UNIVERSITY OF CRETE

DEPARTMENT OF COMPUTER SCIENCE

Performance Analysis and Pricing in Broadband Networks

Vasilios A. Siris

Doctoral Dissertation

Heraklion, Crete

December 1997

Performance Analysis and Pricing in Broadband Networks

Vasilios A. Siris

Doctoral Dissertation

Department of Computer Science
University of Crete
December 1997

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Ανάλυση Επίδοσης και Τιμολόγηση σε Ευρυζώνια Δίκτυα

Διατριβή που υποβλήθηκε από τον

Βασίλειο Α. Σύρη

ως μερική απαίτηση για την απόκτηση του
ΔΙΔΑΚΤΟΡΙΚΟΥ ΔΙΠΛΩΜΑΤΟΣ

Ηράκλειο, Δεκέμβριος 1997

Συγγραφέας:

Εξεταστική Επιτροπή:

Κώστας Κουρκουμπέτης, Καθηγητής, Πανεπιστήμιο Κρήτης, Επόπτης

Jean Walrand, Professor, University of California, Berkeley

Πάνος Κωνσταντόπουλος, Καθηγητής, Πανεπιστήμιο Κρήτης

Βασίλειος Μάγκλαρης, Καθηγητής, Εθνικό Μετσόβιο Πολυτεχνείο

Λάζαρος Μεράκος, Καθηγητής, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Χρήστος Νικολάου, Αναπληρωτής Καθηγητής, Πανεπιστήμιο Κρήτης

Απόστολος Τραγανίτης, Αναπληρωτής Καθηγητής, Πανεπιστήμιο Κρήτης

Δημήτρης Σερπάνος, Επίκουρος Καθηγητής, Πανεπιστήμιο Κρήτης

Δεκτή:

Πάνος Κωνσταντόπουλος, Καθηγητής, Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Acknowledgements

I am very grateful to my supervisor, Professor Costas Courcoubetis, for giving me the opportunity to do research in such an exciting area, for the high quality knowledge he has teached me, and for his confidence in me. His enthusiastic and tireless guidance, along with our stimulating discussions, have helped keep my interest and motivation high throughout my Ph.D. studies.

I would like to thank the members of my dissertation committee from the University of Crete, Professors Panos Constantopoulos, Christos Nikolaou, Apostolos Traganitis, and Dimitris Serpanos, as well as the external members, Professors Jean Walrand (University of California, Berkeley, USA), Basil Maglaris (National Technical University of Athens), and Lazaros Merakos (National and Capodistrian University of Athens), for their helpful comments and questions. It is my honor to have them on my committee. I would also like to thank Professors Manolis Katevenis and Christos Nikolaou for serving on my advisory committee.

It is a great honor to know and work with Professors Frank P. Kelly and Richard Weber (both from the University of Cambridge, UK), who have contributed a great deal to this work. I would also like to thank Professor George D. Stamoulis (University of Crete) for a very fruitful collaboration, and for his helpful and precise comments on parts of this dissertation.

Many thanks to the members of the Telecommunications and Networks Group at the Institute of Computer Science, Foundation for Research and Technology - Hellas, and particularly Stelios Sartzetakis, Magda Chatzaki, and Fotis Kitsos, for creating a friendly working environment and for the interesting discussions we had. Thanks also to George Fouskas with whom I had useful discussions during the first years of my Ph.D. studies, when we shared the same office.

I would like to thank the secretaries of the Department of Computer Science and the Institute of Computer Science, as well as the technical staff of the latter, for their assistance.

I would like to acknowledge the support, both financial and in facilities, by the Institute of Computer Science, Foundation for Research and Technology - Hellas.

This dissertation would not have been possible without the continuous support and encouragement from my mother Athanasia and my brother Mark. Last but certainly not least, I would like to thank Sophia for her support, patience, and endurance to all the frustration and long working hours that have been part of this dissertation, and for being there to remind me of the many joys of life.

Στην μνήμη του πατέρα μου, Αναστάσιο
και στην μητέρα μου, Αθανασία

In memory of my father, Anastasios
and to my mother, Athanasia

Performance Analysis and Pricing in Broadband Networks

Vasilios A. Siris

Doctoral Dissertation
Department of Computer Science
University of Crete

Abstract

Effective procedures for performance analysis and pricing are important for controlling the use and dimensioning of broadband networks. These procedures must be accurate, efficient, and take into account the technological characteristics of broadband networks, such as the support for traffic with different characteristics and different Quality of Service (QoS) requirements, as well as the support for services with open loop and closed loop congestion control. The objective of this dissertation is to apply and evaluate large deviation techniques for performance analysis and traffic engineering, to investigate usage-based pricing schemes for network transport services with open loop congestion control, and to investigate pricing and resource sharing for Available Bit Rate (ABR) services, which support closed loop congestion control.

We use the recently developed theory of effective bandwidths where the effective bandwidth depends not only on the statistical characteristics of the traffic stream, but also on the link's operating point through two parameters, the space and time parameters, which are computed using the many sources asymptotic. We show for real broadband traffic that this effective bandwidth definition can accurately quantify resource usage. Furthermore, we show how the values of the space and time parameters can be used to clarify the effects on the link's performance of the time scales of traffic burstiness, of the traffic mix, of the link parameters (capacity and buffer), and of traffic control mechanisms, such as traffic smoothing. We use the many sources asymptotic to simultaneously capture the cell scale and burst scale effects on the buffer overflow probability at an Asynchronous Transfer Mode (ATM) output link with a small buffer that multiplexes a large number of periodic on-off sources. In addition to giving the correct expression of the overflow probability for very small buffers, we give a new qualitative description of how overflow occurs for such buffers.

We investigate usage-based pricing schemes for services with open loop control that are based on bounds of the effective bandwidth. These include time-volume schemes that involve two measurements (the duration and transferred volume) for each connection, and two schemes that involve measurements in distinct time intervals, smaller than the duration of a connection: pricing with renegotiation and the virtual bucket scheme. The schemes are compared according to their ability to capture the relative amount of resources used by connections (fairness). Furthermore, we show, for a particular setup, the incentive compatibility of the time-volume schemes.

Finally, we present an approach to pricing and resource sharing for ABR services with the following three features: (i) prices are adjusted according to the demand, (ii) users declare the price per unit of time according to which they will be charged, and (iii) resource sharing is based on effective bandwidths. The first feature captures the dynamic nature of congestion which is anticipated for ABR services, whereas the second and third features allow users to adequately reveal their preferences for network usage in terms of the price per unit of time, and enable the differentiation of connections based on their mean rates.

Our experiments are performed for link capacities and buffer sizes that will be used in broadband networks, and involve MPEG-1 compressed video with various contents, Internet Wide Area Network (WAN) traffic, and modeled voice traffic.

Supervisor:
Costas Courcoubetis,
Professor of Computer Science,
University of Crete

Ανάλυση Επίδοσης και Τιμολόγηση σε Ευρυζώνια Δίκτυα

Βασίλειος Α. Σύρης

Διδακτορική Διατριβή
Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Εκτενής Περίληψη

Αποτελεσματικές διαδικασίες για την ανάλυση της επίδοσης και την τιμολόγηση είναι σημαντικές για τον έλεγχο της χρήσης και την διαστασιοποίηση (dimensioning) των πόρων σε ευρυζώνια δίκτυα. Οι διαδικασίες αυτές θα πρέπει να είναι ακριβείς, αποδοτικές, και να λαμβάνουν υπόψη τα χαρακτηριστικά των δικτύων αυτών, όπως την μεταφορά κυκλοφορίας με διαφορετικά χαρακτηριστικά (τα οποία μπορεί να μεταβάλλονται με τον χρόνο) και τις διαφορετικές απαιτήσεις των χρηστών σε ποιότητα εξυπηρέτησης (Quality of Service - QoS), καθώς και την υποστήριξη τόσο υπηρεσιών με έλεγχο συμφόρησης ανοιχτού βρόχου όσο και υπηρεσιών με έλεγχο συμφόρησης κλειστού βρόχου.

Ο σκοπός της παρούσας διατριβής είναι να εφαρμόσει και να αξιολογήσει τεχνικές από την θεωρία μεγάλων αποκλίσεων (large deviations) για την ανάλυση επίδοσης και την διαστασιοποίηση των πόρων ενός δικτύου, να ερευνήσει σχήματα τιμολόγησης που βασίζονται στην χρήση για υπηρεσίες με έλεγχο συμφόρησης ανοιχτού βρόχου, και να ερευνήσει την τιμολόγηση και διαμοίραση πόρων για υπηρεσίες διαθέσιμου ρυθμού (Available Bit Rate - ABR), οι οποίες υποστηρίζουν έλεγχο συμφόρησης κλειστού βρόχου.

Ο ρόλος της τιμολόγησης δεν είναι μόνο να δημιουργεί κέρδη για τον παροχέα δικτυακών υπηρεσιών, αλλά και να ελέγχει την χρήση των πόρων ενός δικτύου. Το σχήμα τιμολόγησης πρέπει να δίνει τα σωστά κίνητρα ώστε οι χρήστες να κάνουν αποδοτική, από οικονομικής πλευράς, χρήση των πόρων του δικτύου. Αυτό θα μείωνε τις αρνητικές επιπτώσεις που έχει η συμφόρηση, η οποία είναι ένα από τα μεγαλύτερα προβλήματα στο παγκόσμιο διαδίκτυο Internet. Πολλοί μηχανικοί και οικονομολόγοι πιστεύουν ότι η συμφόρηση στο παγκόσμιο διαδίκτυο οφείλεται στο αναποτελεσματικό σχήμα τιμολόγησης, που βασίζεται κατεξοχήν

στην τιμολόγηση ενιαίου ρυθμού (flat rate pricing) όπου τα μηνιαία τέλη εξαρτώνται μόνο από την ταχύτητα της σύνδεσης του πελάτη με τον παροχέα. Αυτό το σχήμα τιμολόγησης δεν δίνει κίνητρα στους χρήστες να χρησιμοποιούν μικρότερη χωρητικότητα από αυτή της σύνδεσης τους. Επιπλέον, η τιμολόγηση ενιαίου ρυθμού δεν επιτρέπει στους χρήστες να γνωστοποιήσουν τις προτιμήσεις τους. Όλοι οι χρήστες αντιμετωπίζονται με τον ίδιο τρόπο, παρ'όλο που διαφορετικοί χρήστες μπορεί να δίνουν διαφορετική αξία στην ίδια υπηρεσία. Οι δύο παραπάνω περιορισμοί έχουν σαν αποτέλεσμα ένα συμφορημένο δίκτυο του οποίου οι πόροι χρησιμοποιούνται κατά τρόπο που δεν ανταποκρίνεται στις πραγματικές ανάγκες των χρηστών.

Η στατιστική ανάλυση μετρήσεων κυκλοφορίας έχει δείξει ότι πολλά είδη κυκλοφορίας, όπως η κυκλοφορία τοπικού δικτύου Ethernet, η κυκλοφορία δικτύου ευρείας περιοχής (Wide Area Network - WAN), η κυκλοφορία video, και η κυκλοφορία του παγκόσμιου ιστού (World Wide Web), έχουν αυτοπαρεμφερή συμπεριφορά. Κυκλοφορία αυτού του τύπου έχει ισχυρές εξαρτήσεις ή αυτοσυσχέτιση που φθίνει αργά. Σε ποιό βαθμό η ύπαρξη ισχυρών εξαρτήσεων επηρεάζει την επίδοση δικτύων είναι ένα ανοιχτό ερευνητικό θέμα. Όμως, η πρόσφατη έρευνα έχει δείξει ότι οι ισχυρές εξαρτήσεις έχουν δευτερεύουσα σημασία στην πιθανότητα υπερχείλισης σε έναν σύνδεσμο με μικρό μέγεθος ενταμιευτή. Αυτό συμβαίνει όταν υποστηρίζεται επικοινωνία πραγματικού χρόνου. Τα παραπάνω αποτελέσματα δημιουργούν την ανάγκη για διαδικασίες που διευκρινίζουν πώς οι χρονικές κλίμακες φαινομένων που παρατηρούνται σε πραγματική κυκλοφορία επηρεάζουν την επίδοση δικτύων. Μερικά απλά ερωτήματα που πρέπει να απαντηθούν είναι τα εξής: Πώς μεταβάλλεται η πιθανότητα υπερχείλισης του ενταμιευτή ενός συνδέσμου με την αύξηση του μεγέθους του ενταμιευτή ή την αύξηση της χωρητικότητας του συνδέσμου; Πώς επηρεάζει η εξομάλυνση της κυκλοφορίας (traffic smoothing) την ικανότητα πολυπλεξίας ενός συνδέσμου και την ποσότητα πόρων που χρησιμοποιεί μία πηγή;

Ανάλυση επίδοσης σε ευρυζώνια δίκτυα

Το ισοδύναμο εύρος ζώνης είναι μία παράμετρος που εξαρτάται από τις στατιστικές ιδιότητες μιας πηγής και, όπως θα δούμε παρακάτω, από τις στατιστικές ιδιότητες των υπόλοιπων πηγών με τις οποίες πολυπλέκεται και τις παραμέτρους (χωρητικότητα και μέγεθος του ενταμιευτή) του συνδέσμου στον οποίο πολυπλέκεται. Το ισοδύναμο εύρος ζώνης εκφράζει την ποσότητα

πόρων που πρέπει να δεσμευτούν για την πηγή ώστε να ικανοποιηθούν οι απαιτήσεις της σε ποιότητα εξυπηρέτησης. Μία τέτοια ποσότητα επιτρέπει την αναγωγή προβλημάτων (όπως τον έλεγχο αποδοχής συνδέσεων και την δρομολόγηση) ευρυζώνιων δικτύων που πολυπλέκουν εκρηκτικές πηγές σε προβλήματα παραδοσιακών δικτύων μεταγωγής κυκλωμάτων.

Η ασυμπτωτική ανάλυση ασχολείται με τον τρόπο που η πιθανότητα υπερχείλισης του ενταμιευτή ενός συνδέσμου φθίνει όταν κάποια ποσότητα του αυξάνεται. Αν αυτή η ποσότητα είναι το μέγεθος του ενταμιευτή, τότε έχουμε την ασυμπτωτική ανάλυση για μεγάλο ενταμιευτή (large buffer asymptotic). Εάν το ανά πηγή μέγεθος του ενταμιευτή και η ανά πηγή χωρητικότητα του συνδέσμου παραμένουν σταθερά, και μας ενδιαφέρει ο τρόπος που η πιθανότητα υπερχείλισης φθίνει όταν το μέγεθος του συστήματος (που αποτελείται από τον σύνδεσμο και τις πολυπλεγμένες πηγές) αυξάνεται, τότε έχουμε την ασυμπτωτική ανάλυση για μεγάλο πλήθος πηγών (many sources asymptotic).

Ορισμοί του ισοδύναμου εύρους ζώνης που βασίζονται στην ασυμπτωτική ανάλυση για μεγάλο ενταμιευτή μπορεί να υπερεκτιμήσουν ή και να υποτιμήσουν την πραγματική ποσότητα πόρων που χρειάζεται μία πηγή για να ικανοποιηθούν οι απαιτήσεις της σε ποιότητα εξυπηρέτησης. Αυτό συμβαίνει επειδή η ασυμπτωτική ανάλυση για μεγάλο μέγεθος ενταμιευτή δεν λαμβάνει υπόψη το στατιστικό κέρδος όταν ένα μεγάλο πλήθος από ανεξάρτητες πηγές πολυπλέκονται μεταξύ τους. Μόνο πρόσφατα [Kel96a, CKW97] έχει βρεθεί τρόπος να εισαχθεί αυτή η πληροφορία στον ορισμό του ισοδύναμου εύρους ζώνης. Η εργασία αυτή έχει δείξει ότι το ισοδύναμο εύρος ζώνης μιας πηγής εξαρτάται από το σημείο λειτουργίας ενός συνδέσμου μέσω δύο παραμέτρων, της παραμέτρου χώρου και της παραμέτρου χρόνου, οι οποίες με την σειρά τους εξαρτώνται από τους πόρους του συνδέσμου (χωρητικότητα και μέγεθος του ενταμιευτή) και τις στατιστικές ιδιότητες της πολυπλεγμένης κυκλοφορίας. Οι τιμές αυτών των παραμέτρων υπολογίζονται με την ασυμπτωτική ανάλυση για μεγάλο πλήθος πηγών.

Στην παρούσα διατριβή εφαρμόζουμε και αξιολογούμε τα αποτελέσματα του ασυμπτωτικού για μεγάλο πλήθος πηγών για την ανάλυση επίδοσης και την διαστασιοποίηση των πόρων ενός δικτύου, και την χρήση του ισοδύναμου εύρος ζώνης για την ποσοτικοποίηση των πόρων που χρησιμοποιεί μία πηγή. Η ανάλυση γίνεται για χωρητικότητες συνδέσμου και μεγέθη του ενταμιευτή που θα χρησιμοποιηθούν στα μελλοντικά ευρυζώνια δίκτυα και περιλαμβάνουν

πραγματική κυκλοφορία που αποτελείται από MPEG-1[1] συμπιεσμένη κυκλοφορία video και κυκλοφορία Internet σε δίκτυο ευρείας περιοχής, καθώς και κυκλοφορία από μοντέλο φωνής. Τα αποτελέσματα των πειραμάτων δείχνουν ότι, ενώ το ασυμπτωτικό για μεγάλο πλήθος πηγών (με την βελτίωση Bahadur-Rao) μπορεί να υπερεκτιμήσει την πιθανότητα απώλειας κελιού κατά 1-2 τάξεις μεγέθους, επιτρέπει βαθμό εκμετάλλευσης ενός συνδέσμου πολύ κοντά στον μέγιστο βαθμό εκμετάλλευσης. Για παράδειγμα, στην περίπτωση MPEG-1 συμπιεσμένης κυκλοφορίας video που πολυπλέκεται σε σύνδεσμο χωρητικότητας 155 Mbps και μέγεθος ενταμιευτή που αντιστοιχεί σε καθυστέρηση μεγαλύτερη από 1 msec (περίπου 351 κελιά ATM - Ασύγχρονου Τρόπου Μετάδοσης), το ασυμπτωτικό αποτέλεσμα για μεγάλο πλήθος πηγών με την βελτίωση Bahadur-Rao επιτρέπει βαθμό εκμετάλλευσης του συνδέσμου που είναι πάνω από 94% του μέγιστου βαθμού εκμετάλλευσης. Τα αποτελέσματα αυτά δείχνουν ότι το ισοδύναμο εύρος ζώνης που βασίζεται στο ασυμπτωτικό για μεγάλο πλήθος πηγών ποσοτικοποιεί με ακρίβεια την χρήση πόρων.

Όπως αναφέρθηκε παραπάνω, οι παράμετροι χώρου και χρόνου χαρακτηρίζουν το σημείο λειτουργίας ενός συνδέσμου. Στην παρούσα διατριβή ερευνούμε πώς οι τιμές αυτών των παραμέτρων διευκρινίζουν την επίδραση που έχουν στην επίδοση ενός συνδέσμου οι χρονικές κλίμακες φαινομένων εκρηκτικότητας, οι παράμετροι του συνδέσμου (χωρητικότητα και μέγεθος του ενταμιευτή), το μείγμα κυκλοφορίας, και μηχανισμοί ελέγχου κυκλοφορίας, όπως η εξομάλυνση της κυκλοφορίας. Ειδικότερα, η παράμετρος χώρου δείχνει τον ρυθμό με τον οποίο μειώνεται ο λογάριθμος της πιθανότητας υπερχείλισης όταν το μέγεθος του ενταμιευτή αυξάνεται, ενώ η παράμετρος χρόνου δείχνει την κλίμακα χρόνου (time scale) στην οποία οι μηχανισμοί ελέγχου κυκλοφορίας πρέπει να εξεταστούν για να βρεθεί πόσο αυτοί επηρεάζουν την χρήση πόρων του δικτύου. Επιπλέον, η παράμετρος χρόνου δείχνει την διακριτότητα (granularity) που θα πρέπει να έχουν τα ίχνη κυκλοφορίας για να μήν χαθεί πληροφορία σημαντική για την ανάλυση της επίδοσης του συνδέσμου. Για τα μείγματα κυκλοφορίας που εξετάζουμε, οι τιμές των παραμέτρων χώρου και χρόνου, κατά μεγάλο βαθμό, παρουσιάζουν μικρές αλλαγές όταν το ποσοστό των διαφόρων τύπων κυκλοφορίας στο μείγμα αλλάζει λίγο. Επιπλέον, αυτή η εξάρτηση γίνεται ακόμα πιο ασθενής για μεγαλύτερες χωρητικότητες

[1]Το MPEG-1 είναι μία μέθοδο συμπίεσης video και του συνοδευόμενου ήχου που έχει τυποποιηθεί από τον Διεθνή Οργανισμό Τυποποίησης (International Organization for Standardization - ISO). Τα αρχικά MPEG προέρχονται από το Motion Pictures Expert Group η οποία είναι η ομάδα που ανέπτυξε την μέθοδο.

και μεγαλύτερα μεγέθη του ενταμιευτή. Τα αποτελέσματα αυτά αποτελούν ένδειξη ότι συγκεκριμένα ζευγάρια τιμών των παραμέτρων χώρου και χρόνου μπορεί να χαρακτηρίσουν διαφορετικές περιόδους της ημέρας κατά την διάρκεια των οποίων η σύνθεση του μείγματος της κυκλοφορίας παραμένει σχετικά σταθερή.

Πιθανότητα υπερχείλισης σε σύνδεσμο ATM με μικρό ενταμιευτή

Στα δίκτυα Ασύγχρονου Τρόπου Μετάδοσης (Asynchronous Transfer Mode - ATM) η πληροφορία μεταφέρεται σε μικρά, σταθερού μεγέθους (53 bytes) κελιά (cells). Όταν το μέγεθος του ενταμιευτή είναι μικρό, αυτή η διακριτοποίηση της μεταφοράς πληροφορίας μπορεί να οδηγήσει σε συμφόρηση κλίμακας κελιού (cell scale congestion). Η συμφόρηση κλίμακας κελιού συμβαίνει όταν κελιά από διαφορετικές πηγές φθάνουν στον ενταμιευτή την ίδια σχεδόν χρονική στιγμή. Αυτό δημιουργεί, σε μικρό χρονικό διάστημα, έναν συνολικό αριθμό αφίξεων κελιών μεγαλύτερο από την χωρητικότητα του συνδέσμου. Από την άλλη, η συμφόρηση κλίμακας έκρηξης (burst scale congestion) συμβαίνει όταν η είσοδος από τις πολυπλεγμένες πηγές, θεωρούμενη ως μία συνεχή ροή υγρού, έχει ρυθμό μεγαλύτερο από την χωρητικότητα του συνδέσμου για αρκετό χρονικό διάστημα ώστε να γεμίσει ο ενταμιευτής. Το μεγαλύτερο μέρος της ανάλυσης σε δίκτυα ATM έχει γίνει μοντελοποιώντας την ροή κελιών ως μία συνεχή ροή υγρού. Το μοντέλο αυτό συλλαμβάνει το φαινόμενο της συμφόρησης κλίμακας έκρηξης, όχι όμως και το φαινόμενο της συμφόρησης κλίμακας κελιού, και είναι ακριβές όταν το μέγεθος του ενταμιευτή δεν είναι πολύ μικρό. Για παράδειγμα, όταν πολυπλέκονται περιοδικές πηγές με μέγιστο ρυθμό 1 Mbps και μέσο ρυθμό 0.25 Mbps σε σύνδεσμο ATM με χωρητικότητα 622 Mbps, με μέγεθος ενταμιευτή 20 κελιών, και με βαθμό εκμετάλλευσης 0.8, η μοντελοποίηση ροής υγρού μας δίνει πιθανότητα υπερχείλισης μικρότερη από 10^{-8}, ενώ η πραγματική πιθανότητα υπερχείλισης είναι μεγαλύτερη από 10^{-4}.

Στην παρούσα διατριβή χρησιμοποιούμε την ασυμπτωτική ανάλυση για μεγάλο πλήθος πηγών για να μοντελοποιήσουμε ταυτόχρονα την συμφόρηση κλίμακας κελιού και την συμφόρηση κλίμακας έκρηξης σε έναν σύνδεσμο που πολυπλέκει περιοδικές πηγές δύο καταστάσεων (on-off). Με αυτό τον τρόπο δίνουμε την σωστή έκφραση για την πιθανότητα υπερχείλισης και μία καινούργια ποιοτική εξήγηση για το πώς συμβαίνει υπερχείλιση όταν το μέγεθος του ενταμιευτή είναι πολύ μικρό. Συγκεκριμένα, υπερχείλιση συμβαίνει με τον συνδυασμό δύο διαφορετικών φαινομένων: ένα αυξημένο ποσοστό από τις πηγές είναι ενεργές

(βρίσκονται στην κατάσταση ‘‘on’’), και ο συγχρονισμός των αφίξεων κελιών από τις πηγές αυτές δημιουργεί συμφόρηση σε κλίμακα κελιού.

Τιμολόγηση για υπηρεσίες με έλεγχο συμφόρησης ανοιχτού βρόχου

Έχουμε αναφέρει ότι το ισοδύναμο εύρος ζώνης μπορεί με ακρίβεια να αποτιμήσει την ποσότητα των πόρων μιας σύνδεσης. Αυτό δικαιολογεί την χρήση της έννοιας του ισοδύναμου εύρους ζώνης στα σχήματα τιμολόγησης που ερευνούμε. Εκτός από την ακριβή αποτίμηση της ποσότητας των πόρων μιας σύνδεσης, ένα σχήμα τιμολόγησης θα πρέπει να λαμβάνει υπόψη ιδιαίτερα χαρακτηριστικά των μηχανισμών που χρησιμοποιεί το δίκτυο για την παροχή υπηρεσιών, όπως τον έλεγχο συμφόρησης ανοιχτού βρόχου και τον έλεγχο συμφόρησης κλειστού βρόχου, και να είναι απλό και εύκολα υλοποιήσιμο. Είναι βασικό να προσθέσουμε ότι η τελική χρέωση για μία σύνδεση θα εξαρτηθεί, εκτός από την ποσότητα των πόρων που χρησιμοποιεί η σύνδεση, και από οικονομικούς παράγοντες, όπως την διαφήμιση και τον ανταγωνισμό. Στην παρούσα διατριβή ασχολούμαστε μόνο με την κατασκευή του μέρους της συνάρτησης τιμολόγησης που αποτιμά την ποσότητα των πόρων που χρησιμοποιεί μία σύνδεση.

Οι υπηρεσίες με έλεγχο συμφόρησης ανοιχτού βρόχου περιλαμβάνουν συμβόλαιο κυκλοφορίας (traffic contract) μεταξύ του χρήστη και του δικτύου. Το συμβόλαιο κυκλοφορίας περιέχει την συμφωνημένη ποιότητα εξυπηρέτησης και την περιγραφή της μέγιστης κυκλοφορίας που επιτρέπεται να στείλει ο χρήστης. Παραδείγματα τέτοιων υπηρεσιών είναι οι υπηρεσίες σταθερού ρυθμού (Constant Bit Rate - CBR) και οι υπηρεσίες μεταβλητού ρυθμού (Variable Bit Rate - VBR) σε δίκτυα ATM, οι υπηρεσίες εγγυημένης εξυπηρέτησης (guaranteed service) και οι υπηρεσίες ελεγχόμενου φόρτου (controlled-load) που ορίζονται για την αρχιτεκτονική ενοποιημένων υπηρεσίων του Internet, καθώς και οι υπηρεσίες σταθερής πρόσβασης στο Internet.

Τα σχήματα τιμολόγησης για υπηρεσίες ανοιχτού βρόχου που εξετάζουμε περιλαμβάνουν απλά σχήματα τιμολόγησης χρόνου-όγκου (time-volume pricing schemes) που απαιτούν μόνο δύο μετρήσεις για όλη την σύνδεση (την συνολική διάρκεια της σύνδεσης και τον συνολικό όγκο πληροφορίας που μεταφέρθηκε) και σχήματα τιμολόγησης που περιλαμβάνουν μετρήσεις σε χωριστά χρονικά διαστήματα, μικρότερα της συνολικής διάρκειας της σύνδεσης.

Τα σχήματα τιμολόγησης χρόνου-όγκου πρωτοεμφανίστηκαν στην εργασία [Kel94] και

επεκτάθηκαν αργότερα στην εργασία [CKW97]. Τα σχήματα αυτά βασίζονται σε φράγματα του ισοδύναμου εύρους ζώνης που λαμβάνουν υπόψη και a priori πληροφορία που περιέχεται στο συμβόλαιο κυκλοφορίας (όπως τον μέγιστο ρυθμό με τον οποίο επιτρέπεται να στέλνει δεδομένα ο χρήστης) και a posteriori πληροφορία που μετράται κατά την διάρκεια της σύνδεσης (όπως την συνολική διάρκεια της σύνδεσης). Στα σχήματα αυτά το συνολικό τέλος υπολογίζεται σαν μία γραμμική συνάρτηση της συνολικής διάρκειας της σύνδεσης και του συνολικού όγκου πληροφορίας που μεταφέρθηκε. Οι παράμετροι αυτής της συνάρτησης ορίζονται από φράγματα του ισοδύναμου εύρους ζώνης κατά τρόπο που να μειώνεται η χρέωση των χρηστών όταν αυτοί γνωρίζουν προκαταβολικά (ή μπορούν να προβλέψουν με σχετική ακρίβεια) τον μέσο ρυθμό κυκλοφορίας τους. Στην παρούσα διατριβή αναλύουμε την απόδοση της μεθόδου για τρία φράγματα του ισοδύναμου εύρους ζώνης: το φράγμα on-off [Kel94], το ‘‘απλό’’ φράγμα [CKW97], και την προσέγγιση ανεστραμμένου T [CKW97].

Μαζί με τα σχήματα τιμολόγησης χρόνου-όγκου, εξετάζουμε δύο νέα σχήματα τιμολόγησης που βασίζονται σε μετρήσεις που γίνονται σε χρονικά διαστήματα μικρότερα της συνολικής διάρκειας της σύνδεσης: την τιμολόγηση με επαναδιαπραγμάτευση (pricing with renegotiation) και το σχήμα τιμολόγησης εικονικού δοχείου (virtual bucket pricing scheme). Η τιμολόγηση με επαναδιαπραγμάτευση αποτελεί εφαρμογή των σχημάτων χρόνου-όγκου σε χρονικά διαστήματα μικρότερα από την συνολική διάρκεια μιας σύνδεσης, όταν επιτρέπεται η επαναδιαπραγμάτευση των παραμέτρων περιγραφής της κυκλοφορίας. Η τελευταία δυνατότητα υποστηρίζεται από πρόσφατες προσπάθειες τυποποίησης. Το σχήμα τιμολόγησης εικονικού δοχείου δεν επιβαρύνεται με τον επιπλέον φόρτο λόγω επαναδιαπραγμάτευσης, αλλά απαιτεί την δήλωση μιας παραμέτρου (του ρυθμού διαρροής του εικονικού δοχείου) στην έναρξη της σύνδεσης.

Εφόσον η ποσότητα των πόρων που χρησιμοποιεί μία σύνδεση δεν είναι παρά ένα μόνο μέρος της τελικής χρέωσης της σύνδεσης, δεν είναι σημαντική η ακριβής εκτίμηση της απόλυτης τιμή της ποσότητας των πόρων που χρησιμοποιούνται από τις συνδέσεις, αλλά η ακριβής εκτίμηση της σχετικής ποσότητας των πόρων. Την τελευταία ιδιότητα την ονομάζουμε δικαιοσύνη (fairness), και με βάση αυτήν συγκρίνουμε τα προαναφερθέντα σχήματα τιμολόγησης για MPEG-1 συμπιεσμένη κυκλοφορία video με διαφορετικά περιεχόμενα (ταινίες, ειδήσεις, και συζητήσεις) και κυκλοφορία Internet σε δίκτυο ευρείας περιοχής. Τα πειραματικά μας αποτελέσματα

δείχνουν ότι η δικαιοσύνη και η ανθεκτικότητα (robustness) σε μικρές αλλαγές στο σημείο λειτουργίας του δικτύου (που εκφράζεται μέσω των παραμέτρων χώρου και χρόνου) των σχημάτων χρόνου-όγκου είναι μεγαλύτερη για μεγαλύτερη χωρητικότητα του συνδέσμου και μεγαλύτερο μέγεθος του ενταμιευτή, για συνδέσεις με μικρότερη διάρκεια, και για συνδέσεις με λιγότερο εκρηκτική κυκλοφορία. Μάλιστα, για συνδέσμους με μεγάλη χωρητικότητα και συνδέσεις με μικρή διάρκεια, η απόδοση αυτών των σχημάτων πλησιάζει εκείνη των σχημάτων τιμολόγησης που περιλαμβάνουν μετρήσεις σε χωριστά χρονικά διαστήματα. Αυτό δείχνει ότι για συνδέσμους με μεγάλο βαθμό στατιστικής πολυπλεξίας, ορισμένες πολύ απλές προσεγγίσεις επιτυγχάνουν πολύ καλή απόδοση. Από την άλλη, η δικαιοσύνη και ανθεκτικότητα των σχημάτων τιμολόγησης που περιλαμβάνουν μετρήσεις σε χωριστά χρονικά διαστήματα, τα οποία έχουν μεγαλύτερο κόστος υλοποίησης λόγω του μεγαλύτερου αριθμού μετρήσεων που απαιτούν, είναι πάντα μεγαλύτερες των σχημάτων τιμολόγησης χρόνου-όγκου. Επιπλέον, η απόδοση του σχήματος εικονικού δοχείου, το οποίο δεν επιβαρύνεται με τον επιπλέον φόρτο λόγω επαναδιαπραγμάτευσης, είναι πολύ κοντά στην τιμολόγηση με επαναδιαπραγμάτευση.

Εκτός από τις λεπτομερείς αριθμητικές συγκρίσεις των σχημάτων τιμολόγησης, δείχνουμε, για συγκεκριμένο σενάριο, την συμβατότητα κινήτρων του δικτύου και των χρηστών (incentive compatibility) των σχημάτων τιμολόγησης χρόνου-όγκου, και μελετούμε κατά πόσο το σημείο ισορροπίας ενός δικτύου επηρεάζεται από την εξομάλυνση κυκλοφορίας. Ο ρόλος ενός σχήματος τιμολόγησης που έχει την ιδιότητα της συμβατότητας κινήτρων είναι να μετακινεί αργά το σύστημα, που αποτελείται από το δίκτυο και τους χρήστες, σε ένα ευσταθές και αποδοτικό σημείο λειτουργίας. Η κίνηση αυτή είναι αποτέλεσμα του βρόχου χρέωσης που περιλαμβάνει το δίκτυο και τους χρήστες: Το δίκτυο αναγγέλει τιμές οι οποίες επηρεάζουν τα συμβόλαια κυκλοφορίας των χρηστών που επιδιώκουν να μειώσουν την χρέωση τους. Οι αποφάσεις αυτών θα επηρεάσουν το σημείο λειτουργίας του δικτύου, οπότε και τις τιμές του, κ.ο.κ. Σχήματα τιμολόγησης που δεν έχουν την ιδιότητα της συμβατότητας κινήτρων μπορεί να οδηγήσουν το δίκτυο σε ασταθή και μή αποδοτική λειτουργία.

Τιμολόγηση και διαμοίραση πόρων για υπηρεσίες ABR

Η υπηρεσία διαθέσιμου ρυθμού (Available Bit Rate - ABR) σε δίκτυα ATM υποστηρίζει έλεγχο συμφόρησης κλειστού βρόχου. Στην υπηρεσία αυτή δεν υπάρχει εγγυημένη ποιότητα εξυπηρέτησης (πλήν της εγγύησης ενός ελάχιστου ρυθμού, Minimum Cell Rate - MCR, με τον

οποίο ένας χρήστης μπορεί να στέλνει δεδομένα) ούτε συγκεκριμένη περιγραφή κυκλοφορίας στην οποία ο χρήστης θα πρέπει να συμμορφώνεται. Αντί αυτών, ο χρήστης προσαρμόζει τον ρυθμό κυκλοφορίας σύμφωνα με σήματα ανάδρασης που λαμβάνει από το δίκτυο. Τα σήματα ανάδρασης μπορεί να δηλώνουν, μέσω ενός bit, την ύπαρξη ή μή συμφόρησης, ή μπορεί να περιέχουν τον μέγιστο ρυθμό με τον οποίο μπορεί να στέλνει δεδομένα ο χρήστης. Στην παρούσα εργασία μας αποσχολεί η δεύτερη περίπτωση. Εφόσον το δίκτυο είναι εκείνο που αποφασίζει τον μέγιστο ρυθμό ενός χρήστη, είναι λογικό το σχήμα τιμολόγησης να επηρεάζει τον τρόπο με τον οποίο αυτό μοιράζει τους πόρους του.

Στην παρούσα διατριβή παρουσιάζουμε ένα σχήμα τιμολόγησης και διαμοίρασης πόρων για υπηρεσίας ABR που έχει τα ακόλουθα χαρακτηριστικά: (i) οι τιμές προσαρμόζονται σύμφωνα με την ζήτηση, (ii) οι χρήστες δηλώνουν την τιμή ανά μονάδα χρόνου με βάση την οποία θα χρεωθούν, και (iii) η διαμοίραση των πόρων γίνεται με βάση το ισοδύναμο εύρος ζώνης. Το πρώτο χαρακτηριστικό αντιμετωπίζει τον δυναμικό χαρακτήρα που αναμένεται να έχει η συμφόρηση για υπηρεσίες ABR. Το δεύτερο και τρίτο χαρακτηριστικό επιτρέπουν στους χρήστες να δηλώνουν τις προτιμήσεις τους μέσω της τιμής που θα χρεωθούν, και οδηγούν στην διαφοροποίηση συνδέσεων με διαφορετικό μέσο ρυθμό. Η τελευταία ιδιότητα συνεισφέρει στην συμβατότητα κινήτρων του προτεινόμενου σχήματος, και λείπει από υπάρχουσες διαδικασίες ελέγχου συμφόρησης σύμφωνα στις οποίες το δίκτυο ελέγχει τους μέγιστους ρυθμούς κυκλοφορίας, χωρίς να λαμβάνει υπόψη τους μέσους ρυθμούς κυκλοφορίας. Με πειράματα προσομοίωσης επιδεικνύουμε την σύγκλιση των δυναμικά μεταβαλλόμενων τιμών και την διαφοροποίηση συνδέσεων με διαφορετικό μέσο ρυθμό κυκλοφορίας. Μία σημαντική ιδιότητα του προτεινόμενου σχήματος είναι ότι μπορεί να υλοποιηθεί με τους υπάρχοντες μηχανισμούς ελέγχου συμφόρησης που έχουν οριστεί στα πρότυπα για τις υπηρεσίες ABR.

Επόπτης:
Κώστας Κουρκουμπέτης,
Καθηγητής Επιστήμης Υπολογιστών,
Πανεπιστήμιο Κρήτης

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Effective procedures for performance analysis and pricing are important for controlling the use and dimensioning of broadband networks. These procedures must be accurate, efficient, and take into account the technological characteristics of such networks. The objective of this dissertation is to apply and evaluate large deviation techniques for performance analysis and traffic engineering, to investigate usage-based pricing schemes for network transport services with open loop congestion control, and to investigate pricing and resource sharing for Available Bit Rate (ABR) services, which support closed loop congestion control.

In this section we present the motivation of our research and summarize its contributions, survey related work, and give an outline of the rest of the dissertation.

1.1 Motivation

The networking scene has drastically changed over the last few years. Due to advances in technology and progress in standardization, Asynchronous Transfer Mode (ATM) technology [I.321, ATMF94] has found its way inside the campus backbone, and Wide Area Network (WAN) ATM services are slowly being deployed. Networks based on ATM technology will carry traffic with different characteristics, which may vary in time, and different Quality of Service (QoS) requirements. At the same time, the Internet, once primarily an academic and research network, is rapidly being commercialized and standards are being developed to augment its service model, which currently contains only best-effort service, with services providing guarantees in terms of throughput and

delay [BCS94, She95a]. Finally, bold steps towards deregulation are increasing the competitive nature of the telecommunications market [GTR96]. In such an environment, economic theory suggests that prices should reflect resource usage.

The above changes have generated the need for simple and efficient pricing schemes, and simple and accurate procedures for performance analysis and traffic engineering.

1.1.1 Usage-based Pricing

The role of pricing is not only to generate income for the service provider, but also to control the use of network resources. The pricing structure must provide the right incentives for users to efficiently, in an economical sense, use network resources. This would reduce the negative effects of congestion, which is currently one of the most intense problems of the Internet. Many engineers and economists believe that it is due to the Internet's ineffective pricing structure, which is based primarily on *flat rate* pricing where prices depend only on the rate of the access pipe which connects the customer to his provider [MMV94, Fir97]. Such a pricing scheme provides no incentives for users to use less bandwidth than the rate of their access pipe. Furthermore, flat rate pricing does not enable users to adequately reveal their preferences for network usage. All users are treated the same, even though different users might have a different value for the same service. Both of these limitations result in a congested network where resources are not used according to the actual needs of users.

The focus of this dissertation is on pricing network *transport* services, i.e., pricing connections which simply move bits, without having knowledge of the information they are carrying. The problem of pricing such services involves many disciplines - *Economics*, *Marketing*, and *Network Engineering* are perhaps the most important (Figure 1.1).

Economic theory argues that usage-based pricing is needed in order for network resources to be used efficiently [MMV95a]. Furthermore, there is a large number of well understood economic models and a large amount of empirical results for networks that offer a limited range of services, such as the telephone network [MV91]. An important feature of such networks is that once a user is granted a connection, the resources associated with the connection remain reserved throughout its duration. The case of broadband networks is far more complex [CSSM95]. Traffic has different characteristics, which may vary in time, and different QoS requirements expressed in terms of loss

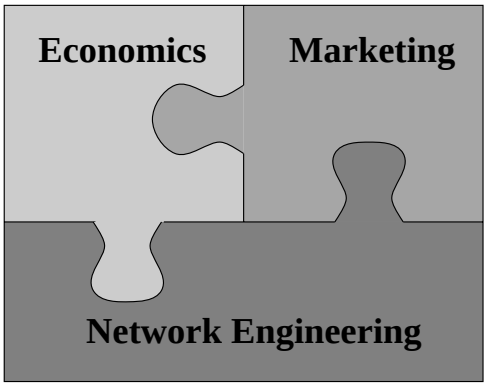


Figure 1.1: **The pricing network transport puzzle.** An important part of the puzzle, which is a focus of this dissertation, is the problem of quantifying and efficiently measuring resource usage for creating pricing schemes which take into account relevant technological characteristics and provide the right incentives for efficient and stable network operation.

probability, delay, and delay variance. Quantifying the amount of resources used by bursty traffic is not an easy matter, and has been a subject of much research.

In the area of marketing, there are many open questions regarding broadband network services. Hence, while there is a large amount of empirical data regarding the demand for telephone services, such data does not exist for broadband services. The reason behind this is that broadband network technology is still being developed and there is still is no wide scale deployment of broadband services. Indeed, the market structure of such services is currently unknown [Fau95, Var96], and current research efforts seek to investigate and measure the user demand as a function of quality of service, pricing, and application [INDEX].

What can network engineering do in an area where there are many well understood models (e.g., social welfare economics) but also many unknowns (e.g., market structure of broadband networks)? Recall that in order for pricing to lead to efficient use of network resources, it must reflect the actual amount of resources used by connections, hence an important issue is to *quantify* this amount. The actual charge for a network service, in addition to the amount of resources used, will also depend on other factors, such as competition and marketing. The focus of this dissertation is solely on functions for measuring the *usage component* of the charge, which we will refer to simply as ‘‘price’’. Furthermore, whatever the market structure for broadband services turns out to be, quantifying resource usage in specific market segments will still be important, if pricing within

each segment is to reflect resource usage.

In addition to accurately reflecting resource usage, usage-based pricing must take into account the *technological characteristics* of the offered services, and must be *simple* and *efficient* to implement.

Unlike traditional telephone networks that offer a single service (telephone) which occupies a constant amount of bandwidth (56 or 64 Kbps), broadband networks will support traffic with different characteristics, which may vary in time, and different Quality of Service (QoS) requirements expressed in terms of loss probability, delay, and delay variance (jitter). The amount of resources used by such bursty traffic will depend, in addition to its statistical characteristics and QoS, also on how it is multiplexed inside the network and how the network controls the use of its resources. If time division multiplexing is used, then the network would allocate a fixed amount of bandwidth (equal to the peak rate) for each connection. In this case, due to the burstiness of the traffic, network resources would be underutilized. The latter is the motivation for *statistical multiplexing* under which there is no fixed allocation of network resources, rather these are shared *on demand*. Statistical multiplexing is a feature of both ATM (Asynchronous Transfer Mode) networks and IP (Internet Protocol) networks, such as the Internet. With statistical multiplexing, the amount of resources used by a bursty traffic stream varies between its mean rate and its peak rate. The notion of effective bandwidth, which is central to our work, summarizes this quantity into a single scalar.

Broadband networks will offer two types of services which differ in the traffic control mechanisms they use: services with *open loop* control and services with *closed loop* control. Services with open loop control go through *Connection Admission Control (CAC)*. A user sends a connection request to the network containing his desired QoS and a description of his traffic. The network uses this information to check if it has enough resources to satisfy the request. If the network accepts the request, then the requested QoS and the description of the user's traffic (connection traffic descriptor) form the *traffic contract* of the connection. Based on this contract, the network commits to guarantee the QoS of the connection, provided the user's traffic is within (or *conforms to*) the traffic description contained in the contract.

The *Constant Bit Rate (CBR)* and *Variable Bit Rate (VBR)* service categories [ATMF96] defined by the ATM Forum[1] are examples of services with open loop control (Table 1.1). These services are very close to the *Deterministic Bit Rate* and the *Statistical Bit Rate* services defined by

[1]The ATM Forum is an international organization formed in 1991 with the objective of accelerating the use of ATM products and services through a rapid convergence of interoperability specifications.

the ITU-T[2] for Broadband Integrated Services Digital Networks (B-ISDNs) [I.371]. The *guaranteed* service [SPG97] and the *controlled-load* service [Wro97], which are currently being defined by the IETF[3] for Internet's integrated services architecture [BCS94], are also services with open loop control (Table 1.2).

Service name	Usage
Constant Bit Rate (CBR)	real-time applications requiring constant bandwidth (e.g., circuit emulation)
real-time Variable Bit Rate (rt-VBR)	real-time applications requiring variable bandwidth (e.g., compressed video/audio transfer)
non-real-time Variable Bit Rate (nrt-VBR)	non-real-time applications requiring variable bandwidth (e.g., transaction processing systems)
Available Bit Rate (ABR)	non-real-time rate-adaptive applications (e.g., data transfer, video with adaptive rate compression)
Unspecified Bit Rate (UBR)	non-real-time applications (e.g., traditional data transfer)

Table 1.1: **ATM Forum service architecture.**

Service name	Usage
Guaranteed	applications requiring deterministic delay guarantees (e.g., hard real-time video/audio)
Controlled-Load	applications sensitive to overloaded conditions (e.g., adaptive video/audio applications)
Best-effort	applications with no delay bounds (e.g., traditional data transfer)

Table 1.2: **IETF integrated services model.**[4]

Services with closed loop control do not support specific QoS requirements, nor is there a traffic description negotiated at connection setup to which the users must conform. Rather, the user adjusts his traffic rate based on *feedback signals* he receives from the network. These feedback signals provide an indication of the congestion inside the network. If the feedback signals indicate the presence of congestion, then the user must decrease his traffic rate, whereas if they indicate that

[2]The ITU-T is the Telecommunications sector of the International Telecommunications Union (ITU) which is responsible for setting network standards for public telecommunication. ITU-T is the new name for CCITT (International Consultative Committee on Telegraphy and Telephony) .

[3]The Internet Engineering Task Force (IETF) is a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet.

[4]Both the guaranteed and controlled-load service definitions have the status of Internet Drafts. As such, they are still under development.

there is no congestion, then the user is allowed to increase his traffic rate. Because of the required adaptation on the part of the users (applications), the latter have been termed *elastic* in the Internet community [BCS94, She95b]. An example of a service with closed loop control that we will focus on is the Available Bit Rate[5] (ABR) service in ATM networks [ATMF96, CLS96].

How do the two different control mechanisms used in services with open loop and closed loop congestion control affect the way such services are priced? In services with open loop control, there is a traffic contract which contains the agreed QoS and a traffic descriptor. The user is free to send any traffic, as long as it conforms to the traffic descriptor. Since the amount of resources used by a connection might depend on the traffic contract parameters, it is reasonable for a pricing scheme to take into account these parameters. Furthermore, since information between the network and the user occurs only at connection setup[6], and in order to have predictable charges, the user must know at connection setup how he will be priced for the traffic he sends.

