

Организация администрирования компьютерных систем

Лекции.

Тема 4. Функции повышения
надежности и производительности

В настоящее время для повышения надежности и производительности каналов связи в распоряжении интеграторов и сетевых администраторов имеется целый набор протоколов и функций.

Наиболее распространенным является создание резервных связей между коммутаторами на основе двух технологий:

- резервирование соединений с помощью протоколов семейства Spanning Tree;
- балансировка нагрузки, обеспечивающая параллельную передачу данных по всем альтернативным соединениям с помощью механизма агрегирования портов.

4.1. Протоколы Spanning Tree

Протокол связующего дерева Spanning Tree Protocol (STP) является протоколом 2 уровня модели OSI, который позволяет строить древовидные, свободные от петель, конфигурации связей между коммутаторами локальной сети.

Помимо этого, алгоритм обеспечивает возможность автоматического резервирования альтернативных каналов связи между коммутаторами на случай выхода активных каналов из строя.

В настоящее время существуют следующие версии протоколов связующего дерева:

- IEEE 802.1D Spanning Tree Protocol (STP);
- IEEE 802.1w Rapid Spanning Tree Protocol (RSTP);
- IEEE 802.1s Multiple Spanning Tree Protocol (MSTP).

Spanning Tree Protocol (STP)

Если для обеспечения избыточности между коммутаторами создается несколько соединений, то могут возникать коммутационные петли. Петля предполагает существование нескольких маршрутов по промежуточным сетям, а сеть с несколькими маршрутами между источником и приемником отличается повышенной отказоустойчивостью. Хотя наличие избыточных каналов связи очень полезно, петли, тем не менее, создают проблемы, самые актуальные из которых:

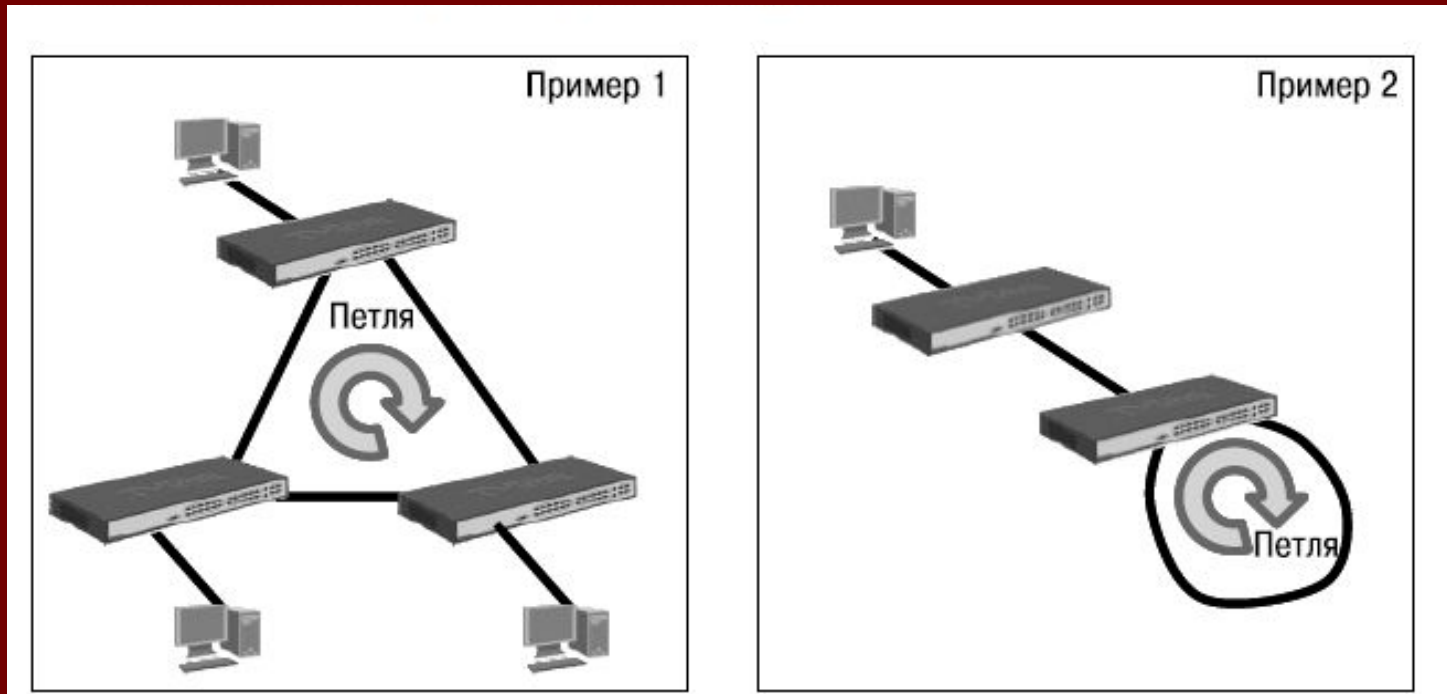
- широковещательные штормы;
- множественные копии кадров;
- множественные петли.

Для решения этих проблем и был разработан протокол связующего дерева, который был определен в стандарте IEEE 802.1D-1998.

Коммутаторы, поддерживающие протокол STP, автоматически создают древовидную конфигурацию связей без петель в компьютерной сети. Такая конфигурация называется связующим деревом — Spanning Tree (иногда ее называют остовым или покрывающим деревом). Конфигурация связующего дерева строится коммутаторами автоматически с использованием обмена служебными кадрами, называемыми *Bridge Protocol Data Units (BPDU)*

Широковещательный шторм

- Распространение широковещательных сообщений в сетях с петлями представляет серьезную проблему. Предположим, что первый кадр, поступивший от одного из узлов, является широковещательным. Тогда все коммутаторы будут пересылать кадры бесконечно (пример 1), используя всю доступную полосу пропускания сети и блокируя передачу других кадров во всех сегментах.



Множественные копии кадров

- Еще одна проблема заключается в том, что коммутатор нередко получает несколько копий одного кадра, одновременно приходящих из нескольких участков сети.
- В этом случае таблица коммутации не сможет определить расположение устройства, потому что коммутатор будет получать кадр из нескольких каналов. Может случиться так, что коммутатор вообще не сможет переслать кадр, т.к. будет постоянно обновлять таблицу коммутации.

Множественные петли

- Одна из самых сложных проблем — это множественные петли, образующиеся в объединенной сети. Возможно появление петли внутри других петель. Если за этим последует ширококвещательный шторм, то сеть не сможет выполнять коммутацию кадров

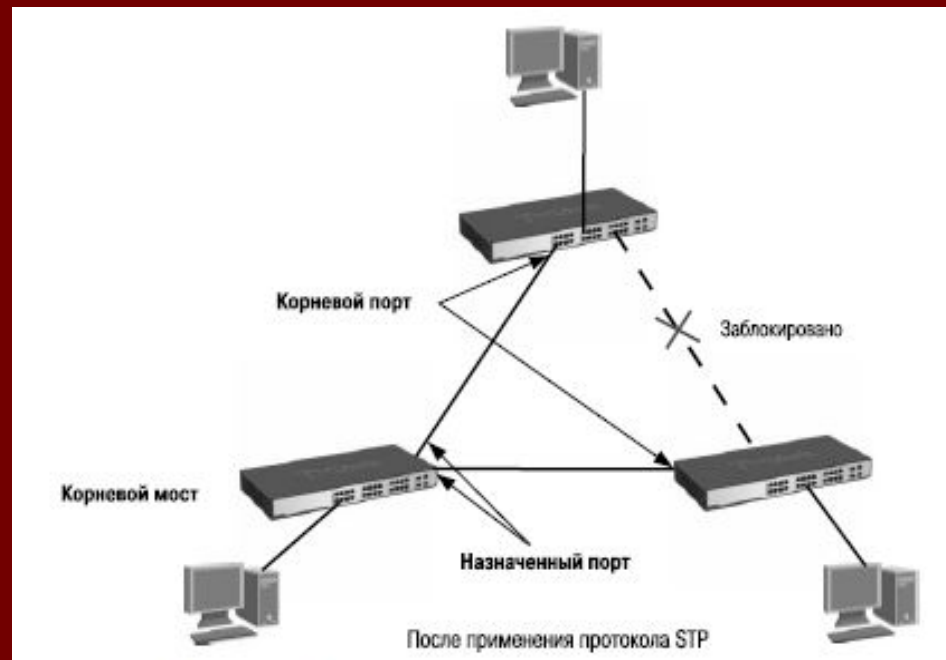
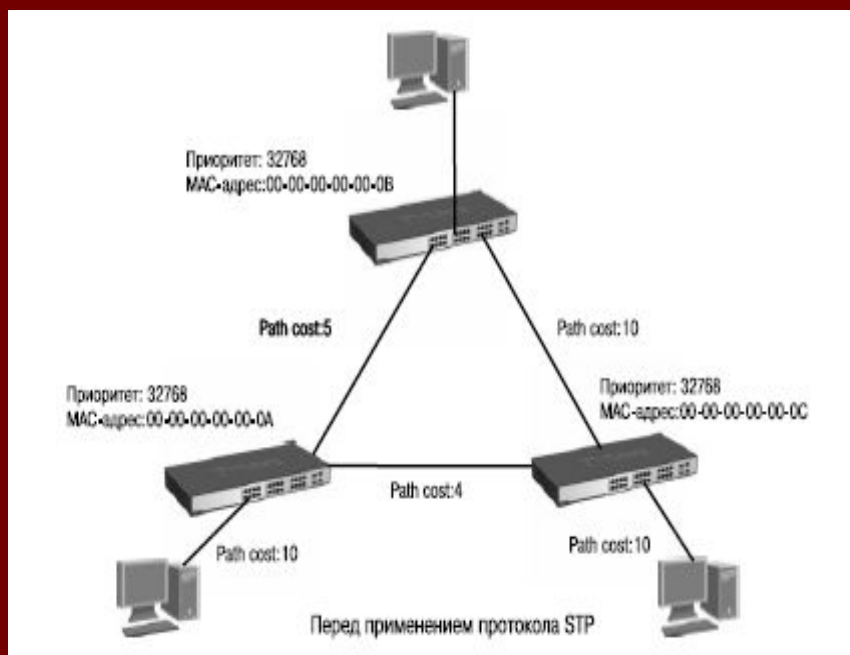
Построение активной топологии связующего дерева

- Для построения устойчивой активной топологии с помощью протокола STP необходимо с каждым коммутатором сети ассоциировать уникальный *идентификатор моста (Bridge ID)*, а с каждым портом коммутатора ассоциировать *стоимость пути (Path Cost)* и *идентификатор порта (Port ID)*.
- Реализуется в три этапа:
 - Вычисление связующего дерева
 - Выбор корневых портов (*Root Port*).
 - Определение назначенных портов (*Designated Port*)

Процесс вычисления связующего дерева

- **Процесс вычисления связующего дерева начинается** с выбора корневого моста (*Root Bridge*), от которого будет строиться дерево. В качестве корня дерева выбирается коммутатор с наименьшим значением идентификатора моста. Идентификатор моста — это 8-байтное поле, которое состоит из 2-х частей: приоритета моста (2 байта), назначаемого администратором сети, и MAC-адреса блока управления коммутатора (6 байт). При сравнении идентификаторов двух коммутаторов сначала сравниваются значения приоритетов. Корневым мостом становится коммутатор с наименьшим значением приоритета. Если они одинаковы (по умолчанию приоритет равен 32768), то корневой мост определяется по наименьшему MAC-адресу.
- Для того чтобы в качестве корневого моста было выбрано определенное устройство (исходя из структуры сети), администратор может вручную назначить соответствующему коммутатору наименьший приоритет.

Пример функционирования протокола STP



Выбор корневых портов

- **Второй этап** работы STP — выбор корневых портов (*Root Port*).
- Когда процесс выбора корневого моста завершен, оставшиеся коммутаторы сети определяют стоимость каждого возможного пути от себя до корня дерева. Стоимость пути рассчитывается как *суммарное условное время* на передачу данных от порта данного коммутатора до порта корневого моста. Условное время сегмента рассчитывается как время передачи одного бита информации через канал с определенной полосой пропускания. Стоимости пути по умолчанию для каждого канала определены в стандарте IEEE 802.1D-1998.
- Сравнив стоимости всех возможных маршрутов до корня, каждый коммутатор выбирает среди них один с наименьшим значением стоимости. Порт, соединяющий коммутатор с этим маршрутом, становится корневым портом. В случае если минимальные стоимости пути нескольких маршрутов окажутся одинаковыми, корневым портом станет порт, имеющий наименьшее значение идентификатора порта.

Определение назначенных портов

- **Третий шаг** работы STP — определение назначенных портов (Designated Port).
- Каждый сегмент в коммутируемой сети имеет один назначенный порт. Этот порт функционирует как единственный порт моста, т.е. принимает кадры от сегмента и передает их в направлении корневого моста через корневой порт данного коммутатора. Коммутатор, содержащий назначенный порт для данного сегмента, называется назначенным мостом (*Designated Bridge*) этого сегмента. Назначенный порт сегмента определяется путем сравнения значений стоимости пути всех маршрутов от данного сегмента до корневого моста. Им становится порт, имеющий наименьшее значение стоимости, среди всех портов, подключенных к данному сегменту. Если минимальные значения стоимости пути окажутся одинаковыми у двух или нескольких портов, то для выбора назначенного порта сегмента STP принимает решение на основе последовательного сравнения идентификаторов мостов и идентификаторов портов.
- У корневого моста все порты являются назначенными, а их расстояние до корня полагается равным нулю. Корневого порта у корневого моста нет.
- После выбора корневых и назначенных портов все остальные порты коммутаторов сети переводятся в состояние Blocking («Блокировка»), то есть такое, при котором они принимают и передают только кадры BPDU. При таком выборе активных портов в сети исключаются петли, и оставшиеся связи образуют связующее дерево.

Bridge Protocol Data Unit (BPDU)

- Вычисление связующего дерева происходит при включении коммутатора и при изменении топологии. Эти вычисления требуют периодического обмена информацией между коммутаторами связующего дерева, что достигается при помощи специальных кадров, называемых блоками данных протокола моста — BPDU (Bridge Protocol Data Unit).
- Коммутатор отправляет BPDU, используя уникальный MAC-адрес порта в качестве адреса-источника и многоадресный MAC-адрес протокола STP 01-80-C2-00-00-00 в качестве адреса-приемника. Кадры BPDU помещаются в поле данных кадров канального уровня, например, кадров Ethernet.
- **Внимание:** иногда, с целью повышения безопасности, сетевым администраторам необходимо отключать возможность передачи кадров BPDU на граничные коммутаторы сети, чтобы избежать получения случайных кадров BPDU клиентскими портами, которые могут распространить вычисления STP по клиентским сетям. Управляемые коммутаторы D-Link поддерживают возможность включения и отключения передачи кадров BPDU для каждого порта

Типы кадров BPDU

- Существует три типа кадров BPDU:
 - Configuration BPDU (CBPDU) — конфигурационный кадр BPDU, который используется для вычисления связующего дерева (тип сообщения: 0x00);
 - Topology Change Notification (TCN) BPDU — уведомление об изменении топологии сети (тип сообщения: 0x80);
 - Topology Change Notification Acknowledgement (TCA) — подтверждение о получении уведомления об изменении топологии сети.
- Коммутаторы обмениваются BPDU через равные интервалы времени (по умолчанию 2 сек.), что позволяет им отслеживать состояние топологии сети.

