Journal of
Computer Science & Systems Biology

Sobiya and Gomathy Nayagam, J Comput Sci Syst Biol 2014, 7:3
DOI: 10.4172/jcsb.1000142

Research Article
Open Access

Dominating Set Based Content Cloud Architecture for Video Distribution Services

Sobiya* and Gomathy Nayagam M

Department of Computer Science, Einstein College of Engineering, Tirunelveli, India

Abstract

Cloud computing is a new technology that delivers services to end users on the promise of supporting on-demand in a flexible manner by scheduling bandwidth, storage and compute resources on the fly. Content delivery networks (CDN) based computers are considered as the best solutions to deliver User-generated content (UGC). But none of the existing CDN based solutions can support all the required features in UGC delivery. In this paper we analyses the various mechanism for deploying video in geographically distributed data center and optimal utilization of bandwidth, storage and other computing resources. Optimal Deployment can be done by using Dominating Set Algorithm.

Keywords: Content delivery networks; User generated content; Dominating set; Video on-demand

Introduction

This Cloud computing has recently emerged as a new computing paradigm for organizing a shared pool of servers in data centers into a cloud infrastructure that can provide on-demand server utilities (CPU, storage, bandwidth, etc.) to users anywhere anytime [1-3]. Live Media Streaming is deployed using CDN and peer-peer network [4]. The former achieves high availability and short start-up latencies, but suffers from excessive costs for deploying dedicated servers. This is particularly severe if the user demand fluctuates significantly and the servers have to be over-provisioned for peak loads [5].

Video on demand (VOD) are systems which allow users to select and watch/listen to video or audio content on demand manner [6]. IPTV technology is use for video on demand services in televisions and personal computers [7]. Catch up TV is an example of video on demand [8]. Television Video On Demand systems such as stream content through a set-top box, a computer or other device, allowing the user to view in real time, or download it to a device such as a computer, digital video recorder (also called a personal video recorder) or portable media player for user to view at any time. The majority of service providers offer both VOD streaming, including pay-per-view and free content, whereby a user buys or selects a content i.e., video services and it begins to play on the device set almost instantaneously, or downloading to a DVR rented from the provider, or downloaded onto a PC, for viewing in the future [9]. Internet television, using the Internet, is an increasingly popular form of video on demand [10].

A content delivery network or content distribution network (CDN) is a large distributed system of servers deployed in multiple data centers across the Internet. CDN is used to serve content to end-users with high availability and high performance [11,12]. CDNs serve a large fraction of the Internet content today, including web objects (text, graphics and scripts), downloadable objects (media files, software, and documents), applications (e-commerce, portals), live streaming media, on-demand streaming media, and social networks [13]. Some general-purpose CDNs also provide CDN services for on-demand video content [14].

Challenging task of video service provider (VSP) optimally deploy its distribution infrastructure over multiple geo-distributed cloud data centers [15]. A VSP needs to minimize the operational cost induced by the rentals of cloud resources without sacrificing user experience in all regions. The geographical diversity of cloud resource prices further makes the problem complicated [16].

The scope of our effort is to construct an overlay network based on Dominating Set Theory to optimize the number of nodes for large data transfer. In order to offload popular servers and improve end user experience, copies of popular content are often stored in different locations. With mirror site replication, documents from a primary site are proactively replicated at secondary sites. When a copy of the same document exists at multiple servers, choosing the server that provides the best response time is not trivial and the resulting performance can dramatically vary depending on the server selected. By placing UGC that is lecture video on different cloud data center using cloud service provider user from different region can utilize the video. Here we are considering best trade-off between user experience and operational cost, deployment of UGC video on different cloud data centers. Deployment of VSP over multiple geo-distributed cloud data centers in optimized manner using dominating set and we can achieve the best trade-off between operational cost and user experience. Video streaming from different user region handled on-Demand basis.

Related Work

Wendell and Freedman introduces Flash crowd [17]. Flash crowd: As a period over which request rates for a particular fully qualified domain name are increasing exponentially. The behaviour of thousands of flash crowds on coral CDN is characterized and quantified. Coral CDN is an open content distribution network running at several hundred POPs.

