Simple Economic Management Approaches of Overlay Traffic in Heterogeneous Internet Topologies

European Seventh Framework Project FP7-2008-ICT-216259-STREP

D1.1 Annex
SmoothIT Related Work

The SmoothIT Consortium

University of Zürich, UZH, Switzerland
DoCoMo Communications Laboratories Europe Gmbh, DoCoMo, Germany
Technische Universität Darmstadt, TUD, Germany
Athens University of Economics and Business - Research Center, AUEB-RC, Greece
PrimeTel Limited, PrimeTel, Cyprus
Akademia Gorniczo-Hutnicza im. Stanislawa Staszica W Krakowie, AGH, Poland
Intracom S.A. Telecom Solutions, ICOM, Greece
Julius-Maximilians Universität Würzburg, UniWue, Germany
Telefónica Investigación y Desarollo, TID, Spain

For more information on this document or the SmoothIT project, please contact:

Prof. Dr. Burkhard Stiller
Universität Zürich, CSG@IFI
Binzmühlestrasse 14
CH—8050 Zürich
Switzerland

Phone: +41 44 635 4355
Fax: +41 44 635 6809
E-mail: info-smoothit@smoothit.org

Document Control

Title: D1.1 Annex - SmoothIT Related Work

Type: Internal

Editor(s): Haris Neophytou

E-mail: harisn@prime-tel.com

Author(s): Wolfgang Kellerer, Burkhard Stiller, Maximilian Michel, Sergios Soursos, Ioanna Papafili, Neophytou Haris, Frank Lehrieder, Simon Oechsner, Dirk Staehle, Tobias Hoßfeld, Konstantin Pussep, Fabio Hecht, Juan Fdez-Palacios, María Ángeles Callejo Rodríguez, George D. Stamoulis, Jan Derkacz, Rafal Stankiewicz, Krzysztof Wajda, Neofytos Gerosavva

Doc ID: D1.1 Annex-v1.3.doc

AMENDMENT HISTORY

Version	Date	Author	Description/Comments
0.1	April 17, 2008	Wolfgang Kellerer	First version, provide TOC
0.2	June 27, 2008	Maximilian Michel	Provide related work descriptions (DoCoMo)
0.3	July 3, 2008	Sergios Soursos	Added papers & descriptions (AUEB)
0.4	Sep 1, 2008	Frank Lehrieder, Simon Oechsner, Dirk Staehle, Tobias Hoßfeld, Jan Derkacz, Rafal Stankiewicz, Krzysztof Wajda	UniWue and AGH Contribution.
0.5	Sep 5, 2008	Konstantin Pussep	TUD Contribution
0.6	Sep 14. 2008	Haris Neophytou	Merging contributions together.
1.0	Oct 27, 2008	Sergios Soursos, Fabio Hecht, George Stamoulis, Juan Fdez-Palacios, María Ángeles Callejo Rodríguez	Providing additional contribution for Section 2.1
1.1	Oct 28, 2008	Haris Neophytou	Updating and editing document.
1.2	Nov 17, 2008	Burkhard Stiller	Update of structure and content
1.3	Jan 05, 2009	Haris Neophytou	Added paper summary, expanded executive summary

Table of Contents

1 Executive Summary

This document reports on part of the work performed within Work Package WP1, *Application and Traffic Requirements* and serves as Annex to deliverable D1.1. “*Requirements and Application Classes and Traffic Characteristics (Initial Version)*” aiming to provide a comprehensive collection and discussion of the relevant state of the art for the SmoothIT project.

The SmoothIT consortium identified and analysed relevant research results on European and global level, which were collected by a broad literature survey (scientific press and internet sources) in order to compile and present the current state of knowledge.

The deliverable contents include an overview of major and up-to-date work carried out on a global scale in relation to SmoothIT goals and objectives. Topics vary from novel frameworks for effective traffic control between network providers and P2P applications like the P4P, to well-thought algorithms and techniques that intend to improve locality awareness.

The content of this document can be used for various forthcoming deliverable and publications of SmoothIT.

2 SmoothIT Related Work

The work related to the SmoothIT project covers a smaller number of directly related projects and approaches as well as a larger number of aspects in various other works, which has been investigated by the SmoothIT consortium.

To ensure that a clear technical, economic, and organizational understanding of each of these related approaches can be achieved, the following sections of this Annex to the major starting deliverable of SmoothIT D1.1. “Requirements and Application Classes and Traffic Characteristics (Initial Version)” – termed “D1.1 Annex: SmoothIT Related Work” – do address typically firstly the key contributions of the approach investigated. Secondly, project goals or targets are listed in short and, if available, more detailed aspects of the architecture, protocols in use, or models applied/developed are added. Finally, this objective presentation of material obtained through published papers, magazines, or articles as well as personal discussion with other projects may be followed by a SmoothIT-related discussion, which may be already enhanced by a comparison to the SmoothIT architecture as it stands at the end of project year 1.

3 Proactive Network Provider Participation for P2P

P4P stands for “Proactive network Provider Participation for P2P” and has been developed at Yale University and is being promoted by the P4P Working Group [1] of the Distributed Computing Industry Association (DCIA). The related paper can be found in [2].

Core members of the WG are:

- ISPs such as *AT&T* and *Verizon.*
- P2P software developers such as *BitTorrent*, *Joost* and *Vuze*.
- Telecom suppliers such as *Cisco*.
- Academics: *Washington University*.

For a full list, visit http://www.pandonetworks.com/p4p

3.1 Project Goals

The main goal is the cooperation between ISPs and P2P software distributors in order to accelerate content distribution and optimize the utilization of ISP network resources. Investigate and establish *voluntary* best practices for collaboration. The P2P traffic should be optimized and become more manageable to ISPs before it becomes a real problem. One example is the decreased number of hop counts for P2P content delivery, which should be good both for ISPs and P2P users.

P4P, as stated in the mission statement of the P4P Working Group, is “*a set of business practices and integrated network topology awareness models designed to optimize ISP network resources and enable P2P based content payload acceleration.*”

Therefore, P4P proposes a network aware P2P that envisages maximizing delivery, reduce costs and improve performance. P4P sets to allow the P2P networks to optimize traffic within each ISP, which aims not only to reduce the volume of data traversing the ISP’s infrastructure, but to also create a more manageable flow of data.

