Sustainability of Service Provisioning Systems under Attack

Georgios S. Paschos  
MIT, USA  
CERTH-ITI, Greece  
gpasxos@mit.edu

Leandros Tassiulas  
University of Thessaly, Greece  
leandros@uth.gr

ABSTRACT

We propose a resource allocation model that captures the interaction between legitimate users of a distributed service provisioning system with malicious intruders attempting to disrupt its operation. The system consists of a bank of servers providing service to incoming requests. Malicious intruders generate fake traffic to the servers attempting to degrade service provisioning. Legitimate traffic may be balanced using available mechanisms in order to mitigate the damage from the attack. We characterize the guaranteed region, i.e., the set of legitimate traffic intensities that are sustainable given specific intensities of the fake traffic, under the assumption that the fake traffic is routed using static policies. This assumption will be relaxed, allowing arbitrary routing policies, in the full version of this work.

Categories and Subject Descriptors

H.1 [Information Systems Applications]: Models and Principles; Miscellaneous;

General Terms

Algorithms, Reliability, Theory

Keywords

Service provisioning system; guaranteed sustainability; stability

1. SYSTEM MODEL AND DEFINITIONS

Consider a set \( N \triangleq \{1, \ldots, N\} \) of parallel servers with constant service rates \( \mu_n, n \in N \). The servers are fed by a set of legitimate streams \( L \triangleq \{1, \ldots, |L|\} \) of traffic, each stream \( l \in L \) associated with traffic intensity \( a_l \) and a set of reachable servers \( S_l \subseteq N \). The traffic arriving from a stream \( l \) is routed to some of the servers in \( S_l \).

A malicious system launches a Degradation of Service attack (a type of Denial of Service attack) in order to disrupt the operation of the system. In particular, the malicious system has a set \( M \triangleq \{1, \ldots, |M|\} \) of malicious traffic streams, where the stream \( m \in M \) generates fake traffic with intensity \( b_m \) and is capable of routing it towards a subset of servers \( Q_m \subseteq N \). See Figure 1 for an example of the studied system in terms of a bipartite graph.

We assume the operation of two controllers with conflicting interests. Controller 1 splits legitimate traffic to al-

![Figure 1: An example of the system for 2 servers, 3 legitimate streams and 2 malicious streams. Also, \( S_1 = Q_1 = \{1\} \), \( S_2 = Q_2 = \{1, 2\} \) and \( S_3 = \{2\} \).](image)

Copyright is held by the author/owner(s).
SIGMETRICS ’13, June 17–21, 2013, Pittsburgh, PA, USA.
ACM 978-1-4503-1900-3/13/06.
The guaranteed region is parametrized by the fake traffic intensity, $b$. For $b$ large enough, $\Lambda_b$ might contain only the zero element vector $0 \triangleq (0, 0, \ldots, 0)$, which implies that there is a malicious policy $\phi$ such that even arbitrarily small legitimate traffic intensities are not sustainable, regardless of the legitimate policy $f$ used. In practical terms, we can think of such a situation as a DoS attack. The DDoS attack, on the other hand, corresponds to cases where the guaranteed region is not degenerated and legitimate traffic can still be sustained despite the attack, albeit in smaller intensities.

2. MAIN RESULT

First, we fix a malicious policy $\phi$ and study the sustainable region of traffic intensities under this policy. Let $r_n(\phi) \triangleq (\mu_n - \sum_{m \in M} \phi_{mn})^+ \geq b$ be the available resource of server $n$ after the traffic arriving from malicious streams under $\phi$ is subtracted. We use $(\cdot)^+ \triangleq \max\{\cdot, 0\}$. Using the stability definition, we conclude that the system is stable iff there exists a legitimate policy $f$ such that

$$\sum_{i \in L} f_n \leq r_n(\phi), \text{ for all } n \in \mathcal{N}. \quad (1)$$

In what follows, we will express the sustainable region $\Lambda_b^\phi$ in terms of traffic intensities $a, b$ and service rates $\mu$. For an arbitrary non-empty subset of the servers $N \subseteq \mathcal{N}$ consider the induced subsets $L, M$, where

- $\mathcal{L} = \{i \in \mathcal{L} : S_i \subseteq \mathcal{N}\}$ is the set of legitimate traffic streams that must direct all traffic to some of the servers in $\mathcal{N}$
- $\mathcal{M} = \{m \in \mathcal{M} : Q_m \cap \mathcal{N} \neq \emptyset\}$ is the set of fake traffic streams that can direct fake traffic to some of the servers in $\mathcal{N}$.

