



BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM  
TÁVKÖZLÉSI ÉS MÉDIAINFORMATIKAI TANSZÉK

CSOMAGKAPCSOLT HÁLÓZATOKBAN ALKALMAZHATÓ  
MINŐSÉGBIZTOSÍTÁSI ARCHITEKTÚRÁK  
TELJESÍTMÉNYVIZSGÁLATA

Martinecz Mátyás

Tézisfüzet

Tudományos vezető

Dr. Bíró József és Dr. Heszberger Zalán  
Nagysebességű Hálózatok Laboratóriuma  
Távközlési és Médiainformatikai Tanszék  
Budapesti Műszaki és Gazdaságtudományi Egyetem

Budapest

2008

# 1 Bevezetés

Napjainkban a különböző célú hálózatok konvergenciájának lehetünk tanúi. Ezen hálózatok közös nevezője valószínűleg a csomagkapcsolt átvitel lesz, amely sok előnye mellett néhány megoldásra váró problémát is felvet. Az egyik legfőbb kihívás a szolgáltatás minőségére vonatkozó minőségi garanciák (ún. QoS garanciák) nyújtása, ezek nélkül ugyanis az olyan új, értéknövelt szolgáltatások, mint pl. a Voice over IP (VoIP) vagy a Video on Demand (VoD) elterjedése és piacképessége elképzelhetetlen. Ez tehát egy olyan, fontos terület, amely nem csak a kutatók, de a távközlési hálózatok üzemeltetőinek és a tartalomszolgáltatóknak megkülönböztetett figyelmében kell, hogy részesüljön.

Sokan lebecsülik a QoS biztosítási technikák fontosságát, különösen a "nagy sávszélesség" filozófiájának követői. Ők a csomagkapcsolt hálózatok túlméretezésére voksolnak, azt állítva, hogy ezzel implicit módon a megfelelő átviteli minőség is garantálható. Úgy tűnik, mostanáig nekik volt igazuk, hiszen az átviteli technológiák evolúciójának köszönhető átviteli sebesség növekedés üteme rendre felülmúlta a nagy sávszélességet igénylő alkalmazások iránti igény gyarapodását. Ez a helyzet azonban a mobil és vezeték nélküli technológiák térhódításával megváltozhat a jövőben, hiszen ezeknél a sávszélesség bővítés vagy lehetetlen (a véges frekvenciatartomány miatt) vagy túl költséges ahhoz, hogy piacképes alternatívaként szolgáljon.

A fent elmondottakat figyelembevéve egy hatékony QoS biztosítási mechanizmus, pl. egy beengedést szabályozó algoritmus, amely egy kellően pontos, ugyanakkor gyors és egyszerű erőforrásigény becslő eljárásra épít - olcsóbb megoldást jelenthet, mint a hálózat egyszerű túlméretezése. A beengedett hálózati folyamatok erőforrásigényének megfelelően pontos ismerete más esetekben is nagyon hasznosnak bizonyulhat, pl. amikor valamilyen előre nem látott vagy kritikus szituáció adódik (pl. linkek meghibásodása, hirtelen megugró igény egyes szolgáltatások iránt, stb.).

Disszertációmban hatékony erőforrásigény becslő eljárásokat mutatok be, amelyek a Nagy Eltérések Elméletében (Large Deviation Theory - LDT) gyökereznek. Ezekkel a módszerekkel megbecsülhető egy adott aggregált forgalmi folyam számára minimálisan szükséges átviteli kapacitás vagy tárolóigény abban az esetben, amikor a minőségre vonatkozó előírás az adatvesztésre vonatkozik. Bemutatok továbbá olyan kombinált méretezési módszereket, amelyekkel mód nyílik egyidejű rendelkezésre állási és adatvesztési előírások teljesítésére. Ezeken túl ismeretetek egy olyan minőség-biztosítási architektúrát is, amelyet elsősorban hozzáférési hálózatokban lehet alkalmazni transzparens minőségi garanciák nyújtására.

## 2 Célkitűzések

Célul tűztem ki olyan technikák megalkotását, melyekkel csomagkiszolgálók fő paramétereit - tehát a tárolóméretet és a kiszolgálási sebességet - hatékonyan lehet méretezni abban az esetben, ha a maximálisan megengedett adatvesztés mértéke munkavesztési aránnyal (Workload Loss Ratio - WLR) jellemzett. Ezen egyszerűbb technikák alapján szerettem volna olyan kombinált, ún. többszintű QoS-t biztosító eljárásokat megalkotni, amelyek nem csak csomagszintű (pl. WLR), hanem hívásszintű (pl. blokkolási valószínűség) QoS igényeket is képesek egyidejűleg figyelembe venni. Végső célom egy olyan megbízható, könnyen megvalósítható minőségbiztosítási architektúra kidolgozása volt, amely nagyban épít a fent említett méretezési módszerekre és képes megfelelő QoS-t biztosítani értéknövelt szolgáltatások számára elsősorban nagysebességű hozzáférési hálózatokban.

A disszertációmban található kutatási eredményeimet három téziscsoportba szerveztem. Az első téziscsoportban olyan új ekvivalens kapacitásbecslőket mutatok be, amelyekkel kiszámítható egy adott hálózati folyamat számára az előírt WLR biztosításához minimálisan szükséges kiszolgálási kapacitás. Ezek a becslők fontos szerepet játszhatnak azon hálózatméretezési feladatoknál, amelyeknél a hálózat üzemeltetőjének fő célja megbízható adatvesztésre vonatkozó garanciák nyújtása. A bemutatott technikák alapvetően bonyolultságukban és pontosságukban különböznek egymástól, például egyes formuláktól pontosabb becslést remélhetünk, amennyiben a vizsgált hálózati folyamat jól jellemezhető Gauss folyamattal, míg más formulák pontosatlanabb, de egyszerűbben kiértékelhető közelítést valósítanak meg.

A második téziscsoportban hatékony tárolóigény becslő eljárásokat ismertetek, amelyekkel kiszámítható egy adott hálózati folyamat számára minimálisan szükséges tárméret abban az esetben, ha az átviteli kapacitás rögzített és WLR-re vonatkozó garanciát szeretnénk nyújtani. Ezek a formulák az első téziscsoportban bemutatott eljárások kiegészítő technikáiként is felfoghatók abban az értelemben, hogy alkalmazásuk elsősorban akkor javasolt, amikor a kiszolgálás sebessége nem növelhető tetszőleges mértékben (pl. ha az átviteli link véges frekvenciasávval gazdálkodó rádiós interfészen valósul meg), mégis szeretnénk adatvesztési garanciát nyújtani. Az itt ismertetésre kerülő formulák teljesítménye jelentősen eltérhet egymástól az alkalmazási kör függvényében, ezért ezeknél kitérek majd arra, hogy melyik technikát milyen esetekben érdemes használni.

A harmadik téziscsoport három új eredményt tartalmaz. Ezekből az első kettő kombinált, többszintű QoS garanciákat nyújtó méretezési technika, amellyel kiszámítható egy adott forgalom számára minimálisan szükséges átviteli kapacitás vagy tárolóméret abban az esetben, ha a QoS kritériumok WLR-re és blokkolási valószínűsége (Time Blocking Probability - TBP) vannak megadva. Ezek a méretezési eljárások képezik az utolsó tézisben ismertetésre kerülő minőségbiztosítási architektúra magját.

A minőségbiztosítási architektúra elsősorban nagysebességű hozzáférési hálózatokban történő működésre lett kitalálva, ahol a megfelelő átviteli minőség biztosítása az érték-növelt szolgáltatások számára kiemelt fontosságú feladat. Az architektúra fő építő elemeinek, az alkalmazott hívásengedélyezési szabályrendszernek, továbbá a működéshez szükséges protokolloknak a leírása szintén itt kapott helyet.

### **3 Kutatási módszertan**

Kutatási munkám során olyan LDT alapú modelleket alkalmaztam, amelyek alkalmasak egyetlen tárolós kimeneti linkkel rendelkező csomagkiszolgáló csomópont leírására. Ezzel a megközelítéssel a valós csomagkiszolgálók (routerek) teljesítményének vizsgálatára alkalmas modellezési keretrendszerhez jutottam.

A disszertációban található eredményeket alapvetően analitikus módszerrel értem el. Mivel az analitikusan levezetett formulák hatékony kiértékeléséhez közelítésekkel kellett élnem, melyek hatása analitikusan nehezen becsülhető, kiterjedt numerikus vizsgálatokat végeztem az új eredmények működési teljesítményének felmérésére.

