SOEN 321
Information Systems Security

Firewalls, IDS, DDoS

Note: Many of these slides are collected/modified from different books and online resources

Firewall Goals

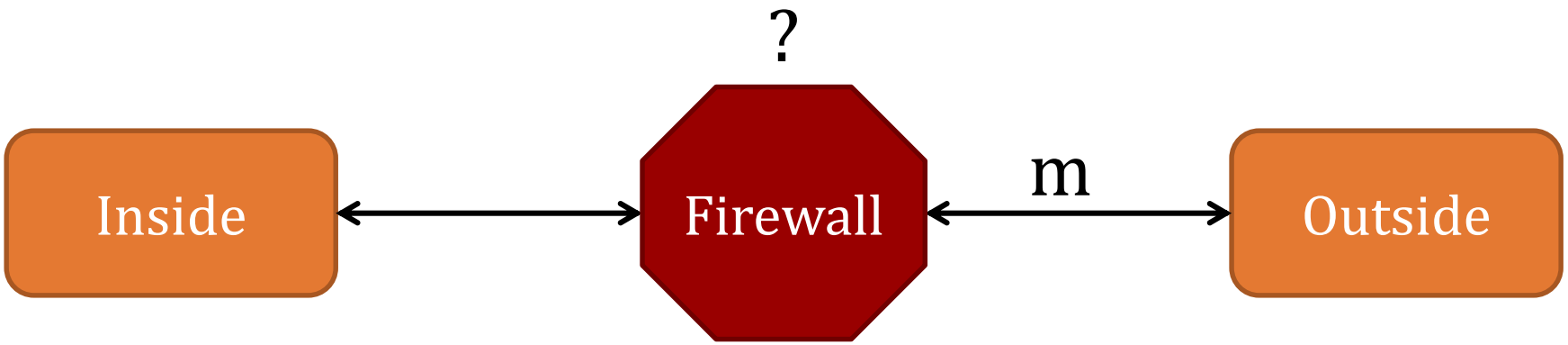
Provide defense in depth by

1. Blocking attacks against hosts and services
2. Control traffic between zones of trust

Firewalls Dimensions

1. Host vs. Network
2. Stateless vs. Stateful
3. Network Layer

Logical Viewpoint

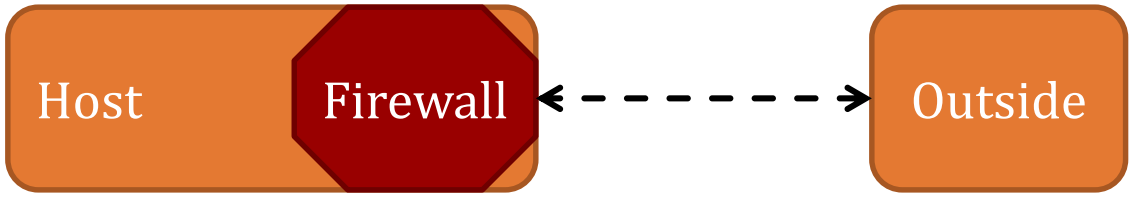


For each message m, either:

- Allow
- Block (by dropping or by dropping and sending rejection notice)
- Queue

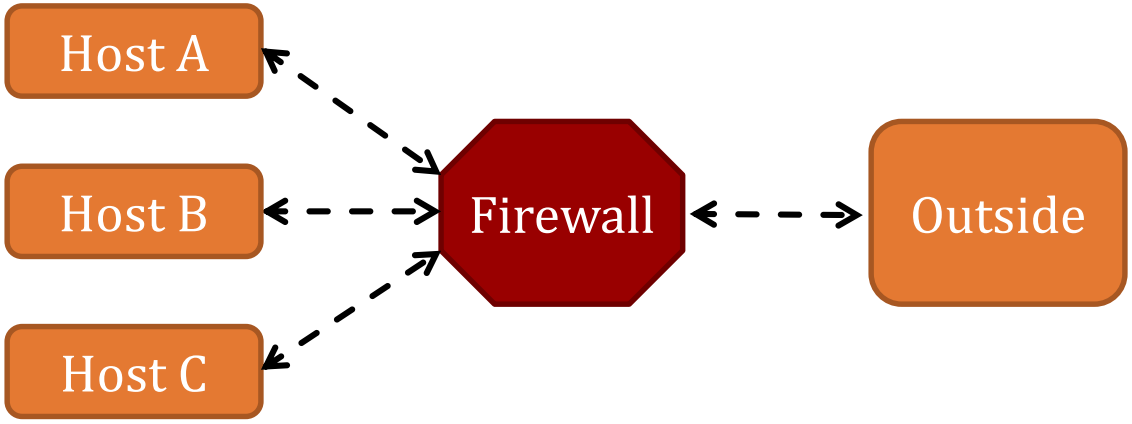
Placement

Host-based Firewall



Features:
- Faithful to local configuration
- Travels with you

Network-Based Firewall



Features:
- Protect whole network
- Can make decisions on all of traffic (traffic-based anomaly)

Parameters

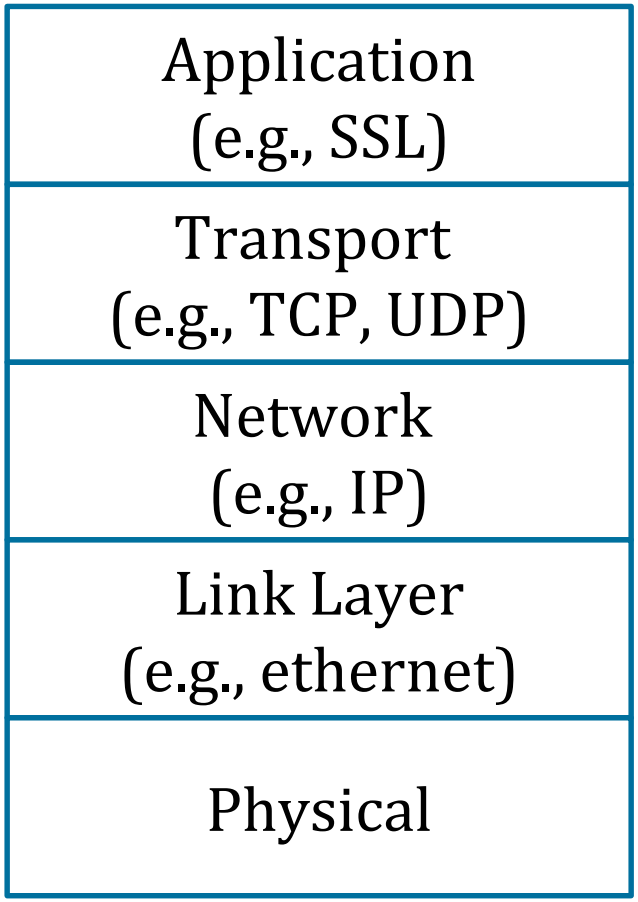
Types of Firewalls

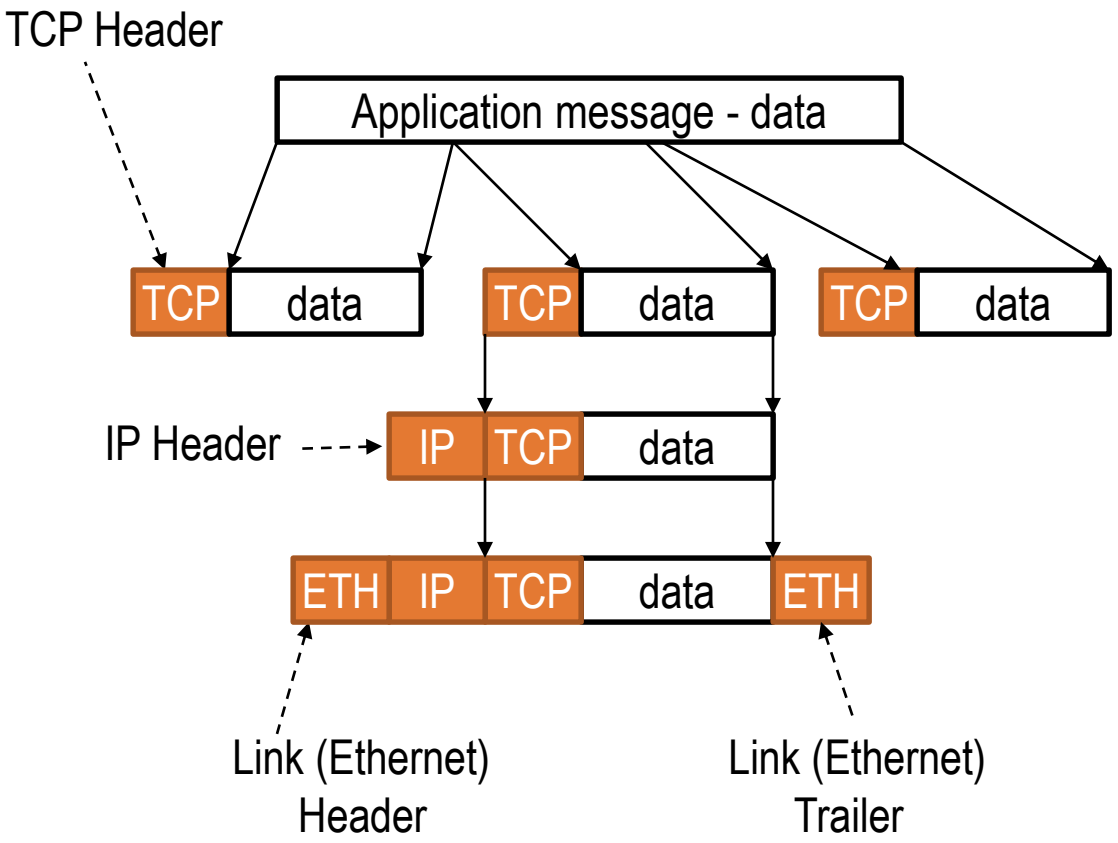
1. Packet Filtering
2. Stateful Inspection
3. Application proxy

