Method of reliability and availability analysis

From the dynamic properties of routing and forwarding paths

Invited Paper

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—**Confronted to the increasing dynamic of Internet routing system and its underlying topology, we propose a reliability and availability analysis method based on the characterization of the dynamic properties (in particular, the stability properties) of routing paths and their corresponding forwarding paths. The key driver underlying this method is that transient but frequent changes in the spatio-temporal properties of routing paths can affect the performance and operating conditions of the corresponding forwarding paths; hence, their reliability. The results obtained by means of this method verify that, although the main cause of instability results from the forwarding plane dynamics, a second order effect relates forwarding and routing path instability events. Applying our analysis method reveals that the dominant source (main cause) of instability originates indeed from the forwarding plane. This result which confirms previous studies from 2003 further corroborates the assumption that the dynamic properties of routing system are mainly driven by its adaptation to the forwarding system (adaptive routing). However, even if the likelihood of forwarding instability becomes the prominent behavior (cause), about 50% of them induce routing path instability whereas the corresponding forwarding path remains unstable. This observation suggests that 50% of the reactive decisions performed by the BGP routing system (reactive routing) tend to further delay convergence of the forwarding paths. In turn, this observation provides the first indication that simple causal effects can't explain anymore the occurrence of instability. Moreover, more elaborated analysis techniques (such as the one proposed in this paper) are required to explain the inter-dependent routing and forwarding paths dynamics which affects their reliability and availability.**

Keywords—measurement; reliability; availability; analysis; routing; forwarding; paths; dynamics; Internet

I. INTRODUCTION

Internet routing path dynamics has been subject to many studies since middle of the 90's following the study of routing path behavior by V.Paxson [1]. The prominent research efforts [2] [3] [4] [5] conducted over last fifteen years to understand the root causes of Border Gateway Protocol (BGP) routing paths dynamics led to classify them as policy-induced or protocol-induced. This broad classification accounts for the distinction between BGP operations and the inherent behavior of the underlying path-vector routing algorithm which adapts its decisions upon routing topology changes. Because of BGP's slow convergence (phenomenon observed by C.Labovitz et al. [6]) and path exploration (behavior intrinsic to the shortest-path

vector algorithm), a single path change can typically induce the exchange of a large number of BGP messages.

Over time, many techniques have been proposed to mitigate the effects of (local) instabilities and/or to accelerate the convergence of the routing path(s) after occurrence of various perturbation events [7] [8] [9] [10]. Nevertheless, as recorded during the Routing and Addressing workshop held by the Internet Architecture Board (IAB) in 2006 [11], stability remains a key criterion to be met by the Internet routing system and its underlying Border Gateway Protocol (BGP). Moreover, the motivation for characterizing the dynamics of the Internet routing system does not only relate to the Internet routing system properties but also to 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 table 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.

On the other hand, many studies have been conducted to characterize the properties of the Internet forwarding paths including their length, their average length and the variation of their spatial properties. These studies are often realized by means of traceroute [12] and its variants such as the Paris traceroute [13] which improves the accuracy of traceroute. Tracetree tools [14] aims at improving the probing frequency, an important property when sources have a limited probing capacity which prevents them from issuing traceroute probes frequently enough to observe changes on all paths (probing all forwarding paths at the same frequency wastes probes on paths that are not changing while potentially missing changes in other paths). The DTRACK tool [15] which separates path change detection from path remapping performed only once a change is detected, aims at providing both accuracy and frequency tuning enabling in turn to track a larger number of paths. The same objective underlies the recently developed RADAR tool [14] using a set of distributed monitors.

The proposal to improve the availability of the Internet connectivity by means of reactive routing (compared to adaptive routing, which relies on routing information (ex)changes to select between paths) led to study the relationships between forwarding paths and routing paths properties. As observed by Z.M.Mao et al. [16], routing and

forwarding paths do not always match due in particular to route aggregation and forwarding anomalies. In [17], N.Feamster et al. studied the relationships between forwarding path failure and BGP instability induced by messaging. This study observed that i) forwarding path failures often precede BGP instability as BGP messages may indeed follow the appearance of a failure due to the slow adaptation of the BGP routing system upon failure occurrence but ii) BGP instability may also precede forwarding path failures when routing path changes signal external events such as network maintenance or where routing protocol engines themselves are the cause of such failures. The present paper aims at measuring and analyzing the dynamic properties between forwarding and routing paths. In particular, we show that simple causality effect as assumed in [17] is not verified anymore. Our analysis determines that the main cause of instability results from the forwarding plane, confirming the results reported in [17] and corroborating the assumption that the dynamic properties of the routing system are mainly driven by its adaptation to the forwarding system. However, about 50% of the forwarding path instabilities induce routing path instabilities whereas the corresponding forwarding path remains unstable. This observation suggests that reactive-like routing systems are now in place but 50% of their decisions tend to delay the convergence of forwarding paths (instead of only delaying convergence of routing paths through adaptive routing as reported by Z.M.Mao et al. [18]).

