Impact of Denial of Service Solutions on Network Quality of Service

Scott Fowler  
Linköping University  
Department of Science and Technology  
SE-601 74, Norrköping, Sweden

Sherali Zeadally  
Department of Computer Science and Information Technology  
University of the District of Columbia  
Washington DC, USA

Naveen Chilamkurti  
Department of Computer Science and Computer Engineering  
La Trobe University, Melbourne  
Victoria, Australia-3086

Abstract—The Internet has become a universal communication network tool. It has evolved from a platform that supports best-effort traffic to one that now carries different traffic types including those involving continuous media with Quality of Service (QoS) requirements. As more services are delivered over the Internet, we face increasing risk to their availability given that malicious attacks on those Internet services continue to increase. Several networks have witnessed Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks over the past few years which have disrupted QoS of network services, thereby violating the Service Level Agreement (SLA) between the client and the Internet Service Provider (ISP). Hence DoS or DDoS attacks are major threats to network QoS. In this paper we survey techniques and solutions that have been deployed to thwart DoS and DDoS attacks and we evaluate them in terms of their impact on network QoS for Internet services. We also present vulnerabilities that can be exploited for QoS protocols and also affect QoS if exploited. In addition, we also highlight challenges that still need to be addressed to achieve end-to-end QoS with recently proposed DoS/DDoS solutions.

Index Terms—MPLS, RSVP, DiffServ, QoS, Denial of Service, DoS, DDoS, Protocol, Performance, Security

I. INTRODUCTION

The Internet has become a universal communication network tool. It has evolved from a platform that used to support primarily best effort traffic to one that is now delivering all kinds of traffic types (including those involving continuous media such as audio and video). The support of multimedia traffic on the Internet has been an active topic of research because of their stringent Quality of Service (QoS) requirements such as bandwidth, delay, etc. It remains a significant challenge to support traffic with QoS requirements and simultaneously enable secure transmissions of such traffic types. It is difficult to overcome this challenge since the Internet is very susceptible to attacks such as Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks. Indeed, the goal of DoS or DDoS attacks is to reduce the availability of network resources such as CPU or buffers, resulting in the disruption of services provided to legitimate sources. Consequently, DoS and DDoS attacks affect network QoS and violate Service Level Agreement (SLA) between the client and the Internet Service Provider (ISP). Such attacks are major threats to network QoS [1], [2].

Recent security research has focused a lot on privacy or authentication. Many networking security systems and mechanisms are designed to provide more secure infrastructures, such as firewalls, authentication and authorization protocols, packet filtering, source identification and cryptographic algorithms. These security methods require extra storage or computational resources, and the resulting limited resources become easy targets for DoS attacks [29].

DoS attacks cannot be simply addressed using traditional security approaches with or without the above mentioned storage or computational resource issues [29]. For example, the Internet Key Exchange protocol (IKE) used for key establishment and Security Associations (SA) parameter used by IPsec are susceptible to DoS attacks (Figure 1) [3], [29] since the server has to create states for SA and computes Diffie-Hellman exponential generation [2], [3]. It is worth noting that in Figure 1 the Certificate from the client is never sent, resulting in the Server waiting for a response or re-initialization of the IKE exchange certificates. According to the latest data publicly available from Carnegie Mellon’s CERT® Coordination Center report [7], 21% of organizations which responded to the survey reported that financial loss caused by electronic crimes (e-crimes) had increased over the year (2007). Among the e-crimes, 49% of the organizations had experienced DoS attacks, which is the fifth most common e-crime after malicious code crimes, unauthorized access, spam email and spyware [7].

This work demonstrates that DoS/DDoS attacks are major threats to network QoS. Although several review papers have provided various taxonomies of DoS/DDoS attacks, the impact of such attacks on QoS has not really been studied and investigated. Indeed, it remains a significant challenge to ensure and maintain network QoS of different types of traffic while maintaining secure network infrastructures. In this context, the
main contributions of this work are as follows:

- Many kinds of defensive measures have been proposed to thwart DoS/DDoS attacks. However many of these defensive techniques do not take into account the effects they have on network QoS delivery. In this work, we investigate recently proposed DDoS solutions and we systematically evaluate their impact on network QoS.
- In addition to the overview of the each defensive technique based on their effectiveness on the security, we will extensively discuss on how it may counteract QoS provision.
- The vulnerabilities of QoS protocols: Several QoS approaches and protocols have been proposed and implemented to support Multimedia over networks. Among them, several traffic engineering technologies including Multi-Protocol Label Switching (MPLS), Differentiated Services (Diffserv), and Resource Reservation Protocol (RSVP) have gained wide acceptance. These protocols, however, are prone to DoS/DDoS attacks. Therefore, we will summarize each QoS protocol and thoroughly describe its vulnerability to DoS/DDoS attack.

The rest of this paper is organized as follows. Section II describes elementary models of DoS and DDoS attacks. Section III presents five basic DoS attack strategies. In Section IV, various defense measures aimed at preventing DDoS attacks and their effects on network QoS are presented. In Section V, we identify DDoS attacks that can exploit vulnerabilities of various QoS protocols. In Section VI, we present future works. Finally, in Section VII we make some concluding remarks.

II. BACKGROUND ON BASIC DoS/DDoS MODELS

According to the World Wide Web Security Consortium [33], [68] “Denial of Service (DoS) is an attack designed to render a computer or network incapable of providing normal services” (Figure 2). DoS attacks accomplish this by attacking a target computer’s (victim) network bandwidth or connectivity. With bandwidth attacks, the network is flooded with a high volume of malicious traffic, resulting in all the available network resources to be consumed and legitimate user requests to be denied.

The WWW Security Consortium also defines [33], [68], “A Distributed Denial of Service (DDoS) attack as one that uses many computers to launch a coordinated DoS attack against one or more targets. Using client/server technology, the perpetrator is able to multiply the effectiveness of the Denial of Service significantly by harnessing the resources of multiple unwitting accomplice computers which serve as attack platforms” (Figure 3).

