## HY537: Έλεγχος Πόρων και Επίδοση σε Ευρυζωνικά Δίκτυα

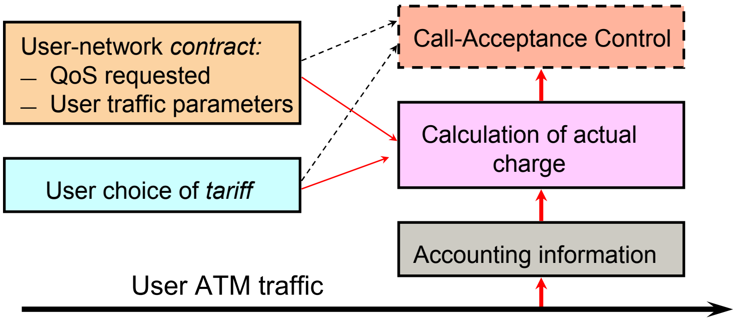
Βασίλειος Σύρης

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης
Εαρινό εξάμηνο 2008

## Contents

## Outline of the VBR-Charging Approach

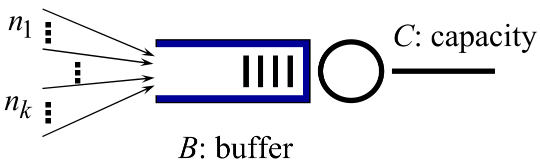
- Total charge for a call is $a(\ldots)T+b(\ldots)V+c(\ldots)$ where
  - $T$ = duration of call (e.g. seconds)
  - $V$ = volume of call (e.g. Mbits or Mcells)
  - $a(\ldots)$, $b(\ldots)$, $c(\ldots)$ capture QoS, UNI choices (peak rate, etc), user's choice of tariff






## Effective Bandwidths Reminder

- $k$ traffic classes
- class $i$ contributes $n_i$ sources




- Which $(n_1,\ldots,n_k)$ do not violate QoS constraint (e.g. CLP $\leq e^{-\gamma}$) ?

➔ **Acceptance Region:** linear constraints of the form

$$n_1 \cdot \alpha_1 + \ldots + n_k \cdot \alpha_k \leq C^* \quad \text{where} \quad \alpha_j(s,t) = \frac{1}{st} \log E\left[e^{sX_j[0,t]}\right]$$

## Effective Bandwidth Formula

- Effective bandwidth of a source of type $j$

$$\alpha_j(s,t) = \frac{1}{st} \log E\left[e^{sX_j[0,t]}\right]$$

$X_j[0,t]$: load produced by source of type $j$ in window $t$

- The effective bandwidth $\alpha_j(s,t)$ quantifies resource usage for a *particular operating point* $(s,t)$
- Effective bandwidth for connection k of type j can be computed using empirical estimate

$$\hat{\alpha}_{jk}(s,t) = \frac{1}{st} \log\left( \frac{1}{T_k} \sum_{i=1}^{T_k/t} e^{sX_{jk}[(i-1)t,it]} \right)$$

  - Tk: duration of connection k
  - t: interval length

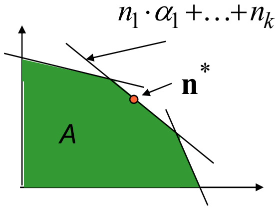


## Economic theory reminder

- Social welfare maximisation:

$$\max_{\mathbf{n}} U(\mathbf{n})$$
$$s.t.\ \mathbf{n} \in A$$

$$n_1 \cdot \alpha_1 + \ldots + n_k \cdot \alpha_k \le C^*$$




- Prices defined by shadow costs of effective bandwidth constraints

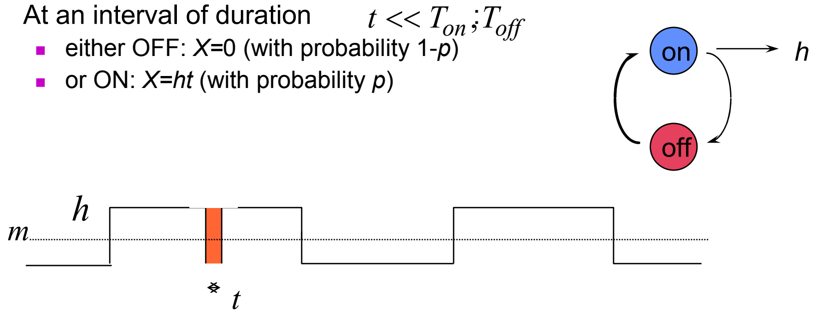
$$\frac{p_i}{p_j} = \frac{eb_i}{eb_j}$$

- Pricing in proportion to effective bandwidths:
  - captures QoS requirements
  - incentive compatible
  - content independent
  - well understood theory

➔ IS IT OBVIOUS?

## The "slow" ON/OFF Source Model

- At an interval of duration $t << T_{on}; T_{off}$
  - either OFF: *X*=0 (with probability 1-*p*)
  - or ON: *X*=*ht* (with probability *p*)




- mean rate $m = p*h$
- effective bandwidth $\alpha_{on/off}(s,t) = \frac{1}{st}\log\left[1+\frac{m}{h}\left(e^{sth}-1\right)\right]$

➔ **ON/OFF has the *largest* possible effective bandwidth (over *all* traffic models) for *given m* and *h* and operating point**

Guaranteed Services- 7

## Charging VBR Services - Idea 1

- Apply CAC with effective bandwidths

and

- Charge (per unit time) each VBR call proportionally to effective bandwidth
  - Pro: Theoretically justifiable and fair
  - Con: Requires accurate *a priori* knowledge of complex traffic statistics
    - If *empirical* estimate from sampling previous calls is used, users have the *wrong* incentive to reach this value
    - user will tend to *overload* the network => similar to overeating in all-you-can-eat restaurants

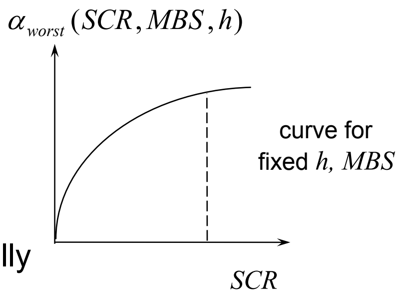
Guaranteed Services- 8

## Charging VBR Services - Idea 2

- Use traffic contracts specifying:
  - peak rate
  - sustainable cell rate, max burst size
- Apply CAC with the *worst*-case effective bandwidths for contracts

and

- Charge each VBR call proportionally to worst-case effective bandwidth




- Pro: No complicated traffic statistics
- Cons: *Unfair* to users with mean rate < SCR, or peak rate < *h*
  Provides the *wrong* incentive to increase traffic