On the other hand, in services with closed loop congestion control, the user adjusts his traffic rate based on feedback signals he receives from the network. Hence, the network is the one who decides how much traffic a user can send. For this reason, it is reasonable that pricing affects how the network shares its resources.

Usage-based pricing requires that resource usage is measured for each connection individually. If, as anticipated, the number of connections supported by broadband networks is large, then it would be costly to perform detailed statistical measurements and complex computations for each connection. This prompts the need for simple pricing schemes which require quantities that can be measured efficiently. Simplicity is also required for a pricing scheme to provide the right incentives for efficient use of network resources, since a complex pricing scheme would make it difficult for users to understand how their actions (e.g., decreasing their peak rate) will affect their charge.

1.1.2 Performance Analysis and Traffic Engineering

The problem of quantifying resource usage is closely related to the performance analysis of broadband networks. We will concentrate solely on the buffer overflow probability as a measure

[5] Although ABR services do not support specific QoS, they do support a Minimum Cell Rate (MCR) which represents the minimum rate at which the user can always send traffic (see Sections 1.3.4 and 5.1).

[6] Current work in standardization is adding the capability of renegotiating the traffic contract parameters [Q.2963, ZDE+ 93]. The information exchanged during renegotiation will be similar to that exchanged during connection setup.

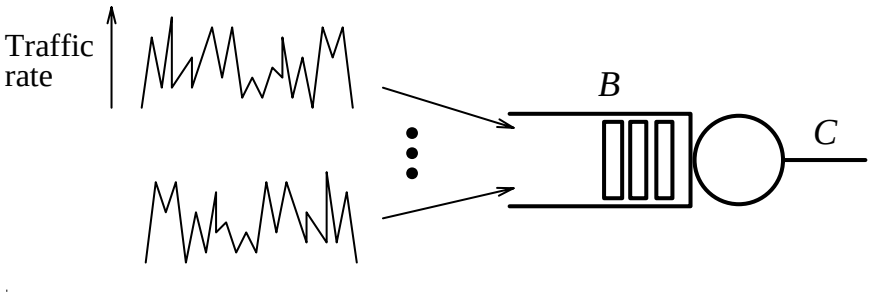


Figure 1.2: **Broadband link.** A basic problem in performance analysis of networks supporting statistical multiplexing is to accurately estimate the overflow probability at a link with a buffer of size B served at rate C (link capacity) which multiplexes bursty traffic.

of Quality of Service (QoS). A basic problem in performance analysis of networks supporting statistical multiplexing is to accurately estimate the overflow probability at a link with a buffer of size B served at rate C (link capacity) which multiplexes bursty traffic (Figure 1.2). The dual problem is to determine the combination of various types of sources that can be multiplexed, while satisfying a target (maximum) overflow probability.

The statistical analysis of traffic measurements has shown that traffic has a self-similar or fractal behavior. This has been shown for many traffic types such as Ethernet traffic [LTWW93, FL91], Internet WAN traffic [PF95], video traffic [BSTW95, GW94], and World Wide Web traffic [CB96]. Such traffic exhibits *long range dependence* or slowly decaying autocorrelation. The implications of such long range dependence is still an open issue (e.g., see [ENW96, GB96, RE96, WTSW97, JLS97] and the references therein). However, recent work [GB96, RE96] has shown that the implications of long range dependence on the buffer overflow probability is of secondary importance when the buffer size is small; this is the case when real-time communication is supported.

The above results motivate the need for procedures to clarify the effects of the time scales of burstiness in real broadband traffic and the effect of link parameters (capacity and buffer) on the performance of a link. Some basic question that such procedures must answer are the following (refer to Figure 1.2): How much does the overflow probability decrease if we increase the buffer size or the capacity of a link? How does traffic shaping affect the multiplexing capability of a link and the amount of resources used by a bursty source?

Information in ATM networks is carried in small, fixed size (48 bytes payload + 5 bytes header) cells. For very small buffer sizes, the discretization of information transfer becomes important for

buffer overflow and can lead to *cell scale* congestion. The latter occurs at a much smaller time scale compared to *burst scale* congestion, where the cell flow is viewed as a continuous fluid flow. It is important to understand when cell scale effects stop to be important and how overflow occurs in very small buffers. Such is the case when ATM switches are designed to fit on a single chip where the total buffer might amount to a few hundred cells to be shared by all links [KSV96]. Cases of very small buffers also arise when delay and jitter requirements are very strict. For example, in order to support two-way high-end video, public carrier switches deployed for WAN services are required to have a cell transfer delay (99th percentile) equal to 150 microseconds per switch [KM94, page 117]. For an output buffered switch with 155 Mbps links, this gives an upper bound on the buffer size of approximately 53 cells.

1.2 Summary of Contributions

The contributions made by this dissertation are related to performance analysis and traffic engineering, usage-based pricing for network services with open loop control, and pricing and resource sharing for ABR services. These contributions are summarized below.

Performance analysis and traffic engineering:
We apply and evaluate the many sources asymptotic for performance analysis and traffic engineering, and the effective bandwidth for quantifying resource usage. Our experiments are performed for link capacities and buffer sizes that will be used in broadband networks, and involve real traffic such as MPEG-1 compressed video and Internet Wide Area Network (WAN) traffic, as well as modeled voice traffic. Our approach demonstrates how the many sources asymptotic can be used to provide important engineering insight to the complex phenomena that occur at a link which multiplexes a large number of sources.

- *Overflow probability and link utilization:* Although the many sources asymptotic with the Bahadur-Rao improvement overestimates the cell loss probability by 1-2 orders of magnitude, the utilization it achieves is very close to the maximum achievable utilization. For example, when MPEG-1 compressed video is being multiplexed in a link with capacity 155 Mbps and buffer larger than 1 msec ($\approx$ 351 cells), the many sources asymptotic with its Bahadur-Rao improvement can achieve a utilization that is over 94% of the maximum utilization. Hence,

the effective bandwidth can accurately quantify resource usage in broadband networks.[7]

- *Application of space and time parameters to traffic engineering:* According to the theory of effective bandwidths and many sources asymptotic, a link's operating point is characterized by two parameters: the space and time parameters. Our simulation results verify the interpretation of the space and time parameters given by theory. Furthermore, we show how the values of these parameters can be used to clarify the effects on the link's performance of the time scales of traffic burstiness, of the traffic mix, of the link parameters (capacity and buffer), and of traffic control mechanisms, such as traffic smoothing.

- *Effects of the traffic mix:* For the traffic mixes considered, the space and time parameters are, to a large extent, insensitive to small variations of the percentage of different source types. This indicates that particular pairs of these parameters can be used to characterize different periods of the day during which the traffic mix remains relatively constant. We propose that the values for these parameters be computed off-line from traffic traces taken at that particular period.

- *Overflow probability for ATM links with a small buffer:* We use the many sources asymptotic to simultaneously capture the cell scale and burst scale effects on the overflow probability at an ATM link with a small buffer which multiplexes periodic on-off sources. This allows us to give the correct expression for the overflow probability for very small buffers, and to investigate the boundary in the link parameter (bandwidth and buffer) space where cell scale effects stop to be important. In addition to accurately estimating the overflow probability, we give a new qualitative description of how overflow occurs in very small buffers.

Usage-based pricing for network services with open loop control:
We investigate usage-based pricing schemes that are based on bounds of the effective bandwidth. These schemes include previously proposed time-volume schemes that involve two measurements (time and volume) for the whole duration of a connection, and two new schemes that involve measurements in distinct time intervals, smaller than the duration of a connection: pricing with

[7]It is important to note that the overestimation of the cell loss probability is not due to the inaccuracy of the many sources asymptotic, but is due to the fact that the asymptotic computes the probability that in a link with infinite buffer the buffer occupancy becomes greater than some threshold, rather than the cell loss probability.

renegotiation and the virtual bucket scheme. The pricing schemes are compared according to their fairness, i.e., their ability to capture the relative amount of resources used by connections. The comparisons are performed for link capacities and buffer sizes that will be used in broadband networks and involve MPEG-1 compressed video with various contents (movies, news, and talk shows) and Internet WAN traffic. The results of our investigations are the following:

- *Time-volume pricing schemes:* In general, the fairness and robustness (with respect to small variations of the link's operating point parameters) of the time-volume pricing schemes is higher for larger link capacities and buffer sizes, for smaller call durations, and for smoother (less bursty) traffic. Indeed, for high capacity links (over 622 Mbps) and short duration connections, the performance of these schemes approaches that of the schemes which involve measurements in distinct time intervals. This indicates that for high bandwidth links with a high degree of statistical multiplexing, simple approximations can achieve good performance.

- *Pricing schemes which involve measurements in distinct time intervals:* Such schemes can achieve higher fairness and robustness, at the cost of higher accounting overhead, compared to the time-volume schemes. Furthermore, the performance of the virtual bucket scheme, which does not entail the signaling overhead due to renegotiation, is very close to the performance of the pricing with renegotiation scheme.

- *Incentive compatibility of the time-volume pricing schemes:* We show, for a particular setup, the incentive compatibility of the time-volume schemes. Furthermore, we compare the equilibrium operating point of a link, in the presence of pricing, for the case of deterministic and statistical multiplexing, and investigate to what extent it is affected by traffic shaping.

Pricing and resource sharing for Available Bit Rate (ABR) services:

We present an approach to pricing and resource sharing for ABR services with the following three features: (i) prices are adjusted according to demand, (ii) users declare the price per unit of time according to which they will be charged, and (iii) resource sharing is based on effective bandwidths. The first feature captures the dynamic nature of congestion which is anticipated for ABR services, while the second and third features allow users to adequately reveal their preferences for network usage in terms of the price per unit of time they are charged, and enable the differentiation of

connections based on their mean rates. Such a property of incentive compatibility is lacking in the existing flow control procedures where the network allocates in a fair way peak rates with no reference to the corresponding mean rates.

An important part of our work involves detailed numerical and simulation experiments. For this reason, we have developed a set of software tools for the performance analysis experiments and for the pricing experiments. For the former, we have developed programs for numerically solving the `supinf` formula given by the many sources asymptotic, by which we compute the values of the space and time parameters and the overflow probability, as well as a program for trace-driven simulation of a single link (see Appendix B for details). The numerical solution of the `supinf` formula receives input from traffic traces, except for the case of periodic on-off sources where it uses an analytic expression for the moment generating function (see Appendix C for details).

For the pricing experiments, in addition to the software implementation of the investigated pricing schemes, we have developed programs for computing the optimal selection of traffic parameters (e.g., mean rate and leaky bucket parameters) as well as a program for simulating a network of switches that is based on a ‘‘fluid’’ model (Appendix B.3); the latter is used for the ABR pricing and resource sharing experiments.

In addition to the above software, we have written a set of flexible scripts for performing experiments for a range of link capacities and buffer sizes, and for processing and presenting the results in the form of graphs.

1.3 Related Work

In this section we survey related work, identifying where it differs from our research. References to specific results in the literature will be given, where appropriate, at later parts of the dissertation.

1.3.1 Performance Analysis and Traffic Engineering

The overflow probability in broadband networks will typically be less than 10^{-6}, hence the overflow probability is a *rare event*. For this reason, the theory of *large deviations* [Wei95, SW95], which is a collection of techniques for estimating properties of rare events, lends itself naturally to the study of the overflow probability in broadband networks. Large asymptotics is an application of large

deviations which studies the tail of the overflow probability of a link when some quantity increases. When this quantity is the size of the buffer, we have the *large buffer asymptotic*. When the capacity and buffer per source are kept constant and we study the overflow probability when the size of the system (the link and the multiplexed sources) increases, we have the *many sources asymptotic*.

According to the large buffer asymptotic [CW95, dVW95, EM93, KWC93, and the references therein], the logarithm of the overflow probability decreases linearly with the buffer size B, i.e.,

$$P(\text{overflow}) \approx e^{-\delta B}\,, \tag{1.1}$$

where δ is a positive constant which depends on the link capacity and the statistical properties of the multiplexed traffic.

The large buffer asymptotic takes into account only the *time averaging* of the fluctuations of bursty sources and not the *statistical averaging* when many independent sources are multiplexed together [HW94]. Because of this, it can be overly conservative or too optimistic [CLW94, Kni97]. Furthermore, for traffic with long range dependence, equation (1.1) does not hold. In this case, the logarithm of the overflow probability does not decrease linearly with the buffer size, rather it decreases at a sub-linear rate [DO96, DLO95].

On the other hand, the many sources asymptotic [CW96, BD95] (a similar result is proved in [SG95] for the case of on-off fluid sources) assumes only the stationarity of the multiplexed traffic streams. The overflow probability using the many sources asymptotic can be approximated by

$$P(\text{overflow}) \approx e^{-NI}\,, \tag{1.2}$$

where N is the number of sources and I is the asymptotic *rate function* which depends on the link capacity and buffer size, and on the statistical properties of the multiplexed traffic.

An improvement to the many sources asymptotic can be obtained using the Bahadur-Rao theorem which adds to equation (1.2) a prefactor term of the form $A/\sqrt{N}$, where A depends only on the characteristics of the multiplexed sources and is independent of the number of sources N. The work in [LM97] extended the proof in [CW96] to obtain the improvement, while [HW94, MdV96] introduced the improvement earlier as a heuristic and applied it for on-off Markov fluid sources. A similar improvement has also been used for the large buffer asymptotic [EHL+95], and shown to give good results for video teleconference traffic [EHL+95], but not very good results for MPEG-1 traffic [Kni97].

The work in [RE96, GB96] investigated the importance of long-range dependence on the overflow probability. This work has shown that correlations only up to a certain time scale have an affect on buffer overflow. The work in [MdV96] also noted the importance of the time scales of buffer overflow for traffic characterization, and relates the time parameter of the many sources asymptotic with the results from frequency domain approaches to traffic characterization. The analysis in [RE96, MdV96] used the many source asymptotic.

The *effective bandwidth*, first introduced in [Kel91, GH91] ([Hui88, GAN91] also contained related notions), is a scalar which gives the amount of resources that must be reserved for a source in order to satisfy its QoS requirements; such a scalar enables problems (e.g., admission control, routing) in broadband networks carrying bursty traffic to be mapped into problems in traditional circuit-switched networks. In [Kel96b] a unifying definition of the effective bandwidth is introduced. The definition summarizes the statistical characteristics of traffic over different time and space scales, and builds on earlier work on effective bandwidths and asymptotic models. Illustrative examples which demonstrate the unifying property of the definition include periodic sources, fractional Brownian input, policed and shaped sources, and deterministic multiplexing.

Central to our work is the notion of effective bandwidth, as defined in [Kel96b], and the many sources asymptotics [CW96, BD95] which can compute the space and time parameters that appear in the effective bandwidth definition (see Section 2.1). We apply and evaluate these techniques for performance analysis of broadband links and for quantifying resource usage for real broadband traffic which includes MPEG-1 compressed video and Internet Wide Area Network (WAN) traffic. Particular emphasis is given on how traffic burstiness over multiple time scales affects the values of the space and time parameters, and how the latter can be used for traffic engineering.

1.3.2 Overflow Probability at an ATM Link with a Small Buffer

Much of the research activity towards estimating the overflow probability at an output link has been conducted by modeling the cells that enter the buffer as a *continuous fluid.* Under such a model, when the buffer size is small, it can be a good approximation to assume that each source being multiplexed is either fully on or fully off over the typical period during which the buffer fills; this gives the simplification that the overflow probability depends on the source characteristics only through the peak and mean rates, hence the bufferless on-off fluid model can be used [Hui88, Kel91].

In ATM networks, when the buffer size is small, overflow can occur due to cell scale congestion; the latter is related to the transfer of information in small, fixed size cells. The fluid model fails to capture such congestion. Prior work [RSKJ91, NRSV91, FLVO94] has modeled the burst and the cell scale component separately at a time before the large deviation analysis of buffer overflow was well understood.

We apply the many sources asymptotic to *simultaneously* capture the effects of both the cell scale and burst scale components on the cell loss probability. In addition to giving the correct expression for the overflow probability, we give a new qualitative description of how overflow occurs for small buffers.

1.3.3 Usage-based Pricing for Network Transport Services

Economic models have been extensively used to study the problem of pricing guaranteed services.[8] In [LV93], such an approach is considered in a context where users separately request bandwidth and buffers at each link along the route of their connection. Under conditions, the state which maximizes the network revenue also maximizes the social welfare of the system containing the network and its users. Equilibrium prices can be computed using a distributed algorithm. In [dVB95], using the notion of effective bandwidth, the model of [LV93] is extended to account for statistical multiplexing and to include best-effort services. In [MMP94], a distributed pricing scheme for allocating bandwidth in virtual paths[9] is considered. It is assumed that there is a single virtual path from every source to its destination. Prices are adjusted iteratively such that the link bandwidth constraints are always satisfied. In [SFY95], QoS provisioning is related to pricing. As in [LV93], resources (buffer and bandwidth) are priced separately. Pareto efficient allocation is proved for specific utility functions. Finally, equilibrium prices are computed using an iterative procedure. In [WPS96], the problem of optimal pricing of guaranteed and best-effort services in ATM networks is addressed. Based on the maximum number of calls that can be carried and the desired blocking probability, optimal prices are computed. For best-effort services, per cell spot pricing is considered. [OS97] considers a network offering multiple service classes to its users and investigates the existence and computation of prices which induce a user to select the service class

[8] Economic models have also been applied for flow control and load balancing, e.g., see [FYN88, San88, WHH+92] and the references therein.

[9] A virtual path is a collection of independent connections (virtual circuits) which are treated similar.

which the network considers most appropriate for that user's traffic type.

A common feature of the above work is that prices are adjusted according to the user demand. There is also work on static pricing schemes. [CESZ91, She95b] argues and presents results showing that QoS sensitive and usage-based pricing can increase the efficient use of network resources. In [CSEZ93], the roles of pricing policies and service disciplines in obtaining efficient use of network resources are investigated. Flat rate pricing and usage-based priority pricing (with static prices) are compared using simulation. [PKF92] investigates the problem of pricing in integrated services networks based on the amount of reserved resources. Simulation results showing the effects of pricing on blocking probability, total revenue, and network utilization are presented. In [PF92], the authors consider a pricing scheme that takes into account bandwidth, buffer, and delay, without however presenting a method for combining all these quantities into a single measure of resource usage.

In addition to how prices are set, the pricing architecture is also important. In [SCEH96], the authors present a critique on the optimality paradigm much of the research on pricing has focused on. It is argued that other structural and architectural issues, such as multicast pricing and receiver pricing, are more important. An architecture, called *edge pricing*, where prices are determined locally at the access point to a network, is presented. Along the same lines, [Cla96b, Cla96a] describes an approach for discriminating users at times of congestion. The scheme allows different users to obtain a different share of bandwidth at times of congestion, based on some profile which can be expressed, e.g., in terms of a leaky bucket. Also, [EMV95] considers the implementation of a billing system for TCP (Transmission Control Protocol) traffic and demonstrates that such a billing system is practical and may be used for controlling congestion by rescheduling time insensitive traffic to less congested times of a day.

From the above work, only [dVB95] uses the effective bandwidth for measuring resource usage. The other work either does not address the issue or assumes that resources, such as bandwidth and buffers, are priced independently. [Kel94, Kel96b, CKW97] consider how pricing schemes linear in time and volume measurements can be created based on bounds of the effective bandwidth. The methods take into account both *a priori* parameters that are declared at connection setup, and *a posteriori* parameters that are measured during the call. An important feature of the approach is that it provides incentives for users to truthfully declare their traffic characteristics, such as their mean

rate.

We investigate usage-based pricing schemes that are based on bounds of the effective bandwidth. The investigated schemes include time-volume schemes where prices are linear in measurements of the duration and transferred volume of a connection [Kel94, Kel96b, CKW97] and schemes which involve measurements in distinct time intervals, which are smaller than the duration of the connection. The above schemes are compared in terms of their fairness, i.e., their ability to capture the relative amount of resources used by connections, for MPEG-1 compressed video traffic with various contents and Internet WAN traffic, and for link capacities and buffer sizes that will be used in broadband networks.

1.3.4 Pricing Best-Effort and Available Bit Rate (ABR) Services

In addition to pricing guaranteed services, economic models have also been considered for pricing best-effort services, such as the service currently offered by the Internet. In [MMV95c], a ''smart market'' approach to pricing is considered. Each packet contains a ''bid'' which indicates how much a user is willing to pay for the transmission of the packet. Routers serve packets in the order of decreasing bids. In [MMV95b], a general framework for pricing congestible network resources is considered and the implications of pricing models on capacity expansion under different competition environments are discussed. In [GSW94b, GSW94a], the Internet is modeled as a collection of servers offering services with different delay (priorities). Prices are adjusted in a decentralized manner based on measurements of the demand (average flow rate) and the observed delay. [WV96, Section 8.4] investigates time-of-day pricing (with two periods) for a single resource, and discusses how the aggregate demand for each of the two periods is affected by prices. [HS95] considers a model of a single link which multiplexes heterogeneous users and presents numerical results showing the effect of prices on network revenue and performance (delay).

In [Odl97], an approach for pricing the Internet, called Paris Metro Pricing (PMP), is presented. The idea is to partition the Internet into several logical networks, each with separate and non-sharable resources. The (fixed) price of bandwidth would be different for each logical network, and the expectation would be that higher priced networks would be less congested than lower priced networks. Hence, the scheme allows a user requiring better performance to switch to a higher priced and less congested logical network.

In [Low97], a general model where users separately request for bandwidth and buffers is considered. The model differs from the one in [LV93] in that users can request both a fixed and a variable amount of each resource. The price the user will pay depends on the reservation he has made for both of these resources. The main results are that, under assumptions for the users' utility, a competitive equilibrium exists and, more importantly, at this equilibrium all users request strictly positive amounts of variable bandwidth and buffer. The latter suggests that in a competitive environment ABR services will indeed be attractive.

In [Kel96b], a unified approach to price guaranteed services and ABR services is considered. The approach charges a higher price for cells that are sent up to the Minimum Cell Rate[10] (MCR) which a user has declared at connection setup. To avoid having users split their traffic, it is assumed that link bandwidth is allocated to ABR connections in proportion to their MCR.

The author in [Kel97] considers an approach which hides the dynamic nature of prices from users, while still achieves the social welfare optimum of the system containing the network and its users. According to the approach, the network allocates resources based on a price per unit of time that users declare at connection setup and dynamic prices which are *internal* to the network. [KMT98] contains rigorous proof that, under conditions, using an iterative update of prices the approach in [Kel97] will converge to the social welfare optimum. Furthermore, the convergence is robust with respect to stochastic effects and propagation delays.

Our approach also uses prices that are adjusted according to the demand and extends the ideas of [Kel97, KMT98] to take into account the bursty nature of traffic. Hence, our approach allows both smooth and bursty users to adequately reveal their preferences through a single variable (the price per unit of time according to which they will be charged). Such a feature of incentive compatibility is lacking in the existing flow control procedures where the network allocates in a fair way peak rates with no reference to the corresponding mean rates. Furthermore, we discuss how our approach can be implemented within the ABR rate-base congestion control framework.

[10]The Minimum Cell Rate (MCR) for an ABR connection is the minimum rate at which the user of the connection is always allowed to send traffic with.

1.4 Dissertation Outline

The rest of the dissertation is structured as follows. Chapter 2 first reviews basic results from the theory of effective bandwidths and many sources asymptotics (Section 2.1). The application of this theory to traffic engineering is discussed in Section 2.2, and Section 2.3 applies and evaluates the theory for real broadband traffic.

Chapter 3 uses the many sources asymptotic to study the overflow probability at an ATM link with a small buffer. In Section 3.2 we study the case where the link multiplexes constant bit rate sources. In Section 3.3 we study the case where the link multiplexes periodic on-off sources and we present a simple but sound argument which, in addition to being accurate, enables us to qualitatively explain the nature of buffer overflow when the buffer is very small and cell scale effects cannot be ignored.

Chapter 4 investigates usage-based pricing schemes for network services with open loop congestion control that are based on bounds of the effective bandwidth. The schemes we investigate include pricing schemes that are linear in time and volume (Section 4.2), and schemes which involve measurements in distinct time intervals, which are smaller than the duration of the call: pricing with renegotiation and the virtual bucket scheme (Section 4.3). Section 4.4 discusses the fairness criteria according to which we compare the pricing schemes (Section 4.5). Section 4.6 shows, for a particular setup, the incentive compatibility of the time-volume schemes for deterministic and statistical multiplexing, and investigates the effect of traffic shaping on a network's equilibrium operating point.

Chapter 5 investigates the problem of pricing and resource sharing for Available Bit Rate (ABR) services. Section 5.2 discusses how dynamic pricing can be used to maximize the social welfare of a system (the network and its users) and how it can be implemented within the ABR congestion control framework. Next, we discuss an approach that hides the dynamic nature of prices from users, which simply declare a price per unit of time according to which they will be charged. Section 5.3 presents a resource sharing approach that is based on effective bandwidths. The implementation of the approach withing the ABR rate-based congestion control framework is discussed, and its ability to differentiate connections with different mean rates is demonstrated through a comprehensive set of simulation experiments.

Chapter 6 summarizes the dissertation and identifies areas for future work.

Finally, there are three appendices. Appendix A presents the traffic sources used in the numerical and simulation experiments, while Appendix B presents some details of these experiments. Appendix C contains the analytical expressions for the logarithmic moment generating function of periodic on-off sources; these are used in the numerical experiments of Chapter 3.

Chapter 2

Performance Analysis and Traffic Engineering

This chapter serves two purposes. First, it reviews basic results from the theory of effective bandwidths and many sources asymptotic which form the mathematical background of our work. Second, it applies and evaluates these techniques for performance analysis and traffic engineering, and for quantifying the amount of resources used by sources. Our experiments involve MPEG-1 compressed video, Internet Wide Area Network (WAN) traffic, and modeled voice traffic.

The *effective bandwidth*, first introduced in [Kel91, GH91] ([Hui88, GAN91] also contained related notions), is a scalar which depends on the statistical properties of a source and, as we will discuss later, the statistical properties of the other traffic it is multiplexed with and the parameters (capacity and buffer) of the multiplexing link. The effective bandwidth gives the amount of resources that must be reserved for the source in order to satisfy its Quality of Service (QoS) requirements; such a scalar enables problems (e.g., admission control, routing) in broadband networks carrying bursty traffic to be mapped into problems in traditional circuit-switched networks. In particular, consider a broadband link which multiplexes N_j sources of type j (Figure 2.1), each having effective bandwidth α_j. The QoS of all the sources will be met if the following linear constraint is satisfied:

$$\sum_{j=1}^{J} N_j \alpha_j \leq C^* , \tag{2.1}$$

where J is the number of source types and C^* is the amount of ''effective capacity'' which depends on the link capacity and buffer, the QoS, and the traffic mix (percentage of the various types of

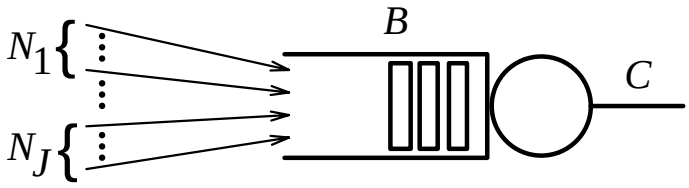


Figure 2.1: **Broadband link model.** There are J types of traffic (N_j sources for each type j), multiplexed in a buffer of size B which is served at rate C.

traffic and their statistical properties).

Asymptotic analysis is concerned with how the buffer overflow probability decays as some quantity increases. If this quantity is the size of the buffer, then we have the *large buffer asymptotic* [CW95, dVW95, EM93, KWC93]. If the buffer per source and capacity per source are kept constant, and we are interested in how the overflow probability decays as the size of the system (the link and the multiplexed sources) increases, then we have the *many sources asymptotic*; this asymptotic regime has been investigated in [CW96, BD95, SG95].

Effective bandwidth definitions based on the large buffer asymptotic were found, in some cases, to be overly conservative or too optimistic [CLW94, CSS97a, Kni97]. This occurs because the large buffer asymptotic does not take into account the gain when many independent sources are statistically multiplexed together. Only recently [Kel96a, CKW97] has it been understood how to incorporate such information into the definition of the effective bandwidth. This work has shown that the effective bandwidth of a source depends on the link's operating point through two parameters, called the *space* and *time* parameters, which in turn depend on the link resources (capacity and buffer) and the statistical properties of the multiplexed traffic. These parameters can be computed using the many sources asymptotic and, as we will demonstrate with real broadband traffic, have important applications to traffic engineering.

2.1 Large Deviation Techniques

This section reviews basic results from the theory of effective bandwidths and many sources asymptotic (Section 2.1.1), and the Bahadur-Rao improvement for the latter (Section 2.1.2).

2.1.1 Effective Bandwidths and Many Sources Asymptotic

Suppose the arrival process at a broadband link is the superposition of independent sources of J types. Let $N_j = N\rho_j$ be the number of sources of type j, and let $\rho = (\rho_1, \rho_2, \ldots, \rho_J)$. This system can be viewed as having N sources of the same type, which consists of a proportion of the J source types and can be characterized by the vector ρ. The broadband link has a shared buffer of size $B = Nb$ and link capacity $C = Nc$. The parameter N is the scaling parameter (size of the system), and parameters b, c are the buffer and capacity per source, respectively. We suppose that after taking into account all economic factors (such as demand and competition) the proportions of traffic of each of the J types remains close to that given by the vector ρ, and we seek to understand the relative usage of network resources that should be attributed to each traffic type.

Let $X_j[0,t]$ be the total load produced by a source of type j in the time interval $(0,t]$, which feeds the above link. We assume that $X_j[0,t]$ has stationary increments. Then, the effective bandwidth of a source of type j is defined as [Kel96a]

$$\alpha_j(s,t) = \frac{1}{st} \log E\left[e^{sX_j[0,t]}\right] , \tag{2.2}$$

where s, t are system parameters which are defined by the context of the source, i.e., the characteristics of the multiplexed traffic and the link resources (capacity and buffer). Specifically, the *time* parameter t (measured in, e.g., milliseconds) corresponds to the most probable duration of the busy period of the buffer prior to overflow. The *space* parameter s (measured in, e.g., kb^{-1}) corresponds to the degree of multiplexing and depends, among others, on the size of the peak rate of the multiplexed sources relative to the link capacity. In particular, for links with capacity much larger than the peak rate of the multiplexed sources, s tends to zero and $\alpha_j(s,t)$ approaches the mean rate of the source, while for links with capacity not much larger than the peak rate of the sources, s is large and $\alpha_j(s,t)$ approaches the maximum value of $X_j[0,t]/t$ over all random variables $X_j[0,t]$.

Let $Q(Nc, Nb, N\rho) = P(\text{overflow})$ be the probability that in an infinite buffer which multiplexes $N\rho = (N\rho_1, N\rho_2, \ldots, N\rho_J)$ sources and is served at rate $C = Nc$, the queue length is above the threshold $B = Nb$. The following holds for $Q(Nc, Nb, N\rho)$ [CW96]:

$$\lim_{N\to\infty} \frac{1}{N} \log Q(Nc, Nb, N\rho) = \sup_t \inf_s \left[st \sum_{j=1}^{J} \rho_j \alpha_j(s,t) - s(ct+b) \right] = -I , \tag{2.3}$$

where I is called the asymptotic *rate function*. The last equation is referred to as the *many sources*

asymptotic, and has been proved for continuous time in [BD95] and for on-off fluid sources in [SG95]. A similar asymptotic holds for the proportion of workload lost through the overflow of a finite buffer of size Nb. Due to equation (2.3), the overflow probability can be written as $P(\text{overflow}) = e^{-NI+o(N)}$, which leads to the following approximation when N is large:

$$P(\text{overflow}) \approx e^{-NI} . \tag{2.4}$$

The accuracy of the above approximation and, more importantly, the achievable link utilization for MPEG-1 compressed video and Internet WAN traffic is investigated in Section 2.3.1.

Consider the QoS constraint on the overflow probability to be $P(\text{overflow}) \leq e^{-\gamma}$, and assume $\gamma = N\gamma_0$. Let $A(N\gamma_0, Nc, Nb)$ be a subset of $\mathbf{Z}_+^J$ such that $(N\rho_1, N\rho_2, \ldots, N\rho_J) \in A(N\gamma_0, Nc, Nb)$ implies $\log P(\text{overflow}) \leq -N\gamma_0$, i.e., the QoS constraint on the overflow probability is satisfied. Because a point in $A(N\gamma_0, Nc, Nb)$ corresponds to a traffic mix under which the overflow probability is below the maximum acceptable value, $A(N\gamma_0, Nc, Nb)$ is called the *acceptance region*. The region $A(N\gamma_0, Nc, Nb)$ is hard to compute. However, the following holds [Kel96a]:

$$\lim_{N\to\infty} \frac{A(N\gamma_0, Nc, Nb)}{N} = A ,$$

where

$$A = \bigcap_{0<t<\infty} A_t , \tag{2.5}$$

and A_t is given by

$$A_t = \left\{ (\rho_1, \rho_2, \ldots \rho_J) : \inf_s \left[st \sum_{j=1}^{J} \rho_j \alpha_j(s,t) - s(ct+b) \right] \leq -\gamma_0 \right\} . \tag{2.6}$$

Hence, the scaled acceptance region $A(N\gamma_0, Nc, Nb)/N$, for large N, can be approximated by A.

If $(\rho_1, \rho_2, \ldots, \rho_J)$ is on the boundary of the region A and the boundary is differentiable at that point, then the tangent plane determines the half-space [Kel96a]

$$\sum_{j=1}^{J} \rho_j \alpha_j(s,t) \leq c + \frac{1}{t}\left(b - \frac{\gamma_0}{s}\right) = c^* , \tag{2.7}$$

where (s,t) is an extremizing pair in equation (2.3) and c^* is the ''effective capacity'' per source. The case for two source types $(J = 2)$ is shown in Figure 2.2.

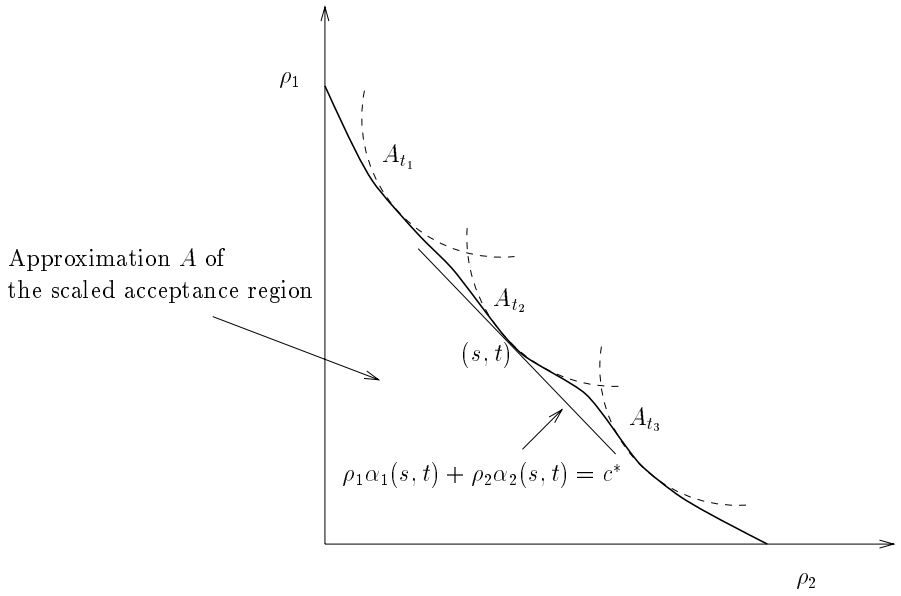


Figure 2.2: **Acceptance region for two source types.** The figure shows the approximation A of the scaled acceptance region $A(N\gamma_0, Nc, Nb)/N$ for three values of t in (2.5). The actual acceptance region can be derived by multiplying with the size N of the actual system.

Since $A(N\gamma_0, Nc, Nb)$ can be approximated by NA, it follows from (2.7) that a point $(N_1, N_2, \ldots, N_J) = (N\rho_1, N\rho_2, \ldots, N\rho_j) \in A(N\gamma_0, Nc, Nb)$ can be taken to satisfy

$$\sum_{j=1}^{J} N_j \alpha_j(s,t) \leq C + \frac{1}{t}\left(B - \frac{\gamma}{s}\right) = C^*\,, \tag{2.8}$$

where, as in (2.7), (s,t) is an extremizing pair in equation (2.3) and C^* is the ''effective capacity'' of the system at the operating point (s,t).

According to (2.8), the effective bandwidth $\alpha_j(s,t)$ provides a measure of resource usage *for a particular operating point* of the link, expressed through parameters s, t. For example, if a source of type j_1 has twice as much effective bandwidth as a source of type j_2, then for this particular operating point of the link, one source of the first type uses twice as much resources than a source of the second type. Note that the QoS guarantees are encoded in the effective bandwidth definition through the value of γ, which appears on the right hand side of (2.8) and affects the shape of the acceptance region.

The asymptotics underlying the above results assume only stationarity of sources. Illustrative examples discussed in [Kel96a] and [RE96] include periodic sources, fractional Brownian input, policed and shaped sources, and deterministic multiplexing.

Unlike the above definition of the effective bandwidth which takes into account the effects of statistical multiplexing, the effective bandwidth based on the large buffer asymptotic depends solely on the characteristics of the source and the QoS constraint. Specifically [CW95, dVW95, EM93, KWC93], consider the QoS constraint $P(\text{overflow}) \leq e^{-\delta B}$, where B is the total buffer. Then, the effective bandwidth of a source of type j based on the large buffer asymptotic is

$$\alpha_j(s) = \frac{1}{s} \lim_{t\to\infty} \frac{1}{t} \log E\left[e^{sX[0,t]}\right] , \tag{2.9}$$

and the constraint on the effective bandwidths is

$$\sum_{j=1}^{J} N_j \alpha_j(\delta) \leq C . \tag{2.10}$$

Observe that (2.9) is a special case of (2.2) for $t \to \infty$. Indeed, equation (2.9) becomes accurate when the buffer is large, in which case the time parameter t becomes large. However, for finite buffer sizes, equation (2.9) can lead to significant underutilization [CLW94, CSS97a, Kni97] or even overutilization of link capacity [CLW94].

2.1.2 Bahadur-Rao Improvement

In this section we discuss an improvement of (2.4), due to [LM97], which is based on the Bahadur-Rao theorem. Similar ideas were introduced as heuristics in [HW94, MdV96]. Then we present an effective bandwidth constraint similar to (2.8) [CSS97a], which takes into account this improvement. An important result is that both the effective bandwidth formula (2.2) and the computation of parameters s, t remains the same, i.e., it uses the `supinf` formula (2.3).

Recently the authors of [LM97] extended the proof in [CW96] to obtain the following:

$$\lim_{N\to\infty} P(\text{overflow}) = \frac{1}{\sqrt{2\pi N\sigma^2 s^2}} e^{-NI} \left(1 + O\left(\frac{1}{N}\right)\right) , \tag{2.11}$$

where (s, t) is an extremizing pair of (2.3) and σ^2 is given by

$$\sigma^2 = \frac{\partial^2}{\partial s^2} \log E\left[e^{sX_j[0,t]}\right] = \frac{M''(s)}{M(s)} - (ct + b)^2 ,$$

and $M(s) = E\left[e^{sX_j[0,t]}\right]$ is the moment generating function. Based on (2.11), we have the following approximation for large N:

$$P(\text{overflow}) \approx \frac{1}{\sqrt{2\pi N\sigma^2 s^2}} e^{-NI} = e^{-NI - \frac{1}{2}\log(2\pi N\sigma^2 s^2)} . \tag{2.12}$$

The term $\frac{1}{2}\log(2\pi N\sigma^2 s^2)$ can be approximated by $\frac{1}{2}\log(4\pi NI)$ [MdV96]. Hence, equation (2.12) does not require any additional computations compared to (2.4).

Next we derive the effective bandwidth constraint similar to (2.8) using the Bahadur-Rao improvement (2.12). If the number of sources $N\rho = (N_1, N_2, \ldots, N_J)$ is such that the overflow probability given by (2.12) is equal to the target overflow probability $\epsilon^{-\gamma}$, then we have

$$-NI - \frac{1}{2}\log(2\pi N\sigma^2 s^2) = -\gamma\,.$$

Substituting $\frac{1}{2}\log(2\pi N\sigma^2 s^2)$ with $\frac{1}{2}\log(4\pi NI)$ in this equation gives

$$-NI - \frac{1}{2}\log(4\pi NI) = -\gamma \Leftrightarrow NI = \gamma - \frac{1}{2}\log(4\pi NI)\,.$$

By setting $NI = \gamma + \epsilon$ in the last equation and taking the expansion of the logarithm on the right-hand side, i.e., $\log(4\pi NI) = \log\left(4\pi(\gamma+\epsilon)\right) \approx \log(4\pi\gamma) + \epsilon/\gamma$, we obtain

$$\begin{aligned}
\gamma + \epsilon &\approx \gamma - \frac{1}{2}\log(4\pi\gamma) - \frac{1}{2\gamma}\epsilon \Leftrightarrow \\
\left(1 + \frac{1}{2\gamma}\right)\epsilon &\approx -\frac{1}{2}\log(4\pi\gamma) \Leftrightarrow \\
\epsilon &\approx -\frac{\frac{1}{2}\log(4\pi\gamma)}{1+\frac{1}{2\gamma}}\,.
\end{aligned}$$

Substituting the last equation in $NI = \gamma + \epsilon$ gives

$$NI \approx \gamma - \frac{\frac{1}{2}\log(4\pi\gamma)}{1+\frac{1}{2\gamma}} = \gamma'\,.$$

Combining the last equation with (2.3) gives the following effective bandwidth constraint in the neighborhood of the extremizing pair (s,t) of (2.3):

$$\sum_{j=1}^{J} N_j\alpha_j(s,t) \le C + \frac{1}{t}\left(B - \frac{\gamma'}{s}\right) = C^*_{\text{B-R}}\,. \tag{2.13}$$

It is important to note that the same formula for the effective bandwidth, given by equation (2.2), is used in both (2.8) and (2.13). The Bahadur-Rao improvement only affects (increases) the amount of effective capacity $C^*_{\text{B-R}} > C^*$. Furthermore, the parameters s, t are computed as before (using formula (2.3)).

2.2 Implications to Traffic Engineering

In this section we discuss the interpretation of the space parameter s and the time parameter t, and how they can be used for traffic engineering [CSS97b]. In Section 2.3 we will apply these ideas to a link multiplexing real broadband traffic.

For any traffic stream, the effective bandwidth $\alpha_j(s,t)$ given by (2.2) is a template that must be filled with the system operating point parameters s,t in order to provide the correct measure of effective usage *for that particular operating point*. Furthermore, as we will see in Section 2.3.3, the values of s,t are, to a large extent, insensitive to small variations of the traffic mix. Since during different times of day, the traffic mix at a link multiplexing a large number of sources is anticipated to remain relatively constant, we can assign particular pairs (s,t) to different periods of the day. The values of s,t corresponding to a particular period of the day can be computed off-line using the `supinf` formula (2.3) and the effective bandwidth formula (2.2), where the expectation in (2.2) is replaced by the empirical mean; the latter can be computed from traffic traces taken during that period of the day.

Recall that the time parameter t corresponds to the most probable duration of the buffer busy period prior to overflow. As we will discuss now, this parameter also identifies the time scale that is important for buffer overflow. Assume that a link is operating at a particular operating point s,t. The effective bandwidth (2.2) depends on the statistical properties of the source through $X[0,t]$, which is the amount of workload produced by the source in a time interval of length t. If two sources have the same distribution of $X[0,t]$ for this value of t, then they will both have the same effective bandwidth. A case of practical interest where this result can be applied is traffic smoothing: To have an effect on the amount of resources used by a connection, traffic smoothing must be performed on a time scale larger than t, since only then does it affect the distribution of $X[0,t]$. We will investigate this with real broadband traffic in Section 2.3.3.

From the above discussion, it follows that the time parameter t also indicates the granularity that traffic traces must have. Hence, it is sufficient to choose the granularity to be a few times smaller than the value of the time parameter. Traditionally, the granularity of traces was chosen in a rather ad-hoc manner.

Now we turn our attention to the space parameter s, which is measured in, e.g, kb^{-1}. By setting $\gamma = -\log P(\text{overflow})$ equal to the right hand side of (2.3) and taking the derivative with respect to

B, we get the following expression for the space parameter s [CKW97]:

$$s = \frac{\partial \gamma}{\partial B} . \tag{2.14}$$

Thus, parameter s is interpreted as the rate at which the logarithm of the overflow probability decreases with the buffer size.

Similar to the above, setting $\gamma = -\log P(\text{overflow})$ equal to the right hand side of (2.3) and taking the derivative with respect to C, we get the following expression for the product st [CKW97]:

$$st = \frac{\partial \gamma}{\partial C} . \tag{2.15}$$

Thus, the product st is interpreted as the rate at which the logarithm of the overflow probability decreases with the link capacity, while the buffer size remains the same.

2.3 Experiments with Real Traffic

In Section 2.1 we presented a mathematical framework for performance analysis of broadband links and for quantifying the amount of resources used by sources, and in Section 2.2 we discussed how this framework can be used for traffic engineering. In this section we apply and evaluate this framework for broadband traffic which includes MPEG-1 compressed video, Internet WAN traffic, and modeled voice traffic [CSS97b]. (Appendix A contains a description of the traffic sources.) The specific issues we address are the following:

- Compare the overflow probability and link utilization using the many sources asymptotic and its Bahadur-Rao improvement to the actual cell loss probability and the maximum utilization. (Section 2.3.1)

- Compare the values of parameters s, t computed by theory to the values estimated using simulation. (Section 2.3.2)

- Estimate and interpret typical values of parameters s, t for real broadband traffic. (Section 2.3.2)

- Investigate how the values of s, t, and subsequently the effective bandwidth depends on the traffic mix. (Section 2.3.3)

We consider link capacities 34 Mbps, 155 Mbps, and 622 Mbps, which correspond to an E-1 physical interface and a Synchronous Digital Hierarchy (SDH) STM-1 and STM-4 physical interface respectively, and a range of buffer sizes which are anticipated for broadband networks. For example, acceptable delay for two-way real-time video is a few hundred milliseconds (e.g., see [Onv94, page 81] and [HW94]). If we assume that the target delay is 200 msec (100 msec one-way) and that there are 10 switches between the source and destination, then the maximum queueing delay[1] at each switch must be in the range $3 - 8$ msec. This value can be higher for one-way transmission and Internet traffic, because of the less stringent requirements in these cases. In our investigations, we consider a buffer size up to 50 msec for MPEG-1 traffic, and a buffer size up to 250 msec for Internet WAN traffic.

2.3.1 Overflow Probability and Link Utilization

In this section we compare the overflow probability and link utilization using the many sources asymptotic and its Bahadur-Rao improvement to the actual cell loss probability and the maximum utilization estimated using simulation. We also derive a simple heuristic for computing the cell loss probability from the overflow probability.

Overflow probability

Figure 2.3 shows, for a fixed number of sources, the overflow probability estimated using the many sources asymptotic and its Bahadur-Rao improvement, and the cell loss probability and frame overflow probability estimated using simulation; the latter is the probability that at least one cell of a frame is lost. Both the cell loss probability and the frame overflow probability are measured using a discrete time simulation model with an epoch equal to one frame time (see Appendix B.2). For each method, the base-10 logarithm of the overflow probability is plotted against the buffer size (measured in milliseconds), while the link utilization remains the same.

In Figure 2.3, we first observe that for small buffer sizes there is a relatively fast decrease of the overflow probability as the buffer size increases. However, for some buffer size (e.g., $5 - 8$ msec for a 155 Mbps link), this stops to occur and increasing the buffer size has a smaller effect on the overflow probability. In addition, the logarithm of the overflow probability in both of these regimes

[1] We have not taken into account propagation and packetization delay [WV96, page 204].

is almost linear with the buffer size. The above indicate that there are two dominant time scales for MPEG-1 traffic: fast time scales which are important for overflow when the buffer size is small and slow time scales which are important for overflow when the buffer size is large. Increasing the buffer size has a greater effect on the overflow probability in the former case.

Second, observe that even though the many sources asymptotic overestimates the Cell Loss Probability (CLP) by approximately 2-3 orders of magnitude, it qualitatively tracks its shape very well. The many sources asymptotic with the Bahadur-Rao improvement overestimates the CLP by 1-2 orders of magnitude. On the other hand, the large buffer asymptotic, in addition to overestimating the CLP by many orders of magnitude, also fails to track its shape.

