Survivability of P2P Multicasting

Krzysztof Walkowiak

Chair of Systems and Computer Networks
Wroclaw University of Technology
Wroclaw, Poland
krzysztof.walkowiak@pwr.wroc.pl

Abstract— **At the present time there are many Peer-to-Peer (P2P) multicasting systems supporting live streaming, i.e. real-time dissemination of various content issued at a source node towards a set of receivers. The objective of content distributing can be twofold: maximization of the system throughput (i.e. the streaming rate) and minimization of the streaming cost while guaranteeing the particular streaming rate. In this we paper we focus on P2P multicasting applied for delivering very important content, which need delivery guarantees, e.g. weather forecast, hurricane warnings, distribution of security updates, stock exchange data, etc. To protect the system against failures several (at least two) multicasting trees are established. We tackle the question of how additional survivability constraints guaranteeing failure-disjoint trees influence the cost and throughput of the P2P multicasting system. We consider failures of the following network elements: overlay link, upstream node and ISP link. To investigate the problem we use both simulations and offline optimization methods. According to our results, the additional requirements necessary for protection do not have a substantial impact on the P2P multicasting system performance.**

Keywords-P2P multicasting, protection, survivability

I. INTRODUCTION

Recently, many new services and applications are introduced to the Internet. Many of them are related to streaming of various content to many users, e.g. VoD, IPTV, radio, news channels, etc. In many cases – to reduce the investment costs needed to deploy these services – the concept of P2P multicasting is applied. Overlay P2P multicasting uses a multicast delivery tree constructed among peers (end hosts). Different to traditional IP multicast, the uploading (non-leaf) nodes in the tree are normal end hosts [7]. The content to be distributed through P2P multicasting can be divided into two categories: elastic content (e.g. data files), and streaming content with specific bit rate requirements (e.g. media streaming) [23], [24]. It should be noted that in many previous works there is an assumptions that peers of the P2P multicasting system are potentially vulnerable to frequent leaving the multicast session. In this work we assume a different scenario, i.e. the P2P multicasting system is relatively *static* – all peers interested in participating in the system do not leave the session so frequently as in P2P file sharing systems. Examples of such static systems are:

- P2P multicast system applied for dissemination of critical information (stream or data), e.g. weather forecast (hurricane warnings), security software updates, stock exchange data, traffic information, etc.

- IPTV system using STBs (set-top boxes) [6].

- CDN (Content Delivery Network).

Consequently, we can focus on network failures and do not consider issues related to dynamics of P2P systems. Since, the transmitted content is assumed to be of great importance, the P2P multicasting must be provided with delivery guarantees. Our motivation comes from survivability mechanisms developed for connection-oriented network using unicast transmission. Two basic approaches were proposed to provide network survivability: protection and restoration. The distinction between protection and restoration consists in the different time scale in which they operate. Protection needs preallocated network resources while restoration applies dynamic resource establishment. The main advantage of protection is quick reaction to the failure guaranteeing small restoration time. Examples of protection methods are: 1+1 (dedicated protection), 1:1 protection, backup paths, p-cycles, [8], [21]. To protect the P2P multicasting system we propose to use an approach similar to 1:1 protection and establish two (or more) failure disjoint P2P multicast trees streaming the same content. As possible failure scenarios we want to focus on overlay link failure, uploading node failure and ISP (Internet Service Provider) link failure.

The major goal of our work is to examine how the additional constraints following from survivability requirement influence the P2P multicasting in terms of two metrics: the streaming cost and the system throughput. The first criterion can be interpreted in many ways, e.g. as network delay, bandwidth consumption, etc. The second function denotes the throughput of the content distribution session, i.e., the receiving rate of each peer. To verify our approach we apply simulations and optimization of Mixed Integer Programming (MIP) models using CPLEX solver [10].

In this work we examine the P2P multicasting on the flow level. We do not address in detail such issues as signaling and management protocols, packet level behavior of the network, join and leave procedure. However, a lot of research on these issues can be found in previous works (see section II and references therein). In spite of these simplifying assumptions, we believe that results presented in this work are universal and can be valuable for many P2P multicasting systems independent of technical details related to implementation.

The problem of resilient P2P multicasting was a topic of many papers (see next section). However, most of previous papers focus on the issues of special coding systems used in P2P multicasting and providing many disjoint parents for each peer. Moreover, the problem of path diversity is usually limited to assumption that the set of ancestors of a node in each tree should be as disjoint as possible [4], [14], [18]. Thus, overlay links and ISP links are not assured to be disjoint. To our best survey, there are not many works that thoroughly examine survivability constraints in the optimization of P2P multicasting flows with cost and throughput objectives. It should be underlined that the P2P multicasting protection can be used in parallel with other proactive and reactive methods proposed in the literature, i.e. our approach of failure-disjoint trees can be complemented with other mechanisms proposed for resilient P2P multicasting.

The main contributions of this paper are as follows: (1) Description of a protection method for P2P multicasting. (2) Simulation evaluation of how additional survivability constraints influence the streaming cost and the throughput. (3) Formulation of MIP problems modeling survivable P2P multicasting including proposals of cut inequalities. (4) Numerical experiments on MIP models showing the impact of survivability constraints on P2P multicasting in terms of the streaming cost and the system throughput.

The rest of the paper is organized in the following way. In Section II we present briefly previous research on P2P multicasting with a special focus on survivability. Section III presents the protection method for P2P multicasting including results of simulations. In Section IV we introduce MIP models of survivable P2P multicasting and show some results. Finally, last section concludes this work.

