A Mathematical Computational Design of Resource-Saving File Management Scheme for Online Video Provisioning on Content Delivery Networks

Dr.M.Upendra Kumar[#1], Dr.A.V.Krishna Prasad[*2], Dr.D.Shravani[#3]

[#1]*Professor Dept. of CSE M.V.S.R.E.C., Osmania University Hyderabad TS India*

[*2]*Associate Professor Dept. of CSE M.V.S.R.E.C., Osmania University Hyderabad TS India*

[#3]*Associate Professor Dept. of CSE SCETW, Osmania University Hyderabad TS India*

Abstract *This research paper discusses a mathematical computational design of resource-saving file management scheme for online video provisioning on content delivery networks.*

Keywords — *Mathematical Computational Design, File management, Online Video, Content Delivery Networks.*

I. INTRODUCTION

Content delivery networks (CDNs) have been widely implemented to provide scalable cloud services. Such networks support resource pooling by allowing virtual machines or physical servers to be dynamically activated and deactivated according to current user demand. This paper examines online video replication and placement problems in CDNs. An effective video provisioning scheme must simultaneously (i) utilize system resources to reduce total energy consumption and (ii) limit replication overhead. We propose a scheme called adaptive data placement (ADP) that can dynamically place and reorganize video replicas among cache servers on subscribers' arrival and departure. Both the analyses and simulation results show that ADP can reduce the number of activated cache servers with limited replication overhead. In addition, ADP's performance is approximate to the optimal solution.

In recent years, content delivery networks (CDNs) services have been widely implemented to provide scalable cloud. Fig. 1 shows a typical local CDN whose servers are located in the same place. The request of each visitor (A) is first processed and identified by a gateway server (B) and then directed to a cache server (CS) (C) in the server farm. CSs are typically virtual machines and can dynamically provide various services by executing different contents loaded from a backhaul database (D).

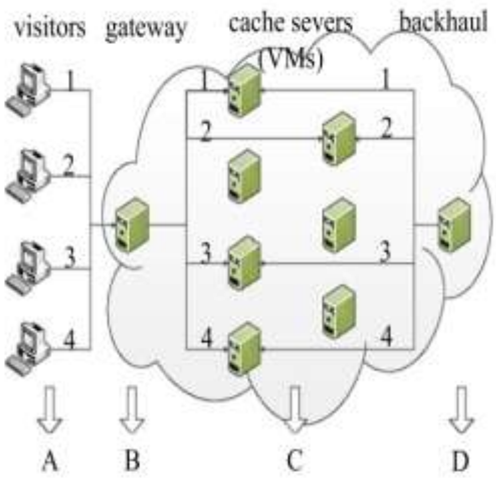

Fig. 1. Illustration of a local CDN.

This cloud-based architecture can provide a high degree of scalability and flexibility for service provisioning because it adaptively utilizes storage space, computing power, and network bandwidth by activating different numbers of CSs. The use of CSs is critical because the average loading of a single CS is typically substantially lower than its maximum capability. Therefore, minimizing the number of activated CSs is correlated to, if not equal to, minimizing the total energy consumption because of two reasons. First, supporting an activated virtual/physical machine requires considerable power compared with the dynamic workload of visitors. Second, the system resources (e.g., network bandwidth and CPU time) and power consumption required by each visitor are almost identical. Many related studies, have focused on analyzing or reducing the number of activated CSs in a CDN. CDNs are effective platforms for providing various types of services. Among them, on-demand video provisioning is a popular application, allowing numerous users to arbitrarily request videos from a massive database. Video websites such as YouTube, Vimeo, and DailyMotion are examples. When a visitor arrives and requests a video clip, the system must assign a serving CS and copy a replica of the clip from the backhaul database if the CS does not cache the clip for other visitors. Because each CS has limited capability, the total number of video clips it stores and the total outgoing bandwidth of subscribers it bears is limited by its space and bandwidth constraints.

