An Efficient CAC Scheme for IP Traffic over Wireless ATM Networks

H. Zainol Abidin[1], N Fisal[2], M.S.Ansari [3]
Faculty of Electrical Engineering,
University Teknologi Malaysia,
81310 Johor Bahru, Johor Darul Ta'zim.
Email: [2]sheila@suria.fke.utm.my

Abstract— **Traffic control and resource management are complex issues in wireless asynchronous transfer mode (WATM) due to broad range quality-of-services (QoS), limited bandwidth spectrum and susceptibility to error. In WATM, traffic control is essential in order to protect the network from congestion and to achieve realistic network efficiency in compliance with the QoS. Call Admission Control (CAC) is one of the critical mechanisms in providing QoS. This paper describes bandwidth allocation algorithms for IP traffic over WATM network based on VCT cell cluster threshold. Simulation work using Network Simulator (NS2) is also carried out to observe the network performance based on the proposed bandwidth allocation.**

Keywords-Wireless ATM; CAC; QoS; effective bandwidth; IP

I. INTRODUCTION

Asynchronous Transfer Mode (ATM) is a data transport technology that supports a single high-speed infrastructure for integrated broadband communications involving voice, video and data. Bandwidth efficiencies are achieved through statistical multiplexing of transmission bandwidth. ATM networks are characterized by virtual channel connections (VCCs) that carry small, fixed size packet (53 bytes) called cells within the network irrespective of the applications being supported. ATM is able to handle a wide range of information bit rate together with various types of real-time and non real-time service classes with different traffic attributes and quality of service (QoS) guarantees at cell and call levels.

Call Admission Control (CAC) is one of the critical mechanisms in providing QoS. In CAC, network attempts to deliver required QoS by allocating an appropriate amount of resources such as bandwidth. Bandwidth allocation in ATM is usually done by obtaining the effective bandwidth of the sources.

The virtual connection tree (VCT) concept as described in [1,2] reduces the call setup and routing load on the network call processor (NCP), while maintaining the QoS of the handoff calls. The VCT is a collection of cellular base stations (BSs), which is called the cell cluster, while the backbone network or the wired network, which consists of switching nodes and links as shown in Figure 1.

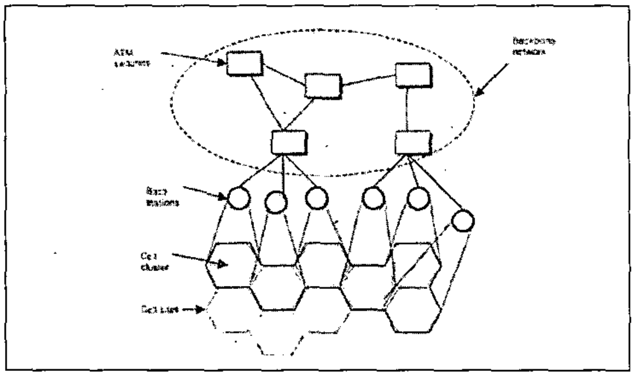

Figure 1. Structure of VCT

This paper presents and analyzes the bandwidth allocation algorithms for IP traffic over WATM. We assumed the network operates based on VCT cell cluster threshold. The rest of the paper will be organized as follows. Section 2 will focus on wireless ATM (WATM). Then, description on Call Admission Control (CAC) including statistical multiplexing and IP traffic will be introduced in section 3. Section 4 outlines the proposed bandwidth allocation technique while section 5 will briefly discuss the simulation work. Finally, section 6 concludes the paper.

II. WIRELESS ATM (WATM)

Wireless communication offer borderless freedom to customers to using mobile devices such as phones, television and wireless computers. According to some forecast the number of mobile phone subscribers may reach one billion by the year 2010 and surpass fixed phone lines [3]. As market demands for wireless communication continue to grow rapidly, it is apparent that a new generation of wireless networks will be needed for two main reasons:

- To enable different wireless technologies to interworking seamlessly with existing wired networks;
- To meet the diverse traffic demands (e.g. voice, video, data) required by current and future customers.

The convergence of ATM in the wired and wireless domains serves as an effective platform in achieving these objectives. This trend can be found in several projects within research program of the European Community (EC). There, a new generation of mobile communications technology had been introduced as the Universal Mobile Telecommunication System (UMTS) for the RACE project, which focus on the design of the Mobile Broadband System (MBS), a wireless cellular network fully integrated into B-ISDN. This mobile application can reach an approximate range of 2 to 155 Mb/s and RACE mobile is now studying a Mobile Broadband System operating in 60 GHz bands [4].

