Optimal Bandwidth Assignment for Packet Rings

Cunqing Hua, Tak-Shing Peter Yum and Cheng Li

Abstract: **The network throughput is an important performance criteria for the packet ring networks. Since maximizing the network throughput can lead to severe bias in bandwidth allocation among all flows, fairness should be imposed to prevent bandwidth starvation. The challenge here, therefore, is the joint optimization of the network throughput and fairness. In this paper, we present the optimal bandwidth assignment scheme to decompose this optimization problem into two tasks, one for finding fair bandwidth assignment and the other for finding the optimal routing. The network throughput is maximized under the fairness constraints when these tasks are performed iteratively.**

Index Terms: **Bandwidth assignment, fairness, optimization, packet rings**

I. Introduction

In the past few years, the rapidly increasing volume of data traffic has driven the evolution of packet-based transport in Metropolitan Area Networks(MAN). However, current MAN technologies such as SDH/SONET and Gigabit Ethernet (GigE) have their limitations. For example, SDH/SONET can provide point-to-point guaranteed bandwidth through fixed circuit allocation, but it is not efficient for bursty data traffic because the bandwidth reserved for one flow cannot be used by others. Gigabit Ethernet offers statistical multiplexing and better bandwidth utilization, but it lacks the provision of fairness in bandwidth assignment.

Resilient Packet Ring (RPR) [1] is a standard (IEEE 802.17) targeted for solving the inherent limitations of existing MAN technologies. It is essential a MAC layer protocol over various physical layer technologies. It offers services to various transport layers so that both the networks and the services can be made scalable. Thus, operators can deploy a single technology that is both efficient (due to statistical multiplexing capability) and reliable (due to the fault tolerant dual ring topology).

RPR adopts a dual counter-rotating ring topology. Both rings are used for transporting data and control packets [1], [2]. Routing in RPR appears to be simple, packets either go clockwise or counter-clockwise, which is determined by the shortest path routing algorithm. A salient feature of RPR is that it offers spatial reuse through destination-stripping. That is, packets are removed from the ring at the destination station so that the complementary ring segments can be used for transmitting packets of other non-overlapping flows.

This work was supported in part by the Hong Kong Research Grants Council under Grant CUHK 4220/03E.

Cuning Hua is with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, email: chua0@ie.cuhk.edu.hk

Tak-Shing Peter Yum is with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, email: yum@ie.cuhk.edu.hk

Cheng Li is with the NXP Semiconductors. email: cli0@ie.cuhk.edu.hk

This spatial reuse feature can improve the link utilization by allowing transmission of nonoverlapping flows. However, it may lead to the starvation problem whereby some heavy-loaded nodes can dominate the ring links and leave the other nodes unserved. This problem is critical since the bandwidth starvation of a station may affect the service of a large number of associated users. To address this problem, Gandalf [3] and Aladdin [4] protocols were proposed to coordinate the fair distribution of bandwidth across all stations on the ring. However, it was found that these two protocols may exhibit severe oscillation in unbalanced traffic scenarios[5], [6], [7]. Addressing this problem, the *Distributed Virtual-time Scheduling in Rings(DVSR)* algorithm was proposed in [5], whereby each node computes a lower bound of temporally and spatially aggregated virtual time by counting the packet arrivals. By exchanging this information, each ingress node can approximate the Generalized Processor Sharing (GPS) [8] and determine the per-destination fair rate for the incoming flows. In [6], the *Distributed Bandwidth Reallocated in Rings (DBRR)* algorithm was proposed such that each node combines the local and remote information and transmits the results to the downstream nodes for calculating the fair rate per-destination flows. A rate-based fairness control algorithm adopting the local fairness concept in [9] was proposed in [7] and the end-to-end delay bound was derived.

In this paper, we consider the bandwidth assignment problem for a generic dual ring network rather than specific for RPR. We assume that the offered traffic is characterized by a traffic *demand matrix*. Multipath routing is used so that traffic can be split and delivered over both inner and outer rings. The problem is to find the optimal bandwidth assignment that maximizes the network throughput while at the same time providing some notations of fairness. The solution involves determining the maximum bandwidth for each flow as well as determining which ring to take under the fairness constraints. This is therefore a joint *flow-routing-fairness* control problem.