## 4 Új eredmények

### 4.1 Ekvivalens kapacitás becslő technikák

Az újszerű értéknövelt szolgáltatások elterjedése a kialakulóban lévő sokszolgáltatásos (multiservice) hálózatokban csak akkor lehetséges, ha megbízható QoS garanciákat tudunk számukra nyújtani. Ez a kijelentés vitán felül áll, azonban a végső, univerzális módszer a vonatkozó "hogyan" kérdés megválaszolására mindmáig ismeretlen. A fő probléma ezzel kapcsolatban a megvalósítandó rendszer költséghatékonyságával függ össze: az erőforrás-megtakarítás és rendszerüzemeltetési költségek (illetve rendszerkomplexitás) által kijelölt függvényen keressük a rendszer optimális munkapontját.

Egy jól alkalmazható minőségbiztosítási mechanizmus "szíve és lelke" egy olyan hatékony erőforrásigény-becsülő metódus, amely képes megállapítani sokféle forgalomtípus erőforrás-szükségleteit megfelelő pontossággal. Alapvetően az erőforrásigény két paraméterrel írható le: átviteli kapacitás és tárolókapacitás. A két mennyiség nyilvánvalóan hat egymásra, de általában mindig létezik egy vezérelv vagy éppen-séggel egy kényszerfeltétel, amely eldönti, melyik paraméter rögzített és melyik változtatható. Sok esetben a tároló méretét a választott hardver, vagy az egy adott forgalomtípus által elviselt maximális késleltetés mértéke korlátozza, ezért az esetek többségében az adott átviteli minőséget garantáló minimális átviteli kapacitást szeretnénk meghatározni

Számos cikk foglalkozik a kapacitásbecslés problémakörével. Alapműnek számító munkák a témában Kelly [7] és Roberts [8] cikkei. Courcobetis és mások [14] úgynevezett ekvivalens kapacitás formulákat dolgoztak arra az estére, amikor a QoS előírás a tároló túlsordulási valószínűségére vonatkozik. Seres és mások [3] olyan technikákat alkottak meg, melyekkel a Courcobetis formulákkal megegyező eredményhez jutunk, viszont kisebb számításigényűek azoknál. Likhanov és Mazumdar [5] egzakt aszimptotikus becselőket adott meg a WLR értékére, míg Jamin és mások [12], továbbá Courcobetis és mások [27] olyan eljárásokat dolgoztak ki, melyekkel az ekvivalens kapacitás számításhoz szükséges bemenő paraméterek mérés alapú módszerekkel határozhatók meg.

A tézisfüzet ezen szakaszában három kapacitásbecslő eljárást ismertetek, melyek alkalmasak egy adott forgalom számára az előírt munkavesztési arány teljesítéséhez szükséges minimális átviteli sebesség kiszámítására. A fő különbség, amely az itt bemutatandó módszereket megkülönbözteti a többi, szakirodalomban fellelhető módszertől az, hogy a QoS előírás itt a munkavesztési arányra (WLR), míg azoknál legtöbbször a tároló túlsordulási valószínűségére vonatkozik, továbbá az általam kidolgozott formulák gyorsabban kiszámíthatóak a már ismert, a WLR definícióján

alapuló eljárásoknál. Az ismertetésre kerülő kapacitásbecslők egymástól pontosságban, számításigényben és javasolt alkalmazási körben térnek el egymástól.

Jelölje az  $X[0, t)$  sztochasztikus folyamat az  $N$  független forrásból, a  $[0, t)$  időintervallumban, a  $B$  tárolóval rendelkező,  $C$  átviteli sebességű kommunikációs linkhez érkező forgalmat.  $N$ -et skála tényezőnek tekintve definiálható  $b = \frac{B}{N}$  és  $c = \frac{C}{N}$ , az egy forrásra jutó tárolóméret és kiszolgáló kapacitás. Tegyük fel továbbá, hogy  $X[0, t)$  stacioner növekményű.

A puffertúlsordulási valószínűség,  $P(Q > B)$  a fent leírt véges tárolójú rendszerben leírható az idő azon részével, amikor a  $Q(C, N)$  sorhossz  $B$ -nél nagyobb a vonatkozó végtelen tárolóméretű rendszerben. Az LDT keretrendszeren belüli úgynevezett sokforrásos aszimptotikus (Many Sources Asymptotics - MSA) egyenlet értelmében a puffertúlsordulási valószínűség logaritmusának lecsengése aszimptotikusan lineáris  $N$ -ben, ha az egy forrásra eső rendszerparamétereket konstans értéken tartjuk [16] [17]:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log P(Q(N, cN) > bN) = \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s, t)}{N} - s(b + ct) \right\} \stackrel{\text{def}}{=} -I, \quad (1)$$

ahol  $\Lambda(s, t) \stackrel{\text{def}}{=} \log E[e^{sX[0, t)}]$  az  $X[0, t)$  kumulánsképző függvénye,  $s$  és  $t$  szabad paraméterek. (1)-re gyakran hivatkoznak úgy, mint az LDT sokforrásos aszimptotikus egyenletére.

A fent leírt egyenlet gyakorlati következménye, hogy nagy  $N$  értékekre a puffertúlsordulási valószínűség egyszerűen közelíthető a következő összefüggéssel:

$$P(Q(N, C) > B) \approx e^{-NI}, \quad (2)$$

ahol  $-NI$  a következőképpen számolható:

$$-NI = \sup_{t>0} \inf_{s>0} \{ \Lambda(s, t) - s(B + Ct) \}. \quad (3)$$

A munkavesztési arányra vonatkozó QoS előírás a következő formában írható:

$$WLR \stackrel{\text{def}}{=} \frac{E[Q - B]^+}{E[X]} \leq e^{-\gamma}, \quad (4)$$

ahol  $X$  az időegység alatt érkező munkamennyiséget leíró valószínűségi változó.

Azt a minimális kiszolgálási sebességet, amely még éppen biztosítja az előírt WLR korlát teljesülését ekvivalens kapacitásnak nevezzük és  $C_{equ, WLR}$ -rel jelöljük. A mennyiség definíciója:

$$C_{equ, WLR} \stackrel{\text{def}}{=} \inf \{ C : WLR \leq e^{-\gamma} \}. \quad (5)$$

**Tézis 1.** [J2] [C4] Olyan becslő technikákat dolgoztam ki, amelyekkel WLR előírás esetére lehet kiszámítani egy adott aggregált forgalom ekvivalens kapacitását. Bebizonyítottam, hogy az általam javasolt megoldások megőrzik, sőt közülük néhány túl is szárnyalja az LDT sokforrásos aszimptotikus egyenletének, (1)-nek a megoldására jellemző logaritmikus aszimptotikus tulajdonságot. Arra is rámutattam, hogy az általam kidolgozott formulák gyorsabban vezetnek eredményre, mint a szakirodalomban fellelhető módszerek.

#### 4.1.1 Elméleti ekvivalens kapacitás becslő eljárás

**Tézis 1.1.**  $C_{equ,WLR}$  közelíthető a következő fix-pont típusú egyenlet megoldásával:

$$c = \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s,t) + F(\gamma,c)}{st} - \frac{B}{t} \right\} \stackrel{\text{def}}{=} f(c), \quad (6)$$

amely megoldandó  $c$ -re, ahol  $F(\gamma,c) = \gamma - \log P\{Q > 0\} + \log WLR(0)$ . Analitikus módszerrel megmutattam, hogy (6) megoldása megőrzi (1) megoldásának logaritmikus aszimptotikus tulajdonságát.

A  $c = f(c)$  fix-pont típusú egyenletet iteratíván kell megoldani, azaz  $c_{n+1} = f(c_n)$ , ahol  $n$  az iterációs lépés sorszáma. Az iterációs eljárást mindaddig folytatni kell, amíg  $c$  és  $f(c)$  relatív különbsége a megadott  $\varepsilon$  küszöb alá csökken. Az algoritmus ezen leállási feltétele formálisan így írható:

$$\frac{|c_n - f(c_n)|}{\max\{c_n, f(c_n)\}} \leq \varepsilon. \quad (7)$$

Az (6) kifejezés kiértékeléséhez egy dupla optimalizálási feladatot kell megoldani minden egyes iterációs lépésben. Ez a megközelítés kellően pontos becslést képes adni kevesebb lépésben, mintha a (5) definíció alapján akarnánk kiszámolni az ekvivalens kapacitást.

#### 4.1.2 Praktikus, konzervatív ekvivalens kapacitás becslési módszer

Az 1.1 tézisben megfogalmazott ekvivalens kapacitás becslő technika alkalmazásához szükség van  $P(Q > 0)$  meghatározására minden egyes iterációs lépésben. Ennek a mennyiségnek a becslése azonban messzemenően nem triviális feladat, ezért javasolt annak kikerülése. Egyes szerzők (pl. Shroff és mások [6]) állítása szerint bizonyos puffertúlsordulási valószínűség becslő technikák elég pontos eredményt adnak  $B = 0$  esetén is, azonban az általuk említett technikák paraméterigénye jelentős lehet, ami megnehezíti alkalmazásukat. Ezért szükséges valamilyen kerülőutat választani, és

$P(Q > 0)$  kifejezést olyan mennyiséggel helyettesíteni, amely kellően pontos becslést ad, de ugyanakkor könnyen és kevés paraméterből számítható.

