

ЛОГИЧЕСКАЯ СТРУКТУРИЗАЦИЯ СЕТИ

Преимущества сети, построенной на общей разделяемой среде

- + Простая топология сети
- + Легкое наращивание числа узлов (в определённых пределах)
- + Отсутствие потерь кадров из-за переполнения буферов коммутационных узлов
- + Простота протоколов, обеспечивающая низкую стоимость оборудования

Ограничения сети, построенной на общей разделяемой среде

- Ограничения на число узлов
 - Ethernet: 1024
- Ограничения на максимальную длину каналов
- Невозможность организации резервных каналов
- Снижение эффективности алгоритмов доступа к общей среде с увеличением нагрузки

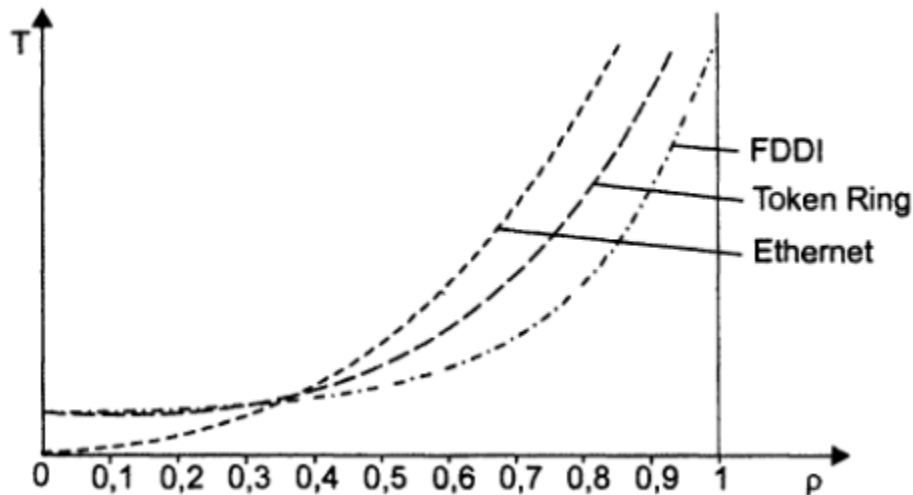
Зависимость эффективности алгоритмов общего доступа к среде от нагрузки

Коэффициент использования сети

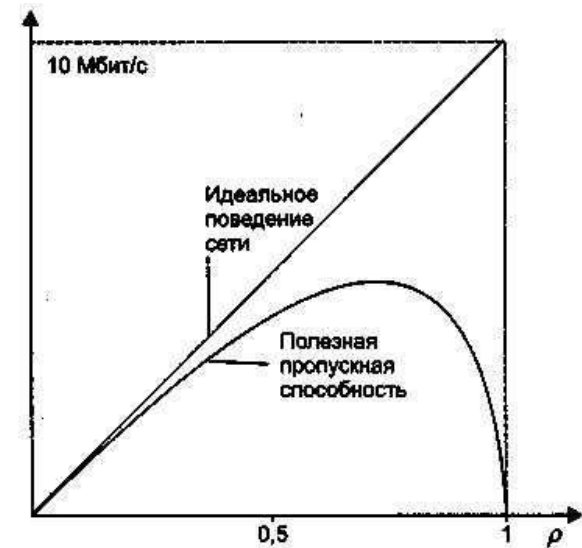
$$\rho = V_{curr} / V_{max},$$

где V_{curr} - текущая нагрузка на сеть, V_{max} - максимальная пропускная способность.

Задержки



Полезная пропускная способность



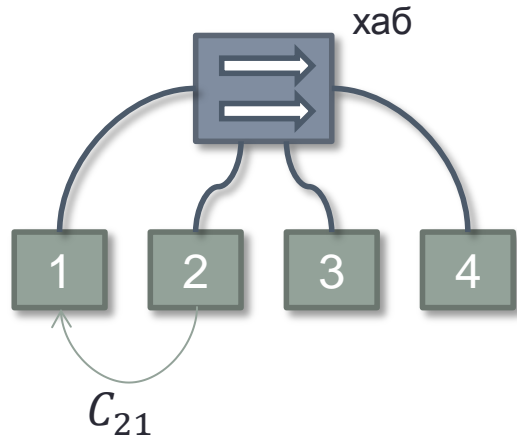
Логическая структуризация сети

- Проблемы, возникающие при использовании большой общей разделяемой среды можно решить, разделив сеть на несколько разделяемых сред и соединив их с помощью мостов или коммутаторов.
- Такой подход называют логической структуризацией сети.
- Получившиеся фрагменты с общей разделяемой средой называют
 - Логическими сегментами
 - Доменами коллизий

Преимущества логической структуризации сети

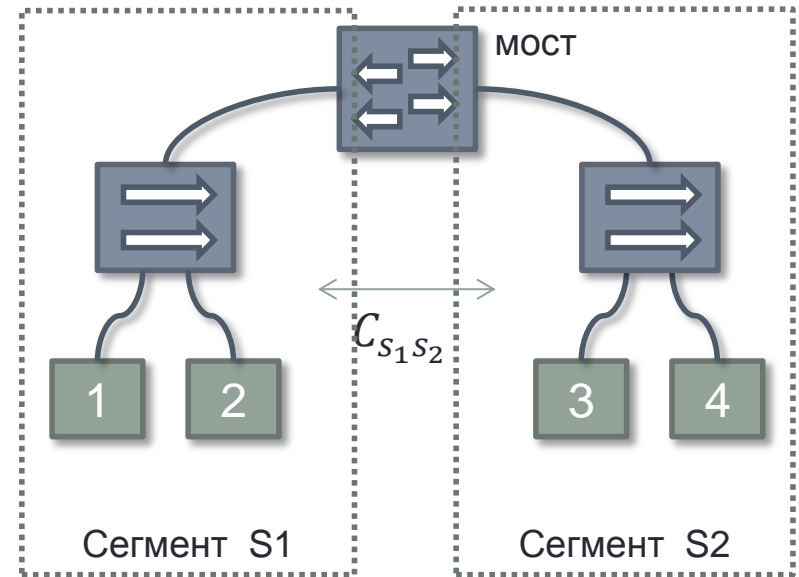
- + Повышается эффективность работы сети
- + Увеличивает степень защиты данных
- + Увеличивает гибкость построения сети
- + Способствует локализации неисправностей и проблем в сети
- + Облегчает управление сетью

