Autonomous System Isolation under BGP Session Attacks with RFD Exploitation*

Kotikalapudi Sriram†, Doug Montgomery, Oliver Borchert, Okhee Kim, and Rick Kuhn
National Institute of Standards and Technology (NIST)
Gaithersburg, Maryland 20899
Email: {ksriram, dougm, borchert, okim, rkuhn}@nist.gov

September 12, 2005

Abstract

There is a growing apprehension in the Internet community that there are potentially significant vulnerabilities in the deployed Border Gateway Protocol (BGP) routing system. Researchers speculate and debate the potential of targeted attacks to trigger large scale, potentially cascading, failures and persistent instability in the global routing system. To date, most modeling and analysis of BGP behavior under threatening scenarios has been limited to post mortem analysis of global routing exchanges during worm and virus attacks of Internet hosts; but these are not attacks focused on BGP. In this paper, we present results from our effort to conduct "what if" analyses of yet unseen attacks and to develop means to characterize the impact of various attacks on a distributed BGP routing system. In particular, we present a detailed study of the impact of BGP peering session attacks and the resulting exploitation of RFD that cause network-wide routing disruptions. Analytical results provide insights into the nature of the problem and impact of the attacks. Detailed packet level simulation results complement the analytical results and provide many useful insights as well. We also quantify the effect of BGP Graceful Restart mechanism on partial mitigation of the BGP vulnerability to peering session attacks.

Keywords: Border Gateway Protocol (BGP), Route Flap Damping (RFD), Internet Routing Protocol Security, Performance Modeling, BGP Graceful Restart

---

*This research was supported by the Department of Homeland Security under the Secure Protocols for the Routing Infrastructure (SPRI) program and the NIST Information Technology Laboratory Trustworthy Networking Program.

†Corresponding author: Dr. Kotikalapudi Sriram, 100 Bureau Drive, Stop 8920, NIST, Gaithersburg, MD 20878. Tel: +1 301 975 3973, Email: ksriram@nist.gov

1 Introduction

There is a growing apprehension in governments and the Internet industry that there are potentially significant vulnerabilities [1]-[17] in the deployed Border Gateway Protocol (BGP) routing system [18][19]. While to date there have been few, if any, serious focused attacks on the BGP infrastructure, researchers speculate and debate the potential of targeted attacks to trigger large scale, potentially cascading, failures and persistent instability in the global routing system [2]-[12]. In response to this situation, numerous proposals have been developed that attempt to provide varying levels of protection and assurance to various aspects of BGP's operation [20]-[23]. Each of these proposals implicitly embodies a somewhat different view of the attributes of the problem space and the practical constraints of the solution space. Unfortunately, the lack of a shared understanding of both the risks associated with focused attacks and the cost-benefit tradeoffs of various mitigation techniques will likely doom prospects for the rapid development and wide spread adoption of a comprehensive set of solutions. It may also be noted that there are some efforts to even fundamentally rethink the design of BGP and the control plane design from a security point of view [24][25].

Our efforts at NIST are addressing this issue by developing tools and techniques to help the community identify and characterize the risks associated with focused attacks on the BGP infrastructure, and to evaluate the effectiveness and impact of various proposed mitigation techniques.

To date, most modeling and analysis of BGP behavior under threatening scenarios has focused on post mortem analysis of global routing tables during worm and virus attacks of Internet hosts [26]-[28]. Unfortunately (or fortunately, depending upon your perspective), there are no known live data or traces from large-scale attacks that were targeted at BGP itself. In order to fill this void, we have developed a simulation capability to model large scale attacks specifically focused on the BGP infrastructure. Our goal is to conduct "what if" analyses of yet unseen attacks and to develop means to characterize the impact of various attacks on a distributed BGP routing system. Of particular interest is the discovery of potential global emergent behaviors (e.g., cascading failures, persistent oscillations, permanently degraded routing) induced by successful local attacks, and the identification and evaluation of new BGP threat scenarios. We have extended the SSFNet BGP simulation modeling tools [29][30] to include an attack-modeling framework capable of generating arbitrary attacks with parameterized form, intensity, behavior, extent and duration. In addition, we have developed metrics and an attack analysis framework capable of characterizing the impact of successful attacks in terms of their effects on global routing and the detailed operation of the BGP protocol [16][31]. Our analysis framework supports both macroscopic characterizations of network wide cumulative response and microscopic characterizations of "peering session" level transient behaviors.

In [16] we presented an overview of our attack modeling framework and simulation tools, and some preliminary results and observations from our studies of large scale BGP peering session attacks. The simulation tool has the capability to simulate several hundreds of Autonomous Systems (AS). In this paper, we study the impact of

focused BGP peering session attacks and present simulation and analytical results. Through these results, it is revealed that malicious attackers could exploit Route Flap Damping (RFD) mechanisms to amplify the duration of AS-to-AS or AS-to-prefix isolations. RFD is a method for receiver side route monitoring and suppression in the event of frequent updates [32][33]. However, that benefit has the flipside in that by sustained peering session attacks into various BGP sessions in an AS-path or into a portion of a network with many AS paths, attackers can cause isolation of ASes at the two ends of the attack region. We show that this is potentially a serious type of denial of service (DOS) attack, which is amplified by the particulars of BGP behavior (namely, RFD), and present a detailed quantitative analysis of its impact. Our analytical results provide insights into the nature of the problem and impact of the attacks. Detailed packet level simulation results complement the analytical results and provide many useful insights as well. We also quantify the effect of BGP Graceful Restart mechanism on partial mitigation of the RFD-based BGP vulnerability.

In Section 2, we present an understanding of how random peering session attacks may trigger the RFD penalty and cause routes to enter the RFD suppression state. In Sections 3 and 4, the analytical model of RFD behavior during BGP session attacks and numerical results based on the analysis are presented, respectively. An analysis of the benefits of using BGP Graceful Restart (BGP-GR) mechanism [34] is presented in Section 5. Our SSFNet-BGP based attack simulation framework and models are described in Section 6. Section 7 deals with simulation results and their discussion. We state our conclusions in Section 8.

2 BGP Attacks with Exploitation of Route Flap Damping

We start here by providing brief introductions to the principles of BGP Minimum Route Advertisement Interval (MRAI) and the Route Flap Damping (RFD). These collectively play a role in the models we develop to characterize the impact of peering sessions attacks and concomitant RFD exploitation. MRAI is a sender-side peering discipline designed to control the BGP update-processing load. Values of MRAI are randomly chosen in the range of 22.5 s to 30 s on per peer basis. As shown in Fig. 1, router R1 receives two BGP updates, U-A and U-B, in succession from one or more peers. These updates may arrive temporally close to each other at R1, but the MRAI at R1 causes them to be separated from one another by at least the MRAI time in their propagation to peer router R2. This may result in R1 receiving and processing additional updates from other peers during the time when it has sent U-A and is waiting to send U-B. Thus, there is a chance for multiple incoming updates to be coalesced into a single outgoing route update to upstream peers. Consequently, the quantity and rate of updates are potentially reduced.

RFD is a method for receiver side route monitoring and suppression of oscillations or unstable paths. An upstream router assigns an incremental RFD penalty to a peer and destination (i.e., prefix) combination each time it receives a BGP update pertaining to that combination. If the RFD penalty exceeds a preset cutoff threshold, then the route is suppressed and withdrawals are sent to neighbors about the prefix in question. The

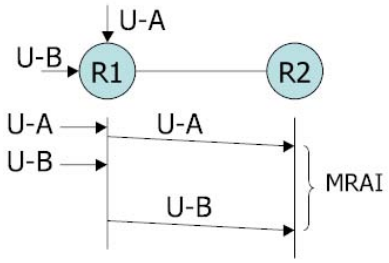

Figure 1: The role MRAI in update propagation.

RFD penalty is allowed to decay exponentially with a chosen halftime (i.e., decay constant). When it drops below a chosen reuse threshold, then the route is no longer suppressed and updates are once again processed for the peer-prefix combination in question. Table 1 shows the values of various RFD parameters for two common commercial implementations labeled as Vendors A and B; the two sets of numbers are later used in this paper for a sensitivity study relative to RFD parameters.

There are many different attack possibilities on the BGP routing infrastructure; a BGP attack enumeration is provide in [17]. We focus on attacks that cause the BGP peering sessions to be reset. A common way to reset a BGP peering session is to reset or attack the underlying Transmission Control Protocol (TCP) connection. There are several known vulnerabilities associated with TCP and the Internet Control Message Protocol (ICMP), which could be exploited to cause TCP connection-reset attacks [11]-[15]. On example is the "slipping in the window" TCP reset attack, which received a lot of attention recently [14]-[15]. The success of this attack depends on the attacker's ability to correctly guess a TCP sequence number within a TCP flow control window. Spoofed ICMP error messages to cause TCP reset have also been brought to attention recently [11]. The ICMP based attacks causing TCP resets do not require guessing the TCP sequence number. Hard or soft ICMP error messages can be potentially spoofed to cause TCP resets. The details regarding ICMP attacks against TCP are discussed in [11]. Fig. 2 illustrates how random BGP peering session attacks can lead to

Table 1: RFD parameter values.

RFD Parameter	Vendor A	Vendor B
Withdrawal penalty	1000	1000
Re-advertisement penalty	0	1000
Attribute change penalty	500	500
Cutoff threshold	2000	3000
Half time	900 s	900 s
Reuse threshold	750	750
Max suppress time	3600 s	3600 s
Max penalty	12000	12000

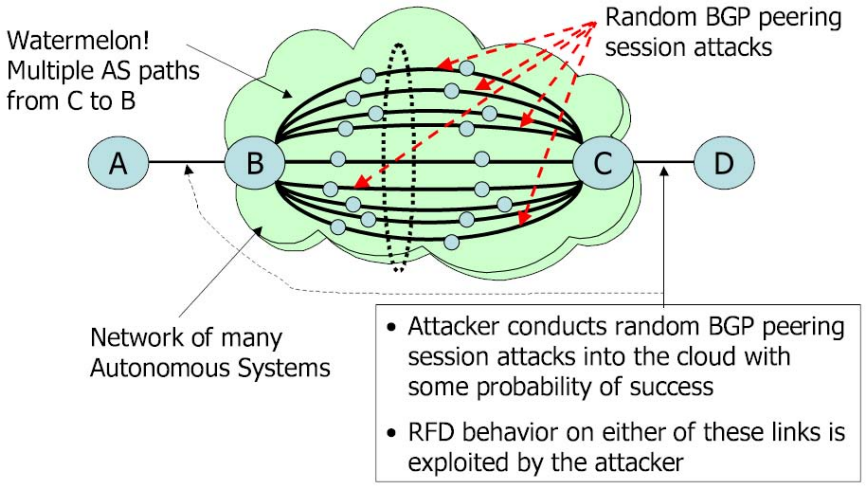

Figure 2: Illustration of random BGP peering session attacks.

exploitation of the RFD by attackers, and cause prolonged AS-prefix isolations. Here each node in the network represents an AS as well as a prefix (i.e., destination). The figure illustrates that there may be multiple AS paths between ASes B and C. What is in the cloud is a network of many ASes. The attackers are assumed to have some capability available that they use to launch BGP peering session attacks into the network, and depending on their method and collective resources, there is a measurable probability of success for each peering session attack. A successful attack would cause a peering session to be terminated, and cause the affected ASes to send withdrawals about all the prefixes in their routing tables that are rendered unreachable. The RFD behavior on links A-B and C-D would be exploited due to the attacks, and major outages (isolations) can result due to: (1) D putting (C,B) peer-prefix combination and all prefixes reachable via B under RFD suppression, and (2) likewise A putting (B,C) peer-prefix combination and all prefixes reachable from C under RFD suppression.

