INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Consideration of Multicast Channel Announcement Architecture

Hitoshi Asaeda — Vincent Roca

N° 4762

March 2003

THÈME 1

ISSN 0249-6399 ISRN INRIA/RR--4762--FR+ENG

Consideration of Multicast Channel Announcement Architecture

Hitoshi Asaeda * , Vincent Roca †

Thème 1 — Réseaux et systèmes
Projet Planete

Rapport de recherche n° 4762 — March 2003 — 25 pages

Abstract: In this document, we propose a new multicast session directory system, "Channel Reflector". One of the goals of this system is to provide a feasible channel announcement mechanism for Source-Specific Multicast (SSM) environment without using traditional SAP advertisement. The main advantage of it is to provide a channel announcement policy and a data distribution area by effective scoping technique. Channel Reflector does not require any protocol change to an end user, therefore, the easy deployment is also one of the benefits.

The channel information including the policy configuration is written in XML format. Due to an XML's property, network administrators can flexibly prepare independent channel information, and the end user retrieves available channel information as a human-readable information transparently.

Key-words: multicast, SSM, channel directory, session announcement

* Hitoshi.Asaeda@sophia.inria.fr
† Vincent.Roca@inrialpes.fr

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Considération sur une architecture d'annonce de canal multicast

Résumé : Dans ce document, nous proposons un nouveau système d'annuaire de sessions multicasts, "Channel Reflector". Un des objectifs de ce système est de proposer un mécanisme d'annoncement de canal pour les environnements Multicast de Source Spécifique (SSM) sans utiliser les traditionnelles annonces SAP. Le principal avantage de ce mécanisme est de proposer une politique d'annonce de canal ainsi que d'une aire de distribution de données, grâce à une technique efficace de vérification de rayon d'action. Channel Reflector ne requiert pas de changement de protocole pour un utilisateur final, en conséquence de quoi un autre bénéfice est un déploiement simple.

Les informations des canaux, incluant la configuration des politiques, sont décrites en XML. Grâce à une propriété du XML, les administrateurs réseaux peuvent préparer des informations de manière flexible pour des canaux indépendants, et les utilisateurs finaux peuvent consulter les informations sur les canaux de manière transparente en accédant à une présentation compréhensible et formattée des données.

Mots-clés : multicast, SSM, annonce de session, annuaire, rayon d'action des données multicast

1 Introduction to Multicast Session Announcement over the Internet and Related Works

1.1 Current Session Announcement Techniques

Multicast communication techniques are highly beneficial for the large scale distribution of multimedia content: audio and video streaming, video-conferencing and distance-learning. Due to the multicast addressing architecture [1, 2], transient multicast addresses which are dynamically assigned are usually used for these sessions. These addresses only exist as long as some traffic is sent. This is a major difference with unicast addresses that are assigned to individual nodes for a long span of time. The direct consequence is that an end user who wants to join a multicast session must first resolve the transient multicast address used by the session he's interested in.

There are two multicast session discovery approaches: the *"invitation model"* and the *"announcement model"*. In the invitation model, a user is explicitly invited by another user to join an on-going session. Session Initiation Protocol (SIP) [3] provides mechanisms for such invitations, as well as for user location discovery, a negotiation of session parameters, and so on. Although this approach works well within a small domain, it is not suited to large multicast sessions, since the inviter must know the unicast addresses of all participants to be invited beforehand.

The session directory system *sdr* follows the announcement model. It has been intensively used in the Multicast Backbone (MBone) [4], the world-wide experimental multicast routing infrastructure. This directory system can advertise information for all current or future sessions to other directory systems, and can assist end users to select the data flows they want to receive.

The session information can be described using the Session Description Protocol (SDP) [5] in both invitation and announcement models. SDP describes valuable information including session name, session time, sender and multicast addresses, format of the media, etc., and the information becomes the key of joining and participating in the session.

But SDP does not specify how the information is transported. This is the role of the Session Announcement Protocol (SAP) [6]. When the multicast application starts sending session data, the sender announces its media-specific information to prospective participants using SAP. Therefore SAP is currently one of the necessary components of a session directory system following the announcement model. With SAP, an instance of a session directory system to which a sender registers a session periodically multicasts packets containing this session description to a well-known multicast group. These advertisements are received by other session directory instances so that each potential remote participant has enough information to join a session. Whenever a session is deleted or modified, SAP messages are multicast similarly to keep all the session directory instances synchronized. While this manner frees the users from the cumbersome task of specifying the application arguments, periodic session announcements lead to scalability problems when the number of sessions increases. Additionally, since the SAP announcement use the standard non reliable best-effort

UDP/IP semantics, improving SAP robustness in front of packet losses requires transmitting several times SAP announcement. Although this strategy keeps the protocol implementation simple, it creates additional transmission overhead and further reduces scalability.

1.2 Overview of Other Information Distribution Systems

Domain Name System (DNS) is undoubtedly a successful information distribution system of the current Internet. In this approach, a hierarchy of DNS servers keeps all the information, and each prospective client can consult them whenever it is required to obtain to desired information. From this point of view, DNS is a potential candidate to create a multicast session announcement system. Yet two reasons prevent its use:

- precisely because DNS is already largely deployed, it is difficult to change all DNS systems including the client resolver to support new record types;
- moreover the first-hop DNS server does not necessarily access the original database upon each client request since it caches information locally for a defined period.

Hence the DNS system cannot manage highly dynamic multicast services that are launched and stopped more frequently than its cache refresh period.

Other alternative ways of conveying session descriptions would include e-mail and the World Wide Web (WWW). Both applications are of wide use and are flexible enough to carry many kinds of information. Especially, the WWW can easily provide human-readable session information which is highly valuable for the end user. To provide a multicast announcement service, however, either approach requires that a central mail or WWW server be used. Remember that a data sender or a network administrator possibly defines a policy region called *scope* [7] to limit the data distribution area in order that only an end user which belongs to the region can receive the session data. Concerning point is that the announcements of multicast sessions simply made by such central servers do not indicate the property that the receiver cannot receive the session because the multicast sessions may be restricted in a scope, and reception of e-mail or access to the WWW server is possible outside this scope.