The actual cell loss probability differs from the overflow probability estimated using the many sources asymptotic and its Bahadur-Rao improvement because the latter is not a measure of the CLP, but a measure of the probability that in an infinite buffer the queue length becomes greater than B. This probability is equal to the frame overflow probability, i.e., the probability that at least one cell of a frame is lost. Indeed, as Figure 2.3 shows, the overflow probability estimated using the many sources asymptotic with the Bahadur-Rao improvement is very close to the frame overflow probability.

Next we derive a simple expression for the cell loss probability in terms of the frame overflow probability $P_{\rm f}$. If one observes a large number of frames, say M, the average number of frames in which we have at least one lost cell is $MP_{\rm f}$. Let $\bar{x}$ be the average number of cells that are lost when a frame overflow occurs. Then the average number of cells that are lost in M frames is $MP_{\rm f}\bar{x}$. Now we can approximate the cell loss probability with the percentage of lost cells, i.e.,

$$\text{CLP} \approx \frac{MP_{\rm f}\bar{x}}{M\bar{F}} = \frac{\bar{x}}{\bar{F}}P_{\rm f}\,, \tag{2.16}$$

where $\bar{F}$ is the average number of cells in a frame.

From the last equation we see that the cell loss probability differs from the frame overflow probability by a correction term $L_{\rm c} = \bar{x}/\bar{F}$. We assume that when an overflow occurs only a few cells are lost. This is expected for small loss probabilities, since it is exponentially harder to lose cells in a buffer of size $B+\epsilon$, than to lose cells in a buffer of size B. In particular, we will assume that only one cell is lost, hence $L_{\rm c} \approx 1/\bar{F}$; since the average number of cells in one frame is 25, we have $L_{\rm c} \approx 1/25 = 10^{-1.4}$. This heuristic agrees with the difference between the frame overflow probability and the cell loss probability that we observe in Figure 2.3.

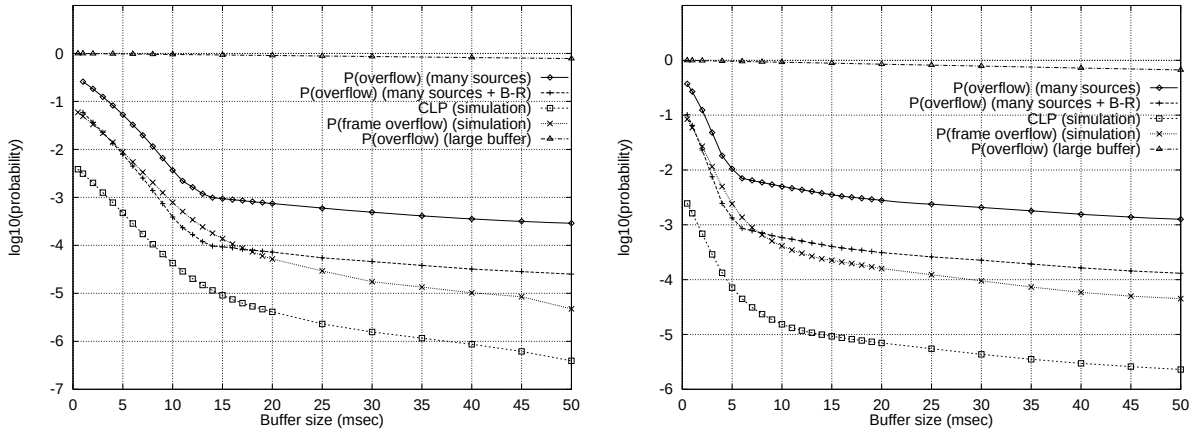
Figure 2.4 shows the cell loss probability estimated using (2.16), where $L_c = 10^{-1.4}$. Observe that the cell loss probability using this heuristic matches the cell loss probability estimated using simulation extremely well.

Link utilization

Let $\varrho = Nm/C$ be the link utilization, where N is the number of sources, m is the mean rate, and C is the link capacity. Figure 2.5 compares, for a range of buffer sizes and for overflow probability 10^{-6}, the link utilization using the many sources asymptotic and its Bahadur-Rao improvement to the maximum achievable utilization (estimated using simulation). Similar to the previous observations regarding the overflow probability, there are significant gains (higher link utilization) in increasing the buffer only up to a certain value. Increasing the buffer above this value has a very small effect on the link utilization.

Recall that the many sources asymptotic overestimated the CLP by 2-3 orders of magnitude. However, as Table 2.1 shows, it performs much better in estimating the maximum utilization. In particular, for $C = 34$ Mbps the many sources asymptotic achieves a utilization which is approximately 79% of the maximum utilization (for a 1 msec buffer). The Bahadur-Rao improvement increases the percentage to 88%. Furthermore, these percentages increase for larger link capacities; e.g., for $C = 155$ Mbps and $B = 1$ msec, the many sources asymptotic achieves a utilization which is 90% of the maximum (Table 2.1(b)). Of course, as Figure 2.6 shows, using the heuristic described above achieves a utilization which almost coincides with the maximum utilization.

Finally, Figure 2.7 shows the link utilization in the case of Internet WAN traffic. Observe that while for *Star Wars* traffic the gains (higher link utilization) in increasing the buffer decrease abruptly, for Internet WAN traffic the gains in increasing the buffer decrease smoother as the buffer size increases. This indicates that there are more time scales in Internet traffic which, for different buffer sizes, become important for overflow.



(a) $C = 34$ Mbps, $\varrho = 0.83$ (b) $C = 155$ Mbps, $\varrho = 0.93$

Figure 2.3: **Overflow probability: theory vs. simulation for Star Wars traffic.** The many sources asymptotic tracks the cell loss probability very well. Note, however, that it overestimates it by 2-3 orders of magnitude. The Bahadur-Rao improvement overestimates the CLP by 1-2 orders of magnitude. Finally, observe that the large buffer asymptotic, in addition to overestimating the CLP by many orders of magnitude, also fails to track its shape.

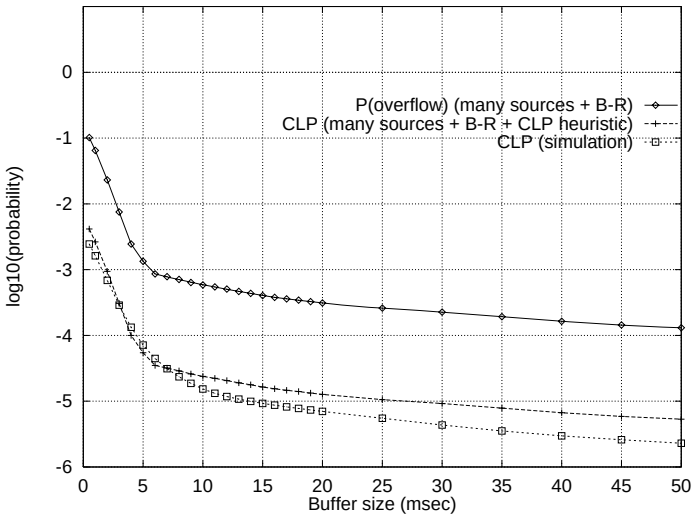
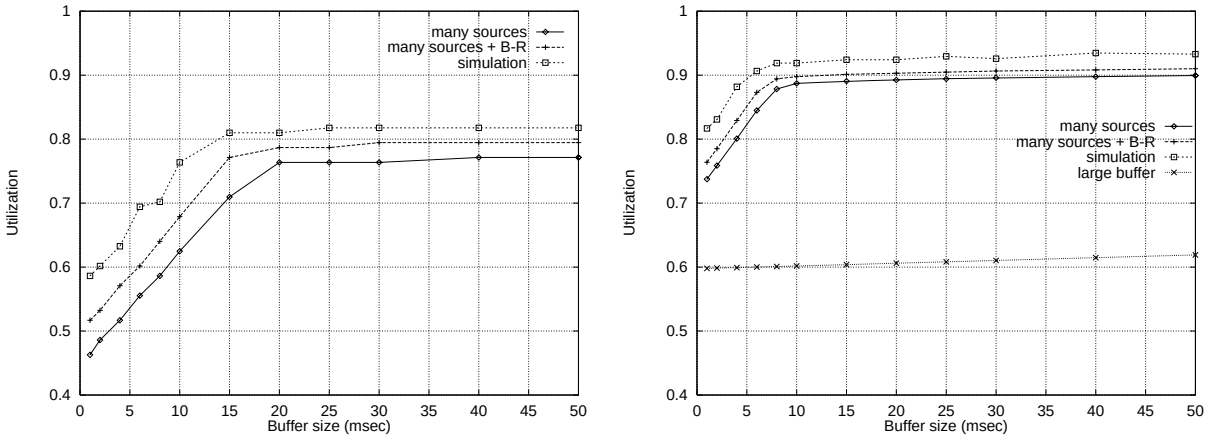


Figure 2.4: **Overflow probability using the many sources asymptotic with the Bahadur-Rao improvement and CLP heuristic (Star Wars traffic).** Observe that the many sources asymptotic with the Bahadur-Rao improvement and CLP heuristic matches the CLP estimated using simulation extremely well. [$C = 155$ Mbps, $\varrho = 0.93$]



(a) $C = 34$ Mbps (b) $C = 155$ Mbps

Figure 2.5: **Link utilization: theory vs. simulation for Star Wars traffic.** The many sources asymptotic with the Bahadur-Rao improvement performs better in utilizing a link than it does in estimating the cell loss probability (Figure 2.3). [$P(\text{overflow}) \leq 10^{-6}$]

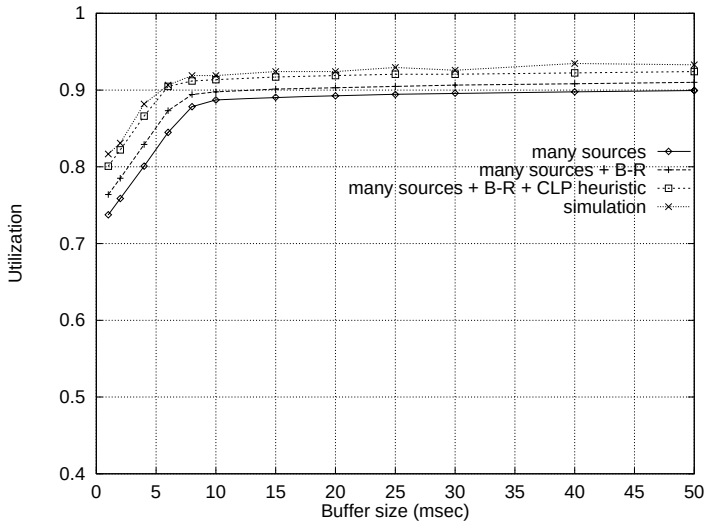


Figure 2.6: **Link utilization using the many sources asymptotic with the Bahadur-Rao improvement and CLP heuristic (Star Wars traffic).** Observe that the many sources asymptotic with the Bahadur-Rao improvement and CLP heuristic achieves practically the maximum utilization (estimated using simulation). [$C = 155$ Mbps]

Buffer msec (cells)	Utilization ϱ		
	Simulation	Many sources	Many sources + B-R
1 (80)	0.57	0.46 (79 %)	0.52 (88 %)
8 (641)	0.70	0.59 (84 %)	0.64 (91 %)
16 (1282)	0.81	0.71 (88 %)	0.77 (96 %)

(a) $C = 34$ Mbps

Buffer msec (cells)	Utilization ϱ		
	Simulation	Many sources	Many sources + B-R
1 (351)	0.82	0.74 (90 %)	0.76 (94 %)
8 (2811)	0.92	0.88 (96 %)	0.89 (97 %)
16 (5622)	0.92	0.89 (96 %)	0.90 (98 %)

(b) $C = 155$ Mbps

Table 2.1: **Link utilization: theory vs. simulation for Star Wars traffic.** The parenthesis contain the percentage of the maximum utilization (second column). Observe that the many sources asymptotic with the Bahadur-Rao improvement achieves a link utilization which is a large percentage of the maximum achievable utilization. [$P(\text{overflow}) \leq 10^{-6}$]

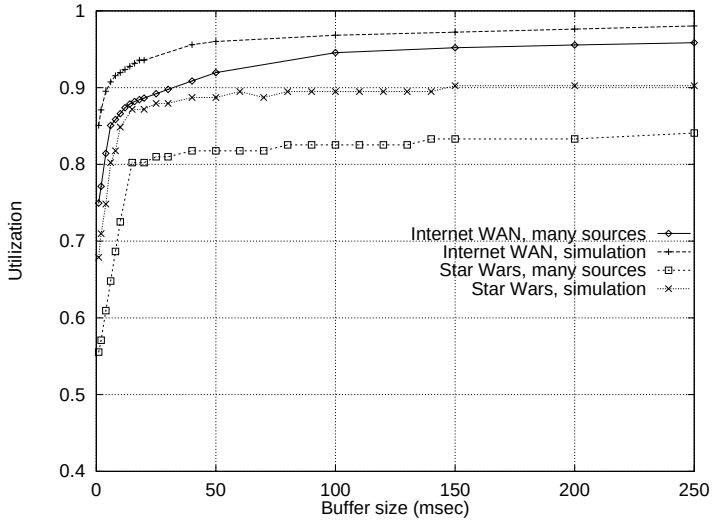


Figure 2.7: **Link utilization for Internet WAN and Star Wars traffic.** As was the case for MPEG-1 traffic, for Internet WAN traffic the many sources asymptotic achieves a high link utilization. However, while for *Star Wars* traffic the gains in increasing the buffer decrease abruptly (there is a sharp knee in the corresponding curve), for Internet WAN traffic the gains in increasing the buffer decrease smoother as the buffer size increases. This indicates that there are more time scales in Internet traffic which, for different buffer sizes, become important for buffer overflow. [$C = 34$ Mbps, $P(\text{overflow}) \leq 10^{-4}$]

2.3.2 Space and Time Parameters

The space s and time t parameters characterize a link's operating point and depend on the characteristics of the multiplexed traffic. In this section we compare the values of parameters s, t computed using the `supinf` formula (2.3) to the corresponding values estimated using simulation. Furthermore, we compute and interpret typical values for these parameters, demonstrating how they can be used for traffic engineering.

Space and time parameters: theory vs. simulation

Recall from the discussion in Section 2.2 that the space parameter s is interpreted as the rate at which the logarithm of the overflow probability decreases with the buffer size, equation (2.14). Motivated by this equation, we can estimate s using the following ratio of differences:

$$s = \frac{\Delta\gamma}{\Delta B}\,, \tag{2.17}$$

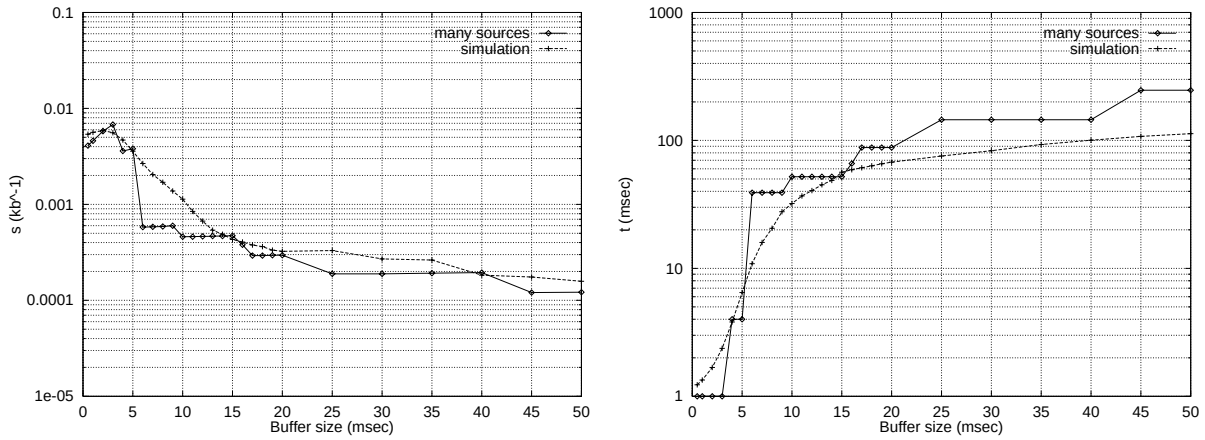
where $\gamma = \log(\text{CLP}_{\text{sim}})$, and CLP_{sim} is the cell loss probability estimated using simulation. Figure 2.8(a) shows that the value of parameter s computed using the `supinf` formula (2.3) is in good agreement with the value computed using (2.17). Note that the ''steps'' in the value computed using the `supinf` formula are expected since the many sources asymptotic (and large deviations theory in general) captures only the most probable way overflow can occur [Wei95, SW95]. On the other hand, the curve labeled ''simulation'' in Figure 2.8(a) includes all the events that contribute to buffer overflow.

Recall from the discussion in Section 2.1.1 that the time parameter t can be interpreted as the most probable duration of the busy period prior to buffer overflow. Figure 2.8(b) compares the value of parameter t computed using the `supinf` formula to the average value of the busy period prior to buffer overflow. As was the case for parameter s, the agreement between the two curves is good.

Typical values and interpretation of the space and time parameters

Next we investigate how parameters s, t and the product st depend on the link capacity and buffer size. The values of s, t are computed using the `supinf` formula for target overflow probability 10^{-7}.

Figure 2.9 shows the values of parameter s as a function of the buffer size, for various link capacities (Figure 2.9(a)) and for various video contents (Figure 2.9(b)). In Figure 2.9(a), the



(a) Parameter s (b) Parameter t

Figure 2.8: **Parameters s, t: theory vs. simulation for Star Wars traffic.** The values of parameters s, t computed by the many sources asymptotic using the `supinf` formula (2.3) are in good agreement with the values estimated using simulation. [$C = 155$ Mbps, $\varrho = 0.93$]

explanation of the steep decrease of the value of s is similar to the explanation of the knee of the curves in Figures 2.3 and 2.5. Specifically, according to equation (2.14), s represents the rate at which the logarithm of the overflow probability decreases with the buffer size. Up to some value, the buffer is used to smooth the fast time scales of the multiplexed traffic. Thus, increasing the buffer has a large effect on the overflow probability, and the value of s is high. Once the fast time scales have been smoothed, the slow time scales govern the buffer overflow. Thus, increasing the buffer has a very small effect on the overflow probability, and the value of s is small. Also, observe in Figure 2.9(a) that the steep decrease of the value of s occurs for smaller buffer sizes (measured in milliseconds) as the link capacity increases. Finally, see Figure 2.9(b), similar behavior is observed for MPEG-1 traffic with various contents. This indicates that the dependence of s on the link capacity and buffer size is related to the MPEG-1 compression algorithm.

The dependence of parameter t on the buffer size is shown in Figure 2.10(a). Observe that the steep increase in its value occurs for the same buffer sizes that the decrease of the value of s occurs (Figure 2.9(a)). The large values of t correspond to the regime where the slow time scales are important for buffer overflow, whereas the small values of t correspond to the regime where fast time scales are important for buffer overflow.

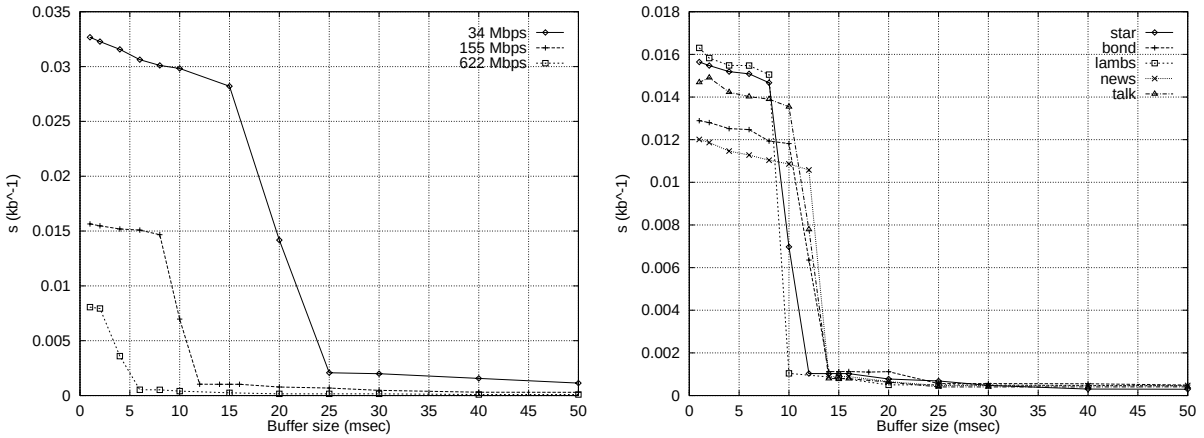
The product st, for different values of buffer size, is shown in Figure 2.10(b). Once again we observe a steep increase in its value, occurring at the same buffer sizes where the changes in the

values of s and t occurred. However, the explanation for the increase in the value of st is more subtle than the explanation for the behavior of s, t. Recall from our discussion in Section 2.2 that the st is interpreted as the rate at which the logarithm of the overflow probability decreases with the link capacity, while the buffer size remains the same; see equation (2.15). Comparing Figure 2.10(a) with Figure 2.10(b), we observe that the large values of st correspond to large values of t. Larger values of t result in an averaging effect in the amount of load $X[0, t]$ that appears in the effective bandwidth formula (2.2). Hence, for the overflow phenomenon, the traffic appears smoother. But for a link which multiplexes smooth traffic and is operating with a cell loss probability greater than zero, a change of the link capacity has a greater effect on the overflow probability compared to a link which multiplexes more bursty traffic. This gives the intuition of why the value of st increases.

Figure 2.11 shows the values of s, t for *Star Wars* and voice traffic (modeled using an on-off Markov fluid, see Appendix A.3). Figure 2.11(a) shows that as the buffer size increases, the value of s for voice traffic decreases smoothly. Furthermore, the rate of decrease is smaller for larger buffer sizes. Comparing the value of s for MPEG-1 and voice traffic we conclude that for buffer sizes up to 2 msec and above 10 msec, increasing the buffer has a larger effect for a network carrying voice traffic compared to a network carrying MPEG-1 traffic. This demonstrates how the values of the space parameter s can be used for buffer dimensioning.

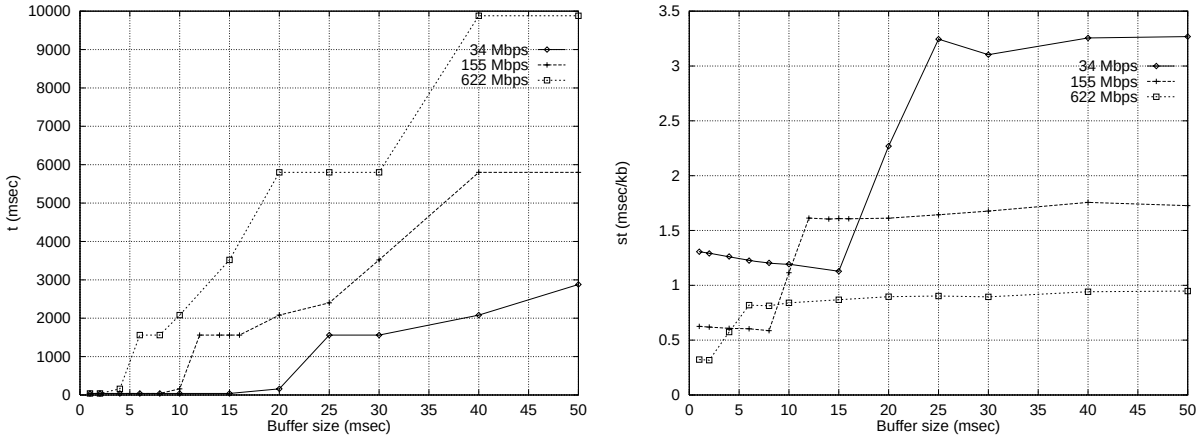
Figure 2.11(b) shows that for voice traffic the time parameter t increases almost linearly with the buffer size. This is expected since for a high degree of multiplexing, voice sources (which are modeled as on-off Markov fluids) behave as Gaussian sources. For such sources, it can be shown [CW96] that the time parameter t increases linearly with the buffer size.

Figure 2.12(a) shows the value of parameter s for *Star Wars* and Internet WAN traffic. For MPEG-1 traffic, the values of s form two distinct sets corresponding to the two distinct time scales that are important for buffer overflow. On the other hand, for Internet traffic the values of s form more than two groups, indicating that for such traffic there are more time scales which, for different buffer sizes, become important for buffer overflow. Recall that this is also the reason behind the smoother dependence of the link utilization on the buffer size for Internet WAN traffic compared to *Star Wars* traffic (Figure 2.7). Finally, Figure 2.12(b) shows that parameter s has different values for different Internet traffic segments, illustrating that different segments can have different statistical properties.



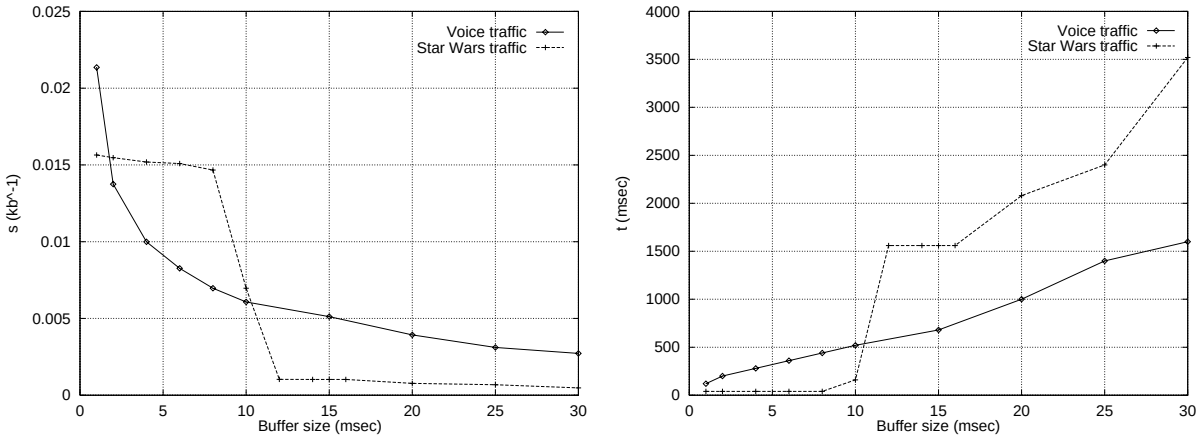
(a) *Star Wars* traffic (b) MPEG-1 traffic with various contents

Figure 2.9: **Parameter s for MPEG-1 traffic.** The value of the parameter s decreases abruptly at some buffer size; this occurs because the buffer has absorbed the fast time scales and only the remaining slow time scales contribute to buffer overflow. The buffer size (measured in msec) for which this occurs decreases as the link capacity increases (left curve). Similar behavior is observed for MPEG-1 traffic with various contents (right figure). [$P(\text{overflow}) \leq 10^{-7}$]



(a) Parameter t (b) Product st

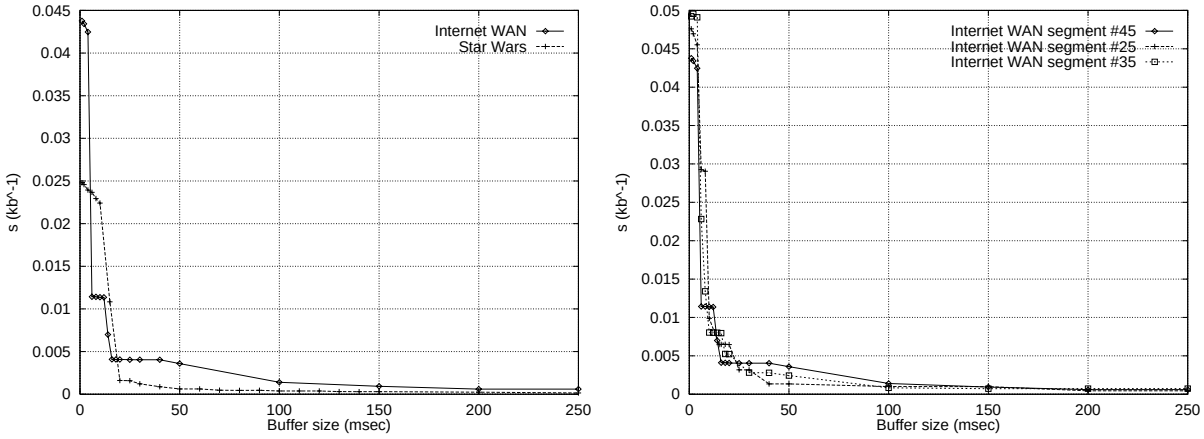
Figure 2.10: **Parameters t and st for Star Wars traffic.** The time parameter t increases as the buffer size increases, indicating that slower time scales become important for buffer overflow (left figure). Observe that the product st abruptly increases for some buffer size (right figure); this coincides with slow time scales becoming important for buffer overflow. [$P(\text{overflow}) \leq 10^{-7}$]



(a) Parameter s

(b) Parameter t

Figure 2.11: **Parameters s, t for Star Wars and voice traffic.** Whereas for MPEG-1 traffic the space parameter s abruptly decreases for some buffer size (indicating that slow time scales become important for buffer overflow), for voice traffic it gradually decreases, with a rate that also decreases as the buffer size increases (left figure). This indicates a smoother change of the time scales for voice traffic. The time parameter t for voice traffic increases linearly with the buffer size, unlike the case of MPEG-1 traffic where it exhibits abrupt jumps. [$P(\text{overflow}) \leq 10^{-7}$]



(a) Parameter s for Internet WAN and *Star Wars* traffic.

(b) Parameter s for different Internet WAN segments.

Figure 2.12: **Parameters s for Internet WAN and Star Wars traffic.** Whereas for MPEG-1 traffic the values of the space parameter s form two distinct groups, corresponding to the two distinct time scales that are important for buffer overflow, for Internet WAN traffic the values of parameter s form more than two groups, indicating that for such traffic there are more than two time scales which, for different buffer size, become important for buffer overflow (left figure). Also, different segments of Internet traffic give different values for parameter s (right figure). [$C = 34$ Mbps, $P(\text{overflow}) \leq 10^{-4}$]

2.3.3 Effects of the Traffic Mix

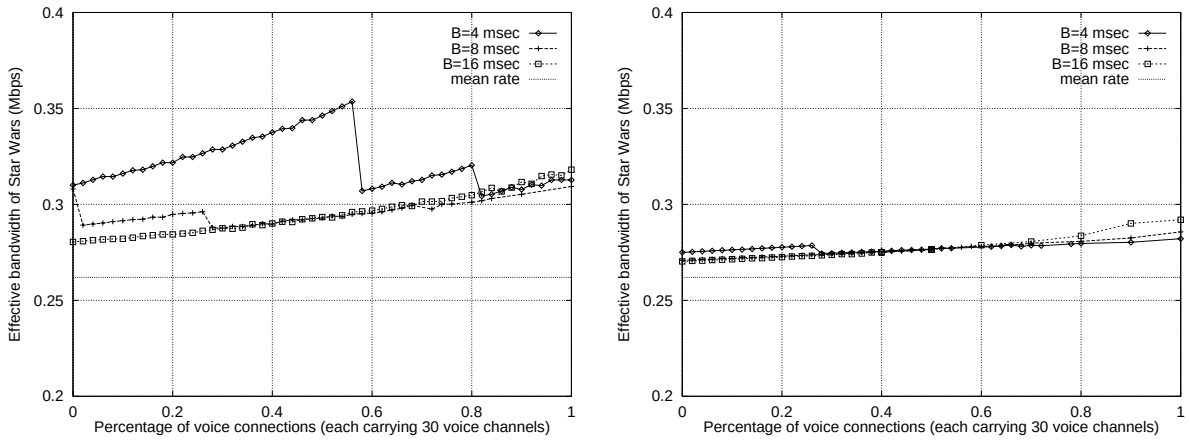
As discussed in Section 2.1.1, particular pairs (s, t) will characterize periods of the day when the traffic mix remains relatively constant. In this section we investigate the dependence of these parameters, hence of the effective bandwidth, on the traffic mix. The traffic mix we consider consists of different types of traffic (MPEG-1 compressed video and voice), traffic with the same structure but different contents (MPEG-1 compressed video with various contents), and smoothed/unsmoothed traffic of the same type and content.

Traffic mix containing Star Wars and voice traffic

We first consider the traffic mix containing *Star Wars* and voice traffic. The horizontal axis in Figures 2.13(a) and 2.13(b) depicts the percentage of voice connections, each containing 30 individual voice channels (modeled using an on-off Markov fluid). The vertical axis of the above figures depicts the effective bandwidth of the *Star Wars* traffic stream. Observe that (1) the effective bandwidth, to a large extent, changes slowly with the traffic mix, (2) the dependence of the effective bandwidth on the traffic mix is smaller for larger capacities and larger buffer sizes, and (3) there are cases where increasing the percentage of voice connections leads to sharp decreases in the value of the effective bandwidth.

The first observation supports the argument that pairs of (s, t) can be assigned to periods of the day where the traffic mix remains relatively constant. However, the third observation indicates that there are percentages of the traffic mix where the effective bandwidth exhibits sharp changes. If the link's operating point is close to such a percentage, we can consider the average effective bandwidth of a source as a linear combination of the effective bandwidth to the left and to the right of the sharp change. The coefficients of the linear combination would be determined by the percentage of the time the network was operating on the left and on the right of the change. Such a linear combination would, on the average, reflect the relative amount of resources used by a source.

The sharp decrease in the value of the effective bandwidth identified above is due to the change of the time scales that are important for buffer overflow. In particular, Figure 2.14 shows that for capacity 155 Mbps and buffer size 4 msec, the time parameter t increases for the same percentage of voice connection at which the sharp decrease in the value of the effective bandwidth occurs in Figure 2.13(a). The increase of t produces an averaging effect (also discussed in Section 2.3.2) in the



(a) $C = 155$ Mbps (b) $C = 622$ Mbps

Figure 2.13: **Dependence of the effective bandwidth on the traffic mix containing Star Wars and voice traffic.** The effective bandwidth of the *Star Wars* sequence changes slowly for certain ranges of the traffic mix. Furthermore, the sensitivity of the effective bandwidth on the traffic mix decreases as the buffer size and link capacity increases. [$P(\text{overflow}) \leq 10^{-7}$]

amount of workload $X[0,t]$ that appears in the effective bandwidth formula (2.2); this averaging results in a smaller effective bandwidth.

Next we show that the value of the time parameter t is related to the MPEG-1 frame pattern, which for the *Star Wars* MPEG-1 sequence is IBBPBBPBBPBB. The average size of the three frame types is largest for I frames and smallest for B frames (Figure 2.15). In Figure 2.14, the graph corresponding to $B = 4$ msec shows that for a small percentage of voice connections $t = 1$, indicating that the buffer overflows when I frames from different sources are synchronized. For some percentage of voice traffic we get $t = 4$, indicating that overflow occurs when IBBP or PBBI patterns from different sources are synchronized (both frame patterns being equally likely). Because of the very small size of B frames (Figure 2.15), the probability of overflow for $t = 4$ is higher than the probability of overflow for $t = 2$ (due to the synchronization, e.g., of IB patterns) or $t = 3$ (due to the synchronization, e.g., of IBB patterns). Finally, for even larger percentage of voice traffic we get $t = 7$, indicating that overflow occurs when IBBPBBP, PBBPBBI, or PBBIBBP patterns from different sources are synchronized. We note that it is a coincidence that $t = 7$ for a link with $C = 155$ Mbps and $B = 4$ msec carrying only voice traffic (Figure 2.14).

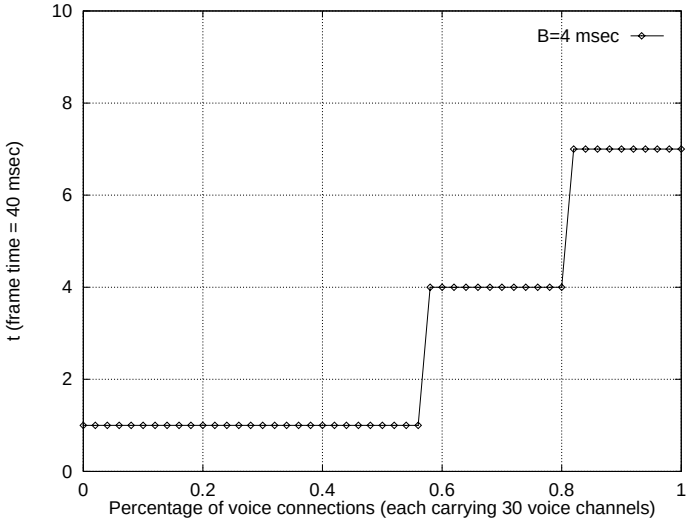


Figure 2.14: **Dependence of parameter t on the traffic mix containing Star Wars and voice traffic.** For a 4 msec buffer the values of the time parameter, 1, 4, and 7 (measured in frames), are closely related to the MPEG-1 frame pattern (Figure 2.15). [$C = 155$ Mbps, $P(\text{overflow}) \leq 10^{-7}$]

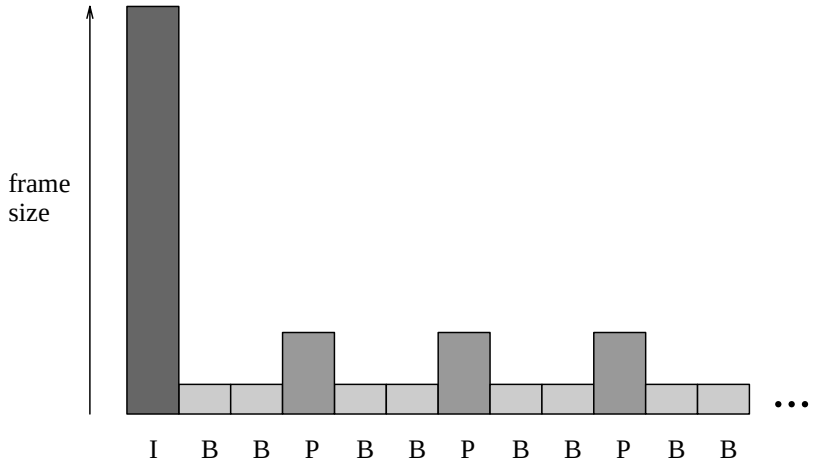
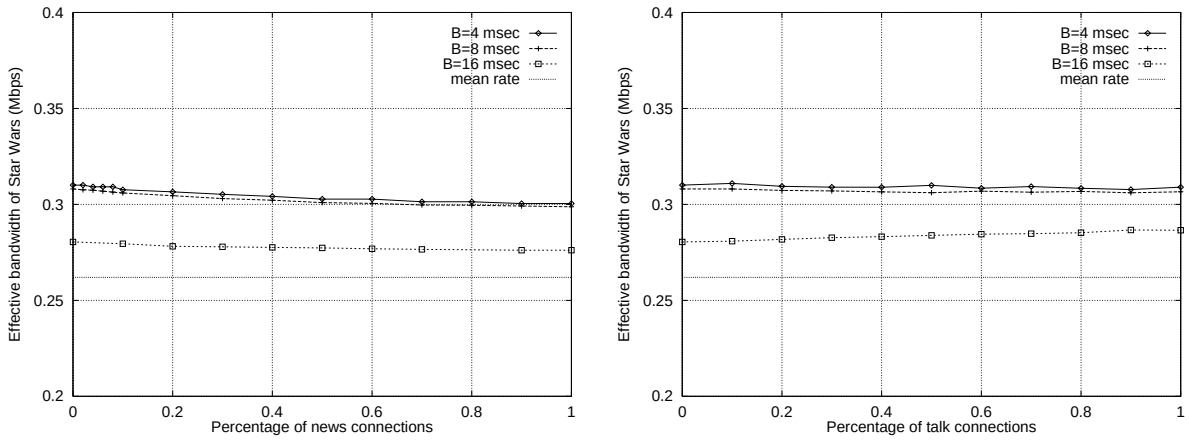


Figure 2.15: **MPEG-1 frame pattern for the Star Wars sequence.** The MPEG-1 frame pattern is IBBPBBPBBPBB (12 frames). The average cells in I, P, and B frames is 115, 26, and 12 cells, respectively. Because of the small size of B frames, for $B = 4$ msec (Figure 2.14), the time parameter takes values $t = 1$ (i.e., overflow is due to the synchronization of I frames), $t = 4$ (i.e., overflow is due to the synchronization of sequences of 4 frames beginning and ending in I and P frames) and $t = 7$ (i.e., overflow is due to the synchronization of sequences containing 7 frames that include one I frame and begin and end with I or P frames).



(a) *Star Wars* + news

(b) *Star Wars* + talk show

Figure 2.16: **Dependence of the effective bandwidth on the traffic mix containing Star Wars + news/talk show traffic.** The figures show that the content of MPEG-1 traffic has a small effect on the effective bandwidth. It is the MPEG-1 frame structure that has an important effect. [$C = 155$ Mbps, $P(\text{overflow}) \leq 10^{-7}$]

Traffic mix containing MPEG-1 traffic with different contents

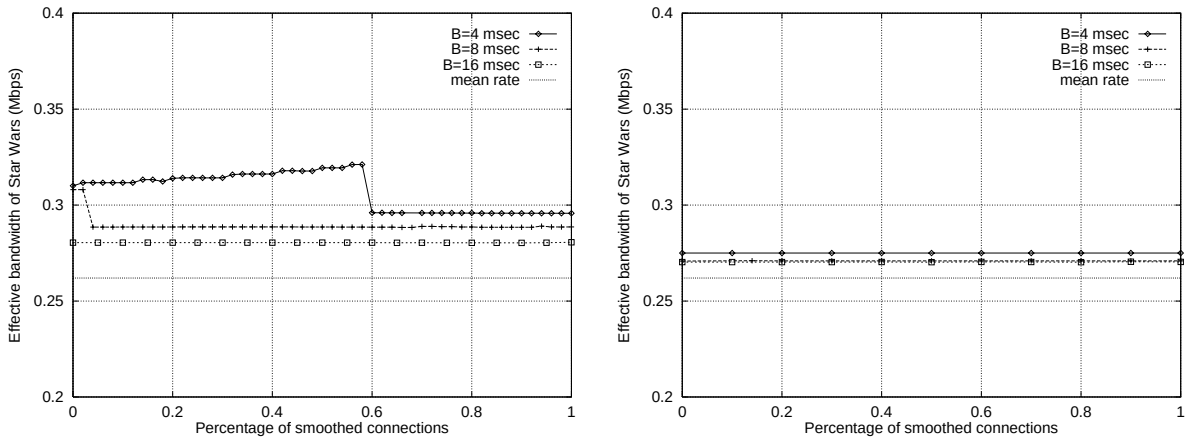
Up to now we investigated the case where the traffic mix contains traffic with different structure. Next we investigate the case where the traffic mix contains MPEG-1 video traffic with the same encoding parameters but with different contents. Figures 2.16(a) and 2.16(b) show the effective bandwidth of the *Star Wars* sequence as a function of the percentage of news and talk show sources, respectively. Observe that the content has a small effect on the effective bandwidth (and on parameters s, t).

Traffic mix containing unsmoothed and smoothed Star Wars traffic

Our final investigation deals with an important question in traffic engineering: How does traffic smoothing affect the link's operating point and the amount of resources used by a source? We will see that parameter t shows the minimum time scale at which smoothing must be performed in order for it to affect resource usage.

Figure 2.17 shows the effective bandwidth of the unsmoothed *Star Wars* traffic stream for different percentages of a traffic mix which consists of unsmoothed and smoothed *Star Wars* traffic; the latter is created by evenly spacing the cells belonging to *two* consecutive frames. First, observe that the effects of the traffic mix on the effective bandwidth decreases when the link capacity and buffer size increases. Second, observe that for some buffer sizes, smoothing has no effect on the effective bandwidth (and on the link's operating point), e.g., in Figure 2.17(a) the curve for $C = 155$ Mbps and $B = 8$ msec, and the curves for $C = 622$ Mbps and $B = 4, 8$, and 16 msec are *flat*. Third, observe that there are cases where increasing the buffer size has a very small effect on the effective bandwidth, e.g., at $C = 622$ Mbps the curves for $B = 8$ msec and $B = 16$ msec practically coincide.

Figure 2.18 shows the effective bandwidth for both the smoothed and unsmoothed *Star Wars* streams, when $C = 155$ Mbps and $B = 4$ msec. When the percentage of smoothed traffic is small, the time parameter t ($= 40$ msec) is smaller than the time interval over which smoothing was performed (80 msec). For this reason, the amount of workload $X[0, t]$ that appears in the effective bandwidth formula (2.2) exhibits less variability for the smoothed stream than for the unsmoothed stream. Hence, the effective bandwidth of the smoothed stream is smaller than the effective bandwidth of the unsmoothed stream. For some percentage of smoothed traffic ($\approx 60\%$), the time parameter t ($= 80$ msec) is no longer smaller than the time interval over which smoothing was performed (80 msec). Because of this, the amount of workload $X[0, t]$ is the same for both the smoothed and unsmoothed streams. Hence, the effective bandwidth of both streams is the same. This also explains why the curve in Figure 2.17(a) that corresponds to buffer $B = 4$ msec is flat when the percentage of smoothed traffic is larger than 60%.



(a) $C = 155$ Mbps (b) $C = 622$ Mbps

Figure 2.17: **Dependence of the effective bandwidth on the traffic mix containing Star Wars and smoothed Star Wars traffic.** Observe that (1) the effects of the traffic mix on the effective bandwidth decreases when the link capacity and buffer size increases, (2) for some buffer sizes, smoothing has no effect on the effective bandwidth, e.g., $C = 155$ Mbps, $B = 8$ msec (left graph) and $C = 622$ Mbps and $B = 4, 8,$ and 16 msec (right graph), and (3) there are situations where increasing the buffer size has a very small effect on the effective bandwidth, e.g., at $C = 622$ Mbps the curves for $B = 8$ msec and $B = 16$ msec practically coincide. [$P(\text{overflow}) \leq 10^{-7}$]

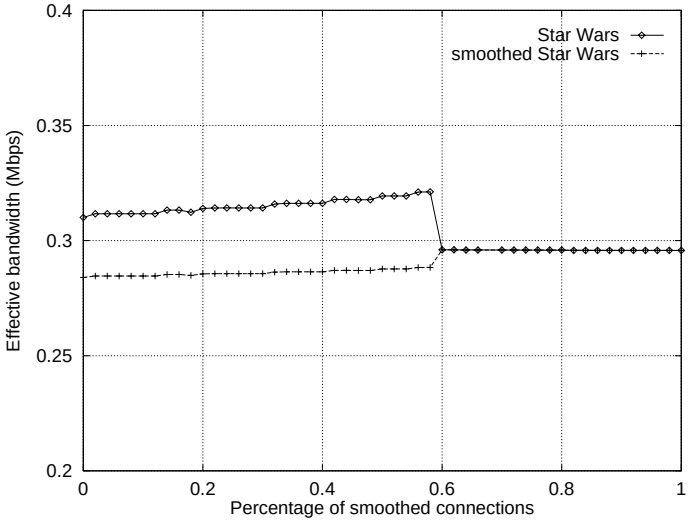


Figure 2.18: **Effective bandwidth for Star Wars and smoothed Star Wars traffic.** When the percentage of smoothed traffic is small, the time parameter t is small and the amount of workload $X[0,t]$ that appears in the effective bandwidth formula (2.2) exhibits less variability for the smoothed stream. Hence, the effective bandwidth of the smoothed stream is smaller than the effective bandwidth of the unsmoothed stream. When the percentage of smoothed traffic increases, parameter t becomes larger than the time interval over which smoothing was performed. Hence, smoothing has no effect, and both unsmoothed and smoothed traffic have the same effective bandwidth. [$C = 155$ Mbps, $B = 4$ msec, $P(\text{overflow}) \leq 10^{-7}$]

2.4 Conclusions

In this chapter we have presented key results from the theory of effective bandwidths and many sources asymptotic which form the mathematical background of our work, and the results from a series of experiments where we apply and evaluate this framework, demonstrating how it can be used to provide important engineering insight to the complex phenomena that occur at a link with a high degree of multiplexing. The experiments involved MPEG-1 compressed video, Internet WAN traffic, and modeled voice traffic. The main results of our investigations are the following:

- *Overflow probability and link utilization:* Although the many sources asymptotic with the Bahadur-Rao improvement overestimates the cell loss probability by 1-2 orders of magnitude, the utilization it achieves is very close to the maximum utilization. For example, in the case of MPEG-1 compressed video that is multiplexed in a link with capacity 155 Mbps and buffer larger than 1 msec ($\approx$ 351 cells), the many sources asymptotic with its Bahadur-Rao improvement can achieve a utilization that is over 94% of the maximum utilization. Hence, the effective bandwidth can accurately quantify resource usage in broadband networks.

