When Is Service Really Denied?
A User-Centric DoS Metric

Jelena Mirkovic
University of Delaware
sunshine@cis.udel.edu

Alefiya Hussain
SPARTA, Inc.
Alefiya.Hussain@sparta.com

Brett Wilson
SPARTA, Inc.
Brett.Wilson@sparta.com

Sonia Fahmy
Purdue University
fahmy@cs.purdue.edu

Wei-Min Yao
Purdue University
wmyao@cs.purdue.edu

Peter Reiher
UCLA
reiher@cs.ucla.edu

Stephen Schwab
SPARTA, Inc.
Stephen.Schwab@sparta.com

Roshan Thomas
SPARTA, Inc.
Roshan.Thomas@sparta.com

Categories and Subject Descriptors: C.4 Performance of systems: Measurement techniques
General Terms: Measurement, security, standardization.
Keywords: Denial of service, metrics.

1. INTRODUCTION

Accurately measuring DoS impact is essential for evaluation of potential DoS defenses. Current approaches to quantify the impact of DoS attacks involve a collection of one or several traffic measurements (e.g. legitimate traffic’s request/response delay, transaction duration, loss, etc.) and a comparison of their first order statistics or the value distributions in the baseline, the attack-only and the attack-with-defense case. This measurement approach causes the results to be incomplete, as each independent traffic measurement captures only one aspect of the service denial, and measurements collected in different scenarios cannot be compared. Comparisons of measurement statistics or distributions among test cases result in imprecise metrics. These can only express that network traffic behaves differently under attack, but do not accurately measure which services have been denied and how severely.

We propose a novel, user-centric approach to DoS impact measurement that holistically captures a user’s QoS perception during a test scenario. We define QoS requirements for a large range of popular Internet applications and identify traffic parameters and corresponding thresholds that define good service range. For each legitimate transaction during a testbed experiment or a simulation, we measure the selected traffic parameters and compare them to their corresponding thresholds. Transactions that do not meet all the requirements are considered failed. We aggregate the information about transaction failure into two composite metrics to expose the precise interaction of the DoS attack with the legitimate traffic: the impact of attack on various applications and times when failures occur. We illustrate the inadequacy of the existing metrics and the utility of our proposed metrics in live testbed experiments on the DETER testbed [2].

2. PROPOSED DOS IMPACT METRICS

We propose to measure the impact of a DoS attack on network services by directly measuring the quality of service experienced by end users. For the popular applications in today’s Internet, we identified the traffic parameters whose values indicate if the particular application’s service was denied. We further defined a series of thresholds for relevant parameters that, when breached, indicate poor service quality. When defining these thresholds we were guided by the existing findings in the QoS research [3, 1, 6] and efforts of large standard bodies to define QoS requirements for next generation telecommunication networks [5]. Table 1 lists the proposed QoS requirements.

<table>
<thead>
<tr>
<th>Category</th>
<th>One-way delay</th>
<th>Req/res delay</th>
<th>Loss</th>
<th>Dur</th>
<th>Jitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email</td>
<td>&lt; 150 ms</td>
<td>whole, RTT &lt; 4 s</td>
<td>&lt; 3%</td>
<td>&lt; 50 ms</td>
<td></td>
</tr>
<tr>
<td>Chat, typ</td>
<td>&lt; 150 ms</td>
<td>whole, RTT &lt; 4 s</td>
<td>&lt; 3%</td>
<td>&lt; 50 ms</td>
<td></td>
</tr>
<tr>
<td>Chat, vod</td>
<td>&lt; 200 ms</td>
<td>whole, RTT &lt; 4 s</td>
<td>&lt; 3%</td>
<td>&lt; 50 ms</td>
<td></td>
</tr>
<tr>
<td>Web</td>
<td>&lt; 150 ms</td>
<td>whole, RTT &lt; 4 s</td>
<td>&lt; 3%</td>
<td>&lt; 50 ms</td>
<td></td>
</tr>
<tr>
<td>FTP</td>
<td>&lt; 200 ms</td>
<td>whole, RTT &lt; 4 s</td>
<td>&lt; 3%</td>
<td>&lt; 50 ms</td>
<td></td>
</tr>
<tr>
<td>DNS</td>
<td>&lt; 200 ms</td>
<td>whole, RTT &lt; 4 s</td>
<td>&lt; 3%</td>
<td>&lt; 50 ms</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Application categories and their QoS requirements