Эффективность логической структуризации сети



Суммарный трафик

$$C = \sum_{i \neq j} C_{ij}$$



Трафик сегмента S1

$$C_{S1} = C_{S1} + C_{S1S2} = C - C_{S2},$$

трафик сегмента S2

$$C_{S2} = C_{S2} + C_{S1S2} = C - C_{S1},$$

где

$$C_{S_k} = \sum_{i \neq j, i \in S_k, j \in S_k} C_{ij}.$$

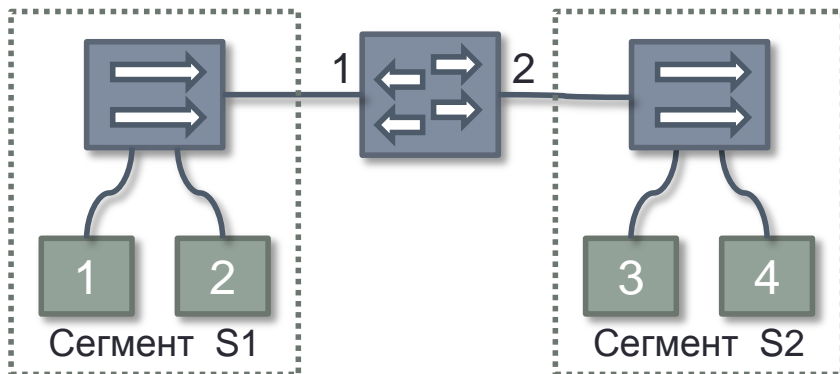
СЕТЕВЫЕ УСТРОЙСТВА КАНАЛЬНОГО УРОВНЯ

Коммутаторы и мосты

- Коммутаторы и мосты работают с кадрами канального уровня, принимая решения о передаче или не передаче кадров на основе MAC адресов.
 - Коммутатор является более высокопроизводительным устройством, обрабатывая множество кадров одновременно.
 - Мосты, как правило, обрабатывают кадры последовательно и используются для соединения двух сегментов.
- Порты коммутаторов работают как конечные узлы сегмента за исключением того, что они не имеют MAC адресов – они работают в неразборчивом режиме.
 - Адреса, как MAC, так и сетевые, могут иметь блоки управления коммутатором, если они присутствуют.
- Коммутаторами и мостами используются следующие алгоритмы
 - Алгоритм прозрачного моста IEEE 802.1D
 - Алгоритм коммутации от источника Source Routing

Алгоритм прозрачного моста IEEE802.1D

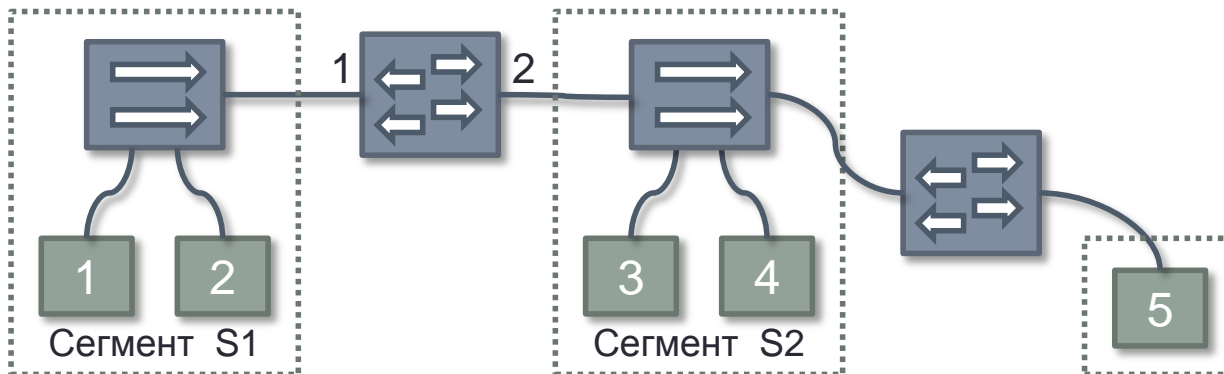
- Для принятия решения о действии с кадром коммутатор использует таблицу коммутации (таблица фильтрации, forwarding table)
- Записи в таблице связывают MAC адрес устройства, и порт коммутатора к которому это устройство подключено.
- Может быть два вида записей
 - Статические записи – создаются вручную, не имеют срока жизни.
 - Динамические записи – автоматически добавляются коммутатором на основании анализа обратных адресов принимаемых кадров, по истечении определённого таймаута удаляются.



MAC адрес	Порт
1	1
2	1
3	2
4	2

Алгоритм прозрачного моста IEEE802.1D

- Для принятия решения о действии с кадром коммутатор использует таблицу коммутации (таблица фильтрации, forwarding table)
- Записи в таблице связывают MAC адрес устройства, и порт коммутатора к которому это устройство подключено.
- Может быть два вида записей
 - Статические записи – создаются вручную, не имеют срока жизни.
 - Динамические записи – автоматически добавляются коммутатором на основании анализа обратных адресов принимаемых кадров, по истечении определённого таймаута удаляются.



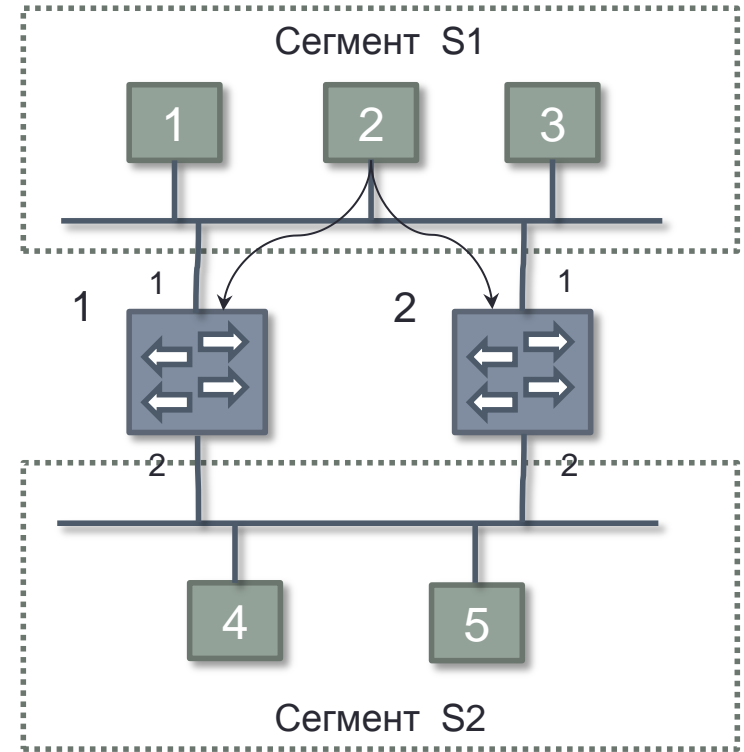
MAC адрес	Порт
1	1
2	1
3	2
4	2
5	2

Алгоритм прозрачного моста IEEE802.1D

1. Кадр принимается и запоминается в буфере входящего порта.
2. Создаётся запись в таблице коммутации с адресом отправителя кадра и номером порта, на котором он был принят.
3. Коммутатор ищет в таблице коммутации запись с MAC-адресом получателя:
 - Запись найдена – анализируем порт
 - Порт получателя кадра совпадает с портом, указанным в таблице
⇒ Кадр удаляется из буфера – **фильтрация / filtering**
 - Кадр получен на порт, не совпадающий с указанным в таблице
⇒ Кадр передаётся через порт, указанный в таблице – **продвижение / forwarding**
 - Запись не найдена
⇒ Кадр передаётся на все порты коммутатора – **затопление / flooding сети**
 - Широковещательные пакеты передаются в режиме затопления сети.

