

híradástechnika

1945 VOLUME LXII. 2007

hírközlés ■ informatika



Szolgáltatásminőség

QoS – QoSE

Szolgáltatásminőség DSL rendszerekben

GMPLS

2007/4

**A Hírközlési és Informatikai Tudományos Egyesület folyóirata
a Nemzeti Hírközlési és Informatikai Tanács együttműködésével**

Tartalom

SZOLGÁLTATÁSMINŐSÉG	1
Nándorfiné Somogyvári Magdolna QoSE: négy betű, vagy annál sokkal több?	2
Gál Zoltán, Balla Tamás A QoS hatása az infokommunikációs alkalmazásokra	7
Vonderviszt Lajos Szolgáltatásminőség és a szolgáltatás minősége – szempontok az Internet-elérési szolgáltatások értékeléséhez	17
Nagy Tibor Minőségi szolgáltatások ADSL környezetben	23
Kőrösi Attila, Székely Balázs, Lukovszki Csaba, Dang Dihn Trang DSL hozzáférési hálózatokban alkalmazott csomagütemezők sorbanállási modellezése és analízise teljes és részleges visszautasítás esetére	29
Zahemszky András, Császár András, Tóth Gábor, Takács Attila Átjáró szerver választása a GMPLS PCE-architektúrában	34
Tóth Katalin, Schulcz Róbert, Imre Sándor Ütközésfeloldás RFID rendszerekben	39
Kovács Benedek, Fülöp Péter Mobilitás-menedzsment algoritmusok numerikus vizsgálata	47
Horváth László S. Kari, Károly, Charles, Karel...	54

Címlap: A Nemzetközi Űrállomás (ISS) 15. expedíciójának startja az oroszországi Bajkonurból, a fedélzetén Charles Simonyival.

Védnökök

SALLAI GYULA a HTE elnöke és DETREKŐI ÁKOS az NHIT elnöke

Főszerkesztő

SZABÓ CSABA ATTILA

Szerkesztőbizottság

Elnök: ZOMBORY LÁSZLÓ

BARTOLITS ISTVÁN
BÁRSONY ISTVÁN
BUTTYÁN LEVENTE
GYŐRI ERZSÉBET

IMRE SÁNDOR
KÁNTOR CSABA
LOIS LÁSZLÓ
NÉMETH GÉZA
PAKSY GÉZA

PRAZSÁK GERGŐ
TÉTÉNYI ISTVÁN
VESZELY GYULA
VONDERVISZT LAJOS

Szolgáltatásminőség

vonderviszt.lajos@nhh.hu

A szolgáltatásminőség, illetve az egyes minőségi paraméterek kutatása a hírközlési hálózatok fejlesztésének és üzemeltetésének egyre fontosabbá váló kérdése, különös tekintettel a minőségi garanciákat alapfilozófiájukat tekintve nem tartalmazó IP alapú technológiák terjedésére. E számunk főtémájaként öt cikk foglalkozik a szolgáltatásminőség különböző vonatkozásaival.

Nándorfiné Somogyvári Magdolna cikkében az érzékelt szolgáltatásminőség kérdését állítja a középpontba, megvizsgálva a fogalom eredetét és az IPTV terjedésében játszott szerepét. A szolgáltatók előszeretettel adnak meg könnyen mérhető vagy névleges szolgáltatásminőségi paramétereket, miközben a felhasználók számára az érzékelt szolgáltatásminőség jelenti az „igazi” szolgáltatásminőséget. A hazai és nemzetközi példák azt mutatják, hogy megtörténtek az első lépések a felhasználói szempontok érvényesítése érdekében, azonban még további kutatások és fejlesztések szükségesek ahhoz, hogy kielégítően lehessen a szolgáltatás minőségi paramétereiből (QoS) az érzékelt szolgáltatásminőséget megjósolni.

Vonderviszt Lajos cikke az érzékelt szolgáltatási minőség (QoSE) és a szolgáltatásminőségi paraméterek (QoS) közötti összefüggésekre világít rá. Példákon keresztül megmutatja a fogyasztók érdekében tett hazai és nemzetközi lépéseket, valamint a fejlődési irányokat.

Gál Zoltán és Balla Tamás az interneten hagyományosan működő hálózati alkalmazások viselkedését elemezték egyetlen QoS tartományon belül szabályozott paraméterek függvényében, a legelterjedtebb második rétegbeli hálózati protokoll (ethernet) feletti IP és további protokollok (TCP, UDP) esetében.

Nagy Tibor a szolgáltatások minőségének kérdését a felhasználó oldali (ADSL) kapcsolat környezetében vizsgálta, különös tekintettel a hálózati szolgáltatásban résztvevő eszközök hangolási, beállítási kérdéseire.

Kőrösi Attila és társainak cikke DSL hozzáférési hálózatokban alkalmazott csomagütemezők sorbanállási modellezésével és analízisével foglalkozik. A QoS biztosítása érdekében különböző prioritású osztályba sorolt csomagok a kiszolgálás során a különböző osztályokhoz rendelt pufferekben tárolódnak és többnyire az osztályok között szigorú prioritásos ütemezést használnak. Egy ilyen szigorú prioritáson alapuló DSL ütemező teljesítmény analízisét végezték el a szerzők.

Számunk további, beküldött cikkei közül Zahemszky András és társainak cikke bemutatja, hogy a GMPLS új útvonalválasztó architektúrájára, a Path Computation Element-re épülő modell hogyan alkalmazható átjáró szerverek választására, hogy elkerüljük a szerverek túlterheléséből adódó hívás blokkolást, és minimalizáljuk a transzport hálózat összerhelését. A bemutatott algoritmusok figyelembe veszik a szerverek terheltségét, illetve a hálózati topológiát az átjáróválasztásnál, melyeket a szerzők hálózatszimuláció segítségével hasonlítanak össze.

Tóth Katalin, Schulcz Róbert és Imre Sándor cikke az RFID környezetben felmerülő többes leolvasási problémák megoldására rendelkezésre álló ütközésfeloldó algoritmusok ismertet, kitérve az olvasók és a transzponderek ütközésére egyaránt. Szimulációs eredményeket közöl az ISO 18000-3 Mode 1 ütközésfeloldásra használt algoritmusára.

Fülöp Péter és Szálka Tamás cikkében különböző mobilitási stratégiákat tanulmányoznak a hálózati jelzésforgalom, az egyes csomópontokban fellépő feldolgozás és a rádiós interfész költsége alapján. A végső cél olyan irányelvek létrehozása, amelyek segítséget nyújtanak újgenerációs, mobil hálózati menedzsmentrendszer megtervezéséhez.

Vonderviszt Lajos
vendégszerkesztő

Nemzeti Hírközlési Hatóság

Szabó Csaba Attila
főszerkesztő

QoSE: négy betű, vagy annál sokkal több?

NÁNDORFINÉ SOMOGYVÁRI MAGDOLNA

Nemzeti Hírközlési Hatóság
nandorfi@nhh.hu

Kulcsszavak: szolgáltatásminőség, felhasználó által tapasztalt QoS, elvárás, elégedettség, IPTV

Ebben a cikkben megvizsgáljuk a QoSE (Quality of Service Experienced = a felhasználó által tapasztalt szolgáltatásminőség) fogalom eredetét, mérhető jellemzőit, azok meghatározását egy példa segítségével. Valamint megtaláljuk azokat a tényezőket, amelyek révén ez a fogalom az érdeklődés középpontjába került. Szintén példa alapján megmutatjuk, hogy miként tudja ez a fogalom egy egyre népszerűbbé váló szolgáltatás, az IPTV szolgáltatónak a versenyelőnyért folytatott küzdelmét segíteni.

1. A QoSE eredete

Az 1. ábrán látható modell az ETSI-től származik (ETR 003). A 90-es években hozták létre. A szolgáltatás minőségét részben a szolgáltató, részben a szolgáltatást használó részéről vizsgálja. A felhasználónak vannak elvárásai egy szolgáltatással kapcsolatban. A szolgáltató ezeket is figyelembe véve tervezi meg szolgáltatását, amit felajánl a felhasználóknak. A szolgáltató által megvalósított és nyújtott szolgáltatás már különbözik a tervezettől. Ezt a kapott szolgáltatást érzékeli és tapasztalja a felhasználó. A szolgáltató által tervezett és felajánlott szolgáltatás objektív és mért paramétereivel jellemezhető. A felhasználó követelményei verbálisan megfogalmazottak és az érzékelt és megtapasztalt szolgáltatás minősége szubjektív paraméterekkel jellemezhető és szubjektív módszerekkel mérhető. Ezt a modellt később az ITU is átvette és következetesen alkalmazza (2001 ITU-T Recommendation G.1000).

2007 februárjában az ITU 2-es számú Tanulmányi Bizottsága elfogadta a korábban QoE rövidítéssel és Quality of Experienced megnevezéssel használt mutatót, melyet főleg a multimédiás gyakorlatban alkalmaznak és jelölését QoSE-re változtatta, ez a rövidítés tehát a QoS felhasználó által tapasztalt értékét jelenti, amely azonosan egyenlő a fenti modellben szereplő érzékelt szolgáltatásminőséggel és mérési módszereik is azonosak.

Az ITU által elfogadott definíció a következő:

QoSE = egy alkalmazásnak vagy szolgáltatásnak teljes elfogadottsága, ahogy azt a végfelhasználó szubjektíven érzékeli vagy érzékelt.

Megjegyzések:

1. A QoSE tartalmazza a teljes végpontok közötti rendszer hatásait (végfelhasználó, végberendezés, hálózat, infrastruktúra stb.),

2. A teljes elfogadottságot befolyásolhatják a felhasználók elvárásai és a körülmények.

Fontos megállapítás, hogy az ETSI/ITU modellnek a „a felhasználók által érzékelt QoS” eleme azonosan egyenlő a fent definiált „felhasználó által tapasztalt QoS”-sel.

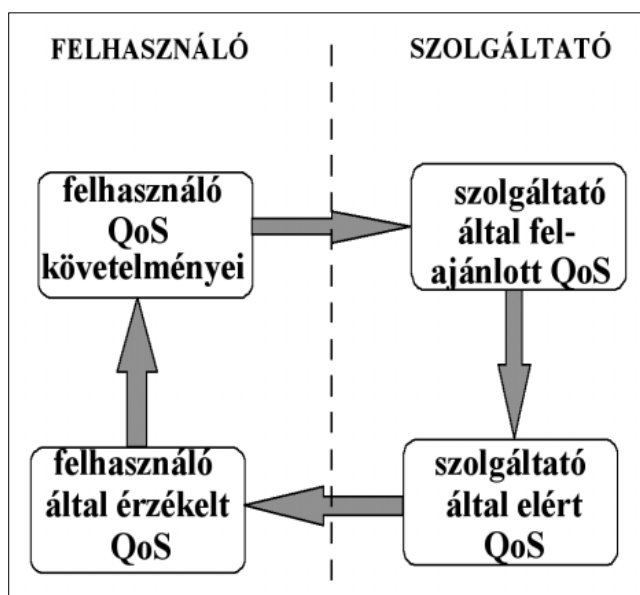
2. A QoSE mérhető jellemzői

A minőség és így a QoS megítélése, meghatározása jellege szerint lehet objektív és szubjektív. A minőség objektív értéke jellemző paramétereinek fizikai mérése útján határozható meg. A minőség szubjektív értéke számszerűsíthető mutatóinak meghatározásával, felhasználói vélemények alapján számítható. Röviden csak objektív és szubjektív szolgáltatásminőségről (QoS) szoktunk beszélni.

A szubjektív szolgáltatásminőség Lewis és Booms definíciója szerint annak mértéke, hogy a szolgáltatás milyen mértékben illeszkedik a felhasználó követelményeihez.

A szubjektív minőség mérhető jellemzői egyrészt a felhasználó QoS követelményei (lásd az 1. ábrát). Ezt a jellemzőt *elvárásnak* is szokás nevezni. Az elvárás a felhasználó elképzelése arról, hogy mit kell, hogy nyújtson a szolgáltatás vagy, hogy mit képes nyújtani a szolgáltatás.

1. ábra Az ETSI QoS-modellje



A szubjektív minőség másik mérhető jellemzője a *felhasználó által érzékelt QoS*, illetve az ITU-T jelöléseit használva: QoSE, amit másképpen *elégedettségnek* nevezünk. A felhasználó elégedettsége, azaz az általa érzékelt minőség Parasuraman és Lewis szerint a szolgáltatással kapcsolatos elvárások megerősítése, illetve meg nem erősítése, vagy Smith és Houston szerint az éppen aktuális szolgáltatás teljesítése és a felhasználók elvárásainak összehasonlításából adódik.

A szolgáltatások szubjektív minőségének leghatékonyabb vizsgálati módszerét három amerikai kutató, Parasuraman, Zeithaml és Berry dolgozta ki. Csoportokat hoztak létre különböző szolgáltatások igénybevevőiből, és az így nyert adatok elemzése alapján fogalmazták meg az úgynevezett SERVQUAL módszer lényegét. A SERVQUAL tíz alapvető dimenziót fogalmaz meg, amelyek a felhasználó elvárásait fejezik ki. Ezek az alábbiak: megbízhatóság, reagálási készség, szakértelem, udvariasság, hitelesség, biztonságérzet, elérhetőség, kommunikáció képesség, egyedi probléma-kezelés, kézzelfoghatóság/dologi tényezők. Vagyis a felhasználó elvárásai részben a szolgáltatásra, részben a szolgáltatóra, vagy annak személyzetére vonatkoznak.

Zeithaml, Parasuraman és Berry ezen tíz dimenzió alapján az elvárásokat az alábbi 5 paraméterrel jellemzik:

1. **Megebízhatóság:** magában foglalja a teljesítmény és igénybevehetőség állandóságát. Azt jelenti, hogy a szolgáltató tiszteletben tartja ígéreteit, pontos a számításban, a kért szolgáltatást időben teljesíti,

2. **Fogékonyság/érzékenység:** azt jelenti, hogy a szolgáltató akarja és kész arra, hogy biztosítsa a szolgáltatást és kellő segítséget nyújtson. Figyel a pontos időbeli teljesítésekre, vagyis az ügyletek esetleges időbeli csúszását azonnal rendezzi, ha szükséges, akkor az ügyfelet azonnal visszahívja és a megrendeléseket a lehető leggyorsabban teljesíti,

3. **Biztonságosság:** egyrészt azt jelenti, hogy maga a nyújtott szolgáltatás biztonságos, valamint, hogy a szolgáltató kialakítja a felhasználó biztonságérzetét azáltal, hogy hozzáértő, udvarias, hiteles,

4. **Empátia:** a szolgáltató megérti a felhasználó elvárásait és azokat, valamint a felhasználó egyéni kívánságait is tiszteletben tartja, jól kommunikál a felhasználóval, annak kérdéseit jól és gyorsan megválaszolja, kelőképpen tájékoztatja,

5. **Kézzelfoghatóság:** azt jelenti, hogy az alkalmazott berendezések korszerűek, jól menedzselik azokat, ügyelnek mind az eszközök, mind a személyzet külső megjelenésére az ügyféllel való találkozáskor, gondoskodnak mind a szolgáltatás, mind a szolgáltató megfelelő helyen és időben és kis várakozással való elérhetőségéről.

A szerzők a modell leírásakor a jellemzők méréséről és a mért eredmények kiértékeléséről is beszámolnak a hivatkozott irodalomban [3].

Sok mérés tapasztalataiból (amelyek különböző szolgáltatásokra vonatkoztak és különböző számú megkérdezett válaszait tartalmazzák) Parasuraman a fenti öt paraméter relatív fontosságáról az alábbiakat állapította meg:

A válaszadók számára a szolgáltatásminőség fenti öt mutatójának fontossága az alábbi eloszlást mutatta:

- | | |
|----------------------|-----|
| 1. Megbízhatóság: | 32% |
| 2. Fogékonyság: | 22% |
| 3. Biztonságosság: | 19% |
| 4. Empátia: | 16% |
| 5. Kézzelfoghatóság: | 11% |

A fontosság szintén egy lényeges jellemző a szubjektív QoS szempontjából, mert a gyakorlatban az elégedettséget a fontosság függvényében azzal együtt szokták vizsgálni. A módszerre példát a következő fejezetben mutatunk.

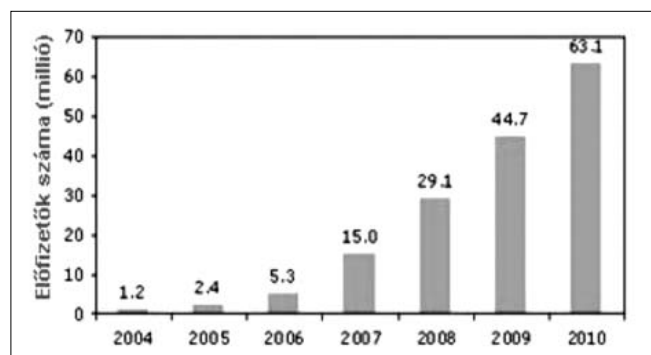
3. A QoSE a gyakorlatban

A szolgáltatásminőség felhasználó általi megítélése egyre inkább érdekli a szolgáltatókat, mert az utóbbi években számos új szórakoztató célú hírközlési szolgáltatás terjedt el a piacon és nagy a küzdelem a felhasználóért. Ezen szolgáltatások közé tartozik az IPTV is.

Az iSuppli amerikai piacelemző csoport szerint az IPTV szolgáltatások előfizetőinek száma várhatóan 26-szorosára növekszik 2005-2010 között, és az évtized végére meghaladhatja a 63 milliót. Ez azt jelenti, hogy az évi átlagos növekedési ütem több mint 92% a 2005-ös 2,4 millióról indulva, ahogy azt a 2. ábra mutatja.

2. ábra

Az IPTV-szolgáltatás előfizetőszámának előrejelzése



Az iSuppli az IPTV szolgáltatás fejlődését három fázisra osztja. A jelenlegi IPTV piac az első fázis elején tart, az alapszolgáltatás telepítése folyik. A második fázis már értéknövelt és interaktív szolgáltatások sorát adja az alapszolgáltatáshoz. A harmadik fázis erőteljes fejlődést hoz majd az integráció és interaktivitás terén.

Kétségtelen, hogy a technológia és a szélessávú Internet-hozzáféréshez kapcsolódó fejlesztések az elkövetkező években drámaian meg fogják változtatni a TV-nézési szokásokat. A fogyasztók nemcsak azt fogják eldönteni, hogy mikor, mit és hol akarnak megnézni, de olyan személyreszabott programokat élvezhetnek majd, melyek magas szintű interaktivitást kínálnak.

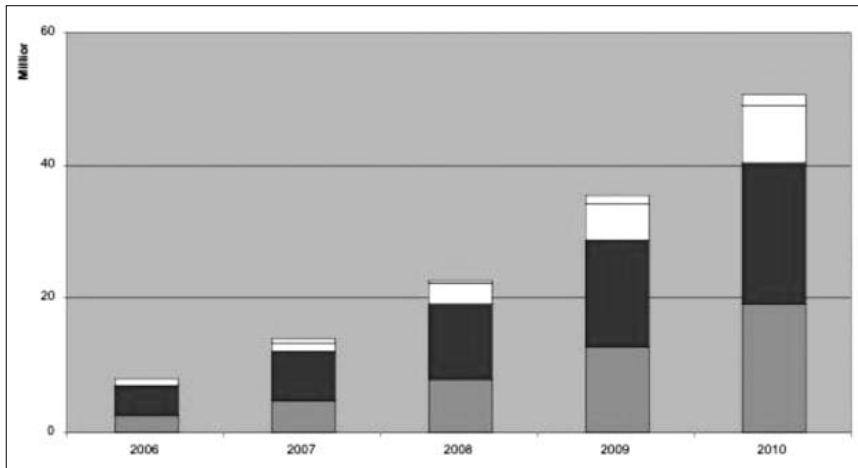
Várhatóan hogyan alakul a világ globális IPTV piaca?

Az IPTV előfizetőszámok alakulásának előrejelzését kontinensenként a 3. ábra mutatja. Az oszlopokban

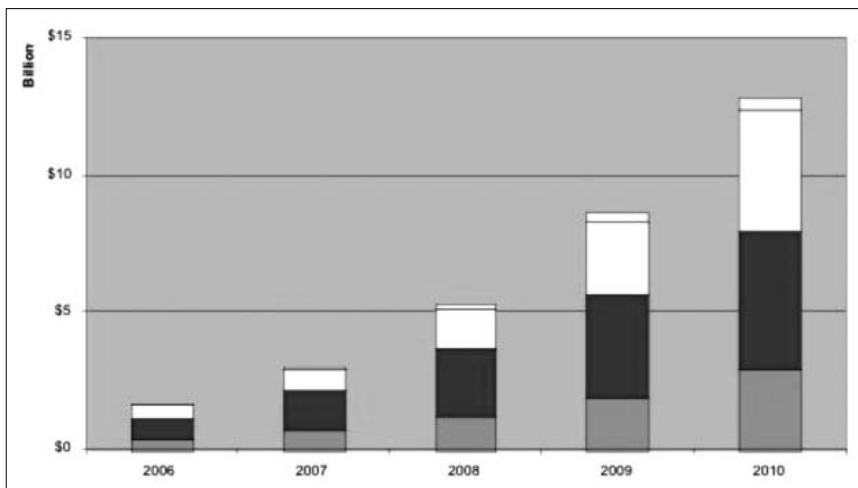
alulról felfelé sorrendben a európai, ázsiai, észak-amerikai előrejelzett adatokat látjuk. A legfelső réteg a Föld összes többi területére vonatkozó IPTV-előfizetőszámot prognosztizálja.

Az IPTV-szolgáltatásból származó bevétel alakulásának előrejelzését kontinensenként a 4. ábra mutatja.

3. ábra Az IPTV előfizetőszámok alakulásának előrejelzése kontinensenként



4. ábra Az IPTV-szolgáltatásból származó bevétel alakulásának előrejelzése



Az előfizetők számát és a bevételeket tekintve a globális IPTV piacon korábban az európai volt a legerősebb, most Ázsia mutatja a leggyorsabb növekedést az összes kontinens közül és ez év végére itt lesz a legnagyobb az előfizetők száma.

Ami az európai helyzetet illeti, várhatóan az IPTV Franciaországban, Spanyolországban és Olaszországban a fizetős piac 16-20%-t fogja elérni 2010-re. Ugyanerre az időpontra Angliában ez az érték körülbelül 7,5% lesz.

Magyarországon 2006-ban kezdte meg két hazai Internet-szolgáltató IP-alapú televíziós szolgáltatások biztosítását. Az új technológia a tévézési lehetőségén túl olyan emelt szintű tartalmi szolgáltatásokat is nyújt, mint a digitális videotéka, az elektronikus műsorújság, a szünetfunkció, vagy a gyerekzár.

A T-Online tavaly év végén startolt IPTV-szolgáltatása Budapesten és több vidéki kis- és nagyvárosban

(például Debrecenben, Mosonmagyaróváron, Nyíregyházán, Pécsen, Székesfehérváron és Zalaegerszegen) már összesen közel ezer felhasználóval rendelkezik. Az újfajta tévézést kínáló T-Home tavasszal további, az igény szerinti tévézés élményét kínáló kényelmi funkciókkal bővíti.

A nagy cégek az IPTV szolgáltatást általában a triple-play szolgáltatás részeként ajánlják.

A továbbiakban egy példát közlünk egy feltételezett QoSE felmérésről és annak kiértékeléséből levonható következtetésekről. Sajnos a szolgáltatás, mint láttuk, korai fejlődési fázisában van és így nagy mennyiségű adatot feldolgozó tényleges kutatások még nem állnak rendelkezésünkre.

A következőkben bemutatásra kerülő példák közül az első olyan esetre vonatkozik, amikor a felhasználókat arról faggatjuk, hogy az IPTV szolgáltatás jellemzőit milyen fontossági sorrendbe állítják és ezekkel a jellemzőkkel mennyire elégedettek.

A második példában ugyanezeket a kérdéseket tesszük fel, de a megkérdezettek három különböző szolgáltató ügyfelei.

1. példa

Az IPTV szolgáltatás különböző szolgáltatáselemeinek fontossága a felhasználók véleménye szerint az 1. táblázatban látható.

A táblázat első oszlopa a szolgáltatáselemeket tartalmazza. A szolgáltatáselemek azok a szolgáltatásjellemzők, amelyek fontosak a felhasználó számára a szolgáltatás használata során.

A táblázat adatai a következő kérdésre adott válaszok: „Hogyan tudja jellemezni az IPTV szolgáltatás egyes szolgáltatás elemeinek fontosságát?” Egy-egy válaszadó az egyes kérdésekre 1 és 5 közötti egész számot adhat, az alábbi megfeleltetés szerint: 5 = nagyon fontos, 4 = fontos, 3 = közepesen fontos, 2 = nem nagyon fontos, 1 = egyáltalán nem fontos.

Az adott egész pontszámok számtani átlaga van a második oszlopban. A harmadik oszlopban szereplő top (%) azt jelenti, hogy a válaszadók hány százaléka tette az illető szolgáltatás elemét az első helyre.

A felhasználók elégedettsége (QoSE) az IPTV-vel (az IPTV szolgáltatás különböző szolgáltatás elemeivel) a 2. táblázatban látható.

A táblázat adatai a következő kérdésre adott válaszok: „Mennyire elégedett az IPTV szolgáltatással, annak alábbi szolgáltatás elemeivel?” A táblázat adatai iránymutatóak azon IPTV szolgáltató számára, amely-

nek ügyfelei körében történt ez a véleménykutatás. Megmutatja számára, hogy milyen szolgáltatás elemek területén kell fejlesztenie, javítania.

2. példa

Egy más típusú véleménykutatás más kérdésekre ad választ. Ugyanezeknek a szolgáltatáselemeknek a vizsgálatát elvégezve, három egymással versenyben álló IPTV szolgáltató esetén a fontosságok és elégedettségek ismeretében az 5. ábrát kapjuk.

Az itt következő ábrázolási mód, elnevezések és a mögöttes filozófia a Bell Research-tól ered. Az y tengelyen ábrázoljuk a növekvő fontosságot. A maximális fontosság 4,89, a minimális 2,36. Az elégedettséget (QoSE) az x tengelyen találjuk. A maximális elégedettség 4,71, a minimális 3,38. Az így kapott téglalapban ábrázolunk 42 pontot (14 szolgáltatáselemre vonatkozó fontosság és elégedettség, mint koordináták 14 pontot határoznak meg egy szolgáltató esetén, de három versengő szolgáltató lévén, 42 pontot kapunk). Az átlagos fontosság és elégedettség értékénél a tengelyekkel párhuzamos egyenesek a téglalapot négy szektorra bontják: a jobb felső szektorban lévő pontok az alapkövetelmények teljesülését jelzik, ugyanis itt található az az az szolgáltatáselemek, amelyeket nagyon fontosnak tekintettek az ügyfelek és egyúttal nagyon elégedettek is

1. táblázat

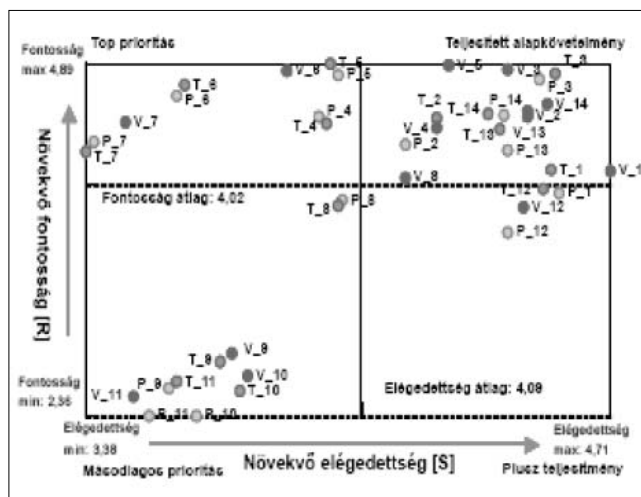
A különböző szolgáltatáselemek fontossága

Szolgáltatáselemek	átlagérték (pontszám)	top (%)
Mozgókép-minőség	4,86	94,4
Állókép-minőség	4,81	94,3
Hangminőség	4,73	90,2
Beruházási költség	4,54	86,6
Díjak	4,48	83,5
Díjfizetési módszer	4,44	83,0
Számlázás	4,37	81,8
Távírányítás	4,31	76,1
Freemail -képesség	4,08	67,8
Teletex	3,90	66,5
Időeltolási funkció	3,84	64,3
Szolgáltatóváltás	3,68	28,7
Iwiw-elérhetőség	3,47	24,7

2. táblázat

A felhasználók elégedettsége az IPTV szolgáltatás különböző szolgáltatáselemeivel

Szolgáltatáselemek	átlagérték (pontszám)	top (%)
Mozgókép-minőség	4,06	72,8
Állókép-minőség	4,53	88,2
Hangminőség	3,67	53,2
Beruházási költség	4,46	73,7
Díjak	4,04	60,5
Díjfizetési módszer	4,28	73,4
Számlázás	4,46	83,9
Távírányítás	3,40	42,1
Freemail-képesség	4,60	85,5
Teletex	4,05	61,7
Időeltolási funkció	4,51	84,5
Szolgáltatóváltás	3,66	23,5
Iwiw-elérhetőség	3,54	23,7



5. ábra Szolgáltatáselemek a fontosság-elégedettség koordináta-rendszerben

voltak velük. A bal felső doboz jelzi a top-prioritásokat, mert itt olyan szolgáltatáselemek találhatók, amelyeket nagyon fontosnak tartottak a felhasználók, de nem voltak velük nagyon elégedettek. A bal alsó téglalap a másodlagos prioritással bíró szolgáltatáselemeket tartalmazza. Az elnevezés onnan ered, hogy ebben a térfélben lévő pontok alacsony fontosságúak. Végül a jobb alsó doboz elemeire jellemző, hogy ezekbe a szolgáltatás elemekbe plusz energiákat fektet be a szolgáltató, ugyanis bár az ügyfelek nem tekintik nagyon fontosnak, mégis nagyon elégedettek velük.

Az egyes pontokat a megfelelő szolgáltató megjelölésével paraméterezve, megkönnyítik a potenciális új ügyfelek számára a szolgáltatóválasztást.

4. Összefoglalás

Az előzőek is bizonyítják, hogy a QoSE hajtóerő a hírközlési szolgáltatások piacán. Meghatározása számos értékes információt nyújt a szolgáltatónak és döntési helyzetben segítségül szolgál a felhasználó számára.

Irodalom

- [1] ITU-T, SG 2: Progress Report for Q.4/2, Geneva, 30 January-8 February 2007.
- [2] ITU-T Recommendation G.1000, Communications Quality of Service: A framework and definitions 2001.
- [3] Parasuraman, TRI/SERVQUAL/LibQUAL+™ University, Miami, 2002.
- [4] IPTV Global Forecast – 2006 to 2010, 2006 október.
- [5] IT3 Tanulmány – dr. Bartolits István, Technológiai jelenségek részletes elemzése; NHIT kiadvány, 2005.
- [6] Dr. Hosszú Gábor, Az internetes kommunikáció informatikai alapjai.
- [7] Bell Research, Magyar Infokommunikációs Jelentés 2005.

A QoS hatása az infokommunikációs alkalmazásokra

GÁL ZOLTÁN, BALLA TAMÁS

Debreceni Egyetem TEK, Információtechnológiai Központ
zgal@unideb.hu, ballat@delfin.unideb.hu

Kulcsszavak: QoS, IPv4, IPv6, TCP, UDP, torlódás, jitter, VoIP, codec, H.323, H.261/H.264, ATM

A hálózati szolgáltatásminőség (Quality of Service) olyan funkció, amely segítségével a forgalom kezelése történik az alkalmazói programok számára. Ehhez alapvető forgalomkezelési mechanizmusokra, valamint ezeket ellenőrző algoritmusokra van szükség. A QoS-funkcionalitás egyrészt a hálózati alkalmazásokat, másrészt pedig a hálózati adminisztrátorokat szolgálja ki. Addig, amíg a hálózati adminisztrátor korlátozza az erőforrásokat, az alkalmazás az erőforrások minél szélesebb körét próbálja igénybe venni.

A QoS az Internet technológiák környezetében a VoIP (Voice over IP) megoldás telefonbeszélgetés költségcsökkentő hatásának következtében jelent meg és terjed el egyre inkább. Az Internet- és intranet-alapú új, sávszélességigényes alkalmazások, valamint az adat-, hang- és videoforgalom IP-infrastruktúra feletti konvergenciája ugyancsak a QoS iránti igényeket hangsúlyozza. A jelenlegi alkalmazások több mint 95%-a Ethernet csomópontokban végződik, így a csomagok ezen átviteltechnikán való homogén továbbítása költségcsökkentést jelent, mivel nem szükséges protokoll-konverzió az adatok továbbítása során. A cikkben a QoS mechanizmus L2 és L3 rétegekben kifejtett hatását vizsgáljuk meg egyetlen QoS tartományon belül szabályozott paraméterek segítségével, H.261 és H.264 videó codec alkalmazása mellett. Konkrét mérések alapján az Interneten hagyományosan működő hálózati alkalmazások viselkedését tanulmányozzuk néhány QoS paraméter módosítása esetén. Vizsgáljuk az UDP és a TCP hálózati erőforrás kihasználását és számszerű mérési módszert javasolunk a videóműsor minőségének elemzésére, valamint megvizsgáljuk, hogy IEEE 802.3 környezetben milyen feltételek mellett képesek a valós idejű és a hagyományos adatátviteli szolgáltatások együttműködni.

1. Bevezetés

A hálózatba kapcsolt számítógépes alkalmazások legegyszerűbb megközelítése szerint az alkalmazói program a másik gépen futó alkalmazói programmal úgy kommunikál, hogy az operációs rendszernek adja át az adatokat. Ahogy az adat az operációs rendszerhez jut, hálózati forgalmat generál. A hálózati szolgáltatásminőség (QoS) a hálózat azon tulajdonsága, amely segítségével a forgalom kezelése történik az alkalmazói program számára. Ehhez alapvető forgalomkezelési mechanizmusokra, valamint ezeket ellenőrző algoritmusokra van szükség. A QoS-funkcionalitás egyrészt a hálózati

alkalmazásokat, másrészt pedig a hálózati adminisztrátorokat szolgálja ki. Addig amíg a hálózati adminisztrátor korlátozza az erőforrásokat, az alkalmazás az erőforrások minél szélesebb körét próbálja igénybe venni. A QoS az Internet-technológiák környezetében a VoIP (Voice over IP) megoldás telefonbeszélgetés költségcsökkentő hatásának következtében jelent meg és terjed el egyre inkább [1]. Az Internet- és intranet-alapú új, sávszélesség igényes alkalmazás, valamint az adat-, hang-, videoforgalom IP infrastruktúra feletti konvergenciája ugyancsak a QoS iránti igényeket hangsúlyozza [2]. A jelenlegi alkalmazások több mint 95%-a Ethernet csomópontokban végződik, így a csomagok ezen átviteltechnikán való homogén továbbítása költségcsökkentést jelent, mivel nem szükséges protokoll-konverzió az adatok továbbítása során [3].

A különböző alkalmazások egymástól eltérő követelményeket támasztanak az adatforgalmat továbbító hálózat felé. A generált forgalom erőforrásigénye időben változó és általában szükséges, hogy a hálózat megfeleljen ennek az igénynek. Bizonyos alkalmazások többé vagy kevésbé toleránsak a forgalom késleltetésére, valamint a késleltetés változásra. Továbbá néhány alkalmazás képes elviselni korláton belül adatvesztést, míg mások nem. Ezek a követelmények a következő négy QoS-jellegű paraméter segítségével kerülnek kifejezésre. *Sávszélesség*: az alkalmazás forgalmának továbbítási sebessége; *lappangási idő*: az a késleltetés, amit egy alkalmazás a csomag kézbesítésénél képes elviselni; *jitter*: a lappangási idő szórása; *adatvesztés*: az elvesztett adatok százalékos aránya [4]. Ha végtelen méretű hálózati erőforrásaink lennének, akkor az alkalmazások forgalma a szükséges sávszélességen, nulla lappangási idővel, nulla jitterrel és nulla adatvesztéssel lenne jellemezhető. Mivel azonban a hálózati erőforrások korlátozottak, a rendszer bizonyos részein időtől függően az igények nem teljesíthetők. A QoS mechanizmusok az alkalmazások szolgáltatásigényének függvényében a hálózati erőforrások foglaltságát szabályozzák.

A hálózati végpontok közötti kapcsolatokhoz különböző hálózati eszközök szükségesek. Mindezek hálózati interfészekkel rendelkeznek, amelyek véges rátával képesek forgalmazni. Ha az adatforgalom olyan irányba halad, ahol az interfész továbbítási rátája kisebb, ott torlódás lép fel. Ezen jelenség kezelésére a köztes eszközök várakozási sorokat alkalmaznak, ezáltal lehetőség nyílik a forgalom eldobására, illetve a torlódás enyhítésére. Emiatt az alkalmazások változó lapangási időt, illetve adatvesztést tapasztalnak. Az interfészek adattovábbítási képessége, valamint a várakozási sor ideiglenes tárolási tulajdonsága az a két alapvető erőforrás, amely az alkalmazások forgalma számára biztosítani tudja QoS-t. A köztes eszközök működési mechanizmusa az egyes forgalmak számára tulajdonképpen meghatározza ezen erőforrásokhoz való hozzáférés sorrendjét, azaz a szolgáltatás minőségét.

Torlódás esetén az erőforrás-kritikus kapcsolatokhoz tartozó csomagok prioritást élveznek az egyéb csomagokhoz képest. Ehhez a köztes hálózati eszközöknek intelligens módon kell az erőforrásokat kezelniük. A különböző prioritások kialakításához az eszköz memóriája meghatározó szerepet játszik. Az erőforrások allokációjához szükséges a különböző típusú forgalmak azonosítása. A forgalom hálózati eszközhöz érkezésekor megtörténik a csomagok osztályozása és különböző adatfolyamokhoz való rendelése. Az eszközön belül minden egyes típusú adatfolyam a kimenő interfész egy-egy várakozási sorába kerül. Ezen várakozási sorok kezelését speciális mechanizmusok végzik, amelyek meghatározzák, hogy az egyes várakozási sorokból külön-külön milyen legyen az interfészen továbbított adatsebesség. A forgalmak típusának meghatározását és a várakozási sorok kezelését együtt az adatforgalom-kezelési mechanizmusok végzik. Fontos megjegyezni, hogy az adatfolyam több módon is definiálható. Egyik lehetséges mód a forrás és a cél logikai címe, a forrás és a cél socket-száma, valamint a session-

azonosító kombinációja. Másik lehetséges mód az adott alkalmazástól érkező adatok vagy adott interfészről érkező adatok beazonosítása. A gyakorlatban bármely típusú azonosítást alkalmazhatónak tekintik. A klasszikus hálózati alkalmazások jellemzőit, illetve ezek erőforrás igényének összefoglalóját az 1. táblázat tartalmazza [5].

A legfontosabb forgalomkezelő mechanizmusok az IEEE 802.1p, a DiffServ (Differentiated Service), az IntServ (Integrated Services), az ATM/ ISSLOW és mások. Ezek mindegyike speciális környezetben képes kifejteni hatását optimálisan.

Az IEEE 802.1p forgalomkezelő mechanizmus

A legtöbb LAN az IEEE 802 (Ethernet, FDDI, Token-Ring stb.) vagy más osztott közeget használó technológiára épül. Az IEEE 802.1p az L2 protokoll adatelem fejrészében egy mezőt alkalmaz, amelyben nyolc prioritás szint fér el. A végfelhasználói csomópontok és a routerek a LAN-ba küldött forgalom kereteiben megadják a prioritás értékét. Az adatkapcsolati eszközök (switch, bridge) a kereteket a prioritásnak megfelelő várakozási sorok segítségével kezelik. A mechanizmus csak alhálózaton belül működik, különböző hálózatok között nem érvényesül.

A DiffServ forgalomkezelő mechanizmus

Ez egy OSI 3 szintű QoS mechanizmus, amelyet annak ellenére, hogy több éve létezik, csak az utóbbi időben kezdtek el alkalmazni. A DiffServ az L3 protokoll adatelem fejrészében DSCP (DiffServ CodePoint) nevű mezőt helyez el. A végfelhasználói csomópontok és a routerek a DiffServ hálózatba küldött forgalom minden egyes csomagját a megfelelő DSCP értékkel látják el. A DiffServ/hálózatban lévő routerek minden csomagra a DSCP érték alapján történő osztályozás szerint specifikus PHB (Per-Hop Behavior) várakozásisor-kezelő algoritmust vagy ütemezőt alkalmaznak. Például az EF (Expedited-Forwarding) PHB limitált adatrátát esetén a

1. táblázat

A hálózati forgalmak jellemzői

Forgalom-típus	Jellemző	Hálózati erőforrásigény	Alkalmazásigény
Hang	Párbeszédes, borsztős. Kézbetűzési késleltetés <100 msec. Minden csomag sorrendben, jitter nélkül, de csomagvesztés lehetséges.	Csomagok routolása minimális késleltetéssel. Prioritást igényel. L4 szintű kezelhetőség.	Csomagok visszaállítása sorrendben, jittermérő puffer igénye, későn érkező csomagok eldobása.
Adat	Borsztős, nagyon hosszú, folyamatos küldés. Késleltetés elviselhető, de adatvesztés nem.	Csomagok routolása hibaellenőrzéssel L4 szinten. Hangsúlyozottan hibamentes késleltetés.	Csomagok visszaállítása hibaellenőrzéssel, hiba vagy túl sok késés esetén újraküldés.
Videó	Folyamatos. Csomagvesztés kevésbé tolerálható. Kismértékű késleltetés elfogadható, kivéve a videokonferencia, ahol az igények a hangéval azonosak.	Csomagok routolása hibamentesen.	Csomagok visszaállítása jitter mérő pufferrel. Videokonferencia esetén csomagok pufferelese és késleltetés csökkentése kései csomagok törlésével.
Játék	Borsztős adat, hang és folyamatos videó. Valós idejű kézbetűzés, csak alacsony mértékű csomagvesztéssel.	Csomagok routolása minimális késleltetéssel. Prioritást igényel. L4 szinten támogatás szükséges.	Csomagok visszaállítása sorrendben, jittermérő puffer, csökkentett késleltetés. Hiba vagy túl sok késés esetén újraküldés.

bemeneti és kimeneti pontok között nagyon alacsony lappangási időt biztosít. Más PHB rögzítheti bizonyos csomagok más csomagokhoz viszonyított relatív prioritását. Ez a prioritás vonatkozhat az átlagos átviteli sebességre, az eldobási sorrendre anélkül, hogy a lappangási időre megkötés létezne. A PHB-k a routerek önálló viselkedési szabályai. Kizárólagosan PHB-k segítségével nem lehetséges végponttól-végpontig típusú QoS garanciát nyújtani.

