

Beágyazott rendszerek illesztése információs rendszerekhez BMEVIMIM343

dr. Kovácsházy Tamás
Ethernet



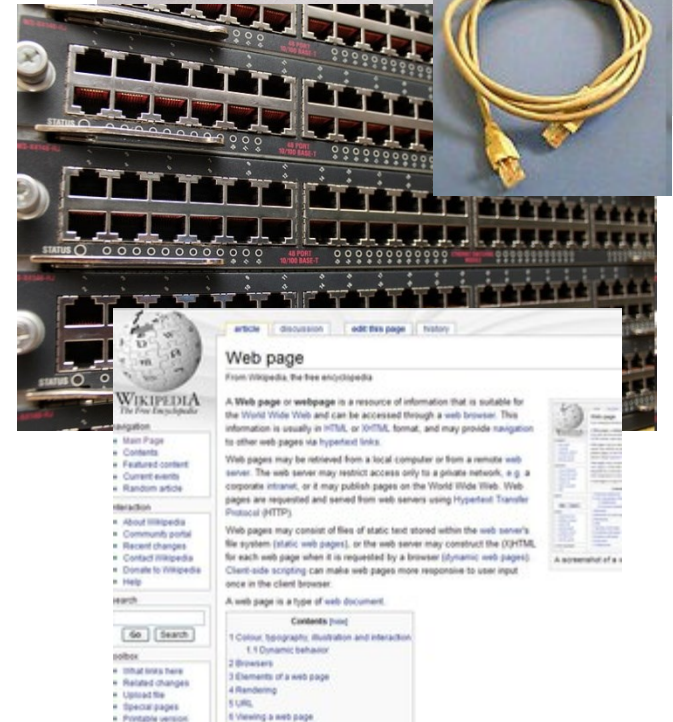
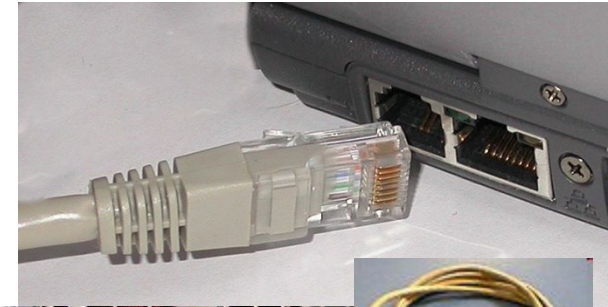
Méréstechnika és
Információs Rendszerek
Tanszék

A „gordiuszi csomó”

- Az Ethernet és a TCP/IP összetett
 - Az ismerkedés fájdalmas
 - Minden mindennel összefügg
 - **Kérdés: Hol kezdjem el a gordiuszi csomó kioldást?**
 - **Kardot ránts! (Éppen mint Nagy Sándor tette.)**
 - **Válasz: Bárhol, majd összeáll a kép idővel...**
- Személyes tapasztalat
 - 1993?: „Számítógépes hálózatok” tárgy heti 4 órában (egy félév)
Visszagondolva gőzöm nem volt a témához a vizsga után, ha jól emlékszem 4-est kaptam.
 - 1994-től foglalkozom Ethernet-tel és TCP/IP-vel
 - 1998-tól dolgozok ki és tartok ilyen tárgyakat és tanfolyamokat (SZIF, BME, cégek)
 - 1998 óta beágyazott területen is
 - **Ma is meg tudnak lepni és új dolgokat tanulok a területen!**
 - Ez az anyag szerintem a legjobb út a megismerési folyamat első lépéseinek megtámogatásához

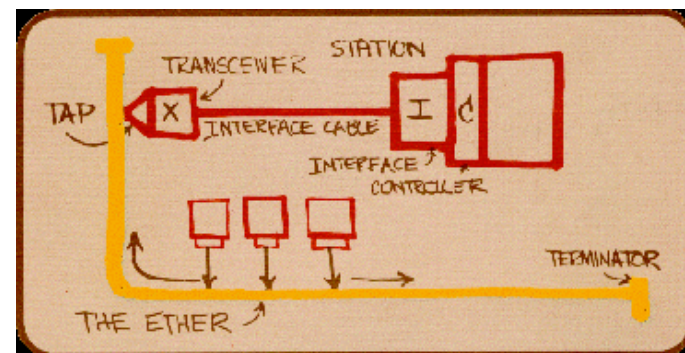
Ethernet és TCP/IP

- Mindenkinnek van felhasználói szintű tapasztalata
- Ethernet (IEEE 802.3)
 - Egyeduralkodó technológia
 - RJ45-os csatlakozó és a CAT5-os kábel az összekapcsolhatóság szimbóluma
 - Ezen kívül optikai közeget is támogat
 - A vezetékes telekommunikációs rendszerekben is majdnem egyeduralkodó technológia
 - A beágyazott területen is gyorsan terjed
 - Több milliárd eladott „port”
- TCP/IP protokollcsalád
 - Az Internet alapja
 - Több milliárd végpont
- Fontos: A WI-FI (IEEE 802.11) nem Ethernet!



Az Ethernet fejlődése

- 1973–1975: Xerox PARC, 3 Mb/s, 8 bites címek
 - Ether : A fényt továbbító közeg (korábbi fizikai elméletekben)
- 1979: 3COM-ot megalapítja Robert Metcalfe (XEROX)
- 1980: DIX (Digital/Intel/Xerox), 10 Mb/s, 48 bites címek, Metcalfe vezetésével, alapértelmezett keretformátum azóta nem változik
 - 1981-: DIX/Ethernet 3COM hálózati eszközök
- 1980: IEEE 802 hálózati szabványosítási csoport elindul
- 1982 vége: IEEE 802.3 CSMA/CD szabvány
 - Számos versenytárs (azóta feledésbe merültek)
- 1995: 100 Mb/s sebességű szabványok
- 1997: Full-duplex működés szabványosan megjelenik
- 1999: 1000 Mb/s sebességű szabványok
 - A CSMA/CD megtalálható a szabványban, de az eszközök nem használják (full duplex kommunikáció)
- 2002- : 10Gb/s szabványok, CSMA/CD nem támogatott
 - Azóta is újabb fizikai rétegek jelennek meg 10 Gb/s-en
- 2003: Power over Ethernet (15.4 W max.)
- 2006: 10GBase-T, 10 Gb/s CAT6a csavart érpár vagy annál jobb felett
- 2009: Power over Ethernet (25W max.)



IEEE 802 LAN/MAN Standards Committee

- 1980-tól IEEE 802 LAN/MAN Standards Committee (LMSC)
- Fontosabb szabványok
 - IEEE 802.1 Bridging (networking) and Network Management
 - IEEE 802.3 CSMA/CD Ethernet
 - IEEE 802.3-2008 jelenleg, és számos munkacsoport dolgozik új kiegészítéseken
 - Kb. 40 Mbyte, és több mint 3000 oldal
 - IEEE 802.11 a/b/g/n Wireless LAN & Mesh (Wi-Fi certification)
 - IEEE 802.15 Wireless PAN (Bluetooth és ZigBee)
 - IEEE 802.16 Broadband Wireless Access (WiMAX certification)
- Get IEEE 802®, a szabványok ingyenesen hozzáférhetőek
 - <http://standards.ieee.org/getieee802/index.html>
 - Érdemes megnézni őket...

IEEE 802.3 CSMA/CD?

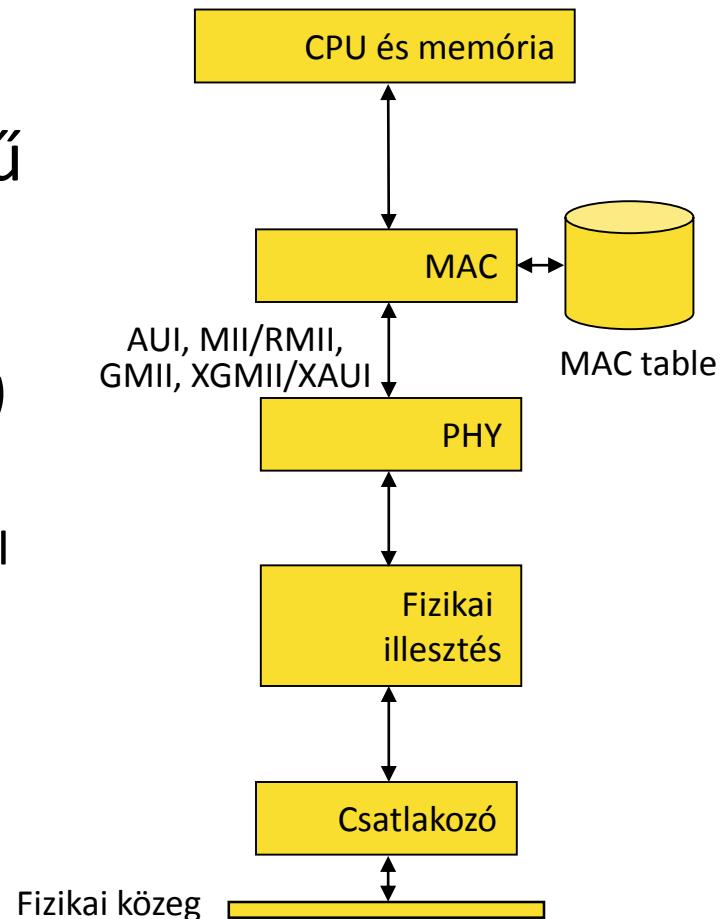
- Eredetileg az Ethernet üzenetszórásos közeget feltételezett
 - CSMA/CD: Carrier sense multiple access with collision detection közeg-hozzáférési protokoll (versenyzéses)
 - Az adni kívánó állomás az adás előtt ellenőrzi a közeg használatát:
 - Ha használt a közeg megvárja a közeg felszabadulását
 - Ha a közeg szabad, adni kezd
 - Mivel a hálózatban az információ terjedési ideje nem nulla, lehetséges, hogy több állomás kezd adni közel egy időben
 - Ezt az állomások érzékelik (ütközés), és befejezik az adást
 - Ezek után véletlen (binary exponential backoff) ideig várnak, majd ismét adni próbálnak (ellenőrizve a közeg használatát)
 - Egyébként újra ütköznének...
 - Hálózatméret és forgalom függő, nem determinisztikus működés
- Az optikai kábel és a csavart érpár full-duplex kommunikációt tesz lehetővé (külön adás és vétel ág)
 - 1997: Full-duplex működés engedélyezése a szabványban
 - 2002: CSMA/CD teljes elvetése a 10 Gb/s Ethernet szabványban
- Nem foglalkozunk a CSMA/CD-vel (gyakorlatban nem használják)

IEEE 802.3 változása 28 év alatt

- Nem változott (csak kompatibilisen bővült)
 - Fizikai rétegek
 - Keretformátum
 - Címek és kiosztásuknak a módja
 - Ethernet hidak (bridge) működése
- Elavult és ezért eltűnt
 - CSMA/CD
 - Ethernet ismétlők és HUB-ok
 - Néhány fizikai réteg technológia
- Rétegszerkezet változik, ahogy a sebesség nő...

