Box2Box - A P2P-based File-Sharing and Synchronization Application

Andri Lareida, Thomas Bocek, Sebastian Golaszewski, Christian Lüthold, Marc Weber
University of Zurich, Department of Informatics (IFI), Communication Systems Group (CSG), Zurich Switzerland
Email: [lareida|bocek]@ifi.uzh.ch, [sebastian.golaszewski|christian.luethold|marc.weber]@uzh.ch

***Abstract*—Due to an increasing number of devices connected to the Internet, data synchronization becomes more important. Centrally managed storage services, such as Dropbox, are popular for synchronizing data between several devices. P2P-based approaches that run fully decentralized, such as BitTorrent-Sync, are starting to emerge. This paper presents Box2Box, a new P2P file synchronization application which supports novel features not present in BitTorrent-Sync. Box2Box is demonstrated in several use cases each targeted at another feature.**

I. Introduction

P2P systems are still popular and account for a large portion of Internet traffic [1], [2]. Most of this P2P traffic is related to file sharing (BitTorrent [3] (BT)), but also new types of P2P applications are emerging. Another trend is that the number of devices connected to the Internet is increasing [4], especially mobile devices [5]. Therefore, synchronization between these devices becomes more important since users tend to have more than one device on which data is accessed, modified, or created.

Centralized systems, such as Dropbox [6] or Google-Drive [7], offer synchronization solutions which enable multiple devices to synchronize their data. However, users are bound to their pricing and terms of service, and lose control over their data when uploading to one of these centralized solutions. Furthermore, recent events leading to the shut down of the Megaupload service [8] show that the single point of failure property of centralized system is a problem. BitTorrent-Sync [9] (BT-Sync) on the other hand offers a decentralized solution for synchronizing files of any size among several devices. However, BT-Sync does not currently offer any versioning features allowing reproduction of falsely deleted or modified content. Furthermore, synchronization is only possible if the devices are online. To ensure privacy, BT-Sync uses secrets which are folder-based and can be used to share folders among several devices or users. These secrets have to be exchanged out-of-band.

Box2Box is a P2P solution similar to BT-Sync, but supports novel features like friend lists and recommendation, versioning, and support for high availability peers running on user controlled nano data centers (UNaDa). A UNaDa uses home routers which are always online to offer services to the user. It differs from the nano data center approach [10] in the way that it is controlled by the user instead of the service provider.

II. Design

Designing a distributed sharing and synchronization system imposes challenges, such as consistency, conflict resolution, versioning, and management of shared files, which are harder to solve compared to a centrally managed system. The focus of this section is the design of the following mechanisms: synchronization and sharing, versioning, friend recommendation, and deploying a stable peer on a UNaDa.

To share and synchronize files in the P2P network, meta data about the file, its location, and its version is created and stored in the DHT. The key to store the meta data in the DHT is randomly selected to prevent guessing of keys. The peer that stores the meta data is responsible to notify observing peers of changes in sharing and versioning information on a per-file basis. In case a (new) file needs to be stored in the network, the peer keeping the meta data is instructed to update its meta data and add the requesting peer to its observer list. If the file upload succeeds, the peer with the meta data updates the location of this file and notifies all observing peers, including all peers of the file owner and friends that share this file. A new version will be stored in a different location, thus, the old version remains accessible.

Consistency and conflict resolution is achieved on a best-effort basis. When a user modifies a file, a modification containing the new and the previous version information of this file is announced to the responsible peer. If a second user tries to update the same file during the procedure, another modification with the same previous version is announced. In that case, a version has two or more successors, and the user, trying to update as second, receives a conflict report. Furthermore, the file is marked as such at this peer. Although a conflict is reported, Box2Box can cope with it and leaves it to the user to decide which version should be used for further modifications.

A peer in Box2Box is able to share its files with a friend. A friend can be added in two different ways: add a friend via the GUI, which will trigger a friend request. If such a friend request is acknowledged the two peers are in a friend relation and can share files. A friend can also be recommended based on the friends-of-friends approach. A peer ranks unknown friends of friends according to the number of occurrences in the friend lists. Peers exchange their friend list upon request. The user can then decide to send a friend request.

The master peer feature of Box2Box allows a user to deploy a high availability peer on a UNaDa. This super peer is responsible for storing all the user's files (large and small) since this peer has high availability and large disk space for storage. A distinction of large and small files is made as in the DHT large files may not be stored due to bandwidth restrictions. Thus large files are stored on UNaDas exclusively,

978-1-4799-0521-8/13/$31.00

while the small files can be stored in the DHT and on a UNaDa.