Internet Media Guides (IMGs) [8] has been recently proposed in IETF MMUSIC WG. IMGs allow users to initiate streaming media sessions, schedule delivery of downloadable or multicast content or listen to live multicast sessions. The fact that it relies on WWW collaboration schemes and that it creates a media directory system for the Internet is very interesting. Unfortunately, however, there is no concrete architecture yet and it does not provide any effective scoping technique or policy management whereas we consider these aspect as central to our own needs. Furthermore, since there is no realistic protocol design, it seems that SAP will be used continuously.

1.3 Realistic Scope Definition for Session Announcement

In the MBone, TTL scoping has been used to control the distribution of multicast traffic with the objective of limiting the stress on scarce resources (e.g. bandwidth), or to achieve

some kind of improved privacy or scaling properties. However, defining a scope using TTL is obsolete, since TTL scoping has proven to be difficult to implement reliably and the resulting schemes have often been complex and difficult to understand.

It has been recognized that administratively scoped IP multicast [7] can provide clear and simple semantics for scoped IP multicast, since;

- packets addressed to administratively scoped multicast addresses do not cross configured administrative boundaries, and
- administratively scoped multicast addresses are locally assigned, and hence are not required to be unique across administrative boundaries.

SAP has been focusing these demands, since SAP announcement is multicast with the same scope as the session it is announcing to ensure that the recipients of the announcement are within the scope of the session the announcement describes. However, this situation is about to change again, since using multicast addresses to define the scope would be weak and not so reasonable. Some network administrators need to distinguish service domains per session not only by the traditional scope boundary. For this purpose, configuring packet filters for independent boundaries to all multicast routers really increases administration cost. Furthermore, nowadays Source-Specific Multicast (SSM) [9] using IGMPv3 [12] and MLDv2 [13] is recognized as the most feasible multicast communication model for a wide use throughout the Internet, since it eliminates many complexities associated with the traditional Any-Source Multicast (ASM) routing protocols and communication model. But SAP cannot make a scope combining with a different address range, e.g., an administrative scope definition using SSM address range [10], even with MZAP [11]. And within an SSM capable network, the source address may also be used as a keyword to set up a scope, whereas SAP does not support this possibility.

1.4 Goals of This Work and Organization of the Paper

In this document, we propose a new multicast session directory system, called "Channel Reflector". One of its goals is to provide a channel announcement mechanism for an SSM environment without using traditional SAP advertisements, and to enable administrators to finely specify channel announcement policies and data distribution scoping.

The remainder of this article is organized as follows: section 2 details the features that a multicast session directory system should provide. Section 3 introduces the general architecture of our "Channel Reflector" proposal, and section 4 shows the underlying protocol. Configuration aspects are introduced in section 5. Finally section 6 concludes this paper.

2 Properties of a New Multicast Session Announcement Scheme

A new multicast session announcement mechanism should comply with the following key features:

- Scalability
  SAP makes periodic session announcement to all potential recipients. In this approach, if the target scope is world-wide region, all data senders' session information is flooded to whole Internet through many-to-many multicast communication topology. The ideal announcement scheme should be more effective. In general, a session information is only needed when the receiver decides to trigger a join, therefore using a receiver-driven query mechanism may be a better solution.

- Easy deployment
  A new session announcement mechanism must be adapted to any kind of end user and any kind of equipment. From this viewpoint, it should be minimized to change current client environment and protocols. Avoiding complexity should be also considered for the new mechanism.

- Flexible policy and scope controls
  Most network administrators would want to advertise only acceptable sessions to their users. The decision of the acceptable sessions may be resolved by the bandwidth, contents, sender's location, and so on. And if they have some local channels, they may not want to advertise their channels out of the associated service domain. These scope boundaries may not be simply defined by network address prefix, TTL, etc. A new approach is thus required to enable a flexible and precise configuration of policies and scopes.

Obviously, a session announcement mechanism is closely related with a session directory system, and it is important to consider the applicability to all users. The following properties should be respected:

- Easy use
  Since providing a well visualized directory system is an important item for easy use, collaborative WWW schemes may operate in the fashion. But we should take care that a non real-time WWW system does not cause any contradiction to a dynamic session announcement.

- Adaptation to the receiver's environment
  Users can be highly heterogeneous. Some of them using regular PCs, while others use lightweight mobile equipments. To cope with this heterogeneity, SDP [5] can use layered encoding schemes for session classification, but *sdr* will still show flat session information. The receivers need to select each appropriate session based on

their equipment and network environment. A better solution would be that users only obtain the appropriate session information when accessing the directory system transparently.

- Conformance to SSM-only network
  SAP communication assumes that the whole network implements the traditional many-to-many multicast communication model, i.e. ASM. This is required by the fact that every data sender becomes a source in the SAP announcement group. This is a major requirement and we can conceive that some Internet Service Providers only support the SSM model, since SSM is recognized as the most feasible multicast communication model[1]. Several operating systems are already SSM capable [14, 15, 16, 17, 18], and assuming an SSM-only network is realistic. In that case, using a well-managed directory system that does not rely on the ASM model is more realistic.

[1] For instance the IETF MAGMA WG currently assumes that Inter-domain IPv6 multicast routing relies on the SSM model.

3 Multicast Channel Advertisement Architecture

3.1 Concept

In order to comply with all the properties stated above we propose the "Channel Reflector" (CR) architecture. This is a directory system that advertises available multicast channel information, including the (S, G) pair (sender and multicast addresses), the media information, timing information, etc. End users can just access it like a regular WWW server.

Due to a policy control and a scope control of the multicast channel advertisements, the architecture is based on a hierarchy of "site CRs" rooted at a "primary CR". This primary CR gives the information associated to the global multicast channels, and each site CR gives the information of the channels available locally, in the associated domain. This focused domain conforms to an area where the channel information can be referred.

Each CR has a "parent-and-child" relation. A primary CR is maintained by some authorized or well-known site. Here, we assume it has been maintained by the server on "ChannelReflector.net". All channels globally available in the Internet should be registered in the server. This primary CR has no parent CR but has multiple site CRs as its child CRs.