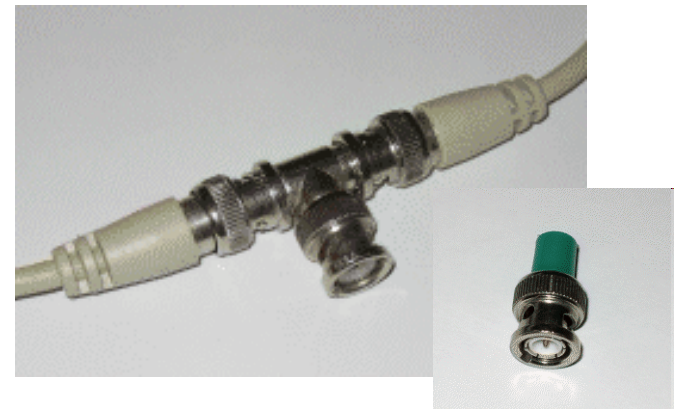
Ethernet rétegek

- A rétegszerkezet változik a sebesség növekedésével
- Visszafelé mindig kompatibilis
 - 10/100/1000 Mb/s sebességű 1000Base-T NIC-ek
 - Jellegzetes felépítés
 - Media Access Controller (MAC)
 - Szabványos interfész
 - AUI, MII/RMII, GMII, XGMII/XAUI
 - Physical Layer Device (PHY)
 - Fizikai illesztés (impulzus transzformátor vagy optoelektronika)
 - Csatlakozó
 - Fizikai közeg



Ethernet fizikai közegek (korai technológiák, busz topológia)

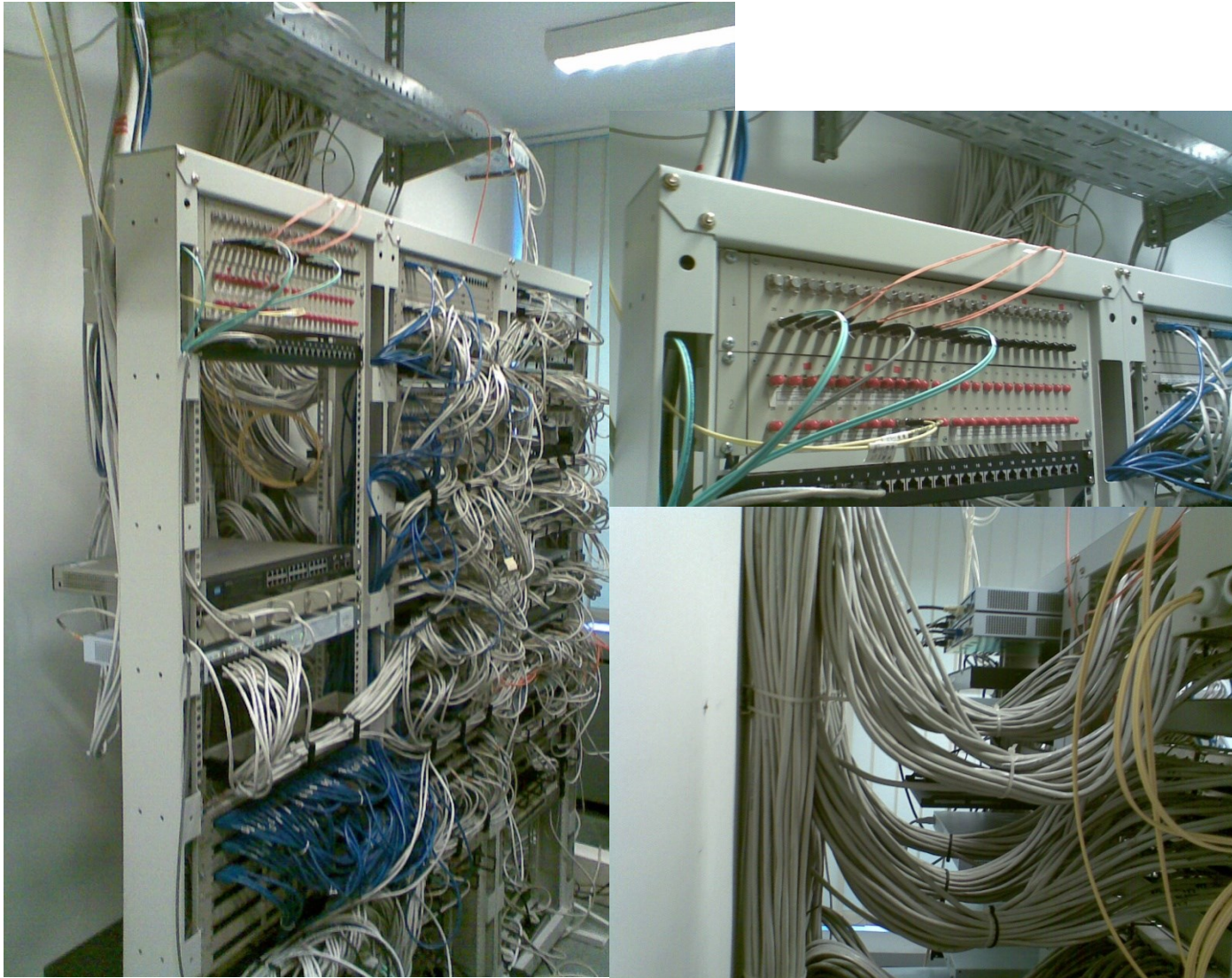
- Vastag koax 10Base-5
 - Első széles körben elterjedt technológia (1980-as évek)
 - Drága, nehezen telepíthető
- Vékony koax 10Base-2
 - Olcsó, könnyen telepíthető
 - 50 Ohm koax kábel, max. 185m kábelhossz
 - Meghibásodási gyakoriság riasztóan magas
 - Lezáró ellenállással rohangáló rendszergazdák és a dolgozni nem tudó munkatársak látványa
 - Karbantarthatatlan
 - Növekszik a kábelhossz
 - Egy idő után bizonyos állomások nem tudnak egymással kommunikálni



CATx kábelezés

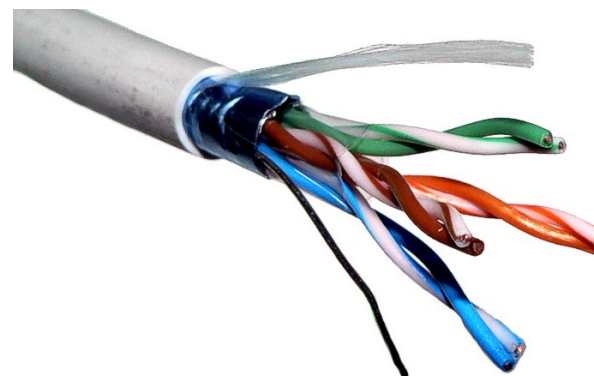
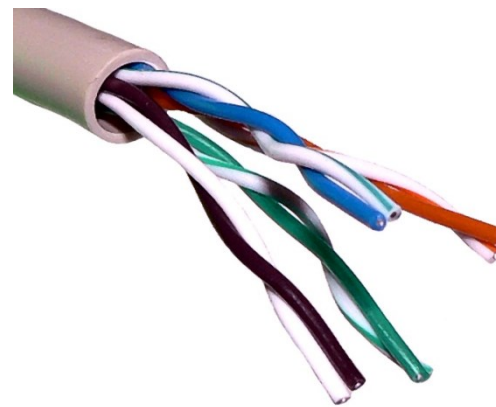
- 10Base-2 megbízhatatlan és menedzselhetetlen
- 1990-es évek elején elterjedt a CAT3 strukturált kábelezési rendszer az irodaházakban
 - TIA/EIA-568-A.1-1991 szabvány
 - Csillag topológia
 - Egy eszköz meghibásodása nem okozza feltétlenül a többi működésképtelenségét
 - Kábelrendező szoba aktív eszközökkel (hozzáférés kontrollálható)
 - Fali csatlakozók a munkahelyeknél (kulturált kialakítás)
 - 8P8C (RJ45 hibásan) csatlakozók használata (telefon, ISDN, stb. beköthető rá, a csatlakozók kompatibilisek, pl. RJ11 a telefon esetén)
 - 4 csavart érpár érpáronként 16MHz sávszélességgel
- A számítógépeknek is képesnek kell lenniük ezt használni
 - Galvanikus leválasztás impulzus transzformátorral
- Fejlődés azóta:
 - CAT5, CAT5e (100 MHz)
 - 52 pF/m : 100 m kábel 5.2 nF kapacitású, ezen kell átvinni a jelet...
 - CAT6 (250 MHz)
 - CAT6a (500 MHz), 10GBase-T minimum követelmény
 - CAT7, CAT7a (600/1000 MHz, 8P8C vagy GG45 csatlakozók)

Strukturált kábelezés?



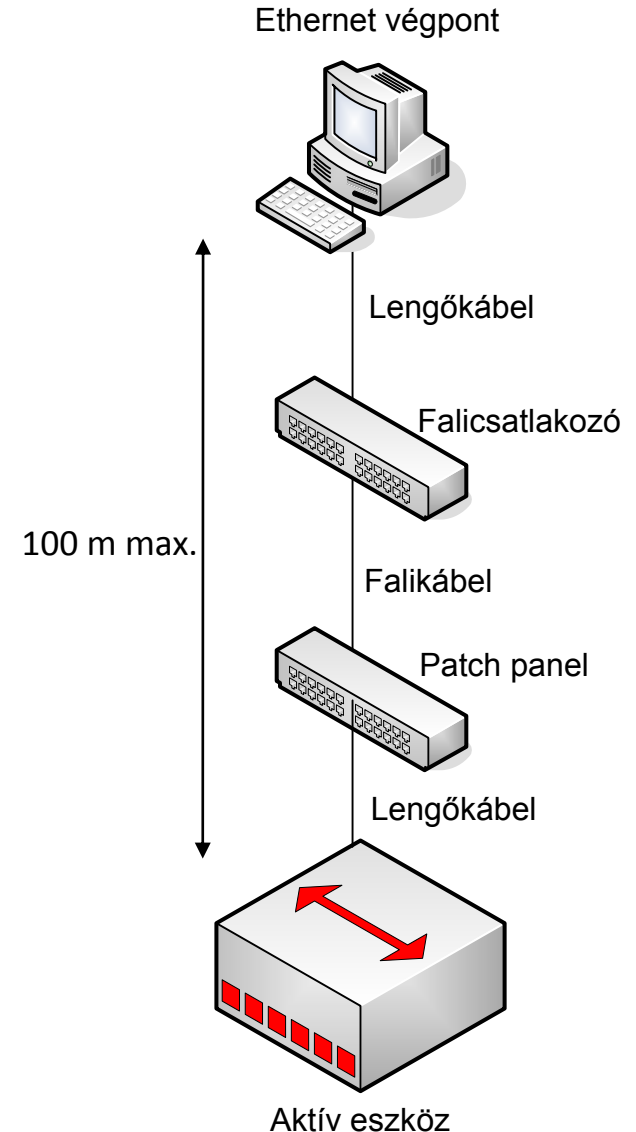
Csavart érpár

- 4 érpár egy kábelben
- 100 Ohm hullám impedancia
- Unshielded twisted pair (UTP)
- Foiled Twisted Pair (FTP)
- Screened Fully shielded Twisted Pair (S/FTP)
- Az RJ45-os csatlakozónak is van árnyékolt változata, amit az FTP és S/FTP kábelnél alkalmazni kell



10Base-T

- Minimum CAT3 minőségű csavart érpárt használó strukturált kábelezés
 - Maximum 100m kábelhossz
 - 90m falikábel
 - 10m lengőkábel (2 db-ban)
 - 2 érpárt használ a kábelben lévő négyből
- RJ45-os csatlakozók
- Csillag vagy fa topológia
- Aktív hálózati eszközök a csillag vagy fa elágazásaiban
- Full-duplex kommunikáció lehetséges
- Napjainkban ritka a kizárólagos használata új eszközökben
 - Microchip ENC28J60 MAC+PHY és PIC18F97J60 MCU+MAC+PHY van gyártásban (csak 10Base-T-t tudnak)



100Base-TX

- A 10Base-T alapján
 - Más, hatékonyabb kódolás, 125 Mbaud
- CAT5 UTP kábelezés + RJ45 csatlakozó
 - Nagyobb sávszélességű kábel
 - 16 MHz helyett 100 MHz
 - Kisebb érpárok közötti áthallás engedélyezett
 - A 4 érpár különbözően csavarva az áthallás csökkentésére
 - Precízebb mechanikai kivitel
 - Szimmetria az érpár kábeleik között pontosabban tartandó
 - 100 Ohm-os hullám impedancia pontosabb betartása
 - Opcionális árnyékolás (FTP, S/FTP)
- Automatikus opció egyeztetés
 - 10/100, half-duplex/full-duplex, flow-control

MDI vagy MDIX (crossover)

