CSc 466/566

Computer Security

24 : Network Security — Spoofing

Version: 2019/11/25 11:30:22

Department of Computer Science
University of Arizona

collberg@gmail.com

Christian Collberg

Outline

1. Introduction
2. The Link Layer
   - Hubs and Switches
   - Ethernet Frames
   - ARP Spoofing
3. The Network Layer
   - ICPM
   - IP Spoofing
4. The Transport Layer
   - TCP Session Hijacking
5. Summary

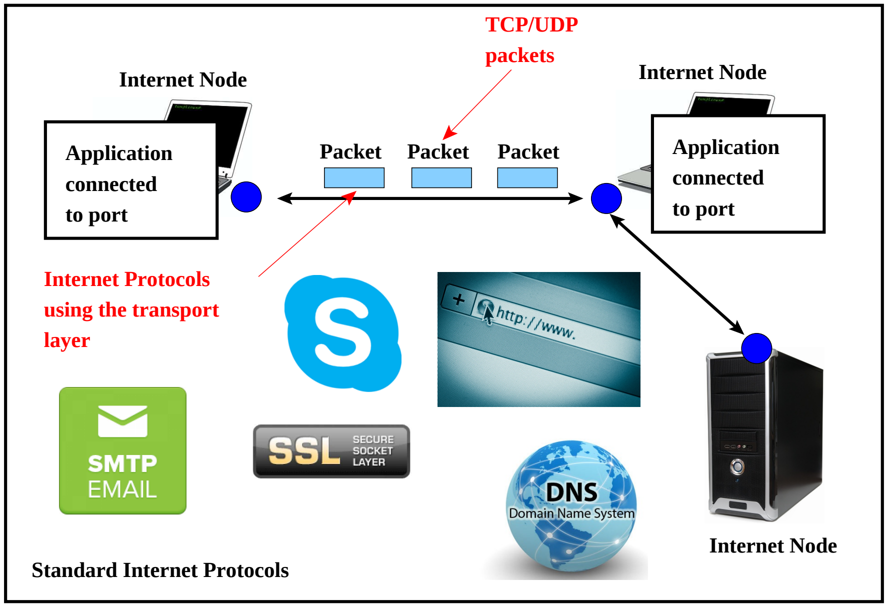
Spoofing Attacks

Definition (Spoofing Attack)

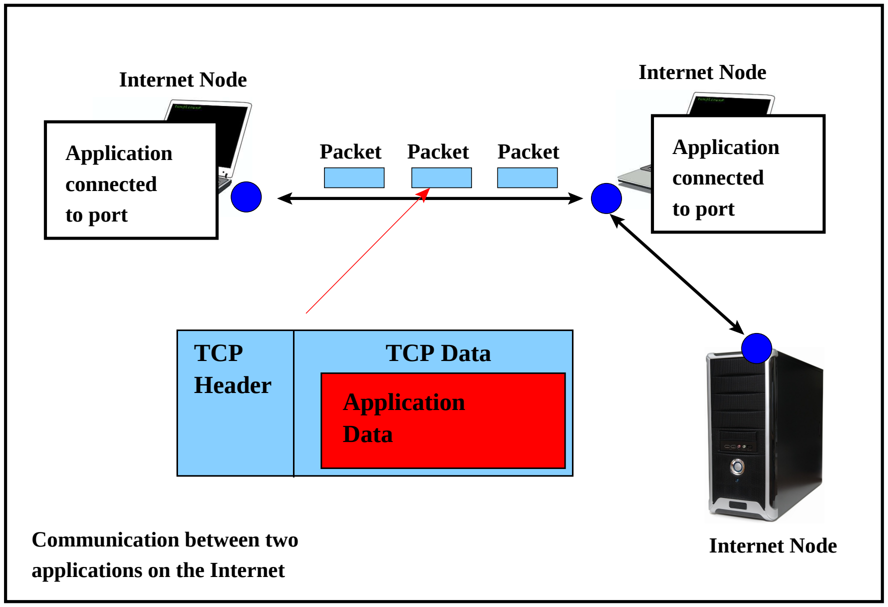
A situation in which a user or program impersonates another device or user on a network.

- Email spoofing: Send a message from a bogus e-mail address or the address of another user.
- IP spoofing: Hiding or faking a computer's IP address.
- . . .