III. CALL ADMISSION CONTROL (CAC)

Call admission control (CAC) is the first line of defense for a network in protecting itself from excessive loads. In WATM, the CAC function is located in the BS. Usually, the CAC function has to be called before a handoff procedure can be completed. In order to guarantee a certain degree of mobility for the wireless ATM terminals, the CAC function needs to reserve some of the available radio bandwidth for handoff calls. A handoff call drop results an interruption of on going call and occurs if the bandwidth required by a connection cannot be supplied in the new cell site location.

In virtual connection tree (VCT), the CAC function is a mechanism that limits the incoming calls into the network in order to protect the already connected calls from being interrupted. An accepted connection that is requested by a mobile terminal (MT) will be given pre-established virtual connections to all the BSs in the tree so that the MT can freely move within the tree. The VCT network also monitors the existing connections in the network in order not to exceed the pre determined threshold limit. Moreover, it uses a scheme that gives priority to the handoff calls where a percentage of channels are reserved for handoff calls whenever the channels in the BSs are busy.

A. *Statistical Multiplexing and IP Traffic*

In ATM network, several sources will superimpose on a single link such as trunk line that may carry hundreds of videophone calls. Furthermore, the efficiency is gained by relying on the statistical multiplexing effect sources, on the condition that enough sources are multiplexed and they are not correlated.

The issue of resource allocation to connection is a challenge, due to the lack of physical separation between resources used by different connection in ATM based networks. The resources are actually the carrier system, which are for high capacity networks designed for the digital transmission of voice, video and data. In order to offer efficient resource allocation, we could apply different bandwidth allocation techniques for different types of IP traffics which will be outlined in next section.

As we are focusing on the provisioning QoS of IP traffic over WATM, we proposed that IP traffic is classified as in Differentiated Services (DiffServ) manner.

DiffServ offers QoS by allowing priority scheduling to facilitate the multimedia applications over the Internet. Figure 2, shows the structure of DiffServ domain. The DiffServ specifications refer to the forwarding treatment provided at a router as per-hop behavior (PHB). PHB must be available at all routers and normally PHB is the only part of DiffServ implemented in core router.

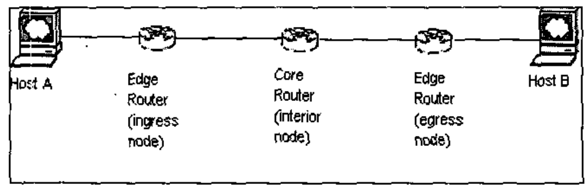

Figure 2. DiffServ domain

The Internet Engineering Task Force (IETF) has currently specified two different PHBs known as Expedited Forwarding (EF) and Assured Forwarding (AF) [5]. EF traffics are normally given strict priority over the traditional best-effort (BE) traffic inside the DiffServ domain. However, each flow has to specify the required bandwidth in advance so that the appropriate resources can be reserved inside the network. Furthermore, the edge router will police each flow and the non-conformant packets will either be dropped or shaped. Since AF does not offer hard QoS guarantees, IETF has specified four different AF class. Each class is assigned a certain amount of bandwidth at each node. When the amount of traffic exceeds this bandwidth, packets will be dropped according to their drop precedence value.

In this paper, EF, AF and BE traffic will be represented by CBR, VBR and ABR or non real-time sources respectively. It is assumed that the BS functions similar to ingress node in DiffServ domain while the wired ATM network does not require PHB at all their nodes. It is also assumed that BS is able to recognize EF, AF and BE as CBR, VBR and ABR respectively.

IV. BANDWIDTH ALLOCATION

WATM allows transfer of various types of multimedia traffic and ensured guaranteed QoS. A CAC in WATM will determine the amount of bandwidth to be allocated for each source accepted into the network. The ATM Forum has categorized the various types of source traffic into CBR, VBR (rt-VBR and nrt-VBR), UBR and ABR. The bandwidth allocation mechanisms differ for each time of demand. The following sub-section explains the bandwidth allocation techniques for CBR, VBR and UBR ad ABR

A. *Peak Bandwidth for EF (CBR) Sources*

A deterministic rate such as voice sources, usually hold one unit of source for the whole duration of the connection. Figure

3 illustrates the state-transition-rate diagram for m-server (time-slot) loss system with Markov arrival and service process.

