсутствии реализации протокола spanning-tree возникает петля 2-го уровня. Петли уровня 2 могут привести к нестабильности таблицы MAC-адресов, перегрузке каналов и высокой загрузке ЦП на коммутаторах и конечных устройствах, в результате чего сеть становится непригодной для использования.
- Уровень 2 Ethernet не включает в себя механизм распознавания и устранения бесконечно зацикливающихся кадров. Некоторые протоколы 3-го уровня (IP как наиболее распространенный) используют механизмы времени жизни (TTL), которые ограничивают количество попыток повторной передачи пакетов сетевыми устройствами 3-го уровня. Маршрутизатор уменьшает TTL (Time to Live) в каждом пакете IPv4 и поле Hop Limit в каждом пакете IPv6. Когда эти поля уменьшаются до 0, маршрутизатор отбрасывает пакет. Протокол Ethernet не имеет сопоставимого механизма ограничения числа передач кадра уровня 2 коммутатором. STP был разработан специально в качестве механизма предотвращения петли для Ethernet как протокола уровня 2 модели OSI.

Назначение протокола STP

Петли уровня 2

- Без включения STP могут формироваться петли уровня 2, что приводит к бесконечному циклу широковещательных (broadcast), многоадресных (multicast) и неизвестных одноадресных (unknown unicast) кадров. Это может быстро разрушить сеть.
- При появлении петли возникает возможность постоянного изменения таблицы MAC-адресов на коммутаторе обновлениями из кадров широковещательной рассылки (broadcast), что приводит к нестабильности базы данных MAC-адресов. Это может привести к высокой загрузке ЦП, что приводит коммутатор в нерабочее состояние.
- Неизвестный одноадресный (unknown unicast) кадр с коммутатора формируется, когда у коммутатора нет MAC-адреса назначения в таблице MAC-адресов, и он должен переслать этот кадр со всех своих портов, за исключением входного порта.

Назначение протокола STP

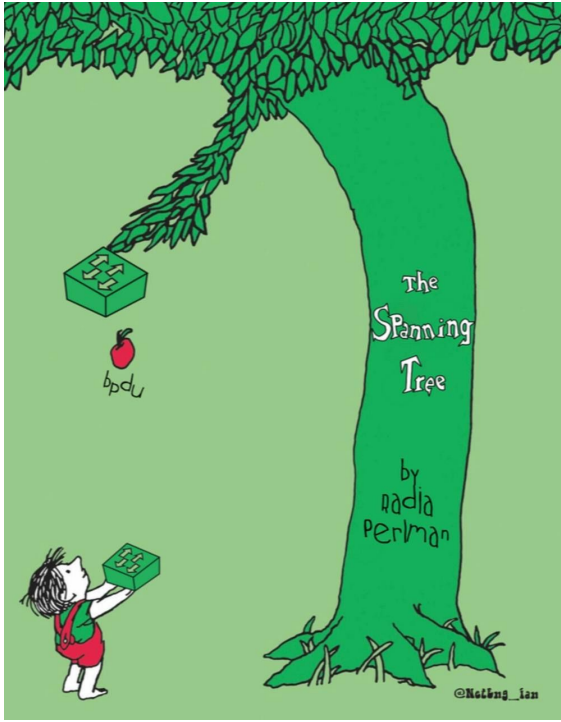
Широковещательный шторм

- Широковещательный шторм - это ненормально большое количество широковещательных передач (broadcast), подавляющих сеть в течение определенного периода времени. Широковещательные штормы могут отключить сеть за считанные секунды, перегружая коммутаторы и конечные устройства. Широковещательные штормы могут быть вызваны аппаратными проблемами, такими как неисправный сетевой адаптер или петля 2-го уровня в сети.
- Широковещательные рассылки уровня 2 в сети, такие как ARP-запросы, очень распространены. Многоадресные рассылки (multicast) второго уровня обычно пересылаются так же, как и широковещательные рассылки коммутатором. Пакеты IPv6 никогда не пересылаются как широковещательная передача уровня 2, ICMPv6 Neighbor Discovery использует многоадресную рассылку (multicast) уровня 2.
- Узел, участвующий в сетевой петле, недоступен для других узлов в сети. Кроме того, вследствие постоянных изменений в таблице MAC-адресов коммутатор не знает, из какого порта следует пересылать кадры одноадресной рассылки (unicast).
- Во избежание подобных проблем в сети с избыточностью, на коммутаторах должны быть включены определённые типы протокола spanning-tree. Протокол spanning-tree по умолчанию включен на коммутаторах Cisco, предотвращая, таким образом, возникновение петель 2-го уровня.

Назначение протокола STP

Алгоритм STP

- Протокол STP основан на алгоритме, изобретенном Радией Перлман (Radia Perlman) во время ее работы в Digital Equipment Corporation и опубликованном в статье 1985 г. «Алгоритм распределенного вычисления протокола связующего дерева в расширенной сети LAN» (An Algorithm for Distributed Computation of a Spanning Tree in an Extended LAN). Ее **алгоритм связующего дерева (STA)** создает топологию без петли, выбрав один корневой мост, где все остальные коммутаторы определяют один путь с наименьшей стоимостью.
- Протокол STP предотвращает возникновение петель за счет настройки беспетлевого пути в сети с использованием портов, стратегически настроенных **на заблокированное состояние**. Коммутаторы, использующие протокол STP, могут компенсировать сбои за счет динамической разблокировки ранее заблокированных портов и разрешения передачи трафика по альтернативным путям.



Назначение протокола STP

Алгоритм STP (продолжение)

Как STA создает топологию без петли?

- **Выбор корневого моста (root bridge)**: этот мост (коммутатор) является **опорной точкой** для всей сети для построения STP.
- **Блокирование резервных путей**: Протокол STP обеспечивает наличие **только одного логического пути** между всеми узлами назначения в сети путем намеренного блокирования резервных путей, которые могли бы вызвать петлю. Порт считается заблокированным, когда заблокирована отправка и прием данных на этот порт.
- **Создание топологии без петли**: заблокированный порт **не пересылает** кадры между двумя коммутаторами. Это создает топологию, в которой каждый коммутатор имеет только один путь к корневому мосту, аналогично ветвям дерева, которые подключаются к корню дерева.
- **Пересчет в случае сбоя соединения**: физические пути по-прежнему используются для обеспечения избыточности, однако **эти пути отключены** в целях предотвращения петель. Если путь потребуется для компенсации неисправности сетевого кабеля или коммутатора, протокол STP повторно рассчитывает пути и снимает блокировку с требуемых портов, чтобы разрешить активацию избыточного пути. Перерасчет STP может происходить в любой момент, когда новый коммутатор или новый межкоммутационный канал добавляется в сеть.

Принципы работы STP

Принципы работы STP

Шаги к беспетельной топологии

- Используя STA, STP строит топологию без петель в четыре этапа:
    1. Выбор корневого моста (Root bridge).
    2. Выбор корневых портов (Root ports).
    3. Выбор избранных (назначенных) портов (Designated ports).
    4. Выбор альтернативных (заблокированных) портов (Alternate and Backup ports).

- При работе STA и STP коммутаторы используют **блоки данных протокола моста (Bridge Protocol Data Unit, BPDU)** для обмена информацией о себе и своих каналах. BPDU используются для выбора корневого моста, корневых портов, назначенных портов и альтернативных портов.

- Каждый BPDU содержит идентификатор **Bridge ID (BID)**, который определяет коммутатор, отправивший BPDU. BID участвует в принятии многих решений STA, включая роли корневого моста и портов.

- Идентификатор BID содержит значение **приоритета, MAC-адрес отправляющего коммутатора и дополнительный расширенный идентификатор системы**. Самое низкое значение BID определяется комбинацией значений в этих трех полях.

Принципы работы STP

Формат Идентификатора моста

Идентификатор моста или Bridge ID (BID) - **8 байт**:

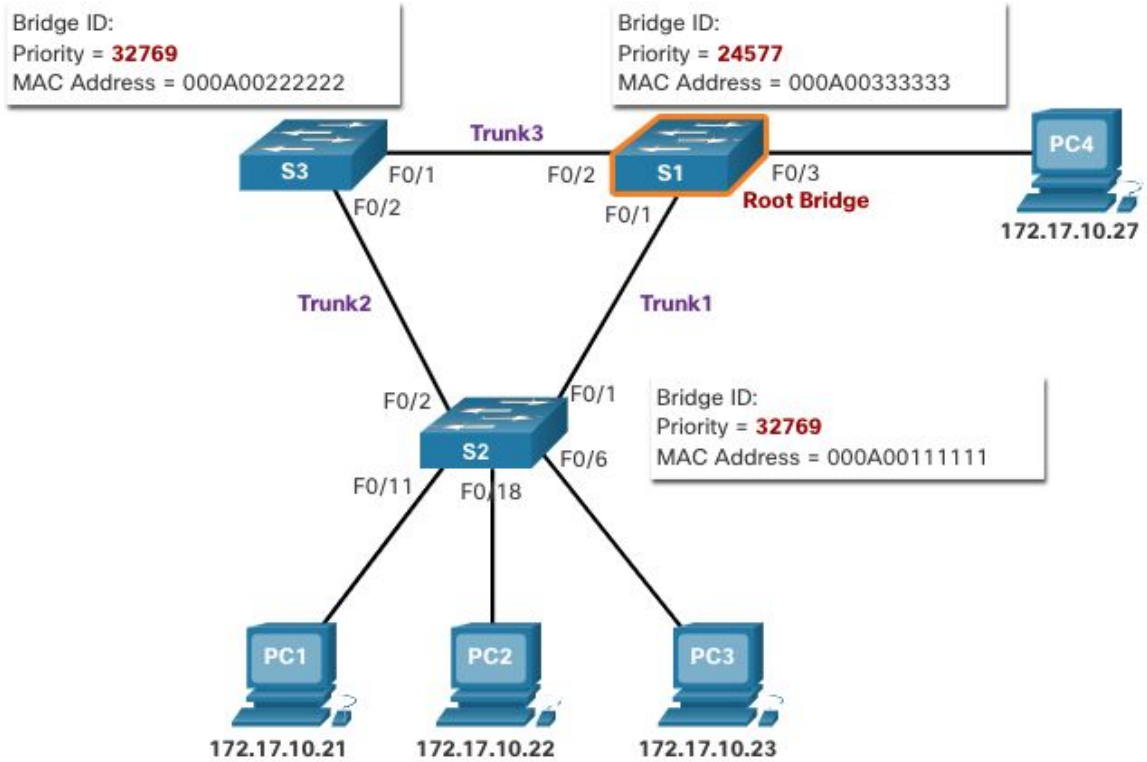
- **Приоритет моста (Bridge Priority, 4 бита)**: Значение приоритета по умолчанию для всех коммутаторов Cisco равно десятичному значению 32768. Значения варьируются в диапазоне от 0 до 61440 с шагом в 4096. Предпочтительнее более низкий приоритет моста . Приоритет моста 0 имеет преимущество по сравнению со всеми остальными значениями приоритета моста.
- **Значение расширенного идентификатора системы (System Extension ID, 12 бит)** — это десятичное значение, добавляемое к значению приоритета моста в BID для определения приоритета и сети VLAN кадра BPDU.
- **MAC-адрес (48 бит)**: Если два коммутатора настроены с одинаковым приоритетом, и у них одинаковый расширенный идентификатор системы, то коммутатор с наименьшим значением MAC-адреса получит меньший идентификатор BID. Используется “системный” MAC (иногда называют “burned-in address, BIA” или “Base MAC” или “Main MAC”, часто печатается на стикерах на коммутаторах и как правило не может быть изменен).

