

Analysis of a tandem network model of a single-router Network-on-Chip

Paul Beekhuizen · Dee Denteneer · Ivo Adan

Published online: 20 March 2008
© The Author(s) 2008

Abstract We study a single-router Network-on-Chip modelled as a tandem queueing network. The first node is a $geo^K/D/1$ queue (K fixed) representing a network interface, and the second node is a $.G/1$ queue representing the packet switch. If $K > 1$ we have train arrivals at the second node. If $K = 1$ the arrival process of the second node reduces to a Bernoulli process. In the latter case, routers have been studied extensively as part of ATM and LAN networks under the assumption that the number of input ports N tends to infinity. In Networks-on-Chips N is usually 4 or 5 and results for ATM and LAN routers lead to inaccurate results. We introduce a new approximation scheme that yields accurate results for small switches. In addition to this we analyse the tandem network, both for $K = 1$ and $K > 1$, and we approximate the mean sojourn time in the switch and the mean end-to-end delay. If $N = 4$ our approximation has a relative error of only 4.5% if $K = 6$ and 1% if $K = 1$.

1 Introduction

Systems on Chips typically consist of on-chip modules such as processors, memories, etc. These modules are traditionally connected via single buses. Because the wires cannot be used by multiple modules simultaneously, communication difficulties arise as the number

P. Beekhuizen (✉)
EURANDOM, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven,
The Netherlands
e-mail: beekhuizen@eurandom.tue.nl

P. Beekhuizen
Digital Signal Processing Group, Philips Research, High Tech Campus 36, 5656 AE, Eindhoven,
The Netherlands

D. Denteneer
Connectivity Systems and Networks Group, Philips Research, Eindhoven, The Netherlands

I. Adan
Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven,
The Netherlands

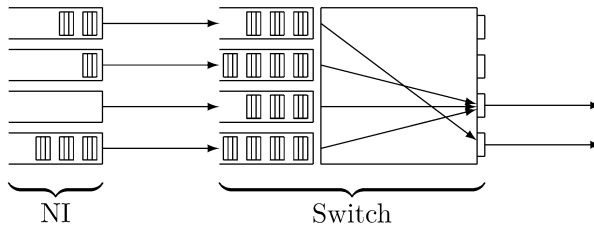


Fig. 1 The network model of this paper with $N = 4$ NIs. Packets arrive at the NIs according to a Bernoulli process with parameter p and are sent to the switch flit-by-flit. The switch sends the flits of the packets to the right output port using wormhole routing, with the restriction that each output port can only be used by one flit at a time. Contention is resolved randomly

of modules increases. This is one of the reasons why Networks on Chips (NoCs) are emerging as a paradigm for the connection of on-chip modules (González Pestana et al. 2004; Rădulescu et al. 2004; Rijpkema et al. 2003). In NoCs, routers are used to transmit packets to their destination, and as such they resemble other packet-switching networks. For a more elaborate introduction to NoCs, the interested reader is referred to Dally and Towles (2001).

In this paper we analyse the mean delay of a NoC consisting of N Network Interfaces (NI, see Rădulescu et al. 2004) and an $N \times N$ router, where each NI is connected to a unique input port of the router (see Fig. 1). In NoCs, NIs act as bridges between on-chip modules such as processors, memory modules, etc. and the routers of the NoC. We model one NI and an input queue of the router together as a discrete-time tandem queue with two different nodes: The first node represents the NI and the second node the router.

We assume that packets of fixed size K arrive at the NIs according to i.i.d. Bernoulli processes with parameter p . The size of packets is measured in flits, where a flit is precisely the amount of data that can be transmitted in one time slot. The arriving packets are stored in the NIs and the flits are then sent to the switch one-by-one, so the NI can be seen as a $geo^K/D/1$ queue. In particular we stress that the output process of this queue (and thus the input process of the switch) is an on-off process, whose on-period is equal to the busy period of the $geo^K/D/1$ queue. Arrival processes of this form are sometimes called ‘train arrivals’ (Xiong and Bruneel 1993).

The input ports of the switch are equipped with FIFO-queues to store incoming packets, and we assume that these buffers are infinite. We say that the first position of the queue is the HOL-position (*Head-Of-Line* position). Because of the FIFO policy only flits in the HOL-positions of the queues can be sent to an output port. The probability that a packet arriving at input port i has destination j is $1/N$ for all i and j . Note that each flit in a packet has the same destination.

The switch of our model uses wormhole routing: If the first flit of a packet (the header) is transmitted through a certain output port, that output port remains reserved until all flits of the packet have been transmitted. Because each output port can only be used once per time slot, contention occurs if there are multiple headers in HOL-positions with the same destination. In this case, the switch selects one of the contending headers at random, each with the same probability.

Because the switch transmits flits, it can be seen as a server. We therefore define the service time of a flit as the time spent in the HOL-position. By doing so, we include the amount of contention in the service time distribution. We stress that the service time of a header is determined by the number of consecutive output conflicts plus its transmission time of 1 slot, whereas the service time of a non-header flit is always equal to its transmission

time of 1 slot. The service time is thus always at least equal to 1 slot. For example, if a flit arrives at the HOL-position at the beginning of time slot $(t, t + 1]$ and leaves the switch at the end of time slot $(u, u + 1]$, we say that its service time is $u + 1 - t$. Note that we assume that flits arrive at the beginning of time slots, and leave at the end of them (the so-called early-arrival model, see e.g. Gravey and Hébuterne 1992; Takagi 1993).

If the packet size $K = 1$, an interesting special case of our model arises; we no longer have wormhole routing and the arrival process at the switch simplifies to a Bernoulli process. Because of this, a distinction is made between the situations $K = 1$ and $K > 1$ throughout this paper.

If $K = 1$ we have an ordinary packet switch with Bernoulli arrivals. Such packet switches have been studied extensively because of their presence in ATM and LAN networks (see e.g. Eng et al. 1989; Hluchyj and Karol 1988; Karol et al. 1987). In these studies, it is common to assume that $N \rightarrow \infty$, where N is the number of input and output ports of the switch. This assumption is justified by the size of switches in LAN and ATM networks. For instance, in Eng et al. (1989), the authors consider an ATM switch with $N > 1000$.