Policies

1. Default allow
2. Default deny

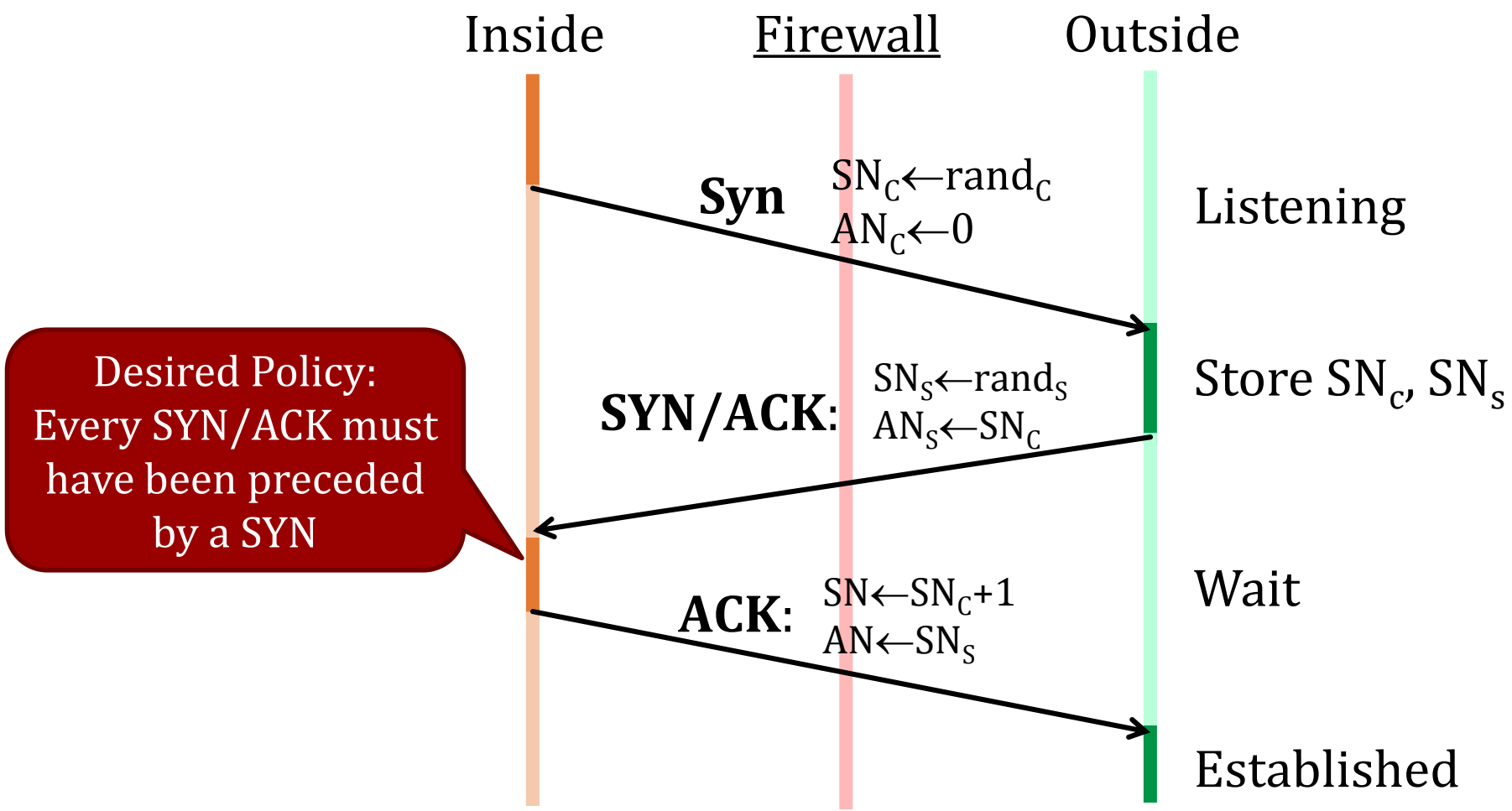
Recall: Protocol Stack





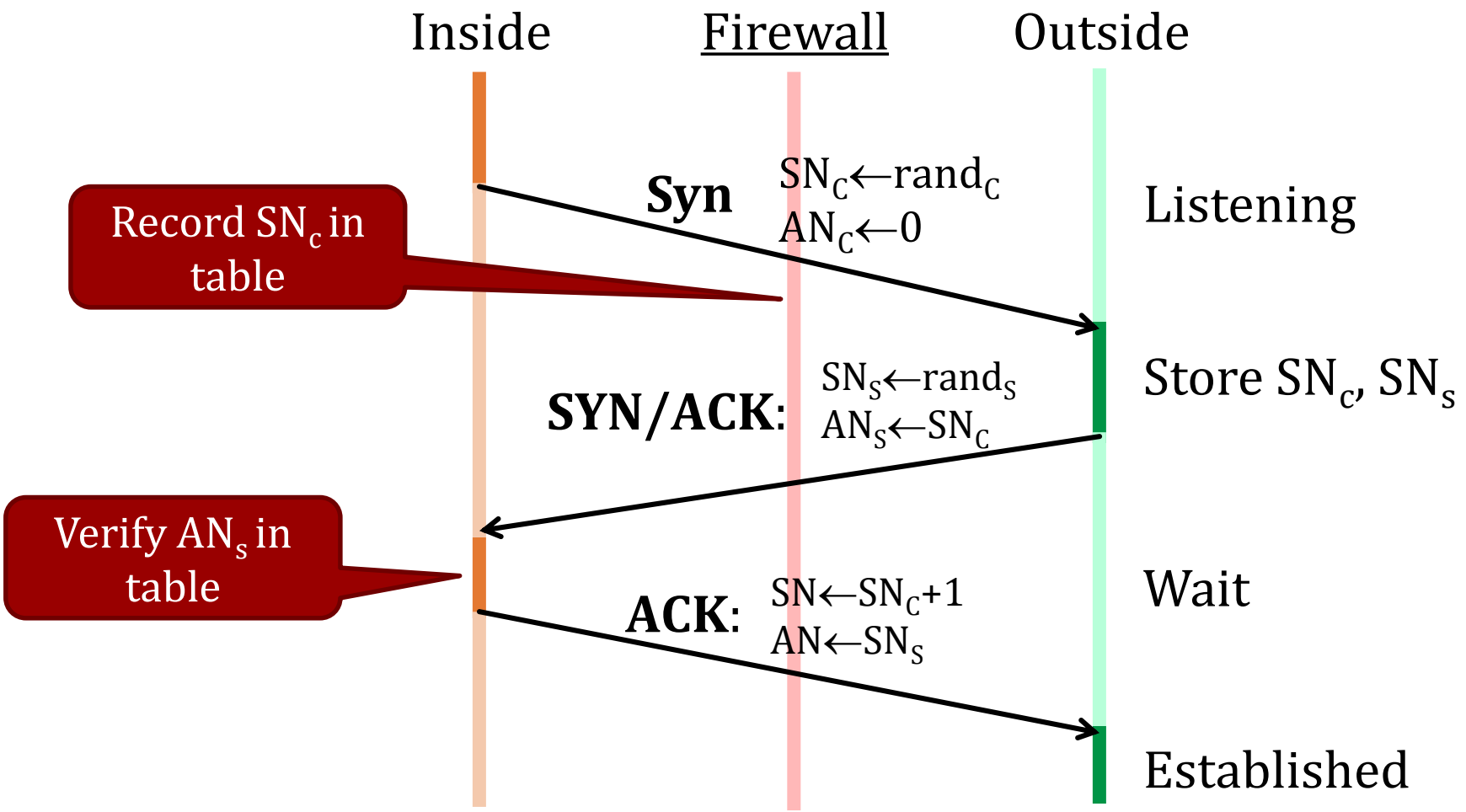
The Need to Keep State

Example: TCP Handshake



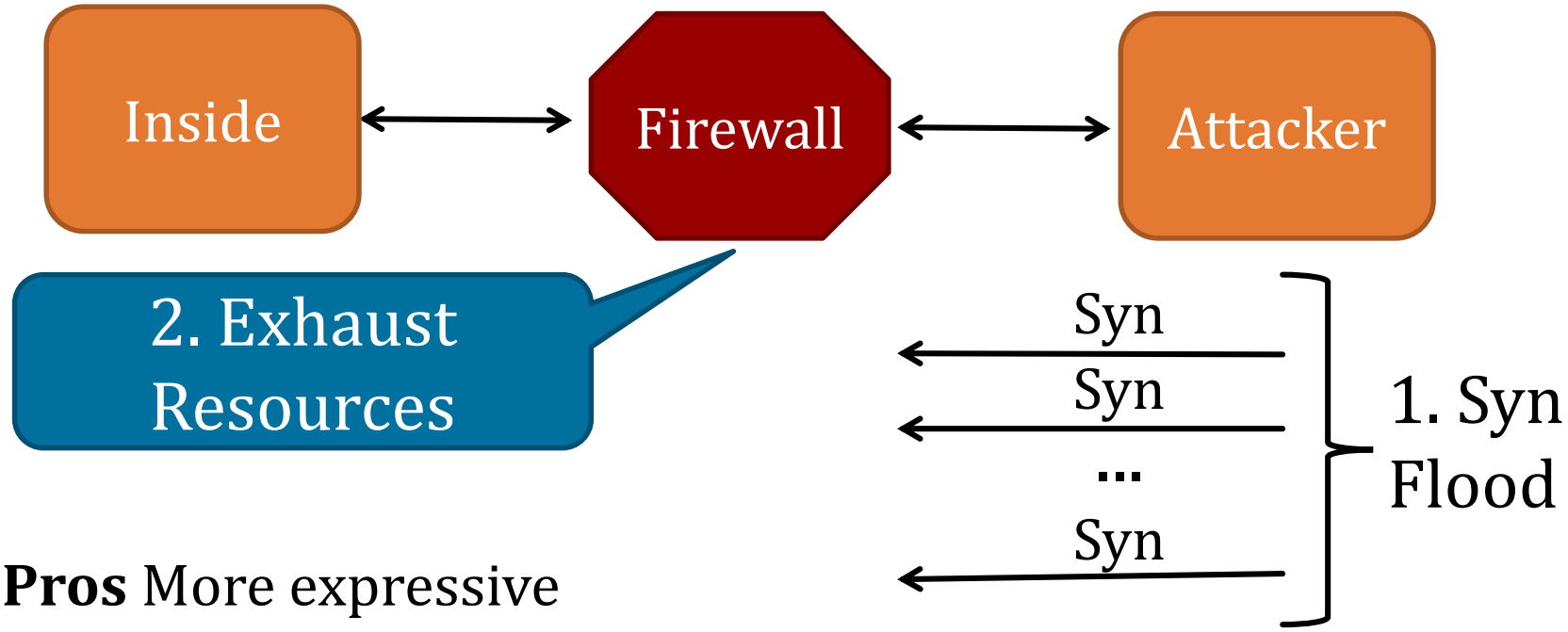
Stateful More Expressive

Example: TCP Handshake



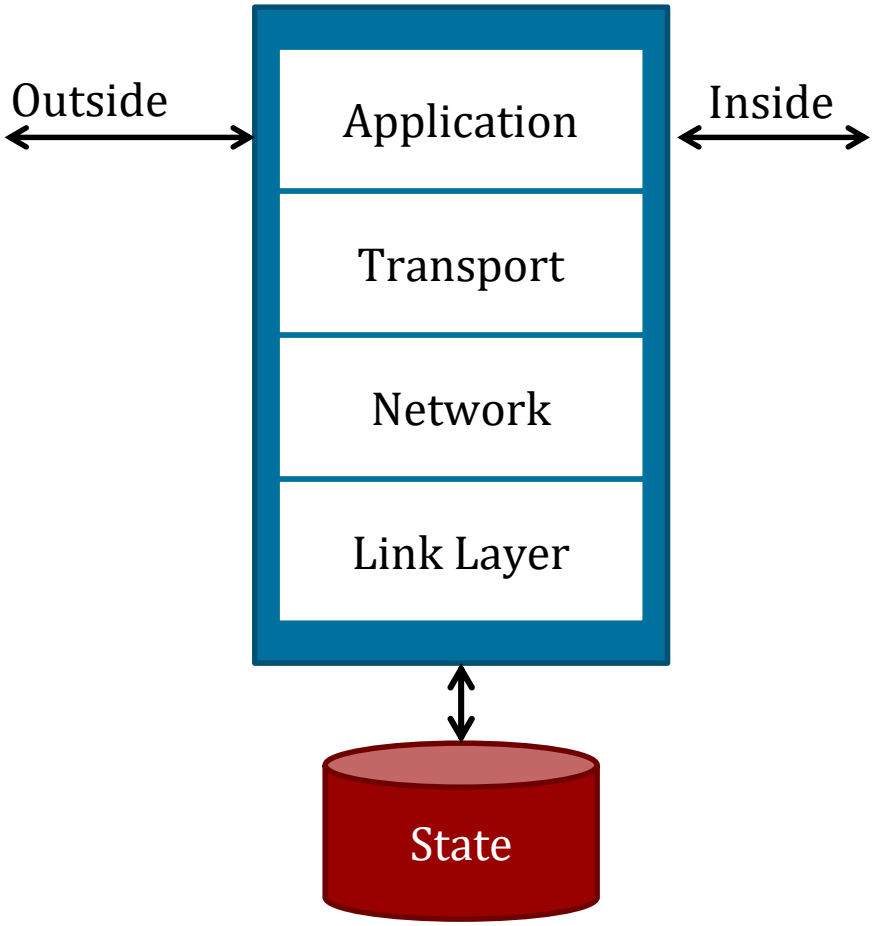
State Holding Attack

Assume stateful TCP policy



- **Pros** More expressive
