A New Scalable Multicast Solution in MPLS Networks*

Lie Qian1, Yiyang2, Yuke Wang1, Bashar Bou-Diab3, and Wladex Olesinski3

1. Department of CCPS
Southeastern Oklahoma State University
Durant, OK 74701, USA
1qian@sou.edu

2. Department of Computer Science
University of Texas at Dallas
Richardson, TX 75080 USA
{yiyian, yuke}@utdallas.edu

3. Research and Innovation, Alcatel
Ottawa, Ontario, K2K 2E6, Canada
{Bashar.Bou-Diab, Wladyslaw.Olesinski}@alcatel.com

Abstract — Multicast is an efficient mechanism for delivering data to multiple receivers. However, conventional multicast routing suffers from a linear scalability problem that impedes the deployment of IP multicast in the Internet. The number of forwarding states maintained in the network layer increases linearly with the number of existing multicast channels in the network. MPLS coexists with many existing network protocols and can provide scalability for how these protocols are used in today's networks. In this paper, we propose a new scalable multicast solution in MPLS networks to reduce the number of forwarding states in the network layer. In previously existing solutions, forwarding states are removed from non-branching routers on multicast trees and packets are label switched by MPLS LSPs in non-branching routers. Our new solution proposes a new algorithm for multicast tree construction and a new channel sharing scheme to further reduce the number of forwarding states in branching routers. Simulation results show that the new solution can achieve 85%-97% reduction in the number of multicast forwarding states in MPLS networks over network layer multicast schemes, and 45%-90% reduction over the existing MPLS Multicast Tree-based multicast.

I. INTRODUCTION

Multicast is an efficient mechanism for the delivery of data from one or multiple sources to a set of destinations (receivers) identified by a multicast channel (group) [1]. A multicast channel is defined as a collection of packets identified by the same channel ID. In case multiple sources exist, the multicast channel ID is a multicast IP address. In case of a single source, a channel ID is a (S,G) address pair [2][3], where S is the source's network address and G is a multicast group address. Multicast routing protocols establish logical multicast tree(s) for a channel to connect the source(s) with the destinations. Packets are delivered to destinations following the constructed multicast trees. There are two main variants of multicast tree. The first one is the Source Specific Multicast Tree for a channel with a single source, which is a multicast tree rooted at its source and having all destinations as leaves [2]. The second is the Shared Multicast Tree for channels with multiple sources, which is rooted at a Rendezvous Point and has all destinations as leaves [4]. The rendezvous point collects packets from multiple sources and then delivers packets to destinations following the shared multicast tree. In this paper, our discussion focuses on single source multicast with source specific multicast trees.

In the conventional network layer multicast scheme, every router R on the multicast tree of channel c maintains a forwarding state of channel c for routing in the network layer, which contains the child routers of R on the multicast tree. Therefore, conventional multicast has limited scalability because the number of maintained forwarding states in a router increases linearly with the number of multicast channels passing through it. Different schemes have been proposed to reduce the forwarding states in routers and to improve multicast scalability. Proposed scalable schemes could be categorized into 1) aggregation based schemes, 2) smart packet based schemes, and 3) tunnel based schemes.

Multi-protocol label switching (MPLS) [5] is a network technology that promises to offer high speed packet forwarding, QoS, traffic engineering and many other new and timely features to the current best-effort, IP-based Internet. MPLS is an important enabling technology for the next generation of Internet devices such as IP terabit routers. MPLS can coexist with many existing network layer and data link layer protocols and provide scalability for how these protocols are used in today's networks.

In this paper, we propose a new scalable multicast solution to reduce the number of forwarding states in MPLS networks. In previously existing solutions, forwarding states are removed from non-branching routers on multicast trees and packets are label switched by static or dynamic MPLS Label Switched Paths (LSPs) in non-branching routers. Our new solution proposes a new multicast tree construction algorithm and a new channel sharing scheme to further reduce the number of forwarding states in branching routers. Simulation results show that our new solution can achieve 85%-97% reduction in the number of multicast forwarding states in MPLS networks over network layer multicast, and 45%-90% reduction over the existing scalable multicast scheme in MPLS networks, such as MPLS Multicast Tree (MMT) [6][7]. Furthermore, the new solution is efficient in the usage of MPLS labels. It consumes only 10% of the MPLS labels required in pure MPLS layer multicast routing scheme. With our new solution, through controlling the channel sharing criteria, network operators can adjust the scalability burden between MPLS layer and network layer.

