Multicast Routers Cooperating with Channel Announcement System

Hitoshi Asaeda
INRIA
Project PLANETE
2004, Route des Lucioles, BP 93,
06902 Sophia Antipolis, France
Hitoshi.Asaeda@sophia.inria.fr

Walid Dabbous
INRIA
Project PLANETE
2004, Route des Lucioles, BP 93,
06902 Sophia Antipolis, France
Walid.Dabbous@sophia.inria.fr

Abstract

Source-Specific Multicast (SSM) has been discussed as realistic multicast communication architecture used in the Internet. Due to the SSM conformance, source discovery procedure is eliminated from multicast routing protocol, and the protocol scalability is improved. However, because of no source address validation mechanism on the router side, multicast router cannot recognize un-trustworthy join requests and cannot avoid invalid or unavailable routing tree construction.

In this document, after studying the problems of SSM deployment, we propose a new communication model between multicast routers and multicast session directory systems. As a candidate of the session directory system, we propose to enhance Channel Reflector, which is a multicast channel announcement system maintaining (S,G) channel information for end users. Since it provides an effective policy and scope management, multicast routers cooperating with this system can validate (S,G) join and optionally translate (,G) join to appropriate (S,G) joins.*

1. Introduction

1.1. Multicast Service Model and Routing Architecture

Traditional IP multicast routing protocols have been focusing many-to-many communication model. Unfortunately, this service model and protocol architecture, called Any-Source Multicast (ASM), has run into significant barriers for the wide-scale deployment. Mainly, these barriers are rooted at problems to build efficient multicast routing trees for dynamic group memberships. More precisely, explicit join type multicast routing protocols, e.g., PIM-SM [1], use a core router (Rendezvous Point) and maintain "core-rooted" multicast routing tree (Shared Tree), in order to find available source addresses using corresponding multicast addresses. They require complex routing algorithms to construct and maintain multicast routing tree including the mechanism to switch to optimized "source-rooted" routing tree (Shortest Path Tree (SPT)).

However, considering live streaming or contents distribution style multicast applications used in the Internet, one-to-many or few-to-many communication is mostly sufficient. Within this communication model, once an end-node requests to receive data, it can tell both of source address(es) and group address to an upstream router as group membership information. In this communication model called Source-Specific Multicast (SSM) [2], multicast routers do not need to discover the source addresses, and they can construct "source-rooted" multicast routing tree with no Shared Tree coordination. Multicast scalability problems are enormously reduced by eliminating a core router in the network, therefore an SSM is recognized as a reasonable architecture for most Internet multicast applications, and it provides feasible Inter-domain multicast routing.

While an SSM requires to make a join or leave process with specifying the pair of interesting source address(es) and multicast address as known as "(S,G) multicast channel(s)", IGMPv3 [3] implementation for IPv4 and MLDv2 [4] implementation for IPv6 are indispensable to both host and router sides. With regard to these protocol specifications, an end-node also needs to specify "filter mode", which is either `INCLUDE` or `EXCLUDE`, for each request. In `INCLUDE` mode, reception of packets sent to the specified multicast address is requested only from those IP source addresses listed in the source-list parameter, and in `EXCLUDE` mode, reception of packets sent to the given multicast address is requested from all IP source addresses except those listed in the source-list parameter. Based on the semantics of an SSM, only the request of `INCLUDE` filter mode prompts SSM communication, and traditional "non-source-specific join/leave, expressed by "(*,G) join/leave",

Proceedings of the 2004 International Symposium on Applications and the Internet (SAINT'04)
0-7695-2068-5/04 $20.00 © 2004 IEEE

are handled as a join/leave of EXCLUDE filter mode with null source address list.

1.2. Problem Statements for SSM Deployment

An SSM can provide promising approach to deploy realistic multicast services in the Internet. However, precisely because there is no channel validation system in a router side, Shortest Path Tree coordination triggered by an SSM capable end-node may bring security problems. For example, PIM-SM initially constructs a Shared Tree in order to find available sources for requested multicast address, and switches to each SPT for active sources. This implies PIM-SM router working for an ASM model does not voluntarily construct a *non-active* SPT. On the contrary, an SSM capable router constructs an SPT with no Shared Tree coordination. Thus, even if an end-node triggers invalid or unavailable (S,G) joins, the upstream router starts establishing all SPTs with no intellectual decision. By using some timer mechanism to monitor the data flow, it would be possible to prune unavailable (S,G) entries from the routing table. But it is neither a great deal of the solution for tens of thousands of bogus requests.