Előfordulhat olyan eset is, amelynél a végponttól-végpontig QoS szolgáltatást az útvonal menti összes routerben azonos PHB beállításokkal biztosítjuk. Ilyenkor a logikai kapcsolat béreltvonal jellegű összeköttetést ad, amely képes megfelelni akár interaktív hangkapcsolat vagy videó lejátszás számára is.

Az IntServ forgalomkezelő mechanizmus

Ez két modulból álló szolgálathalmaz, részei a Guaranteed Load, vagy Controlled Load (garantált, illetve ellenőrzött terhelés) szolgáltatások. A garantált szolgáltatás a forgalom számára kvantálható mértéket és korlátos lappangási időt biztosít. Az ellenőrzött terhelésű szolgáltatás megadott mértékű forgalom számára terheletlen hálózati környezetet emulál. Ezek a szolgáltatások kvantálhatók abban az értelemben, hogy bizonyos forgalom mennyiség számára szabályozható a QoS. Az IntServ-szolgáltatások többsége az RSVP jelzésrendszerre épül. Mindegyik IntServ-szolgáltatás beengedés-szabályozási algoritmusokat definiál, amelyek az adott eszközön befogadott forgalom mennyiségét határozzák meg anélkül, hogy romolna a szolgáltatás minősége. Az IntServ szolgáltatások nem használnak várakozásisor-algoritmusokat.

2. A QoS értelmezése az infokommunikációs rendszerekben

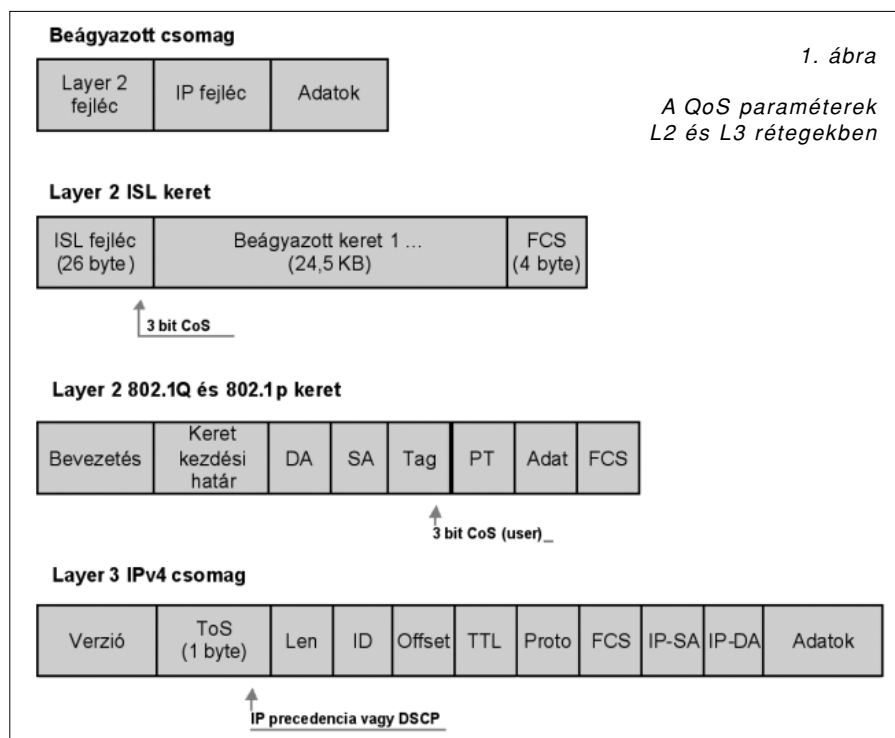
A hálózatok tipikusan átviteli kötelezettség nélkül (best-effort) kézbesítenek, ami azt jelenti, hogy minden forgalom azonos prioritású, és egyenlő esélye van arra, hogy a kézbesítése egy bizonyos időintervallumon belül megtörténjen. Torlódás esetén viszont minden csomagnak azonos esélye van az eldobásra is. Amikor QoS-t konfigurálunk, ki lehet választani azokat a specifikus hálózati forgalmakat, amelyeket prioritással kezelünk, majd ezekhez használhatunk torlódásvezérlési és torlódást elkerülő technológiákat. A QoS technológia használata a hálózat teljesítményét skálázhatóvá, a sávszélesség kihasználtságát pedig hatékonyabbá teszi.

Az IETF szabványai szerint a QoS alkalmazása leggyakrabban a DiffServ architektúrán alapul [6]. Ez előírja minden csomag osztályba sorolását a hálózaton belül. Az osztályozást az IP csomag fejlécében fenntartott hat bites szolgálat típus (TOS – Type of Service) mező értéke teszi lehetővé. Lehetséges azonban az osztályozás az L2 rétegben szállított minőségi jellemzők alapján is. Ezen speciális biteket az L2 és L3 protokoll adatelemek esetén az 1. ábra mutatja be.

Az L2 Inter-Switch Link (ISL) kereteknek létezik egy 1 bájtos felhasználó (User) mezője, ami az utolsó három biten magában hordozza az IEEE 802.1p CoS (Class of Service) értéket. Az L2 ISL trónk interfészek ISL kereteket továbbítanak. Az IEEE 802.1Q keret egy 2 bájtos TCI (Tag Control Information) mező segítségével szállítja a CoS értéket az utolsó három, User Priority biten. Az ilyen L2-es trónkon minden forgalom 802.1Q kereteket tartalmaz, kivéve a Native VLAN esetében. Más fajta kerettípusok nem tudják szállítani a második rétegbeli CoS értékeket. A CoS értékei 0-7 tartományból vehetnek növekvő prioritású értéket.

Az L3 IP csomagok vagy az IP precedencia, vagy a DSCP (Differentiated Services Code Point) értéket továbbítják. A QoS támogatja mindkét fajta értékelést, mert a DSCP értékek kompatibilisek az IP precedencia értékekkel. Az IP precedencia értékek 0-7, míg a DSCP értékek 0-63 tartományban léteznek.

Minden switch és router amely az Interneten forgalmaz, a csomagokat osztály információval látja el, amely segítségével az azonos osztályhoz tartozó csomagok azonos kezelésben, a különböző osztályhoz tartozó csomagok pedig különböző kezelésben részesülnek. Az osztály információt a csomagokban végfelhasználói csomópontok, vagy switch, illetve router köztes csomópontok is meghatározhatják, függően, a helyi policy-tól,



a részletes csomagvizsgálattól vagy mindkettőtől. A részletes csomagvizsgálat tipikusan a hálózat hozzáférési pontban történik, annak érdekében, hogy a meghatározó switchek és routerek ne legyenek túlterhelve. A switchek és routerek az útvonalon használhatják az osztály információt, hogy meghatározzák a rendelkezésre álló erőforrás készletet a forgalmi osztályok számára. Egy egyedülálló eszköz DiffServ architektúra szerinti forgalom kezelő viselkedését PHB-nak (Per-Hop Behavior) nevezik. Ha a továbbítási útvonal mentén az összes eszköz konzisztens per-hop módszerrel működik, akkor végponttól-végpontig típusú QoS megoldás érvényesül. A QoS bevezetése a hálózaton függ az aktív eszközök QoS sajátosságától, a forgalom típusától és mintájától a hálózaton, valamint a bejövő és kimenő forgalomra alkalmazott vezérlés részletességétől.

3. A QoS struktúra modellje az L2/L3 rétegeknél

Az osztályozás a forgalmak típus szerinti szétválasztását biztosítja. A bemeneti (ingress) eszköz működése magába foglalja a forgalomosztályozás (classification), a szabályozás (policing), a jelölés (mark), a sorbarendezés (queuing) és az ütemezés (scheduling) feladatokat. A QoS alapmodellt a 2. ábra mutatja be. Bemeneti interfészek az osztályozás szétválogatja a különböző típusú forgalmakat [7].

A folyamat készít egy belső DSCP-t csomagonként, ami meghatározza a továbbítás közben végrehajtandó QoS tevékenységeket. A szabályozás meghatározza, hogy a csomag szerepel-e a bekonfigurált profilban összehasonlítva a belső DSCP-t a beállított szabályozókkal (policer), amelyek az adatfolyam által felhasznált sávszélességet korlátozzák. A jelölő (marker) kiértékeli a szabályozót és az interfész szintű konfigurációs információt, majd megvizsgálja azt az előírást, ami szerint kell eljárnia. Ha a csomag a profilon kívül esik, átengedi a csomagot módosított DSCP értékkel vagy eldobja. A sorbarendezés (queuing) megvizsgálja a DSCP vagy a CoS értéket, és ez alapján eldönti, hogy a csomag melyik bemeneti várakozási sorba kerüljön a kettő közül.

Kimeneti interfészek az osztályozás (queuing) kiértékeli a belső DSCP-t és meghatározza, hogy a 4 kimeneti sor közül melyikbe tegye a csomagot. Erre azért van szükség, mert torlódás alakulhat ki eszközön belül, ha a két bemeneti várakozási sor egyszerre küldi az adatot a kimeneti interfész felé. Gyakran torlódás megelőző technológiákat alkalmaznak (WRED – Weighted

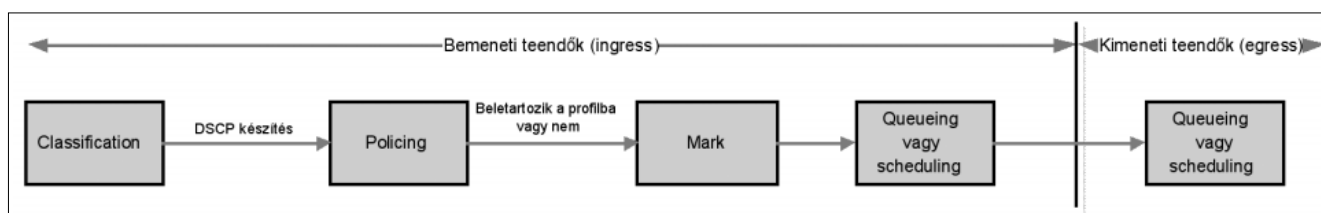
Random Early Detection, és tail drop) a gigabit képes ethernet portok, illetve egy küszöbértékes „tail drop” mechanizmust a 10/100 Mbps-os ethernet portok. Ütemezés (scheduling) a négy kimeneti sor közül egy maximális előnyben (expedite) részesül, így ebbe a sorba kerülő csomagok mindegyike továbbításra kerül mielőtt bármely másik sor tartalma kiszolgálásra kerülne.

IP-től különböző forgalom esetén, ha a beérkező csomag nem rendelkezik CoS értékkel, akkor a bemeneti interfészen érvényes helyi fix beállítás érvényesül. Ha a beérkező keret rendelkezik CoS értékkel, akkor a kimeneti interfész alkalmazza a CoS-DSCP térképet, ami alapján a kerethez rendeli a belső DSCP értéket. Ha a beállítások MAC szűrő listát (ACL – Access Control List) tartalmaznak, a forrás-, a cél cím, illetve a keret típusa alapján történik a DSCP értékének beállítása. Ha nincs ACL, akkor a csomag DSCP=0 értéket kap, azaz „best-effort” alapján továbbítódik.

IP forgalom esetén eszközön belül a beérkező csomagban lévő DSCP használható. Az IETF a ToS mező hat legfontosabb bitjét a DSCP-ként értelmezi. A prioritást a 0-63 intervallumban lévő DSCP érték fejezi ki. Különböző QoS zónák közötti fizikai kapcsolatot biztosító interfészek a DSCP-DSCP mutációs összerendelés alapján megváltoztathatják a két zóna között továbbított csomag DSCP értékét. Lehetőség van a beérkező csomag IP precedencia mezőjének kiértékelésére is, ami alapján a DSCP érték hozzárendelése az IP precedencia-DSCP táblázat alapján történik. Az IPv4 a ToS mező három legnagyobb helyiértékű bitjét használja a precedencia tárolására. Ha a csomagban jelen van a CoS (Class of Service) érték, akkor a DSCP érték a CoS-DSCP táblázatból áll elő. Konfigurált szabványos vagy kiterjesztett IP ACL esetén az IP csomag fejrészében lévő különböző mezők értékei azonosíthatók be. Szűrési találat esetén a szűrőhöz előírt DSCP érték hozzárendelődik az adott csomaghoz. Ha nem létezik ACL, akkor a csomag DSCP=0 értékkel halad tovább.

Az osztályozás-összerendelés (class map) mechanizmus arra használható, hogy egy speciális adatfolyam beazonosítható és megkülönböztethető legyen más adatfolyamoktól. Ez a mechanizmus az adatfolyam további kategóriákba sorolását teszi lehetővé, amihez a döntést az ACL szerinti illeszkedés, DSCP listához vagy IP precedencia listához való tartozás biztosítja. További adatfolyam osztályozásához egy-egy további eltérő nevű osztály-összerendelés lehet készíteni. Ha a csomag egyezik az osztály-összerendelés szabállyal, akkor a policy-összerendelés segítségével megtörténik a kategóriába sorolása. A policy-összerendelés

2. ábra A QoS modell elemei



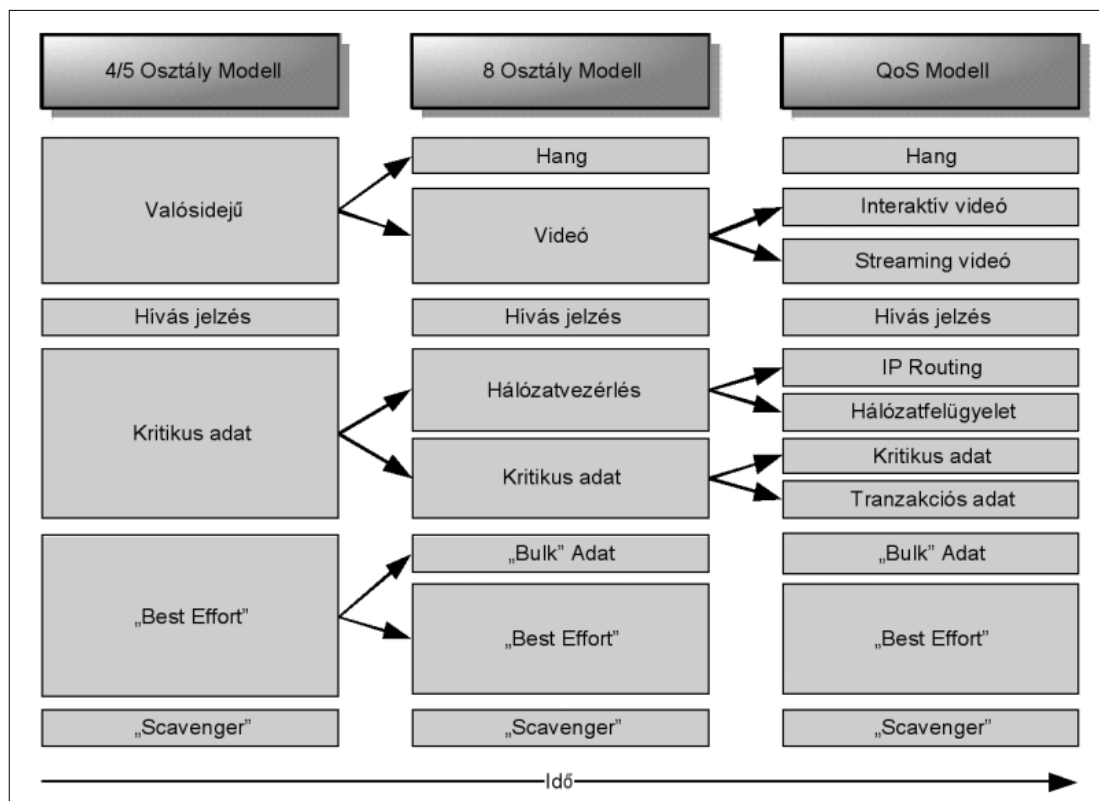
interfészen történő aktivizálása következményeként az alábbi tevékenységek történhetnek: a CoS, DSCP, IP precedencia értékek kiértékelése; a DSCP vagy az IP precedencia érték beállítása; az adatfolyam sávszélességének korlátozása; olyan tevékenység elvégzése, amely a forgalom profile illesztése esetén szükséges. A szolgálat osztályok felsorolását, illetve időben történő kialakulásukat a 3. ábra mutatja [8]. A „Bulk” adat a háttérben futó nagyméretű adatletöltések jellemzője. A „Scavenger” adattípus az IPv6 esetén kerül előtérbe, amely még a „best-effort” adatnál is könnyebben eldobható.

A csomag eszközön belüli DSCP értékének meghatározása után a policing és jelölés események következnek. A policing a forgalom sávszélességének szabályozását lehetővé teszi policerek segítségével. A policerek minden egyes csomagot megvizsgálhatnak és eldöntik, hogy megfelel-e a profilnak vagy sem. A profilnak megfelelő szabályozásokat a jelölő végzi, amely dönt a csomag kézbesítése vagy eldobása felől. A policer lehet egyedi vagy aggregált. Az egyedi QoS policer a sávszélesség korlátokat a megfelelő forgalom osztályok alapján alkalmazza. Az aggregált QoS policer a beállításokat globálisan kezeli, minden forgalmat megvizsgál. A policer zsetonos vödör (token bucket) algoritmust alkalmaz. Ez hasonlít az ATM átviteltechnikánál alkalmazott megoldásra, de itt csak egy edény és csak egy szivárgó lyuk létezik [9]. Minden beérkező keret esetén a vödörbe egy zsetont helyez el. A beállított sávszélességnek megfelelő ritmusban a vödörből a zsetonok kiszivárognak. Amikor a zseton a vödörbe kerül, a kapcsoló eszköz előzetesen ellenőrzi a vödörben lévő üres helyet. Ha nincs elegendő hely a zseton

számára, akkor a csomag nem megfelelő jelölést kap és az annak megfelelő policer intézkedés következik be. Ez lehet a csomag eldobása vagy a DSCP értékének lecsökkentése is. A vödör telítődésének gyorsaságát a vödör mérete (börst [bájt]) a vödör szivárgásának mértéke (bitráta [bps]) és az átlagos bitráta feletti börst időtartama befolyásolja. A vödör mérete a börst hosszát korlátozza és meghatározza az eszközben a bementi pont és a kimeneti pont között továbbítható keretek darabszámát. Alacsony forgalom esetén az adatfolyam nem befolyásolódik. Ha a börst hosszú és magas bitrátájú, a vödör túlcsoordulása miatt a kerettel szemben policer intézkedés lép érvénybe.

A kapcsoló a sorbaállítás és az ütemezés folyamat során torlódás menedzsment célból kimeneti várakozási sorokat, valamint WRR (Weighted Round Robin) mechanizmust használ. Minden Gigabit Ethernet port 4 darab várakozási sorral rendelkezik, amelyek közül egyik kiemelt prioritásúként működhet. A várakozási soroknak két-két küszöbértéke van. A DSCP-küszöbérték táblázat alapján történik a csomag „tail-drop” vagy WRED algoritmus szerinti kezelése. A várakozási sor mérete, a küszöbérték, a „tail-drop” vagy WRED algoritmus és a DSCP-küszöbérték táblázat együtt befolyásolja, hogy a küszöbérték meghaladásakor mikor és melyik csomag eldobása következik be. A kimenő interfész fizikai sávszélessége együttesen képezi a négy várakozási sor számára rendelkezésre álló sávszélességet.

A „tail-drop” a Gigabit Ethernet interfészek alapértelmezett torlódást megelőző mechanizmusa. A csomagok addig kerülnek a várakozási sorokba, amíg a küszöbértéket el nem érik. Ilyen esetben az első küszöbér-



3. ábra

A QoS szerinti hálózat-szolgálati osztályok

tékhöz rendelt csomagok eldobása mindaddig ismétlődik, amíg a forgalom a küszöbérték alá nem csökken. Az elsőnél nagyobb második küszöbértékhez rendelt csomagok ilyenkor is továbbítódnak mindaddig, amíg a forgalom második küszöbértékét el nem éri. A küszöbértékek százalékosan a sorok lefoglaltságát mutatják. A „tail-drop” és a WRED két olyan mechanizmus, amely közül csak az egyik működhet egyidőben az interfészen.

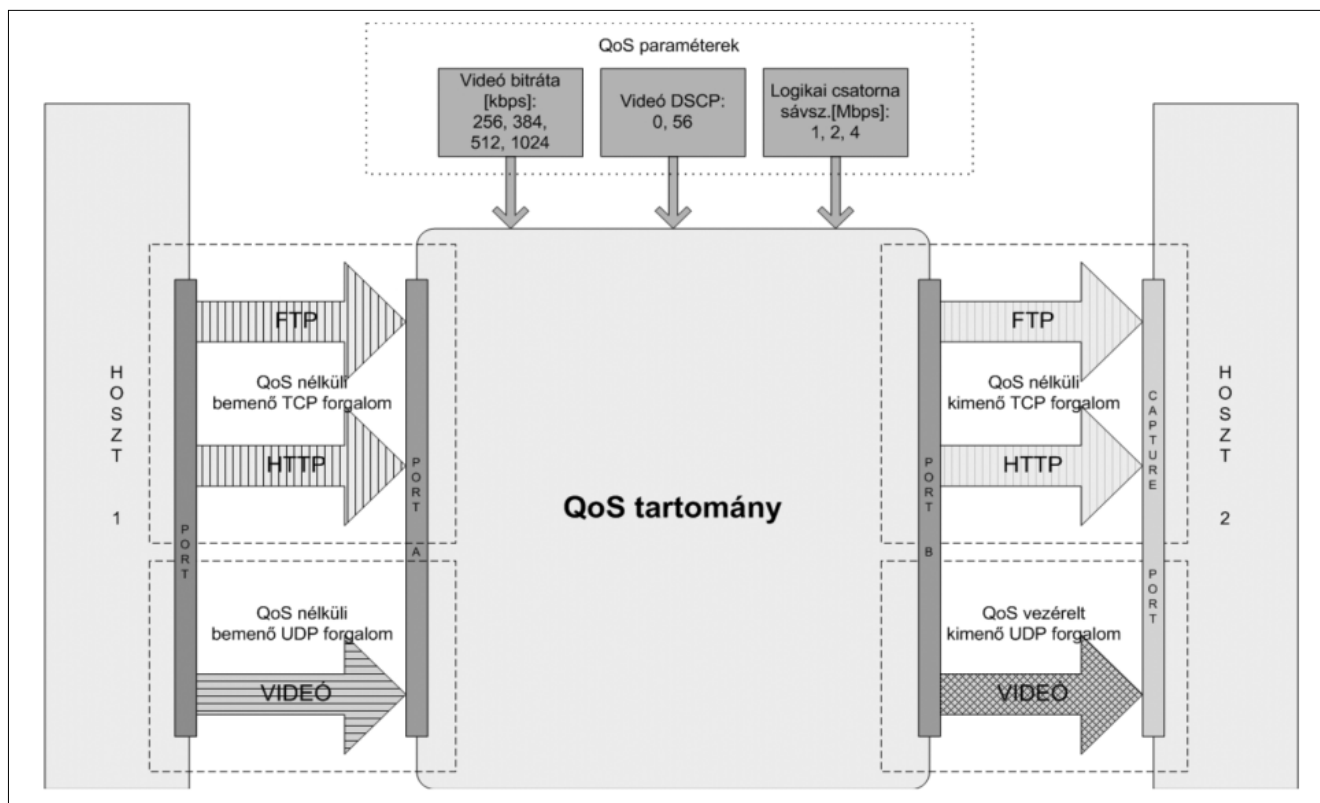
A WRED (Weighted Random Early Detection) abban különbözik más torlódás feloldó mechanizmustól, hogy a torlódás kezelése helyett megpróbálja megelőzni annak kialakulását. A WRED felhasználja a TCP azon torlódásvezérlési tulajdonságát, hogy a TCP a várakozási sora méretének szabályozásával képes ideiglenesen leállítani az üzenetek küldését. A WRED véletlenszerűen eldob csomagokat azelőtt, hogy erős torlódás lépne fel, így a forrás TCP protokoll entitás csökkenti a küldési sebességét, és az L3 rétegben megelőzhető a torlódás. A véletlenszerű csomageldobás lehetővé teszi, hogy a „tail-drop” algoritmussal ellentétben ne kelljen sok csomagot eldobni, ugyanakkor a fizikai csatorna jobb kihasználására nyílik lehetőség. A WRED a nagyobb rátájú forgalmakból többet dob el, mint az alacsony rátájúakból. A kimenő interfész mind a négy várakozási sora itt is rendelkezik egy-egy küszöbértékkel. Ennek meghaladása esetén kezdődik véletlenszerűen a forgalom csomagjainak eldobása. Minél jobban meghaladja a küszöbértéket a forgalom, annál több csomagot dob el. A csomagok kezelése a DSCP-küszöbérték táblázat alapján történik.

A csomag QoS miatti módosítása különböző esetekben következik be:

- i) IP csomagnál az osztályozás alapján DSCP érték rendelődik a csomaghoz. Előfordulhat, hogy ilyenkor a csomag nem módosul, de a DSCP hozzárendelés megtörténik. Ennek az oka, hogy mivel a QoS osztályozás és az ACL szűrőlista illesztése egyidőben történik, az ACL miatt szükség lehet a csomag kiválasztására. Ilyenkor a csomag az eredeti DSCP értékével a kapcsoló CPU-jához kerül, ahol a routing miatt újból ACL illesztés következhet. Az útvonal elemzése az osztályozott DSCP-re épül.
- ii) IP-től különböző csomag esetén nem létezik DSCP, így az osztályozás a csomaghoz egy belső DSCP-t rendel. A belső DSCP alapján a csomagot CoS osztályba sorolja és annak megfelelő módon processzálja.
- iii) „Policing” fázisban az IP és a nem IP csomagok DSCP értéke módosulhat, ha az előírt profil nem illeszthető. Ilyenkor a módosítást a lejelölés (markdown) funkció végzi el.

Az L4-L7 rétegek esetén is van lehetőség a QoS szabályozására [10]. Ebben az esetben megfelelő mechanizmusok segítségével figyelni lehet statikusan és dinamikusan a TCP és az UDP portok használatának statisztikáját; az UDP, illetve a TCP-től eltérő protokollok alkalmazásának arányát; lehetséges továbbá alport szerinti osztályozás, amely a csomag mélyebb szintű elemzésére épít.

4. ábra A mérési környezet és az adatfolyamok



A HTTP forgalmak osztályozását URL, hoszt, illetve MIME alapján végzik. A valós idejű multimédiás hálózati alkalmazások által használt RTP (Real Time Protocol) kontroll modulja páratlan szám azonosítójú portot, az adat modulja pedig páros szám azonosítójú portot alkalmaz [11].

Az RTP időzítés szabályozást, adatvesztés érzékelést, adatvédelem és tartalomazonosítást biztosít. A hasznos teher osztályozása során a hang, a videó, a sűrített vagy sűrítetlen videó, a codec beazonosítására nyílik lehetőség. A felhasználói egyéb alkalmazások statikus porthozzárendelés alapján azonosíthatók be.

A peer-to-peer fájlmegosztó protokollok (Gnutella, FastTrack stb.) erőteljes erőforrás igénye miatt egyrészt a statikus port értéke, másrészt a generált forgalom dinamikája alapján történhet a szabályozás.

4. A mérési környezet és a mért értékek ismertetése

Egyetlen QoS tartományhoz tartozó forgalmakat vizsgáltunk meg. A Hoszt1 géptől a Hoszt2 gép felé egy időben TCP, illetve UDP forgalmat generáltunk. A TCP FTP és HTTP letöltéseket, az UDP pedig interaktív videó átvitelt végzett. A mérési környezetet a 4. ábra mutatja. Mivel az FTP és a HTTP az Ethernet 1500 bájtos MTU-jánál nagyobb méretű IP csomagokkal forgalmaz, ezeknél fragmentáció lépett fel. A QoS paramétereket fragmentum csomagoknál nem lehetséges kezelni, ezért a TCP forgalom „best-effort”, azaz DSCP = 0 érték mellett zajlott.

A videokonferencia UDP forgalom minden csomagja elfér egy-egy Ethernet keretben, így ennek prioritását a TCP forgalom prioritása fölé lehetett emelni az UDP DSCP = 56 értékének beállítása segítségével. A Hoszt1 a lehető legnagyobb rátával küldi a videót a QoS tartományba, ez azonban a forrás fizikai kapacitása miatt legfeljebb 1 Mbps lehet. A Hoszt2 csomópontnál külön a TCP, és külön az UDP hálózati forgalom mérése TCPDump program segítségével történt. A QoS tartományon belül a videó bitrátát és a videó prioritását a Port_A, míg az UDP és a TCP forgalom számára közösen rendelkezésre álló csatorna kimeneti sávszélességét a Port_B pontokban szabályoztuk. A Port_B sávszélességét interfész szintű globális QoS paraméterrel befolyásoltuk. A mérésnél alkalmazott paramétereket a 2. táblázat tartalmazza.

2. táblázat A mérésnél alkalmazott paraméterek

Paraméter	Értékek
Videó bitrátá [kbps]	256, 384, 512, 768, 1024
Videó kódolási algoritmus	H.261, H.264
Videó adatfolyam DSCP értéke	0, 56
Logikai csatorna sávszélesség [Mbps]	1, 2, 4
Adatkapcsolat	IEEE 802.3 (Full Duplex)

Jellemző	Definíció
Átlag (Mean) [kbit/sec]	$\text{Mean}(Z_p) = \bar{Z}_p = \frac{\sum_{t=1}^{T_p} Z_p(t)}{T_p}$
Szórás (Standard Deviation) [kbit/sec]	$\text{STD}(Z_p) = \sqrt{\frac{1}{T_p} \sum_{t=1}^{T_p} (Z_p(t) - \bar{Z}_p)^2}$
Relatív szórás (Relative STD) [%]	$\text{RSD}(Z_p) = \frac{\text{STD}(Z_p(t))}{\bar{Z}_p}$
Ferdeség (Skewness) [%]	$\text{Skewness}(Z_p) = \frac{\sqrt{T_p} \sum_{t=1}^{T_p} (Z_p(t) - \bar{Z}_p)^3}{\left(\sum_{t=1}^{T_p} (Z_p(t) - \bar{Z}_p)^2 \right)^{3/2}}$

3. táblázat Az idősorok jellemzői

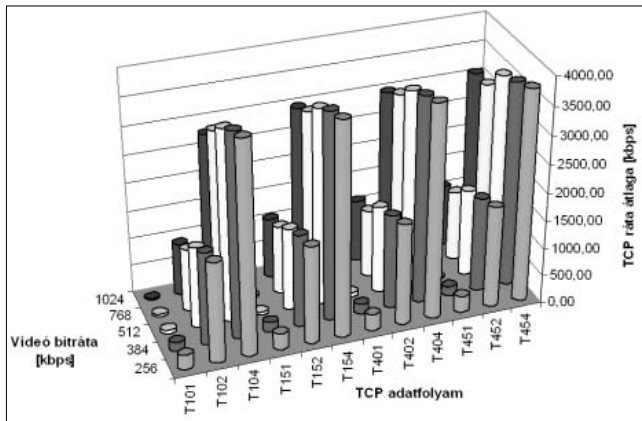
Külön a TCP és külön az UDP forgalmakra az L2 keretek méretét és időpontját rögzítettük. Egy mérés 60 másodpercig tartott, összesen 60 mérés készült. Minden egyes mérésnél ugyanazt a műsort forgalmaztuk a két hoszt között. Az első harminc mérésnél a forrás H.261, a továbbiaknál H.264 kodeket használt. Minden egyes idősorra 100 msec-os mintavételezéssel meghatároztuk a bitrátát. Az így nyert újabb két idősor halmaz $Z_p^T(t)$ ($p=1,2,\dots,60$ TCP esetén), illetve $Z_p^U(t)$ ($p=1,2,\dots,60$ UDP esetén) minden eleme $T_p = 60.000$ értéket tartalmaz. Ezen idősorokat elemeztük matematikai statisztikai szemszögből. A megvizsgált jellemzők a bitráta átlaga, szórása, relatív szórása, valamint a ferdesége. Ezek definícióját a fenti, 3. táblázat mutatja be.

5. A mérési eredmények elemzése és értelmezése

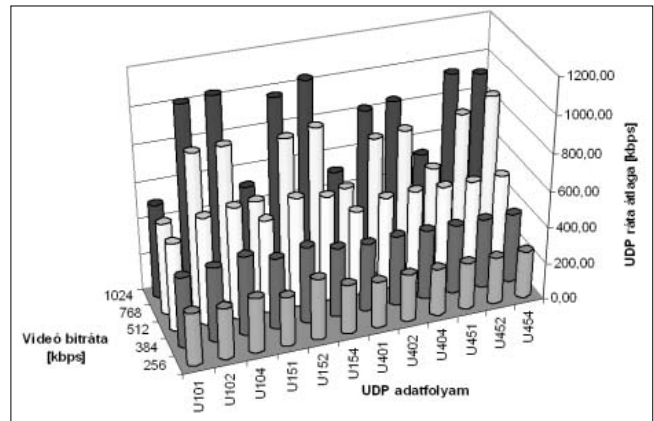
Az 5-12. ábrák az UDP, illetve a TCP forgalmak bitrátájának átlagát, szórását, relatív szórását, illetve ferdeségét mutatják.

A különböző adatfolyam halmazokat $abc=(101\dots454)$ index segítségével jelöltük, ahol $a=(H.261, H.264)$, $b=(DSCP1=0, DSCP=56)$, $c=(\text{logikai csatorna sebesség}=1 \text{ Mbps}, 2 \text{ Mbps}, 4 \text{ Mbps})$. Így például az $abc=454$ indexű idősor halmaz H.264 codec, DSCP=56 érték és a Port_B logikai csatornájának átviteli sebessége=4 Mbps esetben készült.

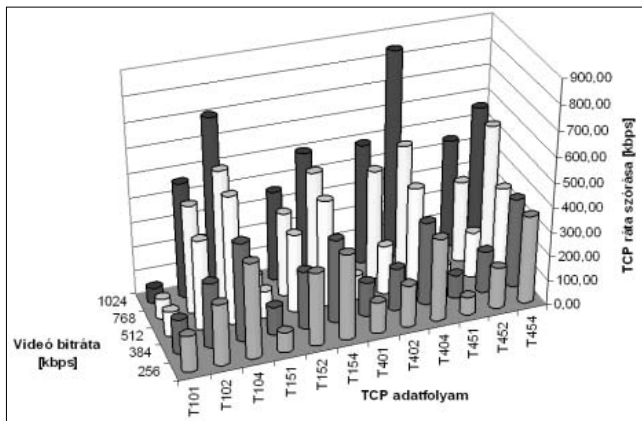
Adott halmaz elemei 256, 384, 512, 768 és 1024 Kbps-os bitráták mellett készültek. A TCP és az UDP „best-effort” jellegű ($b=0$ halmazok) forgalmazása esetén a különböző adatfolyamok átlagosan kitöltik a rendelkezésre álló sávszélességet, és mindegyik alkalmazás működik. QoS segítségével történő videó bitráta növelése miatt az FTP és a HTTP forgalmak átlagát a TCP automatikusan visszaszabályozza. A H.261 codec a H.264-hez képest átlagosan több UDP adatot képes továbbítani annak ellenére, hogy régebbi algoritmus. Ennek oka, hogy a rendelkezésre álló videó bitráta nem haladja meg az 1 Mbps értéket, ami alatt a H.264 nem működik optimálisan. A forgalmak bitrátájának szórása azt mutatja, hogy nagyobb rátánál növekszik a szórás



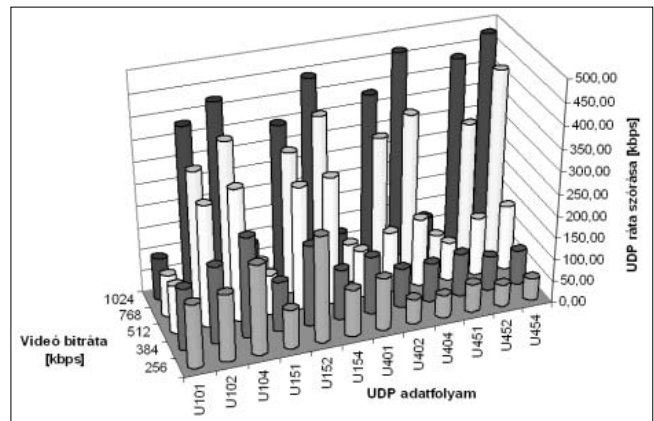
5. ábra A TCP bitráta átlaga



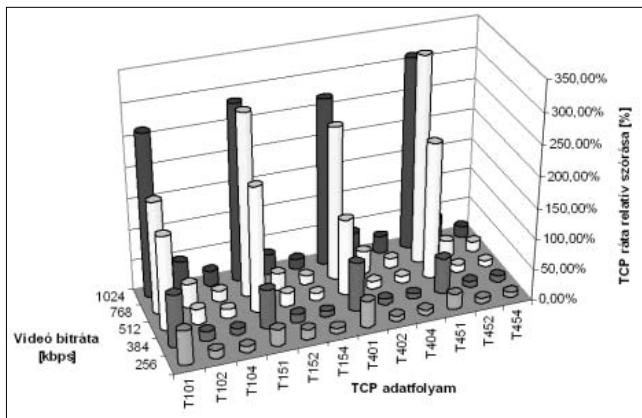
6. ábra Az UDP bitráta átlaga



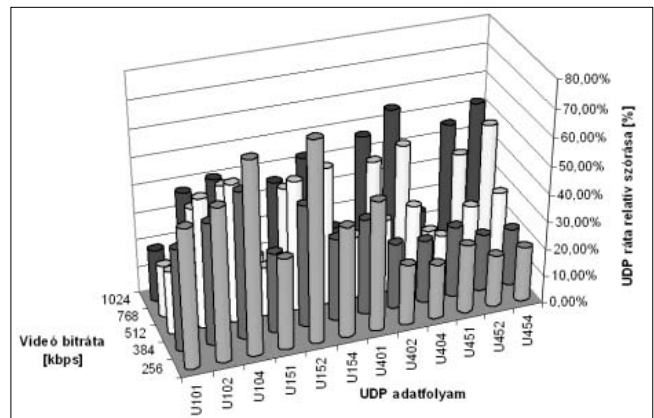
7. ábra A TCP bitráta szórása



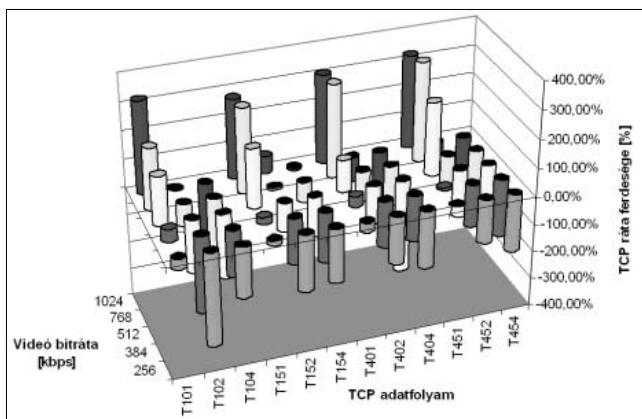
8. ábra Az UDP bitráta szórása



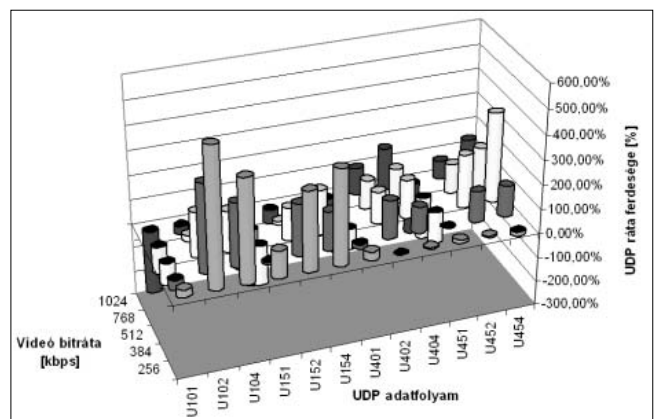
9. ábra A TCP bitráta relatív szórása



10. ábra Az UDP bitráta relatív szórása



11. ábra A TCP bitráta ferdesége



12. ábra Az UDP bitráta ferdesége

Dinamika	Darabosság	Átlapoltság	Színhűség	Értelmezhetőség	OS
Álló	Nagyon	Igen	Nem	Nem	1
Akadozó	Nagyon	Igen	Nem	Nem	2
Akadozó	Közepes	Igen	Nem	Kevésbé	3
Akadozó	Közepes	Igen	Nem	Közepes	4
Akadozó	Közepes	Igen	Nem	Közepes	5
Mozgó	Kevésbé	Nem	Nem	Közepes	6
Mozgó	Nem	Nem	Nem	Közepes	7
Mozgó	Nem	Nem	Nem	Elfogadható	8
Mozgó	Nem	Nem	Nem	Jó	9
Mozgó	Nem	Nem	Igen	Nagyon jó	10

4. táblázat A videó minőségének vélemény-érték (OS) metrikája

is, viszont alacsony videó bitráta esetén a H.264 codec kevesebbet szór, mint a H.261. 4 Mbps csatorna esetén a T104 és T404 erőteljesen szór, ha a videó bitráta 1 Mbps-on van. Ez azt jelenti, hogy a videó legjobb minősége esetén a TCP a maradék 3 Mbps sávszélességen egyre inkább börsztösi az adatátvitelét. A TCP bitráta relatív szórása alapján látható, hogy 1 Mbps-os csatorna sávszélesség esetén a TCP erőteljesebben szór, és eléri a 300%-ot is, míg a többi esetben ez jóval 40% alatt marad. Az UDP relatív szórása kis videó bitráta esetén viszonylag magas, de még így is csak 70% alatti.

A H.264 relatívan is kevesebbet szór alacsony videó bitrátánál, függetlenül a csatorna fizikai sávszélességétől. A bitráta ferdesége azt jelenti, hogy a periodogram a súlyvonalához képest balra (negatív) vagy jobbra (pozitív) ferdül el. Negatív ferdeség azt jelenti, hogy az átlagos bitrátánál sokszor kisebb a forgalom, viszont a forgalomban ritkábban börsztös jelenségek léteznek. A pozitív ferdeség azt mutatja, hogy az átlagos bitrátánál gyakran nagyobb a forgalom, de csak kis mértékben, és léteznek hosszabb időszakok, amikor az átvitel szünetel. A TCP bitrátája csak 768 Kbps-nál nagyobb videó bitráta beállítás esetén pozitív ferdeségű, a többi esetben negatív. Az UDP bitrátája ezzel ellentétben csak a H.261 codec és magas videó bitráta esetén negatív ferdeségű, a többi esetben pozitív, sőt gyakran megközelíti az 500%-ot is.

