Computer Communications 33 (2010) 1557–1565

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Multiple path selection algorithm for DiffServ-aware MPLS traffic engineering

Wu-Hsiao Hsu [a], Yuh-Pyng Shieh [a,*], Jenhui Chen [b]

[a] Department of Computer Science and Information Engineering, Ming Chuan University, Taoyuan, Taiwan
[b] Department of Computer Science and Information Engineering, Chang Gung University, Taoyuan, Taiwan

ARTICLE INFO

Article history:
Received 5 September 2008
Received in revised form 7 April 2010
Accepted 9 April 2010
Available online 21 April 2010

Keywords:
DiffServ-TE
MSA
CT
LSP

ABSTRACT

This paper proposes a new per-class bandwidth constraint algorithm, called the multipath selection algorithm (MSA), for a DiffServ-aware traffic engineering (DiffServ-TE). The MSA comprises three steps. First, a given source uses the MSA to find multiple label switch paths (LSPs) from the source to a destination for a specific class type (CT). Second, the source uses the available bandwidth of the CT on all the links along these LSPs to allocate the initial traffic to the selected LSPs. Third, the source dynamically adjusts traffic to these LSPs based on individual round trip time. Simulation results indicate that the proposed algorithm offers better performance than existing approaches in average transmission time, average packet loss rate, average throughput, and available bandwidth variance for each link.

1. Introduction

The current approach in providing quality of service (QoS) on the Internet is based on the DiffServ protocol [1]. DiffServ divides traffic into a small number of classes and allocates network resources on a per-class basis. In this architecture, packets are marked with different DiffServ code points (DSCP) at edge routers. Each DSCP is associated with a particular QoS class characterized by per-hop-behavior (PHB), and the core routers treat each packet with a specific PHB based on the DSCP carried by the packet. The PHB is achieved through a combination of scheduling and queue management schemes. The DiffServ architecture includes several standardized PHBs. The expedited forwarding (EF) PHB provides a low-loss, low-latency, low-jitter, and assured bandwidth service. The assured forwarding (AF) PHB supports services in which the customers are likely to get the negotiated service level agreement (SLA) without any guarantees.

Multiprotocol label switching (MPLS) traffic engineering (MPLS-TE) [2] makes it possible to establish bandwidth guaranteed label switched paths (LSPs) using constraint-based routing algorithm. However, since MPLS-TE operates without referring to different classes, it may not be optimal in a DiffServ network. Several analyses of integrating DiffServ and MPLS-TE can be found in [3–8]. The studies [3,4] introduce the concept of DiffServ-aware traffic engineering (DiffServ-TE). DiffServ-TE makes separate bandwidth reservations for different classes of traffic. Hence, traffic flows toward a given destination can be forwarded on separate LSPs based on class. For this purpose, the studies define the concept of a class type (CT) as the set of traffic trunks crossing a link, which is governed by a specific set of bandwidth constraints. A given traffic trunk belongs to the same CT at all links. The TE class is introduced as a pair of a CT and a preemption priority allowed for that CT. The IETF requires the support of up to eight CTs, referred to as CT0 through CT7. By definition, each CT is assigned to either a bandwidth constraint (BC), or a set of BCs. A CT represents a class in the DiffServ-TE architecture much like PHB represents a class for DiffServ. Note that flexible mappings between CTs and PHBs are possible.

The study [5] analyzes the QoS performance for different types of services in a DiffServ-TE network, including VoIP, real time video, and best effort data traffic. The study [6] proposes an architecture for the MPLS restoration routing of DiffServ traffic. This architecture, called per class aggregate information with preemption (CAIP), provisions two key QoS features for multimedia traffic: prioritized guaranteed bandwidth and fast restoration in the event of an element failure.

The study [7] proposes a new preemption policy that includes an adaptive scheme aimed at minimize rerouting. This policy combines the three main preemption optimization criteria: number of LSPs to be preempted, priority of the LSPs, and preempted bandwidth. All the studies above focus on preemption policy and restoration routing.

The study [8] proposes a Max–Min bandwidth constraint model that guarantees each CT without causing resource fragmentation. This paper also develops three new bandwidth preemption

* Corresponding author. Address: Department of Computer Science and Engineering, Ming-Chuan University, 5 Teh-Ming Rd., Gwie Shan District, Taoyuan, Taiwan. Tel.: +886 3 3507001/3429; fax: +886 3 3593874.
E-mail address: arping@gmail.com (Y.-P. Shieh).

0140-3664/$ - see front matter Crown Copyright © 2010 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2010.04.015

algorithms for three bandwidth constraint models, respectively, and focuses on how to design a bandwidth manager to support DiffServ-TE.

This paper proposes a new per-class bandwidth constraint algorithm, called the multipath selection algorithm (MSA), for a DiffServ-TE network. Unlike previous studies, which focus on preemption policy, restoration routing, or bandwidth manager, the proposed MSA finds multiple LSPs per-class and allows flexible division of traffic over these LSPs. The MSA comprises three steps. First, a given source uses the MSA to find multiple LSPs from the source to a given destination for a CT. Second, the source uses the available bandwidth of the CT on all the links along these LSPs to allocate initial traffic. Third, the source dynamically adjusts traffic for these LSPs based on individual round trip time.