- Medium dependent interface (MDI)
 - Lényegében a csatlakozó és annak kiosztása
 - A 10Base-T és 100Base-TX
 - Adó érpár
 - Vevő érpár
 - Az adó és a vevő érpárat a kommunikációhoz valahol meg kell cserélni (mint az RS232-nél)
- Ethernet megoldása a problémára az egyenes bekötésű kábel
 - MDI : magasabb rétegben dolgozó eszközök
 - Számítógép, router, stb.
 - MDIX (MDI-X) vagy auto-MDI/MDIX: ethernet aktív eszközök
 - Régi eszközökön esetleg kapcsoló „MDI/MDIX” vagy „uplink/downlink” felirattal
 - Crossover kábel
 - Piros kábel vagy törésgátló, X jelzés valahol a kábelén
- 1000Base-T esetén nincs szükség rá
 - Minden érpárat egy időben használ adásra és vételre
 - 10/100 módban is automatikusan megtalálja a érpárokat

100Base-T

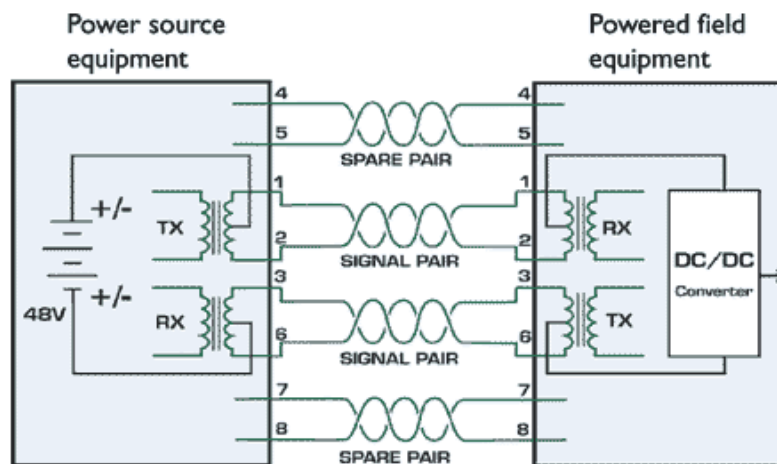
- CAT5e kábelezés szükséges
 - A kábelek közötti távoli végből származó áthallás (far end crosstalk, FEXT) pontosabb specifikálását jelenti
- 125 Mbaud, akár csak a 100Base-TX
 - PAM-5 moduláció a csatornán
 - 4 érpáron egy időben adás és vétel
 - Csatorna kompenzálás, near end echo cancellation és near end crosstalk (NEXT) cancellation
 - Bonyolult digitális jelfeldolgozás szükséges
 - Fogyasztás ezért emelkedik (hűtőborda volt az első NIC-eken)
 - A kábel átviteli tulajdonságait mérni kell
 - Tanulási fázis bekapcsoláskor
 - A hibaaarány határérték fölé növekedése esetén újratanulás
 - Valós idejű működés hogyan biztosítható?
- Visszafelé kompatibilis a 10Base-T és 100Base-TX szabvánnyal
 - 10/100/1000

10GBase-T

- IEEE 802.3an-2006 szabványban jelent meg
- CAT6a kábelezés minimum
 - Jobb minőségű csatlakozók
 - Speciális kábelek
 - Meglévő épületek újrakábelezése szükséges
- Nagy fogyasztás
 - 2010: 6-8 W portonként csak a PHY fogyasztása 2010-ben nagy sorozatban szállított eszközök esetén (nagy hűtőborda)
 - 2012: 2-3 W portonként az új PHY-k fogyasztása (Energy-Efficient Ethernet nélkül)
- Nagy késleltetés (nagyobb, mint a 1000Base-T késleltetése) a komplex jelfeldolgozás miatt
- A végberendezés oldalon ez a sávszélesség nem szükséges
 - Esetleg sokprocesszoros szervereknél lehet rá szükség igazán
 - Szervertermek, gerinchálózatok használják
- Az optikai rendszerek ára csökken, nem biztos, hogy ez a szabvány a korábbi mértékben fog terjedni
 - Újrakábelezés esetén az optikai kábelek beszerelése jobb döntés lehet
 - Ekkora sebességnél a beágyazott rendszerekben az optikai fizikai rétegek alkalmazása előnyösebb lehet

Power over Ethernet (intro)

- Az Ethernet kábelben az információ mellett tápfeszültség szolgáltatása is
 - Közös módusú jelként akár a kommunikációra is használt érpárokon
 - egyik érpár a pozitív, a másik a negatív tápfesz
 - Power sourcing device
 - Fel kell ismernie a tápfeszültséget kérő eszközt
 - 15.4 W DC (minimum 44V , 350 mA)
 - Powered device
 - 12.95 W a kábel és egyéb veszteségek miatt
 - Opció egyeztetés (szükséges teljesítmény)
 - Galvanikus leválasztás a tápfeszültségre is
- IEEE 802.3at (2009/09/11) 25W-ra növeli a teljesítményt



Optikai kábel

- Üveg optikai szál
 - Optikai hullám vezető
 - Alapvetően két fő típusa különböztethető meg
 - Multimodusú (50/100 μm vagy 62.5/125 μm mag átmérővel)
 - Olcsó
 - Kedvezőtlenebb tulajdonságok
 - » Kis sávszélesség távolság szorzat (300-400 Mb/s 1 kilométerre)
 - Elsősorban a diszperzió miatt (a jel gyorsan torzul)
 - » A modulusok különböző sebességgel haladnak...
 - Mono vagy single modulusú
 - Nagyon sok típus létezik, különböző árakon és tulajdonságokkal
 - Épületen belül és kontinensek között is
 - Léteznek optikai erősítők és egyéb feldolgozó eszközök is
 - 15.5 Tb/s 7000 km-re az aktuális rekord (2009/09/27)
 - » Bell Labs (elméleti határok feszegetése)
 - A gyakorlatban használt rendszerekben a 400 Gb/s a maximum

Optikai kábelek, előnyök és hátrányok

■ Előnyök

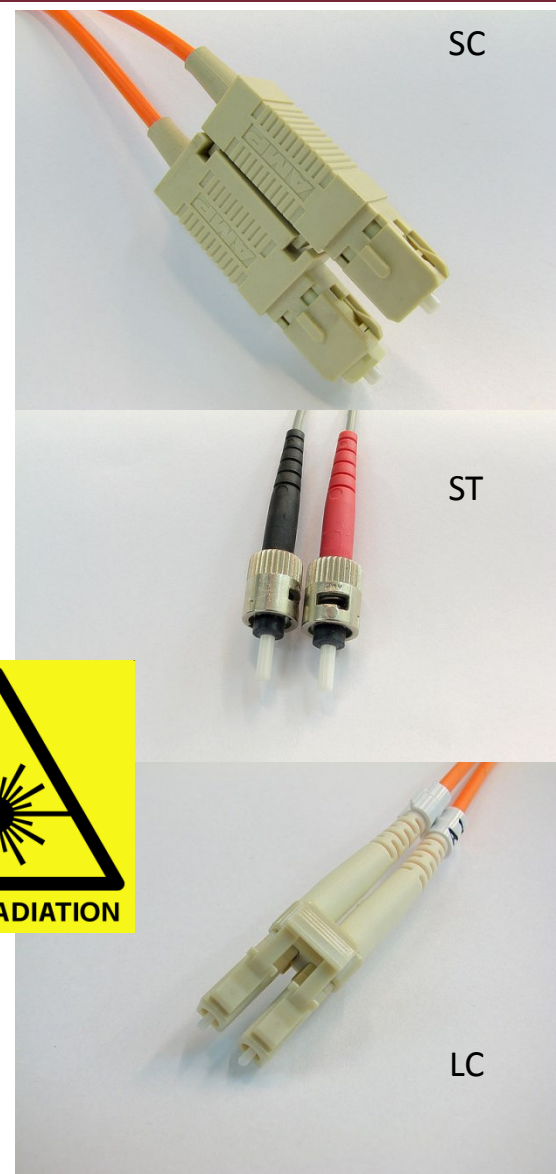
- Galvanikus leválasztás megoldott
- EMC immunitás
- Nagy sávszélesség

■ Hátrányok

- Magas ár
- Speciális eszközök szükségesek a szereléséhez, méréséhez, vizsgálatához
- Nagyobb fogyasztás és méretek (még 1 Gb/s sebességnél igaz, de 10 Gb/s-nél már nem biztos)
- A csatlakozó és kábeltípusok nagy száma
- Ethernet esetén az eltérő fizikai rétegek nagy száma
 - 1 Gb/s és 10 Gb/s esetén is 7 eltérő szabványos fizikai réteg került kidolgozásra, és folyamatosan újak jelennek meg

Optikai kábelezés, optikai csatlakozók

- Strukturált kábelezési rendszer
- Optikai csatlakozók
 - Különböző típusok
 - SC, LC eszközökön elsősorban
 - ST fali csatlakozókon és patch paneleken (eszközön soha)
 - MT-RJ (RJ45 alapján) rövid ideig elérhető volt, már ritka
- Fali (sok ér) és lengőkábelek (érpár)
- A víz bejutása a kábelt tönkreteszi
 - 850, 1300 and 1550 nm hullámhosszokat is ezért választották (víz szennyezés)
- Porvédők felhelyezése ajánlott, ha kihúzzuk a kábelt
- Lézerveszély (kábelvégbe tilos belenézni)



100Base-FX és 100Base-SX

■ 100Base-FX

- Napjainkban még gyártják
 - Elsősorban az ipari irányítástechnikában és mérőrendszerekben használják
- CSMA/CD esetén 400m max. kábelhossz
- Full-duplex működés esetén 2 km
- 1300 nm-es infravörös lézer és multimódusú kábel
- SC csatlakozó az aktív eszközökön
 - Adás és vétel ág megkülönböztetése nehéz

■ 100Base-SX

- 850 nm-es infravörös LED
- Max. 300m-es kábelhossz
- Nem terjedt el

1000Base-SX és 1000Base-LX

- Az aktív eszközök esetén modulok
 - Nem tudják előre az igényeket
 - GBIC SC csatlakozóval (elavult)
 - SFP LC csatlakozóval
- 1000Base-SX
 - 770 - 860 nm lézer
 - 220 m 62.5/125 μm -es, 550m 50/100 μm -es multimódusú optikai kábel
 - Egyeduralkodó épületen belül
 - 30-40 eFt egy minőségi modul (olcsó)
- 1000Base-LX
 - 1270 - 1355 nm lézer
 - 5 km 10 μm -es monomódusú kábel
 - Épületek között „szabad/sötét kábel”
 - Drága (kb. 100-150 eFt egy modul)

GBIC



SFP



Az Ethernet jövője, Sebesség...

- 10GBase-XX
 - Szabványos megoldások (12 különböző fizikai réteg)
 - Optikai kábel
 - Hullámhossz multiplexelés is alkalmazásra kerül
 - Single mode kábellel korábbi távolságok
 - Multimode kábellel rövid távolságok
 - Csavart érpár 10GBase-T
 - Cat5e kábelezés nem használható, új kábelezés és patch kábelek szükségesek
 - Nagy fogyasztás
 - Twin-ax kábel kábelrendezőn belül, rövid távolságra (nem terjedt el)
 - Csak full-duplex működés, a CSMA/CD nem támogatott
- 40 Gb/s és 100 Gb/s fejlesztések
 - Az első szabványok megjelentek (IEEE 802.3ba)
 - Jelentős energiafogyasztás
 - Elsősorban a gerinchálózatban várható az elterjedése
 - Nem várható a beágyazott területen a gyors megjelenése
 - Elsősorban a nagy adatközpontok igénylik (Google, YouTube, stb.)
 - Nekik már kevés a 10 Gb/s...

Az Ethernet jövője, energia fogyasztás

- Energy Efficient Ethernet, IEEE P802.3az (2010 október)
 - Az Internet kommunikációs részének fogyasztása kb. 6 TWh évente
 - Ethernet portok fogyasztása független a forgalomtól (idle szimbólumot küld forgalom nélkül)
 - 3 TWh lenne megtakarítható a becslések szerint
 - Low Power Idle küldése periodikusan, ha nincs küldendő információ
 - Fogyasztás csökkenés az adó és a vevő oldalon is
 - A link karbantartása szükséges (refresh)
 - Beágyazott rendszerekben is nagyon fontos lesz
- Egyéb nem szabványos, de a szabvánnyal nem ellentétes megoldások
 - A kábelhossz figyelembe vétele az adó oldalon
 - Automatikus sebesség váltás a kihasználtság függvényében (késleltetés)

Ethernet jövője (fizikai szint), autóipar 1.