Формат кадра BPDU

	Байты
Идентификатор протокола (Protocol Identifier)	2
Версия протокола (Protocol Version Identifier)	1
Тип BPDU (BPDU Type)	1
Флаги (Flags)	1
Идентификатор корневого моста (Root Identifier)	8
Расстояние до корневого моста (Root Path Cost)	2
Идентификатор моста (Bridge Identifier)	8
Идентификатор порта (Port Identifier)	2
Время жизни сообщения (Message Age)	2
Максимальное время жизни сообщения (Max Age)	2
Время приветствия (Hello Time)	2
Задержка смены состояний (Forward Delay)	2

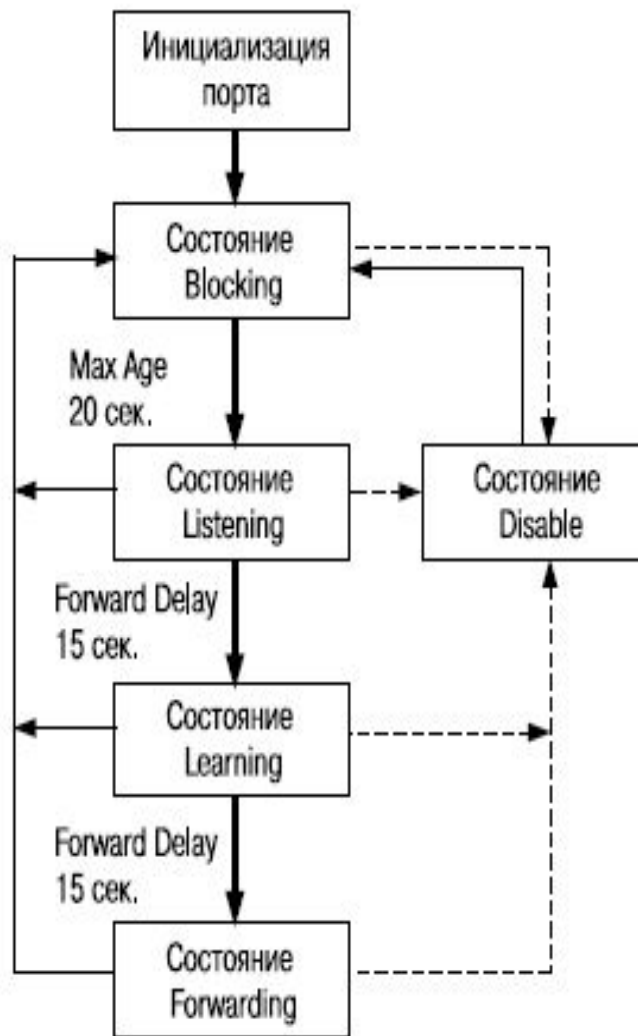
Поля кадра BPDU

- идентификатор протокола (Protocol Identifier) — 2 байта. Значение всегда равно 0;
- версия протокола STP (Protocol Version Identifier) — 1 байт. Значение всегда равно 0;
- тип BPDU (BPDU Type) — 1 байт. Значение «00» — конфигурационный BPDU, «01» — изменение топологии;
- флаги (Flags) — 1 байт. Бит 1 — флаг изменения топологии, бит 8 — флаг подтверждения изменения топологии;
- идентификатор корневого моста (Root Identifier) — 8 байтов. Идентификатор текущего корневого моста;
- расстояние до корневого моста (Root Path Cost) — 2 байта. Суммарная стоимость пути до корневого моста;
- идентификатор моста (Bridge Identifier) — 8 байтов. Идентификатор текущего моста;
- идентификатор порта (Port Identifier) — 2 байта. Уникальный идентификатор порта, который отправил этот BPDU;

Поля кадра BPDU

- время жизни сообщения (Message Age) — 2 байта. Нефиксированный временной интервал в секундах, прошедший с момента отправки BPDU корневым мостом. Служит для выявления устаревших сообщений BPDU. Первоначальное значение равно нулю. По мере передачи кадра BPDU по сети каждый коммутатор добавляет ко времени жизни сообщения время его задержки данным коммутатором. По умолчанию оно равно 1 сек. Значение параметра Message Age должно быть меньше значения таймера Max Age;
- максимальное время жизни сообщения (Max Age) — 2 байта. Временной интервал в секундах, определяющий максимальное время хранения конфигурации STP, прежде чем коммутатор ее отбросит;
- время приветствия (Hello Time) — 2 байта. Временной интервал в секундах, через который посылаются кадры BPDU;
- задержка смены состояний (Forward Delay) — 2 байта. Временной интервал в секундах, в течение которого порт коммутатора находится в состояниях «Прослушивание» и «Обучение».

Состояния портов



Состояния портов

- *Blocking* («Блокировка») — при инициализации коммутатора все порты (за исключением отключенных) автоматически переводятся в состояние «Заблокирован». В этом случае порт принимает и обрабатывает только кадры BPDU. Все остальные кадры отбрасываются;
- *Listening* («Прослушивание») — в этом состоянии порт продолжает принимать, обрабатывать и ретранслировать только кадры BPDU. Из этого состояния порт может перейти в состояние «Заблокирован», если получит BPDU с лучшими параметрами, чем его собственные (стоимость пути, идентификатор моста или порта). В противном случае, при истечении периода, установленного таймером задержки смены состояний (Forward Delay), порт перейдет в следующее состояние «Обучение»;
- *Learning* («Обучение») — порт начинает принимать все кадры и на основе MAC-адресов источника строить таблицу коммутации. Порт в этом состоянии все еще не продвигает кадры. Порт продолжает участвовать в работе алгоритма STP и при поступлении BPDU с лучшими параметрами переходит в состояние «Заблокирован». В противном случае, при истечении периода, установленного таймером задержки смены состояний, порт перейдет в следующее состояние «Продвижение»;

Состояния портов

- *Forwarding* («Продвижение») — в этом состоянии порт может обрабатывать кадры данных в соответствии с построенной таблицей коммутации. Также продолжают приниматься, передаваться и обрабатываться кадры BPDU;
- *Disable* («Отключен») — в это состояние порт переводит администратор. Отключенный порт не участвует ни в работе протокола STP, ни в продвижении кадров данных. Порт можно также вручную включить, и первоначально он перейдет в состояние «Заблокирован».

В процессе нормальной работы корневой мост продолжает генерировать служебные кадры BPDU, а остальные коммутаторы продолжают их принимать своими корневыми портами и ретранслировать назначенными. Если по истечении максимального времени жизни сообщения (по умолчанию — 20 секунд) корневой порт любого коммутатора сети не получает служебный кадр BPDU, то он инициализирует новую процедуру построения связующего дерева.

Таймеры STP (используются для того, чтобы все коммутаторы сети имели возможность получить точную информацию о конфигурации связующего дерева)

- *Hello Time* — это интервал времени, через который корневой мост отправляет конфигурационные BPDU. Значение таймера Hello Time, настроенное на корневом мосте, будет определять значения таймеров Hello Time на всех некорневых коммутаторах, т.к. они просто пересылают конфигурационные BPDU, когда получают их от корня. Значение таймера Hello Time по умолчанию 2 секунды: диапазон возможных значений от 1 до 10 секунд.
- *Forward Delay* — это интервал времени, в течение которого порт коммутатора находится в состояниях «Прослушивание» и «Обучение». Такая задержка смены состояний необходима, чтобы исключить возможность временного возникновения альтернативных маршрутов при одновременной смене состояний портов во время реконфигурации. Значение таймера Forward Delay по умолчанию 15 секунд. диапазон возможных значений от 4 до 30 секунд.

Таймеры STP

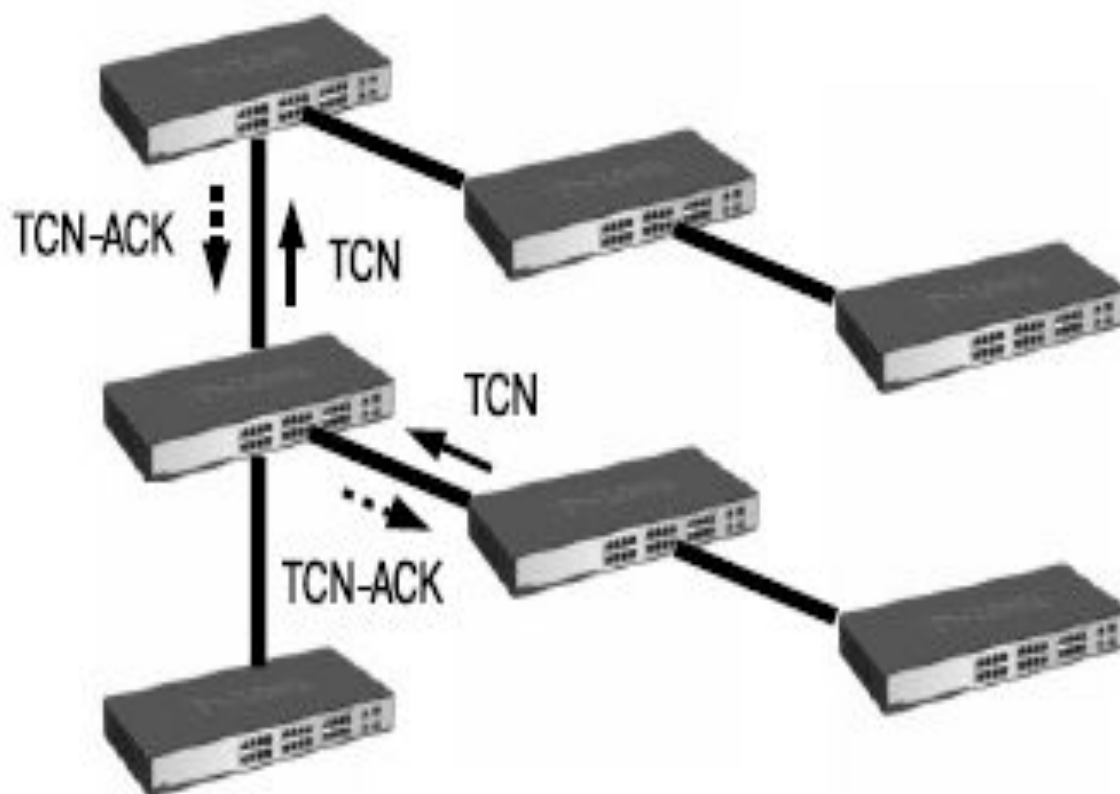
Max Age — это интервал времени, в течение которого коммутатор хранит параметры текущей конфигурации связующего дерева. Значение таймера *Max Age* устанавливается корневым мостом и позволяет гарантировать, что все коммутаторы сети обладают одинаковой информацией о времени хранения конфигурации STP. Если период времени, определенный таймером, истек, а коммутатор за это время не получил кадр BPDU от корневого моста, то он начинает считать себя корневым мостом и рассылает свои собственные BPDU всем коммутаторам сети, инициируя новую процедуру построения связующего дерева. Значение таймера *Max Age* по умолчанию 20 секунд, диапазон возможных значений от 6 до 40 секунд.

Значения таймеров *Hello Time*, *Forward Delay* и *Max Age* могут быть вручную настроены администратором сети на коммутаторе. Обычно эти настройки выполняются только на коммутаторе, являющемся корневым для данной топологии связующего дерева. При настройке важно помнить, что неправильно подобранные значения таймеров могут значительно увеличить время сходимости топологии STP и снизить производительность сети, поэтому рекомендуется использовать значения таймеров по умолчанию.

Изменение топологии

- Коммутатор отправляет BPDU с уведомлением об изменении топологии (Topology Change Notification BPDU, TCN BPDU) в случае возникновения одного из следующих событий:
 - некорневой мост получает сообщение TCN BPDU на свой назначенный порт;
 - после истечения времени, определенного таймером Forward Delay, порт переходит в состояние Forwarding, но коммутатор уже имеет назначенный порт для данного сегмента;
 - порт, находившийся в состоянии Forwarding или Listening, переходит в состояние Blocking (в случае проблем с каналом связи);
 - когда коммутатор становится корневым мостом.

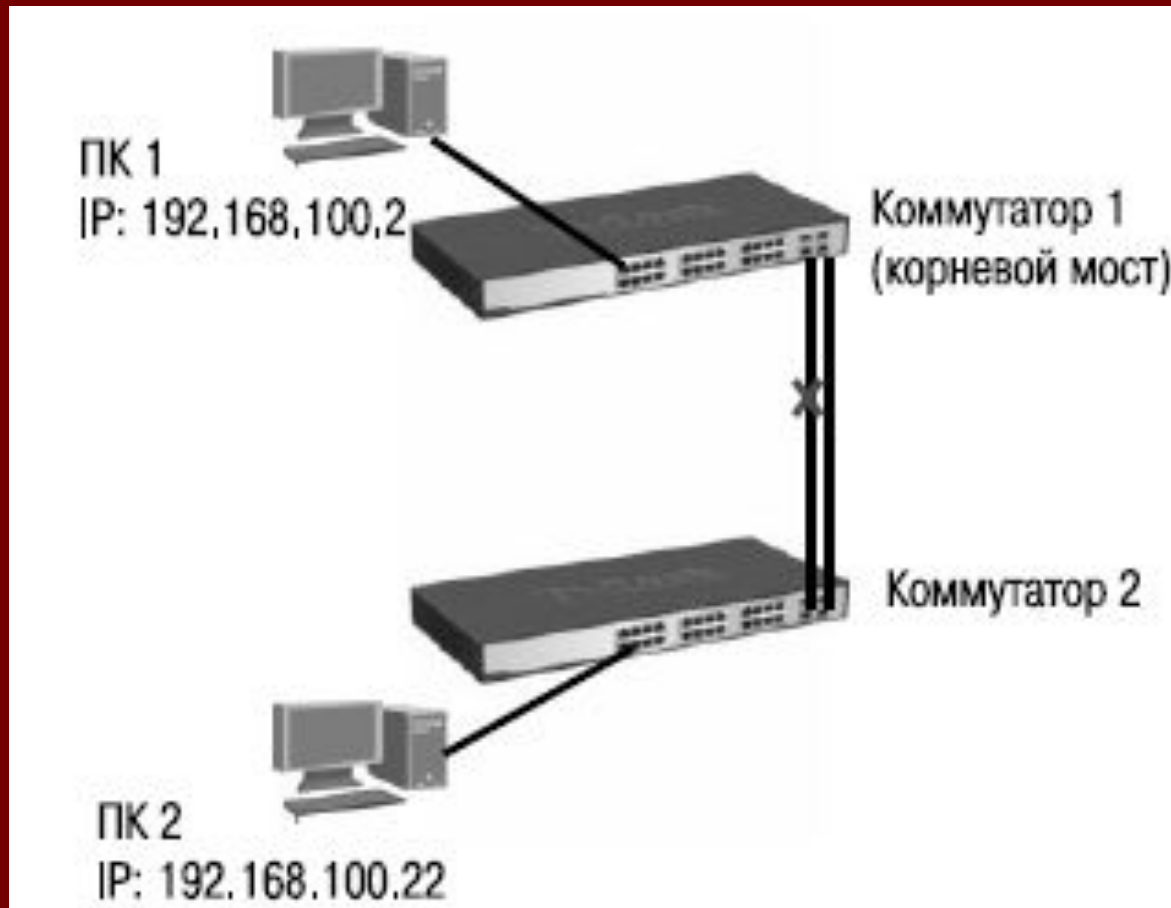
Процесс уведомления об изменении ТОПОЛОГИИ



Изменение топологии

- TCN BPDU отправляется коммутатором в тот сегмент сети, к которому подключен его корневой порт. Эти BPDU будут передаваться через интервал Hello до тех пор, пока коммутатор не получит подтверждение Topology Change Notification Acknowledgement (TCN-ACK) от вышестоящего коммутатора. Соседний коммутатор продолжит трансляцию TCN BPDU через свой корневой порт в направлении корневого моста сети, используя такую же процедуру. Этот процесс будет продолжаться до тех пор, пока TCN BPDU не достигнет корневого моста.
- Когда корневой мост получает TCN BPDU или сам изменяет топологию, он устанавливает во всех передаваемых конфигурационных BPDU флаг изменения топологии (Topology Change, TC) на период времени, равный сумме значений таймеров Forward Delay и Max Age. Когда нижележащие коммутаторы получат конфигурационные BPDU с флагом Topology Change, они установят значения таймеров старения записей адресных таблиц (Aging Timer) равными длительности таймера задержки передачи Forward Delay.
- Управляемые коммутаторы D-Link при настройке функции STP позволяют включать и отключать на каждом порте возможность приема TCN BPDU с помощью параметра *restricted_tcn*. По умолчанию параметр *restricted_tcn* отключен. Использование данного параметра позволяет избежать сетевых атак, связанных с отправкой ложных кадров TCN BPDU.

