Computer Networks 138 (2018) 149–163

Contents lists available at ScienceDirect

Computer Networks

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

NEMO: Novel and efficient multicast routing schemes for Hybrid Data Center Networks

Xiaofeng Gao[a,*], Tao Chen[a], Zongchen Chen[b], Guihai Chen[a]

[a] Shanghai Key Laboratory of Scalable Computing and Systems,Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

[b] Algorithms, Combinatorics, and Optimization, Georgia Institute of Technology, Atlanta, USA

ARTICLE INFO

Article history:
Received 6 September 2017
Revised 10 March 2018
Accepted 26 March 2018
Available online 4 April 2018

Keywords:
Data center network
Multicast routing
60 GHz wireless

ABSTRACT

Multicast places an important role in group communication in data center networks (DCN's). As the rapid development of DCN's, researchers elegantly integrate wireless connections with wired DCN to form a new structure, Hybrid Data Center Network (HDCN), which can improve the performance of multicast significantly with wireless gains. In this paper, we present two novel and efficient multicast routing schemes (named *NEMO-Group* and *NEMO-Cluster*) for HDCN to achieve fast multicast tree construction and simple network traffic routing. We theoretically analyze the expected performance improvement of the two designs with different wireless transmission ranges. To the best of our knowledge, we are the first to discuss the sender-driven polynomial-time multicast routing schemes for HDCN, and to estimate the upper and lower bound of wireless gains. Moreover, our numerical results show that with the same construction cost, NEMO's reduce 15%–40% traffic and improve 10%–40% throughput than the existing multicast protocols.

1. Introduction

Group communication has become increasingly popular in the last decade, such as VoIP, multiplayer online games, Internet live video, etc. As the rapid development of cloud computing, the aforementioned applications have been widely deployed in the cloud and been supported by many data centers [1,2]. Moreover, group communication is often used for serializable sessions/data synchronization, file chunk replication, executive binary delivery in map-reduce alike distributed computation, as well as OS/application upgrading in data centers [3,4]. Noticing its importance, Li et al. [5,6] demonstrated that the traditional Internet-oriented multicast technologies are not appropriate to form multicast trees in Data Center Networks (DCN's), since DCN's contain multiple equal-cost paths and low-end commodity switches, bringing difficulties to receiver-driven multicast schemes, whose objective is to avoid the problem of link wastage and memory shortage for storing routing tables. They thereby constructed a sender-driven network-level multicast scheme named ESM, saving 40%–50% network traffic and doubling the application throughputs.

After the raise of 60 GHz wireless technologies in data center environment [7], researchers implement wireless connections into wired DCN to form a new structure, Hybrid Data Center Network (HDCN), to achieve higher bisection bandwidth and overall throughput. Till now, HDCN has become a popular design and future trend for modern DCN architecture [8–11]. However, Li's scheme does not fit the property of HDCN well due to lack of consideration of the wireless connectivity and the localized topological benefits. Consequently, a new multicast scheme is desiderative to overcome the weakness of Internet-oriented multicast technologies and utilize the architectural features of HDCN.

In this paper, we present two novel and efficient multicast routing schemes named *NEMO-Group* and *NEMO-Cluster* for HDCN to achieve fast multicast tree construction and simple traffic routing. We theoretically analyze the expected performance of NEMO with respect to the wireless topology of HDCN, demonstrating an efficient physical placement of HDCN to improve the efficiency of NEMO. We also prove upper and lower bounds of performance improvement with wireless gains, since the newly inserted wireless topologies have localized characteristics. Furthermore, we illustrate by simulations that with the same construction cost, NEMO reduces traffic by 15%-40% and improve 10%-40% throughput compared to ESM. In all, our contributions are listed as follows:

- To the best of our knowledge, we are the first to discuss the sender-driven polynomial-time multicast routing scheme for HDCN, which has both theoretical and practical significance.
- NEMO's are efficient and dynamic multicast schemes for HDCN,

* Corresponding author.
E-mail addresses: gao-xf@cs.sjtu.edu.cn, gaoxiaofeng@sjtu.edu.cn (X. Gao).

https://doi.org/10.1016/j.comnet.2018.03.028
1389-1286/

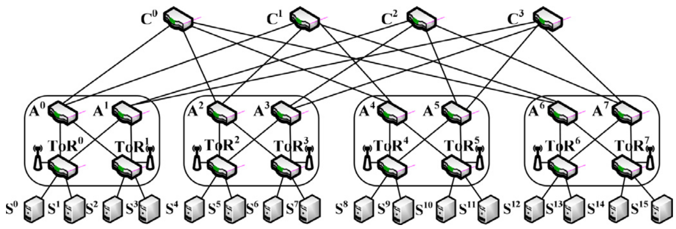

Fig. 1. An example of an HDCN with Fat-Tree topology (4 pods).

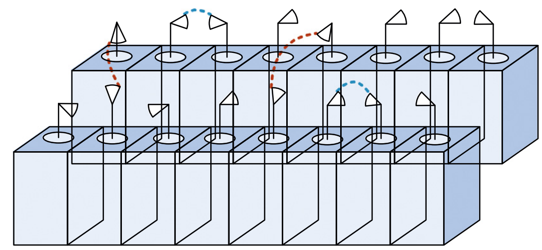

Fig. 2. An example of wireless connections in an HDCN.

which keep the benefits of ESM and utilize the feature of wireless connections. Simulation results show that NEMO's can save around 15%–40% traffic and improve 10%–40% throughput than ESM under different scenarios.

