

10 Gigabit Ethernet

JÁKÓ ANDRÁS

BME Egyetemi Informatikai Szolgáltató Központ
jako.andras@eik.bme.hu

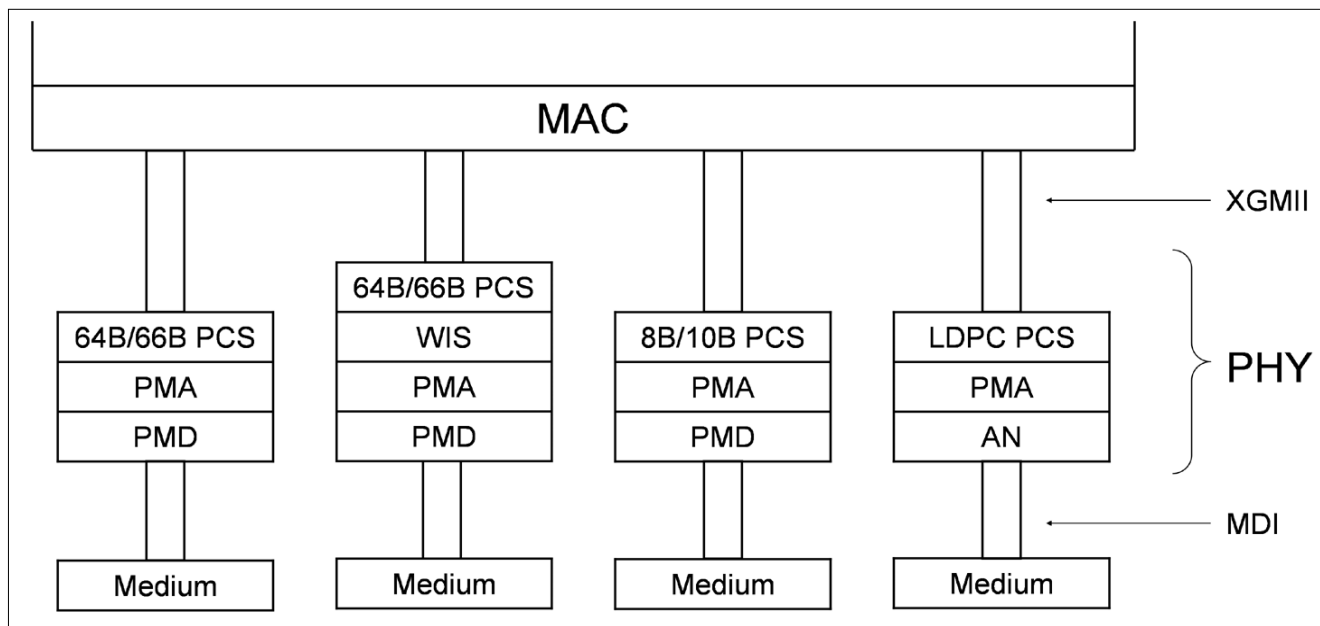
Kulcsszavak: Ethernet, 802.3, lokális hálózat (LAN), 10GBASE-T eszközök

A 10 Gigabit Ethernet¹ szabványosítása már jó néhány éve folyik. Több ajánlás megjelent már, de néhány ajánlás most is dolgozik az IEEE – ezek várhatóan idén vagy jövőre készülnek el. Ebben a cikkben a 10GigabitEthernet újdonságait mutatja be a szerző az Ethernet család korábbi, kisebb sebességű tagjaihoz képest.

1. Bevezetés

Az Ethernet családot leíró IEEE² 802.3 szabvány [1] folyamatosan hízik, immár több mint húsz éve. A 802.3 szabványhoz első megjelenése óta számos kiegészítés készült, mint például a 802.3u jelű 1995-ben, ami a 100BASE-TX fizikai réteget specifikálta (ez a ma is használatos csavart érpáros FastEthernet), vagy 2003-ban a csavart érpáron keresztüli tápellátásról szóló 802.3af. Most is több, pillanatnyilag nyolc ilyen kiegészítésen dolgoznak az IEEE munkacsoportjai. Ezek a kiegészítések aztán belekerülnek a szabványba, amit az IEEE néhány évente újra kiad – a legutolsó változat a tavaly megjelent IEEE 802.3-2005.

1. ábra 10 Gigabit Ethernet komponensek



A legfrissebb már elkészült és a jelenleg kidolgozás alatt álló kiegészítések között is több olyan van, ami a 10 Gbit/s sebességű működést, és az arra alkalmas fizikai rétegeket írja le. Ezzel tehát ismét egy nagyságrenddel gyorsabb tagokkal bővült az Ethernet család a korábbi legnagyobb, 1 Gbit/s sebességű GigabitEthernethez képest.

2. MAC

A 10GigabitEthernet MAC³ legfontosabb újdonsága, hogy ezen a sebességen csak full-duplex működést enged meg.

A busz topológiájú Ethernet fizikai rétegek (10BASE2, 10BASE5) esetén a CSMA/CD⁴ protokoll és a half-duplex működés elengedhetetlen volt. Az igen ritkán hasz-

1 Az IEEE 802.3 szabványban használt írásmód szerint a különféle sebességű Ethernetek megnevezése külön szóba írandó (pl. „Fast Ethernet” vagy „10 Gigabit Ethernet”), de ezeket a könnyebb értelmezés és az egyértelműség kedvéért a továbbiakban egybe fogom írni („FastEthernet”, illetve „10GigabitEthernet” stb.)

2 IEEE: Institute of Electrical and Electronics Engineers

3 MAC: Media Access Control sublayer, közeghozzáférési alréteg

4 CSMA/CD: Carrier Sense Multiple Access with Collision Detection

nált 100BASE-T4 esetén a half-duplex működésre már nem a topológia miatt volt szükség, hanem azért, mert nem használtak hibrid áramköröket, így egy érpáron egyszerre csak egy irányba lehetett adatokat továbbítani, viszont a 100 Mbit/s sebesség eléréséhez a Cat3 minőségű kábelben egyszerre három érpárra volt szükség.