Настройка STP на коммутаторах D-Link (**Внимание:** по умолчанию протокол STP на коммутаторах D-Link отключен)



Настройка коммутаторов

Настройка коммутатора 1

Активизировать STP

- **enable stp**
- **config stp version stp**

Установить коммутатору 1 наименьшее значение приоритета, что бы он был выбран корневым мостом (приоритет по умолчанию равен 32768)

- **config stp priority 4096 instance_id 0**

Настроить порты STP

- **config stp ports 1-24 edge true**

Настройка коммутатора 2

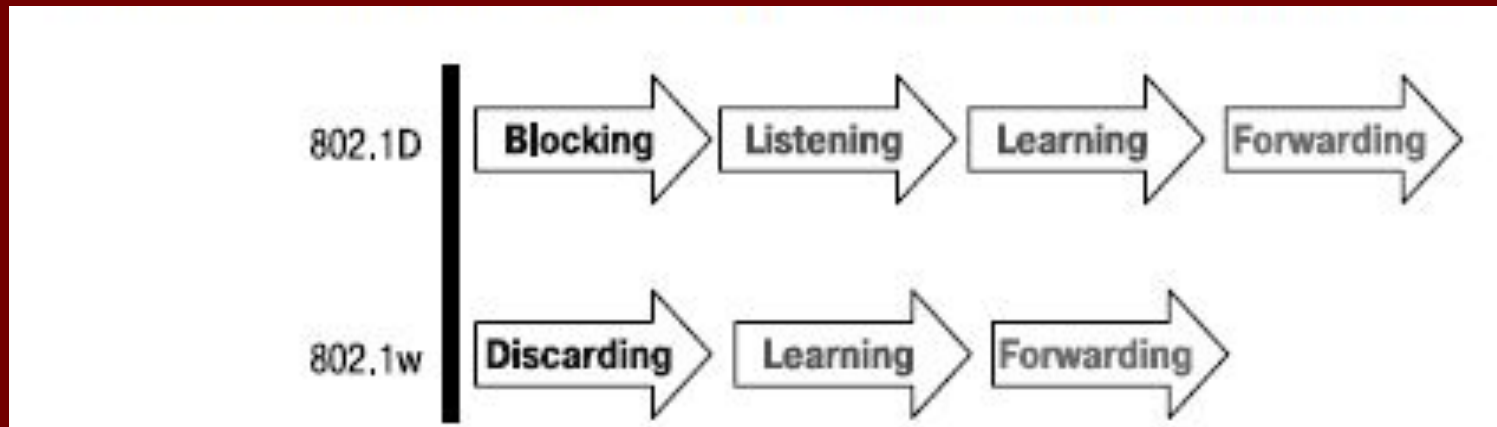
- **enable stp**
- **config stp version stp config stp ports 1-24 edge true**

4.3. Rapid Spanning Tree Protocol

- Протокол Rapid Spanning Tree Protocol (RSTP) является развитием протокола STP и в настоящее время определен в стандарте IEEE 802.1D-2004 (ранее был определен в стандарте IEEE 802.1w-2001). Он был разработан для преодоления отдельных ограничений протокола STP, связанных с его производительностью. Протокол RSTP значительно ускоряет время сходимости коммутируемой сети за счет мгновенного перехода корневых и назначенных портов в состояние продвижения.
- RSTP может работать с оборудованием, поддерживающим STP, однако все преимущества от его использования будут потеряны.

Состояния портов протоколов STP и RSTP

- Основные понятия и терминология протоколов STP и RSTP одинаковы (рис.4.7, табл. 4.1). Существенным их отличием является способ перехода портов в состояние продвижения и то, каким образом этот переход влияет на роль порта в топологии. RSTP объединяет состояния Disabled, Blocking и Listening, используемые в STP, и создает единственное состояние *Discarding* («Отбрасывание»), при котором порт не активен.



Различия между состояниями портов в STP и RSTP

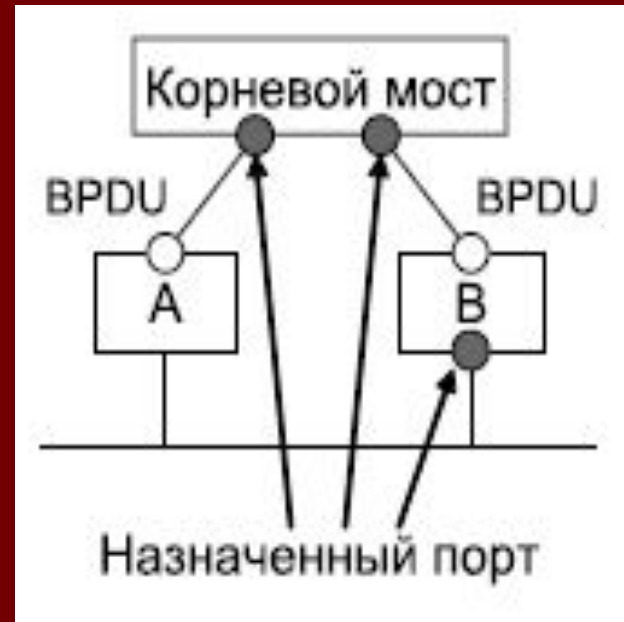
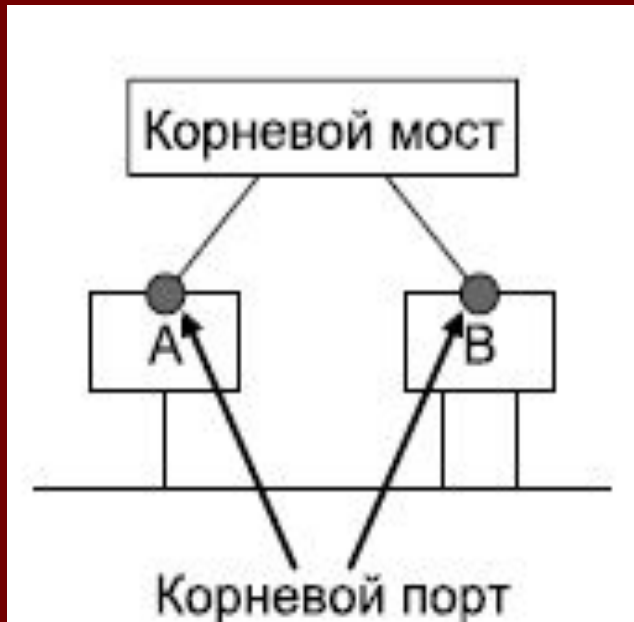
Состояние порта STP	Административное состояние порта коммутатора	Порт изучает MAC-адреса?	Состояние порта RSTP	Роль порта в активной топологии
Disable	Disabled	Нет	Discarding	Исключен (Disabled)
Disable	Enabled	Нет	Discarding	Исключен (Disabled)
Blocking	Enabled	Да	Discarding	Исключен (Alternate, Backup)
Listening	Enabled	Да	Discarding	Включен (Root, Designated)
Learning	Enabled	Да	Learning	Включен (Root, Designated)
Forwarding	Enabled	Да	Forwarding	Включен (Root, Designated)

Роли портов

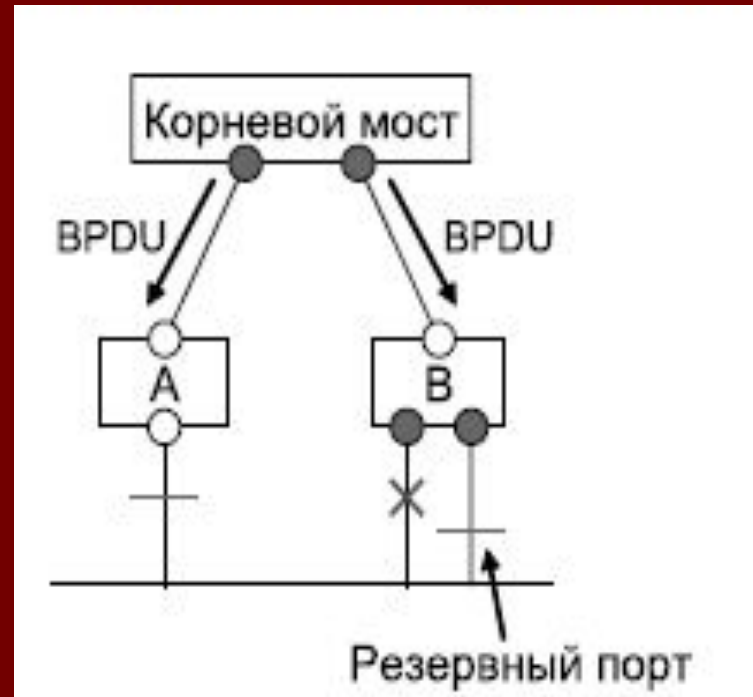
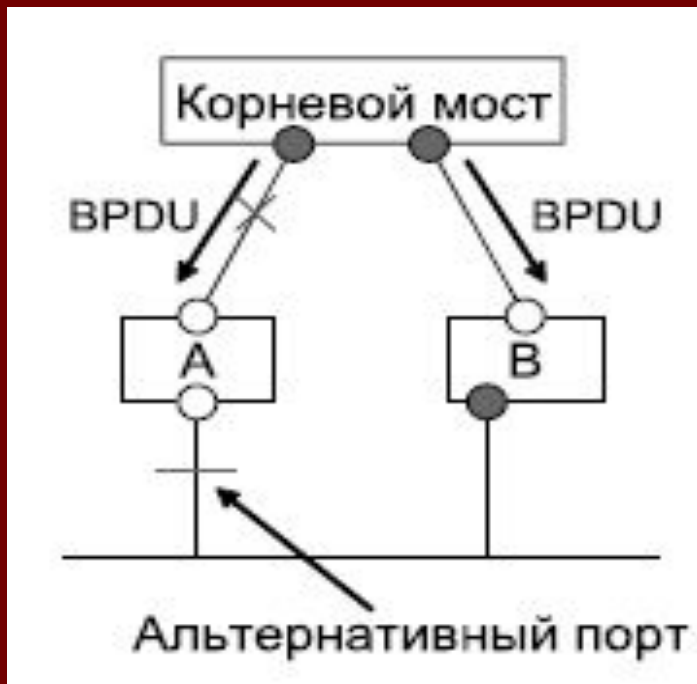
- корневой порт (Root Port) — это порт коммутатора, который имеет по сети кратчайшее расстояние (в терминах стоимости пути) до корневого коммутатора;
- назначенный порт (Designated Port). Порт является *назначенным*, если он посылает BPDU с наилучшими параметрами в тот сегмент, к которому подключен. Роли «корневой порт» и «назначенный порт» включают порт в активную топологию;
- альтернативный порт (Alternate Port) предлагает альтернативный основному маршруту путь в направлении корневого моста и может заменить корневой порт в случае выхода его из строя;
- резервный порт (Backup Port) предназначен для резервирования пути, предоставляемого назначенным портом в направлении сегментов сети, и не может гарантировать альтернативное подключение к корневому мосту. Резервные порты существуют только в конфигурациях, где есть два или более соединения данного моста с данной сетью (сегментом сети)

В RSTP существуют две дополнительные роли — альтернативный порт (*Alternate*) и резервный порт (*Backup*), соответствующие состоянию «Заблокирован» в STP и исключающие порт из активной топологии

Корневой порт (слева) и назначенный порт (справа)



Альтернативный порт (слева) и резервный порт (справа)



Формат кадра BPDU протокола RSTP

	Байты
Идентификатор протокола (Protocol Identifier)	2
Версия протокола (Protocol Version Identifier)	1
Тип BPDU (BPDU Type)	1
Флаги (Flags)	1
Идентификатор корневого моста (Root Identifier)	8
Расстояние до корневого моста (Root Path Cost)	2
Идентификатор моста (Bridge Identifier)	8
Идентификатор порта (Port Identifier)	2
Время жизни сообщения (Message Age)	2
Максимальное время жизни сообщения (Max Age)	2
Время приветствия (Hello Time)	2
Задержка смены состояний (Forward Delay)	2
Длина версии 1 (Version 1 Length)	1