Because of the policy and scope requirements, multiple site CRs are located in the Internet. A site CR has one single parent CR, which is either a primary CR or another site CR, and it may have one or more child CR(s).

Each parent-and-child relation is configured statically. This connection can be recognized as a *hard-state* model in which they do not need to be refreshed periodically. While the primary CR defines "global scopes", each site CR defines at least two independent scopes, the FQDN scope (e.g. cr.example.com), and the second-level domain name scope (e.g. example.com). Every multicast channel entry has an associated scope information (i.e. a scope label), and therefore is registered only on the CRs that provide the corresponding scope label.

For each end-node, a site CR is assigned by the site administrator beforehand. If an end-node wants to become a receiver, it consults an available channel information to the CR. If an end-node behaves as a data sender, instead of sending the multicast session information over SAP, it registers its channel entry to the assigned CR by using modified *sdr*, CGI, e-mail, etc. when it starts or schedules transferring the multicast data. Due to the security consideration, sender authorization/authentication mechanism would be enabled on the CR. As one of the simplest way, setting up the Access-Control-List (ACL) for valid senders is included in CR's basic component as described in Section 5.3.3, but additional strong secure mechanism, like using IP Authentication Header (AH) [19], would be encouraged.

After channel entry is registered correctly, it is transferred to its parent and child CRs, and each CR forwards it to the parent and child CRs over HTTP. This *hop-by-hop* data transfer is the fundamental premise driving the channel advertisement, since it is easy to maintain and authenticate parent-and-child relation, and it minimizes the number of data exchange. Of course, the most effective mechanism for such hop-by-hop data transfer is multicast. But data distribution tree is dependent on the scope label, therefore, for our

case, multicast is not suitable since maintaining and reconstructing a number of multicast routing trees based on each scope label may be complex. Furthermore, there may be a large number of site CRs in the Internet, and every CR has a possibility to become a channel entry sender. Rather than using ASM architecture for this situation, simple hop-by-hop data forwarding is appropriate in CR's environment.

3.2 Policy Control

The available channel information on a site CR is comprised of independent site-local channel entries and imported channel entries. As for independent site-local channel entries, since the site CR shows them on its directory but does not advertise them to other CRs, only an end-node accessing this CR can know these channels locally. On the other hands, the decision to select which channel entries are imported from other CRs is complied with the site-local policy configuration. This policy configuration is equivalent to a channel filtering mechanism, and many of these filtering keywords are given from SDP syntax. Each site CR obtains channel entries advertised by the parent CR. Each parent CR also has own policy configuration, and accepted channel entries are decided by the policy, therefore it results that the site CR inherits its parent CR's policy. And when these entries are sent to the site CR, an independent policy of the site CR selects and imports accepted channel entries. It finally results that the site CR shows the channel information only which is accepted by its parent CR and by itself.

Let's see Figure 1. There are one primary CR and multiple site CRs in the Internet. Each solid line shows a policy relation called "policy tree". This tree is established by the static configurations of the fully qualified domain names (FQDNs) of each CR.

Each CR does not have a restriction of geographical topology, but usually it follows Autonomous System (AS) or other network topology. In this topology map, both Domain-C and D inherit policy definitions, which are configured by CR-A and B respectively, and overlap each individual policy definition. Domain-E inherits CR-A's policy but does not have any policy relation with CR-B. In this case, if CR-A filters out some channel information, CR-B, C, D and E do not show it on their channel information directories.

As another story, if site-local administrators want to prohibit for the end-node to join some channel announced from the parent CR, they can hide it as an invisible channel. For instance, S1 registers the channel entry to its site CR. This channel entry is transferred to other CRs hop-by-hop within the scope area (detailed in a next section). But if this data stream plans to consume 1Mbps bandwidth and if CR-E's administrators configure that the bandwidth of acceptable sessions must be lower than 512kbps, S1's channel entry will not appear on CR-E's channel list and its child CRs, since CR-E discards the channel entry.

3.3 Scope Control

A CR introduces a new scoping architecture, which uses neither multicast address prefix nor TTL. Each CR retains at least two scope labels, which are the FQDN and the second-level domain name. Although other scope labels, e.g., a country name, Autonomous System

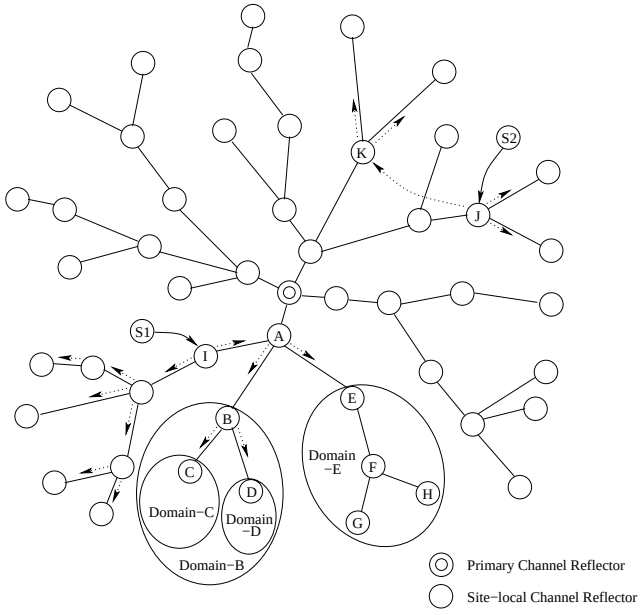

Figure 1: Sample topology of Channel Reflectors

(AS) number, may be possible to be used, it would require some additional procedures to synchronize or check a scope label duplication within a global scope (Section 6). Therefore, we've considered CR's FQDN and the second-level domain name as scope labels at this moment.

When a data sender or the administrator (hereafter, referred to as a registrant) registers the channel entry to the site CR, he/she can specify one or more target scope label(s) in order to limit the data distribution area. If the registrant specifies a second-level domain name as the scope, the channel entry is registered only on CRs which have a same second-level domain name and on their child CRs. If the scope label is an FQDN, the channel entry is registered either only on the CR or on the CR and its child CR(s). The behavior whether the target scope includes the child CRs or not is based on the registrant's decision, which is expressed by "exact match" field in a channel announcement message explained in Section 5.3.5. These manners result that only an end-node accessing these corresponding CRs can belong to the scope and can know these channels. If the registrant wants to distribute the data to whole Internet, that target scope label would become "world-wide" which every CR implicitly belongs to.