A rendelkezésre álló négy érpáron tehát nem maradt más lehetőség, mint két érpár felváltott használata egyszer az egyik, egyszer a másik irányban. A többi 1 Gbit/s-nál lassabb fizikai rétegnél maga a médium lehetővé tette a full-duplex működést is. Az IEEE 802.3 szabványba csak 1997-re került bele a full-duplex működés leírása (a 802.3x kiegészítésben), de full-duplex Ethernet termékek már valamivel korábban is léteztek (ami nem meglepő, különösen ha arra gondolunk, hogy a full-duplex működés a half-duplexhez képest szinte csak egyszerűsítés). A kábeleket összekapcsoló berendezések ára gyakran mégis indokoltá tette a jobb műszaki lehetőségek ellenére is a half-duplex mód használatát még éveken át. A switch-ek kezdetben ugyanis annyival drágábbak voltak a repeaterknél vagy a huboknál, hogy az csak kevés felhasználónak volt kifizetődő.

A GigabitEthernet fizikai rétegek mindegyike lehetővé tette a full-duplex működést, a szabvány alkotói mégis sok energiát fektettek abba, hogy megtartsák a half-duplex működés lehetőségét is. Ehhez az 1 Gbit/s sebesség miatt a carrier extension és a frame bursting bevezetésére volt szükség. A gyakorlat azonban azt mutatta, hogy – főleg a switch portok árának csökkenése folytán – a GigabitEthernetnél már senkinek sem volt szüksége a half-duplex működésre. Tehát az ahhoz szükséges két fent említett kiegészítés teljesen feleslegesen bonyolítja el a 802.3 szabványt.

A 10GigabitEthernet fizikai rétegben a GigabitEthernethez hasonlóan csak pont-pont összeköttetések lehetségesek, és minden 10GigabitEthernet fizikai réteg alkalmas full-duplex átvitelre. Ennek, és a GigabitEthernetnél tapasztalt fent említett tendenciáknak megfelelően a 10GigabitEtherneten már nincs half-duplex működés és az ahhoz szükséges CSMA/CD protokoll is hiányzik.

3. A 10 Gigabit Media Independent Interface (XGMII)

Az XGMII a MAC és a PHY (fizikai réteg) közti chip-to-chip interfész, azaz egy jól definiált elektromos felület a MAC és a PHY között, hogy azokat könnyű legyen egymástól függetlenül kialakítani. Az XGMII használata opcionális, de az IEEE szabvány az XGMII meglétét feltételezve specifikálja a különböző komponenseket. Az XGMII fizikailag tipikusan ASIC-en belüli logikai egységek között, vagy nyomtatott áramkörön valósulhat meg.

Ez utóbbi esetben az XGMII által összekötött MAC IC és PHY IC távolsága körülbelül 7 cm lehet, és az összeköttetésen HSTL (High Speed Transceiver Logics) jelszinteket kell használni.

Az XGMII irányonként (adás, vétel) egy-egy 32 bit széles adatbuszból, továbbá 4 vezérlő-, és egy órajelből áll. A jelzési sebesség az adatbuszon (és a vezérlőjeleknél) 312.5 Mbaud. A vezérlőjelekből derül ki, hogy a hozzájuk tartozó 8 bit az adatbuszon az Ethernet keret egy byte-ja, vagy valamilyen vezérlő byte (ami például a keret elejét, végét, vagy a keretek közti szünetet jelzi).

3.1. XGMII Extender, XAUI

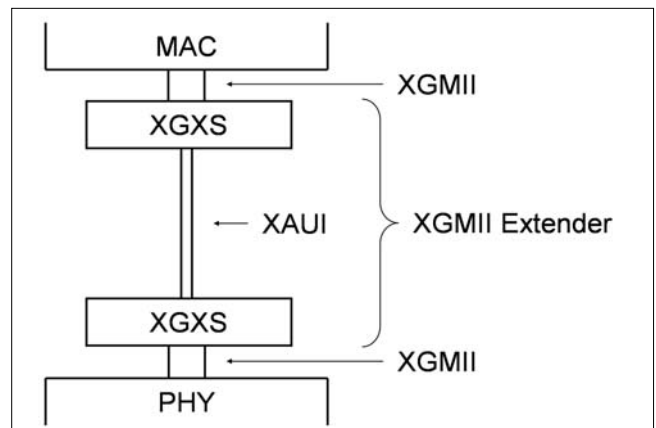
Az XGMII Extender célja az XGMII által nyomtatott áramkörön áthidalható kb. 7 cm távolság növelése (de még mindig csak egy nyomtatott áramkör méreteiben gondolkodva, hiszen így is csak kb. 50 cm lehet a távolság a MAC és a PHY között), valamint a párhuzamos vezetékek számának csökkentése. Az XGMII Extender nem más, mint a 10 Gigabit Attachment Unit Interface (XAUI) két végén kiegészítve egy-egy XGMII Extender Sublayer (XGXS) komponenssel. Ez a két XGXS biztosítja, hogy az XGMII Extender komplexum egészében kívülről mindkét vége felől úgy „nézzen ki”, mint egy egyszerű XGMII.

Az XAUI irányonként mindössze 4 vonalból áll (szemben az XGMII irányonkénti 37 vonalával). A vonalak differenciális meghajtásúak, a jelzési sebesség 3.125 Gbaud. Ez a jelzési sebesség a GigabitEthernetnél megismert 8B/10B kódolás használatából, és a vonalankénti 2.5 Gbit/s adatsebességből adódik. A 8B/10B kódolás nagyobb szimbólumterének köszönhetően nincs szükség külön vonalakra a jelzésekhez, és az órajel is visszanyerhető a jelfolyamból. A 8B/10B kódolást és dekódolást a két XGXS végzi, csak úgy, mint az órajel előállítását a 8B/10B dekódolási oldalon.