As another aspect, because of the scalability and manageability reason, we can expect that some Internet Service Providers (ISPs) decide the policy that an ASM is not allowed but only SSM communication is permitted in their networks. In this case, as a customer-side issue, both of an application and a kernel on an end-node must support IGMPv3/MLDv2. In fact, some applications and kernels today have fully or partially implemented IGMPv3 and MLDv2 [5, 6]. Nevertheless, if ISPs must wait to start SSM services until *every* end-node is ready for *everything*, they cannot begin their attractive services for a while.

From a routing protocol's point of view, in IPv4 environment, PIM-SM working with MSDP [7] is recognized as a current Inter-domain multicast routing protocol. PIM-SSM is a subset of PIM-SM and works for an SSM address range [8]. Based on the SSM specification [2], if PIM-SM or PIM-SSM router receives (*,G) join/leave[1] whose group address range is in an SSM range, it discards the request. According to this rule, a non-SSM capable node cannot trigger an SSM join/leave to the upstream router. However, in mixed environment that SSM capable nodes and non-upgraded nodes coexist on a LAN, although non-upgraded nodes cannot trigger any join whose group address range is in an SSM range, it can receive the data after another SSM capable node triggers (S,G) join using the same multicast address. This is because the upstream router discards (*,G) join using an SSM address range and can warn it, but it does not suppress flowing the data to any node. This situation makes a general user confuse and increases an administrator's work.

[1] Hereafter, (*,G) join/leave and (S,G) join/leave mean IGMP or MLD join/leave. For other join/leave, e.g., PIM (*,G) join, it will be mentioned precisely.

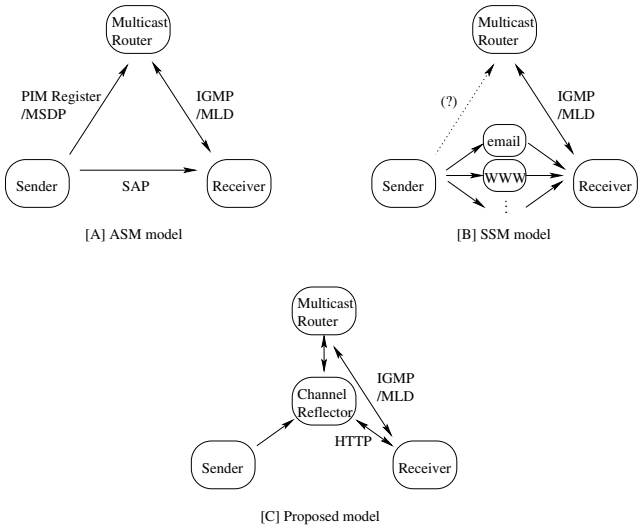

Figure 1. Cooperations of multicast data sender, receiver and router

2. SSM Communication Model

The essential multicast communication is formed by a data sender, a receiver and a multicast router. Let's see Figure 1. As stated in Section 1.2, a PIM-SM router working for an ASM model does not voluntarily construct a non-active Shortest Path Tree. The core router (Rendezvous Point) can recognize active data senders by PIM Register or MSDP messages and other routers indirectly discover these source addresses (Figure 1 [A]). On the contrary, in an SSM model, a data sender and a multicast router do not communicate each other (Figure 1 [B]). For a communication between a multicast router and a data receiver, IGMP and MLD work for both of an ASM and an SSM. With regard to an announcement of all available multicast addresses to whole networks, an ASM uses traditional Session Announcement Protocol (SAP) [17]. SAP is a sender-initiate protocol, and a data receiver can resolve (S,G) address pairs when he receives SAP messages. However, because every data source must be a source of a SAP announcement group for this mechanism, SAP cannot work in SSM-only network. Thus in SSM environment, (S,G) channel information should be announced by email, WWW, and so on.

Here we summarize the motivation of our study and try to find key components to make the SSM deployment advance.