Two studies that deal with the case $K = 1$ are of particular importance for our paper, that of Karol et al. (1987) and that of Kim et al. (2000). In Karol et al. (1987), the service time, delay and throughput of a homogeneous switch is analysed, again under the assumption that $N \rightarrow \infty$, and the famous result that the asymptotic throughput is $2 - \sqrt{2} \approx 0.586$ is established. Kim et al. (2000) also analyse the service time and delay of a homogeneous switch for $N \rightarrow \infty$, but they make an approximation assumption that the number of consecutive output conflicts follows a geometric distribution. These studies will be discussed in more detail in Sects. 2.1 and 2.2 respectively.

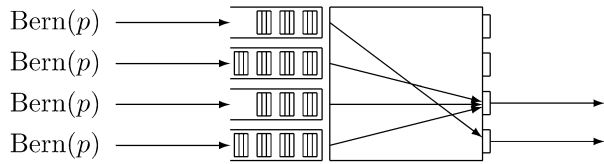
To summarise, the NoC can be modelled as a discrete-time tandem network with two different nodes representing the NI and the router. The first node is a $Geo^K/D/1$ queue with unit service times, and the second node is a discrete-time $.G/1$ with an unknown service time distribution. The results of the studies of the case $K = 1$ would in principle allow us to approximate the unknown service time distribution of the second node. However, in NoCs the size of routers is often $N = 4$ or $N = 5$, and we will see in Sect. 5 that these asymptotic approximations give significant errors. In order to obtain more accurate results, we also devise a new approximation of the service time distribution specifically aimed at small switches. Its results are compared favourably to the results of Karol et al. (1987) and Kim et al. (2000).

The purpose of our paper is thus twofold: first, we provide a new approximation of the service time distribution which is more accurate for small switches. Second, we analyse the tandem network model and we use our service time approximation to approximate the delay in this network.

The organisation of our paper is as follows: Sect. 2 is devoted to approximations for the case $K = 1$. In Sects. 2.1 and 2.2 we describe the approximations of Karol et al. and Kim et al. respectively. In Sect. 2.3 we introduce our new $Geo/Geo/1$ approximation. In Sect. 3 we extend this approximation to the case $K > 1$. In order to approximate the mean sojourn time in the switch if $K > 1$ we analyse the tandem network in Sect. 4.

The second part of our paper contains more experimental results. First, we analyse the performance of our approximation for $K = 1$ and $K > 1$ with respect to simulation outcomes in Sect. 5. For $K = 1$, we also compare our approximation to that of Karol et al. and Kim et al. Because we use a $Geo/Geo/1$ model for $K = 1$, we implicitly assume that the number of output conflicts can be approximated by a geometric distribution. This particular assumption is explored in more depth in Sect. 6. Finally, we present the conclusions of our research in Sect. 7.

Fig. 2 A switch in isolation with $N = 4$. Packets of size 1 arrive at the input queues of the switch according to i.i.d. Bernoulli arrival processes with parameter p



Throughout this paper, standard results on the $Geo^X/G/1$ queue will be used, see e.g. Takagi (1993). For the sake of completeness, we will state the two most important results for our paper here. We denote the waiting time by W , the service time by B , the sojourn time by S , and the size of a batch by X . Note that we assume that a batch of size X arrives *each* time slot (with $\mathbb{P}(X = 0) > 0$), so the interarrival times of non-zero batches are geometrically distributed with parameter $\mathbb{P}(X > 0)$. The mean waiting time of an arbitrary customer is given by

$$\mathbb{E}[W] = \frac{\rho}{1 - \rho} \left(\frac{\mathbb{E}[B^2]}{2\mathbb{E}[B]} - \frac{1}{2} \right) + \frac{1}{1 - \rho} \left(\frac{\mathbb{E}[X^2]}{2\mathbb{E}[X]} - \frac{1}{2} \right) \mathbb{E}[B], \tag{1.1}$$

where $\rho = \mathbb{E}[X]\mathbb{E}[B]$. The waiting time of the *first* customer of a batch (W_1) will also turn out to be important, and its mean is given by

$$\begin{aligned} \mathbb{E}[W_1] &= \mathbb{E}[W] - \left(\frac{\mathbb{E}[X^2]}{2\mathbb{E}[X]} - \frac{1}{2} \right) \mathbb{E}[B] \\ &= \frac{\rho}{1 - \rho} \left(\frac{\mathbb{E}[B^2]}{2\mathbb{E}[B]} - \frac{1}{2} \right) + \frac{\rho}{1 - \rho} \left(\frac{\mathbb{E}[X^2]}{2\mathbb{E}[X]} - \frac{1}{2} \right) \mathbb{E}[B]. \end{aligned} \tag{1.2}$$

The corresponding mean sojourn times can of course be found using $S = W + B$.

2 Approximations for $K = 1$

For now, we assume that the packet size $K = 1$. Most importantly, this entails that the NI effectively only serves as a mechanism that stalls each flit for exactly one time slot and as a result the output process of the NI is reduced to a Bernoulli process. We can therefore model the input queue of the switch directly as a $Geo/G/1$ queue, with an a priori unknown service time distribution.

This particular model of a packet switch has been studied before, by Karol et al. (1987) and Kim et al. (2000). Both of these studies are based on the limit of $N \rightarrow \infty$. The main difference is that in Karol et al. (1987) the first two moments of the service time are determined exactly, whereas in Kim et al. (2000) a geometric distribution is assumed and fitted to the mean determined in Karol et al. (1987). We will refer to the first model as the KHM model and to the second model as the KKL model. Both of these models are briefly surveyed in this section.

We also introduce a new approximation in which the service time is assumed to be geometrically distributed, as in Kim et al. (2000), but we use a quadratic approximation of the service rate based on light and heavy traffic limits. Because our model is not based on $N \rightarrow \infty$ we find a different mean service time and obtain better results for small switches.

2.1 The KHM model

The crucial observation of Karol et al. is that for $N \rightarrow \infty$ the number of packets with the same destination arriving at the *HOL-positions* follows a Poisson(p) distribution. Because the switch randomly selects flits, this selection process itself can be viewed as a queuing process with random order of service and unit service times. In other words, the service time of a flit is equal to the sojourn time in a $Geo^X/D/1$ queue with random order of service and batch arrivals governed by a Poisson distribution. Let \tilde{S} , \tilde{B} , \tilde{W} , and \tilde{X} denote the sojourn time, service time, waiting time, and batch size in this $Geo^X/D/1$ queue. We thus have

$$B \stackrel{D}{=} \tilde{S}, \tag{2.1}$$

where B is the service time in the switch and $\stackrel{D}{=}$ denotes equality in distribution.

