G. Cristallo, G. Gastaud
,D. Ooms, D.Galand, C. Preguica
ALCATEL
A.Baudot
FRANCE TELECOM
G. Diribarne
UDcast

Internet Draft

Document: <draft-many-ngtrans-connect-ipv6-igp-01.txt>
Feb 2002,
Expires August 2002

Connecting IPV6 islands within an IPV4 AS

Status of this Memo

1. Abstract

This draft proposes a mechanism that allows an incremental deployment of IPv6 inside an IPv4 AS. It aims at automatically set up tunnels, using IGP standard capabilities, in order to gradually introduce IPv6 connectivity within an initially IPv4 AS. As specified in draft-ietf-ngtrans-introduction-to-ipv6-transition-07.txt, this technique must be considered as a solution with a scope of domain.

The connectivity between IPv6 islands within a domain is obtained by the introduction of Routing Protocols Gateways (RPG). IPv6-specific routing information is transferred over the IPv4 routing infrastructure and allows the automatic creation of tunnels between the IPv6 islands.

First, the general mechanism of an RPG is explained. Then, an instantiation of an RPG for OSPFv2/OSPFv3 is detailed and the structure of a generic opaque LSA is specified.
Then, an instantiation of an RPG for IS-ISv6/IS-ISv4 and the structure of a generic TLV is specified.

2. Conventions and terminology used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [2].

Terminology:

o Routing Protocols Gateway RPG

A routing protocols gateway is a component able to transfer routing information (e.g. IPv6 prefixes, metrics...) from a routing protocol to another one.

3. Introduction

This draft proposes a mechanism that allows an incremental deployment of IPv6 inside an IPv4 AS. The mechanism is intended as a transition tool used during the period of co-existence of IPv4 and IPv6. In a late transition phase, the same mechanism can be applied to IPv4 islands in an IPv6 domain.

In this document, an IPv6 island is a connected set of IPv6 nodes surrounded by an IPv4 network, separated from other IPv6 networks. The nodes inside an IPv6 island can be IPv6-only nodes, Dual Stack nodes, or both can be mixed. Internal nodes should only use their IPv6 capabilities. An island will be connected to the IPv4 routing infrastructure with dual stack (IPv4/IPv6) routers.

4. Specification

4.1 General case

As previously said, this technique of transition is used to provide connectivity between IPv6 islands within the same IPv4 AS. The figure below illustrates the situation.

```
+---------------------------------+
|                                 |
|   +--------+         +--------+ |
|   |IPv6   * -  -  - -* IPv6  | |
|   |island |          |island | |
|   +--------+         +--------+ |
|                                 |
|                                 |
```

```
                    |       IPv4 AS domain           |
                    +--------------------------------+
```

    Note : The dual stack routers integrating Routing Protocols gateways
are   represented with "*".

These IPv6 islands are connected to the IPv4 AS with dual stack (IPv4/IPv6) routers containing a RPG.

RPGs could be developed for a lot of routing protocol combinations: OSPFv3, IS-ISv6, etc running in the IPv6 islands and OSPFv2, IS-ISv4 in the IPv4 AS domain.

The IPv4 IGP will flood sufficient information in the AS provide connectivity between these " IPv6 islands " to set up tunnel.

>From a scope point of view, two cases are possible :

An IPv6 island is considered as a part of an IPv4 area :
    In this case, the information flooded through the network (via IPv4 IGP messages) should be broadcasted with an intra area scope. The border area routers will have to process the IPv4 IGP message (with intra area scope) and flood another IPV4 IGP message with inter area scope

An IPv6 island is considered as a complete IPv6 area :
   In this case, the information flooded through the network (via IPv4 IGP messages) must be broadcasted with an inter area scope.

4.2 Connecting IPv6 islands running OSPFv3 over an IPv4 AS running OSPFv2.

4.2.1 Case description

This case is depicted below:

```
                    +--------------------------------+
                    |                                |
                    |   +--------+         +--------+ |
                    |   |IPv6   *- - - - - *IPv6   | |
                    |   |island |          |island | |
                    |   |OSPFv3 |          |OSPFv3 | |
                    |   +--------+         +--------+ |
                    |                                |
                    |                                |
                    |   IPv4 AS domain  (OSPFv2)     |
                    +--------------------------------+
```

RPGs transfer specific information arriving via OSPFv2 to OSPFv3 and vice versa. In this way, it passes information from one OSPFv3 island to another OSPFv3 island crossing one or multiple OSPFv2 areas.

