

Katedra softwarového inženýrství,
Matematicko-fyzikální fakulta,
Univerzita Karlova, Praha



Lekce 4: Ethernet - II

J. Peterka, 2010

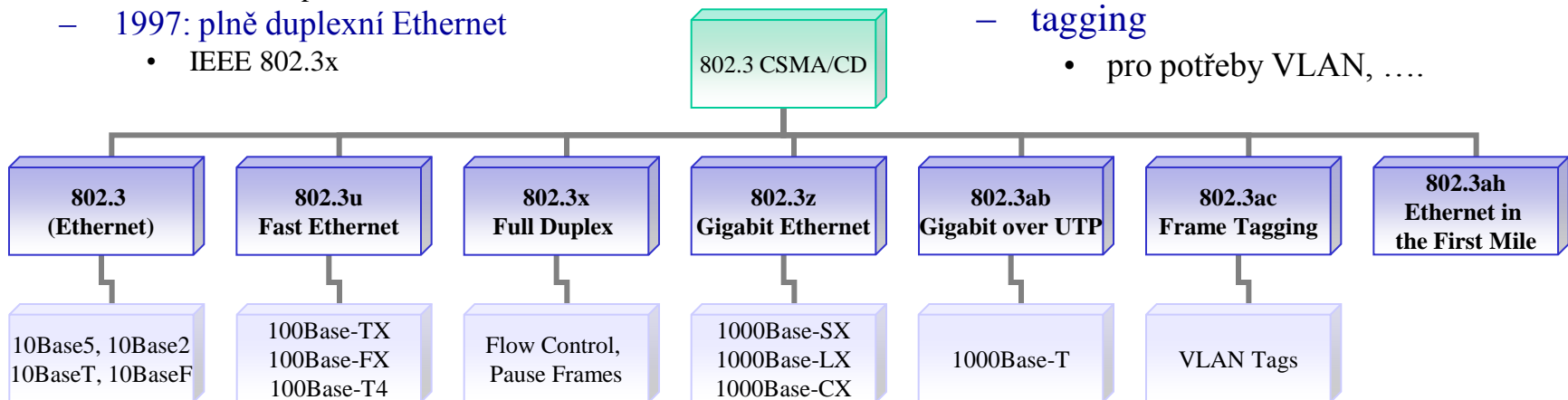
vývoj Ethernetu

- možnost použití jiných přenosových médií
 - viz předchozí přednáška
 - 1980/83: tlustý koax. kabel
 - 10Base5
 - 1985: tenký koax. kabel
 - 10Base2
 - 1990: kroucená dvoulinka
 - 10BaseT
 - 1992: optická vlákna
 - 10BaseF
- jiný způsob fungování
 - 1984: první transparentní most
 - firma DEC
 - 1991: první přepínač
 - firma Kalpana
 - 1997: plně duplexní Ethernet
 - IEEE 802.3x

- zvýšení rychlosti

10 Mbps	1976/80
100 Mbps	1995
1 Gbps	1998
10 Gbit/s	2002
100 Gbit/s	?? 2010 ??
1 Tbit/s	
10 Tbit/S	

- další
 - řízení toku
 - rámce PAUSE
 - tagging
 - pro potřeby VLAN,



stomegabitový Ethernet

- je výsledkem snahy zrychlit Ethernet 10x
- prosadily se 2 odlišné přístupy:
 - ponechat vše tak jak je, a pouze vše 10x zrychlit
 - ponechat jen takové vlastnosti, které se ukázaly jako výhodné, ostatní změnit (a 10x zrychlit)
- IEEE 802.3 dostala v roce 1992 na stůl dva návrhy
 - na 100 Mbps Ethernet „beze změn“
 - od firem Grand Junction, 3Com, SynOptics, Intel ...
 - na 100 Mbps Ethernet „se změnami“
 - od firem Hewlett Packard, IBM,
- návrh 100 Mbps Ethernetu „beze změn“ předpokládal:
 - že je třeba v maximální možné míře zachovat všechny vlastnosti Ethernetu
 - kvůli návaznosti na 10Mbps řešení
 - včetně zachování přístupové metody CSMA/CD
- návrh 100 Mbps Ethernetu „se změnami“ předpokládal:
 - že je vhodné zachovat to, co se ukázalo jako šikovné, a pozměnit to ostatní
 - za „nepříliš šikovný“ byl považován především nedeterministický charakter Ethernetu
 - návrh předpokládal změnu přístupové metody

v roce 1993 vzniká sdružení Fast Ethernet Alliance, s cílem prosadit jednotný standard Fast Ethernetu

filosofie stomegabitového Ethernetu

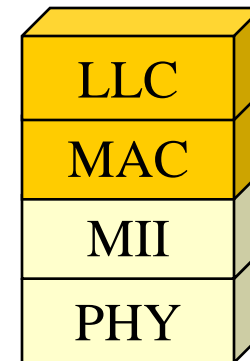
- návrh 100 Mbps Ethernetu „beze změn“ předpokládal:
 - že není nutné reagovat na principiální změnu danou kabeláží na bázi kroucené dvoulinky
 - že původně sdílené přenosové médium už je v zásadě dedikované
 - návrh se snaží nadále přistupovat k přenosovému médiu spíše jako ke sdílenému
 - a vyšší efektivity se snaží dosáhnout switchováním
- oba návrhy se sešly v komisi IEEE 802.3
 - ta rozhodla, že návrh „Ethernetu se změnami“ již není Ethernetem
 - kvůli tomu, že nepoužívá metodu CSMA/CD
 - standardizovala návrh „Ethernetu beze změn“
 - jako tzv. Fast Ethernet, neboli 100BaseT
 - skrze standard [IEEE 802.3u](#), schválený [v červnu 1995](#)
- návrh 100 Mbps Ethernetu „se změnami“ předpokládal:
 - že je vhodné reflektovat na principiální změnu vlastností kabeláže
 - že každý uzel má svou vlastní dedikovanou přípojku
 - a využít ji k dosažení celkového determinismu a vyšší efektivity
- návrh „Ethernetu se změnami“ ale nebyl zcela smeten ze stolu
 - IEEE 802 pro něj vytvořila samostatnou pracovní skupinu
 - IEEE 802.12
 - a ta přijala návrh jako svůj standard
 - v červnu 1995
 - nikoli pod názvem „Ethernet“
 - ale jako **100VG Any-LAN**

100BaseT, 802.3u

- 10x násobné zrychlení se dosáhlo:
 - 10x násobným zkrácením bitového intervalu
 - zkrácením maximálního dosahu kabelových segmentů
 - efektivnějším kódováním
 - 10 Mbps: kódování Manchester
 - 2 změny na 1 bit
 - 100 Mbps: kódování 4B/5B a NRZI/MLT-3
 - 5 změn na 4 bity
 - další změny oproti 10Mbps verzi:
 - zavedení mechanismu pro detekci rychlosti (auto-negotiation of media speed)
 - umožňuje to vyrábět síťové karty pro 10/100 Mbps, které samy rozpoznají rychlost a přizpůsobí se
 - nejkratší možný odstup mezi rámci (IPG, Inter Packet GAP) se zmenšil desetkrát
 - z 9,6 μ sec. na 0,96 μ sec.
 - beze změny naopak zůstal:
 - formát linkových rámců
 - používají se přesně stejné rámce
 - linkové adresy
 - 48-bitové ethernetové adresy
 - přístupová metoda
 - CSMA/CD – zůstala, ačkoli nemusela
 - varianta "beze změn"
 - ...
 - důsledek:
 - migrace z 10 Mbit/s na 100 Mbit/s je snadná, stejně jako koexistence obou rychlostních verzí
 - mohou existovat přepínače s porty 10/100 Mbit/s
 - nebo rozhraní, která se přizpůsobí svou rychlostí
-

fyzická vrstva 100BaseT

- „fyzická vrstva“ 100 Mbps Ethernetu se rozdělila na dvě podvrstvy
 - Medium Independent Interface (MII)
 - Physical Layer Device (PHY)
- zavedla se možnost používání různých druhů kabeláže
 - dvoulinky (UTP) kategorie 5
 - UTP kategorie 3
 - optických vláken
- konkrétní řešení (standarty) pro UTP kat. 5 a optická vlákna převzaty z FDDI
 - jde o standardy ANSI X3T9.5
 - TP-PMD, resp. CDDI (pro TX)
 - SMF-PMD (pro FX)