Because the mean service and sojourn time do not depend on the order of service, it follows from (1.1) that

$$\mathbb{E}[B] = \mathbb{E}[\tilde{S}] = \frac{\tilde{\rho}}{1 - \tilde{\rho}} \left(\frac{\mathbb{E}[\tilde{B}^2]}{2\mathbb{E}[\tilde{B}]} - \frac{1}{2} \right) + \frac{\mathbb{E}[\tilde{X}(\tilde{X} - 1)]\mathbb{E}[\tilde{B}]^2}{2\mathbb{E}[\tilde{X}](1 - \tilde{\rho})} + \mathbb{E}[\tilde{B}], \tag{2.2}$$

where $\tilde{\rho} = \mathbb{E}[\tilde{X}]\mathbb{E}[\tilde{B}]$. In our case, $\tilde{X} \sim \text{Poisson}(p)$ and $\tilde{B} \sim \text{Det}(1)$, which implies

$$\mathbb{E}[B] = \frac{2 - p}{2(1 - p)}. \tag{2.3}$$

The mean service time in a switch can thus be approximated by (2.3). In order to approximate the mean sojourn time in a switch, we also need $\mathbb{E}[B^2] = \mathbb{E}[\tilde{S}^2]$. Karol et al. (1987) provide a numerical procedure to find the distribution of \tilde{S} , which in particular allows us to approximate $\mathbb{E}[\tilde{S}^2]$. Using this procedure we can approximate the mean sojourn time in the switch by

$$\mathbb{E}[S] = \frac{\rho}{1 - \rho} \left(\frac{\mathbb{E}[B^2]}{\mathbb{E}[B]} - \frac{1}{2} \right) + \mathbb{E}[B], \tag{2.4}$$

where $\rho = p\mathbb{E}[B]$.

2.2 The KKL model

Kim et al. (2000) also model the switch as a $Geo/G/1$ model for $N \rightarrow \infty$, but they assume a geometric service time distribution with parameter q . We only briefly discuss the outcome of the modelling here. For a more comprehensive analysis, the reader is referred to Kim et al. (2000).

Kim et al. use the mean service time found by Karol et al. (1987) as the parameter that describes the geometric service time distribution. Thus,

$$q = \frac{2(1 - p)}{2 - p} = 1 - \frac{p}{2 - p}.$$

For the $Geo/Geo/1$ model with arrival rate p and service rate q , the mean sojourn time can be derived using (1.1), which results in

$$\mathbb{E}[S] = \frac{1 - p}{q - p}.$$

With our specific q , this implies

$$\mathbb{E}[S] = \frac{(1 - p)(2 - p)}{p^2 - 4p + 2}.$$

Note that the model is based on a non-saturated switch, which means that the results are only valid for $p < 2 - \sqrt{2}$.

2.3 Geometric approximation

We propose a new approximation in which we assume that the service time is geometrically distributed. In particular we want to apply this model to small switches, so an approximation based on the limit of $N \rightarrow \infty$ would give inaccurate results. Instead, we find an approximation \hat{q} of the service rate q and we use this approximation as the success probability of the geometric distribution. We regard \hat{q} and q as functions of p , so we write $\hat{q} = \hat{q}(p)$ and $q = q(p)$ in this section.

The throughput of a switch will also play a key role in our approximation. It is defined as the expected number of served flits per time slot, divided by N . It has been well established that switches have a maximal throughput (Beekhuizen 2005; Boot 2005; Karol et al. 1987), which we refer to as saturation throughput and denote by $T_{\text{sat}}(N)$. The throughput of a switch is independent of the packet size, as long as it is deterministic (see Sect. 3). Table 1 shows the saturation throughput for some values of N . For small N , the throughput can be determined using a Markov chain approach. For large N we have to resort to simulation approximations, or to its limiting value $2 - \sqrt{2} \approx 0.586$ (see Karol et al. 1987).

It readily follows that

$$q(p) = T_{\text{sat}}(N) \quad \text{for } T_{\text{sat}}(N) \leq p \leq 1, \tag{2.5a}$$

$$\lim_{p \downarrow 0} q(p) = 1, \tag{2.5b}$$

where the latter expression follows from the observation that $\lim_{p \downarrow 0} q(p)$ represents the service rate if all flits arrive at an empty switch.

In order to derive a good approximation to $q(p)$ for all p , we will interpolate between the two values found in (2.5). It appeared that a linear interpolation does not yield a very good approximation, so we propose a quadratic interpolation in which $q'(0)$ is also used. In order to find such an approximation, we analyse the situation in which $p > 0$ is (arbitrarily) close to zero, while neglecting $\mathcal{O}(p^2)$ terms. If we neglect $\mathcal{O}(p^2)$ we can determine the entire light traffic service time distribution, which also gives us $q = 1/\mathbb{E}[B]$.

We consider an arbitrary time slot t and an arbitrary flit which we tag. Suppose that the tagged flit arrives at a non-empty switch. Because there is at least one flit present from slot

Table 1 Saturation throughputs for several N . Note that $N = 1$ represents a single wire, so obviously $T_{\text{sat}}(1) = 1$

N	$T_{\text{sat}}(N)$	N	$T_{\text{sat}}(N)$
2	0.75	8	0.6184
3	0.6825	9	0.6146
4	0.6552	10	0.6116
5	0.6399	11	0.6091
6	0.6302	12	0.6071
7	0.6238	∞	0.586

$t - 1$, there must have been at least two flits present in that time slot. This in turn implies that at some point in time there must have been two simultaneous arrivals. Since this happens with a probability of $\mathcal{O}(p^2)$, we may ignore the situation in which a flit arrives at a non-empty switch.