Let's recall Figure 1. S1 registers the channel entry to its site CR, CR-I, with specifying the scope label with A's FQDN. In this case, S1's channel entry is notified to the parent CR and the child CRs and forwarded hop-by-hop, but after it reaches CR-A, CR-A forwards

the entry only to the child CRs not to the parent CR. This scoping mechanism is, of course, independent on a policy control stated above, therefore CR-E can stop forwarding the entry based on its configuration and CR-F, G and H do not receive the entry. And also, since CR's rule does not restrict any geographical network topology, it is possible for a registrant to make a *spot* announcement. Here, S2 registers its channel entry on CR-J with specifying the scope label used by CR-J and K. In this case, CR-J directly advertises the channel information to CR-K. As the result, only end-nodes located under each domain are able to know the channel entry. But we should remember this spot announcement requires several rules because of a security reason (Section 4.2).

3.4 Consideration of High Availability

When we consider an availability and a performance advantage of each CR, preparing multiple servers would be preferable. As a matter of fact, there is one negative situation on a single server. For each policy and scope inheritance, if a parent CR goes down, it's impossible for the child CR(s) to send and receive channel information and scope labels to/from other CRs. This is one of the characteristics of hop-by-hop data transfer model adapted by CR's communication architecture. This situation must be solved by some mechanism, but we do not want to bypass this trouble by using alternative parent CR discovery or selection mechanism, since such technique would be complex and non-scalable in general.

Taking a good balance for our goal, it is certainly sufficient to prepare multiple CRs something like mirror servers. It is natural that a WWW mirroring system as a regular HTTP server is simple for this kind of distribution style, and moreover it is quite easy to maintain the system. The mirroring servers can be configured as a single server as specified in Section 5.3.1. If the master server does not make any response, another mirroring server quickly take its role seamlessly.

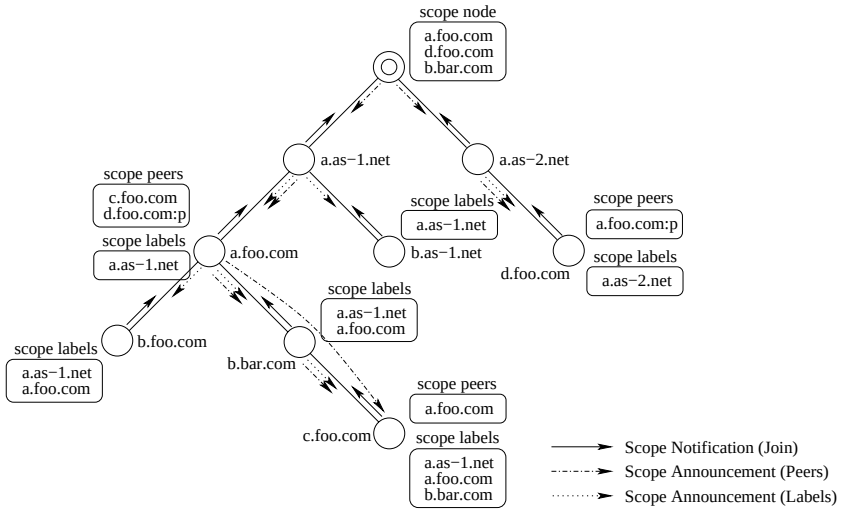

Figure 2: Scope management by Scope Notification and Scope Announcement messages

4 Multicast Channel Advertisement Protocol

Regarding well-managed data transmission between a parent CR and child CRs, scope label synchronization and channel information distribution are the main portion for each CR's role. Considering the transport protocol, since HTTP would be commonly available in every equipment on the Internet, we propose that each end-node and CR itself communicate with any CR by using HTTP. As described in Section 3.1, there is no special format for end user's request, but for the communication between each CR, new messages, which are embedded in an HTTP or SOAP [20] message, are required. For scope label synchronization, "Scope Notification" and "Scope Announcement" messages are newly defined. For channel information distribution, "Channel Announcement" message is used. Although these are talked in following sections, additional fields especially for security related messages would be included in near future.

4.1 Scope Label Synchronization

To control the multicast data distribution area using the scoping mechanism, synchronizing the scope label of all CRs is required. "Scope Notification" and "Scope Announcement" messages are used for scope label synchronization within a global scope. This is also done hop-by-hop, so each CR must not accept these messages coming from non-parent or non-child CR.

"Scope Notification" message is used to tell an FQDN to join or leave a policy tree. This message has a type field to specify "Join" and "Leave". "Scope Announcement" message is used to advertise scope peer nodes and scope labels with filling "Peers" or "Labels" type field.

Let's see Figure 2. When a site CR initially comes up, it sends Scope Notification Join message including the new FQDN to its parent CR. After the parent CR verifies that the message sender is its child CR, it sends back Scope Announcement Labels message including all scope labels kept by the parent CR. If the parent CR recognizes the second-level domain name of the FQDN is not same of the parent CR's domain name, it forwards the message to its parent CR. This message transfer is done hop-by-hop until it reaches a CR having the same second-level domain name or the primary CR.

If some CR having the same second-level domain name receives the Join message,

1. it registers the FQDN as its "scope peers",
2. it sends Scope Announcement Peers message including its own FQDN to the message originator, and
3. it stops forwarding the Scope Notification Join message.

And after the Scope Notification message originator receives the Scope Announcement Peers message, it registers the message sender's FQDN in its scope peers.

If there is no CR having the same second-level domain name along the path to the primary CR, the message finally reaches to the primary CR. When the primary CR receives the message from its valid child CRs,

1. it registers the FQDN as its "scope nodes",
2. if there are other scope nodes having same second-level domain name, it sends Scope Announcement Peers message including these FQDNs to the message originator and to these FQDN nodes, and
3. it stops forwarding the Scope Notification Join message.

