Routing Arbiter Working Group   John Chu, editor
IBM Corp.
Version 1.0
November 17, 1994

Route Server Design

i. Scope of this Document

The Routing Arbiter Architecture evolved from the Interdomain Routing Architecture. By illustrating this progression, this document establishes the continuity from what is already well understood, i.e., the Interdomain Routing Architecture, to what is to be constructed, i.e., the Routing Arbiter.

By establishing such continuity, one can determine what is already available in terms of know-how and resources, so that further work can be formulated.

This document covers the Route Server, which is a major component of the Routing Arbiter Architecture. Works needed for the other components, namely the Routing Policies Analysis, the Routing Information Database (or the Routing Registry Database,) and the Routing Arbiter Network Operations Center, are to be covered by separate documentation.

This document is a design specification, to formulate the work needed for the construction of Route Server.

The reader is assumed to have a good understanding of the Interdomain Routing Architecture, and be familiar with the Routing Arbiter Architecture and its intented purposes. Please see the Reference section of this document for more background information.

The following Acronyms are used through out this document:

AS	- Autonomous System
BGP	- Border Gateway Protocol [2]
BR	- Border Router
FIB	- Forwarding Information Base
IDRP	- Inter-domain Routing Protocol [3]
NAP	- Network Access Point [1]
NSP	- Network Service Provider [1]
RA	- Routing Arbiter [1]
RIB	- Routing Information Base
RIDB	- Routing Information Database [1]

RRDB - Routing Registry Database, an implementation of RIDB
RS - Route Server

ii. Acknowledgement

This document is the result of extensive discussions with Jim Rubas (IBM), under the leadership of Yakov Rekhter (IBM), for the implementation of the Route Server.

Kannan Varadhan (ISI) provided refreshing perspectives in understanding of the routing architectures and offered many valuable comments.

Members of the Merit Internet Engineering Group (MERIT) provided many valuable comments on the drafts of this document. Special thanks go to Elise Gerich (MERIT) and Jessica Yu (MERIT) for their comments and assistance.

Dennis Ferguson (MCI), who implemented BGP-4 in gated, offered valuable suggestions at the early stage of the Route Server development.

1. From Border Router to Route Server

In this section, the similarities and differences between a BR and an RS are examined.

1.1 The Border Router

A Border Router receives routes from one or more of its neighboring ASes by speaking a border gateway protocol (BGP/IDRP) with those ASes. What routes to accept and how the accepted routes are arranged depend on the local AS's import policies. The routes, filtered by the policies, are then stored in the BR's RIB and FIB. Figure 1 models this part of the route flow in the BR.

```
                                        BR
                            import
                            policies
                               /|   ____________________
                              //   /                   /
             ASa ------------>  //   /       RIB        /
                  route flow   |/  /___________________/
                                             |
                                             v
                                            FIB
```

Figure 1

Once stored in the RIB, the routes may then be propagated to other neighboring ASes. What routes and to which ASes the routes are propagated are determined by the local AS's export policies. Figure 2 models this part of the route flow.

```
                             BR            export
                                           policies
                     ____________________   /|
                    /                   /  //
                   /       RIB         /  //  ------------>  ASa+1
                  /___________________/  |/    route flow
                             |
                             v
                            FIB
```

Figure 2

The plane representing the RIB in Figure 1 or 2 can be thought of as the view of the outside world as seen by the BR, biased by its import policies governed by local interest.

1.2 The Route Server

A Route Server receives routes from the ASes attached to the NAP by speaking a border gateway protocol (BGP/IDRP) with those ASes. So, an RS appears to those ASes just like a BR does.

An RS differs from a BR by maintaining separate views for different sets of ASes. The Routing Policies Analysis segregates the ASes into a number of sets with each set containing one or more ASes and each AS is a member of only one set. How the Routing Policies Analysis determines the sets is based on the commonalities in the import policies submitted by the ASes, and its details are covered by the documentation for the Routing Policies Analysis component. The RS serving that NAP is responsible for providing a separate view for the ASes in each different set.

The differences among the views are the result of different import policies submitted by the ASes and examined by the Routing Policies Analysis, which in turn groups the ASes that share the same import policies into one set. The sets of ASes are determined and the policies for the sets are stored in the Routing Information (or Routing Registry) Database.

When a route is received from any AS, the RS applies to it the import policies of all the sets independently to build corresponding view for each set. Figure 3 models this process.

```
                                       RS
                           import                        view
                           policies
                           per view                       ^
                              /|  _____________________   |
                             //  /                   /    |
                     +->    //  /       VIEWv+2     /     |
                     |     |/  /___________________/      |
                     |                                    |
                     |        /|  _____________________   |
                     |       //  /                   /    |
                     +->    //  /       VIEWv+1     /     |
                     |     |/  /___________________/      |
                     |                                    |
                     |        /|  _____________________   |
                     |       //  /                   /    |
         ASa --------+->    //  /       VIEWv       /     +----------->
             route flow    |/  /___________________/        route flow
```

Figure 3

The concept of views in RS can be thought of as extensions orthogonal to the route flow, as indicated on the right-hand side of Figure 3.

