# CS 524 A

Introduction to Cloud Computing

## Lecture 8: More on Network Appliances: Domain Name System (DNS) and Firewalls

# Outline

- DNS
  - History
  - DNS components
  - Domain name servers
  - Record structure and examples
  - Discovery
  - Security problems
  - The Role of ICANN
- Firewalls
  - The concept
  - Firewalls in Virtual Private Networks (VPN)
  - *Ingress* and *Egress* firewalls
  - Application Gateways
  - Stateful firewalls
  - Network Zoning

# Purpose

- We need to access *resources* (hosts, mailboxes, web pages, files)
  - Can we do that by specifying an *IP* address?
  - What happens if you send e-mail to ifaynber@155.246.14.121?
  - (By the way, I used *ping* to find "the address" of *mail.stevens.edu* )
- A few reasons to use *something* translatable to an IP address as a resource name instead of the IP address itself:
  - to use an application-layer naming scheme
  - to support resource mobility
  - to enable load balancing
  - to use a name that is easy (for a human) to remember
- **But it is not about translations only—it is also used for *service discovery***
- **The infrastructure to respond to such needs is central to Cloud Computing**

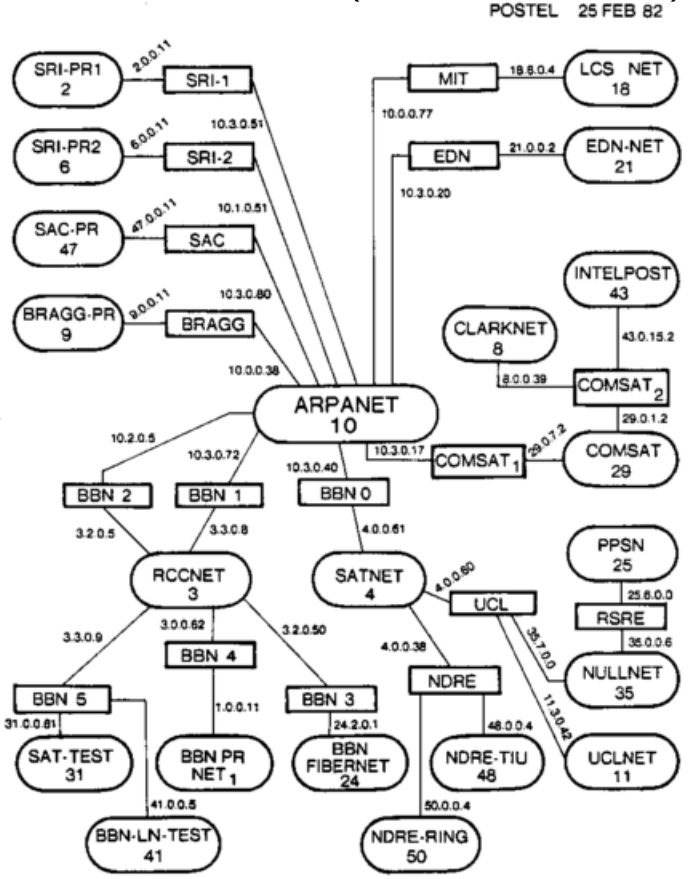
# SOME HISTORY

•Initially, the hosts names along with their addresses were kept in a file called *hosts.txt*

•But then the Internet looked like that…

**Jon Postel (1943-1998) was the first .us top name domain administrator**

Jon Postel's map of the Internet in 1982
(Source: Wikimedia)



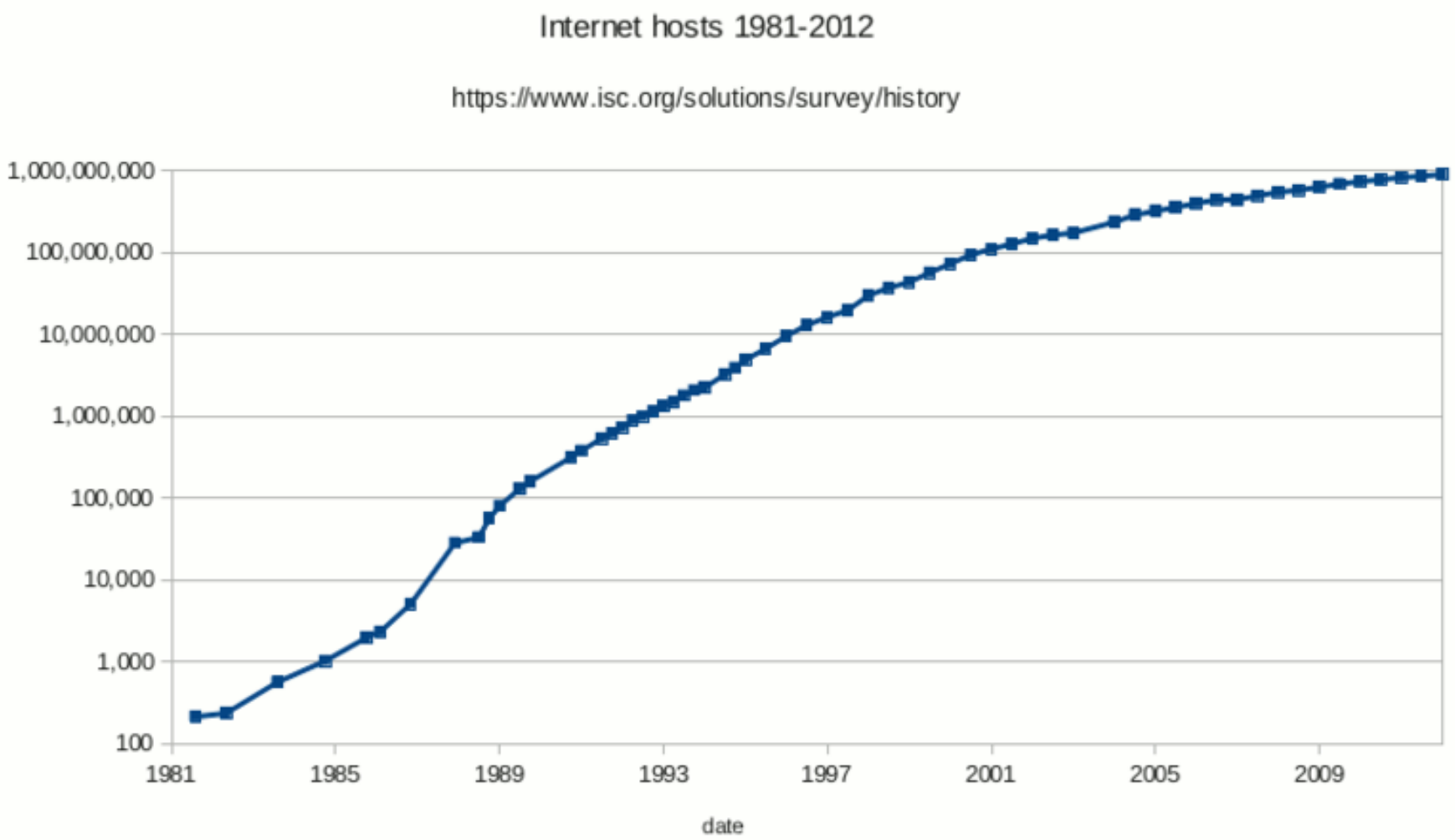
# The scale of the problem (https://www.isc.org/services/survey/)

- As of January 2019, there have been *1,03 billion*

hosts on the Internet that advertise their services!

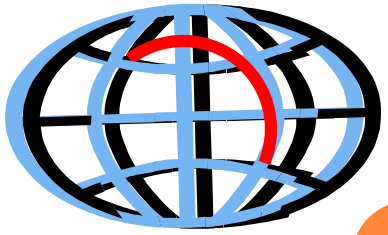
  - In January 2011, *818.374.269*
  - In January 2010, *732,740,444*
  - In January 2002, *147,344,723*
  - In January 1993, *1,313,000*
- And then a host may have multiple IP addresses (multi-homing, which we discussed earlier)
- Good questions:
  - How can the names be mapped to addresses?
  - How can *any* such scheme be managed?

# One picture is better than…



# Requirements for a possible scheme

- There must be no name conflicts
- The translations must happen fast, independent of geography
- The system must transcend international politics
- The system must always be on



# The scheme (Invented by Paul Mackapetris in 1980s)

- Is a hierarchical, multi-part, naming scheme
- Is implemented as a distributed database
- Maps host names, e-mail addresses, E.164 telephone numbers (fairly recently), and possibly other strings into **IP addresses** of the hosts that provide respective services
- Employs UDP (port 53) as transport protocol for queries smaller than 512 bytes (Why? The maximum UDP message size is 65507!) and TCP for larger queries
- Is standardized by the IETF (starting from RFC 1034 and RFC 1035)

RFC 1034 has been updated by RFC 1101, RFC 1183, RFC 1348, RFC 1876, RFC 1982, RFC 2065, RFC 2181, RFC 2308, RFC 2535,RFC 4033, RFC 4034, RFC 4035, RFC 4343, RFC 4035, RFC 4592, and RFC 5936.

