A Technical Report of ITTC's
Networking and Distributed Systems Laboratory

Evaluation of Media Access Control (MAC) Protocols for Broadband Wireless Local Loop (BWLL) Applications and Recommended Design Improvements

Mihir Thaker
Joseph Evans
James Roberts

Technical Report
ITTC-FY2000-TR-15663-03

Information and Telecommunication Technology Center
The University of Kansas
Lawrence, Kansas

November 2000

Project Sponsor:
Sprint Corporation

Abstract

Because of the tremendous growth of the Internet, the existing network infrastructure is unable to handle the large volume of traffic flowing across it. Advances in the field of wireless communications have yielded Broadband Wireless Local Loop (B-WLL) as a solution to alleviate the problem of providing Internet access to a large number of users for the "last mile." Utilization of the available bandwidth depends upon the efficiency of the Media Access Control (MAC) scheme used in B-WLL. Each application that used over the Internet has a different traffic pattern. Because of its architecture, the behavior of a given MAC protocol varies with the type of application running over it. An Internet Service Provider (ISP) can decide upon the best-suited MAC for its operation, if it has information regarding the dominant types of applications used on the Internet. This report evaluates the performance of two widely used MAC protocols, namely Reservation-TDMA (R-TDMA) and Multi-Frequency Polling (MF-Polling) based on the dominant applications used on the Internet, namely File Transfer Program (FTP) and Web Browsing using Hyper Text Transfer Protocol (HTTP). Aggregate throughput, average queuing delay and the supported user population have been chosen as output parameters for this evaluation. Design improvements have been suggested to reduce the average queuing delay of the protocol. It was observed that HTTP based applications are suited to an R-TDMA system while FTP based applications perform better with MF-Polling system. Future work could involve making the MAC scheduler QoS-aware and thereby provide efficient real-time services.

Contents

List of Figures

Chapter 1

Introduction

The Internet has experienced tremendous growth over the last decade. With the Internet growing at a rate of more than 200 percent per year in terms of user population, severe demands are being placed on the Internet Service Providers (ISPs) to handle this phenomenal growth. The existing wireline infrastructure is unable to grow as rapidly as the user base due to several constraints [CoA98]. This has led the industry pundits to look upon wireless media as a cost effective and viable alternative to deliver the same services being offered by wireline counterparts. There are several wireless spectrum blocks in the 2.1 to 2.7 GHz band that can be used for Internet services, including Multi-point Distribution Service (MDS, 2.150-2.162 GHz), Multi-channel Multi-point Distribution Service (MMDS, 2.596-2.644 GHz and 2.686-2.689 GHz).

The Federal Communications Commission (FCC) has divided the spectrum into bands that are either licensed or unlicensed. The licensed bands require an explicit license from the FCC before they can be used. This may not be suitable for research applications where the focus may lie on rapidly developing a prototype. Unlicensed bands are more suited to such applications and tend to be less expensive. To meet the increasing demands of the ISPs and effectively utilize the radio spectrum, the FCC created the Unlicensed National Information Infrastructure (U-NII) band by expanding and dividing the 5.8 GHz Industrial, Scientific and Medical band (also known as the ISM band) into three 100 MHz bands. The frequency ranges include 5.15-5.25 GHz, 5.25-5.35 GHz, and 5.725-5.825 GHz. The FCC has also eased broadcasting restrictions on the MDS and MMDS bands and two way broadcasting is now permitted in these bands. Similarly, there are no restrictions on the U-NII bands except for the maximum effective radiated power. Such change in policies has fostered growth in the field of wireless communications and is catching industry attention.

One of the biggest technological challenges in convincing the consumers to switch over from wireline to wireless systems is a question of quality. It is imperative that, in order to compete with their traditional counterparts, the wireless services must be of comparable quality. Faster access to information is always welcome to the end users. This implies that the ISPs must be able to provide greater bandwidths or make best use of the available resources. The wireless systems will have to cope up with this requirement and provide the necessary bandwidth if they are to compete with the wireline systems. Many wireless technologies are emerging to satisfy the consumer's quest for bandwidth. These include WLL (Wireless Local Loop) and WLAN (Wireless Local Area Network). Over the last few years, much attention has also been focussed on the extension of ATM to wireless communications, thus leading to WATM technology.

1.1 Motivation

Currently, there are two dominant methods for accessing the Internet. The first involves a gateway from a high-speed LAN and is representative of the access mode enjoyed by large business users. The advantage of this mode is high-speed access (multi-megabit/sec) from each computer attached to that LAN.

The second access mode is that suffered by small business and residential users. Here ordinary telephone calls are routed to the telephone network via narrowband dial-up circuits provisioned by the telephone plant Central Office (CO). The ISP provisions connections to the Internet. While the data path is high speed and packet switched on the Internet side of the ISP router, it is via a dial-up narrowband circuit provisioned through the CO on the subscriber side. Thus, the same telephone plant used for narrowband, circuit-switched voice is also used for access to the Internet. This is the so-called "last mile" problem: the CO-based switching apparatus and the twisted-pair copper wiring connecting the subscriber to the CO today prevents the realization of the full, bandwidth-upon-demand, bandwidth-intensive multimedia potential of the Internet. It is this last mile dilemma which has caught the attention of researchers and the industry alike.

A third set of access technologies that enable high speed Internet access has emerged since then and is already being deployed. DSL (Digital Subscriber Line) and WLL technologies are the most prominent technologies from this set. DSL is a new modem technology which can permit multi-megabit access directly to Internet packet switches over existing copper wiring, provided that the span is sufficiently short and the wiring is of sufficiently high quality (newly installed, few splices). It is estimated that fewer than 40% of American homes are candidates for DSL [CoA98]. Cable modem technology exploits the relatively high bandwidth of the existing coaxial cable plant, originally intended for the one-way distribution of entertainment video. Suitably upgraded for two-way communications, and suitably equipped with premises-based cable modems, this plant has the potential to provide 50%-60% of American homes with multi-megabit packet service for Internet access [CoA98].

The other access strategy is that of WLL to provide broadband services. In telephony, the term "loop" is defined as the circuit connecting a subscriber's station (e.g., a telephone set) with the line terminating in a CO (e.g. a switch in a telephone network). These circuits are broken down into several smaller bundles of circuits and eventually separated into individual drops for the residences. By definition, WLL (a.k.a. Fixed Radio Access) is a system that connects subscribers to the PSTN using radio signals as substitute for the existing copper wire for all part of the connection between the subscriber and the switch. WLL is concerned only with the connection from the distribution point to the house; all other parts of the network remain unaffected. Hence, apart from a radio and antenna, the home subscriber does not notice any difference. The central office here, termed as "Headend," has a similar radio and antenna pair.

Of the three access technologies, the wireless approach offers the distinct advantage of being completely *tetherless*, that is, service is supplied directly to wireless terminals, rather than to the limited set of fixed DSL and cable modem locations terminating the copper wiring and coax drop, respectively [Aca99].

Since the FCC has permitted long distance carriers to offer direct access to residential users, WLL has gained popularity as a viable alternative to the expense of running a wire to the home. This implies very low deployment time and cost for the ISPs as well as the end users.

It must be noted that in order to reap the benefits of the WLL technology, the radio resources must be utilized efficiently. The ISP must take into account two important factors before deploying any WLL system, namely, the supported user population and the data rate offered to the each end user. The former forms the revenue generating part that every ISP aims to maximize, and the latter forms the Quality of Service (QoS) aspect dictated to the end users. It can be seen that the above-mentioned factors bear an inverse relationship with each other. Figure 1.1. shows a typical WLL setup.

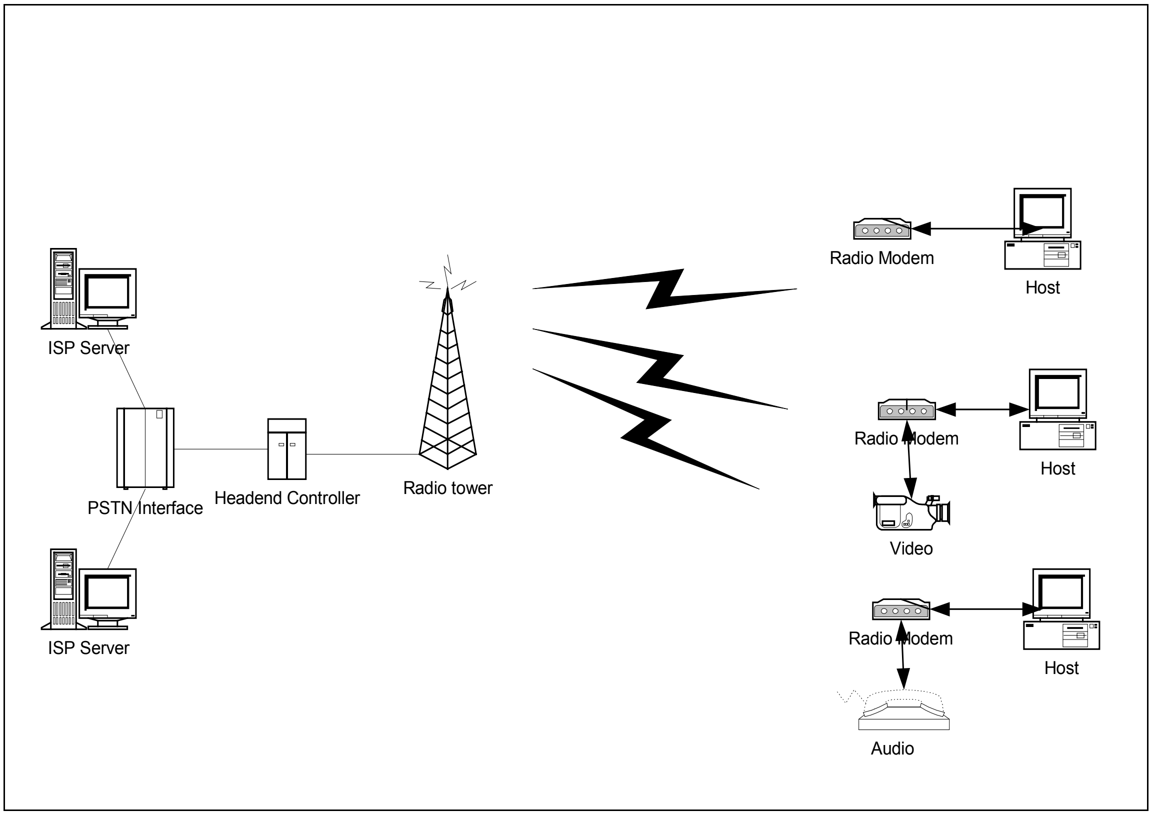

Fig. 1.1. Wireless Local Loop Setup

In order to support a large user population and utilize the radio resources effectively, there arises a need for a mechanism that would regulate user traffic on the radio channel. Such mechanisms or protocols are commonly termed as Media Access Control (MAC) protocols. The following section provides a brief overview of the various types of MAC protocols.

1.2 Concept

MAC protocols are channel allocation schemes that control shared resources (in this case, a radio channel) in order to make a transmission successful. The need for multiple access protocols arises not only in communication systems but also in many other systems such as computer systems, storage facilities, and servers of any kind where a resource is shared (and thus accessed) by a number of independent users.

There are various ways to classify MAC protocols. Examples of such classifications appear in [KuS84]-[Sac88]. One way of classifying MAC protocols could be as non-centralized access and centralized access protocols. In non-centralized access protocols, all nodes behave according to the same set of rules. In particular there is no single node coordinating the activities of the others (whose protocol differs from the rest). These are also known as *Contention based* protocols. ALOHA, Slotted ALOHA, Carrier Sense Multiple Access with Collision Detection (CSMA/CD) are examples belonging to this class of MAC protocols. These mechanisms, however, suffer from the problem of high collision and low bandwidth efficiency. In centralized access protocols, a central node controls the flow of data in the network. This requires some complex scheduling mechanism in the central node, which makes the implementation of a centralized access protocol relatively difficult. However, the advantages are a large supported user population and high bandwidth efficiency. Bandwidth can be managed either in the time domain or in the frequency domain. Time Division Multiple Access (TDMA) and Frequency Division Multiple

Access (FDMA) are the two most prominent types of mechanisms for centralized access.

It is not possible for any system provider to know beforehand which type of access protocol would be best suited for deployment unless the characteristics of each type of protocol are fully known. In order ensure efficient utilization of the available bandwidth, any protocol must take into account the user traffic profiles. Allocation values larger than those demanded by the user profiles would lead to inefficient utilization, thereby lowering system capacity. On the other hand, values smaller than those demanded by the user would place the end user at a disadvantage, rendering the choice of using WLL useless. Thus, it is desirable that a MAC protocol adapt to changing user traffic. However, user traffic type is unpredictable in nature. Also, every MAC protocol must able to support the standard user applications like File Transfer Program (FTP), Internet browsing which uses Hypertext Transfer Protocol (HTTP), and multi-media applications. Having known the dominant type of application that a given user base uses, it is possible to simulate different scenarios with different MAC protocols.

This report summarizes simulations and comparisons of the performance of two centralized access MAC protocols namely Reservation TDMA (R-TDMA) and Multi-Frequency Polling (MF-Polling), which are variations of TDMA and FDMA respectively, [Tha00]. The R-TDMA protocol is used by Adaptive Broadband in their wireless local loop products, and the MF-Polling protocol is used by Hybrid Networks in their solutions. Upon identification of the superior technique, we can map different applications to the best-suited MAC protocol. The report also recommends design improvements to the protocols that can improve performance.

A comparative model has been developed to achieve the above mentioned objectives. This model is as shown in Figure 1.2.

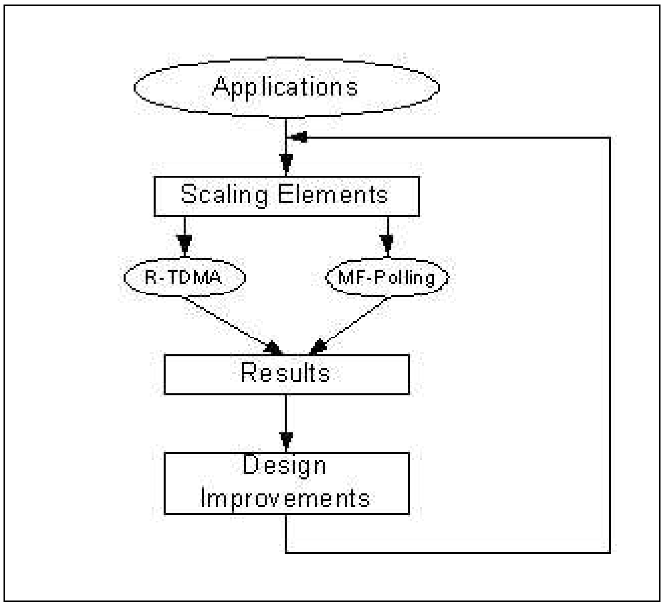

Fig. 1.2. Comparative Model

As seen from the above figure, our comparative model runs the same set of applications over the two protocols under consideration. In order to compare the two MAC protocols, it is necessary that the operating conditions be the same for each one of them. These operating conditions would include the available bandwidth, user traffic patterns, Transport Control Protocol (TCP) parameters, modulation schemes etc. We are ensuring that the parameters, which might affect the output, are scaled in both the protocols by means of the Scaling Elements block that is discussed in Section 2.2.3. Having observed the results we can compare and contrast the obtained values. Improvements can be suggested to the protocol design and the above process can be repeated to obtain improved results. OPNET Modeler/Radio^{TM} has been used as the simulation tool to conduct all the tests. This tool allows users to model radio links with the desired characteristics of bandwidth, modulation schemes, transceiver frequency etc. Various types of networks can also be modeled using this tool. The end-to-end behavior of any wireline or wireless system can be simulated and its performance measured.

1.3 Overview

The rest of this report is organized as follows:

Chapter 2 deals with the technical aspects of different MAC protocols, concentrating more on R-TDMA and Multi-frequency Polling. The system design used for simulation is also discussed in greater detail. Our basis for comparing the two different types of protocols is justified in Section 2.3, where we discuss about the specific parameters that affect the output performance. *Chapter 2* also gives a brief overview of the related efforts in this field highlighting the significance and uniqueness of this research work.

Chapter 3 discusses the test scenarios used in performance evaluation of the two MAC protocols. Here we discuss the rationale behind the various types of applications used. In terms of the volume of traffic sent and received by the end user, we classify the applications as being *Symmetric* or *Asymmetric*. Standardized traffic patterns, which govern such type of applications, are also discussed.

Chapter 4 presents the simulated test results and provides a comparative study of the same. Having conducted the test scenarios mentioned in *Chapter 3,* we analyze the graphs obtained and also discuss the reasons for the bias of each protocol towards a particular type of application. The significance of using robust protocol architecture is made clear in the ensuing discussion.

Chapter 5 suggests design improvements to the existing MAC protocols and re-evaluates the performance of the improvised MAC protocols for the same test scenarios. We specifically consider the effects of varying the number of contention slots for R-TDMA protocol from frame to frame depending upon the number of collisions observed. Similarly, for MF-Polling we reduce the queuing delay by decreasing the maximum window size required for contention. The effect for reducing the polling cycle time is also evident from its impact on the queuing delay.

Chapter 6 summarizes the performance evaluation and suggests future work. Future work can include improvisation to the protocols by providing support for QoS requirements of the end users. A strong foundation can be laid for combining the best of R-TDMA with the fixed channel sizes of MF-Polling. A good framework can be built for designing MF-TDMA systems.

Chapter 2

Media Access Control Technologies

2.1 Introduction to Wireless MAC protocols

The MAC protocol deals with the problem of ensuring coherent communication between end systems. In designing a MAC protocol, one must always keep in mind that all the users in a given system may not have data at all times. Traffic patterns are unpredictable in nature and such variations in traffic conditions have led the research community to develop separate classes of MAC protocols. These protocols can be classified to be either *structured* or *probabilistic* in nature. Structured protocols are based on the concept of monitoring and regulating traffic on the channel. This requires the presence of a monitoring node in the system which imposes certain constraints on the system and affects the system latency (Latency is defined as, "the amount of time taken by a single bit to propagate from one end of a network to another" [PeD96]). However, the advantages are fewer collisions, higher throughput, and large supported user population. TDMA, FDMA and Code Division Multiple Access (CDMA) are examples in this category.

On the other hand, probabilistic class of protocols takes into account the randomness and probability associated with data being present in the user queue. There is no one unit that controls the data transmission. These systems suffer from poor bandwidth efficiency and hence can support only a low user population [Rap99]. ALOHA, Slotted ALOHA, Carrier Sense Multiple Access with Collision Detection

(CSMA/CD) are examples of this class. All other types of protocols are variations and hybrids of the above two classes [RoS90]. These variations have been quite popular in the industry. Their popularity can be attributed to the fact that they are not only efficient but can also be scaled as per the requirements of the target users and the offered service mix. R-TDMA, Polling, and Demand Assignment Multiple Access (DAMA) are examples of such protocols. A distinct feature of these protocols is that it requires a scheduler to maintain the desired access scheme. This is perfectly suitable to WLL as it is inherently based on the principle of centralized access (i.e. point-to-multi-point links). We shall use the terms "headend" and "host" throughout this document to represent the central host and end user respectively.

2.2 Protocol Specifications

2.2.1 Reservation TDMA (R-TDMA)

Time Division Multiple Access (TDMA) is a technology for delivering digital wireless service using Time Division Multiplexing (TDM). TDMA works on the principle of assigning the entire frequency band to various users for a fixed amount of time (known as "slot"). Figure 2.1. further elaborates upon this concept.

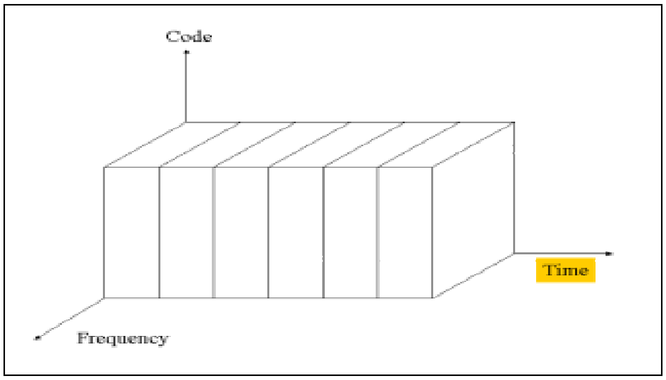

Fig. 2.1. Time Division Multiple Access

Utilization of given radio frequency band can be achieved by choosing any one of the axes parameters. TDMA chooses the "Time" axis as the basis for providing multiple access and is thus named so. It can be seen that for any given value of time on the "Time" axis, the "Frequency" value remains unchanged. In this manner a single frequency band can support multiple users spaced in time. Since the division is based entirely on time slots, each user enjoys maximum data transfer rate supported by the system for that particular slot.

R-TDMA is a variation of the TDMA protocol, which is used to provide slots-on-demand type of service. This has the advantage that the bandwidth is utilized only when required, and hence is efficient. Such a protocol employs a frame structure, which is repetitive in nature and is as shown in Figure 2.2.

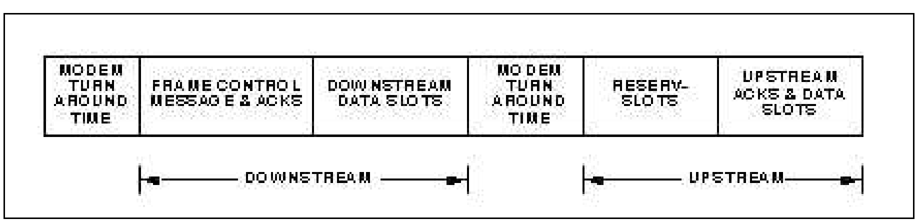

Fig. 2.2. Reservation TDMA Frame Structure

The headend maintains the frame structure and also manages the scheduling and assignment of slots in each frame. The frame structure can be divided into *downstream* and *upstream* parts respectively. The *downstream* here is constituted by flow of traffic from the headend to the host. The downstream part of the frame consists of *frame control* messages generated by the headend. These are used to convey the current frame information to the users. The current frame information is usually the assignment of slots to different users who have either requested slots for transmission in the upstream or those users who would be receiving data from the headend in downstream.

Slots can be requested by either contending for them during the reservation period in the upstream portion of the frame or by implicitly tagging the request information along with the upstream data. The latter provides a mechanism by which the user remains in the system and does not fall into contention mode, thereby allowing new users to enter in the system which further reduces the number of contending users. Slotted ALOHA with binary exponential backoff has been used as the contention mechanism in this evaluation. A user automatically falls into contention mode if it does not request for additional data slots. This ensures that the bandwidth is utilized efficiently.