A következő tézisben egy olyan becslő eljárást ismertetek, amely már nem tartalmazza  $P(Q > 0)$  mennyiséget, helyette a jóval könnyebben számolható  $P(X > C)$  kifejezés jelenik meg a fix-pont típusú egyenletben. Azt állítom, hogy a  $P(X > C)$ -t használó technikával az (6) egyenlet megoldásának egy konzervatív, felső becslését kapjuk. Jelöljük az 1.1 tézisben szereplő fix-pont típusú egyenlet megoldását  $\tilde{C}_{equ,WLR}$ -l, formálisan leírva:

$$\tilde{C}_{equ,WLR} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} c_n, \quad (8)$$

ahol  $c_n$  ( $n$  pozitív egész) jelöli a (6) kifejezés értékét az  $n$ . iterációs lépésben.

**Tézés 1.2.** *A következő praktikus fix-pont típusú egyenlet megoldásával becsülhető  $C_{equ,WLR}$ :*

$$\hat{c} = \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s, t) + \hat{F}(\gamma, \hat{c})}{st} - \frac{B}{t} \right\}, \quad (9)$$

amelyet  $\hat{c}$ -re kell megoldani, ahol  $\hat{F}(\gamma, \hat{c}) = \gamma - \log P\{X > \hat{c}\} + \log WLR(0)$ .

Megmutattam, hogy (9) megoldása nem kisebb  $\tilde{C}_{equ,WLR}$ -nél, tehát ez a módszer alkalmas annak konzervatív becslésére. Rámutattam arra is, hogy amennyiben az aggregált forgalom néhány tíz vagy száz ITU-T G.729 kódolású VoIP vagy MPEG-4 high-quality video folyamból áll, az 1%-os relatív hibához kötött leállási feltétel néhány iterációs lépés alatt elérhető.

Jelölje (9) megoldását  $\hat{C}_{equ,WLR}$ , formálisan:

$$\hat{C}_{equ,WLR} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \hat{c}_n, \quad (10)$$

ahol  $\hat{c}_n$  ( $n$  pozitív egész) jelöli (9) kifejezés értékét az  $n$ . iterációs lépésben.

A fent ismertetett eljárás alkalmazásánál  $P(X > C)$  és  $WLR(0)$  értékét kell meghatározni minden egyes iterációs lépésben. Mivel  $P(X > C)$  becslésére számos technika fellelhető a szakirodalomban, ezért könnyű ezek közül egy olyat választani, ami alacsony számítás- és paraméterigényű. Bár a  $WLR(0)$  mennyiség kiszámításával viszonylag kevesen foglalkoztak, mégis léteznek kis paraméterigényű és kellően pontos technikák, amelyek számításigénye sem jelentős. Ezekről egy színvonalas összefoglaló értekezés található [31]-ben.

Ahogy azt már az előző szakaszban említettem az (6) fix-pont típusú egyenlet kevesebb lépésben képes kellően pontos eredményt szolgáltatni, mint az ekvivalens kapacitás definícióján alapuló módszerek. Ez a kijelentés igaz marad az 1.2 tézisben bemutatott fix-pont típusú egyenletre is, azonban a becslés pontossága rosszabb lehet mind az (6) egyenlet megoldásánál, mind a definíció alapján számolt ekvivalens kapacitásnál a  $P(Q > 0) \approx P(X > C)$  közelítés miatt.



### 4.1.3 Egzakt aszimptotikus ekvivalens kapacitás becslési eljárások

A (6) és (9) formulák a sokforrásos aszimptotikus egyenletben gyökerező puffersaturációs valószínűség becslésre épülnek, amely becslés azonban csak logaritmikusan aszimptotikus tulajdonságokkal bír. Léteznek olyan puffersaturációs valószínűséget közelítő módszerek is, amelyek egzakt aszimptotikus tulajdonságokkal bírnak, ezek Bahadur és Rao eredményein alapulnak [15]. A Bahadur-Rao előtag felhasználásával előnyösebb aszimptotikus tulajdonságokkal bíró becslő eljárások konstruálhatók, ennek ára azonban az, hogy ismerni kell a kumulánsképző függvény második deriváltját is. Ez a tényező azonban egy Montgomery és De Veciana által javasolt alkalmas becsléssel kiiktatható a formulákból [13].

Courcobetis és mások [14], valamint Seres és mások [3] eredményei alapján két olyan ekvivalens kapacitás becslő módszert dolgoztam ki, amelyek paraméter- és számításigénye gyakorlatilag azonos az előző tézisekben megfogalmazott formulákéval, viszont az általuk szolgáltatott eredmény kedvezőbb (egzakt) aszimptotikus tulajdonságokkal bír. Ezek a becslők tehát elvben jobb eredményt adhatnak az előbbieknél, különösen akkor ha a vizsgált forgalom Gauss-i tulajdonságokkal bír (mert ilyenkor az egyik közelítés hibátagja nulla).

**Tézis 1.3.** *A következő fix-pont típusú egyenletek megoldásaival közelíthető  $C_{equ,WLR}$ :*

$$c^{B \cdot R} = \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s, t) + G(F(\gamma, c^{B \cdot R}))}{st} - \frac{B}{t} \right\} \quad (11)$$

és

$$\hat{c}^{B \cdot R} = \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s, t) + G(\hat{F}(\gamma, \hat{c}^{B \cdot R}))}{st} - \frac{B}{t} \right\} \quad (12)$$

melyek megoldandók  $c^{B \cdot R}$ -re és  $\hat{c}^{B \cdot R}$ -re, ahol  $G(x) \stackrel{\text{def}}{=} x - \frac{\frac{1}{2} \log 4\pi x}{1 + \frac{1}{2x}}$ .

*Analitikus módszerrel igazoltam, hogy a fenti egyenletek megoldásai Gaussi-i forgalmak esetén egzakt aszimptotikus tulajdonságokkal rendelkeznek és ebben az értelemben pontosabb becslést adhatnak (6) megoldásánál. Azt is megmutattam, hogy (12) megoldása (11) megoldásának egy könnyeben számolható felső korlátja.*

Jelöljük a (11) és (12) egyenletek megoldásait rendre  $\tilde{C}_{equ,WLR}^{B \cdot R}$ -vel és  $\hat{C}_{equ,WLR}^{B \cdot R}$ -vel, formálisan:

$$\tilde{C}_{equ,WLR}^{B \cdot R} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} c_n^{B \cdot R} \quad (13)$$

és

$$\hat{C}_{equ,WLR}^{B \cdot R} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \hat{c}_n^{B \cdot R}, \quad (14)$$

ahol jelölje rendre  $c_n^{B \cdot R}$  és  $\hat{c}_n^{B \cdot R}$  ( $n$  pozitív egész) (11) és (12) kifejezés értékét az  $n$ . iterációs lépésben.

A (11) és (12) egyenletekben megfogalmazott ekvivalens kapacitás becselő módszereket elsősorban Gauss-i tulajdonságokat mutató aggregált forgalmak esetén érdemes használni, mert ebben az esetben pontosabb becslést adhatnak az eddig ismertetett módszereknél. A sok, független, azonos és véges szórású sebességeloszlással bíró forrás összegeként előálló aggregált folyamatok Gauss-i folyamattal jellemezhetők.

## 4.2 Tárolóigény becselő eljárások

A kapacitásméretezési technikák nem mindig az egyetlen lehetséges módszerei az átviteli minőségre vonatkozó garanciák nyújtásának. Sok esetben nem lehetséges vagy nem gazdaságos (pl. rádiós hálózatokban) a megnövekedett forgalmi terhelésre válaszul a fizikai csatorna sávszélességét növelni. Még gyakrabban - különösen borsztös forgalmak esetén - jelentős átviteli kapacitás takarítható meg ha a kiszolgálási sebesség helyett inkább a tároló méretét növeljük,

Természetesen a tárolóméret növelésének megvannak a maga korlátai: egy megfelelő méretű és sebességű puffer alkalmazása pénzügyi és technológia korlátokba ütközhet, de ami még ennél is fontosabb, a nagyobb méretű tároló megnövelheti a folyamatok által érzékelt késleltetést, ami végső soron élvezhetetlenné tehet egyes értéknövelt szolgáltatásokat. Ez a probléma különösen érzékenyen hat például a valós-idejű alkalmazásokra, mint amilyen pl. a VoIP vagy a VoD. Ezeket a korlátokat tehát szem előtt kell tartani, amikor a hálózattervezés során az egyes rendszerparaméterek változtatásáról döntünk.

