Impact of BGP Dynamics on Router CPU Utilization

Sharad Agarwal[1], Chen-Nee Chuah[2], Supratik Bhattacharyya[3], and Christophe Diot[4]

[1] University of California, Berkeley, USA, sagarwal@cs.berkeley.edu
[2] University of California, Davis, USA, chuah@ece.ucdavis.edu
[3] Sprint ATL, Burlingame, USA, supratik@sprintlabs.com
[4] Intel Research, Cambridge, UK, christophe.diot@intel.com

1 Introduction

The Internet is an interconnection of separately administered networks called Autonomous Systems or ASes. To reach entities outside the AS, the inter-domain routing protocol used today is the Border Gateway Protocol or BGP [1]. It has been approximately 15 years since BGP was deployed on the Internet. The number of ASes participating in BGP has grown to over $16,000$ today. However, this growth has been super-linear during the past few years [2]. With this sudden growth there has been concern in the research community about how well BGP is scaling. In particular, it has been noted that there is significant growth in the volume of BGP route announcements (or route flapping) [3] and in the number of BGP route entries in the routers of various ASes [2].

For every BGP routing update that is received by a router, several tasks need to be performed [4]. First, the appropriate RIB-in (routing information base) needs to be updated. Ingress filtering, as defined in the router's configuration, has to be applied to the route announcement. If it is not filtered out, the route undergoes the BGP route selection rules and it is compared against other routes. If it is selected, then it is added to the BGP routing table and the appropriate forwarding table entry is updated. Egress filtering then needs to be applied for every BGP peer (except the one that sent the original announcement). New BGP announcements need to be generated and then added to the appropriate RIB-out queues.

These actions can increase the load on the router CPU. Long periods of high router CPU utilization are undesirable due to two main reasons. High utilization can potentially increase the amount of time a router spends processing a routing change, thereby increasing route convergence time. High route convergence times can cause packet loss by increasing the window of time during which the route for a particular destination is unavailable. Further, high router CPU utilization can disrupt other tasks, such as other protocol processing, keep alive message processing and in extreme cases, can cause the router to crash.

In this work, we answer the question "Do BGP routing table changes cause an increase in average router CPU utilization in the Sprint IP network?". Sprint operates a "tier-1" IP network that connects to over $2,000$ other ASes. Thus, we believe that it is a suitable point for studying the impact of BGP on average router CPU utilization. There has been prior work [5, 6] in analyzing BGP protocol behavior during worm propagation. To the best of our knowledge, there has been no prior published work on analyzing the relationship between BGP protocol behavior and router CPU utilization.

We examine BGP data from multiple routers in the network. We correlate this with SNMP data on CPU utilization for about 200 routers on different days inside Sprint. The findings of our work are:

- On average, BGP processes consume the majority of router CPU cycles. For short periods of time, we observe very high router CPU utilization due to BGP processes.
- We find that during normal network operation, there is some correlation between short increases in the number of BGP routing table changes and increases in average router CPU utilization and vice versa, at the time granularity of 5 minutes. However, over all the instances we observed, the impact was negligible.
- We find that during the SQL Slammer worm attack in January 2003, there was a tremendous increase in the amount of BGP changes that lasted for several hours and there was a correlation with average router CPU utilization. However, we find that the increase in utilization during this time was under 20% for almost all routers.

There are some limitations of our work:

- Since the Sprint IP network consists purely of Cisco routers, we do not measure how other router architectures react to BGP protocol behavior.
- We do not benchmark router CPU performance in a laboratory testbed. Our emphasis is on actual performance in an operational environment.
- We do not attempt to create a detailed model of how the router CPU should react to different kinds and volumes of BGP messages. Such a model will be heavily influenced by the operating system source code, which we do not have access to.

2 Analysis Data

To understand how BGP routing table changes impact average router CPU utilization, we need to analyze three kinds of data from the operational network. We now describe the routers that we access, the interactive session data, the SNMP (Simple Network Management Protocol) data and the BGP data that we analyze.

