ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

Cost-Minimizing Dynamic Migration of Content Distribution Services into Hybrid Clouds

Rashmi Rajendra Kurkure [1], Madhuri Zawar [2]

P.G. Student, Department of Computer Engineering, GF's Foundation Godavari College of Engineering, Jalgaon, Maharashtra, India[1]

Associate Professor, Department of Computer Engineering, GF's Foundation Godavari College of Engineering, Jalgaon, Maharashtra, India[2]

ABSTRACT: End-users are provided with plentiful contents using a type of Internet cloud service called Content Service. A technology known as Content Multi-homing is used to ensure high performance in offering content services to end-users. Multiple geographically distributed data centers generate contents and the generated contents are delivered by multiple distributed content distribution networks (CDNs). The major contributors for the cost of content service are electricity costs for DC and usage costs for CDNs. Scheduling data centers and CDNs has a tremendous consequence for optimizing content service cost, as usage costs ands electricity prices vary across DC and CDNs. A content service provider chooses a data center to generate contents based on data availability, performance, cost and other reasons. Second, contents can also be delivered by a group of CDNs distributed over the Internet. To minimize the content service cost a framework called MCSCIDCUM is used which dynamically balances end-user's loads among DCs and CDNs. Real life electricity prices and CDN traces are used as input. Experiments demonstrate that MCSCIDCUM effectively reduces the cost of content service by more than 40%.

KEYWORDS: Content Multihoming, content distribution network.

I. INTRODUCTION

Multihoming refers to a computer or device connected to more than one computer network. It can be used, for example, to increase the reliability of an Internet Protocol network, such as a user served by more than one Internet service provider. Contents are generated from multiple geographically distributed Data centers and delivered by multiple distributed by CDNs. Recent years have witnessed the proliferation of cloud computing. Content service, such as video streaming, is a type of cloud-computing service that provides end-users plentiful contents. Content service utilizes two types of infrastructure data center and content distribution network. A data center is the origin of contents it serves the end users' requests, generates contents, and sends them back to end-users.

A CDN is responsible for efficient delivery of contents it replicates contents originated from a data center, and uses the locally stored contents to serve end-users' requests. To ensure high performance for content delivering, content service utilizes a technology known as content multi-homing for content generation and distribution. First, contents can be generated from a group of data centers distributed in different regions. For instance, it was reported that Google operated 19 data centers throughout the US.

A content service provider chooses a data center to generate contents based on data availability, performance, cost and other reasons. Second, contents can also be delivered by a group of CDNs distributed over the Internet. For instance, it was reported that to deliver video contents, Hulu and Netflix used three CDNs Akamai, Limelight, and

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

Level3. A content service provider chooses a CDN based on performance, cost and other reasons. As data centers develop rapidly to meet the soaring cloud-computing demands, the consumption and cost of energy by data centers are skyrocketing.

Recent studies showed that large-scale data centers consumed dozens of Megawatt and incurred a total cost as high as 56M dollars per year; data centers consumed about 1.3% of the total worldwide supplied electricity and this fraction would grow to 8% by 2020.

Electricity cost minimization for data centers is an important research problem. There is growing interest in how to choose CDNs for content delivery and minimize the CDN usage costs while the performance requirement is satisfied. A few recent studies took a constrained optimization approach to the usage cost minimization for CDNs. While minimizing electricity costs for data centers and minimizing usage costs for CDNs are both good for reducing the content service cost, most recent works took either the former or latter approach.

II. RELATED WORK

Content Delivery Networks (CDNs) deliver web content to end-users from a large distributed platform of web servers hosted in data centers belonging to thousands of Internet Service Providers (ISPs) around the world. The bandwidth cost incurred by a CDN is the sum of the amounts it pays each ISP for routing traffic from its servers located in that ISP out to end-users. A large enterprise may also contract with multiple ISPs to provide redundant Internet access for its origin infrastructure using technologies such as multihoming and mirroring, thereby incurring a significant bandwidth cost across multiple ISPs. The project Optimizing Algorithm initiates the formal algorithmic study of bandwidth cost minimization in the context of a large enterprise or a CDN, a problem area that is both algorithmically rich and practically very important.

A. Existing system

Electricity cost minimization for data centers is an important research problem. Most existing work tackles this problem by exploiting the electricity price variation in location or time, and scheduling end-users' load among geographically distributed data centers or among time slots with different electricity prices.

