Use of Peer-To-Peer Technology in Internet Access Networks and its Impacts (IPDPS 2010 PhD Forum)

Peter Danielis, Dirk Timmermann
University of Rostock
Institute of Applied Microelectronics and Computer Engineering
18051 Rostock, Germany
Tel.: +49 (381) 498 -7272
Email: peter.danielis@uni-rostock.de

Personal Website: http://www.imd.uni-rostock.de/ma/pd032
Internetworking Project: http://www.imd.uni-rostock.de/networking

***Abstract*—Objectives of the dissertation are impacts of Peer-to-Peer (P2P) traffic on Internet core networks as well as novel approaches for using P2P technology in Internet access networks. Thereby, challenges of P2P computing concerning topology awareness, scalability, and fault-tolerance are analyzed. The thesis' first part focuses on improving insufficient scalability and fault-tolerance properties of present-day Internet services like the Domain Name System (DNS) by using available resources in access networks. The second part addresses P2P mechanisms for a highly scalable, resilient, and distributed storing and computing solution in the access network. Finally, a new algorithm for allowing topology awareness in P2P networks is proposed. For evaluation, a prototype for a P2P-based DNS has been developed and another prototype for a P2P-based store and compute platform is currently set up.**

Currently, the author is in his fourth PhD year and will finish in 2010. Further information, full references, and papers can be obtained from the websites given below the affiliation.

***Keywords*-Peer-to-Peer; Access Networks; Internet; Topology Awareness; Scalability; Fault-Tolerance**

I. Introduction

Due to the Internet's increasing complexity and the steadily growing number of users, many Internet core services like the Domain Name System (DNS) and Internet Service Provider (ISP) storage solutions show severe problems concerning scalability and fault-tolerance. On the one hand, they depend on old (centralized) service structures, which have not been designed for the dimension and complexity of today's and future Internet. On the other hand, cost-intensive infrastructure has to be provided to compensate for insufficient scalability and fault-tolerance.

Recently, Peer-to-Peer (P2P)-based communication infrastructures have emerged, which provide an excellent technical basis for the realization of efficient decentralized solutions to replace or at least complement existing structures. Thereby, the authors understand P2P as a new networking paradigm rather than as synonym for today's incriminated applications such as file sharing. P2P networks show intrinsic features like high scalability and fault and attack resilience, which can therefore be used at no extra costs.

As one important part of the Internet, the thesis focuses on Ethernet-based, packet-switched access networks. Access networks comprise, among other things, access nodes. These access nodes have a certain available storage and computing capacity, which may be used at no extra costs, assuming some idle time or spare capacity being left in an average access node. Therefore, they shall serve as member of the P2P network and are connected by a self-organizing distributed hash table-based P2P network. Each access node contributes a part of its storage and computing capacity to the distributed hash table network.

As one exemplary result of the research work, a P2P-based DNS is briefly outlined in Section I-A. Section I-B shortly reports selected aspects of a P2P-based store and compute platform.

However, a major drawback of P2P traffic is its dominance of Internet traffic today and the associated needs for traffic engineering. Section I-C introduces a new algorithm to prevent ISPs' core networks from being overburdened with P2P data.

A. P-DONAS: A P2P-based DNS in Access Networks

DNS, which represents one of the Internet core services, has not been designed for the dimension and complexity of the Internet today. Therefore, scalability is one major challenge. Nowadays, ISPs need to provide and operate regional DNS server farms for resilience and load sharing. Addressing scalability properly means investment in additional equipment, i.e., DNS server farms, efforts to operate and manage them, and energy to have server farms run 24/7.

To facilitate, among other things, cost reductions for ISPs, P-DONAS–a P2P-based Domain Name System–has been developed. Access nodes of an ISP's access network are organized into a distributed hash table-based P2P network. As high scalability and resilience are intrinsic features of distributed hash tables, P-DONAS provides these features as well. Each access node acts as traditional DNS server and stores a piece of DNS data it is responsible for and shares it with all other access nodes. Thereby, available memory

and computing resources of the access nodes are utilized. DNS requests issued to an access node are resolved via P2P lookups while maintaining full compatibility with and access to the traditional DNS. Consequently, P-DONAS is an innovative way of providing the DNS service and saves DNS server farms.