And after these scope nodes receive the Scope Announcement Peers message, they register all announced FQDNs in their scope peers. This is occurred only when the message is not registered in any CR's scope peers. Because of these manners, it is supported that multiple CRs having a same second-level domain name are located in different policy trees.

In order to classify that Scope Announcement Peers message has been announced by the primary CR or other CRs, each FQDN in scope peers has a "primary bit" to indicate "it has been announced by the primary CR". This is used when some CR leaves from a policy tree or modifies its scope labels, FQDN, etc.

If some CR wants to change the parent CR, the CR and its child CRs first need to leave from current policy tree. To leave from the policy tree, the CR sends Scope Notification Leave message to all scope peers and its parent and child CRs.

Talking about CR's concept in Section 3.1, we mentioned that IP Authentication Header (AH) should be used for a multicast data registrant to register channel entry to its site CR.

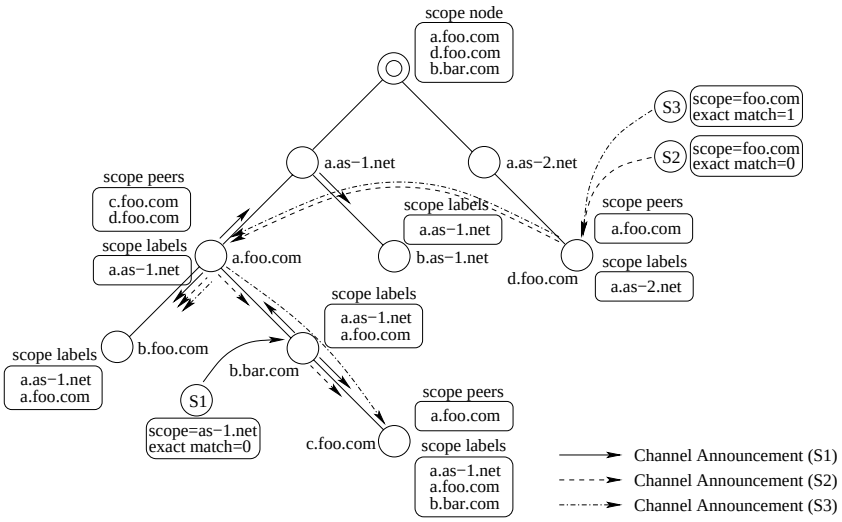

Figure 3: Channel information distribution by Channel Announcement message

As well, for channel entry and scope labels distributions done by each CR, it's now obvious that AH would be necessary.

4.2 Channel Information Distribution

Channel Announcement message is used for channel information distribution. This message is also transferred over HTTP or SOAP. The message format follows SDP syntax, but it additionally includes a origin field and a sequence number field. In the origin field, CR's FQDN is inserted. In the sequence number field, some unique number generated by each CR is inserted. It is required to check the uniquness of the message within a ring peer topology explained in a next section.

received channel entry may be unaccepted channel by the CR's policy configuration. In this case, if the scope is larger than the CR, the channel entry is only forwarded to the parent CR. Otherwise, it's just rejected.

4.3 Consideration of Inter-domain Channel Management

For the sake of availability and in order to better define new policies and scope regions, a network administrator may divide one single scope into multiple scopes. The way to divide a scope depends on the administrator's decision. New divided CRs may attach with a new parent-and-child relation. Or these CRs may not have any parent-and-child relation but these parent CR is the same CR. It's up to the situation. This kind of action does not lead

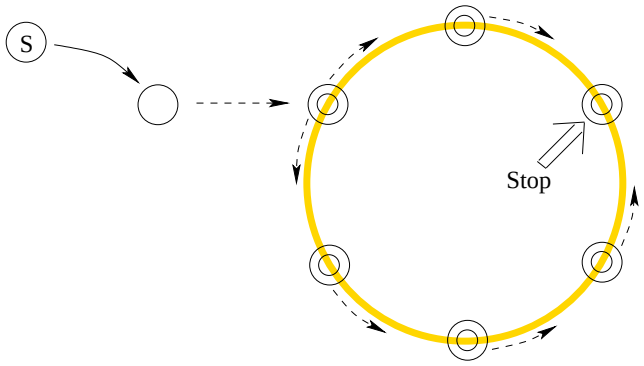

Figure 4: Data distribution over ring peer relation

any complex procedure. However, if we try to divide a primary CR as well, the condition is changed. In fact, it would relatively happen when CR's topology and its policy tree enormously grow up in the Internet. In this case, it may be impossible to handle all channel information and to control all scopes over the Internet by using a single primary CR. It would be indispensable to reform a world-wide topology map by multiple primary CRs. Each primary CR has a same role to provide all channel globally available in the Internet. Each one is registered as one of hosts of ChannelReflector.net. In this section, we talk about scalable channel advertisement for Inter-domain communication.

At the beginning, remember that an individual condition which has to be taken into account for multiple primary CRs environment is involved with a synchronization procedure of channel information and scope labels among divided scopes. This comes from same thought of regular policy control and scope control, but there is no parent-and-child relation for each primary CR. Every primary CR is on even ground. Now we propose making a peering relation between each primary CR. With establishing the peering connection statically, each primary CR exchanges its own world-wide scoped channel information with Channel Announcement message and preserving scope labels with Scope Announcement Peers message along the peering path. After receiving such information, the CR deals with it as native information which has been created independently.

How does it make each peering relation effectively? Information exchange using such peering connection might remind us of MSDP [21]. MSDP has been used with PIM-SM [22] in order to discover multicast data sender addresses outside of PIM domain. Unfortunately, however, MSDP is recognized as non-scalable protocol. One of the main reasons is MSDP needs to flood Source Active (SA) messages to all peer Rendezvous Points (RPs) whenever PIM-SM router receives multicast data. It has made several unwished troubles caused by unauthorized multicast data to every MSDP peer [23, 24]. Study of the MSDP DoS vulnerabilities makes sense of a critical need in well-designing and securing the multicast infrastructure. This input implies that the validation of data sender must be required.

