Канално ниво в локалните мрежи

7. Методи на достъп до съобщителната среда в Ethernet
8. Управление на канала в Ethernet. Превключватели и мостове. Виртуални локални мрежи и протокол Spanning Tree.

LANs - мрежите с общодостъпно предаване

Мрежите с общодостъпно предаване се характеризират с общ комуникационен канал, който се споделя от всички машини, включени в мрежата.

Всеки изпратен кадър минава през общия канал и достига до всички машини в мрежата. Адресно поле в кадъра посочва за кой е предназначен този кадър.

Когато една машина получи кадър, тя проверява дали той е предназначен за нея. Ако това е така, кадърът се приема и обработва, в противен случай се отхвърля.

Мрежи с общодостъпно предаване

При мрежите с общодостъпно предаване основен проблем е да се определи кой да започне да използва канала, дали да има състезание или поредност.

Протоколите, които разрешават този проблем се отнасят към подниво на каналния слой, наречено **подниво за достъп до средата** (medium access control - MAC). Нариичат се още протоколи за множествен достъп (Multiple Access)

Регионалните мрежи използват връзки "точка-точка" (point-to-point), докато общодостъпни многоточкови (multipoint) канали се използват най-вече при локалните мрежи.

Мрежи с общодостъпно предаване

Протоколите (процедурите) за достъп до канала се делят на две основни групи:

- детерминирани и
- състезателни

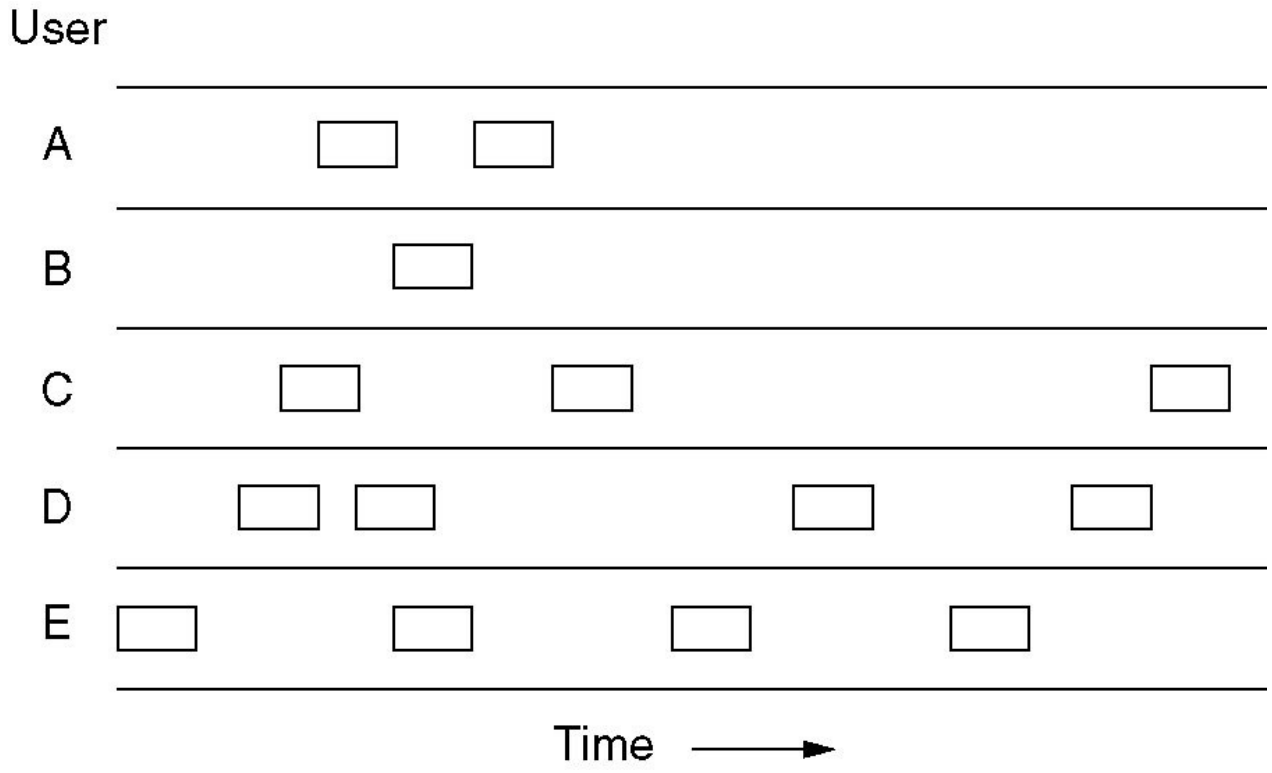
От първите най-известни са Token Ring (разработка на IBM) и FDDI. Те могат да се сравнят с кръгово кръстовище, регулирано със светофари.

Поради сложността им бяха изместени изцяло от състезателните. По-нататък ще се занимаваме с тях.

“Чиста” ALOHA

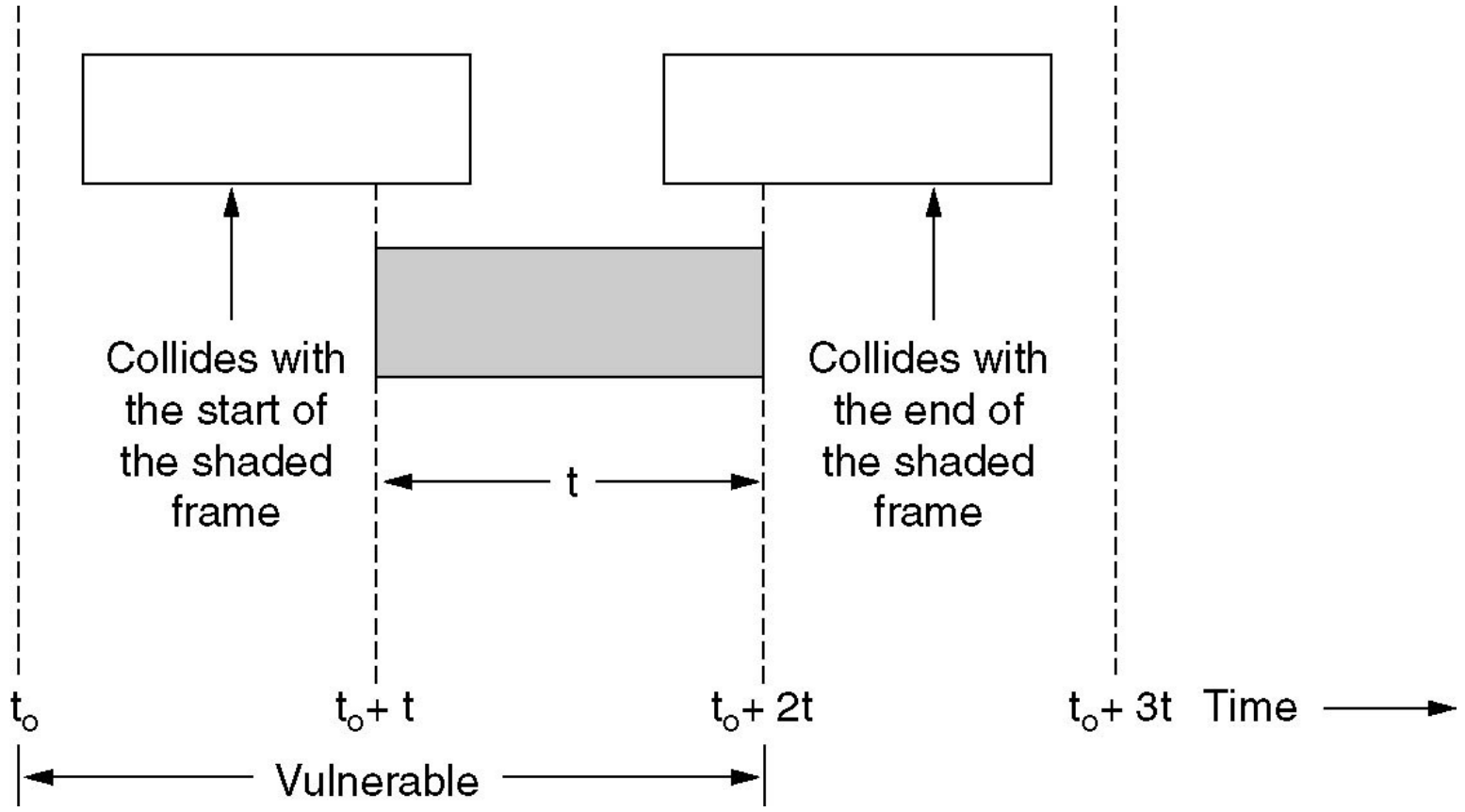
Идва от мрежата в Университета в Хонолулу – Хавайските острови.
Множество радиостанции, разположени на различните острови.
Всяка предава, когато “си поиска”, без да се съобразява с другите.
Aloha си е Multiple Access (MA) и
Съответства на “нерегулируемо кръстовище”

Чиста ALOHA



Кадрите се предават в произволно време.

Чиста ALOHA. Колизии.

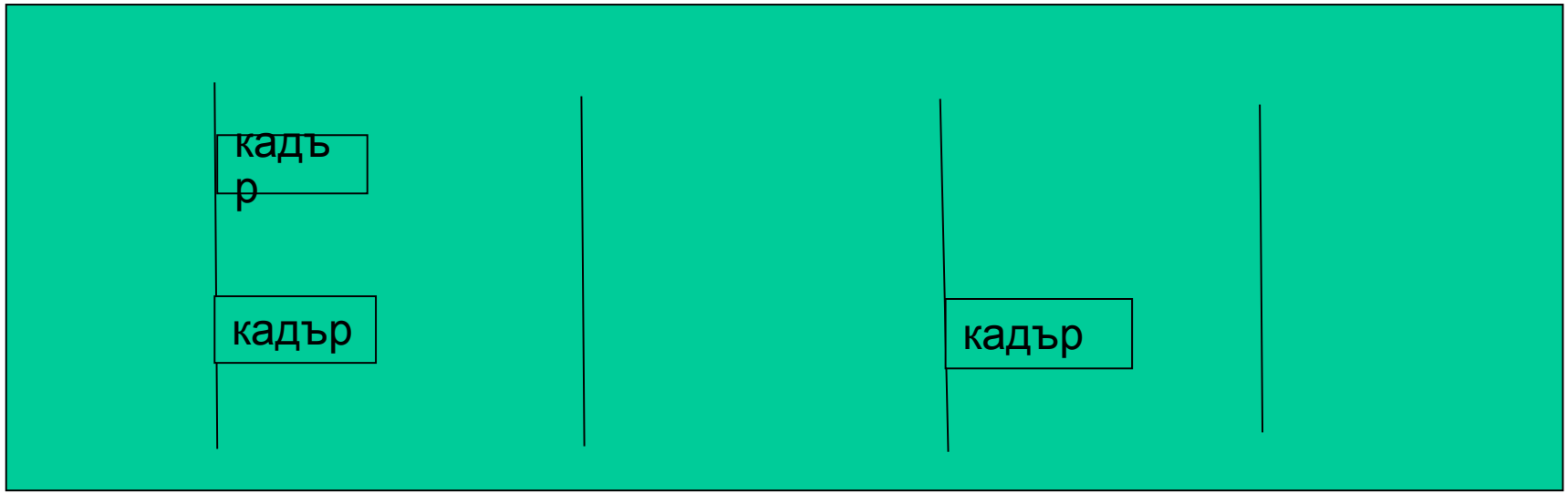


Колизии с началото и края на долния кадър.

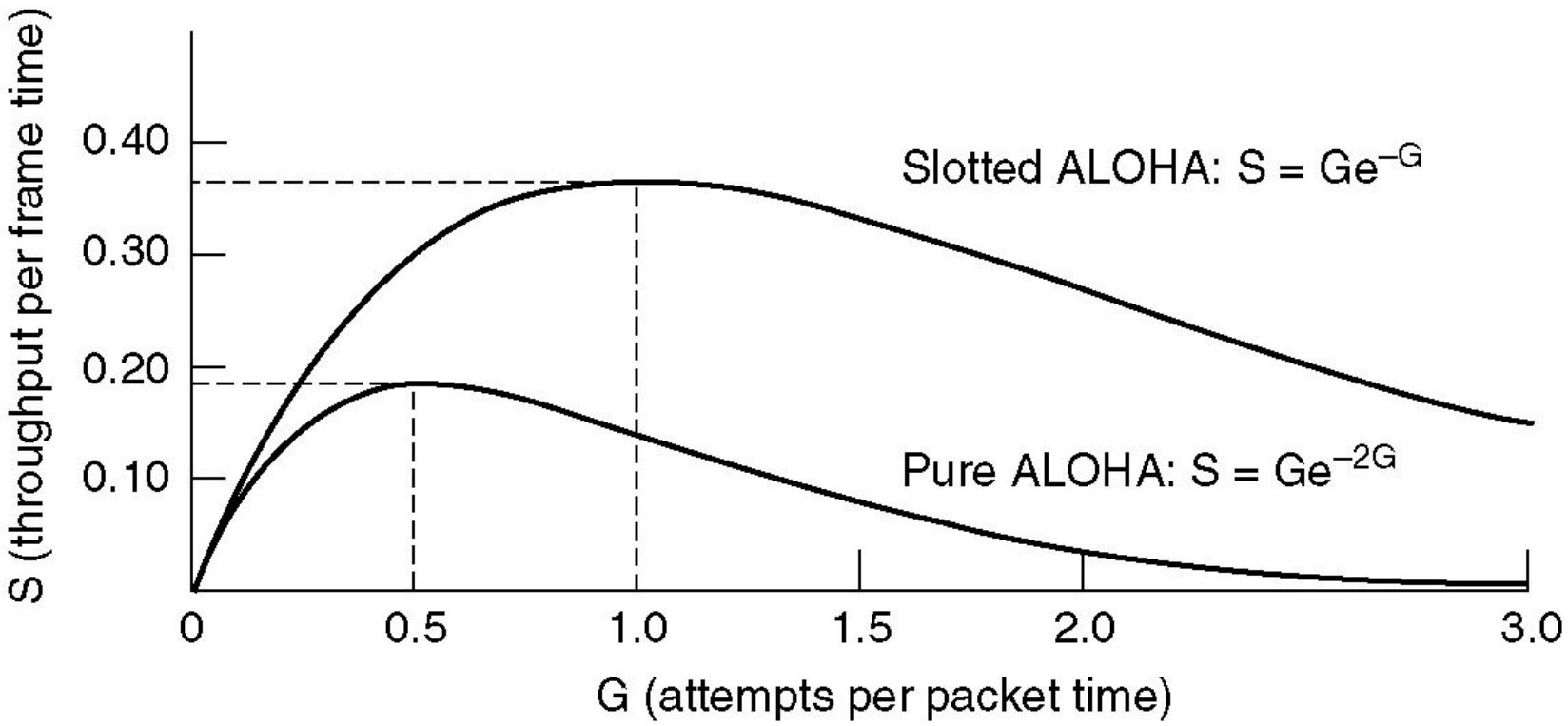
Slotted ALOHA

Предава само в началото на синхронизирани отрязъци от време - “slot times”

Колизиите се ограничават само във времето на предаване на един кадър

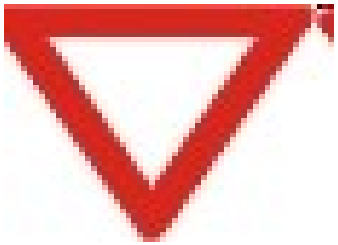


Pure vs. Slotted ALOHA



Пропускателна способност спрямо ниво на трафика

Carrier Sense Multiple Access (CSMA)



Можем да го сравним със знака „Пропусни движещите се по пътя с предимство!“

Протоколите, които прослушват носещата, се наричат **carrier sense multiple access (множествен достъп с откриване на носещата – МДОН)**.