The rest of this paper is organized as follows. In Section II, we review the network model, MPLS networks and related works in scalable multicast.

II. BACKGROUND

In this section, we review the network model, MPLS networks and related works in scalable multicast.

A. Network Model

In this paper, we focus on the single source multicast model, in which a multicast tree is constructed for a channel with destinations as leaves and with the source as the root. On a multicast tree, there are three types of routers 1) Branching Routers (BRs) – root router and routers that have more than one child router on the multicast tree, 2) Non-Branching Routers (NBR) - routers that have only one child router and
one parent router on the multicast tree, and 3) Destination designated Routers (DRs) - leaves on the multicast tree that deliver the multicast packets to connected hosts or receivers. Packets are only duplicated at branching routers and destination routers on a multicast tree. Given a branching router BR1, its next hop BRs are the branching routers on the tree reachable from BR1 via non-branching routers only. BR1’s next hop DRs are the destinations’ designated routers on the tree reachable from BR1 via non-branching routers only. If BR2 is BR1’s next hop BR on the tree of channel c, then we say BR1 is BR2’s parent branching router on the tree, denoted as parent_BR(c, BR2).

B. MPLS Network

MPLS [5] is a versatile data transport solution that addresses network problems such as scalability, QoS management, and traffic engineering. Label Edge Routers (LERs) in MPLS networks compute and prepend 32-bit long MPLS shim headers that includes 20-bit long MPLS labels, to incoming packets based on Forwarding Equivalence Classes (FECs). An MPLS Label Switching Router (LSR), maintains a MPLS label switching table, with each entry specifying an ingress MPLS label $L_{in}$, one or multiple egress MPLS labels $L_{out}$ and egress interface ID(s) $I_{out}$. When packet $P$ carrying MPLS label $L_{in}$ arrives at LSR $R$, $R$ searches for an entry with ingress label $L_{in}$ in the MPLS label switching table. For each pair ($L_{out}$, $I_{out}$) in the matched entry, $R$ swaps $L_{in}$ in $P$ with $L_{out}$ and forwards packet $P$ to neighbor router through egress interface $I_{out}$. Such process is called “label switching” of the packet. The specific path through the MPLS network that a packet follows based on its MPLS labels is called Label Switched Path (LSP). LSPs could be either point to point or point to multi-point.

C. Related Works

Various solutions have been proposed to improve the scalability of multicast through reducing the number of forwarding states in the network layer. Network layer aggregation solutions [8]-[16] reduce the storage and routing complexity in routers through replacing multiple forwarding states in a router with one forwarding state. Due to the non-hierarchical allocation of multicast IP addresses, existing network layer aggregation solutions yield either insignificant reduction in the number of forwarding states or leaky bandwidth, in which multicast packets are delivered to routers that are not members of the group.

Another scalable multicast solution category is smart packet solutions [17]-[21]. In smart packet solutions, the information of packet forwarding is stored in the packet header to eliminate multicast forwarding states in routers. Routers forward packets based on information extracted from the packet headers. Smart packet solutions require changes in the format of packet headers, limit the number of branching routers, and cause problems due to a lot packet fragmentation.

IP encapsulation solutions [22]-[25], belonging to network layer tunnel based solutions, have been proposed for scalable multicast. In IP encapsulation solutions, a packet is transported from one branching router to its next hop BR (or DR) through unicast IP encapsulation, in which the original packet is encapsulated in a new packet that has the next hop BR’s address as destination address. Forwarding states are removed from non-branching routers. IP encapsulation solutions involve resource intensive extra IP packet en/decapsulation at branching routers.

A framework that explains IP multicast deployment in MPLS environment was proposed by Ooms et al [26], which is a pure MPLS layer multicast routing solution. In this framework, a point to multipoint LSP is used for a multicast tree. One point to multipoint LSP could be shared by multiple trees only if they have the same tree structure. Therefore, large number of multicast channels with various tree structures consumes too many MPLS labels to establish large number of point to multipoint LSPs and introduce scalability problem in MPLS layer. Another study about MPLS and multicast proposed by Farinacci et al. [27] explains how to use PIM to distribute MPLS labels for multicast routes.

