Stability metrics and criteria for path-vector routing

Dimitri Papadimitriou
Alcatel-Lucent Bell Labs
Antwerp, Belgium
E-mail: dimitri.papadimitriou@alcatel-lucent.com

Albert Cabellos-Aparicio, Florin Coras
Technical University of Barcelona
Barcelona, Spain
E-mail: {acabello,fcoras}@ac.upc.edu

***Abstract*—Since so far, most studies on path-vector routing stability have been conducted by means of ad-hoc analysis of Border Gateway Protocol (BGP) data traces. None of them consider the specification of an analytic method including the use of stability metrics for the systematic analysis of BGP traces and associated meta-processing for determining the local state of the routing system. In this paper, we define a set of stability metrics that characterize the local stability properties of path-vector routing such as BGP. By means of these metrics, we derive a stability decision criterion that can be applied during the BGP route selection process. Results obtained using real BGP datasets show that 90% of the routes are not affected by a path length increase when selected based on this criterion. Moreover, among the remaining 10%, a significant fraction of the routes is covered by a path length increase of one-hop. These results corroborate the assumption that enforcing stability would not come at the detriment of increasing the stretch of the routing paths.**

Keywords-component; path-vector, routing, stability, metrics

I. INTRODUCTION

Prominent research efforts to understand Border Gateway Protocol (BGP) instability led to classify them as policy- or protocol-induced to account for the distinction between protocol operations and the inherent behavior of the underlying path-vector routing algorithm.

Policy-induced instabilities: solving the routing stability problem consistently with planned BGP routing policy requires to prevent and/or to eliminate conflicting policy interactions, in particular those leading to unintended unstable routing states. Griffin et al.'s seminal work [1] modeled BGP as a distributed algorithm for solving the Stable Paths Problem, and derived a general sufficient condition for BGP stability, known as "No Dispute Wheel". This sufficient condition guarantees the existence of a stable solution to which BGP always converges. Informally, this sufficient condition allows nodes to have more expressive and realistic preferences than always preferring shorter routes to longer ones. The game theoretic approach introduced in [2] relies on the best-reply BGP dynamics: a convergence game model in which each Autonomous System (AS) is instructed to continuously execute the following actions: i) receive update messages from BGP peering nodes announcing their routes to the destination, ii) choose a single peering node whose route is most preferred to send traffic to, iii) announce the new route to peering nodes. However, as proved in [2], best-reply BGP dynamics is not incentive-compatible even if No Dispute Wheel condition holds: even if all but one AS are following the BGP rules, the remaining AS may not have the incentive to follow them. Interestingly, as demonstrated in [2], incentive compatibility of best-reply BGP dynamics requires combining an additional global condition (Route Verification) together with the "No Dispute Wheels" to guarantee stability. Consequently, all known conditions for global stability define sufficient but not necessary conditions (checking them is an NP-hard problem and enforcing them requires a global deployment of an additional mechanism); on the other hand, local instability effects have yet to be characterized.

Protocol-induced instabilities: BGP, the inter-AS path-vector routing protocol of the Internet, is prone to Path Exploration, phenomenon characterizing any routing protocol that relies on the path-vector algorithm. Indeed, BGP routers may announce as valid, routes that are affected by a topological change and that will be withdrawn shortly after subsequent routing updates. This phenomenon is the main reason for the large number of routing updates received by BGP routers which exacerbate the inter-domain routing system instability and processing overhead [3]. Both result in delaying BGP convergence time upon topology change/failure [4]. Several mitigation mechanisms exist to partially limit the effects of path exploration; however, none actually eliminate them. Hence, BGP is intrinsically subject to instability.

BGP stability has also been reported in RFC4984, outcome of the Routing and Addressing Workshop held by the Internet Architecture Board (IAB), remains a key criterion to be met by the Internet routing system. It is also important to underline that the dynamics of the Internet routing system determines the resource consumption of routing engines, in particular, in terms of memory and CPU. System resource consumption depends on the size of the routing state space but also on the number of BGP peering relationships between routers. Indeed, the increasing dynamics of the exchanges of routing information updates between all BGP peerings, increases the memory and CPU requirements for the operations of the routing protocol. The objectives for investigating path-vector routing stability are to 1) Develop a method to systematically process and interpret the data part of BGP routing information bases in order to identify and characterize occurrences of BGP routing system instability from its routing paths properties; 2) Define a consistent set of stability metrics and related processing methods to better understand the BGP routing system's stability; 3) Exploit some of these metrics as route selection criteria. Overall, the proposed analytical method aims to bring rigor and consistency to the study of the stability properties of routing paths as locally experienced by routers.