Предложени са от Kleinrock и Tobagi (1975), които са анализирали техни варианти.

Един от тях се нарича **1-persistent CSMA** (1 настойчив).

Когато станция има данни за предаване,прослушва канала.

Ако е зает, чака, докато се освободи. Когато открие свободен канал, предава кадъра. Ако настъпи колизия, изчаква произволен период от време и започва отново.

Протоколът се нарича 1-persistent, защото станцията започва да предава с вероятност 1, ако има свободен канал.

Nonpersistent CSMA

Този протокол не е толкова “лаком”. Станцията прослушва канала, ако никой не предава, започва тя.

Ако каналът е зает, станцията не продължава да прослушва, а изчаква произволен период от време, след което повтаря алгоритъма.

Постига се по-добро оползотворяване на канала от 1-persistent CSMA.

p-persistent CSMA се отнася към канали с времеделене (time slot).

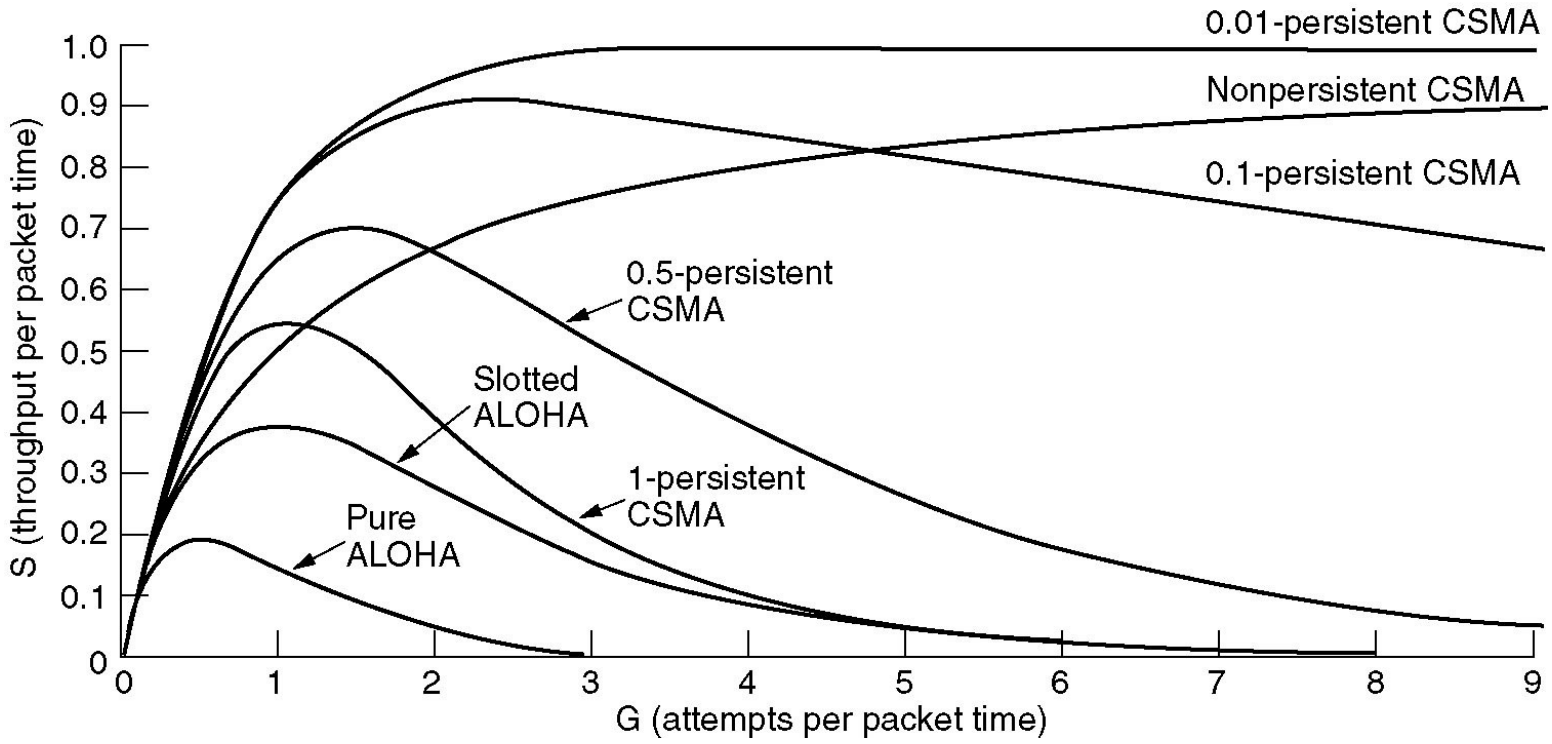
Ако каналът е свободен готовата станция започва да предава с вероятност p. С вероятност $q = 1 - p$ отлага за следващия слот. Ако и той е свободен, или предава, или отлага с вероятност p или q.

Процесът продължава, докато кадърът се предаде или друга станция започне да предава.

В последния случай неуспялата действа като при колизия (изчаква произволно време).

Ако станцията открие зает канал, изчаква следващия слот и прилага горния алгоритъм.

Persistent и Nonpersistent CSMA



Използване на канала спрямо натоварването

CSMA плюс Collision Detection

Друго подобрение е станциите да прекратятват предаването в момента, когато “забележат” колизия.

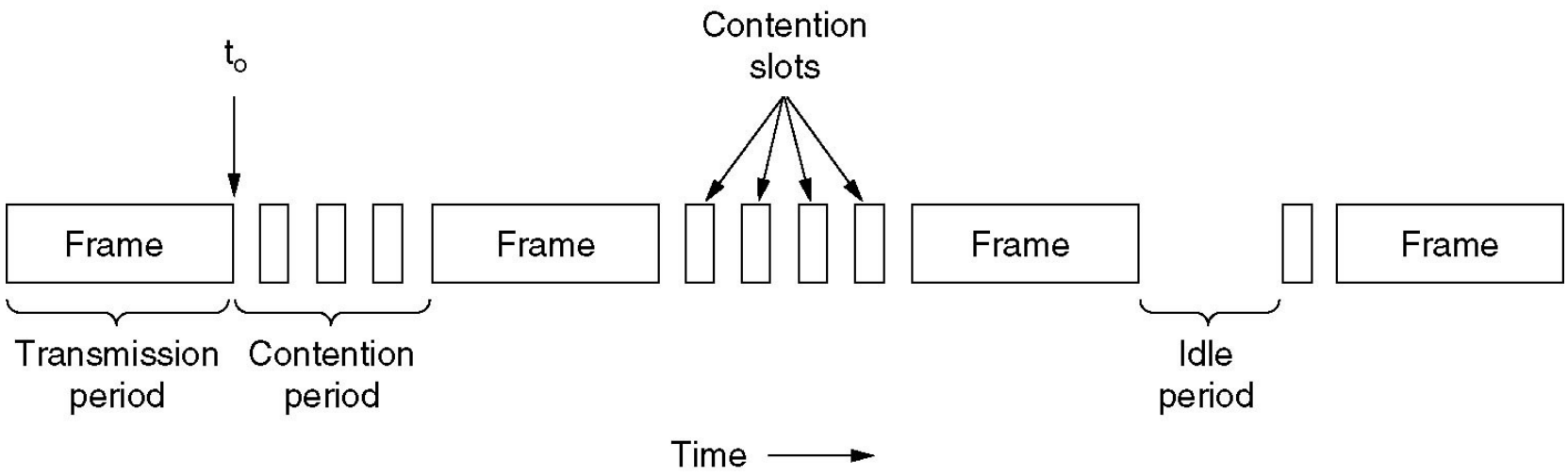
Т.е, ако две станции открият свободен канал и започнат да предават едновременно, едновременно ще разпознаят и колизията.

Те спират предаването веднага щом разпознаят колизията.

Този протокол е **CSMA/CD** (**CSMA with Collision Detection**) и се използва в LAN Ethernet.

CSMA/CD като състезателна процедура използва концептуалния модел, показан по-долу:

CSMA/CD



CSMA/CD е в едно от трите състояния: състезание, предаване свободно.

CSMA/CD

Малко по-подробно за състезателната процедура. Нека двете станции започнат да предават в момент $t0$. Колко време им трябва да разберат за колизията?

Минималното време е времето, за което сигналът пропътува от една станция към друга.

Създава се илюзията, че е достатъчно станцията да не е чула колизия за времето за пропътуване на сигнала през целия кабел, за да си мисли, че е "хванала" канала.

Но знаят ли другите станции за това. НЕ.

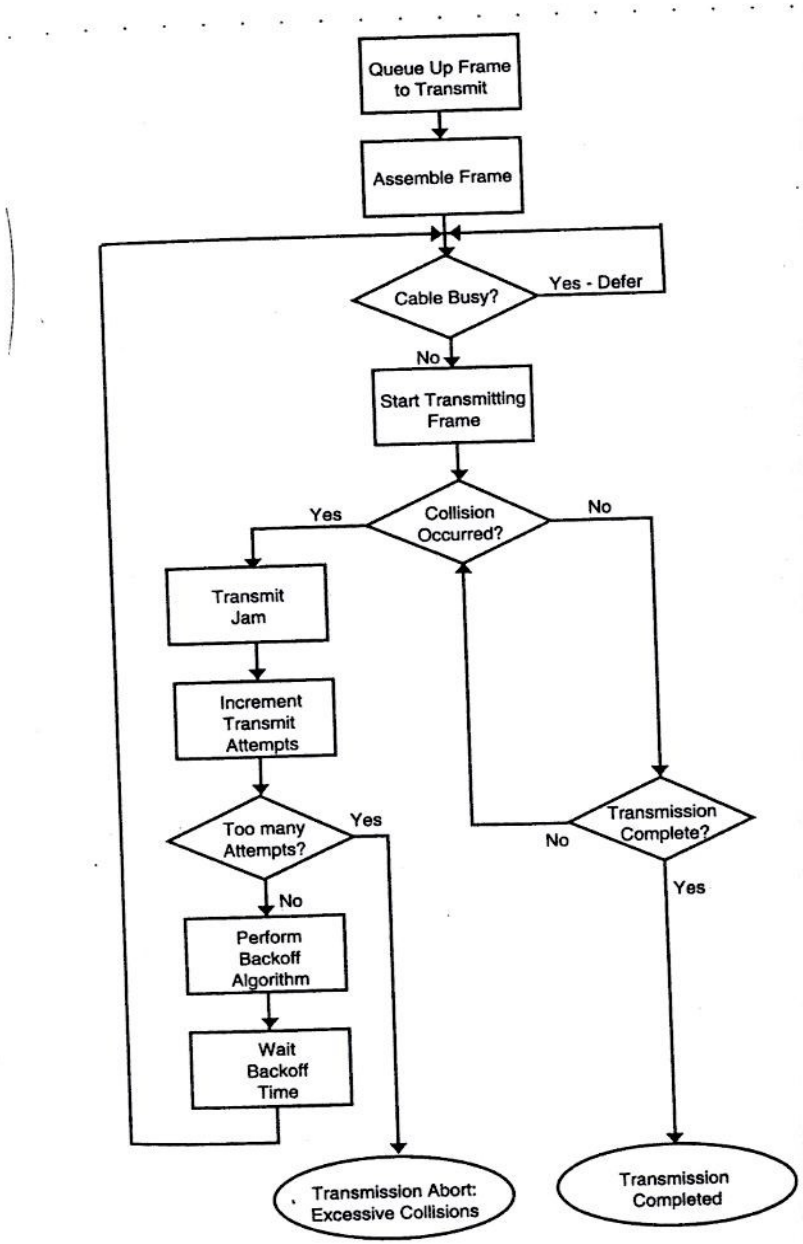
Най-лошият случай. Нека времето за пътуване между двете най-отдалечени станции е τ. В $t0$ една станция започва да предава. В $\tau - \varepsilon$ най-отдалечената също започва да предава. Тя веднага разпознава колизията и спира, но

Шумът от колизията достига до оригиналната за време $2\tau - \varepsilon$.

Т.е, в най-лошият случай една станция не може да е сигурна, че е "хванала" канала, докато не е предавала за 2τ, без да е чула колизия.

В 1-km коаксиален кабел $\tau \approx 4.8$ μsec.

CSMA/CD в най-разпространената LAN



Robert M. "Bob" Metcalfe

Откривателят на Ethernet

Най-разпространената LAN Ethernet

Описана в стандарта IEEE (Institute of Electrical and Electronic Engineers) 802.3, издаден през 70-те години.

Един персонален компютър се свързва в Ethernet мрежа с помощта на NIC (Network Interface Card) - това е каналната станция, която осъществява обмена по Ethernet канала.