3.2 P4P Architecture Overview

The P4P framework consists of a control-plane component and an optional data-plane component. In the control plane, P4P introduces iTrackers to provide portals for P2P to communicate with network providers. iTrackers allow P4P to divide traffic control responsibilities between P2P and providers, and to also make P4P incrementally deployable and extensible. iTrackers are light-weight, maintaining necessary states only and therefore providing scalability.

iTrackers communicate with appTracker in P2P systems to obtain important information and suggestions related to the peering decision making process.

iTracker implements cache discovery providing users with a list of “good peers” (e.g. users in same PoP range, users with higher uplink capacities, etc) and cache servers within the ISP cloud. P4P allows routers on the data plane to give highly granular feedback to P2P and enable more efficient usage of network resources.

The Information to be exchanged between P2P (overlay) applications and ISPs is summarized below.

ISPs offer iTrackers which provide (on request) the overlay following information:

- Network topology & status.
- Provider's policy (traffic ratio balance policy etc).
- Capabilities (DiffServ etc).

The overlay collects this info from different iTrackers (1 tracker per ISP) and modifies its traffic accordingly. Additionally the overlay can provide its swarm IDs (e.g. an ID per BitTorrent Tracker) so the ISP can optimize the P2P traffic and *suggest* the new topology to the overlay.

Each overlay provides a designated appTracker to communicate with iTrackers.

The benefits to expect from this approach are:

- Faster downloads for P2P (e.g. BitTorrent).
- Less interdomain traffic for ISPs (due to optimization).

Open Issues:

- ISPs offer too much information (may be misused).
- How up-to-date is the optimization info received by an overlay from an iTracker?

3.3 Discussion

One of the main issues that arise is that with P4P the ISP tracks and aids illegal file sharing. Furthermore, clients following P4P could have less success on non-P4P trackers or non-P4P aware ISPs. Also, P4P might slow down transfers of non-adopters. Such an observation raises Network Neutrality issues and provides incentives for the ISPs to ban P4P-aware clients.

Another important point is that the required cooperation (sharing of information) between ISP and P2P provider is taken for granted and no incentives that suggests both players to do are introduced. Combined with the previous observation, this provides a disincentive for the ISPs. The simulation done by the WG assume same goals for the ISP and overlay provider; a case that is not always true. Moreover, the fact that the mechanism proposed is swarm-aware raises many scalability and privacy issues. How the mechanism deals with the large number of active swarms at any given moment, for a single P2P application?

Based on the mathematical formulation provided, P4P mainly deals with file sharing applications, since it tries to achieve bandwidth maximization. But, for other types of applications, like P2P streaming applications, bandwidth is not the only important property. Delay and jitter are most crucial and thus, the model should be extended to handle these cases as well.

Finally, the inter-domain case is not yet considered and the effects of such a mechanism in the interconnection relations among the ISPs are not studied. From the economic point of view, it remains open to further research whether such a mechanism increases intra-ISP congestion and cost, at the sake of locality-awareness. Additionally, there is a possibility that non-P2P applications incur deteriorated performance. Such a phenomenon should also be further studied.

3.4 Comparison to the SmoothIT Architecture

P4P and SmoothIT share similar key objectives: better P2P performance and more efficient network resource usage. Thus, similar functionality is expected to be provided by the architectural components. Based on the P4P Overview paper, the description of the P4P framework is not detailed enough to allow for a precise comparison with SmoothIT preliminary design. In fact, at this stage SmoothIT also has not provided a full architecture description to allow for a thorough comparison. However, some differences at this high level of design can be identified:

- P4P does not consider inter-domain interactions among iTrackers of different ISPs, while SmoothIT has as a main objective to promote inter-domain communications between the architectural components. P4P relies on the appTracker, normally a BitTorrent tracker, to collect information from the iTrackers, either through the peers or directly. The appTracker plays the role of a decision point about which other peers a peer should connect to. In SmoothIT, there is no notion of appTracker. We try to be as transparent in the overlay layer as possible, introducing entities that are treated the same with the rest of p2p entities. Of course, some entities will have augmented functionalities (e.g., exchange of information with the underlay), but this will not affect the existing p2p protocol.
- There is no description of the protocols used to access the interfaces of iTrackers and how well they perform. So far, there is no evaluation of the communication overhead. SmoothIT is working on minimizing the overhead involved in the explicit communication.
- SmoothIT has not yet considered some information or services that are considered by P4P:
    - iTracker policy interface allows a peer to obtain network policies, e.g., time-of-day link usage policy, as the incentive for the application is not clear – unless there is a corresponding charging policy.
    - Request to reserve a node in an ISP to act as a peer, e.g., to accelerate content distribution

    - Security (access control) is mentioned but not described. SmoothIT must provide a good security framework, going beyond the notion of “if you think it is not secure, don't use it”.
    - Privacy for both ISP and end users: SmoothIT will consider privacy of different players.
- SmoothIT is also investigating alternative charging schemes to provide further incentives for end users and ISPs to cooperate.
- P4P assumes that the answer given by the iTracker is always beneficial to the P2P application. However, there are cases in which this may not be true, and SmoothIT will account for those situations.

3.5 Summary and Conclusions

Even though both SmoothIT’s and P4P’s main goal is to improve the performance of overlay applications by means of the active collaboration between ISP and overlay applications, the technical strategy to achieve this objective is different in both cases. In particular, as detailed in the previous sections, the following main differences (in terms of incentives and architecture designs) can be considered:

- P4P is mainly focused on P2P File sharing applications (P2P aims to optimize the download time) while SmoothIT also aims to provide incentives to other overlay applications, such as P2P Streaming based applications. Therefore, SmoothIT architecture must provide an interface able to interact with multiple overlay applications.
- P4P technical solution mainly focuses on providing unconditional locality awareness mechanisms while SmoothIT is looking into adaptive solutions and is also aiming at providing new incentives such as QoS enforcement mechanisms.
- SmoothIT is also focusing on inter-domain solutions that could allow the interaction between different SmoothIT based architectures that are deployed in different ISPs domains.

Taking into account this set of main differences (among others that are also described in the previous sections), even though SmoothIT and P4P have similar target goal, they are providing different approaches to achieve this general objective.