The other information carried by the control message is the sequence of *acknowledgements (ACKS)* to users who have sent data upstream in the previous frame. It also carries information about the users who would be receiving data in the current frame. It is followed by the downstream acknowledgement sequence to the users notified by the control message.

The *upstream* portion of the frame consists of *reservation slots* wherein, users who have not been allocated any slots in the current frame can request for the same by contending with other users. The users send data upstream in accordance with the current frame plan received from the headend. Users also acknowledge the received data in the upstream by means of *upstream ACKS*. Separate slots are allocated for this purpose.

It can be seen that the headend and user radio switches from transmit to receive mode in each frame. In order to allow for the propagation delay for the farthest located user and modem turn time, no transmission is allowed for a specific period between upstream and downstream part of the frame. Downstream data is managed entirely by the headend and does not require any user intervention or request. The frame efficiency and aggregate system throughput can be measured as follows,

Let U_{tot} and D_{tot} represent the total upstream and total downstream data in bits. The entire frame size in bits can be represented as F_{tot}. This includes the data bits as well as the overhead message bits. Thus, the frame efficiency, η, is computed as

$$\eta = \frac{U_{tot} + D_{tot}}{F_{tot}} \tag{1}$$

The *link rate* of a channel is maximum allowable data rate on the given channel. Similarly, *throughput* is the amount of data that can be transmitted on the channel.

We can measure the R-TDMA system throughput T (bits/sec), from the link rate L (bits/sec) and the frame efficiency from (1) as

$$T = L \times \eta \tag{2}$$

Quadrature Phase Shift Keying (QPSK) has been used as the modulation scheme. Since a wireless channel is inherently noisy and is prone to burst errors, an optimum tradeoff is required between the maximum supported data rate and the available noise margin. QPSK provides a spectral efficiency of 2 bits/s/Hz and has a noise margin 3-dB lesser than Binary Phase Shift Keying (BPSK). Using an appropriate error detection and correction scheme can adequately compensate this 3-dB penalty in the signal-to-noise ratio (SNR). Thus, for this evaluation, QPSK has been employed. This sets the link rate (bits/sec) to twice the available bandwidth (Hz).

There exists a certain amount of delay before a user can actually transmit data. This delay is due to the contention faced by the user before being granted slots for transmission by the headend. This is known as the *contention delay*. The amount of time that a data packet spends in the queue before being actually transmitted, known as the *queuing delay*, is directly affected by the contention delay. The important parameters that need to be observed are the *throughput*, *queuing delay* and *number of users supported*.

2.2.2 Multi Frequency Polling (MF-Polling)

As shown in Fig. 2.3., the other method of managing bandwidth is by allocating different users to different frequency bands. This is known as Frequency

Division Multiple Access (FDMA). For any value of frequency on the "Frequency" axis the value of "Time" remains constant. This implies that at any point of time in its operation, the user has full access to that particular frequency slot. The maximum allowable data rate is limited to the channel bandwidth and the type of modulation scheme used. Such a protocol has the problem of co-channel interference and the capacity of the system is fixed according to the number of sub-channels in the frequency spectrum. However, it has the advantage that it does not require any synchronization like TDMA and thus can have up to 100% bandwidth efficiency [Kas98].

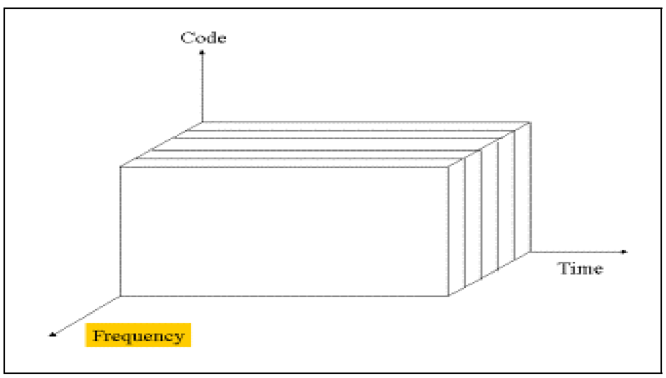

Fig. 2.3. Frequency Division Multiple Access

FDMA also suffers from the problem of under-utilized links when the user does not have data to send.

MF-Polling is a variation of FDMA, which attempts to utilize the bandwidth and support a larger user population by dynamically allocating the resources. The given radio spectrum B can be divided explicitly and equally into two channels as shown in Figure 2.4.

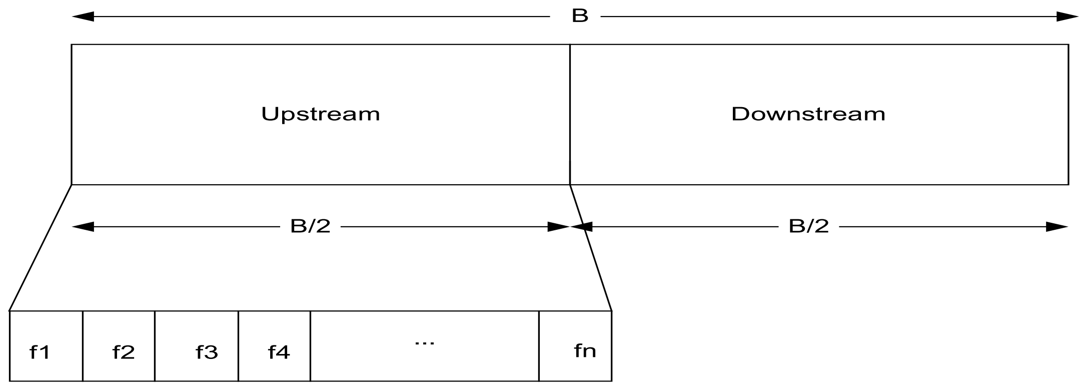

Fig. 2.4. MF-Polling Bandwidth Allocation Scheme

The upstream band is further sub-channeled into multiple fixed sized units. Each sub-channel represents a unique transmission frequency. Some of the channels are reserved for contention purposes (i.e. only request message and grant messages can be transmitted on these channels). These channels are polled at regular intervals, and users respond to these messages if they are currently not in the system and have data to transmit. In case of a collision, the users employ binary exponential backoff. The remaining channels are used for purposes of polling. Successful users from the contending pool of users are added to the polling channels.

Polling is achieved by means of a grant message, which is circulated in a round robin fashion to all the users existing on that particular frequency channel. The user in turn sends a stipulated amount data on the upstream and then ends its transmission by acknowledging the grant message. The maximum number of users that can be polled on a particular channel is decided by the maximum allowable latency on the system. It is likely that a user can have a large amount of data queued up which eventually leads to high queuing delays. To avoid this, the user can request for additional polling cycles in its acknowledgement. The headend recognizes such a user and temporarily moves the user to an empty polling channel where it can burst till its queue is empty. Users are aggressively moved out of the system if they fail to send data and are idle for more than a predefined time limit. This again ensures efficient bandwidth utilization. As in case of R-TDMA, we evaluate the same output parameters, namely, *aggregate throughput*, *queuing delay* and *number of users supported*.

From the above description it can be seen that, since we are observing the same output parameters for both the protocols, it is necessary to ensure that the input system parameters which affect the results be the same in both cases. Hence, we need to scale the protocols and operating conditions accordingly. Some parameters that directly affect the output have been chosen for purposes of scaling.

2.2.3 Scaling Elements

Since, we are comparing the *aggregate throughput*, *queuing delay* and *number of users supported* by the system, the following have been chosen as the common input parameters.

- *Available bandwidth* – The amount of available bandwidth combined with the modulation scheme used dictates the maximum link rate on the channel. For our simulation the currently popular radio bands for WLL (i.e. MDS1 and MDS2 aggregated) have been used.
- *User queue size* – The simulation tool (i.e. OPNET) suggests some standard queue sizes based on their research and industry specifications. We have used the same values for our simulation.

- *Modulation scheme* – QPSK has been used as the modulation scheme in this case.
- *Contention protocol* – Slotted ALOHA with exponential backoff has been used for R-TDMA protocol while exponential backoff has been used for MF-Polling. Since R-TDMA and Slotted ALOHA are based on the principle of using slots for communication, it is easier to incorporate the latter as the contention scheme for the former. Since MF-Polling is an FDMA based system, special design considerations have to be made to accommodate Slotted ALOHA as the contention scheme. Similarly, exponential backoff is more compatible with MF-Polling rather than with R-TDMA.

  Exponential backoff has higher efficiency than Slotted ALOHA [BeG92]. The efficiency of an exponential backoff scheme depends upon the product of average delay involved between retransmissions in case of a collision and the channel bandwidth. In the current implementation of the MF-Polling scheme, even though the channel bandwidth is low, the average delay is dependent upon the time required for each broadcast polling cycle. Since the polling cycle time is high compared to the average delay on a wireline system, we have a high delay-bandwidth product and effectively a low value of efficiency. The resulting protocol efficiency is comparable in value to that of Slotted ALOHA. Hence we are justified in using two different contention mechanisms, as their operating efficiency is about the same for the wireless systems under consideration.
- *User traffic profile* – Standard traffic profiles for FTP and HTTP suggested by OPNET are used for the simulation. Apart from this, other customized traffic

patterns have also been used. The customizations include varying the packet inter-arrival times, the packet generation pdf's (probability distribution functions), changing the application request pdf's such that they would emulate the real systems as closely as possible. Real-time traffic patterns have been collected from the data flowing across the Internet [PeD96].

- *Simulation time* – Simulation run time has been chosen long enough for the system to reach a steady state. This ensures that the results are not an effect of any transient conditions that might occur.
- *Error Detection* – Since contention messages have a higher probability of collision, error detection is performed only for such messages. It has been assumed that data and control messages do not undergo collisions because of the structured nature of protocols used. As contention delay directly affects the user data queuing delay, it is necessary that the erred contention messages do not go undetected. Also, we are primarily interested in observing the ad-hoc behavior of the given MAC protocols and not that of the radio channel. For simulation purposes the channel can be assumed to be error free.

2.3 Related Work

There has been ongoing research in simulation and comparison of wireless MAC protocols over the past few decades. Several MAC protocols have been developed and their behavior has been compared in order to test their effectiveness in various environments. Because of the increasing demands being placed on expanding the scope of services offered by the wireless systems, MAC protocols need to be customized according to the requirements of the applications and services. The onus to achieve this objective has been placed equally on both the industry and researchers. This has undoubtedly fostered innovation in this industry.

Wang and Wen [WaW96], have compared the queuing delay and throughput of R-TDMA for both TDD (Time Division Duplex) and FDD (Frequency Division Duplex) schemes. We have utilized a similar approach in the choice of our output parameters. However, our measurements also include the delay characteristics at the application layer.

Rom and Sidi [RoS90] have done a comprehensive analysis and comparison of TDMA and FDMA protocols. They have applied the concepts of queuing theory in order to derive a relationship between the expected delay, normalized throughput and the number of users in the system. Pure TDMA and FDMA systems can be treated as M/D/1 queues and thus, a relationship can be established between the expected delay of a TDMA and FDMA system. The results from this relationship show that for the number of users greater than 2, the TDMA expected delay always remains lesser than that of FDMA. Our implementation deals with a specific variation of the two protocols, which are dynamic in nature as opposed to the above-mentioned implementations. Thus, we expect to see some variation in the expected delay, rather than dominance of a particular protocol over the other.

Another approach towards such a comparison has been to treat the overall wireless system as Client-Server setup and observe its behavior as an Open Loop and

Closed Loop model. This unique approach was formulated by LaMaire et. al. [LKA93]. An Open Loop model is one in which a user can generate reservation requests for the central node irrespective of whether it has received a response from the central node. LaMaire's work bases it throughput performance on the probability of transmission during the contention slots of an R-TDMA system. This implementation, though being an Open Loop model implements, a transmission scheme in which the choice of the transmission slot is exponentially distributed and has a probability of transmission as 1 for that particular slot. This has been the most widely used mechanism in the industry.

Mikhailov and Tsybakov [MiT81], Molle [Mol82], Berger and Tszan [BeT85] have contributed significantly towards the development of a theoretical basis for system capacity of a radio channel. The channel capacity of three multiple access schemes viz., FDMA, TDMA and CDMA is compared under the assumption of both forward and reverse link power control. K.V. Ravi [Rav94] has also contributed towards this effort. The channel capacity is one of the parameters that we have taken into consideration while comparing the two protocols. Our measure of system capacity is based on the service that an ISP is willing to offer. Thus, we have presented comparative graphs of queuing delay vs. number of users and queuing delay vs. system throughput for both the protocols. This would thus give a broader picture to the ISP in deciding the operating point while deploying the protocol.

Le, Babak and Aghvami [LBA98] have carried out comparison of TDMA and R-TDMA. This is useful in understanding the behavior of a dynamic centralized scheme that has been used for this evaluation. This also allows us to comment upon the ad-hoc behavior of the protocol.

The ongoing efforts to improvise MAC protocols are dauntingly large to be mentioned here, as each of them cater to a specific constraint observed in the protocol. Though, significant work has been done in comparing and evaluating the performances of TDMA, FDMA, CDMA protocols, not much has been done in identifying the MAC protocol that is more suited to a particular application type. In this report, we have focussed our attention on this issue and have identified a mapping between a MAC protocol and an application type. In this respect, this report is oriented towards providing an ISP with performance chart based on empirical results. Though the design improvement suggested in this report addresses only the issues of two specific protocol architectures, the concept is still general in nature and can be applied to other protocols as well.

Chapter 3

Test Scenarios

3.1 Nature of Applications

Efficient and robust deployment of a WLL system hinges on the estimates of user traffic patterns. Since the primary goal of a WLL system is to provide high speed Internet access to the end user, it is imperative that the traffic patterns of the users be studied closely. For example, consider a system that has been designed to support a large user population with the tradeoff being an overall higher average queuing delay. If a user runs voice based applications over such a system, performance would be quite poor and nevertheless annoying. Even though the link capacity of the "last mile" is not as high as that of the backbone telecommunication networks, the system can be designed to handle various traffic conditions if the design is based on traffic estimation and forecasting. The MAC protocol design can be scaled and/or changed to accommodate the user's requirements.

Applications can be classified on the basis of constraints posed by their usage or on the symmetry of traffic flowing upstream and downstream. The former class includes voice, video and other real-time applications which have a maximum bound on the delay and jitter. Currently system designs are being tested in order to support such applications. Such systems would require a QoS-aware scheduler at the central node. The other method of classifying applications takes into account the volume of traffic flowing in either direction. It can be further classified as s*ymmetric* or a*symmetric*. *Symmetric* applications have equal volume of traffic flowing in both directions and hence rely heavily on the fairness of the MAC protocol to achieve this symmetry. Applications like Chat, E-mail etc. belong to this category. *Asymmetric* applications have traffic dominant in one direction and short requests or acknowledgements are sent in the opposite direction. Applications that involve file transfers and Internet browsing are such examples. Researchers have put in a lot of effort to build a mathematical model or distribution that would closely resemble the existing traffic conditions on the Internet. Though there is no standard model that represents the Internet traffic accurately, it has been observed that the Internet traffic is mostly asymmetric in nature [Cla98]. There is always more traffic in the downstream direction than in the upstream. Traffic measurement studies conducted on the Internet also show that TCP packets have maximum occupancy on any given link and that HTTP traffic dominates the available bandwidth, followed by FTP and Simple Mail Transfer Protocol (SMTP). Since any WLL service provider would take into account these factors while designing the system we have based our tests mainly on HTTP and FTP traffic patterns and observe the performance of the two MAC protocols using these patterns.

3.2 Traffic Patterns

FTP and HTTP use TCP as their transport mechanism. Since TCP is a connection-oriented protocol (i.e. it requires an explicit setup and teardown phase for communication to take place), it relies on the acknowledgements received for each packet that it sends. There exists a bound on the delay before which the expected acknowledgement must arrive. The underlying MAC protocol should be designed such that it accommodates the TCP requirement. In the following sub-sections we would be discussing about the various types of traffic conditions and their standard patterns. The standard pattern here refers to the output that has been generated using a wireline system with a single host and server. Data rate on the wireline system is same as the maximum data rate that can be achieved using QPSK modulation on the available radio bandwidth. The graphs in the following sub-sections have been generated using OPNET™.

3.2.1 FTP Traffic Pattern

OPNET™ classifies FTP traffic into various categories, each specifying a particular set of parameters. Each parameter can be explained in brief as follows.

- *Command Mix*: This parameter specifies the percentage of "get" commands executed to the total commands executed during each FTP session. The remaining commands are "put" commands. "Get" command refers to a "download" on part of the user, while a "put" command refers to an upload on part of the user. Hence, we can also specify this parameter as percent upload and percent download.
- *File Transfer Rate*: This parameter (measured in files/hour) specifies the average number of files transferred in one hour of a given FTP session.
- *Average File Size*: This specifies the average size of the file transferred in bytes.

We can thus categorize the different FTP traffic patterns as:

1. FTP Low Download

FTP Low Download is characterized by the following parameters:

Command Mix: 100%

File Transfer Rate (Files/hour): 1.0

Average File Size (bytes): 10,000

Fig 3.1. shows the server load on the Y-axis, which can be measured in terms of sessions/sec and requests/sec. Each FTP session is initiated by a user request and eventually leads to a session between the client and server.

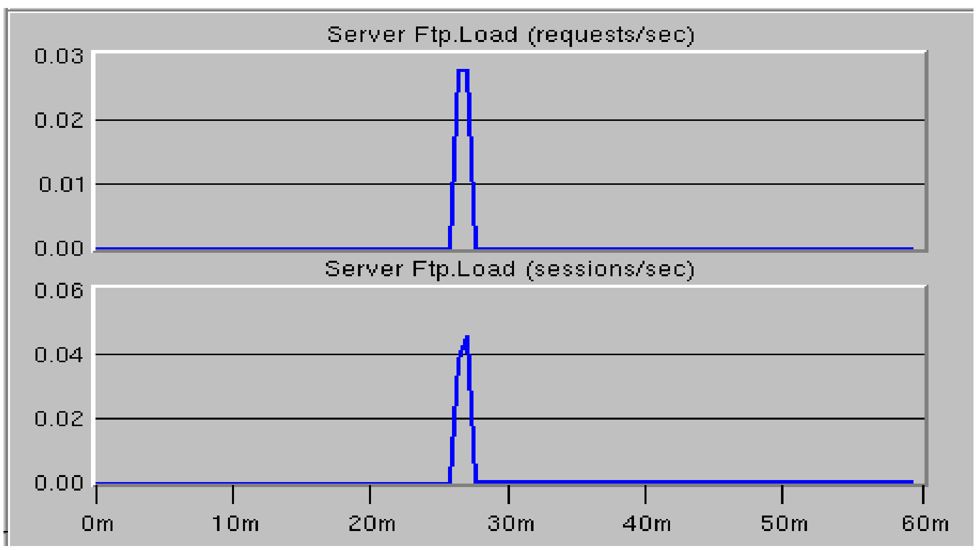

Fig. 3.1. FTP Low Download - Server Load

The server load is a measure of the load conditions on the FTP server. The "processing time" (i.e. the time required to process any service request) depends upon the server load. Higher the server load, higher is the processing delay and consequently the end user queuing delay. The X-axis represents the time in minutes over which the simulation has been performed.

Fig 3.2. shows the average data sent and received by the server. In case of a single user the traffic received by the server is same as the traffic sent by the user.

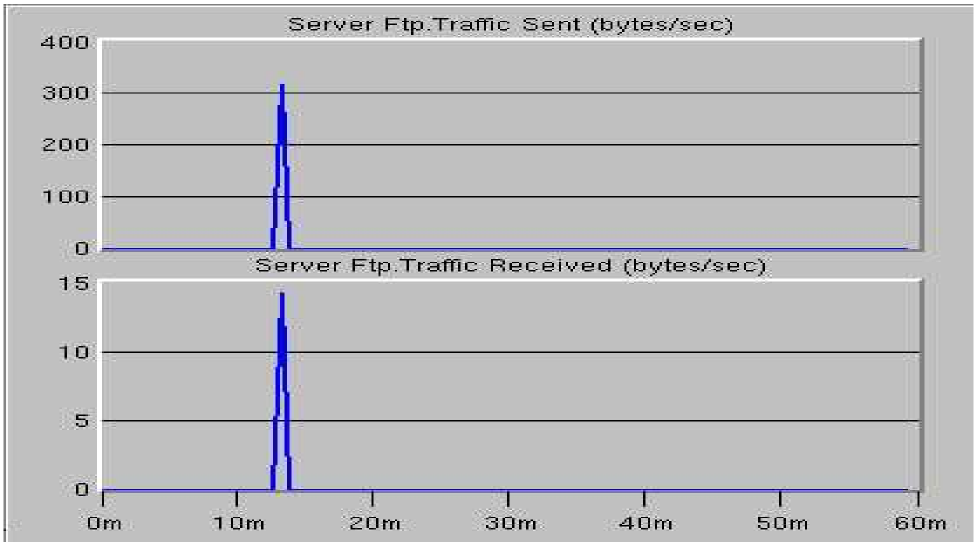

Fig. 3.2. FTP Low Load – Server Traffic Received and Sent

Fig. 3.3. shows the processing delay for each request received by the server.

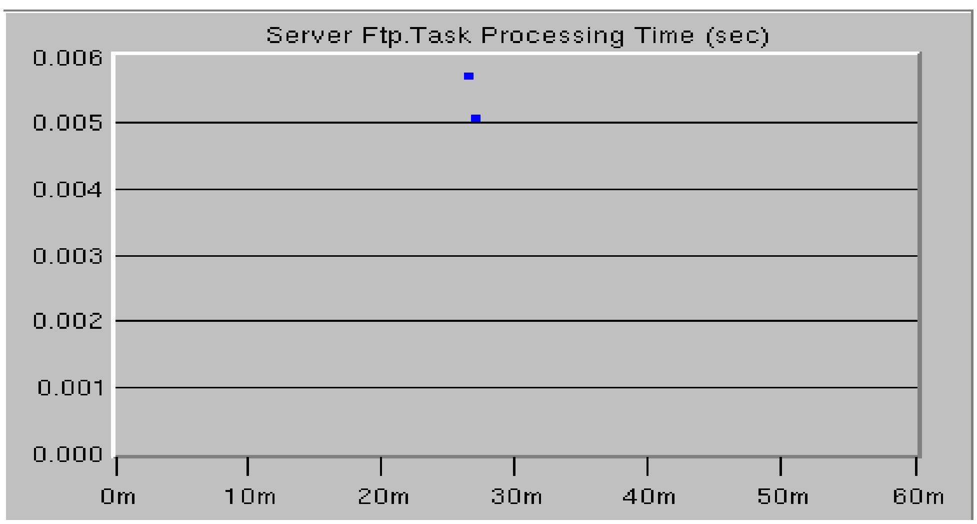

Fig. 3.3. FTP Low Load - Server Task Processing Time

Thus, for an FTP session the total file transfer time is a sum of the processing delay and the actual time required to transfer the given file.