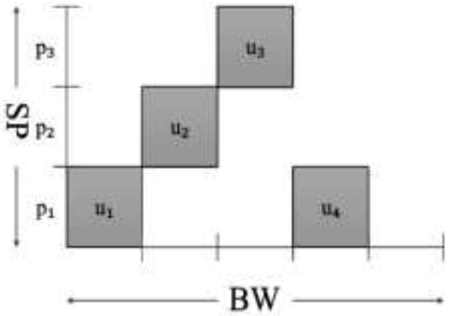

Fig. 2. Example of a CS in a CDN.

Fig. 2 illustrates an example of this two-dimensional resource allocation problem, where the length and width respectively express the space and bandwidth capacity of a CS, and u and p respectively illustrate users and their requested programs. Here, Visitors 1, 2, and 3 require Video Programs A, B, and C respectively. When User 4 is later directed to this server, an additional unit of bandwidth is required because each video is independently transmitted/played by each user. However, Visitor 4 does not occupy an additional unit of storage space because Program A has already been requested by Visitor 1. Because of the many time-variant requirements of video clips, intelligently placing videos among CSs and determine their serving subscribers without violating capacity and bandwidth limits is challenging. Typical CDN management schemes in data centers fail to address this video provisioning problem for three reasons. First, rather than being separately required by a specific user/subscriber, a video clip may be simultaneously accessed by numerous online subscribers. Second, the two resource requirements (i.e., bandwidth and storage size) exhibit different characteristics (i.e., when different visitors request the same clip, their storage requirements can be combined, whereas their bandwidth requirements are independent). Third, because video placements and server selections are conducted online as visitors arrive and depart; the migration/management of clip overhead should be limited to provide real-time responsiveness.

II. A MATHEMATICAL COMPUTATIONAL DESIGN OF PROPOSED MODEL

Fig 3 provides Class Diagram of Proposed Model. Fig 4 provides Usecase Diagram of Proposed Model

