Избыточность LAN

Избыточность сети — ключ к обеспечению надёжности сети. Избыточные маршруты обеспечиваются за счёт нескольких физических каналов между устройствами. Таким образом, сеть может продолжать работу даже в случае сбоя одного канала или порта. Также по избыточным каналам можно распределить нагрузку трафика, что позволяет увеличить емкость.

Во избежание возникновения петель 2 уровня требуется управление несколькими маршрутами. Выбираются оптимальные маршруты, и альтернативный маршрут должен быть незамедлительно доступен в случае сбоя основного маршрута. Протоколы STP используются для управления избыточностью 2 уровня.

Избыточные устройства, например, многоуровневые коммутаторы или маршрутизаторы, предоставляют клиентам возможность использования альтернативного шлюза по умолчанию в случае сбоя основного шлюза по умолчанию. Таким образом клиент сможет использовать несколько путей к нескольким возможным шлюзам по умолчанию.

Протоколы обеспечения избыточности на первом хопе (FHRP) используются для управления назначением клиенту шлюза по умолчанию, а также для предоставления возможности использования альтернативного шлюза по умолчанию в случае сбоя основного шлюза по умолчанию.

Понятия протокола spanning-tree

Предназначение протокола spanning-tree

Избыточность 1 и 2 уровней модели OSI

Трехуровневая иерархическая модель сети, которая использует уровни ядра, распределения и доступа с избыточностью, призвана устранить единую точку отказа в сети. Использование нескольких физически подключенных каналов между коммутаторами обеспечивает физическую избыточность в коммутируемой сети.

Это повышает надёжность и доступность сети. Наличие альтернативных физических каналов для передачи данных по сети позволяет пользователям получить доступ к сетевым ресурсам даже в случае сбоя одного из каналов.

Вставить видео

Данная анимация, демонстрирует на избыточность и демонстрирует следующее:

1. PC1 взаимодействует с PC4 через избыточную топологию сети.
2. Когда в сетевом канале между S1 и S2 происходит сбой, путь между PC1 и PC4 автоматически корректируется, чтобы компенсировать сбой.
3. Если сетевое соединение между S1 и S2 восстановлено, путь повторно корректируется для маршрутизации трафика непосредственно от S2 к S1 для его доставки на PC4.

Для многих организаций доступность сети является важнейшим фактором обеспечения соответствия требованиям бизнеса. Таким образом, модель инфраструктуры сети является критически важным для бизнеса компонентом.

Избыточность маршрута предоставляет решение, обеспечивающее необходимую доступность нескольких сетевых служб за счёт устранения потенциальной единой точки отказа.

Важной частью иерархической архитектуры является избыточность, использование которой позволяет предотвратить перебои в обслуживании оконечных пользователей. Для работы избыточных сетей требуются физические маршруты, однако и логическая избыточность также должна быть частью архитектуры. Тем не менее, избыточные маршруты в коммутируемой сети Ethernet могут привести к возникновению физических и логических петель 2 уровня.

Вследствие работы коммутаторов, особенно в процессе получения данных и пересылки, могут возникать логические петли 2 уровня. При наличии нескольких путей между двумя устройствами и отсутствии реализации протокола spanning-tree возникает петля 2 уровня. Как показано на рисунке (Слайд №7), петля 2 уровня, как правило, приводит к трем проблемам.

- **Нестабильность базы данных MAC-адресов**. Нестабильность содержимого таблицы MAC-адресов возникает из-за создания копий одного кадра, полученных на разных портах коммутатора. Если коммутатор использует все ресурсы для преодоления последствий нестабильности таблицы MAC-адресов, эффективность пересылки данных может быть снижена.

- **Широковещательные штормы**. Без использования надлежащего процесса предотвращения петель каждый из коммутаторов может бесконечно выполнять широковещательную рассылку. Обычно такую ситуацию называют широковещательным штормом.

- **Множественная передача кадров**. На станции назначения могут доставляться несколько копий одноадресных кадров. Многие протоколы предполагают получение только одной копии каждого передаваемого блока. Прием нескольких копий одного кадра может привести к неустранимым ошибкам.

Проблемы с избыточностью 1 уровня. Нестабильность базы данных MAC – адресов.

В отличие от IP-пакетов, кадры Ethernet не содержат атрибут «время жизни» (TTL). Как результат, если не используется механизм блокирования постоянного распространения этих кадров в коммутируемой сети, кадры продолжают распространяться между коммутаторами бесконечно или до тех пор, пока не произойдет сбой канала, в результате чего петля будет прервана.

Такое постоянное распространение между коммутаторами может привести к нестабильности базы данных MAC-адресов. Это может произойти вследствие пересылки широковещательных кадров.

Широковещательные кадры пересылаются из всех портов коммутатора, за исключением исходного входного порта. Это гарантирует, что все устройства в домене широковещательной рассылки могут получить кадр.

При наличии нескольких путей для пересылки кадров может возникнуть бесконечная петля. В случае возникновении петли таблица MAC-адресов на коммутаторе может постоянно изменяться за счёт обновлений от широковещательных кадров, что приводит к нестабильности базы данных MAC-адресов.

Вставить видео

Содержание анимации:

1. PC1 отправляет широковещательный кадр на S2. S2 принимает широковещательный кадр на интерфейс F0/11. Когда S2 принимает широковещательный кадр, он обновляет свою таблицу MAC-адресов, чтобы зарегистрировать доступность PC1 на порте F0/11.
2. Поскольку этот кадр — широковещательный, S2 пересылает кадр из всех портов, включая Магистраль 1 и Магистраль 2. Когда широковещательный кадр поступает на S3 и S1, их таблицы MAC-адресов обновляются относительно PC1, который доступен на порту F0/1 на S1 и на порту F0/2 на S3.
3. Поскольку этот кадр является широковещательным, S3 и S1 пересылают кадр из всех портов, за исключением исходного входного порта. S3 отправляет широковещательный кадр с PC1 на S1. S1 отправляет широковещательный кадр с PC1 на S3. Все коммутаторы обновляют свою таблицу MAC-адресов с учетом неправильного порта PC1.
4. Все коммутаторы снова пересылают широковещательный кадр из всех портов, за исключением входного порта. Это приводит к тому, что оба коммутатора пересылают кадр на S2.
5. Когда S2 получает широковещательные кадры от S3 и S1, таблица MAC-адресов снова обновляется, в этот раз с учетом последней записи, полученной от двух других коммутаторов.

Этот процесс повторяется до тех пор, пока петля не будет прервана путем физического отключения соединений, вызывающих ее, или отключения питания одного из коммутаторов в петле. При этом создается высокая нагрузка на ЦП на всех коммутаторах, участвующих в петле.

Поскольку между всеми коммутаторами в петле постоянно передаются одни и те же кадры, ЦП коммутатора приходится обрабатывать большой объём данных. При этом снижается производительность коммутатора при поступлении допустимого трафика.

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. В вышеуказанном примере для PC1 перечислены неправильные порты.

Любой кадр одноадресной рассылки, предназначенный для PC1, участвует в петле, как и кадры широковещательной рассылки. Из-за возрастающего числа кадров, циклически распространяемых в сети, постепенно создается широковещательный шторм.

Проблемы с избыточностью 1 уровня. Широковещательный шторм

Широковещательный шторм возникает в случае, когда в петлю на 2 уровне попадает столько кадров широковещательной рассылки, что при этом потребляется вся доступная полоса пропускания. Соответственно, для легитимного трафика нет доступной полосы пропускания, и сеть становится недоступной для обмена данными. Описанная ситуация — эффективный отказ в обслуживании.

Широковещательный шторм неизбежен в сети, где возникла петля. По мере того, как все больше устройств отправляют широковещательные рассылки по сети, все больше трафика попадает в петлю и потребляет ресурсы. В конечном счете это создает широковещательный шторм, что приводит к сбоям в сети.

Вставить
видео

После просмотра анимации о широковещательном шторме рассмотрим ее содержимое:

1. PC1 передает кадр широковещательной рассылки в сеть, где возникла петля.
2. Кадр широковещательной рассылки циклически передается между всеми соединенными друг с другом коммутаторами в сети.
3. PC4 тоже отправляет кадр широковещательной рассылки в сеть, где возникла петля.
4. Кадр широковещательной рассылки PC4 также попадает в петлю между всеми соединенными друг с другом коммутаторами, как и кадр широковещательной рассылки PC1.
5. По мере того, как все больше устройств отправляют широковещательные рассылки по сети, все больше трафика попадает в петлю и потребляет ресурсы. В конечном счете это создает широковещательный шторм, что приводит к сбоям в сети.
6. Когда сеть полностью насыщена трафиком широковещательной рассылки, который циклически передается между коммутаторами, новый трафик отбрасывается коммутатором, поскольку он не в состоянии его обработать.

Проблемы с избыточностью на 1 уровне. Дублирование одноадресные кадры.

Кадры широковещательной рассылки являются не единственным типом кадров, на которые влияет возникновение петель. Кадры одноадресной рассылки, отправленные в сеть, где возникла петля, могут стать причиной дублирования кадров, поступающих на устройство назначения.

ВСТАВИТЬ видео

Содержание анимации:

1. PC1 отправляет кадр одноадресной рассылки, предназначенный для PC4.
2. S2 не содержит в своей таблице MAC-адресов записи для PC4, поэтому выполняет лавинную рассылку этого кадра из всех портов коммутатора, пытаясь найти PC4.
3. Кадр поступает на коммутаторы S1 и S3.
4. S1 содержит в таблице MAC-адресов записи для PC4, поэтому он отправляет кадр на PC4.
5. S3 также содержит в таблице MAC-адресов запись для PC4, поэтому отправляет кадр одноадресной рассылки из порта Магистраль 3 на S1.
6. S1 принимает дублированный кадр и отправляет его на PC4.
7. Таким образом, PC4 принимает два одинаковых кадра.

Большинство протоколов верхнего уровня не предназначены для распознавания или устранения проблемы дублированной передачи. Как правило, протоколы, использующие механизм нумерации последовательности, предполагают, что произошел сбой передачи, и номер последовательности переходит в другой сеанс обмена данными.

Остальные протоколы пытаются передать дублированные данные соответствующему протоколу верхнего уровня для обработки и, возможно, отбрасывания.

Протоколы LAN 2 уровня, например, Ethernet, не поддерживают механизмы распознавания и предотвращения бесконечных циклических кадров. Некоторые протоколы 3 уровня используют механизмы времени жизни (TTL), которые ограничивают количество попыток повторной передачи пакетов сетевыми устройствами 3 уровня.

В отсутствие такого механизма устройства 2 уровня будут производить трафик в бесконечном цикле. Механизм предотвращения петли 2 уровня (STP) разработан как раз для решения данных проблем.

Во избежание подобных проблем в сети с избыточностью, на коммутаторах должны быть включены определённые типы протокола spanning-tree. Протокол spanning-tree по умолчанию включено на коммутаторах Cisco, предотвращая, таким образом, возникновение петель 2 уровня.

Принцип работы STP

Алгоритм протокола spanning – tree protocol

Избыточность повышает доступность топологии сети посредством защиты сети от единой точки отказа — например, неисправного сетевого кабеля или коммутатора. При реализации в проектировании физической избыточности возникают петли и дублирование кадров.

Петли и дублированные кадры являются причиной серьезных неполадок в коммутируемой сети. Протокол STP разработан для решения подобных проблем.

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. К таким данным не относятся кадры BPDU, которые используются протоколом STP для предотвращения петель.

Для предотвращения петель в сети чрезвычайно важно блокировать избыточные пути. 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.

Вставить видео

В рассматриваемом примере протокол STP включен на всех коммутаторах:

1. PC1 отправляет широковещательную рассылку в сеть.

2. S2 настроен с использованием протокола STP, и для порта Магистраль 2 задано состояние блокировки.

3. S1 принимает кадр широковещательной рассылки и пересылает его из всех портов коммутатора, откуда он поступает на PC4 и S3. S3 пересылает кадр из порта для Магистраль 2, и S2 пропускает этот кадр. Возникновение петли 2 уровня предотвращено.

Вставить видео

В приведённом примере:

1. PC1 отправляет широковещательную рассылку в сеть.
2. После этого широковещательная рассылка пересылается по сети, как показано в предыдущей анимации.
3. Возникает сбой в транковом канале между S2 и S1, что приводит к прерыванию предыдущего пути.
4. S2 снимает блокировку с предварительно заблокированного порта для Магистраль 2 и разрешает передачу трафика широковещательной сети по альтернативному пути, обеспечивая дальнейший обмен данными.

Если этот канал снова работает, выполняется повторное схождение протокола STP, а порт на S2 снова блокируется.

Протокол STP предотвращает возникновение петель за счёт настройки беспетлевого пути в сети с использованием портов, стратегически настроенных на заблокированное состояние. Коммутаторы, использующие протокол STP, могут компенсировать сбои за счёт динамической разблокировки ранее блокированных портов и разрешения передачи трафика по альтернативным путям.

До сих пор использовался термин Spanning Tree Protocol (протокол spanning-tree) и аббревиатура STP. Однако использование этого термина и этой аббревиатуры может быть двусмысленным. Многие специалисты используют данный термин и аббревиатуру для обозначения различных реализаций протокола spanning-tree, например протокола Rapid Spanning Tree Protocol (RSTP) и протокола Multiple Spanning Tree Protocol (MSTP).