- *Application of space and time parameters to traffic engineering:* Our simulation results verify the interpretation of the space and time parameters given by theory. In particular, the space parameter s and the product st indicate how much the logarithm of the overflow probability decreases (i.e., the QoS improves) when the buffer size and link capacity increase, respectively. Moreover, the time parameter t indicates the time scale where traffic control mechanisms, such as smoothing, need to be studied in order to find their impact on resource usage. The time parameter also indicates the granularity that traces must have in order to capture the statistical information that is important for performance analysis. We show how the values of these parameters can be used to clarify the effects on the link's performance of the time scales of traffic burstiness, of the traffic mix, of the link parameters (capacity and buffer), and of traffic control mechanisms, such as traffic smoothing.

- *Effects of the traffic mix:* For the traffic mixes considered, the space and time parameters are, to a large extent, insensitive to small variations of the percentage of different source types. Furthermore, the dependence decreases for larger link capacities and buffer sizes. These results indicate that particular pairs (s, t) can characterize different periods of the day during

which the traffic mix remains relatively constant. We propose that values of these parameters be computed off-line using traffic traces taken during the corresponding period.

An area for further work includes the application of the above approach for traffic engineering and network dimensioning in a real multi-service network environment that involves a large number of different source types. An interesting question is whether the number of discontinuities of the time parameter, that we have identified for a traffic mix containing two source types, increases with the number of source types, and how the size of these discontinuities are affected. A second area for further work lies in the area of traffic modeling. Rather than developing general models that try to emulate real traffic in any operating environment, a new approach would be to develop models which emulate real traffic *for a particular operating point* of a link, expressed through the pair (s, t). Such models would be simple and efficient to implement, and can be the basis for fast and flexible traffic generators.

Chapter 3

Overflow Probability at an ATM Link with a Small Buffer

One of the key issues discussed in the previous chapter was that of identifying the time scale that is important for buffer overflow. In ATM networks, information is transferred in small, fixed size (53 bytes) cells. When the buffer size is small, this discretization of information transfer can lead to *cell scale* congestion. Cell scale congestion occurs when sources deliver cells to the buffer at nearly the same time, thus producing an aggregate cell arrival rate, in a small interval of the order of the cell inter-arrival time, greater than the link capacity. On the other hand, *burst scale* congestion occurs when the rate of the combined input from the multiplexed sources, which is viewed as a *continuous fluid flow*, is greater than the link capacity.

In this chapter we use the many sources asymptotic to simultaneously capture the cell scale and burst scale effects at an ATM link with a small buffer. This allows us to give the correct expression of the overflow probability for very small buffers, and study the boundary in the link parameter space where cell scale effects stop to be important. Furthermore, we give a new qualitative description of how overflow occurs in very small buffers.

3.1 Cell Scale and Burst Scale Congestion

In this section we present the model we will be using in this chapter, and discuss the two phenomena that will concern us: cell scale and burst scale congestion.

Consider N identical sources, where N is some large number. Each source is a periodic on-off source with deterministic ''on'' and ''off'' phases of duration T_{on} and T_{off}, respectively. During each ''on'' phase the source produces cells at rate h, i.e., when h is measured in cells per second, it produces one cell every $\tau = 1/h$ seconds. Cells enter a shared buffer of size $B = Nb$ which is served at rate $C = Nc$, i.e., when c is measured in cells per second, one cell is served every $1/(Nc)$ seconds. Parameters b, c are the buffer and capacity per source, respectively.

We assume that $\tau = 1/h$ is very small compared to T_{on} and T_{off}. For each source, the start of the first ''on'' phase following time 0 is uniformly distributed on $(0, T_{\text{on}} + T_{\text{off}}]$, independently of other sources. Let $\bar{p}N$ be the average number of sources that are ''on'' and let m be the mean rate of each source, i.e., $\bar{p} = T_{\text{on}}/(T_{\text{on}} + T_{\text{off}})$ and $m = \bar{p}h$. Figure 3.1 shows the model for $N = 4$.

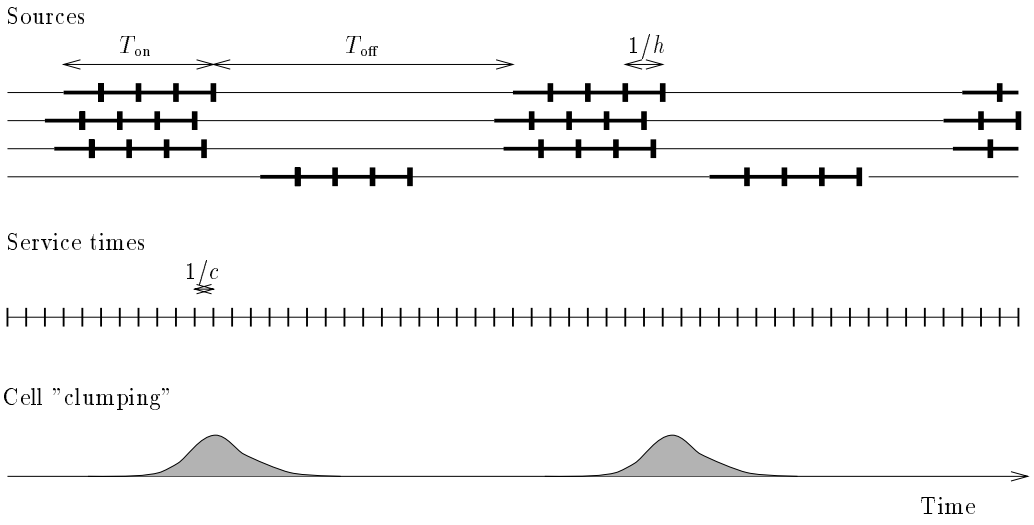


Figure 3.1: **Cell scale congestion.** Cell scale congestion occurs when cell arrivals from different streams become synchronized, and produce ''clumps'' of cells; this produces an aggregate cell rate, in an interval of the order of the cell inter-arrival time, larger than the link capacity.

Cell loss at an ATM output link can arise from both *cell scale* and *burst scale* effects [Rob91]. A burst scale effect takes place when the number of sources that are in their ''on'' phase, say aN, is such that their combined rate, which is viewed as a continuous fluid flow, exceeds the server capacity, i.e., $aNh > Nc$; in this case, the buffer will start to fill. A cell scale effect takes place when the synchronization of the sources which are ''on'' is such that they deliver cells to the buffer at nearly the same time, e.g., as the top three sources in Figure 3.1. Such a synchronization of a large number of sources creates ''clumps'' of cells, which produce an aggregate arrival rate, in a small interval of time of the order of the cell inter-arrival time, greater than the link capacity C; this can

occur even when only the average number of sources $\bar{p}N$ are ''on'', producing an aggregate arrival rate $\bar{p}Nh < Nc$.

Much of the analysis of queueing and multiplexing in ATM networks has been conducted using a model in which the bursts of discrete cells are replaced by a *continuous fluid*. This fluid is either ''on'', at a constant rate h, or ''off''. Such a model captures the burst scale effects, but ignores any cell scale effects. Whether or not this approximation is good depends on the length of the typical time over which the buffer content increases from 0 to B just prior to cell loss. If this time t is large in comparison to the cell scale $1/h$, then the fluid model is good. Typically, this is the case when b is large. If t is large compared to $1/h$ but small compared to T_{on} and T_{off}, then it can be a good approximation to assume that each source is fully ''on'' or fully ''off'' over the typical period during which the buffer fills; this gives the simplification that the overflow probability depends on the source statistics only through the peak and mean rates, and the bufferless on-off fluid model can be used [Hui88, Kel91].

If the time for buffer overflow t is in the order of magnitude of $1/h$, which occurs when the buffer per source b is very small, cell scale effects can not be disregarded; in this case, the bufferless fluid model may underestimate the overflow probability by several orders of magnitude. As an illustration, suppose we have a link with capacity $C = 622$ Mbps and total buffer $B = 20$ cells, which multiplexes identical on-off sources with peak rate $h = 1$ Mbps and mean rate $m = 0.25$ Mbps such that the link utilization is 0.8. The bufferless on-off fluid approximation gives an overflow probability less than 10^{-8}, when in fact the actual overflow probability is greater than 10^{-4}.

Figure 3.2 shows the logarithm of the overflow probability plotted against the buffer size. Two distinct regimes are immediately evident, each dominated by cell scale and burst scale effects [Rob91]. It is interesting to note that Figure 3.2 is similar to Figure 2.3 (page 33) which shows the logarithm of the overflow probability plotted against the buffer size for MPEG-1 traffic. Both figures show two distinct regimes, where different time scales are important for buffer overflow. In the previous chapter we showed that the many sources asymptotic could identify the time scales that are important for buffer overflow in the case of real broadband traffic and for buffer sizes such that overflow is due to burst scale effects. One of the objectives of this chapter is to show that the many sources asymptotic can also identify the relevant time scales for much smaller buffer sizes, where cell scale effects become important.

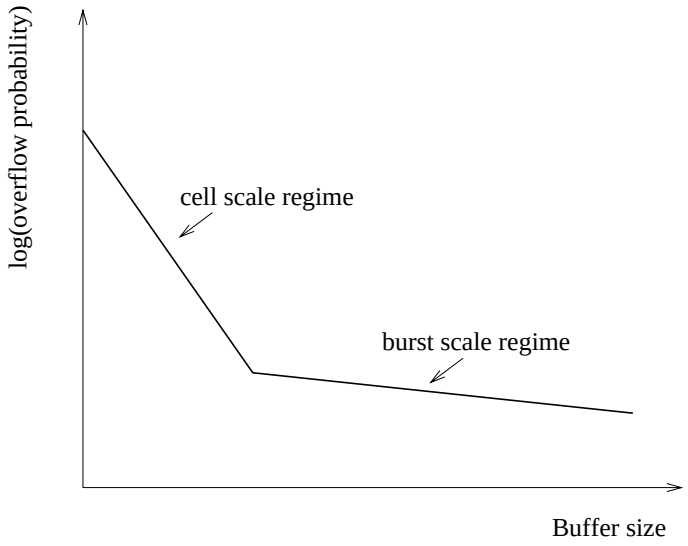


Figure 3.2: **Cell scale and burst scale regimes.** The dependence of the logarithm of the overflow probability on the buffer size shows two distinct regimes where overflow is governed by cell scale and burst scale effects, respectively. Also note that the dependence in each of the two regimes is linear.

Our approach simultaneously captures the effects at the cell scale and burst scale, and accurately computes the cell loss probability. This contrasts with other work which has addressed either the cell scale or the burst scale alone. Using standard ideas from large deviations theory, we are able to investigate the qualitative nature of cell loss for very small buffers [CSW97]. We show that there is a critical buffer size above which cell scale effects are no longer active, in which case it is valid to adopt the fluid model. However, when the amount of buffer per source is very small, both burst and cell scale effects are present (see Figure 3.3). The burst scale effect occurs when the empirically observed proportion of sources which are in their ''on'' phase is above average; the cell scale effect occurs when the ''on'' sources of the above burst are ''in phase'', creating a large cell arrival rate for some short interval of time. In Figure 3.3, these are labeled *empirical mean deviation* and *random phase* effects, respectively.

3.1.1 Asymptotic Approximation of the Overflow Probability

The output link model considered in the previous section is the same as the one shown in Figure 2.1 (page 22), with the only difference that here we assume that all sources are of the same type. Recall from Section 2.1.1 that, based on the many sources asymptotic, the overflow probability can be

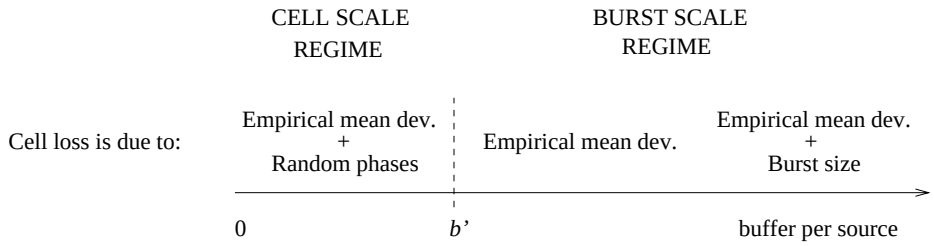


Figure 3.3: **Qualitative description of overflow for small buffers.** For buffer per source values less than b', overflow is due to a combination of two events: a deviation that takes the number of sources that are ''on'' above its mean value, and a synchronization of the random phases of the ''on'' sources. The value of b' depends on the link capacity and the characteristics of the multiplexed traffic (mean and peak rate).

approximated by

$$P(\text{overflow}) \approx e^{-NI} . \tag{3.1}$$

The above formula is asymptotically exact in the sense that $\log_{N\to\infty} \frac{1}{N} \log P(\text{overflow}) = -I$. The rate function I is given by (2.3) which for the homogeneous case we consider reduces to

$$I = -\sup_t \inf_s \left[\varphi(s,t) - s(ct+b)\right] , \tag{3.2}$$

where $\varphi(s,t) = \log E\left[e^{sX[0,t]}\right]$ is the logarithmic moment generating function, and $X[0,t]$ is the number of cells generated by a single source in an interval of length t.

3.2 Constant Bit Rate Sources

In this section we consider independent and identical *constant bit rate* sources whose phases are randomly distributed. This is the model introduced in Section 3.1 with $T_{\text{off}} = 0$. We suppose source i has a rate of h cells per second and generates a new cell at each of the times $\ldots, Y_i - 1/h, Y_i, Y_i + 1/h, Y_i + 2/h, \ldots$, where the Y_i's are independent and uniformly distributed on $[0, 1/h]$. For stability, we require that the utilization is less than one, i.e., $\varrho = h/c < 1$.

3.2.1 Bufferless Case

We first consider the case where the buffer is zero. The probability that cell loss occurs during an interval of length $\Delta = 1/C$, i.e., an interval during which the server can serve exactly one cell, is equal to the probability that the N sources together produce more than one cell during this interval.

A single source produces either 0 or 1 cell during such an interval with probabilities $1 - \varrho/N$ and ϱ/N, respectively. Since the sources are independent, the total number of cells that they produce in an interval of length $1/C$ has the binomial distribution $B(N, \varrho/N)$, whose limit is the probability that a Poisson random variable with mean ϱ is greater than 1, i.e.,

$$P(\text{overflow}) \stackrel{N\to\infty}{\to} 1 - (1+\varrho)e^{-\varrho} .$$

Hence, in the zero buffer case $P(\text{overflow}) \not\to 0$ as $N \to \infty$. The fluid model would give $P(\text{overflow}) = 0$.

3.2.2 Buffered Case

We now consider the case where the server has a buffer. During the interval $(0, t]$, the expected number of cells produced by a single source is ht. Let $k = \lfloor ht \rfloor$ be the greatest integer not exceeding ht. The actual number of cells which are produced by one source in this interval is either k or $k+1$, depending on the phase of the source relative to time 0. The logarithmic moment generating function $\varphi(s,t)$ for random phase constant bit rate sources is

$$\begin{aligned} \varphi(s,t) &= \log E\left[e^{sX[0,t]}\right] \\ &= \log\left[(k+1-ht)e^{sk} + (ht-k)e^{s(k+1)}\right] . \end{aligned} \tag{3.3}$$

Now we show that the extremizing t in (3.2), where $\varphi(s,t)$ is substituted by the above, cannot be greater than $1/c$. Assume that $t > 1/c$, and let $t' = t - 1/c$. The following holds for all s:

$$\begin{aligned} \log E\left[e^{sX[0,t']}\right] - s(ct'+b) &= \log E\left[e^{sX[0,t]-sX[t',t]}\right] - s(ct+b) + s \\ &\geq \log E\left[e^{sX[0,t]}\right] - s - s(ct+b) + s \\ &= \log E\left[e^{sX[0,t]}\right] - s(ct+b) . \end{aligned}$$

The above inequality holds since, by stability, a single source cannot produce more than one cell in an interval of length $1/c$, i.e., $X[t',t] = X[0,1/c] \leq 1$. Hence, we have shown that for every $t > 1/c$ we can find a $t'(= t - 1/c)$ which gives a larger value for $-I$.

If we set $k = 0$ in (3.3) and combine it with (3.2), then we get the following expression for the rate function:

$$I_p(\varrho, b) = -\sup_{0<t\leq 1/c} \inf_s \left\{\log\left[(1-ht) + hte^s\right] - s(ct+b)\right\}$$

$$= \quad - \sup_{0<t\leq 1} \inf_{s} \{\log\left[(1-\varrho t) + \varrho t e^{s}\right] - s(t+b)\}\ , \tag{3.4}$$

where now it is convenient to show I_p as a function of the utilization ϱ and buffer per source b. The suffix p is a reminder that this rate function is for constant bit rate sources and measures a cell scale effect due to their random phases.

3.2.3 Comparison with other Approximations

In this section we compare the overflow probability estimated using the many sources asymptotic (equations (3.1) and (3.4)), with the values obtained using the exact computation and two other approximations.

Recall that $P(\text{overflow})$ is probability that in a system with an infinite buffer the queue length becomes larger than B. The exact formula for this probability is [NRSV91]:

$$P(\text{overflow}) = \sum_{Nb<i\leq N} \left(\begin{array}{c} N \\ i \end{array} \right) \left(\frac{i-Nb}{1/\varrho}\right)^{i} \left(1 - \frac{i-Nb}{1/\varrho}\right)^{N-i} \frac{1/\varrho - N + Nb}{1/\varrho - i + Nb}\,. \tag{3.5}$$

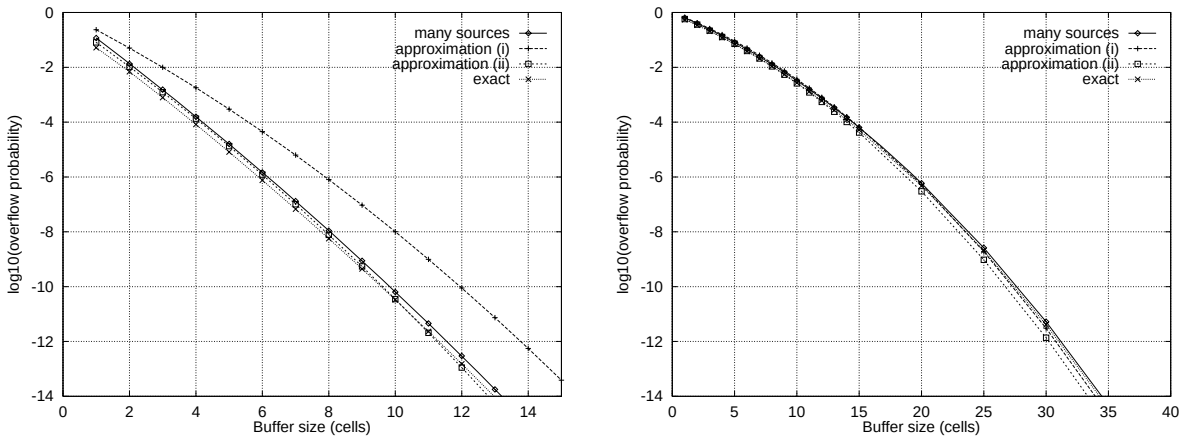
However, this formula gives little insight on how overflow occurs. We also have the following two approximations, which are not asymptotically exact:

$$\text{(i)} \qquad P(\text{overflow}) \approx e^{-\frac{2B^2}{N} - 2B(1-\varrho)} = e^{-N(2b^2+2b(1-\varrho))}\,, \tag{3.6}$$

$$\text{(ii)} \qquad P(\text{overflow}) \approx -\frac{1-\varrho}{\log\varrho} e^{-\frac{2B^2}{N} - B(1-\varrho-\log\varrho)} = -\frac{1-\varrho}{\log\varrho} e^{-N(b^2 - b(1-\varrho-\log\varrho))}\,. \tag{3.7}$$

The first of is based on a Brownian bridge approximation and holds for large N and ϱ close to one [DRS91, NRSV91]. The second appeared in [FLVO94].

Figure 3.4 compares (3.5), (3.6), and (3.7) with the many sources asymptotic approximation based on (2.4) and (3.4). All approximations are close when N is large and ϱ is close to one. Indeed one can show by power series expansion that (3.4), (3.6), and (3.7) differ only by $O((1-\varrho)^2)$. Equation (3.6) is inaccurate when the utilization is small and (3.7) is accurate for the range of overflow probabilities that is of practical concern, i.e., less than 10^{-10}. Even for larger buffer sizes, (3.7) only slightly underestimates the overflow probability.



(a) $\varrho = 0.3$ (b) $\varrho = 0.8$

Figure 3.4: **Overflow probability at a link multiplexing constant bit rate sources.** Both the many sources asymptotic and approximation (ii) accurately determine the overflow probability. Approximation (i) is accurate when the utilization ϱ is close to one. [$C = 155$ Mbps, $h = 1$ Mbps]

3.3 Periodic On-Off Sources

In this section we investigate the overflow probability at a server multiplexing a number of identical periodic on-off sources. Each source is either ''on'' (producing cells at rate h) or ''off'' (not producing cells). The probability that a source is ''on'' is $\bar{p} = T_{\text{on}}/(T_{\text{on}} + T_{\text{off}})$ and $m = \bar{p}h$ is the source's mean rate. The system utilization is $\varrho = m/c$ where, as previously, c is the capacity per source.

We consider two approaches. The first approach is based on a simple, yet sound, intuitive argument which, in addition to being accurate, enables us to *qualitatively* explain the nature of buffer overflow when b is small. The second approach is a direct application of (3.1) and (3.2), where the logarithmic moment generating function in (3.2) is calculated analytically (see Appendix C).

3.3.1 Overflow Probability for Small Buffers

We approximate the overflow probability for small buffers with the probability of the most likely of the following two events: (a) the event of overflow when some number of sources are ''on'' (*empirical mean deviation*) and cells from these sources arrive very close together (*random phases*),

and (b) the event of overflow when the number of ''on'' sources are enough to produce overflow (i.e., without the random phases component of (a)).

We approximate event (a) with the most probable way it can happen. If aN sources are ''on'' during the time which the buffer fills, then overflow occurs according to the constant bit rate model discussed in Section 3.2.2, where the ''effective utilization'' is $\varrho_a = ah/c = a\varrho/\bar{p}$ and the ''effective buffer'' per ''on'' source is $b_a = b/a$. This is the case because, as we have already discussed in the Section 3.1, we assume that T_{on} and T_{off} are large compared to $1/h$ and, in the small buffer regime we are investigating, the time for buffer overflow is of the order of $1/h$. Hence, the probability of (a) is

$$P(\text{event (a)}) = e^{-N \min_a [a I_p(\varrho_a, b_a) + I_f(a)]} . \tag{3.8}$$

The term $e^{-NI_f(a)}$ is the large deviation approximation of the probability that at least aN out of N sources are ''on''. The rate function $I_f(a)$ is given by

$$I_f(a) = a \log \frac{a}{\bar{p}} + (1-a) \log \left(\frac{1-a}{1-\bar{p}} \right) , \tag{3.9}$$

where $\bar{p}$ is the probability of a source being on. The suffix f in the last equation indicates that I_f is the rate function used to estimate the overflow probability in a fluid model.

The probability of event (b) is

$$P(\text{event (b)}) = e^{-NI_f(c/h)} , \tag{3.10}$$

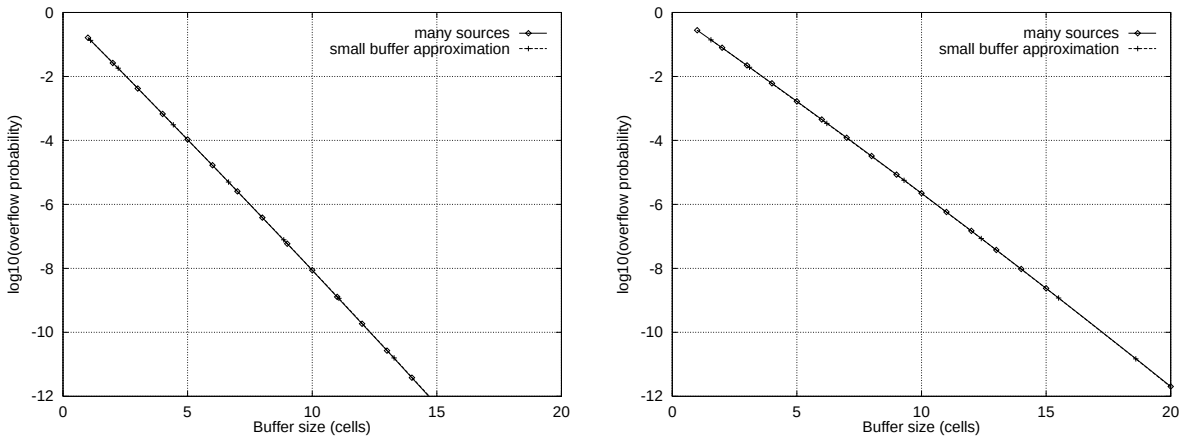
where the rate function I_f is given by (3.9).

Hence, the probability of overflow in a small buffer is approximated by taking the maximum of (3.8) and (3.10)

$$\begin{aligned} P(\text{overflow}) &= \max \left\{ e^{-N \min_a [a I_p(\varrho_a, b_a) + I_f(a)]}, e^{-NI_f(c/h)} \right\} \\ &= e^{-N \min\{\min_a [a I_p(\varrho_a, b_a) + I_f(a)], I_f(c/h)\}} . \end{aligned} \tag{3.11}$$

Accuracy of the small buffer approximation

Figure 3.5 shows that the overflow probability given by the above approximation coincides with the overflow probability computed using the direct application of the many sources asymptotic, i.e., equations (3.1) and (3.2), where the moment generating function $\varphi(s,t)$ for periodic on-off sources is calculated analytically (see Appendix C).



(a) $c/h = 0.7$ ($\varrho = 0.36$) (b) $c/h = 0.5$ ($\varrho = 0.5$)

Figure 3.5: **Overflow probability at a link multiplexing periodic on-off sources.** The overflow probability computed using the small buffer approximation is equal to the overflow probability computed using direct application of the many sources asymptotic. [$C = 155$ Mbps, $h = 1$ Mbps, $\bar{p} = m/h = 0.25$]

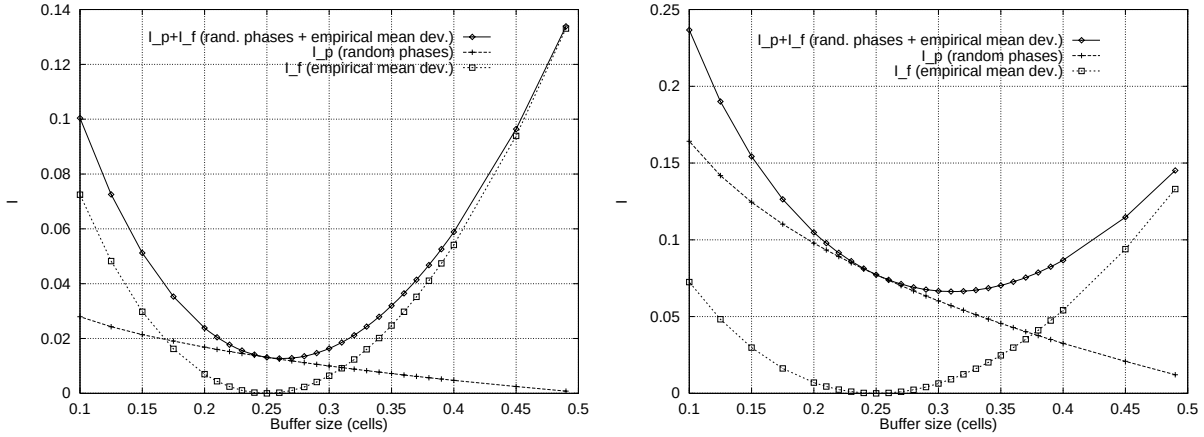
Dependence of a^* on the buffer size

Denote $a^* = \text{argmin}_a[aI_p(\varrho_a, b_b) + I_f(a)]$. Figures 3.6 shows that a^* increases with b. Furthermore, for small buffer sizes, $a^* \approx \bar{p}$, i.e., the number of ''on'' sources is close to the average; in this case overflow is almost completely due to random phase effects.

3.3.2 Boundary between Cell Scale and Burst Scale Regimes

In the bufferless on-off fluid model, cell loss occurs as soon as the aggregate rate of incoming cells exceeds the capacity of the link, i.e., if aN sources are ''on'' and a is such that $aNh > C$, or equivalently $a > c/h$. Thus $P(\text{overflow}) \approx e^{-NI_f(c/h)}$, where I_f is given by (3.9). At the boundary of the cell scale and burst scale regimes, $\min_a\{aI(ah/c, b/a) + I_f(a)\} = I_f(c/h)$; this boundary is shown in Figure 3.7.

Observe that for overflow probabilities of practical interest, the buffer per source is small. Specifically, for sources with $h = 1$ Mbps and $m/h = 0.25$, and for $P(\text{overflow}) = 10^{-8}$, the buffer size above which cell scale effects are no longer important (refer to Figure 3.3), is $b' \approx 0.02$ cells for $C = 622$ Mbps and $b' \approx 0.05$ cells for $C = 155$ Mbps. For such small buffers, $a^* \approx \bar{p} = m/h$, hence



(a) $b = .01$ cells (b) $b = .05$ cells

Figure 3.6: **Contribution of the empirical mean deviation and random phase events to buffer overflow.** As the buffer per source b increases, the contribution due to the empirical mean deviation increases. For small b, the most probable way for overflow to occur is to have the mean number of sources ''on''. [$\bar{p} = m/h = 0.25$, $c/h = 0.5$]

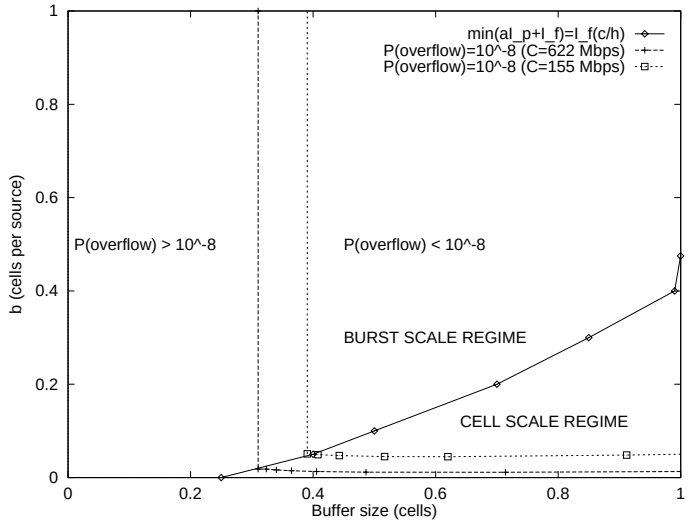


Figure 3.7: **Boundary between cell scale and burst scale regimes.** Cell scale effects are important in the lower right area of the graph. On the boundary between the two regimes $\min\{aI(ah/c, b/a) + I_f(a)\} = I_f(c/h)$. The area where the cell scale effects are important and which corresponds to the range of overflow probabilities of practical interest is located near the x-axis. [$h = 1$ Mbps, $m/h = 0.25$]

from (3.11) the overflow probability is

$$P(\text{overflow}) \approx e^{-NI}, \text{ where } I = \bar{p} I_p(\bar{p}h/c, b/\bar{p}) ,$$

and using (3.7) (since it is the more accurate than (3.6)) from which we drop the prefactor in front of the exponent, we get ($\varrho = h\bar{p}/c$)

$$I = 2b^2/\bar{p} + b(1 - h\bar{p}/c - \log(h\bar{p}/c)) . \tag{3.12}$$

In [RSKJ91, NRSV91, FLVO94], the cell scale component when N periodic on-off sources are multiplexed is approximated by the cell loss in a system of N constant bit rate sources, each with peak rate $h' = m$. The utilization in this system takes the correct value, i.e., $\varrho' = h\bar{p}/c$, but otherwise this approximation has no particular justification. However, (3.7) gives

$$I = 2b^2 + b(1 - h\bar{p}/c - \log(h\bar{p}/c)) . \tag{3.13}$$

Equations (3.12) and (3.13) differ in one term: the first term on the right-hand side of (3.13) is $2b^2$, where it appears that the correct asymptotic has $2b^2/\bar{p}$. However, recall from our previous discussion that for overflow probabilities of practical interest, the buffer per source is small, e.g., $b \approx 0.013$ cells for $C = 622$ Mbps, $P(\text{overflow}) = 10^{-8.4}$, utilization $\varrho = 0.6$, and $m = 0.25$ Mbps. For such values, the contribution of the first term of (3.13) is very small. In conclusion, for overflow probabilities of interest, the rate function can be approximated by

$$I = b(1 - h\bar{p}/c - \log(h\bar{p}/c)) = b(1 - \varrho - \log \varrho) . \tag{3.14}$$

3.3.3 Numerical Results

In this section we present numerical results which show, in practical situations, the buffer size where cell scale effects are no longer important, hence the bufferless on-off fluid approximation can be used. We also show that the many sources asymptotic, in addition to being able to capture both cell scale and burst scale effects, can also capture the effects of different burst sizes.

We consider periodic on-off sources with $h = 1$ Mbps and $m/h = 0.25$, and compare the overflow probability estimated using a bufferless on-off fluid model with the overflow probability estimated using the many sources asymptotic. The results for link capacities 155 Mbps and 622 Mbps are shown in Table 3.1.

Buffer (cells)	$\varrho = 0.4, N = 248$		$\varrho = 0.6, N = 372$		$\varrho = 0.8, N = 496$	
	b	$\log_{10}[P(\text{overflow})]$	b	$\log_{10}[P(\text{overflow})]$	b	$\log_{10}[P(\text{overflow})]$
5	0.020	-3.57	0.013	-2.10	0.010	-0.97
10	0.040	-7.25	0.027	-4.29	0.020	-2.02
15	0.060	-11.05	0.040	-6.57		
20	0.081	-14.96	0.054	-8.94		
on-off fluid		-33.69		-10.70		-2.14

a) $C = 155$ Mbps

Buffer (cells)	$\varrho = 0.4, N = 995$		$\varrho = 0.6, N = 1492$		$\varrho = 0.8, N = 1990$	
	b	$\log_{10}[P(\text{overflow})]$	b	$\log_{10}[P(\text{overflow})]$	b	$\log_{10}[P(\text{overflow})]$
5	0.005	-3.53	0.0034	-2.07	0.0025	-0.95
10	0.010	-7.09	0.0067	-4.16	0.0050	-1.91
15	0.015	-10.67	0.0101	-6.28	0.0075	-2.90
20	0.020	-14.29	0.0134	-8.41	0.0101	-3.90
25	0.025	-17.93	0.0168	-10.57	0.0126	-4.92
30			0.0201	-12.75	0.0151	-5.96
35					0.0176	-7.02
40					0.0201	-8.10
on-off fluid		-135.20		-43.00		-8.60

b) $C = 622$ Mbps

Table 3.1: **Overflow probability for small buffers.** The overflow probability depends primarily on the total buffer and the link utilization. Furthermore, for larger link capacities, the buffer per source for which cell scale effects stop to be important decreases. [$h = 1$ Mbps, $m/h = 0.25$]

Observe that, in the cell scale regime, the overflow probability depends primarily on B and ϱ, and is independent of C. This agrees with (3.14) that leads to an approximation of the overflow probability given by $P(\text{overflow}) \approx e^{-NI} = e^{-B(1-\varrho-\log\varrho)}$ which depends only on B and ϱ.

For fixed utilization and overflow probability, a greater C allows a lesser buffer per source b. This is due to the more efficient statistical multiplexing for larger link capacities. The total amount of buffer for which the overflow probability is less than 10^{-8} grows with the link capacity and the link utilization, and amounts to 40 cells for a 622 Mbps link with $\varrho = 0.8$.

Finally, Figure 3.8 shows the base-10 logarithm of the overflow probability for different buffer sizes. This figure shows the two distinct regimes where overflow is due to cell scale and burst scale effects. Furthermore, the figure also shows that the burst size affects only the burst scale regime. Specifically, smaller burst sizes lead to smaller overflow probabilities and a larger slope for the overflow probability in the burst scale regime.

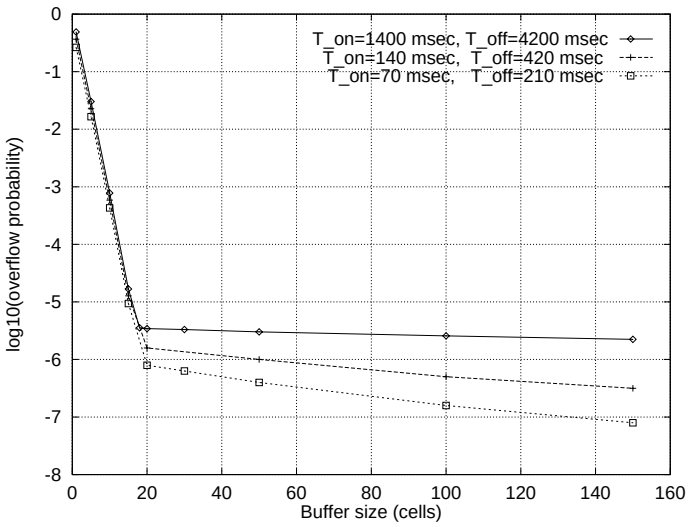


Figure 3.8: **Overflow probability for different buffer and burst sizes.** The many sources asymptotic captures both the cell scale and burst scale components, as well as the effects of the burst size. Observe that the cell scale component is not affected by the burst size. On the other hand, decreasing the burst size leads to lower overflow probabilities and a larger slope for the overflow probability in the burst scale regime. [$C = 155$ Mbps, $h = 1$ Mbps, $m/h = 0.25$, $\varrho = 0.7$]

3.4 Conclusions

In this chapter we have used the many sources asymptotic to investigate the cell scale and burst scale effects on the overflow probability at an ATM outlink link with a small buffer which multiplexes constant bit rate and periodic on-off sources. Our results can be summarized as follows:

- *Capturing cell scale and burst scale effects:* We are able to simultaneously capture both the cell scale and burst scale effects on the overflow probability. This allows us to give the correct expression for the overflow probability. Complemented with the results of the previous chapter, we have shown that the many sources asymptotic is accurate and identifies the time scales that are important for buffer overflow when the buffer size ranges from a few cells, in which case cell scale effects are important for buffer overflow, to thousands of cell, in which case burst scale effects are important for buffer overflow.

- *Qualitative description of buffer overflow:* Applying a simple and sound heuristic, which is motivated by standard ideas of large deviation theory, we are able to give a new qualitative description of the way overflow occurs in small buffers, namely, that overflow occurs due to a combination of two events: a deviation that takes the number of ‘‘on’’ sources above its

mean value, and a synchronization among the random phases of these sources.

- *Boundary between cell scale and burst scale regime:* By simultaneously capturing both the cell scale and burst scale effects we are able to investigate the boundary where cell scale effects stop to be important. We show that for overflow probabilities of practical interest, the buffer per source where cell scale effects stop to be important is very small, e.g., for $C = 155$ Mbps this value is typically less than 0.05 cells per source.

Chapter 4

Usage-based Pricing for Network Services with Open Loop Control

Usage-based pricing enables network providers to recover costs from their users (customers) in a fair manner, hence provide the right incentives for efficient use of network resources. In Chapter 2 we investigated the use of the effective bandwidth for *quantifying* resource usage. In addition to accurately quantifying resource usage, usage-based pricing schemes must be efficient to implement, must be easily understood by users, and must provide the right incentives for efficient and stable network operation.

In this chapter we investigate usage-based pricing schemes for network services with open loop control that are based on bounds of the effective bandwidth. The actual charge for a network service, in addition to the amount of resources used, will also depend on economic factors (e.g., time-of-day pricing) and marketing (e.g., discount for bulk sales), and might contain fixed components, such as fixed connection charges. Our focus is solely on the usage component of the charge, which we will refer to simply as ''price''. The schemes we investigate include simple pricing schemes (called *time-volume* pricing schemes) that are linear in two measured quantities (time and volume), and schemes that involve measurements in *distinct time intervals*, which are smaller than the duration of the call: pricing with renegotiation and the virtual bucket scheme. An important property of usage-based pricing schemes is *fairness*, i.e., their ability to capture the *relative* amount of resources used by connections. Based on this property we evaluate the schemes for MPEG-1 compressed video with various contents (movies, news, talk shows) and Internet WAN traffic, and for link capacities

and buffer sizes that will be used in broadband networks. Finally, we investigate the incentive compatibility of the time-volume pricing schemes for the case of deterministic and statistical multiplexing, and the effect that traffic shaping has on resource usage.

4.1 Technological Characteristics and Desirable Properties of Pricing

In this section we first discuss two important properties of pricing: *incentive compatibility* and *fairness*. Then we discuss the technological characteristics of services with closed loop control, such as the *traffic contract* that is agreed upon by the user and the network at connection setup. Finally, we discuss the relation between pricing and Connection Admission Control (CAC).

The role of pricing is not only to generate income for the service provider, but also to control the use of network resources. For example, prices may be set in a manner which allows users to reduce their charges by shaping their traffic; hence, they can trade additional delay (due to shaping) for smaller charges. This is the key idea behind *incentive compatibility* [Var92, page 441]: Prices should induce users to select network services that better suit their needs, and by doing so they use resources in a way that is good for overall network performance, e.g., in a way that increases social welfare. Pricing schemes which are not incentive compatible give wrong signals to users and result in inefficient use of network resources. For example, in December 1996 America Online switched to flat rate pricing, where users were charged a fixed monthly fee for unlimited access [BR97]. Because of the fixed prices, users had no incentive to disconnect, even when they were not using their connection; this resulted in modems being tied up, blocked from other people who really wanted to use them.

The effect of incentive compatible pricing schemes is to slowly move the global system, which consists of the network and its users, to a stable and efficient operating point. This movement occurs as a result of the pricing loop involving the network and its users: The network announces tariffs which affect the traffic contract selection of users (who seek to minimize their charge), which in turn affects the network's operating point, hence its tariffs, and so on. Badly designed pricing schemes may lead to unstable operation and poor network performance, as was the case for America Online's flat rate pricing scheme.

Well designed pricing schemes should have the *fairness* property. By fairness we mean that prices

should reflect the user's *relative* resource usage, so that a user who makes less use of the network is charged less. This raises an interesting question of when one pricing scheme has better performance compared to another, where the performance is not a measure of the accuracy in determining the absolute amount of resource usage, but is a measure of the ability of prices to capture the relative amount of resources used by users. The issue of fairness will be discussed further in Section 4.4.

In Chapter 2 we investigated the use of the effective bandwidth for *quantifying* resource usage. The effective bandwidth of a source of type j is defined as

$$\alpha_j(s,t) = \frac{1}{st} \log E\left[e^{s X_j[0,t]}\right] , \tag{4.1}$$

where $X_j[0,t]$ is the amount of workload produced in a time interval of length t. The space parameter s (measured in, e.g., kb^{-1}) and the time parameter t (measured in, e.g., milliseconds) are system defined parameters which depend on the characteristics of the multiplexed traffic and the link resources (capacity and buffer). The experiments in Section 2.3 show that, for the traffic mixes considered, these parameters are, to a large extent, insensitive to small variations of the traffic mix. Hence, particular pairs (s,t) can be used to characterize periods of the day during which the traffic mix remains relatively constant.

One approach to pricing would be to directly use formula (4.1). However, this has the following disadvantages. First, it is costly because it requires, for each connection, traces with granularity at most t (the time parameter of the link) and the computation of the logarithmic moment generating function. Second, such an approach leads to a complicated pricing scheme which makes it difficult for users to determine the effect of their decisions (e.g., the reduction of their peak rate) on prices. This would make it difficult for them to behave ''rationally'', which is a requirement for a pricing scheme to lead to stable and efficient network operation.

The pricing schemes we investigate in this section are targeted for services with open loop control. In such services, at connection setup the user and the network agree on a traffic contract. This contract contains the agreed QoS and a traffic descriptor which specifies the maximum traffic the user is allowed to send into the network. As we will see in Section 4.2, pricing schemes can more accurately reflect resource usage if they take into account this traffic descriptor. Examples of services with open loop control are the Constant Bit Rate (CBR) and Variable Bit Rate (VBR) services defined by the ATM Forum [ATMF96], and the guaranteed service [SPG97], and controlled-load service [Wro97] defined by the IETF. The schemes we investigate can also be applied for computing

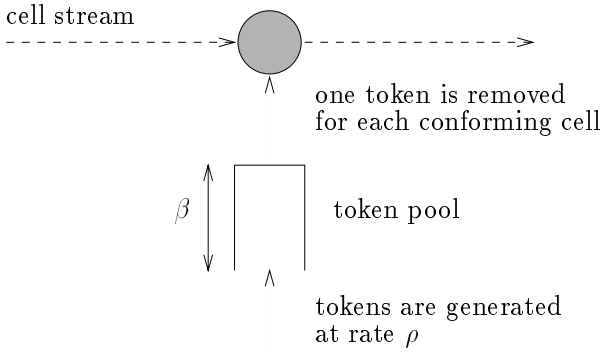


Figure 4.1: **Leaky bucket algorithm.** A *leaky bucket* (β, ρ) consists of a token pool of size β which is filled at rate ρ. One token is removed from the token pool for each conforming cell.

the usage price of a contract among a large user and his Internet service provider.[1] Tariffs for such contracts depend on parameters such as the access speed and volume of the transferred traffic and are currently determined in a fairly ad-hoc manner [Wex96, UUNET97].

A popular algorithm for a traffic descriptor is the *leaky bucket algorithm* [Tur86]. A leaky bucket (see Figure 4.1) consists of a token pool of size β (*bucket size*) which fills at rate ρ (*leak rate*) measured in tokens per second. The leaky bucket is used to police a cell stream in the following way. While the token pool is non-empty, one token is removed from the token pool for each cell. Such a cell is said to be *conforming* to the leaky bucket (β, ρ). On the other hand, if the token pool is empty, the cell is said to be *non-conforming*. Using the leaky bucket, an upper bound $\bar{X}[0,t]$ on the traffic that a source can send in an interval of length t is $\bar{X}[0,t] = \beta + \rho t$.

Our consideration of the leaky bucket is motivated by its use as a traffic descriptor both in ATM networks, since the Generic Cell Rate Algorithm (GCRA) defined in [ATMF94, ATMF96] is equivalent to a continuous state leaky bucket, and in the Internet's intergrated services architecture [CWSA95, BCS94]. Furthermore, a leaky bucket descriptor can be used in contracts among large customers and their service providers (e.g., see [Cla96b, SCEH96]).

[1]The main difference between this contract and, e.g., the ones for CBR and VBR connections is that the network does not commit to provide a certain level of performance.

4.1.1 Pricing and Connection Admission Control (CAC)

In addition to the QoS requirements and the traffic descriptor, the amount of resources used by a connection also depends on how the network controls its resources. For services with open loop control, this is done through Connection Admission Control (CAC). The role of CAC is to determine, at connection setup, whether there are enough resources to accept a connection. In this section we discuss the relation between pricing and CAC.

There are two approaches to connection admission control: *static* and *dynamic*. In order to decide whether to accept or reject a connection request, static CAC relies only on information contained in the traffic contract of connections. Hence, for static CAC the amount of resources used by a connection is determined entirely by its traffic descriptor. Because traffic descriptors, such as the leaky bucket, can provide only a crude characterization of a user's traffic, static CAC can result in low utilization.

Static CAC can be complemented with a pricing scheme that provides the users with the incentive to reveal additional information about their traffic. The network can use this information to improve its resource utilization. As an example, consider a pricing scheme that provides the incentive for users to know their mean rates as accurately as possible (we will discuss such pricing schemes in the next section). This information can be combined with the traffic descriptor of connections to obtain a more accurate estimate of the resources used by all connections, hence can be used by the service provider to achieve higher utilization. In addition to the a priori information contained in the traffic contract, such a pricing scheme can depend on a posteriori information, e.g., the actual mean rate of a connection. Hence, even though the user is free to send traffic at a different mean rate than the value he declared through pricing, doing so would lead to a charge which is higher than if he had declared his actual mean rate from the beginning.

In the case of dynamic CAC, the amount of available resources is estimated using traffic measurements of the connections that have been accepted. Hence, once a connection is accepted, the resources it uses depends on the traffic it sends. A pricing scheme that determines resource usage solely from the traffic description of a connection would be less competitive than a pricing scheme which takes into account the actual amount of traffic that is sent over the connection.