- **Cons** State-holding attack

Application Firewall



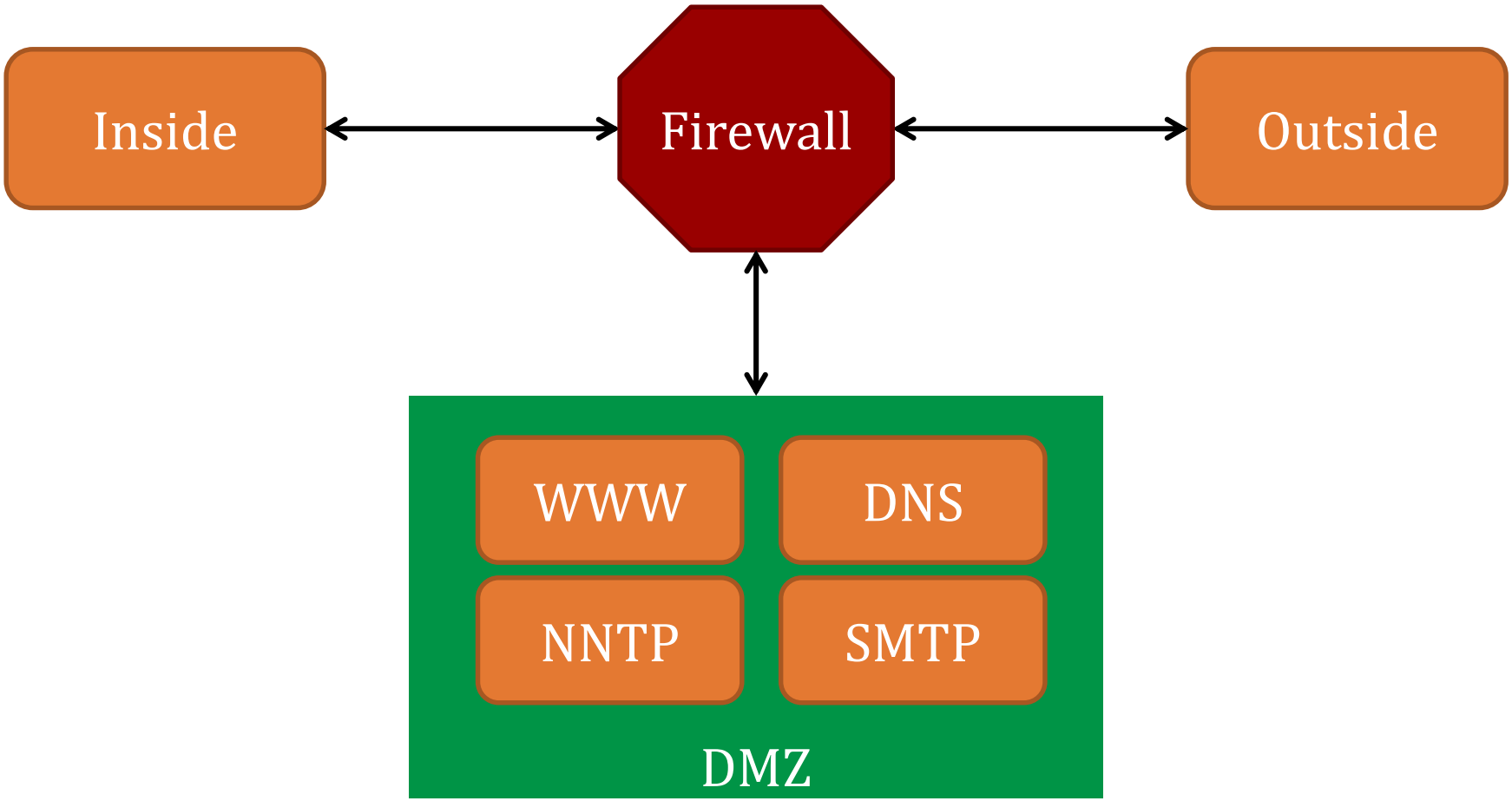
Check protocol messages directly

Example:

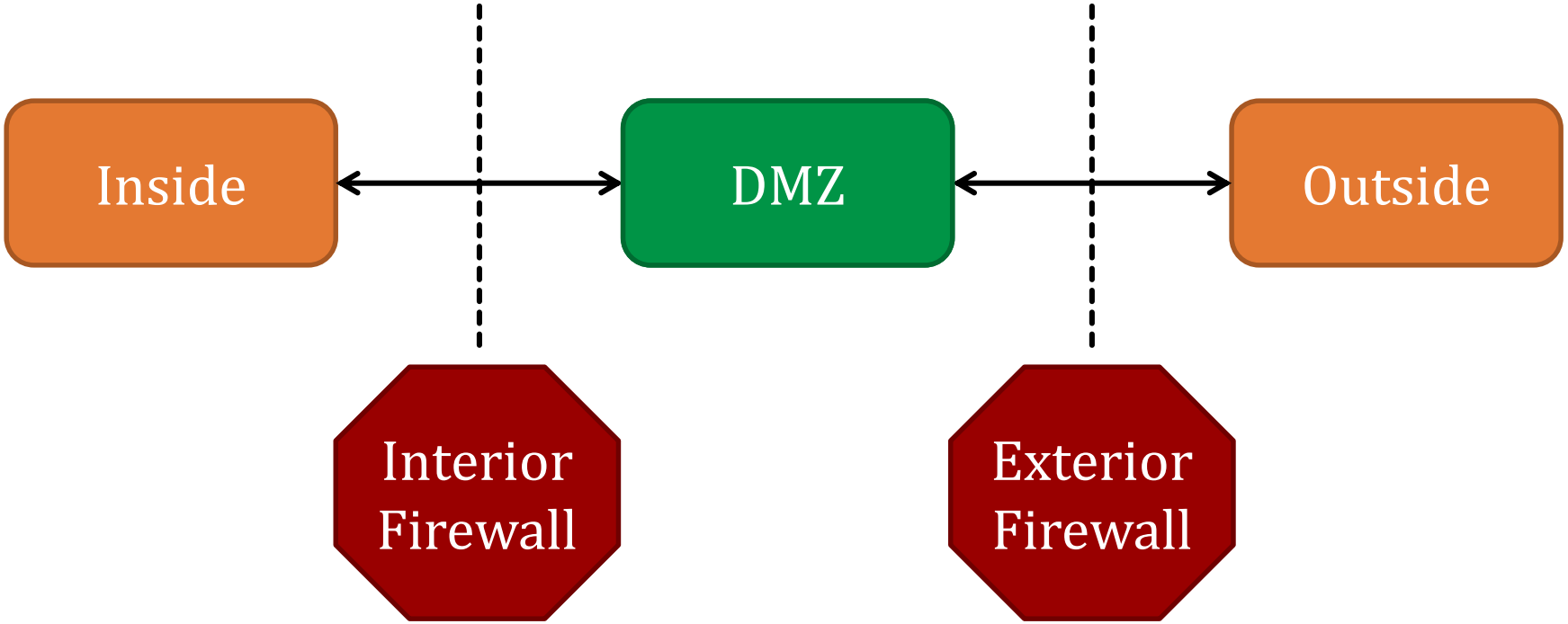
– SMTP virus scanner

Firewall Placement

Demilitarized Zone (DMZ)

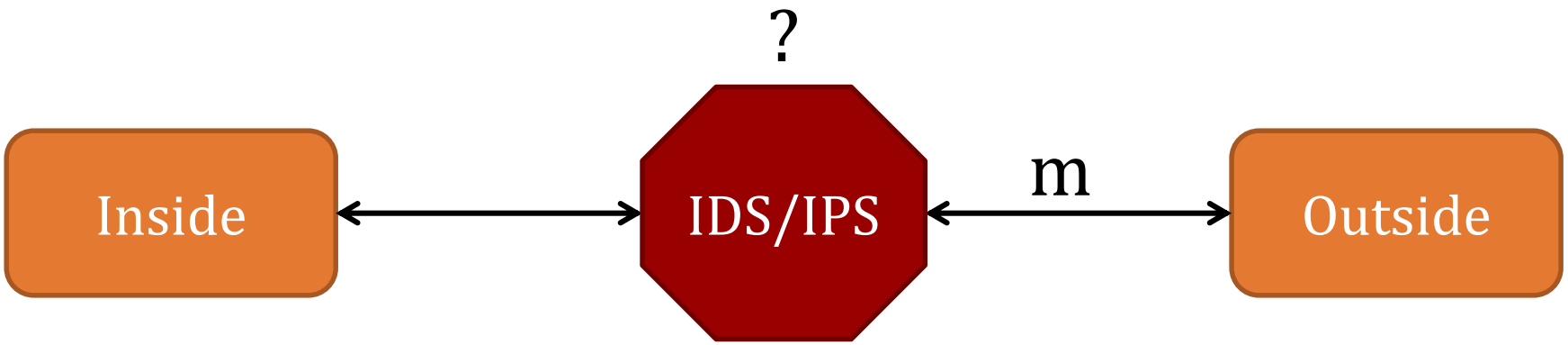


Dual Firewall



Intrusion Detection and Prevetion Systems

Logical Viewpoint



For each message m, either:

- Report *m* (IPS: drop or log)
- Allow *m*
- Queue

IDS Overview

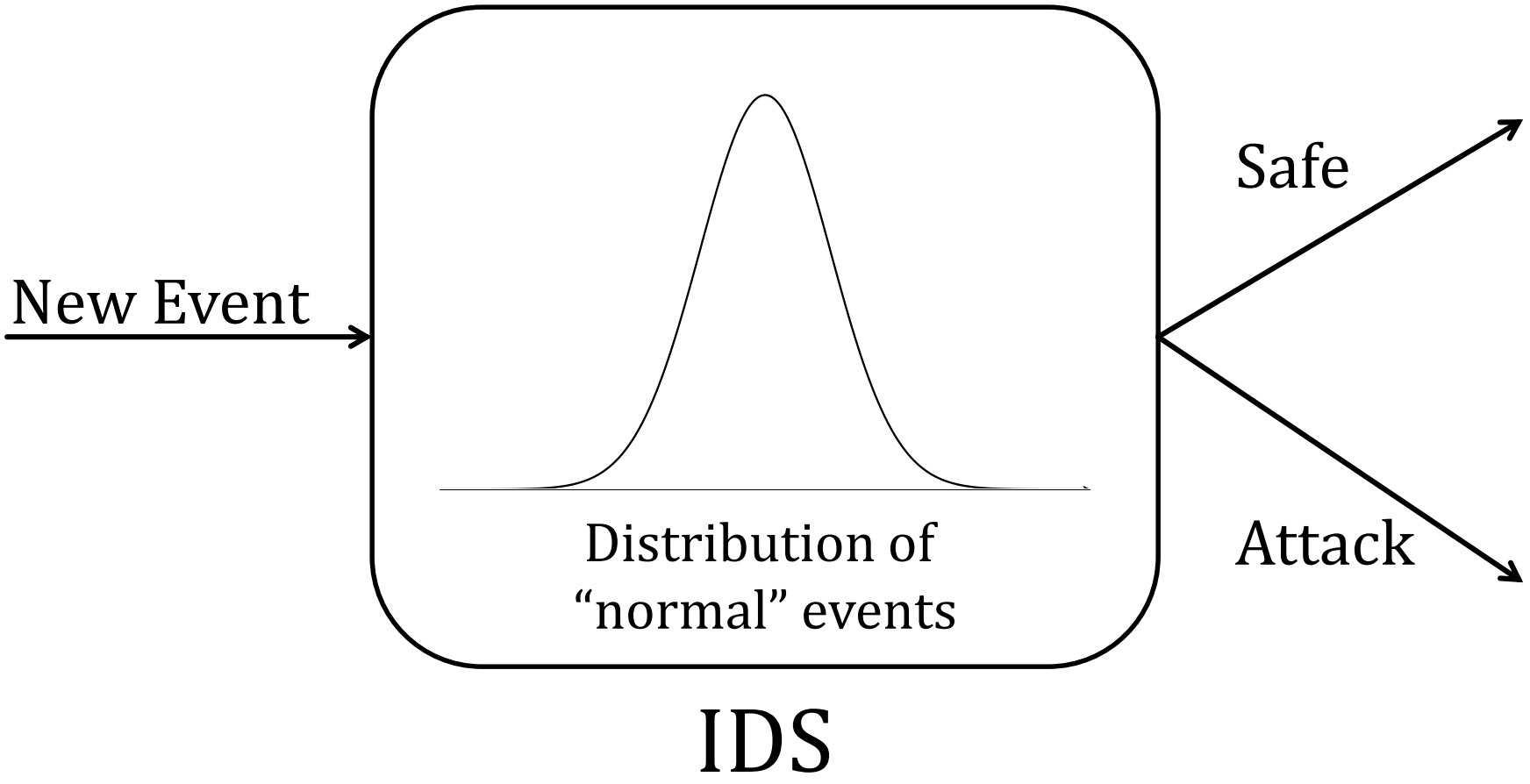
- Approach: Policy (Signature) vs. Anomaly
- Location: Network vs. Host
- Action: Detect vs. Prevent

Policy-Based IDS

Use pre-determined rules to detect attacks

Examples: Regular expressions (snort),
Cryptographic hash (tripwire, snort)

Anomaly Detection



Anomaly Detection

Pros

- Does not require pre-determining policy (an “unknown” threat)

Cons

- Learning distributions is hard

Detection Theory

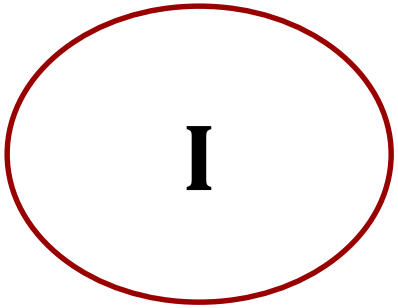
Ω

Let Ω be the set of all possible events.
For example:

- Audit records produced on a host
- Network packets seen

Ω

Example: IDS Received 1,000,000 packets. 20 of them corresponded to an intrusion.
The *intrusion rate* Pr[I] is:
Pr[I] = 20/1,000,000 = .00002

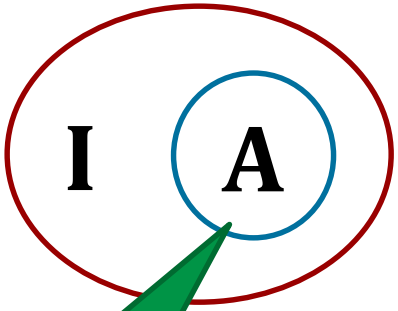


Set of intrusion events **I**

Intrusion Rate:

$$\Pr[I] = \frac{|I|}{|\Omega|}$$

Ω

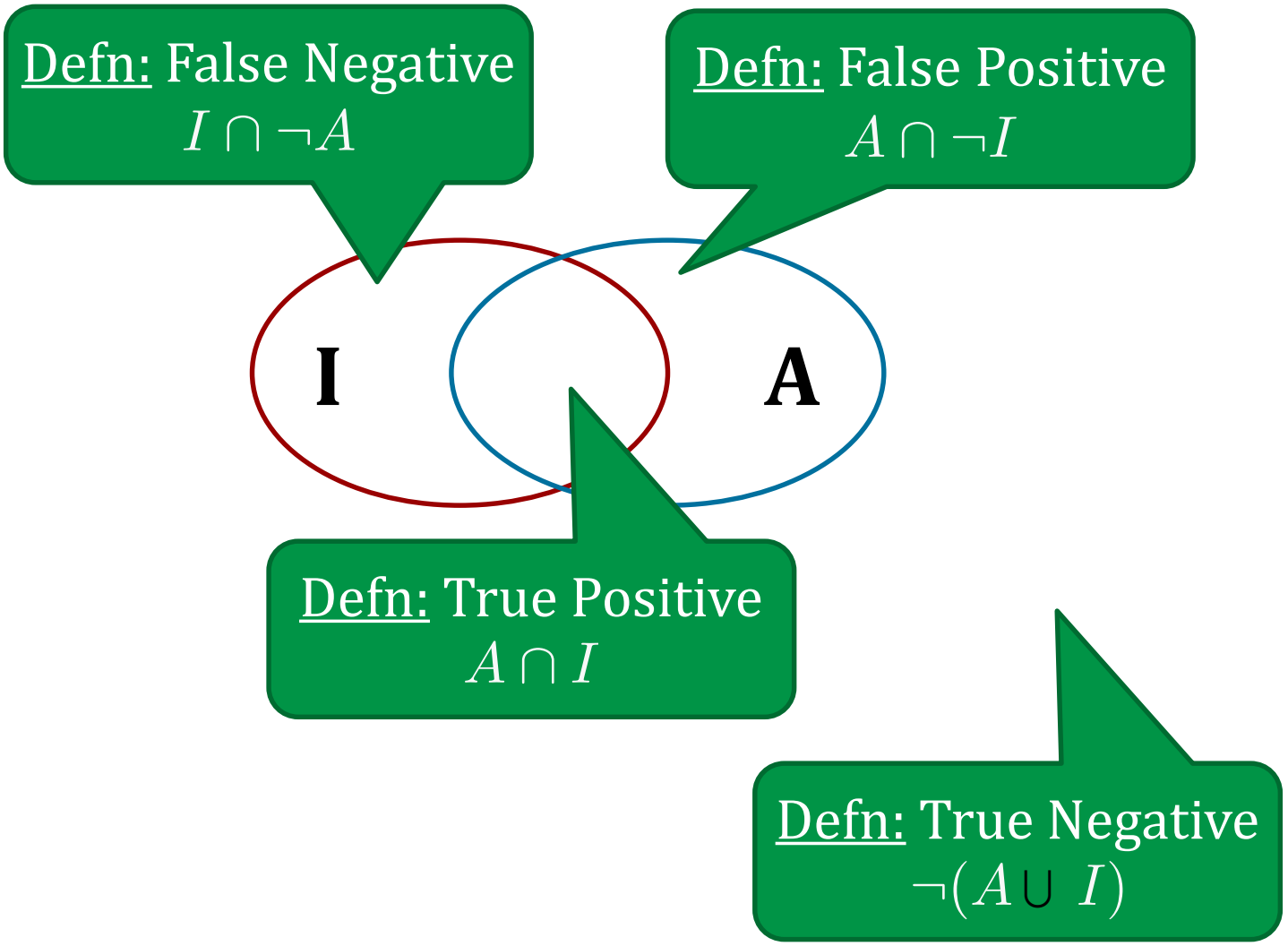


Set of alerts **A**

Alert Rate:

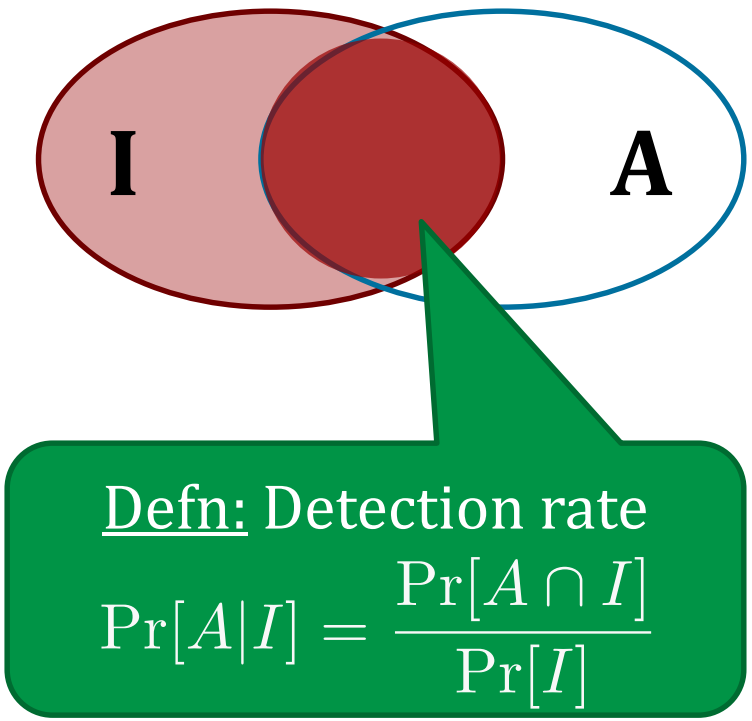
$$\Pr[A] = \frac{|A|}{|\Omega|}$$

Ω


Defn: False Negative
$I \cap \neg A$
Defn: False Positive
$A \cap \neg I$
I
A
Defn: True Positive
$A \cap I$
Defn: True Negative
$\neg(A \cup I)$

Ω

Think of the detection rate as the set of *intrusions raising an alert* normalized by the *set of <u>all</u> intrusions*.