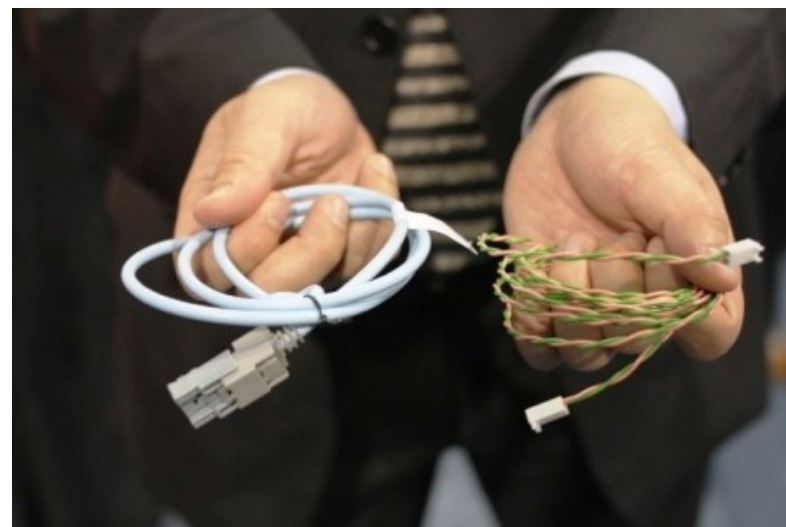
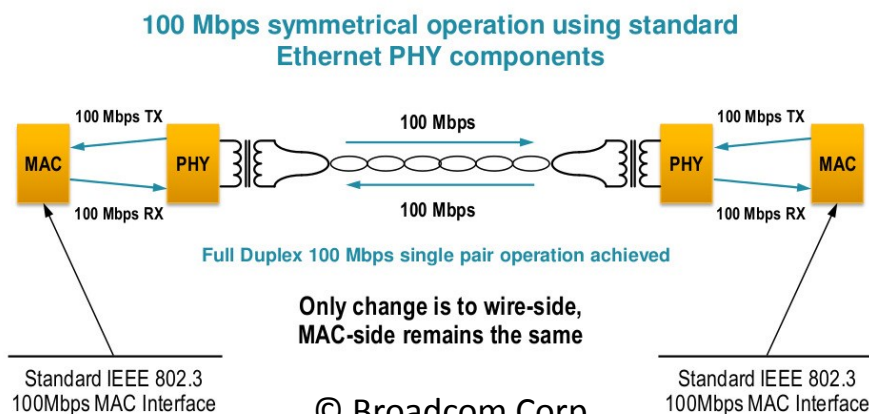
■ BroadR-Reach szabvány (Broadcom, NXP, Freescale)

○ Mi a gond az Ethernetnel az autóiparban?

- 4 érpár, mechanikailag gyenge csatlakozók
- 100 m-es távolságra nincs szükség
- Nagy felhasznált sávszélesség a frekvenciatartományban
- Nem kompatibilis az autóipar EMC/EMI szabványokkal

○ Mi a megoldás?

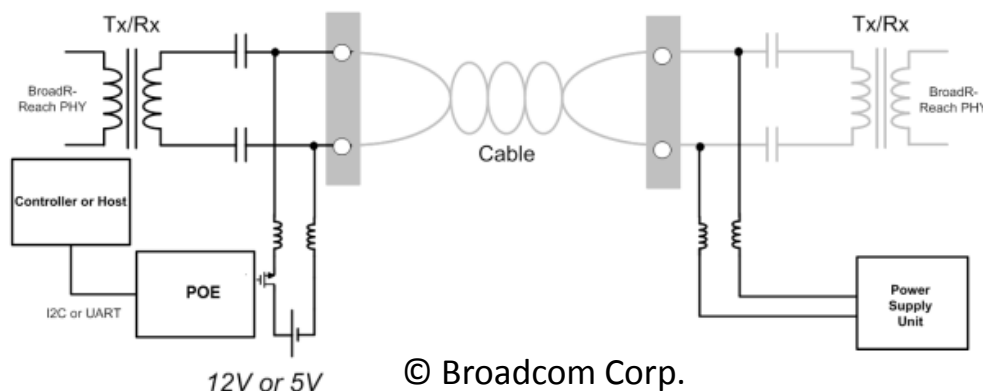
- 1 érpár, 15 m max. kábelhossz, 100 Mbps (1000Base-T szerű működés)
- Olcsó, autóipari csatlakozók
- Szükséges sávszélesség: 33 MHz
- EMI/EMC kompatibilis



Ethernet jövője (fizikai szint), autóipar 2.

■ Power over BroadR-Reach Ethernet

- Ne kelljen extra vezeték a táp miatt, de a standard PoE túl bonyolult egy autóhoz...



■ IEEE 802.3bp RTPGE (Reduced Twisted Pair Gigabit Ethernet)

- 1 Gbps egy érpáron
- 15m-es maximális távolság (tipikus 3.5m) az autóiparban, de 40m-re kiterjeszthető (busz és teherautó, vasút, repülőipar, hajózás)
- Egyelőre még gondok vannak ez EMC/EMI-vel
- Cél: Autó gerinchálózata...
- Hogy szükség lesz-e rá, az jó kérdés, de dolgoznak rajta...

- Archaikus okokból (CSMA/CD-ből levezethető) a kerethossz:
 - 64-1518 byte (octet-ben van megadva a szabványban)
 - Az Ethernet az LSB bitet/szimbólumot (0. bit) küldi először!
 - Nem tartozik a keretbe (Ethernet bit sorrend)
 - 7 byte Preamble (vevő órájának szinkronizálása) 7x10101010
 - 1 byte Start of Frame (10101011)
 - Keret
 - 6 byte Cél Ethernet címe
 - 6 byte Forrás Ethernet címe
 - 2 byte Típus/Hossz mező
 - 46-1500 byte hosszú adattartalom
 - A 46 byte-nál rövidebb információt fel kell tölteni (de miből?)
 - 4 byte hosszú CRC
 - Nem tartozik a keretbe
 - 12 byte Interframe gap

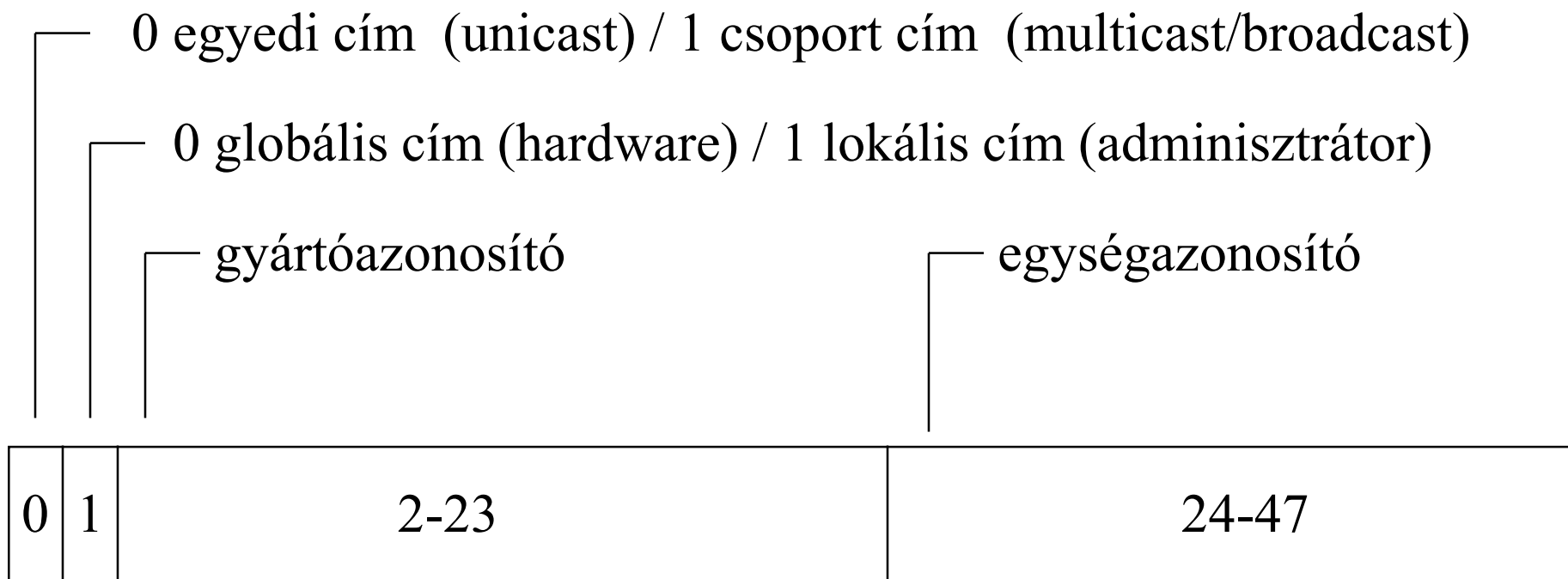
Ethernet keret formátum Ábra

Preamble	SF	Célcím	Forráscím	T/L	Adat	CS
----------	----	--------	-----------	-----	------	----

Megjelölés	Név	Hossz byte-ben (ocetet-ben)
Preamble	Előtag	7
SF	Start of frame	1
Célcím	Célállomás címe	6
Forráscím	Küldő állomás címe	6
T/L	Típus (Ethernet) Adatmező Hossz (IEEE 802.3)	2
Adat	Adat (Hasznos teher)	46-1500
CS	CRC (hibadetektálás)	4

IEEE 802.3 / Ethernet keretformátum

Ethernet címek



6 octet → 48bit

- 48 bites cím, 46 bit használható
 - 2^{46} különböző globális egyedi cím (elméletileg)
 - 2^{22} (4194304) gyártó azonosító (Organizationally Unique Identifier, OUI)
 - Az IEEE osztja ki őket (\$1,650)
 - Igazolni kell, hogy 95%-ban elhasználta a cég a korábbi tartományát
 - <http://standards.ieee.org/regauth/index.html>
 - 2^{24} (16777216) egyedi azonosító egy gyártó azonosítón belül
 - Egyediségükért a gyártó felel
 - Opcionálisan: Individual Address Block (IAB) 12 bit
 - » Akinek 4096 cím elég (\$550)
 - » Egy speciális OUI tartományban, relatív drága
 - » 00-50-C2 gyártóazonosító alatt
 - Lokális egyedi cím használata ritka
 - Fejlesztés alatt álló HW esetén
 - Bizonyos archaikus hálózati protokollok (DECnet) ezeket használva alakítottak ki hierarchikus címezést

OUI visszakeresése, pl. NI

- 00-80-2F (hex)
 - NATIONAL INSTRUMENTS CORP. 11500 North Mopac Expressway AUSTIN TX 78759-3504 UNITED STATES
- <http://standards.ieee.org/cgi-bin/ouisearch?abcdef>
 - on-line keresés
 - ahol „abcdef” a gyártóazonosító
- kb. 14650 bejegyzés a nyilvános OUI adatbázisban 2011/02/13-án
- kb. 15260 bejegyzés 2012/09/14-én
 - <http://standards.ieee.org/regauth/oui/oui.txt>
 - Van privát is, de az drágább, és évente meg kell újítani
 - Pl. 13790 volt 2009 őszén (jó három éve)
 - 2^{22} (4194304) gyártóazonosító van, nem fog hamar elfogyni...
 - Van már 64 bites MAC cím is (EUI-64), pl. ZigBee azt használja

Típus vagy hossz mező?

