# Content Distribution for Mobile Internet: A Cloud-based Approach

Zhenhua Li · Yafei Dai · Guihai Chen
Yunhao Liu

# Content Distribution for Mobile Internet: A Cloud-based Approach

Springer

Zhenhua Li
Tsinghua University
Beijing
China

Guihai Chen
Shanghai Jiao Tong University
Shanghai
China

Yafei Dai
Peking University
Beijing
China

Yunhao Liu
Tsinghua University
Beijing
China

ISBN 978-981-10-1462-8 ISBN 978-981-10-1463-5 (eBook)
DOI 10.1007/978-981-10-1463-5

Library of Congress Control Number: 2016943326



Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

# Preface

*Content distribution* (also known as *content delivery*) is the most fundamental function of the Internet, i.e., distributing *digital content* from one *node* to another node or multiple nodes. Here digital content includes webpage, image, software, audio, video, and so on; a node can be a large server cluster, a personal computer, a smartphone, a tiny sensor, *etc.* Typical content distribution systems include Akamai, BitTorrent, eMule, Skype, Xunlei, QQXuanfeng, and so forth. Since Amazon's launch of Elastic Compute Cloud (EC2) in 2006 and Apple's release of iPhone in 2007, Internet content distribution has illustrated a strong trend of *polarization*. On one hand, great fortune has been invested in building heavyweight and integrated data centers, in order to achieve the economies of scale and the cost efficiency for content distribution. On the other hand, end user devices have been growing more lightweight, mobile, and heterogeneous, thus posing rigorous requirements on the traffic usage, energy consumption, speed, and latency of content distribution.

Through comprehensive real-world measurements, we observe that existing content distribution techniques often exhibit poor performance under the settings of cloud computing and mobile Internet. To address the issue, this book investigates content distribution for mobile Internet with a cloud-based approach, by designing novel traffic-saving, energy-efficient, high-speed, and delay-tolerant content distribution techniques and frameworks that automatically adapt to mobile scenarios.

The major content of this book is organized in six parts, which are further elaborated into ten chapters. Specifically, we start with the background and overview in Part I. Then, since cellular traffic is the most expensive among all Internet traffic, its cloud-based optimization is first explored in Part II. Next, video content dominates the majority of Internet traffic, whose delivery deserves deep investigation in Part III. Moreover, P2P content distribution incurs little infrastructure cost and can scale well with the user base; however, its working efficacy can be poor and unpredictable without the assistance of cloud platforms, which is carefully addressed in Part IV. In addition, as an advanced paradigm of content distribution, cloud storage services like Dropbox and Google Drive have quickly gained

enormous popularity in recent years, which are widely studied in Part V. At last, we summarize the major research contributions and discuss the future work in Part VI.

To be more specific, we summarize the main body of the book as follows:

Part II *Cloud-based Cellular Traffic Optimization.* As the penetration of 3G/4G/5G data networks, cellular traffic optimization has been a common desire of both cellular users and carriers. Together with the Baidu PhoneGuard team, we design and deploy TrafficGuard, a third-party mobile traffic proxy widely used by over 10 million Android devices (Chap. 2). TrafficGuard effectively reduces cellular traffic using a network-layer virtual private network (VPN) that connects a client-side proxy to a centralized traffic processing cloud. Most importantly, it works transparently across heterogeneous apps, so it is not constrained to any specific app.

Part III *Cloud-based Mobile Video Distribution.* Driven by the special requirements of mobile devices on video content distribution, we measure and analyze the industrial "cloud downloading" (Chap. 3), "cloud transcoding" (Chap. 4), and "offline downloading" (Chap. 5) services based on the Tencent Xuanfeng system and popular smart home routers. In particular, we diagnose their respective performance bottlenecks and propose the corresponding optimization schemes.

Part IV *Cloud-assisted P2P Content Distribution.* Through large-scale measurements and analysis of industrial cloud-assisted peer-to-peer (P2P) systems like QQXuanfeng and Xunlei, we extract the basic model of "cloud tracking" content distribution in Chap. 6. Further, we design the "cloud bandwidth scheduling" algorithm to maximize the cloud bandwidth multiplier effect in Chap. 7.

Part V *Cloud Storage-oriented Content Distribution.*We are the first to discover the "traffic overuse problem" that pervasively exists in today's cloud storage services. Also, we propose and implement a variety of algorithms to address the problem, such as BDS—batched data sync, IDS—incremental data sync, ASD—adaptive sync defer (Chap. 8), and UDS—update-batched delayed sync (Chap. 9).

In this book, we provide a series of useful takeaways and easy-to-follow experiences to the researchers and developers working on mobile Internet and cloud computing/storage. Additionally, we have built an educational and experimental cloud computing platform (http://www.thucloud.com) to benefit the readers. On top of this platform, the readers can monitor (virtual) cloud servers, accelerate web content distribution, explore the potentials of offline downloading, acquire free cloud storage space, and so forth. Should you have any questions or suggestions, please contact the four authors via *lizhenhua1983@gmail.com*, *dyf@pku.edu.cn*, *gchen@cs.sjtu.edu.cn*, and *yunhaoliu@gmail.com*.