Should a pricing scheme take into account the traffic description contained in the contract? The answer to this question depends on whether such information increases the performance of pricing

and CAC. For example, the time-volume schemes described in Section 4.2 can more accurately reflect resource usage if they take into account the traffic description parameters (peak rate and leaky bucket parameters). This increased accuracy in determining resource usage can also lead to improved CAC [Kel96b]. On the other hand, if the CAC mechanism does not use any such information, e.g., the approach in [CKR+95], a more competitive pricing scheme would take into account only the actual traffic that is sent by the users. Such a pricing scheme could directly use the effective bandwidth formula (4.1) which, as discussed previously, has a high implementation cost and is difficult for users to understand. An alternative would be to take into account part of the information contained in the traffic description. An example of such an approach is the virtual bucket scheme described in Section 4.3.2 which considers only the peak rate of the traffic descriptor. As we will discuss in the following sections, this scheme has a lower implementation cost than the direct application of formula (4.1), but a higher implementation cost compared to the time-volume schemes.

What combinations of CAC and pricing are most appropriate will depend on the target market and the capabilities of the underlying technology. For example, for connections used by large customers whose traffic is an aggregation of the traffic produced by smaller users, it might make sense to use a static CAC mechanism, and price according to a connection's traffic descriptor. This is motivated by the fact that large customers will typically use all the bandwidth that is allowed by their traffic description. On the other hand, pricing schemes that more accurately reflect the actual amount of resources used by connections would make a network provider more competitive. Of course such pricing schemes could have a high implementation cost, which must be weighed against the advantages of the increased competitiveness.

4.2 Pricing Schemes Linear in Time and Volume

The pricing schemes discussed in this section were initiated by the work in [Kel94]. This work presented an approach for creating pricing schemes linear in measurements of time and volume which are based on the effective bandwidth and give users the incentive to provide the network with an accurate estimate of their mean rate. These schemes were later extended in [CKW97] to include an arbitrary number of measurements (taken for the whole connection), and to include both *static*

parameters which are known a priori (e.g., the peak rate and leaky bucket parameters), and *dynamic parameters* which are measured a posteriori (e.g., the duration and total volume of a connection).

The pricing schemes of this section have the following simple form:

$$\text{Price} = aT + bV \ , \tag{4.2}$$

where T, V are the duration and transferred volume of a connection, respectively. Parameters a, b correspond to the user's tariff selection: At connection setup, given his traffic contract, the user is offered a set of possible tariff pairs (a, b) to choose from. A rational user will select the pair (a, b) which minimizes the a priori expected value of his price. According to the theory developed in [Kel94, CKW97], the tariff pairs (a, b) can be appropriately defined so that the expected price for a rational user is $\bar{\alpha}T$, where $\bar{\alpha}$ is some bound of the effective bandwidth. Let $\bar{\alpha}(m, \mathbf{h})$ be an upper bound on the effective bandwidth subject to the mean rate being m and the traffic being within the description contained in the traffic contract $\mathbf{h}$. The function $\bar{\alpha}(m, \mathbf{h})$ is concave in m [CKW97]. If (a, b) are such that $a + bm$ are tangents of the bound $\bar{\alpha}(m, \mathbf{h})$ (see Figure 4.2), then the user will minimize his average price if he selects the tariff pair[2] (a, b) which corresponds to the tangent of $\bar{\alpha}$ at point M, where $M = V/T$ is the user's mean rate. In this case, his price will be $(a + bM)T = \bar{\alpha}(M, \mathbf{h})T$, and his price per unit of time will be $a + bM = \bar{\alpha}(M, \mathbf{h})$.

Note that $\bar{\alpha}T$ is a user's average price only if he knows his mean rate. If he is inaccurate in determining his mean, then his price will be higher. Hence, constructing tariffs using the above approach provides the user with the incentive to know his mean rate as accurately as possible. The user reveals to the network the estimate of his mean rate indirectly through his tariff selection. In the rest of this chapter we assume that users are rational and that they know the exact value of their mean rate. Under these assumptions, the comparison of the pricing schemes reduces to the comparison of the bounds of the effective bandwidth on which they are based. The comparison of pricing schemes is discussed in Section 4.4.

We will consider two bounds[3] of the the effective bandwidth: the on-off bound [Kel94], the simple bound [CKW97], and one approximation of such a bound: the inverted $\mathtt{T}$ approximation [CKSW98].

[2] We assume that users can select from a continuum of tariff pairs (a, b). In practice there will be a small set of tariff pairs, e.g., three pairs corresponding to a small, medium, and large mean rate.

[3] To keep a uniform notation, we will use the same notation $\widetilde{\alpha}$ for both bounds and approximations of a bound of the effective bandwidth.

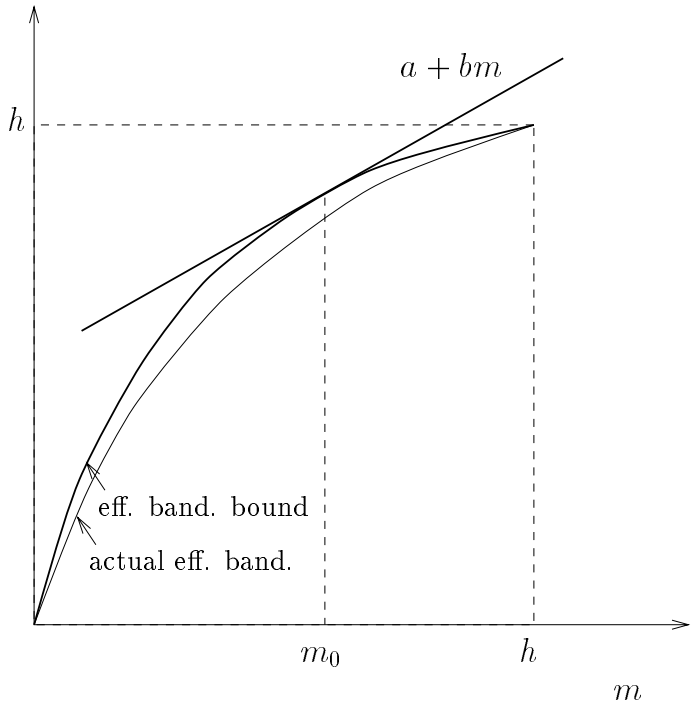


Figure 4.2: **Pricing based on the tangent of an effective bandwidth bound.** The user minimizes his charge if he selects the pair (a, b) which corresponds to his actual mean rate M (i.e., $m_0 = M$). We assume that users are rational and know the exact value of their mean rate. Under these assumptions, the comparison of pricing schemes reduces to comparison of the effective bandwidth bounds on which they are based.

4.2.1 On-Off Bound

The on-off bound [Kel94] depends solely on a connection's peak rate h, which we assume is policed, and its mean rate m, and is given by

$$\widetilde{\alpha}_{\text{on-off}}(m,h) = \frac{1}{st}\log\left[1 + \frac{m}{h}\left(e^{sht} - 1\right)\right] . \tag{4.3}$$

The above bound follows from the constraint $0 \leq X[0,t] \leq ht$ and the convexity of the exponential function in the effective bandwidth formula (4.1) [CKW97], and becomes accurate when the traffic rate takes two extreme values, a minimum of zero (''off'' phase) and a maximum of h (''on'' phase), and the duration of the ''on'' and ''off'' phase is large relative to the most probable buffer busy period prior to overflow.

4.2.2 Simple Bound

When a connection, in addition to having his peak rate policed, is policed by a leaky bucket (β,ρ), a tighter bound of the effective bandwidth is the following [CKW97]:

$$\widetilde{\alpha}_{\text{sb}}(m,\beta,\rho,h) = \frac{1}{st}\log\left[1 + \frac{tm}{\bar{X}[0,t]}\left(e^{s\bar{X}[0,t]} - 1\right)\right] , \tag{4.4}$$

where $\bar{X}[0,t] = \min\{ht, \beta+\rho t\}$, i.e., $\bar{X}[0,t]$ is the maximum amount of workload that the source can produce in a window of time t. The above bound follows from the constraint $0 \leq X[0,t] \leq \bar{X}[0,t]$ and the convexity of the exponential function in (4.1) [CKW97]. Observe that if $\min\{ht, \beta+\rho t\} = ht$, then equation (4.4), which we will refer to as the ''simple bound'', coincides with the on-off bound (4.3).

4.2.3 Inverted T Approximation

The third and more accurate approximation which we will investigate is motivated by investigations in [CKW97] which showed that for a given pair (s,t), in many cases, the worst case output of a leaky bucket consists of blocks of an inverted T shape which repeats periodically or with random gaps. Here, we consider the periodic inverted T pattern shown in Figure 4.3 [CKSW98, CS98]. Sources are assumed independent, hence their corresponding patterns have random phases, i.e., their start is randomly distributed in the time interval $2t + t_{\text{off}}$.

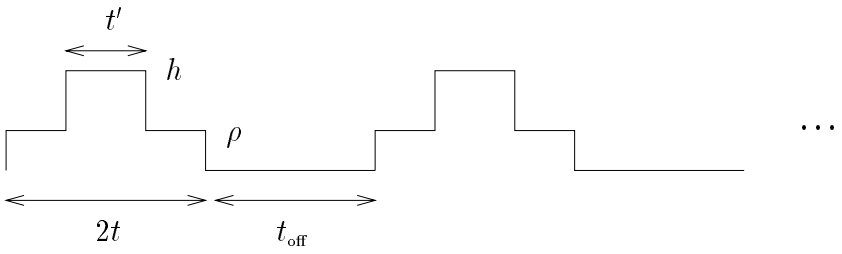


Figure 4.3: **Periodic pattern for the inverted $\top$ approximation.** $t' = \frac{\beta}{h-\rho}$, $t_{\text{off}} = \frac{(2t-t')\rho + t'h}{m} - 2t$

Based on the periodic pattern in Figure 4.3, we have the following effective bandwidth approximation:

$$\widetilde{\alpha}_{\perp}(m, \beta, \rho, h) = \frac{1}{st} \log E\left[e^{sX_{\perp}[0,t]}\right], \tag{4.5}$$

where $X_{\perp}[0,t]$ denotes the amount of workload produced by the inverted $\top$ pattern in an interval of length t. Note that $X_{\perp}[0,t]$ depends on the peak rate h and the leaky bucket (β, ρ); see Figure 4.3. (For simplicity, we have dropped this dependence in the notation.) The expected value in the right-hand side of (4.5) is computed analytically.

4.3 Pricing Schemes Involving Measurements in Distinct Time Intervals

In the previous section we discussed simple pricing schemes (*time-volume* pricing schemes) which involve only two measurements: the duration of the connection and the total volume transferred. In this section we describe two schemes that involve measurements in distinct time intervals, which are smaller than the duration of a connection: pricing with renegotiation and the virtual bucket scheme. Such approaches are interesting for two reasons. First, it is important to investigate how traffic parameter renegotiation affects pricing, since such a capability is being added to signaling protocols in B-ISDN [Q.2963], while the Resource Reservation Protocol (RSVP) [ZDE+93], which will be used in Internet's integrated services architecture, has inherent support for renegotiation. Second, it is important to investigate how the performance of such pricing schemes compares to that of the simple time-volume schemes and the ''optimal'' pricing scheme (the latter will be discussed in the next section).

The first scheme we investigate applies to connections that support traffic parameter renegotiation. The scheme is an application of the time-volume pricing schemes described in the previous section for measuring resource usage in each renegotiation interval, rather than for the whole

duration of the connection [CS98]. Note, however, that a straightforward without renegotiation extension of the simple schemes to include measurements of time and volume in intervals smaller than the duration of a connection would give the same prices as the time-volume schemes; this occurs because (4.2) is linear in T, V and the pair (a, b) is the same for all the intervals. It is the ability to renegotiate traffic contract parameters (hence have different parameters a, b in different intervals) that can lead to a better approximation. Such an approach will be advantageous when the connection has different ''modes'' of operation, such as movie scenes with high and low action. In this case, the leaky bucket produces a tighter bound when the traffic it characterizes has a single operating mode.

The second pricing scheme that we investigate, the *virtual bucket* scheme [CS98], requires measurements in a priori fixed time intervals. However, the same leak rate of a ''virtual'' leaky bucket is used to measure the corresponding bucket size in each time interval. Hence, this scheme does not entail the renegotiation overhead of the previous approach. Furthermore, the scheme uses only the peak rate from the connection's traffic descriptor, but requires the declaration of a leak rate at connection setup.

4.3.1 Pricing with Renegotiation

For services that support renegotiation, the schemes described in Section 4.2 can be used to measure resource usage for each renegotiation interval, where now the peak rate h_k and the leaky bucket parameters (β_k, ρ_k) are renegotiated in each interval k [CS98]. In what follows, we will use the inverted T approximation since it is more accurate than both the simple bound and the on-off bound (this will be shown in Section 4.5). Assume, for simplicity, that the renegotiation intervals have equal length T_{p}, and let V_k denote the volume transferred during interval k. If K is the total number of renegotiation intervals, then the average price per unit of time is[4]

$$\widetilde{\alpha}_{\mathrm{ren}}\left(\{m_k\},\{\beta_k\},\{\rho_k\},\{h_k\}\right) = \frac{\sum_{k=1}^{K} \widetilde{\alpha}_{\perp}(m_k,\beta_k,\rho_k,h_k)}{K}\,, \tag{4.6}$$

where $\{m_k\}, \{\beta_k\}, \{\rho_k\}$, and $\{h_k\}$ are sets containing the mean rate $m_k = V_k/T_{\mathrm{p}}$, the bucket size β_k, the leak rate ρ_k, and the peak rate h_k respectively, for all the renegotiation intervals $k = 1, \ldots, K$.

[4]Equations (4.6) and (4.7) assume that the duration of a connection is an integer multiple of the renegotiation or pricing interval. They can be extended to cover the general case.

4.3.2 The Virtual Bucket Scheme

Unlike the previous scheme, the *virtual bucket* scheme [CS98] does not require renegotiation of traffic parameters. Let r be the leak rate of a ''virtual'' leaky bucket and h the peak rate, both declared at connection setup, e.g., indirectly through the tariff selection. In fixed, non-overlapping time intervals of length T_{p} (pricing interval), two measurement are made for each connection: the volume V_k and the minimum bucket size b_k such that the connection's traffic in interval k conforms to the leaky bucket (b_k, r). The quantity b_k can be measured using two counters: B_1 and B_2. Counter B_1 decreases at rate r (measured in cells per second), and increases by one for every cell received. At each update of B_1, its value is compared to B_2. If $B_1 > B_2$, then B_2 is set to B_1. At the end of each pricing interval, the value of B_2 will be the bucket size b_k we are looking for. If K is the total number of pricing intervals, the average price per unit of time is

$$\widetilde{\alpha}_{\mathrm{vb}}\left(\{m_k\},\{b_k\},r,h\right) = \frac{\sum_{k=1}^{K}\widetilde{\alpha}_{\perp}(m_k,b_k,r,h)}{K}\,, \tag{4.7}$$

where $\{m_k\}$ and $\{b_k\}$ are sets containing the mean rate $m_k = V_k/T_{\mathrm{p}}$ and the bucket size b_k for all the renegotiation intervals $k = 1,\ldots,K$.

4.4 Fairness of Pricing Schemes

In this section we discuss the fairness property of pricing schemes and present the specific metrics according to which the pricing schemes discussed in the previous two sections will be compared.

As discussed in the introduction, the fairness property of a pricing scheme is its ability to capture the *relative* amount of resources used by connections. Suppose $\widetilde{\alpha}$ is an approximation of the bound of the effective bandwidth α and let, with a slight abuse of notation, $\alpha(x)$ and $\widetilde{\alpha}(x)$ be the corresponding charges for some connection x. A pricing scheme which is based on $\widetilde{\alpha}$ is fair if for any two connections x and y, we have $\widetilde{\alpha}(y)/\widetilde{\alpha}(x) \approx \alpha(y)/\alpha(x)$. As a measure of the *unfairness* of a pricing scheme over a set of connections we take the following [CKSW98, CS98]:

$$\text{Unfairness} := \frac{\sigma}{\mu}\,,$$

where σ is the standard deviation of $\widetilde{\alpha}(x)/\alpha(x)$ and μ is the average of $\widetilde{\alpha}(x)/\alpha(x)$, as x ranges over the connection set. Note that we expect $\mu \geq 1$ since we consider upper bounds (and approximations

of upper bounds) of the effective bandwidth. A small value for the unfairness indicator means that pricing schemes based on approximation $\widetilde{\alpha}$ will tend to provide users with the right incentives, i.e., a user who chooses a tariff which results in a smaller price is actually making less use of network resources.

The above definition refers to unfairness among connections with the *same type* of traffic. However, it is also important that a pricing scheme is fair among connections with a *different type* of traffic. For example, consider a pricing scheme which is based on some approximation of the effective bandwidth $\widetilde{\alpha}$, and two types of traffic j_1, j_2. Assume that the pricing scheme is equally fair for both types of traffic, i.e., it has the same unfairness indicator for both types of traffic. If the pricing scheme has a smaller average value μ of the ratio $\widetilde{\alpha}(x)/\alpha(x)$ for connections of type j_1 than for connections of type j_2, then it might happen that two connections of type j_1 and j_2 have the same price, although on the average the actual amount of resources used by the connection of type j_2 is smaller than the amount used by the connection of type j_1. Hence, the scheme will be unfair for connections of type j_2 compared to connections of type j_1.

In summary, the fairness of the pricing schemes will be investigated using two metrics:

- *Fairness over the same type of traffic:* We will use the unfairness index σ/μ, where σ is the standard deviation of $\widetilde{\alpha}(x)/\alpha(x)$ and μ is the average of $\widetilde{\alpha}(x)/\alpha(x)$ as x ranges over the connection set. For the same connection set, if the unfairness index of pricing scheme p_1 is smaller than the unfairness index of pricing scheme p_2, then pricing scheme p_1 is fairer than pricing scheme p_2 for that connection set.

- *Fairness over different types of traffic:* We will use the average ratio μ of the ratio $\widetilde{\alpha}(x)/\alpha(x)$ as x ranges over the set of connections carrying the same type of traffic. For two connection sets with different types of traffic, if pricing scheme p_1 has the same average ratio μ, while pricing scheme p_2 has a different average ratio, then pricing scheme p_1 is fairer than pricing scheme p_2 for connections belonging to the two sets.

We conclude this section with a discussion of how the value of the effective bandwidth $\alpha(x)$ that appears in the ratio $\widetilde{\alpha}(x)/\alpha(x)$ should be computed. First, note that we wish to compare the pricing schemes with the theoretically ‘‘optimal’’ pricing scheme. Such an optimal pricing scheme would directly use the effective bandwidth formula (4.1). But, this formula can be used to measure

resource usage in any time interval, and not necessarily for the whole duration of a connection. What should this time interval, say T_{x}, be? At one extreme T_{x}, can be very small, in which case effective bandwidth equals the connection's mean rate. However, measuring resource usage in such a small interval will not be advantageous to the network, if the latter cannot control resource usage in such an interval. This suggests that T_{x} must be the time interval that the network makes resource allocation decisions. The latter cannot be smaller than the connection inter-arrival/inter-departure time.

Hence, the ''optimal'' pricing scheme would measure resource usage in a time interval T_{x} equal to the connection inter-arrival/inter-departure time. Due to the lack of such information, we instead assume that T_{x} is equal to the pricing/renegotiation interval T_{p} of the virtual bucket and pricing with renegotiation schemes. Hence, the value of $\alpha(x)$ is computed by taking the average of the effective bandwidth estimates $\alpha_k(x)$ for each time interval k, whose length is equal to T_{p} (the duration of the pricing/renegotiation interval):

$$\alpha(x) = \frac{\sum_{k=1}^{K} \alpha_k(x)}{K},$$

where K is the total number of pricing intervals, i.e., $K = T/T_{\mathrm{p}}$ with T the total duration of the connection. The values of $\alpha_k(x)$ are computed using the effective bandwidth formula (4.1), where the expectation is replaced by the empirical mean, i.e.,

$$\alpha_k(x) = \frac{1}{st} \log \left[\frac{1}{T_{\mathrm{p}}/t} \sum_{i=1}^{T_{\mathrm{p}}/t} e^{sX[(k-1)T_{\mathrm{p}}+(i-1)t,(k-1)T_{\mathrm{p}}+it]} \right],$$

where $X[t_1, t_2]$ is the amount of workload connection x produces in the time interval $[t_1, t_2]$.

4.5 Experiments with Real Traffic

In this section we evaluate the time-volume pricing schemes (Sections 4.2) and the pricing schemes with measurements in distinct time intervals (Section 4.3), based on the unfairness criteria discussed in the previous section. Our investigations involve broadband traffic which consists of long MPEG-1 compressed video and Internet WAN traffic [CS98, CKSW98]. Details of the traffic sources can be found in Appendix A. The specific issues we investigate are the following:

- Effect of the pricing/renegotiation interval on the performance of the pricing with renegotiation and virtual bucket schemes

- Effect of the connection (or call) duration on the fairness of the pricing schemes
- Unfairness of the pricing schemes for various link capacities and buffer sizes
- Robustness for small variations of the link's operating point parameters s, t
- Unfairness of the pricing schemes for various traffic contents
- Effect of traffic smoothing on the unfairness of the pricing schemes

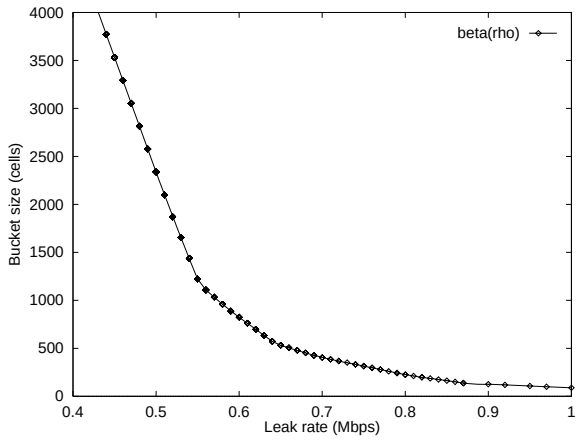
We assume that users have full knowledge of their traffic and seek to minimize their charge. The rate at which such users will be charged, as discussed in Section 4.2, is the value of the effective bandwidth bound used by the pricing scheme. Therefore, the comparison of the pricing schemes reduces to the comparison of the effective bandwidth bounds on which they are based.

Furthermore, since users seek to minimize their charge, they select the leaky bucket parameters (or only the leak rate for the virtual bucket scheme) which yield the smallest effective bandwidth bound. This is illustrated in Figure 4.4 which corresponds to a user whose traffic is a 3 minute segment from the *Star Wars* sequence (Appendix A.1). The user selects the leaky bucket parameters (β, ρ) which minimize his charge (Figure 4.4(b)), among the pairs (β, ρ) to which his traffic conforms (Figure 4.4(a)). Since we assume users have full knowledge of their traffic, the (β, ρ) pairs to which a user's traffic conforms are computed beforehand for the whole sequence of his traffic. Note that as the duration of the sequence increases, the (β, ρ) curve in Figure 4.4(a) moves towards the top-right corner.

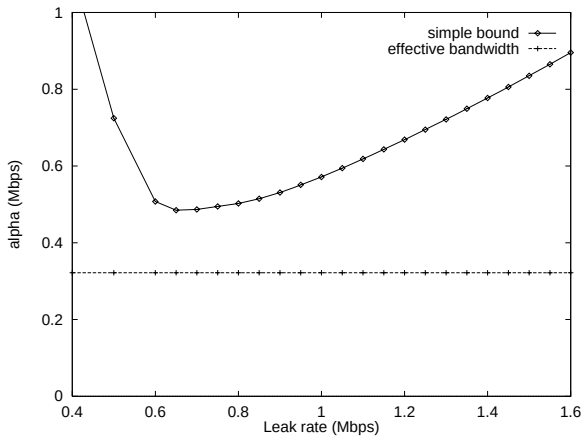
Our investigations involve link capacities and buffer sizes that will be encountered in practice.[5] In particular, the link capacities considered are 34 Mbps (E-3 physical interface), 155 Mbps (STM-1 Synchronous Digital Hierarchy - SDH - physical interface), and 622 Mbps (STM-4 SDH physical interface). For the investigations involving MPEG-1 video, we consider buffer sizes in the range $1 - 16$ msec. These fall within the expected values of switch buffers that are to support interactive real time transmission. For the investigations involving Internet traffic, we consider buffer sizes up to 400 msec, since such traffic has less stringent delay requirements.

All our investigations assume ''typical'' values of the parameters s, t which correspond to the link capacity, buffer size, and traffic content of each case. In practice, these values will be computed

[5]For example, FORE System's ASX-200BX and ASX-1000 can support up to 13312 cells or 166 msec for a single 34 Mbps E-3, and up to 32768 cells or 93 msec for a single 155 Mbps STM-1 interface.



(a) Leak rate/bucket size tradeoff



(b) Simple bound approximation for different (β, ρ) pairs that lie on graph (a)

Figure 4.4: **Leaky bucket selection for a rational user.** A rational user selects the leaky bucket parameters (β, ρ) which minimize his charge (right graph), among the pairs (β, ρ) to which his traffic conforms (left graph). [$C = 155$ Mbps, $B = 16$ msec (≈ 5622 cells), *Star Wars* segment with duration ≈ 3 minutes]

off-line using the `supinf` formula (2.3) (page 23) and the effective bandwidth formula (2.2) (see Appendix B.1). For MPEG-1 traffic, we consider a target overflow probability of 10^{-7}, while for the Internet traffic the overflow probability is 10^{-6}.

In order to make the section easier to read, all the figures have been placed at the end of the section (page 85).

4.5.1 MPEG-1 Traffic

In this section we evaluate the pricing schemes for MPEG-1 compressed video traffic (see Appendix A.1 for details of the traffic used).

Effect of the pricing/renegotiation interval

Figure 4.5 shows, for *Star Wars* traffic, the effect of the pricing/renegotiation interval on the performance of the pricing with renegotiation and virtual bucket schemes. As expected, the accuracy of these two schemes is much better than both the simple bound and the inverted `T` approximation. Furthermore, the accuracy of the virtual bucket scheme is close to that of the pricing with renegotiation scheme.

In the case of the time-volume pricing schemes, observe that the unfairness decreases when the pricing interval increases. This occurs because the mean rate fluctuates more when it is measured over smaller intervals. The latter results in a larger variance of the ratio $\widetilde{\alpha}/\alpha$, where $\widetilde{\alpha}$ is an approximation of the effective bandwidth α. Finally, note that the on-off bound (not shown in Figure 4.5) coincides with the simple bound. This occurs because, for the link capacity (155 Mbps) and buffer size (8 msec $\approx$ 2811 cells) considered, the time parameter t is such that $\min\{ht, \beta + \rho t\} = ht$, hence the on-off bound (4.3) becomes identical to the simple bound (4.4).

Figure 4.6 shows the performance of the pricing schemes for *Aliki in the Navy* traffic. Observe that the accuracy of the time-volume pricing schemes is much better than it was for the *Star Wars* sequence. This is because the *Aliki in the Navy* sequence is less bursty than the *Star Wars* sequence (it has a smaller peak rate to mean rate ratio, see Table A.2). Hence, the performance of the time-volume pricing schemes increases for less bursty traffic. On the other hand, the performance of the virtual bucket and pricing with renegotiation schemes is not affected much by the burstiness of traffic.

Effect of the call duration on fairness

Figure 4.7 shows that the average ratio μ for the simple bound and the inverted $\mathtt{T}$ approximation increases when the duration of the connection (call) increases. Hence, these schemes are less fair to long duration connections. This is expected because the leaky bucket description is less tight when the duration of a connection is large, in which case it includes a large number of scenes, compared to when the duration is small, in which case it includes a smaller number of scenes.

On the other hand, because both the pricing with renegotiation and virtual bucket schemes involve measurements in fixed time intervals, the average ratio μ is not affected by the call duration.

In the remaining investigations with MPEG-1 video, the traffic used consists of three sets: *movies*, *news*, and *talk shows* (see Appendix A.1 for details). The duration of the segments in each of these sets is approximately 3 minutes. Since the encoding parameters for all three types of traffic are the same, any differences among the three sets are due solely to the traffic content.

Unfairness for various link capacities and buffer sizes

Figure 4.8 shows the unfairness of the pricing schemes for various link capacities ($34, 155$, and 622 Mbps) and buffer sizes ($1, 4, 8,$ and 16 msec). The comparison also includes the unfairness of the

pricing with renegotiation and virtual bucket schemes for different pricing/renegotiation intervals (10, 20, and 30 seconds).

First, observe that even for pricing/renegotiation interval $T_{\mathrm{p}} = 30$ seconds, the unfairness of the pricing schemes with measurement in distinct intervals is much better than the time-volume pricing schemes. However, the difference decreases as we move to larger link capacities.

Second, observe that for small buffer sizes, the on-off bound coincides with the simple bound. In particular, the unfairness of the two bounds is the same for buffer sizes up to 16 msec when $C = 34$ Mbps, up to 8 msec when $C = 155$ Mbps, and up to 4 msec when $C = 622$ Mbps. This happens because, for these values of buffer sizes, the time parameter t is such that $\min\{ht, \beta + \rho t\} = ht$, hence the on-off bound (4.3) becomes identical to the simple bound (4.4).

Third, observe that the unfairness of the pricing schemes with measurements in distinct intervals is less affected by the buffer size. This is not the case for the time-volume pricing schemes.

Finally, observe that as the link capacity increases, the unfairness of all the pricing schemes increases. This occurs because as the link capacity increases, statistical multiplexing becomes more efficient and the effective bandwidth approaches the mean rate.

Robustness for small variations of parameters s, t

Next we investigate the unfairness of the pricing schemes for small variation of the link's operating point parameters s, t. Such variations can be due, e.g., to small changes of the traffic mix. Figure 4.9 shows how robust the schemes are for values of s, t in a range that is typical for link capacity 155 Mbps and buffer 8 msec ($\approx$ 2811 cells). Observe that the two pricing schemes with measurements in distinct time intervals are more robust than the time-volume pricing schemes (simple bound and inverted $\mathtt{T}$ approximation). Comparing Figure 4.9 with Figure 4.10, which shows the robustness for $C = 155$ Mbps and $B = 16$ msec ($\approx$ 5622 cells), we see that the robustness of the pricing schemes increases when the buffer size increases.

Effect of the traffic content

Figure 4.11 shows that the unfairness of the pricing schemes with measurements in distinct time intervals is smaller and less affected by the traffic content, compared to the time-volume pricing schemes.

Now we concentrate on the time-volume pricing schemes. Figure 4.11 shows that the unfairness

is higher for action movies, and lower for talk shows. For news, the unfairness is somewhere in the middle. Figure 4.12 shows the average of the ratio of the effective bandwidth approximation over the actual effective bandwidth. Based on these observations we can make the following remarks on how the video content affects the performance of the time-volume pricing schemes, or equivalently the effective bandwidth approximations on which they are based.

Talk shows: For talk shows, the average ratio of the time-volume schemes is the smallest (Figure 4.12). Furthermore, for talk shows the unfairness is also very small (Figure 4.11). Hence, the time-volume schemes are *uniformly accurate* for talk shows.

Movies: For movies, the average ratio of the time-volume schemes is almost as small as that for talk shows (Figure 4.12). However, the unfairness is highest for movies (Figure 4.11). Hence, for some movie segments the time-volume schemes are accurate, while for others they are very inaccurate.

News: For news, the average ratio is the largest among the three types of traffic (Figure 4.12). However, the unfairness index for news is smaller than that for movies (Figure 4.11). Hence, the time-volume schemes are *uniformly inaccurate* for talk shows.

Of course the above observations refer to the specific video traces, and a generalization would require more detailed comparisons with traces encoded using different parameters.

Effect of traffic smoothing on unfairness

In Figure 4.13, the traces are smoothed such that cells produced during two consecutive frames are evenly spaced throughout the duration of the two frames ($= 80$ msec). Comparing Figure 4.13 with Figure 4.11, we observe that traffic smoothing has a noticeable effect on the unfairness of the time-volume pricing schemes only up to a certain buffer size, whose value depends on the link capacity. In particular, traffic smoothing has a noticeable effect on unfairness for a buffer size up to 8 msec ($\approx$ 2811 cells) for capacity 155 Mbps, and up to 4 msec ($\approx$ 5660 cells) for capacity 622 Mbps. For capacity 34 Mbps, smoothing affects all the buffer sizes shown, i.e., up to 16 msec ($\approx$ 1280 cells). The effect of smoothing is related to the time scale (expressed through parameter t) that is important for buffer overflow (see also Sections 2.2 and 2.3). For large buffer sizes, t is large (i.e., slow time scales are important for overflow), and smoothing traffic in a smaller time interval has no effect on the effective bandwidth.

4.5.2 Internet WAN Traffic

In this section we evaluate the pricing schemes for Internet Wide Area Network (WAN) traffic. The traffic considered was a set of 15 trace segments, each with a duration of approximately two hours, that was created from a larger trace (see Appendix A.2 for details).

Effect of the pricing/renegotiation interval

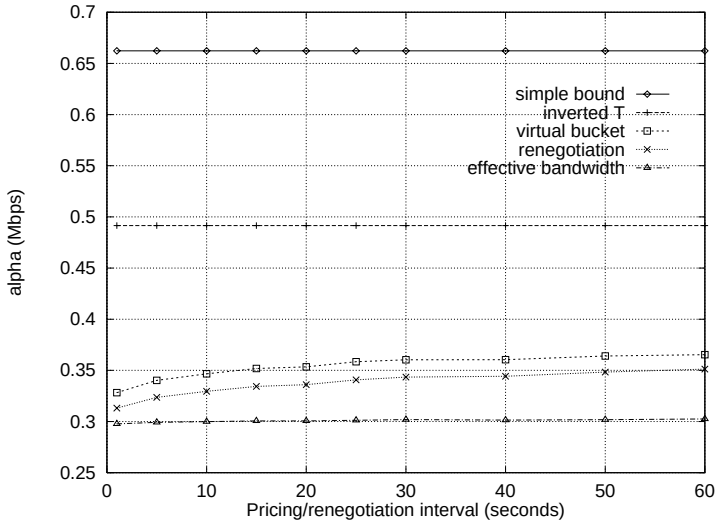
Figure 4.14 shows the effect of the pricing/renegotiation interval on the performance of the virtual bucket and pricing with renegotiation schemes. The graph refers to a 2 hour trace of Internet WAN traffic. As was the case for MPEG-1 traffic, the accuracy of the pricing with renegotiation and virtual bucket schemes is much better than the time-volume pricing schemes, and the performance of the virtual bucket scheme is close to that of the pricing with renegotiation scheme.

Unfairness for various link capacities and buffer sizes

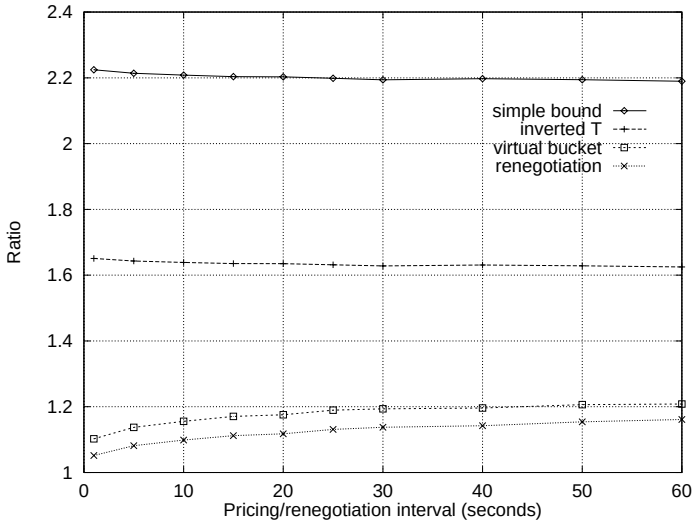
Figure 4.8 shows the unfairness of the pricing schemes for capacity 34 Mbps and 155 Mbps. Our conclusions are similar to the ones for MPEG-1 video: The unfairness of the two schemes involving measurements in distinct time intervals is better than the unfairness of the time-volume pricing schemes. However, the difference decreases as the capacity and buffer size increases.

Robustness for small variations of parameters s, t

Figure 4.16 shows the robustness of the pricing schemes for small variations of parameters s, t, when these fall within the range which is typical for link capacity 34 Mbps and buffer 50 msec. As was the case of MPEG-1 video, the robustness of the two schemes with measurements in distinct time intervals is better than the time-volume pricing schemes. Furthermore, the robustness for all schemes increases for larger buffer sizes (compare Figure 4.16 with Figure 4.17).

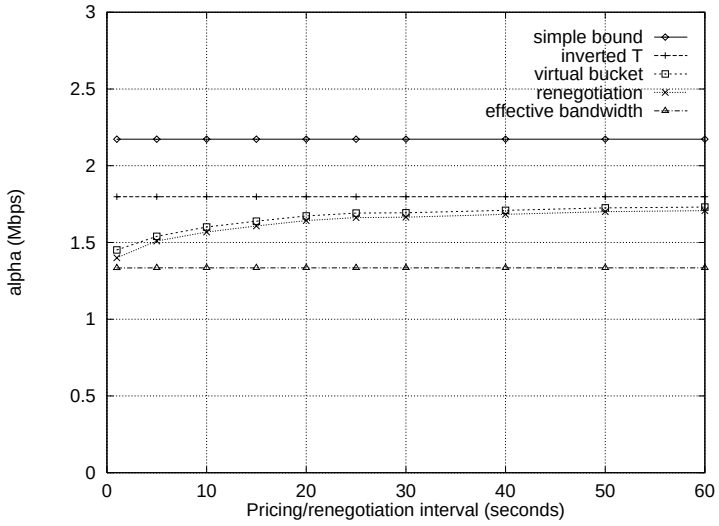


(a) Effective bandwidth

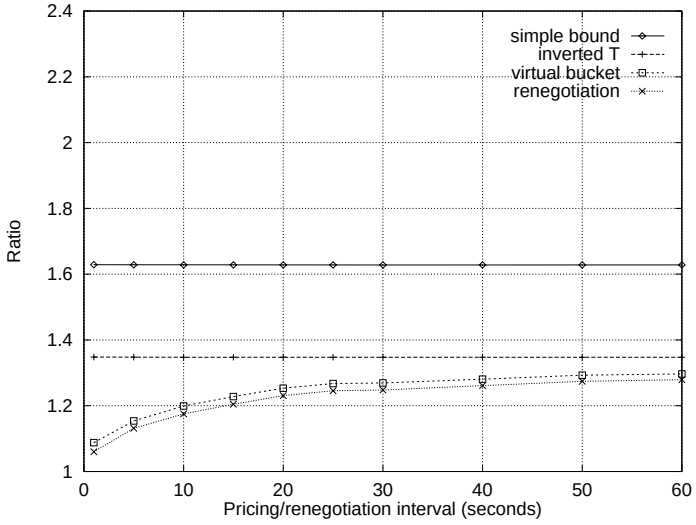


(b) Ratio $\widetilde{\alpha}/\alpha$

Figure 4.5: **Effect of the pricing/renegotiation interval (Star Wars traffic).** The pricing schemes with measurements in distinct time intervals (virtual bucket and renegotiation) are more accurate then the time-volume pricing schemes (simple bound and inverted $\mathtt{T}$), i.e., they have a smaller average μ of the ratio $\widetilde{\alpha}/\alpha$ (where $\widetilde{\alpha}$ is an approximation of the bound of the effective bandwidth α). Also, the performance of the virtual bucket scheme is close to that of the pricing with renegotiation scheme. [$C = 155$ Mbps, $B = 8$ msec (≈ 2811 cells), *Star Wars* (duration ≈ 30 minutes)]



(a) Effective bandwidth



(b) Ratio $\widetilde{\alpha}/\alpha$

Figure 4.6: **Effect of the pricing/renegotiation interval (Aliki in the Navy traffic).** The pricing schemes with measurements in distinct time intervals (virtual bucket and renegotiation) are more accurate then the time-volume pricing schemes (simple bound and inverted T). However, the difference is much smaller than in Figure 4.5. This is because the *Aliki in the Navy* stream is less bursty than the *Star Wars* stream (see Table A.2, page 137). Also, observe that the performance of the virtual bucket scheme is close to that of the pricing with renegotiation scheme. [$C = 155$ Mbps, $B = 8$ msec (≈ 2811 cells), *Aliki in the Navy* (duration ≈ 30 minutes)]

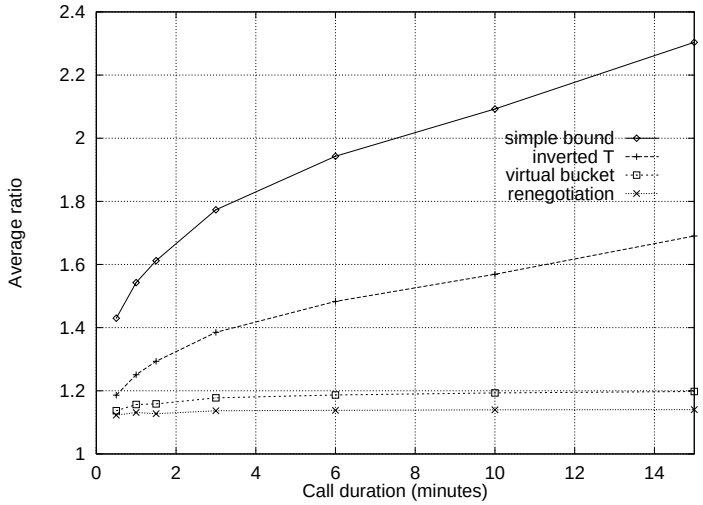
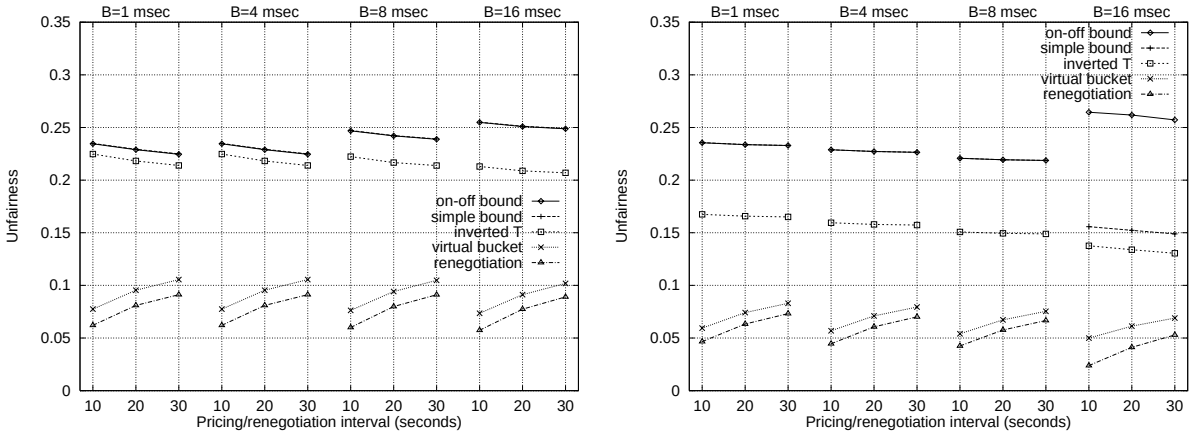
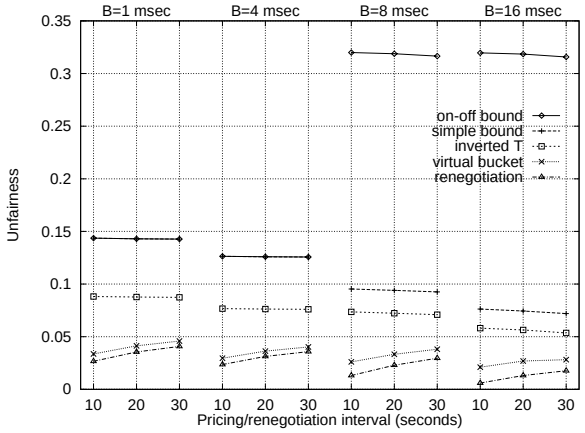


Figure 4.7: **Effect of the call duration on the accuracy of the pricing schemes.** The horizontal axis depicts the average μ of the ratio $\widetilde{\alpha}/\alpha$ (where $\widetilde{\alpha}$ is an approximation of the bound of the effective bandwidth α). The accuracy of the simple pricing schemes (simple bound and inverted $\mathtt{T}$) decreases when the call duration increases; hence, they are unfair to long duration calls compared to short duration ones. On the other hand, the pricing schemes that involve measurements in distinct time intervals (virtual bucket and renegotiation) are not affected by the call duration. [$C = 155$ Mbps, $B = 8$ msec (≈ 2811 cells), $T_{\mathrm{p}} = 10$ seconds, MPEG-1 movie traffic]



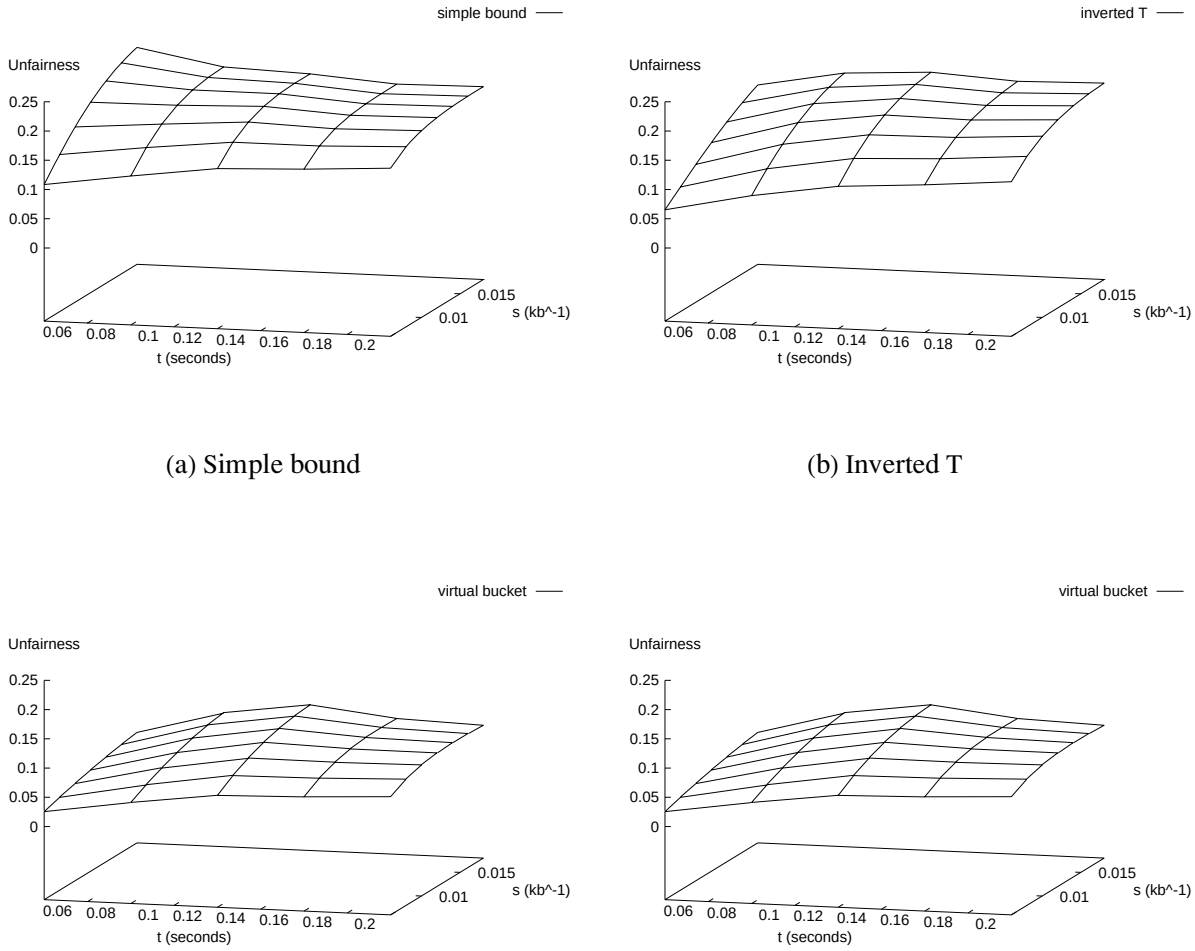
(a) $C = 34$ Mbps

(b) $C = 155$ Mbps



(c) $C = 622$ Mbps

Figure 4.8: **Unfairness of the pricing schemes for various link capacities and buffer sizes.** The unfairness of the pricing schemes with measurements in distinct time intervals (virtual bucket and renegotiation) is smaller compared to the unfairness of the time-volume pricing schemes (on-off bound, simple bound, and inverted T). However, the difference decreases when the capacity increases. Furthermore, the link capacity and buffer size has a greater effect on the time-volume pricing schemes than on the pricing schemes with measurements in distinct time intervals. [MPEG-1 movie set (54 segments each with duration $\approx$ 3 minutes)]