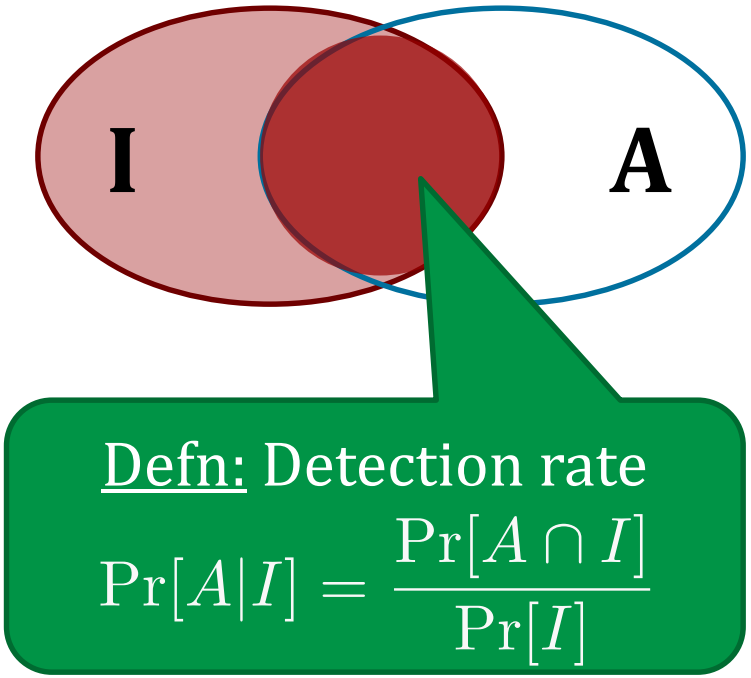
Suppose:

$$|\Omega| = \frac{20}{1,000,000}, |I| = 20$$

$$|I \cap A| = 18, |A| = 22$$

What is the detection rate?

Ω

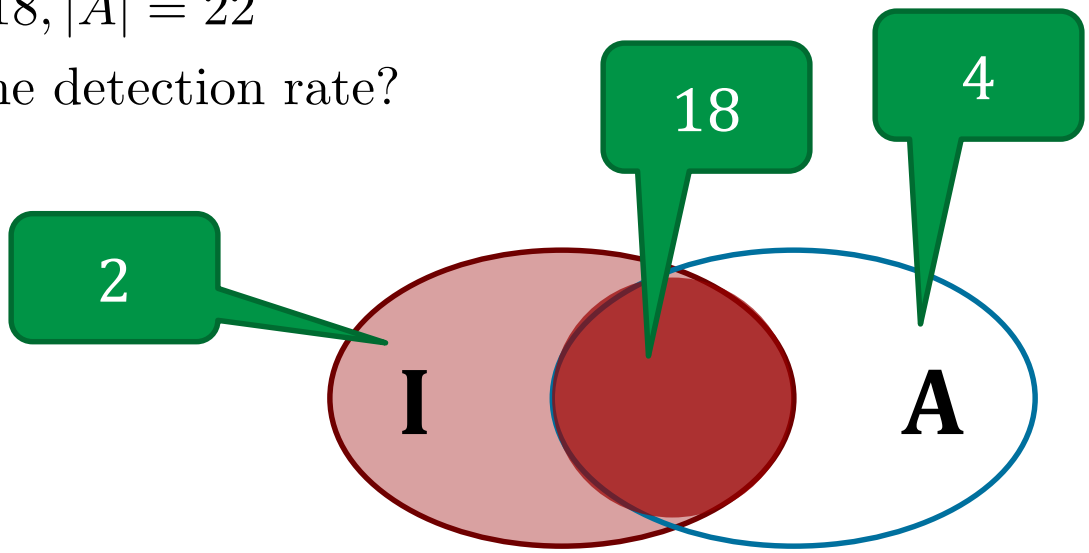


Suppose: |Ω|=1,000,000, |I|=20

Ω

$|I \cap A| = 18, |A| = 22$

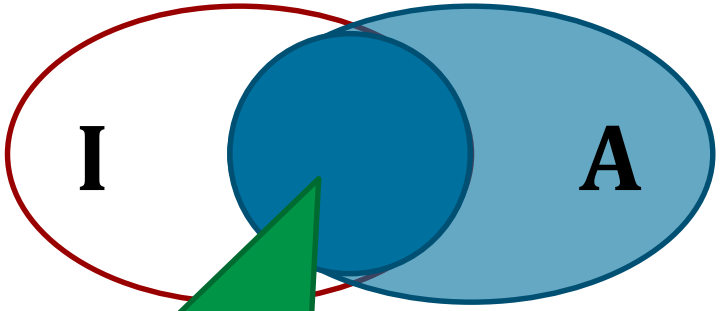
What is the detection rate?



$$\Pr[A|I] = \frac{\Pr[A \cap I]}{\Pr[I]} = \frac{18/1,000,000}{20/1,000,000} = .90 = 90\%$$

Ω

Think of the Bayesian detection rate as the set of *intrusions raising an alert* normalized by the *set of all alerts*. (vs. detection rate which normalizes on intrusions.)



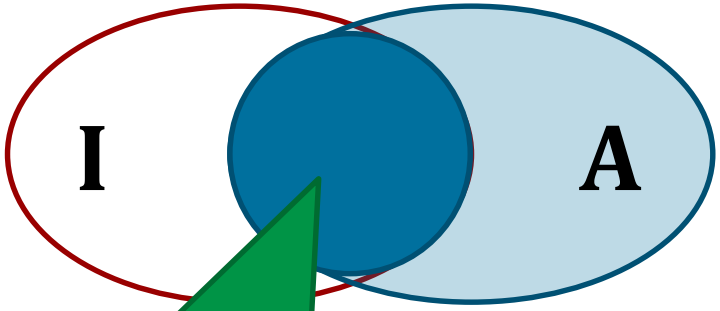
Defn: *Bayesian* Detection rate

$$\Pr[I|A] = \frac{\Pr[A \cap I]}{\Pr[A]}$$

! Crux of IDS usefulness

Ω

Bayesian detection rate is the probability an alert signifies an intrusion.
The crux of deciding whether a detection system is useful often comes down to this equation.



Defn: *Bayesian* Detection rate

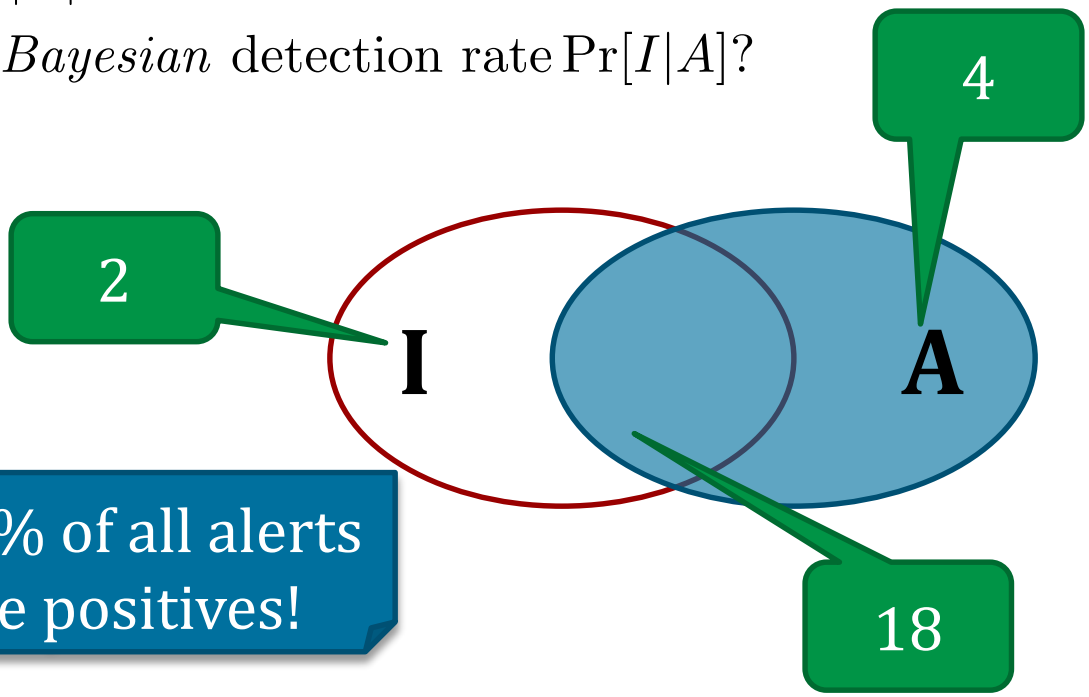
$$\Pr[I|A] = \frac{\Pr[A \cap I]}{\Pr[A]}$$

Suppose:

$|\Omega| = 1,000,000, |I| = 20$

$|I \cap A| = 18, |A| = 22$

What is the *Bayesian* detection rate $\Pr[I|A]$?



