Computer Networks 62 (2014) 43–54

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Achieving destination differentiation in ingress aggregated fairness for resilient packet rings by weighted destination based fair dropping

Mete Yilmaz [a,*], Nirwan Ansari [b]

[a] Edge Routing Business Unit, Cisco, CA 95134, USA
[b] Department of Electrical and Computer Engineering, New Jersey Institute of Technology, NJ 07102, USA

ARTICLE INFO

Article history:
Received 29 July 2012
Received in revised form 16 December 2013
Accepted 6 January 2014
Available online 10 January 2014

Keywords:
MAN
WAN
Ring networks
Spatial reuse
Fairness
Active queue management

ABSTRACT

The IEEE 802.17 is a standardized ring topology network architecture, called the Resilient Packet Ring (RPR), to be used mainly in metropolitan and wide area networks. This paper introduces destination differentiation in ingress aggregated fairness for RPR and focuses on the RPR MAC client implementation of the IEEE 802.17 RPR MAC in the aggressive mode of operation. It also introduces an enhanced active queue management scheme for ring networks that achieves destination differentiation as well as higher overall utilization of the ring bandwidth with simpler and less expensive implementation than the generic implementation provided in the standard. The enhanced scheme introduced in this paper provides performance comparable to the per destination queuing implementation, which is the best achievable performance, while providing weighted destination based fairness as well as weighted ingress aggregated fairness. In addition, the proposed scheme has been demonstrated via extensive simulations to provide improved stability and fairness with respect to different packet arrival rates as compared to earlier algorithms.

1. Introduction

The IEEE 802.17 Resilient Packet Ring (RPR) is a standardized network architecture based on ring topology deliberated by the IEEE LAN/MAN Standards Committee [1]. The ring is established by using point-to-point bidirectional connections between stations in RPR. The RPR network has two counter-rotating rings called ringlets. The RPR protocol utilizes both ringlets to overcome connection disturbances, and hence provides resilience. In addition, the RPR protocol provides a fairness mechanism among stations in terms of sharing the overall ring bandwidth applied to the best effort traffic.

The ring network is simpler to operate and administer than a complex mesh or an irregular network while virtually providing connection from each station to every other station [2]. SONET based ring networks have been already deployed by service providers in MANs or WANs. However, in these networks, many SONET rings consist of a dual-ring configuration in which one of the rings is used as the back-up ring and remains unused during normal operation to be utilized only in the case of a failure of the primary ring [2]. The static bandwidth allocation and network monitoring requirements increase the total cost of a SONET network. While plain Ethernet does not require static allocation and provides cost advantages, it cannot provide desired features such as fairness and auto-restoration [2]. In addition, the RPR standard is a media independent standard which can be used on top of SONET or Ethernet network to obtain the above advantages.

---

* Corresponding author. Tel.: +1 4088530767.
E-mail addresses: myilmaz@cisco.com (M. Yilmaz), nirwan.ansari@njit.edu (N. Ansari).

1389-1286/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bjp.2014.01.003

The fairness aspects of RPR have been investigated in depth in light of interesting scenarios as those described in [3–5]. Improvements for the current fairness algorithm of the IEEE 802.17 have also been proposed as reported in [3,4,6–8]. In addition, in our earlier work we introduced weighted fairness and its use cases in [9,10]. The basic algorithm relies on identifying a single congestion point on the ring. This decreases the ring utilization and spatial reuse in certain traffic scenarios. The standard provides means to keep track of all the congested downstream nodes and their fair rates. A generic way to utilize this mechanism fully is to implement 255 separate queues (one queue per destination, including the multicast queue) within the MAC client. Supporting multiple destination queues at the MAC layer is not efficient and requires additional memory or complex shared memory queuing implementations. In addition, this may not be even possible to implement on the already deployed hardware. This paper presents a simpler and less expensive MAC client implementation that does not require separate queues while providing ring utilization comparable to the generic implementation. Our previous work [11] showed that it is possible to utilize an active queue management scheme at the MAC client level. However, the implementation that will be presented in this article achieves destination differentiation at the source stations by allowing different weights per destination while achieving RPR ingress aggregated fairness. In addition, the proposed implementation provides a higher degree of stability under different packet arrival rates and patterns. This is achieved by penalizing the flows which exceed their fair shares more than the others.

The rest of this paper is organized as follows. In Section 2, the RPR operation and fairness algorithm will be discussed with emphasis on MAC client implementations. Section 3 will discuss the MAC client implementation utilizing our proposed active queue management scheme based on the approximate fair dropping algorithm [12]. In Section 4, behavior of the current (aggressive) fairness algorithm with MAC client implementations of single queue, multiple queues, and two other algorithms with active queue management of a single queue will be illustrated through simulations. Section 5 provides simulation results for the modified scenario that provides destination differentiation. In Section 6, comparison of a large ring hub scenario will be provided. Finally, the conclusion will be drawn in Section 7.

In order to demonstrate different operational modes, some performance figures of merits are included and discussed. The scenarios have been executed on the RPR simulator model developed at Cisco during the RPR standardization process. The simulation model is implemented in Opnet. The suggested modification has been incorporated into the simulator model and its behavior is verified through simulations.

2. RPR operation and fairness

The operation of the RPR protocol has been discussed in detail in [2]. In this paper, a brief overview based on an example will be provided.

An RPR network can have a maximum of 255 stations. Each station on the ring removes a frame from the ringlet if that frame is destined for that station (except multicast frames, which are removed by the sender). This facilitates the spatial reuse property of an RPR network since the bandwidth will not be consumed by that frame around the ringlet.

Fig. 1 shows an example scenario. Stations *4* and *3* are transmitting to Station *2* on the inner ringlet, while Stations *2* and *1* are transmitting to Station *7*. In addition Station *2* has traffic destined to Station *1*. Each time a frame is received at Station *2* from the ring, the frame will be checked to see if that frame is destined to the station itself or not. The frame will be removed from the ring if it is destined to the station (in this case, any frame from Station *3* or *4*). In parallel, Station *1* will receive frames from Station *2*, and Station *7* will be able to receive the frames from Stations *1* and *2* without being impacted by the traffic generated at Stations *3* and *4*.

The objective of the fairness algorithm is to distribute the unallocated bandwidth around the ring among stations in a fair manner. In the case of Fig. 1 (assuming that all the stations have equal weights), Stations *3* and *4* will get an equal amount of the link bandwidth on the link between Stations *2* and *3*, while Stations *1* and *2* will get an equal amount of the link bandwidth on the link between Stations *1* and *7*. More detailed explanation of the standardized fairness algorithms can be found in [1,2].