- Source address discovery and validation
  Active (S,G) information is not flooded to all routers. This is the fundamental concept of explicit join type multicast routing protocols. There is no core router to notify available source address within an SSM network. This provides a remarkable benefit to SSM communication architecture. However, in order to admit correct (S,G) joins, it is indispensable for a multicast router to validate available (S,G) entries. Yet, the source address discovery mechanism should not be embedded in multicast routing protocols because of the scalability reason. Effective solution would be encouraged.

- Receiver address validation
  If a multicast router receives unavailable joins from downstream nodes, the router should discard these invalid joins no matter whether they are intentionally or accidentally requested. Although IETF MSEC WG [9] has been trying to standardize frameworks for securing group communication over the Internet, current IGMPv3 and MLDv2 do not have a standard mechanism to validate requested joins. Therefore, some mechanism to recognize valid join requests must be encouraged before new standard securing mechanism is embedded in each protocol.

- ASM-to-SSM translation
  There are many Internet users whose operating systems do not support SSM. They do not have any chance to legally receive SSM channel data. But if the requested group address is in an SSM address range and if the first-hop router knows corresponding data sender's address for the group address beforehand, it would be possible to translate (*,G) join to one or more appropriate (S*n*,G) joins, and any kinds of attached end-nodes can request SSM communication and the router can construct each SPT without Shared Tree coordination.

3. Channel Validation Mechanism for Multicast Routers

3.1. New Communication Model

Regarding all of concerning points stated in above section, we propose a new communication model; "multicast routers cooperate with a multicast channel announcement system". The fundamental idea comes from the viewpoint that every available and valid (S,G) entry is registered to a multicast session announcement system by the data sender or network administrator, and the announcement system maintaining the information of each accurate source and multicast address pair can give the information to nodes, including multicast routers. Once the multicast router consults the channel information to the session announcement system, it can validate the source and multicast address pair whenever it receives (S,G) join.

As such well-designed session announcement system, Channel Reflector (CR) [10] becomes the suitable candidate for our demands. It has been recently proposed for an alternative multicast session directory system used in SSM communication environment. CR is the concrete architecture on top of the Internet Media Guide (IMG) framework [11] specialized in the IETF MMUSIC WG. It supports not only an available or scheduled channel information distribution to an end-node but an effective management policy and scoping technique as the Inter-domain channel announcement system. In this document, we enhance the functions of Channel Reflector for multicast routers to discover available and valid (S,G) entries through Channel Reflector.

3.2. Channel Reflector

In this section, we pick up the important functions of Channel Reflector and briefly explain its basic components.

While Session Announcement Protocol has been used in ASM environment, it has no user authentication inside the model since it makes periodic session announcement to all potential recipients. CR is a session directory system collaborative with WWW schema, thus it would also great help for receiver's validation. For instance, CR shows available channel information only to the accepted receivers by having an Access-Control-List (ACL) as a regular WWW server.

In order to provide policy and scope control features, CR's architecture is composed of *"primary CR"* and *"site CRs"*. A primary CR maintains only the globally available multicast channels. A site CR maintains the globally available multicast channels and multicast channels available in the associated domain. It is assumed that at least one site CR exists in each administrative controlled domain. This controlled domain forms the scoping area where each channel information belongs to.

Each CR has a *parent-and-child* relation and forms a "policy tree" (Figure 2). A primary CR is the root of the policy tree and has no parent CR. Each site CR has one parent CR which is either the primary CR or another site CR and has zero, one or more child CRs. Each parent-and-child relation is configured statically. This relation is a *hard-state* connection, meaning that in the absence of any major update event requiring a tree update (e.g. a CR failure), so the parent-and-child state will remain unchanged for an unbounded period.