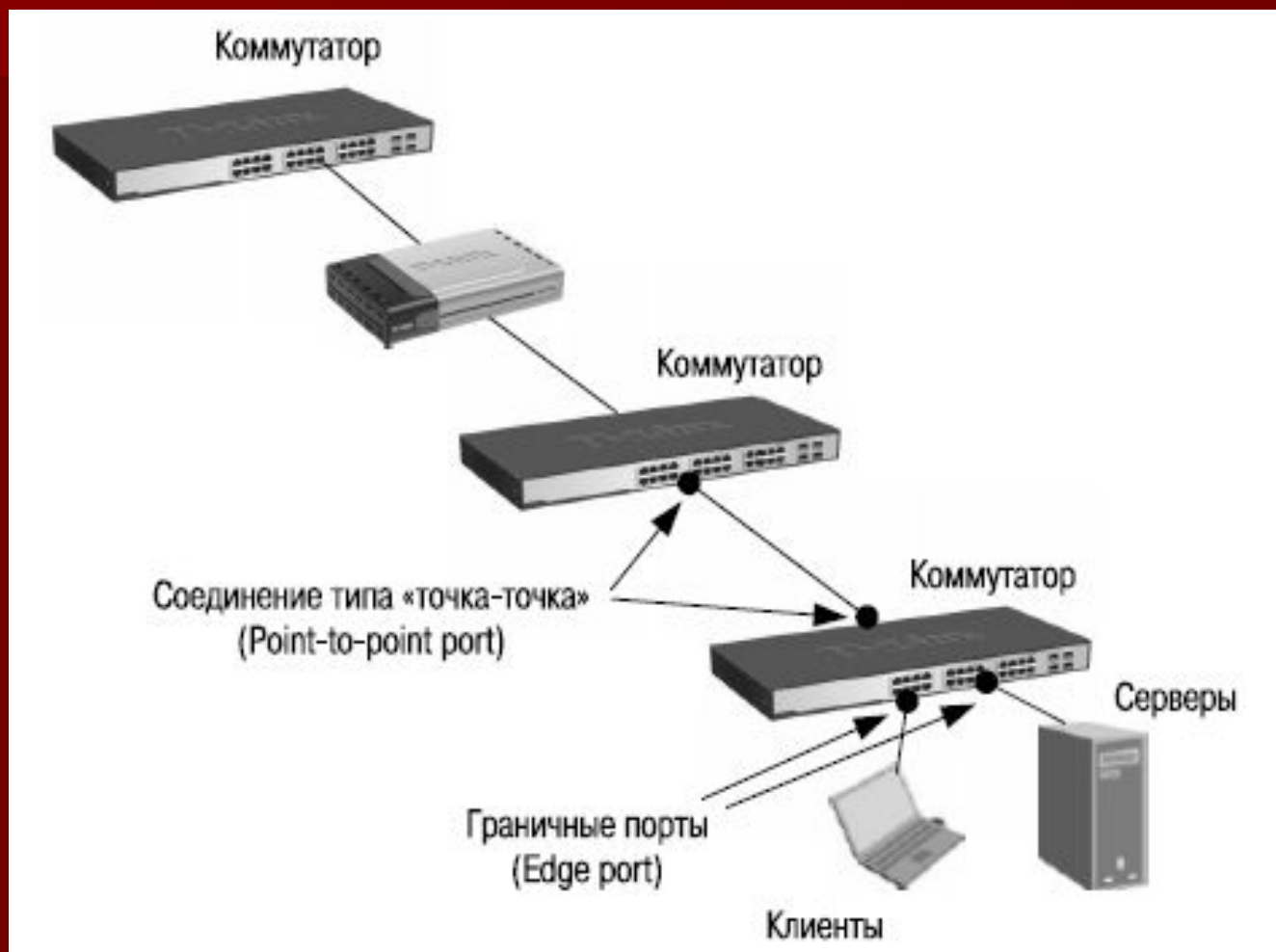
Формат BPDU

- поля версии протокола и типа BPDU RSTP содержат значение 2;
- в поле Flag BPDU протокола STP используются только два бита, которые определяют флаги изменения топологии TC и подтверждения TC (TCA). В поле Flag протокола RSTP используются все 8 бит. Бит 1 — флаг изменения топологии (*Topology Change*), бит 2 — флаг предложения (*Proposal*), биты 3 и 4 предназначены для кодирования роли порта (*Port Role*), бит 5 — флаг изучения (*Learning*), бит 6 — флаг продвижения (*Forwarding*), бит 7 — флаг соглашения (*Agreement*), бит 8 — флаг подтверждения TC (*Topology Change Acknowledgment*).
- кадр BPDU протокола RSTP имеет дополнительное поле *Version 1 Length* длиной 1 байт. Это поле содержит значение 0000 0000 и показывает, что BPDU не содержит никакой информации протокола STP версии 1.

Быстрый переход в состояние продвижения

- Процесс построения связующего дерева у протоколов STP и RSTP одинаков. Однако при работе RSTP порт может перейти в состояние продвижения значительно быстрее, т.к. он больше не зависит от настроек таймеров. Протокол RSTP предоставляет механизм предложений и соглашений, который обеспечивает быстрый переход корневых и назначенных портов в состояние Forwarding, а альтернативных и резервных портов в состояние Discarding. Для этого протокол RSTP вводит два новых понятия: граничный порт и тип соединения.
- Граничным портом (*Edge Port*) объявляется порт, непосредственно подключенный к сегменту сети, в котором не могут быть созданы петли. Например, порт подключен к рабочей станции, которая может периодически включаться или выключаться и активизировать механизм уведомления об изменении топологии или чтобы избежать распространения вычислений STP по клиентским сетям, с целью повышения безопасности. Граничный порт мгновенно переходит в состояние продвижения, минуя состояния прослушивания и обучения. Граничный порт теряет свой статус и становится обычным портом связующего дерева в том случае, если получит кадр BPDU.
- При работе протокола RSTP назначенный порт может выполнять быстрый переход в состояние продвижения в соединениях типа «точка — точка» (*Point-to-Point, P2P*), т.е. если он подключен только к одному коммутатору

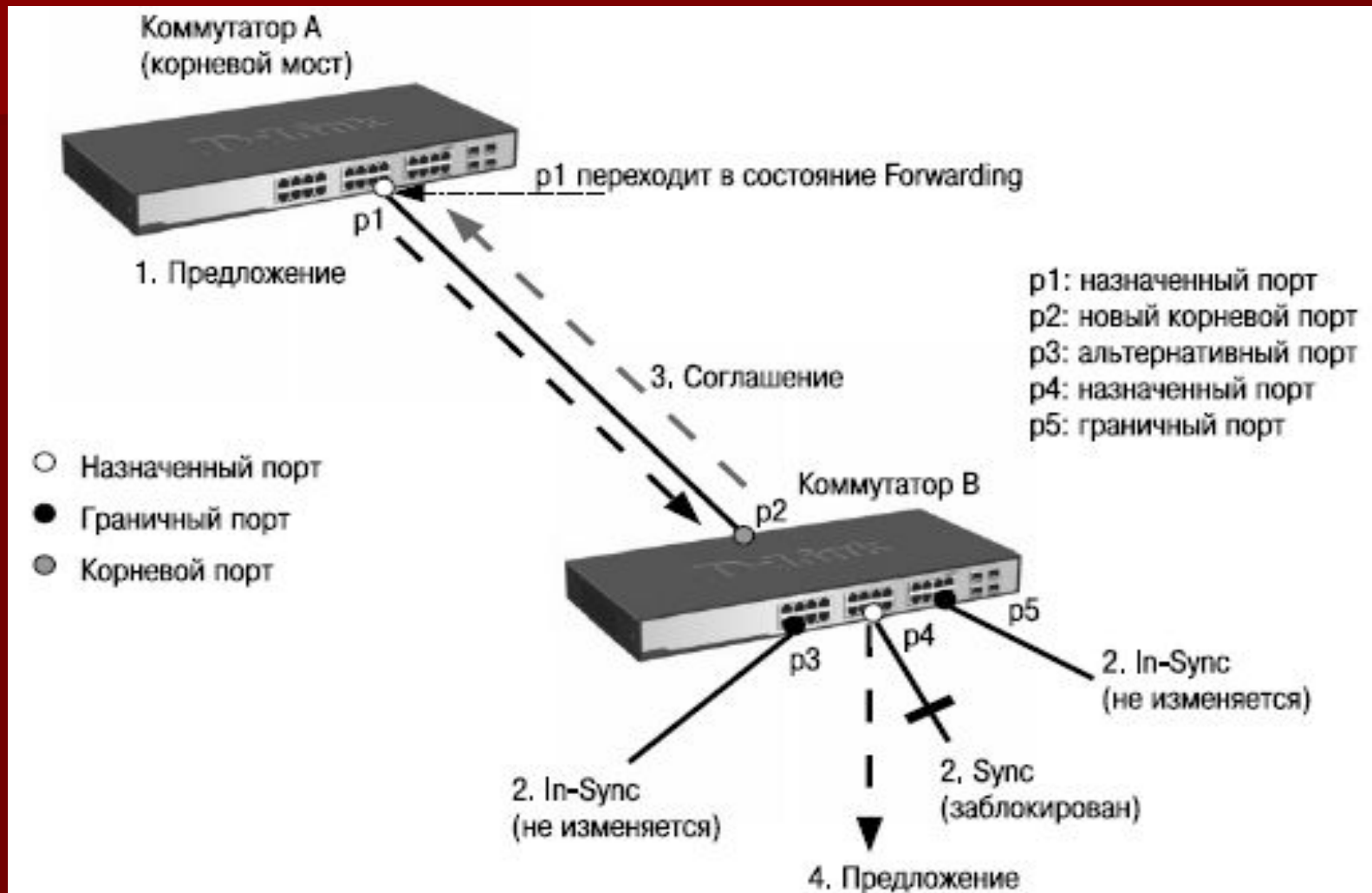
Граничные порты и порты «точка — точка»



Порты, удовлетворяющие, по крайней мере, одному из следующих условий, автоматически рассматриваются протоколом RSTP как порты P2P:

- порт принадлежит агрегированному каналу связи;
- на порте включена функция автосогласования и она определила работу в полнодуплексном режиме;
- работа в полнодуплексном режиме на порте была настроена вручную администратором сети.
- Администратор сети может вручную включать или выключать статусы Edge и P2P либо устанавливать их работу в автоматическом режиме, выполнив соответствующие настройки порта коммутатора.

Механизм предложений и соглашений



Механизм предложений и соглашений

- Коммутаторы А и В соединены между собой каналом типа «точка — точка». Предположим, что коммутатор А является корневым мостом сети. Коммутатор А посылает коммутатору В кадр BPDU с установленным флагом Proposal (шаг 1), предлагая себя в качестве назначенного моста этого сегмента (BPDU-предложение будет передаваться только в том случае, если порт находится в состоянии Discarding или Learning). После получения предложения коммутатор В выберет в качестве нового корневого порта тот порт, через который этот BPDU был получен (порт p2), и переведет все неграничные порты в заблокированное состояние. Все остальные порты будут синхронизированы с новой информацией, чтобы иметь непротиворечивую информацию о топологии сети.
- Порт является синхронизированным «*in-sync*», если он удовлетворяет следующим критериям:
 - он находится в заблокированном состоянии (это состояние Discarding в стабильной топологии);
 - он является граничным портом.

Механизм предложений и соглашений

- Чтобы продемонстрировать действие метода синхронизации на различные типы портов, предположим, что в коммутаторе В имеются граничные порты р3 и р5 и назначенный порт р4. Порты р3 и р5 уже удовлетворяют одному из условий синхронизации. Чтобы находиться в режиме синхронизации (шаг 2), коммутатору В необходимо заблокировать порт р4, переведя его в состояние Discarding.
- После того, как коммутатор В убедится, что все порты синхронизированы, он разблокирует свой новый корневой порт (шаг 3) и отправит через него коммутатору А согласие на предложение. Это сообщение является копией BPDU-предложения, в котором вместо бита Proposal установлен бит Agreement. Благодаря этому порт р1 коммутатора А точно знает, какому предложению соответствует полученное согласие. После этого коммутатор А мгновенно переведет свой назначенный порт р1 в состояние продвижения.
- Находясь в заблокированном состоянии порт р4 коммутатора В начнет отсылать предложения нижележащему коммутатору и пытаться быстро перейти в состояние продвижения (шаг 4).

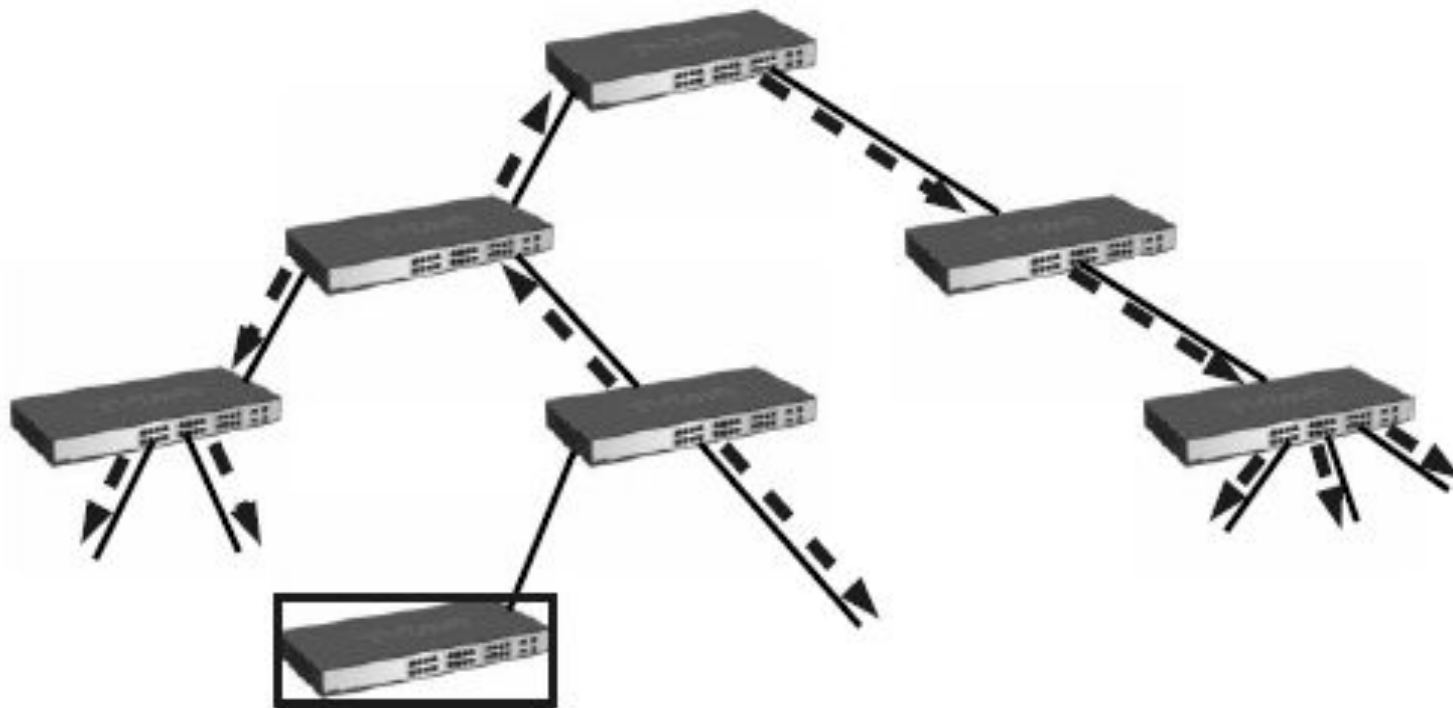
Новый механизм изменения топологии

1. Определение изменений топологии.

В протоколе RSTP только неграничные порты, переходя в состояние продвижения, могут вызвать процесс изменения топологии. Это означает, что разрыв соединения больше не рассматривается как изменение в топологии, в отличие от протокола STP, т.е. при переходе порта в заблокированное состояние соответствующий коммутатор не генерирует TCN BPDU. Когда мост RSTP обнаруживает изменение топологии, происходит следующее:

- коммутатор устанавливает начальное значение таймера TC While равным удвоенному интервалу Hello для всех неграничных назначенных портов и корневого порта. While Timer — это интервал времени, в течение которого мост RSTP активно информирует остальные мосты в сети об изменении топологии;
- удаляет MAC-адреса, ассоциированные со всеми неграничными назначенными портами и корневым портом;
- до тех пор, пока не истечет время, установленное таймером TC While, запущенным на порте, в BPDU, отправляемых через него, будет установлен бит TC.