In short, the fairness algorithm provides a fair sharing of the link bandwidth according to the weights of the stations when there is more traffic than that can be transmitted through that link. RPR fairness is based on ingress aggregation. This fairness is referred to as “Ring Ingress Aggregated with Spatial Reuse” (RIAS) fairness in an earlier article [3]. This definition follows the same methodology used in [13] for max–min flow control. The RIAS fairness definition, however, does not include the station weights in the generalized formula while this is included in the IEEE 802.17 standard in the calculation of the estimated fair rate of a node. In addition, the RIAS definition assumes equal sharing of the bandwidth among flows originating from the same station, while the standard does not require that. In our previous works [9,10], we have expanded this definition with the inclusion of source station weights. In this section, a more general definition will be provided with the inclusion of the destination station weights along with

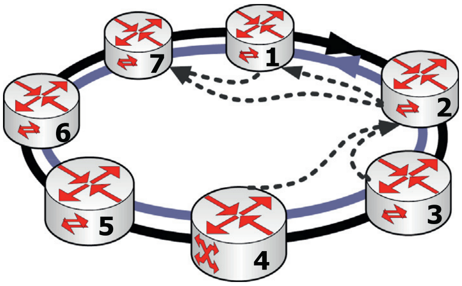

Fig. 1. Destination stripping and spatial reuse illustrated on the inner ring (ringlet 1).

the source station weights to provide a more comprehensive representation of fairness as in the IEEE 802.17 standard.

Denote N as the total number of stations on a ringlet. Let the capacity of each link on the ringlet be C. Each Station s on the ringlet is given a weight w_s for providing the weighted fairness. On this ringlet, a flow vector is defined by $\boldsymbol{F} = \{f_{st}\}$, in which each flow from Station s to Station t is denoted by f_{st} which is also referred to as the path of the flow. For each flow a weight vector is defined by $P = \{\rho_{st}\}$, in which the weight of each flow from Station s to Station t will be denoted by ρ_{st}. A fair rate vector is defined by $\mathbf{R} = \{r_{st}\}$ in which the fair rate of flow f_{st} is denoted by r_{st}. By using the above definitions, the total allocated rate on link n of the ringlet is given by (1).

$$T_n = \sum_{\forall s,t:\ link\ n \in f_{st}} r_{st} \tag{1}$$

On this ringlet, the vector $\mathbf{R}$ is said to be feasible if the following conditions in (2) and (3) are met.

$$r_{st} > 0 \quad \forall s,t:\ f_{st} \in \boldsymbol{F} \tag{2}$$

$$T_n \leqslant C \quad \forall n \in \mathrm{N}:\ 0 < n \leqslant N \tag{3}$$

The sum of all flows originating from Station s and passing through link n is

$$A_n(s) = \sum_{\forall t \in N:\ link\ n \in f_{st}} r_{st} \tag{4}$$

For a feasible vector $\mathbf{R}$, the link n is a bottleneck link, $\mathbf{B_n}(s,t)$, with respect to $\mathbf{R}$ for f_{st} crossing link n if the following conditions in (5)–(7) are met with respect to all flows $f_{s't'}$ crossing link n.

$$T_n = C \tag{5}$$

$$r_{s't'} \leqslant r_{st} \quad \forall s',t':\ s = s' \,\&\, t' \neq t \,\&\, link\ n \in f_{s't'} \tag{6}$$

$$A_n(s') \leqslant A_n(s) \quad \forall s',t':\ s' \neq s \,\&\, link\ n \in f_{s't'} \tag{7}$$

Note that if there are no other flows originated from any station other than Station s going through link n, $A_n(s')$ will be zero and (7) will be satisfied by default.

The vector $\mathbf{R}$ is said to be "weighted" ingress aggregated fair with destination differentiation if it is feasible as defined in (2) and (3) and if for each f_{st}, r_{st} cannot be increased while maintaining feasibility without decreasing the fair rate $r_{s't'}$ of some flow $f_{s't'}$ for which

$$\frac{r_{s't'}}{\rho_{s't'}} \leqslant \frac{r_{st}}{\rho_{st}} \quad \forall s',t':\ s' = s \,\&\, f_{s't'} \in F \tag{8}$$

$$\begin{gathered} \frac{A_n(s')}{w_{s'}} \leqslant \frac{A_n(s)}{w_s} \\ \forall s',t',n,:\ s' \neq s \,\&\, f_{s't'} \in F \,\& \\ link(n) \in f_{s't'} \,\&\, link(n) \in f_{st} \end{gathered} \tag{9}$$

Eq. (8) ensures the fairness among the flows originating from the same station with destination differentiation, while (9) ensures the fairness among ingress aggregated flows. The weights ρ and w are used to normalize the comparisons, and hence to achieve the weighted fairness for both destination flows and ingress aggregation, respectively.

For the scenario given in Fig. 1, if Station *4* is weighed two times that of Station *3*, it will get two times more bandwidth out of the ring than Station *3*. In this case, if Station *3* increases its share, (9) will not be satisfied. Destination differentiation can be provided at Station *2* if the destination weights (ρ_{st21} and ρ_{27}) are adjusted so that the destination Station *1* has two times more weight than Station 7. Then the Station *1* will receive two times more traffic from Station *2* compared to Station *7* in order to satisfy Eq. (8).

The standard defines two mechanisms to distribute the fairness information around the ring. The first mechanism is used to distribute the fair rate of the nearest congested station to the upstream stations. This fairness message is called the single-choke fairness frame (SCFF) in the standard [2]. The second mechanism is used to propagate the fair rate of each station to all the other stations. This fairness message in RPR is called the multi-choke fairness frame (MCFF) [2]. Each station on the ring broadcasts how much of its output link is used by the station itself in its MCFF and sends it to all the stations on the ring. A receiving station may collect these messages and build a global view of the congestion situation on the ring, and schedule the traffic to add to the ring accordingly. As shown in our previous article [16], this information can even be used to support non-uniform capacity links in an RPR network efficiently.

Fig. 2 shows the separation of RPR MAC and its client. RPR MAC transfers MCFF and SCFF messages via the control path indication, while the MAC client transmits and receives the packets via data path request and indication messages, respectively. MAC Datapath Sublayer handles the transmission and reception of the frames to and from the dual ringlets. The implementation details of an RPR MAC client layer is not part of the standard; however, in Appendix J of the standard [1], some examples are provided for a single queue implementation that utilizes SCFF and another that utilizes MCFF with virtual output queues. By shaping traffic according to the MCFF messages at the MAC client, one can increase the bandwidth utilization by avoiding single congestion points. In Section 3, an RPR

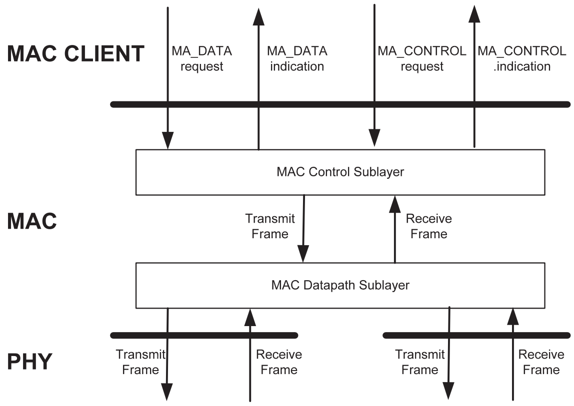

Fig. 2. RPR MAC services model.

MAC client implementation that utilizes a modified Approximate Fair Dropping (AFD) algorithm [12] will be shown.

