Peer-to-peer Files Sharing Networks

Submitted by: Anas AbuDaqa

Definition

Peer-to-peer (P2P) files sharing networks have been one of the most reliable platforms to realize the potential of network coding, since end hosts at the edge of the Internet have abundant computational resources with modern processors. In this survey paper, we extensively summarize, assess, compare, and classify the most recently used techniques to improve P2P content distribution systems performance using network coding. To the best of our knowledge, this survey is the first comprehensive survey that specifically focuses on the performance of network coding based P2P file sharing systems.

1. Introduction

The Peer-to-Peer (P2P) architecture has triggered a technical revolution of large content distribution service on the Internet. The P2P model is the antithesis of the classic client–server architecture, in which each node can behave either as a server or a client. In the client–server architecture, data stored on the server is sent to clients on a request basis. However, for a large number of clients, the workload on the server may be too heavy, leading to extremely low download rates. Moreover, for this architecture, the server represents a single point of failure. On the contrary, in the P2P architecture, each node is called a servent [1][2] which means that a node may behave as a server as well as a client at the same time. This leads to higher bandwidth utilization, more availability of content, and reduced download times [3]. The BitTorrent protocol [4] is considered to be one of the most successful P2P file sharing protocols. BitTorrent-like systems work in two phases, the first phase utilizes a discovery protocol, in which peers discover and establish connections to each other. Detailed discussions on the discovery protocol can be found in [5][6][7][8][9][10]. The second phase utilizes file dissemination protocol, in which peers start to download and upload the chunks of a file until they get the complete file.

2. Network Coding Based P2P File Sharing

In the file dissemination protocol, the algorithm used for the selection of the pieces to download is highly important to achieve an acceptable level of system performance. In fact, wrong selection choices may lead to a situation where pieces of a file owned by a peer are not required by any other peer. Conversely, a piece which is needed by many peers is either very rare or not available within the network. Consequently, the peer requires a specific policy for downloading the pieces of the file. The original specification of the BitTorrent protocol [4]] requires that clients are able to download pieces in a purely random way. Subsequently, new selection techniques were introduced to improve performance [11][12]. In general, an involved peer in any P2P file sharing network must answer the following questions when requesting a file: (1) which pieces should be downloaded, and (2) from which peers? This is known as piece scheduling [13], piece selection [14]], or the coupon collector's problem [15]]. The other important challenge, especially in large-scale content distribution networks, is that the departure, or the failure of a peer is naturally dynamic as the peer may depart or fail without notice, which is known as dynamics peer participation or churn [16]]. Subsequently, pieces that were owned by the departing peers may become rare or unavailable.

With the assistance of network coding [17]], the aforementioned issues can be partially resolved. Rather than distributing the original data pieces, a peer shares encoded pieces, where each encoded piece is a linear combination of the original pieces. In this manner, there is no need for the requester peer to ask for special data pieces since all encoded pieces are almost equally likely to be beneficial to the peer. A peer in the network can blindly download encoded pieces from other peers. When a sufficient number of encoded pieces are owned, the original file can be reconstructed. Using this scheme, the rarest piece problem is avoided, and the download time of the file is potentially reduced [18]. Another feature of network coding is the robustness against sudden peer departure. In the original BitTorrent, if a peer who is the single owner of a piece, leaves the swarm, the complete original file cannot be collected by the remaining peers. However, with

network coding, the concern regarding the absence of certain data pieces is no longer an issue. Since the pieces are linearly combined together, each of them is available in a large number of encoded pieces in the network.

The main contributions of this paper are as follows. First, we provide a detailed tutorial of network coding implementation for P2P file sharing. Second, we extensively survey the most important and recent network coding based P2P file sharing protocols and provide a classification for those protocols. Third, this paper, to the best of our knowledge, is the first paper that is specific to network coding based P2P content distribution systems.

References

1. I. Stoica; R. Morris; D. Liben-Nowell; D.R. Karger; M.F. Kaashoek; F. Dabek; H. Balakrishnan; Chord: a scalable peer-to-peer lookup protocol for internet applications. *Power Prefixes Prioritization for Smarter BGP Reconvergence* **2003**, *11*, 17-32, 10.1109/tnet.2002.808407.

2. Francis Nwebonyi; Rolando Martins; Manuel Eduardo Correia; Reputation based approach for improved fairness and robustness in P2P protocols. *Peer-to-Peer Networking and Applications* **2018**, *12*, 951-968, 10.1007/s12083-018-0701-x.