Jin and Wen proposed Content Delivery as a service (CODaaS)

***Corresponding author:** Sobiya, Department of Computer Science, Einstein College of Engineering, Tirunelveli, India, Tel: 9487493495; E-mail: sobiyaraj99@gmai.com

Received March 13, 2014; **Accepted** April 03, 2014; **Published** April 05, 2014

Citation: Sobiya, Gomathy Nayagam M (2014) Dominating Set Based Content Cloud Architecture for Video Distribution Services. J Comput Sci Syst Biol 7: 089-094. doi:10.4172/jcsb.1000142

for video distribution. CODaaS is an innovative idea to enable on-demand virtual content delivery service (vCDS) overlays for UGC providers to deliver their contents to a group of designated consumers. It is a context aware media processing rendering capabilities for better quality of experience (QDE) and social networking function. A VOD provider can make reservations for bandwidth guarantees from cloud service providers to guarantee the streaming performance in each video channel [18]. The predictable anti-correlation between demands to enhance resource utilization and derive optimal load direction that minimizes the bandwidth resource reservation while confining under provision risks.

In [1] the optimal deployment and joint optimization problem for distributing video in different geographically distributed cloud data centers. A video Service Provider (VSP) needs to minimize the operational cost induced by the rentals of cloud resources without sacrificing user experience in all regions [19]. Optimal deployment problem of cloud assisted video distribution service is handled by Nash bargaining Solutions. The problems here addressed are,

1. Optimal deployment of cloud resources in different data centers.
2. Minimum network flow model formulates cloud deployment.
3. Nash bargaining deals with video placement strategies.

These are applied in centralized and peer contributions.

Wang and Liu proposed CALMS. CALMS adaptively leases and adjusts cloud server resource in a fine granularity to accommodate temporal and spatial dynamics of demands from live streaming users. To mitigate the overall system deployment costs and yet provides users with satisfactory streaming and latency rate CALMS is used [20]. Presenting optimal solutions to deal with cloud servers with diverse capabilities and lease price as well as potential latencies in initiating and terminating leases in real world cloud platforms [20].

An open content distribution network running at several hundred POPs. In this paper, measurements shows that crowds vary in their amenability to effective caching and that cooperation between caching proxies minimizes origin load for some crowds but not all [21]. Some crowds grow too quickly for dynamic resource allocation to keep up with demand. Third party portals contribute to substantial number of crowds and documents that sharing of links through social media and uncoordinated use of search engines lead to flash crowd formation [17].

Basic Problem

Internet Video Service Provider generally resorts to CDN to conduct Large-Scale video distribution. CDN solutions are inadequate for emerging video traffic growth. CDN offers only Semi-Static resource Provisioning. On-demand resource provisioning is not effectively defined in CDN. Emerging UGC in video is long-tail in nature. So handling such videos in cloud environment is difficult to handle in an on-demand manner. Bandwidth and storage cost are very sensitive to usage variations [22]. For efficient use of Cloud CDN requires “replica placement” and “user redirection”. Replica Placement: Placing content in different machines for 24×7 availability, easy access and failure recovery [23]. User Redirection: When the requested service is not available in the data center then it must have to redirect the request to another data center. Volatile request rates complicate the delivery of Internet video services. A VOD provider makes sure that sufficient server bandwidth is provisioned to sustain continuous media delivery to end users [24]. To move UGC to clouds, to partition UGC and assign them into number of cloud servers.

Implementation Overview

To distribute video across multiple geographically distributed data center user has to clear about Cloud Service Provider (CSP), Video Service Provider (VSP), Data center and User region (Figure 1).

Cloud services

Services that are made available to users on demand via the Internet from a cloud computing provider’s servers and it is opposed to provide from a company’s own on-premises servers.

Cloud service provider

A service provider that offer customers storage or software services available via a private (private cloud) or public network (cloud). Usually, it means the storage and software is available for access via the Internet.

