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

http://www.springer.com/978-981-10-1462-8

Content Distribution for Mobile Internet: A Cloud-based Approach

Li, Z.; Dai, Y.; Chen, G.; Liu, Y.

2016, XIII, 231 p. 146 illus., 90 illus. in color., Hardcover
ISBN: 978-981-10-1462-8