A Novel Approach for Enhancing Performance of VoD Systems

Sukumar Ramanujam, Mahesh T R
Department of Computer Science & Engineering,
School of Engineering & Technology, Jain University,
Bengaluru, India
rsugumdevi@gmail.com, maheshtr.1978@gmail.com

Mouleeshwaran S K
Department of Computer Science & Engineering,
Dayanand Sagar University,
Bengaluru, India.
meetmoulee@gmail.com

Abstract—Multimedia applications such as Video-on-Demand (VoD), Live streaming, Internet stock quotes, Internet radio, audio/music delivery, video surveillance are of growing interest among general public. Existing systems that support these kinds of applications such as centralized server, independent server nodes, and proxy incur significant delay and serve only less number of videos. In this research, a multi-server system that utilizes a split and merge scheme is proposed to reduce the waiting time. This system helps us to achieve load balancing, while increasing the number of videos being served. Our simulation model consists of a single main multimedia server and a set of streaming servers. The performance of the proposed system for various K values is evaluated in the VoD scenario. The results show that the proposed multi-server scheme performs better in terms of initial latency and number of videos being served, compared to the other existing schemes. Index Terms—Component, formatting, style, styling, insert.

Keywords: Multi-Server System, Vod, Multicast Bandwidth, Scalability, Wireless Networks

I. Introduction

Recently, There Has Been An Increasing Demand For Video Service Applications Related To Education, Information, And Entertainment. Many Universities And Colleges Are Moving Towards The Concept Of Virtual Class Rooms And On-Line Education (H. Mauve Et Al 2003). Such Advancements In Technologies Would Not Only Reduce The Cost Of The Services Offered To The End Users, But Also Benefit People Living In Remote Places. A Recent Study Reveals That Video-On-Demand (Vod) Is A Service That Will Generate Huge Amount Of Revenue For Private Network Companies, But Which Needs Performance-Effective Architecture Solutions That Provide High Scalability.

In an on-demand video system (i.e., videos are displayed upon user request with negligible delay), a number of storage servers such as tertiary libraries or jukeboxes (collectively referred to as a repository) store all the video contents of interest to a large number of geographically distributed users (C.-C. Shen and C. Jaikaeo 2005). If videos were to be streamed directly to the users, the user capacity in the system would be limited by the streaming capacity of the repository. Such capacity can be increased by using a ladder of servers, in which multiple streaming servers cache the movies delivered from the repository and stream them to the users. If the streaming servers were co-located with the repository, the transmission cost incurred in streaming videos to remote users might be high. To overcome this problem, as well as to take advantage of the access locality and demand characteristics of the user pool, the streaming servers may be placed close to the user groups, thus forming a "multi-server architecture." Such an architecture is able to achieve scalable storage and streaming capacities by introducing more repository servers and local servers as the traffic increases.

In a wireless environment, users may enjoy ubiquitous entertainment services (T.- C. Chiang and Y.-M. Huang 2004). For example, they may watch videos or play games of their interest online wherever they are. This chapter focuses on technologies that enable (VoD) services over wireless networks. A VoD system is an interactive multimedia system working much like a cable television, the difference being that the client can select a movie from a large video database stored in a distant video server (D. A. Maltz, et al 1999). Varied from other video services such as pay-per-view (PPV) and video in demand (VID), individual VoD clients in an area can watch different programs whenever they wish to, not just in time as in VID or pre-scheduled as in PPV. A VoD system is, therefore, a realization of the video rental shop brought into the home.

II. Problem Formulation

A wireless VoD system has many practical applications (C. Gui and P. Mohapatra 2003). For instance, airlines could provide VoD services in airport lounges to entertain passengers on their own PDA (personal digital assistant) while they are waiting for a flight; a museum could provide video information on the exhibits on demand over the wireless network; in education, a university could also install such a system on campus to allow students to watch videos recorded earlier from lectures they were not able to attend.