So we consider a tagged flit arriving at an empty system. The tagged flit is almost always switched except if another flit arrives simultaneously with the same destination and wins contention. Note that the probability of two or more *other* arrivals is $\mathcal{O}(p^2)$, which we neglect. The probability that one other flit arrives is $(N - 1)p + \mathcal{O}(p^2)$ because there are $N - 1$ remaining input streams and at each input stream an arrival happens with probability p . The two flits have the same destination with probability $\frac{1}{N}$, in which case the tagged flit is not switched with probability $\frac{1}{2}$. Multiplying these probabilities gives us that the tagged flit is switched with probability $1 - \frac{1}{2} \frac{N-1}{N} p + \mathcal{O}(p^2)$. Since another arrival in the next time slot would induce another factor p , we know that the tagged flit is switched in time slot $t + 1$. Altogether we obtain

$$B = \begin{cases} 1 & \text{w.p. } 1 - \frac{1}{2} \frac{N-1}{N} p + \mathcal{O}(p^2), \\ 2 & \text{w.p. } \frac{1}{2} \frac{N-1}{N} p + \mathcal{O}(p^2), \end{cases} \tag{2.6}$$

which implies that $\mathbb{E}[B] = 1 + \frac{1}{2} \frac{N-1}{N} p + \mathcal{O}(p^2)$ and

$$q(p) = \frac{1}{\mathbb{E}[B]} = \frac{1}{1 + \frac{1}{2} \frac{N-1}{N} p + \mathcal{O}(p^2)} = 1 - \frac{1}{2} \frac{N-1}{N} p + \mathcal{O}(p^2). \tag{2.7}$$

As a result $q'(0) = -\frac{1}{2} \frac{N-1}{N}$, which results in the following quadratic approximation:

$$\widehat{q}(p) = \begin{cases} 1 - \frac{1}{2} \frac{N-1}{N} p + \left(\left(1 + \frac{1}{2} \frac{N-1}{N} \right) T_{\text{sat}}(N) - 1 \right) \frac{p^2}{(T_{\text{sat}}(N))^2}, & \text{for } 0 \leq p < T_{\text{sat}}(N) \\ T_{\text{sat}}, & \text{for } p \geq T_{\text{sat}}(N). \end{cases} \tag{2.8}$$

As stated in the introduction, the switch is modelled as a *Geo/Geo/1* queue, for which

$$\mathbb{E}[S] = \frac{1 - p}{q - p}, \tag{2.9}$$

where S is the sojourn time. Of course the mean service time is given by

$$\mathbb{E}[B] = \frac{1}{q}. \tag{2.10}$$

Both quantities can be approximated by substituting \widehat{q} for q .

Remark 2.1 Perhaps the accuracy of the approximation of q can be improved by taking higher order terms into account. Most importantly, however, this would prevent us from considering only flits that arrive at an empty system, which complicates the analysis. Although incorporating higher order terms constitutes an interesting research option, the present approximation is sufficiently accurate for our purposes, as is shown in Sect. 5.

3 Approximation for $K > 1$

In this section we extend our approximation to the case $K > 1$. From the analysis of simulation results, we can infer that there is a certain periodicity in the service time distribution.

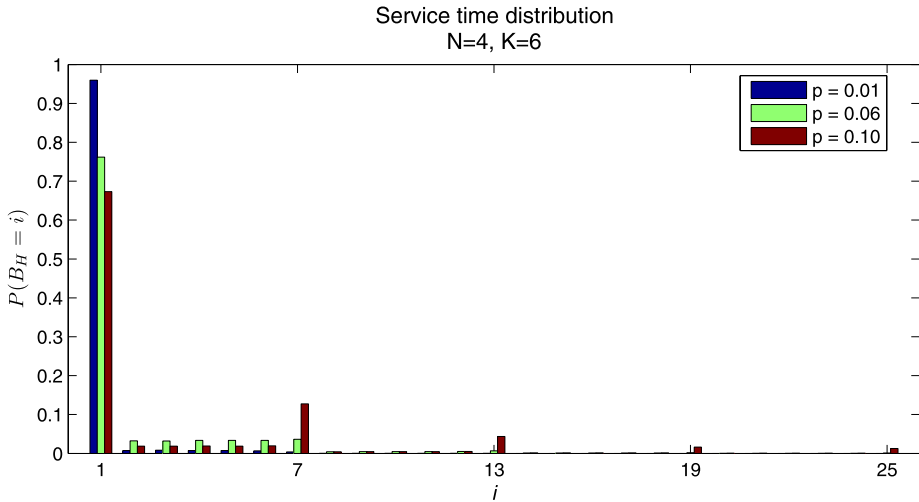


Fig. 3 Simulation results for the distribution of the header service time, denoted by B_H . Note that $p = 0.1$ means that on average 0.6 flits arrive at the switch per time slot, which is close to the saturation value of 0.655242 flits per time slot

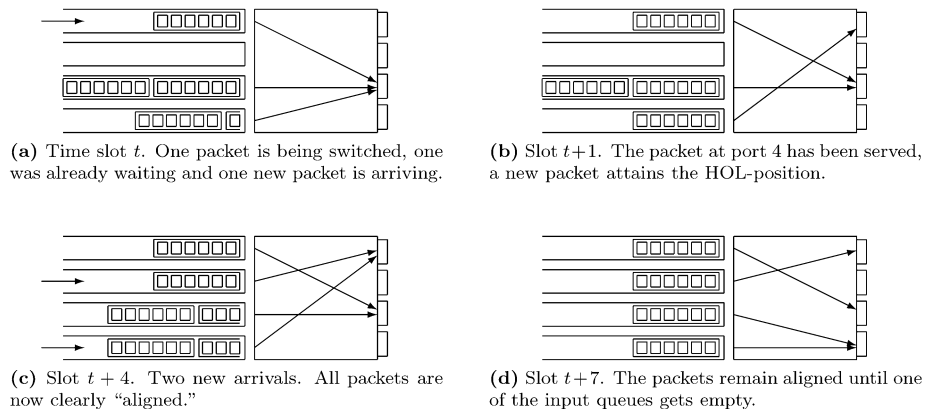


Fig. 4 Alignment of packets. In the pictures the packets arrive at the switch in their entirety but this does not fundamentally change the alignment concept

This periodicity is perhaps best explained in Fig. 3 where the service time distribution is plotted for $K = 6$. There are large peaks for $i = 1, 7, 13, 19, \dots$, especially for $p = 0.1$, while for $i = 2, \dots, 6, i = 8, \dots, 12, i = 14, \dots, 18$, and $i = 20, \dots, 24$, the service probabilities seem to be uniform. Note that on average Kp flits arrive at the switch per time slot, so $p = 0.1$ implies that the arrival rate is close to the saturation value of 0.655242.