This paper is structured as follows. Section II provides an overview on prior work concerning the BGP routing system stability. In Section III, we define the proposed routing stability metrics and detail the corresponding computational procedures. In this section, we also derive a stability decision criterion that can be applied during the BGP route selection process. In Section IV, we document the measurement methodology and the BGP datasets considered to evaluate the applicability of these metrics. Section V reports on the measurement results and analysis the applicability of the proposed stability-based criterion using real BGP datasets. Finally, Section VI draws conclusion from this study and outlines possible future work.

II. Prior Work

Beside the references cited in Section I, there have been numerous studies of BGP dynamics properties over the years. Work began in the early 1990s on an enhancement to the BGP called Route Flap Damping (RFD). The purpose of RFD was to prevent or limit sustained route oscillations that could potentially put an undue processing load on BGP. At that time there was a belief that the predominate cause of route oscillation was due to BGP routing sessions going up and down because they were being carried on circuits that were themselves persistently going up and down [3]. This would result in a constant stream of route updates from the affected BGP sessions that could propagate through the entire network due to the network's flat addressing architecture. The first version of the RFD algorithm specification appeared in October 1993, updates and revisions lead to the publication of RFC 2439 in November 1998.

Mao et al. [5] published in August 2002 a paper that discussed how the use of RFD, as specified in RFC 2439, can significantly slowdown the convergence times of relatively stable routing entries. This abnormal behavior arises during route withdrawal from the interaction of RFD with "BGP path exploration" (in which in response to path failures or routing policy changes, some BGP routers may try a sequence of transient alternate paths before selecting a new path or declaring the corresponding destination unreachable). Bush et al. [6] summarized the findings of Mao et al. [5] and presented some observational data to illustrate the phenomena. The overall conclusion of this work was to avoid using RFD so that the overall ability of the network to re-converge after an episode of "BGP path exploration" was not needlessly slowed.

More recently solutions such as the enhanced path vector routing protocol EPIC [7] propose to add a forward edge sequence numbers mechanism to annotate the AS paths with additional “path dependency’’ information. This information is combined with an enhanced path vector algorithm to limit path exploration and to reduce convergence time in case of failure. EPIC shows significant reduction of convergence time and the number of messages in the fail-down scenario (a part of the network is disconnected from the rest of the network) but only a modest improvement in the fail-over scenario (edges failures without isolation). The main drawback of EPIC is the large amount of extra information stored at the nodes and the increase of the size of messages. Another solution, BGP with Root Cause Notification (RCN) [8] proposes to reduce the BGP convergence delay by announcing the root cause of a link failure location. This solution also offers a significant reduction of the convergence time in the fail-down scenario. However, the convergence time improvement achieved with RCN is modest on the Internet topology compared to legacy BGP (in the fail-over scenario). More advanced techniques such as the recently introduced Path Exploration Damping (PED) [9] augments BGP for selectively damping the propagation of path exploration updates. PED selectively delays and suppresses the propagation of BGP updates that either lengthen an existing AS Path or vary an existing AS-path without shortening its length.

All these approaches try to mitigate the effects and/or to accelerate the convergence of the BGP routing entries after occurrence of a perturbation event, but none of them ask the fundamental question why selecting a route subject to path exploration at first place. The answer is essentially because these mechanisms rely on the network-wide quality criteria that are primarily based on the spatial properties of the AS-path.

III. Routing Stability and Metrics

A. Preliminaries

The autonomous system (AS) topology of the routing system is described as a graph G = (V,E), where each vertex (or node) $u \in V$, $|V| = n$, represents an AS, and each edge $e \in E$, $|E| = m$, represents a link between an AS pair denoted (u,v), where $u, v \in V$. At each node $u \in V$, a route r per destination d ($d \in D$) is selected and stored as an entry in the local routing table (RT) whose total number of entries is denoted by N, i.e., $|RT| = N$. At node u, a route r_i to destination d at time t is defined by $r_i(t) = \{d, (v_k=u, v_{k-1},\ldots,v_0=v), A\}$ with $k > 0 \mid \forall$ j, $k \geq j > 0$, $\{v_j, v_{j-1}\} \in E$ and $i \in [1,N]$, where $(v_k=u, v_{k-1},\ldots,v_0=v)$ represents the AS-path, v_{k-1} the next hop of v along the AS-path from node u to v, and A its attribute set. Let $P_{(u,v),d}$ denote the set of paths from node u to v towards destination d where each path p(u,v) is of the form $\{(v_k=u, v_{k-1},\ldots,v_0=v), A\}$. A routing information update leads to a change of the AS-path $(v_k, v_{k-1},\ldots,v_0)$ or an element of its attribute set A. Next, a withdrawal is denoted by an empty AS-path (ε) and $A = \varnothing$: $\{d,\varepsilon,\varnothing\}$. According to the above definition, if there is more than one AS-path per destination d, they will be considered as multiple distinct routes.