Video service provider

The Video Service Provider is an entity that provides video programming to the end user. Some of the examples are the satellite, cable, telco and free-over-the-air broadcasters.

User region

Different users are grouped as a region based on their coverage area. Each and every user region is connected to data center (DC). Data center may be connected to one or more user regions. Users can request video to video service provider by means of Data centers.

Video Segmentation

Segmentation is the process of partitioning a piece of information into meaningful elementary parts termed as segments.

Considering still images, (spatial) segmentation means partitioning the image to a number of arbitrarily shaped regions, each of them typically being assumed to constitute a meaningful part of the image, i.e. to correspond to one of the objects depicted in it or to a part of one such object. Considering moving images, i.e. video, the term segmentation is used to describe a range of different processes for partitioning the video to meaningful parts at different granularities. Segmentation of video can thus be temporal, aiming to break down the video to scenes or shots,

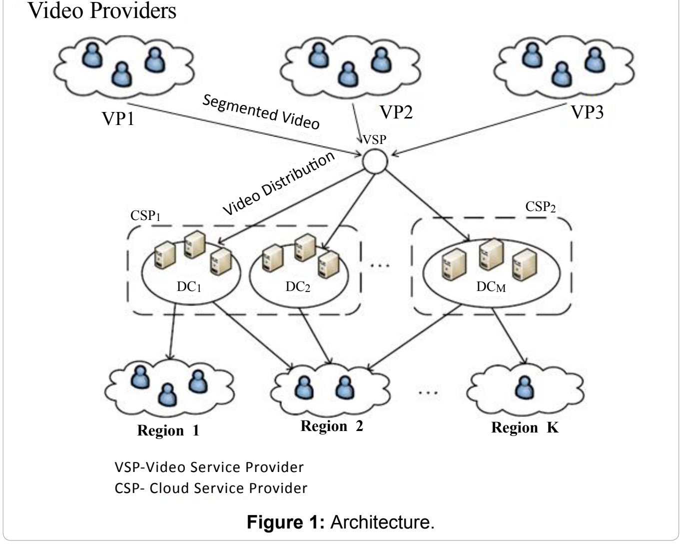

Figure 1: Architecture.

spatial, addressing the problem of independently segmenting each video frame to arbitrarily shaped regions, or spatio-temporal, extending the previous case to the generation of temporal sequences of arbitrarily shaped spatial regions. The term segmentation is also frequently used to describe foreground/background separation in video, which can be seen as a special case of spatio-temporal segmentation [17]. Regardless of the employed decision space, i.e. 1D, 2D or 3D for temporal, spatial and spatio-temporal segmentation, respectively, the application of any segmentation method is often preceded by a simplification step for discarding unnecessary information (e.g. low-pass filtering) and a feature extraction step for modifying or estimating features not readily available in the visual medium (e.g. texture, motion features etc., but also colour features in a different colour space etc.), as illustrated in for a variety a segmentation algorithms.

Segmentation of images and video is generally an ill-posed problem, i.e. for a given natural image or image sequence, there exists no unique solution to the segmentation problem; the spatial, temporal or spatio-temporal segments that should ideally be formed as a result of segmentation largely depend on the application under consideration and most frequently on the subjective view of each human observer.

Commonly considered applications of segmentation include region-based image and video description, indexing and retrieval, video summarization, interactive region-based annotation schemes, detection of objects that can serve as cues for event recognition, region-based coding, etc.

Video segmentation is a ways of dividing a movie into segments that is having some sort of information about the video. In the context of video capture, segmentation is best applied to capture screen presentation, for easy use of the presenter goes through slide after slide.

In computer vision, video segmentation is the process of partitioning a video into multiple segments (sets of frames, also known as super frames). The goal of segmentation is to simplify and/or change the representation of a video into something that is more meaningful and easier to analyze. Video segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in Video. More precisely, Video segmentation is the process of assigning a label to every frame in a Video such that frames with the same label share certain visual characteristics.

