

Best Effort Flow Control in Network-on-Chip

Mohammad S. Talebi¹, Fahimeh Jafari^{1,2}, Ahmad Khonsari^{3,1},
and Mohammad H. Yaghmaee²

¹ IPM, School of Computer Science, Tehran, Iran

² Ferdowsi University of Mashhad, Mashahhad, Iran

³ ECE Department, University of Tehran, Tehran, Iran

mstalebi@ipm.ir, jafari@ipm.ir, ak@ipm.ir, hyaghmae@ferdowsi.um.ac.ir.

Abstract. Research community has recently witnessed the emergence of Interconnect networks methodology based on Network-on-Chip (NoC). Employing end-to-end congestion control is becoming more imminent in the design process of NoCs. This paper presents a centralized congestion scheme in the presence of both elastic and streaming flow traffic mixture. In this paper, we model the desired BE source rates as the solution to a utility maximization problem while preserving GS traffics services requirements at the desired level. We propose an iterative algorithm as the solution to the maximization problem which has the benefit of low complexity and fast convergence.

Keywords: Flow Control, Utility-based optimization, Gradient Method

1 Introduction

Quality-of-Service (QoS) provisioning in NoC's environment has attracted many researchers and currently it is the focus of many literatures in NoC research community. NoCs are expected to serve as multimedia servers and are required not only to carry Elastic Flows, i.e. Best Effort (BE) traffic, but also Inelastic Flows, i.e. Guaranteed Service (GS) traffic which requires tight performance constraints such as necessary bandwidth and maximum delay boundaries [1].

Flow control is still a novel problem in NoCs and to the best of our knowledge only few works has been carried out in this field. So far, several works have focused on this issue for NoC systems. In [2], a prediction-based flow-control strategy for on-chip networks is proposed in which each router predicts the buffer occupancy to sense congestion. This scheme controls the packet injection rate and regulates the number of packets in the network. In [3] link utilization is used as a congestion measure and a Model Prediction-Based Controller (MPC), determines the source rates. Dyad [4] controls the congestion by using adaptive routing when the NoC faces congestion.

In this paper, we focus on the flow control for Best Effort traffic as the solution to a utility-based optimization problem. To the best of our knowledge, none of the aforementioned works have dealt with the flow control problem through utility optimization approach.

This paper is organized as follows. In Section 2 we present the system model and formulate the underlying optimization problem for BE flow control. In section 3 we

propose the optimal flow control algorithm. Section 4 presents the simulation results. Finally, section 5 concludes the paper and states some future work directions.

2 System Model and Flow Control Problem

We model the flow control in NoC as the solution to an optimization problem. As in [5], we consider NoC as a network with a set of bidirectional links L and a set of sources S . A source consists of Processing Elements (PEs), routers and Input/Output ports. Each link $l \in L$ is a set of wires, busses and channels that are responsible for connecting different parts of the NoC and has a fixed capacity of c_l packets/sec. We denote the set of sources that share link l by $S(l)$.

We denote the set of sources with BE and GS traffic by S_{BE} and S_{GS} , respectively. Each link l is shared between the two aforementioned traffics. GS sources will obtain the required amount of the capacity of links and BE sources benefit from the remainder. Our objective is to choose source rates with Best Effort traffic so that to maximize the sum of the logarithm of the BE source rates as following [5]:

$$\max_{x_s} \sum_{s \in S_{BE}} \log x_s \quad (1)$$

subject to:

$$\sum_{s \in S_{BE}(l)} x_s + \sum_{s \in S_{GS}(l)} x_s \leq c_l \quad \forall l \in L \quad (2)$$

$$x_s > 0 \quad \forall s \in S_{BE} \quad (3)$$

Problem (1) is a convex optimization problem and therefore admits a unique maximizer [6]. Due to coupled nature of problem (1), it should be solved using centralized methods like Interior Point method which poses great computational overhead onto the system [6] and hence is of little interest. One way to reduce the computational complexity is to transform the constrained optimization problem into its *Dual*, which can be defined to be unconstrained. According to the Duality Theory [6], each convex optimization problem has a dual, whose optimal solution, leads to the optimal solution of the main problem.

3 Optimal Flow Control Algorithm

In order to solve problem (1), we have obtained its Dual function through which the Dual problem can be achieved. As the Dual problem is unconstrained, it can be solved using simple iterative methods such as Gradient Projection Method [6]. We have solved the Dual problem using Projected Gradient Method, in an iterative manner.