MPLS Multicast Tree (MMT) solution was proposed in [6][7] for scalable multicast in MPLS networks, in which packets are transported from one branching router to the next hop BR (or DR) by MPLS label switching. Forwarding states are only stored in branching routers. However, MMT made no effort to reduce the forwarding states in branching routers. MMT2, proposed in [30], uses channel bounded MPLS label as identifier for each channel to reduce network layer forwarding states in branching routers. Similar as Ooms’ pure MPLS layer multicast routing solution in [26], MMT2 consumes too many MPLS labels to identify large number of channels (such consumption only exist in branching router in MMT2 and exist in all on-tree routers in Ooms’ solution) and cause scalability problem in MPLS layer at branching routers.

III. New Scalable Multicast Solution in MPLS Networks

In this section, we present our newly designed scalable multicast solution in MPLS networks. In the new solution, multicast packets are transported between branching/destination routers through dynamic or static tunneling LSPs. Compared with the MMT solution that also uses LSP for tunneling between branching/destination routers, we further improve the scalability by proposing a new Total Match SubTree (TMST) based multicast tree construction algorithm and a new MPLS layer channel sharing scheme to reduce the number of forwarding states in branching routers.

In our solution, we use a Centralized Network Information Management System (CNIMS) to compute multicast trees in a network. CNIMS collects information about network layout and channel membership. Then CNIMS computes the multicast tree for each channel. The computation for a channel discovers all branching routers. The CNIMS, having a complete overview about the network and multicast trees, uses the TMST based multicast tree construction algorithm to discover branching routers and corresponding next hop BRs (DRs) for each channel. Channel sharing scheme is implemented in routers to reduce the number of forwarding states in branching routers. Our new solution can be generally depicted as in Figure 1.

A. TMST based Multicast Tree Construction Algorithm

The TMST multicast tree algorithm tries to find TMSTs from existing multicast trees and use them to construct multicast trees for new channels.
We first define the child destinations of a branching router BR1 on channel d, denoted as Dest(d, BR1), to be a set containing all destinations of channel d that are reachable from BR1 on its multicast tree. We use $T_{sub}(d, BR1)$ to represent a subtree of channel d’s multicast tree. $T_{sub}(d, BR1)$ is rooted at BR1 and contains all the destinations in Dest(d, BR1). If another channel c has destination set equal to Dest(d, BR1), we say $T_{sub}(d, BR1)$ is a Total Match SubTree (TMST) of channel c.

CNIMS searches the existing multicast trees to find a TMST for channel c:

- If no TMST exists for channel c,
  - A multicast tree rooted at source S connecting all destinations of channel c is computed using any existing multicast tree algorithm (e.g., shortest path tree, Steiner tree)
- If there exists one or more such TMSTs,
  - For each found TMST, say $T_{sub}(d, BR1)$, CNIMS computes a new multicast tree by connecting the $T_{sub}(d,BR1)$ with the shortest path from S to BR1
  - CNIMS selects one multicast tree from the set above with the minimum tree cost and uses it as the multicast tree for channel c. We call such multicast tree a TMST multicast tree

CNIMS sends messages to all branching routers on the new constructed tree to inform them about their next hop BRs/DRs. On receiving this message, every branching router on the new tree creates a multicast forwarding state for multicast channel c.

Once branching routers are identified, packets will be sent from a branching router to another until they reach their address of their parent branching router. Each forwarding state contains information including channel ID, addresses of next hop BRs, addresses of next hop DRs, and the address of the parent branching router.

The example in Figure 2 shows an existing multicast tree of channel $(S1, G1)$ with destinations $\{d1, d2, d3, d4, d5\}$, TMST multicast tree for channel $(S2, G2)$ with destinations $\{d2, d3, d4, d5\}$ and shortest path tree for $(S2, G2)$. In the TMST multicast tree of $(S2, G2)$, three branching routers $\{R3, R4, R5\}$ contain the same set of next hop BRs/DRs as channel $(S1, G1)$, which gives a higher likelihood of channel sharing in these branching routers, which we will discuss in following subsections. It is noticeable that the TMST multicast trees are not guaranteed to be SPT or Steiner tree. We will evaluate the extra multicast tree cost in Section IV through simulations.

