e-ISSN: 2349-9745
p-ISSN: 2393-8161

Scientific Journal Impact Factor (SJIF): 1.711

International Journal of Modern Trends in Engineering and Research

www.ijmter.com

MINIMIZING CONTENT SERVICE COST IN DATA CENTRE [CDN] USING CONTENT MULTIHOMING

B.M. Brinda[1], R.Jayashree[2], N.Nanadhini[3]

[1]*Assistant professor, Dept of Information Technology, Vivekanandha college of Engineering for Women*
[2]*Dept of Information Technology, Vivekanandha College of Engineering for Women*
[3]*Dept of Information Technology, Vivekanandha College of Engineering for Women*

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 Distribution Network (CDNs), Data Centre (DC), Multihoming,Optimization, Replica Server.

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 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 FORMULATION

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 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 real-world 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.

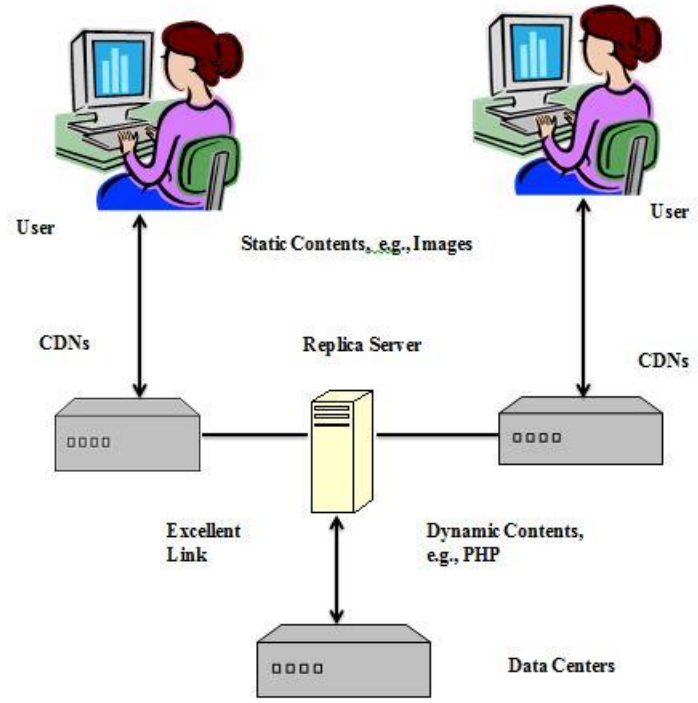

Fig 1.1 Procedures To Serve End users Request in a CDN

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 u_l requests the dynamic content data d_m through CDN c_j periodically. Usually, CDN c_j will cache the static content in their own servers, and only the dynamic content data in data center is focused here. Hence, CDN c_j will distribute its requests to origin data centers which has the data. Here it is assumed as data center n_k receives this request, and n_k will serve this request and sends the desired data d_m to CDN c j. After receiving the data, c_j will send it to u_l . In this request distribution process, the number of requests from user group u_l to CDN c_j for the data d_m is represented by the function of $f_{uc}(u_l ; c_j; d_m)$; similarly, the function of $f_{cd}(c_j; n_k; d_m)$ represents the number of requests which CDN c_j distributes to data center n_k for the data d_m.

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-multihoming 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 real-world 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-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.

V. EXPECTED RESULTS

- Provide reliable service to end users.
- Optimizing electricity cost for data centers,
- Optimizing cost for CDNs
- Satisfy both the provider and the end user.

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 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.

