ELSEVIER

Journal of Systems Architecture 47 (2001) 529–541

JOURNAL OF SYSTEMS ARCHITECTURE

www.elsevier.com/locate/sysarc

Buffer management control in data transport network node

Ivan Hanuliak

University of Žilina, Moyzesova 20, 010 26 Žilina, Slovakia

Received 1 August 1999; received in revised form 1 April 2001; accepted 4 June 2001

Abstract

The application of analytical queuing theory results in behaviour analysis of a distributed computer network or mobile data system (data transport network). It belongs to the preferred method in comparison to the simulation method. The use of analytical methods allows us to calculate effectively various values of parameters in equilibrium, including the total input intensity of data units to every node of transport network even for the more realistic models than the M/M/1 systems. However, these results are derived assuming an infinite buffer in size at a given node. For practical application, we need to project the concrete number of buffers in every node. This paper describes the method of buffer management control for each decomposed network's node of a data transport network in two real cases. For this purpose, the linear dependence between buffer memory size and input queue size at each node of a data transport network was used.

For these two real statistical distributions of incoming data units closed expressions are derived, enabling to calculate the required queue size for both queue size limitation methods (assumption of the unlimited and limited queue sizes). For practical use, a very efficient way of computing queue overflow probabilities was developed.

Keywords: Data transport network; Distributed computer network; Queuing theory system; Buffer management; Queue overflow; Kendall classification; Overflow probability

1. Introduction

During the recent years, we have been able to see rapid development of distributed computer networks (e.g., distributed personal computer networks in our country) and of mobile data systems. Both of them consist of two parts: the powerful one, created by host computers (host processors – HP) as the providers of the network services (server, fileserver, print server, fax server, etc.), and the connected communication part, providing necessary network connections and communications (e.g., the communication computers in the form of routers, etc.).

The structure of a typical communication interface is illustrated in Fig. 1. The primary input message queue, with one or more of the same processors $P_1, P_2, \ldots, P_r$ (symmetrical multiprocessor system), is modelled as one queuing theory system, and each node's communication equipment to communication among nodes, as another one. The communication interface together with the set of point-to-point transmission lines constitute the communication network (data transport system) of a distributed computer network or of a mobile data system.

E-mail address: hanuliak@frtk.utc.sk (I. Hanuliak).

1383-7621/01/$ - see front matter
PII: S 1 3 8 3 - 7 6 2 1 (0 1) 0 0 0 2 4 - 8

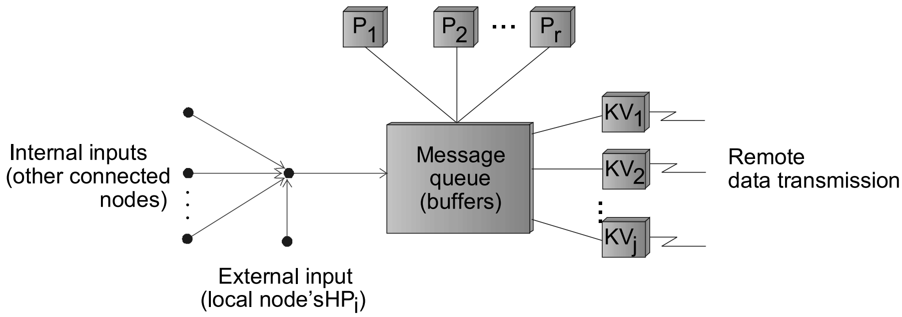

Fig. 1. Block structure of a communication interface.

2. Modelling of the transport system

Data units switching store-and-forward (S&F) communication network (transport network) is considered. The network is represented by a weighted graph. Messages arrive at random at a source station and follow a specific route in the networks towards their destination station. Message lengths are considered to be random variables, following generally an arbitrary distribution. The message is divided at the source station into fixed length sub-messages, called data units (packets, words, characters, bytes, bits). Those data units are then sent independently through the transport network nodes towards the destination station. At each node a queue of data units is served according to a first-come first-served (FCFS) discipline. All analytical models for S&F communication networks, utilising an "independence assumption" (which requires re-choosing the message length, at random at any node). This intuitive assumption was verified by Kleinrock [16].

If we define the individual communication network nodes as graph nodes and their mutual communication lines as graph edges, we generally can get an oriented graph with U-nodes for the transport system with U-communication computers (Fig. 2) where:

- $\gamma_1, \gamma_2, \ldots, \gamma_u$ represent total intensity of input data stream to the given node (the summary input stream from the connected input terminals and host computers to the given node computer). It is given as Poisson input stream with intensity λ demands in time units.
- r_{ij} are given as the relation probabilities from node i to the neighbouring connected nodes j.
- $\beta_1, \beta_2, \ldots, \beta_u$ correspond to the total external output stream of data units from the individual nodes (the total output stream to the connected host computers and terminals of the given node computer).

2.1. Transport network analysis methods

Two basic differentiating principles have been developed for computer network analysis:

- Analytical methods using the queuing mathematical theory [2,4–6,9–13,15–17,19,20].
- Discrete simulation [3,8,14,21].

Discrete simulation can give a very interesting insight to the behaviour analysis of smaller transport networks, but it is fully unusable to analyse large, complex transport networks. It is very useful mainly in cases when analytical methods do not exist, and it is the only analytical tool.

The theoretical approach based on mathematical queuing theory application is a very effective and practical tool for the analysis of large, complex data transport networks.

Concrete analytical methods are as follows:

- Kleinrock analytical model based on Jackson theorem (decomposition to the mutually independent network nodes on the basis of the M/M/1 systems with infinite buffer). This, together with the verified independence assumption [16], reduces a very difficult problem to an open network of independent queues.