Примечание: Source MAC адрес фрейма BPDU в заголовке Ethernet - это MAC адрес порта, с которого коммутатор послал данный фрейм (не Main MAC).

Принципы работы STP

1. Выбор Root bridge

- STA назначает один из коммутаторов в качестве корневого моста и использует его как точку привязки для расчёта всех путей. Коммутаторы обмениваются BPDU для создания беспетельной топологии, начиная с выбора корневого моста.
- Все коммутаторы в домене широковещательной рассылки участвуют в процессе выбора. После загрузки коммутатора они начинают рассылать кадры BPDU с интервалом в две секунды. Эти кадры BPDU содержат BID передающего коммутатора и BID корневого моста, известный как Root ID или Root Bridge ID (RID).
- Коммутатор с самым низким значением идентификатора моста (BID) становится корневым мостом. Сначала все коммутаторы объявляют себя корневым мостом с собственным BID, установленным в качестве корневого идентификатора. В конце концов коммутаторы узнают через обмен BPDU, какой коммутатор имеет самый низкий BID. Этот коммутатор и будет считаться корневым мостом (Root Bridge).

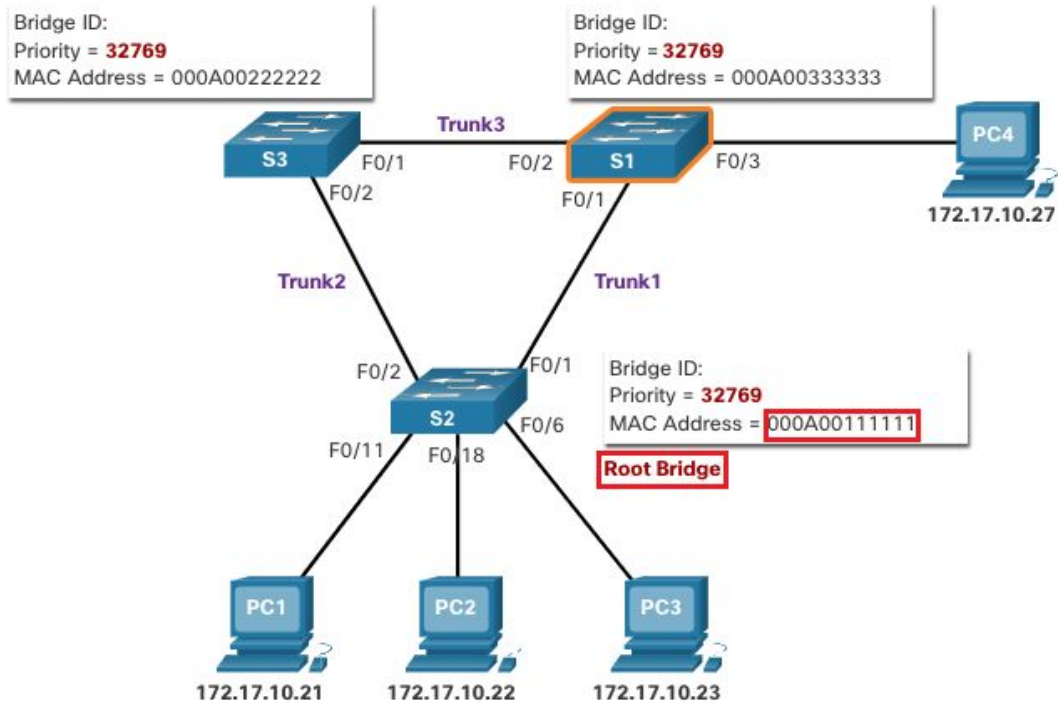


Принципы работы STP

Влияние BID по умолчанию

- Поскольку значение BID по умолчанию равно 32768, два или более коммутаторов могут иметь одинаковый приоритет. В этом сценарии, где приоритеты одинаковы, коммутатор с меньшим MAC-адресом станет корневым мостом. Администратор должен настроить требуемый корневой коммутатор с более низким приоритетом.
- На рисунке все коммутаторы настроены с одинаковым приоритетом 32769. Здесь MAC-адрес становится решающим фактором в отношении того, какой коммутатор становится корневым мостом. MAC-адрес с самым низким шестнадцатеричным значением считается предпочтительным корневым мостом. В этом примере S2 имеет наименьшее значение MAC-адреса и, следовательно, назначается корневым мостом для этого экземпляра протокола spanning-tree.

Примечание: для всех коммутаторов используется значение 32769. Это значение основано на значении приоритета по умолчанию 32768 и назначении сети VLAN 1, связанном с каждым из коммутаторов (32768+1).



Принципы работы STP

Определение стоимости корневого пути (Root path cost)

- Если корневой мост выбран для экземпляра протокола spanning-tree, STA начинает процесс определения оптимальных путей к корневому мосту от всех некорневых коммутаторов в домене широковещательной рассылки (broadcast domain). Информация о пути, известная как стоимость **внутреннего корневого пути** (root path cost), равна сумме стоимости отдельных портов на пути от коммутатора к корневому мосту.
- Когда коммутатор получает блок BPDU, он добавляет стоимость входного порта сегмента для определения своей стоимости для **внутреннего корневого пути**.
- Стоимость портов по умолчанию определяется скоростью работы порта и рассчитывается по формуле: **cost = Reference / Bandwidth**, где Reference - это некое фиксированное значение, а Bandwidth - скорость работы порта (например, 1 Gbps). В таблице показаны расходы на порты по умолчанию, предложенные IEEE. Коммутаторы Cisco по умолчанию используют значения, определенные стандартом IEEE 802.1D, также известные как стоимость короткого пути, как для STP, так и для RSTP.
- Хотя с портами коммутатора связано значение стоимости пути по умолчанию, значение стоимости порта можно настроить. Возможность настройки отдельных портов предоставляет администратору необходимую гибкость при контроле путей протокола spanning-tree к корневому мосту.

Скорость канала	Стоимость STP: IEEE 802.1D-1998	Стоимость RSTP: IEEE 802.1w-2004
10 Гбит/с	2	2 000
1 Гбит/с	4	20 000
100 Мбит/с	19	200 000
10 Мбит/с	100	2 000 000
	ref = 1 Gbit/s	**ref = 20 Tbit/s**

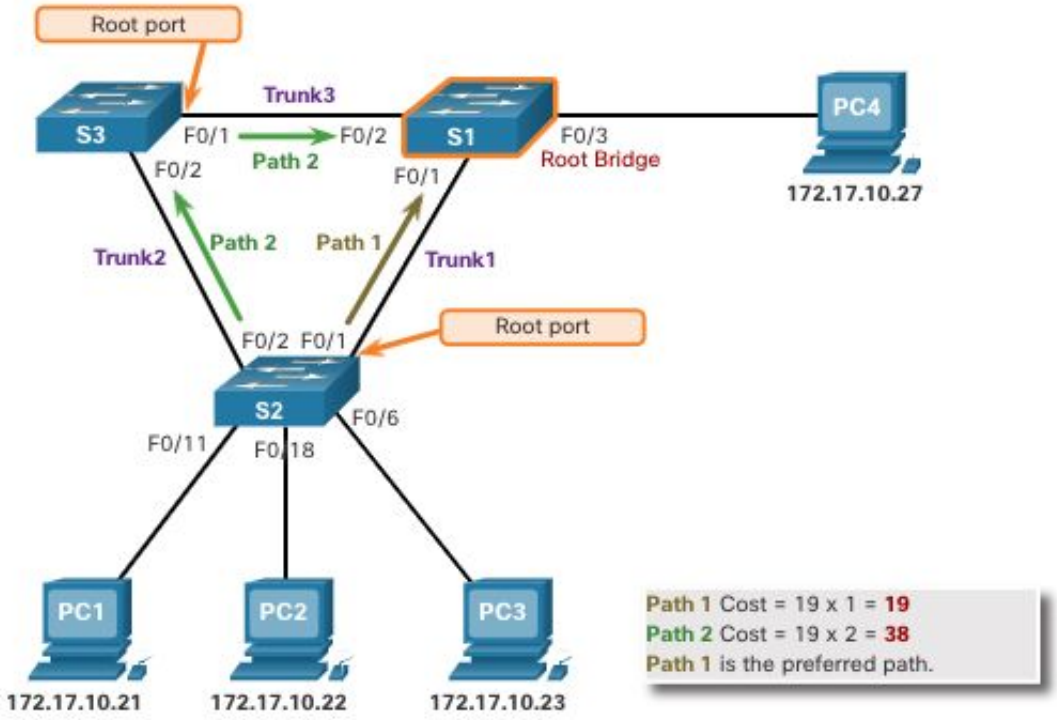
Принципы работы STP

2. Выбор корневых портов (Root ports)

После определения корневого моста для выбора корневого порта используется алгоритм STA. Каждый некорневой коммутатор выбирает **один корневой порт**. Корневые порты — порты коммутатора, ближайшие к корневому мосту с точки зрения общей стоимости маршрута к нему. Эта общая стоимость известна как стоимость пути до корневого моста (Root path cost).