4 Relation of P2P Technology and IPTV Services

In the work of [3], it is compared, how the effect of different network models on the possible advantages of P2P technology for content distribution (IPTV) looks like. They conclude that most of the existing work in this field uses a too abstract view on the network infrastructure (cloud model) and thus overestimates the advantages of P2P technology. Further, a number of different possible P2P incentive strategies is presented and evaluated.

This work emphasizes the need for a deeper look on the underlying networks of the internet by using a physical model which considers also bandwidth and link constraints, to see if different P2P technologies are beneficial in the real networks.

5 Locality Aware Peer Assisted Delivery: The Way to Scale Internet Video to the World

This work in [4] motivates the adoption of locality aware peer assisted content distribution with the lack of the current Internet backbone to serve much further growth in video-streaming with central servers. If BitTorrent-like P2P content distribution systems are aware of the topology and each peer is aware of its position within the hierarchical Internet model (levels: AS, POP neighborhood, branch office), the majority of the peers to download from can be chosen from the neighborhood and thus traffic stays local and does not need to traverse more expensive and limited Internet backbone links. Location can be measured by the peers itself using the IP address and the subnet mask of the WAN link.

An important contribution of this paper is the further refinement of the level of granularity in terms of locality. While comparable work just mostly differentiates between inside and outside ISP peers, here several levels of locality are introduced.

6 LiveSwarms: Adapting BitTorrent for End Host Multicast

In [5] a system called LiveSwarms is presented. Its main target is to provide application layer real-time multicast delivery. The system is based on BitTorrent but incorporates some protocol and policy modifications, such as switching from PULL-based to PUSH-based messaging for chunk exchange, to adapt to streaming by higher upload utilization.

The authors show the feasibility of such a mesh-based streaming system with trace-based experiments for the order of hundreds of users for heterogeneous links and in case of churn and flash crowds. Especially interesting is the observation, that BitTorrent’s tit-for-tat preserves some self-healing capabilities for the system in that sense, that nodes, whose insufficient upload decreases the system’s resources, will be separated and do not interrupt the service for all participating nodes.

7 Improving Traffic Locality in Bittorrent via Biased Neighbor Selection

The approach in [6] compares the Biased Neighbor Selection algorithm as a way to promote locality awareness in BitTorrent to existing approaches such as throttling, gateway peers and caching. This new approach is based on selecting a high number of neighbors from within the same ISP, which does not hurt the performance of the BitTorrent system in terms of download time. At the same time the rate of traffic heading for peers which are not in the same ISP network is reduce.

Furthermore it is shown, that those existing approaches can also benefit from biased neighbor selection. This work demonstrates the feasibility of a win/win solution, where neither the user nor the network provider will suffer from an improvement for the other part. It also presents a direct approach towards this goal, which could be implemented unobtrusively into existing systems without a demand for additional hardware at the ISP or even without changed BitTorrent software for the P2P users.

8 On Cooperative Content Distribution and the Price of Barter

The paper [7] deals with the problem of content distribution. First, the Binomial Pipeline algorithm is introduced, which is optimal in terms of required download time for cooperative content distribution systems (all participants are inclined to give their full upload speed without any reciprocation to the system). Then strict, credit-limited and triangular barter strategies are presented and explored to find solutions for the case of content distribution, where peers will not cooperate without incentives. The authors state, that regarding BitTorrent other current work often use too simplified assumptions about the content availability at specific nodes and therefore overestimate its efficiency.

9 On the Benefits and Feasibility of Incentive-based Routing Infrastructure

In this work of [8], the authors analyze the advantages and disadvantages of the coupling (or lack) between the Internet's underlying business relationships and technical routing protocols. They point out the delineation of the core conflict between user-choice routing and current routing practices and business relationships, by the use of some examples. They observe that both the current disadvantages and the now-developing conflict can be mitigated by a single, simple idea: the introduction of explicitly represented incentives (as prices) into the routing fabric.

They continue by describing possible application of the approach to example routing frameworks, showing that simple changes can address the key conflict and simplify the technical routing environment. They also introduce "The Price of Tyranny," a framework for representing and analyzing the mechanisms discussed.

10 Network Topology Information Desk Service (NTIDS)

The work published in [39] addresses application layer networks (or Overlay Applications) using the Internet, which exist in multiple use cases: Telephony, publishing information, offering file access, DNS and much more in the field of P2P, file sharing, and video streaming sharing applications. Every service is using it's own protocol, forming a network on top of the Internet. Often identical information is published twice or more within such an Overlay Network. Especially files, which are often very large, are published at twenty or more hosts. The search engine finding them, typically does not rate those documents due to remoteness. However, users decide by their own, where they download the file. Thus there is the need to develop a decision metric to optimize load balancing within the Internet, primarily avoiding long distance transfers.

The new approach of merging overlay network distribution with network layer network topology information does this. The approach describes all details of the Network Topology Information Desk Service (NTIDS) and the relating message exchange protocol.

NTIDS is a mixture of functions and data types, which are all experimental in the sense of the IETF. In the sense of this paper some functions and data types define a protocol and some remain still experimental. The official protocol parts include standard queries, responses and the Internet class RR data (*e.g.*, host addresses). The key goal of the service is to offer a mechanism for knowledge transfer between different protocols about the protocol dependent network architecture. NTIDS was originally developed to decrease the lack of knowledge about network layer topology of the IP at the application layer.

NTIDS assumes that the network architecture or the decisive parts of it are stored in a central database or are accessible by a public index or search key.

From the users point of view it is helpful for certain actions to have information about the specific status or behavior of a network or a part of it. Especially information about the (network related) distance or connection status between two hosts within the network architecture is often very decisive. For instance connection or link related problems could be omitted in advance instead of waiting for a timeout. Another advantage is traffic management. A reduction of the number of ways twice traveled is possible.

11 Economics of Overlay Networks: An Industrial Organization Perspective on Network Economics

In the paper [9], the theory of Industrial Organization is used to demonstrate the demand and cost conditions on entry and scaling incentives of Internet overlay networks, restricting the problem to Internet content distribution. The author describes the structure of the content delivery market, the stakeholders, and their business relationships, the nature of demand and the cost-allocation mechanism in both wholesale and retail markets. It is shown how the end-to-end coordination failures of the ISPs and content providers has resulted in wholesale market failures, inducing entry by propriety content distribution “middleboxes” that act as intermediaries, coordinating unmet demands.