Electricity costs for data centers and usage costs for CDNs are another major contributor to the content service cost. There is growing interest in choosing CDNs for content delivery and minimize the CDN usage costs while the performance requirement is satisfied.

Drawbacks of Existing system

- There is growing interest in how to choose CDNs for content delivery
- Minimize the CDN usage costs while the performance requirement is not satisfied
- Speed is not much satisfiable

III. PROBLEM DEFINITION

We discuss how to choose data centres for content generation and CDNs for content delivery so as environment. The minimization of content service cost is formulated from service provider's perspective and it can significantly reduce the operation cost so as to maximize the profits.

Note that the content services we discuss are not real-time, such as VoIP. Hence, contents can be replicated in CDNs without compromising the quality of service. To this end, we propose a novel load scheduling framework named MCSCIDCUM (Cost Optimization for Internet Content Multihoming). MCSCIDCUM takes a holistic approach to the content service cost minimization by formulating an optimization problem that minimizes the sum of electricity costs for data centres and usage costs for CDNs as well as guaranteeing service performance requirements.

The contributions of this paper are twofold. We study an important research problem: the content service cost minimization. To our best knowledge, our work is the first that takes a holistic approach by covering the content service cost from the content generation to the content delivery, i.e., the electricity costs for data centres and the usage costs for

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

CDNs; Our extensive experiments show that MCSCIDCUM is effective in reducing the content service cost. Moreover, MCSCIDCUM is proposed with the real-world practicality in mind. MCSCIDCUM takes as inputs the real-time electricity prices on data centre sites and the real-time usage costs of CDNs, and satisfies the realworld constraints, such as the processing capacities of data centres and CDNs, and the data availability situation in data centres. Thus, MCSCIDCUM is amenable to deployment in the real world.

Advantages

- MCSCIDCUM effectively reduces the content service cost by more than 40%.
- Optimizing electricity cost for data centers.
- Optimizing cost for CDNs.

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy.

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs.

In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user.

IV. SYSTEM FRAMEWORK

The real time electricity prices and CDN traces are used as input. For instance To explain the designing method, whole process is given about how a request of dynamic content data is processed.

The user group ul requests the dynamic content data dm through CDN cj periodically. Usually, CDN cj will cache the static content in their own servers, and only the dynamic content data in data center is focused here. Hence, CDN cj will distribute its requests to origin data centers which has the data.

Here it is assumed as data center nk receives this request, and nk will serve this request and sends the desired data dm to CDN c j. After receiving the data, cj will send it to ul . In this request distribution process, the number of requests from user group ul to CDN cj for the data dm is represented by the function of fuc(ul ;cj;dm); similarly, the function of fcd(cj;nk;dm) represents the number of requests which CDN cj distributes to data center nk for the data dm.
Few studies have investigated the content service cost minimization problem and took both electricity costs for data centers and usage costs for CDNs into consideration. Therefore it would be interesting to study the content service cost minimization problem from the perspective of scheduling end-users' loads among data centers and CDNs to minimize the sum of electricity costs for data centers and usage costs for CDNs. In MCSCIDCUM, a study is undergone to choose data centers for content generation and CDNs for content delivery so as to minimize the content service cost in a content-multi homing environment. The minimization of content service cost is formulated from service provider's perspective and it can significantly reduce the operation cost so as to maximize the profits. Hence, contents can be replicated in CDNs without compromising the quality of service. To this end, a load scheduling framework named MCSCIDCUM. MCSCIDCUM takes a holistic approach to the content service cost minimization by formulating an optimization problem that minimizes the sum of electricity costs for data centers and usage costs for CDNs as well as Guaranteeing service performance requirements. The contributions of MCSCIDCUM paper are twofold an important research problem the content service cost minimization. MCSCIDCUM project is the first that takes a holistic approach by covering the content service cost from the content generation to the content delivery, i.e., the electricity costs for data centers and the usage costs for CDNs; The project show that MCSCIDCUM is effective in reducing the content service cost. Moreover, MCSCIDCUM is proposed with the realworld practicality in mind. MCSCIDCUM takes as inputs the real-time electricity prices on data center sites and the real-time usage costs of CDNs, and satisfies the real-

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

world constraints, such as the processing capacities of data centers and CDNs, and the data availability situation in data centers. Thus, MCSCIDCUM is amenable to deployment in the real world.