Visual simulation output provided a very good explanation for this periodicity. Because the packet sizes are the same for all inputs, the services of the packets become gradually aligned, as shown by Fig. 4. Once the packets are aligned, the alignment can only be broken if one of the queues gets empty, which explains why the phenomenon is more apparent for large p .

The alignment phenomenon also implies that the throughput does not depend on K . If the load exceeds the saturation load the packets remain aligned because the queues do not become empty with probability 1. In this case K is effectively only a time scaling factor. Because the throughput is defined as the expected number of switched flits in steady state, this implies that the throughput is equal for all values of K as long as K is deterministic.

As a result of the alignment, we know that the geometric approximation of $K = 1$ is equivalent to that of $K > 1$ once alignment has occurred. We therefore extend the geometric approximation by ignoring the possibility that packets are not aligned and by maintaining the assumption that the number of successive output conflicts is geometrically distributed. However, the parameter of the geometric distribution may depend on K , so we denote it by q_K . Because of the equivalence when alignment occurs, we approximate q_K by $\hat{q}_K = \hat{q}(Kp)$, with \hat{q} as in (2.8).

By assuming this geometric distribution of the number of successive output conflicts, we in fact assume that an output conflict is lost with probability $1 - q_K$, and won with probability q_K . If an output conflict is lost, the flit has to wait for an additional K time units, until the other packet has fully completed its transition. After this, the header of the packet is again switched with probability q_K and so on. Recall that wormhole routing is used, so non-header flits are always immediately switched. The resulting distribution of B_H , the service time of a header, can thus be characterised as follows:

$$\mathbb{P}(B_H = mK + 1) = q_K(1 - q_K)^m \quad \text{for } m \geq 0. \quad (3.1)$$

We approximate this by substituting \hat{q}_K for q_K .

In (3.1) we essentially ignore the possibility that the service time of a header is not of the form $mK + 1$, with $m \in \mathbb{N}$. This is obviously not entirely realistic, but in Sect. 5 it is shown that the probability that the service time is not of the form $mK + 1$ is rather small.

In order to analyse the mean sojourn time in the switch we shall now analyse the tandem network.

4 Network analysis

In this section we analyse the network under the assumption that $K > 1$. As a result, the input process of the switch is no longer Bernoulli and the switch can therefore no longer be seen as a *Geo/G/1* queue. Instead, there are train arrivals where each train has a length equal to the busy period in a *Geo/D/1* queue. The number of empty slots between two successive trains still follows a geometric distribution with parameter p .

This particular output process clearly complicates the analysis of the network. This complication, however, can be circumvented by first regarding an artificial model in which packets arrive in their entirety at the switch. This model is studied in Sect. 4.1. The results of Sect. 4.1 are then used to determine the mean sojourn time and end-to-end delay in the switch in the original network in Sect. 4.2.

4.1 Arrivals at the switch

In this section we consider an artificial situation in which packets arrive at the switch in their entirety. In other words, each time slot a packet of size $K > 1$ arrives at each input queue according to i.i.d. Bernoulli processes with parameter p . The switch itself can now be seen as a *geo/G/1* queue. Because the results of this subsection will be used in the analysis

of our network model in the next subsection, we denote all performance measures of this subsection by tildes above the normal letters.

In Sect. 3 we assumed that

$$\mathbb{P}(B_H = mK + 1) = q_K(1 - q_K)^m \quad \text{for } m \geq 0,$$

where B_H denotes the service time distribution of a header. We cannot use this distribution in the $Geo/G/1$ model directly, because the header is a special first customer in each batch; the service time of each flit behind the header is always equal to 1 whereas the header suffers from contention. This difficulty can be resolved by viewing the entire packet as one ‘super-customer’.

The service time \tilde{B} of a super-customer is equal to the time it takes the header to win the output conflict and an additional $K - 1$ time slots for the non-header flits to be transmitted, i.e. $\tilde{B} = B_H + (K - 1)$. In other words

$$\mathbb{P}(\tilde{B} = mK) = q_K(1 - q_K)^{m-1} \quad \text{for } m \geq 1. \tag{4.1}$$

Obviously, $\mathbb{E}\tilde{B} = K/q_K$. For the second moment of \tilde{B} , we get

$$\mathbb{E}[\tilde{B}^2] = \frac{2 - q_K}{q_K^2} K^2.$$

The mean sojourn time $\mathbb{E}\tilde{S}$ in the switch can be found by applying (1.1) and using $S = W + B$:

$$\mathbb{E}[\tilde{S}] = \frac{pK}{q_K - pK} \left(\frac{K}{q_K} - \frac{1}{2}(K + 1) \right) + \frac{K}{q_K}. \tag{4.2}$$

4.2 Arrivals at the NI

We will now study the original model with arrivals at the NI again. Our goal is to determine the mean sojourn time in the switch (S) and the mean sojourn time of a packet in the network (D). The sojourn time of a packet in the network (or delay) is defined as the time between its arrival at the NI and the departure of its last flit from the switch.

The sojourn time of a packet in the network consists of the sojourn time of the header in the NI, the sojourn time of the header in the switch, and the time difference between the departures of the header and the last flit from the switch, which is equal to $K - 1$. We thus obtain

$$D = S_{NI} + S + (K - 1), \tag{4.3}$$

where S_{NI} is the sojourn time of a header in the NI.

It can easily be seen that $D = \tilde{S} + 1$ where \tilde{S} is the sojourn time in the switch in the model of Sect. 4.1, so

$$\mathbb{E}[D] = \mathbb{E}[\tilde{S}] + 1 = \frac{pK}{q_K - pK} \left(\frac{K}{q_K} - \frac{1}{2}(K + 1) \right) + \frac{K}{q_K} + 1. \tag{4.4}$$

To find the mean sojourn time in the switch $\mathbb{E}[S]$ we thus only have to determine $\mathbb{E}[S_{NI}]$.