Due to lack of enough space we have omitted mathematical derivations. We direct serious readers to the extended version of this work [7]. The iterative solution of the Dual problem can be used as an iterative flow control algorithm for BE traffic, which is listed below as algorithm 1.

Algorithm 1: Flow Control for BE Traffics in NoC

Initialization:

1. Initialize c_l of all links.
2. Set link shadow price vector to zero.
3. Set the ε as the stopping criteria.

Loop:

Do until $(\max |x_s(t+1) - x_s(t)| < \varepsilon)$

1. $\forall l \in L$: Compute new link prices:

$$\lambda_l(t+1) = \left[\lambda_l(t) - \gamma \left(c_l(t) - \sum_{s \in \mathcal{S}_{CS}(l)} x_s(t) - \sum_{s \in \mathcal{S}_{BE}(l)} x_s(\lambda(t)) \right) \right]^+$$

2. Compute new BE source rates as follows

$$x_s(t+1) = \frac{a_s}{\sum_{l \in L(s)} \lambda_l(t+1)}$$

Output:

Communicate BE source rates to the corresponding nodes.

4 Simulation Results

In this section we examine the proposed flow control algorithm, listed above as Algorithm 1, for a typical NoC architecture. In our scenario, we have used a NoC with 4×4 Mesh topology which consists of 16 nodes communicating using 24 shared bidirectional links; each one has a fixed capacity of 1 Gbps. In our scheme, packets traverse the network on a shortest path using a deadlock free XY routing.

As a performance metric, the relative error with respect to optimal source rates for different step sizes is shown in Fig. 1. The relative error is averaged over all active sources. This figure reveals that for $\gamma = 1.05$ the algorithm leads to less than 10% error in average just after about 13 iteration steps, and after 20 steps the average error lies below 5%. However, for a smaller step size, e.g. $\gamma = 0.2$, these thresholds are about 60 and 75 steps, respectively.

5 Conclusion and Future Works

In this paper we addressed the problem of flow control for BE traffic in NoC systems. Flow control was considered as the solution to a utility maximization

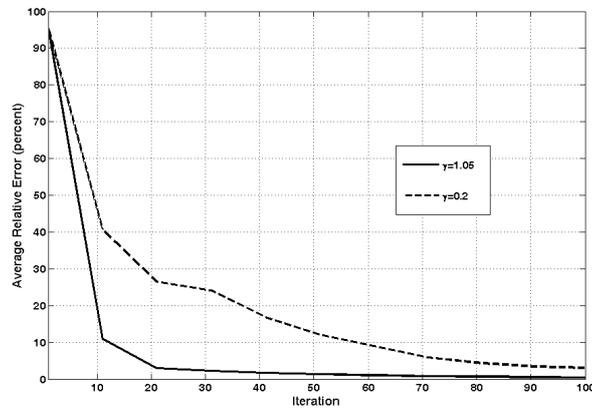


Fig. 1. Average Relative Error for $\gamma = 1.05$ and $\gamma = 0.2$

problem which was solved indirectly through its dual using gradient projection method. This was led to an iterative algorithm that can be used to determine optimal BE source rates. The algorithm can be implemented by a controller which poses a light communication and communication overhead to the system. Further investigation about the effect of delay incurred by the proposed algorithm and convergence behavior are the main directions of our future studies.

References

1. Benini, L., DeMicheli, G.: Networks on Chips: A New SoC Paradigm. In: Computer, vol. 35, no. 1, pp. 70-78. (2002)
2. Ogras, U. Y., Marculescu, R.: Prediction-based flow control for network-on-chip traffic. In: Proceedings of the Design Automation Conference, (2006)
3. Van den Brand, J. W., Ciordas, C., Goossens, J. W., Basten, T.: Congestion-Controlled Best-Effort Communication for Networks-on-Chip. In: Proceedings of Design, Automation and Test in Europe Conference, pp. 948-953. (2007)
4. Jingcao, H., Marculescu, R.: DyAD - smart routing for networks-on-chip. In: Design Automation Conference, pp. 260- 263. (2004)
5. Low, S. H., Lapsley, D. E.: Optimization Flow Control, I: Basic Algorithm and Convergence. In: IEEE/ACM Transactions on Networking, vol. 7, no. 6, pp. 861-874. (1999)
6. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, (2004)
7. Talebi, M. S., Jafari, F., Khonsari, A., Yaghmaee, M. H.: Weighted Proportionally Fair Flow Control for Best Effort Traffic in Network-On-Chip. In: To appear in proceedings of IPDPS, (2007)