An Efficient CDN Placement Algorithm for the Delivery of High-Quality TV Content

Adrian J. Cahill
Department of Computer Science
University College Cork
Cork, Ireland
cahill@cs.ucc.ie

Cormac J. Sreenan
Department of Computer Science
University College Cork
Cork, Ireland
cjs@cs.ucc.ie

1. ABSTRACT

Personal Video Recorders (PVRs) such as TiVo have become very popular in recent years due to their ability to intelligently record TV content and make it accessible in an on-demand fashion to its users. Our research looks at providing a globally accessible storage architecture where all content broadcast over a period of time is available for streaming. Our architecture consists of idle ISP servers, that can be rented and released dynamically as the load requires. In this paper we focus on managing the resources required to provide TV on Demand (TVoD) system, and develop a cost function which examines the resources required to serve a client and identifies the optimal proxy to serve the client.

Categories and Subject Descriptors: H.3.2: Information Storage

General Terms: Algorithms, Management, Economics

Keywords: Video on Demand, Resource management, P2P management

2. PROBLEM STATEMENT

There are a number of problems associated with delivering high-quality TVoD to a large number of users. (i) Startup Latency; clients will not be satisfied if the video does not begin streaming almost immediately, (ii) Metadata; suitable metadata will need to be stored along with the content to ease search and retrieval; and most importantly (iii) Resource Management; due to the size of these high-quality video files, and the time taken to stream them to the users, server resources will become a bottleneck, possibly resulting in Denial Of Service for some clients.
Due to space limitations we will be focusing on the resource management problem.

2.1 Resource Management

System Resources such as server storage space, disk and network I/O have always been an important factor in determining the scalability of a media distribution architecture. One of the key differences with the distribution of web documents (HTML pages and images) is the relatively short time the server resources are required in the delivery process, unlike streaming large video objects which require server resources to be tied-up for much larger periods of time. Optimizing the resources required to serve a number of clients involves optimizing (i) the content placement and (ii) the number of replicas of an object. But these parameters are influenced by the client request patterns which are constantly changing, therefore it will be necessary to frequently re-evaluate the resource usage.

MM'04, October 10-16, 2004, New York, New York, USA.
ACM 1-58113-893-8/04/0010.

3. OUR CONTRIBUTION

In this section we briefly look at using Content Distribution Networks (CDNs) and Peer-to-Peer networks (P2P) as possible distribution architectures. We identify limiting factors with both of them and propose a hybrid architecture to reap the benefits of both architectures. We re-examine the resource management problem, providing a solution in the form of a cost function which only places content on a server if it results in reduced resource usage.

3.1 Design Considerations

CDNs relieve bottlenecks in the Internet by replicating popular content to a number of surrogate servers, thus distributing the client load and providing replicas close to clients which results in lower latencies for client. Extending a CDN's resources is typically achieved by adding new hardware to the system, though recent work by the IETF [1] has looked at developing a framework for Content Internetworking(CDI) to facilitate the offloading of client request in the event of overloading of servers. We believe that if CDNs are to replace the existing TV broadcasting architecture, it should include an approach similar to CDI to deal with unexpectedly large loads such as those that may occur at a time of a big sporting, for example *The World Cup.*

Another popular distribution scheme is P2P networks. P2P networks such as bitTorrent have become very popular in recent months and are used extensively for the distribution of large objects for example movies and *CD Image files* which are typically in the range of 400-700MBs. Distributing streaming objects over a P2P network (clients act as peers) has a number of drawbacks (i) latency (ii) changes in the distribution tree and (iii) retrieval of unpopular content can be difficult. Many of the existing P2P implementations suffer from digital rights issues.

In previous work by the authors [2] a Video Content Distribution Network (VCDN) architecture is proposed. The VCDN architecture is a hybrid CDN-P2P network, utilizing the dynamic nature of a P2P architecture while providing a CDN service on-top of this overlay network. In doing this we hope to overcome the limitation of CDN extendability, while also minimizing the overall resources required to serve a given client request pattern. Idle ISP servers can advertise their willingness to partake in VCDN by acting as peers, for which micro-payments will be made according to the server resource usage. It is hoped that by agreeing to partake in the CDN, servers will not leave the CDN without providing adequate notice, thus allowing the system to find an alternative source for the interrupted clients.

