MSAP - a Protocol for Efficient Multimedia Content Search and Distribution on the Internet

Qiang Nelson Yin Chung-Horng Lung
{nqyin, chlung}@sce.carleton.ca
Department of Systems and Computer Engineering
Carleton University, Ottawa, Canada

Abstract—**More and more music or movie consumers are using Internet to download multimedia content. Consumers have more to consider than just the name of the artist when they buy multimedia contents online. With thousands of web sites that are available on the Internet, it is time consuming for consumers to obtain the desired multimedia content, starting from the beginning of searching a title to the end of placing a purchase order. In this paper, we propose a protocol MSAP (Multimedia Service Announcement Protocol) that enables consumers to search and download multimedia content more effectively and faster. It also helps to establish a secure channel for the producers of the multimedia content to distribute their original multimedia contents to consumers. We have implemented the initial MSAP prototype in our test bed and the results indicate that our approach is efficient and accurate.**

Keywords-MSAP, Multimedia Service Announcement Protocol, Multimedia Content Search, Multimedia Content Distribution

I. INTRODUCTION

Nowadays having a hi-speed Internet connection means you can find and download files fast. Users with cable, DSL or high speed connection to the Internet will have the best set up to download music or movie from the Internet. For an Internet user, it seems that it is very easy to do so, because all they have to do is just to run their Internet browser and use Multimedia Search Engine, such as Singingfish, to search multimedia contents. However, the Internet also creates new challenges for multimedia content retrieval.

The biggest problem facing users of web search engines today is the quality of the results they get back, since search algorithms that rely on keyword matching usually return too many low quality matches [1]. As a result, the number of items returned is huge and many of them actually only provide links to the content provider's web site so that users have to verify the authenticity or accuracy and then go through a purchasing process before downloading anything. One approach to mitigate the problem is to enhance the search algorithm. However, the fundamental challenge is that the multimedia contents are scattered all over the Internet and there are a large number of possible variabilities, e.g., content providers, identical or similar titles, multimedia formats (MP3, WMP, RealOne, and QuickTime), and so on.

Our goal is to develop a network solution that can be used to effectively and efficiently search and download multimedia content. Not much research has been conducted in this area, especially from the network solution perspective. Our solution includes an architecture and a new supporting protocol called MSAP. The essence of MSAP is to enable multimedia content providers to announce their presence and attributes of their services and then advertise the information to multimedia distributors who will eventually deliver the information to multimedia consumers. In our research, we have defined the entities and components that are involved and their interactions. We believe our solution not only is faster, but also more accurate, reliable and scalable.

This paper is organized as follows. Section 2 identifies and illustrates the pitfalls of search engines. Section 3 introduces MSAP protocol, which addresses many problems of existing search engines. Section 4 presents the implementation of the MSAP in a network test bed environment. Lastly, the paper ends with a short conclusion.

II. CHALLENGES OF MULTIMEDIA CONTENT SEARCH

Currently, the amount of multimedia content on the Internet as well as the number of inexperienced users for search is growing rapidly. The most commonly used method to search and determine the location of multimedia content, for example, a song or a movie, is to use the search engine such as yahoo or multimedia content search engine like Singingfish.

Typically pull mode is used for the multimedia content search. In this model, an Internet user starts with the search of a title or/and an artist's name, trying to pull the information out of Internet and to discover the websites containing the title or the relevant hyperlinks. After screening all the search results or obtaining the services, user will determine if he will be a subscriber of an ICP (Internet Content Provider) or not. Generally, the whole process will take about 10-30 minutes to accomplish.

Reference [3] points out that a typical multimedia search engine has the abilities to: a) Select media types; b) Select, deselect and prioritize domains; c) Expose a variety of categories of interest (music, news, sports, etc.); d) Provide custom results ordering; and e) Add new media types as they are developed. However, we identified that a typical search engine also has the following pitfalls:

- It doesn't have the ability to verify if the results are reliable and accurate. Most of the search results are not instantly downloadable. To illustrate the pitfall, using Singingfish [2], we started our experiment with executing 8 search test cases. For 8 test cases, there are totally 282,871 search results. After verifying 20 top returned items for each case, more than 20% of the search results are actually links to the content provider's web sites from where end user has to repeat some steps to narrow down the scope of the search, e.g., the price and format etc, before finally determining which one to go for. Also,

about 34% are completely different multimedia content with the same title. Some results are invalid anymore. For example 6% of them are broken links because some content are only available for a period of time and they are obsolete by the time they get searched. Only 1.3% are instantly downloadable, but they are low quality or pirated contents.