Алгоритм прозрачного моста и петли

1. Узел 2 отправляет широковещательный пакет

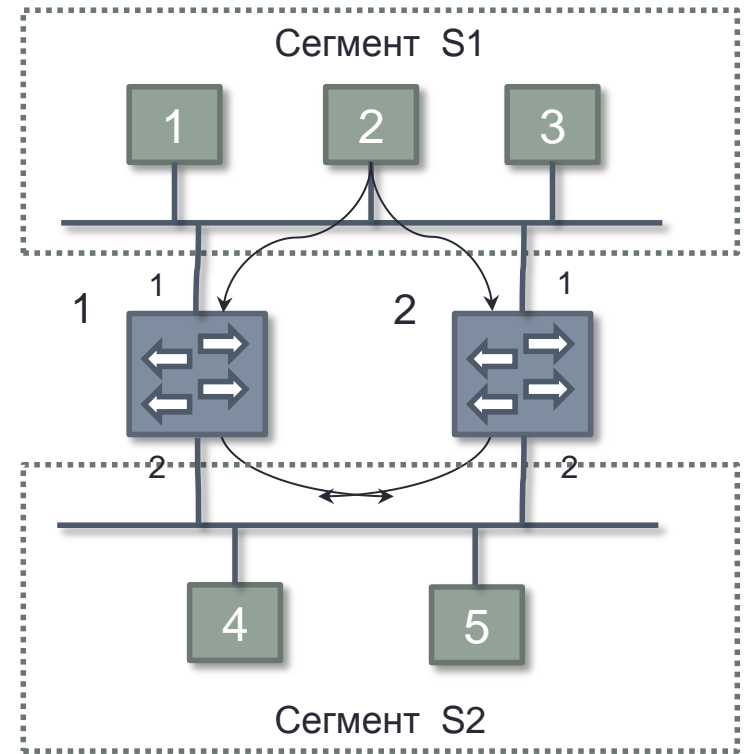


MAC	Порт

MAC	Порт

Алгоритм прозрачного моста и петли

1. Узел 2 отправляет широковещательный пакет
2. Оба коммутатора продвигают кадр в сегмент S2 и запоминают, что узел 2 подключен к порту 1.

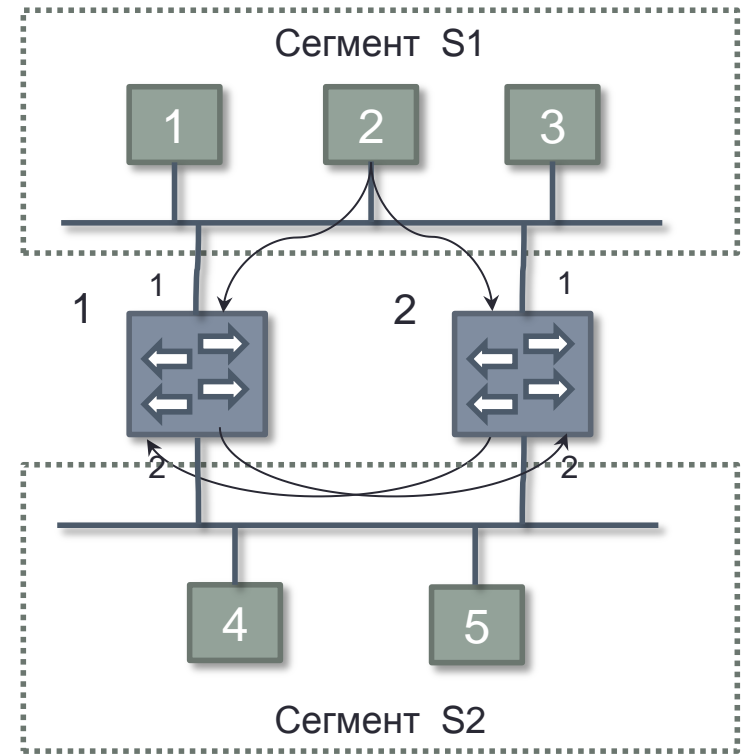


MAC	Порт
2	1

MAC	Порт
2	1

Алгоритм прозрачного моста и петли

1. Узел 2 отправляет широковещательный пакет
2. Оба коммутатора продвигают кадр в сегмент S2 и запоминают, что узел 2 подключен к порту 1.
3. Коммутаторы принимают кадр, запоминают, что узел 2 подключен к порту 2 и продвигают кадр в сегмент S1.

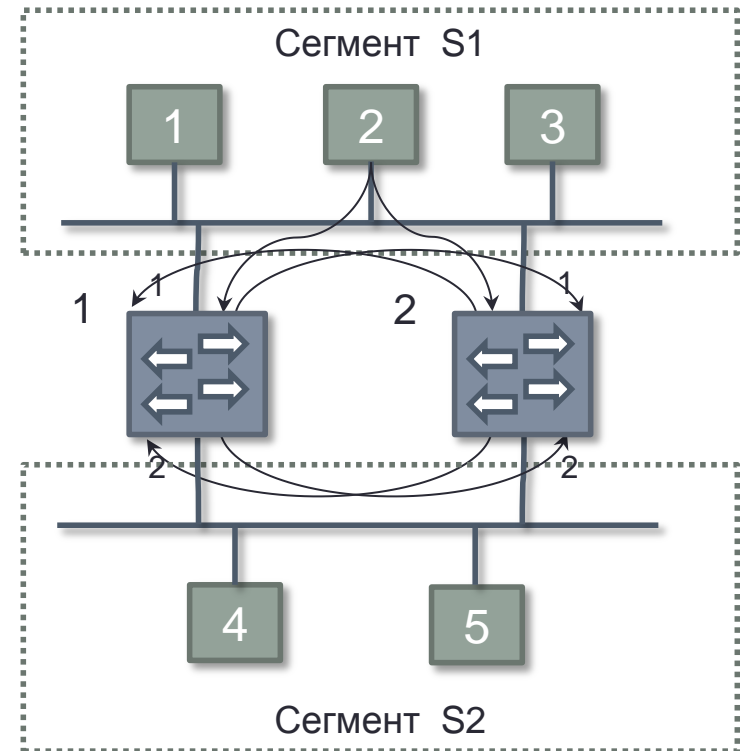


MAC	Порт
2	2

MAC	Порт
2	2

Алгоритм прозрачного моста и петли

1. Узел 2 отправляет широковещательный пакет
2. Оба коммутатора продвигают кадр в сегмент S2 и запоминают, что узел 2 подключен к порту 1.
3. Коммутаторы принимают кадр, запоминают, что узел 2 подключен к порту 2 и продвигают кадр в сегмент S1.
4. Коммутаторы принимают кадр из сегмента S1, запоминают что узел 2 подключен к порту 1 и передают кадр в сегмент S2.
5. ...

