Computer Networks 65 (2014) 167–182

Contents lists available at ScienceDirect

Computer Networks

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

# Multi-objective traffic engineering for data center networks

Trong-Viet Ho [*], Yves Deville [1], Olivier Bonaventure [2]

*ICTEAM Institute, Université catholique de Louvain, Place Sainte Barbe 2, 1348 Louvain-la-Neuve, Belgium*

ARTICLE INFO

*Article history:*
Received 12 May 2012
Received in revised form 7 February 2014
Accepted 18 March 2014
Available online 27 March 2014

*Keywords:*
Traffic engineering
Multi-objective optimization
Data center traffic
Local search
Multiple Spanning Tree Protocol

ABSTRACT

Data centers are now the basis for many Internet and cloud computing services. The Spanning Tree Protocol and its variants have been widely used in data center networks for a couple of decades. An efficient use of the limited spanning tree links would enable to solve the traffic engineering problem in data centers. In this paper, we propose five local search approaches for generating a good set of spanning trees in data centers using multiple VLANs. The quality of these algorithms is evaluated by different multi-criteria assessment methods. The performance of each algorithm is assessed based on three standard measures: maximal link utilization, sum load, and the number of used links.

## 1. Introduction

Currently, data centers containing thousands of servers are used to support a growing number of services. In many of these environments, the required bandwidth is growing quickly and network operators need to find solutions which ensure that their network can sustain the traffic demand without having overloaded links.

Current enterprise data centers contain a large number of Ethernet switches [29–31]. The IEEE 802.1s Multiple Spanning Tree Protocol (MSTP) [1,2] is widely deployed in large Ethernet networks. Network operators often divide the network into different regions, called Virtual Local Area Networks (VLANs) [3]. Note that two VLANs can share some nodes. MSTP configures the spanning tree for each VLAN by computing the least-weight path from each switch to an elected root switch. From a traffic engineering (TE) viewpoint, the major drawback is that MSTP itself cannot ensure a good set of spanning trees for a given set of traffic demand matrices.

Solving the TE problem in large Ethernet networks can be defined as the optimization of the network based on many evaluation metrics, such as link utilization, link delay, sum load, and number of used links. It is impossible to optimize all these metrics at the same time because an improvement in one metric can degrade other metrics. In practice, network operators would like to provide the best network performance based on a compromise among a given set of metrics. Multi-objective optimization [4] is an interesting approach for solving this kind of TE problem, where each objective is related to one metric. Solving the single objective problem is already very complex, as the search space is exponential even for networks of reasonable size. Approximate methods, such as Constraint-Based Local Search (CBLS) [5], have been proved to be well suited for solving complex combinatorial problems with huge search spaces such as this TE problem. In this paper, we focus on approximate methods based on local search.

Online TE [6] approaches are supposed to dynamically adapt the routing when traffic changes in real-time. However, the network devices are required to be more

---

[*] Corresponding author. Tel.: +32 (0)10/47 31 50; fax: +32 (0)10/45 03 45.

*E-mail addresses:* trong.ho@uclouvain.be (T.-V. Ho), yves.deville@uclouvain.be (Y. Deville), olivier.bonaventure@uclouvain.be (O. Bonaventure).

[1] Tel.: +32 (0)10/47 20 67; fax: +32 (0)10/45 03 45.

[2] Tel.: +32 (0)10/47 90 12; fax: +32 (0)10/45 03 45.

http://dx.doi.org/10.1016/j.comnet.2014.03.018
1389-1286/

intelligent and support more flexibles policies rather than standard switches to perform the local adjustments due to the actual traffic situation of the network (e.g. OpenFlow switches [8]). Traditional "offline" TE is used to improve the performance based on the given traffic demand matrices. The goal is to provide global solutions for the network design and configuration. The advantage is that the implementation of offline TE approaches is often perfomed on cheap and normal switches. Hence, we choose an offline approach for solving the TE problem in enterprise data center networks. Offline TE approaches are suitable in public datacenters where companies rent a large number of servers. In such datacenters, each customer defines its contractual traffic matrix and the network needs to be configured to sustain this matrix. The traffic generated by each customer can vary within the bounds of the contracted traffic matrix. The Offline Traffic Engineering techniques proposed in this paper are suitable for environments were the traffic demand is stable or can be correctly estimated. Such environments include enterprise datacenter that are running long batch jobs such as scientific computing. Public cloud providers could also benefit from the techniques. In public clouds, the virtual machines that are used by the customers are often group in one VLAN per customer for traffic isolation reasons. Furthermore, many cloud providers use shaping to limit the bandwidth consumed by each virtual machine [7]. The shaping configuration of each virtual machine could provide the traffic demand matrix used by our Offline TE technique to optimise the network.