About 18% of all alerts are false positives!

$$\Pr[I|A] = \frac{\Pr[A \cap I]}{\Pr[A]} = \frac{0.000018}{0.000022} = \overline{.81} \approx 82\%$$

Challenge

We're often given the detection rate and know the intrusion rate, and want to calculate the Bayesian detection rate

- 99% accurate IDS

Calculating Bayesian Detection Rate

Fact: $\Pr[A] = \Pr[I] * \Pr[A|I] + \Pr[\neg I] * \Pr[A|\neg I]$

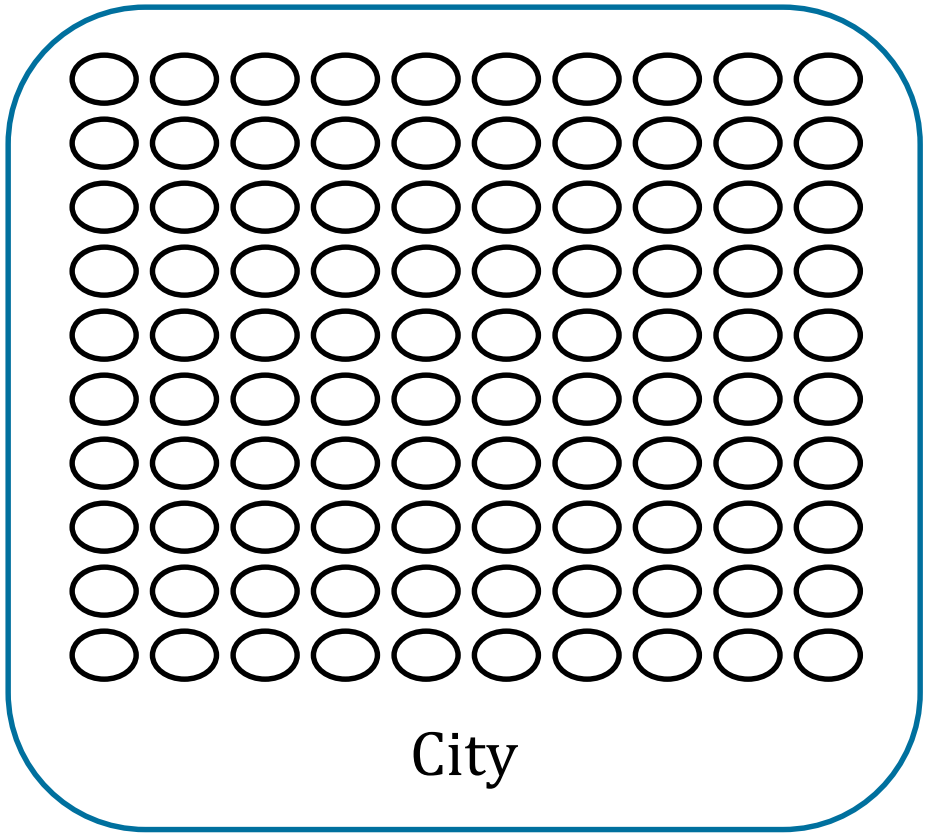
So to calculate the Bayesian detection rate:

$$\Pr[I|A] = \frac{\Pr[A \cap I]}{\Pr[A]}$$

One way is to compute:

$$\Pr[I|A] = \frac{\Pr[A \cap I]}{\Pr[I] * \Pr[A|I] + \Pr[\neg I] * \Pr[A|\neg I]}$$

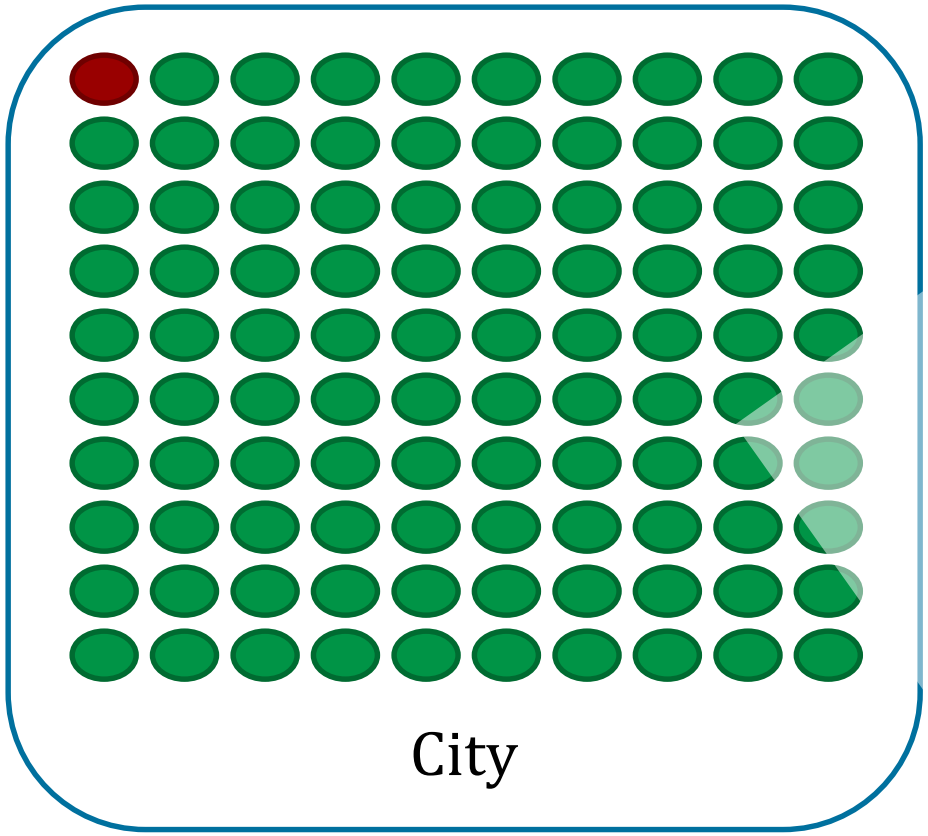
Example



(this times 10)

- 1,000 people in the city
- 1 is a terrorists, and we have their pictures. Thus the *base rate* of terrorists is 1/1000
- Suppose we have a new terrorist facial recognition system that is 99% accurate, i.e.:
  - 99/100 times when someone is a terrorist there is an alarm
  - For every 100 good guys, the alarm only goes off once.
- An alarm went off. Is the suspect really a terrorist?

Example

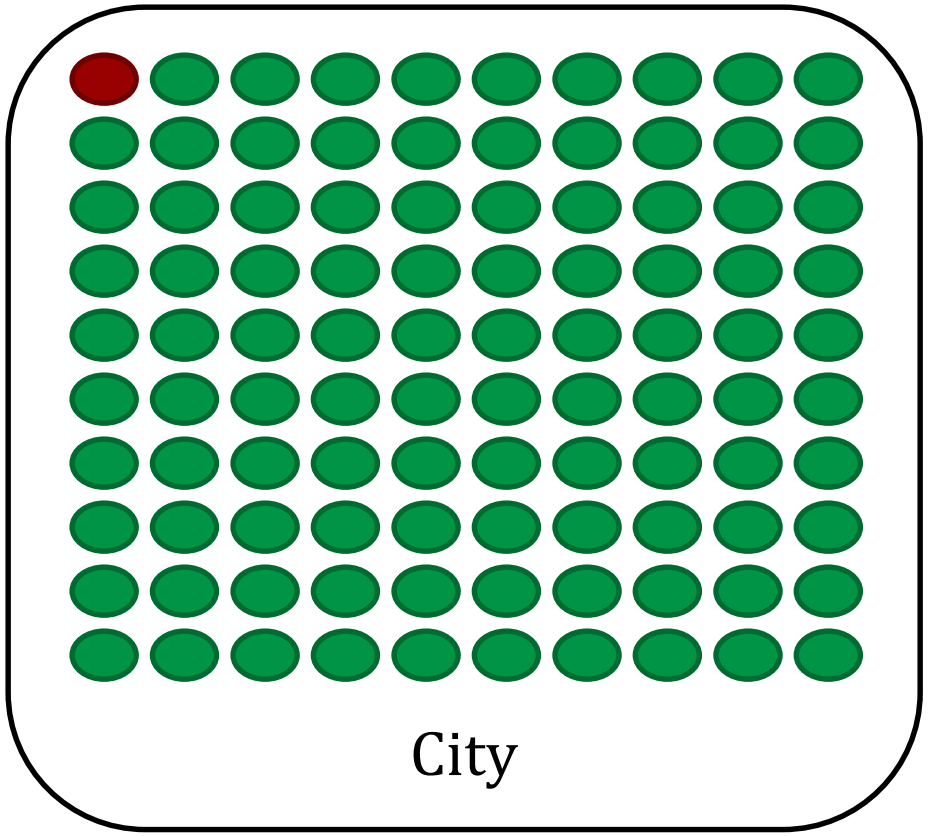


(this times 10)

Answer: The facial recognition system is 99% accurate. That means there is only a 1% chance the guy is not the terrorist.

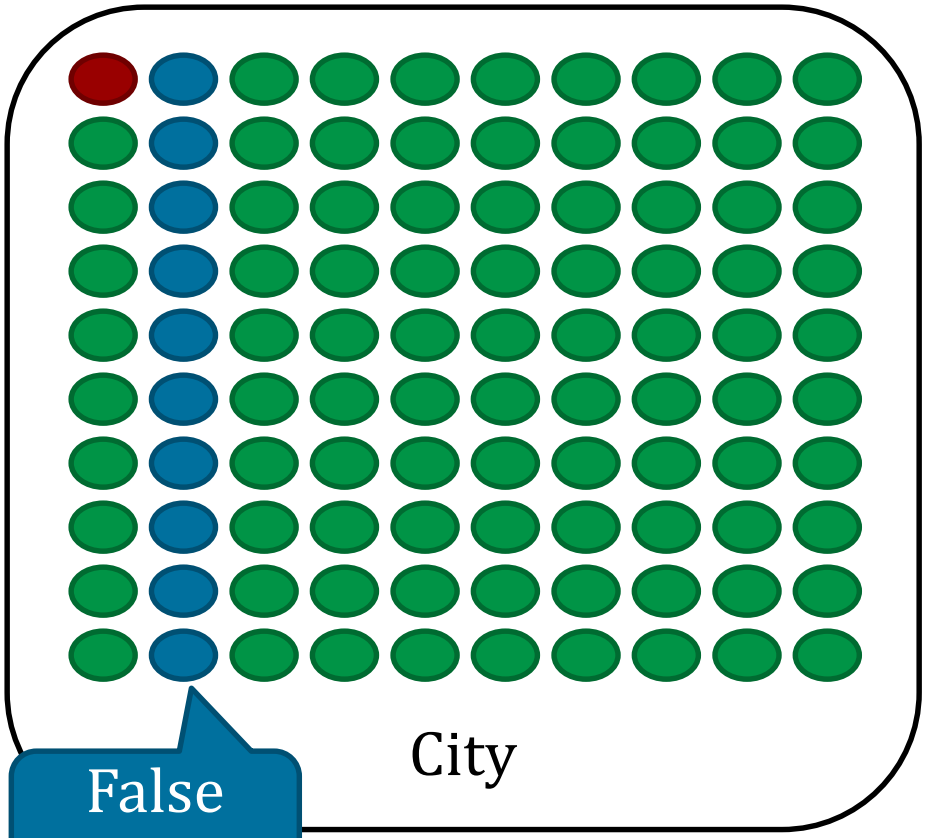
Wrong!

