Performance Analysis of Distributed Queueing Random Access Protocol--DQRAP

Xianyu Zhang, Graham Campbell

Department of Computer Science
Illinois Institute of Technology
March 3, 1993

Abstract— Distributed Queueing Random Access Protocol (DQRAP) is a multi-access protocol for a broadcast channel that provides performance superior to any other protocol with respect to stability, delay, and throughput. DQRAP achieves this performance with a transmission queue TQ, a collision resolution queue RQ, and ternary feedback from as few as three minislots. Xu and Campbell introduced DQRAP and provided a mathematical model of the throughput. This paper extends the throughput analysis of Xu and Campbell to show that a throughput of one is achieved with a load of one or greater. A mathematical model of the delay is introduced. Stability is discussed. The average delay and throughput of DQRAP as obtained by theoretical analysis are compared with simulation results. Comparisons of DQRAP with ALOHA and Tree protocols show the superiority of DQRAP with respect to both delay and throughput.

I. INTRODUCTION

For decades there has been a search for the ideal multi-access protocol whose performance would approach the theoretical performance of an idealized M/D/1 queue. This ideal system would provide: i) immediate access, ii) throughput equal to input rate up to $\lambda = 1$, and then equal to 1 at $\lambda > 1$, iii) first-come, first-served. The search commenced with ALOHA that achieves a maximum throughput of 0.184 and 0.368 [6] with instability for the pure and slotted versions respectively. Since then, other members of the ALOHA family such as CSMA and CSMA/CD have been introduced with the goal of improving throughput and stability. The family of Tree (also called Splitting) protocols solve the instability problem by restricting uncommitted stations from transmitting until the stations involved in a collision transmit their packets successfully. The basic Tree approach has been improved to allow a maximum stable throughput of 0.46 [7]. The upper bound of the throughput of basic tree protocols is 0.568 [4].

To improve the throughput characteristics of a tree protocol, a control minislot method (CMS) was applied to resolve collisions. The CMS approach uses minislots as a reservation mechanism to to reduce the number of wasted data slots and it does improve delay and throughput. The best of the protocols with CMSs from a theoretical point view is the Announced Arrival Random Access (AARA) protocol of Towsley and Vales [8] with a maximum throughput of 0.853 using three CMS. However, in a real AARA system, the theoretical maximum throughput can never be achieved because when the input rate reaches 0.7 the average delay approaches infinity. A real AARA system with 3 minislots would not prove practical at input rates of $\lambda > 0.7$.

DQRAP was introduced in 1990 [1], [2], [16]. The protocol is different from both the ALOHA and Tree families of protocols. However, DQRAP uses elements of each such as immediate access (ALOHA), splitting CRA (Tree), and CMSs to control collision. The major innovation is the introduction of two queues at each station that respectively buffer i) the packets involved in collisions in CMSs, and ii) the packets that achieve permission to transmit. The Contention Resolution Queue, named RQ, is designed to resolve collisions among stations attempting to successfully obtain a minislot. The Transmission Queue, TQ, is used to buffer the data packets that have obtained permission to transmit and are awaiting their scheduled time of departure using a FCFS discipline. Note that the TQ and RQ are described as buffering packets. This is for clarity in explaining the mathematical model, in practice the RQ and TQ are implemented as simple counters. This paper shows that with an input rate of 0.95, DQRAP achieves a stable throughput of 0.9215, excluding the overhead of 3 CMSs ($\alpha = 0.03$), with an average delay of less than 14 time units (slots). The delay of an idealized M/D/1 system at an input rate of 0.95 is 11 time units (slots). When the input rate reaches 1, the throughput of DQRAP, excluding overhead, approaches 1 with just 3 minislots.

Three terms are used in the paper to describe data traffic flowing through a system: Input Rate λ, Offered Load G, and Throughput S. To avoid confusion, the terms are defined as follows: i). Input Rate λ is the ratio of the average number of newly arrived packets per slot to the transmission capacity of the system; ii). Offered Load G is the ratio of the total amount of traffic per slot loaded over a system, including both new and backlogged traffic ,to the transmission capacity of the system; iii). Throughput S is the ratio of the average number of packets transmitted successfully per slot by a system to the transmission capacity of the system.

II. DQRAP CRA AND STATISTICAL MODELS

A DQRAP environment is considered as a communication system serving a large population of bursty users who communicate over a multi-access broadcast channel. The channel is divided into time slots of equal length. Each slot consists of M control minislots (CMS) and a single data slot (DS). The length of a slot is assumed to be one. The size of a CMS, ε, may be different based on both channel synchronization requirements and the requirement of obtaining ternary feedback, but is assumed small compared with the length of a slot, e.g. $\varepsilon << 1$. In this paper, we use $\alpha = M\varepsilon$ to describe the ratio of the total overhead of M CMSs to the size of a slot.

The feedback information from CMSs is used to show the state of the channel in either the current slot or a slot passed some time ago depending on the delay characteristics of the channel. We use the terms BF and TF to express Binary Feedback and Ternary Feedback respectively. IF and DF(n) are used to represent Immediate Feedback and Delay Feedback respectively. IF means feedback is available immediately after transmission on each CMS while the feedback of DF(n) can be received at the end of the nth slot after the current slot. When n is equal to one, the feedback is available at the end of current slot. When n is greater than 1, the feedback is carried by a slot in a channel with an interleaving factor of n. The interleaving feedback is applicable to systems with values of "a" > 0.5 where a = ratio of propagation delay to transmission time. In this paper, we consider the case of ternary feedback with IF and DF(1) feedback types. The IF case is optimal and obviously easier to analyze than DF(1). The cases of binary feedback and DF(n), n > 1, will be addressed in future research.

A detailed explanation of The Collision Resolution Algorithm (CRA) of DQRAP is presented in [16]. Here, we restrict ourselves to the important steps of the CRA. Each station connected to the channel monitors the feedback of each CMS. The feedback indicates the position the packet will take in the RQ or TQ. The newly arrived packets, called group G, are split equally into M CMSs and form groups, G_1, G_2, ..., G_k, ..., $G_{M'}$, M'<=M when the previous collisions are resolved, i.e., the RQ is empty. If there is a single packet in a group, the packet obtains permission to transmit and queues in the TQ with FCFS discipline. The collided groups "return" to the RQ and the position of the groups in the RQ follows their group number. Next time slot, the group of packets at the head of the RQ split into M CMSs. The procedure repeats until all the packets in group G obtain permission to transmit. In each slot, a packet is picked from the top of the TQ for transmission. A special case is when the TQ is empty: the packets in the current contention group are transmitted immediately in the current slot. These packet(s) are transmitted in parallel with the regular CMS contention activity. This case represents the immediate access feature of DQRAP, i.e., if there is no packet waiting in the TQ then new arrivals transmit immediately. Multiple arrivals obviously collide, the only instance of data collision in DQRAP.

Fig. 1 shows the statistical delay and throughput models of DQRAP. There are two components to the DQRAP models: one of them is the CR component that addresses RQ management and resolves collisions; the other is the DT component that manages the TQ and controls Data Transmission. The two components form a serial system in which traffic flows through the CR component seeking permission to move to the DT component for data transmission.

The CR component in the throughput model, CR(s), is an unknown type queue (RQ) with the feedback probability of $(1 - P_t(G))$ and the input rate λ. The DT(s) component consists of an instant server with the probability of $(P_c + P_d)$ and a G/D/1 queue with the input rate of $\lambda_{tq} = P_t(G)*G$ and the service rate $\mu_{tq} = 1$ respectively. In next section, we will see that in some circumstances, e.g. M >= 3, the $\lambda_{tq} = P_t(G)*G = \lambda$.