## Discussion

- Previous approaches only based on *static* a priori variables, determined by traffic contract
  - Provided *wrong* incentives to *exhaust* the resource usage permissible by contract
- Charging should also employ *dynamic* a posteriori variables, *measured* during call

## Charging VBR Services - Idea 3

- *Measure* the effective bandwidth during each VBR call, and charge proportionally to it
  - Pro: Provides the *right* incentives
  - Con: Incompatible with *static CAC*
    - If eb estimated empirically, difficult for users to understand charges
    - users may possibly *not* pay for all the resources *reserved* by CAC
      e.g. a user with no traffic, would face 0 charge

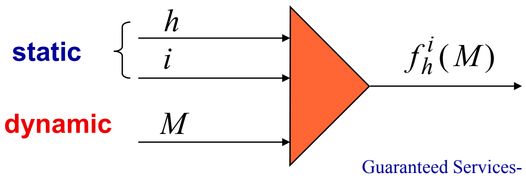


## What is Needed ?

- Charge according to both:
  - static variables reflecting traffic contract and resources reserved by CAC
  - dynamic variables reflecting actual usage
- Final charge should be close to *actual* effective bandwidth
- Allow the user to select a tariff:
  - selection should reveal some important additional *information* to the network

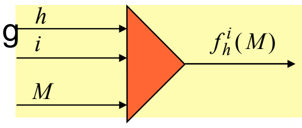
## Defining the problem

- **A reasonable model**
  - *H* = set of traffic contracts
  - *M* = traffic parameters measured (mean, etc.)
  - for each $h \in H$ the available tariffs are
    - functions of measurements *M* , indexed by *i* $F_h = \{f_h^i(M)\}$
- **User chooses at call set up**
  - the traffic contract *h*
  - the best tariff from $F_h$ by solving $\arg\min_i Ef_h^i(M)$
- Network computes at end of call






## Defining the problem (cont.)

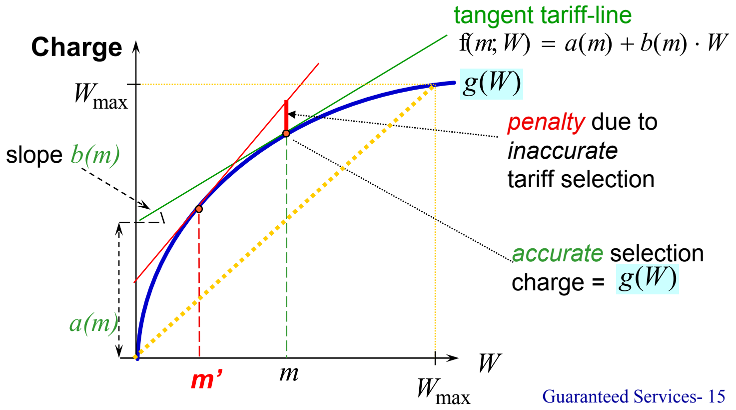



- What is a reasonable choice of charging functions $f_h^i$
  - charge corresponds to effective bandwidth
  - choice of tariff **reveals** useful information
  - required measurements are simple (eg. T,V)
- Not obvious!!
  - *Use only static parameters: f* = worst case eb (*peak rate*)
    - bad incentives, “all you can eat” problem
  - *Use only measurements*: *f* = actual eb of traffic sent
    - no account of resource reservation at call set up
  - **Combine static with dynamic parameters**
    - *f* = worst case eb given *h*, *M*

# An interesting structural property for tariffs

- Boxes arrive to be packed
  - a priori info: $W_{max}$
  - a posterior info: $W$
- **Can tariffs make users reveal $W$ ?**

⇨ **charge = concave function of weight**




Guaranteed Services- 15

# Application: Simple Charging Scheme for VBR

- Charging of VBR calls
  - a priori information: *peak* rate *h* (policed)
  - a posterior information: *mean* rate *M* (measured)
- Use as charging function an effective bandwidth *eb*(*M*)
  - for example,
  - concave in *M*
  - charge = worst case eb given *h*, *M*

$$\alpha_{on/off}(M) = \frac{1}{st}[1 + \frac{M}{h}(e^{ht} - 1)]$$

- Construct tariffs as tangents to points *m* of the eb curve

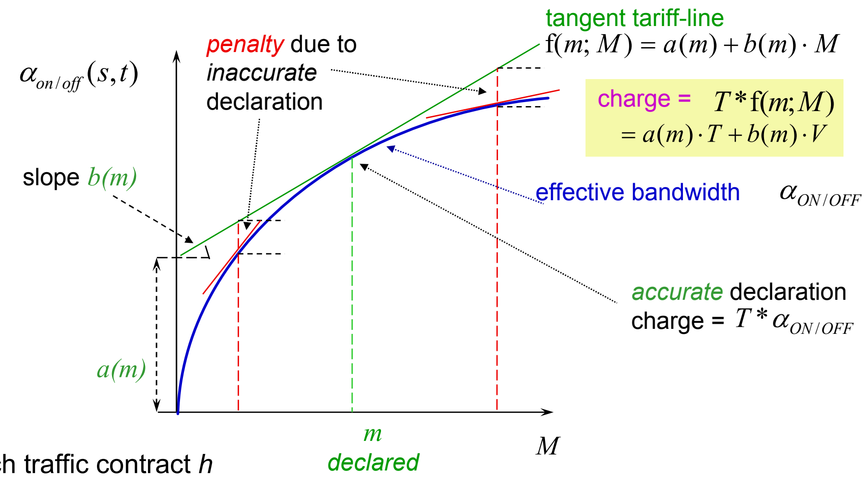
$$f(m; M) = a(m) + b(m) \cdot M$$

- $M$ → *actual* mean rate, measured as $M = V/T$ = *Volume / Time*
- $m$ → user *prediction* of mean rate - defines tariff

- Implicit declaration of *M* by users

Guaranteed Services- 16

## Simple Charging Scheme for VBR




- Each traffic contract *h* defines a family of tariff lines



## Properties of Simple Charging Scheme

- Total charge $= T \cdot [a(m) + b(m) \cdot M]$

$$= a(m) \cdot T + b(m) \cdot V$$

- Accounts both for
  - resource reservation => time-component
  - actual usage => volume-component
- Simple Accounting
  - Requires only *simple* measurements: *T* and *V*
- **Flexibility added to traffic contracts**
- Rational users pay in proportion to their effective use
  - Tariff coefficients depend on traffic contract parameters