Once each of the views has been updated, depending on the contents of the import policies for each view, it may or not may produce a route to be propagated to the ASes for which that view has been maintained by the RS. Furthermore, even if such a route can be produced, whether it will get propagated to an AS is subject to the export policies of that AS. Figure 4 gives an example for this process.

```
                              RS
                                       export
                                       policies
                                       per as
                                             /|
                                            |/   --------->  ASa+3
                  ____________________    /|
                 /                   /   |/                ASa+2
                /       VIEWv+2     /  /|
               /___________________/  |/   ---------->  ASa+6

                  ____________________
                 /                   /
                /       VIEWv+1     /  /|
               /___________________/  |/                ASa+5

                  ____________________    /|
                 /                   /   |/   --------->  ASa+7
                /       VIEWv       /  /|
               /___________________/  |/   ---------->  ASa+1
                                           route flow
```

Figure 4

In the example illustrated in Figure 4, the RS maintains VIEWv for {ASa+1, ASa+7}, VIEWv+1 for {ASa+5}, and VIEWv+2 for {ASa+6, ASa+2, ASa+3}, respectively. VIEWv produced a route, and the export policies for ASa+1 or ASa+7 allowed its propagation. VIEWv+1 did not produce a route. VIEWv+2 produced a route, and the export policies for ASa+6 or ASa+3 allowed its propagation, but the export policies for ASa+2 did not allow.

An RS also maintains a separate view just for itself. The output of that view is used to update the RS’s FIB.

2. Identification of Views

The concept of views is present in each of the components of the Routing Arbiter Architecture.

In Routing Policies Analysis, views are the sets of characteristics that resulted from analyzing the import policies submitted by the NSPs and applying constraints in order to provide equitable routing services to the NSPs and to achieve a consistent and balanced routing pattern. NSPs (ASes) are placed into the sets according to the characteristics they possess. Each set carries along the policies that reflect those characteristics.

In Routing Information Database (or RRDB), the views are recorded as the sets of ASes and their respective policies. These data are fed into tools that produce configuration files understood by the RS.

In the configuration file, each AS is identified as a member of a particular set. Each set of ASes is mapped to the view for that set. Import policy statements are specified for each view, and export policy statements are specified for each AS.

In Route Server, each view is represented by a route table, which records the routes along with the policies specified for that view. Only the ASes in the set for which this view is maintained are sent route updates if the route table generates one, filtered by the export policy for that AS.

3. Configuration of Views

The RS is configured by a configuration file generated by the Routing Information Database (or RRDB.) In this file, statements are made about the following facts:

- Identification of a neighboring AS and the membership of the set it belongs to. The set is represented by its view.

- Preference assigned to a route received from a neighboring AS and maintained in a view.

- Preference assigned to a route received from a neighboring AS, for a specific destination, and maintained in a view.

- Preference assigned to a route carrying a specific pattern of ASPATH and maintained in a view.

We will examine each of these statements in this section.

3.1 Neighboring ASes and Their Set Membership

The ASes that are peering with the RS through BGP/IDRP are identified in the RS's configuration file. The view serving each AS is also identified. This statement can be defined as the following:

```
peeras <as> ofview <view>
    [...]
        {...};
```

where <as> is the AS number, and <view> is a number identifying the view serving AS <as>. This statement says: AS <as> is a peer of the RS and the RS maintains view <view> for this AS.

This statement may be repeated for different values of <as> and different values of <view>. For each value of <view>, there may be multiples of this statement each of which has a different value of <as>. For each value of <as>, there is one and only one value of <view>. This reflects the fact that one view can serve one or more ASes, and one AS is only served by one view.

3.2 Preference Assignment According to the Source of a Route

The source AS, from which a route is received, can determine the degree of preference associated with that route in a view. This statement is nested in the peeras statement as defined in 3.1 above so that the source AS is identified. It can be defined as follows:

```
peeras <as> ofview <view>
    preference <preference> forview <view+v> [...]
        {...};
```

where <preference> is a numerical value specifying the degree of preference, and <view+v> is a number identifying the view. This statement says: when a route is received from AS <as> and is placed in view <view+v>, assign it preference <preference>. The preference statement can be repeated for all values of v.

3.3 Preference Assignment According to the Source of a Route and its Reachability

In addition to the source AS, the destination in the route can also determine the degree of preference associated with that route in a view. This statement can be defined as follows:

```
import as <as+a> forview <view+v> {
    prefix <prefix> preference <preference>; [...]
};
```

where <prefix> is the destination/mask that specifies a reachability. This statement says: when a route is received from AS <as+a> and its reachability is <prefix>, assign preference <preference> before placing it in view <view+v>. The import statement can be repeated for all values of a and v.

3.4 Preference Assignment according to AS Path Attribute

Instead of the source AS, the AS Path Attribute carried in a route can determine the degree of preference associated with that route in a view. This statement can be defined as follows:

```
import aspath <aspath_spec> forview <view+v>
    preference <preference>
        {...};
```

where <aspath_spec> is a regular expression specifying an AS path or a part of it. This statement says: if a route carries an AS Path Attribute that matches <aspath_spec>, assign preference <preference> before placing it in view <view+v>. The import statement can be repeated for all values of v.

The policy statements specified in this section place constraints on the attributes of a route, and they are listed in the order with progressively more specific constraints. The preference assignment statements nested in a policy statement with more specific constraints overwrite those nested in a policy statement with less specific constraints.

References

[1] Estrin, D., Postel, J., and Rekhter, Y., "Routing Arbiter Architecture", University of Southern California and Information Sciences Institute, Information Sciences Institute, T.J. Watson Research Center, IBM Corp., May, 1994.

[2] Rekhter, Y., and Li, T., "A Border Gateway Protocol 4 (BGP-4)", Internet Draft, T.J. Watson Research Center, IBM Corp., cisco Systems, May, 1994.

[3] Information Processing Systems - Telecommunications and

Information Exchange between Systems - Protocol for Exchange of Inter-domain Routeing Information among Intermediate Systems to Support Forwarding of ISO 8473 PDUs, ISO/IEC 10747, October 18, 1993.

Editor’s Address

John Chu
IBM Corp.
P.O. Box 704
Yorktown Heights, NY 10598

Phone: +1 914 784 7839
Email: jychu@watson.ibm.com