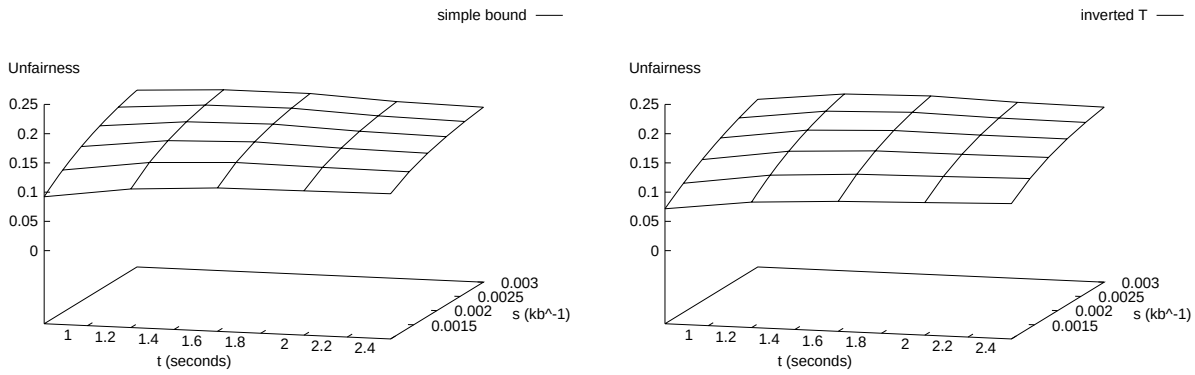
(a) Simple bound

(b) Inverted T

(c) Virtual bucket

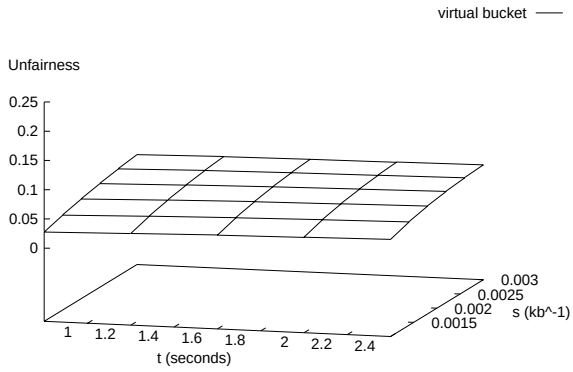
(d) Renegotiation

Figure 4.9: **Robustness of the pricing schemes for values of s, t that are typical for $C = 155$ Mbps and $B = 8$ msec.** The pricing schemes with measurements in distinct time intervals are more robust (their surface is more ''flat'') compared to the time-volume pricing schemes. [$T_{\rm p} = 10$ seconds, MPEG-1 movie set (54 segments each with duration $\approx$ 3 minutes)]

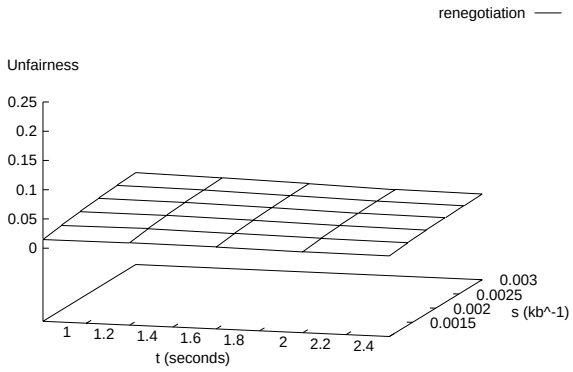


(a) Simple bound

(b) Inverted T

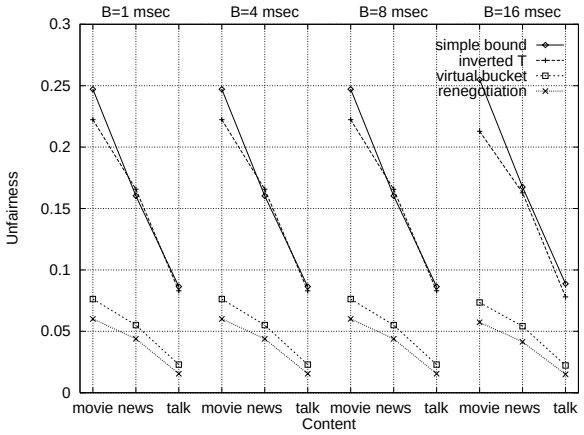


(c) Virtual bucket

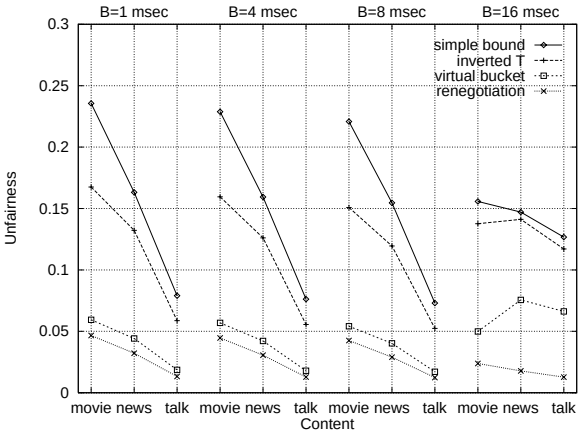


(d) Renegotiation

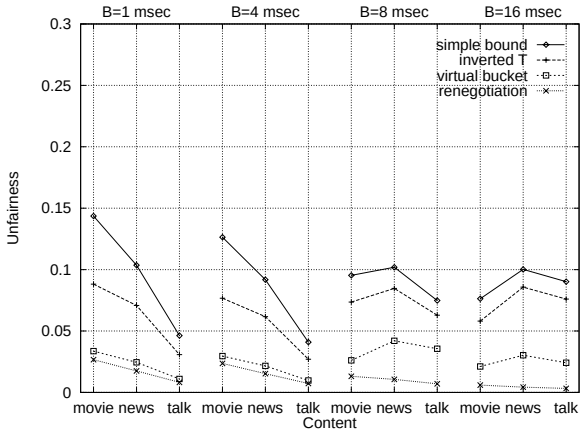
Figure 4.10: **Robustness of the pricing schemes for values of s, t that are typical for $C = 155$ Mbps and $B = 16$ msec.** The robustness for all the pricing schemes has increased compared to the robustness for $B = 4$ msec (previous figure). However, the pricing schemes with measurements in distinct time intervals (virtual bucket and renegotiation) are still more robust (their surface is more ''flat'') compared to the time-volume pricing schemes. [$T_{\mathrm{p}} = 10$ seconds, MPEG-1 movie set (54 segments each with duration ≈ 3 minutes)]



(a) $C = 34$ Mbps

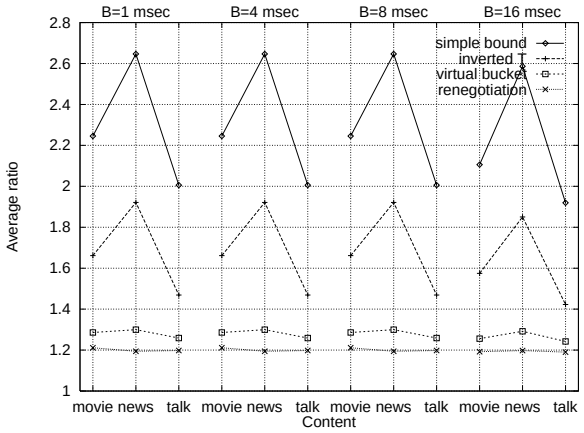


(b) $C = 155$ Mbps

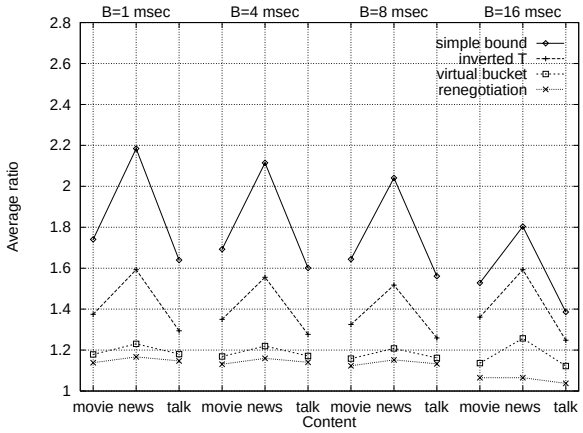


(c) $C = 622$ Mbps

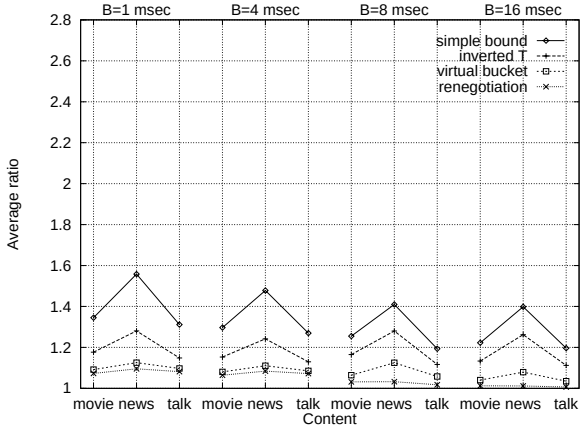
Figure 4.11: **Unfairness of the pricing schemes for various video contents.** The unfairness of the pricing schemes with measurements in distinct time intervals is smaller and less affected by the video content compared to the unfairness of the time-volume pricing schemes. However, the difference is smaller for larger link capacities. [$T_{\mathrm{p}} = 10$ seconds, MPEG-1 video traffic (each segment has duration ≈ 3 minutes)]



(a) $C = 34$ Mbps

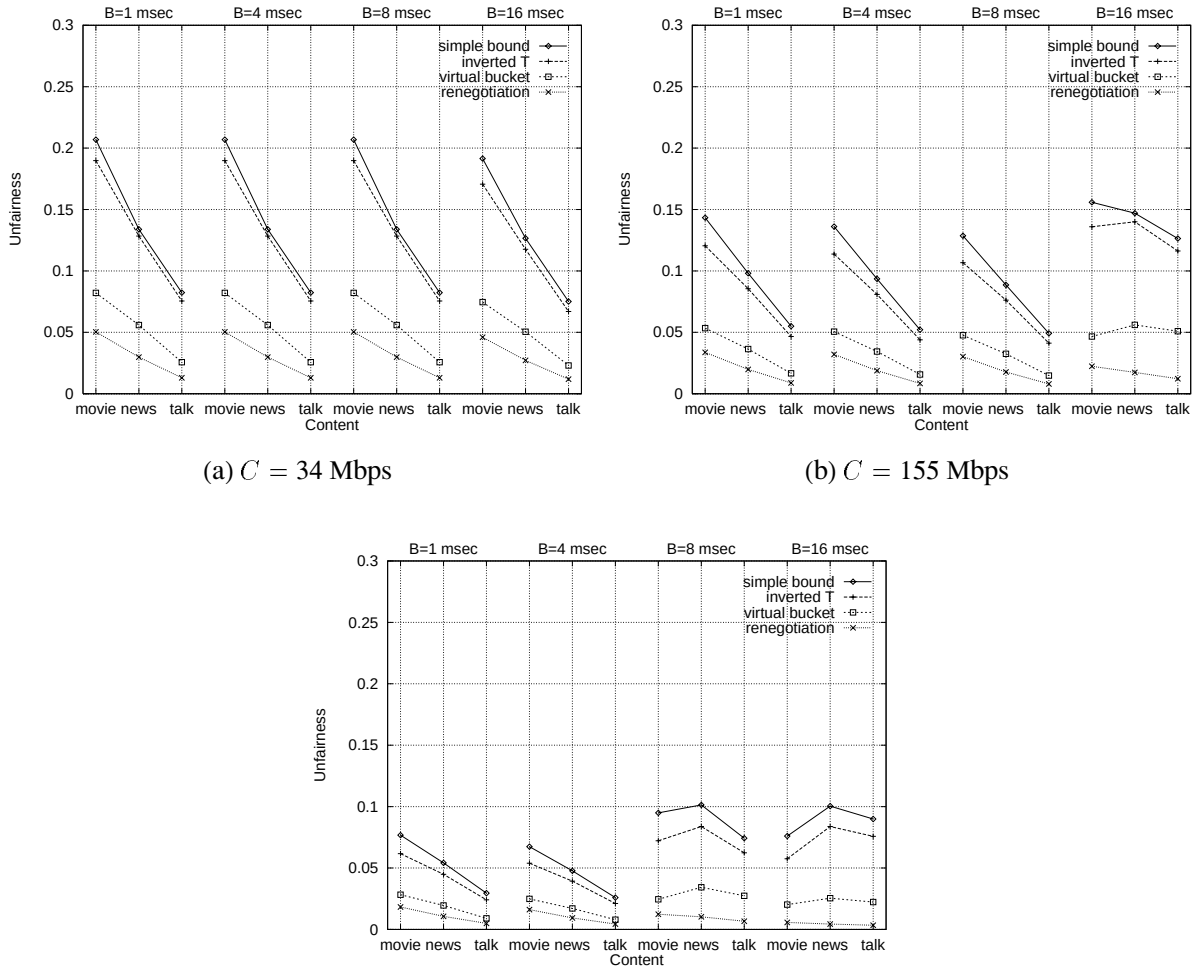


(b) $C = 155$ Mbps



(c) $C = 622$ Mbps

Figure 4.12: **Average ratio of the pricing schemes for various video contents.** The average μ of the ratio $\widetilde{\alpha}/\alpha$ (where $\widetilde{\alpha}$ is an approximation of the effective bandwidth α) of the pricing schemes with measurements in distinct time intervals is smaller and less affected by the video content, compared to the time-volume pricing schemes. However, the difference is smaller for larger link capacities. [$T_{\mathrm{p}} = 10$ seconds, MPEG-1 video traffic (each segment has duration ≈ 3 minutes)]

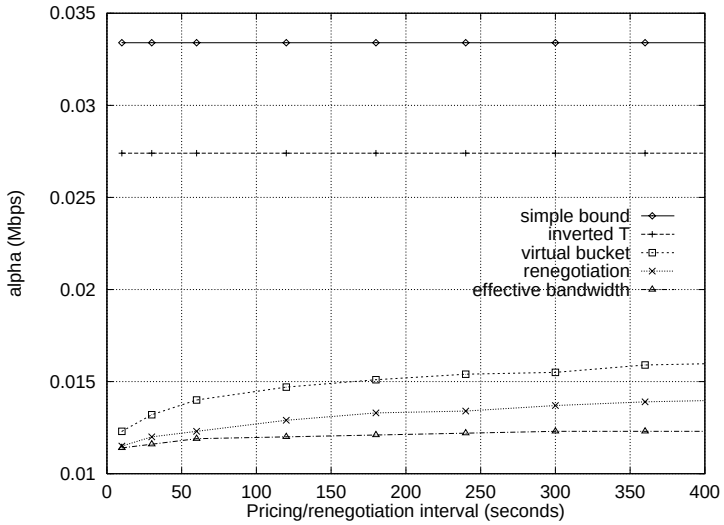


(a) $C = 34$ Mbps

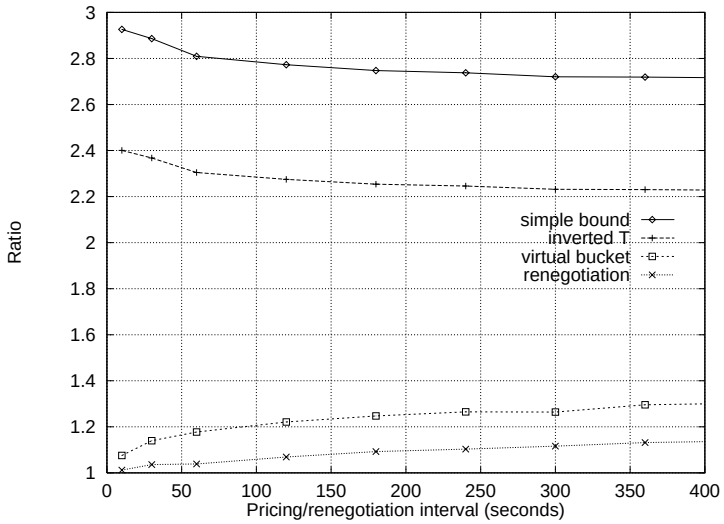
(b) $C = 155$ Mbps

(c) $C = 622$ Mbps

Figure 4.13: **Unfairness of the pricing schemes for smoothed traffic with various contents.** Comparison with Figure 4.11 shows that traffic smoothing decreases the unfairness of the time-volume pricing scheme up to a certain buffer size, which depends on the link capacity. The buffer sizes for which smoothing has a very small effect are such that the time parameter t is larger than the smoothing interval (which is equal to two frame times or 80 msec). [$T_{\mathrm{p}} = 10$ seconds, MPEG-1 video traffic (each segment has duration ≈ 3 minutes)]

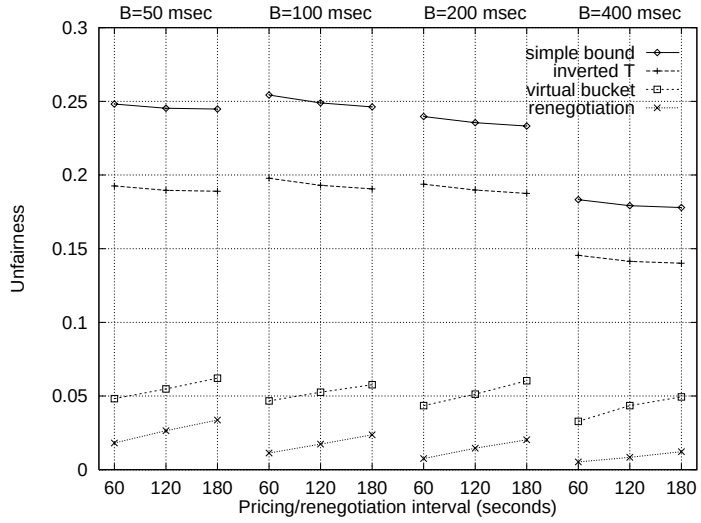


(a) Effective bandwidth

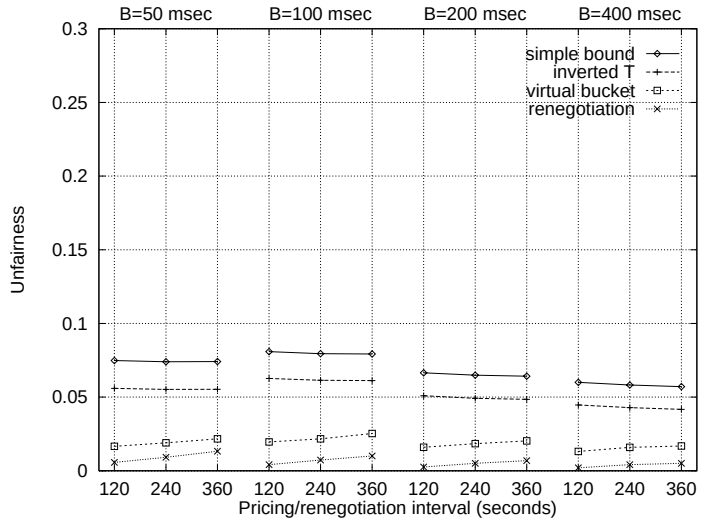


(b) Ratio $\widetilde{\alpha}/\alpha$

Figure 4.14: **Effect of the pricing/renegotiation interval (Internet WAN traffic).** The pricing schemes with measurements in distinct time intervals are more accurate than the time-volume pricing schemes, i.e., they have a smaller average μ of the ratio $\widetilde{\alpha}/\alpha$ (where $\widetilde{\alpha}$ is an approximation of the effective bandwidth α). Also, the performance of the virtual bucket scheme is close to that of the pricing with renegotiation scheme. [$C = 34$ Mbps, $B = 100$ msec, Internet WAN traffic (duration $\approx$ 2 hours)]

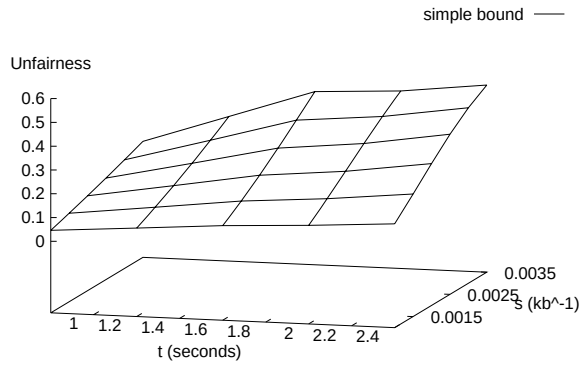


(a) $C = 34$ Mbps

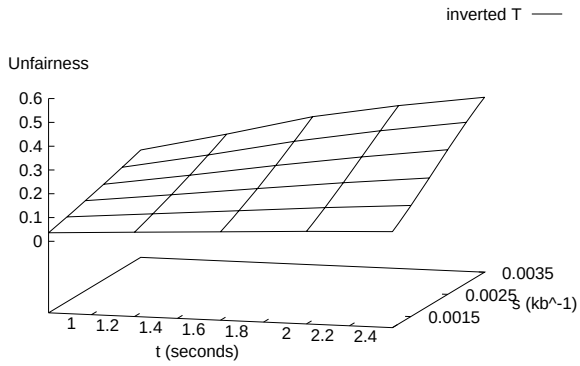


(b) $C = 155$ Mbps

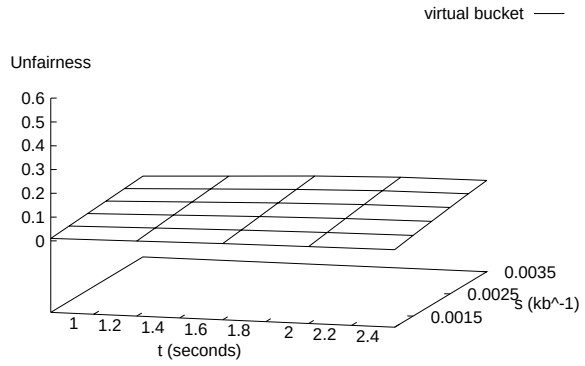
Figure 4.15: **Unfairness of the pricing schemes for various link capacities and buffer sizes**. Similar to the results for MPEG-1 traffic, the pricing schemes with measurements in distinct time intervals have smaller unfairness, which is less affected by the link capacity and buffer size compared to the time-volume schemes. The difference decreases as the link capacity and buffer size increases. [Internet WAN traffic set (15 segments each with duration $\approx$ 2 hours)]



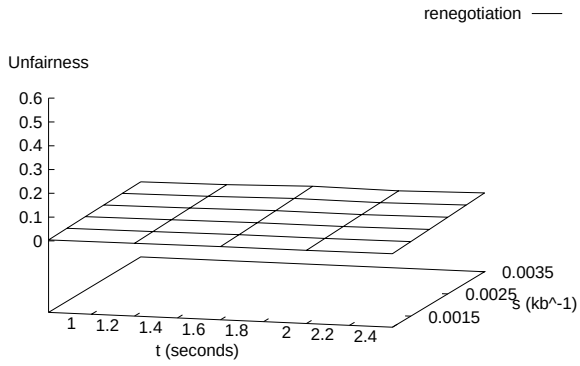
(a) Simple bound



(b) Inverted T

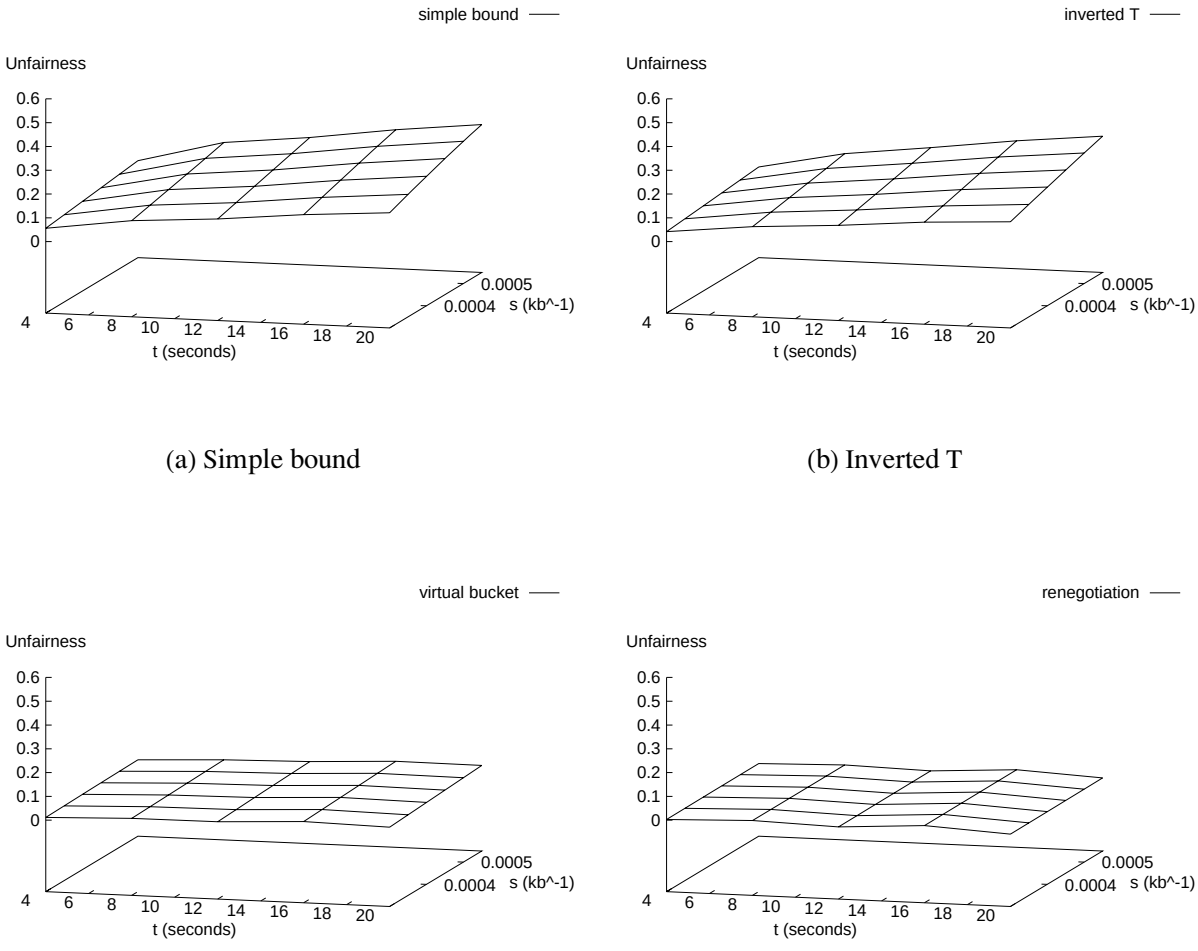


(c) Virtual bucket



(d) Renegotiation

Figure 4.16: **Robustness of the pricing schemes for values of s, t that are typical for $C = 34$ Mbps and $B = 50$ msec.** Similar to the results for MPEG-1 traffic, the two pricing schemes with measurements in distinct time intervals are more robust (their surface is more ''flat'') than the time-volume pricing schemes. [$T_{\mathrm{p}} = 60$ seconds, Internet WAN traffic set (15 segments each with duration ≈ 2 hours)]



(a) Simple bound

(b) Inverted T

(c) Virtual bucket

(d) Renegotiation

Figure 4.17: **Robustness of the pricing schemes for values of s, t that are typical for $C = 34$ Mbps and $B = 200$ msec.** The robustness of all pricing schemes is better than the robustness for $B = 50$ msec (previous figure). However, the pricing schemes with measurements in distinct time intervals are still more robust (their surface is more ‘‘flat’’) compared to the time-volume pricing schemes. [$T_{\rm p} = 60$ seconds, Internet WAN traffic set (15 segments each with duration $\approx$ 2 hours)]

4.6 Incentive Compatibility

As we have already mentioned in Section 4.1, the operating point of the link and the posted tariffs are interrelated in a circular fashion. The network operator posts tariffs that have been computed for the current operating point of the link, expressed through the parameters s, t. These tariffs provide *incentives* for the users to change their contracts in order to minimize their anticipated charges. Under these new contracts, the operating point of the system will move, since the network operator must guarantee the performance requirements of these new contracts. Hence, the network operator will calculate new tariffs for the new operating point. This interaction between the network and the users will continue until an equilibrium is reached. We validate below, for a simple example, that if the network operator uses the charging scheme based on the simple bound, then an equilibrium does exist and at the equilibrium the number of users admitted to the system is maximized [CKSW98].

For simplicity, we assume that all users have identical profiles, are policed by a single leaky bucket, and have an identical leak rate/bucket size tradeoff expressed by the function $G = \beta(\rho)$. We assume that G is convex, tends to infinity when ρ goes to the mean rate m, and is zero for $\rho \geq h$. The network consists of a shared link with capacity C and buffer B. We also assume that there is enough demand to fill the link and that the iterations between the network (that sets prices) and the users (that respond by selecting new tariffs) proceed in lock-step.

In the case of deterministic multiplexing (zero cell loss), the effective bandwidth theory [Kel96a, CKW97] suggests that the value of the space parameter s should be ∞ (this follows from (2.8), page 25, when $\gamma = \infty$), and that the worst case effective bandwidth of a connection of type j which is policed by (β_j, ρ_j) is $\alpha_i(\infty, t) = \rho_j + \frac{\beta_j}{t}$ for $t > 0$ and $\alpha_i(\infty, 0) = \beta_j$.

Recall from Section 2.1.1 that the acceptance region can be approximated by (2.5): $NA = \bigcap_{0<t<\infty} NA_t$, where NA_t is given by (2.6):

$$NA_t = \left\{ (N_1, N_2, \ldots N_J) : \inf_s \left[st \sum_{j=1}^{J} N_j \alpha_j(s,t) - s(Nb + Nct) \right] \leq -\gamma \right\} . \tag{4.8}$$

In the case of deterministic multiplexing, the above becomes [Kel96a]

$$NA = NA_0 \cap NA_\infty ,$$

where

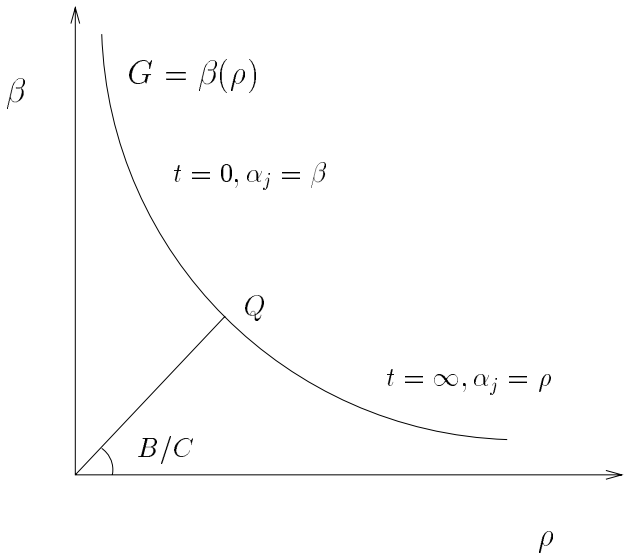


Figure 4.18: **Incentive compatibility in the case of deterministic multiplexing.** For points above Q, users will tend to decrease their bucket β, whereas for points below Q, users will tend to decrease their leak rate ρ.

$$NA_0 = \left\{ (N_1, N_2, \ldots N_J) : \sum_{j=1}^{J} N_j \rho_i \leq C \right\} , \; NA_\infty = \left\{ (N_1, N_2, \ldots N_J) : \sum_{j=1}^{J} N_j \beta_j \leq B \right\} , \quad (4.9)$$

with N_j being the number of users of type j, and J is the total number of user types.

Since the system proceeds in lock-step and users have identical requirements, at any point in time their choices will coincide, hence the constraints in (4.9) become $N\rho \leq C$ and $N\beta \leq B$.

Consider the point $Q \in G$ where the two constraints coincide and the number of users N is maximized, i.e., the social welfare is maximized. This point lies on the intersection of curve G with the line with slope B/C that passes from the origin (see Figure 4.18). For any initial choice of (β, ρ) by users which corresponds to a point M on G which is below Q, the system will fill so that the active constraint will have a corresponding value $t = \infty$ for the calculation of the effective bandwidth, whereas if M lies above Q, then $t = 0$.

Assume now that the time-volume pricing approach is used by the network. If the users choose a point M below Q, then the first constraint will be active, hence $t = \infty$ and the charge will be proportional to the effective bandwidth which is ρ; this will guide the users to reduce ρ and move towards Q. If the users choose M above Q, then the charge will be proportional to β; this will guide the users to reduce β and move towards Q. Assuming that, in order to avoid oscillations, users are allowed to make small changes of their tariffs, the point Q will eventually be reached.

Since both constraints are active at point Q, the price at that point will be proportional to a linear combination $\lambda_1\rho + \lambda_2\beta$ of the effective bandwidths corresponding to the active constraints at Q, where λ_1, λ_2 are the shadow prices of the optimization problem which maximizes the number of users under constraints (4.9). One can check that the above prices correspond to the tangent of G at Q, hence Q is an equilibrium since the user minimizes his charge by remaining there.

In the case of statistical multiplexing, the above arguments can be extended to show a similar user-network behavior. We assume that the network charges using the simple bound (4.4):

$$\widetilde{\alpha}(s,t;m,\beta,\rho,h) = \frac{1}{st}\log\left[1 + \frac{tm}{\bar{X}[0,t]}\left(e^{s\,\bar{X}[0,t]} - 1\right)\right], \tag{4.10}$$

where $\bar{X}[0,t] = \min\{ht, \rho t + \beta\}$. A rational user will always seek to minimize his charge. Hence, he will select the pair $(\beta,\rho) \in G$ that minimizes his charge which is given by (4.10). Hence, the user performs the following optimization:

$$\text{User: } \min_{(\beta,\rho)} \widetilde{\alpha}(s,t;m,\beta,\rho,h) \quad \text{such that} \quad (\beta,\rho) \in G\,, \tag{4.11}$$

where m, h is a user's mean rate and peak rate, respectively.

On the other hand, the network tries to maximize the number of users (all of the same type) that it can accept while satisfying its QoS constraint $P(\text{overflow}) \leq e^{-\gamma}$. Using the many sources asymptotic (2.3), page 23, the network performs the following maximization:

$$\text{Network: } \max N \quad \text{such that} \quad \sup_t \inf_s \left[stN\widetilde{\alpha}(s,t;m,\beta,\rho,h) - s(Ct+B)\right] < -\gamma\,, \tag{4.12}$$

where m, h is a user's mean rate and peak rate respectively and β, ρ are a user's leaky bucket parameters.

Performing the maximization in (4.12), the network computes a new pair (s,t), i.e., the link's operating point moves. The new values of (s,t) will affect a user's price, since (4.10) is a function of s,t. The users now perform the maximization in (4.11) and select a new leaky bucket (β,ρ), which in turn will affect the link's operating point, hence the tariffs, and so on.

We have calculated such equilibria for a link capacity 34 Mbps, target overflow probability 10^{-6}, and different buffer sizes, when the user traffic consists of Internet WAN traffic. Table 4.1 compares the corresponding equilibria for statistical and deterministic multiplexing. The substantial increase of the utilization in the case of statistical multiplexing is expected, since the mean rate of the traffic we considered (Internet WAN traffic) was very small relative to its peak rate, and the peak rate was very small relative to the link capacity (9 Kbps vs. 34 Mbps).

Buffer (bytes)	Deterministic mult.			Statistical mult.		
	ρ (Mbps)	β (bytes)	$N_{\max}$	ρ (Mbps)	β (bytes)	$N_{\max}$
0.5×10^6	0.615	10600	33	0.475	29100	1530
1×10^6	0.553	18300	54	0.399	52800	1650
5×10^6	0.373	62500	80	0.202	175500	2070
10×10^6	0.285	95500	105	0.162	341100	2170
15×10^6	0.246	124300	121	0.126	490500	2270

Table 4.1: **Comparison of the equilibrium under deterministic and statistical multiplexing.** As expected, statistical multiplexing leads to a higher utilization than deterministic multiplexing. Also observe that the values of β in the statistical multiplexing case are high. This has interesting implications of how traffic shaping affects the amount of resources used by connections. [$C = 34$ Mbps, $P(\text{overflow}) \leq 10^{-6}$ (for the statistical multiplexing case), Internet WAN traffic]

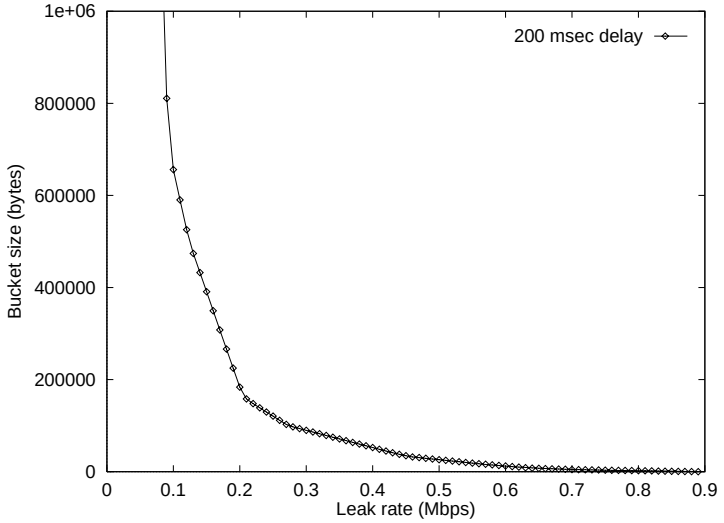
4.6.1 Effect of Traffic Shaping

We are now in a position to make some interesting observations about the effect of users delaying their traffic into the network. As we will argue, for the anticipated buffer sizes, shaping has a surprisingly small effect on the overall multiplexing capability of the network.

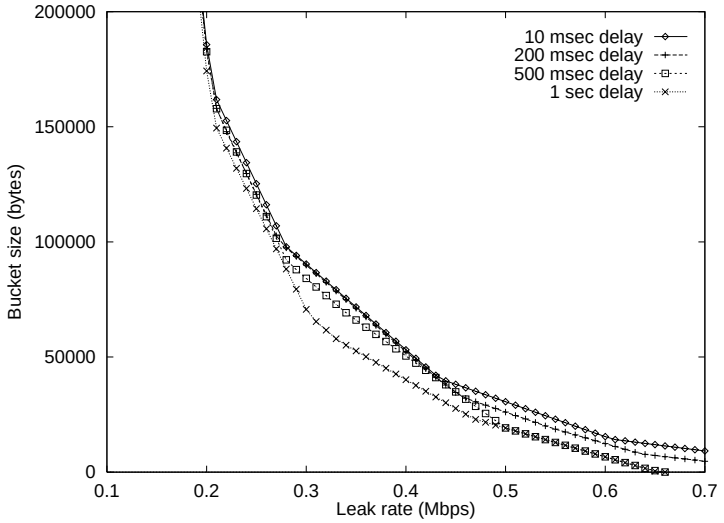
First, observe in Figure 4.19 that for large values of β, the indifference curve $G(d)$ is not greatly affected when the shaping delay is smaller than 500 msec. Second, observe in Table 4.1 that for the case of statistical multiplexing and for total buffer greater than 1×10^6 bytes, at the equilibrium we have $\beta > 50000$ bytes. Combining these two observations we see that when the shaping delay is less than 500 msec and the buffer size is larger than 1×10^6 bytes, the equilibrium point is not affected by traffic shaping.

Of course a user can use shaping to make a contract with a lower peak rate. However, contrary to the intuition, this will not affect his effective bandwidth as seen by the network, since the time parameter t at the equilibrium is always large enough so that $ht > \beta + \rho t$. In this case, the effective bandwidth is determined largely in terms of the values (β, ρ) (e.g., if the user sends traffic close to the maximum amount allowed by the simple bound (4.10)) which, as argued previously, remain practically unaffected by shaping.

The above discussion demonstrates (see also Section 2.3) how the performance evaluation framework described in Section 2.1 clarifies the effects of various time scales and the importance of the various traffic and network parameters on the amount of resources used by connections.



(a) Leak rate/bucket size tradeoff for 200 msec shaping delay.



(b) Leak rate/bucket size tradeoff.

Figure 4.19: **Leak rate/bucket size tradeoff for Internet WAN traffic.** The graphs show the (β, ρ) pairs such that all traffic is compliant. In the bottom curve observe that traffic smoothing has a small effect for large values of the bucket size. [Internet WAN traffic segment with duration $\approx$ 2 hours]

4.7 Conclusions

This chapter has dealt with one important aspect of pricing network services: efficiently measuring the resource usage component of a charge. The schemes investigated are based on bounds of the effective bandwidth, and included time-volume pricing schemes that involve two measurements (time and volume) for the whole duration of a connection, and schemes that involve measurements in distinct time intervals, smaller than the duration of a connection (pricing with renegotiation and the virtual bucket scheme).

The pricing schemes were evaluated for MPEG-1 compressed video with various contents and Internet WAN traffic according to their fairness, i.e., their ability to capture the relative amount of resource used by connections. The results from our experiments can be summarized as follows:

- *Time-volume pricing schemes:* In general, the fairness and robustness of the time-volume pricing schemes is better for larger link capacities and buffer sizes, for smaller call durations, and for smoother (less bursty) traffic. Indeed, for large capacity links (over 622 Mbps) the performance of these schemes approaches that of the schemes which involve measurements in distinct time intervals. This indicates that for high bandwidth links with a high degree of statistical multiplexing, very simple approximations can achieve good performance.

- *Pricing schemes which involve measurements in distinct time intervals:* Such schemes can achieve higher performance in terms of fairness and robustness to small variations of the link's operating point (which is characterized by the space and time parameters) compared to the time-volume schemes which involve measurements for the whole duration of a connection. This comes at the cost of higher accounting overhead. Furthermore, the performance of the virtual bucket scheme, which does not entail the signaling overhead due to renegotiation, was found to be very close to the pricing with renegotiation scheme.

- *Incentive compatibility of the time-volume pricing schemes:* We demonstrated, in a particular setup, the incentive compatibility of the time-volume schemes. The role of incentive compatibility is to slowly move the global system (the network and its users) to an efficient and stable operating point. Furthermore, we compared the equilibrium operating point of a link, in the presence of pricing, for the case of deterministic and statistical multiplexing and investigated to what extent it is affected by traffic shaping.

In our experiments we assumed that users have full knowledge of their traffic. Even with the existence of sophisticated user agents, there will inevitably be some uncertainty in determining quantities such as the mean rate and leaky bucket parameters for the time-volume schemes, and the leak rate for the virtual bucket scheme. The degree and effects of this uncertainty remains an open issue.

Chapter 5

Pricing and Resource Sharing for Available Bit Rate (ABR) Services

In Chapter 4 we investigated usage-based pricing schemes for services with *open loop congestion control*, such as the Constant Bit Rate (CBR) and Variable Bit Rate (VBR) services in ATM networks, and Internet access services. In the case of network services with *closed loop congestion control*, such as the Available Bit Rate (ABR) service in ATM networks, there is no guaranteed QoS[1], nor a traffic description to which the user must conform. Rather, the network sends congestion control signals to the user, based on which he increases (if there is no congestion) or decreases (if there is congestion) the rate at which he is sending traffic. Applications that are intended to use such services can be easily rescheduled for different times of a day, hence static time-of-day pricing would not be effective for controlling congestion. What is needed is a *dynamic* account of the level of congestion. For the above reasons, pricing ABR services poses different challenges and problems compared to pricing services with open loop control.

In this chapter we investigate the problem of pricing ABR services, which is closely related to how resources are shared among ABR connections. The key features of our approach are the following: (i) dynamic prices are used to capture the dynamically varying level of congestion, (ii) users declare a price per unit of time which buys them some amount of information transfer capability, and (iii) resource sharing is based on effective bandwidths which take into account the burstiness of traffic. In addition, the approach can be cast onto the existing rate-based congestion

[1]ABR services also support a Minimum Cell Rate (MCR). We will discuss this issue later.

control framework of ABR services. The prices are internal to the network. The user only receives, as in normal ABR operation, the maximum rate he can send traffic with. The second and third feature allow users to adequately reveal their preferences by declaring a single parameter (the price per unit of time). The key properties of the approach, namely dynamic prices and resource sharing based on effective bandwidths, are demonstrated through a comprehensive set of simulation experiments.

5.1 Technological Characteristics and Desirable Properties of Pricing

The Available Bit Rate (ABR) service category is intended for applications that can gracefully adapt their sending rate to time-varying available bandwidth [ATMF96, CLS96]. Such applications have been termed *elastic* in the Internet community [BCS94, She95b]. Although ABR does not support specific Quality of Service (QoS), other than the guarantees provided by a Minimum Cell Rate (MCR), it is expected that sources which adapt to the congestion control signals will experience small loss probability and small delay. ABR connections will use the bandwidth that is left over from guaranteed services, such as CBR and VBR (Figure 5.1), with - possibly - some minimum amount of bandwidth reserved for ABR services. Hence, ABR has the potential of efficiently utilizing link bandwidth, while satisfying the performance requirements of many applications.

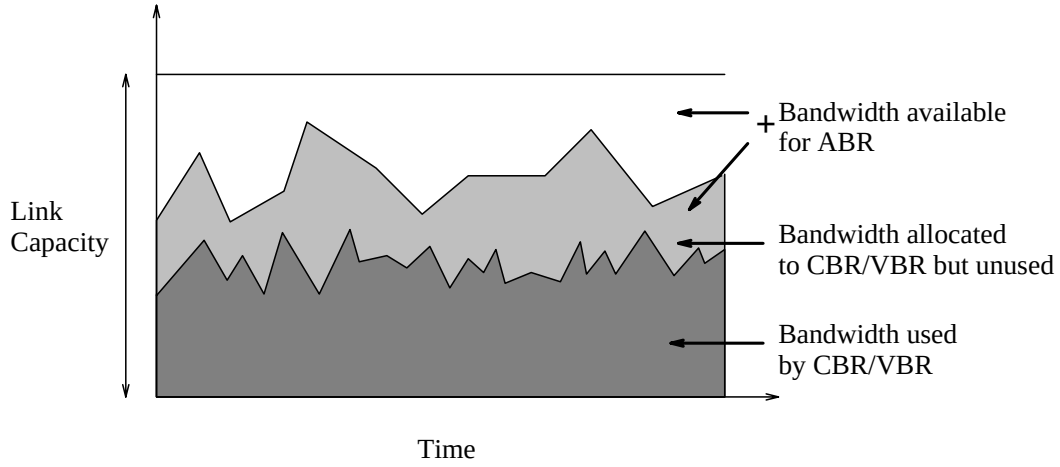


Figure 5.1: **Bandwidth available for ABR connections.** ABR connections will use the bandwidth that is not used by CBR and VBR connections, with - possibly - some minimum amount of bandwidth reserved for ABR services.

The dynamic nature of the bandwidth that will be available for ABR services requires a feedback mechanism which informs the users about the level of congestion so they can increase their rate

when their is unused bandwidth, or decrease their rate during congestion. This is achieved through a closed loop rate-based congestion control mechanism, which comes in different variants (or *modes*) depending on the amount of information in the feedback signal and how this information is conveyed from the network to the users [ATMF96]. At one extreme, called *Explicit Forward Congestion Indication (EFCI) marking*, the feedback is sent using a single bit in the header of ATM cells and consists of a binary signal which indicates the presence or absence of congestion. At the other extreme, called *Explicit Rate marking*, the feedback is sent using special control cells, called *Resource Management* (RM) cells, and consists of an Explicit Rate (ER) which is the maximum traffic rate the source can have. We will deal with the latter of the two modes.

Although not part of the specifications, switch algorithms for computing explicit rates share link capacity according to some policy or *fairness criteria*. The most common fairness criteria is *max-min fairness* which gives higher priority to connections with smaller rates. A property of such an allocation policy is that all connections with equal rates are treated the same. However, in an environment where different users value the same service differently, switch algorithms must be able to provide a different level of performance to different connections.

Applications that are intended the use ABR services tend to be more flexible, since their use of the network can be easily rescheduled for different times of a day. This has implications on the requirements of a pricing scheme which seeks to reduce the negative effects of congestion. For example, consider using time-of-day pricing with prices lower during the night. Facing such a pricing scheme, users would schedule their use of the network (e.g., file transfers or transfers of interesting web pages) for that time, thus shifting congestion to the night hours. Hence, static time-of-day pricing would be ineffective for controlling congestion. The above discussion leads us to a first requirement for ABR pricing, the need for a *dynamic* account of the level of congestion.

A second requirement for ABR pricing is that it should be incentive compatible and enable users to reveal their preferences for network usage. Consider two users, one bursty and one smooth, who are paying the same price. Also assume a scheme where prices depend solely on the explicit rate. Because the bursty user doesn't use his explicit rate at all times, he is actually using less resources than the smooth user who is using his explicit rate at all times. Since both users are paying the same, the pricing scheme is unfair to the bursty user. Facing such tariffs, the bursty user does not have the incentive to use less than his explicit rate at all times, even though he actually needs less. As a

result, network resources are used inefficiently. This discussion shows that prices should reflect not simple measures of flow, such as average or peak rate, but the actual amount of resources used.

