IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 10, Issue 1, Ver. II (Jan - Feb. 2015), PP 66-74
www.iosrjournals.org

A new approach to fight against the inter-domain BGP oscillations

[1]Alphonse Binele Abana, [2]Emmanuel Tonye
[1&2](Department of Electrical and Telecommunication Engineering, Polytechnic National Advanced School of Engineering/ University Of Yaoundé 1, Cameroon)

Abstract: *The current inter-domain routing protocol, Border Gateway Protocol (BGP4), has evolved over the past decade and now constitutes a critical part of the Internet infrastructure. Despite impressive progress in characterizing the various ills of the Border Gateway Protocol (BGP), many problems remain unsolved, and the behavior of the routing system is still poorly understood. The BGP routing protocol is used to interconnect the various Internet operators together. BGP enables ASes to make local decisions based on private routing policies, and forms global routes from these local decisions. At a very high level, the execution of BGP requires ASes to continuously make independent greedy route selections, based on their local preferences over routes, and announce these choices to all neighboring ASes. The inconsistencies of these policies can cause route instabilities also called oscillations.*
In this article, we discuss the different theoretical models of bgp protocol fighting against inter-domain routes oscillations (between ASes).
Keywords: *Autonomous System, BGP, Internet, Path, inter-domain oscillations.*

I. Introduction

The substantial complexity of inter-domain routing in the Internet comes from the need to support flexible policies while scaling to a large number of Autonomous Systems. An autonomous system is a set of routers under a single network administration. As an example of AS in Cameroon, we have CAMTEL, Orange, and MTN. Each AS decide its internal routing protocol (RIP, OSPF ...) and its external routing policy. The BGP routing protocol, defined in RFC 4271 [1], is used to interconnect the various Internet operators between them. It is thus possible to model web as a graph where each autonomous system is represented by a node and the edges model the communication links (physical) between the ASes. Consequently, BGP is a protocol which directly influences the Internet. BGP has been designed for two purposes: to best meet the requirements imposed by the operator's routing policies and respect the principle of confidentiality of such policies. Inconsistencies of these policies cause route oscillations. Many works on inter-domain oscillations are found in the literature on the occurrence those of Griffin. Al [7] which is the first to study the root causes of the problem.

In this article, we describe the selection process of a path in a bgp router, and then we describe the inter-domain oscillation phenomenon before presenting some solutions to this problem; among these solutions, we present a new approach to prevent this phenomenon.

II. Path Selection Process In A Bgp Router

Each router learns from its neighbors a path to a destination. A path P received by a router R in an AS v contains the following attributes:

- **LOCAL_PREF**: preference value indicating the classification of the choice of the path P in the local routing policy of the AS v;
- **AS path:** AS sequence along the path to reach the destination of the current AS v;
- **MED:** Used to discriminate links by associating to each a degree of preference when two ASes are interconnected using several links. A small value of the MED indicates a greater preference for the link;
- **Next_hop:** The IP address of the border router "next hop" along the path P. If he traffic of the router R along the path P passes through other routers before leaving AS v, then next_hop is the IP address of the border router that is the exit point of the AS v. If the traffic of the router R along the path P goes directly from R to a neighboring router in another AS, then next_hop is the IP address of the neighboring router.

For each AS, a router receives a path (potentially empty) to reach the destination. From this set of paths, the router must choose the best path and adopt it as its own way. The best path is selected according to the algorithm implemented. If a router adopts a new path, if the best path is not the path chosen before, the router informs all of his peers on the newly chosen path.

III. Inter-Domain Oscillations

Remember that the path preferences are chosen locally at each AS. If the path preferences of a neighbor AS are in conflict with those of the other, it is not possible to maintain a stable path to the destination. That is to say that the paths chosen by some AS oscillate (divergent) continuously even if neither the interconnection graph of AS paths nor choice policies are changed. Many works on inter-domain oscillations are found in the literature on the occurrence those of Griffin. Al [7] which is the first to study the root causes of the problem. He presented a theoretical model of the problem known as the Stable Path Problem (SPP) and provided sufficient conditions for the convergence of ASes. Consider the graph of AS in figure 1 also called "Bad Gadget" [5].

Assume that each AS prefers longer paths relative to the shortest paths. u prefers the longer path (u, v, d) rather than the shortest path (u, d). Let's see it closely.