Стоимость **внутреннего корневого пути** равна сумме стоимостей путей от всех портов к корневому мосту, как показано на рисунке. Пути с наименьшей стоимостью становятся предпочтительными, а все остальные избыточные пути блокируются.

В этом примере стоимость внутреннего корневого пути от S2 до корневого моста S1 по пути 1 равна 19, а стоимость внутреннего корневого пути для пути 2 равна 38. Поскольку общая стоимость пути 1 к корневому мосту ниже, именно этот путь является предпочтительным.



Принципы работы STP

3. Выбор назначенных портов (Designated)

Каждый сегмент между двумя коммутаторами будет иметь один назначенный порт (Designated port). Назначенный порт имеет наилучший путь для приема трафика, ведущего к корневому мосту.

То, что не является корневым или назначенным портом, становится альтернативным или заблокированным портом.

Все порты на корневом мосте являются назначенными портами.

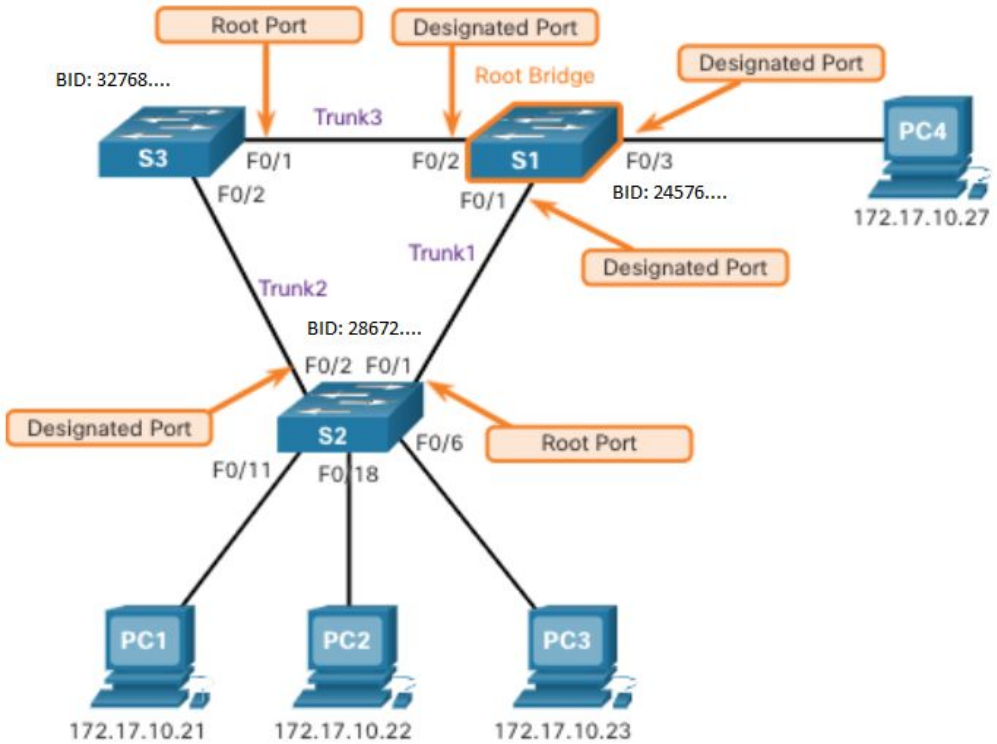
Если на одном конце сегмента находится корневой порт, на другом конце будет назначенный порт.

Все порты, подключенные к конечным устройствам, являются назначенными портами.

На сегментах между двумя коммутаторами, где ни один из коммутаторов не является корневым мостом, порт коммутатора с **наименьшей стоимостью пути к корневому мосту** является назначенным портом.

Если коммутатор имеет **несколько путей равной стоимости** к корневому мосту, коммутатор определяет порт, используя следующие критерии:

- Самое **низкое значение идентификатора моста-отправителя** (lowest Bridge ID), соседнего коммутатора
- Самое **низкое значение идентификатора порта-отправителя** (lowest Port ID), порта соседнего коммутатора, не своего собственного

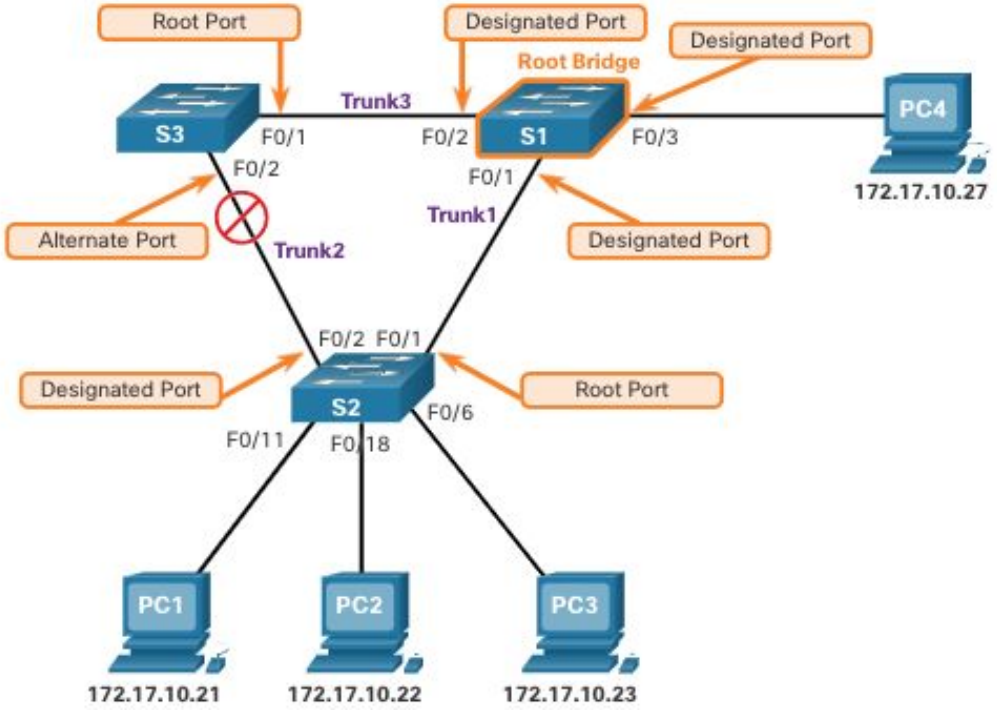


Принципы работы STP

4. Выбор альтернативных (заблокированных) портов

Если порт не является корневым или назначенным портом, он становится альтернативным (или резервным) портом. Альтернативные порты находятся в состоянии отклонения или блокирования для предотвращения петель.

На рисунке STA настроил порт F0/2 на коммутаторе S3 в роли альтернативного порта. Порт F0/2 на S3 находится в блокирующем состоянии и не будет пересылать кадры Ethernet. Все остальные порты между коммутаторами находятся в состоянии пересылки. Он работает как часть STP для предотвращения образования петель.



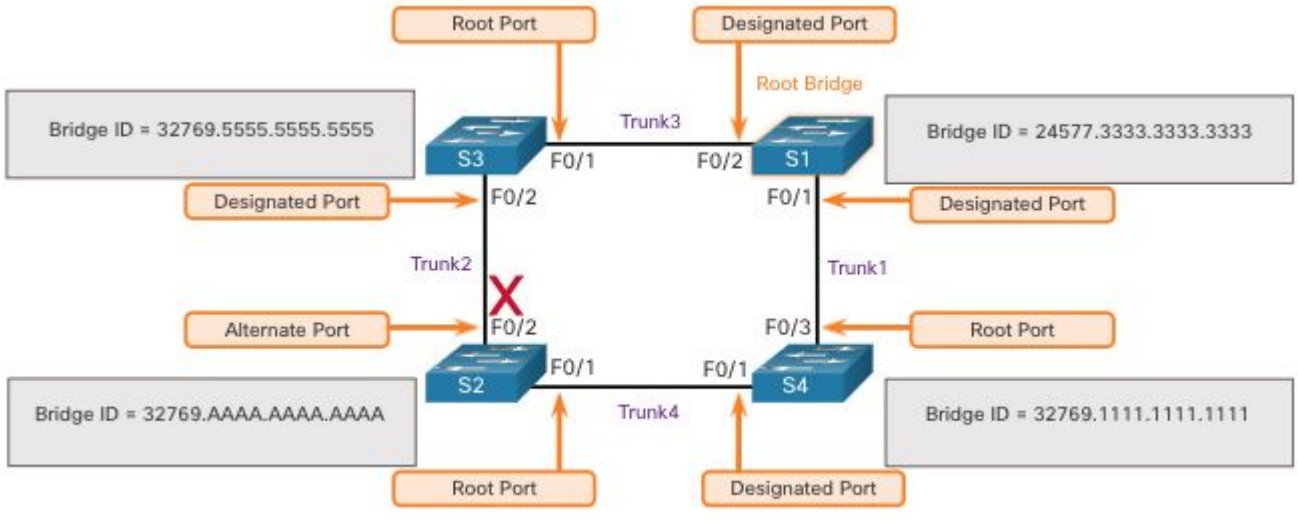
Принципы работы STP
Выбор корневого порта из нескольких путей равной стоимости

Если коммутатор имеет **несколько путей равной стоимости** к корневому мосту, коммутатор определяет порт, используя следующие критерии:

- Самое низкое значение идентификатора моста-отправителя (lowest Bridge ID), т.е. соседнего коммутатора
- Самое низкое значение идентификатора порта-отправителя (lowest Port ID), т.е. порта соседа, а не своего собственного

Принципы работы STP
Выбор корневого порта из нескольких путей равной стоимости

Самый низкий BID отправителя: эта топология имеет четыре коммутатора с коммутатором S1 в качестве корневого моста. Порт F0/1 на коммутаторе S3 и порт F0/3 на коммутаторе S4 были выбраны в качестве корневых портов, поскольку они имеют стоимость корневого пути к корневому мосту для соответствующих коммутаторов. S2 содержит два порта — F0/1 и F0/2 — с путями равной стоимости к корневому мосту. Идентификаторы моста S3 и S4 будут использоваться для дальнейшего определения роли. Это называется BID отправителя. S3 имеет BID 32769.5555.5555.5555, а S4 имеет BID 32769.1111.1111.1111. Поскольку значение BID для S4 меньше, корневым портом будет порт коммутатора S2 F0/1, подключенный к S4.

