

Internet routing paths stability model and relation to forwarding paths

Dimitri Papadimitriou

Alcatel-Lucent
Antwerp, Belgium

Davide Careglio

Technical University of
Barcelona
Barcelona, Spain

Fabien Tarissan

UPMC - LIP6
Paris, France

Piet Demeester

Ghent University -
iMinds
Gent, Belgium

Abstract—Analysis of real datasets to characterize the local stability properties of the Internet routing paths suggests that extending the route selection criteria to account for such property would not increase the routing path length. Nevertheless, even if selecting a more stable routing path could be considered as valuable from a routing perspective, it does not necessarily imply that the associated forwarding path would be more stable. Hence, if the dynamics of the Internet routing and forwarding system show different properties, then one can not straightforwardly derive the one from the other. If this assumption is verified, then the relationship between the stability of the forwarding path (followed by the traffic) and the corresponding routing path as selected by the path-vector routing algorithm requires further characterization. For this purpose, we locally relate, i.e., at the router level, the stability properties of routing path with the corresponding forwarding path. The proposed stability model and measurement results verify this assumption and show that, although the main cause of instability results from the forwarding plane, a second order effect relates forwarding and routing path instability events. This observation provides the first indication that differential stability can safely be taken into account as part of the route selection process.

Keywords—component; Internet, path-vector routing, stability, parametric, metrics

I. INTRODUCTION

Following the Routing and Addressing Workshop held by the Internet Architecture Board (IAB) in 2006 [1], stability remains a key criterion to be met by the Internet routing system and its underlying Border Gateway Protocol (BGP). The prominent research efforts [2] [3] [4] [5] conducted over last fifteen years to understand BGP instabilities led to classify them as policy-induced or protocol-induced to account for the distinction between BGP protocol operations and the inherent behavior of the underlying path-vector routing algorithm. Following these studies, stability of the individual local routing states and associated routing path should remain (at least marginally) stable upon occurrence of perturbation resulting from i) the exploration of the routing state space due to the BGP path exploration phenomenon that is intrinsic to the shortest-path vector algorithm, and ii) the BGP routing policies interactions due to which among other can lead to "dispute wheels", i.e., non-deterministic unintended but unstable states. In this context, it is important to underline that the dynamics of the Internet routing system determines the resource consumption of local routing engines, in particular, in terms of memory and CPU. System resource consumption depends not

only 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 overall 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 detect, 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 possible route selection criteria. The method proposed in [5] aims to bring rigor and consistency when studying the stability properties of routing paths as locally experienced by routers. The experimental results reported show that this method enables to locally detect instability events that are affecting routing tables' entries, and derive their impact on the local stability properties of the routing tables. From the metrics defined in [5], a differential stability-based decision criterion is derived that can be taken into account as part of the BGP route selection process [6]. Results show that 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).

Nevertheless, even if selecting a more stable routing path could be considered as valuable from a routing level perspective, it does not necessarily imply that the corresponding forwarding path(s) would be itself more stable. In this work, our first objective consists thus in determining if the dynamics of the Internet routing and forwarding system (through the analysis of routing and forwarding path instability) show different properties. If this assumption is verified then as one can not straightforwardly derive the one from the other; our second objective becomes to investigate the relationship between the stability of the forwarding path followed by the traffic and the corresponding routing path as selected by the path-vector routing protocol. For this purpose, we extend our stability model and locally relate the stability measurements carried on forwarding paths with the measurement performed for the corresponding routing paths following the method developed in [6].

The remainder of this paper is structured as follows. Section II provides an overview on prior work concerning the BGP routing system stability. In Section III, we extend the routing stability model and metrics proposed in [6] and detail the corresponding computational procedures. We document in Section IV the measurement and processing methodology together with the real datasets onto which these metrics have been applied. Section V reports on the measurement results and analysis obtained. Finally, Section VI draws conclusion from this study and outlines possible future work.

II. PRIOR WORK

Numerous studies on BGP dynamics properties have been conducted over last twenty 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, the predominant cause of route oscillation was assumed to result from BGP sessions going up and down because established on circuits that were themselves persistently going up and down. This would lead to a constant stream of BGP update messages from the affected BGP sessions that could propagate through the entire network. The first version of the RFD algorithm specification appeared in 1993, updates and revisions lead to RFC 2439 in 1998 [7].

Mao et al. [8] published in 2002 a paper that studied 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. [9] summarized the findings of Mao et al. [8] 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) [10] 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) [11] 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) [12] 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 (locally observed) instability effects and/or to accelerate the convergence of the routing state(s) after occurrence of a perturbation event. However, none of them ask the fundamental question why selecting a route subject to path exploration at first place. The answer is essentially because none of these techniques rely on the actual quantification of the instability effect and still use network-wide spatial criteria to perform route selection instead of combining them with temporal criteria.