There are two types of DDoS attacks that have emerged, the Handler-based and Internet Relay Chat (IRC)-based [4]–[6], [54]. In the Handler-based model (Figure 4): A handler is a special program that can be deployed throughout the Internet. Each handler is capable of controlling multiple agents. An agent is a compromised system that is responsible for generating a stream of packets directed toward the intended victim. Sometimes, the terms “handler” and “agent” are referred as “master” and “daemon” [54].

The IRC-based model (Figure 5) is similar to the Handler-based model, but instead of the Handler program an IRC (Internet Relay Chat) communication channel is employed to link the client to the various agents. The use of an IRC commu-
nication channel allows an attacker to exploit “legitimate” IRC ports to send commands to the various agents [4], [54] making DDoS packets more difficult to be identified. In addition, the client does not need to keep a record of agents because it only has to log on to the IRC server and check an inventory of accessible agents [4], [54]. In an IRC-based DDoS attack architecture, the agents are often referred to as “Zombie Bots” or “Bots” [54].

One of the network architectures, namely, Peer-to-Peer (P2P), has been receiving a lot of attention recently. P2P enables a highly scalable, decentralized, and robust distribution service for live multimedia streaming applications (such as Skype used for voice or video) compared to the classical Client-Server network paradigm. P2P is also prone to DoS attacks. Typical DoS attacks [76], [77] on a P2P include: a) a peer stores and retrieves data repeatedly, b) a peer joins and leaves the P2P network rapidly, c) a peer sends unsolicited messages to other peers repeatedly, and d) a peer disturbs the routing information of the P2P network.

III. BASIC DoS ATTACK STRATEGIES

DoS attack strategies can be divided up into 5 categories [54], [68]. These five categories can be summarized as follows:

1) Network Level: This type of attack is accomplished by taking advantage of a bug(s), weaknesses in the software, or trying to exhaust the hardware resources of the network. One example of an exploited network service is the Secure Socket Layer (SSL). The Cisco CSS 11500 Series Content Services Switches [69] when configured with SSL termination services were vulnerable to DoS attacks when processing malformed client certificates.

2) Operating System Level: Exploits the way the operating system executes its protocol. An example is the Ping of Death (or Ping Flood) [54], [70]. The attack generates a large number of ICMP packets (ping) that exceed the maximum size of 65535 bytes. The victim receives the ping in fragments and starts reassembling the packet. Unfortunately, once the packet is reassembled, it is too big for the buffer and overflows it, resulting in reboots or the system hanging.

3) Application Level: A bug (running on the target host (victim)) is exploited at the application level to consume the victim resources. By exploiting bugs via an application, this attack allows an intruder to disrupt services by causing excessive processing on the target host. One such example of this is the finger bomb attack which redirects the finger command to remote sites. In the use of finger username@@host1 [71], the @ causes the finger to recursively finger the same machine (itself) repeatedly, thereby exhausting all of the resources on the host.

4) Attacks based on protocol features: This type of attack takes advantage of the standard protocol characteristic. For example IP spoofing is used to gain unauthorized access to computers by imitating a trusted host IP address. Another attack is Synchronization (SYN) Flooding which takes advantage of a flaw in such away that hosts implement the TCP three-way handshake. When host B receives the SYN request from a malicious host, host B must keep track of the partially opened connection in a “listen queue” for at least 75 seconds. The malicious host can exploit the small size of the listen queue by sending multiple SYN requests to host B, but never replying to the SYN & ACK messages that host B sends back (Figure 6). By doing so, host B’s queue is quickly filled up, leading to rejection of new connections until a partially opened connection in the queue is completed or the timeout has been reached.

5) Data Flooding: Attempts to oversaturate the bandwidth or host by generating huge volumes of traffic.

IV. DEFENSE TECHNIQUES AGAINST DoS/DDoS ATTACKS

There are many kinds of defensive measures that have been proposed to thwart DoS/DDoS attacks [11]–[29], [51], [52], [55]–[60], [62], [63], [65]–[67], [72], [74], [75], [78]–[81], [83], [85], [86], [89], [90], [92]. However, these methods have not considered the impact on network QoS. In this section we discuss several defense techniques against DoS/DDoS and their impact on the QoS delivery over the network.

A. Egress/Ingress filtering

Filtering is the process of separating traffic that can be defined as invalid or inappropriate for a specific part of the network. This filtering will take place at the entry and/or exit points of the network. In other words, egress/ingress filtering is effective in blocking packet with a spoofed source address [29], [59], [72].

While egress/ingress traffic filtering drastically reduces the success of spoofing, there are potential problems when supporting QoS along with the filtering mechanism.

First, egress/ingress filtering does not preclude an attacker from using a forged source address of another host within the permitted prefix filter range. As a result, on large scale

![Figure 6: Transmission Control Protocol (TCP) Three-way-handshaking attack.](image_url)
backbone networks with complex topology, egress/ingress filtering has difficulty differentiating between incoming traffic and outgoing traffic. Therefore, the use of filtering is not very scalable, and this unscalability can affect support for large-scale QoS. Scalability is not an issue when we need to handle only QoS traffic inside the egress/ingress filtering. However, when QoS methods need to be applied outside egress/ingress filtering, QoS traffic and DoS attacks have to be differentiated so that QoS traffic to be recognized as legitimate traffic. Second, networks with filtering routers do not allow a mobile host to send packets directly from the foreign network to the correspondent node using its home address as the source address. To be more specific, traffic to the mobile node is tunneled, but traffic from the mobile node is not tunneled. This results in packets from the mobile node which have source addresses that do not match with the network where the station is currently attached [9]. To accommodate the filtering constraint, reverse tunneling was proposed to establish a correct reverse tunnel from the care-of-address of the mobile node to the home agent [8]. This reverse tunnel enables packets to be sent from the mobile node to its home agent. The packet is then forwarded to the correspondent node after de-tunneling. Unfortunately, this causes triangular routing in the reverse direction and it makes current route optimization techniques become unidirectional or asymmetric [10]. For asymmetric traffic, filtering methods require extra lookups at the router table (on the source address), which in turn results in an increase in delay for QoS traffic.