CR offers effective *"domain name base"* scope mechanism. While the primary CR belongs to a "global scope"

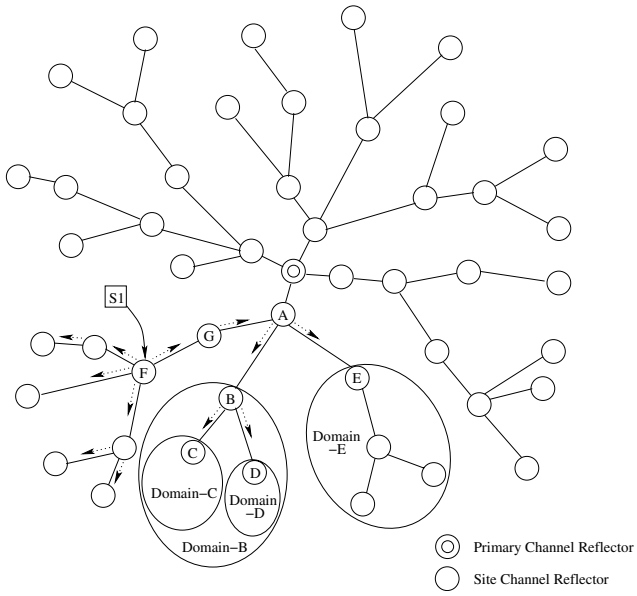

Figure 2. Policy tree of Channel Reflectors

(i.e., world-wide scope region), each site CR belongs to two scopes: (1) its fully qualified domain name (FQDN) (e.g. cr.example.com), and (2) its second-level domain name (e.g. example.com). Both of the FQDN and the domain name become *"scope labels"*. Every multicast channel entry has at least one associated scope label, and it is registered on the CRs which have the corresponding scope labels, and on their child CRs according to the scope management condition.

Let's see an example of CR's mechanism. Although CR's topology does not rely on any geographical condition, it usually adheres to Autonomous System (AS) and other network topology boundaries. For instance, in Figure 2, the source S1 registers the channel entry to his site CR (CR-F). Here we assume S1 specifies the channel scope label to CR-A's FQDN. This channel information is transferred to the next-hop CRs, which are its parent CR (CR-G) and child CRs. This advertisement process continues within the associated scope area hop-by-hop. Since CR-A is the scope boundary for S1's channel, CR-A forwards the channel information to CR-B and CR-E (i.e. its child CRs except this information forwarder), but does not advertise it to its parent CR (primary CR). In this example, both of Domain-C and Domain-D inherit CR-A's and B's policy definitions, on top of which CR-C and CR-D add their own local policy definitions. Because of the CR-E's policy, if CR-E filters out S1's channel information, then this channel entry will appear neither on CR-E's channel list nor on any child CRs' channel list in Domain-E. One reason for filtering the S1's channel is for instance when the data stream plans to consume a bandwidth larger than the maximum threshold permitted by CR-E's administrator.

3.3. Communication with Channel Announcement System

Figure 1 [C] shows the proposed communication model of a data sender, a receiver and a multicast router. CR maintains all available channel information based on its own policy and scoping technique. As well as a DNS server's address known by every end node, an address of a site CR is configured on the associated multicast routers by the site-local administrator beforehand. After that, these routers can validate channel senders by accessing the CR.

This communication model requires that CR is deployed in the Internet and multicast router supports the extension to access assigned CR; nevertheless multicast routing protocol itself is not changed, and multicast router can well manage an administrative scope safer than the current situation [12].

Let's see Figure 3. When a company or an organization starts or schedules the multicast data transmission, the data sender or the site administrator registers the channel information to the site CR by one of many possible means: a modified *sdr*, a CGI, an email, etc. Due to the availability and performance advantage, multiple redundant CRs are the necessary components for each CR as WWW and DNS prepare mirroring server and secondary server respectively.

After a multicast router gets available channel information from the site CR, it starts maintaining the Shortest Path Tree. On this condition, the administrators can instruct their site-local policy to multicast routers through CR. For instance, if the administrators want to prohibit for downstream nodes to join some channels, they can set up these channels as invalid channels through the site CR.

This communication model also greatly helps avoiding several multicast DoS attacks we have recently encountered [13, 14]. In our new infrastructure, invalid multicast channels are not registered by the administrator, therefore even if the first-hop router receives bogus data packets, it can easily discard them and does not forward to any network.