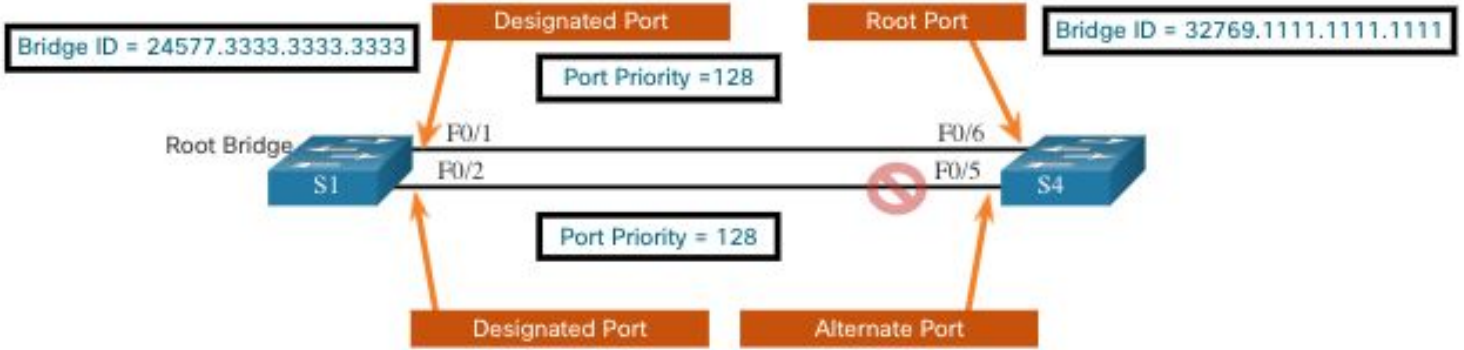


Принципы работы STP
Выбор корневого порта из нескольких путей равной стоимости

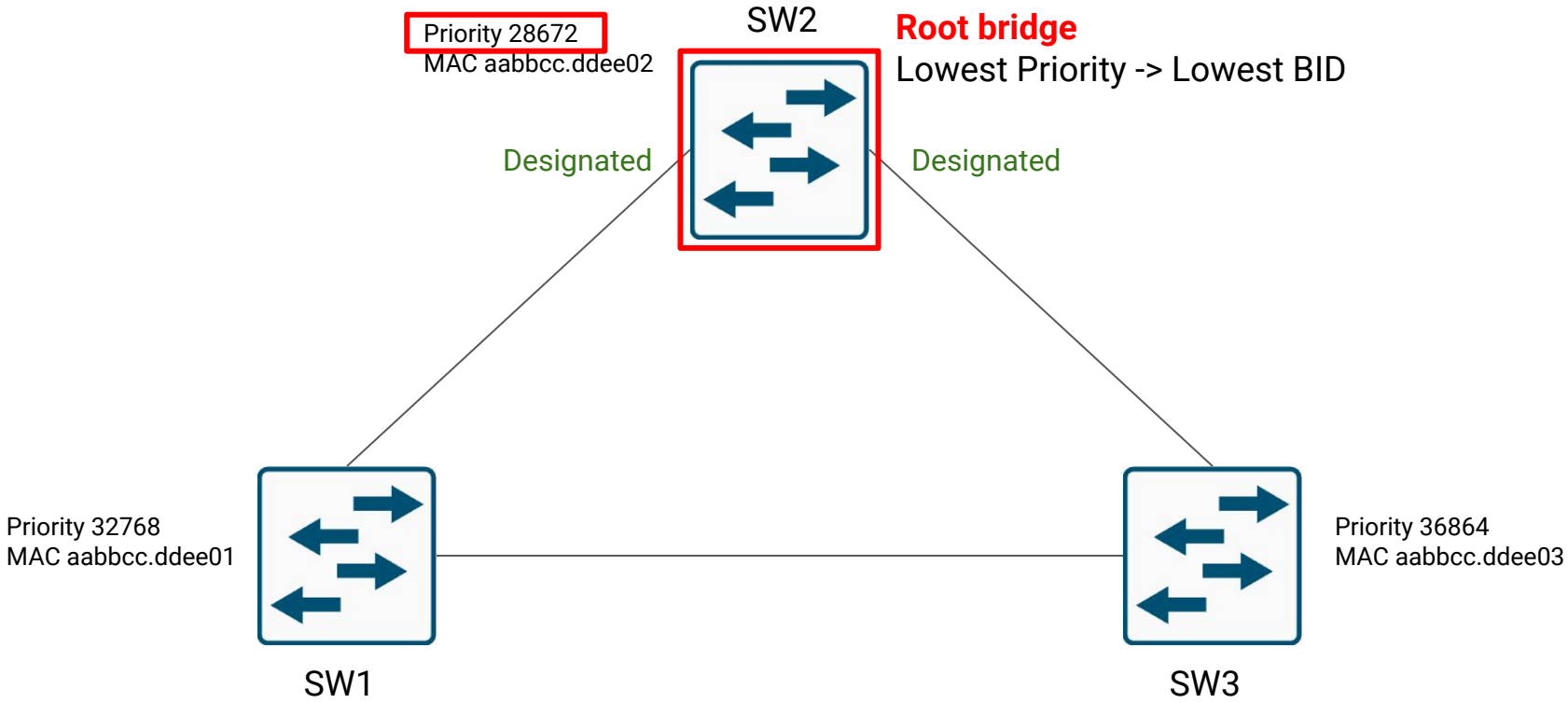
Самый низкий приоритет порта отправителя: Эта топология имеет два коммутатора, которые связаны между собой двумя равноправными путями. S1 является корневым мостом, поэтому оба его порта являются назначенными портами.
S4 имеет два порта с равными по стоимости путями к корневому мосту. Поскольку оба порта подключены к одному коммутатору, BID отправителя (S1) равен. Итак, первый шаг - ничья.
Далее — приоритет порта отправителя (S1). Приоритет порта по умолчанию равен 128, поэтому оба порта S1 имеют одинаковый приоритет порта. Это тоже ничья. Однако, если любой порт на S1 настроен с более низким приоритетом порта, S4 помещал бы свой смежный порт в состояние пересылки. Другой порт на S4 будет блокирующим состоянием.

Принципы работы STP
Выбор корневого порта из нескольких путей равной стоимости

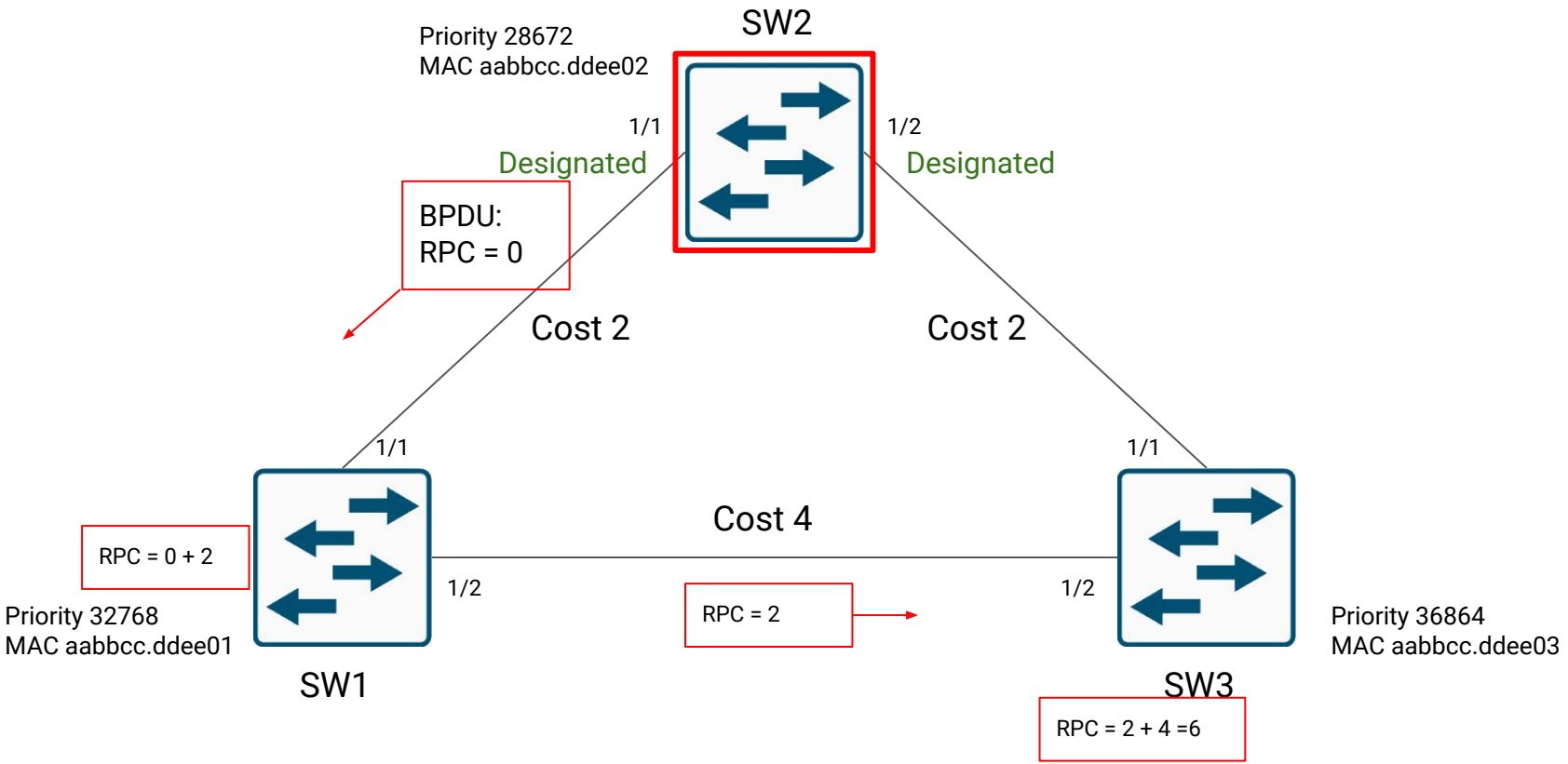
Самый низкий идентификатор порта отправителя: Последний определитель - является самым низким идентификатором порта отправителя. Коммутатор S4 получил BPDU от порта F0/1 и порта F0/2 на S1. Решение основано на идентификаторе порта отправителя, а не на идентификаторе порта получателя. Поскольку идентификатор порта F0/1 на S1 меньше, чем порт F0/2, порт F0/6 коммутатора S4 будет корневым портом. Это порт на S4, который подключен к порту F0/1 на S1.
Порт F0/5 на S4 станет альтернативным портом и будет помещен в состояние блокировки.