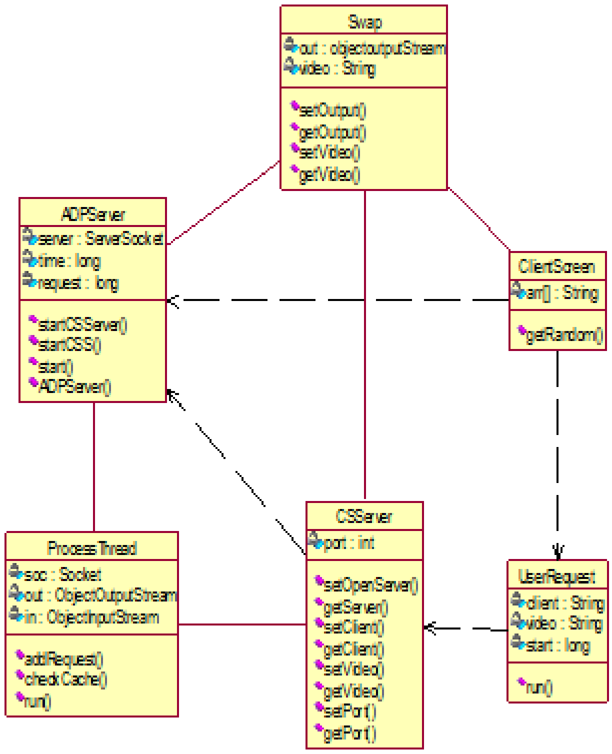

Fig. 1 Class Diagram of Proposed Model

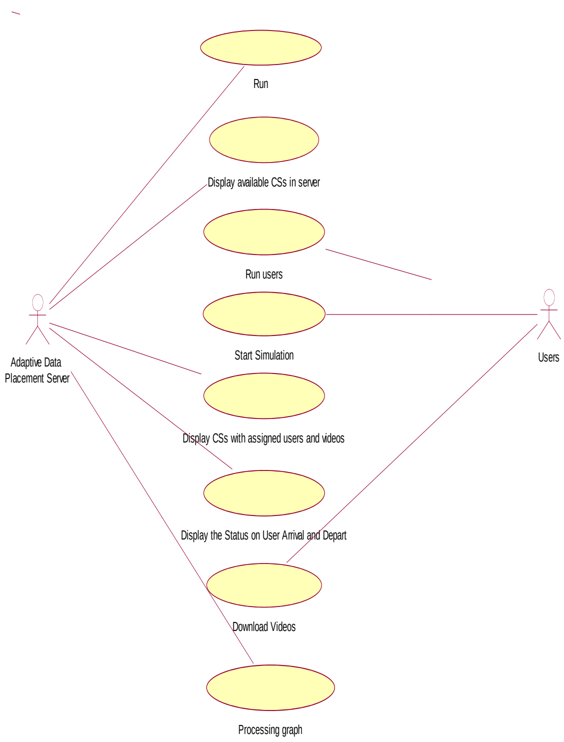

Fig. 4 use caseDiagram of Proposed Model

Fig 5 provides Sequence Diagram of Proposed Model. Fig 6 provides Component Diagram of Proposed Model. Fig 7 provides Deployment Diagram of Proposed Model.