- The current model is time consuming and inconvenient for end user. End users have to go through many hyperlinks and purchasing process in order to purchase or download multimedia content due to the accuracy and reliability of the search results. Users have to repeatly verify the results until they can actually place an order. Further, it requires a few minutes for users to fill out the personal and payment information. From the beginning of search to the start of download, it takes approximately 10-20 minutes. For a subscriber of a content provider, the entire process may be shorter. The shortcomings with this approach include: a) Users have to determine a content provider. b) A content provider has limited capacity. c) The length of time for downloading and uninterrupted downloaded process cannot be guaranteed.
- Other pitfalls of multimedia search engines also included: a) Web site registration and web site placement techniques can determine the order of results being presented by search engines [4]. b) It does not present the search results with pricing and content provider's information. c) It does not validate the authenticity of the content,

Due to the length of this paper, we only discuss the first two problems identified.

III. MSAP

Using SIP (Session Initiation Protocol) [5] as a reference mode, we propose a protocol called MSAP (Multimedia Service Announcement Protocol). It is a service announcement protocol for Internet multimedia content providers to advertise the multimedia contents and supports query management on the Internet. MSAP is a peer-to-peer application-layer protocol that can be used to discover, acquire, and advertise the multimedia contents by multimedia content distributors and consumers. Content advertisement allows valid multimedia content information to be carried across networks. MSAP has the following key capabilities:

- Advertise the multimedia content information from different content sources to a defined group of Internet users.
- Search and determine the target location of the multimedia content source. Handle the launch and termination of quiries and requests.
- Provide generic multimedia content information of the source location via Simple Multimedia Description Protocol (SMDP) [6]. SMDP describes the multimedia content that content providers wish to provide to consumers, such as the price to purchase and the format of content.
- Determine the authenticity, availability of the multimedia contents and the source locations. If a query cannot be successfully processed due to unauthentic source location or location unavailability, MSAP will determine if there are other available sources for the same contents and if it is possible to redirect the query.
- Determine the originating and target end points to establish a secure session between the content provider and consumers. MSAP will signal the lower layer protocol, such as RSVP to establish a secure or non-secure session between the end points.

Detailed description of MSAP can be found in [7].

A. Entities of MSAP

MSAP contains three entities, as shown in Figure 1.

- Multimedia Service Providers (MSPs) are those ICPs or ASPs (Application Service Providers) who are specialized in producing or providing multimedia contents. MSP uses the SMDP to describe their available multimedia contents and propagates that information to MCDs (Multimedia Content Distributors).
- Multimedia Content Distributors (MCDs) are responsible for advertising multimedia contents, discovering any contents requested by consumers, creating a secure channel to consumers, as well as delivering the content to consumers on behalf of MSPs. MCDs authenticate the source, authenticity, availability, location of those multimedia contents and then advertise a group of registered consumers. And they also response to any query and request from MCCs and determine the best MCP for the interest of consumer. For MCDs, MCPs are provider clients and MCCs are called consumer clients.
- Multimedia Content Consumers (MCCs) are end-users who register to the MCD and wish to obtain multimedia content. They can customize their queries and send them to the MCD. They can select the wanted multimedia content and decide from whom they will get the multimedia content.

B. Architecture of MSAP

From the architecture aspect, the physical components of the MSAP protocol can be grouped into two categories: clients and servers. There are two types of clients in MSAP: MSP Provider Clients and MCC Consumer Clients. MSAP servers include proxy server, redirect server, registry server as well as gateway server. In our research, different types of servers and clients, their roles and interactions have been identified. Scenarios that cover multiple servers and related message types have been identified and exercised. Detailed discussion is presented in [7]. Figure 2 illustrates the architecture of MSAP.

C. MSAP Components

MSAP is a peer-to-peer protocol. The peers in a MSAP session are called User Agents (UA). A client agent can function in one of the following roles:

- Consumer Client Agent (CCA): A client application that initiates the MSAP query/request for multimedia content.
- Provider Client Agent (PCA): A client application that constantly provides MCD the information of their

multimedia content. They provide or may request content distributor the information they wish to distribute.

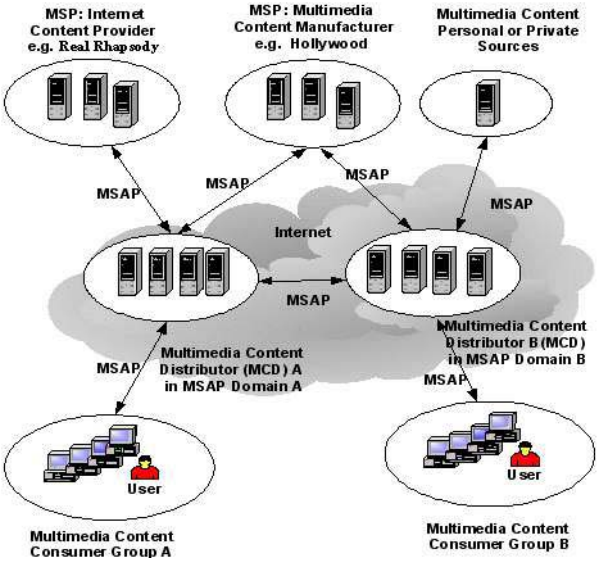

Figure 1. Entities in MSAP

- Distribution Server Agent (DSA): A server application that contacts the consumer when an MSAP query and request are received and returns a response or provides consumer the content on behalf of the MSP. Content distributor uses DSA.