1- Initially u, v and w choose paths (u, v, d), (v, d), and (w, d), respectively as shown in Fig1 (a)

2. v finds that w selected the path (w, d). Thus v changes his current path (v, d) to the path with greater local preference (v, w, d). This in turn forces v to change its path in (v, d) as shown in Fig1 (c).

3. Finally, u finds that v has chosen the path (v, d). Hence, u changes its path to (u, v, w). This strengthens w to change its path to (w, d), and the system is back to the original state. Convergence to an equilibrium state is very sensitive to the paths rankings. For example, by reversing the ranking of paths at AS u we ensure that the system reaches a steady state.

IV. Solutions To Bgp Oscillations Problem

BGP has many interesting properties that allow it to be always preferred today. Any solution to the problem of divergence or bgp oscillations must therefore maintain these properties. These properties are:

- **Scalability and efficiency:** given the global Internet, the proposed solution must be efficient and scalable;
- **No global coordination:** Since the number of ASes increases, any solution that requires global coordination between ASes cannot be scalable;
- **No restrictions on AS Policies:** Each AS should independently be able to control its routing policies. Thus, the solution should as far as possible avoid the restriction of routing policies;
- **The minimal changes:** due to the widespread deployment of BGP, any proposed solution must change current behavior of the BGP protocol as minimum as possible;
- **Confidentiality constraints:** an AS should not transmit the aggregated information that can allow his neighbor to rebuild its routing policy.

The solutions to the inter-domain oscillations can be divided into 3 categories: the first category requires global coordination between ASes to avoid any conflict of routing policies [9]; The second category establishes convergence by limiting the type of routing policies may adopt an AS [10]; The third category avoids oscillations by detecting conflicts during the execution time. For this third category, there are three approaches:

The first approach announces historic path [11] with each of BGP UPDATE message and conflicts are detected by observing the cycles in the received stories.

The second approach detects conflicts through the diffusion of computations [12] and prevents an AS to choose a path which conflicts with other ASes.

In the third approach, each AS maintains a metric value [13], which grows boundless during the divergence (oscillation). The divergence is avoided by eliminating the policy only when the value of the metric reaches above a certain threshold.

Below, we present solutions in each of the three categories above. Each solution has advantages and disadvantages. At present, some service providers check the statistics of conflicts in routing policies whenever possible, using the global coordination. It is yet to be determined if any of the solutions outlined below will get wide acceptance on the part of service providers.

I.1 Static checking of routing policies

Conflicts in routing policies can be avoided by collecting routing policies of all AS in one place, and analyzing these policies to check for conflicts before they are implemented. An example of this category is given by the project "Routing Arbiter," [9] where Govindan et Al. has developed an inter-domain routing architecture for gathering multiple AS routing policies and checking if there are conflicts. The architecture is to describe routing policies using a common language, and storing them in a global database. A set of software tools are provided to analyze the routing policies in the global database and attempt to find conflicts between them. A disadvantage of this solution is that the ASes are often reluctant to share about their local routing policies with others for privacy reasons. More importantly, Griffin et al. [5] showed that deciding whether a set of routing policies can lead to a divergent behavior is untreatable, more specifically, it is NP-hard.

I.2 Restricting routing policies

Gao et al. [10] proposed a set of guidelines for the choice of routing policies based on the hierarchical structure of commercial relations between the ASes. These relationships include customer relationships, and peer to peer. In general, trade relations are based on the size of the AS. In a customer-supplier relationship, data from customer AS transit through a supplier AS (which is larger than the customer AS) to reach the rest of the Internet. The supplier AS could itself be a client of an even bigger AS. In a peer-to-peer relationship, two similarly sized AS mutually use their network resources to connect to the Internet.

Each AS exports its routes and those learned from his client AS to all supplier ASes. Each AS also exports its routes and those learned from its suppliers and peers to all its customers. Peer AS export their routes and paths learned from their respective clients each of other. The guidelines for the selection of paths are each AS prefers paths via its customer AS than via peer AS or provider AS. These restrictions ensure that the routing policies are conflict-free, so the convergence is ensured.