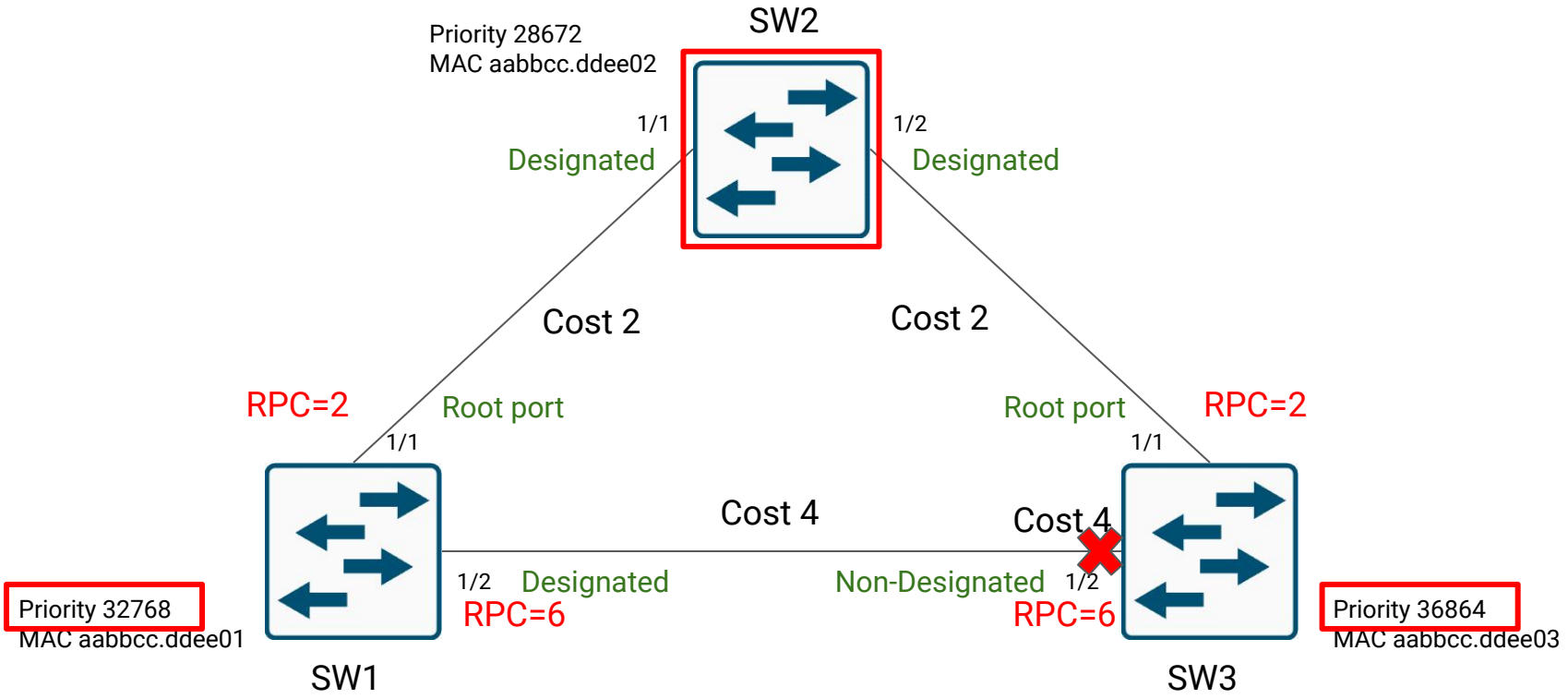
Выбор Root bridge



Выбор Designated port



Выбор Designated port



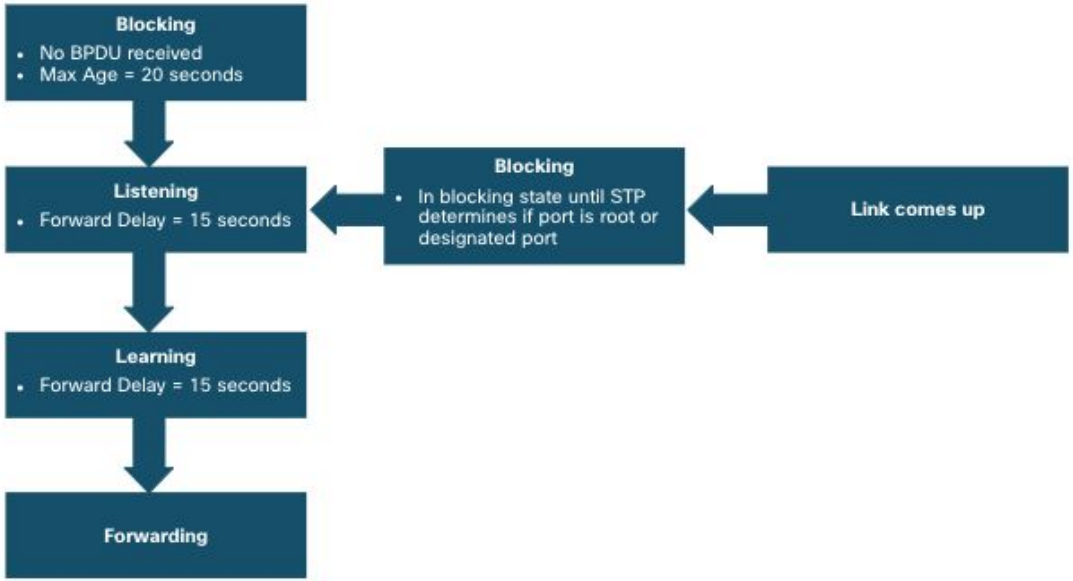
Принципы работы STP
Таймеры STP и состояния портов

Для конвергенции (сходимости) STP требуется три таймера, а именно:

- **Hello Timer** - время приветствия - это интервал между BPDU. По умолчанию это значение равно 2 секундам, но его можно настроить в диапазоне от 1 до 10 секунд.
- **Forward Delay Timer** -Таймер задержки пересылки (Forward Delay Timer) (15 секунд) — время, проводимое в состояниях прослушивания (listening) и обучения (learning). Значение по умолчанию составляет 15 секунд, но может быть изменено на 4-30 секунд.
- **Max Age Timer** - это максимальное время ожидания коммутатора перед попыткой изменения топологии STP. По умолчанию это значение равно 20 секундам, но его можно настроить в диапазоне от 6 до 40 секунд.

Примечание. Время по умолчанию может быть изменено на корневом мосту, который определяет значение этих таймеров для домена STP.

Принципы работы STP
Таймеры STP и состояния портов



Протокол STP упрощает создание логического беспетлевого пути по домену широковещательной рассылки с помощью данных, полученных в процессе обмена кадрами BPDU между соединенными друг с другом коммутаторами.

Если порт коммутатора переходит непосредственно из состояния блокирования в состояние пересылки, не получив информацию о полной топологии в процессе перехода, он может временно создать петлю. По этой причине STP имеет пять состояний портов, четыре из которых являются рабочими состояниями портов, как показано на рисунке. Отключенное состояние считается неработоспособным.

Принципы работы STP
Эксплуатационные данные каждого состояния порта

В таблице приведены рабочие подробности каждого состояния порта.

Состояние порта	BPDU	Таблица MAC-адресов	Пересылка кадров данных
Блокирующий режим	Только получение	Без обновления	Нет
Режим прослушивания	Получение и отправка	Без обновления	Нет
Обучение	Получение и отправка	Обновление таблицы	Нет
Режим пересылки	Получение и отправка	Обновление таблицы	Да
Отключено	Не отправлено или получено	Без обновления	Нет

Состояния портов (states)