The authors also shows that the intermediary has incentives to strategically use the Internet cost allocation mechanism to internalize and scale using these indirect externalities from trade. The scaling of incentives in such overlay markets is also briefly compared with bargaining institutions implemented by P2P content networks. The main contribution of the paper is the in-depth analysis of the markets and the flow of data and money, as long as some insight on the Two-Sided Market theory.

12 LiteLoad*: Content-unaware Routing for Localizing P2P Protocols*

LiteLoad is an ISP-based approach for localizing P2P traffic without caching or packet load inspection [10]. The main idea is to intercept session initiation messages of P2P protocols. If such a message is sent to a destination peer external to the ISP then it is redirect to a local peer. The redirection is done by rewriting the packet header and therefore the ISP does not have to inspect the (probably encrypted) packet payload and is not aware of the shared content. If local peers are not able to answer the query, the requesting peer will try to query further peers. This behavioral pattern is recognized by LiteLoad and such repeated search queries are not redirected. Therefore, the solution is stateful because the LiteLoad server has to keep track of intercepted queries.

The approach is especially applicable to super-peer based overlays, such as KaZaA and eDonkey. Here only the search messages sent to super-peers need to be intercepted.

The system comprises two entities at the ISP side: Filter analyses the destination peer of messages and can be implemented in a router. The detected messages to external peers are forwarded to the actual LiteLoad node. This node is responsible for managing a replacement pool of candidate peers, recognizing the request pattern and replacing the destination address in the message header.

13 Taming *the Torrent*

The paper [11] proposes an approach (called Ono) to reduce cross-ISP traffic. It takes advantage of the fact that content distribution networks (CDNs), spread through the AS on the Internet, and may provide accurate and easily available information about proximity. This system has been implemented as a plug-in for BitTorrent client Azureus. It works by executing periodical DNS lookups to CDN names in order to build and maintain proximity maps. This information of proximity returned by CDNs will reflect the redirection behavior performed and thus an up-to-date view on the Internet. Every peer will compare other's maps with its own and consider the peers with the largest overlap as closest.

The Ono approach is purely overlay-based, i.e. no cooperation from the ISP or CDN owner is required. Moreover, Ono is already being used by more than 100.000 users. The overhead (both for overlay and CDN owner) is minimal and the results show that inter-AS traffic is minimized. The main incentive to use this plugin for users is the improved download performance.

14 Can ISPs and P2P Users Cooperate for Improved Performance?

The work [12] proposes to install an extra component called oracle, located at the ISP. This oracle would take part in making a decision on neighbor selection of peers. The oracle works as a Web server that queries a dynamic database. When a peer has several possible sources to download content, it sends the list to the oracle, which will sort the list of addresses according to the ISP metrics. Usually this means according to locality information, which would be the most profitable for the ISP (simplest case: same ISP or not).

This solution allows localizing the traffic inside the ISP. The approach was applied to Gnutella overlay, where the peer receiving multiple QueryHits selects the source based on the sorting performed by oracle.

15 BASS: BitTorrent Assisted Streaming System for Video-on-Demand

The paper [13] proposes a hybrid server/P2P system. The standard BitTorrent client is used as P2P system. The only modification to BitTorrent is that it does not download chunks which already expired. In addition, the BASS application downloads missing chunks with near playout deadlines from a central media server in order to guarantee a continuous playback of the video. The authors investigate to which extent BASS can reduce the load on the central media server. Their simulations show that BASS achieves savings of 34% compared to pure server-based system.

16 Extending BitTorrent for Streaming Applications

The author of [14] describes necessary adaptations to BitTorrent so that it is capable of delivering VoD streaming data with soft quality of service (QoS) guarantees. He proposes to adapt (1) the chunk selection so that it takes into account whether a chunk is useful and (2) the peer selection so that it picks peers that provide higher QoS.

Furthermore, he proposes QoS extensions to the protocol messages, e.g., “I am interested in pieces 3-12, at 25 kbps sustained rate,…”. However, the paper does not provide any simulation or evaluation of the proposed adaptations.

17 BiToS: Enhancing BitTorrent for Supporting Streaming Applications

The paper [15] studies the impact of the chunk selection mechanism on BitTorrent’s VoD streaming capabilities. The authors argue that only the chunk selection mechanism needs to be adapted. They propose to divide the set missing pieces into a “high priority set” and a “remaining pieces set”. The chunk selection mechanism chooses a piece from the high priority set with probability p and a piece of the remaining pieces set with 1-p.
The authors claim that only the chunk selection mechanism needs to be adapted. However, from my personal point of view, it is questionable if the adaptation of the chunk selection mechanism is sufficient to support streaming which is shown in other studies.

18 LiveBT: Providing Video-on-Demand Streaming Service over BitTorrent Systems

The authors of [16] identify four major challenges when BitTorrent is used as a streaming service: (1) the least-shared-first chunk selection, (2) the unstable client performance when the state of neighbors changes, (3) the download schedule mechanism in combination with the diversity of the capabilities of the neighbors, and (4) the tit-for-tat incentive mechanisms. The paper provides solutions to these challenges and compares them to Bitcomet, a BT client which claims to support video-on-demand.

Simulations show that with LiveBT the download speed is roughly as twice as high as with Bitcomet.

19 Improving VoD Server Efficiency with BitTorrent

Like BASS, the investigated system of [17] called TOAST is a hybrid server/P2P-system. The focus of the investigation is the impact of chunk selection mechanisms on the savings of server-bandwidth. Four chunk selection algorithms are studied: (1) random, (2) in-order, (3) beta, and (4) hybrid. The random strategy and the in-order strategy are straight forward. The beta strategy preferentially requests chunks with near playback deadlines and the hybrid strategy is a combination of in-order and beta.

Compared to BASS, TOAST achieves higher savings. The concrete values depend on the clients' upload bandwidth, on when clients leave the system, and on the used chunk selection mechanism.