II. RELATED WORK

Authors of [18] address the resilient P2P streaming problem. They propose to apply multiple, diverse distribution trees to provide redundancy in network paths and multiple description coding (MDC). A simple tree management algorithm that provides the necessary path diversity MDC mechanisms is developed. Simulation using real date is conducted to evaluate the approach. The objective is to improve the peak signal-to-noise ratio (PSNR) parameter.

A proactive tree recovery mechanism to make the overlay multicast resilient to peer failures is proposed in [7]. To improve the system resilience, each non-leaf node (uploading node) precalculates a parent-to-be for each of its children in spite of the reactive approach in which downstream nodes try to find a new parent after a node departure. If the uploading node is broken, all its children can find their new parents immediately. Authors use extensive simulations to prove that the proactive approach can recover from node failures much faster than reactive methods, keeping the cost of recovery on reasonable level.

Authors of [5] introduce Nemo – a new P2P multicast protocol that provides high delivery ratio without sacrificing end-to-end latency or incurring additional costs. Nemo uses two basic techniques: co-leaders to minimize dependencies and triggered negative acknowledgments (NACKs) to detect lost packets. Results of simulation comparative evaluation of the protocol show that Nemo can achieve delivery ratios similar to other protocols under high failure rates, but provides a reduction of the cost in terms of duplicate packets and control-related traffic.

Liang and Nahrstedt develop in [14] DagStream – a new P2P streaming framework called. The DagStream creates of structure of peers in the form of a directed acyclic graph (DAG). To guarantee system resilience each node maintains at least k parents. The experimental part of the work contains both simulations and wide area environment and shows that DagStream protocol assures that peers can quickly self-organize into a locality aware DAG.

Wu and Li address the problem of bandwidth allocation for P2P streaming sessions [25]. Authors present a decentralized algorithm that can assign the bandwidth of each overlay link to particular streams, such that the required streaming rate is always guaranteed, without depending on any a priori knowledge of available peer upload or overlay link bandwidth. Simulation study is used to verify the proposed algorithm.

In [1] an optimal distributed rate allocation scheme to solve the rate allocation problem in overlay multirate video multicasting systems is proposed. The goal is to optimize the overall video quality perceived within the multicast group members that organize themselves into an overlay tree. Results of simulations are provided.

Zhu *et al.* introduce a model of an overlay network which includes correlated link capacities and linear capacity constraints (LCC) in order to consider hidden shared bottlenecks [27]. New metrics are defined to estimate overlay quality in terms of its accuracy and efficiency. Analysis and results of simulations indicate that LCC-overlay is accurate and hence provides much higher efficiency than the inaccurate independent overlay. For more information on multicasting refer to [2], [16].

III. PROTECTION FOR P2P MULTICASTING

A. Concept

P2P multicasting is an efficient method to provide streaming of various content over the Internet. When the transferred content is of great significance and delivery guarantees are needed, some protection methods are required. Taking the inspiration from protection methods developed for computer networks including technologies like SONET/SDH, MPLS, ATM, DWDM we propose to use failure-disjoint trees transmitting the same content simultaneously. Thus, each receiver (peer) downloads the stream using at least two failure-disjoint trees. Consequently, in the case of a failure peers affected by the failure of one of the trees can use another tree(s) to receive the required data. This procedure guarantees very low restoration time. Most of previous work on P2P multicasting resilience propose reactive methods that characterizes with much larger restoration time. Note that in the case of streaming content special coding can be used and each tree broadcasts different layer of the original stream [1], [24].

We consider three kinds of failures: overlay link failure, uploading node failure and ISP link failure. In the case of the overlay link failure, a pair of peers is disconnected. If there was a transfer on this link, some downstream nodes are affected by the failure. The overlay link failure comprises failure of both directed links. This follows from the fact that usually a network failure influences the transfer in both directions. The second failure – uploading node failure – impacts all successors of the failed peer in the tree. Therefore, we focus only on failure of nodes that have some children. Leaf node failure affects only this one node. P2P multicasting is usually used in the Internet, which consists of many ISP operators. Each peer is connected to a particular ISP. A failure of cross ISP link means, that all overlay links between peers of one ISP and peers of the second ISP are not available. We do not address the problem what can be a reason of the network failure, but this information can be found in [8], [21] and references therein.

On Fig. 1a we present a simple example to illustrate our concept. Two trees *A* and *B* are established in the overlay network connecting 8 nodes *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*. Peer *a* is the root of tree *A* and peer *b* is the root of tree *B*. Other nodes *c*, *d*, *e*, *f*, *g*, *h* are receivers of the signal. In the case of tree *A* nodes *a*, *c* and *f* are uploading (non-leaf) nodes, while the remaining nodes *d*, *e*, *g*, *h* are leafs. Nodes *a*, *c*, *d* and *e* belong to ISP 1. Nodes *b*, *f*, *g* and *h* are assigned to ISP 2. We use the term of *level* to describe the nodes [3], [14], [18]. For instance, the node *a* is on level 1 of tree *A*, nodes *c* and *f* are on level 2 of tree *A*, nodes *d*, *e*, *g* and *h* are on level 3. The tree *A* has 3 levels of nodes and the tree B is of 4 levels. Notice that the overlay link (*c*,*d*) belongs to both trees, so in the consequence of link (*c*,*d*) failure, the node *d* will be connected to none tree. Peer *c* is an uploading node in both trees. Therefore, the failure of *c* will disconnect all successors of *c*. Finally, the link between ISP 1 and ISP 2 is shared by both trees – in tree *A* link (*a*,*f*) and in tree *B* link (*b*,*c*). Fig. 1b shows survivable configuration of trees.