More generally, the overall motivations for studying the relationships between the dynamic properties (in particular, the (in)stability properties) of the Internet routing paths and their corresponding forwarding paths stem from three main reasons. The first translates the fact that transient but frequent changes in the spatio-temporal properties of routing paths may affect the performance and operating conditions of the corresponding forwarding paths (hence, their reliability). The second reason is that frequent instabilities when observed for the same (subset of) path(s) that can be attributed to a spatially localized portion(s) of the Internet may reveal that the underlying physical topology is more prone to failures (hence, showing limited reliability). The third results from the increasing operational need to provide for a longer term estimation of the Internet routing-forwarding system performance and operating conditions using well-proven statistical analysis accounting for recurrence of events and correlation between instability events.

The remainder of the paper is organized as follows. In Section II we outline the prior work realized in the characterization of routing and forwarding Internet paths and position our contribution. We detail the proposed processing and analysis method in Section III. The next Section IV provides the results obtained by the application of the method documented in Section III. A statistical analysis of the observation data is detailed in Section V together with the description of the nonparametric analysis method. Finally, Section VI draws conclusion from this study and outlines possible future work.

II. PRIOR WORK AND OUR CONTRIBUTION

A. Preliminaries

Consider a network topology modeled by an undirected graph $G = (V, E)$ where, the vertex set V ($|V| = n$) represents the finite set of nodes and the edge set E ($|E| = m$) represents the finite set of links. For $u, v \in V$, the loop-free path $p(u, v)$ from vertex u to v is defined as the finite sequence $[x_0(= u), x_1, \ldots, x_{i-1}, x_i, \ldots, x_p(= v)]$ such that the vertex x_{i-1} is adjacent to x_i, $\forall (x_{i-1}, x_i)_{i=1,\ldots,p} \in E$. The *length* $\ell(u, v)$ of the path $p(u, v)$ is defined as the finite number of edges the path $p(u, v)$ traverses from vertex u to v. The *diameter* $\delta(G)$ of the graph G denotes the maximum length of the shortest topological path between any two vertices $u, v \in V$, i.e., $\delta(G) = max_{u,v}\{\ell(u, v)|u, v \in V\}$. The *average path length* is defined as the average of the shortest path length, averaged over all pairs of nodes; thus, for an undirected graph G comprising n nodes, the average path length is provided by the formula $\frac{1}{n(n-1)}\sum_{\substack{u \neq v \\ u,v \in V}} \ell(u, v)$.

Distinction is made between topological, forwarding and routing paths. A *topological path* from vertex u to v is defined as the shortest path on the graph G where vertices u and v represent nodes (gateways, routers, etc.) and its edges represent the interconnection between abstract nodes (Autonomous Systems or AS) or links between routers. The *routing path* r_i towards destination i (where, $i = 1, \ldots, |D|$, D being the set of destinations) denotes the path $p(u, i)$ produced at node u by the routing algorithm using as input the information of the graph G. The routing topology defines thus a sub-graph of G representing the actual nodes and links along the paths as selected by the routing algorithm (and stored in local routing tables). A *forwarding path* f_i denotes the path $p(u, i)$ followed by the data traffic from node u to destination i. The forwarding path is derived at each node from the local routing table information. The forwarding topology defines thus a sub-graph of G representing the actual nodes and links as selected by each router's forwarding decision.