Application Layer



Transport Layer



Internet Node

Application connected to port

Packet Packet Packet

Internet Node

Application connected to port

TCP Header

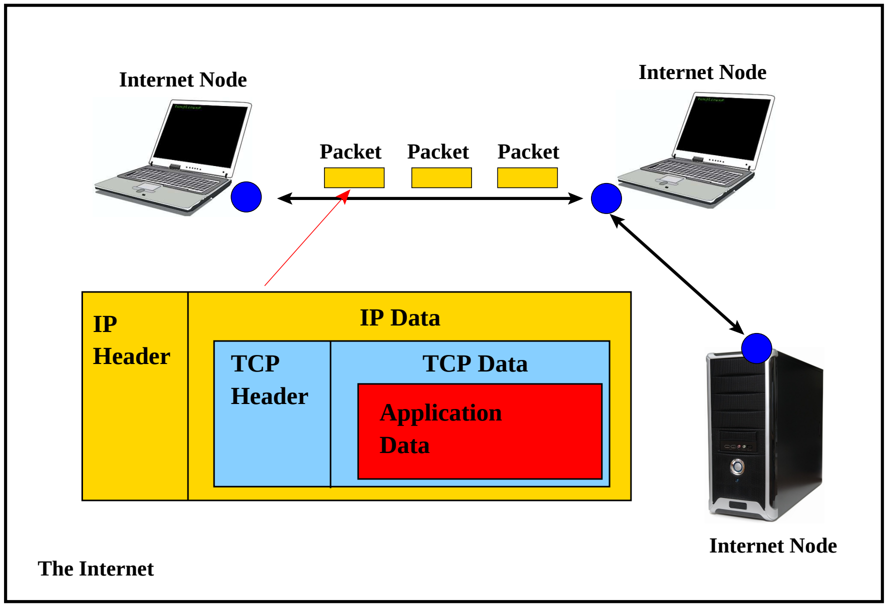
TCP Data

Application Data

Communication between two applications on the Internet

Internet Node

Internet Layer



Internet Node

Packet Packet Packet

Internet Node

IP Header

IP Data

TCP Header

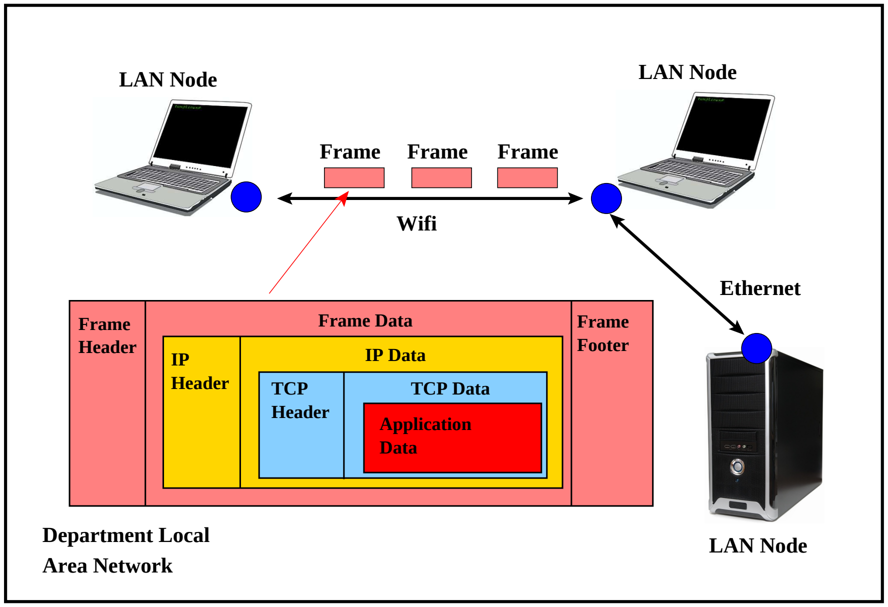
TCP Data

Application Data

The Internet

Internet Node

Link Layer



LAN Node

Frame Frame Frame

Wifi

LAN Node

Ethernet

Frame Header

Frame Data

Frame Footer

IP Header

IP Data

TCP Header

TCP Data

Application Data

Department Local Area Network

LAN Node

Packet Sniffing Animation

- http://williams.comp.ncat.edu/IA_visualization_labs/security_visual_tools/VisTools.html → *Packet sniffer simulator.*
- NOTE: Use Firefox, not Safari.
- Maybe gone?

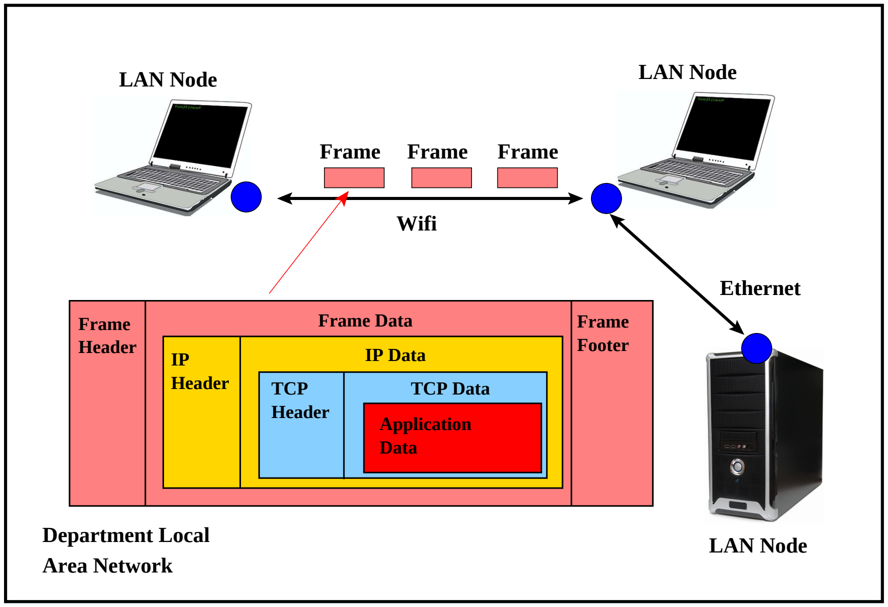
Outline

1. Introduction
2. The Link Layer
   - Hubs and Switches
   - Ethernet Frames
   - ARP Spoofing
3. The Network Layer
   - ICPM
   - IP Spoofing
4. The Transport Layer
   - TCP Session Hijacking
5. Summary

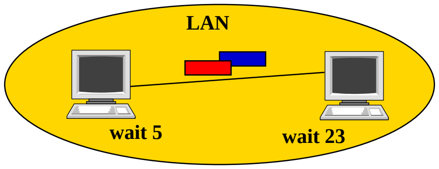
The Link Layer

- The Link Layer sits on top of the physical layer.
- Ethernet — IEEE 802.3.
- Ethernet cables connect computers on a LAN.
- Collision: Two computers on the same network segment send a packet at the same time.
- History of Ethernet: http://www.youtube.com/watch?v=g5MezxMcRmk.

Internet Protocol Layers — Link Layer



Ethernet Collision

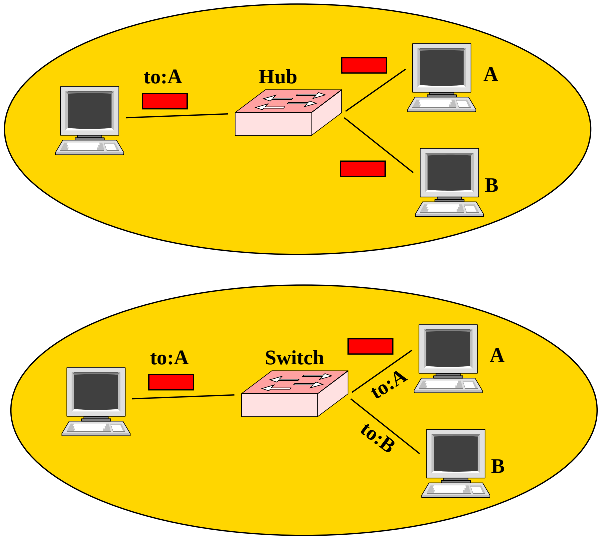


- Collision algorithm:
  1. Each computer waits a random length of time;
  2. Retransmit!
  3. Another collision? Repeat from 1!

Hubs and Switches

- Hubs and Switches connect devices on a LAN.
- Ethernet Hub:
  - Forward all frames to all attached devices.
  - Lots of extra traffic: all frames are duplicated!
  - All devices are on the same network segment, and must do collision avoidance.
- Ethernet Switch:
  - Initially works like a hub.
  - Over time, learns the addresses of attached devices.
  - Eventually, only forwards a frame to the destination device.
  - Fewer collisions.

Ethernet Hub vs Switch



MAC Addresses

- MAC address: 48 bits assigned to network interface.
- Example: `00:A0:B9:14:C8:30`
- MAC structure:

locally assigned (1 bit)	manufacturer (23 bits)	unique number (24 bits)

- Software (Unix: `ifconfig`) can change a device's MAC: *locally assigned*=1.