Let's briefly talk about the specification of MSDP's peer connection, comparing with proposed design of peering relation for multiple primary CRs.

- No authentication mechanism
  This is the origin of above MSDP problems. As an ideal way, if multicast router receives unavailable joins which could be intentionally or accidentally requested from downstream nodes, it should discard these joins silently. Although IETF MSEC WG [25] has been trying to standardize protocols for securing group communication over the Internet, current IGMPv3 and MLDv2 do not have any particular mechanism to validate sender and group addresses of requested joins.
  But for CR's environment, channel registrant is authenticated by each site CR, and each site CR is authenticated by its parent CR as well. Therefore, if peering primary CRs can be authenticated, it results that the data coming from the peering CR has been validated.

- Mesh peer establishment
  MSDP conceptually configures mesh peers for each RP. Mesh peer has one benefit that data transfer can be done quickly, but since the data is flooded to all peer nodes with no intelligence, it wastes network bandwidth and may affect troubles to all peer nodes rapidly.
  For CR's environment, although quick data distribution is one of the important conditions, we propose to adapt another peering style, "two-ways ring peer" (Figure 4). On the ring peer topology map, each peer node has definitive two neighbor peer nodes on a logical flat ring. When the node needs to forward data coming from outside of the ring, it distributes the data to both neighbor peer nodes, and the data is transferred by the neighbor nodes hop-by-hop. If some peering node receives the data twice, then the node stops forwarding since it suggests the data has been gone through the ring and has been already received by every node on the ring. Although the original idea comes from the primary analysis of possible overlay topologies used by Host-Based Multicast (HBM) [26], we modified a regular ring topology, which can be called "one-way ring peer". Two-ways ring peer is suitable for CR's environment, since it is also realized hop-by-hop basis as well as the basic concept (Section 3.1) and effectively works on a flat topology[2].

Although each ring peer CR is based on a static configuration, increasing peering CR gives no problem. Showing concrete example would help understanding the situation. In Figure 5, there are several domains rooted at individual primary CRs. Initially, four primary CRs (AS-A, Region-A, Country-A and Domain-A) have made ring peers. After that, with preparing each primary CR, AS-B and C want to join the ring peer, and Domain-A want to divide itself to Domain-B and C. Due to these demands, after new primary CRs set up to connect to the ring peer, what they should do is just to make neighbor primary CRs change their neighbor peers. Their join does not affect anything to other primary CRs.

[2] Hereafter, "two-ways ring peer" is referred to as just "ring peer". "One-way ring peer" is precisely specified when needed.

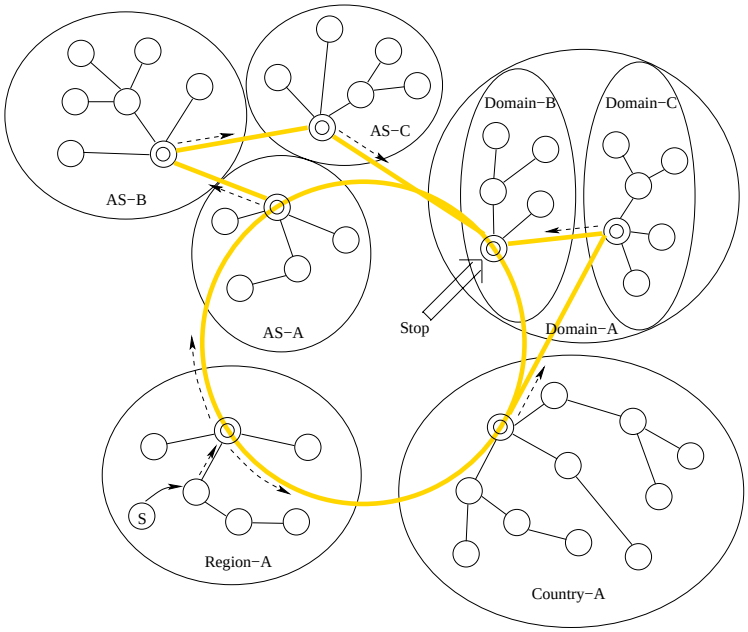

Figure 5: Divided world-wide scopes and peering relations with multiple primary CRs

5 Configuration of Channel Reflector

5.1 XML Formatted Information

A CR needs to keep a large number of channel entries. These entries are indivisual since they are dependent on the CR's policy and scope configurations. In addition, as well as making these configurations, the network administrators would need to configure accepted sources, permitted clients, etc., based on their demands. In these situations, at the view of the management cost, providing easy syntax for their configurations would be highly encouraged.

A CR basically drives SDP syntax and inherits all keywords. Underlying this concept, using SDPng [27] is also the possibility. This story implies that, as well as the channel information, any policy and scope configurations can be written in a same syntax whose definition is bound in well-known tag format. Now, we can say the eXtensible Markup Language (XML) is a suitable description language for such thought. XML is an emergent set of open standards and gains widespread support, therefore it is not only suitable to declare visible structures but also useful to manage and deliver multiple (S,G) entries mixed with multiple policies. And of course, since XML has a good affinity to HTTP server, from a viewpoint of deployment, it also has a remarkable benefit.

In this section, because an XML syntax to describe channel information itself is already written in SDPng document [27], we explain the fundamental concept and concrete entries which a CR specifically used.

5.2 Configuration of Channel Lists

A primary CR only provides a general configuration quoted by "Primary Channel Reflector" entry (Figure 6). This configuration is prepared by XML format (Figure 7). Although only IPv4 channel information is indicated in Figure 6, channels using other protocols, e.g., IPv6, can be inputted in a same manner.

This Primary Channel Reflector entry can be divided to "SSM Address Range Entry" and "Non-SSM Address Range Entry" per protocol. Each entry consists of same entries. "Receiver Lists" entry quotes data listeners' IP addresses with network prefixes or host names, which are permitted/denied to get any channel lists from this CR. The keyword, "permit" or "deny", must be mentioned as its attribute. If there is no access policy for a client, this Receiver Lists can be omitted. To specify channel lists, "Group Address" quotes several information. "The Number of Source Addresses" indicates the total number of "Source Address" entries including "AS Number", "Contact Person", and so on. This Group Address entry indicates available (S*n*,G) channel information itself and is heir to many keywords from SDP [5].

