Сеть Ethernet Построение коммутируемой сети

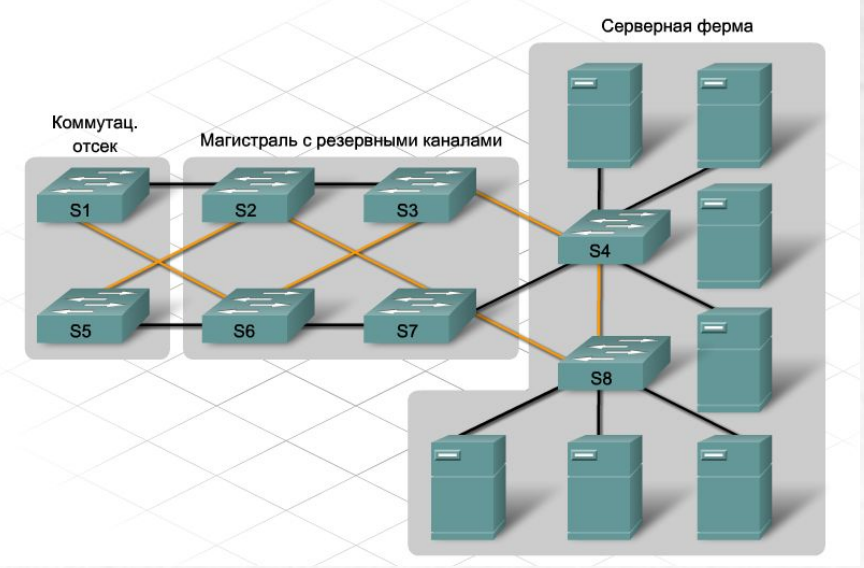
- **Холопов Владимир Анатольевич**

  к.т.н., доцент.

Резервирование в коммутируемой сети

Отказ одного сетевого канала, одного устройства или важного порта коммутатора может стать причиной простоя сети. Чтобы исключить критические точки отказа и обеспечить высокую надежность, в сетевую архитектуру необходимо ввести резервирование. Резервирование реализуется путем установки дублированного оборудования и сетевых устройств в важных областях.

Иногда полное резервирование всех каналов и устройств становится неоправданно дорогим. Сетевые инженеры часто вынуждены искать компромисс между затратами на резервирование и требованиями к доступности сети. Простой сети преобразуется в потенциально катастрофические потери бизнеса, прибыли и доверия заказчиков.



Резервирование в коммутируемой сети

Резервирование обозначает наличие двух разных путей к одному месту назначения. Примеры резервирования в несетевых средах: две дороги в один город, два моста через реку или два выхода в здании. Если один путь заблокирован, второй остается доступным.

Резервирование коммутаторов реализуется путем создания нескольких каналов между ними. Резервные каналы в коммутируемой сети снижают перегрузку и поддерживают высокую доступность и распределение нагрузки.

Однако соединение коммутаторов может стать причиной проблем. В частности, широковещательная природа трафика Ethernet приводит к образованию петель коммутации. Широковещательные кадры циклически распространяются во всех направлениях, вызывая "шторм" широковещательных пакетов. Широковещательные штормы занимают всю доступную полосу пропускания, блокируют создание новых сетевых подключений и разрывают существующие подключения.

Резервирование в коммутируемой сети

Широковещательные штормы — не единственная проблема, обусловленная резервными каналами в коммутируемой сети. Кадры одноадресной пересылки могут вызывать такие проблемы, как множественная передача кадров и нестабильность базы данных MAC-адресов.

Множественная передача кадров

Если узел посылает одноадресный кадр узлу назначения и MAC-адрес не представлен ни в одной из таблиц MAC-адресов подключенных коммутаторов, все коммутаторы выполняют лавинную рассылку этого кадра из всех портов. В сети с петлями кадр может вернуться к исходному коммутатору. Этот процесс повторяется, что приводит к образованию нескольких копий кадра в сети.

В результате узел назначения получает несколько копий кадра. Это становится причиной трех проблем: неэффективное расходование полосы пропускания, неэффективное расходование циклов ЦП и потенциальное дублирование транзакционного трафика.

Нестабильность базы данных MAC-адресов

Коммутаторы в резервируемой сети могут получать неверные данные о положении узла. Если в сети присутствует петля, один коммутатор может связать MAC-адрес назначения с двумя портами. Это приведет к путанице и неоптимальной пересылке кадров.

