Available online at www.sciencedirect.com
ScienceDirect
Int. J. Electron. Commun. (AEÜ) 62 (2008) 185 – 192
International Journal of Electronics and Communications AEÜ
www.elsevier.de/aeue

Bandwidth allocation and pricing in multinode network

Jyrki Joutsensalo*, Ari Viinikainen, Mika Wikström, Timo Hämäläinen

Department of Mathematical Information Technology, University of Jyväskylä, P.O. Box 35 (AGORA), 40014 University of Jyväskylä, Finland

Received 28 April 2006; accepted 15 March 2007

Abstract

This paper presents adaptive resource sharing model that uses a revenue criterion to allocate network resources in the optimal way. The model ensures QoS requirements of data flows and, at the same time, maximizes the total revenue by adjusting parameters of the underlying scheduler. Besides, the adaptive model eliminates the need to find the optimal static weight values because they are calculated dynamically. The simulation consists of several cases that analyse the model and the way it provides the required QoS guarantees. The simulation reveals that the installation of the adaptive model increases the total revenue and ensures the QoS requirements for all service classes.

Keywords: Pricing; Revenue optimization; Bandwidth allocation; Packet scheduling; Quality of Service (QoS)

1. Introduction

Today's Internet supports QoS for the flows that are sensitive to effects like delay and jitter. It must ensure that all the flows receive their reserved resources while QoS is also maintained. To ensure this, there must be mechanisms to give the guaranteed bandwidth and computational resources to incoming flows. However, the allocation of bandwidth and CPU resources are interdependent and maintaining fairness in one resource allocation does not necessarily entail fairness in another resource allocation. Therefore, for the better maintenance of QoS guarantees and overall fairness in resource allocations for the contending flows, the processor and bandwidth scheduling schemes should be integrated. A significant amount of work has been done in resource scheduling for traditional networks.

The fair allocation of bandwidth is typically achieved by using per-flow queueing mechanisms that are complex to implement such as Fair Queueing [1,2] and its many variants [3–5]. However, these mechanisms require that each arriving packet has to be classified into a flow and the router must perform several operations based on the updated per-flow state variables. Several methods have been presented to reduce the complexity of the per-packet operations, such as [6–12]. In [13] the complexity is localized in the edge routers, where the per-flow information is computed, while the core routers just use first-in-first-out (FIFO) queueing and keep no per-flow state. However, it still remains unclear if these algorithms can be cost-effectively implemented.

In this paper we present a low complexity packet scheduling algorithm (to be used, e.g. in edge routers) for the allocation of a fair share of bandwidth to different service classes while providing optimal revenue to the service provider. We extend our previous studies [14,15], where delay was considered as QoS parameter.

The rest of the paper is organized as follows: Section 2 discusses network model and bandwidth formulation. In Section 3 the pricing function and the optimal weights for revenue maximization are presented. We present also an approximation of the optimal algorithm in Section 4, that

* Corresponding author. Tel.: +358 142603296; fax: +358 142602545.
E-mail addresses: jyrkij@mit.jyu.fi (J. Joutsensalo), arjuvi@mit.jyu.fi (A. Viinikainen), wikstrom@mit.jyu.fi (M. Wikström), timoh@mit.jyu.fi (T. Hämäläinen).

1434-8411/$ - see front matter
doi:10.1016/j.aeue.2007.03.019

achieves near optimal results with lower computational requirements. Section 5 considers implementation issues and computational complexity. In Section 6 the operation of the algorithm is simulated. Finally, the conclusions are discussed in Section 7.

2. Multinode network and bandwidth

In this section, we formulate expression for bandwidth (bit rate) of the data traffic in the multinode network. Although the algorithm operates in the general multinode system with an arbitrary number of nodes, we present it by using the four-node case to avoid complicated notation. Consider a network with four-edge nodes (Fig. 1). Data are transmitted from nodes 1 and 2 to either nodes 3 and 4. There are three service classes, namely gold, silver, and bronze. The customers of the gold class pay more than the customers of the silver and bronze classes – in our case for the available bandwidth – but on the other hand, they get more bandwidth.