In a simple RPR MAC client, a single low priority queue handles the packets as they are waiting to enter the ring. After a station sends its fair share of traffic to the network, the RPR MAC will not allow any more packets to be sent and this will cause head of line (HOL) blocking. The RPR MAC client can also support separate queues per destination for low priority packets to eliminate HOL blocking. This multi-queue implementation of an RPR MAC client keeps track of SCFF and MCFF messages, and schedules packets in a deficit round robin fashion as long as they are allowed by the RPR MAC. During this scheduling, the important point to consider is to avoid sending more packets beyond the congested nodes than allowed. Otherwise, the RPR MAC will not allow transmission of any packet for a period of time. However, implementing multiple queues at an RPR MAC client is costly. This article proposes an RPR MAC client with an active queue management scheme and will be referred to as weighted Destination Based Fair Dropping (wDBFD). Similar to a multi-queue RPR MAC client, this MAC client also needs to keep track of SCFF and MCFF messages. Instead of implementing multiple queues, only a single queue is needed, and packets are probabilistically dropped before they are accepted into the queue instead of waiting for tail drops based on queue occupancy and RPR fairness messages. Furthermore, by adding weights to the drop probabilities among different destination stations, destination differentiation is achieved. This algorithm is more stable in terms of fairness when subject to different packet arrival rates than the algorithm proposed in [11] as shown in Section 4.

3. Weighted destination based fair dropping for RPR

Similar to the other active queue management techniques (e.g., RED [14]), weighted DBFD also utilizes the queue size to accept or drop packets on arrival. However, wDBFD not only relies on past measurements of the queue size but also on recent observed rates of flows. By using this additional information wDBFD can provide fairness among different flows [15].

Approximation of virtual output queuing using single FIFO queue with active queue management is shown with the simulation results in Section 4. The implementation of wDBFD in RPR requires a modification of the AFD algorithm since in RPR the received fair rate from a congested station changes the packet drop probabilities of all flows destined to stations downstream to that congested station (excluding the congested station itself) once the aggregated rate of flows exceeds the received fair rate. Therefore, the MAC client needs to actively adjust the drop probability of each packet to each destination by considering the received fair rates from congested stations. While providing fairness among destinations, the MAC client will not need to implement 255 destination queues with the wDBFD algorithm. Instead, this scheme utilizes per destination counters in the MAC client (most of which are already necessary for a multi-queue implementation of the standard). Thus, the hardware implementation is simplified and allows microcode based implementations on presently deployed hardware.

For the wDBFD algorithm, consider a ring with the source Station s and the destination Station d as shown in Fig. 3. On this ring, assume Station i as the station that has the minimum fair rate in between Station s and Station d. Also define Station j as any arbitrary downstream station beyond Station s.

On this ringlet, $\mathbf{U}_s$ is denoted as the received fair rate vector at Station s, where $\mathbf{U}_s$ is the set of fair rates obtained from SCFF or MCFF sent by all the downstream stations in between Station s and Station d at time t. If u_i is the minimum of all fair rates received, then Station i is the most congested station in between the source Station s and the destination Station d.

$$u_i = \min(\mathbf{U}_s) \tag{10}$$

Define a flow vector, $\mathbf{F} = \{f_{sd}\}$, in which the arrival rate of a flow to Station s destined to Station d is denoted by f_{sd}. Define another flow vector, $\mathbf{R} = \{r_{sj}\}$, in which a flow sourced by Station s destined to Station j is denoted by r_{sj}. By using the above definitions, the total traffic sourced by Station s that is destined beyond Station i is given by (11).

$$k_s(i) = \sum_{\forall j:\ \text{Station}\, j > \text{Station}\ i} r_{sj} \tag{11}$$

Define Q_{target}, α_1, α_2 as arbitrary constants, $dbfd'_s$ as the previous value of the wDBFD rate, q_s as the length of the queue at Station s in bytes, and q'_s as the previous value of the queue length, then the current wDBFD rate at Station s is given by (12). Note that the details of adjusting parameters α_1 and α_2 can be found in [15].

$$dbfd_s = \begin{cases} Q_{\text{target}} & q_s = 0 \\ dbfd'_s - \alpha_1(q_s - Q_{\text{target}}) + \alpha_2(q'_s - Q_{\text{target}}) & q_s \neq 0 \end{cases} \tag{12}$$

Define β as an arbitrary constant to normalize the wDBFD rate with respect to the rate measurement, which is equal to LINK_RATE [1] divided by Q_{target}. Define f_{sd} as the arrival rate of a flow at Station s destined to Station d, then the drop probability of a packet destined to Station d will be given by (13).

$$p_s(d) = \begin{cases} 1 & w_s \cdot u_i < k_s(i) \\ 1 - \frac{\rho_{sd}\cdot\beta\cdot dbfd_s}{f_{sd}} & (w_s \cdot u_i \geqslant k_s(i)) \text{ and } (f_{sd} \geqslant \rho_{sd}\cdot\beta\cdot dbfd_s) \\ 0 & (w_s \cdot u_i \geqslant k_s(i)) \text{ and } (f_{sd} < \rho_{sd}\cdot\beta\cdot dbfd_s) \end{cases} \tag{13}$$

As defined in (13), if the received fair rate (u_i) is less than the total traffic sourced by Station s destined beyond Station i, then the traffic destined beyond Station i will be dropped. If the Station s is still allowed to send more traffic beyond Station i, then based on the wDBFD rate and the arrival rate (f_{sd}) the traffic may be randomly dropped. The

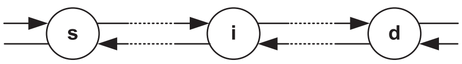

Fig. 3. Multi destination scenario.

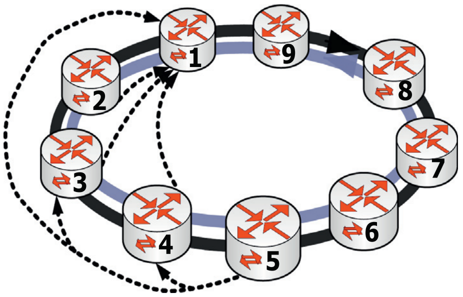

Fig. 4. Multi destination scenario.

drop probability goes up if the arrival rate of a flow is much higher than the wDBFD rate. If a flow continues to send at the higher rate, it will be penalized more by increasing the drop probability. This allows the algorithm to be more stable with respect to different packet arrival rates. If the arrival rate of traffic is higher than the departure rate from the queue, the algorithm will operate at buffer occupancy of Q_{target}. The algorithm will establish a certain packet mix in the buffer so that the ratio of packets destined to each station will correspond to the ratio of allowed fair rates to each destination. In addition, the drop probabilities are also adjusted by weight ρ_{sd} to allow different destinations receive different proportions of the available bandwidth.

The current single queue implementation of RPR may result in severely underutilized rings in some scenarios, and oscillations in utilization are also observed in those scenarios [3]. With the proposed mechanism, the ring utilization can easily approach the theoretical limit of the number of links multiplied by the bandwidth with a relatively simple implementation. The advantage of this algorithm for RPR is that it improves the performance of an RPR ring, and it is backward compatible with the standard which is not true for the previously proposed solutions. In addition, it does not require 255 independent queues to be implemented in the scheduling hierarchy. In general adding an additional level to the scheduling hierarchy is not possible without requiring new hardware.