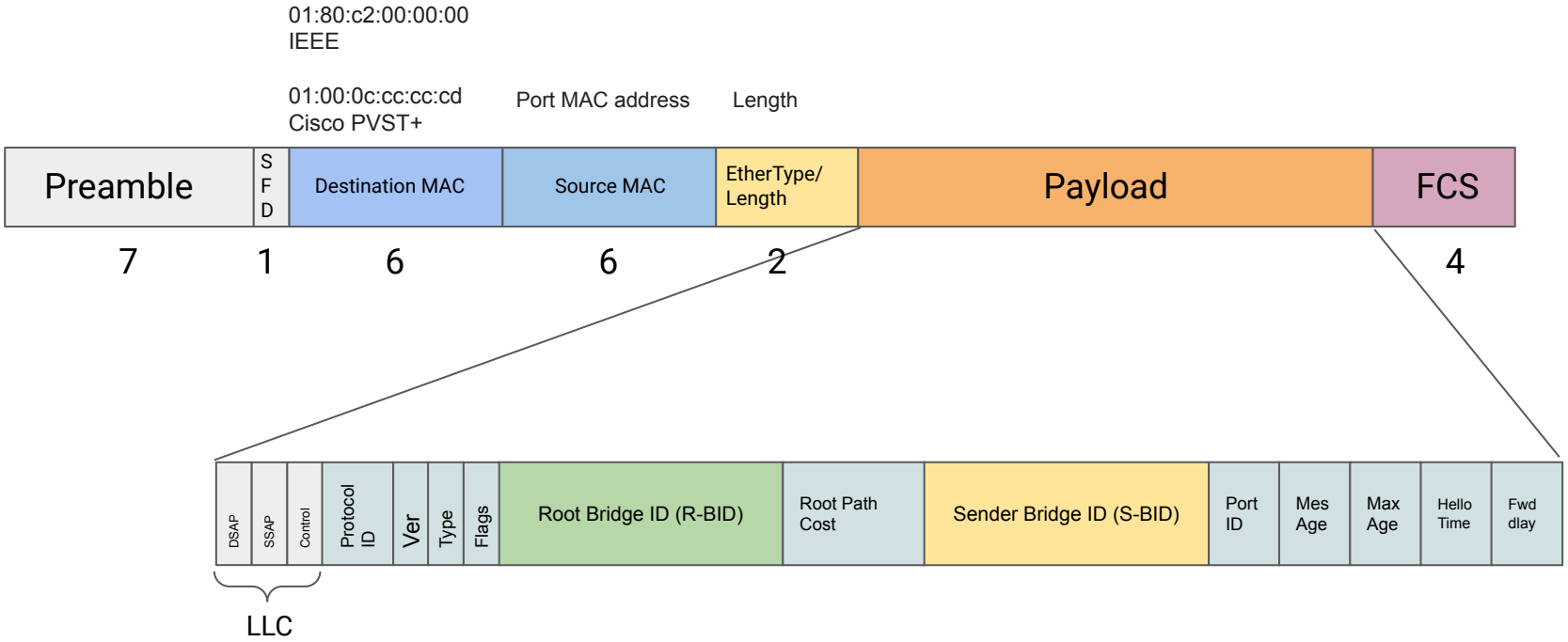
Порт коммутатора с точки зрения STP может находиться в одном из следующих состояний:

- Disabled - порт в состоянии shutdown, трафик не передается (Link down)
- Blocking - порт не передает трафик, принимаются только BPDU, другие фреймы сбрасываются (drop), MAC адреса не изучаются (Learning)
- Listening (переходное состояние) - порт может принимать **и передавать** BPDU, другие фреймы сбрасываются (drop), MAC адреса не изучаются (Learning)
- Learning (переходное состояние) - то же, что в Listening, но включается MAC learning
- Forwarding - рабочее состояние порта (принимается/передается трафик, изучаются MAC адреса)

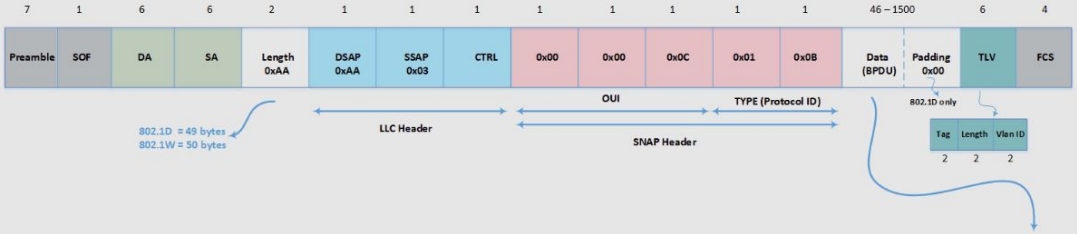
Формат BPDU

Принципы работы STP

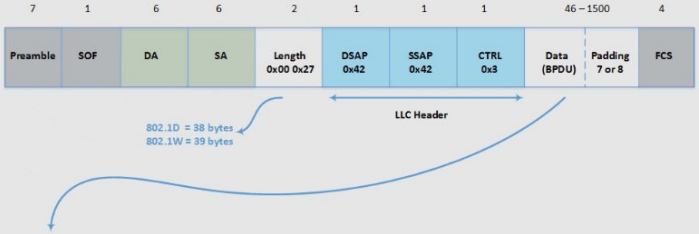
Формат сообщений BPDU



Cisco BPDU (Native vlan)
PVST+/RPVST+
Multicast : 0100.0ccc.cccd
802.3/802.2 SNAP Frame



IEEE BPDU
802.1D/802.1w
Multicast : 0180.C200.0000
802.3/802.2 LLC Frame



BPDU

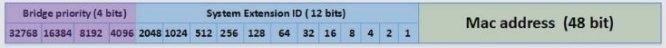
Configuration BPDU	0x00
Topology change BPDU	0x80
RSTP	0x02

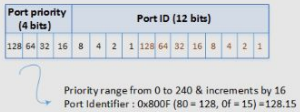
IEEE 802.1W

Bit	Function
0	Topology Change
1	Proposal
2-3	Port Role : 00 Unknown, 01 Alternative, 10 Root Port, 11 Designated port
4	Learning
5	Forwarding
6	Agreement
7	Topology Change Acknowledment

IEEE 802.1D

Bit	Function
0	Topology Change
1	Unused
2	Unused
3	Unused
4	Unused
5	Unused
6	Unused
7	Topology Change Acknowledment





By www.abdorefky.com

Understand Rapid Spanning Tree Protocol (802.1w) - Cisco

Принципы работы STP

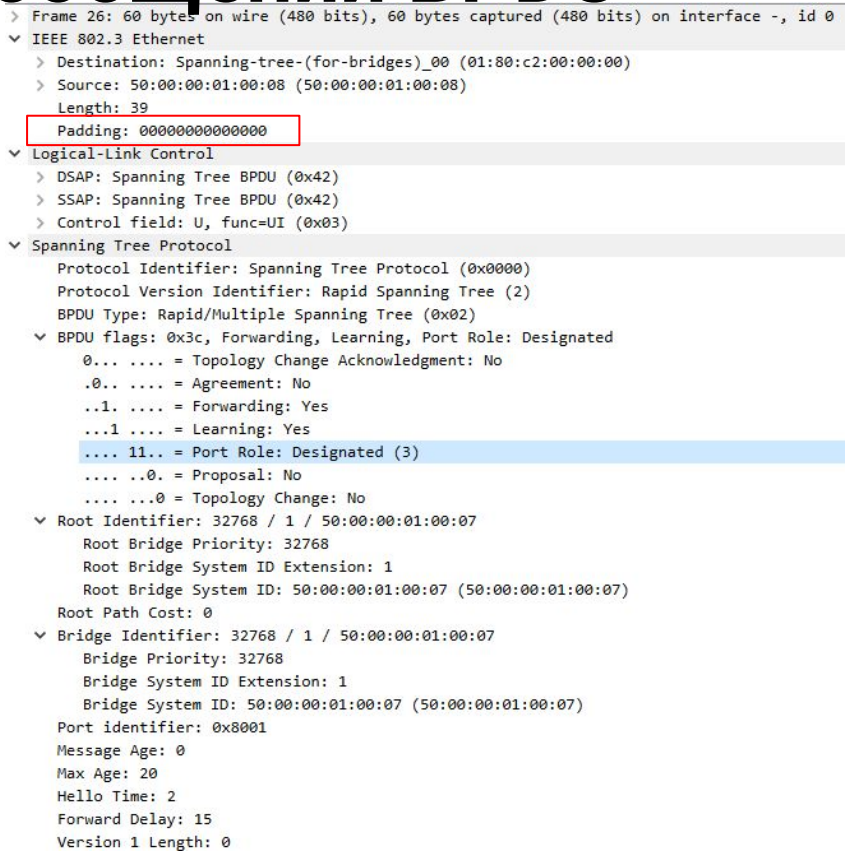
Поля в сообщениях BPDU

- **LLC** - т.к. BPDU инкапсулируются в фреймы Ethernet 802.3, нет отдельного поля EtherType. Для обозначения протокола верхнего уровня используются заголовки
  - LLC в случае стандартного RSTP с DSAP=SSAP=0x42
  - SNAP в случае Cisco PVST с DSAP=SSAP=0xAA, Protocol=0x010B
- **Protocol Identifier** = Spanning Tree Protocol (0x0000)
- **Protocol Version Identifier** = Rapid Spanning Tree (2)
- **BPDU Type** = Rapid/Multiple Spanning Tree (0x02); в STP применялись Configuration и TSN
- **BPDU flags** - флаги. Определены первый (MSB) Topology Change Acknowledgment и последний (LSB) Topology Change. Остальные 6 - Cisco proprietary
- **Root Bridge Id** - BID корневого коммутатора
- **Root Path cost** - стоимость пути до корневого коммутатора
- **Bridge ID** - BID посылающего коммутатора
- **Port Id** (2 байта) - составное поле:
  - Port Priority (4 бита) настраивается в конфигурации порта, по умолчанию b1000
  - Собственно идентификатор порта (12 бит) - может меняться от платформы к платформе, основная задача - должен быть уникальным на устройстве
- **Message Age** - использовался в STP, оставлен для совместимости - время жизни BPDU в сети, STP коммутатор увеличивает поле на 1 при каждой пересылке BPDU, сгенерированного корневым коммутатором. Неприменимо к RSTP, т.к. в нем BPDU не пересылаются, а генерируются каждым коммутатором
- **Max Age** - максимальный Message Age. В STP BPDU считается валидным (Max Age - Message Age) секунд. Неприменимо к RSTP, там BPDU - это keepalive механизм, и время жизни равно 3*Hello time
- **Hello Time** - частота (период) формирования BPDU, по умолчанию 2 с
- **Forward Delay** - использовался в STP для нахождения порта в Listening и Learning состояниях
- **Version 1 Length** = 0 фиксированное значение (указывает длину заголовка BPDU Type = 1, который не определен; определены 2 - RSTP, 0 - STP, 128 - STP TCN)

Примечание: далее могут идти или проприетарные заголовки, например, **Originating VLAN (PVID)** у Cisco RSTP, или MST заголовки