Formalization



(this times 10)

- 1 is terrorists, and we have their pictures. Thus the *base rate* of terrorists is 1/1000. **P[T] = 0.001**
- 99/100 times when someone is a terrorist there is an alarm. **P[A|T] = .99**
- For every 100 good guys, the alarm only goes off once. **P[A | not T] = .01**
- Want to know **P[T|A]**



- 1 is terrorists, and we have their pictures. Thus the *base rate* of terrorists is 1/1000. **P[T] = 0.001**
- 99/100 times when someone is a terrorist there is an alarm. **P[A|T] = .99**
- For every 100 good guys, the alarm only goes off once. **P[A | not T] = .01**
- Want to know **P[T|A]**

Guesses?

$$\text{Have: } \Pr[T] = 0.001$$

$$\Pr[A|T] = .99, \Pr[A|\neg T] = .01$$

$$\text{Want to calculate: } \Pr[T|A] = \frac{\Pr[T \cap A]}{\Pr[A]}$$

Unknown

Unknown

$$\text{Have: } \Pr[T] = 0.001$$

$$\Pr[A|T] = .99, \Pr[A|\neg T] = .01$$

$$\text{Want to calculate: } \Pr[T|A] = \frac{\Pr[T \cap A]}{\Pr[A]}$$

Unknown

✓

$$= \frac{\Pr[T \cap A]}{\Pr[T] * \Pr[A|I] + \Pr[\neg T] + \Pr[A|\neg T]}$$

$$\text{Have: } \Pr[T] = 0.001$$
$$\Pr[A|T] = .99, \Pr[A|\neg T] = .01$$

$$\text{Want to calculate: } \Pr[T|A] = \frac{\Pr[T \cap A]}{\Pr[A]}$$

$$= \frac{\Pr[T \cap A]}{\Pr[T] * \Pr[A|I] + \Pr[\neg T] + \Pr[A|\neg T]}$$

$$= \frac{\Pr[A|T] * P[T]}{\Pr[T] * \Pr[A|I] + \Pr[\neg T] + \Pr[A|\neg T]}$$

$$\text{Have: } \Pr[T] = 0.00001$$
$$\Pr[A|T] = .99, \Pr[A|\neg T] = .01$$

$$\text{Want to calculate: } \Pr[T|A] = \frac{\Pr[A|T] * P[T]}{\Pr[T] * \Pr[A|T] + \Pr[\neg T] * \Pr[A|\neg T]}$$
$$\frac{.99 * .001}{.001 * .99 + .999 * .01}$$
$$= 0.\overline{09} \approx 9\%$$

Summary of IDS and Firewall Goals

Effectiveness: How well does it detect attacks while avoiding false positives?

Efficiency: How many resources does it take, and how quickly does it decide?

Ease of use: How much training is necessary? Can a non-security expert use it?

Security: Can the system itself be attacked?

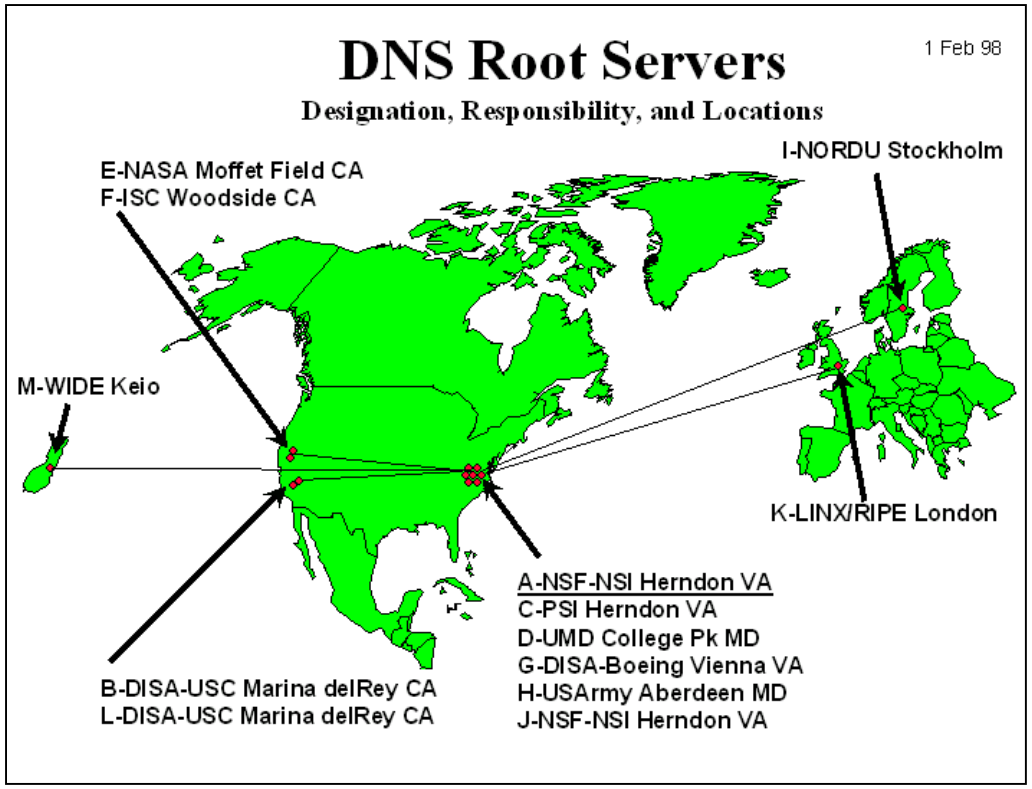
Transparency: How intrusive is it to use?

Expressiveness: What kinds of policies can we write?

Domain Name Service (DNS) Abuse

DNS Root Name Servers

- DNS maps symbolic names to numeric IP addresses
- 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 Sunday night.

The diversion was the result of the group's attack on computers that hold web address information.

Real URL names were deliberately mistranslated into the IP address of the hackers' site.

No data from the seven victims was lost or compromised as a result of the attack.

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

This page greeted many visitors to the sites of Vodafone, The Telegraph and others

Related Stories

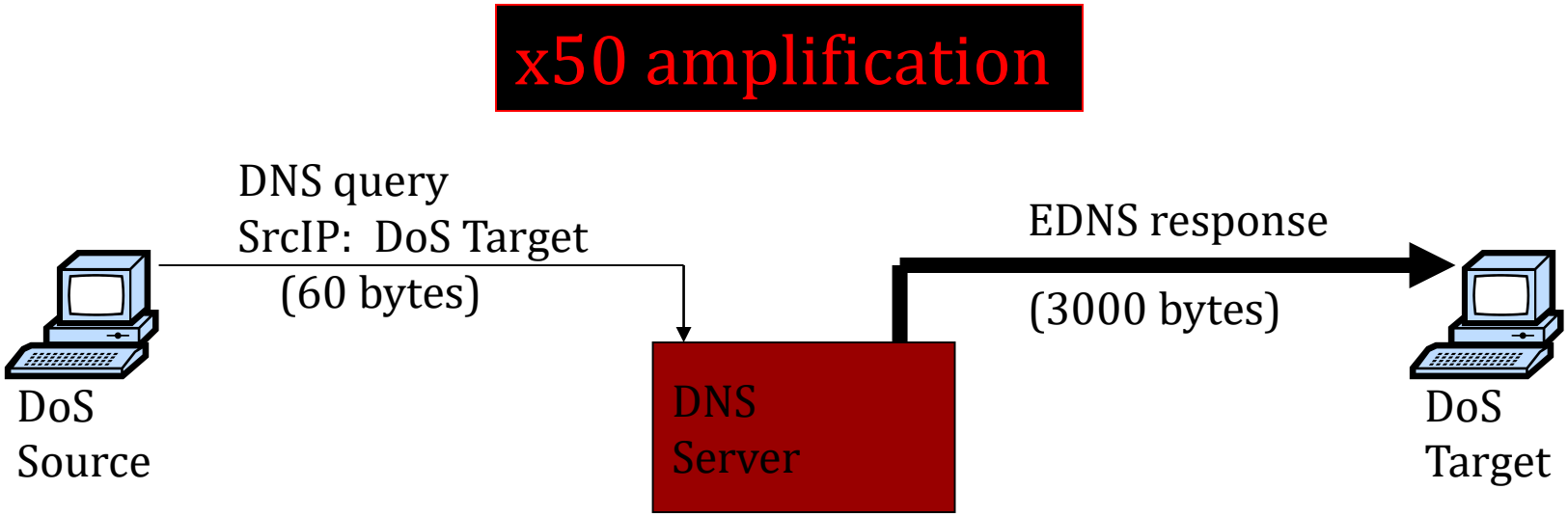
Nokia's developer

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

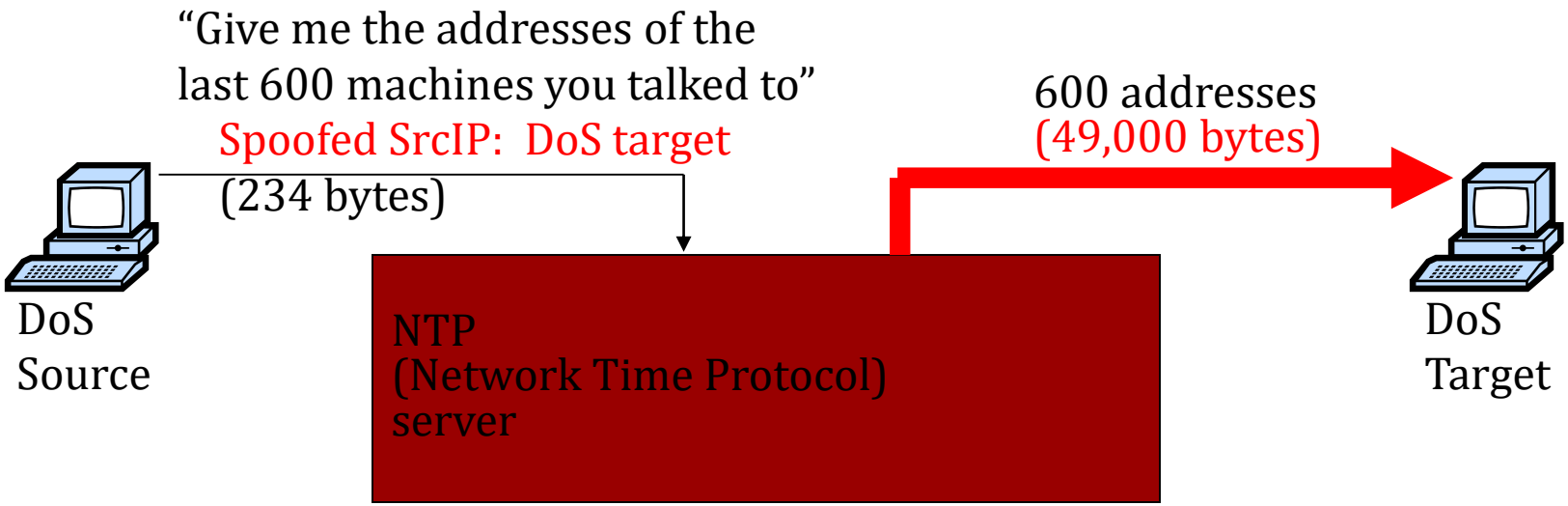


2006: 0.58M open resolvers on Internet
2013: 21.7M open resolvers (openresolverproject.org)