The details of how this suppression works are further explained with the help of Fig. 3, where a linear topology (representing a single AS path), consistent with the individual alternate AS paths between B and D in Fig. 2, is considered. In Fig. 3, the progression of BGP updates horizontally from left to right is in the spatial dimension (hop to hop), and vertically from top to bottom is the progression of time. The figure shows three BGP peering session attacks happening on three different hops (B-I, J-K, K-L) in the AS path at different times. It shows the flow of updates, classified as either Withdrawals (WD), Re-advertisements (Re-Adv), or Attribute-Change (AttrCh)). The BGP nodes along the way cause the updates to be separated by MRAI intervals. It is assumed that the peering session that was attacked is able to recover within a short time as compared to the MRAI. This quick recovery can occur, for example, when BGP session is forced to terminate by a TCP reset attack. The BGP session is automatically reestablished immediately after the TCP connection is restored between the affected peers. An attack causes a withdrawal to be sent to neighbors and a recovery causes a Re-Adv to be sent. When C receives a withdrawal about B, if an alternate path is available (see Fig. 2), then C sends an Attribute-Change update to D informing D of an alternate route to B. When C receives a Re-Advertisement

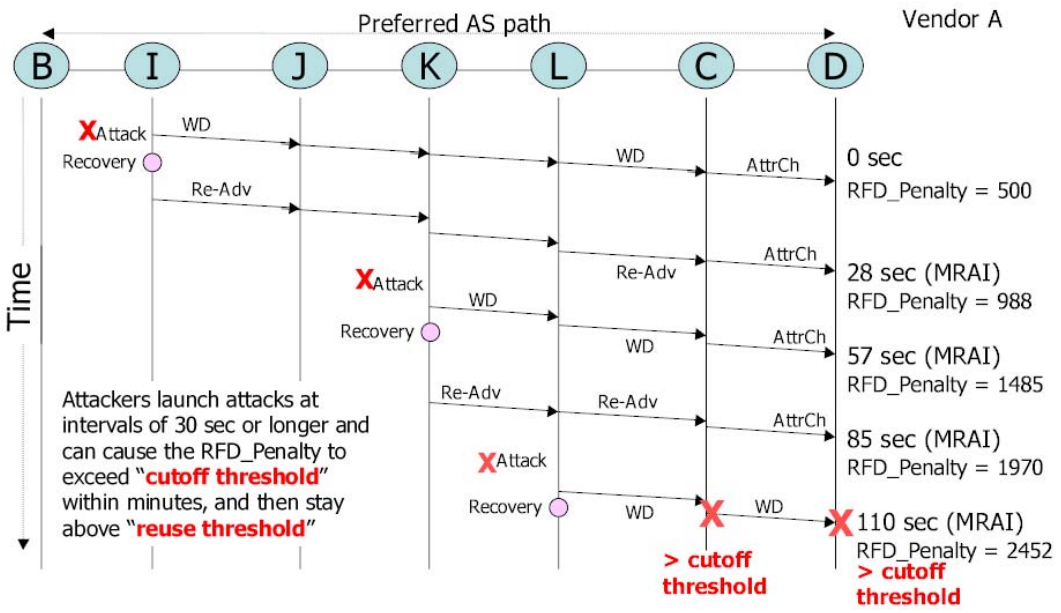

Figure 3: Illustration of update message propagation and RFD penalty accumulation for the preferred AS path between nodes B and D.

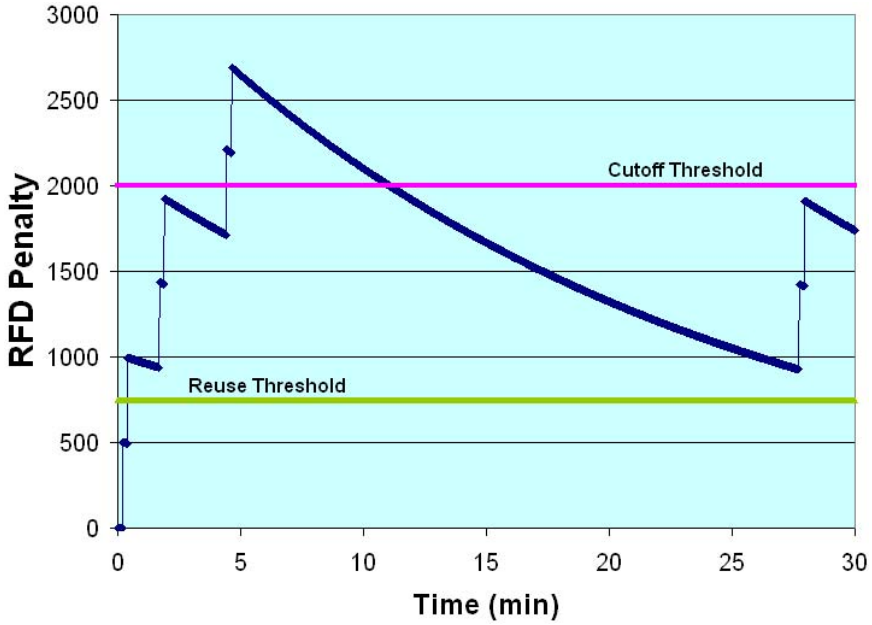

Figure 4: Illustration of RFD penalty based cutoff and recovery.

about B, it reverts to the previous path, and again sends an Attribute-Change update to D informing D of the reversal to the previous path. While these updates about B from peer C are received at D, the RFD penalty for B at D via peer C increases, and exceeds the cutoff threshold following just three attacks as shown in the Figs. 3 and 4. B and all prefixes reachable via B are suppressed at D; this essentially isolates D from B and prefixes via B until the RFD penalty decays below the reuse threshold. In most implementations of BGP, if

another attack is launched along the AS path before the RFD penalty reaches below reuse threshold, then the suppression (and isolation) continues even longer (see Fig. 4). When an update is regarded as a flap, the RFD penalty will be incremented due to the update even when the RFD is in a decay mode.

3 Analytical Model for Peering Session Attacks Triggering RFD Cutoff

The purpose of the analytical model presented here is to predict the probability of AS-prefix isolation under suitable assumptions regarding the attack characteristics. As described earlier, AS-prefix isolations are the result of the RFD penalty exceeding the cutoff on all alternative paths between the AS and the prefix. Here we assume that the attacks happen independently on any of the BGP peering sessions in the cloud of Fig. 2. Thus chances are that the RFD penalty will exceed the cutoff on different AS paths between BGP routers B and C in quick succession of one another. Thus, it is a reasonable approximation if we derive the probability of RFD penalty exceeding the cutoff for the longest AS path between B and C, and approximate that to the AS-prefix isolation probability in question. It can be expected that this would be in fact a fairly accurate and slightly conservative approximation.

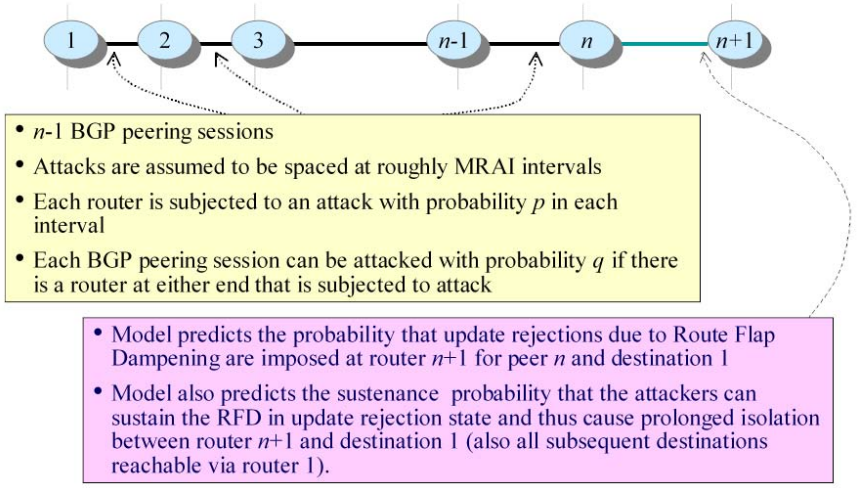

Figure 5: Analytical model for AS-prefix isolation probability.

As shown in Fig. 5, we model the AS path between the end-points of interest as n-1 BGP peering sessions (BGP router 1 to BGP router n). For purposes of modeling, we assume (with out loss of generality) that the peering session between BGP routers n and n+1 does not come under attack. The RFD penalty at BGP router n+1 is to be modeled in order to determine the probability of isolation at BGP router n+1 in relation to the peer router n and destination router 1 and prefixes reachable via 1. Attacks spaced closer than MRAI interval do not speed up the time to isolation and hence it is meaningful to assume that the attacks would be

spaced approximately at MRAI time intervals. We assume that a router's control plane may be compromised with probability p and an associated BGP peering session may be then attacked with probability q. Thus the probability of a successful BGP peering session attack is $Q = pq$. The model we describe below derives the probability that update rejection and hence withdrawal happen due to RFD at BGP router n+1 for peer n and destination 1. The model also predicts the probability that the attackers can sustain the RFD in update rejection state and thus cause prolonged isolation between router n+1 and destination 1.

