Poznańskie Warsztaty Telekomunikacyjne
Poznań 6 - 7 grudnia 2007

POZNAN UNIVERSITY OF TECHNOLOGY ACADEMIC JOURNALS

Tomasz BARTCZAK*
Tomasz SZEWCZYK**
Piotr ZWIERZYKOWSKI***

ASSURING QUALITY OF SERVICE FOR IP MULTICAST TRANSMISSION IN ISP NETWORKS

The paper, discusses current state of art of multicast transmission in packet networks with assured QoS. Multicast transmission presents one of the most important challenges in multi-service backbones. This work proposes combining heuristic algorithms developed for building multicast distribution trees together with signaling protocols used in MPLS networks. The aim of the article is to show that these two mechanisms should cooperate to provide high quality multicast transmission in networks with complex sources and receivers distribution structure.

Keywords: multicast transmission, MPLS, heuristic algorithm.

1. INTRODUCTION

More and more applications in the Internet that reflect new tendencies in telecommunications make use of the one-to-many transmission model. This paradigm can be implemented in many ways, among which multicast transmission is particularly notable as it allows to limit significantly network traffic. Packets sent in this mode reach user groups without a necessity to replicate the data stream by the source.

It is easily noticeable that the application of multicast transmission effects in a far more efficient use of network resources. It is particularly significant if we take into account the current predictions indicating that as much as 90 % of packet traffic can be linked to multicast transmissions [1].

Modern packet networks offer a wide range of services that make use of multicast transmission [11]. Services that require particularly defined QoS parameters are particularly common. In the case of Internet TV and its ever-growing popularity, the QoS parameters define the maximum bandwidth, maximum delay, maximum level of lost packets, etc. Delay is the constraint that is most frequently taken into

* Nokia Siemens Networks
** Poznańskie Centrum Superkomputerowo-Sieciowe
*** Politechnika Poznańska

consideration because of its great influence upon the user's reception of the transmission quality. This means that the algorithms that construct the multicast transmission tree should be aimed at constructing optimal multicast transmission trees from the viewpoint of the use of network resources and, simultaneously, take into consideration metrics defining quality requirements of the service.

The essential problem in data distribution is the optimal selection of paths. In the case of multicast transmission, Karp has proved that a construction of optimal multicast transmission trees is a NP-complete problem [2]. This means that algorithms determining the optimal multicast transmission tree can be applied only in very small networks. Therefore, many approximate solutions that are characterized by much reduced computational complexity have been devised [13] [14] [15].

Heuristic algorithms view the problem of building multicast transmission trees as a mathematical problem and, as a rule, do not take into consideration practical aspects related to their potential use in routing protocols for packet networks [13] [15]. In all works that are devoted to possible implementations of heuristics in existing packet networks that are known to the authors such considerations, however, have not been included. Thus, a presentation of technical possibilities of the implementations of heuristics in multicast transmission seemed to the authors to be a laudable task.

The article is divided into 6 chapters. Chapter 2 presents formally the problem of building multicast transmission trees. Chapter 3 includes exemplary heuristics resolving the problem of building multicast transmission trees. The following chapter discusses problems and obstacles related to practical implementations of heuristics. Chapter 5 presents the proposed solution. The last chapter provides a conclusion following a discussion on the presented problems and indicates the directions of further research.

2. THE PROBLEM OF OPTIMIZATION OF MULTICAST TRANSMISSION TREES

The problem of building optimal multicast transmission trees can be formally presented in the following way. Let us assume that we have a transmission network denoted by the graph G = (E,V), where E is a set of edges and V is a set of nodes. To determine the quality of the set up connections of each of the edges l from the set E, the cost $c(l)$ must be associated. This cost can correspond, for example, to the amount of the fee for the use of the link or to other preferences of the operator. Additionally, the node s which is the source of multicast transmission and the set of receivers M such that $M \in V$ and $s \notin M$ are identified in the graph.

With the above assumptions, the multicast transmission tree is defined as a tree rooted in node s that connects all nodes from the set M and its cost satisfies the following condition:

$$C(MT) = \min \sum_{l \in MT} c(l). \tag{1}$$

Meeting the additional conditions expressed by metrics, for instance in the case of the maximum delay, the constraint is expressed by the following formula:

$$\forall_{p \in MT} \sum_{e \in p} D(e) \leq \Delta, \tag{2}$$

where *D(e)* is the delay of the edge *e.*

3. EXEMPLARY ALGORITHMS OF MULTICAST TRANSMISSION TREES BUILDING

There are many heuristics solving the problem of building multicast transmission trees with constraints. In the authors' opinion the most representative algorithms in question: KPP, CSPT, CCPT, DCMA, XCG, MMR. The present article, however, does not aim at a complex presentation of heuristics for multicast transmission trees, especially that there are many works devoted to the problem [13] [14] [15]. Nevertheless, to better illustrate the problems to be addressed and solved in order to make the application of the above heuristics possible in practice, one of them will be discussed here.