5.3 Site-local Policy Configuration

Globally available channel information is inherited from the primary CR to site CRs. And also, site CRs can maintain additional site-local channel information and policies defined by each administrator (Figure 8).

5.3.1 Parent and Child Channel Reflector

"Channel Reflector Address" entry in "Parent Channel Reflector" entry specifies a parent CR address. This entry can be specified multiple times due to the care of the availability. Each CR must be a pair of mirroring servers in order to keep a consistency of policy and scope configurations (Section 3.4). And also, this site CR must be permitted as a client by specified parent CR. For "Child Channel Reflector" entry, more than one child CR address can be specified.

"IPv4 Non-SSM Address Range" and "IPv6 Non-SSM Address Range" specify whether the channel lists of non-SSM address range should be imported or not. If these entries are omitted, the site CR implies to import *only* the channel lists of SSM address range. This means the default values of these attributes are "no".

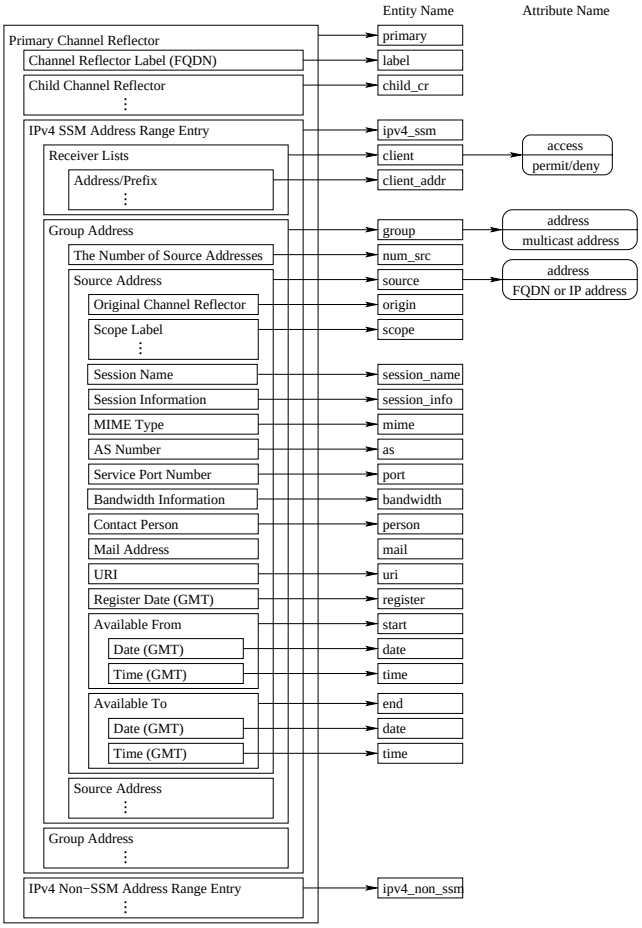

Figure 6: Structure of the primary Channel Reflector and defined entities

5.3.2 Receiver Lists

Each site CR prepares the lists of child CRs. This description is same of primary CR's Receiver Lists.

5.3.3 Sender Lists

Basically, sender selection mechanism on the primary CR would be unneeded since every node except its child CRs may be prohibited to access this directory system as a regular client. However, regarding site CRs, it should specify valid or invalid sender addresses or

```
<ipv4_ssm>
    <client access="deny">
        <client_addr>10.0.0.0/8</client_addr>
        <client_addr>172.16.0.0/12</client_addr>
        <client_addr>192.168.0.0/16</client_addr>
    </client>
    <group address="232.1.1.1">
        <num_src>1</num_src>
        <source address="server.example.com">
            <as>1111</as>
            <port>2222</port>
            <person>
                <given_name>Jane</given_name>
                <family_name>Doe</family_name>
            </person>
            <mail>j.doe@example.com</mail>
            <register>
                <year>2003</year>
                <month>12</month>
                <day>24</day>
            </register>
            <start>
                <year>2003</year>
                <month>12</month>
                <day>25</day>
                <hour>20</hour>
                <minute>00</minute>
            </start>
            <end>
                <year>2004</year>
                <month>01</month>
                <day>01</day>
                <hour>23</hour>
                <minute>59</minute>
            </end>
        </source>
    </group>
</ipv4_ssm>
```

Figure 7: Example of channel description using XML format

prefixes in "Sender Lists" entry. This is useful, for example, when a site-local administrator decides only a host belonging to the same domain can send multicast data to the site.

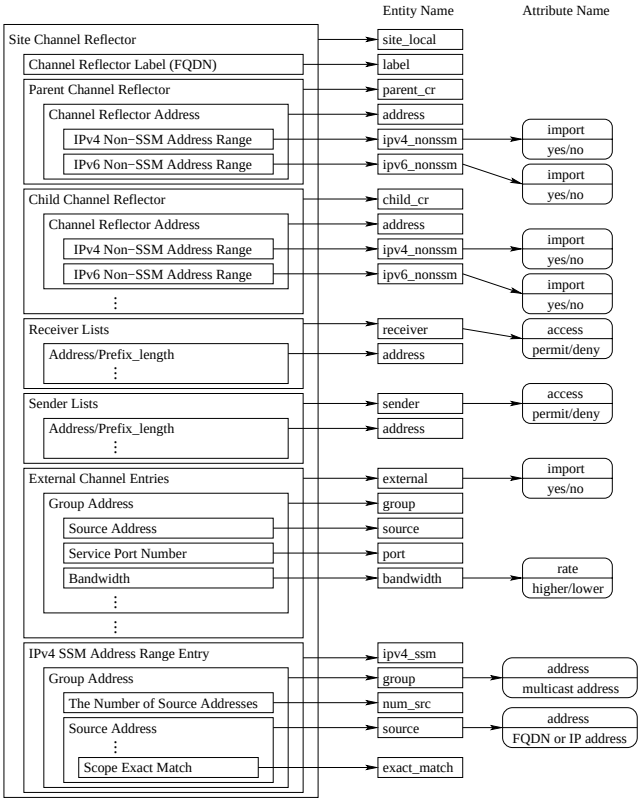

Figure 8: Structure of a site Channel Reflector and defined entities

5.3.4 External Channel Entry

This part is equivalent to a channel filtering configuration. A site-local network administrator can define what kinds of channel should be imported or not in this part.