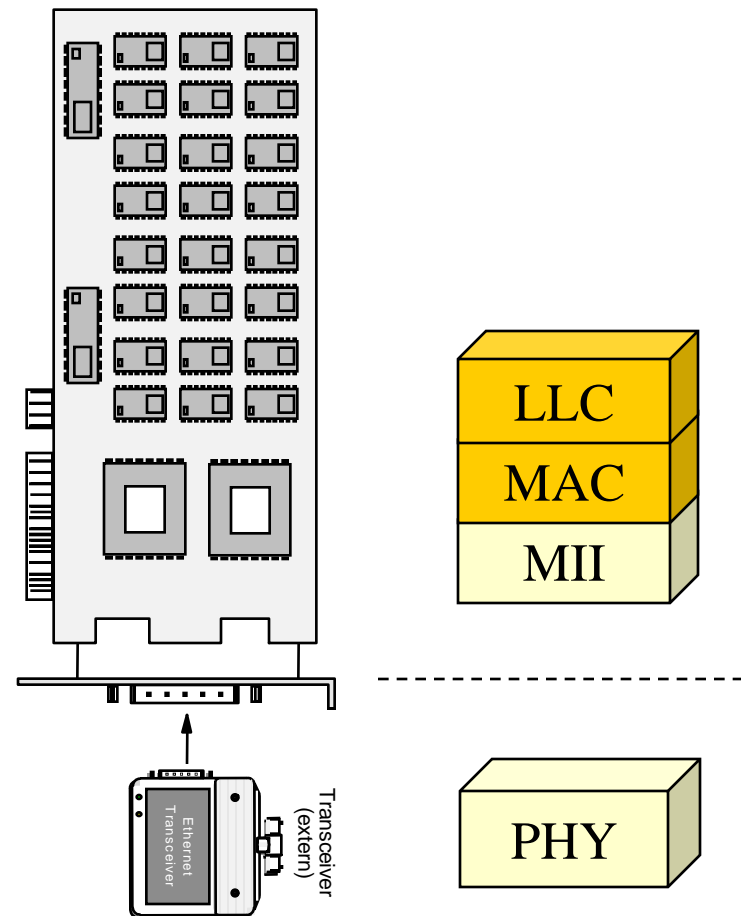


- **100 Base TX**
 - říká jak provozovat 100Mbps Ethernet nad 2 páry dvoulinky kategorie 5
- **100 Base FX**
 - dtto, pro optická vlákna
- **100 Base T4**
 - říká jak provozovat 100 Mbps Ethernet nad 4 páry dvoulinky kategorie 3 („telefonní“)
- existuje též varianta 100BaseT2
 - pro 2 páry dvoulinky kat. 3

souhrnně označováno
jako 100BaseX

vrstva MII (Medium Independent Interface)

- díky rozdělení fyzické vrstvy je opět možné "osamostatnění" transceiveru
 - a jeho propojení se síťovou kartou pomocí drop kabelu
 - max. 0,5 metru
 - častěji je ale transceiver integrován na kartě
- důvod:
 - podvrstva PHY je z velké části analogová
 - podvrstva MAC je již digitální
- podvrstva PHY je implementována v transceiveru
 - liší se pro TX, T4 a FX
- vrstva MII zajišťuje přizpůsobení mezi PHY a řídicími obvody Ethernetu
 - které mj. implementují přístupovou metodu CSMA/CD
 - MII lze chápat jako náhradu AUI



kódování bitů

10Base5&2

přenosová rychlost: 10 Mbit/s

vezme se

1bit

vytvoří se

1bit

zakódují se pomocí

kódování Manchester
(2 úrovně signálu: +,-)

frekvence přenášeného signálu:
10 MHz

100BaseTX

přenosová rychlost: 100 Mbit/s

4bity

kódování 4B/5B

$2^5 = 32$ možností, vybírají se pětice bitů alespoň se 2 jedničkami

5bitů

modulační rychlost: 125 MBaud

kódování MLT-3
3-úrovně signálu
(-,0,+)

frekvence signálu: 31,25 MHz

100BaseT4

přenosová rychlost: 100 Mbit/s

8bitů

kódování 8B/6T

$3^6 = 729$ možností, vybírají se šestice ternárních stavů signálů (-, 0, +), v každé šestici musí být alespoň dvě změny napětí

6 změn signálu
(-,0,+) na 8 bitů

frekvence signálu: 37,5 MHz

rozkládá se do 3 vodičů,
každý s $f = 12,5$ MHz
indikace kolize

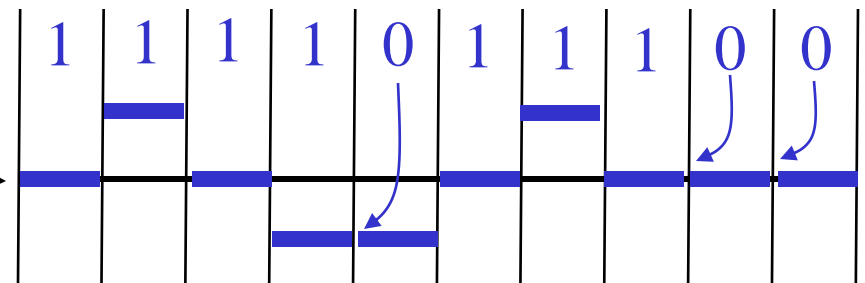
dosah 100BaseT

- důsledek zachování metody CSMA/CD:
 - minimální velikost linkového rámce zůstává stejná
 - 512 bitů (64B)
 - doba pro detekci kolize (ne)zůstává stejná
 - stále: 64B, resp. 512 x 1 bitový interval
 - ale doba přenosu 1 bitu se 10x zkrátila !!!
- důsledek pro maximální dosah (velikost kolizní domény):
 - zmenšuje se !!!
 - nikoli 10x !!!
- důsledek různého kódování u 100BaseTX, FX a T4:
 - existují dva druhy opakovačů
- obecné zásady:
 - žádný segment z kroucené dvoulinky nesmí být delší než 100 metrů
 - žádný optický segment nesmí být delší než 412 metrů
 - drop kabely (MII kabely, mezi transceiverem a kartou) nesmí být delší než 0,5 m
- další omezení se uplatňují při použití opakovačů
 - nelze mechanicky sčítat délky segmentů
 - existují konkrétní pravidla pro spojování různých segmentů pomocí různých opakovačů

100BaseTX - kódování

- v praxi dnes nejčastěji
- používá 2 páry kroucené dvoulinky kategorie 5
 - souběžný přenos po obou párech představuje kolizi
- představa:
 - místo každé čtveřice bitů se přenáší pětice bitů
 - 4B/5B
 - volí se tak, aby v pěti bitech bylo nejméně 2 jedničky
 - pak nastupuje kódování MLT-3 (Multi-Level Transition)
 - 0 nechává signál beze změny
 - 1 znamená změnu signálu
 - změna není mezi 0 a 1, ale na další prvek v posloupnosti 0, -, 0, + atd.

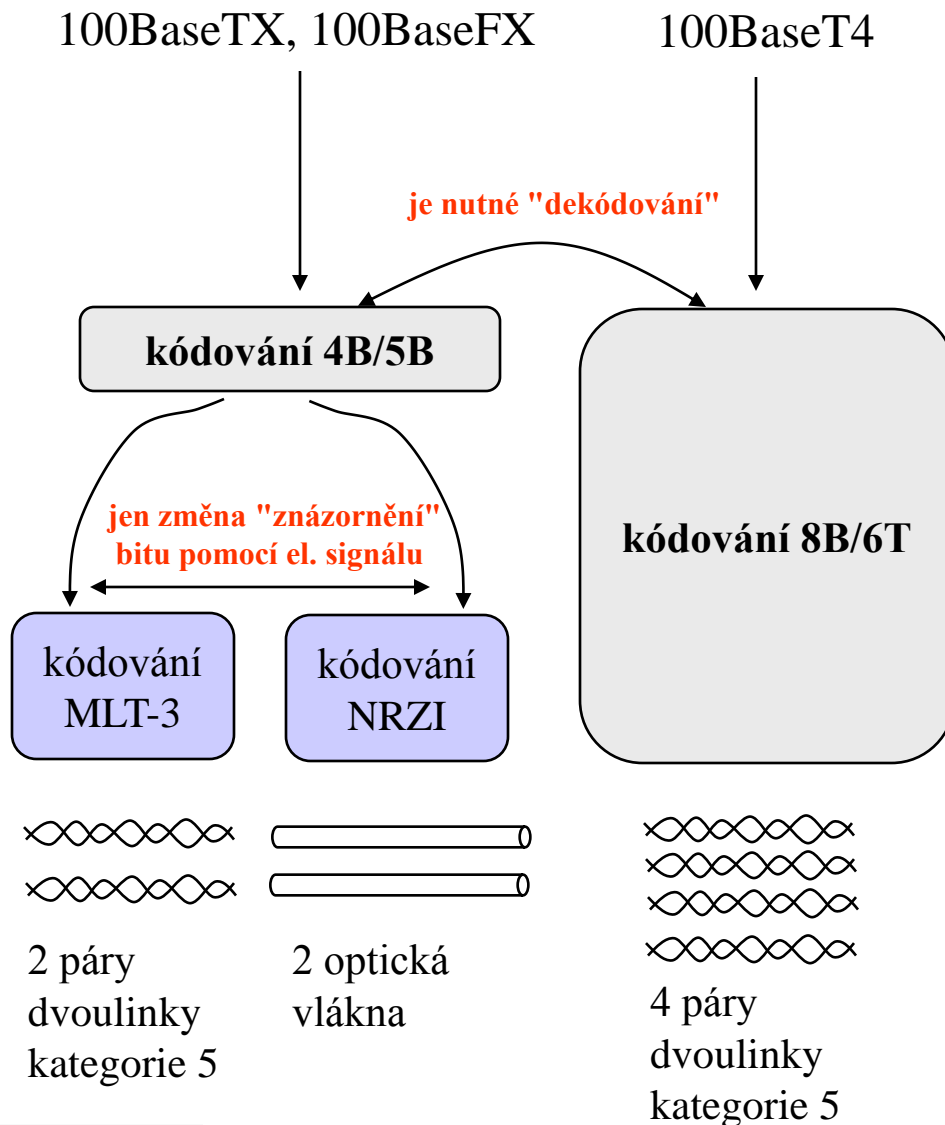
Příklad: 0000 $\xrightarrow{4B/5B}$ 11110
1110 $\xrightarrow{4B/5B}$ 11100