The important parameters that can be measured at the client end are the average download response time. The download response time for the FTP session is as shown in Fig. 3.4. The response time is a measure of the total time required to complete the file transfer. This also includes the associated queuing and contention delays.

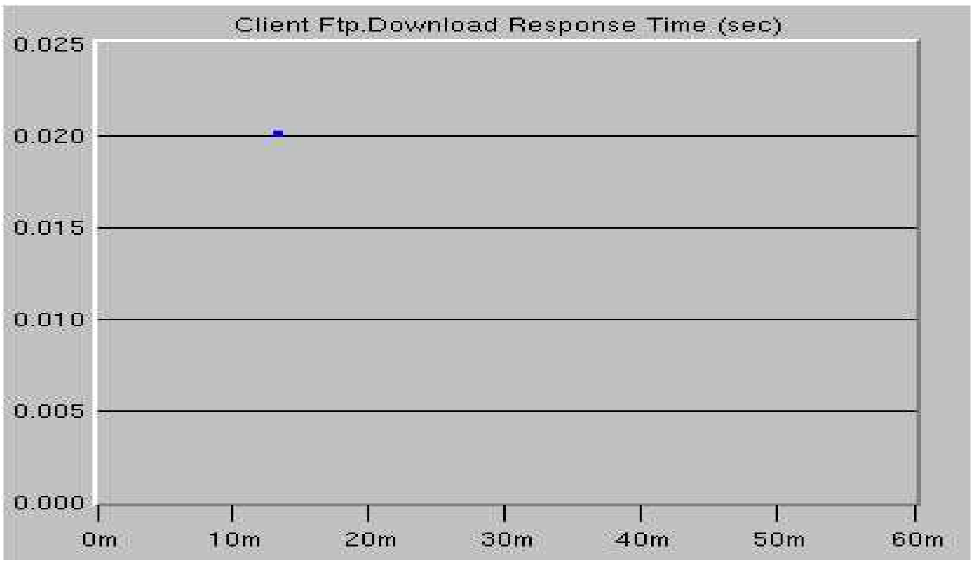

Fig. 3.4. FTP Low Load - File Download Response Time

For a wireless channel the headend has total control over the downstream transmission and users have to contend for transmission in the upstream direction. This slightly increases the values of upload and download response times as compared to a wireline system. Even though the download is controlled totally by the headend, it still has to wait for the upstream acknowledgements to arrive before further transmissions can take place.

2. FTP Heavy Download

FTP Heavy Load can be characterized by the following parameters:

Command Mix: 100%

File Transfer Rate (Files/hour): 10

Average File Size (bytes): 100,000

Fig. 3.5. shows the average server load on the Y-axis in terms of the requests/second and sessions/sec. The X-axis shows the time over which the simulation has been performed.

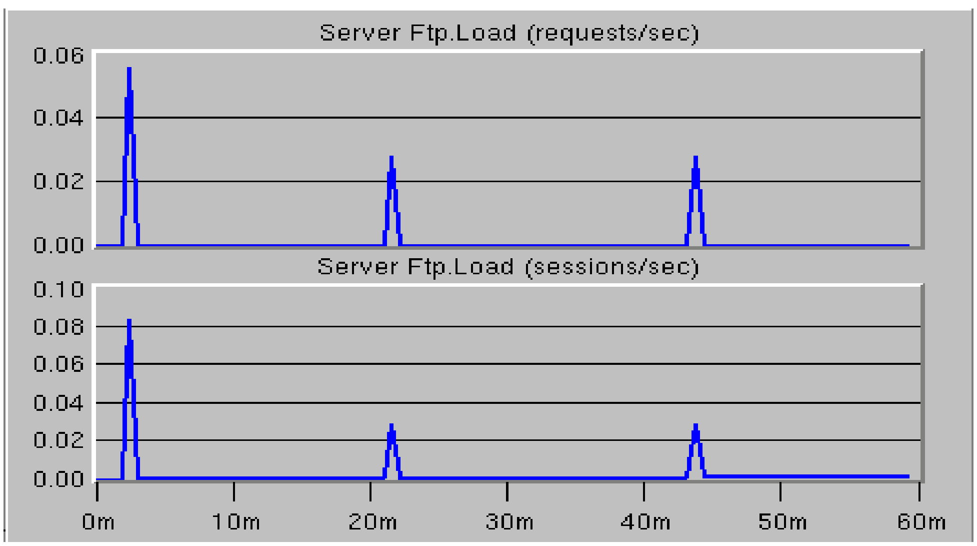

Fig. 3.5. FTP Heavy Load - Server Load

From Fig. 3.5. it can be seen that the average number of sessions and requests is much higher than for FTP low load condition shown in Fig 3.2.

The average traffic received and sent by the FTP server is shown in Fig. 3.6.

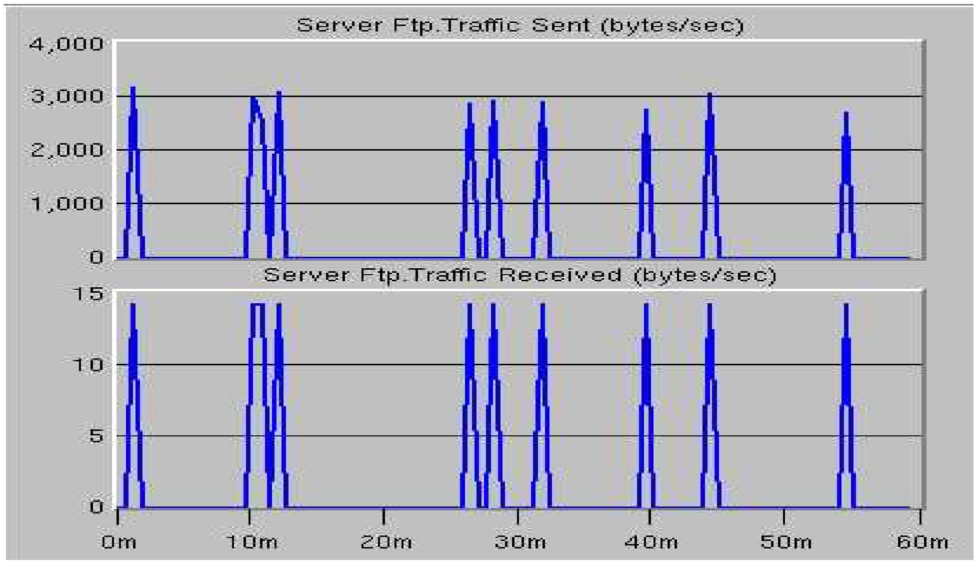

Fig. 3.6. FTP Heavy Load – Server Traffic Received and Sent

The server task processing time is as shown in Fig. 3.7.

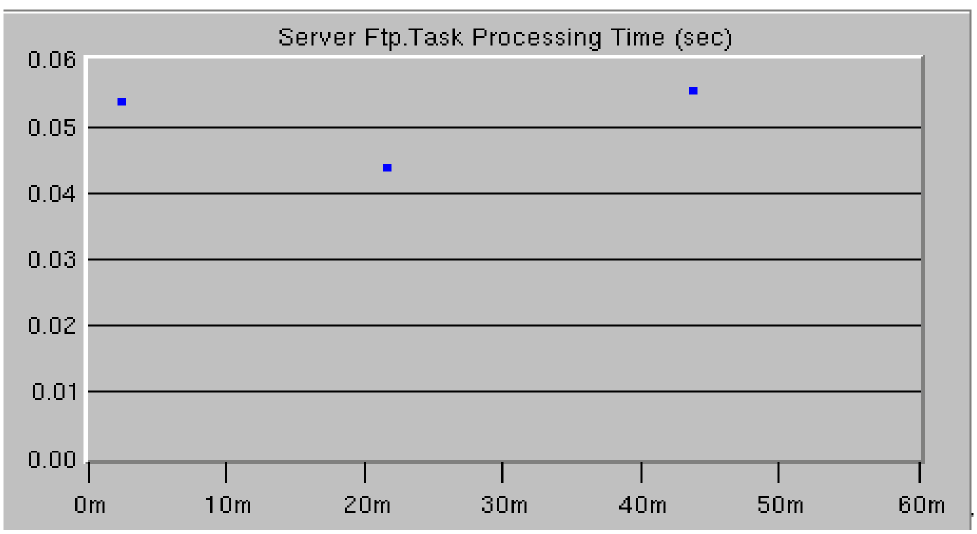

Fig. 3.7. FTP Heavy Load – Server Task Processing Time

The server processing time is higher than that of low load condition as shown in Fig. 3.3. This is due to the nature of requests received.

Fig. 3.8. shows the average file response time for each file download.

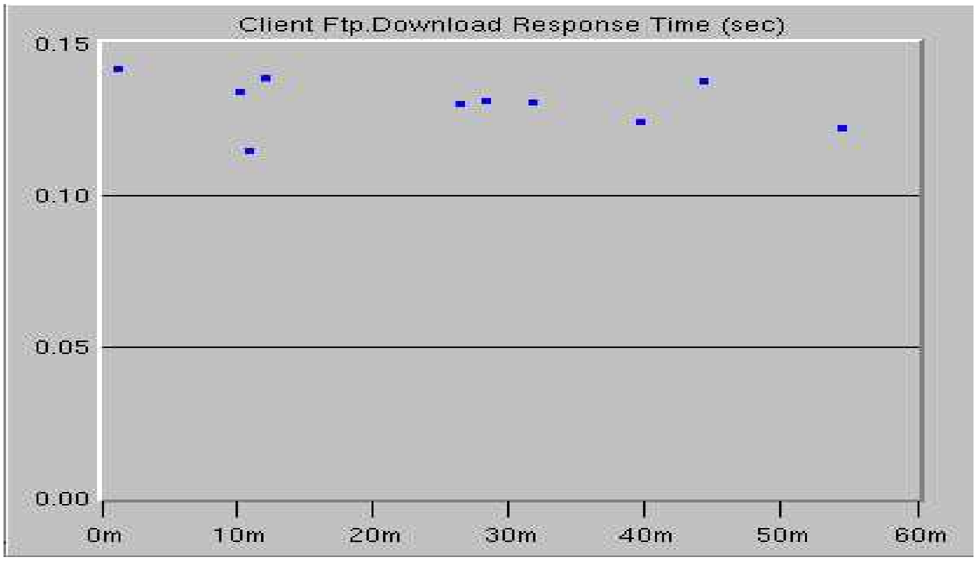

Fig. 3.8. FTP Heavy Load – File Download Response Time

For heavy load conditions, the file size is larger and the download response time is higher than that of low load conditions, which is shown in Fig. 3.4.

3.2.2 HTTP Traffic Pattern

OPNET™ classifies HTTP traffic into various categories, each specifying a particular set of parameters. Each of the parameter can be explained in brief as follows.

- *Page Rate:* This parameter defines the average number of pages downloaded per hour. Each page can be considered to be composed of objects having different sizes.
- *Page Size:* This is a measure of the average number of objects per page. Greater the number of objects, heavier is the download in terms of the page contents and higher the response times.

- *Average Object Size:* This is defined as the average size in bytes of each object associated with a particular page.

We can thus discuss the various HTTP traffic patterns as:

1. HTTP Light Browsing:

HTTP Light Browsing is characterized by the following parameters:

Page Rate (Pages/hour): 5

Page Size (Objects/page): 10

Average Object Size (bytes/object): 12,000

Fig. 3.9. shows the average server load on the Y-axis in terms of sessions/sec and requests/sec. The X-axis shows the simulated time in minutes.

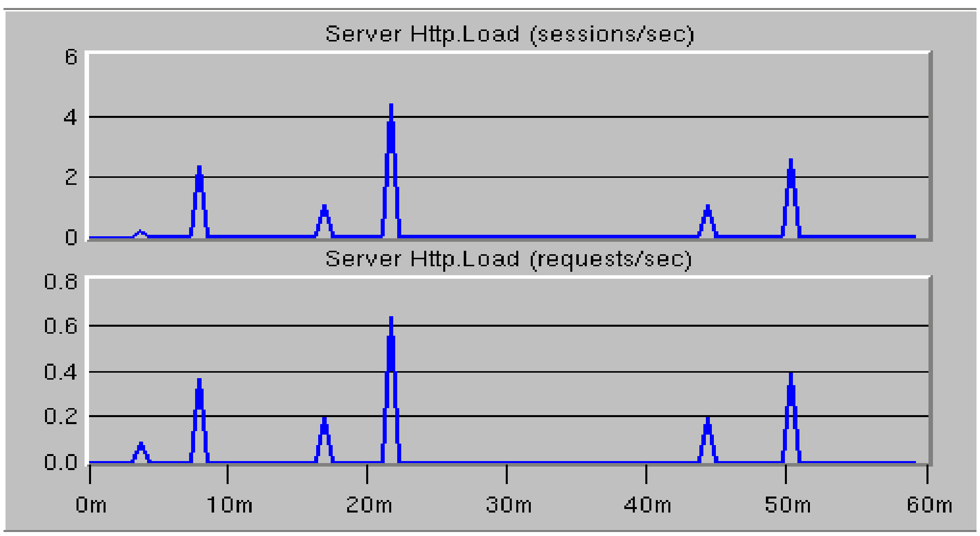

Fig 3.9. HTTP Light Browsing - Server Load

The average traffic received and sent by the server is as shown in Fig. 3.10.

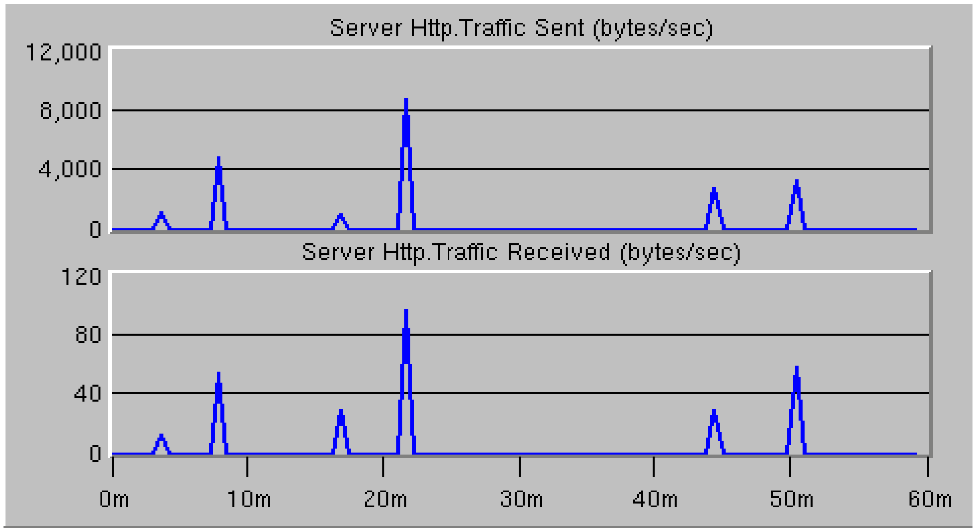

Fig 3.10. HTTP Light Browsing – Server Traffic Received and Sent

It can be seen from the above figure that the traffic sent is much higher than the traffic received by the server. This primarily due to the fact that HTTP traffic is asymmetric in nature with more data in the downstream than in the upstream. The upstream traffic mainly consists of request and acknowledgements.

The server task processing time is shown in Fig 3.11.

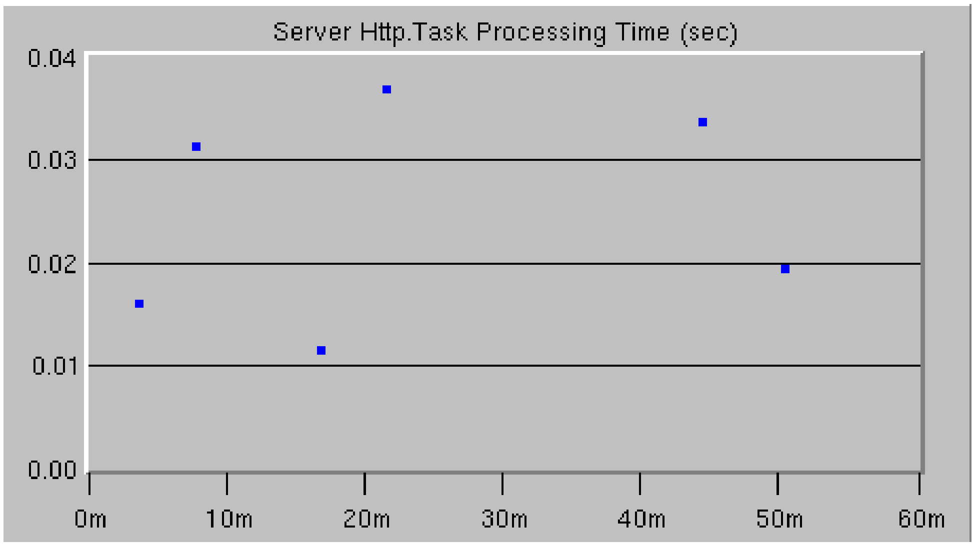

Fig 3.11. HTTP Light Browsing – Server Task Processing Time

The interesting parameters that can be observed on the client side are object response time, page response time. Another important parameter is the time required by the client to download an entire page. This is also a measure of the contention and queuing delay experienced by the user.

The average page and object download time is as shown in Fig. 3.12.

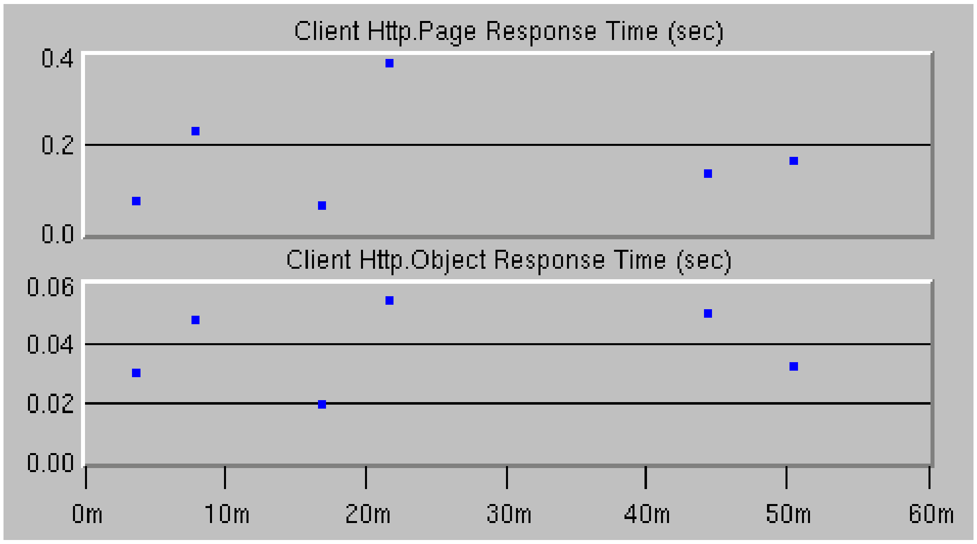

Fig. 3.12. HTTP Light Load – Average Page and Object Response Time

2. HTTP Heavy Browsing

HTTP Heavy Browsing is characterized by the following parameters:

Page Rate (Pages/hour): 60

Page Size (Objects/page): 10

Average Object Size (bytes/object): 12,000

It should be noted that heavy browsing differs from light browsing only in the Page Rate. Fig. 3.13. shows the server load on the Y-axis in terms of sessions/sec and requests/sec. The X-axis shows the simulation time in minutes.

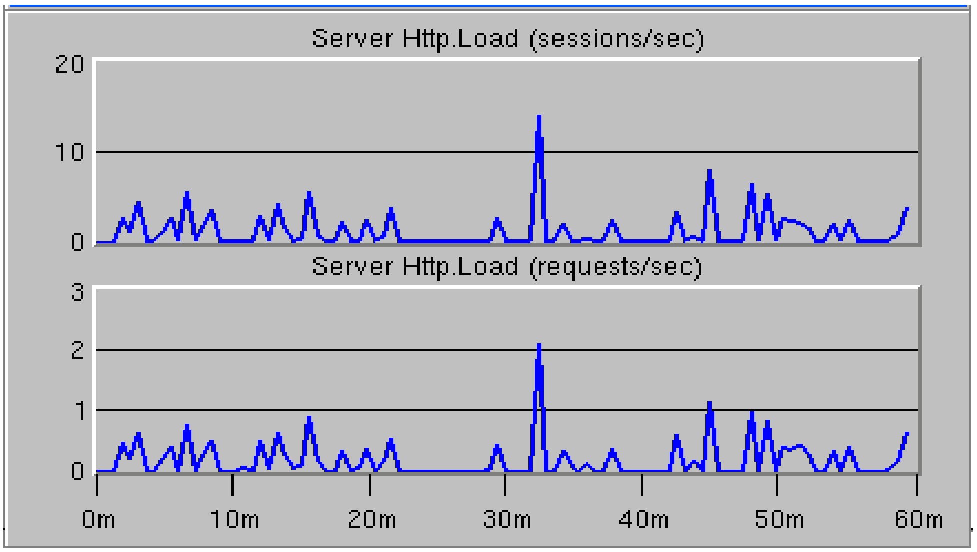

Fig. 3.13. HTTP Heavy Browsing – Server Load

Fig 3.14. shows the average traffic sent and received by the server.