III. ROUTING STABILITY AND METRICS

A. Preliminaries

The autonomous system (AS) topology underlying the routing system is described as a graph $G = (V, E)$, where each vertex (or abstract 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 $u, v \in V$. Each AS comprises a set of physical nodes referred to as routers; the AS representation of the topology combines thus both its partitioning and its abstraction. The subset of physical nodes of interest for this paper comprises the routers running the path-vector algorithm (typically sitting at the periphery of each AS). At each of these routers, a route r per destination d ($d \in D$) is selected and stored as an entry in the local routing table (RT). The total number of routing table entries is denoted by N , i.e., $|RT| = N$. A route r_i to destination d at time t is defined by $r_i(t) = \{d, (v_k=u, v_{k-1}, \dots, v_0=v), A\}$ with $k > 0 \mid \forall j, k \geq j > 0, \{v_j, v_{j-1}\} \in E$ and $i \in [1, N]$, where $(v_k=u, v_{k-1}, \dots, v_0=v)$ represents the AS-Path from the abstract node u to v , v_{k-1} the next hop of v along this AS-Path, 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 $p(u,v) = \{(v_k=u, v_{k-1}, \dots, v_0=v), A\}$. A routing information update leads to a change of the AS-Path $(v_k, v_{k-1}, \dots, v_0)$ or an element of its attribute set A . Next, a withdrawal is denoted by an empty AS-Path (ε) and $A = \emptyset: \{d, \varepsilon, \emptyset\}$. According to the above definition, if there is more than one AS-Path per destination d , they will be considered as multiple distinct routes.

Even if the stability model and associated metrics detailed in this paper applies to any path-vector routing protocol, BGP is in the context of this paper the path-vector routing protocol under consideration. Thus, we further detail its storage data structures, referred to as Routing Information Bases (RIBs), used to store its routes $r_i(t)$. At each BGP speaker, the RIB consists of three distinct parts: the Adj-RIB-In, the Loc-RIB, and the Adj-RIB-Out. The Adj-RIB-In contains unprocessed routing information that has been announced to the local BGP speaker by its peers. The Loc-RIB which corresponds to the BGP local routing table (RT) contains the routes that have been selected following the local BGP speaker's decision process. Finally, the Adj-RIB-Out organizes the routes for announcement to specific (downstream) peers. When a router

receives a route announcement, it first applies inbound filtering process (using some import policies) to the received routing information. If accepted, the route is stored in the Adj-RIB-In. The collection of routes received from all neighbors (external and internal) that are stored in the Adj-RIB-In defines the set of candidate routes (for that destination). Subsequently, the BGP router invokes a route selection process - guided by locally defined policies - to select from this set a single best route for each destination. After this selection is performed, the selected best route is stored in the Loc-RIB and is subject to some outbound filtering process and then announced to all the router's neighbors. Importantly, prior to being announced to an external neighbor, but not to an internal neighbor in the same AS, the AS path carried in the announcement is prepended with the ASN of the local AS.

B. Routing Path Stability

The stability of a routing path is characterized by its response (in terms of processing of routing information) to inputs of finite amplitude. Inputs affecting routing path states 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 (originating router), etc. during their route selection process.

Measuring the magnitude of the output from the received input and current local routing state is the main purpose of the metric referred to as stability of the selected route stored in the Loc_RIB (that corresponds to the BGP routing table). For this purpose, we define the criteria for qualifying the effects of a perturbation on the entries of the local routing table so as to locally characterize the stability properties of the routing paths. More precisely, let $|\Delta(r_i(t+1) - r_i(t))|$ be the magnitude of the change experienced by the route r_i between time $t = t_0 + k$ and time $t + 1 = t_0 + (k+1)$, 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. Following the BGP specification, 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. Note that a similar reasoning to the one applied for the stability of the routes stored in the Loc_RIB can also be applied to the routes stored in the Adj_RIB_In (see Section III.A).

1) Equilibrium States

Consider the nonlinear discrete time autonomous system described by $r_i(k+1) = r_i(k) + h(r_i(k)) + u_i(k)$, where, $r_i(k)$ is the discrete state associated to the route r_i at time k , $h(r_i(k))$ is a nonlinear function of $r_i(k)$ and $u_i(k)$ is the control input for the

corresponding route at time k . We distinguish three different equilibrium states associated to a given route r_i :

Definition 1: the equilibrium state r_i^* associated to the route r_i is considered to be *stable* if $\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0$, such that if $|\Delta(r_i(t_0) - r_i^*)| < \delta$ then $|\Delta(r_i(t_0+k) - r_i^*)| < \varepsilon$ for all $k \geq 0$. If this condition is met, the stability of the route r_i as locally observed at time t_0 remains for all $k \geq 0$ close to its initial equilibrium value r_i^* at time t_0+k .

Conversely, the equilibrium state r_i^* associated to the route r_i is considered to be *unstable* if $\exists \varepsilon > 0, \forall \delta(\varepsilon) > 0$, such that if $|\Delta(r_i(t_0) - r_i^*)| < \delta$ then $|\Delta(r_i(t_0+k) - r_i^*)| > \varepsilon$ for some $k \geq 0$.

Definition 2: the equilibrium state r_i^* associated to the route r_i is considered to be *asymptotically stable* if: i) it is stable and ii) $\exists \gamma > 0$ such that $|\Delta(r_i(t_0) - r_i^*)| < \gamma$ implies that $|\Delta(r_i(t_0+k) - r_i^*)| \rightarrow 0$ for $k \rightarrow \infty$. If these two conditions are met, the stability of the route r_i as locally observed at time t_0 converges to its initial equilibrium value r_i^* for $k \rightarrow \infty$.

