

## #2 High speed LANy

### LAN Bridging

- dostupné od 80-tych rokov, prepájali rámce medzi homogénnymi sieťami
- teraz sú dostupné aj bridge medzi rôznymi sieťami

#### transparent bridging

- hlavne v Ethernet sieťach
- prepája sa na základe cieľovej adresy a lokálnej tabuľky

#### source-route bridging

- hlavne v Token ringoch
- prepája na základe cesty, ktorá je celá obsiahnutá v samotnom rámci

#### translation bridging

- preklad formátov a princípov medzi rôznymi, spravidla Ethernet a Token ring,
- toto už teraz robia routre

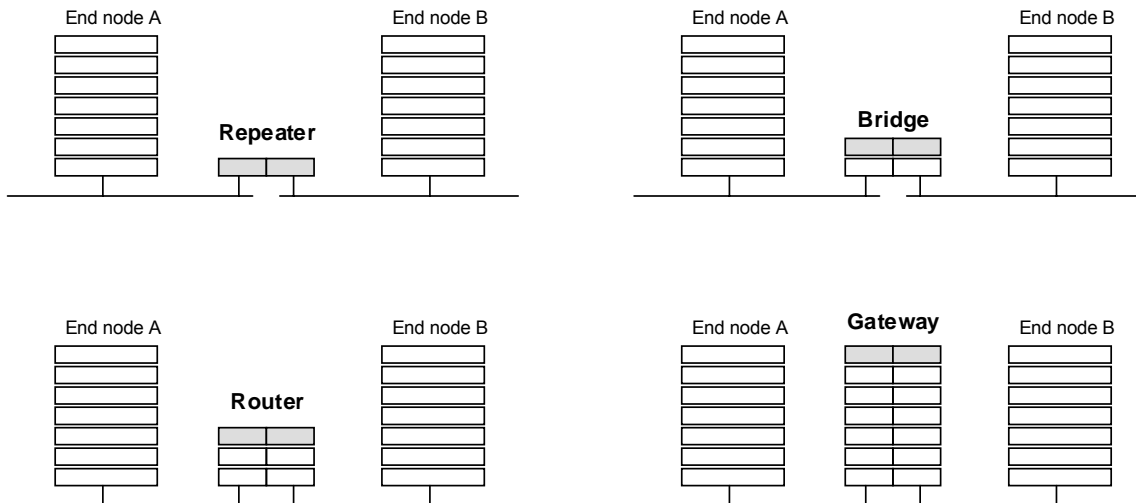
#### source route translation bridging

- kombinácia algoritmov transparentného a source-route bridgingu
- umožňujú komunikáciu medzi Ethernetom a Token Ringom

dnes už obyčajne funkciu bridgov zastupujú routre, avšak na špeciálne prípady sa hodia aj špecializované zariadenia

na prepájanie sietí sa používajú tieto zariadenia:

repeater (Layer1)  
bridge (Layer2)  
router (Layer3)  
gateway (Layer7).



**Bridge** je zariadenie, ktoré prepája siete, resp. segmenty sietí na druhej vrstve RM OSI.

Bridge môžu byť lokálne, alebo vzdialené (remote). Lokálne prepájajú segmenty v tej istej oblasti, vzdialené spájajú viaceré segmenty LAN v rôznych oblastiach obyčajne cez telekomunikačné linky. Preto musia prispôbovať rýchlosti medzi LAN a WAN, obyčajne silným bufferingom.

Vzhľadom na štandardy 802.x a ich spoločnú LLC vrstvu, sú bridge pracujúce na MAC podvrstve, vtedy spájajú rovnaké siete, a sú aj bridge pracujúce na LLC podvrstve, spájajúce aj nehomogénne siete – napr. 802.3 a 802.5. Samozrejme bridgovanie rozdielnych sietí nie je dokonalé, napr. Ethernet nemá priority.