The result of Video segmentation is a set of segments that collectively cover the entire Video, or a set of contours extracted from the Video. Each of the frames in a region is similar with respect to some characteristic or computed property.

Algorithm

Input:

Video (any format)

Output:

Segmented video in .jpeg format

• Initialization step: define frame length f.

• Let S be shots and E be edits and V as frame.

• S= $(s_1, s_2, \ldots s_n)$ and E= $(e_1, e_2, \ldots e_n)$

• S= (t_b,t_e) where t_b, t_e is starting and ending time of frame.

• Compute V using $V=s_1Oe_1(s_1,s_2)Os_2O\ldots s_n$ where O is concatenation operation.

• Segment the video using clone method.

• Increment f->f+1 to obtain next frame.

• Return f and V value.

Consider a set of shots $S = (s_1, s_2, \ldots s_n)$ with cardinality N. Each shot Si c S can be represented by a closed time interval:

$Si=[t_b/t_e]$,

Where t_b is the time at which the shot begins and t_e is the time at which the shot ends. Before the final cut is made t_b=0. S since there is no relative ordering of the shots. Let $E=(e_1, e_2, \ldots e_n)$ be the set edits available. Consider a video V. Let V be composed of n shots taken from the set S. Then V can be represented as follows.

$$\boldsymbol{V}=\mathbf{s_1\ o\ e_{12}}(s_1, s_2)\ \mathbf{o\ s_2\ o\ e_{23}(s_2, s_3)\ o\ldots o\ s(_{n-1})\ o}\ \boldsymbol{e(_{n-1})_n(s_{n-1},s_n)}\ \mathbf{os_n}$$

Where S_iєS, the subscript i denotes the temporal position of the shot in the sequence (i.e. if i < j shot Si appears before shot Sj in the final cut) and Eij denotes the edit transition between shots Si and Sj. The o denotes the concatenation operation and Eij є E.

Video Distribution

The advent of user-generated content (UGC) has re-shaped the online video market enormously. Nowadays, hundreds of millions of Internet users are self-publishing consumers. The content length is shortened by two orders of magnitude and so is the production time. By placing UGC that is video on different cloud data center using cloud service provider user from different region can utilize the video. Here we are considering best trade-off between user experience and operational cost, deployment of UGC video on different cloud data centers. Deployment of VSP over multiple geo-distributed cloud data centers in optimized manner using dominating set and we can achieve the best trade-off between operational cost and user experience. Video streaming from different user region handled on-Demand basis.

Dominating Set

A dominating set (DS) is a subset of nodes such that every node in the graph is either in the set or is adjacent to a node in the set. If the sub graph induced from a Dominating Set of the node is connected, the Dominating Set is a Connected Dominating Set. These sets are needed when the mobile network becomes larger. These connected dominating sets are formed based on neighbour node‟s locations (hops). For distributing video on cloud data centers we are using mathematical model called dominating set. Here dominating set modelling is used for placing cloud servers and for more coverage area. In graph theory, a dominating set for a graph $G=(V, E)$ is a subset D of V such that every vertex not in D is adjacent to at least one member of D. The domination number $\alpha(G)$ is the number of vertices in a smallest dominating set for G. The dominating set problem concerns testing whether $\alpha(G) \le S$ for a given graph G and input S; it is a classical NP-complete decision problem in computational complexity theory. It is believed that there is no efficient algorithm that finds a smallest dominating set for a given graph (Figure 2).

Figure 2a–2c on the right show three examples of dominating set for a graph. In example, each white vertex is adjacent to at least one red vertex, and it is said that the vertex white is dominated by the red vertex. The domination number of graph is two: the examples (2b) and (2c) show that there is a dominating set with two vertices, and it can be checked that there is no dominating set with only 1 vertex for this graph.