### B. Tunneling LSPs between Branching/Destination Routers

On the constructed multicast tree, the forwarding states in the network layer are only installed in branching routers and packets are transmitted between branching/destination routers by using MPLS LSPs in a manner similar to MMT [6][7] solution. The LSPs used for label switching between branching/destination routers could be either static tunneling LSPs or dynamic tunneling LSPs. Static tunneling LSPs are pre-established to reduce LSP setup latency in the multicast tree establishment process. Dynamic tunneling LSPs are dynamically established to avoid labels being wasted in tunneling LSPs not used by any multicast channel.

### C. Channel Sharing Scheme in Branching Routers

The purpose of channel sharing scheme is to further reduce the number of forwarding states in branching routers. Without channel sharing, packets are routed in branching routers based on forwarding states and label switched in non-branching routers. Each forwarding state contains information including channel ID, addresses of next hop BRs, addresses of next hop DRs, and the address of the parent branching router.

Suppose BR1 is a branching router for both channel c and channel d. Channel c and channel d are related at BR1 if they have exactly the same set of next hop BRs (DRs) and for each common next hop BR (DR), packets from both channels are forwarded using the same LSP in BR1. The set of all channels that are related to channel c in BR1 is denoted as $Relate(c, BR1)$.

Suppose $BR1$ is a branching router of both channel c and channel d. If packets from these two channels are label switched by the same MPLS point-to-multipoint label switching entry of the MPLS table in BR1, we say channel c is shared with channel d in BR1. This could happen only after the channel sharing process in $BR1$ on channel c and d as described below. The set of all channels that are shared with channel c in $BR1$ is denoted as $Share(c, BR1)$.

Given a related set $Relate(e, BR1)$ at $BR1$, the process in $BR1$ to share channels is described by the following steps:

**Step 1:** For each channel e in $Relate(c, BR1)$

- $BR1$ requests parent_BR(e, BR1), the parent branching router of $BR1$ on channel e, to return the channels in set
Share(e, parent_BR(e, BR1)), which are the channels shared with e in parent_BR(e, BR1).

- If any channel in Share(e, parent_BR(e, BR1)) is not in \(\text{Relate}(e, \text{BR1})\), channel \(e\) is removed from the set \(\text{Relate}(e, \text{BR1})\).

**Step 2:** Based on the number of channels left in \(\text{Relate}(e, \text{BR1})\), BR1 decides if the sharing process needs to continue, which could be used by the network operator to control the extent of sharing. If continue, BR1 selects an unused label \(L_{in}\) and creates an entry in its MPLS label switching table for channel sharing in \(\text{Relate}(e, \text{BR1})\).

- The created entry has
  - Ingress label \(L_{in}\)
  - Multiple egress interface-label pairs \((L_{out1}, L_{out2}), (L_{out2}, L_{out3}), \ldots\) from LSPs used to label switch packets of channels in \(\text{Relate}(e, \text{BR1})\) at BR1.
  - Action to swap \(L_{in}\) with \(L_{out1}, L_{out2}\), or \(L_{out}\) and send packet out on corresponding \(L_{out1}, L_{out2}\), or \(L_{out}\) ... 

**Step 3:** For each \(BR_{o}\) which is in the parent BRs set \(\{\text{parent_BR}(e, \text{BR1}); e \in \text{Relate}(e, \text{BR1})\}\),

- BR1 sets up a point to point LSP on the path from \(BR_{o}\) to \(BR1\) using a receiver-Initiate LSP setup mechanism such as that provided by RSVP with an appropriate traffic specification that includes all channels in the subset that should share the LSP \(\{e : e \in \text{Relate}(e, \text{BR1}) \text{ and parent_BR}(e, \text{BR1}) = BR_{o}\}\). The entry added in MPLS label switching table at \(BR1\)'s parent router for this LSP must have \(L_{in}\) as egress label. Therefore, packets of channels in \(\text{Relate}(e, \text{BR1})\) will arrive to \(BR1\) carrying the same label \(L_{in}\).

- \(BR_{o}\) uses this LSP to label switch packets of all related channels in the set \(\{e : e \in \text{Relate}(e, \text{BR1}) \text{ and parent_BR}(e, \text{BR1}) = BR_{o}\}\) to \(BR1\). Packets of the channels, which are in \(\text{Relate}(e, \text{BR1})\) and have \(BR_{o}\) as parent branching router, will then be label switched to \(BR1\) by the same LSP.