B. Tree Construction Strategies

In this section we consider a case when the P2P multicast trees are created in a distributed fashion, i.e. each node connects to the tree by its own. We consider several strategies used by peers in the tree construction process. The general idea follows from P2P systems. The first one is called RSL (Random Selection with Level control), the parent node selects at random its children among all nodes requesting to join the tree. The goal is to limit the number of tree levels and make the tree as short as possible [18]. The motivation is to minimize the consequences of the node failure. For each tree, we start at the root (level) and connect successors until free upload capacity of the root is available. Next we repeat the procedure for the next level of nodes. The new tree level is started if there is not enough spare capacity on the current level. Each node knows its level in each tree (it is simply the parent level plus 1) and the current level of each tree. All not-connected nodes can contact only uploading nodes located on the current tree level. Therefore – as in [18] – some centralized control is required in the system. For instance, the tree root node monitors the tree level and informs all nodes about the change of the current tree level.
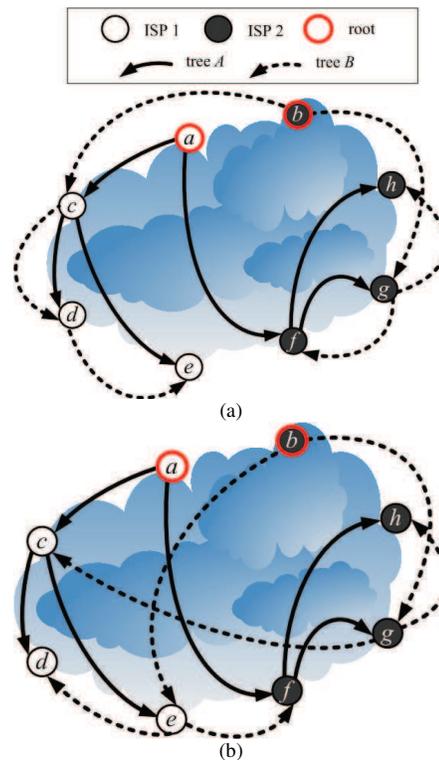

(a)

(b)

Figure 1. Examples of P2P mulitcast trees

The next strategy – CSL (Cost Selection with Level control) – uses as the objective the network cost. The parent node selects a cheapest peer (in term of the overly link cost) among all requesting peers. Analogous to RSL, the CSL tries to minimize the depth of the tree by controlling the tree level.

Based on the CSL method we develop three strategies taking into account survivability constraints according to failure scenarios presented in the previous section: CSL-LD (link disjoint), CSL-ND (node disjoint) and CSL-ID (ISP link disjoint). In each particular strategy the upstream node selects the cheapest downloading peer trying to assure that multicast tree is disjoint according to particular failure scenario. To enable this functionality, each peer must monitor and store all information on created overlay links. Note that the objective is to create trees covering all peers. Thus, if it is not possible to establish a new overlay link satisfying survivability constraints (e.g. link disjoint), the uploading peer selects a new downloader without taking into account disjoint constraints.

C. Results

In this section we present results of simulations. The simulator is relatively simple – it concentrates on the creation process of trees. However, in our opinion this approach is sufficient to examine the influence of survivability constraints on the P2P multicasting. We focus on two criteria related to P2P multicasting: cost and throughput.

In the first experiment there are randomly generated 5000 peers belonging to 10 ISPs. 10% of nodes (including roots) are connected to the Internet by Ethernet links (10Mb/s). Other nodes use asymmetric links with upload bandwidth in the range

256-1536 Kb/s and the download capacity in the range 2048-12288 Kb/s. The overlay link cost is from 3 to 100. There are two multicast trees, each tree distributes the data with the rate 256Kb/s. We consider two cases: (1) root nodes are located in separate ISPs, (2) roots are connected to the same ISP.

On Tables I and II we report the performance of tested strategies in terms of: cost (second column), number of common overlay links in both trees (third column), number of common uploading nodes in both trees (fourth column) and number of common ISP links on both trees (fifth column). The results are average – for each selection strategy 100 tests were repeated for various networks generated according to assumptions given above. We can easily notice that RSL strategy (second row) approach yields the worst results in terms of the cost. The difference between four other strategies (rows 3-6) is not large. Surprisingly, the survivability-aware strategies provide the average value of the cost slightly lower than CSL approach. This follows from the node selection strategy using the tree level control. In the case of CSL method, all uploading nodes located on the current level must saturate the uploading capacity before the next tree level is started. Consequently, some relatively "expensive" (in terms of the link cost) nodes can be connected to the tree. If additional survivability constraints are introduced, the new tree level can be started earlier, thus some expensive transfers can be avoided. Tables I and II prove that strategies CSL-LD, CSL-ND and CSL-ID significantly outperform CSL approach in terms of survivability criteria expressed as the number of common network elements. Comparing Table I against Table II do not show considerable differences between the two scenarios of root nodes location (separate ISP against common ISP).