To create tunnels between IPv6 islands, the IPv6 prefixes of a particular island will be passed by OSPFv3 to OSPFv2 via the RPG. The RPG will construct an OSPFv2 message and will flood it to all other OSPFv2 nodes. An OSPF v2 handling specific routing protocol gateway message will receive this message and pass it to OSPF v3 via the RPG.
This technique will automatize the creation of tunnel in the IPV4 domain.

The IPv6 prefixes transferred over the IPV4 IGP can be of any types (e.g. ISATAP).

This naturally drives to the definition of an OPAQUE-LSA for OSPF v2 that carries IPv6-specific information. The next section defines a general opaque LSA format that will be used for this case.

4.2.2. Definition of a Multi-protocol opaque LSA

As specified in the RFC 2370, opaque LSAs consist of standard LSA header followed by application specific information. Opaque LSAs provide a generalized mechanism to allow for the future extensibility of OSPF. The information contained in Opaque LSAs may be used directly by OSPF or indirectly by any application that wishes information to be distributed throughout the OSPF domain.

This draft applies for types (10 ,11) of opaque LSAs, according to the selected scope.

The type 10 is used to connect two islands included in the same area.

The type 11 is used to connect two island included in the same Autonomous System

According to the packet format of the opaque LSA given in appendix A.2 of the RFC 2370 :

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |            LS age             |     Options   |      10 or 11 |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |  Opaque Type  |               Opaque ID                       |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                      Advertising Router                       |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                      LS Sequence Number                       |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |         LS checksum           |       Total Length            |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |             AFI               |    SAFI       | orig Rout prot|
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      | Type          |  Length       | Value ...                     |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                              ...                              |
```

the solution is to use:

o The Opaque Type to define the type of information carried by the opaque LSA. Its value is registered by IANA. In that case it indicates that it carries a tunnel end point value and transferred prefixes.

o The opaque ID field is TBD

o The AFI and SAFI contains the Address Family and Subsequent Address Family Identifiers that are carried by the opaque LSA. It allows to carry multi-protocol as defined by MP_BGP. These fields are already defined by IANA and are re-used in the opaque LSA, since it fits the same purpose.

The AFI and SAFI will express the characteristics of the originating island.

o The Originating Routing Protocol (8 bits) specifies the kind and the version of the routing protocol running inside the origin islands (e.g. OSPF v3,RIPng etc). An opaque LSA contains only one value specifying only one routing protocol. The value 0 is used when no IPv6 routing protocol runs inside the IPv6 island. Other values should be defined for the other originating routing protocols.

o Different TLVs described here under.

For all TLV, eight bits are used to define Type, 8 bits are used to define Length.

These TLVs are used:

Type 1: It corresponds to the tunnel end point address. The address carried is unicast address, with a length corresponding to the address length (128 for IPv6, 32 for IPV4)

Type 1 is mandatory in this opaque LSA and must appear only once in the opaque LSA.

Type 2 :

The format of this type is :

```
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |    Type=2     |TLV Leng=1byte | cost (3 bytes)      ...       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | ... cost      | prefix length |    prefix of the island...    |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

This type is used to transfer one prefix of the island with the related cost. This prefix will be ADDED in the routing database of the remote island(s). The type 2 is optional in this opaque LSA. A cost field has been defined to contain cost from originating routing protocols. If AFI = IPV6, the prefix of the island will be encoded in a 128 bits structure. If AFI = IPV4, the prefix of the island will be encoded in a 32 **bits structure.**

Type 3: To transfer and withdraw one prefix of the island. This prefix will be WITHDRAWN in the routing database of the remote IPv6 island.

New TLVs can be added later.

4.2.3 Role of routing protocols gateways at emission and reception of opaque LSA

The triggers for the emission of this opaque LSA for OSPF v2 are implementation dependant and are not tackled in this document. Nevertheless, the operations that have to be performed at emission and reception of this opaque LSA are briefly detailed.

4.2.3.1 Emission of opaque LSA

When the application of the IPv6 egress
router decides to send LSAs, it includes all the information inside the opaque LSA with respect to the framework detailed previously specifying the opaque ID and, the opaque Type. This opaque LSA is broadcasted throughout the network (with the related scope in order to reduce the level of information flooded in the network).

Emission of WITHDRAWN happens when an OSPF node is not visible anymore. The RPG is in charge of sending a specific WITHDRAWN with the impacted prefixes.