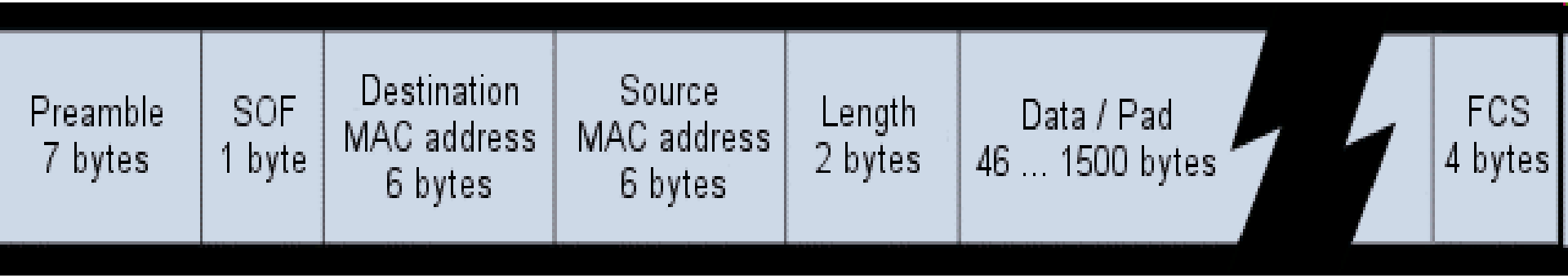
Преди да изпрати кадър, каналната станция проверява състоянието на канала. Ако той е свободен, тя веднага започва предаване.

Ако каналът не е свободен (т.е. предава друга станция), то станцията изчаква неговото освобождаване.

След като започне предаването, каналната станция продължава да подслушва канала. Ако се открие изкривяване на предавания сигнал, това означава, че по същото време е започнала да предава друга станция и е настъпила **колизия**.

В този случай двете станции спират предаването и всяка от тях изчаква случаен интервал от време преди да предава отново.

802.3 Кадър



Preamble = 56 бита 0-и и 1-ци.

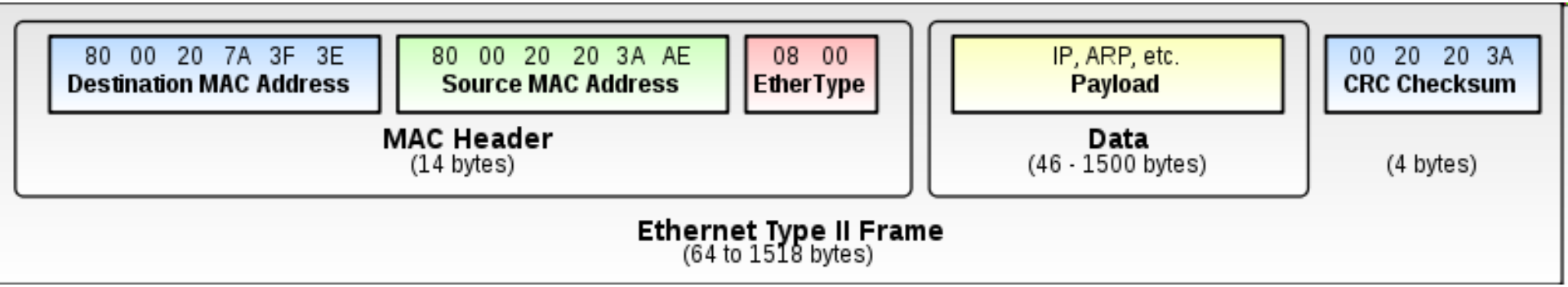
SOF = Start of frame: "10101011"

Data / Pad = ако няма достатъчно данни (payload), полето за данни се допълва, за да имаме минимален размер на кадъра

FCS = Frame check sequence – CRC

Днес се използва **Ethernet II frame**, **DIX** frame (DEC, Intel и Xerox); директно от Internet Protocol.

Ethernet II кадър



Destination address съдържа адресът на получателя на кадъра

Source address - адресът на изпращача на кадъра.

Най-младшият бит на най-старшия байт на адреса на получателя е 0 за нормален адрес и 1 за групов адрес. При групов адрес, кадърът е предназначен за група станции (multicast). Адрес на получател, състоящ се само от 1 означава, че кадърът е предназначен за всички станции (broadcast).

Полето *EtherType*: 0x0800 кадърът носи IPv4 дейтаграма; 0x0806 - ARP, 0x8100 - IEEE 802.1Q и 0x86DD - IPv6.

Данните се съдържат в полето *Data* и максималната им дължина е 1500 байта. Освен максимална дължина на кадъра има и минимална дължина на кадъра.

Формат на кадрите в Ethernet

Когато една предаваща станция разбере за конфликт, тя веднага спира предаването, като орязва настоящия кадър.

За да може да се прави разлика между валидни и орязани кадри, дължината на кадъра трябва да е поне толкова голяма, че да може предаването да не е завършило, преди станцията да разбере за конфликта.

В стандарта 802.3 минималната дължина на кадъра е 64 байта.

Ако данните са по-малко от 46 байта, то се използва полето *Pad* за запълване на кадъра до 64 байта.

Полето *Checksum* е контролна сума, която се използва за откриване на грешки при предаването.

Maximum Transmission Unit (MTU)

В компютърните мрежи MTU в протокол на даден слой е размера (в байтове) на най-големия протоколен блок за данни (PDU), който може да понесе дадения слой. Т.е максималния payload.

По-голям MTU означава по-високаа ефективност:

- един пакет носи повече потребителски данни;
- по-малко служебна информация (overhead).

Но, по-големите пакети окупират за по-голям период бавните линии. Например, 1500-байтов Ethernet кадър “захваща” за цяла секунда 14.4k модемна линия. Затова се налага фрагментиране

Ефективност и нетна скорост

$$\text{Efficiency} = \frac{\text{Payload size}}{\text{Frame size}}$$

Максимална ефективност се постига с максимален payload:

$$\frac{1500}{1538} = 97.53\%$$

за untagged ethernet кадри и е $\frac{1500}{1542} = 97.28\%$

за 802.1Q VLAN tagging.

Net bit rate: $\text{Net bit rate} = \text{Efficiency} \times \text{Wire bit rate}$

Maximum net bit rate for 100BASE-TX Ethernet without 802.1Q is 97.53 Mbit/s.

MTU. Jumbo Frames.

jumbo frames са Ethernet кадри с дължина по-голяма от 1500 байта payload (MTU). Приема се, че jumbo frames носят до 9000 bytes.

Много, не и всички, Gigabit Ethernet суичове и карти поддържат jumbo frames, но всички Fast Ethernet поддържат само стандартните 1500 байта.

Дължина на Ethernet кадъра от 1518 байта е избрана въз основа на оценка нанадеждността и скоростта на канала.

От друга страна, ако увеличим размера, по-големи обеми от данни ще се предадат с по-малко усилия:

- по-малко CPU цикли;
- по-малко прекъсвания;
- CPU се съсредоточава върху потребителските данни.

Jumbo Frames.Super Jumbo Frames

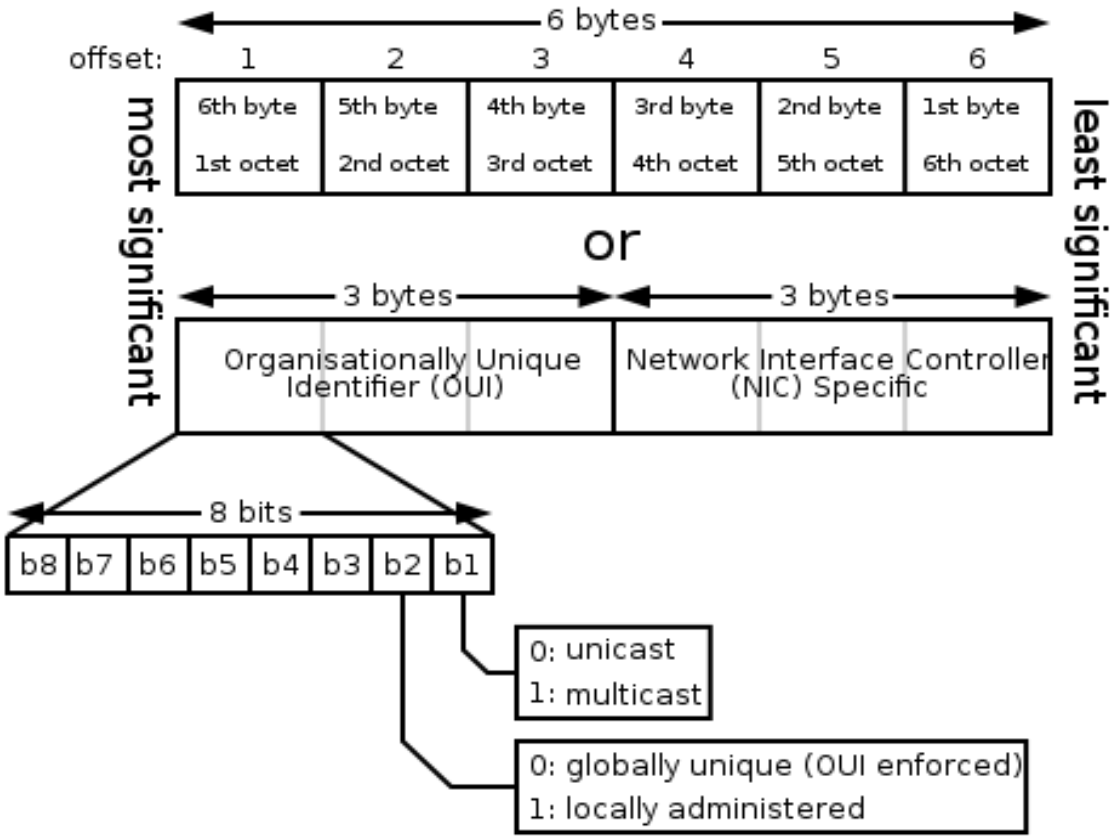
9000 байта като предпочитан размер на jumbo frames е резултат от споразумение между Joint Engineering Team of Internet2 и правителствените мрежи в САЩ.

Super jumbo frames (**SJF**s) са кадри с дължина над 9000 байта.

С растежа на скоростта на линията пропорционално би трябвало да расте и payload. Това обаче зависи от възможностите на логическите схеми, обработващи пакетите.

Колкото и да са трудни преговорите в тази насока, възможно е да се достигне дължина от 64000 байта.

Формат на MAC адрес



Формат на MAC адрес

Media Access Control адресът (MAC адрес), Ethernet Hardware Address (EHA) или хардуерен адрес, адрес на адаптера или физически адрес е квазиуникален идентификатор, присвоен на мрежов адаптер или NIC от производителя. Втози случай MAC адресът съдържа закодиран идентификатора на производителя.

IEEE дефинира три схеми за формулиране на MAC адрес: MAC-48, EUI-48 и EUI-64. Търговски марки на IEEE са "EUI-48" и "EUI-64“ (EUI - Extended Unique Identifier). Разликата между EUI-48 и MAC-48 е чисто семантична (но не и синтактическа): MAC-48 се използва за мрежов хардуер, а EUI-48 идентифицира други устройства и софтуер.

Макар че е смятан за перманентен и глобално уникален, днес е възможно да се смени MAC адреса (т.е не е “прогорен”) - MAC spoofing.

Оригиналният IEEE 802 MAC произлиза от Xerox Ethernet. Съдържа 2^{48} или 281,474,976,710,656 възможни адреси.

Адресите могат да бъдат “универсално администрирани" или “локално администрирани“.

Формат на MAC адрес

Универсално администриран е присвоен от производителя, още “прогорен” - "burned-in addresses" (BIA). Първите три октета показват организацията, издала идентификатора - Organizationally Unique Identifier (OUI).

Следващите три октета (MAC-48 и EUI-48) или пет (EUI-64) се дават от самата организация.

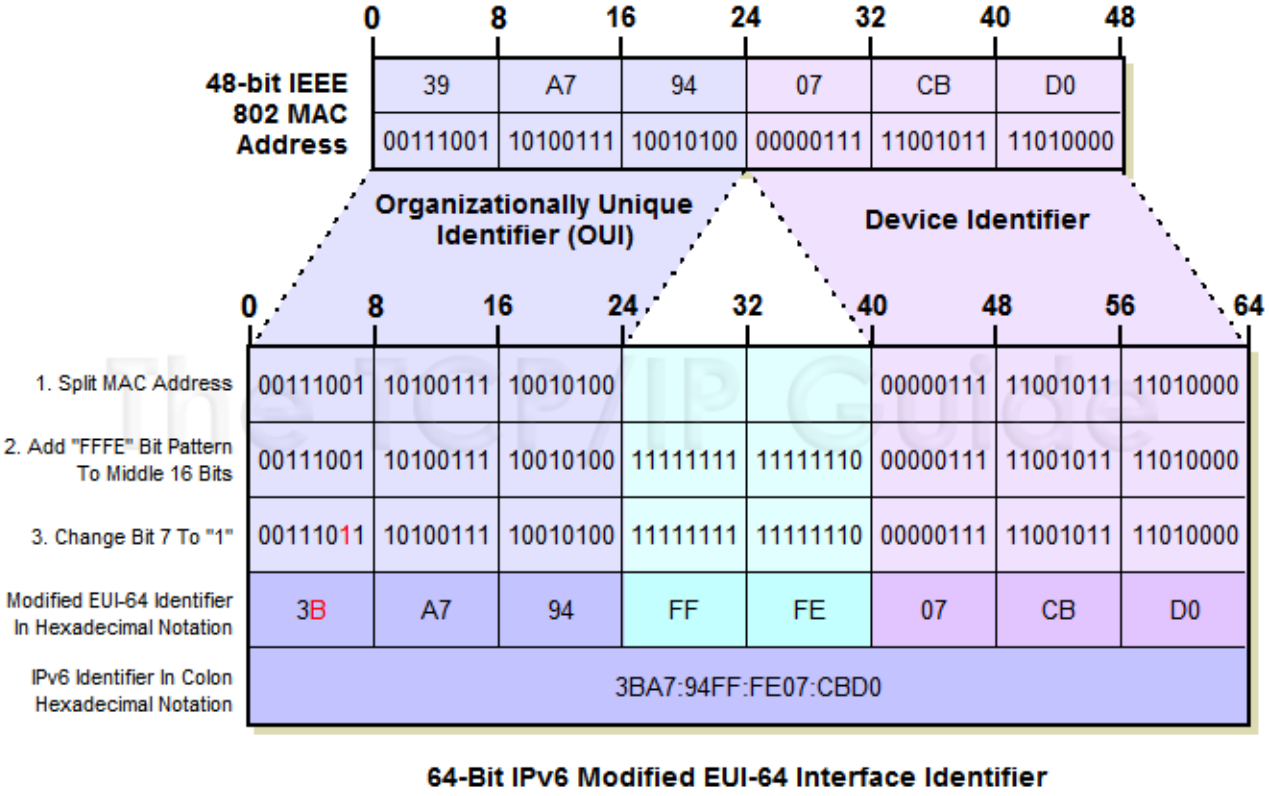
Според IEEE MAC-48 пространството няма да се изчерпи до 2100 г.