4. Fizikai réteg

Az itt tárgyalt fizikai rétegek közül a 10GBASE-T és a 10GBASE-LRM jelenleg kidolgozás alatt áll, még nem része a szabványnak, tehát ezekről még nincs véglegesnek tekinthető információ. A többi már megtalálható a szabvány aktuális változatában (IEEE 802.3-2005).

2. ábra XGMII Extender



4.1. Taxonómia

4.1.1. LAN és WAN fizikai rétegek

A 10GigabitEthernetnél megkülönböztethetünk úgynevezett LAN⁵ és WAN⁶ fizikai rétegeket. A LAN PHY (10GBASE-R, 10GBASE-T, és 10GBASE-X) az Ethernetnél már megszokott megoldás, amikor a médiumra az Ethernet keretek közvetlenül kerülnek (természetesen megfelelő kódolással), az adatkapcsolati rétegben szomszédos Ethernet állomások pedig kábelek és esetleg a fizikai rétegben működő aktív Ethernet hálózati eszközök (hub, repeater, media converter) segítségével kapcsolhatók össze egymással. A 10GigabitEthernet WAN PHY (10GBASE-W) ezzel szemben adáskor SDH/SONET⁷ keretkezéssel látja el az Ethernet kereteket, így két ilyen Ethernet állomás közé nem Ethernet aktív eszközöket, hanem SDH/SONET hálózatot, pontosabban egy SDH/SONET hálózaton megvalósított pont-pont összeköttetést „tehetünk”. A keretkezébeli különbségek miatt LAN és WAN PHY nem köthető össze egymással.

A LAN PHY elnevezés ne tévesszen meg senkit, itt a hálózat kiterjedését tekintve nem feltétlenül lokális hálózatról van szó, hiszen a LAN PHY-k között van olyan, ami 40 km hosszú szakasz áthidalására is képes – és akárhogy is nézzük, viszonylag kevés olyan LAN létezik a világon, amiben ekkora hálózati szegmensek is előfordulnak.

4.1.2. Soros és párhuzamos fizikai rétegek

Osztályozhatjuk a fizikai rétegeket aszerint is, hogy hány jelfolyamot használnak a médiumon. A 10GBAS-

E-R és a 10GBASE-W PHY-k egyetlen, körülbelül 10 Gbit/s sebességű jelfolyamot használnak – ezeket soros PHY-nek szokás nevezni. Ezzel szemben a 10GBASE-T és a 10GBASE-X PHY-k négy kisebb sebességű párhuzamos jelfolyammal érik el a két szomszédos Ethernet állomás közt a 10 Gbit/s adatátviteli sebességet. A négy párhuzamos jelfolyam optikai szálon hullámhossz multiplexálással, réz kábelben pedig külön vezetők használatával valósul meg.

4.1.3. Átviteli közeg

A harmadik lehetséges szempont a 10GigabitEthernet fizikai rétegek osztályozására az átviteli közeg (médium). Ez lehet többféle optikai szálpár, lehet úgynevezett twinaxiális réz kábel, vagy lehet csavart érpáros réz kábel.

4.2. Soros fizikai rétegek

A soros (10GBASE-R és 10GBASE-W) fizikai rétegek nagy részének specifikációja 2002-re készült el. Ezeket az IEEE 802.3 szabvány 802.3ae jelű kiegészítése írta le (ami ma már része a szabvány 2005-ös kiadásának). Az egyetlen kivétel a jelenleg kidolgozás alatt álló 802.3aq jelű kiegészítésben [2] specifikált 10GBASE-LRM.

4.2.1. 64B/66B PCS

A 64B/66B PCS (Physical Coding Sublayer) használatos mindkét soros PHY csoportnál, azaz a WAN PHY-k (10GBASE-W) és a soros LAN PHY-k (10GBASE-R) esetén is.

1. táblázat
10GigabitEthernet
taxonómia

	10GBASE-SR	10GBASE-LRM	10GBASE-LR	10GBASE-ER	10GBASE-SW	10GBASE-LW	10GBASE-EW	10GBASE-LX4	10GBASE-CX4	10GBASE-T
	R				W			X		T
soros		■				■				
párhuzamos								■		■
LAN		■						■		■
WAN						■				
optikai szál		■				■		■		
twinax									■	
csavart érpár										■

5 LAN: Local Area Network, lokális hálózat

6 WAN: Wide Area Network, nagy távolságú hálózat

7 SDH: Synchronous Digital Hierarchy / SONET: Synchronous Optical Network

szinkron bitje „00” vagy „11” értékű nem lehet. Ez alapján, hogy a bitfolyamban pontosan 66 bitenként mindig vagy a „01” vagy az „10” minta fordul elő, a vevő oldali PCS azonosítani tudja a bitfolyamban a 66 bites blokkok határát.

A 66 bites blokkok hátsó 64 bitjén egy önszinkronizáló folyamkódoló alkalmazása biztosítja a megfelelő gyakoriságú szintváltásokat, hogy a vevő oldalon a jelfolyamból az órajel könnyen kinyerhető legyen.

Összegezve a fentieket, adáskor első lépés a 64 bites blokk összeállítás, utána következik a folyamkódolóval való „megkeverés”, majd végül a blokk tartalmának megfelelő két bites szinkron fejléc hozzáadása történik. Vételkor először a 66 bites blokkok határát kell megkeresni a szinkron fejlécek segítségével, utána visszaállítani a „megkevert” 64 bitet, és végül a szinkron fejléc tartalmának megfelelően kinyerni a 64 bites blokból az adatokat és/vagy a vezérlőinformációkat.

LAN PHY esetén a 10GigabitEthernet adatsebesség (a MAC és a PHY között) 10 Gbit/s (azaz 10^{10} bit/s). A fizikai közegen a soros LAN PHY 64B/66B kódolásából adódóan így 10.3125 Gbaud a jelzési sebesség.

4.2.2. WAN PHY

10GBASE-W PHY használatakor az adatkapcsolati rétegben egymással szomszédos 10GigabitEthernet állomások közé SDH vagy SONET hálózatot, pontosabban egy SDH VC-4-64c vagy SONET STS-192c pont-pont összeköttetést kapcsolhatunk. (Ezeket konyhanyelven STM-64, illetve OC-192 néven szokás említeni.)

