ATNP/WG2/12
WP/_391
9 June 1997

AERONAUTICAL TELECOMMUNICATIONS NETWORK PANEL

WORKING GROUP TWO

Langen, June/July 1997

The IDRP Route Server alternative to a full mesh routing

Presented by Arnaud Dedryvère

Prepared by Stéphane Tamalet

(France)

SUMMARY

This Information Paper describes the IDRP Route Server function and the potential benefits that may represent the use of Route Servers in the ATN.
The use of Route Servers may reduce overhead and management complexity of maintaining numerous direct BIS-BIS connections which otherwise might be required or desired among BISs in different domains that are connected to a common subnetwork.

TABLE OF CONTENTS

1. Introduction

The aim of this Information Paper is to provide WG2 participants with food for thought about the Route Server IDRP function, that is an optional IDRP mechanism, totally compatible with the ATN SARPs, and the use of which could be proved very useful in an ATN network for reducing overhead, optimizing the routing, and simplifying routers administration.

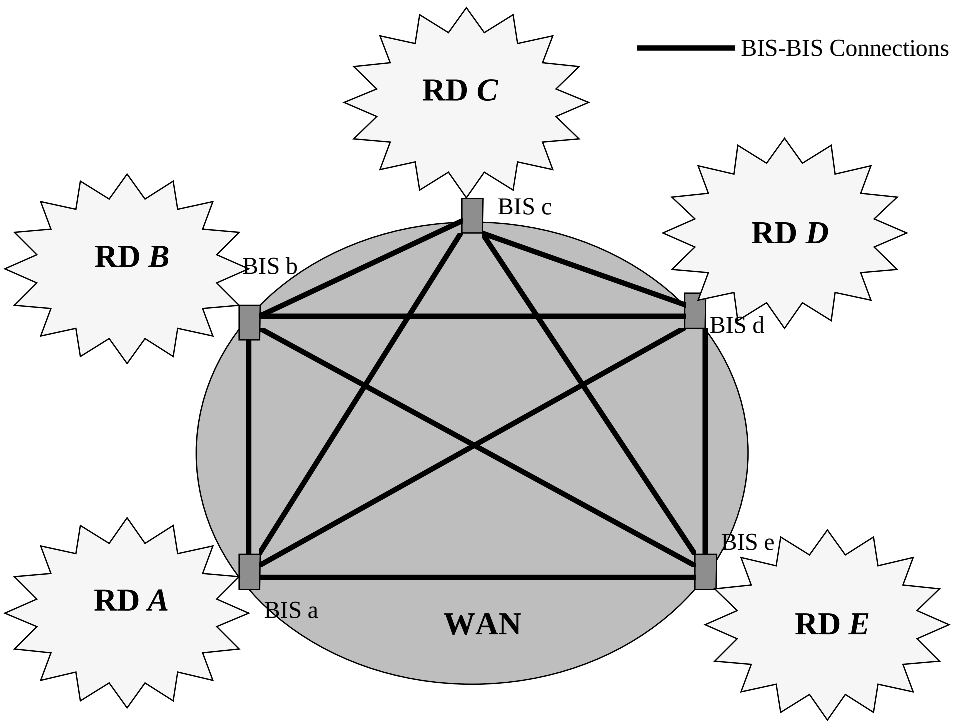
This Paper describes the IDRP Route Server function and the potential benefits that may represent the use of Route Servers in the ATN.

2. Problem Statement

Consider an international Wide Area Network, such as an international ATM or X.25 subnetwork and a set of BISs connected to it. These routers are administered by different organizations and exchange routing information across organizational boundaries (by participating in inter-domain routing).

Any BIS connected to the international WAN is potentially able to reach directly any other BIS connected to the network. Thus, in principle it should be possible to arrange the inter-domain routing over such a network in such a way that any IDRP routes crossing this network would have to go at most through 2 BISs connected to the network. In other words, any CLNP packets should be able to go from any RD to any other RD connected to the same international WAN in one single hop through the shared media.

In the most simple scenario, the solution to minimize the number of hops between BISs in different domains and connected to a common WAN is that each BIS maintain direct routing peering with each other. This way each BIS acquires routing information about destinations reachable through every BIS connected to the WAN. Such direct peering allows to acquire « first hand » information about destinations which are directly reachable through adjacent routers and select the optimum direct paths to these destinations. A single CLNP hop objective would be accomplished.



