Attacks on TCP/IP, BGP, DNS Denial of Service

Vitaly Shmatikov

Reading Assignment

- "SYN cookies" by Bernstein
- "IP spoofing demystified" from Phrack magazine
- "It's the end of the cache as we know it" by Kaminsky (BlackHat 2008)

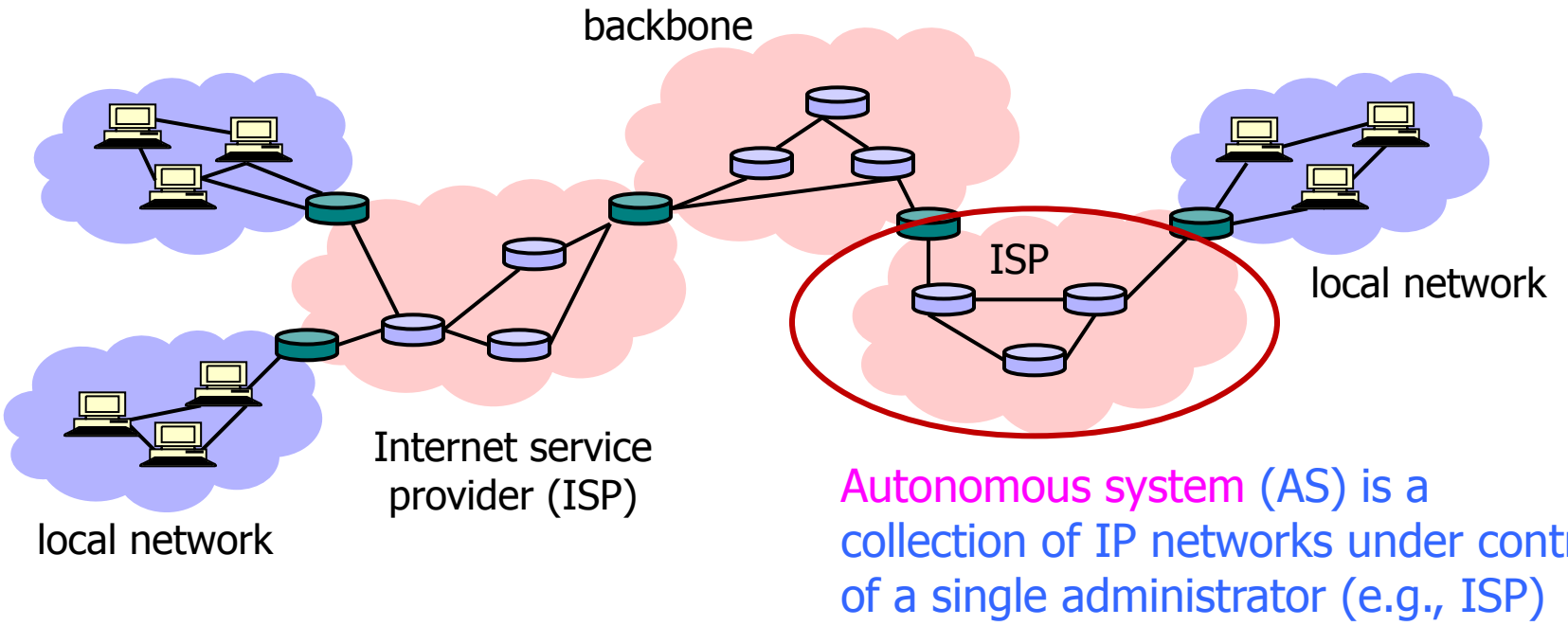
Warm Up: 802.11b

- NAV (Network Allocation Vector)
  - 15-bit field, max value: 32767
  - Any node can reserve channel for NAV microseconds
  - No one else should transmit during NAV period

    … but not followed by most 802.11b cards
- De-authentication
  - Any node can send deauth packet to AP
  - Deauth packet unauthenticated

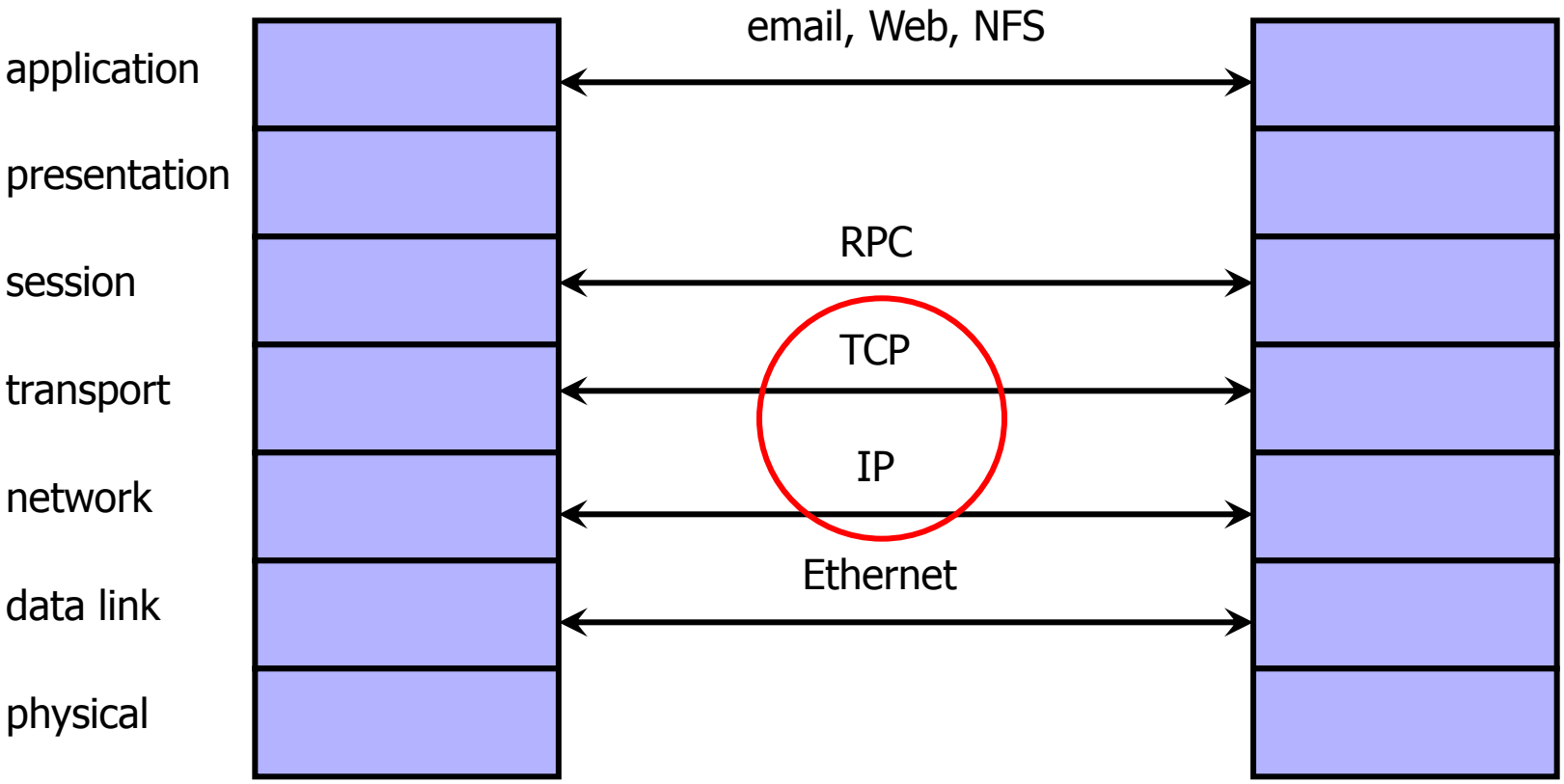
    … attacker can repeatedly deauth anyone

Internet Is a Network of Networks

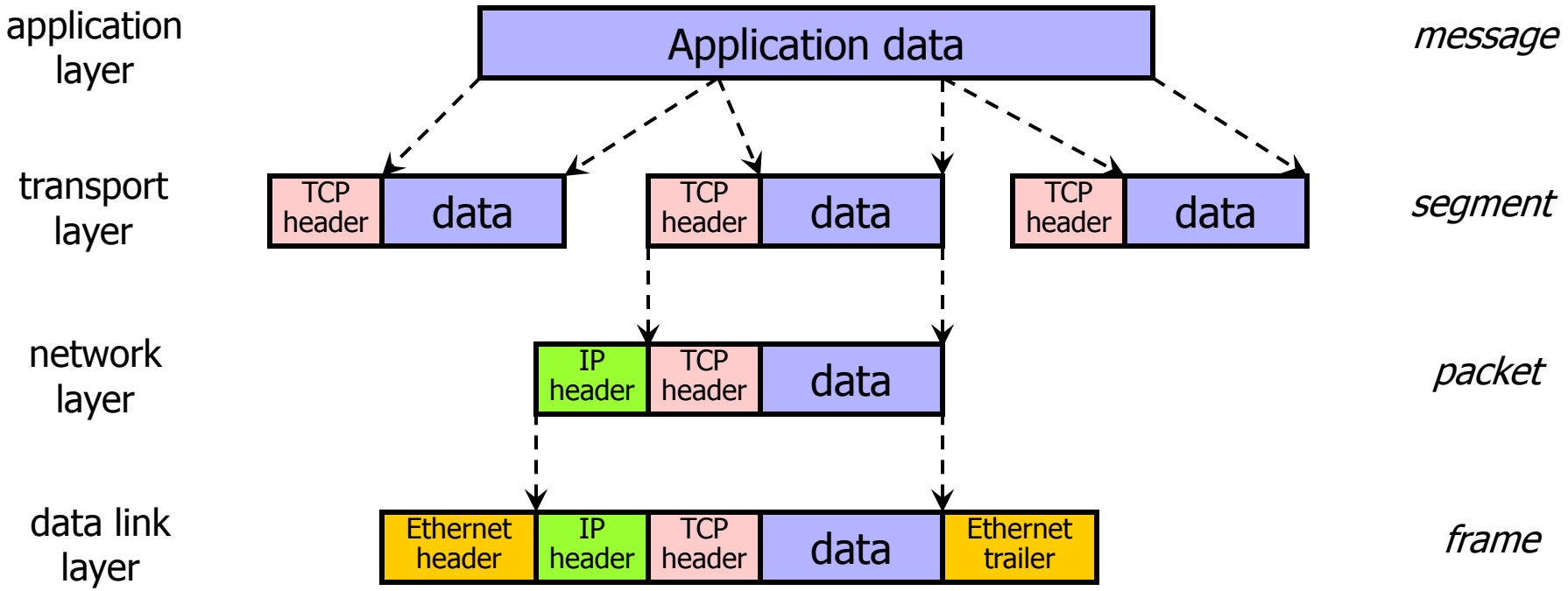


- TCP/IP for packet routing and connections
- Border Gateway Protocol (BGP) for route discovery
- Domain Name System (DNS) for IP address discovery

