DDoS and IP Traceback

Dr. Arjan Durresi
Louisiana State University, Baton Rouge, LA
70803
durresi@csc.lsu.edu

Overview

- Distributed Denial of Service (DDoS)
- Proposed solutions
- Autonomous System (AS) based solution
- Conclusions
Security

- People can justifiably rely on computer-based systems to perform critical functions
  - national scale infrastructures: water, power, communication, transportation, ...
  - localized systems: cars, homes, workplaces, ...
- People can justifiably rely on systems processing sensitive information about them to conform to public policy
  - health, banking, libraries, e-commerce, government records, ...
- Without fear of sudden disruption by cyber attacks

Denial Of Service

- The goal of a denial of service attack is to deny legitimate users access to a particular resource.
- An incident is considered an attack if a malicious user intentionally disrupts service to a computer or network resource.
- Resource exhaustion
Resource Exhaustion

- Disk Space
- CPU Cycles
- Memory
- Network Bandwidth
- Application Resources
  - TCP Stack
  - Web Connections

What’s the Harm?

- Financial loss can be difficult to estimate
  - Lost business
  - Bad publicity and damaged reputation
- 2002 CSI/FBI Survey
  - 40% of reported attacks are DOS
  - Average cost per attack is >$1 million
- Distributed DOS attacks (February 2000)
  - Amazon, CNN, E-Trade, eBay, etc...
  - Estimated losses were “several millions to billions of dollars”
- DOS can also be used to cover-up “real” attacks
- Nations critical infrastructure is also at risk
Denial of Service Attacks

- Most involve either resource exhaustion or corruption of the operating system runtime environment.
- UDP bombing
- tcp SYN flooding
- ping of death
- smurf attack

Distributed Denial of Service Attacks (DDoS)

- Attacker logs into Master and signals slaves to launch an attack on a specific target address (victim).
- Slaves then respond by initiating TCP, UDP, ICMP or Smurf attack on victim.
Denial of Service in pervasive networks
- Power-draining attacks
- Bandwidth-usage attacks
- CPU-usage attack
Why are DOS attacks possible?

- IP employs an open architecture
  - No authentication of sender’s IP address
  - Easy to forge any address, hard to detect offender
  - IP traceback, ingress/egress filters (later)
- No resource regulation in the network
  - Employ QOS techniques to mitigate impact (later)

Security Mechanisms

- Normally, not a single silver bullet
- Develop multiple layers of defense
- Employ as many layers of defense as needed - risk, resource profiles
- Castle, moat, drawbridge, mountain-top lookout, perimeter wall, inner wall, ruler decoy etc.
- Firewall, resource managers, app. Level security, logging, antivirus, remote backups, egress filters...
Security Analogy

Two Security Philosophies

Super Protection – very expensive, could be broken

Prevention Power of punishment
**DOS attacks**

- All DOS attacks consume resources
  - Bandwidth in UDP floods
  - Processing power in CGI bin attacks
  - Memory in fragmentation attacks
- Can we detect and contain attacks if we kept good accounting of resources?

**Resource Accounting**

- Monitor network bandwidth, processor time and memory usage per process at server
- Regulate processes exceeding preset thresholds
- Problems: Hard to identify the process to whom resource usage needs to be charged
  - Interrupts, context-switches
  - A packet arrives at network interface
### Ingress filtering

- Campus
- Internet
- Victims
- Core router
- Ingress router
- Egress router

**DOS attacks**

- Ingress filtering is not widely employed
  - Can be expensive in transit and backbone networks
- How to effectively trace back the source of the attack?
- If successful, may be able to throttle attack traffic at the network ingress
**ICMP traceback (Bellovin, IETF)**

- Generate ICMP packets with packet header, router and its neighbors ids
- Do this with low probability 1/20,000
- These ICMP packets can be used to trace the source
- More likely to get packets from routers closer to destination, rather than source

**IP Traceback**

What is IP Traceback?
Getting back to the attacker, by identifying the attack path.
**IP Traceback**

- Probabilistic Packet Marking (PPM)
  - No of attack packets required is 1000s
  - Difficult to handle DDoS attacks (too complex to construct attack path).
- ICMP Traceback or iTrace - Overhead
- Controlled Flooding - a form of DoS itself
- Hash-Based IP traceback
  - Less space needed and No eavesdropping
- IP Traceback with IPSec
  - Poor scalability
  - ISP need to update topology to all end users
  - End users need to know network topology

- Not practical to assume that all routers in the Internet will participate in marking scheme
- When some routers don’t participate in marking, not sure if the last router in the constructed path is the true origin
- To be protected against single attacker that insert false information into the path the marking probability should be more than 0.5
  - Very high number (thousands) of packet to be analyzed by the victim
IP traceback
(Savage...Sigcomm00)

- Exact Traceback
  - R_6, R_3, R_2, R_1
- Approximate Traceback
  - Valid path suffix
  - R_5, R_6, R_3, R_2, R_1

- Attacker can generate any packet
- Attackers may conspire
- Aware of the tracing mechanism
- Attackers send lots of packets
- Packets may be lost, reordered
- Routes are pretty stable
- Routers are memory, CPU limited
**IP traceback - Node Append**

- Attach each router’s IP address to the packet
  - Like IP record route option
- Every packet will have path info
- Too expensive
- Could lead to fragmentation problems

