

# Collaboration between ISPs for Efficient Overlay Traffic Management

Eleni Agiatzidou and George D. Stamoulis

Athens University of Economics and Business  
{agiatzidou, gstamoul}@aueb.gr

**Abstract.** As peer-to-peer (P2P) applications (e.g. BitTorrent) impose high costs to Internet Service Providers (ISPs) due to the large volumes of inter-domain traffic generated, extensive research work has been done concerning locality-awareness approaches. Although such approaches are based on properties of the physical network topology, they take very little consideration of the real business relationships between ISPs (peering and transit agreements) as well as of the heterogeneous peer distributions among ISPs. In this paper, we propose an innovative way to exploit the business relationships between ISPs of either the same or different Tiers, by introducing new collaborative approaches for overlay traffic management on top of the locality-awareness ones. By means of simulations, we show that win-win situation (reduced transit traffic for ISPs and better performance for users) can mostly be achieved under certain complicated collaborative approaches. Also we show how the collaboration between the upper Tier ISPs can affect favorably both the inter-domain traffic in the transit links, where charging applies, and the performance of their customers ISPs and how the latter can benefit from their collaboration with the transit ISPs.

**Keywords:** Overlay Applications, Peering Agreements, Peer-to-Peer Networks, Traffic Management.

## 1 Introduction

Peer-to-peer (P2P) applications, particularly those based on BitTorrent, have become very popular recently, since they allow the users to distribute collaboratively large volumes of content, which account for 43% to 57% of the Internet traffic depending on the geographic area [1]. Despite the advantages that the ISPs and the users gained through the increasing demand for such applications, a significant problem has arisen for the ISPs. Indeed, overlay applications exchange their traffic through logical overlay connections, which are usually agnostic to the physical structure of the Internet. This imposes traffic engineering difficulties and high costs to the ISPs due to the inter-domain traffic exchanged. Extensive research work has been carried out aiming at the reduction of the transit inter-domain traffic and, consequently, of the associated costs of the ISPs. This work led to overlay traffic management techniques, referred to as locality-awareness, which are successful from the ISP's point of view

since they increase the level of locality of the traffic. In Section 2, we present an overview of such approaches. In this paper though, we build on the approach developed by EU-funded project SmoothIT [2], which deals with the incentive-based management of overlay traffic so that this is done beneficially for both the ISP (wrt inter-domain traffic and charges) and users (wrt overlay performance). In most of the locality-awareness approaches (including that of SmoothIT), ISPs provide information about the physical topology to the overlay application clients, thus helping peers to choose their overlay neighbors in a more localized way. However, neither this approach nor the other ones overviewed in Section 2 distinguish non-local peers according to the business relationships [8] among the ISPs. In our opinion, such a distinction is very important, because it may lead to a significant reduction of the transit inter-domain traffic where charging applies, as will be seen in our work.

The specific Locality-Aware BitTorrent peer selection algorithm employed by SmoothIT is based on a combination of Biased Neighbor Selection (BNS) and Biased Unchoking (BU), as presented in [4] and [5], and has already been evaluated successfully for a wide variety of simulated scenarios; see also Section 3. We introduce an innovative variation to this that employs the notion of interconnection agreements in the ranking method of the peers that is used by BNS and BU. Our objective is to investigate the benefits of employing this information for the ISPs and the peers. Moreover, we introduce and evaluate an innovative traffic management approach on top of the locality-awareness techniques that also exploits the business relationships though in a different way. This approach referred to as Splitting of Chunks, aims to avoid the download of redundant (duplicate) content via the costly inter-domain links, while allowing more traffic to be exchanged through the peering links. This approach requires detailed coordination between the collaborating ISPs at the level of swarm, which can be considered as not realistic for practical cases and possibly not acceptable by users. Nevertheless, Splitting of Chunks represents the “ultimate” way for ISPs to collaborate in order to reduce traffic redundancy in the charged transit links, which motivates our studying it. We present the specification of the BNS-BU algorithms that are based on the agreements of the ISPs (Collaborative BNS-BU approaches) and the Splitting of Chunks approach and we compare their performance to the standard BNS-BU algorithm. Note that the practical application of our proposed approaches does not involve any additional overhead compared to SmoothIT BNS-BU, except for a modification in the information already exchanged between the overlay client and the ISP server. To the best of our knowledge, there is no published work up to date that takes into account interconnection agreements in locality promotion or employs some other collaboration between ISPs.