To take full advantage of egress/ingress filtering technology (spoof prevention, compatibility to currently available protocols, routers, and network infrastructure), we need to investigate how to apply it while simultaneously satisfying multiple QoS requirements. Thus, methods to reduce delays, such as the best possible distribution of filters, rather than simply increasing filters in a large scale network needs to be studied, especially in relation to QoS traffic.

B. Firewall

Firewalls provide a defense mechanism that relies on a well-defined boundary between the inside of the network domain and the outside network domain. Thus, firewalls mediate the passage of information. While firewalls can be very valuable if employed properly, but they are limited in their ability to protect a network [28], [58]. Firewalls will become less effective over time as users tunnel protocols through them, allowing for inadequate security on the tunnel endpoints. Hence a DDoS attack will be able to target an end node supporting QoS. If the tunnels are encrypted, there is no way for the firewall to censor the tunnel or DDoS attacks [28].

To defend against the SYN flooding attack, a firewall may work as a proxy for TCP connection requests. It may also act as a TCP connection request monitor that sends a third message to the destination host to release its resources [29]. Unfortunately, a firewall requires a significantly higher amount of processing time than routing [30], [31]. Processing time increases as the rule set increases in length and complexity [32]. As a result, a firewall can easily become a bottleneck and also becomes susceptible to DoS attacks [29], [32], [57]. These attacks merely delay or prevent legitimate packets from being processed for legitimate connections, such additional delays degrade the QoS of traffic ongoing connections. Robust efficient methods are needed to improve firewall performance for QoS traffic when subjected to security threats.

C. Honeypots

Honeypot act as a decoy by being placed on a network to be attacked so that we can monitor how the network is being compromised from an attack [73], [74]. Usually it consists of a computer that looks as if it is part of a network but limits the attacker’s access to the entire network since it is actually isolated from the network. A honeypot is intended to entice an attacker to install either handler(s) or agent code(s), which enables the honeypot to track the behavior of handler or agent before it presents a problem for the network [64]. The issue with honeypots is how they can be deployed to detect, identify, and capture attacks and whether the honeypot should be placed inside or outside of a firewall. Given that honeypots are only capable of handling attacks directed to themselves, sophisticated attackers can easily avoid honeypots deployed at fixed locations on the network [87]. In other words, they can even attack network QoS with little effort by simply evading the location where honeypots are deployed.

D. IP traceback

IP traceback methods provide the victim the capability to identify the actual true source of the packets causing the DoS attacks [12]. Given the complexity of the current Internet, it is difficult for the victim to determine the actual source of the attack because the attacker routinely counterfeits its IP source address (spoofing). IP traceback is an important concept for restoring and averting reoccurrences of an attack, particularly when an attacker spoofs (counterfeits) its IP source address. Simply discovering the attacker might seem like a narrow goal, but the essential clues it provides may help discover the actual attacker [13]. The challenge of IP traceback is to find an effective and scalable way to track the sources of an IP packet. Hence the IP traceback should minimize time that routers spend on tracking and the storage that is used to keep the tracking information. In this section we discuss four popular IP traceback methods and evaluate them according to how they can affect network QoS.

1) IP traceback by Packet Marking: This approach lets the routers insert additional information into an IP packet so that the victim is able to deduce the path the traffic has taken (Figure 7). Consideration of the robustness of this method needs to be addressed. For example, when the packet marking is not secure or inadequate, an attacker can generate false packet markings (also called spoofed marking) [11], [84]. This spoofed marking results in the victim receiving several false paths, and the victim becomes unable to determine who is the attacker. To make this method more effective, it is necessary to reduce the packet size unless we have downstream
fragmentation. A drawback with downstream fragmentation, however, is that it increases the number of packets into the network. These packets lead to increased network traffic which in turn will affect network QoS. If a larger packet size is allowed instead, it introduces an increase in latency, also affecting network QoS. Furthermore, packet marking has to be fast enough to allow for real-time packet marking. Otherwise, we will likely experience additional delays which will also impact QoS.

Another proposal to minimize the router overheads is Probabilistic Packet Marking (PPM) [12], [84]. This method includes operations such as node append, probabilistic node sampling, and edge sampling. PPM does not incur any storage overhead with routers and the marking procedure (checksum update) can be easily and efficiently executed at current routers. Unfortunately, PPM is not scalable. When there are 25 or more attacking sources, path rebuilding takes considerable amount of time [34]. Therefore, for real-time traffic with QoS requirements this is not a desired solution.

2) **IP traceback by Logging:** One solution for tracking an attacker is to log packets at various key routers throughout the Internet and use data-mining methods to extract information to form a path of the attacker [13]. This solution appears appealing because it allows for precise analysis of the attacker. However, it requires a large amount of processing and storage overhead to save the various logs. Sharing the log information among various ISP, presents a problem of legal issues and privacy concerns. Probabilistic sampling and compression will reduce the various logs, but these methods present considerable demands on the network resources [13]. One logging method proposed is Source Path Isolation Engine (SPIE) [14], [79]. SPIE only hashes relevant portions of a packet in an efficient memory structure by means of bloom filters. Overlay methods also aim to decrease the log, but it presents a burden on network resources such as bandwidth, processing capability from the establishment and maintenance of tunnels [11].