A vonali sebesség az optikai szálon ilyenkor 9.95328 Gbit/s, az adatsebesség pedig 9.58464 Gbit/s (az SDH/SONET fejlécek miatt). A MAC és a PHY között az átviteli sebesség ennél nagyobb, pontosan 10 Gbit/s. A különbség áthidalása érdekében WAN PHY esetén a MAC úgynevezett „IFS stretch” módban működik, azaz adáskor keretenként a keret méretével arányos hosszúságú extra szünetet (idle szimbólumokat) iktat be, tehát az interframe gap a minimálisan előírtnál hosszabb lesz. 104 bitenként 8 bit extra interframe gap szükséges a sebességkülönbség kompenzálásához.

Ezeket az idle szimbólumokat a PHY eldobja, így a VC-4-64c, illetve az STS-192c adatsebességnek megfelelő továbbítandó jelfolyamot kap.

Vételkor ennek a fordítottja történik, azaz a PHY szűr be szünetet a vett keretek közé, hogy a MAC és a PHY között meglegyen a 10 Gbit/s sebesség. A MAC-nek vételkor ezzel semmi különleges teendője nincs, hiszen az egyébként is teljesen normálisnak számít, hogy a vett keretek közt kisebb-nagyobb szünetek vannak.

A WAN Interface Sublayer (WIS) a 10GBASE-W fizikai réteg fontos komponense, melynek feladata adáskor a 64B/66B PCS-től kapott bitfolyam SDH/SONET keretkezéssel való ellátása, illetve vételkor az adatok kinyerése az SDH/SONET keretéből. A keretkezéssel együtt természetesen a WIS feladata az SDH/SONET fejlécekben szereplő jelzések, riasztások kezelése is.

4.2.3. Soros PMD

Soros 10GigabitEthernet fizikai rétegek többféle fizikai médiummal és többféle PMD-vel (Physical Medium Dependent sublayer) léteznek. A fizikai médium minden soros PHY esetén optikai szálpár.

Multimódusú szálpáron működik a Short Wavelength Serial PMD (10GBASE-SR és 10GBASE-SW PHY) 850 nm névleges hullámhosszon, valamint a 10GBASE-LRM PMD 1310 nm névleges hullámhosszon. Az áthidalható távolság a fényvezető szál megfelelő hullámhosszon mérhető módális sáv szélességétől függ. Ezeket a 2. táblázat foglalja össze (a könnyebb összehasonlítás érdekében szerepel a táblázat utolsó oszlopában a 10GBASE-LX4 párhuzamos PMD is, amiről bővebben a következő fejezetben lesz szó).

A 10GBASE-LRM PMD a 10.3125 Gbaud jelzési sebesség mellett 220, illetve 100 méteres távolság elérése érdekében a vevő oldalon elektronikus diszperzió kompenzációt alkalmaz.

1310 nm-es 10GBASE-LRM PMD multimódusú szálon való használata esetén úgynevezett „mode-conditioning patch” kábel beiktatása szükséges az adónál (hasonlóan ahhoz, mint például amikor nagyobb távolságú multimódusú szakaszon használunk 1000BASE-

2. táblázat

Multimódusú optikai szálon áthidalható távolság a kábel típusának függvényében (az első sorban szereplő 160/500 MHz*km módális sáv szélességű kábel az ún. „FDDI grade” kábel)

mag átmérő [µm]	módális sáv szélesség [MHz*km]		áthidalható távolság [m]		
	850 nm	1300 nm	10GBASE-SR, -SW (850 nm, ~10 Gbaud)	10GBASE-LRM (1310 nm, ~10 Gbaud)	10GBASE-LX4 (1310 nm, ~3 Gbaud)
62.5	160	500	26	220	300
	200	500	33	220	300
50	400	400	66	100	240
	500	500	82	220	300
	2000	500	300	220	300

LX PMD-t). A probléma forrása a régi multimódusú szálaknál a gyártási technológia, ami miatt a sugár szerinti törésmutató profil a 0 környékén (a szál közepén) gyakran nem eléggé egyenletes, azaz viszonylag nagy ugrások lehetnek benne [3].

Ez a régebben használt – csak kisebb jelzési sebességekre alkalmas – LED fényforrás esetén nem okozott gondot, hiszen a széles szögtartományban világító LED-ből a szál magját betöltő fény (OFL, Overfilled Launch) energiájának csak kis hányada jut a szál hibás részére. Viszont a sokkal kisebb keresztmetszeten és szögtartományban sugárzó LASER fényforrással belevilágítva ugyanezen szálak közepébe az energiának igen nagy része jut a mag hibás középső részére, és ez könnyen elviselhetetlenül nagy DMD-hez (Differential Mode Delay) vezethet.

A probléma elkerülésére alkalmazott speciális patch kábel egy monomódusú és egy graded index multimódusú szakaszból áll, amelyek szándékosan excentrikusan vannak egymáshoz illesztve. A 10GBASE-LRM adó oldalára kerül a monomódusú szakasz, amiből a LASER által kibocsátott fénysugár a multimódusú kábelbe nem a közepén, hanem kicsit arrébb jut be, és ezáltal a régi multimódusú optikai kábelszakaszon elkerüli a mag feltehetőleg hibás középső részét.

Monomódusú szálpáron, 1310 nm névleges hullámhosszon működik a 10GBASE-LR és a 10GBASE-LW fizikai rétegben használt Long Wavelength Serial PMD. Az ezzel áthidalható távolság 10 km.

A legnagyobb áthidalható távolságot, 40 km-t az Extra Long Wavelength Serial PMD nyújtja. Ebben a csoportban is megvan a LAN (10GBASE-ER) és a WAN (10GBASE-EW) PHY. A használt médium természetesen monomódusú optikai szálpár, a névleges hullámhossz pedig 1550 nm. Ennél a PMD-nél az optikai szakasz csillapítása legalább 5 dB kell, hogy legyen. Ha ez nem adódik ki a kábel és a csatlakozók csillapításából, akkor megfelelő méretű csillapító tag beépítése szükséges.