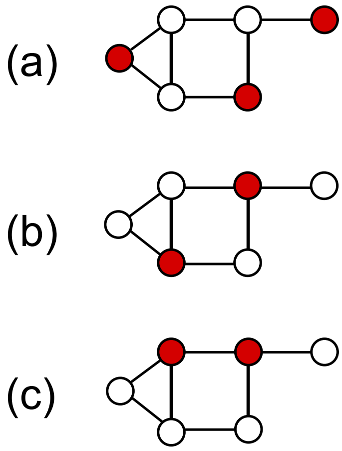

Figure 2: Dominating node.

Placing video in data center based on dominating set for video distribution

Data center selection Network 'G' can be defined as follows. G={V, E}, where V={V1, V2, V3, ..Vn} be the set of servers and E is the set of edges between i^{th} server and j^{th} server i.e. E=(Vi, Vj) such that Vi ≠Vj.

Let D be the dominating set of G and D ⊂ G, the data center not in D is adjacent to at least one data center in D. Hence, all the servers are either member of D or V\D. Equitable Dominating set D is a set of 'r' dominating vertices in V since *D*=r and V\D is the set of all the adjacent vertices of dominating server set D such that the difference between the degrees of all the vertices in D can differ utmost by 1. Each vertex v in D has more or less same number of neighbour nodes which are members of V\D. So contents are only replicated in the set of dominating data center D which contains 'r' data centers or less than 'r' number of data center's i.e. *D* ⊂ *V*.

Algorithm for Forming Dominating Set Based Network

Step 1: The algorithm begins by marking all the vertices of the graph white.

Step 2: Select the vertex with the maximal number of white neighbours.

Step 3: Selected vertex is marked black and its neighbours are marked red.

Step 4: Iteratively scans the red nodes and their white neighbours, and selects the red node or the pair of nodes (a grey node and one of its white neighbours), whichever has the maximal number of white neighbours.

Step 5: The selected node or the selected pair of nodes is marked black, with their white neighbours marked red.

Step 6: Once all the vertices are marked red or black, the algorithm terminates. All the black nodes form a connected dominating set (CDS).

Step 7: After forming the CDS, check the degree of each vertices of the connected dominating set.

Step 8: If the degree of any vertex vary more than one then mark that vertex red and find the suitable alternate vertex as the member of the dominating set and mark it black. If no alternate node is found then leave as it is (Figure 3).

Performance Analysis

Performance metrics for fast replica

Let Time denote the transfer time of file F from the original node N0 to node Ni as measured at node Ni. Two performance metrics: average and maximum replication times are considered. In idealistic setting all the nodes and links are homogeneous, and let each node can support n network connections to other nodes at B bytes/sec. Then,

Time distribution=Size (F)/(nxB)

Time collection=Size (F)/(nxB)

For Fast Replica,

Time FR=Time distribution+Time collection=2 x Size (F)/(nxB)

For Multiple Unicast,

Time MU=Size (F)/B

Replication Time Speedup=Time MU/Time FR=n/2.

In this work, the following distribution schemes are compared.

Fast Replica-files are replicated from many hosts in the system based upon which nodes download these files.

Sequential Unicast-approximates distribution via IP multicast, measures transfer time of entire file from the source to each recipient independently.

Multiple Unicast-simultaneously transfers the entire file to all the recipient nodes by using concurrent connections.

Performance analysis of video segmentation

Segmentation is the process of partitioning a piece of information into meaningful elementary parts termed as segments (Table 1, Figures 4 and 5).

In Figure 5 comparisons between different approaches are made. Here the Max-central and Max-CDN approaches yields maximum cost compared to CALMS and Dominating set. CALMS achieve 33.5%-45.1% of the Max-Central approach and 30.5%-39.6% of the Max-CDN approach, respectively. By using Dominating set we can achieve 10%-12% of CALMS approach. P2P locality is the lowest cost approach. There is no streaming at all direct communication is possible.