In the next experiment we change the percentage of Ethernet nodes among all 5000 peers from 0% to 45% (Fig. 2). It can be seen that the streaming cost grows with the increase of the percentage of Ethernet peers in the network, which bring more abundant upload capacities. This again can be explained by the level-oriented selection strategies applied in our approach – uploading nodes must be fully saturated, sometimes generating relatively expensive transfers. The difference between particular selection strategies are not substantial. Fig. 3 shows the streaming cost as a function on the number of nodes. Obviously, with the increase of nodes the system cost increases. But the gap between particular curves representing selection strategies is relatively small irrespective of the number of nodes. Figures 2-3 present results of the common ISP case (roots are in the same ISP), however in the second case of separate ISPs results were similar.

The next goal of experiments was to verify how the throughput objective is influenced by survivability constraints. We tested 5 networks consisting of 5000 nodes. For each network we calculated the upper bound of the throughput using formula (16) and methodology presented in Section IV.B. Next we run simulations with this maximum value of the throughput to check if strategies CSL, CSL-LD, CSL-ND and CSL-ID can find a feasible solutions, i.e. construct two P2P multicast trees. For each network we repeat the simulations 10 times and in at least one case a feasible solution was found.

TABLE I.   SELECTION STRATEGIES, ROOTS IN SEPARATE ISPs

Strategy	Cost	Common links	Common node	Common ISP links
RSL	530511	1.4	356.2	45.0
CSL	221415	723.0	652.8	40.6
CSL-LD	219318	0.0	607.4	41.2
CSL-ND	214344	1.3	0.0	39.8
CSL-ID	220787	3.8	286.3	12.2

TABLE II.   SELECTION STRATEGIES, ROOTS IN THE SAME ISP

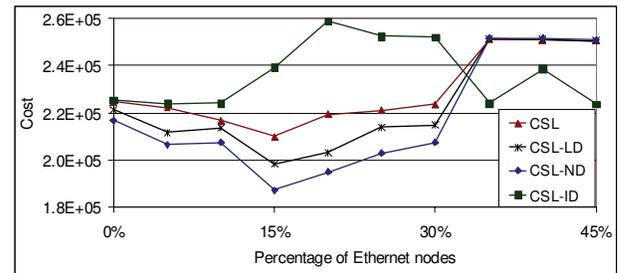| Strategy	Cost	Common links	Common node	Common ISP links
RSL	530487	1.2	353.5	45.0
CSL	222106	800.8	629.5	41.6
CSL-LD	220706	0.0	593.6	41.6
CSL-ND	215952	1.1	0.0	40.8
CSL-ID	216934	0.2	115.6	3.5

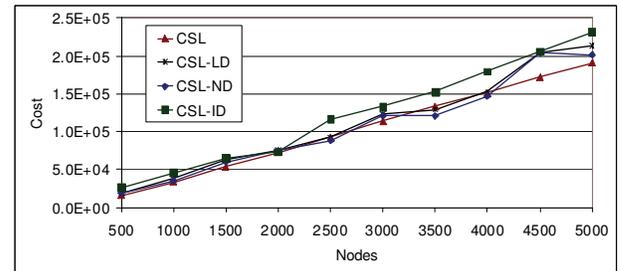Figure 2. Streaming cost as a function of Ethernet nodes percentage

Figure 3. Streaming cost as a function of the number of nodes

IV. INTEGER PROGRAMMING MODELS

A. Streaming Cost

In this section we present a MIP model of the survivable P2P multicasting in overlay networks. The offline optimization can be used to find lower bound of other solution approaches. Moreover, some P2P multicasting systems are static (e.g. data distribution in CDN), and results of offline optimization can be used in such systems to improve the system performance. The first considered problem uses as the objective streaming cost as proposed in [24]. Assumptions of the model follow from real systems and previous works. We consider only node capacity constraint nodes, which according to [27] are typically sufficient in overlay networks, since in the concept of overlay networks usually the underlay core network is overprovisioned and the only bottlenecks are access links [1], [26].

A common approach in the literature is to use Steiner or spanning trees to model multicasting. However, the canonical

Integer Programming (IP) formulation of trees includes a large number of cut constraints that guarantees that the tree is connected [11]. Authors of [13], [23], [24] use a *conceptual flows* formulation to model multicast trees. However, we apply another formulation – *level flows* – first time proposed in our work [22]. The main advantage of level flows formulation over conceptual flows formulation is that we can limit the number of tree levels and make the tree as short as possible.

Now we introduce the notation indispensable to present the level flows formulation. Let indices $v,w = 1,2,\dots,V$ denote peers – nodes of the overlay network that are participants of the P2P multicasting streaming. Each node has node capacity constraint – constants u_v and d_v denote the available upload and download capacity of node v, respectively. There are T streaming trees indexed $t = 1,2,\dots,T$. If a peer v is the root of the tree (i.e. v is a streaming node), then $r_{vt} = 1$, otherwise $r_{vt} = 0$. For simplicity of notation we assume that peer $w = (V - T + 1)$ is the root of tree $t = 1$, peer $w = (V - T + 2)$ is the root of tree $t = 2$, etc. All peers that want to receive the stream and that are not roots in any tree are denoted using index $k = 1,2,\dots,K$. Note that the number of receiving nodes is $K = (V - T)$.

We assume that the root of the tree is located on level 1. All children of the root (peers that have a direct link from the root) are located on level 2, etc (see Fig. 1 and Section II.A). The proposed notation enables us to set the value of L as a limit on the maximal depth of the tree.