The approach we present in this chapter satisfies the two requirements discussed above. In addition, it can be cast onto the existing ABR rate-based congestion control framework. The key features of the approach are the following:

- The price of network resources is adjusted according to the demand. Hence, the scheme can capture the dynamically varying level of congestion [CSS96, Kel97, LV93, MMP94, MMV95b].
- Users declare the price per unit of time according to which they will be charged. The amount of resources allocated to a user depends on this price and the (dynamic) price of network resources [Kel97].
- Resources are shared based on the effective bandwidth of a user. Hence, the scheme takes into account the bursty nature of traffic and allows users to adequately reveal their preferences through a single parameter (the price per unit of time) [CS97].

The first two features are discussed in the next section, while the third feature is discussed in Section 5.3.

We end this section with a discussion of the Minimum Cell Rate (MCR) that ABR services can support. The MCR represents the minimum cell rate with which the user can always send traffic. We will not discuss the problem of pricing connections with MCR > 0. We note, however, that two connections using the same amount of bandwidth but having a different MCR should be charged differently, since the connection with higher MCR has the guarantee of always being able to send traffic at that rate. A simple approach would be to use the pricing scheme described in this chapter, adding to it a price which depends on the value of MCR. The price per unit of MCR will depend on the price for CBR and VBR services, since it is expected that the requirements of many applications will be satisfied by any of these three services (i.e., CBR, VBR, and ABR with MCR > 0).

5.2 Pricing for Social Welfare Maximization

In this section we apply the theory of social welfare maximization to the problem of pricing ABR services. This leads us to prices which vary with the demand. The implementation of such

dynamic prices within the framework of ABR congestion control is discussed in Section 5.2.1, and related simulation results are presented in Section 5.2.2. In Section 5.2.3 we discuss a user-network interaction, first proposed in [Kel97], that hides the dynamic nature of prices from users which simply declare a price per unit of time according to which they will be charged.

Consider a network with a set of links L, and let C_l be the bandwidth available for ABR services on link l. We assume that this bandwidth is either constant or varies slowly. As discussed in Section 5.1, link capacity in broadband networks will be shared among guaranteed services and services with no performance guarantees, such as ABR services, with higher priority given to guaranteed services. As a result, bandwidth available for ABR services will vary depending on the aggregate traffic of guaranteed services. However, due to the aggregation of a large number of connections, which is anticipated in the case of high capacity broadband links, the bandwidth available for ABR services is likely to vary slowly.

Let $\mathcal{C}$ be the set of connections (users) in the network. Denote R_i and $U_i(x_i)$ the route and utility of user i, where x_i is the rate at which the user is allowed to send traffic. We assume that utility functions are increasing, strictly concave, and differentiable. According to economic theory, the socially optimal allocation of rates is given by the solution to the following problem:

$$\max_{x_i, i \in \mathcal{C}} \sum_{i \in \mathcal{C}} U_i(x_i), \text{ such that } \sum_{i: l \in R_i} x_i \leq C_l \, , \forall \, l \in L \, . \tag{5.1}$$

The sum of all the utilities is the *social welfare* of the system (the network and its users). The above maximization involves utility functions which are unlikely to be known by the network, making a centralized solution to the problem infeasible. What is needed is a decentralized approach for solving (5.1). Such problems are quite common in economic theory [Var92], and have found a number of applications in computer networks (e.g., see [CSS96, Kel97, LV93, MMP94, MMV95b, SFY95, San88, and the references therein]). A distinguishing feature of the problem we are studying is that an ABR connection will be allocated the same amount of bandwidth on all links it traverses. The standard approach is to introduce prices λ_i per unit of rate, and allow users, who seek to maximize their utility, to select the amount of resource (rate in our case) they wish to buy. The prices λ_i are set by the network such that the optimal selection of rates for the user is also the optimal action from the viewpoint of the network [Var92].

A user i wishing to maximize his utility is faced with the following problem:

$$\max_{x_i \geq 0} U_i(x_i) - \lambda_i x_i \,. \tag{5.2}$$

The quantity $\lambda_i x_i$ represents the charge for user i who uses rate x_i. It can be proved [Kel97] (also see [LV93]), that there exist prices $(\lambda_i, i \in \mathcal{C})$ such that the vector of rates $x = (x_i, i \in \mathcal{C})$ which solves (5.2) for all users i, also solves the social welfare maximizing problem (5.1). Furthermore, the vector x also maximizes the network revenue $\sum_{i \in \mathcal{C}} \lambda_i x_i$. The theory also tells us that the following hold for the social welfare maximum:

$$\lambda_i = \sum_{l \in R_i} \mu_l \,, \qquad \forall\, i \in \mathcal{C} \tag{5.3}$$

$$\mu_l \left(\sum_{i: l \in R_i} x_i - C_l \right) = 0 \,, \qquad \forall\, l \in L \tag{5.4}$$

where $(\mu_l, l \in L)$ are the *shadow prices* of the maximization problem (5.1).

From equation (5.3), the price per unit of rate λ_i for connection i is the sum of the shadow prices μ_l on all links along the route R_i of connection i. A second observation, which is important for the implementation of dynamic prices, is that in the state of maximum social welfare the prices are determined by the congested links. This agrees with (5.4) which holds if $\sum_i x_i < C_l$ and $\mu_l = 0$, i.e., the price for a non-congested link is zero, or if $\sum_i x_i = C_l$ and $\mu_l \neq 0$, i.e., the price is non-zero for congested links.

5.2.1 Dynamic Pricing

The solution to problems (5.1) and (5.2) can be performed decentralized if at each link l, the price per unit of rate μ_l is increased or decreased if $C_l < \sum_i x_i$ or $C_l > \sum_i x_i$, respectively. This can be written as

$$\frac{d}{dt}\mu_l(t) = \kappa \left(\sum_{i: l \in R_i} x_i(t) - C_l \right) \,, \tag{5.5}$$

where κ is some small constant. Prices are updated in time intervals (we will assume that they all have the same length) whose duration depends on how fast the aggregate demand $\sum_i x_i(t)$ changes. This interval (price update interval) should not be less than the interval from the time new prices are posted until the time the users' responses to these prices are received by the network, i.e., one round-trip delay.

Next we discuss how to implement dynamic prices within the framework of ABR congestion control.

According to ABR rate-based congestion control [ATMF96], every source sends special control cells, called Resource Management (RM) cells, either periodically or after a specific number of data cells carrying user information (Figure 5.2). When the destination receives an RM cell (called forward RM cell), it copies its information to a *backward* RM cell which it sends back to the source. When an intermediate switch detects congestion on one of its links, it sets a congestion indication bit and places a rate in the Explicit Rate (ER) field of all backward RM cells traversing that link.[2] While the source does not receive RM cells, or while it receives RM cells with the congestion bit set, it decreases its cell rate by some percentage (multiplicative decrease). Furthermore, its cell rate must always be smaller than the explicit rate in the received RM cell. If the source receives an RM cell with the congestion indication bit cleared, it is allowed to increase its cell rate by some additive quantity (additive increase).

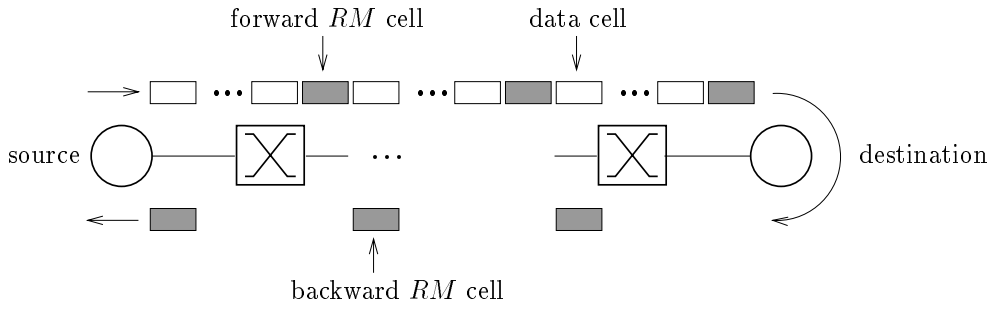


Figure 5.2: **ABR congestion control loop**. Dynamic prices can be implemented by having every source i place its demand x_i in the ER field of the forward RM cells, and having switches add the price μ_l for link l in a new price field P in the backward RM cells. When a backward RM cell reaches the source, P will contain the value $\lambda_i = \sum_{l \in R_i} \mu_l$.

To implement the dynamic pricing scheme, switches adjust the price of bandwidth for each of their output links using a discrete version of equation (5.5). This equation requires that the total demand is known to the switches. The latter can be achieved by having every source place their demand in the ER field of the forward RM cells.

We also need a way to form the price $\lambda_i = \sum_{l \in R_i} \mu_l$ for each connection i. This can be done by adding a new price field P in the RM cell. When a backward RM cell is sent by the destination, the value of P is zero. Each switch adds the price μ_l to the amount contained in the price field P of

[2] In ATM networks, the forward and backward RM cells follow the same route.

the backward RM cells that traverse link l. When a backward RM cell reaches the source, P will contain the value $\lambda_i = \sum_{l \in R_i} \mu_l$.

Recall from the previous section that prices are non-zero only for congested links. It is very common that a network consists of a number of high bandwidth *internal* links which are not congested and a small number of low bandwidth *external* links which are congestion (Figure 5.3). In such a case, only the price of the external links will be non-zero, hence only the switches directly connected to these links need to implement dynamic prices.

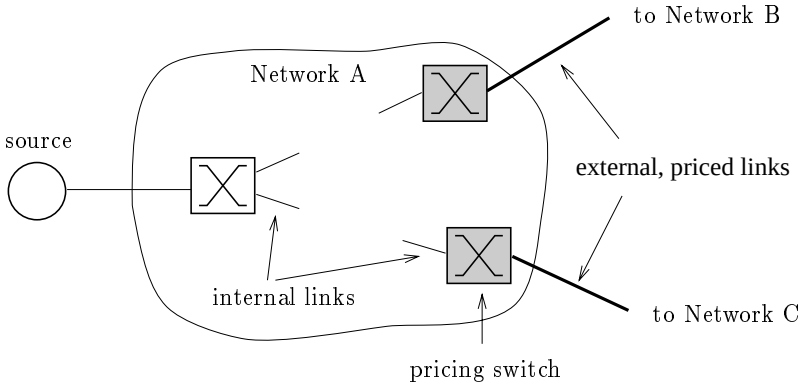


Figure 5.3: **Implementation of dynamic prices in a network.** Only congested links (thick lines) have non-zero prices. Hence, only the switches (pricing switches) connected to these link need to implement dynamic prices.

5.2.2 Simulation Results

The goal of our simulation experiments[3] is to demonstrate the convergence properties and transient behavior of dynamic prices [CSS96]. Rather than simulating the network at the cell level, we have modeled the propagation of rate changes and the propagation of resource management cells (see Appendix B.3). The simulated network is shown in Figure 5.4. All link rates are 155 Mbps. We consider three different values for the distance between switches: 1 km, 100 km, and 1000 km. These values represent typical distances in local and wide area networks, and correspond to propagation delays 5 μsec, 500 μsec, and 5000 μsec respectively, if we assume that the propagation delay for optical signals is approximately 5 μsec/ km [WV96, page 205].

We assume that switches update the prices $\mu_l(t)$ in discrete time intervals using the following

[3] In [CKS+97] we discuss some of the issues that arise for conducting ABR pricing experiments in a real network.

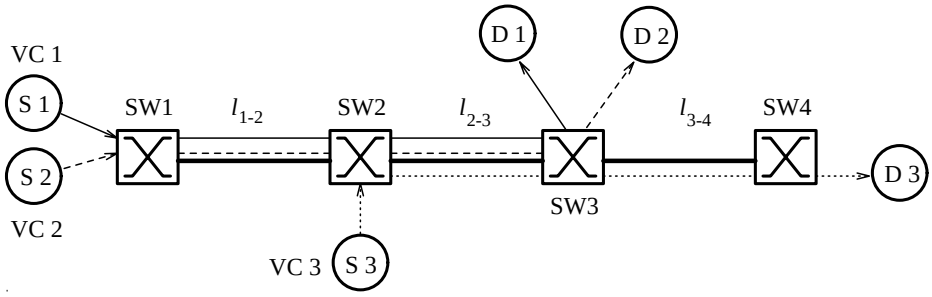


Figure 5.4: **Network topology for the dynamic pricing simulation experiments.** There are three Virtual Connections (VCs). The distance between switches is 1 km, 100 km, and 1000 km.

discrete time version of (5.5):

$$\mu_l(t) = \mu_l(t-1) + \kappa \left(\sum_{i:l \in R_i} x_i(t-1) - C_l \right) , \tag{5.6}$$

where C_l is the capacity of link l (we assume that all of this capacity is used by ABR services), $x_i(t-1)$ is the demand for bandwidth by connection i during the price update interval $t-1$. We assume that $\mu_l(0) = 0$ for all $l \in L$. In [CSS96] we have experimented with a price update function where κ is a function of $\mu_l(t-1)$.

For simplicity, we assume that sources send RM cells in fixed time intervals, equal to 200 μsec (interval-based behavior), rather than after a specific number of data cells (counter-based behavior). Furthermore, we assume that all sources i have a demand for bandwidth D_i which decreases exponentially with the price per unit of rate λ_i, i.e., $D_i(\lambda_i) = v_i e^{-\lambda_i}$, where $v_i = 155$ Mbps.

In the first experiment, the distance between switches was 1 km and the price update interval was 200 μsec. The value of parameter κ in (5.6) was set to 0.0005 Mbps^{-1}. This parameter determines the number of round-trip times needed for prices to converge (on the contrary, the number of round-trips is independent of the round-trip delay). Larger values of κ could lead to oscillations. In general, the selection of κ depends on the number of multiplexed Virtual Connections (VCs), the magnitude of changes relative to link capacity, the network topology, and the sources' demands. Figure 5.5 shows that prices converge quickly after changes of the input traffic.

In the second experiment, the distance between switches was 100 km and the price update interval was 2.5 msec. Parameter κ in equation (5.6) was 0.001 Mbps^{-1}. From Figure 5.6(a), we see that convergence times ($\approx$ 30 msec) are greater than the case of 1 km links ($\approx$ 5 – 10 msec).

In practice, this will not be a problem since in wide area networks, due to the high aggregation of traffic, changes are expected to be smaller and less frequent.

Figure 5.6(b) shows the dynamic behavior of prices when the distance between switches was 1000 km. The price update interval in this case was 21 msec. In the experiments of Figures 5.6(a) and 5.6(b) the network reached the equilibrium state after the same number of iterations (approximately 15). However, the convergence time is longer when the distance between switches is larger. This is due to the larger pricing interval (21 msec for 1000 km links compared to 2.5 msec for 100 km links) which is required because of the longer round-trip delay.

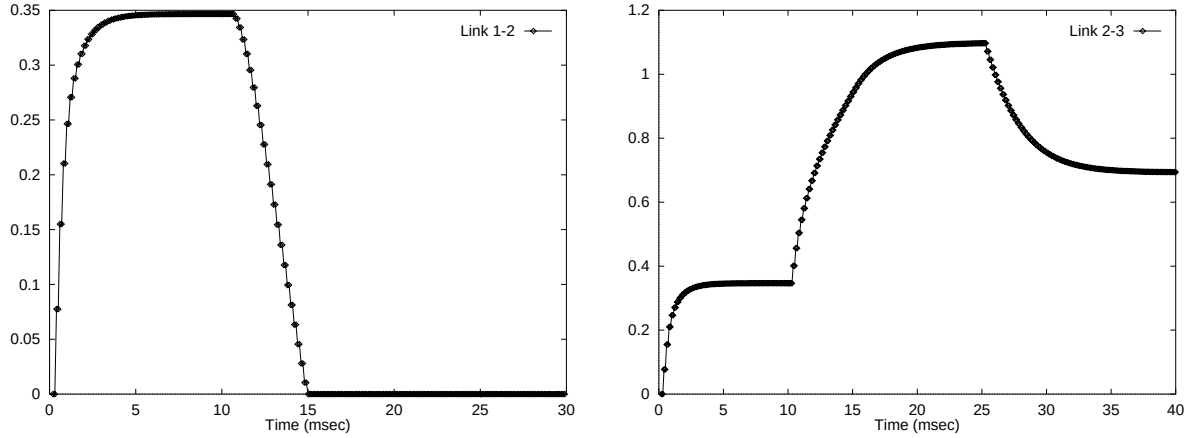
5.2.3 An Alternative User-Network Interaction

In the simulation experiments of the previous section, we assumed that users directly react to prices. With such a scenario there are two closed loop controls to which users need to react: pricing and congestion control. In this section we describe an alternative user-network interaction, first introduced in [Kel97], which combines the two control loops, while still achieving the social welfare maximum of (5.1). This has the advantage that the user behaves as he would under normal ABR congestion control operation.

The basic idea behind the scheme is to have users select, at connection setup or during renegotiation, the price per unit of time according to which they will be charged and have the network share its resources based on the users' selection of prices. A user who selects a higher price will be offered a larger share of bandwidth compared to a user who selects a lower price.

The scheme has the property that the user does not know *a priori* the performance he will receive; the latter depends on the other users who are also using the system. Similar uncertainty exists with Internet's flat rate pricing schemes: The price is fixed, but the performance one receives depends on who else is using the network at the same time. However, all users necessarily receive the same performance. On the other hand, the scheme described in this section allows users who wish to receive better performance to do so by paying more.

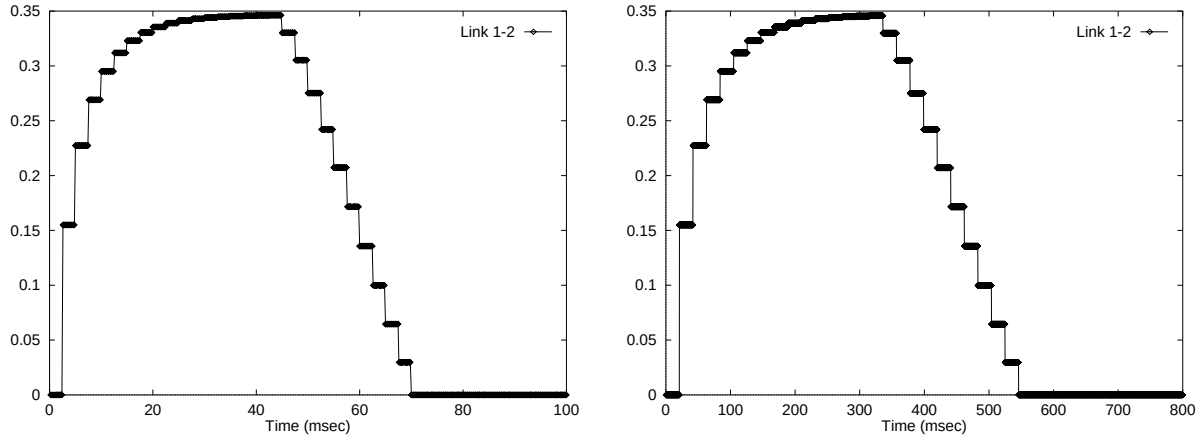
The pricing scheme introduced above has similarities with the pricing of transportation services. First, consider the bus. Although the fare is the same, the service one receives in not always the same. If one takes the bus during rush hours it is very likely that he will not find a seat and will have to stand, in a very crowded bus, for the whole duration of his trip. If one takes the bus during



(a) Price at link $l_{1\rightarrow 2}$

(b) Price at link $l_{2\rightarrow 3}$

Figure 5.5: **Dynamic pricing simulation results for 1 km switch distances.** VC 1 started at 0 msec and terminated at 25 msec, VC 2 started at 0 msec, and VC 3 at 10 msec. Converge time is approximately 5 – 10 msec.



(a) Results for 100 km switch distances. Convergence time is approximately 30 msec.

(b) Results for 1000 km switch distances. Convergence time is approximately 300 msec.

Figure 5.6: **Dynamic pricing simulation results for 100 km and 1000 km switch distances.** Price at link $l_{1\rightarrow 2}$ when the distance between switches is 100 km and 1000 km, respectively. VC 1 and VC 2 started at 0 msec, and VC 3 at 40 msec (for 100 km switch distances) or 300 msec (for 1000 km switch distances).

the non-rush hours of the day, then it is likely that he will travel in an uncrowded bus and might even find a seat. Finally, if one takes the bus late at night, then he will almost surely find a seat. The service differs also in terms of the time to reach the final destination. During rush hours, because of traffic, the trip will take a longer time than it takes during non-rush hours. The above is an example of a service which has a fixed associated cost, but whose service characteristics differ rather dynamically. Despite the dynamic nature of the offered service, one can make a reasonably good guess of the service performance he will receive.

Similar to buses, taxis and private cars also have an (almost) constant cost. Furthermore, using a car is more expensive than using a taxi, which in turn are more expensive than using a bus. Still the service characteristics are not known a priori. During rush hours, the duration of one's trip will be longer. Indeed, all these services share one common resource: roads. If roads are congested, e.g., due to cars, then the service of both buses and taxis will be effected.

Hence, there are different levels of transportation services (bus, taxi, and car) with fixed costs, but with varying service characteristics. Furthermore, because they all share a common resource (roads), the service a customer of one level receives depends not only one how many other customers are using the same service level, but also on how many customers are using the other service levels.

Rate control algorithm

Next we describe the rate control algorithm according to which the network shares link capacity using the above pricing scheme [Kel97, KMT98].

Suppose that user i chooses a price per unit of time w_i at connection setup or at renegotiation, if the latter is supported. Then the network allows him to send traffic with rate x_i. Hence, $\lambda_i = w_i/x_i$ is the price per unit rate. User i, in order to maximize his benefit, will solve the following maximization problem:

$$\max_{w_i \geq 0} U_i\left(\frac{w_i}{\lambda_i}\right) - w_i\,. \tag{5.7}$$

It is important to note that the user does not have knowledge of the prices λ_i, which are *internal* to the network. All he receives from the network is a rate x_i which is the maximum rate with which he can send traffic.

Suppose now that the network knows the vector of prices $(w_i, i \in \mathcal{C})$ and performs the following

maximization problem:

$$\max_{x_i \geq 0, i \in \mathcal{C}} \sum_i w_i \log x_i \text{ such that } \sum_{i: l \in R_i} x_i \leq C_l, \forall l \in L\,. \tag{5.8}$$

In [Kel97] it is shown that the decomposition of the social welfare maximization problem (5.1) into (5.7) and (5.8) is sound, i.e., there exist vectors $\lambda = (\lambda_i, i \in \mathcal{C})$, $w = (w_i, i \in \mathcal{C})$, $x = (x_i, i \in \mathcal{C})$ satisfying $w_i = \lambda_i x_i$ for all $i \in \mathcal{C}$, such that w solves (5.7), x solves (5.8), and x is the unique solution of (5.1). The latter means that the vector x maximizes the social welfare of the global system (the network and its users).

The above results imply that if the users, independently of the network, optimize their benefits by solving (5.7) and communicate to the network the amounts w, and if the network solves (5.8) and communicates to the users the flows x, then the system will have as an equilibrium the social welfare maximum of (5.1). The critical issue is the solution of (5.8) by the network.[4]

In [KMT98] it is shown that, for appropriate values of parameter κ in (5.5), if prices μ_l at each link l are updated according to (5.5), where now the rates x_i are computed by the network using

$$x_i(t) = \frac{w_i}{\sum_{l \in R_i} \mu_l(t)}\,, \tag{5.9}$$

then system will iteratively converge to the solution of (5.1), (5.7), and (5.8). Furthermore, the convergence is robust with respect to stochastic effects and propagation delays.

5.3 Resource Sharing Based on Effective Bandwidths

The previous section discussed how prices are updated (equation (5.5)) and how rates are allocated to users (equation (5.9)). Interpreting these rates, or information transfer capabilities, as traditional measures of flow, such as the average or maximum rate, does not capture the requirements of *bursty* users. Such users, in addition to the average rate, also value the amount of distortion of their traffic due to the limited rate at which the network allows them to send traffic. In this section we motivate and present an approach that allocates rates according to the effective bandwidth of users. One of the important properties of the approach is that it can be cast onto the current rate-based congestion control framework of ABR services, according to which the network sends to the source an explicit rate (ER) which is the maximum rate at which the source can send traffic.

[4] Solving (5.8) is interesting by itself, since the resulting allocation of flows has important fairness properties [Kel97].

The network transfer capability that a user receives depends, in addition to the average rate, also on the explicit rate ER assigned by the network. Consider a bursty source which produces traffic with mean rate m and peak rate h. Assume that the network shares resources using a control procedure which operates at a time scale larger than the time scales of the bursts of the input sources, and which chooses for each source the maximum rate ER at which the source is allowed to send traffic. If $ER \geq h$ then the source's bursts are not distorted at the entrance of the network. If $m \leq ER < h$, then many bursts will be flattened out (by buffering at the source), resulting in less bursty traffic. If $ER \leq m$, then the traffic enters the network at a constant rate.

A ''burst-sensitive'' user values highly a network service which allows him to send his bursts into the network unaffected. Hence, such a user is willing to pay a high price for maintaining high values of ER. On the other hand, a ''burst-insensitive'' user does not care whether his bursts are flattened out before entering the network, hence will be indifferent to the value of ER; his only requirement is that it is greater than his mean rate.

A pricing scheme should allow both burst-insensitive and burst-sensitive users to adequately reveal their preferences for network usage. To do this, prices must reflect the actual amount of resources that are used, hence must take into account the burstiness of traffic. In Chapter 2 and 4 we have argued for using the effective bandwidth for quantifying resource usage. This motivates us to interpret the information transfer capability x_i assigned to user i using (5.9) as the effective bandwidth of that user, which now is controlled by the network through the explicit rate ER it sends to the user. Hence, x_i is an ''effective rate'' for user i. A simple formula we propose to use for quantifying x_i is the on-off bound (Section 4.2.1):

$$\widetilde{\alpha}(m,h) = \frac{1}{st} \log \left[1 + \frac{m}{h} \left(e^{sht} - 1 \right) \right] , \tag{5.10}$$

where the parameters s, t characterize the link's operating point, m is the user's mean rate, and h is the peak rate the user is allowed to send with.

Contrast the way (5.10) was used for pricing services with open loop control (Chapter 4) with the way we propose to use it for ABR services: For services with open loop congestion control, the peak rate is part of a user's traffic contract. The user's mean rate is measured and, if he has made the optimal tariff selection, his price would be given by (5.10). For ABR services, the user declares a price per unit of time w_i at connection setup or during renegotiation. This price will buy him some amount of effective rate x_i determined by (5.9). Throughout the duration of his connection, the

user's mean rate m_i is measured and using (5.10) with $\widetilde{\alpha}(m_i, h_i) = x_i$, the value of h_i is computed; this will be the explicit rate ER that the network sends to the user.

By relating prices to the effective bandwidth, we are able to provide the users with the right incentives for efficient use of network resources. A user that does not tolerate having his bursts smoothed will have a higher price w_i compared to a user with the same mean rate, but who does not value burstiness. Furthermore, the bursty user would not want to increase his mean rate, since that would result in a decrease of his explicit rate. Such a feature of incentive compatibility is lacking in the existing flow control procedures where the network allocates in a fair way peak rates with no reference to the corresponding mean rates (e.g., see [CRL96, JKG+97, KVR95, and the references therein] and [ATMF96, Appendix I.5]).

5.3.1 Controlling Effective Rates

In this section we investigate implementation issues of the effective flow control procedure whose mathematical underpinnings are based on the pricing mechanism described by equations (5.5) and (5.9), and the effective bandwidth approximation (5.10). Our goal is to use mechanisms already proposed for ABR rate-based congestion control. An important feature of our approach is that the source behavior is the same as that of a normal ABR source [ATMF96].

We first need to define the following notation. Let $\widetilde{\alpha}_l(m, h)$ be the effective bandwidth using the on-off approximation (5.10) of a connection with mean m and peak h at link l, hence l captures the link's operating point parameters s, t. Let $\widetilde{\alpha}_l^{-1}(m, x)$ be the value of the peak rate h for which $\widetilde{\alpha}_l(m, h) = x$, for some effective rate x.

The network can control a bursty source so that its effective rate through link l becomes equal to x by measuring the source's mean rate m and limiting the source's peak rate to the value $\widetilde{\alpha}_l^{-1}(m, x)$. This can be achieved with ABR rate-based congestion control by setting the explicit rate ER of the source equal to $\widetilde{\alpha}_l^{-1}(m, x)$. To allow for efficient implementation, simple tables can be used for computing $\widetilde{\alpha}_l^{-1}(m, x)$. An accurate estimate of the short term mean for each source is not necessary, since ABR users are more interested in their average explicit rate over a large time interval. Furthermore, the mean can be measured once at the first switch of the network and circulated to the intermediate switches via RM cells.

We start by defining the simplest version of our effective flow control implementation when all

links of the network have similar operating characteristics, and hence the values of s, t are identical throughout the network. Under the above simplification, we can drop the link subscript from the effective bandwidth formula, which becomes $\widetilde{\alpha}(m, h)$ for all links.

The implementation of the scheme is now straightforward. In the RM cells of a connection i we introduce a new field W storing the value of the willingness to pay w_i and, as described in Section 5.2.1, a price field P storing the sum of the per unit of bandwidth prices on all links the connection traverses, i.e., the value $\sum_{l \in R_i} \mu_l$. This sum is created by having each switch add the price μ_l to the price field P in the backward RM cells (refer to Figure 5.2). When the first switch connected to the source receives the backward RM cell, it computes the value of the effective rate $x_i = w_i / \sum_{l \in R_i} \mu_l$ by dividing the two fields W and P, and sets the ER field of the backward RM cells of connection i equal to

$$ER = \widetilde{\alpha}^{-1}(m_i, x_i)\,, \tag{5.11}$$

where m_i is the mean rate of connection i. Such a procedure will provably converge to a fair allocation of the effective rates (as defined by the formula (5.9)) which will solve the optimization problem (5.1).

Now we discuss the general situation where the operating conditions of the various links are different, and hence $\widetilde{\alpha}_l(m, h)$ possibly differs for different links l. In this case, the prices at each link are computed as before, but every link l must compute for each connection i the value of its explicit rate $ER' = \widetilde{\alpha}_l^{-1}(m_i, x_i)$, where $x_i = w_i / \sum_{l \in R_i} \mu_l$. The latter requires that all switches know the sum $\sum_{l \in R_i} \mu_l$. This can be achieved by having the first switch connected to the source copy the value of the price field P of the backward RM cells to the price field P of the forward RM cells, thus making it available to all the switches along the route of connection i. After computing ER', the switch sets the value of the ER field of the backward RM cells belonging to connection i equal to the minimum of its current value and ER'.

The achievable utilization using the above procedures relies on one hand on the degree of multiplexing that takes place, and on the other hand on the accuracy of the effective bandwidth approximation (5.10). Indeed, as we have seen in Section 4.5, the formula (5.10) tends to be conservative, hence the system will be underutilized. A simple way to remedy this inefficiency, while keeping the essential merits of the approach, is to couple it with direct measurements of the cell loss probability (or equivalently, of the probability that cells are delayed more than some target

value). Hence, while the operating point (link prices) is defined as above, we slowly perturb it in order to increase the efficiency. This can be done by having each link l measure the buffer overflow probability (or some ''proxy'' of it), and compute a load factor z_l reflecting the underutilization or overutilization of the link resources. The effective capacity C_l in (5.5) is replaced by the product $z_l C_l$. While the cell loss probability is higher or lower than the prespecified maximum value, the value of z_l is decreased or increased. One anticipates that changing the load factor z_l at a slower time scale compared to the time it takes for prices to converge will eventually stabilize the system at an operating point where all links have the desired maximum utilization.

The procedures that we have described in this section for implementing the effective bandwidth ABR (EB-ABR) resource sharing scheme are summarized in Table 5.1. Note that these procedures refer solely to the operation of switches. The source behavior is the same as that of a normal ABR source [ATMF96]. The user is only required to declare at connection setup the price per unit of time w_i.

/* Link parameters */
κ : price update parameter
T_{p} : price update interval
T_{m} : mean rate measurement interval
C_l : link capacity available for ABR connections

/* Link variables */
μ_l : price per unit of bandwidth
z_l : link load factor

/* Connection variables */
w_i : price per unit of time for connection i
λ_i : price per unit of bandwidth for connection i
V_i : total number of cells of connection i since last mean rate measurement

/* RM cell fields */
ER : Explicit Rate field
P : Price field
W : Willingness to pay field

/* functions performed for each link l */

if (forward RM cell received for connection i)
$\lambda_i = P$
$w_i = W$

if (backward RM cell received for connection i)
$P = P + \mu_l$
$x_i = w_i/\lambda_i$
$ER' = \widetilde{\alpha}^{-1}(m_i, x_i)$
$ER = \min(ER', ER)$

if (time to perform price update) /* performed once every T_{p} */
$\mu_l = \mu_l + \kappa \left(\sum_{i:l \in R_i} w_i/\lambda_i - z_l C_l \right)$

if (time to measure mean rate) /* performed once every T_{m} */
$m_i = V_i/T_{\mathrm{p}}$

Table 5.1: **Switch operation for the effective bandwidth ABR (EB-ABR) resource sharing scheme.** The first switch to which a source is connected to, in addition to the above, copies the value of the price field P of the backward RM cells (i.e., the value $\lambda_i = \sum_{l \in R_i} \mu_l$) to the price field P of the forward RM cells. The link load factor z_l changes in a smaller time scale (compared to the price updates), based on whether the loss probability (or probability that the queue length becomes greater than some threshold) is higher or lower than some target value.

5.3.2 Simulation Results

In this section we present numerical results comparing the effective bandwidth ABR (EB-ABR) resource sharing scheme described in the previous section with the ERICA (Explicit Rate Indication for Congestion Avoidance) scheme [JKG+97], and demonstrate how our approach differentiates connections with different mean rates. For simplicity, we assume that all sources are paying the same price, hence they are all entitled to the same effective rate.

As discussed in the previous section, our approach relies on a high degree of multiplexing. Here, we show that our scheme can achieve high allocation of explicit rates for capacity 155 Mbps (we assume that all of this capacity is used by ABR connections), and for traffic sources with bursts of approximately 943 cells or 360 kb. This value is large compared to the most common size of WWW documents which was found to be less than 80 kb [TMW97].

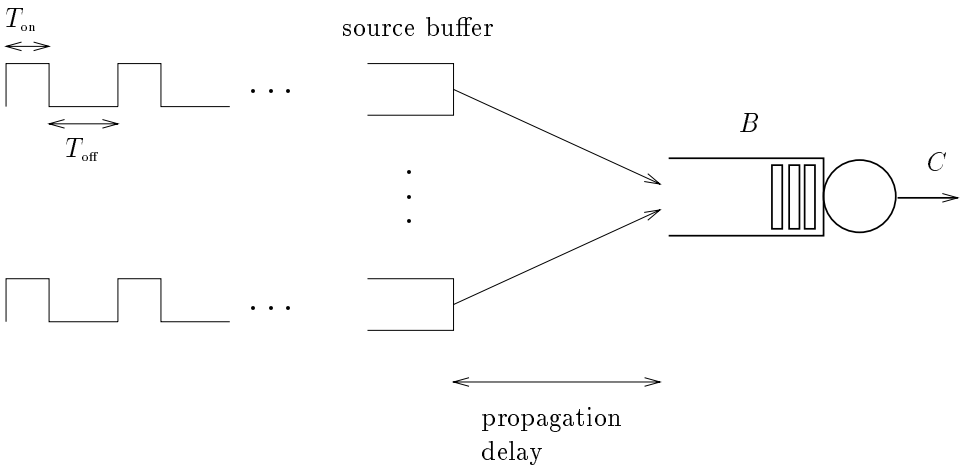


Figure 5.7: **Simulation model for the ABR resource sharing experiments.** Sources are periodic on-off with the same peak rate but different ''on'' and ''off'' durations.

Our simulation model consists of a single link with buffer B and capacity C which multiplexes N bursty sources (Figure 5.7). Similar to the dynamic pricing experiments, rather than simulating the network at the cell level, we have used a ''fluid'' simulator where the events are the change of rates, the arrival and departure of resource management cells, and the buffer becoming empty (see Appendix B.3).

Each source has a buffer for storing cells that are delayed due to the limited traffic rate of

the source. For simplicity, we assume that each source sends RM cells in fixed time intervals (interval-based behavior). However, our approach also works with a counter-based behavior where one RM cell is sent after some number of data cells [ATMF96]. Bursty sources are modeled as periodic on-off sources with deterministic ''on'' and ''off'' phases, denoted T_{on} and T_{off} respectively. While in the ''on'' state, the source produces cells at a fixed rate h. In our experiments we consider sources of two types, both of which have the same peak rate (4 Mbps) but different mean rates (0.8 Mbps and 0.89 Mbps). We consider three different values for T_{on} and T_{off} which, along with the other parameters of our simulation experiments, are shown in Table 5.2.

For the EB-ABR scheme, we assume that the link's operating point is such that the percentage of cells delayed in the shared buffer by more than 20 msec is less than 10^{-4}. This is equivalent to considering an overflow probability less than 10^{-4} in a shared buffer of approximately 7030 cells; the latter corresponds to a 20 msec delay in a 155 Mbps link. The link parameters s, t which appear in (5.10) are computed the same way as in Chapter 4, i.e., using the `supinf` formula (2.3), page 23. As we discussed in Section 2.1.1 and investigated in Section 2.3.3, particular pairs (s, t) will characterize different periods of the day during which the traffic mix remains relatively constant. The values of these parameters can be computed off-line using traffic traces taken during the corresponding period.

If m_i is the measured mean rate for connection i, then the explicit rate ER returned to that connection is given by (5.11), i.e., $ER = \widetilde{\alpha}^{-1}(m_i, x_i)$. Since we assume that all connections are paying the same price, the effective rate x_i of all connections is the same.

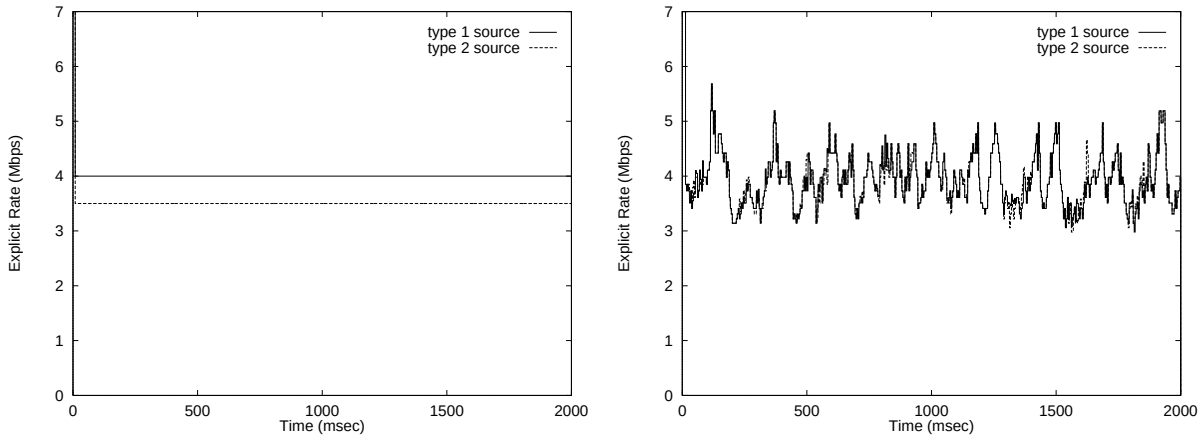
Figure 5.8 shows the explicit rate using the EB-ABR scheme and the ERICA congestion avoidance scheme. The latter detects congestion based on measurements of the aggregate load over a small time interval (measurement interval, Table 5.2) and supports a fair share policy [JKG+97]. Figure 5.8(b) shows that the explicit rate of connections fluctuates. This occurs because the ERICA scheme detects congestion based on measurements of aggregate load over a small time interval. Also observe that connections with the same peak rate but different mean rates are treated similarly. On the other hand, Figure 5.8(a) shows that the EB-ABR scheme remains the same for a long time interval. This occurs because the EB-ABR scheme shares capacity on a time scale much larger than the round trip delay. More important is the following: Because bandwidth is shared according to the effective bandwidth, which takes into account the mean rate, connections with a lower mean rate are entitled

a higher explicit rate. For example, in Figure 5.8(a)) the mean rate for type 1 sources is 0.8 Mbps, which is lower than the mean rate (0.89 Mbps) for type 2 sources, hence type 1 sources are entitled a higher explicit rate.

Figures 5.9 and 5.10 show the explicit rate when the sources are 2 and 4 times slower than the ones of Figure 5.8, while the peak and mean rates are kept the same. The explicit rates with the ERICA scheme are similar to the case of faster sources; the only difference is that they fluctuate less rapidly. Figure 5.10(a) shows that the explicit rates with EB-ABR are smaller than those of Figures 5.8(b) and 5.9(b). This is expected since it is more difficult to multiplex slow sources than it is to multiplex fast sources. Nevertheless the EB-ABR scheme achieves a high allocation of explicit rates when the ''on'' phase is 100 msec, which corresponds to bursts of 943 cells or 360 kb (the peak rate is 4 Mbps).

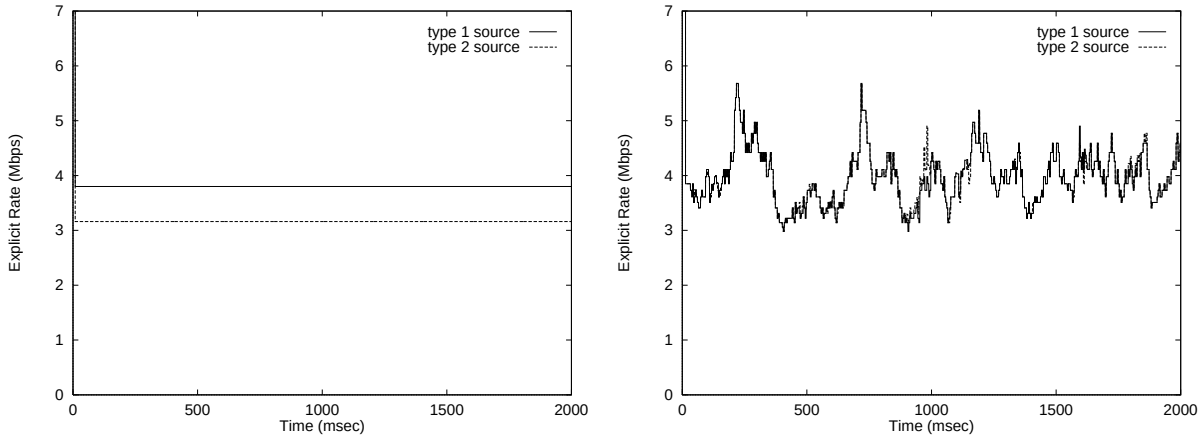
Network parameters	
Propagation delay	5000 μsec (1000 km)
Capacity	155 Mbps
Source parameters	
Peak h	4 Mbps
T_{on}	50 msec, 100 msec, 200 msec
T_{off} (type 1)	200 msec, 400 msec, 800 msec
T_{off} (type 2)	175 msec, 350 msec, 700 msec
Mean m (type 1)	0.8 Mbps
Mean m (type 2)	0.89 Mbps
RM interval	4 msec
Number of sources N	140
Percentage of type 2 sources	30%
EB-ABR parameters	
$P(\text{cell delay} > 20 \text{ msec}) < 10^{-4}$	
ERICA parameters	
Measurement interval	1 msec
Target utilization	90%

Table 5.2: **Simulation parameters for the resource sharing experiments.** The propagation delay corresponds to a wide area network. The average link utilization is approximately 0.86.



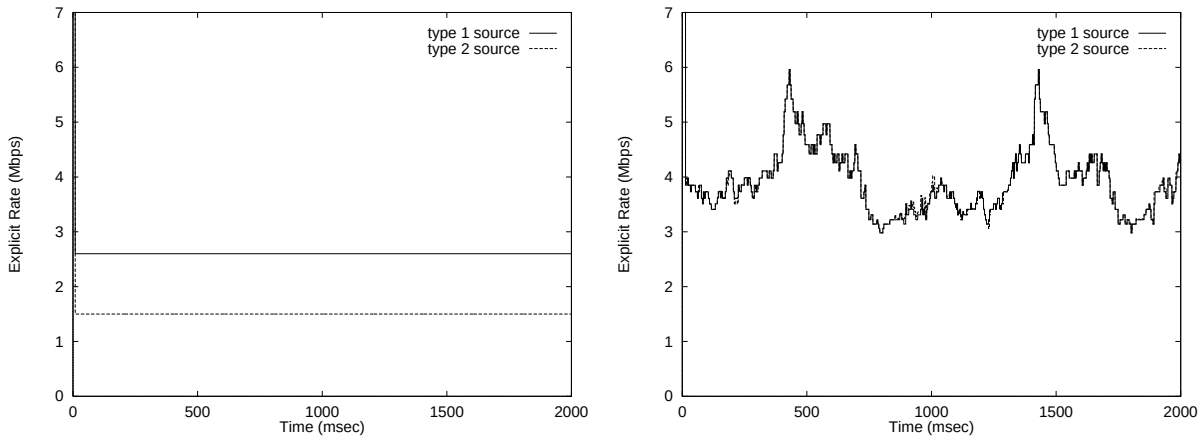
(a) EB-ABR (b) ERICA

Figure 5.8: **Explicit rate allocation results.** Type 1 source: $h = 4$ Mbps, $T_{\text{on}} = 50$ msec, $T_{\text{off}} = 200$ msec ($m = 0.8$ Mbps), Type 2 source: $T_{\text{on}} = 50$ msec, $T_{\text{off}} = 175$ msec ($m = 0.89$ Mbps). With the EB-ABR scheme, the source with the smaller mean rate receives a higher explicit rate. On the other hand, with the ERICA scheme both sources are treated the same.



(a) EB-ABR (b) ERICA

Figure 5.9: **Explicit rate allocation results.** Type 1 source: $h = 4$ Mbps, $T_{\text{on}} = 100$ msec, $T_{\text{off}} = 400$ msec ($m = 0.8$ Mbps), Type 2 source: $T_{\text{on}} = 100$ msec, $T_{\text{off}} = 350$ msec ($m = 0.89$ Mbps). With the EB-ABR scheme, the source with the smaller mean rate receives a higher explicit rate. On the other hand, with the ERICA scheme both sources are treated the same.



(a) EB-ABR (b) ERICA

Figure 5.10: **Explicit rate allocation results.** Type 1 source: $h = 4$ Mbps, $T_{\text{on}} = 200$ msec, $T_{\text{off}} = 800$ msec ($m = 0.8$ Mbps), Type 2 source: $h = 4$ Mbps, $T_{\text{on}} = 200$ msec, $T_{\text{off}} = 700$ msec ($m = 0.89$ Mbps). The explicit rates with the EB-ABR scheme are smaller than they were for the experiments shown in Figures 5.8 and 5.9. This is because the sources are less bursty and it is more difficult to multiplex such sources.

5.4 Conclusions

We have presented an approach to pricing and resource sharing for ABR services which has the following features:

- *Dynamic prices:* The price for network resources are adjusted according to the demand, hence can account for the dynamic nature of congestion which is anticipated for ABR services. The dynamic prices are not seen by the user, but are internal to the network.

- *Users declare the price per unit of time according to which they will be charged:* This price buys users a certain information transfer capability or effective rate.

- *Resources are shared based on effective bandwidths:* The information transfer capability (or effective rate) is interpreted as the effective bandwidth, which is controlled through the explicit rate that the network sends to the user. Hence, the approach allows bursty users to adequately reveal their preferences for network usage in terms of the price per unit of time they are charged.

We have discussed in detail how our approach can be implemented within the current ABR rate-based congestion control framework. Note, however, that the underlying ideas are independent of the specific congestion control mechanism. An issue for further research would be to investigate how these ideas can be implemented in, e.g., networks with credit-based flow control or networks based on the Internet Protocol (IP).

Finally, we have presented a comprehensive set of experiments which illustrate the convergence and transient behavior of dynamic prices, and which compare our effective bandwidth ABR (EB-ABR) resource sharing scheme with the ERICA (Explicit Rate Indication for Congestion Avoidance) scheme, demonstrating how our approach differentiates connections based on their mean rates. Such a feature of incentive compatibility is lacking in the existing flow control procedures where the network allocates in a fair way peak rates with no reference to the corresponding mean rates.

Our approach to resource sharing for ABR services operates at time scales much larger than the round trip delay in networks. This can be advantageous for wide area networks which span large distances. An issue for further investigation is to compare, for various propagation delays, the performance our approach with traditional congestion control schemes in terms of link utilization and buffer occupancy.