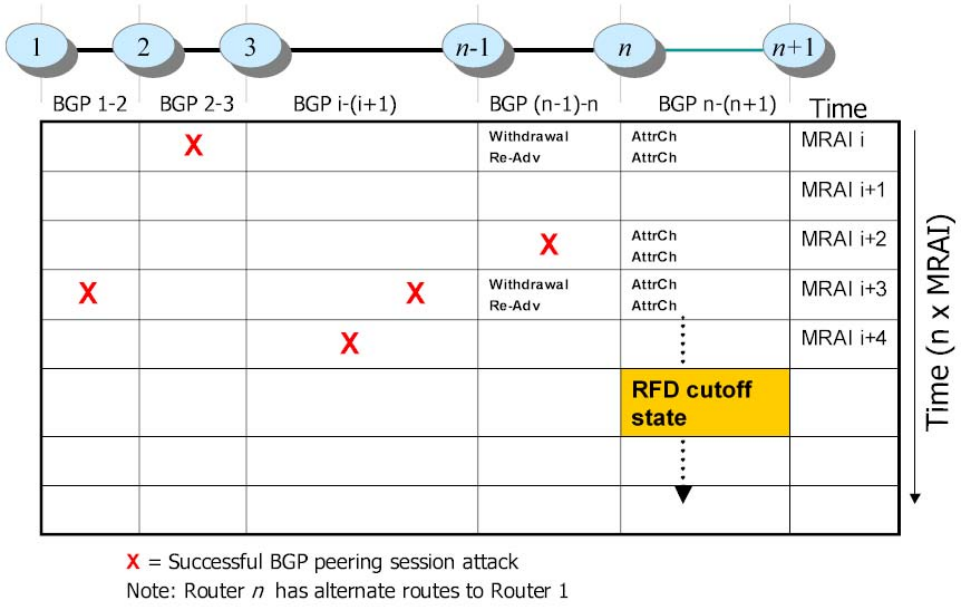

Figure 6: Time-space model for relating RFD penalty accumulation to attacks and derivation of probability of successful AS-prefix isolation.

The basic principle of the time-space model of the attacks and RFD penalty accumulation can be explained with the help of Fig. 6. In the time-space matrix illustrated in Fig. 6, the cells represent a hop (or BGP peering session) location and the time in multiples of MRAI. Although the MRAI is a variable in the range 22.5 s to 30 s, here we assume it be a constant with a fixed value of 26 s. The X's in the figure represent peering session attacks. Just to get a conservative estimate of how soon the RFD penalty could accumulate, we make two reasonable assumptions: (1) attacks occur at intervals of MRAI (speedier attacks do not benefit the attackers from an RFD cutoff point of view because MRAI flow controls the updates), and (2) When an attack happens, no other updates have recently happened within approximately an MRAI time so that the updates resulting from the attack in consideration propagate quickly across from left to right in Fig. 6 well within an MRAI time interval. The second assumption is aided in its accuracy partially due to the first assumption. The goal of the modeling is to probabilistically predict the time (in multiples of MRAI intervals) in which the RFD penalty of interest exceeds the cutoff threshold at node n+1.

To describe the stochastic model, let us define the following parameters:

C = cutoff threshold,

R = reuse threshold,

H = halftime (decay parameter),

T = MRAI time,

P = incremental RFD penalty incurred per successful attack event,

n = number of BGP nodes in the AS path subject to attacks (see Fig. 6),

Q = Prob.{a BGP peering session attack is successful},

θ = Prob.{AS path of n ASes is successfully attacked at one or more BGP peering sessions},

E = Elapsed time from the time of beginning of BGP session attacks (in multiples of MRAI),

$R_P(n+1; n, 1; iT)$ = RFD penalty at router n+1 for peer n and destination 1 at time iT,

$\alpha(n, k)$ = Prob.$\{R_P(n+1; n, 1; iT) > C$ for some $i \in (0, k) \mid E = kT\}$.

In the above definitions, the incremental RFD penalty, P, incurred per successful attack is assumed to be one number even though Table 1 shows different penalty values for different types of updates. This is because in the system we are modeling, each BGP session attack eventually produces a pair of attribute-change updates between nodes C and D (see Fig. 3) or between nodes n and n+1 in our analytical model in Fig. 6. These attribute-change updates are still decipherable by the receiving peer as corresponding to withdrawal (implicit) and re-advertisement. In effect, corresponding to each successful attack, the net incremental penalty, P, will be 1000 (=1000 + 0) for the case of Vendor A and 2000 (=2*1000) for the case of Vendor B. The key performance metric of interest in this analysis is $\alpha(n, k)$, the probability that the attackers can cause RFD triggered AS-prefix isolations in kT time interval or less.

From Fig. 6, given that there are n BGP peering sessions in the AS-path of interest, it can be deduced that the probability, θ, of a successful BGP attack on the (n-1)-hop AS path in an MRAI interval is given by

$$\theta = 1 - (1 - Q)^{n-1}. \tag{1}$$

The probability, $\beta_i(n, k)$, that there are i successful attacks on the (n-1)-hop AS path in time kT (equivalently, k successive MRAI intervals), is given by

$$\beta_i(n, k) = \frac{k!}{i!(k-i)!}\theta^i(1-\theta)^{k-i}, \quad i \leq k. \tag{2}$$

Let us define k_m as the absolute least number of successful BGP attacks needed on the AS-path in consideration to cause the RFD penalty to exceed the cutoff threshold, C. Then, we have

$$k_m = \lceil \frac{C}{P} \rceil. \tag{3}$$

Clearly, if k_m consecutive MRAI intervals have successful peering session attacks, and if there were no exponential decay, then the cumulative penalty would meet or exceed the cutoff threshold, C. However, in reality, the exponential decay of RFD penalty must be taken into account (see Fig. 7). If the attacks are

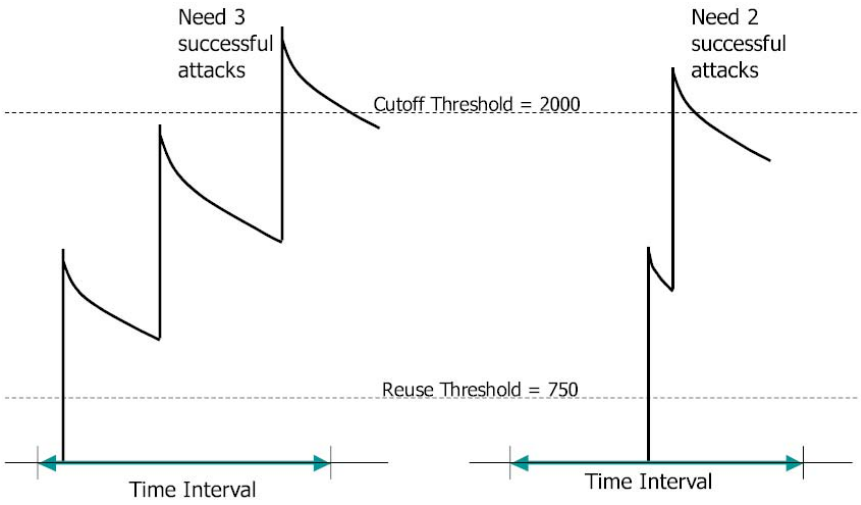

Figure 7: Model for estimation of attacks needed to push penalty above cutoff.

bunched together (closely spaced), and located towards the beginning or the end of the k-MRAI interval, then the decay would be too much or too little, respectively, and would not lead to a realistic estimate of minimum number of attacks needed to exceed the cutoff. Thus, while taking the exponential decay into consideration, it is reasonable to assert that the attacks in the k-MRAI time period can be spread nearly evenly to derive a realistic estimate of the minimum number of attacks, $j_{min}(k)$, needed to meet or exceed the cutoff threshold, C (see Fig. 7). Thus, for a given k, the $j_{min}(k)$ can be estimated by finding the smallest integer j for which the following inequality is satisfied:

$$P\sum_{i=0}^{j-1} 2^{\{-\frac{ikT}{(j-1)H}\}} > C, \quad k \geq k_m. \tag{4}$$

Once $j_{min}(k)$ is known, then the key performance of metric interest, $\alpha(n,k)$, is derived as follows:

$$\alpha(n,k) = \sum_{j_{min}(k)}^{k} \beta_i(n,k), \quad k \geq k_m. \tag{5}$$

Based on available BGP protocol descriptions in the literature[32][33], it appears that the RFD specifications require the penalty be incremented even when the RFD is in a cutoff (suppression) state, provided that the received update is a flap. As a result, it is possible that, if another attack is launched along the AS path before the decaying RFD penalty reaches below the reuse threshold, then the RFD penalty may be incremented further until the maximum penalty value (12000) is reached (see Table 1). Thus, route suppression and AS-prefix isolation continue even longer (as previously noted in Fig. 4). We define the probability of sustenance, P_{sus}, as the probability that the AS-prefix isolation, once reached, is sustained further by launching at least one additional successful peering session attack on the AS-path in consideration before the RFD penalty goes below the reuse threshold. This probability of sustenance, P_{sus}, is given by

$$P_{sus} = 1 - (1-\theta)^{\lceil \frac{H(log_2 \frac{C}{R})}{T} \rceil}. \tag{6}$$

This equation essentially estimates the probability that at least one successful attack can be launched on the AS-path in consideration during the decay time from the cutoff threshold, C, to the reuse threshold, R.

4 Numerical Results from Analytical Model

In this section, we present some numerical results based on the analytical model of the preceding section. The probability that AS-prefix isolation happens in time t s or less is shown for the case of $n = 5$ hops in Fig. 8. In this plot, we also show the sensitivities to the vendor parameters as well as to the probability of success of a peering attack, Q. The performance (or vulnerability) is worse for Vendor B because the incremental penalty per successful attack (withdrawal plus re-advertisement) is much higher for Vendor B (2000) vs. Vendor A (1000). This effect is dominating even though Vendor B has a higher cutoff threshold than Vendor A (see Table 1). As would be expected, Fig. 8 also shows that higher probability of AS-prefix isolation occurs for higher values of Q.

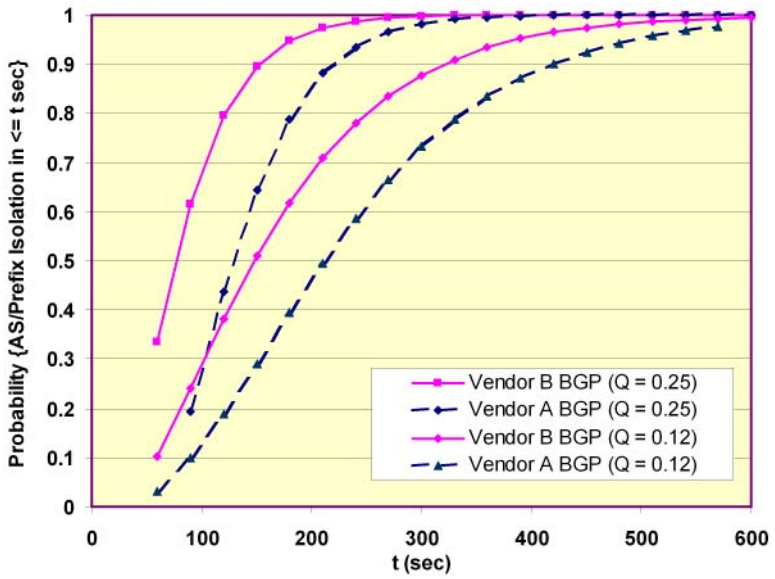

Figure 8: Probability of AS-prefix isolation: Sensitivity to vendor parameters and probability of successful session attack ($n = 4$).