In Section II, we first formulate the optimal bandwidth assignment problem and discuss the tradeoff of maximizing network throughput and fairness provisioning. We then decompose the problem into the iteration of two tasks, one for fair bandwidth allocation and the other for balanced routing, and present the formulation for these two tasks. In Section III, the details of the iterative algorithm for throughput maximization are described. A simple example is provided in Section IV to illustrate the bandwidth assignment procedure. Performance evaluation is presented in Section V and finally we conclude this paper in Section VI.

II. Optimal Bandwidth Assignment Problem

A. Network Model and Notations

Consider a N-node ring network with two counter-rotating rings as shown in Fig. 1. Let f_{ij} denote the flow from an ingress

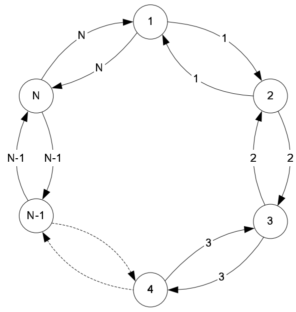


Fig. 1. Packet ring network model

node i to an egress node j. We will use the following notations throughout this paper.

- $\mathbf{C} = [\mathbf{C_1}, \mathbf{C_2}, \cdots, \mathbf{C_N}]$: inner ring link capacity.
- $\mathbf{D} = [\mathbf{D_1}, \mathbf{D_2}, \cdots, \mathbf{D_N}]$: outer ring link capacity.
- $\mathbf{R} = [\mathbf{r_{ij}}]_{\mathbf{N \times N}}$: the bandwidth demand matrix where r_{ij} is the bandwidth demand of flow f_{ij}.
- $\mathbf{X} = [\mathbf{x_{ij}}]_{\mathbf{N \times N}}$: inner ring bandwidth assignment matrix where x_{ij} is the inner ring bandwidth assigned to flow f_{ij}.
- $\mathbf{Y} = [\mathbf{y_{ij}}]_{\mathbf{N \times N}}$: outer ring bandwidth assignment matrix where y_{ij} is the outer ring bandwidth assigned to flow f_{ij}.

B. Problem Statement

The network throughput S is the total assigned bandwidth to all flows in both inner and outer rings, which is given by

$$S = \sum_{i,j \in N, i \neq j} f_{ij} \tag{1}$$

where $f_{ij} = x_{ij} + y_{ij}$ is the total flow rate between node i and j over both rings. The conventional wisdom is to allocate the bandwidth to all flows so that the network throughput *S* is maximized. However, it is well known that by maximizing *S*, some flows may be starved if fairness constraints are not imposed[10], [11]. In this paper, we want to find a bandwidth assignment scheme that maximizes the throughput while at the same time satisfying the fairness requirement for a given bandwidth demand.

The bandwidth demand can be simply specified through external service agreement or online measurement between each ingress-egress node pair, but fairness constraints need to be carefully defined. There are various fairness criterion proposed and debated in the literature[12], [13]. In our study, we assume all flows have the same right to access the bandwidth of the ring, so we choose to use the *max-min* fairness criterion[10]. This fairness criterion protects the lower demand flows from the domination of high-demand flows. Specially, it tries to maximize the minimum flow bandwidth allocation among all flows subject to link capacity constraints and the demand constraints.

Since the objective of throughput maximization and fairness cannot be achieved simultaneously, we choose to decompose the problem into two tasks. The first task is to find the *max-min* Fair Bandwidth Assignment(FBA) under the *max-min* fairness constraint. This assignment, however, may lead to unbalanced traffic load distribution across the network. So the second task is Balanced Routing(BR), which is to reduce the loading of the bottleneck links by optimally distributing the traffic across the inner and outer rings. After that, more bandwidth can be allocated to flows with demands not yet satisfied. This additional allocation, however, also needs to be fair and can be found by performing the FBA task again on the *residual ring* with the residual bandwidth demand as the input. The *residual ring* is defined as the ring with the residual link capacity(i.e., leftover capacity after the last bandwidth assignment). These two tasks will be repeated until no further improvement is possible. Finally, the total bandwidth of each flow is the sum of assigned bandwidth in all iterations, and the total network throughput is simply the sum of the assigned bandwidth all flows.

