Cloud Download: *Using Cloud Utilities to Achieve High-quality Content Distribution for Unpopular Videos*

Yan Huang, Zhenhua Li, Gang Liu, and Yafei Dai

Tencent Research, and Peking University

Abstract

Video content distribution contributes to a dominant portion of the Internet traffic today. The state-of-the-art techniques (i.e., CDN and P2P) generally work well in distributing popular videos, but do not provide satisfactory content distribution service for **unpopular videos** due to low **"data health"** or **low data transfer rate**. To address this problem, we propose and implement a **"cloud download"** scheme which achieves *high-quality video content distribution* by using cloud utilities to **guarantee the data health** and **enhance the data transfer rate**. Specifically, a user sends his video request to the cloud; subsequently, the cloud downloads the video from the Internet on behalf of the user and stores a copy in the cloud cache. Then the user can retrieve his requested video (whether popular or unpopular) from the cloud with high data rate. Cloud download also brings about considerable **user-side energy efficiency**, which greatly benefits those battery-operated mobile devices. Running logs of our deployed commercial system (called "VideoCloud") confirm the effectiveness of cloud download. We find that the users' average data transfer rate of unpopular videos reaches 2.1 Mbps, 81% more than 300 Kbps -- the basic playback rate of online videos. In addition, measurement results show that the user-side energy consumption is reduced by 89%.

Motivation

Video content distribution dominates Internet traffic:

➢ *Cisco report: **~90%** of consumer IP traffic is due to video content distribution, 2012*

YouTube NETFLIX PPTV

High-quality video content distribution is of great significance

➢*high data health —– data redundancy level of a video file*

➢*high data transfer rate —– enables online video streaming*

State-of-the-art Techniques: CDN and P2P

CDN (Content Distribution Network)

➢*Strategically deploying **edge servers***

➢*Cooperate to replicate or move data according to **data popularity** and **server load***

➢*User obtains copy from a nearby edge server*

➢*Limited storage and bandwidth*

➢***Not cost-effective** for CDN to replicate **unpopular videos** to edge servers*

➢***Charged facility** only serving the content providers who have paid*

P2P (Peer-to-Peer)

➢*End users forming P2P data swarms*

➢*data directly exchanged between peers*

➢*Real strength shows for **popular file sharing***

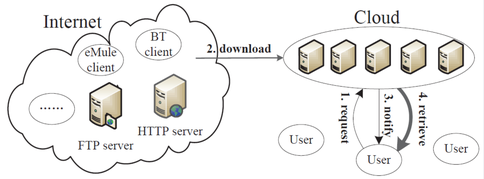
➢*Poor performance for unpopular videos*

➢*Too few peers → low data health*

➢*Too few peers → low data transfer rate*

Cloud Download

Using cloud utilities to guarantee the data health and enhance the data transfer rate



Business model

➢*CDN: serves paid content providers*

➢*Cloud download: charges users for better obtaining content*

Video accommodation

➢*CDN: one movie - **numerous copies** at edge servers*

➢*Cloud download: one movie - only **two copies** (one for user access, the other for redundancy) (unpopular video → two copies are enough)*

➢*Cloud download can accommodate many more videos than CDN, with the same storage capacity*

User-side Energy Efficiency

➢Commonly download an unpopular video

- *A common user keeps his computer powered-on for long hours*
- *Much Energy is wasted while waiting*

➢Cloud download an unpopular video

- *The user can just be "offline"*
- *When video is ready, quickly retrieve it in short time*
- ***User-side energy efficient !***

Drawback: View Startup Delay

➢For some videos

- *Anxious user must wait for the cloud to download it*
- *Thus can't view it at once*
- *The waiting time is **view startup delay***

➢This drawback is effectively alleviated

- *By the implicit and secure data reuse among users*
- *The cloud only downloads a video when it is requested for the first time*
- *Subsequent requests directly satisfied*
- *Secure because oblivious to users*

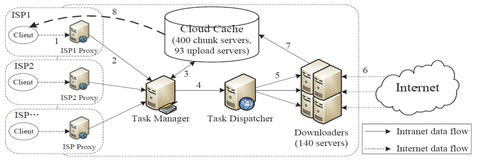
Cloud cache

Cache hit rate: 87% !

Commercial System —— VideoCloud

System Architecture:

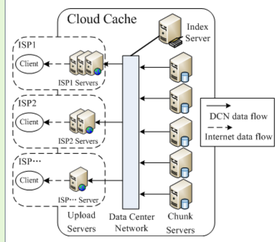
➢649 commodity servers deployed in 10 ISPs in China



Hardware Composition:

Building Block	Number of servers	Memory	Storage	Bandwidth
ISP Proxy	6	8 GB	250 GB	1 Gbps (Intranet), 0.3 Gbps (Internet)
Task Manager	4	8 GB	250 GB	1 Gbps (Intranet)
Task Dispatcher	3	8 GB	460 GB	1 Gbps (Intranet)
Downloaders	140	8 GB	460 GB	1 Gbps (Intranet), ~0.325 Gbps (Internet)
Cloud Cache	400 chunk servers, 93 upload servers, and 3 index servers	8 GB	4 TB (chunk server), 250 GB (upload server)	1 Gbps (Intranet), ~0.3 Gbps (Internet)