The remainder of this paper is organized as follows: In the next section, we overview approaches for locality promotion proposed in the literature, while in Section 3 we describe briefly the BNS and BU algorithms. In Section 4, we present the simulation setup used for our experimental evaluation. Sections 5 and 6 follow the same structure; the first part presents the analysis of the Collaborative BNS-BU approaches and Splitting of Chunks mechanism respectively, while the second part is dedicated to the simulation results of the aforementioned mechanism. Finally in Section 7 we present our conclusions and future work.

## 2 Related Work

Extensive research work has already been done in the literature on the promotion of locality in P2P networks. In this section we present an overview of proposals for locality awareness mechanisms as well as research work that indicates arising issues resulted from these mechanisms that provided an extra motivation for our work.

Aggrawal et al. propose an Oracle service hosted by the ISPs in order for them to cooperate with P2P users [3]. The Oracle ranks a list of potential neighbors based on locality (same AS) and proximity (AS hops distance). The evaluation results showed that the properties of the overlay topologies, such as small diameter, small mean path lengths and node degree are not affected by the use of the oracle service, while at the same time the network locality increases. A similar approach where again an overlay entity called iTracker communicates either with the peers or with application trackers providing them information about the underlay, is P4P project [6]. The simulations in PlanetLab and real networks showed a reduction in transit inter-domain traffic and download times as well. Bindal et al. [4] in their Biased Neighbor Selection (BNS) algorithm presume that a peer can be provided by a biased list of peers (in order to connect to). The simulation results over 14 different ASes showed that the transit inter-domain traffic is reduced significantly, while the download times of the peers are not influenced much. Oechsner et al. propose in [5] a new algorithm, referred to as Biased Unchoking (BU) and they combine it with the aforementioned BNS algorithm (BNS-BU). BU influences, based on the ranked list, the optimistic unchoking algorithm of a peer that indicates to which neighbor to upload. This list is received from the tracker, unlike SmoothIT and our approaches, where it is received from an ISP-owned overlay server; see Section 3. The experimental evaluation and comparison of plain BitTorrent, BNS, BU and BNS-BU in [5] showed that the two complementary mechanisms should be used together in order to achieve better performance in terms of transit inter-domain traffic while the performance of the peers remains unaffected. A different implementation of the same idea is presented by Choffnes and Bustamante in [11]. Biased neighbor selection is based on the information that is collected from DNS lookups on popular CDNs names. The similarity of the lookups of two peers determines their level of proximity.

In [13], Piatek et al. present three pitfalls to ISP-friendly P2P design. One of them concerns the conflicting interest that different Tier ISPs have while inter-domain traffic is reduced causing less costs for some and less revenues for others. Wang et al. in [9] examine the different contents and peer properties in regards to the locality issues. They conclude that the peers belonging to a few large AS clusters are more eligible to be affected by a locality mechanism; thus, a selective locality mechanism is more promising to optimize the overhead and the robustness of BitTorrent. Lehtieder et al. in [10] investigate the way BNS and BU mechanisms affect the performance of the users. For scenarios with homogeneous peer distributions and all peers having the same access speed a win-no lose situation arises, i.e. reduction of transit inter-domain traffic – no deterioration of performance of users. Changing the percentage of local peers used for either BNS or BU, they argue that the actual impact for a specific peer depends heavily on that percentage as well as the topology used.

### 3 Locality-Aware Mechanisms: BNS-BU

In this section we briefly present the two main approaches of locality-awareness, Biased Neighbor Selection (BNS) and Biased Unchoking (BU) [4], [5] since we build our collaborative mechanisms on top of their combination (referred to as BNS-BU), as implemented by SmoothIT [2]. Both approaches are based on the existence of an overlay server that resides in each AS and accepts from peers, lists with remote peers acquired from a tracker. Its main role is to rank the remote peers in the list according to locality (same AS with the requesting peer, i.e. local ones), peering agreements and proximity (AS hops away from requesting peer's AS) [2]. Thus, the overlay server maps a specific value to every local peer, forming the *highest* or *1<sup>st</sup>-value* group and another value to every peer that belongs to the peering ASes, forming the *2<sup>nd</sup>-value* group. The rest of the peers are ranked according to the proximity of their ASes based on BGP hops. More AS hops mean that the peers belong to *smaller-value* groups.