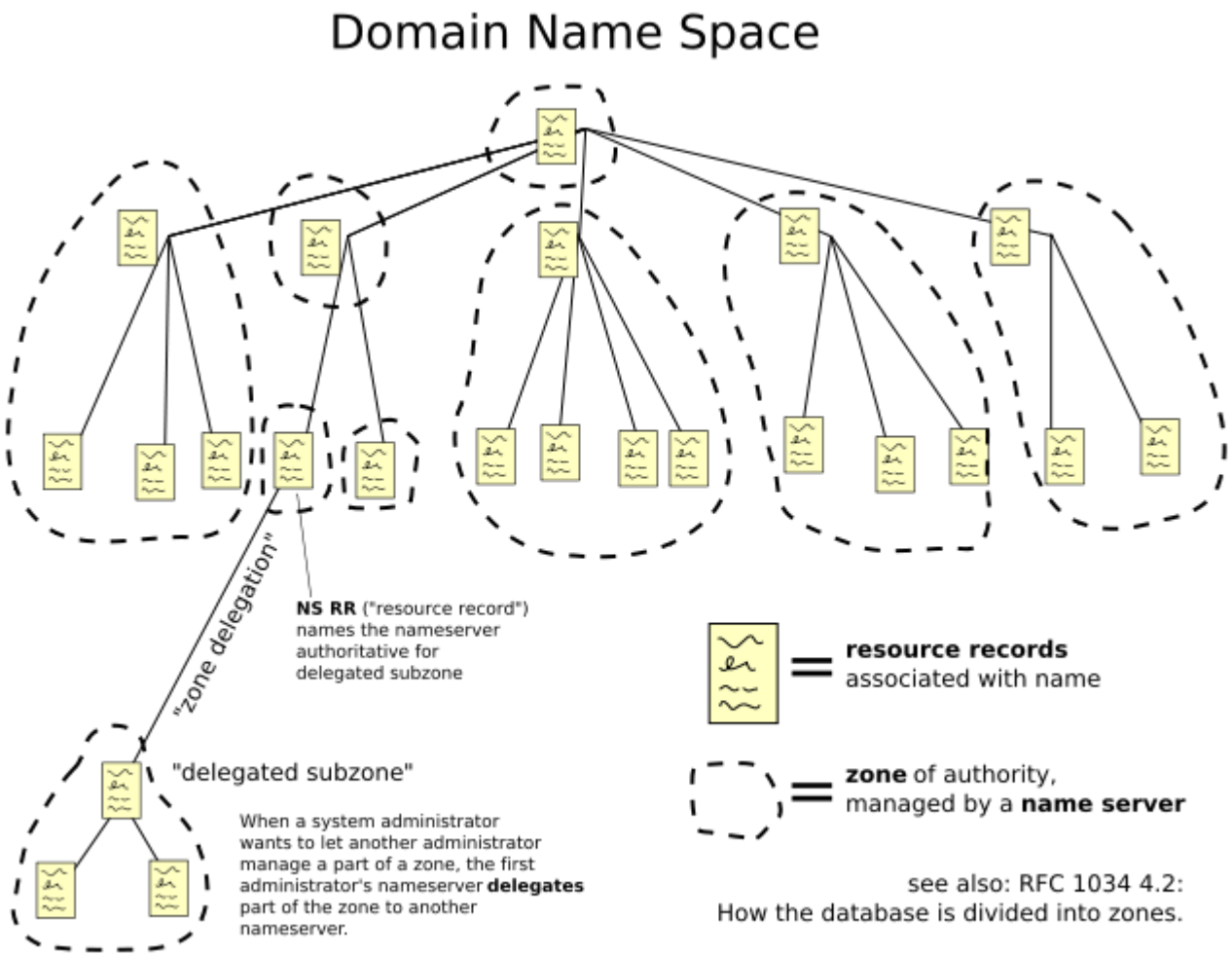
# DNS Components (after RFC 1034)

1. The Domain Name Space and Resource Records are specifications for the domain tree information
2. ***Name Servers*** hold information about its domain tree's structure. A name server is an ***authority*** for its part of the name space. Authoritative information is organized into units called **zones**
3. ***Resolvers*** extract information from the cache (if available) or name servers in response to a client request. Resolvers are expected to reside on the client host, and are typically accessed via system calls

# THREE VIEWS

1. The user's point of view:
   - the domain system is accessed through a call to a local resolver
   - the domain space is seen as a single tree
   - the user can request information from any section of the tree.
2. The resolver's point of view:
   - the domain system is composed of a set of name servers.
   - each server has one or more pieces of the whole domain tree's data
   - the resolver views each of these databases as *static*
3. A name server's point of view:
   - the domain system consists of separate ***zones***.
   - the name server must concurrently process queries that arrive from resolvers and it may delegate them

# The Domain Name Space




Source: Wikimedia

# SYNTAX

- A domain name is a sequence of *labels* (www.cs.stevens.edu), separated by *dots (.),* read from right to left
- The right-most label corresponds to the *top-level domain* (*edu* in www.stevens.edu*, right?*)
- *Interesting!:* The *root domain's* name is an empty string, and so every proper *fully-qualified* domain name must end with "." (pronounced *dot*) In other words, a complete name of the *Stevens* www server is *www.stevens.edu***.** ←
- Of course, *relative* names are processed if the context is known
- A label on the left specifies a *subdomain* of the domain to the right
- Labels are case-insensitive (try *www.Cs.StEVenS.eDU*)
- A *hostname* is a domain name that has at least one IP address associated with it. Not every domain name is a hostname (e.g., *com* is not*)*

# Some domain names

- .**com** (as in *commercial*)
  Originally intended only for for-profit businesses, but now anyone can register there, and there is much *squatting*

- **.biz** (for *business*)—created to relieve *.com*
  Anyone can register, but registrations may be challenged later

- **.org** (as in *organization*)
  Originally intended for non-profit organizations, but open to all

- **.net** (as in *network*)
  Originally intended only for telecommunications services providers, but now open to all
- **.us** (as in ***US***)
  Mostly used by local governments in the United States, but technically is unrestricted
- **.xxx** (as in ***xxx-rated***)
  This involves a huge controversy and is a case study for important business and legal issues.
  Look, for example, at *www.kite.xxx*. And then, look at *PopeBenedict.xxx* and read the
  March 19, 2012 article in *The Register*®:
  www.theregister.co.uk/2012/03/19/pope_benedict_cybersquatter

**Looks like everything is open to everyone…**

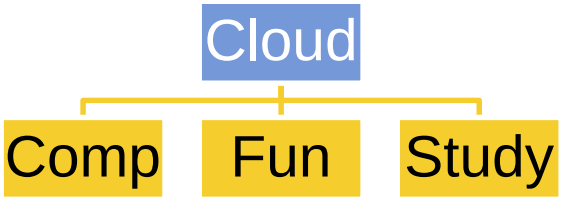
# …But some domains are strict

**USA Top-Level Domains**

- **.edu** (as in *education*)
  Limited to accredited colleges and universities
- **.gov** (as in *government*)
  Limited to US governmental agencies
- **.mil** (as in *military*)
  Limited to the US military
- **.int** (as in *international*)
  Limited to organizations and programs endorsed by treaties (e.g., *itu.int*)
- **.museum**
  Limited to legitimate museums

# Subdomain Creation Permission

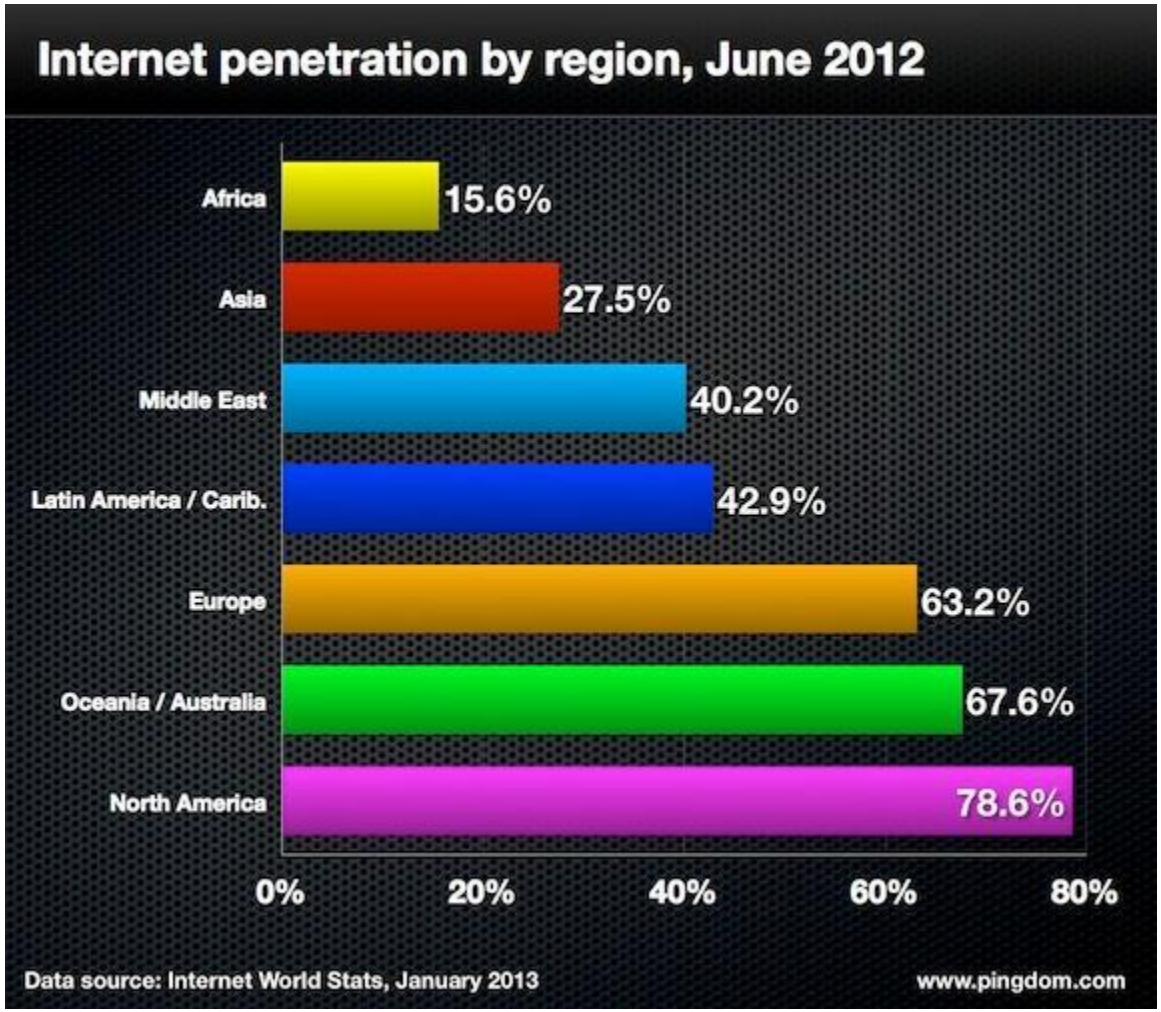
- is up to the domain administrator: To create a subdomain *cloud* at cs.stevens.edu, one needs to get a permission of the *cs* subdomain administrator
- Once a subdomain is created, its administrator can create other subdomains at will:



# And then there are country code (ISO 3166-1) top level domains

- Some require citizenship for registration (e.g., *.al* for *Albania*)
- Some require only local administrative contact (e.g., *.ee for Estonia* or *.de* for Germany)
- Some are reported to be inactive (e.g., *Eritrea* offers no services to register under the domain *.er*)
- Some restrict registration to specific categories only (e.g., *.com.br*, *gov.br* in Brazil)
- Some are sold to companies (e.g., *.tk [Tokelau]* sold to *DOT.TK*) (see http://m.zdnet.com.au/pacific-atoll-a-phishing-haven-339313909.htm)

# The Internet penetration differs by region



# THECOUNTRY CODE DOMAIN NAMES HAVE BEEN INTERNATIONALIZED!

| DNS Name | Domain Name | Country |
|---|---|---|
| xn--lgbbat1ad8j | .الجزائر | Algeria |
| xn--fiqs8s | .中国 | China |
| xn--fiqz9s | .中國 | China |
| xn--h2brj9c | .भारत | India |
| xn--h2brj9c | .भारत | India |
| xn--mgbbh1a71e | .بهارت | India |
| xn--fpcrj9c3d | .భారత్ | India |
| xn--gecrj9c | .ભારત | India |
| xn--s9brj9c | .ਭਾਰਤ | India |
| xn--xkc2dl3a5ee0h | .இந்தியா | India |
| xn--45brj9c | .ভারত | India |

# ASCII-COMPATIBLE ENCODING (ACE) LABELS (RFC 3490)

- indicate the internationalized names
- are used in prefixes to indicate conversions between ASCII and non-ASCII forms of a domain name, accomplished by algorithms called *ToASCII* and *ToUnicode*.
  - *ToASCII* translates the name to ASCII using *punicode encode* algorithm and then prepends the four-character ACE prefix ("xn--“).
  - *ToUnicode* strips off the ACE prefix and applies the *punycode decode* algorithm

# Some Security Problems right here?

- Many people use www.*paypal.com*
- What if someone is directed to *www.paypa1.com*?