III. Demo Scenario

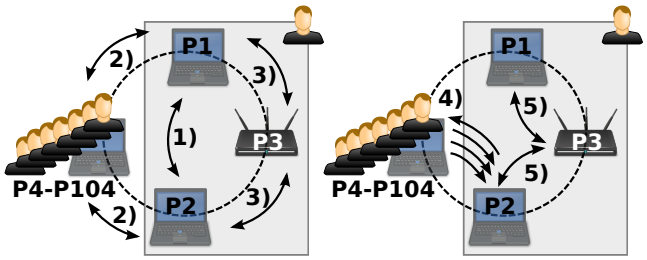
The scenario is based on one user running three peers on three devices, two peers on a laptop (P1 and P2) and one peer on a UNaDa (P3). In addition, a large network of 100 peers is running in the background on a single laptop. In this scenario the demo presents five use cases. These use cases are depicted in Figure 1 and show the setup of the demonstration and the peer interaction in each use case. The gray box that includes P1, P2, and P3 represents one user. The peers P4-P104 have one user per peer, resulting in 100 users. The GUI and console log of P1 and P2 will be shown on screens.

1) P1 is online and P2 is offline. The user, on P1, adds a large file to B2B, since P2 is offline no synchronization is possible. P1 goes offline and P2 comes online. Still, no synchronization will happen, since no direct transfer is possible. P2 will be notified about the new file though. P1 comes online again and the large file is transferred to P2, similar to BitTorrent-Sync (cf. Figure 1a). The synchronized file appears on P2 and the synchronization process can be observed in the log files.
2) Only P1 is online to the network and adds a small file to B2B. This file is encrypted and stored in the underlying P2P-network (P1 - P104). P1 goes offline and P2 comes online. Because the small file is stored in the network it can be transferred to P2 without P1 being online. Therefore, P2 downloads and decrypts the file (cf. Figure 1a).
3) The user installs Box2Box on his UNaDa (P3), acting as the super peer which stores all files (cf. Figure 1a). This means that all the files of the user are directly stored on the super peer.
4) The user on P2 requests a friend recommendation and receives a list of firends of friends. After the two users have mutually established their friend relation they can share files. The user on P2 shares a file with the selected friend. Since it is a small file it can be synchronized offline (cf. Figure 1b).
5) The two peers P1 and P2 update a file at the same time resulting in a conflict. The peer that was first to upload his file is continuing as usual. The Peer that was second receives a conflict notification (cf. Figure 1b). After the user resolves the conflict a new version of the file is created and uploaded.

IV. Future Work

The next step is to release Box2Box to the public as open source software. For the initial release several improvements are still necessary.

Future work includes backup storage on friend peers. The user will be able to set the redundancy ratio and split the backup in a way that allows to reconstruct the backup if a given fraction of the friends are online (*e.g.*, backup can be reconstructed if 3 of 5 friends are online). Furthermore, friends will have the status as trusted peers and storage of data can



(a) Use Cases 1-3 (b) Use Cases 4-5

Fig. 1. Demonstration Setup of the five Box2Box Use Cases.

prefer to use trusted peers to store data to mitigate security concerns.

Currently a JavaFX front-end is used, which will be replaced by a web-based GUI that is work in progress. For future work, a desktop integration that allows the user to use Box2Box transparently is foreseen. Furthermore, mobile versions with a reduced set of functions will be supported.

Besides encryption, future work investigates more security aspects and potential attack scenarios, such as access restriction or colluding peers. An interesting security aspect is to consider friend peers as trusted entities. Relying on these trusted peers can mitigate many attack scenarios.

Acknowledgments

This work was supported partially by the SmartenIT and the FLAMINGO projects funded by the EU FP7 Program under Contract No. FP7-2012-ICT-317846 and No. FP7-2012-ICT-318488, respectively.

References

[1] Sandvine Inc. UCL, "Global Internet Phenomena Report1H," http://www.sandvine.com/downloads/documents/Phenomena_1H_2013/Sandvine_Global_Internet_Phenomena_Report_1H_2013.pdf, last visited: 30.5.2013, May 2013.

[2] CISCO. (2013, Jan.) Cisco Visual Networking Index: Forecast and Methodology, 2011-2016. [Online]. Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html

[3] B. Cohen, "Incentives Build Robustness in BitTorrent," in *1st Workshop on Economics of Peer-to-Peer Systems (P2PECON)*, Berkeley, CA, USA, June 2003.

[4] Miniwatts Marketing Group, "Internet Growth Statistics," http://www.internetworldstats.com/emarketing.htm, last visited: 5.8.2009, January 2008.

[5] M. Meeker and L. Wu. (2013, May) Internet trends. [Online]. Available: http://allthingsd.com/tag/mary-meeker/

[6] Dropbox Inc. (2013, May) Dropbox. [Online]. Available: http://www.dropbox.com/

[7] Google Inc. (2013, May) Google drive. [Online]. Available: http://drive.google.com

[8] BBC. (2013, May) Megaupload file-sharing site shut down. [Online]. Available: http://www.bbc.co.uk/news/technology-16642369

[9] BitTorrent Labs. (2013, May) BitTorrent Sync. [Online]. Available: http://labs.bittorrent.com/experiments/sync.html

[10] V. Valancius, N. Laoutaris, L. Massoulié, C. Diot, and P. Rodriguez, "Greening the internet with nano data centers," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*. ACM, 2009, pp. 37–48.