- Hossz értelmezés
 - $0x0800$ alatt, IEEE 802.3 szerinti beágyazás
 - Az Ethernet keretet egy Logical Link Control fejrész követi
 - Ritkán használják
- Típus értelmezés ($0x0800$ és felette, 2048)
 - Az IEEE osztja ki (mint az OUI-t), Ethernet beágyazás
 - <http://standards.ieee.org/develop/regauth/ethertype/eth.txt>
 - Hogyan tudja meg a vevő a keret végét?
 - IDLE szimbólum jön (magától kiderül)
 - Megadja, hogy milyen felsőbb réteg kezeli a csomagot
 - EtherType Field (\$2,500.00 egy regisztrálása)
 - $0x0800$ Internet Protocol, Version 4
 - $0x0806$ Address Resolution Protocol
 - $0x8100$ VLAN-tagged frame (IEEE 802.1Q)
 - $0x86DD$ Internet Protocol, Version 6
 - $0x88F7$ Precision Time Protocol (IEEE 1588)

IEEE 802.1q

- IEEE 802.1q VLAN Tagging szabvány
- Virtuális hálózat és prioritás megadására
 - VLAN logikailag eltérő LAN ugyanazon a fizikai hálózaton
 - Különböző VLAN-ban lévő állomások nem képesek egymással az adatkapcsolati rétegben kommunikálni
 - Prioritás az Ethernet eszközökben és a számítógépekben történő feldolgozás sorrendjének befolyásolására (8 szint)
- Típus `0x8100` VLAN-tagged frame (IEEE 802.1Q) Ethertype
- 16 bites mező a VLAN és prioritás megadására
 - 3 bit Prioritás (8 szint)
 - 1 bit Canonical Format Indicator (0 Ethernet esetén)
 - 12 bit VLAN tag (VLAN tag = 0, nincs VLAN-ban)

Keretmérettel mi lesz 802.1q-nál?

- A maximális keret méret 1522 byte-ra nő
 - Előzetes vizsgálatok szerint ez nem okoz problémát a szabvány előtti eszközökben
 - Az Ethernet eszközök továbbítják (nem tartják hibásnak)
 - Hibajelzés 2048 (2^{11}) byte vagy 1536 ($48 \cdot 32$) byte hosszú keretek felett, de egyes eszközök 8192 byte hosszhatárig működtek
 - Néha jól jön a „hanyag” implementáció

IEEE 802.1q Ethernet keret formátuma

Preamble	SF	Célcím	Forráscím	Tag	T/L	Adat	CS
----------	----	--------	-----------	-----	-----	------	----

Célcím	Célcím	6
Forráscím	Forráscím	6
Tag	IEEE802.1Q mező, 81h-00h-XXh-XXh, ahol az XXh-XXh mező (16 bit) hordozza a VLAN azonosítót, és a szolgáltatási osztályt, felépítése: 3 bit: Priority Code Point (PCP) 1 bit: Canonical Format Indicator (CFI), 0 for Ethernet 12 bit: VLAN Identifier (VID), 0 not VLAN, 0xFFF reserved	4
T/L	Típus (Ethernet) Adatmező Hossz (IEEE 802.3)	2
Adat	Adat (Hasznos teher)	46-1500
CS	CRC (hibadetektálás)	4

IEEE 802.3 / Ethernet keretformátum IEEE802.1Q alkalmazása esetén

Jumbo keretek

- 1518 byte hosszú maximális keretméret a korai rendszerek korlátai miatt alakult ki (válaszidő elfogadható értéken tartása, 1.2 ms)
- Nagysebességű hálózatokon ez nem kedvező
 - 1 Gb/s környékén (1990-es évek közepe) komoly teljesítmény problémák merültek fel a végeszközökben (Szerverek és nagyteljesítményű munkaállomások)
 - Keretképzés és fogadás overhead-je elviselhetetlen volt
 - A szűk keresztmetszet a NIC-HW-OS interfész
- Jumbo (óriás) keretek
 - 9000 byte maximális hossz, kisebb terhelés a NIC-HW-OS-nél
 - Nem része az IEEE 802.3 szabványnak (elutasították a szabványosítását)
 - Egyes eszközök mégis támogatják
 - Egyre kevesebb
- Napjainkban nem sok előnye van
 - Architektúrális változások: Megváltozott a NIC-HW-OS interfész működése (optimalizálták)
 - Sokkal gyorsabbak a mai számítógépek

Ethernet ismétlők: A veszélyeztetett állatfaj

- Fizikai rétegben működő eszköz
 - Ismétlő, repeater, HUB, több portos eszköz
- Feladata:
 - Jel regenerálása
 - Vétel (egy porton) - digitális bit/szimbólum stream – adás (minden porton)
 - A CSMA/CD működése során létrejövő ütközéseket is továbbítja
 - A hibás kábel szegmenseket érzékeli, és leválasztja a többiről, majd javítás után visszakapcsolja
- Csak azonos sebességű eszközöket tud összekapcsolni
 - Dual speed HUB tartalmaz egy Ethernet hidat is!
- Ma már nem gyártanak ilyen eszközöket és ritkán is használják őket
- Speciális esetben hasznos, ha van egy tartalékban:
 - Ismétlők, vagyis lehetővé teszik a hálózat megfigyelését
 - Drasztikusan beavatkoznak a hálózat működésébe (hátrány)

Ethernet hidak (bridge)

- Adatkapcsolati rétegben működő eszköz
 - IEEE 802.1D szerinti transzparens híd
 - Megtanulja a hálózat topológiáját (fordított tanulás)
 - Automatikusan megtanulja, hogy egy adott állomás melyik kimeneti porton található (egress port)
 - Ezek alapján dönt a beérkező keretek továbbításáról
- Ethernet keretek szintjén működik
- „Tárold és továbbítsd” működés (régen volt más elv is)
 - Veszi a teljes keretet egy bemeneten (ingress port)
 - Döntést hoz
 - Ha szükséges, továbbítja a megfelelő kimenetre vagy kimenetekre (egress port)
 - A kimeneten van egy vagy több (prioritás) várakozási sor
- Eltérő sebességű eszközöket is képes összekapcsolni
 - Folyamszabályzás szükséges lehet
 - Egy gyors adó túlterhelheti a lassú vevőt (PAUSE frame)

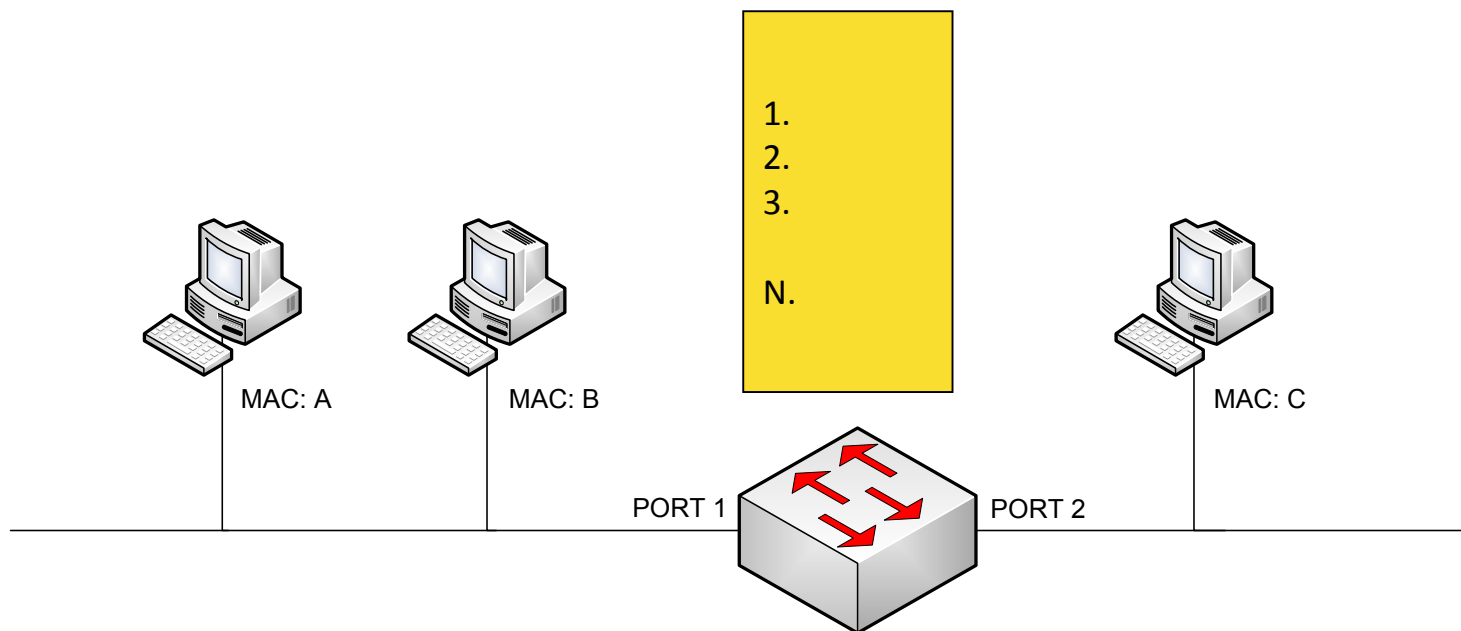
Fordított tanulás

- A beérkező keretek forrás címek alapján tanul
 - Feltételezés: Az adott porton beérkező keretekben megadott forráscímű gép az adott porton található
 - Az információ alapján egy „forwarding database” jön létre a hídban
 - Az adatbázis mérete korlátozott
 - Néhány száz bejegyzés SOHO eszközökben
 - Néhány tízezer vállalati eszközökben
 - Néhány százezer gerinchálózati eszközökben
 - Az adatbázis elemeit öregíti (ageing) a rendszer
 - 30 perc az alapértelmezett felejtési idő
 - A tábla telítődése esetén a legrégebben használt bejegyzés íródik felül
 - Változó címmel küldő hibás eszköz komoly problémákat tud okozni (ritka meghibásodás)

Kerettovábbítás

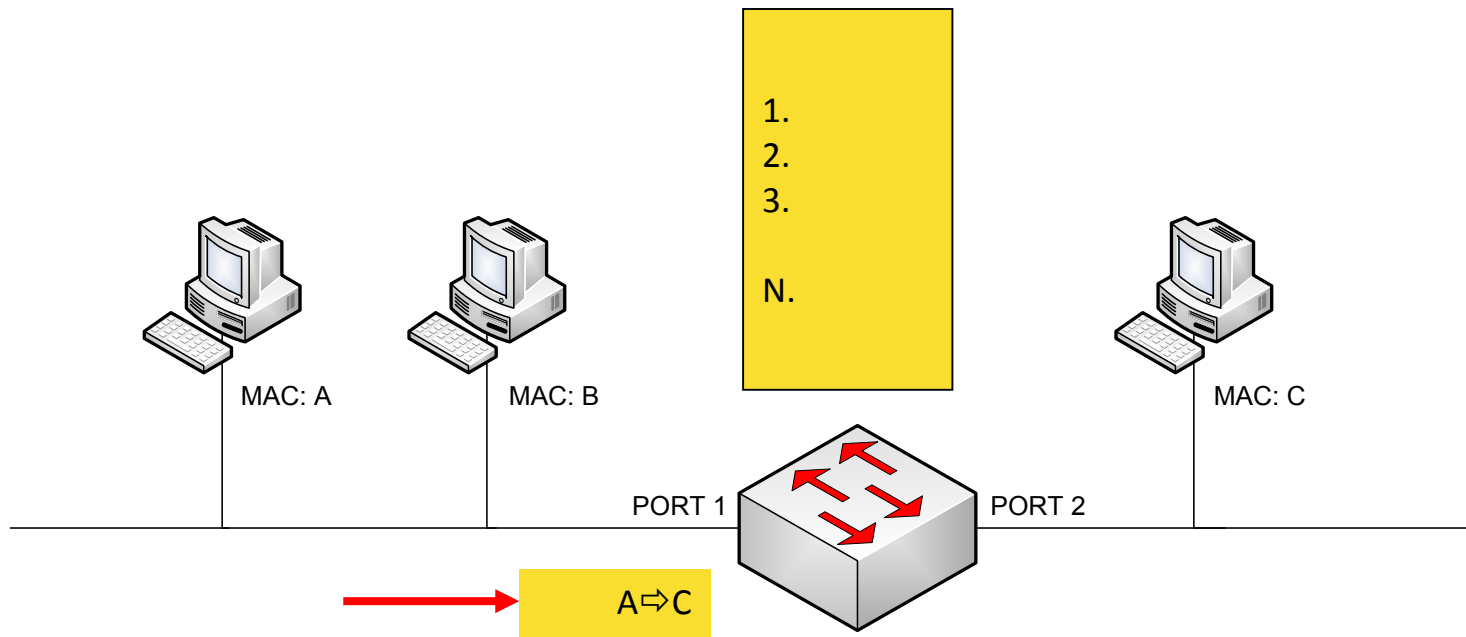
- A „forwarding database” alapján:
 - Keret beérkezik az ingress porton
 - Adatbázis frissítése a forráscím alapján
 - Továbbítási döntés a célcím alapján:
 - Ismert unicast cím: Az adott unicast címhez tartozó portra kiküldésre kerül, kivéve ha az az ingress port (szűrés)
 - Ismeretlen unicast cím: Minden portra kiküldésre kerül kivéve az ingress port (elárasztás), hátha válaszol rá a kérdéses gép (és akkor meg lehet tanulni)
 - Multicast cím: Alapesetben elárasztás, multicast tanulás esetén minden az adott multicast címre regisztrált portra kiküldésre kerül kivéve az ingress port
 - Broadcast cím: Elárasztás
- VLAN-ok konfigurálása esetén a kerettovábbítás összetettebb:
 - VLAN-ok azonossága is feltétel

Kerettovábbítás példával 1.