3. Valentino Pacifici; Frank Lehrieder; György Dán; Cache Bandwidth Allocation for P2P File-Sharing Systems to Minimize Inter-ISP Traffic. *Power Prefixes Prioritization for Smarter BGP Reconvergence* **2014**, *24*, 437-448, 10.1109/tnet.2014.2367021.

4. Cohen, B. The BitTorrent Protocol Specification. 2008. Available online: http://bittorrent.org/beps/bep_0003.html (accessed on 24 March 2020)

5. Lareida, A.; Bocek, T.; Waldburger, M.; Stiller, B. RB-tracker: A fully distributed, replicating, network-, and topology-aware P2P CDN. In Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), Ghent, Belgium, 27–31 May 2013; pp. 1199–1202.

6. D. Wu; P. Dhungel; Xiaojun Hei; C. Zhang; K. W. Ross; Understanding Peer Exchange in BitTorrent Systems. *2010 IEEE Tenth International Conference on Peer-to-Peer Computing (P2P)* **2010**, *null*, 1-8, 10.1109/p2p.2010.5569967.

7. Fabio V. Hecht; Thomas Bocek; Burkhard Stiller; B-Tracker: Improving load balancing and efficiency in distributed P2P trackers. *2011 IEEE International Conference on Peer-to-Peer Computing* **2011**, *null*, 310-313, 10.1109/p2p.2011.6038749.

8. Daniel S. Menasché; A.A. De A. Rocha; Bin Li; Don Towsley; Arun Venkataramani; Content Availability and Bundling in Swarming Systems. *Power Prefixes Prioritization for Smarter BGP Reconvergence* **2012**, *21*, 580-593, 10.1109/tnet.2012.2212205.

9. Fry, C.P.; Reiter, M.K. Really Truly Trackerless BitTorrent; School of Computer Science, Carnegie Mellon University, Tech. Rep: Pittsburgh, PA, USA, 2006; pp. 6–148. [Google Scholar]

10. Nima Jafari Navimipour; Farnaz Sharifi Milani; Erratum to: A comprehensive study of the resource discovery techniques in Peer-to-Peer networks. *Peer-to-Peer Networking and Applications* **2015**, *9*, 254-254, 10.1007/s12083-015-0335-1.

11. Jiaqing Luo; Bin Xiao; Kai Bu; Shijie Zhou; Understanding and Improving Piece-Related Algorithms in the BitTorrent Protocol. *IEEE Transactions on Parallel and Distributed Systems* **2013**, *24*, 2526-2537, 10.1109/TPDS.2013.8.

12. Yusuo Hu; Dafan Dong; Jiang Li; Feng Wu; Efficient and incentive-compatible resource allocation mechanism for P2P-assisted content delivery systems. *Future Generation Computer Systems* **2013**, *29*, 1611-1620, 10.1016/j.future.2012.08.008.

13. Chiu, D.M.; Yeung, R.W.; Huang, J.; Fan, B. Can network coding help in P2P networks? In Proceedings of the 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, Boston, MA, USA, 26 February–2 March 2006; pp. 1–5

14. Jeng-Long Chiang; Yin-Yeh Tseng; Wen-Tsuen Chen; Interest-Intended Piece Selection in BitTorrent-like peer-to-peer file sharing systems. *Journal of Parallel and Distributed Computing* **2011**, *71*, 879-888, 10.1016/j.jpdc.2010.12.011.

15. Fragouli, C.; Le Boudec, J.Y.; Widmer, J.; Network coding: An instant primer. *ACM SIGCOMM Comput. Commun. Rev.* **2006**, *36*, 63–68.

16. Håkan Terelius; Karl Henrik Johansson; Peer-to-Peer Gradient Topologies in Networks With Churn. *IEEE Transactions on Control of Network Systems* **2018**, *5*, 2085-2095, 10.1109/tcns.2018.2795704.

17. Feng, C.; Li, B. Network coding for content distribution and multimedia streaming in peer-to-peer networks. In Network Coding; Elsevier: Amsterdam, The Netherlands, 2012; pp. 61–86

18. Qing-Chao Cai; Kwok-Tung Lo; Two Blocks Are Enough: On the Feasibility of Using Network Coding to Ameliorate the Content Availability of BitTorrent Swarms. *IEEE Transactions on Parallel and Distributed Systems* **2012**, *24*, 1682-1694, 10.1109/TPDS.2012.211.

Keywords