## Incentives to the User

- Provides the user with the *right incentives.* In particular:
  - Incentive to accurately declare the mean rate $M$, if known a priori
  - For random mean rate $M$ : *Expected* charge is *minimised* for

    $$m=E[M]$$

    → user has the incentive to *estimate* this (from empirical information), and declare it to the network

  - Incentive to *shape* traffic, thus reducing peak rate $h$ and the charge



## Incentive Compatibility

- User's optimal declaration of $m$ is *informative* to the network provider

→ Can be used by network provider in more efficient allocation of resources, thus improving operation of the network

- User's incentive to shape traffic reduces burstiness, thus also leading to more efficient operation of the network

## Computation of *a(m)* and *b(m)*

- Both *a(m)* and *b(m)* can be expressed in *closed* form

  If t=1, then

$$b(m) = \frac{e^{sh} - 1}{s\left[h + m\left(e^{sh} - 1\right)\right]}, \qquad a(m) = \alpha_{\text{on-off}}(m) - mb(m)$$

- Appropriate values of *s,t* can be derived numerically



## Examples of Tariffs

| $h = 3$ Mbps $st = 1$ sec/Mbit | | |
|---|---|---|
| *M* | *a(m)* | *b(m)* |
| 0.20 Mbps | 0.26 | 2.80 |
| 0.75 Mbps | 0.93 | 1.10 |
| 1.50 Mbps | 1.46 | 0.60 |
| 2.25 Mbps | 1.81 | 0.41 |
| 2.80 Mbps | 1.98 | 0.34 |

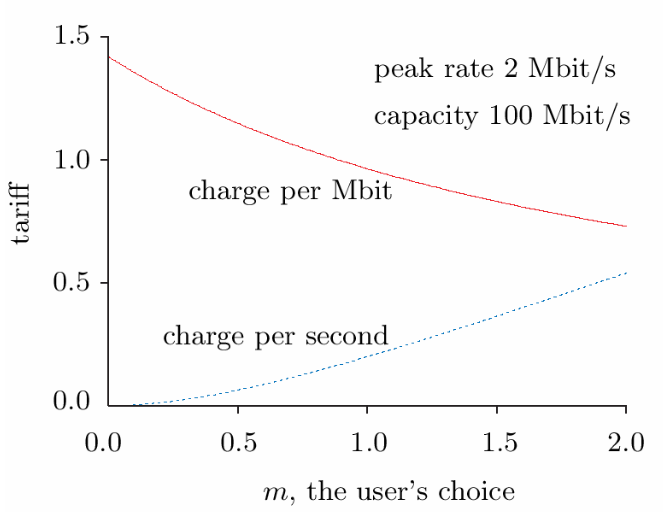
| $h = 1.5$ Mbps $st = 1$ sec/Mbit | | |
|---|---|---|
| *M* | *a(m)* | *b(m)* |
| 0.20 Mbps | 0.06 | 1.59 |
| 0.75 Mbps | 0.37 | 0.85 |
| 1.50 Mbps | 0.72 | 0.52 |

| $h = 3$ Mbps $st = 2$ sec/Mbit | | |
|---|---|---|
| *M* | *a(m)* | *b(m)* |
| 0.20 Mbps | 1.18 | 2.41 |
| 0.75 Mbps | 1.82 | 0.66 |
| 1.50 Mbps | 2.16 | 0.33 |
| 2.25 Mbps | 2.36 | 0.22 |
| 2.80 Mbps | 2.46 | 0.18 |

***a(m)* => $/sec**

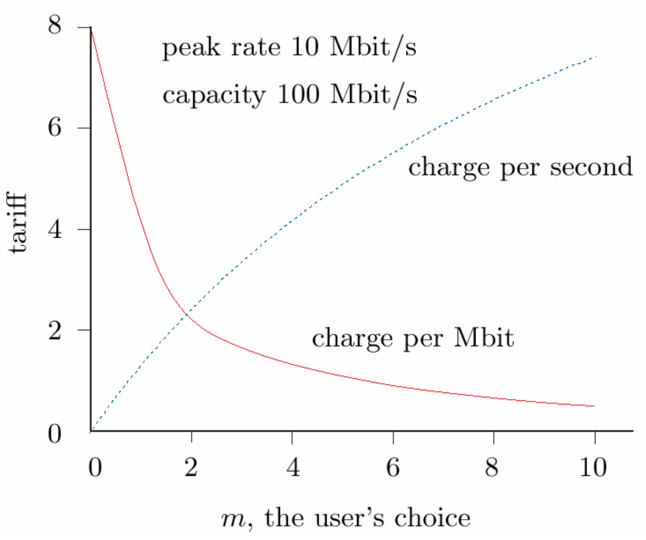
***b(m)* => $/Mbit**

## How *a(m)* and *b(m)* Vary




Guaranteed Services- 23

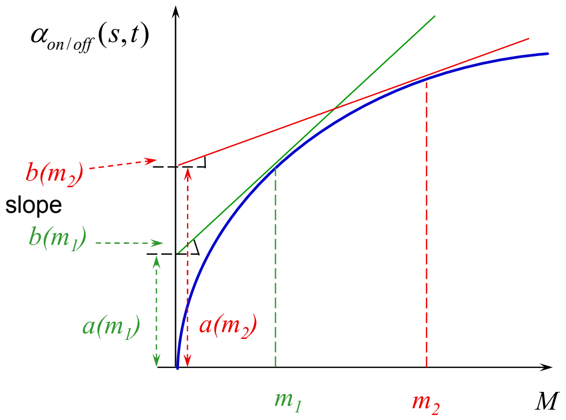
## How *a(m)* and *b(m)* Vary




Guaranteed Services- 24

## How $a(m)$ and $b(m)$ Vary

- For fixed $h, s, t$, as $m$ increases:
  - $a(m)$ *increases*
  - $b(m)$ *decreases*
  - ➔ charge for time *increases*, while charge for volume *decreases*, because the ability for multiplexing diminishes