Ethernet Frame Format

Preamble (7 bytes)
Start-of-Frame delimiter (1 byte)
MAC destination (6 bytes)
MAC source (6 bytes)
Ethertype/length (2 bytes)
Payload (45-1500 bytes)
CRC-32 Checksum (4 bytes)
Interframe Gap (12 bytes)

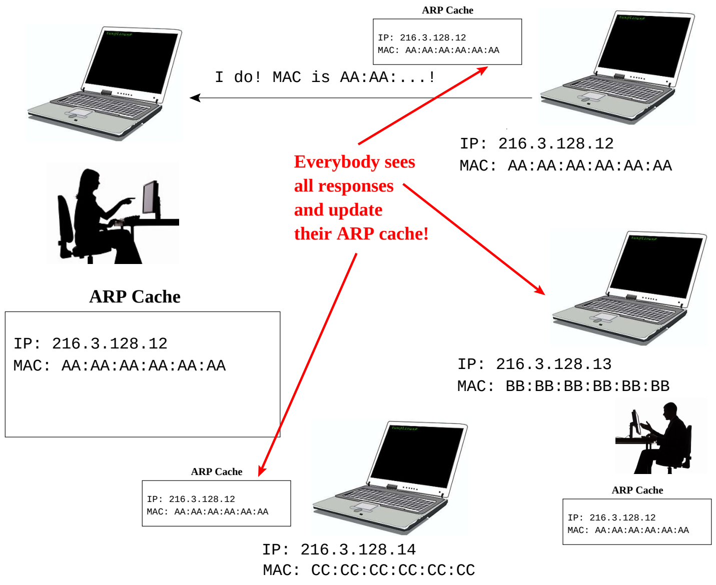
Ethernet Frame Format. . .

- The CRC-32 checksum can catch simple transmission errors.
- Switches learn the location of network devices from the MAC addresses.

Address Resolution Protocol

- Address Resolution Protocol (ARP): Find the MAC address given the IP address.
- Algorithm (Alice wants to know the MAC address of IP address *A*):
  1. Broadcast to all network interfaces: Who has IP address *A*?.
  2. Wait for a response *A* is at MAC address *M*! from the devices with IP address *A*.
  3. Store *A* ↔ *M* in the ARP cache.
- Problem: no authentication.

Address Resolution Protocol. . .



ARP Cache

IP: 216.3.128.12
MAC: AA:AA:AA:AA:AA:AA

I do! MAC is AA:AA:...!

IP: 216.3.128.12
MAC: AA:AA:AA:AA:AA:AA

Everybody sees all responses and update their ARP cache!

ARP Cache

IP: 216.3.128.12
MAC: AA:AA:AA:AA:AA:AA

IP: 216.3.128.13
MAC: BB:BB:BB:BB:BB:BB

ARP Cache

IP: 216.3.128.12
MAC: AA:AA:AA:AA:AA:AA

IP: 216.3.128.14
MAC: CC:CC:CC:CC:CC:CC

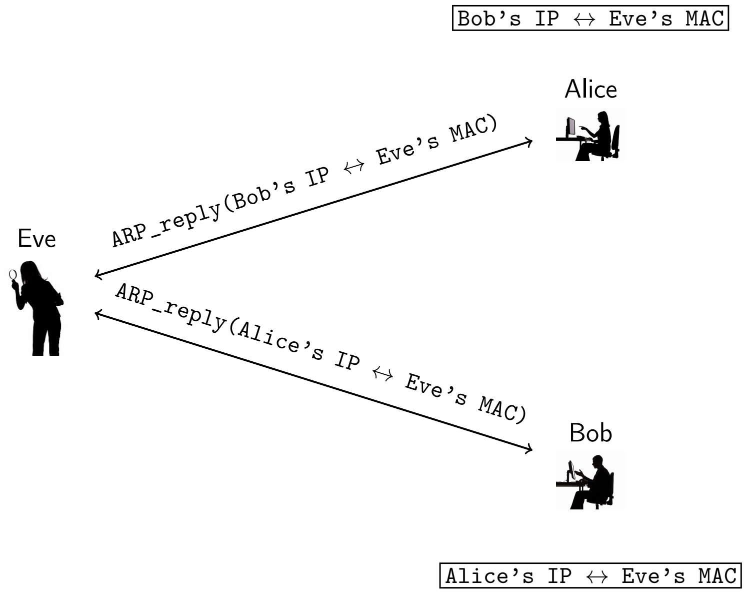
ARP Cache

IP: 216.3.128.12
MAC: AA:AA:AA:AA:AA:AA

ARP Spoofing

- Any computer on the network could claim to have a particular IP address.
- Machines will update their ARP cache whenever they see an ARP reply — even if there was no corresponding ARP request!
- Attack:
  1. Eve sends `ARP_reply(Bob’s IP ↔ Eve’s MAC)` to Alice.
  2. Alice puts `Bob’s IP ↔ Eve’s MAC` in her ARP cache.
  3. Eve sends `ARP_reply(Alice’s IP ↔ Eve’s MAC)` to Bob.
  4. Bob puts `Alice’s IP ↔ Eve’s MAC` in his ARP cache.

ARP Spoofing



Bob's IP ↔ Eve's MAC

Alice

Eve

ARP_reply(Bob's IP ↔ Eve's MAC)

ARP_reply(Alice's IP ↔ Eve's MAC)

Bob

Alice's IP ↔ Eve's MAC

ARP Spoofing...

- After the ARP cache poisoning all traffic between Alice and Bob is routed through Eve:
  1. MITM attack;
  2. Denial of Service attack.

ARP Spoofing — Visualization

- ⇒
  http://williams.comp.ncat.edu/IA_visualization_labs/security_visual_tools/VisTools.html
- *Wireless Network Attacks → ARP Cache Poisoning.*
- Maybe gone?

1. Bob makes an ARP request to Alice.
2. Alice sends `ARP-RESPONSE(Alice`$_{\mathtt{IP}}$ ↔ `Bob`$_{\mathtt{MAC}}$`)` to Bob.
3. John sends forged `ARP-RESPONSE(Bob`$_{\mathtt{IP}}$ ↔ `John`$_{\mathtt{MAC}}$`)` to Alice.
4. Alice adds `Bob`$_{\mathtt{IP}}$ ↔ `John`$_{\mathtt{MAC}}$ to her ARP cache.
5. Alice sends a message to Bob, but it gets sent to John.