Новый механизм изменения топологии



Коммутатор определил
изменение топологии

1. Запускается таймер TC While и удаляются MAC-адреса, ассоциированные со всеми неграничными портами.
2. Коммутатор отправляет BPDU с установленным битом TC через все свои неграничные порты.

Новый механизм изменения топологии

2. Распространение информации об изменении топологии.

Когда коммутатор получает от соседа BPDU с установленным битом TC, происходит следующее:

- коммутатор удаляет все MAC-адреса, изученные его неограниченными назначенными портами и корневым портом, за исключением того порта, который получил информацию об изменении топологии;
- коммутатор запускает таймер TC While и отправляет BPDU с установленным битом TC через все неограниченные порты (RSTP не использует специальные TCN BPDU, за исключением случаев, когда требуется уведомить коммутатор, поддерживающий только протокол STP).

Коммутатор-отправитель BPDU с битом TC непосредственно распространяет информацию об изменении топологии через всю сеть (в отличие от STP, где это может выполнить только корневой мост). Этот механизм распространения информации об изменении топологии быстрее, чем его аналог в протоколе STP, т.к. нет необходимости ждать, когда будет уведомлен корневой мост, и потом поддерживать состояние изменения топологии для всей сети в течение периода времени, равного сумме значений таймеров Forward Delay и Max Age.

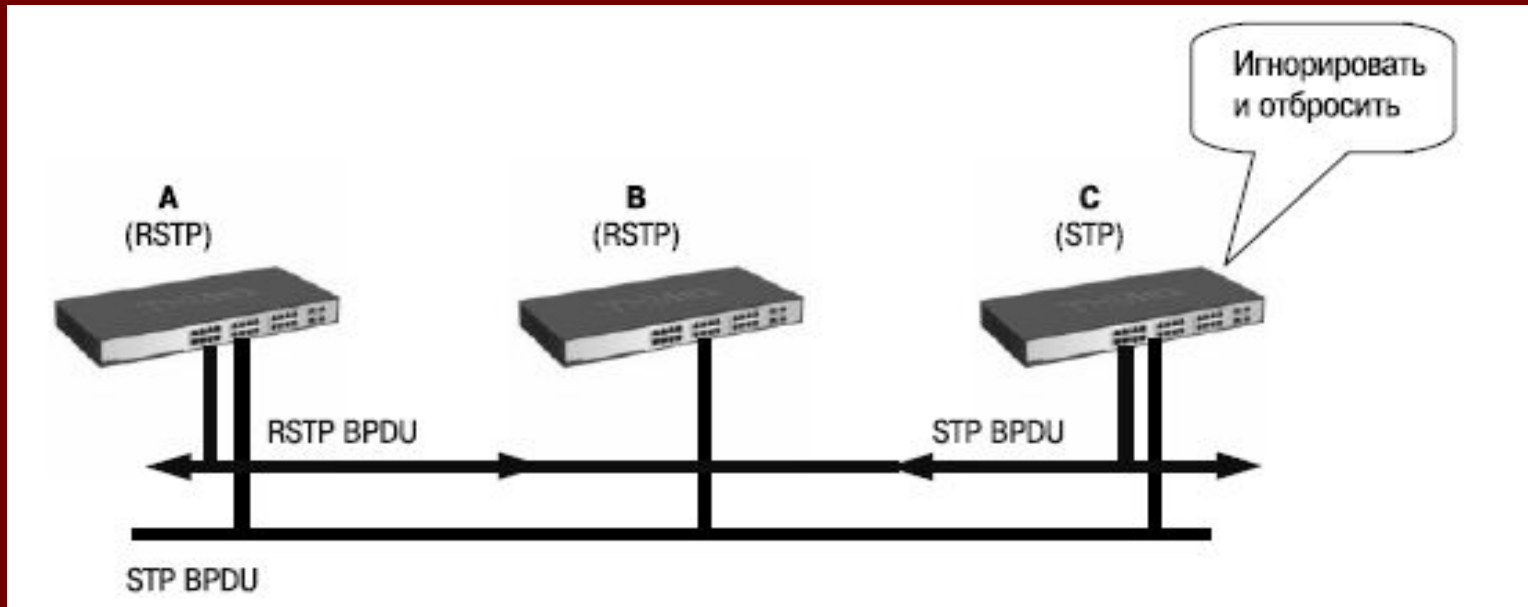
Стоимость пути RSTP.

Протокол RSTP определяет следующие рекомендованные значения стоимости пути по умолчанию для портов коммутаторов.

Параметр	Скорость канала	Рекомендованное значение	Рекомендованный диапазон	Диапазон значений
Стоимость пути	10 Мбит/с	2 000 000	200 000– 20 000 000	1– 200 000 000
Стоимость пути	100 Мбит/с	200 000	20 000– 2 000 000	1– 200 000 000
Стоимость пути	1 Гбит/с	20 000	2 000–200 000	1– 200 000 000
Стоимость пути	10 Гбит/с	2 000	200–20 000	1– 200 000 000

Совместимость с STP

- Протокол RSTP может взаимодействовать с оборудованием, поддерживающим STP, и, если необходимо, автоматически преобразовывать кадры BPDU в формат 802.1D. Однако преимущество быстрой сходимости RSTP (когда все коммутаторы быстро переходят в состояние пересылки или блокировки и обладают тождественной информацией) теряется.



Совместимость с STP

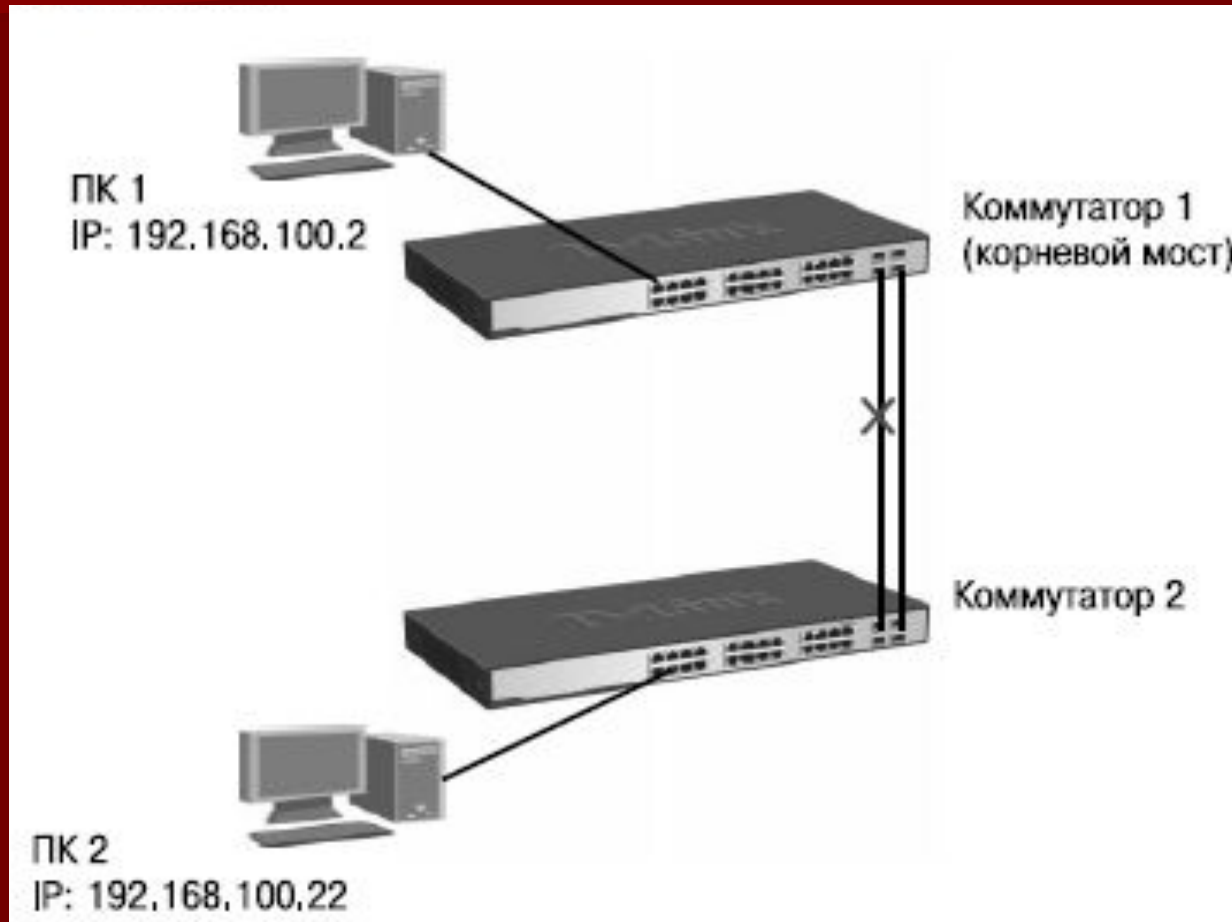
- Каждый порт хранит переменную, определяющую тип протокола, используемого в соответствующем сегменте. При включении порта активизируется таймер задержки миграции (*Migration delay timer*) длительностью 3 секунды. При запуске этого таймера текущий режим (STP или RSTP) ассоциированный с портом, блокируется. Как только истечет время задержки миграции, порт начнет работать в режиме, соответствующем типу следующего полученного им BPDU. Если в результате получения BPDU порт изменил свой режим работы, таймер задержки миграции запускается вновь, что позволяет ограничить частоту возможной смены режимов.
- Предположим, что коммутаторы А и В работают в режиме RSTP. Коммутатор А является выделенным мостом этого сегмента. К существующему каналу связи подключается коммутатор С, который является коммутатором с поддержкой протокола STP. Так как коммутаторы STP игнорируют BPDU протокола RSTP и отбрасывают их, то коммутатор С считает, что в этом сегменте сети больше коммутаторов нет и начинает отправлять BPDU формата 802.1D.

Совместимость с STP

- Коммутатор А получает эти BPDU и после истечения периода времени, установленного таймером задержки миграции, переходит на этом порте в режим работы STP. В результате коммутатор С начинает понимать BPDU коммутатора А и признает его назначенным коммутатором этого сегмента.
- Следует отметить, что если бы в этом частном случае коммутатор С был удален из сегмента, то коммутатор А остался бы работать в режиме STP на этом порте, хотя он мог бы эффективно работать в режиме RSTP со своим единственным соседом коммутатором В. Т.е. у коммутатора А нет возможности узнать, что коммутатор С удален из этого сегмента. В этом частном случае для перезагрузки протокола, используемого на порте коммутатора, требуется вмешательство администратора сети. Когда порт находится в режиме, совместимом с 802.1D, он также может обрабатывать уведомления об изменении топологии TCN BPDU с установленными битами ТС и ТСА.

Настройка RSTP

(настройка на коммутаторах D-Link аналогична настройке протокола STP)



Настройка коммутаторов

Настройка коммутатора 1

Активизировать RSTP

- **enable stp**
- **config stp version rstp**

Установить коммутатору 1 наименьшее значение приоритета, чтобы он был выбран корневым мостом (приоритет по умолчанию равен 32768)

- **config stp priority 4096 instance_id 0**

Настроить граничные порты RSTP

- **config stp ports 1-24 edge true**

Настройка коммутатора 2

- **enable stp**
- **config stp version rstp**
- **config stp ports 1-24 edge true**

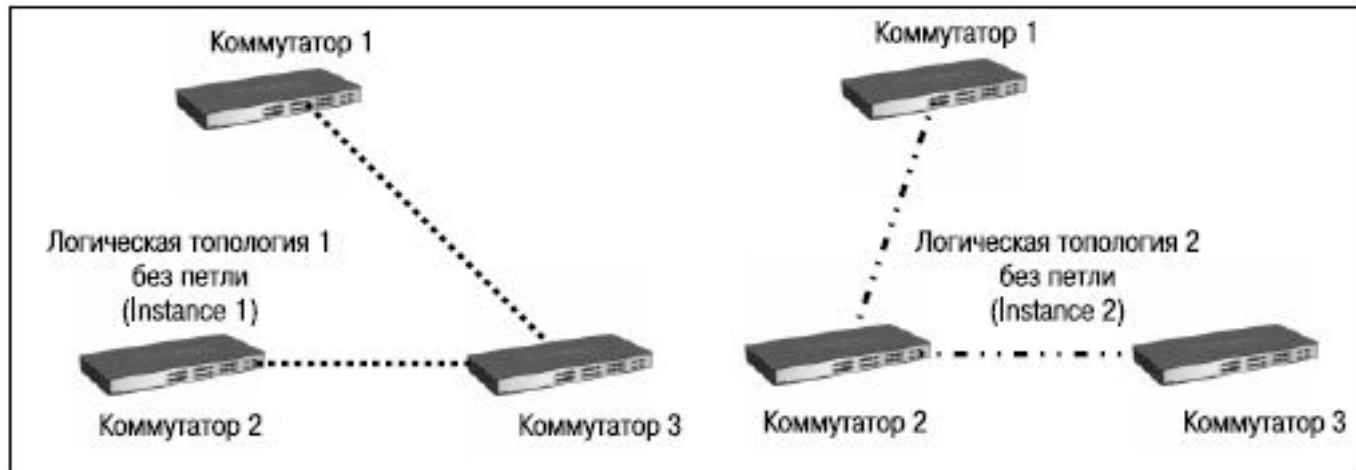
4.4. Multiple Spanning Tree Protocol

- Несмотря на то, что протокол RSTP обеспечивает быструю сходимость сети, он так же, как и протокол STP, обладает следующим недостатком — не поддерживает возможность создания отдельного связующего дерева для каждой VLAN, настроенной в сети. Это означает, что резервные каналы связи не могут блокироваться на основе VLAN и все VLAN образуют одну логическую топологию, не обладающую достаточной гибкостью.
- Протокол Multiple Spanning Tree Protocol (MSTP), являющийся расширением протокола RSTP, преодолевает это ограничение. В дополнение к обеспечению быстрой сходимости сети он позволяет настраивать отдельное связующее дерево для любой VLAN или группы VLAN, создавая множество маршрутов передачи трафика и позволяя осуществлять балансировку нагрузки. Первоначально протокол MSTP был определен в стандарте IEEE 802.1s, но позднее был добавлен в стандарт IEEE 802.1Q-2003. Протокол MSTP обратно совместим с протоколами STP и RSTP.