3) **IP traceback by hashing:** To distribute the overhead throughout the network, hash-based IP traceback [14], [78] has been proposed. This method consists of the following components: Data Generation Agents (DGAs), SPIE Collection and Reduction Agents (SCARs), and SPIE Traceback Manager (STM). DGA is applied to each router with bloom filters so that every router deterministically logs some information about each packet traversing the router. SCAR is responsible for one area of the network and is linked to all DGAs inside of its area. The STM is responsible for dealing with the victim’s requests and assembling the path information from the various SCARs (Figure 8).

When the victim’s Intrusion Detection System (IDS) recognizes features of the attacking packets, the IDS will report these features back to the STM (Figure 8). The STM forwards the request to the appropriate SCARs. The various SCARs in the STM area collect the data logged by each router having a DGA. Next, the SCAR reconstructs the attack path inside its area and submits the path reconstruction result back to the STM. The drawback of SPIE is that it incurs heavy computational calculations, distributed management, and storage overheads [11]. The computational burden is also distributed over the network (SCARs and STM) and a high communication burden is placed on the network bandwidth. Since the victim is affected by the high communication overheads, its QoS will also suffer as a consequence.

4) **IP Traceback by ICMP:** The Internet Engineering Task Force (IETF) proposed an ICMP (Internet Control Message Protocol) traceback technique [16], [86] called iTrace as a possible DDoS solution. With this approach, when forwarding packets, a router copies the contents of the packet to a specific type of ICMP traceback message containing information about either the adjacent upstream or downstream routers, or both adjacent upstream and downstream routers. Based on this information, the ICMP traceback message is forwarded towards the same destination as the original packet. An ICMP traceback message to the destination is generated for every 20,000 packets (Figure 9). The ICMP traceback packet includes the identity of the router, contents of the packet, and information about the adjacent routers. During an attack, after receiving a considerable number of traceback messages, the victim can identify the approximate source of the attack by tracing the entire path taken by the attacker from the information received in the ICMP traceback packets.

The basic idea with this scheme is that every router samples one of the packets with low probability (e.g., 1/20,000) by copying contents of the packet into a special ICMP traceback
was introduced to identify and regulate an aggregate traffic
E. Pushback by Aggregate-based Congestion Control
performances [11], [13].
1/20,000 for all paths may seriously degrade the network
result, simply using a predetermined marking probability of
an ICMP traceback message [13]. The proposed method
module would select the kind of packet to use next to generate
information supplied by the routing table, in which a decision
determined by the victim [11]. This is accomplished by explicit
the data obtained for path reconstruction may be promptly
adds intelligence to the marking process. This means that
intention-driven ICMP traceback [15]. The proposed method
from the neighboring routers but very few attackers [13],
attack packets. If all the packets traveling over the network are
attack packets, a victim needs to receive at least 19,999 attack
packets to obtain one ICMP traceback message. Considering
the fact that ICMP packets are not allowed sometimes, the
actual number of attack packets received will take a long
time to reach 20,000. By the time enough ICMP traceback
messages would have been generated, the QoS of ongoing
network services would have already been degraded.
It should be noted, if each DDoS slave only contributes
a small amount of attack traffic, then the probability for a
nearby router picking the right attack packet can be very
small [15]. The routers closer to the victims tend to have more
iTrace messages (or ICMP traceback messages) forwarded to
them because most of the attack traffic has been aggregated.
The victim probably will get many ICMP traceback messages
from the neighboring routers but very few attackers [13],
[15]. A method proposed to address this problem is called
intention-driven ICMP traceback [15]. The proposed method
adds intelligence to the marking process. This means that
the data obtained for path reconstruction may be promptly
determined by the victim [11]. This is accomplished by explicit
information supplied by the routing table, in which a decision
module would select the kind of packet to use next to generate
an ICMP traceback message [13].
Even with the above proposed method, there are several
attacks on different paths with dissimilar lengths. As a re-
result, simply using a predetermined marking probability of
1/20,000 for all paths may seriously degrade the network
performances [11], [13].
E. Pushback by Aggregate-based Congestion Control
The Aggregate-based Congestion Control (ACC) method
was introduced to identify and regulate an aggregate traffic
stream at a single router by means of a pushback mechanism.
The primary goal of ACC is to protect the network and the
rest of the traffic from severe congestions caused by high-
bandwidth aggregates [62], [75]. A router implementing ACC
monitors its level of congestion (Figure 10). When a host is
not the attacker, it decreases its transmission rate. ACC adds
rate-limiting functionality to routers for detecting and dropping
suspicous packets. When the router discovers sustained severe
congestion, it tries to identify the aggregate(s) responsible
for such congestion, using either drop history or random
samples. With this method, the pushback mechanism begins
only when congestion occurs. Actually, the upstream routers
will be notified by the pushback request message to limit the
number of packets destined to the victim to let other legitimate
traffic move because the downstream routers will drop those
packets later anyway. The concept of pushback usually uses
the destination prefix as a congestion signature to distinguish
and protect the legitimate traffic within the aggregate. The
goal is to identify the malicious connections by examining
the destination IP addresses of the dropped packets [62], [65].
When a host decreases its transmission rate, the number of
packets dropped will decrease, resulting in the host being
regarded as a legitimate user. If a large portion of the dropped
packets are identified as having the same IP address, these
packets are considered to belong to the DDoS attack and their
shared destination IP address is used as the attack signature.
Once the malicious traffic is identified, the proposed solution
is to block the malicious traffic [35]–[38]. Unfortunately, the
pushback cannot distinguish between good and bad traffic
going to the destination, and will drop them equally [61], [92].
This could result in the legitimate multimedia traffic being
taken for malicious traffic. For instance, those packets from
hosts that are close to the attackers would also be punished
whether they belong to best effort or multimedia traffic.
F. Authentication
Authentication has long been used in networks to defend
against various attacks. For example, IP Security (IPSec) [39]
provide services such as data source authentication, data
integrity, confidentiality, protection against replay attack, data
privacy, access control, and end-to-end security for IP packets
by means of establishing shared keys between the source
and destination. Hence, various authentication protocols (e.g.
IPSec) are used to verify the identities of the communicating
parties. The integration of authentication protocol to effec-
tively resist network DoS attacks and other types of common
attacks have been proposed [17]–[22], [60].