- jiná představa: jde o sinusovku o $\frac{1}{4}$ frekvenci
 - 1 ji "nechává běžet"
 - 0 ji zastavuje
- výsledkem je efektivní zpomalení frekvence přenášeného signálu
 - na $\frac{1}{4}$
 - 125 Mbit/s dat se přenáší pomocí signálu o frekvenci 31,25 MHz

důležité kvůli vyzařování !!!

rozdíly v kódování



- 100BaseTX a 100BaseFX používají velmi podobné kódování
 - stejné 4B/5B
 - jiné "fyzické" kódování
 - MLT-3 vs. NRZI
 - přechod mezi TX a FX je jednoduchý a rychlý
 - jen na úrovni "fyzického kódování"
- 100BaseT4 používá úplně jiné kódování
 - 8B/6T místo 4B/5B+MLT-3/NRZI
- přechod mezi 100BaseT4 a TX/FX je složitější
 - vyžaduje "dekódování" – převod různých velkých sekvencí bitů
 - zabere to mnohem více času

2 druhy opakovačů 100BaseT

- Class I

- tzv. "**Translational Repeater**"
- „dekóduje“ jednotlivé bity
 - překládá mezi různými druhy kódování (různými médii)
- umožňuje přechod mezi různými přenosovými médii
 - např. 100BaseTX a FX
- generuje zpoždění v délce 140 bitů
 - bitových intervalů
- v kolizní doméně smí být jen 1x
 - protože je pomalý ...

- Class II

- tzv. "**Transparent Repeater**"
- „nedekóduje“ jednotlivé bity
 - pouze „vyhlazuje“ signál, nesnaží se jej interpretovat ani na úrovni bitů
- dokáže propojit jen segmenty se stejným způsobem (logického) kódování
 - TX a TX, nebo TF a TF, nebo T4 a T4, nebo TX a TF
 - nikoli "křížem,, mezi TX/FX a T4
- v kolizní doméně smí být až 2 tyto opakovače
 - ale max. 5 metrů mezi nimi

max. velikost kolizní domény

POUŽITÝ OPAKOVAČ	TWIST	OPTICKÉ VLÁKNO	T4 + FX	TX+FX
Žádný	<i>100m</i>	<i>412m</i>	<i>N/A</i>	<i>N/A</i>
1x Class I	<i>200m</i>	<i>272m</i>	<i>231m</i>	<i>260m</i>
1x Class II	<i>200m</i>	<i>320m</i>	<i>N/A</i>	<i>308m</i>
2x Class II	<i>205m</i>	<i>228m</i>	<i>N/A</i>	<i>216m</i>

migrace z 10BaseT na 100BaseT

- přechod v zásadě znamená:
 - výměnu síťových karet (NIC)
 - výměnu aktivních síťových prvků (opakovačů, přepínačů)
- kabeláž může zůstat (často) beze změny
 - dvoulinka kat. 5 je zcela beze změny
 - dvoulinka kat. 3 „spotřebuje“ dvojnásobek párů
- díky vlastnostem 100BaseT je možná koexistence 10Mbps a 100Mbps segmentů
 - k propojení nestačí opakovače, jsou nutné alespoň přepínače
- migraci lze dělat postupně
- je relativně snadná a běžná
 - od začátku se s ní počítá jako s možností
 - nebo koexistence segmentů různých rychlostí může být záměrná!!!
 - může být cílovým stavem
- existují (vyrábí se) kombinované (10/100Mbps) ethernetové přepínače
 - všechny porty 10/100 Mbps
 - rychlost poznají samy, pomocí autodetekce
 - dnes nejčastější
 - některé porty 10 Mbps, jiné 100Mbps
 - porty 10Mbps, tzv. uplink 100Mbps
 - uplink je spoj k vyššímu uzlu ve smyslu stromovitého uspořádání
- existují (vyrábí se) kombinované síťové karty
 - samy poznají, jakou rychlostí mají komunikovat
 - nebo se ručně nastaví
- možná strategie:
 - servery se připojí na 100Mbps segmenty, stanice na 10Mbps segmenty
 - karty se kupují a instalují už jen kombinované
 - nejsou o moc dražší než normální pro 10Mbps

příklad – koexistence 10BaseT a 100BaseT

přepínač 10/100 Mbps

100 Mbps

10 Mbps

10 Mbps

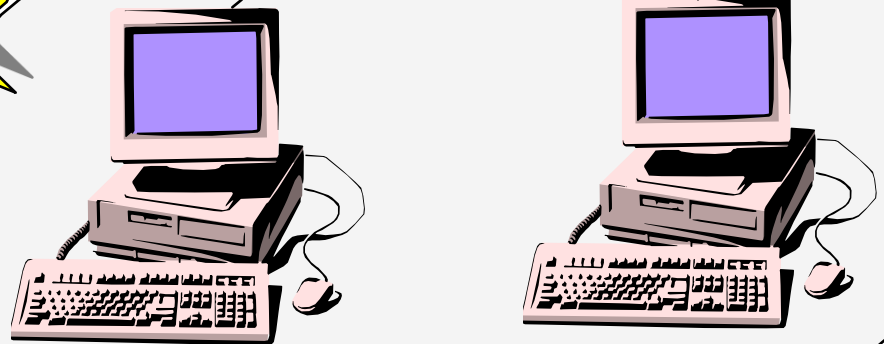
opakovač 10 Mbps

10 Mbps

10 Mbps

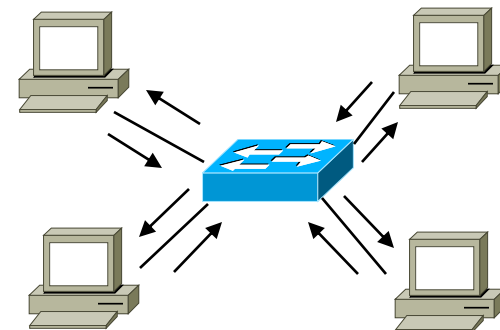
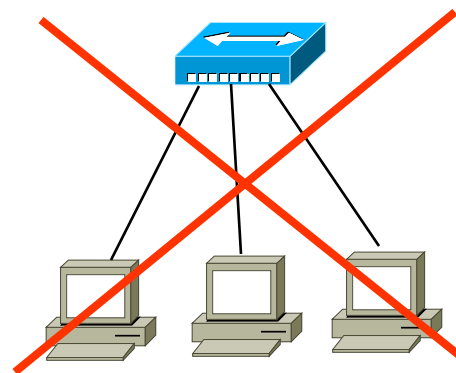
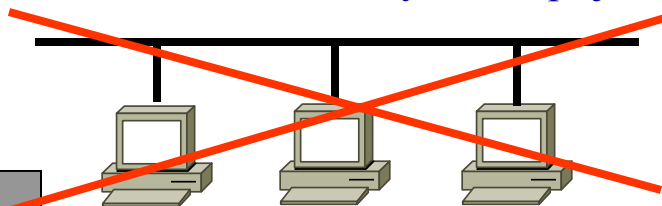
Kolizní doména
„klasického“ Eth.

Server(y) lze připojit
na segmenty 100Mbps



plně duplexní Ethernet

- „běžný“ Ethernet je poloduplexní
 - dokáže přenášet data v obou „směrech“, ale ne současně
 - je to dáno vlastnostmi původní koaxiální kabeláže
 - na poloduplexním charakteru komunikace je postavena i celá přístupová metoda CSMA/CD
 - dnes používaná kabeláž umožňuje plně duplexní provoz
 - dva páry kroucené dvoulinky (optických vláken) mohou sloužit pro příjem i vysílání současně
 - u poloduplexního Ethernetu souběh signalizuje kolizi
 - myšlenka plně duplexního Ethernetu:
 - umožní se současné vysílání i příjem
- podmínky pro plný duplex:
 - žádné sdílené segmenty
 - žádné koaxiální kabely s odbočkami
 - žádné opakovače, jen přepínače
 - mikrosegmentace (jen vyhrazená přenosová kapacita)
 - v každém segmentu jen 1 uzel
 - každý uzel musí být schopen současně vysílat i přijímat
 - bez interference (vzájemného ovlivňování)
 - všechna síťová rozhraní musí podporovat a být nastavena na plný duplex