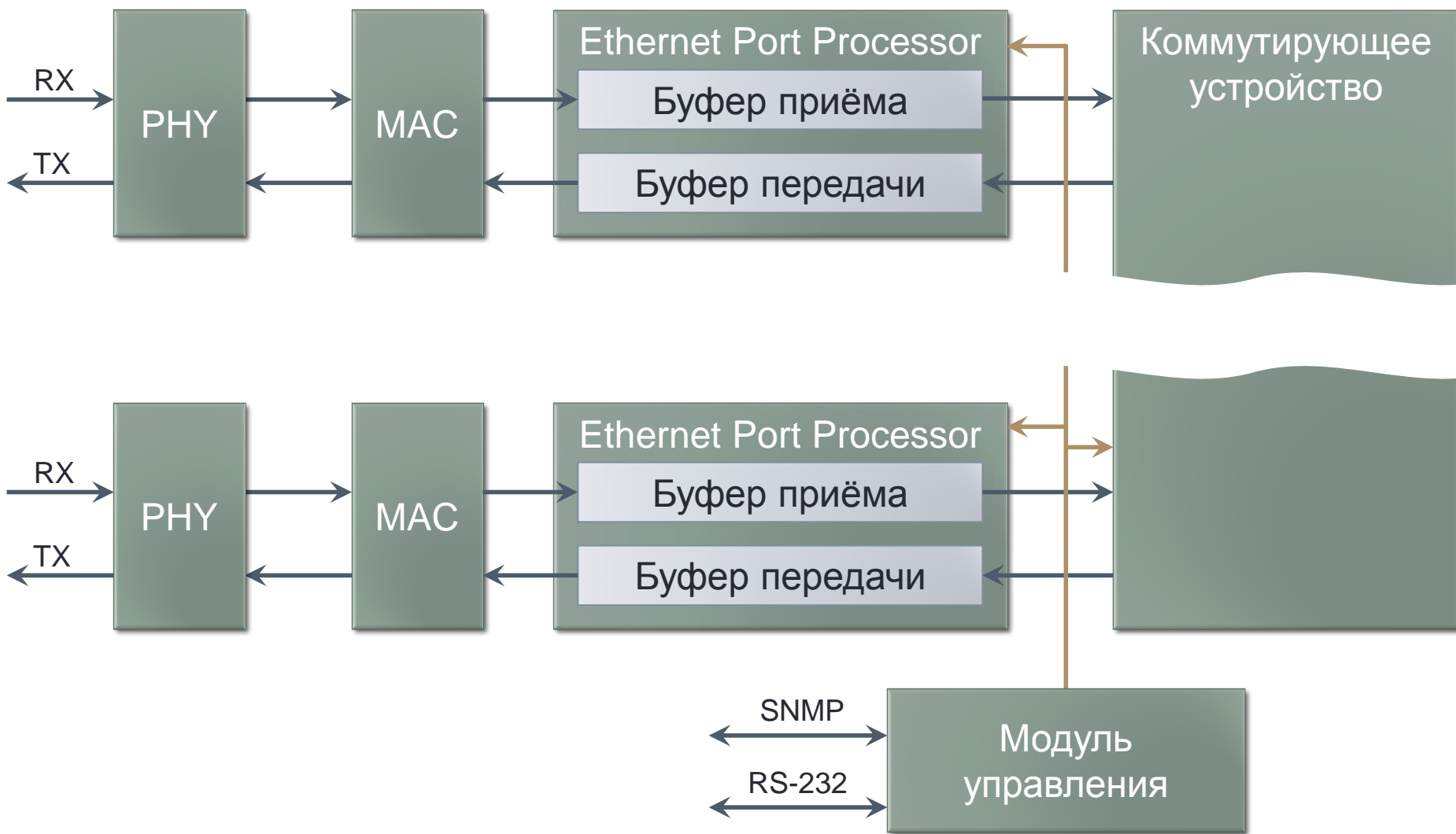


MAC	Порт
2	1

MAC	Порт
2	1

- Пакеты дублируются
- Пакеты бесконечно циркулируют по сети
- Записи таблицы коммутации постоянно меняются.

Архитектура коммутаторов



Производительность коммутаторов

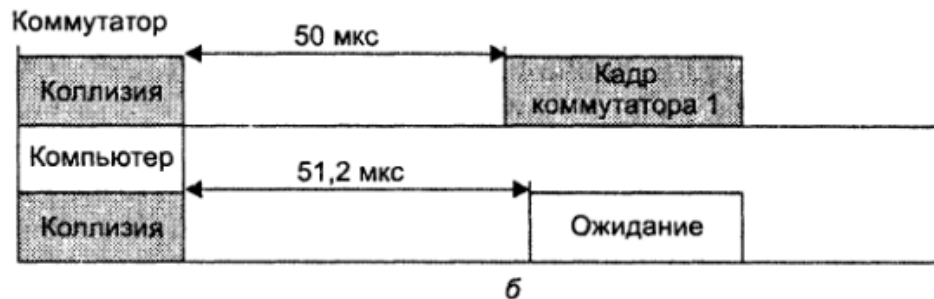
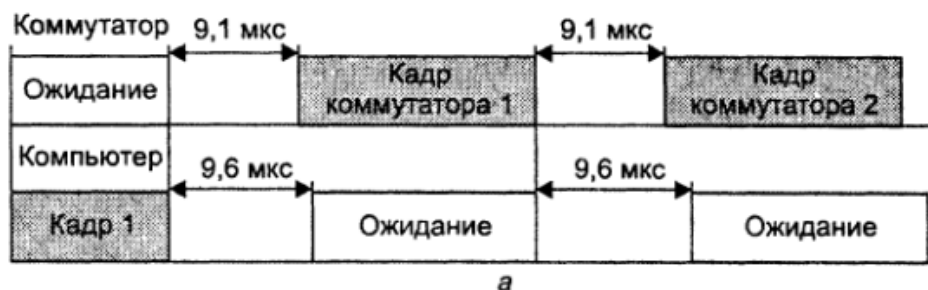
- Показатели производительности коммутаторов:
 - скорость фильтрации кадров;
 - скорость продвижения кадров;
 - пропускная способность;
 - задержка передачи кадра.
- Основные характеристики, влияющие на производительность:
 - тип коммутации — сквозной или с полной буферизацией;
 - размер буфера (буферов) кадров;
 - производительность внутренней шины;
 - производительность процессоров;
 - размер таблицы коммутации.

Режимы работы коммутаторов

- С промежуточным хранением (Store and Forward).
 - Коммутатор читает всю информацию в кадре, проверяет его на отсутствие ошибок, выбирает порт коммутации и после этого посылает в него кадр.
- Сквозной (cut-through).
 - Коммутатор считывает в кадре только адрес назначения и после выполняет коммутацию. Этот режим уменьшает задержки при передаче, но при этом нет возможности выполнять контроль ошибок и использовать некоторые другие специальные функции.
- Бесфрагментный (fragment-free) или гибридный.
 - Этот режим является модификацией сквозного режима. Передача осуществляется после фильтрации фрагментов коллизий (кадр размером 64 байта обрабатываются по технологии store-and-forward, остальные по технологии cut-through).

Борьба с перегрузками

- Метод обратного давления
 - Создание искусственных коллизий в сегменте, который посылает слишком много кадров в коммутатор.
- Агрессивный захват среды