A MSAP end point is capable of working as both a PCA and a CCA, but functions only as one or the other per transaction. Whether the endpoint functions as a PCA or a CCA depends on the UA that initiated the request. Provider Client can act only as PCA and consumer client can act as either a CCA or PCA.

D. How MSAP Works

MSAP is a simple protocol above TCP or UDP layer. It uses requests and responses to establish communication among the various components in the network. The MSP provider client utilizes the MSAP messages to update MCD their content information and advertise news. MCD servers will be distributing the MSP's endpoint information, multimedia content news to MCCs as well as response to MCC's request. MCCs initiate the MSAP queries for content search and requests for content's downloading or real-time streaming. The following nine steps illustrate some typical MSAP messages running between PCA and DSA:

1) A MSP PCA initiates a multicasting GNS (Get Nearest Server) message trying to find the nearest proxy server from a MCD in a distribution domain.
2) The MCD's DSA replies with the "nearest" proxy server's address and a list of other proxy servers that are available in the same distribution domain. The primary proxy server will be on the top of the list. The second "nearest" server will be the backup proxy server for a content provider.
3) From the list of proxy servers, PCA requests to join the primary proxy server. After being authenticated, PCA will be confirmed whether the join request is approved by the proxy server. A session will be established and maintained between PCA and DSA.

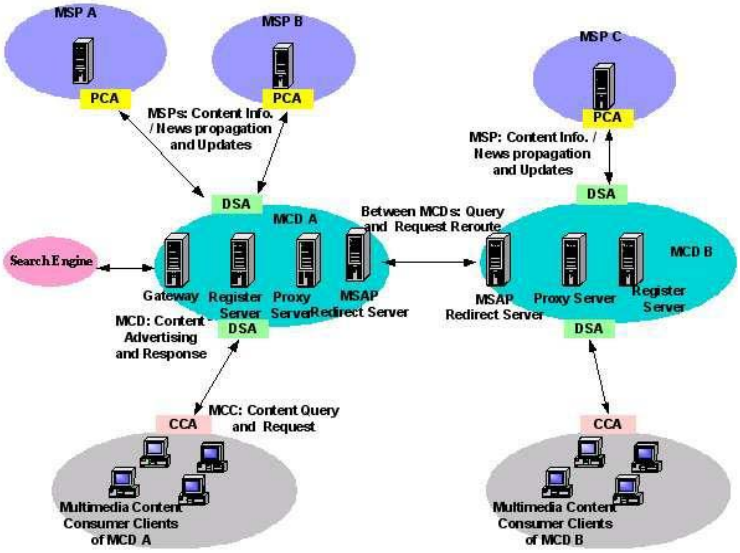

Figure 2. Architecture of MSAP

4) PCA sends the IMCA (Initial Multimedia Content Advertisement) message to MSAP proxy server, announcing the available multimedia services to proxy server.
5) The server replies with IMCA Confirm message after receiving IMCA message. And it will use the IMCA to construct a MST (Multimedia Service Table). Essentially, MST is a database in which every data entry contains generic information of a movie or song (e.g. the price, format , size, producer, artist and title) provided by a MSP and information of MSP's (e.g. protocol and endpoint to download, the total number of movies and songs). Reference [6] provides detailed information about MST.
6) Afterward, any change to the available content such as content server's URL, password and removal of the content will be carried by MSA (Multimedia Content Advertisement) message and be advertised to proxy server to update the MST. Also, any information, such as a movie news that MSPs wish to advertise, will be sent to the proxy server with the incremental MSA message.
7) MSAP server replies with "Update Confirm" message after updating the MST.
8) If a MSP wants to leave the distribution domain, the PCA initiates a "Leave" request and it will be confirmed by the MSAP server with the "Leave Confirm" message. The session between PCA and DSA will be torn down.

A consumer client in MSAP is identified by unique MSAP address defined in [6]. Consumer clients register with a register server using their assigned MSAP addresses in a distribution domain. The register server provides this information to the proxy server upon request. Using the CCA, a Consumer Client is able to:

- Fully customize a query and request, for example, a movie title, format, desired price and when to be advertised.