4.3. 10GBASE-X fizikai rétegek

A két 10GBASE-X fizikai réteg közül a 10GBASE-LX4 az IEEE 802.3 szabvány 802.3ae jelű, 2002-es kiegészítésében szerepel. A 10GBASE-CX4 specifikációja 2004-re készült el, a szabvány 802.3ak jelű kiegészítéseként. Ma már mindkettő része a szabvány 2005-ös kiadásának.

4.3.1. 10GBASE-X PCS

A 10GBASE-X fizikai rétegek a GigabitEthernetből ismert 8B/10B blokk kódolást használják, amely 8 bites adatblokkokat képez le 10 bites szimbólumokra. A 10 bites kódszavak között az adat byte-ok megfelelőin kívül van még néhány érvényes, jelzésre használt kódszó is. Ezeket a speciális kódszavakat használja az 10GBASE-X PCS például a keretek elejének és végének megjelölésére, vagy a keretek közti szünetek kitöltésére. A 8B/10B kódolás lehetővé teszi bithibák de-

tektálását. A kód úgy van kialakítva, hogy minden 10 bites kódszóban legalább 3 szintváltás legyen, így az órajel kinyerhető magából a csatornán vett adatfolyamból. Szintén detektálható egyéb vezérlőjelek nélkül – az idle szimbólumok segítségével – a kódszavak határa is.

A 10GBASE-X PCS feladata adáskor a MAC felől érkező 32 bites adatblokkok és vezérlőinformációk 8B/10B kódolása. Az így kapott négy kódszó 4 külön vonalon kerül a csatornára. Az idle szimbólumokat a PCS úgy választja meg, hogy az elektromágneses spektrum egyenletes, a fehér zajhoz hasonló legyen, tehát ne okozzon kellemetlen interferenciát.

Vételkor a PCS feladata a 10 bites kódszóhatárok detektálása vonalanként, majd a 4 kódszó egymáshoz igazítása, hiszen a négy vonalon keletkezhet átvitel közben egy kis csúszás a túloldalon egyszerre leadott szimbólumok közt. Ezután az így kapott négy 10 bites szimbólumot a PCS dekódolja, és továbbítja a MAC felé.

Mivel a 10GBASE-X PMD-k ugyanazt a kódolást használják, mint az XAUI, és mivel az XGMII interface használata opcionális, ezért 10GBASE-X PMD és XAUI együttes használata esetén a PHY és az XAUI közti XGXS, valamint a PHY PCS és PMA komponensei kihagyhatók, hiszen ezek ilyenkor feleslegesen konvertálnák a jeleket egyik formából a másikba és vissza. Ilyenkor tehát az XAUI közvetlenül csatlakozhat a PMD-hez, a PCS és a PMA funkcióját pedig az XAUI és a MAC közti XGXS látja el.

Ebben az esetben azonban szükséges lehet az XAUI-n keletkező esetleges jitter kompenzálására a PMD és az XAUI között.

4.3.2. 10GBASE-X PMD

A MAC és a LAN PHY közti 10 Gbit/s adatsebesség négy párhuzamos jelfolyamra bontása és a 8B/10B kódolás alkalmazása folytán a 10GBASE-X PMD jelzési sebessége 3.125 Gbaud.

A 10GBASE-LX4 PMD optikai szálpáron működik. Használható egyaránt mono- és multimódusú optikai szálakon is. Az áthidalható maximális távolság monomódusú szálon 10 km, 500 MHz*km modális sáv szélességű multimódusú szálon 300 m, 400 MHz*km sáv szélességűn pedig 240 m (lásd 1. táblázatot). Multimódusú kábel esetén itt is szükséges a 10GBASE-LRM PMD-nél tárgyalt mode-conditioning patch kábel használata. A négy jelfolyam átvitele az optikai szálon hullámhosszosztásos multiplexálással (WDM) történik. A vivők (konyhanyelven lambdák) névleges hullámhosszai 24.5 nm-enként helyezkednek el az 1310 nm körüli optikai ablakban.

A 10GBASE-CX4 réz kábelben működik. Tipikusan géptermi alkalmazásokra készült, a vele áthidalható távolság mindössze 15 m. Az átviteli közeg 100 Ω impedanciájú, úgynevezett twinax kábel. A twinaxiális kábel felépítésében hasonlít a koaxiálisra, de a belsejében nem egy, hanem két – külön-külön szigetelt – vezető van, melyeket a 10GBASE-CX4 PMD differenciálisan hajt meg. A négy jelfolyam átvitele négy twinaxiális kábelben

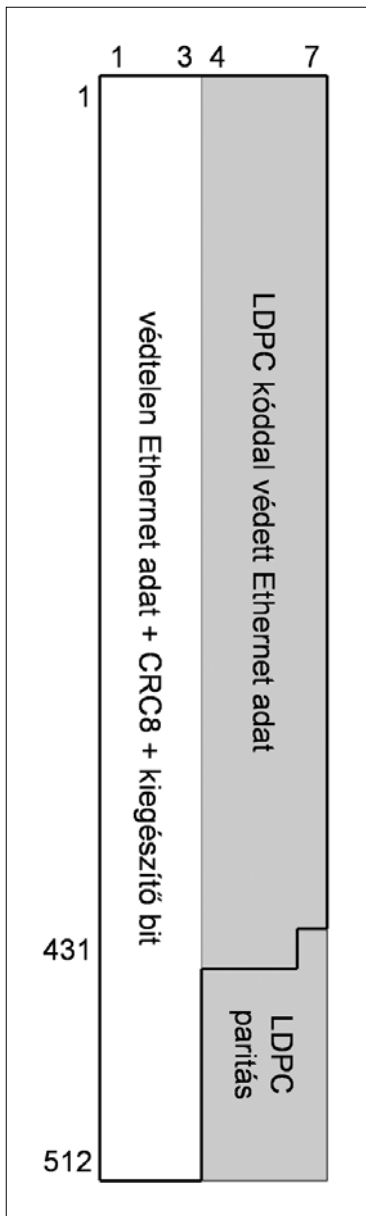
történik, így az adáshoz és vételhez egyszerre 8 twin-axiális kábelre van szükség. Az érintkezők száma 25, mivel a 8 érpár külön-külön, és a teljes kábel együtt is árnyékolva van. Az MDI csatlakozója az IEC 61076-3-113 előírásainak megfelelő Infiniband 4x típusú.

