

# PIM-DM Cost Analysis in Loop Free Topologies

G. Maciá<sup>1</sup>, J. E. Díaz-Verdejo<sup>2</sup>, J.M. Estévez-Tapiador<sup>3</sup>

<sup>1,2</sup>Dpto. Teoría de Señal, Telemática y Comunicaciones

Universidad de Granada – Granada (Spain)

jedv@ugr.es

<sup>3</sup>Departamento de Informática

Universidad Carlos III de Madrid – Madrid (Spain)

jestevez@inf.uc3m.es

## Abstract

*This paper presents an approach to estimate the cost of the PIM-DM protocol in terms of the number of packets, both for data and control traffic. The proposed approach assumes a loop-free network topology and that all links have equal parameters. Although restrictive at a first glance, the results show a good performance in simulated real networks when mean values for the parameters are used. The expressions are deduced from the protocol functioning, overcoming limitations and approximations of previously published works.*

## 1. Introduction

Multicast is an IETF standard [1] with many applications, among others, the experimental Mbone network that has been operational from 1992 [2]. This technology provides a theoretical advantage when compared to unicast for applications designed for providing service to a large, distributed amount of clients or receivers of information. This is due to the fact that only a single multicast packet is sent to a group, no matter how many receivers have joined it. Packet duplication only occurs when paths to multiple receivers diverge, which implies overall bandwidth reduction when compared to unicast. Nevertheless, if multicast transmission is considered in other scenarios, such as small networks with few receivers, its advantage is not as clear due to the cost involved in multicast tree management. In fact, tree construction introduces an unavoidable cost that is not present in unicast. Our work is focused on the problem of finding scenarios in which differences in transmission costs will not be very high, in such a way that unicast

transmissions could be a better choice than multicast. In order to achieve this purpose, an estimation of the cost introduced by the PIM-DM (Protocol Independent Multicast – Dense Mode) protocol [3] is required. This estimation could be used to make a decision among unicast vs. multicast transmission in a given environment. Furthermore, this estimation can be used to switch from an unicast to a multicast approach on the fly.

The remaining of this paper is organized as follows. Section II reviews previous work and identifies some effects not considered in the bibliography that can become relevant in cost analysis. Section III proposes new expressions that collect the previously explained behaviors, yielding a more adequate description for loop-free topologies. Section IV will show the applicability of those expressions to real networks. Finally, Section V concludes summarizing the benefits of the introduced work as well as future research objectives.

## 2. Previous work

In [4], several expressions were proposed to calculate the cost introduced by unicast and multicast PIM-DM protocol. The notion of cost is modeled through a hop-based approach; that is, the proposed expressions calculate the overall number of hops traversed by all the packets generated in a multicast transmission. This concept is split by [4] in two terms: data cost and control overhead cost. Data cost is related to the hops traversed by packets that are traveling across the multicast tree, while control overhead cost consider the effect of the flood-and-prune process, where data packets are periodically flooded to the entire network (overhead cost), and the branches are

pruned where there are no downstream receivers (control cost).

Assuming that  $C$  denotes the cost associated to a multicast source during time  $T_{\text{transmission}}$ , the proposed expression is<sup>1</sup>:

$$C = CD_{DM} + CO_{DM} = \alpha \cdot L_m + 2(L'_m - L_m) \cdot F \quad (1)$$

where  $CD_{DM}$  is the data cost;  $CO_{DM}$  is the control and overhead cost;  $\alpha$  is the number of packets transmitted by the source;  $L_m$  is the number of links in the multicast tree;  $L'_m$  is the length, in number of links, of the broadcast tree (i.e., the number of links of the whole network); and  $F$  represents the number of floodings during the observation time. Obviously,  $F$  can be expressed as:

$$F = \text{ceil} \left( \frac{T_{\text{transmission}}}{\tau_{dm}} \right) \quad (2)$$

where  $\tau_{dm}$  is the time between periodic floodings, also called *pruning timer*.

These expressions are adequate to estimate the cost in large networks, mainly in scenarios where the data cost is greater than the control and overhead cost. Nevertheless, the expression for control and overhead cost does not consider two key aspects of the behavior of the PIM-DM protocol.

First, the expression (1) assumes that only a multicast packet and a prune response are transmitted for all links that do not belong to the multicast tree. Obviously, this consideration implies that the number of overhead packets corresponds to twice the number of those links. However, when a router starts to send multicast packets, it does not stop until a prune response is received. Therefore, and assuming that a router transmits at constant bit rate, more than one packet could be eventually sent until the router will receive the prune response.