Beijing, China
May 2016

Zhenhua Li
Yafei Dai
Guihai Chen
Yunhao Liu

# Acknowledgments

First of all, we deeply appreciate the Chinese Association for Artificial Intelligence (CAAI) for the precious CAAI Ph.D. Thesis Award. In 2015, only eight Ph.D. theses across China received this award. To my knowledge, it was this award that drew the initial attention of Springer on my academic research results. After that, the editors Xiaolan Yao, Jane Li, and Celine Chang have kept in contact with me and provided plenty of useful information and materials for my preparing the book. Specifically, I wish to thank the following people for their contributions to this book:

- Prof. Christo Wilson (@ Northeastern University) guided me closely in my four best researches published in NSDI'16, IMC'15, IMC'14, and Middleware'13. I have learnt so much from him regarding how to properly tell a story, organize a manuscript, present a viewpoint, draw a figure, list a table, and so forth.
- Dr. Tianyin Xu (@ UCSD) helped me closely in conducting the researches published in NSDI'16, IMC'15, and IMC'14. I sincerely appreciate his devotion, honesty, patience, and selflessness during our collaboration.
- Zhefu Jiang (@ Cornell University and Google) helped develop the Linux kernel extension to our proposed UDS (update-batched delayed synchronization) middleware for cloud storage services. In addition, he provided valuable suggestions on the design and implementation of TrafficGuard.
- Weiwei Wang, Changqing Han, Junyi Shao, Xuefeng Luo, Min Guo, Cheng Peng, and Tianwei Wen (@ Baidu Mobile Security) helped develop the TrafficGuard system. Meanwhile, Xin Zhong, Yuxuan Yan, and Jian Chen (@ Tsinghua University) helped evaluate the client-side latency penalty and battery consumption of TrafficGuard.
- Yan Huang, Gang Liu, and Fuchen Wang (@ Tencent Research) helped develop the Xuanfeng system. Meanwhile, Yinlong Wang and Zhen Lu (@ Tsinghua University) helped evaluate the performance of multiple offline downloading systems including smart home routers.

- Prof. Yao Liu (@ SUNY Binghamton) helped evaluate the performance of Dropbox and WeChat in the US. I have learnt a great deal from her concerning measurement-driven research.
- Prof. Ben Y. Zhao (@ UCSB) guided me in conducting the researches published in NSDI'16 and Middleware'13. Also, I used to be a teaching assistant for his summer holiday course at Peking University.
- Prof. Zhi-Li Zhang (@ University of Minnesota) offered me freedom, trust, and valuable guidance during my 1-year visiting life at the University of Minnesota. Also, Dr. Cheng Jin (@ University of Minnesota) helped evaluate the performance of multiple cloud storage services in the US.
- Bob and Joanne Kraftson, as well as Mark and Jaycelyn Colestock, helped me a lot when I was visiting the University of Minnesota. They brought me warm, happy, and unforgettable memories.
- Prof. Xiang-Yang Li (@ University of Science and Technology China and Illinois Institute of Technology) provided useful suggestions on my research published in NSDI'16. I'm impressed by his hard working and diligence.
- Prof. Roger Zimmermann (@ National University of Singapore) helped me with the research published in JCST'15. I am impressed by his patience and elegance.
- Dr. Tieying Zhang (@ Chinese Academy of Sciences) helped me with the research published in IWQoS'12. Also, we collaborated on a number of other researches.
- Dr. Ennan Zhai (@ Yale University) kept in touch with me on a number of topics. Recently, our collaboration produced the first INFOCOM paper in my career.
- Dr. Yiyang Zhao, Dr. Jian Li, Zihui Zhang, Zhinan Ma, Lin Cui, Shuang Xin, Kun Duan, Songyang Wang, and Yong Yi (@ Tsinghua University) helped me a great deal in dealing with numerous affairs and issues in both my work and life.

This work is supported in part by the High-Tech Research and Development Program of China ("863—China Cloud" Major Program) under grant 2015AA01A201, the National Basic Research ("973") Program of China under grants 2014CB340303 and 2012CB316201, the National Natural Science Foundation of China (NSFC) under grants 61432002 (State Key Program), 61471217, 61232004, 61472252 and 61133006, and the CCF-Tencent Open Fund under grant IAGR20150101.

Beijing, China Zhenhua Li
May 2016

# Contents

# Acronyms

| | |
|---|---|
| AP | Access point, or WiFi home router |
| ASD | Adaptive synchronization defer |
| BDS | Batched data synchronization |
| BT | BitTorrent |
| C/S | Client/Server content distribution |
| CCN | Content centric networking |
| CDN | Content distribution (delivery) network |
| CloudP2P | Hybrid cloud and P2P content distribution |
| DASH | Dynamic adaptive streaming over HTTP |
| EC2 | Elastic compute cloud |
| HLS | HTTP live streaming |
| IDS | Incremental data synchronization |
| NDN | Named data networking |
| P2P | Peer-to-Peer content distribution |
| P2SP | Peer-to-Server&Peer content distribution |
| PUE | Power usage effectiveness |
| QoS | Quality of service |
| RFID | Radio frequency identification device |
| S3 | Simple storage service |
| SDN | Software-defined networking |
| TUE | Traffic usage efficiency |
| UDS | Update-batched delayed synchronization |
| VBWC | Value-based web cache |
| VoD | Video on demand |
| VPN | Virtual private network |
| WWW | World Wide Web, or “World Wide Wait” |