An Analytical Model of Adaptive Wormhole Routing with Deadlock Recovery

Mohamed Ould-Khaoua and Ahmad Khonsari

Computer Science Department, Strathclyde University, Glasgow, U.K.
{mohamed, ak}@cs.strath.ac.uk

Abstract: This paper proposes a new analytical model to predict the mean message latency in k -ary n -cubes with Compressionless routing, a recovery-based fully-adaptive routing proposed by Kim *et al* [2].

1 Introduction

Deadlock recovery as a viable alternative to deadlock avoidance has recently gained consideration in the scientific community. It has been shown that deadlocks are quite rare except when the network is close to saturation [2]. Thus the hardware dedicated for deadlock avoidance is not necessary most of the time. This consideration has motivated the authors in [2] to introduce Compressionless Routing as a framework for developing recovery-based fully-adaptive routing algorithms. This paper proposes the first analytical model of compressionless routing in k -ary n -cubes.

2 The Analytical Model

Compressionless routing and the router structure are described in detail in [2]. The model is based on the following assumptions [3].

- i) Message destinations are uniformly distributed. Nodes generate traffic independently of each other, following a Poisson process with a mean rate of λ_g messages/cycle. The message length is m flits, where m is a random variable with a mean $\bar{m}$. A flit requires one-cycle transmission time across a physical channel.
- ii) The timeout period is τ cycles. The probability of timeout at a channel is independent of the subsequent channels. When a transmission failure occurs due to timeout, a message is retransmitted after a time gap of V cycles, where V may be any random variable with a mean $\bar{V}$. Successive message re-transmissions are independent of each other.
- iii) L ($L \geq 1$) virtual channels are used per physical channel.

The mean message latency is composed of the mean network latency, $\bar{t}_r$, and the mean waiting time, $\bar{w}_s$, in the source node. However, to capture the effects of virtual channels multiplexing, the mean message latency has to be scaled by a factor, $\bar{l}$, representing the average degree of virtual channels multiplexing. Therefore, we can write [3]

$$Latency = (\bar{t}_r + \bar{w}_s)\bar{l} \quad (1)$$

The average number of channels that a message crosses along each of the n dimensions and across the network, $\bar{k}$ and $\bar{d}$ respectively, are given by [3]

$$\bar{k} = (k-1)/2 \quad (2)$$

$$\bar{d} = n\bar{k} \quad (3)$$

Consider a message that has been transmitted successfully through the network. Since a message crosses, on average, $\bar{d}$ channels, the distribution of the network latency in the case of a successful transmission can be written as

$$T_S(x) = \text{Prob}(m + \bar{d} + d \cdot w \leq x) \quad (4)$$

where w denoting the waiting time at a channel. Since the waiting times at two successive channels are independent of each other and since the Laplace-Stieltjes Transform (LST) of the sum of two independent random variables is equal to the product of their LSTs [1], the LST of $T_S(x)$ is given by

$$T_S^*(s) = \int_0^\infty e^{-sx} dT_S(x) = M^*(s)e^{-s\bar{d}}W^*(s)^d \quad (5)$$

Consider a message that experiences a timeout at the i -th hop channel. The LST of the network latency in the case of a transmission failure due to a timeout at the i -th hop channel can be written as

$$T_{F_i}^*(s) = e^{-s(2(i-1)+\tau)}W^*(s)^{(i-1)} \quad (6)$$

Let ψ_i be the number of channels that a message can select at its i -th hop, and, P_i be the probability that message experiences a timeout at channel. If P_l denotes the probability that l virtual channels at a physical channel are busy, the probability, P_{F_i} , that a message suffers a timeout at the i -th hop channel can be expressed as

$$P_{F_i} = P_t^{\psi_i} P_L^{\psi_i} \prod_{j=1..i-1} (1 - P_t^{\psi_j} P_L^{\psi_j}) \quad (7)$$

Given that a transmission failure can be caused by a timeout at any of the $\bar{d}$ channels along the message path, the LST of the network latency in case of a transmission failure is given by

$$T_F^*(s) = \sum_{i=1..\bar{d}} P_{F_i} T_{F_i}^*(s) \quad (8)$$

The probability of a successful transmission, P_S , and that of transmission failure, P_F , are simply given by

$$P_S = \prod_{i=1..\bar{d}} (1 - P_t^{\psi_i} P_L^{\psi_i}) \quad (9)$$

$$P_F = 1 - P_S = 1 - \prod_{i=1..\bar{d}} (1 - P_t^{\psi_i} P_L^{\psi_i}) \quad (10)$$

The network latency of a single transmission try, taking into the network latency in case of a transmission success and transmission failure, can be written as

$$T^*(s) = P_S T_S^*(s) + \sum_{i=1..\bar{d}} P_{F_i} T_{F_i}^*(s) \quad (11)$$

Let t_r be a random variable denoting the network latency seen by a message after r re-transmission attempts. Since a new re-transmission is delayed for a time gap of v cycles and it is independent of the previous re-transmissions, we can write

$$T_r^*(s) = \sum_{r=1.. \infty} P_S P_F^{r-1} T^*(s)^r V^*(s)^{r-1} = P_S T^*(s) / (1 - P_F T^*(s) V^*(s)) \quad (12)$$