Notation $N_j^{i\to p}$ means that there are $N_j^{i\to p}$ such connections (customers) in the jth class which transfers data through nodes i and p. For example, in node 1 in the queue of class j, there are $N_j^{1\to 3}$ connections with next destination node 3 and $N_j^{1\to 4}$ connections with next destination node 4. The total number of connections in the node i and class j is denoted by N_{ij}, and it obeys the condition

$$N_{ij} = \sum_{p=1}^{n} N_j^{i\to p}, \tag{1}$$

where n denotes the number of the nodes. In our example case, $n = 4$, and

$$N_{1j} = N_j^{1\to 3} + N_j^{1\to 4}, \tag{2}$$

$$N_{2j} = N_j^{2\to 3} + N_j^{2\to 4}, \tag{3}$$

$$N_{3j} = N_j^{1\to 3} + N_j^{2\to 3}, \tag{4}$$

$$N_{4j} = N_j^{1\to 4} + N_j^{2\to 4}. \tag{5}$$

Here, e.g. $N_1^{1\to 2} = 0$. Generally $N_{ij} = N_j^{i_1\to i_2} + N_j^{i_1\to i_3}$.

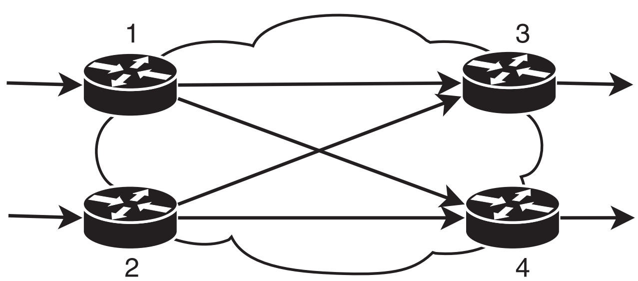

Fig. 1. Four edge node system, with two ingress and egress nodes.

Let us consider the bandwidth in the node i. Let the processing time of the data be T [s/bit] in the packet scheduler. There are N_{ij} connections or packets in class j. Let us denote the packet size b_{ijk} [bits] or [kbytes] in the node i, $i = 1, \ldots, n$, class $j = 1, \ldots, m$ and the connection $k = 1, \ldots, N_{ij}$. Variable w_{ij} is the weight allocated for class j in node i. Constraint for weights w_{ij} is

$$\sum_{j=1}^{m} w_{ij} = 1, w_{ij} > 0. \tag{6}$$

Variables w_{ij} give weights, how long time queues (i, j) are served per total time. It is easy to see that bit rate of the connection (i, j, k) is

- linearly proportional to the packet size b_{ijk},
- linearly proportional to the weight w_{ij},
- inversely proportional to the processing time T, and
- inversely proportional to the total sum of the packet lengths b_{ijk}, $k = 1, \ldots, N_{ij}$, because other packets occupy the same band in a time-divided manner.

Therefore, the expression for the bit rate of the connection (i, j, k) is

$$B_{ijk}[\text{bits/s}] = \frac{K b_{ijk} w_{ij}}{T \sum_{l=1}^{N_{ij}} b_{ijl}} = \frac{b_{ijk} w_{ij}}{\sum_{l=1}^{N_{ij}} b_{ijl}}, \tag{7}$$

where the processing time T in the denominator and the proportionality factor K in the numerator can be scaled $K/T = 1$, without the loss of generality. If processing times differ from node to node, notation is more complicated, but there is no critical difference on our formalism.

3. Pricing model and revenue optimization

We concentrate on the pricing and fair resource allocation from the point of view of the customers. On the other hand, from the point of view of the service provider, we try to maximize revenue. First, we introduce the concept of *pricing function*. Naturally, the price is usually concave with respect to the bandwidth; for example, when images are transferred, the price may be twice compared to the price when voice is transferred, while the bandwidth ratio – image bandwidth/voice bandwidth – is much more than two.

Consider the price paid by customers in class j to the service provider. It depends on the used bit rate. The price $r_j(B)$ is increasing with respect to the bit rate B, and it is concave. We use the *power* pricing model, because using that model, one can get a closed form approximating solution and the model is increasing and concave.

Definition. The pricing model

$$r_j(B) = r_j B^p, \tag{8}$$