C. Fair Bandwidth Assignment (FBA)

Given the bandwidth demand traffic matrix $\mathbf{R}$, the FBA task is to find the bandwidth assignment matrices $\mathbf{X}$ and $\mathbf{Y}$ so that the minimum assigned bandwidth is maximized. More formally, we can formulate it as a $\max - \min$ optimization problem as follows

$$\text{maximize} \quad \min_{i \neq j,\, \forall i,j \in N} \{x_{ij} + y_{ij}\} \tag{2}$$

$$\text{w. r. t.} \quad X = [x_{ij}] \text{ and } Y = [y_{ij}] \tag{3}$$

$$\text{subject to} \quad U_k(\mathbf{X}) = \sum_{\mathbf{i \neq k}} \sum_{\mathbf{j=k}}^{\mathbf{i+1}} \mathbf{x_{ij}} \leq \mathbf{C_k}, \forall \mathbf{k} \in \mathbf{N}, \tag{4}$$

$$V_k(\mathbf{Y}) = \sum_{\mathbf{i \neq k+1}} \sum_{\mathbf{j=k+1}}^{\mathbf{i-1}} \mathbf{y_{ij}} \leq \mathbf{D_k}, \forall \mathbf{k} \in \mathbf{N}, \tag{5}$$

$$0 \leq x_{ij} + y_{ij} \leq r_{ij}, \quad i \neq j,\ \forall i,j \in N, \tag{6}$$

$$x_{ij} \geq 0 \text{ and } y_{ij} \geq 0, \quad i \neq j,\ \forall i,j \in N. \tag{7}$$

The objective function (2) represents the max-min fair bandwidth assignment for all flows. Constraints (4) and (5) are the the capacity constraints on the inner ring and the outer ring respectively. Constraint (6) stipulates that the allocated bandwidth will not exceed the demand, and (7) states that any bandwidth assignment should be nonnegative values.

D. Balanced Routing (BR)

Given the fair bandwidth assignment by FBA, the balanced routing task is to distribute the traffic of each flow across the inner and outer rings so that the traffic loading on the ring network is as balanced as possible. Let a_{ij} denote the fraction of the traffic of flow f_{ij} that goes through the inner ring. The problem is then to find $\mathbf{A} = [\mathbf{a_{ij}}]$ that minimizes the maximum link loading. That is,

$$\text{minimize} \quad \max_{\forall i,j \in N} \{U_i(\mathbf{A}), \mathbf{V_j}(\mathbf{A})\}$$

$$\text{w. r. t.} \quad \mathbf{A} = [\mathbf{a_{ij}}]$$

subject to

$$U_k(\mathbf{A}) = \sum_{i \neq k} \sum_{j=k}^{i+1} a_{ij}(x_{ij} + y_{ij}) \leq C_k, \forall k \in N, \tag{8}$$

$$V_k(\mathbf{A}) = \sum_{i \neq k+1} \sum_{j=k+1}^{i-1} (1 - a_{ij})(x_{ij} + y_{ij}) \leq D_k, \forall k \in N,$$

$$0 \leq a_{ij} \leq 1, \quad \forall i,j \in N \text{ and } i \neq j.$$

The capacity constraints are the same as the FBA problem.

III. Optimal Bandwidth Assignment Algorithm

The network throughput can be progressively maximized by iteratively performing the two tasks as discussed in the previous section. Specifically, the algorithm first finds the max-min bandwidth assignment based on current ring condition by performing the FBA task. It then performs the BR task to balance the traffic loading over inner and outer rings. As a result, some bandwidth can be released from previously saturated links. This process is repeated until either no further bandwidth assignment is feasible or the demands of all flows are satisfied. The detailed algorithm is as follows.

INPUT : link capacity $\mathbf{C}$ and $\mathbf{D}$, and bandwidth demand matrix $\mathbf{R}$.
OUTPUT : bandwidth assignment matrix $\mathbf{X}$ and $\mathbf{Y}$.
VARIABLE : residual link capacity $\mathbf{C_r}$ and $\mathbf{D_r}$, residual bandwidth demand matrix $\mathbf{R_r}$ and bandwidth splitting fraction matrix $\mathbf{A}$.