The CR component in the delay model, CR(d), which is different from that of the throughput model, is simply an M/M/1 queue with an average input rate λ and average service rate $\mu_{rq} = \ln(1 / (1 - P_t(\lambda)))$. The DT(d) component is an instant server with the probability of $(P_c + P_d)$ and an M/D/1 queue (TQ) with $\lambda_{tq} = P_t(G)*G$. The throughput analysis in the next section proves that when the number of CMSs in a slot is greater than 2, the total input traffic λ will pass through the TQ. Therefore, the $\lambda_{tq} = \lambda \mid M >= 3$.

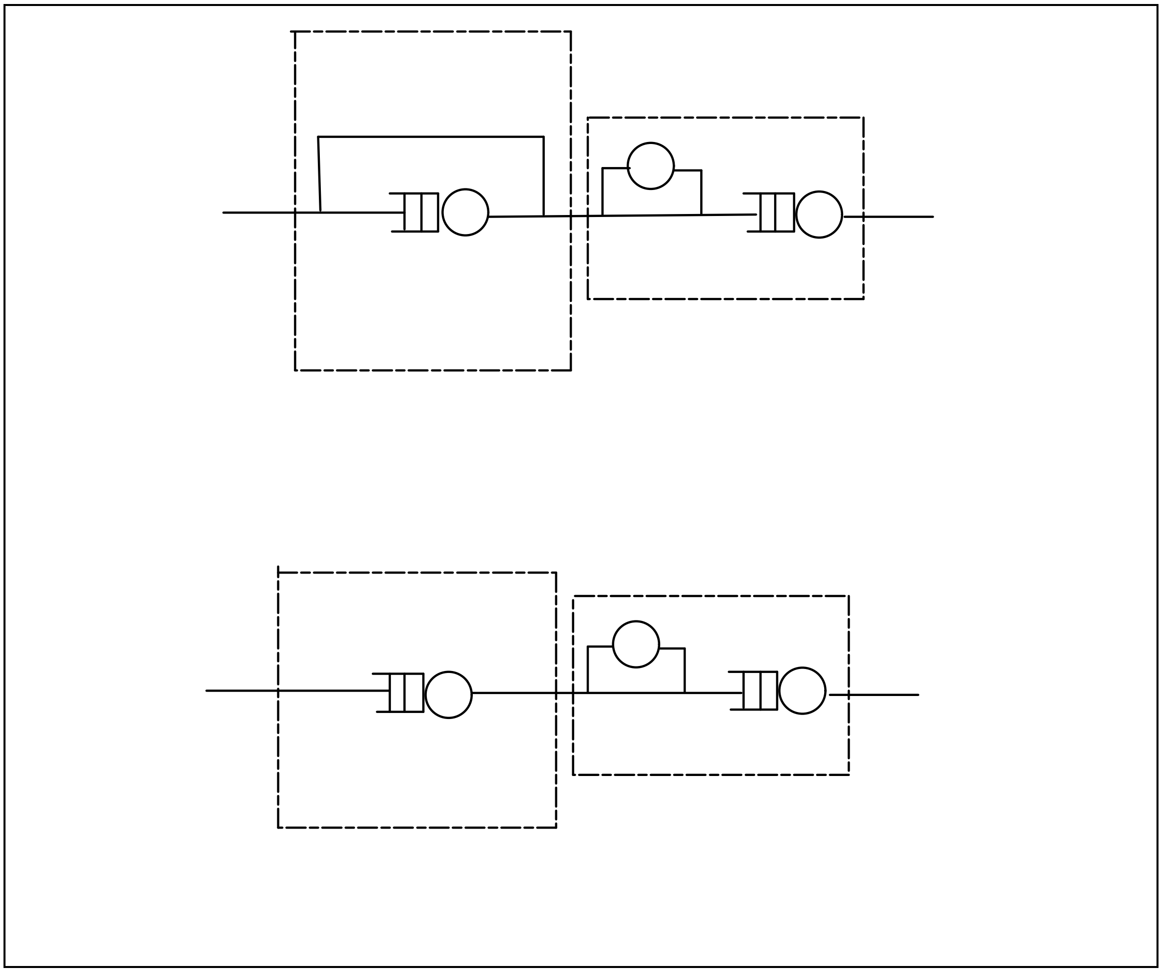


Figure 1 Statistical Model of DQRAP

Notice the difference between the throughput model and the delay model. The former does not care which kind of queues the TQ and RQ are, it requires only the traffic flow. The latter requires a precise queuing model for each of TQ and RQ in order to calculate the delay time.

A brief explanation of the approximate delay model is presented: we can with confidence make the assumption that the input traffic of CR(d) is a Poisson process with input rate λ. Then, based on the rules of DQRAP, the probability of a packet succeeding in a CMS of a slot, $P_t(\lambda)$, is constant with an input rate λ. From the point of view of a discrete random process, the number of unit time (slots) or say RQ service time a packet requires to be successful in a CMS (RQ) follows a geometric distribution with the probability of success of approximately $P_t(\lambda)$. The pdf (probability density function) of the RQ service time from a continuous random process perspective follows an exponential distribution. In other words, the output process of the CR(d) component is a Poisson process. Because the output process of the CR(d) component is the input process of the DT(d) component, the traffic into the DT(d) component or TQ also follows a Poisson distribution. Therefore, we have the model for the delay time.

III. PERFORMANCE ANALYSIS

1. Proof of The Delay Model of DQRAP

In the previous section, the statistical delay model of the protocol was presented. The CR(d) model assumes the RQ service time for a packet follows an exponential distribution. Consequently, the output process of the CR(d) component is a Poisson process. Now, let us prove that this is a reasonable assumption. Each packet succeeds in a CMS with the probability of $P_t(\lambda)$; therefore according to the CRA of the protocol, we can model the RQ service time as a discrete random variable X_i that follows a geometrical distribution. So, we have its pmf (probability mass function):

$$b_{RQ}(X_i) = (1 - P_t)^{Xi-1} * P_t(\lambda) \qquad (3\text{-}1)$$

and its probability distribution function, PDF:

$$B_{RQ}(t) = \sum_{X_i=1}^{[t]} b_{RQ}(X_i)$$

$$= \sum_{X_i=1}^{[t]} (1 - P_t)^{Xi-1} * P_t(\lambda)$$

$$= \frac{P_t(\lambda)}{(1 - P_t(\lambda))} \sum_{X_i=1}^{[t]} (1 - P_t(\lambda))^{Xi}$$

$$= \frac{P_t(\lambda)}{(1 - P_t(\lambda))} \frac{(1 - P_t(\lambda))(1 - (1 - P_t(\lambda))^{[t]}}{(1 - (1 - P_t(\lambda)))}$$

$$= 1 - (1 - P_t(\lambda))^{[t]}$$

$$= 1 - (1 - P_t(\lambda))^{t} \quad t > 0, \qquad (3\text{-}2)$$

In the case where the RQ service time is a continuous random variable, t, we get its pdf by taking the differential of equation (3-2):

$$b_{RQ}(t) = \frac{\partial B_{RQ}(t)}{\partial t}$$

$$= \left(\ln \frac{1}{(1 - P_t(\lambda))} \right) (1 - P_t(\lambda))^t \qquad (3\text{-}3)$$