Локално администриран се присвоява от мрежовия администратор, отменяйки “прогорения”. Те нямат OUI.

Разпознават се по bit 2 (2^1) в най-старшия октет на MAC-а. Ако е 0, адресът е универсален. Ако е 1, адресът е локален. Т.е е 0 на всички OUI-та.

Ако най-младшият бит – bit 1 (2^0) е 0, кадърът е предназначен за конкретна NIC - unicast. Ако е 1, кадърът трябва да достигне няколко (група) NIC-ве. Нарича се групов - multicast.

EUI-64 формат



64-Bit IPv6 Modified EUI-64 Interface Identifier

EUI-64 формат

EUI-64 се използват:

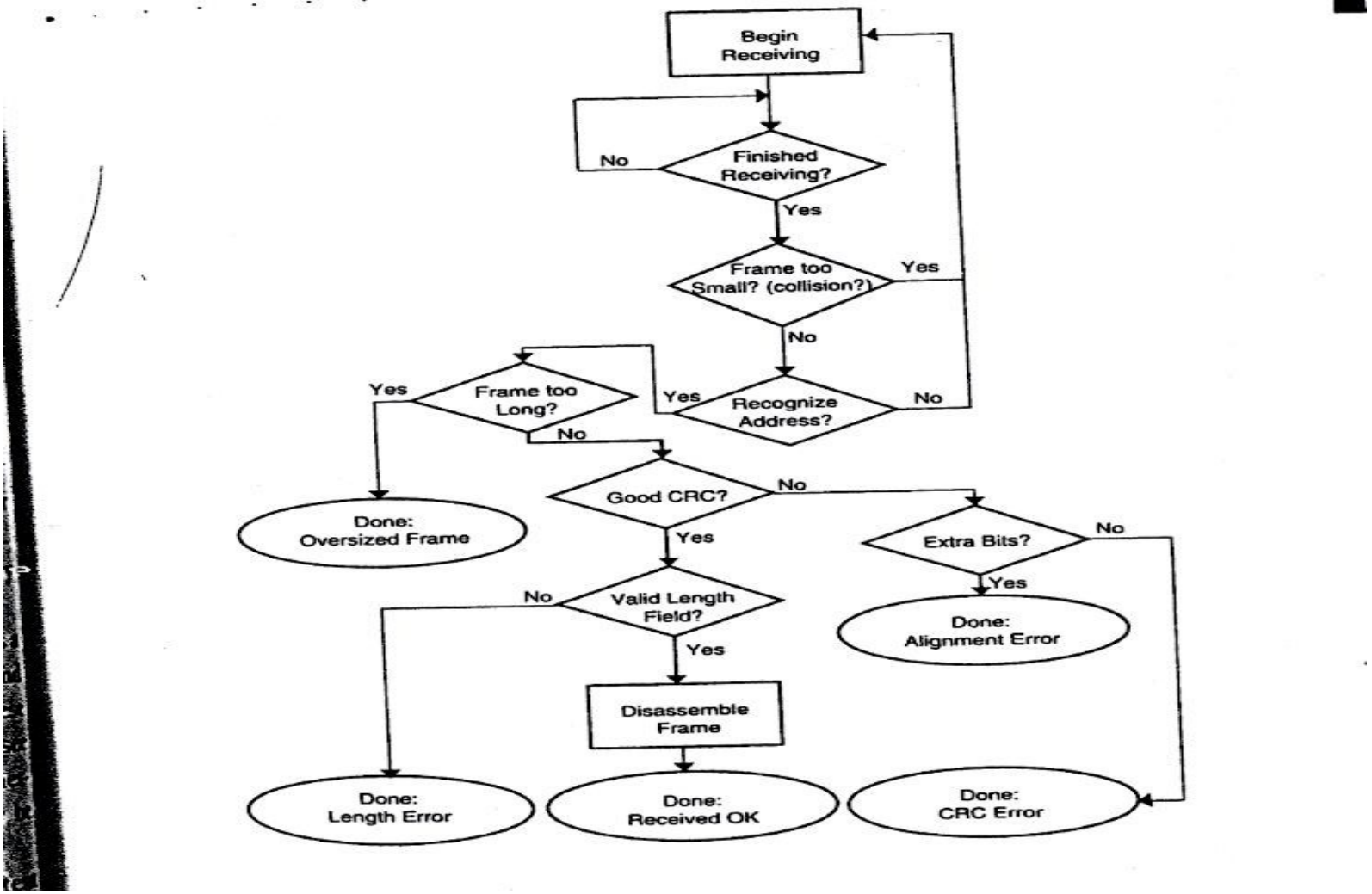
* FireWire
* IPv6 (младшите 64 бита в unicast мрежов адрес или link-local адрес)

Преобразуване на 48-бит MAC адрес в IPv6 модифициран EUI-64 идентификатор:

1. Вземаме 24-бит OUI частта и я поставяме в най-левите 24 бита на interface ID. А 24-бит локална част слагаме в най-десните 24 бита на interface ID.
2. В оставащите в средата 16 бита на interface ID поставяме стойността “11111111 11111110” (“FFFE” hex).
3. Така адресът ни е в EUI-64 формат. Променяме “universal/local” бита (бит 7 отляво) от 0 на 1.

И получаваме модифицирания EUI-64 interface ID.

Приемане на Ethernet кадри



Манчестърски код

При физическото предаване кадрите се кодират в **манчестерски код** (Manchester Encoding).

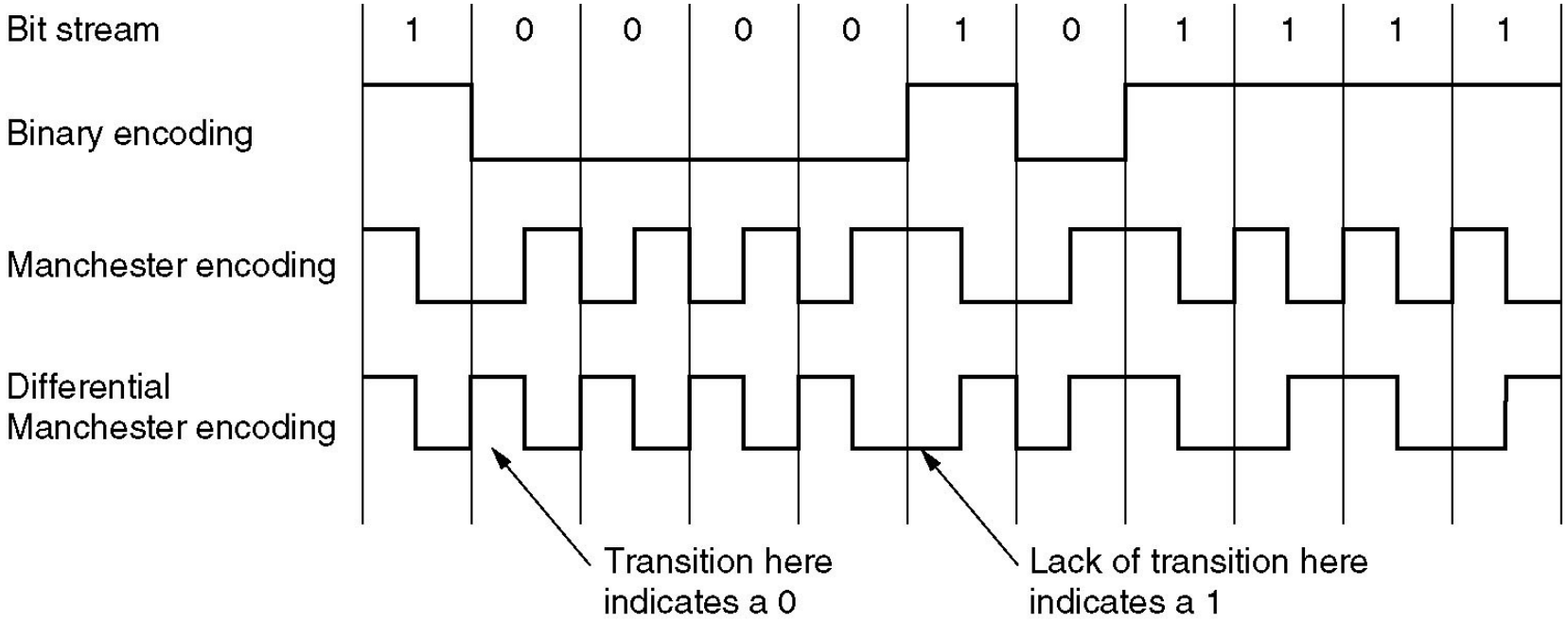
Периодът за предаване на един бит се разделя на две равни части.

Логическа 1 се кодира високо напрежение в първия период и ниско напрежение във втория период.

Логическа 0 се кодира с ниско напрежение в първия период и високо напрежение във втория период.

Преходът в средата на периода служи за синхронизация. По този начин няма нужда от паралелен синхронизиращ сигнал.

Манчестърски код



(a) Двоичен код, (b) Манчестерско кодиране,
(c) Диференциално Манчестерско кодиране.

Ethernet кабели и топологии

Name	Cable	Max. seg.	Nodes/seg.	Advantages
10Base5	Thick coax	500 m	100	Original cable; now obsolete
10Base2	Thin coax	185 m	30	No hub needed
10Base-T	Twisted pair	100 m	1024	Cheapest system
10Base-F	Fiber optics	2000 m	1024	Best between buildings

100BASE-TX: Използва 2 чифта по Category 5 (IEEE 802.3u).

100BASE-FX: 100 Mbit/s Ethernet по FO.

1000BASE-T: 1 Gbit/s over Category 5e copper cabling (802.3ab).

1000BASE-SX: 1 Gbit/s по MM FO.

1000BASE-LX: 1 Gbit/s по SM FO (големи разстояния).

10GBASE-LX4: WDM - 240 m i 300 m по MM FO. 10 km по SM FO (802.3ae).

10GBASE-LR и 10GBASE-ER: 10 km и 40 km по SM FO.

10GBASE-SW, 10GBASE-LW и 10GBASE-EW. Върху WAN PHY

10GBASE-T: меден кабел Категория 6a (802.3an)

Ethernet кабели и топологии

Ethernet кабели и топологии

В началото в Ethernet се използва коаксиален кабел и скоростта на предаването е достигала 10 Mb/s.

По-нататък се въвежда използването на **хъбове** (hub). При окабеляване 100Base-T4 каналните станции се свързват към хъба чрез четири усукани двойки UTP Category 3

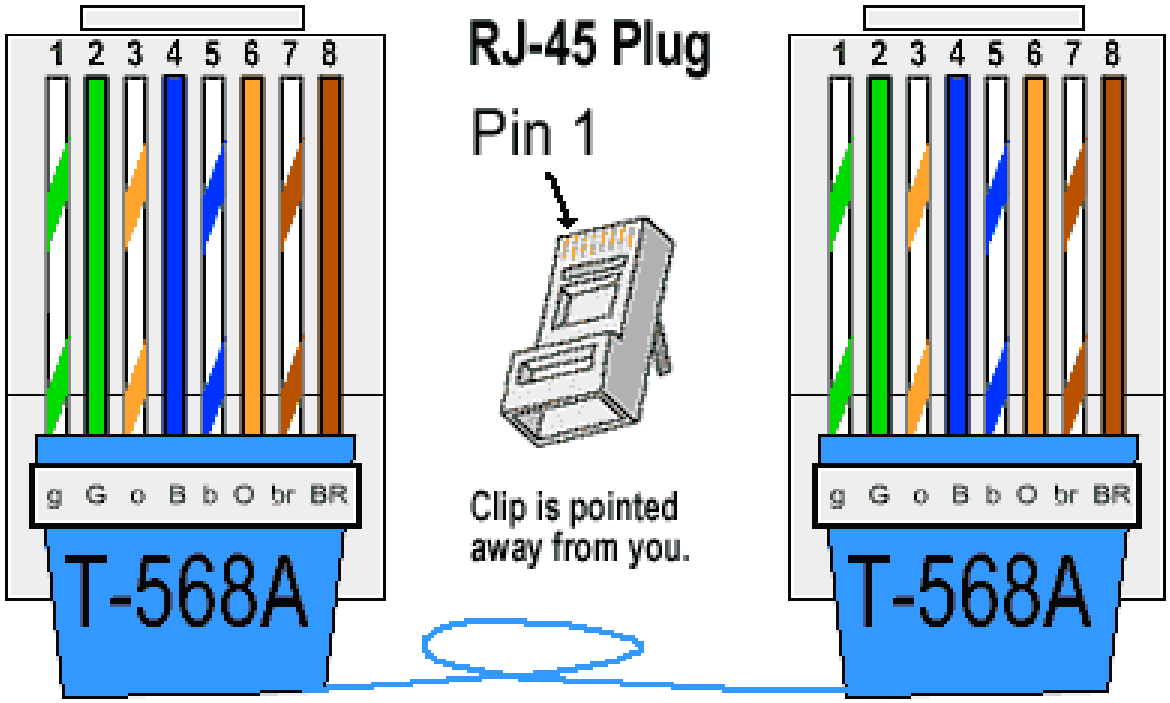
100Base-TX чрез две усукани двойки (UTP Category 5). По една от усуканите двойки се предава към хъба, а по другата се приема от него (при 100Base-T4 останалите две усукани двойки се превключват по посока на предаването). Скоростта на предаване достига 100 Mb/s