1) $k \leftarrow 1, \mathbf{X^k} \leftarrow \emptyset, \mathbf{Y^k} \leftarrow \emptyset, \mathbf{C_r^k} \leftarrow \mathbf{C}, \mathbf{D_r^k} \leftarrow \mathbf{D}, \mathbf{R_r^k} \leftarrow \mathbf{R}$;
2) Perform FBA task with $\mathbf{C_r^k}$, $\mathbf{D_r^k}$ and $\mathbf{R_r^k}$ as inputs, the outputs are bandwidth assignment $\mathbf{X^k}$ and $\mathbf{Y^k}$. If $\mathbf{X^k}$ and $\mathbf{Y^k}$ are zeros, go to step 6;
3) Perform BR task with $\mathbf{C_r^k}$, $\mathbf{D_r^k}$, $\mathbf{X^k}$ and $\mathbf{Y^k}$ as inputs, the output is $\mathbf{A^k}$;
4) With $\mathbf{X^k}$, $\mathbf{Y^k}$ and $\mathbf{A^k}$, do the following:
   (a) Compute the residual bandwidth demand : $\mathbf{R_r^{k+1}} \leftarrow \mathbf{R_r^k} - (\mathbf{X^k} + \mathbf{Y^k})$;
   (b) Compute the residual link capacity: $\mathbf{C_r^{k+1}} \leftarrow \mathbf{C_r^k} - \mathbf{U}(\mathbf{A^k})$ and $\mathbf{D_r^{k+1}} \leftarrow \mathbf{D_r^k} - \mathbf{V}(\mathbf{A^k})$, where $U(A^k)$ and $V(A^k)$ are given in (8);
   (c) Compute aggregated assigned bandwidth of each flow in inner and outer rings:

$$x_{ij} \leftarrow x_{ij} + a_{ij}^k (x_{ij}^k + y_{ij}^k)$$
$$y_{ij} \leftarrow y_{ij} + (1 - a_{ij}^k)(x_{ij}^k + y_{ij}^k)$$

5) If the residual bandwidth demand $\mathbf{R_r^{k+1}}$ is nonzero, let $k \leftarrow k+1$ and go to step 2.
6) return $\mathbf{X}$ and $\mathbf{Y}$ and stop;

Note that step 2 of the algorithm is for finding the fair bandwidth assignment based on the current network condition and the bandwidth demand. If no feasible assignment is found, the algorithm terminates, otherwise it goes to step 3 and performs the BR task to balance the network loading. Step 4 computes the residual link capacity and bandwidth demand. In step 5, if it finds that bandwidth demand is not satisfied, it returns to step 2 for the next iteration, otherwise the algorithm goes to step 6, which returns the overall bandwidth assignment results and the algorithm stops. The *FBA* and *BR* tasks can be solved efficiently using linear programming (LP) [14]. In particular, there is a simple algorithm for the *FBA* task with running time of $O(2N)$[10], where N is the number of nodes on the ring.

IV. Illustrative Example

We now illustrate the procedure of the optimal bandwidth assignment algorithm with a simple example. Consider a 4-node ring network. Let the bandwidth of all links be 100 units. Let there be four active flows with bandwidth demands of $r_{13} = 120, r_{14} = 30, r_{23} = 70$ and $r_{43} = 40$.

