Comparison of IBGP Topologies

Virginie Van den Schrieck
Advisor : Olivier Bonaventure
Dept CSE, Université Catholique de Louvain (UCL)
Louvain-la-Neuve, Belgium
vvandens@info.ucl.ac.be

Categories and Subject Descriptors: C.2.2 [Computer-Communication networks]: Network Protocols

General Terms: Measurement, Performance, Design.

Keywords: iBGP, route reflection, topology, performances.

1. INTRODUCTION

BGP is the de facto interdomain routing protocol in today's Internet [1]. Its role is to allow ASes to exchange routing information with each other. It is divided in two components : eBGP rules the communications between ASes, while iBGP allows one router to transmit routing information received from its eBGP peers to the other BGP routers of its AS. The original specification of iBGP requires a full-mesh of sessions between BGP routers to prevent loops of the advertisements. This constraint means that there will be (n*n-1)/2 iBGP sessions maintained inside the AS, which does not scale well in the case of ASes.

One alternative that has been proposed is to use route reflection [2]. In this scheme, one (or more) BGP router plays the role of *route reflector* for some of the others, which are its *clients*. This router reflects the routes it receives to its clients, and announces its routes and those of its clients to its non-client peers. With those rules, a client router must only maintain one session with its route reflector, and the iBGP topology is much lighter. However, this alternative has drawbacks. For example, as route reflectors have to go through their decision process before announcing routes to their clients, these only receive the routes selected by the reflector. The selection performed by the route reflector may differ from the selection that its clients would make in the case of an iBGP full-mesh, which can lead to sub-optimal routing.

This paper presents a method to evaluate iBGP topologies with route reflection, and to measure their performances compared to the full-mesh, in order to choose the best one. First, we describe the procedure and the criteria used for the comparison. After that, we briefly present the result of an application of this method to the GEANT network [3]. This analysis allows us then to isolate some of the factors that are influencing the quality of an iBGP topology.

CoNEXT'05, October 24–27, 2005, Toulouse, France.
ACM 1-59593-035-3/05/0010.

2. COMPARISON METHOD

The comparison process can be divided in two steps : the choice of the topologies that are evaluated and the evaluation itself. To build an iBGP topology, we need some routers that can play the role of route reflectors. In our heuristic we choose the routers located in the center of the network, i.e. that minimize the sum of IGP distances to the other routers, in order to get an iBGP topology similar to the physical topology. Then, once the route reflector candidates are selected, we build different types of topologies. We study topologies with only one or two reflectors, and with sessions of type *client-reflector* and *reflector-reflector*. Topologies with two reflectors are with and without redundancy, and there are no clients on the physical path between the reflectors. We can then start the evaluation process, based on simulations using the C-BGP tool [4] with a set of routes announced to the AS by eBGP peers. We then compare the results based on four metrics.

The first metric measures the benefit of the introduction of route reflection on the iBGP topology, which is simply reflected by the sum of the IGP length of the paths along all the iBGP sessions. The second criteria is the size of the routing tables. Indeed, route reflection allows them to be much smaller, as the reflectors are only announcing one route for each destination, while with a full-mesh, every router receives all possible routes to each destination. Next, we measure the loss of optimality in the routing tables. This loss come from the fact that it is the reflector that selects the best route for each destination, not the client, thus packets may follow a longer route through the AS. We measure this loss of optimality by taking, for each route chosen by a router, the distance from this router to the exit point of the AS. Our third metric will then be the sum of all those distances. Finally, if the reduction of the routing tables mentioned earlier is interesting to reduce memory usage, it also has drawbacks. If a client only has one route for a destination, and this route is invalidated, it will have to wait for a new route from its reflector. But if it knows two or more different routes for this destination, it becomes possible for him to directly choose another one. Redundancy in the routing tables is then desirable, and we measure this parameter by counting the number of destinations having two or more different routes inside the tables of each router.

Once we have made the simulations and obtained values for all metrics, we are able to compare our topologies. The best of them will typically have large value of redundancy inside the routing information tables and low values for the other metrics.

3. RESULTS

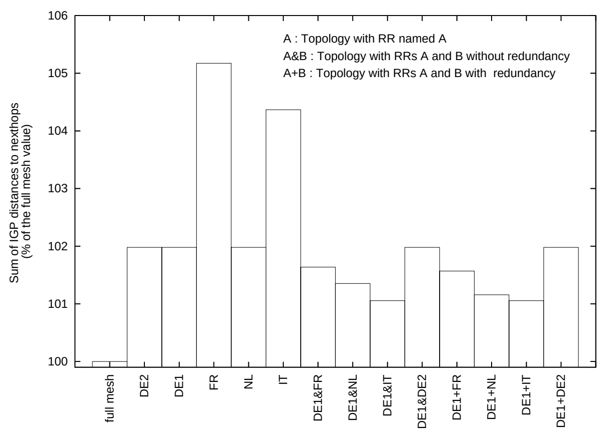

Figure 1: Loss of optimality in the routing tables for each topology, compared to the full mesh value.