Second, a flooding every  $\tau_{dm}$  seconds is supposed according to expression (2). This implies that the pruning timer is initiated at the reception of the first prune, and that subsequent prune receptions do not reset it. Nevertheless, the timer is effectively reset at the reception of any prune message [3].

Consequently, while expressions given by [4] provide a good estimation in scenarios where the overhead cost is not relevant, it is necessary to modify them for those situations in which this cost becomes significant.

<sup>1</sup> Expressions given by [4] are actually expressed in terms of rate instead of in number of packets.

### 3. Cost estimation for PIM-DM algorithm

In this work, the cost of the PIM-DM algorithm will be estimated under the following assumptions:

1. There is only one multicast source that is continuously transmitting packets at a constant bit rate (CBR packets) during a time  $T_{\text{transmission}}$ .
2. There is neither loss of packets nor congestion in any part of the network.
3. Link delay time and bandwidth are equal along the links in the network. Although this assumption is not realistic, in section IV we will point out some annotations that will show how to adapt this work to realistic environments.

In order to simplify the explanations, the estimation of the overhead cost will be split in two terms: the cost due to the generation of prune messages, termed "*prune cost*" ( $CO_{DM}^{\text{prune}}$ ), and the additional cost produced by useless CBR packets in periodic floods of the protocol, termed "*CBR cost*", ( $CO_{DM}^{\text{cbr}}$ ):

$$CO_{DM} = CO_{DM}^{\text{cbr}} + CO_{DM}^{\text{prune}} \quad (3)$$

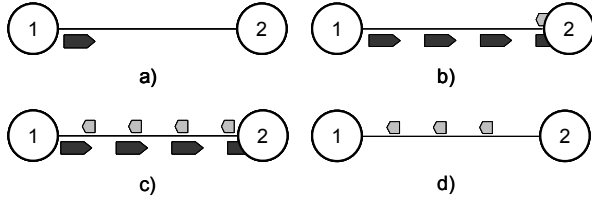
#### 3.1 Prune cost estimation for n-hop branches composed trees

In order to estimate the prune cost we will start from a simple model, represented in Figure 1. In this figure we can see two routers in a point to point link. One of them (node 1) is acting as a multicast source and the other (node 2) as a network element that does not belong to the multicast group that we consider. The behaviour of this model is as follow. In  $t = 0$ , the source begins to send packets (Figure 1.a). When the first multicast packet arrives node 2 (Figure 1.b), it generates a prune answer to this message that is sent back to node 1. When this answer reaches node 1 (Figure 1.c), it stops sending multicast packets, and starts  $\tau_{dm}$  timer. Of course, each multicast packet that arrives at node 2 produces the same behaviour (Figure 1.d).

Let  $t_1$  be the time a multicast packet requires to travel from one node to its neighbor, and  $t_2$  the time used by the generated prune response to make the reverse path. Let  $t_0$  be the addition of these components:

$$t_0 = t_1 + t_2 = 2 \cdot t_p + \frac{L_{\text{cbr}} + L_{\text{prune}}}{BW} \quad (4)$$

where  $t_p$  represents the link propagation time;  $L_{\text{cbr}}$  and  $L_{\text{prune}}$  are, respectively, the CBR and prune packets sizes; and, finally,  $BW$  is the link bandwidth.



**Figure 1: Flood-and-prune mechanism in a single-hop branch**

Obviously, aspects as buffer delays and time processing in routers are not considered.

By a simple examination of the model in Figure 1, we can observe that the number of prune messages generated is equal to the number of CBR messages sent by the source. Therefore, the number of prune messages  $N$  created during a flooding for this example is:

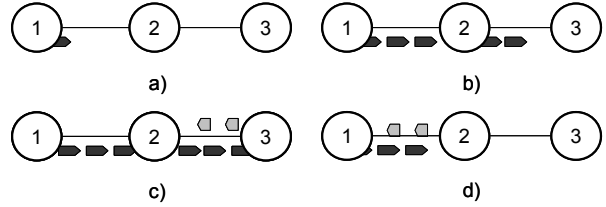
$$N = \text{ceil} \left[ \frac{t_0}{t_{cbr}} \right] \quad (5)$$

where  $t_{cbr}$  is the time between two consecutive sendings of multicast packets. Consequently, the number of packets produced during  $F_1$  floods that happen during a  $T_{\text{transmission}}$  period can be calculated as:

$$CO_{DM}^{prune} = F_1 \cdot \text{ceil} \left[ \frac{t_0}{t_{cbr}} \right] \quad (6)$$