One problem with this scenario is the number of peers each BIS has to maintain. For a WAN with N BIS, the scheme would require each such router to maintain (N-1) peering relations. And considering that each ATN BIS has to support 2 Adj-RIB-Ins and 2 Adj-RIB-Outs per connected adjacent BIS, the maximum number of connections that an ATN BIS can support may become a limiting factor that makes the approach impractical.

The second problem with this scenario is the routing instability problems inherent to fully meshed topology, due to the great number of alternative paths that exist between each pair of BISs. IDRP Convergence modelling studies have shown that the convergence delay measured on a fully meshed Backbone RDC consisting of several BISs increases with the number of BISs.

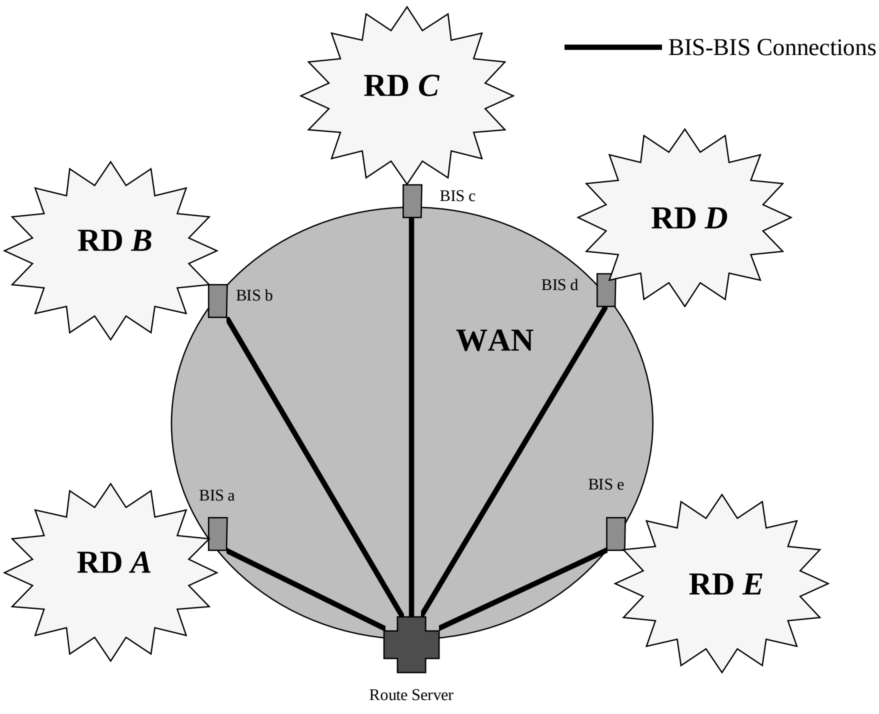
Thirdly, for a switched media as ATM, X.25, or Frame Relay network which may inter-connect a large number of routers, the large number of subnetwork connections that would be needed to maintain a full mesh direct peering between the routers, makes also this approach impractical.

This paper describes a solution, based on a standard IDRP function, which allows to alleviate the « full mesh » problem.

3. Indirect Peering Through an IDRP Route Server

3.1 Principle

A way to reduce the number of peering relations is to make the exchange of routing information among the BISs through indirect peering, where a BIS connected to the common WAN would acquire the routing information provided by all other BISs connected to the same network by peering with a particular BIS called "Route Server".