1.1. Bandwidth constraint models

The maximum allocation model (MAM) [9], the Russian doll model (RDM) [10], and the maximum allocation with reservation (MAR) [11] are three IETF-proposed bandwidth constraint models for supporting DiffServ-TE. The author of [12] compared these three models and concluded that the RDM best matches DiffServ-TE. Hence, the MSA proposed in this study uses the RDM as a bandwidth constraint model. The RDM provides each class with a minimum amount of bandwidth, but lower priority classes can use the bandwidth of higher priority classes when that bandwidth is available.

1.2. The link state interior gateway protocol (IGP)

In the proposed MSA, each node in a DiffServ-TE network works in conjunction with the extensions of the open shortest path first (OSPF) protocol [13]. In the extended OSPF protocol, each node running a link state QoS routing protocol uses reliable flooding to exchange link state advertisements (LSAs) with its neighboring routers. Each LSA must advertise the available bandwidth per-CT on every link. Based on the reliable flooding of LSA, all nodes build identical link state databases that depict the entire OSPF network topology interconnected by a group of nodes. When the available bandwidth per-CT of one or more links changes, the link state database in each node must be updated immediately.

The remainder of this paper is organized as follows. Section 2 introduces the notation and problem in this study. Section 3 presents the proposed algorithm. Section 4 describes the simulation model and results, and Section 5 provides some conclusions.

2. Notation and problem description

2.1. Notation

Before formally introducing the proposed algorithm, the notation used throughout this paper will first be described. Let $G=(V,L)$ denote a DiffServ-TE network, where V is the set of nodes and L is the set of links. Suppose that a node represents a router. A path p from a node x to a node y is a sequence of nodes and links $x=v_0, l(v_0\ \ v_1), l(v_1\ \ v_2), \ldots, l(v_k\ \ v_{k+1}),\ y=v_{k+1}$ and is denoted by $p(v_0, v_1, v_2, \ldots, v_k, v_{k+1})$. $AB(p)$ represents the maximum available bandwidth of path p, and is defined as $AB(p)=\min\{l_a^{(v_i,v_{i+1})}(CT_j)|0 \leqslant i \leqslant k\}$. A link capacity between nodes x and y is denoted by $l_c(x,y)$. Let n denote the number of CTs, and BC_i denote the bandwidth reserved by CT_i, $0 \leqslant i \leqslant n$. The available bandwidth of a link between nodes x and y for the CT_j can be denoted by $l_a^{(x,y)}(CT_j)$, $0 \leqslant j \leqslant n$. Since the proposed MSA uses the RDM as the bandwidth constraint model, the link available bandwidth $l_a^{(x,y)}(CT_j)$ can be computed as follows,

$$l_a^{(x,y)}(CT_j) = \sum_{i=j}^{n} l_a^{(x,y)}(CT_i) + \left(l_c(x,y) - \sum_{i=0}^{n} l_a^{(x,y)}(BC_i) \right). \tag{1}$$

2.2. Problem description

In Fig. 1, the number indicated at each link represents the current link available bandwidth for a CT_j (for example, $l_a^{(A,B)}(CT_j)=8$), and the bandwidth requested for a particular LSP from the source A to the destination H is 10 Mbps. Clearly, no single path has enough bandwidth to meet the 10 Mbps requirement. If the round trip time from source A to the destination H is also considered, finding a path which meets the request and has the minimum round trip time is a NP problem [14]. In fact, the network architecture in Fig. 1 shows that there is more than one path from the source A to the destination H. Thus, concurrent multi-path transmission can meet the request when a single path transmission cannot.

3. The multipath selection algorithm (MSA)

The main purpose of the MSA is to meet the requested bandwidth by finding multi-LSPs for a CT in a DiffServ-TE network. The MSA procedures can be divided into two parts:

(1) Finding multi-LSPs for the request.
(2) Allocating network traffic to the selected LSPs.

3.1. Finding multi-LSPs

The MSA uses two metrics to find multi-LSPs: the path round trip time and the available bandwidth of each link comprising the path. Since router queuing delay completely dominates the path transmission delay from a source to a destination, the source selects the path with the fewest hops as a LSP. The reason for selecting a path with as few hops as possible is that more links on a path not only consume more network resources, but also increase propagation delay [15]. In other words, a LSP with fewer hop counts (HCs) will have a shorter path round trip time.

The three principles of finding the multi-LSPs are as follows:

(I) All the found LSPs have no loop.
(II) The source selects the path with sufficient available bandwidth and the fewest number of nodes as a LSP. Any link in the LSP can be selected repeatedly by other LSPs as long as the link has enough available bandwidth. When a link no longer has enough bandwidth to carry more LSP traffic, these links are not selected.
(III) If nodes want to keep the most current view of the available bandwidth on all links in the network, they must update the link state database frequently. However, frequent link state database updates are neither scalable nor practical every

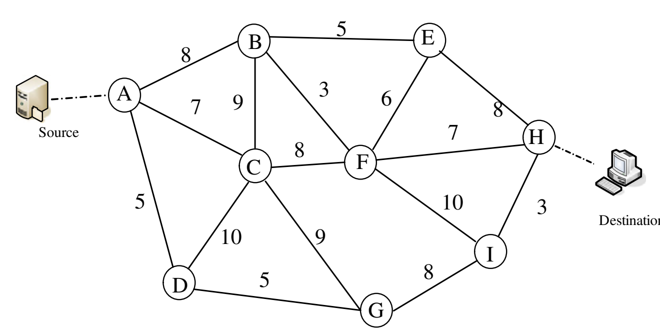

Fig. 1. A simple network architecture.