**Lemma 1** (Cut constraints). The traffic intensities $a$ are sustainable under $\phi$ if and only if

$$\sum_{i \in \mathcal{L}} a_i \leq \sum_{n \in \mathcal{N}} r_n(\phi), \text{ for all } \mathcal{N} \subseteq \mathcal{N}.$$

2.1 Guaranteed region $\Lambda_b$

Consider an auxiliary network $G(\mathcal{N}) = (\mathcal{V}, \mathcal{E})$. We define the set of nodes as $\mathcal{V} \triangleq \{s, t, u_i, v_j : i \in \mathcal{M}, j \in \mathcal{N}\}$, where $s$ is the source node, $t$ is the sink, nodes $u_i, v_j \in \mathcal{M}$ correspond to members of $\mathcal{M}$ and nodes $v_j, j \in \mathcal{N}$ correspond to members of $\mathcal{N}$. The set of links consists of three subsets $\mathcal{E} = \mathcal{E}_a \cup \mathcal{E}_b \cup \mathcal{E}_q$, where each subset consists of directional links defined as follows

$$\mathcal{E}_a \triangleq \{(s, u_i) : i \in \mathcal{M}\}, \mathcal{E}_b \triangleq \{(v_j, t) : j \in \mathcal{N}\}$$

$$\mathcal{E}_q \triangleq \{(u_i, v_j) : i \in \mathcal{M}, j \in \mathcal{Q} \}.$$

A link $(s, u_i)$ has capacity $b_i$, a link $(v_j, t)$ has capacity $\mu_j$, while all links in subset $\mathcal{E}_q$ have infinite capacity. Let $M_{\text{max}}(\mathcal{N})$ denote the maximum $s$-$t$ flow of network $G(\mathcal{N})$.

**Definition 3** (Guaranteed region $\Lambda_b$). The guaranteed region $\Lambda_b$ of the system attacked by a malicious adversary with available traffic intensities $b$, is the set of all $a$ for which there exists a legitimate policy $f \in \Pi_1$, such that the system remains stable under any selection $\phi \in \Pi_2$.

The guaranteed region is parametrized by the fake traffic intensity, $b$. For $b$ large enough, $\Lambda_b$ might contain only the zero element vector $0 \triangleq (0, 0, \ldots, 0)$, which implies that there is a malicious policy $\phi$ such that even arbitrarily small legitimate traffic intensities are not sustainable, regardless of the legitimate policy $f$ used. In practical terms, we can think of such a situation as a DoS attack. The DDoS attack, on the other hand, corresponds to cases where the guaranteed region is not degenerated and legitimate traffic can still be sustained despite the attack, albeit in smaller intensities.

3. DISCUSSION OF THE DYNAMIC CASE

In the followup work we extend the study to the case of dynamic routing policies. In case controller 2 is static, Join the Shortest Queue turns out to be the optimal policy for the legitimate controller and it can be shown that the guaranteed region described here is achieved by this policy. If, however, the controller 2 is allowed to allocate bogus jobs in a dynamic fashion, the guaranteed region changes drastically. Below we demonstrate this in an example.

3.1 An example with two servers

Consider two servers with unit service rate fed by one legitimate stream with traffic $a$ and one malicious with traffic $b$. Traffic can be routed to both servers. Using the results of the previous section, we conclude that $a + b \leq 2$ is a necessary and sufficient condition for guaranteed sustainability as long as the malicious intruder is constrained to static routing policies. We call this region $\Lambda_b^{st}$, see Figure 2.

**Definition 4** (Conditions C.1). The following inequality is satisfied

$$\sum_{i \in \mathcal{L}} a_i \leq \sum_{n \in \mathcal{N}} \mu_n - M_{\text{max}}(\mathcal{N}), \text{ for all } \mathcal{N} \subseteq \mathcal{N}. \quad (2)$$

**Theorem 1.** (Guaranteed region) Conditions C.1 are necessary and sufficient to guarantee sustainability for the traffic intensity $a$ under any $\phi$.

4. ACKNOWLEDGMENTS

The work of G. Paschos is supported by the WiNC project of the Action:Supporting Postdoctoral Researchers, funded by national and Community funds (European Social Fund).