důsledky plně duplexního Ethernetu

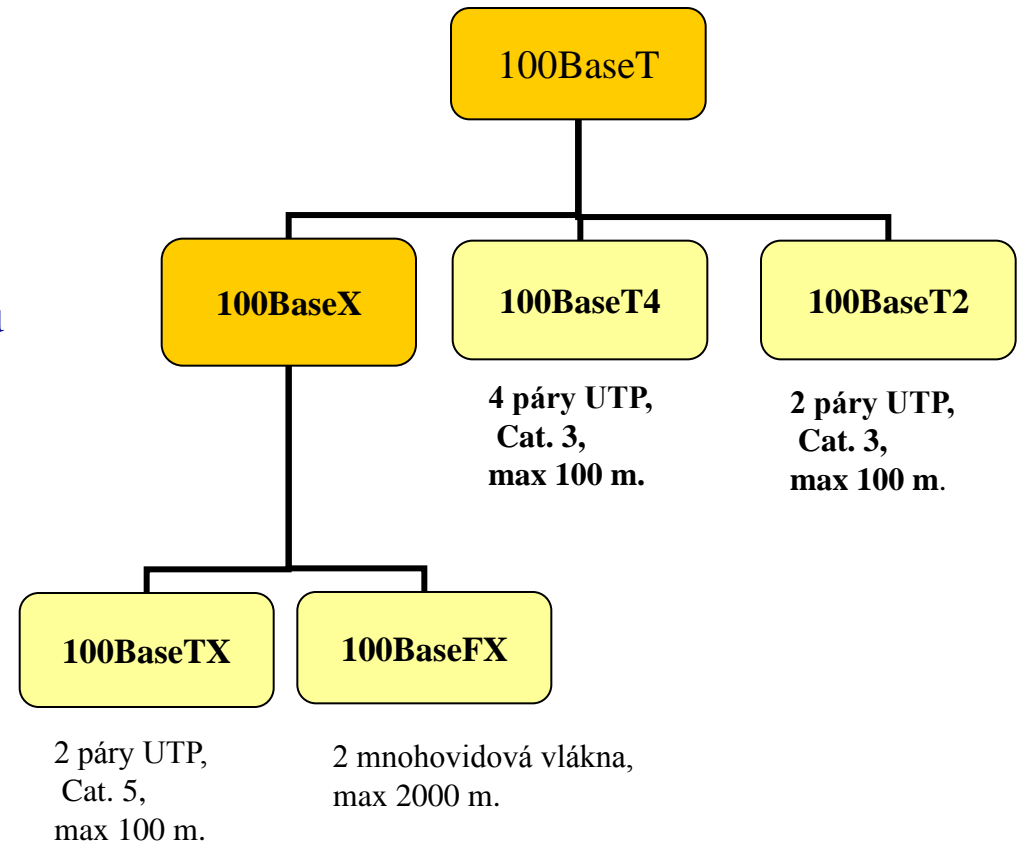
- není sdílené médium = není potřeba řídit k němu přístup
 - jen vyhrazená přenosová kapacita = je plně k dispozici (výhradně svému vlastníkovi)
 - uzlu, který je (jako jediný) připojen k (mikro)segmentu
 - odpadá možnost kolizí !!!
- odpadá potřeba přístupové metody
 - už žádná CSMA/CD !!!!
- odpadají omezení existující kvůli přístupové metodě
 - hlavně: dosah
 - již není nutné, aby se kolize rozšířila do všech částí kolizní domény v čase t (51,2 μ sec.)
 - minimální velikost rámce (512b, 64B)
 - původně nutná kvůli tomu, aby se stihla zaznamenat kolize
 - přesto se dodržuje i u plně duplexního Ethernetu, kvůli kompatibilitě
- dosah plně duplexního Ethernetu není apriorně omezen !!!!
 - fakticky je omezen obvodovými vlastnostmi média
- zvětšení dosahu otevírá dveře k tomu, aby Ethernet přestal být technologií pro síť LAN, a pronikl i do oblasti metropolitních sítí (MAN) a rozlehlých sítí (WAN)
 - dokáže překlenout kilometry ...
- zvětšuje se efektivní propustnost
 - teoreticky na dvojnásobek
 - prakticky záleží na druhu provozu
 - jednotlivé uzly obvykle generují spíše poloduplexní provoz
 - buďto vysílají, nebo přijímají
 - větší efekt je u agregovaného provozu
 - jaký generují například přepínače, servery, ...

100BaseT4

- plně duplexní Ethernet je možný u 10 megabitového, 100 megabitového Ethernetu, i u rychlejších verzí
- výjimky:
 - 10 Base5 a 10Base2
 - jejich kabeláž vytváří sdílenou přenosovou kapacitu
 - 100BaseT4
 - kvůli způsobu kódování a využití přenosového média
- 100BaseT4
 - vyžaduje 4 páry kroucené dvoulinky kategorie 3
 - používá kódování 8B/6T
 - vezme vždy 8 bitů, zakóduje je pomocí 6 změn ternárního signálu
 - celková (střední) frekvence přenášeného signálu vychází 37,5 MHz
 - data, přenášená v jednom směru, "rozkládá" do 3 párů
 - kvůli minimalizaci vyzařování
 - signál na jednom páru má frekvenci 12,5 MHz
 - již nezbývá kapacita pro přenos v opačném směru
 - čtvrtý pár se využívá pro signalizaci kolize

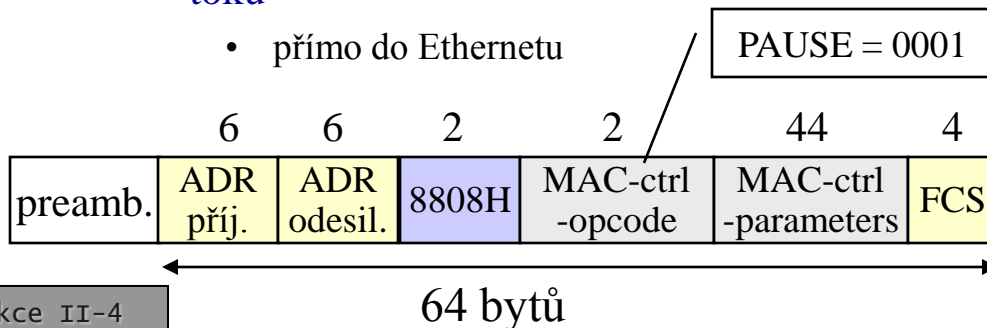
100BaseT2

- málo používaná varianta 100 Mbit/s Ethernetu
 - definuje standard IEEE 802.3y
- vyžaduje pouze 2 páry kroucené dvoulinky kat. 3
 - umožňuje plný duplex
 - v každém páru přenáší data v obou směrech současně
 - rozkládá tok do obou párů
 - používá 4-stavovou PAM
 - pulsně-amplitudová modulace
 - data kóduje do 5 různých stavů přenášeného signálu
 - přenášený signál má frekvenci 12.5 MHz
 - přijatelné vyzařování



řízení toku v Ethernetu

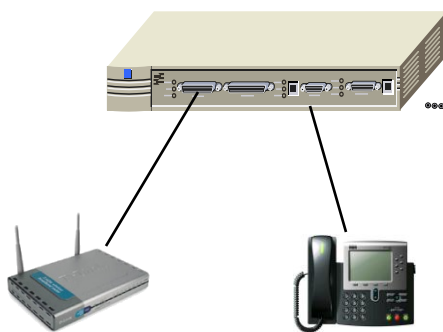
- původně:
 - Ethernet pracuje stylem "best effort"
 - nemá žádné mechanismy pro řízení toku
 - Ethernet je málo zatěžován
 - absence řízení toku nevadí
- dnes:
 - přepínače (switch-e) jsou hodně zatížené
 - řízení toku na síťové či transportní vrstvě nestačí
 - není dostatečně efektivní
 - vzniká potřeba dodatečného zpracování mechanismů pro řízení toku
 - přímo do Ethernetu
- řeší IEEE 802.3x
- princip fungování:
 - příliš zatížený přepínač pošle odesilateli dat rámeček PAUSE
 - parametr: **n**
 - říká, na kolik časových jednotek se má zdroj odmlčet
 - velikost časové jednotky:
 - 512 bitů pro 10-100 Mbit/s
 - 4096 bitů pro 1 Gbit/s
 - odesílatel pokračuje v odesílání:
 - až po vypršení n časových jednotek, nebo
 - po příjmu rámce PAUSE 0
 - rámeček PAUSE může být poslán
 - jednomu konkrétnímu odesilateli
 - unicast
 - všem odesílatelům
 - broadcast
 - na speciální multicastovou adresu 01-80-C2-00-00-01
 - šíří se jen v daném segmentu, neprochází dál přes mosty/přepínače



mechanismus autodetekce

- v 10BaseT existuje mechanismus **NLP**
 - Normal Link Pulse
 - každé zařízení vysílá každých 16 milisekund puls o délce 100 nanosekund
 - informuje "o své existenci"
 - absence tohoto pulsu na druhé straně je interpretována jako přerušení spoje (segmentu)
- 100BaseTX používá mechanismus **FLP**
 - Fast Link Pulse
 - dávka 17 až 33 NLP pulsů
 - posílá se jen na začátku
 - když dojde k propojení/zapojení kabelu
- autodetekce rychlosti
 - jen 1x NLP = 10 Mbit/s
 - celá dávka 17-33 NLP = 100 Mbit/s
- podrobnější negociace
 - pokud stojí proti sobě dvě zařízení 100 Mbit/s, mohou se domlouvat podrobněji
 - zda podporují half/full duplex
 - ... (o dalších rozšířeních)
 - dávka NLP obsahuje zakódovaný popis schopností odesílatele
 - dochází k výměně NLP
 - postupně se domluví na největší společné podmnožině