The Laplace transform of $b_{RQ}(t)$ is:

$$L_{BRQ}(s) = \int_0^{\infty} e^{-st}\, b_{RQ}(t)\, dt$$

$$= \int_0^{\infty} e^{-st} \left(\ln \frac{1}{(1 - P_t(\lambda))}\right) (1 - P_t(\lambda))^t\, dt$$

$$= \frac{\ln 1/(1 - P_t(\lambda))}{\ln 1/(1 - P_t(\lambda)) + s} \qquad (3\text{-}4)$$

Equation (3-4) is an interesting result. Recall that a feature of Laplace transforms is that if two functions have the same Laplace transform then the two functions are identical. We know the Laplace transform of the pdf with exponential distribution has the form μ / (μ + s). Therefore, equation (3-4) shows that the RQ service time for a packet follows an exponential distribution with an average service rate of

$$\mu_{rq} = \ln (1 / (1 - P_t(\lambda))) \qquad (3\text{-}4.1)$$

Here, $P_t(\lambda)$ is the probability that a packet sends a signal in a CMS successfully. Based on the assumption that the traffic arriving at the RQ follows a Poisson distribution, we have,

$$P_t(\lambda) = P(k=0) + P(k=1)M(1/M)(1-1/M) + P(k=2)M(1/M)(1-1/M)^2 +$$

$$... + P(k=n)M(1/M)(1-1/M)^n + ...$$

$= (\lambda^0/0!)e^{-\lambda} + (\lambda^1/1!)e^{-\lambda}((M-1)/M) + (\lambda^2/2!)e^{-\lambda}((M-1)/M)^2 +$

$\ldots + (\lambda^n/n!)e^{-\lambda}((M-1)/M)^n + \ldots$

$= e^{-\lambda} (1 + (1/1!)((M-1)/M)\lambda + (1/2!)(((M-1)/M)\lambda)^2 +$

$\ldots + (1/n!)(((M-1)/M)\lambda)^n + \ldots$

$= e^{-\lambda}e^{-((M-1)/M)\lambda}$

$$= e^{-\lambda/M} \qquad (3\text{-}4.2)$$

The result can be explained intuitively. If the input rate to the system is λ the load of each CMS is λ/M. So, the probability that a CMS is empty is $e^{-\lambda/M}$.

We next consider the departure traffic process of the RQ. Let a(t), $b_{RQ}(t)$ and $d_{RQ}(t)$ denote the pdfs of the interarrival interval, the service time and the interdeparture interval of RQ_i respectively; and let $L_{ARQ}(s)$, $L_{BRQ}(s)$ and $L_{DRQ}(s)$ denote the Laplace transforms of the pdfs respectively. We can present the following equations:

$$d_{RQ}(t) = \rho\, b_{RQ}(t) + (1 - \rho)\, a(t)\, b_{RQ}(t) \qquad (3\text{-}5)$$

and

$$L_{DRQ}(s) = \rho\, L_{BRQ}(s) + (1 - \rho)\, L_A(s)\, L_{BRQ}(s) \qquad (3\text{-}6)\ [9]$$

here, $$L_A(s) = \frac{\lambda_{rq}}{\lambda_{rq} + s} \qquad (3\text{-}7),\ [9],\ [10]$$

$\lambda_{rq} = \lambda$. Substituting (3-4) and (3-7) for $L_A(s)$ and $L_{BRQ}(s)$ in equation (3-6),

$$L_{DRQ}(s) = \rho \frac{\ln 1/(1-P_t(\lambda))}{\ln 1/(1-P_t(\lambda)) + s} + (1-\rho) \frac{\lambda_{rq}}{\lambda_{rq} + s} \frac{\ln 1/(1-P_t(\lambda))}{\ln 1/(1-P_t(\lambda)) + s}$$

$$= \frac{\lambda_{rq}}{\lambda_{rq} + s}$$

$= L_A(s)$ (3-8)

Equation (3-8) shows that the interdeparture interval of the output traffic of RQ follows an exponential distribution with average departure rate of $\lambda_{rq} = \lambda$. The output process which is the same as the departure process thus follows a Poisson distribution. Since the input traffic process of the DT(d) component and the output traffic process of the CR(d) component are identical, the input traffic process of TQ is a Poisson process.

2. Delay Analysis

We consider packet i in the system. After arriving and spending a residual time t_{rdi} waiting for an incoming slot, packet i joins the RQ. After some waiting time, the packet obtains service in the RQ. The total time spent by packet i in the RQ is denoted by t_{rqi}, which is the RQ waiting time plus the RQ service time. According to the CRA rules, the packet can be transmitted in the current slot right after its RQ service. A collision will occur in the current data slot if the length of TQ is zero and more than one packet is output from the RQ. If so, the slot is wasted. The time when the collision occurs is denoted by t_{ci}. Finally, the packet is successfully transmitted after spending t_{tqi} in the TQ. Here, t_{tqi} represents the TQ waiting time plus the TQ service time.

So, the total delay for packet i, t_i is:

$$t_i = t_{rdi} + t_{rqi} + t_{ci} + t_{tqi} \qquad (3\text{-}9)$$

Taking the expected value of (3-9) and averaging over i, we obtain

$$E[t] = E[t_{rd}] + E[t_{rq}] + E[t_c] + E[t_{tq}] \qquad (3\text{-}10)$$

where E[t] is the average delay time of the system, $E[t_{rd}]$ is the average residual time, which is equal to 0.5 unit time (slot), $E[t_r]$ is the average delay time of the RQ, $E[t_c]$ is the average delay time for the collided data slots, and $E[t_{tq}]$ is the average delay time of the TQ.

Here, $E[t_c]$ is different depending on when the feedback information is available. If the type of feedback is IF (Immediate Feedback at the end of a CMS or DS), $E[t_c]$ is zero because at the end of a CMS, every station obtains the feedback and knows the number of successful packets in the current slot. If more than one packet leaves the RQ in the current slot and there is no packet in the TQ, the order of transmission in the current slot follows the sequence in the CMSs which means that the packet succeeding in the first CMS has the highest priority of transmission. Thus, no collision will occur in a data slot. To distinguish between IF and DF(1), we rewrite equation (3-10) as two equations corresponding to the IF and DF cases respectively:

$$E_{IF}[t] = E[t_{rd}] + E[t_{rq}] + E[t_{tq}] \qquad (3\text{-}11)$$

and

$$E_{DF1}[t] = E[t_{rd}] + E[t_r] + E[t_c] + E[t_t] \qquad (3\text{-}12)$$

In this paper, we describe the IF and DF(n=1) cases and defer the case of DF (n>1), the so-called interleaving feedback for further study.