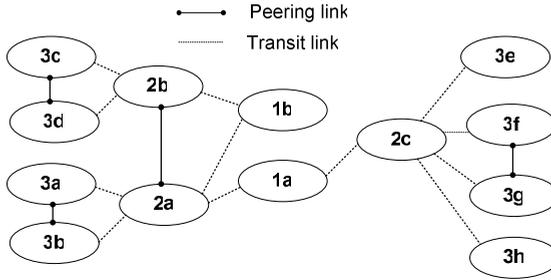
BNS enables a peer that had already received a ranked list from the overlay server to connect to the most preferable peers. In our implementation, as also implemented in [5], 90% of a peer's neighbors are chosen based on their ranking. The peer starts connections to the *1<sup>st</sup>-value* group until it reaches the threshold of 90%. If the *1<sup>st</sup>-value* group contains fewer peers than those needed to reach 90%, then the peer iteratively connects to peers from the next value group, i.e. from the peering AS or *2<sup>nd</sup>-value* group, then from the ASes one hop away etc. After reaching 90% of the connections, it randomly chooses the remaining neighbors from the rest of the groups.

In BU a peer chooses to unchoke one of his interested (in the currently available chunks) neighbors that belong to the *highest-value* group, i.e. to the same AS according to the metrics that obtained from the ranked list. If there are no interested peers in this group, then the peer chooses randomly from the next one.

### 4 Simulation Setup

In this section we present the basic configuration of our simulations that we used to compare the BNS-BU mechanism with our variations that introduce collaboration between ISPs from the same or different Tiers and with Splitting of Chunks approach.

In Figure 1 we defined the topology for our simulations. Each AS is distinguished through an identification code, which consists of a number that indicates the Tier that this ISP belongs to, and a letter that distinguishes it from others of the same Tier. There are eight Tier 3, three Tier 2 and two Tier 1 ISPs. Tier 2 ISPs have twice as many peers as Tier 3 ones, a set up that realistically reflects Internet. To be able to show the function of the collaboration mechanisms we introduced five peering agreements (links) between ISPs of the same Tier. Furthermore each ISP contains an overlay entity that except for ranking the remote peers of a requesting peer is aware of the business relationships of the ISP and adjusts its policy to the various ranking approaches analyzed in the next session. Moreover, this server assists peers to select which part of the content will download from each peer in the Splitting of Chunks approach. This topology is rich and representative enough to motivate our mechanisms and better exploit their potential by employing the business relations



**Fig. 1.** Topology used in Simulation

among ISPs of two different Tiers; this would not have been possible with a topology comprising two Tiers only. At the same time the topology is simple enough for the results to be easy to analyze and comprehend.

In our experiments we used the SmoothIT simulator, which is based on Protopeer [7]. In order to incorporate the approaches introduced in this paper, we extended the simulator accordingly. The peers are connected to the ASes with access speed equal to 16Mbps for download and 1Mbps for upload, roughly the average speed values of DSL lines today. The network delay among all peers is 10msec regardless their physical location. All inter-domain links have symmetrical capacities, while the access links of peers constitute the only bottlenecks in the network. All peers exchange a file whose size is 154.6MB. For simplicity reasons we simulate one single swarm for evaluating our mechanisms. Peers arrive in the swarm following a deterministic distribution. Every 30 and 60 sec Tier 2 and Tier 3 ISPs respectively acquire one more peer. Also, all peers after downloading the whole file, they seed it for 3 minutes before they exit the system. Thus, as measured during the simulations, the swarm contains about 100 to 200 peers on average, a typical size for medium-sized swarms, as reported in [12].

The main metrics of interest are the mean inter-domain traffic exchanged between ISPs 2a and 3a and between ISPs 1a and 2a, as well as the mean download times of the peers in 3a and 2a. The former is used to assess the savings for the ISPs, while the latter to assess the impact of the mechanism on the user performance. We calculated the mean values per minute for 10 runs of each experiment and the confidence intervals for a confidence level of 95%. The size of confidence intervals was below 6% of the corresponding mean value.

## 5 Collaborative BNS-BU Approaches

### 5.1 Approach

Transit agreements between ISPs are used in order to provide connectivity to the rest of the Internet. Peering agreements on the other hand, help ISPs to reduce their transit inter-domain traffic and, therefore, the implied costs, which are comparatively greater than the cost of maintaining a peering link [9]. In this paper we focus on transit and peering agreements between different or same Tier ISPs.