While the implementation of wDBFD algorithm does not require 255 queues, it requires additional calculation stages on top of the current RPR fairness algorithm. Specifically, the station will need to keep track of fairness messages from all other stations in a 255 entry array. In addition, at each RPR parameter calculation interval called decay interval, the “dbfd_fair” rate needs to be decayed and smoothed out. The highest computational complexity comes from the case when a queuing decision needs to be made at each packet arrival. One can employ the sampling algorithm as proposed in [12] to allow rate estimations at certain intervals so as not to burden the system with calculating the rates at each packet arrival.

4. Performance evaluation of weighted destination based dropping

The example scenario shown in Fig. 4 will be used to compare the performance of different MAC client implementations. In this scenario, Stations *2*, *3*, *4* and *5* have traffic destined to Station *1*. Station *5* has also traffic destined to Stations *3* and *4*. Stations *2*, *3*, and *4* start sending traffic to Station *1* at time 1 s. Station *5* starts sending traffic to Stations *1* and *4* at time 1 s. Station *5* then starts sending traffic to Station *3* at time 2 s. In this scenario, the traffic demand of all but one flow at each node is OC12 rate per activeflow from one station to another. Station 5 receives two times more traffic to destination Station 3 than the other stations in order to demonstrate the stability of the wDBFD algorithm.

Per RIAS fairness [3], the 620 Mbps bandwidth on the link between Stations *1* and *2* should be equally shared among four stations resulting in 155 Mbps per station. Station *5* can utilize more bandwidth without impacting this fairness by utilizing the unused 465 Mbps bandwidth on link between Stations *5* and *4* and also a maximum available bandwidth of 310 Mbps on link between Stations *4*

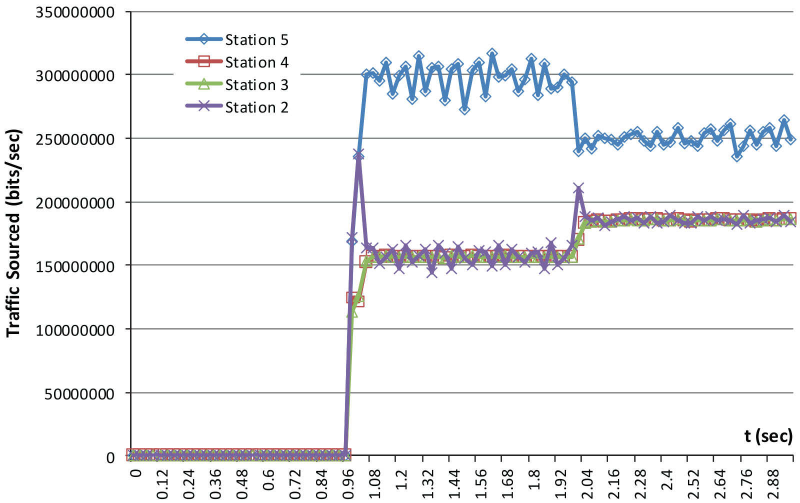

Fig. 5. Actual traffic sourced at Stations *2*, *3*, *4* and *5* with single MAC client queue.

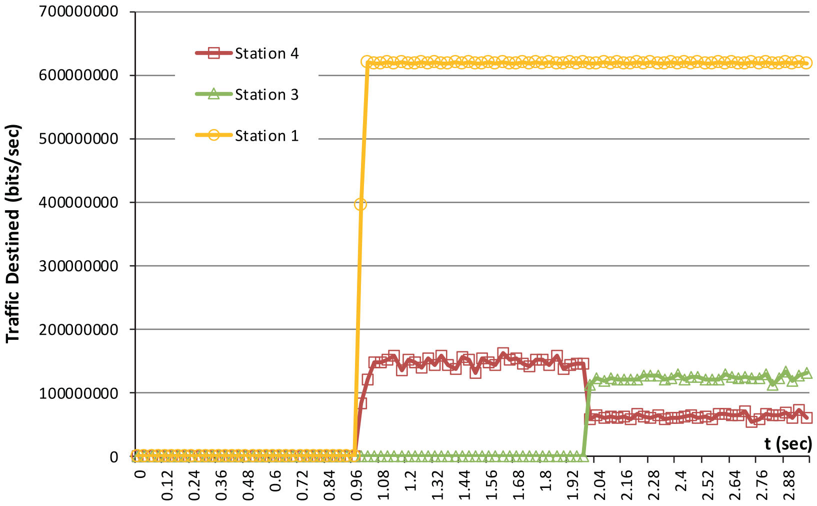

Fig. 6. Traffic received at Stations *1*, *3* and *4* with single MAC client queue.

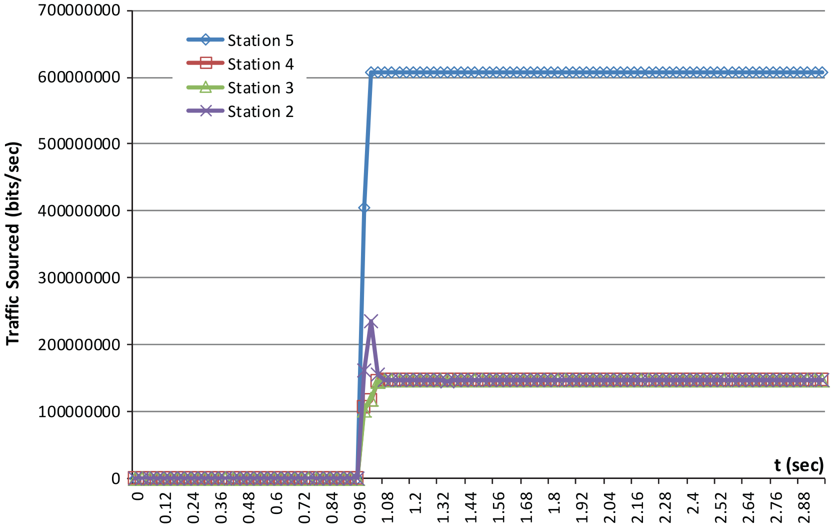

Fig. 7. Actual traffic sourced at Stations *2*, *3*, *4* and *5* with VoQ.

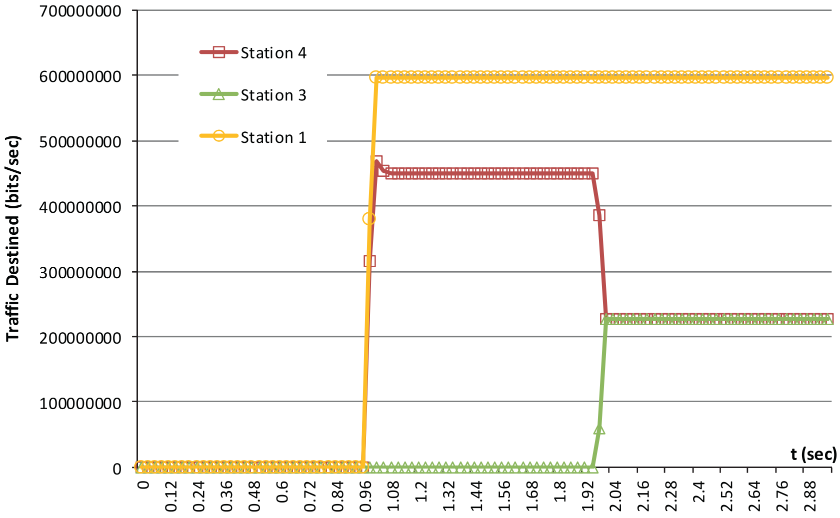

Fig. 8. Traffic received at Stations *1*, *3* and *4* with VoQ.