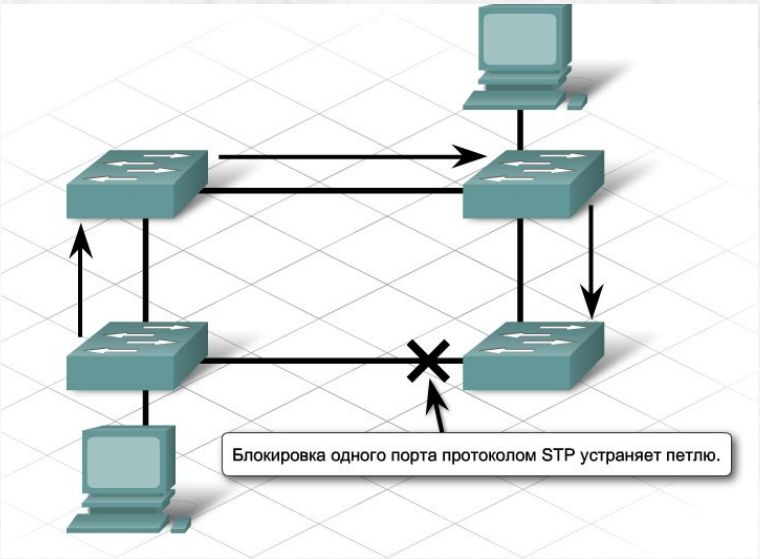
Протокол STP

Протокол STP обеспечивает механизм отключения резервных каналов в коммутируемой сети. **STP** позволяет использовать резервирование, необходимое для надежной эксплуатации, без создания петель коммутации.

STP основывается на открытых стандартах и используется для создания логической топологии без петель коммутации.

Протокол STP относительно самодостаточен и требует минимальной настройки. При первом включении коммутаторы с поддержкой STP проверяют коммутируемую сеть на наличие петель.

Коммутаторы, обнаруживающие потенциальную петлю, блокируют некоторые из подключенных портов, оставляя другие порты активными для пересылки кадров.



Протокол STP

STP задает дерево, которое охватывает все коммутаторы в топологии "иерархическая звезда". Коммутаторы постоянно проверяют сеть, чтобы гарантировать отсутствие петель и эффективную работу всех портов.

Чтобы предотвратить образование петель, протокол STP:

- переводит часть интерфейсов в резервный или заблокированный режим;
- оставляет другие интерфейсы в режиме пересылки;
- перенастраивает сеть, активируя соответствующий резервный путь, если путь пересылки становится недоступным.

В терминологии STP термин "коммутатор" часто заменяется термином "мост". Например, корневой мост — это основной мост или центральная точка в топологии STP. Корневой мост взаимодействует с другими коммутаторами с помощью блоков данных протокола моста (BPDU). BPDU — это кадры, которые рассылаются другим коммутаторам каждые 2 секунды.

Протокол STP

BPDU содержат следующие сведения:

- идентификатор коммутатора-источника;
- идентификатор порта-источника;
- стоимость порта-источника;
- значение таймеров устаревания;
- значение таймера приветствия.

Структура BPDU

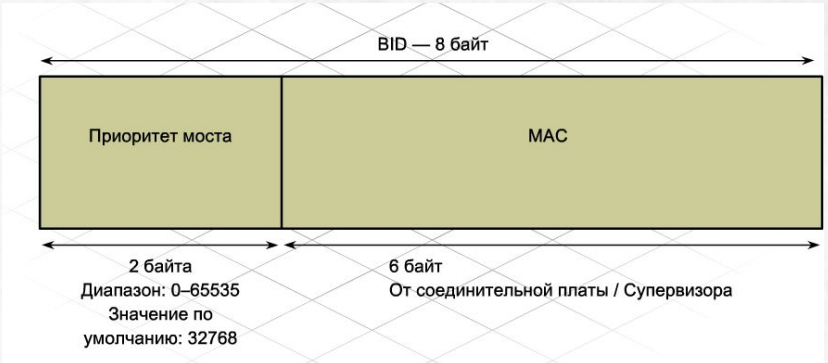
Идентификатор протокола	Версия	Тип сообщения	Флаги	Идентификатор корневого моста	Стоимость корневого пути
Идентификатор моста	Идентификатор порта	Возраст сообщения	Максимальный возраст	Время приветствия	Задержка при пересылке

Корневые мосты протокола STP

Коммутаторы в сети определяют коммутатор, который является центральной точкой сети, чтобы протокол STP мог функционировать. STP использует центральную точку сети, которая называется корневым мостом или корневым коммутатором, для определения портов, которые необходимо блокировать, и портов, которые следует перевести в режим пересылки. Корневой мост рассылает кадры BPDU с информацией о топологии сети всем остальным коммутаторам. Эта информация обеспечивает перенастройку сети в случае отказа.