Чтобы правильно объяснять принципы протокола spanning-tree, важно понимать, о какой конкретно реализации или стандарте идет речь в данном контексте.

В новейшей версии документации IEEE по протоколу spanning-tree (IEEE-802-1D-2004) говорится: «Протокол STP в настоящее время заменен протоколом Rapid Spanning Tree Protocol (RSTP)»; можно заметить, что в IEEE термин «STP» используется для обозначения исходной реализации протокола spanning-tree, а «RSTP» — для описания версии протокола spanning-tree, указанной в IEEE-802.1D-2004.

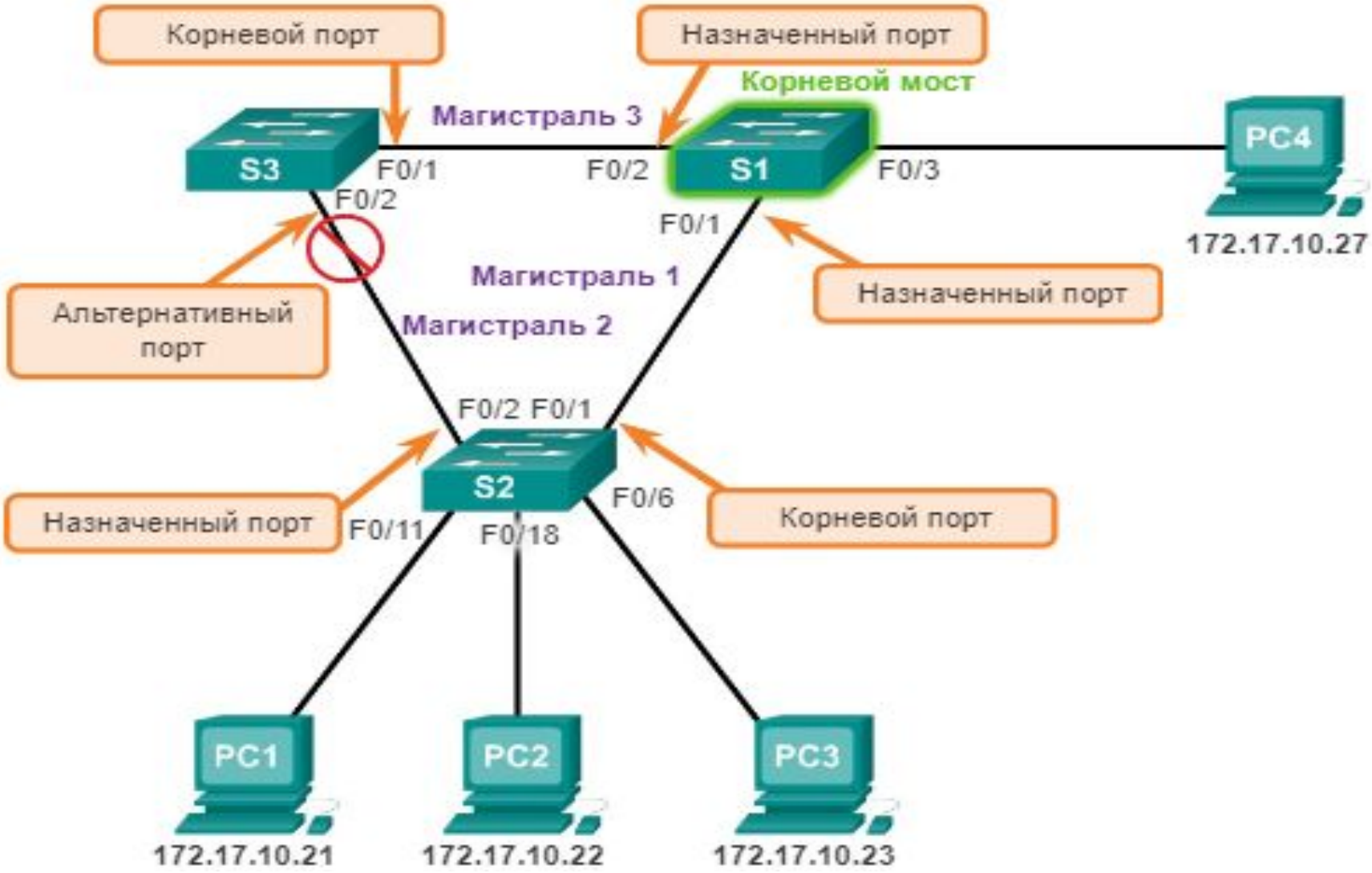
В рамках данной программы, если в контексте обсуждения речь идет об исходном протоколе STP, то во избежание расхождений используется фраза: «исходный протокол spanning-tree 802.1D».

Алгоритм протокола spanning – tree protocol. Роли портов.

EEE 802.1D STP использует алгоритм протокола spanning-tree (STA), чтобы определить, какие порты коммутаторов в сети должны быть переведены в состояние блокировки во избежание возникновения петель.

STA назначает один из коммутаторов в качестве корневого моста и использует его как точку привязки для расчёта всех путей.

На рисунке корневой мост (коммутатор S1) выбран с помощью специального процесса выбора.



Все коммутаторы, участвующие в STP, обмениваются кадрами BPDU, чтобы определить, какой коммутатор имеет самое низкое значение идентификатора моста (BID) в сети. Коммутатор с наименьшим значением BID автоматически становится корневым мостом для расчётов STA.

BPDU представляет собой кадр обмена сообщениями, которым обмениваются коммутаторы для STP. Каждый BPDU содержит идентификатор BID, который определяет коммутатор, отправивший BPDU. 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.

После определения корневого моста STA рассчитывает кратчайший путь до него. Все коммутаторы используют STA для определения портов, подлежащих блокировке. Пока STA определяет оптимальные пути до корневого моста для всех портов коммутатора в домене широковещательной рассылки, пересылка трафика по сети заблокирована.

При определении портов, подлежащих блокировке, STA учитывает стоимость как пути, так и порта. Стоимость портов рассчитывается с помощью значений стоимости порта, зависящей от скорости каждого порта коммутатора на данном маршруте. Сумма значений стоимости порта определяет общую стоимость пути до корневого моста. Если для выбора доступно несколько путей, STA выбирает путь с наименьшей стоимостью.

Определив наиболее предпочтительные пути для каждого коммутатора, алгоритм STA назначает роли участвующим портам коммутаторов.

Роли портов описывают их связь с корневым мостом в сети, а также указывают, разрешена ли для них пересылка трафика:

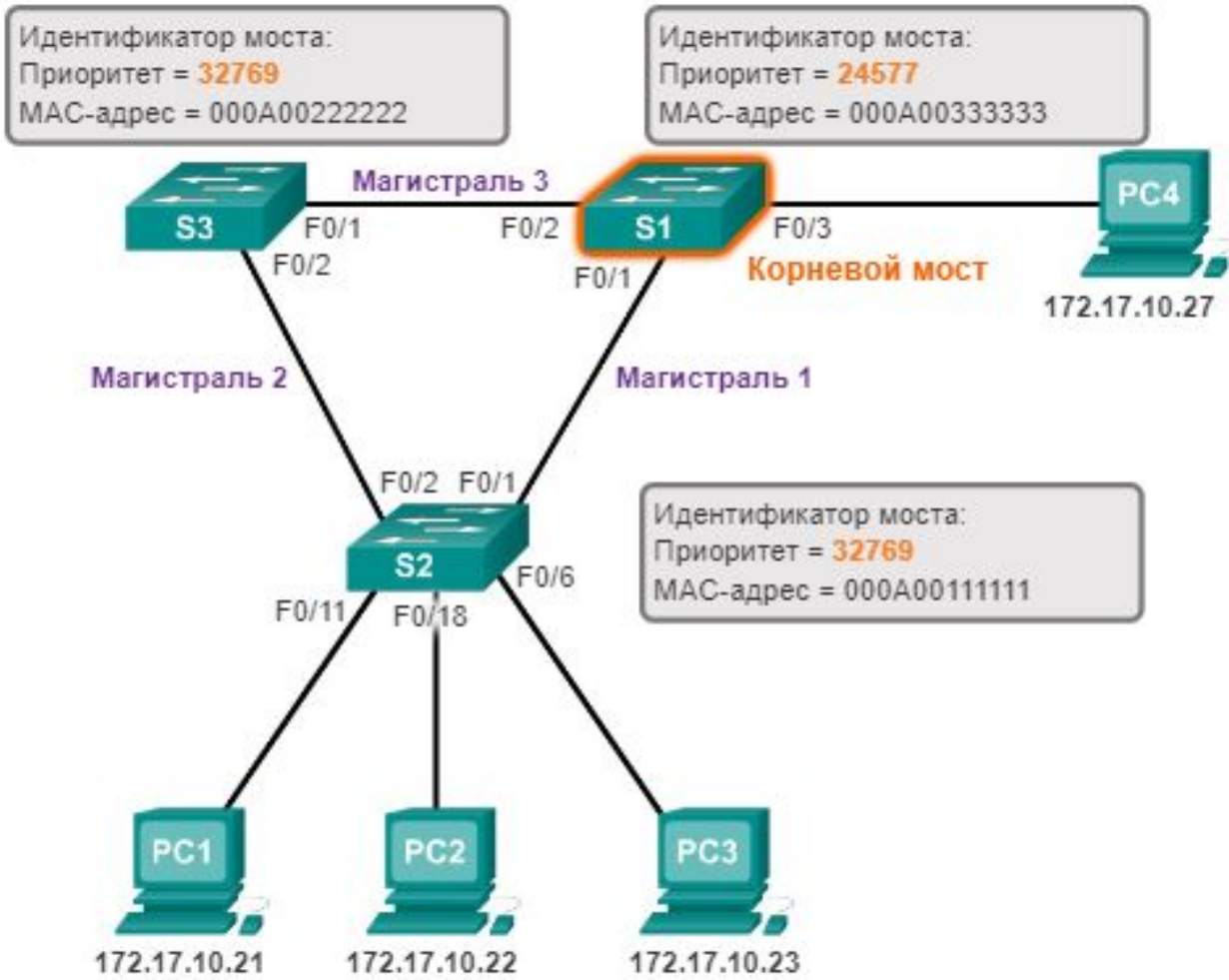
1. **Корневые порты** – порты коммутатора, находящиеся максимально близко к корневому мосту.
2. **Назначенные порты** — все некорневые порты, которым, тем не менее, разрешено пересылать трафик по сети.
3. **Альтернативные и резервные порты** – альтернативные и резервные порты настраиваются в состояние блокировки во избежание возникновения петель.
4. **Отключенные порты** – отключенным называется порт коммутатора, питание которого отключено.

Алгоритм протокола spanning – tree protocol. Корневой мост.

Как показано на рисунке, все экземпляры протокола spanning-tree (коммутируемая сеть LAN или домен широковещательной рассылки) содержат коммутатор, назначенный в качестве корневого моста.

Корневой мост служит точкой привязки для всех расчётов протокола spanning-tree, позволяя определить избыточные пути, которые следует заблокировать.

Процесс выбора определяет, какой из коммутаторов станет корневым мостом.



На рисунке показаны поля BID. Идентификатор BID состоит из значения приоритета, расширенного идентификатора системы и MAC-адреса коммутатора.

Все коммутаторы в домене широковещательной рассылки участвуют в процессе выбора. После загрузки коммутатора они начинают рассылать кадры BPDU с интервалом в две секунды.

Эти BPDU содержат идентификатор BID коммутатора и идентификатор корневого моста.

Когда коммутаторы пересылают свои кадры BPDU, смежные коммутаторы в домене широковещательной рассылки считывают из них данные об идентификаторе корневого моста.

Если идентификатор корневого моста полученного кадра BPDU имеет меньшее значение, чем идентификатор корневого моста на принимающем коммутаторе, то в этом случае принимающий коммутатор обновляет свой идентификатор корневого моста, указывая смежный коммутатор в качестве корневого моста.

Фактически это может быть не смежный коммутатор, а любой другой коммутатор в домене широковещательной рассылки. Затем коммутатор пересылает новые кадры BPDU с меньшим значением идентификатора корневого моста на другие смежные коммутаторы.

Постепенно коммутатор с наименьшим значением идентификатора BID определяется в качестве корневого моста для экземпляра протокола spanning-tree.

Алгоритм протокола spanning – tree protocol. Стоимость пути.

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.

Сведения о пути определяются посредством суммирования значений стоимости отдельных портов на пути от некорневого коммутатора к корневому мосту. Каждый «адрес назначения», по сути, является портом коммутатора.

Стоимость портов по умолчанию определяется скоростью работы порта. Как показано на рисунок, значение стоимости портов Ethernet 10 Гбит/с равно 2, портов Ethernet 1 Гбит/с — 4, портов Fast Ethernet 100 — 19, а портов Ethernet 10 Мбит/с — 100.