OSI Protocol Stack

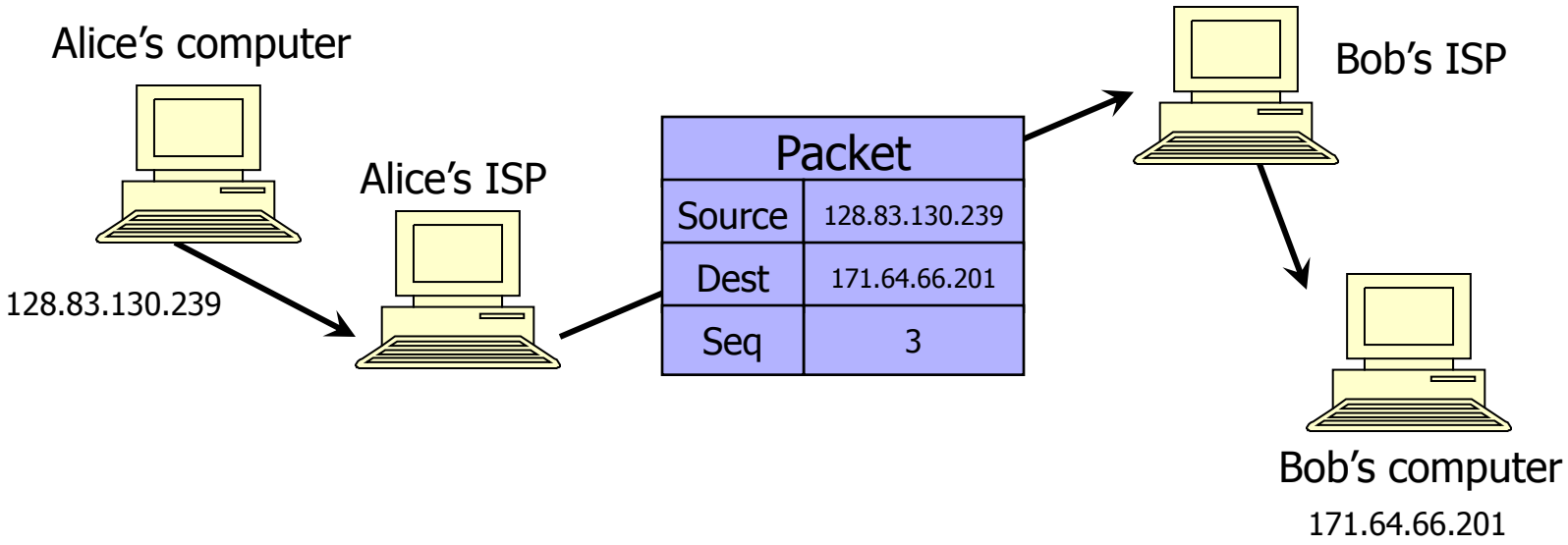


Data Formats



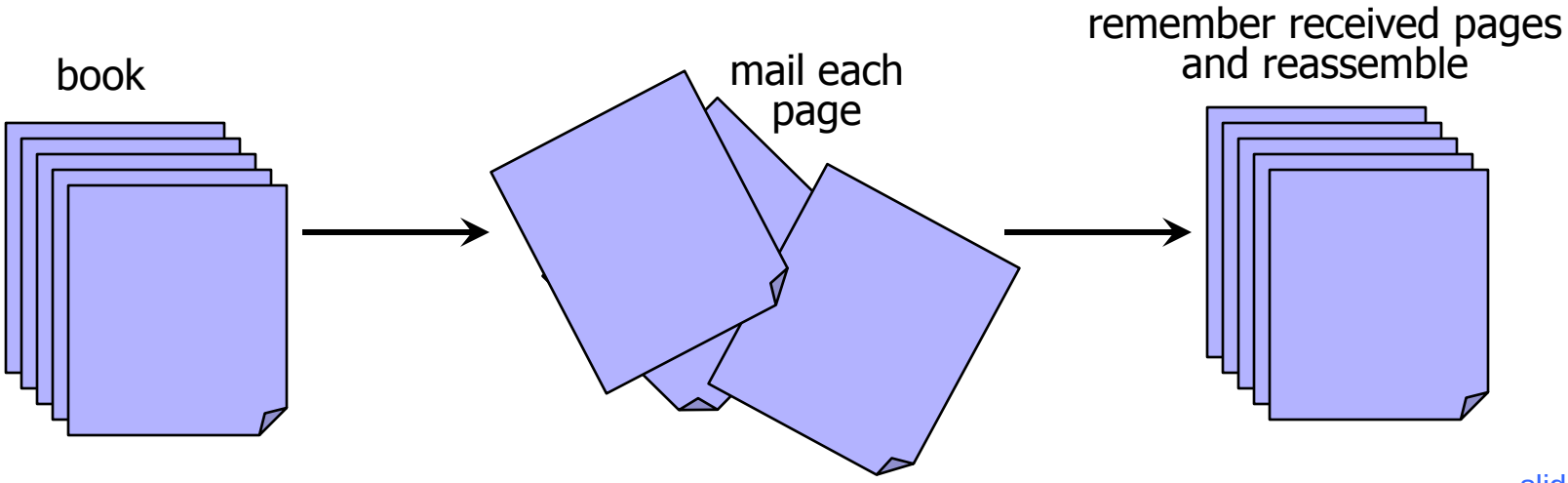
IP (Internet Protocol)

- Connectionless
  - Unreliable, "best-effort" protocol
- Uses numeric addresses for routing
- Typically several hops in the route



TCP (Transmission Control Protocol)

- Sender: break data into packets
  - Sequence number is attached to every packet
- Receiver: reassemble packets in correct order
  - Acknowledge receipt; lost packets are re-sent
- Connection state maintained on both sides

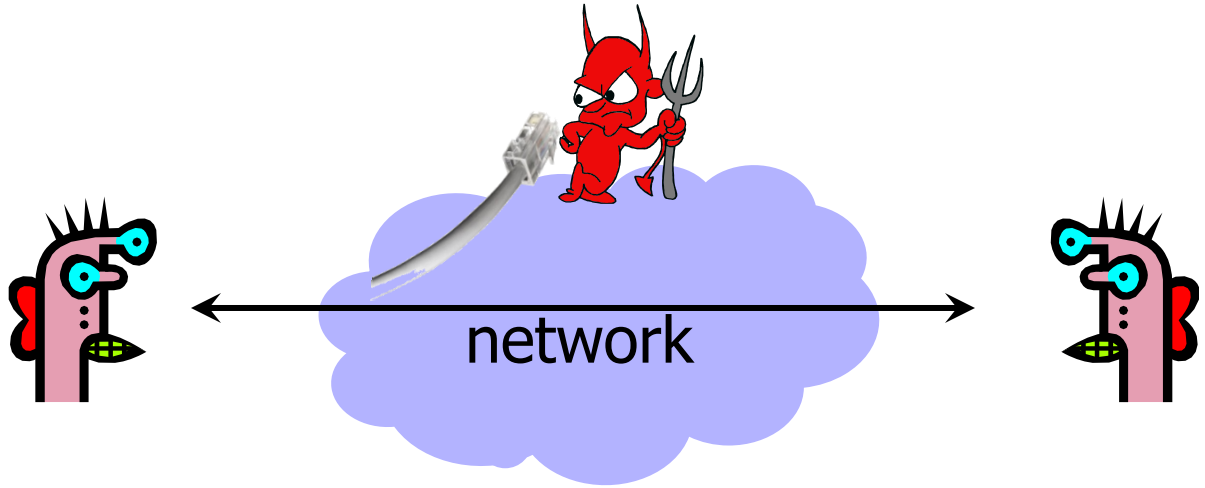


ICMP (Control Message Protocol)

- Provides feedback about network operation
  - "Out-of-band" messages carried in IP packets
- Error reporting, congestion control, reachability…
  - Destination unreachable
  - Time exceeded
  - Parameter problem
  - Redirect to better gateway
  - Reachability test (echo / echo reply)
  - Message transit delay (timestamp request / reply)

Packet Sniffing

- Many applications send data unencrypted
  - ftp, telnet send passwords in the clear
- Network interface card (NIC) in "promiscuous mode" reads all passing data



Solution: encryption (e.g., IPsec, HTTPS), improved routing

"Ping of Death"

- If an old Windows machine received an ICMP packet with a payload longer than 64K, machine would crash or reboot
  - Programming error in older versions of Windows
  - Packets of this length are illegal, so programmers of Windows code did not account for them

Solution: patch OS, filter out ICMP packets

“Teardrop” and “Bonk”

- TCP fragments contain Offset field
- Attacker sets Offset field to overlapping values
  - Bad implementation of TCP/IP will crash when attempting to re-assemble the fragments
- … or to very large values
  - Bad TCP/IP implementation will crash

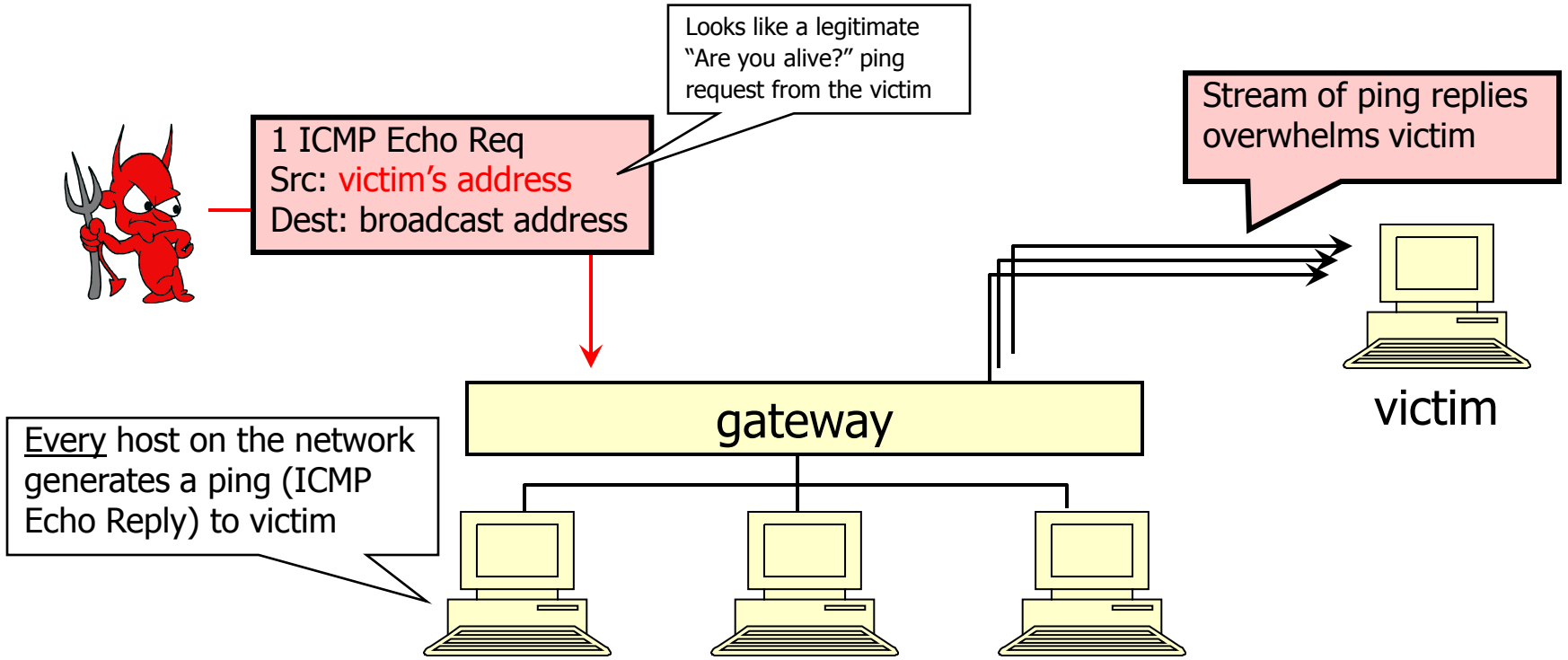
Solution: use up-to-date TCP/IP implementation

"LAND"

- IP packet with source address, port equal to destination address, port; SYN flag set
- Triggers loopback in the Windows XP SP2 implementation of TCP/IP stack, locks up CPU