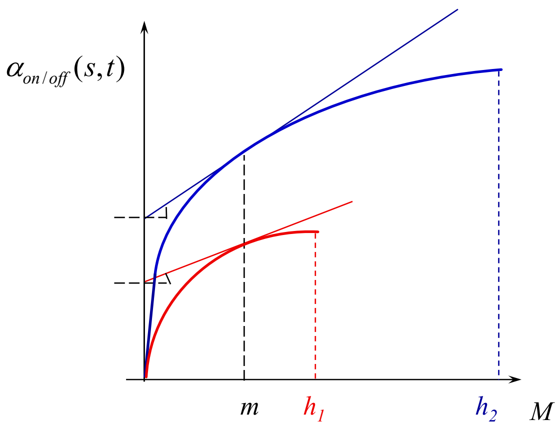



-----



## How $a(m)$ and $b(m)$ Vary (continued)

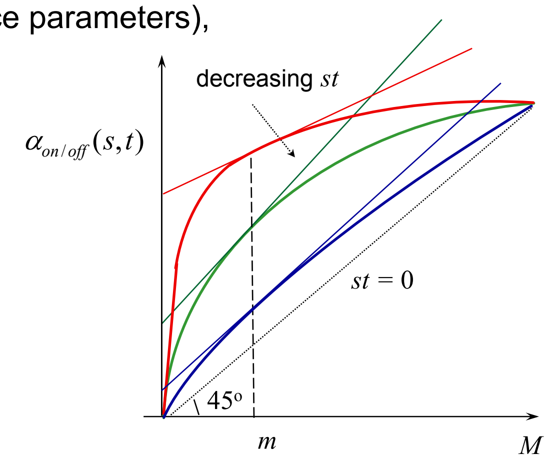
- For fixed $m, s, t$, as $h$ increases :
  - both $a(m)$ and $b(m)$ *increase*
  - ➔ the source is *more bursty*, thus reserving and using more resources




-----

## How $a(m)$ and $b(m)$ Vary (continued)

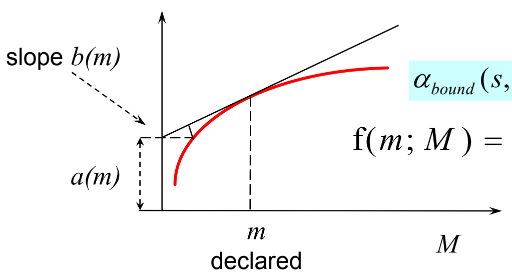
- For fixed $m$, $h$ (i.e. source parameters), when $st$ decreases :
  - $a(m)$ *decreases*
  - $b(m)$ *increases* and then *decreases* to 1
  - ➔ for small $st$ time-charge *decreases*, and volume-charge *increases*, because the ability for multiplexing increases




Guaranteed Services- 27

## Improving accuracy of Simple Charging Scheme

- The simple charging scheme bounds the effective bandwidth according to the ON/OFF bound
  - does not capture general traffic contracts for VBR
- *Other* bounds can also be used
  - **functions of mean rate *and the LBs* of the traffic contract**
  - *Same* approach: charge per unit time derived according to the *tangent* selected by the user



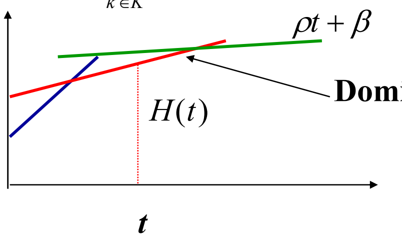

$$f(m; M) = a(m) + b(m) \cdot M$$

Guaranteed Services- 28

## Taking into account leaky bucket constraints

- ON/OFF bound corresponds to a single leaky bucket, constraining **only** the peak rate $\alpha_{on/off}(s,t)=\frac{1}{st}\log\left[1+\frac{m}{h}\left(e^{sth}-1\right)\right]$
- For traffic contracts involving multiple leaky buckets, we can use the *tighter* bound $\alpha_{lb}(s,t)=\frac{1}{st}\log\left[1+\frac{tm}{H(t)}\left(e^{sH(t)}-1\right)\right]$

  where $H(t):=\min_{k\in K}\{\rho_k t+\beta_k\}$

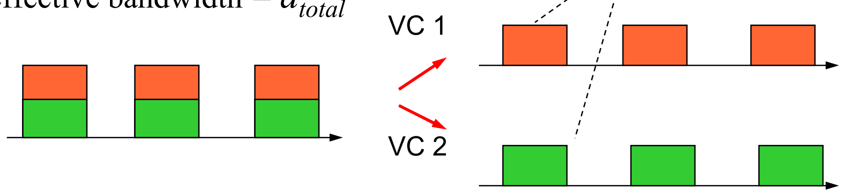





## Splitting of traffic

peak rate = $h$ mean rate = $m$
effective bandwidth = $a_{total}$

peak rate = $h/2$ , mean rate = $m/2$
effective bandwidth = $a_{split}$




- Splitting can be *beneficial* to the user => possibly less total charge, because

$$2a_{split} < a_{total}$$

- correlated traffic streams are erroneously charged as independent ones

## Discouraging splitting - fixed charge

- Traffic splitting is undesirable to provider, because:
  - may lead to reduced revenue
  - set of available VPI/VCI may be exhausted
  - increased signalling overhead for setting more VCs

→ **Splitting should be discouraged => add a *fixed* charge per VC**

→ Total Charge = $a(m) \cdot T + b(m) \cdot V + c(m)$

- However, traffic splitting could be beneficial to provider, if substreams can only be accommodated through *different routes*



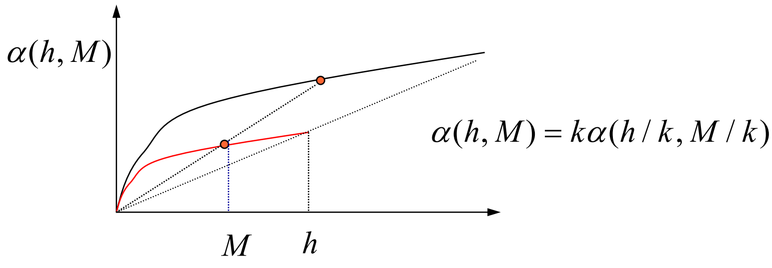
## Examples of $a$, $b$, $c$ Tariffs

| $h = 3$ Mbps $st = 1$ sec/Mbit | | | |
|---|---|---|---|
| *M* | *a(m)* | *b(m)* | *c(m)* |
| 0.20 Mbps | 0.26 | 2.80 | 1.26 |
| 0.75 Mbps | 0.93 | 1.10 | 5.65 |
| 1.50 Mbps | 1.46 | 0.60 | 8.30 |
| 2.25 Mbps | 1.81 | 0.41 | 10.05 |
| 2.80 Mbps | 1.98 | 0.34 | 10.90 |

- Fixed charge *c(m)*
  - is expressed in $
  - was taken (in the examples) as *a(m)**5sec+1

## Discouraging splitting of traffic (cont.)

- Use **homothetic** tariffs




- **Pros:** convexity makes users reveal their mean rates, no incentive to split
- **Cons:** charge not proportional to eb (but close!)



