Interdomain Routing

Project Report

Network Infrastructure improvement proposal To Company A

Team 4:
Zhang Li
Bin Yang
Md. Safiqul Islam
Saurabh Arora

Network Infrastructure Improvement
Interdomain routing Project Report

Saurabh Arora Bin Yang Md. Safiqul Islam Zhang Li
{arora, binyang, islam3, lizhang } @ kth.se

Abstract

Network Infrastructures comprising of high performance and powerful end nodes may operate poorly in environments characterized by very long delay paths and frequent network partitions. These problems are exacerbated by poor design policies and inefficient setup of network architecture. To achieve interoperability between them, we propose a network architecture and related design traffic controlling policies. This report presents the solution for company A's network distributed infrastructure spread across in two cities, to better the enablement of efficient network flow. This report also proposes issues in scalability, performance, robustness and load balancing of the company's network. The model we propose consists of changes in the network infrastructure design as well as implementations of new routing policies based on various routing protocols.

1 Introduction

The performance of IP based corporate networks depends on a wide variety of dynamic conditions. Traffic shifts, equipment failures, planned maintenance, and topology changes in other parts of the Internet can all degrade performance. To maintain good performance, corporate networks must be continually reconfigured for the routing protocols. Network controllers configure BGP to control how traffic flows to neighboring Autonomous Systems (ASes), as well as how traffic traverses their networks. However, because BGP route selection is distributed, indirectly controlled by configurable policies, and influenced by complex interactions with intra-domain routing protocols, it can inadvertently result in degrading network performance.

The current network infrastructure design of company A is suboptimal, as few areas inside the network are growing larger and loosing structure. Moreover, no redundancy is provided with the physical links currently in place. Finally, the only routing protocol running is OSPF, which does not allow us to apply routing policies specific to company A. Therefore, we propose the creation of new areas and also suggest that BGP is applied between the core routers as well as between areas. This way, using BGP attributes, we will be able to enforce the required routing policies, in other words, manipulate the flow of the traffic over different links. In each of the following sections, we will point out the problems with each area, propose and justify our solution to the problem and finally explain how the new solution can be implemented. Finally, in the conclusion we will summarize our work.

2 Network Infrastructure

Company A is divided into 2 geographical locations; Stockholm and Copenhagen. In previous infrastructure, all network run OSPF, each stub area