Cain & Abel ARP Attack Tool

> *Cain & Abel is a password recovery tool for Microsoft Operating Systems. It allows easy recovery of various kind of passwords by sniffing the network, cracking encrypted passwords using Dictionary, Brute-Force and Cryptanalysis attacks, recording VoIP conversations, decoding scrambled passwords, recovering wireless network keys, revealing password boxes, uncovering cached passwords and analyzing routing protocols.*
>
> *The latest version is faster and contains a lot of new features like APR (Arp Poison Routing) which enables sniffing on switched LANs and Man-in-the-Middle attacks.*

http://www.oxid.it/cain.html

ARP Spoofing — Countermeasures

1. Restrict LAN access to trusted users.
2. Check for multiple occurrences of the same MAC address on the LAN.
3. Static ARP tables: the system adminstrator manually sets up the routers' ARP caches.
4. Inspect all ARP packets, detecting attempted spoofing.

Exercise I

1. The *Address Resolution Protocol (ARP)* finds the
   ✎

   address given an
   ✎

   address.
2. *Static ARP tables* (i.e. the system adminstrator manually sets up the routers' ARP caches) is a defense against ARP spoofing attacks (TRUE/FALSE):
   ✎

Exercise II

Alice, Bob, and Eve have the following IP and MAC addresses:

Who	MAC	IP
Alice	0x40	150.135.68.100
Bob	0x41	150.135.68.101
Eve	0x42	150.135.68.102

1. *Explain* how Eve can use ARP Spoofing (AKA ARP Cache Poisoning) to launch a Man-in-the-Middle attack against Alice and Bob.
2. *Draw a figure* showing the flow of information on the network during the attack, and the content of each person's ARP Cache at the end of the attack.

Exercise II: Explain

✎

Exercise II: Draw

✎

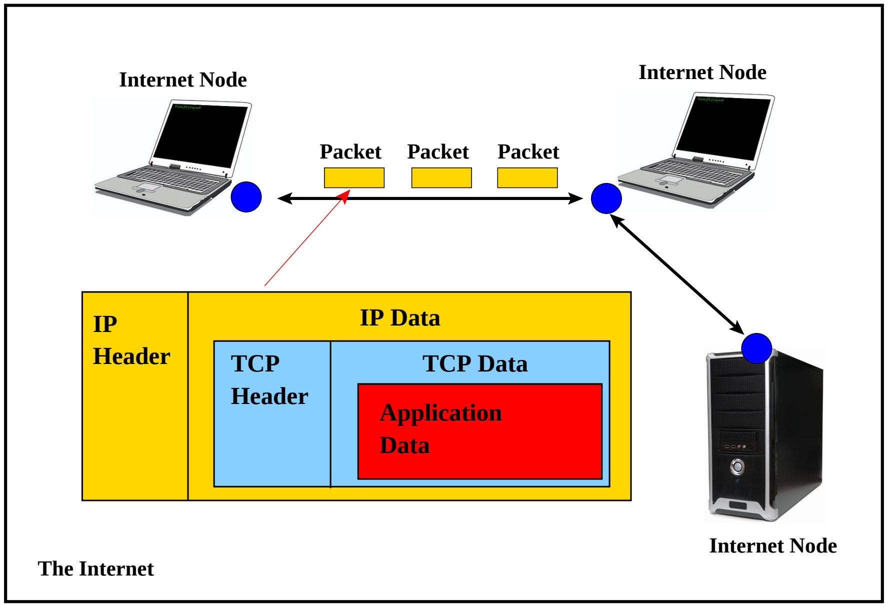
Outline

1. Introduction
2. The Link Layer
   - Hubs and Switches
   - Ethernet Frames
   - ARP Spoofing
3. The Network Layer
   - ICPM
   - IP Spoofing
4. The Transport Layer
   - TCP Session Hijacking
5. Summary

The Network (Internet) Layer

- Best effort routing of packets between any two hosts on the Internet.
- Abstraction:
  1. Source/Destination: Internet nodes
  2. Data: IP packets
  3. Addressing: Internet Protocol (IP) addresses.
- IPv4 — 32-bit addresses, IPv6 — 128-bit addresses.
- No guarantees a packet will be delivered.

Internet Protocol Layers — Internet Layer



IPv4 Packet Format

Version (4 bits)

Header length (4 bits)

Service type (8 bits)

Total length (16 bits)

Identification (16 bits)

Flags (3 bits)

Fragment offset (13 bits)

Time-to-Live (8 bits)

Protocol (8 bits)

Header Checksum (16 bits)

Source Address (32 bits)

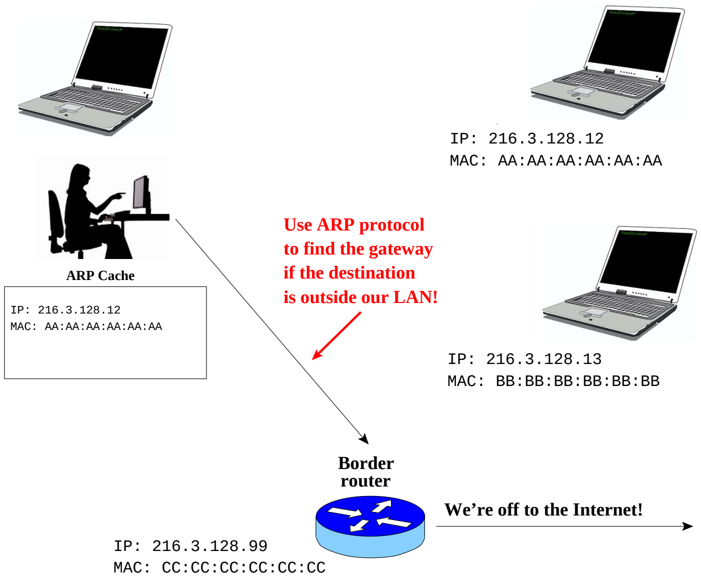
Destination Address (32 bits)

Payload

Routing Algorithm — From a Host Node

- Sending a packet P from a host node N:
  1. If P's destination is on this LAN:
     - Use the ARP protocol to find the MAC address,
     - deliver directly.
  2. Otherwise:
     - use the ARP protocol to find the MAC address of the gateway,
     - forward.

Routing — From a Host Node . . .



ARP Cache

```
IP: 216.3.128.12
MAC: AA:AA:AA:AA:AA:AA
```

```
IP: 216.3.128.12
MAC: AA:AA:AA:AA:AA:AA
```

```
IP: 216.3.128.13
MAC: BB:BB:BB:BB:BB:BB
```

Use ARP protocol to find the gateway if the destination is outside our LAN!