Moreover, it is often useful to estimate how fast a given route r_i approaches its equilibrium state. For this purpose, the notion of exponential stability is introduced:

Definition 3: the equilibrium state r_i^* is considered to be *exponentially stable* if i) it is stable and ii) $\exists \gamma > 0$ such that $|\Delta(r_i(t_0) - r_i^*)| < \gamma$ implies $|\Delta(r_i(t_0+k) - r_i^*)| \leq \kappa |\Delta(r_i(t_0) - r_i^*)| e^{-\lambda t}$ for all $k > 0$, where $\kappa > 0$ and $\lambda > 0$ is often called the exponential convergence rate.

Note that exponential stability implies asymptotic stability and asymptotic stability implies stability in the sense of definition 1. The reverse implications are not necessarily true. Consequently we can classify equilibrium states as either asymptotically stable or not. The latter subdivides into marginally stable and unstable states:

Definition 4: the equilibrium state r_i^* associated to the route r_i is considered to be *marginally stable* if: i) it is stable and ii) $\exists \delta > 0$ and $\exists \alpha > 0, \beta > 0$ with $\alpha < \beta < \infty$ such that $|\Delta(r_i(t_0) - r_i^*)| < \delta$ implies $\alpha < |\Delta(r_i(t_0+k) - r_i^*)| < \beta$ for all $k \geq 0$. If these two conditions are met, the stability of the route r_i as locally observed at time t_0 oscillates around the equilibrium state r_i^* .

2) Parametric Stability

The actual values of the above parameters α, β, γ and δ depend on several factors. Among them, the MRAI value and the integer k that determines the number of MRAI time units that have elapsed since the beginning of the observation sequence. Their computation is explained in Section III.C.

In the present context, the equilibrium states shift over time (when the path vector routing algorithm converges to a new routing state). Hence, the standard approach that comprises the following steps i) locate the equilibrium state, ii) select the one that is of interest, iii) translate it to the origin (justified by the fact that a stability analysis can be developed "without loss of generality" for the equilibrium at the origin and then universally used for other equilibria of model), and iv) lastly determine its stability property breaks down when parametric uncertainties are present. Each time a route parameter is changed the original equilibrium r_i^* may shift to another

equilibrium state making the stability analysis of the translated equilibrium at the origin at best imprecise and at worst inappropriate. For this purpose, we adapt the concept of parametric stability introduced in [13]; this concept embodies in a single definition the feasibility and the stability of dynamic nonlinear systems with uncertain parameters $\pi \in \mathcal{R}$. More precisely, to account for the possible change of the parameter $\pi^* \rightarrow \pi$, such that $\pi \in \mathcal{N}(\pi^*)$, where $\mathcal{N}(\pi^*)$ is defined as a (small) neighborhood of π^* , we consider the following changes compared to the equations and definitions provided in Section III.B.1: i) the nonlinear discrete time autonomous system is now described by $r_i(k+1) = r_i(k, \pi) + h(r_i(k), \pi) + u_i(k)$ ¹ and ii) we replace the following elements included in the Definitions 1, 2, 3, and 4 provided in Section III.B.1: $r_i^* \rightarrow r_{i,eq}(\pi)$, $\delta \rightarrow \delta(\pi)$, $\gamma \rightarrow \gamma(\pi)$. In turn, the above definitions translate respectively the parametric stability, parametric asymptotic stability, and parameter marginal stability of the system at $\pi^* \in \mathcal{R}$.

- The route $r_i(\pi^*)$ is said to be *parametrically stable* at π^* if $\exists \mathcal{N}(\pi^*)$ such that $\forall \pi \in \mathcal{N}(\pi^*)$: i) \exists equilibrium $r_{i,eq}(\pi)$ and ii) $\forall \varepsilon > 0$, $\exists \delta(\varepsilon, \pi) > 0$, such that if $|\Delta(r_i(t_0) - r_{i,eq}(\pi))| < \delta$ then $|\Delta(r_i(t_0+k) - r_{i,eq}(\pi))| < \varepsilon$ for all $k \geq 0$.
- The route $r_i(\pi^*)$ is said to be *parametrically asymptotically stable* at π^* if: i) it is parametrically stable at π^* and ii) $\forall \pi \in \mathcal{N}(\pi^*)$, $\exists \gamma(\pi) > 0$ such that $|\Delta(r_i(t_0, \pi) - r_{i,eq}(\pi))| < \gamma$ implies that $|\Delta(r_i(t_0+k, \pi) - r_{i,eq}(\pi))| \rightarrow 0$ for $k \rightarrow \infty$.
- The route $r_i(\pi^*)$ is said to be *parametrically marginally stable* at π^* if: i) it is parametrically stable at π^* and ii) $\exists \delta(\pi) > 0$ and $\exists \alpha, \beta > 0$ with $\alpha < \beta$ such that $|\Delta(r_i(t_0) - r_{i,eq}(\pi))| < \delta$ implies $0 < \alpha < |\Delta(r_i(t_0+k) - r_{i,eq}(\pi))| < \beta < \infty$ for all $k \geq 0$.