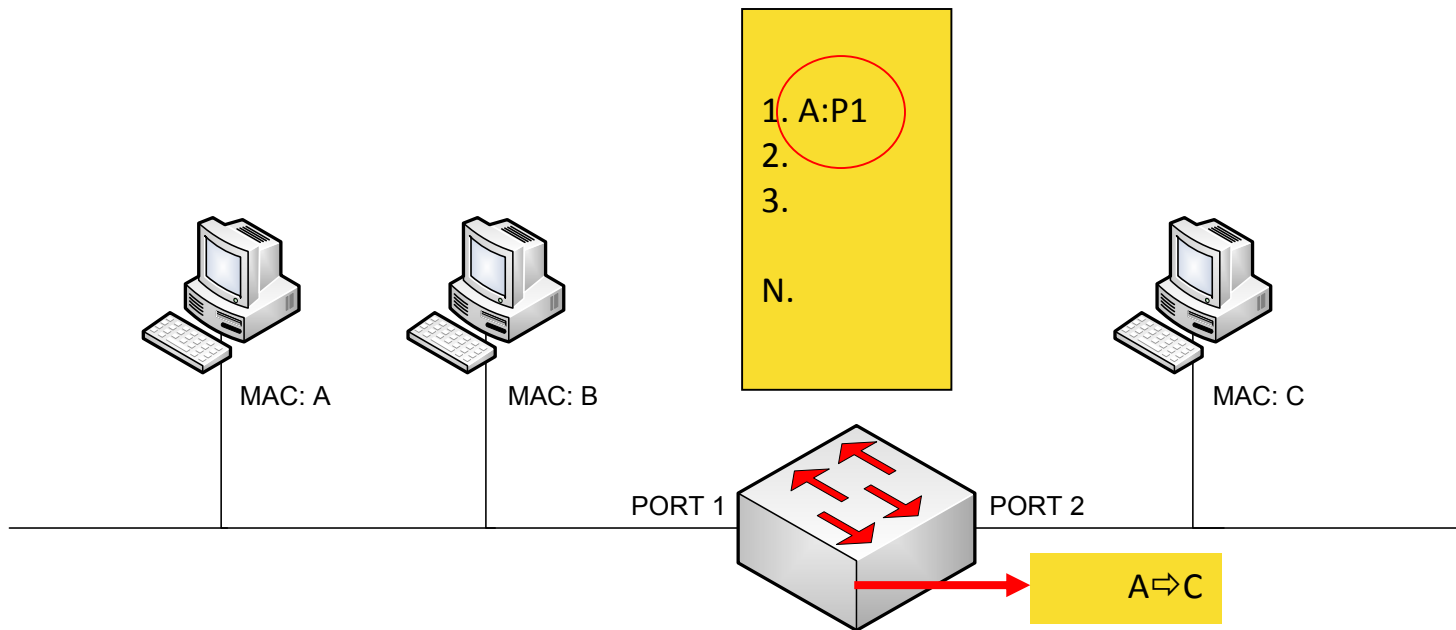
Bekapcsolás utáni hálózat

Kerettovábbítás példával 2.



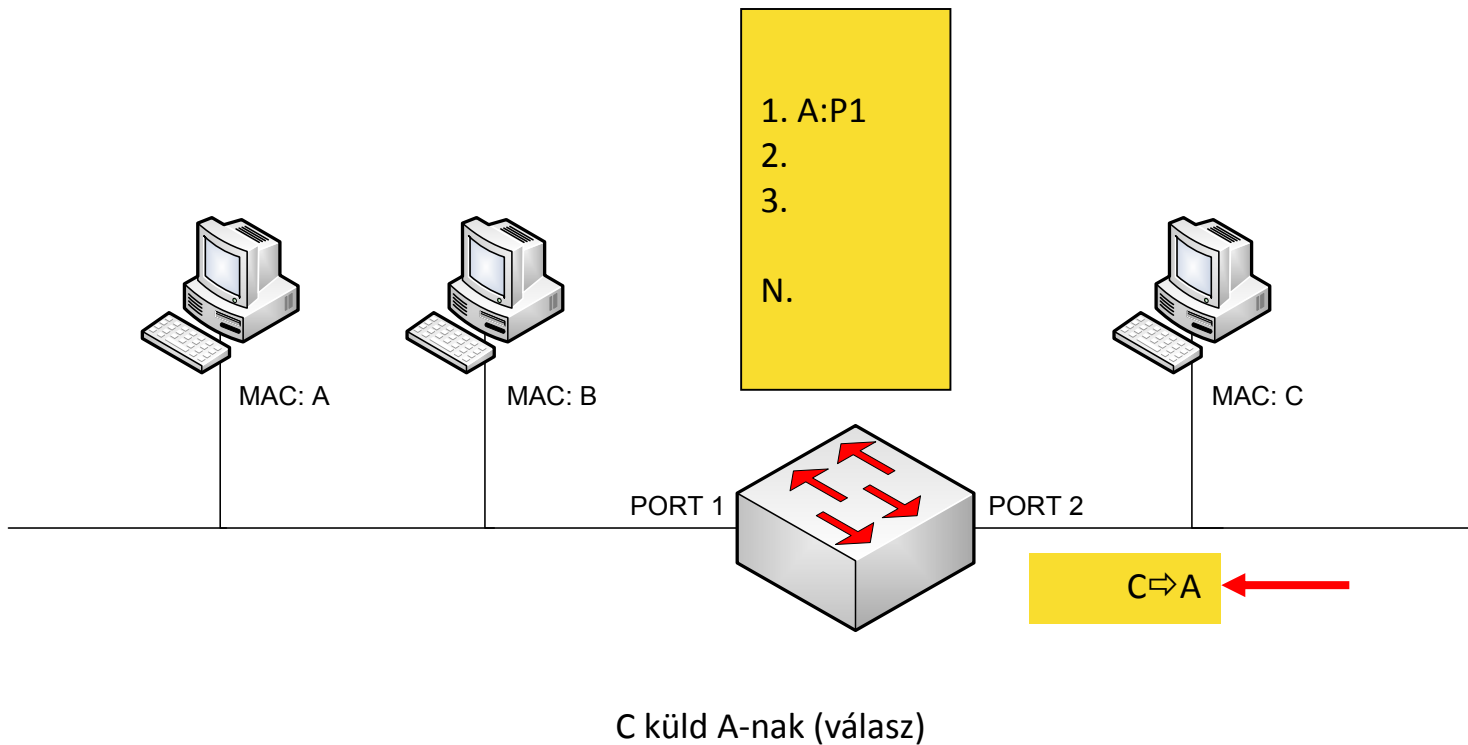
A küld C-nek, az adatbázis üres

Kerettovábbítás példával 3.

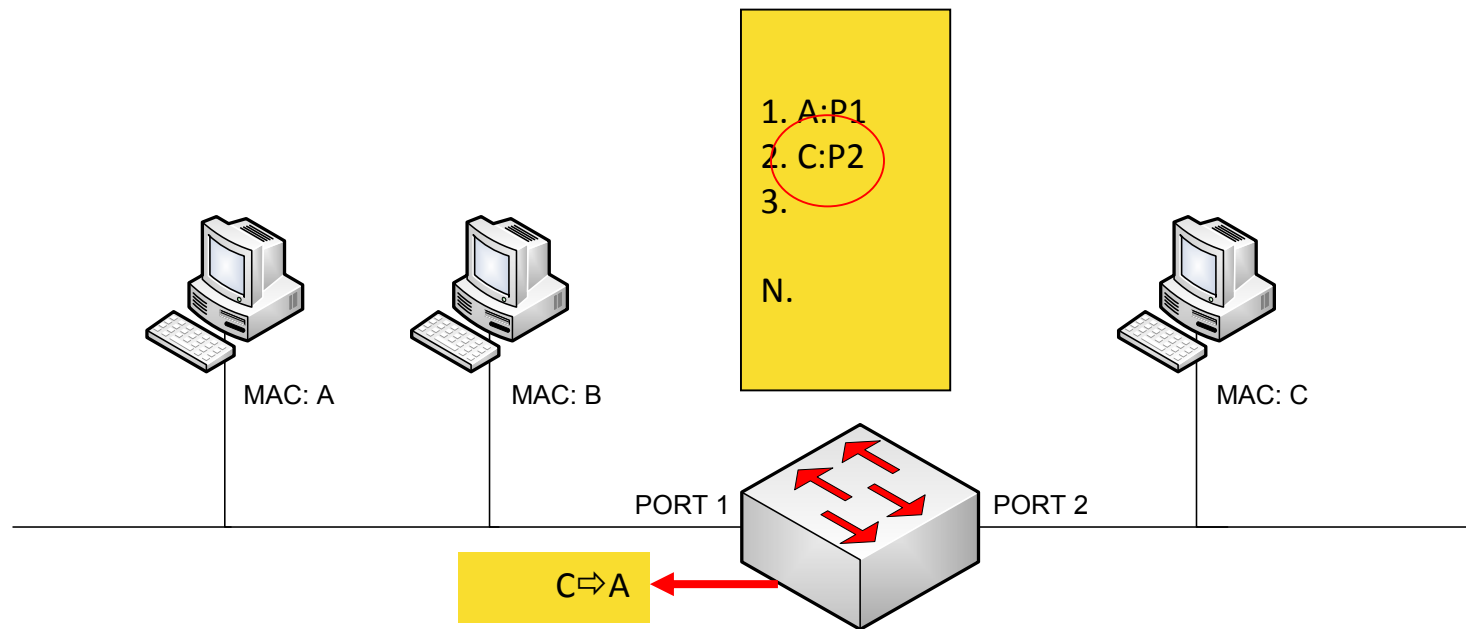


C? ⇒ Elárasztás ⇒ Port 2-ön kiküldjük

Kerettovábbítás példával 4.