The advantage of this solution is that it requires no changes to the current BGP protocol. There are some drawbacks, however. Trade relations between the ASes are not always clearly defined and all ASes may not wish to restrict their routing policies that way. Thus, the freedom to define the original BGP routing policies is lost. Also, if the routing policy is poorly accidentally configured on a router, then the conflict may occur.

I.3 Policy Analysis at run time

In this section, we present three solutions that do not restrict routing policies. Instead, they solve the problem of divergence by reading the routing policy conflicts at run. These three solutions assume that each AS is seen as a single router.

I.3.1 Policy analysis at run time via an historical path

The RFC 4271 [1] describes the current operations of the protocol bgp as it is currently implemented in routers. The SPP (Stable Path Problem) gives a simplified view of the instability of the paths in routing protocols such as BGP. Introduced by Griffin et al., it allows focusing on the main points causing instabilities and separate BGP features that play no role in this problem (eg MED, route aggregation ...).

Let G = (V, E) be a graph such as V = the set of ASes and E= the set of BGP links. Each AS determines a list of paths ordered by order of preference. Each AS defines a set "choice" containing all possible paths to a destination. The "best "function will return the route with the highest preference possible. So we have: best (u) = max (choice (u)) (u is an AS). A solution to SPP is an assignment of a path for each AS, such as the preference ranking of these paths is the highest possible.

Griffin [10] developed the SPVP (Safe Path Vector Protocol) which is an abstract simulation of BGP behavior. Three versions of SPVP have been proposed. The first SPVP1 represents the current operation of BGP. The second SPVP2, is an improvement from the previous version. It offers a dynamic management of messages through a historical paths described by a path historic structure used to detect cycles. But these two versions are not reliable in the sense that they do not resolve the instabilities. The third version, SPVP3, is a reliable extension of SPVP2.

The historic of a path is the concatenation of a local event with the historic of the path advertised by the neighbor. Specifically, the historical path is a sequence of path change events. If an AS changes its path from Pold to Pnew at the arrival of an UPDATE message from a neighbor with the historic Hist, then there are two choices: if the AS prefers Pnew instead of Pold, it inserts event (+,Pnew) at the beginning of Hist. Otherwise, if the AS prefers Pold instead of Pnew, it inserts the event (-, Pold) at the beginning of Hist. The figures fig 2, fig 3 and fig 4 summarizes the SPVP1, SPVP2 and SPVP3 Griffin algorithms.

The benefit of path historics is that the path causing the cycle is stored in the background and can be used later to analyze the problem.

When an AS sends a route, it also announces the historic associated to this route. At the reception, the recipients AS must update their own historic. All these operations require lot of memory space and disrupt network performance.

The oscillation is resolved by removing a path from the set of paths allowed to AS, and thus break the cycle in the historic of the path. However, after a sequence of topology changes, the path removed may be essential to maintain connectivity if it becomes the only available path to the destination.

In addition, the act of communicating historic to each path announcement (allowing all AS to obtain global information about other AS) provides an opportunity for an AS to rebuild the policy of their neighbors. Thus, the policy privacy constraint is not met.

I.3.2 Restricting path choice by diffusing computations

In [12], we can observe that if any time an AS changes his path to a destination, he is assured of a path with a local preference value at least as high as that of its current path, then a set of Stable paths is guaranteed to

be reached. In other words, when the local preference value of the chosen path in all AS monotonically increases, the system cannot oscillate.

Since the preferences of the paths are chosen independently at each AS, an additional restriction is necessary to ensure the monotony of LOCAL_PREF values (local preference) as follows: before an AS q adopts a new path, q asks to any other AS p (whose path is currently traversing q) if this path changing in q will lower the local preference value of the current path at p. If it is the case, q refrains from adopting the new path.

Coordination between q and p is via diffusing computation [15] along the routing tree. The routing tree is defined as follows: for each AS p and next neighbor next-hop q along its path to d, consider the link-oriented (p, q). The union of these entire links-oriented on all ASes forms a routing tree. If there is a path from p to q along the routing tree, then p is a descendant of q and q is an ancestor of the p.

The above technique has the advantage of forcing convergence whatever routing policy selected at each AS. While convergence is assured the freedom to define the BGP routing policies is removed. In addition, the protocol may prevent certain sequences of path changes not necessarily causing the divergence of the system. Finally, the diffusing computation causes additional overhead messages.