4. Configuration of Multicast Router and Channel Reflector

4.1. XML Formatted Information

Considering the deployment, in order to avoid additional burden of yet another round of modifications to the existing infrastructure, interoperation with Channel Reflector and multicast router must be simple and independent on the routing protocols. In addition, since each CR may need to keep a large number of channel entries, easy policy configuration and management would be encouraged. According to

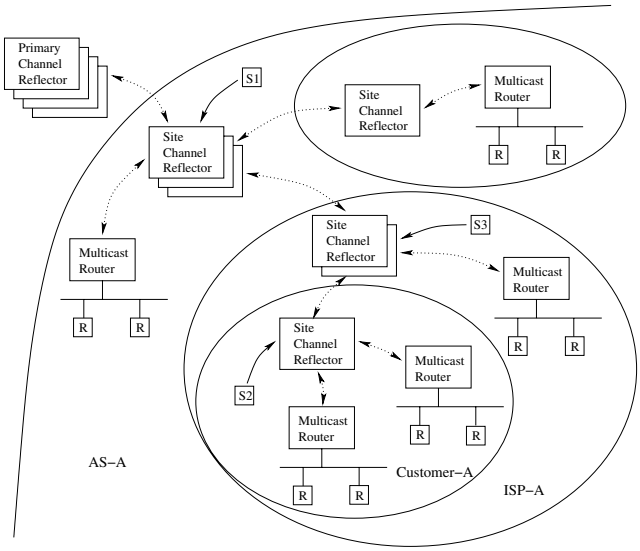

Figure 3. Channel validation by Channel Reflectors and multicast routers

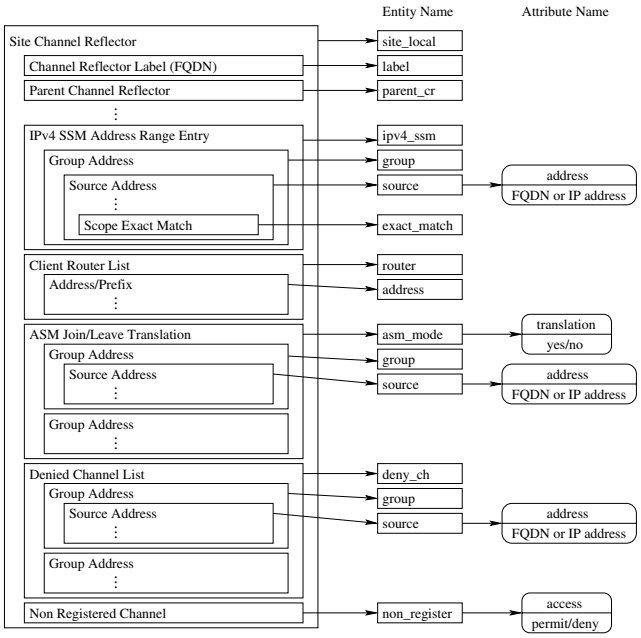

Figure 4. Configurations of a site Channel Reflector to cooperate with multicast routers

the channel management concept, each CR announces independent channels based on each site-local policy and scope configurations.

The CR approach basically drives a Session Description Protocol (SDP) syntax [15] and inherits all of its keywords. Underpinning this concept, SDPng [16] has the advantage of making use of the Extensible Markup Language (XML), which is an emergent set of open standards and gains widespread support. In our context, XML can be efficiently used to describe the (S,G) entries and policy descriptions. Furthermore, since XML has an excellent affinity to HTTP and HTTP would be commonly available in every equipment on the Internet, we propose that each multicast router communicate with the site CR by using SOAP [18] over HTTP.

4.2. Site-local Policy Configuration

In addition to globally available channel information, a site CR can maintain additional site-local channel information and policies defined by each administrator. While all basic entries that CR proposes for channel information description and its local configuration (e.g., scope labels, accepted receivers list) are explained in [10], several new entities should be defined to CR in order to cooperate with multicast router (Figure 4).

- Client Router List
  Each site CR prepares the lists of client routers and child CRs. This "Client Router List" entry quotes permitted router IP addresses with network prefixes.

- ASM Join/Leave Translation
  As described in Section 1.2, a non-SSM capable node cannot trigger SSM join/leave to the upstream router. But if we use Channel Reflector, (*,G) join can be translated to appropriate (S,G) join requests, and a non-upgraded node can be supported in SSM network. This behavior is decided by "ASM Join/Leave Translation" entry.
  If "translation" attribute of this entry is "yes" and;

  – if "Group Address" entry is empty, every (*,G) join/leave is translated to every registered (S,G) joins.