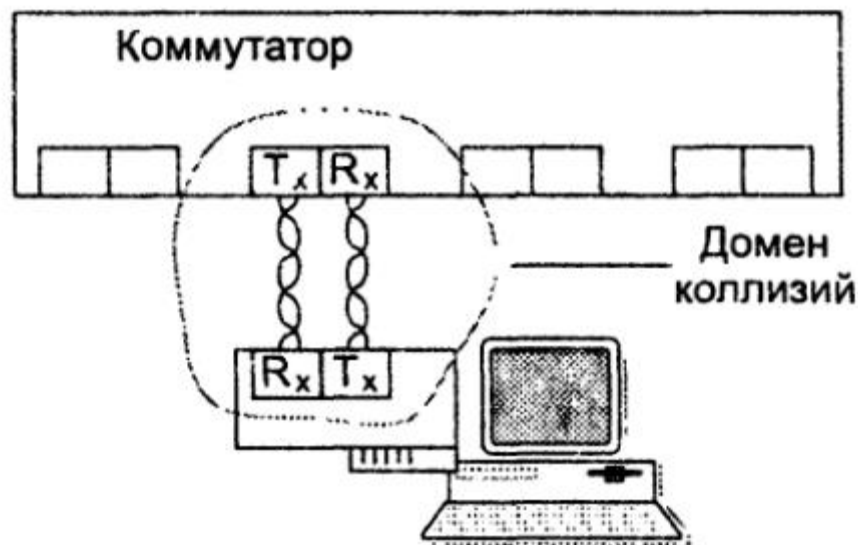


Дополнительные функции коммутаторов

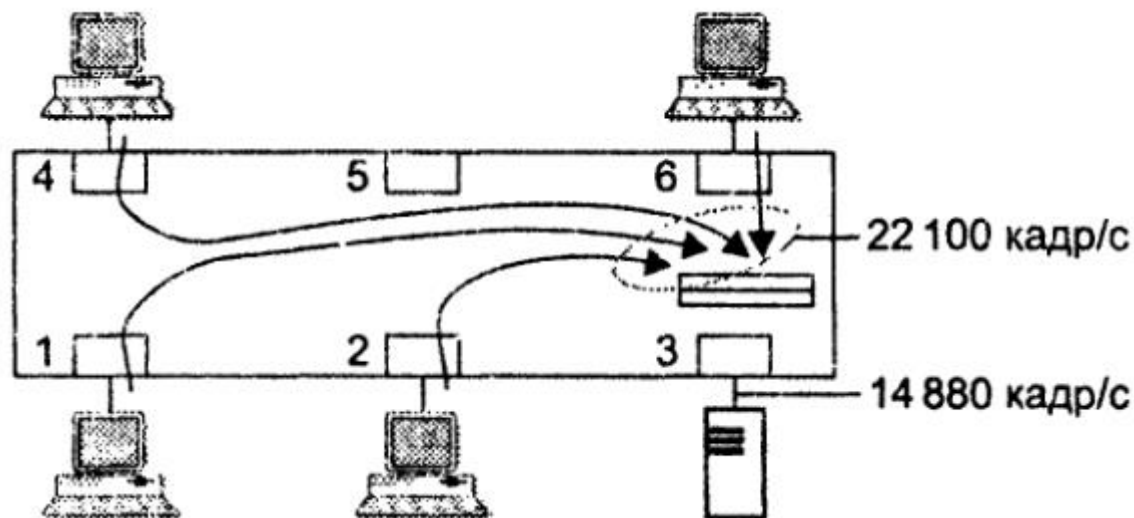
- Трансляция протоколов канального уровня
- Удалённое управление
- Зеркалирование / Port mirroring
 - Дублирование определённого трафика на выбранный порт (с целью мониторинга)
- Агрегирование / Link aggregation, Port trunking, bonding
 - Объединение нескольких физических каналов в один логический. IEEE 802.3ad Link aggregation for parallel links
- Поддержка QoS
- Поддержка фильтрации
- Аутентификация и управление доступом
- Поддержка протокола Spanning Tree
- Поддержка VLAN
- Поддержка мультикаста / IGMP Snooping

Полнодуплексный режим работы

- Микросегментация – подключение компьютеров непосредственно к портам коммутатора.



Полнодуплексный режим работы Перегрузки



- Механизм обратной связи – IEEE802.3x (1997 год)
 - Введён подуровень управления уровнем MAC
 - Кадр подуровня управления – пауза
 - Используется инкапсуляция LLC
 - В качестве адреса назначения используется зарезервированный мультикастовый адрес 01:80:C2:00:00:01
 - Длина паузы измеряется в 512 битных интервалах, диапазон 0...65535
 - Не работает с QoS / VLAN

Протокол Spanning Tree (STP)

1. Определить корневой коммутатор / Root switch
 - Выбирается коммутатор с наименьшим идентификатором
2. Найти минимальный путь от всех коммутаторов и сегментов до корневого коммутатора
 - Root port (RP) – Порт, соединяющий коммутатор с корневым.
 - Designated port (DP) – Порт, соединяющий сегмент с корневым коммутатором.
3. Отключить все порты, не задействованные в минимальных путях
 - Blocked port (BP) – заблокированный порт.

Протокол Spanning Tree

Расстояние до корня

- Расстояние до корня определяется как суммарное условное время на передачу данных от порта данного коммутатора до порта корневого коммутатора.
- Задержки на передачу пакетов коммутатором игнорируются.
- Учитываются только условное время сегмента: время, затрачиваемое на передачу одного бита информации в наносекундах между непосредственно связанными по сегменту сети портами.

Data rate	STP Cost (802.1D-1998)	RSTP Cost (802.1W-2001)
4 Mbit/s	250	5,000,000
10 Mbit/s	100	2,000,000
16 Mbit/s	62	1,250,000
100 Mbit/s	19	200,000
1 Gbit/s	4	20,000
2 Gbit/s	3	10,000
10 Gbit/s	2	2,000

Протокол Spanning Tree

- При включении каждый коммутатор считает себя корневым
- Корневой коммутатор рассылает пакеты Bridge Protocol Data Unit (BPDU), содержащие идентификатор корневого маршрутизатора, расстояние до корня (0) и идентификатор порта, через который отправлен BPDU.
- При получении BPDU с меньшим идентификатором коммутатор перестаёт считать себя корневым и начинает ретранслировать другие BPDU
 - Расстояние до корня увеличивается на условное время сегмента, из которого был принят BPDU
- Определение корневого порта:
 - Порт, по которому был принят BPDU с минимальным расстоянием до корня
- Определение назначенного порта:
 - Не рассматриваем корневой порт
 - Выбираем порт, у которого минимальное расстояние в принятых BPDU больше, чем расстояние для корневого порта
- Остальные порты объявляются отключенными
 - Используются только для приёма BPDU

Протокол Spanning Tree

- Разрешение коллизий:
 - В случае, у коммутатора есть несколько путей минимальной длины до корневого, то он выбирает путь, который проходит через соседний коммутатор с минимальным идентификатором.
 - В случае, если имеется несколько кандидатов на роль назначенного порта для сегмента, выбирается порт коммутатора с наименьшим идентификатором.
 - В случае, если два коммутатора соединены несколькими кабелями, выбирается кабель проходящий через порт соседа с минимальным номером.
- Поддержание работы сети
 - Корневой маршрутизатор продолжает рассылать BPDU во время работы сети.
 - В случае изменения конфигурации сети обнаруживший изменение коммутатор посылает сообщение Topology Change Notification (TCN) корневому.
 - Корневой коммутатор получив TCN начинает рассылать BPDU с флагом Topology Change, что уведомляет коммутаторы о начале перестроения дерева.

Протокол Spanning Tree

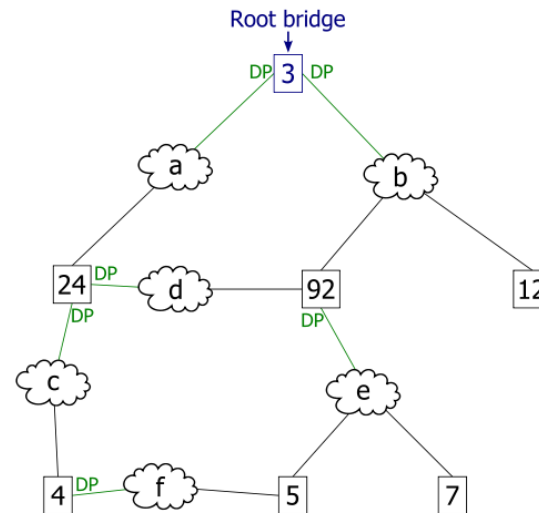
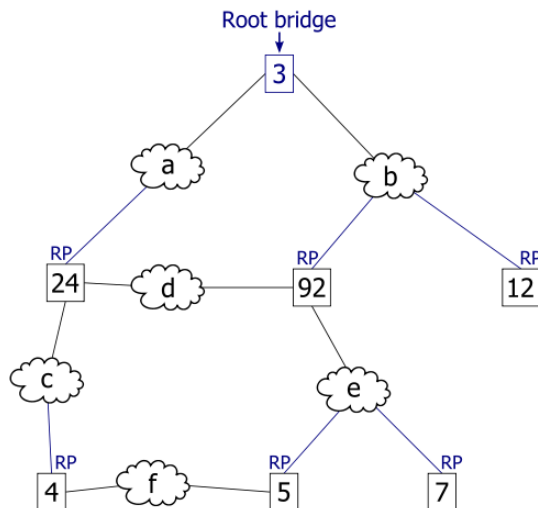
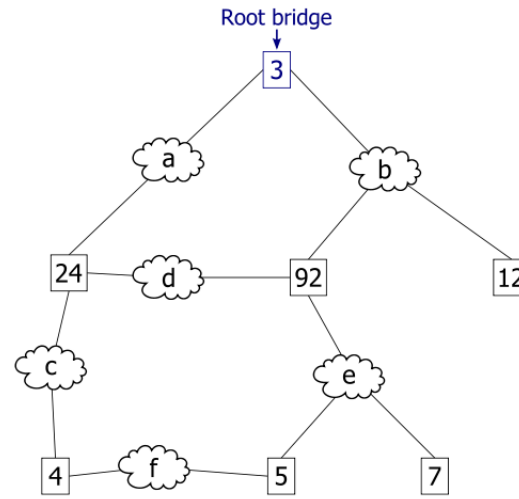
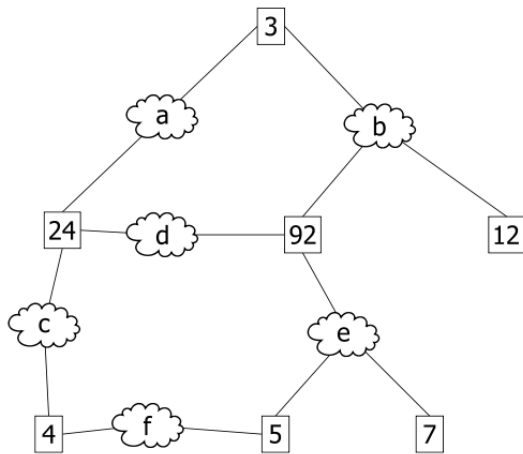
Формат BPDU