In this section we propose the modification of the BNS-BU approach in such a way that takes into consideration the business relationships of the ISPs in ranking the peers in a list upon such a request. To explain this better, let us assume that the overlay entity returns to the requesting peer a pair of values i.e. {remote peer IP address, remote peer group value} ( $a$ ,  $g$ ), one for each peer in the list. Thus, the requesting peer uses  $g$  to identify the group to which remote peer  $a$  belongs, as mentioned in Section 3. The overlay entity will map the highest  $g$  to local peers (relatively to the requesting peer), the second higher  $g$  to peers from the peering AS and then classifies the rest of the peers according to BGP hops; more hops away means smaller values. Nevertheless, if two ASes are the same number of hops away, this does not necessarily indicate that the exchange of traffic with each one of them will cost the same. For example, a peer in 2a will receive the same  $g$  value (thus, forming one group) for peers from 3c, 3d, and 2c. However, exchanging traffic with 3c and 3d is cheaper than exchanging traffic with 2c, since, as shown in Figure 1, 2a and 2b have a peering link with each other. Preferably the peer could connect to as many as possible peers from 3c and 3d before connecting to 2c. This knowledge that results from the business agreements of the ISPs, is not available explicitly to ISP 2a, but since 2a and 2b collaborate with each other, they can exchange the preference values (from where this knowledge is extracted) that the overlay entity of each ISP gives to other ASes. In such cases peer selection can be optimized using the knowledge regarding the business relationships among ISPs in order for groups to be formed differently. We propose two collaborative approaches that take advantage of this knowledge, referred to as Collaborative BNS-BU and Layered Collaborative BNS-BU.

Collaborative BNS-BU takes into consideration the business relationships of the ISPs in order to form the groups of peers differently from BNS-BU. In the  $1^{st}$ -value group belong local peers or peers from the peering AS, thus forming a larger group than in BNS-BU. The  $2^{nd}$ -value group consists of peers from the customer ASes of the ISPs. Larger groups were chosen in order to provide peers with more resources, boosting this way their performance. In case that ISPs 2a and 2b collaborate using this approach, their  $1^{st}$ -value group consist of peers from 2a and 2b while the  $2^{nd}$ -value group is formed by peers from the customer ASes, i.e. 3a, 3b, 3c, 3d. The rest of the groups are formed using AS hops. Tier 3 ISPs, when collaborate with each other cannot affect much the grouping since they know only their own business relationships as opposed to the Tiers above. Thus only the *highest-value* group of locals and peers from peering ASes is formed and the rest stay unaffected. On the other hand, if Tier 3 overlay entities collaborate with Tier 2 one, the  $2^{nd}$ -value group of Tier 3 ISPs 3a and 3b consists of peers from 2a, 2b, 3c and 3d.

Layered Collaborative BNS-BU uses a slightly refined technique in order to form as many groups as possible according to the business relationships. This approach differs from the previous one in the number of groups. Thus, under this mechanism a peer first connects to all resources found locally before it chooses a peer from the next group, i.e. from the peering ISP. Thus, the first group, as formed in the Collaborative BNS-BU case, is split in two groups, exactly as in the plain-BNS-BU. The  $2^{nd}$ -value group of the previous approach is split to multiple smaller groups according to BGP hops. Smaller groups are used in order to further promote proximity. Indeed, due to such small groups, a peer will connect to all of the peers in a 'closer' group before moving to the next one. As an example, a peer in ISP 2a would receive the following

group of peers:  $1^{st}$  group - local peers,  $2^{nd}$  group - peers from 2b,  $3^{rd}$  group - peers from customers of 2a (3a, 3b) and  $4^{th}$  group - peers from the customers of 2b (3c, 3d). The rest of the ASes are formed into groups according to the BGP hops. It is worth noticing that for Tier 3 ISPs this approach forms groups that are identical to the plain-BNS-BU groups unless cross-Tier collaboration exists (e.g. Tier 2-Tier 3).

## 5.2 Experimental Evaluation

In Table 1 we present the effect of the aforementioned approaches to the transit inter-domain traffic (in MB/s) of 1a2a and 2a3a links, as well as to the download times (in min) that peers in ISPs 2a and 3a experience. Each column of the table corresponds to a different scenario; *BNS-BU*, *Collaborative BNS-BU* and *Layered Collaborative BNS-BU*. Each scenario is separated to multiple columns, each one corresponding to groups of simulations; for each group a different pair of ISPs runs the specific scenario. For each scenario the rest of the ISPs run BNS-BU since we consider that is in the interest of ISPs to run those two algorithms. For example, *Collaborative BNS-BU Tier 2* means that ISPs 2a and 2b use the Collaborative BNS-BU approach while all other ISPs run BNS-BU. *Collaborative BNS-BU Tier 3* means that Tier 3 ISPs 3a and 3b only use the specific approach. *Tier 2 & Tier 3* refers to the case where 2a and 2b as well as their customers follow the specific approach. Each row comprises the mean values of a specific metric, i.e. the mean inter-domain traffic 1a2a, the mean inter-domain traffic 2a3a, and the mean download times in 2a and in 3a.