4.4. Csavart érpáros fizikai réteg

A GigabitEthernethez hasonlóan a 10 Gbit/s sebességű fizikai rétegek között is jóval később fog elkészülni a csavart érpáron működő változat, mint például az optikai szálpáron működő társai.

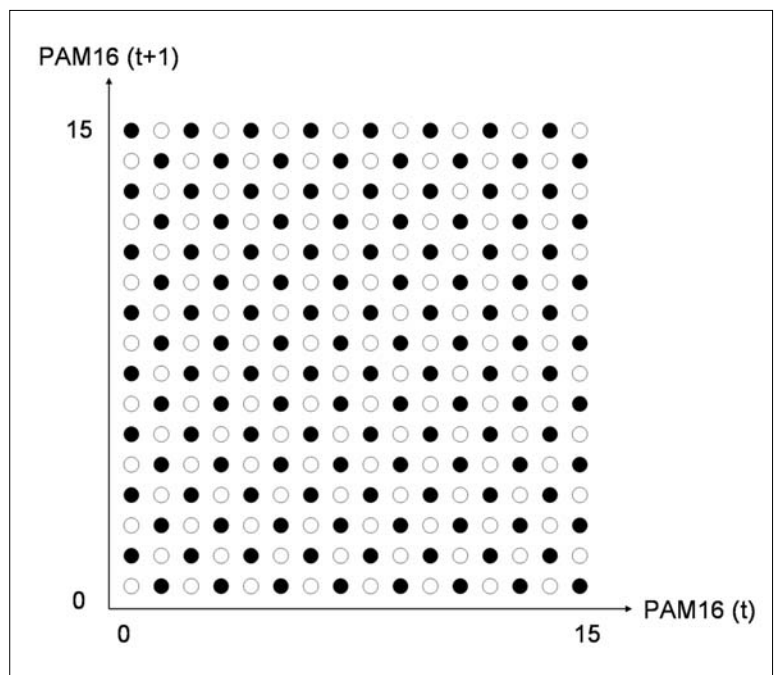
Ez azzal magyarázható, hogy a csavart érpáros kábel sáv szélesség kapacitása jóval kisebb, mint egy optikai szálé, ezért komoly kódolási és jelfeldolgozási arzenál bevetésére van szükség a nagy adatátviteli sebesség eléréséhez. A 802.3 szabvány csavart érpáros (10GBASE-T) fizikai réteget leíró kiegészítése (802.3an [4]) jelenleg kidolgozás alatt áll, de várhatóan hamarosan elkészül.

A 10GBASE-T az Ethernet család kisebb sebességű tagjaihoz hasonlóan 4 csavart réz érpárból álló kábelt használ, bár a 10 Gbit/s sebességhez szükséges kábel valamivel bonyolultabb felépítésű – és ebből adódóan kedvezőbb átviteli tulajdonságokkal rendelkezik – mint a kisebb sebességekhez szükséges csavart érpáros kábelek.



3. ábra
Az LDPC kódoló alkalmazása

4. ábra
DSQ128 konstelláció



4.4.1. Auto-negotiation

A 10GBASE-T a többi csavart érpáros Ethernet fizikai rétegnél megszokott auto-negotiationt használ. Az auto-negotiation alapvető funkciói nem változtak a 1000BASE-T-hez képest. Az eljárás ki van egészítve a 10 Gbit/s sebességű kapcsolathoz szükséges elemekkel, és néhány kisebb optimalizáláson is átesett, de természetesen a korábbi változatokkal felülről kompatibilis maradt. (A duplexitás egyeztetésére 10 Gbit/s sebesség esetén nincs szükség, de mivel a sebességet és a duplexitást úgy is egyszerre, egy paraméterben kezeli az auto-negotiation, ezért ez sem jelent lényegi változást.)

4.4.2. PCS

A kódolási lépések sorában adáskor az első a 64B/65B kódolás. Ez majdnem azonos a soros fizikai rétegeknél használatos 64B/66B kódolással. A különbség annyi, hogy itt kettő helyett csak egy bites a 65 bites blokkok fejléce, ami jelzi, hogy az utána következő 64 biten csak adat byte-ok vannak-e, vagy vezérlőinformációk is. (Az 1 bites fejléc értelemszerűen nem játszik szerepet a blokkhatárok detektálásában.)

A következő lépésben az adó összegyűjt 50 darab ilyen 65 bites blokkot, amihez 8 bit CRC-t (Cyclic Redundancy Check) ad, majd további 1 (hálózati információt nem hordozó) bittel egészíti ki. Ez utóbbi kiegészítésre azért van szükség, hogy összeálljon egy pontosan 3259 bites blokk, ami a kódolás következő lépéséhez kényelmes méret. Ebből a blokkból 3*512 bit (köztük az 1 kitöltő bit is) megmarad hibajavító kódolás nélkül, a másik 1723 bit mellé pedig LDPC (Low-Density Parity-Check) lineáris szisztematikus blokk kódolót [5] alkalmazva 325 bit paritás kerül.

Így az adó összesen 1723+325=2048=4*512 hibajavító kóddal védett bithez, továbbá 3*512 korábban „félretett” védtelen bithez jut. Ez együtt 512 db olyan 7

bites szóba rendezhető, ahol minden szó első 3 bitje védtelen, a második 4 bitje pedig hibajavító kóddal védett.