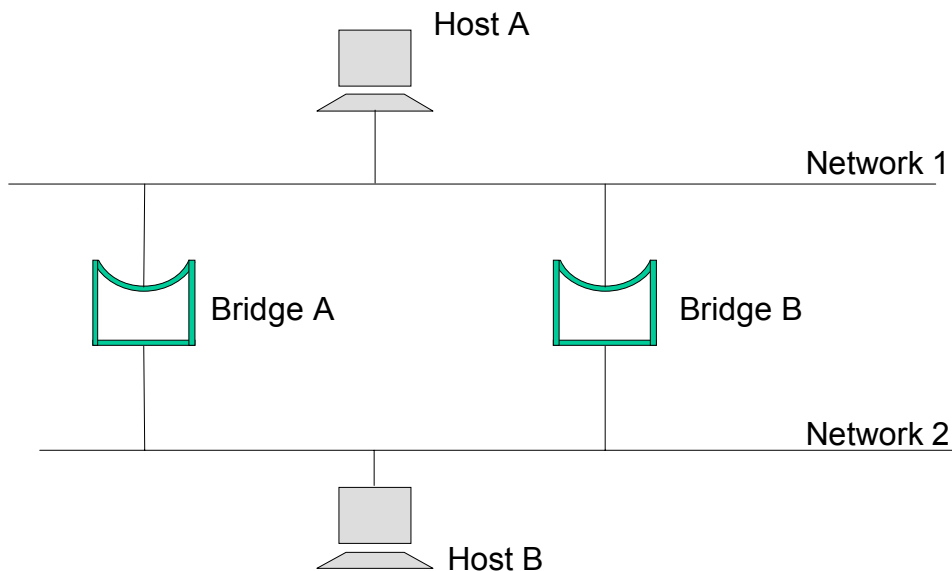
### Transparentné bridge

Prvé boli vyvinuté v DEC-u (Digital) začiatkom 80-tych rokov. Neskôr boli zahrnuté do štandardu 802.1. V Ethernet/802.3 sieťach sú veľmi rozšírené.

Pomenovanie transparentné je preto, lebo samotná operácia je pre hostov v sieti transparentná. Keď sa TB zapnú, naučia sa topológiu na základe analýzy zdrojových adries prichádzajúcich rámcov. Keď príde rámec na bridge, zistí, či pozná cieľovú adresu – tak, že sa pozrie sa do svojej tabuľky. Ak pozná, pošle rámec na príslušný interface, ak nie pošle na všetky (flooding). V každom prípade si však do tabuľky vloží zdrojovú adresu a zviaže ju s interfejsom, z ktorého rámec prišiel. TBs účinne izolujú trafiku vrámci jednotlivých segmentov, čo môže vylepšiť ich priepustnosť.

### Bridging loops

Bez dobrého bridge-to-bridge protokolu, transparent bridge algoritmus zlyhá pri viacnásobných cestách, resp. slučkách. Nasledovný obrázok opisuje prípad vzniku nekorektného správania sa (učenia a následného forwardingu) transparentných bridgov zapojených do slučky.



Čo sa vlastne stane:

- Host A pošle frame na Host B
- Oba bridge dostanú tento rámec, zapíšu si, že Host A je na sieti 1 a (oba!!) pošlú rámec na sieť 2 (hostu B)
- Sieť 2 je zdieľané médium, takže to pošlú postupne, povedzme, že Bridge A pošle rámec pre B skôr
- vtedy tento rámec dostane aj Bridge B, a zo zdrojovej adresy sa naučí, že Host A je na sieti 2 (lebo odtiaľ prišiel rámec, nevie, že prišiel z bridgu), cieľová adresa – teda host B je na sieti 2, tak rámec neforwardne – nechá ho tak
- na Bridge B pritom ešte čaká na forwardnutie na sieť 2 pôvodný rámec z Hosta A na Host B, tak ho teda pošle na sieť 2
- Teraz sa zase Bridge A chybné naučí, že Host A je na sieti 2
- Keď host B chce poslať odpoveď na Host A, bridge rámce nepošlú na sieť 1, lebo si myslia, že Host A je na sieti 2