With longer duration of attacks or larger area of vulnerability in the network, the attackers have a greater chance to be successful. This is illustrated in Fig. 9 where the probability of AS-prefix isolation goes higher as the number of hops, n, in the AS-path increases. It is generally known that the typical AS-path length in the Internet is about 4. The 3-D plot of Fig. 10 further illustrates how the probability of AS-prefix isolation increases with the number of hops as well as the time duration of attacks.

In Section 3, we also derived the probability of sustenance of AS-prefix isolation after such isolation has been achieved once by the attacker. Fig. 11 shows a plot of P_{sus} (derived in Eq. 6) as a function of Q. It can be

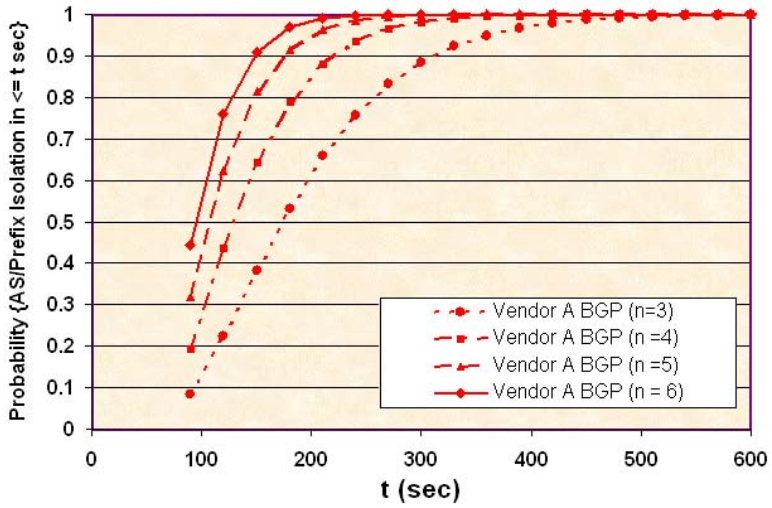

Figure 9: Probability of AS-prefix isolation: Sensitivity to vendor parameters and number of hops in AS path ($Q = 0.25$).

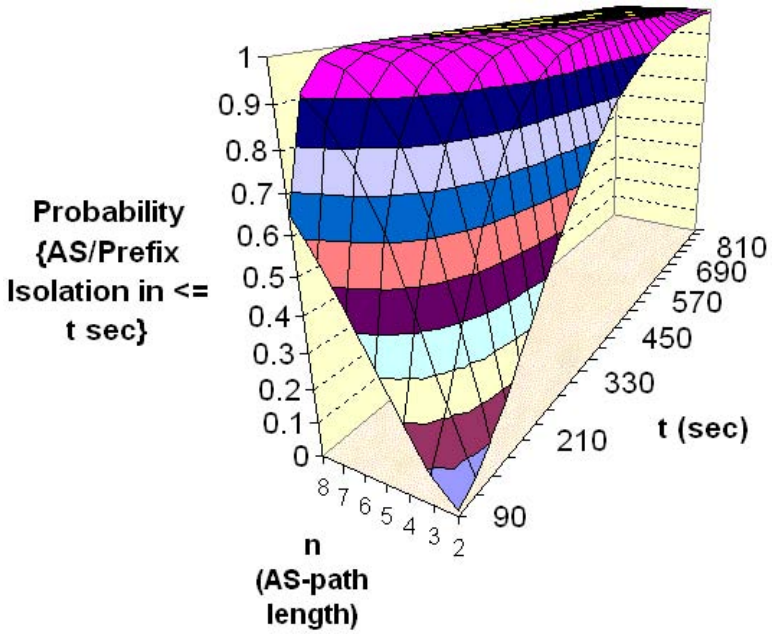

Figure 10: Probability of AS-prefix isolation as a 3D-function of time and AS-path length ($Q = 0.25$).

observed from Fig. 11 that it is much easier to sustain the AS-prefix isolation as compared to achieving the isolation initially. This is because the exponential decay time is very long as compared to BGP session recovery time. Any update that can be classified as a flap would increase the RFD penalty even though the RFD penalty in question is in a state of recovery (exponential decay). Thus, the malicious attacker has the luxury of time to slow down (or use fewer resources for the attacks) in order to sustain the isolation.

The analytical results presented here highlight the ease with which BGP peering session attacks can exploit

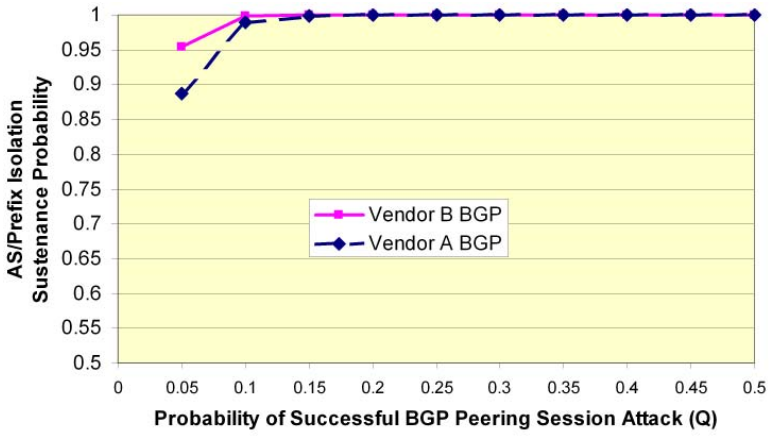

Figure 11: Probability of sustenance of AS-prefix isolation as a function of success of peering session attack (n = 4).

RFD and cause AS-prefix isolations, and sustain the same for prolonged periods. In Section 7, numerous complementary results based on simulations will be presented.

5 Benefit of BGP Graceful Restart

The BGP graceful restart (BGP-GR) mechanism [34] gives the downed router (downed only in the control plane) time to restart without peers withdrawing its routes. This option is negotiated between peers at the time of BGP peering session establishment. Two flags bits used in capability advertisement during BGP-GR negotiations are: 1) Restart bit - used to indicate that router has restarted; 2) Forwarding bit - used to tell a peer router that the capability exists to preserve forwarding state through a restart period. Once a router has announced its BGP-GR capability, during its restart (of BGP or BGP peering session) its neighbors do not immediately delete routes via that peer so that undue route flapping is prevented. A restart timer is used at each peer to determine how long it would wait before deleting stale neighbor routes. If the BGP open message is not received from the restarting router before the expiry of the restart timer, then the restart is presumed failed, routes previously announced by that peer are deleted, and withdrawals are sent.

Without BGP-GR, it is expected that the peering session attacks exploiting RFD behavior would be much more feasible. BGP-GR helps mitigate the effects of this type of attack. We have extended the analysis of Section 3 to model the impact of peering session attacks when BGP-GR is used. The analysis of Section 3 has been slightly modified to start a with a Poisson attack arrival model for the BGP peering session attacks. As previously stated, each attack would have a given probability of success. From these assumptions the probability of a successful session attack, Q, on a BGP peering session in an MRAI interval can be determined. Thus, Q would now be a function of the attack arrival rate. Now the equations of Section 3 can be readily used to derive performance metrics of interest as a function of the attack arrival rate.

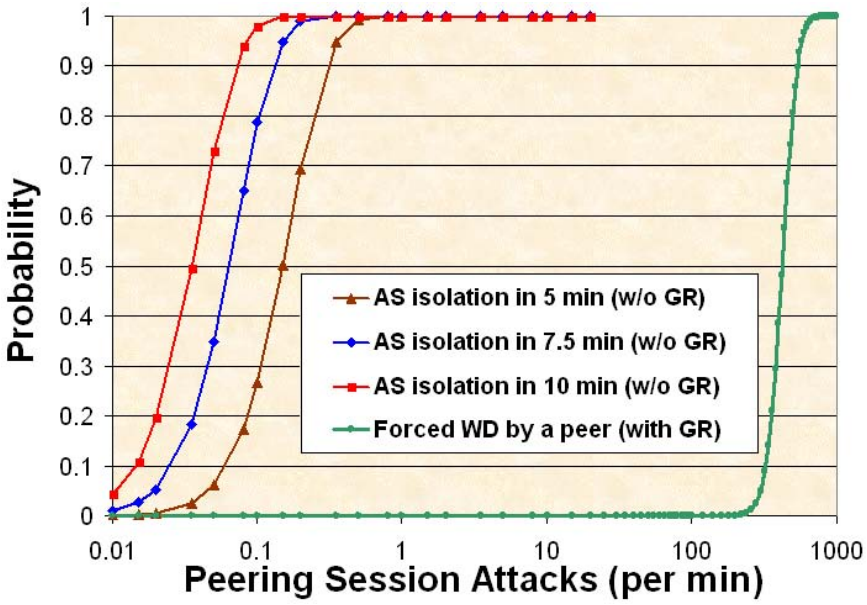

Figure 12: Probability of adverse impact on routing as a function of peering session attack rate for BGP with/without GR ($Q = 0.25$, $n = 4$).

To model the effect of BGP-GR, it is to be noted that BGP-GR allows a router's control plane (or BGP peering sessions) to be attacked many times over the duration of the restart timer without causing any disruptions in the forwarding plane. The attackers have to persistently attack within each BGP session recovery time over the duration of the restart timer in order to cause the peers to send withdrawals. We capture this in our analytical model of the attack effects with BGP-GR. The plots shown in Fig. 12 comparing the performance with and without BGP graceful restart (GR) are very instructive. We assume here that the BGP-GR restart timer is 120 s and the BGP session recovery time is about 4 s (much smaller in reality when TCP reset attack causes BGP session closure). The three plots to the left in Fig. 12 show the probability that AS-prefix isolation can be achieved in 5 min, 7.5 min, and 10 min, respectively, at comparatively low rates of peering session attacks when BGP-GR is not used. But at least two orders of magnitude higher rates of attacks are required in order to even cause forced withdrawals by peers when BGP-GR is used.

There seem to be some serious practical concerns about use of BGP-GR. Most newer BGP routers in the service provider networks have BGP-GR capability but it is very rarely (if at all) turned on. A BGP best practices document from the National Infrastructure Security Co-ordination Centre (NISCC, UK Govt.) notes that "Several providers (US) suggest that the cost of implementing this feature [BGP-GR] outweighs the benefits" (see note 5 on pg. 8 of [15]). In an informal survey of several commercial ISPs, we found that carriers' customers seem to prefer, in the case of multi-homing, that routing be done via a completely healthy BGP router (including control plane) rather than use BGP-GR over a route where the control plane is compromised. Their preference seems to be in that they wish to avoid a BGP router that is in recovery because it could have

stale routing information in its Forwarding Information Base (FIB).