- We provide theoretical analysis and practical tests to depict the relation among different parameters in HDCN, which may influence the performance of NEMO.
- We estimate the expected performance improvement if we append a new wireless topology onto DCN. Such estimation depicts an upper and a lower bound of wireless gains, reflecting the capacity and limitation of wireless technologies to multicast applications. It can be a baseline to justify and evaluate the benefits of wireless gains.

The rest of our paper is organized as follows. Section 2 describes the problem formulations and system architecture. Section 3 proposes the construction of NEMO-Group and NEMO-Cluster algorithms. Section 4 analyzes the performance of NEMO's theoretically, and provides the expected upper and lower bounds of wireless gains. We also discuss the updating strategies and the implementation on other topologies for NEMO's in Section 5. Section 6 illustrates further optimization by rack rearrangement. Section 7 evaluates the performance of NEMO with ESM by mass number of simulations. Section 8 summarizes related work. Finally, Section 9 gives the conclusion and future work.

2. Problem formulation

2.1. System architecture

A Hybrid Data Center Network (HDCN) is a typical wired DCN plus high-throughput, 2D/3D beamforming wireless links in the 60 GHz band. In a common wireless deployment, wireless radios with horn antennas are placed on top of each rack to connect pairs of top-of-rack (ToR) switches [11]. We first take a Fat-Tree topology [12] as the underlying wired DCN topology, and later introduce the generalization to other wired topologies in Section 5.

For a Fat-Tree structure, we place a 60 GHz antenna with 8-element Phocus arrays (can be switched within 250 μs [10]) on each ToR switch to achieve complete local wireless connections. Both the wired and wireless paths are controlled by a centralized multicast manager, which is also in charge of the multicast address allocations, the multicast tree construction, and the routing configuration distributions [13]. Data center is an environment with high controllability and manageability, and this sort of centralized controller have been widely used across switch-centric and server-centric data center networks, such as Fat-Tree [12], VL2 [14], Aspen tree [15] and BCube [16]. For larger-scale DCN, a group of servers can be adopted to work cooperatively, such as the logically centralized controllers in Software-Defined Networks [17] and the directory servers in VL2 [14]. Flyway system [10] also adopts a central controller to schedule the wireless links between ToR switches.

Fig. 1 is an example of HDCN with 16 servers (S^0–S^{15}), 8 ToR switches (ToR^0–ToR^7, each connected with an antenna, the same below in this paper), 8 aggregation switches (A^0–A^7), 4 pods

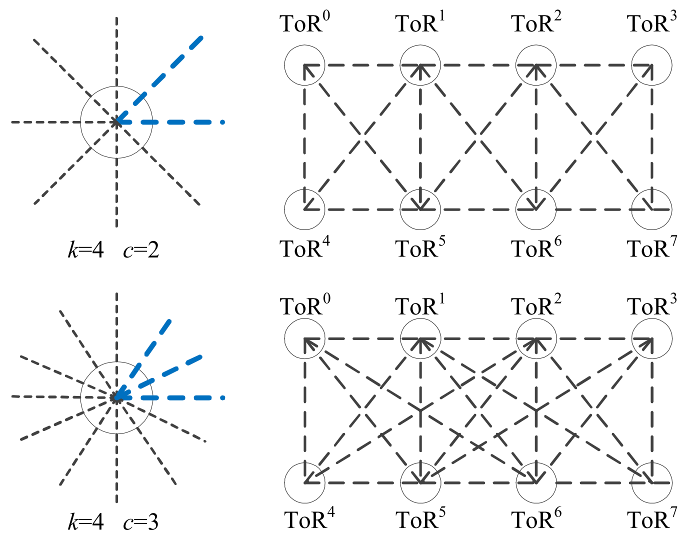

Fig. 3. An illustration of c.

(2 ToR's connected with 2 aggregation switches), and 4 core switches (C^0–C^3).

Due to the limitation of wireless technologies, current indoor 60 GHz technologies can achieve a stable propagation range of nearly 10 m with directional antennas [10]. By adjusting the directions of antennas electronically or with rotators, adjacent racks or racks between two rows can be connected flexibly on demand by wireless links without any obstacles. Fig. 2 is an example of wireless connections in a DCN, in which dashed lines represent connections between ToR's in different rows and connections between neighbor ToR's in the same row. Correspondingly, we assume a regular physical placement for racks in a data center, in which a k pod Fat-Tree is placed in $\frac{k}{2}$ rows, with k racks in each row.

We define a parameter c to denote the strength of wireless connections, i.e., the number of connectable ToR's for a ToR by wireless connections. If we consider a ToR as the origin of a coordinate axis, then c represents the number of ToR's it can connect wireless links in the first quadrant (regardless of the axis and the marginal effect). An simple illustration of c is shown in Fig. 3, which draws two wireless connections with respect to the Fat-Tree topology in Fig. 1. The left upper figure shows the first quadrant with $c = 2$, in which two bold dashed lines are highlighted. Correspondingly, the right upper figure shows the wireless subnetwork of $c = 2$ with boundary effects. Comparatively, the bottom two figures show the situation of $c = 3$. If we take ToR^5 as an example, when $c = 2$, it can connect two links (ToR^5, ToR^2), (ToR^5 ,ToR^6) in the first quadrant. When $c = 3$, it can connect three links (ToR^5, ToR^2), (ToR^5 ,ToR^3) and (ToR^5 ,ToR^6) in the first quadrant.

Appending wireless connections into a DCN can reduce the network traffic significantly. For instance, in Fig. 1, if server S^9 wants to communicate with server S^2, one possible path is $S^9 \rightarrow ToR^4 \rightarrow A^5 \rightarrow C^2 \rightarrow A^1 \rightarrow ToR^1 \rightarrow S^2$. However, we can immedi-