A binary variable x_{vwtl} is used to model the multicast tree. We assume that $x_{vwtl} = 1$ if in the multicast tree t there is a link from node v to node w (no other peer nodes in between) and the upstream node v is located on level l of tree t; 0 otherwise (binary). Additional binary variable y_{vwt} is 1 if overlay link (w,v) (no other peer nodes in between) is used in multicast tree t; 0 otherwise. As in [23] and [24], we want to minimize the total cost of streaming, while guaranteeing the constant streaming rate for all trees. A constant c_{wv} is applied to denote the streaming cost on an overlay link (w,v). The objective is to minimize the overall streaming costs (including all trees) $\sum_w\sum_v\sum_t y_{wvt}c_{wv}$. Since the streaming rate is the same in all trees, we do not include the streaming rate in the objective function.

To formulate the problem we use the notation as in [20].

Streaming Cost Minimization (SCM)

indices

$v,w = 1,2,\dots,V$	overlay nodes (peers)
$k = 1,2,\dots,K$	receiving nodes (peers)
$t = 1,2,\dots,T$	streaming tree index
$l = 1,2,\dots,L$	levels of the tree

constants

d_v download capacity of node v (kb/s)

u_v upload capacity of node v (kb/s)

r_{vt} = 1 if node v is the root (streaming node) of tree t; 0 otherwise

q the streaming rate (kb/s)

c_{wv} streaming cost on overlay link from node w to node v

M large number

variables

x_{wvtl} streaming rate on an overlay link (w,v) (no other peer nodes in between) in multicast tree t and w is located on level l of tree t; (continuous, non-negative)

y_{wvt} = 1 if link from node w to node v (no other peer nodes in between) is in multicast tree t; 0 otherwise; (binary)

objective

$$\text{minimize} \quad \sum_w\sum_v\sum_t y_{wvt}c_{wv} \tag{1}$$

constraints

$$\sum_w\sum_t\sum_l x_{wvtl} = 0 \quad v = 1,2,\dots,V \quad t = 1,2,\dots,T \quad r_{vt} = 1 \tag{2}$$

$$\sum_w\sum_l x_{wktl} = q \quad k = 1,2,\dots,K \quad t = 1,2,\dots,T \tag{3}$$

$$\sum_v x_{wvt1} \le u_w r_{wt} \quad w = 1,2,\dots,V \quad t = 1,2,\dots,T \tag{4}$$

$$x_{wvt(l+1)} \le \sum_b x_{bwtl} \quad v = 1,2,\dots,V \quad w = 1,2,\dots,v \\ t = 1,2,\dots,T \quad l = 1,2,\dots,L-1 \tag{5}$$

$$\sum_w\sum_t\sum_l x_{wvtl} \le d_v \quad v = 1,2,\dots,V \tag{6}$$

$$\sum_v\sum_t\sum_l x_{wvtl} \le u_w \quad w = 1,2,\dots,V \tag{7}$$

$$\sum_v\sum_t\sum_l x_{vvtl} = 0 \tag{8}$$

$$\sum_l x_{wvtl} \le qy_{wvt} \quad v = 1,2,\dots,V \quad w = 1,2,\dots,v \\ t = 1,2,\dots,T \tag{9}$$

$$y_{wvt} \le \sum_l x_{wvtl} \quad v = 1,2,\dots,V \quad w = 1,2,\dots,v \\ t = 1,2,\dots,T \tag{10}$$

$$\sum_w y_{wkt} = 1 \quad k = 1,2,\dots,K \quad t = 1,2,\dots,T \tag{11}$$

The objective function (1) is the streaming cost. Constraint (2) forces the download flow to be zero for the root node of each tree. To meet the requirement that each receiving node $k = 1,2,\dots,K$ must be connected to each streaming tree we add condition (3). Constraint (4) assures that node w can be a parent of the first level link in tree t, only if it is the root node ($r_{wt} = 1$). Condition (5) guarantees that each node w cannot upload to any other peer v on level $(l + 1)$ more than it downloads on level l. (6) and (7) are download and upload capacity constraints, respectively. Constraint (8) is in the model to assure that the node internal flow is zero. To bind variable x_{wvtl} and y_{wvt} we introduce constraints (9) and (10). Finally, condition (11) guarantees that each receiving peer has exactly one parent node.

Notice that in the model (1)-(11) multicast flow cannot be split, i.e. each receiving peer downloads the stream in each tree

from exactly one parent node. Consequently, variables x_{wvtl} are either 0 or q (streaming rate). The formulation of conceptual flows proposed in [13], [23], [24] allows to split flows. It should be underlined that in the latter case the management of the streaming system is more complex.

Since the SCM problem is a complex and NP-complete problem (it can be reduced to a knapsack problem) the only way to find optimal solution is branch-and-cut (B&C) method. To facilitate the high computation complexity caused by the NP-completeness in solving the MIP model (1)-(11), we introduce additional cut inequalities that can be applied in construction of branch-and-cut algorithm. Cut inequalities are introduced into the optimization problem, enabling the branching phase to utilize this information in calculation of more effective bounds. We consider the cut-and-branch variant of the B&C algorithm, in which cut inequalities are added to the root node of the solution tree. It means that all generated cuts are valid throughout the whole B&C tree [17].

We propose a cut inequality using the mixed-integer rounding (MIR) approach [9], [15]. The major idea of our MIR procedure is to find the closest lower bound of download and upload capacity for each node, what is the multiple of q. For instance, if $u_v = 384$ and $q = 256$, then we can set $u_v = 256$.

