



HAL
open science

Software-defined Adaptive Resource Management for Cloud-hosted Group Communication Applications

Prithviraj Patil, Akram Hakiri, Aniruddha Gokhale

► **To cite this version:**

Prithviraj Patil, Akram Hakiri, Aniruddha Gokhale. Software-defined Adaptive Resource Management for Cloud-hosted Group Communication Applications. DEBS '17 , Jun 2017, Barcelone, Spain. 10.1145/3093742.3095087 . hal-01633335

HAL Id: hal-01633335

<https://hal.science/hal-01633335>

Submitted on 12 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Poster: Software-defined Adaptive Resource Management for Cloud-hosted Group Communication Applications

Prithviraj Patil*
Vanderbilt University
Nashville, Tennessee, USA 37235
prithviraj6116@gmail.com

Akram Hakiri
ISSAT Mateur, Univ de Carthage
Mateur, Bizerte, Tunisia 7050
akram.hakiri@gmail.com

Aniruddha Gokhale
Dept of EECS, Vanderbilt University
Nashville, Tennessee, USA 37235
a.gokhale@vanderbilt.edu

ABSTRACT

Many cloud-hosted applications and cloud infrastructure management services with publish/subscribe semantics rely heavily on the use of group communications. Although multicast is useful for efficient group communications, traditional IP multicast has seen very low adoption in cloud data center networks (DCNs) due to issues with its stability and scalability. The introduction of Software-defined Networking (SDN) in DCNs has provided new opportunities for re-engineering and effectively utilizing multicast capabilities that can overcome existing impediments to the adoption of IP multicast. To that end this paper presents an adaptive and flexible middleware solution called SDN-based Multicast (SDMC), which provides group communication capabilities in DCNs in a way that is both network load-aware and switch memory-efficient. Applications using SDMC remain agnostic to the underlying group communication semantics, which SDMC provides efficiently by dynamically adapting between unicast and multicast in accordance with changing network bandwidth and switch memory usage.

CCS CONCEPTS

•Networks →Network resources allocation; Cloud computing; Programmable networks; Data center networks; •Computer systems organization →Cloud computing; •Computer systems organization →Fault-tolerant network topologies;

KEYWORDS

Group Communications, Unicast/Multicast, SDN, Adaptive, Cloud

ACM Reference format:

Prithviraj Patil, Akram Hakiri, and Aniruddha Gokhale. 2017. Poster: Software-defined Adaptive Resource Management for Cloud-hosted Group Communication Applications. In *Proceedings of ACM International Conference on Distributed and Event-based Systems, Barcelona, Spain, June 19–23, 2017 (DEBS’17)*, 2 pages.
DOI: 10.1145/3093742.3095087

1 INTRODUCTION

Group communication semantics play a key role in cloud data center networks (DCNs) for a variety of cloud-hosted applications and

*Work performed by the first author while being a graduate student.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

DEBS’17, Barcelona, Spain

© 2017 Copyright held by the owner/author(s). 978-1-4503-5065-5/17/06...\$15.00
DOI: 10.1145/3093742.3095087

infrastructure-level tasks that inherently follow the event-based publish/subscribe communication model, which in turn requires one-to-many or many-to-many communication semantics. Although IP multicast (IPMC) is a candidate technology to implement group communications in DCNs, it incurs several stability and scalability problems [1], which calls for a prudent and careful solution for deployment in DCNs. Unfortunately, any such effort has shown to have required substantial manual efforts on the part of network operators [5]. A further downside with multicast concerns its address structure, which cannot be hierarchically aggregated and hence the reconstruction of multicast trees becomes an expensive process in terms of network overhead [7].

Apart from the known IPMC limitations, two key considerations dictate the performance of group communications in DCNs. First, the router/switch’s available network capacity (i.e., available bandwidth) is required to estimate the bandwidth available for data traffic and control messages that can be forwarded along their multiple ports concurrently [8]. Hence, the ability to instrument the router’s (or switch’s) network capacity is crucial to estimating the network load in terms of the overall bandwidth a router/switch is able to support. Further, since the memory of the routers/switches (including their queuing buffers) holds packets and connection state information of the traffic transiting across the network devices, it is important to maintain the router/switch memory utilization under a given threshold to avoid buffer overflows. Software-Defined Networking (SDN) [4] is a promising technology to address many of these problems.

Although recent efforts [2, 3] have used SDN for efficient bandwidth management during the creation of multicast trees by using topology information in DCNs, these approaches cannot adapt to the network load and router (or SDN switch) memory utilization. To that end we propose SDN-based Multicast (SDMC) for flexible, network load-aware, and switch memory-efficient group communications in DCNs. SDMC is a middleware that intelligently uses a combination of unicast and software-defined multicast, toggling between them at runtime to adapt to load while ensuring that applications remain transparent to these adaptive changes. Thus, SDMC effectively manages the network bandwidth and switch memory utilization while providing scalable and efficient group communications in SDN-based DCNs thereby enabling a large number of applications and services to avail of our adaptive resource management solution.

2 SDMC DESIGN

Figure 1 depicts the architecture of our SDN-based multicast approach called SDMC. At the control plane level, a (logically) centralized SDN controller communicates with the SDN routers using

the OpenFlow channel. In the data or forwarding plane, these routers are connected to host machines (physical or virtual) that host SDMC's SDN middleware. The SDMC logic is made available in a distributed manner as follows: the core intelligence of SDMC is made available as a suite of SDN applications (called SDN NetApps in SDN parlance) that execute inside the controller as shown in the figure; and as a SDN middleware layer available on each host where the senders and receivers use the APIs provided by SDMC. The use of the SDN principles allows SDMC to be flexible and dynamic.