As proof of concept, a software prototype has been developed in C++, which emulates an access node with P-DONAS functionality. Currently, this prototype is ported to a Xilinx evaluation board. It will be extended by a VHDL hardware component for high-speed DNS answers from a cache memory. As popular DNS entries are stored in and answered from this cache, traffic caused by lookups on the distributed hash table ring is kept at a minimum.

B. A P2P-based Store and Compute Platform in Access Networks

ISPs have to store session data of their customers for operation, management, and control tasks. Thereby, each access node of an access network has to keep track of session data of all customers, which are connected to it via physical ports. Session data comprises, among other things, IP addresses, physical ports, MAC addresses, and lease times of IP addresses. This data has to be stored persistently as it has to be reloaded after an access node's restart or crash, i.e., it is needed for session data recovery. Today, an access node's flash memory is used for the storage of session data. However, this memory is limited in its availability and rewritability and is intented for other purposes.

Therefore, access nodes shall be organized into a distributed hash table-based P2P network to share their available RAM resources. Thereby, the distributed hash table network serves as semi-permanent distributed memory for a structured redundant and interleaved storage of session data. After a restart or crash, an access node performs session data recovery by selectively reading required data from the distributed hash table network. As data is stored interleaved by means of erasure resilient codes, all data can be restored even in case of a high number of unavailable access nodes. Thereby, much less storage overhead is created than using pure data replication.

Currently, a P2P-based store and compute prototype is set up in C++, which runs both under Windows and on a Xilinx evalution board.

Prospectively, an access node will be used as a distributed computing resource as well. From the stored data, some useful statistics will be calculated and again stored in a distributed manner among access nodes of the distributed hash table network. Thereby, this computation task will be given to all access nodes, which contribute a part of their available computing power during idle periods, thereby utilizing existing and unused capacity more effectively.

C. Considering Physical Proximity in P2P Networks

P2P file sharing generates by far the most Internet traffic reaching up to 70 % in some regions of the world. These data volumes pose a significant challenge to ISPs regarding traffic engineering. Because P2P routing is usually agnostic of the underlying topology, traffic engineering abilities of ISPs are inhibited and their core networks are overburdened with P2P data. This problem is known as mismatching problem between the logical P2P overlay topology and the underlying physical network topology [1].

To disburden ISPs' core networks, a new algorithm is proposed in order to improve peer selection. P2P users are provided with accurate information on the hop counts to other peers to select close peers in unstructured P2P networks. Thereby, the initial Time-To-Live value (TTL) of outgoing IP packets is copied and inserted as part of the P2P protocol's payload. At the packet's destination, the hop count is calculated as the difference between the copied TTL and the TTL of the IP header. Using the hop count, a relationship between the logical overlay and the physical network is established to perform traffic engineering. By selecting proximate peers, physical proximity of P2P traffic can be increased. Therefore, less bandwidth is consumed, avoiding traffic congestions when the load of the network is already heavy.

Currently, simulations are carried out for the BitTorrent algorithm to show that ISPs benefit from a modified BitTorrent using the hop count as additional selection criterion for download partners. Moreover, the simulations are expected to show that using the hop count in the modified BitTorrent algorithm does not deteriorate the users' Quality of Experience compared to standard BitTorrent.

II. Related Work

P2P-based DNS: There are already approaches to design a P2P-based DNS [2], [3]. The solution in [2] depends on the cooperation of end nodes, which impedes the adoption of such a system. Moreover, it leaves a lot of questions concerning the implementation unsolved. As opposed to this approach, P-DONAS is a trustworthy and reliable ISP-provided DNS service and provides more insight in the technical realization. The approach in [3] is dependent on globally distributed servers, which have to be contributed to form a globally shared DNS cache. Contrary to this approach, P-DONAS focuses on providing DNS functionality by using available resources of access nodes in the access network rather than additional servers.

P2P-based store and compute platforms:

[4], [5], [6] propose to use distributed hash table-based solutions for distributed storage of data. In [4], globally distributed untrusted servers shall be used. In [5], the authors suggest a public data management system for Web applications. The solution in [6] focuses on a scalable storage platform for storing IP flow records. In [7], data sets are

organized into redundancy groups. Thereby, the focus is on high availability of data in P2P-based applications with a high churn of nodes. In contrast to the author's approach, none of the cited works uses available resources of trusted reliable infrastructure to provide a storage platform for general data.