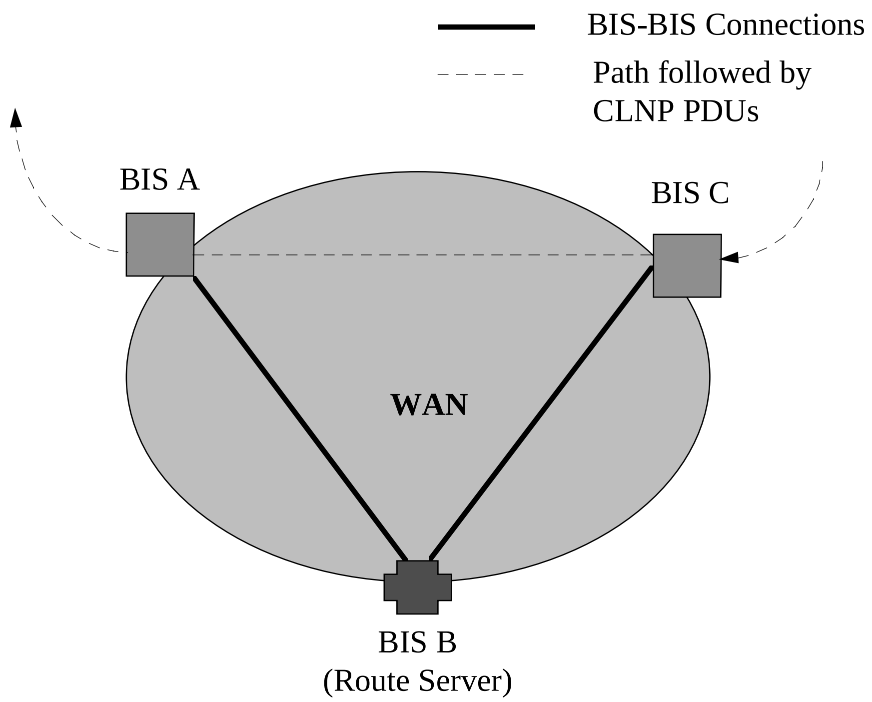


A Route Server is a system that participates in IDRP, but doesn't participate in the actual CLNP packet forwarding. A Route Server is a BIS dedicated to the processing of routes: it acquires routing information from all the BIS connected to a common WAN, performs decision process over this information, and then redistributes the results to the routers.

When a Route Server acquires a route from a BIS, and passes the route to other BISs the Route Server specifies the address of the router from which the route was acquired in the NEXT_HOP path attribute. This way CLNP traffic bypasses the Route Server, and flows directly between the routers.

To better understand the route server function, consider the following example:

- Three BISs A, B and C are all attached to the same WAN
- Through the WAN, A can reach directly B and C, B can reach directly A and C and C can reach directly A and B
- A and B share a BIS-BIS connection; B and C share a BIS-BIS connection
- B is a route server



if A propagates an UPDATE PDU to B, the route server function of B will make B readvertise this route to C with an additional NEXT_HOP attribute, indicating that the true Next Hop BIS for this route is A. As a result, all CLNP packets following this route will directly be forwarded by C to A and this even if A and C do not share a BIS-BIS connection.

It is easy to see that by using Route Server when N BISs are connected to a common WAN, instead of (N-1) peering relations, each BIS has to maintain only 1 peering relations -- the relation with the Route Server. Thus, this approach presents a clear improvement (in terms of the required number of peering relations) over the approach outlined in the previous section.

In addition to improved scaling with respect to the number of peering IDRP connections, the approach presented in this section also improves scaling with respect to the volume of routing information. In presence of multiple routes to a destination only the Route Server has to maintain all these routes, while other routers (on a common WAN) that peer with the Route Server have to maintain at most one route -- the one selected by the Route Server.

Note that the Route Server is still expected to be capable of peering with all the routers connected to a common WAN; the Route Server is also expected to be capable of storing all the routes received from all the routers. So, the approach improves scaling on most, but not all the routers. Specifically, the approach reduces the load on all the routers, except for the Route Server, -- the load on the Route Server is the same as on any router in the case of a complete mesh peering. However, the Route Server is only assumed to handle IDRP Traffic; it does not participate in the CLNP forwarding. It can therefore be considered that a Router alleviated from the task of forwarding packets will be in a position to manage more routing traffic than a classical BIS both forwarding CLNP packet and processing IDRP routes.

To eliminate a single point of failure associated with a single Route Server, a small number of Route Servers (e.g. 2-3) can be deployed. This way each BIS will maintain a peering relation with every Route Server connected to the WAN. If one of the Route Servers goes down, the information provided by the other Route Servers should be sufficient to preserve routing.

3.2 Drawback of the approach

The approach presented in this section has one drawback, as compared to the approach of the previous section. With complete mesh peering (as described in the previous section) each BIS acquires routing information from all of the other adjacent BIS, and performs the selection process (selecting best routes) based on its local route selection criteria using the acquired information as an input. As a result, with the complete mesh approach there is no interdependencies on the selection criteria among different BIS. With the approach outlined in this section the selection is done by the Route Server. Thus, if two different BIS connected to the same WAN have different selection criteria, then the Route Server can accommodate either of them, but not both.

For example, if a destination X is reachable via two BISs, E1 and E2, connected to the common WAN, and some other BIS A (connected to the network) prefers the route through E1, while some other BIS B (also connected to the network) prefers the route through E2, then the Route Server can make its selection process consistent with either A (thus advertising the route through E1 to all of its peers, including A and B), or with B (thus advertising the route through E2 to all of its peers, including A and B), but not both.