Based on the delay model of the protocol, we can treat the CR component (RQ) as an M/M/1 system. Apply the average service rate μ_{rq} derived in equation (3-4.1) and the input rate λ to the M/M/1 queue, we achieve the average delay of RQ [10],

$$E[t_{rq}] = \frac{1}{\ln 1 / (1 - P_t(\lambda)) - \lambda} \qquad (3\text{-}13)$$

We have proved that the input traffic process to TQ, or say the output traffic from RQ, is a Poisson process. We also assume the size of a slot is one. For TQ, the service time for a packet is constant, one slot. So, following M/D/1 queue analysis [10], we obtain $E[t_{tq}]$ immediately,

$$E[t_t] = \left(\rho + \frac{\rho^2}{2 (1 - \rho)} \right) \frac{1}{\lambda_{tq}}$$

$$= 1 + \frac{\lambda_{tq}}{2 (1 - \lambda_{tq})} \qquad (3\text{-}14)$$

A data slot DS will be wasted by the collision occurring when more than one packet transmits in the DS with DF(1) feedback. The collision in a data slot occurs if and only if the length of the TQ is zero and more than one packet departs from the RQ. So, the probability that a packet will collide in a data slot after departing from RQ, P_c, is:

$$P_c = P (k>0 \mid RQ (k,t)) P (TQ=0) \qquad (3\text{-}15)$$

where, $P(k>0 \mid RQ(k,t))$ is the probability of packets, other than the packet under consideration, leaving the RQ and arriving in the slot. Here, we must note that when a collision occurs in a DS the length of the TQ should be greater than zero at the end of the slot. Then, we have $P(TQ>0) =$

P(TQ Transmission) + P_c = ρ + P_c, and P(TQ>0) + P(TQ=0) = 1. P(TQ Transmission), which is the probability that a packet transmitted in a DS is equal to ρ. So, P (TQ=0) is:

$$P(TQ=0) = 1 - P (TQ>0) = 1 - \rho - P_c \qquad (3\text{-}16)$$

Substituting equation (3-16) in (3-15):

$$P_c = P (k > 0 \mid RQ (k, t)) (1 - \rho - P_c) \text{ and}$$

$$P_c = \frac{(1 - \rho) P (k > 0 \mid RQ (k, t))}{1 + P (k > 0 \mid RQ (k, t))} \qquad (3\text{-}16.1)$$

We also know

$$P(k > 0 \mid RQ (k, t)) = (1 - e^{-\lambda tq}) \qquad (3\text{-}16.2)$$

Substitute equation (3-16.2) in (3-16.1) and note that $\rho = \lambda_{tq}$, we obtain

$$P_c = \frac{(1 - \lambda_{tq}) (1 - e^{-\lambda tq})}{1 + (1 - e^{-\lambda tq})} \qquad (3\text{-}16.3)$$

When a packet leaves the RQ, it will delay one more slot with the probability of P_c derived in equation (3-16.3). Also, if a packet is not involved in a data slot collision but at the time of its leaving the RQ the TQ is busy, the packet still faces the chance of delay because of a slot collided by earlier packets. The probability is,

$$P_d = (1 - P_c) * P \text{ (TQ Transmission)} * P_c \qquad (3\text{-}16.4)$$

where P (TQ Transmission) = ρ = λ_{tq}. The average number of the wasted slots is,

$$E[t_c] = (P_c + P_d) * 1 + (1 - P_c - P_d) * 0$$

$$= (1 + (1 - P_c) P(\text{TQ Transmission})) P_c$$

$$= (1 + (1 - P_c)\, \lambda_{tq})\, P_c$$

$$= \left(1 + \lambda_{tq} \left(1 - \frac{(1 - \lambda_{tq})\,(1 - e^{-\lambda tq})}{1 + (1 - e^{-\lambda tq})} \right)\right) \frac{(1 - \lambda)\,(1 - e^{-\lambda tq})}{1 + (1 - e^{-\lambda tq})} \qquad (3\text{-}17)$$

Substitute 0.5 slot for $E[t_{rd}]$, the RQ delay time in equation (3-13) for $E[t_{rq}]$, the TQ delay time in equation (3-14) for $E[t_{tq}]$ and the collided DS delay time in equation (3-17) for $E[t_c]$ in equation (3-11) and (3-12) we obtain the average delay time for IF and DF(1) feedback cases of DQRAP:

$$E_{IF}[t] = \frac{(3 - 2\lambda)(\ln 1 / (1 - e^{-\lambda/M}) - \lambda) + 2(1 - \lambda)}{2(1 - \lambda)(\ln 1 / (1 - e^{-\lambda/M}) - \lambda)} \qquad (3 - 17.1)$$

$$E_{DF1}[t] = \frac{(3 - 2\lambda)(\ln 1 / (1 - e^{-\lambda/M}) - \lambda) + 2(1 - \lambda)}{2(1 - \lambda)(\ln 1 / (1 - e^{-\lambda/M}) - \lambda)} +$$

$$\left(1 + \lambda \left(1 - \frac{(1 - \lambda_{tq})\,(1 - e^{-\lambda})}{1 + (1 - e^{-\lambda})}\right)\right) \frac{(1 - \lambda)\,(1 - e^{-\lambda})}{1 + (1 - e^{-\lambda})} \qquad (3 - 17.2)$$

G	λ	$E_{DF(1)}[t]$	$E_{sim}[t]$	$E_{m/d/1}[t]$	S	S_{sim}
0.104	0.100	1.9428	1.7157	1.5560	0.0970	0.0968
0.215	0.200	2.1628	1.9658	1.6250	0.1940	0.1932
0.336	0.300	2.3825	2.2530	1.7140	0.2910	0.2900
0.468	0.400	2.6274	2.5929	1.8330	0.3880	0.3875
0.614	0.500	2.9296	2.9836	2.0000	0.4850	0.4820
0.776	0.600	3.3427	3.4851	2.2500	0.5820	0.5800
0.966	0.700	3.9806	4.1875	2.6670	0.6790	0.6745

1.188	0.800	5.1575	5.3790	3.5000	0.7760	0.7710
1.468	0.900	8.2909	8.3146	6.0000	0.8730	0.8715
1.642	0.950	13.8603	13.6809	11.0000	0.9215	0.9200

Table 1. Avg. Delay and Throughput of M/D/1 and DQRAP with M = 3, α = 0.03 and DF(1).

Table 1 shows the average delay and the throughput of both DQRAP with DF(1) feedback and M = 3 and an idealized M/D/1 queue, using both the theoretical analysis and simulation. Table 2 is for the case of IF (immediate feedback).

G	λ	$ED_{IF}[t]$	$E_{sim}[t]$	$E_{m/d/1}[t]$	S	S_{sim}
0.104	0.100	1.8574	1.6534	1.5560	0.0970	0.0968
0.215	0.200	2.0185	1.7826	1.6250	0.1940	0.1932
0.336	0.300	2.2014	1.9092	1.7140	0.2910	0.2900
0.468	0.400	2.4279	2.4123	1.8330	0.3880	0.3875
0.614	0.500	2.7278	2.7645	2.0000	0.4850	0.4820
0.776	0.600	3.1529	3.3097	2.2500	0.5820	0.5800
0.966	0.700	3.8168	3.9214	2.6670	0.6790	0.6745
1.188	0.800	5.0337	4.9657	3.5000	0.7760	0.7710
1.468	0.900	8.2215	8.2156	6.0000	0.8730	0.8715
1.642	0.950	13.8236	13.5934	11.0000	0.9215	0.9200

Table 2. Avg. Delay and Throughput of M/D/1 and DQRAP with M = 3, α = 0.03 and IF.

The simulation figures support the theoretical analysis. Under a light load (load < 0.3), the results of the delay analysis are slightly higher than the simulation figures. This is due to the approximation of the exponential distribution of RQ service time (refer to Equation (3-2)) which at light loads indicates the upper boundary of the delay of the system. Fig. 2 is the diagram of the tradeoff between delay and throughput. An interesting observation is that even though the IF system is collision free the performance is only marginally better than the DF(1) system.

