Deadline-aware and Energy-Efficient Dynamic Flow Scheduling in Data Center Network

Zan Yao, Ying Wang, Junhua Ba, Junran Zong, Sixiang Feng, Zhanwei Wu
State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Beijing, China
Email: {yao3617, wangy}@bupt.edu.cn

***Abstract*—The construction of energy-efficient network and achievement of green communication have garnered great attention as a promising a way to reduce network operating costs and C emissions. Moreover, recently the deadline-aware and energy-efficient routing and scheduling algorithms in data center network have been attracting a broad attention. However, the dynamic scheduling for flows has not been explicitly studied by the existing research. In this paper, we investigated the dynamic flow scheduling in data center network, and propose a deadline-aware and energy-efficient dynamic flow scheduling (DEDFS) algorithm, assuming the path of the flow could be calculated in advance and pre-stored. In addition, the number of mouse flows in data center network accounts for main proportion, but consumption is very small. In order to achieve the balance of energy-saved and efficiency, mouse flows will be directly transferred, while elephant flows will be scheduled by the Most-Critical-First static strategy based dynamic scheduling algorithm. It selects the interval of largest energy consumption density as the critical interval, and all of the flows in this critical interval will be preferentially scheduled. Finally, the feasibility and validity of the algorithm are verified by simulation.**

Keywords—data center network; deadline-aware; energy-efficient; dynamic flow scheduling

I. INTRODUCTION

The power consumption of enormous network devices in data center network has emerged as a big concern to data center operators. Moreover, providing high performance is the primary goal in a network, energy-efficiency technology with considering the network performance (e.g. throughput, delay) is important in practice.

SDN have been applied to data center network to enable traffic engineering by provide flexible control with a global network view. Except the equipment-level technology, network-level energy-efficient technology is mainly through the establishment of new routing and scheduling mechanism to achieve energy-efficient [1]. Some research focused on the topology-aware energy-efficient routing, which is not suitable for network-limited flows that require huge bandwidth to finish traffic transmission within deadline [2-6]. To solve the challenge of energy-efficient path and deadline satisfied for network-limited flows. Some research focused on traffic-aware energy-efficient routing methods [7-10]. Ref. [11] taked into account energy efficiency and link utilization and formulated the routing problem as a 0-1 integer-programming problem, and proposed a 2-phase energy efficient flow routing algorithm. The energy-efficient problem of deadline restricted flow is described and modeled for the first time in Ref. [12], and a bandwidth-aware energy-efficient routing algorithm (BEERS) based on SDN controller is proposed. BEERS adopts non-preemptive scheduling and multiple traffics can share the same link at different traffic rates. Because the data center network structure is fixed, the path of the flow usually could be calculated in advance and pre-stored. In order to reduce complexity, we only focus on the dynamic flow scheduling problem in data center network. Ref. [13] proposed a preemptive Most-Critical-First scheduling algorithm for the flow with strict delay desire in data center network, and every flow could exclusively take up the link. It selects the interval with the largest energy consumption density as the critical interval, and all of the flows in this critical interval will be preferentially scheduled by EDF policy. Compared with Ref. [12]. The preemptive and exclusively scheduling algorithm could increase the effectiveness of energy saving. Moreover, it assumed that the switch could automatically adjust the rates of ports, which could affect the power consumption of ports.

However, because only the flows of current moment is known, Ref. [13] assumed that different arrival times of flows have been known and could be unified scheduled, but it is unrealistic. Moreover, the algorithm is static and cannot meet the desire of real-time. Based on the Most-Critical-First static scheduling algorithm, we propose a dynamic scheduling algorithm. Moreover, in order to achieve the balance of energy saving and efficiency, we propose different scheduling strategies for elephant flows and mouse flows [14]. The elephant flows are scheduled by the dynamic scheduling algorithm, while mouse flows are transferred directly.

In summary, we mainly study the deadline-aware and energy-efficient dynamic flow-scheduling problem in data center network. Our contributions are two folds:

(1) We propose a deadline-aware and energy-efficient dynamic flow scheduling (DEDFS) algorithm. Compared with the static scheduling scheme, it is more practical and feasible.

(2) The proposed DEDFS is a scheduling algorithm with traffic classification, which could achieve the balance of energy saving and efficiency.

The rest of the article is arranged as follows, Section Ⅱ describes the power consumption model of data center network and formulates the energy-efficient scheduling problem for flows with deadline guaranteed. The dynamic scheduling mechanism is developed in Section Ⅲ. Simulation results are presented in Section Ⅳ, followed by conclusion in Section Ⅴ.