Considering physical proximity in P2P networks: Many approaches to construct unstructured topology-aware overlay networks do exist, e.g., [8] and [9]. In contrast to these approaches, the author's proposed algorithm does not intervene with the construction of unstructured P2P networks and no additional packets have to be sent to determine the distance, i.e., the hop count between peers.

Also, there are approaches to shape P2P traffic in a more efficient way with the support of the ISP [10], [11], [12]. Contrary, the author's approach does not require network support but does the necessary modifications solely in the application.

III. Conclusion

This brief outline described the main aspects and recent results of the author's research work. Major terms are topology awareness, scalability, and fault-tolerance in the area of P2P networks. Thereby, the focus is on impacts of P2P technology and traffic on access and core networks. Selected aspects of the IPDPS's scope are thus touched in this proposal, e.g., distributed algorithms focusing on scalability, applications of P2P computing, and parallel architectures for storage systems.

Acknowledgment

The authors would like to thank the Broadband Access Division of Nokia Siemens Networks GmbH & Co. KG in Greifswald, Germany, for their inspiration and continued support in this project. This work is partly granted by Nokia Siemens Networks.

References

[1] P. Danielis, H. Widiger, S. Kubisch, J. Schulz, D. Timmermann, D. Duchow, and T. Bahls, "A Conceptual Framework for Increasing Physical Proximity in Unstructured Peer-To-Peer Networks." Princeton, NJ, USA: IEEE Sarnoff Symposium 2008, 2008.

[2] R. Cox, A. Muthitacharoen, and R. T. Morris, "Serving DNS using a Peer-to-Peer Lookup Service." IPTPS, 2002, pp. 155–165.

[3] V. Ramasubramanian and E. G. Sirer, "The Design and Implementation of a Next Generation Name Service for the Internet." ACM SIGCOMM, 2004, pp. 331–342.

[4] J. Kubiatowicz, D. Bindel, Y. Chen, S. Czerwinski, P. Eaton, D. Geels, R. Gummadi, S. Rhea, H.Weatherspoon, W.Weimer, C.Wells, and B. Zhao, "OceanStore: An Architecture for Global-Scale Persistent Storage." ASPLOS, 2000, pp. 190–201.

[5] M. Karnstedt, K.-U. Sattler, M. Richtarsky, J. Mueller, M. Hauswirth, R. Schmidt, and R. John, "UniStore: Querying a DHT-based Universal Storage." International Conference on Data Engineering, 2007, pp. 1503–1504.

[6] C. Morariu, T. Kramis, and B. Stiller, "DIPStorage: Distributed Storage of IP Flow Records." 16th IEEE Workshop on Local and Metropolitan Area Networks, 2008, pp. 108–113.

[7] Q. Xin, T. Schwarz, and E. L. Miller, "Availability in Global Peer-To-Peer Storage Systems." Distributed Data and Structures, Proceedings in Informatics, 2004.

[8] S. Merugu and E. Zegura, "Adding Structure to Unstructured Peer-to-Peer Networks: The Use of Small-World Graphs." JPDC, 2005, pp. 142–153.

[9] Y. Liu, X. Liu, L. Xiao, L.M.Ni, and X. Zhang, "Location-Aware Topology Matching in P2P Systems." INFOCOM, 2004, pp. 2220–2230.

[10] IETF, "Application-Layer Traffic Optimization (alto)," 2009. [Online]. Available: http://www.ietf.org/html.charters/alto-charter.html

[11] H. Xie, Y. R. Yang, A. Krishnamurthy, and Y. L. A. Silberschatz, "P4P: Provider Portal for Applications." ACM SIGCOMM, 2008, pp. 351–362.

[12] E. Leonardi, M. Mellia, A. Horvath, L. Muscariello, S. Niccolini, and D. Rossi, "Building a cooperative P2P-TV application over a wise network: The approach of the European FP-7 strep NAPA-WINE." IEEE Communications Magazine 46 (4), 2008, pp. 20+22.