**Step 4:** BR1 removes forwarding states of channels in \(\text{Relate}(e, \text{BR1})\).

- A channel sharing record is added in \(BR1\) for \(\text{Relate}(e, \text{BR1})\) in the network layer.

- The channel sharing record maintains
  - Channel ID and the address of \text{parent_BR}(e, BR1) of each channel \(e\) in \(\text{Relate}(e, \text{BR1})\)
  - The next hop BRs(DRs) of channels in \(\text{Relate}(e, \text{BR1})\) at \(BR1\)
  - The entry of MPLS label switching table in \(BR1\) with \(L_{in}\) as ingress label

Figure 3 shows an example of channel sharing at router \(BR1\) for two related channels \((S1, G1)\) and \((S2, G2)\). Both channels are not shared with any other channel in \(BR2\) and \(BR3\). When the shared channels need to be separated, they are removed from the channel sharing record and their forwarding states are restored in the network layer based on the information in the channel sharing record.

IV. SIMULATION RESULTS

We perform simulations to evaluate the new scalable multicast solution with dynamic tunneling scheme, TMST multicast tree construction algorithm and channel sharing scheme. Simulations use randomly generated topology with 4000 routers and 5000 multicast channels with maximum of 1000 destinations in each channel.

In our simulations, we evaluate the scalability in the network layer through measuring the number of forwarding states and the consumed memory storage. The measured results are compared with those for MMT and for the network layer multicast scheme. Network layer multicast schemes, such as PIM [28] and CBT [29], maintains forwarding states in the network layer of all routers on multicast trees for multicast routing. In the computation of the storage, we assume an IP address takes 32 bits, a port (interface ID) takes 16 bits and a MPLS label takes 20 bits. We consider both shortest path tree (Figure 4) and Steiner tree cases (Figure 5) under different TMST matching rates.

Under shortest path trees case, our new solution achieves 60%~90% reduction in the number of forwarding states over MMT, 87.5%~97% over network layer multicast scheme. In terms of consumed storage in the network layer, our solution achieves 9%~26% reduction over MMT, and 47.5%~57% reduction over network layer multicast scheme. Because the shortest path trees are not optimal in tree cost, the TMST multicast trees have their tree cost 3% less than that of shortest path trees.

Under Steiner tree case, the new solution achieves 45%~84% reduction in the number of forwarding states over MMT, and 84.7%~95.5% over network layer multicast scheme. In terms of consumed storage in the network layer, our solution achieves 2%~5% reduction over MMT, and 51%~54% reduction over network layer multicast scheme. The TMST multicast trees are not optimal in tree cost and have 16% more tree cost over Steiner trees. Such overhead should be evaluated by the network operator to decide if the extra tree cost is affordable to improve scalability.

In the simulations of our new solution, an average of 46~66 MPLS labels are used in every router in the network for multicast routing. Compared with 540~580 MPLS labels used...
in the framework proposed in [29], in which multicast routing is realized entirely in the MPLS layer without any forwarding state maintained in the network layer, our solution achieves about 88%–92% reduction in the number of consumed MPLS labels for multicast routing. Therefore, our solution balances the scalability in the MPLS layer and the network layer through tunneling and channel sharing. By controlling the channel sharing criteria, which is described in Section IIIC, step 2, network operators can adjust the scalability burden between the MPLS layer and the network layer based on different network topologies.

![Figure 4 SPT Simulation Results](image)

![Figure 5 Steiner Tree Simulation Results](image)

V. CONCLUSION

In this paper, we propose a novel scalable multicast solution in MPLS networks. In the new solution, forwarding states are removed from non-branching routers on multicast trees and packets are label switched by static or dynamic MPLS LSPs in non-branching routers. We propose a new multicast tree construction algorithm and a new channel sharing scheme to further reduce the number of forwarding states in branching routers. Simulation results show that our new solution can achieve 85%–97% reduction in the number of multicast forwarding states in MPLS networks over network layer multicast and 45%–90% reduction over existing scalable multicast scheme in MPLS networks. Our solution consumes 90% less MPLS labels in multicast routing than pure MPLS layer multicast routing. The new solution provides network operators the capability to adjust the scalability burden between MPLS layer and network layer by controlling the channel sharing criteria.

REFERENCES