II. POWER CONSUMPTION MODEL OF DATA CENTER NETWORK AND PROBLEM FORMULATION

A. Power Consumption Model of Data Center Network

The data center network can be modeled as G = {V, E}, where V = {switch}, E = {link}, and it is an undirected graph. The ports of switches and links are the main power consumers, we adopt the power consumption model.The power consumption function $f(x_e)$ is given to uniformly characterize the manner, in which the energy of each link $e \in \varepsilon$ and the ports at the ends of the link is being consumed with respect to the transmission rate x_e [13, 15].

$$f(x_e) = \begin{cases} 0, & x_e = 0 \\ \sigma + \mu x_e^{\alpha}, & 0 \le x_e \le C \end{cases} \tag{1}$$

where the three parameters σ, μ and α are constants associated with the link type. C is the maximum transmission rate of a link. $f(.)$ is superadditive, i.e., $\alpha > 1$.

B. Problem Formulation

Before studying the dynamic scheduling problem of elephant flows, we first study static Deadline-Constrained Flow Scheduling (DCFS) problem in this section [13].

Assume that at the current time $t = t_0$, the set of elephant flows that need to be transmitted is $J(t=t_0) = \left\{ j_{(t_0,1)}, j_{(t_0,2)}, \ldots, j_{(t_0,i)}, \ldots, j_{(t_0,n)} \right\}$. For each flow $j_{(t_0,i)} = \left\{ p_{(t_0,i)}, q_{(t_0,i)}, r_{(t_0,i)}, d_{(t_0,i)}, w_{(t_0,i)}, P_{(t_0,i)} \right\}$, $j_{(t_0,i)} \in J(t_0)$, where $p_{(t_0,i)}, q_{(t_0,i)}$ denotes the source and destination nodes of flow $j_{(t_0,i)}$, respectively; $r_{(t_0,i)}, d_{(t_0,i)}$ represent the start time and the deadline of the flow respectively; and $w_{(t_0,i)}$ means the data size that needs to be transmitted; $P_{(t_0,i)}$ is the set of links through which j_i is routed, and it is calculated by the shortest path principle and is stored.

It has been proved that the minimum energy scheduling is to select a minimum transfer rate for each flow, while ensuring that the deadlines of all flows are satisfied [13]. It is not difficult to prove that in each calculation process, the rate of each flow is constant, so the rate of each flow $j_{(t_0,i)}$ is defined as $s_{(t_0,i)}$, the problem can be modeled as a convex optimization problem:

$$\text{Min} \sum_{e \in \varepsilon_a} \sum_{j_{(t_0,i)} \in J(t_0)} w_{(t_0,i)} \cdot \mu s_{(t_0,i)}^{\alpha-1} \tag{2}$$

Subject to:

$$\sum_{j_i \in J'} \frac{w_{(t_0,i)}}{s_{(t_0,i)}} - \left(\max_{j_i \in J'} d_{(t_0,i)} - t_0 \right) \le 0, J' \subseteq J_e(t_0) \tag{3}$$

$$s_i > 0, \quad \forall j_i \in J \tag{4}$$

The first constraint is that all flows in any subset J' on an arbitrary link must be completed within the deadlines of the flows in Eq. (3). The second constraint is that the transfer rate of each flow must be bigger than zero in Eq. (4). It is easy to prove that objective function in Eq. (2) is a convex function, since all constraints are linear.

III. DEADLINE-AWARE AND ENERGY-EFFICIENT DYNAMIC FLOW SCHEDULING ALGORITHM

In this section, we propose the deadline-aware and energy-efficient dynamic flow scheduling (DEDFS) algorithm with traffic classification to saving energy and satisfying the delay requirement. It considers the characteristics of elephant flow and mouse flow, and adopts different scheduling mechanism. The 20% of the link bandwidth will be reserved for mouse flows, while the elephant flows will be transferred through 80% of the link bandwidth by the dynamic scheduling algorithm[15]. The pseudo code of DEDFS algorithm is shown in as below.

When a flow comes, the algorithm first detects whether it is a mouse one or an elephant one. Then, if it is the former (the size of flow is smaller or equal to a constant a), it will be directly transferred; and if it is the latter, the elephant flow will be added in a flow set $J'(t_0)$ in lines 2-8. Except the new flows, we also need to reconsider the previous flows, which are not transmitted completely. Before adding the ones that are not transmitted completely into the flow set $J'(t_0)$ and it will be needed to

ALGORITHM : DEDFS

Input: network topology $G=(V,E)$, the flow set $J(t)=\{j_{(t,1)}, j_{(t,2)},\dots, j_{(t,n)}\}$, where $j_{(t,i)}=\{p_{(t,i)}, q_{(t,i)}, r_{(t,i)}, d_{(t,i)}, w_{(t,i)}, P_{(t,i)}\}$.