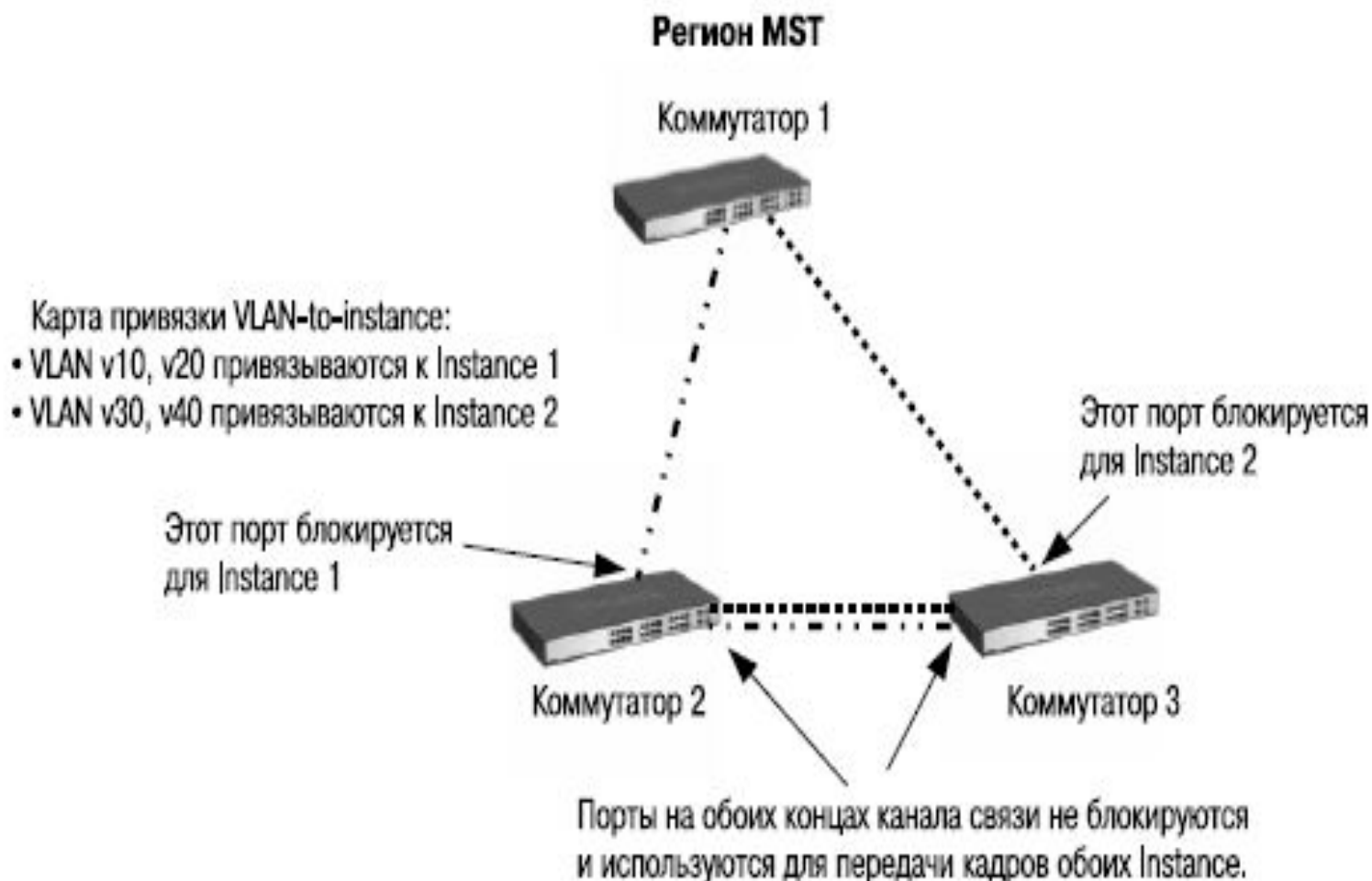
Логическая структура MSTP

- Протокол MSTP делит коммутируемую сеть на регионы MST (*Multiple Spanning Tree (MST) Region*), каждый из которых может содержать множество копий связующих деревьев (*Multiple Spanning Tree Instance, MSTI*) с независимой друг от друга топологией. Другими словами, регион MST, представляющий собой набор физически подключенных друг к другу коммутаторов, делит данную физическую топологию на множество логических (следующий слайд № 53)
- Для того чтобы два и более коммутатора принадлежали одному региону MST, они должны обладать одинаковой конфигурацией MST (слайд 54). Конфигурация MST включает такие параметры, как номер ревизии MSTP (*MSTP revision level number*), имя региона (*Region name*), карту привязки VLAN к копии связующего дерева (*VLAN-to-instance mapping*).
- Внутри коммутируемой сети может быть создано множество MST- регионов (слайд 55)

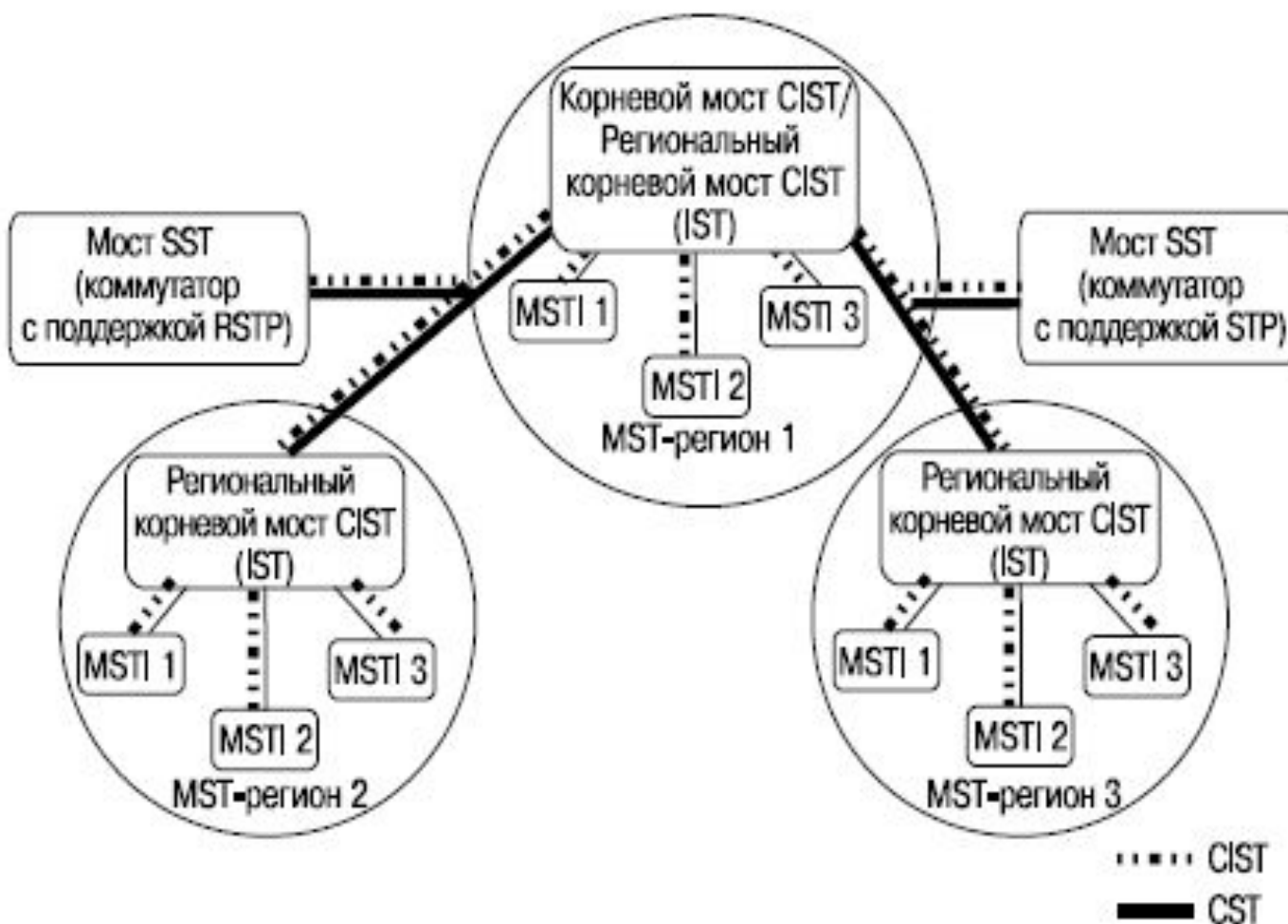
Физическая и логическая топология региона MST



Регион MST



Логическая структура MSTP



Протокол MSTP определяет следующие ТИПЫ СВЯЗУЮЩИХ ДЕРЕВЬЕВ:

- *Internal Spanning Tree (IST)* — специальная копия связующего дерева, которая по умолчанию существует в каждом MST-регионе. *IST* присвоен номер 0 (Instance 0). Она может отправлять и получать кадры BPDU и служит для управления топологией внутри региона. По умолчанию все VLAN одного региона привязаны к *IST*. Если в регионе создано несколько *MSTI*, то VLAN, не ассоциированные с ними, остаются привязанными к *IST*. Динамические VLAN, созданные с помощью протокола GVRP, также ассоциируются с *IST*;
- *Common Spanning Tree (CST)* — единое связующее дерево, вычисленное с использованием протоколов STP, RSTP, MSTP и объединяющее все регионы MST и мосты SST (Single Spanning Tree (SST) Bridge);
- *Common and Internal Spanning Tree (CIST)* — единое связующее дерево, объединяющее *CST* и *IST* каждого MST-региона;
- *Single Spanning Tree (SST) Bridge* — мост, поддерживающий только единственное связующее дерево, *CST*. Это единственное связующее дерево может поддерживать протокол STP или протокол RSTP.

Multiple Spanning Tree Instance (MSTI)

- По умолчанию все VLAN данного MST-региона назначены в *IST*. Помимо *IST*, в каждом MST-регионе может быть дополнительно создано множество связующих деревьев с независимой друг от друга архитектурой — MSTI. К каждой MSTI администратор сети может вручную привязать соответствующие сети VLAN.
- MSTI обладают следующими характеристиками:
 - MSTI является копией связующего дерева, существующей только внутри региона;
 - MSTI не может отправлять BPDU за пределы своего региона (отправлять и получать BPDU может только IST);
 - все MSTI внутри региона могут нумероваться 1, 2, 3, 4 и т.д. (максимальное количество MSTI зависит от модели коммутатора и версии программного обеспечения);
 - MSTI не отправляет индивидуальные BPDU. Вся информация о данной MSTI помещается в конфигурационное сообщение MSTI (MSTI Configuration Message, M-record), которое инкапсулируется в кадры MSTP BPDU, рассылаемые IST.
- Для того чтобы каждая MSTI представляла собой отдельную от IST логическую топологию, администратор сети может присвоить коммутаторам и портам внутри MSTI собственные значения приоритетов и стоимости пути.

Формат MSTP BPDU

Роли портов MSTP

- Формат MSTP BPDU аналогичен формату RSTP BPDU, за исключением полей, предназначенных для передачи информации об IST, каждой MSTI (если они созданы в регионе) и конфигурации MST.
- Протокол MSTP определяет роли портов, которые участвуют в процессе вычисления активной топологии CIST и MSTI, аналогичные протоколам STP и RSTP (рис. 4.21):
 - корневой порт (Root Port);
 - назначенный порт (Designated Port);
 - альтернативный/резервный порт (Alternate/Backup Port).
- Дополнительно в MSTI используется еще одна роль, которая может быть присвоена порту, — мастер-порт (Master Port).

Роли портов RSTP определяют роли каждого порта коммутатора, участвующего в построении активной топологии RSTP.

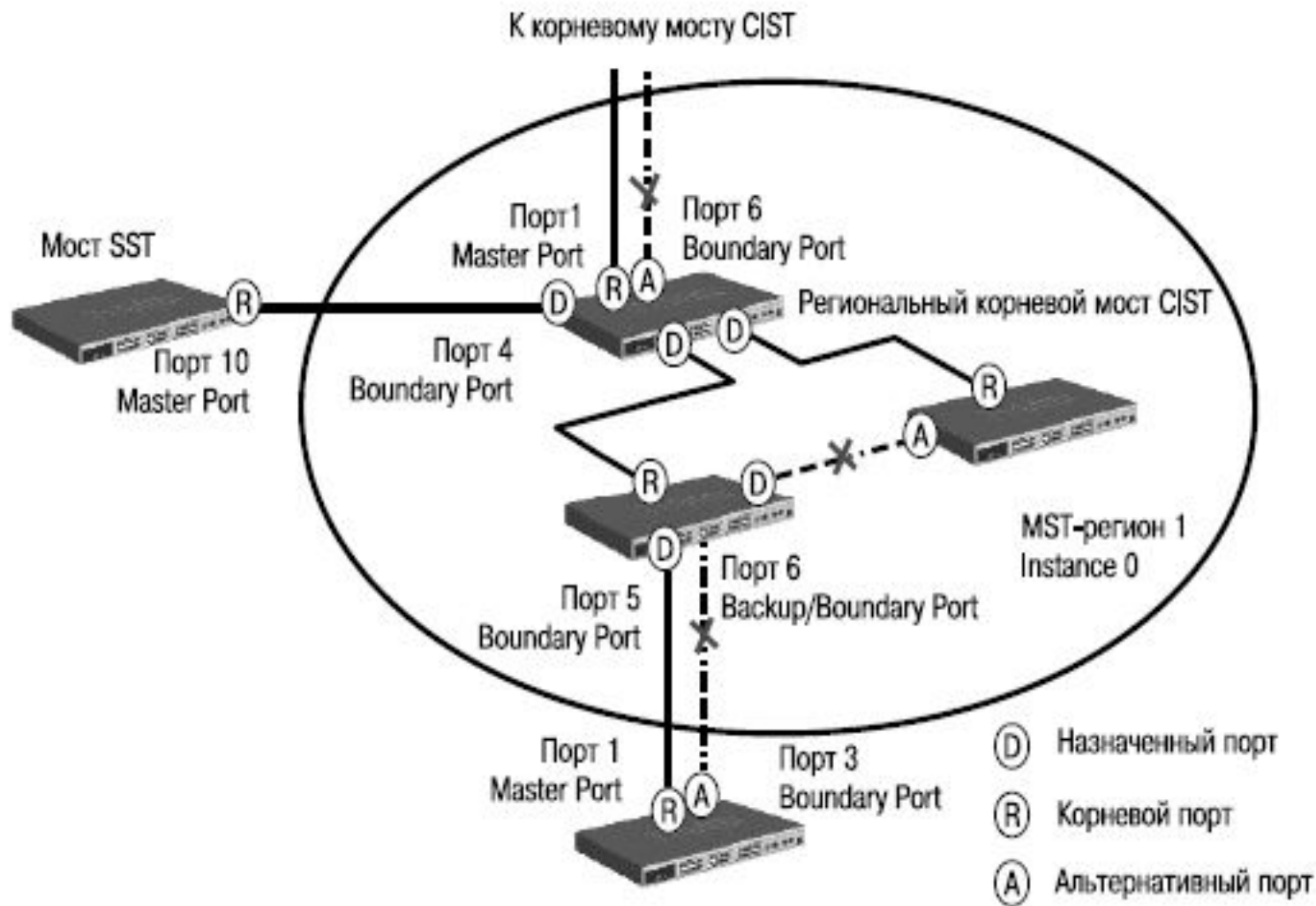
- *Корневой порт (Root Port)* — это порт, который обладает минимальной стоимостью пути от коммутатора до корневого моста RSTP (в случае если мост не является RSTP Root) через региональный мост (в том случае, если коммутатор не является региональным корнем RSTP).
- *Назначенный порт (Designated Port)* — это порт, обладающий наименьшей стоимостью пути от подключенного сегмента сети до корневого моста RSTP.
- *Альтернативный/резервный порт (Alternate/Backup Port)* — это порт, который обеспечивает подключение, если происходит потеря соединения с какими-либо коммутаторами или сегментами сети.