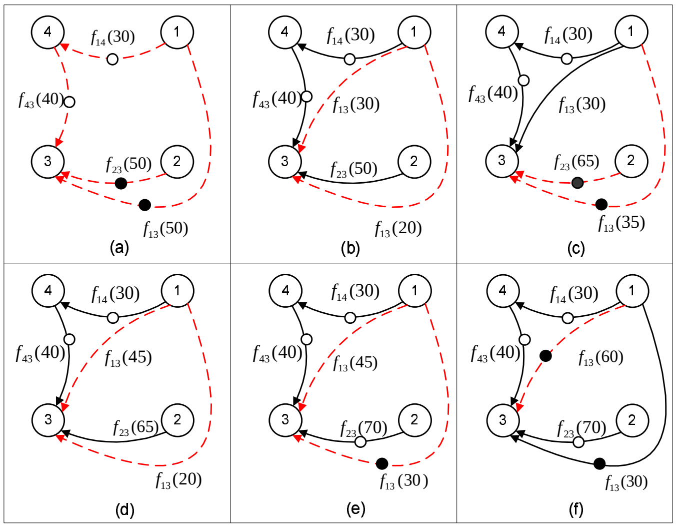


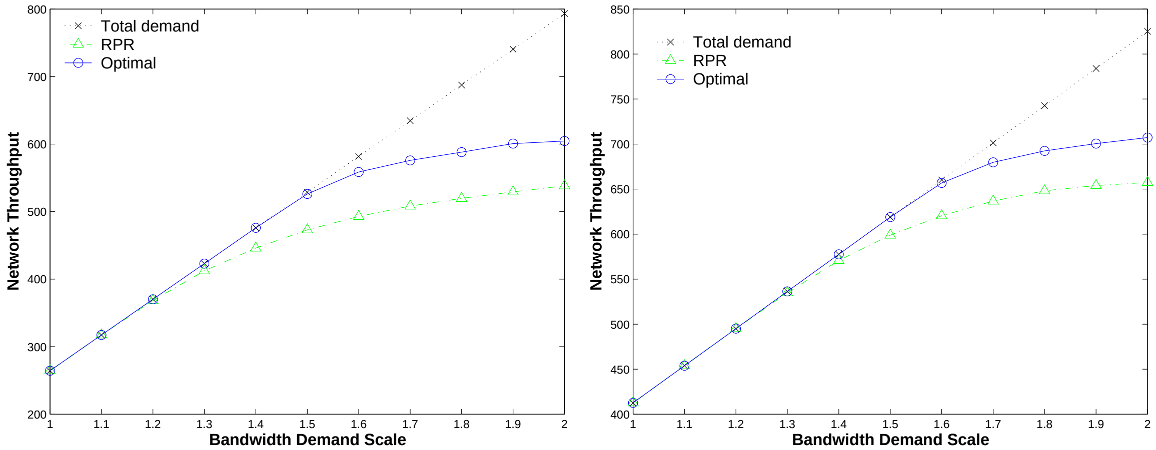
Fig. 2. Bandwidth assignment for a four-node ring with four active flows.

In the beginning, all flows are initialized with zero bandwidth. Then in step 2, Task 1 is performed and the max-min fair bandwidth assignment is found based on the shortest path routing. The results are shown in Fig. 2(a). Here, flows f_{14} and f_{43} are satisfied(indicated by white circle marks). Flows f_{13} and f_{23} need to pass link(2,3), so each is assigned 50 units of bandwidth. As link(2,3) is saturated after this assignment, we go to step 3. In this step, the traffic loading on link(2,3) is reduced by splitting the assignment of flow f_{13} to the inner and outer rings, The bandwidth of flow f_{13} is 30 units in the inner ring and 20 units in the outer ring as shown in Fig. 2(b)(indicated by dashed curves). The residual capacity of link(2,3) is $100 - 50 - 20 = 30$ units. Since flow f_{13} and f_{23} are not yet satisfied, the algorithm enters the second iteration. By repeating step 2, the residual bandwidth of link(2,3) is shared by these two flows and they each get 15 units of bandwidth. Fig. 2(c) shows the results after this step. In step 3, the splitting of bandwidth assignment across inner and outer rings is performed again. The bandwidth of flow f_{13} is 45 units in the inner ring and 20 units in the outer ring respectively. The results are shown in Fig. 2(d). After that, link(2,3) has 15 units of residual bandwidth.

In iteration 3, flow f_{23} is satisfied when 5 more units of bandwidth from link(2,3) are assigned to it. The remaining 10 units of bandwidth of link(2,3) are assigned to flow f_{13} as shown in Fig. 2(e). Since flow f_{13} is still not satisfied, so the remaining 15 units of bandwidth of link (4,3) are all assigned to it. The final results are shown in the Fig. 2(f). Here, flows f_{14}, f_{23} and f_{43} are satisfied, only f_{13} is not satisfied and gets 90 units of bandwidth out of a demand of 120. This, however, is the maximum achievable throughput for f_{13} under the max-min fairness constraints.