**Table 1.** Collaborative BNS-BU approaches

Metric / Scenario	BNS-BU	Collaborative BNS-BU			Layered Collaborative BNS-BU	
		Tier 2	Tier 3	Tier 2&3	Tier 2	Tier 2&3
Traffic 1a2a (MB/s)	1.09	0.8	1.01	0.61	0.8	0.63
Traffic 2a3a (MB/s)	11.11	11.29	11.03	11.54	11.02	10.86
DT 2a (min)	2.92	2.99	2.98	2.03	3.06	3.03
DT 3a (min)	8.95	8.85	9.01	9.11	8.96	9.06

As shown in Table 1, when only ISPs 2a and 2b run Collaborative BNS-BU (*Collaborative BNS-BU Tier 2*), they manage to reduce their transit inter-domain traffic by 26% essentially without affecting peers' performance, thus attaining a win-no lose situation. Traffic between 2a and 3a and download times for peers in 3a are marginally affected (1.6% and 1.1% respectively). Larger download times for peers in 3a are caused by the use of BNS-BU mechanism. Thus, for *Collaborative BNS-BU Tier 2* there is a win-no lose situation for ISPs in both Tiers (2a and 3a). For ISPs 3a and 3b, the transit inter-domain traffic is not significantly affected (*Collaborative BNS-BU Tier 3*); at the same time this has very little effect on peers' performance in both Tier ISPs and on the inter-domain traffic 1a2a. However, if ISPs 3a and 3b run Collaborative BNS-BU and 3c and 3d do so too, then Tier 2 ISPs 2a and 2b have a clear incentive to also run Collaborative BNS-BU. Indeed, this extra collaboration between them will improve the transit inter-domain traffic 1a2a by 44% and considerably improve the download times of peers in 2a (30.5%), (see column *Collaborative BNS-BU Tier 2&3* in Table 1). Since in this case Tier 3 ISPs are

indifferent in adopting the approach, ISP 2a can incite it to its customers by sharing its own benefits with them. Similar results are obtained from the implementation of the Layered Collaborative BNS-BU for Tier 2 ISPs and for Tier 2 & 3 ISPs, due to the large amount of resources from the local and the peering ISPs. However, download times deteriorate slightly more in this case, due to the extra constraints imposed by the finer layering. Thus, we conclude that especially large ISPs, such as Tier 2 ones, gain from deploying Collaborative algorithms for BNS-BU approaches due to fact that their peers can find larger amounts of resources locally, from their peering ISP and their customers ISPs; thus, inter-domain traffic exchanged through links for which the Tier 2 ISPs are charged can be reduced.

## 6 Splitting of Chunks

### 6.1 Approach

On top of the aforementioned mechanisms we develop an innovative traffic management approach, referred to as Splitting of Chunks. This mechanism is based on the collaboration of ISPs that have a peering agreement with each other; this provides them the opportunity to further collaborate in order to reduce the redundancy in the downloading content from non-local peers (or ISPs). The proposed mechanism suggests that the peering ISPs can share costs by deciding to split and download different parts of content through their transit links and then exchange the rest of the parts via their peering link. Indeed, as each peering ISP will download different chunks using its transit link, it will be charged less by the ISP of the upper Tier.

Peering ISPs first decide which chunks of the content will download from remote ASes. The chunks can be partitioned in two subsets of equal size according to their ids, e.g. in even and odds or first half and second half or any other similar way. The rest of the chunks that each ISP does not download from the remote ASes will be retrieved from the peering link or from local peers. As in the plain BNS-BU, the peer may identify its neighbors by the ranked list of peers that it retrieves from the overlay server. Since we consider that the peers already run BNS-BU algorithms only few modifications have to be implemented. In particular, the ranked list needs to contain another field, which indicates the type of the chunks (or in other words the set of ids) that the peer is allowed to download from every peer in the list. In our evaluation we split the chunks according to even and odd ids. In this case the list contains triplets of  $\{address - group's\ value - chunkTypeID\}$ ,  $(a, g, c)$ . Thus, the querying peer uses  $g$  to identify proximity of peer  $a$  for BNS and BU algorithms and  $c$  to identify which chunks it can download from this particular peer. Let us suppose that ISPs 3a and 3b (as in Figure 1) are collaborating using the proposed mechanism and have decided that peers from 3a will download the even chunks of the content ( $c=2$ ) while peers from 3b the odd ones ( $c=1$ ). Then a peer from 3a receives three triplets, one for a peer from 3a,  $(a_1, g_1, c_1)$ , one for a peer from 3b  $(a_2, g_2, c_2)$  and one for a peer from 2a  $(a_3, g_3, c_3)$ . The first peer belongs to 3a, thus  $g_1=1$  and  $c_1=0$ , which indicates that  $a_1$  peer is local and the requesting peer can download every chunk therefrom. For the second peer (in 3b)  $g_2=2$  and  $c_2=0$ , which indicates that  $a_2$  peer belongs to a peering ISP and the peer can download every chunk too. Finally for the third peer (in 2a)  $g_3=3$  and  $c_3=2$ , which indicates that  $a_3$  peer belongs to an ISP 1 hop away and the peer can download only even chunks.