In this paper, we extend the TE problem defined in [9], but with two bi-objective optimization problems. The first one is the minimization of the maximal link utilization and the sum load. The second one is the minimization of the maximal link utilization and the number of used links. The maximal link utilization appears in both problems because it is the essential metric for assessing the quality of a TE technique. In this paper, we propose different local search algorithms which are evaluated on different topologies. Our contribution is mainly focused on the design, development, analysis, and experimentation of multi-objective methods for the TE problem in data centers.

This paper is organized as follows. We first present related research in Section 2. We describe the theoretical background of MSTP, Local Search, and Multi-objective optimization in Section 3. Section 4 presents the problem statement. Section 5 describes our five proposed local search algorithms. Our evaluation is presented in Section 6 with an analysis of the experimental results, and we conclude the paper in Section 7.

## 2. Related work

Many TE techniques have been proposed for improving the performance of 802.1s MSTP in switched Ethernet networks. The proposed approaches can be divided into four main classes. First, [10–12] map a set of VLANs to a given number of spanning trees. Second, [13–15] assign a set of flows to a set of given spanning trees. Third, [16] addressed the TE problem by building multiple source-based spanning trees (constructing a spanning tree for each of the given source nodes). Last, [17,18] proposes solving the TE for MSTP by finding the set of spanning trees described by customer traffic demands and a given network topology. But these techniques only cope with networks of small size, containing fewer than 30 switches. Our local search algorithm in [9] follows the same approach as in [17,18], but aims at minimizing the maximal link utilization in large Ethernet networks containing more than 500 switches. This paper is an expanded version of [9,19].

From the multi-objective optimization view point, [11] took into account link load balancing and admission fairness. The framework in [12] considered both network throughput and delay. Recently, [15] aimed at minimizing the $n$ worst link loads (with $n$ ranging up to the total number of network links) and the average link load in the network. However, the final output of these approaches is a unique solution where the second metric is only considered as a constraint or a lexicographical objective. In practice, the outcome of multi-objective algorithms is often a set of non-dominated solutions, called a Pareto front [4]. Many assessment methods for evaluating the quality of a Pareto front have been presented, such as in [20–22]. In the present paper, our local search algorithms will be evaluated using five different assessment methods (described in Section 6).

## 3. Background

Before describing, analyzing, and evaluating our multi-objective algorithms, we would like to clarify the concepts of Multiple Spanning Tree Protocol 802.1s, local search (LS), and multi-objective optimization (MO), which will be used throughout this paper.

### 3.1. Multiple Spanning Tree Protocol 802.1s

The multiple Spanning Tree Protocol allows the utilization of several spanning trees inside a single Ethernet network. It can be used in large networks such as data centers relying on switched Ethernet. We restrict here our analysis to VLANs and Ethernet in data centers. Recently, the TRILL [23] (TRansparent Interconnection of Lots of Links) protocol has been proposed by Perlman: it has the ease of configuration of the Ethernet, while benefitting from the routing techniques provided at Layer 3 with the ability to multicast and broadcast. TRILL is currently supported by some high-end switches. In contrast, the Spanning Tree Protocol is implemented in all commodity Ethernet switches. In this section, we describe the Spanning Tree Protocol and its variants.

First, the IEEE 802.1d Spanning Tree Protocol (STP) [24] reduces the network topology to a single spanning tree. To compute the spanning tree, special messages called Bridge Protocol Data Units (BPDUs) are exchanged between switches. Each BPDU message contains three pieces of information: the root ID (ID of the switch assumed to be the root), the transmitting bridge ID (ID of the switch transmitting the BPDU), and the weight (the weight of the least-weight path to the root from the transmitting switch). By default, the STP assigns a weight to each port