Online video using BitTorrent and HTML5 applied to Wikipedia

Arno Bakker*, Riccardo Petrocco*, Michael Dale†, Jan Gerber, Victor Grishchenko*, Diego Rabaioli*, Johan Pouwelse*
* Department of Software Technology
Delft University of Technology
Mekelweg 4, 2628 CD, Delft
The Netherlands
Email: j.a.pouwelse@its.tudelft.nl
† Wikimedia Foundation Inc.
149 New Montgomery Street, 3rd Floor
San Francisco, CA 94105
USA
Email: mdale@wikimedia.org

***Abstract*—Wikipedia started a project in order to enable users to add video and audio on their Wiki pages. The technical downside of this is that its bandwidth requirements will increase manifold. BitTorrent-based peer-to-peer technology from P2P-Next (a European research project) is explored to handle this bandwidth surge. We discuss the impact on the BitTorrent piece picker and outline our "tribe" protocol for seamless integration of P2P into the HTML5 video and audio elements. Ongoing work on libswift which uses UDP, an enhanced transport protocol and integrated NAT/Firewall puncturing, is also described.**

I. Introduction

The peer-to-peer idea of a community of users cooperating to provide a useful service lies at the heart of Wikipedia. Users contribute their knowledge so that others can benefit. Although the knowledge comes for free, hosting it in the MediaWiki software does not, and Wikipedia has to rely on donations to keep its service operational.

Recently, Wikipedia has started a project to enable users to add video and audio on their Wiki pages. The technical downside of that is that its bandwidth requirements will increase manifold. To accommodate this increase Wikipedia decided to take the peer-to-peer idea one step further and explore peer-to-peer not only for content creation but also for content delivery.

The aim of the P2P-Next project financed by the European Union's 7th Framework is to develop an open and open-source platform for content delivery based on peer-to-peer technology for scalability [1]. The openness of the platform and software and the successful trials with users in the wild turned P2P-Next into an attractive partner. In December 2010 Wikipedia and P2P-Next represented by the Delft University of Technology started cooperating to bring peer-to-peer delivered media to Wiki pages. We would like to demonstrate the results of this cooperation at P2P10.

II. Video via P2P-Next

The platform developed by P2P-Next enables peer-to-peer based delivery of video-on-demand and live streaming [2], [3] in a single protocol, based on BitTorrent [4]. The Give-to-Get video-on-demand algorithm discourages freeriding by rewarding peers which forward data to others. Peers are encouraged to compete to forward as much data as possible, since the peers that forward the most data will be provided with a better quality of service by their neighbours. Give-to-Get also includes a novel piece-picking policy in which the set of pieces required for playback is divided into three subsets: a high, a mid, and a low-priority set. Such a division allows an easy-to-implement yet graceful transition between downloading pieces required on the short-term and those required on the long term with a distinct piece-picking policy within each priority set [5].

The live streaming algorithm of P2P-Next is also designed as an extension to the BitTorrent protocol. We added a rotating sliding window over BitTorrent's normally fixed set of pieces to provide an infinite video stream. We replaced the original mechanism for end-to-end integrity checking based on *a priori* computed hashes to a public-key based scheme. The original tit-for-tat mechanism of BitTorrent was reused to counter freeriding. The result, tested in a public trial with 4555 users is a live streaming protocol with very low prebuffering time.

As the next step in peer-to-peer communication, we are looking at simplifying the P2P/TCP/IP protocol stack. *swift* is a lightweight, content-centric, multiparty transport protocol [6]. As a multiparty protocol it is geared towards delivering the same data to a group of receivers (currently video-on-demand, in the future also live). By pruning of unneeded functions and aggressive layer collapsing we have greatly simplified the protocol. For example, it drops TCP's abstraction of sequential reliable data stream delivery. In multiparty communication,out-of-order data is not wrong and missing data can be received from other peers. Swift is implemented

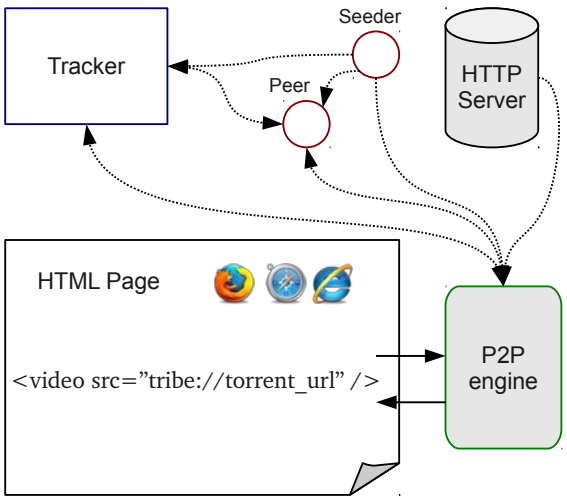

Fig. 1. Architecture overview

on top of UDP and implements user-friendly bandwidth usage via LEDBAT [7] and firewall puncturing is work in progress.