A csatornán, vagyis a csavart érpáros kábelben minden érpáron 16 szintű PAM (Pulse Amplitude Modulation) szimbólumokat küld az adó. Két, időben egymást követő, egyazon érpáron leadott PAM16 szimbólum 256 lehetséges kombinációjából az adó egy 2 dimenziós, 16*16-os mátrix minden második elemét sakktableszerűen elhagyó 128 elemű, maximális távolságú konstellációt használ. Az így kijelölt 128 db PAM16 szimbólumpárt a szabványtervezet DSQ128 (double square 128) szimbólumoknak nevezi.

Ezek a DSQ128 szimbólumok 8 db, egyenként 16 elemű, „csúcsára állított” négyzet alakú csoportba vannak osztva. A csoportok a 16*16-os mátrix széleinél a szemközti oldalon folytatódnak.

A fent említett 512 db 7 bites szó mindegyike egy DSQ128 szimbólumot jelent a csatornán. A 128 elemű konstellációban a csoportot a 7 bites szó első 3 védtelen bitje jelöli ki, a 16 elemű csoporton belüli elemet pedig a szó hátsó 4 védett bitje határozza meg. Ez a módszer biztosítja, hogy a védtelen bitek esetleges módosulása a csatornán a nagy távolság miatt ne maradjon észrevétlen, illetve könnyen javítható legyen. Az LDPC kóddal védett 4 bit módosulása esetén pedig a hibajavító kódolás segíthet a szimbólumok közti kis távolság ellenére is.

Összegezve a fentieket, az adó 50*64 bitből képez 512 db 7 bites szót, majd minden szót két PAM16 szimbólumban ad le. Ez 10 Gbit bemenetre $3.2 \cdot 10^9$ PAM16 szimbólumot jelent, ami 4 érpárra elosztva $800 \cdot 10^6$ szimbólumot eredményez érpáronként. A jelzési sebesség tehát 800 Mbaud.

Vételkor az LDPC dekódolás soft decision módszerrel, esetlegesen több iterációs lépésben történik. A dekóder bemenetére tehát még a 16 szintű A/D konverzió

előtti, sokkal finomabb felbontású információ kerül, így gyakorlatilag nem vesz el az az információ, amit az analóg jelszintek hordoznak. A dekóder a 16 diszkrét érték helyett – amivel hard decision módszer esetén dolgozna – valószínűségi változóként kezel minden PAM16 szimbólumot, és ezek eloszlását a vonalon vett, finom felbontású jelszint szerint határozza meg. Ezekhez a valószínűségi változókhöz próbál meg iterációs lépésekben olyan lehetséges értékeket keresni, aminek egyrészt nagy a valószínűsége, másrészt helyes LDPC paritást ad.

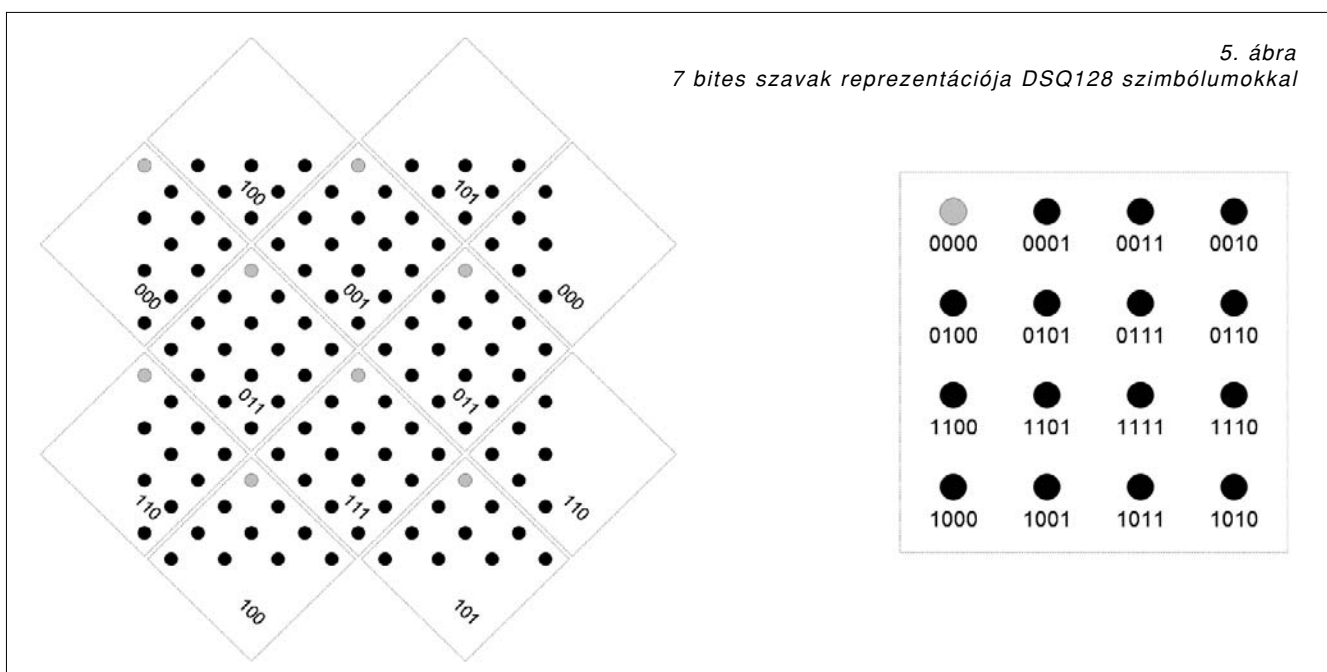
A vevő a potenciális dekódolt blokkot a 8 bites CRC-vel ellenőrzi, hogy vajon valóban sikerült-e a dekódolás, vagy csak véletlenül talált egy olyan bitsorozatot, amire az LDPC kód stimmel.

4.4.3. PMA

A csavart érpáros Physical Medium Attachment alrétteg főként DSP (Digital Signal Processing) technológiákra épül.