Authentication based on the public-key infrastructure is computationally expensive. Cryptographic operations often involve extensive computations such as the modular exponentiation. Even with the use of cryptography in authentication to protect against DoS attacks, it is likely to cause fragmentation (due to the packet changing in size) causing an increase in the traffic load on the bandwidth [40], [41] which will affect network QoS. When public-key based authentication is applied, the computation complexity of encryption/decryption consumes time and power [40]. The transmission and encryption/decryption of credentials can affect several QoS parameters such as delay, call dropping probability, and throughput [40], [41].

G. Statistical Analysis

Instead of defending against DDoS attacks by means of authentication, another method proposed for defending against DDoS attacks is statistical analysis [23]–[27], [90].

Statistical methods use the characteristic of the packet header such as IP address or Time-To-Live (TTL) for statistical calculations. Based on the statistical calculation and measurement, packets which are considered to be attack packets are dropped. Such methods preclude the traffic distinctiveness that are naturally constant with standard network processes [24], [27]. As a result, when a DDoS attack occurs, based on the traffic distinctiveness and standard network processes, irregular traffic distinctiveness will be considered to be illegitimate traffic, resulting in the illegitimate traffic to be dropped by the filter that guards the victim’s network [44]. An example of the usage characteristic of the packet is hop-count filtering [25], [63]. This method relies on the TTL field in an IP packet to provide information on the number of hops between the source and destination that are stored in a table with its respective IP address. When an IP packet is received and there is an inconsistency with its hop count and the value stored in the previously built table, the IP packet will be dropped. These methods depend heavily on assumptions of traffic characteristics and probabilistic methods. Sophisticated hackers may use IP addresses with relevant hop count, making this defense strategy altogether ineffective. Furthermore, these methods consider only best effort traffic ignoring QoS traffic. For QoS, the traffic characteristics or the state where the network is in should be considered. For example, when congestion occurs on the network, QoS traffic may be routed to less congested paths, resulting in the number of hops being altered. When calculating statistics of packet attributes, additional latency may be introduced if the calculation is complex resulting in taking more time and power. An end-to-end coordination of QoS traffic is required for statistical calculations of QoS traffic.

H. Overlay Networks

An overlay network is an isolated virtual network deployed over an existing network. It is a collection of varies hosts (originator of the packet sources), routers, and tunnels. Tunnels are network paths in the underling network, providing links in the overlay network. Overlay networks have received a lot of interest for addressing DoS attacks because of link-testing (or hop-by-hop tracing) [13], [56], [80], [91]. Link-testing is done by testing network links between routers in order to identify the source of the attacker. A link-testing method that uses overlay networks is called CenterTrack [51] (Figure 11). With this method, a particular router called a Tracking Router (TR) or a group of TRs are used for input debugging [11], [67]. The input debugging allows the network administrator to decide on incoming network links for specific packets [11]. The TR is directly connected to each ingress\(^2\) and egress\(^3\) edge router of the protected network through tunnels. Once the attacker is detected, TR is able to locate the ingress router of the attacker traffic flow since the ingress edge router may be seen as only one hop away from the TR [11]. However, CenterTrack imposes an overhead management load on the network resources such as bandwidth and processing capability, due to establishment and maintenance of tunnels. This management overhead degrades the network QoS. CenterTrack can determine the ingress edge router being attacked, but unfortunately, it is incapable of identifying the path only belongs to its domain [11]. Therefore, it is not very scalable as the local domain increases. Scalability in the network is important for QoS, otherwise, QoS method would only work for certain situations.

A malicious user may send spoofed packets in order to attack the defense of a node, or can inject attacking traffic to the overlay network. To avoid such problems, another overlay network called Secure Overlay Service (SOS) was proposed. The SOS architecture utilize intensive filtering and anonymity (“secret servlets”) [52], [55]. A secret servlet node computes keys for well-known hash functions based on the site’s IP address. One example is Secure Overlay Access Point (SOAP), which is a delegated router allowing a packet to enter the overlay only after its identity is validated [81], [88]. Hence, the packets must be authenticated and authorized by the SOS architecture before traffic is allowed to traverse through the overlay network to the end user. SOS therefore uses a combination of secure overlay tunneling, routing via

\(^2\)Ingress: is the entry point (label edge router) into the overlay network.

\(^3\)Egress: is the exit point (label edge router) into a network.
consistent hashing, and filtering. However, the insensitive filtering introduces latency and increases the traffic delays [52]. The exchange of keys consumes time and power which affects many QoS parameters such as delay and throughput.

I. Network Forensics

The use of Network Forensics against DDoS attacks is briefly summarized in this section. Network forensics monitor network traffic and determine if an anomaly exists in the traffic, and if so, whether the anomaly can be an attack. With the use of custom-made sniffers and scripts, skilled programmers can design a technique to identify the DDoS attacks. An example of the use of forensics to other areas includes load balancing and throttling [88]. By applying load balancing, ISP providers can increase the network bandwidth of QoS connections which and can prevent QoS from being affected by a DDoS attack. However, the drawback of such an approach is that QoS will be degraded if the attack occurs before the maximum bandwidth is achieved. Throttling is a method to control the bandwidth that a network application can use. During throttling, incoming traffic is adjusted to acceptable levels that can be managed to the server. This technique may be utilized to mitigate DDoS attacks. Nonetheless, currently available technologies cannot always distinguish between legitimate traffic from malicious traffic correctly. Further studies are needed to investigate the impact of a practical implementation of this method.