PoE: Power over Ethernet

- záměr:
 - řada "malých" zařízení
 - např. webových kamer, VOIP telefonů,
 - má jen malý příkon, ale ten je řešen samostatně
 - baterie, síťový adaptér, ...
 - další kabeláž, ...
 - idea: zkusit je napájet z jiného zařízení, po "datových" kabelech
 - první záměr standardu
 - v roce 1999
 - standard PoE (IEEE 802.3af) schválen v červnu 2003
- 
- PSE**
(Power Sourcing Equipment)
-
- PD**
(Powered Device)
- lze použít na kabelech s kroucenou dvoulinkou kategorie 5
 - varianta 1, 4 páry:
 - pro data jsou využívány dva střední páry,
 - pro napájení slouží 2 krajní páry
 - varianta 2:
 - 2 střední páry jsou využívány současně pro data i pro napájení
 - příkon zařízení:
 - až 13 W při napětí 48V
 - zařízení si samo transformuje na nižší napětí, pokud potřebuje
 - je zajištěno, že zařízení nepodporující PoE nebude poškozeno
 - součástí PoE je "otukávání"
 - PSE nejprve opatrně zkusí, zda PD podporuje PoE
 - napájí jej nejprve bezpečným malým napětím a vyhodnocuje odezvu
 - teprve když zjistí že zařízení podporuje PoE, "přidá" napájení

- Iso-Ethernet, alias
 - Isochronous Ethernet Integrated Services
 - existuje ve formě standardu IEEE 802.9a
- jde o kombinaci klasického (10Mbps) Ethernetu a kanálů ISDN
 - 1x 10Mbps Ethernet
 - 96x 64 kbps ISDN kanál B
 - 1x 64 kbps ISDN kanál D
 - celkem 10 + 6,144 Mbps
- cíl Iso-Ethernetu:
 - vyjít vstříc jak potřebám klasických „datových“ přenosů
 - kterým dobře vyhovuje paketový charakter Ethernetu
 - tak i potřebám aplikací v reálném čase, které vyžadují garanci služeb
 - např. přenosu zvuku, obrazu, telefonování, telemetrii apod.
 - těmto službám dokáže lépe vyhovět ISDN s kanály na principu přepojování okruhů
 - „propojit“ sítě LAN a WAN
 - například „rozvést“ ISDN do lokální sítě
- filosofie Iso-Ethernetu:
 - zkombinovat dvě technologie:
 - paketovou technologii Ethernetu (fungující na principu přepojování paketů)
 - technologii ISDN, fungující na principu přepojování okruhů
 - udělat to tak, aby
 - si každá technologie zachovala své přednosti
 - neovlivňovala nepříznivě „tu druhou“
 - bylo to laciné, efektivní,

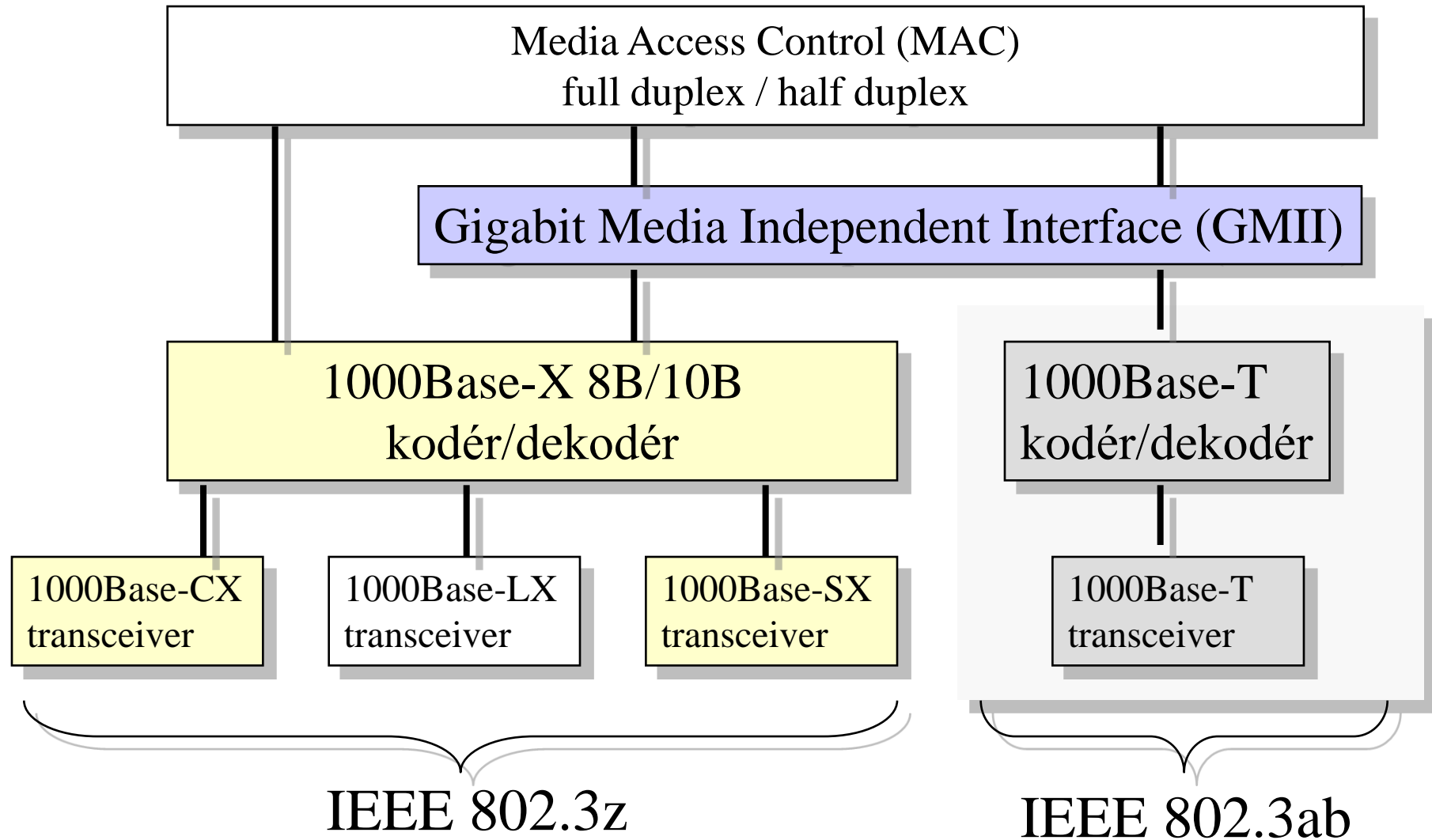
- migrace:
 - kabeláž zůstane
 - síťové adaptéry musí být nové
 - aktivní prvky (huby, přepínače, ...) musí být nové
- technické řešení:
 - časový multiplex (TDM)
 - je vytvořen 1 ISDN kanál typu P (10 Mbps), který se chová jako běžný Ethernet (10BaseT)
 - „zbytek“ je použit pro ISDN kanály B a D
 - aktivní prvky (opakovač, switch) jsou v principu kombinací
 - běžného Ethernetového přepínače či opakovače
 - ISDN ústředny
- Iso-Ethernet může pracovat ve třech režimech:
 - „Ethernet only“
 - pouze 10Mbps kanál, jako běžný Ethernet
 - Multiservice mode
 - Ethernet i kanály ISDN
 - „Isochronous only“
 - celá přenosová kapacita využívána pro isochronní přenos, tj. pro 248 B-kanálů ISDN
- Ethernetový kanál je sdílený
 - chová se jako klasický 10BaseT



gigabitový Ethernet

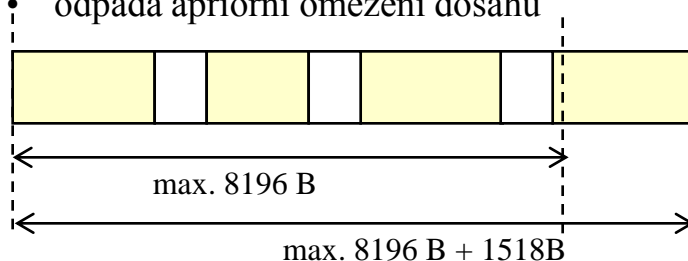
- potřeba zvýšit propustnost Ethernetu neskončila se zavedením verze 100 Mbit/s
 - snahy o zrychlení pokračují
 - nyní na 1000 Mbit/s
- gigabitový Ethernet
 - první záměr v roce 1995
 - práce na standardu začaly v roce 1996
 - založena Gigabit Ethernet Alliance
 - standard schválen v červnu 1998
- jako přenosové médium se používá
 - optické vlákno (1000Base-X)
 - konkrétní řešení převzato z Fiber Channel
 - kroucená dvoulinka (1000Base-T, 1000Base-CX)
 - dosah 25 až 100 metrů
 - zde byly potřebné standardy nově vyvinuty
- varianty standardů gigabitového Ethernetu
 - 1000Base-SX
 - 850 nm laser na mnohovidovém vlákně
 - dosah 300 metrů s využitím vlákna průměru 62,5 μm , nebo až 550 metrů s vláknem 50 μm
 - 1000Base-LX
 - 1300 nm laser na mnohovidovém nebo jednovidovém vlákně
 - dosah 550 metrů na mnohovidovém vlákně, až 3 km na jednovidovém vlákně
 - 1000Base-CX
 - STP (Shielded Twisted Pair) cable
 - dosah max. 25 metrů
 - 1000Base-T
 - UTP (Unshielded Twisted Pair, kat. 5)
 - dosah max. 100 metrů – 4 páry !!!