A TCP „best-effort” és az UDP QoS vezérlése (b=5 idősorok) esetén a bitráták átlagának grafikonja nem teszi lehetővé a homogén „best-effort” jellegű forgalmaktól való különbségtételt annak ellenére, hogy a tapasztalt videó minősége lényegesen eltér egymástól a QoS jelenléte, illetve hiánya esetekben.

A TCP mindig a maradék rendelkezésre álló sávszélességet használja ki, mivel az L4 forgalomszabályozás ezt lehetővé teszi. Az UDP a TCP-hez viszonyítva csak 50%-os bitráta szórást okoz,

azonban az 1 Mbps-os (c=1 síkok) UDP a 2 Mbps és 4 Mbps csatornákkal ellentétben a videó bitráta növelésével csökkenti a szórást. Erre az a magyarázat, hogy nagyon alacsony csatorna sávszélességnél a TCP radikálisan csökkenti a forgalmát, így a videó továbbítása kevésbé börsztösen lehetséges.

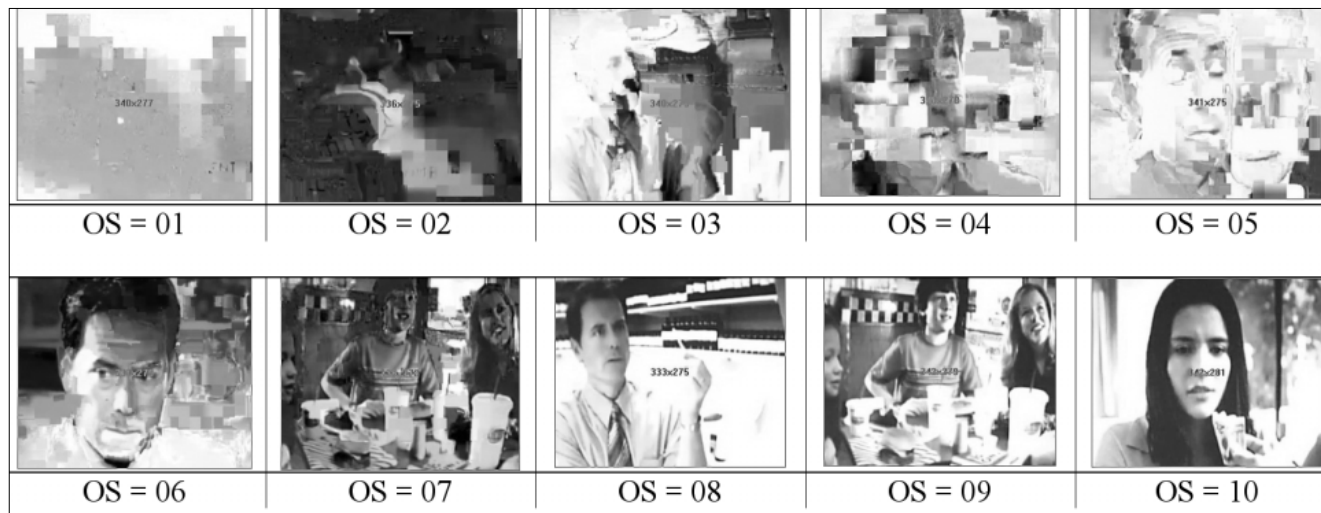
A H.264 relatív szórása független a QoS-tól és növekszik a videó bitrátával, ami a H.264 codec dinamikusabb működését igazolja. A TCP bitráta ferdeségére nincs hatással a videó QoS beállítása, mivel a TCP a maradék sávszélességet használja fel. Az UDP bitráta ferdeségét a QoS beállítások kis mértékben csökkentik, de még mindig a pozitív tartományban tartják.

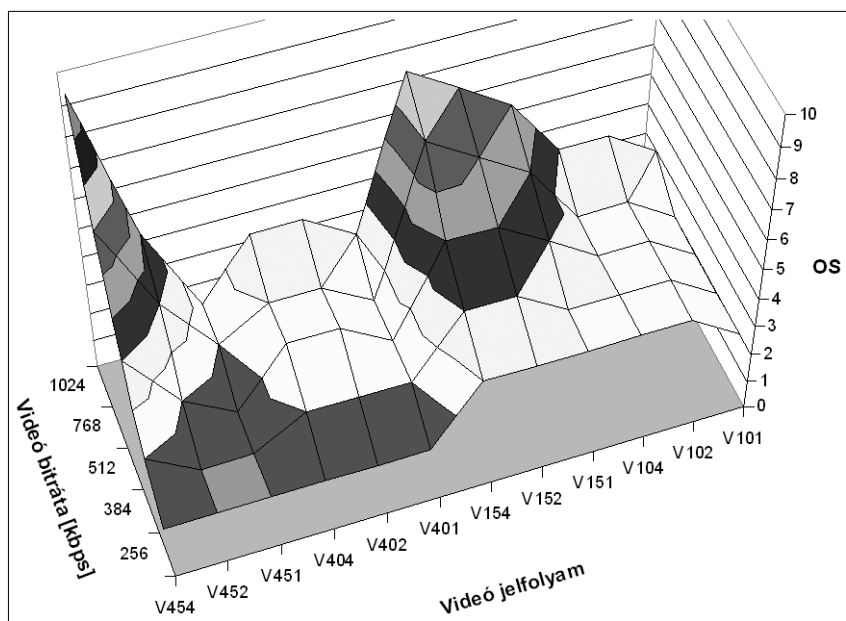
Egy tíz pontos tartományban mérő vélemény-érték (OS – Opinion Score) saját metrikát képeztünk adott videó műsor minőségének globális számszerűsítéséhez. Őt alapvető minőségi szempontot javasolunk, amelyeket a 4. táblázat mutatja be.

Az OS segítségével a QoS mechanizmus és a codec együttes hatását objektív módon mérhetjük.

Adott videó műsor globális OS, vélemény értékekhez tartozó képeket az 5. táblázat tartalmazza. Az így összeállított lista lehetővé teszi a videó műsor minőségének számszerű értékelését és kategóriákba sorolását is. Alacsony OS értékek gyenge minőséget, nagy OS értékek jó minőséget képviselnek.

5. táblázat A vélemény-értékek (OS) szerinti videó képek





13. ábra

A videó forgalmak vélemény-értéke (Opinion Score)

A különböző QoS paraméterekkel szabályozott videó jelfolyamok (V101...V454) egymáshoz viszonyított globális OS, vélemény értékét a 13. ábra mutatja be. Megfigyelhető, hogy a videó átvitel vegyes terhelésű hálózaton nagymértékben függ a QoS beállításoktól.

A homogén „best-effort” módszer szerinti forgalomtovábbítás megosztja az erőforrásokat a különböző adatfolyamok között, míg a QoS mechanizmus alacsony csatorna sávszélességnél a valós idejű alkalmazásokat megszakítja, nagyobb csatorna sávszélesség esetén pedig erőteljesebb különbséget tesz az eltérő típusú adatfolyamok között. A videó kapcsolat kielégítően jó minőségű átviteléhez végponttól-végpontig minimum 1 Mbps-ra van szükség. Alacsony sávszélességen a H.261 codec jobb minőséget ad, mint a H.264, viszont utóbbi képes akár HDTV minőségű műsor továbbítására is 2 Mbps-nál nagyobb sebességű összeköttetés esetén [11]. QoS mechanizmus működtetése mellett jelentős minőségi ugrást az 500 Kbps-nál nagyobb sebességű adatkapcsolat esetén tapasztalhatunk, amit a felhasználók véleményének az OS tízes skáláján a első tartományban való elhelyezése tükröz.

6. Összefoglalás

Jelen cikkben a QoS mechanizmus L2 és L3 rétegekben kifejtett hatását vizsgáltuk meg egyetlen QoS tartományon belül szabályozott paraméterek segítségével, H.261 és H.264 videó codec alkalmazása mellett. A mérések alapján kijelenthető, hogy a QoS mechanizmus aktivizálása jelentősen megváltoztatja az Interneten hagyományosan működő hálózati alkalmazások viselkedését. Az UDP egyenletesebb adatfolyamot biztosít, míg a TCP a maradék hálózati erőforrás teljes kihasználására is képes. QoS mechanizmusok segítség-

gével hatékonyan differenciálni lehet a különböző típusú adatforgalmak között, így a valós idejű hálózati alkalmazások (VoIP, video, játék stb.) kielégítő minőségben képesek együttműködni a hagyományos adatátviteli szolgáltatásokkal. Ez jelentős beruházási megtakarításokat jelent a jövőben, hiszen a meglévő infrastruktúra teljes lecserélése nélkül a QoS mechanizmusokkal lehetőség van a hang-, videó-, adatátvitel integrációjának folytatására.

További vizsgálatok szükségesek az egyéb QoS paraméterek, egyetlen, illetve több QoS tartományon átívelő multimédiás kapcsolatok viselkedése, valamint az L4-L7 rétegek működésének minőségi befolyásolhatósága témakörökben.

Irodalom

- [1] Luis F. Ortiz:
„Solving QoS in VoIP: A formula for explosive growth?”
Brooktrout Technology
- [2] „AutoQoS for Voice Over IP (VoIP)”,
White Paper, Cisco Systems Co.,
<http://www.cisco.com>
- [3] „Advanced QoS”,
White Paper, Allied Telesis,
<http://www.alliedtelesis.com>
- [4] „Quality of Service”,
Technical White Paper, Microsoft Co.,
<http://www.microsoft.com/technet/prodtechnol/windows2000serv/plan/qosover.msp>
- [5] „QoS” – White Paper, Allied Telesis,
<http://www.alliedtelesis.com>
- [6] „Configuring QoS” – Catalyst 3550 Multilayer Switch
Software Configuration Guide, Cisco Systems Co.,
<http://www.cisco.com>
- [7] „Configuring QoS” – Catalyst 3750 Multilayer Switch
Software Configuration Guide, Cisco Systems Co.,
<http://www.cisco.com>
- [8] „Enterprise QoS Solution Reference” –
Network Design Guide, Cisco Systems Co.,
<http://www.cisco.com>
- [9] Zoltán Gál, Csaba Szabó:
„Migration to ATM in an Academic MAN Environment
– Network Design Considerations and a Case Study”,
8th IEEE LAN/MAN Workshop Proceedings,
Berlin, 25-28 August, 1996.
- [10] „Network-Based Application Recognition and
Distributed Network-Based Application Recognition”
– Network Design Guide, Cisco Systems Co.,
<http://www.cisco.com>
- [11] Gál Zoltán, Karsai Andrea, Orosz Péter:
„A WiFi rendszerek multimédiás alkalmazásokra
gyakorolt hatása”,
Híradástechnika, 2006/6, pp.15–23.

Szolgáltatásminőség és a szolgáltatás minősége – szempontok az Internet-elérési szolgáltatások értékeléséhez

VONDERVISZT LAJOS

Nemzeti Hírközlési Hatóság
vonderviszt.lajos@nhh.hu

Kulcsszavak: QoSE, QoS, Internet, sebességmérés

Az érzékelt szolgáltatási minőség (QoSE) és a szolgáltatásminőségi paraméterek (QoS) között az Internet-felhasználók szempontjából közötti távolság. A cikk megkísérli vázolni a problémát, példákon keresztül megmutatni a fogyasztók érdekében tett hazai és nemzetközi lépéseket, valamint a fejlődési irányokat.

1. Bevezetés

Az Internet-szolgáltatást igénybe vevő átlagos felhasználó jellemzően nem tudja értelmezni a hálózati szolgáltatás technológiai paramétereit, csak a végeredményképpen kapott szolgáltatási minőséget érzékeli szubjektíven. De bevallhatjuk őszintén, a mindennapi gyakorlatban nem ezzel a témával foglalkozó szakember is bajban lenne, ha megkérdeznénk, a szolgáltatótól kapott 2 Mbit/s-es ADSL szolgáltatás sávszélességét hogyan kell értelmezni, ez vajon adatkapcsolati rétegre vonatkoztatott sávszélesség-e, tartalmazza-e a második rétegbeli protokoll járulékos adatok (pl. fejlécszek) miatti többletét (overhead), mennyi hasznos adatot (mekkora fájlt) lehet átvinni ezzel időegység alatt a legjobb esetben, nyugtázásos protokollt (TCP) vagy megbízhatatlan protokollt (UDP) használva, a garantált legkisebb sávszélesség mely pontok között érvényes, vagy milyen módon lehet verifikálni a szolgáltató állításait a megadott szolgáltatási paraméterekkel kapcsolatban.

A felhasználók a jellegzetes szolgáltatásminőségi (QoS) paraméterek (pl. sávszélesség, késleltetés, késleltetés ingadozás, bithiba-arány) helyett komplex, nem algoritmizálható módon érzékelik a szolgáltatás minőségét (service quality), ezen belül a számukra érzékelhető paraméterek (pl. letöltési sebesség) „fuzzy” minősítésével (lassú, gyors) jellemzik a szolgáltatás aktuális állapotát.

Nem segítenek az igazán objektív kép kialakításában az elérhető – több szolgáltató által elérhetővé tett – letöltési sebességet mérő alkalmazások, hiszen a méréshez használt környezet ezekben az esetekben nem jól definiált.

A példánál maradva, a szolgáltató a felhasználóval kötött szerződésében egy adott hálózati sávszélesség (hálózati minőség-paraméter) biztosítását ajánlja fel általában a felhasználóoldali hozzáfé-

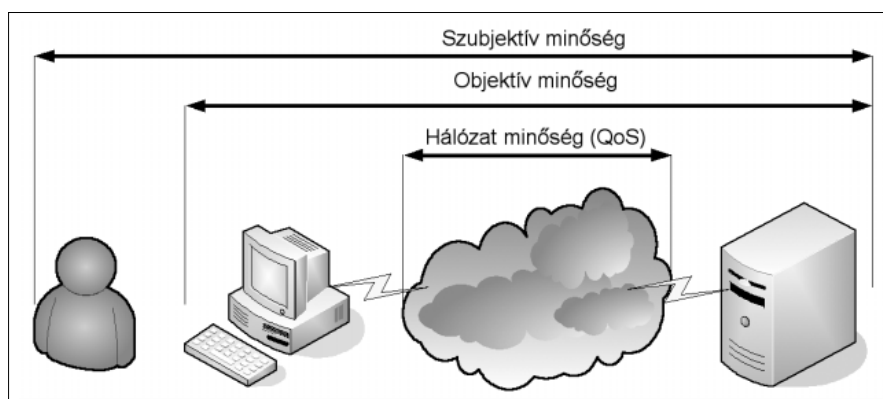
rési pont és a szolgáltatóoldali kiszolgáló eszköz között, illetve a ennél szokásosan kisebb sávszélesség biztosítását garantálja a szolgáltató belföldi kapcsolataira (Magyarországon a BIX – Budapest Internet eXchange). A felhasználó a saját rendszere, a futó szerveroldali alkalmazás, valamint az ezeket összekötő hálózat, mint komplex rendszer sebességét érzékeli, az olyan paraméterek pontos ismerete nélkül, mint például – a teljesség igénye nélkül – a lokális számítógépre telepített vírusirtó, tűzfal, a többi futó program, az operációs rendszer, a számítógép-hardver, a hálózati kártya, az otthoni router, a szerveroldali hálózati kártya, operációs rendszer, a szerver CPU terhelése, a szerveren futó védelmi programok, a mérő alkalmazás lassító hatása.

A rendszer komplexitásából adódóan a szerződésben rögzített paramétert a felhasználó közvetlenül képtelen meghatározni, egy ilyen típusú mérés során meghatározhatja a teljes rendszer objektív minőségi paramétereit (1. ábra), de a hálózati minőségre csak az állítható bizonyossággal, hogy ez jobb, mint a mért érték.

Kérdés viszont az, hogy valóban olyan paraméterekkel kell-e jellemezni a hálózatokat, amelyek a felhasználók számára nem mérhetők.

Ha elfogadjuk, hogy a felhasználó által érzékelt szolgáltatás minősége nem származtató közvetlenül a szolgáltatásminőségi paraméterekből – bár ezekkel kapcsolatban van – megnyílik az út a továbblépéshez.

1. ábra Szubjektív és objektív minőség



2. Az érzékelt szolgáltatásminőség

A felhasználó által érzékelt szolgáltatásminőség (QoSE – Quality of Service Experienced) [4] szubjektív módon származik a technológia által biztosított, adott esetben számszerűen mérhető szolgáltatásminőségi paramétereiből és olyan egyébként sem objektív tényezőkből, mint például az ár, vagy a márkanév (2. ábra).

A felhasználó joggal várja el, hogy a szolgáltatás várható és teljesült minőségi paramétereiről tájékoztatást kapjon, méghozzá olyan adatokkal, amelyeket ő is képes értelmezni illetve érzékelni, amelyek összehasonlíthatóvá teszik a különböző szolgáltatók szolgáltatásait (ilyenformán technológiafüggetlenek), amelyek értelmezhetők a szolgáltatók és a felhasználók számára egyaránt (ilyenformán tárgyalási alapot képeznek) és minden lényeges szempontot figyelembe vesznek.

A szolgáltatók és a felhasználók között a szolgáltatásminőségi paramétereket a szerződés részét képező, azon belül gyakran nem nevesített szolgáltatási-szint-egyezmény (SLA – Service Level Agreement) rögzíti [2], amely mérhető jellemzőkkel definiálja a teljesítmény, rendelkezésre állás, konfiguráció, beszámolás és ügyféltámogatás elvárható és minimális szintjét. Az ábra alapján könnyen levonható a következtetés, hogy az SLA körébe tartozó jellemzők közül a teljesítmény és a rendelkezésre állás azok, amelyek a legnehezebben foghatók meg együtt a szolgáltatók és a felhasználók számára.

3. A magyarországi gyakorlat

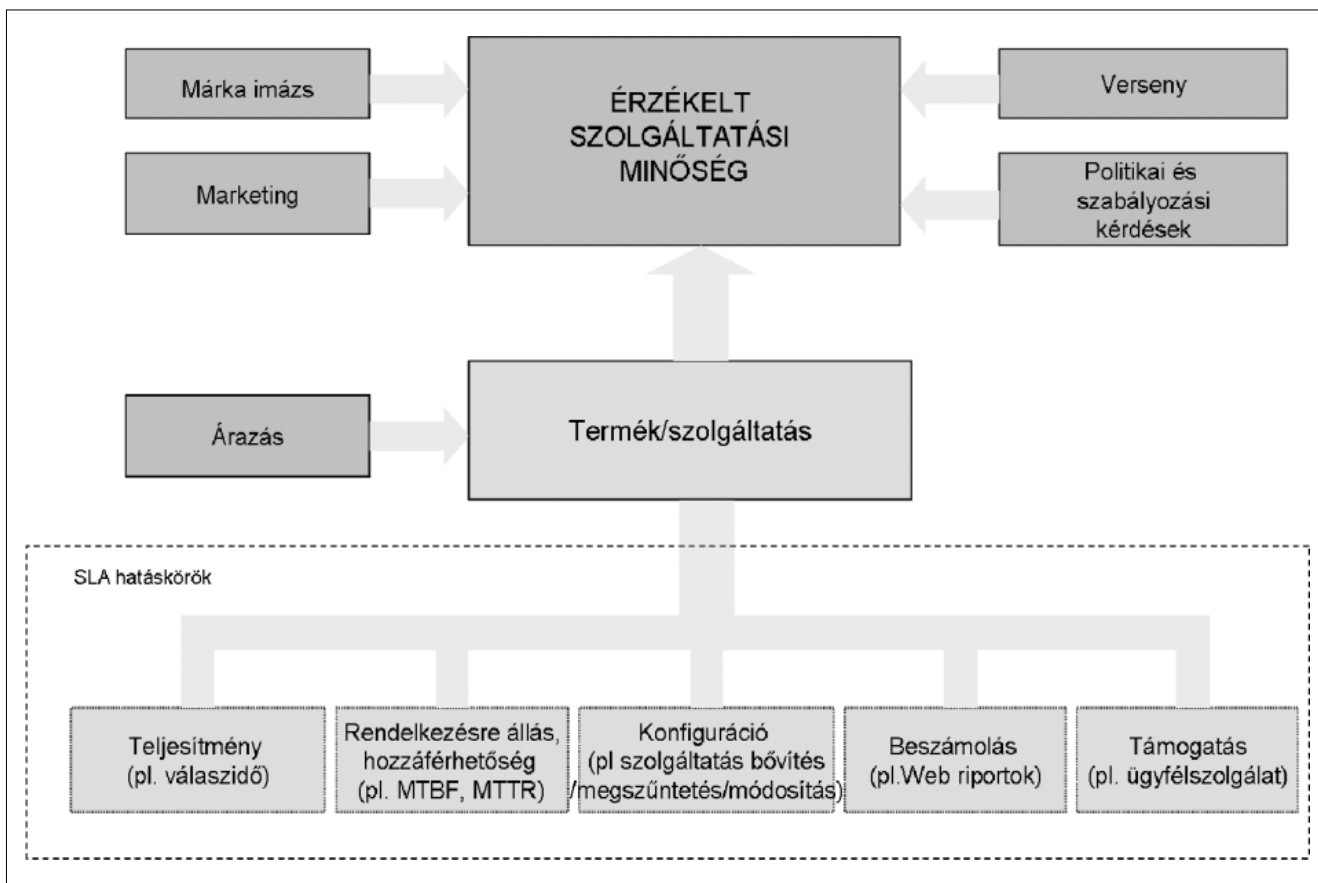
A „345/2004. (XII. 22.) Korm. rendelet az elektronikus hírközlési szolgáltatás minőségének a fogyasztók védelmével összefüggő követelményeiről” rendelkezik arról, hogy a szolgáltatók kötelesek minden, az adott helyzetben általában elvárható intézkedést megtenni annak érdekében, hogy szolgáltatásuk megfeleljen a szolgáltatásminőségi követelményeknek, amelyet ugyanezen rendelet a következőképpen definiál:

„Szolgáltatásminőség: a szolgáltatási képességek azon együttes hatása, amely a szolgáltatás fogyasztójának elégedettségi fokát meghatározza. A gyakorlatban a szolgáltatási képességekre vonatkozó, objektívan mérhető minőségi mutatókkal jellemezhető.”

A szolgáltató az általa nyújtott szolgáltatások kormányrendelet szerinti megfelelőségének hitelt érdemlő igazolását szolgáltatásonként a Nemzeti Hírközlési Hatóság által rendszeresíteni kívánt elektronikus adatlap és a megfelelőség igazolás (nyilatkozat és/vagy tanúsítványok, felülvizsgálati jelentés) megküldésével (benyújtásával) teljesítheti a tárgyévét követő év január 31-ig.

A megadandó minőségi paraméterek az egyes szolgáltatásokra (Telefon-, Internet és egyéb adatátviteli, Műsorelosztó- és Egyéb szolgáltatás) vonatkozóan kategorizáltak (3. ábra), és a szolgáltatók által megadott adatok illetve a belőlük származó statisztikák megtekinthetők a Hatóság honlapján (<https://www.nhh.hu/adatlap345/stat/statmenu.jsp>).

2. ábra Az érzékelt szolgáltatásminőség



A jogszabály által meghatározott paraméterek túlnyomó része konkrét minőségi jellemzőkre kérdez rá, ezen belül a technológiával szoros kapcsolatban álló teljesítmény és rendelkezésre állási paraméterekre.

Minőségi paraméter / szolgáltatás típus	T	I	M	E
1. Új hozzáférés létesítési idő	√	√	√	√
2. Minőségi panasz hibaelhárítási ideje	√	√	√	√
3. Számlapanasz kivizsgálási és elintézési ideje	°	°	°	°
4. A szolgáltatás rendelkezésre állása	√	√	√	√
5. A sikertelen hívások aránya	√	-	-	-
6. A hívás felépítési ideje	√	-	-	-
7. Kezelő, hibafelvevő válaszüzeje	√	-	-	-
8. Bithiba arány hozzáférési vonalanként másodlagos felhasználás esetén	√	-	-	-
9. A szolgáltatási terület egészét érintő szünetelés	-	√	-	-
10. Az előfizetők legalább 10%-át érintő szünetelés	-	√	-	-
11. A garantált (átlagosan elvárható) le- és feltöltési sebesség	-	√	-	-
12. Bithiba arány hozzáférési vonalanként	-	√	-	-
13. Vívószint az előfizetői hozzáférési ponton	-	-	√	-
14. Jelszintkülönbségek az előfizetői hozzáférési ponton	-	-	√	-
15. Vívó/zaj viszony az előfizetői hozzáférési ponton	-	-	√	-
16. Előfizetői panaszok száma	√	√	√	√
17. Minőségi panaszok száma	√	√	√	√
18. A jogos minőségi panaszok száma	√	√	√	√
19. A számlapanaszok száma	°	°	°	°
20. A jogos számlapanaszok száma	°	°	°	°
21. Az ügyintézés elleni panaszok száma	√	√	√	√

3. ábra A szolgáltatások minőségi paraméterei

A Hatóság által közzétett kézikönyv [1] szerint a szolgáltatónak a szolgáltatással kapcsolatosan célértéket és minimál értéket kell az Általános Szerződési Feltételek (ÁSZF) dokumentumában szerepeltetnie és a hatóságnak megadnia. A szolgáltatás célértékének teljesülése vagy túlteljesülése esetén az adott szolgáltatás megfelelőnek minősül. A minimálértéket nem elérő szolgáltatás viszont a szolgáltatás „igénybe nem vehető” minősítést kap. A két érték között a szolgáltatás csökkentett értékűnek kell tekinteni.

A szolgáltatóknak a ténylegesen teljesített értékeket méréssel kell igazolniuk, és az adatszolgáltatás során a teljesítési érték oszlopban szerepeltetniük. A kézikönyv szerint:

„Internet szolgáltatás esetén a tesztméréseket a BIX hálóra kapcsolt referencia szerver(ek)en kell elvégezni.

A referencia szerverre mesterséges tesztforgalmat kell generálni kiválasztott helyeken telepített hozzáférési pontokról. A tesztforgalom egyszerű TCP (FTP) forgalomból állhat. Nem szükséges a használt hálózati protokollok (http, mail stb.) pontos lemásolása.

A tesztvizsgálat időpontja: a mérést forgalmas időszakban kell végezni.

A tesztvizsgálat időtartama: 3 perc.

A referenciaszerver(ek)et független szervezetnek kell működtetnie. Amíg referenciaszerver nem áll rendelkezésre, a szolgáltatónak kell gondoskodnia vizsgálószerverről. A referenciaszerver rendelkezésre állását követő harminc napig a szolgáltató által biztosított vizsgálószerverrel

elvégzett vizsgálatok eredményei is elfogadhatók.” Illetve „Az előfizetői csatlakozások, szolgáltatási terület és szolgáltatáscsomagok tekintetében a mintavételi eljárásnak reprezentatívnak kell lennie, a vizsgálathoz felhasznált teszt mérések minimális darabszáma: az átlagos előfizető szám 1%-a, de legfeljebb 50 db.”

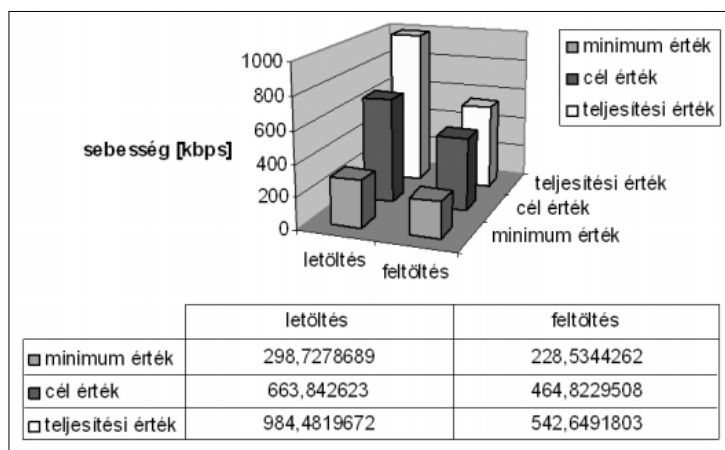
Mint látható, a mérés olyan módon kívánja a szolgáltatók által megajánlott sávszélességet megállapítani, amely figyelembe veszi a felhasználók szokásait, azaz megfelel a QoSE modellnek. A cikk írásának pillanatában a referenciaszerver még nem áll rendelkezésre, így a lekérdezhető adatok a szolgáltatók méréseiből származnak.

A 2006-os adatokat vizsgálva 320 szolgáltató közlését figyelembe véve, elmondható, hogy a szolgáltatók a vállalt letöltési sávszélességet több, mint 50%-kal túlteljesítik (4. ábra). Tovább árnyalja viszont

a képet, hogy az adatsorok elemzése szerint 20 szolgáltató mintegy hétszeresét éri el a célértéknek, ami adott esetben nem a szolgáltatás minőségét jelzi, hanem azt, hogy a szolgáltató nem vállal felelősséget a felhasználó előtt „megcsilllogtatott” nagy sávszélességért. Meg kell jegyezni, hogy a hirdetésekben, weboldalakon a szolgáltató általában az adott előfizetői hozzáféréseken elérhető maximális átviteli sávszélességet adja meg, amelyet a felhasználó a protokollok sajátosságai miatt gyakorlatilag soha nem tapasztal.

A szolgáltatási minőség árnyaltabb érzékelése érdekében, a verseny növeléséhez és az árazási pozíciók áttekintéséhez a Hírközlési Fogasztói Jogok Képviselője a <http://tantusz.nhh.hu> oldalon kereshetővé és összehasonlíthatóvá tette a szélessávú szolgáltatók adatait.

4. ábra Letöltési sebességek



4. Egy gyakorlati példa az érzékelt minőség mérésére

A portugál Nemzeti Hírközlési Hatóság (Autoridade Nacional de Comunicações) a fentiek figyelembevételével végzett értékelést Internet-hozzáférési szolgáltatásra az öt legnagyobb internet szolgáltató esetében 2005 során [3]. A mérések a felhasználókat területileg jól reprezentáló célcsoport kiválasztásával, összesen száz háztartás megkeresésével, ellenőrzött körülmények között történtek. A teszt letöltési sebesség- (rendelkezésre álló sávszélesség-) paraméterek mérését tűzte ki célul a következő módszerrel:

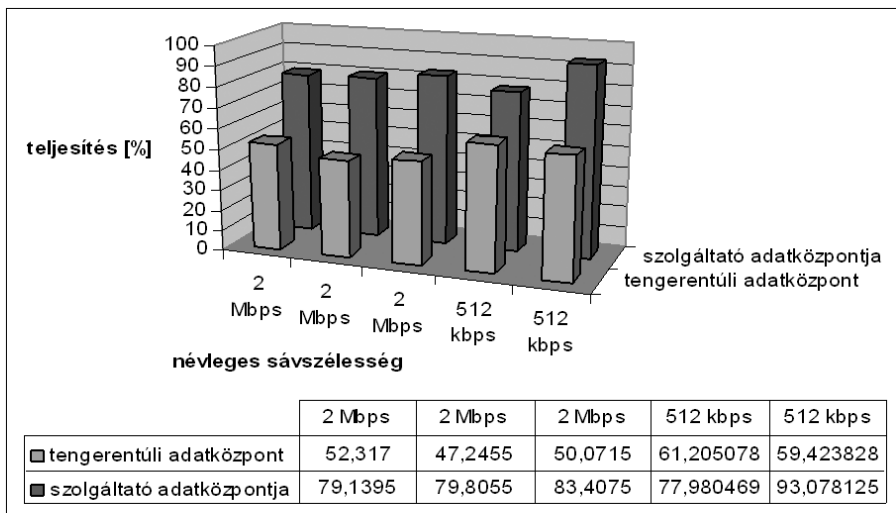
- 1 Mbyte-os weboldal letöltése a szolgáltató adatközpontjából, másik helyi szolgáltató adatközpontjába illetve tengerentúli szerverről.
- 0.5 Mbyte-os fájl feltöltése a szolgáltató adatközpontjában levő szerverre, másik helyi szolgáltató adatközpontjába és tengerentúli adatközpontba.
- 0.5 Mbyte-os fájl letöltése a szolgáltató adatközpontjában levő szerverről, másik helyi szolgáltató adatközpontjából és tengerentúli adatközpontból.

A fenti paraméterek egyértelműen letöltési sebesség (rendelkezésre álló sávszélesség) mérését teszik lehetővé a szolgáltató saját hálózatán belül, belföldi kicserélő ponton keresztül és a leggyakrabban használt nemzetközi kapcsolatot (USA) használva a felhasználók jellegzetes tevékenységének (böngészés, fájlok letöltése/feltöltése) figyelembevételével és azt is szem előtt tartva, hogy a felhasználók minőség érzékelésének részét képezi a szolgáltatók sorozatán keresztüli információeléréssel kapcsolatos észlelés is.

Az 5. ábrán öt szolgáltató esetében mutatjuk be a felhasználóknak „ígért” és a ténylegesen mért letöltési sebességet a szolgáltató saját adatközpontja tengerentúli adatközpontok elérése esetén. Látható, hogy a felhasználó által érzékelhető sávszélességek optimális esetben sem érik el a névleges sávszélességet, azaz a szolgáltatók az ígért szolgáltatási minőséget nem tudják teljesíteni.

5. ábra

Ígért és tényleges letöltési sebességek öt szolgáltatónál



5. A letöltési sebesség mérésének problémái

Ha a letöltési sebesség mérésénél a mérést végző rendszerek által okozott minden járulékos hatást ki tudnánk küszöbölni, akkor is komplex mérőszámot kapnánk, hiszen a mérés eredménye függ a rendelkezésre álló sávszélességtől, a csomagtovábbítás során igénybe vett útvonalon rendelkezésre álló sávszélességtől, a terhelés dinamikus változásától, a bithibák és esetleges elárasztási szituációk okozta csomagvesztéstől, a késleltetési időtől, a használt protokoll sajátosságaitól.

A felsorolt QoS paraméterek mérése egyenként megoldható, viszont az intenzív kutatás ellenére a modellek (például [7]) nem tudják 10%-on belüli pontossággal jóslni ezek tetszőleges kombinációja esetén a felhasználó által is érzékelhető letöltési sebességet. Ennek megfelelően e paraméterek mérése vagy bizonyítása a felhasználói oldalról nézve nem igazán releváns.

Amennyiben a letöltési sebességet referenciaszerver segítségével kívánjuk mérni, ahogy azt a fenti példákban láttuk, szintén komoly problémákba ütközünk.

A felhasználóknak ugyanis jogos elvárása, hogy a mérés általa is végrehajtható legyen, azaz valamilyen módon ellenőrizni tudja, hogy a szolgáltató teljesíti-e a vállalt minőségi paramétereket.

Ahhoz, hogy a mérés ne befolyásolja lényegesen a hálózat teljesítményét, az adott hálózati csomópontra jellemző aggregált sávszélesség 5%-át nem haladhatja meg a mérés okozta terhelés. A BIX „legerősebb” kicserélő központját (<http://www.bix.hu>) nézve tehát nem szabad 13 Gbit/s-nél nagyobb összforgalmat generálni a mérő szerveren, de ha azt is figyelembe vesszük, hogy egy interfészen sem célszerű megengedni, hogy a méréssel okozott terhelés 5% fölé kerüljön, a mérésekre lefoglalható tényleges sávszélesség ennél is kisebb.

Az elfogadható pontossághoz és a rendszerekből adódó járulékos hatásokat kiküszöbölő három perces méréseket feltételezve [1] egy 1 Mbit/s sávszélességű interfészen 22.5 Mbyte adat tölthető le.

Természetesen fennáll annak a lehetősége is, hogy a mérést felhasználói szempontból ugyanolyan szolgáltatásnak tekintsük, mint bármelyik webszolgáltatást és ne korlátozzuk az általa használható sávszélességet. Ennek hátránya az, hogy a mérés statisztikai jellemzői eltérnek a szokásos felhasználástól, aminek következtében megváltoznak azok a feltételek, amelyek segítségével a szolgáltató a vállalt minőségi paramétereket kalkulálta.

A felhasználó oldali névleges hozzáférési sávszélességek már jelen pillanatban elérik a 18 Mbit/s-t (lásd Tantsz weboldal), ami azt jelenti, hogy egy 10 Gbit/s sávszélességű interfésszel és megfelelő

processzási sebességgel rendelkező szerver elvben egyidőben mintegy 550 ilyen kapcsolat mérésére alkalmas. Ha azt is figyelembe vesszük, hogy az egyes szolgáltatók felé az interfészek sávszélessége nem haladja meg a 10 Gbit/s-t, úgy az egy szolgáltató oldalán az egyidejű méréseket a jelen állapotban is korlátozni kéne 33-ra, illetve a csak 1 Gbit/s sávszélességgel rendelkező szolgáltatók esetén 3-ra. (Ha a jellemző hozzáférési sávszélességet 1 Mbit/s-re becsüljük – ami a jelenlegi helyzetet bár nem sokáig, de jobban jellemzi –, akkor az említett számok 18-szorosát kell figyelembe venni. A nagyobb értékek mellett szól azonban, hogy a referencia szerver szintjén nem lehet elkülöníteni a nagyobb sávszélességgel rendelkező üzleti felhasználókat, valamint az egyetemi hálózatban lévő 100 Mbit/s hozzáférési sávszélességgel rendelkező százezres nagyságrendű felhasználói tömeget.)

A helyzetet tovább bonyolítja, hogy a mérés jellegéből adódóan a mérő szerver elérési sávszélességét nem lehetne korlátozni, amiből következően a felhasználók sorba rendelését kellene megoldani, ami azért nem felhasználóbarát megoldás, mert a felhasználók általában egy konkrét szituációban (időpontban) szeretnének meggyőződni a szolgáltatás minőségéről, jelsül akkor, amikor valamilyen problémát, „lassulást” tapasztalnak. Ebben az esetben viszont a mérési igény tömegesen jelentkezik.

Ha a portugál példából okulva a felhasználók azon igényét is figyelembe vesszük, hogy szolgáltatókon átnyúló szolgáltatások minőségi paraméterei is elfogadható tartományon belül legyenek, tovább csökken az egy időben kiszolgálható felhasználók száma, hiszen ezek forgalma is a kicserélő központon menne át.

A másik lehetőség – ahogy azt a példákban is láttuk – a független szervezet segítségével való mérés. Ennek tagadhatatlan előnye a mérések kézben tartása, viszont hátránya a felhasználók és a szervezet közötti koordináció szükségessége, a reprezentatív minta kiválasztásának problémája, a végezhető mérések relatíve kis száma, valamint a mérések és a felhasználói érzékelés elkülönültsége különös tekintettel a mérésben nem résztvevő felhasználókra. Mindezek mellett továbbra is igaz, hogy a független méréseknek bizalomerősítő és a szolgáltatók magatartását jótékonyan befolyásoló hatása van.

6. A fejlődési irányok

Az Internet folyamatos terjedése, a technológia fejlődésének következményei és a konvergencia a szolgáltatásminőséget jellemző paraméterek iránti igényekre is kihatnak. A szolgáltatás jól jellemezhető volt sávszélesség (letöltési sebesség) jellegű adatokkal egészen

addig, amíg a felhasználók adatállományokat töltöttek le, akár html-oldalak, akár elektronikus levelek formájában. A mindennapi jelenünkhöz tartozó kommunikációs alkalmazások megjelenése, illetve az adatfolyam jellegű (streaming) médiaszolgáltatások elterjedése szükségessé teszi további QoS paraméterek, mint például a késleltetés (delay), a késleltetésingadozás (jitter), a csomagvesztés (packet loss) értékek figyelembe vételét.

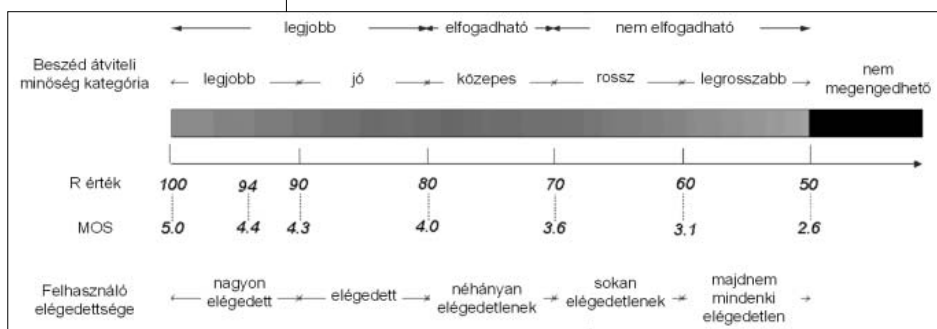
Mivel ezek jellegzetesen olyan paraméterek, amelyeket a fogyasztó csak közvetve érzékel, különösen fontos, hogy világos legyen hatásuk a fogyasztó által érzékelt minőségre. Jó példa az e területen folyó kutatásra az úgynevezett R érték (6. ábra) kidolgozása, amely világos kapcsolatot teremt a hangátvitel minőségét befolyásoló paraméterek és a felhasználók szubjektív minőségérzékelésén alapuló MOS (Mean Opinion Score) között az ITU G.107 és G.108 szabványában meghatározottak szerint a következők figyelembe vételével [4],[6]:

$$R = R_0 - I_S - I_D - I_E + A,$$

ahol R_0 kapcsolatos az elektronika és környezet okozta jel/zaj viszonytal, I_S jellemzi a digitalizálás és a csomagvesztés okozta torzítást, I_D tartalmazza a késleltetés és a visszhang okozta minőségvesztést, I_E pedig kifejezi a kódoló/dekódoló alacsony bitsebességét, valamint a késleltetés ingadozást, végül A reprezentálja azokat a tényezőket, amelyek kompenzálni képesek az előzőekben figyelembe vett hatásokat (például mobilítás).

A technológiai fejlődés és a fogyasztói igények növekedése újabb „killer”, azaz erőforrásigényes alkalmazások bevezetését eredményezi. [6] szerint a ma már

6. ábra Az R-érték

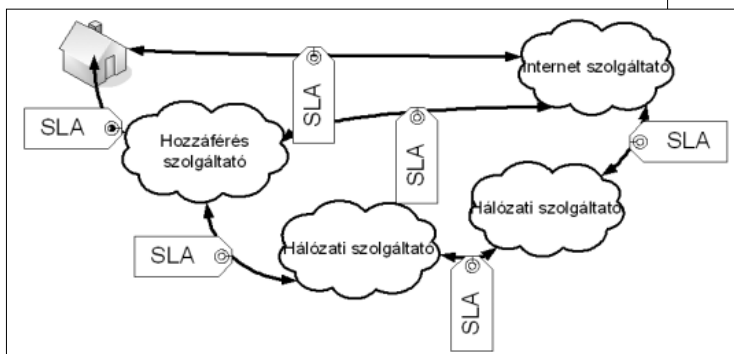


elterjedt párbeszédes alkalmazások mellett teret fognak hódítani olyan alkalmazások, mint

- minőségi, több forrásból származó audió, ahol a kis késleltetés és a szinkronitás fontos,
- professzionális, legalább CD minőségű audió,
- minőségi audió- és videokonferencia és együttműködési platform,
- minőségi, több forrásból származó videó,
- orvosi alkalmazások, mint a távsebészet vagy a távdiagnózis,
- valódi interakciót lehetővé tevő távoktatási alkalmazások,
- nagy adatmennyiségeket megmozgató grid-számítási alkalmazások,

- videó közvetítés, videó jelfolyamok, igény szerinti videó (VOD),
- táv-együttlét (például 3D közvetítések),
- eszközök távoli vezérlése

A felsorolt alkalmazástípusok más és más QoS-paraméterkombinációk, illetve adott esetben valószínűségi eloszlások mellett érhetnek el adott érzékelt szolgáltatásminőség-szinteket. Mivel ezek jelentős része még nem képezi részét a mindennapi gyakorlatnak, a mérhető QoS jellemzők és az érzékelt szolgáltatási minőség közötti összefüggés nem tisztázott.