I.3.3 Policy analysis at run time via a metric

In this part, we present our contribution to find solutions to the problem of bgp inter-domain oscillations. From observations made on the operation of bgp as described by Griffin in his SPVP1 algorithm, we found that if there is oscillation of routes to a destination, then each AS involved in this circuit will announce the same routes to this destination in each cycle while the bad route is not forbidden in the cycle.

So to stop an oscillation we have to ban bad routes in a cycle by limiting multiple advertisements of routes to a threshold value; to stop the oscillations in bgp, we propose to modify the bgp protocol by adding some attributes to take into account the threshold value of multiple advertisements of a route at an AS (these attributes are not sent to neighbors during routes advertisements). Thus, we propose the SPVPOC algorithm (Safe Path Vector Path Occurrences) by adding the following attributes:

- **Path_occurence:** it indicates the number of times a route has already been advertised by a router in an AS. With the exception of direct routes to that destination, the number of routes will be counted.
- **Max_path_occurence:** indicates the maximum number of times a route can be advertised before being banned.

This algorithm is described in fig 5.
Let be the fig 6 representing a scenario of 5 ASes facing the problem of routes instability also called the “Bad Gadget with 5 ASes.
Table 1 presents the state of the system at each stage and Table 2 shows the numbers of occurrence of different paths at each stage.

If the attribute ”Max_path_occurence” is set to 2, then the path 210 at the AS 2 will not be advertised to its neighbors and AS 2 will retain its best path 20 to the destination 0. The system will then converge. Given that the SPVPOC is inspired from the SPVP1 without touching the fundamental properties of bgp described above, then it inherits all the fundamental properties of bgp and avoids oscillations. BGP provides a variable length in UPDATE messages reserved for possible additions of new attributes; and therefore our approach proposes adding two attributes: path_occurence and Max_path_occurence.

V. Comparison of two algorithms: SPVP3 and SPVPOC

We have developed a test bed that simulates different algorithms of Griffin (SPVP1, SPVP2 and SPVP3) and our algorithm SPVPOC.

Consider the Bad Gadget with 4 AS represented in fig 6. At AS 1, the possible best paths allowed to reach one destination in the AS 0 are 10 and 130; At AS 2, the possible best paths allowed to reach one destination in AS 0 are 20 and 210;at AS 3, the possible best paths allowed to reach one destination in the AS 0 are 30 and 320. All the conditions to obtain this situation are resumed in table 3. The input interface looks like the one in fig 8; every algorithm has its input interface for simulation.

After running the SPVP3 algorithm, it detected an oscillation at the AS 2 and prevented the announcement of the 20 which is considered as the path causing the instability. The system converges to step 4 and bgp routing table after convergence is as follows in each AS:

- **AS 1:** 130 is the route to a destination in AS 0;
- **AS2:** No route to a destination in AS 0.
- **AS 3:** 30 is the route to a destination in AS 0;

After running SPVPO algorithm, when at step 5 AS 3 receives the path 20 of its neighbor 2, the path 320 formed at the AS 3 has a local preference (100) higher than the one of its current best path 30 (50). Thus, it

will announce the route 320 to its neighbors. In step 6, the AS 1 will receive the path 320 from its neighbor 3 and the potential best path 1320 is a path forbidden at the AS 1 therefore cannot be accepted; therefore, the AS 1 will leave his current best path 130 since it is no longer reachable; it will therefore take the direct route 10 and will announce it to its neighbors. Our algorithm prevented the execution of step 7. Indeed, in step 7, AS 2 would have received the path 10 from its neighbor AS 1 and would have left his best current path 20 to consider the path 210 because the local preference of the path 210 (100) is greater than that of the path 20 (50). Also, the value of the attribute path_occurence of the path 210 was already at 1 and it would have been increased and would have become 2 in step 7. As the value of Max_path_occurence is fixed to 2, the SPVPOC algorithm will therefore consider that the wrong path is 210 and therefore will not accept it; AS 2 will therefore retain its former best path 20. Hence the system converges at step 7 and bgp routing table is as follows in each AS:

- **AS 1:**best path to a destination in AS 0 is 10;
- **AS 2:** best path to a destination in AS 0 is 20;
- **AS 3:** best path to a destination in AS 0 is 320;

From that simulation, we can see that:

- Operations of SPVP3 can cause connectionless at some ASes: in for example, in the previous simulation, after the convergence, it is impossible to reach one destination in AS 0 from AS 2.
- SPVPOC takes much time than SPVP3 to reach a stable state: in fact, it took 4 steps for SPVP3 to converge while SPVPOC took 7 steps.
- SPVP3 violates the routing policy confidentiality rule of BGP and need to do many changes in actual BGP protocol.
- SPVPOC integrates all the BGP properties list above which are: scalability and efficiency, no global coordination, no restrictions on AS Policies, minimal changes, and confidentiality constraints.