Border router

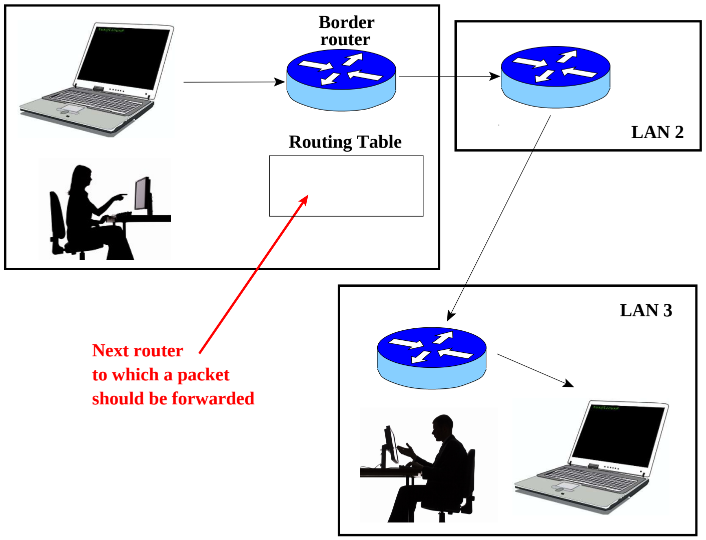
```
IP: 216.3.128.99
MAC: CC:CC:CC:CC:CC:CC
```

We're off to the Internet!

Routing Algorithm — From a Router

- Router — gateways and other network nodes that handle routing of packages on the Internet.
- A router typically connects two or more LANs.
- Routing tables describe the next router to which a packet should be forwarded.

Routing — From a Router

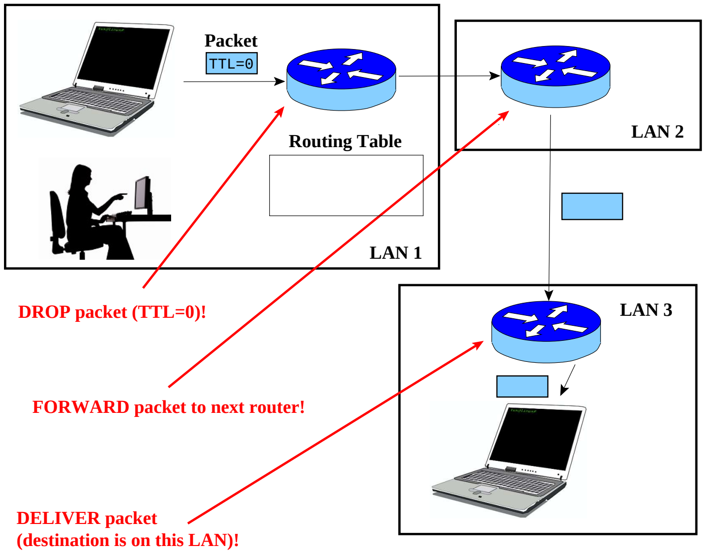


Next router to which a packet should be forwarded

Router Operations

- For each packet, the router decides whether to
  1. Drop — expired packets (TTL=0) are dropped.
  2. Deliver — if the packet is going to a machine on this LAN, deliver it.
  3. Forward — otherwise, send to neighboring router.
- TTL (time to live): a field in the IP header, decremented by each router, used to prevent packets from living forever.

Router Operation



DROP packet (TTL=0)!

FORWARD packet to next router!

DELIVER packet
(destination is on this LAN)!

Routing Table Protocols

- Open Shortest Path First (OSPF) — how should packets be routed *within* an autonomous system?
  - packets should travel along shortest paths.
- Border Gateway Protocol (BGP) — how should packets be routed *between* autonomous systems?
  - packets are routed based on contractual agreements.
- Routing animation: http://www.youtube.com/watch?v=RbY8Hb6abbg

Routing vs. Switch

- Switch:
  - forwards packets on a single LAN.
  - learns routes over time.
- Router:
  - can belong to multiple LANs.
  - uses routing tables to forward packets.

IP Address Format

- IPv4 address: 32 bits.
- IPv4 address structure:

network portion	host portion

- Network portion: IP prefix for all machines on a network.
- Host portion: identifies a particular device
- ⇒ Peter Packet & Subnetting:

http://www.youtube.com/watch?v=x-QC6l9KhQY&feature=related

IP Address Classes

Class	Leading bits	Size of network number bit field	Size of rest bit field	Number of networks	Addresses per network
A	0	8	24	2^7	2^{24}
B	10	16	16	2^{14}	2^{16}
C	110	24	8	2^{21}	2^8

- Class A — Reserved for government organizations, telcos.
- Class B — Reserved for ISPs, large businesses.
- Class C — Reserved for smaller organizations.

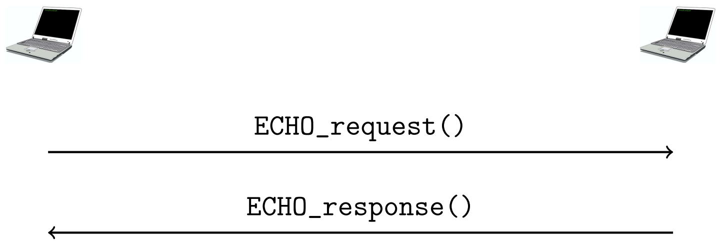
IP Address Classes. . .

Class	Start address	End address
A	0.0.0.0	127.255.255.255
B	128.0.0.0	191.255.255.255
C	192.0.0.0	223.255.255.255

Internet Control Message Protocol

- Internet Control Message Protocol (ICMP) — used for network diagnostics.
- ICMP messages:
  1. Echo request: please acknowledge receipt of packet.
  2. Echo response: packet receipt is acknowledged.
  3. Time exceeded: notify that packet has expired (TTL=0).
  4. Destination unreachable: notify that packet could not be delivered.
- https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol

Ping Protocol



ECHO_request()

ECHO_response()

- Diagnostic tool too see if a host is working.

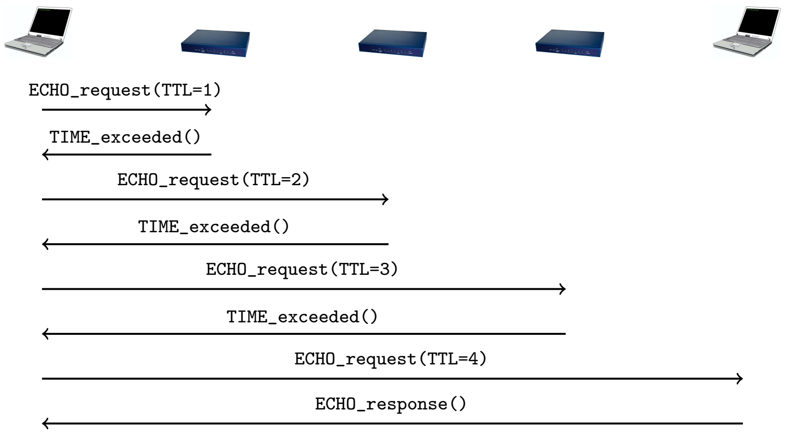
Traceroute Protocol

- How do we find the path a packet takes to a node N?
- Algorithm:
  1. Send `ECHO_request(TTL=1)` to N.
  2. A router that receives `ECHO_request(TTL=1)` responds with `TIME_exceeded()`.
  3. Send `ECHO_request(TTL=2)` to N.
  4. Repeat, increasing TTL each time, until N is reached, responding with `ECHO_response()`.