Other methods which have been applied in forensic analysis include traceback, event logs, packet sniffers, filtering, and firewalls [82], [91]. We have already discussed various challenges associated with these approaches in the previous sections.

V. QoS Protocol Vulnerabilities Prone to DoS/DDoS Attacks

Multimedia support over networks has been extensively studied over the last decade. Several QoS approaches and protocols have been proposed and implemented. However, very few of these proposed approaches have been incorporated into commercial products. Consequently, these techniques have not been widely deployed. However, there are some traffic engineering technologies that have gained wide acceptance namely, Multi-Protocol Label Switching (MPLS), Differentiated Services (DiffServ) and Resource Reservation Protocol (RSVP). In this section we discuss possible DoS attacks that can exploit some of the vulnerabilities of these QoS protocols.

A. MPLS

To support QoS over IP networks, Traffic Engineering (TE) has introduced Multi-Protocol Label Switching (MPLS). MPLS is being widely deployed in the Internet backbone [53]. In MPLS, the packet forwarding process is done by means of label swapping. Since labels are short and have fixed length, MPLS can achieve high efficiency compared to conventional IP routing where the longest prefix matching is used [53]. Similar to IP spoofing attacks, where an attacker fakes the source IP address of a packet, it is also theoretically possible to spoof the label of an MPLS packet [42], [89], assuming the MPLS core network may be trusted. However, one should not assume that the core MPLS network will provide a level of security and intrusion invisibility [42]. As the core network is part of the end-to-end QoS path, protection against an insecure core is also required. An example of the vulnerability of MPLS to DoS attacks was revealed by Cisco MPLS router [43].

It is important to prevent theft of services from an Internet Service Provider (ISP). Therefore, a cryptographic protocol to protect MPLS header needs to be investigated. For the usage of cryptographic protocol, it is necessary to ensure that the protocol will not interfere with the QoS provided by an ISP on MPLS network. For example, addition of 128 bits to each MPLS header will result in the header becoming four times larger. This would add more processing delay for each packet. Furthermore, the key exchange algorithm also needs to be aware of the computation burden it imposes since this also affects QoS [42].

Standard secure protocols (e.g. IPSec), operate at layer 3 (compared to MPLS which operates at layer 2). Many of the MPLS network designs require the exchange of labeled packets. This creates opportunities for a third party to introduce labeled packets, which, if correctly crafted, might be associated with certain Virtual Private Networks (VPNs) on MPLS networks and could effectively introduce false packets into a VPN [42]. The design of a Security Label Distribution Protocol (LDP) is an open issue for future study [42]. In particular, is defense of privacy or authentication against DoS attacks require further investigation.

Various attacks can also be launched on layer 2 protocols technologies (for instance, Address Resolution Protocol (ARP) attacks). Since LDP uses the ARP to establish a Label Switched Path (LSP), ARP spoofing allows an attacker to redirect traffic between two routers through the device of the attacker. As a result, the attacker gains access to all packets between the two routers [42]. When many Label Switching Routers (LSRs) share a common layer 2 network, a third party can inject packets into such networks. The label forwarding method used in MPLS is affected by the vulnerability of MPLS to DoS attacks such as label spoofing, traffic insertion, and disabling of the ingress. This, in turn, results in the degradation of MPLS labels fast forwarding and ultimately degrades the network QoS.

B. ReSource ReserVation Protocol (RSVP)

ReSource ReserVation Protocol (RSVP) [50] was engineered to accomplish the creation and maintenance of distributed reservation states across a set of multicast or unicast delivery paths. The RSVP method allows the reservation of bandwidth for a path from the sender to the receiver that meets the traffic QoS requirements. To establish an explicit route computed by constraint based routing, as well as to reserve resources along that route for QoS, some kind of path set up signaling is needed. RSVP communicates with two basic types of messages, PATH and RESV (Figure 12). PATH
messages flow from a sender to the receiver. Upon receipt of the PATH message from the sender, the receiver sends an RESV message in return and reserves network resource for the path. Maintaining an RSVP path requires the retransmission of refresh messages. The establishment and maintenance of an RSVP path use IP datagrams to transport messages between peers, which is based on UDP instead of TCP. Without TCP, a peer must handle the loss of distribution messages by itself. RSVP is a signaling mechanisms, which does not contain methods to verify that request to conforms to resource allocation policies (which makes RSVP vulnerable to DoS attacks) [48], [85] before admitting a flow and allocating the resources contained in the RSVP message.

Figure 12: Basic Resource Reservation Protocol (RSVP) Model.

Path message holds explicit route <A,B,RECEIVER>
Generate Reservation message for required traffic/QoS constraint
Per hop reservation and refresh message unless suppressed

For authorization and identification of the sender or receiver, one might consider the use of public key cryptography. The use of public-key-based authentication techniques is, however, still considered to be too heavy-weight (computationally and from a bandwidth perspective) to be used for per-flow signaling [49]. It also requires a substantial amount of processing and memory requirement which can deplete resources available for QoS traffic.

The attacker can potentially intercept or drop all or some of the reservation messages (such that the QoS reservation ones) and channel establishments can be maliciously delayed in a persistent way [49]. When RSVP is communicating with PATH (a PATH message is sent from the sender along the data path and stores the path state in each node along the path) and a RESV (a RESV message is sent from the receiver to the sender along the reverse data path), an attacker can prevent RESV messages from being sent. As a result, PATH messages will be sent again, thereby creating additional traffic on the network which will likely affect QoS of ongoing traffic.

C. Differentiated Services (DiffServ)

Differentiated Services (DiffServ), networks provide QoS guarantees by policing traffic using a fixed number of pre-existing classes. DiffServ is a scalable method for providing QoS to multimedia traffic. DiffServ supports three different Per Hop Behavior (PHB) as it transport traffic. These PHB traffic flows are Expedited Forwarding (EF), Assured Forwarding (AF), and Best Effort (BE). There are methods which have been explored to defend against DoS attacks while maintaining QoS by using DiffServ [45]–[47], [66], [83].