In the context of path vector routing, the parameter π translate the variation of the preference rules associated to the selected route r_i . Using these parametric stability conditions one can then provide an answer to the following critical problems in the context of the analysis of local routing states: does there exist a new equilibrium state $r_{i,eq}(\pi)$ and how far is it from the equilibrium state $r_i(\pi^*)$? If $r_{i,eq}(\pi)$ exists, is it stable as $r_i(\pi^*)$ was or are its stability properties modified by the change of parameter from π^* to π ? Moreover, one can estimate the extent of a region $\mathcal{P} \in \mathcal{R}$ in the parameter space where a given route $r_i(\pi)$ is parametrically stable by verifying the above mentioned conditions $\forall \pi \in \mathcal{P}$. This leads to the following theorem:

Theorem: if the stability regions \mathcal{P}_j for the route r_i along each of its node j define a non-empty intersection \mathcal{P}^* (i.e., $\forall j$, $\exists \pi \in \mathcal{P}_j$ such that $\pi \in \bigcap_{k \neq j} \mathcal{P}_k$) then $\forall \pi \in \mathcal{P}^*$, the route $r_i(\pi)$ is parametrically stable.

Proof: if $\exists \pi \in \mathcal{P}^*$ such that the route $r_i(\pi)$ is parametrically unstable then there is no intersection \mathcal{P}^* such that $\mathcal{P}^* \neq \emptyset$, i.e., it is impossible to find an intersection $\mathcal{P}^* \supset \pi$ such that the route $r_i(\pi)$ is parametrically stable.

The route $r_i(\pi^*)$ is said to be parametrically unstable at π^* if $\forall \mathcal{N}(\pi^*)$, $\exists \pi \in \mathcal{N}(\pi^*)$ for which either there is no equilibrium $r_{i,eq}(\pi)$ or there is an equilibrium $r_{i,eq}(\pi)$ which is unstable per Section III.B.1. Moreover, the route $r_i(\pi^*)$ is parametrically unstable with respect to any region $\mathcal{P}_\theta = \{\pi \in \mathcal{P} : |\pi - \pi^*| < \theta, \theta > 0\}$ if \exists (at least one) $\pi' \in \mathcal{P}$ such that the route $r_i(\pi')$ is parametrically unstable.

We have thus two possible alternatives: either $\pi' \in \mathcal{P}_j \wedge \pi' \notin \bigcap_{k \neq j} \mathcal{P}_k$ or $\pi' \in \mathcal{P}_j \wedge \pi' \in \bigcap_{k \neq j} \mathcal{P}_k$. In the former case, the parameter π' does not belong to the intersection \mathcal{P}^* ; in the second case, the intersection \mathcal{P}^* is not empty ($\mathcal{P}^* \neq \emptyset$) but it does not define a stability region since route $r_i(\pi')$ is parametrically unstable at π' . Therefore, it is not possible to find an intersection $\mathcal{P}^* \supset \pi$ such that the route $r_i(\pi)$ is parametrically stable. Hence, the intersection $\mathcal{P}^* = \emptyset$. \square

Consequently, the set of all parameters $\pi \in \mathcal{P}^*$ such that $\pi \in \mathcal{P}_j, \forall j$, define the stability region \mathcal{P}^* for the route r_i . In other terms, if the parameters π associated to the preference rules along each node j of the routing path r_i could be determined, adjusting (locally) these parameters would guarantee stability. However, these parameters are only known locally, i.e., at the AS to which they apply, and inferring them is usually impossible outside of direct AS relationships.

C. Routing Path Stability Computation

It is thus necessary to locally quantify the deviation of the routing state (associated to the each route) upon perturbation of the initial condition at which that routing state (and thus the corresponding route) is assumed to be in equilibrium. For this purpose, we determine the stability of the route $r_i(t)$, $i \in [1, |D|]$, that is stored at time t in the Loc_RIB ($|\text{Loc_RIB}| = N$) by means of the quantity $\phi_i(t)$ defined as the *stability metric* of the selected route $r_i(t)$. Upon perturbation by an external and/or internal event of the state associated to the route $r_i(t)$, the metric $\phi_i(t+1)$ quantifies 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 and the integer k accounts for the number of MRAI time units that have elapsed since the starting time of the measurement sequence. In other terms, this metric accounts for all change(s) experienced by the route r_i with a periodicity determined by the MRAI time.

Moreover, it is also interesting to measure the instability induced by the BGP route selection process itself. The latter, referred to as the differential stability metric $\delta\phi_i$ [5] [6], enables to quantify the difference between the most stable route in the Adj_RIB_In and the selected route stored in the Loc_RIB for the same destination d . Using this metric as part of the decision criteria enables to prevent replacement (in the Loc_RIB) of more stable routes by less stable ones but also to select more stable routes than the currently selected routes.