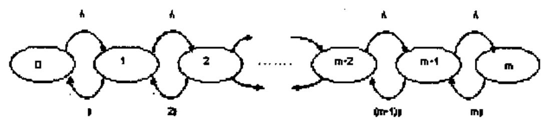

Figure 3. State-transition-rate diagram for M/M/m/m

The probability that the systems having k calls, p_k can be obtained as follows:

$$p_k = \begin{cases} p_0\left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!} & k \le m \\ 0 & k > m \end{cases} \tag{1}$$

where $p_0 = \left[\sum_{k=0}^{m}\left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}\right]^{-1}$, λ is the call arrival rate and μ is the call departure rate. Hence, the fraction of time that all m timeslots are busy, p_m is determined as;

$$p_m = \frac{(\lambda/\mu)^m/m!}{\sum_{k=0}^{m}(\lambda/\mu)^k/k!} \tag{2}$$

This probability expression is known as Erlang's loss formula. Thus, Erlang's loss formula is generally derived from M/M/m/m queue. This system has a blocked calls cleared situation in which there are available m servers (timeslots). Normally, CBR connection is given a fixed peak bit rate such as 64kb/s for voice [6]. The Erlang's loss formula also indicates B_p which is equal the peak rate will be allocated for EF sources.

B. *Effective Bandwidth for AF Sources*

Finding the effective bandwidth in ATM connection is important in order to maintain the QoS of the connection and to ensure that the connections are used efficiently for variable bit rate traffic. AF sources such as video traffic can be represented by Gaussian distribution. The effective bandwidth for a Gaussian i.i.d random process, *x(n)* can be calculated by using the following equation [7].

$$B_E = \mu + \frac{\sigma^2}{2}\delta \tag{3}$$

where μ is mean arrival rate and σ^2 is the variance. δ is determined by using the following equation

$$\delta = \left[\frac{\ln(\alpha) - \ln(\varepsilon)}{B}\right] \quad , \qquad 0 < \alpha \le 1 \tag{4}$$

Typically, α is set to 1. ε is the loss probability and B is the buffer size. Calculation is done by considering the following parameters. $\alpha = 1$, $\varepsilon = 10^{-9}$ and B varies from 100kB to 300kB. σ and μ will be shown in the table in the form of (σ,μ) in kB. (σ,μ) values are taken from reference [8]. Table I shows the effective bandwidth for AF sources by applying equations (3) and (4).

TABLE I. EFFECTIVE BANDWIDTH FOR VBR SOURCES USING GAUSSIAN I.I.D

(σ,μ)	B		
	100kB	*200kB*	*300kB*
8kB, 21kB	27.62kB	24.32kB	23.21kB
5.7kB, 16.3kB	19.66kB	17.98kB	17.42kB
18.1kB, 37.6kB	71.88kB	54.76kB	49.04kB

The result shows that Gaussian i.i.d seems to be not conservative in providing effective bandwidth. Therefore, a more conservative effective bandwidth derivation has been adopted based on self-similar traffic known as Pareto distribution. Pareto distribution is suitable to represent modern network traffic such as VBR, Web etc. The effective bandwidth, B_E can be determined by using Pareto distribution as follows:

$$B_E \ge \mu + \left[\frac{-2a\mu \ln\varepsilon}{B^{2(1-H)} f(H)}\right]^{\frac{H}{2}} \tag{5}$$

where, μ is the mean arrival rate of the traffic stream in bps, a is the variance coefficient which is calculated as the variance to mean ratio, B is the buffer size, H is the Hurst value and finally ε is the lost probability. Function of Hurst parameter, $f(H)$ could be determined by using the following equation.

$$f(H) = \frac{1}{(1-H)^{2(1-H)} H^{2H}} \tag{6}$$

By using the same parameters and $H = 0.85$, the effective bandwidth required by AF traffic calculated by using equations (5) and (6) are shown in Table II.

TABLE II. EFFECTIVE BANDWIDTH FOR VBR SOURCES USING PARETO DISTRIBUTION

(σ,μ)	B		
	100kB	*200kB*	*300kB*
8kB, 21kB	22.63kB	22.49kB	22.41kB
5.7kB, 16.3kB	17.52kB	17.42kB	17.36kB
18.1kB, 37.6kB	40.87kB	40.60kB	40.44kB

From the above results, it is proven that Pareto distribution tends to be more conservative in allocating resources when network traffic exhibits long-term memory. Therefore, bandwidth that will be allocated in our network would be based on the effective bandwidth obtained from the Pareto distribution calculation.

C. *Mean Bandwidth for BE Sources*

Best effort IP traffic represents many of the non real-time applications such as data coming from LAN. The arrival and service time of IP packets can be approximated to Markov birth and death process in an infinite queuing system. The birth-death process of such queuing system can be illustrated as state-transition-rate diagram of Markov Chain in Figure 4.