In order to characterize the dynamic properties of a given routing path r_i at time t, we use the quantity $\varphi_i(t)$ as defined in [19]. At time t, φ_i, referred to as the *stability metric* of the routing path r_i, quantifies the change(s) experienced by the routing path r_i from time $t - 1 = t_0 + k$ to $t = t_0 + (k + 1)$. The measurement time interval is thus determined by the integer k which accounts for the number of Minimum Routing Advertisement Interval (MRAI) time units (set by default to 30s) that have elapsed since the starting time of the measurement sequence t_0. In other terms, the metric φ_i accounts for all change(s) experienced by the routing path r_i with a periodicity determined by the MRAI time. For forwarding paths, an analogous quantity $\xi_i(t)$ is defined which characterizes the stability of forwarding path f_i at time t. At time t, ξ_i quantifies the change(s) experienced by the forwarding path f_i from time $t - 1 = t_0 + l$ to $t = t_0 + (l + 1)$, where the integer l accounts for the number of time units that have elapsed since the starting time of the measurement sequence t_0.

B. Prior Work

Many studies have been dedicated to determine the properties of the forwarding paths by means of traceroute [12] and its variants such as the Paris traceroute [13]. The latter improves the accuracy of traceroute by controlling the header fields of probe packets such that all probes towards a given

destination follow the same path in the presence of per-flow load balancing. Indeed, load balancing creates multiple simultaneous paths from a source to a given destination leading to traceroute errors and misinterpretation of path changes. On the other hand, tracetree tools aim at improving the probing frequency. The latter characterizes an important property of traceroute tools when monitoring sources have a limited probing capacity which prevents them from issuing traceroute probes frequently enough to observe changes on all paths, i.e., probing all paths at the same frequency wastes probes on paths that are not changing while potentially missing changes in other forwarding paths. The DTRACK tool [15] aims at providing both accuracy and frequency tuning by optimizing probing rate according to the likelihood of path changes. In turn, this tuning process enables to track a larger number of paths). For this purpose, the tool separates the tracking of path changes into path change detection and path remapping performed only once a change is detected. The same objective underlies the recently developed RADAR tool [14] which uses a set of distributed monitors as further explained in this paper.

Following the seminal paper of V.Paxson [1], prominent research efforts [2] [3] [4] [5] have been conducted to understand the dynamic properties of BGP routing paths and, in particular, the root cause(s) of their instabilities. Based on the proposal to improve the availability of Internet connectivity by means reactive routing instead of adaptive routing (such as BGP, which relies on routing information changes to select between paths), N.Feamster et al. reported in [17] several results on the relationships between forwarding path failure and BGP instability. In reactive routing, active measurements of network-layer path characteristics such as reachability, loss, and latency are used in combination with passive traffic monitoring to decide which paths are better; the differences are in how they take advantage of alternate paths. This study further observed that i) forwarding path failures often precede BGP instability as BGP messages may indeed follow the appearance of a failure due to the slow adaptation of the BGP routing system upon failure occurrence but ii) BGP instability can also precede forwarding path failures. BGP messages may actually precede a failure when routing changes signal external events such as network maintenance or when routing problems themselves are the cause of failures.

C. Our contribution

The study [17] reports that 80% of the forwarding path failures precede BGP messages by 2 to 4 minutes (even though BGP messages sometimes trail failures by 15 minutes or longer) and that 20% of the forwarding path failures follow BGP messages within 15 minutes. These observations are drawn from statistics where about 90% of the failures last less than 15 minutes and about 70% of all failures last less than 5 minutes. One can thus question if this temporal correlation between forwarding path failure and BGP messages (indicating routing path change) can be explained by a simple causal effect. On the one hand, the present study confirms the expected behavior: slow BGP convergence implies that the majority of the observed forwarding path instability is temporally correlated with routing path instability though over much larger time periods than the average reported in [17]. On the other hand, we don't observe the same fraction of routing

path instabilities preceding forwarding path instabilities but more importantly, our analysis shows that both instabilities can coexist as second order effect. Even if further evidences would be required to determine if this behavior is commonly observable, it provides a counter-example of the simple causality assumption elaborated in [17]. This implies that the causality effect as assumed by the above-mentioned study does not find anymore a simple explanation. Indeed, although the main cause of instability results from the forwarding path dynamics, about 50% of the forwarding path instabilities induce routing path instabilities whereas the corresponding forwarding path remains unstable. This observation suggests that reactive-like systems are now in place but 50% of their decisions tend to delay convergence of the forwarding paths (instead of only delaying the routing path convergence as discovered by Z.M.Mao et al. [18] when extending the adaptive capabilities of BGP to prevent routing path oscillations).