Скорость канала	Стоимость (по изменённой спецификации IEEE)	Стоимость (по предыдущей спецификации IEEE)
10 Гбит/с	2	1
1 Гбит/с	4	1
100 Мбит/с	19	10
10 Мбит/с	100	100

Чтобы продемонстрировать постоянные изменения, связанные с высокоскоростными сетевыми технологиями, коммутаторы Catalyst 4500 и 6500 поддерживают метод стоимости более длинного пути. Например, 10 Гбит/с имеет значение стоимости пути 2000, 100 Гбит/с — 200, а 1 Тбит/с — 20.

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.

Чтобы настроить стоимость порта интерфейса, введите команду **spanning-tree cost** *value* в режиме конфигурации интерфейса. Это значение может находиться в диапазоне между 1 и 200 000 000.

Настроить стоимость порта

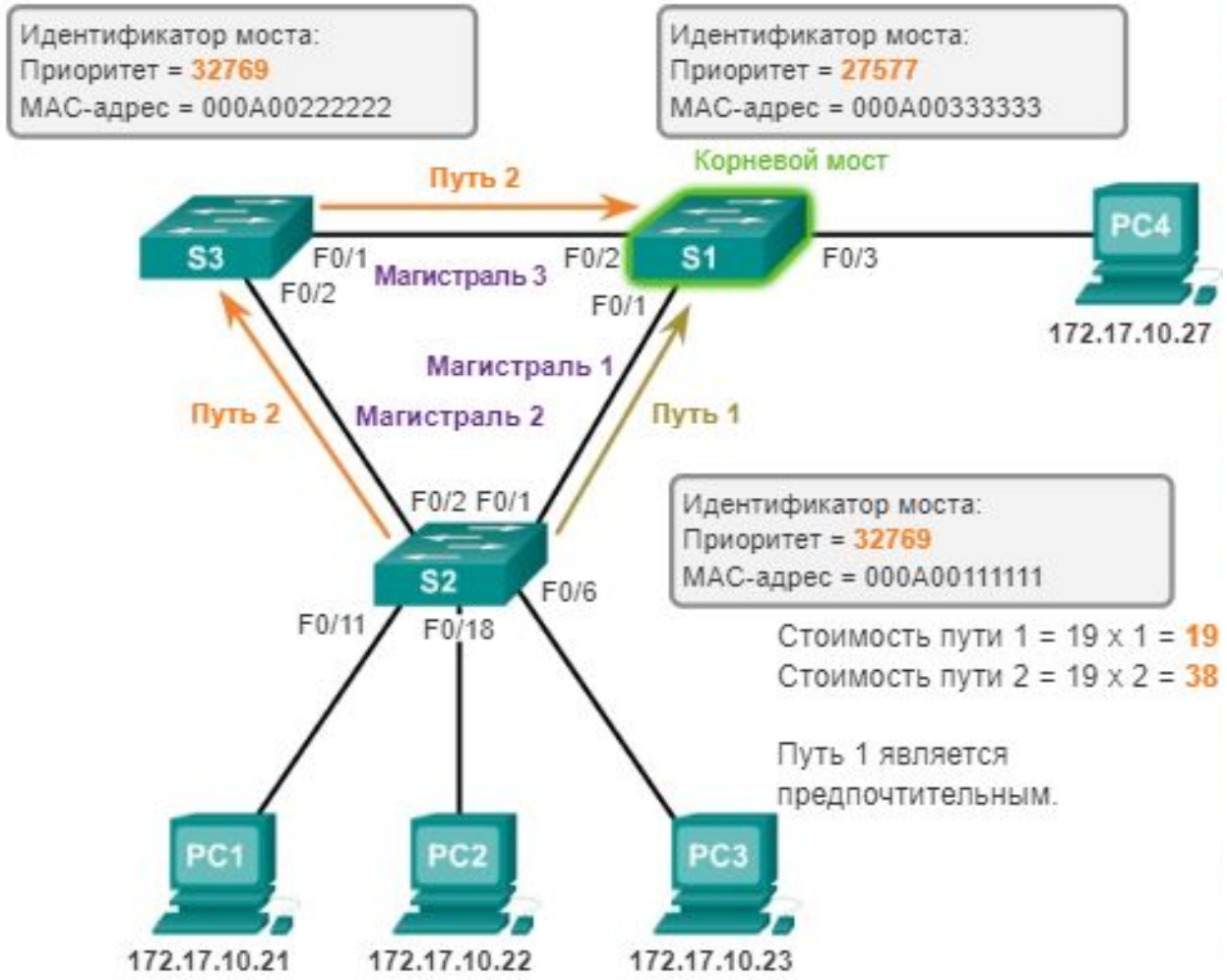
```
S2# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
S2(config)# interface f0/1
S2(config-if)# spanning-tree cost 25
S2(config-if# end
S2#
```

Сбросить стоимость порта

```
S2# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
S2(config)# interface f0/1
S2(config-if)# no spanning-tree cost
S2(config-if)# end
S2#
```

Чтобы восстановить значение стоимости порта по умолчанию 19, введите команду режима конфигурации интерфейса **no spanning-tree cost**.

Стоимость пути равна сумме всех значений стоимости порта по пути к корневому мосту.



Пути с наименьшей стоимостью становятся предпочтительными, а все остальные избыточные пути блокируются. В данном примере значение стоимости пути от S2 к корневому мосту S1 равно 19 по пути 1 (с учетом отдельных значений стоимости порта, указанных в стандарте IEEE) и 38 по пути 2.

Поскольку общая стоимость пути 1 к корневому мосту ниже, именно этот путь является предпочтительным. После этого протокол STP осуществляет блокирование избыточного пути для предотвращения возникновения петли.

Чтобы проверить стоимость порта и стоимость пути к корневому мосту, введите команду **- show spanning-tree**

```
S2# show spanning-tree

VLAN001
   Spanning tree enabled protocol ieee
   Root ID     Priority  24577
               Address      000A.0033.3333
               Cost         19
               Port         1
               Hello Time   2 sec Max Age 20 sec Forward Delay 15 sec

   Bridge ID   Priority      32769  (priority 32768 sys-id-ext 1)
               Address 000A.0011.1111
               Hello time 2 sec Max Age 20 sec Forward Delay 15 sec
               Aging Time 300

Interface    Role   Sts  Cost    Prio.Nbr  Type
-----------  -----  ---- -----   --------- -----------------
F0/1         Root   FWD   19      128.1     Edge P2p
F0/2         Desg   FWD   19      128.2     Edge P2p
```

Поле Cost в верхней части выходных данных содержит итоговое значение стоимости пути к корневому мосту.

Это значение меняется в зависимости от количества портов коммутатора, которые должны быть пройдены на пути к корневому мосту.

В выходных данных все интерфейсы также определены со значением стоимости отдельного порта 19.

Формат кадра BPDU 802.1D

Алгоритм протокола spanning-tree зависит от обмена кадрами BPDU, выполняемого для определения корневого моста.

Кадр BPDU содержит 12 отдельных полей, которые содержат сведения о пути и приоритете, используемые для определения корневого моста и путей к нему.

Номер поля	Байты	Поле
1-4	2	Protocol ID
	1	Version
	1	Version
	1	Flags
5-8	8	Root ID
	4	Root ID
	8	Bridge ID
	2	Bridge ID
9-12	2	Message age
	2	Max age
	2	Hello time
	2	Forward delay

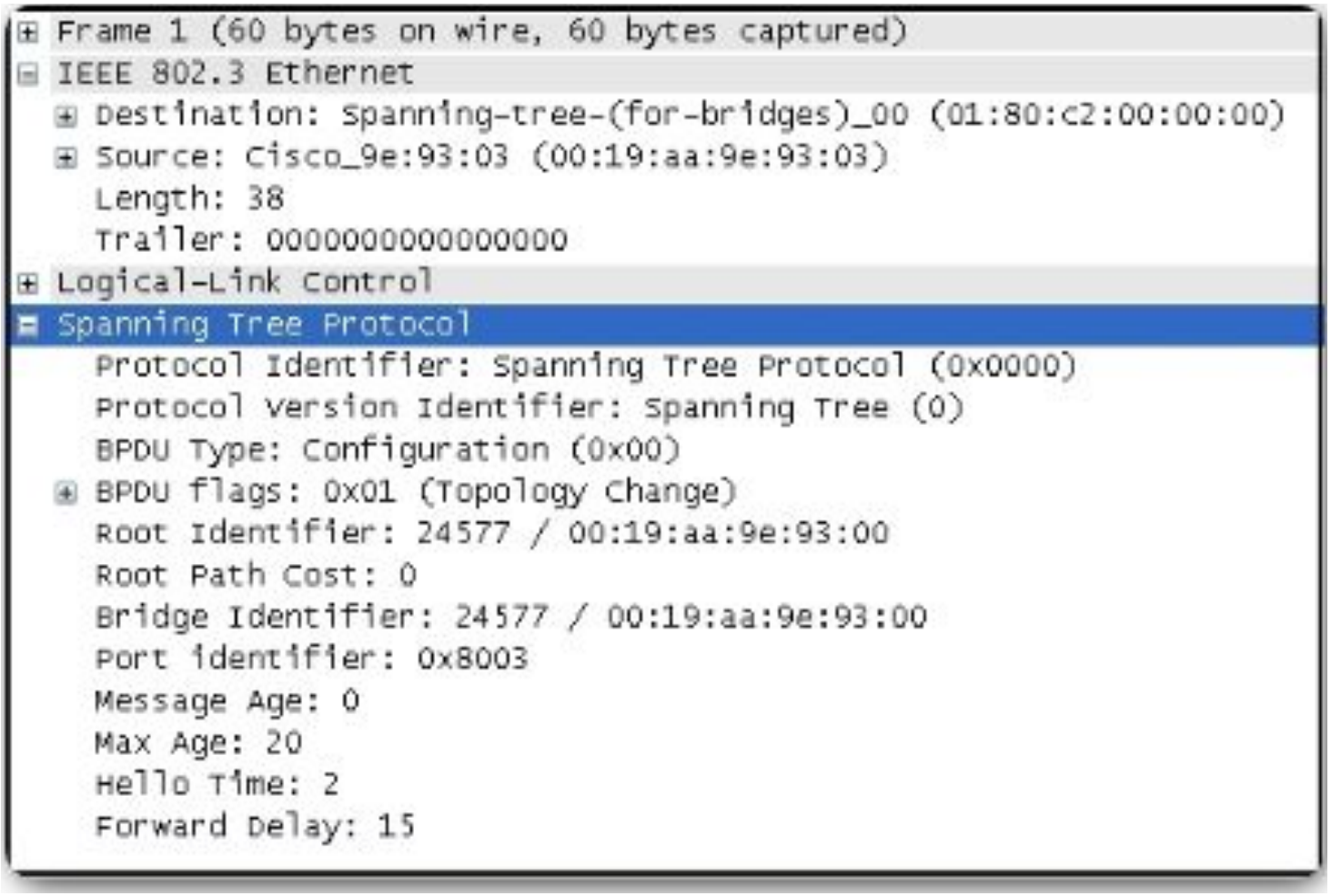
Protocol ID

В поле Protocol ID «Идентификатор протокола» указан тип используемого протокола. Данное поле содержит нулевое значение.

Поля BPDU:

1. В первых четырех полях указаны протокол, версия, тип сообщения и флаги состояния.
2. Следующие четыре поля используются для определения корневого моста и стоимости пути к нему.
3. Последние четыре поля являются полями таймера, которые определяют интервал отправки сообщений BPDU и продолжительность хранения данных, полученных посредством процесса BPDU (см. следующий раздел).

На рисунке показан кадр BPDU, полученный с помощью программы Wireshark.



В этом примере кадр BPDU содержит большее количество полей, чем описано выше. Сообщение BPDU при передаче по сети инкапсулируется в кадр Ethernet. Заголовок 802.3 указывает адреса источника и назначения кадра BPDU.

Кадр содержит MAC-адрес назначения 01:80:C2:00:00:00, который является адресом групповой рассылки для группы протокола spanning-tree.

При адресации кадра с использованием этого MAC-адреса все коммутаторы, настроенные для протокола spanning-tree, принимают и считывают данные из кадра. Все остальные устройства в сети игнорируют кадр.

В этом примере идентификатор корневого моста в полученном кадре BPDU совпадает с идентификатором BID.

Это указывает на то, что кадр получен из корневого моста. Все таймеры настроены с использованием значений по умолчанию.

Распространение и процесс BPDU

Изначально каждый коммутатор в домене широковещательной рассылки считает себя корневым мостом для экземпляра протокола spanning-tree, поэтому отправленные кадры BPDU содержат идентификатор BID локального коммутатора в качестве идентификатора корневого моста.

По умолчанию после загрузки коммутатора кадры BPDU отправляются с интервалом в две секунды; то есть значение таймера приветствия по умолчанию, указанное в кадре BPDU — две секунды. Все коммутаторы предоставляют сведения о собственном идентификаторе BID, идентификаторе корневого моста и стоимости пути к корневому мосту.

Когда смежные коммутаторы принимают кадр BPDU, они сопоставляют содержащийся в нем идентификатор корневого моста с локальным идентификатором корневого моста. Если идентификатор корневого моста в кадре BPDU имеет меньшее значение, чем локальный идентификатор корневого моста, коммутатор обновляет локальный идентификатор и тот идентификатор, который содержится в сообщениях BPDU.