Another method proposed is differential packet filtering [45]. In this method, traffic conditions for each incoming packet are considered safe, normal, or risky according to a continuously updated history table of trusted IP sources. When an attack occurs, some of the risky IP packets are dropped. Next, some of the safe and normal IP packets will be downgraded or discarded. The differential packet filtering is a probabilistic method used to determine risky packets. Unfortunately, it is possible that some of the legitimate traffic would also be dropped. Such packet loss is unacceptable for QoS traffic. This method also relies on history tables which further intensify the amount of processing and storage overhead required to save history tables.

DiffServ uses the Type of Service (ToS) field in IPv4 and the Traffic Type field in IPv6 to identify the QoS of traffic. A malformed packet may set the optional field to some fake QoS format. As a result DiffServ will exhaust processing to determine whether it is a malformed packet or an actual QoS packet [64]. DiffServ also identifies traffic based on IP addresses. Therefore it is susceptible to another type of DoS attacks, called IP address attack. In an IP address attack, the packet has the same source and destination IP address [64]. As a result, it confuses the network/operating system resulting in a network or user operating system to crash [64] and it in turn will affect the QoS of network services. Generally, IP addresses of QoS packets are not protected with encryption because of the high processing overheads associated with encryption/decryption, and the exchange of security keys also consumes time and power which will in turn affect QoS.

[46], [47] proposed resource isolation for TCP ACK, ICMP, TCP-SYN messages. When a packet arrives, it determines the type of traffic (BE, EF or AF). Next, the packet is divided into an UDP aggregate and a TCP aggregate (Figure 13). The TCP message carries data or control segments (ICMP, ACKs, SYN, etc). Traffic flows are isolated and the core is not affected by connection attacks (for example TCP-SYN attacks). However, the problem with this approach is that once the connection is attacked, it will prevent other connections from being set up. In addition, this method does not address the issue when the attack is on QoS traffic flows in AF or EF classes.

Figure 13: Proposed DoS with DiffServ.
VI. Future work

The task of mitigating DoS and DDoS attacks and the configuration of QoS requirements on networks should be transparent to end-users. This is a highly desirable goal, but much work remains to make it a reality. As pointed out in this paper, several issues involved in supporting seamless security with end-to-end QoS support need to be addressed. Among those issues include:

**Authentication:** Data integrity, confidentiality, privacy, and end-to-end security are important to prevent routing information from being compromised. Authentication will affect the QoS performance. Thus, it needs to be simple enough not to affect QoS traffic, but advanced and sophisticated enough to protect against DoS/DDoS attacks.

**Logging:** Whether it is used for statistical analysis, traceback, or filtering, considerations on how many logging tables and how much information need to be addressed because they affect the amount of storage required to store such information. Since network storage resources are limited, smaller logging tables are required. Having log tables record information make them also prone to attacks from intruders.

Complex computations (whether statistical analysis or key exchange) consume time and power to validate the identity of the user. This issue needs to be addressed in future works. If unnecessary computations are performed, QoS will be affected because resources are used instead to perform such tasks. Any new security method should not increase the size of original packet nor have complex control message exchange.

Changing the size of packet, which could result in the packet being fragmented, or having complex signaling protocols, which takes time to process, will cause greater delays for QoS traffic due to the decrease in bandwidth because of the additional traffic load and processing overheads required. While the network is congested, control messages sent from the defensive agents may be dropped or lost. This presents a problem as control messages have to be arrived safely for most QoS traffic.

**MPLS:** As mentioned in the previous section, we should not assume that the core MPLS network will provide a high level of security and intrusion invisibility \(^4\) [42]. Since the core network is part of the end-to-end QoS path, protection against an insecure core is required. The use of cryptographic protocol with MPLS and current secure protocols (such as IPSec) are particularly important areas of focus in the future.

Cryptographic protocols are used to prevent theft of services. It is necessary to ensure that such protocol will not interfere with the QoS on an MPLS network.

Standard secure protocols (such as IPSec) operate at layer 3 and not at layer 2 as MPLS does. Thus, MPLS architectures require the exchange of labeled packets information whether it is forwarding a label or establishing a LSP. This may enable a third party to introduce labeled packets, which, if correctly crafted, could effectively introduce false packets into a VPN [42]. A secure Label Distribution Protocol (LDP) remains an open issue for future investigations [42].

**Traceback:** To accomplish IP traceback, we need to reach the host where the attack originated. It is not easy to trace packets through firewalls because the last-traced IP address might be the firewall’s address. Another drawback to the usage of traceback systems is that generally traceback methods require changing the network, as well as adding router functions and changing packets. Guidelines for dealing with tracing back an attack packet through various networks are needed because it infringes on the privacy of those customers paying for QoS services as well as those not paying for QoS. To promote traceback methods with QoS support, their effects on QoS traffic should be taken into account. In addition, further consideration need to be given to the use of information about an attacker’s source identified by IP traceback.

**DiffServ:** DiffServ provides QoS guarantees by policing the traffic as it enters the network. This is done by classifying and conditioning traffic to conform to a specific behavior aggregate based on the Service Level Agreement (SLA) between the source and the DiffServ provider into a fixed number of pre-existing classes. An attacker may be able to obtain better service by modifying the DiffServ field or by injecting packets to the DiffServ classes. A DoS attack will occur when the modified or injected traffic depletes the resources available for forwarding packets. Hence, ingress nodes are the primary line of defense against DoS attacks. Ingress nodes must condition all inbound traffic to ensure that the DiffServ codepoints are acceptable. Packets found to have unacceptable codepoints must be either discarded or must have their DiffServ codepoints modified to acceptable values before being forwarded. An important instance of an ingress node is that any traffic originating node in a DiffServ domain is the ingress node for that traffic, and must ensure that all traffic coming from it carries acceptable DiffServ codepoints. The proposed countermeasure should be combined with other security protocols if both QoS guarantee and security assurance are required. This technique allows the DiffServ network to distinguish valid traffic from malicious traffic.