B. Routing Stability

The stability of a routing system is characterized by its response (in terms of processing of routing information) to inputs of finite amplitude. Routing system inputs may be classified as i) internal system events such as changes in the routing protocol configuration or ii) external events such as those resulting from topological changes. Both types of events lead to the exchange of routing information updates (or simply routing updates) that may result in routing states changes. Indeed, BGP and in general any path-vector routing, does not differentiate routing updates with respect to their root cause, their identification (origin), etc. during its selection process.

Definition 1: Let $r_i(t)$ represent the route r_i at some time t as stored in the routing table (RT). At time t+1, $r_i(t+1) = r_i(t) \oplus \Delta r_i(t+1)$, where $\Delta r_i(t+1)$ accounts for all changes experienced by the route r_i from time t to t+1.

This research work is conducted and funded by the European Commission through the EULER project (Grant No.258307) part of the Future Internet Research and Experimentation (FIRE) objective of the Seventh Framework Programme (FP7).

Definition 2: Let RT(t) represent the routing table at some time t. At time t+1, RT(t+1) = RT(t) ⊕ ΔRT(t+1) where, RT(t) is the set of routes that experience no change between time t and t+1, and ΔRT(t+1) accounts for all route changes (additions, deletions, and changes to previously existing routes) between time t and t+1.

The magnitude of the output of a stable routing system should be small whenever the input is small. That is, a single routing update shall not result in output amplification. Equivalently, a stable system's output will always decrease to zero whenever the input events stop. A routing system, which remains in an unending condition of transition from one state to another when disturbed by an external or internal event, is considered to be unstable. In this context, provide means for measuring the magnitude of the output is the main purpose of the metric referred to as "stability of the selected route". For this purpose, we define the criteria for qualifying the effects of a perturbation on the local routing table entries so as to locally characterize the stability properties of the routing system. More precisely, let |ΔRT(t+1)| be the magnitude of the change to the routing table (RT) between time $t = t_0 + k$ to $t + 1 = t_0 + (k+1)$, where t_0 is the starting time of the measurement sequence, and k the integer that determines the number of Minimum Routing Advertisement Interval (MRAI) that have elapsed since the starting time of the measurement sequence. The MRAI determines the minimum amount of time that must elapse between an advertisement and/or withdrawal of routes to a particular destination by a BGP speaker to a peer. The MRAI does not limit the rate of the route selection process but only the rate of route advertisements. Hence, using the MRAI as time unit ensures to record at most one routing update per destination (per BGP peer) per sampling period. Using these definitions, we distinguish three different equilibrium states for the routing table:

Definition 3: when disturbed by an external and/or internal event, a RT is considered to be *stable* if: $|\Delta RT(t+1)| \leq \alpha$, $t \rightarrow \infty$, where $\alpha > 0$ is small. If this condition is met, the routing system (as locally observed) returns to its initial equilibrium state, and is considered to be (asymptotically) stable.

Definition 4: when disturbed by an external and/or internal event, a RT is considered to be *marginally stable* if: $\alpha < |\Delta RT(t+1)| \leq \beta$, $t \rightarrow \infty$, where $\beta > 0$ is small, $\alpha < \beta$. If this condition is met, the routing system (as locally observed) transitions to a new equilibrium state, and is considered to be marginally stable.

Definition 5: when disturbed by an external and/or internal event, a RT is considered to be *unstable* if: $|\Delta RT(t+1)| > \beta$, $t \rightarrow \infty$. If this condition is met, the routing system (as locally observed) remains in an unending condition of transition from one state to another, and is considered to be locally unstable

The actual values of the parameters α and β depend on several factors. Among them, the MRAI value and the integer k that determines the number of MRAI that have elapsed since the beginning of the observation sequence. Other factors influencing these parameters are explained in Section III.C.

Note that a similar reasoning to the one applied for the Loc_RIB stability (that corresponds to the BGP routing table) can be applied to the Adj_RIB_In, which stores the incoming routes from neighbors. It is also interesting to measure the instability induced by the BGP selection process itself.

C. *Stability Metrics*

To measure the degree of stability of the Loc_RIB, and the Adj_RIB_In, the following stability metrics are introduced.

The *stability* $\varphi_i(t)$ of the selected route $r_i(t)$ characterizes the stability of the route $r_i(t)$ ($i \in [1,|D|]$) stored at time t in the Loc_RIB (|Loc_RIB| = N). The value $\varphi_i(t)$ quantifies the magnitude of the change(s) experienced by the route r_i from time $t = t_0+k$ to time $t+1 = t_0+(k+1)$, where t_0 is the starting time of the measurement sequence (time units are counted by default in terms of MRAI), and the integer k accounts for the number of MRAI times that have elapsed since the starting time of the measurement sequence. This metric quantifies thus the magnitude of the change(s) experienced by the route r_i with a periodicity determined by the MRAI time. This metric can be computed by using the procedure described in Fig.1. Upon creation of a new routing table entry associated to the route r_i, the value $\varphi_i(t)$ is initialized together with the parameters α and β (see Section III.B). These parameters can be derived from this procedure on a per individual route basis.

```
/* Initialization when route ri is created */
φi(t) ← 0
αmin,i = αi ← 0
βmax,i = βi ← 0

/* Measurement during k * MRAI time units */
While k > 0
if Δri(t+1) ≠ 0
   /* ri experiences an AS-path change
      or ri experiences an attribute change */
then φi(t+1) ← φi(t) + 1
     βi ← φi(t+1)
     if βmax,i < βi then βmax,i ← βi
     end if
else /* ri experiences no change: Δri(t+1)=0 */
     if φi(t) > 0
     then φi(t+1) ← φi(t) - 1
          αi ← φi(t+1)
          if αi > αmin,i then αmin,i ← αi
          end if
     else φi(t+1) ← 0
     end if
end if
k ← k - 1
end k loop
```