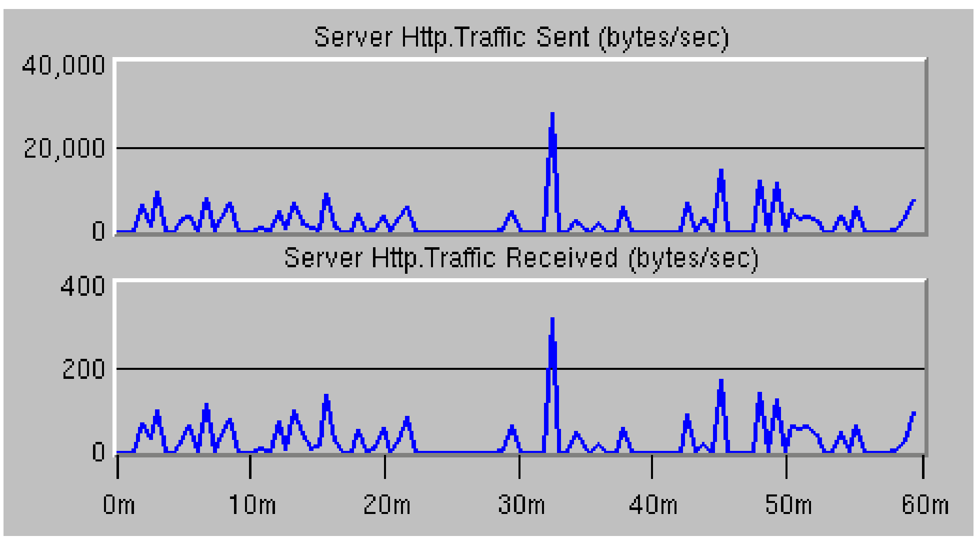

Fig. 3.14. HTTP Heavy Load – Server Traffic Received and Sent

The server task processing time is as shown in Fig. 3.15. It can be seen that increasing the load does not necessarily increase the processing time. This is because the requests are spread over time.

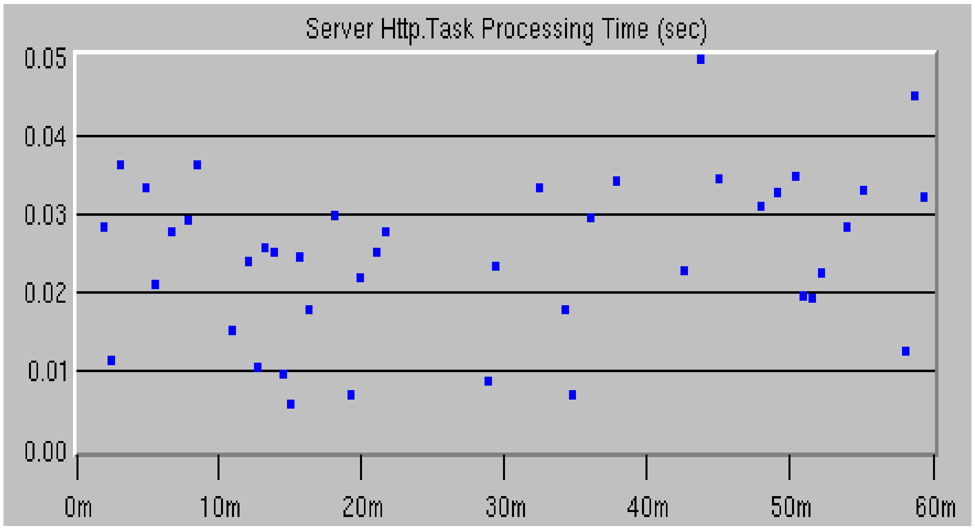

Fig. 3.15. HTTP Heavy Load – Server Task Processing Time

On the client side, we can observe the same parameters as in case of HTTP Light Browsing. Fig. 3.16. shows the relevant graph for the response time.

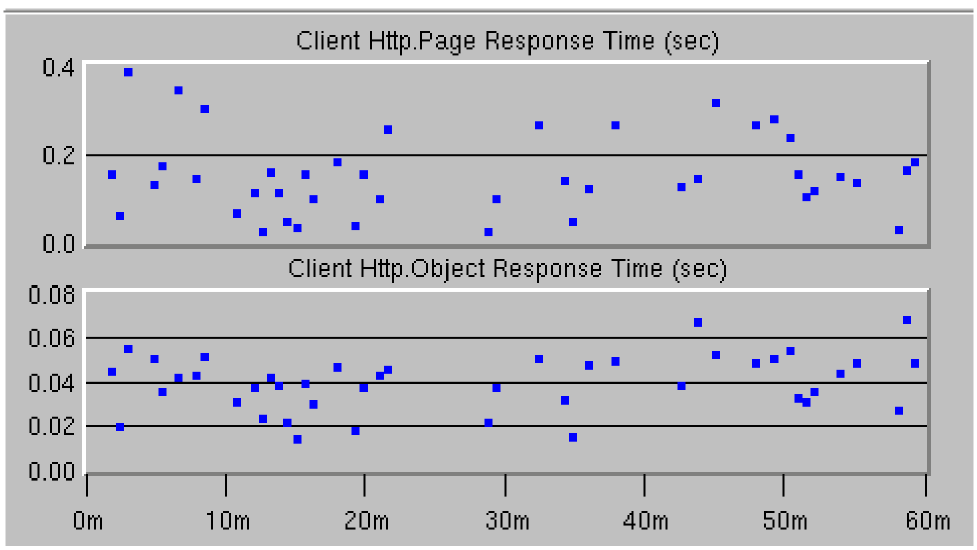

Fig 3.16. HTTP Heavy Load – User Page and Object Response Time

There also exist other traffic patterns that are combinations of FTP and HTTP patterns and hence have performance characteristics which are hybrid of these two patterns. We have used the patterns discussed earlier for all the tests conducted, the difference being that our tests were conducted over a wireless channel with a gradual increase in the number of users.

3.3 Bound Measurements

Apart from the tests discussed in sections 3.2.1 and 3.2.2, other tests were also conducted to assess the upper bound on a particular protocol's performance in terms of throughput and delay. In order to achieve this, a queuing model was built to emulate the MAC layer and data was bursted into the queue at a rate marginally equal to the channel link rate. These models (known as "node models") were developed for both the client and server. Since the models were built to emulate the real world system, their physical locations were scaled according to the real distances. A maximum separation of 5 Km. was used between the client and server during the course of simulation. The setup is as shown in Fig. 3.17

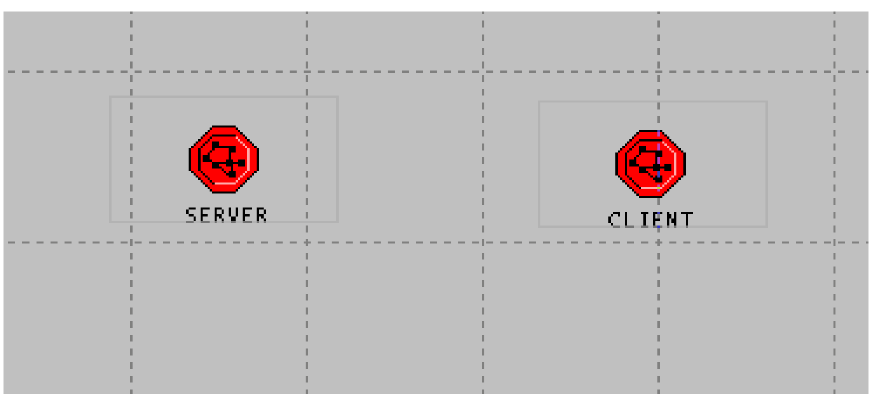

Fig. 3.17. Client-Server Setup

The above figure shows a two dimensional view of the placement of the server node and the client node which emulates a real world scenario. The grid can be viewed as an XY plane with the location of the two nodes being marked by their XY coordinates. The axes are calibrated in Km. Thus, the client and the server node can be placed such that their separation is 5 Km. The server node consists of the headend radio and an Ethernet/ATM server. The client node consists of an Ethernet/ATM workstation and can the host radio. This is shown in Fig.3.18. and Fig 3.19.

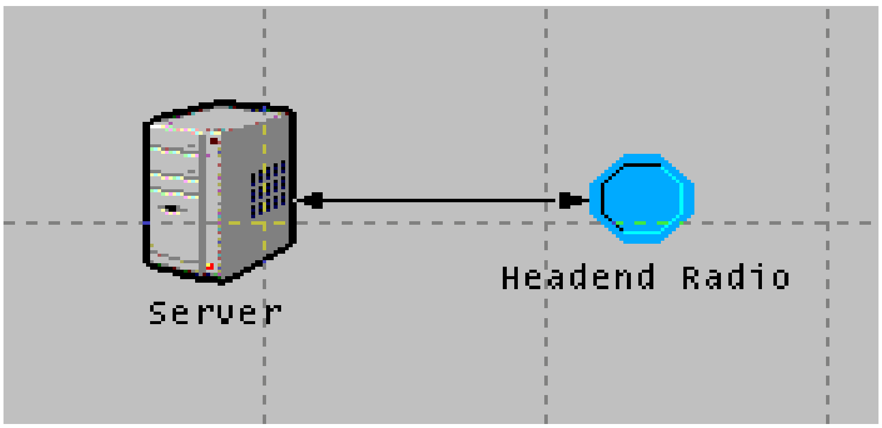

Fig 3.18. Server Node Setup

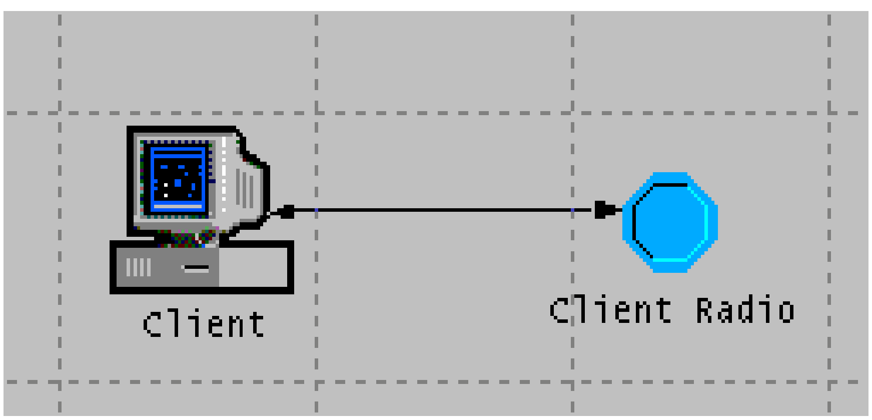

Fig 3.19. Client Node Setup

The headend and client radios differ only in their functionality. It can be shown as in Fig 3.20. The radio basically consists of a pair of transmitter and receiver of which one is used for communication on the radio channel. It can be configured for the given data rate, modulation scheme, operating frequency and bandwidth. The other pair is used for communication with the corresponding server or client workstation.

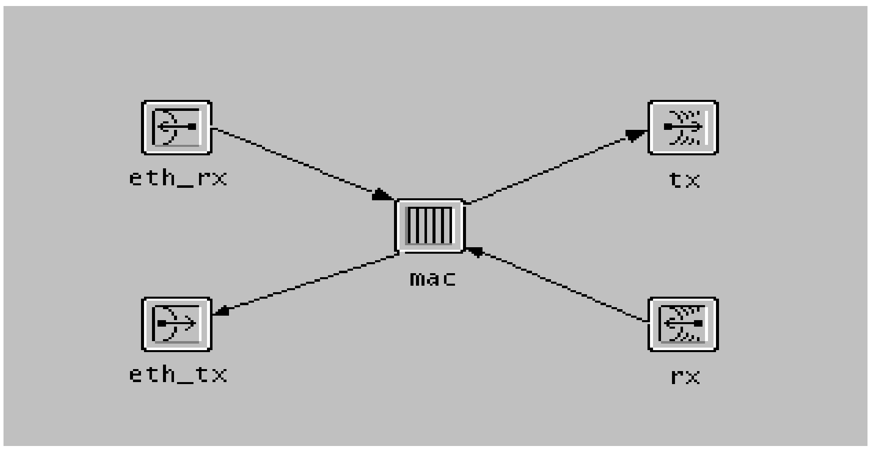

Fig. 3.20. Radio Node Setup

The MAC layer has been modeled as queue with infinite capacity for simulation purposes. It possible to replace the server and client shown in Fig. 3.18. and Fig. 3.19 by simple packet generators with a suitable inter-arrival pdf and inter-arrival rate. The inter-arrival pdf defines a distribution for packet arrivals into the queue. We have used exponential distribution for our simulation. When the inter-arrival rate is made equal to the link rate on the channel the MAC layer always has enough data to transmit and utilize the available channel fully. Thus there are no lull periods in transmission and the observed throughput is the maximum throughput achievable on the channel with that particular MAC protocol. Also, the queuing delay is maintained at its minimum hence the lower bound on the queuing delay can be observed. This measurement gives the upper bound on the system performance. Beyond this the system performance will only degrade.

Thus for our performance evaluation we have conducted the tests discussed in sections 3.2.1, 3.2.2 and 3.3 as these tests show the output performance for varying load conditions and extreme conditions. This would enable system designers and service providers to decide upon the parameters critical to their operation. These tests also help us to analyze the strengths and weaknesses of a given MAC protocol and improvise its design to achieve better performance.

Chapter 4

Test Results

4.1 Overview

This chapter discusses the results of the tests that were conducted for evaluating the performance of the two MAC protocols. These tests included the following:

1. **Packet Generator Test:** A simple packet generator delivers packet to the MAC layer such that the inter-arrival rate equals or is greater than the link rate of the channel. This test serves to measure the maximum achievable throughput of the given MAC protocol. Having the same inter-arrival time for the packet generators at the host as well as headend, symmetric traffic is ensured in the upstream and downstream channels.
2. **FTP Low Download:** Download of a file by the user serves as the load for this test. The file size suggested by OPNET™ is 10 KB. The number of users in the system is gradually increased from one to seventy. A good estimate of the performance can be obtained by running the simulation for this user range.
3. **FTP Heavy Download:** This test is similar to FTP Low Download test except for the fact that the file size is increased to100 KB.
4. **HTTP Light Browsing:** The characteristics for this test are mentioned in Section 3.2.2.

5. **HTTP Heavy Browsing:** Heavy browsing by the user serves as load for this test. The test characteristics are mentioned in Section 3.2.2.
6. **Medium Load:** This test uses an equal share of FTP and HTTP traffic for the duration of operation. The FTP traffic has characteristics of FTP Low Download traffic, while the HTTP traffic has characteristics of HTTP Light Browsing traffic.

This test suite fairly covers the user traffic patterns existent on the Internet today. For all the tests from 2 to 6 we observe the same output parameter viz. aggregate throughput and average queuing delay. Additional tests that have different load conditions can always be performed to observe the output performance.

4.2 Protocol Performance

4.2.1 Packet Generator Test

Fig. 4.1 shows a comparative graph of the maximum achievable throughput on the given channel that has a bandwidth of 12 MHz and uses QPSK as its modulation scheme.

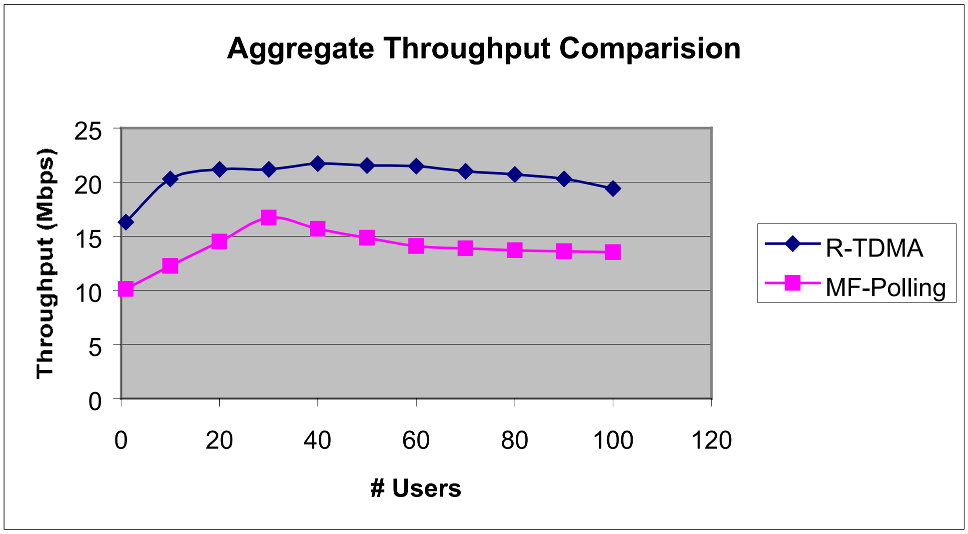

Fig. 4.1. Aggregate Throughput Comparison – Packet Generator Test

From the above figure it can be seen that R-TDMA gives better throughput values than MF-Polling. This can be attributed to the basic architecture of each protocol. In case of R-TDMA protocol, as the number of users is increased, more and more slots in each frame are utilized. This increases the efficiency of the frame and thus the aggregate throughput. The increase in frame efficiency is bound by the number of allowable data slots in the frame. When the number of data slots used reaches it maximum value the efficiency and thereby throughput too reaches it maximum value. Addition of users to the system after this condition still maintains the throughput value constant. The aggregate throughput also depends upon the performance the of contention protocol. Our implementation of R-TDMA uses Slotted ALOHA with exponential backoff as the contention protocol. The efficiency of this contention protocol decreases as the number of users is increased. This leads to degradation in throughput for higher user values in the range from 80 to 100 as seen from Fig. 4.1. Thus, the advantage of increased frame efficiency for larger frames is masked by the decrease in contention efficiency. The degradation is graceful in nature because of the backoff mechanism used for contention. The backoff mechanism spreads the transmission of user requests over multiple frames thereby uniformly distributing the number of requests. Also piggybacking of request with data reduces the number of users contending in each frame. This also helps to maintain the frame efficiency at its maximum value. Thereby the aggregate throughput remains fairly constant over the range of 20 to 80 users in the system.

In case of MF-Polling, not all channels are used for data transmission during the operation time. This is because of the fact that few of channels are reserved for users who are in a bursting mode. This mode requires the user to have data equivalent to one full sized Ethernet frame in its queue. If the user does not have enough data it remains in polling mode where it can send only one data packet at a time when polled. Thus, of the available 12 MHz bandwidth around 1 MHz is utilized only when required. This is the drawback of the MF-Polling protocol architecture. All channels are uniform in nature i.e. each can support up to 256 kbps data rate. One channel is dedicated solely for contention purposes. Exponential backoff is used as the contention scheme. This system does not use any reservation mechanism to conserve bandwidth as in case of R-TDMA. Users are removed from the system based on their data inactivity. If a user remains inactive for period greater than the stipulated number of polling cycles, it is pulled out of the system. The user has to contend again in order to re-enter the system. Thus the throughput is also affected by the inactivity time which is a multiple of the polling cycle time. This also affects the average queuing delay of the system. In order to send a request on the contention channel the user has to wait for an integral number of broadcast polling cycles. R-TDMA has the highest frame time of the order of 3 milliseconds, while MF-Polling has a polling cycle time of 30 ms. The tradeoff here is the number of users supported by both the protocols. The drawback with R-TDMA is that the system attains maximum efficiency when the number of users equals the number of slots. The throughput degrades if more users enter the system. Thus we observe high throughput and low delay for lower user population and gradual degradation in the throughput and delay as the population increases.

For MF-Polling system, number of users supported is decided by the product of number of polling channels and maximum number of users per channel. The maximum number of users per channel is decided by the allowable polling latency of the system. Thus, a MF-Polling system can support a large user population at a sustained value of throughput as compared to R-TDMA system that can support a smaller user population at throughput value higher than that of MF-Polling. As the number of users is increased for the MF-Polling system each of the channels is sequentially occupied. The throughput increases as each channel has only one user and is thus not polled. The throughput reaches a peak value when the number of available channels equals the number of users in the system. As users are continuously added the channels, each channel gets converted from a dedicated channel to a polling channel. The throughput also gradually decreases as all the channels get converted to polling channels. Thereafter the aggregate throughput remains constant till the number of users in the system reaches the maximum supported value.

In order to compare the protocols fairly, a metric can be obtained from Fig 4.1 that will take into account the large user population supported by MF-Polling and high aggregate throughput delivered by R-TDMA. One such metric that can be suggested is the product of aggregate throughput and user population. A comparison of the two protocols based on this proposed metric is as shown in Fig. 4.2

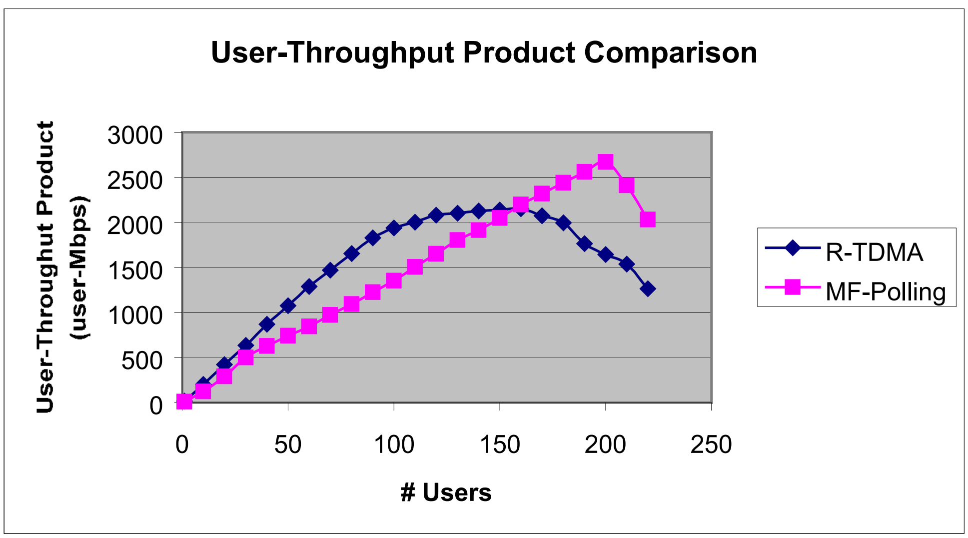

Fig. 4.2. User-Throughput Product Comparison

As the number of users in the system in gradually increased, the R-TDMA product value undergoes a graceful rise and fall. The MF-Polling product value tends to increase linearly as the number of users is increased. This continues till the system reaches its maximum user capacity, after which the product degradation is very sharp. If allowed to continue, the MF-Polling curve will again intersect the R-TDMA curve within a span of few tens of users. Thus, there evidently exists a range of values over which a given protocol performs better than another. However, the tradeoff is increased value of average queuing delay. If a User-Average Queuing Delay product is plotted against the number of users, we would observe constantly upward rising curves for both the protocols wherein the MF-Polling curve having higher values than that of R-TDMA. As MF-Polling has an associated fixed latency with each polling cycle, the average queuing delay increases as the number of users is increased. The same is not the case with R-TDMA, wherein the average user queuing delay is kept in check by allocating only one slot per user per frame.

From a simulation standpoint, it is extremely difficult to maintain the packet generator inter-arrival rate equal to the link rate. This leads to large queue build-up and consequently large associated queuing delays. Hence, the queuing delay prediction is based entirely on the theory and architecture of the MAC protocols under consideration.