It can be observed that DOMINATING SET for video distribution

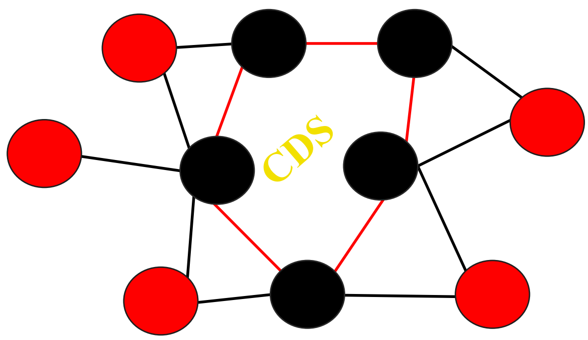

Figure 3: Formation of dominating set based network.

Requirements	without cloud	with cloud
CPU Usage	67	52
Memory	71	63
Disk	4	2

Table 1: Performance Analysis.

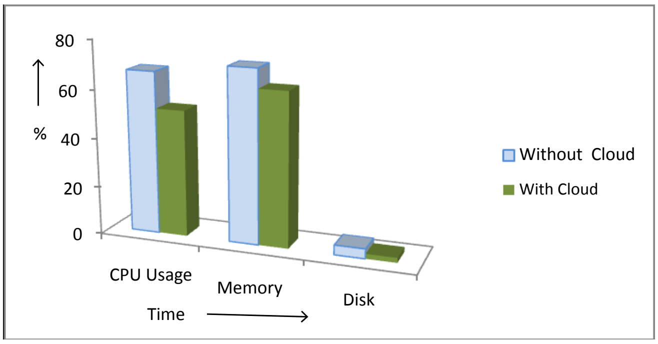

Figure 4: Performance analysis.

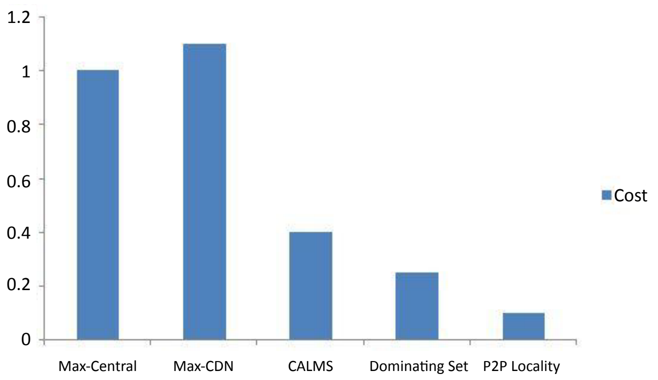

Figure 5: Comparison between different approaches.

can reduce around 80% of the total cost compared with the existing Nash Bargaining Solution. The reduction is about 50% compared with the local-only strategy. Nash bargaining and local-only strategy are not implemented it is a theoretical result. By considering peer contribution, we can further reduce the incurred total cost. To analyze our proposed algorithms in detail, we study the incurred the bandwidth cost and the latency cost. Figure 4 illustrates the performance analysis of video segmentation under cloud and without cloud environment. As expected, the local-only strategy incurs the largest bandwidth cost, and the optimal bandwidth cost can be achieved by the centralized strategy.

Conclusion and Future Work

The on-demand self-service features of cloud computing enables a video service provider to provision its cloud resources adaptively for video distribution [25]. In this paper, the optimal deployment of cloud resources in multiple geographically distributed data center is formulated to improve the user experience and minimize the operational cost simultaneously. Here Dominating Set is used to formulate the cloud deployment problem and derive corresponding optimal bandwidth provisioning and video placement strategies. First the segmentation of video can be done by using frame grabber algorithm and then segmented video is distributed using Dominating Set without cloud environment and with cloud environment. The results are provided in graph. Here we investigated the deployment over a hybrid cloud infrastructure with private clouds. In future, better selection of data centers to optimize video applications with different QOS requirements will be studied. We are also interested in exploring other open issues along this direction, such as designing better user demand and capacity forecast mechanisms as well as extending video distribution to integrate other type of cloud service platform, e.g., Spot Cloud, a platform that allows the users to contribute/sell their idle computing resources and build their own cloud services [26].