  (Well, this has been fixed)
- How about *www.paypal.com*? (Try copying this in your browser.)

Fortunately, this  results in a bad domain name:

```
Network Error (dns_unresolved_hostname)

Your requested host "www.xn--pypl-53dc.com" could not be resolved
by DNS.

For assistance, contact your network support team.
```

# Checking the *Whois*

For example, http://dnsquery.org/whois/ returns the same result for paypal.com and paypa1.com:

Registrant: Domain Administrator eBay Inc. 2145 Hamilton Avenue San Jose CA 95125 US hostmaster@ebay.com +1.4083767400 Fax: +1.4083767514

Domain Name: paypa1.com

Registrar Name: Markmonitor.com Registrar

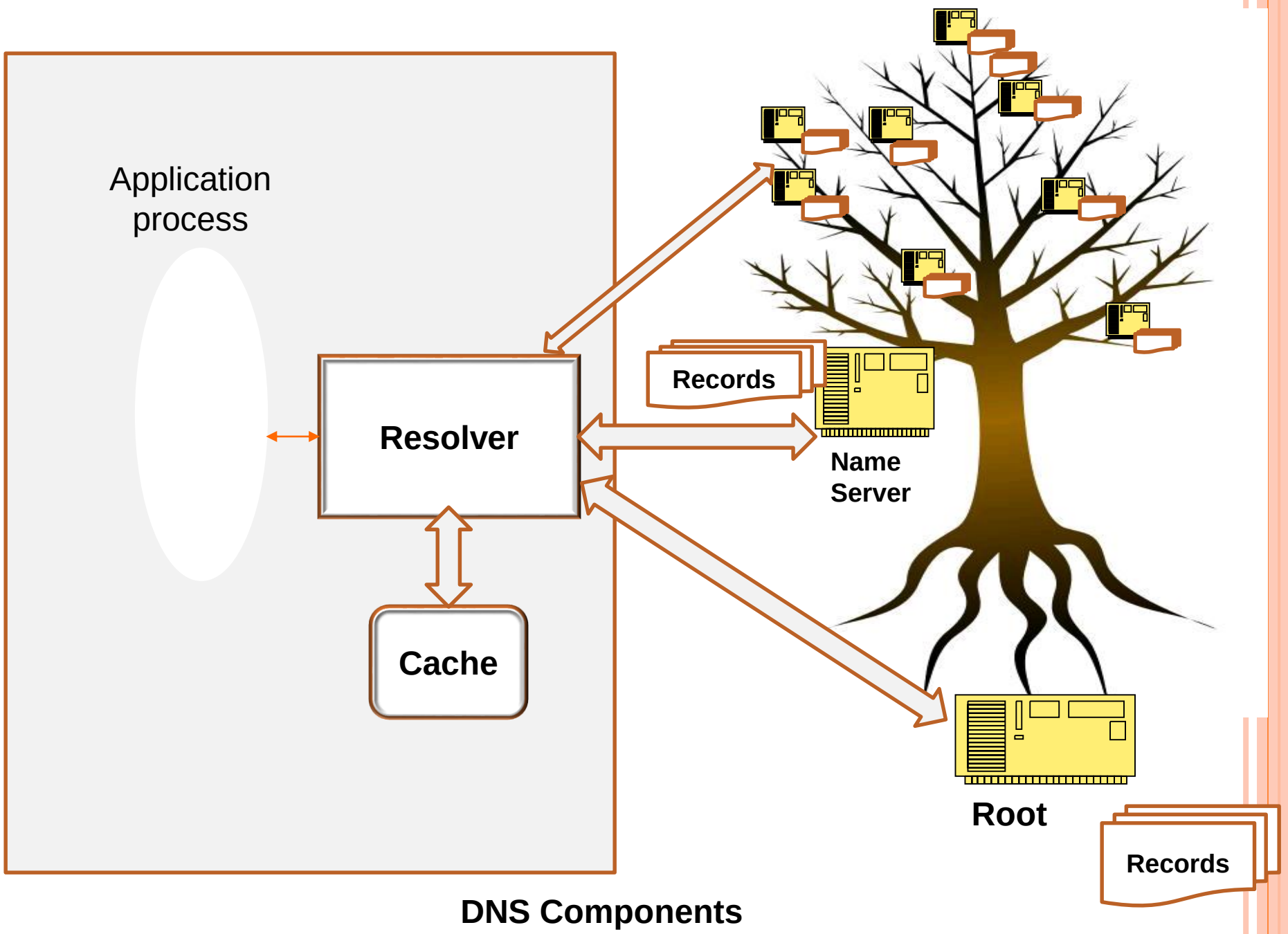
Whois: whois.markmonitor.com

Registrar Homepage: http://www.markmonitor.com

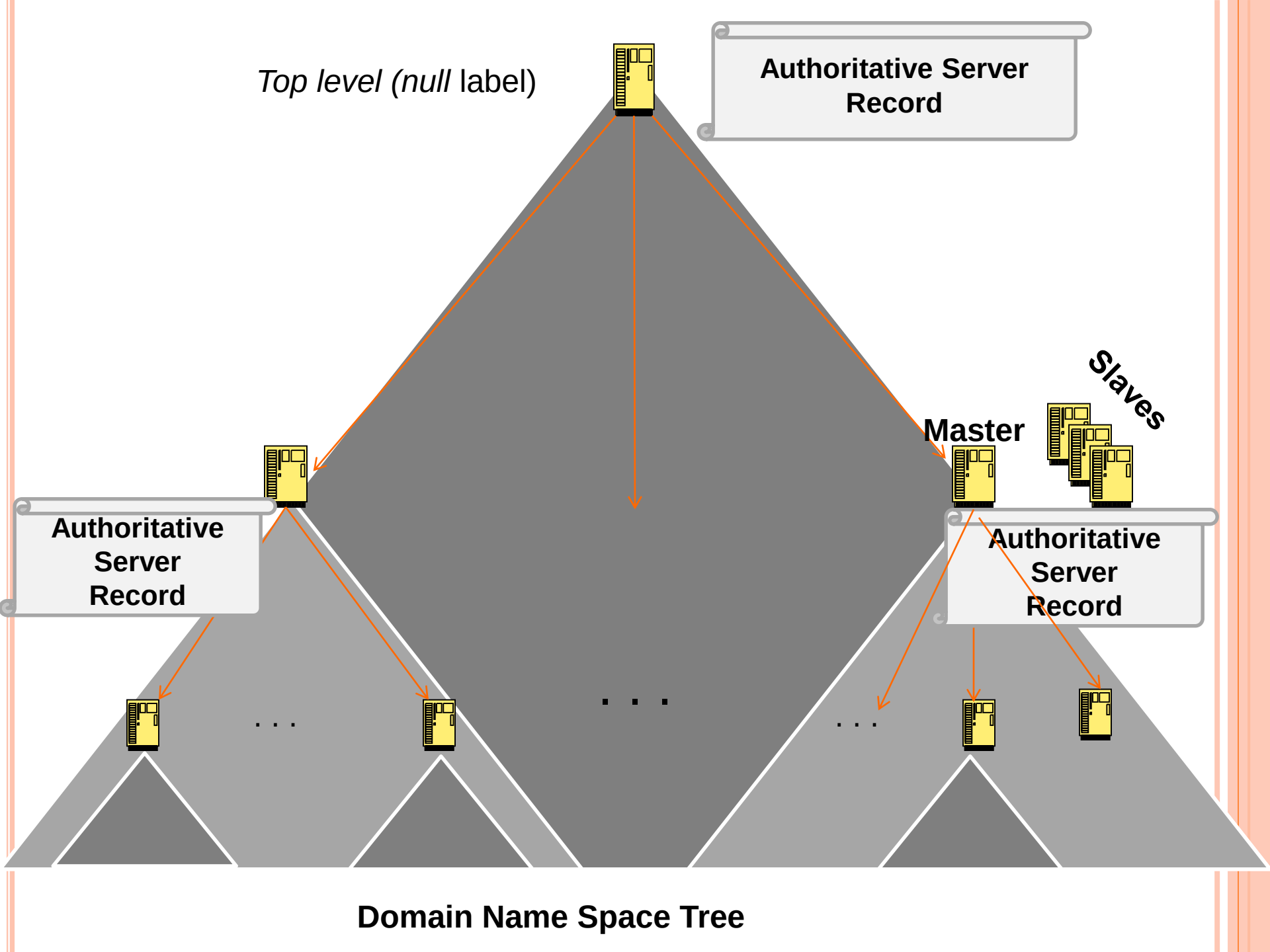
Administrative Contact: Domain Administrator eBay Inc. 2145 Hamilton Avenue San Jose CA 95125 US hostmaster@ebay.com +1.4083767400 Fax: +1.4083767514 Technical Contact, Zone Contact: Domain Administrator eBay Inc. 2145 Hamilton Avenue San Jose CA 95125 US hostmaster@ebay.com +1.4083767400 Fax: +1.4083767514 Created on.....: 2000-10-20. Expires on.....: 2012-10-20. Record last updated on..: 2012-02-25. Domain servers in listed order: ns7.markmonitor.com ns3.markmonitor.com ns6.markmonitor.com ns2.markmonitor.com ns4.markmonitor.com ns5.markmonitor.com ns1.markmonitor.com

# DNS Organization: Top Level Domain (TLD) names

- Top Level Domain (TLD) names correspond to the DNS Root Zone. They are
  - Under the authority of the *The National Telecommunications and Information Administration (NTIA*),agency of the United States Department of Commerce
  - Managed by the Internet Assigned Numbers Authority (IANA)
  - Operated by the *Internet Corporation for Assigned Names and Numbers (ICANN*)
  - Administered by *VeriSign, Inc* (as a contractor)