Also in order for a peer to download content in accordance to Splitting of Chunks, two extra modifications have to be introduced. In particular, the peer once acquiring the message with the list of the available chunks of a neighbor, it has to decide if it is interested in any of the chunks this neighbor has (and eventually download them), as normally happens in BitTorrent mechanism. Therefore it needs to check whether the neighbor possesses any of its missing chunks and also if those chunks belong to the subset of chunks that the field  $c$  indicates for that neighbor. If it does so, then the peer declares its interest to that neighbor and requests the specific chunks.

## 6.2 Experimental Evaluation

In this section, we present the experimental performance of Splitting of Chunks mechanism and its comparison to BNS-BU mechanism with or without collaboration between ISPs. In the first set of simulations, we run Splitting of Chunks on top of BNS-BU combination mechanism in order to evaluate our approach. The columns of Table 2 present the different scenarios; *BNS-BU* and *Splitting of Chunks*. BNS-BU runs in all ISPs of our topology while there are different set of experiments for Splitting of Chunks. *SC Tier2* refers on the sets of experiments where only 2a and 2b run Splitting of Chunks. Respectively, *SC Tier 3* refers to the sets of experiments where 3a and 3b support this approach while in *SC Tier 2 & Tier 3*, 2a, 2b, 3a and 3b support it. For all experiments the rest of the ISPs in our topology support BNS-BU. The rows of Table 2 comprise the mean traffic volume for inter-domain links 1a2a and 2a3a and the mean download times for peers that belong to 2a and 3a ASes.

In *SC Tier 2* scenario, ISP 2a manages to reduce the inter-domain traffic exchanged with its transit ISP (1a) by 15.6%. The traffic exchanged via the transit link with its customers remains unaffected ( $\sim 1\%$ ) as well as the download times for peers in 2a and 3a ( $< 4\%$ ). Also, the inter-domain traffic 2a3a is slightly reduced too. Thus, if ISP 2a runs Splitting of Chunks, there is a win-no lose situation, since it reduces transit traffic without affecting the performance of its users. At the same time there is a win-no lose situation for its customer ISPs, since the action of ISP 2a affects neither their traffic nor their users' performance. In fact, ISP 2a would not like to have caused a significant reduction of the transit traffic of its customers, because this would affect adversely its own revenue.

From Tier 3 ISPs' (3a and 3b) perspective, in case they agree to support the Splitting of Chunks approach, they reduce almost 30% their inter-domain traffic exchanged with their transit ISP. On the other hand, the performance of their users is strongly deteriorated and thus it is highly unlikely for them to proceed with such an approach. In fact, even if Tier 2 ISPs proceed in adopting Splitting of Chunks, Tier

**Table 2.** Splitting of Chunks comparison with BNS-BU mechanism

Metric / Scenario	BNS-BU	SC Tier 2	SC Tier 3	SC Tier 2 & 3
Traffic 1a2a (MB/s)	1.09	0.92	2.2	1.69
Traffic 2a3a (MB/s)	11.11	10.99	7.84	7.74
DT 2a (min)	2.92	3.04	3.00	2.99
DT 3a (min)	8.95	9.25	11.19	11.28

3 ones will not make the same decision, despite the significant gain in their transit inter-domain traffic, since their users' performance is still strongly affected. To reach safe conclusions according to the problem that Tier 3 ISPs face, we conducted experiments using also Collaborative BNS-BU only in Tier 3 ISPs 3a and 3b. In a similar way is not in the interest of a small Tier 3 ISP to proceed to such a collaborative approach, due to the additional delays in download times that its users are experiencing.

Since the introduction of Splitting of Chunks in Tier 2 ISPs may be considered as applicable and promising, we continue our experimental analysis by introducing also Collaborative BNS-BU and Layered Collaborative BNS-BU in Tier 2 ISPs and combining these approaches with Splitting of Chunks. In Table 3 we present the previous results for BNS-BU, Collaborative BNS-BU, Layered Collaborative BNS-BU and the Splitting of Chunks approach on top of those algorithms. The first column of the table comprises *BNS-BU* scenario. The second and the third column refer to two scenarios based on *Collaborative BNS-BU in Tier 2* ISPs 2a and 2b (second column) and *Splitting of Chunks* on top of Collaborative BNS-BU in Tier 2 ISPs 2a and 2b (third column). In the last two columns we replace Collaborative BNS-BU with Layered Collaborative BNS-BU. The rest of the ISPs run BNS-BU.