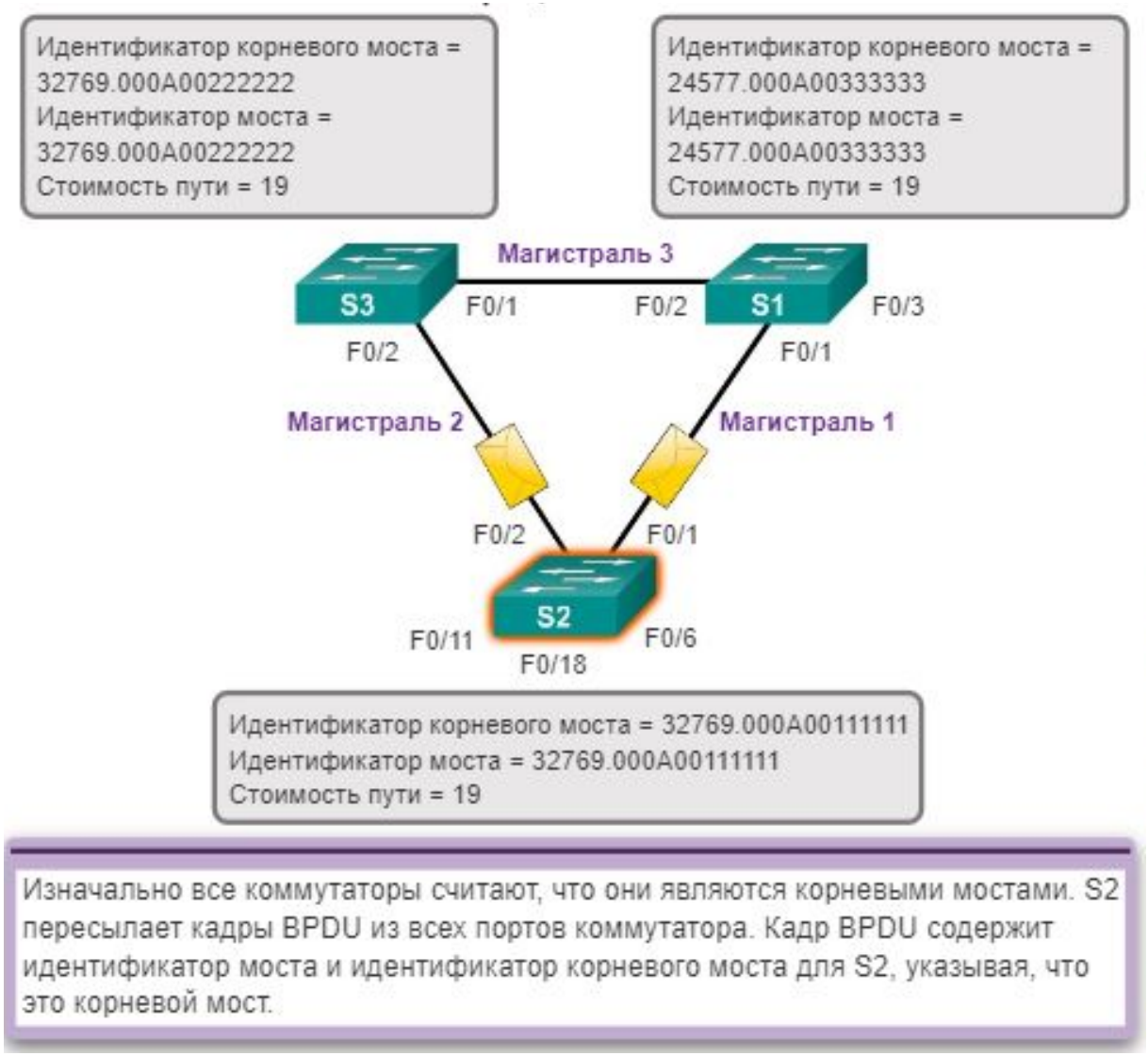
Эти сообщения указывают новый корневой мост в сети. Расстояние до корневого моста также обозначается посредством обновления стоимости пути. Например, если кадр BPDU получен на порте коммутатора Fast Ethernet, стоимость пути увеличивается на 19. Если локальный идентификатор корневого моста имеет меньшее значение, чем идентификатор корневого моста, полученный в кадре BPDU, кадр BPDU отбрасывается.

После обновления идентификатора корневого моста в целях определения нового корневого моста, все последующие кадры BPDU, отправленные с этого коммутатора, будут содержать новый идентификатор корневого моста и обновленное значение стоимости пути.

Таким образом, все остальные смежные коммутаторы могут постоянно видеть самое низкое значение идентификатора корневого моста. По мере прохождения кадров BPDU между другими смежными коммутаторами стоимость пути постоянно обновляется, чтобы указать общую стоимости пути к корневому мосту. Все коммутаторы в протоколе spanning tree используют свой путь для определения оптимального пути к корневому мосту.

Далее представлено краткое описание процесса BPDU:

1. Изначально все коммутаторы определяют себя в качестве корневого моста. Коммутатор S2 пересылает кадры BPDU из всех своих портов.



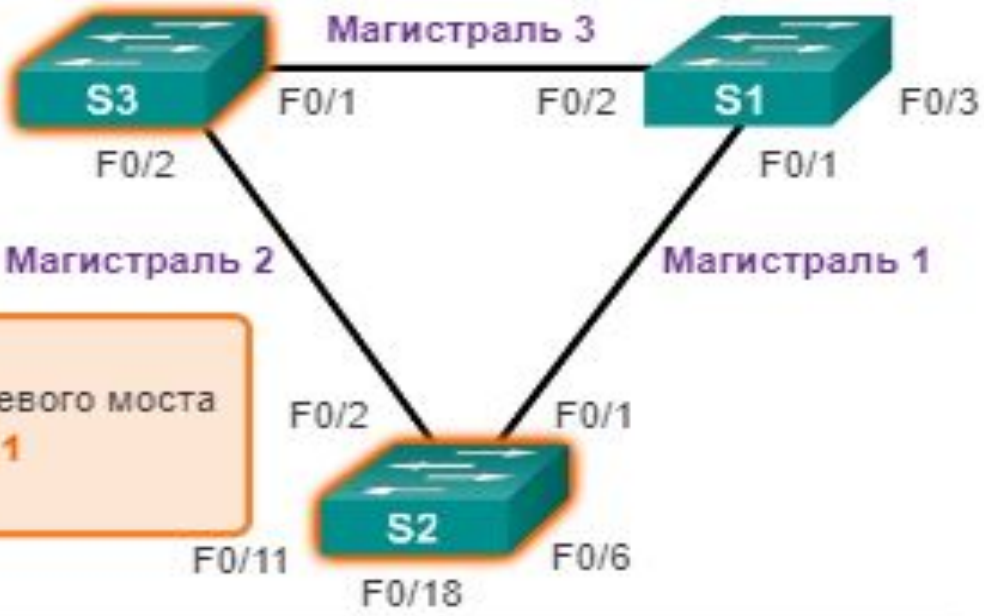
2. Когда S3 получает BPDU от S2, S3 сравнивает свой идентификатор корневого моста с полученным кадром BPDU.

Приоритеты одинаковы, поэтому коммутатор вынужден проверить часть MAC-адреса, чтобы определить, какой из MAC-адресов имеет более низкое значение.

Поскольку S2 содержит меньшее значение MAC-адреса, то S3 обновляет свой идентификатор корневого моста с учетом идентификатора корневого моста S2. На этом этапе S3 считает S2 корневым мостом (Слайд №56) .

Идентификатор корневого моста = 32769.000A00111111
Идентификатор моста = 32769.000A00222222
Стоимость пути = 19

Идентификатор корневого моста = 24577.000A00333333
Идентификатор моста = 24577.000A00333333
Стоимость пути = 19

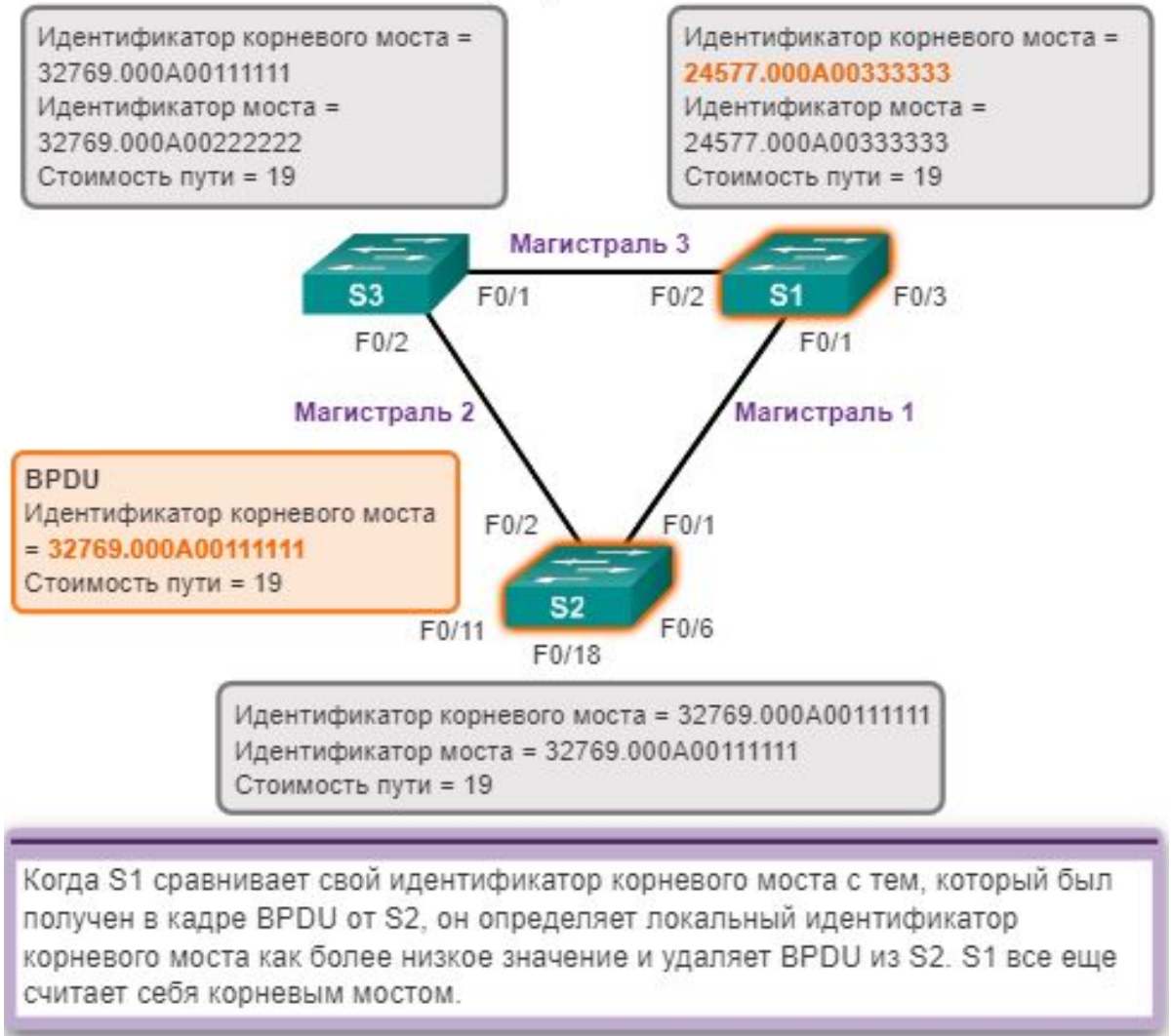


BPDU
Идентификатор корневого моста = 32769.000A00111111
Стоимость пути = 19

Идентификатор корневого моста = 32769.000A00111111
Идентификатор моста = 32769.000A00111111
Стоимость пути = 19

S3 сравнивает полученный идентификатор корневого моста с собственным и определяет S2 как более низкое значение идентификатора корневого моста. S3 обновляет свой идентификатор корневого моста до значения идентификатора корневого моста S2.

3. Когда S1 сравнивает свой идентификатор корневого моста с идентификатором, содержащимся в полученном кадре BPDU, он определяет свой локальный идентификатор корневого моста как меньшее значение и отбрасывает кадр BPDU, полученный от S2.