DNS Components




Domain Name Space Tree

| Domain | Sponsoring organization |
|---|---|
| تونس. | Agence Tunisienne d'Internet |
| .中国 | China Internet Network Information Center |
| .中國 | China Internet Network Information Center |
| рф | Coordination Center for TLD RU (Russia) |
| .ಭಾರತ್ | National Internet Exchange of India |
| .ভারত | National Internet Exchange of India |
| .срб | Serbian National Register of Internet Domain Names (RNIDS) |
| .சிங்கப்பூர் | Singapore Network Information Centre (SGNIC) Pte Ltd |

**Figure 5.10: Examples of internationalized country code top domain names**

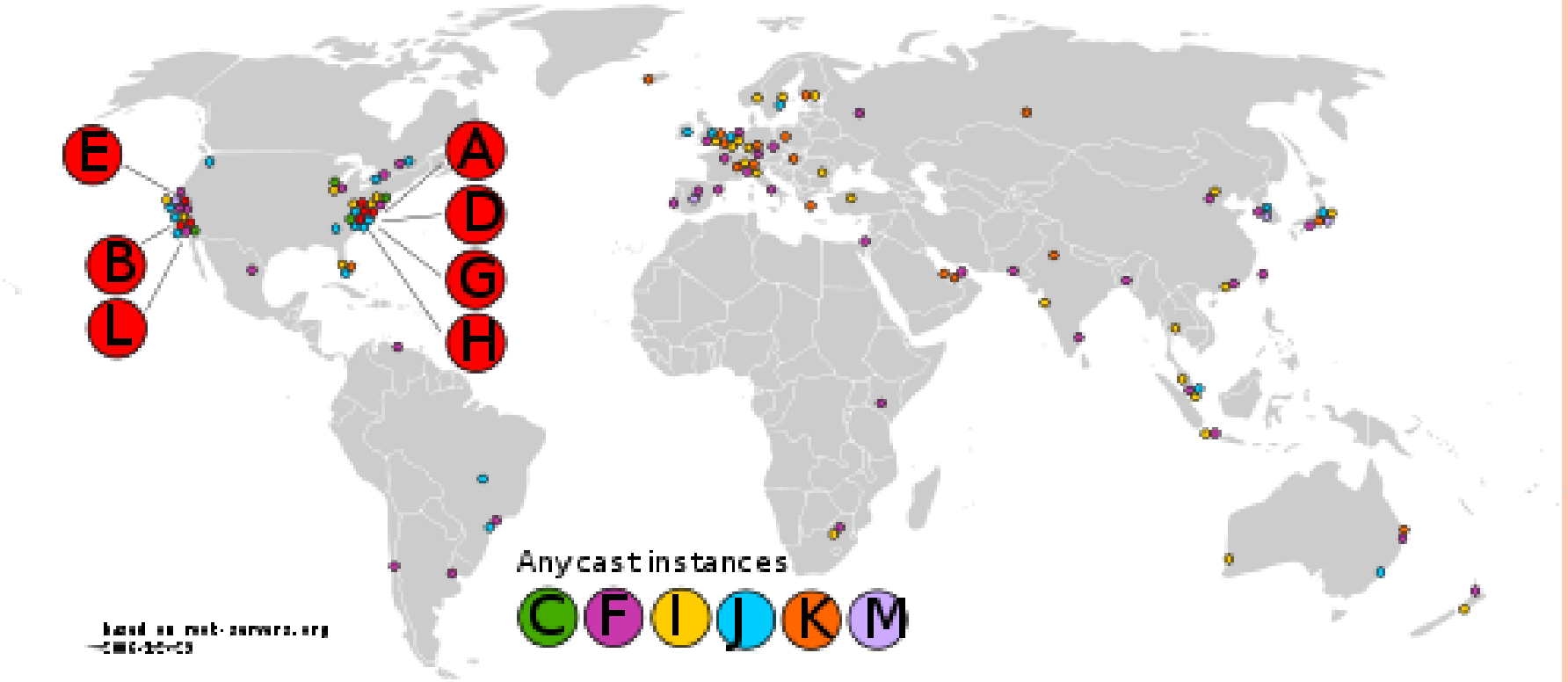
Source: Internet Assigned Number Authority, www.iana.org

# ROOT DOMAIN

- Has an empty string as its label
- Contains all generic and country-specific top-level domains as well as *.arpa* [reverse DNS look-ups and E.164] and *.test* domains. The *current* full list of the domain names is at http://data.iana.org/TLD/tlds-alpha-by-domain.txt
- The root name servers have a form: *<letter>.root-servers.net.* Currently, 13 letters, A through M are used
- Some servers are replicated in multiple locations and announce their addresses in *anycast* form

  (Reminder: Anycast addressing routes datagrams to a single member of a group of potential receivers that are all identified by the same destination address. This is a *one-to-one-of-many* association.)
- …and then there have been alternative namespaces, with another huge controversy involved (see http://news.cnet.com/2100-1023-204904.html)
- www.internic.net/zones/root.zone is the official root zone file with the names and IP addresses of all top-level domain servers

# Root name servers




Complete information at
http://www.root-servers.org/

# DNS Resolver

- Is a client side (in a more general sense) of the DNS
- Maintains a cache of recent queries (why?)
- Issues either a
  - *Non-recursive* query (the DNS server providers a record for which it is an authoritative server)
  - *Recursive* query (the DNS server may need to answer by referring to other servers)

# To expand a bit…

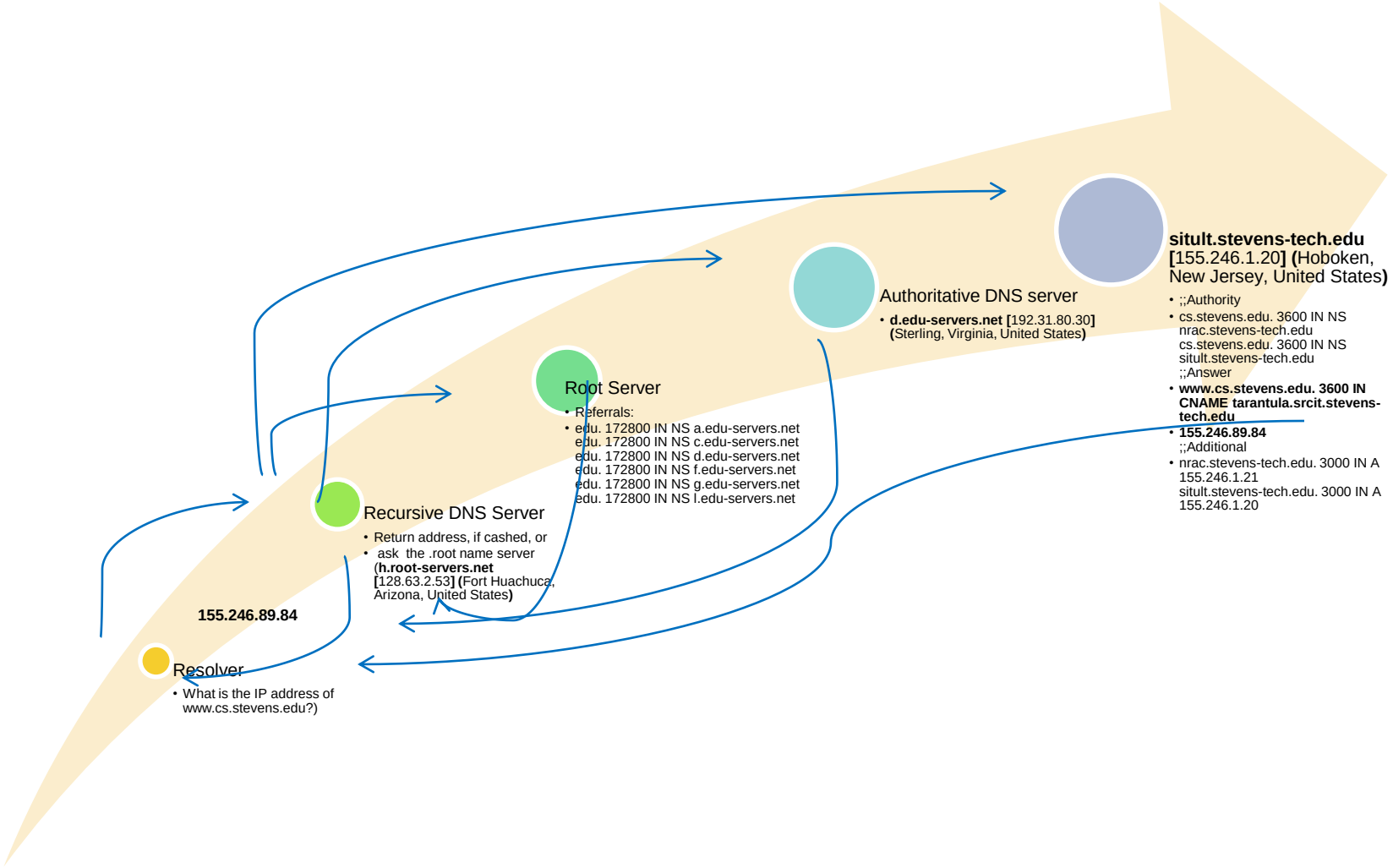
- With a *recursive name query*, the DNS client requires that the DNS server respond with either
  - the requested resource record ; or
  - an error message stating that the record or domain name does not exist. (No referral here!)
- Recursive name queries are supposed to be made by a DNS client to a DNS server .
- An *iterative name query* is one in which a DNS client expects the DNS server to return the best answer it can. In the absence of an exact match, the server may return a *referral* (that is, a pointer to a DNS server authoritative for a lower level of the domain namespace).