Thus, though MF-Polling performs better for a range of users, it suffers from large average queuing delays. A suitable operating point must be chosen by a system provider in order to provide the best of both values. It must also be noted that in this test, the upstream and downstream traffic is independent. The scenario is different when TCP based applications are used. TCP has a finite limit on the maximum time that it waits before it re-transmitting a packet. Thus, the underlying MAC protocol should ideally have its maximum average queuing delay lesser than that of TCP maximum acknowledgement delay.

4.2.2 FTP Low Download

Fig. 4.3. shows a comparative graph of aggregate throughput versus number of users in the system for the two protocols supporting FTP Low Download traffic.

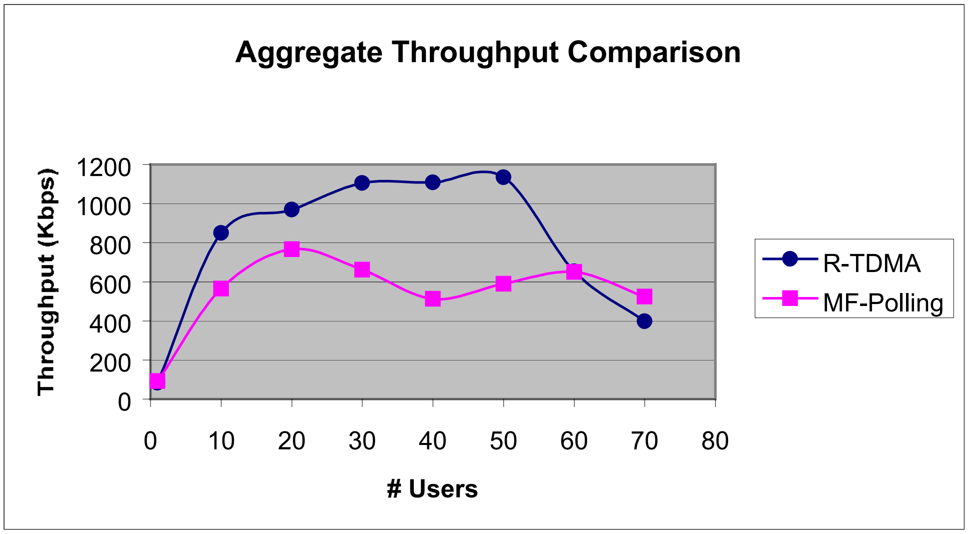

Fig. 4.3. Aggregate Throughput Comparison – FTP Low Download

Low load implies that over the simulation time period data burst periods are much smaller than lull periods. Data transfer is sporadic in nature. Hence, for R-TDMA system most of the frames have little or no data. Thus, at any point of time there are more users who are idle compared to the number of users who are transmitting data. As the user population increases, so does the percentage collision, which leads to degradation in throughput for higher user population. In case of MF-Polling, as the number of user increases, each user is assigned a dedicated channel. Hence, we observe an initial increase in throughput till the number of users equals the number of channels. The throughput tends to decrease as the dedicated channels are converted to polling channels. It must also be noted that the channels, which are reserved for bursting users, are utilized as the number of user increases. This leads to a gradual rise in throughput for higher user values. However, the number of bursting channels is not sufficient to accommodate the traffic generated by large user population and hence, most of the users tend to remain in polling mode. This again leads to a decrease in throughput after a certain range of user population.

The primary reason for MF-Polling protocol's throughput to be lower than that of R-TDMA is the large value of polling cycle time and the limitation of allowing a user to transmit only a single packet for each polling cycle. Also, the inactivity timeout associated with a user being in the polling mode is high compared to the time in which a user actually transmits data. This leads to a user occupying a channel without transmitting any data, which severely affects throughput.

Fig. 4.4. shows a graph of average queuing delay versus number of users for MF-Polling and R-TDMA respectively.

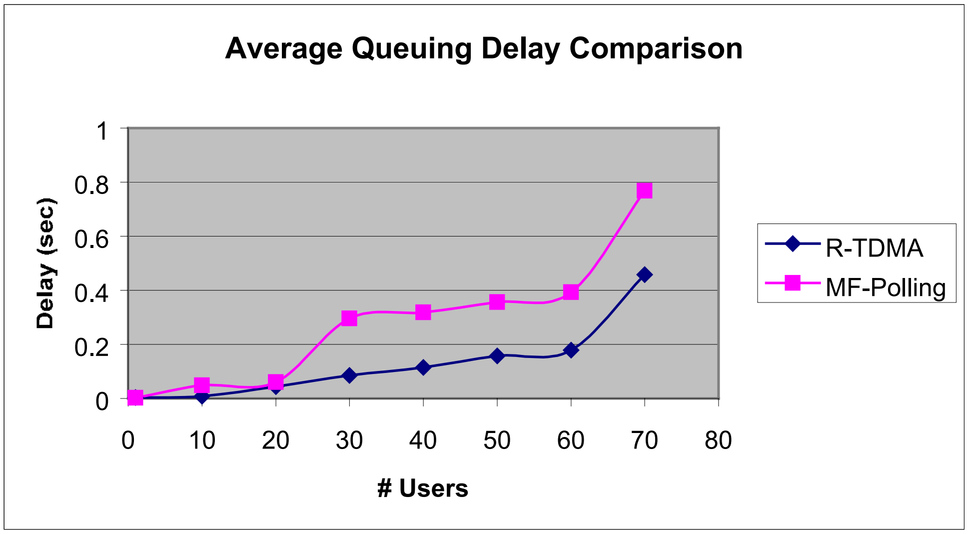

Fig. 4.4. Average Queuing Delay Comparison – FTP Low Download

It can be seen from Fig 4.4. that as long as the number of users in the system is less than the number of channels, the queuing delay of MF-Polling system is comparable to that of R-TDMA. However, the queuing delay increases rapidly once the system moves the users to polling mode. This increase in queuing delay results in an equivalent decrease in throughput as shown in Fig. 4.3. This is true even for R-TDMA system, though it experiences delay which is nearly half as that of MF-Polling.

4.2.3 FTP Heavy Download

FTP Heavy Download is characterized by the parameters mentioned in Section 3.2. Fig. 4.5. shows a comparative graph of FTP Heavy Download aggregate throughput versus number of users in the system for R-TDMA and MF-Polling.

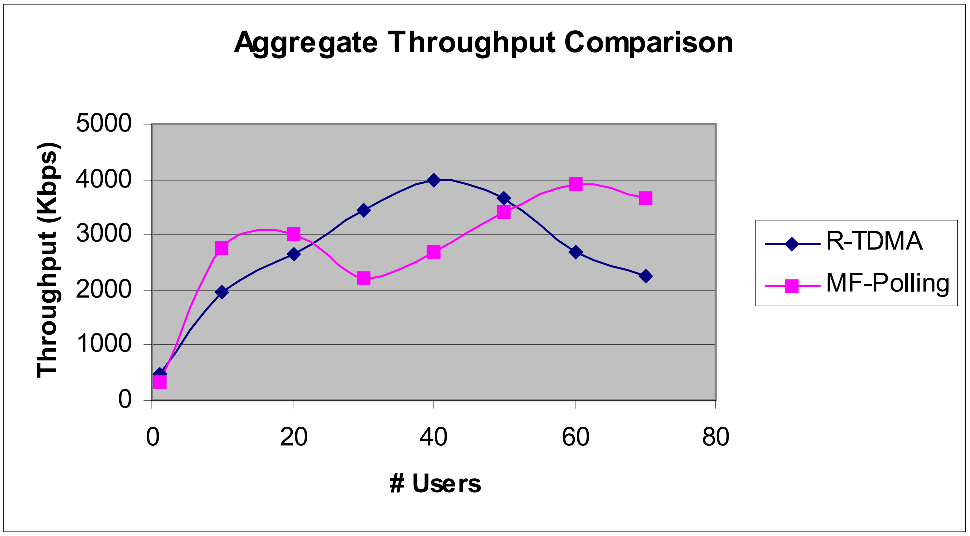

Fig. 4.5. Aggregate Throughput Comparison – FTP Heavy Download

As seen from figure Fig.4.5. R-TDMA has a typical rise and fall type of throughput curve. The throughput values obtained in this test are much higher because of greater amount of data being transferred. This makes the R-TDMA frames more efficient and consequently increases its throughput. However, this protocol faces the problem of high contention delay as more and more users are added to the system. It must be noted that even though the nature of file transfer is downstream in nature, FTP is a TCP based application and hence requires an acknowledgement for every packet sent. The large timeouts associated with MF-Polling for polling channels proves to be an advantage in this case as a user always has enough data to send upstream and is thus not pulled out of the system. This reduces the contention delay experienced by the user for FTP Low Download traffic conditions.

The contention efficiency of R-TDMA system depends entirely on the number of slots available in each frame. The number of slots are varied on a per frame basis such that, one slot is added if collision is experienced in the previous frame and reduced for every non-colliding frame. Users, on the other hand, employ an exponential backoff mechanism and choose a particular frame number from the available window size. The available window size is always less than or equal to the maximum window size. This mechanism however has a subtle flaw. If collision occurs for a user request, the headend increases the contention slot by one. The users correspondingly double their contention window size and randomly skip multiple request opportunities. The headend increases the slot value only for the consequent frame and based on the activity/inactivity increases/reduces the available slots for the subsequent frame. It thus becomes essential that the user should re-contend when the number of slots has been increased. This depends entirely on the random value generated by the user MAC scheduler, which statistically would generate all values lying within the available window size. Thus, a user is more likely to encounter a reduced slot size than an increased one when contending. This severely affects performance for a large user population.

Shown in Fig 4.6. is a comparative graph of queuing delay versus the number of users for FTP Heavy Download conditions.

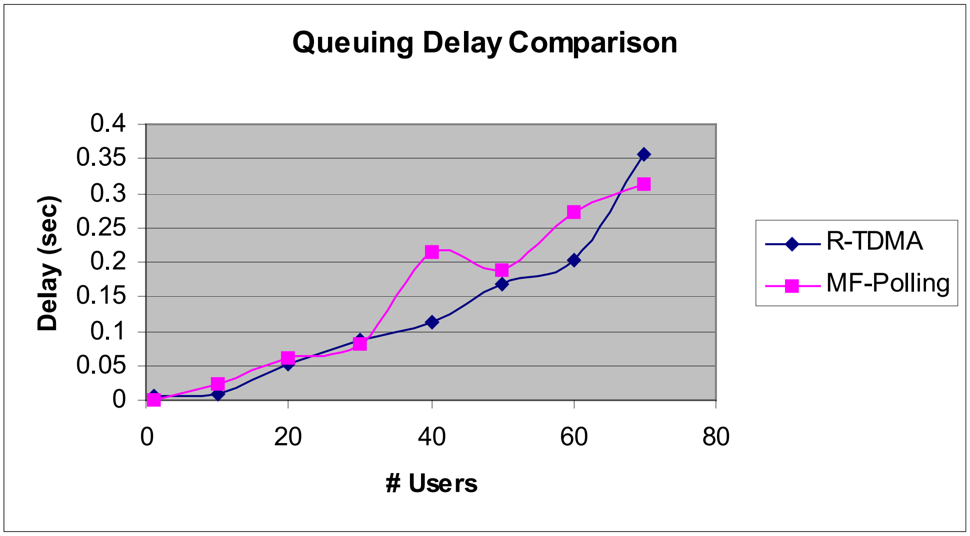

Fig. 4.6. Queuing Delay Comparison – FTP Heavy Download

From Fig.4.6. it can be seen that for both the systems, the increase in queuing delay is approximately linear in nature, and is directly responsible for the degradation in throughput.

4.2.4 HTTP Light Browsing

HTTP Light Browsing characteristics are mentioned in section 3.2.2. HTTP also uses TCP as its transport mechanism. The difference between FTP and HTTP is that the former maintains only one TCP session for each FTP connection, while the latter can maintain multiple TCP sessions for a single connection. This implies that in addition to a large volume of data sent downstream, there also exists a large volume of acknowledgements in the upstream direction.

A comparative graph of aggregate throughput against the number of users for HTTP Light Browsing conditions is shown in Fig. 4.7.

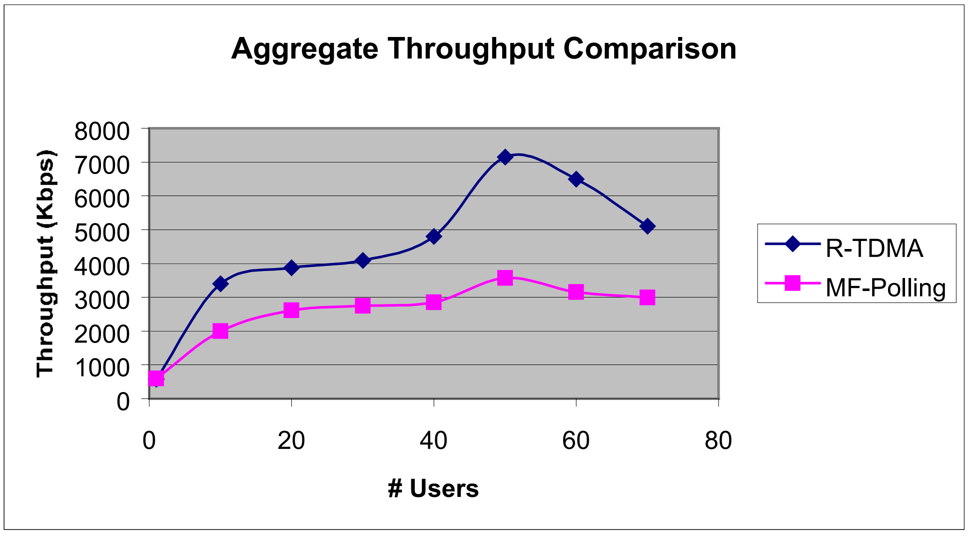

Fig. 4.7. Aggregate Throughput Comparison – HTTP Light Browsing

The above figure clearly shows the effect of symmetric nature of traffic on throughput. In case of R-TDMA system the scheduler is designed to handle symmetric traffic. Because of this, users are able to request for slots upstream by piggybacking their requests on the data. The scheduler assigns only one slot per user per frame. In case of MF-Polling system no scheduling is done in the downstream and data is sent to users in the same order as the headend receives from the application server. R-TDMA sends data downstream in a FIFO (First In First Out) manner. For R-TDMA, as the number of user increases, so does the frame efficiency and consequently the aggregate throughput. The throughput decrease due to decrease in contention efficiency for large number of users is partly compensated by the increase in throughput due to piggybacking of requests. The throughput for MF-

Polling system is limited entirely by the fixed latency associated with each polling cycle. Since our implementation of MF-Polling system uses FEC (Forward Error Correction) for all data transmitted, it has an associated fixed FEC interleaving delay. Such is not the case for R-TDMA wherein ARQ is used. This architectural constraint also affects the average queuing delay as seen from Fig 4.8.

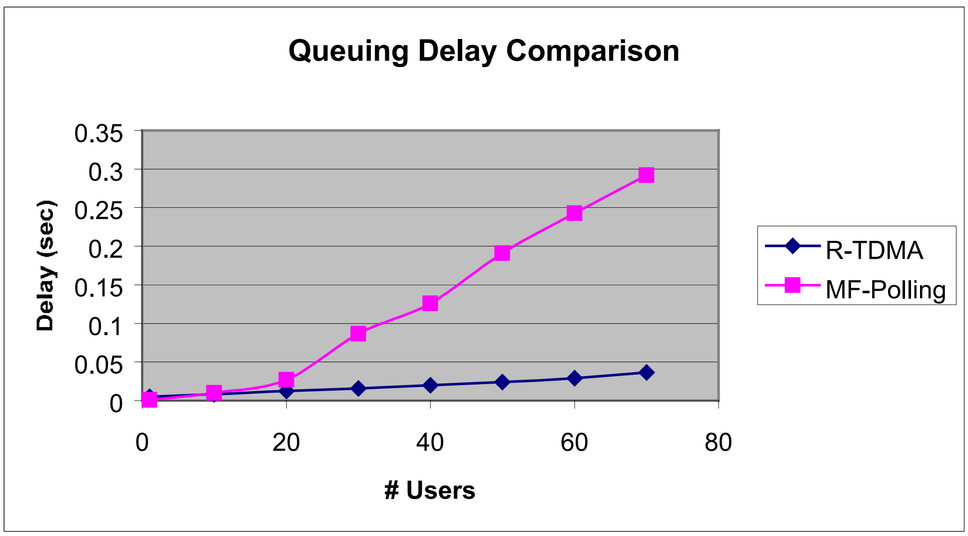

Fig. 4.8. Queuing Delay Comparison – HTTP Light Browsing

Queuing delay consists of two components viz. contention delay and actual wait time before data is transmitted. The latter is the number of frames or slots that a packet has to wait before being transmitted in case of R-TDMA system or is the number of polling cycles in case of MF-Polling. Since MF-Polling system does not have any scheduling mechanism in the downstream users are randomly served which leads to some users being pulled out of the polling channels if they are inactive while other users are busy transmitting upstream. This increases contention in the contention channel. This contention delay is directly dependent on the broadcast cycle time, which is much larger than a full sized R-TDMA frame. Hence, as the user population increases, the contention delay also increases. However, because of piggybacking, the contention delay is greatly reduced for R-TDMA system. The outcome is an overall decrease in average queuing delay and better performance than its MF-Polling counterpart.

4.2.5 HTTP Heavy Browsing

HTTP Heavy Browsing has traffic characteristics as mentioned in Section 3.2.2. Depicted in Fig. 4.9. is a comparative graph of aggregate throughput versus the number of users for the two protocols under consideration

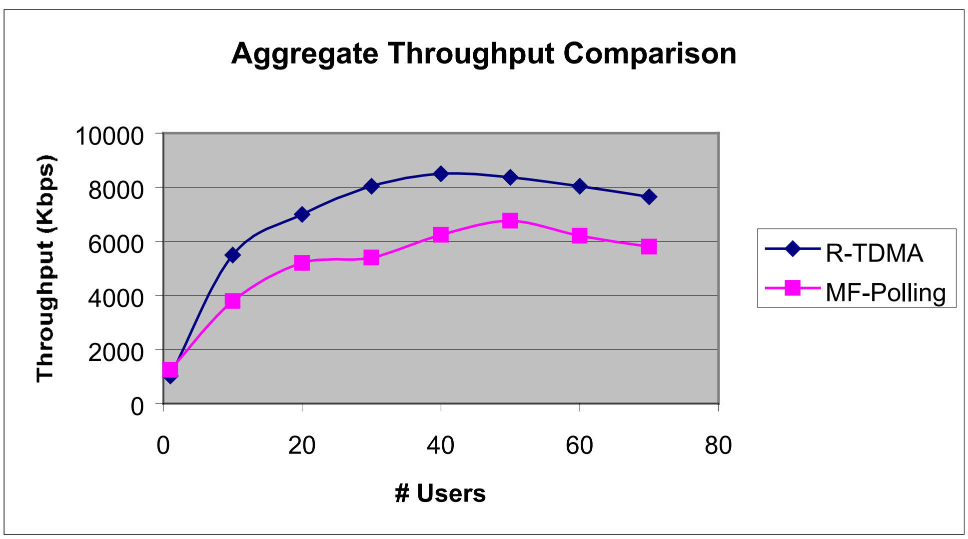

Fig. 4.9. Aggregate Throughput Comparison – HTTP Heavy Browsing

As the load conditions are more severe in this test we see an increased value of maximum throughput as compared to HTTP Light Browsing. The exception in this case being that each HTTP session persists for a longer duration due to a large volume of data. Because of this, for the R-TDMA system the average queuing delay increases, as users who are currently in system tend to reserve the resources (i.e. slots) for more number of frames. This increase in delay is not because of contention but purely because of unavailability of resources. The queuing delay hence becomes comparable to that of MF-Polling, which is as shown in Fig. 4.10.

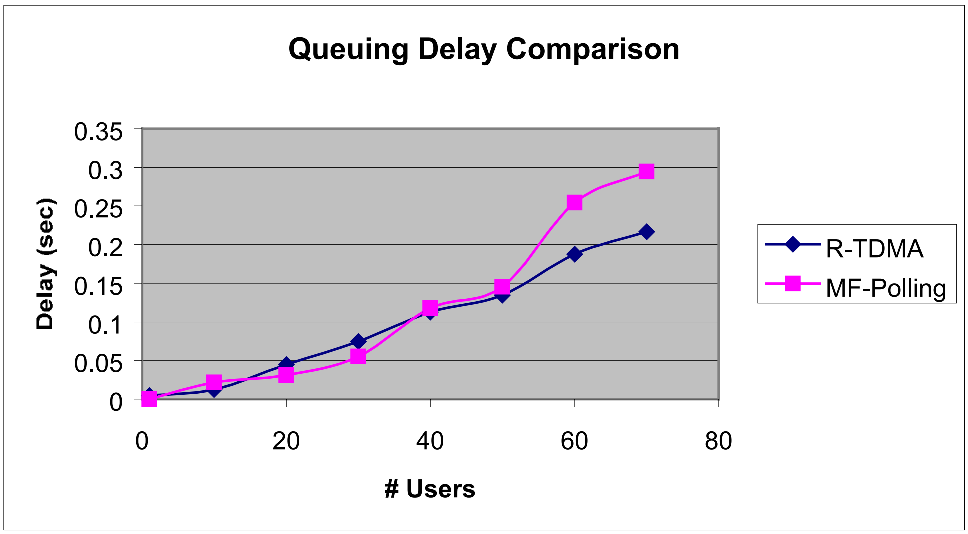

Fig. 4.10. Queuing Delay Comparison – HTTP Heavy Browsing

It can be seen that though the delay in the R-TDMA system is comparable to that of MF-Polling, it still is lower than the latter due to reasons mentioned in the previous test scenario.

4.2.6 Medium Load (FTP Low Download + HTTP Light Browsing)

Combination of FTP Low Download traffic and HTTP Light Browsing traffic can be considered as a medium load on the system. This can be considered to be a medium load on the system. From the previous tests it can be seen that FTP traffic patterns have a sharp rise and fall in its throughput, while it is gradual for HTTP traffic. Thus, it is interesting to observe the system behavior when the traffic is a combination of the two traffic patterns. The overall effect of such a traffic pattern would be a marginal increase in queuing delay as compared to any single traffic type. This is because of the fact that between lull periods of data arrival (as in FTP sessions) there is data in the users queue due to the continuous nature of the other application (in this case HTTP Light Browsing). Such a behavior is depicted in Fig. 4.11., which shows the aggregate throughput versus the number of users in the system for the medium load conditions.