The next cut proposal follows directly from the fact that the number of links in each tree must be $K = (V - T)$, i.e.

$$\sum_w \sum_k y_{wkt} = (V - T) \quad t = 1,2,\ldots,T \qquad (12)$$

On Fig. 4 we present the effectiveness of the proposed MIR cut and (12) in terms of the execution time for 10-node network. We apply the MIP optimizer included in CPLEX [10]. The upload and download capacity is a multiple of 256 Kb/s, the streaming rate is set to 200 Kb/s.

B. Throughput

Now we present the MIP model with the throughput objective, i.e. we want to maximize the aggregate receiving rate at each participating peer [24]. The assumptions and notation are as in previous section.

Throughput Maximization (TM)

variables (additional)

q the streaming rate (kb/s); (continuous, non-negative)

objective

$$\text{maximize} \quad q \qquad (13)$$

constraints (2)-(11)

Notice that in the (TM) problem the left-hand side constraint (3) is not a constant, but a variable.

As in the case of the SCM problem we develop additional cut inequalities to reduce the execution time of B&C algorithm. First we want to find the upper bound of q denotes as q^{UB}. For easy of notation let U denote the minimum upload capacity of the root node for trees $t = 1,2,\ldots T$:
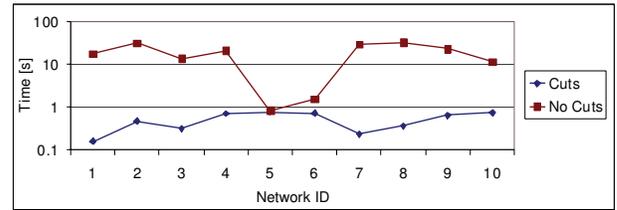

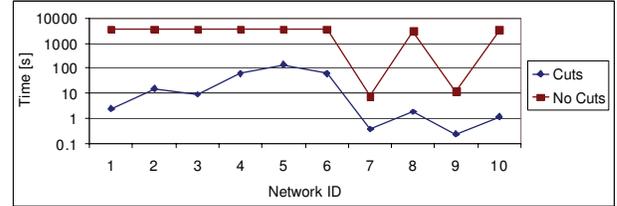Figure 4. Efficiency of cut inequalities for cost objective

Figure 5. Efficiency of cut inequalities for throughput objective

$$U = \min u_v \quad v = 1,2,\ldots,V \quad t = 1,2,\ldots T \quad r_{vt} = 1 \qquad (14)$$

If we assume that the only bandwidth limitations are the node upstream capacities, the maximum achievable streaming rate can be expressed as

$$q^{UB} = \min\left\{U, \frac{\sum_v u_v}{(V - T)T}\right\} \qquad (15)$$

(15) is a small modification of the formula proposed in [12] and [19]. Notice that the upload capacity must be enough to stream the content to $(V - T)$ receiving peers in each t of T trees.

In our formulation we assume that the flow cannot be split, therefore we can include the procedure analogous to MIR cuts. Notice that if the streaming rate is q, then the maximum number of tree links leaving node v is $\left\lfloor \dfrac{u_v}{q} \right\rfloor$. Since we must establish T trees and each tree consists of $(V - T)$ links (every peer expect root must have one father), the number of all links must be $T(V - T)$. Thus we can write the following constraint

$$\sum_v \left\lfloor \frac{u_v}{q} \right\rfloor \geq T(V - T) \qquad (16)$$

Using formula (16) we can easily find the upper bound of the system throughput q^{UB}, which is more restrictive than (15).

To find the lower bound of q denoted as q^{LB} we solve the problem using heuristic algorithm, e.g. we stop the CPLEX solver after a limited execution time.

The following inequality cuts can be added to the TM formulation

$$q \leq q^{UB} \qquad (17)$$

$$\sum_v \sum_t y_{wvt} \leq \left\lfloor \frac{u_w}{q^{LB}} \right\rfloor \quad w = 1,2,...,V \qquad (18)$$

On Fig. 5 we present the effectiveness of the proposed MIR cut in terms of the execution time for 10-node network. The execution time of CPLEX is limited to 1 hour. In 6 of 10 cases, the CPLEX without cuts could not find an optimal solution.

C. Survivability Constraints

In this section we formulate additional constraints that are related to survivability of P2P multicasting. First we address the problem of overlay link failure. To assure protection against this failure, multicasting trees must be link disjoint.

Link Disjoint (LD)

constraints (additional)

$$\sum_t (y_{wvt} + y_{vwt}) \leq 1 \quad v = 1,2,...,V \quad w = 1,2,...,V \quad v < w \quad (19)$$

Notice that in our formulation in the case of overlay link failure both directed links $((w,v)$ and $(v,w))$ are broken. This follows from the fact that usually a network failure influences the transfer in both directions.

Next we formulate additional constraints for the uploading node failure. We introduce to the problem additional binary variable denoting if a particular node is uploading in a given tree.

Node Disjoint (ND)

variables (additional)

y_{vt} $\quad$ =1 if node v is uploading in multicast tree t; 0 otherwise (binary)

constraints (additional)

$$\sum_v y_{wvt} \leq My_{wt} \quad w = 1,2,...,V \quad t = 1,2,...,T \qquad (20)$$

$$y_{wt} \leq \sum_v y_{wvt} \quad w = 1,2,...,V \quad t = 1,2,...,T \qquad (21)$$

$$\sum_t y_{vt} \leq 1 \quad v = 1,2,...,V \qquad (22)$$

Finally we present a model related to ISP link failure.