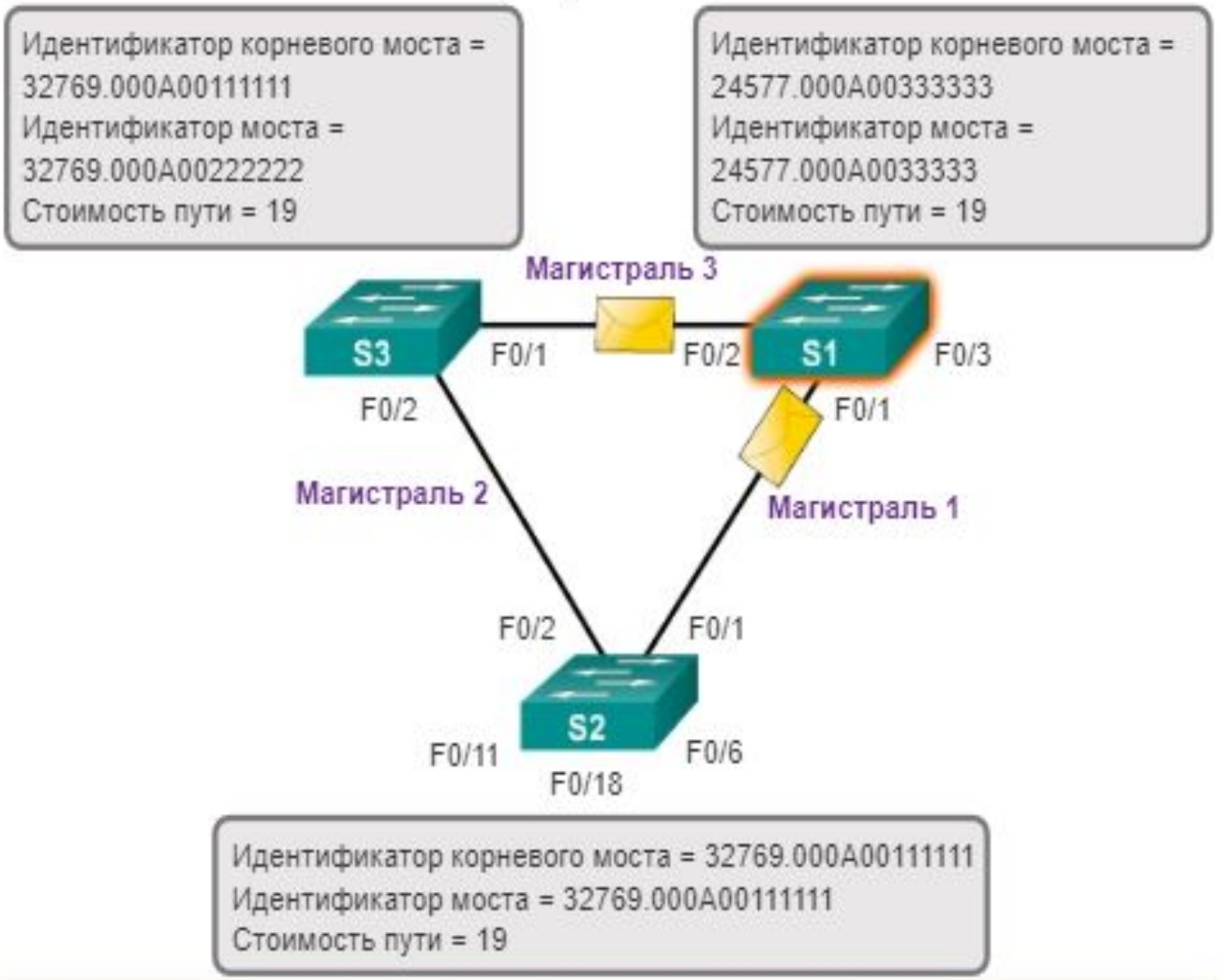
4. Когда S3 отправляет свои кадры BPDU, идентификатору корневого моста, в кадре BPDU содержится идентификатор корневого моста S2.

S3 пересылает кадры BPDU из всех портов коммутатора. Кадр BPDU содержит идентификатор корневого моста S2, указывая, что это корневой мост.

5. Когда S2 получает кадр BPDU, он отбрасывает его после того, как подтвердит, что идентификатор корневого моста в BPDU совпадает с локальным идентификатором корневого моста.

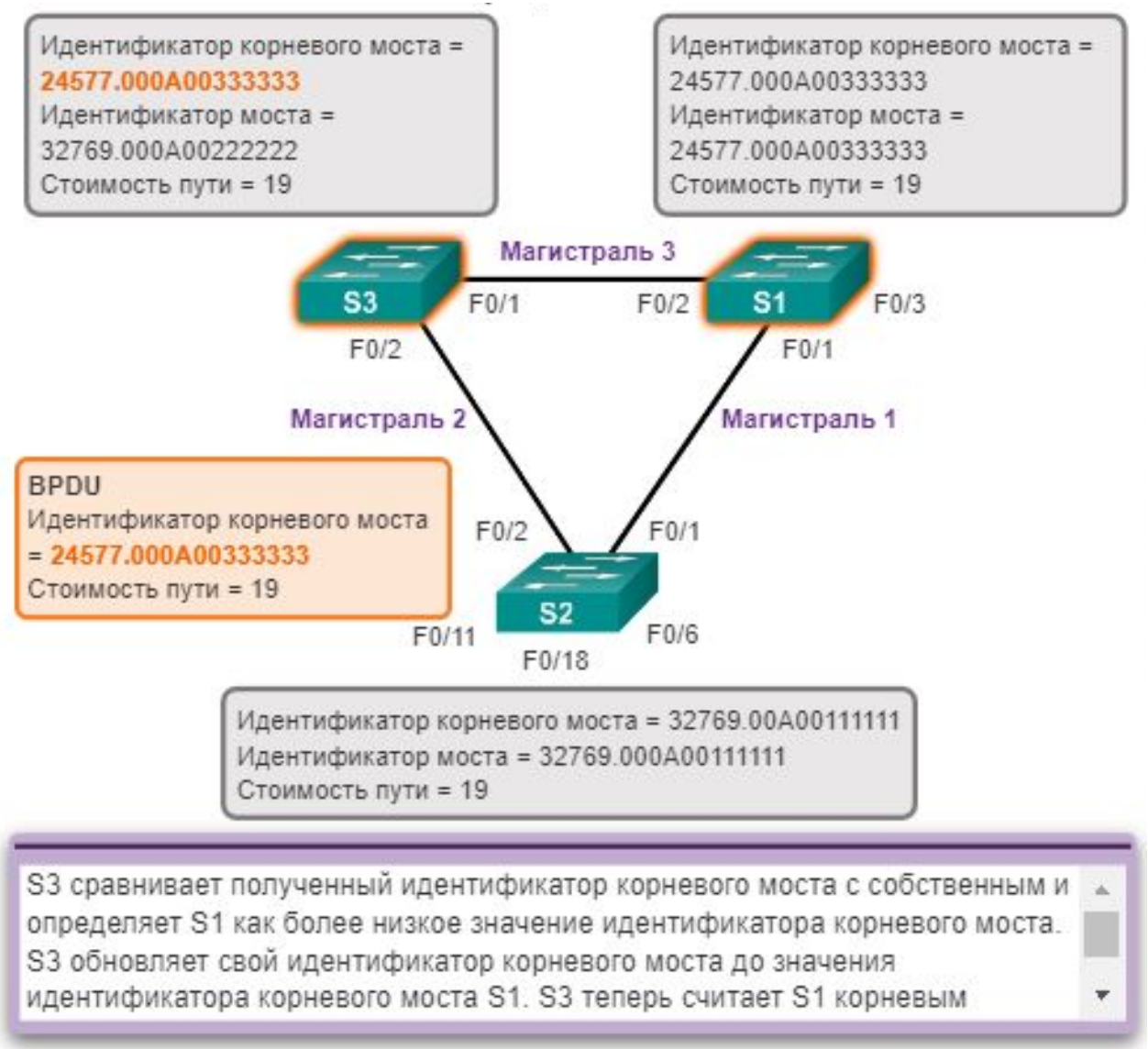
6. Поскольку S1 содержит более низкое значение приоритета в своем идентификаторе корневого моста, он отбрасывает кадр BPDU, полученный от S3.

7. S1 отправляет свои кадры BPDU.

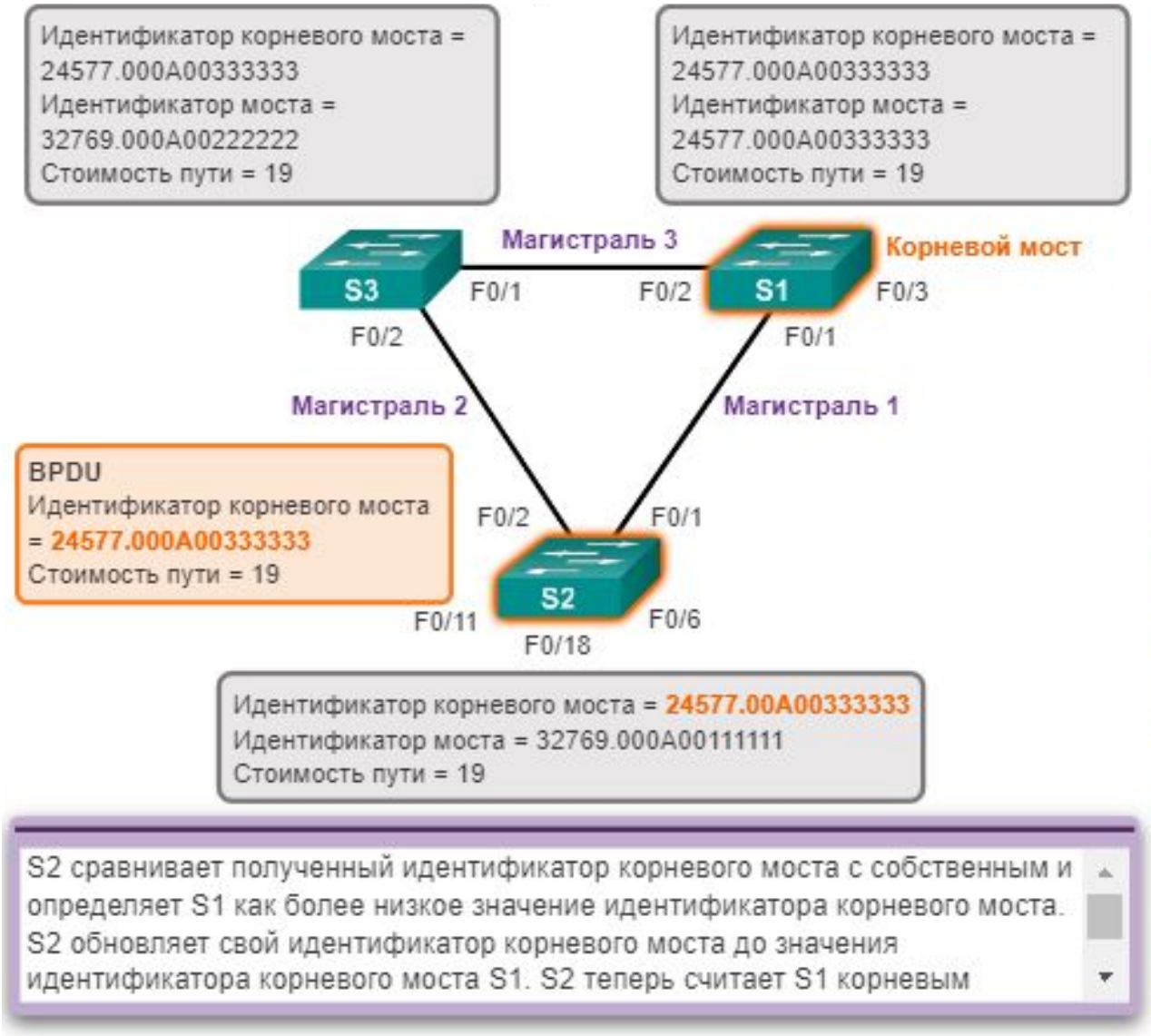


S1 пересылает кадры BPDU из всех портов коммутатора. Кадр BPDU содержит идентификатор моста и идентификатор корневого моста для S1, указывая, что это корневой мост.

8. S3 определяет, что идентификатор корневого моста в кадре BPDU содержит меньшее значение и, следовательно, обновляет свои значения идентификатора корневого моста, указывая, что S1 теперь является корневым мостом.



9. S2 определяет, что идентификатор корневого моста в кадре BPDU содержит меньшее значение и, следовательно, обновляет свои значения идентификатора корневого моста, указывая, что S1 теперь является корневым мостом.



Поля BID кадра BPDU

Идентификатор моста (BID) используется для определения корневого моста в сети.

Поле BID кадра BPDU содержит три отдельных поля:

1. Приоритет моста
2. Расширенный идентификатор системы
3. MAC-адрес

При выборе корневого моста используются все поля.

Приоритет моста

Приоритет моста представляет собой настраиваемое значение, которое можно использовать для определения коммутатора, который станет корневым мостом.

Коммутатор с наименьшим приоритетом, который подразумевает наименьшее значение BID, становится корневым мостом, поскольку преимущество имеет более низкое значение приоритета.

Например, если вы хотите назначить в качестве корневого моста конкретный коммутатор, то для него следует задать более низкое значение приоритета, чем для остальных коммутаторов в сети.

По умолчанию для всех коммутаторов Cisco используется значение приоритета 32768.

Значения варьируются в диапазоне от 0 до 61440 с шагом в 4096.

Допустимые значения приоритета:

0, 4096, 8192, 12288, 16384, 20480, 24576, 28672, 32768, 36864, 40960, 45056, 49152, 53248, 57344 и 61440.

Все остальные значения отклоняются.

Приоритет моста 0 имеет преимущество по сравнению со всеми остальными значениями приоритета моста.

Расширенный идентификатор системы

Благодаря повсеместному использованию сетей VLAN для сегментации сетевой инфраструктуры, 802.1D был расширен с учетом поддержки сетей VLAN. Именно поэтому идентификатор сети VLAN был добавлен в кадр BPDU.