## Simpler Charging: Dispensing with Duration

- **The time-component of charge can be eliminated**
  - total charge = $b \cdot V + c$
  - tariff will be simpler
  - dependence of usage-charge on QoS will be clearer
- **Reasoning:**
  - *c* can be set to account for *typical time-charge*, or
  - we can assume a typical value for *m* and infer $T \approx V/m$, hence

$$a(m) \cdot T + b(m) \cdot V + c(m) \approx a(m) \cdot (V/m) + b(m) \cdot V + c(m) = b'(m) \cdot V + c(m)$$

- **However, users will have *no* incentive to close connections**
  - set of available VPI/VCI may be exhausted
  - provider can limit the maximum number of VPI/VCIs permissible per user

## Charging CBR Services

- **Simple charging scheme can also be applied to CBR services**
  - users should declare $m = h$
  - Total Charge = $a(h)\cdot T + b(h)\cdot V + c(h)$
  - Volume-charge does *not* vanish, because $b(h) \neq 0$
- CBR services should be charged *only* on the basis of time, if their peak rate is really *reserved*, and CBR is *not* multiplexed statistically
  - simpler scheme
  - already adopted in practice

Guaranteed Services- 35

## Charging PVCs

- So far have only dealt with *Switched* VCs for VBR services (SVCs)
- Simple charging scheme can also be applied to *Permanent* VCs (PVCs) for VBR services
- However, PVCs can also be charged *only* on the basis of time, if they are *not* multiplexed statistically, due to their long duration
  - simpler scheme
  - already adopted in practice

Guaranteed Services- 36

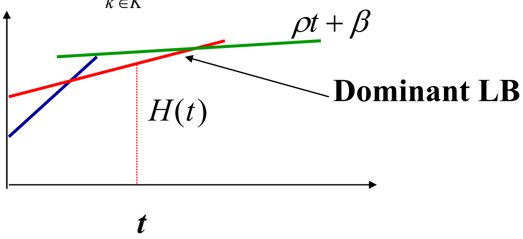
## Taking into account leaky bucket constraints

- ON/OFF bound corresponds to a single leaky bucket, constraining **only** the peak rate

$$\alpha_{on/off}(s,t) = \frac{1}{st}\log\left[1+\frac{m}{h}\left(e^{sth}-1\right)\right]$$

- For traffic contracts involving multiple leaky buckets, we can use the *tighter* bound $\alpha_{lb}(s,t) = \frac{1}{st}\log\left[1+\frac{tm}{H(t)}\left(e^{sH(t)}-1\right)\right]$

where $H(t) := \min_{k\in K}\{\rho_k t + \beta_k\}$






## Effective bandwidth of a traffic contract

- "Simple" bound of a traffic contract's effective bandwidth:

$$\alpha(SCR, PCR, MBS) = \frac{1}{st}\log\left[1+\frac{tSCR}{H(t)}\left(e^{sH(t)}-1\right)\right]$$

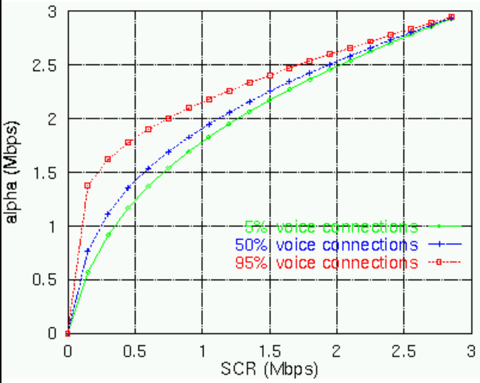
where

$$H(t) = \min\{\rho_0 t + \beta_0, \rho_1 t + \beta_1\}$$

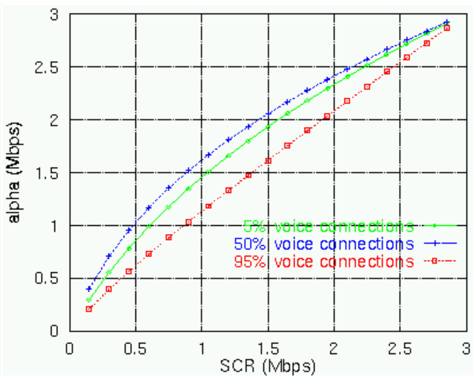
$$(\rho_0, \beta_0) = (PCR, 0)$$

$$(\rho_1, \beta_1) = (SCR, (MBS-1)(1-SCR/PCR)+1)$$

# Effect of traffic mix and link capacity




C=34 Mb/s

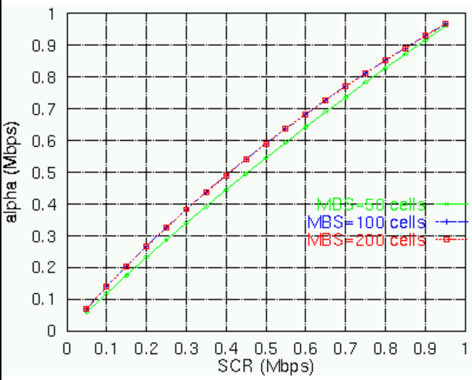



C=155 Mb/s

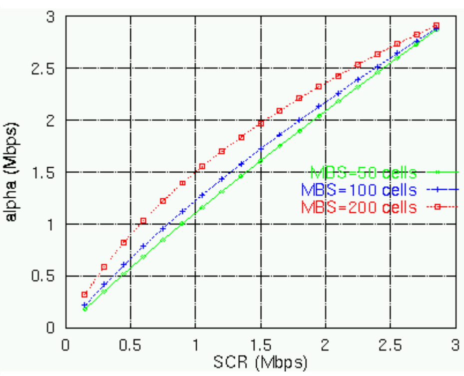
- PCR=3 Mbit/s, MBS=200
- Buffer=4 msec, MPEG-1 & voice traffic



# Effect of PCR, SCR, MBS




PCR=1 Mb/s




PCR=3 Mb/s

- C=155 Mbit/s, Buffer=4 msec, 20% voice connections

## British Telecom's (BT) tariffs

- Based on price "multipliers"

$$P_{VBR} = M_1(PCR/SCR) \cdot M_2(PCR/SCR, MBS) \cdot P_{CBR}$$

*(a) Multiplier $M_1$ for the burst ratio* PCR/SCR.