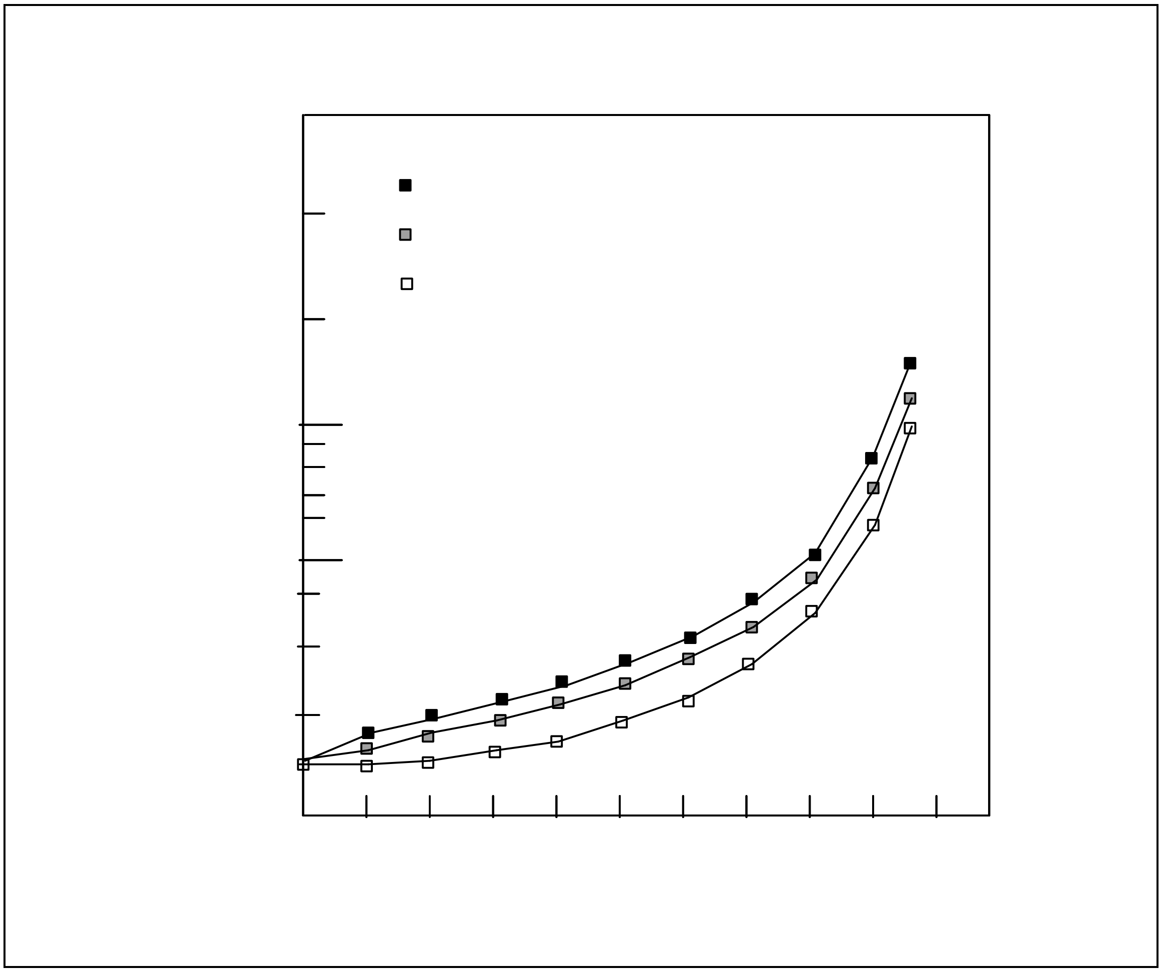


Figure 2 Average Delay vs Througput - M/D/1 and DQRAP

3. Throughput Analysis

We utilize the throughput model introduced by Xu and Campbell [2]. We confirm their results but also make an interesting comparison of throughput, the number of minislots, and Slotted Aloha. The characteristic of the traffic into the DQRAP system is presented by the input rate of λ. The traffic λ and the backlogged traffic from CR component itself form the offered load G of RQ. The traffic G goes through RQ and then splits into two branches of traffic. One returns to CR(s) with the probability of 1 - $P_t(G)$ and the other is the traffic S' which proceeds to the DT(s) component. Refer to the throughput model in Fig. 1, Obviously, S' is a fraction of the offered load G:

$$S' = G * P_t(G) = G * e^{-G/M} \qquad (3\text{-}18)$$

where the derivation of $P_t(G) = e^{-G/M}$ is the same as that of equation (3-4.2). The maximum value of S' can be achieved easily by taking the differential of equation (3-18) and making the result equation equal to zero. So, we obtain:

$$\text{Max } S' = S'(G = M)$$
$$= M * e^{-1}$$

Consider that the throughput of the system is identical with the throughput of the TQ and that a system throughput greater than 1 is unreasonable, we obtain the throughput of the system S:

$$S = \text{Minimum}(S', 1)$$
$$= \text{Minimum}(G * e^{-G/M}, 1) \qquad (3\text{-}18.1)$$

and the maximum throughput of the system is

$$\text{Max } S = \text{Minimum}(\text{Max } S', 1)$$
$$= \text{Minimum}(M*e^{-1}, 1)$$

From the formulas above, we have the following:

$$S = \begin{cases} G * e^{-G/M} & \text{when } S' < 1 \\ 1 & \text{when } S' >= 1 \end{cases} \qquad (3\text{-}18.2)$$

An interesting result. When M = 2, we have $S = G*e^{-G/2}$ from which we obtain a maximum throughput which is twice that of slotted ALOHA! When M = 3, $S = G * e^{-G/3}$ and Max S (M = 3) = Minimum (1.104, 1) = 1. It demonstrates one of the elegant parts of the DQRAP system: a throughput approaching 1 is obtained with 3 CMSs!

Now, consider the question about whether a collision in a data slot with DF(1) feedback has an impact on the throughput. Note that when IF feedback is utilized, DQRAP never wastes a time slot since no data slot collision occurs. There are practical systems which can utilize DQRAP with IF feedback, e..g., radio systems where "a", the ratio of propagation delay to frame size, is very much less than 1. This type of system may be called a lossless system.

In the case of DF(n=1), collision in a data slot may occur. From equation (3-15) and (3-16) in section 3.2, we know the probability of a packet colliding in a data slot is P_c. The debate here is whether or not the collision reduces the throughput of the system. Now, we prove that a collision has no impact on the throughput at any input rate $\lambda <= 1$ by reduction ad absurdum.

First, assume the throughput of the system

$$S' = \frac{N_{pslots}}{N_{sslots} + M} < S = \frac{N_{pslots}}{N_{sslots}} \qquad (3\text{-}18.3)$$

where N_{pslots} is the total number of slots spent on packet transmission; N_{sslots} is the total number of slots spent on the whole system. M is the difference between the number of slots wasted by collision and the number of empty slots. If and only if M > 0, inequality (3-18.3) holds. From Fig. 3, we obtain