Data Transfer Acceleration:



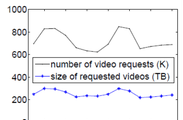
ISPs we support:

1. Telecom
2. Unicom
3. Mobile
4. CERNET
5. Tietong
6. GWBN
7. TBN
8. OCN
9. Teletron
10. Gehua

Cache Capacity Planning:

➢Handle 1.0M daily requests

- *Average video size: 390 MB*
- *Video cache duration: < 12 days*
- *Cloud cache hit rate: > 83%*
- *Thus, C = 390 MB * 1.0M *12 * (1-83%) = 796 TB < **800 TB** ← Current Cache Capacity*

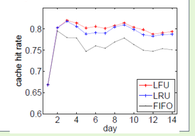


Cache Replacement Strategy:

➢*Trace-driven simulations to see what if?*

- *FIFO vs. LRU vs. LFU*
- *FIFO worst, LFU best!*

← *Unpopular data objects*



Performance Evaluation

Commercial software (http://xf.qq.com)

Dataset

➢*complete running log of the VideoCloud system in 14 days: Sep. 9, 2011 -- Sep. 22, 2011*

➢*10.1M video requests, 1.38M unique videos*

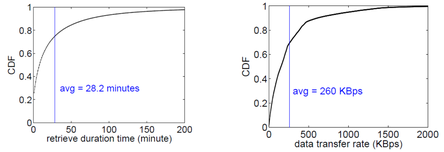
Metrics

➢*Data transfer rate*

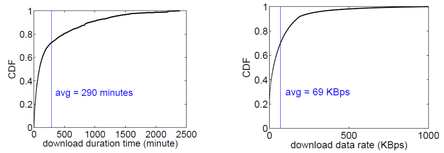
➢*View startup delay*

➢*Energy efficiency*

Data Transfer rate



View Startup Delay



User-side Energy Efficiency

➢E_1: *users' energy consumption using common download*

➢E_u: *users' energy consumption using cloud download*

➢*User-side energy efficiency = $(E_1 - E_u)/E_1$ = 89% !*

Overall Energy Efficiency

➢E_c: *the cloud's energy consumption*

➢E_2: *the total energy consumption of the cloud and users, so $E_2 = E_c + E_u$*

➢*Overall energy efficiency = $(E_1 - E_2)/E_1$ = 85% !*

User Service Policy

Charges users according to their cloud storage/cache capacity, regardless of their bandwidth consumed

➢a user of VideoCloud usually retrieves his requested video from the cloud ***for only once** (to watch)*

➢*his bandwidth consumed is basically dependent on his **storage capacity***

User level	Trial	VIP1	VIP2	VIP3	VIP4	VIP5	VIP6
Storage capacity (GB)	3	10	15	20	25	35	50
Storage duration (day)	7	7	8	9	10	11	12
Max number of concurrent requests	3	3	4	5	6	7	8
Max file size of a single request (GB)	8	8	10	15	20	25	40
Monthly charge ($) (1.57 $ = 10 RMB)	0	1.57	1.57 + 600 EXP	1.57 + 1800 EXP	1.57 + 3600 EXP	1.57 + 6000 EXP	1.57 + 10800 EXP

Related Works

Hybrid CDN-P2P

➢H. Yin, X. Liu, T. Zhan, V. Sekar, F. Qiu, C. Lin, H. Zhang, and B. Li. "Design and deployment of a hybrid CDN-P2P system for live video streaming: experiences with LiveSky," In ACM Multimedia, 2009.

P2SP

➢C. Wu, B. Li, and S. Zhao. "On Dynamic Server Provisioning in Multichannel P2P Live Streaming," to appear in IEEE/ACM Transactions on Networking.

P2P Cloud HD Video Streaming

➢F. Liu, S. Shen, B. Li, B. Li, H. Yin, and S. Li. "Novasky: Cinematic-Quality VoD in a P2P Storage Cloud," In IEEE INFOCOM, 2011.

Thunder Cloud Download: http://vip.xunlei.com/freedom/lixian.html.

Web browser using cloud download

➢UCWeb mobile browser and Amazon Silk browser

Contact

Tencent (Shenzhen, Shanghai, Beijing, Chengdu) is one of the biggest Internet companies in China, whose business involves IM, Web portal, (Micro)Blog, Video Streaming, etc. http://www.tencent.com.

Peking University (Beijing, Shenzhen) is one of the oldest and best comprehensive universities in China, which is founded in 1898.

Yan Huang, *Tencent Research* galehuang@tencent.com
Zhenhua Li, *Peking University* lzh@net.pku.edu.cn
Gang Liu, *Tencent Research* sinbadliu@tencent.com
Yafei Dai, Peking University dyf@net.pku.edu.cn