For a given route r_i , the stability metric ϕ_i can be computed by using the procedure described in Fig.1. Upon creation of a new routing table entry for the route r_i at time $t = t_0$, the value $\phi_i(t=t_0)$ is initialized together with the parameters α and β (defined in Section III.B). The parameters α and β can be

¹ In the remainder of this paper, we assume that $u_i(k) = 0$

derived from this procedure on a per individual route basis. Moreover, following the procedure described in Fig.1, the function $f_i(\phi_i(k), \pi): \mathcal{R}^+ \rightarrow \mathcal{R}^+$, $\phi_i(k), \pi \rightarrow a_f \phi_i(k), \pi + b_f$ accounts for all changes experienced by the route r_i to destination d from time t to time $t+1$. The function $g_i(\phi_i(k), \pi): \mathcal{R}^+ \rightarrow \mathcal{R}^+$, $\phi_i(k), \pi \rightarrow a_g \phi_i(k), \pi + b_g$ operates when the corresponding route r_i undergoes no change during the time interval between t and $t+1$. An appropriate value assignment to the coefficients a_f, a_g, b_f, b_g enables to accommodate any additive ($a_f = 0, b_f > 0$) / multiplicative ($a_f > 0, b_f = 0$) increase, additive ($a_g = 0, b_g > 0$) / multiplicative decrease ($0 < a_g < 1, b_g = 0$) algorithm. In the following, we set these coefficients to operate in the additive increase/additive decrease mode.

```

/* Initialization when route  $r_i$  to destination
 $d$  is created at time  $t = t_0$  */
 $\phi_i(t=t_0=0) \leftarrow 0$ 
 $|\Delta\phi_i(t=t_0=0)| \leftarrow 0$ 
 $\alpha_{min,i} \leftarrow 0$ 
 $\beta_{max,i} \leftarrow 1$ 

/* Measurement during interval  $T$  time units */
While (( $k * \text{MRAI}$ ) <  $T$ )
  if  $|\Delta(r_i(k+1) - r_i(k, \pi))| > 0$ 
    /*  $r_i$  experiences an AS-Path change
       or experiences an attribute change */
  then  $\phi_i(k+1) \leftarrow \phi_i(k, \pi) + f_i(\phi_i(k), \pi)$ 
        $|\Delta\phi_i(k+1)| \leftarrow \phi_i(k+1) / [\phi_i(k+1) + 1]$ 
       if  $\phi_i(k) = 0$ 
       then  $\delta_i \leftarrow |\phi_i(k+1) - \phi_i(t_0=0)|$ 
       else  $\beta_i \leftarrow \phi_i(k+1)$ 
           if  $\beta_{max,i} < \beta_i$  then  $\beta_{max,i} \leftarrow \beta_i$ 
           end if
  else /*  $r_i$  experiences no change
        $|\Delta(r_i(k+1) - r_i(k, \pi))| = 0$  */
       if  $\phi_i(k) > 0$ 
       then  $\phi_i(k+1) \leftarrow \phi_i(k, \pi) - g_i(\phi_i(k), \pi)$ 
            $|\Delta\phi_i(k+1)| \leftarrow \phi_i(k+1) / \phi_i(k, \pi)$ 
            $\alpha_i \leftarrow \phi_i(k+1)$ 
           if  $\alpha_{min,i} > \alpha_i$  then  $\alpha_{min,i} \leftarrow \alpha_i$ 
           end if
       else  $\phi_i(k+1) \leftarrow 0$ 
       end if
  end if
   $k \leftarrow k + 1$ 
end k loop

```

Figure 1. Stability of individual routes

D. Forwarding Path Stability

Even if the selection of a more stable routing path could be considered as valuable from the routing perspective, it does not necessarily imply that the corresponding forwarding path(s) would itself be more stable. Hence, it is also required to locally quantify the stability of the forwarding state associated to the each forwarding path resulting from the selection of a given route. For this purpose, we make use of the RADAR tool that records sequences of IP addresses corresponding to the routers

traversed by the forwarding path and not the AS-path as provided by the RouteView tool. Thus, each forwarding path needs to be associated to the corresponding AS number sequence (corresponding to the routing path) following the procedure documented in Section IV.B. Analogously to the procedure applied to routing paths, the stability of each forwarding path can be computed following the algorithm described in Fig.1. Observe though that as the computation algorithm is applied to the AS sequence instead of the IP address sequence corresponding to each forwarding path, the stability metric identifies "inter-AS" changes in the forwarding path not "intra-AS" changes. Identification of additional "intra-AS" changes is left for future study.

IV. DATASETS AND PROCESSING METHOD

In this section, we describe the processing method applied to the results of the stability metric computation obtained from the application of the algorithm described in Section III.C.

A. Datasets

The computation of the stability of the routing paths relies on the processing BGP update messages as collected by the RouteViews project (www.routeviews.org). The computation of the stability of the forwarding paths makes use of the forwarding paths as recorded by the RADAR tool [14].

1) Routing Path Dataset

RouteViews [15] is a project founded and sponsored by the University of Oregon which consists in a set of BGP routers distributed worldwide. The BGP routing information collected by these routers is stored in BGP feeds. The BGP datasets obtained from these routers can be openly accessed by anyone, interested or involved in the field of Internet (routing) research. This information has led to various noticeable studies including those conducted in [16] and [17]. More precisely, the BGP datasets obtained from RouteViews contain the BGP routing information a monitored router receives from its neighboring BGP speakers. The datasets obtained from these monitored routers comprise the following information: i) the complete Routing Information Base (RIB) entries (updated every two hours) and ii) the received updates from peering ASs separated in files recorded every 15 minutes. The format used to encode the records in these files is MRT [18]. The monitored router is route-views.wide.routeviews.org in order to facilitate the association with the forwarding path dataset described in Section IV.B. We use the tool developed in [5] [6] to process all the BGP data collected from this RouteViews router.