3.2 Resource Management

In this section we outline the costs we believe to be influential in the placement of TV content within the VCDN, and develop a cost function which minimizes these resources while serving a given set of client requests.

Streamed Network Cost β_r A cost associated with *streaming* video content from video proxy *'r'*. For the purpose of our research we consider this to be a cost unit, per byte, per hop.

Bulk Transfer Network Cost $\alpha_{s,d}$ A cost associated with the transfer of a video object from video proxy *'s'* to proxy *'d'*. For the purpose of our research we consider this to be a cost unit, per byte, per hop. This is considered separately to *streamed network cost* as this provides a fine grained cost model which facilitates a different pricing depending on the network requirements of the connection, for example QoS requirements.

Storage Cost λ_d A cost associated with storing content on video server *'d'*. For the purposes of our research we consider this to be a cost unit, per byte, per unit time.

Each cost outlined above could be set for each individual proxy, thus allowing the creation of a competitive model whereby ISPs compete to become active video distributors. This also provides a means where proxy location and specifications can influence their value.

3.2.1 VCDN Cost Function

We have developed a cost function which factors in the resources required to serve a given client request and determines which server requires the least resources to serve the client. Our work focuses on clusters of clients rather than individual clients defined as follows: **A cluster is a group of clients located within the same region of the network, who are all viewing the same content**

The cost of using video server (V_d) to serve 'N' clients is given by the following equations:

$$RC_{s,d,m} = \alpha_{s,d} * \delta_{s,d} * S(m) \quad (1)$$

$$SC_{d,m} = \lambda_d * S(m) * T_{max} \quad (2)$$

$$C_{s,d,m,N} = [RC_{s,d,m}] + SC_{d,m} + \sum_{n=1}^{N} (\beta_r * \delta_{d,c} * B_n) \quad (3)$$

where B_n can be calculated as follows, if the client has just started then B_n = S(m), otherwise B_n = S(m) - *bytes already served.*

V_i	Video Server *'i'*
$S(m)$	Filesize of movie *'m'*
B_n	Bytes remaining to be served in cluster *'n'*
$\delta_{src,dst}$	Num of network hops between *'src'* and *'dst'*
λ_r	Cost of storing 1 byte on Server *'r'* (per unit time)
T_{max}	Time required to server the most recent request for this object
$RC_{s,d,m}$	Cost of replicating movie *'m'* from server *'s'* to server *'d'*
$SC_{d,m}$	Cost of storing movie *'m'* on server *'d'*
$C_{s,d,m,N}$	Total cost of serving movie *'m'* to all *'N'* clients from server *'d'*

Figure 1: Parameters to Cost Function

By periodically executing this cost function at a proxy and altering content placement based on its results, we can ensure that the system always stays in a cost effective state.

4. SUMMARY & FUTURE WORK

In this paper we outline some of the problems with implementing a TVoD architecture, focusing on the resource management problem. We proposed a hybrid CDN-P2P architecture which can dynamically alter the set of video servers to reduce/increase resources depending the client load. To achieve this, a cost function is outlined which calculates the resources required to serve a set of clients using each proxy, and selects the proxy requiring the least resources.
We are currently implementing the VCDN architecture within the Network Simulator (NS2) with the intent to compare the resource usage of VCDN with other media distribution approaches. This is currently work in progress.

5. ACKNOWLEDGEMENTS

The authors would like to thank AT&T Labs-Research for supporting this project. We would also like to thank the anonymous reviewers for their helpful feedback.

6. REFERENCES

[1] M. Day, B. Cain, G. Tomilson, and P. Rzewski, "A model for content internetworking," work in progress, www.ietf.org, May 2002. http://www.ietf.org/internet-drafts/drafts-ietf-cdi-model-02.txt.

[2] A. J. Cahill and C. J. Sreenan, "Vcdn: A content distribution network for high quality video distribution," in *Proc. Information Technology & Telecommunications, Letterkenny, Ireland, Oct 22-23*, 2003.