7. ábra Illusztráció az SLA-egyezményekhez

[5] szerint a professzionális szolgáltatások mellett legalább még két tényező figyelembe vétele indokolt, amelyek egyre fokozódó hangsúlyt kapnak a már hírközlési szektor közép-távú fejlődése során is, ezek a mobilitás és a biztonság (security).

A felhasználóknak jogos igénye lesz az, hogy az előfizetett (adott esetben csak megszokott) szolgáltatásokat megfelelő biztonsággal a világ tetszőleges pontjáról vegyék igénybe. Ennek előfeltétele, hogy a szolgáltatási láncban résztvevő szolgáltatók szolgáltatás-szint-egyezményeket kössenek egymással (7. ábra), de ez követelmény újabb kihívás elé állítja az érzékelt szolgáltatási minőséget hitelesen mérni kívánó szervezeteket is. A biztonság alapvető követelménye, hogy a teljes átvitel során biztosítva legyen az adatok sértetlensége, bizalmassága és letagadhatatlansága. A fentiek mellett a felhasználók magánéletéhez való jogának érvényesítése (privacy) is egyre nagyobb hangsúlyt kap.

A fogyasztói megelégedettség szempontjából és az igények növekedtével egyre fontosabbá válik a fogyasztók tájékoztatása (lásd 2. ábra – Beszámolás), azaz a mért QoS paraméterek fogyasztói érzékeléshez közeli jellemzőkre konvertálása és nyilvánosságra hozása. Ezzel egyidőben egyre fontosabbá válik a fogyasztók képzése annak érdekében, hogy a közölt adatokat helyesen legyenek képesek értelmezni, illetve csökkenjen a szolgáltatás elvárt szintje és tényleges lehetőségei közötti különbség.

7. Összefoglalás

Az Internet-szolgáltatásokat jellemző alapvetően technológiai jellegű minőségi paraméterek nem állnak a fogyasztók által is jól követhető összefüggésben az érzékelt szolgáltatási minőséggel. A hazai és nemzetközi gyakorlat, valamint a kutatás megpróbál lépést tartani a felhasználók azon igényével, hogy a szolgáltatásokat objektíven mérhető, de az érzékelt minőséghez közeli paraméterekkel jellemezze, de már most látható, hogy a felhasználók számára korlátozás nélkül elérhető megfelelő pontosságú mérési lehetőség megvalósítása hosszú távon nem valószínűsíthető.

További problémát jelent, hogy a felhasználói igények és tudatosság növekedésével a mérések iránti igény úgy nő, hogy egyúttal a mérendő paraméterek száma, illetve az alkalmazások komplexitása is növekszik, ezért további kutatás szükséges annak érdekében, hogy a szolgáltatás triviálisabban mérhető adataiból számítással lehessen megállapítani az egyes fogyasztók által érzékelt minőséget jól jellemző adatokat.

Irodalom

- [1] NHH H 345 projekt munkacsoport: Kézikönyv az elektronikus hírközlési szolgáltatás minőségének a fogyasztók védelmével összefüggő követelményeiről szóló 345/2004. (XII. 22.) Korm. rendeletben foglalt követelmények teljesítéséhez, Nemzeti Hírközlési Hatóság, 2005.
<http://www.nhh.hu/dokumentum.php?cid=10000&letolt>
- [2] SLA Management Handbook, Volume 2, Concepts and Principles, Release 2.5. TeleManagement Forum, 2005.
<http://www.itu.int/md/T05-SG04-050920-TD-WP2/en>
- [3] Evaluation of the Internet Access Service, Broadband: ADSL and Cable, 2006.
http://www.anacom.pt/streaming/adslcabo06en.pdf?categoryId=182223&contentId=335185&field=ATTACHED_FILE
- [4] SLA Management Handbook, Volume 4, Enterprise Perspective, The Open Group, 2004.
<http://www.opengroup.org/bookstore/catalog/g045.htm>
- [5] Purple Book v2, Part 2, The CELTIC Initiative, Cooperation for a European sustained leadership in telecommunications, 2003.
<http://www.celtic-initiative.org>
- [6] A survey of network QoS needs of advanced internet applications, 2002. Internet2 QoS Working Group,
<http://qos.internet2.edu/wg/apps/fellowship/Docs/Internet2AppsQoSNeeds.html>
- [7] P. Sevcik, J. Bartlett: Understanding Web Performance, Business Communication Review, 2001.
[http://www.apmadvisors.com/Articles/BCR Article Web Performance FNL.pdf](http://www.apmadvisors.com/Articles/BCR%20Article%20Web%20Performance%20FNL.pdf)

Minőségi szolgáltatások ADSL környezetben

NAGY TIBOR

Cisco Systems, Inc.

tinagy@cisco.com

Kulcsszavak: QoS, ADSL, forgalomszabályozás, triple-play

Az ADSL környezetben megjelenő értéknövelt szolgáltatások (hang, IPTV, video streaming) átvitele új igényeket támaszt a hálózattal szemben. Ezeknek egy része megnövekedett sávszélességigényben jelentkezik, másik része a hálózat késleltetésére, illetve torlódás esetén a csomagvesztésre érzékeny alkalmazások megfelelő szintű kezelésére vezethetők vissza. A jó minőségű átvitel biztosítására a sávszélesség növelése nem minden esetben nyújt kielégítő megoldást, a Cisco Systems QoS tervezésre vonatkozó dokumentumai még 100 Mbit/s-os, sőt annál nagyobb sebességű Ethernetes hálózatokban is javasolják a paraméterek megfelelő hangolását. A sokfelhasználós ADSL-hálózatokban továbbá különösen gyakori igény, hogy a hálózat különböző pontjain megjelenő felhasználói sávszélességeket intelligens módon – a szolgáltatási szerződésnek megfelelően – korlátozni kell. Ezekre az ADSL környezetben jelentkező feladatokra a Cisco Systems egy speciális QoS modellt dolgozott ki, amelynek architektúrájáról és gyakorlati megvalósításáról a cikkben részletes információt adunk.

1. Bevezetés

Az IP hálózatokon – így az Interneten is – alkalmazott technikák fejlődése, illetve a szélessávú hozzáférés elterjedése lehetővé teszi a szolgáltatók számára olyan értéknövelt szolgáltatások kifejlesztését, mint a garantált minőségű hangátvitel, valamint az ADSL környezetben is biztosított IPTV szolgáltatás megjelenése.

Az Internet-szolgáltatás mellett nyújtott telefonkapcsolat és az IPTV új követelményeket támasztanak a meglévők mellé, amelyeket a gyártók speciális QoS technológiák alkalmazásával igyekeznek kielégíteni.

A Cisco Systems már több, mint egy évtizede felismerte a multimédiás alkalmazások jelentőségét és hálózati eszközeiben – útválasztók, kapcsolók, egyéb komponensek – sorra jelentek és jelennek meg olyan mechanizmusok, melyek képesek nagy pontossággal azonosítani (classification), megjelölni (marking), és prioritással kezelni (scheduling with prioritization) késleltetés-, illetve csomagvesztés-érzékeny forgalmakat.

A forgalom precízen kontrollált részének bufferelési technikák nélküli eldobását végző alkalmazását „policing-nek”, buffereléssel megvalósított – a forgalomra nézve kevésbé drasztikus – technológiáját „shaping-nek” nevezzük.

2. Felhasznált hálózati komponensek

A sokunk által használt ADSL környezetben a QoS technológiákat az aggregációs hálózatban a következő komponenseken valósítjuk meg:

- CPE (Customer Premises Equipment, azaz az ügyfélnél elhelyezett eszköz)
- DSLAM (Digital Subscriber Line Access Multiplexer, azaz ADSL vonalakat fizikailag aggregáló eszköz)
- BNG (Broadband Network Gateway,

korábban BRAS - Broadband Remote Access Server, azaz a felhasználói forgalmat logikailag aggregáló eszköz)

- A BNG és a DSLAM közötti napjainkban leginkább Ethernet-alapú aggregációs hálózat

Mivel a megfelelő QoS paraméterek által biztosított szolgáltatás végponttól végpontig értelmezendő, természetesen a hálózat többi elemének (gerinchálózati útválasztók, tartalom szolgáltató szerverek stb.) szintén biztosítani kell a csomagok megfelelő szintű kezelését (megjelölés, priorizálás stb.).

Az aggregációs komponensekre vonatkozóan a DSL világban meghatározó ajánlásokat készítő DSL Fórum számos dokumentuma ad QoS iránymutatást. Ezek közül a javasolt QoS architektúrákról, az elvárt paramétereikről és a technológiai megvalósítási modellekről leginkább a TR-059 [5] ajánlásban (ATM alapú aggregációs hálózat a DSL mögött) és a TR-101 [6] (Áttérés ATM alapú aggregációról Ethernet alapú aggregációra) megfelelő fejezeteiben olvashatunk.

Tekintettel arra, hogy manapság leginkább az Ethernet-típusú uplinkkel rendelkező DSLAM-okat használják a szolgáltatók, az alábbi technológiai áttekintés elsősorban erre a környezetre értelmezendő, habár fontos tudni, hogy a TR-059-ben leírt ATM alapú aggregációs modellnél alkalmazott QoS alapelvek nagy része megfeleltethető az Ethernet modellnél alkalmazottaknak.

Az 1. ábra egy tipikus Ethernet-alapú aggregációs környezetet szemléltet.

3. QoS megvalósításának követelményei az aggregációs hálózatban

Az ügyfélnél kihelyezett eszköztől a logikai aggregációt végző BNG eszközig terjedő komponensek esetében az alábbi követelményekkel kell számolni.

3.1. Általános követelmények

- Kerüljük el a downstream (előfizető felé) irányuló forgalom kontroll nélküli csomagvesztését a DSLAM eszközön.
 - Ez akkor lehetséges, ha a DSLAM támogatja az előfizető irányába néző ADSL interfészén a megkülönböztetett torlódás vezérlést (differential scheduling). A megvalósítás lehet több ATM virtuális áramkörre alapuló (úgynevezett multi-VC megoldás), ahol a különböző típusú forgalmakat különböző ATM PVC-ken visszük az előfizető ADSL CPE-je felé, vagy DiffServ [7] technológiára épülő, ahol kiemelt prioritással (EF) kezeljük az érzékeny forgalmat.
 - Amennyiben a DSLAM nem támogatja az előzőekben említett technológiákat, akkor a DSLAM előtti eszköz(ök)-ben kell megvalósítani azokat (policing/shaping).
- Egy szolgáltatási osztályhoz minimálisan garantált sávszélesség biztosítása
 - Késleltetés és csomagvesztés tekintetében más szolgáltatási osztályoktól független minimum sávszélességet kell megfelelő prioritációval biztosítani, hiszen az előfizető jogosan elvár egy elfogadható le- és feltöltési sebességet.
- Egy szolgáltatási osztályhoz (pl. Internet) a maximális igénybevehető sávszélességre vonatkozó korlátok alkalmazása.

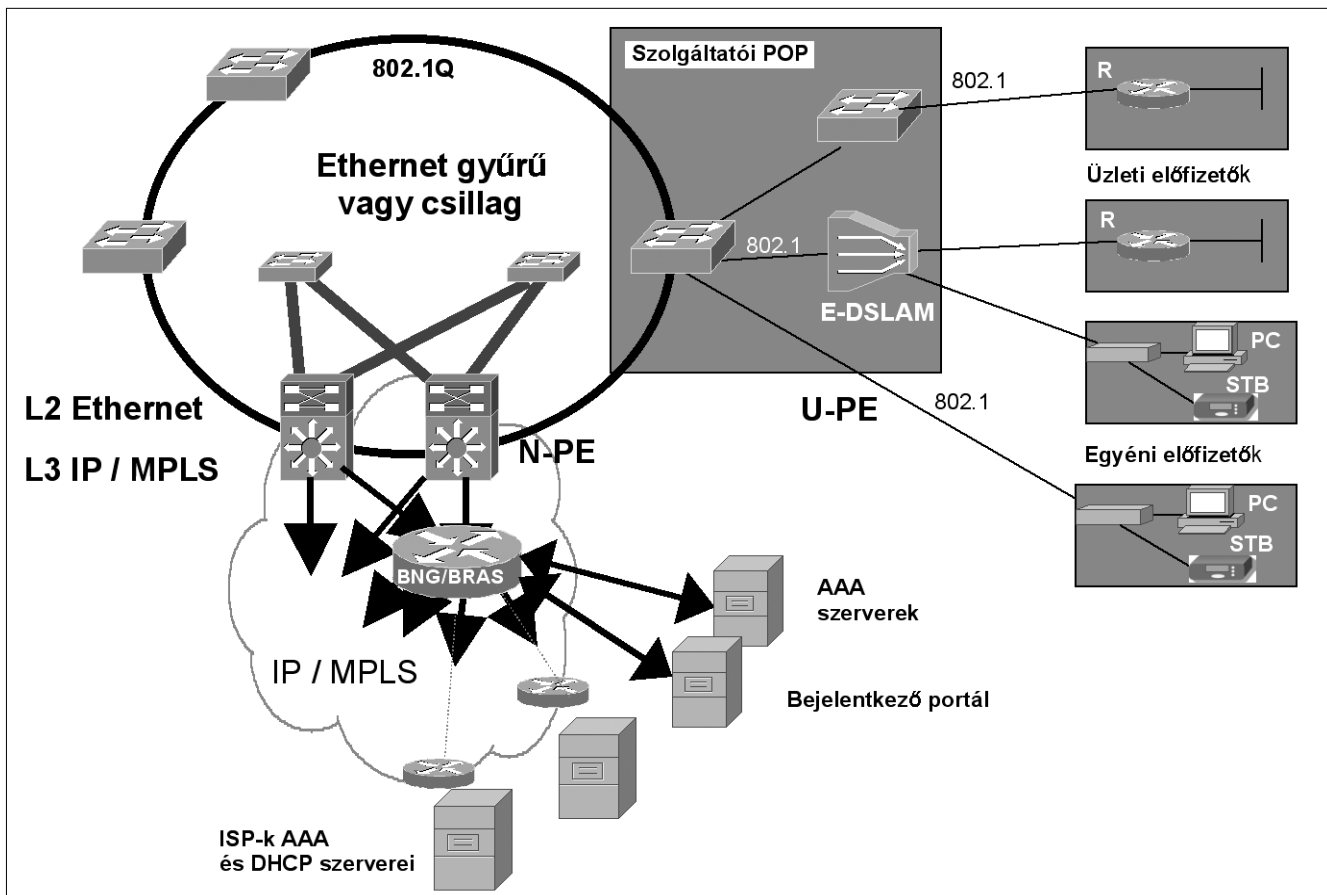
- Forgalmi korlátok érvényesítése a forgalom mindkét irányára alkalmazva (policing/shaping). Ez azért fontos, mert az IP a technológiából fakadóan sávszélességigényes, elveszi, ami csak rendelkezésre áll. Természetesen a szolgáltató nem szeretné, ha valamely előfizetője esetleg más előfizető forgalmának kárára venne igénybe túl nagy sávszélességű szolgáltatást.
- Valamely szolgáltatási osztály által nem használt sávszélesség használatának lehetősége más forgalmi osztályok számára. Ha már IP-ről beszélünk, miért ne alkalmaznánk a technológia által automatikusan nyújtott dinamikus sávszélesség kihasználást, hiszen itt nincs fix sávszélesség-foglalás, mint korábban például az időosztásos (TDM) rendszereknél.

3.2. Üzleti előfizetők követelményei

Az általános követelményeken túl az üzleti előfizetőknek szóló csomagok követelményei az alábbiak.

- Egy adott előfizető maximális aggregált sávszélességének korlátozása
 - Ahhoz, hogy úgy biztosítsuk a maximális sávszélességet, hogy közben szolgáltatási osztályonként a garantált minimális sávszélesség azért rendelkezésre áll, valamint a más szolgáltatási osztályok által nem használt sávszélesség igénybe vehető, hierarchikus QoS implementációra van szükség, amellyel „per előfizető” alapon sha-

1. ábra Ethernet-alapú aggregációs környezet



ping algoritmust, a visszafogott forgalmon belül pedig a forgalmi osztálynak megfelelő prioritizálást hajtunk végre (például a hangcsomagokat előre vesszük a bufferelési technikáknál).

- Támogatni kell az úgynevezett QinQ aggregációs modellt, amely tulajdonképpen a VLAN azonosítóval ellátott forgalom újabb addicionális VLAN azonosítóval történő ellátását jelenti.

- Ez azért lényeges, mert előfordulhat olyan szituáció egy nagyobb hálózatban, hogy az IEEE 802.1q szabványban rögzített 12 bites VLAN azonosító által biztosított maximálisan 4096 azonosító kevésnek bizonyul, mert egy adott aggregációs eszköz Layer3 interfészén több, mint 4096 előfizető ül.

A QinQ technikánál viszont 2x12 bit áll rendelkezésre, az előfizetőt egyértelműen azonosíthatja a belső VLAN azonosító, a külső azonosító pedig az üzleti előfizetők egy csoportját, például a DSLAM-ot jelölheti.

3.3. Egyéni előfizetők követelményei

Az általános követelményeken túl az egyéni előfizetőknek szóló csomagok követelményei:

- Több különböző ponton történő forgalom betáplálása a hozzáférési hálózatba, multicast (IPTV) forgalom továbbításának lehetőségével

- Internet-nagykereskedelmi (wholesale) modell támogatása.

- Hasonló logikai alapon, mint korábban az ATM technológiánál, a Virtális Útvonal (VP) – Virtuális Áramkör (VC) modellnél.

- Per session alapú szolgáltatás kezelés (session-on PPPoE, illetve IP sessiont értünk).

- Előfizetői vonalanként több sessiont is tudni kell kezelni. Gyakran előfordul ugyanis, hogy az Internetes forgalom PPP beágyazással, az IPTV pedig natív IP-ben érkezik az előfizetőhöz.

- VLAN aggregációs modell támogatása.

- Nagy számú előfizető esetén kevés lehet a normál VLAN címke által biztosított 4096 azonosító, mint ahogy azt korábban kifejtettük.

Habár a DSL Fórum TR-101-es ajánlása nagy hangsúlyt fektet a DSLAM QoS képességeire, ezeket a mechanizmusokat a gyakorlati megvalósítás hiánya miatt a szolgáltatók jelenleg nem, vagy csak korlátozott mértékben képesek igénybe venni.

Éppen ezért a DSLAM mögötti aggregációs hálózati komponenseknek (Ethernet switch, BNG/BRAS) kell biztosítaniuk a hiányzó funkciókat.

4. A Cisco QoS modellje DSL architektúrára

A megvalósításra a Cisco számos modellt dolgozott ki, közülük most a klasszikus ATM-alapú aggregációhoz leginkább hasonló (VP és VC szinten is forgalomszabályzott (shaped)) architektúra részletes bemutatása következik.

Nézzük meg, milyen technológiákat alkalmaz a Cisco az aggregációs eszközökben.

• VP shapingnek megfelelő technológia

A felhasználók egy csoportjának forgalmát egy előre definiált értékre szabályozzuk le (shaping), ahol az egyéni előfizetők egy csoportját leginkább egy VLAN azonosítóval, az üzleti előfizetők egy csoportját pedig a QinQ címke külső VLAN értékével azonosítjuk.

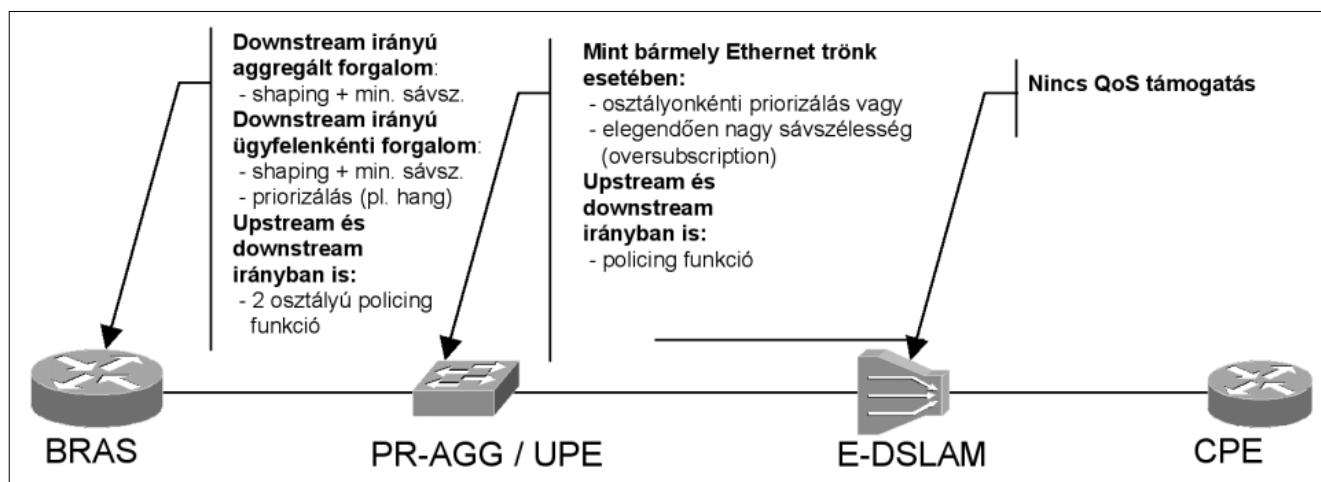
Az oversubscription, azaz vonali túlkönyvelést nem az előfizetői csoportok leszabályzott forgalmának az interfész sebességénél nagyobb értéke adja, az már a VLAN-ok sáv szélességének kialakításakor megtörténik. Mivel nem előfizetőnkénti alapon fogjuk vissza a forgalmat, hanem előfizetői csoportonként, az ADSL fellett lévő ATM technológia használatából fakadó overhead nem jelentős.

• VC shapingnek megfelelő technológia

Az előfizető irányába menő forgalmat a DSL vonal szinkronizálási sebességére, vagy az alá vesszük le.

Az egyes egyéni előfizetői forgalmak azonosítása a DHCP-ben használt 82-es opcióval történhet, az üzleti előfizetőket a többszörös VLAN tag esetén pedig a címkék azonosítják egyértelműen.

2. ábra



- Mivel az interfész-sebességnél nagyobb az előfizetői forgalmak össze-sített értéke, ezért minimális sávszé-lesség-garancia is szükséges.
- A késleltetés-, illetve csomagvesztés-érzékeny forgalmi típusoknak (hang, videó) kétszintű prioritizációt kell biz-tosítani.

Ennek a javasolt megoldásnak fontos előnye, hogy nem kell QoS támogatás a DSLAM eszközben, továbbá támogatja a nagykereskedelmi (wholesale) koncepci-ót, hátránya, hogy Denial-of-Service (DOS) támadások ellen az előfizetőtől jövő for-galmat még külön korlátok között kell tar-tani, például policing funkció segítségével, hogy ne legyen képes egyetlen előfi-zető a szolgáltatásban zavart okozni.

5. Fragmentáció, avagy a csomagok feldarabolásának szükségessége

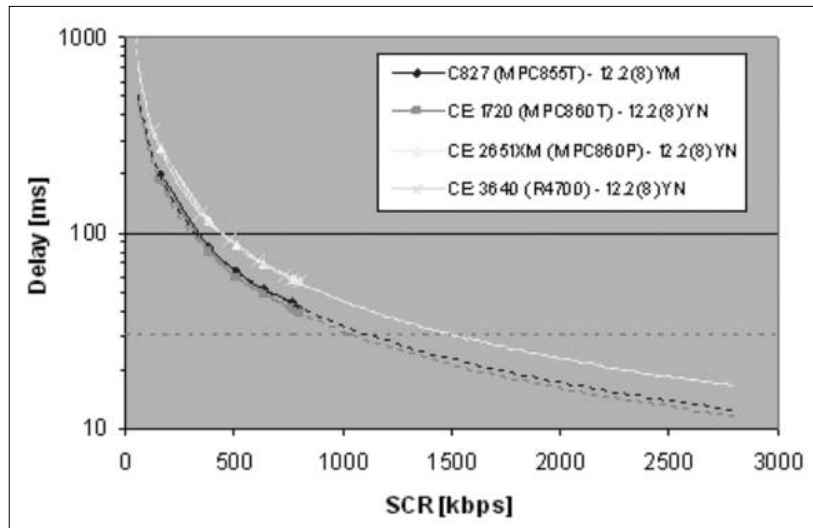
Az ADSL technológia fontos jellemzője az aszimmetria, tehát az előfizető felé (downstream) és az előfizetőtől jövő (upstream) forgalom sávszélességének különbö-zősége. A downstream-forgalom maximuma 8 Mbit/s ADSL és 25 Mbit/s körüli ADSL2+ technológia esetén. Az upstream-forgalom viszont ezzel szemben gyakran csak 512 kbit/s, maximális értéke 1Mbit/s körüli ADSL2+ esetén.

A hangforgalom megfelelő minőségének biztosítása érdekében a késleltetését és késleltetésének változá-sát minimális értéken úgy lehet tartani, ha a végpont-tól-végpontig számított késleltetésben fontos szerepet játszó serialization komponenst, vagyis a csomagok a vonalra helyezésének időtartamát minimális értéken tartjuk.

Könnyen belátható, hogy minél rövidebb csomago-kat kell a vonalra rakni (vagy minél gyorsabb vonalra rakjuk rá őket), annál rövidebb ideig tart ez a művelet, tehát adódik a megoldás, hogy a hosszú – FTP forga-lom esetén Ethernet-becsomagolással 1500 byte-os adatcsomagokat – fel kell darabolni és a feldarabolt adatcsomagok közé hangcsomagokat lehet illeszteni.

Ha ezt a műveletet a BRAS végzi, akkor a végpont-nak (CPE) természetesen vissza kell állítani a feldara-bolt keretkből az eredeti keretet.

A DSL környezetben rendszerint útmutató DSL Fór-um konzervatív módon 30 ms-ben határozta meg a hozzáférési vonal késleltetésének maximumát, ahhoz, hogy elfogadható minőségű hangszolgáltatást lehes-sen nyújtani az adott infrastruktúrán. Annak érdeké-ben, hogy ezt az értéket tartani tudjuk, a Cisco a külön-böző CPE berendezésein (827,1720,2651,3725 stb.) méréseket végzett és arra a következtetésre jutott, hogy a korábban általánosságban hangoztatott 768 kbit/s-os határérték, amely alatt a Cisco mindenképpen javasolt valamilyen vonali fragmentációt, ilyen környe-zetben 1 Mbit/s körüli értékre módosul.



3. ábra Upstream irányú késleltetés alakulása

A 3. ábra a fragmentáció nélküli upstream irányú kés-leltetési értékeket mutatja a Cisco különböző CPE plat-formjain.

5.1. Fragmentációs megoldások DSL környezetben

5.1.1. PPPoA környezet

Az ADSL világban korábban elterjedt PPPoverATM (PPPoA) technológia esetében rendelkezésre áll jól definiált szabvány (DSL Fórum TR-59, WT-92 [8]) és megvalósításra is került a korábban tipikusan bérelt vo-nalakon fregmentációra használt Multilink PPP – Link Fragmentation and Interleaving (MLPPP/LFI) mecha-nizmus, amely pontosan a fenti követelményeknek meg-felelően elvégzi a nagy méretű adatcsomagok feldara-bolását (és visszaállítást), valamint közéjük a kis hang-csomagok beillesztését.

5.1.2. PPPoE környezet

A manapság sokkal elterjedtebb PPP over Ethernet technológia esetében két megoldás is kínálkozik.

• Multi-VC megoldás

Jogos elvárás az Ethernetes DSLAM-okkal szem-ben, hogy ATM QoS-t támogassanak, különböző ATM forgalmi osztállyal működő ATM virtuális áramkörökön különböző típusú forgalmakat szállítva. Ez meg is old-ja a problémát, hiszen ennél a megoldásnál nincs szük-ség a felettes rétegeken fragmentációra, az ATM PVC-k megfelelő módon elkülönítik az Internetes adatforgal-mat a hang-, videó- illetve IPTV-forgalomtól.

Természetesen ehhez arra van szükség, hogy a kü-lönböző forgalmak szétválasztása és megfelelő para-méterekkel ellátott virtuális áramkörökbe (PVC) irányítá-sa mind a DSLAM, mind pedig a CPE oldalon támoga-tott legyen.

Több lehetőség is adott a hozzárendelés megvaló-sítására:

- Az Ethernet fejléc 802.1p mezőjének értéke alapján.
(Ez az úgynevezett VC bundlingnak felel meg.)

- A csomagot fogadó fizikai vagy logikai interfész alapján (tipikusan VLAN azonosító a DSLAM oldalon és fizikai interfész a CPE oldalon).
- A DSLAM-ban, illetve a CPE eszközben lévő Layer 2-es vagy esetleg Layer 3-as csomag-továbbítási logika alapján.

• *MLPPP/LFI PPPoE esetében*

Amennyiben PPPoE fut a CPE és a BRAS között, a MLPPPoE/LFI egy potenciális megoldás lehet, habár ennek a gyakorlati megvalósítása a Cisco platformokon csak limitáltan áll rendelkezésre.

Problémák a „jól bevált” MLPPP-vel Ethernetes ADSL környezetben:

- Priorizált forgalmi osztály esetén a MLPPP egyszerű PPP fejléccet használ, míg nem priorizált forgalom (fregmentált) esetén pedig MLPPP fejléccet. A PPPoE esetében az összes forgalom fregmentált és MLPPPoE fejléccel ellátott.
- A másik probléma lehet, hogy a MLPPP session végponttól végpontig terjed, ezért amennyiben az ügyfél oldalon PC van és nem különálló CPE berendezés, a PC-n futó PPP sztekknek is támogatnia kell a MLPPPoE-t, ami nem általános.
- A hangcsomagok esetében a feldarabolt adat-csomagok közötti illesztéshez valamilyen intelligens bufferelési megoldásra (shaping with queuing) van szükség, ami PPPoA esetében viszonylag egyszerűen megoldott, itt viszont lefelé (downstream) irányban szolgáltatási osztály/felhasználó szintű forgalom korlátozásra (shapingre) van szükség.

• *ADSL2+*

Az ADSL2+ esetében rendelkezésre áll – megfelelő minőségű rézérpár és korlátozott távolság esetén – akkora sáv szélesség az előfizetőtől a DSLAM felé (upstream), hogy ne legyen szükség fregmentációra. A DSLAM és CPE gyártók döntő többsége jelenleg már olyan berendezést gyárt, amely támogatja az ADSL2+ technológiát.

6. A QoS szabályok beállítása a Cisco aggregációs eszközeiben

A CPE berendezésekben beállításra kerülő szabályok (például hangforgalom esetén prioritizáció) legtöbb esetben manuális módon az eszköz beállító felületéről (GUI vagy CLI) könnyen elvégezhetők.

A logikai aggregációt végző BNG/BRAS eszközön az eddigiekben tárgyalt szabályrendszer beállítását és alkalmazását az egyéni illetve üzleti előfizetők forgalmára leggyakrabban dinamikus módszerrel, RADIUS attribútumok automatikus letöltésével célszerű elvégezni.

A nagy előfizetői bázissal rendelkező szolgáltatók szinte mindegyike ezt a módszert használja, azaz az előfizető irányába menő downstream illetve az előfizetőtől jövő upstream forgalomhoz – amely a BNG (BRAS)

eszközben egy virtuális interfészen keresztül halad – dinamikusan rendelik hozzá a megfelelő QoS szabályokat.

Ez a technológia még olyan környezetekben is kiválóan alkalmazható, ahol a hang és adatforgalom, valamint az IPTV külön válik egymástól a hálózatban.

A Cisco ajánlása Triple-play környezetekre disztributált szolgáltatás termináláson és továbbításon alapul, tehát az IPTV forgalomnak nem célszerű a BNG-n keresztül haladnia, hanem önálló, Layer3-as szinten processzált forgalomként „route-olódik” a hálózatban.

A BNG eszközön tipikusan Internet forgalomra beállított QoS értékeket tehát a következő módon, három lépésben állíthatjuk be:

- Forgalmi osztályok kialakítása.
- A forgalmi osztályok és a szabályrendszer összerendelése.
- A RADIUS szerver az adott felhasználói név/jelszó párhoz az előre beállított szabályrendszer nevét tölti le az eszközbe, ami aztán a konfigurációban a szabályrendszer alatt definiált parancsokat (forgalomkorlátozás-shaping, policing stb.) egymás után végrehajtja.

A RADIUS szabvány lehetőséget ad úgynevezett gyártóspecifikus attribútumok használatára (VSA), amelyek esetében a VSA mező értéke (szabályrendszer neve) letöltésre és értelmezésre kerül a hálózati eszközben, jelen esetünkben BNG/BRAS-on.

Létezik upstream (VSA37) és downstream (VSA38) irányú attribútum is a különböző típusú forgalmak kezelésére.

7. Universal Subscriber Edge (USE) architektúra

Az eddig tárgyalt metódusok kitűnően alkalmazhatók a Cisco eszközökön, a beállítások finomhangolásával jó minőségű adat/hang/videó továbbítás érhető el.

A hálózati trendeket követő szakértők az ADSL-en nyújtott szolgáltatások jövőbeli kibővülésére hívják fel a figyelmet, azaz olyan értéknövelt szolgáltatásokra számíthatunk, mint a felhasználó által dinamikusan, webes felületen állítható le- és feltöltési sebesség (Turbo gomb használata), vagy a szintén webes felületen beállítható felhasználói profilok megjelenése, amellyel például a szolgáltatásért fizető szülők egyszerű módon tudnak biztonságos tartalomszűrő feltételeket definiálni családtagjaik részére.

Az ilyen típusú szolgáltatási formákhoz a legtöbb esetben egy újabb komponens működtetése szükséges a hálózatban, amelyet Policy Servernek nevezünk. A policy szerver együttműködve az autentikációt, autorizációt és accountingot végző RADIUS szerverrel képes olyan funkciókat biztosítani, hogy az említett szolgáltatás példák lehetővé váljanak.

A technológiai fejlesztők más területeken is komoly eredményeket mutattak fel ebben a tekintetben, jó példa erre a RADIUS CoA (Change of Authorization) kiter-

jesztés, melynek segítségével egy adott IP vagy akár PPP session QoS paraméterei megszakadás nélkül dinamikusan változtathatók.

8. Összefoglalás

Összefoglalva tehát a szolgáltatók napjainkban izgalmas fejlesztési irányvonalakkal foglalkoznak a tömegekhez eljuttatott szélessávú ADSL szolgáltatás fejlesztésével kapcsolatban, ezek megvalósításához mindenképpen szükség van (lesz) a minőségi szolgáltatás (QoS) paraméterek precíz hangolási lehetőségeire, amelyhez a Cisco eszközökön rendelkezésre álló technológiák megfelelő segítséget nyújtanak.

Irodalom

- [1] Cisco Systems:
DSL Aggregation for Wireline Carriers
http://www.cisco.com/en/US/netsol/ns568/networking_solutions_solution.html
- [2] Cisco Systems:
Video/IPTV Solutions for Wireline Carriers –
http://www.cisco.com/en/US/netsol/ns610/networking_solutions_solution_category.html
- [3] Cisco Systems:
Solution for DSL Forum TR-59 architecture, 2004.
- [4] Cisco Systems:
QoS Models for Ethernet-DSL deployments,
Rev.022, 2005-2006.
- [5] DSL Forum TR-059 DSL Evolution –
Architecture Requirements for the Support of
QoS-Enabled IP Services
<http://www.dslforum.org/techwork/tr/TR-059.pdf>
- [6] DSL Forum TR-101 Migration
to Ethernet-Based DSL Aggregation
<http://www.dslforum.org/techwork/tr/TR-101.pdf>
- [7] IETF RFC 2745
An Architecture for Differentiated Services (Diffserv)
<http://tools.ietf.org/html/rfc2475>
- [8] DSL Forum WT-092
Broadband Remote Access Server (BRAS)
Requirements Document,
<http://www.dslforum.org/techwork/tr/TR-092.pdf>

Szakmai elismerések

Folyóiratunk 2006. évi vendégszerkesztői, szerzői, illetve cikkei közül az alábbiak kaptak Pollák-Virág díjat:

Kántor Csaba – a 2006/4. szám vendégszerkesztője

Csillag Kristóf – Dobrowiecki Tadeusz – Istenes Zoltán:

Bevezetés az érvértékpézetbe
(2006/1. szám, pp.23-28.)

Farkasvölgyi Andrea:

Optikai sávú összeköttetések alkalmazása az úrtávközlésben
(2006/2. szám, pp.17-22.)

Takács György – Tihanyi Attila – Bárdi Tamás – Feldhoffer Gergely – Strancsik Bálint:

Beszédjel átalakítása mozgó száj képévé siketek kommunikációjának segítésére
(2006/3. szám, pp.31-37.)

Pintér István:

Beszédjelek pillanatnyi jellemzőinek becslése a Teager-operátorral és a Hilbert-Huang-transzformációval
(2006/8. szám, pp.28-37.)

Ács Gergely – Buttyán Levente:

Útvonalválasztó protokollok vezetékek nélküli szenzorhálózatokban
(2006/12. szám, pp.3-11.)

A díjazottaknak gratulálunk!

DSL hozzáférési hálózatokban alkalmazott csomagütemezők sorbanállási modellezése és analízise

KÓRÖSI ATTILA, SZÉKELY BALÁZS

BME Matematikai Intézet, Sztohasztika Tanszék, {akorosi, szbalazs}@math.bme.hu

LUKOVSKI CSABA, DANG DIHN TRANG

BME Távközlési és Médiainformatikai Tanszék, {lukovszki, trang}@tmit.bme.hu

Lektorált

Kulcsszavak: DSL, sorbanállásos modellek, csomagütemező

Jelen munkában azt tűztük ki célul, hogy leírjuk egy szigorú prioritáson alapuló DSL ütemező teljesítmény analízisét. A kapcsolódó csomag-szintű modell egy véges pufferes prioritásos sorbanállási rendszer elemzéséhez vezet, melyben az igények csoportokban érkeznek, és a pufferek telítődése esetében teljes vagy részleges visszautasítást alkalmaznak.

1. Bevezetés

Napjaink hozzáférési hálózataiban a DSL (Digital Subscriber Line) technológiák által nyújtott elérési fénykorát éli [1,2]. Habár az ezen technológiák által nyújtott sáv-szélesség egyre növekszik és ugyanakkor az elérési is egyre megbízhatóbbá válik, mégis többnyire ez az összeköttetés bizonyul a legszűkebb keresztmetszetnek. Következésképpen ez a hálózati szegmens a legfontosabb a biztosított minőségi szolgáltatások szempontjából.

A szolgáltatásbiztosításhoz vezető legrövidebb út a folyamatok osztályokba sorolásával valósítható meg. Ennek érdekében, ahogy ezt a 3GPP és ITU-T is javasolja, az igényeket négy különböző prioritású osztályba soroljuk a forgalmak QoS követelményeinek megfelelően. A kiszolgálás során a csomagok a különböző osztályokhoz rendelt pufferekben tárolódnak és többnyire az osztályok között szigorú prioritásos ütemezést használnak a csomagok kiszolgálására napjaink eszköz szintű megvalósításaiban [3].

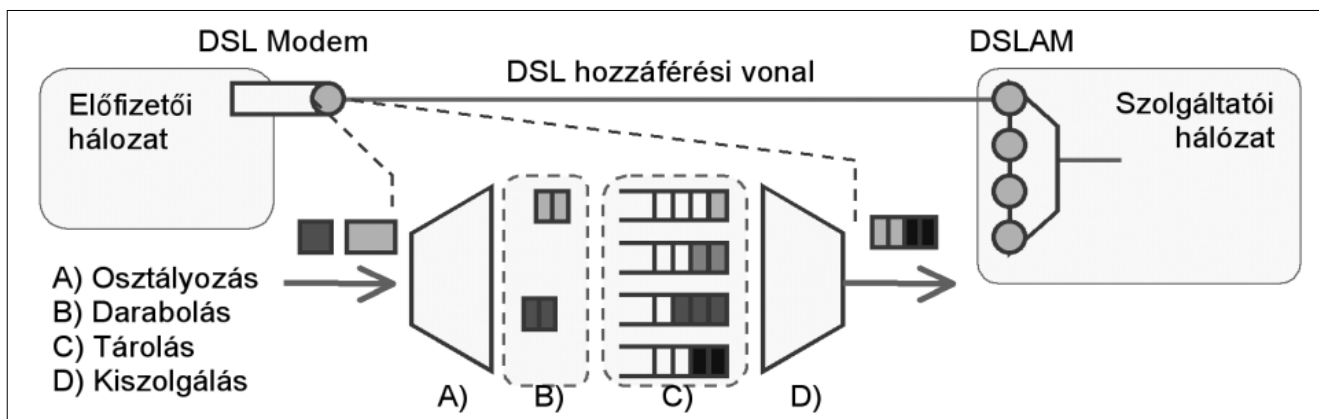
2. DSL-alapú hozzáférési hálózat

A DSL felhasználók DSL modemeket használva kapcsolódnak a hozzáférési hálózatba, melyek a DSLAM (DSL Access Multiplexer) eszközben aggregálódnak, ahogy ezt az 1. ábra is mutatja.

A DSL legújabb szabványai [1,2] két lehetőséget kínálnak a magasabb szintű csomagok továbbítására a DSL vonalakon. Az ATM (Asynchronous Transfer Mode) alapú megoldással a DSL végpontok kisebb cellákba szegmentálják az átküldendő magasabb szintű IP vagy Ethernet csomagokat.