2) Forwarding Path Dataset

The measurements carried out by RADAR are traceroute-like probes initiated from a set of monitoring nodes. Such probes target a large set of IP address prefixes and end-hosts distributed across the Internet. Based on these measurements, the RADAR tool builds ego-centered views of the forwarding topology (in other terms, the initiating router collects traces along the forwarding paths that it probes). A subset of the forwarding paths traced by the RADAR probes corresponds expectedly to the routers monitored by RouteViews; consequently, a subset of the monitored AS-Paths is also monitored by RADAR.

B. Datasets Pre-Processing

The first step to relate the stability of a given forwarding path to the corresponding routing path is to find the possible association between the dataset records provided by the RADAR tool (forwarding paths) and RouteViews (routing paths). Indeed, the data provided by RADAR are sequences of IP addresses while the BGP routes as provided by the RouteViews datasets are sequences of AS numbers (AS-Path). Finding association (or matching) between IP forwarding paths to AS routing paths is thus required. Performing this operation can be obtained by executing the Whois protocol [19]. Whois is a TCP-based transaction-oriented query/response protocol that is widely used for querying databases that store the registered users or assignees of an Internet resource, such as a domain name, an IP address prefix, or an AS. We have used the Whois-based web tool provided by the Team Cymru (<http://www.team-cymru.org/>). This specific tool takes as input a file containing IP addresses and translates them into AS numbers as output of the tool developed in C++ programming language. In total, each sample includes a bit less than 1000 forwarding path - routing path sequences. It is important to mention that performing association between pairs of forwarding and routing paths (per destination) does not require the full identification of data sequences before association but only that a given forwarding path (IP address sequence) can be associated unambiguously to a given routing path (AS number sequence). Hence, the identification problem could be limited to a specific subset of the total number of pairs (those experiencing instability). Moreover, this method working by association is much simpler compared to the one that would require performing full mapping of IP addresses (forwarding paths) and AS (routing paths) before computation and analysis.

On the other hand, the interval of the measurement applied by the RADAR tool and the MRAI time interval used to process the BGP routes (obtained from RouteViews) are basically different. In RADAR, each measurement round takes approximately 4 minutes and 10 minutes elapses between the end of a given round and the beginning of the next one. We thus run two different sets of execution with the routes obtained from the RouteViews data.

- MRAI time interval: The first set of executions uses the actual BGP UPDATE message time interval as determined by the MRAI time; the stability of the routing path is therefore computed (per destination) according to the real BGP UPDATE message period as regulated by the MRAI. Moreover, the value of the forwarding path stability is not computed at this granularity but assumed to remain constant at the value computed at the beginning of each 4 minute interval.
- RADAR time interval: the second set of executions relies on the RADAR iteration (around each 10 minutes). In this case, the routing path stability variation(s) are accounted following the RADAR time interval; thus, in this case the timing at which the stability metric computation is performed is driven by the RADAR tool.

After having associated the forwarding paths to the routing paths (i.e., produce pairs of paths) and scaled their measurement intervals, one can then i) compute the stability

metric (as defined in Section III.C) for each routing path using the collected BGP datasets, ii) compute the stability metric (as defined in Section III.D) for the corresponding forwarding path using the collected RADAR datasets, and iii) derive the associated statistics and evolution over time.

C. Processing Method

Based on the obtained set of forwarding path (FP) - routing path (RP) pairs, a first characterization of the observed instability would consist in determining whether instability events can be detected or not at the forwarding path and/or the routing path level. For this purpose, the set of (FP,RP) pairs can be classified into 4 subsets (or classes) following the type of experienced event:

1. (FP_Stable, RP_Stable): obviously such pair does not require any further processing or analysis.
2. (FP_Stable, RP_Unstable): each pair comprised as part of this class translates routing path instability with forwarding path stability.
3. (FP_Unstable, RP_Stable): each pair comprised as part of this class translates routing path stability without forwarding path stability.
4. (FP_Unstable, RP_Unstable): each pair comprised as part of this class requires identification if a common segment is at the origin of the instability (thus, the AS Path -IP address mapping is required to determine whether there is a common origin to the observed instability).

Over all pairs belonging to each class, we record the minimum and the maximum value of the stability metric in addition to the computation of the average and variance of the stability metric. Pairs part of classes (Class_2) and (Class_3) are also interesting to analyze because they translate routing path instability without forwarding path instability and vice-versa; identifying the origin of the instability for the pairs belonging to corresponding classes can be performed in a second phase of analysis (as more costly).

After a couple of initial executions performed over the whole duration of the measurement period, it became clear that it was simply not possible to classify all routing path - forwarding path pairs by means of a single discriminant. Indeed, some of these pairs can exhibit multiple patterns during the measurement period and capturing this behavior under a single value was not capable of producing sufficient information. Instead, we adapted the procedure and counted the number of events labeled as (FP_Stable, RP_Stable), (FP_Stable, RP_Unstable), (FP_Unstable, RP_Stable), or (FP_Unstable, RP_Unstable) that were observed for each (FP,RP) pair. We then derive a dominant/main trend corresponding to the label with the maximum number of counts and a sub-trend. We also count for each pair the duration (in MRAI time units) associated to the occurrence of the events that are determined by the above-mentioned classes. This additional processing enables to derive for each pair a dominant and a sub-trend with respect to their duration. Indeed, certain pair may have a very few number of counts but certain of them may be very long.