$$M = \sum (m_i - n_i)$$

$$= \sum m_i - \sum n_i$$

$$= M_m - N_n$$

$$= P_c * N_{sslots} - P(TQ=0) * N_{sslots}$$

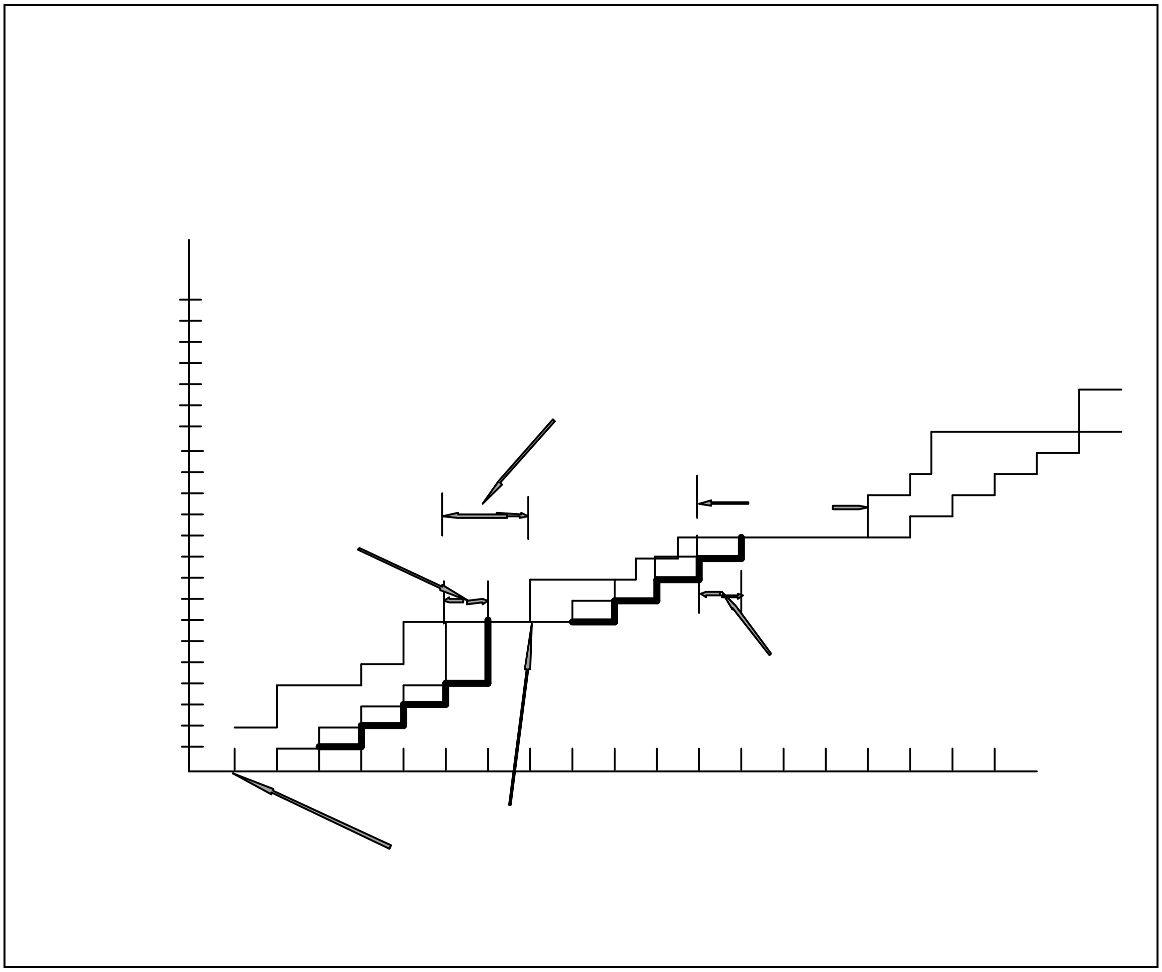


Figure 3 Arrival and Departure of DQRAP

From equation (3-15), we obtain the probability of a collided data slot P_c = P(k>0 | RQ(k,t)) P(TQ=0). Also P(k>0 | RQ(k,t)) is the probability of more than one packet leaving the RQ at a given slot. So, we have,

$$M = P(k>0 \mid RQ (k,t))\, P (TQ=0) * N_{sslots} - P(TQ=0) * N_{sslots}$$

$$= \{P(k>0 \mid RQ(k,t)) - 1\}\, P(TQ=0)\, N_{sslots} \qquad (3\text{-}18.4)$$

In the procedure, the symbols N_m and N_n denote the total number of slots collided and the total number of slots empty respectively. We consider that the part of equation (3-18.4),

$$P(k>0 \mid RQ(k,t)) - 1$$

can never be greater than 0. Therefore, $M \le 0$. From the condition of inequality (3-18.1): $M > 0$, we know inequality (3-18.1) never holds. So, we know the throughput of the system,

$$S' = S \qquad (3\text{-}18.5)$$

Thus DQRAP can achieve 100% throughput with just 3 CMSs! The simulations for both DF(1) and IF feedback strongly surpports the theoretical derivation. From Table 1 and Table 2, we find the throughput of IF feedback which has no collided data slot is exactly the same as that of DF(1) feedback. It further prove form a point of view of simulation the collided data slots have no impact on the throughput when the input rate is less than 1. If we count the overhead of the CMSs and let α denote the fraction, then (3-18.2) can be modified to:

$$S = \begin{cases} (1 - \alpha) * G * e^{-G/M} & \text{when } S' < 1 \\ (1 - \alpha) * 1 & \text{when } S' >= 1 \end{cases} \qquad (3\text{-}18.6)$$

The numerical value of the throughput of the system from both the theoretical calculations and simulation with M = 3 is shown in Table 1. Fig. 4 shows the curve of the throughput S vs. the offered load G with 3 CMSs. In actual applications, whether the medium be fiber optic, copper, or wireless, the CMS overhead will in general be well under 10%.