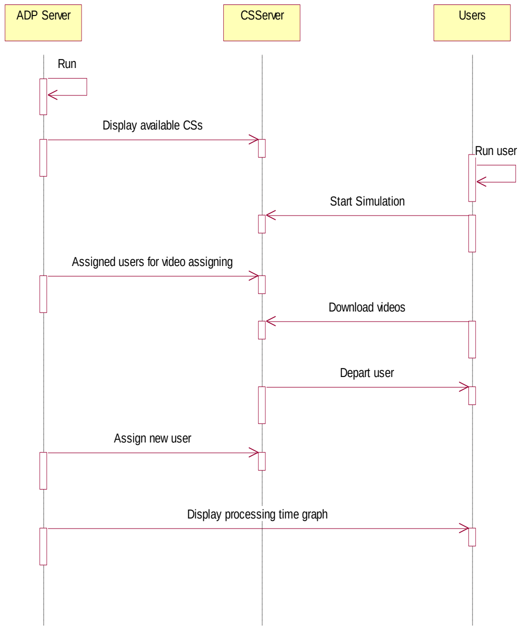

Fig. 5 sequence Diagram of Proposed Model

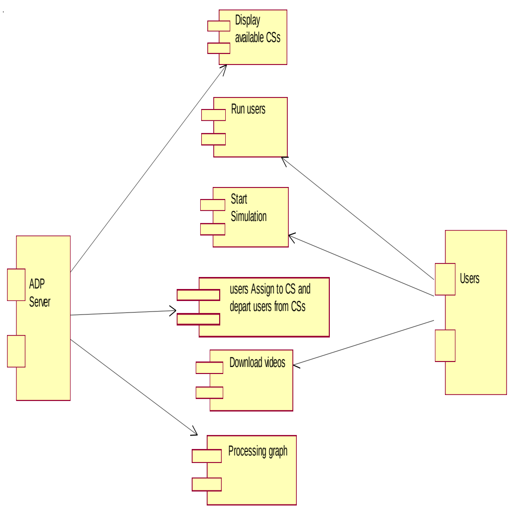

Fig. 6 component Diagram of Proposed Model

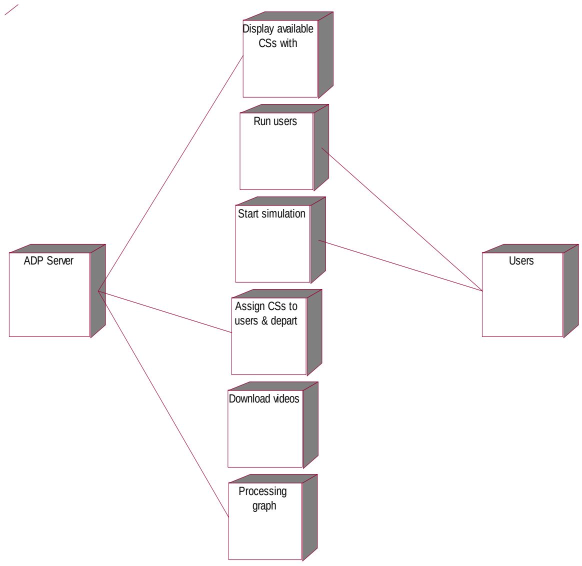

Fig. 7 Deployment Diagram of Proposed Model

Fig 8 to Fig 11 provides Execution Screen shots results of Proposed Model

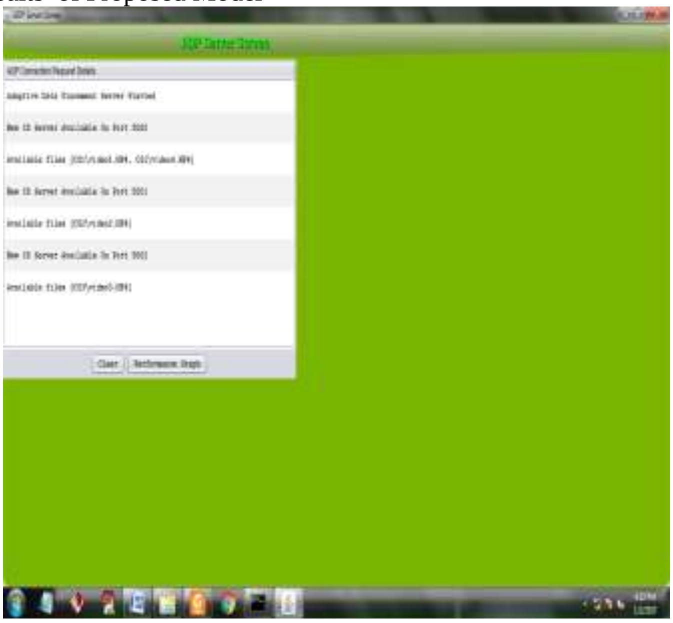

Fig. 8 Execution Screen shots results of Proposed Model

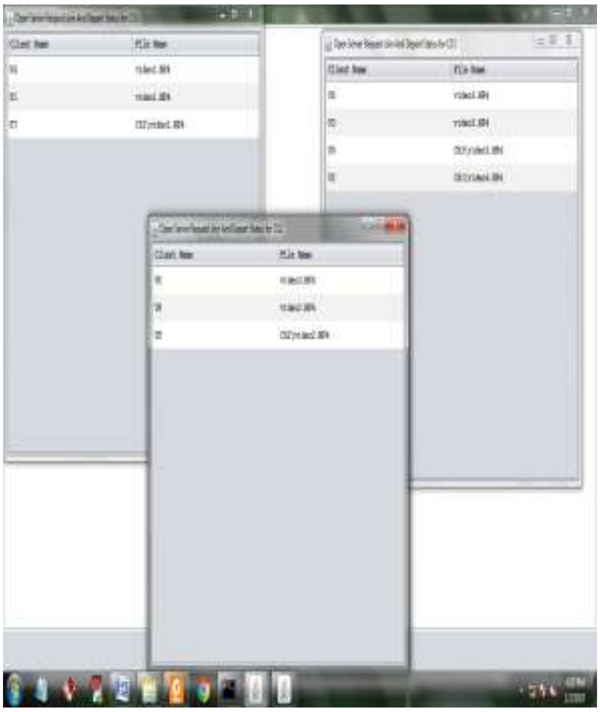

Fig. 9 Execution Screen shots results of Proposed Model

Fig. 10 Execution Screen shots results of Proposed Model

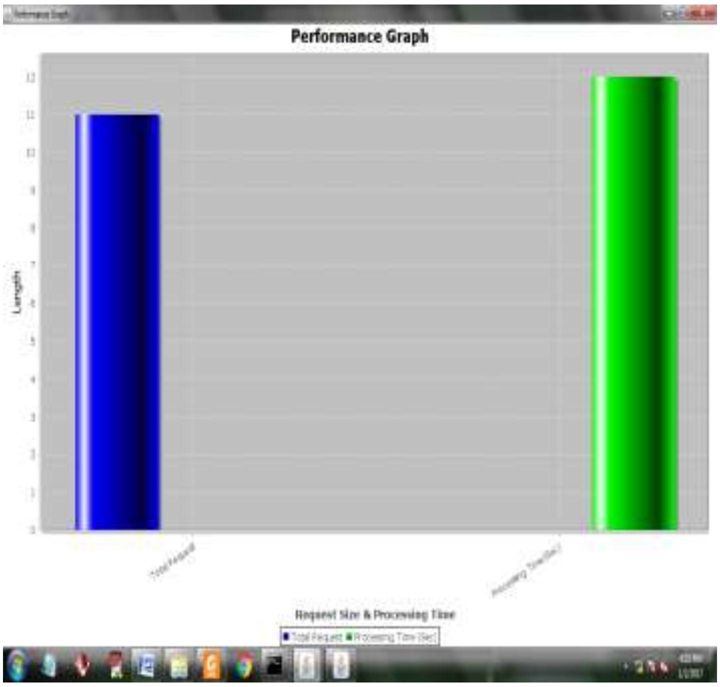

Fig. 11 Execution Screen shots results of Proposed Model

III. CONCLUSIONS

In this paper, we examine an online video placement scheme for superior utilization and energy-saving in cloud delivery networks. We introduce a new problem that dynamically places incoming video subscribers to CSs to limit the number of active machines as well as the replication overhead. This problem considers both transmission bandwidth and storage space constraints and is modeled in a general manner. It can therefore be applied effectively to various types and scales of CDNs. By classifying servers into different types, our proposed ADP scheme places and reorganizes video subscriptions on their arrival and departure. Through analysis, we demonstrate the effectiveness of ADP regarding performance and overhead. The worst-case overhead of ADP is limited, and the performance difference to the optimum is bounded. The outstanding performance of ADP is also evidenced by the simulations. The results show that ADP significantly outperforms the compared scheme under various conditions and maintains performance approximate to the optimal solution. In addition, the replication overhead of the system is also limited. To the best of our knowledge, ADP is the only scheme that addresses this placement problem and provides all the mentioned advantages. We notice that other types of approaches may effectively reduce the server number (although they might produce heavier migration overhead; e.g., by analyzing and predicting the incoming subscriptions or using machine learning to redistribute replicas).