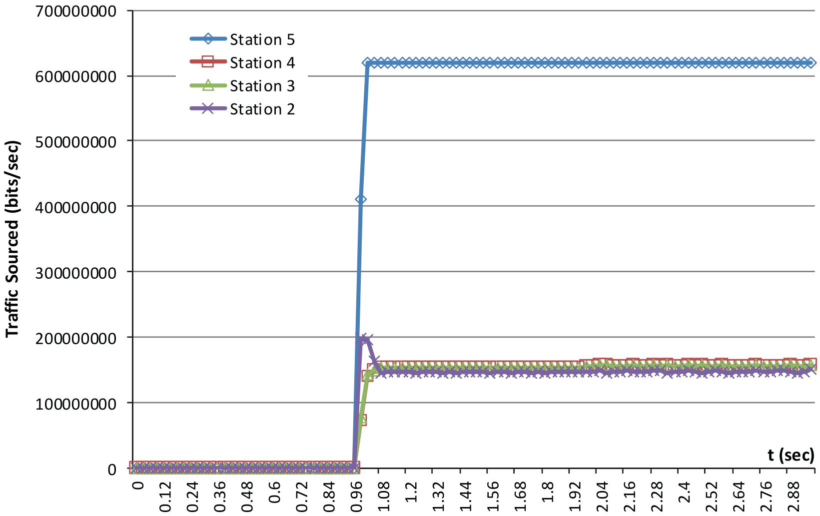

Fig. 9. Actual traffic sourced at Stations *2*, *3*, *4* and *5* with DBFD.

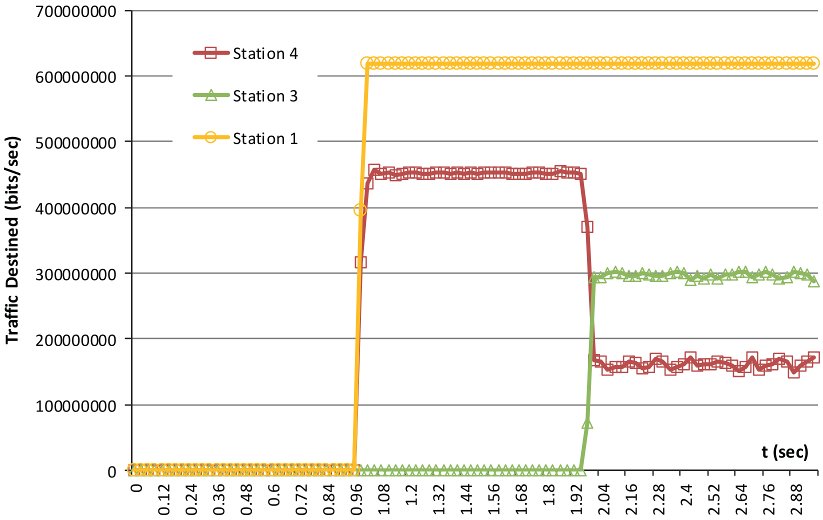

Fig. 10. Traffic received at Stations *1*, *3* and *4* with DBFD.

and *3*. In the ideal case, Stations *1*, *3* and *4* should receive 620 Mbps, 232.5 Mbps, and 232.5 Mbps, respectively. Therefore, the total bandwidth utilization on the ring will be 1085 Mbps based on be "weighted" ingress aggregated fair with destination differentiation definition. This scenario will be used to compare the performance of different MAC client implementations.

The scenario is simulated using the single queue RPR model implemented in the Opnet simulator. An OC12 ring which is composed of 9 stations is created with 20 km of distance between every two adjacent stations. Each station is configured as a dual-queue station with the aggressive fairness mode enabled. The size of the secondary transit queue (STQ) at each station is 512 KB and the "LP_COEFF" [1] parameter of the RPR MAC is set to 4.

Fig. 5 shows the total traffic sourced by Stations *2*, *3*, *4* and *5* to the outer ringlet. As expected, the available bandwidth is being shared equally by Stations *2*, *3*, and *4*, while Station *5* is able to get more bandwidth out of the ring by utilizing the unused bandwidth on the links. However, at time 2 s, once Station *5* starts sending traffic to Station *3*, the fairness message generated by Station 4 limits the total traffic that can be sourced by Station *5*.

As shown in Fig. 6, Station *1* receives the full 620 Mbps of traffic, while as shown in Fig. 5 Station *5* is not able to utilize the unused bandwidth fully. In addition, once Station *5* starts sending packets to Station *3* at time 2 s, the fairness is lost and Station *5* is not able to get its fair share of the bandwidth, and Stations *2*, *3* and *4* start sourcing more traffic to Station *1* than Station *5*. The total ring utili-

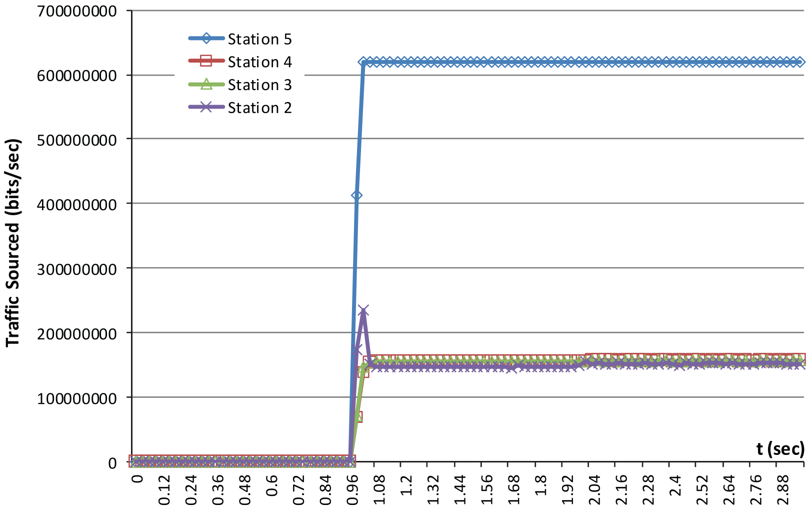

Fig. 11. Actual traffic sourced at Stations *2*, *3*, *4* and *5 with wDBFD*.