Ezen cellák 53 bájtosak, melyből 5 bájt a fejléc, és további 48 bájt a tiszta adat. Ez úgy valósul meg, hogy a SAR (Segmentation and Reassembly) alrétteg a már AAL5 (ATM Adaptation Layer 5)-ös keretbe ágyazott magasabb szintű csomagokat ATM cellákba darabolja, még azelőtt hogy az adatokat átküldené a fizikai vonalon. Az ATM cellái kicsik a többi hasonló rendeltetésű rendszerhez képest és pontosan ezért képes egyidejűleg akár hangot, videót és egyéb dokumentumot is hatékonyan átvinni. Mindezek mellett az ATM alapú átvitelben használt PVC (Permanent Virtual Circuit) megoldás alkalmazása biztosít arról, hogy az átjutó adatfolyamból szétválaszthatók legyenek az eredeti adatok. A módszer rossz hatékonysága a fejlécek méretéből fakad, ugyanis egy, az IP kommunikáció során gyakran használt 64 bájtos Ethernet-keretet csak két részletben lehet átvinni, az egyik 48 bájt + 5 bájt hosszúságú, a másik fele 16 bájt + 5 bájt. Mivel az AAL5 szegmentá-

1. ábra
DSL előfizetői vonalak aggregációja és csomagszintű modellje napjaink hozzáférési hálózataiban



lás miatt a második cella megmaradó részét sem lehet másik csomaggal feltölteni, a 64 bájt átviteléhez összesen 106 bájt átvitele szükséges, ami megközelítőleg 60%-os kihasználtságot eredményez erre a csomagméretre.

Ezzel szemben az Ethernet alapú DSL technológiák egyik legfontosabb újítása a 64/65-ös beágyazás, vagyis 64 bájt információ mellé csak egy bájt fejléc kerül. IEEE 802.3ah egy speciális Ethernet PTM-TC (Packet Transmission Mode – Transmission Convergence) megoldást definiál, mely 64/65-oktett darabolást végez a magasabb szintű csomagokon. A fenti 64 bájt átviteléhez ebben az opcióban csak 65 bájt adatot kell átvinni a fizikai vonalon. Mivel a PTM-TC lényegesen hatékonyabb és a modern IP alapú hozzáférési hálózatokkal való együttműködése is könnyebben kezelhető, minden valószínűség szerint, a VDSL2 elterjedésével az Ethernet-alapú átvitel fog elterjedni az ATM ellenében. Mindkét esetre alkalmazható a modell, miszerint a magasabb rétegekből érkező csomag kisebb adategységekre darabolódnak és ezek a szegmensek a QoS paramétereiknek megfelelő pufferben tárolódnak, miként az az első ábrán is látható. Ez a darabolási eljárás megoldja azt a problémát, mely során a magasabb prioritású csomagoknak meg kell várniuk a nagyméretű alacsonyabb prioritású csomagok teljes kiszolgálódását, ami jelentős késleltetést jelent a magasabb prioritású csomagok átvitelében.

3. Prioritásos sorbanállási rendszerek elmélete

A fent bemutatott ütemező modellezhető egy szigorú prioritást használó sorbanállási rendszerrel. Habár a prioritásos sorok tanulmányozása még mindig népszerű része a sorbanállási rendszereknek, a véges pufferekkel és borszt érkezéssel működő rendszerek pontos leírására eddig nem született eredmény.

A kiszolgálási rendszerek témakörében született monográfiák, például Kleinrock [4] és Tejms [5] nem érintik ezen dolgozatban tárgyalt prioritással rendelkező rendszereket. Daigle [6] 2005-ben megjelent monográfiájában végtelen pufferes, prioritásos rendszerek esetén meghatározta a várakozó csomagok számának, és a várakozási időnek a generátor függvényét. Továbbá ezen mennyiségek várható értékére vezetett le formulákat.

Végtelen pufferek feltételezése sok esetben egyszerűsítést jelent, mert a matematikai eszközök hatékonyabbak ilyen struktúrákon, így tömörebb, sok esetben zárt formulákat kaphatunk eredményül. Mivel a felhasználását tekintve általában nagy pufferek fordulnak elő, ezért jól lehet közelíteni a véges pufferes értékeket a végtelen pufferes megfelelőjéből kiszámolt értékekkel. Takács Lajos [7] korai munkájában, amelyben Poisson beérkezéssel számolt, szükséges és elégséges feltételt adott a várakozási idők stacionárius határeloszlására, majd ezen határeloszlásnak a Laplace-transzfor-

máltjának a segítségével kiszámolta az első három momentumát. 1972-ben Hooke [8] két csatorna esetén az alacsonyabb prioritású osztály várakozási idejére és a teljes rendszerben várakozó csomagok számára bizonyított határeloszlás tételt Laplace-Stieltjes transzformált segítségével. 1999-ben T. Takine [9] munkájában a beérkezések Markov-folyamatot, a kiszolgálási idők általános eloszlást követnek. Generátor függvény módszer és mátrix analízis segítségével meghatározta a sor hosszúságok marginális eloszlását, illetve Laplace-Stieltjes transzformált segítségével elemezte a várakozási idők eloszlását.

2005-ben Xue és Alfa [10] munkájukban szintén Markov-folyamat szerinti beérkezést feltételeztek, továbbá a legfelső csatornába borsztökből érkeztek a csomagok. Eredményeik két csatornás rendszerekre vonatkoztak. Elégséges feltételt adtak arra nézve, hogy a várakozó csomagok száma aszimptotikusan geometriai eloszlás legyen. Ezen feltétel mellett hatékony módszerrel adtak a veszteségi valószínűségek meghatározására. Alfa, Liu és He [11] mátrix geometria módszerrel adtak megoldást egyszerű, markovi beérkezés esetén a rendszer stacionárius eloszlására.

A számítási igények lerövidítése a célja Van der Heijden, van Harten és Sleptchenko [12] munkájának. Írárukban megmutatnak egy módszert, amelyben egy N soros rendszer vizsgálatát visszavezetik N darab két soros rendszer vizsgálatára. Lényegében a második sorba aggregálják az összes többi, és azt közelítik hiperexponenciális eloszlással.

A véges pufferes rendszerrel foglalkozott Sharma és Virtamo [13]. Két sort, Poisson érkezést, folytonos borszt méret eloszlást és tetszőleges kiszolgálást feltételezve. Algoritmust mutattak stacionárius eloszlás kiszámolására, majd határeloszlás tételekre alapozva meghatározták a várakozási idő eloszlását és a csomagvesztések valószínűségeit.

4. A sorbanállási modell

A fentiekben leírt, DSL hálózatok csomagütemezési és kiszolgálási modellje visszavezethető egy sorbanállási rendszer analízisére. Tekintsünk egy kiszolgálós prioritásos sorbanállási rendszert állandó kiszolgálási sebességgel (V bit/s). Egy ilyen rendszerben az ütemezés miatt csak akkor választ kiszolgálandó igényt alacsonyabb pufferből, ha már az összes magasabb prioritású tároló üres. A rendszer nem-megszakításos elvvel (non-preemptive principle, NPRP) működik, ami azt jelenti, hogy egy kiszolgálás alatt álló igény kiszolgálása nem megszakítható.

Feltételezzük, hogy adott rendszernek I darab prioritásos csatornája van, ahol a kisebb indexű csatorna nagyobb prioritást jelöl. Minden egyes csatorna véges pufferrel rendelkezik, amelyeknek hossza rendre $b^{(i)}$, $i = 1, 2, \dots, I$. A pufferekben levő csomagok FCFS (First Come, First Served) elvvel szolgálódnak ki. Az igények a rendszerhez borsztökből érkeznek, amelyek a beérke-

ző csomagok szegmenseit reprezentálják. A borsztők beérkezése Poisson-eloszlás szerint történik $\lambda^{(i)}$, $i = 1, 2, \dots, I$ intenzitással. A borsztők mérete diszkrét $X^{(i)}$ eloszlást követ. Modellünkben ez az eloszlás csatornánként különböző lehet. Egy borszt tetszőleges számú szegmenst tartalmazhat, amelyeknek mérete állandó (L bit). A rendszer teljes vagy részleges visszautasítást alkalmaz, azaz ha egy érkező borszt nem fér a megfelelő pufferbe, akkor az egész borszt eldobódik (teljes visszautasítás), illetve a rendszer feltölti szegmensekkel a puffert és csak a maradék dobódik el (részleges visszautasítás). A rendszer stabilitásának feltétele, hogy $\sum L \cdot \mathbf{E}X^{(i)} \cdot \lambda^{(i)} \leq V$ korlát teljesüljön.

A fentiekben leírt rendszer-modell illusztrációja a 2. ábrán látható.

5. A modell matematikai feldolgozása

Mivel a teljes rendszer leírása kezelhetetlenül nagy állapotteret eredményezne, első lépésként a bemutatott állandó sebességű I sorból álló sorbanállási rendszert visszavezetjük I darab $M/G/1/b$ sorra. Ebben az esetben a korábbi fix (L/V) kiszolgálási idő helyett általános kiszolgálási idő eloszlást feltételezünk, mely kiszolgálási idő az adott sorra van vonatkoztatva. Ehhez figyeljük meg a rendszer működését az i -edik sor szemszögéből. A borsztők L -méretű szegmensekből épülnek fel, és $\lambda^{(i)}$ paraméterű Poisson-folyamat szerint érkeznek.

A létrehozott I darab új sorban ezek után a kiszolgálási idő nem egy szegmens kiszolgálási idejét jelenti, hanem az adott sorra azt az időt, amely egy i -dik sorból választott igény kiszolgálásának megkezdése és a következő olyan időpillanat között telik el, amikor a kiszolgáló újra kész az i -edik sorból származó igény kiszolgálására. Ez az idő magában foglalja azokat az időket is, amelyek azzal telnek el, hogy az adott sor két szegmensének a kiszolgálása között a rendszer kiszolgálja az időközben érkezett magasabb prioritású szegmenseket. Ezt a kiszolgálási időt a továbbiakban jelöljük $S^{(i)}$ -vel. A rendszer megoldásához rekurziót használtunk az $S^{(i)}$ idők eloszlására, mely az előző sor kiszolgálási idején és foglaltsági idején ($T^{(i)}$) alapszik:

$$S^{(i+1)} = S^{(i)} + T^{(i)} [N^{(i)}(S^{(i)})], \quad i = 1, 2, \dots, I.$$

A továbbiakban a rendszer teljes leírásához szükségünk van még az üres sorba érkező borsztők első szegmensének a kiszolgálási idejére (S^*). Ez a valószínűségi változó függ a többi sor állapotától, ennek a kezelése egy, a fentihez hasonló, de bonyolultabb rekurzióval megoldható.

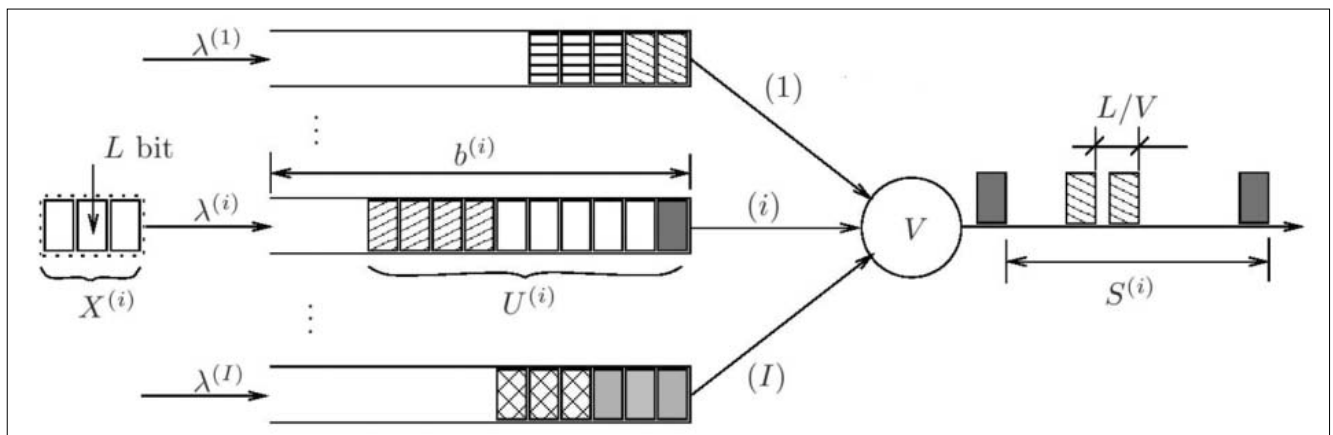
Következő lépésként meghatározzuk egy speciális első kiszolgálási idővel ellátott $M/G/1/b$ sor hosszú távú átlagos eloszlását. Annak a valószínűségét, hogy j szegmens van a sorban jelöljük p_j -vel, és a j szegmenshez tartozó állapotban töltött időhányad határértékével definiáljuk. Ennek a valószínűségnek a meghatározásához a felújítási folyamatok elméletét alkalmazzuk, pontosabban felosztjuk az egész idő intervallumot az adott sor kiszolgálási időire és megvizsgáljuk, egy kiszolgálási idő alatt mennyi időt tölt a rendszer j szegmenssel a sorban. Ennek a várható értékét annak a függvényében számoljuk ki, hogy a kiszolgálás legelején hány szegmens volt a sorban. Jelöljük $t_{i,j}$ -vel azt a várható értéket, amely azt fejezi ki, hogy amennyiben i szegmens volt a kiszolgálás megkezdésekor a sorban, akkor várhatóan mennyi időt tölt a kiszolgálás alatt j állapotban.

Ahhoz, hogy megkapjuk p_j értékét, $t_{i,j}$ értékeket súlyoznunk kell annak a valószínűségével, hogy a kiszolgálás i szegmenssel kezdődött. Ennek megoldásához felírjuk azt a Markov-láncot, amelynek az állapotai a kiszolgálások elején a sorban lévő szegmensek száma. Ehhez szükséges a Markov-lánc állapotátmeneti mátrixa, ami könnyen számolható, ezek után egyszerűen adódik a stacionárius eloszlása ($\mathbf{d}$), ami pont azt fejezi ki, hogy milyen valószínűséggel kezd i szegmenssel egy kiszolgálási idő. A fenti gondolatmenet a következő képletet adja:

$$p_j = \frac{d_0 t_{0,j} + d_1 t_{1,j} + L d_b t_{b,j} + d_{x_1} t_{x_1,j} + L + d_{x_l} t_{x_l,j}}{d_0 \frac{1}{\lambda} + (d_1 + L d_b) \mathbf{E}S + (d_{x_1} + L + d_{x_l}) \mathbf{E}S^*}$$

A $t_{i,j}$ -k számolásának az alapötlete az, hogy ha már a pufferben van j szegmens, akkor ez csak kétféleképpen változhat meg, vagy érkezik egy új borszt, amiből kerül szegmens a pufferbe, vagy véget ér a kiszolgálás és eggyel csökken a szegmensek száma a pufferben.

2. ábra A sorbanállási modell



Új borszt beérkezéséig exponenciális ideig kell várunk, továbbá a kiszolgálás végéig hátralévő idő eloszlását ki lehet számolni i és j függvényében. $t_{i,j}$ ennek a két mennyiség minimumának a várható értéke.

A várakozási időre a fenti eredményeket figyelembe véve könnyen adódik az alábbi összefüggés:

$$W = R + \sum_{k=1}^U S_k + \sum_{k=U+1}^{U+X-1} S_k + \frac{L}{V}.$$

Ennek jelentése az, hogy a borszt X szegmenssel érkezik a sorba, ahol már U szegmens várakozik. Ekkor a borszt teljes kiszolgálásáig meg kell várni az aktuális kiszolgálásból még hátralévő időt (R), a sorban már korábban is várakozó U szegmens kiszolgálását, a saját szegmensei közül az első $X-1$ teljes kiszolgálási idejét és végül az utolsó szegmens kiszolgálását, amely L/V . Mivel ezek az idők függetlenek, egy egyszerű véletlen tagszámú konvolúciót kell alkalmazni, ahol a tagszám eloszlása ismert. Megjegyezzük, hogy az R eloszlásának a számolása nagyon hasonló a $t_{i,j}$ számolásához, azonban R azt fejezi ki, hogy ha j szegmens van a pufferben és előbb érkezik borszt, mint vége lenne a kiszolgálásnak, akkor mennyi idő marad hátra a kiszolgálásból a borszt érkezése után.

A visszautasítási valószínűség kiszámolás egyszerű a sorhossz eloszlás ismeretében:

$$P_{rej} = \sum P(U = j) P(X > b - j)$$

A két különböző visszautasítási szabály alkalmazása az algoritmusokban két fontos különbséget eredményez. Az első abban a mátrixban van, amely egy borszt beérkezése esetén a sorhossz változás valószínűségeit mutatja. Ezt a mátrixot felhasználjuk a kiszolgálási idő eloszlásának és a fent említett Markov-lánc stacionárius eloszlásának a meghatározásához is. Emellett a másik különbség a $t_{i,j}$ és R számolásával kapcsolatos, ugyanis mindkét mennyiség meghatározásában szere-

pet játszik egy exponenciális idő, amely azt mondja meg hogy ha j szegmens van a sorban, akkor mennyi idő múlva érkezik a pufferbe új szegmens. A különbség abban jelentkezik, hogy teljes visszautasítást alkalmazva új borszt érkezés esetén csak akkor kerül szegmens a sorba, ha a borszt teljes egészében befér, tehát ebben az esetben a kérdéses exponenciális idő paramétere nagy j esetén eltér a sorhoz tartozó paramétértől. Azonban részleges visszautasítás esetén, amennyiben nincs tele a puffer és érkezik borszt, akkor kerül új szegmens a sorba. Ezért ebben az esetben, ha nincs tele a puffer, akkor a beérkezés intenzitása minden állapotra meg egyeznek a sorhoz tartozó beérkezési intenzitással.

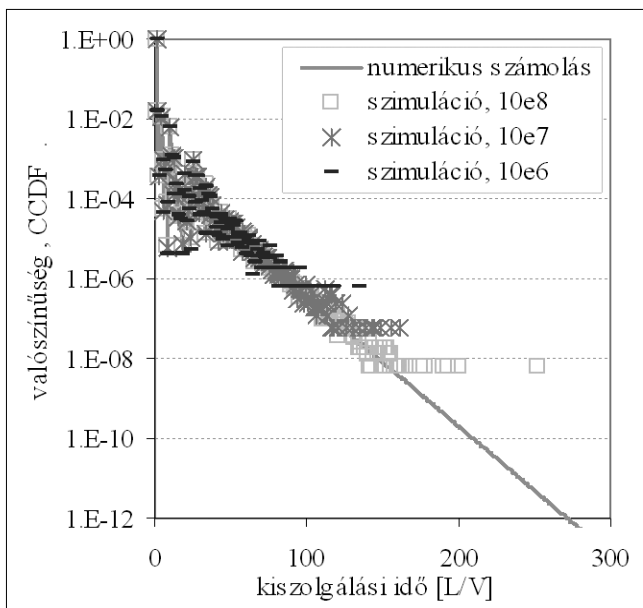
6. Numerikus eredmények

Az előző fejezetben vázolt numerikus algoritmust implementáltuk C nyelvben, és hogy igazoljuk az analízisünk helyességét, valamint pontosságát szimulációs eredményekkel hasonlítottuk össze. A rendszer paramétereit valósághoz közelinek választottuk, a rendszer bemeneti forgalmát a kimeneti link kapacitáshoz mérten 50%, 70% és 90%-ra választottuk, a négy csatorna között felosztva 4%, 12%, 24% és 60% arányban. A csomagméreteket az első sorban hang forgalomnak megfelelően, a többiben internet forgalmat empirikusan leíró IMIX modell alapján választottuk. Továbbá fontos megjegyezni, hogy a numerikus analízis során a végtelen összegzések helyett az eloszlásokat 10^{-6} hibahatárral számoltuk.

A 3. ábrán azt láthatjuk, ahogy a szimulációs eredmények egyre jobban közelítik a numerikus eredményeket. Az eredmény a negyedik csatorna kiszolgálási idejének komplementer eloszlása van ábrázolva 90%-os terheltség mellett. A numerikus eredmények mellett a szimulációs eredményeket is ábrázoltuk 10^6 , 10^7 , illetve 10^8 csomag beérkezésére.

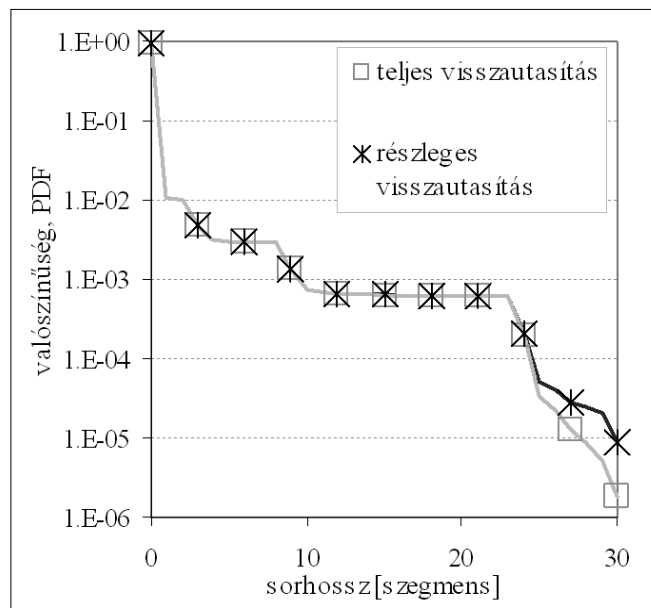
3. ábra

A kiszolgálási idő numerikus és szimulációs eredményei



4. ábra

Sorhossz eloszlás teljes és részleges visszautasítás esetén



Összehasonlításként elmondhatjuk, hogy pontosságban a numerikus eredmények messze felülmúlják a szimulációs eredményeket, hiszen az ábrázolt tartományban a numerikus eredmények 10^{12} csomag beérkezésből számított eloszlás pontosságának felelnek meg.

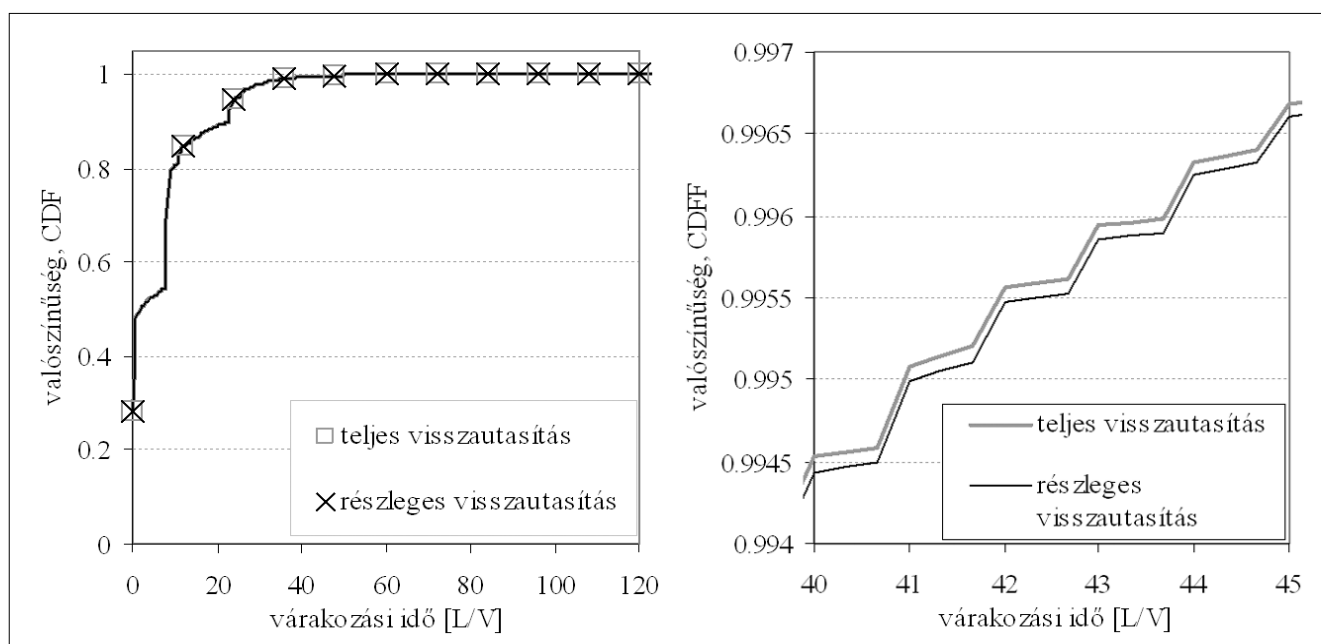
A 4. ábrán a teljes és részleges visszautasítás közötti különbség látható a sorhossz valószínűségeket ábrázolva a második csatornára 50%-os terhelés mellett. Egyértelműen látható a különbség a két eloszlás közelében a puffer telítettség határán, ugyanis részleges visszautasítás esetében a maradék szegmensek bekerülnek a pufferbe ezzel növelve a valószínűséget a puffer méret közeli tartományban.

Az 5. és 6. ábrán a várakozási idő eloszlása látható mind teljes mind részleges visszautasítás esetén. Az 5. ábrán nem látható közöttük nagy különbség, de ha közelebről megvizsgáljuk, mint ahogy az a 6. ábrán is látszik, a teljes visszautasítás eloszlás-függvénye a részleges fölött van.

Irodalom

- [1] Asymmetric digital subscriber line (ADSL) transceivers – extended bandwidth adsl2 (ADSL2+), ITU-T, Recommendation G.992.5., 2005.
- [2] Very high speed digital subscriber line transceivers 2 (VDSL2), ITU-T, Recommendation G.993.2., 2006.
- [3] T. Orphanoudakis, S. Perissakis, K. Pramataris, N. Nikolaou, N. Zervos, M. Steck, C. Baumhof, D. Verkest, C. Ykman-Couvreux, G. Doumenis, F. Karoubalis, I. Theologitou, D. Reisis, G. Konstantoulakis, N. Vogiatzis, Hardware Architectures for the Efficient Implementation of Multi-Service Broadband Access and Multimedia Home Networks, Telecommunication Systems, Springer Netherlands, 23(3-4), pp.351–367.
- [4] Kleinrock, L., Queueing Systems, Vol. II, Computer Applications, Wiley, New York, 1976.
- [5] Tijms, H. C., A First Course in Stochastic Models, John Wiley & Sons, 2003.
- [6] Daigle, J. N., Queueing Theory with Applications to Packet Telecommunication, Springer, 2005.
- [7] Takács, L., Priority queues, Operations Res. 12, pp.63–74., 1964.
- [8] Hooke, J. A., Some heavy-traffic limit theorems for a priority queue with general arrivals, Operations Res. 20, pp.381–388., 1972.
- [9] Takine, T., Hasegawa, T., The nonpreemptive priority MAP/G/1 queue, Operations Research 47(6), pp.917–927., 1999.
- [10] Xue, J., Alfa, A. S., Tail probability of low-priority queue length in a discrete-time priority BMAP/PH/1 queue, Stoch. Models 21(2-3), pp.799–820., 2005.
- [11] Alfa, A., Liu, B. & HE, Q., Discrete-time analysis of MAP/PH/1 multiclass general preemptive priority queue, Naval Research Logistics 50, pp.662–682., 2003.
- [12] van der Heijden, M., van Harten, A., Sleptchenko, A., Approximations for Markovian multi-class queues with preemptive priorities, Oper. Res. Lett. 32(3), pp.273–282., 2004.
- [13] Sharma, V, Virtamo, J. T., A finite buffer queue with priorities, Performance Evaluation 47, pp.1–22., 2002.

5. és 6. ábra Várakozási idő eloszlása teljes és részleges visszautasítás esetén



Átjáró szerver választása a GMPLS PCE-architektúrában

ZAHEMSZKY ANDRÁS,

BME Távközlési és Telematikai Tanszék, za488@hszk.bme.hu

CSÁSZÁR ANDRÁS, TÓTH GÁBOR, TAKÁCS ATTILA

Ericsson Magyarország Kft.

{andras.csaszar, gabor.toth, attila.takacs}@ericsson.com

Lektorált

Kulcsszavak: GMPLS, átjárószerver, útvonalszámító eszköz

Ebben a cikkben bemutatjuk, hogy a GMPLS új útvonalválasztó architektúrája, az Útvonal Számító Eszközre (Path Computation Element – PCE) épülő modell hogyan alkalmazható átjáró szerverek választására. A választás célja egyrészt az, hogy elkerüljük a szerverek túlterheléséből adódó hívás blokkolást, másrészt az, hogy minimalizáljuk a transzport hálózat összerhelését, és ezekkel együttesen elérjük, hogy a felhasználók által érzékelt szolgáltatási minőség ne romoljon. Olyan algoritmusokat mutatunk be, amelyek csak egyik vagy mindkét szempontot (a szerverek terheltsége illetve a hálózati topológia) figyelembe veszik, amikor a felhasználó kérésére átjárót választanak. Hálózatszimuláció segítségével összehasonlítjuk a különböző algoritmusokat és megmutatjuk, hogy alacsony blokkolási arány, illetve megfelelően alacsony hálózati terhelés érhető el azokkal, amelyek mindkét szempontot figyelembe veszik a döntés során.

1. Bevezetés

A mai számítógép-hálózatokban gyakran osztunk szét valamilyen képességet több entitás között. Ilyenkor minden entitás ugyanolyan képességekkel rendelkezik, a felhasználó számára mindegyik pontosan ugyanazt a szolgáltatást tudja nyújtani. Az erőforrások ilyenforma szétosztásával a rendszer skálázhatóságát növeljük, mert elkerüljük, hogy az erőforrások a hálózat szűk keresztmetszetévé váljanak. Mivel több szerver is rendelkezésre áll, biztosítani kell egy választási mechanizmust, mely az adott igényhez egy erőforrást rendel a csoportból.

A választási algoritmusnak több célja is lehet. Egyrészt fontos a szerverek terhelését megfelelően szétosztani, annak érdekében, hogy egyik szerver se legyen még ideiglenesen sem túlterhelt. A túlterhelés ugyanis oda vezethet, hogy a szolgáltatás egy bizonyos ideig a felhasználók egy része számára nem lesz elérhető, illetve bizonyos esetekben a felhasználói adatforgalom is megszakad, így a felhasználók által érzett minőség csökken. Másrészt fontos szempont, hogy a választott szerver a felhasználótól rövid úton elérhető legyen. A rövid út a válaszidő, illetve a késleltetés csökkenését jelentheti a felhasználónak, illetve a torlódás kialakulásának valószínűségét is csökkenti, hiszen kevesebb kapcsolat terhelését növeli egy új adatfolyam. Nem mellékes mellékhatás az sem, hogy a hálózat kihasználtsága is javul.

Hálózati terheléelosztással több szituációban is találkozhatunk. Legkézenfekvőbb alkalmazása a Web [1]. Ekkor több Web szerver áll rendelkezésre, amelyek mind ugyanazt a tartalmat tárolják. Előfordulhat olyan eset, amikor egy weblap látogatottsága hirtelen megnő, s ilyenkor a terhelés megfelelő elosztása nagyon fontossá válik. Cél, hogy a szerverek terheltségét egyenletesre állítsuk be. Erre egy úgynevezett hálózat-alapú

megoldási lehetőség (azaz amikor nem a felhasználó, hanem a hálózat választ) a DNS-szervereken alapuló megoldás. Ennek lényege, hogy a DNS-szerver az IP-cím feloldásakor mindig más és más szervert választhat. Azonban mivel lokálisan is tárolhatjuk a név-IP-cím összerendeléseket, illetve a közbülső DNS-szerverek is megteszik ezt, nem biztos, hogy a kérés minden esetben eljut a megfelelő helyre. A DNS-szerver tipikusan egy egyszerű algoritmus alapján választ szervert, például *körülfordulósos (Round-Robin) stratégiát* alkalmaz.

Hálózati terheléelosztásról beszélhetünk akkor is, amikor egy ad hoc hálózatban a forgalmazni kívánó csomópontok számára átjárókat rendelünk az Internet felé. Az erre vonatkozó kutatások arra jutottak, hogy nem elég csupán a közelségi információkat figyelembe venni, hanem a szerverek terhelésére vonatkozó információk is szükségesek a jó döntéshez [2].

A terhelés minél egyenletesebb elosztására használható az IPv6-ban megtalálható *anycast* címezés is. Ekkor több szervernek is ugyanaz az anycast címe, ami azt fejezi ki, hogy egy csoportba tartoznak. A címezés lényege, hogy az anycast címre küldött csomagot a hálózat legalább egyik tagja meg fogja kapni, de azt a hálózat dönti el, hogy pontosan melyik. A megoldások java része a kérést a legközelebbi szerverhez irányítja, egy másik megoldás viszont a terhelésekre vonatkozó információkat használja fel az aktív útválasztó technológiát használva, azonban itt a közelségi információk figyelembevétele hiányzik [3].

Cikkünkben bemutatjuk, hogy a GMPLS (*Generalized Multiprotocol Label Switching*) PCE-alapú, új útvonalszámítási modellje hogyan használható hálózati terheléelosztásra és ezzel a szolgáltatás minőségének javítására. Az általunk vizsgált hálózatban a külső hálózatot átjáró szerverek egy csoportján keresztül érhetjük el. A felhasználói adatforgalom megkezdésekor a PCE fog átjárót rendelni a végberendezéshez. Sajnos

a szerverek túlterhelésének kivédése és mindig a legközelebbi szerver választása két, egymásnak erősen ellentmondó szempont. Egyszerű választási algoritmusokat javasolunk, amelyek mindkét szempontot figyelembe veszik a döntéskor, és ezeket összehasonlítjuk azokkal, amelyek csak egy szempont alapján döntenek.

2. A GMPLS és a PCE-architektúra

A GMPLS protokollcsalád egy széles körben használt, általános kontrollmegoldás [4]. Az MPLS általánosításaként azon alapszik, hogy tartalmazza az alagutazás (tunneling) koncepcióját: a felhasználói adatforgalom megkezdése előtt a forrás és a nyelő között egy úgynevezett *címkekapcsolt útvonalat* (*Label Switched Path – LSP*) hozunk létre. Az MPLS-ben a felhasználók csomagjait egy plusz fejrészben található címkék azonosítják, és a hálózaton belül a csomagok továbbítása a (címke, bemeneti interfész) pár alapján történik. A GMPLS az MPLS-sel szemben abban igazán új, hogy már nemcsak csomagkapcsolt hálózat felett képes működni, hanem más, például hullámhosszkapcsolt hálózatok felett is, és segítségével különböző, egymástól lényegesen eltérő hálózati technikákat lehet egységesen vezérelni. A GMPLS-ben már általánosított címkéről beszélünk, ami egy virtuális címke, amit például egyes időrések, hullámhosszok azonosíthatnak, és természetesen a korábbi MPLS-címke is GMPLS-címke.

A GMPLS protokollcsalád három protokollt tartalmaz. A jelzési protokoll feladata címkekapcsolt útvonalak létrehozása, módosítása, illetve törlése felhasználói igény felmerülése esetén. Elterjedt jelzési protokoll az *RSVP-TE* (*Resource Reservation Protocol – Traffic Engineering*). A routing protokoll feladata útvonalválasztási információk terjesztése a GMPLS-hálózatban. A GMPLS gyakran használt routing protokollja az *OSPF-TE* (*Open Shortest Path First – Traffic Engineering*). Az OSPF-TE *opaque LSA* (*Link State Advertisement*) [5] üzenete teszi lehetővé, hogy különböző forgalom menedzsmenthez szükséges információkat is elterjeszthessünk a hálózatban (például sáv szélesség vagy képesség-információk). *Opaque LSA* használatával más alkalmazások is elterjeszthetik információikat a hálózatban. Ezekon kívül a GMPLS még tartalmaz egy menedzsment protokollt is (ez tipikusan az *LMP – Link Management Protocol*), aminek feladata vezérlő csatornák létrehozása szomszédos vezérlő entitások között, hibák észlelése és lokalizálása.

LSP kiépítése előtt a jelzési folyamatot megkezdő csomópont általában a jelzési üzenetbe ágyazza a cél-csomópont felé haladó út csomópontjainak listáját, azaz forrás-vezérelt útvonalválasztásról beszélünk. Gyakran előfordul azonban olyan eset, amikor a teljes útvonal nem áll helyben rendelkezésre, tipikusan olyan esetben, amikor a cél csomópont a hálózat egy másik tartományában helyezkedik el, hiszen a csomópont csak a saját tartományáról rendelkezik teljes képpel. Ilyenkor egy lehetőség, hogy a forrás csomópont egy távoli en-

titáshoz fordul, amely képes – szükség szerint további entitásokkal együttműködve – útvonalat számolni a forrás és a cél-csomópont között. A távoli entitás neve PCE (*Path Computation Element*). Egy tartományú esetben is van értelme PCE használatának: a csomópont feltételezheti, hogy a PCE további szempontokat is figyelembe tud venni a döntéskor, és így valamilyen értelemben jobb utat tud számolni.

A PCE definíciója nagyon általános: a PCE egy olyan entitás, amely képes egy, a hálózatot reprezentáló gráfon útvonalak kiszámítására, és eközben képes a megfelelő kényszerek figyelembevételére is [6]. Azt a csomópontot, amely útvonalszámítási kéréssel fordul a PCE-hez, *PCC-nek* (*Path Computation Client*) nevezzük. A PCC, illetve a PCE közötti kommunikáció egy egyszerű, kliens-szerver protokoll segítségével történik: ideális esetben a PCC útvonalszámítási kérelmét tartalmazó üzenetére, mely a cél-csomópont címét és esetleges kényszereket tartalmaz, a PCE olyan üzenettel válaszol, mely tartalmazza a részleges, vagy teljes útvonal-listát a cél-csomópont felé.

A PCE-k számát, elhelyezkedését illetve együttműködését illetően a PCE-architektúrának több modellje létezik. Elhelyezkedés szempontjából a PCE lehet az útválasztó egy komponense, ilyenkor a PCC, illetve a PCE ugyanabban a hálózati csomópontban helyezkedik el. Másrészt a PCE lehet egy távoli hálózati komponens. Ez a megkülönböztetés nem éles, hiszen előfordulhat, hogy egy csomópont egy másik csomópont számára PCE-ként funkcionál, azonban bizonyos más helyzetekben kliensként fog működni.

Nem szükséges azonban, hogy egy útvonalat egyetlen PCE számoljon ki. Ilyenkor több PCE-általi útvonalszámításról beszélünk. Itt is két esetet különböztünk meg aszerint, hogy van-e PCE-PCE kommunikáció, vagy nincs.

A modellek elosztottság szempontjából is csoportosíthatók. A centralizált modellben egy tartományon belül az útvonalszámítási felelősség egy csomópontté, míg az elosztott modellben egy tartomány több, egyszerre működő PCE-t is tartalmaz.

Vizsgálatainkban a centralizált modellre szorítkozunk, és a PCE minden csomópont számára távoli hálózati komponens. Az általunk használt modell az 1. ábrán látható.

3. Átjáró-választási algoritmusok

Ez a fejezet a PCE által használt egyszerű átjáró-választási algoritmusokat mutatja be. Míg az átjáró-választási stratégia minden esetben más és más, az útvonalválasztás mechanizmusa mindig ugyanaz: a PCE a PCC és az átjáró közötti legkisebb lépésszámú utat fogja választani. A bemutatott algoritmusok két csoportba oszthatók. Az egyik csoportba azok az algoritmusok tartoznak, amelyek egy fázisban választanak átjárót, míg a másikba azok, amelyek két körben döntenek – azaz egymás után két szempontot is igyekeznek figyelembe venni.

Az egyik legegyszerűbb algoritmus a UNIFORM. Ekor az átjárók között egyenletes eloszlással véletlen választunk, és az egymást követő választások függetlenek egymástól.

A ROUND-ROBIN algoritmus egyszerű körbefordulós stratégia alapján rendel átjárót a felhasználóhoz.

A következő két algoritmus már figyelembe vesz az átjárók elhelyezkedésére, illetve terhelésére vonatkozó információkat. Így a NEAREST algoritmus mindig a legközelebbi átjárót fogja a PCC-hez rendelni. Ez lehetséges, hiszen a tartomány teljes topológiáját a topológialváltozások utáni rövid tranziensek kivételével pontosan ismeri. Amennyiben több átjáró is pontosan ugyanannyi lépésre lenne a PCC-től, akkor a döntő logika egy előre definiált, önkényes sorrendet figyelembe véve oldja fel a holtversenyt. Így vegyük észre, hogy ezen algoritmus, amennyiben nem történik topológia változás, az egyes PCC-k számára minden esetben ugyanazt az átjárót választja.

A GREEDY algoritmus olyan értelemben mohó, hogy mindig a legkisebb terhelésű átjáróhoz irányítja a felhasználót. Azonban itt, nem úgy, mint a topológia-információk figyelembevételekor, már nem garantált, hogy a PCE pontos képpel rendelkezik a hálózatról. A PCE által a szerverek terheltségéről látott képet két dolog befolyásolja döntően: egyrészt az átjárók által opaque LSA-kba ágyazott és rendszeresen küldött terheltségi információk, másrészt az a tény, hogy a PCE a saját döntéseit képes nyomon követni, így az átjárók állapotában történő változásokat bizonyos mértékben figyelembe tudja venni, azonban a felhasználói forgalom befejezéséről nem értesül. Emiatt fontos tényező, hogy az átjárók milyen gyakran frissítik a terhelési információkat. Amennyiben az átjárók gyakran küldenek üzenetet állapotukról, akkor a PCE-k helyes képpel fognak rendelkezni a hálózatról, azonban ennek ára sok vezérlő üzenet, ami a skálázhatóság rovására megy. Ezzel elmentében, hosszú frissítési periódus esetén a PCE-k által látott kép elavulttá válhat, ekkor viszont sokkal kevesebb vezérlő üzenetre van szükség.

A következő két algoritmus két szempontot is figyelembe vesz. A LOADNEAR(d) az első körben a d darab legkevesbé terhelt átjárót engedi át a rostán, majd ezek közül a PCC-hez legközelebbit választja. Észrevehető, hogy az algoritmus $d=1$ paraméter esetén éppen úgy működik, mint a GREEDY, míg ha az első körben egyetlen átjárót sem szűrünk ki ($d=\infty$), a NEAREST algoritmushoz jutunk.

A NEARLOAD(d) ettől csak a szempontok figyelembevételének sorrendjében különbözik: először a topológia-információk alapján engedünk át d szervert a szűrőn, majd ezek közül a legkisebb terhelésűt választjuk. Természetesen a PCE által látott kép helyességére való fenti megfontolás ennél a két algoritmusnál is érvényes.

4. Szimulációs eredmények ismertetése

A szimulációkhoz az ns-2 hálózati szimulátor alkalmazást [7] bővítettük ki a szükséges funkciókkal.

A vizsgálatokat többszintű, összekapcsolt gyűrűkből álló topológián végeztük. Az első ábrán az egyszerűség kedvéért csak egy kétszintű hálózatot tüntettünk fel, azonban a munka során háromszintű hálózattal dolgoztunk, amely összesen 84 csomópontot tartalmazott. A csomópontok egyben PCC-k is, melyekhez a felhasználók kapcsolódnak. A hálózatban egy központi PCE-t helyeztünk el, átjáró szerverek pedig az első és a második hierarchiaszint gyűrűihez csatlakoznak, azaz számuk összesen 20.