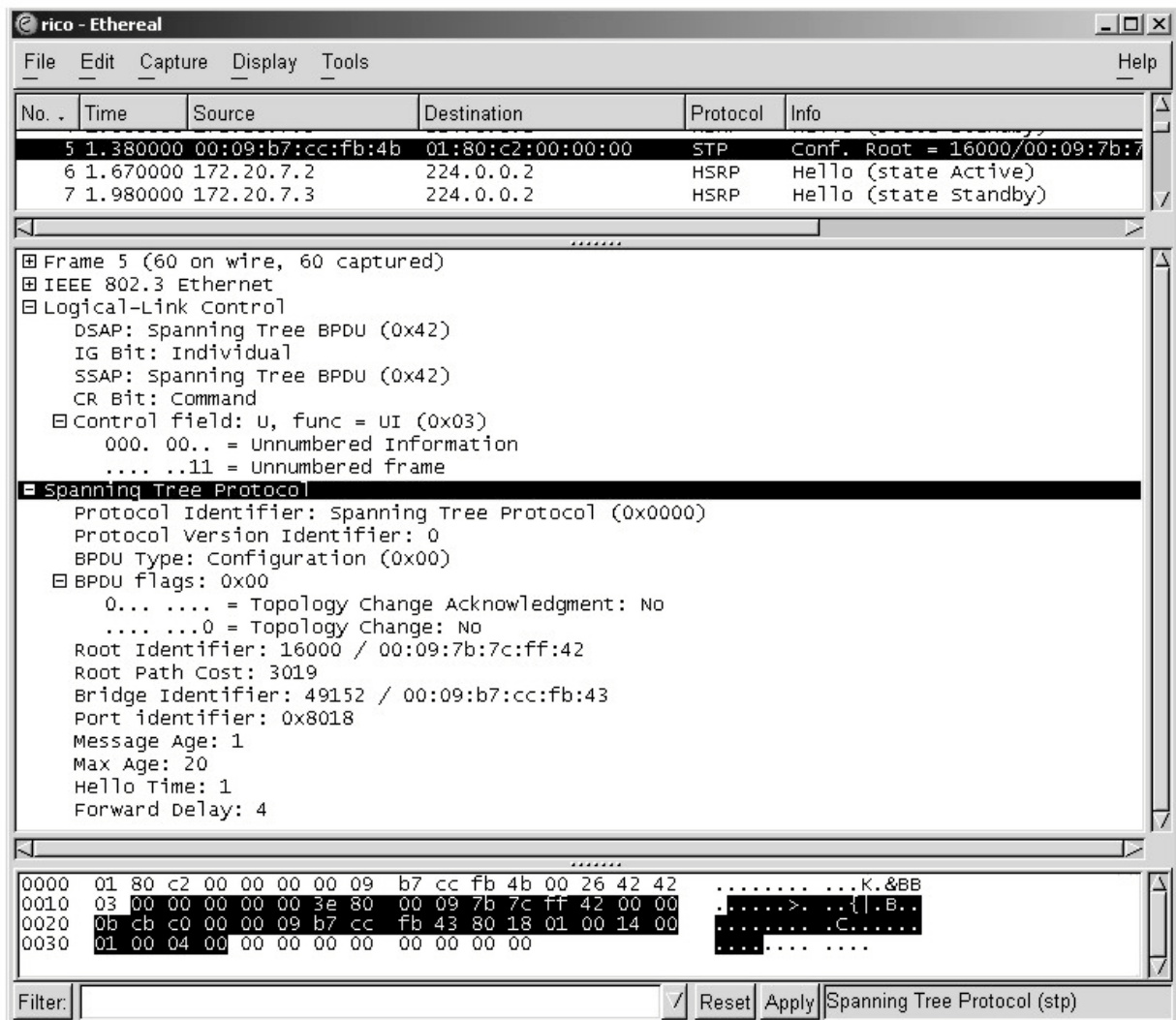
V prípade, že by sa jednalo o broadcast rámec, tak by bridge donekonečna preposielali jeho kópie a naplno by zaťažili sieť – nastal by tak kolaps siete.