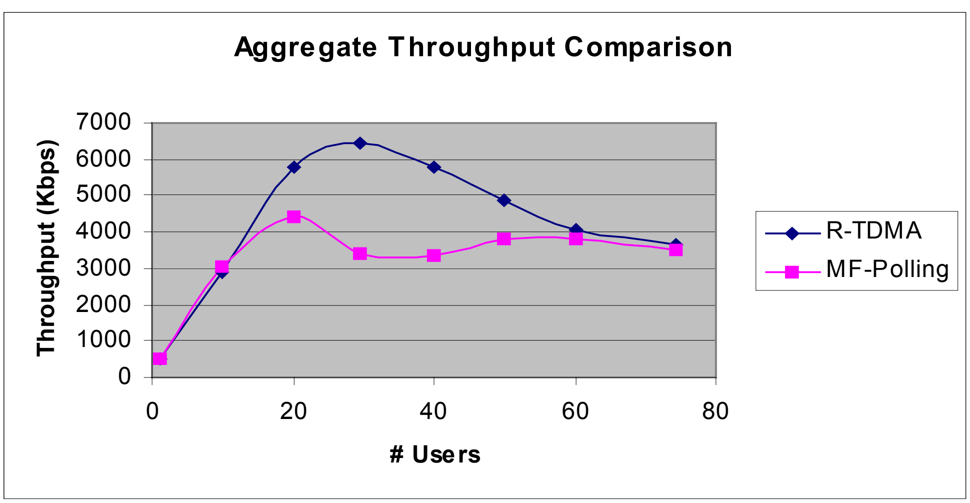

Fig. 4.11. Aggregate Throughput Comparison – Medium Load

It can be seen that MF-Polling is more suited to this combination load as its output remains stable over a large range of users. For larger user population the performance of MF-Polling system is comparable to that of R-TDMA. The average queuing delay is as shown in Fig. 4.12.

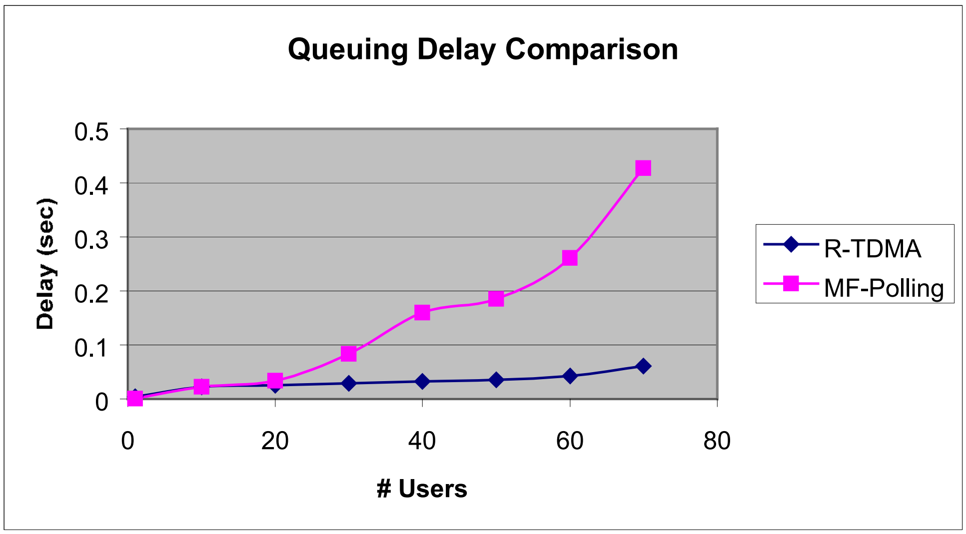

Fig. 4.12. Average Queuing Delay – Medium Load

Fig. 4.12 is similar to Fig. 4.8. As traffic increases, so does the session duration for any single user. This leads to a marginally higher time period before which data can be transmitted. It can be argued that the average queuing delay for R-TDMA with a medium load is lower than that for the FTP low download case. This could be attributed to the fact that with FTP low download, a given user tends to fall out of the system more often than the times for which it transmits data. This leads to contention delay whenever a users needs to transmit data. However, such a condition does not arise very often for the medium load. This is because of the larger amount of data present at the user queue. This leads to the piggybacking scenario and consequently reduced delay.

From the test results obtained we can analyze the behavior of each MAC protocol with different types of applications running over it and thereby decide which protocol is more suited for a given application. Based on the test results we can infer the following:

1. Output performance is certainly dependent on the protocol architecture. This is clearly seen from the Packet Generator Test wherein the channel bandwidth and the total number of channels available for utilization limit the maximum throughput of the MF-Polling system. Also there is no provision for reserving bandwidth as in case of R-TDMA. Thus, the MF –Polling architecture inherently under utilizes the available resources.
2. For a large user population, the MF-Polling system tends to perform marginally better than R-TDMA system for FTP based application. The reason being that the MF-Polling system architecture aims to accommodate more number of users at a fairly constant throughput values. With FTP based applications because the traffic is bursty and short lived in nature, a user in the R-TDMA system needs to contend and re-enter the system for every transfer. However, the advantages are nullified by the degrading performance of the contention protocol for a larger user population. Unlike the R-TDMA system, MF-Polling has large values of timeouts associated with a user being inactive in the system. Thus, contention is

avoided and with light load conditions the polling cycle time is reduced because of low data volume at the users queue.

3. It can be seen that the R-TDMA system is dependent on continuous availability of data to maintain reservation and avoid contention. As per the traffic pattern suggested by OPNET™ (Refer to Section 3.2.2) it can be seen that HTTP traffic is continuous is nature. Test results obtained for HTTP Light and Heavy Browsing confirm that R-TDMA gives better performance than MF-Polling for HTTP traffic. Contention is decreased in the R-TDMA system as slot reservation comes into effect. Though MF-Polling maintains constant throughput, it suffers from large queuing delay because of the increased number of polling cycles per channel.
4. It is not surprising to see that for medium load conditions, both the protocols have comparable performance for larger user population. The presence of each type of application equally enhances and degrades a protocol's performance. Hence, we see a uniform throughput and delay behavior.

Based on the performance of the protocols, a service provider can now decide upon the MAC protocol to be deployed depending upon the criteria which may be of greater concern and the possible tradeoffs between throughput, queuing delay, and the number of supported users.

We have also observed that the MF-Polling system suffers from large average queuing delays. This is primarily because of the contention delay component of queuing delay. Similarly, R-TDMA faces severe degradation in its throughput for large number of users. This can be attributed to contention faced by users because of varying number of slots for contention. Thus, contention efficiency is a parameter which is critical to the performance of a given MAC protocol, and it is necessary that this parameter be controlled in order to achieve better performance. The following chapter describes design improvement to the protocol architecture in order to reduce the average queuing delay.

Chapter 5

Design Improvements and Re-evaluation

5.1 Motivation

From the test results discussed in the previous chapter it is evident that the performance of a MAC protocol is limited by the average queuing delay experienced by a user. Performance can be improved by fine-tuning certain parameters, one of them being the contention delay. It must be noted that unlike a wireline system, a wireless system relies totally on the efficiency of contention mechanism used to deliver its promised services. The user arbitration mechanism is thus a critical component in providing real-time services. To enhance the contention efficiency it is necessary to clearly understand the behavior of the contention mechanism. The following sub-sections briefly describe the contention mechanism for each MAC protocol.

5.1.1 Exponential Backoff with Slotted ALOHA

R-TDMA employs exponential backoff with slotted ALOHA for providing users access to the channel in the upstream direction. The central idea to such a combination is to efficiently utilize the R-TDMA frame structure and accommodate contention slots as part of the frame. This greatly simplifies the scheduling mechanism at the headend. Certain number of slots can be provided in each frame, which will be utilized by users to transmit their requests for data slots. This is essentially the Slotted ALOHA mechanism. The number of slots in each frame is varied depending upon the traffic condition. Initially the number of slots is kept at a minimum (i.e. one). As more and more users transmit their requests, the number of colliding requests increases. This condition is detected by the headend receiver which in turn increases the number of slots in the subsequent frame. Depending on the collision status in the next frame, the headend either increases or decreases the number of slots in the subsequent frame. The user generates requests only when it has data in its queue and when it has not requested data slots by piggybacking its request in its previous data transmission. The efficiency of the contention mechanism can be increased by scheduling the transmission of request over a range of frames rather than in the frame immediately following the arrival of data at the user's queue. Initially the user can transmit its request immediately. If it is not granted a data slot in the next frame, it assumes that its request has collided. The Available Contention Window Size (ACWS) decides the upper bound on the number of frames to skip before attempting another transmission. The user doubles this value for every retransmission that it needs to make. It then chooses a random value, which is lower than ACWS and schedules its retransmission after those number of frames have passed. The ACWS can never exceed the Maximum Contention Window Size (MCWS) and is equal to 32. This value provides a maximum wait time equal to the largest possible frame time. This essentially is exponential backoff mechanism at the user end.

In order to reduce the probability of collision, it is necessary that a request transmission of request occur when the number of available slots is maximum. Such a condition occurs in the frame that follows the frame in which collision has occurred. Due to the exponential backoff mechanism, the user tends to double its window size and choose a random value which decides the number of frames to skip before retransmission. The probability that it would choose the immediate frame for retransmission is inversely proportional to the ACWS. The probability decreases as ACWS increases. The headend on the other hand reduces the number of slots when it does not observe any activity in the contention slots. Thus, there tends to be a mismatch between the number of slots available for contention and user's selection of a frame for request retransmission. This eventually affects the system performance and the performance degradation is severe when the number of users in the system is increased. Our design improvement focuses on this constraint and provides a solution for the same.

5.1.2 Exponential Backoff

MF-Polling system employs exponential backoff mechanism for user arbitration in the upstream direction. MF-Polling has a channel reserved for the users to contend. Control messages are sent downstream by the headend periodically. Users respond to these messages with their request packets, if they have data to transmit. Each polling cycle represents an "opportunity" to transmit a request. Along with the control messages the headend also notifies the users regarding the collision status during the previous broadcast period. By observing this status the user knows whether its request transmission was successful or not. The MCWS is maintained at 1024 in this case. As the number of users is increased the ACWS rapidly reaches the

MCWS. Though this mechanism is similar to exponential backoff employed in Ethernet, its behavior is controlled by the polling cycle time and ACWS. Also, because of this, the efficiency of this protocol is nearly same as that of Slotted ALOHA. It can be seen that the maximum time between two retransmissions cannot exceed the product of MCWS and polling cycle time. In this case, for a polling cycle time of 30 milliseconds and a MCWS of 1024, the product yields a maximum time for retransmission as 30.72 seconds, which is comparably large. Section 5.2.2 describes a simple rationale for controlling this value.

5.2 Design Improvements

5.2.1 Design Improvement for R-TDMA system

As explained in Section 5.1.1, the contention efficiency of the R-TDMA system depends on the number of slots in each frame and is also affected by the lack of synchronization between the increased slot value and which frame a user selects for retransmission. In order to improve the effects of such behavior the number of slots in each frame could be kept constant at its maximum permissible value. This can be based on the logic that even though the frame time is increased by the presence of additional slots, the only factor that would now affect the contention efficiency would be which frame the user selects for contention. Thus the probability of contention depends only on the ACWS. The efficiency is also increased because of the greater number of available slots per frame.

5.2.2 Design Improvement for MF-Polling system

As described in Section 5.1.2, the ACWS combined with the polling cycle time is responsible for comparatively large values of contention delay experienced by the user. The ACWS is bounded by the value of MCWS used. The MCWS value may not be the most appropriate one for the number of users currently being served by headend. This is because every collision tends to double the window size and consequently doubles the queuing delay. This is acceptable as long as the increase in window size reduces collisions and does not increase the queuing delay. Now if the MCWS is reduced, there exists an upper bound on the maximum wait time that a user would experience. It can be argued that reducing the MCWS will increase the number of collisions, which certainly is true. However, even though collisions increase, the average queuing delay is reduced. In order to understand the system behavior at a reduced MCWS, we observed the number of retransmits on the contention channel. The MWCS was reduced from 1024 to 32. A graph of retransmits versus number of users is as shown in Fig. 5.1.

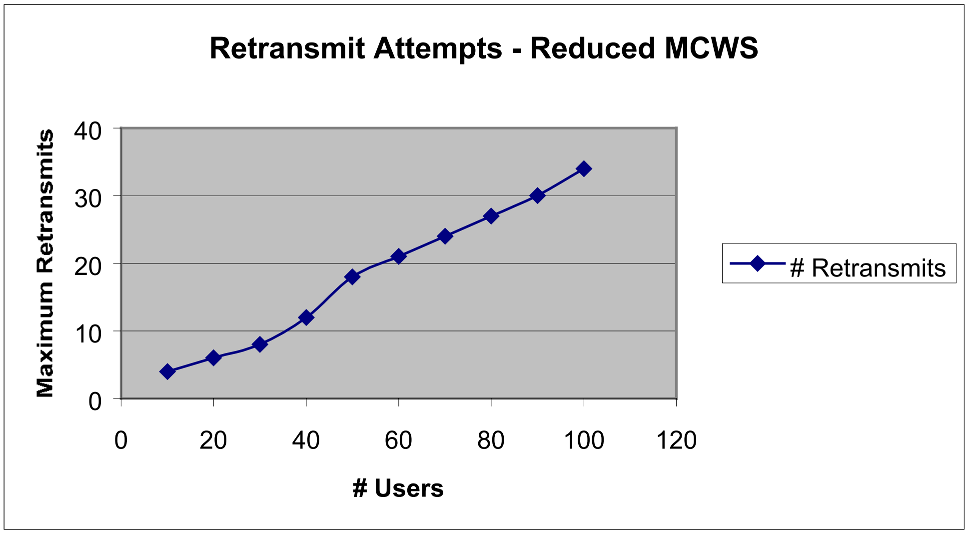

Fig. 5.1. Retransmission Attempts by a single user for reduced MCWS

It can be seen that there is a linear increase in the number of retransmits as users are added to the system. Also the number of retransmits crosses the MCWS value when the number of users in system is well above 100, which is a large value for a simultaneously bursting user population. We can now compute the worst-case maximum waiting time for a given user in this case.

Let us assume that the user under consideration, decides to skip MCWS number of transmit opportunities for all the 32 retransmits. Also let $MCWS_{prev}$, T_p and T_{pmwt} represent the Previous MCWS value, Polling Cycle Time and Previous Maximum Waiting Time respectively. Thus we have,

$MCWS_{prev} = 1024, T_p = 30ms$

We can thus compute T_{pmwt} as,

$$T_{pmwt} = MCWS_{prev} \times T_p \qquad (1)$$

which yields,

$T_{pmwt} = 30.72\,ms$

Now for the case with design improvement, let $MCWS_r$, N_{wr} and T_{wt} represent the Reduced MCWS, Number of Worst Case Retransmits and Maximum Worst Case Waiting Time respectively. Thus we have,

$MCWS_r = N_{wr} = 32$
$T_p = 30\,ms$

Now we can compute T_{wt} as,

$$T_{wt} = MCWS_r \times N_{wr} \times T_p \qquad (2)$$

which yields,

$T_{wt} = 30.72\,ms$

Thus,

$T_{pmwt} = T_{wt}$

This proves that the waiting time is lower for lesser number of retransmits, which is the case in practice. Hence, we have reduced the MCWS to 32 to achieve better performance.

5.3 Test Results with Design Improvements

5.3.1 Packet Generator Test

Since this test was conducted primarily to observe the throughput bound of the system no queuing delay measurements were made. The design improvement that aims to decrease the queuing delay is therefore not applicable in this case.

5.3.2 FTP Low Download

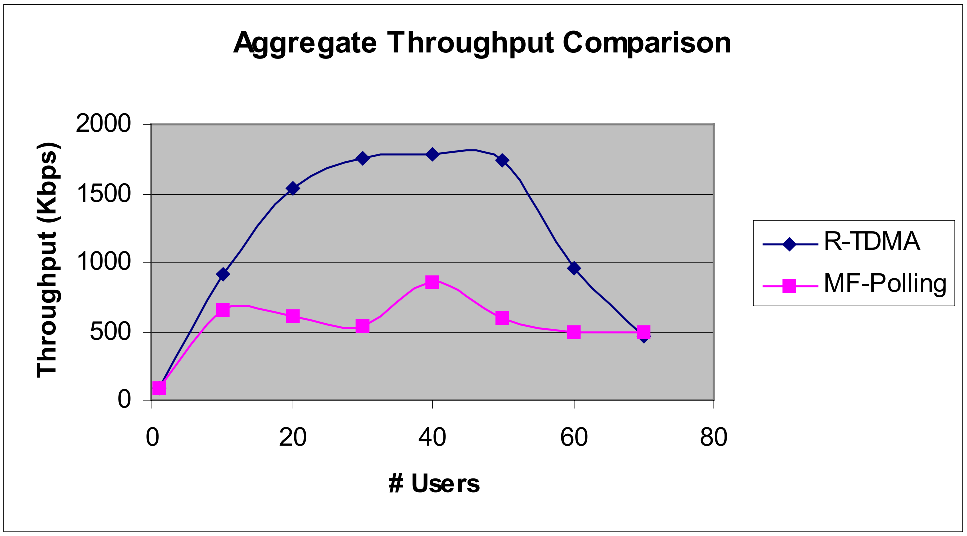

Fig 5.2. Aggregate Throughput Comparison – FTP Low Download (Improved)

Fig. 5.2 shows the graph with improved values of aggregate throughput. This figure can be best explained by observing the comparative graph of queuing delay against the number of users shown in Fig. 5.3.

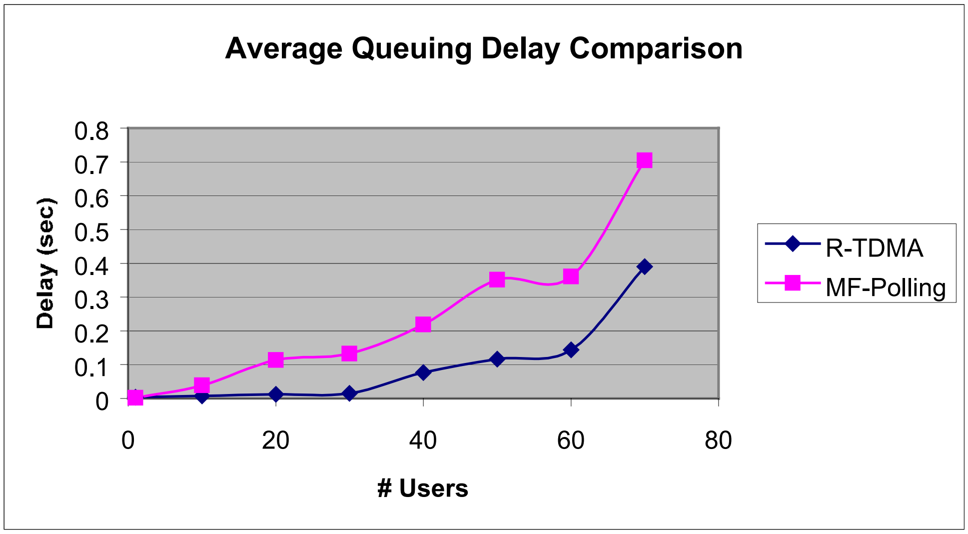

Fig. 5.3. Average Queuing Delay Comparison – FTP Low Download (Improved)

It is interesting to observe the improvement in delay for each of the MAC protocols compared to the values obtained earlier. Fig 5.4 and Fig. 5.5 show such a comparison for R-TDMA and MF-Polling respectively.

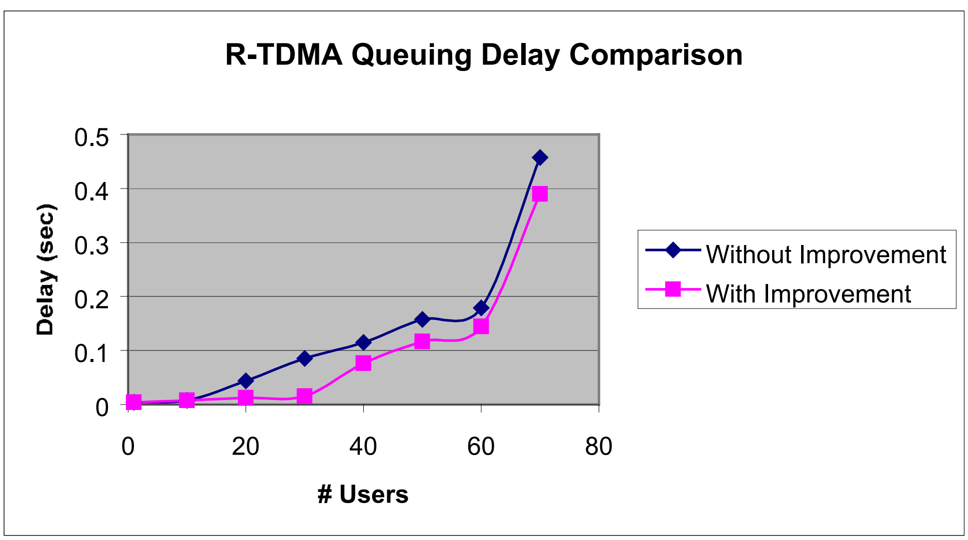

Fig. 5.4. R-TDMA Queuing Delay Comparison – FTP Low Download

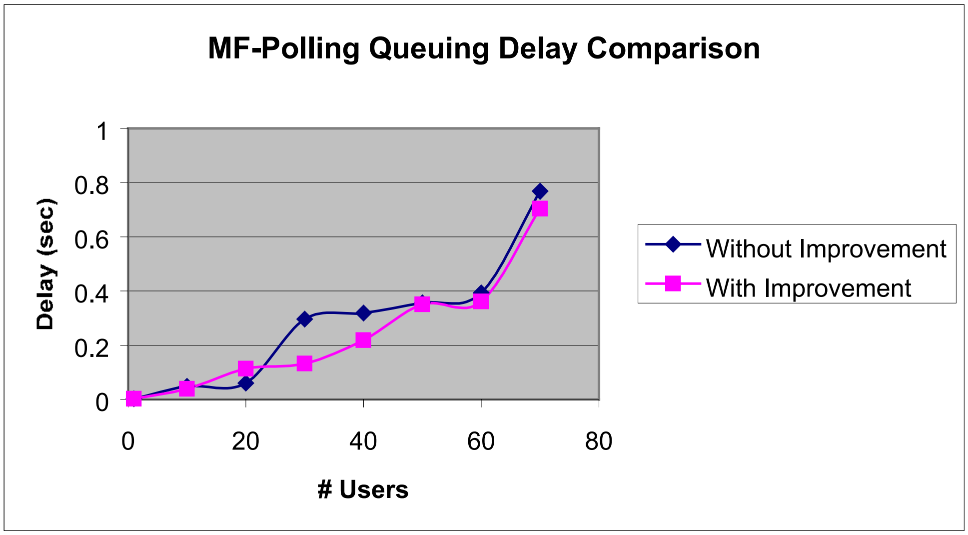

Fig. 5.5. MF-Polling Queuing Delay Comparison – FTP Low Download

It can be seen that the improvement in queuing delay values is not very significant. This is due to the fact that reduction in contention delay does not affect the actual waiting time for data transmission, which depends upon the availability of data at users queue. Hence, for light load conditions the improvement in queuing delay is not significant for either of the protocols. However, R-TDMA still shows better performance in terms of throughput and delay compared to MF-Polling.