Ethernet кабели и топологии

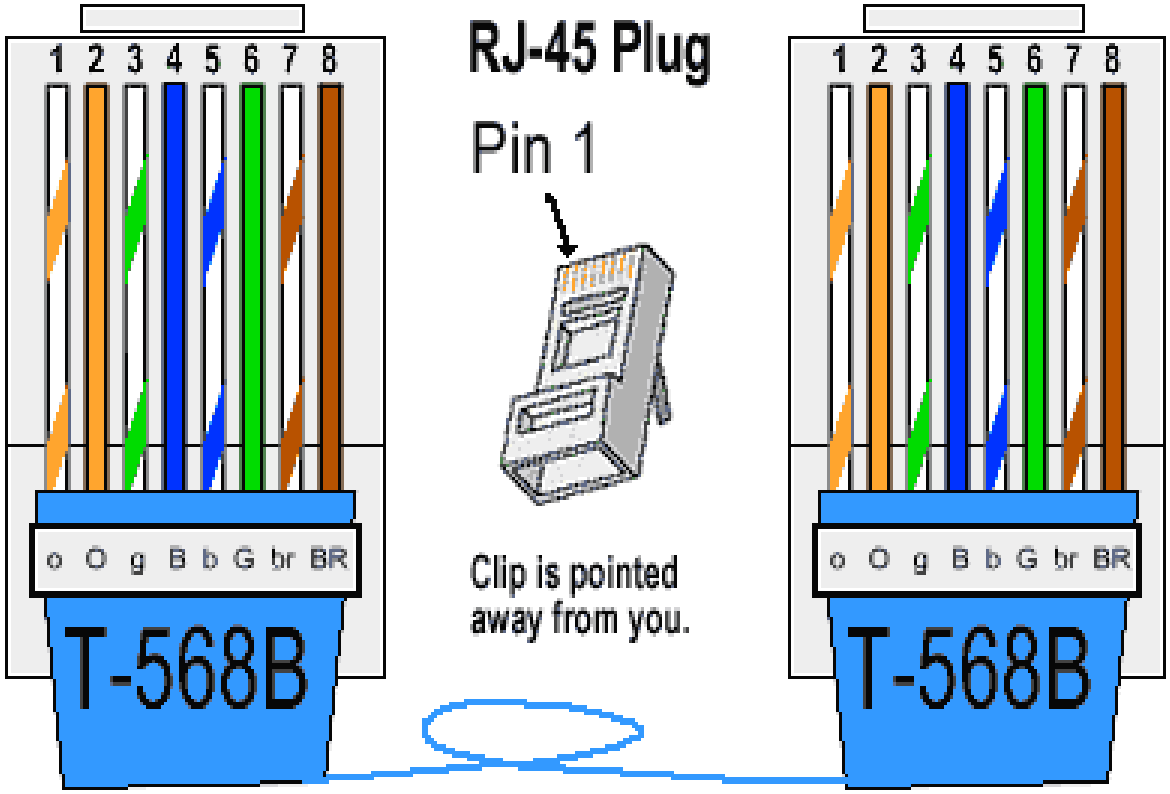
Станциите се свързват към хъба в **прав кабел**, т.е. предаващата двойка на всяка станция съответства на предаващата двойка на хъба и съответно приемащата двойка на всяка станция съответства на приемащата двойка на хъба.

При свързване на два хъба чрез усукана двойка, обаче, се използва **кръстосан** (**cross**) **кабел**, т.е. предаващата двойка на единия хъб се свързва с приемащата двойка на другия хъб и обратно.

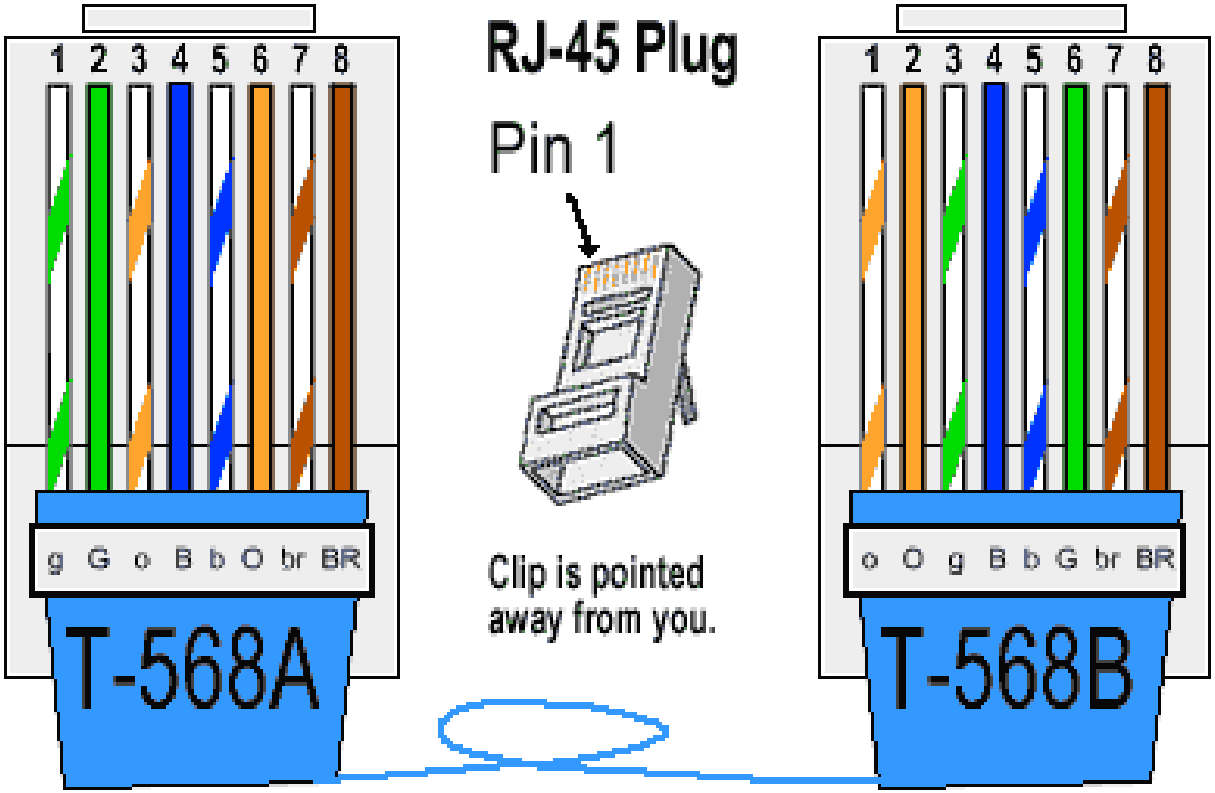
Прав (Straight-Through) кабел



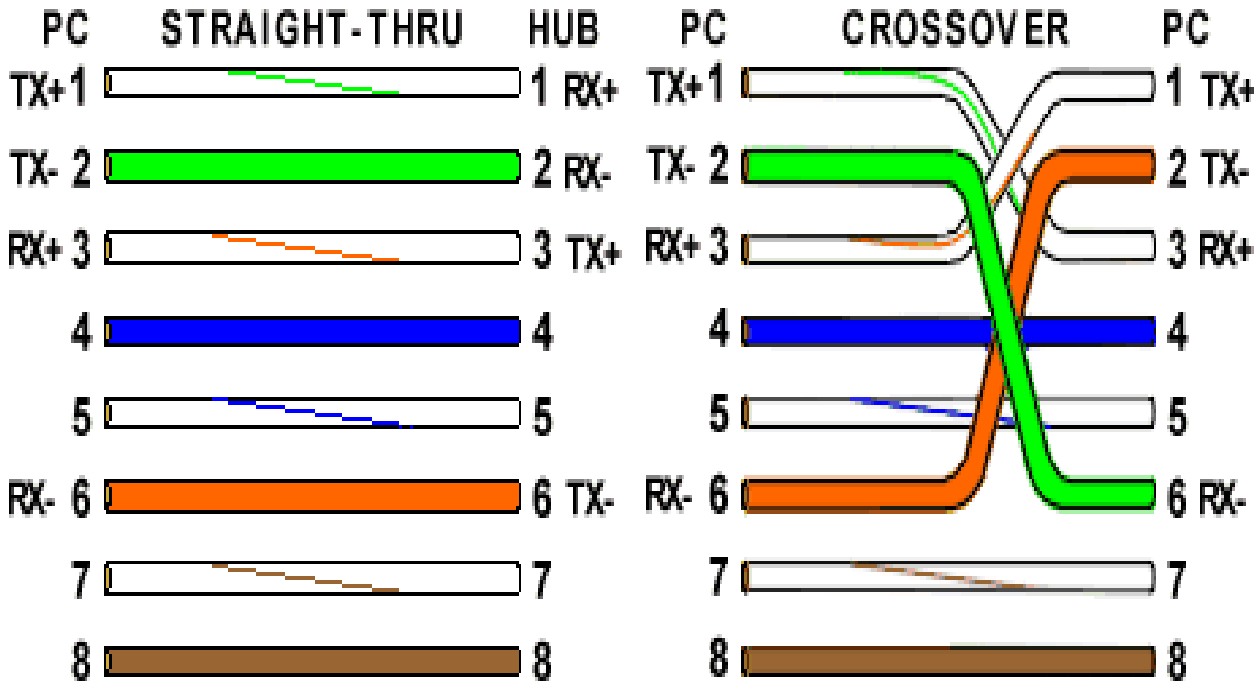
Прав (Straight-Through) кабел



Кръстосан (Crossover) ксбел



Straight vs. Cross (теория)



Хъб и повторител

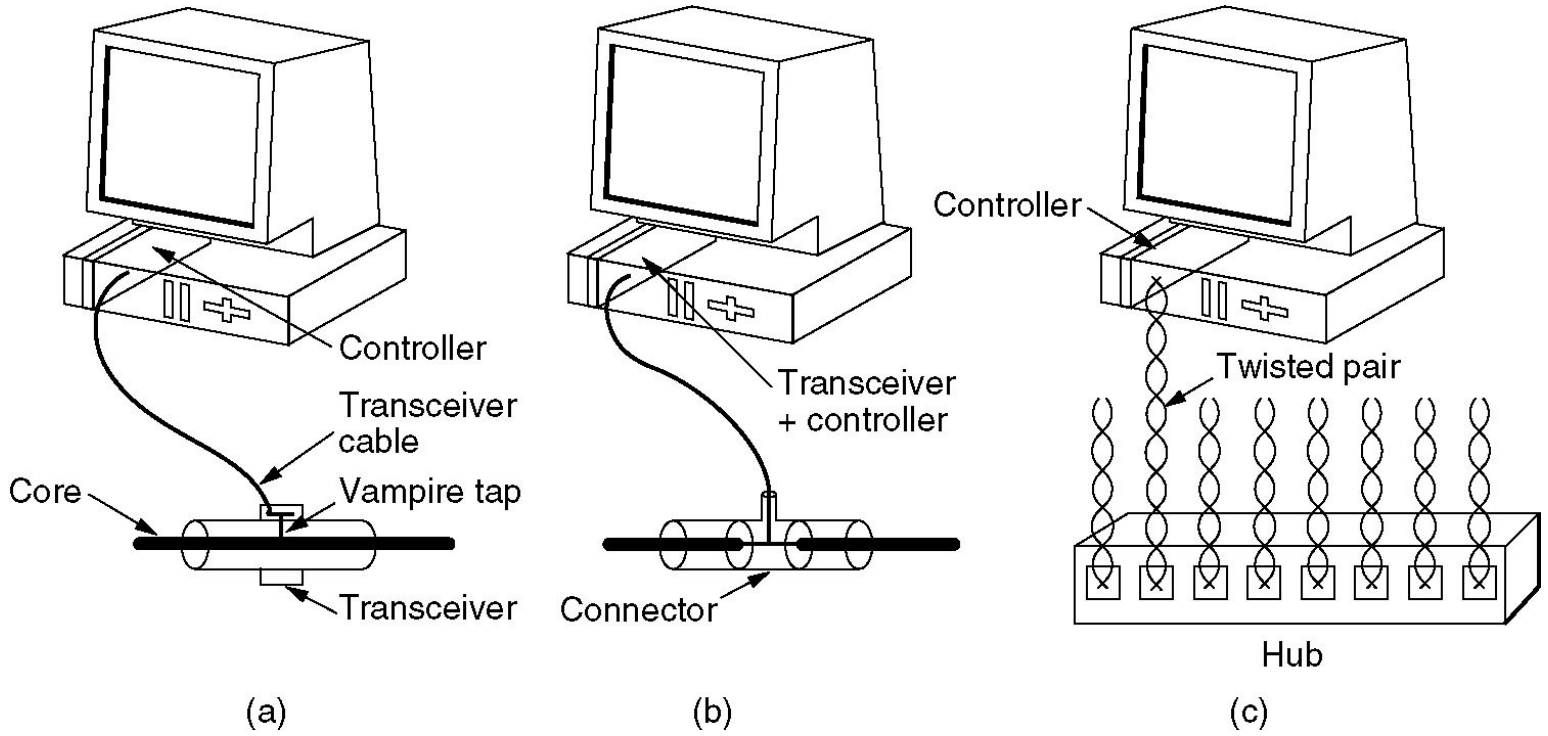
Ако хъбът получи кадър по някоя линия, той изпраща този кадър по всички останали линии. Хъбът не знае адресите на каналните станции.

Хъбът е пример за устройство, чрез което се препредават кадри от един кабел към друг. Той работи на физическо ниво.

Друго подобно устройство на физическо ниво е **повторителят** (repeater).