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

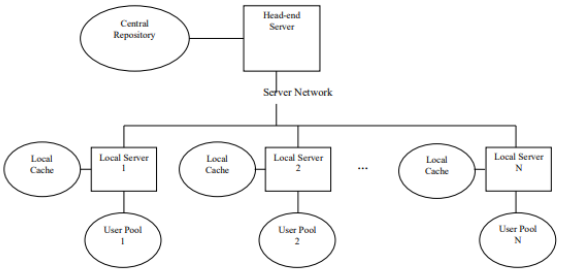


Fig. 1.Multi-server network architecture for VoD application

Depending on how the servers are operated and maintained, there are three different approaches of VoD: centralized, independent server nodes, and proxy. The centralized server approach uses only one server with a huge storage. But the drawbacks of this approach are its single point of failure and the lack of scalability (J. Wu et al 2015). The independent servers approach utilizes n different servers; each has a dedicated storage facility. But it suffers from the problem of poor fault tolerance and high storage requirement (J. J. Garcia-Luna-Aceves and E. L. Madruga 1999). The proxy approach makes use of the existing web proxies' capacity to cache media data, which is cost-effective. The scalability of this approach is bounded by the bandwidth availability of the centralized server. In a multiserver architecture the workload is divided evenly among all the servers and scalability is achieved.

III. Problem Solution

A service in the multi-server system, is often provided by one or more servers running on different machines with the service functions distributed and/or replicated - primarily for higher availability and performance. The servers are organized as members of a server group which exports the service through a generic runtime interface with the distribution and replication of the functions hidden from clients. An application is then implemented as one or more clients communicate with the server group.

Existing VoD architectures do not support, frequent member join/leave and content exchanges, due to limited processing capacity and network support (Bin Li et al 2016). Each of the local servers, therefore, receives its movies only from the repository and works independently from the other, and the repository has to unicast the movies to the local servers. Furthermore, network limitations such as disjoint networks and limited network capacity also lead to performance instability (S. Lee et al 2000). On the other hand, the multiserver approach for video streaming applications consists of a network of servers, where each maintains a dedicated local cache. This allows the local server to cache the full video as compared to the partial cache facility supported by the previous schemes. This enhances the availability and performance while minimizing the delay.

Two types of services are possible with the multi-server network model: local service and remote service (Tang D et al 2013). In a local service, a client can directly receive video information from its local server-node. Local service requires only localnetwork resources and server-node resources (E. M. Royer and C. E. Perkins 1999). If the video requested by a client is not available in the local service the remote service is used. Thus, a remote service requires the resource of an inter-connection network. In order to minimize the resource requirements of remote services, video-mapping strategies have been developed to calculate the optimal number of replicated popular videos in each node.

Table 1:Comparison of waiting time for various VoD schemes

No. of requests	Client waiting Time(s) Centralized	Client waiting Time(s) Independent server nodes	Client waiting Time(s) Proxy	Client waiting Time(s) Multi-server system with K=2	Client waiting Time(s) Multi-server system with K=3	Client waiting Time(s) Multi-server system with K=4
50	201	156	52	55	35	21
100	235	159	64	76	38	26
150	288	158	86	45	40	39
200	321	162	98	58	46	41
250	325	162	134	62	51	47

In order to reduce the delay in receiving the data from the repository always, the local servers can be connected by a network so that they can communicate and exchange their data with each other. This will also decrease the long-haul transmission cost of the requested video from the central storage to the local server. It is possible for servers to be co-located in a campus area network, with an entertainment network of cooperative service providers, or a private enterprise network. In general, the transmission cost between the local servers is low, relative to that from the repository. The local cache connected to each local server helps to buffer the requested video in advance. This, in turn, helps the server network to satisfy any subsequent local request(s) among themselves, thus saving bandwidth from the central storage to the server network. The central storage may multicast movies to the local servers, or it may unicast the movie to a specific local server, which in turn, unicasts/multicasts the movie to the other servers using the network between the local servers.

IV. Results And Discussion

In applications where there are many users and frequent additions or deletions to the group membership, key management schemes can reduce significantly the communication burden (B. O. V. Lashkari and M. Dehghan 2007). In this thesis, an attempt is made to reduce such a communication burden by improving their efficiency in the face of QoS parameters using a multi-server approach. Having met the basic design goal, there is a need to examine the performance of the multi-server approach in realistic scenarios.

Our simulation model consists of a single main multimedia server and a set of streaming servers. Each streaming server is connected to the next level of servers which are interconnected among themselves using the IEEE 802.11. The following are the assumptions made in our simulation model. The user requests for the video that follows the Zipf law

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

of distribution. The number of streaming servers in each group is assumed to be same and the sizes of the videos are uniformly distributed over a range. The values considered for simulation are as follows. The sizes of videos are 350MB to 550MB and each simulation is carried out for 1500 seconds. The performance parameters are client waiting time for the requested videos and number of videos served.