ISP Link Disjoint (ID)

indices (additional)

$p,m = 1,2,...,P$ $\qquad$ ISPs

constants (additional)

$\alpha(v,p)$ $\qquad$ = 1 if node v belongs to ISP p; 0 otherwise

variables (additional)

z_{pmt} $\quad$ = 1 if in multicast tree t there is at least one link from a node located in ISP p to a node located in ISP m or in opposite direction; 0 otherwise (binary)

constraints (additional)

$$\sum_{w:\alpha(w,p)=1}\sum_{v:\alpha(v,m)=1} (y_{wvt} + y_{vwt}) \leq Mz_{pmt} \quad p = 1,2,...,P$$
$$m = 1,2,...,P \quad p \neq m \quad t = 1,2,...,T \qquad (23)$$

$$z_{pmt} \leq \sum_{w:\alpha(w,p)=1}\sum_{v:\alpha(v,m)=1} (y_{wvt} + y_{vwt}) \quad p = 1,2,...,P$$
$$m = 1,2,...,P \quad p \neq m \quad t = 1,2,...,T \qquad (24)$$

$$\sum_t z_{pmt} \leq 1 \quad p = 1,2,...,P \quad m = 1,2,...,P \qquad (25)$$

Note that the ISP link failure includes the failure of both directed ISP links, i.e. (p,m) and (m,p).

TABLE III. MIP MODELS, ROOTS LOCATED IN SEPARATE ISPS

Model	Cost	Common links	Common node	Common ISP links
SCM	705	8.7	2.7	2.8
SCM-LD	794	0.0	1.5	2.8
SCM-ND	780	4.5	0.0	2.6
SCM-ID	902	5.8	1.7	0.0

TABLE IV. MIP MODELS, ROOTS LOCATED IN THE SAME ISP

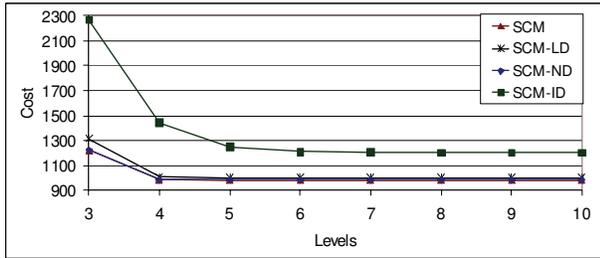| Model	Cost	Common links	Common node	Common ISP links
SCM	831	6.5	1.5	3.6
SCM-LD	888	0.0	0.7	3.3
SCM-ND	862	4.1	0.0	3.5
SCM-ID	1127	4.8	1.8	0.0

Figure 6. Streaming cost as a function of level number

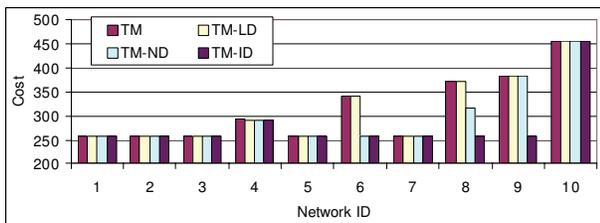Figure 7. Streaming cost as a function of streaming rate

Figure 8. Throughput for various MIP models

2009 7th International Workshop on the Design of Reliable Communication Networks

D. Results

To solve problems presented in previous section we apply the MIP solver included in CPLEX 11.0 library [10]. Since we want to obtain optimal results, the size of tested networks was limited to 20 nodes what guaranteed that the execution time was from seconds to hours. Nodes are located in 5 ISPs, the link costs are in the range 3-109, the number of tree levels is 4. 10 random networks were generated. Nodes (including roots) are either symmetric (2048 Kb/s or 4096 Kb/s) or asymmetric (2048/256 Kb/s or 4096/512 Kb/s or 6144/768 Kb/s). The streaming rate was set to 256 Kb/s. In Table III and IV we present the average results obtained for cost models in terms of: cost, number of common overlay links in both trees, number of common uploading nodes in both trees and number of common ISP links on both trees. The gap between model SCM and models SCM-LD and SCM-ND is less than 10%. Only in the case of SCM-ID model the corresponding gap is larger. Fig. 6 shows the streaming cost as a function of the number of levels. We can watch that the cost drops with the increase of levels, however starting from 5 levels the cost converges to stable value for each curve. On Fig. 7 we report the cost as a function of streaming rate. Note that in the case of the SCM-ID model for streaming rate larger than 256Kb/s no solution was found. Figures 6-7 present results of the common ISP case (roots are in the same ISP), however in the second case of separate ISPs results were comparable. Fig. 8 present the system throughput. We can easily notice that in 6 of 10 networks all models provide the same throughput. The TM-LD model in all 10 cases yields the same solution as model TM. The average gap to TM model for TM-ND and TM-ID models is 5.15% and 12.89%, respectively.

V. CONCLUDING REMARKS

The paper has addressed the problem how the additional survivability requirements impact the P2P multicasting system in terms of two criteria: cost and throughput. For this purpose, we have proposed a simulation framework and formulate the a MIP optimization problem. Based on results of simulations and results yielded by MIP solver, we can conclude that the requirement to construct failure-disjoint trees do not influence significantly the objective functions of cost and throughput. Consequently, the P2P multicasting system can be enhanced with additional survivability criteria without substantial degradation of the system in terms of the streaming cost and the system throughput. In the ongoing work, we plan to develop efficient heuristic algorithms solving MIP formulations for larger networks and to develop a highly reliable P2P multicasting framework using dual-homing for each participant.