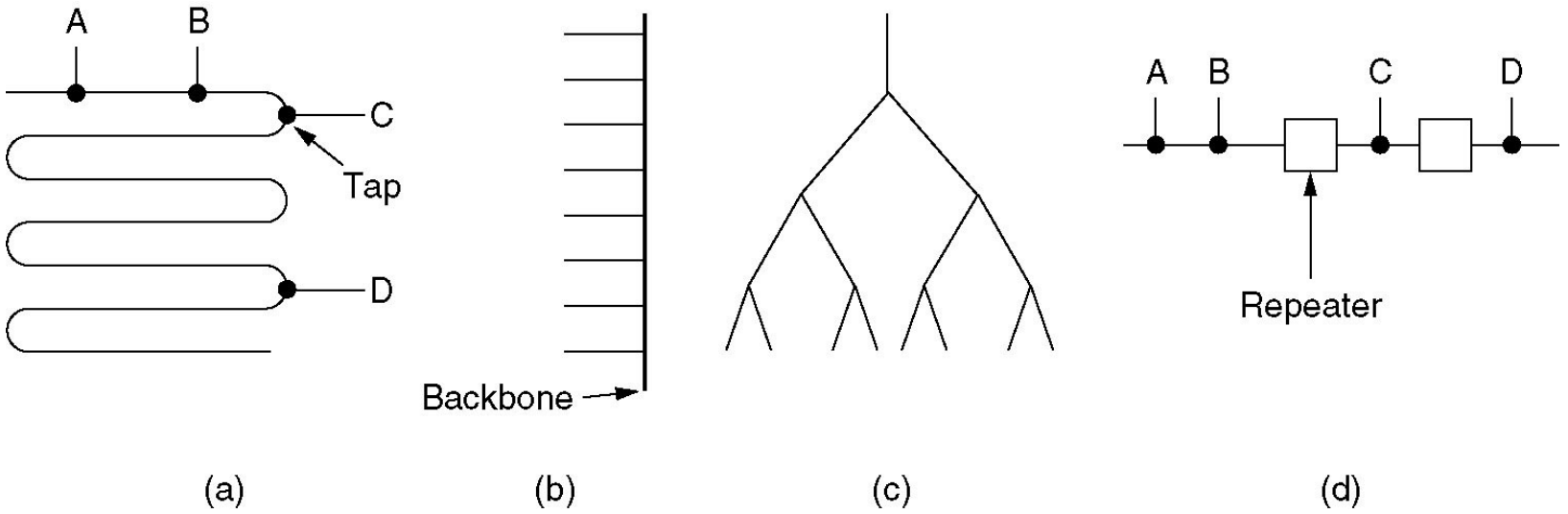
Той приема сигнал на единия си порт, усилва го и предава сигналът на другия си порт. По този начин може да се увеличи максималната дължина на кабела в една локална мрежа

Коаксиални кабели и хъб



Три вида Ethernet cabling.
(a) 10Base5, (b) 10Base2, (c) 10Base-T.

Топологии на Ethernet окабеляване



(a) Шина, (b) вертикално, (c) Дървоводна, (d) Сегментирана.

Bridge и switch

Мостът (bridge) работи на канално ниво и служи за свързване на две локални мрежи. За разлика от повторителите и хъбовете, мостът анализира получените кадри.

Той прочита адреса на получателя и по него определя към коя изходна линия да изпрати кадъра (за целта се поддържа специална таблица).

Мостът предава кадъра само към определената от него изходна линия, а не по всички изходни линии.

Подобно устройство е **превключвателят** (switch) – многопортов мост. Той също прочита адресите на постъпилите в него кадри.

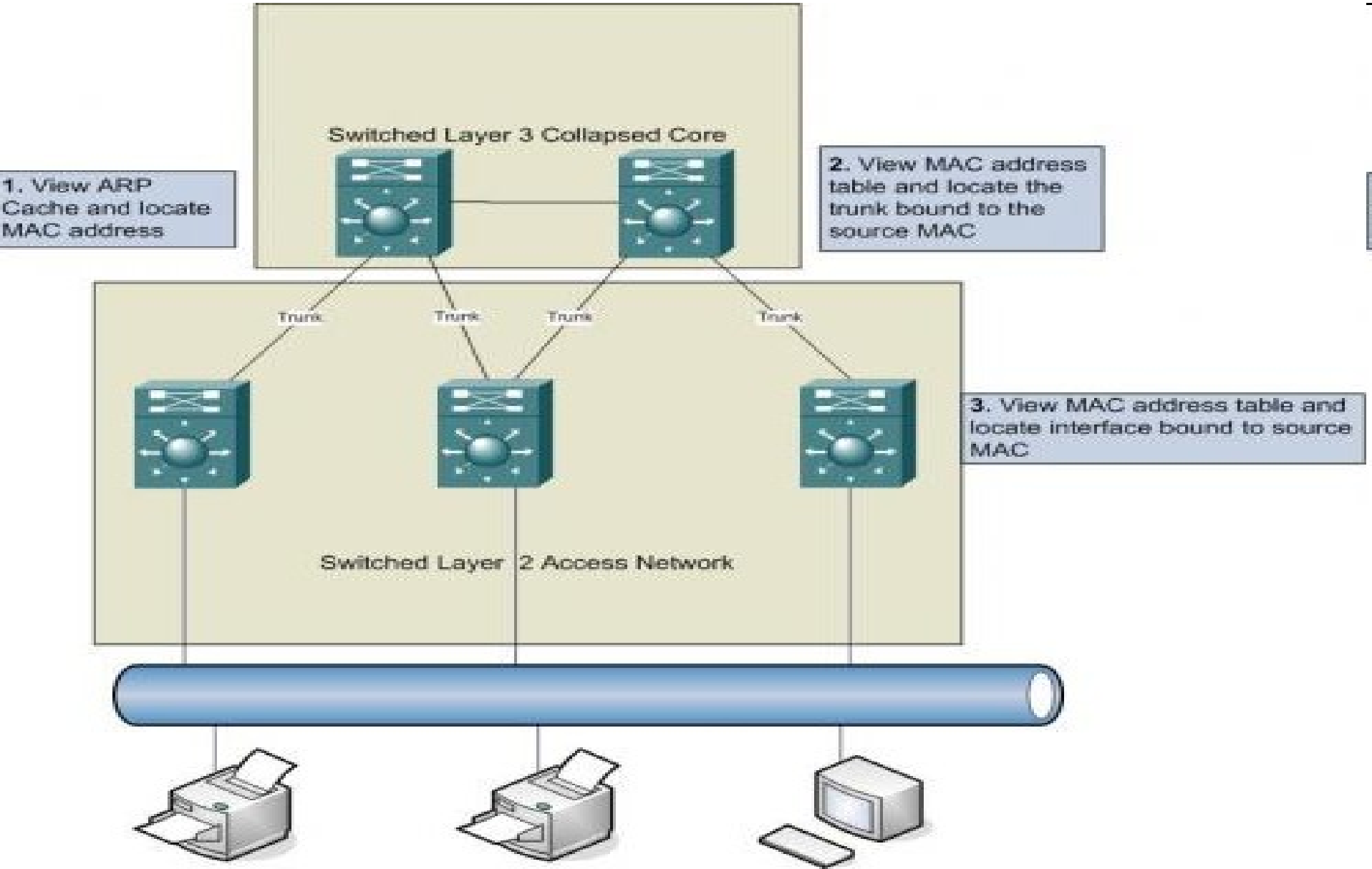
Bridge и switch

Всяка линия (порт) е самостоятелна и представлява отделна област – домейн, на колизиите. Това се нарича още микросегментиране.

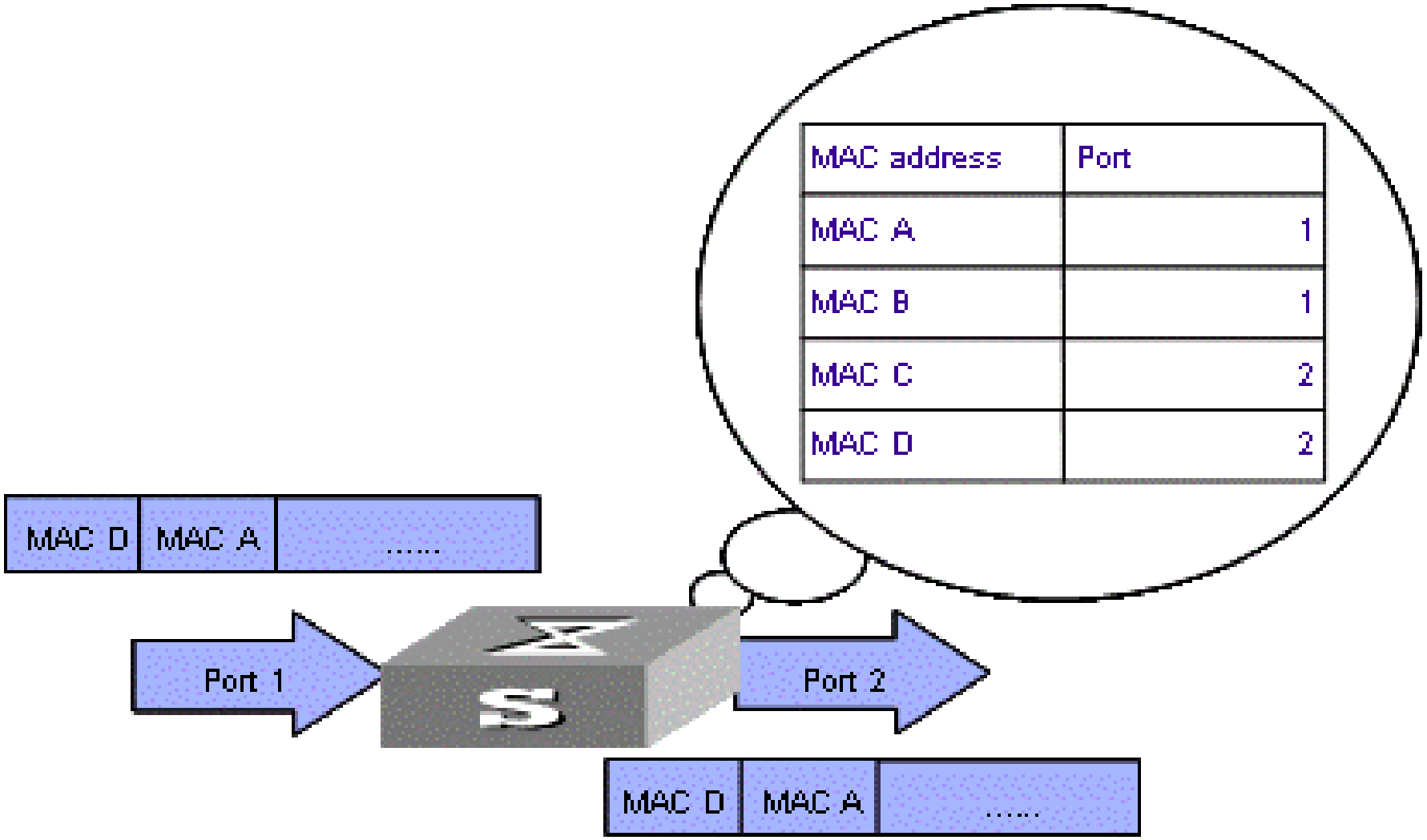
При превключване между сегменти кадри не могат да бъдат изгубени поради колизии.

За целта превключвателя трябва да има достатъчно буферно пространство за да може да се препращат кадрите.

Switched Ethernet



MAC Address Table



Три режима на превключване

С пълно буфериране (**store and forward**). В буферната памет се записва целия кадър и чак след това се превключва към изходния порт. Внася се закъснение и изисква повече памет.

Cut-through – Суичът прочита адреса на получателя при получаване на кадъра. Започва прехвърлянето към изходящия порт, преди да получи пълния кадър. Така се намалява закъснението. Имаме две форми на cut-through:

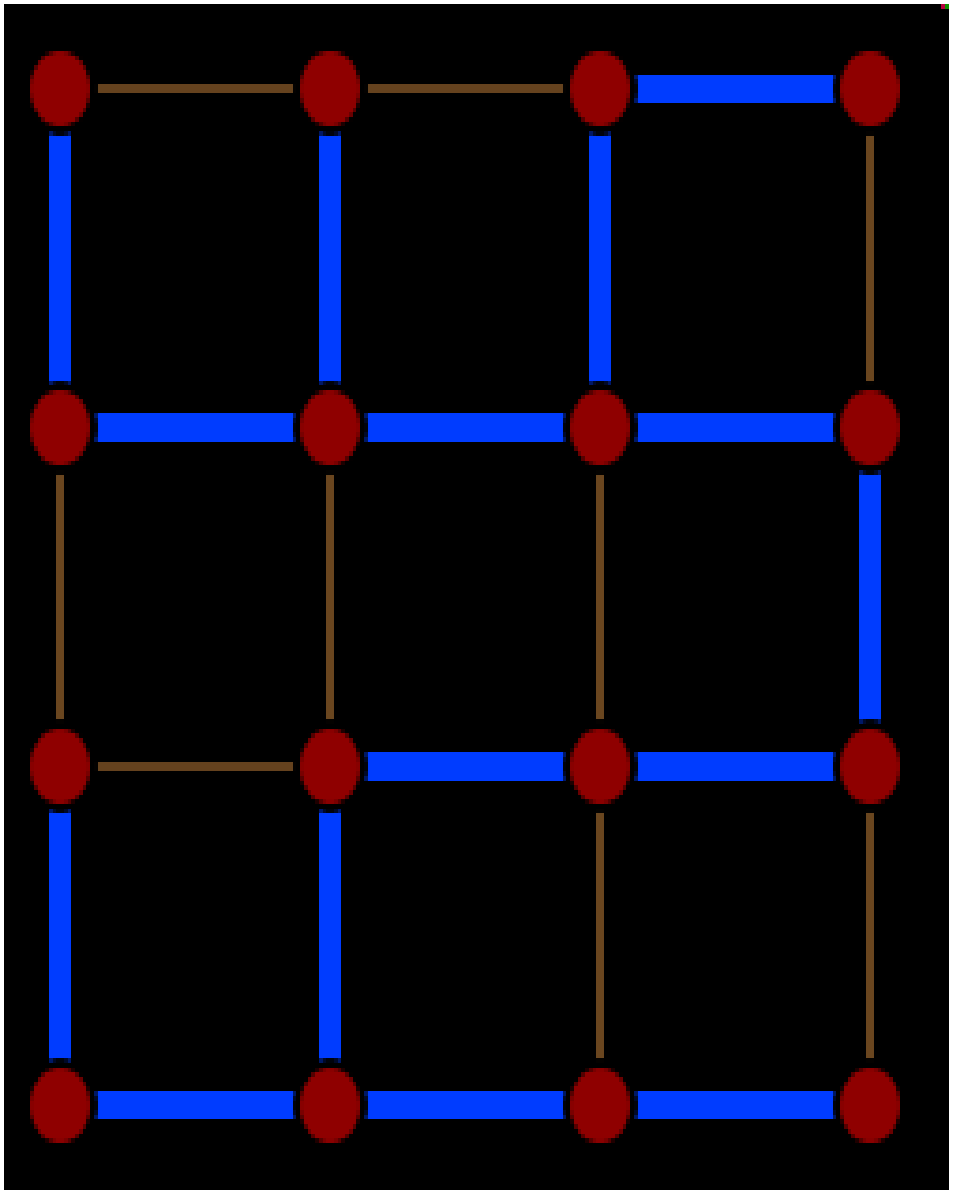
Fast-forward – С най-ниско закъснение, веднага превключва кадъра след приемане на адреса на получателя. Има проблеми с откриването на грешки.