The contribution of this paper is threefold. First, we develop a systematic method to characterize the dynamic properties (in particular, the stability properties) of the routing paths and their corresponding forwarding paths. For this purpose, we associate tracetree-like measures of the forwarding paths data obtained using the RADAR monitoring tool [14] to the corresponding AS routing paths derived from the local BGP routing information bases. Indeed, following the observation reported in [16] that routing and forwarding AS paths do not always match due to route aggregation and forwarding anomalies, the complete mapping technique often leads to errors. Secondly, we analyze the resulting data recorded over long observation periods by means of multi-label classification. Finally, we perform a non-parametric statistical analysis of the data using the Recurrent Event Data Analysis (RDA) technique to determine the trend of observed rate of instability event as a function of time.

III. METHOD

In this section, we describe the proposed method which comprises three steps: the pre-processing of data sequences (Section III.A), the computation of the characteristic quantity translating path changes (Section III.B), and the analysis of the sequences of computed values (Section III.C).

A. Pre-Processing

Pre-processing is required to find the association between forwarding path data being sequences of IP addresses and routing path data being sequences of AS numbers (AS-Path attribute of BGP routes). To map the IP address (forwarding path) to the corresponding AS Number (routing path), we use the Whois [20] tool provided by the Team Cymru (http://www.team-cymru.org/). This tool, developed in C++ programming language, takes as input a file containing IP addresses and translates them into AS numbers as output.

It is important to mention that finding the association per destination between a forwarding and a routing path does not require performing the full mapping of data sequences but only to unambiguously associate a given forwarding path (IP address sequence) to its routing path (AS number sequence). Hence, pre-processing by association is simpler compared to methods that require performing full mapping of IP addresses

(forwarding paths) and AS number (routing paths). Moreover, this technique prevents introducing additional mapping errors when such operation is not strictly required.

B. Computational procedure

After having associated each forwarding path to its corresponding routing path (i.e., produce pairs of paths) one can then execute the following steps:

1. Compute for each routing path r_i the stability metric $\varphi_i(t)$ (as defined in Section II.1) at each MRAI time unit increment from time $t = t_0$ to $t_0 + K$, where the integer $K \geq 1$ defines the entire duration of the measurement period in MRAI time units;

2. Compute for the corresponding forwarding path f_i the stability metric $\xi_i(t)$ (as defined in Section II.1), from time $t = t_0$ to $t_0 + L$, where the integer $L \geq 1$ defines the entire duration of the measurement period;

3. Derive the associated statistics including the minimum and the maximum value of the stability metrics φ_i and ξ_i, in addition to their median, average and variance.
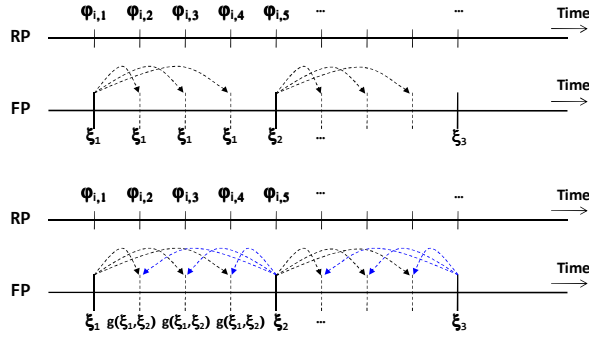

Fig. 1. Measurement rescaling technique

In case the periodicity between forwarding and routing path measurements differs, rescaling of the measurement time intervals is required. We first observe that the routing path measurement interval determined by the MRAI timer (set by default to 30s) is shorter than the forwarding path probe frequency which is of the order of 60s or higher. The naïve technique (as illustrated in the top of Fig.1) consists in extrapolating the value of the forwarding path stability $\xi_{i,1}$ as computed at time t during the entire time interval $[t, t+1]$ whereas the value φ_i associated to its corresponding routing path may change during this time interval. The problem with this approach is twofold. First, the effect(s) of the value computed at time t may decrease during the time interval between t and $t+1$ (attenuation effect). In addition, the value computed at the next iteration (i.e., at time $t+1$) could have a backward effect on the previous time interval between t and $t+1$. For this purpose, the proposed technique relies on the two-variable function $g: \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+: (\xi_1(t_0 + l), \xi_2(t_0 + l + 1)) \to (1-a)\xi_1(t_0 + l) + a\,\xi_2(t_0 + l + 1)$ where, the real parameter $a \in [0,1]$. At each intermediate time increment, the function $g(\xi_1, \xi_2)$ enables to combine the extrapolation of the value computed at time t (i.e., $\xi_1(t_0 + l)$) with the value

computed at time $t + 1$ (i.e., $\xi_2(t_0 + l + 1)$). This technique is further illustrated at the bottom of Fig.1.