**RSVP:** RSVP is a signaling mechanisms, which does not contain methods to verify the request to conform to resource allocation policies [48], [85] before admitting the flow and allocating the resources contained in the RSVP. This, as a result, makes RSVP vulnerable to DoS attacks. For authorization and identification of the sender or receiver, one might consider the use of public key cryptography. The use of public-key-based authentication techniques is, however, still considered to be too heavy-weight (computationally and from a bandwidth perspective) to be used for per-flow signaling [49]. It requires a significant amount of processing and memory requirement and will result in fragmentation thereby allowing DoS attacks. Trade-offs between performance and security has been a recurrent theme. As signaling messages are transmitted at a low rate, the protection of these messages is usually not a problem. For an RSVP router, even a high volume of messages

\(^4\)Similar to IP spoofing attacks, where an attacker fakes the source IP address of a packet, it is also theoretically possible to spoof the label of an MPLS packet.
does not cause performance problems due to the efficiency of the keyed message digest routine [49]. It is dynamic key manage-
ment that is computationally more demanding, especially when it comes to scalability. Since RSVP does not state a particular key exchange protocol, it is hard to approximate the effort needed to create the required security associations. Furthermore, the number of key exchanges to be triggered depends on security policy issues and the authentication mode used by the key exchange protocol. Hence the greater the security applied to RSVP, the greater is the impact on the QoS performance delivered by RSVP. Thus, the security method used in conjunction with RSVP needs to be simple enough so as not to affect the QoS traffic, but robust as well to provide protection against possible DoS attacks.

VII. Conclusion

There are many kinds of defensive measures that have been proposed to reduce the possibility of DoS/DDoS at-
tacks. However, these security techniques have given little consideration to the performance impact they have on network QoS. We provided an overview of several defense mech-
nisms against DoS/DDoS and we evaluated their impact on the QoS delivered by the network. In addition, we also discussed and presented vulnerabilities associated with popular QoS protocols that have been proposed (some of which are already widely deployed in current networks) that are prone to DoS/DDoS attacks. Each of the DoS/DDoS defense mechanisms described in this work has its own limitations when it comes to their impact on network QoS. DoS/DDoS attacks can seriously affect QoS traffic, thereby breaking the Service Level Agreement (SLA) between the client and the Internet Service Provider (ISP). Hence, DoS/DDoS attacks are major threats to network QoS. It is our hope that the results of this work will motivate researchers and designers to investigate novel, comprehensive DoS/DDoS solutions that not only protect networks and systems but are also able to be deployed in actual networking environments with negligible impact on network QoS. Such solutions will lead to adaptable, secure network systems that can also efficiently support end-to-end QoS.

Acknowledgements

We thank the anonymous reviewers for their constructive suggestions that helped us to improve the quality of the paper. Scott Fowler was partially supported by the Swedish Excellence Center at Linköping - Lund in information Technology (ELLIIIT). Sherali Zeadally was supported by an NSF award (No. 0911969) during this work.

References

[1] CERT® Coordination Center (CERT/CC). CERT® Advisory CA-
1996-21 TCP SYN Flooding and IP Spoofing Attacks. [ONLINE] 
sion against Denial of Service Attacks. IEEE International Symposium 
and Keromytis A. Efficient, DoS-resistant, secure key exchange for 
internet protocols. ACM Conference on Computer and Communications 
Security (CCS), November 2002.
Denial of Service Attack Technology. [ONLINE] 
[5] Carnegie Mellon’s CERT® Coordination Center. Types of In-
truder Attacks. [ONLINE] http://www.cert.org/present/cert-overview-
ties and Attack Methods. [ONLINE] http://www.cert.org/present/cert-
[7] Carnegie Mellon’s CERT® Coordination Center. 2007 E-
CrimeWatch Survey™ Summary of Findings. [ONLINE] 
[9] Ferguson P and Senie D. Network Ingress Filtering: Defeating De-
nial of Service Attacks which employ IP Source Address Spoofing. 
route optimization in mobile IP over wireless LAN. IEEE Semiannual 
131.
for IP Traceback. IEEE/ACM Transactions on Networking, June 2001; 
http://www.cs.columbia.edu/~smb/papers/draft-ietf-itrace-04.txt, Febru-
ary 2003.
[17] Lee M and Fung C-K. A denial-of-service resistant public-key authen-
tication and key establishment protocol. IEEE International Perfor-
171–178.
[18] Choi S. Denial-of-Service Resistant Multicast Authentication Protocol 
with Prediction Hashing and One-way Key Chain. IEEE International 
[19] Pietro R, Durante A and Mancini L. A Reliable Key Authentication 
Schema for Secure Multicast Communications. International Symposi-
um on Reliable Distributed Systems, October 2003, pp. 231–240.
[20] Saxena A and Soh B. Distributed Denial of Service Attacks and Anony-
mous Group Authentication on the Internet. International Conference 
460–464.
Design and analysis of a replicated elusive server scheme for mitigating 
denial of service attacks. Elsevier Journal of Systems and Software 
Roaming Honeypots for Mitigating Service-Level Denial-of-Service 
Attacks. IEEE International Conference on Distributed Computing 
International Conference on Network Protocols (ICNP), September 
approaches to DDoS attack detection and response. IEEE DARPA 
Information Survivability Conference and Exposition, April 2003, pp. 
303–314.
Defense Against Spoofed DoS Traffic. ACM International Conference
[77] ITU-T Recommendation X.1161. Framework for secure peer-to-peer communications. to be published.