5.3.3 FTP High Download

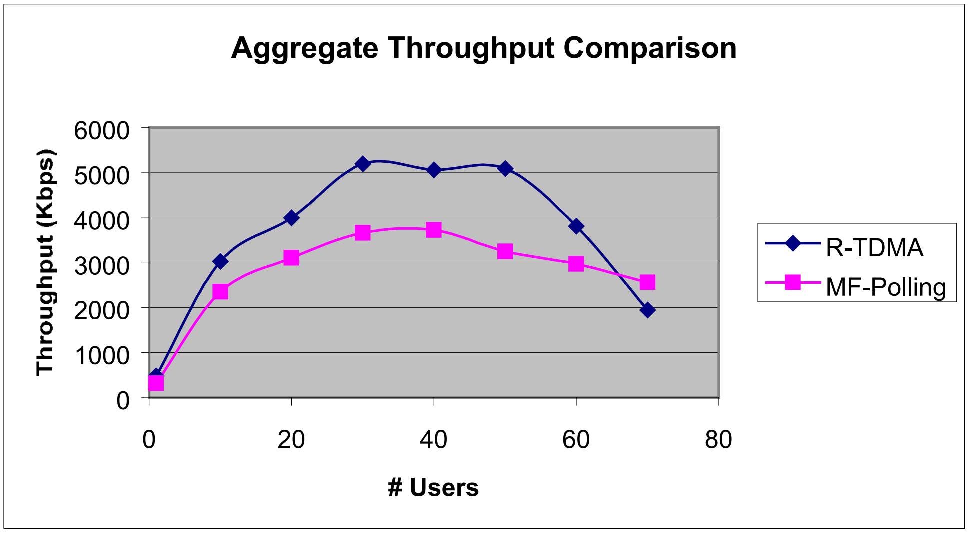

Fig. 5.6. Aggregate Throughput Comparison – FTP High Download (Improved)

From Fig. 5.6 it can be seen that both MAC protocols show an improvement in performance compared to the values depicted in Fig. 4.5. It can also be seen that because of reducing the MCWS, data transmission is more continuous in nature, and this smoothes out the MF-Polling curve. The advantage with fixed slots for R-TDMA is that it maintains its throughput performance better than that of MF-Polling for higher user values. The effect of reduced MCWS for MF-Polling is seen for a large number of users where its throughput performance is better than its counterpart.

The queuing delay comparison is as shown in Fig. 5.7 while figures 5.8 and 5.9 show the improvement in delay performance over its original value for both the protocols.

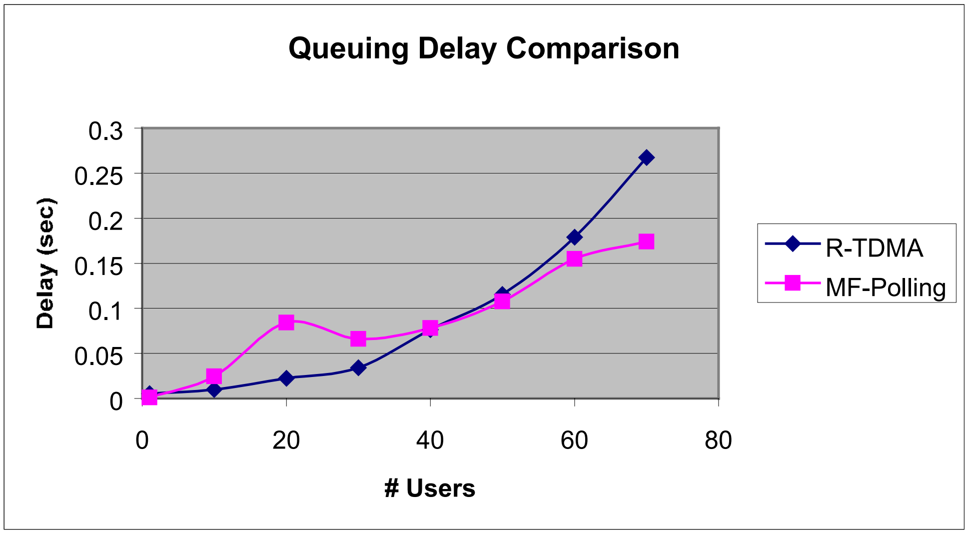

Fig. 5.7. Queuing Delay Comparison – FTP High Download (Improved)

When compared to Fig. 4.6, Fig.5.7. shows that MF-Polling performs better for a larger user population than R-TDMA even with the design improvements.

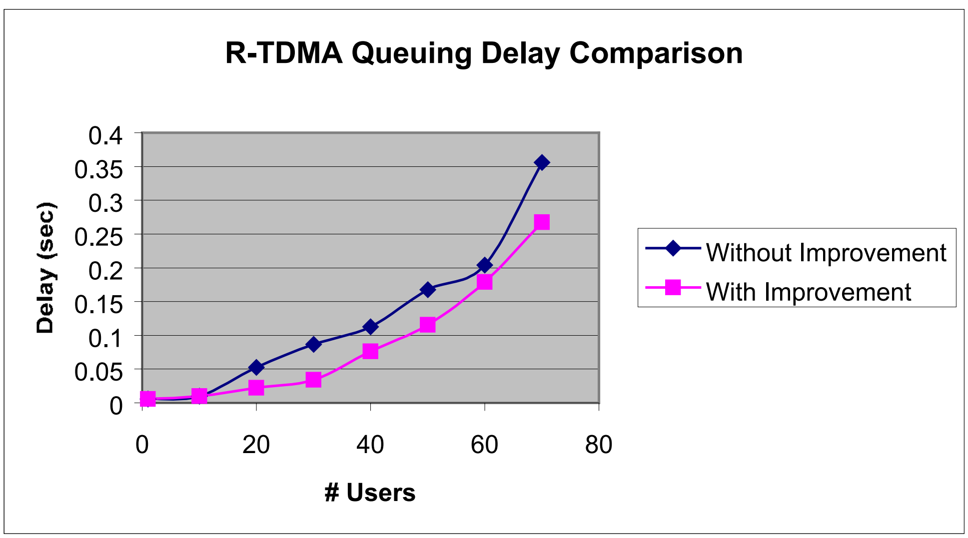

Fig. 5.8. R-TDMA Queuing Delay Comparison – FTP High Download

Because of the higher volume of data the difference in delay value is much more pronounced in this case as compared to the FTP low download case which can be seen from Fig 5.4.

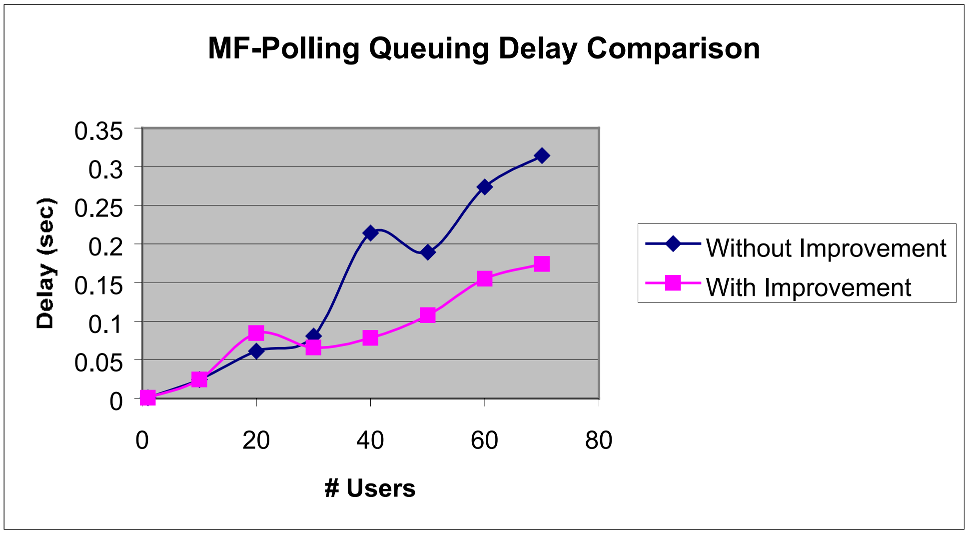

Fig. 5.9. MF-Polling Queuing Delay Comparison – FTP High Download

The effect of the design improvement is clearly visible from the Fig.5.9. where the improved queuing delay is nearly half of the original value.

5.3.4 HTTP Light Browsing

We can see the effect of the design improvement for this test case in Fig. 5.10, which is a comparative graph of aggregate throughput versus number of users in the system.

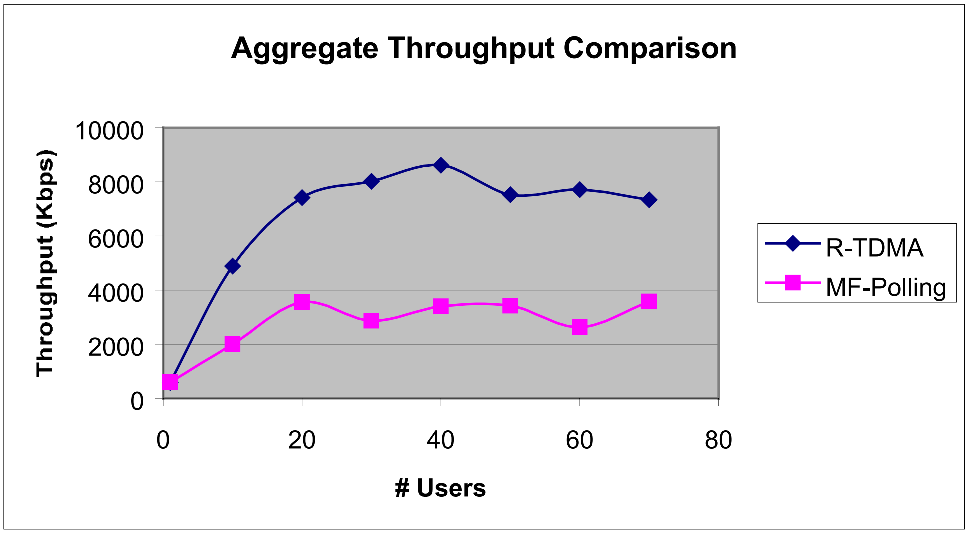

Fig. 5.10 Aggregate Throughput Comparison – HTTP Light Browsing (Improved)

Comparing the above figure with Fig. 4.7 it can be observed that the performance of R-TDMA has improved and has become more stable over a larger range of users. The performance characteristics still remain better than that of the MF-Polling system. Both the protocols show an increase in their respective throughput values. The reason for this improvement remains the same as mentioned in the previous test case. Fig. 5.11 shows the improved average queuing delay comparison for the two protocols.

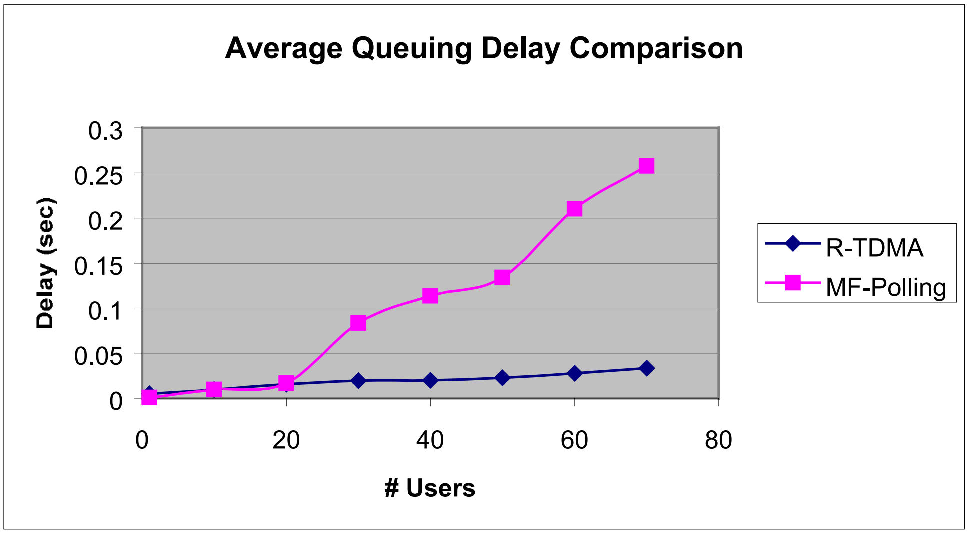

Fig. 5.11. Average Queuing Delay Comparison – HTTP Light Browsing (Improved)

The graph in Fig.5.11. is similar to Fig. 4.8, but has lower values of queuing delay. Fig.5.12. and Fig. 5.13. show the performance improvements for R-TDMA and MF-Polling respectively.

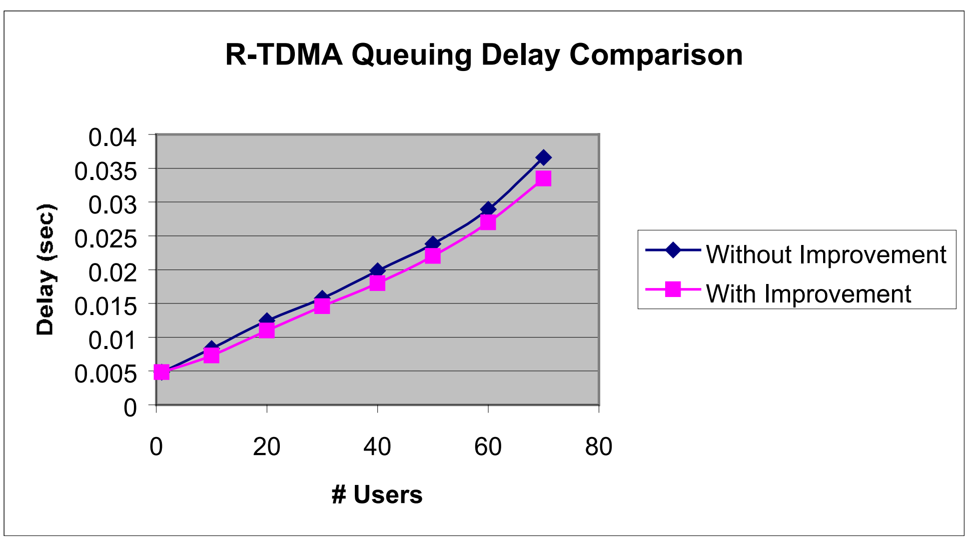

Fig. 5.12. R-TDMA Queuing Delay Comparison – HTTP Light Browsing

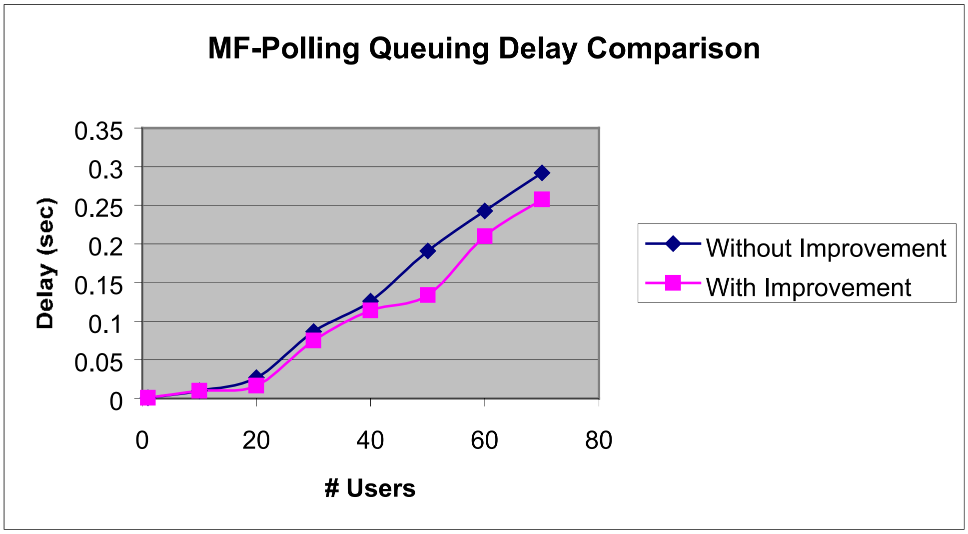

Fig. 5.13. MF-Polling Queuing Delay Comparison – HTTP Light Browsing

5.3.5 HTTP Heavy Browsing

An aggregate throughput comparison can be done for the improved versions of the two protocols. This is as shown in Fig. 5.14.

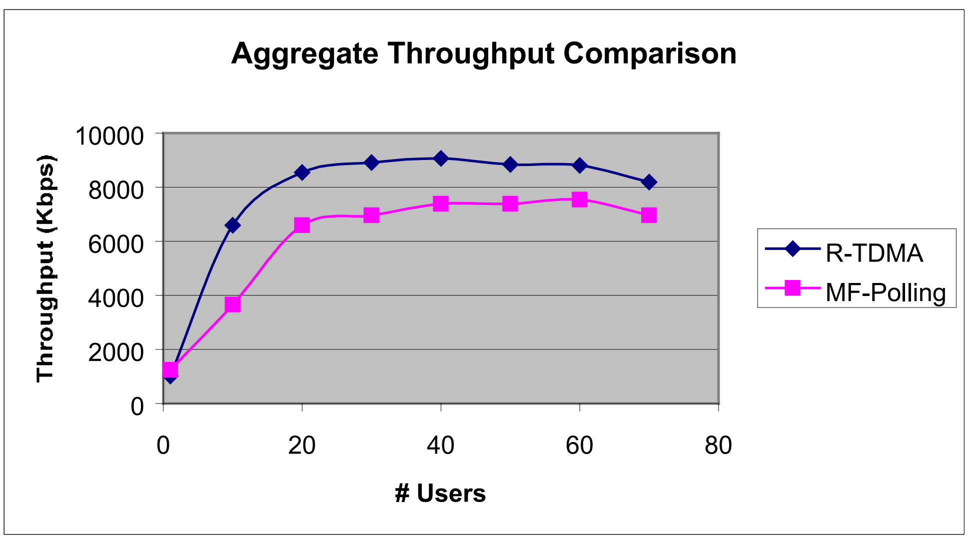

Fig. 5.14. Aggregate Throughput Comparison – HTTP Heavy Browsing (Improved)

Similar to the reasoning mentioned for Fig. 4.9., the throughput performance for the R-TDMA system is because of the prolonged data sessions in which more and more users are not granted slots for transmission. Hence, the design improvement does not affect the R-TDMA system values as it affects the MF-Polling system, which now shows an improvement in its throughput.

The queuing delay performance can be observed from Fig.5.15, Fig 5.16 and Fig.5.17.

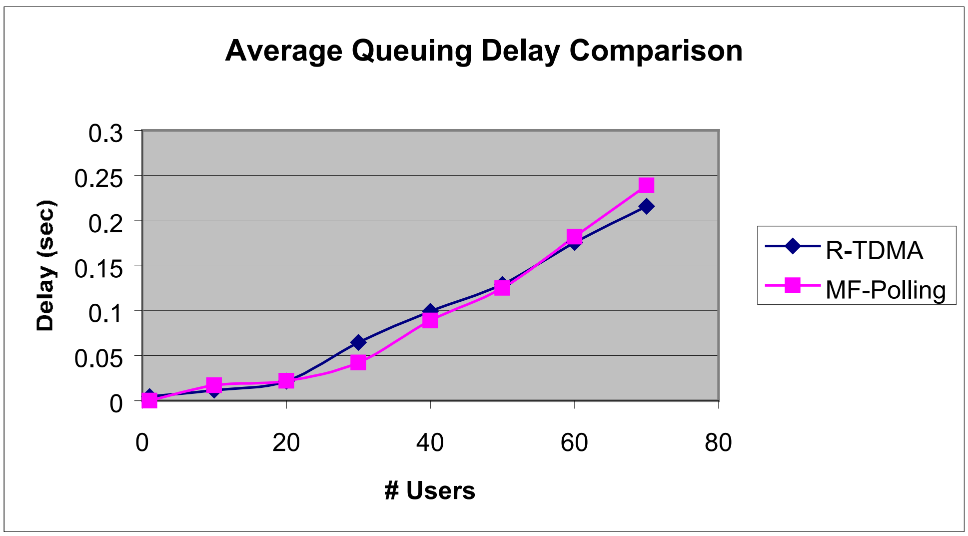

Fig. 5.15. Average Queuing Delay Comparison – HTTP Heavy Browsing (Improved)

The design improvement reduces the MF-Polling average queuing delay and makes it comparable to that of R-TDMA. Figures 5.16. and 5.17. compare the improved values of queuing delay with their original values.

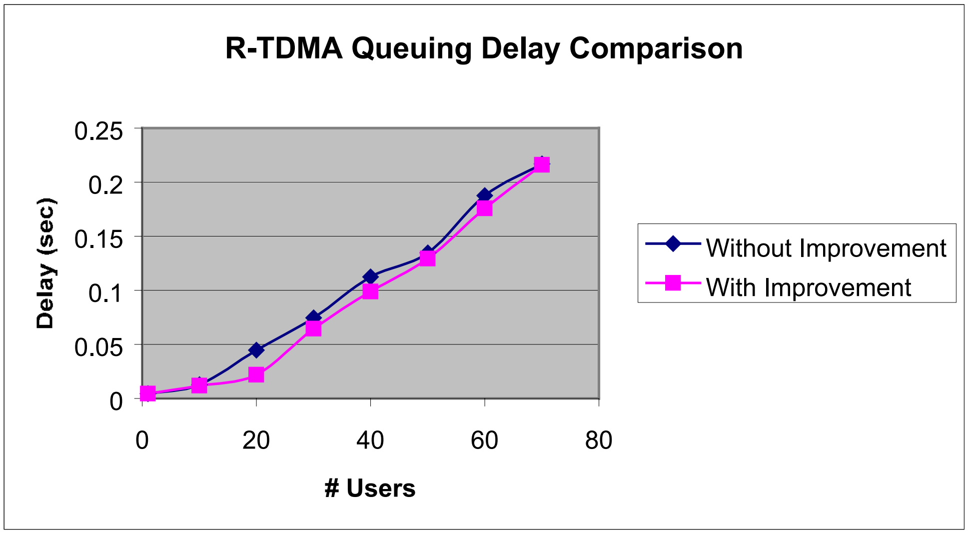

Fig. 5.16. R-TDMA Queuing Delay Comparison – HTTP Heavy Browsing

It can be seen that design improvement has not significantly affected the queuing delay for R-TDMA system as the resources are unavailable to many users due to reasons mentioned in Section 4.2.5 hence, the queuing delay is not affected by contention.