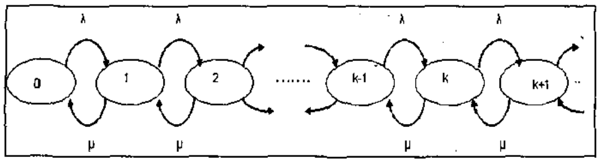

Figure 4. State-transition-rate diagram for infinite queuing system

From the state-transition-rate diagram, the equilibrium different equation for state k can be determined as follows [10]:

$$\lambda_{k-1}p_{k-1} + \mu_{k+1}p_{k+1} = (\lambda_k + \mu_k)p_k \tag{7}$$

where λ is the call arrival rate, μ is the call departure rate and p_k is the probability that the systems with k members. Hence, p_k can be simplified as follows:

$$p_k = p_o \prod_{i=0}^{k-1} \frac{\lambda}{\mu} \quad , k \geq 0 \tag{8}$$

where

$$p_0 = \frac{1}{\left[1 + \sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^k\right]}$$

Since $\lambda < \mu$, the summation will converge, therefore

$$p_0 = 1 - \frac{\lambda}{\mu} \tag{9}$$

From the stability conditions, the utilization, ρ should be $0 \leq \rho < 1$ *to ensure that* $p_0 > 0$. *The steady-state probability of* finding k customers in the system is:

$$p_k = (1-\rho)\rho^k \quad , k = 0,1,2,.. \tag{10}$$

By applying Little's formula, the average delay, *E[t]* is obtained from E[t] = E[n]/λ, where *E[n]* is the average number of customers in the system. Thus,

$$E[t] = \frac{1/\mu}{1-\rho} \tag{11}$$

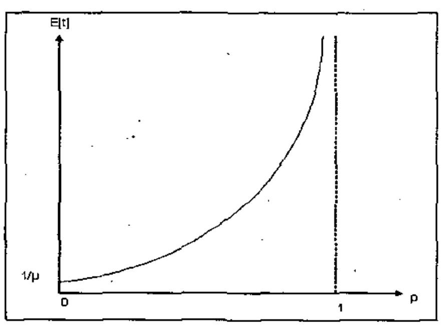

Figure 5. Average delay as a function of ρ for M/M/1

Figure 5 shows that when $\rho = 0$, *E[t]* is exactly the average service time that is expected by a customer. Therefore it is reasonable to allocate mean bit rate, B_m for BE traffic.

V. SIMULATION WORK

Simulation work has been carried out to observe the effect of our proposed bandwidth allocation mechanism to the network performance for AF sources. Figure 6 shows the simulation network model.

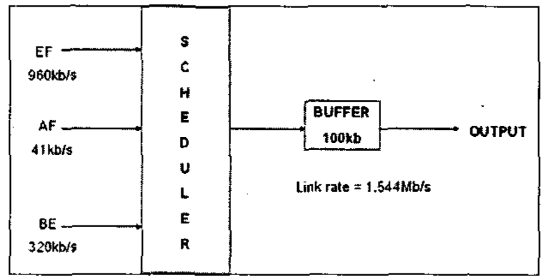

Figure 6. Simulation Network Model

The three types of sources used are as described below:

- Source 1 (EF): CBR traffic is based on UDP transport protocol with a rate of 960kbps and mean packet size of 1500 bytes. This traffic is suitable to represent voice application.
- Source 2 (AF): represents VBR traffic based on UDP transport protocol with a mean rate of 41kbps and mean packet size equal to 1500 bytes. This traffic would be suitable to represent video application.
- Source 3 (BE): represent ABR/UBR traffic for non real-time traffic with rate of 320kbps and packet size equal to 1500 bytes.

In the simulation work, the buffer size is assumed 100kb, link rate is using T1 line equals to 1.544Mb/s and scheduling mechanism is based on Priority Queuing (PQ). A guard channel of a certain percentage is reserved for handoff of on-going calls. The three types of traffic arrived at the BS according to deterministic, Gaussian and Pareto distribution representing IP traffic of type EF, AF and BE. In normal case IP traffic is treated as ABR/UBR traffic in ATM network.

