Simulation Modelling Practice and Theory 64 (2016) 83–98

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

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

Towards efficient operation of internet data center networks: Joint data placement and flow control for cost optimization

Abdallah Khreishah[a,*], Jacob Chakareski[b], Ammar Gharaibeh[a], Issa Khalil[c], Ali Diabat[d]

[a] New Jersey Institute of Technology, United States
[b] University of Alabama, United States
[c] Qatar Computing Research Institute, Qatar
[d] Masdar Institute of Science and Technology, United Arab Emirates

ARTICLE INFO

Article history:
Available online 6 October 2015

Keywords:
Internet datacenter networks
Data placement
Cost efficiency

ABSTRACT

The problem of cost-efficient operation of data center networks used to deliver **file sharing** services is studied. The aggregate costs are split into server-load-related and link-load-related shares. Thus, the problem of interest is formulated as one of joint data placement and flow control, and mixed integer-linear programming is used to compute the optimal solution. The high complexity of the latter motivated us to design two additional sets of strategies, based on data coding and heuristics, respectively. With coding, a distributed algorithm for the problem is developed. In the simulation experiments, carried out based on actual data center information, network topology and link cost, as well as electricity prices, the advantages of data coding, in particular in the context of multicast, and the impact of different factors such as the network topology and service popularity, on the total cost incurred by all considered strategies, are examined. Network coding with multicast is shown to provide cost savings in the order of 30–80%, depending on the specific context under consideration, relative to the other optimization strategies and heuristic methods examined in this work.

1. Introduction

A defining landmark of the present Internet are data center networks – global distributed systems featuring multiple geographically dispersed data centers and high speed data trunks between them. They are used to deliver a variety of services to increasingly bigger client populations online. The two most recognized characteristics of such systems are their sheer scale and power demand [1]. Therefore, cost efficient operation of data center networks is a necessity. There are two types of cost that an Online Service Provider (OSP) operating a data center network needs to bear. The first one is the data transport cost associated with delivering the services to a multitude of demanding customers. The second one is the operating cost of the data centers that is associated with their actual location in the network. As the two cost items are interrelated, optimizing the network operation solely based on one, may lead to very inefficient performance relative to the case when the aggregate cost is taken into consideration. At the same time, the OSP needs to ensure that the services delivered over the network maintain their corresponding performance requirements at the end user. Traditionally, network

* Corresponding author.
E-mail addresses: abdallah@njit.edu (A. Khreishah), jacob@ua.edu (J. Chakareski), amg54@njit.edu (A. Gharaibeh), ikhalil@qf.org.qa (I. Khalil), adiabat@masdar.ac.ae (A. Diabat).
http://dx.doi.org/10.1016/j.simpat.2015.09.010

traffic engineering to date has exclusively focused on load balancing across multiple internal paths in transit ISP networks [2]. Only in the domain of multihoming,[1] we have witnessed interest in minimizing the transport cost of the customer [3] or the average latency of its data flows, for a given cost constraint [4]. Other studies try to minimize the network operational costs or the average latency by caching the files at different locations of the network [5–9]. The most closely related work is [10] that studies the joint data center placement and flow control in online service provider networks, with the goal of minimizing the overall operational cost of such networks, for the given performance guarantees associated with each service. Relative to [10], various unicast and multicast scenarios where the data is network-coded and the server load is taken into consideration are examined.

The work in [11] studies the problem of joint request mapping and response routing for cloud services running on geographically distributed data centers. The objective is to minimize the total costs (electricity and bandwidth costs) of serving clients' requests. Relative to [11], this work targets different types of cloud applications (i.e. file sharing), which do not have strict delay requirements. Moreover, various unicast and multicast scenarios are considered in this paper.

The work in [12] targets online services, such as search and instant messaging, that have strict delay requirements, and tries to find a curve, where each point in the curve represents a trade-off between the performance metric, which is Round Trip Time (RTT), and the operational costs of the Online Service Provider (OSP). Relative to [12], this paper considers file sharing services, which do not have strict delay requirements. Moreover, the work in [12] does not consider the operational costs of data centers.

The work in [13] proposes a two-stage optimization problem. The first stage is an admission control stage, where requests are selectively admitted to maximize the revenue collected by the Cloud Data Center (CDC) provider. The revenue of a request is modeled as a utility function that depends on the average response time. Given the admitted requests from the first stage, the second stage decides which data center will satisfy the request and which path is chosen to deliver the response in order to minimize the total energy and bandwidth cost. Relative to [13], this work targets different types of cloud applications (file sharing) that do not have strict delay requirements. Moreover, the authors in [13] consider that the bandwidth cost of different links belonging to the same Internet Service Provider is the same, while this work considers that the cost of traversing a link depends on the bandwidth and the length of the link.

Other related work includes [14] that considers incorporating renewable energy sources as prospective suppliers of electricity for data center networks and the demand markets that therefore need to be stimulated, in order for the former to become a viable alternative in this context. Refs. [15,16] study the advantages of network coding versus data replication in distributed content distribution systems.

To illustrate the benefits of our approach, assume that our data center is distributed over three sites that can be accessed over the Internet. Assume that they are located in Tokyo, Dublin, and Seattle (these cities could be changed to any other three).[2] Fig. 1 represents our system model in which a distribution center decides the allocation of service requests to data centers. Assume that the three sites are powered via solar energy. However, if solar energy is not available and the data center is running, then it will use brown energy. In a typical summer day, it is fairly reasonable to assume that at any given time the solar energy will be available at two of the sites, and not available at the third one. Assume that a file of 2 GB, is to be available at any given time through the three sites. Assume for simplicity that each site can store only 1 GB. The objective here is to minimize the use of brown energy. Assume that the available space for this file at each site is 1 GB. One approach to distribute this file over the data centers is to keep the first half of the file (A) on one site, say Tokyo, the second half of the file (B) on the second site, say Dublin, and to replicate either the first half of the file (A) or the second half (B) at Seattle. In this way, the total brown energy consumption would be $\frac{1}{3} \times BEE$, where BEE is the brown energy expenditure during a day per GB of data. That is because, assuming we chose A to be at Seattle, the Dublin site that contains B has to be powered all of the time, and in $\frac{1}{3}$ of the time it uses brown energy. With coding, instead of storing either A or B at Seattle, a coded packet $A \otimes B$ is stored, where $\otimes$ is the bitwise XOR operation. Therefore, at any given time accessing any two of the sites is sufficient to retrieve the whole file. Therefore, the brown energy use in this case will be zero. Fig. 2 represents our approach. The same example can be used to minimize the total energy use when a multi-electricity environment is present, in which each site is powered by a different provider, and the cost of electricity at these sites is different and changing over time. In this case, the site where the cost of electricity is the maximum among the different sites is put to sleep [17–20], and the other sites are turned on, if coding is used. In order to do that without coding, the whole file must be stored in the three sites which is expensive. The same approach can be used to minimize the energy used to cool the data centers, as it will change over time depending on the temperature, the pressure, and other factors.

The main contributions of our paper can be summarized as follows:

- Multiple formulations of the problem under consideration are derived that examine the impact of network coding, multicast, deterministic and stochastic settings, and decentralized operation.
- Network coding is demonstrated to reduce the overall system cost and the problem's complexity (NP-complete to polynomial).
- In the multi-cast case, the problem can be solved in polynomial time with very good performance, if network coding is performed in the intermediate nodes.

[1] A customer is connected to the Internet via multiple ISPs.

[2] This is a fairly realistic example, see for example Microsoft's Azure cloud that is distributed over three continents.