Egyik fontos eleme adáskor a Tomlinson-Harashima Precoder használata. Ebben az eljárásban a szomszédos 10GigabitEthernet állomások a csatorna karakterisztikáját speciális teszt szimbólumok segítségével vizsgálják, majd adatátvitelkor equalizerek segítségével úgy torzítják a leadott jelet, hogy az minél jobban kompenzálja a csatorna előzetesen megállapított torzítását.

Az alkalmazott jelfeldolgozási módszerek másik lényeges csoportja a különféle visszhangok és áthallások miatt keletkező zavarok elnyomása. Visszhangok egyrészt a hibrid áramkörök, másrészt a különböző csatlakozókon keletkező reflexiók miatt adódnak. Áthallás keletkezik az adott kábel többi érpárja, valamint a közelben vezetett többi kábel szórt energiájából. Jól kezelhető az áthallások közül például a közeli áthallás (NEXT, Near-End Cross-Talk), mely az ugyanazon kábel



másik érpárjain adott jelek szórt energiáiból származik, hiszen itt a zavaró jel jól ismert, így viszonylag könnyű kompenzálni azt.

4.4.4. Médium

A 10GigabitEthernethez használható csavart érpáros kábel a Cat6a (augmented Category 6) minőségű UTP (Unshielded Twisted Pair) kábel. A Cat6a specifikáció egyelőre nincs készen, de várhatóan még 2006-ban megjelenik.

Az előzetes specifikációnak megfelelő kábelek jellemzően négy fő újítással érik el a jobb átviteli képességet. Az érpárok csavarása sűrűbb, intenzívebb, azaz egységnyi hosszon a korábnál többször keresztezi egymást a pár két tagja. Az érpárok csavarásának intenzitása egymástól eltérő. Az érpárok térbeli pozicionálását rendszerint műanyag terelőidomok segítik a kábelben belül.

A kábel külső burkolata – és így az egész kábel – vastagabb, hogy a szomszédos kábel érpárjaitól való távolság nagyobb legyen. Ezek a változtatások mind a kábelben belüli és a kábelek közti áthallás, valamint az egyéb külső zavarok hatásának csökkenéséhez vezetnek.

4.4.5. A 10GBASE-T fennálló kihívásai

Az elért jelentős eredmények mellett a csavart érpáros 10GigabitEthernet létrehozásában fontos kihívást jelent egyrészt az interface komplexitása és hatékonysága közti megfelelő kompromisszum megtalálása, mind a különböző DSP komponensek, mind a dekóder esetén. Másrészt lényeges feladat az energiaigény leszorítása – ezen a téren mostanában a teljesítmény 10 W alá szorítása a cél.

Ezen fennálló problémák mellett az előrehaladást látva várható, hogy a 10GBASE-T specifikáció a közeljövőben elkészül, és utána hamarosan megjelennek az első termékek is.

4.5. Backplane Ethernet

Szándékosan nem szóltam idáig azokról az Ethernet fizikai rétegekről, amelyek a számítástechnikai készülékek belsejében az egyes modulok egymással való összekötésre kívánnak megoldást nyújtani, hiszen ez viszonylag távol áll az Ethernet megszokott adathálózati felhasználásától. De egy 10GigabitEthernetről szóló cikk nem lehet teljes ezek említése nélkül, hiszen backplane Ethernetről szóló IEEE 802.3ap tervezetben két 10GigabitEthernet fizikai réteg is szerepel.

A cél 1 méter távolság áthidalása az ilyen készülékek hátlapján jellemző környezetben. A kidolgozás alatt álló 10GBASE-KX4, illetve 10GBASE-KR ezt a korábban ismertetett (10GBASE-X vagy 10GBASE-R) jellemzőkkel valósítja meg.

5. Összefoglalás

Az Ethernet család a 10GigabitEthernet megjelenésével ismét egy nagyságrenddel gyorsabb tagokkal bővült a korábbiakhoz képest. Az IEEE 802.3 szabvány fejlődésében is megfigyelhető az a trend, ami az Ethernet alkalmazásában is, hogy az Ethernet egyre nagyobb teret hódít a LAN felhasználáson kívül is. Az átviteli közegeket tekintve a 10GigabitEthernet paletta hasonló, bár valamivel gazdagabb, mint a GigabitEthernet, és az egyes médiumok megjelenésének sorrendje – tükrözvén az adott probléma nehézségét, bonyolultságát – is hasonló.

A jövőt tekintve fontos eseménynek ígérkezik a 10GBASE-T eszközök megjelenése. Az azt követő jelentős újdonság talán az átviteli sebesség ismételt növelése lehet, de egyelőre nem tudni, hogy a következő lépcső a 40 vagy a 100 Gbit/s lesz-e.

Irodalom

- [1] „IEEE Std 802.3-2005, Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications”, 2005. IEEE, ISBN 0-7381-4741-9
- [2] „IEEE Draft P802.3aq/D3.1, Amendment: Physical Layer and Management Parameters for 10 Gb/s Operation, Type 10GBASE-LRM”, 2006. IEEE, ISBN 0-7381-1799-4
- [3] Abott, J.S., et al.: „Analysis of Multimode Fiber Behavior with Laser Sources in the Development of the Gigabit Ethernet Fiber Optic Specification”, 1998. IWCS Proceedings
- [4] „IEEE Draft P802.3an/D3.1, Amendment: Physical Layer and Management Parameters for 10 Gb/s Operation, Type 10GBASE-T”, 2006. IEEE, ISBN 0-7381-4889-X
- [5] Gallager, R.G.: „Low-Density Parity-Check Codes”, 1963. The MIT Press, Cambridge, MA, <http://web.mit.edu/gallager/www/pages/ldpc.pdf>
- [6] Langner, P., Woodruff, B., Kohl, B.: „10 Gigabit Ethernet on Unshielded Twisted-Pair Cabling, Moving 10 Gigabit Ethernet into a Volume Platform”, DesignCon 2006, http://www.ethernetalliance.org/technology/white_papers/