2.1 Routers

The Sprint network (AS 1239) consists of over 600 routers, all of which are Cisco routers. We had access to data from 196 routers, the majority of which are Cisco GSR 12000 series and Cisco 7500 series with VIP interfaces. They all have either about 256 MB or 512 MB of processor memory. The route processor on each of these routers is a 200 Mhz MIPS $R5000$ family processor. The BGP routing protocol runs as part of the operating system. On older routers such as the Cisco 1600, 2500, 4500, 4700, 7200, and 7500, there is a single processor that runs both the operating system (IOS) and packet switching [7]. On newer architectures, such as the Cisco $7500{-}VIP$ and $GSR{-}12000$, the main processor runs IOS and interface card processors perform distributed packet switching. On these routers, the route processor runs all the routing protocols, builds the forwarding, CEF (Cisco Express Forwarding) and adjacency tables and distributes them to the line cards.

There are typically four BGP processes in Cisco IOS [5]. The "BGP Open" process handles opening BGP sessions with other routers. It runs rarely. The "BGP Scanner" process checks the reachability of every route in the BGP table and performs route dampening. It will run once a minute and the size of the routing table will determine how long it takes to complete. The "BGP Router" process receives and sends announcements and calculates the best BGP path. It runs every second. The "BGP I/O" process handles the processing and queueing involved in receiving and sending BGP messages. The frequency of execution of this process will be related to the frequency of BGP updates.

2.2 Interactive Session Data

All the routers in our study allow command line interface (CLI) access via secure shell (SSH). Upon logging into the router, we can issue commands to query the state of the router. We issued the "show process cpu" command to all routers during the study. This command lists all the processes in IOS, along with the CPU utilization of each process [7]. Both the percentage utilization of the CPU (over the last 5 seconds, 1 minute and 5 minutes) and the total number of CPU milliseconds consumed since the last boot of IOS is reported. This data provides an instantaneous snapshot of what processes are currently loading the CPU and which processes have consumed the most CPU resources since boot up. A sample output is below.

```
CPU utilization for five seconds: 99%/2%; one minute: 18%; five minutes: 15%
 PID Runtime(ms)   Invoked      uSecs   5Sec   1Min   5Min TTY Process
   1        3756   1242536          3  0.00%  0.00%  0.00%   0 Load Meter
   2         192        35       5485  0.00%  0.02%  0.03%   2 SSH Process
...
 142    22517688  86537321        260  0.40%  0.39%  0.44%   0 BGP Router
 143    29989784  29380359       1020  0.24%  0.69%  0.70%   0 BGP I/O
 144   720698256   4712209     152943 95.48% 14.07% 11.12%   0 BGP Scanner
...
```

2.3 SNMP Data

Due to the limitation of how frequently we can collect CLI data, we also collect SNMP data. The SNMP protocol allows for a data collection machine to query certain SNMP counters on these routers and store the values in a database. We query the 1 minute exponentially-decayed moving average of the CPU busy percentage as defined in SNMP MIB 1.3.6.1.4.1.9.2.1.57. We query and store this value once every 5 minutes from each one of the 196 routers that we have access to. The network operators cited concerns about affecting the router CPU load if we were to poll this counter more frequently. This data provides us with the CPU utilization at a relatively large time scale granularity and hence does not identify the process responsible for high load. We have collected this data for as long as 2.5 years for some routers.

2.4 BGP Routing Data

In order to know if high CPU utilization is caused by a large number of BGP messages, we also analyze BGP data. We collect iBGP data from over 150 routers in the Sprint

[5] http://www.cisco.com/warp/public/459/highcpu-bgp.html

network, all of which we also collect SNMP data from. We collect BGP data using the GNU Zebra [6] routing software. We connect to about 150 routers which comprise the iBGP route reflector mesh inside the Sprint network. Each PoP announces BGP routes that originate from it to the data collector. We also connect as route reflector clients to two routers in two different PoPs.

3 Results

3.1 Short Time Scale Behavior

We first address the impact of the BGP routing protocol on router CPU utilization in short time scales. We analyze interactive session data here using the "show process cpu" command on routers in the Sprint network. This command was executed during normal network operation when no significant outage or abnormal behavior occurred. Across all 196 routers that we have access to, we see one of two cases when we execute the command. The common case is when the CPU is lightly loaded and no process is consuming a significant percentage of CPU load. In other cases, either the "BGP Scanner" or the "BGP Router" process consumes a significant percentage (sometimes over 95%) of CPU load in the 5 second average, but not in the longer term averages. This indicates that for very short time periods, BGP processes can contribute to high load on a router CPU. Aggregate statistics on how often this short time scale behavior manifests itself are difficult to produce since they are highly dependent on the collection methodology. Since polling techniques may have a lower priority than other routing processes in the operating system, aggregate results for such short time scales may be unreliable. However, for the remainder of this section, we focus on longer time scale behavior which does not suffer from this problem.