Figure 7 shows the end-to-end delay for AF sources using the proposed bandwidth allocation technique while Figure 8 illustrates the end-to-end delay for AF sources that are treated as normal BE traffic. *In the later case the effective bandwidth* of AF sources is B_m

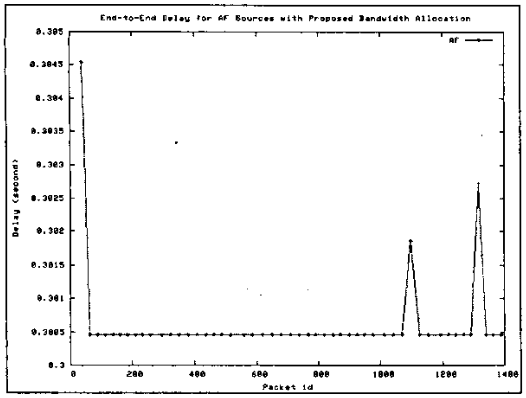

Figure 7. End-to-end delay for AF with proposed bandwidth allocation

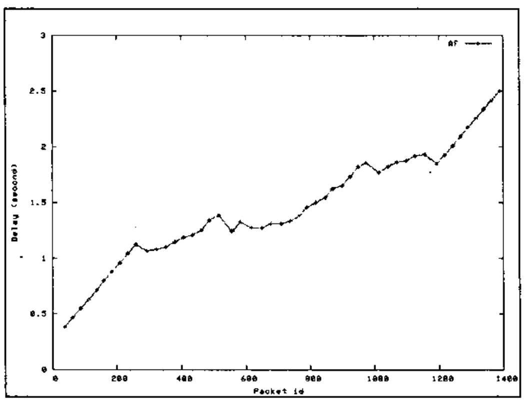

Figure 8. End-to-end delay for AF treated as BE

The two graphs show that our proposed bandwidth allocation technique with PQ scheduling for AF traffic gives almost constant low (about 0.3 second) end-to-end delay compared to treating AF traffic as normal IP BE traffic.

VI. CONCLUSIONS

Call admission control (CAC) is an important mechanism in provisioning QoS for IP traffic over WATM network. This paper has proposed a simple bandwidth allocation technique for various types of IP traffic sources. Our efficient CAC function uses an efficient bandwidth allocation technique as well as efficient scheduling method. Peak bandwidth, effective bandwidth and mean bandwidth are allocated to CBR, VBR and ABR sources respectively. We assume that IP traffic over the WATM network is using similar DiffServ technique at the ingress node (base station). This is to ensure that the different types of traffic are given priority. Thus, DiffServ PHB at the ingress node such as EF, AF and BE is represented by CBR, VBR and ABR sources respectively in WATM. The simulation study has shown that the performance of the network in terms of end-to-end delay is minimized as we allocate effective bandwidth to the traffic.

REFERENCES

[1] S.H.S. Ariffin, N. Fisal and M. Esa, "Quality-of-Service Performance in Micro Cellular Network," Proceedings of World Engineering Congress 1999, Malaysia, pp. 63-67.

[2] S.H.S. Ariffin, N. Fisal and M. Esa, Study of Mobility Management in Wireless ATM Network," Proceedings of NCTT 98, Universiti Putra Malaysia, Malaysia, pp. 125-131.

[3] M. Shafi et. al. "Wireless Communication in the twenty First Century: A perspective". Proceedings of the IEEE. Vol. 85. No. 10. October 1997.

[4] R. S. Swain, "UMTS- A 21st Century System: A RACE Mobile Project Line Assembly Vision". http://www.vtt.fi/tte/nh/UMTS/umts. Html

[5] S. Bakiras and V.O.K. Li, "Efficient Resource Management for End-to-End QoS Guarantees in DiffServ Networks," IEEE International Conference on Communications, 2002. ICC 2002, Vol.2, pp. 1220-1224.

[6] S.H.S. Ariffin, "Quality of Service Provisioning in ATM Wireless Network under Real-Time and Non Real-Time Connections," Master of Electrical Engineering (Telecommunications) Thesis, Fakulti Kejuruteraan Elektrik, Universiti Teknologi Malaysia, 2001.

[7] K. Nagarajan and G.T. Zhou, "A New Resource Allocation Scheme for VBR Video Traffic Sources," Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers, 2000, Vol.2, pp 1245-1249.

[8] P. Orenstein, H. Kim and C.L.Lau, "Bandwidth Allocation for Self-Similar Traffic Consisting of Multiple Traffic Classes with Distinct Characteristic,"IEEE Global Telecommunications Conference 2001, GLOBECOM '01, Vol.4, pp. 2576-2580.

[9] I. Norros, "On the use of Fractional Brownian Motion in the Theory of Connectionless Networks," IEEE Journal on Selected Areas in Communications, vol. 13(6), pp.1028-1038, August 1995.

[10] L. Kleinrock, "Queuing System, Volume 1: Theory," New York, John Wiley, 1975.