V. RESULTS AND ANALYSIS

Table I summarizes the results obtained from the classification of the set of forwarding path (FP) - routing path (RP) pairs using the following class labels: (FP_unstable, RP_unstable), (FP_unstable, RP_stable) and (FP_stable, RP_unstable). The second column indicates the absolute number and the percentage of pairs per class for which at least one instability event has been observed. The third column provides the maximum number of measurement intervals over which the corresponding behavior has been observed together with the median value.

TABLE I: CLASSIFICATION OF (FP,RP) PAIRS

Label	Number and Percent of Pairs	Max.Count - Median
FP unstable - RP unstable	517 - 54%	40 - 2
FP unstable - RP stable	915 - 96%	223 - 47
FP stable - RP unstable	182 - 19%	117 - 2

Table II details the observed dominant behavior/trend and the sub-trend; each computed as follows: a score of 1 is assigned to the dominant behavior and 0 to the others. In case a balance is observed between two (three) classes, a score of 0.5 (0.33) is assigned to each of them. As it can be observed from this table, the majority of the (FP,RP) pairs falls in the (FP_unstable, RP_stable) class. The latter determines the dominant behavior, i.e., the most representative behavior. The total number of instability events observed for about 95% of the pairs results from forwarding path instability. Few pairs (less than 4%) are labeled as (FP_unstable, RP_unstable), meaning that only a small fraction of the routing paths experiencing instability events corresponds to forwarding path instability. The second trend indicates that for about 50% of the pairs the observed instability result from both forwarding and routing path instability.

TABLE II: TREND ANALYSIS

Main trend	Number of Pairs	Score
FP unstable - RP unstable	36	32
FP unstable - RP stable	912	906
FP stable - RP unstable	15	12
Second trend	Number of Pairs	Score
FP unstable - RP unstable	474	444
FP unstable - RP stable	3	3
FP stable - RP unstable	58	12

Figure 2 plots the percentage of the observed instability events (i.e., either the forwarding or the routing path would be unstable) over the entire measurement period in the form of a cumulative distribution function (CDF). From this figure, the following observations can be drawn. The majority of the pairs (around 60%) are labeled as (FP_stable, RP_stable), i.e., the instability events observed for each of these pairs account for less than 10% of the observed events; moreover, for around 75% of the pairs, the percentage of observed instability events is less or equal to 20% of the total number of events. The latter percentage increases to 50% when reaching about 87% of the pairs (i.e., for only 13% of the pairs, the instability events experienced is higher or equal to 50% of total number of events).

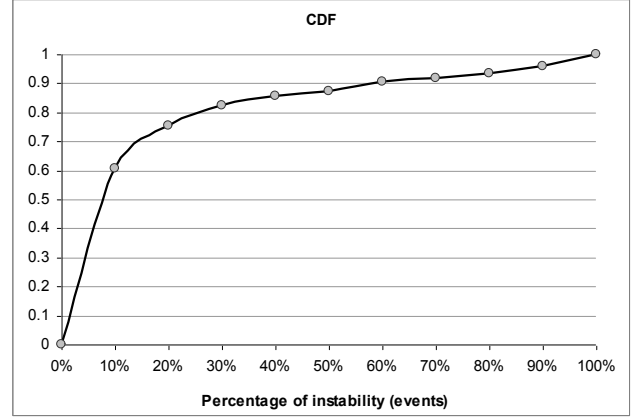


Figure 2. Cumulative Distribution Function vs Perc. of instability (events)

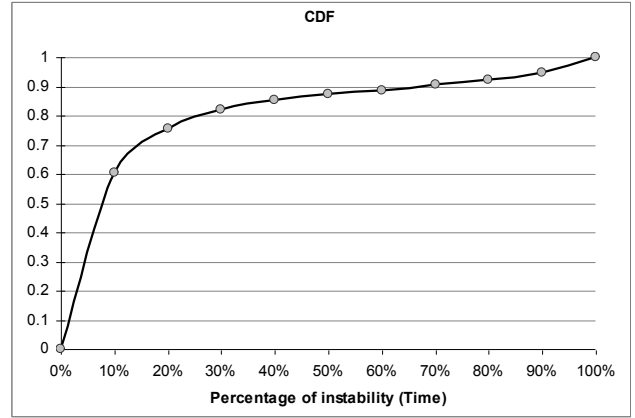


Figure 3. Cumulative Distribution Function vs Perc. of instability (time)

We have also computed the total duration (counted in MRAI time units) of the instability events observed for each pair in order to distinguish the number of transitions to instability events from their actual duration. From this computation, we determine that the instability events observed for about 99% of the pairs results in majority from forwarding path instability. More precisely, the dominant instability behavior over time is characterized by a majority of pairs belonging to the (FP_unstable, RP_stable) class. More generally, the results obtained in terms of duration measurement tend to enforce the main trend observed from

Table II. However, they do not confirm those observed for the second trend: number of (FP_unstable, RP_unstable) pairs 145 vs. 474 and number of (FP_stable, RP_unstable) pairs 793 vs. 58. This seems to imply that for a majority (about 85%) of the pairs the second order temporal effect is dominated by routing path instability without forwarding path instability.