Output: energy consumption function $\phi_g(s)$.

Procedure Minimize $\phi_g(s)$

1. **while** a new flow set $J(t)$ at $t=t_0$ comes **do**
2. **for each** $j_{(t_0,i)} \in J(t_0)$ **do**
3.   **if** $w_{(t_0,i)} \le a$ **then**
4. $j_{(t_0,i)}$ is directly transferred by rate $w_{(t_0,i)}$;
5.   **else**
6. $J'(t_0) \leftarrow j_{(t_0,i)}$;
7.   **end if**
8. **end for**
9. **for** $\forall j_{(t',i)} \in J(t)$, $t' < t_0$ **do**
10.   **if** $\int_{t'}^{t_0} s_{(t',i)} dt \le w_{(t',i)}$ **then**
11. $w_{(t',i)} \leftarrow w_{(t',i)} - \int_{t'}^{t} s_{(t',i)} dt$;
12. $J'(t_0) \leftarrow j_{(t',i)}$;
13.   **end if**
14. **end for**
15. Schedule $J'(t_0)$ by Most-Critical-First algorithm and obtain the transmission rate of every elephant flows at certain interval.
16. **end while**

end procedure

modify their sizes and the beginning times in lines 9-14. Then we schedule $J'(t_0)$ by Most-Critical-First algorithm, and obtain the transmission rate of elephant flows in line 15. Considering the deadline-aware and energy-efficient static scheduling of elephant flows in a specific time is a convex optimization problem, the Most-Critical-First algorithm is used to schedule elephant flows [13]. It is not difficult to see that every elephant flow is likely to be transferred in a few different intervals and the transmission rate may vary in different intervals.

In summary, the DEDFS algorithm with traffic classification not only achieves the energy-efficient targets but also satisfies the performance of both elephant flows and mouse flows. In addition, it could improve the computational efficiency.

IV. SIMULATION

To verify the effectiveness of the proposed dynamic flow scheduling DEDFS algorithm, simulation is conducted in Fat-Tree and Bcube data center network. We build a simulator in JAVA.

A. Simulation Environment

This work selects the commonly used Fat-Tree data center network topology which consists 20 four-point switches, 16 hosts and 48 links, and BCube network topology built by 12 four-point switches, 8 hosts and 12 links. We choose to schedule the flows in total time interval [1,300] ms. The starting times and the deadlines of flows follow a random distribution between 1 and 300. The number of flows varies from 8, 16, 32, and 64, among which 80% are mouse flows, 20% are elephant flows. The size of mouse flows and elephant flows follow the uniform distribution ranging in the [0.1, 3] MB and [40, 200] MB, respectively. The α in the energy consumption function is set 2, and μ is set 1, and sufficient instances are simulated and results are reported below.

B. Simulation Results

First, we verify the performance of our proposed DEDFS scheduling algorithm at different topologies and parameter values by comparing with the DEDFS algorithm without traffic classification and static Most-Critical-First scheduling algorithm. The results in shown in Fig.1-Fig.2.

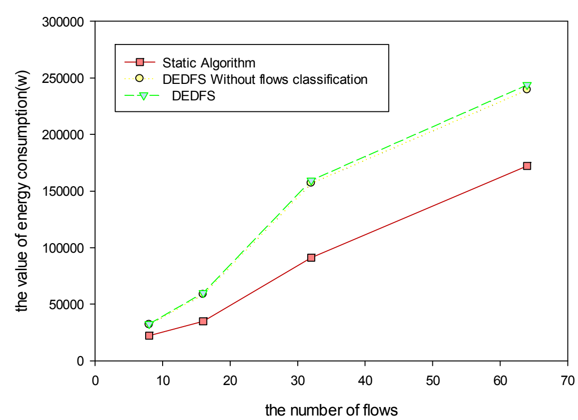

Fig.1 The energy-efficient performance in Fat-Tree with α=2

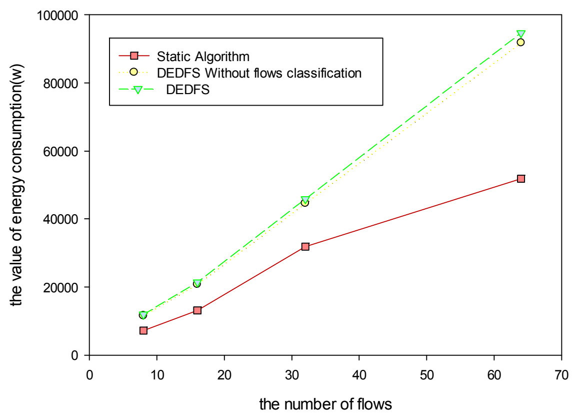

Fig.2 The energy-efficient performance in BCub with $\alpha=2$

As shown in Fig.1-Fig.2, the energy consumption of our proposed DEDFS algorithm is slightly above than static Most-Critical-First scheduling algorithm. Because static Most-Critical-First scheduling algorithm knows the information of all flows in advance, then could make better scheduling.