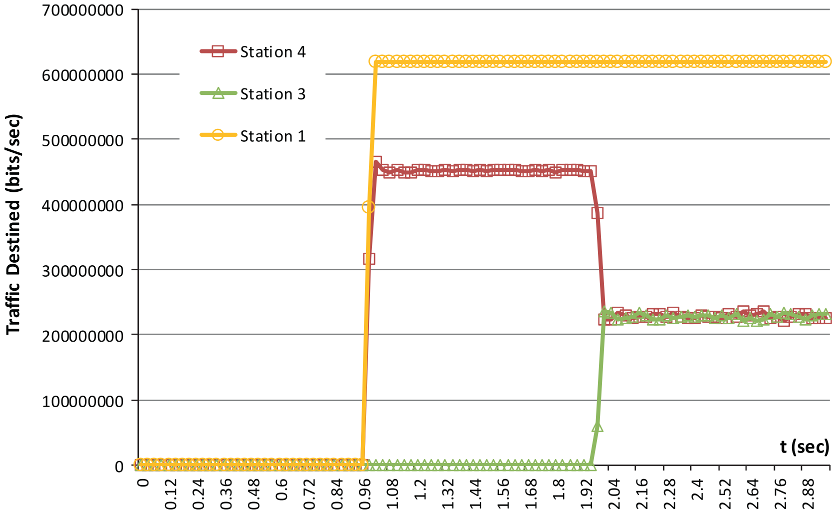

Fig. 12. Traffic received at Stations *1*, *3* and *4 with wDBFD*.

zation is 810 Mbps instead of the expected 1085 Mbps. Oscillations observed around the steady state behavior show less stability.

One other observation is that after time 2 s, Station *3* receives more traffic than Station *4* even though the traffic is sourced by the same Station *5*. The main reason for this behavior is the imbalanced traffic demand used in this scenario and the simple single queue implementation not being able to maintain fairness at the source station per destination. Therefore, destination fairness is not achieved.

The same scenario is simulated with the RPR MAC client model using virtual output queues (VoQ) as explained in the standard. Fig. 7 shows the actual traffic sourced at stations *2*, *3*, *4*, and *5*. In this case, the oscillations are minimized, and a steady response is observed at time 2 s, when Station *5* starts sending traffic to Station *3*.

Fig. 8 shows the traffic received at Stations *2*, *3*, and *4*, respectively. The observed bandwidth matches the expected values, and provides a maximum bandwidth utilization of 1085 Mbps. In addition, the destination fairness is achieved with respect to traffic received at Station *3* and Station *4*, since virtual output queuing is able to maintain destination separation with respect to different packet arrival rates for each destination.

Next, the same scenario is simulated using RPR MAC client with DBFD as explained in [11]. Figs. 9 and 10 show the actual traffic sourced and received from and at respective stations. The steady response is observed after time 2 s when Station *5* starts sending traffic to Station 3 and the total ring utilization reaches up to the expected 1085 Mbps. Similar to the single queue RPR implementation, the destination fairness is not achieved for the traffic sourced by Station *5* to the destination Stations *3* and *4*. Specifically, Station 3 receives almost two times more traffic than Station *4*. Since the packet arrival rate destined to Station *3* after time 2 s is two times more than the packet

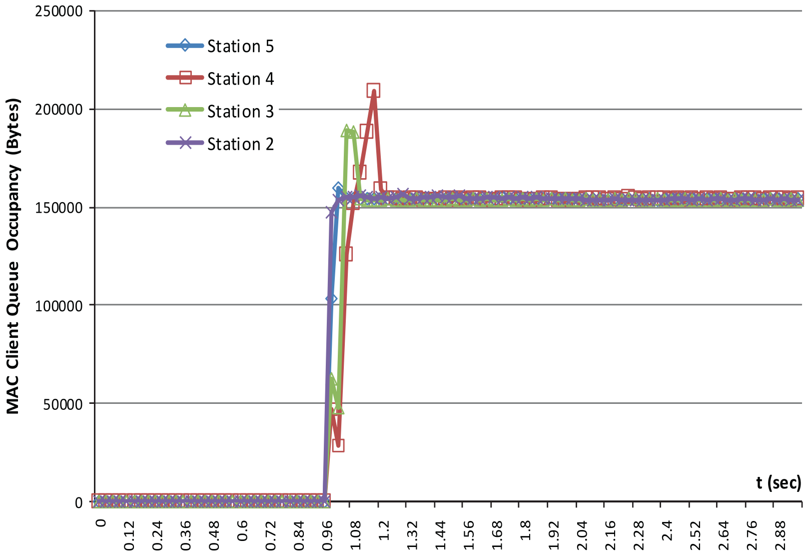

Fig. 13. MAC client queue occupancy of Stations *2*, *3*, *4* and *5 with wDBFD.*

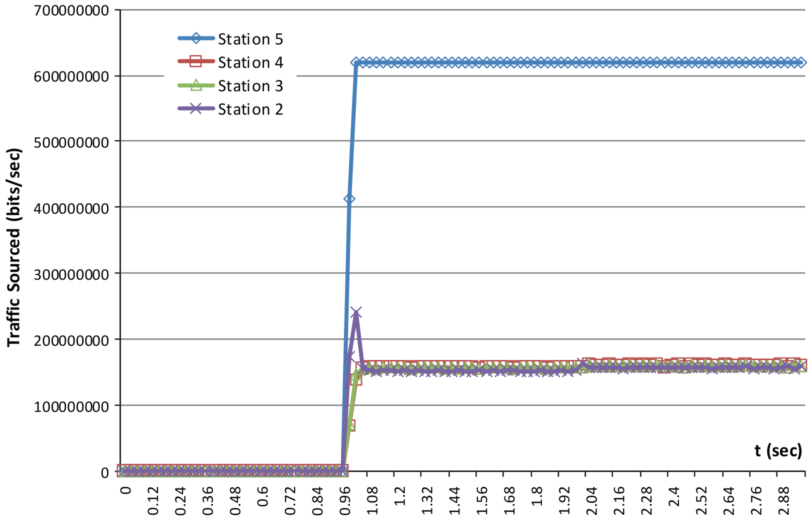

Fig. 14. Actual traffic sourced at Stations *2*, *3*, *4* and *5 with wDBFD.*

arrival rate destined to Station *4* and the arrival rate is not regulated per destination, this undesirable behavior is expected for the DBFD algorithm as well.

Next, the same scenario is simulated with the wDBFD algorithm as explained in Section 3. The results are shown in Figs. 11 and 12.

The results shown in Figs. 11 and 12 are very similar to the ideal case with virtual output queues. Destination fairness at Station *5* is maintained regardless of its packet arrival rate difference from destination Stations *3* and *4*. This is made possible by increasing the probability of packet drop per flow according to the ratio of the arrival rate of a flow to the "DBFD" rate. This provides stability and fair sharing of the queues even when the arrival rates and/or packet sizes are different among flows.

Another aspect to consider is the convergence speed. The interval required for the ring traffic to converge to the fair rate is also impacted by the desired queue occupancy (Q_{target}). Based on the RPR fairness messages and the destination weights, the MAC client queue will have the right mix of packets to match the fair rates to each destination. As the desired queue occupancy increases, the convergence to fair rates will take longer. If the desired queue occupancy (Q_{target}) is set too low, unnecessary packet drops can be observed. This scenario has been tested with different packet arrival patterns per destination. Regardless of the packet arrival rates and packet sizes, the destination stations receive a similar amount of traffic as shown in Figs. 11 and 12.