Сведения о сети VLAN включены в кадр BPDU с помощью расширенного идентификатора системы. Все новые модели коммутаторов по умолчанию используют расширенные идентификаторы системы.

Как показано на рисунке, поле приоритета моста имеет 2 байта или 16 бит в длину; 4 бита указывают приоритет моста, а 12 бит — расширенный идентификатор системы, который определяет сеть VLAN, участвующую в данном процессе STP.

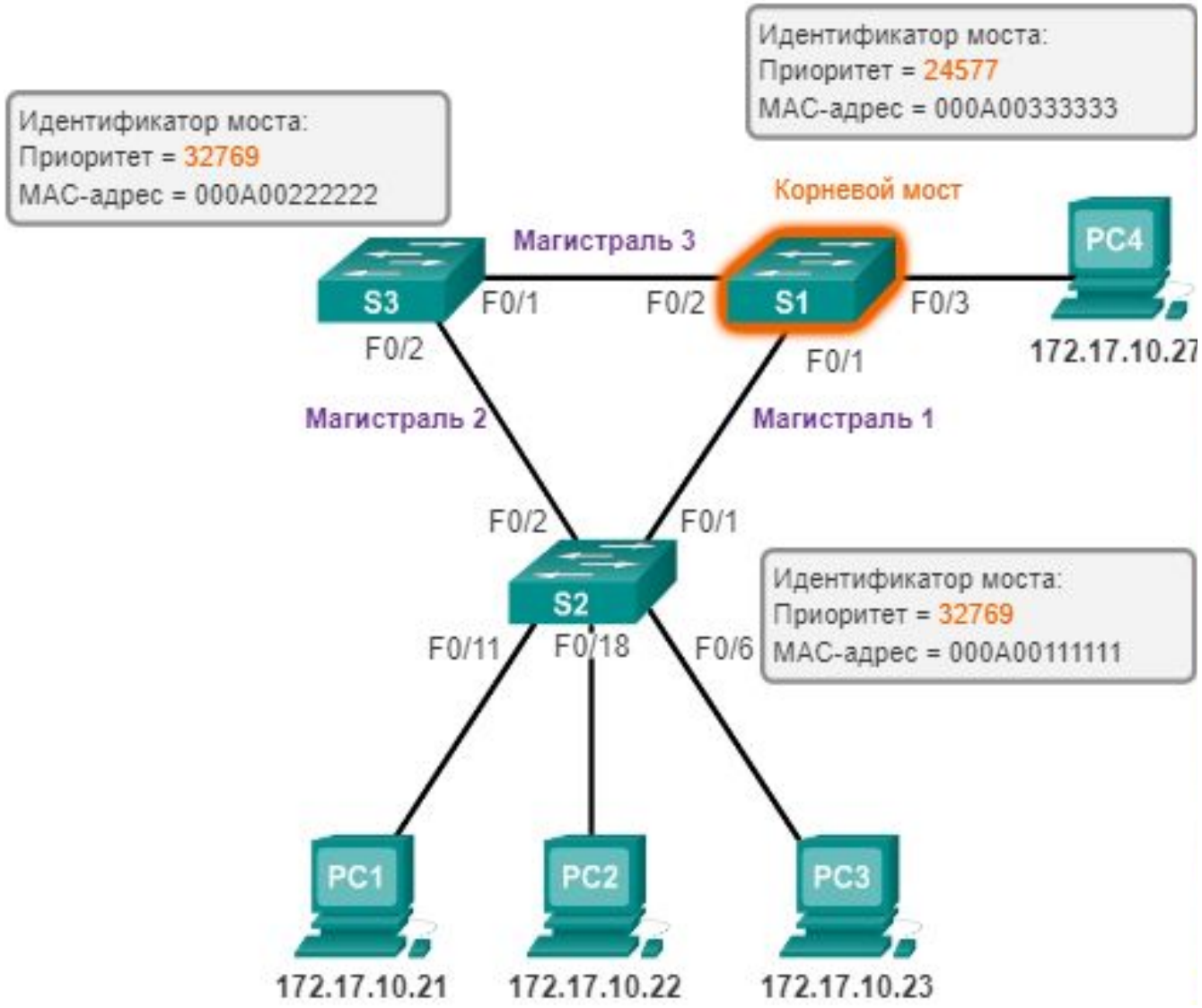
Благодаря тому, что длина расширенного идентификатора системы составляет 12 бит, длина приоритета моста сокращена до 4 бит. В рамках этого процесса крайние правые 12 бит резервируются для идентификатора сети VLAN, а крайние левые 4 бита — для приоритета моста.

Это объясняет, почему значение приоритета моста можно настроить только кратным 4096 или 2^12. Если крайними левыми являются биты 0001, в этом случае значение приоритета моста равно 4096; если крайними левыми являются биты 1111, то значение приоритета моста равно 61440 (= 15 x 4096).

Коммутаторы Catalyst серий 2960 и 3560 не поддерживают настройку приоритета моста равному значению 65536 (= 16 x 4096), поскольку это предполагает использование пятого бита, который недоступен вследствие использования расширенного идентификатора системы.

Для указания приоритета и сети VLAN для кадра BPDU значение расширенного идентификатора системы добавляется к значению приоритета моста в идентификаторе BID.

На рисунке S1 имеет более низкое значение приоритета, чем другие коммутаторы, следовательно, этот коммутатор является предпочтительным в качестве корневого моста для этого экземпляра протокола spanning-tree.



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.

Типы протоколов STP

Краткий обзор

Список протоколов STP

С момента создания исходного стандарта IEEE 802.1D было разработано несколько разновидностей протоколов STP.

К разновидностям протоколов STP относятся следующие:

1. **STP**: исходная версия IEEE 802.1D (802.1D-1998 и более ранние), в рамках которой предоставляется беспетлевая топология в сети с избыточными каналами.
2. **PVST+** является усовершенствованным протоколом компании Cisco, в котором для каждого отдельного VLAN используется отдельный экземпляр RSTP.

3. 802.1D-2004: обновленная версия стандарта STP, в которую входит IEEE 802.1w.

4. Быстрый протокол STP (RSTP) или IEEE 802.1w: доработанный протокол STP, который обеспечивает более быстрое схождение, чем протокол STP.

5. Rapid PVST+: усовершенствованный корпорацией Cisco протокол RSTP, который использует PVST+.

6. Протокол MSTP (несколько протоколов spanning-tree) (MSTP): стандарт IEEE на базе ранее существующей собственной реализации Multiple Instance STP (MISTP) корпорации Cisco.

Характеристики протокола STP

Далее представлены характеристики различных протоколов STP. Выделенные курсивом слова указывают, является ли конкретный протокол STP собственным протоколом Cisco или стандартной реализацией IEEE:

1. **STP**: использует один экземпляр протокола spanning-tree *IEEE 802.1D* для всей коммутируемой сети независимо от количества сетей VLAN. Поскольку используется только один экземпляр, требования к ЦП и памяти для этой версии ниже, чем в отношении других протоколов.

2. **PVST+**: усовершенствованный корпорацией Cisco протокол STP, который предоставляет отдельный экземпляр реализации 802.1D корпорации Cisco для каждой сети VLAN, настроенной в сети.

3. RSTP (или *IEEE 802.1w*): быстрый протокол spanning-tree, обеспечивающий более быстрое схождение, чем исходная реализация 802.1D.

4. Rapid PVST+: усовершенствованный корпорацией Cisco протокол RSTP, который использует PVST+.

5. MSTP: стандарт *IEEE 802.1s*, созданный на основе предыдущей собственной реализации протокола MISTP компании Cisco.

6. MST: реализация Cisco протокола MSTP, которая обеспечивает до 16 экземпляров протокола RSTP (802.1w) и объединяет множество сетей VLAN с идентичной физической и логической топологиями в один общий экземпляр RSTP.

Протокол	Стандарт	Требуемые ресурсы	Сходимость	Расчёт дерева
STP	802.1D	Низкая	Медленная	Все сети VLAN
PVST+	Cisco	Высокая	Медленная	На VLAN
RSTP	802.1w	Средняя	Быстрая	Все сети VLAN
Rapid PVST+	Cisco	Очень высокая	Быстрая	На VLAN
MSTP	802.1s, Cisco	Средняя или высокая	Быстрая	На экземпляр

Каждая реализация поддерживает функции PortFast, BPDU guard, BPDU filter, root guard и loop guard. Требования к ЦП и памяти для этой версии ниже, чем аналогичные требования в отношении протокола Rapid PVST+, но выше, чем для протокола RSTP.

Для коммутаторов Cisco Catalyst по умолчанию используется режим протокола spanning-tree PVST+, включенный на всех портах. PVST+ характеризуется существенно более медленным схождением после изменения топологии, чем Rapid PVST+.

PVST+

Обзор PVST+

Исходный стандарт 802.1D определяет протокол общего spanning-tree (CST), который подразумевает использование только одного экземпляра протокола spanning-tree во всей коммутируемой сети независимо от количества VLAN.

Сеть, использующая CST, имеет следующие характеристики:

1. Распределение нагрузки не поддерживается. Один восходящий канал должен блокировать все сети VLAN.
2. Ресурсы ЦП используются экономно. Требуется вычисление только одного экземпляра протокола spanning-tree.

Корпорация Cisco разработала протокол PVST+ таким образом, чтобы сеть могла использовать независимый экземпляр реализации стандарта IEEE 802.1D для каждой сети VLAN в пределах сети.

PVST+ позволяет одному транковому порту на коммутаторе блокировать отдельную сеть VLAN, не блокируя при этом остальные сети VLAN.

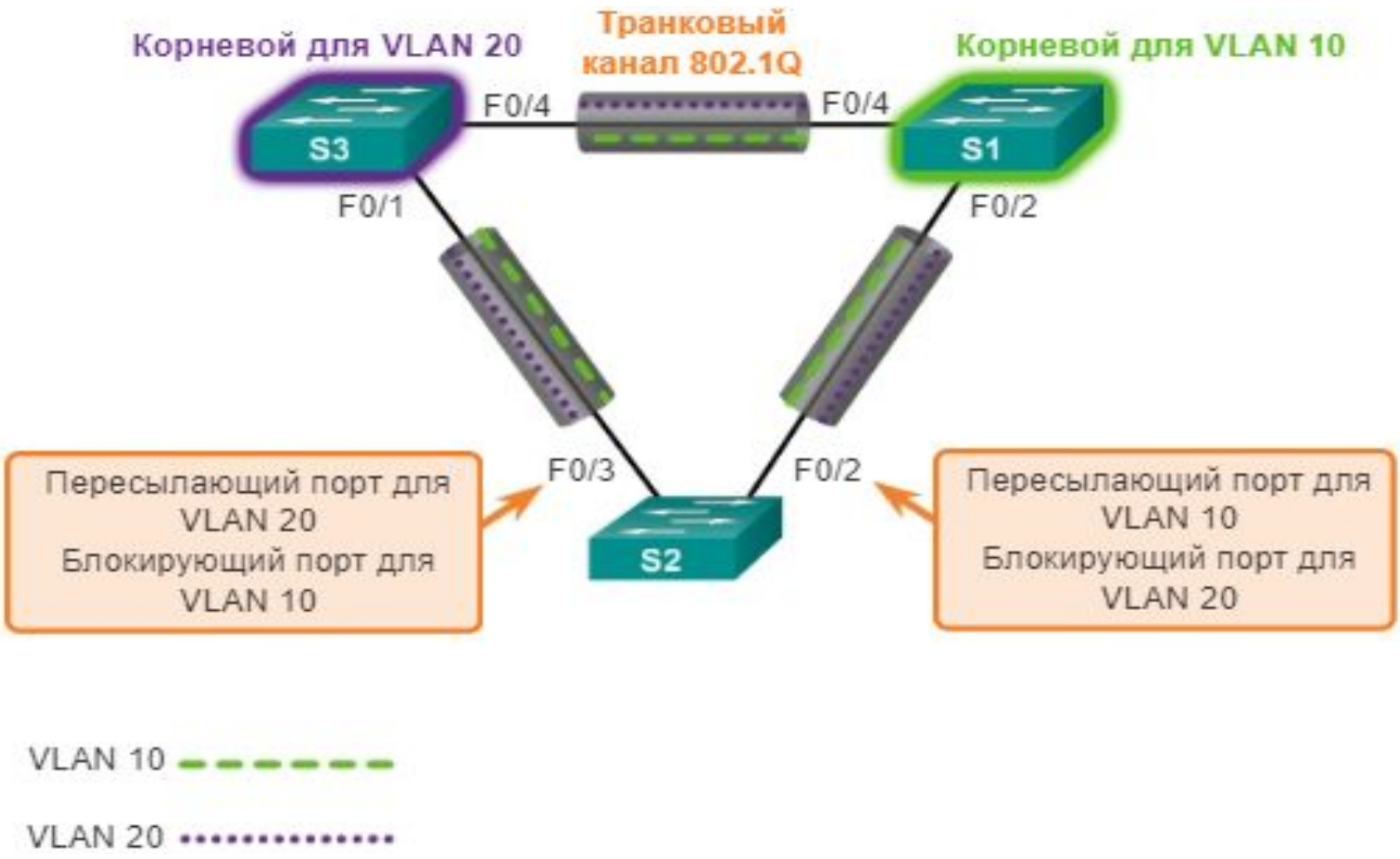
PVST+ можно использовать для распределения нагрузки на 2 уровне. Поскольку все сети VLAN используют отдельный экземпляр STP, коммутаторам в среде PVST+ требуется больший объём ресурсов ЦП и полосы пропускания BPDU, чем в стандартной реализации CST протокола STP.