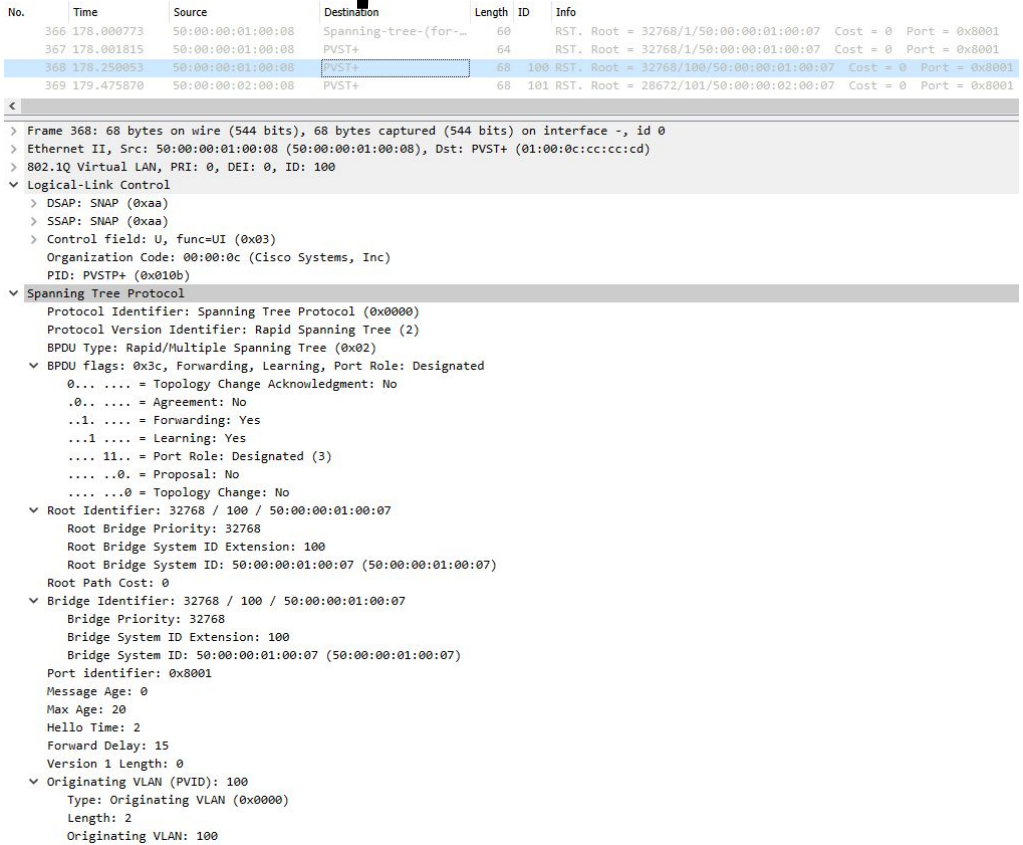
Принципы работы STP

Формат сообщений BPDU



Принципы работы STP

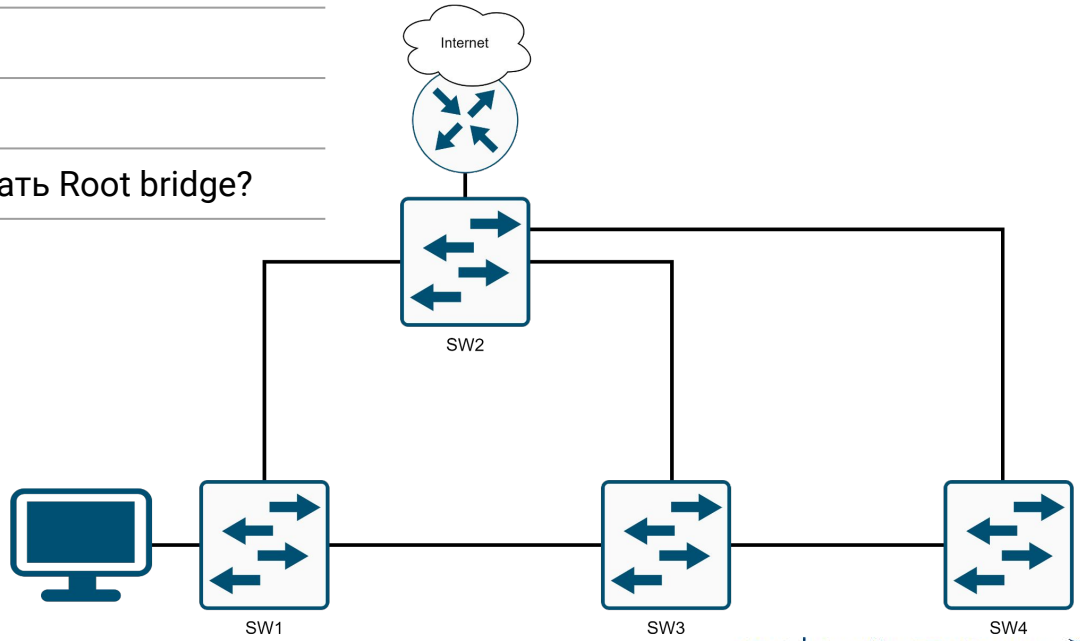
Формат сообщений BPDU



Вопросы для проверки

По пройденному материалу

1. Что такое BPDU?
2. Что такое Root port?
3. Что такое Root bridge?
4. * Где в сети лучше располагать Root bridge?



Эволюция STP

Эволюция STP
Различные версии STP

- Многие специалисты используют термин spanning tree и STP для обозначения различных реализаций протокола spanning-tree, например протокола Rapid Spanning Tree Protocol (RSTP) и протокола Multiple Spanning Tree Protocol (MSTP, MST). Чтобы правильно объяснять принципы протокола spanning-tree, важно понимать, о какой конкретно реализации или стандарте идет речь в данном контексте.
- В новейшей документации IEEE по протоколу связующего дерева (IEEE-802-1D-2004) указано: «STP теперь заменен протоколом Rapid Spanning Tree Protocol (RSTP)». IEEE использует «STP» для обозначения исходной реализации связующего дерева, а «RSTP» — для описания версии связующего дерева, указанной в IEEE-802.1D-2004.
- Так как в этих двух протоколах используется по большей части одинаковая терминология и методы обеспечения пути без петель, основной акцент будет сделан на текущем стандарте и собственных реализациях Cisco для протоколов STP и RSTP.
- Коммутаторы Cisco под управлением IOS 15.0 или более поздней версии по умолчанию запускают PVST+. Эта версия содержит множество спецификаций IEEE 802.1D-2004, таких как альтернативные порты вместо бывших неназначенных портов. Чтобы использовать протокол RSTP, коммутаторы должны быть явно настроены на быстрый режим связующего дерева.

Эволюция STP

Различные версии STP

Варианты STP	Описание
STP	Это исходная версия IEEE 802.1D (802.1D-1998 и более ранняя), которая предотвращает формирование петель в топологии сети с резервными каналами. Общий протокол spanning-tree (CST): предполагает использование только одного экземпляра протокола spanning-tree для всей сети с мостовым соединением независимо от количества сетей VLAN.
PVST+	Per-VLAN Spanning Tree (PVST+): усовершенствованный корпорацией Cisco протокол STP, обеспечивающий отдельный экземпляр связующего дерева 802.1D для каждой сети VLAN, настроенной в сети. Рассматриваемый вариант протокола spanning-tree поддерживает PortFast, UplinkFast, BackboneFast, BPDU guard, BPDU filter, root guard и loop guard.
802.1D-2004	Это обновленная версия стандарта STP, в которую входит IEEE 802.1w.
RSTP	Быстрый протокол STP (RSTP) или IEEE 802.1w: доработанный протокол STP, который обеспечивает более быстрое схождение, чем протокол STP.
Rapid PVST+	Это усовершенствованная технология RSTP Cisco, которая использует PVST+ и предоставляет отдельный экземпляр 802.1w на VLAN. Каждый отдельный экземпляр поддерживает функции PortFast, BPDU guard, BPDU filter, root guard и loop guard.
MSTP	Протокол MSTP (Multiple Spanning Tree Protocol): это стандарт IEEE на базе ранней реализации собственного протокола Cisco с несколькими экземплярами — Multiple Instance STP (MISTP). MSTP сопоставляет несколько сетей VLAN в пределах одного экземпляра протокола spanning-tree.
MST	Реализация Cisco протокола MSTP, которая обеспечивает до 16 экземпляров протокола RSTP и объединяет множество сетей VLAN с идентичной физической и логической топологией в один общий экземпляр RSTP. Каждая реализация поддерживает функции PortFast, BPDU guard, BPDU filter, root guard и loop guard.

Эволюция STP
Принципы работы RSTP

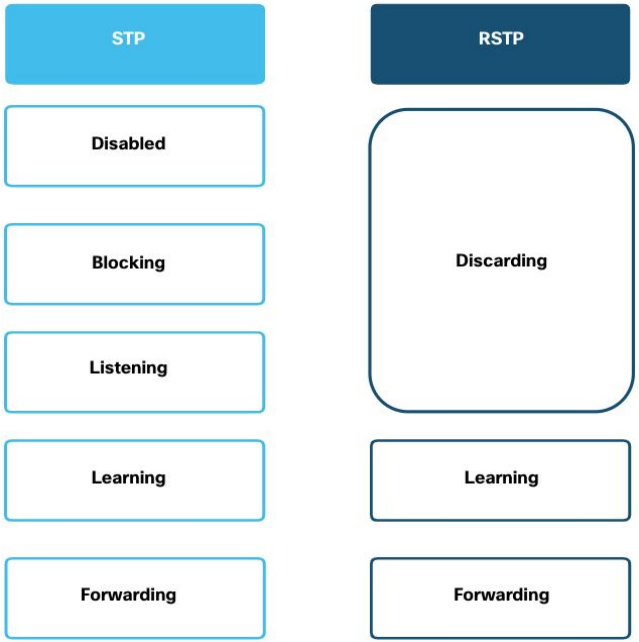
- Протокол RSTP (802.1w) заменяет собой исходный стандарт 802.1D, поддерживая при этом функции обратной совместимости. Терминология, относящаяся к STP 802.1w, остается **в основном той же**, что и для исходного стандарта STP IEEE 802.1D. Большинство параметров остались без изменений. Пользователи, знакомые с исходным стандартом STP, могут легко настроить RSTP. Один и тот же алгоритм связующего дерева используется для STP и RSTP для определения ролей портов и топологии.

- Протокол RSTP ускоряет повторный расчёт протокола spanning-tree в случае изменения топологии сети 2-го уровня. В правильно настроенной сети RSTP может достичь состояния сходимости гораздо быстрее, иногда всего за несколько сот миллисекунд. Если порт настроен альтернативным или резервным, он может немедленно перейти в состояние пересылки без ожидания сходимости сети.