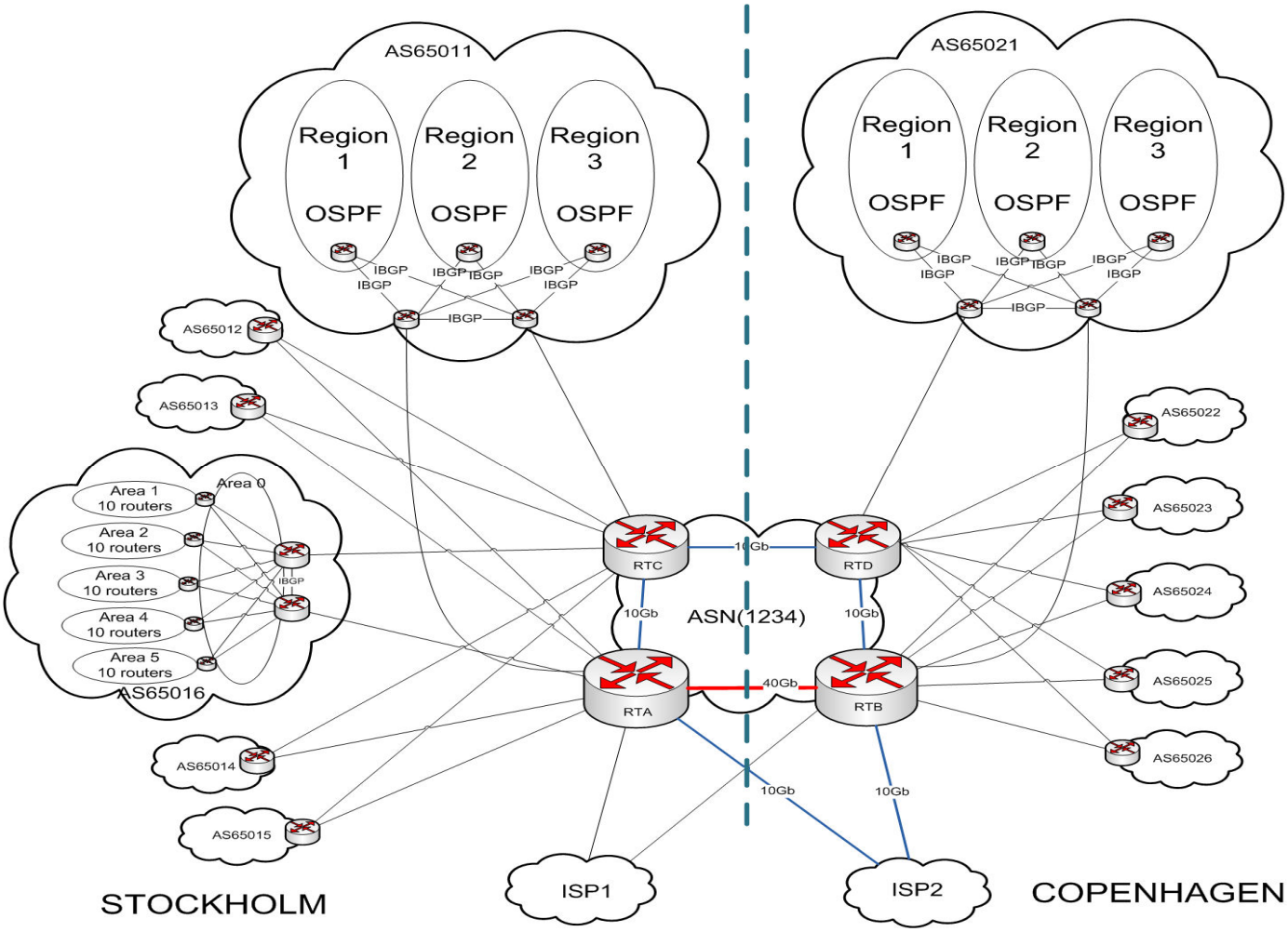

Figure 1. Network Architecture

connects to backbone area 0 by a single link. In this infrastructure, there are lot of routing information in *area 0* and, also no redundancy between *area 0* and *stub areas*. In order to improve whole network infrastructure and have quick convergence and stability, we decide that using EBGP segments into multiple private ASes. We use two high-end core routers and two core routers to compose a new public AS with public AS number *1234*. Each stub area will become private AS. Between private ASes and public *AS 1234*, we run EBGP. There are four routers in public AS, we run OSPF as IGP and run IBGP in each router with logical full mesh. In each private AS, we run respective IGP.

2.1 Public AS 1234

After our improvement, public *AS(1234)* will become core in infrastructure. It is responsible for sending and receiving traffic between private ASs and public AS. It is also responsible for sending and receiving traffic between company A and two ISPs.

Between the two cities, we have reduced the links to only two instead of three links before. One link is 40Gbps, another one is 10Gbps. We will use 40Gbps link between two high-end core routers in order to utilize the available high speed link. Another 10Gbps link as the backup link should be connected between other two core routers. One purpose is that exchanging traffic between company A and ISPs, other purpose is to exchange traffic between two cities. In the new network infrastructure, company has two ISPs. Two high-end core routers will be connected to ISPs with two links by EBGP.

2.2 Private ASes and Policies

In each private AS, there are two links connected to public *AS1234*, instead of signal link connected to *area 0*. We use EBGP to exchange traffic between public AS1234 and each private ASes. Each private AS will advertise an aggregated network to public AS 1234 and learn a default route from public AS1234. So the traffic between private AS and public AS can be controlled by policy like local-preference and MED. On the private AS side, local-preference is used to control outgoing traffic of private AS, making link which is connected to RTA or RTB is primary link. MED is used to control incoming traffic of private AS, making link which is connected to RTA or RTB is primary link.

The best advantage is that we don't need set any policy (for private AS) on public AS. If we use local-preference on public AS side to control incoming traffic of private AS, we need modify these policies in public AS when adding or deleting a private AS.