6 Description of the Simulation Model

Our simulation environment is based on the Scalable Simulation Framework Network (SSFNet) [29], a discrete event simulator widely used for modeling of Internet protocols and networks. SSFNet provides a set of modules to simulate traffic at the IP layer and above. To support our work, we made several extensions and modifications to the TCP/IP and BGP modules. The default version of SSFNet does not come with a complete TCP state machine. Here code was added to produce proper failure messages and warnings as well as the detection of half open TCP connections. This extension allowed the simulation of spoofed TCP session resets, which in turn simulates BGP session attacks.

The existing implementation of the BGP protocol in SSFNet also had to be extended. The set of BGP modules that come with the SSFNet distribution did not include the "uncontrolled shutdown" of a BGP sessions. This scenario occurs if a successful hostile attack on the underlying TCP session results in a breakdown of the transport connection. Here the interaction between BGP and TCP as well as the proper shutdown and restart mechanism within BGP had to be implemented.

In addition to the extension of existing modules, we designed a BGP attack modeling framework. This framework allows the installation and configuration of individual "attacker" modules into each BGP router. All modules have some basic attributes in common, such as attack execution probability, module activation start time, and duration of attack activity. These parameters can be scripted for each AS or AS-group separately and/or for all ASes globally. For this study, we use a single type of attack module, namely, a TCP-Session Attacker. We are currently in the process of designing and running experiments with other types of attack modules, such as BGP message spoofing/tampering attacks and other protocol aware attacks (besides the one presented here). These are being documented elsewhere[31].

We considered both regular and realistic topologies for our study. Here the results are presented for the regular topologies, generally using nxn grids which offers many alternate paths between any two ASes. An 8x8 grid topology is shown in Fig. 13 which consists of 64 BGP nodes each representing an AS. In our experiments, the nodes represent ASes as well as destinations (i.e., prefixes). In our grid topologies, each AS contains one destination or prefix. We also consider a 16x16 grid topology that offers even more alternate AS-paths as compared to the 8x8 grid topology.

In the first of two simulation experiments performed in this study, we used the following network and attack parameters: (1) 8x8 grid topology (64 BGP nodes or routers); (2) Attacks can occur on any of the 112 BGP peering sessions in the entire network; (3) Attack duration of 240 s; (4) Attack duration is divided into 24 intervals of 10 s each, and there is the potential for one BGP session attack per peering session in each interval; (5) The probability of success of each session attack is assumed to be 100 %; (6) The timing of an attack is

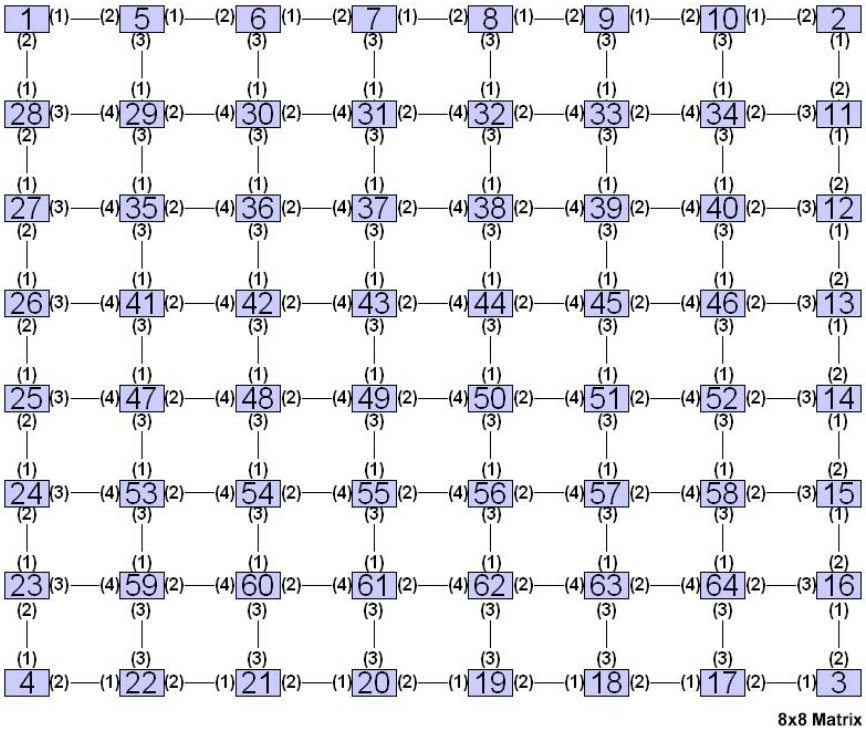

Figure 13: Grid topology of size 8x8.

uniformly random within each segment.

In the second simulation experiment performed in this study, we used the following network and attack parameters: (1) 16x16 grid topology (256 BGP nodes or routers); (2) Attacks can occur only on any of the 40 BGP peering sessions associated with the 4x4 central sub-grid portion of the network; (3) Attack duration of 10 s; (4) Attack duration is divided into 16 segments of 10/16 s each, and there is the potential for one BGP session attack per peering session in each such segment; (5) The probability of success of each session attack is assumed to be 25 %; (6) The timing of the attack is uniformly random within each segment. The second experiment contrasts from the first in that it has a much smaller success rate of attacks, and the attacks are topologically limited to the center $1/16^{th}$ of the network.

The following choices of parameters and protocol features are common to both simulation experiments: (1) Vendor B's RFD parameters are used; (2) MRAI is randomly chosen each time at each BGP router from a uniform distribution over the range 22.5 s to 30 s; (3) MRAI applies to both advertisements and withdrawals; (4) Random tie breaking is used in BGP route computation when two AS-paths are of equal cost.

7 Simulation Results and Discussion

We will now present and discuss simulation results for the two topologies previously described. We first present results for the 8x8 grid and later present results for the 16x16 grid topology. In both sets of results, we provide

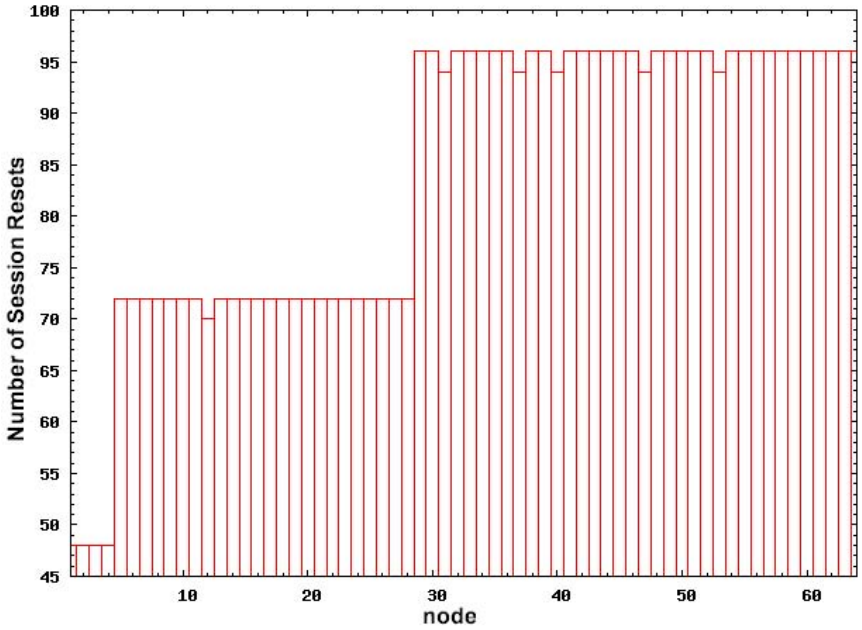

Figure 14: Measured number of BGP session attacks vs. the node ID in the 8x8 grid network.

insights into the impacts due to the attacks in terms of several performance metrics related to route stability and degradation of route quality. In the results that follow, the metrics will be compared for the cases of (a) without RFD and (b) with RFD. The purpose of this comparison is to show how the attacks, when tuned to protocol specifics (RFD, MRAI), can cause much more amplified damage to the routing infrastructure than otherwise.

We allow the network a start-up and initial stabilization time of 500 s, by which time all route have converged to their respective stable routes. The attacks are launched starting at 500 s, and go on for 240 s for the first experiment and for 10 s for the second experiment. It may be noted that our observation interval is on the order of 1000's of seconds while the attack duration is only on the order of 10's of seconds. This is done so that the ripple effects of the attacks are captured well, especially in the case of BGP with use of RFD. When RFD penalty is triggered, the many route restorations to respective stable paths happen after 1000's of seconds following the attacks. Note that the halftime of the exponential decay of RFD penalty is 900 s. Hence the observation time must be at least a few 1000's of seconds.

In Fig. 14, the number of session resets are shown for the first experiment as a function of the node ID. The node ID's are as shown in Fig. 13. The BGP sessions at all nodes are attacked with the same probability (100 % success in this experiment) but the corner and edge nodes show fewer attacks because they do not have as many peers as the interior nodes. The purpose in this experiment is to simulate repeated session attacks throughout the network, and possibly trigger RFD cutoffs and isolations in a predominant manner.

For the purpose of this study, the unreachability time between a node and a prefix is defined as the total time (summed over possibly multiple occurrences) that the prefix is unreachable from the node during an experiment's

observation interval. In Fig. 15, the unreachability times are plotted in 3D as a function of the node and prefix ID's, and a comparison is shown between two cases: (a) without use of RFD and (b) with use of RFD. Because of the intense nature of the attacks (spaced at 10 s time intervals and throughout the network), almost any prefix is rendered unreachable from almost any node for some duration of time during the observation period. What is most striking in Fig. 15 is the fact that the node-prefix unreachability time in the case with RFD is about 10 times larger than that in the case without RFD (approximately 3000 s and 300 s for the cases with RFD and without RFD, respectively). While RFD serves its purpose in terms of damping undesirable route flaps during normal operation of the Internet, it could however, aid the malicious attackers in terms amplifying the impact of focused attacks.

This observation is further reinforced when we look at another important metric in Fig. 16. The count of (i,j)-pairs unreachable is a very useful routing performance metric since it tells us the total number of all such pairs where prefix j is unreachable from node i at a given time. The span of the x-axis (the time span) before the count of (i,j)-pairs unreachable goes down to zero is important to note in interpreting Fig. 16. This time span is about 820 s in the case without RFD while the same is about 3300 s in the case with use of RFD. The attacks last for 240 s. Without RFD it takes 80 s (820 s - 500 s - 240 s) after that for all the routes to converge back to reachable routes (mostly stable routes) while it takes 2560 s (3300 s - 500 s - 240 s) for the same to happen when RFD is in use. There are a total of 64x64 - 64 = 4032 (i,j)-pairs in the network in consideration. In the case of without RFD, a little over 3000 (i,j)-pairs are rendered unreachable by the attacks, whereas in the case of RFD, all 4032 (i,j)-pairs are rendered unreachable lasting in that state for about 1800 s.