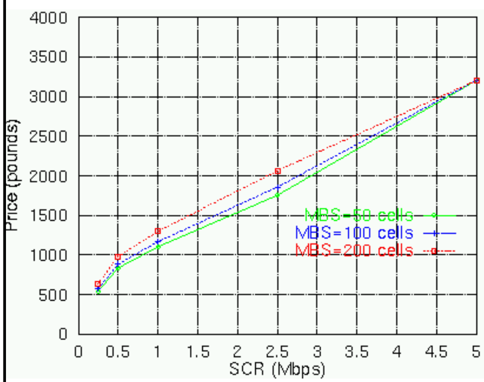
| Burst ratio | 1 | 2 | 5 | 10 | 15 | 20 |
|---|---|---|---|---|---|---|
| Multiplier | 0.9 | 1.1 | 1.5 | 2.0 | 2.3 | 2.5 |

*(b) Multiplier $M_2$ for* MBS.

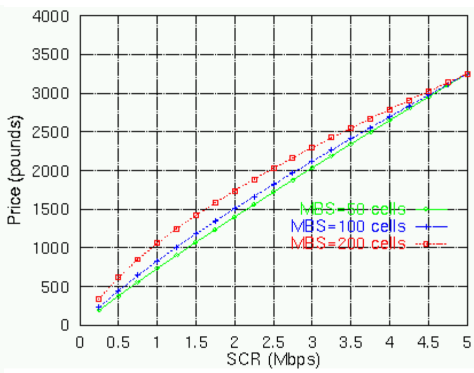
| *MBS* (in cells) | 50 | 100 | 200 |
|---|---|---|---|
| Multiplier ($PCR/SCR \geq 1.8$) | 0.85 | 0.9 | 1.0 |
| Multiplier ($PCR/SCR < 1.8$) | 1.0 | 1.0 | 1.0 |

Guaranteed Services- 41

## BT vs. EB tariffs




BT

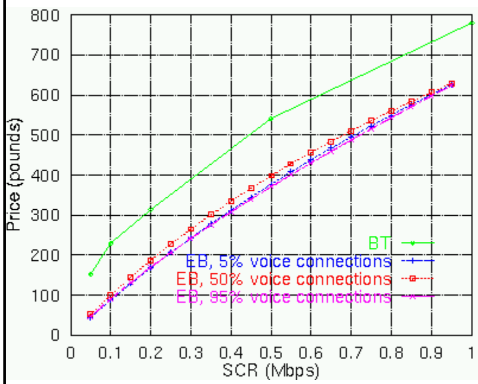



EB

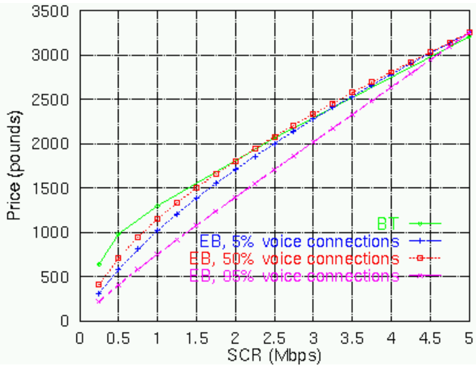
- PCR=5 Mbit/s
- C=155 Mbit/s, Buffer=4 msec, mix with 20% voice traffic

Guaranteed Services- 42

# BT vs. EB tariffs (2)




PCR=1 Mbit/s

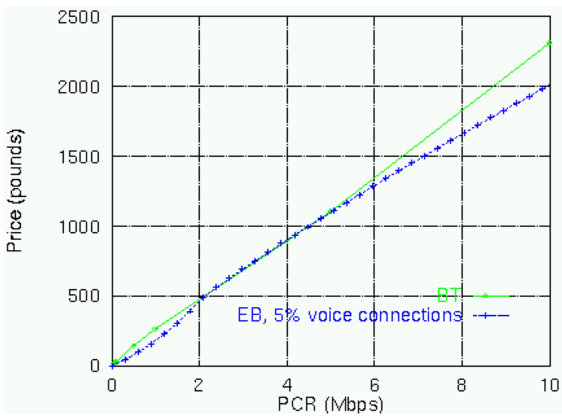



PCR=5 Mbit/s

- MBS=200
- C=155 Mbit/s, Buffer=4 msec

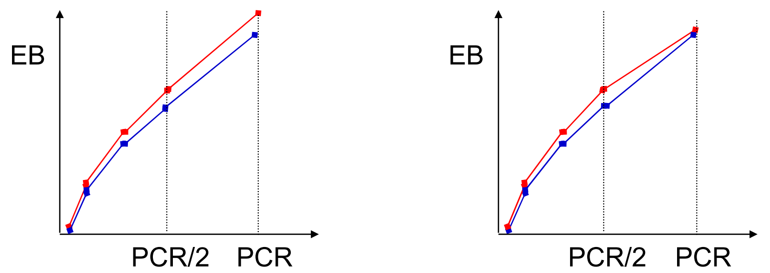


# BT vs. EB tariffs (3)




- PCR/SCR=5, MBS=200
- C=155 Mbit/s, Buffer=4 msec, mix with 20% voice traffic

## The PCR/SCR<1.8 condition




- Left graph is <u>without</u> the condition:
  "if PCR/SCR<1.8 then price as MBS=200"

Guaranteed Services- 45

## Χρέωση ATM VBR VCs βάσει χρόνου

- Τέλος: $a(\mathbf{x})\cdot T$
  - $\mathbf{x} \leftarrow$ παράμετροι σύνδεσης, $T$ = διάρκεια της υπηρεσίας
- $a(\mathbf{x})$: ι.ε.ζ. της μέγιστης κίνησης σύμφωνη με το συμβόλαιο $\mathbf{x}$
  - Για μεγάλο βαθμό στατιστικής πολυπλεξίας, η χρήση πόρων εξαρτάται από το *PCR*, *SCR* μόνο μέσω του λόγου τους *PCR/SCR*
- Υπολογισμός τελών σύνδεσης VBR βάσει των τελών σύνδεσης CBR με την χρήση πολλαπλασιαστών

$$a_{\mathrm{VBR}}(PCR,SCR,MBS) \approx M_1(PCR/SCR)\cdot M_2(MBS,PCR/SCR)\cdot a_{\mathrm{CBR}}(SCR)$$

$M_1$: πολλαπλασιαστής του λόγου εκρηκτικότητας *PCR/SCR*
$M_2$: πολλαπλασιαστής του μεγέθους έκρηξης *MBS*