Figure 1. Stability of individual routes

As described in Fig.2, the computation of the stability metric for an entire routing table can then be derived form the stability of its individual routes. Let $|\Delta r_i(t+1)|$ denote the magnitude of change in terms of stability as experienced by a single route r_i from time t to t+1. The set of values $|\Delta r_i(t+1)|$, i

$\in$ [1,N], are then used to compute the value $|\Delta RT(t+1)|$ defined as the magnitude of change in terms of stability for the entire routing table from time t to t+1. Moreover, $|\Delta RT(t+1)|$ can be normalized so that $0 \leq |\Delta RT(t+1)| \leq 1$, where 0 implies perfect stability, and 1 indicates complete instability.

```
For i=1 to N
   /* N = total number of routes in RT(t+1) */
   if ri(t+1) is a new route
      ∨ [φi(t) = 0 ∧ φi(t+1) = 0]
   then |Δri(t+1)| ← 0
   else if φi(t+1) > φi(t)
        then |Δri(t+1)| ← [φi(t)+1]/[φi(t+1)+1]
        else if φi(t+1) ≤ φi(t)
             then |Δri(t+1)| ← φi(t+1)/φi(t)
             end if
        end if
   end if
end i loop

μ = |ΔRT(t+1)| ← Σi |Δri(t+1)| / N
σ² ← Σi (Δri(t+1) - |ΔRT(t+1)|)² / N
```

Figure 2. Stability metric computation for a set of routing entries

The *most stable route* in the Adj_RIB_In (|Adj_RIB_In| = M) quantifies the relative stability between incoming routes to the same destination d as learned from all upstream BGP peers (i.e., downstream from the point of view of the AS-path towards destination d) and the one amongst them determined as the most stable at time t. For this purpose, let $W_u \subset V$ denote the set of node's u BGP peers, $|W_u| = W \leq M$, and w one of its elements such that $(u,w) \in E$. Let $\varphi_{i,j}(t)$ denote the stability of the route $r_i(t)$ to destination d as received by the peering router j ($j \in [1,W]$) at time t. At node u, $r'_{i,stable}(t)=\min\{\varphi_{i,j}(t), \forall j \in [1,W] \mid \{(v_k=u,v_{k-1}=w,\ldots,v_0=v),A\} \in P_{(u,v);d}, \forall w \in W_u\}$ defines –independently of the BGP route selection rules– the selectable route that is the most stable for destination d at time t. Next, we define $\Delta\varphi_i$ as the relative measure stability $\varphi_{i,j}$ of the route r_i at time t+1 for destination d with respect to the stability $\varphi_{i,stable}$ of the most stable route $r'_{i,stable}$ at time t for the same destination d.

```
For i=1 to N
    /* |dest. in Adj_RIB_In| = |Loc_RIB| */
    For j=1 to |Wu|
        /* number of peers for ith dest. */
        Δφi,j(t+1)←[φi,j(t+1)+1]/[φi,stable(t)+1]
    end j loop
    ΔΦi(t+1) ← Σj Δφi,j(t+1)/|Wu|
end i loop

μ = ΔΦ(t+1) ← Σi ΔΦi(t+1) / N
σ² ← Σi (ΔΦi(t+1) - ΔΦ(t+1))² / N
```

Figure 3. Most stable route

The *best selectable route* from the Adj_RIB_In quantifies the relative stability between incoming routes to the same destination d as learned from all upstream peers and the one amongst them selected by BGP at time t as the best route (thus, following BGP route selection rules). As described in Fig.4, the computational procedure is similar to the one depicted in Fig.3, if one replaces $\varphi_{i,stable}$ by $\varphi_{i,selected}$ during the computation of the $\Delta\varphi_{i,j}$ terms.

```
For i=1 to N
    /* |dest. in Adj_RIB_In| = |Loc_RIB| */
    For j=1 to |Wu|
        /* number of peers for ith dest. */
        Δφi,j(t+1) ← [φi,j(t+1)+1]/[φi,selected(t)+1]
    end j loop
    ΔΦi(t+1) ← Σj Δφi,j(t+1)/|Wu|
end i loop

μ = ΔΦ(t+1) ← Σi ΔΦi(t+1) / N
σ² ← Σi (ΔΦi(t+1) - ΔΦ(t+1))² / N
```