В каждой сети работает только один корневой мост, который выбирается на основании идентификатора моста (BID). BID равняется сумме значения приоритета моста и его MAC-адреса.

Значение приоритета моста по умолчанию равняется 32 768. Если MAC-адрес коммутатора AA-11-BB-22-CC-33, BID будет равен 32768: AA-11-BB-22-CC-33.

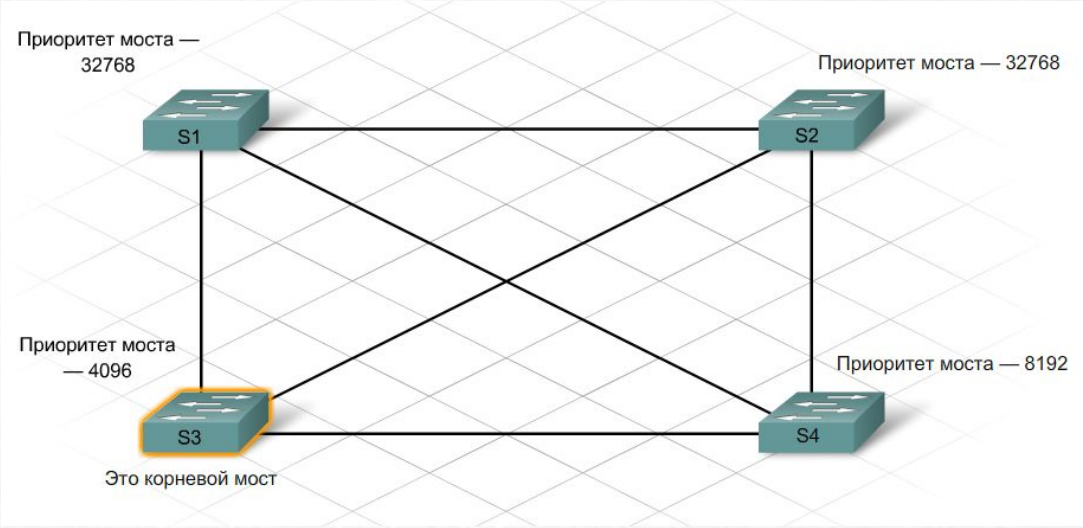


Корневые мосты протокола STP

Мост с наименьшим значением BID становится корневым. Поскольку коммутаторы, как правило, используют одинаковое значение приоритета по умолчанию, коммутатор с наименьшим MAC-адресом становится корневым мостом.

При включении коммутатор предполагает, что является корневым мостом, и рассылает кадры BPDU со своим идентификатором BID. Например, если коммутатор S2 объявляет корневой идентификатор меньше, чем идентификатор S1, S1 прекращает объявление своего идентификатора моста и принимает корневой идентификатор S2. S2 становится корневым мостом.

STP использует три типа портов:
- **корневые порты;**
- **назначенные порты;**
- **заблокированные порты.**



Корневые мосты протокола STP

Корневой порт

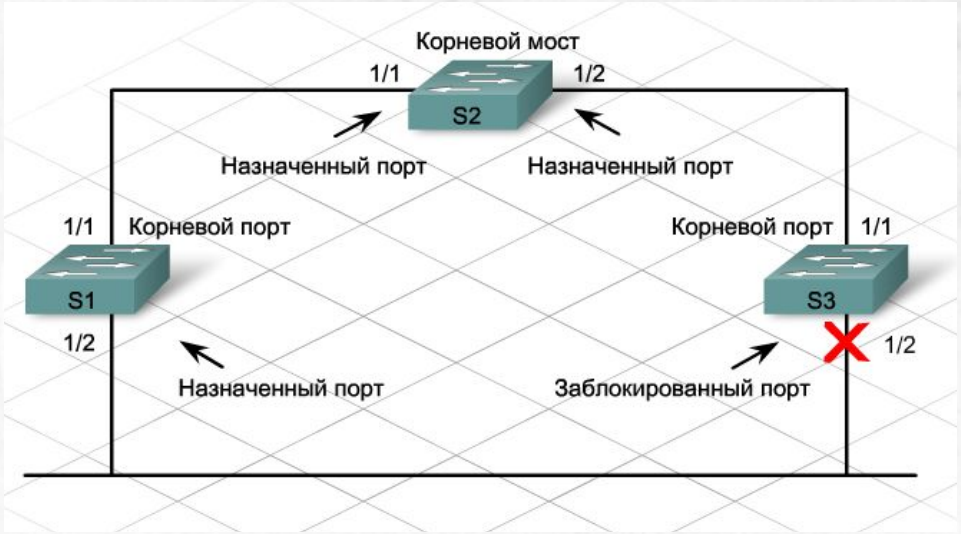
Порт с маршрутом оптимальной стоимости к корневому мосту назначается корневым. Коммутаторы вычисляют путь с наименьшей стоимостью, используя стоимость полосы пропускания каждого канала на пути к корневому мосту.