Field length, in bytes	2	1	1	1	8	4	8	2	2	2	2	2
Protocol identifier	Version	Message type	Flags	Root ID	Root path cost	Bridge ID	Port ID	Message age	Maximum age	Hello time	Forward delay	

- **Protocol Identifier** – Должно быть равно 0.
- **Version** - Должно быть равно 0.
- **Message Type** - Должно быть равно 0.
- **Flag** – флаги:
 - topology-change (TC) сигнализирует о изменении топологии.
 - topology-change acknowledgment (TCA) подтверждает получение BPDU с флагом TC.
- **Root ID** – Идентификатор корневого коммутатора (2 байта приоритет, 6 байт идентификатор).
- **Root Path Cost** – Стоимость пути до корневого маршрутизатора.
- **Bridge ID** – Идентификатор коммутатора, отправившего BPDU.
- **Port ID** – Идентификатор порта, через который отправлен BPDU.
- **Message Age** – Время отправки BPDU.
- **Maximum Age** – Время, в которое BPDU должен быть уничтожен.
- **Hello Time** – Интервал между отправкой BPDU корневым коммутатором.
- **Forward Delay** – Задержка, которую коммутаторы должны выдержать после изменения топологии перед переводом портов в новое состояние.

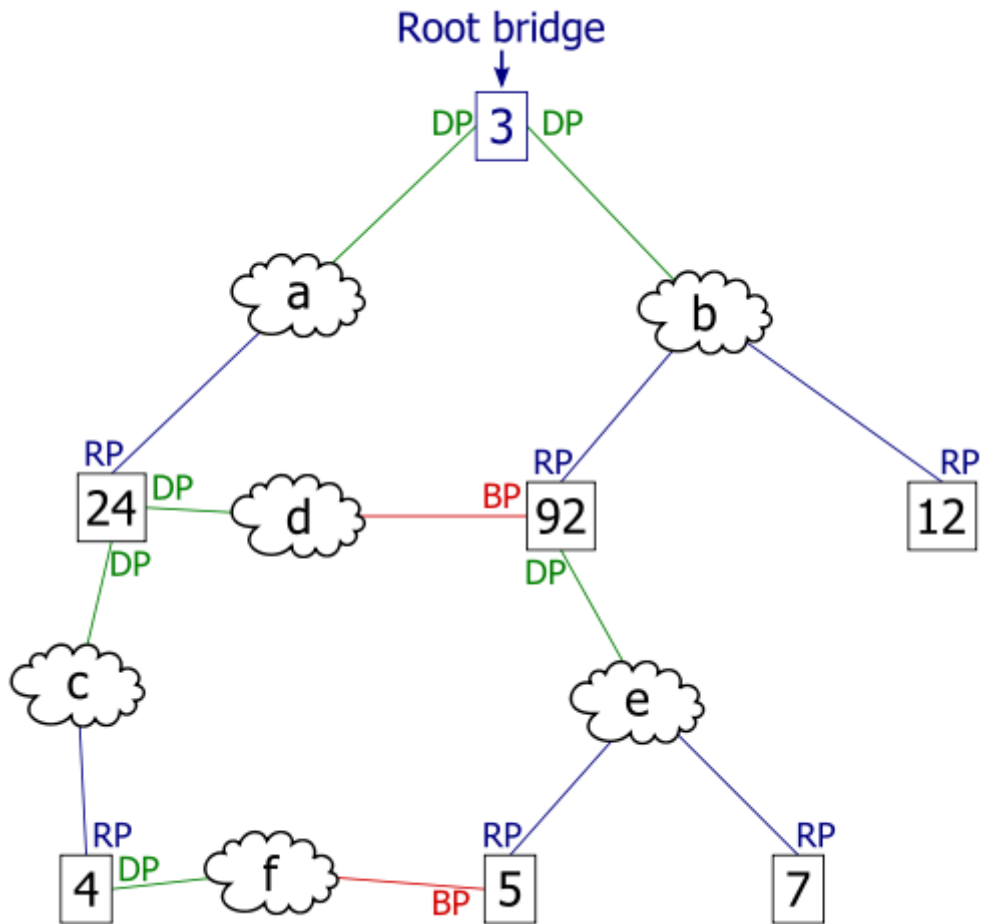
Протокол Spanning Tree

Пример



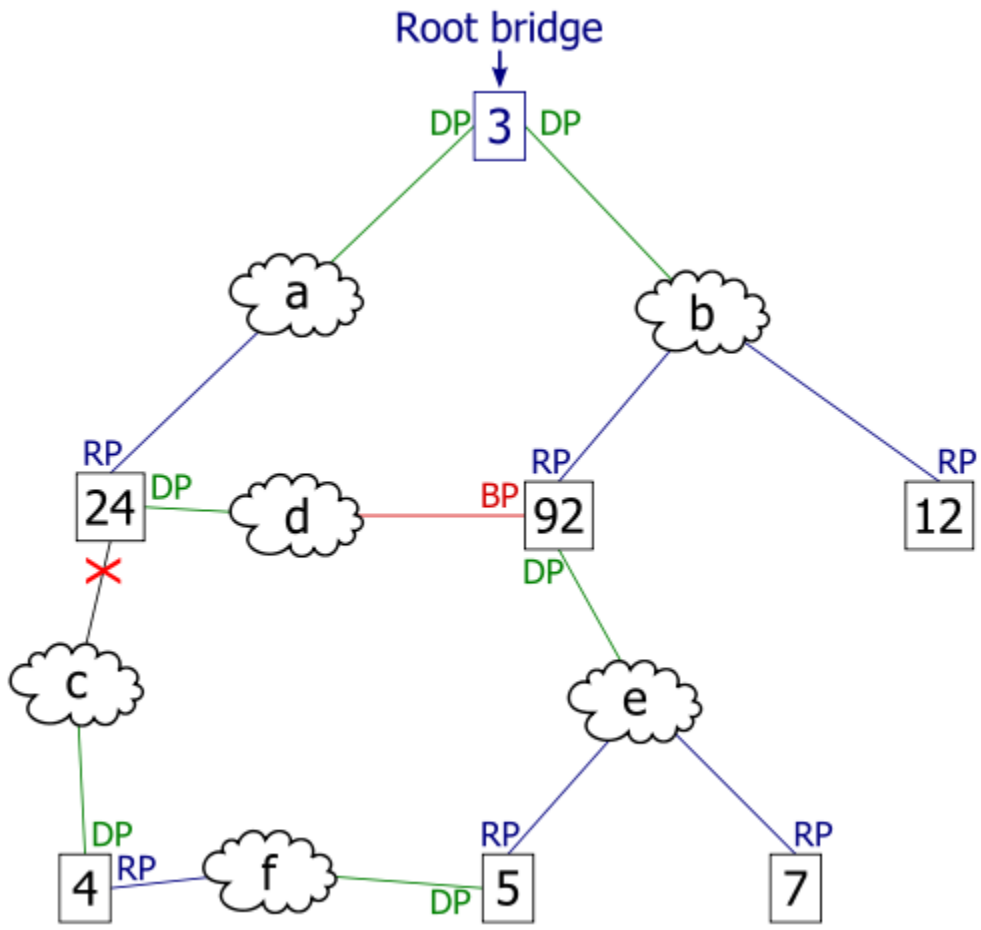
Протокол Spanning Tree

Пример



Протокол Spanning Tree

Пример



Протокол Spanning Tree

Развитие протокола

- 1985 фирмой DEC разработана первая версия протокола
- 1990 Протокол Spanning Tree утверждён стандартом IEEE 802.1D
- 2001 Утверждён стандарт 802.1w Rapid Spanning Tree Protocol (RSTP)
- 2004 в стандарт 802.1D-2004 введён протокол RSTP, протокол STP объявлен устаревшим.
- Протоколы с поддержкой виртуальных локальных сетей
 - Per-VLAN Spanning Tree (PVST) – протокол фирмы Cisco
 - VLAN Spanning Tree Protocol (VSTP) – протокол фирмы Juniper Networks
 - Multiple Spanning Tree Protocol (MSTP) – определён в IEEE 802.1s (1998), затем включен в IEEE 802.1Q-2005

Протокол Spanning Tree

Развитие протокола

- 1985 фирмой DEC разработана первая версия протокола
- 1990 Протокол Spanning Tree утверждён стандартом IEEE 802.1D
