Towards Software-defined and Self-Driving Cloud Infrastructure*

Extended Abstract

Wei Xu
Institute for Interdisciplinary Information Sciences
Tsinghua University, Beijng, China
weixu@tsinghua.edu.cn

CCS CONCEPTS

• **Computer systems organization → Reliability**; *Maintainability and maintenance*;

Traditionally, people abstract away the infrastructure operation, such as power management, network traffic engineering and even the "cloud computing" layers from software developers. This abstraction brings easier application development and maintenance but leads to complexities and inefficiencies for infrastructure operation. Now in the data center community, people have adopted the software-defined paradigm, hoping to bring more flexibility to data center infrastructure, to improve the performance, efficiency, and reliability of the resource-demanding applications.

There are three challenges to design such an infrastructure: 1) how we can improve the flexibility of traditional data center hardware components, especially those with less open protocols and control interfaces, such as power and the networking; 2) how we take advantage of the flexibility to improve software application performance; and 3) how can we debug a system consisting of lots of software and hardware components from different vendors.

In this talk, we present our recent research into this area. Using data center power management, networking as well as cloud software management as examples.

First, we show that we can optimize data center power system, the least software-definable components, purely in software layers. Using feedback control models, we can keep the total power of a data center under the budget and brings zero performance disturbance to the existing jobs. Also, we design an A-B testing framework in data centers, making it possible to evaluate large scale scheduling algorithms in real production clusters. [3]

Second, we extend the popular software-defined networking (SDN) techniques into the physical layer, using flexible optical networking equipment. Using a combination of off-the-shelf Ethernet switches and custom designed wavelength switches, we demonstrated that with an online, full-stack network optimization algorithm, we could take full advantage of the physical-layer flexibility, and reduce the tail latency in data centers [4] and improve data transfer performance on wide-area networks [1].

Third, also for data center networks, we push the software-defined design to an extreme, by moving all the states from network layer into the server software. In our system, DumbNet [2], the switches have no tables, no state, and thus require no configurations. We push almost all control plane functions to hosts: they determine the entire path of a packet. We design a set of host-based mechanisms to make the new architecture viable, from network bootstrapping and topology maintenance to network routing and failure handling. Extensive evaluations show that DumbNet achieves performance comparable to traditional networks, supports application-specific extensions like flowlet-based traffic engineering, and stays extremely simple and easy-to-manage.

Last but not least, layers of software abstractions in a large scale distributed system lead to situations that often confuses administrators, even experienced ones. Using OpenStack, a popular cloud computing software as an example, we show that we can organize system states and events data into a knowledge graph, and turn ad hoc system debugging process into a coherent graph traversal and even support natural-language-based queries [5].

Looking forward, we believe the design paradigm for datacenter infrastructure and applications are shifting with the rapid development of software-defined technologies. Infrastructure levels designs will become more and more agile, and it calls for new tools for developing and maintaining such new architectures.

Acknowledgments

This research is supported in part by the National Natural Science Foundation of China (NSFC) grant 61532001, Tsinghua Initiative Research Program Grant 20151080475 and gift funds from Huawei and Ant Financial.

REFERENCES

[1] Xin Jin, Yiran Li, Da Wei, Siming Li, Jie Gao, Lei Xu, Guangzhi Li, Wei Xu, and Jennifer Rexford. 2016. Optimizing bulk transfers with software-defined optical WAN. In *Proceedings of the 2016 ACM SIGCOMM Conference*. ACM, 87–100.
[2] Yiran Li, Da Wei, Xiaoqi Chen, Ziheng Song, Ruihan Wu, Yuxing Li, Xin Jin, and Wei Xu. 2018. DumbNet: A Smart Data Center Network Fabric with Dumb Switches. In *Proceedings of the Twelfth European Conference on Computer Systems*. ACM.
[3] Guosai Wang, Shuhao Wang, Bing Luo, Weisong Shi, Yinghang Zhu, Wenjun Yang, Dianming Hu, Longbo Huang, Xin Jin, and Wei Xu. 2016. Increasing large-scale data center capacity by statistical power control. In *Proceedings of the Eleventh European Conference on Computer Systems*. ACM, 8.
[4] Da Wei, Lei Xu, Xin Jin, Yiran Li, and Wei Xu. 2016. A 12-rack, 180-server datacenter network (DCN) using multiwavelength optical switching and full stack optimization. In *Optical Fiber Communication Conference*. Optical Society of America, Th5B–6.
[5] Yong Xiang, Hu Li, Sen Wang, Charley Peter Chen, and Wei Xu. 2016. Debugging openstack problems using a state graph approach. In *Proceedings of the 7th ACM SIGOPS Asia-Pacific Workshop on Systems*. ACM, 13.