Fragment-free - Филтрира кадри (фрагменти), претърпели колизии, най-често срещаните грешки. Обикновено това са кадри с дължина, по-малка от 64 байта. Т.е прочита първите 64 бита, за да определи дали това не е колизионен фрагмент, преди да започне превключването.

Spanning Tree (математика)

spanning tree (разперено дърво) на граф *G* е сбор от клони на *G,* които формират дърво, разперващо се *spanning* от всеки връх.

Т.е всеки връх е в дървото, но няма зацикляне (no loops).



Spanning Tree Protocol (STP)

Spanning Tree Protocol (STP) е протокол на 2 слой по модела на OSI, който гарантира топология без зацикляне в Switched LAN. Базира се на алгоритъма на Radia Perlman, който е работил за Digital Equipment Corporation.

Позволява да се включват резервни пътища, които автоматично да се активират при авария в основните без опасност от зацикляне.

Зациклянето в тези мрежи е опасно заради липсата на механизъм TTL, както ще видим в IP протокола на 3 слой.

STP се дефинира в стандарта IEEE 802.1D.

STP - стойности

Скорост (Data rate)	(STP Cost – 802.1D-1998)	(802.1t-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	2000

Spanning Tree - алгоритъм

Spanning Tree алгоритъмът изчислява път без зацикляне.

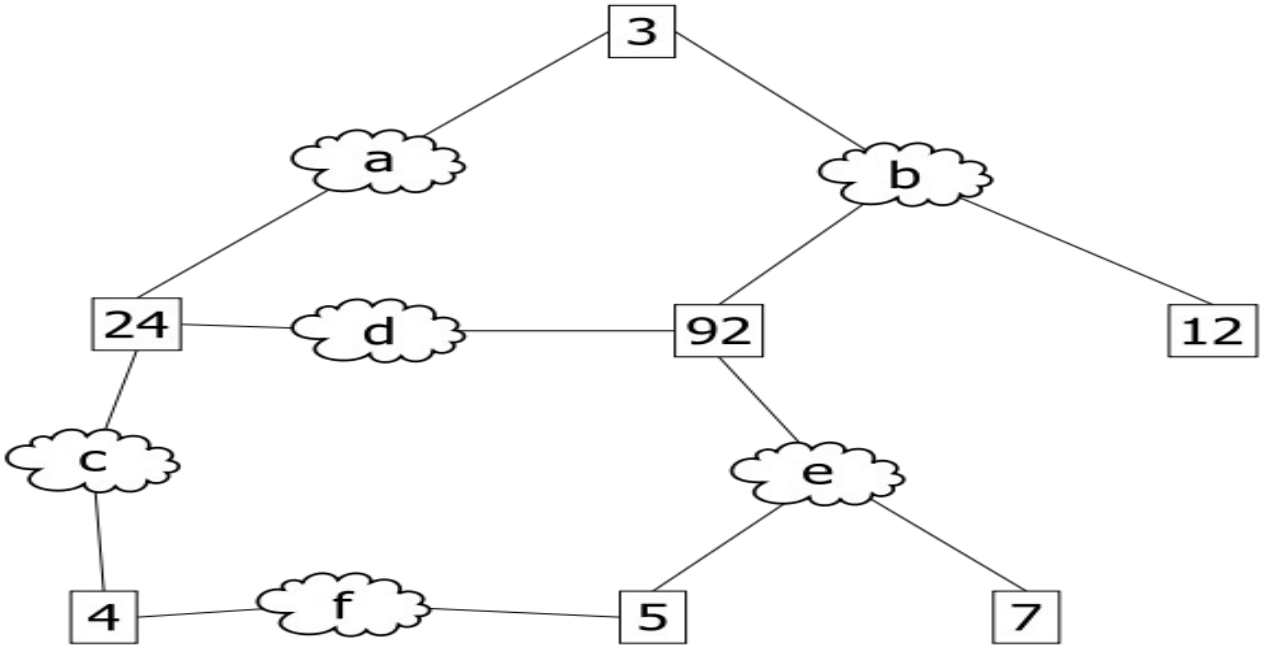
Първоначално всички портове са блокирани. Отнема около 50 s, докато започнат да превключват.

Стъпка 1 : Избор на Root Bridge – с най-нисък приоритет или най-ниско bridge ID (MAC адрес)

Стъпка 2 : Избор на Root Ports – От алтернативните пътища се избират тези с най-малка стойност до Root Bridge.

Стъпка 3 : Избор на Designated Ports – Порт, който праща и получава трафик от Root Bridge – с най-ниска стойност до Root Bridge.

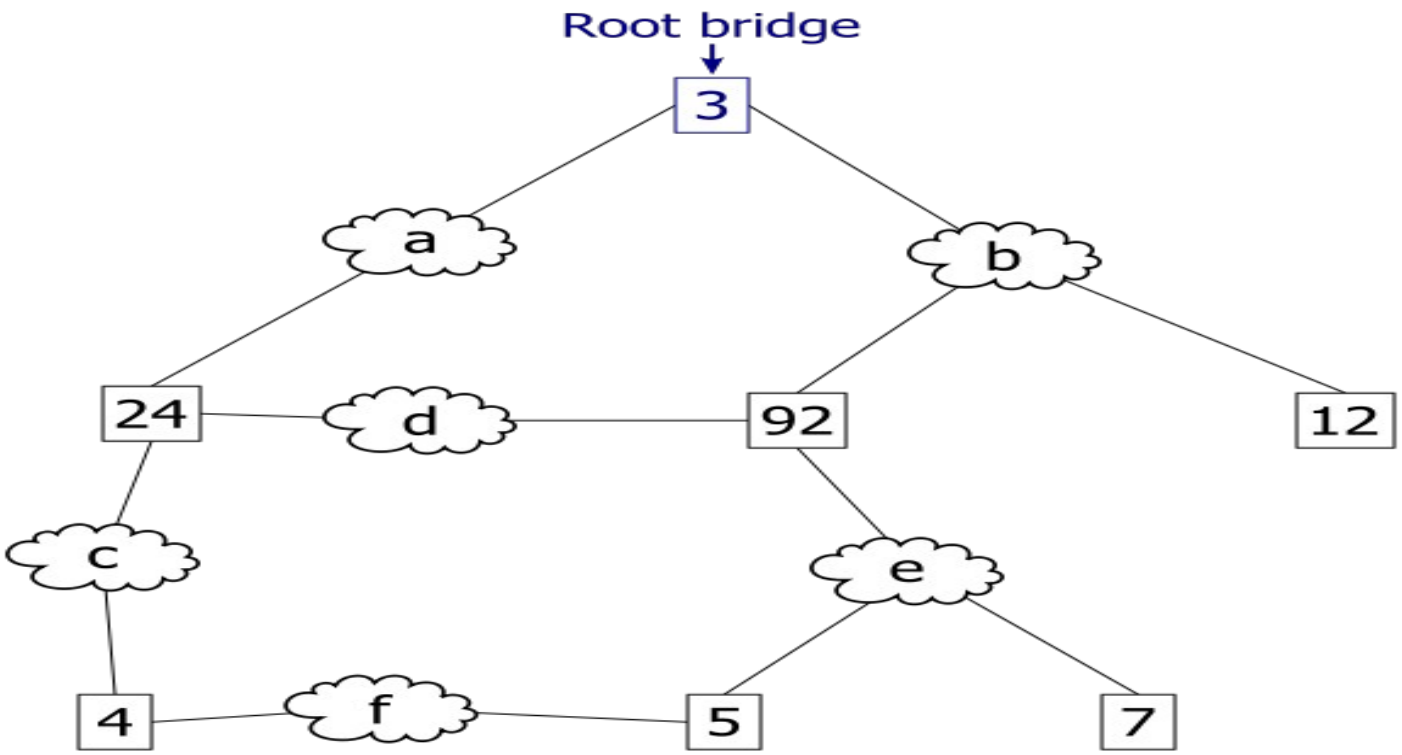
Пример



Номерираните кутийки -**bridge ID**.

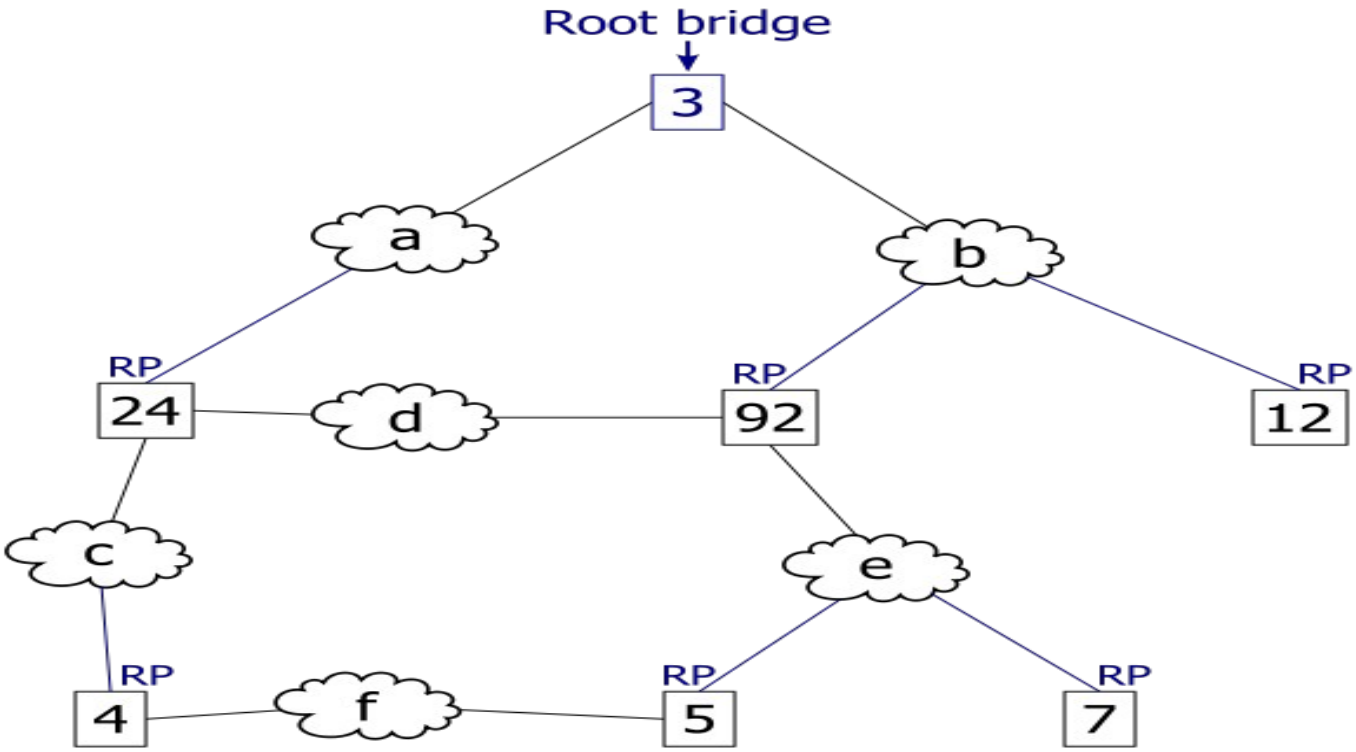
Номерираните облаци – мрежови сегменти.