4.1 Waiting Time

Waiting time is the initial latency experienced by a customer before watching the video (Y. H. Moon et al 2013). It is certainly a significant parameter from the client's perspective, which determines the QoS in a VoD system. The performance analysis on waiting time as a function of the number of requests for various VoD schemes is shown in Fig. 2, 3, and 4.

The numerical comparison in terms of initial latency in all the strategies is shown in Table 1. Very popular videos are batched and then only served, hence incurring initial latency. A significant initial latency is observed in the graph as majority of requests are for very popular videos. It is evident from the graph that in the multi-server approach, initial latency is nearly uniform irrespective of the number of requests, whereas other schemes show a significant increase in initial latency for a large number of requests due to an inefficient buffer management technique implemented in the local server.

Table 2:Comparison of Number of videos served for various VoD schemes

Time(s)	Centralized	Independent server nodes	Proxy	Multi-server system with K=2	Multi-server system with K=3	Multi-server system with K=4
30	35	31	35	54	55	61
60	41	33	41	59	64	65
90	46	47	55	61	68	71
120	49	66	79	62	99	112
150	62	95	106	89	154	134
180	81	156	182	168	198	218
210	135	201	212	196	201	257
240	199	255	270	245	278	298
270	201	289	301	285	325	355
300	225	321	356	311	329	413

A multi-server system achieves negligible initial latency by means of local streaming servers. This advantage, however, comes with the cost of additional servers and storage.

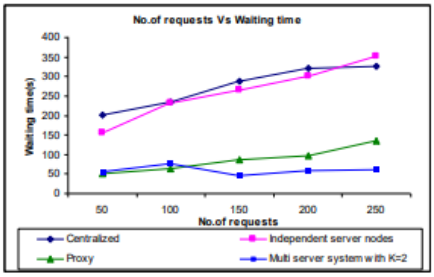


Fig. 2.Waiting time for multi-server system with K=2

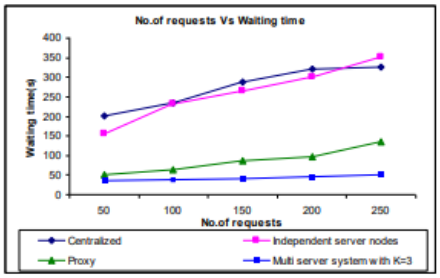


Fig. 3.Waiting time for multi-server system with K=3

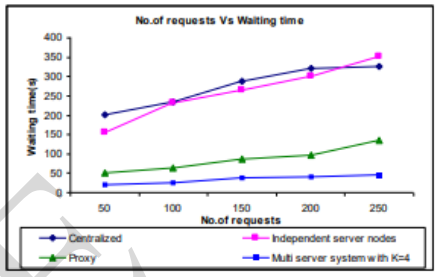


Fig. 4.Waiting time for multi-server system with K=4

4.2 Number Of Videos Served

The performance analysis of the number of requests serviced as a function of time for various VoD schemes is shown in Fig. 5 through 7. In the earlier stages (till 30-60 seconds), all the techniques serve the requests and none of them has completed the service. Later, the number of requests served grows linearly in all the schemes. Since local streaming servers are used, more number of requests are processed in the multi-server system and hence achieve significantly greater throughput than all the other schemes. Requests are served faster using the available resources and the resources which enable other requests to be served are released. For example, at time 270, in the multi-server scheme with K=4, over 250 requests are served, when compared to the other scheme where only 135 requests are served, which clearly indicates that in our approach more than 90% of the requests are served when compared to the centralized server scheme.