REFERENCES

[1] B. Akbari, H. Rabiee and M. Ghanbari, „An optimal discrete rate allocation for overlay video multicasting," Computer Communications, no. 31, 2008, pp. 551-562.

[2] A. Benslimane (ed.), Multimedia Multicast on the Internet, ISTE, 2007.

[3] S. Banerjee, L. Seungjoon, B. Bhattacharjee, A. Srinivasan, A., Resilient multicast using overlays, IEEE/ACM Transactions on Networking, vol. 14, no. 2, 2006, pp. 237-248.

[4] S. Birrer, F. Bustamante, "Resilient peer-to-peer multicast without the cost," Multimedia Computing and Networking 2005. Edited by Chandra, Surendar; Venkatasubramanian, Nalini. Proceedings of the SPIE, Volume 5680, 2004, pp. 113-120.

[5] S. Birrer, D. Lu, F. Bustamante, Y. Qiao and P. Dinda, "FatNemo: Building a Resilient Multi-Source Multicast Fat-Tree," Lecture Notes in Computer Science, vol. 3293, 2004, pp. 182-196.

[6] M. Cha, P. Rodriguez, S. Moon, and J. Crowcroft, "On next-generation telco-managed P2P TV architectures," in Proc. of the 7th International Workshop on Peer-to-Peer Systems (IPTPS '08), 2008.

[7] Z. Fei, M. Yang, "A Proactive Tree Recovery Mechanism for Resilient Overlay Multicast," IEEE/ACM Transactions on Networking, vol. 15, no. 1, 2007, pp. 173-186.

[8] W. Grover, Mesh-based Survivable Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking, Prentice Hall PTR, Upper Saddle River, New Jersey, 2004.

[9] O. Gunluk, "Branch-and-Cut Algorithm for Capacitated Network Design Problems," Math. Programming, 86 (1999), pp. 17-39.

[10] ILOG CPLEX 11.0 User's Manual, France, 2007.

[11] T. Koch and A. Martin, "Solving Steiner tree problems in graphs to optimality," Networks, vol.32, no. 3, 1998, pp. 207-232.

[12] R. Kumar, Y. Lin, and K. Ross, "Stochastic fluid theory for P2P streaming systems," In Proc. of the 26th IEEE International Conference on Computer Communications INFOCOM, 2007, pp. 919-927.

[13] Z. Li, B. Li, D. Jiang, and L. C. Lau, "On Achieving Optimal Throughput with Network Coding," in Proc. of IEEE INFOCOM 2005, vol. 3, 2005, pp. 2184-2194.

[14] J. Liang and K. Nahrstedt, "DagStream: Locality Aware and Failure Resilient Peer-to-Peer Streaming," in Proc. of the 13th SPIE/ACM Multimedia Computing and Networking Conference (MMCN'06), 2006.

[15] H. Marchand, L. Wolsey, "Aggregation and mixed integer rounding to solve MIPs," Operations Research, 49 (2001), pp. 363-371.

[16] D. Minoli, IP Mulitcast with Applications to IPTV and Mobile DVB-H, John Wiley & Sons, 2008.

[17] J. Mitchell, "Branch-and-cut methods for combinatorial optimization problems," in the Handbook of Applied Optimization, Oxford Univ.ersity Press, 2002.

[18] V. Padmanabhan, H. Wang, P. Chou, "Resilient peer-to-peer streaming," in Proc. of the 11th IEEE International Conference on Network Protocols, 2003, pp. 16-27.

[19] F. Picconi, L. Massoulie, "Is There a Future for Mesh-Based live Video Streaming?," In Proc. of the Eighth International Conference on Peer-to-Peer Computing , 2008, pp. 289-298.

[20] M. Pióro and D. Medhi, Routing, Flow, and Capacity Design in Communication and Computer Networks, Morgan Kaufman 2004.

[21] J. Vasseur, M. Pickavet, P. Demeester, Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP and MPLS, Morgan Kaufmann, San Francisco, 2004.

[22] K. Walkowiak, "Network Design Problem for P2P Multicasting," in Proc of the Int. Network Optimization Conference INOC, 2009.

[23] C. Wu and B. Li, "Optimal peer selection for minimum-delay peer-to-peer streaming with rateless codes," in Proc. of the ACM workshop on Advances in peer-to-peer multimedia streaming, 2005, pp. 69-78.

[24] C. Wu and B. Li, "Optimal Rate Allocation in Overlay Content Distribution," in Proc. of the 6th Networking Conf., 2007, pp. 678-690.

[25] C. Wu and B. Li, "On Meeting P2P Streaming Bandwidth Demand with Limited Supplies," in Proc. of the Fifteenth Annual SPIE/ACM International Conf. on Multimedia Computing and Networking 2008.

[26] G. Wu and C. Tzi-cker, "Peer to Peer File Download and Streaming", RPE report, TR-185, 2005.

[27] Y. Zhu and B. Li, "Overlay Networks with Linear Capacity Constraints, IEEE Transactions on Parallel and Distributed Systems", vol. 19, no. 2, 2008, pp. 159-173.