C. Classification

Using the set of forwarding path (FP) - routing path (RP) pairs together with the corresponding sequence of values $\varphi_i(t)$ and $\xi_i(t)$ computed at each time increment, a first characterization of the observed instability events can be performed. It consists in determining whether the observed instability events can be detected or not at the forwarding path and/or the routing path level. For this purpose, each event (recorded for each (FP,RP) pair) is associated to one of the following qualifying labels: (FP_Stable, RP_Stable), $(FP\_Stable, RP\_Unstable)$, $(FP\_Unstable, RP\_Stable)$ and $(FP\_Unstable, RP\_Unstable)$. We then count the number of events for each pair and classify the pairs as follows:

1. (FP_Stable, RP_Stable): for each pair included in this class, most events translate stability of the routing path and its corresponding forwarding path.

2. $(FP\_Stable, RP\_Unstable)$: for each pair included in this class, the majority of the observed instability events are experienced by the routing path while the corresponding forwarding path remains stable (during the same time interval).

3. $(FP\_Unstable, RP\_Stable)$: for each pair included in this class, the majority of the observed instability events are experienced by the forwarding path while the corresponding routing path remains stable (during the same time interval).

4. $(FP\_Unstable, RP\_Unstable)$: for each pair included in this class, the majority of observed instability events involve both routing and forwarding paths. Moreover, further identification is required if a common segment is at the origin of the instability.

However, some (FP,RP) pairs can exhibit multiple instability patterns during the overall measurement period whereas capturing this behavior under a single (multi-value) label results in loss of information when classifying (FP,RP) pairs. Hence, to classify unambiguously and without loss of information all (FP,RP) pairs, the analysis procedure is further adapted to derive a dominant/main trend, i.e., the most representative behavior, and a sub-trend, i.e., the second most representative behavior. Following this procedure, each (FP,RP) pair is classified according to its dominant trend, i.e., associated to the class label for which the maximum number of counts has been computed, and then according to its sub-trend, i.e., associated to the class label for which the second maximum number of counts has been recorded. 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 median, average and variance of the stability metric. 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 additionally derive for each pair a dominant trend and a sub-trend with respect to their duration. Indeed, certain (FP,RP) pairs may have a very few number of counts but certain of them may be very long.

In this section, we detail the results obtained by means of the method documented in Section III and analyze them according to the classification procedure of Section III.C.

A. Datasets

The computation of the stability of the routing paths relies on the processing of the BGP datasets provided by the RouteViews project [21]. These datasets collected from the monitored BGP routers, comprise the following information: i) the complete Routing Information Base (RIB) entries (updated every two hours) and ii) the received BGP routing updates received from peering ASs (separated in files recorded every 15 minutes). This information is encoded by means of the Multi-Threaded Routing (MRT) format [22]. The monitored router is route-views.wide.routeviews.org

The computation of the stability of the forwarding paths is performed on the data recorded by the RADAR tool [14]. The measurements carried out by this tool are traceroute-like probes initiated from a set of monitoring nodes. Such probes target a large set of IP address prefixes distributed across the Internet. Based on these measurements, the RADAR tool builds ego-centered views of the forwarding topology, i.e., the initiating router collects traces along the forwarding paths that it probes). A subset of the forwarding paths traced by the RADAR tool corresponds to the routes obtained from the RouteViews dataset. Consequently, a subset of the monitored AS routing paths is also monitored by the RADAR tool. In total, the analyzed dataset includes a bit less than 1000 forwarding path - routing path pairs covering a monitoring period of 50 days.

B. Observations

The results obtained by applying the computational procedure documented in Section III.B to each (FP,RP) pair are best represented by graphics. For this purpose, we provide in Fig.2 an example of a (FP,RP) pair belonging to the class $(FP\_stable, RP\_unstable)$, in Fig.3 an example of a (FP,RP) pair belonging to the class $(FP\_unstable, RP\_stable)$ and in Fig.4 an example of a (FP,RP) pair belonging to the class (FP_stable, RP_stable). In these figures, the Y-axis represents the stability value computed by means of the procedure specified in [19] and the X-axis the time (10000 time units correspond to a period of 50 days). Their interpretation relies on the value of the Y-axis: as this value increases (decreases), the instability of the corresponding path increases (decreases).