A különböző sorbanállási rendszerek sorhossz-eloszlásával számos tanulmány foglalkozik. A [24] szerzői a cellavesztési valószínűség és a kimeneti puffer mérete között fennálló összefüggést vizsgálják ATM kapcsolókban. Az Appenzeller és mások az általuk publikált könyvben [25] az LDT keretrendszerben megfogalmazott sorhossz-eloszlásra vonatkozó eredményeket összegzik. Ganesh és mások színvonalas cikkükben [26] torlódásvezérelt (alapvetően TCP) folyamatok által igénybevett csomagkiszolgáló tárolójának méretezésére fogalmaznak meg ökölszabályokat.

A következő téziscsoportban olyan technikákat ismertettek, amelyekkel lehetőség nyílik egy rögzített kiszolgálási sebességű csomagkiszolgáló számára az előírt munkavesztési arány teljesítéséhez minimálisan szükséges tárolóméretének a kiszámítására. Az első technika a sokforrásos aszimptotikus egyenlőségre, továbbá Seres és mások [3] illetve Shroff és mások [6] eredményeire épít és a tárolóméret becslés logaritmikus aszimptotikus tulajdonságokkal bír. A második tézisben bemutatásra kerülő formula az elsőnek egy könnyebben kiértékelhető felső korlátját adja eredményül. A harmadik tézisben ismertetésre kerülő módszerekben alkalmazásra kerül a már említett Bahadur-Rao előtag [15], így ezekkel a technikákkal elvben pontosabb, egzakt aszimptotikus tulajdonsággal rendelkező közelítés adható a minimális tárméretre.

Jelölje az  $X[0, t)$  sztochasztikus folyamat az  $N$  darab független forrásból, a  $[0, t)$  időintervallumban, a  $C$  kapacitású linkre érkező forgalmat. Tegyük fel továbbá, hogy  $X[0, t)$  stacioner növekményű.

Definiáljuk a minimális tárolóigényt, mint azt a legkisebb tárméretet, amely még biztosítja az előírt  $e^{-\gamma}$  munkavesztési arány teljesülését. Formálisan:

$$B_{req, WLR} \stackrel{\text{def}}{=} \inf \{ B : WLR \leq e^{-\gamma} \}. \quad (15)$$

**Tézis 2.** *Olyan eljárásokat dolgoztam ki, amellyel megbecsülhető egy adott aggregált forgalom minimális tárolóigénye abban az esetben, ha a QoS előírás WLR-re vonatkozik. Analitikusan igazoltam, hogy egyes eredményeim megőrzik, mások túl is szárnyalják a sokforrásos aszimptotikus egyenlet megoldásának aszimptotikus tulajdonságait. Igazoltam továbbá, hogy az általam javasolt formulák számításigénye alacsonyabb a szakirodalomban fellelhető megoldásokénál.*

#### 4.2.1 Elméleti módszer a tárolóigény becslésére

**Tézis 2.1.** *A következő formulával becsülhető  $B_{req,WLR}$ :*

$$B_{req,WLR} \approx \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s,t) + F(\gamma, C)}{s} - Ct \right\} \stackrel{\text{def}}{=} \tilde{B}_{req,WLR}, \quad (16)$$

ahol

$$F(\gamma, C) = \gamma - \log P\{Q > 0\} + \log WLR(0).$$

*Megmutattam, hogy  $\tilde{B}_{req,WLR}$  megőrzi a sokforrásos aszimptotikus egyenlet megoldásának logaritmikus aszimptotikus tulajdonságát. Igazoltam továbbá, hogy (16) számításigénye alacsonyabb a szakirodalomban található más módszerekénél.*

A (16) kifejezés kiértékeléséhez egy kettős optimalizálási feladatot kell megoldani. Ez a módszer ekvivalens a (15) definíció alapuló implicit módszerrel abban az értelemben, hogy szükségszerűen ugyanarra az eredményre vezet. Azonban az implicit módszernél egy háromdimenziós optimalizálási problémával állunk szemben, hiszen a várható WLR mértékét meg kell becsülni különböző  $C$  értékekre, amely lépés eleve egy kétdimenziós optimumkeresést jelent. Következésképp a (16)-ban leírt módszer számításigény szempontjából jóval előnyösebb a definíció alapuló, implicit becslő technikánál.

#### 4.2.2 Praktikus, konzervatív tárigény becslési eljárás

Jó becslést találni  $P(Q > 0)$  értékére a (16) közelítés kiértékeléséhez problémát jelenthet, ezért egy olyan másikat javasoltam a tárigény becslésére, amelyhez nincs szükség  $P(Q > 0)$  meghatározására, viszont a becslés elfogadhatóan pontos marad és könnyen kiértékelhetővé válik. Az alapötlet itt  $P(Q > 0)$  közelítése annak egy könnyen kiszámítható alsó korlátjával,  $P(X > C)$ -vel, mivel így a módosított eljárás eredménye garantáltan nem lesz kisebb a 2.1 tézisben ismertetett technikával számolt tárigény értékénél.

**Tézis 2.2.** *A következő, könnyen kiértékelhető formulával becsülhető  $B_{req,WLR}$  értéke:*

$$\tilde{B}_{req,WLR} \leq \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s,t) + \hat{F}(\gamma, C)}{s} - Ct \right\} \stackrel{\text{def}}{=} \hat{B}_{req,WLR}, \quad (17)$$

ahol

$$\hat{F}(\gamma, C) = \gamma - \log P\{X > C\} + \log WLR(0).$$

Azt is megmutattam, hogy  $\hat{B}_{req,WLR}$  egy alkalmas felső korlátja  $\tilde{B}_{req,WLR}$ -nek.

A (17) kifejezés kiértékeléséhez szükség van a puffertelen szaturációs valószínűség ismeretére, amely mérték azonban sokkal könnyebben számolható, mint  $P(Q > 0)$ , ezáltal a most bemutatott módszer gyakorlati alkalmazása lényegesen egyszerűbb, mint a 2.1 tézisben tárgyalt technikáé.

### 4.2.3 Egzakt aszimptotikus tárolóigény becslő módszer

A 2.1 és 2.2 tézisekben ismertetett eljárások a sokforrásos aszimptotikus egyenletben gyökerező  $P(Q > B)$  becslére épülnek, aminek köszönhetően öröklik annak logaritmus aszimptotikus tulajdonságát. Ennél erősebb, egzakt aszimptotikus tulajdonság is elérhető azonban, ha az eredményekbe beépítjük a Bahadur-Rao előtagot.

A következő tézisben két olyan tárí igény becslő módszert fogalmazok meg, amelyek az egzakt aszimptotikus tulajdonságú  $P(Q > B)$  becslésre épülnek, így véges számú forrásra a legrosszabb esetet feltételezve pontosabb becsléssel szolgálnak a szükséges tárí igényre vonatkozóan.

Fontos tulajdonsága az ismertetendő formuláknak, hogy továbbra is ugyanabban az egyszerű, explicit alakban számíthatók, mint a 2.1 illetve 2.2 tézisekben megfogalmazott eredmények, így számításigényük is nagyságrendileg azonos.

**Tézés 2.3.**  $B_{req,WLR}$  értéke becsülhető a következő kifejezésekkel:

$$B_{req,WLR} \approx \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s,t) + G(F(\gamma, C))}{s} - Ct \right\} \stackrel{\text{def}}{=} \tilde{B}_{req,WLR}^{B-R} \quad (18)$$

és

$$\tilde{B}_{req,WLR}^{B-R} \leq \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda(s,t) + G(\hat{F}(\gamma, C))}{s} - Ct \right\} \stackrel{\text{def}}{=} \hat{B}_{req,WLR}^{B-R}, \quad (19)$$

ahol  $G(x) \stackrel{\text{def}}{=} x - \frac{\frac{1}{2} \log 4\pi x}{1 + \frac{1}{2x}}$ .

Megmutattam továbbá  $\tilde{B}_{req,WLR}^{B-R}$  egzakt aszimptotikus tulajdonságokkal bír, abban az esetben ha az aggregált forgalom Gauss-i tulajdonságú. Rémutattam arra is, hogy  $\hat{B}_{req,WLR}^{B-R}$  nem kisebb  $\tilde{B}_{req,WLR}^{B-R}$ -nél, tehát annak egy alkalmas, praktikus könnyebben számolható felső korlátja.

Megfigyelhető, hogy míg a (18) és (19) formulák elvben pontosabb becslést adhatnak, mégis a (16) és (17) kifejezésekhez nagyon hasonló formában jelennek meg. Ez burkoltan azt is jelenti, hogy ezen eljárások számításigénye is nagyon közel van egymáshoz. Ennek az előnyös tulajdonságnak az az ára, hogy míg a 2.1 illetve 2.2 tézisek eredményei univerzálisak, addig a 2.3 tézisben tárgyalt formulák alkalmazása elsősorban Gauss-i (pl. fBm modellel jól leírható) forgalmak esetén javasolt, ekkor ugyanis a Bahadur-Rao előtag beépítéséhez felhasznált közelítés hibatagja nullához tart.