We interpret the traffic as series of transactions that represent higher-level tasks whose completion is meaningful to a user, such as browsing one Web page or downloading one file. For each transaction in the experiment, we measure the chosen traffic parameters and compare them to their corresponding thresholds. Transactions that violate at least one of their thresholds are considered failed. Our main DoS impact measure is the percentage of failed transactions (pft) in each application category. We aggregate the measures of transaction success and failure into two composite metrics to capture the DoS impact on network services: (1) The DoS-hist measure is the histogram of pft measures across application categories. (2) The failure ratio measure is the percentage of transactions that are alive in the current interval, but will die in the future, and captures the time-varying nature of some attacks and the timeliness of a defense’s response.

3. EVALUATION RESULTS

We evaluated our metrics with many popular variants of DoS
attacks, and proved that they could accurately capture the attack’s impact. Due to space constraints, we only show two such scenarios.

The experimental topology is shown in Figure 1. We simulate six application types: Web, DNS, FTP, Telnet, IRC and VoIP. The servers for these applications are labeled in the Figure 1. Each attack network hosts two attackers, and each legitimate network hosts two clients and four servers. Clients talk with servers in their own network, and with servers from two out of three external networks. Traffic patterns are shown in Figure 1. Wherever possible, we used real server and client applications to generate traffic, so we could faithfully replicate traffic dynamics and avoid artifacts introduced by traffic generators. File sizes, user request arrivals and transaction durations are drawn from the distributions observed in real-world traffic [4]. We select attack targets from the network Net3. Our first experiment is the UDP bandwidth flood. All four attackers target the DNS/Web server in Net3, and send at the maximum possible speed. Figure 2 shows the DoS-hist measures. Labels on the top of the graph indicate DoS-hist measures that belong to the same source network, x-axis labels denote the destination network, and y-axis shows the pft per application.

The traffic from and to Net1 experiences the largest service denial because the attack from network ANet1 shares the bottleneck link with this legitimate traffic, and completely saturates this link. Almost all traffic from and to Net1 is denied service. Traffic from and to Net3 experiences less service denial, for two reasons: (1) the attack traffic from network ANet3 does not cross the bottleneck link between the Net3 and the core, and (2) the attack traffic from ANet1 arrives to the bottleneck link at a small volume (10 Mbps), because it was shaped by the link connecting Net1 to the core. Similar percentages of outgoing and incoming transactions to Net3 fail.

We show the legacy metric of request/response delay for traffic originated from Net1 in Figure 3. The Figure shows the distribution of this metric in the baseline case and when the attack is present, on a logarithmic scale. While the distribution under attack looks different than the distribution in the baseline case, they are not very far apart. Some transactions that fail have the same or lower request/response delay than transactions that have succeeded, thus indicating that this metric alone cannot accurately capture the DoS impact. We have highlighted one such point A on the figure.

Our second experiment is the UDP pulsing attack, with the same parameters as in the case of UDP bandwidth flood. The pulses start at 195 seconds, last for 20 seconds, with a sleep time between pulses of 100 seconds. There are a total of 5 pulses. Figure 4 shows the failure ratio for transactions originating from Net1 to Net3 during the experiment. The failure ratio oscillates with the attack, but the transactions fail even when the attack is not present, because the periodic loss inflicts significant damage that cannot be compensated until the next pulse’s activation.

4. CONCLUSIONS

Ultimately, DoS attacks are about preventing users from doing what they legitimately and ordinarily want to do. Only a metric that measures a user-level experience can truly capture the effect of a denial of service attack. Our proposed metrics meet this goal by measuring a user’s quality of service experience, and comparing it against application-specific QoS thresholds. While much more work remains on refining the proposed metrics, this paper is the first step towards precise and objective DoS impact evaluation.

5. REFERENCES