Figure 4. Best selectable route

The *differential stability* between the most stable route in the Adj_RIB_In and the selected route stored in the Loc_RIB for the same destination d characterizes the stability of the currently selected routes for a given destination d against most stable routes as learned from upstream neighbors. This metric provides a measure of the stability of the learned routes compared to the stability of the currently selected route. A variant of this metric, denoted $\delta\varphi_i(t)$, $i \in [1,|D|]$, characterizes the stability of the newly selected path p*(u,v) at time t for destination d against the stability of the path p(u,v) that is stored as time t in the Loc_RIB for destination d and that would be replaced at time t+1 by the path p*(u,v): $\delta\varphi_i(t) = \varphi_i(t) - \varphi_i{*}(t)$. In turn, if the differential stability metric $\delta\varphi_i(t) > 0$, then the replacement of route $r_i(t)$ by the route $r_i{*}(t)$ increases the stability of the route to destination d; otherwise, the safest decision is to keep the currently selected route $r_i(t)$ stored in the Loc_RIB.

Application of the differential stability metric $\delta\varphi_i$ during the BGP selection process would prevent replacement (in the Loc_RIB) of more stable routes by less stable ones but also enable selection of more stable routes than the currently selected routes. However, for this assumption to hold, we must also prove the consistency of the stability-based selection with the existing preferential-based route selection model that relies on a path ranking function (i.e., a non-negative, integer-value function λ_u, defined over $P_{(u,v),d}$, such that if $p_1(u,v)$ and $p_2(u,v) \in P_{(u,v),d}$ and $\lambda_u(p_1) < \lambda_u(p_2)$ then $p_2(u,v)$ is said to be preferred over $p_1(u,v)$). The route selection problem is consistent with the stability function $\delta\varphi(t)$, if $\forall$ $u \in V$ and $p_1(u,v)$ and $p_2(u,v) \in P_{(u,v),d}$ (1) if $\lambda_u(p_1) < \lambda_u(p_2)$ then $\delta\varphi(t) = \varphi_1(t) - \varphi_2(t) \geq 0$ and (2) $\lambda_u(p_1) = \lambda_u(p_2)$ then $\delta\varphi(t) = 0$. We show in [10] that if $p_1(u,v)$ and $p_2(u,v) \in P_{(u,v),d} \wedge p_2(u,v)$ is embedded in $p_1(u,v)$, then the route selection problem is consistent with the stability function $\delta\varphi$ and the route selection is not stretch increasing. By stretch decreasing, we mean here that the length $\rho_i{*}(t)$ of the path p*(u,v) (measured in terms of number of AS hops in case of BGP route) associated at time t to the route $r_i{*}$ is smaller than the length $\rho_i(t)$ of path p(u,v) associated at time t to the route r_i: $\delta\rho_i(t) = \rho_i{*}(t) - \rho_i(t) < 0$.

D. Stability Decision Criterion

The BGP selection process enhanced by the stability-based decision criteria, following the differential stability metric defined in Section III.C, would be driven by the following selection rules:

```
if δφi(t) > 0
then if δρi(t) ≤ 0
     then select ri(t) per δφi(t)
     else if δρi(t) < γ
             then select ri(t) per δφi(t)
             else select ri(t) per
                  default BGP selection rules

if δφi(t) ≤ 0
then select ri(t) per
     default BGP selection rules
```

In this selection process, the positive integer parameter γ is determined by the increase of the multiplicative stretch considered as acceptable. Hence, the actual problem becomes to find a mean to actually determine (or at least estimate) the acceptable stretch increase of the routing path that would result from the application of the stability-based decision criteria. Past experiments dedicated to the measure of the BGP AS-path length have shown that even if the average length of AS-paths is relatively stable (about 4 to 5), a significant fraction of AS-paths has a length up to 10 [11]. From this perspective, if we assume that a 10% increase of the multiplicative stretch would be acceptable (resulting multiplicative stretch would be equal to 1.1 instead of 1.0), then routes with an average AS-path length increase of 1 AS-hop would instead be selected. Note that this study does not evaluate the increase in memory consumption required to store the routes with longer AS-path attributes. Moreover, the application of the stability-based decision criterion prevents propagation of the routing updates churn resulting from the occurrence of a path exploration event when the following conditions are met i) the route corresponding to the next stable state is locally stored in the Adj_RIB_In and ii) this route corresponds to the most stable (next) route in the Adj_RIB_In. Indeed, if such event occurs, then the selection of a stable route becomes possible without delaying local convergence resulting from the exploration of all intermediate routing states (e.g., AS-paths of increasing length). Nevertheless, if the path exploration event also affects the route corresponding to the next state corresponding to the most stable next route, then selecting the AS-path that is the least topologically correlated[1] to the previous state provides the safest decision.

