Computer Networks 57 (2013) 2880–2899

Contents lists available at SciVerse ScienceDirect

Computer Networks

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

Greening data center networks with throughput-guaranteed power-aware routing

Mingwei Xu [a,b,c], Yunfei Shang [a], Dan Li [a,*], Xin Wang [d]

[a] Department of Computer Science and Technology, Tsinghua University, Beijing, PR China
[b] State key Lab. of Networking and Switching Technology Beijing University of Posts and Telecommunications, Beijing, PR China
[c] Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing, PR China
[d] Department of Electrical and Computer Engineering, State Univ. of New York at Stony Brook, NY, United States

ARTICLE INFO

Article history:
Received 3 March 2012
Received in revised form 15 October 2012
Accepted 21 December 2012
Available online 28 December 2012

Keywords:
Throughput-guaranteed power-aware routing
Data center network
Network power consumption model

ABSTRACT

Cloud based applications and services require high performance and strong reliability provided by data center networks. To overcome the problem of traditional tree based data center network, recently many new network architectures are proposed, such as Fat-Tree and BCube. They use aggressively over-provisioned network devices and links to achieve 1:1 oversubscription ratio. However, most of the time data center traffic is far below the peak value and a large number of idle network devices and links in data centers consume a significant amount of power, which is now becoming a big problem for many cloud providers.

In this paper, we aim to reduce the power consumption of high-density data center networks from the routing perspective while meeting the network performance requirement. We call this kind of routing *throughput-guaranteed power-aware routing*. The essence of our idea is to use as little network power as possible to provide the routing service, without significantly compromise on the network performance. The idle network devices and links can be shut down or put into the sleep mode for power saving. Extensive simulations conducted in typical data center networks show that our power-aware routing can effectively reduce the power consumption of network devices, especially under low network loads.

1. Introduction

Today's data centers integrate a great number of switches and servers to provide various cloud-based services, such as online search, web mail, e-business, as well as basic computational and storage functions, such as MapReduce [1], GFS [2], and CloudStore [3]. The goal of the data center network (DCN) is to interconnect the massive number of data center servers, and provide efficient and fault-tolerant routing to support upper-layer applications. It has attracted great attention to design a reliable and efficient DCN architecture recently. The traditional tree architecture applied in current data centers [34] is known to face many challenges in supporting bandwidth-hungry communications in data centers. More specifically, the tree structure suffers from low scalability, high cost as well as single point of failure. Hence, recently many advanced network architectures are proposed to mitigate these issues, represented by Fat-Tree [4], BCube [5], etc. These new data center architectures use more network devices and links to effectively overcome the shortcomings of the tree architecture and to enjoy 1:1 oversubscription ratio.

In order to better support data-intensive applications in data centers, these "richly-connected" network architectures are designed with the major purpose of ensuring high communication performance and robustness. These

* Corresponding author at: Department of Computer Science and Technology, Tsinghua University, Beijing, PR China. Tel.: +86 010 62603064; fax: +86 010 62603062.

E-mail addresses: xmw@csnet1.cs.tsinghua.edu.cn (M. Xu), shangyunfei@csnet1.cs.tsinghua.edu.cn (Y. Shang), lidan@csnet1.cs.tsinghua.edu.cn (D. Li), xwang@ece.sunysb.edu (X. Wang).

1389-1286/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.comnet.2012.12.012

architectures have two characteristics in common: *over-provisioned network resources* and *inefficient power usage*. An excessive number of network devices and redundant links are provided aggressively for the busy-hour load. However, most of the time, the traffic in a data center is far below the peak value and varies greatly between day-time and night, which leaves a large number of network devices to stay idle. The goal of network power conservation is to make the power consumption on networking devices proportional to the traffic load [35]. In existing data centers, however, the network at the low load still consumes more than 90% of power used at the busy-hour load [7]. So a large number of idle network devices in high-density networks consume a significant amount of power, which results in an extremely inefficient power usage in data center networks.