Because packets arrive at the NI in batches of size K according to a Bernoulli process, the NI itself can be seen as a $Geo^K/D/1$ queue with unit service times. For the batch size Λ we have

$$\mathbb{P}(\Lambda = 0) = 1 - p,$$

$$\mathbb{P}(\Lambda = K) = p,$$

and the utilisation rate $\rho = \mathbb{E}[\Lambda] = pK$. Furthermore, $\mathbb{E}[\Lambda^2] = pK^2$. The mean waiting time of a header in the NI is the mean waiting time of the first customer in a batch, which is given by (1.2). Hence,

$$\mathbb{E}[S_{NI}] = \mathbb{E}[W_{NI}] + 1 = \frac{pK(K - 1)}{2(1 - pK)} + 1.$$

Finally, we apply (4.3) to obtain

$$\begin{aligned} \mathbb{E}[S] &= \mathbb{E}[D] - \mathbb{E}[S_{NI}] - (K - 1) \\ &= \frac{pK}{q_K - pK} \left(\frac{K}{q_K} - \frac{1}{2}(K + 1) \right) - \frac{pK(K - 1)}{2(1 - pK)} + \frac{K}{q_K} - (K - 1). \end{aligned} \quad (4.5)$$

5 Approximation comparison

In this section, the performance of the approximations of the previous sections is analysed. First we look at the case $K = 1$ and we study the quality of \hat{q} , the approximation of q . We will do so by analysing the approximations of the mean service time, see Fig. 5. It is clear that our approximation (2.10) with $q = \hat{q}$ performs much better than the other approximations if $N = 4$. This result is not very surprising, as our approach is based on small N rather than the limit of $N \rightarrow \infty$. We furthermore see that for $N = 128$ our approximation is equally accurate as the approximations based on $N \rightarrow \infty$.

Perhaps an even more important performance measure is the mean sojourn time, which is analysed in Fig. 6. From Fig. 6a it is again clear that for $N = 4$, approximation (2.9) outperforms the other approximations. In Fig. 6b, the relative error of our approximation is plotted. Until the system approaches saturation, there is a maximum relative error of roughly 1%.

We also see that the asymptotic approximations saturate too early. This can easily be understood if we observe that the asymptotic saturation point is $2 - \sqrt{2} \approx 0.586$ while the

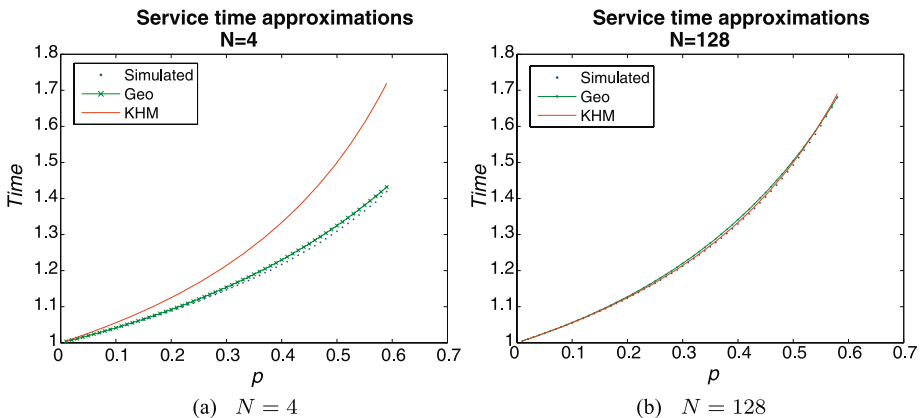


Fig. 5 Mean service time approximations for $N = 4$ and $N = 128$. Recall that the mean service time approximations of the KKL and the KHM model are identical

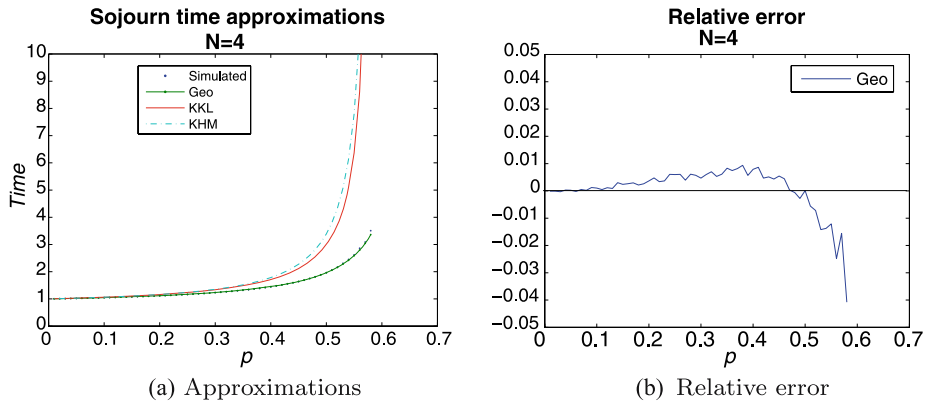
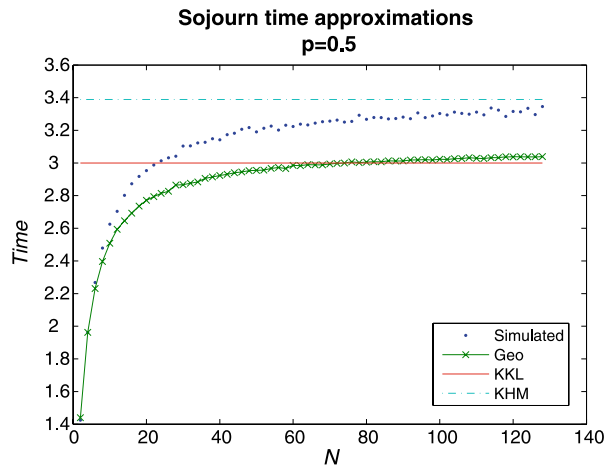


Fig. 6 Mean sojourn time approximations and the relative error of our approximation for $N = 4$

Fig. 7 Mean sojourn time approximations with ρ fixed. Our model clearly performs better for small N , but the simulation results seem to converge to the KHM estimate as $N \rightarrow \infty$. Note that both the KKL and the KHM model are based on $N \rightarrow \infty$, which explains the horizontal line



saturation point for $N = 4$ is 0.655242 (see Table 1). This difference leads to very large errors if ρ is close to saturation. For instance, for $\rho = 0.5$ the mean sojourn time approximation of Karol et al. has a relative error of roughly 70%.