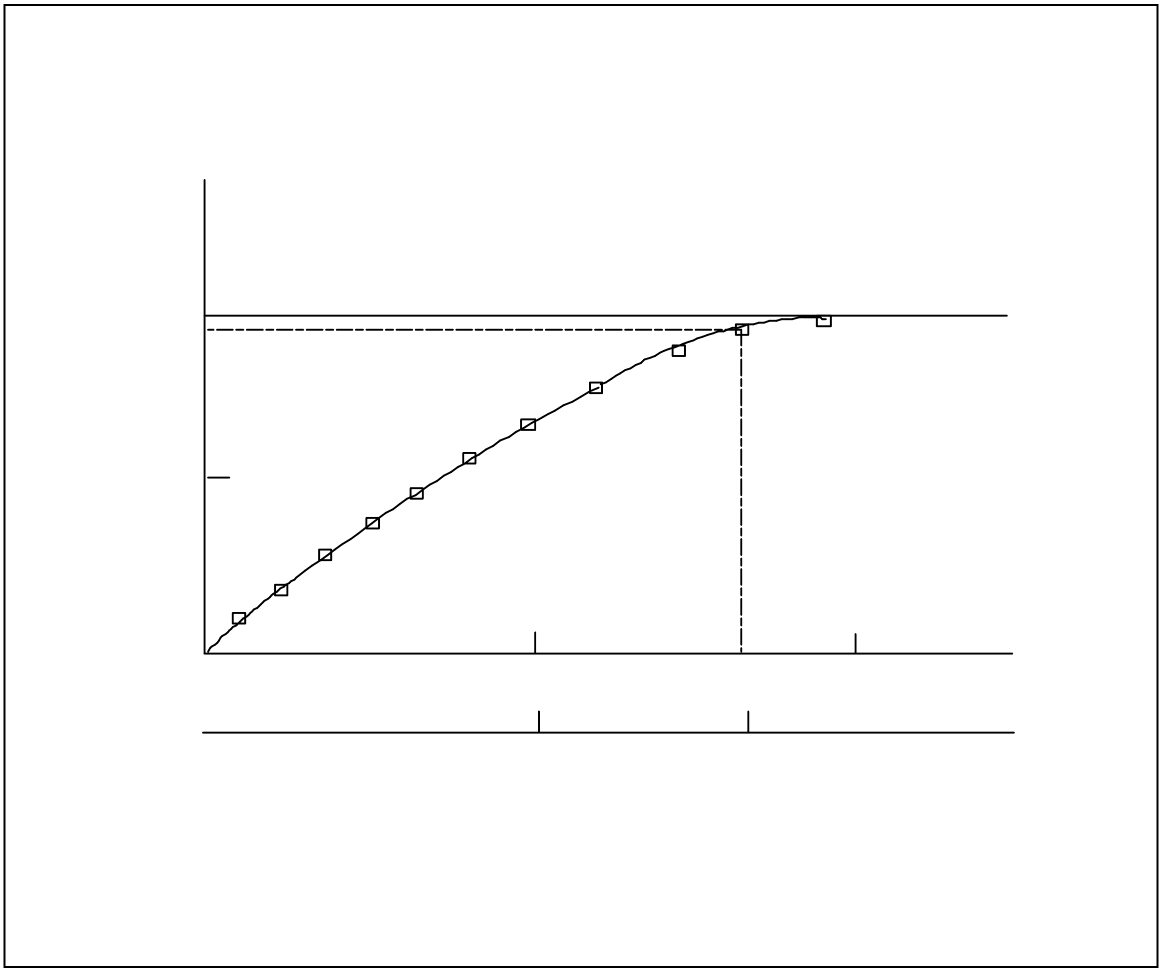


Figure 4 DQRAP Throughput vs Offered Load

When the number of CMSs in a slot increases, the characteristic of the delay time improves since the probability of a packet succeeding in a CMS increases. But, increasing CMSs means reducing the utilization of the channel as the overhead increases. If the system objective is minimum delay to support real-time requirements, the number of CMSs can be increased to achieve the required delay at the expense of throughput. Alternatively if the system priority is high throughput, the number of CMSs, M, may be set at 3. But note that in Fig. 2, which show the average delay with different number of CMSs with DF(1) feedback type, the delay with as few as 3 minislots compares well with the delay of an ideal M/D/1 system.

4. Stability

Stability is discussed from two different points of view. First, we ignore the mathematics and evaluate the protocol using intuition. Saying a system is unstable means that for a particular input rate, the system may transmit backlogged packets continuously but the throughput is almost zero. At the point, a packet may wait an infinite amount of time to complete its transmission. The family of ALOHA protocols belongs in this category. To avoid the unstable point, ALOHAs must introduce a control mechanism to stabilize the system. DQRAP uses the basic splitting concept of Tree protocols. The CRA of a tree protocol can be described thus: when a collision occurs new access to the channel stops and the packets involved in the collision split into two groups and each group then transmits separately. If a collision occurs, the algorithm is applied recursively to each new group until all the packets are transmitted successfully. The difference between the CRA of Tree and DQRAP includes: (1). DQRAP splits the packets into M different groups instead of two; (2). The procedure of the CRA of DQRAP is carried out in CMSs instead of in data slots. These differences are the basis of DQRAPs great performance. Neither a Tree protocol or DQRAP ever enters the situation of no output packets thus both are described as stable multi-access protocols.

Next, we consider the stability of protocols using mathematical analysis. We use terminology similar to that used by Bertsekas and Gallager [7]. For DQRAP, we assume there are n stations connected to the channel and m denotes the number of the backlogged stations. q_a and q_r are the probabilities of a station transmitting a packet for a new station and a backlogged station respectively. Therefore, the offered load for the protocol is

$$G = (n - m) * q_a + m * q_r \qquad (3\text{-}19)$$

The new arrival rate is the first part of (3-19), $(n-m)*q_a$, Fig. 5 shows the relationship between the arrival line and the departure line (the throughput $S = G*e^{-G/M}$). The arrival line is the case of $q_r > q_a$. We are interested in this case because if $q_r < q_a$ the system is always stable. From the curves, we find that as long as the arrival rate stays between 0 and 1 there is a single intersection of the arrival line and the throughput curve. Therefore, DQRAP with 3 or more CMSs is stable at any input rate of $\lambda < 1$.

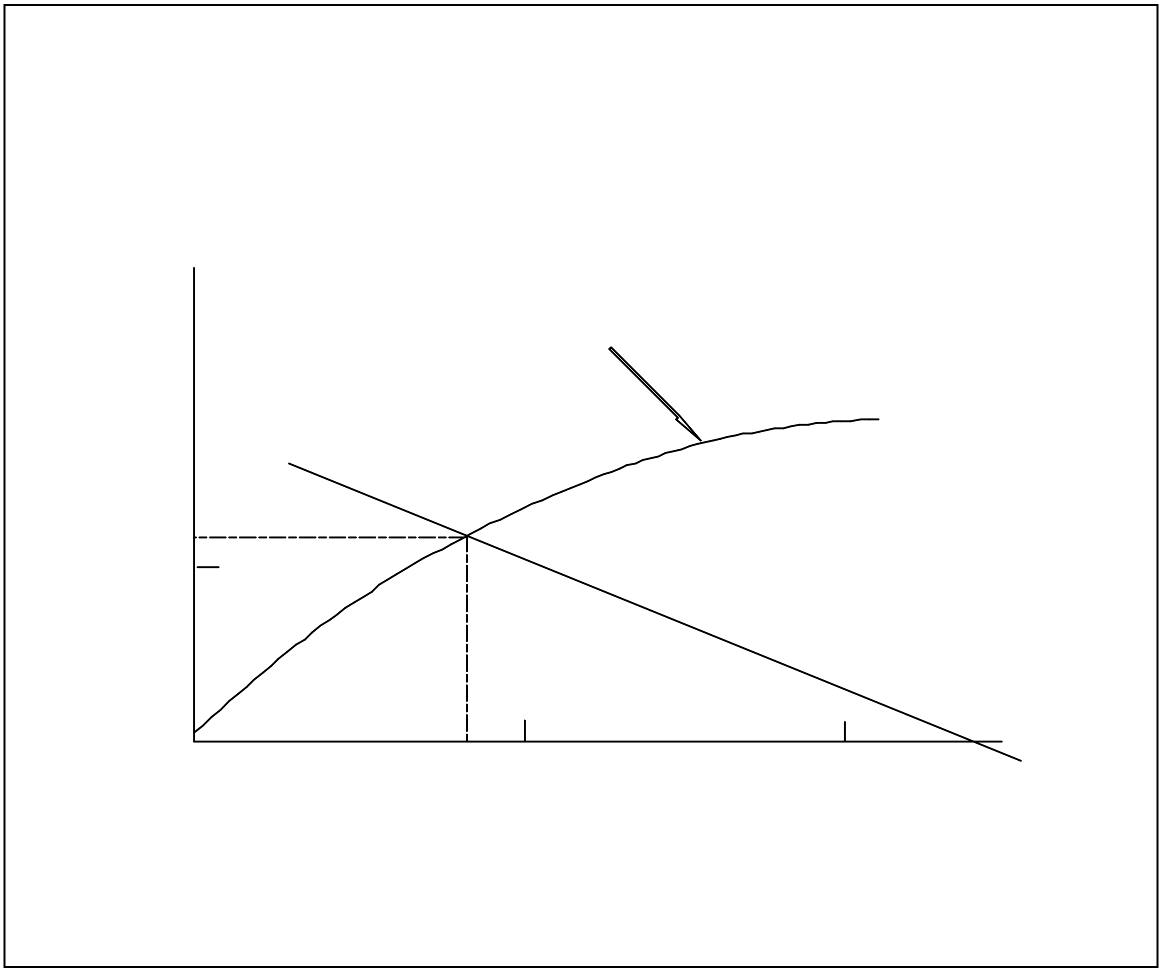


Figure 5 Stability of DQRAP

IV. PERFORMANCE COMPARISON

DQRAP is a stable, high performance multi-access protocol. Fig. 6, Fig. 7 and Fig. 8 compare DQRAP, AARA (a Tree protocol) and CSMA protocols. From both the throughput and average delay point of view, DQRAP is superior to AARA and the CSMA family of protocols. The most attractive feature of DQRAP is that even when its throughput reaches 0.95, the average delay time still is under 14 slots with M, the number of CMSs, equal to 3. From Fig. 7, we note that even though AARA and DQRAP use a similar CMS approach, the delay feature of DQRAP with M = 3 is better than AARA with M=10.