20 Analytical Model for BitTorrent-based Live Video Streaming

This work of [18] investigates a hybrid server/P2P system for live video streaming. The authors show that the savings of the required server-side capacity depend on the upload utilization of the peers.

Furthermore, they provide an analytical model which shows peer group size has no impact on the efficiency for group sizes larger than 8. In contrast, the number of fragments available for sharing is crucial. As a consequence, the authors propose to use smaller fragments than standard BT implementations.

21 Offloading Servers with Collaborative Video on Demand

This paper [19] describes the efficiency of a bartering-based P2P system by its entropy, i.e. the probability that two randomly selected peers have no chunk to exchange. Furthermore the paper proposes the biased random chunk selection strategy and collaborative video-on-demand. The biased random strategy selects a chunk with a probability p indirectly proportional to the number of available copies within the neighborhood of a specific peer.

With collaborative VoD, idle peers download chunks on behalf of other peers and increase thereby the availability of chunks in the system. Finally, the paper presents a simulation study which investigates the impact of the number of helpers on the bandwidth savings of the server.

22 2Fast: Collaborative Downloads in P2P Networks

The paper [20] gives a more detailed description of the collaborative download mechanism. The performance is investigated analytically and by simulation experiments. The authors conclude that using the 2Fast mechanism the bottleneck of the upload capacity can be eliminated. They quantify the speedup of the download by a factor of 3.5 compared to other P2P systems.

23 Give-to-Get: Free-riding-resilient Video-on-Demand in P2P Systems

The paper [21] presents a peer selection strategy which aims to prevent free-riding. Tit-for-tat is not suited for VoD as one of the two bartering partners has a later playback position as the other one in most cases. Therefore, it cannot be served by the peer that has the earlier playback position. As a consequence, give-to-get establishes more complex relations between peers. In particular, choking/unchoking decisions are based on how well a peer forwards data to other peers.

To this end, a peer orders its neighbors according to decreasing forwarding ranks. Simulation experiments that a P2P systems tolerates free-riders only in case of spare upload capacity. If the system load is high, only those peers are served which also upload data to other peers.

24 Hybrid Server/P2P System (1)

The three papers on “Providing Video-on-Demand using Peer-to-Peer Networks” [22], “Is High-Quality Feasible using P2P swarming” [23], and “Exploring VoD in P2P Swarming Systems” [24] are considered jointly. The authors present a hybrid server/P2P system. They investigate chunk selection mechanisms and implement different mechanisms on the server and the clients. The video is divided into segments with a fixed number of blocks. The segments are downloaded in playback order whereas different strategies are presented to choose a block of a given segment. The performance of the whole system decreases if the server uploads chunks according to the clients’ needs. Instead, it should take care of the diversity of the chunks in the system.

Furthermore, clients should request blocks of a given segment that are rare in the system. In addition, pre-fetching helps to enhance the overall performance. Pre-fetching means that clients sometimes request blocks of future segments. Finally, the impact of arriving and departing peers as well as the impact of heterogeneous client is investigated. All papers are very similar. [23] and [24] contain only a subset of the issues treated in [22].

25 Hybrid Server/P2P System (2)

The two papers on “Peer-Assisted VoD: Making Internet Video Distribution Cheap” [25] and “Can Internet Video-on-Demand be profitable?” [26] are considered jointly. [25] investigates a hybrid server/P2P system and investigates different strategies for chunk selection. Three strategies are investigated. With the no-prefetching policy, the used download capacity of the peers equals the bandwidth of the video. Data is downloaded in playback order from the server or other peers which have joined the system earlier. The water-leveling policy intends to use spare upload capacity by downloading later chunks in parallel to near chunks. Thereby, the buffers of the peers are treated as water tanks and the lowest one is filled first. With the greedy policy, each user dedicates its remaining upload bandwidth to the next user in terms of the playback position. The authors evaluate the performance of a P2P assisted VoD delivery using the greedy policy. To this end, they simulate a P2P system using the real-world data of the two most popular videos from MSN in April 2006 and compare the bandwidth requirements on the server.

In addition, [26] investigates ISP-friendly peer-assisted VoD. In particular, the authors examine the required server bandwidth when peers only exchange data with peers of the same AS. The authors conclude that ISP-friendly peer-assisted VoD requires more server bandwidth. However, neglecting the interests of the ISPs would probably cause that ISPs drop P2P VoD traffic or forward it with smaller priorities.

26 Peer-assisted VoD for Set-top Box-based IP Network

The paper [27] gives a high-level overview of an architecture that supports VoD using set-top boxes which are controlled by VoD provider. The authors address chunk selection, locality-aware peer selection, QoS, and admission control.

However, no simulation experiments and no analytical performance evaluations are given.

27 A Survey and Comparison of Peer-to Peer Overlay Network Schemes

The comparison of various overlay networks is presented in the paper [28]. The various schemes of P2P overlay networks are categorized into two groups: structured and unstructured in the design spectrum. The application-level network performance of each group is discussed. In the structured networks the topology is tightly controlled and content is placed not at random peers but at specified locations that will make subsequent queries more efficient. The unstructured P2P networks are composed of nodes joining the network with some loose roles, without any prior knowledge of the topology. The main problem of the unstructured systems is that peers readily become overloaded and thus the system does not scale when handling a high rate of aggregate queries and sudden increases in system size. The structured P2P networks incur significantly higher overheads than unstructured ones for popular content, but they are able to efficient locate rare items since the key-based routing is scalable.

In the category of structured P2P overlay networks the CAN (Content Addressable Network), Tapestry, Chord, Pastry, Kademlia and Viceroy are presented. The described networks assign keys to data items and organize its peers into a graph that match each data key to a peer and enables efficient discovery of data items. In the unstructured overlay networks (Freenet, Gnutella, FastTrack/KaZaA, BitTorrent and Overnet/eDonkey2000) the peers are organized in a random graph in a flat or hierarchical manner and use flooding or random walks or expanding-ring Time-To-Live (TTL) search.

28 A Survey of Peer-to-Peer Content Distribution Technologies

The framework for analyzing peer-to-peer content distribution technologies is presented in [29]. The description and analysis of applications, systems and infrastructures that are based on peer-to-peer architectures are described in details. Content distribution applications typically allow personal computers to function in a coordinated manner as a distributed storage medium by contributing, searching and obtaining digital content. The unstructured (Napster, Gnutella, FastTrack 2003, KaZaA) and structured (Freenet, Chord, CAN, Tapestry) P2P network architectures are presented in details.

