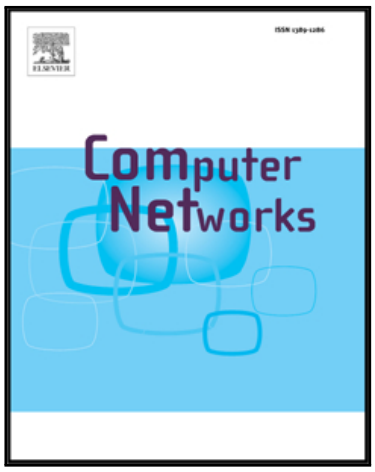

Energy-Efficient Dynamic Virtual Network Traffic Engineering for North-South Traffic in Multi-Location Data Center Networks

Mirza Mohd Shahriar Maswood, Chris Develder, Edmundo Madeira, Deep Medhi

PII: S1389-1286(17)30173-1
DOI: 10.1016/j.comnet.2017.04.042
Reference: COMPNW 6181

To appear in: *Computer Networks*

Received date: 2 November 2016
Revised date: 7 March 2017
Accepted date: 12 April 2017

Please cite this article as: Mirza Mohd Shahriar Maswood, Chris Develder, Edmundo Madeira, Deep Medhi, Energy-Efficient Dynamic Virtual Network Traffic Engineering for North-South Traffic in Multi-Location Data Center Networks, *Computer Networks* (2017), doi: 10.1016/j.comnet.2017.04.042

Energy-Efficient Dynamic Virtual Network Traffic Engineering for North-South Traffic in Multi-Location Data Center Networks

Mirza Mohd Shahriar Maswood[a], Chris Develder[b], Edmundo Madeira[c], Deep Medhi[a,*]

[a] *University of Missouri–Kansas City, Kansas City, MO, USA*
[b] *Ghent University – Imec, Belgium*
[c] *University of Campinas, Brazil*

Abstract

We consider the problem of allocating data center (DC) resources for cloud enterprise customers who require guaranteed services on demand. In particular, a request from an enterprise customer is mapped to a virtual network (VN) class that is allocated both bandwidth and compute resources by connecting it from an entry point of a data center to one or more hosts while there are multiple geographically distributed data centers to choose from. We take a dynamic traffic engineering approach over multiple time periods in which an energy-aware resource reservation model is solved at each review point. For the energy-aware resource reservation problem, we present a mixed-integer linear programming (MILP) formulation (for small-scale problems) and a heuristic approach (for large-scale problems). Our heuristic is is fast for solving large-scale problems where the MILP problem becomes difficult to solve. Through a comprehensive set of studies, we found that a VN class with a low resource requirement has a low blocking even in heavy traffic, while the VN class with a high resource requirement faces a high service denial. Furthermore, the VN class having randomly distributed resource requirement has a high provisioning cost and blocking compared to the VN class having the same resource requirement for each request although the average resource requirement is same for both these VN classes. We also observe that our approach reduces the maximum energy consumption by about one-sixth at the low arrival rate to by about one-third at the highest arrival rate—this also depends on how many different CPU frequency levels a server can run at.

Keywords: Data Center Networks, Resource Optimization and allocation on-demand, Request Blocking, Energy Efficiency, Virtual Network, Dynamic Traffic Engineering

1. Introduction

The increasing growth of cloud based applications such as video streaming, web search, distributed file systems, scientific computations, software libraries and document collection made the data centers (DC) a popular platform in the Internet world. Companies such as Amazon, Google, Facebook, and Yahoo! routinely employ data centers for storage, web services and large-scale computations [1], [2], [3]. With the increase in demand, the size and number of DCs are increasing day by day. Large-scale data centers are set up with a large number of servers that are interconnected through routers, switch, and high speed links [4]. Due to the growing usage of data centers, the expenses of maintenance are also increasing. Power consumption is a major concern in operating data centers as most of the equipment in data centers are temperature sensitive and cooling through air and water is necessary to keep the temperature within an acceptable limit. Moreover, operating the servers, routers and switches also requires a huge amount of power. Data centers in the USA consumed about 91 billion kilowatt hours annually in 2013 and are estimated to consume 140 billion kilowatt-hours of electricity annually by 2020 [5]. Hence, reducing the energy consumption of data centers has been a challenging research problem. The ultimate aim behind designing a data center is reducing the expenses while gaining the highest efficiency.

There has been a number of contributions so far to increase the efficiency of data centers by better utilizing the server resources, applying traffic engineering techniques to reduce the bandwidth and other operational costs. Some of them [6], [7] focus on energy efficient resource provisioning using dynamic traffic engineering. However, to our knowledge, no work has considered how both compute resources at the end hosts and network resources inside the data center are allocated to satisfy the request of virtual network (VN) customers while minimizing both energy consumption and bandwidth cost. Secondly, most work related to traffic engineering of intra-DC networks consider east-west traffic, i.e., the intra data center traffic between hosts. In our work, we focus instead on enterprise customers' requests that result in north-south traffic in data centers requiring both network bandwidth and server resources. In particular, we address serving different enterprise customer groups using VNs at data centers through dynamic traffic engineering by allocating both network bandwidth and processing resources efficiently, while factoring in energy consumption. That is, we

*Corresponding Author
Email addresses: mmnt7@mail.umkc.edu (Mirza Mohd Shahriar Maswood), chris.develder@intec.ugent.be (Chris Develder), edmundo@ic.unicamp.br (Edmundo Madeira), dmedhi@umkc.edu (Deep Medhi)