### 4.3 Többszintű minőségi garanciát nyújtó technikák

Az előző téziscsoportokban bemutatott eljárások kapacitás- és tárolóméretezési problémák megoldására alkalmazhatók, amennyiben a QoS kritérium a munkavesztési arányra vonatkozik. A valós alkalmazások azonban nem csak vesztesre, de gyakran késleltetésre, késleltetésingadozásra vonatkozó elvárásokat is támasztanak a hálózattal szemben. Különösen ez a helyzet az értéknövelt, multimédia szolgáltatások esetén, mint amilyen a napjainkban egyre népszerűbb Video-on-Demand. Az elmondottak fényében szükség van tehát olyan QoS biztosító eljárásokra, amelyek az egyes alkalmazások szerteágazó minőségi kritériumait együttesen képesek figyelembe venni.

Egy másik fontos szempont, amit méretezési feladatok megoldása során érdemes figyelembe venni, az egyes forgalmak hívásszintű dinamikája, tehát a hálózati folyamatok születés-halálzási folyamatának feltérképezése. Ennek az információnak a méretezési módszerbe építésével sokkal pontosabb és gazdaságosabb eljárást kaphatunk, továbbá az egyes szolgáltatások megfelelő mértékű rendelkezése állása is garantálható.

A szakirodalomban található méretezési módszerek általában csak egy bizonyos forgalomgranularitási szinten vizsgálják a hálózati folyamatokat. Például a telefonközpontok (vagy úgy általában a vonalkapcsolt hálózatok) méretezésével foglalkozó művek rendszerint csak a hívásszintű modellekre fókuszálnak és a rendelkezésre állásra vonatkozó minőségi kritériumokat, gyakran a blokkolási valószínűséget (Time Blocking Probability - TBP) igyekeznek megbecsülni. A témában alapvető fontosságúak Erlang [9], Engset [10] és Kleinrock [11] nevéhez fűződő munkák.

A tágabb tématerülettel foglalkozó kutatók egy másik csoportja inkább a hálózati folyamatok belső dinamikáját és azok csomagszintű viselkedését vizsgálja. Az ő céljuk általában valamilyen korlátot vagy becslést adni a vizsgált rendszerben fellépő adatvesztés mértékére. Széles körben elismert művek ebben a témában - természetesen a teljesség igénye nélkül - Kelly [7], Courcobetis [14] és Roberts [8] nevéhez köthetők.

A következő téziscsoportban egy olyan kombinált méretezési eljárást ismertettek, amely alkalmas egyidejű hívás- és csomagszintű QoS garanciák nyújtására. Az általam kidolgozott technika alkalmazásával mód nyílik megfelelő szintű blokkolási valószínűség és munkavesztési arány együttes biztosítására.

Ebben a szakszban bemutatok továbbá egy olyan nagysebességű hozzáférési hálózatokban alkalmazható minőségbiztosítási architektúrát is, amely a kombinált méretezési technikákra nagyban építve alkalmas többszintű QoS garanciák nyújtására. Az architektúra építőelemeit, az alkalmazott protollokat illetve a javasolt hívásengedélyezési startégiát szintén részletesen tárgyalom.

Jellemezzük a hálózatot igénybe vevő szolgáltatások hívásszintű dinamikáját az Engset modellel [10] (lásd 1. ábra). Jelölje  $N$  a kiszolgálókapacitást (a szimultán folyamatok maximális számát),  $P$  a populáció méretét,  $\lambda$  a vizsgált szolgáltatástípus érkezési intenzitását,  $h$  pedig egy hívás átlagos tartási idejét. Tegyük fel, hogy a rendszer minden hívást beenged, ha a már beengedett folyamatok száma kisebb, mint

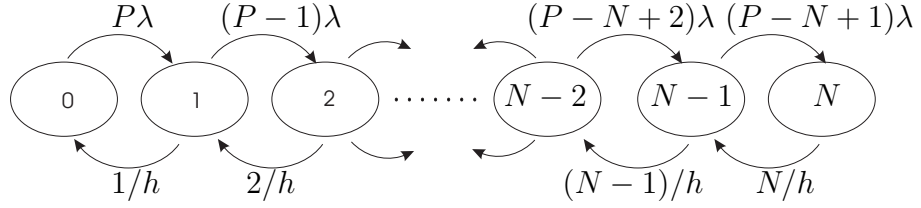


Figure 1: Az alkalmazott hívásszintű (Engset) modell

$N$ , minden más esetben elutasítja az újonnan érkező folyamat.

Definiáljuk a kombinált ekvivalens kapacitást, mint azt a legkisebb átviteli kapacitást, amely rögzített tárolóméret esetén biztosítja a  $WLR \leq e^{-\gamma}$  és a  $TBP \leq \beta$  munkavesztési arányra és blokkolási valószínűsége előírt kritériumok együttes teljesülését.

$$C^* \stackrel{\text{def}}{=} \inf \{C : WLR \leq e^{-\gamma}, TBP \leq \beta\} \quad (20)$$

Hasonlóan definiálhatjuk a kombinált minimális tárolóméretet is:

$$B^* \stackrel{\text{def}}{=} \inf \{B : WLR \leq e^{-\gamma}, TBP \leq \beta\} \quad (21)$$

### Tézis 3. [J1] [C2] [C3]

*Olyan kombinált méretezési technikákat dolgoztam ki, amelyekkel mód nyílik egyidejű hívás- és csomagszintű minőségi kritériumok biztosítására. Megalkottam továbbá egy olyan hibatűrő, többszintű QoS garanciákat megbízhatóan nyújtó minőségbiztosítási architektúrát, amely nagyban épít a kombinált méretezési eljárásokra és elosztott, rugalmas, transzparens hívásengedélyezési módszert valósít meg.*

#### 4.3.1 Többszintű QoS kritériumokat biztosító kapacitásbecslő eljárás

**Tézis 3.1.** *A következő fix-pont típusú egyenlet megoldásával becsülhető a kombinált ekvivalens kapacitás*

$$c^* = \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda_{N^*}(s, t) + K}{st} - \frac{B}{t} \right\}, \quad (22)$$

amely megoldandó  $c^*$ -ra, ahol  $B$  a tárolóméret,  $s$  és  $t$  szabad paraméterek,  $K \in \{F(\gamma, c^*), \hat{F}(\gamma, c^*), G(F(\gamma, c^*)), G(\hat{F}(\gamma, c^*))\}$  és  $\Lambda_{N^*}(s, t)$  az  $X_{N^*}(0, t]$  kumulánsképző függvénye,  $X_{N^*}(0, t]$  az  $N^*$  számú forrásból a  $(0, t]$  időintervallumban érkező munkamenységet leíró valószínűségi változó. A források számának ( $N^*$ ) kiszámítására a



következő módszert javaslom:

$$N^* = \inf \left\{ N : \frac{\binom{P}{N} (\lambda h)^N}{\sum_{j=0}^N \binom{P}{j} (\lambda h)^j} \leq \beta \right\}. \quad (23)$$

Egyszerűen megfogalmazva a 3.1 tézis azt állítja, hogy ha  $P$  méretű populáció  $\lambda$  intenzitással kezdeményez  $h$  átlagos tartási idejű hívásokat, akkor  $N^*$  számú folyamat egyszerre beengedő csatornát kell biztosítanunk a blokkolási valószínűségekre vonatkozó előírás teljesítéséhez. Ha azt is szeretnénk elérni, hogy a beengedett folyamatok által tapasztalt WLR se lépje túl a megengedett értéket, akkor - a legrosszabb esetet feltételezve (tehát amikor valóban  $N^*$  számú forgalom tartózkodik a rendszerben) - az 1. és 2. téziscsoportokban bemutatott módszerek értelmében legalább a (22) formulával kiszámolt kapacitására kell az átviteli csatornát méretezni.

A fent ismertetett eljárás megfelelő működésének kritériuma az alkalmazott modellek pontossága és egy egyszerű, maximum  $N^*$  forgalmat beengedő hívásengedélyezési metódus megvalósítása.

#### 4.3.2 Többszintű QoS garanciákat nyújtó tárolóméretező módszer

Hasonló tézis mondható ki a kombinált minimális tárolóigény,  $B^*$  kiszámítására is.

**Tézis 3.2.** *A következő formulával becsülhető a kombinált minimális tárolóigény:*

$$B^* \approx \sup_{t>0} \inf_{s>0} \left\{ \frac{\Lambda_{N^*}(s, t) + L}{s} - Ct \right\} \quad (24)$$

ahol  $C$  az átviteli kapacitás,  $s$  és  $t$  szabad paraméter,  $L \in \{F(\gamma, C), \hat{F}(\gamma, C), G(F(\gamma, C)), G(\hat{F}(\gamma, C))\}$  és  $\Lambda_{N^*}(s, t)$  az  $X_{N^*}(0, t]$  kumulánsképző függvénye,  $X_{N^*}(0, t]$  az  $N^*$  forrásból az  $(0, t]$  intervallumban érkező munkamenységet leíró valószínűségi változó. A források számát( $N^*$ )-ot a (23) képlettel javaslom kiszámítani.