  – if some multicast addresses are specified in "Group Address" entry, only these multicast groups are translated.
  If source addresses are also specified in "Source Address" entries, only these multicast channels are translated.

  Due to some security reason, administrators may want to require explicit (S,G) join for some channel. "translation" attribute with "no" value configures non-translated channels. This indicates that the specified channel will not be included in the translated join. Therefore, if "translation" attribute of this entry is "no" and;

– if "Group Address" entry is empty, (*,G) joins are not translated to any (S,G) joins.

– if some multicast addresses are specified in "Group Address" entry, these multicast groups are not translated.
If source addresses are also specified in "Source Address" entries, all registered multicast channels except these multicast channels are translated.

For example, if (S1,G1) is configured in this list and "translation" attribute is "no", then downstream (*,G1) join will translate to (S*n*,G1) joins except (S1,G1) join. But remember it does not prohibit that a router constructs an SPT for (S1,G1) when it receives (S1,G1) join. If an administrator wants to discard a join for some channel, the configuration of following "Denied Channel Lists" is required.
If there is no corresponding channel information on the CR, (*,G) join for an SSM address range will be ignored based on the SSM specification [2] and (*,G) join for a non-SSM address range will be processed as an ASM request by routers, no matter whether "ASM Join/Leave Translation" is configured or not.

- Denied Channel List
A site CR can input unacceptable channel lists to routers by using "Denied Channel Lists". If an end-node joins to the corresponding channel entry on this list, the router just ignores the join message and does not create the multicast routing tree.

- Non Registered Channel
From now on, since a CR provides accurate channel lists, non-registered (S,G) address pair might be invalid. By using this entity, an administrator also can set up whether each multicast router accepts non-registered (S,G) join or not.

4.3. Router Configuration

A router configuration is pretty simple. It only needs to specify its site CR addresses and cache expiration period. Different CR addresses can be specified for each interface when the administrator wants to distinguish the policy and scope per interface. Regarding the cache expiration period, although multicast router would be able to access its parent CR per client's join request with no cache, the administrator must know this frequent reference may burden network and hardware resources. This consideration is mentioned in Section 6.

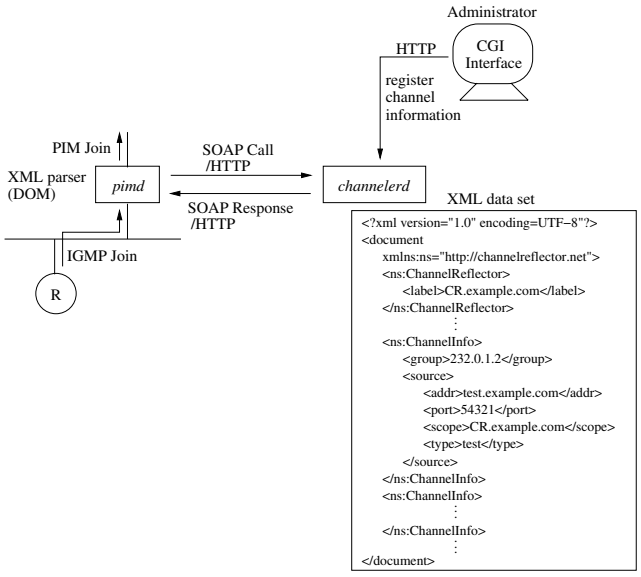

Figure 5. *channelerd* **and** *pimd* **on our test network**

5. Experiences

We have implemented a prototype Channel Reflector, *channelerd*. It works as a regular HTTP server and keeps multicast channel information which the site administrator registers via its CGI interface. On the other hand, it can reply available (S,G) address pair to its client request with SOAP message encoding. Figure 5 shows a part of our test network.

In our test network, USC *pimd* [19], which is a PIM-SM multicast router daemon, has been modified to work as a PIM-SSM router and a *channelerd*'s client. An XML parser, Document Object Model (DOM), is also inside our *pimd*. When this routing engine receives IGMP join message, it sends SOAP Call message with a target multicast address to *channelerd*. After that, it gets available (S,G) address pair(s) by the response of SOAP Response message.