Importantly, the applicability of the stability-based decision criterion does not only depend on the point-value of the differential stability metric but also on its evolution over time. This means in practice that we have also to ensure that when the stability criteria are met at time t, and the corresponding selection rules are applied at time t, they also remain applicable at time t+1, and more generally at time t+Δt, where $\Delta t >> 0$. The reason stems as follows: at a given router once a route is selected at time t based on its stability properties, reverting unilaterally to the default BGP selection rules at time t+Δt can itself increase the instability induced by the concerned routes on its downstream routers. Here again, our stability metrics provide a suitable method to estimate the deviation over time and the robustness of the selection process. Indeed, it suffices to notice that (even if it is impossible to locally anticipate all occurrence of BGP instability events before they occur) these metrics enable to determine over time the candidate replacement routes that are more stable compared to the set of possible alternative routes that do not show the same stability properties. When such alternative route does not exist, the exchange process of BGP routing updates between the local router and its downstream neighbors (with respect to the direction of propagation of the routing updates) requires enhancement in order to enable a smooth transition between the route selection rules. This mechanism performs as follows: anticipatively once no candidate replacement route is available for the route currently selected based on the stability criteria, that route is advertized to downstream neighbors together with the route that would be selected based on the default BGP selection rules. This process enables each downstream router to tune its decision process based on its own selection rules for that route. Note that this process enables to advertize both routes, i.e., the one selected based on the stability criteria and the one selected based on the BGP default rules.

IV. Measurement Methodology and Data Set

We apply the metrics defined in Section III to the BGP updates provided by the Route-Views project [12]. The BGP dataset obtained from this project comprises archives containing BGP feeds from a set of worldwide distributed Linux PCs running Quagga/Zebra[2] [13]. Route Views is a project founded and sponsored by the University of Oregon which consists in a set of routers distributed across the world. The BGP routing information collected by these routers can be openly accessed by anyone, interested or involved in the field of Internet research. This information has led to various noticeable studies including those conducted in [3]. The Route-Views data records contain the BGP information a router receives from its neighboring BGP speakers. That is basically each neighbor route (with its route attributes) to each address prefix the neighbor has knowledge about. With this information, the monitored BGP router can find a route for each IP prefix it needs to send packets to.

Current Route-Views data record format does not provide the Loc_RIB information as stored locally by each router (that is, basically and for our interests, the information about which route BGP selects for each prefix). We must thus derive this information from the Adj-RIB-In table as provided by the Route-Views dataset. For this purpose, we infer the Loc-RIB table from the Adj-RIB-In tables by implementing a selection process based on the algorithm used in Quagga/Zebra routers, which is representative of the BGP selection process

[1] Two AS_paths are topologically correlated if they share at least one common edge, i.e., an AS adjacency.

[2] Quagga is a fork of GNU Zebra which was developed by Kunihiro Ishiguro.

commonly applied on Internet routers. A detailed description of the tool developed in C++ programming language to process the BGP datasets obtained and to derive the value of the metrics (defined in Section III), their associated statistics as well as their evolution over time is available in [14].

V. MEASUREMENT RESULTS AND ANALYSIS

This section presents a set of experimental results obtained by applying the metrics and selection rules defined in Section III to BGP dataset obtained from the Route-Views project [12].