A kombinált tárolóméretezési eljárás akkor működik megbízhatóan, ha a folyamatok hívás- és csomagszintű dinamikáját leíró modellek megfelelő pontosságúak és a hálózatba nem engedünk  $N^*$ -nál több folyamatot egyszerre.

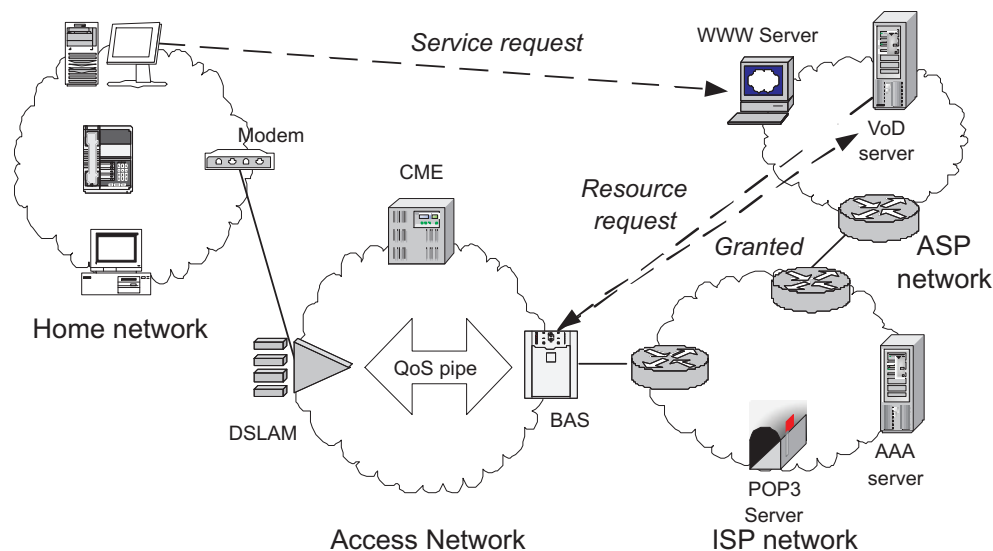


Figure 2: A javasolt QoS architektúra felépítése és működési vázlatja

### 4.3.3 Többszintű QoS garanciákat nyújtó minőségbiztosítási architektúra

**Tézis 3.3.** *Olyan robusztus minőségbiztosítási architektúrát alkottam, amely az előzőekben ismertetett kombinált méretezési eljárások segítségével leegyszerűsíti az erőforráskezeléssel kapcsolatos teendőket és elosztott, transzparens, rugalmas hívásengedélyezési módszer alkalmazásával megbízható, többszintű QoS garanciákat képes nyújtani a hozzáférési hálózatot igénybe vevő értéknövelt szolgáltatások számára.*

Az architektúra egyik legfőbb vezérelve, hogy egy olyan interfészt biztosítson a NAP (Network Access Provider - hozzáférési hálózat üzemeltető) és az ASP (Application Service Provider - Szolgáltatást Nyújtó Fél) hálózata között, amelyen keresztül mind a felhasználó mind az ASP számára transzparens módon lehet QoS kritériumokkal és erőforrásfoglalással kapcsolatos információkat cserélni. A rendszer egyik fő eleme egy központi egység (lásd 2. ábra), amelyet CME-nek (Central Management Entity) nevezünk. Ennek két fő feladata egy logikai ráépülő hálózat létrehozása a fizikai topológia felett, illetve minden erőforrásfoglalással összefüggő tevékenység vezérlése. Az előredefiniált ráépülő hálózat a hozzáférési hálózat határcsomópontjai között létrehozott logikai pont-pont összeköttetésekből áll, amelyeket QoS csöveknek nevezünk, mivel minden QoS osztályhoz és határcsomópont párhoz tartozik egy ilyen csatorna. A minőségbiztosítási architektúra a működéséhez szükséges paraméterek meghatározásánál a napjainkban széleskörben alkalmazott DHCP és RSVP protokollok bizonyos szolgáltatásaira támaszkodik.

A forgalomszabályozási feladatok többségét - a hívásengedélyezést, a forgalom-mérést, forgalomformázást (shaping), forgalomcsonkolást (policing) és számlázást

beleértve - a BAS-ok (Broadband Access Server) látják el. Ezek a csomópontok az NSP és az ISP hálózatának határán találhatóak és így minden forgalom, amely egy bizonyos felhasználóhoz kapcsolható áthalad legalább egy BAS-on. Számuk jóval kevesebb lehet, mint a hozzáférési hálózat másik határán lévő DSLAM-ek (DSL Access Multiplexer) száma, nagyságrendileg annyi BAS egységre van szükség, ahány ISP áll szerződésben a hozzáférési hálózat üzemeltetőjével. A BAS-ok legtöbbször a NAP telephelyén találhatóak és ezek a csomópontok intelligensebbek a DSLAM-eknél. Mindezek fényében kiválóan alkalmasnak tűnnek a fent említett extra feladatok elvégzésére.

A hozzáférési hálózat belsejében ATM vagy Ethernet kapcsolók találhatóak, melyek intelligenciája és funkciókészlete általában limitált. Ezek feladata az egyik portjukon belépő forgalom továbbítása egy másik portjukra. Ezeknek a csomópontoknak képesnek kell lenniük különböző bánásmódban részesíteni a különböző QoS osztályokba tartozó folyamatok adatait. Ugyancsak elvárt funkciójuk ezeknek az eszközöknek, hogy erőforrást tudjanak allokálni a hálózatban kialakított PVP-k (Permanent Virtual Path) vagy VLAN-ok (Virtual Local Area Network) számára. Ez utóbbi funkció egyébként implicit garantomá a korábban említett egymástól független, dedikált erőforrásokkal gazdálkodó QoS csövek megvalósítását.

Az architektúra negyedik építőelemei a már említett DSLAM-ek. Ezek olyan ATM vagy Ethernet kapcsolók, amelyek az előfizetők otthoni hálózata és a hozzáférési hálózat határán találhatóak. Fő feladatuk az egyes felhasználók felől (vagy felé) érkező forgalom aggregálása (vagy szétosztása). Ezeknek a csomópontoknak is tudniuk kell a QoS osztályba tartozás szerint kezelni a forgalmakat és a DSLAM-ek egyben végpontjai is a QoS csöveknek.

Az architektúra működése a következőképpen alakul. Még a hálózat "beindítása" előtt a CME elkészíti a ráépülő hálózatot, tehát létrehozza a QoS csöveket a határ-csomópontok között. Ez gyakorlatilag a DSLAM-ek, kapcsolók és BAS-ok megfelelő felkonfigurálását jelenti. A QoS csövek kezdeti méretét és a hozzájuk rendelt pufferméretet a 3.1 illetve 3.2 tézisekben ismertetett módszerrel számítja ki a CME és ezt az értéket állítja be a hálózat minden érintett csomópontjában.

Amint a logikai ráépülő hálózat elkészült a minőségbiztosítási architektúra működőképes, a továbbiakban felmerülő forgalomszabályozási feladatokat alapvetően a BAS csomópontok látják el. Ezek az eszközök folyamatosan figyelemmel kísérik, mérik az általuk felügyelt QoS csövek aktuális terheltségét és transzparens módon hívásengedélyezési döntéseket hoznak, amikor új igény érkezik. A forgalommérési adatokat a BAS egységek az 1. és 2. téziscsoportban bemutatott ekvivalens kapacitás és minimális tárolóigény becslő eljárások kiértékelésénél használják fel.

A hívásengedélyezési metódus transzparenciáját a BAS egy további funkcióján keresztül lehet megvalósítani. Amennyiben a BAS egységet képessé tesszük DHCP

üzenetek szűrésére és feldolgozására, továbbá RSVP üzenetek elkapására, feldolgozására és megválaszolására a beengedési eljárás átlátszósága garantálható. A DHCP üzenetek felfogásával a BAS egység képes összerendelni az egyes felhasználók IP címét a felhasználó eléréséhez használt DSLAM azonosítójával.

Az RSVP üzenetek elkapásával a BAS a felhasználó helyett tud válaszolni erőforrásfoglalással kapcsolatos kérdésekben. Ez a következő módon valósul meg. A BAS elkapja az előfizetőnek szánt RSVP üzenetet, majd az üzenetben tárolt erőforrásigényt összeveti a megfelelő QoS cső aktuálisan elérhető erőforrásaival. Ezután a BAS a felhasználót megszemélyesítve válaszol az eredeti RSVP üzenetre, amelynek tartalmát az erőforrásigény összevetés eredményével összhangban állítja be, azaz implicit hívásengedélyezést végez.