The reason for the three plateaus in recovery of routes in Fig. 16 can be explained as follows. Each plateau of

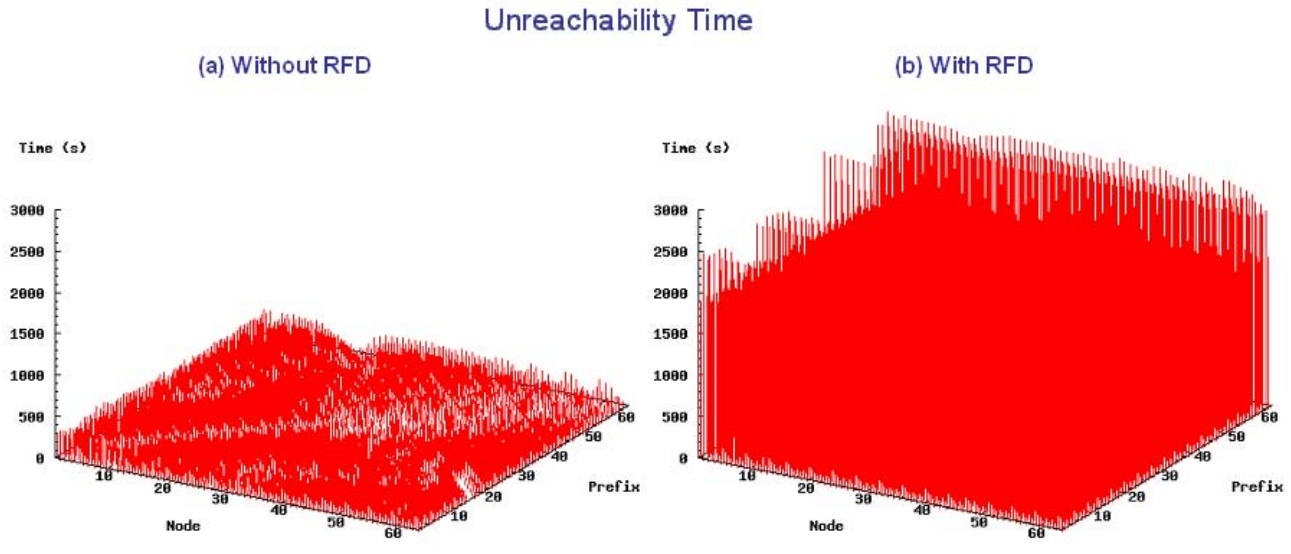

Figure 15: Unreachability time from BGP node i to prefix j: Comparison of (a) Without RFD versus (b) With RFD highlights amplification attributable to RFD.

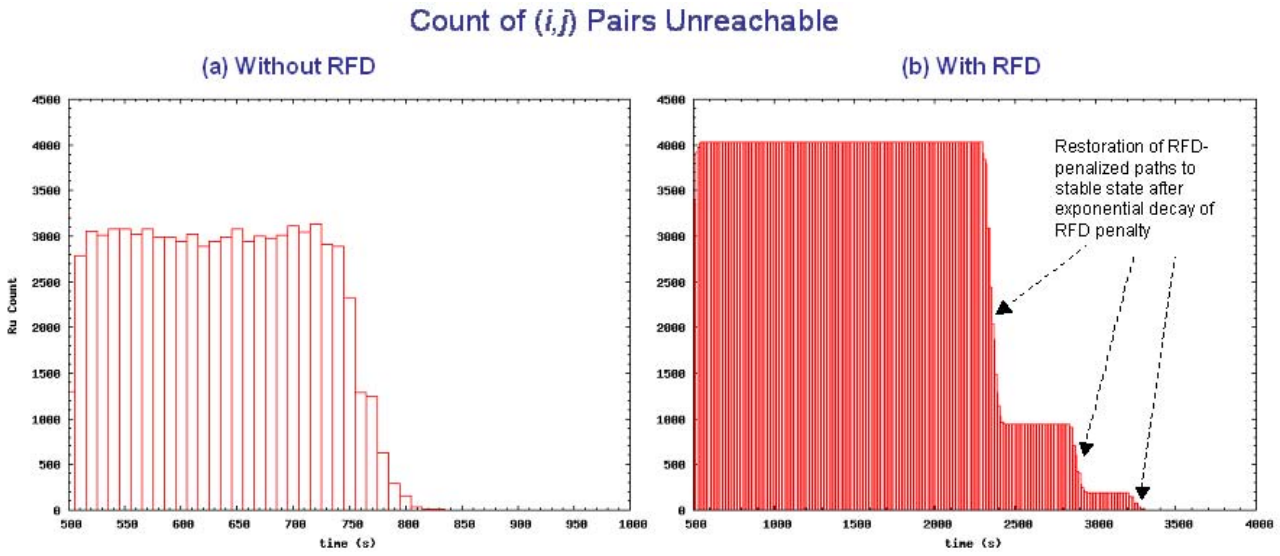

Figure 16: Count of node-prefix pairs unreachable as a function of time: (a) vs. (b) comparison shows that widespread unreachability lasts about 4 times longer in (b) that is attributable to RFD.

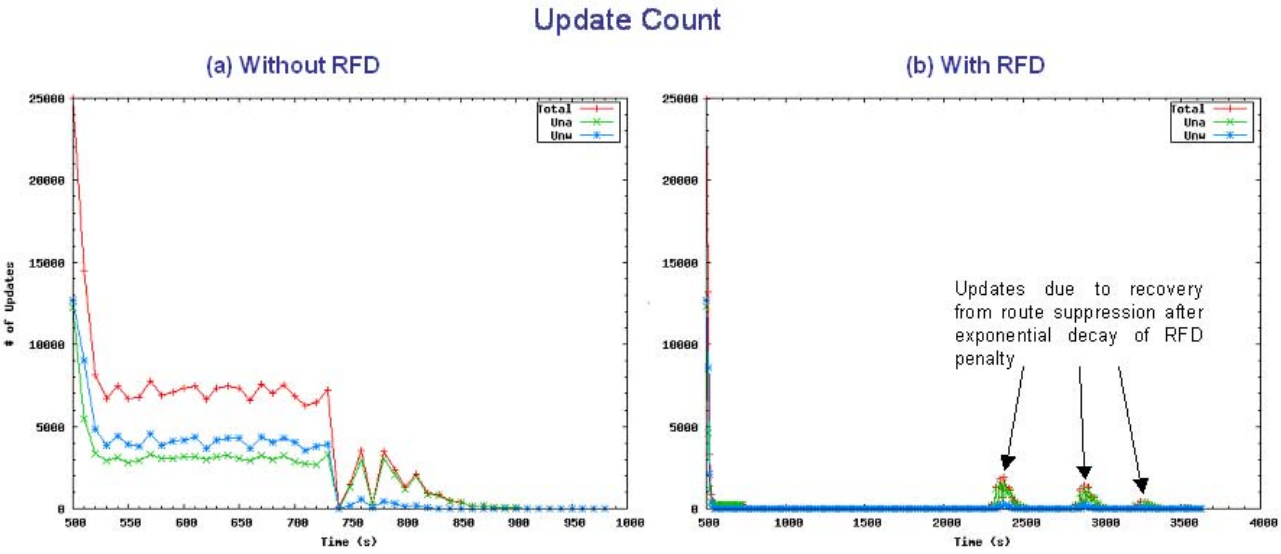

Figure 17: BGP update count as a function of time: (a) vs. (b) comparison shows the suppression followed by flurry of updates after RFD penalty decays.

recoveries corresponds to the level of the accumulated RFD penalty for each route at the onset of RFD cutoff. Of the three sets, the two dominant batches of recoveries can be explained based on RFD decay arguments. Some routes exceed the cutoff threshold just barely in which case they take about 1800 s (given a halftime value of 900 s) to decay from RFD cutoff threshold of 3000 to RFD reuse threshold of 750. Another set of routes exceed the cutoff threshold by almost the amount of one withdrawal/re-advertisement penalty of 1000 in which case they take about 2170 s to decay from an accumulated RFD penalty of about 4000 to RFD reuse

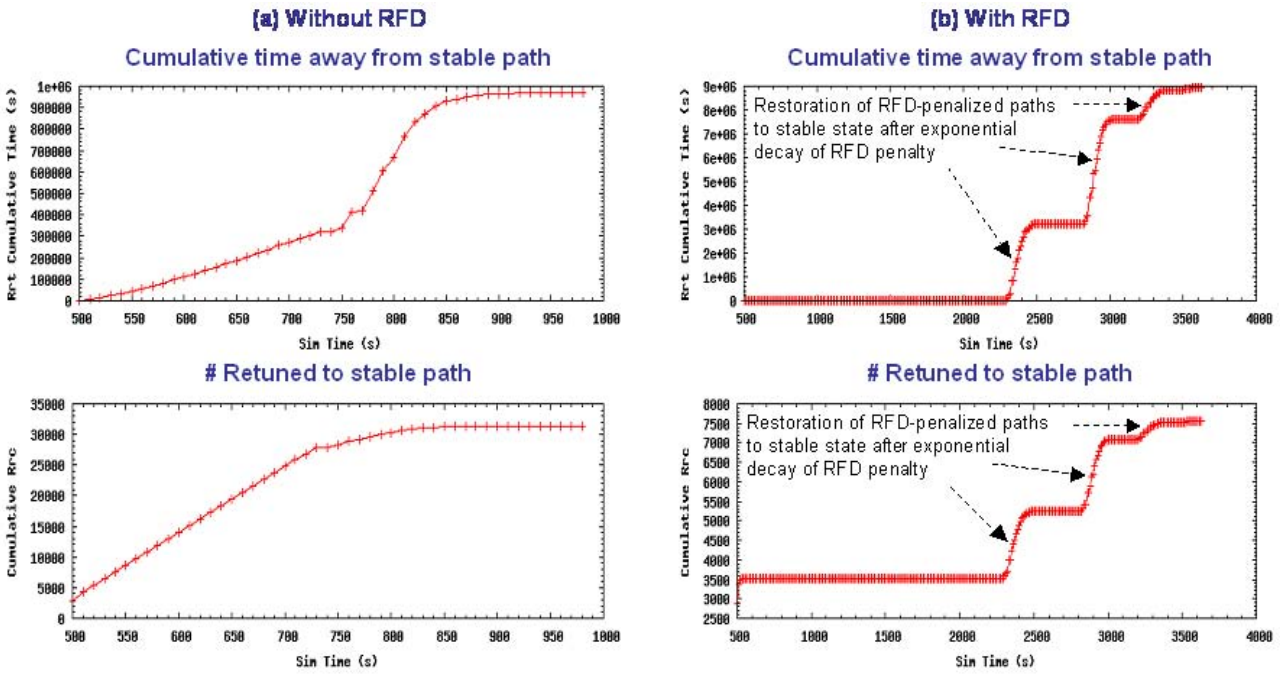

Figure 18: Cumulative route deviation time (away from stable path) and count of route returns to stable path: (a) vs. (b) comparison shows significantly larger time away from stable path although significantly smaller number of deviations (damping effect) when RFD is used.