We will return to this later

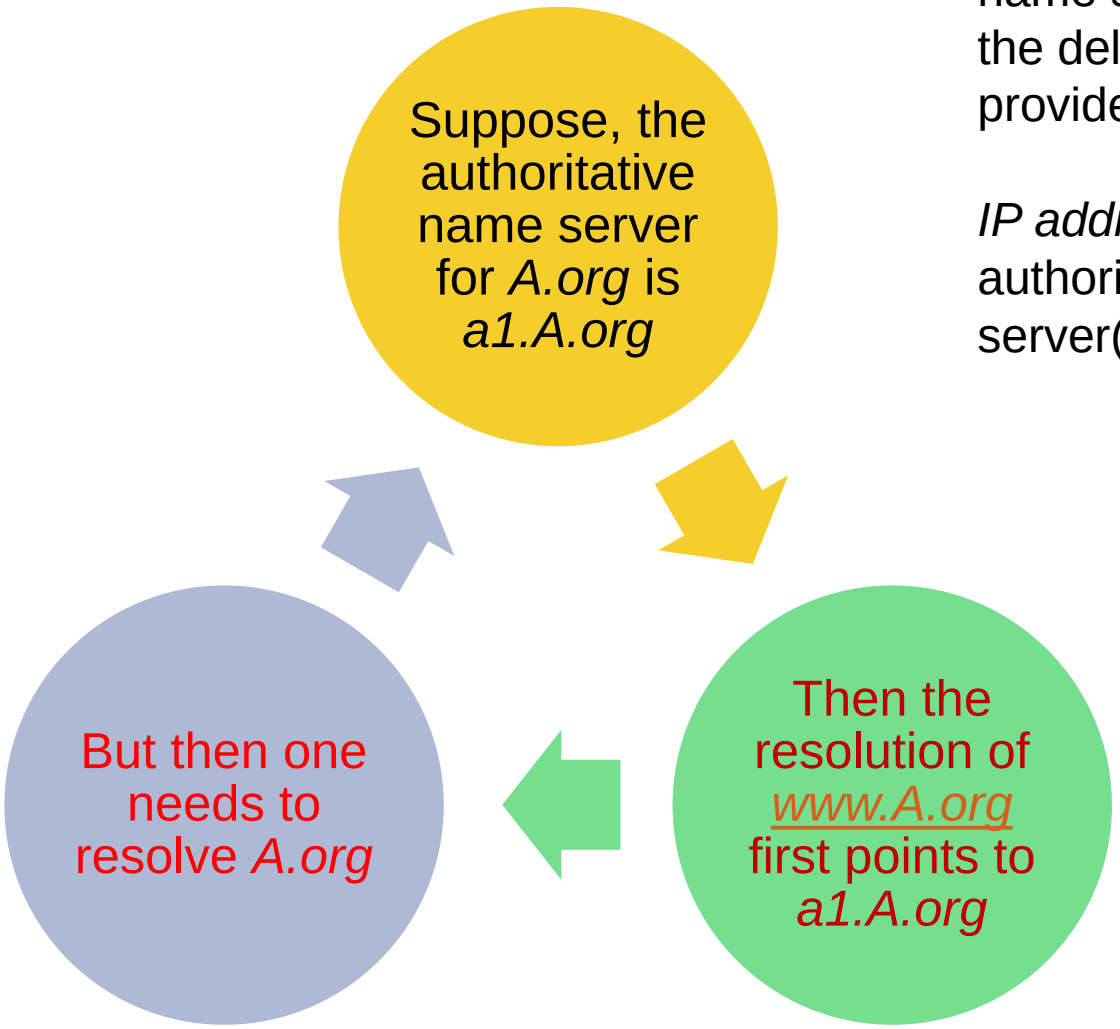
# An Authoritative Name Server

- Is a server that has been configured by the **domain administrator** for each zone
  - *Master Server* contains the original master copies of all zone records
  - *Slave Server* uses the *updating mechanism* of the DNS protocol to maintain an identical copy of the Master server
- Indicates the authority by setting the *Authoritative Answer (AA)* bit in its response

# DNS Process in a nutshell

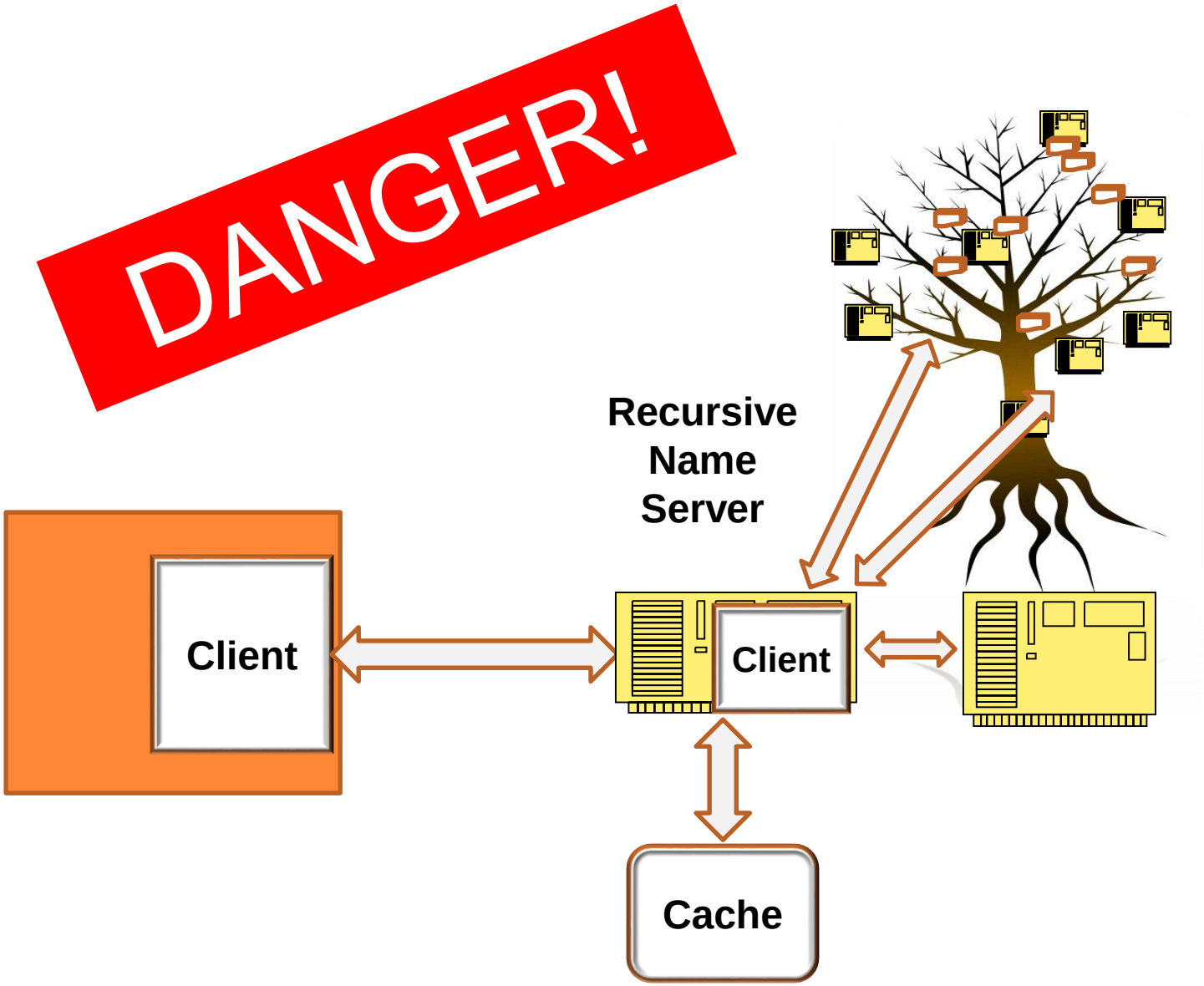


# Circular dependencies

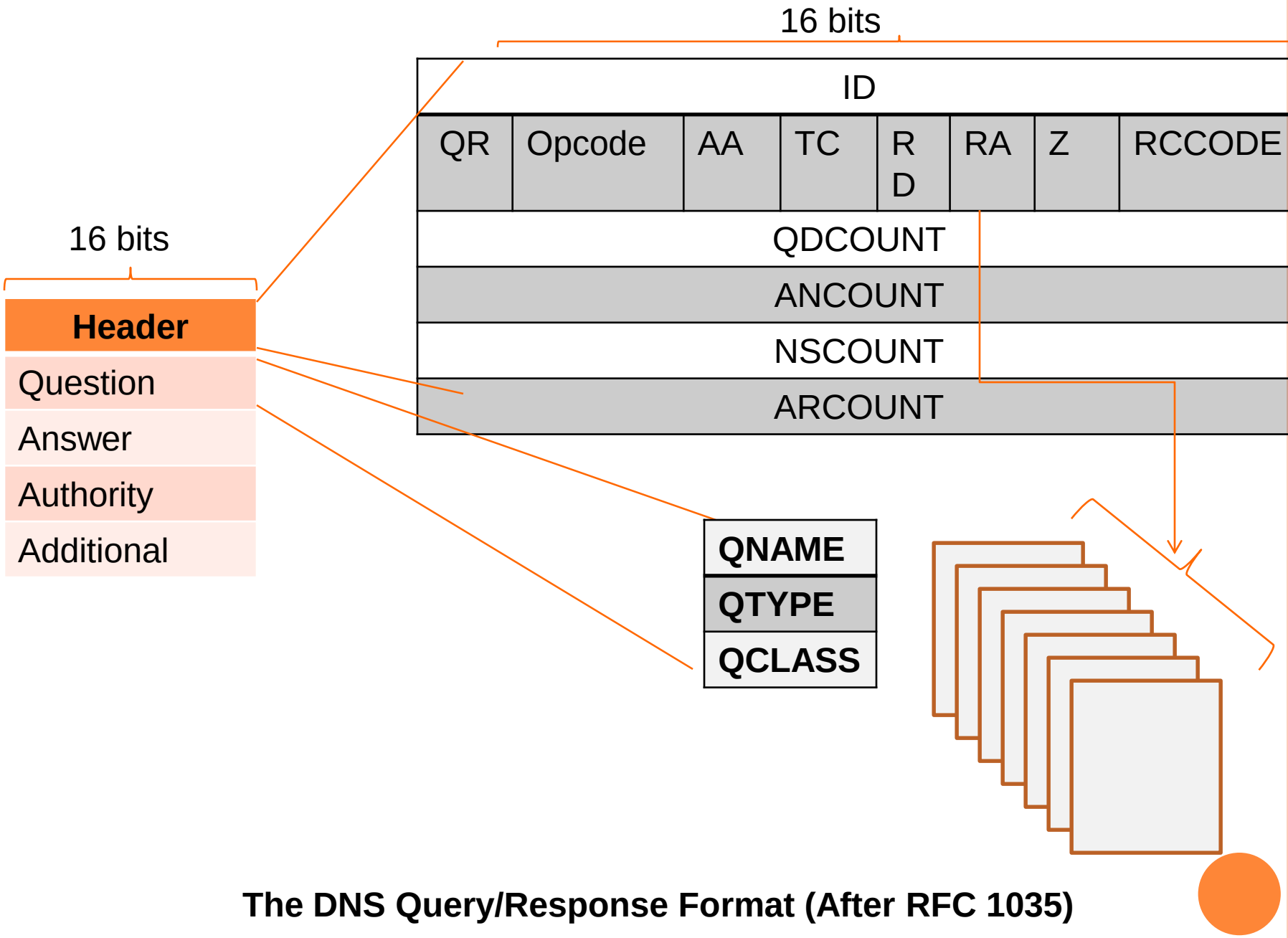



For this reason, the name server providing the delegation MUST provide the *glue:*

*IP address(es)* of the authoritative name server(s) delegated to




A recursive name server (resolver)