A hívásengedélyezési folyamat az alábbi modellt követi:

- *1. lépés* Az előfizető az otthoni hálózatának egyik csomópontján keresztül csatlakozni kíván az Internetre, ezért IP címet kér az ISP-jétől a széleskörben elterjedt DHCP protokollon keresztül.
- *2. lépés* A lezajló DHCP kommunikációt megfigyeli az a BAS egység, amely az előfizető ISP-jének hálózata és a hozzáférési hálózat határán található. A BAS összerendeli az előfizető DSLAM azonosítóját és az újonnan kiosztott IP címét.
- *3. lépés* Az előfizető igénybe kíván venni egy az Internetes értéknövelt szolgáltatást, ezért jelzi erre vonatkozó igényét (pl. a megfelelő hyperlinkre kattintással) az ASP hálózatában található web szervernek.
- *4. lépés* A web szerver továbbítja a kérést a tartalomszolgáltató szervernek, amely tudatában van a kért szolgáltatás erőforrásigényének.
- *5. lépés* A tartalomszolgáltató szerver RSVP üzenetet küld az előfizetőnek, amely tartalmazza az adott szolgáltatás erőforrásigényét.
- *6. lépés* Az RSVP üzenetet elkapja a BAS csomópont, az IP cím és QoS osztály alapján megnézi, mely QoS csőben haladna át az új forgalom, végül összeveti az adott csatorna szabad erőforrásait az igényelt erőforrások mértékével.
- *7. lépés* A BAS burkolt (transzparens) hívásengedélyezést végez. A vizsgált QoS cső szabad erőforrásainak függvényében az alábbi három esetet különböztethetjük meg:
  - Amennyiben a vizsgált QoS csőben van elég szabad erőforrás az új forgalom igényének kielégítéséhez, az új forgalom beengedésre kerül, azaz a BAS nyugtázó jelelű RSVP üzenetben válaszol a tartalomszolgáltató szervernek az előfizető nevében. Ezzel párhuzamosan a BAS frissíti az adott QoS cső szabad erőforrásainak értékeit.

- Amennyiben nincs elég szabad erőforrás a vizsgált logikai csatornában, viszont a QoS cső fizikai útvonalán van elég szabad erőforrás a logikai összeköttetés erőforrásainak megfelelő mértékű felduzzasztásához, úgy a BAS kezdeményez(het)i az adott QoS cső átméretezését a CME-től. Amennyiben az átméretezés sikeres, az új forgalom beengedésre kerül, erről a BAS RSVP üzenetben értesíti a tartalomszolgáltató szervert az előfizető nevében. Ezzel párhuzamosan a BAS frissíti az adott QoS cső szabad erőforrásainak értékeit.
- Amennyiben sem a vizsgált QoS csőben, sem annak fizikai útvonalán nincs elég szabad erőforrás az új igény kielégítéséhez, az új folyam hálózatba lépését elutasítja a BAS, erről egy megfelelő tartalmú RSVP üzenetben tájékoztatja a tartalomszolgáltató szervert az előfizető nevében. Természetesen a QoS csövek aktuális szabad erőforrásainak BAS-ban tárolt értéke ebben az esetben változatlan marad.

## 5 Az eredmények alkalmazhatósága

Az első téziscsoportban javasolt ekvivalens kapacitás becslő technikák hatékonyan alkalmazhatók forgalmi linkek méretezési feladatainál, illetve alapjául szolgálhatnak megbízható hívásengedélyezési eljárásoknak. Ahogy azt láttuk, ezek a technikák fontos alkotóelemei a harmadik téziscsoportban ismertetett kombinált méretezési módszereknek is.

A második téziscsoportban tárgyalt formulák aggregált hálózati forgalmak tárolói-gényének közelítésére alkalmasak. Elsősorban ott lehet érdemes ezeket a módszereket használni, ahol a csomagkiszolgáló átviteli kapacitása korlátozott vagy annak bővítése túl költséges (például rádiós linkek esetén). Az itt bemutatott technikák ilyen értelemben az első téziscsoport eredményeinek kiegészítő módszereiként tekinthetők.

A harmadik téziscsoportban ismertetett kombinált méretezési módszerek olyan hálózatméretezési eszközök, mellyekkel a hálózat operátora képessé teheti hálózatát csomagszintű vesztesre és hívásszintű rendelkezésre állásra vonatkozó elvárások egyidejű teljesítésére. Az itt tárgyalt erőforrásigény becslő technikák képezik a 3.3 tézisben bemutatott minőségbiztosítási architektúra alapját. Az ismeretett minőségbiztosítási architektúra célja, hogy megbízhatóan, hatékonyan és transzparens módon megfelelő rendelkezésre állást és átviteli minőséget garantáljon értéknövelt szolgáltatások számára a hozzáférési hálózatban.

Az 1. és 3. téziscsoport eredményei nemzetközi és magyar fórumokon is publikálásra kerültek, míg a 2. téziscsoport eredményeiből jelenleg csak tervezett publikációk vannak. A megjelent publikációk közül a [C3] konferenciacikk "Best paper award"-ot nyert, míg a [J4] folyóiratcikket a Hírközlési és Informatikai Tudományos Egyesület Pollák-Virág díjjal jutalmazta.

## Felhasznált irodalom

- [1] S.R.S. Varadhan, “Asymptotic probability and differential equations”, *Comm. Pure Appl. Math.* vol. 19, pages 261-286., 1966.
- [2] C. Bouchat, S. van den Bosch, T. Pollet, “QoS in DSL Access”, *IEEE Communications Magazine* vol 41. no. 9, Page(s): 108–114, November 2003.
- [3] G. Seres, A. Szlávik, J. Zátanyi and J. Bíró, “Quantifying resource usage - a large deviation-based approach”, *IEICE Transactions on Communications (Japan)*, E85-B(1):938–952, 2002.
- [4] A. György, T. Borsos, “Estimates on the packet loss ratio via queue tail probabilities“, *Global Telecommunications Conference 2001., GLOBECOM '01.*, IEEE, Volume 4, Issue 2001, Page(s): 2407 - 2411 vol.4, 2001.
- [5] N. Likhanov and R. R. Mazumdar, “Cell loss asymptotics in buffers fed with a large number of independent stationary sources“, *INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies*, Volume 1, Page(s):339 - 346, March 1998.
- [6] H.S. Kim, N.B. Shroff, “Loss probability calculations and asymptotic analysis for finite buffer multiplexers“, *IEEE/ACM Transactions on Networking*, Volume 9, Issue 6, Page(s):755 - 768, December 2001.
- [7] F. P. Kelly. “Notes on effective bandwidths” *Stochastic Networks: Theory and Applications*, vol 4, pp. 141–168, Oxford University Press, 1996.
- [8] T. Donald, P. Olivier, J. Roberts, “Dimensioning high speed IP access networks”, *Proceedings of ITC-18*, pages 241–251, Aug. 2003.
- [9] A. K. Erlang, “Solution of some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges”, *Elektroteknikereren*, vol 13, 1917.
- [10] E. Pinsky, A. Conway, Wei Liu, “Blocking formulae for the Engset model ”, *IEEE Transactions on Communications*, Volume: 42, Issue: 6, page(s): 2213-2214, Jun 1994
- [11] L. Kleinrock, “Queueing Systems. Vol I: Theory”, pp. 93–94, John Wiley & Sons, New York - London - Sydney - Toronto, 1975.
- [12] S.Jamin , P. B. Danzig, S. J. Shenker, L. Zhang , “A measurement-based admission control algorithm for integrated service packet networks”, *IEEE/ACM Transactions on Networking (TON)*, Volume 5 , Issue 1, Pages: 56 - 70, 1997

- [13] M. Montgomery and G. de Veciana, “On the relevance of time scales in performance oriented traffic characterizations”, *Conference on Computer Communications (IEEE Infocom)*, Volume 2, Page(s): 513 - 520, March 1996.
- [14] C. Courcoubetis, V. A. Siris and G. D. Stamoulis, “Application of the many sources asymptotic and effective bandwidths to traffic engineering”, *Telecommunication systems*, 12:167-191, 1999.
- [15] R. R. Bahadur and R. Rao, “On deviations of the sample mean”, *The Annals of Mathematical Statistics*, Vol. 31, No. 4, page(s): 1015-1027. December 1960.
- [16] D. D. Botvich and N.G. Duffield, “Large deviations, the shape of the loss curve, and economics of scale in large multiplexers”, *Queueing Systems*, Vol. 20, page(s): 293-320, 1995.
- [17] C. Courcoubetis and R. Weber, “Buffer overflow asymptotics for a buffer handling many traffic sources”, *Journal of Applied Probability*, Vol. 33, page(s): 886-903, 1996.
- [18] J. Bíró, Z. Heszberger, M. Martinecz, Sz. M. Kis, “Novel Equivalent Capacity Approximation Through Asymptotic Loss Analysis”, *HET-NETs '04, First International Working Conference on Performance Modelling and Evaluation of Heterogenous Networks*, page(s): 123-129, July 2004.
- [19] A. K. Parekh, R. G. Gallager, “A Generalized Processor Sharing approach to Flow control in Integrated Services Networks: The Single Node Case”, *IEEE/ACM Transactions on Networking (TON)*, Volume 1, Issue 3, Pages: 344 - 357, 1993
- [20] J.-Y. Le Boudec, “Application of Network Calculus to Guaranteed Service Networks”, *IEEE Transactions on Information Theory*, 44(3):1087-1096, May 1998.
- [21] J.-Y. Le Boudec and P. Thiran, “Network Calculus: A Theory of Deterministic Queueing Systems for the Internet”, *LNCS: Lecture Notes in Computer Science*, Springer, 2001.
- [22] R. L. Cruz, “A Calculus for Network Delay. Part I: Network Elements in Isolation and Part II: Network Analysis” , *IEEE Transactions on Information Theory*, 37(1):114-141, Jan. 1991.
- [23] M. Fidler, “An End-to-End Probabilistic Network Calculus with Moment Generating Functions ”, *IEEE International Workshop on Quality of Service (IWQoS)*, Jun. 2006.