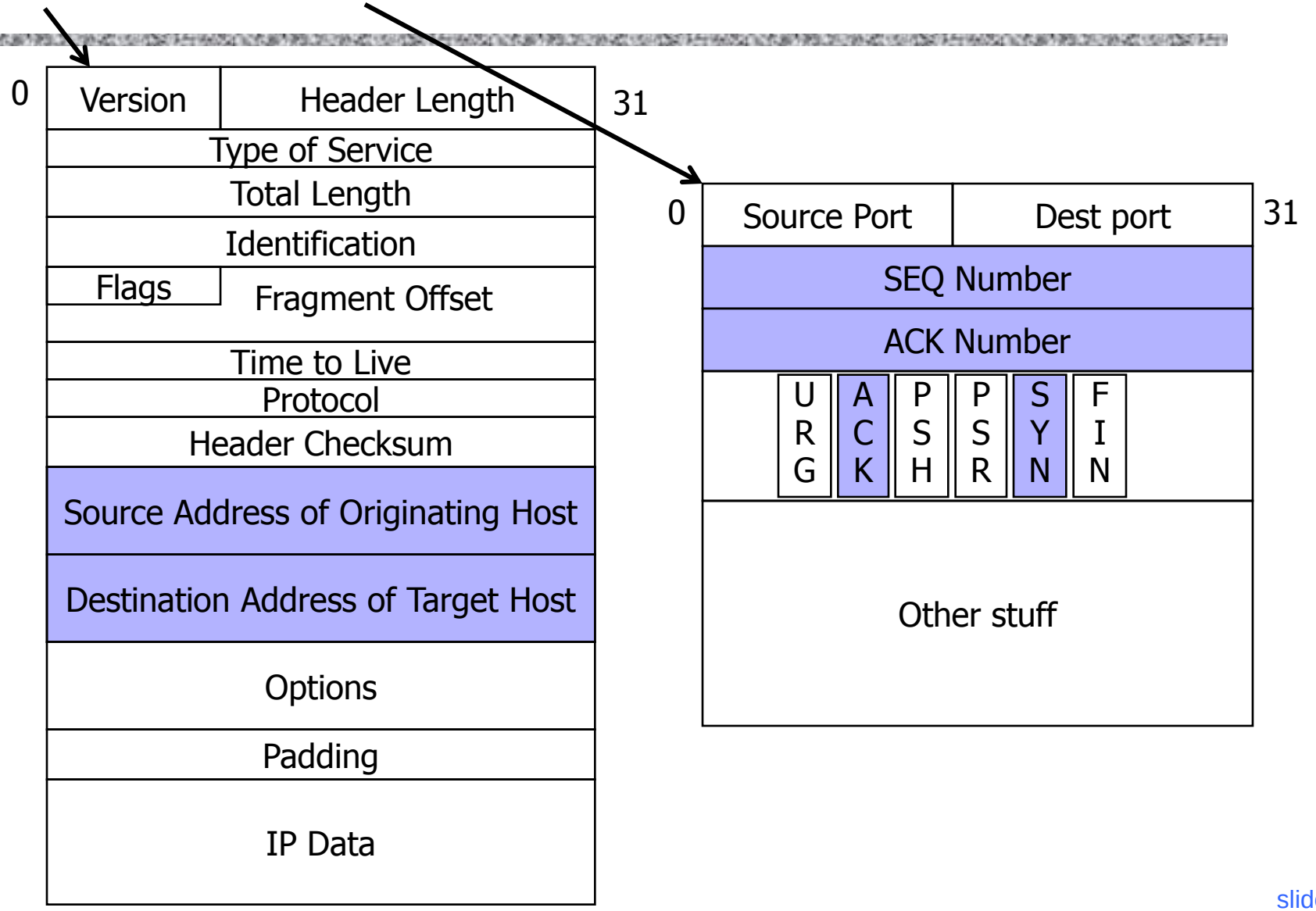
Solution: ingress filtering

"Smurf" Reflector Attack

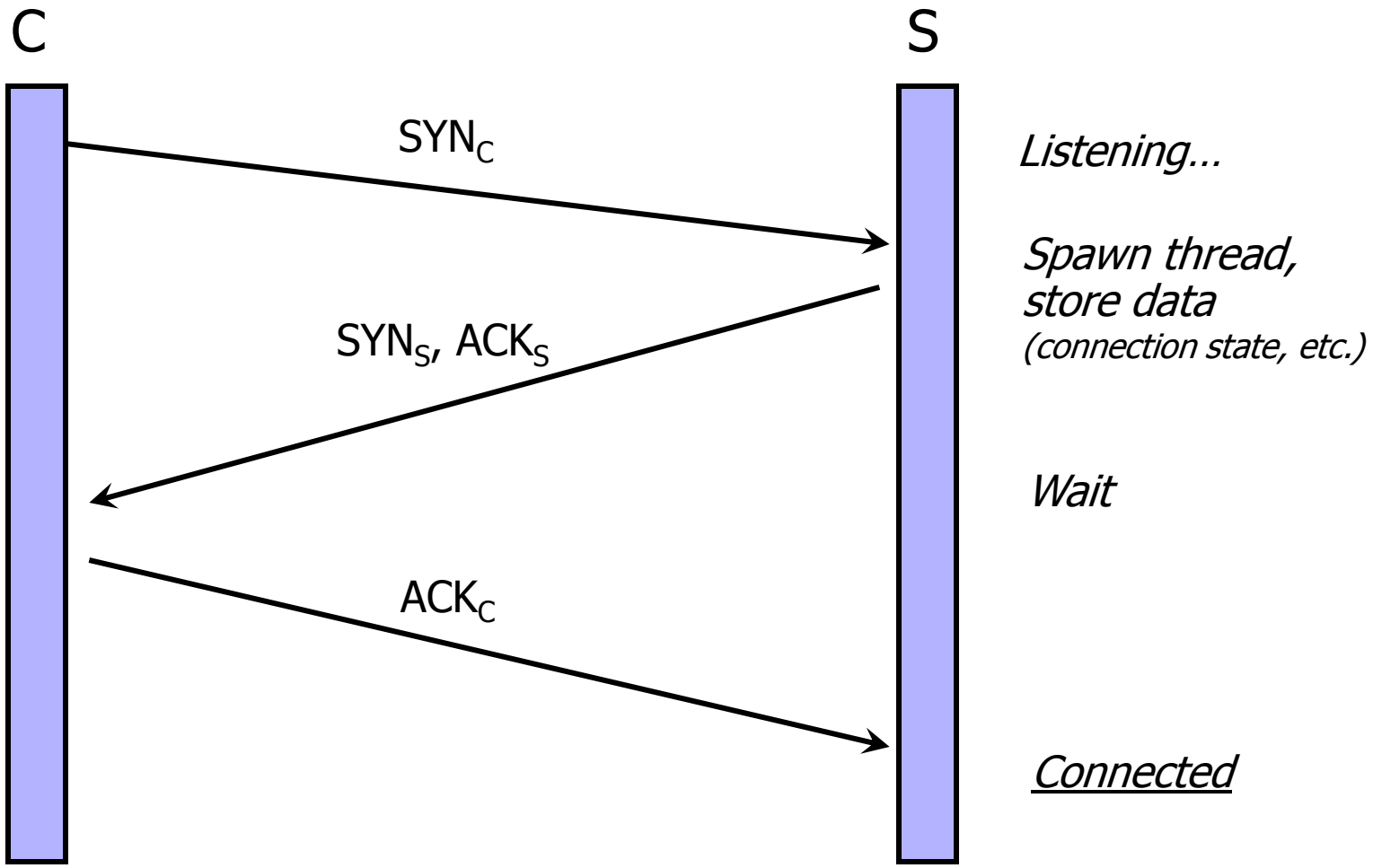


Solution: reject external packets to broadcast addresses

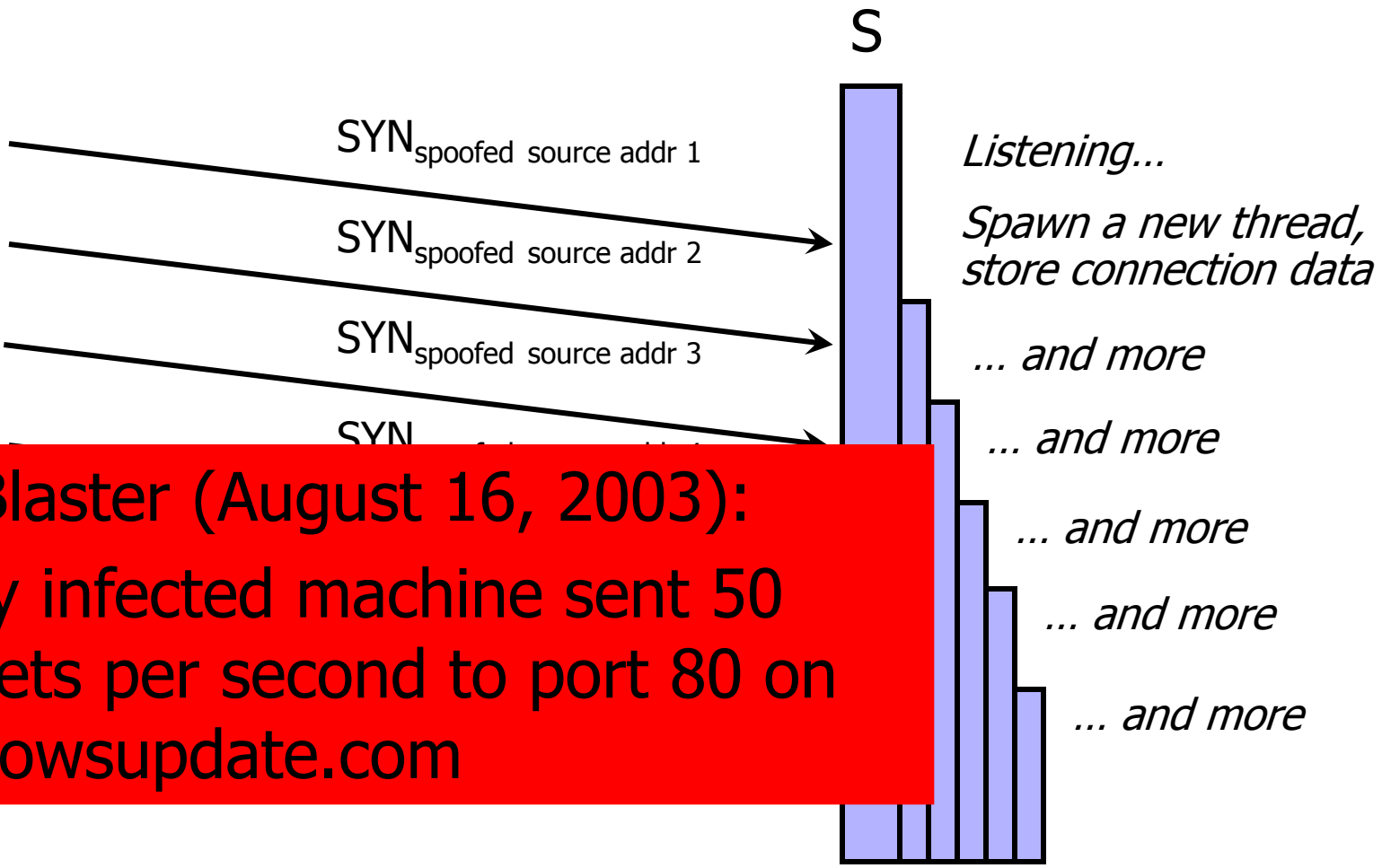
IP and TCP Headers



TCP Handshake



SYN Flooding Attack



MS Blaster (August 16, 2003):

every infected machine sent 50 packets per second to port 80 on windowsupdate.com

SYN Flooding Explained

- Attacker sends many connection requests with spoofed source addresses
- Victim allocates resources for each request
  - New thread, connection state maintained until timeout
  - Fixed bound on half-open connections
- Once resources exhausted, requests from legitimate clients are denied
- This is a classic denial of service pattern
  - It costs nothing to TCP initiator to send a connection request, but TCP responder must spawn a thread for each request - asymmetry!

SYN Floods

[Phrack 48, no 13, 1996]

OS	Backlog queue size
Linux 1.2.x	10
FreeBSD 2.1.5	128
WinNT 4.0	6

Backlog timeout: 3 minutes

Attacker need only send

128 SYN packets every 3 minutes

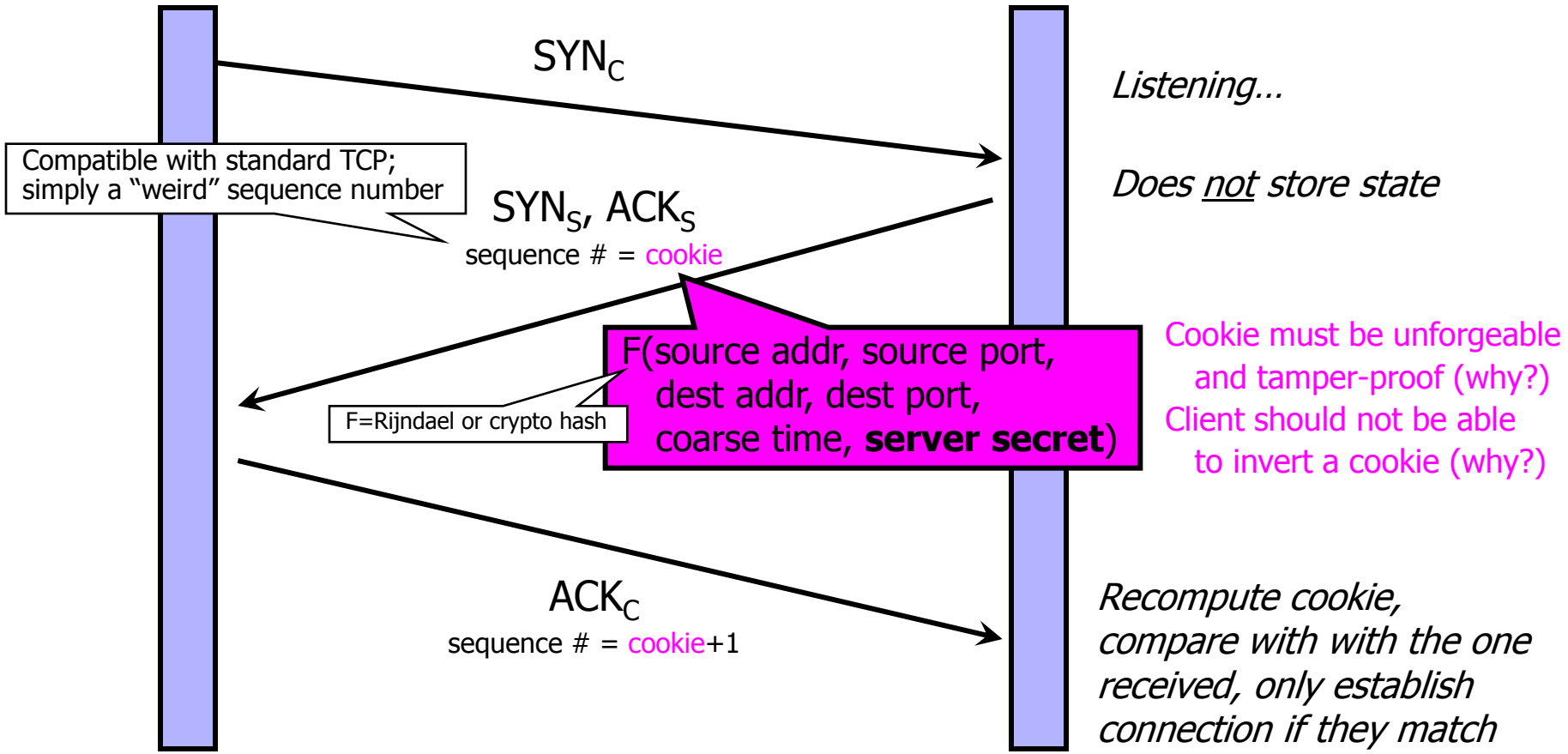
$\Rightarrow$ low-rate SYN flood

Preventing Denial of Service

- DoS is caused by asymmetric state allocation
  - If responder opens new state for each connection attempt, attacker can initiate thousands of connections from bogus or forged IP addresses
- Cookies ensure that the responder is stateless until initiator produced at least two messages
  - Responder's state (IP addresses and ports of the connection) is stored in a cookie and sent to initiator
  - After initiator responds, cookie is regenerated and compared with the cookie returned by the initiator