The power cost brought by hardware devices such as network devices and servers in data centers accounts for a dominant part of the operational costs of data centers, and it may skyrocket as the scale of data centers expands. The power cost is becoming a big burden for many cloud providers. According to the statistics, the total power consumption of global data centers accounts for 1.1–1.5% of the worldwide electricity use in 2011 [8], and the figure will continually increase to 8% by 2020 under the current trend [42]. It has been shown that network devices consume about 20% power in the whole data center [7], and the ratio will grow with the rapid development of power-efficient hardware and power-aware scheduling algorithms on the server side. Therefore, it is of high importance to investigate advanced power conservation technologies for data center networks, which will in turn bring a great benefit in reducing the operational cost of data centers and contribute to the reduction of the carbon footprint.

The objective of this paper is to propose a novel *throughput-guaranteed power-aware routing* algorithm to reduce the total power consumption of network devices, and to make power usage more efficient in "richly-connected" data center networks. The key idea is to use as little network power as possible to provide the routing service, while maintaining the target network throughput. The idle network devices and links can be shut down or put into sleep for power saving. We also consider the tradeoff between power conservation and network fault-tolerance. Our algorithm can flexibly adapt the power-aware routing to meet different reliability requirements. In contrast to previous power-aware routing work, our approach uses a different traffic rate model, in which the transfer rate of each flow is bandwidth-constrained and depends on the network resource competition with other flows, instead of given by the fixed traffic demand in advance. This is because network is becoming the bottleneck for data-intensive distributed computation in data centers.

We make the following contributions in the paper. First, we formally establish the model of throughput-guaranteed power-aware routing problem, which will guide us to effectively analyze and solve the problem. We analyze the time complexity of the power-aware routing problem, and prove that it is NP-hard (Section 2.2 and Appendix A).

Second, we propose a throughput-guaranteed power-aware routing algorithm to achieve our design goal (Section 3). The algorithm works in the following four steps: Step 1, we compute the routing paths and corresponding network throughput with all switches and links in the initial topology of the data center network, which are called *basic routing* and *basic throughput* respectively. Step 2, we introduce an iteration process to gradually remove switches from the basic routing and update the initial topology, while satisfying the predefined performance requirement. Step 3, we remove as many links connected to active switches as possible from the updated topology above while meeting the throughput requirement. Step 4, we further adapt the updated topology to meet the reliability requirement. We can power off the switches and links not involved in the finally updated topology, or put them into the sleep mode to conserve power.

Third, we conduct extensive simulations in typical data center networks to validate the effectiveness of our power-aware routing algorithm under different power-saving granularities, traffic patterns and reliability requirement (Section 4). The results show that our power-aware routing algorithm can effectively reduce the power consumption of network devices in data center networks, especially under low network loads.

The rest of the paper is organized as follows. Section 2 introduces background knowledge and related work, and formally establishes the throughput-guaranteed power-aware routing problem model. Section 3 presents our algorithm design. Section 4 evaluates the algorithm through simulations in typical data center networks. Section 5 discusses practical implementation issues of our algorithm and Section 6 concludes the paper.

2. Background and model

In this section, we first introduce the power consumption models as well as the power conservation strategies of network devices in data centers, and then establish the model of throughput-guaranteed power-aware routing problem. Finally, we present the related work on green Internet and data center networks.

2.1. Network power consumption model

We discuss two types of power consumption models applied to current modular network devices in data centers. The first one, called *General Model*, can accurately calculate the power consumption of network devices, but its dependence on unpredictable network traffic conditions increases the complexity of the computation and thus the model is hardly used in practice. To acquire a practical model, we discuss the *Simplified Model* which is a more feasible model. It can easily compute the power consumption of switches only based on their configurations, regardless of the factor of the network traffic. Note that we are not claiming any novelty in the power consumption model of network devices, but discuss here to make the paper more complete. We will use the network power consumption model to guide the design and evaluate the effectiveness