The effect of Splitting of Chunks on top of Collaborative BNS-BU, either with or without layers, is almost the same. The new approach results in an inter-domain traffic reduction about 9% compared to Collaborative BNS-BU and 33% compared to BNS-BU, without essentially affecting users' performance in both Tiers. Thus, Tier 2 ISPs that have a peering agreement can adopt Splitting of Chunks on top of any BNS-BU approach that they follow.

Since Tier 2 ISPs can be expected to implement Splitting of Chunks on top of Collaborative BNS-BU approaches adopted only by them, we also have to investigate the *reaction* of Tier 3 ISPs. In Table 4 there are two groups of experiments; the one with Collaborative BNS-BU and the second with Layered Collaborative BNS-BU. The first column of each group comprises the corresponding BNS-BU that 2a, 2b and their customers run. The second, third and fourth column refers to Splitting of Chunks approach when only Tier 2 ISPs 2a and 2b, Tier 3 ISPs 3a and 3b and Tier 2 & 3 ISPs 2a, 2b, 3a and 3b run respectively but all of the aforementioned ISPs and their customers run Collaborative BNS-BU; this is different from the corresponding scenario of Table 3, where only 2a and 2b run Collaborative BNS-BU. Also the rest of the ISPs run BNS-BU. When only Tier 2 ISPs 2a, 2b run Splitting of Chunks on top of Collaborative BNS-BU, they manage to reduce their transit inter-domain traffic

**Table 3.** Splitting of Chunks combined with BNS-BU collaboration approaches in Tier 2

Metric / Scenario	BNS-BU	Collaboration		Layered Collaboration	
		BNS-BU Tier 2	SC Tier 2	BNS-BU Tier 2	SC Tier 2
Traffic 1a2a (MB/s)	1.09	0.8	0.73	0.8	0.74
Traffic 2a3a (MB/s)	11.11	11.29	11.16	11.02	11.26
DT 2a (min)	2.92	2.99	2.97	3.06	3.03
DT 3a (min)	8.95	8.85	8.88	8.88	9.02

**Table 4.** Splitting of Chunks combined with BNS-BU collaboration approaches in all Tiers

Metric / Scenario	Collaboration				Layered Collaboration			
	BNS-BU Tier 2&3	SC Tier 2	SC Tier 3	SC Tier 2&3	BNS-BU Tier 2&3	SC Tier 2	SC Tier 3	SC Tier 2&3
Traffic 1a2a (MB/sec)	0.61	0.43	0.8	0.4	0.63	0.61	0.89	0.41
Traffic 2a3a (MB/sec)	11.54	11.9	8.17	11.8	10.86	11.05	7.69	11.15
DT 2a (min)	3.02	2.99	2.95	2.98	3.03	3	2.94	2.93
DT 3a (min)	9.11	8.75	11.1	8.89	9.06	9.06	11.31	8.82

by 30% (60% compare to BNS-BU) essentially without affecting peers' performance, thus attaining a win-no lose situation. Traffic between 2a and 3a and download times for peers in 3a are marginally affected (3% and 3.9% respectively, although download times are better). Thus, there is a win-no lose situation for ISPs in both Tiers.

Tier 3 ISPs will not adopt Splitting of Chunks as in the previous scenarios, due to slower download times of their users. In case that Tier 2 ISPs already deploy Splitting of Chunks on top of Collaborative BNS-BU, Tier 3 ISPs are indifferent in adopting the approach. However, in the case where all adopt Splitting of Chunks we observe the lowest value of inter-domain traffic 1a2a (34%, 63% compared to Collaborative BNS-BU and plain BNS-BU respectively). Again win-no lose situations arise between Tier 2 ISPs and their users and also between Tier 2 ISPs and their customers ISPs. Therefore, ISP 2a may ask its customers to follow this approach sharing the benefits with them. The same observations are applicable to the case of Layered Collaboration BNS-BU. Besides comparing mean download times, we have also measured the difference between the 95<sup>th</sup> percentile value of each scenario and BNS-BU; for all scenarios (except for SC Tier 3 approaches), was below 3%. We also have run experiments where the two large peering ISPs and their customers, have a different amount of peers in the swarm. We have reached the same conclusions as before, though the download times of the peers were deteriorated slightly more due to the lack of resources in the peering ISP.

