Available online at www.sciencedirect.com
ScienceDirect
Computer Communications 31 (2008) 489–505
computer communications
www.elsevier.com/locate/comcom

Multicast algorithms in service overlay networks ☆

Dario Pompili [a,*], Caterina Scoglio [b], Luca Lopez [c]

[a] Rutgers, The State University of New Jersey, Department of Electrical and Computer Engineering, Piscataway, NJ 08854, USA
[b] Kansas State University, Department of Electrical and Computer Engineering, Manhattan, KS 66506, USA
[c] University of Rome La Sapienza, Department of System Engineering, 00184 Rome, Italy

Available online 29 August 2007

Abstract

Overlay routing enhances the reliability and performance of IP networks since it can bypass network congestion and transient outages by forwarding traffic through one or more intermediate overlay nodes. In this paper, two algorithms for multicast applications in service overlay networks are presented. The first algorithm is tailored for source-specific applications such as live video, software and file distribution, replicated database, web site replication, and periodic data delivery; it builds a virtual source-rooted multicast tree to allow one member in the multicast group to send data to the other members. The second algorithm is tailored for group-shared applications such as videoconference, distributed games, file sharing, collaborative groupware, and replicated database; it constructs a virtual shared tree among group members. The objective of both algorithms is to achieve traffic balancing on the overlay network so as to avoid traffic congestion and fluctuation on the underlay network, which cause low performance. To address these problems, the algorithms actively probe the underlay network and compute virtual multicast trees by dynamically selecting the least loaded available paths on the overlay network. This way, network resources are optimally distributed and the number of multicast trees that can be setup is maximized. Both algorithms can offer service differentiation, i.e., provide QoS at application-layer without IP-layer support. The low computational complexity of the proposed algorithms leads to time and resource saving, as shown through extensive simulation experiments.

Keywords: Overlay networks; Multicast; Mathematical programming/optimization; Multirate layering

1. Introduction

Overlay routing has been proposed to enhance the reliability and performance of IP networks since it can bypass network congestion and transient outages by forwarding traffic through one or more intermediate overlay nodes [34,41,42]. Service overlay network is an effective means to address end-to-end Quality of Service (QoS), plaguing the current Internet, and to facilitate the creation and deployment of value-added Internet services such as VoIP, Video-on-Demand, and other emerging QoS-sensitive services. While much of the past research in overlay networking has focused on techniques for building and reconfiguring overlay networks, and evaluating their performance (e.g., [3,17,30,37]), in this paper we present two algorithms to build virtual multicast trees on an overlay network for as many applications as live video, software and file distribution, replicated database, web site replication, periodic data delivery, videoconference, distributed games, file sharing, and collaborative groupware. Moreover, we propose a low complexity multirate layering algorithm to support the virtual multicast algorithms by accommodating heterogeneous receiver requested bit rates.

Throughout this paper, we consider *two layers* of network infrastructure: the *native network*, which includes end-systems, routers, links and the associated routing functionality, and provides best-effort datagram delivery between its nodes; and a *virtual overlay network*, which is formed by a subset of the native layer nodes interconnected

☆ This work was supported by the National Science Foundation (NSF) under Grant No. ANI-0219829.
* Corresponding author. Tel.: +1 732 445 6400x202.
E-mail address: pompili@ece.rutgers.edu (D. Pompili).

0140-3664/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2007.08.023

through overlay links to provide enhanced services. Overlay links are *virtual* in the sense that they are IP tunnels over the native network.

The first proposed algorithm is called *DIfferentiated service Multicast algorithm for Internet Resource Optimization (DIMRO)*. It builds virtual source-rooted multicast trees for source-specific applications. DIMRO takes the virtual link available bandwidth into account to avoid traffic congestion and fluctuation on the underlay network, which cause low performance. The objective is to keep the average link utilization of the overlay network low by fairly distributing data flows among the least loaded links. The second algorithm is called *DIfferentiated service Multicast algorithm for Internet Resource Optimization in Group-shared applications (DIMRO-GS)*. It constructs a virtual shared tree for group-shared applications by connecting each member node to all the other member nodes with a source-rooted tree computed using DIMRO. To support these algorithms, a novel low complexity layering algorithm to accommodate heterogeneous receiver requested bit rates is also proposed, which makes DIMRO and DIMRO-GS native multirate multicast algorithms on overlay networks.