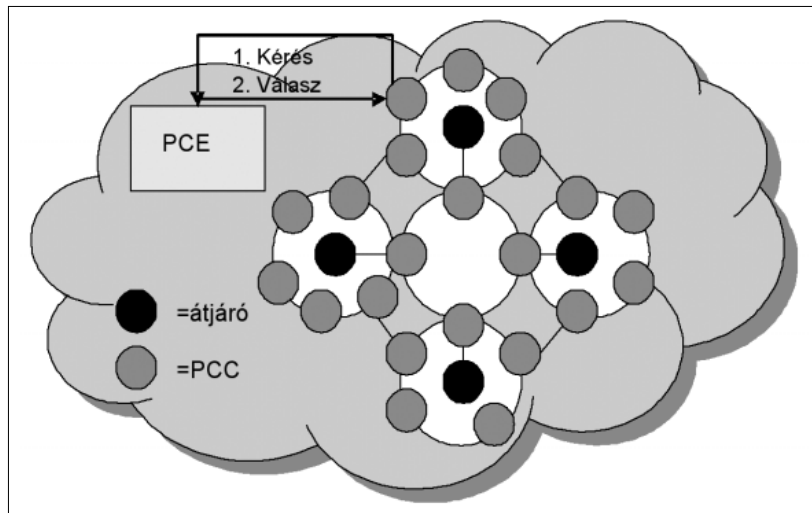
A felhasználói igények érkezési folyamata minden PCC esetében Poisson-folyamat, azonban a köztes idők eloszlása minden PCC-re más és más, választása egy meghatározott intervallumból történik egyenletes eloszlás alapján. Így lesznek olyan csomópontok, amelyek gyakran fordulnak útvonal-választási kérelemmel a PCE-hez, illetve olyanok is, amelyek ritkábban. A hívástartási idő minden felhasználó esetén exponenciális eloszlású, 120 másodperces várható értékkel.

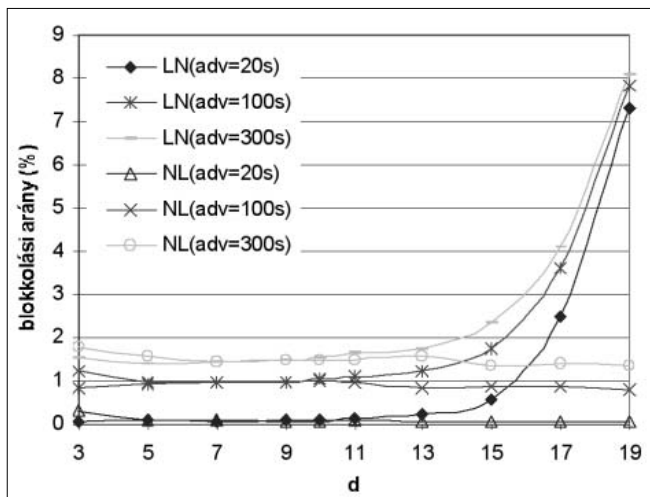
Az alkalmazott modellben az átjárók kapacitása véges. Ez azt jelenti, hogy az általuk egyszerre kiszolgálható felhasználók száma korlátozott. Az egyes átjáró szerverek kapacitását úgy állítottuk be, hogy ideális elosztás esetén az átlagosan felajánlott forgalom 120%-át képesek legyenek elvezetni.

A szimulációk során a blokkolási arányt és az utak átlagos hosszát mértük. Blokkolás akkor léphet fel a rendszerben, ha a PCE olyan átjárót rendel a felhasználóhoz, ami éppen tele van. Az utak átlagos hosszával pedig a transzport hálózat terheltségére, a késleltetésekre és a torlódás valószínűségére következtethetünk.

Először megvizsgáltuk az általunk javasolt két új algoritmus (LOADNEAR, illetve NEARLOAD) viselkedését a d para-

1. ábra A PCE architektúra és egy példa



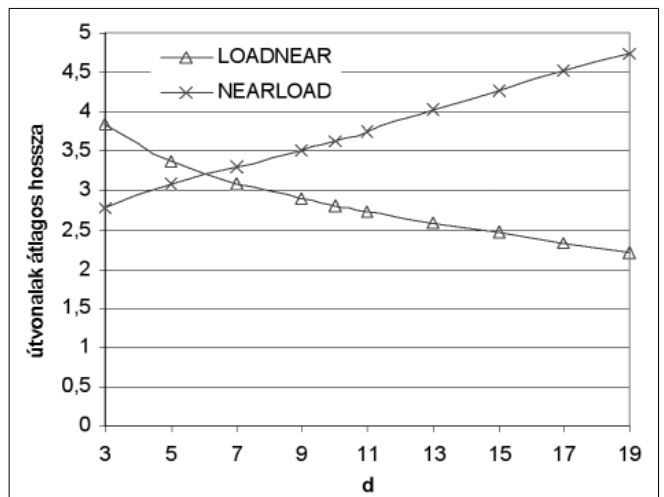


2. és 3. ábra Blokkolási arány és az útvonalak átlagos hossza a d paraméter függvényében

méter, azaz az első fázisban kiválasztott átjárók számának függvényében. Az eredmények a 2. és a 3. ábrán láthatók.

A periodikus frissítés gyakorisága az utak átlagos hosszára nem volt hatással, ezért ezt nem is ábrázoltuk. A d paraméter növelésével a LOADNEAR esetében csökkent, míg a NEARLOAD esetében nőtt a választott útvonalak hossza. Az ábrákról leolvasható, hogy a NEARLOAD algoritmus már $d=3$ esetén is alacsonyban tartja a blokkolást. A LOADNEAR algoritmus ezzel szemben, amíg a d paraméter értéke nem éri el az átjárók számának felét, jól teljesít, később viszont a blokkolási arány egyre rohamosabb ütemben növekszik.

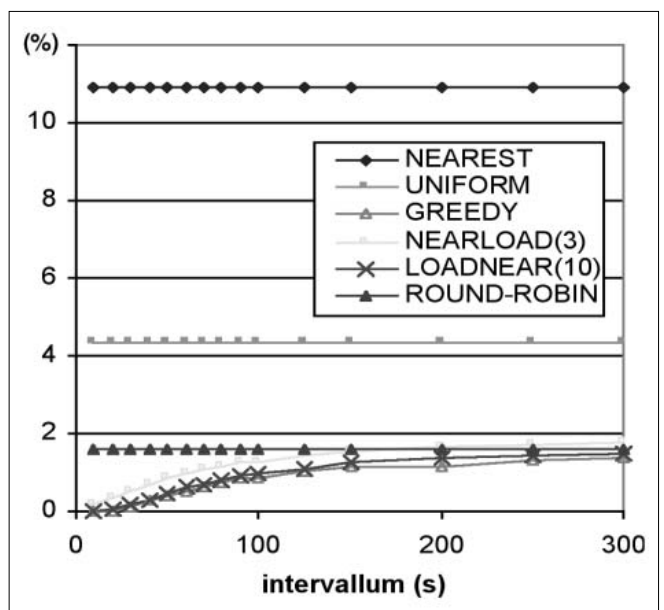
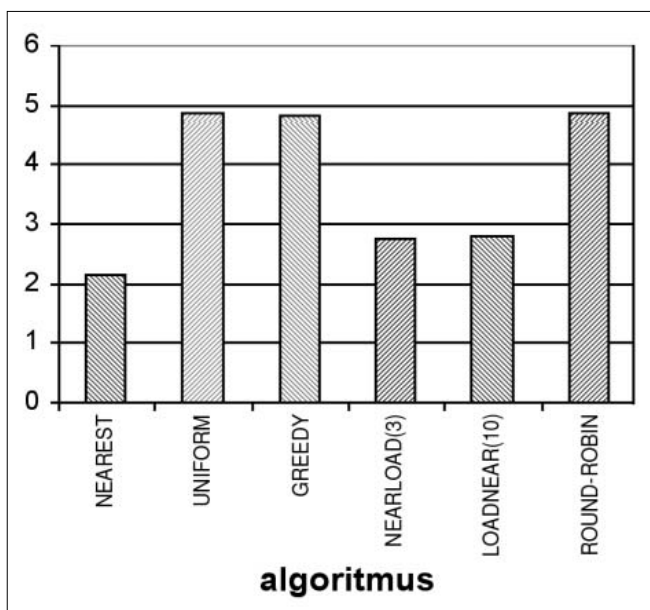
Ezen megfontolásokat megfontolva arra juthatunk, hogy ezen a topológián a NEARLOAD $d=3$, míg a LOADNEAR $d=10$ esetén produkálja a legjobb eredményeket, mindkét szempontot figyelembe véve. A további ábrákon ezért már csak ezen paraméterértékek melletti eredményeket tüntettük fel.



A 4. ábrán láthatóak a különböző algoritmusok által produkált blokkolási arányok az információterjesztés gyakoriságának függvényében. Látható, hogy a legrosszabb eredményeket a NEAREST algoritmus produkálta, azonban az 5. ábra mutatja, hogy sikeresen minimalizálta az utak hosszát. Véletlen algoritmus használatával (UNIFORM) a blokkolási arány csökkenthető, azonban ennek ára az utak megnövekedett hossza. A ROUND-ROBIN algoritmussal az utak hosszát tekintve javulást nem tudunk elérni, azonban a visszautasított hálózati kapcsolatok száma csökken a UNIFORM-hoz képest. Ennek oka abban keresendő, hogy ez az algoritmus mindig új átjárót választ, míg a UNIFORM esetében lehetséges egy átjáró többszöri egymás utáni választása is.

Az intelligensebb algoritmusok (GREEDY, NEARLOAD, LOADNEAR) teljesítményére kihat az információfrissítés gyakorisága. Az eredményekből azt tapasztaltuk, hogy az információküldés ritkításával egyre rosszabb

4. és 5. ábra Az útvonalak átlagos hossza és a blokkolási arány különböző információterjesztési intervallumok esetén



eredményt érnek el, illetve mindhárom teljesítménye a ROUND-ROBIN algoritmus eredményeihez konvergál. Ez azért van így, mert a PCE figyelembe veszi saját korábbi döntéseit, azonban a felhasználók távozásáról nem értesül. Így, ha két frissítés között sok idő telik el, először kiegyenlíti a szerverek terheltségét az általa látott hálózati képben, majd gyakorlatilag körülfordulós módon választ átjárót a következő frissítés vételéig.

Azt kaptuk, hogy a LOADNEAR, illetve a NEARLOAD megfelelő paraméterértékek melletti alkalmazásával – a blokkolási ráta kis növekedése árán – jelentősen csökkenteni tudjuk az átjárókhöz vezető útvonalak átlagos hosszát. A legjobb blokkolási ráta a GREEDY alkalmazásával érhető el, de mivel ez nem veszi figyelembe a topológiát a döntéskor, hosszú útvonalakat fog létesíteni a hálózatban.

5. Összefoglalás

Ha egy bizonyos igénytípust hálózati szerverek egy csoportja képes kiszolgálni, akkor egy konkrét igény kiszolgálását elvégző szerver kiválasztása egy megoldandó feladat. Mivel a hálózati terhelés alacsonyan tartása, illetve a szerverek terheltségének megfelelő elosztása két, egymásnak ellentmondó szempont, így lehetetlen olyan algoritmust találni, amely mindkét szempont szerint optimális eredményt ad. Szerencsére egyszerű algoritmusok használatával, melyek mindkét szempontot figyelembe veszik, jó kompromisszumot lehet elérni. Összességében a paraméterek megfelelő választása mellett a NEARLOAD, illetve a LOADNEAR megfelelő eredményt nyújt.

Irodalom

- [1] Cardellini, V., Colajanni, M., Yu, P.S., „Dynamic load balancing on Web-server systems”, Internet Computing, IEEE, Vol. 3, No.3, pp.28–39., May/June 1999.
- [2] J. Shin, H. Lee, J. Na, A. Park, „Load Balancing among Internet Gateways in Ad Hoc Networks”, in Proc. of Vehicular Technology Conference, 2005. Vol. 3, pp.1677–1680.
- [3] Miura, H., Yamamoto, M., Nishimura, K., Ikeda, H., „Server Load Balancing with Network Support: Active Anycast”. in H. Yasuda (ed.), Proc. of 2nd International Working Conference on Active Networks (IWAN 2000), Springer LNCS 1942, p.371., Tokyo, October 2000.
- [4] Mannie E. (ed.), „Generalized Multi-Protocol Label Switching (GMPLS) Architecture”, RFC3945, October 2004.
- [5] Coltun, R., „The OSPF Opaque LSA Option”, RFC2370, July 1998.
- [6] Farrel, A., Vasseur J.-P., Ash J., „A Path Computation Element (PCE)-Based Architecture”, RFC4655, August 2006.
- [7] Ns-2 honlap, <http://www.isi.edu/nsnam/ns>

Helyreigazítás

Legutóbbi, 2007/3-as számunk 19-23. oldalán, **Csurgai-Horváth László és Bitó János**: „**Fading időtartam-modellezés műholdas földi mozgó rádiócsatornán**” című cikkébe sajnálatos módon több, értelemzavaró hiba került, melyek a következők:

- Az 1. táblázat helyére egy rokon témájú cikk hasonló tartalmú táblázata került. Az ide tartozó táblázat helyesen:
- A (3) és (4) képletekben az eredeti, görög szumma helyett hibás karakter jelent meg.
- A 2. táblázat felirata pedig helyesen: „Paraméterek a (7-8) egyenletekhez”

Műhold neve	MARECS (d=39150 km)
Eleváció	24°
Frekvencia	1.54 GHz
Mintavételezési frekvencia	300.5 Hz
Csatorna ID	14
Környezet	Autópálya
Mérési időtartam	81.2 perc
Jármű sebessége	60 km/h

A szerzőktől és olvasóinktól egyaránt elnézést kérünk!

Ütközésfeloldás RFID rendszerekben

TÓTH KATALIN, SCHULCZ RÓBERT, IMRE SÁNDOR

BME Híradástechnikai Tanszék, Mobil Távközlési Laboratórium
ktoth@mcl.hu, {schulcz, imre}@hit.bme.hu

Lektorált

Kulcsszavak: RFID, olvasók, transzponderek, ütközésfeloldás

Cikkünk az RFID környezetben felmerülő többes leolvasási problémák megoldására rendelkezésre álló ütközésfeloldó algoritmusok ismertetési kitérve az olvasók és a transzponderek ütközésére egyaránt. Továbbá az ISO 18000-3 Mode 1 ütközésfeloldásra használt algoritmusának szimulációs eredményeit is ismertetjük.

1. Az RFID rendszerek és a többes leolvasás problémaköre

Az RFID az angol Radio Frequency Identification, azaz a rádiófrekvenciás azonosítás kifejezés rövidítése. Ez a technológia automatikus azonosításra szolgál, jelleget tekintve a birtok alapú azonosítási rendszerek közé sorolható.

Az RFID rendszerek alapvetően legalább két eszkből állnak: egy azonosítóból és egy azonosítandóból. Az azonosító adatkapcsolatot kezdeményez az azonosítandóval, mely során mindkét vagy csak az egyik irányba adatátvitel történik. Ahogyan a technológia nevéből is kiderül, a kommunikáció rádiófrekvencián zajlik, ezért mind az azonosítónak, mind az azonosítottak rádiós interfésszel kell rendelkeznie.

Az azonosítani kívánt objektumon (pl.: árucikk) egy transzpondert helyeznek el. Az azonosító egy RFID olvasó, amely képes olvasni és/vagy írni a transzpondert.

Egy az előbb vázolt egyszerű RFID rendszer kiegészülhet még alkalmazási környezettől függően vezérlő számítógéppel, mely több olvasó működésének összehangolását vezérli, valamint összeköttetést teremt az olvasók és a háttéradatbázis(ok) között. A háttéradatbázis funkcionálisát tekintve alkalmas az olvasó(k) érzékelési távolságán belüli transzponderekkel kapcsolatos információk tárolására; nem csak lekérdezhetőek, hanem az adatátvitel alatt írhatóak is. Az 1. ábra egy RFID rendszer felépítését mutatja.

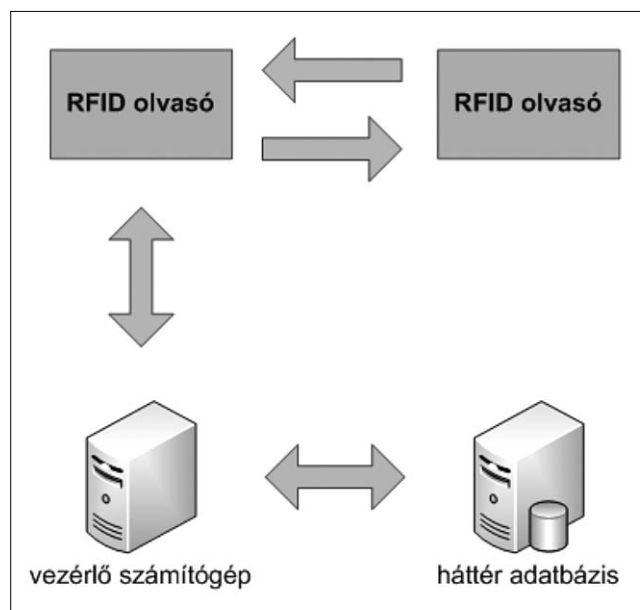
Az RFID alkalmazások az ellátási lánc teljes hosszán, azaz az egyes alkatrészek, alapanyagok előállításától, előkészítésétől egészen a végfelhasználókhoz való eljutásig jelen vannak. Éppen ezért gyakori, hogy egy raktári alkalmazás esetében másodpercenként több száz címkét kell leolvasni. Az alapvető követelmény az RFID rendszerek esetén a minél több címke minél rövidebb idő alatt történő leolvasása nagy megbízhatósággal, azaz ne maradjon ki egyetlen címke sem a leolvasásból. Előfordulhat az is, hogy nem a transzponderek ütköznek, hanem az olvasók. Ilyen környezet képzelhető el egy hipermarketben, ahol sok leolvasó kapu

sorakozik egymás mellett a pénztáraknál. Ebben az esetben ütközés léphet fel az egyes olvasók között, hiszen két vagy több olvasó által lefedett térrészeknek lehetnek közös területei. Ha itt tartózkodik egy transzponder, akkor ütközés lép fel az olvasók között.

Az ütközések legszembetűnőbb negatív hatása a leolvasási sebesség csökkenése. Ennek kiküszöbölésére két módszer kínálkozik: a sávszélesség növelése és az ütközések eliminálása. A sávszélessége növelésének fizikai korlátja mellett szempont az is, hogy ennek a módszernek a kivitelezési költsége jóval nagyobb, mint egy beépített algoritmus futtatásának. Ezért a továbbiakban a különböző ütközésfeloldó algoritmusokat ismertetjük a legelterjedtebb szabványok esetén, továbbá bemutatunk néhány szimulációs eredményt is arra vonatkozóan, milyen paraméterek befolyásolják a leolvasási időt.

RFID rendszerekben ütközésfeloldó eljárásnak nevezzük azt a technikát, vagy protokollt, amely lehetővé teszi az interferencia nélküli többszörös hozzáférést.

1. ábra RFID rendszer felépítése



2. Ütközésfeloldó algoritmusok olvasók ütközésére

Az olvasók ütközésének feloldására három algoritmus terjedt el: NAK-alapú, beacon-alapú megoldások, valamint a változtatható hatótávolságú olvasók alkalmazása.

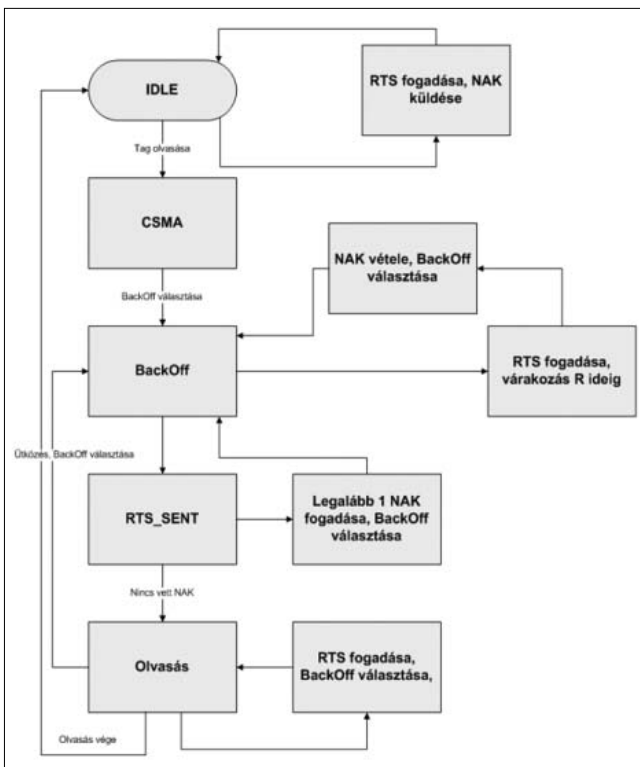
A *NAK-alapú ütközésfeloldás* feltétele, hogy minden olvasó saját vezérlőcsatornával rendelkezzen. E vezérlőcsatorna hatósugara elég nagy ahhoz, hogy az olvasó az összes, a környezetében lévő – az adatcsatornában interferenciát előidéző – olvasóról tudjon. (A NAK a Not Acknowledged, azaz nem nyugtázott kifejezés rövidítése.)

Ennek megfelelően az olvasó először egy RTS üzenetet küld szét a környezetében lévő összes olvasónak. Amelyik épp foglalt, egy NAK üzenetet küld vissza. Ha az olvasó legalább egy NAK-ot vesz, korlátozza az adatcsatornához való hozzáférést. Ezt a folyamatot szemlélteti a 2. ábra.

A *beacon-alapú olvasók közötti ütközésfeloldás* feltételei a NAK-alapú megoldással azonosak: saját, nagy hatósugarú vezérlőcsatorna. Az eljárás szerint az olvasó – olvasás közben – periodikus beacon üzeneteket küld a vezérlőcsatornáján. Ha egy másik olvasó is olvasni szeretne, előbb ellenőrzi az utolsó beacon óta eltelt időt. Ha ez kisebb, mint egy megadott érték (vagyis a beacon periódusideje), az olvasás valószínűleg még folyamatban van, tehát várakozik, míg az előírt idő alatt nem jön újabb beacon. Az algoritmus működése a 3. ábrán követhető nyomon.

Kézenfekvő és egyszerű megoldás az egyes olvasók hatósugarának adaptív változtatása. Az eljárás min-

2. ábra NAK alapú ütközésfeloldás



den n olvasási ciklus után a felügyelt terület duplázásával, ütközés esetén pedig p valószínűséggel ennek csökkentésével optimalizálja a leolvasásokat. Ezzel gyakorlatilag egy SDMA (Space Division Multiple Access), azaz térosztásos többszörös hozzáférés valósul meg.

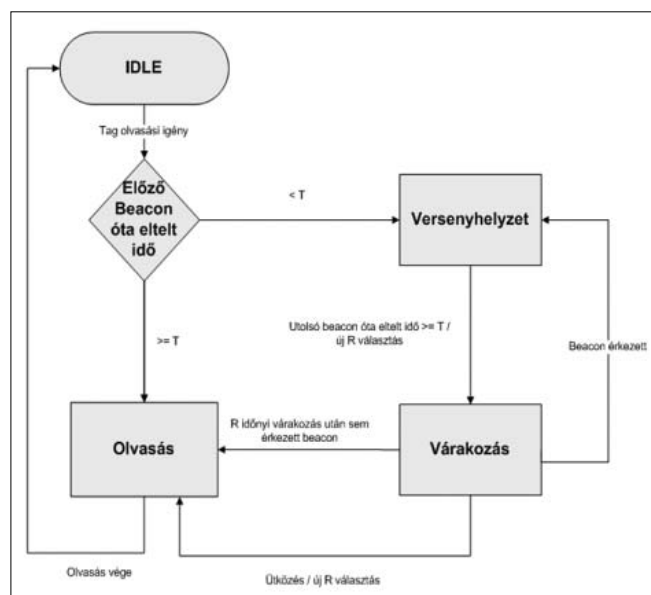
3. Ütközésfeloldó algoritmusok transzponderek ütközésére

Transzponderek ütközése úgy jöhet létre, hogy az olvasó hatósugarában számos tag van jelen. Olvasáskor a leolvasó egység egy request üzenetet broadcastol a tag-eknek. Amennyiben csak egyetlen tag tartózkodik a térben, az olvasó egyetlen választ vesz. Ugyanakkor minden tag visszaküldi a választ az olvasónak, ha több is van belőlük a rendszerben. Ez ütközést eredményez a rádiócsatornán, így válaszüzenetük nem lesz értelmezhető. Ennek kiküszöbölésére fejlesztettek ki számos ütközésselkerülő algoritmust, amelyeket az alábbiakban ismertetünk.

Az ütközések feloldása az esetek nagy részében valamilyen *többszörös közeghozzáférési feladat* megoldására vezethető vissza, így a következőkben ezek a – részben már más területekről ismert – algoritmusok kerülnek bemutatásra.

Az RFID rendszerekben is négyféle többszörös hozzáférési eljárást különböztetünk meg: térosztásos többszörös hozzáférés (SDMA, Space DMA), frekvencia-osztásos többszörös hozzáférés (FDMA, Frequency DMA), időosztásos többszörös hozzáférés (TDMA, Time DMA) és végül a kódosztásos többszörös hozzáférés (CDMA, Code DMA). Fontos megjegyezni, hogy néhány gyártó nem hozza nyilvánosságra saját ütközésfeloldó algoritmusait, míg mások a szabványban is lefektetett, elterjedt megoldásokat használják. Általánosságban elmondható, hogy az egyedi megoldások is ezen szabványokból táplálkoznak némi alkalmazás-specifikus kiterjesztés felhasználásával.

3. ábra Beacon alapú ütközésfeloldás



Az **SDMA** eljárás a rendelkezésre álló területet szeletekre osztva növeli meg a kapacitást.

Egyik lehetséges megoldás az olvasók hatótávolságának jelentős csökkentése, ezzel együtt sok olvasó és antenna elhelyezése a mezőben. Az így létrejövő tömbök nagyjából hézagmentesen fedik le a kívánt területet. Ennek eredményeképpen a szomszédos olvasók számára többszörösen is rendelkezésére áll a csatornkapacitás.

Egy másik lehetőség az *adaptív SDMA* megvalósítása, amikor elektronikusan vezérelt irányított antennát használunk az olvasóban, mely akár mikor az aktuálisan kapcsolatban lévő transzponderre irányítható. Így az egyes tag-ek akár térbeli helyzetük, egymással és az olvasóval bezárt szögük alapján különböztethetők meg. Ilyen összeköttetések 850 MHz feletti RFID rendszerekben valósíthatók meg leginkább. Az eljárás legfőbb hátránya az összetett antenna és antennarendszerek kiépítése, mely csak speciális alkalmazásokban teremt neki létjogosultságot. Adaptív SDMA rendszerekben az elektronikusan vezérelt antenna iránya a fenti módon körforgásban is változtatható, így lépésről lépésre haladva minden transzponder beleesik a nyálkába.

FDMA felhasználása során egyidejűleg több tag számára is rendelkezésre áll egy-egy teljes értékű, más-más frekvenciájú vivő. RFID rendszerekben ez szabadon állítható, alharmonikus vivőket használó transzponderekkel valósítható meg. Mind a broadcast szinkroncsatorna, mind az energiaellátás egy alkalmasan választott f_a frekvencián kerül továbbításra. A transzponderek válaszukat a rendelkezésre álló néhány ($f_1 - f_N$) válaszcsatornán küldik vissza, így teljesen más frekvenciataromány használható up- és down irányokra.

Backscatter és terhelésmódult rendszerek számára egy rendelkezésre álló lehetőség több független vivőfrekvencia használata tag→olvasó irányban. Ezen eljárás hátránya szintén anyagi jellegű. Az olvasókba minden frekvenciához dedikált vevőket kell elhelyezni, ami jelentősen megnöveli az előállítási költségeket. Emiatt az FDMA használata csak egyes speciális alkalmazásokban ajánlott.

A **TDMA** azon eljárások összessége, melyek használata során a teljes rendelkezésre álló csatornkapacitás – a tag-ek valamilyen sorrendjét követve – az egymás után következő tag-ek számára teljes mértékben kiosztásra kerül. Széles körben elterjedt az alkalmazása többek közt a digitális mobil rádiórendszerekben. RFID rendszerekben ütközés feloldó algoritmusokban találkozunk vele. Megkülönböztetünk transzponder-vezérelt és adó/vevő-vezérelt eljárásokat.

A transzponder-vezérelt eljárások aszinkron módon működnek, mivel az olvasó nem szól bele az adatátvitel vezérlésébe (ezen az elven működik a későbbiekben tárgyalt ALOHA eljárás is). Tovább is csoportosíthatjuk az aszinkron módszereket: megkülönböztethetjük az eljárásokat aszerint is, hogy az olvasó küld-e egy „kikapcsol” jelzést a transzpondernek a sikeres adatátvitel után.

A transzponder-vezérelt eljárások nem elég flexibilek, ugyanakkor eredendően lassúak. A legtöbb alkalmazás ezért az adó/vevő-vezérelt eljárásokat használja, mely során az olvasó vezérli az adatátvitelt. E módszerek szinkron eljárásoknak tekinthetők, hiszen minden transzpondert az olvasó ellenőriz és vezérel folyamatosan.

Működése során az olvasó mindig pontosan egy transzpondert választ ki a zónán belül, majd e két fél között megindul az adatátvitel (hitelesítés, írás/olvasás stb.). Kizárólag ezután történhet meg a kapcsolat bontása és az új összeköttetés létesítése egy másik transzponderrel. Ezek az eljárások is tovább csoportosíthatók polling, azaz lekérdezéses, valamint binary search, azaz bináris keresés elven működő algoritmusokká. Minden eljárás közös abban, hogy az egyes tag-eket egy egyedi azonosító alapján választja ki.

Lekérdezéses eljárások

A lekérdezéses eljárásnak a korrekt működéshez feltétlenül szüksége van a kommunikációban vélhetőleg részt vevő tagok azonosítójáról készült listára. A lekérdezéses folyamat során az olvasó egyenként végig lépked ezen a listán, egészen addig, amíg egy transzponder a megfelelő szériaszámmal válaszol. Ez a módszer elég időigényes, csak kis számú transzponder esetén alkalmazható hatékonyan.

A lekérdezés elven működő eljárások közül a következőket fogjuk részletesen ismertetni: bináris keresés, dinamikus bináris keresés, réselt bináris fa és a bitről bite bináris fa.

A *bináris keresés* az egyik legelterjedtebb és legflexibilisebb algoritmusok egyike. Működése során a transzpondert egy olyan csoportból választja ki, melyek azonosítóit szándékosan ütközésbe hozza egymással úgy, hogy a tag-ek egyszerre küldik vissza sorszámaikat az olvasó kérésére válaszolva. Ehhez szükséges, hogy az olvasó képes legyen az egyes bitpozíciók pontos detektálására, melyet a helyes kódolás megválasztásával ér el (pl. Manchester-kódolás).

Az algoritmus előre definiált interakciók (parancsok és válaszok) sorozatát tartalmazza, mely segítségével képes kapcsolatot teremteni az olvasó és a tag-ek között azzal a céllal, hogy közülük majd egyet adatátvitel céljából kiválasszon. A tényleges működéshez egy parancskészletet definiál REQUEST, SELECT, READ_DATA ÉS UNSLECT utasításokkal. Ezen kívül minden transzponder egy egyedi sorozatszámra rendelkezik. A könnyebb megértés érdekében az algoritmus működését egy olyan példával szemléltetjük, melyben 8 bites ID-k szerepelnek, ezzel 256 különböző tag egyediségét garantálva.

A 8 bites sorozatszámokat használó transzponderek a hexadecimális számrendszerben a 00-FFh, a decimális számrendszerben a 0-255, míg a bináris számrendszerben a 00000000-11111111 azonosítókkal jellemezhetők. Az algoritmus első lépéseként az olvasó egy REQUEST parancsot küld broadcast üzenetként (≤ 11111111) argumentummal. A 11111111b-es azono-

sító a legnagyobb értékű a 8 bites sorozatszámok között. Ez azt jelenti, hogy a zónában lévő minden tag válaszolni fog erre a kérésre, hiszen csak kisebb, vagy egyenlő lehet az azonosítójuk, mint 1111111b.

A precíz bitszinkron elengedhetetlen az algoritmus működéséhez, ennek biztosítása kritikus az ütközések elkerülése és felismerése szempontjából.

Az 1. táblázatban látható, hogy a 0., a 4. és a 6. bitpozícióban ütköznek a sorozatszámok, mivel ezeken a helyeken a különböző bitértékek szuperponálódnak. Az ütközés az olvasó számára azt jelenti, hogy legalább két transzponder tartózkodik a zónában. A vett jel a következő lenne: 1X1X001X.

A 6-os bitpozíció a legmagasabb értékű a sorozatszámokon belül, ahol még ütközés történt az első iteráció során. Ez azt jelenti, hogy legalább egy transzponder található az $SNR > 11000000b$ és az $SNR < 10111111b$ tartományokban. Ahhoz, hogy egyetlen transzpontert legyünk képesek kiválasztani az egyelőre ismeretlen számú csoportból, limitálni kell a lehetséges sorszám-tartományt a következő iterációk során.

A folyamatot kezdjük a $< 10111111b$ tartománnyal, vagyis a 6. pozíciót 0-nak rögzítjük, és figyelmen kívül hagyjuk az alacsonyabb tartományt, az adott bitértékek 1-esre állításával.

Amint az olvasó kiadta a REQUEST (≤ 10111111) parancsot, minden – a feltételt kielégítő – transzponder saját sorszámával válaszol. A példában ezek az 1, 2 és 3-as számú transzponderek. Ebből megállapítható, hogy legalább két transzponder lesz jelen a második iterációban is. A vett bitsorozat (101X001X) még mindig négy lehetséges sorozatszámot takar. Az ütközések ismételt jelenléte a 2. iteráció során újabb szűkítést igényel a 3. lépésben.

Ekkor az olvasó a REQUEST (≤ 10101111) parancsot adja ki a 2. táblázat szerinti szabályok alapján. Ezt a feltételt csak a 2-es transzponder teljesíti a 10100011 sorozatszámával. Mivel sikerült egyetlen valós sorozatszámot kiválasztani, további iterációra nincs szükség. Jelenleg tehát kommunikációra vagyunk képesek a 2-es számú transzponderrel. Írhatjuk, olvashatjuk, anélkül hogy a többi tag-gel interferenciába kerülne. Az adatátvitel végeztével egy UNSELECT kiadásával elnémíthatjuk az ismételt REQUEST-ig így biztos inaktív tag-et. A maradék tag-ek kiválasztásához szükséges iterációk száma egyre csökken, ahogy sorban válnak UNSELECT állapotúvá a transzponderek.

Az iterációk átlagos L száma, mely megmondja hány lépésben detektálható egyetlen transzponder egy n tagból álló csoportban könnyen számolható:

$$L(N) = \lg(N) + 1 = \frac{\log(N)}{\log(2)} + 1$$

Transzponder 1	10 11 00 10
Transzponder 2	10 10 00 11
Transzponder 3	10 11 00 11
Transzponder 4	11 10 00 11

Amennyiben egyetlen transzponder tartózkodik a zónában, az iterációk száma természetesen egy, ha ennél több, L könnyen megnövekedhet.

A bináris keresés algoritmus legfőbb negatívuma a rendkívül pazarló működése volt, hiszen mind a keresési/lekérdezési kritériumok, mind a sorozatszám teljes hosszában átvitelre került minden lépésben. Ezt küszöböli ki a dinamikus bináris keresés algoritmus: a teljes sorozatszámok átvitele helyett most azoknak csak egy részét, X -nél vágott hányadukat vesszük át. X jelöli azt a legmagasabb bitpozíciót, ahol az előző iteráció során ütközés volt. Az olvasó tehát csak $(N-X)$ ismert részét küldi el a REQUEST parancs argumentumaként, majd leállítja az adást. Minden transzponder, mely sorozatszám kezdete egyezik $(N-X)$ -szel, válaszol az $(X-1)$ -től 0-ig tartó, fennmaradó bitek átvitelével.

A *reselt bináris fa algoritmus* működése a következő: amennyiben az i . időrésben ütközés történik, minden transzponder, mely nem vett részt benne, vár annak feloldásáig. Az ütköző tag-ek véletlenszerűen két csoportba (nullás, egyes) kerülnek szétválasztásra. Az első (nullás) csoportban lévő tag-ek az $(i+1)$. időrésben adnak újra, míg a másik csoportban lévő transzponderek várnak ennek sikeres végrehajtásáig. Amennyiben az $(i+1)$. időrés üres, vagy sikeres átvitel történt benne, úgy a második csoport elemei az $(i+2)$. időrésben adnak újra. Amennyiben az $(i+1)$. időrés ütközést tartalmaz, úgy az egész eljárás újra megismétlődik.

A szükséges iterációk száma n számú tag esetén a következő egyenlet alapján alakul:

$$I_{SBT} = 1 + \sum_{k=2}^n \binom{n}{k} \frac{2(k-1)(-1)^k}{[1-p^k - (1-p)^k]}$$

Végül a *bitről bitre bináris fa típusú keresést* mutatjuk be az ALOHA típusú ütközéselkerülő algoritmusok ismertetése előtt. A bitről-bitre algoritmus induláskor bekéri a transzponderek azonosítójának 1. bitjét. Miután a tag-ek elküldték a 0-t, vagy az 1-et, az olvasó megnézi, volt-e ütközés, mielőtt bekéri a következő bitet. Amennyiben volt, két csoportba osztja az ütköző tag-eket, majd kiválasztja az egyiket, és csak tőlük bekéri a következő bitpozíció értékét.

Az algoritmus e lépéseket ismétli, míg a teljes – sorozatszámokból felépített – bináris fát be nem járja. Ha ütközés nélkül veszi a k . bitet, annak értékét eltárolja a memóriájában. Ellenkező esetben, ütközés esetén a k . bitet 0-ként menti el, majd inaktív állapotba küldi az összes tag-et, melyek k . bitje 1. Ebben az állapotban a transzponderek ideiglenesen letiltják saját azonosítójuk küldését egészen addig, míg egy tag teljesen azonosításra nem került. Az algoritmus addig ismétli magát, míg ezzel minden tag-et azonosít.

A szükséges iterációk (I_{BBT}) száma: $I_{BBT} = n \times j$.

1. és 2. táblázat

Parancs	1. iteráció	n. iteráció
REQUEST > RANGE	0	Bit(x) = 1, Bit (0-tól (X-1)-ig) = 0
REQUEST < RANGE	SNRmax	Bit(x) = 0, Bit (0-tól (X-1)-ig) = 1

ALOHA-típusú eljárások

A transzponder által vezérelt ütközésfeoldó eljárások ismertetése során az ALOHA-típusú eljárásokról lesz szó.

Az ALOHA eljárás működésének alapját az képezi, hogy amint egy adatcsomag készen áll a küldésre, az nyomban továbbításra is kerül.

Ez a módszer általában csak olvasható transzponderekkel együtt használatos, azok alacsony csomagmérete miatt (jellemzően azonosítók). Ez az adat ciklikus sorozatban kerül küldésre, mely hosszú adások közti szünetekkel jár együtt. Ugyanakkor az egyes tag-ek által használt ismétlési idők nagyban különböznek egymástól. Ez azt jelenti, hogy nagy eséllyel különböző időpillanatokban lépnek fel az egyes tag-ek adatátviteli igényel, kicsi a torlódás/ütközés esélye.

A terhelés természetesen összhangban van az egyes időpillanatokban egyidejűleg adó transzponderek számával. Az átlagos G terhelés egy T megfigyelési időtartamra vett átlag, mely könnyen számolható az adatcsomag τ átviteli idejéből:

$$G = \sum_{i=1}^n \frac{\tau_i}{T} r_i$$

ahol n a rendszerben lévő transzponderek száma, és $r_i = 0, 1, 2, \dots$ az i . transzponder által átvitt adatcsomagok száma. Az s -sel jelzett áteresztőképesség értéke 1 a hibamentes átvitel idejére (ütközésmentes esetben), valamint 0 minden más esetben (például nincs adatátvitel, torlódás miatt nem továbbítható adatcsomag hiba nélkül).

Az átlagos S kibocsátás a felkínált G terhelésből számítható: $S = G \cdot e^{-2G}$.

Ha megvizsgáljuk az S kibocsátást a felkínált G terhelés függvényében, azt találjuk, hogy $G = 0,5$ -nél S -nek egy maximuma van, mely értéke 18,4%.

Kisebbs terhelés esetén a csatorna az idő nagy részében üres lenne; nagyobb terhelés az ütköző transzponderek számának rohamos növekedéséhez vezetne. Ilyen formán a csatorna kapacitásának 80%-a kihasználatlannak látszik. Az ALOHA egyszerű implementációja viszont lehetővé teszi annak ütközés feloldó algoritmusként való használatát. Ezen eljárást használják digitális hálózatokban, mint például a csomagkapcsolt rádió, mely világszerte kedvelt kommunikációs forma a rádióamatőrök között.

A sikeres átvitel q valószínűsége az átlagos G terhelésből és az S kibocsátásból számolható az alábbi képlet alapján:

$$q = \frac{S}{G} = e^{-2G}$$

A k számú adatcsomag T idő alatt sikeres átvitelének $p(k)$ valószínűsége könnyen számolható az átvitelhez szükséges $ridő$, valamint az átlagos G terhelésből. A $p(k)$ eloszlás Poisson-jellegű:

$$p(k) = \frac{(G \frac{T}{\tau})^k}{k!} e^{-(G \frac{T}{\tau})}$$

Az ALOHA relatív alacsony teljesítményének növelésére fejlesztették ki a **réselt ALOHA** (S-ALOHA, Slot-

ted ALOHA) eljárást. A tag-ek, amennyiben ezt az algoritmust használják, csak meghatározott időpillanatokban kezdeményezhetnek adást (szinkronpontok, időrések). A tag-ek szinkronizációja az olvasó feladata, ami következtében ez egy sztochasztikus, adó/vevő-vezérelt TDMA ütközésfeloldó eljárás. Ezt használva ütközés fele akkora periódusban következhet csak be, mint ALOHA-t használva.

Azonos méretű adatcsomagokat (és ezáltal azonos átviteli időt) feltételezve, ALOHA használata során ütközés lép fel, amennyiben két transzponder $T < 2\tau$ időintervallumon belül küld. Réselt ALOHA és a szinkronpontok használatával ez az intervallum $T = \tau$ -ra csökken le. A réselt ALOHA áteresztőképessége ezen megfontolások alapján számolható: $S = G \cdot e^{-G}$.