About exchanging traffic between two cities inside public *AS 1234*, public AS will automatically choose 40Gbps link. Because traffic which enters and exits the public AS will always pass through RTA or RTB in normal conditions.

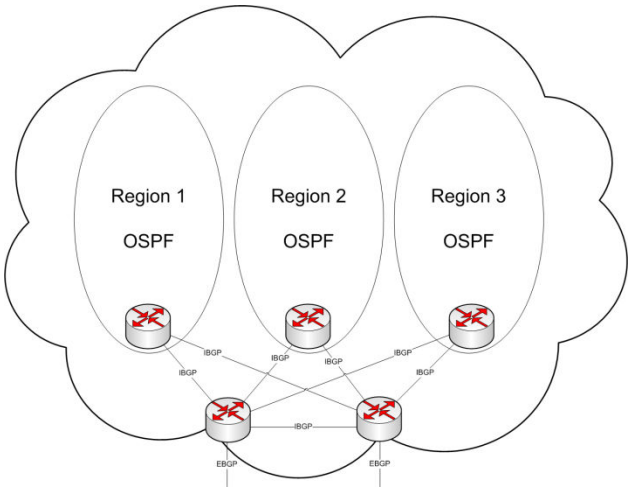

Figure 2. Area 1, 2

So the IGP of the public *AS1234* could choose a shortest path to the exit, which is 40Gbps link. If RTA or RTB goes down, IGP will choose backup link between RTC and RTD.

2.2.1 Area 1 and Area 2

Area 1 and *Area 2* represent internal networks of the headquarters in Stockholm and in Copenhagen respectively. Since the *area 1* and *area 2* have the same architecture, here just take *area 1* as an example. We setup the *area 1* as two tiers deep. As the picture above shows, there are two ABRs (area border routers) which are connected with the core routers with 1-Gbps interface. The *area 1, area 2* with *area 0* communicate via EBGP. The two ABRs communicate with IBGP. In the internal area 1, the area is divided into three *regions*. Each region inside will run *OSPF*. Every region has one router to connect to the two ABRs with IBGP. The router has two links with ABRs. One is primary, and the other is backup. If the primary link broken, the backup link can take the primary's position.

2.2.1 Area 30

Area 30 is the largest single area in this network. It is the area which connects all small branches of company A. There are in total 1200 remote sites, each dual homed, and over 50 routers connecting these dual homed remote sites, all in area 30. If we don't improve the infrastructure, the traffic in this network is too busy to be stable. And there would be a very huge routing table.

In order to avoid the problems mentioned above, we setup this big area as an As. This AS consists of 6 areas. Area 0 is a backbone area, which all other OSPF areas must connect to. All traffic between areas must go through Area 0. In the other 5 areas

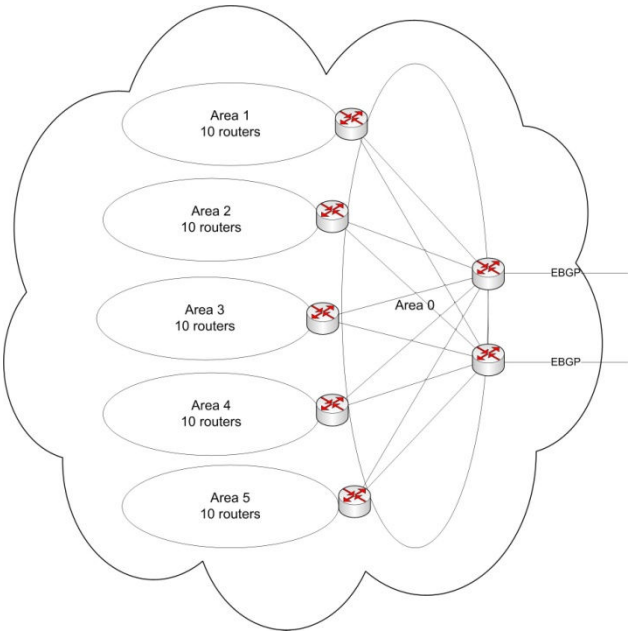

Figure 3. Area 30

which consist of 10 routers for each, they will run OSPF in their own domain. In this As, there are two ABRs which in the area 0 connect outside using EBGP. There are also have 5 routers considered as communicating router to the two ABRs in area 1 to 5 (Each area has one). They also have two links, one is primary and the other is a backup. Once primary down, the backup will take effect. These 5 routers and two ABRs will run OSPF in area 0.