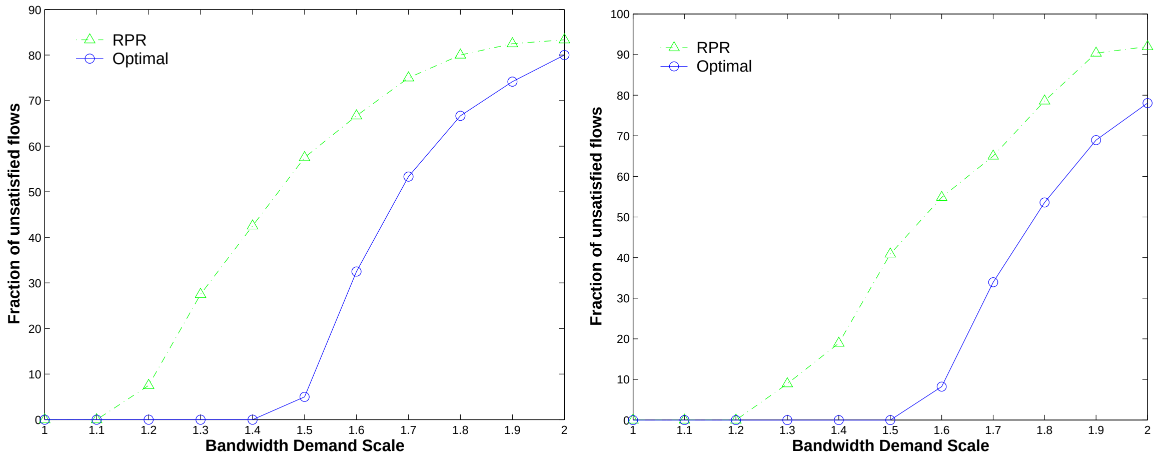
V. Performance Evaluation

In this section we compare the performance of the optimal bandwidth assignment scheme with the default shortest path routing scheme adopted by RPR with a 4-node and a 8-node ring



(a)4-node ring (b)8-node ring

Fig. 3. Comparison of Network Throughput



(a)4-node ring (b)8-node ring

Fig. 4. Comparison of Unsatisfied Flows

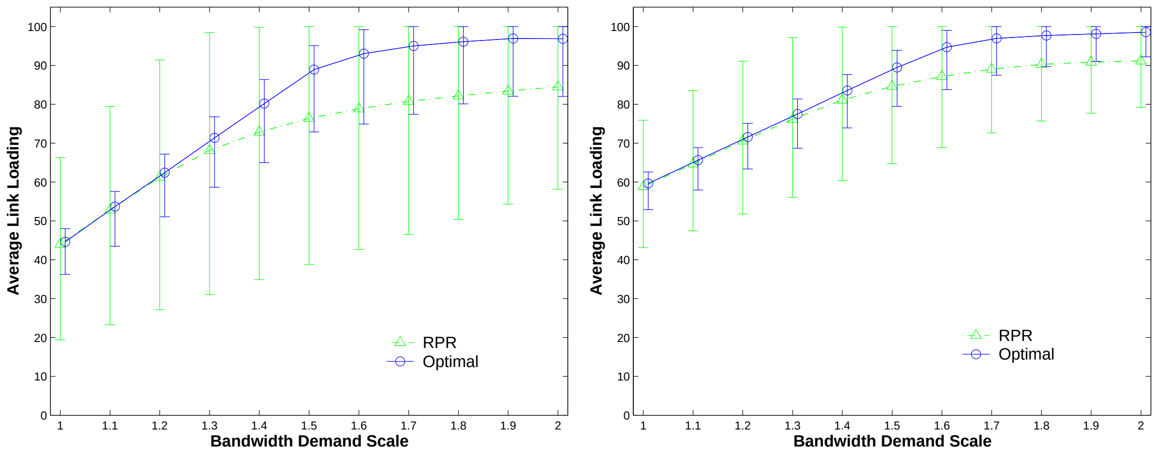
networks. In the experiments, all links have the same bandwidth of 100 units. Let R_0 be a base bandwidth demand matrix, and each entry of R_0 is a random number between 0 and 100 representing the bandwidth demand between a ingress-egress node pair. The actual demand matrix is given as $R = gR_0$, where g is a multiplier ranging between 1.0 and 2.0 to scale the traffic demand. Each data point presented in the plots is the average result obtained from ten experiments with different base demand matrices.