В среде PVST+ параметры протокола spanning-tree можно настроить таким образом, чтобы половина сетей VLAN выполняла пересылку по всем транковым каналам.

На рисунке (Слайд №81) порт F0/3 на коммутаторе S2 является портом, обеспечивающим передачу данных для сети VLAN 20, а порт F0/2 на коммутаторе S2 — является портом, обеспечивающим передачу данных для сети VLAN 10.

Для этого нужно настроить коммутаторы таким образом, чтобы один был выбран в качестве корневого моста для половины сетей VLAN в пределах сети, а второй — в качестве корневого моста для оставшихся сетей VLAN.

На рисунке коммутатор S3 является корневым мостом для сети VLAN 20, а S1 является корневым мостом для сети VLAN 10.



Несколько корневых мостов STP в одной сети VLAN позволяют увеличить объём избыточности в сети.

Сети под управлением PVST+ имеют следующие характеристики:

1. Поддерживается оптимальное распределение нагрузки.
2. Поддержка одного экземпляра протокола spanning-tree для каждой сети VLAN может привести к значительному необоснованному потреблению ресурсов ЦП для всех коммутаторов в сети (помимо ресурсов полосы пропускания, используемых для отправки собственных кадров BPDU каждым из экземпляров). Это нежелательно только в том случае, если настроено большое количество сетей VLAN.

Состояние протокола и работа протокола PVST+

Протокол STP упрощает создание логического беспетлевого пути по домену широковещательной рассылки.

Протокол spanning-tree определяется с помощью данных, полученных в процессе обмена кадрами BPDU между соединенными друг с другом коммутаторами.

Чтобы упростить процесс получения логического протокола spanning-tree, каждый порт коммутатора проходит через пять возможных состояний порта и три таймера BPDU.

На рисунке представлены состояния портов, обеспечивающих отсутствие петель, при формировании логического протокола spanning-tree:

	Состояние порта				
Операция разрешена	**Блокировка**	**Прослушивание**	**Обучение**	**Пересылка**	**Отключён**
Может получать и обрабатывать BPDU	ДА	ДА	ДА	ДА	НЕТ
Может пересылать кадры данных, полученных на интерфейс	НЕТ	НЕТ	НЕТ	ДА	НЕТ
Может пересылать кадры данных, полученные из другого интерфейса	НЕТ	НЕТ	НЕТ	ДА	НЕТ

1. **Блокирование**: порт является альтернативным и не участвует в пересылке кадров.
2. **Прослушивание**: прослушивание пути к корневому мосту.
3. **Изучение**: изучение MAC-адресов. На этапе подготовки к пересылке кадров порт начинает заполнять таблицу MAC-адресов.
4. **Пересылка**: порт считается частью активной топологии. Он пересылает кадры данных, отправляет и принимает кадры BPDU.
5. **Отключенный**: порт 2 уровня не участвует в протоколе spanning-tree и не пересылает кадры.

Обратите внимание, что число портов в каждом из состояний (блокирование, прослушивание, получение данных или пересылка) можно отобразить с помощью команды **show spanning-tree summary**.

Для обеспечения логической беспетлевой топологии сети для каждой сети VLAN в коммутируемой сети протокол PVST+ выполняет четыре действия:

1. **Выбор одного корневого моста**: только один коммутатор может выступать в роли корневого моста (для данной сети VLAN).
2. **Выбор корневого порта на каждом некорневом мосту**: протокол STP устанавливает один корневой порт на каждом некорневом мосту.
3. **Выбор назначенного порта в каждом сегменте**: в каждом канале протокол STP устанавливает один выделенный порт.
4. **Остальные порты в коммутируемой сети являются альтернативными**: альтернативные порты, как правило, остаются в состоянии блокировки, что позволяет логически разорвать петлевую топологию.

Расширенный идентификатор системы и работы PVST+

В среде PVST+ расширенный идентификатор коммутатора обеспечивает уникальный идентификатор BID для каждого коммутатора в каждой из сетей VLAN

Например, сеть VLAN 2 будет использовать идентификатор BID по умолчанию 32770 (приоритет 32768 плюс расширенный идентификатор системы 2). Если приоритет не задан, коммутаторы будут использовать одинаковое значение приоритета по умолчанию, и выбор корневого моста для каждой сети VLAN будет выполняться на основе MAC-адреса. Этот метод служит для произвольного выбора корневого моста.

В отдельных случаях администратор может выбрать отдельный коммутатор в качестве корневого моста. Тому может быть множество причин, среди которых более централизованное расположение коммутатора в модели сети LAN, более высокая мощность обработки коммутатора или просто более удобный доступ и удалённое управление для данного коммутатора.

Для управления процессом выбора корневого моста следует просто назначить более низкий приоритет коммутатору, который должен быть выбран в качестве корневого моста.

Rapid PVST+

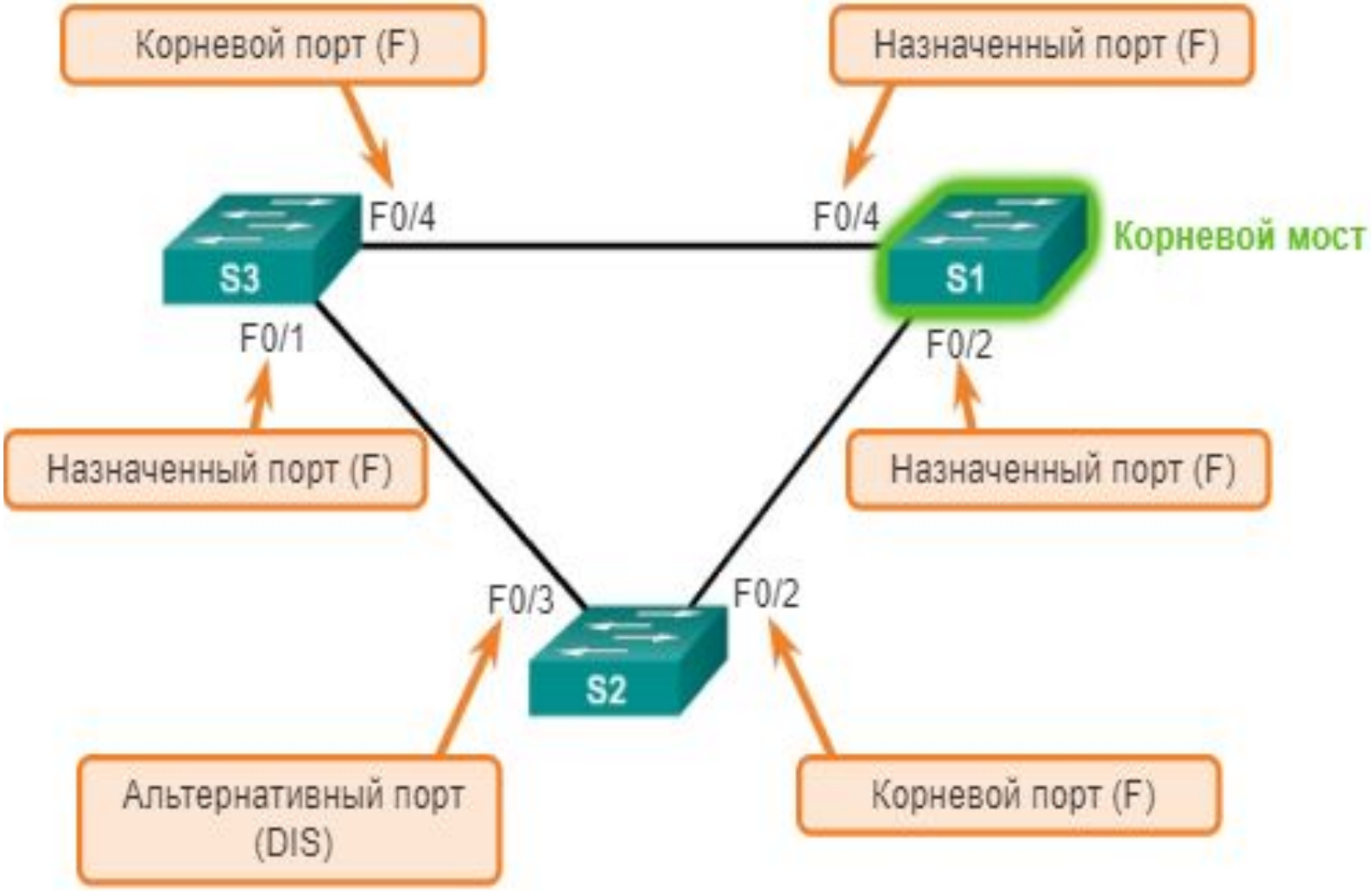
Краткий обзор Rapid PVST+

RSTP (IEEE 802.1w) является развитием исходного стандарт 802.1D; он включен в стандарт IEEE 802.1D-2004. Терминология, относящаяся к STP 802.1w, остается в основном той же, что и для исходного стандарта STP IEEE 802.1D.

Большинство параметров остаются прежними, поэтому пользователи, знакомые с STP, смогут без проблем настроить новый протокол.

Rapid PVST+ — это просто реализация RSTP корпорации Cisco для каждой отдельной сети VLAN. В Rapid PVST+ для каждой сети VLAN запускается самостоятельный экземпляр протокола RSTP.

На рисунке показана сеть под управлением RSTP. S1 является корневым мостом с двумя назначенными портами в состоянии пересылки.



RSTP поддерживает новый тип порта: порт F0/3 на коммутаторе S2 является альтернативным портом в состоянии отбрасывания.

Обратите внимание, что отсутствуют порты, работающие в режиме блокирования. В протоколе RSTP нет состояния блокирования порта. Протокол RSTP определяет следующие состояния портов: отбрасывание, изучение или пересылка.

Протокол RSTP ускоряет повторный расчёт протокола spanning-tree в случае изменения топологии сети 2 уровня. В правильно настроенной сети RSTP может достичь состояния сходимости гораздо быстрее, иногда всего за несколько сот миллисекунд.

Протокол RSTP повторно определяет типы портов и их состояния. Если порт настроен в качестве альтернативного или резервного, он может немедленно перейти в состояние пересылки, не дожидаясь схождения сети.

Далее представлено краткое описание характеристик RSTP:

1. RSTP является предпочтительным протоколом, позволяющим избежать возникновения петель 2 уровня в коммутируемой сети.
2. Проприетарные усовершенствования Cisco для исходного стандарта 802.1D, например, функции UplinkFast и BackboneFast, не совместимы с протоколом RSTP.
3. Протокол RSTP (802.1w) заменяет собой исходный стандарт 802.1D, поддерживая при этом функции обратной совместимости.
4. RSTP сохраняет те же форматы BPDU, что и исходный IEEE 802.1D, за исключением того, что в поле версии установлено значение 2, что указывает на протокол RSTP, а поле флагов задействует все 8 бит.
5. Протокол RSTP может активно подтвердить возможность безопасного перехода порта в состояние пересылки, не полагаясь на конфигурацию таймера.