Bridge sa teda vedú oklamať a obyčajne dôjde k úpnej záplave siete, v menej horšom prípade k mylným informáciám v tabuľkách bridgov a teda "odstrihnutím" niektorých hostov. Riešením je Spanning Tree Algoritmus.

### STA – spanning tree algoritmus

- vyvinutý firmou DEC, prijatý ako štandard IEEE 802.1d (nie úplne totožné, ani nie kompatibilné:)

- princíp je vytvorenie takej štruktúry siete, ktorá neobsahuje slučky – ako vetviaci sa strom (vychádza sa z teórie grafov, podľa ktorej pre každú štruktúru prepojenia nodov existuje stromové zapojenie bez slučiek)
- nadbytočné, redundantné vetvy sú uvedené do stavu blocking, takéto porty môžu byť aktivované, ak hlavná linka padla
- ako to funguje – zvolí sa route, potom najlepšie cesty k nemu, tie sa aktivujú a ostatné sú uvedené do stavu blocking
- je na to samostatný protokol, v hello-time intervale behajú neustále hello pakety, aby sa sieť vedela zrekonfigurovať, konfiguračné správy sa volajú BPDU (bridge protocol data unit)
- všetky rozhodnutia o topológii pri transparentnej topológii sa robia lokálne – každý si rozhoduje sám (neexistuje centrálna autorita) na základe výmeny správ so susedmi.



Root ID – 2B priority + MAC adresa roota

Bridge ID – identifikuje odosielateľa

Root path cost – cena cesty, ktorú má odosielajúci bridge k root bridgu

Port ID- ID portu, z ktorého bola správa odoslaná – využíva sa na odstránenie slučiek

Správa o zmene topológie obsahuje iba prvé 4B (protocol ID, protocol version a message type = 128).

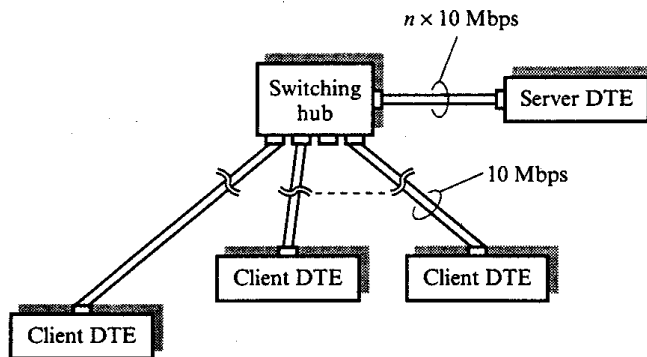
V súčasnosti sú spravidla v Ethernet sieťach používané **switche**, ktoré softvérovo realizujú aj transparent bridging. Napr. Cisco má na svojich zariadeniach implementovaný Spanning Tree Protocol, ktorý je ešte vylepšený o Port fast, Uplink fast a Backbone fast. Normálne mu to trvá 30 sekúnd, kým sa zrekonfiguruje, resp. aktivuje nový port.

## LAN switche

- switche sú funkčne podobné bridgom

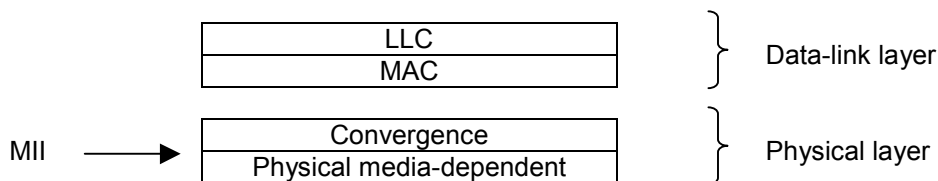
switch	bridge
veľa funkcionality realizovaného na HW	implementované SW
aj cez 500 portov	max. 16 portov
vysoká agregovaná bandwidth	
nízka latencia	