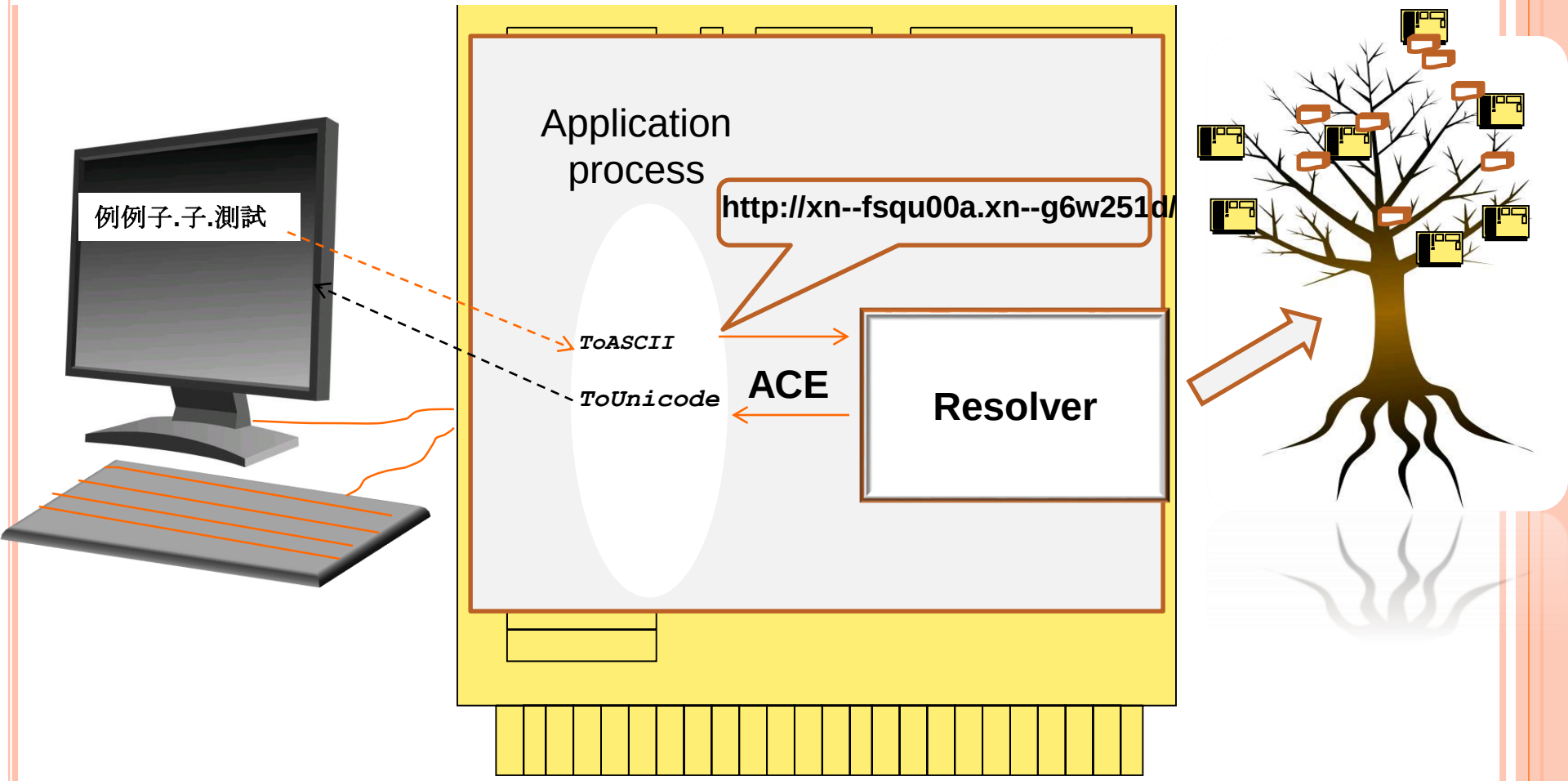

**The DNS Query/Response Format (After RFC 1035)**

# Root Name Systems

| System name | IP Addresses | Manager |
|---|---|---|
| a.root-servers.net | 198.41.0.4, 2001:503:ba3e::2:30 | VeriSign, Inc. |
| b.root-servers.net | 192.228.79.201 | University of Southern California (ISI) |
| c.root-servers.net | 192.33.4.12 | Cogent Communications |
| d.root-servers.net | 199.7.91.13, 2001:500:2d::d | University of Maryland |
| e.root-servers.net | 192.203.230.10 | NASA (Ames Research Center) |
| f.root-servers.net | 192.5.5.241, 2001:500:2f::f | Internet Systems Consortium, Inc. |
| g.root-servers.net | 192.112.36.4 | US Department of Defence (NIC) |
| h.root-servers.net | 128.63.2.53, 2001:500:1::803f:235 | US Army (Research Lab) |
| i.root-servers.net | 192.36.148.17, 2001:7fe::53 | Netnod |
| j.root-servers.net | 192.58.128.30, 2001:503:c27::2:30 | VeriSign, Inc. |
| k.root-servers.net | 193.0.14.129, 2001:7fd::1 | RIPE NCC |
| l.root-servers.net | 199.7.83.42, 2001:500:3::42 | ICANN |
| m.root-servers.net | 202.12.27.33, 2001:dc3::35 | WIDE Project |

Source: Internet Assigned Number Authority, www.iana.org

# The domain name internationalization components (after RFC 3490)



# The Resource Record (RR) structure

- The resource record is a *quintuple*
  1. *Domain Name:* the fully-qualified domain name of the node in the tree
  2. Class: IN (for **Internet hostnames**, servers, and IP addresses); we don't consider others here
  3. Type: the type record (we will address this in a moment)
  4. Record Data: (RDLENGTH, RDATA)
  5. *Time To Live (TTL)* (in seconds): measure of stability of the information in cache
- A simpler view is

  {Domain Name, [Type, Data], TTL,}

16 bits

| | |
|---|---|
| ***Domain Name*** | A fully-qualified domain name of the node in the tree (64 bits) |
| **Class** | IN (for Internet hostnames, servers, and IP addresses) |
| **Type** | Record type (A, AAAA, SOA, …, CERT, …SRV, …) |
| **Record Data** | (RDLENGTH, RDATA) |
| **Time To Live (TTL)** | (in seconds, 32-bit value): a measure of stability of the information in cache (32 bits) |

**The RR structure (After RFC 1035)**

# Some Important Data Types and Their Values (RFC 1035 and RFC 2782)

- Type: *Start of Authority (SOA)*
- Data:
  - MNAME: <domain-name> of the name server that was the original or primary source of data for this zone
  - RNAME: A <domain-name> which specifies the mailbox of the person responsible for this zone
  - SERIAL: The unsigned 32-bit version number of the original copy of the zone
  - REFRESH: A 32-bit time interval before the zone should be refreshed
  - RETRY: A 32-bit time interval that should elapse before a failed refresh should be retried
  - EXPIRE: A 32 bit time value that specifies the upper limit on the time interval that can elapse before the zone is no longer authoritative
  - MINIMUM: The unsigned 32 bit minimum TTL field that should be exported with any RR from this zone

# The result of a look-up

| Name | class | type |
|---|---|---|
| www.alcatel-lucent.com | IN | SOA |

**Answer records**

| Name | class | type | data | TTL |
|---|---|---|---|---|
| www.alcatel-lucent.com | IN | CNAME | portal.alcatel-lucent.com.edgekey.net | 299s |
| portal.alcatel-lucent.com.edgekey.netl | IN | CNAME | e782.b.akamaiedge.net2 | 1600s |

**Authority records**

| Name | class | type | data | TTL |
|---|---|---|---|---|
| liveakamaiedge.net | IN | SOA | internal.akamaiedge.netemail:hostmaster@akamai.com | 18090000s |

# Important RR data types and their values (cont.)

- *Canonical name (CNAME):* a string, to translate aliases into actual host names. While doing the homework you will learn about

- *A:* a 32-bit integer, an IPV4 address of a host
- *AAAA:* a 128-bit integer, an IPV6 address of a host
- *Mail Exchange (MX),* string, a domain willing to accept e-mail accompanied by the priority value (in descending order)
- *Name Server:* string, the name of a name server for this domain

# Important RR data types and their values (cont.)

- *Pointer (PTR):* "A <domain-name> which points to some location in the domain name space." (Reportedly used for reverse look-up)
- *Sender Policy Information (SPF):* a step toward limiting e-mail spam (don't start me on that—you can take a look at US Patent 77,752,440)
- *Service (SRV):* **Service discovery**
  - *service*: the symbolic name of the desired service
  - *proto*: the transport protocol of the desired service; this is usually either TCP or UDP
  - *name*: the domain name for which this record is valid.
  - *TTL*: standard DNS TTL.
  - *class*: standard DNS class field (this is always *IN*).
  - *priority*: the priority of the target host in descending order
  - *weight*: A relative weight for records with the same priority.
  - *port*: the TCP or UDP port on which the service is to be found.
  - *target*: the canonical hostname of the machine providing the service

# And then DNS is used for

- E.164 address for IP telephony
- Data repositories (in conjunction with LDAP)
- Security certificates
- Load balancing (!)
- Content Delivery—we will discuss this later
- …

# More on recursive resolvers: Security Riscs

.

**Generic DOS and DDOS attacks:** Purposefully sending a huge number of requests to overwhelm the server (and thus make it much slower to respond to legidimate queries)

**DOS amplification attacks**: Requests spoofing an IP address result in a flood of results to that address.

**DNS cache poisoning**: If the server does not correctly validate DNS responses (e.g., by using DNSSEC) the server may end up caching the incorrect entries locally and later pass them to other users that make the same request. This is typically used for site hijacking.

**Root name server performance degradation.** When DNS servers are not configured correctly, queries using RFC1918 addressing (also known as "private" addressing) may end up at to the root name servers, causing a degradation in service for legitimate queries to those servers.

