Available online at www.sciencedirect.com
SCIENCE DIRECT
Computer Communications 28 (2005) 1741–1751
computer communications
www.elsevier.com/locate/comcom

Reinforcement learning-based dynamic bandwidth provisioning for quality of service in differentiated services networks

Chen-Khong Tham*, Timothy Chee-Kin Hui

Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Received 3 December 2004; accepted 3 December 2004
Available online 3 February 2005

Abstract

The issue of bandwidth provisioning for Per Hop Behavior (PHB) aggregates in Differentiated Services (DiffServ) networks is imperative for differentiated QoS to be achieved. This paper proposes an adaptive provisioning scheme that determines at regular intervals the amount of bandwidth to provision for each PHB aggregate, based on traffic conditions and feedback received about the extent to which QoS is being met. The scheme adjusts parameters to minimize a penalty function that is based on the QoS requirements agreed upon in the service level agreement (SLA). The novel use of a continuous-space, gradient-descent reinforcement learning algorithm enables the scheme to work effectively without accurate traffic characterization or any assumption about the network model. Using ns-2 simulations, we show that the algorithm is able to converge to a policy that provisions bandwidth such that QoS requirements are satisfied.

Keywords: Continuous-space reinforcement learning; Adaptive bandwidth provisioning; Quality of service; Differentiated services

1. Introduction

Differentiated Services (DiffServ) [1] has been widely accepted as the service model to adopt for providing quality of service (QoS) over next-generation IP networks. This is in contrast to most current day networks that implement QoS, which are mainly circuit-based networks that use the Integrated Services (IntServ) [2] service model. The difference between the service model frameworks is one of paradigm. IntServ provides QoS at a per-flow level by allocating network resources to each individual flow based on their required QoS levels. While this provides good control of QoS levels, it is at the expense of scalability, which is required for next-generation all-IP networks that carry a large number of flows at any one time. The scalability problem stems from the need for per-flow mechanisms, such as per-flow scheduling, queuing and admission control, and per-flow information, such as per-flow states. DiffServ, on the other hand, seeks to develop a trade-off between control and scalability by aggregating flows. Instead of per-flow QoS mechanisms, the DiffServ framework introduces the concept of Per Hop Behaviors (PHBs) [3,4], which provide aggregated levels of QoS to the aggregated flows. In exchange for scalability, the DiffServ paradigm faces the challenging task of maintaining QoS control without per-flow control and per-flow information. We argue that while strict QoS control is good for mission-critical applications, most applications are non-critical; users may specify certain tolerance levels but may not mind the occasional breach of QoS. Examples of such applications are voice over IP (VoIP), online transactions and video conferencing.

DiffServ, as defined in the IETF RFCs, originally provides for QoS in a qualitative way; a higher class of service gets relatively better QoS than a lower class of service in terms of throughput, latency and packet losses [5]. While this may be sufficient for better-than-best-effort requirements, this is certainly not sufficient for next-generation applications that require certain quantitative levels of QoS, as evidenced by service level agreements (SLAs) [6] being drawn up today that specify quantitative tolerance levels. QoS mechanisms therefore require

* Corresponding author.
E-mail addresses: cktham@ieee.org (C.-K. Tham), timhui@ieee.org (T. Chee-Kin Hui).

0140-3664/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2004.12.018

quantitative means of adjustment to meet such QoS requirements.

Most current-day networks provide QoS by over-provisioning their links or by using priority schedulers. The drawback to this is that resources are often over-allocated and wasted. Often, the level of service cannot be well-controlled. For example, when using priority schedulers with multiple classes of traffic, only relative QoS can be achieved. This may mean that a higher class may be receiving exceptionally good QoS at the expense of a lower class, which itself may have some QoS requirements. The over-provisioning case is often seen in leased-line and MPLS networks. For these networks, QoS is exceptionally good at the expense of the links being under-utilized. The solution to this is to have efficient dynamic bandwidth provisioning that is able to juggle resources between classes, such that the QoS requirements of no particular class are breached and no resources are wasted.