3.2 Aggregate Behavior

The main focus of this work is the impact of BGP dynamics on time scales that are long enough to have the potential to increase route convergence times and impact router stability. The first issue to address here is how much the overall contribution of the BGP processes is to the CPU cycles consumed by the router operating system.

The interactive session data shows the number of CPU cycles that each process has consumed since the router was booted. If we add the values for the three BGP processes and compare that to the sum of all the processes, we know the percentage of CPU cycles that the BGP protocol has consumed. On 20 February 2003, we collected a single snapshot of "show process cpu" data from all 196 routers that we have access to. On this day, no significant outage or abnormal network behavior occurred. Over 85% of the routers had been up for over 10 weeks when we collected the snapshots, which we consider to be long enough for a reliable reading. We calculate the percentage of CPU cycles that BGP processes consume and plot the histogram in Figure 1. We see that for the majority of routers, BGP processes consume over 60% of CPU cycles. Most of the routers below 60% were being phased out of the network and thus had few or no

[6] http://www.zebra.org

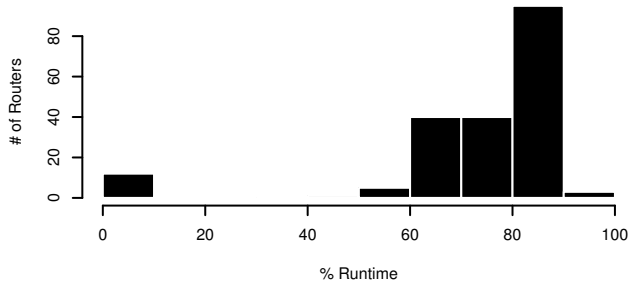

Fig. 1. CPU Utilization due to BGP Processes

BGP adjacencies. This histogram was similar across other days that we collected data on. We find that the BGP processes consume the majority of CPU cycles during router uptime. Given how much resource this protocol consumes, there is significant potential for protocol dynamics to directly impact router CPU load.

We now consider how frequently high CPU load occurs in operational routers over a 2.5 year period. During this time, a CPU load value is reported every 5 minutes via SNMP. Across the routers that we have data for over 2.5 years, roughly 0.6% of the 5 minute samples are missing. This may be due to router reboots and / or losses in SNMP data collection. We find that typically in less than 1% of these samples the CPU load was above 50%. For the vast majority of time across these routers, the CPU load was below 50%.

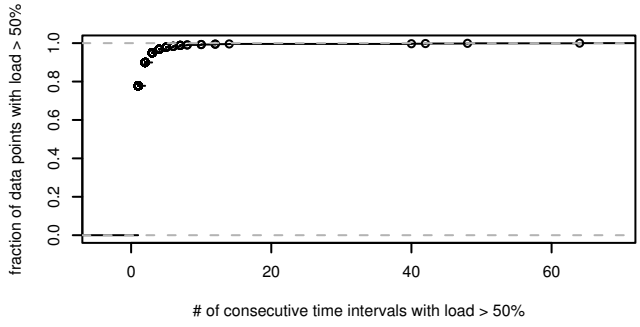

Fig. 2. Cumulative Sum of # of Time Intervals with CPU Load $> 50\%$ (Sept. 2000 to Feb. 2003)

In Figure 2, we only consider time periods when the CPU load was above 50%. We transform the data into the number of consecutive time intervals with high CPU load and plot the cumulative sum of the sorted data. This shows that of the high CPU load occurrences, the majority of them occur for short time periods, but there are some that occur for long periods of time. This behavior is typical compared to the other routers for which we have data. These graphs are over very long periods of time, during which abnormal network conditions may have occurred to cause the long durations of high load. We now examine typical network conditions and abnormal network conditions separately, and specifically consider the impact by BGP.