SYN Cookies

[Bernstein and Schenk]

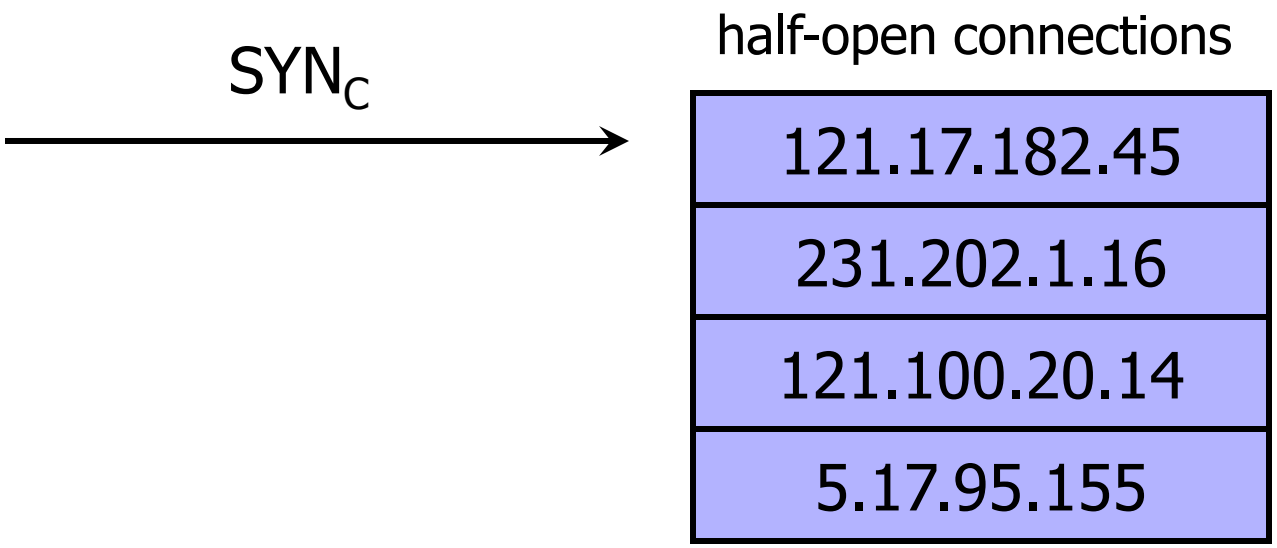


More info: http://cr.yp.to/syncookies.html

Anti-Spoofing Cookies: Basic Pattern

- Client sends request (message #1) to server
- Typical protocol:
  - Server sets up connection, responds with message #2
  - Client may complete session or not - potential DoS!
- Cookie version:
  - Server responds with hashed connection data instead of message #2
  - Client confirms by returning hashed data
    - If source IP address is bogus, attacker can't confirm
  - Need an extra step to send postponed message #2, except in TCP (can piggyback on SYN-ACK in TCP)

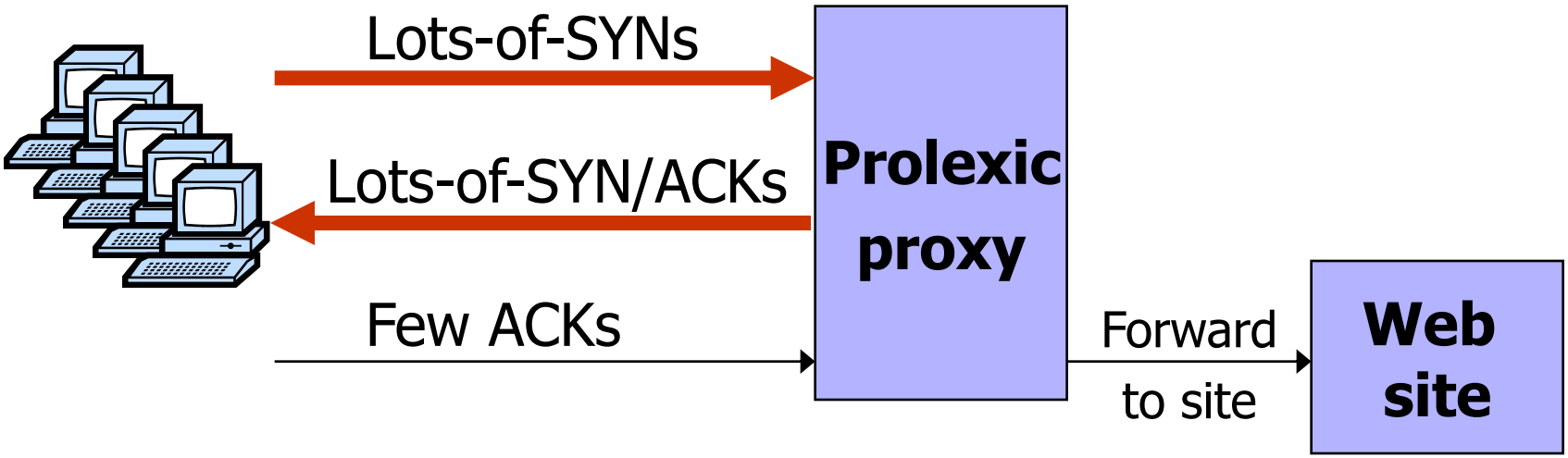
Another Defense: Random Deletion



- If SYN queue is full, delete random entry
  - Legitimate connections have a chance to complete
  - Fake addresses will be eventually deleted
- Easy to implement

Prolexic / Verisign

- Idea: only forward established TCP connections to site



Other Junk-Packet Attacks

Attack Packet	**Victim Response**	**Rate: attk/day [ATLAS 2013]**
TCP SYN to open port	TCP SYN/ACK	773
TCP SYN to closed port	TCP RST	
TCP ACK or TCP DATA	TCP RST	
TCP RST	No response	
TCP NULL	TCP RST	
ICMP ECHO Request	ICMP ECHO Response	**50**
UDP to closed port	ICMP Port unreachable	387

Proxy must keep floods of these away from website

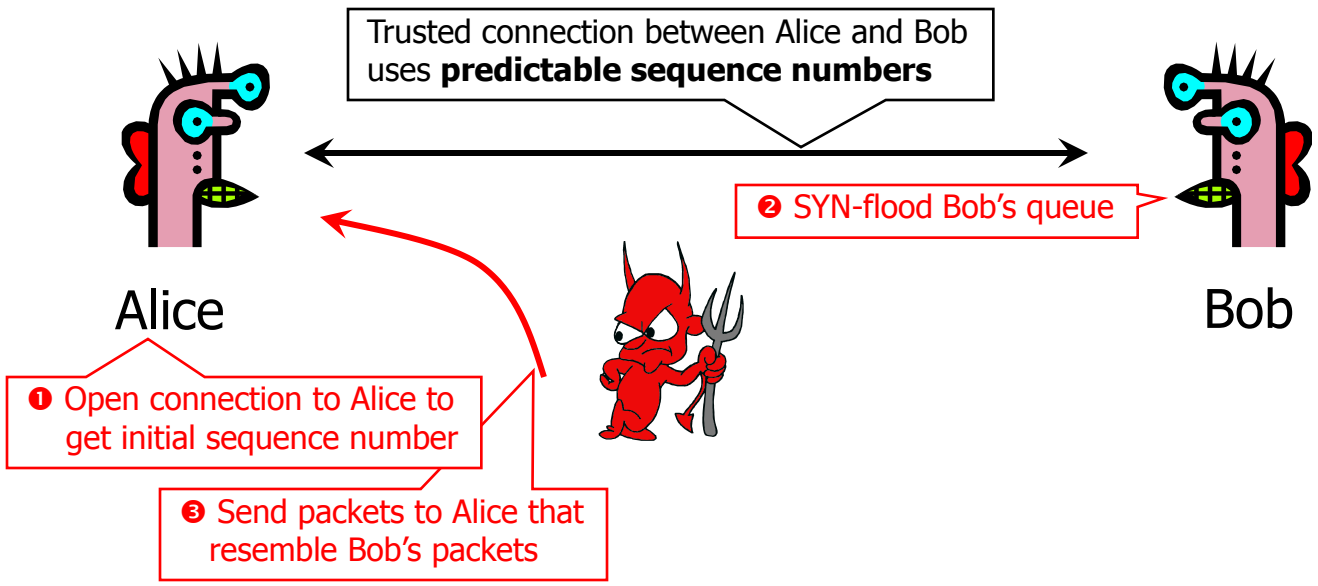
Stronger Attack: TCP Con Flood

- Command bot army to:
  - Complete TCP connection to web site
  - Send short HTTP HEAD request
  - Repeat
- Will bypass SYN flood protection proxy but …
  - Attacker can no longer use random source IPs
    - Reveals location of bot zombies
  - Proxy can now block or rate-limit bots

TCP Connection Spoofing

- Each TCP connection has associated state
  - Sequence number, port number
- TCP state is easy to guess
  - Port numbers standard, seq numbers predictable
- Can inject packets into existing connections
  - If attacker knows initial sequence number and amount of traffic, can guess likely current number
  - Guessing a 32-bit seq number is not practical, BUT…
  - Most systems accept large windows of sequence numbers (to handle packet losses), so send a flood of packets with likely sequence numbers

"Blind" IP Spoofing Attack



- Can't receive packets sent to Bob, but can bypass Alice's IP address-based authentication
  - rlogin and other remote access tools, SPF defense against spam

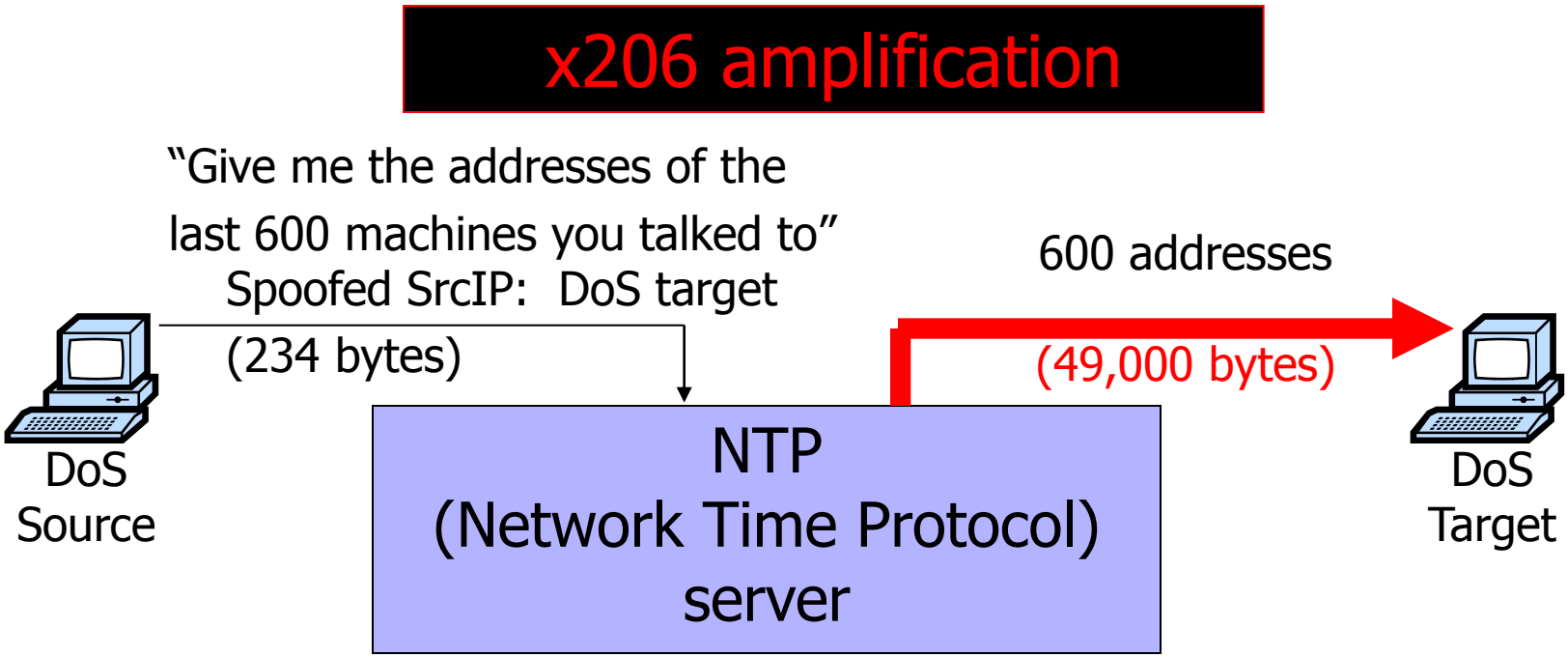
DoS by Connection Reset

- If attacker can guess the current sequence number for an existing connection, can send Reset packet to close it
- Especially effective against long-lived connections
  - For example, BGP route updates