Роли портов MSTI определяют роли каждого порта коммутатора, участвующего в построении активной топологии MSTI внутри границы региона.

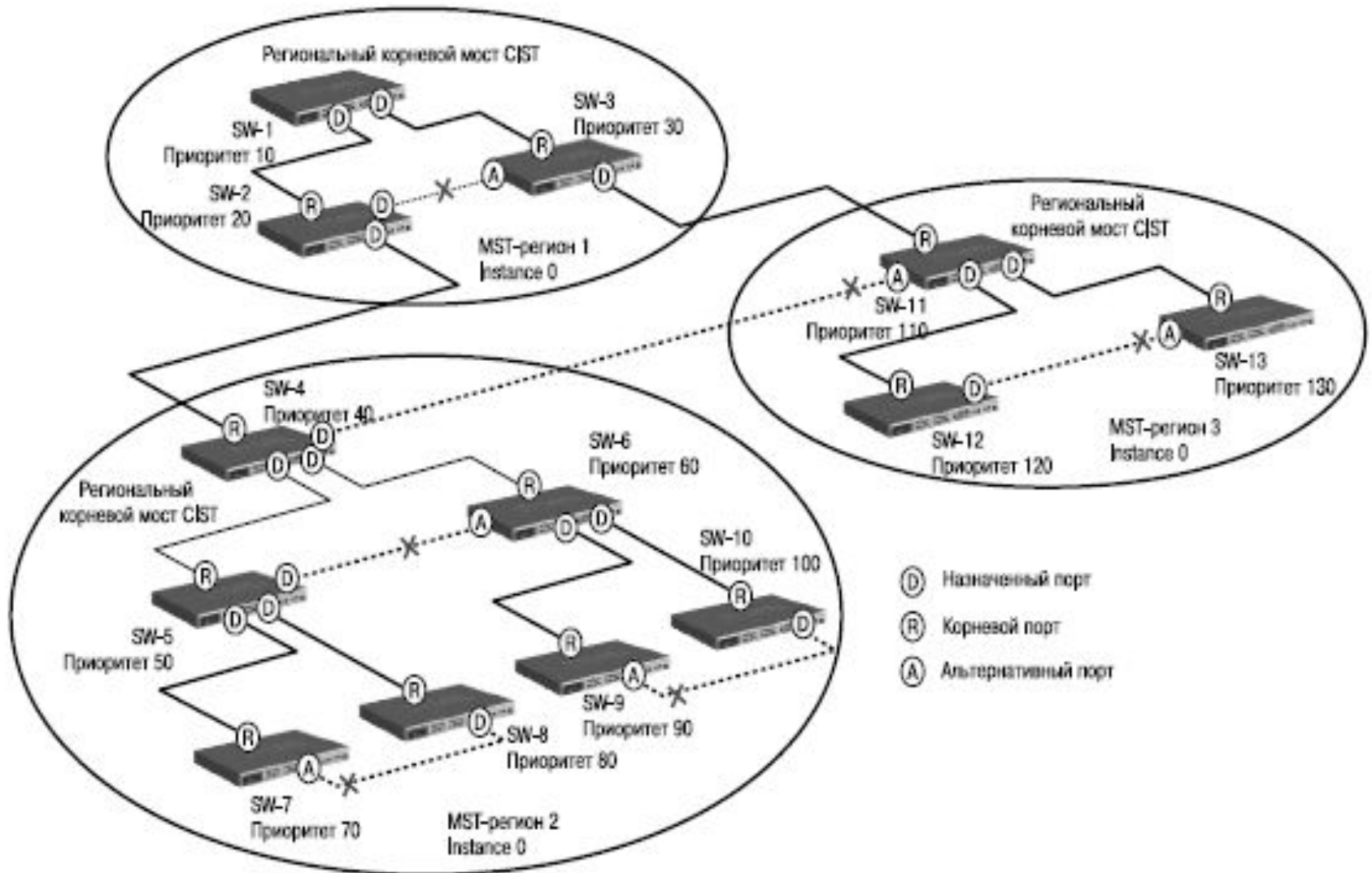
- *Корневой порт (Root Port)* — это порт, который обладает минимальной стоимостью пути от коммутатора до регионального корневого моста MSTI (в случае если мост не является региональным корнем для этой MSTI).
- *Назначенный порт (Designated Port)* — это порт, обладающий наименьшей стоимостью пути от подключенного сегмента сети до регионального корневого моста MSTI.
- *Альтернативный/резервный порт (Alternate/Backup Port)* — это порт, который обеспечивает подключение, если происходит потеря соединения с какими-либо коммутаторами или сегментами сети.
- *Мастер-порт (Master Port)* — это порт, который обеспечивает подключение региона к корневому мосту CIST, находящемуся за пределами данного региона. Корневой порт CIST регионального корневого моста CIST является мастером-портом для всех MSTI.

Протокол MSTP вводит еще одну роль, которая может быть присвоена порту, — *пограничный порт (Boundary Port)*. Пограничным портом является порт, который подключает MST-регион к другому региону или SST-мосту.

Роли портов



Пример топологии MSTP



Сеть разбита на 3 MST-региона, в каждом регионе все коммутаторы ассоциированы с Instance 0.

- Коммутатор 1 (SW-1) выбран в качестве корневого моста CIST, т.к. он обладает наименьшим среди всех коммутаторов сети значением идентификатора моста.
- Коммутаторы 1, 2 и 3 (SW-1, SW-2, SW-3) находятся в одном MST- регионе с номером 1, т.к. обладают одинаковым идентификатором MST-конфигурации. Коммутаторы 2 и 3 находятся в одном регионе с корневым мостом CIST (коммутатор 1), поэтому их внешняя стоимость пути равна 0 и их региональный мост CIST совпадает с корневым мостом CIST.
- Коммутаторы 4-10 (SW-4-SW-10) принадлежат одному региону, т.к. имеют одинаковые идентификаторы MST-конфигурации. Коммутатор 4 (SW-4) является региональным корневым мостом CIST для MST-региона 2, т.к. обладает наименьшей внешней стоимостью пути к CIST Root.
- Коммутаторы 11, 12 и 13 (SW-11-SW-13) принадлежат к MST- региону 3, т.к. обладают одинаковыми идентификаторами MST- конфигурации. Коммутатор 11 (SW-11) выбран в качестве регионального корневого моста CIST для MST-региона 3, т.к. обладает наименьшей внешней стоимостью пути к CIST Root.

Состояние портов MSTP

В протоколе MSTP определены состояния, в которых могут находиться порты, аналогичные протоколу RSTP:

- *Learning* («Обучение») — порт может принимать/отправлять кадры BPDU, изучать MAC-адреса и строить таблицу коммутации. Порт в этом состоянии не передает пользовательские кадры;
- *Forwarding* («Продвижение») — в этом состоянии порт может передавать пользовательские кадры, изучать новые MAC-адреса и принимать/отправлять кадры BPDU;
- *Discarding* («Отбрасывание») — в этом состоянии порт может только принимать кадры BPDU, передача пользовательского трафика и изучение MAC-адресов не выполняется.

4.5. Функции безопасности STP

Из-за ошибок в конфигурации или вредоносных атак в сети может возникнуть ситуация, когда корневой мост получит кадр BPDU, содержащий лучший приоритет, и потеряет свою позицию. При настройке протоколов RSTP или MSTP на управляемых коммутаторах, расположенных на границе сети, с помощью параметра *restricted_role* можно ограничить роли выполняемые портом в активной топологии. При активизации этого параметра порт не будет выбран в качестве корневого порта даже в том случае, если получит BPDU с наилучшим приоритетом. После выбора корневого порта этот порт будет выбран в качестве альтернативного. По умолчанию функция *restricted_role* отключена.

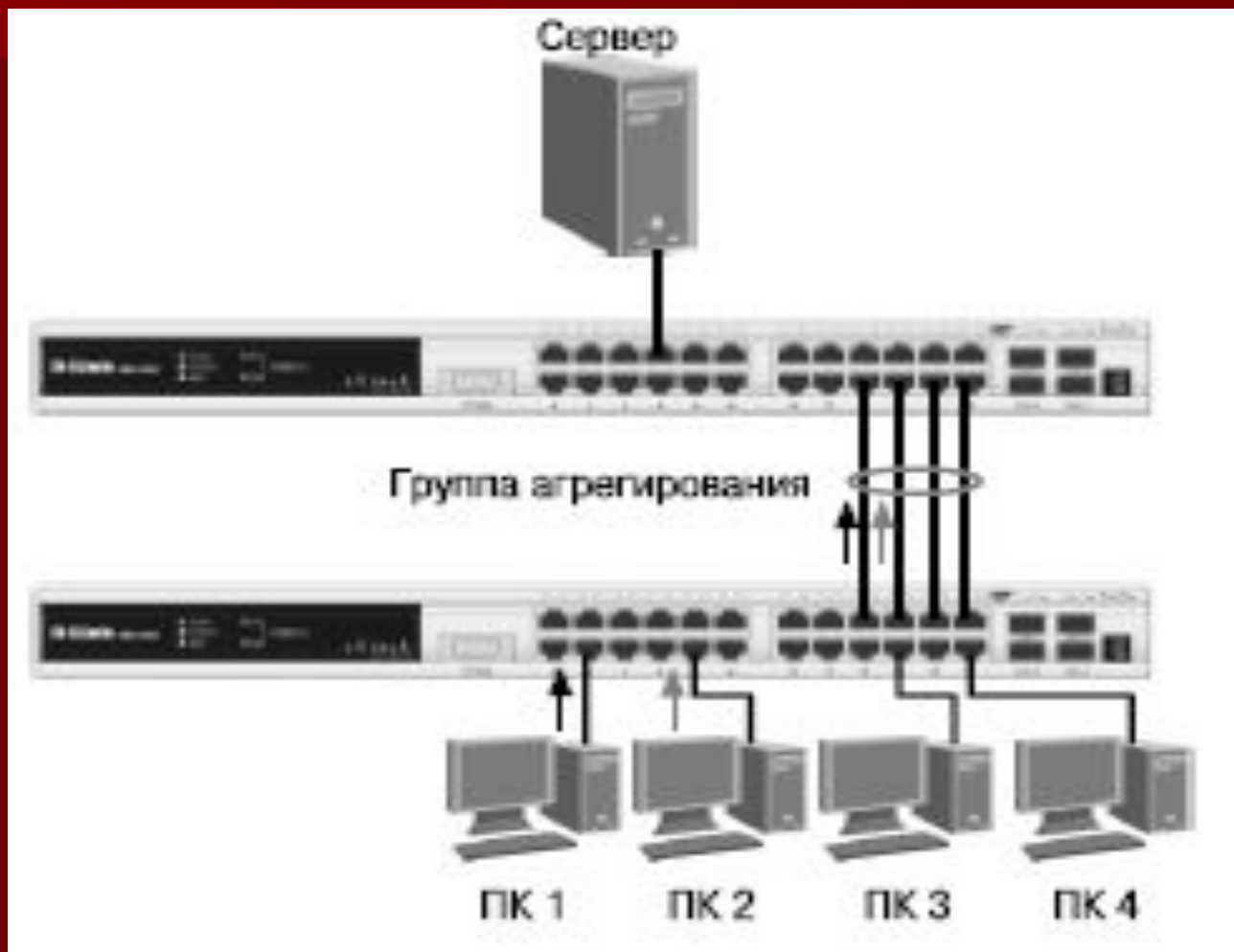
Настройка коммутатора

- **enable stp**
- **config stp version rstp**
- **config stp priority 32768 instance_id 0**
- **config stp ports 1-24 edge true restricted_role true**
- **restricted_tcn true state enable**
- **config stp ports 25-28 edge false state enable fbpdu enable**

4.6. Агрегирование каналов связи

- Агрегирование каналов связи (Link Aggregation) — это объединение нескольких физических портов в одну логическую магистраль на канальном уровне модели OSI с целью образования высокоскоростного канала передачи данных и повышения отказоустойчивости
- В отличие от протокола STP, все избыточные связи в одном агрегированном канале остаются в рабочем состоянии, а имеющийся трафик распределяется между ними для достижения балансировки нагрузки. При отказе одной из линий, входящих в такой логический канал, трафик распределяется между оставшимися линиями.
- Включенные в агрегированный канал порты называются членами группы агрегирования (*Link Aggregation Group*).
- **Внимание:** количество портов в группе зависит от модели коммутатора. В управляемых коммутаторах в группу можно объединить до 8 портов.

Пример агрегированного канала связи между коммутаторами



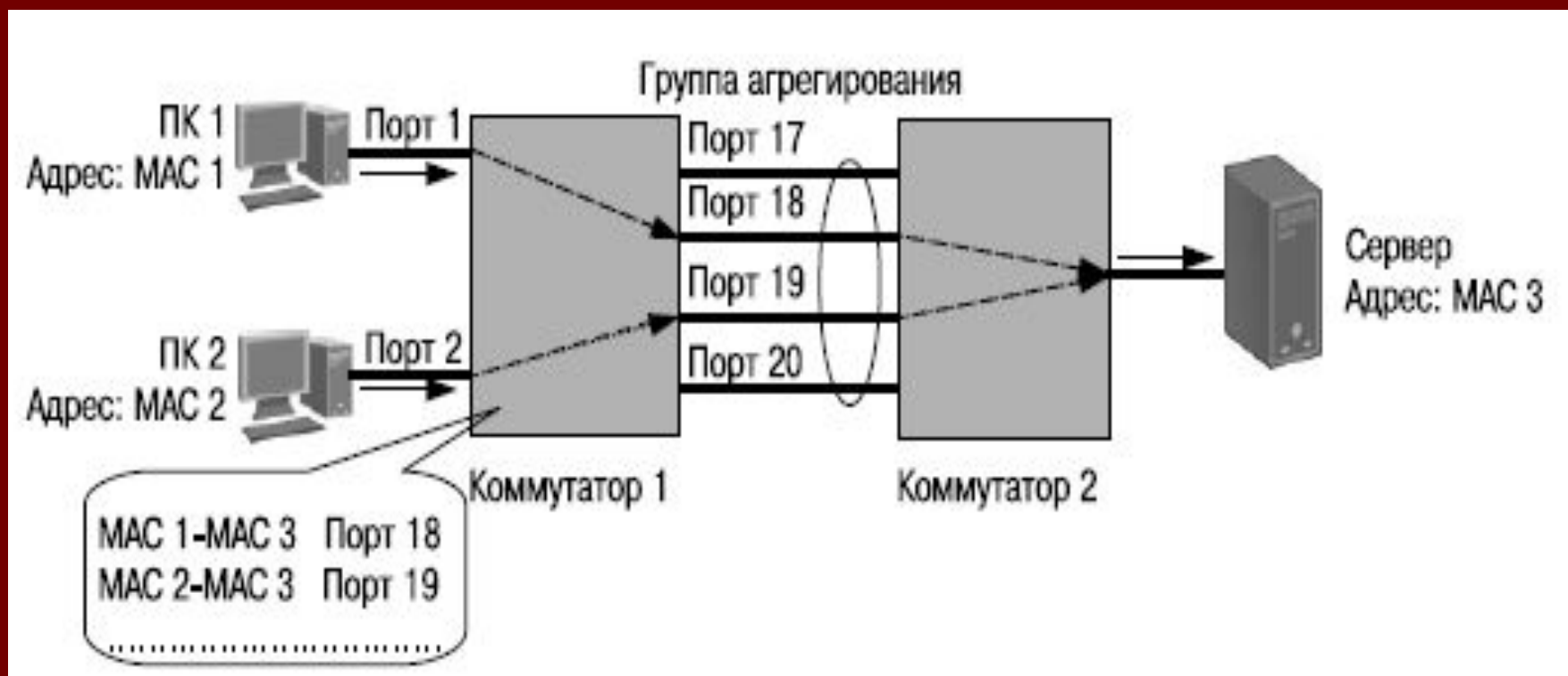
- Один из портов в группе выступает в качестве мастера-порта (*master port*). Так как все порты агрегированной группы должны работать в одном режиме, конфигурация мастера-порта распространяется на все порты в группе. Таким образом, при конфигурировании портов в группе агрегирования достаточно настроить мастер-порт.
- Важным моментом при реализации объединения портов в агрегированный канал является распределение трафика по ним. Если пакеты одного сеанса будут передаваться по разным портам агрегированного канала, то может возникнуть проблема на более высоком уровне модели OSI.
- Например, если два или более смежных кадров одного сеанса станут передаваться через разные порты агрегированного канала, то из-за неодинаковой длины очередей в их буферах может возникнуть ситуация, когда из-за неравномерной задержки передачи кадра более поздний кадр обгонит своего предшественника. Поэтому в большинстве реализаций механизмов агрегирования используются методы статического, а не динамического распределения кадров по портам, т.е. закрепление за определенным портом агрегированного канала потока кадров определенного сеанса между двумя узлами. В этом случае все кадры будут проходить через одну и ту же очередь и их последовательность не изменится. Обычно при статическом распределении выбор порта для конкретного сеанса выполняется на основе выбранного алгоритма агрегирования портов, т.е. на основании некоторых признаков поступающих пакетов.