5.3.5 Site-local Channel Lists

Site-local channels are distinguished from channels transferred from the parent CR, therefore it can provide scope mechanism. This description format follows the rule of primary CR's "Channel Lists". Only the difference is the existence of "exact match". This field directs whether the channel information should be announced only to the CR which has a corresponding scope label or not. If this field is not specified, the CR receiving the channel information sends to its child CRs. Otherwise, it does not forward to any CR.

6 Conclusion and Future Work

In this paper, we propose a new directory system, Channel Reflector. It introduces new policy management and scoping tequnique. The advantage of this approach is not only manageable, but also potential to introduce feasible multicast services to all end users.

In near future, to support strong secure mechanism, authentication key and encapsulation key distributions may be included in Scope Notification Join message. To confirm that Scope Notification message is valid message, embedding reasonable authentication mechanism would be one of our future works.

For the convenience, using other scope labels, e.g., a country or a region name, Autonomous System (AS) number, etc. might be encouraged. It would be able to be supported in this infrastructure, however, as an important feature for it, Channel Reflector must provide additional function to synchronize a scope label within a global scope, since there may be duplicate scope labels which have no relation. The function may cause an additional complexity, therefore further expectation would be required.

As one of interoperability issues, cooperation with SAP can be imagined. However, if a CR needs to have interoperability with SAP and to provide a function to import session entries flooded by SAP announcement, the CR must behave as a SAP client. According to this situation, since each scope concept is different and there is no consistency, we do not offer any function to support it. In fact, if a CR takes any information to all SAP clients, it gives us a contradict situation a CR should not take. We will think about this situation more.

Regarding a performance evaluation, we need to show our simulation result. At this moment, we have finished developing the simulation code, therefore, next step would be for the analysis of the result and for the protocol improvement.

References

[1] S. Deering, "Host Extensions for IP Multicasting", RFC1112, August 1989.

[2] R. Hinden and S. Deering, "IP Version 6 Addressing Architecture", RFC2373, July 1998.

[3] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley and E. Schooler, "SIP: Session Initiation Protocol", RFC3261, June 2002.

[4] S. Casner and S. Deering, "First IETF Internet Audiocast", ACM SIGCOMM Computer Communications Review, pp.92-97, July 1992.

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

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

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

[8] Y. Nomura and H. Schulzrinne, "A Framework for Internet Media Guides", Internet Draft - work in progress, February 2003.

[9] H. Holbrook and B. Cain, "Source-Specific Multicast for IP", Internet Draft - work in progress, November 2001.

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

[11] M. Handley, D. Thaler and R. Kermode, "Multicast-Scope Zone Announcement Protocol (MZAP)", RFC2776, February 2000.

[12] B. Cain et al., "Internet Group Management Protocol, Version 3", RFC3376, May 2002

[13] R. Vida et al., "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", Internet Draft - work in progress, November 2002.

[14] 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.

[15] "IGMPv3 Host-side Implementation for NetBSD", <http://www-sop.inria.fr/planete/Hitoshi.Asaeda/igmpv3>

[16] "MLDv2 Host-side Implementation for NetBSD", <http://www-sop.inria.fr/planete/Hitoshi.Asaeda/mldv2>

[17] "Microsoft TechNet", <http://www.microsoft.com/TechNet/prodtechnol/winxppro/reskit/prcc_tcp_rznx.asp>

[18] "Sprint Labs IGMPv3 Multicast Implementation for Linux", <http://www.sprintlabs.com/Department/IP-Interworking/multicast/linux-igmpv3>

[19] S. Kent and R. Atkinson, "IP Authentication Header", RFC2402, November 1998.

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

[21] D. Meyer and B. Fenner, "Multicast Source Discovery Protocol (MSDP)", Internet Draft - work in progress, November 2001.

[22] B. Fenner, M. Handley, H. Holbrook and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", Internet Draft - work in progress, March 2002.

[23] P. Rajvaidya, K. Ramachandran and K. Almeroth, "Detection and Deflection of DoS Attacks Against the Multicast Source Discovery Protocol", Appeared in Proceedings of IEEE INFOCOM 2003.

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

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

[26] V. Roca and A. El-Sayed, "A Host-Based Multicast (HBM) Solution for Group Communications", Proceedings of IEEE ICN'01, July 2001.

[27] D. Kutscher, J, Ott and C. Bormann, "Session Description and Capability Negotiation", Internet Draft - work in progress, March 2002.

[28] B. Fenner, H. Holbrook and I. Kouvelas, "IPv4 Multicast Source Notification of Interest Protocol (MSNIP)", Internet Draft - work in progress, February 2002.

Contents

1 Introduction to Multicast Session Announcement over the Internet and Related Works **3**
1.1 Current Session Announcement Techniques 3
1.2 Overview of Other Information Distribution Systems 4
1.3 Realistic Scope Definition for Session Announcement 4
1.4 Goals of This Work and Organization of the Paper 5

2 Properties of a New Multicast Session Announcement Scheme **6**

3 Multicast Channel Advertisement Architecture **8**
3.1 Concept 8
3.2 Policy Control 9
3.3 Scope Control 9
3.4 Consideration of High Availability 11

4 Multicast Channel Advertisement Protocol **12**
4.1 Scope Label Synchronization 12
4.2 Channel Information Distribution 14
4.3 Consideration of Inter-domain Channel Management 14

5 Configuration of Channel Reflector **17**
5.1 XML Formatted Information 17
5.2 Configuration of Channel Lists 18
5.3 Site-local Policy Configuration 18
5.3.1 Parent and Child Channel Reflector 18
5.3.2 Receiver Lists 19
5.3.3 Sender Lists 19
5.3.4 External Channel Entry 21
5.3.5 Site-local Channel Lists 21

6 Conclusion and Future Work **22**

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399