3.3 Typical Network Conditions

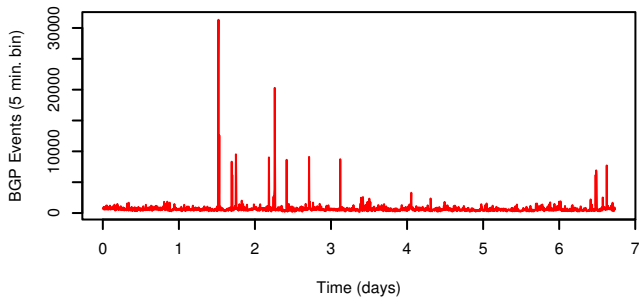

Fig. 3. iBGP Updates During a Typical Week (10-17 Feb. 2003)

We begin by considering a typical week (10-17 February 2003) when no significant network event occurred. We examine if variations in the rate of BGP changes impact the average CPU utilization. In Figure 3, we show the number of BGP routing table changes at a router in the Sprint network. The graph is similar for other routers in the network; this one is picked arbitrarily. Each point in the graph represents the total number of changes to the BGP table during a 5 minute period. We see that on average, there are about 600 routing table changes every 5 minutes, but spikes of much higher rate of change occur. One such spike consisted of over $30,000$ changes, which we denote as "Event A". However, these spikes do not last a long time.

During this same time period, we plot the CPU load in percentage for the same router in Figure 4. Each point shows the percentage of CPU cycles that were consumed by the operating system (the remaining cycles are idle). This value is the 1 minute exponentially-decayed moving average of the CPU load and there is a point every 5 minutes. We see that the load is typically around 25%, and in one case exceeded 45% (which we denote as "Event B").

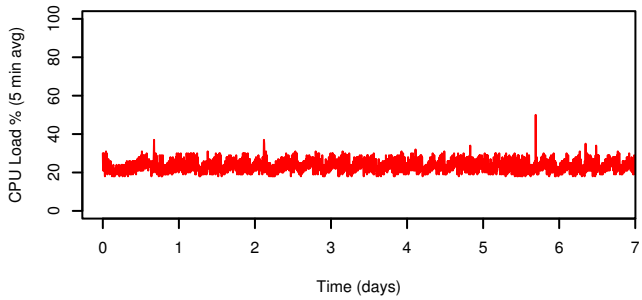

Fig. 4. CPU Load During a Typical Week (10-17 Feb. 2003)

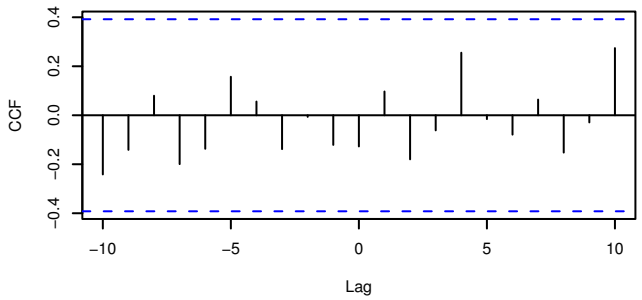

Fig. 5. CCF for 2 Hours Around Event A

Comparing Figure 3 to Figure 4 shows little correlation. There is very little cross correlation between the two time series over the whole week. The CCF (cross correlation function) magnitude is less than 0.1. In Figure 5, we show the cross correlation between the two time series for a two hour period around "Event A". We see that even during this short but significant increase in the number of BGP events, there is only a small correlation with the CPU load (a maximum CCF of about 0.3). In Figure 6, we focus on "Event B" where there was a significant increase in the CPU load. While there is some correlation here (a maximum CCF of about 0.6), when we check Figure 3 around "Event B", we do not see a very large increase compared to normal activity throughout the week.

This behavior we observe is typical across other routers and during other time periods. On average, the cross correlation is below 0.15. In some instances, for two hour periods around specific cases of above average CPU utilization or high BGP activity, the

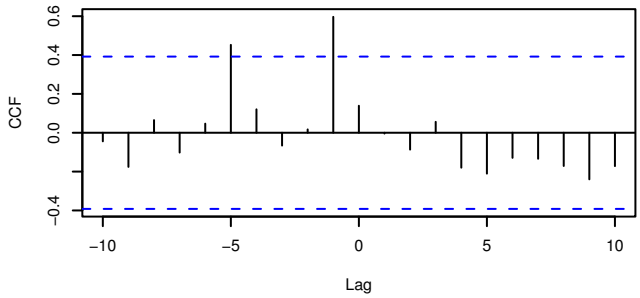

Fig. 6. CCF for 2 Hours Around Event B

cross correlation is around 0.5. However, in none of these instances have we observed *both* high (significantly above average) CPU load and high BGP activity.