In order to get some insight in the role of N , we look at the performance of the approximations while fixing ρ and varying N . This has been done for $\rho = 0.5$ in Fig. 7. Most importantly, we can conclude that our model is a considerable improvement over the other models for small N . For large N , say $N \gtrsim 60$, the numerical procedure of Karol et al. gives the best approximation.

Interestingly enough, our approximation of the mean service time did give quite accurate results for $N = 128$, whereas our mean sojourn time approximation gives a significant error. Apparently the error we make in assuming that the service time distribution is geometric becomes more important for larger N . This conclusion is also backed by the fact that the KKL approximation is quite close to our approximation if N is large. In Sect. 6 we analyse the consequences of our assumption that the service time distribution is geometric in more detail.

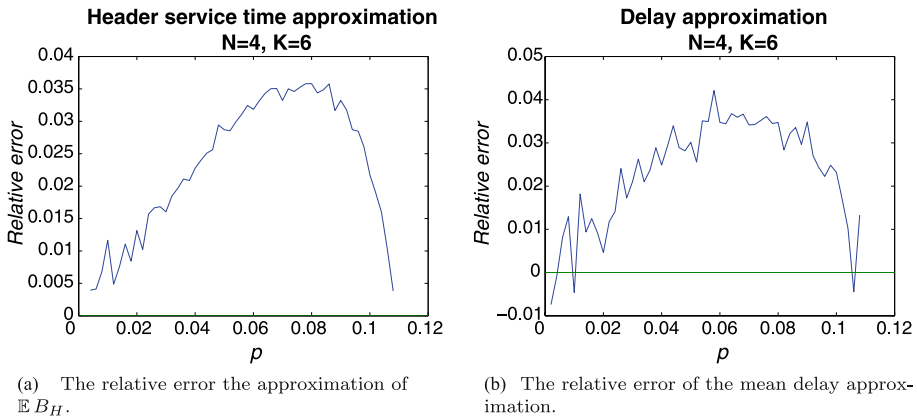


Fig. 8 The relative error of the mean delay and mean header service time approximations

Now that we have looked at the approximation available for the case $K = 1$, we study the quality of our approximation for the case $K > 1$. In this case our approximation still gives very good results as Fig. 8 illustrates: The relative error of the approximation of $\mathbb{E}[B_H]$ compared to simulation outcomes is maximally 3.5%.

The mean delay can be approximated by substituting $\hat{q}(K\rho)$ for q_K in (4.4). The relative error of this approximation can be found in Fig. 8b. Our approximation still gives quite good results since the relative error of the delay approximation is at most 4.5%.

6 Validation of the geometric distribution

In the analysis of Sect. 2.3, we assumed that the service time is geometrically distributed with the service rate q as parameter, which we approximated by \hat{q} . In this section we try to validate the assumption that the service time distribution is geometric. For the sake of clarity we stress that q is defined as the service rate, and \hat{q} is our quadratic approximation given by (2.8). Moreover, in this section all values of q have been determined via simulation.

It can rather easily be argued that the service time distribution is in fact not geometric; the geometric distribution is memoryless, but the switching probability is not. For instance, if all HOL-positions are occupied with packets with the same destination, each packet has probability $\frac{1}{N}$ of being switched. If the newly arriving packet has a different destination, then for all remaining packets, the probability that they are switched in the next time slot is $\frac{1}{N-1}$, and so on. This immediately implies that there is some dependency on the history of the process, yet the precise effect of this dependency is still unclear.

In order to study to what extent the service time distribution deviates from a geometric distribution, we compare the service time distribution obtained via simulation to our geometric approximation, both with parameter q and \hat{q} (see Fig. 9). From this figure it seems that the difference between the approximated distribution and the simulation outcomes is rather limited.

To confirm this suspicion, we look at the first and second moments and the squared coefficient of variation in Table 2. In addition to this, Table 2 shows the value of $\mathbb{E}[B^2]/2\mathbb{E}[B]$ which has a prominent influence on the mean sojourn time approximation. Note that $\mathbb{E}[B^2]/2\mathbb{E}[B] + 1/2$ is the mean residual service time in discrete-time.

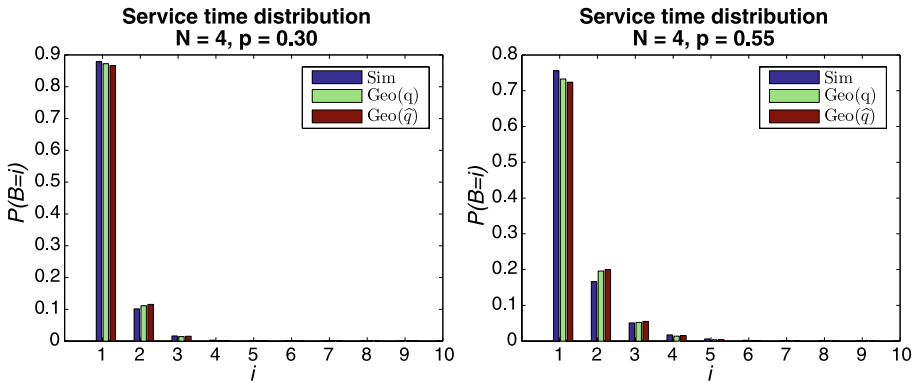


Fig. 9 The service time distribution approximated by a $geo(q)$ and $geo(\hat{q})$ distribution

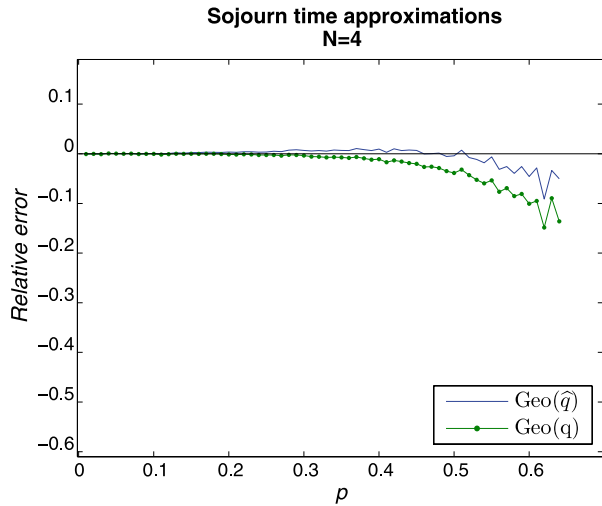
Table 2 The approximations of the first and second moments of the service time, as well as the coefficient of variation, and $\frac{\mathbb{E}[B^2]}{2\mathbb{E}[B]}$, compared to simulation outcomes. In this comparison we set $N = 4$ and $N = 128$ with $p = 0.55$