- [24] H. Bruneel, B. G. Kim, “Discrete-time Models for Communication Systems Including ATM”, *Kluwer Academic Publishers*, ISBN:0792392922, 1993.
- [25] G. Appenzeller, I. Keslassy, N. McKeown, “Sizing Router Buffers”, *ACM SIGCOMM*, Volume 34, Issue 4, Pages: 281 - 292, October 2004
- [26] A. Ganesh, N. O’Connell, D. Wischik, “Big Queues (Lecture Notes in Mathematics)”, *Springer*, ISBN:3540209123, May 1, 2008
- [27] C. A. Courcoubetis, A. Dimakis, G. D. Stamoulis, “Traffic equivalence and substitution in a multiplexer with applications to dynamic available capacity estimation”, *IEEE/ACM Transactions on Networking*, Volume 10, Issue 2, page(s) 217-231, ISSN: 1063-6692, Apr 2002
- [28] G. Mao and D. Habibi. “Loss performance analysis for heterogeneous on-off sources with application to connection admission control” *IEEE/ACM Transactions on Networking (TON)*, 10(1):125–138, 2002.
- [29] I. Norros, “A storage model with self-similar input”, *Queuing Systems*, Volume 16, Numbers 3-4, pages 387-396, September, 1994
- [30] R. J. Gibbens, Y. C. Teh, “Critical time and space scales for statistical multiplexing in multi service networks”, *Proceedings of the International Teletraffic Congress ITC’16*, pages 87-96, Edinburgh, Scotland, 1999.
- [31] Z. Heszberger, J. Zátónyi, and J. Bíró, “Performance bounds for rate envelope multiplexing”, *Performance Evaluation*, Volume 48, Issue 1, May 2002, Pages 87-101. Special Issue on ATM & IP Networks: Performance Modelling and Analysis
- [32] F.H.P. Fitzek and M. Reisslein, “MPEG-4 and H.263 Video Traces for Network Performance Evaluation”, *IEEE Network*, 15(6):40-54., 2001
- [33] András Kern, Mátyás Martinecz and Zalán Heszberger, “Architecture and Configuration of Broadband Access Networks Supporting Multimedia Applications”, *The Tenth IEEE Symposium on Computers and Communications*, June. 2005
- [34] “Dynamic Host Configuration Protocol (DHCP)”, *RFC 2131*, [www.ietf.org](http://www.ietf.org)
- [35] “Resource ReSerVation Protocol (RSVP)”, *RFC 2205*, [www.ietf.org](http://www.ietf.org)
- [36] “An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks”, *RFC 2205*, [www.ietf.org](http://www.ietf.org)

## Publikációk

### Folyóirat cikkek

- [J1] M. Martinecz, A. Kern, Z. Heszberger and J. Bíró, "Architecture and Configuration of Broadband Access Networks Supporting Multimedia Applications", *IJCA Journal Special Issue*. pages 34–46, April, 2007.
  
- [J2] M. Martinecz, J. Bíró and Z. Heszberger, "Novel Techniques for Assessing Resource Requirements in Packet-based Networks", *Híradástechnika - Telecommunications - Hungarian periodical, Selected Papers*. pages 24–29, June, 2005.
  
- [J3] J. Bíró, A. Gulyás and M. Martinecz, "Parsimonious Estimates of Bandwidth Requirement in Quality of Service Packet Networks", *Performance Evaluation*, Volume 59, Issues 2-3, p. 159-178, February 2005.
  
- [J4] M. Martinecz, J. Bíró and Z. Heszberger, "Újszerű erőforrásigény-becslő módszerek csomagkapcsolt hálózatokban", *Híradástechnika - Telecommunications - Hungarian periodical*. pages 13–18, Sept, 2004.

### Konferencia cikkek

- [C1] M. Martinecz, J. Bíró, Z. Heszberger, "A Logical Channel Dimensioning Method Enabling Simple QoS Provision in Aggregation Networks", *World Telecommunications Congress (WTC) 2006*, Budapest, Hungary, May 2006.
  
- [C2] M. Martinecz, A. Kern, J. Bíró, "Multi-level QoS Guarantees in Broadband Access Networks", In *HET-NETs '05, First International Working Conference on Performance Modelling and Evaluation of Heterogenous Networks*, Ilkley, West Yorkshire, U.K., July 2005.
  
- [C3] A. Kern, M. Martinecz and Z. Heszberger, "Architecture and Configuration of Broadband Access Networks", In *Proc. of ISCC'2005, The Tenth IEEE Symposium on Computers and Communications*, volume 1, pages 172–181, La Manga del Mar Menor, Cartagena, Spain, June, 2005.

- [C4] M. Martinecz, J. Bíró, Z. Heszberger, Sz. M. Kis, "Novel Equivalent Capacity Approximation Through Asymptotic Loss Analysis", In *HET-NETs '04, First International Working Conference on Performance Modelling and Evaluation of Heterogenous Networks*, Ilkley, West Yorkshire, U.K., July 2004.
- [C5] J. Bíró, Z. Heszberger, M. Martinecz, N. Felicián and Octavian Pop, "Towards a Framework of QoS Measure Estimates for Packet-Based Networks", In Proc. of *IEEE International Conference on Communications - ICC 2004*, volume 4, pages 2231–2235, Paris, France, June, 2004.
- [C6] J. Bíró, Z. Heszberger and M. Martinecz, "A Family of Performance Bounds for QoS Measures in Packet-Based Networks", *Lecture Notes in Computer Science - LNCS 3042 (Proc of IFIP Networking 2004)*, pages 1108–1119, Athens, Greece, May, 2004.
- [C7] J. Bíró, Z. Heszberger, and M. Martinecz, "Equivalent Capacity Estimators for Bufferless Fluid Flow Multiplexing", In Proc. of *IEEE Global Telecommunications Conference - Globecom 2003*, volume 7, pages 3706–3710, San Francisco, CA, USA, Dec. 2003.
- [C8] J. Bíró, Z. Heszberger, T. Dreilinger, A. Gulyás, and M. Martinecz, "Parsimonious Estimates of Bandwidth Requirement in Quality of Service Packet Networks", In Proc. of *First International Working Conference on Performance Modeling and Evaluation of Heterogenous Networks - HET-NETs '03*, pages 69/1-69/9, Ilkley, UK, 21-23 July 2003.
- [C9] J. Bíró, Z. Heszberger, F. Németh, M. Martinecz, "Bandwidth Requirement Estimators for QoS Guaranteed Packet Networks", In Proc. of *International Network Optimization Conference - INOC 2003*, pages 41–49, Evry-Paris, France, May, 2003.
- [C10] J. Bíró, Z. Heszberger, G. Kún and M. Martinecz, "Advanced QoS Provision for Real-Time Internet Traffic", In *IEEE International Packet Video Workshop - Packet Video 2003*, pages 26–28, Nantes, France, April, 2003.

- [C11] J. Bíró, Z. Heszberger, M. Martinecz, "Measurement-based QoS Provision for VoIP Traffic", In *Proceedings of Transcom 2003 Conference*, pages 195-200, Zilina, Slovakia, June 2003.
- [C12] J. Zátanyi, Z. Heszberger, J. Bíró, M. Martinecz and B. Héjj, "Providing Guaranteed QoS for Real-Time Internet Traffic", In *Proceedings of Polish-Czech-Hungarian Workshop on Circuit Theory and Applications*, pages 65-70, Warsaw, Poland, September 2002.
- [C13] B. Héjj, M. Martinecz, "Minőségbiztosított forgalmak elemzése és szimulációs vizsgálata IP hálózatokban", In *Proceedings of National Student Conference on Information Sciences (OTDK)*, Győr, April 2003.