Назначенный порт

Назначенный порт пересылает трафик к корневому мосту, но не подключен к пути с наименьшей стоимостью.

Заблокированный порт

Заблокированный порт не пересылает трафик.



Протокол STP в иерархической сети

После задания корневого моста, а также корневых, назначенных и заблокированных портов, STP рассылает кадры BPDU по коммутируемой сети с 2-секундным интервалом. STP продолжает отслеживать эти BPDU, чтобы убедиться в отсутствии отказавших каналов и новых петель.

Если происходит отказ канала, STP перерассчитывается путем:

- перевода некоторых портов из блокирующего режима в режим пересылки;
- перевода некоторых портов из режима пересылки в блокирующий режим;
- формирования нового дерева STP для предотвращения образования петель в сети.

Протокол STP не является мгновенным. Когда канал отключается, STP обнаруживает отказ и перерассчитывает наилучшие пути через сеть. Этот расчет и процесс перехода занимают от 30 до 50 секунд для каждого коммутатора. Пользовательские данные не проходят через порты, для которых выполняется перерасчет.

Время ожидания некоторых пользовательских приложений может истечь во время перерасчета, что может привести к снижению производительности и потере прибыли. Частый перерасчет STP негативно влияет на время работы систем.

Протокол STP в иерархической сети

Крупный корпоративный сервер подключен к порту коммутатора. Если для этого порта выполняется перерасчет из-за STP, сервер будет недоступен **в течение 50 секунд.** Трудно представить, какое количество транзакций будет потеряно за это время.

В стабильной сети перерасчеты STP редки. Если сеть нестабильна, необходимо проверить стабильность коммутаторов и изменения их конфигураций. Одна из самых распространенных причин перерасчетов STP — неисправный источник питания или кабель питания коммутатора. Неисправность источника питания вызывает неожиданную перезагрузку устройства.

Протокол STP в иерархической сети

Когда институт IEEE разработал оригинальный протокол STP (Spanning Tree Protocol) 802.1D, период восстановления 1 или 2 минуты был допустимым. Сегодня коммутация уровня 3 и усовершенствованные протоколы маршрутизации обеспечивают более быстрые альтернативные пути к месту назначения. Из-за потребности в передаче трафика, чувствительного к задержкам, например голоса и видео, коммутируемые сети должны поддерживать быструю конвергенцию, чтобы удовлетворять требованиям новых технологий.

Протокол Rapid Spanning Tree Protocol (RSTP), определенный в стандарте IEEE 802.1w, значительно ускоряет перерасчет STP. В отличие от функций PortFast, UplinkFast и BackboneFast, протокол RSTP не является собственным.

Протокол STP в иерархической сети

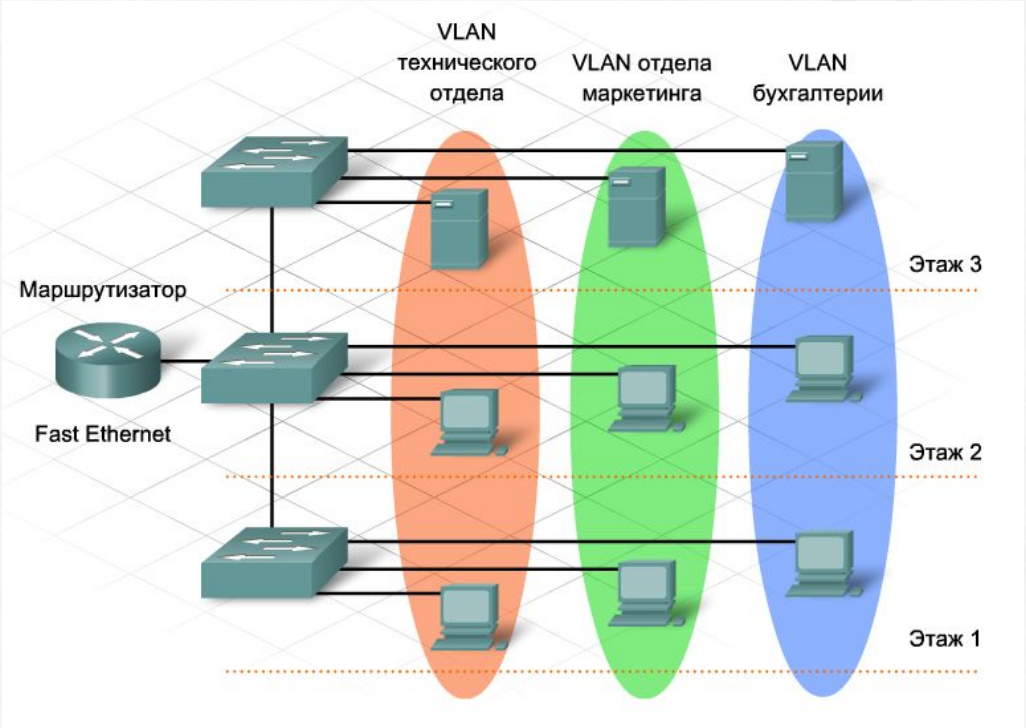
Для обеспечения максимальной скорости переконфигурации протокол RSTP требует полнодуплексного соединения "точка-точка" между коммутаторами. Переконфигурация связующего дерева при использовании протокола RSTP занимает менее одной секунды, аналогичный процесс протокола STP занимает 50 секунд.