A réselt ALOHA rendszerek S kibocsátása maximummal rendelkezik $G=1$ terhelésnél. Ez azt jelenti, hogy pontosan annyi transzponder van a zónában, ahány szabad időrés. Ha túl sok transzpondert helyezünk a térbe, az átbocsátóképesség hamar 0 közelébe esik le, legrosszabb esetben végtelen számú próbálkozás után sem leszünk képesek sorozatszámot detektálni. Ez a jelenség elkerülhető megfelelő számú időrés meghatározásával, ugyanakkor várható az ütközésfeloldó algoritmus teljesítményének csökkenése. Ennek oka az, hogy a rendszernek állandóan figyelnie kell minden elérhető transzponderre a zónában a szeletek teljes időtartama alatt akkor is, ha csak egy transzponder tartózkodik a térben.

A **dinamikus réselt ALOHA** ezen próbál segíteni az időrések számának dinamikus megválasztásával.

Egyik lehetőség erre a rendelkezésre álló időrések számának átvitele, melyet a transzponderek a REQUEST parancs argumentumaként kapnak meg: várakozó üzemmódban az olvasó ciklikus REQUEST parancsokat broadcastol, melyet egy vagy két időrés követ a lehetséges transzponderek számára. Ha a transzponderek nagy száma jelenti a szűk keresztmetszetet az időrésekben, minden REQUEST üzenettel megnő ezek száma, amíg egy egyedi transzponder tisztán detektálhatóvá válik. Ezt követően az olvasó egy BREAK üzenetet küld ki, mely hatására a detektálton kívüli transzponderek blokkolt állapotba kerülnek.

4. SimTag szimuláció

Most, hogy mind az olvasók, mind a transzponderek ütközéseit feloldó alkalmazható algoritmusokat ismertettük, érdemes megvizsgálni, hogy a gyakorlatban milyen paraméterek befolyásolják a többes leolvasás idejét, hatékonyságát. Ehhez egy ISO által kibocsátott szabvány szerinti ütközésfeloldó algoritmus szimulációs eredményeit mutatjuk be.

Az **ISO 18000-3 MODE 1** alapértelmezésben nem definiál ütközésmenedzsmentet. Ez a működés egy sima ALOHA eljárásnak feleltethető meg, mely során az ütköző tag-ek valamikor a későbbiekben újraadnak. A szimulációt a SimTag ütközésfeloldó protokollok szimu-

lációjára kifejlesztett programjával végeztük: mindenre kiterjedő paraméterkínálatával számos különböző szimuláció futtatását teszi lehetővé.

A szimulációs eredmények értelmezésének kulcsa az algoritmus *Main* és *Extended módja* közti különbségek alapos ismerete. A két eljárás közötti szembetűnő különbség az időrések kiosztásában van. Míg a *Main* mód fix 16 időrés/kör paraméterrel dolgozik, addig az *Extended* mód adaptív módon, dinamikusan, a tag-ek és az ütközések számának függvényében állítja ezt a paramétert működés közben. A tartomány határai az előzetes konfiguráció során adhatók meg.

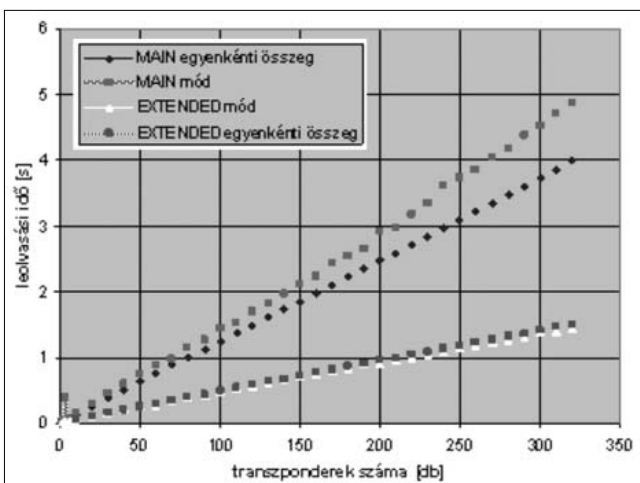
A szimulációkat két fő területre koncentrálva futtatuk le: a teljes leolvasási időt vizsgálva, illetve az időrések számának változtatásával futtatva a programot.

A teljes leolvasási idő vizsgálata

A teljes leolvasási idő szimulációja során megvizsgáltuk, hogyan függ a populáció méretétől az abszolút leolvasási idő rögzített paraméterek mellett *Main*, illetve *Extended* módok használata esetén.

Első lépésként egy tesztet futtattunk *Main* módban, mely 1-től 350-es populációig vizsgálta a transzponderenkénti leolvasási időket. Ezt összegeztük, később – mint egy idealista – viszonyítási alapként szolgált: mekkora lenne a teljes leolvasási idő, ha ez kizárólag a transzponderenkénti leolvasási időkből állna össze, semmiféle járulékos időt nem tartalmazna.

Ezzel szemben áll egy másik teszt eredménye, mely – *Main*, valamint *Extended* módokban lefuttatva a szimulátort – tartalmazza 6 teljes leolvasás átlagolt eredményét a transzponderek növekvő populációját tekintve. A teszt eredménye 1-től 320-as populációig 10-es lépésekben egyenként 6 mérés átlagát tartalmazza.



4. ábra

A leolvasási idő alakulása a transzponderek számának függvényében *MAIN* és *EXTENDED* módban

Az eredmények egymásra helyezett grafikonjai (4. ábra) jól mutatják az egyes algoritmusok közötti alapvető különbségeket.

A transzponderenkénti összeg jó közelítéssel egy lineáris görbe, hiszen a *Main* módban leolvasott transzponderek összege néhány tíz transzponder felett nem

ingadozik nagy mértékben. Ehhez képest a valódi ütközésfeloldó algoritmus (*Main* módban) szintén lineáris, de meredekebb jellegű görbe. A járulékos kommunikációs idők tehát egyre nagyobb mértékben rakódnak rá. Ennek oka a *Main* mód működésében keresendő. Az eljárás fix, 16 időrést használ körönként. Ez a nem dinamikus, adaptív algoritmus egyre rosszabb határfokkal működik, ahogy a populáció mérete nő.

Ennél lényegesen jobb, gyorsabb működésű az *Extended* mód. 300-as populációval 1,36 mp alatt végez, míg a *Main* mód 4,5 mp alatt teszi ezt. Az előbbi algoritmus előnye annál nagyobb, minél nagyobb a populáció, éppen ezért nagy számú tag-eket használó alkalmazásokban érdemes implementálni az *Extended* módot. Az algoritmus adaptív, dinamikusan változtatja a körönkénti időrések számát (természetesen az előzetes konfigurációtól függően), így sokkal flexibilisebb, képes alkalmazkodni nem csak a transzponderek számához, hanem az ütközések arányához is.

Időréseparaméterek hatásvizsgálata

Az időréseparaméterek hatásvizsgálatának szemléltetése előtt érdemes megjegyezni, hogy a szimuláció annak köszönhetően végezhető el, hogy az *Extended* mód három paraméteren keresztül kínál fel lehetőséget az időrések számának befolyásolására. A kezdő, a minimális, valamint a maximális időrések számával tagok között konfigurálható az ütközésfeloldó algoritmus, melynek eredményeit a következőkben ismertetem.

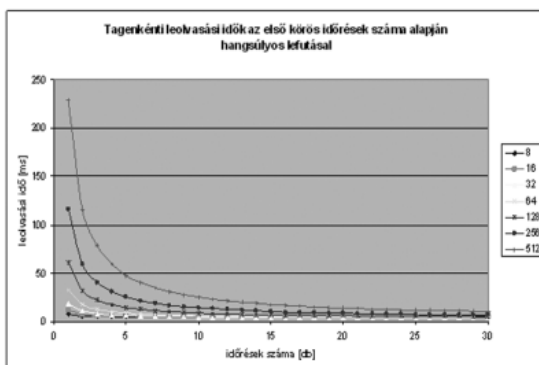
A *kezdő* – vagyis az első körben kiosztásra kerülő – *időrések száma* alapjaiban befolyásolhatja a leolvasási időket. Hogy ez miként függ a zónában lévő transzponderek számától, azt az alábbi szimuláció ismerteti.

Első lépésként hét szimulációt futtattunk 8, 16, 32, 64, 128, 256, valamint 512 kezdő időrés használatával. A tesztek 320-as populációval készültek, egyenként 25-25 mérés átlagaként. Az összesített eredményeket a következő oldalon az 5-8. ábrák szemléltetik.

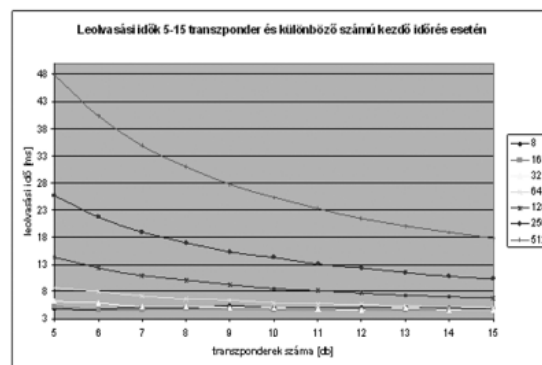
Látható, hogy a több időréses első kör hátránnyal indul kis számú transzponder esetén, ami természetes is, hiszen rengeteg időrés marad kihasználatlanul ez esetben, ami megnöveli az átviteli időt. Ez a hátrány azonban elfogy 150-200-as populációnál, mely felett a több időréses leolvasások jobban teljesítenek. Ugrászerű a növekedés 512 időrést használata esetén, de a 8-hoz képest már a 64-es időrés szám is szembetűnő növekedést jelent. Az időrések számát tehát célszerű az adott alkalmazáshoz igazítani, hogy a kihasználatlan időrések ne rontsák az algoritmus teljesítményét, ezzel pedig az olvasás időszükségletét. Fontos még egyszer megjegyezni, ezek az eredmények az adott populációt alkotó tag-ek egyenkénti leolvasási idejére vonatkoznak, nem pedig a teljes időre.

A körönkénti *minimálisan kiosztandó időrések száma* szintén erős befolyásoló tényező. Túlzottan magas értékével rengeteg kihasználatlan időrés keletkezik, alacsonyra állítása ugyanakkor sok ütközést eredményez. Pontos beállítása szintén szimuláció segítségével történhet a rendszerépítés ezen fázisában.

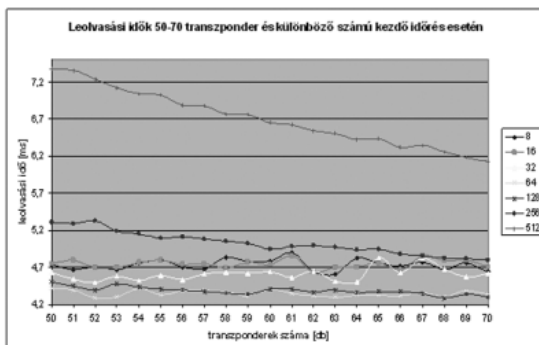
5. ábra
A tag-enkénti
leolvasási idők
alakulása
az első körös
időresek száma
alapján hang-
súlyos lefutással



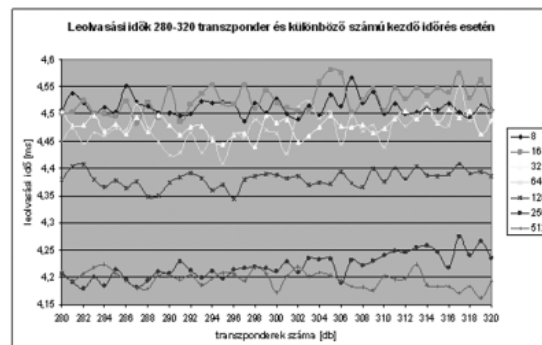
6. ábra
A leolvasási idők
alakulása 5-15
transzponderre
(kezdő időresek
száma szerint)



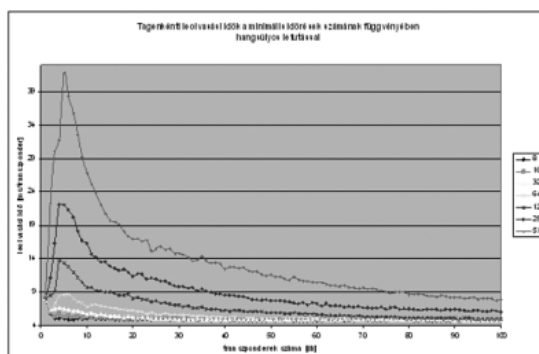
7. ábra
A leolvasási idők
alakulása 50-70
transzponderre
(kezdő időresek
száma szerint)



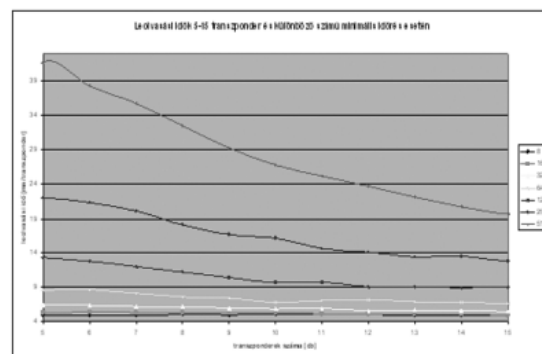
8. ábra
A leolvasási idők
alakulása 280-320
transzponderre
(kezdő időresek
száma szerint)



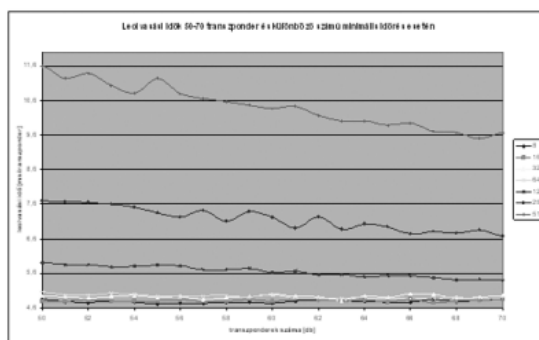
9. ábra
A leolvasási idők
alakulása a transz-
ponderek számának
függvényében a
minimális időresek
száma szerint, hang-
súlyos végértékkel



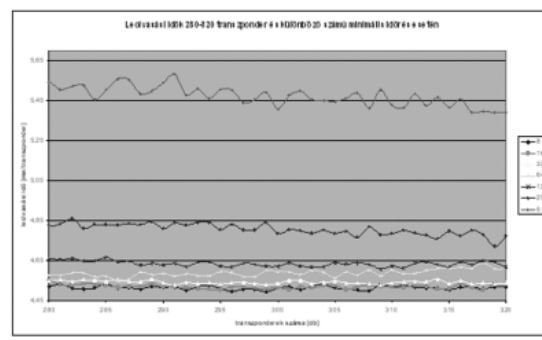
10. ábra
A leolvasási idők
alakulása 5-15
transzponderre
(minimálisan
kiosztandó időresek
száma szerint)



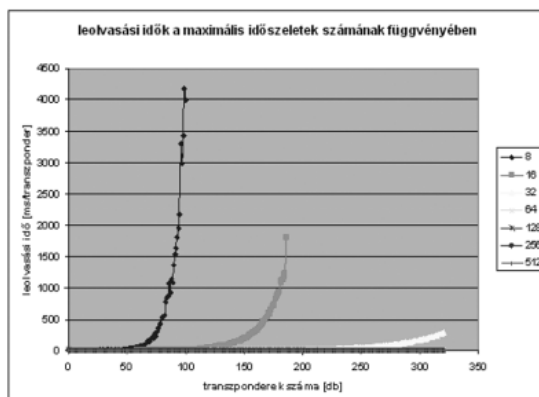
11. ábra
A leolvasási idők
alakulása 50-70
transzponderre
(minimálisan
kiosztandó időresek
száma szerint)



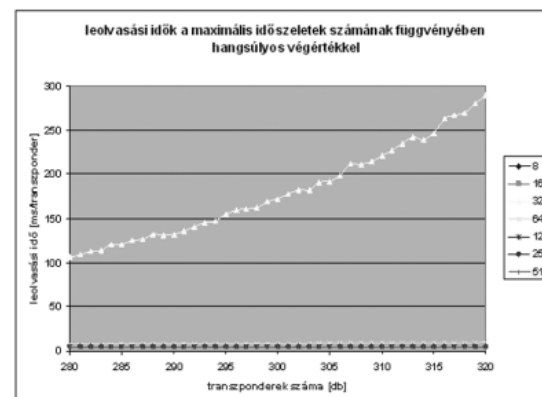
12. ábra
A leolvasási idők
alakulása 280-320
transzponderre
(minimálisan
kiosztandó időresek
száma szerint)



13. ábra
A leolvasási idők
alakulása a transz-
ponderek számának
függvényében a
maximális időresek
száma szerint



14. ábra
A leolvasási idők
alakulása a transz-
ponderek számának
függvényében a
maximális időresek
száma szerint
hangsúlyos
végértékkel



A következő ábrák (9-12.) a minimális időrések számának hatásait illusztrálják, annak különböző értékei esetén. Szintén 320 transzponder adatainak átlagából készültek a grafikonok. A vizsgált tartományban nagy meglepetés nincs: több időrés rosszabbul teljesít, a leggyorsabb leolvasást a legkevesebb időrés használata adja. Figyelem: ez csak egy alsó korlát az adaptív algoritmus számára. Az eljárás maga dönt arról, hogy a keretek között milyen számú időrést választ. Nem meglepő tehát, hogy a legjobb eredmény akkor adódik, ha az algoritmus szabad kezet kap.

A maximálisan kiosztandó időrések számát vizsgáló szimuláció hasonló az előzőhöz, megvizsgálja mi történik, ha az időrések számát meghatározó algoritmus korlátok közé szorul.

Várható, hogy igen leromlik a teljesítmény kis számú maximális időrés használatával, nagy számú korlát esetén viszont az algoritmus maga dönt a jó választásról. A szimuláció eredményei a 13-14. ábrán láthatók.

A grafikonok igazolják a feltételezést: 8, valamint 16 időrés használata 50, illetve 150 transzponder felett gyakorlatilag soha véget nem érő leolvasást eredményez. A minimális leolvasási időket több száz transzponder esetén a 64, 128, 256, valamint 512-es beállítások adják, az ennél kevesebb időrést használó tesztekben az idők exponenciális jelleggel nőnek a transzponderek számának növelésével. Az algoritmus tehát optimálisan működik alaphelyzetben, ha tág határok közt tartjuk a működési területét, ezzel elérve a saját maga általi optimális választást, mely a leghatékonyabb működést eredményezi.

A szimuláció eredményeiből világosan látszik, hogy nagyban befolyásolja a kiosztott időrések száma, illetve a rendszerben lévő transzponderek száma a leolvasási időt. Nyilvánvalóan minél több a transzponder, annál nagyobb az ütközés valószínűsége, és így növekszik a leolvasási idő, illetve minél több időrést oszt ki az algoritmus az egyes transzpondereknek, annál többször lesz kihasználatlan a csatorna, illetve megtalálható az az optimális időrés populációnként, ahogyan a legjobb hatásfokkal futhat az algoritmus.

Fontos megjegyezni, hogy a szoftver működését a Gemplus és a TagSys mérnökei tesztelték, eredményeket saját protokolljaikban is felhasználják. A szimuláció tényleges rendszerépítési- és méretezési kérdésekre is választ ad, ugyanakkor általánosságában is tárgyalja a felmerülő problémakört.

5. ISO és EPCglobal szabványok az RFID területén

Végül egy nagyon rövid, nem átfogó leírást adunk az RFID szabványokkal kapcsolatban.

A két legjelentősebb ellátási láncokkal kapcsolatos szabvány kidolgozója az ISO (International Standardization Organization) és az EPCglobal Inc. A legtöbb nemzeti és ipari szabvány ezekre épül. Az ISO 18000-es számú szabványa foglalkozik. Az EPC szabványok ne-

vüket az Electronic Product Code-ről, azaz az elektronikus termékkódról kapták, mellyel minden terméket egyedileg lehet azonosítani, az EPC eszközöket pedig osztályokba és generációkba sorolták: A Class1 besorolás passzív, csak olvasható, úgynevezett backscatter tag-eket takar egyszer írható, nem felejtő memóriával. A Gen2 generáció pedig a 860-960 MHz-es működési sávra és a 96-256 bites azonosítóformátumra utal.

6. Összefoglalás

A bemutatott ütközésfeloldó algoritmusok jó kiindulási alapként szolgálnak a jövőben még gyorsabb és hatékonyabb többes leolvasási algoritmusok fejlesztésére. Természetesen alkalmazási környezettől függ, hogy milyen algoritmust célszerű használni a speciális igényeknek megfelelően. Az azonban elengedhetetlenül fontos, hogy megfeleljenek a kidolgozott eljárások valamelyik szabványnak, hiszen csak így biztosítható a széleskörű felhasználhatóság. Mivel az EPCglobal Gen2-es szabványát átadták az ISO-nak, hogy az a 18000-es szabványcsalád része lehessen, célszerű ennek alapján kidolgozni, fejleszteni, esetleg javítani a meglévő ütközésfeloldó algoritmusokat.

Ha kellően megbízható és jó teljesítőképességű algoritmus készül el, akkor azt a későbbiekben széles körben lehet használni kezdve egy hipermarketben történő vásárlást követő fizetéstől egy raktárbázis tartalmának lekérdezéséig. Ehhez nagyon jó eszközként szolgálnak a külföldre szimulációs szoftverek, melyekkel akár teljes rendszerek megtervezése is lehetséges. A cél minden RFID-vel kapcsolatos fejlesztés során az, hogy mindenki által használható, együttműködő rendszerek jöjjenek létre, várhatóan tehát ez fogja vezérelni a jövőben a többes leolvasási fejlesztéseket is.

A kutatást a Mobil Innovációs Laboratóriumban végeztük a Nemzeti Kutatási és Technológiai Hivatal Jedlik Ányos programjában a rádiófrekvenciás azonosítás ipari továbbfejlesztési lehetőségeinek témájában.

Irodalom

- [1] K. Finkenzeller, RFID Handbook, Swadlincote: Wiley & Sons Ltd. 2003.
- [2] S. M. Birari, Mitigating the Reader Collision Problem in RFID Networks with Mobile Readers (Február 2006) [Online]. <http://www.it.iitb.ac.in/~shailesh/ThirdStage.ppt>
- [3] S. W. Lee, „A Multiple Access Algorithm for Passive RFID tags,” Thesis, School of Electrical and Electronic Engineering, College of Engineering, Yonsei University, Seoul, South Korea, 2005.
- [4] BME-HIT, RFID kutatócsoport: RFID rendszerek felhasználása, <http://rfid.answare.hu>

Mobilitás-menedzsment algoritmusok numerikus vizsgálata

KOVÁCS BENEDEK, FÜLÖP PÉTER

Budapesti Műszaki és Gazdaságtudományi Egyetem, Híradástechnikai Tanszék
{bence, fepti}@mcl.hu

Lektorált

Kulcsszavak: mobilitás-menedzsment, modellezés, handover

Az értekezésben különböző mobilitási stratégiákat tanulmányozunk a hálózati jelzésforgalom (signaling), az egyes csomópontokban fellépő feldolgozás (processing) és a rádiós interfész (air) költsége alapján. Bevezetünk egy eljárást a kiszolgáló hálózat és a vezeték nélküli csomópont mobilitásának modellezésére, meghatározunk paramétereket a topológiából hogy a vizsgált hálózatot megfelelően leírassuk. Különböző csoportokat alakítunk ki létező és elméletileg lehetséges protokollok alapján, amiket összehasonlítunk különböző mobilitási és hálózati környezetekben. A végső cél egy olyan irányelv létrehozása amely segítséget nyújt újgenerációs, mobil hálózati menedzsmentrendszer megtervezésében.

1. Bevezetés

Az információ mobilitása a modern világ egyik legfontosabb szolgáltatása, köszönhetően elsősorban a sokféle, egyre kedvezőbb mobil hozzáférési technológiának, és az olcsó, többféle vezeték nélküli interfésszel rendelkező „okos” telefonok elterjedésének. Mi sem bizonyítja jobban az információs társadalom „minden időben és helyen bármilyen információt” tendenciáját, hogy szakmai fórumokon rendszeresen jelennek meg az új mobilitás menedzsment protokollok, vagy éppen a régiek módosított változatai.

A mobil szolgáltatásoknak sok különböző követelménye van, és általában a kiszolgáló hálózatban az erőforrás véges és drága. Az első vezeték nélküli protokollok legfőbb célja a jól működő mobilitás megvalósítása volt, és nem helyeztek nagy hangsúlyt a vezetékes kiszolgáló hálózat jelzésforgalmának optimalizálására. Ennek jó példája a világszerte működő Global System for Mobile Telecommunications (GSM) rendszer. Idővel az IP protokoll került előtérbe a távközlésben és így a mobil távközlésben is a Mobile IP (MIP) megjelenésével. Ebből, a hierarchikus szemlélet vagy éppen a location tracking bevezetésével, a MIP különböző, bizonyos szempontok szerint optimálisabb kiterjesztései fejlődtek ki. Kialakultak más mobilitás kezelő megoldások is, például hálózati rétegben a Host Identity Protocol (HIP) vagy szállítási rétegben a Multimedia Stream Transmission Control Protocol (mSCTP). Ezek a megoldások más módon támogatják a mobil felhasználó mozgását, de mégis általános szemlélettel tekintve összehasonlíthatóak és vannak közös vonásaik.

A jövő mobilitás protokolljai talán más-más algoritmusokat alkalmaznak majd eltérő hordozó technológiákon. Ezért, és a már most megtalálható protokollok sokfélesége miatt, az összehasonlíthatóság érdekében, a mobilitást egy absztrakt problémaként kezeljük, függetlenül a technológiától. Ennek megfelelően próbáltuk összehasonlítani munkánk során a mobilitás megoldá-

sokat, megmutatni hogy az egyes hálózati paraméterek hogyan hatnak a protokollok teljesítményére.

Célunk, hogy különböző helyzetekre megtaláljuk a legmegfelelőbb mobilitás megoldást, hogy irányelvet adjunk, melyek követésével egy adott protokollhoz megadható a leoptimálisabb hálózat, vagy éppen egy adott hálózathoz megtalálható a leoptimálisabb protokoll.

2. A mobilitás menedzsment, mint absztrakt probléma

A modellezés során a mobilitást általánosan kezeljük, a pontos, alkalmazott technológiára való tekintet nélkül. Megmutatjuk hogy az egyes létező és lehetséges protokollok hogyan illeszthetők az általunk definiált stratégiák közé. A mobilitás legjellemzőbb tulajdonságait próbáljuk megragadni, és ennek tükrében definiálunk egy *Mobilitás Menedzsment Rendszert*. Ennek elemeit egyértelműen meghatározzuk, definiáljuk a rajtuk futó alkalmazást, mely képes lesz meghatározni egy mobil állomás helyzetét tetszőleges időpillanatban.

A hálózati elemeink a következők:

- *Mobil állomás (Mobile Node, MN)*: mozgó csomópont, amely más mobil vagy fix állomással kommunikál.
- *Mobil hozzáférési pont (Mobile Access Point, MAP)*: az egyetlen olyan típusú, fix entitás a hálózatban, amelyhez a mobil állomás képes kapcsolódni, és rajta keresztül kommunikálni. Szükségszerűen fut rajta az adott mobilitás menedzsment algoritmus. (Nem feltétlenül felel meg ez egy hozzáférési pontnak a valós hálózatban, hanem akár hálózatot, vagy hozzáférési pontok halmazát is jelentheti abban az esetben például ha a mobil egyidejűleg képes több bázisállomással is kapcsolatban állni).
- *Mobil ügynök (Mobile Agent, MA)*: fix csomópont, melyen fut az adott mobilitás menedzsment algoritmus, de mobil nem tud csatlakozni hozzá.

- Egyéb csomópont (Node, N): olyan csomópont a hálózatban, amely nem rendelkezik mobilitás menedzsment szemszögéből semmilyen funkcióval. Ide tartoznak a fix kommunikációs partnerek, routerek és egyéb hálózati elemek.

Ezzel az egyszerű leírással az alapvető funkciók leírhatók, azonban néhány praktikus megkötést még bevezetünk:

- A mobil hozzáférési pont (MAP) minden esetben mobil ügynök (MA) is. (Ahogy már említettük a MAP jelenthet egy alhálózatot is magában foglalva több hozzáférési pontot, és kommunikációs partnereket. Ez teszi lehetővé hogy szükség szerint mikromobilitást vagy éppen magasabb szintű mobilitás modellezésre alkalmazzuk a módszerünket).
- A mobil entitás (MN) legalább egygyel, de akár több MAP-al is kapcsolatba kerülhet. Helyzetéről releváns információt mindig egy ügynök tárol, ezért a feladat hogy megtaláljuk a mobilt, ekvivalens azzal hogy megtaláljuk a hozzátartozó MAP-ot.
- Feltételezzük hogy a csomópontok a hálózatban mindig használnak egy adott metódust, vagy protokollt (például Internet protokoll), amin keresztül kommunikálnak és megtalálják egymást címzés segítségével, ezért ezzel nem foglalkozunk.

Ezzel a mobilitás kezelést MAP (hozzáférési pont) megtalálásának feladatává egyszerűsítettük. Ezen megvizsgálva a különböző menedzsment stratégiákat a jelzés és feldolgozási költségek szamszerűsíthetőek, és egymással összehasonlíthatóak.

A stratégiákat, a *Centralizált*, *Hierarchikus*, *Cellás* és *Tracking* csoportokba osztjuk be. Ennek részletezése az 4. részben olvasható.

Az egyedi vagy más konkrét protokollok modellünkbe történő integrálását az olvasóra bizzuk.

3. A hálózat és a mobil entitás paraméterei

Ebben a fejezetben a mobilitási stratégiákat megvalósító alaphálózat modelljét mutatjuk be. Ahhoz, hogy levezessük a fő paramétereket, a mobil viselkedését kell megfelelően modellezni. Meghatározunk általános és algoritmus specifikus tényezőket, melyeket hálózati struktúrából és a mozgási tulajdonságokból vezetünk. Bevezetjük a három, különböző költség-dimenziót, amit kiszámítunk az egyes stratégia csoportokra, a „jelzés-költséget” (C_{signal}), „processzási költség” ($C_{process}$) és a „rádiós interfész költség” (C_{air}).

3.1. Hálózat modellezése

A hálózat modellezés egyik lehetséges megvalósítása, amikor egy protokoll jellemzéséhez egy konkrét, specifikus hálózat reprezentációt használnak fel. Ennek nyilvánvaló hátránya, hogy más protokollnál, más kontextusban nem használható fel. Emiatt sok érdekezésben a hálózatot egy egyszerű, általános paraméter-

rel írják le, mint például a csomópontok között átlagos távolság. Ezzel bármilyen hálózat leírható, azonban a modell ereje gyenge, ha különböző protokollok viselkedését szeretnénk összehasonlítani az egyes hálózati struktúrákon.

Összegezve a tapasztalatokat, megpróbáljuk ötvözni a két megoldást, megadunk általános jellemzőket, illetve bevezetünk speciális, adott protokoll korrekt leírásához szükséges hálózati paramétereket is.

3.2. Hálózat paramétereinek meghatározása

A mobil entitás két MAP közötti handover frekvenciáját Poisson-folyamat modellezi. Jelölje $B_Q = [b_{ij}]$ a mobil tulajdonságát leíró folytonos Markov lánc $n \times n$ -es ráta mátrixát (más munkákban is jellemző ez a modellezés, például [5]), ahol b_{ij} jelöli a MAP_i és MAP_j közötti átlépési frekvenciát. A Markov lánc n darab csomópontja mutatja a lehetséges MA-kat és MAP-eket. Ahogy már korábban is említettük, a MA általában egy hozzáférési pontot jelöl, de jelölhet egy otthoni ügynököt, vagy akár egy alhálózatot is. A ráta mátrixból levezethető az átlépési B_{Π} matrix, ami MA-hoz tartozó null sorok és oszlopok nélkül irreducibilis és aperiodikus, tehát a mátrix stabil és létezik a határeloszlása, amit jelöljön $\underline{b}$. Ennek elemei határozzák meg azt, hogy a mobil milyen valószínűséggel található meg az egyes MAP-oknál.

A hálózat topológiai felépítését, MA-k közötti kapcsolatokat $n \times n$ A mátrix reprezentálja. Floyd algoritmus felhasználásával a node-ok közötti optimális távolságok kiszámíthatóak, aminek eredményét tároljuk az A_d mátrixban. Ennek a mátrixnak az i . sora jelöli csomópontok távolságát az i . mobilitás ügynöktől (MA). A hazai ügynöktől, mint egy speciális MA-tól számított távolságokat az A_d mátrixból kiemelve az $\underline{a}$ vektor tartalmazza.

Végül vezessük be a w paramétert a hálózat átlagos súlyának jelölésére, amit az A_d mátrix elemeinek összegét elosztva n^2 -tel kapunk meg.

3.2.1. Az m paraméter

Az m paraméter jelöli a mobil átlagos távolságát, az otthoni ügynöktől, azaz a hálózati topológia gráfban az élék átlagos számát az MN és a HA között. Nyilvánvaló hogy a csomópontok átlagos száma ezen az úton $m+1$. Ennek kiszámításához A_d mátrixot, illetve a HA-ra vonatkozó speciális sorát az $\underline{a}$ vektort használjuk fel, majd normalizálunk a korábban bevezetett w értékkel. A súlyozást a Markov-lánc határeloszlásával végezzük el. Így az m a következőképpen alakul:

$$m = \frac{\underline{a} * \underline{b}}{w} \quad (1)$$

A * skaláris szorzatot jelöl. Azok a csomópontok akik nem MAP-ok, a határeloszlásban 0 értékkel szerepelnek, azaz az átlagos távolság kiszámítása során nem kapnak szerepet.

3.2.2. A g_T paraméter

Az m -hez hasonló érték meghatározásának szükségessége is felmerül bizonyos stratégiák esetében, ami azon két MA vagy MAP csomópont átlagos távolsága,

amelyek egy adott mobil handover folyamata során szerepet játszanak. Ezek lehetnek közvetlenül kapcsolatban, de akár nagyon messze is eshetnek egymástól logikailag, mint például vertikális handover esetében. Tehát két tetszőleges MA, vagy MAP átlagos távolságát, amit g_T -vel jelölünk a *Tracking* típusú mobilitáskezelésben betöltött fontossága miatt, a következőképpen számíthatjuk:

$$g_T = \frac{b * tr(A_d \cdot B_H)}{w} \quad (2)$$

A képletben a $tr()$ függvény egy mátrix átlóelemeiből képzett vektort adja eredményül.

3.2.3. A g_H paraméter

Hierarchikus mobilitáskezelés esetében fontos szerepet játszik, hogy a mobil csomópont átlagosan milyen messze esik a legközelebbi hierarchikus „keresztződéstől”, ahol beregisztrálhatja elérhetőségét. Természetesen erről csak akkor beszélhetünk, ha a hálózati gráfot lehetséges lefedni egy olyan fával aminek csúcsában az otthoni ügynök helyezkedik el. A keresett csomópontot az otthoni ügynöktől a mobil régi és új csatlakozási helyéhez vezető út keresztződése determinálja. A g_H paraméter meghatározásának részletes levezetése a [6]-ban olvasható.

3.2.4. A g_C paraméter

A g_C érték speciálisan a *Cellás* típusú megoldásokban jelenik meg. Jelentése egy jól definiált, úgynevezett „paging” területen megtalálható MAP-ok és a terület felett elhelyezkedő gateway MA közötti átlagos távolság. Az optimális paging lefedettség megtalálása NP teljes probléma, azonban az irodalomban megtalálható jónéhány hatékony közelítést nyújtó algoritmus, melyek közül a [12]-t használtuk fel a g_C meghatározására.

3.3. Mobil csomópont modellezése

A mobil viselkedésének egyszerű modellezésére használjuk a ρ -t, ahogyan a [6]-ban, ami az úgynevezett mobilitási viszony, azaz annak a valószínűsége hogy a mobil handovert hajt végre mielőtt egy „hívás” érkezne hozzá. Az érték determinálható a B_Q matrix alapján, amiben a legkülönbözőbb mozgási viselkedések is leírhatók. Ha ennek a mátrixnak az i . sorát összegezzük, megkapjuk hogy a mobil milyen frekvenciával lép el az i . MAP-tól. Ha minden sorra elvégezzük ezt, és elosztjuk n -nel, akkor megkapjuk az átlagos frekvenciát, amit jelölünk λ -val.

A hívás frekvenciáját determinálja μ , ami erősen függ az adott időszaktól, ezért optimálisan, a λ -hoz hasonlóan, az átlagos értéket számoljuk belőle is. Ezek után a ρ meghatározása:

$$\rho = \frac{\lambda}{\lambda + \mu} \quad (3)$$

3.4. Konstans költségek determinálása

Ebben a fejezetben a konstans költségek három nagy csoportját és tagjait vezetjük be. Egyes költségfajták számszerűsítését, változtatását körültekintően kell végrehajtani, mivel ahogyan az a 4. részben látható

lesz, erősen befolyásolhatja a protokollok költségeinek meghatározását és ezzel az összehasonlítás végeredményét is. Ezek meghatározására létező protokollok szimulációs változatait készítettük el amelynek részleteit az [1] tartalmazza.

3.4.1. Vezetékes átviteli közeghez kötődő költségek

- c_u : a helyzetinformáció frissítésének egy linkre eső költsége (update).
- c_d : hívás esetén a mobil felé történő adatforgalmazás egy linkre eső költsége (delivery).

3.4.2. Csomópontokhoz kötődő költségek

- c_r : regisztráció során felmerülő processzási költség az adott MAP-nál, amikor egy MN csatlakozik hozzá (adatbázis írása, IP cím kiosztása stb.)
- c_f : csomag továbbításának költsége egy tetszőleges csomópontnál a hálózatban.
- c_m : MN-hez tartozó bejegyzés módosításának költsége egy MA/MAP-nál.
- c_{ec} : Tunneling esetén az IP becsomagolás processzási költsége MA/MAP-nál (általánosságban jelentheti egy szabványos IP csomag eljuttatásának processzási költségét a mobil állomás felé, ha a csomag cílcíme nem egyezik a mobil aktuális elérhetőségével).
- c_{rc} : Tunneling esetén az IP újracsomagolás processzási költsége MA/MAP-nál.
- c_{dc} : Tunneling esetén az IP kicsomagolás processzási költsége MA/MAP-nál.

3.4.3. Rádiós interfészhez kötődő költségek

- c_{au} : Az uplink irányú üzenet egységnyi költsége az MN és a MAP között.
- c_{ad} : A downlink irányú üzenet egységnyi költsége az MN és a MAP között.

4. Mobilitás kezelési stratégiák

Ebben a fejezetben a négy nagy mobilitás menedzsment stratégiát mutatjuk be, leírjuk alapvető jellemzőiket és bevezetjük a különböző költségfüggvényeiket. Részletesebben a csoportok leírása és költségfüggvények az [1] cikkben találhatóak meg.

4.1. Centralizált

E technológia esetében a mobil állomás minden handovert követően a helyzetinformációját elküldi egy központi management node-nak vagy node csoportnak, aki egy adatbázist tart fenn, amiben tárolja a mobil pontos elérhetőségét. Mivel ez a központi node mindig tudja a mobil állomás helyét, ezért képes továbbítani a mobilnak szóló csomagokat (Mobile IP [10]), vagy éppen meg tudja mondani a mobil állomás elérhetőségét (SIP). (4)

$$C_{CENT}^{signal} = \rho m c_u + (1 - \rho) m c_d$$

$$C_{CENT}^{process} = \rho (c_r + (m - 1) c_f + c_m) + (1 - \rho) (c_{ec} + (m - 2) c_f + c_{dc})$$

$$C_{CENT}^{air} = \rho c_{au} + (1 - \rho) c_{ad}$$

A költségfüggvényei ennek a csoportnak (4. egyenletek) nagyon egyszerűek, emellett a másik előnye a megoldásnak, hogy a hálózatban egy központi ügynök felállításával gyakorlatilag működőképes lehet. Ezek a rendszerek ilyen szempontból egyszerűek, de nagy és felesleges hálózati jelzés overhead-et eredményeznek. Ennek ellenére a legtöbb, létező mobilitási protokoll ebbe a csoportba sorolható.

4.2. Hierarchikus

A központi, globális menedzsmenttel ellentétben itt lokális management node-okat alkalmazunk, amik egy jól definiált területen belül kezelik a mobil mozgását (Hierarchical Mobile IP). Ugyanúgy jelen van egy központi management node is, azonban az ötlet ott rejlik, hogy a helyzetfrissítő információkat csak a központi node felé vezető régi út, és az új út metszésében elhelyezkedő MA/MAP-ig küldi el a mobil (g_C távolságra a mobil állomástól), ennek megfelelően változnak a költségfüggvények is (5. egyenletek):

$$\begin{aligned} C_{HIERARCH}^{signal} &= \rho g_H c_u + (1 - \rho) m c_d \\ C_{HIERARCH}^{process} &= \rho (c_r + (g_H - 1) c_f + c_m) + \\ &\quad + (1 - \rho) (c_{ec} + (m - g_H - 1) c_f + c_{rc} + (g_H - 1) c_f) \\ C_{HIERARCH}^{air} &= \rho c_{au} + (1 - \rho) c_{ad} \end{aligned} \quad (5)$$

A megoldás előnye, hogy nem terheljük az egész hálózatot, azonban bonyolultabb struktúrát eredményez, mint az előző. Legjellemzőbb példája ennek a csoportnak a Hierarchikus Mobil IP (HMIP) [4].

4.3. Cellás

A mobilitáskezelés egyik legrégebbi stratégiái a GSM-ből is jól ismert cellás megoldások. Általában mikromobilitás szintjén alkalmazzák ezt a lehetőséget, de bevezethető magasabb mobilitási hierarchiákban is.

Előnye a gyors handover mechanizmus, hátránya viszont a cellás és cellákat összefogó paging felosztás körültekintő megtervezése, ugyanis suboptimális eset-

ben a handover jelzések különösen nagy, felesleges jelzésforgalmat okozhatnak. Ezek a megoldások hatékonyság szempontjából arra a tényre építenek, hogy a mobilok csak nagyon kis százaléka az aki éppen aktív és csomagot fogad, vagy éppen küld.

Költségfüggvények meghatározásához szükségünk van még a következő hálózati paraméterekre:

n_d : a teljes hálózatban lévő page területek száma.

n_c : egy page területen belül található MAP-ok átlagos száma.

Három altípusát és hozzájuk tartozó költségfüggvényeket mutatjuk be a következő alfejezetekben.