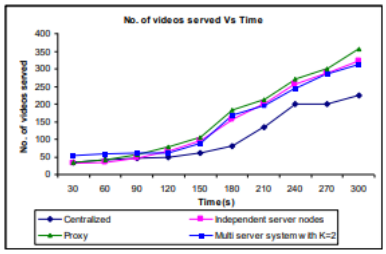


Fig. 5.Number of videos served for multi-server system with K=2

IJARE

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

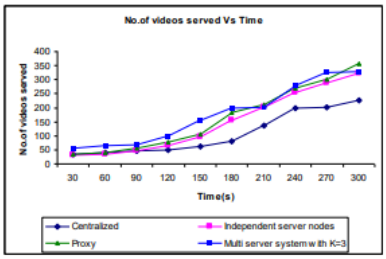


Fig . 6 Number of videos served for multi-server system with K=3

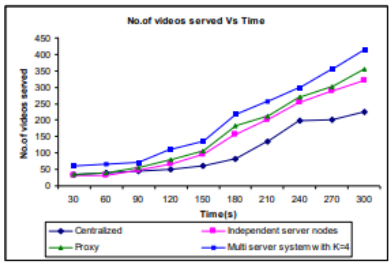


Fig. 7 Number of videos served for multi-server system with K=4

V. Conclusion

Thus the performance of the multi-server system for various K values as the evaluated in the VoD scenario. The performance analysis on waiting time as a function of the number of requests for various VoD schemes has been studied. The performance analysis of the number of requests serviced as a function of time for various VoD schemes has also been studied. The results have shown that the proposed multi-server scheme achieves negligible initial latency by means of local streaming servers. This advantage, however, comes with the cost of additional servers and storage.

References

[1] Bin Li, Hongxiang Li, Ruonan Zhang, 'Adaptive Random Network Coding for Multicasting Hard-Deadline-Constrained Prioritized Data', IEEE Transactions on Vehicular Technology, Vol. 65, No.10, pp. 8739-8744, October, 2016.

[2] Y. H. Moon, K. S. Yoon, S.-T. Park, I. H. Shin, "A new fast encoding algorithm based on an efficient motion estimation process for the scalable video coding standard", IEEE Trans. Multimedia, Vol. 15, No. 3, pp. 477-484, April, 2013.

[3] D. A. Maltz, J. Broch, J. Jetcheva, and D. B. Johnson, "The effects of on-demand behavior in routing protocols for multihop wireless ad hoc," IEEE Journal on Selected Areas in Communications, Vol. 17, No. 8, pp. 1439-1453, August, 1999.

[4] B. O. V. Lashkari and M. Dehghan, "QoS-Aware multicast Ad hoc on-demand distance vector routing," in Proc. World Congress on Engineering, London, U.K., WCE 2007, July 2-4, 2007, pp. 1506-1511.

[5] E. M. Royer and C. E. Perkins, "Multicast operation of the ad hoc on-demand distance vector routing protocol," in Proc. MobiCom' 99, Seattle, Washington, Augugst 15 – 20, 1999, pp. 207- 218.

[6] C. Gui and P. Mohapatra, "Efficient overlay multicast for mobile Ad Hoc networks," in Proc. IEEE WCNC, New Orleans, LA, Mar. 2003, pp. 1118 – 1123.

[7] H. Mauve, H. Füßler, J. Widmer, and T. Lang, "Poster: Position Based multicast routing for mobile ad-hoc networks," in Proc. Fourth ACM International Symposium on Mobile Ad Hoc Networking and Computing: MobiHoc, Germany, 2003, pp. 53-55.

[8] J. J. Garcia-Luna-Aceves and E. L. Madruga, "The core-assisted mesh protocol," IEEE Journal on Selected Areas in Communications, Vol. 17, No. 8, pp. 1380-1384, August, 1999.

[9] Tang D., Lu X., Li J., "Multicast Routing Algorithm Based on Network Coding", In: Liu D., Alippi C., Zhao D., Hussain A. (eds) Advances in Brain Inspired Cognitive Systems. BICS 2013. Lecture Notes in Computer Science, vol. 7888, Springer, Berlin, Heidelberg, 2013, pp. 348-357, 9-11 June 2013.

[10] S. Lee and C. Kim, "Neighbor supporting ad hoc multicast routing protocol," in Proc. 1st ACM International Symposium on Mobile Ad Hoc Networking and Computing, Mobihoc '00, Boston, Massachusetts, USA, 2000, pp. 37 – 44. [11] C.-C. Shen and C. Jaikaeo, "Ad hoc multicast routing algorithm with swarm intelligence," Mobile Networks and Applications, Vol. 10, No. 1-2, pp. 47-59, February, 2005.

[12] T.-C. Chiang and Y.-M. Huang, "Multicast routing representation in ad hoc networks using fuzzy petri nets," in Proc. IEEE 18th International Conference on Advanced Information Networking and Applications, AINA 2004, Fukuoka, Japan, 29 – 31, March, 2004, pp. 1-4.

[13] J. Wu, B. Cheng, C. Yuen, Y. Shang, J. Chen, "Distortion-aware concurrent multipath transfer for mobile video streaming in heterogeneous wireless networks", IEEE Trans. Mobile Comput., Vol. 14, No. 4, pp. 688-701, April, 2015.

IJARE

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)