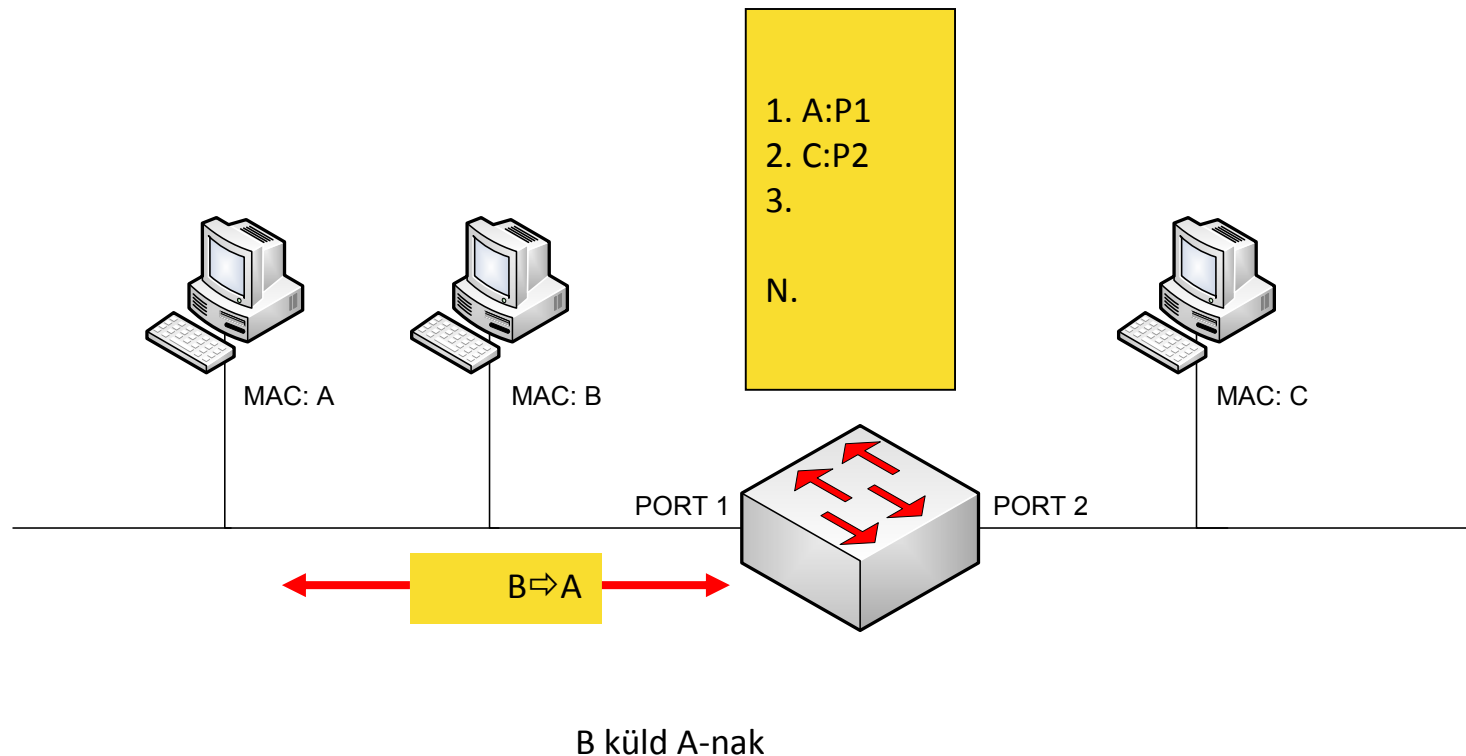


Kerettovábbítás példával 5.

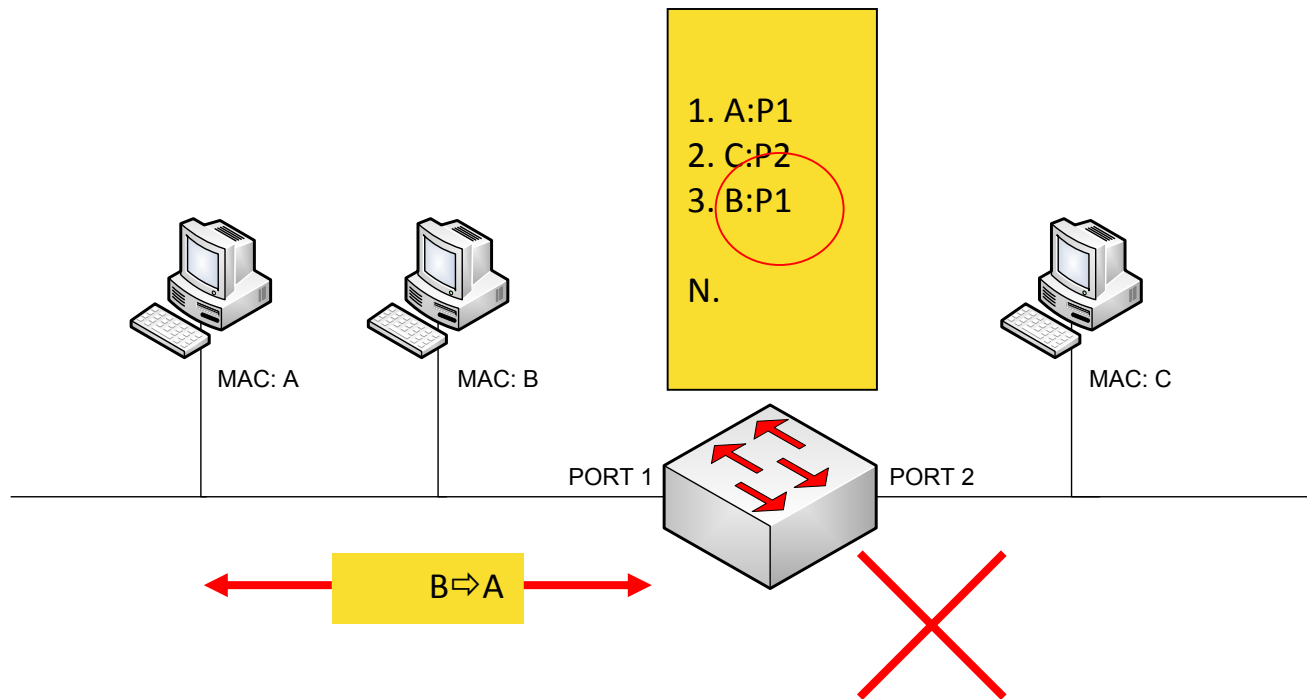


C tanulása, továbbítás A felé (ismert)

Kerettovábbítás példával 6.



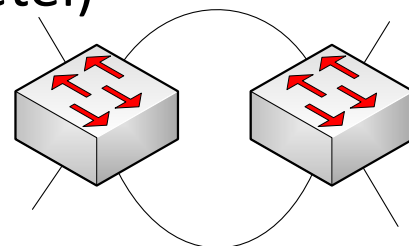
Kerettovábbítás példával 7.



B tanulása, C felé szűrés
(a vizsgált bridge közreműködése nélkül A megkapja B keretét)

Spanning Tree Protocol, STP

- STP/RSTP IEEE 802.3 beágyazást használ (egyik kivétel)
- Spanning Tree Protocol (STP)
 - Adatkapcsolati rétegben működik
 - Redundáns összeköttetések STP nélkül
 - Az Ethernet hálózat „begerjed”
 - Pozitív visszacsatolás a rendszerben
 - A keretek a zárt hurokban körbejárnak, örök életűek
 - Redundáns összeköttetések szükségesek a megbízhatóság javítására, ezért az STP feladata:
 - A redundáns összeköttetések miatti hurkokat tartalmazó gráf topológiában „kikonfigurálni” redundáns összeköttetések közül annyit, hogy egy feszítőfát találjunk (nincs benne hurok)
 - Ha feszítőfa szétesik, akkor ismét összekötni a redundáns „kikonfigurált” összeköttetések közül megfelelő számú visszakonfigurálásával
- Lassú, nagyobb hálózatokban a beálláshoz 30-50 s is szükséges
- Sok más apróbb probléma (tanszéki vPro labor rejtély)



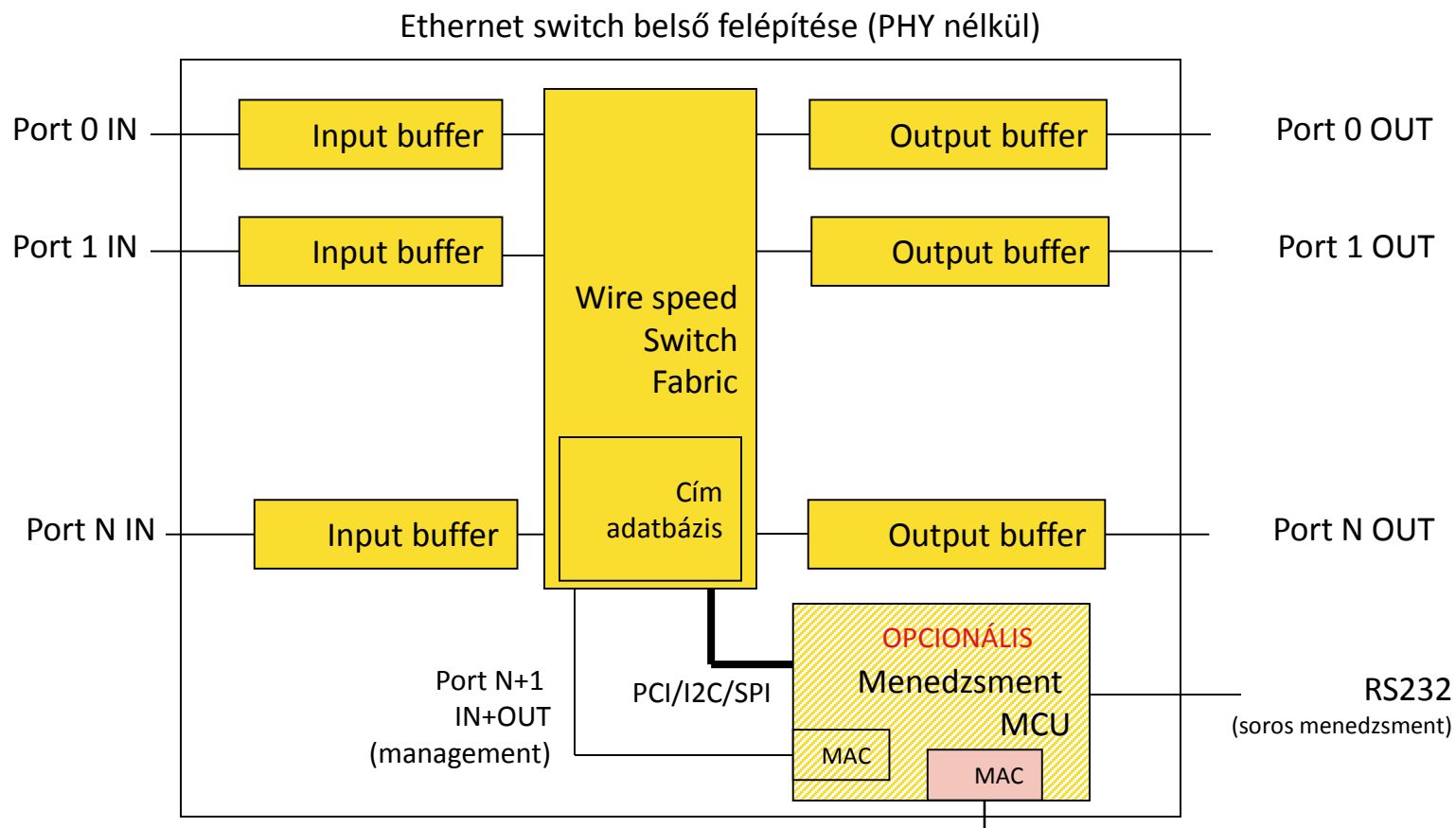
Rapid STP, RSTP

- Rapid STP
 - Az STP lassan konvergál (találja meg a feszítőfát)
 - Az RSTP alkalmazása javasolt (bekapcsolási problémák)
 - VLAN és STP/RSTP konfigurálása nagy körültekintést igényel
- Reakció idő:
 - Detektálható link hiba esetén $n \cdot 1\text{ms}$
 - Egyébként $3 \cdot \text{Hello_Times}$
 - Alapértelmezett Hello_Times : 2 s
- Ha az eszközeink tudják, érdemes beállítani!
 - RSTP sok bajtól tud megkímélni
 - „Rövidrezárt” hálózat katasztrófát okozhat
 - Redundancia
 - STP viszont kerülendő!