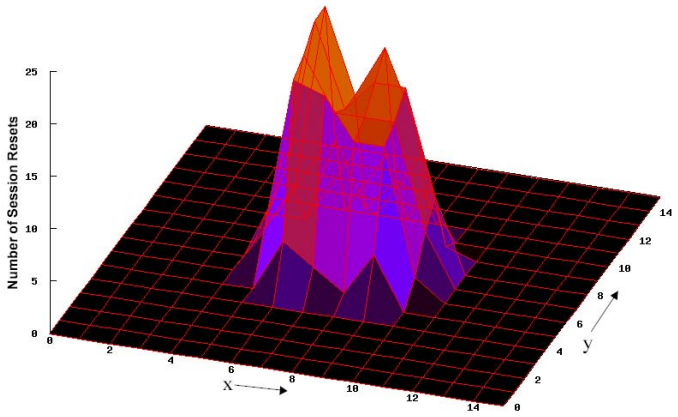

Figure 19: Measured number of BGP session attacks plotted over the topology of the 16x16 grid network; attacks focused over 4x4 sub-grid.

threshold of 750. These estimates closely agree with the timings of the first two (dominant) batches, which are approximately at 1800 s and 2200 s after the onset of attacks at 500 s. The third set seems to be due to additional flapping caused for some routes when the first set is recovering and thus sending updates around the network.

Fig. 17 shows the BGP update count vs. time for cases (a) without RFD and (b) with RFD. The update

count is the number of updates seen network-wide collected in ten-second intervals. The update counts are classified as the number of advertisements Una and withdrawals Unw, respectively. It can be observed that the update activity in Fig. 17 temporally correlates with the rise and fall of number of (i,j)-pairs unreachable in Fig. 16. When session attacks are launched throughout the network over an initial 10 s interval (500-510 s), a sharp spike follows in the number of updates and also in the number of (i,j)-pairs unreachable. The three flurries of updates after the initial big flurry in Fig. 17b can be again explained with the RFD exponential decay arguments (as was done explaining Fig. 16). The updates subside in about 350 s from the onset of attacks in the case of without RFD versus about 3000 s in the case with RFD. It may also be noted in Fig. 17 that during the first 350 s (500s to 850 s), there are many more updates in the case without RFD as compared to the same with RFD. This is understandable because RFD in meant to suppress updates in the event of route flapping. Looking at the details in the first 350 s in the left-hand part (i.e., (a) without RFD), it can be noted that the periodic mini-spikes in the update count happens at intervals of about 20 s to 30 s. This corresponds to the MRAI, which randomly varies from 22.5 s to 30 s. MRAI does cause bunching to occur in the propagation of updates in the network. These MRAI influenced mini-spikes are seen for both advertisements as well as withdrawals.

Two other routing performance metrics of interest are related to route deviations from the stable routes and can be regarded as route quality metrics. These are (1) the cumulative time away from the stable paths over all routes network-wide, and (2) the cumulative number of routes that have returned to their respective stable paths network-wide. In Fig. 18, these metrics are plotted versus time, and compared for the cases of (a) without RFD and (b) with RFD. At the end of the observation interval, the cumulative time away from the stable paths is amplified by a factor of almost 10 for the case with RFD as compared to that for the case without RFD. Looking at the two lower plots in Fig. 18, a little over 30000 deviations from stable paths are observed cumulatively in the case of without RFD while the same is only about 7500 in the case with RFD. In the case with RFD, route deviations are not observed as much because most or all of the routes have already been suppressed early during the attacks and hence are not subject to deviations. However, the suppressed routes have to wait through the exponential RFD decay period before they are restored to their stable paths. This is evident if attention is paid to the time axis in the plots in Fig. 18.

Now we will discuss simulation results based on the second experiment. In the first experiment, the distinction between the case with RFD and the case without RFD was very significant in terms of the unreachability time metric. As we will soon see, the second experiment shows that under some scenarios of session attacks, the cases with RFD and without RFD may not differ much in terms of the unreachability time, but can still have a very huge performance difference in terms of other important metrics.

Fig. 19 shows the total number of session resets per node during the experiment plotted over the topology. The attacks are focused only over a $1/16^{th}$ of the network topology and contained in the center 4x4 sub-grid. The 16x16 grid has many alternative paths between a node and a prefix. In this experiment, there are no policy restrictions on usable paths. Hence, reachability between a node and a prefix is almost assured even

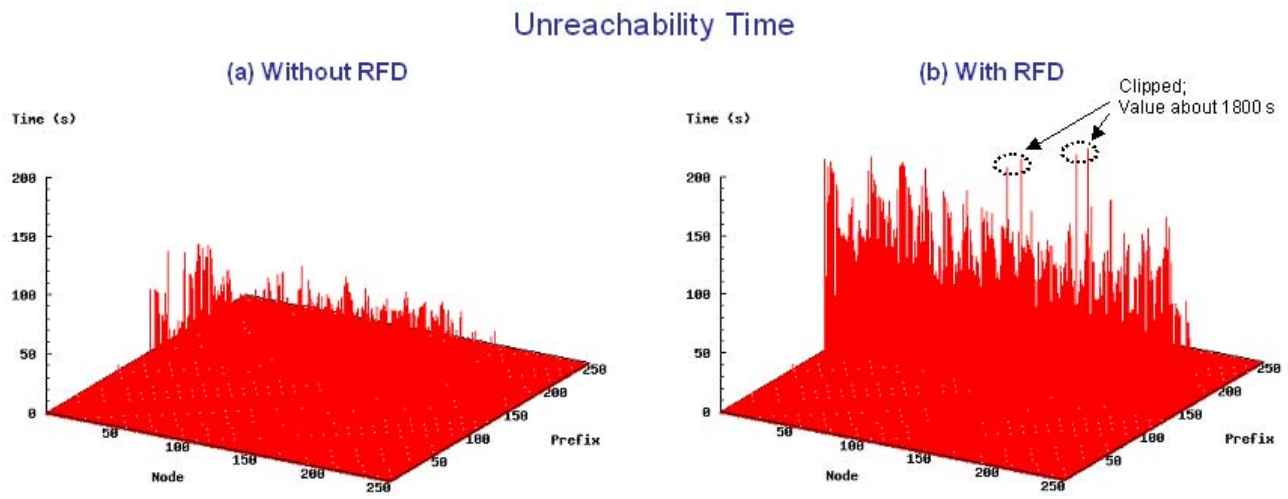

Figure 20: Unreachability time from BGP node i to prefix j: (a) vs. (b) comparison highlights amplification attributable to RFD; Topology of 16x16 grid with 4x4 sub-grid attacks.

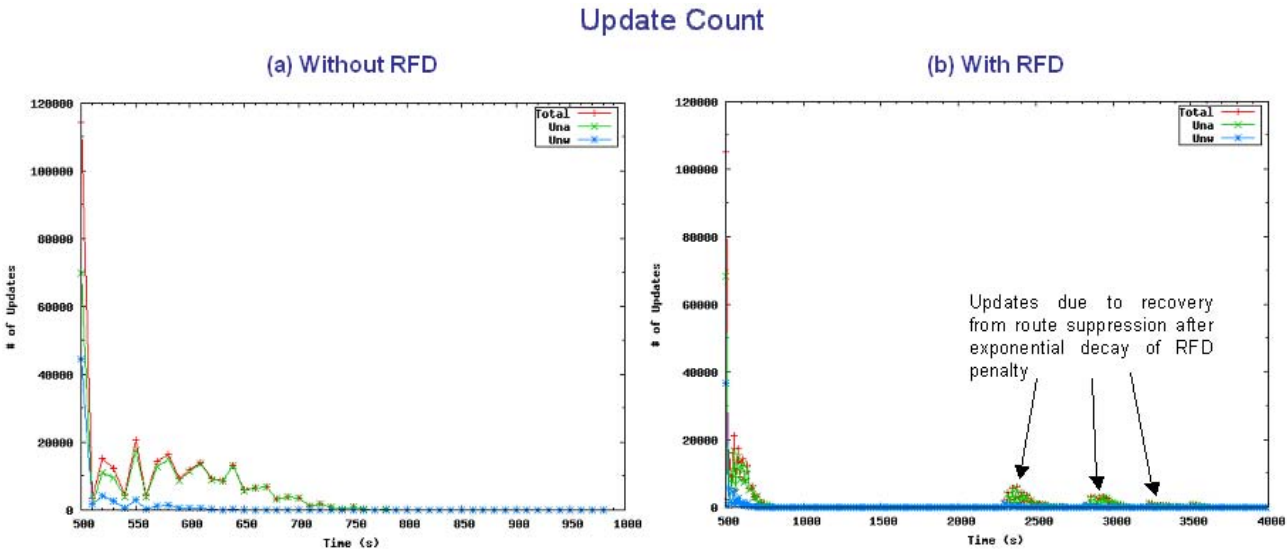

Figure 21: Case of 16x16 grid with center 4x4 sub-grid attacks: BGP update count as a function of time: (a) vs. (b) comparison shows RFD suppression and flurry of updates after RFD penalty decays.

if several of the BGP sessions in the central 4x4 sub-grid are reset at a given time. Because of this, the unreachability time is similar between the cases of without RFD and with RFD as shown in Fig. 20. A few RFD route suppressions occur, but otherwise the unreachability times for the majority of node and prefix pairs are comparable. However, the use of RFD in BGP still shows its adverse effects (or beneficial effects from the malicious attackers' perspective) in terms of other metrics such as the cumulative time away from stable paths and the cumulative count of route returns to stable path.