Traceroute Animation

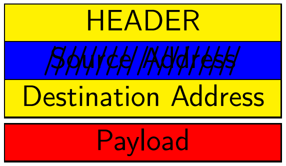
- ⇒ https://www.caida.org/publications/animations/#activemon

Traceroute Protocol. . .



IP Spoofing

- The source address in an IP packet is never checked: overwrite it!
- The sender will never get a response! So, why? Denial of service attack.



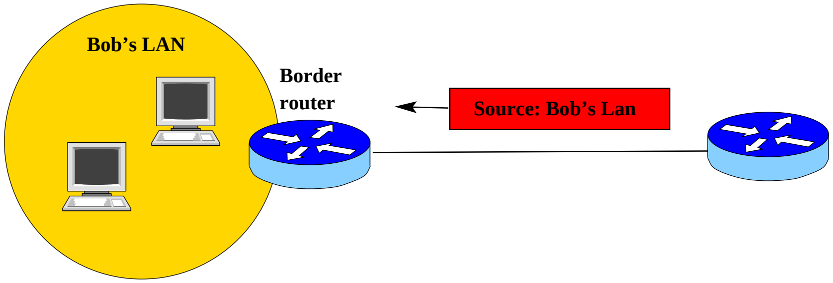
IP Spoofing

Definition (IP Spoofing)

In computer networking, IP address spoofing or IP spoofing is the creation of Internet Protocol (IP) packets with a false source IP address, for the purpose of hiding the identity of the sender or impersonating another computing system. . . .
IP address spoofing is most frequently used in denial-of-service attacks, where the objective is to flood the target with an overwhelming volume of traffic, and the attacker does not care about receiving responses to the attack packets.

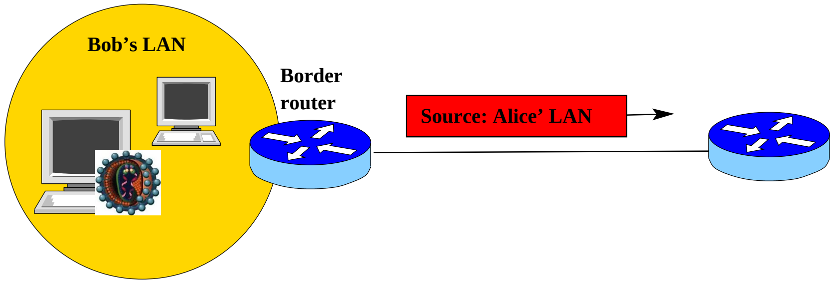
Source: https://en.wikipedia.org/wiki/IP_address_spoofing

Countermeasures: Border Router



- Border router can block packets whose source address appears to be from inside the subnetwork, although they come from outside the subnetwork.

Countermeasures: Border Router. . .



- Border router can block outgoing packets whose source address appears to be from outside the subnetwork.
- Maybe a node has been compromised by malware?

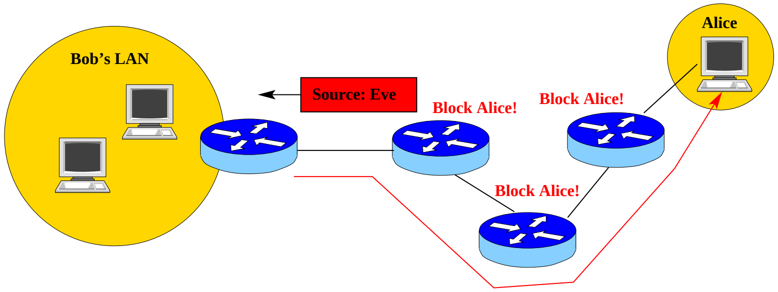
Countermeasures: IP Traceback

Definition (IP Traceback)

IP traceback is any method for reliably determining the origin of a packet on the Internet. The IP protocol does not provide for the authentication of the source IP address of an IP packet. . .
Use of false source IP addresses allows denial-of-service attacks (DoS) . . . . IP traceback is critical for identifying sources of attacks and instituting protection measures for the Internet. . . .

Source: https://en.wikipedia.org/wiki/IP_traceback

Countermeasures: IP Traceback. . .



- IP Traceback — determining the origin of a packet, without using the `source` field.
- Once we know the actual source address, we can ask
  1. the ASs to block packets from this location.
  2. the ISP controlling the source address to block suspicious machines.

IP Traceback: Packet Marking Algorithm

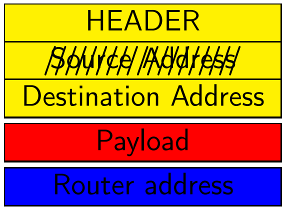
- Routers add information to packets, so that their path can be reconstructed.
- Naive approach: each router adds its address to the end of the packet:

HEADER
~~Source Address~~
Destination Address
Payload
Router 1
Router 2
Router 3
Router 4

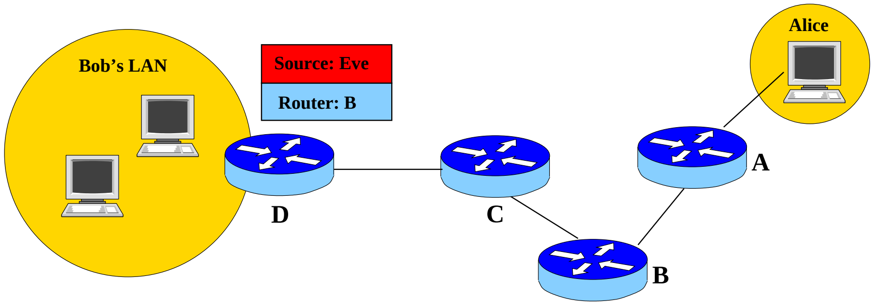
- Advantages: Easy to reconstruct path.
- Disadvantages: Router overhead, how to know if there's space in the packet?, packet fragmentation.

IP Traceback: Node Sampling Algorithm

- Only one router address can be stored in the packet.
- A router writes its address with probability p.



- Given enough packets, the path can be reconstructed.



- Probability the packet will be marked by C: p
- Probability the packet will be marked by B: $p \cdot (1 - p)$
- Probability the packet will be marked by A: $p \cdot (1 - p) \cdot (1 - p)$

IP Traceback Technique — Other Techniques

- Many other techniques have been proposed.
- Most not implemented — require cooperation from Internet routers.