In this way, the network will be more stable since the routing table will be smaller. And the convergence time will be shorter than ever before. If one of the router in Area2 (here is just an example) is down, it would not affect the other areas. The expansibility is also improved.

3 Inter Domain Routing Policies

As a network administrator, we applied inter domain routing policies to the new network infrastructure. Figure 4 shows the connection between ISP1 and ISP2 with our core router.

Company A has two branches. Company A has a public AS number 1234. Also, we divided each region into of each city into private AS and run EBGP in between core router and private AS to apply policy individually to each region. We described it in earlier section.

Both the ISP1 (AS 111) and ISP2 (AS 222) are advertising default route to company A. And Company A have the active network in Stockholm is 192.16.0.0/22, 192.16.4.0/22, 192.16.32.0/20 and active network in Copenhagen is 192.16.16.0/22, 192.16.20.0/22. And we have aggregated address for Stockholm region in 192.16.0.0/21 for the first two addresses and used 192.16.32.0/20 as it is to advertise to ISP1 and ISP2. If we aggregate this address then it will also advertise Copenhagen network and may create a black hole. And for Copenhagen network we aggregated the address to 192.16.16.0/21 network.

3.1 Outgoing traffic

We have used the Local-Preference attribute to achieve our goal for outgoing traffic. We have set local_pref 500 for the default route coming from ISP2 link and local_pref 400 for the default route coming from ISP1 from both the city.

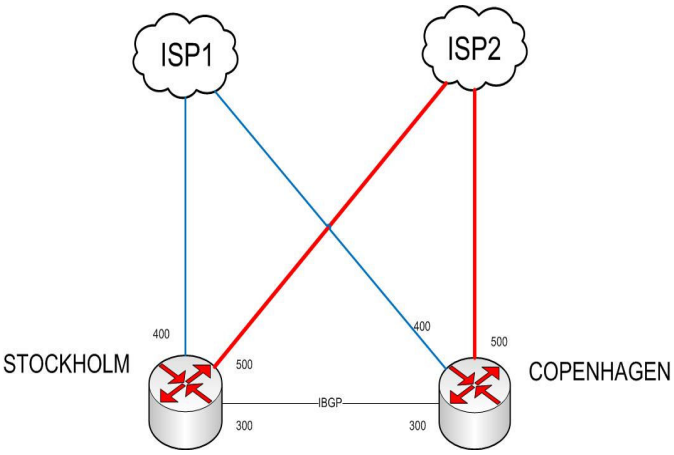

Figure 4. Outgoing traffic behavior for Stocokholm and copenhagen

And we have also applied local_pref 300 for the update coming from Copenhagen router to Stockholm router and Stockholm router to Copenhagen Router. The following diagram illustrates the local_pref settings to different links.

3.1.1 Outgoing Traffic Behavior for Stockholm

According to local_pref as mentioned in *Figure. 4,* for the Stockholm region all outgoing traffic will go via Stockholm ISP2 link. If the link fails then it will choose ISP1 Stockholm link. If both of the links fail then it will send its traffic to Copenhagen link as the local_pref is 300 and Copenhagen ISP2 link will be preferred first and failure of all the links will choose Copenhagen ISP1 link.

3.1.2 Outgoing Traffic Behavior for Copenhagen

For the Copenhagen region all outgoing traffic will go via Copenhagen ISP2 link. If the link fails then it will choose Copenhagen ISP1 link. If both of the links fail then it will send its traffic to Stockholm link as the local_pref is 300 and Stockholm ISP2 link will be preferred first and failure of all the links will choose Stockholm ISP1 link.

3.2 Incoming traffic

We are using AS-Path prepending for the behavior of incoming traffic. The following section will describe the incoming traffic behavior for both the region.

3.2.1 Incoming Traffic Behavior for Stockholm

For Stockholm region to make ISP2 link most preferable we did not prepend any AS-Path. Then we added its as path once to the advertised router of Stockholm. Then we added AS-Path 2 times of Copenhagen for the advertised route of Stockholm to ISP2 and AS-path 3 times for Copenhagen to ISP 1 link.