Now we are ready to extend these results to more complex scenarios. First, we will consider a two hop branch (Figure 2) and afterwards we will extend the conclusions for a  $n$ -hop branch scenario. The behaviour in this two-branches scenario is quite similar to the previous one. Node 1 starts sending multicast packets (Figure 2.a). Upon reaching node 2, they are forwarded to node 3 (Figure 2.b). The reception of a packet in node 3, which has the knowledge of not belonging to the multicast tree<sup>2</sup>, generates a prune answer (Figure 2.c). The first prune answer received in node 2 produce another prune message that is sent from node 2 to node 1 (Figure 2.d). When this prune arrives at node 1, it stops sending multicast packets, starting the pruning timer  $\tau_{dm}$ . We can see that, in this scenario, a prune message is generated in two cases: as a response to a CBR packet from the source, and upon the reception, in an intermediate node (node 2), of a prune message. This second case only produces one packet on each flooding of the protocol, while the first produce  $N$  packets on each link.

<sup>2</sup> This knowledge could have been previously acquired from sources like the IGMP protocol.



**Figure 2: Flood-and-prune mechanism in a two-hops branch.**

Let us suppose a more general scenario composed of  $n$  routers, termed  $r_i$ , where each  $r_i$  and  $r_{i+1}$  are connected by a point to point link. Router  $r_1$  will act as a multicast source, while the rest are elements that do not belong to the multicast group.

In this scenario, a prune message is generated by each node in two cases: as a response to every received CBR packet, whose number is:

$$N = \text{ceil} \left[ \frac{t_0}{t_{cbr}} \right] \quad (7)$$

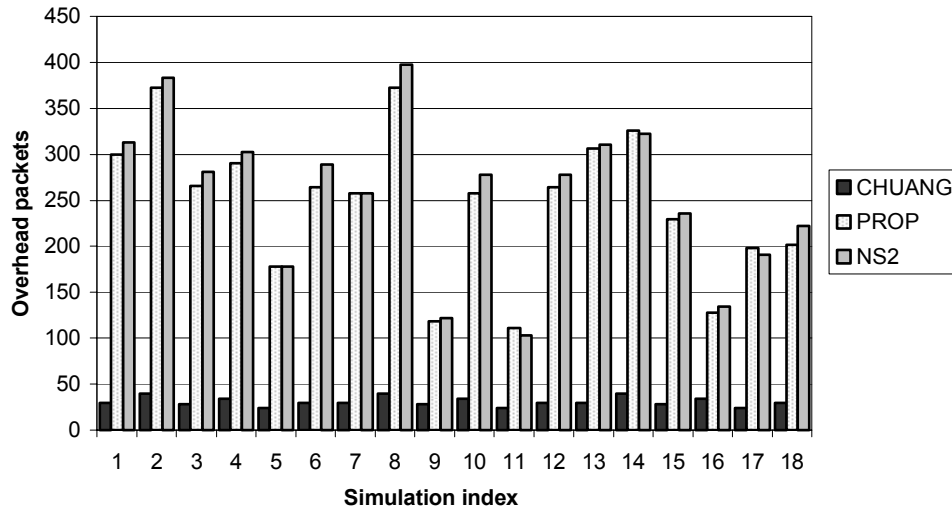
and after the reception, in an intermediate node, of a prune message. This second case produces one single packet on each flooding of the protocol, while the first produces  $N$  packets on each link. Therefore, we conclude that the number of prune packets generated in this scenario is:

$$CO_{DM}^{prune} = \sum_{i=1}^n L_i \cdot F_i \cdot \text{ceil} [i \cdot N + i - 1] \quad (8)$$

where  $L_i$  represents the number of  $i$ -hop branches in the tree, and  $F_i$  is the number of floodings that occur on  $i$ -hop branches during the transmission time. It is important to note that the number of floodings is dependent on the number of hops in the branch due to the fact that the pruning timer is reset on every prune message reception.

### 3.2 CBR cost estimation for $n$ -hop branches composed trees

As an example, consider now the previous scenario with  $n=3$  (Figure 2). In this case, three phases can be distinguished from the point of view of the number of CBR packets generated, that is, intervals  $(0, t_1)$  and  $(2t_1+t_2, 2t_0)$ , in which there is only one CBR packet in the whole branch corresponding to one CBR packet generated by the source, and interval  $(t_1, 2t_1+t_2)$ , where two packets exist for each one generated by the source. Therefore, in this scenario, the number of CBR packets is equal to the number of prune packets only during  $t_0$ , and twice during the same time  $t_0$ . A simple reasoning



**Figure 3: Number of overhead packets for 18 random networks topologies estimated according to [4] (series CHUANG), to the novel expressions (series PROP) and obtained from a simulation by using Network Simulator 2 (series NS2).**

will yield the expression for a generic n-hop branches composed tree:

$$CO_{DM}^{cbr} = \sum_{i=1} F_i \cdot L_i \cdot \sum_{j=1}^i [i \cdot N] \quad (9)$$