- 2001 Утверждён стандарт 802.1w Rapid Spanning Tree Protocol (RSTP)
- 2004 в стандарт 802.1D-2004 введён протокол RSTP, протокол STP объявлен устаревшим.
- Протоколы с поддержкой виртуальных локальных сетей
 - Per-VLAN Spanning Tree (PVST) – протокол фирмы Cisco
 - VLAN Spanning Tree Protocol (VSTP) – протокол фирмы Juniper Networks
 - Multiple Spanning Tree Protocol (MSTP) – определён в IEEE 802.1s (1998), затем включен в IEEE 802.1Q-2005
- 2012 год – утверждён стандарт 802.1aq Shortest Path Bridging, призванный заменить STP, RSTP и MSTP

Особенности протокола RSTP

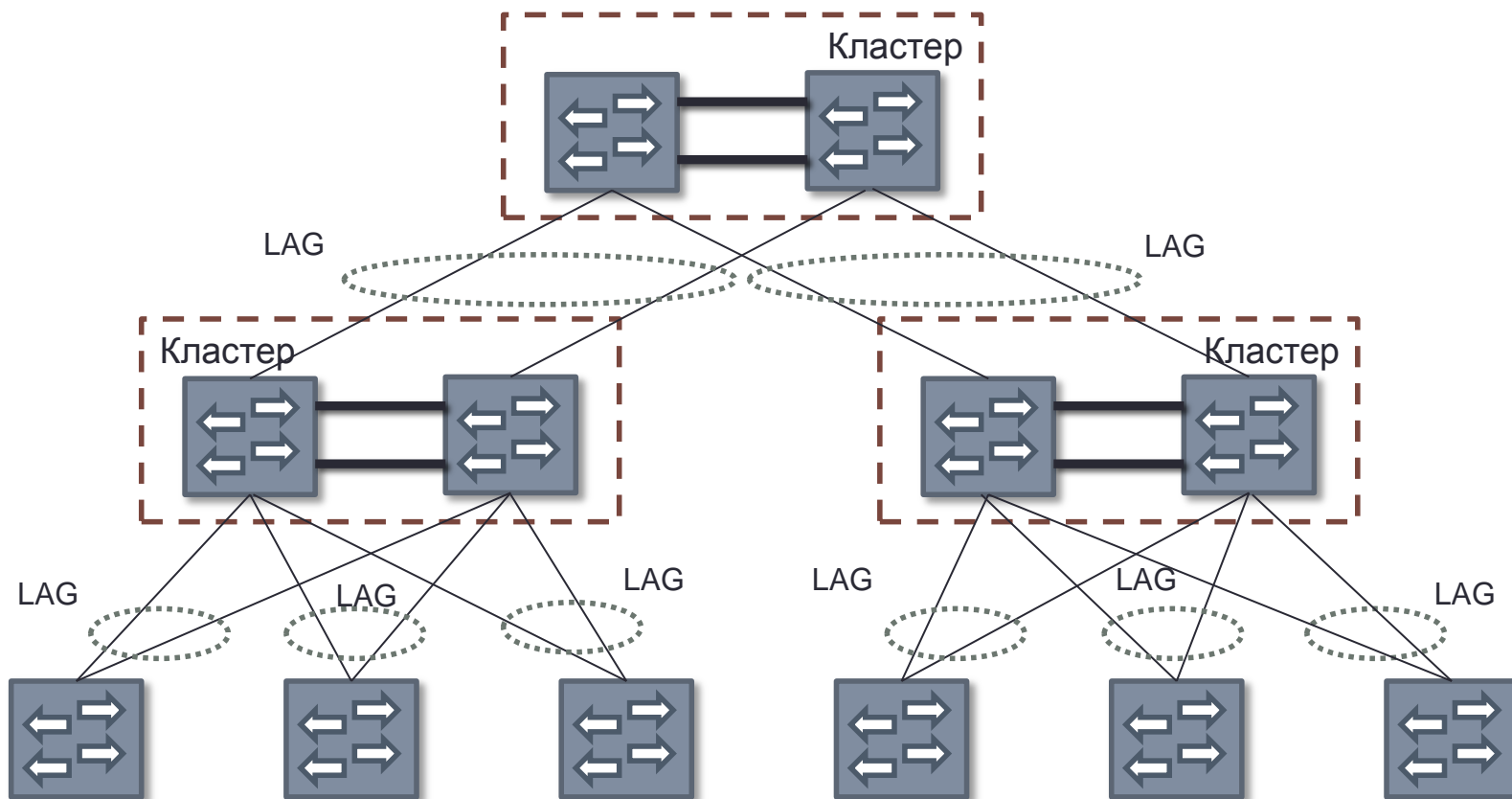
STP (802.1d)	RSTP(802.1w)
Работа в сложившейся топологии	
Корневой коммутатор передаёт BPDU, остальные ретранслируют.	Все коммутаторы шлют BPDU в соответствии с интервалом hello.
Состояния портов	
<ul style="list-style-type: none">- блокировка (blocking)- прослушивание (listening)- обучение (learning)- продвижение (forwarding)- отключен (disabled)	<ul style="list-style-type: none">- отбрасывание (discarding)- обучение (learning)- продвижение (forwarding)
Роли портов	
<ul style="list-style-type: none">- корневой (root)- назначенный (designated)- не назначенный (non-designated)	<ul style="list-style-type: none">- корневой (root)- назначенный (designated)- дополнительный (alternate)- резервный (backup)

Особенности протокола RSTP

STP (802.1d)	RSTP(802.1w)
Действия при изменении топологии	
Обнаруживший изменения извещает корневой коммутатор, тот даёт команду очистить записи о топологии в течение Forward Delay Timer (15 сек).	Обнаружение изменения влечёт немедленную очистку записей.
Начало действий	
Если BPDU не приходил в течение Max Age (20 сек.)	Если BPDU не приходил в течение 3 интервалов Hello (3 x 2 сек.)
Состояния порта при изменении топологии	
<ul style="list-style-type: none">- Blocking (20 сек)- Listening (15 сек)- Learning (15 сек)- Forwarding	Для портов p2p (full duplex, не отмечены как пограничные), используется механизм proposal and agreement для быстрого установления связей.