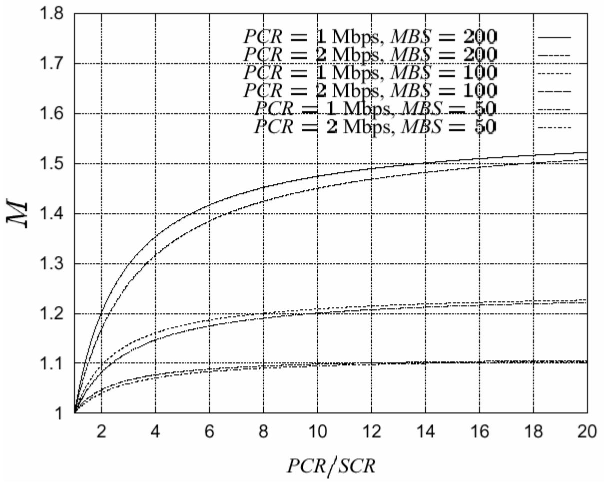
Guaranteed Services- 46

## Price Multipliers for charging

$$P_{\mathrm{VBR}} = \frac{EB(PCR, SCR, MBS)}{SCR} \cdot P_{\mathrm{CBR}}$$
$$= M(PCR, SCR, MBS) \cdot P_{\mathrm{CBR}} ,$$



## Price Multipliers



## Price Multipliers for charging

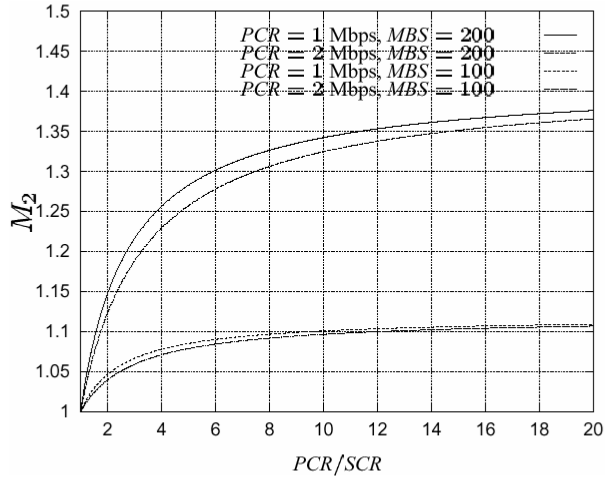
$$\begin{aligned} P_{\mathrm{VBR}} &= \frac{EB(PCR, SCR, MBS)}{SCR} \cdot P_{\mathrm{CBR}} \\ &= \frac{EB(PCR, SCR, MBS = 50)}{SCR} \cdot \\ &\quad \frac{EB(PCR, SCR, MBS)}{EB(PCR, SCR, MBS = 50)} \cdot P_{\mathrm{CBR}} \\ &= M_1 \cdot M_2 \cdot P_{\mathrm{CBR}} \,. \end{aligned}$$

$$M_1(PCR/SCR) = \frac{EB(PCR, SCR, MBS = 50)}{SCR}$$

$$M_2(PCR/SCR, MBS) = \frac{EB(PCR, SCR, MBS)}{EB(PCR, SCR, MBS = 50)}$$

Guaranteed Services- 49

## Price Multiplier M2




Guaranteed Services- 50

# Παραδείγματα χρέωσης ATM VBR VCs βάσει χρόνου

| Burst Ratio (PCR/SCR) | 1 | 2 | 5 | 10 | 15 | 20 |
|---|---|---|---|---|---|---|
| Multiplier | 1 | 1.2 | 1.4 | 1.45 | 1.5 | 1.55 |

| MBS | 50 | 100 | 200 |
|---|---|---|---|
| Multiplier (PCR/SCR≥1.8) | 0.88 | 0.93 | 1.0 |
| Multiplier (PCR/SCR<1.8) | 1.0 | 1.0 | 1.0 |

Μεγάλη στατιστική πολυπλεξία (π.χ. *PCR,SCR* μέχρι 1% της χωρητικότητας)

| Burst Ratio (PCR/SCR) | 1 | 2 | 5 | 10 | 15 | 20 |
|---|---|---|---|---|---|---|
| Multiplier | 1 | 1.2 | 1.7 | 2.2 | 2.5 | 2.7 |

| MBS | 50 | 100 | 200 |
|---|---|---|---|
| Multiplier (PCR/SCR≥1.8) | 0.85 | 0.9 | 1.0 |
| Multiplier (PCR/SCR<1.8) | 1.0 | 1.0 | 1.0 |

Μικρή στατιστική πολυπλεξία (π.χ. *PCR,SCR* μέχρι 10 % της χωρητικότητας)

- Μικρότερη στατιστική πολυπλεξία ⇒ οι πολλαπλασιαστές είναι πιο κοντά στην μονάδα



# British Telecom's (BT) tariffs

- Based on price "multipliers"

$$P_{\mathrm{VBR}} = M_1(PCR/SCR) \cdot M_2(PCR/SCR, MBS) \cdot P_{\mathrm{CBR}}$$

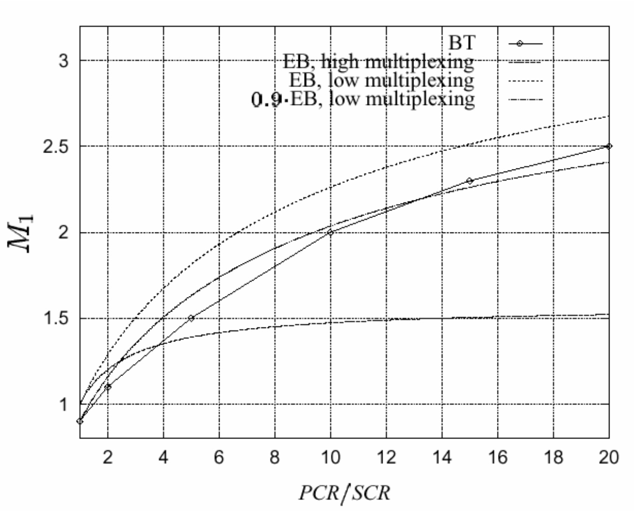
*(a) Multiplier $M_1$ for the burst ratio* PCR/SCR.

| Burst ratio | 1 | 2 | 5 | 10 | 15 | 20 |
|---|---|---|---|---|---|---|
| Multiplier | 0.9 | 1.1 | 1.5 | 2.0 | 2.3 | 2.5 |

*(b) Multiplier $M_2$ for* MBS.