- Use pull mode query. After a consumer client initiates a query, a MSAP request containing information of the client query is sent to a MSAP server. MCD will reply to consumer with all results that exactly match query.
- Use push mode query. Consumer clients also can set up policy-based query and the proxy server will handle consumer's query on behalf of distributor. Once those policies are satisfied, the results of consumer's query will be generated and pushed back to CCA.
- Download the content instantly from CCA's MST. All query results will be displayed in CCA's MST and consumers just need to choose an entry for download.

The next six steps illustrate a typical communication process after a session is set up between CCA and DSA:

1) A consumer defines a query and CCA sends the query to MCD .
2) MCD abstracts all entries that satisfy a query from the proxy server's MST and replies to the consumer with the results that will be presented in CCA's MST. From there, the consumer selects the desired content and a MSP. CCA requests the location (e.g., HTTP address) and protocols such as FTP, RTSP or MMS to communicate with MSP.
3) MCD replies to CCA with the MSPs' video server's location and communication methods.
4) MCD verifies the MCC and MSP's endpoints for communication. The communication channel will be established upon the request of consumer or by MCD.
5) The communication channel between two endpoints is established and the consumer will obtain the content. Finally the channel will be closed after download is completed.
6) CCA uses a "Report" message to report the whole transaction to the MCD.

E. MSAP's Benefits

Compared to those multimedia search engines, MSAP has the following advantages:

1. Search results are more accurate and reliable. That is because all the search results for MCCs are actually abstracted from proxy server's MST, which is kept updated by MSA messages in real time. All the data entries in MST are presented concisely and simply by CCA. They will be 100% reliable and instantly downloadable.
2. Download process is faster. Compared to the download process discussed in section 2, there is no loop in the whole process for both MSAP's first time users and repeated users. The first time MSAP users will be propagated with all MSDs that are available in their local home MSAP domain. With CCA, they can select and subscribe to one or all MCDs before they start to define a query. It only takes about 5 minutes for the process.
3. More flexibility for consumers, since MSAP supports both pull and push modes. The combination of pull mode and push mode enables MCCs to tailor queries and set up policy-based query to receive multimedia news or DSA's response to a query. Moreover, push mode allows MCDs to multicast news or update information to consumers.
4. Other advantages of MSAP also include: a) MSAP is control protocol for the lower network protocols to establish secure tunnels between provider and end-user. b) MSAP is more scalable, since it uses peer-to-peer technology and it allows content providers to share their databases and offers consumers more choices. c) Utilizing SMDP, MSAP also helps to authenticate and promote the originality of the multimedia content.

IV. IMPLEMENTATIONS OF MSAP

The implementation of the entire protocol is divided into a few stages. We developed the prototype of MSAP and some simplest test cases to verify the feasibility and functionality of MSAP[8]. At the current stage, our implementations included:

- Construction of two MSPs databases that have 100 data entries and the propagation of MSA messages.
- Construction of MST table in a MCD proxy server.
- Retrieval of the CCA's MST table and verification of the accuracy of query results and query process.
- Validations of MCC query and download processes as well as query and request's performance measurement.

Next, we are going to compare the performance between the search engine and MASP from the user perspective, and the scalability and robustness of MSAP. We plan to increase the size of MST and the database of MSP from 100 to 10,000, generate queries in thousands from the MCCs simultaneously and also add one more MCD proxy server to test the redirection of query.

V. CONCLUSIONS

This paper presented a new application-layer protocol, MSAP, for efficient multimedia content search and distribution. The paper presented the architecture of MSAP, entities and components of MSAP, the basic mechanisms of MSAP, and the interactions between components of MSAP. We defined messages that dynamically inter-change between MSAP entities. Those messages are used to make the discovery and delivery of multimedia content more accurate and quickly. We also set up a test bed to implement MSAP.

REFERENCES

[1] S. Brin and L. Page, " The Anatomy of a Large-Scale Hypertextual Web Search Engine", http://www-db.stanford.edu/~backrub/google.html, Computer Science Department, Stanford University, 2002

[2] www.singingfish.com

[3] http://www.marketleap.com/singingfish

[4] B. Comings, "Achieving Good Placement in Search Engine Results", http://www.spyrock.com/, 2002

[5] M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg., RFC 2543 "SIP: Session Initiation Protocol", www.ietf.org, March 1999

[6] N. Q. Yin and C.-H. Lung, "SMDP- Simple Multimedia Description Protocol", Technical Report SCE-04-10, Dept. of Systems and Computer Eng., Carleton University , 2004

[7] N. Q. Yin and C.-H. Lung, "MSAP-Multimedia Service Announcement Protocol Specifications Version 1", Technical Report SCE-04-09, Dept. of Systems and Computer Eng., Carleton University, 2004

[8] N. Q. Yin and C.-H. Lung, "Alcatel Broadband Lab Configuration Guide", Dept. of Sys. and Comp. Eng, Carleton University, 2003