C. Analysis

Table I lists the number of (FP,RP) pairs for the dominant behavior/trend and the sub-trend as obtained following the procedure of Section III.C. As it can be observed from this table, the majority of the (FP,RP) pairs experiencing instability belong to the $(FP\_unstable, RP\_stable)$ class, which determines the dominant instability trend. Indeed, the instability events observed for about 95% of the pairs reveal forwarding path instability without routing path instability. Less than 5% of the (FP,RP) pairs are labeled as $(FP\_unstable, RP\_unstable)$. Thus, for the small fraction of (FP,RP) pairs experiencing instability events, the majority of these events corresponds to both forwarding and routing path instability. Interestingly, as recorded in columns 3 and 4 of

Table I, the sub-trend indicates that about 50% of the (FP,RP) pairs experience instability at both forwarding and routing levels. The remaining 50% of the pairs belong to the (FP_stable, RP_stable) class.
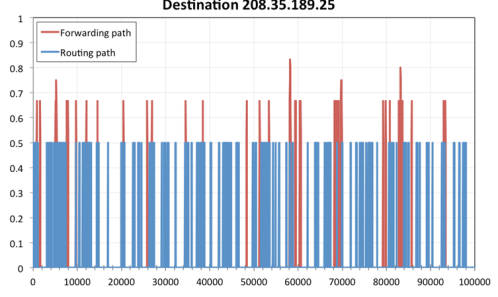

Fig. 2. Example of (FP_stable,RP_unstable) pair

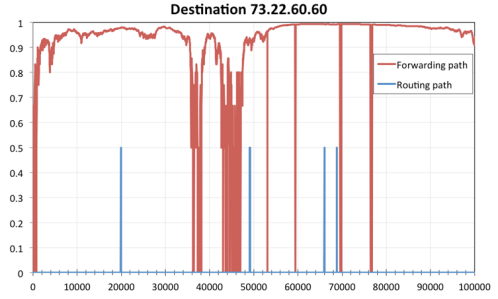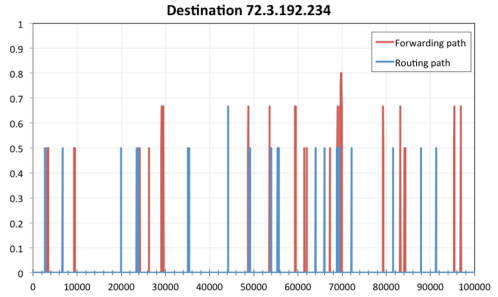Fig. 3. Example of (FP_unstable,RP_stable) pair

Fig. 4. Example of (FP_stable,RP_stable) pair

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 simply 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 instability events. On the other hand, the other 50% of the pairs return to the situation where both forwarding and routing paths are stable.

Certain (FP,RP) pairs may have a very few number of counts but certain of them may last for a (relatively) longer time period compared to others. For this purpose, we also compute for each pair the cumulated percentages of time (in MRAI time units) during which instability events have been observed. Fig.5 plots the cumulated number of (FP,RP) pairs

over the percentage of time of the entire observation period during which instability events have been observed for each pair (i.e., either at the forwarding or the routing path level).

Main trend	Number of Pairs	Second trend	Number of Pairs
FP stable - RP stable	////////	FP stable - RP stable	445
FP stable - RP unstable	12	FP stable - RP unstable	58
FP unstable - RP stable	906	FP unstable - RP stable	3
FP unstable - RP unstable	32	FP unstable - RP unstable	444

We can observe from the Cumulative Distribution Function (CDF) of Fig.5 that for 62% of the (FP,RP) pairs, the cumulated percentage of time during which instability is observed accounts for less than 10% of the total observation time. In other terms, the majority of the (FP,RP) pairs are stable during 90% of the total observation time. From the same figure, we also observe that for about 77% of the pairs, the percentage of time during which they experience instability accounts for up to 20% of the total observation time, i.e., during 80% of the time 77% of the (FP,RP) pairs are stable. The latter percentage increases to 50% when reaching about 50% of the total observation time. Thus, for 10% of the (FP,RP) pairs the percentage of time during which they experience instability is larger than the percentage of time during which they are stable (no instability event is observed).