In this paper, we propose the novel use of a continuous-space, gradient-based reinforcement learning bandwidth provisioning algorithm that adjusts the weights of class-based fair schedulers [7] to changing traffic conditions and congestion. The algorithm uses a penalty function based on the SLA requirements and feedback-based control to adjust bandwidth provisions according to traffic parameters. To minimize the penalty function, our proposed intelligent scheme adaptively finds a policy that balances bandwidth provisions such that QoS for all classes are not breached.

The organization of this paper is as follows. In Section 2, we formulate the bandwidth provisioning problem and show that it is 'hard' to solve and survey some existing methods that have been proposed to solve the problem heuristically. In Section 3, we describe our algorithm called reinforcement learning-based dynamic provisioning (RLDP) to solve the provisioning problem dynamically in discrete time intervals. The impetus for using reinforcement learning is also presented. In Section 4, we present results from ns-2 [8] simulations which demonstrate the algorithm's capability. Finally, we discuss several additional issues related to the implementation of RLDP in actual networks in Section 5 and conclude in Section 6.

2. Bandwidth provisioning in DiffServ networks

2.1. Background

Bandwidth provisioning in DiffServ networks involves the determination of the amount of bandwidth to allocate for each PHB aggregate across each network link. This is usually done at the router's outgoing ports through weighted fair packet scheduling [1]. By provisioning bandwidth, each PHB aggregate shares the bandwidth in a certain proportion as they contend for the use of a network link to transmit data packets from one node to another. By allocating different proportions of bandwidth, the service levels of each PHB can be differentiated. Unfortunately, the proportion of bandwidth to provision is a complex decision due to the interaction of a variety of factors, such as the traffic mix, the level of QoS required and other QoS mechanisms in the network. For this reason, bandwidth provisioning needs to be adaptive.

In the DiffServ framework [1], the amount of bandwidth to provision for each PHB aggregate is determined by the service level requirements as stated in the SLA. Often the amount of bandwidth to provision is proportional to the strictness of the requirements, i.e. the lower the delay bound and the higher the throughput requirement, the more bandwidth needs to be provisioned. The EF PHB is thus given more bandwidth than is needed (over-provisioned) as it has the strictest of requirements. The AF PHB is on the other hand only slightly over-provisioned as it has more elastic requirements. BE traffic usually gets served with the remaining capacity. Although this is a widely-used method, it is almost always used to provide qualitative provisioning. To provision quantitatively requires fairly complex analysis that needs to be based on a variety of factors, such as traffic characteristics, QoS requirements and QoS mechanisms in the network, in order for the correct level of provisioning to be determined [9].

2.2. Formulation of bandwidth provisioning problem

The bandwidth provisioning problem can be broadly formulated as follows:

For all nodes,
Given,
$x_{EF,j}(t)$, $x_{AF,j}(t)$, $x_{BE,j}(t)$: traffic rates of each class entering node destined to leave through link j, where $j \in$ all outgoing links from node.
C_j: capacity of link j.
Select,
$w_{EF,j}(t)$, $w_{AF,j}(t)$, $w_{BE,j}(t)$: Weighted fair proportion of bandwidth for each class on link j.
Constrained by,
$l_{EF} < l^*_{EF}$, $l_{AF} < l^*_{AF}$, $l_{BE} < l^*_{BE}$, $D_{EF} < D^*_{EF}$, $D_{AF} < D^*_{AF}$, where l and D are the loss and delay percentages and l^* and D^* are the loss and delay requirements. The delay requirement is stated as a bound on the percentage of packets that are allowed to exceed the end-to-end latency bound.

Note that the solution space of the problem varies depending on the tightness of the QoS constraints. As such, though a solution can be readily found through some approximate method when the QoS requirements are loose, a solution is not so easily found when the QoS constraints are tight. This is due to some reasons which we now discuss.

Firstly, the problem above is a continuous time problem. This means that at every time instant, there is a different optimization problem to solve due to the changing traffic