# A DDOS ATTACK: DNS REFLECTION

- Made it to the New York Times on March 26, 2013
  http://www.nytimes.com/2013/03/27/technology/internet/online-dispute-becomes-internet-snarling-attack.html?pagewanted=all
  - On March 13, *SpamHaus* (www.spamhaus.org) put *CyberBunker* on its black list.
  - Allegedly, *CyberBunker* retaliated with the *DNS reflection attack*
  - With only 300 bytes of botnet traffic, 3.000 bytes of attack traffic have been generated (TCP responses and, ironically, DNSSEC make traffic even larger)
- Lessons:
  - Recursive Resolvers must reply only to the requests from their own networks
  - Public DNS servers must apply filtering for abusive queries and ensure the frequency of queries is commensurate with the expected volumes.
- For the detail, read this: http://www.spamhaus.org/news/article/695/answers-about-recent-ddos-attack-on-spamhaus

# What is being done

- There is a watch for open recursive resolvers: (http://dns.measurement-factory.com/surveys/openresolvers.html)
- There is a pressure to close them down
- There is a thread at the **North American Network Operators' Group** (**NANOG**) list that discusses the related attacks http://www.merit.edu/mail.archives/nanog/2006-02/msg00579.html
- There is ICANN advisory that recommends source address verification

http://dns.measurement-factory.com/surveys/openresolvers.html

# DNS Software

Most servers run *Berkeley Internet Domain Name (BIND)* DNS software

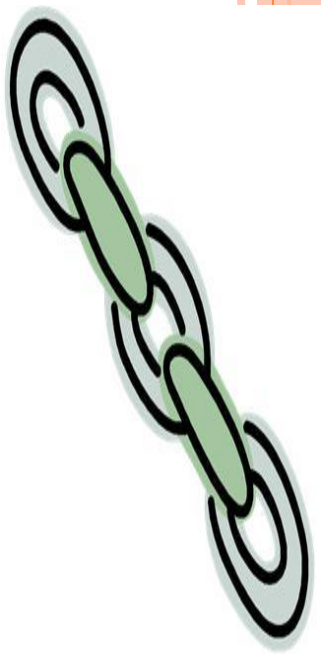
- The first issue of DNS software, *Berkeley Internet Domain Name (BIND)*, was released by University of California at Berkeley.
- The later versions of *BIND* were written in Digital Equipment Corporation
- Then the development of *BIND* has moved into the *Internet Systems Consortium* (http://www.isc.org/). It is an open source project.
- The standards had evolved along with implementations, but security was neither a requirement nor a concern initially—scalability was.
- Security was an add-on, unfortunately, and it was specified as a reaction to serious vulnerabilities (*cache poisoning*, *zone spoofing*—please read RFC 3833 ww.ietf.org/rfc/rfc3833.txt).

# DNS Security Extension (DNSSEC)

- The idea is to authenticate each DNS record so that the client can check its origin. In 1997, the IETF has developed a solution and published the first standard (RFC 2065, now obsolete).
- The work on the standard continued, with the present set of specifications (*DNSSEC-bis*) contained in RFCs 4033-4035.
- The scheme is to use public-key cryptography and verify the chain of trust top-down, starting with the authoritative name server for the root.
- DNSSEC actually provides not only a record's origin authentication by also its *integrity assurance* (against any modification en route to the client). To enable this, DNSSEC adds new resource record types:
  - *Resource Record Signature (RRSIG),*
  - *DNS Public Key (DNSKEY)*
  - *Delegation Signer (DS); and*
  - *Next Secure (NSEC).*
- DNSSEC also modifies the message header by adding new flags.
- Overall, these modifications end up in much larger DNS response messages, the result that can unfortunately be exploited in the denial-of-service attacks.

# TRUST CHAIN

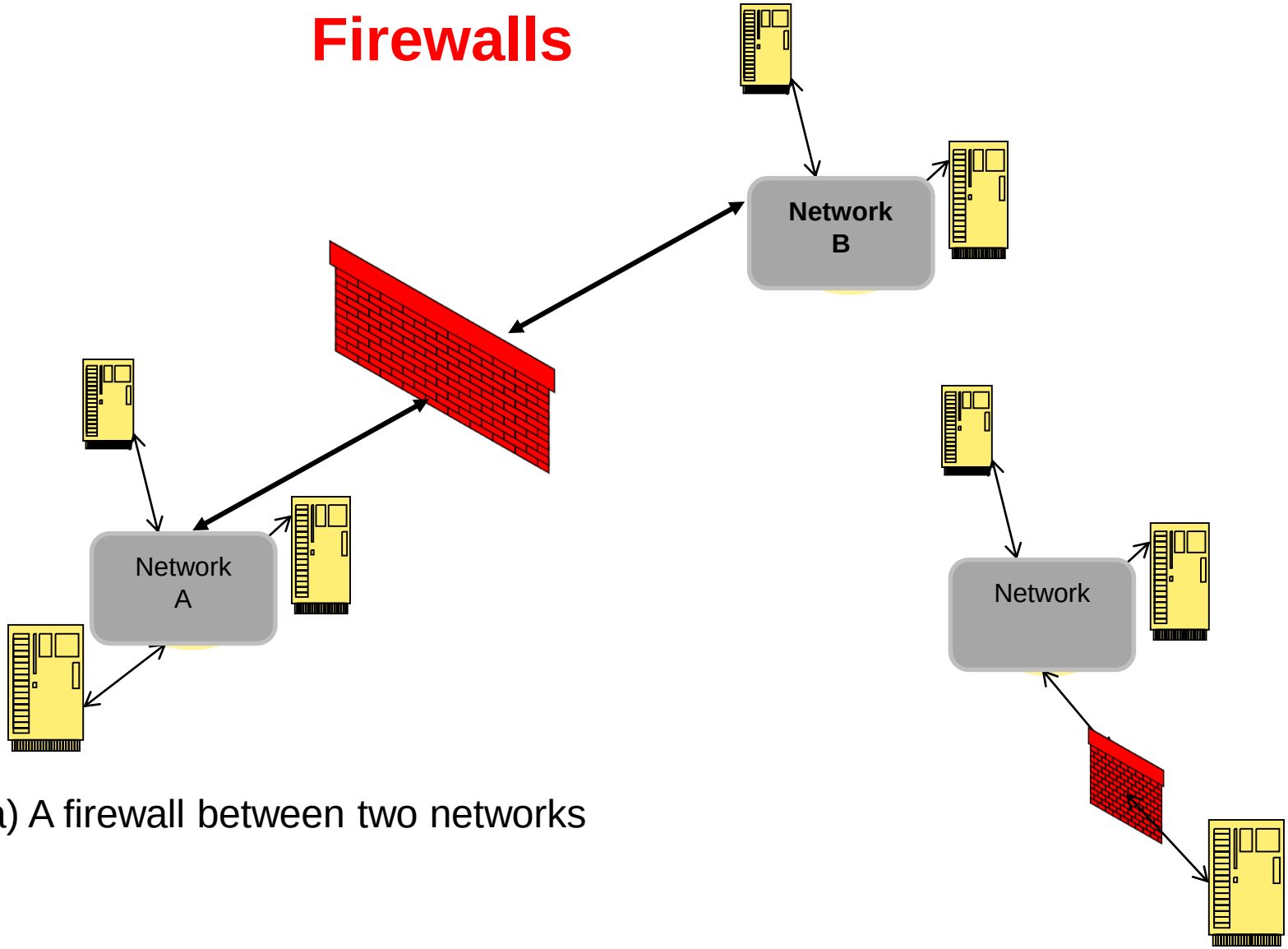
- Security of the top domain name servers is a major concern. ICANN has resolved it by scripting a ceremony (http://data.iana.org/ksk-ceremony/1/ceremony1-script-annotated.pdf) for signing the root
- Multiple *personas* (*Ceremony Administrator, Crypto-Officer, Safe Security Controller, Internal Witness*, and so on) are choreographed into performing the required steps:
  - ensuring that the safe for the key storage is initially empty
  - bringing the key generation equipment into the room
  - generating and signing key—producing the certificates
  - backing the keys up on a smart card,
  - storing the recovery key material (in the tamper-proof hardware security modules), which are then placed in various safe-deposit boxes..
- All of this is done in front of the *auditors*, and each step is carefully documented in one or another log.
- Of course, all participants are themselves authenticated through their respective government-issued IDs.
- Different rooms have different entry permits—not everyone may enter *the safe room*, for example.
- Conversely, until the end of the ceremony or a pre-defined break, no one may leave the ceremony room. There are also procedures for the annual inventories of the recovery material.



# On the role of ICANN

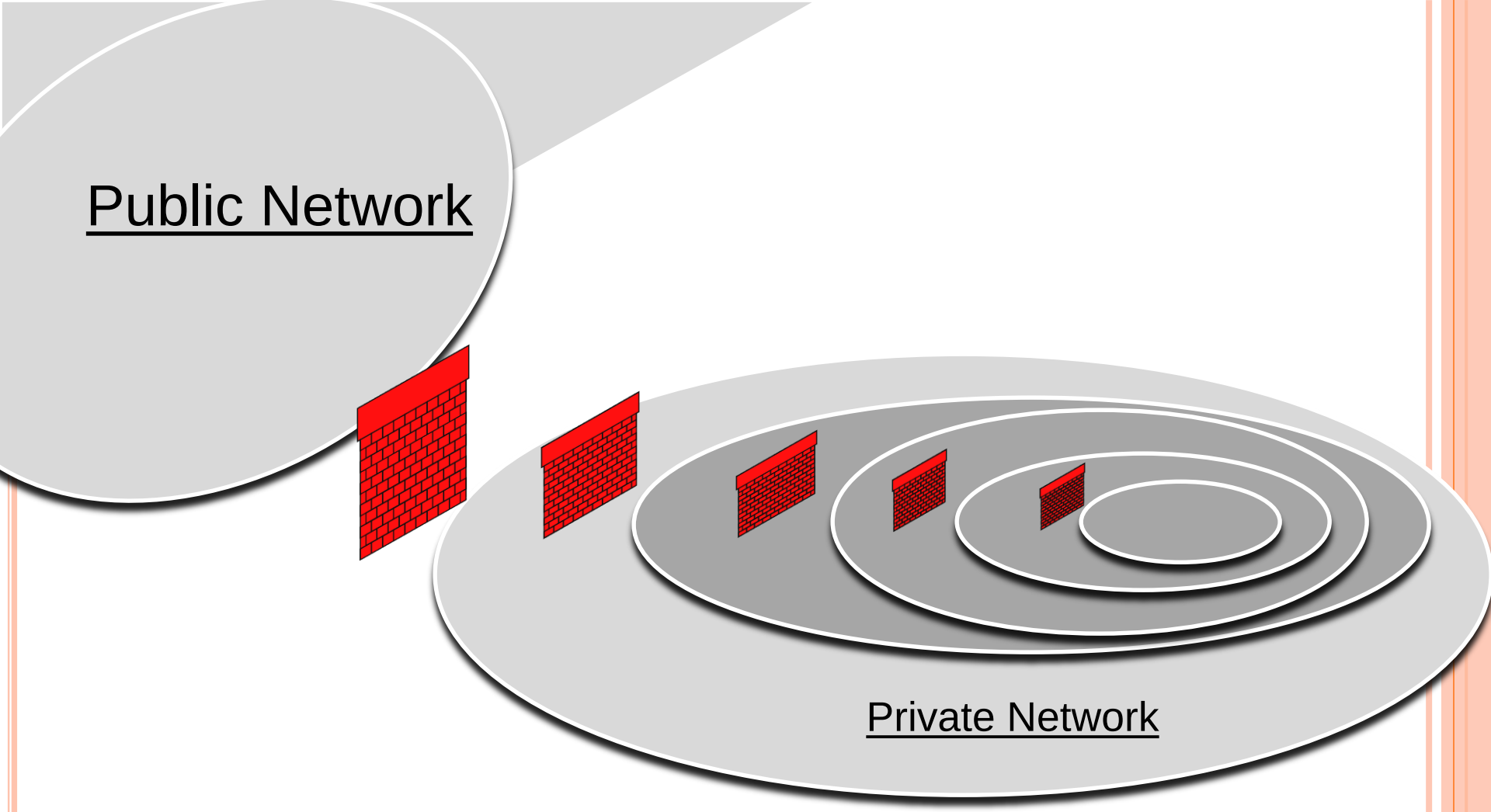
- ICANN has
  - Decentralized the sale and distribution of domain names, resulting in a dramatic drop in the price of registration.
  - Established an effective mechanism for resolving trademark dispute, which, in turn, has diminished the problem of 'cybersquatting'
  - Maintained stability in the naming and numbering system
- Recommended reading: Goldsmith, Jack L. and Tim Wu (2006). *Who Controls the Internet? Illusions of a Borderless World*. Oxford University Press.