- spôsoby prepínania rámcov
  - store and forward celý rámec sa načíta, skontrolua a potom prepošle
  - cut-through odošle rámec okamžite po zistení cieľovej adresy (bez kontroly)
  - fragment free (na cisco), modifikované cut-through, odošle po kontrole prvých 64B
  - treshold detection (adaptive switching) – normálne je to cut-through, keď zistí, že rámce sú pokazené, prepne do Store and forward
- Ethernet switching je spôsob zvýšenia priepustnosti siete
- eliminuje tzv. kolízne domény, t.j. skupiny zariadení zdieľajúcich to isté médium (akoby rovnaký koaxiálny kábel) tým, že ich od seba rozdelí – celkovú agregovanú priepustnosť má obyčajne niekoľkonásobne vyššiu
- umožňuje viacero komunikácií naraz – samozrejme nie na jeden port
- pri portoch s rôznou rýchlosťou umožňuje napr. viacerým 10BaseT staniciam komunikovať naraz so serverom, ktorý je 100BaseT



## High-speed LANs – typu Ethernet

- základom je 10Mbps Ethernet, resp. IEEE 802.3
- fyzická vrstva je rozdelená na Convergence Sublayer a Physical media-dependent sublayer



MII - Medium independent interface

## 100 Mbps Ethernet

IEEE pracovalo na špecifikáciách vysokorýchlostného 100Mbps Ethernetu. Nedohodli sa však úplne na tom, či zachovávajú pôvodnú MAC metódu – CSMA/CD alebo zdefinujú novú. Tak vznikli dve grupy, a následne aj dve špecifikácie – 100BaseT a 100VG-AnyLAN.

### 100BaseT

- špecifikuje 100Mbps Ethernet na UTP a STP
- MAC zostáva kompatibilná s 802.3 MAC sublayer
- štandardizované ako 802.3u

- okrem MAC metódy si zachováva aj formát a veľkosť rámca, error-detection mechanizmus
- podporuje dual speed – 10 a 100 Mbps použitím 100BaseT fast link pulses (FLPs)
- Convergence sublayer – kóduje dáta do 4-bitov pre MII a do MAC prepája carrier sense a collision detection signal generované z PMD
- podporuje dva typy signaizácie: 100BaseX a 4T+
- **signalizácia 100BaseX**
  - má convergence sublayer, ktorá adaptuje full-duplex signalizačný mechanizmus FDDI PMD (physical medium dependent) vrstvy na half-duplex, start-stop signalizáciu Ethernet MAC podvrstvy
  - používa kódovanie 4B5B
  - $2^4 = 16$ ,  $2^5 = 32$  => mám ešte 16 slov navyše, využijem to na link control funkcie
  - je použitá pri **100BaseTX** a **100BaseFX** typoch médií
- **signalizácia 4T+**
  - používa 1 pár vodičov na detekciu kolízie a ďalšie 3 na prenos údajov
  - používa kódovanie 4B6T – teda štvoricu bitov do 6-ich ternárnych
  - $(100 \cdot (6/8)/3) = 25\text{Mhz}$  – frekvenčne v pohode
  - kódujem osmice =>  $2^8 = 256$  slov
  - možných kódových slov je  $3^6 = 729$  => redundancia – použijem na synchronizáciu a DC-free
  - umožňuje fungovanie 100BaseT na UTP category 3
  - je použitá pri **100BaseT4** type média
  - podporuje iba poloduplex
- hlavný rozdiel oproti 10BaseT, okrem rýchlosti je diameter siete – čo je 205 metrov (maximálna vzdialenosť medzi HUBom a DTE je teda cca 100 metrov), čo je asi 10x menej ako pri 10Mbps
- táto redukcia je kvôli tomu, že stanica aj pri prenose najmenšieho rámca (64B) vie zistiť kolíziu – pretože rýchlosť šírenia signálu zostala rovnaká, ale prenosová rýchlosť sa zvýšila 10x, tak aj maximálna vzdialenosť sa musí skrátiť 10x
- Fast link pulses
  - 100BaseT používa tzv. FLP – kontroluje integritu linky medzi HUBom a 100BaseT zariadením
  - FLPs sú spätne kompatibilné s 10BaseT normal-link pulses (NLPs)
  - obsahujú však viac info ako NLP, sú používané pri autonegociačnom procese medzi HUBom a zariadením na sieti
- 100BaseT Autonegotiation Option
  - 100BaseT podporuje Optional feature – autonegotiation
  - umožňuje zariadeniu a HUBu vymieňať informácie (pomocou FLPs) o ich možnostiach a teda vytvorenia optimálneho spojenia
  - speed matching pre 10/100 Mbps zariadenia
  - full-duplex mód
  - automatickú konfiguráciu pre 100BaseT4 a 100BaseTX stanice
- 100BaseT typy médií