Избор на root bridge



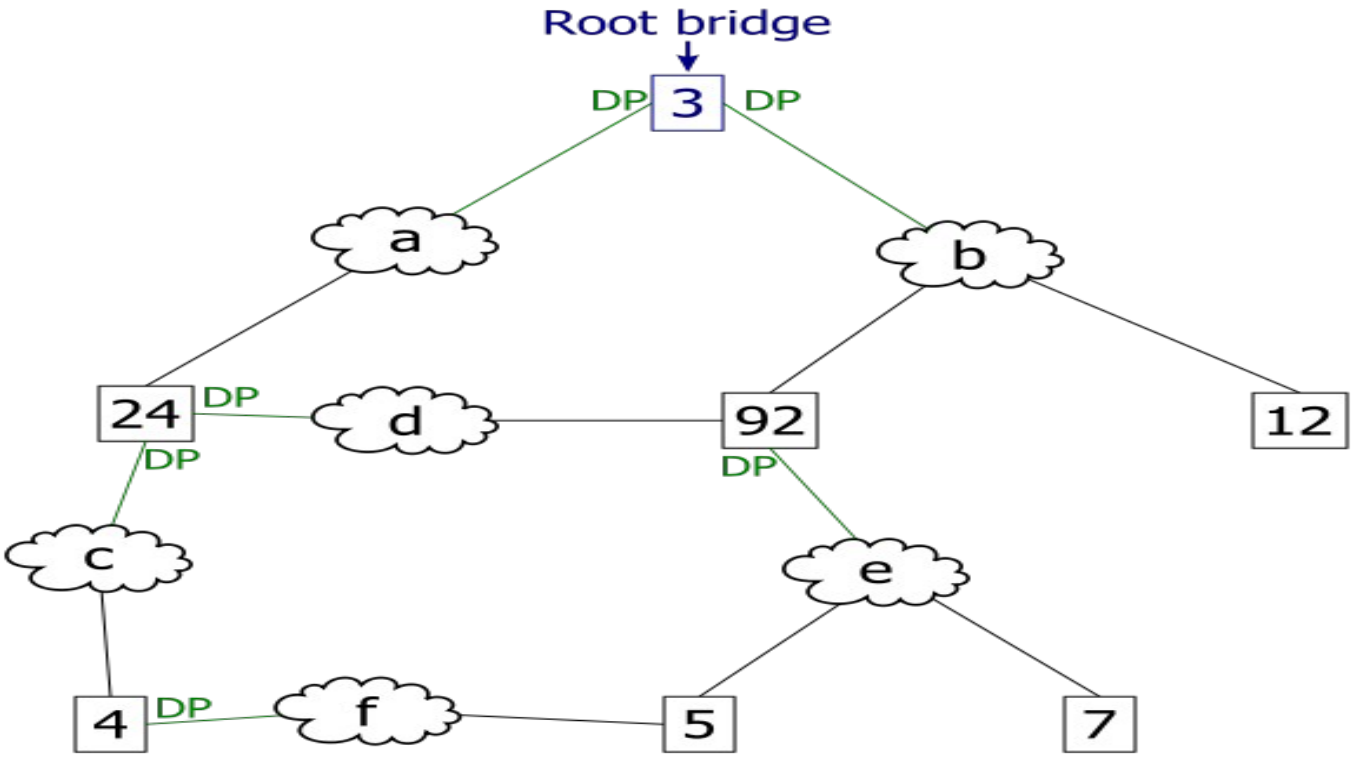
Най-малкият bridge ID е **3**

Избор на root port



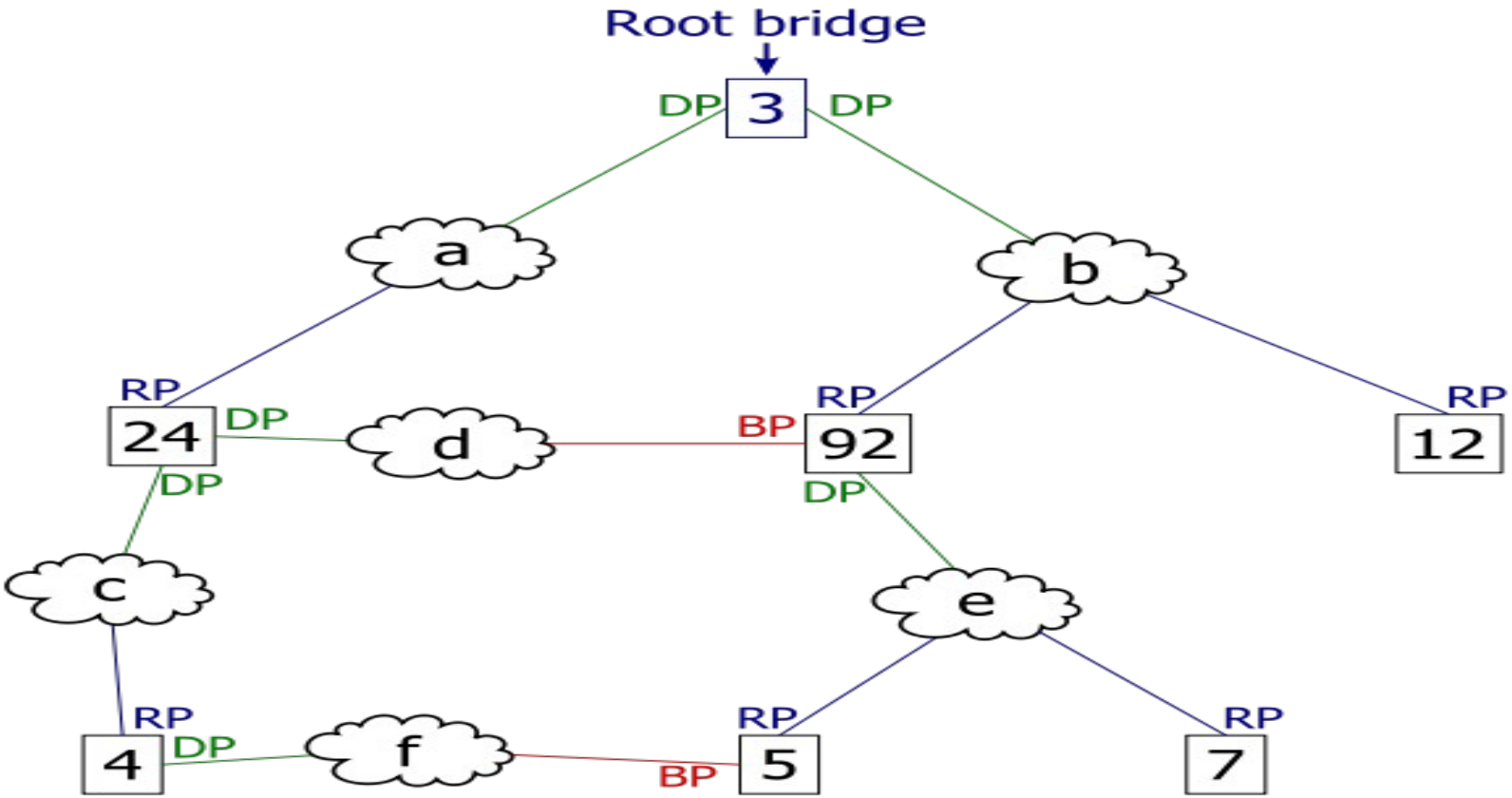
Предполагаме, че стойността на всеки сегмент е **1.** Най-късият път от bridge 4 до root bridge минава през сегмент c.

Избор на designated port



Най-късият (с най-малка стойност) път до root от мрежов сегмент **e** до root минава през bridge 92.

Spanning Tree - резултат



Активни портове, които не са root port или designated port са блокирани (blocked port).

Виртуални ЛМ (Virtual LANs)

VLAN е комутирана мрежа, която е логически сегментирана по някакви функции и не се влияе от физическото разположение на потребителите (по етажи, сгради и т.н.).

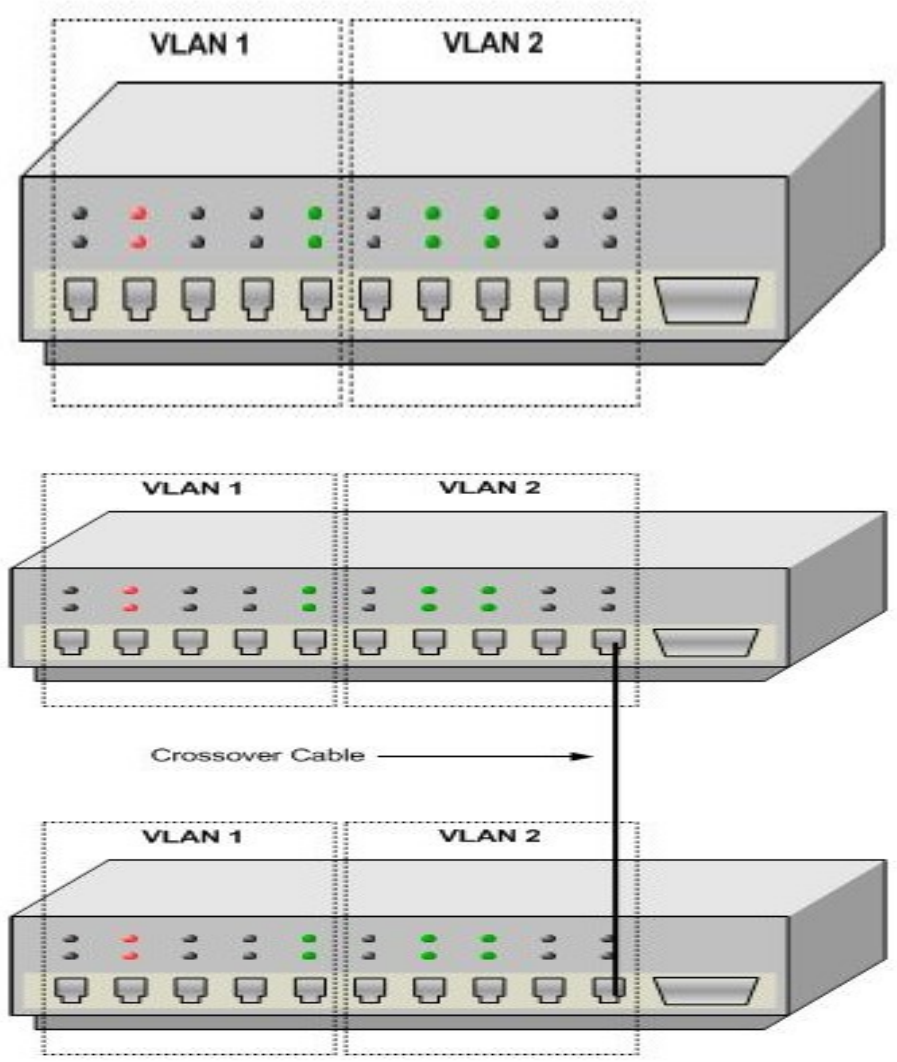
Един VLAN представлява един broadcast domain.

Сигурност. Потребителите на VLANi нямат достъп до машините на VLANj. Това може да стане през рутер.

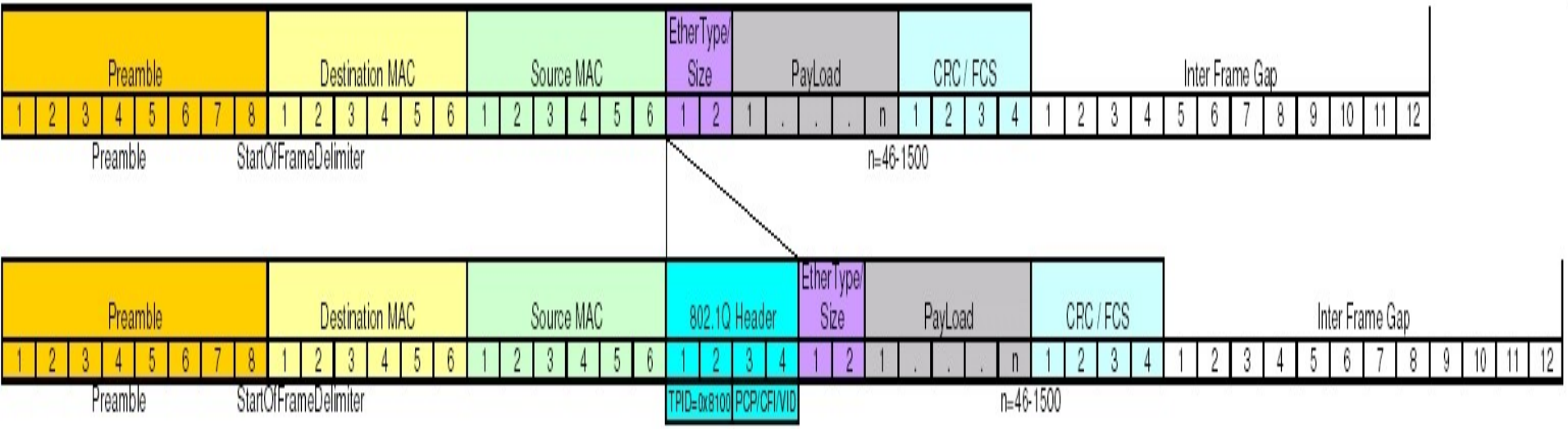
Гъвкавост. Опростява местене, добавяне, премахване на потребителски мшини.

Един порт на суич може да се присвои статично или динамично към VLAN.

Trunk портове за връзка между суичове.



VLANs – 802.1Q Tag



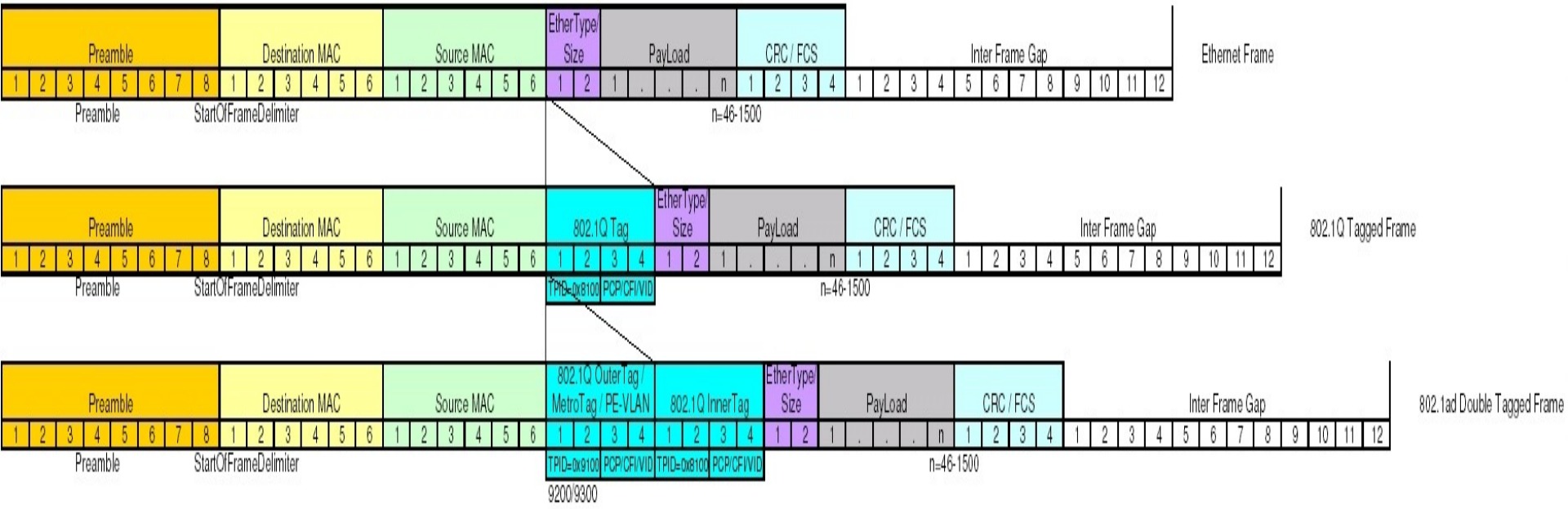
Tag Protocol Identifier (TPID): 16-битово поле: 0x8100 (IEEE 802.1Q)

Priority Code Point (PCP): 3-бита - IEEE 802.1p приоритет: 0 (най-ниско) до 7

Canonical Format Indicator (CFI): 1-бит: “0” за Ethernet суичове

VLAN Identifier (VID): 12 бита.ако е “0”, кадърът не във VLAN; позволява до 4094 VLAN-а. VLAN 1 резервирана за управление.

VLANs – (QinQ)



Double-tagging (QinQ) се използва от ISPs и MAN оператори, както и техните клиенти, да прокарват вътрешни VLAN-и през външен VLAN.

Външен tag предхожда **вътрешен** tag.

TPID - hex 9100, 9200 или 9300 за външния; но 802.1ad определя 88a8 за външни тагове.