4.3.1. Standard cellás megoldás

A *Standard cellás megoldás* alap gondolata jól ismert Cellular IP (CIP)-t [3] követi. Azaz pontosan definiált page területeket vezetünk be, melyen belül a mobil jelzés nélkül végezhet handover-t, viszont page terület átlépésénél minden esetben frissíti a helyzetinformációját. Ennek következményeként a hálózat a mobilnak szóló csomagot teljes biztonsággal csak a page terület gateway MA-jáig tudja eljuttatni, innen vagy ismert a pontos helyzete a mobilnak, vagy elárasztást kell alkalmazni. A költségfüggvények ezek tekintetében a következők (6):

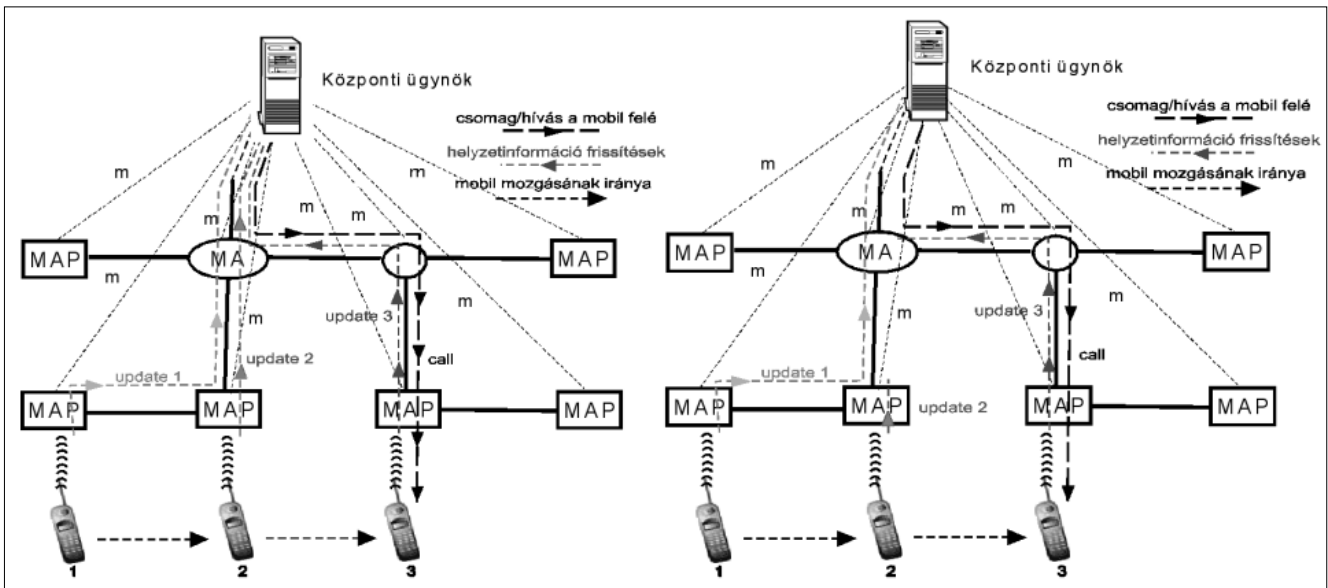
$$\begin{aligned} C_{CELLULAR}^{signal} &= \rho (1 - P_c) g_H c_u + (1 - \rho) ((m - g_C) + (n_c g_C c_d) + g_C c_u) \\ C_{CELLULAR}^{process} &= \rho ((1 - P_c) c_r + g_H c_f + c_m) + \\ &\quad + (1 - \rho) (c_{ec} + (m - g_C - 1) c_f + c_{rc} + (g_C - 1) n_c c_f + c_{dc} c_r) \\ C_{CELLULAR}^{air} &= \rho ((1 - P_c) c_{au}) + (1 - \rho) (n_c c_{ad} + c_{au}) \end{aligned}$$

ahol a P_c annak a valószínűsége hogy a mobil átlép egy másik page-területre.

4.3.2. Hierarchikus paging

A hierarchikus paging [8] csak annyiban különbözik az előzőekben ismertetett standard cellás megoldástól, hogy az alsó szinten használt elárasztást a felső hierarchia szinten is alkalmazza a page gateway MA meg-

1. ábra Centrális és Hierarchikus stratégia



találására. Ezzel jelzéseköltséget spórolhatunk, mivel a helyzetinformációt nem küldjük el egészen a HA-ig, hanem csak a gateway MA-ig, azonban a multilevel elárasztás nagy hálózati terhelést okozhat. Az előzőhöz nagyon hasonló költségfüggvényt a 7. egyenletek mutatják:

$$C_{HP}^{signal} = \rho(1-P_c)g_Cc_u + (1-\rho)((m-g_C) + (n_Cg_Cc_d) + g_Cc_u)$$

$$C_{HP}^{process} = \rho((1-P_c)c_r + g_Cc_f + c_m) +$$

$$+ (1-\rho)(c_{ec} + (m-g_C-1)n_d c_f + c_{rc} + (g_C-1)n_Cc_f + c_{dc}c_r)$$

$$C_{HP}^{air} = \rho((1-P_c)c_{au}) + (1-\rho)(n_Cc_{ad} + c_{au})$$

4.3.3. MANET jellegű

Abban az esetben ha az infrastruktúra kiépítése nehéz, drága vagy éppen felesleges lenne, és a rádiós interfész nagyobb terhelése nem okoz problémát, akkor a MANET (Mobile Ad-hoc Network) [9] mobilitás page-en belüli használata a legjobb megoldás. Alkalmazásánál alapvető feltevés, hogy minden mobil állomás elérhető más mobil állomásokon keresztül, ad-hoc hálózat használatával.

Ebben a menedzsment rendszerben ugyanúgy paging területeket definiálunk, azonban ezekben csak egy MAP található amin keresztül egy optimális MANET [9] algoritmus használatával elérhetőek a mobil állomások, és jelzéseköltségek is csökkenthetők (2. ábra).

A költségfüggvények a következők (8):

$$C_{MANET}^{signal} = \rho(1-P_c)g_Cc_u + (1-\rho)((m-g_C+1) + (P_Mn_Cg_Cc_d) + g_Cc_u)$$

$$C_{MANET}^{process} = \rho((1-P_c)c_r + g_Hc_f + c_m) +$$

$$+ (1-\rho)(c_{ec} + (m-g_C-1)c_f + c_{rc} + P_Mn_Cg_Cc_f + c_{dc}c_r)$$

$$C_{MANET}^{air} = \rho((1-P_c)(g_C-1)c_{au}) + (1-\rho)(P_Mn_Cg_Cc_{ad} + c_{au})$$

ahol P_M annak a valószínűsége hogy a page-en belül minden mobilt érint a csomag továbbítása.

4.4. Tracking típusú

A tracking megoldásokban minden mobilnak ugyanúgy van egy bejegyzése a központi ügynöknél, mint a

korábbi megoldásoknál. A központi állomás annak a MAP-nak a címét tárolja, ahonnan legutóbb kapott helyzetinformáció frissítést a mobiltól. Amikor kérés érkezik a mobilhoz, akkor a központi ügynök az eltárolt MAP-felé továbbítja azt. A mobil állomás vagy még mindig ott tartózkodik, vagy a MAP ismer egy másik MAP-ot, ahol megtalálható lehet a mobil. Így a csomagokat az egyes MAP-ok igyekeznek a mobil állomás után küldeni, míg végül a lánc végén eléri a mobilt.

Tehát handover végrehajtását követően a mobil az új helyzetét a régi hozzáférési csomóponttal közli és nem minden esetben szól feljebb a hierachiában (tracking handover). Bővebb leírás található a [2,6,11] cikkekben.

4.4.1. Vezetéknélküli tracking

Vezetéknélküli tracking esetében a tracking handover során a visszajelzést a régi hozzáférési pontnak az új MAP-ról a rádiós interfészen végzi a mobil állomás. Ennek megfelelően a költségfüggvény (9):

$$C_{WLESSTR}^{signal} = \rho P_H g_H c_u + (1-\rho)(g_H c_d + M[h_r]g_T c_d + (1-P_0)g_H c_u)$$

$$C_{WLESSTR}^{process} = \rho((1-P_H)(c_r + c_m) + P_H(c_r + (g_H-1)c_f + c_m) +$$

$$+ (1-\rho)(c_{ec} + (m-1)c_f + P_0 c_{dc} +$$

$$(1-P_0)M[h_r]((g_T-1)c_f + c_{rc}) + c_{dc} + (g_H-1)c_f + c_m))$$

$$C_{WLESSTR}^{air} = \rho c_{au} + (1-\rho)c_{ad}$$

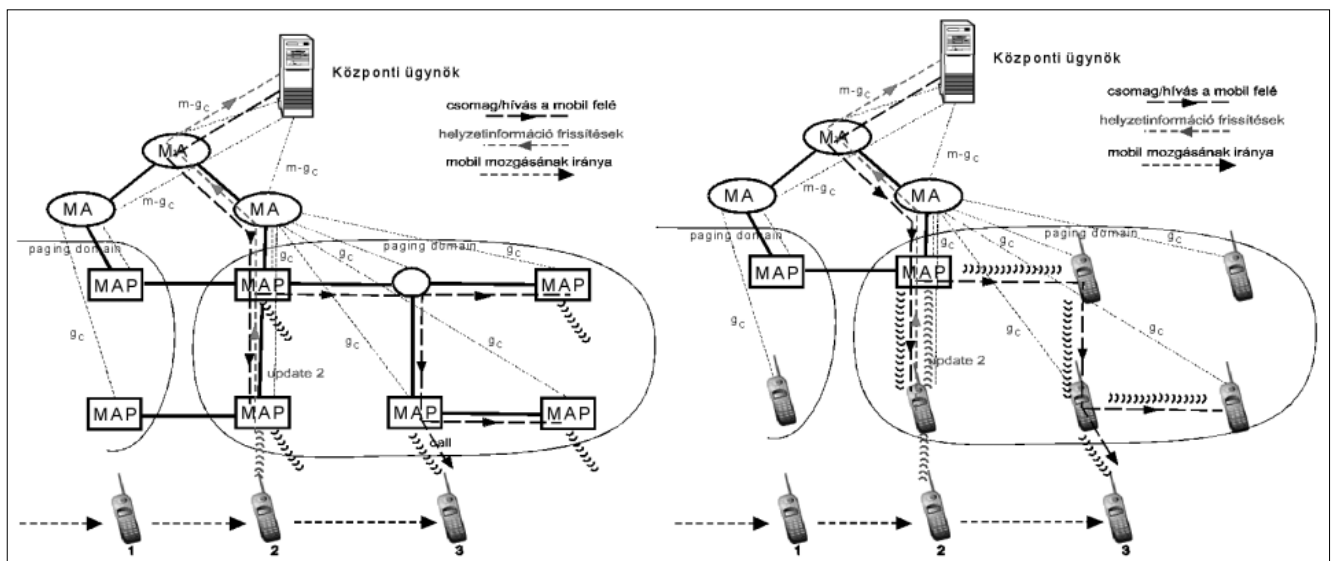
ahol a $M[h_r]$ normál handover követően mobil által végrehajtott tracking handoverek száma, P_H annak valószínűsége hogy éppen az H. állapotban vagyunk a LTRACK Markov modell szerint [6] azaz normál handovernek kell következnie.

4.4.2. Vezetékes tracking

A rádiós interfész kímélése érdekében a tracking handover követő visszajelzéshez a vezetékes hálózatot is igénybeveheti a mobil csomópont, ennek példája a Vezetékes tracking (10):

$$C_{WTRACK}^{signal} = \rho(g_T(1-P_H) + g_H P_H)c_u + (1-\rho)(m c_d + M[h_r]g_T c_d + (1-P_0)g_H c_u)$$

2. ábra Standard cellás és MANET megoldás



$$C_{WTRACK}^{process} = \rho(c_r + (g_T - 1)c_f + c_m) + (1 - \rho)(c_{ec} + (m - 1)c_f + P_0 c_{dc} + (1 - P_0)M[h_r]((g_T - 1)c_f + c_{re}) + c_{dc} + (g_H - 1)c_f + c_m))$$

$$C_{WTRACK}^{air} = \rho c_{an} + (1 - \rho)c_{ad}$$

5. Numerikus eredmények

A fent bemutatott modellezési környezetben implementált protokollok költségeit többféle konkrét hálózaton és hálózati valamint mobilitási paraméterekkel vizsgálhatjuk meg. Ízelítőnek szeretnénk bemutatni néhányat.

Az első két ábra (3. ábra) a mobilitási paraméter függvényében mutatja be az egyes protokollok költségfüggvényeit. Látható, ahogyan hagyományos MIPv4 megoldás költsége lineárisan növekszik, és magas mobilitási érték esetén meghaladja minden más megközelítést.

Külön kiemelnénk a hierarchikus mobil IP megoldást mely költsége ebben az esetben mindig a MIPv4 görbéje alatt marad. Modellünkben ez mindig igaz lesz, amennyiben az utóbbi megoldás tunelling-jellegű költségei nem lesznek túl nagyok. Érdekes megfigyelni, hogy a Tracking típusú megoldások kis mobilitásra hasonló költségűek, mint a hierarchikus (ez azzal magya-

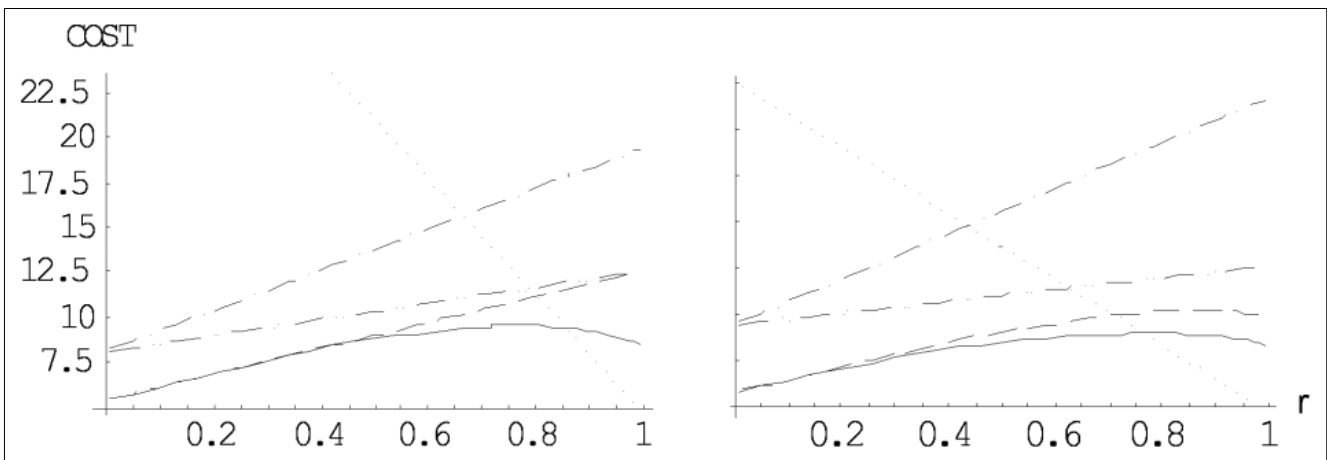
rázható, hogy azonos költségállandók esetén, ha az optimális tracking handoverek száma 0, ezek a protokollok identikusak), nagy mobilitás esetén azonban a tracking handoverek optimális számának növekedésével egyre több a megtakarítás a felesleges update-ek elhagyásával.

Egyértelmű, hogy egy mobilitástól független cellamérettel dolgozó cella alapú megoldás esetén a nagy mobilitási tartományban várunk jó teljesítményt. A két ábra közti különbséget a költség-állandók változtatásával kapjuk. A második esetben ugyanis csökkentettük a delivery költséget az update-tel szemben, így a vezetékes tracking megoldás esetén is hamarabb érzékelhető a tracking handoverek jótékony hatása, a cellás megoldás, a paging költség növekedésével pedig egyértelműen magasabb.

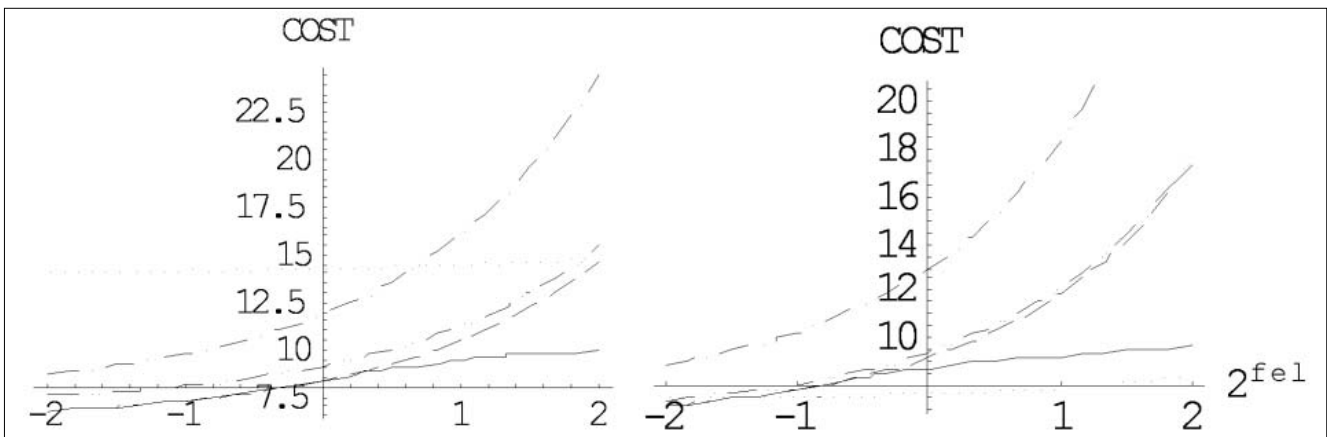
A második két ábrán (4. ábra) az utóbb említett delivery és update költségek arányában ábrázoljuk a protokollok költségeit, logaritmikus skálán a 0.7 és 0.9-es mobilitási szelvényben (az előzőektől különböző hálózati paraméterekkel). Látható az egyértelmű különbség, hogy különböző mértékű mobilitás esetén nem csak az egyes protokollok költségfüggvényei, hanem egymáshoz képesti helyzetük is módosul.

3. ábra

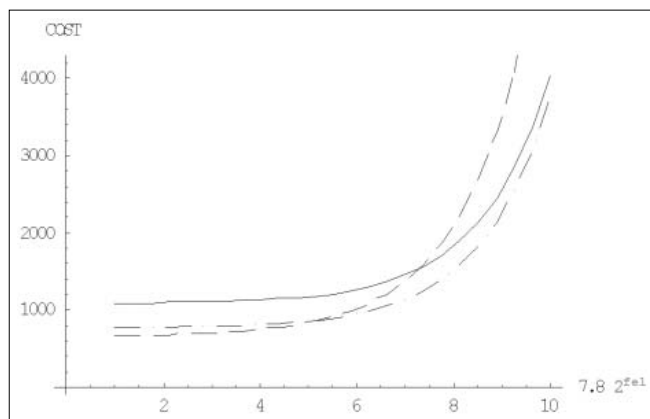
Költségek alakulása mobilitási paraméter függvényében, két különböző hálózaton
(szaggatott egy ponttal – Centralizált; szaggatott két ponttal – Hierarchikus; szaggatott – Vezetéknélküli Tracking; folytonos – Vezetékes Tracking; pontozott – Cellás)



4. ábra Költségek alakulása delivery és update paraméterek függvényében



Az 5. ábra a különböző cellás megoldások költségeinek alakulását mutatja a rádiós interfész költségeinek arányában logaritmikus skálán. A kiindulási érték azért 7.8, mert a szimuláció eredményeképpen a rádiós interfészen a feltöltési költségre ezt az értéket kaptuk. Jól megfigyelhető a függvényeken hogy alacsony költség arányánál a MANET a legoptimálisabb. A Hierarchikus Paging görbéjéhez a Standard Cellás a költségek növekedésével hozzásimul, de soha nem kerül fölé.



5. ábra
Cellás stratégiák költségei a rádiós interfész arányában
(szaggatott vonal – MANET;
szaggatott két ponttal – Standard Cellás;
folytonos – Hierarchikus Paging)

További érdekes számításokat lehetne végezni sokféle képpen, sokféle protokollra. Erre mi egy általunk fejlesztett Mathematica-programcsomagot használtunk, mely tetszés szerint bővíthető.

6. Összefoglalás

A munkákkal rávilágítottunk arra, hogy az IP mobilitás protokollok összehasonlítása sokrétű lehet. Teljesítményük implementáció és hálózathoz kötött, így igazi összehasonlító-elemzést csak speciális esetben végezhetünk. Az általunk bemutatott mobilitás értékelő modell azonban alkalmas arra, hogy megmutassuk, hogy milyen hálózati paraméterek esetén mely megközelítések adhatnak jobb megoldást, ezáltal irányvonalat találhatunk egy környezet és alkalmazás függvényében optimális protokoll megtervezéséhez.

A Mathematica programnyelven, a programcsomaghoz implementált szoftverünk tartalmazza az irodalomban fellelhető legtöbb IP mobilitási megközelítést, de bármikor bővíthető, és új ötletek megvizsgálására alkalmas. Ennek használatával célunk az IP mobilitás további vizsgálata, új megközelítések javaslata.

Irodalom

- [1] Kovács B., Fülöp P., Imre S.: „Study on mobility management modelling methods”, MoMM 2006, Yogyakarta, Indonesia, 2006.
- [2] Abondo, C., Pierre, S.: „Dynamic Location And Forwarding Pointers for Mobility Management”, Mobile Information Systems, IOS Press, 2005. pp.3–24.
- [3] A. T. Campbell, J. Gomez, A. G. Valkó: „An Overview of Cellular IP”, IEEE, 1999. pp.29–34.
- [4] C. Castelluccia, „A Hierarchical Mobile IP Proposal”, Inria Technical Report, 1998.
- [5] Y. Fang and Y. Lin, „Portable Movement Modelling for PCS Networks”, IEEE Transactions on Vehicular Technology, 2000. pp.1356–1362.
- [6] Kovács B., Szalay M., Imre S.: „Modelling and Quantitative Analysis of LTRACK – a Novel Mobility Management Algorithm”, Mobile Information Systems, Vol. 2, Nr.1, 2006. pp.21–50.
- [7] W. Ma, Y. Fang: „Dynamic Hierarchical Mobility Management Strategy for Mobile IP Networks”, IEEE Journal of Selected Areas In Com., 2004.
- [8] Szalay M., Imre S.: „Hierarchical Paging – A novel location management algorithm”, ICLAN'2006 International Conference on Late Advances in Networks, 6-8 December 2006. Paris, France.
- [9] Ashwini K. Pandey, Hiroshi Fujinoki: „Study of MANET routing protocols by GloMoSim simulator”, International Journal of Network Management, 2005.
- [10] C.E. Perkins, „Mobile IP”, IEEE Communications Magazine, 1997.
- [11] R. Ramjee, T. La Porta, S. Thuel, K. Varadhan, L. Salgarelli, „A Hierarchical Mobile IP Proposal”, Inria Technical Report, 1998.
- [12] Simon V., Imre S.: Location Area Design Algorithms for Minimizing Signalling Costs in Mobile Networks, Int. Journal of Business Data Communications and Networking (IJBDCN), 2007.
- [13] Wolfram Research Inc.: „Mathematica”, <http://www.wolfram.com/>

S. Kari, Károly, Charles, Karel...

HORVÁTH LÁSZLÓ

Puskás Tivadar Távközlési Technikum
lacibacsi@puskas.hu

Lassan elmúlnak a szép napok: két átrepülés között egy pár perc (óra) félálom... Az első fiam érkezésére emlékeztet az egész. Talán azért, mert míg 1980. májusában a mi Farkas Bercink az űrben szállt meg Kubaszovval, addig a mi Bálintunkat is meghozta a „gólya”. Keservesen boldog a kispapák élete az első néhány napban: két etetés között mosás, vasalás, főzés és tudósítás a nagyik irányába, aztán már ordít is „őfelsége” és kezdődik minden előről... Ugyanilyen feszítettek a kezdő űrvasdász napjai is. Két átrepülés között hang- és képvágás, pályaadat-egyeztetés, interjú interjú hátán és már jön is az újabb átrepülés. Csak egy-egy szendvics... és néhány óra a hálósákban a rádióállomáson. De nézzük a részleteket...

A gondolat

Gschwindt András docens úr, a műegyetemi rádióamatőrök atyja már Farkas Bercinél is megpróbálkozott rádiókapcsolatot teremteni a Szojuz űrhajóval, de akkor azt biztonsági okokból nem engedélyezték.

Sokat változott az azóta csak „fokozódó” világ. Először a MIR űrállomás légénységével lehetett QSO-zni ékes orosz nyelven egy-egy átrepülés 8-10 perce alatt. Később a Nemzetközi Űrállomáson, az ISS-en is engedélyezték ezt, angolul. A lényegi változást az amerikai űrhajósok tömeges megjelenése hozta. Nekik már a kiképzés során rádióamatőr vizsgát kell tenniük, és mindegyiknek kötelező három amerikai iskola diákjaival elbeszélgetni. (Kevés kisdíák szeretne mérnök, fizikus, űrhajós lenni az USA-ban is, és ez már a gazdaság rovására megy.) Kell a természetudományoknak ott a reklám, elkelne itthon is!

Simonyi Károly is megkapta a kötelező három iskoláját és emellett bevállalt még egy magyart is. Ebben az a csodálatos, hogy mindez közös anyanyelvünkön, vagyis magyarul lett lebonyolítva.

Miért a Puskás Technikum lett a partner?

Szak- és amatőr berkekben a válasz triviális! Nézzük a részleteket: Miután Gschwindt Bandi felkérő levele az összes szűrőn átsusszanva eljutott Simonyi Károlyhoz és ő igent mondott, már csak egy jól működő rádióklubbal rendelkező általános iskolát kellett keresni széles e hazában. Mivel nem találtatott, a következő lépés egy középiskola volt. Itt a Gyáli úton van az ország 150 tagot számláló legnépesebb rádióklubja és itt tananyag az űrtávközlés (Zh-t kell belőle írni, méréseket kell mindkinek elvégeznie és képesítő is, vizsgatétel). Most már csak a NASA-nak kellett elfogadnia a vezető operátor és az intézményvezető pedigéjét (ez egy 20-30 oldalnyi kérdőív volt angolul), majd a diákok kérdéseit kérték angolul és magyarul. (Az ő pedigéjük nem kellett!)

Utólag köszönetet kell mondanom mind az amerikai, mind az orosz szervezőknek, mert a fenti biztonsági intézkedések után a nyílt lejtő (downlink) frekvencia mel-



lett kaptunk egy „titkos” (ez nem vicc!!!) felmenő (uplink) frekvenciát. Így a diákokkal megvalósított kapcsolat zavarmentes volt, míg a előző napi kísérlet kevésbé.

A diákok kiválasztása

Nekünk, 50 év feletti villamosmérnököknek, fizikusoknak, természettudósoknak a Simonyi Károly név a Professzor Urat – a családnak Karit – jelenti a BME-n és szerte az országban.

Először az Elméleti villamosságtan, majd a Fizika kultúrtörténete című fantasztikus könyvét ismertük meg. (Az elmúlt héten éppen azon ment a brusz, „hogymeg az első kiadás van meg 1978-ból”, valakinek meg csak a CD!) A fentiek alapján triviális volt, hogy Károlytól csak olyan kérdezhet, aki Karinál tett valamit az asztalra. A Természet Világának van egy esszépályázata a kultúra egységéről. Hárman ennek győzteseiből kerültek ki, két diák a Simonyi Károlyról elnevezett elektrotechnika verseny nyertese volt az elmúlt években, Magyar Dóra pedig 15 éves korában Kazinczy érmet nyert a Szép magyar beszéd versenyen. Az már csak hab a tortán, hogy azt a kérdést tehette fel Simonyi Károlynak, amit kiképzett űrhajós apjának, Magyar Bélának soha: „Hogy érzi magát a súlytalanságban?”



Felkészülés

Az interneten szép, didaktikus és jó szimulátorok vannak az űrhajók, űrállomások, műholdak követésének, vételének megkönnyítésére. Ezekben a Föld egy henger belső palástjára lett kivetítve, majd kiterítették azt. Az öt kontinens a megszokott, csak a Déli-sarok lett nagyon hosszú. Az űrállomások szinuszos pályát írnak le, és a Föld forgása miatt ezek a szinuszok párhuzamosan tolódnak el. A nappal és az éjszaka is kiválóan látszik. Az idő és a hely is tetszőlegesen beállítható. Ezen program segítségével, az adó és vevőkészülékek ismeretében könnyen megválasztható volt az antenna-rendszer is. Alacsony pályán egy vízszintesen forgatható 5 elemes YAGI-t használtunk, amely kb. 10 dB nyereségű, a meteorológiai műholdak vételénél bevált kb. 5 dB nyereségű úgynevezett QH – négybekezdésű helix – antenna üzemelt.

Itt mindenképpen meg kell állnom, és meg kell említeni *Muhari István* (hívójele: HA5CH) amatőrtársunkat, aki napi rutinnal kommunikál szinte tetszőleges űrállomással, illetve műholdas átjátszóval. Ő tanácsolta a fenti antennákat és tanította be a vele együtt hét fős stábot a feladat végrehajtására. Álljanak itt a nevek: *Tóth István* (HA5OJ) vezető operátor, *Rigó István* (HA5AUC) segédoperátor, *Horváth Márk* (PhD hallgató) a folyamatos internetkapcsolat tartója, *Fazekas Ákos* (villamosmérnök) hangtechnikus, *Varsányi Zsolt* (villamosmérnök) operatőr és *Grátz Márk* (Puskás/Kandó) adásrendező.

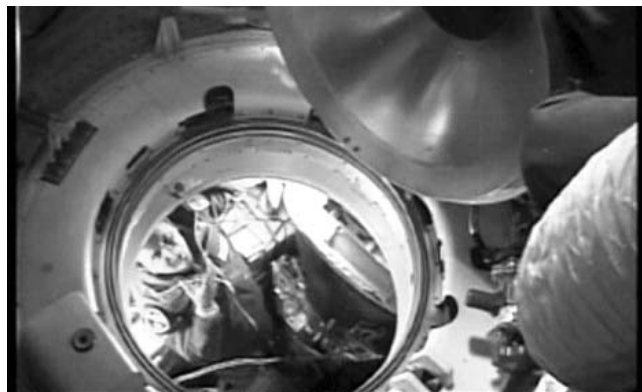
A kísérleti átrepülés

Nehéz leírni azt a várakozással teli feszültséget, amely 2007. április 12-én 0 óra 55 perckor kezdődött és körülbelül 10 perc múlva be is fejeződött. Ez volt a következő napi összeköttetés főpróbája. Nyílt frekvencián 34 magyar amatőr iratkozott fel a Simonyi Károlyhoz eljuttatott listára az űrállomás pályájának megfelelően Soprontól Csongrádig, közöttük ketten a felvidékről, hárman a Vajdaságból. Az űrállomás HA5SIK hívójellel hívta sorba az amatőröket, akik nyugtázták a vételt. A 34 összeköttetés 6-8 perc alatt szép kihívás, már a fele is világsúcs!

A résztvevőkön kívül három TV csatorna stábja és jó néhány fotós, riporter is szorongott a Gyáli úti laboratóriumban. Amikor a szimulátoron az űrállomás Írország fölé ért, elcsendesedett a társaság. A 28.000 km/órás sebességgel száguldó űrállomás Anglia partjainál elérte a 3°-os magasságot, amikor is először angolul megszólalt HA5SIK, az első magyar amatőr az űrből.

Történelmi pillanat volt, majd HA1XY zalaegerszegi amatőrt hívta magyarul a Nemzetközi Űrállomás. Béla barátunkat Budapestről természetesen nem hallottuk, csak az 5SIK nyugtáját: „szervusz Béla, itt Károly beszél QSL”...

És már következett a lista második állomása; Simonyi Károly öreg amatőröket megszégyenítő rutinnal vezényelte le a teljes forgalmat. 20 állomást, vagy nyugtáját mi is hallottuk Budapestről. A hangfelvételek összesítése alapján 25-30 összeköttetés lehetett sikeres. Már a 20 is világrekord!



A stáb nyugodtan vette tudomásul, hogy urala a helyzetnek. Számukra a főpróba tökéletesen sikerült. Talán a sok-sok próbálkozó, tehát a listára be nem jelentkezett hazai és külföldi amatőr okozott csak némi „zajt”.

Az előadás, vagyis amikor a diákok kérdeztek

2007. április 13. 01 óra 15 percre az előző nap tapasztalatai alapján nyugodtan készültünk. Tíz TV stáb egészen jól elfért a laborban, a fotósok szokatlanul viszfogottan vakuzgattak. A mintegy száz fős közönség a nagyelőadóban kivetítőn követte a fejleményeket. Az előző nap eseményeinek felvétele, valamint egy Magyarország feletti átrepülés kommentált levetítése kellően a témára hangolta a nézőket. 0 óra 45 perckor meglepetés-vendékként megjelent *Magyari Béla* kiképzett űrhajós és *Simonyi Tamás* (Károly öccse), aki egyenesen Bajkonurból érkezett. A késői órában a fotósok és riporterek csak lassan eszméltek, habár a Simonyi testvérek hasonlósága igazán feltűnő, meg aztán az űrhajós dzseki a 15. expedíció magyar-orosz-amerikai zászlós jelvényével elegendő volt ahhoz, hogy elszabaduljon a „pokol”. Mindenki felélénkül, fotózott, kérdezett, jegyzetelt. (A kameramanok óvatosságból már a rádióállomáson voltak.)

Az összeköttetés már rutinszerűen zajlott. Kapcsolatfelvétel Anglia partjainál. Kiváló érthetőség, nagy hang-erő, melyből a doppler okozta problémákat az operátor rutinosan „kihangoztatta”. (Ez különben a legmagasabb pályapontján, tehát Magyarország felett volt a legkritikusabb.) A jól felkészített diákok minden félelem nélkül, szépen intonálva tették fel sorban egymás után kérdéseiket, melyekre Simonyi Károly hosszan, érdemben ke-

rek mondatokban, kiváló magyarsággal válaszolt. Kihalt szófűzéséből, hogy hazai középiskolába járt és folyamatosan gyakorolja anyanyelvét.

Talán a legérdekesebb válasz Fekete Soma 24 éves építőmérnök hallgató kérdésére érkezett, aki a Természet Világa pályázatának egyik nyerteseként került a csapatba. A kérdés így szólt: „Melyik könyv és melyik személy volt nagy hatással Önre szakmai pályafutása kezdetén?” A válasz pedig: *„Én Zombok Zoltánt emlétem, aki fiatalon meghalt. Nagyon, nagyon tehetséges mérnök volt, aki engem programozásra tanított még '64-ben és '65-ben. A könyv pedig, ... akkor még elég fiatal voltam, tehát Obádovics József Gyula matematika könyvéből tanultam a matematikát, Számítógépet persze nem lehetett még semmiből sem tanulni.”*

A válaszok elhangzása után a két állomás – a rádió forgalmazás szabályai szerint – nyugtázta az összeköttetést és az amatőr szokásoknál hosszabb búcsúzkodással befejeződött a 6 perc 48 másodperces összeköttetés a legendás Gyáli út és az ISS között, magyar nyelven.

Visszhang

A rádióamatőr mozgalomról az utóbbi 25 évben nem beszéltek annyit, mint ebben az öt napban.

Minden napilap vezető helyen, fotóval hitelesítve hozta az eseményt (lásd a keretes írást).

...és lecsengés

Amikor ezeket a sorokat írom, ma is ugyanúgy repül még felettünk 320 km magasan az ISS. Lassan készül a landolásra... de ez már nem érdekli a médiát. A nap híre: „Nem mi rendezzük a 2012-es foci EB-t a horvátokkal...”, és ez most nagyobb tragédia, mint Isonzó, a Donkanyar, Mohács és Trianon együttvéve...

Köszönöm neked Uram, hogy biztosítottad nekünk egyszerű mérnök-halandóknak ezt a 4-5 balhémentes napot, így jobb híján az újságok címlapjára kerülhettünk.

Mély alázattal a Te szolgálod:

*lacibacsi,
alias a büszke HA5PTL, a Gyáli útról*

Amit senki nem kérdezett meg tőlem...

Ha az utóbbi napokban valaki óvatlanul kinyitotta a „csapot” bárhol az országban, akkor mindenhol a Puskás Technikum, néha Puskás Tivadar Technikum igazgatóját faggatták a Charles Simonyi-féle űrrepülésről. Minden ismertebb csatornán adtam élőben interjút. Reggel, hajnalban és éjszaka is.

Hatalmas élmény volt. Hogy mi?... Nem, nem a rádiókapcsolat az űrállomással... 320 km direkt rálátással a 144 MHz hullámsávban, FM-ben nem egy komoly rádióátviteli probléma. Hanem a riporterek, műsorvezetők, szerkesztők, főszerkesztők (és a csinos sminkesek) kérdései és egy mérnök számára szokatlan megközelítései, valamint racionális gondolkodástól távoli következtetései jelentették számomra az unalmas hétköznapiakon – a dokkolástól a második beszélgetésig – a sikerélményt, a megnyugvást, a feltöltődést.

Csonka Picitől Jakupcseken át Pallagiig, sok-sok kérdést kaptam. Hol 4, hol 8 percben válaszolgattam is rájuk, közben legalább 15 interjút is adtam a legkülönbözőbb újságoknak, csatornáknak... Csak egyet nem kérdezett meg senki: Miért? Kinek a kedvéért? Kinek a dicsőségére csinálom végig ezt a tortúrát? Mikor éppen tavaszi szünet van, én meg négy napja nem láttam a családomat, azóta csak a tornacsarnokban dőlök le néhány órára „hajnalban, hajnal előtt”, mint egy józanul részeg kocsmatöltelék...

Kérdés nélkül mindig csak annyit sikerült elmondanom, hogy egy nagy tudású, humanista tanárom Ó nekem, a Professzor Úr, Simonyi Károly. Én még vizgázhattam, szigorlatozhattam nála „villamosságtanból”, elektronfizikából. Négy aláírása is díszíti az indexemet. Majd 15 évig egy tanszéken dolgoztunk... de ez nem igazán érdekelt senkit...

Tisztelt Professzor Úr!

Ezúton szeretném megköszönni Neked körülbelül húszezer magyar és nem tudom hány ezer német, orosz, román és kínai villamosmérnök nevében, hogy olyan szinten megtanítottad nekünk a szakma alapjait, hogy képesek voltunk évtizedek múltán is a fejünk felett elszáguldo fiaddal biztos rádiókapcsolatot létesíteni.



QoSE: Four letters or much more?

Keywords: QoS, Quality of Service Experienced, expectation, satisfaction, IPTV

In this article we study the origins and the measurable indicators of QoSE (Quality of Service Experienced) and examine the definition of this concept, with the help of an example. We identify the factors that put this concept into the focus. We show, by an other example, that QoSE may be a tool for a popular service (IPTV) in the way, helping service providers in their fight for competition preference.

Effect of QoS on infocommunication applications

Keywords: QoS, IPv4, IPv6, TCP, UDP, congestion, jitter, VoIP, codec, H.323, H.261/H264, ATM

The quality of service (QoS) in network environment is a special function to handle data traffic for network applications. For this purpose we need traffic manipulation and control algorithms. The QoS functionality serve both network applications and network management teams. While network administrators restrict network resources usage, network applications tend to use huge resources of the network. More than 95% of today's applications end in Ethernet nodes. This facility involves service cost reduction because it needs no protocol conversion during the data transfer. In this paper, we analyze the effect of QoS in L2 and L3 layers having H.261 and H.264 coded video streams in a single QoS domain. Executing large number of measurements we evaluate the characteristics of TCP and UDP data streams in function of QoS parameters. We propose quantitative solution to measure the global quality of videoconference based on OS (Opinion Score) metric and give indications for coexistence of different dataflow in IEEE 802.3 LAN/WAN environment.

QoS and QoSE: Guidelines for evaluation of Internet access services

Keywords: QoSE, Internet, speed-metering

From internet users' point of view there is a gap between service quality (quality of service experienced – QoSE) and quality of service (QoS). In this paper we try to characterize the problem, showing the steps that have been taken nationally and internationally in the interest of the customers, and the directions of development.

Quality of Service in ADSL networks

Keywords: ADSL, traffic-shaping, policing, triple-play

The value-added services which appear in ADSL environments eg. voice, IPTV or video streaming need specific technical requirements to be met in the network. On one hand, these applications may require more bandwidth from the network, on the other hand, some of them are very sensitive to network delay, jitter and/or packet loss and these problems also need appropriate treatment. Moreover, in ADSL-based networks, where the number of subscribers can reach high numbers there is a frequent need for limiting the bandwidth of each individual subscriber – according to their service contract – which also require

QoS parameter modifications. These and other aspects of ADSL QoS are discussed in Cisco's special QoS model for ADSL access. This article aims to give detailed information on this architecture and implementation.

Queueing modelling and analysis of packet schedulers of DSL access networks in the case of complete and partial rejections

Keywords: DSL, queueing models, packet scheduler

In this paper an exact analysis is provided of priority queueing system modelling DSL access networks with pre-emption option on packet level. We have shown accuracy and efficiency of our numerical analysis by presenting numerical results based on simulation and numerical analysis both for complete and partial rejections. Consequently, this analysis could be applied for an in-depth packet-level performance evaluation of recent DSL systems.

Gateway selection in the GMPLS PCE architecture

Keywords: GMPLS, gateway, Path Computation Element

In this paper we show how the Path Computation Element-architecture of GMPLS can be used for gateway selection purposes. The goal of selection is to prevent overload in the servers, which may cause blocking of the calls, and minimise the transport load of the network. We propose algorithms that take one or both indicators, distance and gateway load, into consideration when selecting the gateway server. Using network simulations we compare the different algorithms and we show that low blocking ratio and reasonably low network load can be achieved by those algorithms which take both indicators into account.

Conflict resolution in RFID systems

Keywords: RFID, readers, transponders, conflict resolution

The paper introduces conflict resolution algorithms available for solving multiple reading problems in RFID environment and points out to conflicts among readers as well as transponders. Simulation results are presented for the conflict resolution algorithm of ISO 18000-3 Mode 1.

Numeric analysis of mobility management algorithms

Keywords: mobility management, modelling, handover

This paper investigates mobility management strategies from the point of view of their need of signalling and processing resources on the backbone network and load on the air interface. A method is proposed to model the serving network and mobile node mobility in order to be able to model the different types of mobility management algorithms. To obtain a good description of the network we calculate parameters from given topologies that we think are the most important ones. Mobility approaches derived from existing protocols and other, possible mobility scenarios are analyzed and their performance is numerically compared in various network and mobility scenarios. The aim is to give general designing guidelines for the next generation mobility managements on given network and mobility properties.