Fig. 21 shows the number of updates over time in 10 s time granularity. These plots comparing with and

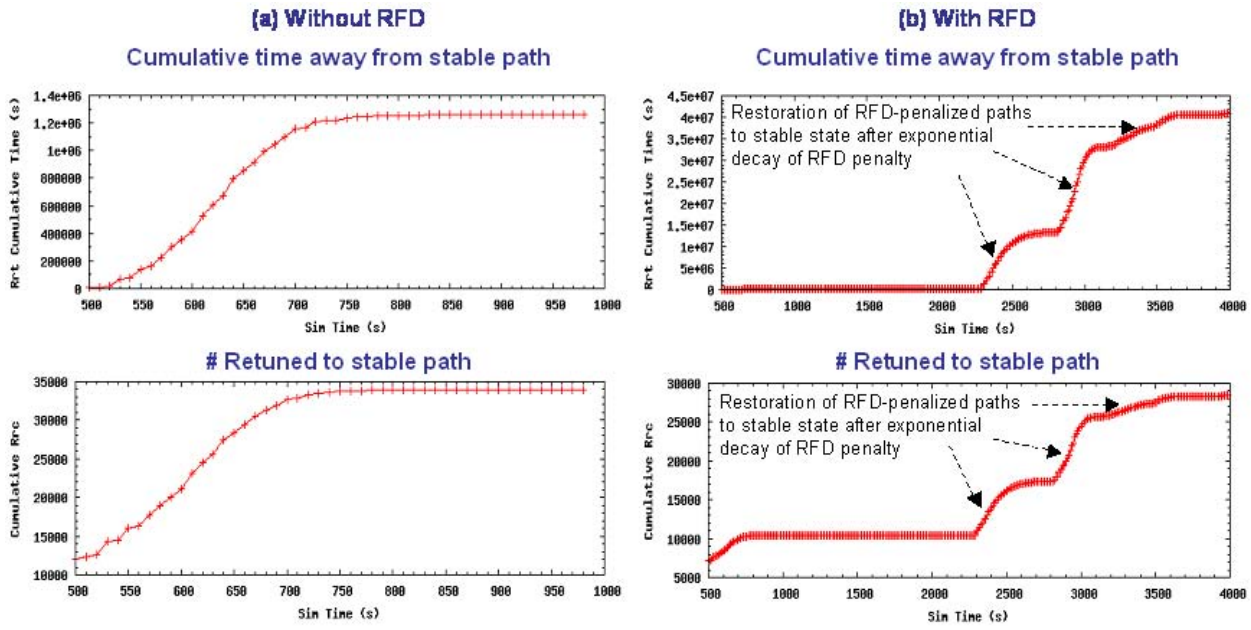

Figure 22: Cumulative route deviation time (away from stable path) and count of route returns to stable path: (a) vs. (b) comparison shows significantly larger time away from stable path although significantly smaller number of deviations (damping effect) when RFD is used; Topology of 16x16 grid with 4x4 sub-grid attacks.

without RFD are very similar to the corresponding plots from the first experiment (see Fig. 17). The second experiment has many more updates at the onset as compared to the first experiment (about 110K vs. 25K) because of the much larger network size in the second experiment. Fig. 22 highlights the difference between the cases without RFD and with RFD by focusing attention on route quality metrics, namely, the cumulative time away from stable paths and the cumulative count of route returns to stable paths. The cumulative time away from stable paths is about 30 times larger for the case with RFD (4.2×10^7) as compared to the case without RFD (1.3×10^6). The lower plots in Fig. 22 further illustrate that all routes converge back to their respective stable paths in about 250 s after the onset of attacks for the case without RFD, but requires about 3200 s in case with RFD.

Thus, the second experiment illustrates that the impact of RFD may not be too evident in the route unreachability metric (under somewhat low profile attack scenarios) but may still be overwhelming in terms of its adverse effects evidenced by prolonged compromise of route quality for a majority of routes in the network.

8 Conclusions

We have shown that routing attacks can be tuned to take advantage of the BGP protocol behavior, and thus significantly amplify the adverse impact of those attacks on the routing infrastructure. RFD was designed to alleviate BGP processing overload and route flapping under non-malicious scenarios. However, under certain

BGP peering session attack scenarios, the RFD facilitates further vulnerability in BGP by allowing severe amplification of unreachability as well as degradation of route quality. We have presented a detailed analytical study of the impact of BGP peering session attacks that exploit the RFD. Our results have revealed that it is possible for the attackers to achieve a high probability of AS-AS and AS-prefix isolation by attacks conducted at a rate roughly equal to once per Minimum Route Advertisement Interval (MRAI), even with a low success rate per BGP session attack. We have also shown that the RFD-based BGP vulnerability can be partially mitigated by using the BGP Graceful Restart (BGP-GR) mechanism.

We have further studied the impact of BGP peering session attacks with RFD exploitation through detailed network simulations. The simulation results confirm that much prolonged isolations between ASes and prefixes are the result of these protocol aware attacks. As part of further research in this area, it would be useful to study whether tuning the RFD parameters can constitute a countermeasure to the potential vulnerability studied here.

The study reported here is a part of a larger effort at NIST to identify and characterize the risks associated with focused attacks of different types on the BGP infrastructure, and to evaluate the effectiveness and impact of various proposed mitigation techniques [16][31]. We expect to be in a position soon to share many detailed simulation results and data for various BGP vulnerability experiments with the R&D community, and are in the process of setting up external web pages to facilitate the same.

References

[1] O. Nordstrom and C. Dovrolis, "Beware of BGP attacks," SIGCOMM Computer Communications Review (2004), Vol. 34 (2), pp. 1-8.

[2] K. Butler, T. Farley, P. McDaniel, and J. Rexford, "A Survey of BGP Security", Technical Report TD-5UGJ33, ATT Labs - Research, Florham Park, NJ, February 2004, (Revised April 2005). http://www.patrickmcdaniel.org/pubs/td-5ugj33.pdf

[3] G. Goth, "Fixing BGP Might Be Difficult–Or Not So Tough," IEEE Internet Computing, vol. 07, no. 3, pp. 7-9, May/June 2003.

[4] Z.M. Mao, R. Govindan, G. Varghese, and R.H. Katz, "Route Flap Damping Exacerbates Internet Routing Convergence," Proceedings of ACM SIGCOMM, Pittsburg, PA, August 2002, pp. 221-233.

[5] Kim, J., S.Y. Ko, D.M. Nicol, X.A. Dimitropoulos, G.F. Riley, "A BGP attack against traffic engineering," Proceedings of the 2004 Winter Simulation Conference, 2004.

[6] S.M. Bellovin and E.R Gansner, "Using Link Cuts to Attack Internet Routing", AT&T Labs Research Technical Report, http://www.research.att.com/ smb/papers/reroute.pdf

[7] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "Observation and analysis of BGP behavior under stress," Proceedings of the Second ACM SIGCOMM Workshop on Internet Measurement Workshop, 2002, Marseille, France, pp. 183-195.

[8] K. Zhang, X. Zhao, and S. F. Wu, "An analysis on selective dropping attack in BGP," Proceedings of IEEE International Performance Computing, and Communications Conference (IPCCC), April 2004, pp. 593-599.

[9] E.G. Coffman Jr., Z. Ge, V. Misra and D. Towsley, "Network Resilience: Exploring Cascading Failures within BGP", Proceedings of the 40th Annual Allerton Conference on Communications, Computing and Control, Monticello, Illinois, October 2002.

[10] K. Varadhan, R. Govindan, and D. Estrin, "Persistent Route Oscillations in Inter-Domain Routing," Computer Networks, vol. 32, no. 1, pp. 1–16, January 2000.

[11] F. Gont, "ICMP Attacks against TCP," IETF Internet Draft, draft-gont-tcpm-icmp-attacks-03.txt, December 2004.

[12] "Flaw Could Cripple Entire Net," Associated Press, April 20, 2004. http://wired-vig.wired.com/news/technology/0,1282,63143,00.html

[13] "NISCC Vulnerability Advisory 236929: Vulnerability Issues in TCP," April 20, 2004.

[14] "CERT Advisory CA-2001-09: Statistical Weaknesses in TCP/IP Initial Sequence Numbers," http://www.cert.org/advisories/CA-2001-09.html, Original date May 2001, last revised February 2005.

[15] NISCC (UK Govt.) Best Practices Guidelines: Border Gateway Protocol, April 2004.

[16] D. Montgomery, K. Sriram, O. Borchert, O. Kim, and R. Kuhn, "Characterizing the Risks and Costs of BGP Insecurity/Security," Presented at the First DHS Workshop on Secure Protocols for the Routing Infrastructure (DHS-SPRI), Washington D.C., March 15-16, 2005 (presentation slides available from authors upon request).

[17] S. Convery, D. Cook, and M. Franz, "An Attack Tree for the Border Gateway Protocol", IETF ID, http://ietfreport.isoc.org/ids/draft-ietf-rpsec-bgpattack-00.txt, February 2004.

[18] Y. Rekhter and T. Li, "A border gateway protocol 4 (BGP-4)", IETF RFC 1771, March 1995.

[19] I. van Beijnum, BGP: Building Reliable Networks with the Border Gateway Protocol, O'Reilly (2002).

[20] S. Kent, C. Lynn, and K. Seo, "Secure Border Gateway Protocol (S-BGP)," IEEEJSAC Special Issue on Network Security, April 2000.

[21] J. Ng, "Extensions to BGP to Support Secure Origin BGP (soBGP)," IETF ID draft-ng-sobgp-bgp-extensions-02.txt, April 2004.

[22] L. Subramanian et al., "Listen and whisper: Security mechanisms for BGP," In First Symposium on Networked Systems Design and Implementation (NSDI'04), 2004.

[23] A. Heffernan, "Protection of BGP Sessions via the TCP MD5 Signature Option," IETF RFC 2385, August 1998.

[24] D. Pei, M. Azuma, D. Massey, L. Zhang, "BGP-RCN: Improving Convergence through Root Cause Notification", Computer Networks, Volume 48, No. 1, May 2005, pp. 175-194.

[25] J. Rexford, A. Greenberg, G. Hjalmtysson, D.A. Maltz, A. Myers, G. Xie, J. Zhan, and H. Zhang, "Network-wide decision making: Toward a wafer-thin control plane," Proc. ACM SIGCOMM HotNets Workshop, November 2004.

[26] J. Cowie, A. Ogielski, B. Premore, and Y. Yuan, "Global Routing Instabilities during Code Red II and Nimda Worm Propagation," http://www.renesys.com/projects/bgp_instability, September 2001.

[27] T. Griffin, "BGP Impact of SQL Worm, 1-25-2003", http://www.research.att.com/ griffin/bgp_monitor/sql_worm.html, January 2003.

[28] I. Dubrawsky, "Effects of Worms on Internet Routing Stability", http://www.securityfocus.com/infocus/1702, June 2003.

[29] SSFNet Gallery of Baseline Networks, http://www.ssfnet.org/Exchange/gallery/index.html

[30] B.J. Premore, "An Analysis of Convergence Properties of the Border Gateway Protocol Using Discrete Event Simulation," Ph.D. thesis, Dartmouth College Department of Computer Science, Technical Report TR2003-452, May 2003.

[31] K. Sriram, D. Montgomery, O. Borchert, O. Kim, and R. Kuhn, "Study of BGP Behavior under Large Scale Attacks," NIST Technical Report (in preparation).

[32] C. Villamizar, R. Chandra, and R. Govindan, "BGP route flap damping," IETF RFC 2439, November 1998.

[33] W. Shen and Lj. Trajkovic "BGP route flap damping algorithms," Proc. SPECTS 2005, Philadelphia, PA, July 2005, pp. 488-495.

[34] S.R. Sangli, Y. Rekhter, R. Fernando, J.G. Scudder, and E. Chen, "Graceful Restart Mechanism for BGP," IETF ID, draft-ietf-idr-restart-10.txt, December 2004.