RSTP устраняет потребность в таких функциях, как PortFast и UplinkFast. RSTP может переключаться в режим STP для обслуживания старого оборудования.

Для ускорения перерасчета число режимов портов протокола RSTP уменьшено до трех: отклонение, обучение и пересылка. Режим отбрасывания аналогичен трем оригинальным режимам STP: блокировки, обучения и "отключен".
Кроме того, в RSTP добавлена концепция активная топология. Все порты, которые не находятся в режиме сброса (не заблокированы), считаются частью активной топологии и немедленно переходят в режим пересылки.

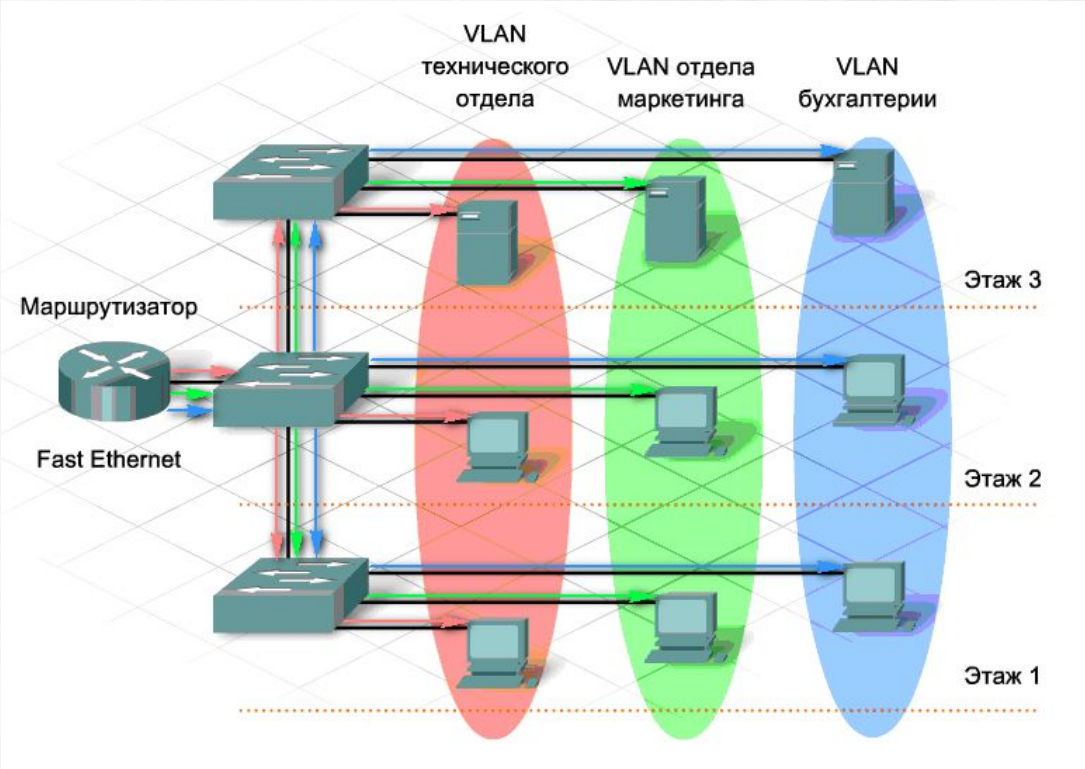
Виртуальные локальные сети (VLAN)

VLAN — это логический домен широковещательной рассылки, который может охватывать несколько физических сегментов LAN. Она позволяет администратору объединять станции по логической функции, проектной группе или приложению независимо от физического положения пользователей.