For this case, the transmit buffer occupancy of Stations *2*, *3*, *4* and *5* are shown in Fig. 13. In this scenario, the desired queue occupancy (Q_{target}) is set to 150000 bytes. As shown in Fig. 13, the MAC client queues converge to the desired queue occupancy after a short transition phase.

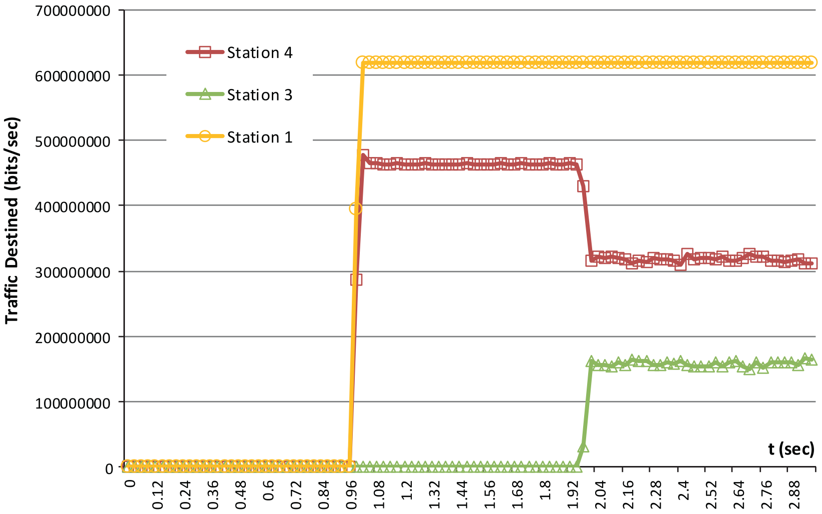

Fig. 15. Traffic received at Stations *1*, *3* and *4 with wDBFD*.

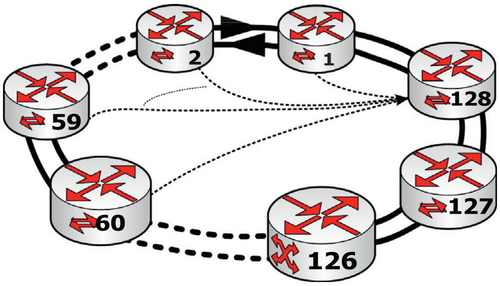

Fig. 16. Large ring hub scenario.

5. Providing destination differentiation through wDBFD

When desired, the proportion of traffic destined from station *s* to station *d* can be adjusted by the ρ_{sd} parameter as described in Section 3 with the wDBFD algorithm presented in this article. This adjustment relies on adjusting the drop probabilities of flows per destination with respect to each other as well as buffer occupancy. The ρ_{sd} parameters will dictate the ratio of packets in the MAC client queue destined to different stations. In this section, the scenario is modified such that the destination Station *4* is given a weight of 2 ($\rho_{54} = 2$) while the flows destined to Stations *1* and *3* are each assigned to weight of 1.

Figs. 14 and 15 show the simulation results with the destination weight adjustment ($\rho_{54} = 2$) utilizing the wDBFD algorithm. In this case, Station *5* can send two times more traffic to Station *4* than the other destinations. Specifically, as shown in Fig. 15, the destination Station *5* is able to transmit approximately 310 Mbps of traffic to Station *4*, while it transmits 155 Mbps of traffic to the Stations *1* and *3*. This shows that the wDBFD algorithm can efficiently provide destination differentiation as required even when the amount of traffic destined to Station *3* is much higher than the amount of traffic destined to Station *4*. Note that while this is a static weight allocation, the wDBFD algorithm does not restrict any bandwidth when

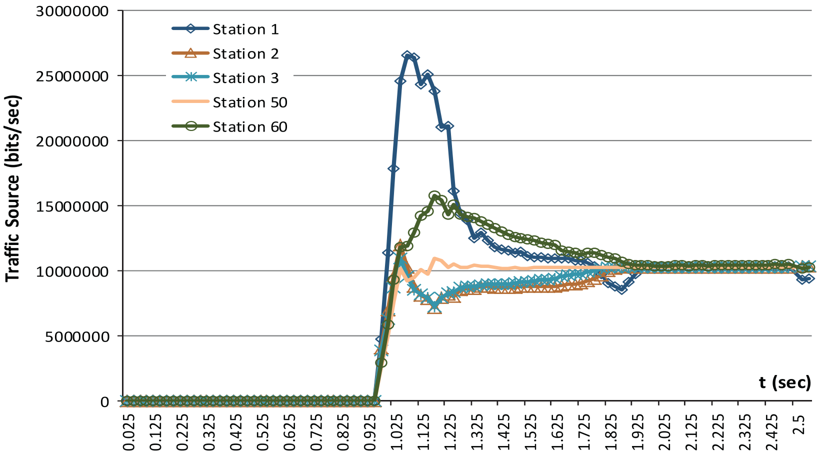

Fig. 17. Actual traffic sourced at Stations *1*, *2*, 3, *50* and *60* with single MAC client queue.

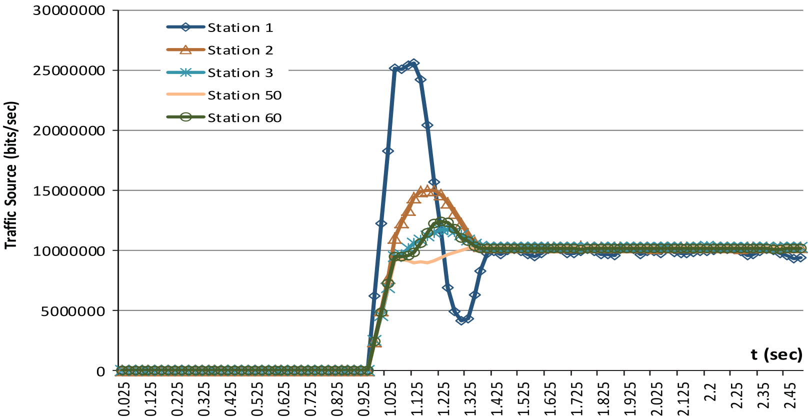

Fig. 18. Actual traffic sourced at Stations *1*, *2*, 3, *50* and *60* with DBFD.

there is no bottleneck and provides fair sharing of the ring bandwidth among destination nodes from one station.

6. Hub scenario in large RING

In this section, performance results will be provided with single MAC queue and DBFD algorithm for a large RPR ring in a hub scenario as shown in Fig. 16.

This scenario requires the control algorithm of RPR to run over the large ring with the ring being composed of 128 stations. The weight (w) of each station is assigned to 1. Each of the stations from Station *1* to *60* has 25.6 Mbps of traffic destined to Station *128* and the ρ_{sd} of each flow is set to 1. On an OC12 ring, each station should be able to transmit 10 Mbps of the traffic destined to Station *128* when bandwidth is shared equally on the ring.