Thus, we conclude that even in case of a different approach such as Splitting of Chunks, large ISPs gain from deploying it due to fact that their peers can find larger amounts of resources locally or from the customers of their ISPs and thus inter-domain traffic exchanged through costly transit links can be reduced.

## 7 Conclusions and Further Work

In this paper, we introduce and evaluate collaborative variations of the BNS-BU (Biased Neighbor Selection - Biased Unchoking) locality promotion approach, namely Collaborative and Layered Collaborative, BNS-BU, as well as the Splitting of Chunks approach. All of them exploit the business relationships between ISPs of either the same or different Tiers in order to reduce the costly transit inter-domain traffic. By means of simulations, we show that large ISPs (Tier 2 in our simulations) can attain this goal, while in general not deteriorate the performance of their users. Indeed, in case peering Tier 2 ISPs collaborate with each other and with their customer ISPs,

running Collaborative BNS-BU, a win-win situation for the large ISPs and their users is arising, without affecting the transit inter-domain traffic and the performance of the users of Tier 3 ISPs. The introduction of Splitting of Chunks on top of the aforementioned scenario leads to additional reduction of the transit inter-domain traffic of Tier 2 ISPs, but just maintains user performance, thus leading to a win-no lose situation. In general, we conclude that large ISPs have the incentives to collaborate with each other, creating large clusters within which they promote locality without affecting the performance of their users, while they can further benefit if they provide their customers ISPs with the incentive to collaborate with them. In future research, we plan to focus on the case of P2P Video-on-Demand applications, and consider different topologies and set-ups for our simulations, e.g. asymmetric cases.

**Acknowledgments.** This work has been performed in the framework of the EU ICT Project SmoothIT (FP7-2007-ICT-216259). The authors would like to thank all SmoothIT partners for useful discussions on the subject of the paper. Also, this work has been partly funded by the European Social Fund and National Resources (Greek Ministry of Education - HERAKLEITOS II Programme).

## References

1. Ipoque. Internet Study (2008/2009), <http://www.ipoque.com/>
2. SmoothIT Simple Economic Management Approaches of Overlay Traffic in Heterogeneous Internet Topologies, <http://www.smoothit.org/>
3. Aggarwal, V., Feldmann, A., Scheideler, C.: Can ISPs and P2P users cooperate for improved performance? ACM SIGGOMM Computer Communications Review 37(3), 29–40 (2007)
4. Bindal, R., Cao, P., Chan, W., Medval, J., Suwala, G., Bates, T., Zang, A.: Improving traffic locality in bittorrent via biased neighbor selection. In: 26th IEEE International Conference on Distributed Computing Systems, p. 66. IEEE, Washington, DC, USA (2006)
5. Oechsner, S., Lehrieder, F., Hofffeld, T., Metzger, F., Pussep, K., Staehle, D.: Pushing performance of Biased Neighbor Selection through Biased Unchoking. In: Peer-to-Peer Computing, pp. 301–310 (2009)
6. Xie, H., Richard, Y., Krishnamurthy, A., Liu, Y.G., S.: P4p: provider portal for applications. SIGCOMM Comput. Commun. Rev. 38, 4 (2008)
7. Protopeer, <http://protopeer.epfl.ch/faq.html>
8. Chang, H., Jamin, S.: To peer or not to peer: Model the evolution of the Internet's AS-level topology. In: Infocom (2006)
9. Wang, H., Liu, J., Xu, K.: On the locality of BitTorrent-based video file swarming. In: 8th International Workshop on Peer-to-Peer Systems, IPTS 2009 (2009)
10. Lehrieder, F., Oechsner, S., Hofffeld, T., Despotovic, Z., Kellerer, W., Michel, M.: Can P2P-Users Benefit from Locality-Awareness? In: Peer-to-Peer Computing (2010)
11. Choffnes, D.R., Bustamante, F.E.: Taming the Torrent: A Practical Approach to Reducing Cross-ISP Traffic in Peer-to-Peer Systems. SIGCOMM (2008)
12. Hofffeld, T., Hock, D., Oechsner, S., Lehrieder, F., Despotovic, Z., Kellerer, W., Michel, M.: Measurement of BitTorrent swarms and their AS topologies. Tech. Rep. 463, University of Wurzburg (2009)
13. Piatek, M., Madhyastha, H.V., John, J.P., Krishnamurthy, A., Anderson, T.: Pitfalls for ISP-friendly P2P design. In: HotNets (2009)