User Datagram Protocol (UDP)

- UDP is a connectionless protocol
  - Simply send datagram to application process at the specified port of the IP address
  - Source port number provides return address
  - Applications: media streaming, broadcast
- No acknowledgement, no flow control, no message continuation
- Denial of service by UDP data flood

NTP Amplification Attack



December 2013 – February 2014:

400 Gbps DDoS attacks involving 4,529 NTP servers

7 million unsecured NTP servers on the Internet (Arbor)

Countermeasures

- Above transport layer: Kerberos
  - Provides authentication, protects against application-layer spoofing
  - Does not protect against connection hijacking
- Above network layer: SSL/TLS and SSH
  - Protects against connection hijacking and injected data
  - Does not protect against DoS by spoofed packets
- Network (IP) layer: IPsec
  - Protects against hijacking, injection, DoS using connection resets, IP address spoofing

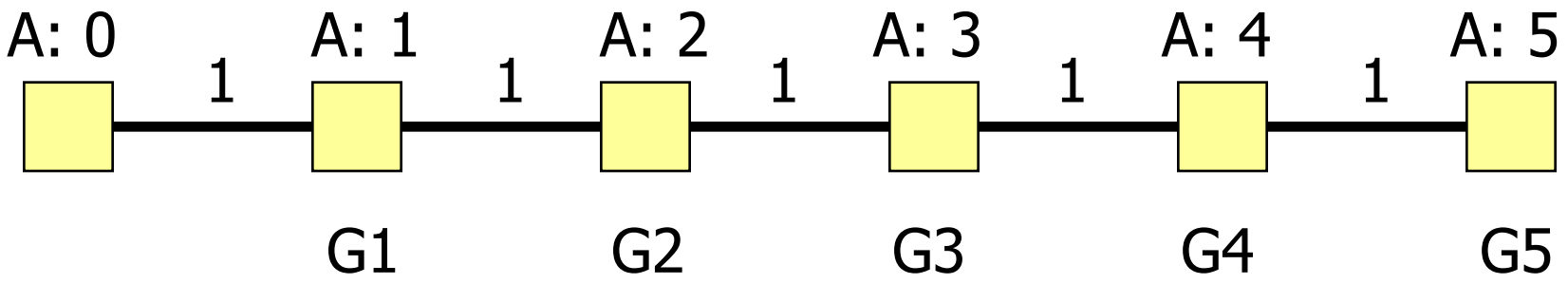
IP Routing

- Routing of IP packets is based on IP addresses
  - 32-bit host identifiers (128-bit in IPv6)
- Routers use a forwarding table
  - Entry = destination, next hop, network interface, metric
  - Table look-up for each packet to decide how to route it
- Routers learn routes to hosts and networks via routing protocols
  - Host is identified by IP address, network by IP prefix
- BGP (Border Gateway Protocol) is the core Internet protocol for establishing inter-AS routes

Distance-Vector Routing

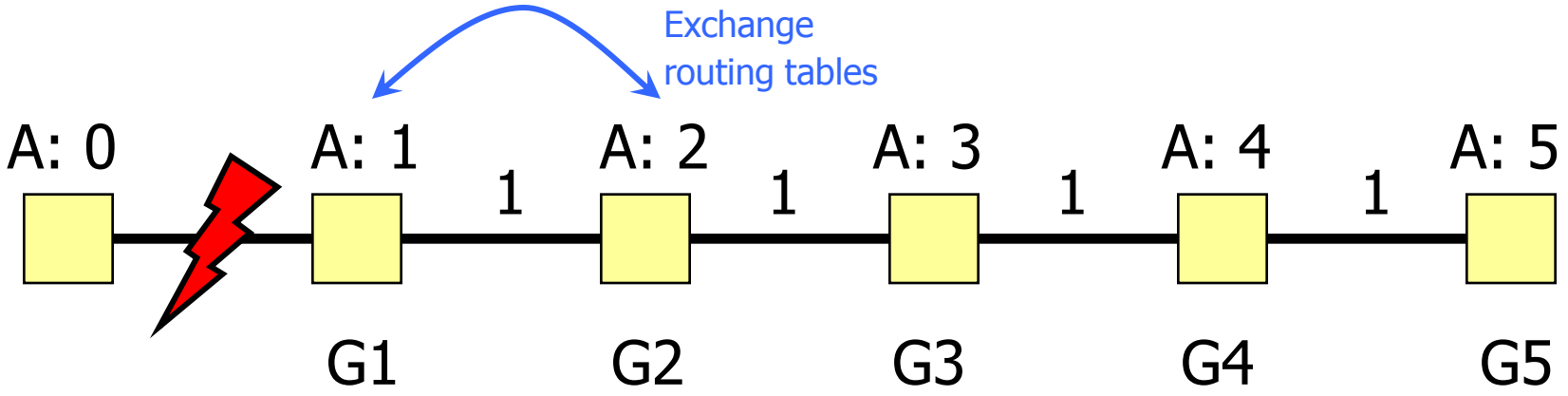
- Each node keeps vector with distances to all nodes
- Periodically sends distance vector to all neighbors
- Neighbors send their distance vectors, too; node updates its vector based on received information
  - Bellman-Ford algorithm: for each destination, router picks the neighbor advertising the cheapest route, adds his entry into its own routing table and re-advertises
  - Used in RIP (routing information protocol)
- Split-horizon update
  - Do not advertise a route on an interface from which you learned the route in the first place!

Good News Travels Fast



- G1 advertises route to network A with distance 1
- G2-G5 quickly learn the good news and install the routes to A via G1 in their local routing tables

Bad News Travels Slowly



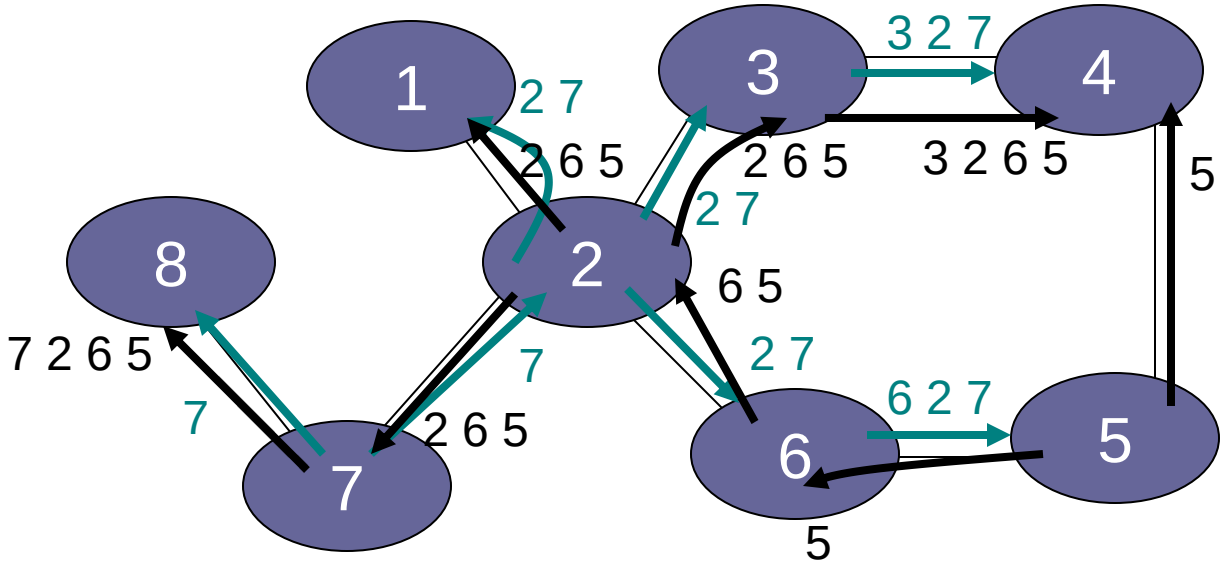
- G1's link to A goes down
- G2 is advertising a pretty good route to G1 (cost=2)
- G1's packets to A are forever looping between G2 and G1
- G1 is now advertising a route to A with cost=3, so G2 updates its own route to A via G1 to have cost=4, and so on
  - G1 and G2 are slowly counting to infinity
  - Split-horizon updates only prevent two-node loops

Overview of BGP

- BGP is a path-vector protocol between ASes
- Just like distance-vector, but routing updates contain an actual path to destination node
  - List of traversed ASes and a set of network prefixes belonging to the first AS on the list
- Each BGP router receives update messages from neighbors, selects one "best" path for each prefix, and advertises this path to its neighbors
  - Can be the shortest path, but doesn't have to be
    - "Hot-potato" vs. "cold-potato" routing
  - Always route to most specific prefix for a destination

BGP Example

[Wetherall]



- AS 2 provides transit for AS 7
  - Traffic to and from AS 7 travels through AS 2

Some (Old) BGP Statistics

- BGP routing tables contain about 125,000 address prefixes mapping to about 17-18,000 paths
- Approx. 10,000 BGP routers
- Approx. 2,000 organizations own AS
- Approx. 6,000 organizations own prefixes
- Average route length is about 3.7
- 50% of routes have length less than 4 ASes
- 95% of routes have length less than 5 ASes

BGP Misconfiguration

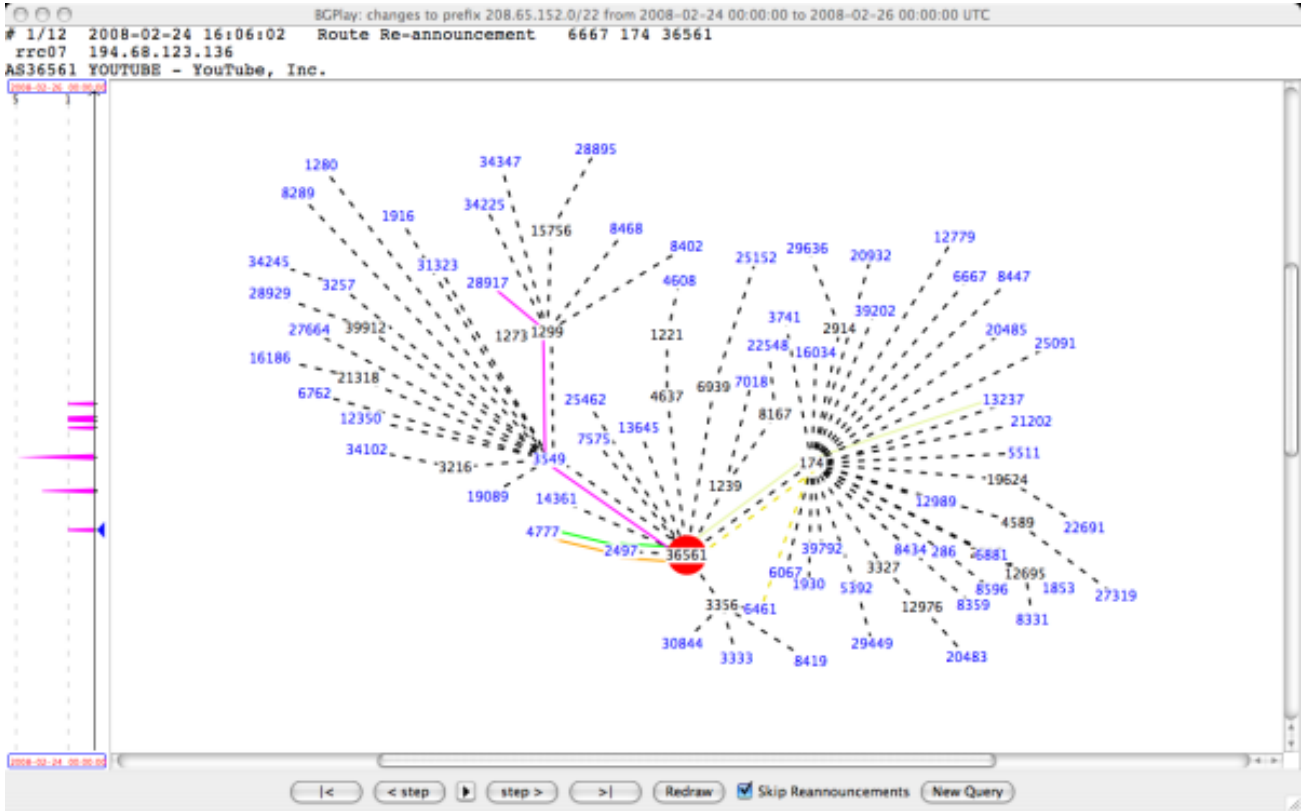
- Domain advertises good routes to addresses it does not know how to reach
  - Result: packets go into a network "black hole"
- April 25, 1997: "The day the Internet died"
  - AS7007 (Florida Internet Exchange) de-aggregated the BGP route table and re-advertised all prefixes as if it originated paths to them
    - In effect, AS7007 was advertising that it has the best route to every host on the Internet
  - Huge network instability as incorrect routing data propagated and routers crashed under traffic