4.2.3.2 Reception of opaque LSA

When receiving this kind of opaque LSA, atunnel is created with the information given by the opaque LSA : the
TLV Type 1 giving the tunnel end point information and the TLVs Type 2 giving the prefixes information.

The way tunnel can be setup and used is implementation dependant.
It can be automatic or configured tunnels.

At reception phase, the same process will be performed when the route has to be withdrawn. The prefixes are withdrawn from the routing table and the WITHDRAWN information is flooded into the island.

When reception of these opaque LSA are successfully completed, each IPv6 island included in this IPv4 AS has the opportunity to communicate with all the other IPv6 islands included in this AS.

4.2.3.3 OSPF cost use

In the Type 2 field, a cost has been designed. This field cost needs to be refined here. Indeed, this field contains a OSPF-based metric, but with modification due to the cross of different routing protocols.

```
+------------------------------------------------------------+
|                                                            |
| IPv6 island A (OSPFv3)            IPv6 island B (OSPFv3) |
|   +------------+                        +------------+  |
|   |            |                        |            |  |
|   |           RPG A                    RPG B         |  |
|   |            *--------------------------*          |  |
|   |   metric A |      metric B            | metric C   |  |
|   +------------+                        +------------+  |
|                                                            |
|                  IPV4 AS domain  (OSPFv2)                  |
+------------------------------------------------------------+
```

RPG A is a member of IPv6 island A. It is an OSPFv2/v3 router. RPG A calculates a path toward destination A in IPv6 island A with a cost equal to metric A. It creates an opaque LSA for advertisement of this prefix to IPv6 island B. This opaque LSA contains a TLV type 2 with prefix A and a "metric A". Metric A is an OSPFv3 metric.

RPG B receives this opaque LSA. RPG B calculates a path to RPG A with a cost equal to metric B: this is an OSPFv2 metric because the path between RPG B and A crosses the IPv4 domain.

The total cost, from RPG B to RPG A is equal to metric A + metric B. Since the two metrics come from different origins, the total cost from RPG B to destination A is set to: metric A + metric B + DELTA. RPG B advertises this prefix A, within IPv6 island B, with the cost: metric A + metric B + DELTA.

DELTA would be a fixed cost that allow to avoid problem when cost that allow to give the RPG prefixes less specific than home network.

The value of DELTA is TBD: it probably will be a configurable parameter.

```
+------------------------------------------------------------+
|                                                            |
| IPv6 island A (OSPFv3)            IPv6 island B (OSPFv3) |
|   +------------+                        +------------+  |
|   |            |                        |            |  |
|   |           RGP A                     RPG B        |  |
|   |  A--met AA-*---------metric BA--------*          |  |
|   |   \-met CA-*---------metric BC------ /|          |  |
```

```
|   |           RPG C                      |            |  |
|   +------------+                         +------------+  |
|                                                          |
|                  IPV4 AS domain  (OSPFv2)                 |
+----------------------------------------------------------+
```

RPG A and RPG C both advertise prefix A with -respectively- metric AA and metric CA. RPG knows the cost of the path to reach A and C -respectively BA and BC. The costs of the two paths to A are:

    via RPG A: metric BA + metric AA + DELTA
    via RPG C: metric BC + metric CA + DELTA

RPG B (or any node within island B) can thus decide which of the two paths is the best one. No loop is introduced and no configuration is needed. RPG advertises the best of the two paths within IPv6 island B.

4.3 Connecting IPv6 islands running IS-IS v6 over an IPv4 AS running IS-IS v4

4.3.1 case description

This case is depicted below:

```
                    +---------------------------------+
                    |                                 |
                    |   +-------+         +-------+ |
                    |   |IPv6   *- - - - - *IPv6   | |
                    |   |island |         |island | |
                    |   |IS-ISv6 |         |IS-ISv6| |
                    |   +-------+         +-------+ |
                    |                                 |
                    |                                 |
                    |   IPv4 AS domain  (IS-ISv4)     |
                    +---------------------------------+
```

RPGs transfer specific information arriving via IS-ISv4 to IS-ISv6 and vice versa. In this way, it passes information from one IS-ISv6 island to another IS-ISv6 island crossing one or multiple IS-ISv4 areas.

This case naturally drives towards the definition of a new TLV for IS-IS v4 to allow the connectivity between these IPv6 islands.