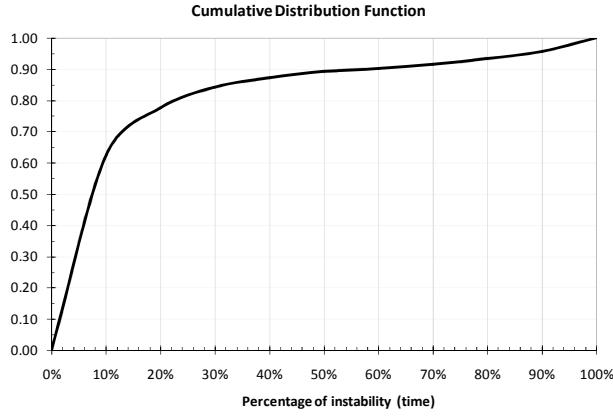**Cumulative Distribution Function**

Fig. 5.   Cumulative distribution function vs. Percentage of instability (time)

V.   NON-PARAMETRIC STATISTICAL ANALYSIS

In our study, the set of (FP,RP) pairs are followed over a period of time, the observation period (of 50 days). For each (FP,RP) pair part of the set, data are collected at multiple time points, i.e., multiple or repeated measurements of the same quantities (φ_i and ξ_i) are performed on each pair and recorded for each pair over a period of time.

Recurrent event data arise when observed elements possibly experience more than one event during the observation period (of 50 days). A key characteristic is that observations for the same pair or subset of pairs may be correlated. Moreover, the temporal trajectories of observed data are often very complex to determine. Consequently, parametric statistical models may not be flexible enough to capture their main features; instead, nonparametric (or semi-parametric) statistical models are particularly attractive in such study. In these models, the mean structures are modeled non-parametrically (or semi-parametrically) and the distributional assumptions are assumed to be non-parametric.

A. Recurrent Event Data Analysis (RDA)

Recurrent Event Data Analysis (RDA) is commonly used in various engineering fields and is particularly useful when performing reliability analysis of repairable systems. Conventional Life Data Analysis (LDA) assumes that events (failures) are independent and identically distributed (i.i.d) whereas in the present case, the instability events are dependent and not identically distributed (common property of repairable system data). Moreover, we are interested in modeling the number of occurrences of instability events over time rather than the length of time prior to the first event as in LDA, which typically focuses on time to event occurrence data. Non-parametric RDA provides a non-parametric graphical estimate of the mean cumulative number of recurrences of events per (FP,RP) pair versus time. This non-parametric analysis method relies on the *Mean Cumulative Function* (MCF).

B. Mean Cumulative Function

When analyzing repairable systems, the simplest plot that can be constructed is a cumulative plot (staircase function), which graphs the number of recurrences of events over time. In the present case, as we have a set of pairs, each pair can be represented by a cumulative plot. It is also possible to represent the behavior of the set of pairs by computing the average of the cumulative number of recurrences of events over time. This average is referred to as the Mean Cumulative Function.

Compared to Mean Time Between Failure (MTBF), the MCF method provides the following advantages i) the MCF is more adequate to represent the event rate because it makes no distributional assumptions (nonparametric method) and ii) the MCF is also more informative because it provides trends in the event rate as a function of time. The MCF reveals also one important aspect of path availability by estimating the expected rate of instability events, i.e., the (cumulated) number of instability events that the paths would experience over time.

1) Computation

To compute the MCF $M(t_i)$ (more precisely, the estimation of the MCF) one proceeds as follows. At each observation time t_i, the number of events n_i that occurred since the previous observation time t_{i-1} is recorded. These events are recurrent (i.e., non-fatal like instability events) and are assumed to occur randomly. The number of events n_i is divided by the number ρ_{i-1} of pairs that are observable at time t_{i-1} (with ρ_1 set to the total number of initially observable pairs). One then computes the MCF estimate using the formula:

$$M(t_i) = \frac{n_i}{\rho_{i-1}} + M(t_{i-1}) \qquad (1)$$

$$M(t_1) = \frac{n_1}{\rho_0} \qquad (2)$$

It is important to note that all observation intervals are taken to be nonrandom, identical for all observations and of equal length. Moreover, since the MCF is an estimate, confidence bounds can also be computed.

2) Analysis

The shape of the MCF plot can reveal several important properties about the behavior of the recurrent events under consideration in this study. Moreover, the MCF versus time curve can be numerically differentiated to obtain the slope, called the *recurrence rate*.

From the shape of the MCF plot, one can then derive the following interpretation assuming that instability events induce transient changes in paths properties that affect their performance and operating conditions (hence, their reliability):

- Constant recurrence rate: the MCF plot increases monotonically, the slope of the MCF remains constant; thus instability events occur at constant rate.