Fig. 3 compares the network throughput under different traffic loading for these two rings. When the offered traffic loading is low, both schemes can allocate the bandwidth to meet the demands of all flows. When the traffic demand is higher, some links are saturated. Since RPR uses the shortest path routing, there is no choice but to block the flows passing through these saturated links. Our scheme, on the other hand, can offer significantly higher throughput with the bifurcated routing over both inner and outer rings.

Fig. 4 shows the fraction of unsatisfied flows under different traffic loading conditions. As seen from this figure, our optimal bandwidth assignment scheme can greatly reduce the unsatisfied flows. For example, when $g = 1.5$, about 60% of flows are not satisfied using RPR scheme in 4-node ring network, while only 5% of flows are not satisfied using our scheme. In the 8-node ring network, the corresponding values are 45% and 0 by using RPR and our scheme respectively.

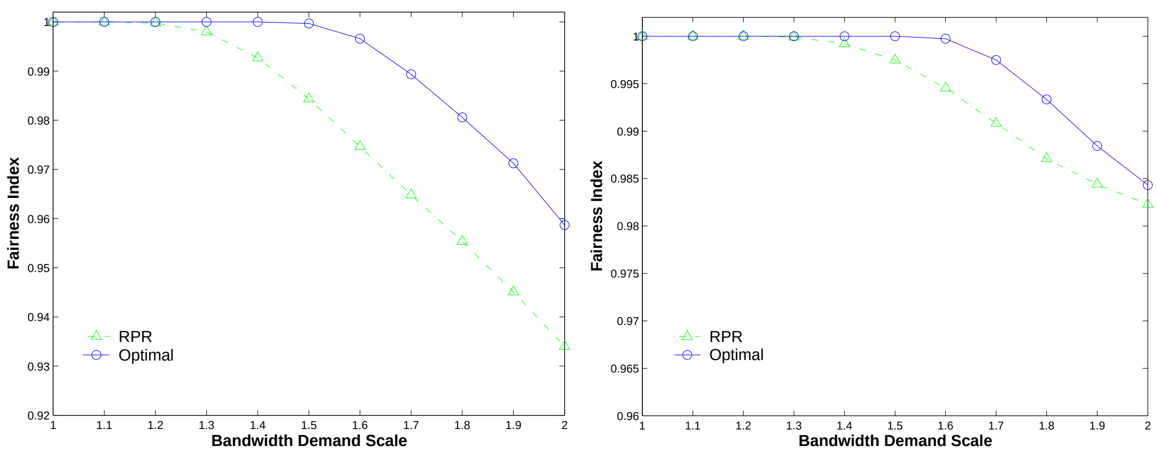
The average link loading under different offered traffic load values are illustrated in Fig. 5 for the two schemes. The vertical bars indicate the maximum and minimum values. We can see that using the optimal scheme, the range of the link loading is much smaller. This shows that the optimal scheme can distribute the offered traffic more uniformly across the rings than the shortest path routing. The optimal scheme, therefore, has better capability to accommodate short-term out-of-profile traffic.

Finally, we show the fairness index[15] of the bandwidth assignment for all flows in Fig. 6. We can see that our scheme can offer better fairness than RPR scheme under different traffic conditions.



(a)4-node ring (b)8-node ring

Fig. 5. Comparison of Link Traffic Load



(a)4-node ring (b)8-node ring

Fig. 6. Comparison of Fairness Index

VI. Conclusions

In this paper, we have presented an optimal bandwidth assignment scheme for generic packet rings. The objective is to maximize the network throughput while at the same time providing fairness guarantee. The proposed scheme consists of the task of finding fair bandwidth assignment and the task of finding balanced routing. By performing these two tasks iteratively, the network throughput is progressively maximized while fairness is guaranteed for the bandwidth allocation. We illustrate the bandwidth assignment procedures with a simple example. Numerical results are provided to show that the proposed scheme performs significantly better than the default RPR routing scheme in terms of network throughput, flow satisfaction, link loading distribution and fairness index.