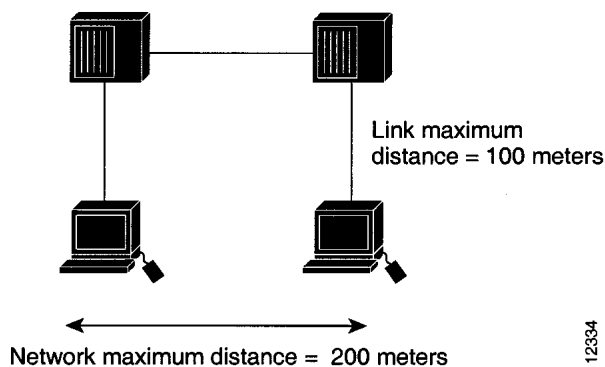
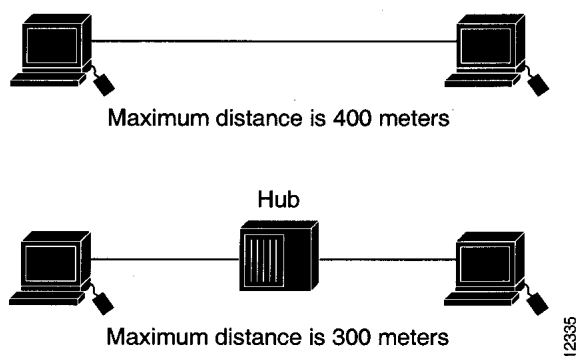
Application SW a Protokoly vyšších vrstiev		
802.3 Media Access Constorl Sublayer		
100BaseTX	100BaseFX	100BaseT4

← 100BaseT  
Physical layer

obrázky

**Table 7-2 Characteristics of 100BaseT Media Types**

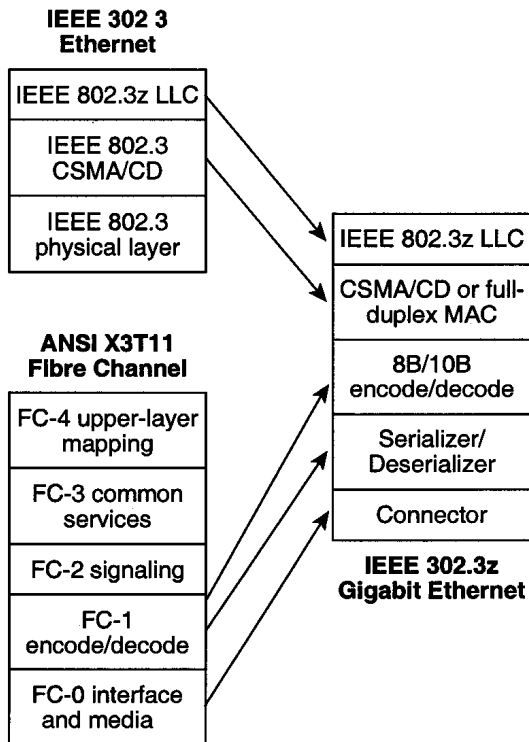
Characteristics	100BaseTX	100BaseFX	100BaseT4
Cable	Category 5 UTP, or Type 1 and 2 STP	62.5/125 micron multi-mode fiber	Category 3, 4, or 5 UTP
Number of pairs or strands	2 pairs	2 strands	4 pairs
Connector	ISO 8877 (RJ-45) connector	Duplex SC media-interface connector (MIC) ST	ISO 8877 (RJ-45) connector
Maximum segment length	100 meters	400 meters	100 meters
Maximum network diameter	200 meters	400 meters	200 meters