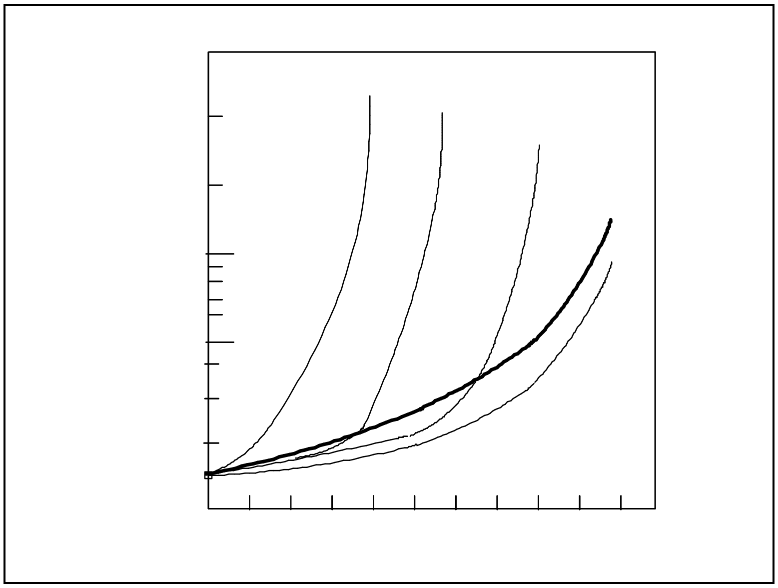
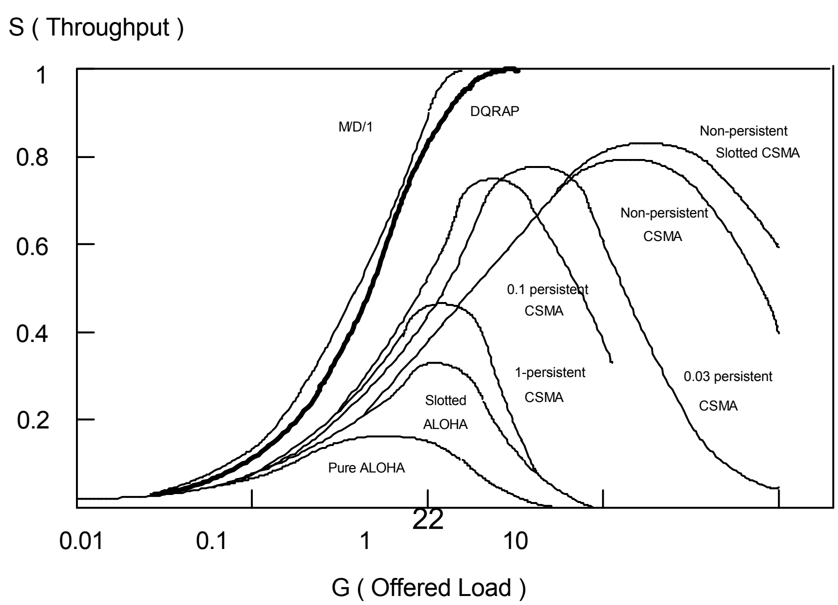


Figure 6 Delay Comparison: CSMA, DQRAP and M/D/1

V. CONCLUSION

The throughput model for DQRAP developed by Xu [2] and described by Xu and Campbell [16] has been verified. A mathematical model of the delay characteristics of DQRAP is introduced. The analytical results are in general agreement with simulation results. This confirms the claim that DQRAP provides a level of performance which is superior to any other practical protocol suitable for a broadcast channel. Research is underway into how DQRAP may be used in satellite communication, high speed WANs, integrated voice and data LANs, packet radio networks, data transmission in CATV systems, high capacity backbones, and parallel busses. Basic research remaining to be carried out on DQRAP includes analysis of binary feedback, robustness, interleaving feedback, and implementation issues for the specific applications.


S (Throughput)
1
0.8
0.6
0.4
0.2
M/D/1
DQRAP
Non-persistent
Slotted CSMA
Non-persistent
CSMA
0.1 persistent
CSMA
1-persistent
CSMA
0.03 persistent
CSMA
Slotted
ALOHA
Pure ALOHA
0.01
0.1
1
10
G (Offered Load)

REFERENCES

[1] Wenxin Xu, Graham Campbell, "A Near Perfect Stable Random Access Protocol For a Broadcast Channel", ICC 1992 Proceedings, Vol 1, pp 0370-0374.

[2] Wenxin Xu, "Distributed Queueing Random Access Protocols for a Broadcast Channel", Ph.D Dissertation, IIT, 1990.

[3] Dipankar Raychaudhuri, "Announced Retransmission Random Access Protocol", IEEE Transactions On Communications, Vol.COM-33, November 1985.

[4] Bernd-Peter Paris, Behnaam Aazhang, "Near-Optimum Control of Multiple-Access Collision Channels", IEEE Transactions On Communications, Vol. 40, August 1992.

[5] F. A. Tobagi, "Carrier Sense Multiple Access with Message-Base Priority Functions", IEEE Transactions On Communications, Vol.COM-30, January 1982.

[6] J. F. Hayes, "Modeling And Analysis Of Computer Communication Networks", Plenum Press, New York and London, 1984.

[7] Dimitri Bertsekas, Robert Gallager, "Data Networks", Prentice-Hall International, Inc., 1987.

[8] Don Towsley, Pedro Orlando Vales, "Announced Arrival Random Access Protocols", IEEE Transactions On Communications, Vol.Com-35, May 1987.

[9] Kishor S. Trivedi, "Probability & Statistics With Reliability, Queuing, and Computer Science Applications", Prentice-Hall, INC., Englewood Cliffs, NJ 07632, 1982

[10] Leonard Kleinrock, "Queuing Systems", John Wiley and Sons INC., NY., 1976

[11] Ralph L. Disney, Peter C. Kiessler, "Traffic Processes in Queuing Networks--A Markov Renewal Approach", The John Hopkins University Press, Baltimore and London, 1987.

[12] Leonard Kleinrock, "Performance Evaluation of Distributed Computer-Communication Systems", Queuing Theory and its Applications, Elsevier Science Publisher B.V., pp 1-57, 1988.

[13] Jean Walrand, "An Introduction to Queuing Networks", Prentice-Hall International Editions, 1988.

[14] Mischa Schwartz, "Computer-Communication Network Design and Analysis", Prentice- Hall, INC., Englewood Cliffs, NY 07632, 1977.

[15] Then Tang Liu, Don Towsley, "Window and Tree Protocol For Satellite Channel", in Proc. IEEE Infocom. Theory, Vol. IT-29, pp. 910-914, November 1983.

[16] Wenxin Xu, Graham Campbell,"A Distributed Queueing Random Access Protocol For a Broadcast Channel", Submitted to SIGCOM 93.