Exercise I

1. What is IP spoofing?

2. Why would someone use IP spoofing?

3. Give an example of an attack that might use IP spoofing as a component?

Exercise II

1. Give a reason why a border router might block an *incoming packet*:
   ✎

2. Give a reason why a border router might block an *outgoing packet*:
   ✎

Exercise III

1. What fundamental network security problem does *IP Traceback* solve?
   ✎

   .

2. How can *IP Traceback* be used to combat a denial-of-service attack?
   ✎

Exercise IV

1. Describe a *naive* algorithm for Packet Marking IP Traceback.
   ✎

Exercise IV. . .

1. What are the advantages of this algorithm?
   ✎

2. What are the disadvantages?
   ✎

Exercise V

1. Describe the *Node Sampling Packet Marking IP Traceback* algorithm.
   ✎

Exercise V. . .

1. What are the advantages of this algorithm?
   ✎

2. What are the disadvantages?
   ✎

Outline

1. Introduction
2. The Link Layer
   - Hubs and Switches
   - Ethernet Frames
   - ARP Spoofing
3. The Network Layer
   - ICPM
   - IP Spoofing
4. The Transport Layer
   - TCP Session Hijacking
5. Summary

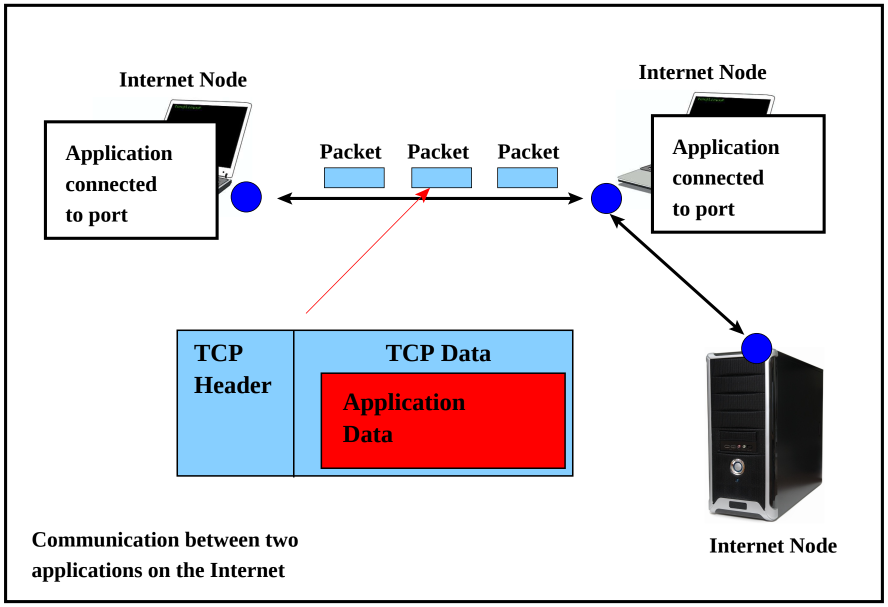
The Transport Layer

- Communication between processes connected to ports.
- Abstraction:
  1. Source/Destination: Ports connected to processes
  2. Data: TCP/UDP packets
  3. Addressing: IP address + port number
- Transmission Control Protocol (TCP) — connection-based protocol; guaranteed and ordered delivery of packets.
- User Datagram Protocol (UDP) — connection-less protocol; quick delivery without guarantees.

TCP Animation

- http://williams.comp.ncat.edu/IA_visualization_labs/security_visual_tools/VisTools.html
  → *SYN Flood Animated Simulator* → *Normal Network Traffic*
- NOTE: Use Firefox, not Safari.
- Maybe gone?

Internet Protocol L. — Transport Layer


Internet Node
Application connected to port
Packet
Packet
Packet
Internet Node
Application connected to port
TCP Header
TCP Data
Application Data
Internet Node
Communication between two applications on the Internet

TCP Packet Format

Source Port (16 bits)

Destination Port (16 bits)

Sequence Number (32 bits)

Acknowledgement Number (32 bits)

Offset (4 bits)

Reserved (4 bits)

Flags (8 bits)

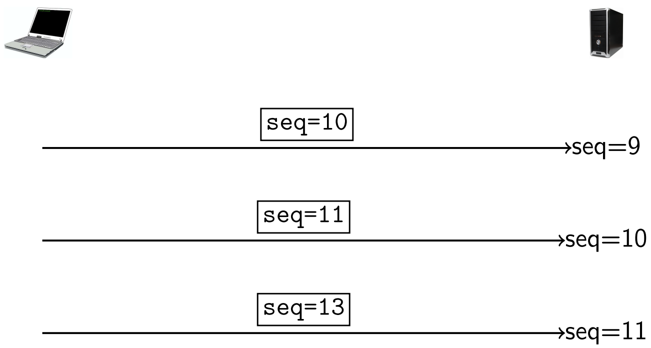
Window size (16 bits)

Checksum (16 bits)

Urgent Pointer (16 bits)

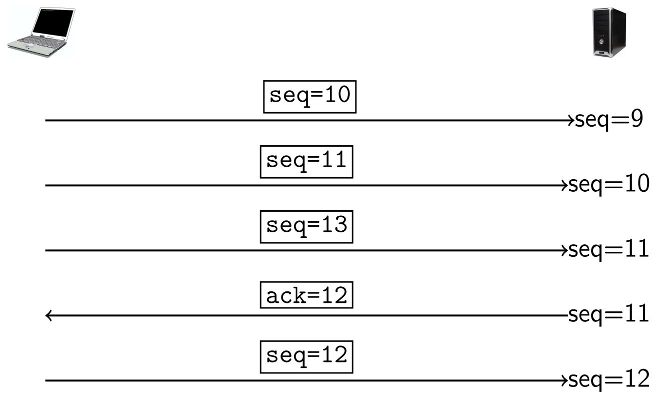
Payload

TCP Sequence Number



- Incremented for every packet by payload length.
- Allows us to determine when packets arrive out of order or don't arrive.

TCP Acknowledgement Number



seq=10

seq=9

seq=11

seq=10

seq=13

seq=11

ack=12

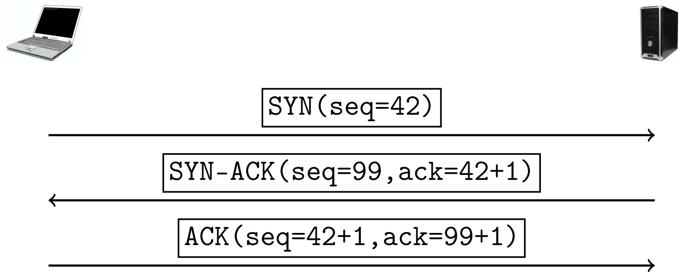
seq=11

seq=12

seq=12

- Receiver sends an acknowledgement package with the sequence number of the next payload byte it wants to receive.