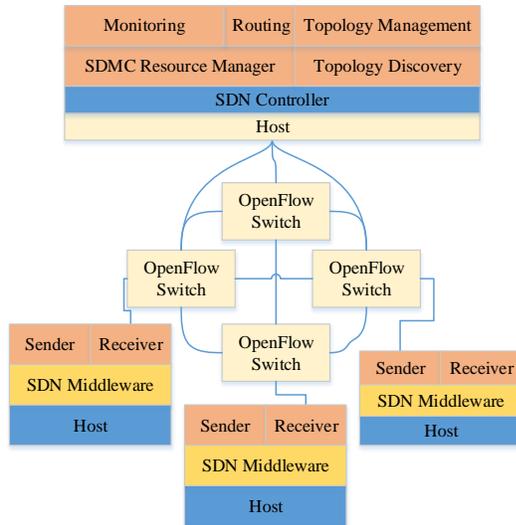


Figure 1: SDMC Architecture

The primary elements of the SDMC architecture include:

SDN controller: The DCN is managed by a (logically) centralized SDN controller, which is placed on a dedicated machine(s) with dedicated out-of-band connections to all the OpenFlow-enabled routers. The controller hosts several network applications (SDN NetApps) which form the bulk of our SDMC contributions. For instance, the topology discovery supports automatic discovery of joining and leaving routers and hosts. The topology management serves to selectively (de)activate the routers and switches. Monitoring is needed to gather runtime statistics of the system. The routing module supports the adaptive multicast intelligence of SDMC. The resource manager keeps track of the resource utilization.

OpenFlow-enabled switches: In a typical DCN, a number of OpenFlow-enabled switches are connected to form a network with topologies like mesh, tree or jellyfish. Each switch contains an OpenFlow client to connect to the SDN controller.

Host machines and SDN middleware: Multicast functionalities, which deal with overlay multicast, are implemented using SDMC's SDN middleware that runs in the host machines. The SDN middleware controls the communication using a host manager service to translate endpoints listening on multicast IDs into multiple unicast IDs, and switch endpoints from unicast to multicast and vice-versa. Those endpoints can decide whether senders or receivers are able to send or listen only on SDMC IDs or not.

Host manager: The host manager module (not shown) keeps track of all machines connected to the network. This application

is used by SDMC to communicate with the SDN middleware of the host machines. This application is required since we build a hybrid multicast protocol by combining application-level multicast (or overlay multicast) and native network-level-multicast.

SDMC participants: The SDMC participants (senders and receivers applications) run on top of the SDN middleware. The SDN middleware hides the adaptive decisions made by SDMC from these user applications.

3 CONCLUSIONS

This paper presented the design of SDN-enabled multicast solution for flexible, network load-aware, switch memory-efficient group communications in data center networks. Our middleware called SDMC uses a combination of unicast and software-defined multicast, and dynamically switches between them while ensuring that applications remain agnostic to the adaptation yet providing superior performance over individual cases. SDMC is a work-in-progress project and several new features are planned. For example, our ongoing work is evaluating SDMC for its scalability and comparing against related research efforts, e.g., [3, 6]. Moreover, SDMC optimization decisions are currently made on a per-switch basis. In future, we seek to provide a system-wide optimization that balances the resource usages across multiple competing flows. SDMC is available in open source and its existing code base is available at <https://github.com/prithviraj6116/sdmc>.

ACKNOWLEDGMENTS

This work is supported in part by NSF US Ignite CNS 1531079. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF.

REFERENCES

- [1] Dmitry Basin, Ken Birman, Idit Keidar, and Ymir Vigfusson. 2010. Sources of Instability in Data Center Multicast. In *Proceedings of the 4th International Workshop on Large Scale Distributed Systems and Middleware (LADIS '10)*. 32–37.
- [2] W. Cui and C. Qian. 2015. Scalable and Load-Balanced Data Center Multicast. In *2015 IEEE Global Communications Conference (GLOBECOM)*. 1–6.
- [3] A. Iyer, P. Kumar, and V. Mann. 2014. Avalanche: Data center Multicast using software defined networking. In *2014 Sixth International Conference on Communication Systems and Networks (COMSNETS)*. 1–8.
- [4] D. Kreutz, F.M.V. Ramos, P. Esteves Verissimo, C. Esteve Rothenberg, S. Azodolmoly, and S. Uhlig. 2015. Software-Defined Networking: A Comprehensive Survey. *Proc. IEEE* 103, 1 (2015), 14–76.
- [5] Xiaozhou Li and Michael J. Freedman. 2013. Scaling IP Multicast on Datacenter Topologies. In *Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies (CoNEXT '13)*. 61–72.
- [6] Cesar AC Marcondes, Tiago PC Santos, Arthur P Godoy, Caio C Viel, and Cesar AC Teixeira. 2012. CastFlow: Clean-slate multicast approach using in-advance path processing in programmable networks. In *Computers and Communications (ISCC), 2012 IEEE Symposium on*. IEEE, 000094–000101.
- [7] Ymir Vigfusson, Hussam Abu-Libdeh, Mahesh Balakrishnan, Ken Birman, Robert Burgess, Gregory Chockler, Haoyuan Li, and Yoav Tock. 2010. Dr. Multicast: Rx for Data Center Communication Scalability. In *Proceedings of the 5th European Conference on Computer Systems (EuroSys '10)*. 349–362.
- [8] Han Wang, Ki Suh Lee, Erluo Li, Chiu Lin Lim, Ao Tang, and Hakim Weatherspoon. 2014. Timing is Everything: Accurate, Minimum Overhead, Available Bandwidth Estimation in High-speed Wired Networks. In *Proceedings of the 2014 Conference on Internet Measurement Conference (IMC '14)*. 407–420.