vrstvový model gigabitového Ethernetu



dosah gigabitového Ethernetu

- 10x násobné zrychlení by znamenalo další 10x násobné zkrácení max. délky segmentu
 - na 10 (20) metrů,
 - to je neúnosné
- principiální možnosti řešení:
 - zachová se poloviční duplex a sdílený charakter
 - zůstává metoda CSMA/CD
 - ale musí se prodloužit "slot time"
 - doba, během které je nutné detekovat kolizi
 - » ta je definována jako doba přenosu nejkratšího možného rámce
 - původně min. délce rámce 64 B (512 bitů)
 - tj. nejkratší rámec se zvětší na 512 B
 - (4096 bitů)
 - zavede se plný duplex
 - odpadá metoda CSMA/CD
 - odpadá nutnost detekovat kolize
 - odpadá apriorní omezení dosahu



- zachování polovičního duplexu:
 - lze realizovat dvěma způsoby
 - Carrier Extension
 - rámce menší jak 512 B jsou "roztaženy" na 512 B doplněním o speciální "vycpávku"
 - to ale plýtvá přenosovou kapacitou, zvláště u malých paketů
 - Frame Bursting
 - v rámci jednoho slotu (512 až 1500 B) může být vysláno více menších rámců
 - max. velikost "dávky" je 8196 B + 1518 B
 - první rámec dávky musí mít nejméně 512 B, ostatní mohou mít nejméně 64 B
- smí být použit jen 1 opakováč v kolizní doméně

gigabitový Ethernet při plném duplexu

- při plně duplexním Ethernetu přestává platit omezení na max. vzdálenost
 - **vyplývající z metody CSMA/CD**
 - protože mizí kolize
 - ale zůstává omezení dané obvodovými vlastnostmi přenosového média
- pro plně duplexní Ethernet nelze použít žádné opakovače (hub-y)
 - **musí být jen přepínače**
 - a to je drahé
- řešení: Buffered Repeater
 - **"full duplex repeater"**
 - zařízení, které funguje jako opakovač (rozesílá rámce na všechny strany)
 - ale bufferuje rámce
 - používá stejné mechanismy řízení toku jako inteligentní switche
 - ale je výrazně lacinější!!
- podle standardů IEEE lze na jednovláknových optických vláknech (1000 Base-LX) dosáhnout až na 3 km
 - **výrobci dnes nabízejí zařízení (přepínače) s gigabitovými porty s dosahem až 100 km!!**
 - např. Cisco
 - je to proprietární řešení, nemusí být kompatibilní s produkty ostatních výrobců
 - **experimentálně bylo již dosaženo vzdálenosti přes 1000 km**
 - **vše se týká dvouvláknových (nesdílených) spojů s plným duplexem**
- **důsledek: Ethernet se stává technologií použitelnou v sítích MAN i WAN**
 - **pro budování páteřních sítí**
 - kde vychází lacinější a jednodušší než např. ATM
 - **pro překonání tzv. poslední míle**

standard 1000BaseT

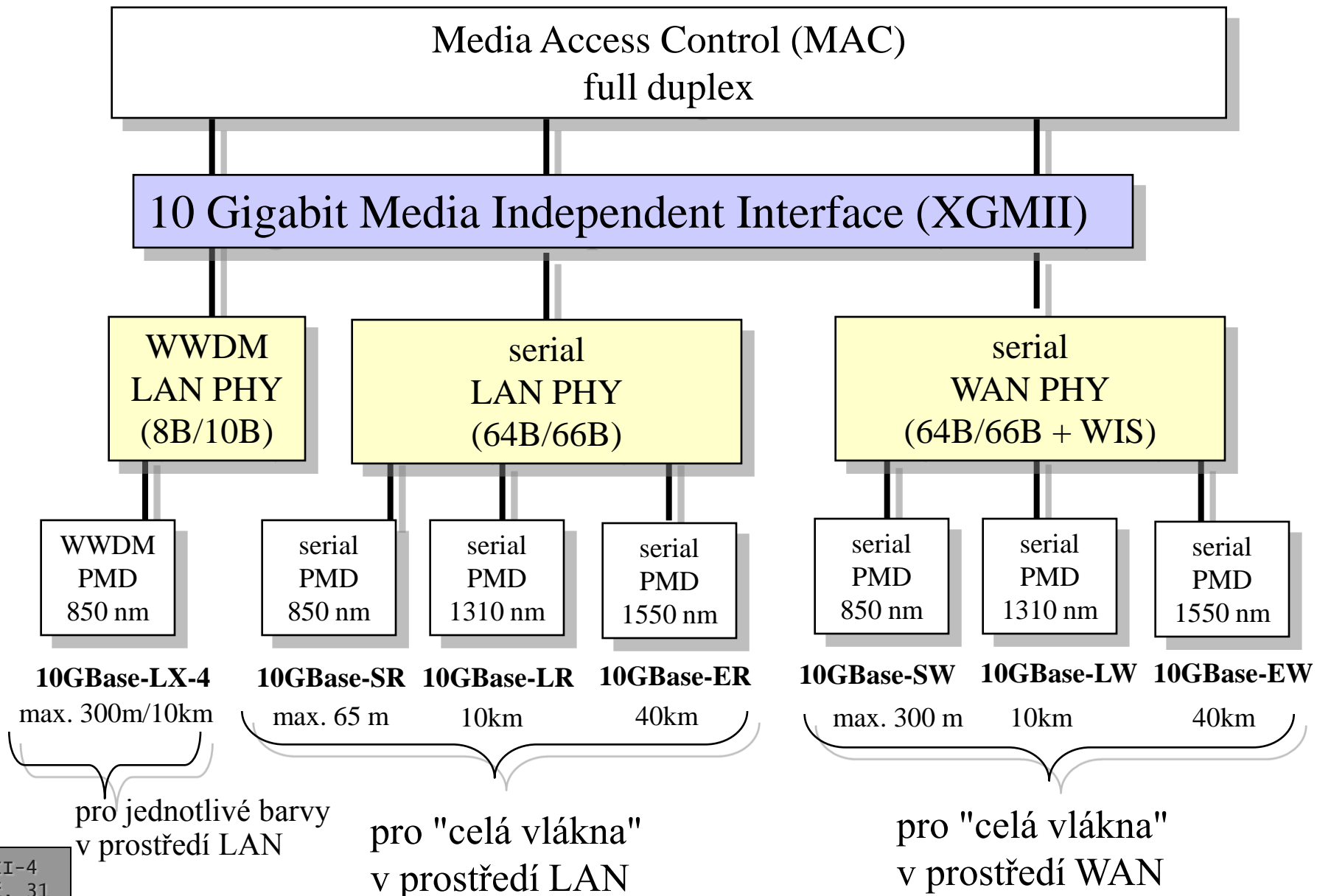
- definuje 802.3ab
 - nabízí plný duplex
 - dosah do 100 metrů
- vyžaduje 4 páry nestíněné kroucené dvoulinky
 - kategorie 5
 - všechny páry se používají pro vysílání i příjem současně
- způsob kódování:
 - datový tok se rozloží do 4 párů vodičů
 - 250 Mbit/s na každý
 - pro znázornění bitů jsou využity 4 různé stavy přenášeného signálu
 - frekvence signálu 31,25 MHz
- autodetekce rychlosti přenosu:
 - existuje i v rámci 1000BaseT
 - zpětná kompatibilita a 100BaseT a 10BaseT
 - řeší se pomocí pulsů NLP a FLP
 - není řešena v rámci optických variant
- negociace
 - domlouvání se na parametrech
 - half/full duplex
 - flow control (ano/ne)
 -
 - existuje u všech verzí gigabitového Ethernetu
 - už nepoužívá pulsy FLP ale speciální rámce



- základní rysy:
 - funguje již jen plně duplexně
 - nebude již omezení na dosah kvůli potřebě detekovat kolize
 - je provozován pouze po optických vláknech
- dosah je až 40 km
 - již není ani CSMA/CD ani kolize
- standard (IEEE 802.3ae) byl schválen v červnu 2002
 - práce na standardu byly zahájeny v roce 1999
 - v rámci 10 GiGabit Ethernet Alliance
- předpokládá se použití "celých" vláken:
 - **mnohovidových:**
 - 850 nm, max. 65 metrů
 - **jednovidových:**
 - 1310 nm, max. 10 km
 - 1550 nm, max. 40 km
- a také jednotlivých "barev" získaných technikou DWDM (4 barvy)
 - **na mnohovidových vláknech:**
 - 1310 nm, max 300 metrů
 - **na jednovidových vláknech:**
 - 1310 nm, max. 10 km
- existují verze 10 GbE pro sítě LAN a pro sítě WAN
 - vybavené dalšími funkcemi, pro provoz nad synchronními digitálními hierarchiemi (WIS, WAN Interface Sublayer)
 - SONET/SDH