Виртуальные локальные сети (VLAN)

Этот пример также демонстрирует другую функцию VLAN. Широковещательные кадры не пересылаются между VLAN, они остаются внутри одной VLAN.



Виртуальные локальные сети (VLAN)

Виртуальная LAN. Общая информация:

- VLAN представляет собой несколько хостов, сгруппированных в логическую цепь, для образования отдельной сети;
- каждая VLAN имеет собственный домен;
- стандарт 802.1q позволяет использовать различные сети VLAN в одном транке;
- элементы сети одной VLAN обладают одним доменом и взаимодействуют между собой без участия маршрутизатора;
- для взаимодействия между VLAN и сетями третьего уровня необходимо дополнительное устройство, к примеру, маршрутизатор.

IEEE 802.1Q — VLAN Tagging

Стандарт IEEE 802.1Q или VLAN Tagging, позволяет передавать множество разделенных между собой сетей, через единое физическое соединение, без обмена информацией между собою.

Порты на коммутаторе могут работать в двух режимах "access“ (Edge) или "trunk":

Access (Edge)

Данное значение следует устанавливать, в случаях, когда порт коммутатора принадлежит к одной VLAN. VID, назначенный порту, добавляется ко всем приходящим на порт пакетам. Этот VID будет присутствовать в пакете до тех пор, пока пакет не дойдет до порта получателя, где метка VID будет удалена из пакета и пакет отправится получателю.

IEEE 802.1Q — VLAN Tagging

Стандарт IEEE 802.1Q или VLAN Tagging, позволяет передавать множество разделенных между собой сетей, через единое физическое соединение, без обмена информацией между собою.

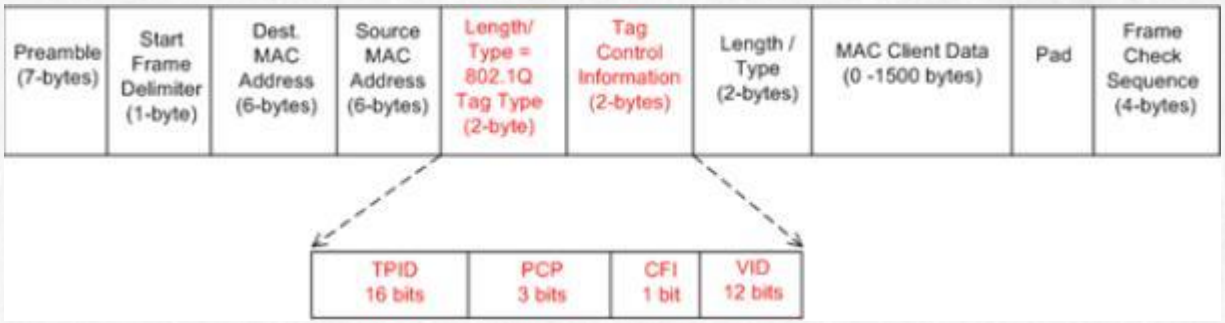
Trunk

В такой конфигурации, порт коммутатора может передавать несколько VLAN сетей, через одно физическое соединение. В таком режиме работы, 802.1Q использует собственный механизм пометки пакетов, который вставляет в оригинальный пакет Ethernet 4-х байтовое поле, между полями адреса отправителя (Source Address) и типа пакета (Type/Length). Поскольку поля изменяются, коммутатор также пересчитывает контрольную сумму всего пакета (FCS). Этот процесс автоматически выполняется коммутатором, перед тем, как пакет отправляется по магистрали. Коммутатор, который получит такой пакет, на другом конце магистрали, удалит из него эту пометку, пересчитает контрольную сумму, а затем отправляется в соответствующую сеть VLAN.

IEEE 802.1Q — VLAN Tagging

802.1Q не выполняет инкапсуляцию пакетов. В место этого, он добавляет в пакет 32-х битное поле, между полями MAC адреса отправителя и полями длинна/тип пакета.

Поле VLAN tag, имеет следующий формат.