March 2013: 300 Gbps DDoS attack on Spamhaus

(Not Just DNS)

x206 amplification



December 2013 – February 2014:
400 Gbps DDoS attacks involving 4,529 NTP servers

7 million unsecured NTP servers on the Internet (Arbor)

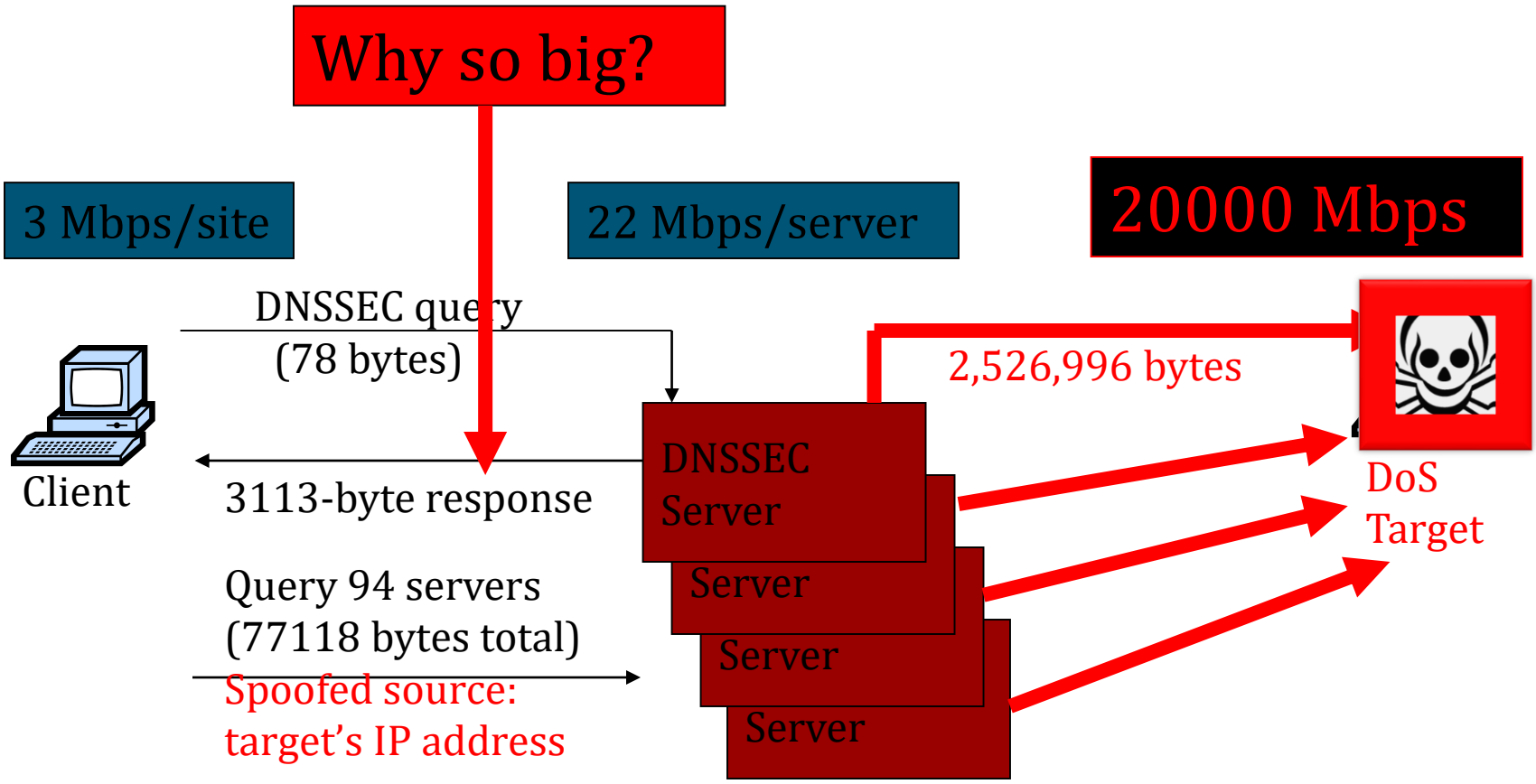
Other DNS Vulnerabilities

- DNS implementations have vulnerabilities
  - Multiple buffer overflows in BIND over the years
  - 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

DNSSEC

- Goals: authentication and integrity of DNS requests and responses
- PK-DNSSEC (public key)
  - DNS server signs its data – 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)

Querying DNSSEC Servers



5 times per second, from 200 sites

Using DNSSEC for DDoS

- RFC 4033 says:

"DNSSEC provides no protection against denial of service attacks"

- RFC 4033 doesn't say:

"DNSSEC is a remote-controlled double-barreled shotgun, the worst DDoS amplifier on the Internet"

Hacking-101

Steps Performed by a Hacker

- **Reconnaissance techniques**
  - Low tech methods (e.g., Social Engineering, Physical Break-In Dumpster Diving)
  - General web searches
  - Whois databases
  - DNS Zone Transfer

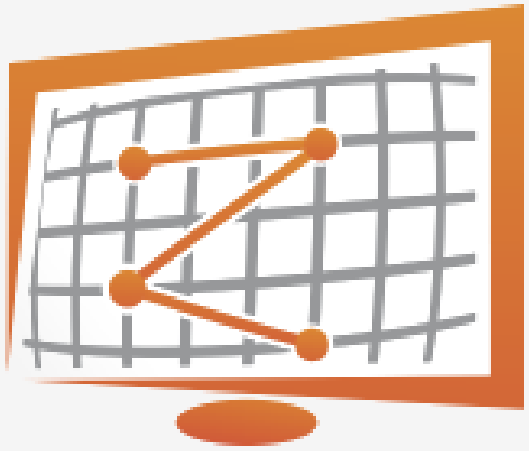
- **Scanning**
  - Network Mapping
  - Port Scanning
  - OS detection
  - Vulnerability assessment
- **Tools**
  - Nmap
  - Zmap
  - Masscan

Carefully & wisely experiment with Kali Linux

Do not just use tools. Develop New ones

Example: Zmap (from University of Michigan)

ZMap is an Internet-wide port scanner capable of scanning at **97% the maximum theoretical speed** of gigabit Ethernet



ZMap completes a single-port TCP SYN scan of **all of IPv4 in forty-five minutes**

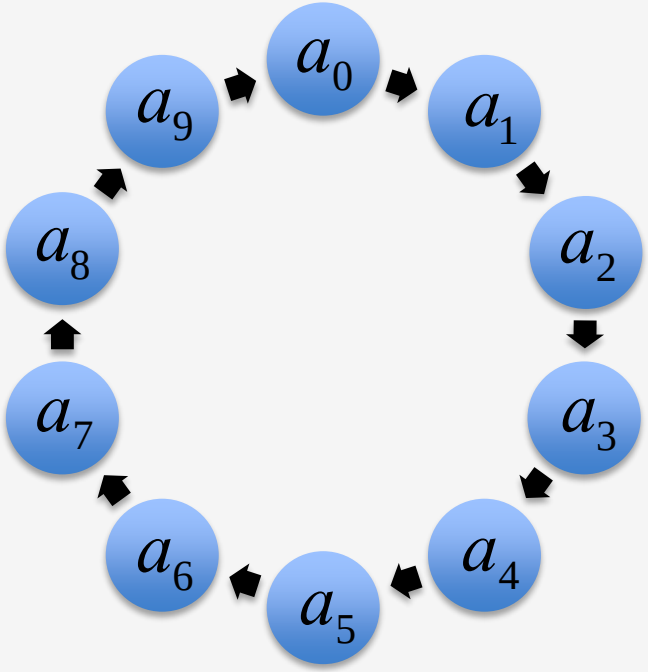
Zippier ZMap

A series of performance enhancements to ZMap, enabling scanning at **95% 10 GigE linespeed**, completing a single-port TCP scan in **under five minutes**

Address Generation

How do we address outgoing packets?

Multithreaded iteration over a cyclic group of integers modulo *p* requires a lock

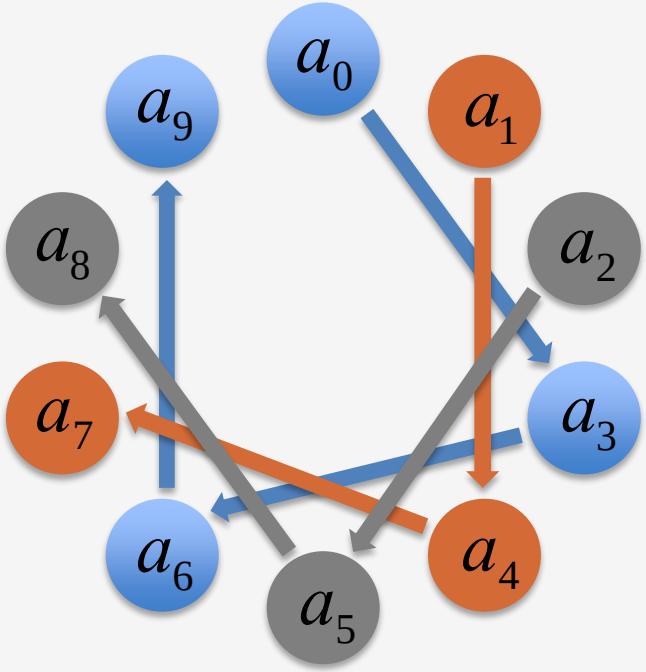


$$a_{i+1} = g \times a_i \bmod p$$

Address Generation

How do we address outgoing packets?

Multithreaded iteration over a cyclic group of integers modulo *p*



$$a_{i+1} = g \times a_i \bmod p$$

$$a_{i+n} = g^n \times a_i \bmod p$$

Shard the cycle into disjoint sets

Address Constraints

Good Internet citizenship demands honoring blacklist requests

1100 entries from 208 organizations on blacklist, 0.15% of IPv4 address space

Use blacklist to exclude IANA-reserved addresses, 14% of IPv4 address space

Optimized Address Constraints

And Remember

- **Ethics** are moral philosophy where a person makes a specific moral choice and sticks to it
- **Law** is a system that comprises of rules and principles to govern a society.
- Though, ethics are based on the goodwill of law, ethics completely differ in their foundation, basis and purpose.