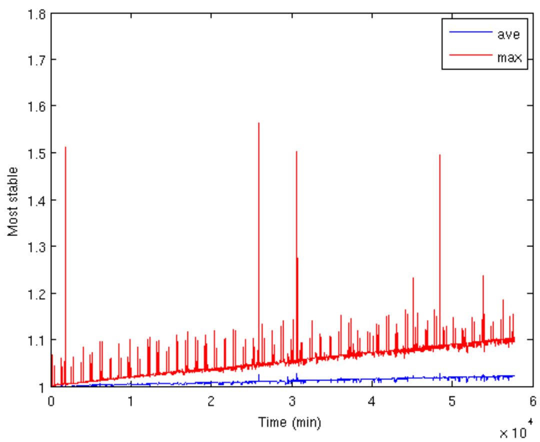

Figure 5. Most stable route metric measure

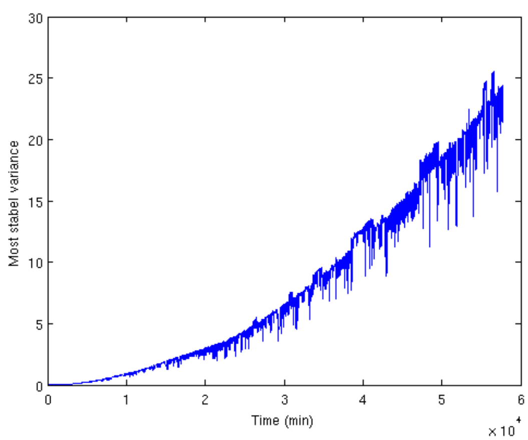

Figure 6. Cumulated variance over time for most stable route

Fig.5 shows that incoming routes stored in Adj_RIB_In have on average slowly decreasing stability compared to the most stable route (a value close to 1 indicates that incoming routes are nearly as stable as the most stable route). As a result, the plot has a small but positive slope. The average of the maximum metric value per destination d shows a positive but larger slope: the most unstable routes have a faster paced decreasing stability (and spiky pattern confirms their unstable behavior). Further, during the entire observation duration (40 days), a subset of routes continuously presented instabilities leading to a monotonic increase of the metric. It can be observed from Fig.5 that the BGP selected route has on average a better stability than the other routes out of which it is selected (a value close to 1 indicates that incoming routes are nearly as stable as the best selectable route). Comparison between Fig.5 and Fig.7 reveals though that local maxima for the selected route exhibits more spaced and less intensive variations than the most stable route (a lower metric value indicates a higher stability). One can also observe the same monotonously increasing trend of this metric for both the average and the maximum, due to routes with sustained instability.

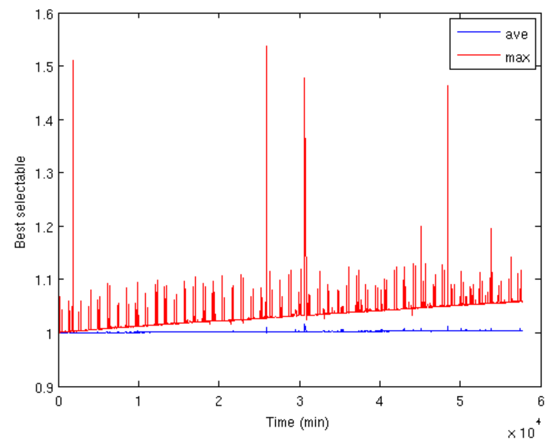

Figure 7. Best selected route metric measure

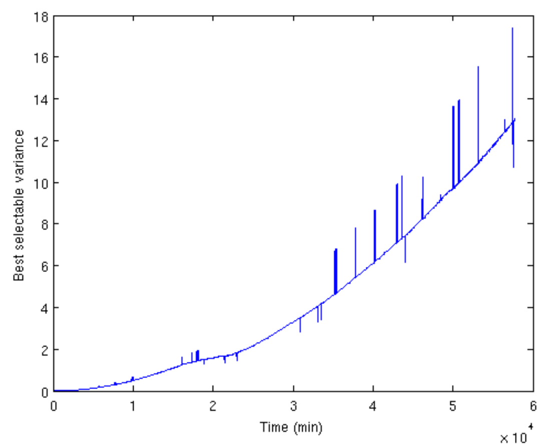

Figure 8. Cumulated variance over time for the best selectable route

Computation of the cumulated variance for the most stable route (Fig.6) shows an increase value over time from about 5 after 20 days and about 25 after 40 days; the slope is super-linear. Nevertheless, we can observe that after one day the variance remains relatively limited, leaving the possibility of selecting the most stable route without incurring significant stability deviation of the entire routing table. Local maxima in Fig.8 indicate large changes in local route stability, i.e., more routes than the average experience instabilities but BGP quickly converges to a new stable state since a part of the affected routes return to their initial state (thanks to the presence of more stable routes in the Adj_RIB_In, as indicated in Fig.6). Interestingly, Fig.8 shows also that the intensity of the instability increases over time indicating that more routes get affected by the change.