IEEE 802.1Q — VLAN Tagging

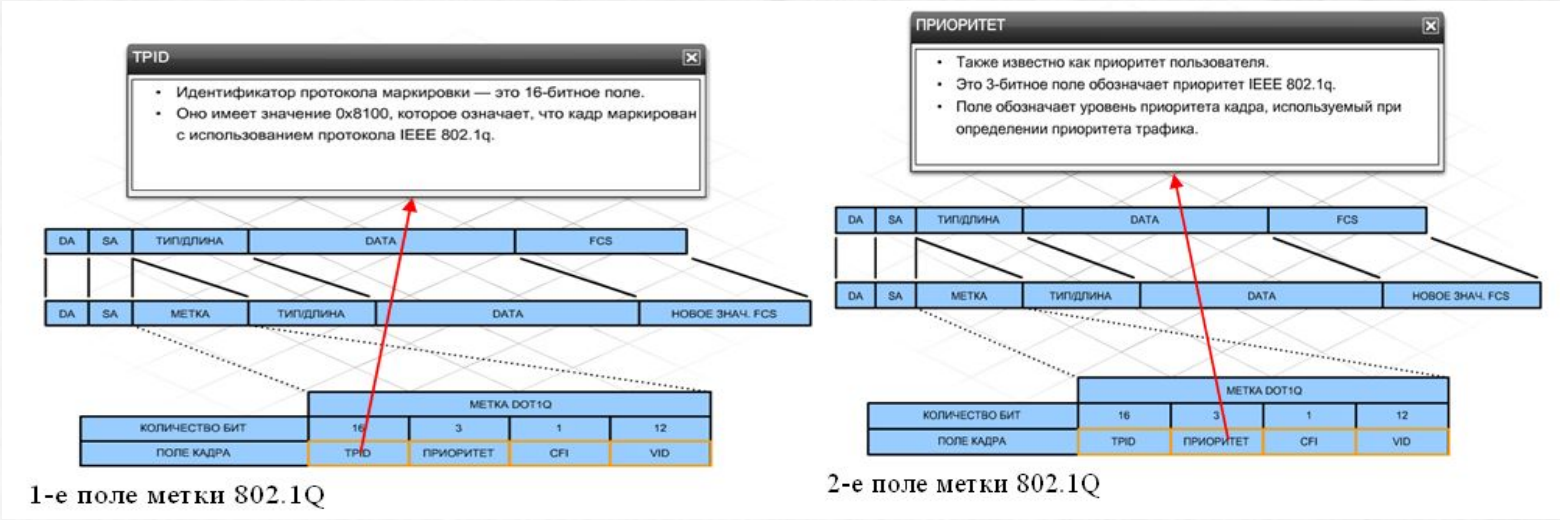
Tag Protocol Identifier (TPID) – 16-ти битное поля идентификатора версии протокола.

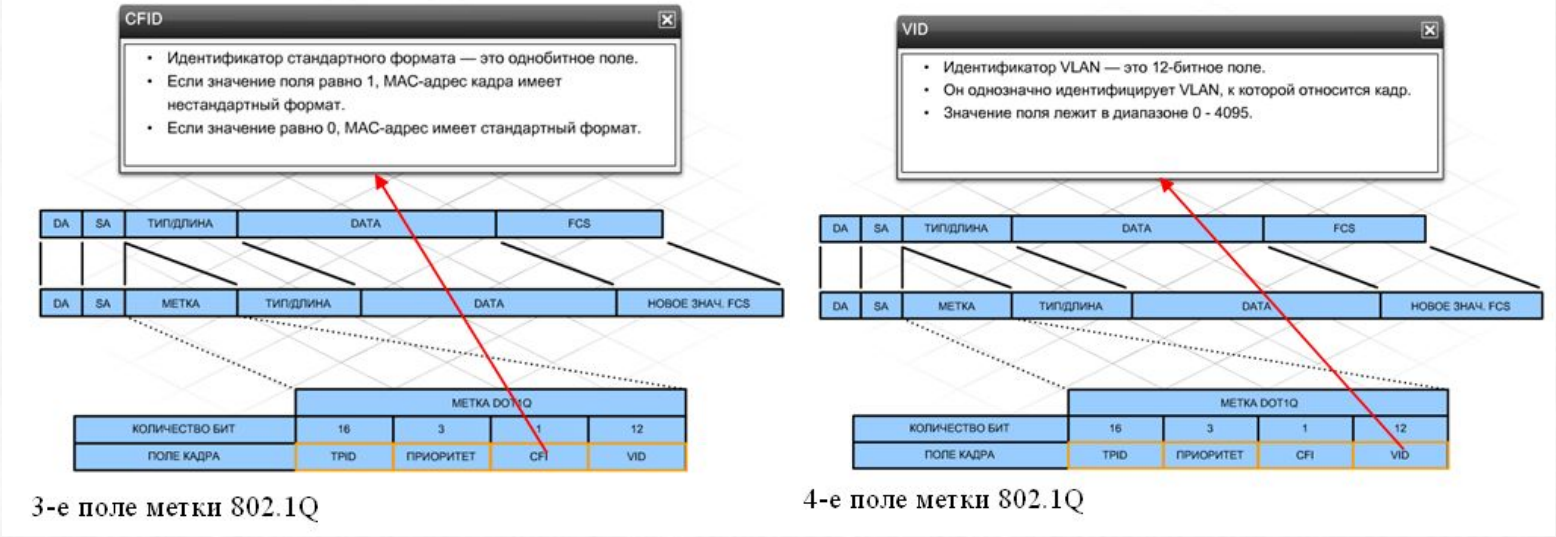
Priority Code Point (PCP) – 3-х битное поле, которое относится к приоритету различных видов трафика, по стандарту IEEE 802.1p, не является обязательным полем для VLAN Tagging.