REFERENCES

[1] L. M. Vaquero, L. R.-Merino, J. Caceres, and M. Lindner, "A break in the clouds: Towards a cloud definition," in Proc. ACM SIGCOMM Comput.Commun. Rev., 2009, pp. 50–55.

[2] L. Tsai and W. Liao, "Cost-aware workload consolidation in green cloud datacenter," in Proc. IEEE CLOUDNET, 2012, pp. 29–34.

[3] L. A. Barroso and U. Holzle, "The case for energy-proportional computing," IEEE Comput., vol. 40, no. 12, pp. 33–37, 2007.

[4] P. Bohrer, E. N. Elnozahy, T. Keller, M. Kistler, C. Lefurgy, C. McDowell, and R. Rajamony, "The case for power management in web servers," Power Aware Computing. Kluwer Academic Publishers, pp. 261–289, 2002.

[5] B. Li, J. Li, J. Huai, T. Wo, Q. Li, and L. Zhong, "EnaCloud: An energy-saving application live placement approach for cloud computing environments," in Proc. IEEE CLOUD, 2009, pp. 17–24.

[6] Y. Ho, P. Liu, and J. Wu, "Server consolidation algorithms with bounded migration cost and performance guarantees in cloud computing," in Proc. IEEE Utility Cloud Comput., 2011, pp. 154–161.

[7] M. Korupol, A. Singh, and B. Bamba "Coupled placement in modern data centers," in Proc. IEEE Int. Parallel Distrib. Processing, 2009, pp. 1–12.

[8] A. Gandhi, M. H.-BALTER, and R. Raghunathan, "AutoScale: Dynamic, robust capacity management for multi-tier data centers," ACM Trans. Comput. Syst., vol. 30, no. 4, p. 14, Nov. 2012.

[9] A. Krioukov, P. Mohan, S. Alspaugh, L. Keys, D. Culler, and R. H. Katz, "NapSAC: Design and implementation of a powerproportional web cluster," in Proc. First ACM SIGCOMM, 2010, pp. 102–108.

[10] Z. Zhou, F. Liu, H. Jin, B. Li, B. Li, and H. Jiang, "On arbitrating the power-performance tradeoff in SaaS clouds," in Proc. IEEE INFOCOM, 2013, pp. 872–880