BGP (In)Security

- BGP update messages contain no authentication or integrity protection
- Attacker may falsify the advertised routes
  - Modify the IP prefixes associated with a route
    - Can blackhole traffic to certain IP prefixes
  - Change the AS path
    - Either attract traffic to attacker's AS, or divert traffic away
    - Interesting economic incentive: an ISP wants to dump its traffic on other ISPs without routing their traffic in exchange
  - Re-advertise/propagate AS path without permission
    - For example, a multi-homed customer may end up advertising transit capability between two large ISPs

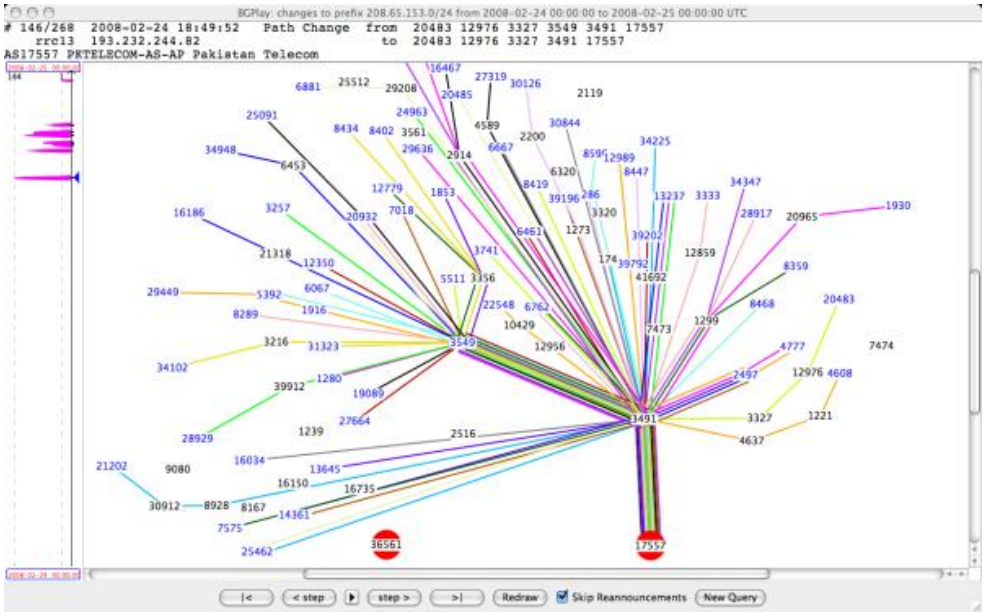
YouTube (Normally)

- AS36561 (YouTube) advertises 208.65.152.0/22



YouTube (February 24, 2008)

- Pakistan government wants to block YouTube
  - AS17557 (Pakistan Telecom) advertises 208.65.153.0/24
  - All YouTube traffic worldwide directed to AS17557



- Result: two-hour YouTube outage

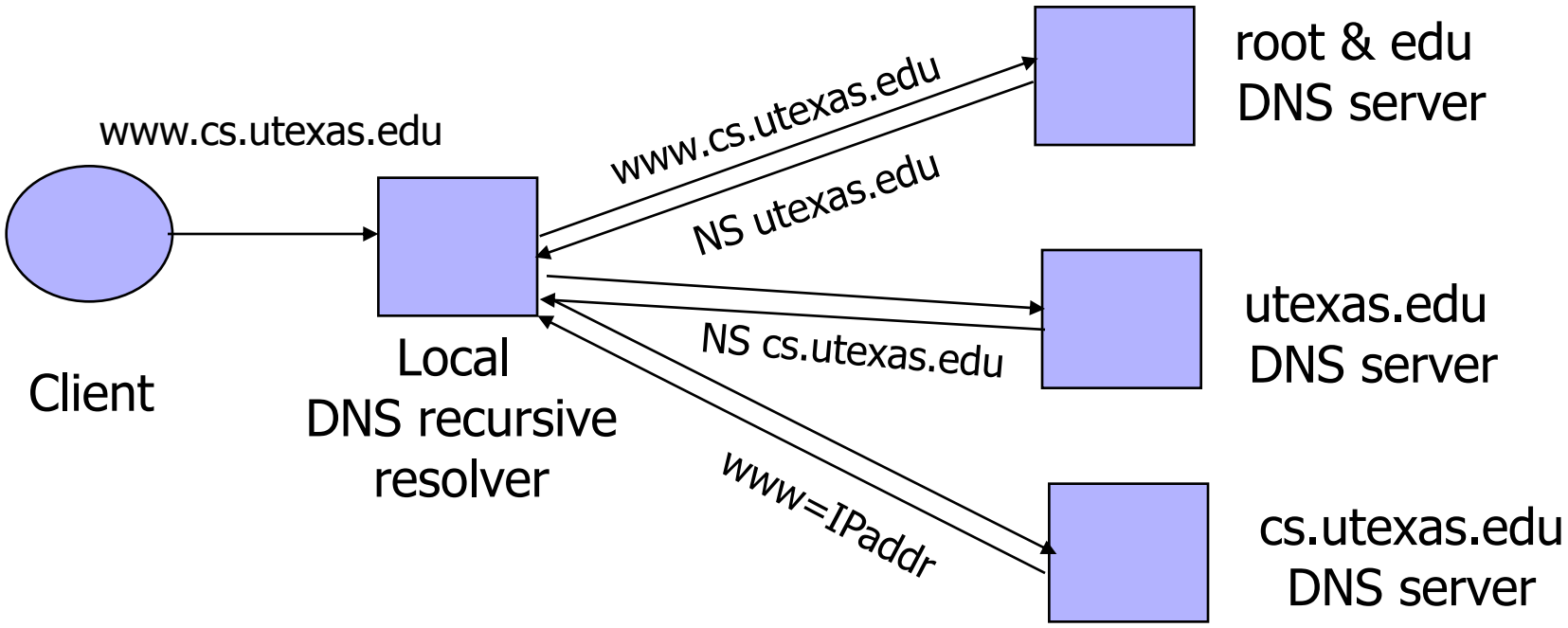
Other BGP Incidents

- May 2003: Spammers hijack unused block of IP addresses belonging to Northrop Grumman
  - Entire Northrop Grumman ends up on spam blacklist
  - Took two months to reclaim ownership of IP addresses
- May 2004: Malaysian ISP hijacks prefix of Yahoo's California data center
- Dec 2004: Turkish ISP advertises routes to the entire Internet, including Amazon, CNN, Yahoo
- Apr 2010: Small Chinese ISP advertises routes to 37,000 networks, incl. Dell, CNN, Apple

DNS: Domain Name Service

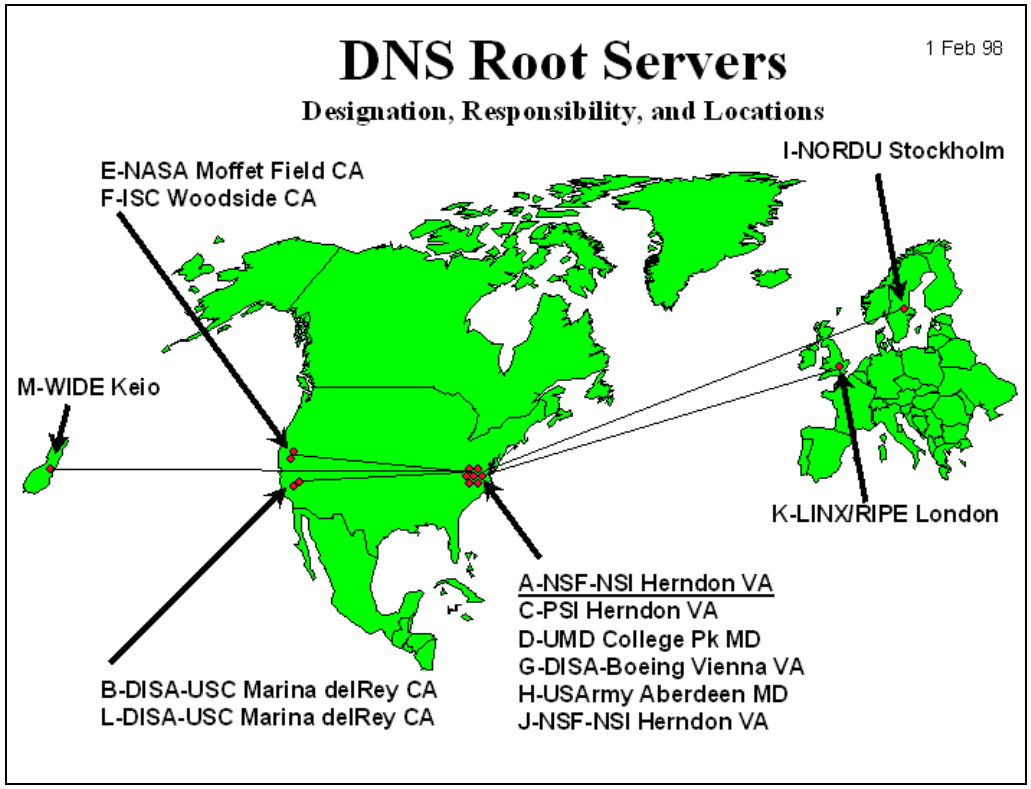
DNS maps symbolic names to numeric IP addresses

(for example, www.cs.utexas.edu ↔ 128.83.120.155)



DNS Root Name Servers

- Root name servers for top-level domains
- Authoritative name servers for subdomains
- Local name resolvers contact authoritative servers when they do not know a name



Feb 6, 2007: Botnet DoS attack on root DNS servers

Turkish net hijack hits big name websites

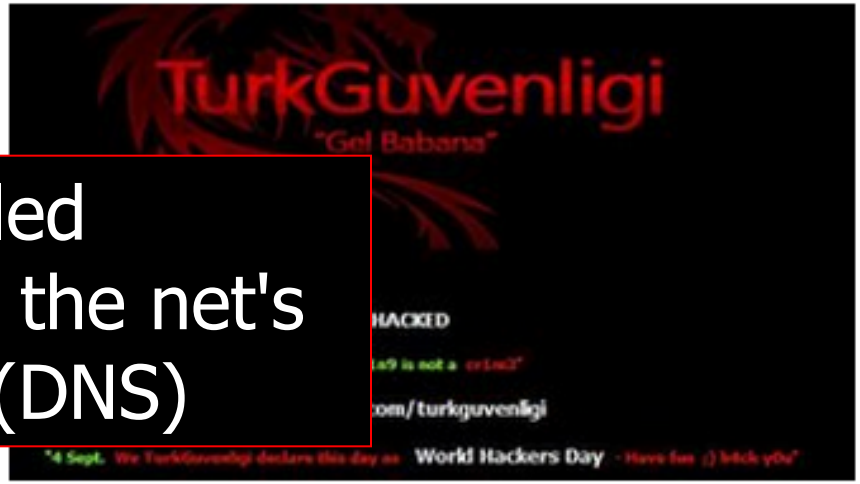
Visitors to the websites of Vodafone, the Daily Telegraph, UPS and four others were re-directed to a site set up by Turkish hackers on Sunda

The di
on cor

Real U
into the IP address of the hackers' site.