### 3.3 Estimation of the number of floodings

We can estimate the time  $T_i$  between the sending of the first CBR packet and the reception of the last prune message in a i-hop branch considering that the maximum time between the first CBR packet sending and the first prune reception is equal to  $t_0$ —prune message from the first node—, and the maximum time from the reception of the first prune message to the last one is i-1 times  $t_0$ —prune message from the deepest node—, if  $\tau_{cbr} > t_0$ . If  $\tau_{cbr} < t_0$ , an additional prune message should be considered from the first node as a response to an in-transit CBR packet. Therefore, this time corresponds to:

$$T_i = \begin{cases} (i+1) \cdot t_0 + \tau_{dm} & \text{if } \tau_{cbr} < t_0 \\ i \cdot t_0 + \tau_{dm} & \text{if } \tau_{cbr} \geq t_0 \end{cases} \quad (10)$$

Finally, the number of floodings in an interval  $T_{transmission}$  is:

$$F_i = \text{ceil} \left[ \frac{T_{transmission}}{T_i} \right] \quad (11)$$

### 3.4 Extension to generic loop free topologies

Expressions (8) and (9) evaluate the costs in single branch topologies. If we consider extending these expressions for generic loop-free topologies with combined branches, each one composed by several hops, we find out that expressions (8) and (9) are still valid, although care about  $L_i$  coefficients is needed.  $L_i$  is the number of branches with i-hops, which can be determined by exploring the network using the following algorithm:

*For (each branch that does not belong to the multicast tree)*  
 {  
   *Take the longest path P (in number of hops) from the source.*  
   *Increase  $L_i$  in one unit, being i the number of hops in path P.*  
   *For each node on path P, go to 1.1.*  
 }

## 4. Cost estimation in loop-free real networks

The expressions (8), (9) and (11), together with the algorithm for evaluating  $L_i$  coefficients, have been experimentally checked by using NS2 simulator [5]. A total number of 20 random topologies were generated by using the topology generator GT-ITM [6], yielding satisfactory results.

In the previous discussion, cost estimation has been obtained under the assumption that all the links in the network have the same bandwidth and delay time. Since real networks do not fulfill these requirements, our next objective will be to obtain an estimation of how thick is the approximation. We afford the task giving these parameters the mean value of all the network links. This way, 18 experimental simulations have been carried out using topologies with a wide range of values for the link delay time and bandwidth. All the simulations present a greater deviation in the CBR cost (variation around 6.5% with variance of  $\pm 7\%$ ) than in the prune cost (mean variations of 1.9%, with low variance of  $\pm 2.5$ ). Anyway, the proposed expressions provide a much more accurate estimation than those in [4], as shown in Figure 3.

#### 4. Conclusions and future work

In this work, we have improved the expressions proposed in [4] for the evaluation of cost of the PIM-DM protocol, in terms of the number of packets. As Chuang's work proved to be inaccurate in scenarios in which the overhead cost is relevant, some unconsidered aspects of PIM-DM have been included, yielding expressions that evaluate this cost in loop-free topologies. Although new expressions were deduced under the assumption of homogeneous link delay time and bandwidth, their application to real loop-free networks shows an important increase in accuracy, when compared with Chuang's.

Further work will be devoted to improve the applicability of the expressions to real networks by including detailed models for delay time and bandwidth. Anyway, a main advantage of the proposed expressions resides in its simplicity, which allows a quick and easy cost evaluation. On the other hand, an important objective to be tackled in future work is to extend the expressions to non loop-free topologies.

#### 5. References

- [1] S. Deering, "Host Extensions for IP Multicasting" RFC 1112, 1989.
- [2] H. Eriksson. "Mbone: the multicast backbone". *Communications of the ACM* 37(8) (1994) 54-60.
- [3] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, A. Helmy, D. Meyer, and L. Wei, "Protocol Independent Multicast Version 2 Dense Mode Specification," Internet Draft, June 1999.
- [4] J. Chuang and M. Sirbu, "Pricing multicast communication: A cost-based approach," in *Proceedings of INET'98*, Geneve, Switzerland, July 1998
- [5] Network Simulator 2. Available at: <http://www.isi.edu/nsnam/ns/>
- [6] K. Calvert, M. Doar, and E. Zegura, "Modeling Internet Topology," in *IEEE Communications Magazine*, Vol. 35. No. 6, Jun. 1997, pp. 160-163.

#### Acknowledgements

This work has been partially supported by the Spanish Government through MCYT (project SERVIRA, TIC2002-02798, FEDER funds 70%).