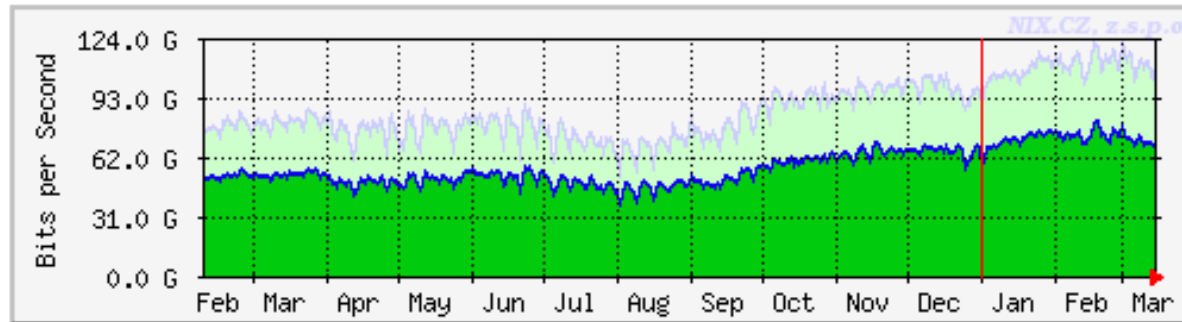
<http://www.10gea.org/>

vrstvý model 10 gigabitového Ethernetu



příklad využití – propojení lokalit NIX.CZ

Roční graf (denní průměr)



	Max	průměr	Current
Příchozí	122.6 Gb/s (21.9%)	56.6 Gb/s (10.1%)	69.3 Gb/s (12.4%)
Odchozí	122.4 Gb/s (21.9%)	56.6 Gb/s (10.1%)	69.3 Gb/s (12.4%)

- Zeleně:** příchozí provoz v bitech za sekundu
- Modře:** odchozí provoz v bitech za sekundu
- Světle zeleně:** maximální pětiminutový příchozí provoz
- Světle modře:** maximální pětiminutový odchozí provoz

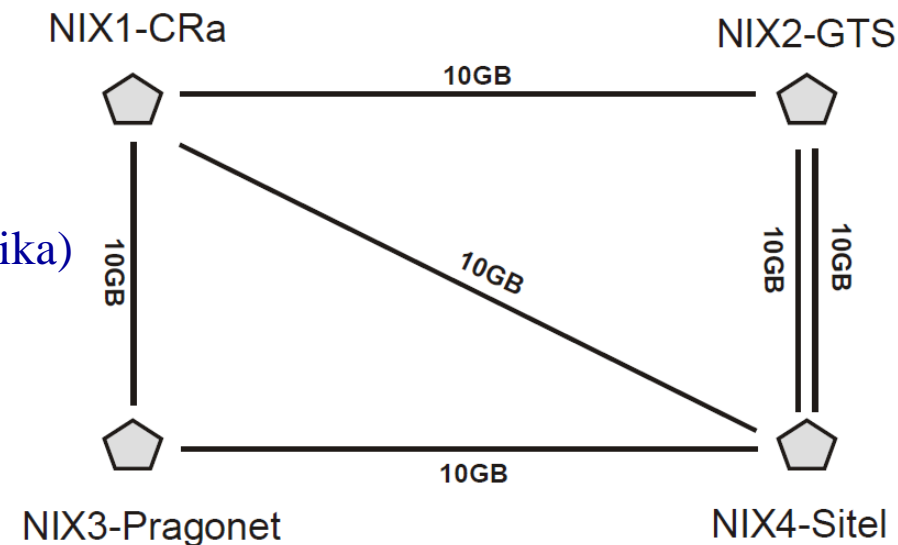
ani 10 Gb/s Ethernet už nepostačuje dalšímu nárůstu datových toků

- stejný problém mají další IX-y
- perspektivně: nutnost upgrade na rychlejší technologie

• kde je 100 Gb/s Ethernet?

• možnosti napojení ISP přes Ethernet v jednotlivých lokalitách:

- 10/100/1000 Mbit/s, konektor RJ45 (metalika)
- 1 Gb/s, konektor SFP (Small Form Factor Pluggable, optika)
- 10Gb/s, konektor XENPAK



další vývoj Ethernetu

- dalšímu zvyšování rychlosti Ethernetu nestojí v cestě žádné principiální bariéry
 - po přechodu jen na plný duplex
 - kdy odpadly kolize
 - po přechodu (ne)jen na 2-bodové spoje
 - obecnější topologie jsou realizovány pomocí prepínačů
- ale (neprincipiální) překážky se překonávají čím dál tím obtížněji
 - například útlum, zkreslení, vyzařování, přeslechy mezi páry vodičů atd.
 - proto další zrychlování není tak snadné, jak se předpokládalo
- jaké jsou další směry vývoje?
 - Ethernet pro přístupové sítě
 - EFM (Ethernet in the First Mile)
 - připravilo sdružení EFMA (Ethernet First Mile Alliance)
 - řešení převzalo IEEE a vydalo jako standard IEEE 802.3ah, v červnu 2004
 - Ethernet pro sítě MAN a WAN
 - tzv. metropolitní Ethernet (Metro Ethernet)
 - část standardů definuje IETF
 - (větší) část standardů připravuje Metro Ethernet Forum
 - Ethernet over MPLS, over DWDM
 - Ethernet over IP
 - Ethernet pro internetové providery a telekomunikační operátory
 - tzv. "carrier-grade" Ethernet, Carrier Ethernet
 - má podporu QoS

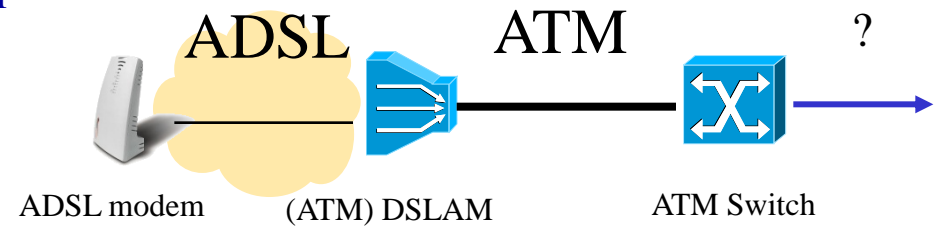
100 Gigabitový Ethernet

- předpoklad:
 - 100Gb/s Ethernet bude používána hlavně v rozlehlejších (WAN) optických sítích
 - asi pro propojování datových center/telehousů/hostingových středisek
- původní očekávání:
 - "ještě rychlejší Ethernet" by měl být dimenzován podle normovaných rychlostí v oblasti telekomunikací
 - podle rychlostí digitálních hierarchií, zejména SDH
 - důsledek:
 - uvažovaly se rychlosti **40 Gb/s, 80 Gb/s, 120 Gb/s**
- dnes:
 - "ještě rychlejší" Ethernet bude respektovat "své" normované rychlosti
 - důsledek: bude jen **100 Gb/s Ethernet**
- vývojem se zabývá IEEE 802.3
 - skupina Higher Speed Study Group (HSSG)
- teprve v prosinci 2006 byl oficiálně potvrzen záměr IEEE připravit standard 100 Gb/s Ethernetu
 - předpoklad: reálná dostupnost produktů v letech 2010/2011
- cíle 100 Gb/s Ethernetu:
 - dosah 100m na mnohovidových optických vláknech
 - dosah 10 km na jednovidových vláknech
 - možnost využití kroucené dvoulinky se zatím zvažuje
 - velké problémy s vyzařováním a přeslechy mezi páry vodičů

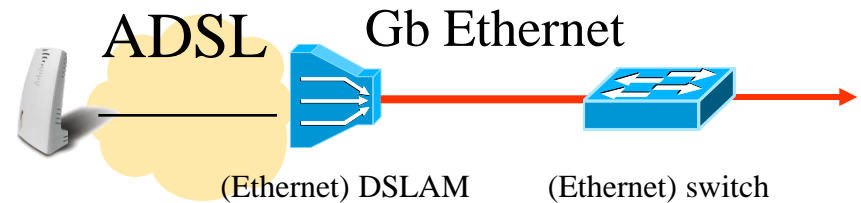
Ethernet in the First Mile (EFM)

- cíl:
 - používat i v přístupové síti takové technologie, jaké se (nejvíce) používají i "na koncích"
 - tj. Ethernet
 - snaha eliminovat režii, která vzniká propojováním a kombinováním různých technologií (ADSL, ATM, SHD, SONET atd.)
 - snaha zvýšit rychlost a dosah
 - nahradit xDSL
 - dříve se používal i termín LRE
 - **Long Reach Ethernet**
- snaha využít:
 - metalická vedení (místní smyčky)
 - optická vlákna P-P (point-to-point)
 - pasivní optické sítě (point-to-multipoint)

současný stav – ADSL a ATM



↓ Ethernet pro napojení DSLAMů na páteřní síť



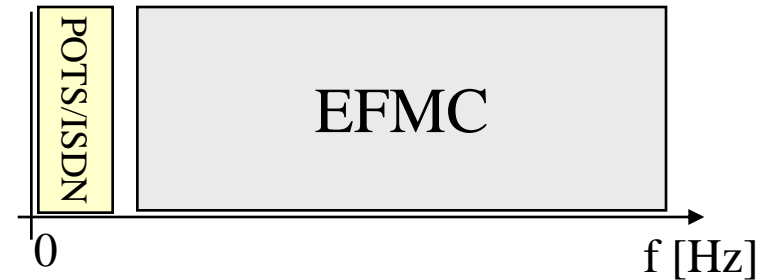
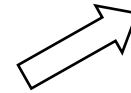
↓ nasazení EFM také na místní smyčky