No data from
compromise

The hacking
System (DNS



This page greeted many visitors to the sites of
iers

ted Stories

's developer

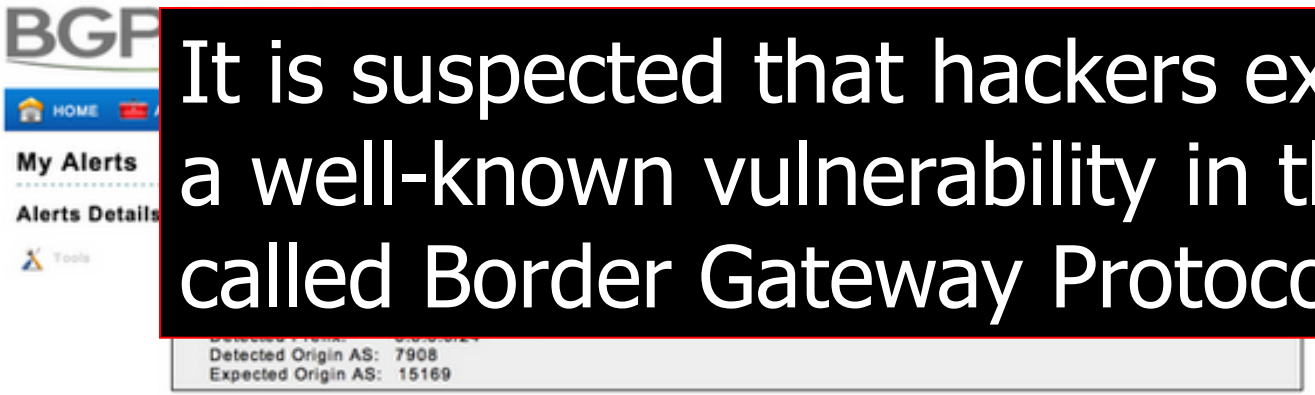
The hacking group, called Turkguvenligi, targeted the net's Domain Name System (DNS)

Turkguvenligi revealed that it got access to the files using a well-established attack method known as SQL injection

Google DNS 8.8.8.8/32 was hijacked for ~22min yesterday, affecting networks in Brazil & Venezuela #bgp #hijack #dns pic.twitter.com/wlBuui8dwO

March 16, 2014

Reply Retweet Favorite More



It is suspected that hackers exploited a well-known vulnerability in the so-called Border Gateway Protocol (BGP)

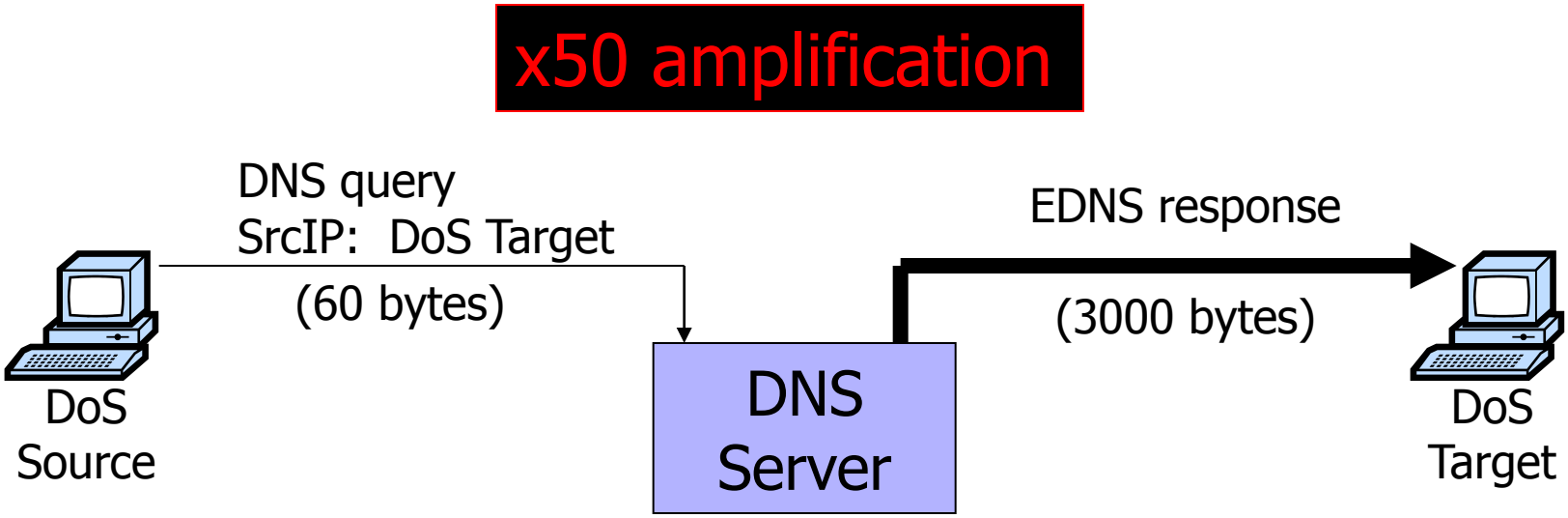
RETWEETS 805 FAVORITES 156

Turkey (2014)


Engin Önder
Follow
#twitter blocked in #turkey tonight. folks are painting #google dns numbers onto the posters of the governing party.
pic.twitter.com/9vQ7NTgotO


DNS: 8.8.8.8
Alternatif: 8.8.4.4

DNS Amplification Attack



2006: 0.58M open resolvers on Internet (Kaminsky-Shiffman)

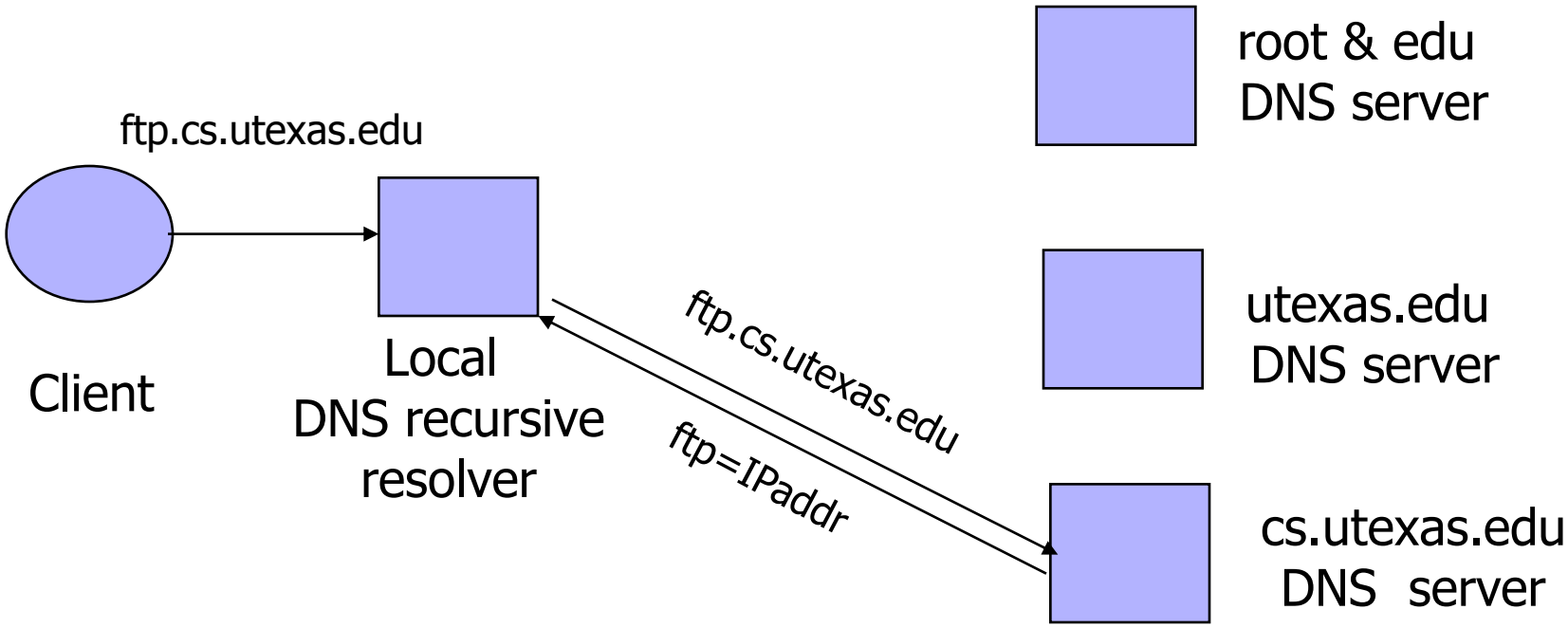
2013: 21.7M open resolvers (openresolverproject.org)

March 2013: 300 Gbps DDoS attack on Spamhaus

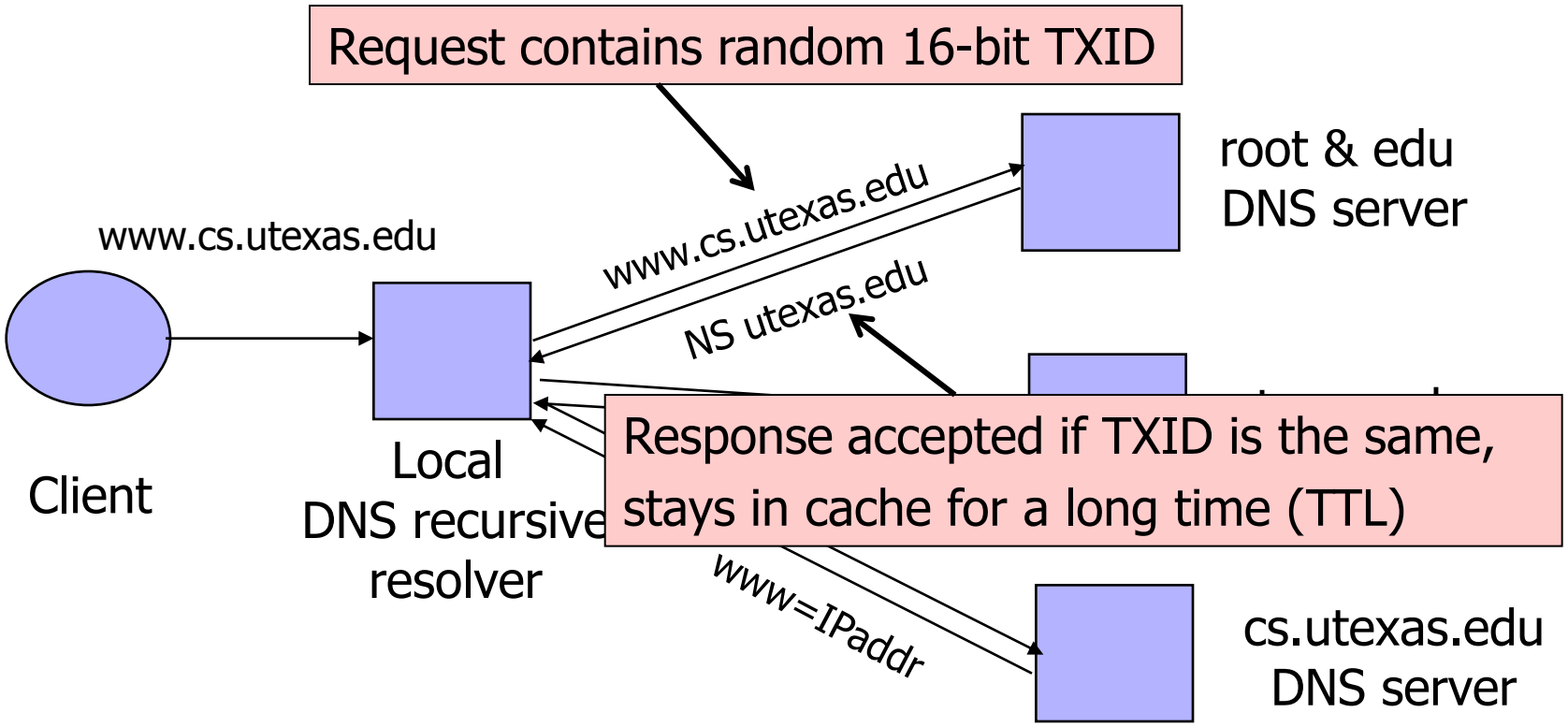
DNS Caching

- DNS responses are cached
  - Quick response for repeated translations
  - Other queries may reuse some parts of lookup
    - NS records identify name servers responsible for a domain
- DNS negative queries are cached
  - Don’t have to repeat past mistakes (misspellings, etc.)
- Cached data periodically times out
  - Lifetime (TTL) of data controlled by owner of data, passed with every record