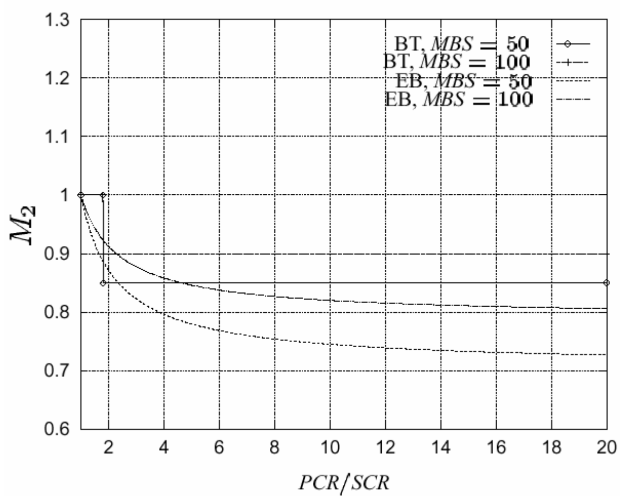
| *MBS* (in cells) | 50 | 100 | 200 |
|---|---|---|---|
| Multiplier ($PCR/SCR \geq 1.8$) | 0.85 | 0.9 | 1.0 |
| Multiplier ($PCR/SCR < 1.8$) | 1.0 | 1.0 | 1.0 |

## Comparison of EB and BT price multipliers




Guaranteed Services- 53

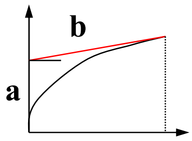
## Comparison of EB and BT price multipliers (2)



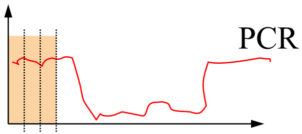

Guaranteed Services- 54

## Charging and CAC

- **Consistency of CAC and charging function:**
  - Natural to charge with the eb used in CAC
  - Suppose CAC according to PCR. How to charge?
    - Not a competitive CAC
    - Better provide incentives to reduce volume




$$aT + bV + c,\ a >> b$$




- Suppose "perfect" **dynamic** CAC is used
  - call arrival and departures occur every *T*
  - control mechanism (by blocking calls) achieves QoS at all times
  - ⇨ effective bandwidth of a call

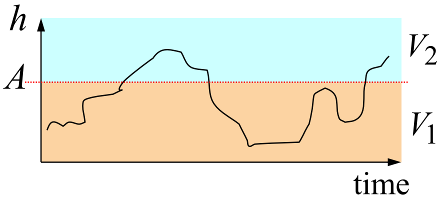
    = average of **actual** effective bandwidth in each period *T*

    = *almost the mean rate*!!



## More general charging schemes

- *2-tax band* scheme



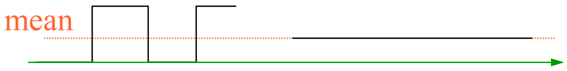

- V1: volume of data during intervals with less than A bytes
- V2: volume of data during intervals with more than A
- Charge is

$$f(V_1, V_2) = a_0 + a_1 V_1 + a_2 V_2$$

$$\text{where} \quad a_1 < a_2$$

## More general charging schemes

- Simple scheme can not distinguish users having the same mean




- Need for more detailed traffic measurements
- Consider the general linear tariff

$$f(X) = a_0 + a_1 g_1(X) + \ldots + a_L g_L(X)$$

  - $X = X_1, \ldots, X_T$ , $g_i(X) =$ measurement function $(= \frac{1}{T} \sum_{j=1,T} X_j)$
- Other possible functions: $(= \frac{1}{T} \sum_{j=1,T} 1\{X_j > 0.9h\}$
- Evaluate implementation cost vs accuracy gain



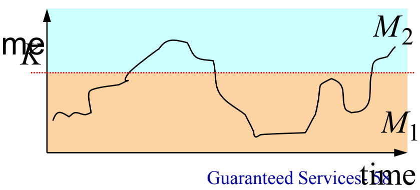
## More general charging schemes (cont.)

- Approach used in Simple Charging Scheme can be extended
  - Define the effective bandwidth to be the function

$$\alpha(h, M) = \sup_{X_t} \left\{ \frac{1}{st} \log E e^{sX[0,t]} \right\}$$

$$s.t.\ Eg(X) = M,\ X_t \in \Xi(h)$$

  - concave in *M*
- Construct linear tariffs = tangent hyperplanes to $\alpha(M)$
- Example: the *2-tax band* scheme



## Time-of-day pricing

- Assume two periods, off-peak and on-peak, t=1 and 2
- User utility $u_i(x_1^i, x_2^i)$
- Social welfare maximization problem

$$\max_{\{x_1^i, x_2^i\}} \sum_{i=1}^{N} u_i(x_1^i, x_2^i), \quad \text{subject to} \quad \sum_{i=1}^{N} x_t^i \le C_t, \; t = 1,2$$

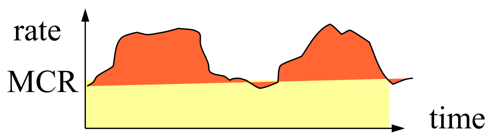
- User problem

$$\max_{x_1^i, x_2^i} [u_i(x_1^i, x_2^i) - p_1 x_1^i - p_2 x_2^i]$$

- p1, p2 set using a tatonnement



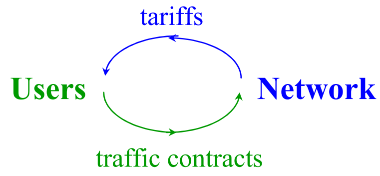
## Simple Time-Volume scheme for ABR




- $aT + bV + c$ can be used for ABR
- User buys an amount *m* of MCR at posted price $p_{MCR}$
- Network charges for a period of usage *T*:
  - $p_{MCR} \times m \times T$ for the data sent within MCR
  - $p_{UBR} \times V$ where *V* is the volume sent **on top** of MCR
  - *c* (for signalling congestion, discourage splitting)
- No incentive for splitting connections if excess capacity allocated proportionally to the amounts of MCR (note that *c*>0)

## Incentive compatibility and user-network interaction

- Operating point of a link and tariffs are inter-related in a circular fashion:




- *Incentive compatibility*: tariffs induce users to select contracts that minimize their charges, and lead to social welfare maximum of the system



## Equilibrium under deterministic multiplexing

- We assume identical users & single leaky bucket
- Users-network interact in lock-step fashion
- For deterministic multiplexing (zero CLP):

- In both cases, operating point moves towards $Q$
- $Q$: maximum number of users and $C\beta_Q=B\rho_Q$