In addition, our proposed DEDFS algorithm is compared with the proposed dynamic scheduling algorithm without traffic classification. As the shown in Fig.1-Fig.2, the computational complexity could be greatly reduced, but there is not much difference between the energy consumption of both algorithms.

In summary, although the energy-saving performance of our proposed DEDFS algorithm is not better than static Most-Critical-First scheduling algorithm, but it does not need the information of all flows in entire intervals in advance. What is more, the computational efficiency is greatly improved through the classification of flows.

V. Conclusions and Future Work

In this paper, we have proposed a dynamic deadline-aware and energy-efficient flow scheduling algorithm. It not only satisfies energy-efficient and delay performance desires of all flows, but also is practical and feasible. In addition, we select different scheduling schemes for elephant flows and mouse flows to obtain the balance of the effectiveness and performance. Simulations have also shown that the proposed DEDFS algorithm is feasible and effective. In the future, we will research how to design the route and scheduling together for saving energy with the certain QoS guaranteed.

Acknowledgment

This work was supported by the National High Technology Research and Development Program of China (2013AA013502) and the National Natural Science Foundation of China (61501044).

References

[1] Yunfei Shang. Research on Power-aware Routing in Data Center Networks [D]. Beijing: Tsinghua University, 2013.

[2] Zhang M, Yi C, Liu B, et al. GreenTE: Power-aware traffic engineering[C]//Pro of the 18th IEEE Int Conf on Network Protocols(ICNP'10). Piscataway, NJ: IEEE, 2010:21-30.

[3] Cianfrani A, Eramo V, Listanti M, et al. An energy saving routing algorithm for a green OSPF protocol[C] //Proc of the 29th IEEE INFOCOM 2010 Workshops. Piscataway NJ: IEEE, 2010: 1-5.

[4] Cianfrani A, Eramo V, Listanti M, et al. An OSPF enhancement for energy saving in IP networks[C] //Proc of the 30th IEEE INFOCOM 2011 Workshops. Piscataway, NJ: IEEE,2011:325-330.

[5] Cuomo F, Cianfrani A, Polverini M, et al. Network pruning for energy saving in the Internet [J]. Computer Networks, 2012, 56(10):2355-2367.

[6] Wang L, Zhang F, Arjona A J, et al. GreenDCN: A general framework for achieving energy efficiency in data center networks[J]. IEEE Journal on Selected Areas in Communications, 2014,32(1):4-15.

[7] Chabarek J, Sommers J, Barford P, et al. Power awareness in network design and routing [C] //Proc of the 27th IEEE INFOCOM'08. Piscataway, NJ: IEEE, 2008:457-465.

[8] Gelenbe E, Silvestri S. Optimisation of power consumption in Heterrogeneous Networks, 2009, 22(12):717-729.

[9] Vasic N, Kostic D. Energy-aware traffic engineering[C] //Proc of the 1st Int Conf on Energy-Efficient Computing and Networking. New York: ACM, 2010: 169-178.

[10] Xu M, Shang Y, Li D, et al. Greening data center networks with throughput-guaranteed power-aware routing [J]. Computer Networks, 2013, 57(15):2880-2899.

[11] Qianqian Zhang, Xiaoyong Zhang, Jun Peng , et al. Energy Optimization by Flow Routing Algorithm in Data Center Network Satisfying Deadline Requirement , Asia-Pacific Services Computing Conference APSCC 2016: Advances in Services Computing pp 117-129.

[12] Xu G, Dai B, Huang B, et al. Bandwidth-aware energy efficient flow scheduling with SDN in data center networks[J]. Future Generation Computer Systems, 2016, 68:163–174.

[13] Wang L, Zhang F, Zheng K, et al. Energy-Efficient Flow Scheduling and Routing with Hard Deadlines in Data Center Networks , 2014 IEEE 34th International Conference on Distributed Computing Systems. 2014:248-257.

[14] Benson T, Akella A, Maltz D A. Network traffic characteristics of data centers in the wild[C]. Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM, 2010: 267-280.

[15] M. Andrews, A. F. Anta, L. Zhang, and W. Zhao, "Routing for power minimization in the speed scaling model," IEEE/ACM Trans. Netw.,2012,20(1):285–294.