TCP Connections



SYN(seq=42)

SYN-ACK(seq=99,ack=42+1)

ACK(seq=42+1,ack=99+1)

- TCP uses a 3-way handshake to set up a connection.
- The protocol includes a random initialization of the sequence number.

TCP Connections. . .

- After the 3-way handshake a connection has been established.
- This means that Alice and Bob have agreed on two sequence numbers:
  - one sequence number that Alice will use when she sends packets to Bob (to allow him to detect when a packet has been lost or has arrived out of order),
  - and one sequence number that Bob will use when he sends packets to Alice (to allow her to detect the same problems).

TCP Session Hijacking

- TCP Session Hijacking — an attacker
  1. hijacks another user's TCP connection, or
  2. alters another user's TCP connection.

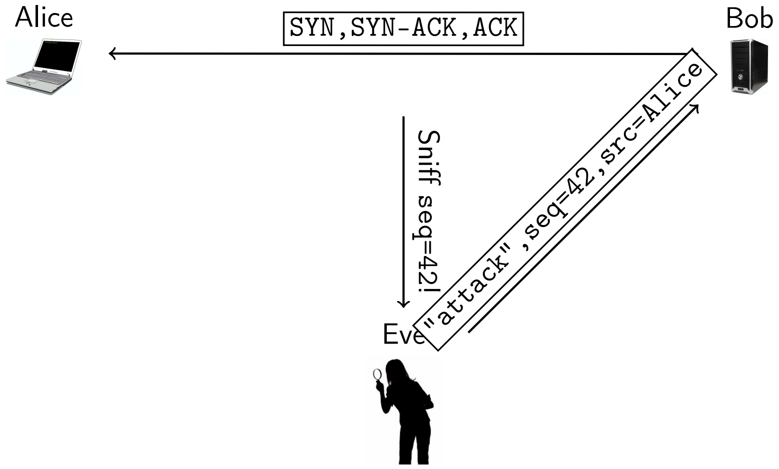
TCP Sequence Prediction Attack

- Session spoofing — The attacker is able to create a TCP session with a server, who thinks it is talking to another client.
- Early TCP implementations had easily guessable sequence numbers.
- Blind injection attack: Eve won't receive replies from Bob.

Attack 1: Complete Session Hijacking

- Eve is on the same network segment as Alice and Bob, and packet sniffs on them as they establish their TCP connection.
- Eve guesses the next sequence number and sends a spoofed attack command to Bob, appearing to be Alice.

Attack 1: Complete Session Hijacking. . .



Alice

`SYN,SYN-ACK,ACK`

Bob

Sniff seq=42!

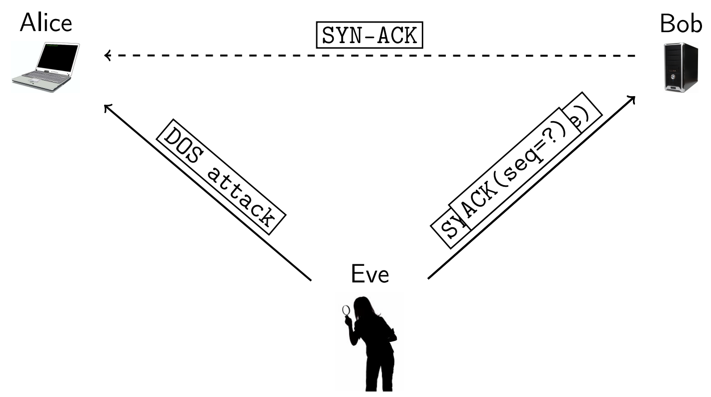
`"attack",seq=42,src=Alice`

Eve

Attack 2: TCP Session Spoofing

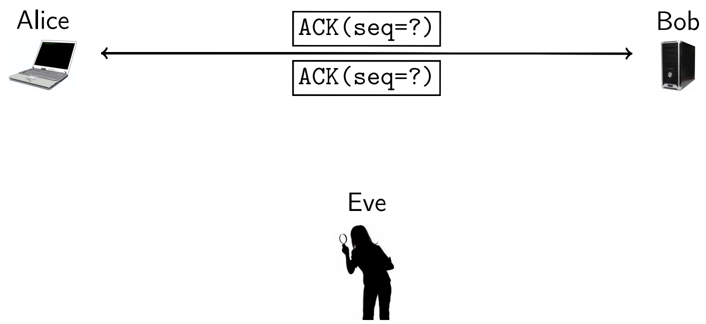
1. Eve launches a denial-of-service attack against Alice so she can't interfere with the attack.
2. Eve sends a `SYN(src=Alice)` to Bob.
3. Bob responds with a `SYN-ACK` to Alice, who cannot respond since she's under attack.
4. Eve guesses N, Bob's next sequence number.
5. Eve sends an `ACK(seq=N)` to Bob.
6. Eve talks to Bob as if she is Alice.

Attack 2: TCP Session Spoofing. . .



- Eve establishes a TCP connection with Bob, who thinks he's talking to Alice.
- Eve needs to guess the next sequence number Bob will use.

Attack 2: Detecting ACK Storms



- Blind injection attacks can cause an ACK Storm, when the client and server try to resynchronize their sequence numbers.
- A firewall can, eventually, detect the ACK Storm.

Countermeasures

- Don't use predictable sequence numbers.
- Encrypt at the network layer (IPsec).
- Encrypt at the application layer (https).

Exercise I

1. What *addressing* information is part of the TCP header?
2. How are random changes to the packet in transit detected?
3. What is the purpose of the *TCP 3-Way Handshake protocol*?

Exercise II

1. Why does the TCP protocol need sequence numbers?
   ✎

2. Why is the initial sequence number random? Why not just start at 0?
   ✎

3. In the *TCP 3-Way Handshake protocol*, Alice starts by sending a
   ✎ , Bob answers with a
   ✎ packet, and Alice concludes the exchange with an
   ✎ packet.

Exercise III

1. Why does the TCP protocol a *3-way* handshake? Why isn't 2 enough?
   ✎

2. In pseudocode, explain the 3-way handshake protocol!
   ✎

Outline

1. Introduction
2. The Link Layer
   - Hubs and Switches
   - Ethernet Frames
   - ARP Spoofing
3. The Network Layer
   - ICPM
   - IP Spoofing
4. The Transport Layer
   - TCP Session Hijacking
5. Summary

Readings and References

- Chapter 5 in *Introduction to Computer Security*, by Goodrich and Tamassia.