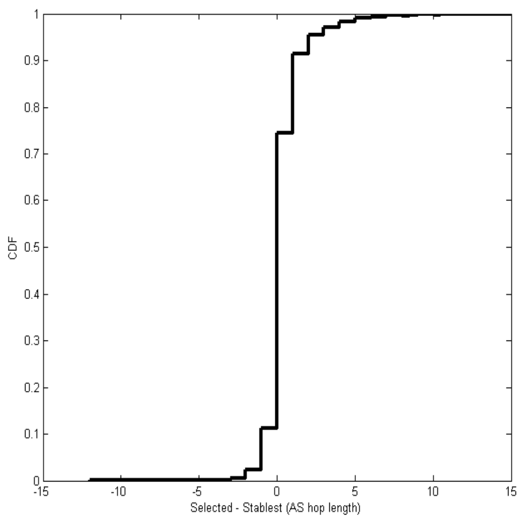

Figure 9. Number of routes vs diff. in AS-path length

Fig.9 shows the cumulated percentage of routes with respect to the AS-path length difference between the selected and the most stable route. A positive difference indicates that the replacement of the selected route (using the BGP path ranking function) by the most stable route would decrease the AS-path length compared to the selected route ($\delta\rho < 0$). A negative difference indicates that such replacement would increase the AS-path length ($\delta\rho > 0$). From this figure, we can deduce that such replacement would be advisable for about 90% of the selected routes since $\delta\rho \leq 0$. Moreover, for 25% percent of the routes, this replacement would also lead to an AS-path length decrease since for these routes $\delta\rho < 0$. Interestingly, only 10% of the routes would be affected by a length increase if they would be selected based on the stability criteria since for these routes $\delta\rho > 0$. Among this percentage of 10%, we can also observe from this figure that a significant fraction of the routes would be covered if an AS-path length increase of one-hop would be considered as acceptable (in average $\delta\rho \cong 1.15$). These observations corroborate the fact that the stability-based selection rule does not lead to a stretch increase for a significant fraction of the routes (90%). On the other hand, by admitting a stretch increase corresponding to one additional AS-hop in the AS-path, only a minor fraction of the routes (about 2%) would be penalized by a higher stretch increase of two AS-hops (and above for a fraction of routes << 1%). This observation can be seen as the experimental evidence that enforcing stability would not come at the detriment of increasing the stretch of the AS-paths.

VI. Conclusion

In this paper, we have defined several stability metrics to characterize the local effects of BGP policy- and protocol-induced instabilities on the routing tables. Our experimental results show that the proposed method enables to locally detect instability events that are affecting routing tables' entries, and deriving their impact on the local stability properties of the routing tables. We have also determined a differential stability-based decision criterion that can be taken into account as part of the BGP route selection process. A significant fraction of the routes (90%) selected by means of this process is not stretch increasing. Moreover, if one would admit an AS-path length increase of one AS-hop, only a minor fraction of the routes (about 2%) would be penalized by a higher stretch increase (two AS-hops and above).

Future work includes verifying the general trade-offs between stability-based route selection and the resulting stretch increase/decrease on the selected routing paths. Moreover, the relationship between local and global stability will be further elaborated to characterize the effects on the global stability of the routing system resulting from the selection of a route that is more stable locally. The idea here is to determine the necessary but sufficient conditions for preventing potential oscillations to occur (as the local action of selecting a more stable route shall not induce unwanted perturbation(s) on neighboring routing states). Finally, the model will be extended to locally discriminate between protocol- and policy-induced instabilities.

References

[1] T.Griffin, F.B.Shepherd, and G.Wilfong, The Stable Paths Problem and Interdomain Routing, IEEE/ACM Transactions on Networking, 10(1):232-243, April 2002.

[2] H.Levin, M.Schapira, and A.Zohar, Interdomain routing and games, Proc. of ACM Symposium on Theory of Computing (STOC), 2008.

[3] G.Huston, Damping BGP, RIPE 55, Routing WG, October 2007.

[4] C.Labovitz, A.Ahuja, A.Bose, and F.Jahanian,, Delayed Internet Routing Convergence, IEEE/ACM Transactions on Networking, 9(3):293-306, June 2001.

[5] Z.M.Mao, R.Govindan, G.Varghese, and R.Katz, Route Flap Damping Exacerbates Internet Routing Convergence, Proc. of ACM SIGCOMM 2002, August 2002.

[6] R.Bush, T.Griffin, and Z.M.Mao, Route flap damping harmful?, NANOG-26, 28 October 2002.

[7] J.Chandrashekar, Z.Duan, Z.-L.Zhang, and J.Krasky, Limiting path exploration in BGP, Proc. of IEEE INFOCOM 2005, Miami, Florida, March 2005.

[8] D.Pei, M.Azuma, D.Massey, and L.Zhang, BGP-RCN: improving BGP convergence through root cause notification, Computer Networks, ISDN Syst. vol. 48, no.2, pp.175-194, June 2005.

[9] G.Huston, M.Rossi, and G.Armitage, A Technique for Reducing BGP Update Announcements through Path Exploration Damping, IEEE Journal on Selected Areas in Communications, vol.28, no.8, October 2010.

[10] D.Papadimitriou, A.Cabellos, and F.Coras, Path-vector Routing Stability Analysis, Proc.13th Workshop on MAthematical Performance Modeling and Analysis, ACM SIGMETRICS, San Jose (CA), USA, June 2011.

[11] B.Huffaker, M.Fomenkov, M.Plummer, D.Moore, and k.claffy, Distance Metrics in the Internet, IEEE International Telecommunications Symposium (ITS), Brazil, pp.200-202, September 2002.

[12] Univ. of Oregon. RouteViews. Available at: http://www.routeviews.org

[13] Quagga Routing Software Suite, GPL licensed, Available at: http://www.nongnu.org/quagga/

[14] EULER FP7 Project, Measurement-based Topology Modeling, Technical report, Deliverable D3.2, Available at: https://www-sop.inria.fr/mascotte/EULER/wiki/pmwiki.php/Main/Deliverables