Ethernet kapcsolók

- Funkciójukban teljesen azonosak a hidakkal
 - A klasszikus dedikált Ethernet hidak az 1990-es évek közepéig voltak használatban (10 Mb/s sebességig, jellegzetesen 2 porttal)
 - Napjainkban ismét előkerülnek tisztán SW megvalósítások:
 - Virtuális gépekben a hálózat elérésének egyik módja (VMWARE, VirtualBox, Microsoft Virtual PC)
- Az Ethernet kapcsoló (L2 switch) gyors dedikált HW implementáció
 - Speciális ASIC végzi az összes funkciót
 - Néhány IC-ből össze lehet rakni egy $48 \times 10/100/1000 \text{ Gb/s} + 2 \times 10 \text{ Gb/s}$ Ethernet kapcsolót
 - „Wire speed” (nem veszít csomagot legitim forgalomnál)
 - Multicast tanulás (Internet Group Management Protocol, IGMP) alapján
 - A multicast forgalom nem jut ki ekkor minden portra (szűrés)
 - Broadcast Storm Control
 - Megadható a broadcast csomagok maximális továbbítási sebessége $N \text{ frame/s}$
 - Bizonyos alkalmazásokkal nem kompatibilis (Norton Ghost Enterprise)
 - Port mirroring (majd beszélünk róla később)

Belső felépítés



Dedikált menedzsment port
(out of band, high end eszközök)

Link aggregation

- IEEE 802.3-2005 43-as bekezdés
 - IEEE 802.1AX-2008-ban általánosítva
- Több Ethernet link összekapcsolása egyetlen logikai link-be
 - Sáv szélesség összeadódik
 - Forgalom megosztás a fizikai linkek között
 - Egy „kapcsolat” csak egy fizikai linket használhat
 - Ethernet cím alapján
 - Hálózati cím alapján (IP cím)
 - Hálózati cím és alkalmazás alapján (IP cím és TCP/UDP port)
 - Redundancia (megbízhatóság nő)
- Konfigurálása
 - Manuális konfigurálás
 - Link Aggregation Control Protocol (LACP)
 - VLAN és Link aggregation együttes konfigurálása nagy körültekintést igényel

Rétegeken kell
átlátni hozzá!
Nem szép...

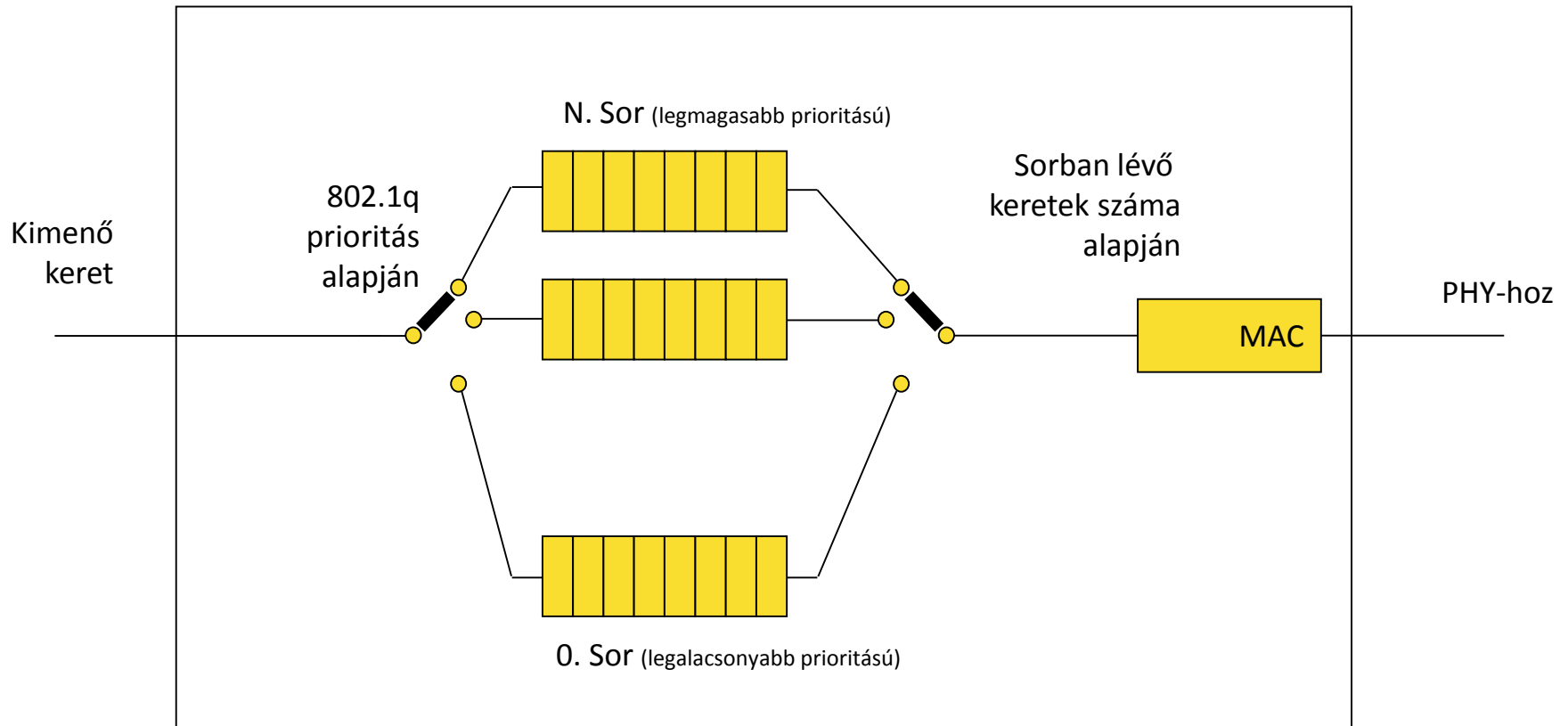
Még rondább,
mint az előző!

IEEE 802.1q alkalmazása

- Prioritások
 - IEEE 802.1q: 3 bites prioritás mező
 - 0: legalacsonyabb, 7: legmagasabb
 - Switch egress porton $1 \leq N \leq 8$ kimeneti várakozási sor
 - 1. kérdés: Egy adott prioritású keret melyik sorba menjen
 - 2. kérdés: Milyen algoritmus szerint ürítjük a várakozási sort
- Bizonyos switch-ek képesek Ethertype, IP cím, TCP/UDP port alapján IEEE 802.1q mezőt beszúrni, akár a meglévő mező által megadott prioritás figyelmen kívül hagyásával is
 - Ez nem adatkapcsolati réteg funkció az Ethertype-ot kivéve
 - Az architektúra nem tiszta, a rétegek közötti funkciók összekeverednek
 - Ugyanakkor így garantált a prioritás (a kliens gép nem feltétlenül van erős felügyelet alatt)
 - Beágyazott rendszerekben jól használható funkció
 - A rendszer állítja be a prioritást, nem a felhasználó

802.1q prioritás érvényesítése, A kimenet blokkvázlata

Output 802.1q esetén

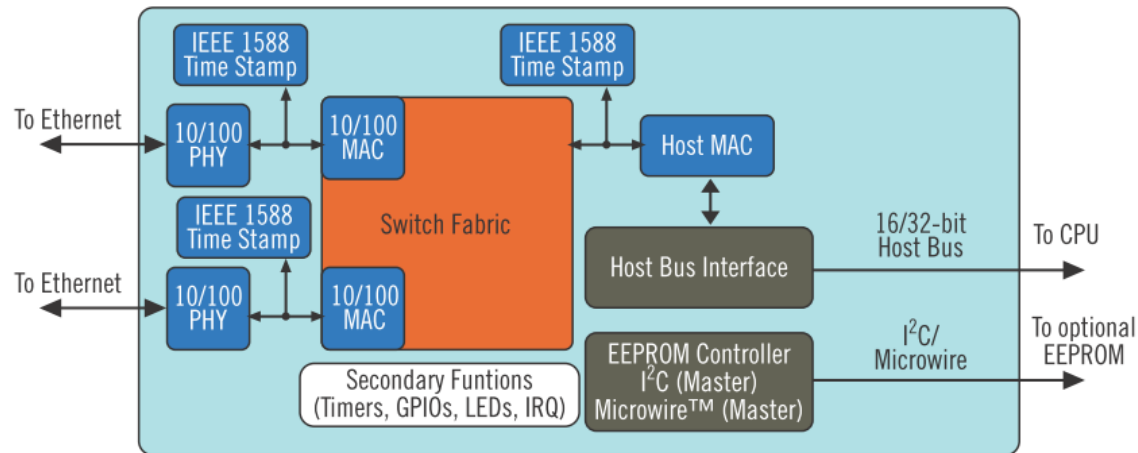


Ethernet switch chip gyártók...

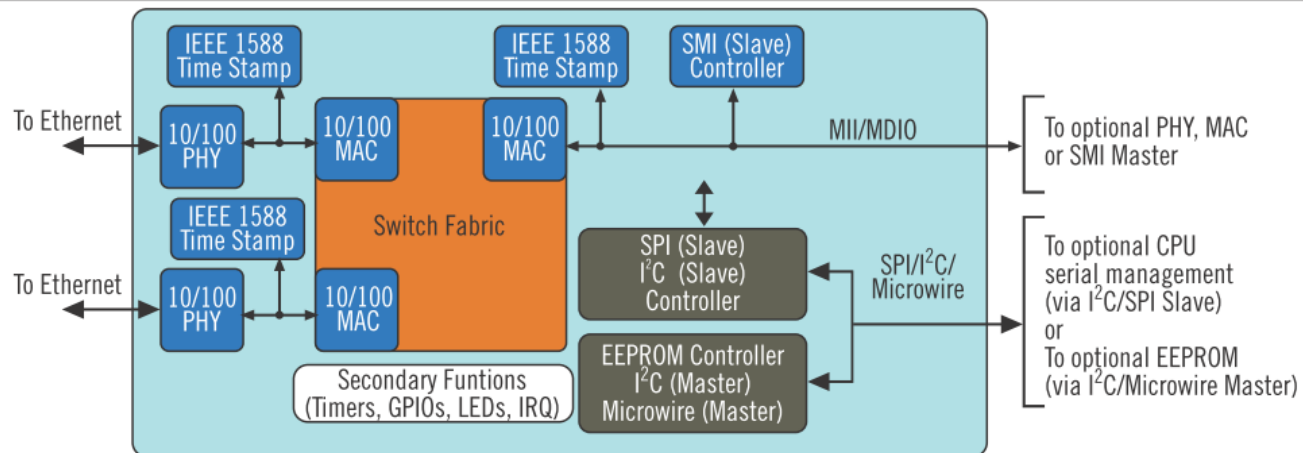
- Broadcom
 - http://www.broadcom.com/products/?industry_id=3
- Marvell
- Vitesse
- Redback Networks (Ericsson)
- Atheros
 - <http://www.atheros.com/>
- Intel
- SMSC , Realtek és egyéb kis gyártók (3 és több port, SOHO, stb.)
 - <http://www.smsc.com/index.php?tid=147>
- Cisco + Juniper + Huawei? (high end, not for resale)

3 portos példa, SMSC LAN931x

LAN9311/9312 BLOCK DIAGRAM



LAN9313 BLOCK DIAGRAM



Vállalti switch példa, BCM56512

- BCM 56512 - 24-Port GbE Multilayer Switch with Two 10 GbE/HiGig+™ Ports (switch)
 - PCI menedzsment interfész
- BCM 5836P - Single-Core 32-bit MIPS® Control Plane Processor with PCI (menedzsment)
 - FLASH + DDR RAM kell még hozzá
 - UART, 2 10/100 Ethernet with MII
 - 3DES, AES, DES, SHA-1, MD5 HW gyorsítás
- 6 db BCM5464R - Quad-Port 10/100/1000BASE-T Gigabit Copper Transceiver (fizikai réteg vezérlők)
 - 256-pin BGA
 - 750mW/port
 - Nincs szükség hűtőbordára
- Adatlapok csak NDA aláírása után

High-end enterprise switch

- BCM56820
 - 24-Port 10-GbE and 4-Port 1 GbE Multilayer Gigabit Ethernet Switch

Autóipari switch

- BCM89500

- 4 vagy 5 BR downlink port, 3 vagy 2 uplink port (max. 1 Gb/s, xMII interfész)
- Cortex R4 menedzsment processzor memóriával
- 802.1AS támogatás és AVB támogatás