The number of channels that a message can use at its i -th hop is found to be [3]

$$\psi_i = \sum_{l=0..j} l P_i^j \quad j\bar{k} + 1 \leq i \leq (j+1)\bar{k} \quad (0 \leq j \leq n-1) \quad (13)$$

$$P_i^j = \binom{n}{j} N_0^{\bar{k}-1} (i-1-j\bar{k}, n-j) / \sum_{l=0..j} \binom{n}{l} N_0^{\bar{k}-1} (i-1-l\bar{k}, n-l) \quad (14)$$

$$N_p^{p+q-1}(r, m) = \sum_{l=0..m} (-1)^l \binom{m}{l} \binom{r-mp-lq+m-1}{m-1} \quad (15)$$

Modelling a channel as an M/M/1 queue with impatient customers and with deterministic impatient time yields a simple and practical model that exhibits a reasonable degree of accuracy. The rate of traffic on each channel is given by

$$\lambda_c = \sum_{i=0..d-1} \prod_{j=1..i} (1 - P_t^{\psi_j} P_L^{\psi_j}) \lambda_0, \text{ where } \lambda_0 = \lambda_g / nP_S \quad (16)$$

Following the suggestions proposed of [4], the mean waiting time and probability of timeout at a given channel can be approximated as

$$\bar{w} = \frac{\frac{\lambda_c \bar{t}^2}{(1 - \lambda_c \bar{t})} - \left(\frac{\bar{t}}{1 - \lambda_c \bar{t}} + \lambda_c \bar{t} \tau \right) e^{\frac{-(1 - \lambda_c \bar{t}) \tau}{\bar{t}}}}{(1 - \lambda_c \bar{t})} - P_t \tau}{\left(1 - \lambda_c^2 \bar{t}^2 e^{\frac{-(1 - \lambda_c \bar{t}) \tau}{\bar{t}}} \right) (1 - P_t)} \quad (17)$$

$$P_t = \left[(1 - \lambda_c \bar{t}) \lambda_c \bar{t} e^{\frac{-(1 - \lambda_c \bar{t}) \tau}{\bar{t}}} \right] / \left[1 - \lambda_c^2 \bar{t}^2 e^{\frac{-(1 - \lambda_c \bar{t}) \tau}{\bar{t}}} \right] \quad (18)$$

Modelling the local queue in the source as an M/M/1 queue with a mean arrival rate λ_g / L and a mean service time $\bar{t}_r$ yields the mean waiting time as [1]

$$\bar{w}_s = (\lambda_g / L) \bar{t}_r^2 / (1 - (\lambda_g / L) \bar{t}_r) \quad (19)$$

The probability, P_l , that l virtual channels at a given physical channel are busy is determined using a Markovian model (see [3] for a more detailed discussion). In the steady state, the model yields the following probabilities.

$$q_l = \begin{cases} 1 & l = 0 \\ q_{l-1} \lambda_c \bar{t} & 0 < l < L \\ q_{l-1} \lambda_c / (1 / \bar{t} - \lambda_c) & l = L \end{cases} \quad (20)$$

$$P_l = \begin{cases} 1 / \sum_{l=0..L} q_l & l = 0 \\ P_{l-1} \lambda_c \bar{t} & 0 < l < L \\ P_{l-1} \lambda_c / (1 / \bar{t} - \lambda_c) & l = L \end{cases} \quad (21)$$

The average degree of virtual channels multiplexing is given by [3]

$$\bar{l} = \sum_{i=1..L} i^2 P_i / \sum_{i=1..L} i P_i \quad (22)$$

There are inter-dependencies between the different variables of the model, and these are solved iteratively [1]. Fig. 1 depicts latency results predicted by the model and simulation in the 2-D torus (with $N=64$ nodes; $M=32$ and 64 flits; $L=1$ and 2 virtual channels; $\tau=32$ and 64 cycles; $V=32$ and 64 cycles). The model predicts latency with a good accuracy under light and moderate traffic. However, its accuracy degrades near the saturation point due to the approximations that were used to develop the model, e.g. assumption (ii).

Fig. 1: Latency from the model and simulation in the 2-D torus. a) $L=1$, b) $L=2$.

3 Conclusion

This paper has presented a new analytical model for fully-adaptive routing with deadlock-recovery, based on the Compressionless routing framework [2], to predict message latency in wormhole-routed k -ary n -cubes. Simulation experiments have revealed that the analytical model predicts latency with a good degree of accuracy.

References

- [1] L. Kleinrock, *Queueing Systems: Theory*, vol. 1, J. Wiley, New York, 1975.
- [2] J. Kim, A. Chien and Z. Liu, Compressionless routing: A framework for adaptive and fault-tolerant routing, *IEEE TPDS* 8(3), 1997, pp. 229- 244.
- [3] M. Ould-Khaoua, An analytical model of Duato's fully-adaptive routing in k -ary n -cubes, *IEEE Trans. Computers*, 44(12), 1999, pp 1-8.
- [4] H.C. Tijms, *Stochastic modelling and analysis: A computational approach*, J. Wiley, 1986.