VI. Figures And Tables

I.4 Figures

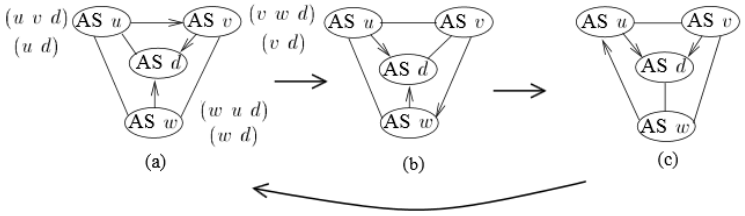

Figure 1: 4 Ases bag gadget evolution

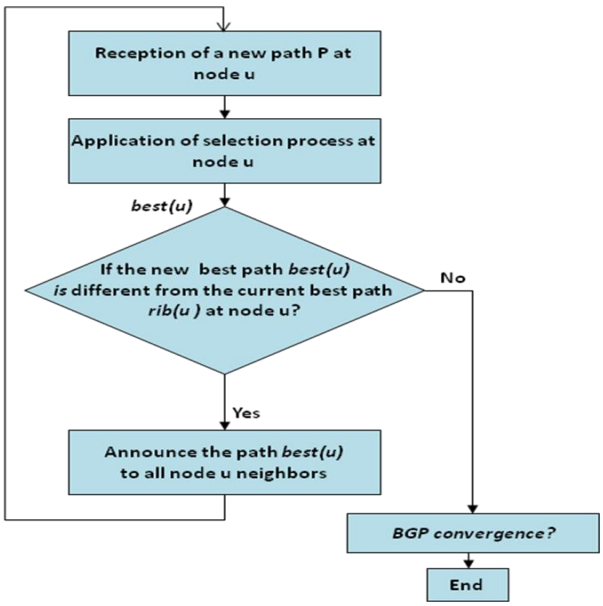

Figure 2: SPVP1 algorithm

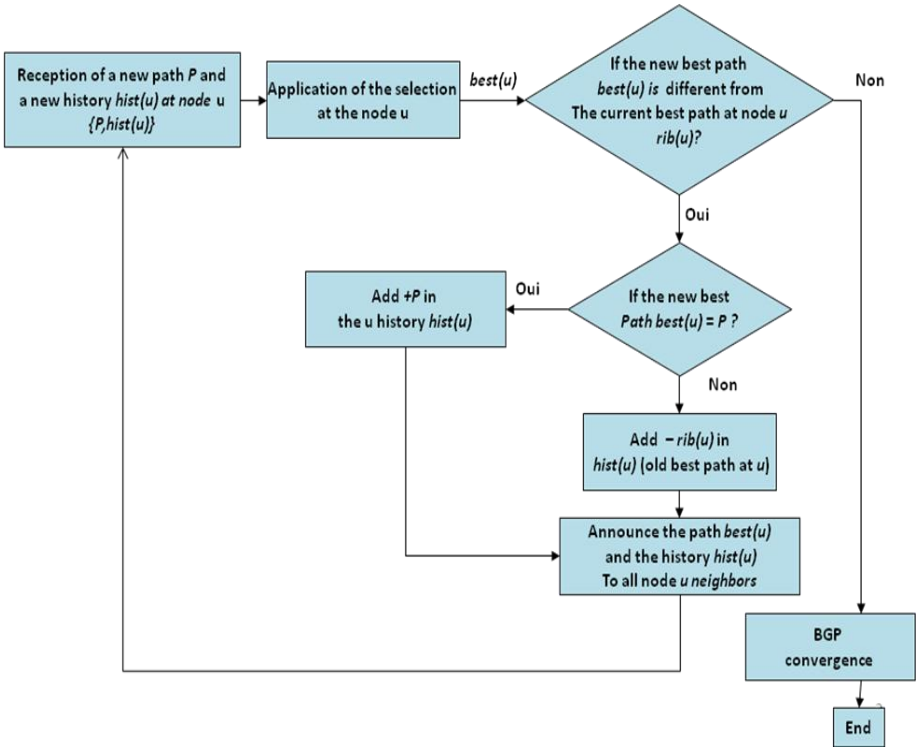

Figure 3: SPVP2 algorithm

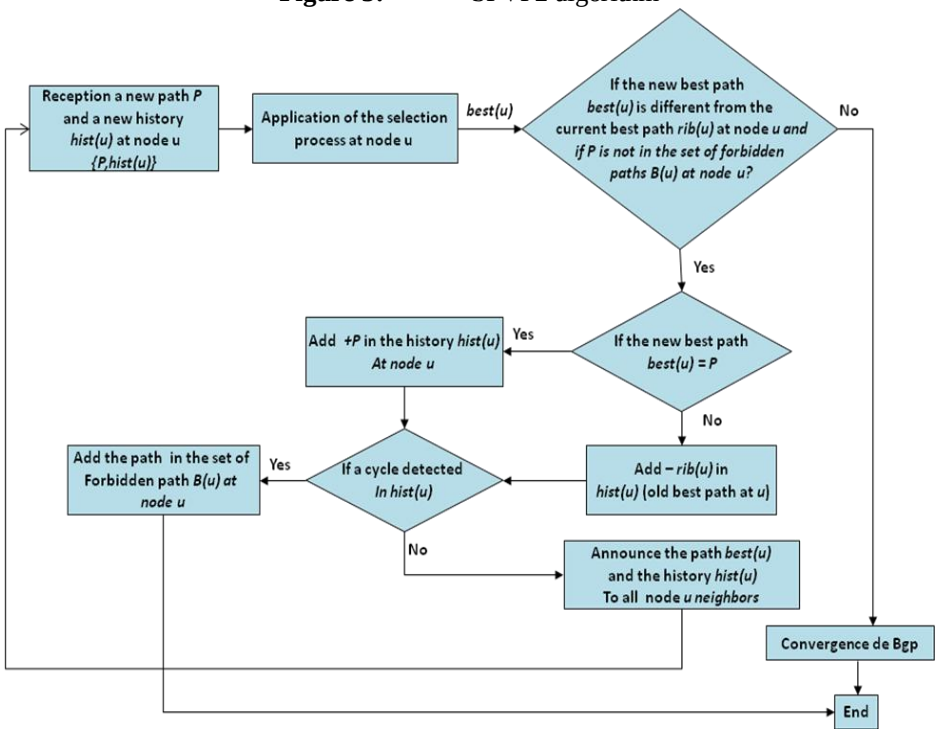

Figure 4: SPVP3 algorithm

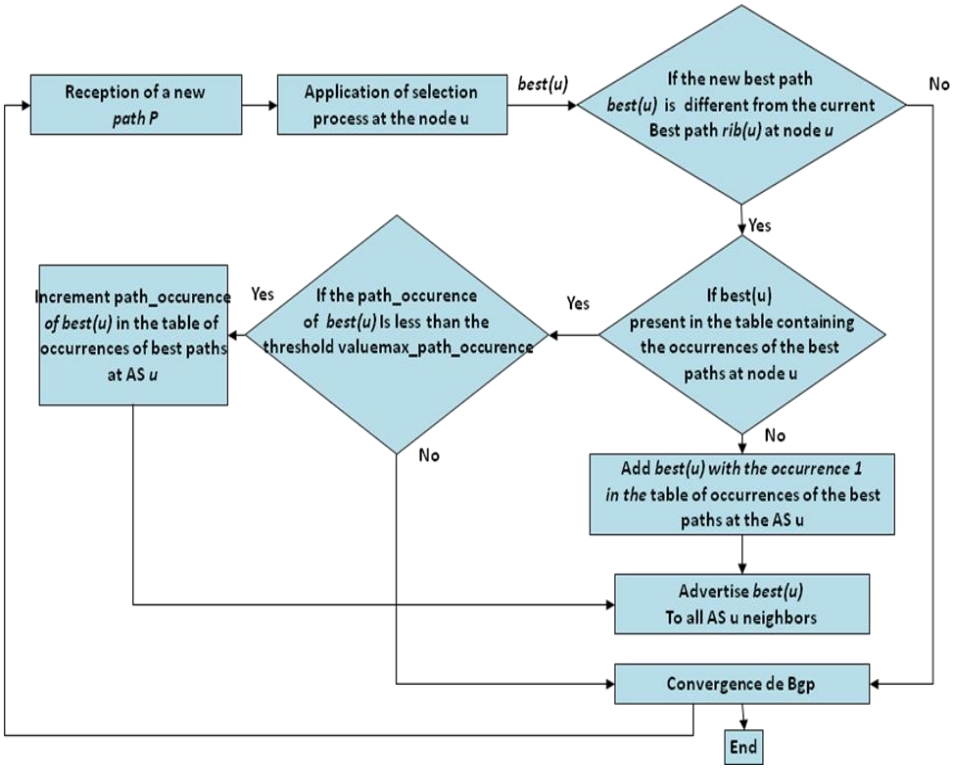

Figure 5: SPVPOC algorithm