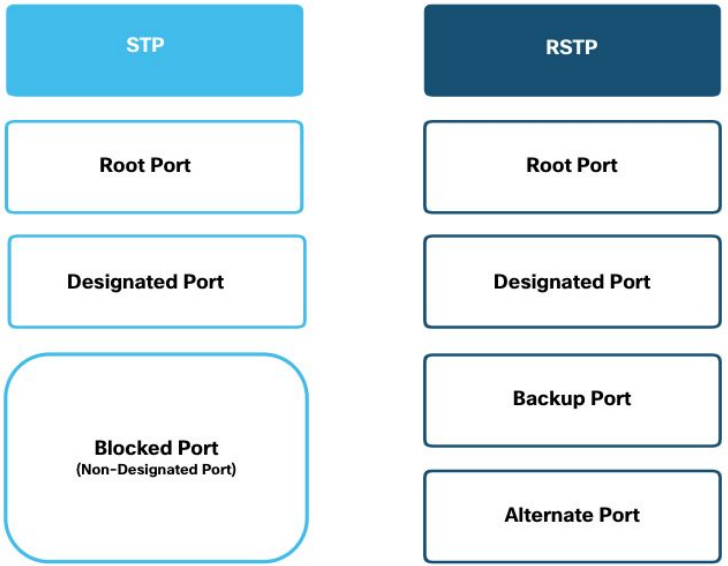
Примечание: Rapid PVST+ представляет собой реализацию RSTP Cisco на основе отдельных VLAN. Для каждой VLAN запускается независимый экземпляр RSTP.

Эволюция STP
RSTP состояния и роли портов

Существует только три состояния порта, которые соответствуют трем возможным рабочим состояниям STP. Состояние отключения, блокировки и прослушивания 802.1D объединяются в уникальное состояние отказа 802.1w.



Корневые порты и назначенные порты одинаковы для STP и RSTP. Тем не менее существует две роли порта RSTP, которые соответствуют состоянию блокировки STP. В STP заблокированный порт определяется как не являющийся назначенным или корневым портом. Для этой цели RSTP имеет две роли портов.

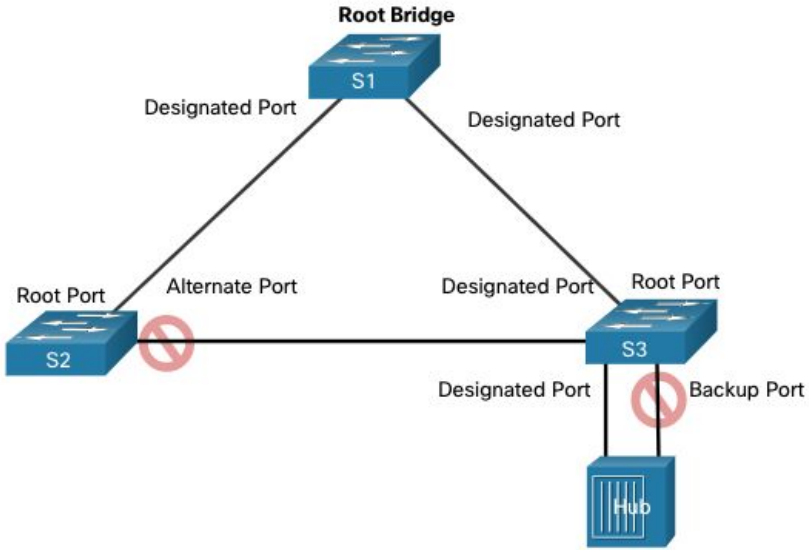


Эволюция STP

RSTP состояния и роли портов

Альтернативный порт имеет альтернативный путь к корневому мосту. Резервный порт является резервным к общей среде, в такой как концентратор (hub).

Резервный порт менее распространен, поскольку концентраторы в настоящее время считаются устаревшими устройствами.



Эволюция STP
Edge ports (PortFast) и BPDU Guard

- Когда устройство подключено к порту коммутатора или когда коммутатор включается, порт коммутатора проходит как прослушивание (Listening) (в случае STP), так и обучение (Learning), каждый раз ожидая истечения срока действия таймера задержки вперед (Forward delay timer). Эта **задержка составляет 15 секунд** для каждого состояния в общей сложности 30 секунд. Это может вызвать проблему для DHCP-клиентов, пытающихся обнаружить DHCP-сервер, поскольку для процесса DHCP может истечь время ожидания. В результате клиент IPv4 не получит действительный адрес IPv4.

- Когда порт коммутатора настроен с помощью PortFast, этот порт переходит из состояния блокировки **в состояние пересылки немедленно**, избегая 30 (15) секундной задержки. Можно использовать PortFast для портов доступа, чтобы устройства, подключенные к этим портам, могли немедленно получить доступ к сети. PortFast следует использовать только для **портов доступа**. Если функция PortFast включена на порте, подключенном к другому коммутатору, возникнет риск возникновения петли протокола spanning-tree.

- Порт коммутатора с включенной функцией PortFast никогда не должен получать BPDU, поскольку это указывает на то, что к порту подключен коммутатор , что может вызвать петлю. Коммутаторы Cisco поддерживают функцию **BPDU guard**. Когда функция BPDU guard включена, при получении блока BPDU она переводит порт в состояние **errdisabled** (error-disabled — отключение из-за ошибки). Это защищает от потенциальных петель, эффективно отключив порт. Администратор должен вручную вернуть интерфейс в эксплуатацию.

Эволюция STP
Link types

- Наряду с Edge ports, в RSTP очень важным нововведением является тип линка между устройствами
- Коммутатор пытается “угадать” какого типа линк используется:
  - Если линк full duplex, то линк - point-to-point
  - Half-duplex - shared
- Можно задать вручную - `(config-if)#spanning-tree link-type [point-to-point|shared]`
- Коммутатор использует тип линка и порта (edge port) в механизмах ускорения сходимости, поэтому важно настраивать эти типы правильно и полно
- Подробности работы механизмов RSTP описаны в статье Understand Rapid Spanning Tree Protocol (802.1w) - Cisco

Эволюция STP
Некоторые отличия STP и RSTP

Операция	**STP (IEEE 802.1D-1998)**	**RSTP (IEEE 802.1D-2004, 802.1w)**
Формирование BPDU	Только Root, остальные только пересылают BPDU, меняя S-BID, Root Path Cost, Message Age, PortID	Каждый коммутатор формирует и рассылает свои BPDU, т.е. BPDU используются как механизм keepalive
Обнаружение изменений	• любое изменение вызывает пересчет топологии • Root коммутатор распространяет информацию по всему домену • все порты проходят все этапы	• изменение топологии только при изменениях на non-Edge ports, причем только если порт переходит в non-Blocking состояние • распространяются всеми коммутаторами • затрагивает только non-Edge порты
Ускорение сходимости	Нет	Edge ports (PortFast), Link types
Фильтрация BPDU	Нет	BPDU Filter, BPDU Guard
Дополнительная защита	Нет	UplinkFast, BackboneFast

Understand Rapid Spanning Tree Protocol (802.1w) - Cisco

Эволюция STP
Альтернативы STP

- С годами организациям требовалась повышенная отказоустойчивость и доступность локальной сети. Сетевые сети Ethernet перешли от нескольких взаимосвязанных коммутаторов, подключенных к одному маршрутизатору, к сложной иерархической структуре сети, включающей коммутаторы доступа, распределения и основного уровня.

- В зависимости от реализации уровень 2 может включать не только уровень доступа, но и распределение или даже уровни ядра. Эти топологии могут включать **сотни коммутаторов с сотнями или даже тысячами VLAN**. STP адаптировалась к дополнительной избыточности и сложности благодаря усовершенствованиям, как часть RSTP и MSTP.

- Важным аспектом проектирования сети является **быстрая и предсказуемая сходимость** при сбое или изменении топологии. Связующее дерево **не обеспечивает такую же эффективность и предсказуемость**, которая обеспечивается протоколами маршрутизации на **уровне 3**.

- Маршрутизация уровня 3 позволяет создавать избыточные пути и петли в топологии **без блокировки портов**. По этой причине некоторые среды переходят на уровень 3 везде, за исключением тех случаев, когда устройства подключаются к коммутатору уровня доступа. Другими словами, соединения между коммутаторами уровня доступа и коммутаторами распределения будут иметь уровень 3, а не уровень 2.

Эволюция STP
Альтернативы STP

- TRILL (TRansparent Interconnection of Lots of Links)
- SPB (Shortest Path Bridging)
- L3 :)

Key takeaways

Основные тезисы

- STP - протокол, целью которого является обнаружение и блокирование **избыточных L2** линков для получения **беспетлевой (древовидной)** топологии

- Термины:
  - BPDU (Bridge Protocol Data Unit) - сообщения протокола STP, рассылаются каждый Hello интервал
  - Root bridge (корневой коммутатор) - вершина дерева
  - Cost, root path cost - “стоимость” линка (больше - хуже), cost = reference/bandwidth или вручную
  - Root port - порт на non-root bridge с наименьшей стоимостью до root bridge
  - Designated port - порт, через который фреймы могут быть доставлены к root bridge, подключен в сегмент, где нет root bridge
  - Alternate port - все остальные (не root и не designated) порты, не передают трафик, не учат MAC адреса, но принимают BPDU

- Сходимость:
  1. Выбор root bridge - коммутатор с минимальным BID
  2. Выбор root port:
     a. Минимальный root path cost
     b. Минимальный BID соседа
     c. Минимальный Port ID соседа
  3. Выбор Designated port:
     a. Минимальный root path cost (если у коммутатора “лучшая” стоимость доставки до root bridge в сегменте, т.е. остальные коммутаторы в сегменте присылают “худшие” BPDU)
     b. Минимальный BID соседа
     c. Минимальный Port ID соседа
  4. Выбор Alternate port - все остальные

- Современное оборудование использует RSTP. Важно настраивать:
  - Edge ports
  - Link types
  - Таймеры Hello
  - Вендорские улучшения/фичи

Домашнее задание