Figs. 17 and 18 show the results for traffic sourced by each station from Station *1* to *60* with single MAC client queue and DBFD algorithms, respectively. Both of the implementations are able to converge to the desired throughput of 10 Mbps. The details of the parameters that can affect convergence were studied earlier in [5,10]. During the initial start of the traffic the Station *1* is able to transmit up 25 Mbps, which is close to the maximum traffic sourced by Station *1* to Station *128*. However, the DBFD algorithm shows a quicker convergence enabled by the active packet dropping mechanism at the MAC client level based on the system load instead of relying solely on the RPR MAC algorithm convergence.

7. Conclusion

In this paper, we have introduced a new fairness scheme that works with the ingress aggregated RPR fairness by including destination differentiation at the source station. In addition, this paper has discussed the proposed implementation of an efficient active queue management mechanism to utilize multi-choke fairness in an RPR network while providing weighted destination based fairness. We also showed the stability of this algorithm in a very large ring as well in Section 6. This algorithm and the principles can be applied to other ring networks to provide fair distribution of bandwidth across the ring as long as the fairness information is distributed among stations. As compared to our earlier implementation, this algorithm provides better isolation of flows with respect to different arrival rates. As shown in this paper, while preserving fairness among stations, this approach has improved the utilization of the underlying network as compared to the single queue implementation of the standard. In addition, the mechanisms discussed in this article do not require any modifications to the standardized IEEE 802.17 RPR fairness mechanism and allow a simpler and computationally less intensive implementation than the generic multi queue implementation discussed in the standard.

References

[1] IEEE Standard 802.17-2004, Resilient Packet Ring.
[2] F. Davik, M. Yilmaz, S. Gjessing, N. Uzun, IEEE 802.17 resilient packet ring tutorial, IEEE Commun. Mag. 42 (3) (2004) 112–118.
[3] V. Gambiroza, P. Yuan, L. Balzano, Y. Liu, S. Sheafor, E. Knightly, Design, analysis, and implementation of DVSR: a fair high-performance protocol for packet rings, IEEE/ACM Trans. Netw. 12 (1) (2004) 85–102.
[4] F. Davik, A. Kvalbein, S. Gjessing, Performance evaluation and improvement of non-stable resilient packet ring behavior, in: Proc. Part II of the 4th International Conference on Networking (ICN'05), Ser. LNCS 3421, Reunion Island, April 17–21 2005, pp. 551–563.
[5] F. Davik, A. Kvalbein, S. Gjessing, An analytical bound for convergence of the resilient packet ring aggressive mode fairness algorithm, in: Proc. 40th Annual IEEE International Conference on Communications (ICC'05), Seoul, Korea, May 16–20 2005.
[6] F. Alharbi, N. Ansari, Distributed bandwidth allocation for resilient packet ring networks, Comput. Netw. 49 (2) (2005) 161–171. <http://www.science direct.com>.
[7] F. Alharbi, N. Ansari, SSA: simple scheduling algorithm for resilient packet ring networks, IEE Proc. Commun. 153 (2) (2006) 183–188.
[8] W. Tang, C. Chang, A Fuzzy inter-ring route control with PRNN predictor bridged resilient packet rings, in: Proc of IEEE Globecom, December 2010.
[9] M. Yilmaz, N. Ansari, Weighted fairness in resilient packet rings, in: Proc. of the 2007 IEEE International Conference on Communications (ICC'07), Glasgow, June 24–28 2007, pp. 2192–2197.
[10] M. Yilmaz, N. Ansari, Weighted fairness and correct sizing of secondary transit queue in resilient packet rings, J. Opt. Commun. Netw. 2 (11) (2010) 944–951.
[11] M. Yilmaz, N. Ansari, J.H. Kao, P. Yilmaz, Active queue management for MAC client implementation of resilient packet rings, in: Proc. of the International Conference on Communication (ICC'09), Dresden, Germany, June 14–18 2009, pp. 1–5.

[12] R. Pan, L. Breslau, B. Prabhakar, S. Shenker, Approximate fairness through differential dropping, SIGCOMM Comput. Commun. 33 (2) (2003) 23–39.
[13] D. Bertsekas, R. Gallager, Data Networks, Prentice-Hall, 1987.
[14] S. Floyd, V. Jacobson, Random early detection gateways for congestion avoidance, IEEE/ACM Trans. Netw. 1 (4) (1993) 397–413.
[15] C. Hollot, V. Misra, D. Towsley, W. Gong, On designing improved controllers for AQM routers supporting TCP flows, in: Proc. of IEEE INFOCOM, April 2001.
[16] M. Yilmaz, N. Ansari, Resilient packet rings with heterogeneous links, in: Proc. of the 2012 IEEE Symposium on Computers and Communications (ISCC'12), Cappadocia, July 1–4 2012.

Mete Yilmaz received his B.S. (1995) and M.S. (1999) degrees from Bogazici University, Istanbul, Turkey, in electrical engineering and computer engineering, respectively. He received his Ph.D. in computer engineering from New Jersey Institute of Technology, Newark, in 2013. He works at Cisco Systems, San Jose, California as a Technical Leader with responsibilities ranging from architecting new line cards to developing packet processing FPGAs for Cisco's high-end routers. He has been working on communication networks for fifteen years. His research interests include quality of service, bandwidth reservation, and fairness algorithms. He holds three patents on fairness algorithms.

Nirwan Ansari received the B.S.E.E. (summa cum laude, gpa = 4.0/4.0) from the New Jersey Institute of Technology (NJIT), Newark, in 1982, the M.S.E.E. degree from University of Michigan, Ann Arbor, in 1983, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1988. He joined NJIT's Department of Electrical and Computer Engineering as Assistant Professor in 1988, tenured Associate Professor in 1993, and Full Professor since 1997. He has also assumed various administrative positions at NJIT. He authored Computational Intelligence for Optimization (Springer, 1997, translated into Chinese in 2000) with E.S.H. Hou, and edited Neural Networks in Telecommunications (Springer, 1994) with B. Yuhas. His current research focuses on various aspects of broadband networks and multimedia communications. He has also contributed over 350 technical papers, over one third of which in widely cited refereed journals/magazines. For example, one of his seminal works was the sixth most cited article published in the IEEE Transactions on Parallel and Distributed Systems. He was/is serving on the Advisory Board and Editorial Board of eight journals, including as a Senior Technical Editor of IEEE Communications Magazine (2006–2009). He had/has been serving the IEEE in various capacities such as Chair of IEEE North Jersey COMSOC Chapter, Chair of IEEE North Jersey Section, Member of IEEE Region 1 Board of Governors, Chair of IEEE COMSOC Networking TC Cluster, Chair of IEEE COMSOC Technical Committee on Ad Hoc and Sensor Networks, and Chair/TPC Chair of several conferences/symposia. He has been frequently invited to deliver keynote addresses, distinguished lectures, tutorials, and talks. Some of his recent awards and recognitions include an IEEE Fellow (Communications Society), IEEE Leadership Award (2007, from Central Jersey/Princeton Section), the NJIT Excellence in Teaching in Outstanding Professional Development (2008), IEEE MGA Leadership Award (2008), the NCE Excellence in Teaching Award (2009), and designation as an IEEE Communications Society Distinguished Lecturer (2006–2009, two terms).