The KPP algorithm [10] has been chosen as an example of heuristics. This particular algorithm has been indicated because of its efficiency and its wide use. Moreover, the algorithm fully illustrates the problems to be addressed and solved in order to make the application of heuristic techniques possible in the existing network protocols. Also, the MLRA algorithm is very effective for large networks and, just as the KPP algorithm, is characterized by low computational complexity [12].

The KPP algorithm is used to construct multicast transmission trees with one constraint which is the maximum admissible delay. According to the constraint pattern, none of the paths connecting individual receivers with the source should exceed the assumed delay. Figure 1 shows the operation of the KPP algorithm for an exemplary network with the maximum delay equal to 10.

The first step in the KPP algorithm determines a set of the cheapest paths for each of the receivers that connects them to the source and to other multicast transmission receivers with the total delay values from 0 to Δ. For example, if Δ

is 10, then the cheapest paths with delays of respectively 1, 2,…,10 are determined for each of the receivers.

The second step determines the complete graph that connects the receivers and the source s. The graph is composed of the cheapest paths determined in the previous step (Fig. 1b)

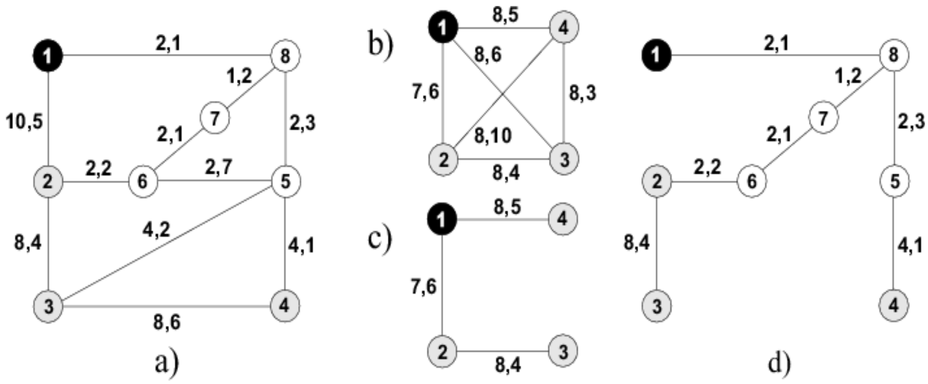

Fig. 1 An example of the operation of KPP algorithm

In the next step of the algorithm, paths that connect the source with particular receivers are determined. These paths are constructed according to (3) that builds multicast transmission for the complete graph. To illustrate the above, this tree is presented in Fig. 1c.

$$\begin{cases} C(v, w) & \text{if} \quad P(v) + D(v, w) < \Delta \\ \infty & \text{dla pozostałych} \end{cases} \qquad \textbf{(3)}$$

In the last stage of the KPP algorithm, the multicast transmission for the complete graph is mapped back onto the original network. If there are any loops, they are removed with the help of Dijkstra algorithm that uses delay as a criterion. Fig. 1d shows the multicast transmission tree determined by the KPP algorithm for the case under scrutiny.

4. TECHNICAL ASPECTS OF ROUTING PROTOCOLS DEPLOYMENT

Nowadays, the most commonly used multicast transmission protocols are PIM SM [17] and PIM DM [16]. The multicast transmission tree determined by these protocols consists of the shortest paths (in terms of their cost) connecting the source with individual receivers. This approach, as compared with the KPP heuristics discussed above, is though characterized by two essential drawbacks: In building multicast transmission trees PIM protocols do not take into consideration constraints and do not tend to build optimal multicast trees that make it possible to use network resources effectively and in the most favorable way.

PIM-based protocols represent a very simple approach to building multicast transmission trees. Routers that are involved in multicast transmission do not know the network structure, distribution of receivers and the source of transmission. This knowledge, however, is required in the case of the discussed heuristics because only the overall or global picture of the network makes it possible to optimize the process of building multicast transmission trees.

In the case of PIM SM and PIM DM protocols, particular elements of the multicast transmission tree are determined independently by each of the routers that are involved in multicast traffic. With the application of heuristics the calculation of the multicast transmission tree must be centralized. Firstly, heuristics require large computational resources and performing them for each of the trees in all of the routers is not desirable from the point of view of the load of network devices. Secondly, making calculations centralized in one point of the network limits the excessive traffic as the information on the receivers involved in the transmission is forwarded only to the central point.