To accommodate certain diversity among the route selection criteria of the BISs connected to a common WAN the scheme presented in this section could be augmented with selective direct peering between BISs. This way most of the routing information will still be acquired via the Route Server, while "exceptions" (due to the diversity of route selection criteria) could be handled via direct peering. Therefore, the scheme outlined in this section is appropriate as long as route selection criteria among all the BISs connected to the WAN are fairly consistent among themselves.

Another solution could be that the Route Server does not apply any selection criteria to the routes received from the BISs for the purpose of distributing all these routes to its clients. In such a case, all routes acquired from the BISs would be relayed to all BISs. With such an approach however, the BISs connected to the WAN, although alleviated from the task of managing a large number of BIS-BIS connections, will not be alleviated from the route pre-selection task that could have been made by the route server.

3.3 Required IDRP mechanisms

The Route Server approach described in the previous section relies on the use of an optional feature of the ISO10747 (IDRP) standard:

- The BISs which are client to the route server shall support « the generation of the NEXT_HOP attribute in support of route servers » options

- The Router Server is a BIS which shall support the « propagation of the NEXT_HOP attribute in support of route servers »

Note: the support on receipt of the NEXT_HOP attribute is a mandatory IDRP function and is therefore assumed to be supported by every standard/ATN IDRP implementation..

The Route Server approach relies therefore on standard mechanisms and can be used in the ATN provided that the options mentioned above are implemented.

The use of Route Servers does not a priori conflict with any of the SARPs requirements.

3.4 Interest of the Approach in the ATN case

The ATN ICS SARPs and Guidance Materials already define ATN routing concepts (such as the concepts of ATN Backbone, ATN Islands and ATN home) relying on properties of the IDRP routing architecture and which will allow to make the ATN a scaleable network.

In this context the interest of new concepts such as the use of Route Servers may appear questionable.

On the other hand, modelling studies have already shown that bottlenecks in the ATN Architecture are possible if no care is made during network architecture design. The IDRP Convergence modelling study made by CENA and Eurocontrol, has shown that the reduction of IDRP convergence delays compelled to break down large ATN Islands into several smaller Islands with small Backbone RDC consisting of not more than 3 BISs. These topological constraints may become cumbersome particularly when the ATN network will expand and include more and more BISs. The expansion of the network could compel to break islands in several smaller Islands or in introducing new hierarchical levels such as sub-Island RDCs.

Considering the requirements that will exist for ground-ground ATN traffic between adjacent ATC centres of the same or different countries, between ATC centres and adjacent airports in the same or different countries, and between adjacent airports of the same or different countries while in the same time large international X.25, ATM or Frame Relay networks can be made available to the ATN community, we can foresee that certain portion of the ATN will require direct Routing Domain to Routing Domain interconnections (without going through a common backbone). This may be the case in France, for instance, where we could imagine that 5 Routing Domains will be formed (one Routing Domain around each of the five French ATC centres), and where direct RD to RD communications will be desired.

Another portion of the ATN where a fully meshed topology could be required, is the ATN Island Backbone RDCs. And it is likely that there will be a requirement for fully meshed backbones consisting of more than 3 BISs. Redundancy of Backbone BISs, and the possible limitations on adjacency number of Backbone BISs, will indeed certainly necessitate the multiplication of BISs in the Backbone.

In those portions of the ATN where a fully meshed topology is required, it is believed that the use of Route Servers will be of interest. The use of Route Servers will allow the implementation of virtually meshed topologies, taking advantage of the large scale subnetwork technology which might be available and optimizing the routing (1 CLNP hop over these network) while avoiding the routing stability problems inherent to true fully meshed topologies. By reducing the number of BIS-BIS connections, this approach should additionally simplify administrative issues such as the setting up of bilateral agreements.

4. Conclusion

The Route Server function appears to be of interest and should be more deeply investigated with for instance simulations studies.

BIS implementers should consider the related IDRP options, as functions that will likely be required by a number of administrations willing to develop an ATN ground infrastructure.

Recommendation: If this approach is considered appropriate to WG2 for ATN network design, Guidance Materials should be drafted about this route server architecture.

5. References

[REF1] Yakov Rekhter - « Inter-Domain Routing over ATM networks », Internet Draft, Ferbruary 1995

[REF2] N. Bui, T. Kircher - « ATN Islands and Homes IDRP Convergence Modelling Study - Final Report » - CENA/RT96014 - January 1997