The content caching, replication and migration issues are described as well as security possibilities related to P2P overlay networks. The very interesting sections of provisions for anonymity and deniability describe the main approaches currently adopted in P2P networks. The information given in the paper may not be used directly in the design process of ETM, but we have to aware of some aspects, e.g., resource management of the content.

29 Resilient Overlay Networks

The Resilient Overlay Network (RON) architecture is proposed in [30]. It allows for distributed Internet applications for decreasing the detection and recovery from path outages and periods of degraded performance time from several minutes to several seconds. The RON nodes monitor the network and use the obtained information to decide whether to route packets directly over the Internet or by way of other RON nodes, optimizing application-specific routing metrics.

The Internet performance studies and network-layer and overlay-based techniques are briefly described in the paper. The design goals of RON (failure detection and recovery in less than 20 seconds, tighter integration of routing and path selection with the application and expressive policy routing) are presented in details. In the main part of the paper the design, implementation and evaluation of RON are presented. The simulation results show that the process from detection to recover from the failure takes 18 sec in the simulated environment, what is a significant advantage with relation to currently used techniques.

30 Overlay Networks – A Scalable Alternative for P2P

This short tutorial-ish paper [31] explains the difference between flooding-style communication (e.g. Gnutella) and overlay networks (based on distributed hash table) presenting the advantages of the latter. It is a good position to start the venture in overlay networks and can be recommended only for the beginners.

Nonetheless upon reading, it is quite easy to understand the difference between this two types of networks. Unfortunately, being published in August 2003, the insight on P2P systems is a bit outdated.

31 An Overlay Architecture for High-Quality VoIP Streams

The paper [32] presents the utilization of the overlay networks for improving the quality of VoIP connections. The authors claim that by combining two mechanisms, namely: real-time packet recovery protocol and specially designed routing protocol, the overall quality of VoIP streams can easily be enhanced. This claims were proved by the extensive simulations, both locally and globally, i.e., by means of PlanetLab. Above that, the solution is viable and relatively easy to be introduced, since it requires no changes to the underlying infrastructure.

Although this work relates to VoIP connections only, the methods described might be used for any kind of real time transmissions in the overlay networks based on the SPINE topology. Unfortunately, BitTorrent is not built on SPINE, still, this paper is an interesting read.

32 Efficient and Secure Peer-to-Peer Overlay Network

The paper [33] proposes a new approach to build P2P overlay network: in contrast to current designs, which focus on their own system overlays, this approach focuses on the physical network, and builds the efficient and secure P2P overlay network following it. The paper proposes a practical and self-certifying identifier for nodes under the open Internet environment, based on node's physical network properties. Additionally, the authors describe secure routing procedure which can be audited. Also, the designed overlay network has the ability to repair itself by ejecting the malicious nodes.

33 Security Issues in Peer-to-Peer Systems

The paper [34] introduces various security issues, methods for protection, and suggestions for further enhancements on the security mechanism in P2P systems. The authors describe threats and various security methods which can be summarized as follow: software and hardware, internal security, trusts issues between peer entities, and various security protocols. Especially, they pay attention to the trust in P2P networks: the guaranteed infrastructure and network applications as well as secure connections between entities in P2P network.

34 Managing P2P Security

The paper [35] presents the security issues of P2P systems from network administrators and designers point of view. It is a survey of managing security in P2P networks. The authors describe many practices and recommendations for managing P2P, for example: establish proper security policy, block access to some resources, perform regular audits, install updates, firewalls, and anti-virus programs.

35 A Proximity-aware Peer Clustering System in Large-scale BitTorrent-like Peer-to-peer Networks

In the paper [36] the authors present a cluster based approach to a large-scale BitTorrent-like peer-to-peer file sharing system. The proposed modification of BitTorrent network is to help to obviate two problems. The first problem is related with random connections among peers in the overlay network, it is very probable that unfavorably distant peers are connected which results in increase of downloading time. The second problem is overwhelmed trackers. Peers are free to join and leave from a torrent, the tracker needs to update its record about the state of peers. The frequent change of peers and requests from new comers can result in the tracker overload in the case of large number of peers.

In the paper author propose the system called Clustered Bit Torrent (CBT) in which they introduce hierarchical architecture.

Peers that are close by each other are grouped in individual cluster. If the number of peers is small the system consists of one cluster (fundamental) and the torrent tracker is the head of the cluster. The tracker is always the head of the fundamental cluster. When the number of peers in the fundamental cluster exceeds a critical value the fundamental cluster is divided and a super-peer is selected. This super-peer plays the role of a local tracker for new cluster (the super-peer is a cluster head). The super-peers are connected with the original tracker and they form a backbone overlay network. A peer joining the system with a few clusters gets a list of all super-peers from the tracker. The new peer joins the cluster in which the cluster super-peer is the nearest to it. The choice is based on the proximity measure in the underlay network, they use in the algorithm the combination of RTT and TTL. In the article they propose the peer joining algorithm and the super-peer selection algorithm and also criteria for the cluster maintenance.

The authors developed a fluid model to compare the performance of the CBT system with the original BitTorrent system. The simulation results show that the CBT system achieves better results then the original BItTorrent system, it is more efficient and it posses better scalability properties. The results demonstrate also that the number of clusters is very crucial for the super-peer load-balancing but too many clusters with the small number of peers can increase downloading time. These two parameters have to be appropriately chosen.

36 Hash-based Proximity Clustering for Efficient Load-balancing in Heterogeneous DHT Networks

In the paper [37], the authors present a hashed-based proximity approach for load-balancing in heterogeneous DHTs. In the presented approach, nodes are classified as supernodes with high capacity and fast connections and as regular nodes with low capacity and slower connections. Supernode network in DHTs is an auxiliary expressway for fast routing between supernodes. Each supernode plays the role of a server for its regular nodes. The routing tables define the interconnections between supernodes and their associated regular nodes. The authors distinguish two types of connections physical and virtual. A physical cluster is a structure in which each node is connected to its physical closest supernode and all supernodes form a DHT. A virtual cluster is a structure in which each node is connected to logically closest supernode in their ID space. The land-marking method is used to represent closeness on Internet by indices. These two structures are used in locality-aware load balancing procedures. The supernodes form a self-organized and churn-resilient auxiliary network for load balancing. The hierarchical structure is used by locality-aware randomized (LAR) load balancing algorithm. A supernode gathers load information of physically close nodes and transfers workload between physically close heavy nodes and light nodes. Hash-based proximity clustering generates a resilient auxiliary supernode network, in which supernodes process load balancing on behalf of their assigned regular nodes.