COMIC Framework

In this section, we give an overview of our proposed COMIC framework to optimize the sum of electricity costs for data centers and usage costs for CDNs through content multihoming. The entities in the COMIC framework mainly include four parts: user group, data, CDN and data center.

1) **User group**: A user group is the set of users who are represented by the one and same identity in the COMIC framework. For example, the users physically close to each other and served by the same regional ISP may be aggregated into a user group. In the following section, we do not distinguish between user and user group.

2) **Data:** Data are the collection of content objects that the users request through the Internet. A content provider can have several types of content objects, such as a video in an online video website or an email or a message. Note that there are both static content (e.g., images) and dynamic content (e.g., PHP) in a request. However, the static content is usually cached in the replica server of CDNs, we only consider the dynamic content in data center1.

3) **CDN:** a CDN is a large distributed system of servers which replicates contents originated from data centers, and delivers the locally stored contents to end-users through network. In small geographical regions, such as a city, the pricing function of a CDN may be the same. However, different geographical regions in a CDN or the same geographical region in different CDNs may charge differently. We use charging region to refer to the geographical region which has the same pricing function within a certain CDN.

4) **Data center**: a data center is a collection of servers in a certain location to serve the data requests from users. Note that the electricity cost of a data center depends on not only the size of arrival requests, but also the local real-time electricity price in the modern power grid.

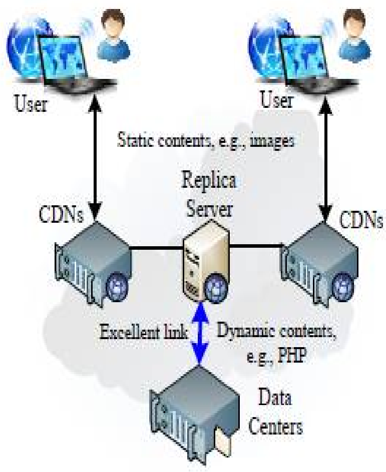

Fig: Architecture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

Algorithm 1 Pseudo-code of generating non-zero variables.

Input: UC_l: the set of CDN c_j the user group u_l connects; CN_j: the set of data center n_k CDN c_j connects; ND_k: the set of data d_m data center n_k could handle; U_s: the set of user groups; C_s: the set of CDNs.

Output: F_{uc}: the set of non-zero variables of f_{uc}; F_{cd}: the set of non-zero variables of f_{cd}.

for all $c_j \in C_s$ **do**
  $CD_j = \emptyset$ //CD_j: the set of data d_m CDN c_j could handle;
  for all $n_k \in CN_j$ **do**
    for all $d_m \in ND_k$ **do**
      if $d_m \notin CD_j$ **then**
        add d_m in CD_j
      end if
    end for
  end for
end for
for all $u_l \in U_s$, $c_j \in UC_l$, $d_m \in CD_j$ **do**
  calculate F_{uc} consisting all the permutations as non-zero variables of f_{uc}
end for
for all $c_j \in C_s$, $n_k \in CN_j$, $d_m \in ND_k$ **do**
  calculate F_{cd} consisting all the permutations as non-zero variables of f_{cd}
end for

V. EXPERIMENTAL SETUP

Our experimental setup includes three parts: data centers and workload data, the usage pricing for CDNs, and electricity pricing on data center sites.