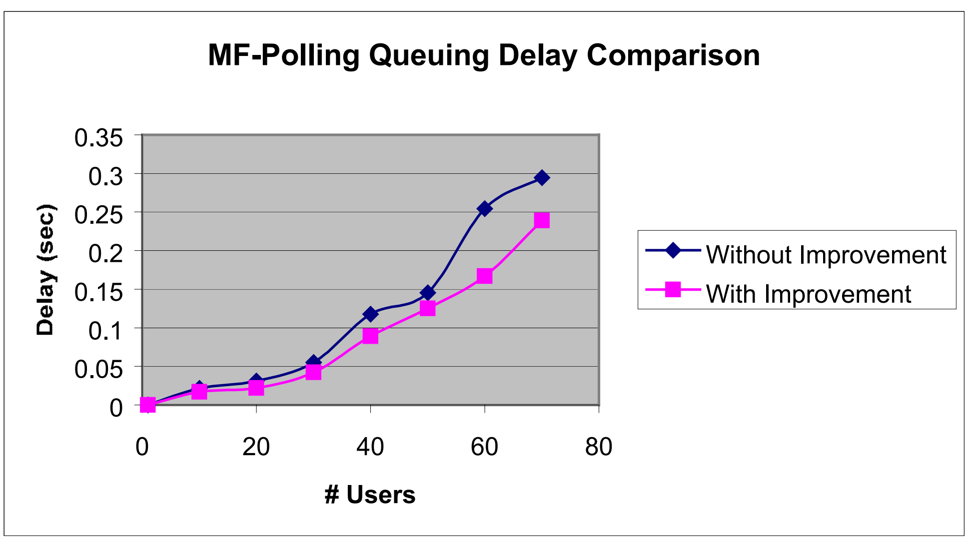

Fig. 5.17. MF-Polling Queuing Delay Comparison – HTTP Heavy Browsing

Improvement in the MF-Polling case can be attributed to the same reasons as mentioned in the previous tests.

5.3.6 Medium Load

This test has traffic pattern which is a combination of FTP Low Download and HTTP Light Browsing. The following figures show the effect of design improvement on the aggregate throughput and queuing delay. Fig. 5.18. depicts the effect of design improvement on the aggregate throughput.

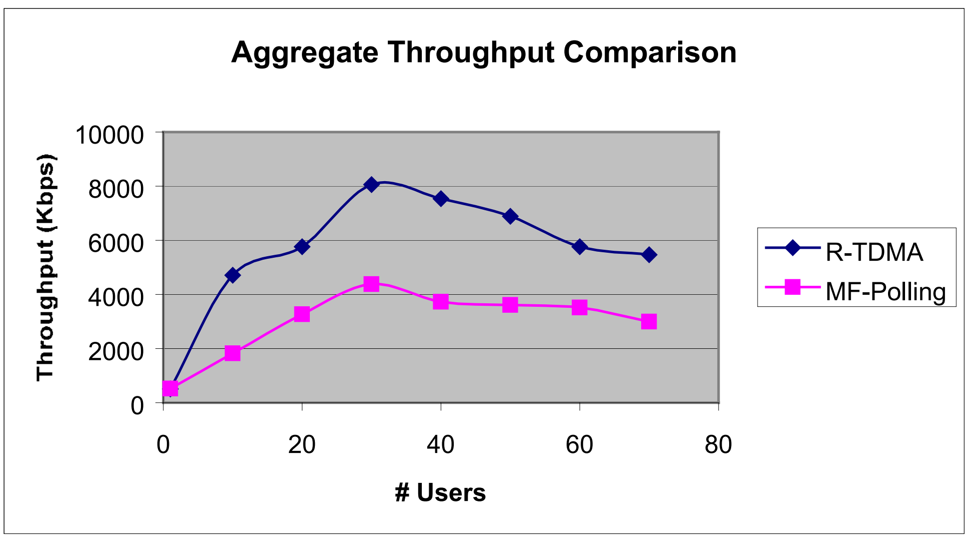

Fig. 5.18. Aggregate Throughput Comparison – Medium Load (Improved)

When compared to Fig. 4.11. we can see significant improvement in the performance of R-TDMA system and the throughput is comparably improved for larger user population. The queuing delay comparison is as shown in Fig. 5.19.

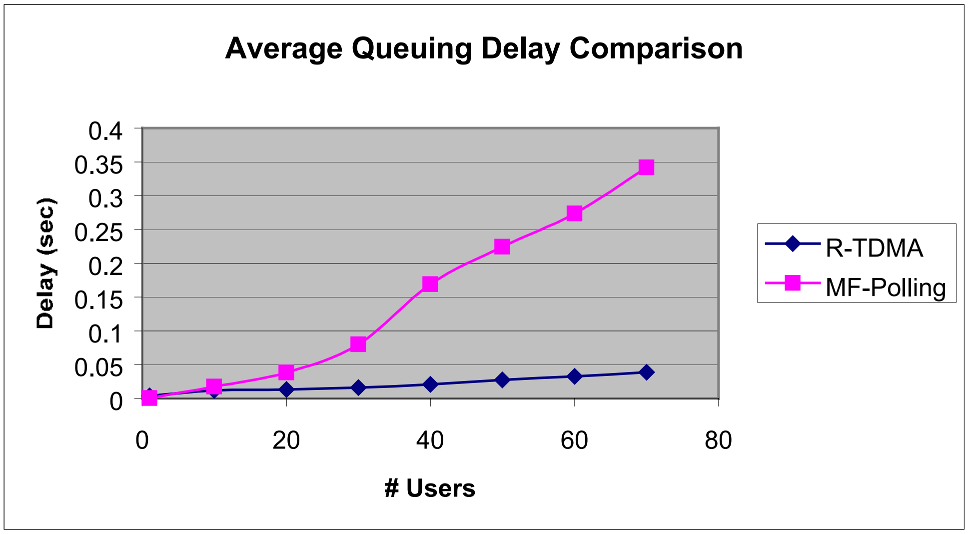

Fig. 5.19. Average Queuing Delay Comparison – Medium Load (Improved)

The figure shown above is similar to Fig. 4.12. with marginally improved queuing delay values because of the design improvement. The individual graphs with improvements are as shown in Fig. 5.20. and Fig. 5.21.

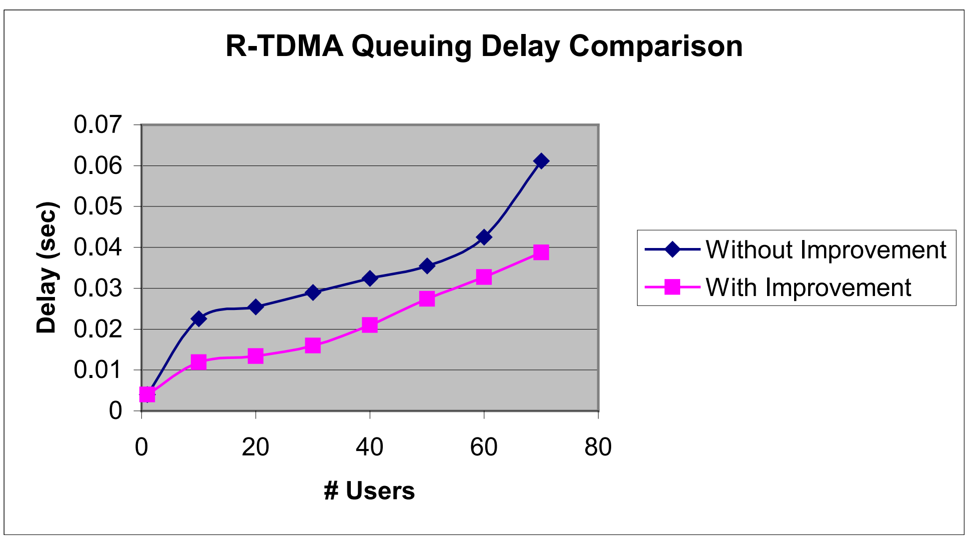

Fig. 5.20. R-TDMA Queuing Delay Comparison – Medium Load

The improvement in queuing delay is evident from Fig. 5.20, which consequently leads to improved throughput performance.

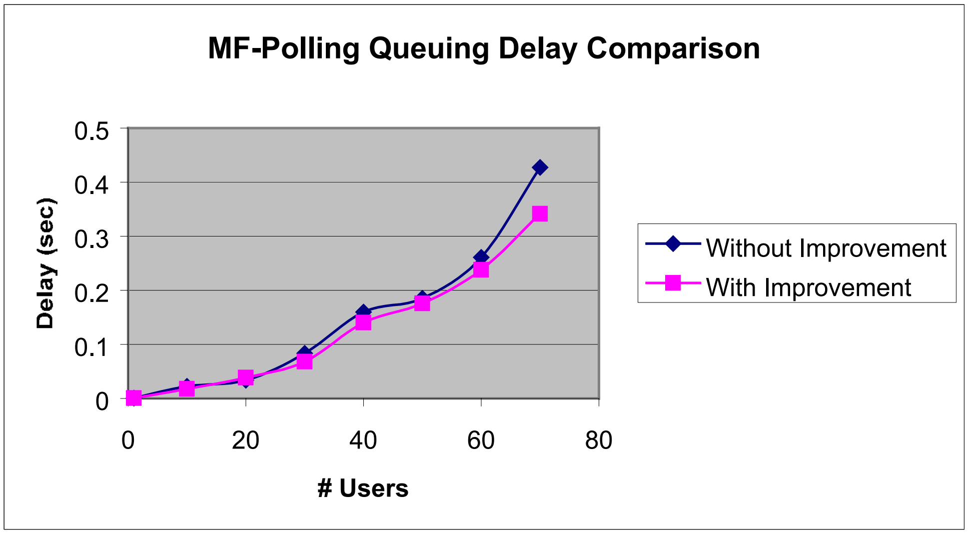

Fig. 5.21. MF-Polling Queuing Delay Comparison – Medium Load

The improvement in queuing delay and throughput is marginal compared to the previous test cases.

From the tests results obtained in this chapter we can see that:

1. Because of the design improvements, better performance is observed for both MAC protocols under consideration as compared to their original performance.
2. The effect of the design improvement is more pronounced in the case of MF-Polling compared to the R-TDMA system. This is observed from the delay performance which is now comparable to that of R-TDMA.
3. MF-Polling offers a greater range of users over which its throughput remains stable and has delay characteristics comparable to that of R-TDMA.
4. Because of the improved contention mechanism, R-TDMA performs better for light load conditions and offers much higher throughput than the MF-Polling system over a large range of users.

5. MF-Polling is suited for FTP based applications with large user population compared to R-TDMA.
6. HTTP based applications still perform better with R-TDMA system over an increased range users with stable throughput.

In general, from the tests results discussed in the previous and current chapter, we conclude that the R-TDMA system is better suited to HTTP based applications than FTP based applications. Similarly, the MF-Polling system is better suited to FTP based applications than HTTP based applications.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This chapter discusses the conclusions of this performance evaluation and the future work that can be carried out in this area. The primary objective of this report was to provide a B-WLL service provider with a comparative performance evaluation of the commonly used MAC protocols in terms of the throughput, queuing delay, and the number of users supported for different types of applications. A large number of applications could have been considered for evaluation, but the considerable time required for each simulation run limited the number of cases that could be investigated. Hence, based on the available data regarding the dominant Internet applications and traffic patterns as suggested by OPNET™, we conducted a set of tests with cases deemed most relevant.

The other objective was to suggest design improvements to the existing protocol architecture. The outcome of this exercise was twofold. First, we were able to judge the performance of each protocol in relation to various traffic conditions that it needed to support. Second, having done the design improvement we were able to re-evaluate the protocol performance and draw general conclusions regarding the type of application or traffic to which the protocol was more suited. Such an evaluation benefits the service provider and system designer equally well. As proposed, the protocol performance was evaluated, design improvements were suggested, and performance was re-evaluated. This exercise led to the following conclusions:

1. The performance of any given MAC protocol stems from its basic architecture. Each protocol is inherently biased towards a particular output parameter. This could either be throughput or queuing delay, which is because of dynamic bandwidth management as in case of R-TDMA or the number of users that can be supported as in case of MF-Polling. The system designer needs to trade off one output parameter for another based on the criterion that is of greater concern. From the test results we can say that in general R-TDMA is more efficient in terms of bandwidth utilization, while MF-Polling can support a larger user population.
2. An application's traffic pattern affects the behavior of the underlying MAC protocol. From the test results, we infer that MF-Polling is better suited to FTP applications, whereas R-TDMA is better suited to HTTP applications.
3. The type of contention mechanism used does dictate the system output performance. For R-TDMA, which employs Slotted ALOHA with exponential backoff, we observe that by keeping the number of slots fixed for each frame, we could improve the contention efficiency of the system. Similarly, for MF-Polling, which employs exponential backoff, a reduction in contention window size enhances the system performance.
4. Real-time services that are critically dependent upon low delay values, low jitter and frequently high throughput (e.g. Voice, Video Conferencing, NetMeeting)

would perform better with R-TDMA, while non-real-time services (e.g. FTP, Telnet, E-mail) would perform better with MF-Polling. The explicit difference being the nature of traffic generated by each class of application.

5. Based on an extrapolation of the results obtained, we can deduce that MF-Polling offers a sustained value of throughput for a larger number of users compared to R-TDMA, which trades off the number of users supported with a higher aggregate throughput.

6.2 Future Work

Because of the various types of services being offered on the Internet, ranging from simple chat services to live webcast services, there is a constant need to efficiently utilize every available unit of the bandwidth. Since more and more ISP's are migrating their services and efforts towards providing access with B-WLL systems, it is necessary that the underlying MAC protocol is made "aware" of the demands placed by the applications. This can be achieved by making the MAC scheduler QoS-aware. R-TDMA systems may be made QoS-aware by categorizing users into various classes (classification may be based either on applications or on some pre-assigned criteria) and scheduling their requests accordingly.

A new MAC scheduler could be developed which would convert the MF-Polling system to a MF-TDMA system with R-TDMA employed on each channel. This would combine the advantages of both the MAC protocols.

Another interesting study would be to observe the effects of using a collision-sensing mechanism as a contention protocol for the system.

Bibliography

[Aca99] Anthony Acampora, "Overview of UCSD's Center for Wireless Communications," at http://www-cwc.ucsd.edu/tony.html, 1999.

[BDLN96] Bolla, Davoli, Lazzaro, Nobile, "Dynamic Bandwidth Sharing between Voice and Data Traffic in a RRA Multiple Access Scheme for Cellular Mobile Radio Networks," *Annual International Conference on Universal Personal Communications - Record*, 2 Sep 29-Oct 2 1996, pp. 553-557

[BeG92] D. Bertsekas and R. Gallager, *Data Networks*, Prentice Hall, New Jersey (1992).

[BeT85] T. Berger and T. S. Tsazan, "An Improved Upper Bound for the Capacity Random Multiple Access System," in *Problemy Peredachi Informatsii*, 21(4) pp. 83-87, January 1985.

[Cla98] K. Claffy, "The Nature of the Beast: Recent Traffic Measurements From an Internet Backbone," at http://www.isoc.org/inet98/proceedings/6g/6g_3.htm

[CoA98] K. G. Coffman and Andrew Odlyzko 1998. "The Size and Growth Rate of the Internet," *First Monday* at http://www.firstmonday.dk/issues/coffman

[Kas98] Kizito Tshilumba Kasengulu, "A Comparison Study between *TDMA and FDMA* in Digital Wireless Systems," in *Spread Spectrum Scene* at http://sss-mag.com/newiss.html, April 1998.

[KSY84] J. F. Kurose, M. Schwartz, and Y. Yemini, "Multiple Access Protocols and Time Constrained Communication," in *Computer Survey* 16(1) pp. 43-17, March 1984.

[LBA98] Le, Babak, Aghvami, "Performance Comparison of new Centralized PRMA and TDMA systems for Wireless ATM Networks," *Proc. IEEE*

Vehicular Technology Conference, v. 2 May 18-21, 1998, pp. 1466-1470.

[LKA93] R. O. La Marie, A. Krishna, H. Ahmadi, "Analysis of a wireless MAC protocol with client-server traffic," *Proc. IEEE INFOCOM*, v. 2, Mar 30–Apr 1 1993, pp. 429-438.

[MiT81] V. A. Mikhailov and B. S. Tsybakov, "Upper Bound for the Capacity of a Random Multiple Access System," *Problemy Peredachi Informatsii*, 17(1) pp. 90-95, January 1981.

[Mol82] M. L. Molle, "On the Capacity of Infinite Population Multiple Access Protocols," *IEEE Trans. on Information Theory*, IT-28(3), pp.396-401, May 1982.

[PeD96] Larry L. Peterson and Bruce S. Davie, *Computer Networks: A Systems Approach,* Morgan Kaufmann Publishers, Inc. San Francisco, California (1996)

[Rap98] Theodore S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice-Hall, New Jersey, pp. 395-409 (1998).

[RoS90] R. Rom and M. Sidi, *Multiple Access Protocols: Performance and Analysis*, Springer-Verlag, New York, (1990)

[Rav94] K. V. Ravi, "Comparison of Multiple-accessing Schemes for Mobile Communication Systems," *Proc. IEEE International Conference on Personal Wireless Communications*, Aug 18-19, 1994, pp. 152-156.

[Sac88] S. R. Sachs, "Alternative Local Area Network Access Protocols," *IEEE Communications Magazine* 26(3) pp. 25-45, March 1998.

[Tha00] M. Thaker, *Performance Evaluation and Design Improvement of Media Access Control Protocols for Broadband Wireless Local Loop*, MSEE Thesis, Department of Electrical Engineering and Computer Science, the University of Kansas, Lawrence, Kansas, August 2000.

[Wan96] Wen Wang, "Performance comparisons of STDD and FDD schemes for TDMA systems using NC-PRMA Protocol," *Proc. IEEE*

International Symposium on Personal, Indoor and Mobile Radio Communications, v. 2, Oct 15-18, 1996, pp. 382-386.

Appendix A

Abbreviations

A

ACWS	Available Contention Window Size
ATM	Asynchronous Transfer Mode

B

BPSK	Binary Phase Shift Keying
B-WLL	Broadband Wireless Local Loop

C

CDMA	Code Division Multiple Access
CO	Central Office
CSMA/CD	Carrier Sense Multiple Access with Collision Detection

D

DSL	Digital Subscriber Line

F

FCC	Federal Communications Commission
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FIFO	First In First Out
FTP	File Transfer Program

H

HTTP	Hyper Text Transfer Protocol

I

ISM	Industrial, Scientific and Medical
ISP	Internet Service Provider

L

LAN	Local Area Network

M

MAC	Media Access Control
MCWS	Maximum Contention Window Size
MDS	Multipoint Distribution Service
MF-Polling	Multi-Frequency Polling
MMDS	Multichannel Multipoint Distribution Service

Q

QoS	Quality of Service
QPSK	Quadrature Phase Shit Keying

R

R-TDMA	Reservation Time Division Multiple Access

S

SMTP	Simple Mail Transfer Protocol
SNR	Signal to Noise Ratio

T

TCP	Transport Control Protocol
TDD	Time Division Duplex
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access

U

U-NII	Unlicensed National Information Infrastructure

W

WATM	Wireless Asynchronous Transfer Mode
WLAN	Wireless Local Area Network
WLL	Wireless Local Loop

Appendix B

Computation of Average Throughput

Aggregate throughput is a measure of the total system throughput i.e. uplink and downlink. The traffic generated for simulation is dependent upon the seed value used for that particular run. Hence, there is some amount of randomness associated with the values of throughput obtained. Ensemble averages thus need to be taken to compute a true throughput. The following describes the procedure used to derive the ensemble averages.

1. For a single simulation run, the simulation run time was long enough such that the system reaches a steady state. The final value thus obtained is the steady state value which has been averaged over the simulation run time.
2. Multiple graphs can be obtained with different seed values for each simulation run. For the test cases used in this report, approximately 35-40 graphs were obtained for each test.
3. Each graph can then be sampled at a fixed interval and discrete values can thus be obtained. Same number of samples should be obtained from each of the graph.
4. An averaging operation can be performed across all the similar numbered samples from the entire number of graphs. This results in an ensemble average of the throughput.

Appendix C

Glossary of Terms

Available contention Window Size (ACWS): A value used by the user performing exponential backoff as an upper limit to compute a random number.

Binary Phase Shift keying (BPSK): A modulation scheme used in digital communications.

Broadband Wireless Local Loop (B-WLL): Wireless Local Loop system that can provide access at speeds greater than 1.5 Mbps.

Code Division Multiple Access (CDMA): A multiple access scheme that utilizes the available bandwidth simultaneously by using unique user codes.

Carrier Sense Multiple Access with Collision Detect (CSMA/CD): A multiple access scheme in which users sense the collision and thus avoid further collision by transmission of their data.

Federal Communications Commission (FCC): A government body responsible for the management of wireless spectrum in the United States.

Frequency Division Duplex (FDD): A multiple access scheme that allows simultaneous data transmission in upstream and downstream but on different frequencies.

Media Access Control (MAC): A set of rules that controls the transmission of data by users onto a shared resource.

Maximum Contention Window Size (MCWS): The maximum value that ACWS is allowed to take.

Quality of Service (QoS): A broad based term used to define the various parameters like throughput, delay, jitter etc. collectively.

Quadrature Phase Shift Keying (QPSK): A modulation scheme used in digital communications.

Signal to Noise Ratio (SNR): A ratio of signal power to noise power generally expressed as log value.

Time Division Duplex (TDD): A multiple access scheme in which transmission occurs on the same frequency in uplink and downlink, but is spaced in time.