Figure 3 plots the percentage of time during which the instability events have been observed (i.e., either the forwarding or the routing path would be unstable) over the entire measurement period in the form of a cumulative distribution function (CDF). The following observations can be drawn from this figure. The majority of the pairs (about 60%) can be labeled as (FP_stable, RP_stable); the cumulated time of instability accounts for less than 10% of the total duration, i.e., during the remaining 90% of the time the observed events are labeled as (FP_stable, RP_stable). Moreover, for about 75% of the pairs, the observed instability events account for up to 20% of the total duration, i.e., during 80% of the time the observed events are labeled as (FP_stable, RP_stable). The latter percentage increases to 50% when reaching about 87% of the pairs.

These observations combined with the fact that the main cause of instability results from the forwarding plane corroborates the assumption that the dynamic properties underlying the forwarding and the routing system are different. Henceforth, it is impossible to derive one behavior from the other. Nevertheless, it can be observed that a second order effect correlates the forwarding and routing path instability for about 50% of the observed events of instability.

VI. CONCLUSION

In this paper, by means of the stability analysis of routing and forwarding paths, we have captured first evidence that the observed dynamics of the Internet routing and forwarding system show different properties. Hence, one can not straightforwardly derive the stability behavior of the one from the other. We further investigate the relationship between the stability of the forwarding path followed by the traffic and the corresponding routing path as selected by the path-vector routing protocol. For this purpose, we locally relate, at the router level, the stability measurements carried on forwarding paths with the corresponding routing paths by means of the method developed in [4]. Our subsequent analysis shows that the main cause of instability results from the forwarding plane as the dominant instability behavior is characterized by a majority of (FP_unstable, RP_stable) events. This observation further corroborates the assumption that the dynamic properties of the forwarding and the routing system are different. However, it can also be observed that a second order effect relates forwarding and routing path instability events. This observation provides first indication that a BGP route selection process based on differential stability decision criteria (see [3]) can safely be taken into account as part of the

BGP route selection process. In future work we will extend our stability model to account for the influence of variable rates of received routing updates on routing path stability; as these rates depend on the number of AS relationships, we intend to refine our routing path stability model with respect to the AS position in the Internet topology.

REFERENCES

- [1] D.Meyer, L.Zhang, and K.Fall, Report from the IAB Workshop on Routing and Addressing, Internet Engineering Task Force (IETF), RFC 4984, September 2007.
- [2] C.Labovitz, R.Malan, and F.Jahani, Origins of Internet Routing Instability, Proc. of IEEE INFOCOM 1999, pp.218-226, New York (NJ), USA, March 1999.
- [3] C.Labovitz, A.Ahuja, A.Bose, and F.Jahani, Delayed Internet Routing Convergence, IEEE/ACM Transactions on Networking, 9(3):293-306, June 2001.
- [4] 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.
- [5] D.Papadimitriou, A.Cabellos, and F.Coras, Path-vector Routing Stability Analysis, Proc.13th Workshop on MAThematical Performance Modeling and Analysis, ACM SIGMETRICS 2011, San Jose (CA), USA, June 2011.
- [6] D.Papadimitriou, A.Cabellos, and F.Coras, Stability metrics and criteria for path-vector routing, To appear in Proc. of IEEE International Conference on Computing, Networking and Communication (ICNC) 2013, San Diego (CA), USA, January 2013.
- [7] C.Villamizar, R.Chandra, and R.Govindan, BGP Route Flap Damping, Internet Engineering Task Force (IETF), RFC 2439, November 1998.
- [8] Z.M.Mao, R.Govindan, G.Varghese, and R.Katz, Route Flap Damping Exacerbates Internet Routing Convergence, Proc. of ACM SIGCOMM 2002, Pittsburgh (PA), USA, August 2002.
- [9] R.Bush, T.Griffin, and Z.M.Mao, Route flap damping harmful?, NANOG-26, 28 October 2002.
- [10] J.Chandrashekar, Z.Duan, Z.-L.Zhang, and J.Krasky, Limiting path exploration in BGP, Proc. of IEEE INFOCOM 2005, Miami (FL), USA, March 2005.
- [11] 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.
- [12] 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 (JSAC), vol.28, no.8, October 2010.
- [13] M.Ikeda, Y.Ohta, D.D.Siljak, Parametric Stability, in G.Conte, A.M.Perdon, B.Wyman (Eds), New Trends in System Theory, Birkhäuser, Boston, 1991, pp.1-20.
- [14] M.Latapy, C.Magnien and F.Ouédraogo, A Radar for the Internet, Complex Systems, vol. 20, no. 1, pp.23-30, March 2011.
- [15] Univ. of Oregon, RouteViews. Available at: <http://www.routeviews.org>
- [16] G.Huston, Damping BGP, RIPE 55, Routing WG, October 2007.
- [17] 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.
- [18] C.Blunk, M.Karir, and C.Labovitz, Multi-Threaded Routing Toolkit (MRT) Routing Information Export Format, Internet Engineering Task Force (IETF), RFC 6396, October 2011.
- [19] L.Daigle, WHOIS Protocol Specification, Internet Engineering Task Force (IETF), RFC 3912, September 2004.