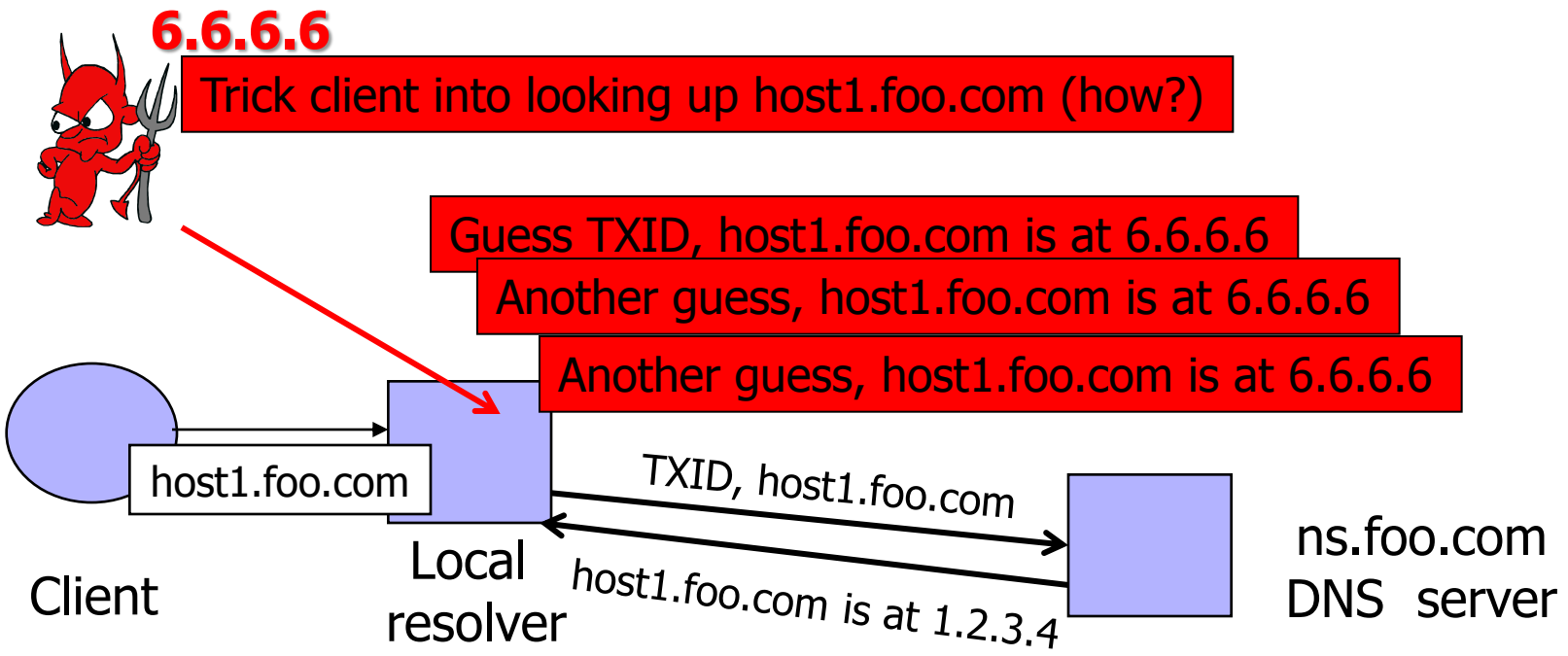
Cached Lookup Example



DNS "Authentication"



DNS Spoofing



Several opportunities to win the race.

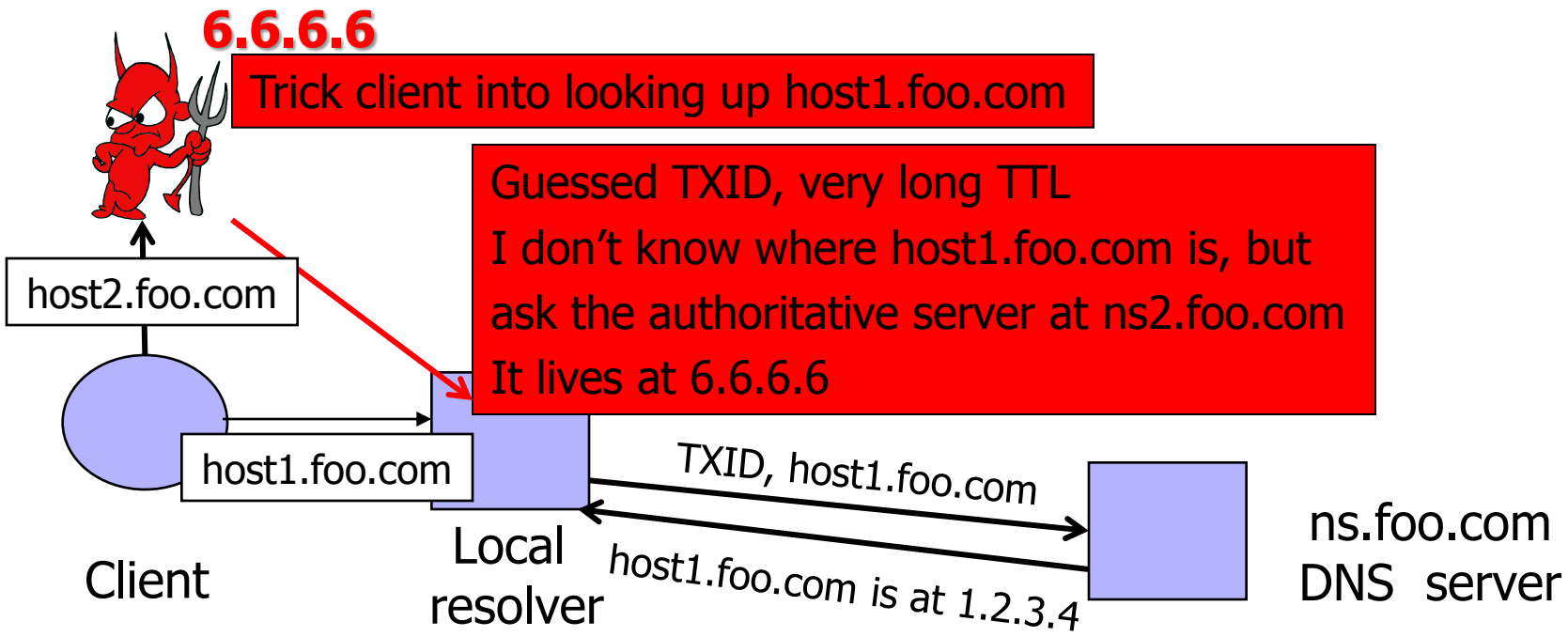
If attacker loses, has to wait until TTL expires…

… but can try again with host2.foo.com, host3.foo.com, etc.

… but what's the point of hijacking host3.foo.com?

Exploiting Recursive Resolving

[Kaminsky]



If win the race, any request for XXX.foo.com will go to 6.6.6.6

The cache is poisoned… for a very long time!

No need to win future races!

If lose, try again with <ANYTHING>.foo.com

Triggering a Race

- Any link, any image, any ad, anything can cause a DNS lookup
  - No JavaScript required, though it helps
- Mail servers will look up what bad guy wants
  - On first greeting: HELO
  - On first learning who they're talking to: MAIL FROM
  - On spam check (oops!)
  - When trying to deliver a bounce
  - When trying to deliver a newsletter
  - When trying to deliver an actual response from an actual employee

Reverse DNS Spoofing

- Trusted access is often based on host names
  - Example: permit all hosts in .rhosts to run remote shell
- Network requests such as rsh or rlogin arrive from numeric source addresses
  - System performs reverse DNS lookup to determine requester's host name and checks if it's in .rhosts
- If attacker can spoof the answer to reverse DNS query, he can fool target machine into thinking that request comes from an authorized host
  - No authentication for DNS responses and typically no double-checking (numeric → symbolic → numeric)

Pharming

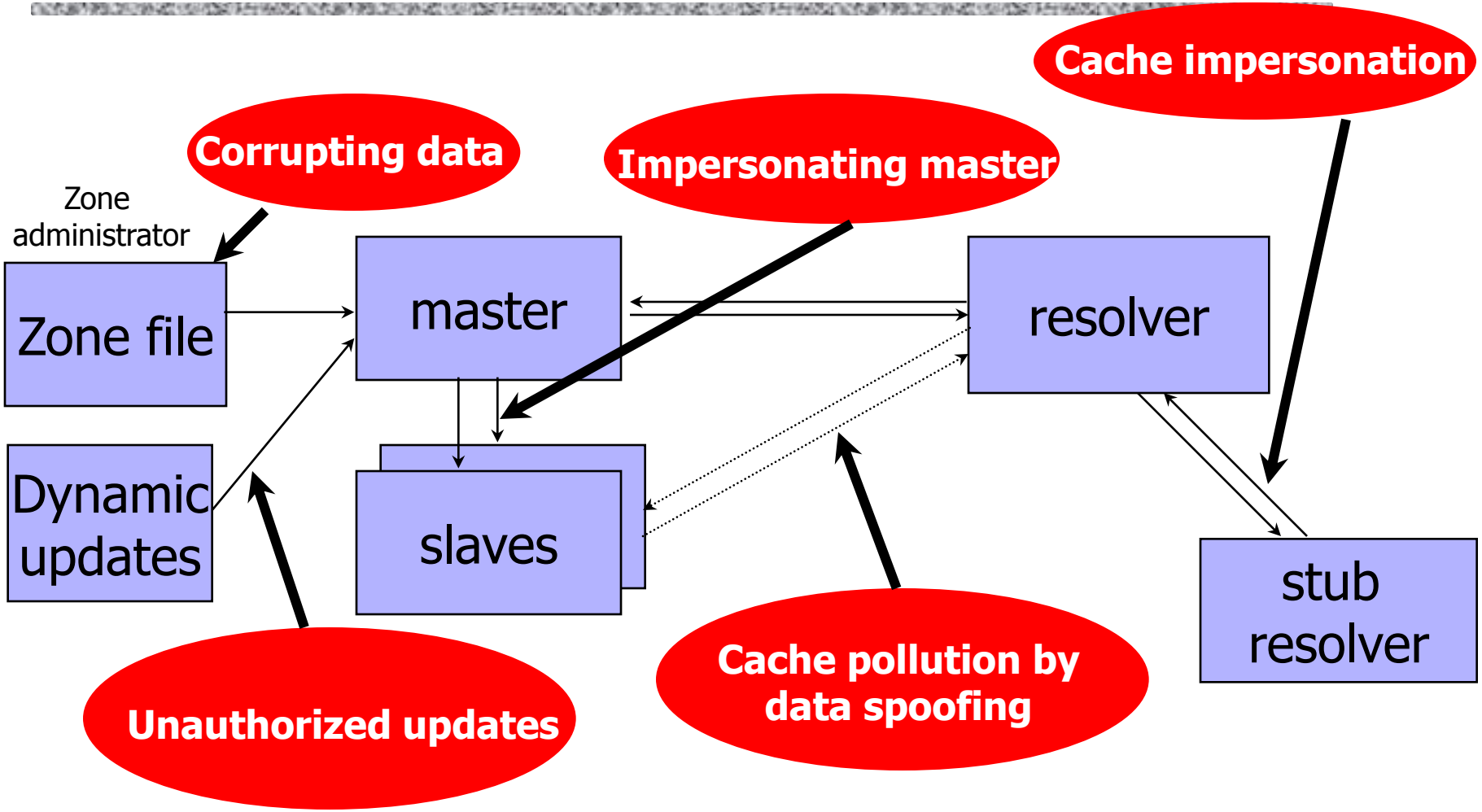
- Many anti-phishing defenses rely on DNS
- Can bypass them by poisoning DNS cache and/or forging DNS responses
  - Browser: "give me the address of www.paypal.com"
  - Attacker: "sure, it's 6.6.6.6" (attacker-controlled site)
- Dynamic pharming
  - Provide bogus DNS mapping for a trusted server, trick user into downloading a malicious script
  - Force user to download content from the real server, temporarily provide correct DNS mapping
  - Malicious script and content have the same origin!

Other DNS Vulnerabilities

- DNS implementations have vulnerabilities
  - Reverse query buffer overrun in old releases of BIND
  - MS DNS for NT 4.0 crashes on chargen stream
- Denial of service
  - Oct '02: ICMP flood took out 9 root servers for 1 hour
- Can use "zone transfer" requests to download DNS database and map out the network
  - "The Art of Intrusion": NYTimes.com and Excite@Home
  - Solution: block port 53 on corporate name servers

See http://cr.yp.to/djbdns/notes.html

DNS Vulnerabilities: Summary



Solving the DNS Spoofing Problem

- Long TTL for legitimate responses
  - Does it really help?
- Randomize port in addition to TXID
  - 32 bits of randomness, makes it harder for attacker to guess TXID+port
- DNSSEC
  - Cryptographic authentication of host-address mappings

DNSSEC

- Goals: authentication and integrity of DNS requests and responses
- PK-DNSSEC (public key)
  - DNS server signs its data (can be done in advance)
  - How do other servers learn the public key?
- SK-DNSSEC (symmetric key)
  - Encryption and MAC: $E_k(m, MAC(m))$
  - Each message contains a nonce to avoid replay
  - Each DNS node shares a symmetric key with its parent
  - Zone root server has a public key (hybrid approach)

Domain Hijacking and Other Risks

- Spoofed ICANN registration and domain hijacking
  - Authentication of domain transfers based on email addr
  - Aug ’04: teenager hijacks eBay’s German site
  - Jan ’05: hijacking of panix.com (oldest ISP in NYC)
    - "The ownership of panix.com was moved to a company in Australia, the actual DNS records were moved to a company in the United Kingdom, and Panix.com's mail has been redirected to yet another company in Canada."
  - Many other domain theft attacks
- Misconfiguration and human error