- Increasing recurrence rate: the MCF plot is convex (the slope of the MCF increases), the recurrence rate increases over time which reveals system performance degradation over time.

- Decreasing recurrence rate: the MCF plot is concave (the slope of the MCF decreases), the recurrence rate decreases over time which reveals maintenance improvement over time (decreasing repair rate).

3) Results

Applying the procedure documented in Section V.B.1 to our dataset yields the MCF plot of Fig.6. This result indicates that both forwarding and routing paths experience instabilities at "constant" recurrence rate (though the rate experienced by the forwarding paths is about 10 times higher than the rate experienced by the routing paths).

Mean Cumulative Function

$y = -2E-06x^2 + 0.052x + 0.013$
$R^2 = 0.998$

— Forwarding Paths
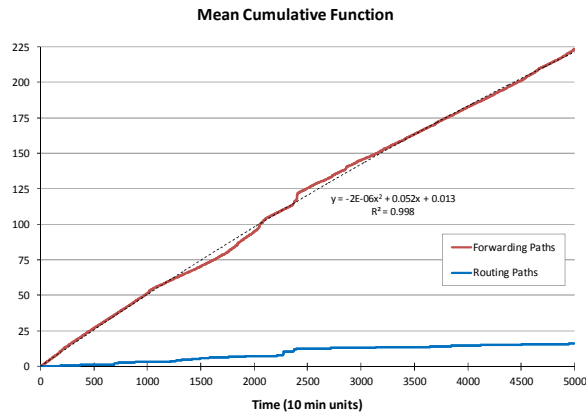— Routing Paths

Time (10 min units)

Fig. 6. Mean Cumulative Function (MCF)

VI. CONCLUSION

In this paper, by analyzing the routing and forwarding paths stability properties, we have captured further evidence that the observed dynamics of the Internet routing and forwarding system show different properties. Hence, one can't derive the stability behavior of the one from the other. We further investigate the relationships between the stability of the forwarding paths followed by the traffic and the corresponding routing paths as selected by the path-vector routing protocol. For this purpose, using the method developed in [19], we relate at the router level the stability measurements carried on the forwarding paths with the corresponding routing paths.

Our 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, confirming the results reported in [17]. This observation further corroborates the assumption that the dynamic properties of the routing system are mainly driven by its adaptation to the forwarding system. However, it can also be observed that a second order effect relates forwarding and routing path instability events. This observation shows that the causality effect assumed by [17] does not find anymore a simple explanation. Indeed, forwarding paths become the dominant source of instability; however, about 50% of the forwarding path instability events induce routing path instability whereas the forwarding path remains unstable. This observation suggests that reactive-like systems being now in place, 50% of their decisions tend to delay the convergence of the forwarding path (instead of only delaying the routing path convergence through adaptive routing).

As part of our future work, we will extend the analysis method to localize and characterize intra-AS instabilities. We will also further improve the procedure for computing path dynamics properties in order to integrate the spatial effects resulting from the AS/node position in the Internet topology.

REFERENCES

[1] V.Paxson, End-to-End Routing Behavior in the Internet. IEEE/ACM Transactions on Networking, vol.5, no.5, pp.601–615, 1997.

[2] C.Labovitz, R.Malan, F.Jahanian, Origins of Internet Routing Instability, Proc. of IEEE INFOCOM 1999, New York (NJ), March 1999.

[3] C.Labovitz, A.Ahuja, A.Bose, F.Jahanian, Delayed Internet Routing Convergence, IEEE/ACM Transactions on Networking, vol.9, no.3, pp.293-306, June 2001.

[4] T.Griffin, F.B.Shepherd, G.Wilfong, The Stable Paths Problem and Interdomain Routing, IEEE/ACM Transactions on Networking, vol.10, no.1, pp.232-243, April 2002.

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

[6] C.Labovitz, A.Ahuja, F.Jahanian, Experimental Study of Internet Stability and Wide-Area Backbone Failures, Proc. of 29th International Symposium on Fault-Tolerant Computing, June 1999.

[7] C.Villamizar, R.Chandra, R.Govindan, BGP Route Flap Damping, Internet Engineering Task Force (IETF), RFC 2439, November 1998.

[8] J.Chandrashekar, Z.Duan, Z.-L.Zhang, J.Krasky, Limiting path exploration in BGP, Proc. of IEEE INFOCOM 2005, Miami (FL), March 2005.

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