The authors present also the algorithms for joining and living the clusters and the routing algorithm.

For performance evaluation, authors implemented simulator in Java using Chord DHT and compared the results with two other load balancing approaches CRA and KTree.

The simulation results show the superiority of the authors approach in comparison with the other analyzed randomized and proximity-aware load balancing algorithms. Benefits of proximity clustering come at the cost of cluster maintenance.

37 Stochastic Fluid Theory for P2P Streaming Systems

The paper [38] establishes a reasonable foundation for simplified description, but still giving a meaningful insight into behavior of P2P streaming systems. As is mentioned by the authors this is a first paper dealing with the theory of P2P streaming systems and relevant bibliographical literature is also given.

Specific feature of fluid models is to use rates (either dynamically changed or static in time) instead of typical packet level description. Usually fluid-type models can be solved by differential equations however for static cases also closed-form solution by simple algebraic transformations can be obtained.

In the paper the basic model is the P2P streaming system with peer churn (i.e. in which peers are joining or leaving). This model is extended to a churnless system. Relevant equations are derived to describe an asymptotic behavior of large P2P streaming system. Impact of playback buffering and lag at peers is then studied. Finally, there is an attempt to define an admission policy for P2P streaming systems.

The main results of the paper include:

- A clear and simplified description how to model P2P streaming system in simplified way and to transfer system mechanisms into equations.
- A definition of a critical value (expressed by derived formula) which is applicable to describe a system performance.
- Numerical results giving insights into system behavior, which are obtained from either equations and carried simulation works.

38 P4P: Provider Portal for Applications

The P4P paper [40], which stands for Provider Portal for Applications, presented at ACM SIGCOMM 2008 - a premier computer networking conference in August 2008 in Seattle - proposes a simple architecture called P4P to allow for more effective cooperative traffic control between applications and network providers. The P4P framework aspires to both reduce the cost to ISPs and improve the performance of P2P applications. P4P provides multiple interfaces for networks to communicate with applications regarding: (a) static network policy; (b) p4p distances reflecting network policy and network status, and (c) network capabilities. These interfaces preserve network provider privacy and allow network providers and applications to jointly optimize their respective performance.
The evaluation methodology (BitTorrent simulations, BitTorrent clients runs on PlanetLab nodes in Abilene and conducting field tests using Pando clients) followed and presented in the paper demonstrates that the P4P framework can be a simple, flexible and promising approach to improve both application performance and provider efficiency.

39 References

[1] http://www.dcia.info/activities/#P4P

[2] Xie, H.; Krishnamurthy, A.; Silberschatz, A.; Yang, Y.R.: P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers, under submission, available online: http://www.dcia.info/documents/P4P_Overview.pdf

[3] Y. Chen, Y. Huang, R. Jana, H. Jiang, M. Rabinovich, B. Wei, Z. Xiao: "When is P2P Technology Beneficial for IPTV Services?", IEEE Transactions on Multimedia, 2007.

[4] J. Li: "Locality Aware Peer Assisted Delivery: The Way to Scale Internet Video to the World", Packet Video 2007.

[5] M. Piatek, C. Dixon, A. Krishnamurthy T. Anderson: "LiveSwarms: Adapting BitTorrent for end host multicast", Technical Report UW-CSE-06-11-01, 2006.

[6] Bindal et al.: "Improving Traffic Locality in BitTorrent via Biased Neighbor Selection", ICDCS 2007.

[7] P. Ganesan, M. Seshadri: "On Cooperative Content Distribution and the Price of Barter", ICDCS 2005.

[8] M. Afergan and J. Wroclawski, "On the benefits and feasibility of incentive based routing infrastructure," PINS '04, Portland, Oregon

[9] P. Faratin, "Economics of Overlay Networks: An Industrial Organization Perspective on Network Economics," (NetEcon+IBC), in conjunction with ACM Conference on Electronic Commerce (EC'07) San Diego, California, June '07

[10] S. Horovitz, D. Dolev, "LiteLoad: Content unaware routing for localizing P2P protocols", IPDPS 2008.

[11] D. Choffnes, F. Bustamante, "Taming the Torrent: A practical approach to reducing cross-ISP traffic in P2P systems" *ACM SIGCOMM* 2008

[12] V. Aggarwal, A. Feldmann, C. Scheideler, "Can ISPS and P2P users cooperate for improved performance?" SIGCOMM Computer Communication Revue, Vol. 37, No. 3. (2007), pp. 29-40.

[13] C. Dana, D. Li, D. Harrison, and C. .-N. Chuah, "BASS: BitTorrent Assisted Streaming System for Video- on-Demand," in Proc. IEEE 7th Workshop on Multimedia Signal Processing, 2005, pp. 1-4.

[14] D. Erman, "Extending BitTorrent for Streaming Applications," in *4th Euro-FGI Workshop on "New Trends in Modelling, Quantitative Methods and Measurements"*, 2007.

[15] A. Vlavianos, M. Iliofotou, and M. Faloutsos, "BiToS: Enhancing BitTorrent for Supporting Streaming Applications," in *Proc. 25th IEEE International Conference on Computer Communications INFOCOM 2006*, 2006, pp. 1-6.

[16] J. Lv, et al., "LiveBT: Providing Video-on-Demand Streaming Service over BitTorrent Systems," in *Proc. Eighth International Conference on Parallel and Distributed Computing, Applications and Technologies PDCAT '07*, 2007, pp. 501-508.

[17] Y. R. Choe, D. L. Schuff, J. M. Dyaberi, and V. S. Pai, "Improving VoD server efficiency with bittorrent," in *MULTIMEDIA '07: Proceedings of the 15th international conference on Multimedia*, 2007, pp. 117-126.