So, ISP2 link of AS-path will be preferred for the Stockholm advertised network. If this link fails then Stockholm ISP1 link will be preferred because of one AS-path prepending and if both of the links fail then Copenhagen ISP2 link. If all the links fail then Copenhagen ISP1 link will be chosen.

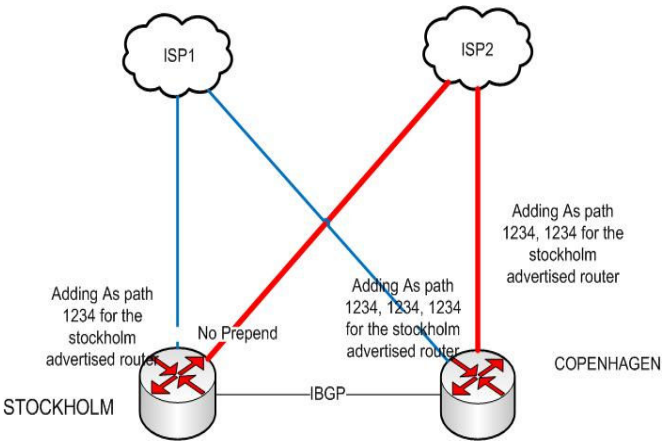

Figure 5. Incoming Traffic - Stockholm

3.2.2 Incoming Traffic Behavior for Copenhagen

For Copenhagen region to make ISP2 link most preferable we did not prepend any AS-Path. Then we added its own as path once to the advertised route of Copenhagen to ISP1 link. Then we added AS-Path 2 times of Stockholm for the advertised route of Copenhagen to ISP2 and AS-path 3 times of Stockholm to ISP 1 link.

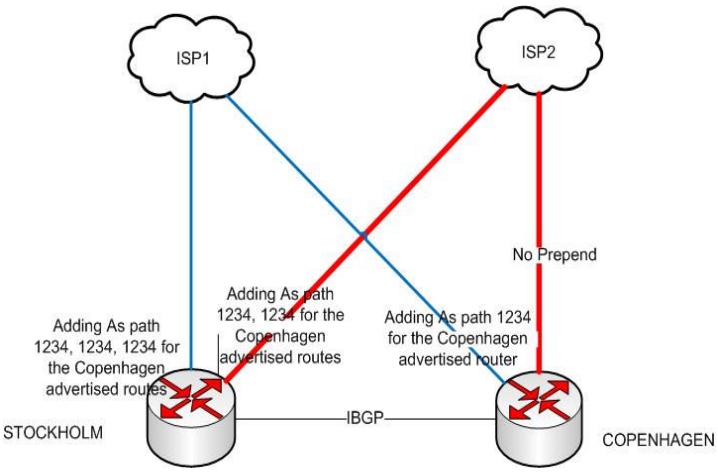

Figure 6. Incoming Traffic Copenhagen

So, ISP2 link of Copenhagen will be preferred for the Copenhagen advertised network. If this link fails then Copenhagen ISP1 link will be preferred because of one AS-path prepending and if both of the links fail then Stockholm ISP2 link. If all the links fail then Stockholm ISP1 link will be chosen.

4 Conclusion

The proposal we have presented is a necessary step for advancing the state of the art of company A's network. We believe that our model and BGP routing policies eliminates crucial problems hindering the performance of company network. We have carefully dealt with various important factors like robustness, redundancy, scalability and performance and paved the way for upcoming future expansions.

5 References

[1] Internet Routing Architectures Second Edition by Sam Halabi (ISBN 1-57870-233-X).

[2] Stril Network's router guide - http://www.tslab.ssvl.kth.se/courses/file.php/20/labs/StrilRouterGuide.pdf

[3] BGP case studies - http://www.tslab.ssvl.kth.se/courses/file.php/20/labs/bgp-toc.pdf

[4] CISCO router guide - http://www.cisco.com/application/pdf/en/us/guest/products/ps5855/c1031/cdccont_0900aecd8019dc1f.pdf

[5] A model of BGP routing for network engineering. N Feamster, J Winick, J Rexford - ACM SIGMETRICS **Performance** Evaluation Review, 2004