In this section we present the results of the application of our evaluation process to the GEANT network [5]. The simulations are made based on the routes announced to the network, after having selected five routers as route reflector candidates. This gives us fourteen topologies to test. Results show that the metric measuring the sum of the IGP length of the paths along all the iBGP sessions is minimal for topologies without redundancy. This is not surprising, as topologies with redundancy require that every client maintains two iBGP sessions, one with each reflector. Variations among topologies of the same type are relatively small and dependant of the physical topology of the AS, so even if these values are interesting for the operator, we won't consider them in this paper. Similarly, the size of routing tables are dramatically decreasing with the introduction of route reflection, showing that this practice allows less memory usage inside routers. Values for topologies with redundancy are roughly twice higher than those for topologies without redundancy, because clients are receiving routes from two reflectors instead of one.

The analysis of results for the sum of the distances to next hops are more interesting. This metric typically measures the differences between the decision process of the reflectors and the one of their clients. Logically, topologies with two reflectors should be better than those with only one, because the two reflectors may make different choices while selecting routes. Figure 1 confirms this tendency, and shows that there are high variations between topologies with one reflector : Routers FR and IT are obviously making a choice that is not optimal for their clients. Differences of values between topologies with two route reflectors are bounded to differences between route choices made by the reflectors. For example, topology with reflectors DE1 and DE2 is less performant than the three others, because those two routers are located close to each other, and are therefore making similar route choices. Among topologies with two reflectors selecting different routes, the best one are those located close to border routers that are source of routes globally preferred by the clients.

We can analyse the value of the metric measuring redundancy of routes, i.e the robustness of the topology, in a similar way. Redundancy is observed when we use topologies with two redundant reflectors, because clients are receiving two routes for each destination and these routes may be different. Figure 2 confirms this tendency. There are large values of route redundancy for three of the four topologies with double route reflection. The fourth one with reflectors DE1 and DE2 has low values because these routers are making similar route choices.

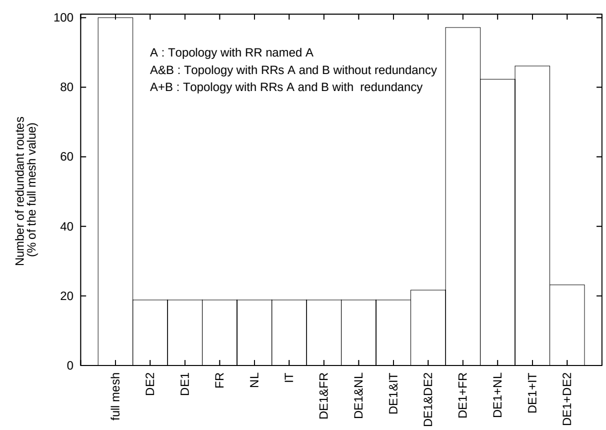

Figure 2: Redundancy inside the routing tables for each topology, compared to the full mesh value.

4. CONCLUSION

The result of our two last metrics have shown that it is usually a good idea to use topologies with two redundant route reflectors and to locate those reflectors in such a way that their route choices are as different as possible. Moreover, our comparison scheme give us a tool allowing an operator to test some iBGP topologies so that he can choose the most efficient one.

This tool can of course be improved. It is for example not yet able to cope with large ASes, for which it can be interesting to consider topologies with more than one level of route reflection. Our heuristic for the choice of route reflectors is also still simple, and we can improve it by using clustering algorithms, so that we choose as route reflectors routers that are in the center of a group of routers instead of in the center of the whole AS. This improvement will for example allow us to consider ASes with routers spread on several continents.

Acknowledgments

This work was partially supported by the E-Next Network of Excellence. We would also like to thank Nicolas Simar for providing the information about the GEANT network.

5. REFERENCES

[1] Y. Rekhter, "A Border Gateway Protocol 4 (BGP-4)", IETF, Internet Draft, 2002, draft-ietf-idr-bgp4-26.txt.

[2] T. Bates, R. Chandra, and E. Chen, "BGP Route Reflection - An Alternative to Full Mesh IBGP", IETF, RFC 2796, April 2000.

[3] http://www.geant.net.

[4] B. Quoitin, C-BGP : http://cbgp.info.ucl.ac.be/.

[5] Tool available at http://totem.info.ucl.ac.be/tools.html.