**Node Sampling**

- Reserve a node field
- Routers write their IP address with probability $p$
- Prob. Of receiving id from $d$ hops
  - $p(1-p)^{d-1}$
- $p > 0.5$, robust against attacker spoofing
- Routers far away from victim don’t send many packets
  - $d=15$, $p=0.51$, expectation = 42,000 packets
Edge Sampling

- Encode edges of path
  - Rather than single nodes
- Employ three fields
  - Start, end, distance
- With probability p, write Router IP address in start, make distance = 0
- Else, (a) if start already marked, distance=0, put your id in end and
  - (b) increment distance

- Tree construction starting from victim (distance =0, 1,...)
- Time for convergence
  - furthest router: \( p(1-p)^{d-1} \)
- Can use any p, spoofed attacker packets distance field longer
- Robust against multiple attackers
  - Edges are different, linear complexity
- Takes many bits \(-32+32+8? = 72\)
Edge Sampling --encoding

- Use XOR of addresses
- R1, 0
- R1 XOR R2, 1
- R1 XOR R2 XOR R3, 2
- Uses roughly half the space

---

Edge Sampling—Fragment Sampling

- Address
- Hash(Address)
- BitInterleave
- Send k fragments into network
Fragment Sampling

- Can compress information into 16 bits
- Use IP fragment identifier space
- Expensive to compute
- Nor robust against large DDOS
Advanced Marking Scheme
Song & Perrig, Infocom01

AMS

- Use two hash functions \( h \) and \( h' \)
- Encode \( h(\text{start}) \) XOR \( h'(\text{end}) \)
- Use 11-bits for hash, 5bits for length
- If you know upstream routers, few choices for \( h(s) \), when we know \( h'(e) \)
- Tolerate multiple attackers
  - Upto 60
  - Main limitation: hash collisions

Figure 2: Encoding in Advanced Marking Scheme I
Use two sets of hash functions

Main intuition:
- Probability of collision with 11 bits $1/2^{11}$
- Probability of collision with $m$ hashes of 11 bits $= 1/(2^{11})^m$
- Multiple hash functions reduce Collisions

Where did we see that before?
AMS-II

- Tries to work within the space of 11 bits
  - While identifying the hash function
- Easier than FSM
- Much more robust than FSM

FMS False positives
AMS & AMS-II

Figure 7: False Positives for Advanced Marking Scheme

Figure 8: False Positives for Advanced Marking Scheme II

FMS Path reconstruction time

Reconstruction Time (Seconds)

Number of Attackers

Fragment Marking Scheme
Traceback is an interesting idea
- Allows us to trace the origin of the attack
- Threat of Identification leads to reduction in attacks
- What about the viruses?
  - Innocent attackers
**Autonomous System - Traceback**

Autonomous System Boarder Router

**Autonomous Systems - AS**

- AS is a group of IP networks managed by one network operator
- AS - set of routers using the same external routing policy
- Number of AS - 14,000, number of hosts - 200M
- In 99.5% of cases, a packet passes less than AS before reaching destination
- Network Operators may not always like to disclose their network details
- AS number is 16 bits compared to IP address 32 bits (IPv6 - 128 bits)
Autonomous System Marking

- Marking by ASBR
- Marking scheme similar to node sampling scheme
- 16 bits for ASN and 3 bits for AS_distance
- A packet is marked only if it leaves the AS
- A packet is marked with a probability \( p \) and the distance is set to zero
- If the ASBR chose not to mark, it increments the distance field

Marking procedure at router \( R \) with AS Number \( R_{AS} \):

for each packet \( w \)

let \( x \) be a random number from \([0, 1)\)
if \( x < p \) then,
- write \( R_{AS} \) into \( w\.AS \)
- set \( w\.AS\_distance = 0 \)
else
increment \( w\.AS\_distance \)
Number of Marked Packets

- If $p = \frac{1}{d_{AS}}$ => 25 packets needed
- $d_{AS} = 7, p=0.51$ => 141 packets needed

Authenticated Marking Scheme

- We assume the presence of a symmetric key infrastructure within each AS
- Each ASBR that belongs to the AS or connected to the AS know the secret key $K_i$
- Use one-way hash chains to generate session keys
  - $h_0, h_1, \ldots, h_n$ where $h_i = H(h_{i-1})$
  - Initially distribute $h_0$
  - Each ASBR computes the chain
  - Use the keys starting from the right to left
Authenticated AS Marking Algorithm

Marking procedure at router R with ASN $R_{AS}$:

- $K_{AS}$ is the symmetric key of $R_{AS}$
- $K'_{AS}$ is the symmetric key of the next AS in the path.

for each packet $w$
  
  Compute $D$(AS Marking, $K_{AS}$)
  
  if (Redundancy Predicate is not fulfilled )
    
    Set AS Marking to $E($ASN || RP $,$ $K'_{AS})$
  
  else
    
    let $x$ be a random number from $[0, 1)$
    
    if $x < p$ then,
      
      Set AS marking to $E($ASN || RP $,$ $K'_{AS})$
    
    else
      
      Set AS marking to $E$(AS Marking, $K'_{AS}$)

Authenticated AS Traceback

- Victim obtains the AS symmetric key of the current session and computes AS marking
- Victim can reconstruct the path
- Victim can use the symmetric key to compute the keys of previous sessions but not any future sessions
  - A compromised victim doesn't affect the security of the mechanism
Summary

- Presented two schemes:
  - Autonomous System based Traceback
  - Authenticated Marking Scheme
- Only ASBR participate in marking
- Low marking overhead
- Enables to reconstruct the AS attack graph in real time
- Authenticated scheme prevents compromised routers from forging ASBR marking

Thank You!