Ethernet in the First Mile - varianty

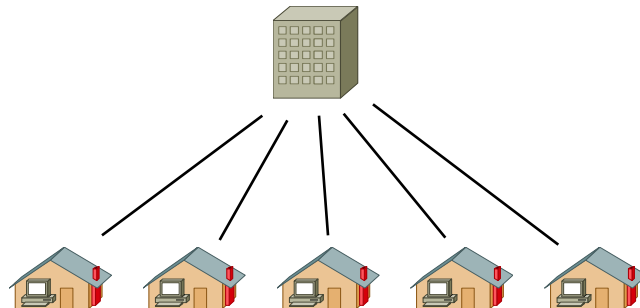
- **EFMC** (EFM over Copper)

- dvoubodové spoje po metalických vedeních
 - plně kompatibilní s ISDN a PSTN
 - tj. může koexistovat spolu s nimi "na stejném drátě"
- symetrické přenosové rychlosti
 - nikoli asymetrické jako ADSL
- alespoň 10 Mbps na 750 metrů
 - po kroucené dvoulince kat. 3
 - EMFC SR (Short Range)
- alespoň 2 Mbps na 2700 metrů
 - EMFC LR (Long Range)



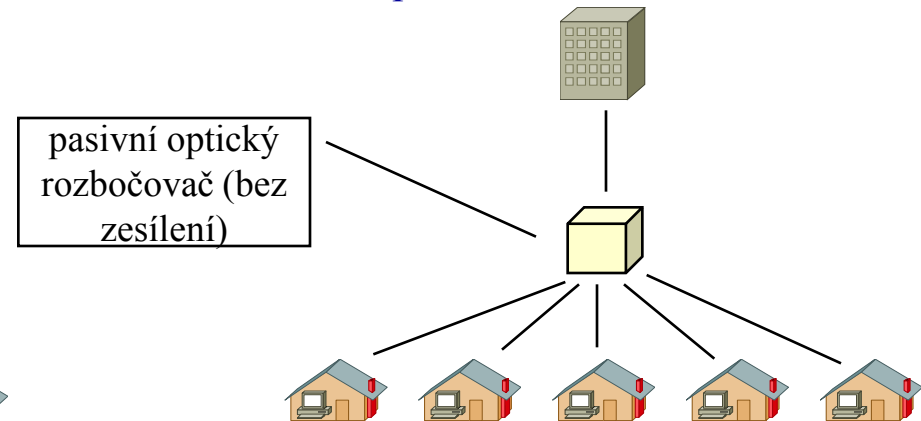
- **EFMF** (EFM over Fiber)

- dvoubodové spoje po jednovidových optických vláknech
 - plně symetrické, plný duplex
- 100 Mbps / 1 Gbps, nejméně na 10 km



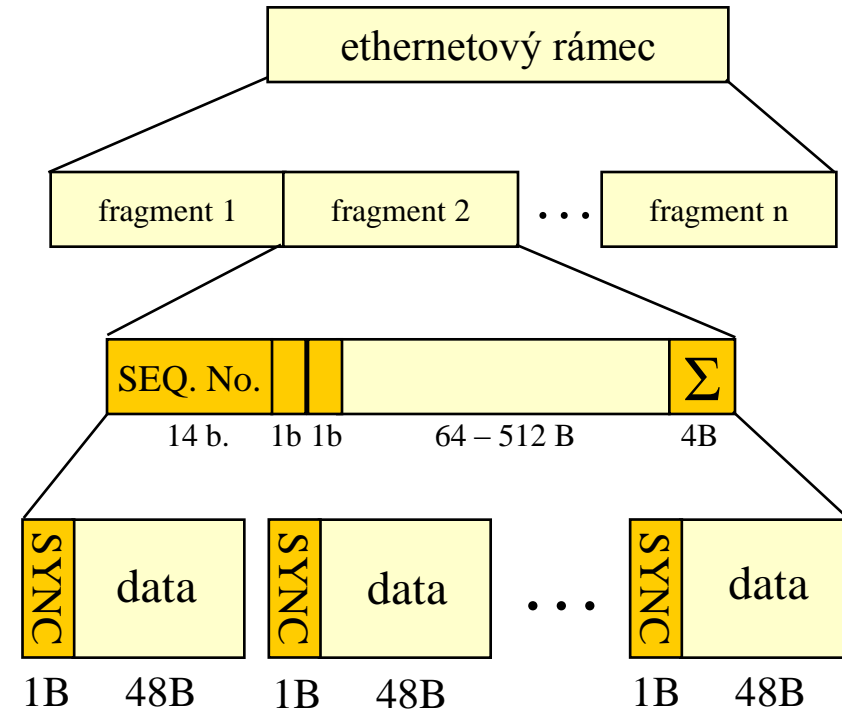
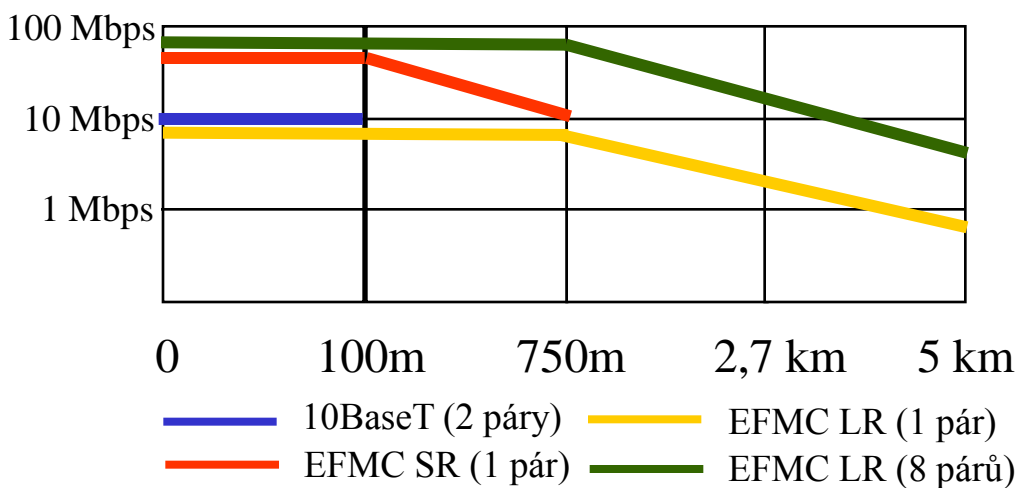
- **EFMP** (EMF PON, EMF over Passive Optical Networks)

- vícebodové (Point-to-Multipoint) spoje po optických vláknech, s pasivním rozbočením
 - šetří se optická vlákna
- symetrické přenosové rychlosti
- 1 Gbps, až na 20 km



EFMC (EFM over Copper)

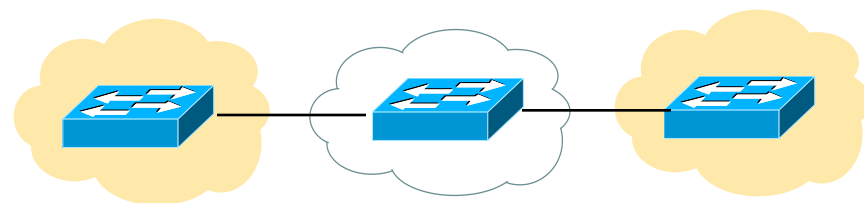
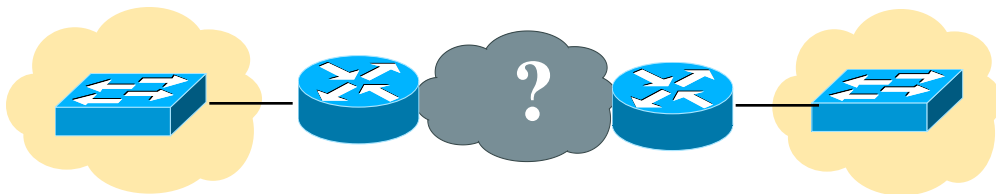
- využívá modulační techniky, používané u DSL
 - DMT (OFDM) a QAM
- může využívat více párů kroucené dvoulinky (kategorie 3)
 - skrze techniku inverzního multiplexu
 - zvyšuje tím rychlost / zlepšuje dosah
- dosah a rychlost závisí na vzdálenosti a počtu použitých párů
 - na krátké vzdálenosti bývají rychlosti i vyšší než 1/10 Mbps



- ethernetové rámce se přenášejí "rozdělené" na fragmenty
 - o velikosti 64 bytů
 - +1 byte režie, kódování 64B/65B
 - celková režie cca 5%
 - podstatně menší než u ADSL

Metropolitní Ethernet

- podobný cíl jako u EFM:
 - používat "uprostřed" stejné technologie, jako "na koncích"
 - "uprostřed" = v síti MAN či WAN
 - snížit režii, zvýšit rychlost, usnadnit škálování,
- dříve:
 - při propojování sítí LAN se muselo "přecházet" přes různé jiné technologie
 - s využitím směrovačů (a síťové vrstvy)



- záměr Metro Ethernetu
 - nabídnout propojení na úrovni linkové vrstvy
 - s možností přenosu ethernetových rámců (bez nutnosti přechodu přes síťovou vrstvu)
 - navíc s podporou VLAN, QoS atd.
- zákazník se připojuje přes standardní ethernetové rozhraní
 - přenos skrze MAN či WAN ve skutečnosti nemusí být realizován po Ethernetu
 - ale např. po ATM či po MPLS
 - pro zákazníka je to neviditelné