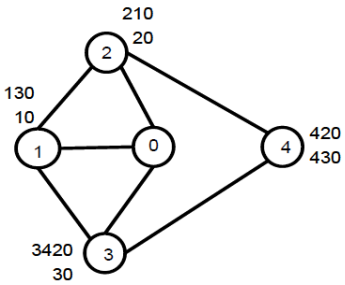

Figure 6: Bad Gadget with 5 ASes

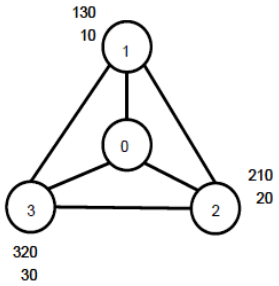

Figure 7: Bad Gadget with 4 ASes

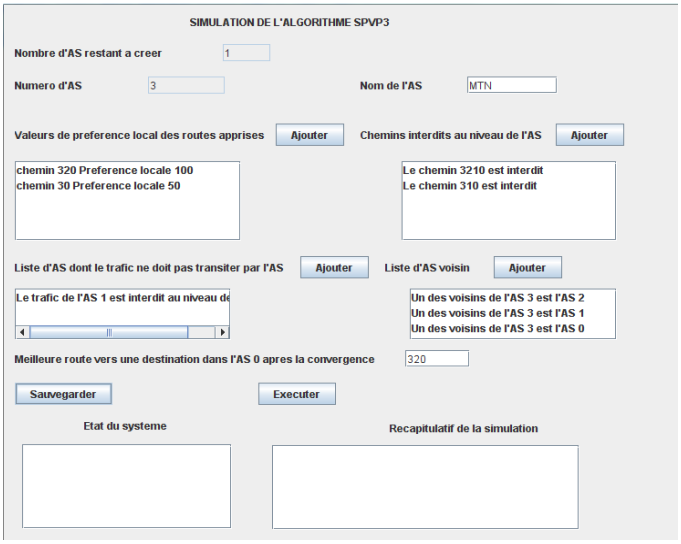

Figure 8: input interface of the SPVP3 algorithm

I.5 Tables

Table 1. System status at each step

u	Steps	0	1	2	3	4	5	6	7	8	9	10
1		10	10	10	10	10	130	130	130	130	10	10
2		20	210	210	210	210	210	20	20	20	20	210
3		3420	3420	3420	30	30	30	30	30	3420	3420	3420
4		420	420	Ø	Ø	430	430	430	420	420	420	420

Table 2. Path occurrences at each step

Steps	Paths	130	10	210	20	30	3420	420	430
0		0	0	0	0	0	0	0	0
1		0	0	1	0	0	0	0	0
2		0	0	1	0	0	0	0	0
3		0	0	1	0	1	0	0	0
4		0	0	1	0	1	0	0	1
5		1	0	1	0	1	0	0	1
6		1	0	1	1	1	0	0	1
7		1	0	1	1	1	0	1	1
8		1	0	1	1	1	1	1	1
9		1	1	1	1	1	1	1	1
10		1	1	2	1	1	1	1	1