6. Conclusions and Future Works

After studying current multicast routing protocols and the SSM deployment problems, we propose a new communication model; multicast routers cooperate with a multicast channel announcement system. The advantage of this approach is not only manageable and feasible to construct appropriate routing trees, but potential to introduce SSM multicast services to all end users. Since Channel Reflector would be designed to be used by both of an end-node and a router, it is also an advantage that SSM channel informa-

tion and administrative policy can be integrated to the single configuration semantics using standardized SDPng and XML.

Our future work will address the deployment aspects. One open issue is related to the performance consideration for the proposed communication model. The key point behind this discussion is that channel information may be changed frequently. Here there is a trade-off between *scalability* and *preciseness*. If channel information is cached within some defined period, e.g., for one minute, although router's scalability would be increase, there may be some information inconsistency, and it may not be perfect for channel validation purpose. If multicast router accesses Channel Reflector per every join request, such inconsistency would not be happened, but both of routers and Channel Reflector may be burdened with a large number of information exchange. Our prototype implementation makes us confirm the basic components with access-per-request basis, but we would measure the scalability with additional simulations.

Other considerations are related to multicast router's behavior based on "ASM Join/Leave Translation". First point is that non-SSM capable group members cannot reply IGMPv3/MLDv2 Group-and-Source-Specific Query messages after (*,G) join has been translated to (S,G) join(s). Second point is that the join translation and leave translation would be inconsistent because of the time lag as stated above. Although there are several possible ideas to take into account these points, we would like to examine detail situation in our test network or wider networks and evaluate what kinds of solutions would be appropriate.

References

[1] B. Fenner, M. Handley, H. Holbrook and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", Internet Draft - work in progress, October 2003.

[2] H. Holbrook and B. Cain, "Source-Specific Multicast for IP", Internet Draft - work in progress, October 2003.

[3] B. Cain, S. Deering, I. Kouvelas, B. Fenner and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC3376, October 2002.

[4] R. Vida and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", Internet Draft - work in progress, June 2003.

[5] "SSM-capable DVTS", <http://www-sop.inria.fr/planete/Hitoshi.Asaeda/dvts>

[6] H. Asaeda and S. Suzuki, "MLDv2 Protocol Design, Implementation and Evaluation for Source-Specific Multicast over IPv6", Proceedings in SAINT 2003 Workshops, pp.244-249, January 2003.

[7] B. Fenner and D. Meyer, "Multicast Source Discovery Protocol (MSDP)", RFC3618, October 2003.

[8] Z. Albanna, K. Almeroth, D. Meyer and M. Schipper, "IANA Guidelines for IPv4 Multicast Address Assignments", Internet Draft - work in progress, March 2002.

[9] "Multicast Security (msec) Charter", <http://www.ietf.org/html.charters/msec-charter.html>

[10] H. Asaeda and V. Roca, "Consideration of Multicast Channel Annoucement Architecture", INRIA Research Report, RR-4762, March 2003.

[11] Y. Nomura, R. Walsh, H. Asaeda and H. Schulzrinne, "A Framework for the Usage of Internet Media Guides", Internet Draft - work in progress, October 2003.

[12] D. Mayer, "Administratively scoped IP multicast", RFC2365, July 1998.

[13] P. Rajvaidya, K. Ramachandran and K. Almeroth, "Detection and Deflection of DoS Attacks Against the Multicast Source Discovery Protocol", UCSB Technical Report, July 2002.

[14] "Sapphire Worm", <http://www.nmsl.cs.ucsb.edu/mantra/ries/sapphire>

[15] M. Handley and V. Jacobson, "SDP: Session Description Protocol", RFC2327, April 1998.

[16] D. Kutscher, J, Ott and C. Bormann, "Session Description and Capability Negotiation", Internet Draft - work in progress, October 2003.

[17] M. Handley, C. Perkins and E. Whelan, "Session Announcement Protocol", RFC2974, October 2000.

[18] "Simple Object Access Protocol (SOAP) 1.1", <http://www.w3.org/TR/SOAP>

[19] "PIM-SM Implementation", <http://netweb.usc.edu/pim/pimd>