References

1. Niu D, Xu H, Li B, Zhao S (2012) Quality-assured cloud bandwidth auto-scaling for video-on-demand applications. INFOCOM 12: 460-468.
2. Shen H, Zhao L, Stokes J, Li J (2012) Toward P2P-Based Multimedia Sharing in User Generated Contents. IEEE Transactions on Parallel and Distributed Systems 23: 966-975.
3. Jian Hey, Di Wuy, YupengZengy, XiaojunHeiz, YonggangWenx (2013)Towards Optimal Deployment of Cloud-Assisted Video Distribution Services.Circuits and Systems for Video Technology 23: 1717-1728.
4. Yin H, Liu X, Zhan T, Sekar V, Qiu F, et al (2010) Live Sky: Enhancing CDN with P2P. ACM Trans Multimedia ComputCommun App 6: 1-19.
5. China Telecom homepage.
6. Jin Y, Wen Y, Shi G, Wang G, Vasilakos A (2012) CoDaaS: An experimental cloud-centric content delivery platform for user-generated contents.In Proc. of IEEE International Conference on Computing, Networking and Communications 934-938.
7. Aggarwal V, Chen X, Gopalakrishnan V, Jana R, Ramakrishnan K, et al. (2011) Exploiting virtualization for delivering cloudbased IPTV services.Computer Communications Workshops (INFOCOM WKSHPS) IEEE 637-641.
8. Amazon EC2 pricing.
9. Binmore K, Rubinstein A, Wolinsky A (1986) Thenash bargaining solution in economic modelling.The RAND Journal of Economics 17:176-188.
10. Wang F, Liu J, Chen M (2012) CALMS: Cloud-assisted live media streaming for globalized demands with time/region diversities. INFOCOM 199-207.
11. Google map homepage.
12. Palomar D, Chiang M (2006) A tutorial on decomposition methods for network utility maximization.IEEE Journal on Selected Areas in Communications 24: 1439-1451.
13. Bertsekas D (2009)Convex Optimization Theory. Athena Scientific 1-256.
14. Niu D, Liu Z, Li B, Zhao S (2011) Demand forecast and performance prediction in peer-assisted on-demand streaming systems. INFOCOM 421-425.
15. Wendell P, Freedman M (2011) Going viral: Flash crowds in an open CDN. Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference 549-558.
16. Huang Z, Mei C, Li L, Woo T (2011) CloudStream: delivering highquality streaming videos through a cloud-based SVC proxy. INFOCOM 201-205.
17. Zhu W, Luo C, Wang J, Li S (2011) Multimedia cloud computing:Directions and applications.Signal Processing Magazine 28: 59-69.
18. Niu D, Feng C, Li B (2012) A theory of cloud bandwidth pricing for video-on-demand providers. INFOCOM 711-719.
19. Cheng X, Liu J (2011) Load-balanced migration of social media to content clouds. NOSSDAV 51–56.
20. Liu F, Li B, Zhong L, Li B, Jin H, et al. (2012) Flash Crowd in P2P Live Streaming Systems: Fundamental Characteristics and Design Implications. IEEE Transactions on Parallel and Distributed Systems 23: 1227–1239.
21. Gursun G, Crovella M, Matta I (2011) Describing and forecasting videoaccess patterns. INFOCOM 16-20.
22. Suh K, Diot C, Kurose JF, et al. (2007) Push -to-peer video-on-demand system: Design and evaluation [J]. IEEE J Selected Areas in Comm 25: 1706-1716.

23. Wu Y, Wu C, Li B, Zhang L, Li Z, et al. (2012) Scaling social media applications into geo-distributed clouds. INFOCOM 684 - 692.

24. Jiang W, Zhang-Shen R, Rexford J, Chiang M (2009) Cooperative content distribution and traffic engineering in an ISP network.SIGMETRICS 37: 239-250.

25. ZhangG, ZhaoX, HeiX, ChengW (2012) A comparison measurement study of hybrid CDN-P2P VoD systems. EIE Department, Huazhong University of Science and Technology.

26. Spot Cloud.