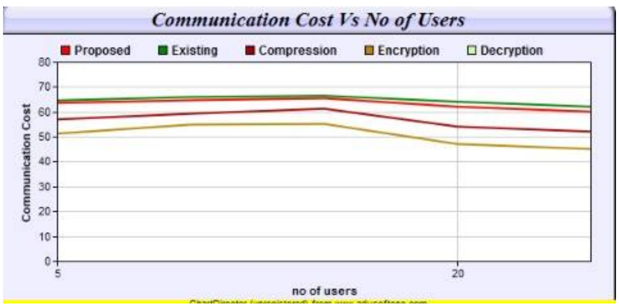

Fig a: Communication cost Vs no of users graph

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

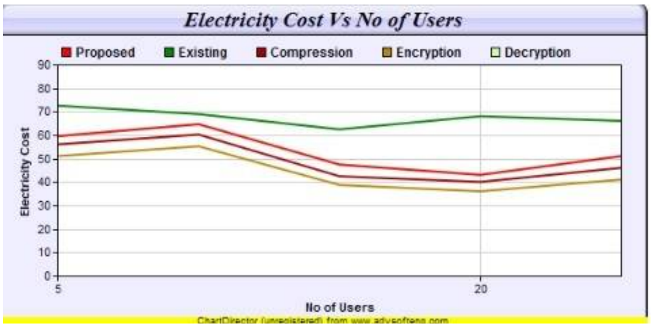

Fig b: Electricity cost Vs no of users graph

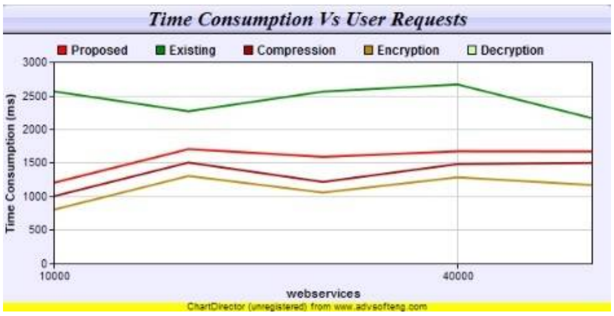

Fig c: Time consumption Vs user requests

This graphs are taken on following configuration-

Windows Edition: Windows 8
Processor: Intel(R) core (TM) i3-3110M CPU @ 2.40GHz
RAM: 4.00 GB
System Type: 64-bit operating system, x64 based processor

Results may vary according to the configuration.

VI. CONCLUSION AND FUTURE SCOPE

Internet content service involves content generation from data centers and content delivery by CDNs.

Electricity costs for data centers and usage costs for CDNs are two major contributors to content service cost. MCSCIDCUM involves the minimization of content service cost. A framework named MCSCIDCUM is proposed in this paper and a centralized optimization algorithm is developed to dynamically balance end-users' loads among data centers and CDNs so as to minimize the content service cost in the content multi-homing environment. Real-life electricity prices and CDN prices are used to evaluate the effectiveness of MCSCIDCUM. Extensive experiments demonstrate that MCSCIDCUM effectively reduces the content service cost by more than 20%**.**

In future, this work may be extended at high level and can also be implemented in colleges and Net centers to avoid over usage of internet. In colleges it can be used to provide only useful and necessary contents for students. Private pages can be accessed only by authorized user and there can be privacy in using private web-sites. The user and the provider both are benefited by using this MCSCIDCUM technique. This can be further extended even

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

more to provide reliable service to the entire state or a country as a whole. The recent experiments proved that the service cost for content delivery is being reduced by more than 20%. This framework can also reduce the cost by more than 40% in upcoming era.

REFERENCES

1. Jianguo Yao, Haihang Zhou, Jianying Luo, Xue Liu, and Haibing Guan (2014), “COMIC Cost Optimization for Internet Content Multihoming,” in IEEE Transactions on Parallel and Distributed System.
2. AdhikariV, Guo Y,HaoF, HiltV, and Zhang Z L (2012), “A tale of three CDNs: An active measurement study of Hulu and its CDNs,” in Proceedings of IEEE Conference on Computer Communications Workshops (INFOCOM WORKSHOPS).
3. Adler M, Sitaraman R K, and Venkataramani H (2011), “Algorithms for optimizing the bandwidth cost of content delivery,” Computer Networks.
4. Almeida J, Eager D, Vernon J, and Wright S (2004), “Minimizing delivery cost in scalable streaming content distribution systems,” IEEE Transactions on Multimedia.
5. Fan X, Weber W D, and Barroso L P (2007), “Power provisioning for a warehouse-sized computer,” in Proceedings of the 34th Annual International Symposium on Computer Architecture (ISCA’07), San Diego, California, USA.
6. GuoY and FangY(2013), “Electricity cost saving strategy in data centers by using energy storage,” IEEE Transactions on Parallel and Distributed Systems.
7. Kayaaslan E, Cambazoglu B B, Blanco R, Junqueira F P and Aykanat C, (2011), “Energy-price-driven query processing in multi-center web search engines,” in Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information (SIGIR’11), New York, NY, USA.
8. Koomey J(2011), “Growth in data center electricity use 2005 to 2010,” Analytics Press.
9. Liu H H, Wang Y, Yang Y R, Wang Y, and Tian C (2012) , “Optimizing cost and performance for content multihoming,” SIGCOMM Comput. Commun.
10. Luo J, Rao L, and Liu X (2013), “Temporal load balancing with service delay guarantees for data center energy cost optimization,” IEEE Transactions on Parallel and Distributed Systems.