Both DIMRO and DIMRO-GS algorithms offer *service differentiation*, i.e., provide QoS at application-layer without IP-layer support. Consequently, multicast group members with less stringent QoS requirements manage to reuse resources already exploited by members with more stringent requirements. This results in a better utilization of network bandwidth, and in an improved QoS as perceived by multicast group members.

To summarize, the main contributions of this work are:

(1) Development of QoS-aware multicast algorithms on overlay networks, which are able to provide application-layer QoS for multicasting without IP-layer support.
(2) Design of low-complexity overlay multirate algorithms to leverage the network bandwidth resources and improve QoS as perceived by multicast group members.

The remainder of the paper is organized as follows: in Section 2, we provide a brief background on multicasting, while in Section 3 we review the relevant literature on overlay networks. In Section 4, we formulate the logical channel rate assignment problem and describe DIMRO, while in Section 5 we introduce DIMRO-GS. In Section 6, we show numerical results through extensive simulation experiments. Finally, in Section 7, we draw the main conclusions and point future work.

2. Background on multicasting

In unicast transmissions, the sender transmits data to a single receiver and, if multiple receivers want to receive the same data content, the sender has to transmit multiple copies of data. In multicast transmission, conversely, the sender transmits only one copy of data that is delivered to multiple receivers. This allows to efficiently exploit Internet resource in as many applications as live video, software and file distribution, replicated database, web site replication, periodic data delivery, videoconference, distributed games, file sharing, and collaborative groupware. One of the most challenging objective in multicasting is to minimize the amount of network resources utilized to compute and setup multicast trees [32,33]. In multicast communication, the routing problem is to find the *minimum-weight tree* that spans all the nodes in the multicast group [15,36,40]. Multicast communication can be classified into two types, i.e., *source specific* and *group shared*. In source-specific multicast communication, only one node in the multicast group sends data, while all the other member nodes receive data. In group-shared multicast communication, each node in the multicast group wants to send/receive data to/from member nodes. A tree that spans all member nodes is called *multicast tree*. Consequently, based on the communication strategy, multicast trees can be classified into two types, i.e., *source-rooted trees* and *shared trees*. A source-rooted tree has the source node as root, and is optimized for source-specific multicast communications, whereas a shared tree is optimized for group-shared communications, and connects each group member with all the other group members.

The classical optimization problem in source-specific multicast communications is the *Steiner tree Problem in Networks* (SPN) [2], whose objective is to find the least-cost tree connecting the source and the group of destinations with the minimum total cost over all links. If each destination has a bandwidth requirement, then the problem is to find the *least-cost* tree that complies with the bandwidth requirements on each path from the source to the receiver. It can be shown that both these problems are at least as complex as the *geometric connected dominating set* problem, which is proven to be NP-complete [21,28]. Hence, both these problems are NP-complete, and efficient algorithms to solve these problems in polynomial time attain only approximate solutions in most cases [2,36].

In group-shared communication, one of the earliest protocol to build a shared tree was the Core-Based Protocol (CBT) [5,6] and its enhanced version (CBTv2) [4], which connect each member node to a *core node* using bi-directional shortest paths. The main drawback is the traffic concentration occurring in the core node. To avoid traffic concentration, a shared tree can be computed by finding as many source-rooted trees as the number of member nodes, each of them with all the other member nodes as leaves. The FTM algorithm [29,38], Feasible solutions using adapted Takahashi–Matsuyama (TM) algorithm, follows this approach. The source-rooted tree connects the root node with the member node by using the path with the greatest capacity. The bottleneck of a path is characterized as the link with the minimum available bandwidth, and the path bandwidth capacity is defined as the available bandwidth of its bottleneck. The path with the greatest