III. Video for Wikipedia

At P2P10 we would like to demonstrate P2P-Next software hosting Wikipedia video content. An important component which we developed jointly as part of this cooperation is an extension to (initially) the Firefox browser, called `tribe`. This extension adds the P2P-Next streaming solutions as a transport protocol to the browser. By offering it as a transport protocol like HTTP, Web developers do not have to worry about the delivery mechanism while authoring, and switching from HTTP to peer-to-peer delivery is just a simple URL-rewriting step in publishing pages.

The tribe protocol is specifically intended to be used with the new HTML5 media elements `<video>` and `<audio>` [8] which allow direct rendering of video and audio in a standardized way (cf. `<img>` for images.) The protocol is the first peer-to-peer video protocol to be integrated with a Web browser in this way. In the future, we intend to add the tribe protocol to any browser that allows dynamic addition of transport protocols. Our P2P10 demonstrations will include the tribe protocol using both the Give-to-Get and swift backends.

Using HTML5 will enable Wikipedia to create richer user experiences in a standardized way. Audio/video elements can be scaled and animated. Also, HTML5 allows overlays to be placed over the video elements, and provides exact information about the playback position. These two features can be used to superimpose subtitles over the video, which was not possible before in HTML. These subtitles could be created in a Wiki fashion with different people contributing subtitles in different languages. We plan to demonstrate this capability at P2P10.

Initially, the Wikipedia content will be hosted on regular Web servers. To support this situation we modified the P2P-Next software to do hybrid content delivery. First it tries to obtain the requested data from the network of peers, and if the peers do not deliver, or are expected to deliver too late for playback, the data is retrieved from the Wikipedia Web servers. Net benefit is that the data is now hosted by the downloading peer and can be delivered via peer-to-peer from then on.

The `<video>` and `<audio>` elements also allow users to seek to arbitrary positions in the (on-demand) content. We adjusted our Give-to-Get implementation to enable this seeking at the peer-to-peer level by shifting the priority sets.

The media elements also allow rendering of live sources in the Ogg video container format [9]. This would enable Wiki pages dedicated to a live event. Unfortunately, Ogg is not a transport-stream format like the MPEG-TS format that allow players to tune in at any point without any additional metadata. For live streaming clients require setup headers before they are able to decode the stream. The P2P Net software was modified to store the required metadata (Ogg headers) in the `.tstream` metadata file that clients use to join in a peer-to-peer broadcast. We plan to demonstrate live streaming at P2P10.

Acknowledgments

Our research has received funding from the European Community's Seventh Framework Programme in the P2P-Next project under grant agreement no 216217.

References

[1] P2P-Next, "Shaping the Next Generation of Internet TV," http://www.p2p-next.org/, Jul. 2009.

[2] J. Mol, J. Pouwelse, M. Meulpolder, D. Epema, and H. Sips, "Give-to-Get: Free-riding Resilient Video-on-demand in P2P Systems," in *Proceedings Multimedia Computing and Networking conference (Proceedings of SPIE Vol. 6818)*, San Jose, California, USA, Jan. 2008.

[3] J. Mol, A. Bakker, J. Pouwelse, D. Epema, and H. Sips, "The design and deployment of a bittorrent live video streaming solution," in *Proceedings IEEE International Symposium on Multimedia 2009*, San Diego, CA, USA, Dec. 2009.

[4] B. Cohen, "Incentives to Build Robustness in BitTorrent," in *Proceedings 1st Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, Jun. 2003, http://bittorrent.com/bittorrentecon.pdf.

[5] J. Mol, "Free-riding Resilient Video Streaming in Peer-to-Peer Networks," Ph.D. dissertation, Department of Software Technology, Delft University of Technology, Delft, The Netherlands, Jan. 2010.

[6] V. Grishchenko, "swift: the multiparty transport protocol," http://libswift.org/, Mar. 2010.

[7] S. Shalunov and M. Sridhavan, "IETF Working Group on Low Extra Delay Background Transport (LEDBAT)," http://tools.ietf.org/wg/ledbat/, Mar. 2010.

[8] I. Hickson and D. Hyatt, "HTML5: A vocabulary and associated APIs for HTML and XHTML. W3C Working Draft 4 March 2010," http://www.w3.org/TR/html5/, Mar. 2010.

[9] Xiph.org Foundation, "The Ogg container format," http://xiph.org/ogg/, May 2010.