[18] S. Tewari and L. Kleinrock, "Analytical Model for BitTorrent-Based Live Video Streaming," in *Proc. 4th IEEE Consumer Communications and Networking Conference CCNC 2007*, 2007, pp. 976-980.

[19] P. Garbacki, D. H. J., J. Pouwelse, and M. v. Steen, "Offloading Servers with Collaborative Video on Demand," in *7th International Workshop on Peer-to-Peer Systems (IPTPS'08)*, 2008.

[20] P. Garbacki, A. Iosup, D. H. J., and M. v. Steen, "2Fast: Collaborative Downloads in P2P Networks," in *6th IEEE Int'l Conference on Peer-to-Peer Computing (P2P2006)*, 2006.

[21] J. J. D. Mol, J. A. Pouwelse, M. Meulpolder, D. H. J. Epema, and H. J. Sips, "Give-to-Get: Free-riding-resilient Video-on-Demand in P2P Systems," in *Multimedia Computing and Networking 2008*, vol. 6818, 2008.

[22] S. Annapureddy, C. Gkantsidis, and P. Rodriguez, "Providing video-on-demand using peer-to-peer networks," in *WWW '07: Proceedings of the 16th international conference on World Wide Web*, 2006, pp. 238-247.

[23] S. Annapureddy, S. Guha, C. Gkantsidis, D. Gunawardena, and P. R. Rodriguez, "Is high-quality VoD feasible using P2P swarming?," in *Is high-quality VoD feasible using P2P swarming?*, 2007, pp. 903-912.

[24] S. Annapureddy, S. Guha, C. Gkantsidis, D. Gunawardena, and P. Rodriguez, "Exploring VoD in P2P Swarming Systems," in *Proc. INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, 2007, pp. 2571-2575.

[25] C. Huang, J. Li, and K. Ross, "Peer-Assisted VoD: Making Internet Video Distribution Cheap," in *Proceedings of Sixth Internaltional Workshop on Peer-to-Peer Systems*, 2007.

[26] C. Huang, J. Li, and K. W. Ross, "Can internet video-on-demand be profitable?," in *SIGCOMM '07: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*, 2007, pp. 133-144.

[27] V. Janardhan and H. Schulzrinne, "Peer assisted VoD for set-top box based IP network," in *P2P-TV '07: Proceedings of the 2007 workshop on Peer-to-peer streaming and IP-TV,* 2007, pp. 335-339.

[28] Lua, Crowcroft, Pias, Sharma et al "A Survey and Comparison of Peer-to Peer Overlay Network Schemes" IEEE Communications Surveys & Tutorials, 2005

[29] Androutsellis-Theotokis, Spinellis "A Survey of Peer-to-Peer Content Distribution Technologies", ACM Computing Surveys, December 2004

[30] Andersen D., Balakrishnan H., Kaashoek F. and Morris R., “Resilient Overlay Networks”, 18th ACM Symp. on Operating Systems Principles (SOSP), October 2001, Banff, Canada.

[31] Doval, O'Mahony "Overlay Networks, A Scalable Alternative for P2P" IEEE Internet Computing, 2003

[32] Amir, Danilov, Goose et al. "An Overlay Architecture for High-Quality VoIP Streams"

[33] Honghao Wang, Yingwu Zhu and Yiming Hu, "An Efficient and Secure Peer-to-Peer Overlay Network", *Proceedings of the IEEE Conference on Local Computer Networks 30th Anniversary* (LCN'05), 2005

[34] Jung-Tae Kim, Hae-Kyeong, Park, Eui-Hyun Paik, "Security Issues in Peer-to-Peer Systems", *7th International Conference on Advanced Communication Technology*, 2005, ICACT 2005

[35] Jason E. Bailes and Gary F. Templeton, "Managing P2P Security", *COMMUNICATIONS OF THE ACM*, September 2004/Vol. 47, No. 9

[36] Jiadi Yu, Minglu Li CBT: "A proximity-aware peer clustering system in large-scale BitTorrent-like peer-to-peer networks", *Computer Communications*, Feb. 2008/vol.31

[37] Haiying Shena, Cheng-Zhong Xub "Hash-based proximity clustering for efficient load balancing in heterogeneous DHT networks", *Journal of Parallel and Distributed Computing*, 2008/vol.68

[38] Rakesh Kumar,Yong Liu, Keith Ross, "Stochastic Fluid Theory for P2P Streaming Systems", Infocom, 2007.

[39] S. Brecht: Investigation of Application Layer Routing; Diploma Thesis No. 17/2003, University of Federal Armed Forces, Munich, Germany, Supervisors: Burkhard Stiller, Peter Racz

[40] Haiyong Xie, Yang Richard Yang, Arvind Krishnamurthy, Yanbin Liu, and Avi Silberschatz, "P4P: Provider Portal for Applications", in Proceedings of ACM SIGCOMM, Seattle, WA, August 2008.

40 Abbreviations

AGH	Akademia Gorniczo-Hutnicza im. Stanislawa Staszica W Krakowie
AS	Autonomous System
BT	BitTorrent
CDN	Content Delivery Network
DCIA	Distributed Computing Industry Association
DiffServ	Differentiated Services
DHT	Distributed Hash Tables
DSL	Digital Subscriber Line
ETM	Economic Traffic Management Mechanisms
ICOM	Intracom
IP	Internet Protocol
IPTV	Internet Protocol Television
ISP	Internet Service Provider
P2P	Peer-to-peer
P4P	Proactive network Provider Participation for P2P

RON	Resilient Overlay Network (RON)
POP	Point of presence
QoS	Quality of Service
SmoothIT	Simple Economic Management Approaches of Overlay Traffic in Heterogeneous Internet Topologies
STREP	Specific Targeted Research Project
TID	Telefónica Investigacíon y Desarollo
TUD	Technische Universität Darmstadt
UniWue	Julius-Maximilians Universität Würzburg
UZH	University of Zürich
VoD	Video on Demand
VoIP	Voice over Internet Protocol

41 Acknowledgements

This deliverable was made possible due to the large and open help of the WP1 and WP2 team of the SmoothIT team within this STREP, which includes besides the deliverable authors as indicated in the document control, a number of additional people as well. Many thanks to all of them.