The application of heuristics makes it possible to use network devices in the most favorable way but, first of all, to construct multicast transmission trees with QoS requirements taken into account. In order to translate a tree determined by a heuristics into a real network, it is important to ensure traffic to follow the planned paths and to secure resource reservation indispensable to carry on with the transmission.

Another important issue is to provide routers with the necessary information on the current use of the resources. If, for example, we have a link with the flow capability of 10 Mb/s. and to build a multicast transmission tree A, 8 Mb/s will be reserved, then it cannot be used to build another multicast transmission tree B that requires 6Mb/s.

5. IMPLEMENTATION OF MULTICAST TRAFFIC DISTRIBUTION WITH PREDEFINED QOS PARAMETERS

As noticed in the previous chapter, heuristic algorithms allow us to evaluate suboptimal multicast distribution tree on single network node. To build distribution tree fulfilling requested QoS parameters, it is required to provide to this node the knowledge about the entire network topology, bandwidth reservation status on all links and location of nodes with directly connected receivers.

The OSPF protocol with Traffic Engineering extension [5] provides knowledge about network topology to all nodes, builds a traffic engineering database, and thereby reports on the reservation state of links. Having this knowledge, a node A can compute a path to a destination node B. This path may be subjected to various constraints on the attributes of the links and nodes that the path traverses. For instance, only links that have required amount of unreserved bandwidth can be subjected to path computation process.

Since IP routing is performed in hop-by-hop manner, where the destination interface is selected basing on the state of a local route table of the node on which the packet has been received, to force a selection of pre-computed path, the data packet must be transmitted through the network in a slightly different way. Before sending the data packet to appropriate neighbor, it is mapped to a MPLS tunnel which is established according to the results of the path computation process. The mechanisms for MPLS tunnel establishment are provided by RSVP-TE protocol [6] and its modifications [7].

In the case of point-to-multipoint multicast transmission (P2MP), a label switched path (LSP) must be established. This P2MP LSP should allow to deliver the data from the source node to a group of destination nodes. The extension to RSVP-TE protocol supporting P2MP LSP is described in [8]. This proposed standard introduces the functionality of building distribution trees comprised of multiple source-to-leaf (S2L) sub-LSPs[1] to MPLS protocol stack. The advantage of this solution is that one *Path* message may signal one or multiple S2L sub-LSPs for a single P2MP LSP.

Let us take a look at RSVP-TE signaled P2MP LSP set up in an MPLS network. Such LSP must be explicitly defined from the root of the tree. Introduced in [6] *Path* message contains a *Session* object with a tree identifier for each P2MP LSP and explicit routes for all branches of the tree. The structure of the P2MP *Session* object is identical to those defined in [6] for traditional P2P LSP, except that the Tunnel Endpoint Address field is replaced by the P2MP ID field. The tree specification is based on a set of secondary explicit route objects (SEROs), describing branches stemming from primary path. Primary path is simply first sub-LSP and is defined in explicit route object (ERO). ERO is represented by [R_1, R_2, …,R_n], where R_n is n-th router in the path from a source node to the leaf. Since the tree is built with a set of EROs, in SERO it is not required to include the whole path from the root to the target leaf. It is important to note that after creating P2MP LSP, new branches can be added or withdrawn from the tree.

[1] The general idea of P2MP LSP is that packets transmitted through "root" LSP from the source, are replicated on nodes to which branches (sub-LSPs) are connected to. The sub-LSP can have additional branches to, and again packets are replicated only on nodes to which subsequent sub-LSPs are connected.

This brief overview of P2MP LSP shows that it can be used to implement the distribution tree computed by heuristic algorithm in the MPLS network. Moreover, such a tree is not a static entity, but can adopt changes (e.g. Fast ReRoute) to the network topology by using existing resiliency procedures, or change its structure depending on the distribution of the receivers. The most obvious application of P2MP LSPs is distribution of audio and video streams. Usually in this case, the placement the source (root of the tree) is not changing, so the tree structure will be computed on a single node. One of serious disadvantages of this solution is the scalability problem when more than one source and a multicast group appears in the network. In this case, for every (S,G) pair, the separate tree must be computed and signaled. It seems, at this moment, that heuristics can also be used also for binding some groups together to limit signaling states in the network and minimize wasted bandwidth.