	Sim	$geo(\hat{q})$	$geo(q)$
(a) $N = 4$			
$\mathbb{E}[B]$	1.3649	1.3813	1.3649
$\mathbb{E}[B^2]$	2.4712	2.4346	2.3609
$\frac{\text{Var}(B)}{(\mathbb{E}B)^2}$	0.3265	0.2760	0.2673
$\frac{\mathbb{E}[B^2]}{2\mathbb{E}[B]}$	0.9053	0.8813	0.8649
(b) $N = 128$			
$\mathbb{E}[B]$	1.4930	1.5049	1.4928
$\mathbb{E}[B^2]$	3.3362	3.0243	2.9641
$\frac{\text{Var}(B)}{(\mathbb{E}B)^2}$	0.4968	0.3355	0.3301
$\frac{\mathbb{E}[B^2]}{2\mathbb{E}[B]}$	1.1173	1.0049	0.9928

We see in the table that the simulated values of these quantities lie closer to the $geo(\hat{q})$ distribution than to the $geo(q)$ distribution, except for the mean service time. In other words, the approximation with \hat{q} actually performs better than the approximation with q . In addition to this, we see that the difference between the simulated service time distribution and the geometric distributions becomes larger as N increases.

We also compare the mean sojourn time approximations if a $geo(q)$ and $geo(\hat{q})$ service time distribution is assumed. In Fig. 10, the relative errors of these sojourn time approximations are plotted. We can clearly see the influence of the error in $\mathbb{E}B^2$; for a large range of p (say $p \gtrsim 0.35$), our approximation clearly performs better if \hat{q} is used instead of q . Apparently the error in approximating q compensates to some extent for the error in the geometric distribution assumption. This can be explained by the fact that assuming a geometric service time distribution with parameter q leads to an underestimation of the $\mathbb{E}[B^2]$ (see Table 2). Because the second moment of a geometric distribution increases as its parameter decreases, and because q is underestimated by \hat{q} , we again get a more accurate approximation.

Fig. 10 Relative errors of sojourn time approximations with a $geo(q)$ and $geo(\hat{q})$ service time distribution



Remark 6.1 For $N = 2$ under saturation, the service time distribution is geometric. It can rather easily be argued that each flit has a probability of $\frac{3}{4}$ of being switched, regardless of what happened in previous time slots.

7 Conclusion

If $K = 1$, the service time distribution of the switch can be approximated by a geometric distribution with parameter \hat{q} , where \hat{q} is a quadratic approximation of the service rate. For small switches, such as those in Networks-on-Chips, this approximation is a significant improvement over the approximations of Karol et al. (1987) and Kim et al. (2000). For $N = 4$, our sojourn time approximation has a relative error of maximally 1% if the switch is not saturated, whereas the KHM approximation already has a relative error of 70% for $p = 0.5$ and $N = 4$. The approximation \hat{q} of the service rate q is very accurate in general, even for large N . Nevertheless, the error of the mean sojourn time approximation becomes larger if N increases. In particular, this is due to the fact that the service time distribution becomes 'less geometric' if N increases.

If $K > 1$, we can maintain the geometric approximation because packets become 'aligned'. The second complication that arises if $K > 1$ is that there are train arrivals at the second node of our network. This can be overcome by analysing an artificial model in which packets arrive in their entirety at the router. In this artificial model we can determine the mean delay and sojourn time and using them we can determine the mean delay and sojourn time in our original network. The resulting approximation also has a very small relative error, of maximally 4.5% if $N = 4$ and $K = 6$.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Beekhuizen, P. (2005). *Mathematical modelling of networks-on-chips with best effort traffic*. Master's thesis, Eindhoven University of Technology.

- Boot, N. (2005). *Throughput and delay analysis for a single router in networks on chip*. Master's thesis, Eindhoven University of Technology.
- Dally, W. J., & Towles, B. (2001). Route packets, not wires: on-chip interconnection networks. In *DAC proceedings* (pp. 684–689).
- Eng, K. Y., Karol, M. J., & Yeh, Y. S. (1989). A growable packet (ATM) switch architecture: Design principles and applications. In *Proc. of IEEE GLOBECOM* (pp. 1159–1165).
- González Pestana, S., Rijpkema, E., Rădulescu, A., Goossens, K., & Gangwal, O. (2004). Cost-performance trade-offs in networks on chips: A simulation based approach. In *DATE proceedings*.
- Gravey, A., & Hébuterne, G. (1992). Simultaneity in discrete-time single server queues with Bernoulli inputs. *Performance Evaluation*, 14, 123–131.
- Hluchyj, M. G., & Karol, M. J. (1988). Queueing in high-performance packet switching. *IEEE Journal on Selected Areas in Communications*, 6(9), 1587–1597.
- Karol, M. J., Hluchyj, M. G., & Morgan, S. P. (1987). Input versus output queueing on a space-division packet switch. *IEEE Transactions on Communications*, 35(12), 1347–1356.
- Kim, H., Kim, K., & Lee, Y. (2000). Derivation of the mean cell delay and cell loss probability for multiple input-queued switches. *IEEE Communications Letters*, 4(4), 140–142.
- Rădulescu, A., Dielissen, J., Goossens, K., Rijpkema, E., & Wielage, P. (2004). An efficient on-chip network interface offering guaranteed services, shared memory abstraction, and flexible network configuration. In *DATE proceedings*.
- Rijpkema, E., Goossens, K., Rădulescu, A., Dielissen, J., van Meerbergen, J., Wielage, P., & Waterlander, E. (2003). Trade-offs in the design of a router with both guaranteed and best-effort services for networks on chip. In *DATE proceedings*.
- Takagi, H. (1993). *Queueing analysis: A foundation of performance evaluation: Vol. 3. Discrete-time systems*. Amsterdam: North-Holland.
- Xiong, Y., & Bruneel, H. (1993). Buffer contents and delay for statistical multiplexers with fixed-length packet-train arrivals. *Performance Evaluation*, 17(1), 31–42.