3.4 Abnormal Network Conditions

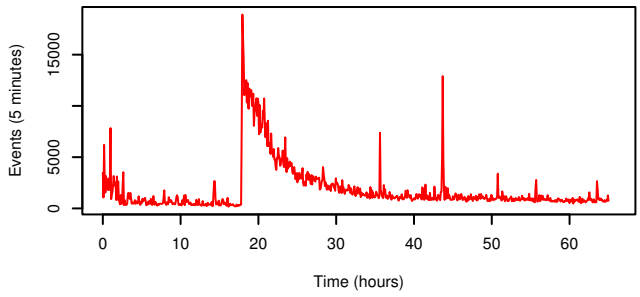

Fig. 7. iBGP Updates During the SQL Worm Attack (24-26 Jan. 2003)

We now focus on abnormal network conditions. Around 05:30 UTC on 25 January 2003, the Sapphire/Slammer SQL worm attacked various end hosts on the Internet. While routers were not targeted, the additional traffic generated by the attack caused various links on the Internet to get saturated. This caused router adjacencies to be lost due to congestion, resulting in a withdrawal of BGP routes. These withdrawals propagated across the Internet. Upon withdrawal of these routes, congestion would no longer

occur on these links and BGP sessions would be restored, causing BGP routes to be re-added. This cycle repeated until filters were applied to drop the attack traffic. As a result, in Figure 7, we see a significant amount of BGP traffic. While the peak value of about $20,000$ route changes in a 5 minute window is not significant compared to the peaks in Figure 3, the length of time is. This large amount of activity took many hours to abate, instead of just a several minute spike. While this graph shows the number of changes at a particular router in the Sprint network, all routers that we have BGP data from experienced similar activity.

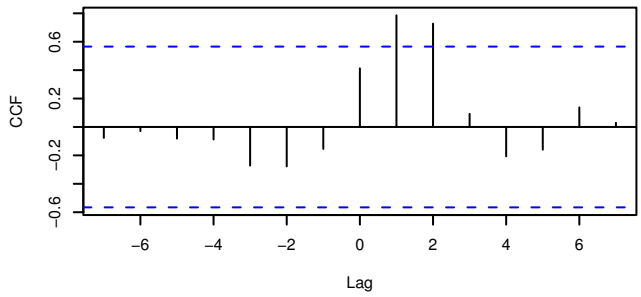

Fig. 8. CCF During the SQL Worm Attack (24-26 Jan. 2003), One Hour Around Event

We correlate this time series with the CPU load percentage for the same router. Across the 65 hour period, there is a correlation of 0.5 around time lag of 0 to 5 minutes. When we focus on 12 hours before and 12 hours after the start of the attack, we see a stronger correlation of 0.6. For a period of 30 minutes before and 30 minutes after the event in Figure 8, a maximum correlation of 0.7 is observed. We plot the maximum correlation between BGP changes and CPU load across all 196 routers that we have access to as a histogram in Figure 9. We see that most routers had a correlation of over 0.5 during this abnormal event.

However, even with a strong correlation, we need to consider the magnitude of the impact on the router CPU load. In Figure 10, we show a histogram of the increase in CPU load during the 65 hour period that we consider across the 196 routers. The value that we show is the difference between the lowest CPU load and highest CPU load that each router experienced during the 3 day period. We see that despite the strong correlation, for most routers, there was less than a 20% increase in the router CPU load. In Figure 11 we show the histogram of the highest CPU utilization experienced by each router at any time during the 3 day period. We see that in most cases, the maximum load was below 50%. A few outliers above 50% exist in the data set, but manual inspection revealed that these few routers underwent scheduled maintenance during the increase in CPU load.