**Figure 7-10 The 100BaseTX is limited to a link distance of 100 meters.****Figure 7-11 The 100BaseFX DTE-to-DTE limit is 400 meters.****100VG-AnyLAN**

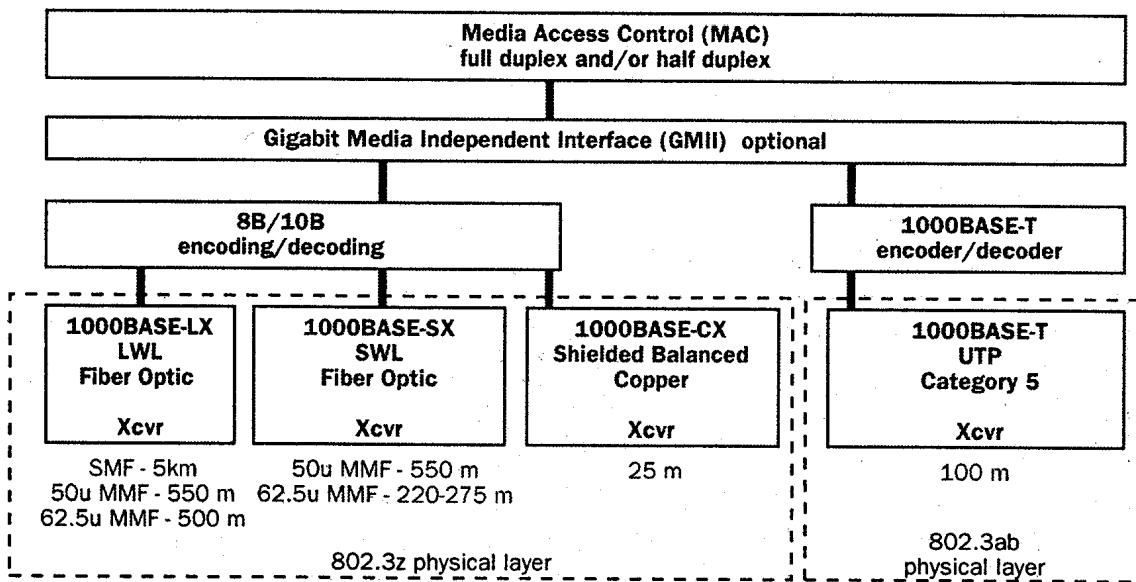
- špecifikácia pre 100Mbps Ethernet a Token Ring na štvorpárovom UTP
- MAC nie je kompatibilná s 802.3
- bola vyvinutá v HP pre nové, time-sensitive aplikácie ako upgrade z Ethernetu a TokenRingu
- štandardizovaná ako 802.12
- podporuje rôzne druhy médií (4páry UTP cat3, 2 páry UTP cat4 alebo cat5, STP, Fiber optic)
- 100VG-AnyLAN HUBy sú v hierarchickej štruktúre
- používa demand-priority access method – eliminuje kolízie, môže byť viac zaťažený ako 100BaseT

### Gigabit Ethernet

- rozšírenie štandardu IEEE 802.3 na rýchlosť 1Gbps
- zlúčením štandardov IEEE 802.3/Ethernet (LLC, prístupová metóda a frame format) a ANSI X3T11 Fibre Channel (médium, interface, encoding) vznikol štandard IEEE 802.3z Gigabit Ethernet. 802.3ab definuje 1Gb Ethernet na UTP
- bolo potrebných niekoľko zmien na fyzickom interface
- od data-link vrstvy je identický s Ethernetom



jednotlivé vrstvy pri štandardoch Gigabitového Ethernetu vyzerajú nasledovne:



## Gigabit Ethernet prenosové médiá

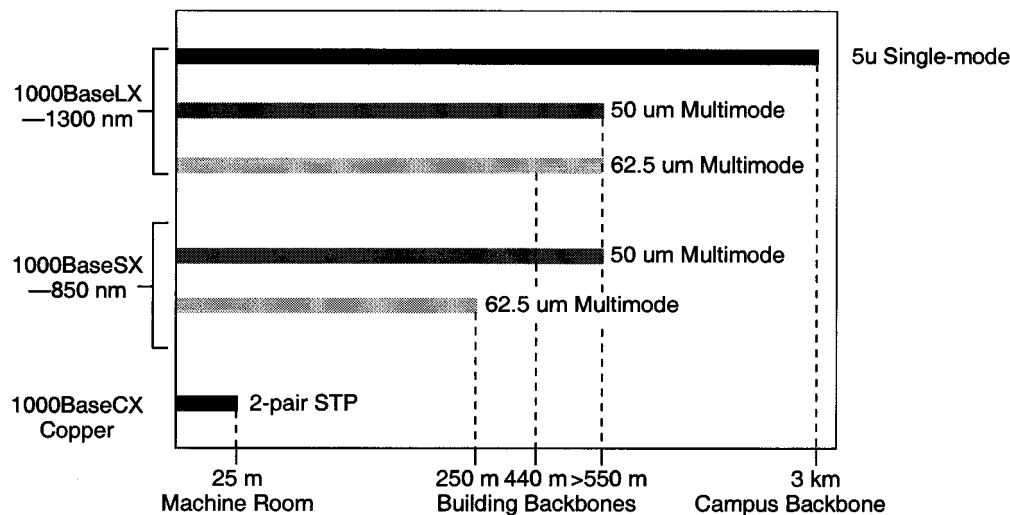
optické médiá, koaxiál

- Physical media attachment (PMA) sublayer je zhodná s PMA pre Fiber Channel, iba rýchlosť signalizácie bola zvýšená z 1,062 gigabaud na 1,25 Gbaud
- 8B/10 encoding

twisted pair

- 1000Base-T – UTP
- kódovanie 5 level PAM – naraz kódujem 2 bity, 1 úroveň mám na FEC, pri 125Mbaud

**Figure 7-18 The Gigabit Ethernet draft specifies these distance specifications for Gigabit Ethernet.**



## Gigabit Ethernet Interface Carrier (GBIC)

- GBIC interface umožňuje použiť zásuvné moduly do switchov
- podľa daného média si môžeme zakúpiť samostatné moduly
- kombinácií je veľa, preto vymysleli moduly (SW-LW na multimód (50 alebo 62,5), singlemód,...)

MAC layer

- podobná ako u Ethernetu a Fast Ethernetu (formát rámcov, collision detection, network diameter)
- half-duplex
  - rovnaké CSMA/CD
  - musí byť zachovaný slot time kvôli možnosti detekcie najkratšieho rámca – preto sa musia pridať bity za rámec
  - pri gigabite umožňuje aj frame bursting ako optional feature (stanica môže naraz vyslať celý burst rámcov)
  - veľmi neefektívne z hľadiska využitia kapacity, prakticky ani nepoužívané
- full-duplex
  - typicky medzi dvomi switchmi, servermi a pod.
  - agregovane teda kapacita 2Gbps, pri použití bundlov až 8Gbps
  - eliminovanie kolízií – ako prístupová metóda a flow control sa nepoužíva CSMA/CD
  - full-duplex flow control metóda je pod štandardom 802.3x –formalizuje full-duplex technológiu
  - 802.3x štandard nie je aplikovaný v 100Mbps kartách (tie boli skôr)
  - môže byť zriadený iba na point-to-point linkách
  - keď je prijímacia stanica zahltená, posiela späť pause frames

Jednotlivé slot time hodnoty:

10Mbps	51,2μs
100Mbps	5,12μs (pri half duplexe sa pridávajú hluché dáta)
1000Mbps	0,512μs