Построение эффективных отказоустойчивых сетей с избыточными каналами

- Используя кластеризацию коммутаторов и агрегацию связей (LAG) можно эффективно использовать дополнительные каналы



Агрегирование портов

- Стандарты
 - 2000 год – 802.3ad
 - 2008 год – 802.1AX
- Требования к каналам
 - полный дуплекс
 - точка-точка
 - одинаковая скорость передачи
- Распределение кадров
 - По пакетам
 - Должно обеспечить:
 - Отсутствие дублирования
 - Сохранение порядка кадров для любого «разговора»

Агрегирование портов

- Ручная настройка
- Использование протокола LACP
 - Коммутатор передаёт кадры LACPDU по всем портам, на которых включен LACP в активном режиме
 - Если включен пассивный режим LACP, то коммутатор будет только отвечать на поступление кадров LACPDU
 - При обнаружении параллельных каналов автоматически включается агрегирование
 - Позволяет выявлять отказ каналов

Агрегирование портов

- + Не изменяет формат кадров
- + Не требует дополнительных буферов
- + Не изменяет порядок кадров
- + Не добавляет заметных задержек

- Не всегда приводит к линейному росту пропускной способности
 - Не увеличивает пропускную способность для одного «разговора»
- Стоимость растёт линейно.

Виртуальные локальные сети VLAN

- Виртуальной сетью (VLAN) называется группа узлов сети, трафик которой, в том числе и широковещательный, на канальном уровне полностью изолирован от других узлов сети.
- Виртуальная сеть образует домен широковещательного трафика.
- Объединение виртуальных сетей осуществляется с помощью сетевого уровня.
- Введение виртуальных сетей позволяет:
 - повысить производительность каждой виртуальной сети;
 - изолировать сети для:
 - управления доступом;
 - защиты от широковещательных штормов.
 - управлять логической структурой сети без изменения физической структуры.

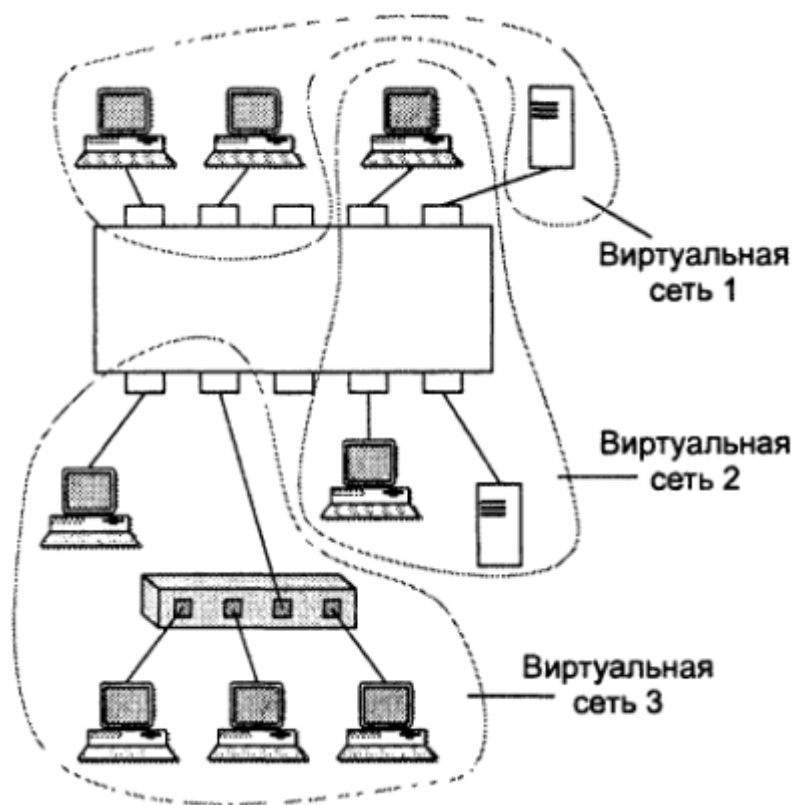
Виртуальные локальные сети VLAN

Способы построения

- Группировка портов.
- Группировка MAC-адресов.
- Использование меток в дополнительном поле кадра - частные протоколы и спецификации IEEE 802.1 Q/P.
- Спецификация LANE для ATM-коммутаторов
- Использование сетевого уровня.

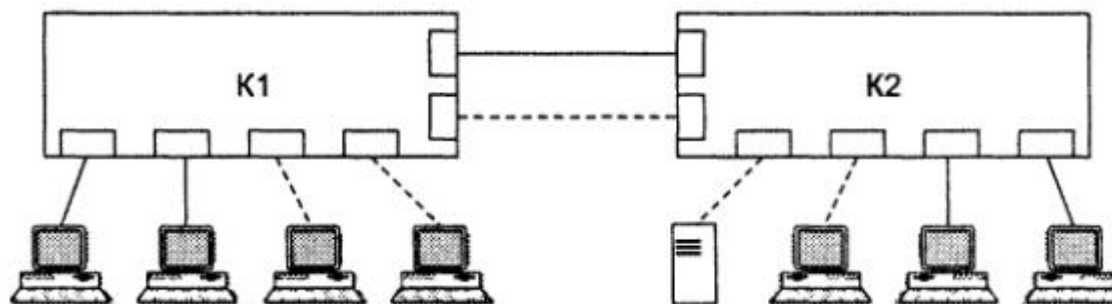
Виртуальные локальные сети

Группировка портов



Виртуальные локальные сети

Группировка портов



Виртуальные локальные сети

Использование адресов

- Использование MAC адресов
 - Позволяет избежать дублирования каналов в случае использования нескольких коммутаторов
 - Много ручной работы по настройке.
- Использование полей (адресов) сетевого уровня
 - Требуется, чтобы коммутатор понимал поля протоколов 3-го уровня

Виртуальные локальные сети

Метки в дополнительных полях кадра

- В кадр добавляется специальное поле, идентифицирующее VLAN
- Такие кадры передаются только между коммутаторами, поддерживающими такую технологию
 - Ingress Traffic – входящий трафик, нужно пометить тегом
 - Progress Traffic – внутренний тегированный трафик
 - Egress Traffic – исходящий трафик, нужно убрать тег
- Стандарт 802.1Q/802.1p
 - 4 байтный тег:
 - 2 байта Tag Protocol Identifier (TPID) - 0x8100
 - 2 байта Tag Control Information (TCI):
 - 3 бита приоритет
 - 1 бит Canonical Format Indicator (CFI)
 - 12 бит идентификатор VLAN
 - Протокол 802.1p - Generic Attributes Registration Protocol (GARP)
 - GARP Multicast Registration Protocol (GMRP)
 - GARP VLAN Registration Protocol (GVRP)