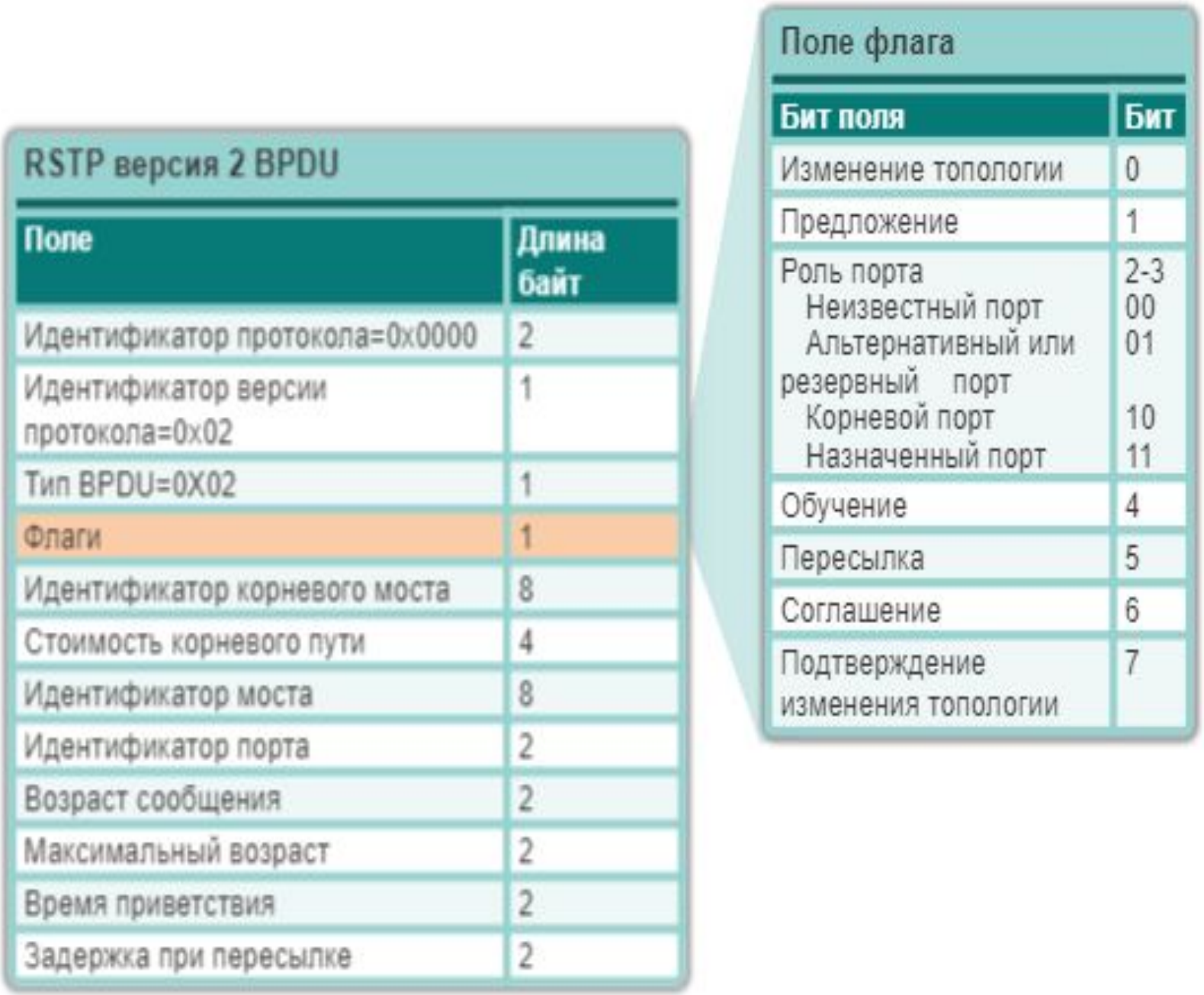
RSTP BPDU

Протокол RSTP использует BPDU типа 2 версии 2. Исходный стандарт 802.1D используется BPDU типа 0 версии 0. Тем не менее, коммутатор под управлением RSTP может осуществлять обмен данными непосредственно с коммутатором под управлением исходного протокола STP 802.1D.

Протокол RSTP отправляет кадры BPDU и заполняет байт флага несколько иначе, чем исходный стандарт 802.1D:

1. Данные протокола на порте могут устареть сразу же, если пакеты приветствия не приняты три раза подряд (по умолчанию — в течение шести секунд) или по истечении максимального времени существования.
2. Поскольку BPDU используется в качестве механизма keepalive, три подряд пропущенных BPDU указывают на потерю соединения между мостом и его соседним корневым мостом или выделенным мостом. Быстрое устаревание данных позволяет быстро обнаруживать сбои.

Как показано на рисунке, протокол RSTP использует байт флага BPDU версии 2:



RSTP версия 2 BPDU

Поле	Длина байт
Идентификатор протокола=0x0000	2
Идентификатор версии протокола=0x02	1
Тип BPDU=0X02	1
Флаги	1
Идентификатор корневого моста	8
Стоимость корневого пути	4
Идентификатор моста	8
Идентификатор порта	2
Возраст сообщения	2
Максимальный возраст	2
Время приветствия	2
Задержка при пересылке	2

Поле флага

Бит поля	Бит
Изменение топологии	0
Предложение	1
Роль порта Неизвестный порт Альтернативный или резервный порт Корневой порт Назначенный порт	2-3 00 01 10 11
Обучение	4
Пересылка	5
Соглашение	6
Подтверждение изменения топологии	7

1. Биты 0 и 7 используются для изменения топологии и подтверждения их поступления в исходный 802.1D.
2. Биты 1 и 6 используются для процесса согласования предложения (для быстрого схождения).
3. Биты со 2 по 5 выполняют кодирование роли и состояния порта.
4. Биты 4 и 5 используются для кодирования роли порта с использованием 2-битного кода.

Пограничные порты

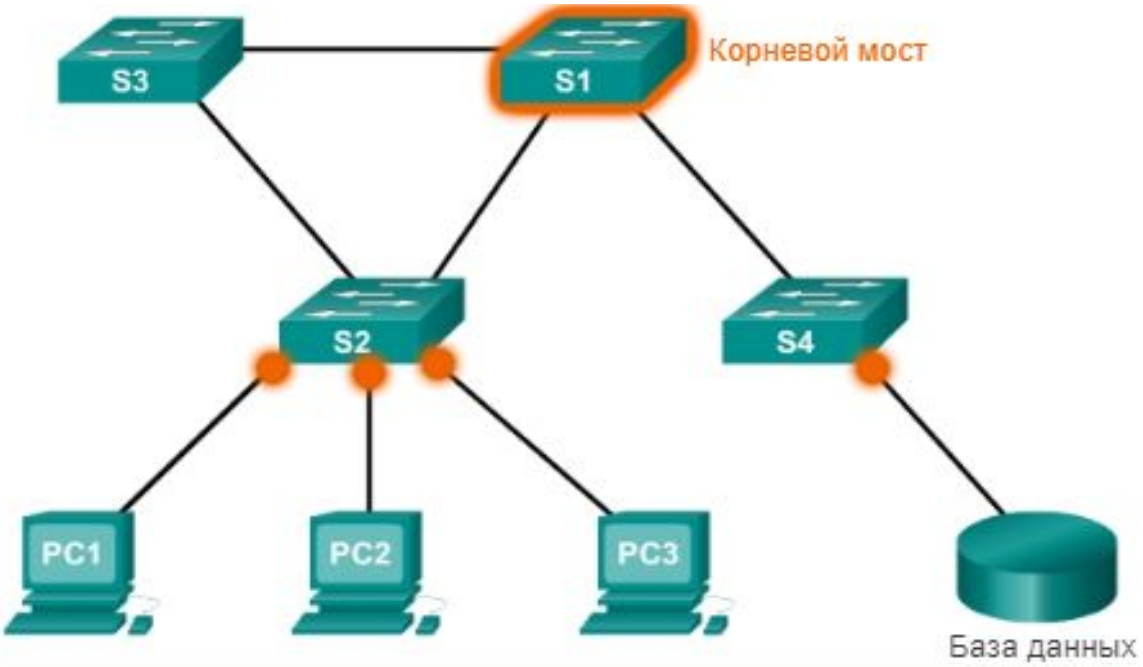
Пограничный порт под управлением RSTP представляет собой порт коммутатора, который никогда не планируется подключать к другому устройству коммутации. После включения он сразу же переходит в состояние пересылки.

Концепция пограничного порта RSTP соответствует функции PVST+ PortFast. Пограничный порт напрямую подключен к конечной станции и предполагает, что к нему не подключено ни одно из устройств коммутации.

Пограничные порты RSTP должны немедленно перейти в состояние пересылки, пропуская, таким образом, состояния прослушивания и изучения исходного 802.1D, которые занимают много времени.

Реализация RSTP Cisco, Rapid PVST+ поддерживает использование ключевого слова PortFast с помощью команды настройки граничного порта **spanning-tree portfast**. Таким образом, переход от STP к RSTP осуществляется без проблем.

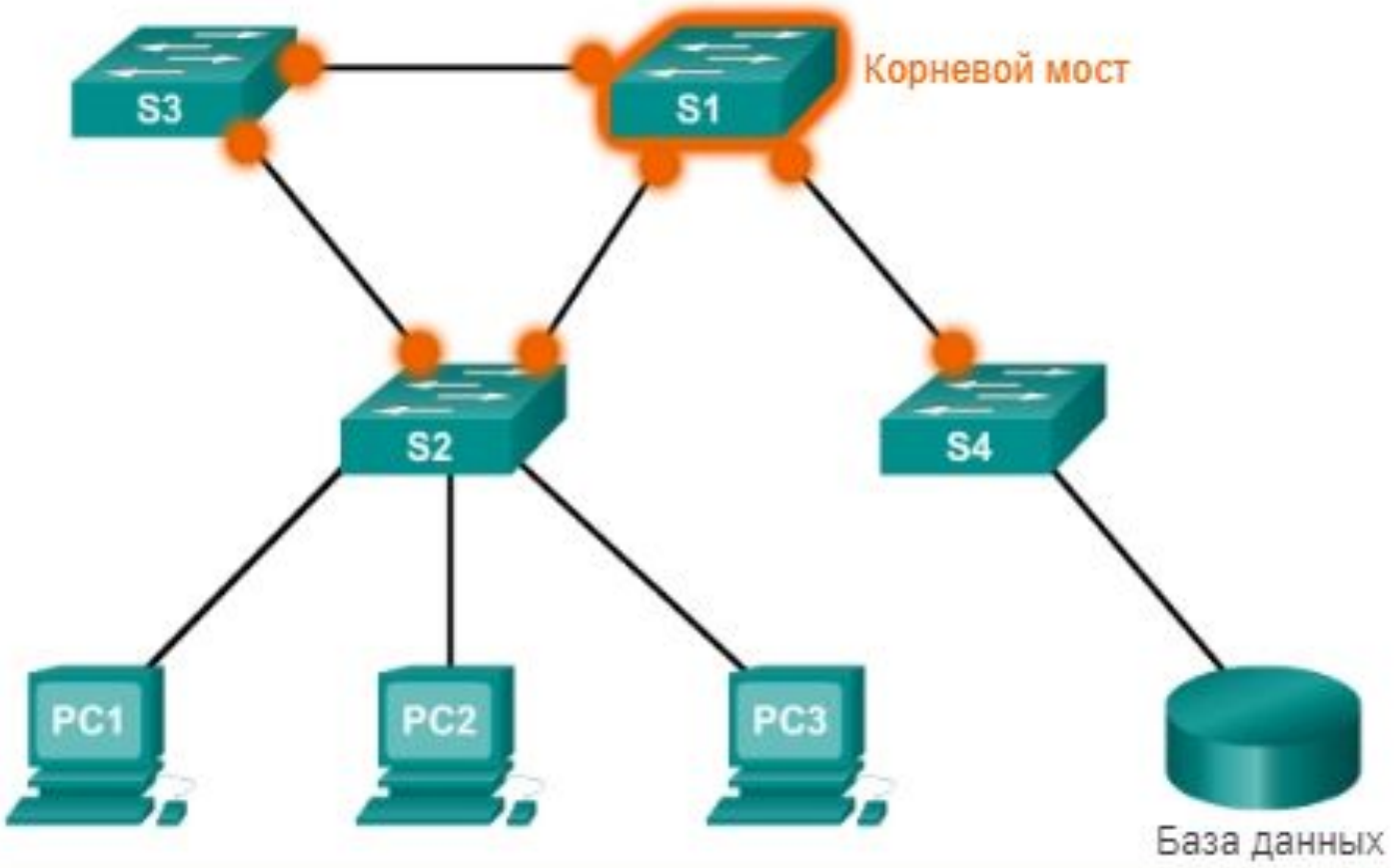
На рисунке показаны примеры портов, которые можно настроить в качестве граничных.



Пограничные порты

- Никогда не будет подключен к коммутатору
- Немедленно переходит в режим пересылки
- Функционирует аналогично порту, настроенному с использованием Cisco PortFast
- На коммутаторе Cisco, настроенном с помощью команды `spanning-tree portfast`

На рисунке показаны примеры портов, отличных от граничных.



Негранничные порты

Порты, которые могут быть соединены с другими устройствами коммутации и не должны быть настроены как граничные порты.

Типы каналов

Тип канала позволяет распределить по категориям каждый порт, участвующий в RSTP на основе дуплексного режима порта.

В зависимости от того, какие устройства подключены к каждому из портов, можно выделить два различных типа каналов:

1. **Точка-точка**: порт, работающий в полнодуплексном режиме; как правило, соединяет два коммутатора и является кандидатом на быстрый переход в состояние пересылки.
2. **Общий**: порт, работающий в полудуплексном режиме; соединяет коммутатор с концентратором, объединяющим несколько устройств.


Корневой порт (F)
Корневой мост
S3
S1
Назначенные порты (F)
Альтернативный порт (DIS)
H1
S2
Корневой порт (F)
Граничный порт
PC1
PC2
PC3
R1

Тип канала позволяет определить, может ли порт сразу перейти в состояние пересылки при условии выполнения определённых условий. Для граничных и неграничных портов требуются разные условия.

Неграничные порты распределяются по категориям в двух типах каналов («точка-точка» и «общий»). Тип канала определяется автоматически, но его можно переопределить с помощью явной конфигурации порта, используя команду **spanning-tree link-type** *parameter*.

Подключения к граничному порту и соединения «точка-точка» претендуют на быстрый переход в состояние пересылки. Тем не менее, прежде чем рассматривать параметр типа канала, RSTP должен определить роль порта.

К характеристикам ролей порта с учетом типов канала относятся следующие:

1. Корневые порты не используют параметр типа канала; корневые порты могут осуществлять быстрый переход в состояние пересылки после синхронизации порта;
2. Альтернативные и резервные порты в большинстве случаев не используют параметр типа канала;
3. Назначенные порты максимально эффективно используют параметр типа канала.

Быстрый переход в состояние пересылки для назначенного порта выполняется только в том случае, если для параметра типа канала установлено значение *point-to-point*.