REFERENCES

[1] White paper. an introduction to resilient packet ring technology, Aug. 2001. available at http://www.rpralliance.org.

[2] White paper. spatial reuse protocol technology, cisco systems, 2000. available at http://www.cisco.com/warp/public/cc/techno/wnty/dpty/tech/srpmc_wp.htm.

[3] J.Kao et al. Gandalf proposal for ieee standard 802.17, draft 0.4, Nov. 2001.

[4] A. Mekkittikul et al. Alladin proposal for ieee standard 802.17, draft 1.0, Nov. 2001.

[5] V. Gambiroza, Y. Liu, P. Yuan, and E. Knightly. High performance fair bandwidth allocation for resilient packet rings. In *Proc. ITC Specialist Seminar on Traffic Engineering and Traffic Management*, 2002.

[6] P. Yue, Z. Liu, and J. Liu. High performance fair bandwidth allocation algorithm for resilient packet ring. In *AINA'03*, Mar. 2003.

[7] C.-C. Huang, H. Peng, F.-J. Yuan, and J. Hawkins. A steady state bound for resilient packet rings. In *Globecom'03*, Dec. 2003.

[8] A.K. Parekh and R. G. Gallagher. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *IEEE/ACM Tran. on Networking*, 1, 1993.

[9] J. Chen, I. Cidon, and Y. Ofek. A local fairness algorithm for gigabit lan's/man's with spatial reuse. *IEEE JSAC*, 11(8), Oct. 1993.

[10] Dimitri P. Bertsekas. *Data Networks*. Englewood Cliffs, N.J. : Prentice Hall, 1992.

[11] H.-Y. Luo, S.-W. Lu, V. Bharghavan, J. Cheng, and G. Zhong. A packet scheduling approach to qos support in multihop wireless networks. *ACM Journal of Mobile Networks and Applications (MONET)*, 9(3), June 2004.

[12] L. Massoulie and J. Roberts. Bandwidth sharing: Objectives and algorithms. In *INFOCOM'99*, pages 1395–1403, 1999.

[13] Richard J. La and V. Anantharam. Utility-based rate control in the internet for elastic traffic. *IEEE/ACM Trans. on Networking*, 10(2):272–286, 2002.

[14] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2003.

[15] R. K. Jain, D.-M. Chiu, and W. Howe. A quantitative measure of fairness

and discrimination for resource allocation in shared systems. In *DEC Res. Rep. TR-301*, Sep. 1984.

Cunqing Hua received the B.S degree in Electronics Engineering from University of Science and Technology of China in 2000, MPhil. and Ph.D. degree in Information Engineering from The Chinese University of Hong Kong in 2002 and 2006 respectively. He is now a Postdoctoral Research Fellow at the Department of Computer Science, University of Houston. His current research interests include protocol and algorithm design, analysis and optimization, and performance evaluation for computer networks, wireless mesh networks and sensor networks.

Tak-Shing Peter Yum (SM'86) was born in Shanghai. He received the B.S., M.S., and Ph.D. degree from Columbia University, New York, NY in 1974, 1975 and 1978 respectively. He joined Bell Telephone Laboratories in April 1978 working on switching and signaling systems. In 1980, he accepted a teaching appointment at the National Chiao Tung University, Taiwan. Then, in 1982, he joined the Chinese University of Hong Kong, where he is now Dean of Engineering and Professor of Information Engineering. He has published widely in Internet research with contributions to routing, buffer management, deadlock handling, message resequencing and multiaccess protocols. He then branched out to work on cellular network, lightwave networks, video distribution networks and 3G networks. His recently interest is in RFID and sensor network technologies and applications.

Cheng Li received the B.Eng degree in Electronics Engineering from Harbin Institute of Technology, Harbin, China, and M.Eng. degree in Information Engineering from The Chinese University of Hong Kong, Hong Kong, China, in 2000 and 2003, respectively. Currently he works for NXP Semiconductors.