Chapter 6

Summary and Future Work

The objective of this dissertation was to apply and evaluate large deviation techniques for performance analysis and traffic engineering, to investigate usage-based pricing schemes for network transport services with open loop control, and to investigate pricing and resource sharing for Available Bit Rate (ABR) services, which support closed loop congestion control.

We applied and evaluated the many sources asymptotic for performance analysis and traffic engineering using MPEG-1 compressed video, Internet WAN traffic, and modeled voice traffic. Although the many sources asymptotic with its Bahadur-Rao improvement can overestimate the cell loss probability by 1-2 orders of magnitude, the utilization it achieves is very close to the maximum utilization. Hence, the effective bandwidth can accurately quantify resource usage. The actual cell loss probability differs from the overflow probability estimated using the many sources asymptotic and its Bahadur-Rao improvement because the latter is not a measure of the CLP, but a measure of the probability that in an infinite buffer the queue length becomes greater than some target value.

According to the theory of effective bandwidths and many sources asymptotic, a link's operating point is characterized by a pair of parameters, the space and time parameters. Our simulation results verify the interpretation of these parameters given by theory. In particular, the space parameter indicates the rate at which the logarithm of the overflow probability decreases (i.e., the QoS improves) when the buffer size increases, while the product of the space and time parameters indicates the rate at which the logarithm of the overflow probability decreases when the link capacity increases. The time parameter indicates the time scale where traffic control mechanisms,

such as smoothing, need to be studied in order to find their impact on resource usage. Moreover, the time parameter also indicates the granularity that traffic traces must have in order to capture the statistical information that is important for performance analysis. We showed how the values of these parameters can be used to clarify the effects on the link's performance of the time scales of traffic burstiness, of the traffic mix, of the link parameters (capacity and buffer), and of traffic control mechanisms, such as traffic smoothing. For the traffic mixes considered, the space and time parameters are, to a large extent, insensitive to small variations of the traffic mix. This indicates that particular pairs of these parameters can be used to characterize different periods of the day during which the traffic mix remains relatively constant. The values for these parameters for a particular period of the day can be computed off-line using traffic traces taken during that period.

We used the many sources asymptotic to simultaneously capture the cell scale and burst scale effects on the overflow probability at an ATM link with a small buffer which multiplexes a large number of periodic on-off sources. This allowed us to give the correct expression of the overflow probability for very small buffers and to investigate the boundary where cell scale effects stop to be important. In addition to accurately estimating the overflow probability, we give a new qualitative description of how overflow occurs in very small buffers.

We investigated usage-based pricing schemes for services with open loop control that are based on bounds of the effective bandwidth. The schemes included time-volume schemes that involve two measurements (the duration and transferred volume) for the whole duration of a connection, and schemes that involve measurements in distinct time intervals, smaller than the duration of a connection: pricing with renegotiation and the virtual bucket scheme. The pricing schemes were compared according to their fairness, i.e., their ability to capture the relative amount of resources used by connections. These comparisons were performed for link capacities and buffer sizes that will be used in broadband networks, and involved MPEG-1 compressed video with various contents (movies, news, and talk shows) and Internet WAN traffic.

Our experiments indicate that the fairness and robustness (with respect to small variations of the link's operating point parameters) of the time-volume pricing schemes are better for larger link capacities and buffer sizes, for smaller call durations, and for smoother (less bursty) traffic. Indeed, for high capacity links (over 622 Mbps) and small duration connections, the performance of these schemes approaches that of the schemes which involve measurements in distinct time intervals.

This indicates that for high bandwidth links with a high degree of statistical multiplexing, simple approximations can achieve good performance. Pricing schemes which involve measurements in distinct time intervals can achieve, at the cost of higher accounting overhead, higher fairness and robustness compared to the time-volume schemes which involve measurements for the whole duration of a connection. Furthermore, the performance of the virtual bucket scheme, which does not entail the signaling overhead due to renegotiation, was found to be very close to the performance of the pricing with renegotiation scheme. Whether the increased performance of these schemes outweighs the increased accounting overhead is an open issue. In addition to a detailed comparison of the aforementioned pricing schemes, we showed, for a particular setup, the incentive compatibility of the time-volume schemes. Furthermore, we compared the equilibrium operating point of a link, in the presence of pricing, for the case of deterministic and statistical multiplexing, and investigated to what extent it is affected by traffic shaping.

In Available Bit Rate (ABR) services, the user adjusts his traffic rate based on feedback signals he receives from the network. Since the network is the one who decides how much traffic a user can send, it is reasonable that pricing affects how the network shares its resources. Along these lines, we presented an approach to pricing and resource sharing for ABR services with the following three features: (i) prices are adjusted according to demand, (ii) users declare the price per unit of time according to which they will be charged, and (iii) resource sharing is based on effective bandwidths. The first feature captures the dynamic nature of congestion which is anticipated for ABR services, whereas the second and third features allow users to adequately reveal their preferences for network usage in terms of the price per unit of time they are paying, and enable the differentiation of connections based on their mean rates. Such a property of incentive compatibility is lacking in the existing flow control procedures where the network allocates in a fair way peak rates with no reference to the corresponding mean rates.

6.1 Future Work

Next we discuss some issues for future work. Some of these have already been identified in the conclusions section of the relevant chapters.

In the area of traffic engineering, an important issue is to apply our traffic engineering approach

in a real multi-service network environment, such as the one that is currently being developed in campus backbones which will support Internet, videoconference, and voice traffic.

The approach to characterize a link's operating point through the space and time parameters also opens up new possibilities for traffic modeling. Rather than developing general models that try to emulate real traffic in any operating environment, a new approach would be to develop models which emulate real traffic *for a particular operating point* of the network. Such models would be simple and efficient to implement, and can be the basis of fast and flexible traffic generators.

Regarding pricing, two issues that remain open are how to price multicast connections and how to price caches in the World Wide Web. Multicast has the potential of increasing network efficiency by allowing multiple receivers share the same delivery tree. The problem of pricing such connections is related to how one can assess the total amount of resources used by a multicast tree, how these are allocated to individual receivers, and what accounting mechanisms are required (e.g., see [HSE95]). Furthermore, the problem of pricing multicast connections is also related to the higher layer costs, such as the cost of content, which may outweigh the costs of network transport.

Caching is very important for improving the performance of the World Wide Web, and much research is being conducted on where to place and how to manage the information contained in the caches. A particular application is to use caching to decrease the load on congested, high priced international links. A direction for further work would be to apply economic modeling approaches to this problem, while taking into account its specific technological characteristics and the large amount of empirical results on caching methods.

Our approach to resource sharing for ABR services operates at time scales much larger than the round trip delay in networks. This can be advantageous in the case of wide area networks which span large distances. An item for further research would be to compare, for various propagation delays, the performance our approach in terms of link utilization and buffer occupancy, with that of traditional congestion control schemes.

Finally, our ideas on pricing ABR services are not limited to the specific mechanisms of ABR. An issue for further research would be to investigate how these ideas can be implemented in, e.g., networks with credit-based flow control or networks based on the Internet Protocol (IP). Indeed, many large institutions and organizations are building ATM backbones over which they will offer, in addition to Internet services, other services such as videoconferencing and voice services. An

important issue is to investigate how the additional mechanisms for network management and accounting that the ATM backbone provides can be used to control resource usage across different departments and user groups within an organization.

Appendix A

Traffic Sources

In this appendix we describe the traffic that was used in the numerical and simulation experiments of this dissertation. This traffic includes MPEG-1 compressed video sequences, Internet Wide Area Network (WAN) traces, and modeled voice traffic.

A.1 MPEG-1 Compressed Video Traffic

MPEG (Motion Pictures Expert Group) is an ISO (International Organization for Standardization) group working on standards for video and associated audio encoding. The first phase of its work was completed in 1991 with development of the MPEG-1 standard which was targeted at coding video and its associated audio at rates up to about 1.5 Mbps [Gal91]. The MPEG-2 standard is a superset of MPEG-1, with main differences being the support for interlaced pictures and layered encoding.

The traffic used in this dissertation was encoded using the MPEG-1 standard. An MPEG-1 compressed video consists of a sequence of frames. The number of frames per second (as well as the lines per second), depend on national standards. For Europe, there are 25 frames per second. There are three frame types, ''I'', ''P'', and ''B'', which differ in their encoding scheme. I (Interlaced) frames use intra-frame coding for reducing spatial redundancy. Intra-frame coding is based on the Discrete Cosine Transform (DCT) and the quantization of DCT coefficients. P (Predicted) frames use similar coding to I frames, with the addition of motion compensation for temporal redundancy. Motion compensation is performed with respect to the previous I or P frame. Finally, B frames are

similar to P frames, only that motion compensation can be done both with respect to the previous I or P frame, and the next I or P frame. Typically, the size of I frames is larger than the size of P frames, which in turn is larger than the size of B frames. The sequence of encoded frames follows a periodic pattern which is called Group Of Pictures (GOP) pattern.

The sequences used in this dissertation are shown in Table A.1. All of the movie sequences, except the sequence *Aliki in the Navy*, have been made available[1] by O. Rose [Ros95]. The frames were digitized at 384 x 288 pels and were encoding using the UC Berkeley MPEG-1 software encoder with the frame pattern IBBPBBPBBPBB. Each sequence contained 40000 frames (approximately 30 minutes), except for the *news2* sequence which contained 31515 frames.

The sequence *Aliki in the Navy*[2] used the same encoder input size (384 x 288) and also contained 40000 frames with the same frame pattern (IBBPBBPBBPBB), but had a smaller compression ratio. In particular, the quantization scale here was 3-4, whereas for the other sequences it was 10,14, and 18 (for the I, P, and B frames respectively). We used a Silicon Graphics O2 workstation equipped with an O2Video video processor for image capture and encoding. The image was captured from a VCR (VHS format) at 25 frames per second in motion JPEG format[3], and converted to MPEG-1 using the SGI Media Convert program.

The traffic used in our experiments was produced from the above MPEG-1 sequences by segmenting each frame into ATM cells, each containing a 48 byte payload. Cells belonging to the same frame are assumed to be evenly spaced within the frame, which had duration 40 msec (25 frames per second). Some statistics of the ATM cell streams are shown in Table A.2. The standard deviation is $\sigma = \sqrt{\mathrm{Var}\{X[0,T_\mathrm{f}]/T_\mathrm{f}\}}$, where $X[0,T_\mathrm{f}]$ is the number of cells in one frame time $T_\mathrm{f} = 40$ msec.

Three sets of video traffic were created for the numerical investigations in Section 4.5.1: *movies*, *news* and *talk shows*. These were created by breaking the cell streams containing the MPEG-1 traffic into non-overlapping segments, each with a duration of approximately 3 minutes (4500 frames). The resulting *movies* set (this did not include the *Aliki in the Navy* sequence) contained 54 segments, the *news* set contained 16 segments, and the *talk shows* set contained 18 segments.

[1]The sequences are available at <http://www.ftp-info3.informatik.uni-wuerzburg.de/pub/>.
[2]This segment was taken from the Greek film *Aliki in the Navy* (*Η Αλίκη στο Ναυτικό*).
[3]JPEG (Joint Photographic Experts Group) is an ISO standard for still image compression.

Movies	
bond	James Bond: Goldfinger
dino	Jurassic Park
lambs	The Silence of the Lambs
movieprev	Movie Preview
star	Star Wars
terminator	Terminator II
aliki	Aliki in the Navy
News	
news1	News
news2	News
Talk Shows	
talk1	Talk show
talk2	Talk show

Table A.1: **MPEG-1 sequences.**

Sequence	Mean (Mbps)	Peak (Mbps)	Peak/Mean	Std. dev. (Mbps)
bond	0.68	6.75	9.99	0.71
dino	0.37	3.31	9.03	0.41
lambs	0.21	3.71	17.92	0.31
movieprev	0.40	4.77	11.94	0.52
star	0.26	3.46	13.18	0.36
terminator	0.31	2.20	7.20	0.28
aliki	1.28	3.65	2.86	0.40
news1	0.43	5.25	12.23	0.54
news2	0.58	5.37	9.34	0.72
talk1	0.41	2.96	7.27	0.46
talk2	0.50	3.67	7.34	0.50

Table A.2: **Statistics of the cell streams created from the MPEG-1 sequences.**

A.2 Internet WAN Traffic

For Internet Wide Area Network (WAN) traffic we used the Bellcore Ethernet trace BC-Oct89Ext made available[4] by W. Leland and D. Wilson [LW91]. The trace had duration 122797.83 seconds.

For the numerical investigations in Section 4.5.2, a set of 15 non-overlapping segments, each with a duration of approximately two hours, was created from the original trace.

A.3 On-Off MMF Model for Voice Traffic

For voice traffic an on-off Markov Modulated Fluid (MMF) model was used. The latter is a continuous time two state Markov chain which produces traffic at a constant rate while in the ''on'' state, and does not produce traffic while in the ''off'' state. The time the system remains in the ''on'' and ''off'' states is exponentially distributed. For modeled voice traffic, the traffic rate in the ''on'' state is 64 Kbps and the average time spent in the ''on'' and ''off'' states are 352 msec and 650 msec, respectively [Bra69, HS92]. This gives an average rate of 22.48 Kbps.

[4]The trace is available at The Internet Traffic Archive, <http://www.acm.org/sigcomm/ITA>.

Appendix B

Details of the Numerical and Simulation Experiments

In this appendix we present some details of the numerical solution of the `supinf` formula (2.3) (Chapters 2-5), the trace-driven simulations of a single link (Chapters 2 and 4), and the network simulations for the ABR pricing and resource sharing experiments (Chapter 5).

B.1 Numerical Solution of the `supinf` Formula

The `supinf` formula (page 23) is given by

$$-I = \sup_t \inf_s \left[st \sum_{j=1}^{J} \rho_j \alpha_j(s,t) - s(ct+b) \right] , \tag{B.1}$$

where ρ_j is the percentage of sources of type j, c and b are the capacity and buffer per source, s, t are the space and time parameters, and $\alpha_j(s,t)$ is the effective bandwidth which is given by

$$\alpha_j(s,t) = \frac{1}{st} \log E\left[e^{s X_j[0,t]} \right] , \tag{B.2}$$

where $X_j[0,t]$ is the load produced by a source of type j in a time interval of length t. The numerical solution of the `supinf` formula (B.1) was implemented for traffic in the form of traces, i.e., number of cells produced in epochs of equal duration (equal to 40 msec for MPEG-1 traffic), and for traffic containing periodic on-off sources with deterministic ‘‘on’’ and ‘‘off’’ periods. For the latter, the effective bandwidth (B.2) can be calculated analytically (see Appendix C).

Buffer msec (cells)	Time in seconds	
	Enterprise 3000	Ultra-1
1 (351)	270	353
10 (3514)	440	549
30 (10542)	806	967
50 (17570)	1144	1563

Table B.1: **Average computation time for the numerical solution of the** `supinf` **formula.** The time column shows the sum of the user and system time given by the Unix `time` command. These values are for Sun's Ultra Enteprise 3000 server with two UltraSPARC processors at 170 MHz and Ultra-1 workstation with one UltraSPARC processor at 170 MHz. [$C = 155$ Mbps, MPEG-1 traffic with duration ≈ 30 minutes]

In the case of traces, the expectation in (B.2) is replaced by the empirical average

$$\alpha_j(s,t) = \frac{1}{st}\log\left[\frac{1}{T/t}\sum_{i=1}^{T/t} e^{sX_j[(i-1)t,it]}\right], \tag{B.3}$$

where T is the duration of the trace.

The search for the extremizing t is done for values of t with a granularity equal to one frame time. This was not a limitation since the cells in a single frame were assumed to be evenly spaced throughout the duration of the frame. The search window was selected empirically and used the observation that, for the same capacity, the extremizing value of t was larger for larger buffer sizes. On the other hand, the search for the extremizing value of s was performed in a window whose size and granularity Δ_s was dynamically rescaled. This allowed the value of s to be determined with a specified accuracy, e.g., $Y\%$, by requiring, when the extremizing value s is found, that the ratio Δ_s/s was smaller than $Y/100$. The average computation time required to solve the `supinf` formula is shown in Table B.1.

The experiments in Chapter 2 were of two types:

1. *Fixed utilization:* In these experiments the utilization, or equivalently the number of sources, was kept constant and equations (B.1), (B.3) were solved once to find the values of s, t and the overflow probability e^{-NI}, or $e^{-NI-\frac{1}{2}\log(4\pi NI)}$ for the many sources asymptotic with the Bahadur-Rao improvement, where N is the number of sources.

2. *Fixed overflow probability:* In these experiments there was a target overflow probability, say $e^{-\gamma}$. The number of multiplexed sources was increased in order to find the maximum number

such that the overflow probability e^{-NI}, or $e^{-NI-\frac{1}{2}\log(4\pi NI)}$ for the many sources asymptotic with the Bahadur-Rao improvement, was less than the target $e^{-\gamma}$. As before, the rate function I is computed using (B.1), (B.3).

B.2 Simulation of a Single Link

The simulation model for a single link was a discrete time simulator, with an epoch equal to one frame time (equal to 40 msec for the case of MPEG-1 traffic). We assumed that the input traffic was aligned on frame boundaries. The starting frame for each source was randomly chosen; when the end of the trace was reached, the frame pointer was advanced to the first frame of the trace.

The Cell Loss Probability (CLP) was computed using

$$\text{CLP} = \frac{\text{\# of lost cells}}{\text{total \# of cells}}, \tag{B.4}$$

and the frame overflow probability was computed using

$$P(\text{frame overflow}) = \frac{\text{\# of frames with at least one cell loss}}{\text{total \# of frames}}. \tag{B.5}$$

The results reported are the average from a total of 100 independent simulation runs, each with a random selection of the starting frame for every source. Each run had duration 5 times the size of the trace. Both the number of independent runs and the duration of each run were chosen empirically.

For the experiments with a fixed overflow probability, similar to the numerical solution of the `supinf` formula, the number of multiplexed sources was increased in order to find the maximum which satisfied the target overflow probability.

B.3 Network Simulation for the ABR Pricing and Resource Sharing Experiments

The ABR pricing and resource sharing simulations used a discrete event fluid simulation, where events are not the arrival/departure of cells but the change of cell rates, the arrival/departure of resource management (RM) cells, and the event of a buffer becoming empty. Such a model results in much smaller simulation times compared to a cell level simulation, and provides sufficient

detail for studying the convergence behavior of prices (Section 5.2.2) and the Explicit Rate (ER) allocations for the ABR resource sharing experiments (Section 5.3.2).

Appendix C

Logarithmic Moment Generating Function for Periodic On-Off Sources

In this section we derive analytical expressions for the logarithmic moment generating function $\varphi(s,t) = \log E[e^{sX[0,t]}]$ for periodic on-off sources. The peak rate of the source is h, measured in cells per unit of time, and $\tau = 1/h$. The duration of the ''on'' and ''off'' periods is T_{on} and T_{off}, respectively, and $T_{\text{total}} = T_{\text{on}} + T_{\text{off}}$. The maximum number of back-to-back cells is $M = T_{\text{on}}h + 1$. We restrict ourselves to the case $T_{\text{on}} < T_{\text{off}}$.

Discrete form

First we compute the discrete form of the logarithmic moment generating function assuming that $t < T_{\text{total}}$.

Case: $t < \tau$

$$\begin{aligned}\varphi(s,t) &= \log E\left[e^{sX[0,t]}\right] \\ &= \log\left[e^s \frac{Mt}{T_{\text{total}}} + \frac{(M-1)(\tau - t) + T_{\text{off}} - t}{T_{\text{total}}}\right]\end{aligned}$$

Case: $k\tau \le t < (k+1)\tau < T_{\text{on}}, T_{\text{off}},$ where $k = 1, 2, 3, \ldots$

$$\begin{aligned}\varphi(s,t) = \log\Bigg[&\frac{T_{\text{off}} - t}{T_{\text{total}}} + \frac{2\tau}{T_{\text{total}}} \sum_{i=1}^{k} e^{is} \\ &+ e^{(k+1)s} \frac{(M-k)(t - k\tau)}{T_{\text{total}}} + e^{ks} \frac{(M-k-1)((k+1)\tau - t)}{T_{\text{total}}}\Bigg]\end{aligned}$$

Case: $T_{\text{on}} \leq t < T_{\text{off}}$

$$\varphi(s,t) = \log\left[\frac{T_{\text{off}}-t}{T_{\text{total}}} + \frac{2\tau}{T_{\text{total}}}\sum_{i=1}^{M-1} e^{is} + e^{Ms}\frac{t-T_{\text{on}}}{T_{\text{total}}}\right]$$

Case: $T_{\text{off}} \leq t < T_{\text{total}}$ and $k\tau \leq t - T_{\text{off}} < (k+1)\tau$

$$\varphi(s,t) = \log\left[\frac{2\tau}{T_{\text{total}}}\sum_{i=k+2}^{M-1} e^{is} + e^{Ms}\frac{t-T_{\text{on}}}{T_{\text{total}}} + e^{(k+1)s}\frac{(k+2)((k+1)\tau-(t-T_{\text{off}}))}{T_{\text{total}}} + e^{(k+2)s}\frac{(k+1)((t-T_{\text{off}})-k\tau)}{T_{\text{total}}}\right]$$

Now we generalize the above formulae to include the case where $t \geq T_{\text{total}}$. In what follows, l is such that $lT_{\text{total}} \leq t < (l+1)T_{\text{total}}$ and $t' = t - lT_{\text{total}}$.

Case: $t' < \tau$

$$\varphi(s,t) = \log\left[e^{(lM+1)s}\frac{Mt'}{T_{\text{total}}} + e^{lMs}\frac{(M-1)(\tau-t')+T_{\text{off}}-t'}{T_{\text{total}}}\right]$$

Case: $k\tau \leq t' < (k+1)\tau < T_{\text{on}} < T_{\text{off}}$

$$\varphi(s,t) = \log\left[\frac{T_{\text{off}}-t'}{T_{\text{total}}} + \frac{2\tau}{T_{\text{total}}}\sum_{i=lM+1}^{lM+k} e^{is} + e^{(lM+k+1)s}\frac{(M-k)(t'-k\tau)}{T_{\text{total}}} + e^{(lM+k)s}\frac{(M-k-1)((k+1)\tau-t')}{T_{\text{total}}}\right]$$

Case: $T_{\text{on}} \leq t' < T_{\text{off}}$

$$\varphi(s,t) = \log\left[\frac{T_{\text{off}}-t'}{T_{\text{total}}} + \frac{2\tau}{T_{\text{total}}}\sum_{i=lM+1}^{lM+M-1} e^{is} + e^{(lM+M)s}\frac{t'-T_{\text{on}}}{T_{\text{total}}}\right]$$

Case: $T_{\text{off}} \leq t' < T_{\text{total}}, k\tau \leq t' - T_{\text{off}} < (k+1)\tau$

$$\varphi(s,t) = \log\left[\frac{2\tau}{T_{\text{total}}}\sum_{i=lM+k+2}^{lM+M-1} e^{is} + e^{(lM+M)s}\frac{t'-T_{\text{on}}}{T_{\text{total}}} + e^{(lM+k+1)s}\frac{(k+2)((k+1)\tau-(t'-T_{\text{off}}))}{T_{\text{total}}} + e^{(lM+k+2)s}\frac{(k+1)((t'-T_{\text{off}})-k\tau)}{T_{\text{total}}}\right]$$

Fluid approximation

Next we derive the formula for $\varphi(s,t)$ assuming that the information flow is a continuous fluid flow. This captures only the burst scale component of buffer overflow. We first consider the case $t < T_{\text{total}}$.

Case: $t < T_{\text{on}}$

$$\varphi(s,t) = \log\left[e^{sht}\frac{T_{\text{on}}-t}{T_{\text{total}}} + \frac{T_{\text{off}}-t}{T_{\text{total}}} + \int_0^t e^{shx}\frac{2dx}{T_{\text{total}}}\right]$$

Case: $T_{\text{on}} \le t < T_{\text{off}}$

$$\varphi(s,t) = \log\left[e^{shT_{\text{on}}}\frac{t-T_{\text{on}}}{T_{\text{total}}} + \frac{T_{\text{off}}-t}{T_{\text{total}}} + \int_0^{T_{\text{on}}} e^{shx}\frac{2dx}{T_{\text{total}}}\right]$$

Case: $T_{\text{off}} \le t < T_{\text{total}}$

$$\varphi(s,t) = \log\left[e^{shT_{\text{on}}}\frac{t-T_{\text{on}}}{T_{\text{total}}} + e^{s(t-T_{\text{off}})h}\frac{t-T_{\text{off}}}{T_{\text{total}}} + \int_{T_{\text{on}}}^{t-T_{\text{off}}} e^{shx}\frac{2dx}{T_{\text{total}}}\right]$$

For the general case where $t > T_{\text{total}}$, we set $t' = t - lT_{\text{total}}$, where l is the integer satisfying $0 \le t - lT_{\text{total}} < T_{\text{total}}$.

Case: $t' < T_{\text{on}}$

$$\varphi(s,t) = \log\left[e^{shlT_{\text{on}}}\left(e^{sht'}\frac{T_{\text{on}}-t'}{T_{\text{total}}} + \frac{T_{\text{off}}-t'}{T_{\text{total}}} + \int_0^t e^{shx}\frac{2dx}{T_{\text{total}}}\right)\right]$$

Case: $T_{\text{on}} \le t' < T_{\text{off}}$

$$\varphi(s,t) = \log\left[e^{shlT_{\text{on}}}\left(e^{shT_{\text{on}}}\frac{t'-T_{\text{on}}}{T_{\text{total}}} + \frac{T_{\text{off}}-t'}{T_{\text{total}}} + \int_0^{T_{\text{on}}} e^{shx}\frac{2dx}{T_{\text{total}}}\right)\right]$$

Case: $T_{\text{off}} \le t' < T_{\text{total}}$

$$\varphi(s,t) = \log\left[e^{shlT_{\text{on}}}\left(e^{shT_{\text{on}}}\frac{t'-T_{\text{on}}}{T_{\text{total}}} + e^{s(t'-T_{\text{off}})h}\frac{t'-T_{\text{off}}}{T_{\text{total}}} + \int_{T_{\text{on}}}^{t'-T_{\text{off}}} e^{shx}\frac{2dx}{T_{\text{total}}}\right)\right]$$

Bibliography

[ATMF94] The ATM Forum. *ATM User-Network Interface (UNI) Specification Version 3.1.* AF-UNI-0010.002, 1994. Also published by Prentice Hall.

[ATMF96] The ATM Forum. *Traffic Management Specification Version 4.0.* AF-TM-0056.000, April 1996.

[BCS94] R. Braden, D. Clark, and S. Shenker. Integrated services in the Internet architecture: an overview. RFC 1633. ISI, MIT, and Xerox PARC, June 1994.

[BD95] D. D. Botvich and N. G. Duffield. Large deviations, the shape of the loss curve, and economies of scale in large multiplexers. *Queueing Systems*, 20:293--320, 1995.

[BR97] A. Barrett and K. Rebello. AOL: For $ 19.95 A Month, Unlimited Headaches For AOL. *Dataquest*, February 28, 1997.

[Bra69] P. T. Brady. A model for generating ON-OFF speech patterns in two-way conversations. *Bell Syst. Tech. J.*, 48, September 1969.

[BSTW95] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger. Long-rage dependence on variable-bit-rate video traffic. *IEEE Trans. Commun.*, 43(2/3/4):1566--1578, February/March/April 1995.

[CB96] M. E. Crovella and A. Bestavros. Self-similarity in World Wide Web traffic, evidence and possible causes. In *Proc. of ACM SIGMETRICS'96*, pages 160--169, Philadelphia, PA, USA, May 1996.

[CESZ91] R. Cocchi, D. Estrin, S. Shenker, and L. Zhang. A study of priority pricing in multiple service class networks. In *Proc. of ACM SIGCOMM'91*, pages 123--130, Zürich, Switzerland, September 1991.

[CKR+95] C. Courcoubetis, G. Kesidis, A. Ridder, J. Walrand, and R.R. Weber. Admission control and routing in ATM networks using inferences from measured buffer occupancy. *IEEE Trans. Commun.*, 43:1778--1784, 1995.

[CKS+97] C. Courcoubetis, F. P. Kelly, V. A. Siris, G. D. Stamoulis, and R. Weber. ABR pricing experiments in a real network. In *Proc. of International Conference on Telecommunications (ICT'97)*, Melbourne, Australia, April 1997.

[CKSW98] C. Courcoubetis, F. P. Kelly, V. A. Siris, and R. Weber. A study of simple usage-based charging schemes for broadband networks. In *Proc. of IFIP International Conference on Broadband Communications (BC'98)*, Stuttgart, Germany, April 1998. To appear.

[CKW97] C. Courcoubetis, F. P. Kelly, and R. Weber. Measurement-based charging in communications networks. Technical Report 1997-19, Statistical Laboratory, University of Cambridge, 1997.

[Cla96a] D. D. Clark. Adding service discrimination to the Internet. *Telecommunications Policy*, 20:169--181, 1996.

[Cla96b] D. D. Clark. A model for cost allocation and pricing in the Internet. In L. W. McKnight and J. P. Bailey, editors, *Internet Economics*. MIT Press, Massachusetts, 1996.

[CLS96] T. M. Chen, S. S. Liu, and V. K. Samalam. The available bit rate service for data in ATM networks. *IEEE Commun. Mag.*, pages 56--70, May 1996.

[CLW94] G. L. Choudhury, D. M. Lucantoni, and W. Whitt. On the effectiveness of effective bandwidths for admission control in ATM networks. In *Proc. of the 14th International Teletraffic Congress (ITC - 14)*, pages 411--420, North Holland, 1994. Elsevier Science B. V.

[CRL96] A. Charny, K. K. Ramakrishnan, and A. Lauck. Time scale analysis and scalability issues for explicit rate allocation in ATM networks. *IEEE/ACM Trans. on Networking*, 4(4):569--580, August 1996.

[CS97] C. Courcoubetis and V. A. Siris. An approach to pricing and resource sharing for Available Bit Rate (ABR) services. Technical Report No. 212, ICS-FORTH, November 1997. Also submitted for publication.

[CS98] C. Courcoubetis and V. A. Siris. An evaluation of pricing schemes that are based on effective usage. In *Proc. of IEEE International Conference on Communications (ICC'98)*, Atlanta, Georgia, USA, June 1998. To appear.

[CSEZ93] R. Cocchi, S. Shenker, D. Estrin, and L. Zhang. Pricing in computer networks: Motivation, formulation, and examples. *IEEE/ACM Trans. on Networking*, 1(6):614--627, November 1993.

[CSS96] C. Courcoubetis, V. A. Siris, and G. D. Stamoulis. Integration of pricing and flow control for available bit rate services in ATM networks. In *Proc. of IEEE GLOBECOM'96*, pages 644--648, London, UK, November 1996.

[CSS97a] C. Courcoubetis, V. A. Siris, and G. D. Stamoulis. Many sources asymptotic and effective bandwidths: Investigation with MPEG traffic. Presented at the *2nd IFIP Workshop on Traffic Management and Synthesis of ATM Networks*, Montreal, Canada, September 1997. Extended version submitted for publication.

[CSS97b] C. Courcoubetis, V. A. Siris, and G. D. Stamoulis. Application and evaluation of large deviation techniques for traffic engineering in broadband networks. Submitted for publication, October 1997.

[CSSM95] C. Courcoubetis, V. A. Siris, G. D. Stamoulis, and Y. Markopoulos. Charging issues for services in broadband networks. In *OECD Conference on the Economics of the Information Society*, Istanbul, Turkey, December 1995.

[CSW97] C. Courcoubetis, V. A. Siris, and R. Weber. Investigation of cell scale and burst scale effects on the cell loss probability using large deviations. In *Proc. of the 15th UK Workshop on Performance Engineering of Computer and Communication Systems (UKPEW'97)*, Ilkley, UK, July 1997.

[CW95] C. Courcoubetis and R. Weber. Effective bandwidths for stationary sources. *Prob. Eng. Inf. Sci.*, 9:285--296, 1995.

[CW96] C. Courcoubetis and R. Weber. Buffer overflow asymptotics for a switch handling many traffic sources. *Journal of Applied Probability*, 33:886-903, 1996.

[CWSA95] J. Crowcroft, Z. Wang, A. Smith, and J. Adams. A rough comparison of the IETF and ATM service models. *IEEE Network*, pages 12--16, November/December 1995.

[DLO95] N. G. Duffield, J. T. Lewis, and N. O'Connell. Predicting quality of service for traffic with long-range fluctuations. In *Proc. of IEEE International Conference on Communications (ICC'95)*, pages 473--477, Seattle, WA, USA, September 1995.

[DO96] N. G. Duffied and N. O'Connell. Large deviations and overflow probabilities for the general single-server queue, with applications. In *Math. Proc. Camb. Phil. Soc.*, 1996.

[DRS91] L. G. Dron, G. Ramamurthy, and B. Sengupta. Delay analysis of continuous bit rate traffic over an ATM network. *IEEE J. Select. Areas Commun.*, 9(3):402--407, April 1991.

[dVB95] G. de Veciana and R. Baldick. Resource allocation in multi-service networks via pricing. Technical report, University of Texas, Austin, January 1995.

[dVW95] G. de Veciana and J. Walrand. Effective bandwidths: Call admission, traffic policing and filtering in ATM networks. *Queueing Systems*, 20:37--59, 1995.

[EHL+95] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra, and A. Weiss. Fundamental bounds and approximations for ATM multiplexers with application to video teleconferencing. *IEEE J. Select. Areas Commun.*, 13(6):1004--1016, August 1995.

[EM93] A. Elwalid and D. Mitra. Effective bandwidth of general Markovian traffic sources and admission control of high speed networks. *IEEE/ACM Trans. on Networking*, 1(3):329--343, October 1993.

[EMV95] R.J. Edell, N. Mckeown, and P.P. Varaiya. Billing users and pricing for TCP. *IEEE J. Select. Areas in Commun.*, 13(7):1162--1175, September 1995.

[ENW96] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Trans. on Networking*, 4(2):209--223, April 1996.

[Fau95] G. R. Faulhaber. Pricing the Net: What economists do. Presented at the *Stanford mini-conference on data network pricing*, 1995. Available at <http://rider.wharton.upenn.edu/ faulhabe/>.

[Fir97] E. Firdman. Rx for the Internet: Usage-Based Pricing. *Data Communications*, January 1997. Available at <http://www.data.com/business_case/rx_internet.html>.

[FL91] H. Fowler and W. Leland. Local area network traffic characteristics, with implications for broadband congestion management. *IEEE J. Select. Areas Commun.*, 9(7):1139--1149, September 1991.

[FLVO94] G. Fiche, W. Lorcher, R. Veyland, and F. Oger. Study of multiplexing for ATM traffic sources. In *Proc. of the 14th International Teletraffic Congress (ITC - 14)*, pages 441--452, North Holland, 1994. Elsevier Science B. V.

[FYN88] D. Ferguson, Y. Yemini, and C. Nikolaou,. Microeconomic Algorithms for Load Balancing in Distributed Computer Systems In *Proc. of International Conference on Distributed Systems (ICDCS'88)*, pages 491--499, 1988.

[Gal91] D. Le Gall. MPEG: A video compression standard for multimedia applications. *Communications of the ACM*, 34(4):46--58, April 1991.

[GAN91] R. Geurin, H. Ahmadi, and M. Naghshineh. Equivalent capacity and its application to bandwidth allocation in high-speed networks. *IEEE J. Select. Areas Commun.*, 9(7):968--981, September 1991.

[GB96] M. Grossglauser and J-C. Bolot. On the relevance of long-range dependence in network traffic. In *Proc. of ACM SIGCOMM'96*, pages 15--24, Stanford, CA, USA, August 1996.

[GH91] R. J. Gibbens and P. J. Hunt. Effective bandwidths for multi-type UAS channel. *Queueing Systems*, 9:17--28, 1991.

[GSW94a] A. Gupta, D. O. Stahl, and A. B. Whinston. An economic approach to networked computing with priority classes. Technical report, University of Texas, Austin, December 1994.

[GSW94b] A. Gupta, D. O. Stahl, and A. B. Whinston. Managing the Internet as an economical system. Technical report, University of Texas, Austin, July 1994.

[GTR96] Global telecommunications reform. *Data Communications*, September 21, 1996.

[GW94] M. W. Garrett and W. Willinger. Analysis, modeling, and generation of self-similar VBR video traffic. In *Proc. of ACM SIGCOMM'94*, pages 269--280, London, UK, August 1994.

[HS92] I. W. Habib and T. N. Saadawi. Multimedia traffic characteristics in broadband networks. *IEEE Commun. Mag.*, pages 48--54, July 1992.

[HS95] M. L. Honig and K. Steiglitz. Usage-based Pricing of Packet Data Generated by a Heterogeneous User Population. In *Proc. of IEEE INFOCOM'95*, pages 867--874, Boston, MA, USA, April 1995.

[HSE95] S. Herzog, S. Shenker, and D. Estrin. Sharing the "cost" of multicast trees:an axiomatic analysis. In *Proc. of ACM SIGCOMM'95*, pages 315--327, Cambridge, MA, USA, August 1995.

[Hui88] J. Y. Hui. Resource allocation for broadband networks. *IEEE J. Select. Areas Commun.*, 6:1598--1608, 1988.

[HW94] I. Hsu and J. Walrand. Admission control for ATM networks. In *IMA Workshop on Stochastic Networks*. Springler-Verlag, 1994.

[INDEX] INDEX: The Internet Demand Experiment. Departments of Electrical Engineering & Computer Science and Economics, University of California at Berkeley, CA. URL: <http://www.INDEX.Berkeley.EDU/public/>.

[I.321] International Telecommunication Union Telecommunications Sector (ITU-T) Recommendation I.321. B-ISDN protocol reference model and its application. Geneva, Switzerland, 1991.

[I.371] International Telecommunication Union Telecommunications Sector (ITU-T) Recommendation I.371. Traffic control and congestion control in B-ISDN. Geneva, Switzerland, November 1995.

[JKG+97] R. Jain, S. Kalyanaraman, R. Goyal, S. Fahmy, and R. Vismanathan. The ERICA switch algorithm for ABR traffic management in ATM networks: Part I: Description. The Ohio State University, Department of CIS, January 1997.

[JLS97] P. R. Jelenković, A. A. Lazar, and N. Semret. The effect of multiple time scales and subexponentiality in MPEG video streams on queueing behavior. *IEEE J. Select. Areas in Commun.*, 15(6):1052--1071, August 1997.

[Kel91] F. P. Kelly. Effective bandwidths at multi-class queues. *Queueing Systems*, 9:5--16, 1991.

[Kel94] F. P. Kelly. On tariffs, policing and admission control for multiservice networks. *Operations Research Letters*, 15:1--9, 1994.

[Kel96a] F. P. Kelly. Notes on effective bandwidths. In F. P. Kelly, S. Zachary, and I. Zeidins, editors, *Stochastic Networks: Theory and Applications*, pages 141--168. Oxford University Press, 1996.

[Kel96b] F.P. Kelly. Charging and accounting for bursty connections. In J. P. Bailey and L. Mcknight, editors, *Internet Economics*, Massachusetts, 1996. MIT Press.

[Kel97] F. P. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33--37, January 1997.

[KM94] R. Keinath and D. Minoli. *Distributed Multimedia through Broadband Communications.* Artech House, Boston, 1994.

[KMT98] F. P. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49, 1998

[Kni97] E. Knightly. Second moment resource allocation in multi-service networks. In *Proc. of ACM SIGMETRICS'97*, pages 181--191, Seattle, WA, USA, 1997.

[KSV96] M. Katevenis, D. Serpanos, and P. Vatsolaki. ATLAS I: A general-purpose, single-chip ATM switch with credit-based flow control. In *IEEE Hot Interconnects IV Symposium*, Stanford, CA, USA, August 1996.

[KVR95] L. Kalampoukas, A. Varma, and K. K. Ramakrishnan. An efficient rate allocation algorithm for ATM networks providing max-min fairness. In *Proc. of 6th IFIP International Conference on High Performance Networking (HPN'95)*, September 1995.

[KWC93] G. Kesidis, J. Walrand, and C.-S. Chang. Effective bandwidths for multiclass Markov fluids and other ATM sources. *IEEE/ACM Trans. on Networking*, 1(3):424--428, October 1993.

[LM97] N. Likhanov and R. R. Mazumdar. Cell loss asymptotics for buffers fed with a large number of independent stationary sources. Preprint. To appear in *Proc. of IEEE INFOCOM'98*, San Francisco, USA, March 1998.

[Low97] S. H. Low. Equilibrium bandwidth and buffer allocations for elastic traffics. Preprint, May 1997. Partial and preliminary results of this paper have been presented in the *International Conference on Telecommunications (ICT'97)*, Melbourne, Australia, April 1997.

[LTWW93] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. ''On the Self-Similar Nature of Ethernet Traffic''. In *Proc. of ACM SIGCOMM'93*, pages 183--193, Ithaca, NY, USA, September 1993.

[LV93] S. H. Low and P. P. Varaiya. A new approach to service provisioning in ATM networks. *IEEE/ACM Trans. on Networking*, 1(3):547--553, October 1993.

[LW91] W. E. Leland and D. V. Wilson. High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection. In *Proc. of IEEE INFOCOM'91*, pages 1360--1366, Bal Harbour, FL, USA, April 1991.

[MdV96] M. Montgomery and G. de Veciana. On the relevance of time scales in performance oriented traffic characterizations. In *Proc. of IEEE INFOCOM'96*, pages 513--520, San Francisco, CA, USA, March 1996.

[MMP94] J. Murphy, L. Murphy, and E. C. Posner. Distributed pricing for embedded ATM networks. In *Proc. of the 14th International Teletraffic Congress (ITC - 14)*, pages 1053--1063, North Holland, 1994. Elsevier Science B. V.

[MMV94] J. K. Mackie-Mason and H. R. Varian. Some FAQs about Usage-Based Pricing. Available at <http://www.spp.umich.edu/ipps/papers/info-nets/useFAQs/useFAQs.html>, September 1994.

[MMV95a] J. K. Mackie-Mason and H. R. Varian. Economic FAQ About the Internet. *Journal of Economic Perspectives*, June 1995. Available at <http://www.spp.umich.edu/ipps/papers/info-nets/Economic_FAQs/FAQs/FAQs.html>.

[MMV95b] J. K. Mackie-Mason and H. R. Varian. Pricing congestible network resources. *IEEE J. Select. Areas in Commun.*, 13(7):1141--1149, September 1995.

[MMV95c] J. K. Mackie-Mason and H. R. Varian. Pricing the Internet. In B. Kahin and J. Keller, editors, *Public Access to the Internet*. Prentice Hall, Englewood Cliffs, NJ, 1995.

[MV91] B.M. Mitchell and I. Vogelsang. *Telecommunications pricing: Theory and Practice*. Cambridge University Press, 1991.

[NRSV91] I. Norros, J. W. Roberts, A. Simonian, and J. T. Virtamo. The superposition of variable bit rate sources in an ATM multiplexer. *IEEE J. Select. Areas Commun.*, 9(3):378--387, April 1991.

[Odl97] A. Odlyzko. A modest proposal for preventing Internet congestion. Preprint, September 1997. AT&T Labs - Research. Available at <http://www.research.att.com/ amo>.

[Onv94] R. O. Onvural. *Asynchronous Transfer Mode Networks: Performance Issues*. Artech House, Boston, 1994.

[OS97] A. Orda and N. Shimkin. Incentive pricing in mutli-class communication networks. In *Proc. of IEEE INFOCOM'97*, Kobe, Japan, April 1997.

[PF92] C. Parris and D. Ferrari. A resource based pricing policy for real-time channels in a packet-switching network. Technical report, International Computer Science Institute, Berkeley, CA, 1992.

[PF95] V. Paxson and S. Floyd. Wide-area traffic: The failure of Poisson modeling. *IEEE/ACM Trans. on Networking*, 3(3):226--244, June 1995.

[PKF92] C. Parris, S. Keshav, and D. Ferrari. A framework for the study of pricing in integrated networks. Technical Report TR-92-016, International Computer Science Institute, Berkeley, CA, March 1992.

[Q.2963] International Telecommunication Union Telecommunications Sector (ITU-T) Draft Recommendation Q.2963. Geneva, Switzerland, 1991.

[RE96] B. K. Ryu and A. Elwalid. The importance of the long-range dependence of VBR video traffic in ATM traffic engineering: Myths and realities. In *Proc. of ACM SIGCOMM'96*, pages 3--14, Stanford, CA, USA, August 1996.

[Rob91] J. W. Roberts. Variable-bit-rate traffic control in B-ISDN. *IEEE Commun. Mag.*, pages 50--56, April 1991.

[Ros95] O. Rose. Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems. Technical Report 101, University of Wuerzburg, February 1995.

[RSKJ91] C. Rasmussen, J. H. Sorensen, K. S. Kvols, and S. B. Jacobsen. Source-independent call acceptance procedures in ATM networks. *IEEE J. Select. Areas Commun.*, 9(3):351--358, April 1991.

[San88] B. A. Sanders. An asynchronous, distributed flow control algorithm for rate allocation in computer networks. *Trans. on Computers*, 37:779--787, July 1988.

[SCEH96] S. Shenker, D. Clark, D. Estrin, and S. Herzog. Pricing in computer networks: Reshaping the research agenda. *ACM Computer Communication Review*, 26(2):19--43, 1996.

[SFY95] J. Sairamesh, D. F. Ferguson, and Y. Yemini. An approach to pricing, optimal allocation and quality of service provisioning in high-speed packet networks. In *Proc. of IEEE INFOCOM'95*, pages 1111--1119, Boston, MA, USA, April 1995.

[SG95] A. Simonian and J. Guibert. Large deviations approximations for fluid queues fed by a large number of on-off sources. *IEEE J. Select. Areas Commun.*, 13(7):1017--1027, August 1995.

[She95a] S. Shenker. Fundamental design issues for the future Internet. *IEEE J. Select. Areas Commun.*, 13(7):1276--1188, September 1995.

[She95b] S. Shenker. Service models and pricing policies for an integrated services Internet. In B. Kahin and J. Keller, editors, *Public Access to the Internet*. Prentice Hall, Englewood Cliffs, NJ, 1995.

[SPG97] S. Shenker, C. Partridge, and R. Guerin. Specification of guaranteed quality of service. Draft RFC, Integrated Services Working Group: draft-ietf-intserv-guaranteed-svc-08.txt. Xerox/BBN/IBM, February 1997.

[SW95] A. Shwartz and A. Weiss. *Large Deviations for Performance Analysis*. Chapman and Hall, NY, 1995.

[TMW97] K. Thompson, G. J. Miller, and R. Wilder. Wide-area Internet traffic patterns and characteristics. *IEEE Network*, November/December 1997.

[Tur86] J. Turner. New directions in communications (or which way to the information age?). *IEEE Commun. Mag.*, pages 8--15, October 1986.

[UUNET97] UUNET Technologies. Burstable T-3 service pricing, 1997. Available at <http://www.us.uu.net/html/t-3.html>.

[Var92] Hal R. Varian. *Microeconomic Analysis*. W. W. Norton & Company Inc., 1992.

[Var96] H. R. Varian. Economic issues facing the Internet. School of Information Management and Systems, University of California at Berkeley, June 1996. Available at <http://www.sims.berkeley.edu/ hal/>.

[Wei95] A. Weiss. An introduction to large deviations for communication networks. *IEEE J. Select. Areas Commun.*, 13(6):938--952, August 1995.

[Wex96] J. Wexler. Usage-based 'Net access pricing in vogue. *Network World*, March 4, 1996.

[WHH+92] C. A. Waldspurger, T. Hogg, B. A. Huberman, J. O. Kephart, and W. S. Stornetta. Spawn: A distributed computational economy. *IEEE Trans. on Software Engineering*, 18(2):103--117, February 1992.

[WPS96] Q. Wang, J. M. Peha, and M. A. Sirbu. The design of an optimal pricing scheme for ATM integrated-services networks. In J. P. Bailey and L. Mcknight, editors, *Internet Economics*, Massachusetts, 1996. MIT Press.

[Wro97] J. Wroclawski. Specification of the controlled-load network element service. Draft RFC, Integrated Services Working Group: draft-ietf-intserv-ctrl-load-svc-05.txt. MIT LCS, May 1997.

[WTSW97] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Trans. on Networking*, 5(1):71--86, February 1997.

[WV96] J. Walrand and P. Varaiya. *High-Performance Communication Networks.* Morgan Kaufmann Publishers, Inc., San Francisco, CA, 1996.

[ZDE+93] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala. RSVP: A Resource ReSerVation Protocol. *IEEE Network Mag.*, pages 8--91, July 1993.