Canonical Format Indicator (CFI) – Коммутаторы Ethernet заполняют это поле нулями. Данное поле, используется для организации совместимости между сетями Ethernet и Token Ring.

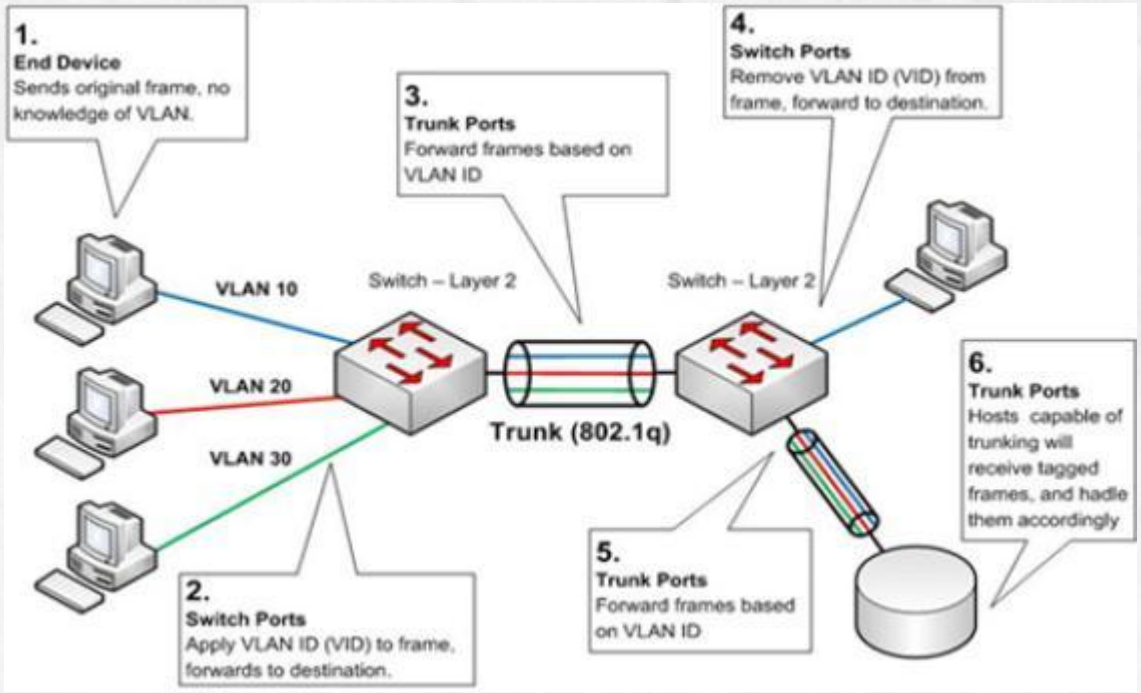
VLAN ID (VID) – Это 12-битное поле, идентификатора VLAN (1 - 4094) к которой принадлежит пакет. Если значение установлено в ноль, это означает, что пакет, не принадлежит ни к какой VLAN и метка 802.1Q указывает только на приоритет. Значение FFF, зарезервировано. На шлюзах, VLAN 1 используется для управления.

IEEE 802.1Q — VLAN Tagging





IEEE 802.1Q — VLAN Tagging



802.1Q не помечает оригинальные VLAN пакеты, передаваемые через магистраль в режиме Trunk. Он помечает все другие пакеты, передаваемые через магистраль. Когда настраивается магистраль 802.1Q, коммутаторы на обеих сторонах, соединения должны работать в стандартном режиме формирования магистрали VLAN.

IEEE 802.1Q — VLAN Tagging

Типы портов VLAN			
Тип порта	Untagged	Tagged	Передаваемые пакеты
Edge Access	Да	Нет	Обычные пакеты своего VLAN
Edge Access	Нет	Да	Tag пакеты своего VLAN
Trunk	Да	Нет	Обычные пакеты своего VLAN и Tag пакеты других VLAN
Trunk	Нет	Да	Tag пакеты своего и других VLAN

Спасибо за внимание

Холопов В.А.