4.3.2 Definition of a TLV for IS-IS v4

An IS-IS LSP is composed of a fixed header and a number of tuples. This part details a new TLV and associated sub-TLVs.

The new TLV called IGP-tunnel TLV (type value: To be registered by IANA) contains information in order to provide connectivity between IS-IS v6 islands included in one IPv4 AS running IS-IS v4.

IGP-tunnel TLV
     Type (1 byte) : TBD
     Length (1 byte): Total length of value field
     Value :
          2 bytes : AFI
          1 byte : SAFI
          1 byte : originating routing protocol
          1 byte of length of sub-TLVs
          0-248 bytes of sub-TLV

In order to transfer all the important information, it is necessary to use Sub TLVs.

The list of sub-TLV is presented here :

Sub TLV number
1 : tunnel end point
2 : prefix of the island to be added
3 : prefix of the island to be removed

sub TLV 1 :
    Type (1 byte) : 1
    Length (1 byte). If AFI = IPv4, the prefix of the island to be
    added (or removed), the prefix will be encoded with 4 bytes. Otherwise
    it will be encoded with 16 bytes.
    Value : IPv4 address or IPv6 address

Sub TLV 2 :
    Type (1 byte) : 2
    Length (1 byte) : Total length of value field
    Value :
          3 bytes for the cost
          1 byte for the prefix length
          4-16 bytes for the prefix islands to be added

Sub TLV 3 :
This sub TLV is encoded in a same way than Sub TLV2.

4.4 Routers requirements

All the routers present in the IPv4 AS, must be opaque-LSA/IS-IS TLVs capable in order to assure the transfer of the opaque LSAs/IS-IS TLVs. All the dual-stack routers (RPG) must be able to decode the content of these opaque LSAs/IS-IS TLVs.

4.5 Mid-term evolution for this draft

For the moment, this technique is used to provide connectivity between IPv6 " islands " included in an IPv4 Autonomous System. In the future, it will be possible to connect IPv4 islands included in an IPv6 Autonomous System.

Combining IGP tunnel technique and BGP tunnel technique will allow to provide connectivity between IPv6/IPv4 islands included in different IPv4/IPv6 autonomous system.

5. IANA Considerations

Number for opaque Type and number for the new IS-IS TLV need to be registered at IANA

6. Security considerations

The security mechanisms of the native routing protocols are used.

7. References

8. Author's Addresses

Alain Baudot
FRANCE TELECOM R&D
FTR&D/DMI/SIR/IPI
42 rue des coutures
BP 6423
F-14066 CAEN Cedex 4
Phone: +33 2 31 75 94 27
Email: alain.baudot@francetelecom.com

Geoffrey Cristallo
ALCATEL
Fr. Wellesplein 1, 2018 Antwerp, Belgium
Email: Geoffrey.Cristallo@alcatel.be

Gilles Diribarne
UDCAST
24, Route des Dolines
BP 355
06906 SOPHIA ANTIPOLIS
Phone: +33 4 93 00 16 60
Email: Gilles.Diribarne@udcast.com

Damien Galand
ALCATEL R&I
Route de Nozay,
F-91461 Marcoussis Cedex
FRANCE
Phone: +33 1 69 63 41 84

Email: Damien.Galand@alcatel.fr

Gerard Gastaud
ALCATEL, 10 rue latecoere 78140 VELIZY, FRANCE
Phone: +33 1 30 77 42 20
Email: Gerard.Gastaud@alcatel.fr

Dirk Ooms
ALCATEL
Fr. Wellesplein 1, 2018 Antwerp, Belgium
Email: Dirk.ooms@alcatel.be

Christophe Preguica
ALCATEL
Route de Nozay
F-91461 Marcoussis Cedex
Phone: +33 1 69 63 44 09
Email: Christophe.Preguica@alcatel.fr

1 J Moy, "OSPF Version 2", RFC 2328, April 1998

2 Rekhter, Y. and Li, T., "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995

3 Coltun, R., Ferguson, D. and Moy, J., "OSPF for IPv6", RFC 2740, December 1999

4 Templin, Fred L., "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)", draft-ietf-ngtrans-isatap-02.txt, March 2001

5 **R Callon, "Use of OSI IS-IS for Routing in TCP/IP and Dual** Environments", RFC 1195, Dec 1990

6 **Christian E Hopps, "Routing IPv6 with IS-IS",** draft-ietf-IS-IS-IPv6-02.txt, April 2001