6. PIM PROTOCOL MODIFICATIONS

Before the distribution tree can be evaluated, the source node has to know the location of the receivers. Depending on the application of multicast transmission in the network this goal can be achieved in may ways. In this paper we focused on multicast transmission in “public” IP networks with MPLS technology. The interesting way to solve this problem was proposed in already expired IETF internet draft [9]. It seems that currently the device vendors are more interested in multicast transmission in VPN networks than in the public Internet. That is why this solution was no longer developed. However, it provides good environment for research related to combining heuristic protocols together with multicast signaling protocols.

The [9] draft presents procedures for IP multicast over a MPLS TE core. It is focused on the case where MPLS TE is used in the network core and the edge routers are participating in multicast routing with other routers. The described solution is based on the following topics:

- Exchanging routing information between the edge routers,
- P2MP LSP endpoints discovery by root of the tree,
- Association of the appropriate multicast routing table entry and forwarding state with P2MP LSP.

The initial assumption for this protocol modification is the knowledge about source of multicast transmission. Simply, the end host needs to use IGMPv3, and PIM neighbor has to send source specific *Join* messages. Let us consider the network shown in Fig. 2. Assume there is existing P2MP LSP consisting of two branches: primary [SPE5, RPE4, RPE7] and secondary [SPE5, PE6, RPE0]. It means that multicast traffic sent from a source connected to SPE5 will be delivered to receivers connected to RPE0 , RPE4 and RPE7. RPE1 receives PIM (S1,G1)

Join (or IGMPv3 report). The [9] draft authors proposed to resolve address of the router connected to S1, by resolving it onto next-hop address that is advertising S1. This problem can also be solved by unicasting modified PIM message towards the S1. If SPE5 receives such a message and finds S1 at directly connected networks, then it reads the message. After receiving PIM *Join* message from remote neighbor, SPE5 assumes that RPE1 can be treated as P2MP LSP leaf. Using procedure described in [8], SPE5 adds subsequent branch to P2MP LSP. Next, it maps (S1,G1) received from RPE1 to P2MP LSP. The selection of P2MP LSP and mapping the multicast traffic to it is local to SPE5. Many schemes may be considered at this point and will be subject of further studies and research. Let us assume that each (S,G) entry is mapped on a single P2MP LSP.

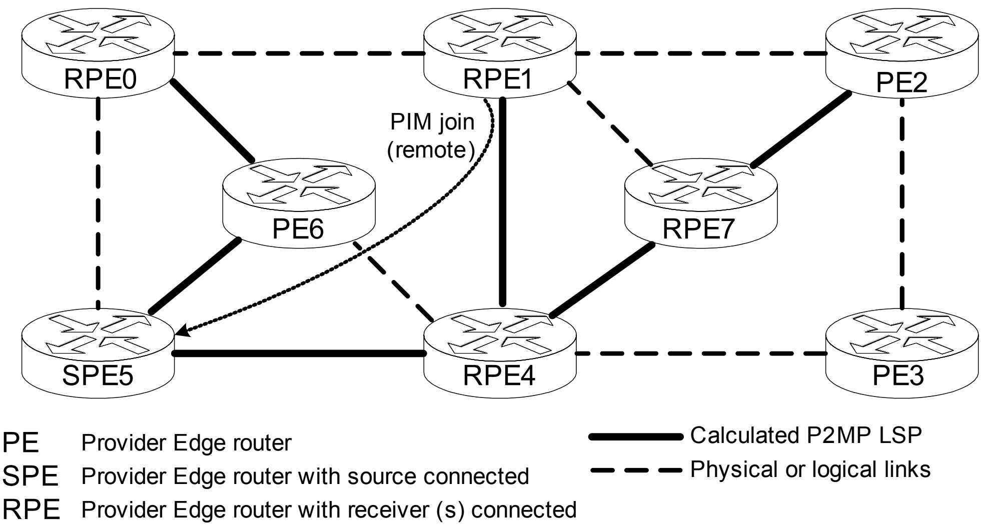

Fig. 2 PIM operation in MPLS network

Before RPE1 can receive traffic from P2MP LSP it has to associate appropriate P2MP LSP to the RPF interface for a given (S,G) entry[2]. For this purpose [9] introduces a new PIM message called *Join Acknowledge*. This message can be sent by SPE5 to RPE1 and convey the P2MP LSP identifier associated with (S1, G1) entry.