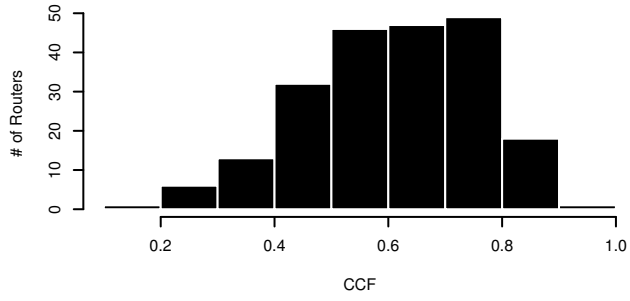

Fig. 9. Max. CCF During the SQL Worm Attack (24-26 Jan. 2003), 1 Hour Around Event

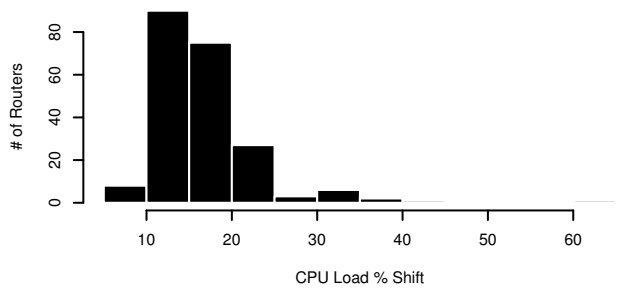

Fig. 10. CPU Utilization Increase During the SQL Worm Attack (24-26 Jan. 2003)

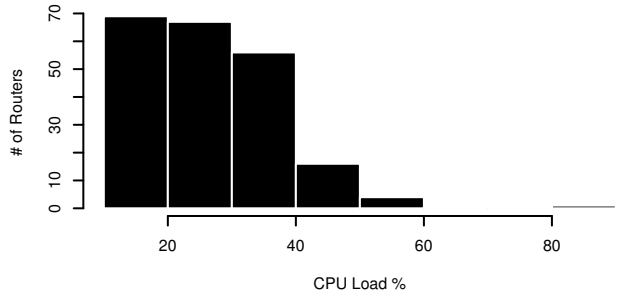

Fig. 11. Maximum CPU Utilization During the SQL Worm Attack (24-26 Jan. 2003)

4 Conclusions

Router CPU load is a significant concern in the operation of an ISP. Long periods of high CPU load can increase route convergence times, and has some correlation with router instability. The BGP routing protocol has a potential for significantly impacting CPU load due to the nature of the protocol. Several actions need to be taken upon receiving a route announcement, and on average more route announcements are seen by an ISP today than ever before. In this study, we examine the impact of BGP activity on operational routers in the Sprint IP network.

We find that on average, BGP processes tend to consume over 60% of a router's non-idle CPU cycles. During short time scales (5 seconds), we have observed BGP processes contributing almost 100% CPU load. During longer time scales ($1 - 5$ minutes), we see a weaker correlation. During normal network operation, we find that there is some correlation between increased BGP activity and router CPU load, but the impact is small. During an abnormal network event that lasted over 10 hours, we find a correlation. However, the increase in CPU load was under 20% for most routers.

BGP processes tend to run frequently for very short intervals, during which the CPU can reach the maximum utilization. Due to this, BGP consumes the majority of CPU cycles over a very long period of time (weeks). The quantity of BGP messages received during a particular cycle can increase the CPU load. However, this increase is not consistently large enough to cause concern about the operation of the router. A possibility is that certain kinds of BGP messages may increase the load more than others. However, without a detailed understanding of the specific implementation of BGP in these routers, we cannot speculate on such specific behavior.

Thus we conclude that during normal operation, CPU load is not significantly impacted by BGP activity in the time scale of minutes. Short term impact in the time scale of seconds in not likely to significantly impact convergence times or router stability. During abnormal events of the magnitude of the SQL Slammer worm, router CPU is not likely to increase significantly.

References

1. Stewart, J.W.: BGP4: Inter-Domain Routing in the Internet. Addison-Wesley (1998)
2. Huston, G.: Analyzing the Internet's BGP Routing Table. Cisco Internet Protocol Journal (2001)
3. Bu, T., Gao, L., Towsley, D.: On routing table growth. In: Proc. IEEE Global Internet Symposium. (2002)
4. Halabi, S., McPherson, D.: Internet Routing Architectures. Second edn. Cisco Press (2001)
5. Cowie, J., Ogielski, A., Premore, B., Yuan, Y.: Global Routing Instabilities during Code Red II and Nimda Worm Propagation. Draft paper (2001)
6. Wang, L., Zhao, X., Pei, D., Bush, R., Massey, D., Mankin, A., Wu, S.F., Zhang, L.: Observation and analysis of BGP behavior under stress. In: Proc. Internet Measurement Workshop. (2002)
7. Bollapragada, V., Murphy, C., White, R.: Inside Cisco IOS Software Architecture. Cisco Press (2000)