В коммутаторах D-Link поддерживается 9 алгоритмов агрегирования портов:

- mac_source — MAC-адрес источника;
- mac_destination — MAC-адрес назначения;
- mac_source_dest — MAC-адрес источника и назначения;
- ip_source — IP-адрес источника;
- ip_destination — IP-адрес назначения;
- ip_source_dest — IP-адрес источника и назначения;
- l4_src_port — TCP/UDP-порт источника;
- l4_dest_port — TCP/UDP-порт назначения;
- l4_src_dest_port — TCP/UDP-порт источника и назначения.

В коммутаторах D-Link по умолчанию используется алгоритм mac_source (MAC-адрес источника)

Распределение потоков данных по каналам агрегированной линии связи для алгоритма mac_source_dest



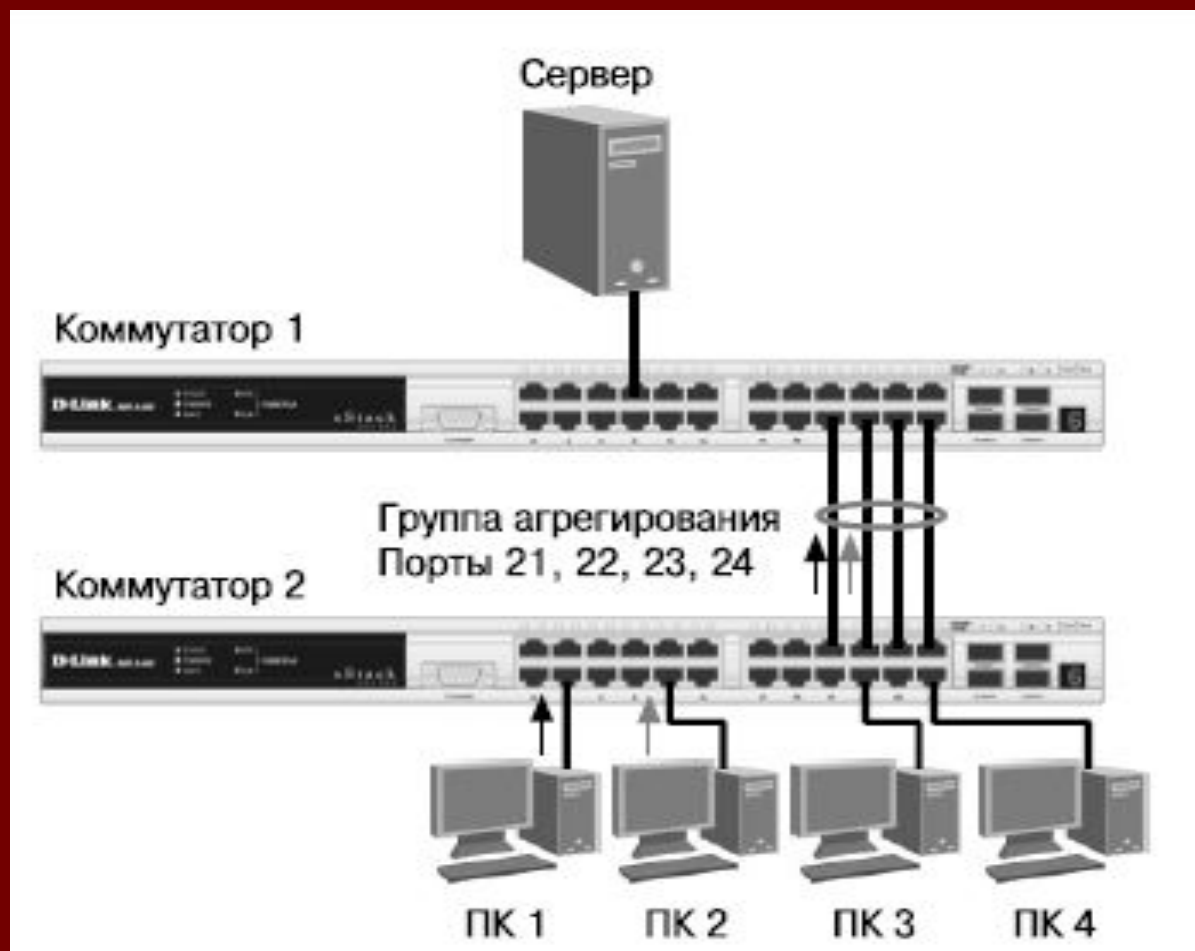
- Объединение каналов следует рассматривать как вариант настройки сети, используемый преимущественно для соединений «коммутатор – коммутатор» или «коммутатор – файл-сервер», требующих более высокой скорости передачи, чем может обеспечить одиночная линия связи. Также эту функцию можно применять для повышения надежности важных каналов связи. В случае повреждения линии связи объединенный канал быстро перенастраивается (не более чем за 1 сек.), а риск дублирования и изменения порядка кадров незначителен.
- Программное обеспечение коммутаторов D-Link поддерживает два типа агрегирования каналов связи:
 - статическое;
 - динамическое, на основе стандарта IEEE 802.3ad (LACP).

- При статическом агрегировании каналов (установлено по умолчанию), все настройки на коммутаторах выполняются вручную, и они не допускают динамических изменений в агрегированной группе.
- Для организации динамического агрегирования каналов между коммутаторами и другими сетевыми устройствами используется протокол управления агрегированным каналом — Link Aggregation Control Protocol (LACP).
- Протокол LACP определяет метод управления объединением нескольких физических портов в одну логическую группу и предоставляет сетевым устройствам возможность автосогласования каналов (их добавления или удаления) путем отправки управляющих кадров протокола LACP непосредственно подключенным устройствам с поддержкой LACP. Пакеты LACP отправляются устройством через все порты, на которых активизирован протокол. Порты, на которых активизирован протокол LACP, могут быть настроены для работы в одном из двух режимов: активном (*active*) или пассивном (*passive*).
- При работе в активном режиме порты выполняют обработку и рассылку управляющих кадров протокола LACP.
- При работе в пассивном режиме порты выполняют только обработку управляющих кадров LACP.

- Для того чтобы динамический канал обладал функцией автосогласования, рекомендуется порты, входящие в агрегированную группу, с одной стороны канала настраивать как активные, а с другой — как пассивные.
- Следует отметить, что у портов, объединяемых в агрегированный канал, нижеперечисленные характеристики должны обладать одинаковыми настройками:
 - тип среды передачи;
 - скорость;
 - режим работы — полный дуплекс;
 - метод управления потоком (Flow Control) .
- При объединении портов в агрегированный канал на них не должны быть настроены функции аутентификации 802.1X, зеркалирования трафика и блокировки портов.

Настройка статических агрегированных каналов

(для повышения пропускной способности канала связи между коммутатором 1, к которому подключен сервер, и коммутатором 2, к которому подключены пользователи, требуется объединить порты коммутаторов в статический агрегированный канал)



Настройка коммутаторов

Настройка коммутатора 1

Создать группу агрегирования (тип канала Static) и задать алгоритм агрегирования.

- **create link_aggregation group_id 1 type static**
- **config link_aggregation algorithm mac_destination**

Включить порты 21, 22, 23, 24 в группу и выбрать порт 21 в качестве мастера-порта.

- **config link_aggregation group_id 1 master_port 21 ports 21, 22, 23, 24 state enabled**

■

Настройка коммутатора 2

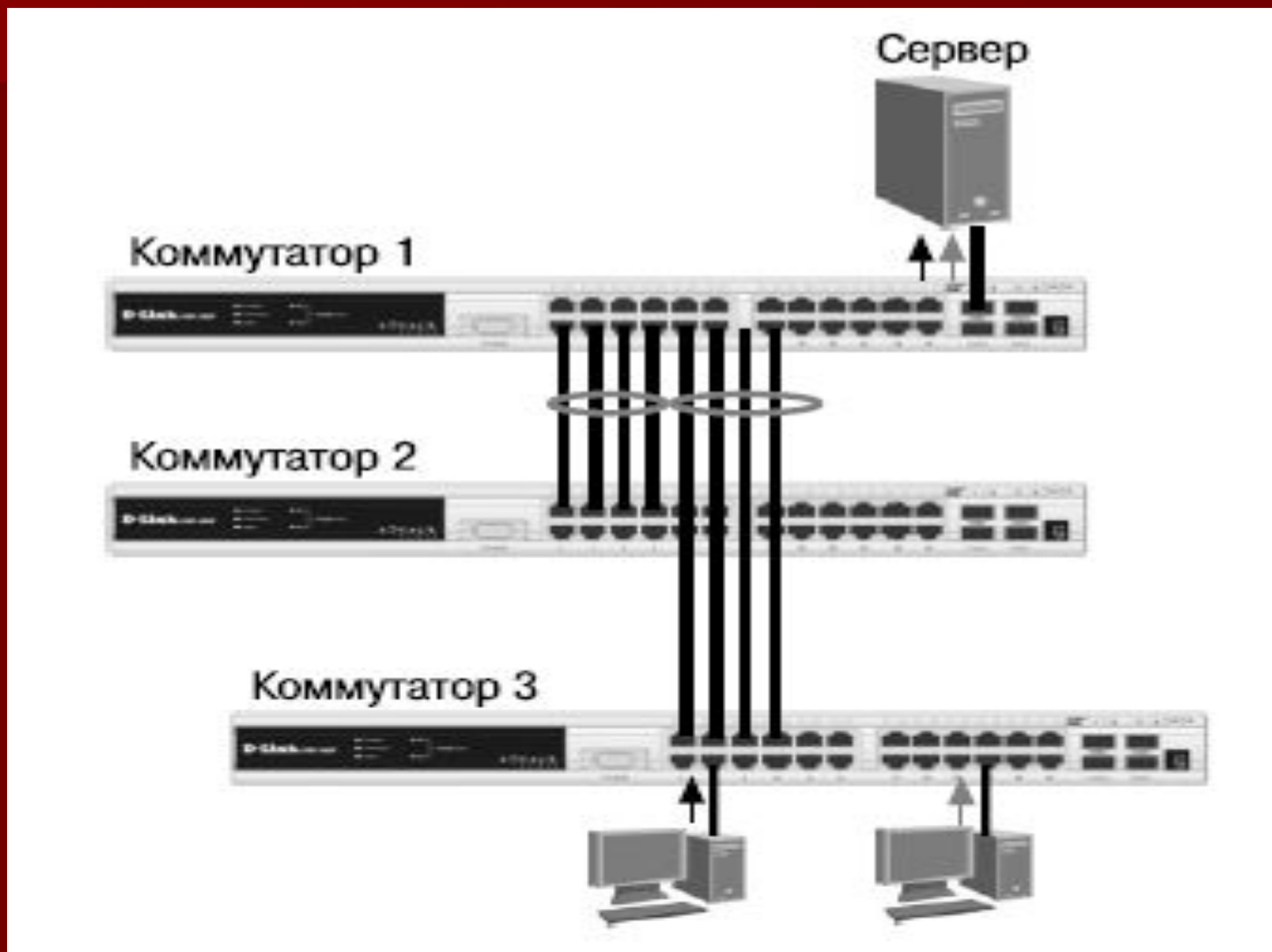
Создать группу агрегирования и задать алгоритм агрегирования.

- **create link_aggregation group_id 1 type static**
- **config link_aggregation algorithm mac_source**

Включить порты 21, 22, 23, 24 в группу и выбрать порт 21 в качестве мастера-порта.

- **config link_aggregation group_id 1 master_port 21 ports 21, 22, 23, 24 state enabled**

Настройка динамических агрегированных каналов (тип канала LACP)



Настройка коммутатора 1

Создать группы агрегирования (тип канала LACP) и задать алгоритм агрегирования.

- **create link_aggregation group_id 1 type lacp**
- **create link_aggregation group_id 2 type lacp**
- **config link_aggregation algorithm mac_destination**

Включить порты 1, 2, 3, 4 в группу 1 и выбрать порт 1 в качестве мастера-порта.

- **config link_aggregation group_id 1 master_port 1 ports 1-4 state enabled**

Включить порты 5, 6, 7, 8 в группу 2 и выбрать порт 5 в качестве мастера-порта.

- **config link_aggregation group_id 2 master_port 5 port 5-8 state enabled**

Настроить для портов 1-8 активный режим работы.

- **config lacp_port 1-8 mode active**

Настройка коммутаторов 2 и 3 (на портах 1-4 ЭТИХ коммутаторов включено автосогласование)

- **create link_aggregation group_id 1 type lacp**
- **config link_aggregation algorithm mac_source**
- **config link_aggregation group_id master_port 1 ports 1-4 state**
- **enabled**

Внимание:

- Если один конец агрегированного канала настроен как LACP, другой конец должен также иметь тип LACP. Если один конец имеет тип LACP, а другой Static, то соединение установлено не будет.
- Если коммутатор с поддержкой LACP требуется подключить к коммутатору, поддерживающему только статическое агрегирование, то тип агрегированного канала на коммутаторе LACP необходимо установить в Static.
- Не соединяйте физически соответствующие порты устройств до тех пор, пока не настроено агрегирование каналов, т.к. в коммутируемой сети может возникнуть петля.