After receiving *Join* message from new remote PIM neighbor SPE5 can run heuristic algorithm to calculate distribution tree. Fig. 2 shows the example network

[2] Since the P2MP LSP provides multicast loop free environment, the RPF mechanism could be disabled for this case. However, it seems that it is easier to modify PIM protocol operation rather than changing RFP check procedure, which is fundamental for multicast.

where KPP algorithm was used. Note that after receiving each Join message, the P2MP LSP reflecting the distribution tree calculated by heuristic algorithm, can be significantly modified. This should not have an impact to network traffic unless new P2MP LSP will be completely established before switching traffic to it. The procedure for selecting conditions which needs to be met before starting P2MP LSP recalculation rather than adding new branch to it, will be subject of further studies.

7. CONCLUSIONS

The implementation of multicast transmission in IP networks has evolved substantially over the past years. Today many IPSs willing to provide triple play services are interested in implementing this way of data delivery in their networks. It seems that assuring appropriate QoS parameters for multicast transmission is one of the most important challenges in the development of multiservice backbone networks. On the one hand, MPLS technology allows to assure required QoS parameters, but on the other hand, the implementation of multicast transmission in such networks is very complex [1]. Many issues remain still open and require additional research [4]. This article briefly describes the most important problems which will be the subject of further studies. The article focuses on IP multicast transmission, but most of the presented topics are also essential for the emulation of Ethernet multicast transmission in multi-site L2VPNs.

The solution presented in this paper gives a possibility to perform research on merging heuristic protocols used for multicast distribution tree computation with signaling protocols used to implement this tree in the network. We believe that it gives a good background for planning further simulations which would allow to rate particular algorithms and protocols.

REFERENCES

[1] Hermsmeyer Ch., Hernandez-Valencia E., Stoll D., Tamm O.: Ethernet aggregation and core network models for effcient and reliable IPTV services: Research Articles, Bell Lab. Tech. J., 12, 2007, pp. 57-76.

[2] Karp R.,: Reducibility among combinatorial problems, Complexity of Computer Computations, 1972, pp. 85–104.

[3] Matrawy A., Yi W., Lambadaris I., Lung C.: MPLS-based Multicast Shared Trees, Proceedings of the 4th Annual Communication Networks and Services Research Conference (CNSR'06), 2006, pp. 227-234.

[4] Martinez-Yelmo, I. Larrabeiti, D. Soto, I. Pacyna, P.: Multicast traffic aggregation in MPLS-based VPN networks, Communications Magazine, IEEE, October 2007, pp. 78-85.

[5] Katz D., Kompella K., Yeung D.: Traffic Engineering (TE) Extensions to OSPF Version 2, RFC 3630, IETF, September 2003.

[6] Awduche D., Berger L., Gan D., Li T., Srinivasan V., Swallow G.: RSVP-TE: Extensions to RSVP for LSP Tunnels, RFC 3209, IETF, December 2001.

[7] Farrel A., Papadimitriou D., Vasseur J.-P., Ayyangar A.:Encoding of Attributes for Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Establishment Using Resource ReserVation Protocol-Traffic Engineering (RSVP-TE), RFC 4420, IETF, February 2006.

[8] Aggarwal R., Papadimitriou D., Yasukawa S.: Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs), RFC 4875, IETF, May 2007.

[9] Aggarwal R., Pusateri T., Farinacci D., Wei L.: IP Multicast With PIM-SM Over a MPLS Traffic Engineered Core, Internet Draft, IETF, work in progress.

[10] Piechowiak M., Zwierzykowski P.: "Analiza efektywności algorytmów heurystycznych dla połączeń rozgałęźnych w sieciach pakietowych", Krajowe Sympozjum Telekomunikacji 2005, pp. 192-200.

[11] Zwierzykowski P., Szewczyk T., Piechowiak M., Bartczak T.: "Multicast IP technologia przyszłości", Networld, Number 6 2007, pp. 65-68.

[12] Piechowiak M, Zwierzykowski P.: A new multicast routing algorithm with lagrange relaxation, submited for Poznanskie Warsztaty Telekomunikacyjne, December 2007.

[13] Bartczak T., Zwierzykowski, P.: Algorytmy genetyczne dla połączeń rozgałęźnych, Poznańskie Warsztaty Telekomunikacyjne, December 2005.

[14] Piechowiak M., Zwierzykowski P., Efficiency Analysis of Multicast Routing Algorithms in Large Networks, Poznan University of Technology Academic Journals, 2007 pp. 173-180.

[15] Crawford J.S., Waters A.G.: Heuristics for ATM Multicast Routing, Sixth IFIP Wokshop on Performance Modelling and Evaluation of ATM Networks, July 1998.

[16] Adams A., Nicholas J., Siadak W.: Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification, RFC 3973, 2005.

[17] Fenner B., Handley M., Holbrook H., Kouvelas I.: Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification, RFC 4601 2006.