Table 3. Bad Gadget of 4 ASes data

AS number	**BGP routing table**		**Forbidden paths**	**Neighbors AS**	**AS whose traffic is forbidden to transit**	**Curent best path**
	Possible best paths	**Local preferences**				
1	10	50	120	0 ;2 ;3	2	10
	130	100	1320			
2	20	50	230	0 ;1 ;3	3	20
	210	100	2130			
3	30	50	310	0 ;1 ;2	1	320
	320	100	3210			

VII. Conclusion

BGP is an important protocol in that it is the heart of the Internet functioning as it connects the Internet operators ASes between them. BGP suffers from divergence problems due to the free application of routing policies in every internet operator. In this article, we discussed the various BGP inter-domain divergence problems, then we presented some existing solutions before offering our approach that boils down to the proposition of a new algorithm for bgp named SPVPOC (Safe Path Vector Path Occurrence). This solution inherits all the properties of the current BGP and is more stable. A perspective would be to study the behavior of this protocol in the context of intra-domain oscillations (iBGP).

References

[1]. REKHTER,Y.,LI,T.,AND HARES, S. A Border Gateway Protocol 4 (BGP-4). Internet Draft draft-ietf-idr-bgp4-25.txt, September 2004.

[2]. Timothy G. Griffin, F. Bruce Sherpherd, and Gordon Wilfong. Policy disputes in pathvector protocols. Proc. 7th Int. Conf. Network Protocols (ICNP'99), pages pp. 21–30,1999.

[3]. Kannan Varadhan,Ramesh Govindan and Deborah Estrin. Persistent route oscillations in inter-domain routing. Elsevier Science B.V., 2000. [http://research.cens.ucla.edu/people/estrin/resources/journals/2000jan-Varadhan-Govindan-Persistent.pdf]

[4]. Jerome BARBOU, Alpha Amadou DIALLO, Greg´ ori FABRE, Damien FONTAINE, Intissar GALL -Universite Paris VI - M2-Rout - Routage Interdomaines : Limites et solutions http://tounkan.free.fr/Web/presentations/rout_pres01.pdf

[5]. T. G. Griffin, F, B. Shepherd, and G. Wilfong, "The stable paths problem and interdomain routing,"IEEE/ACM Trans. Networking, vol. 10 no. 2,pp. 232-243,2002

[6]. R. Musunuri Texas Univ., Dallas, TX, USA, "An overview of solutions to avoid persistent BGP divergence", in IEEE Network: The Magazine of Global Internetworking vol 19, 2005,pp28-34 [http://www.utdallas.edu/~jcobb/PublishedPapers/Journals/IEEE-Network-BGP/musunuri_final.pdf]

[7]. T. G. Griffin, F, B. Shepherd, and G. Wilfong, "Policy disputes in path vector protocols,"in Proc. of IEEE ICNP conference, 1999, pp. 21-30.

[8]. T. G. Griffin, F, B. Shepherd, and G. Wilfong, "An analysis of BGP convergence properties,"in Proc. of INFOCOM conference,, 1999, pp. 277-288.

[9]. IEEE, "A method to eliminate BGP divergence based on AS relationships", in Circuits,Communications and System (PACCS), Second Pacific-Asia Conference on vol 1, 2010, pp 31-34.

[10]. L. Gao and J. Rexford, "Stable Internet routing without global coordination," IEEE/ACM Trans. Networking, vol. 9, no. 6, pp681-692,2001

[11]. T.G. Griffin, F.B. Shepherd, and G.Wilfong, "A safe path vector protocol," in Proc. of INFOCOM conference,2000,pp. 490-499

[12]. J.A Cobb, M.G. Gouda, and R. Musunuri, "A stabilizing solution to the stable paths problem," in Proc. of Symp. on self-stabilizing systems, Springer-Verlag Lecture Notes in Computer Science, vol. 2704, 2003, pp169-183

[13]. J.A. Cobb and R. Musunuri, "Convergence of inter-domain routing," in Proc. of IEEE GLOBECOM conference, 2004, pp1353-1358

[14]. J.J Garcia-Lunes-Aceves, "Loop-free routing using diffusing computations," IEEE/ACM Trans. Networking, vol11 no.1, pp130-141, 1993.