# Firewalls

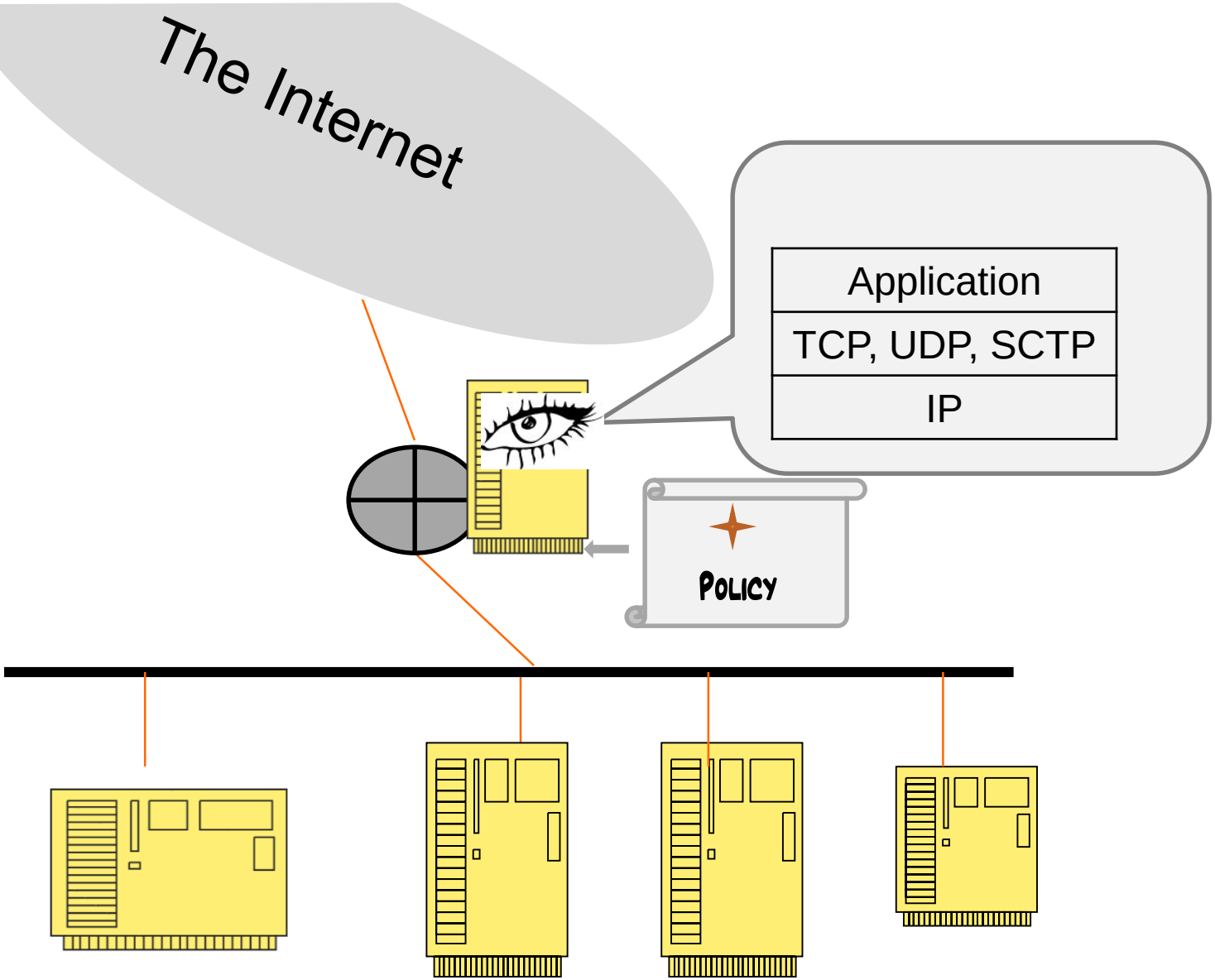



a) A firewall between two networks

b) A firewall protecting a single host

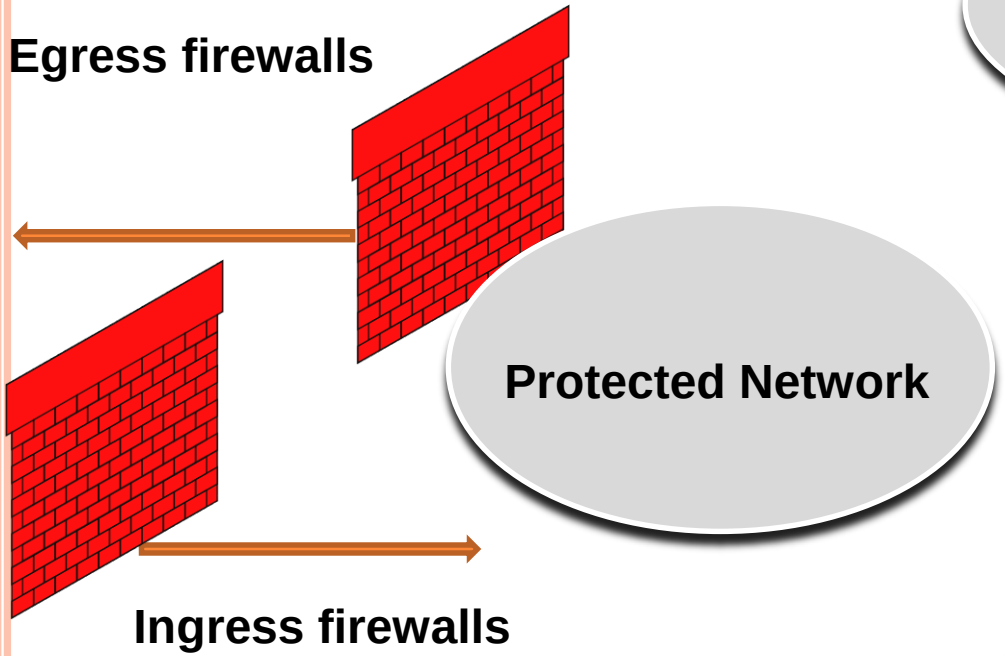



**Interconnecting networks with different security postures**

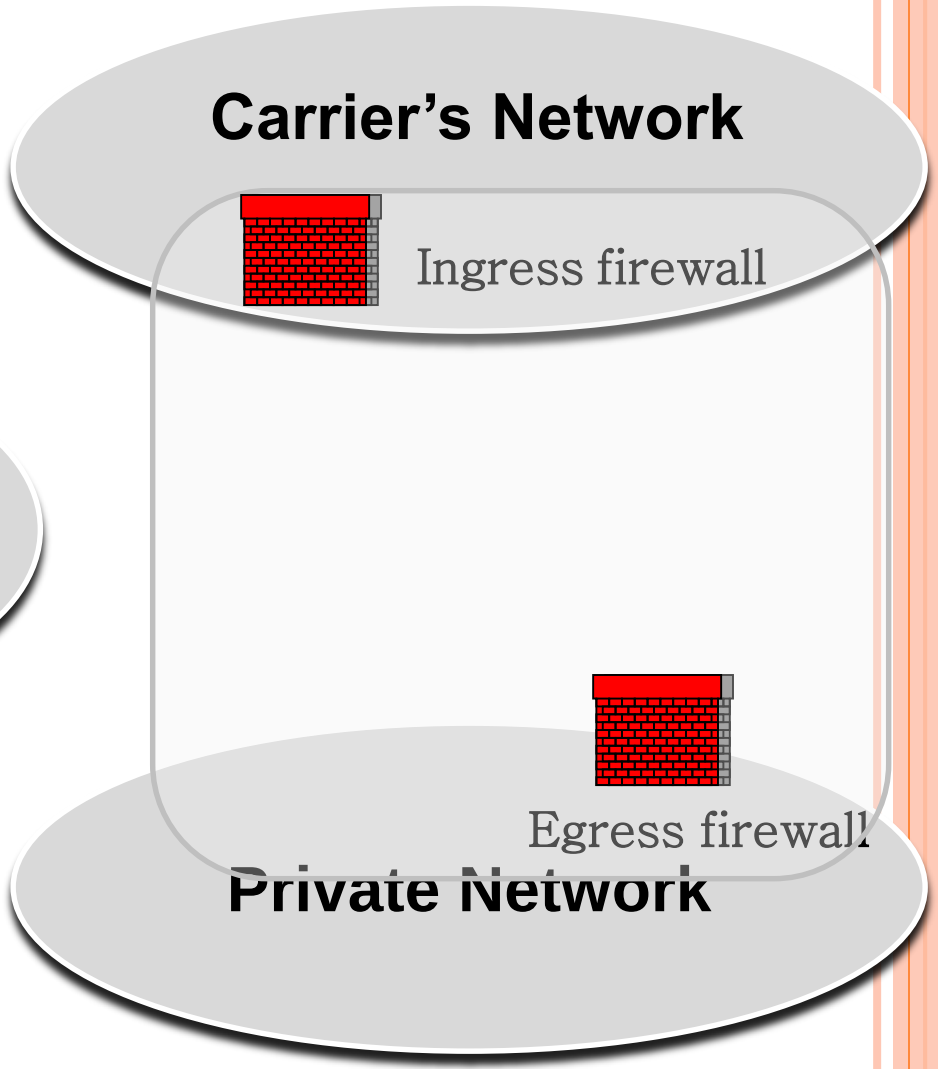



**The application gateway**

# Ingress and egress firewalls

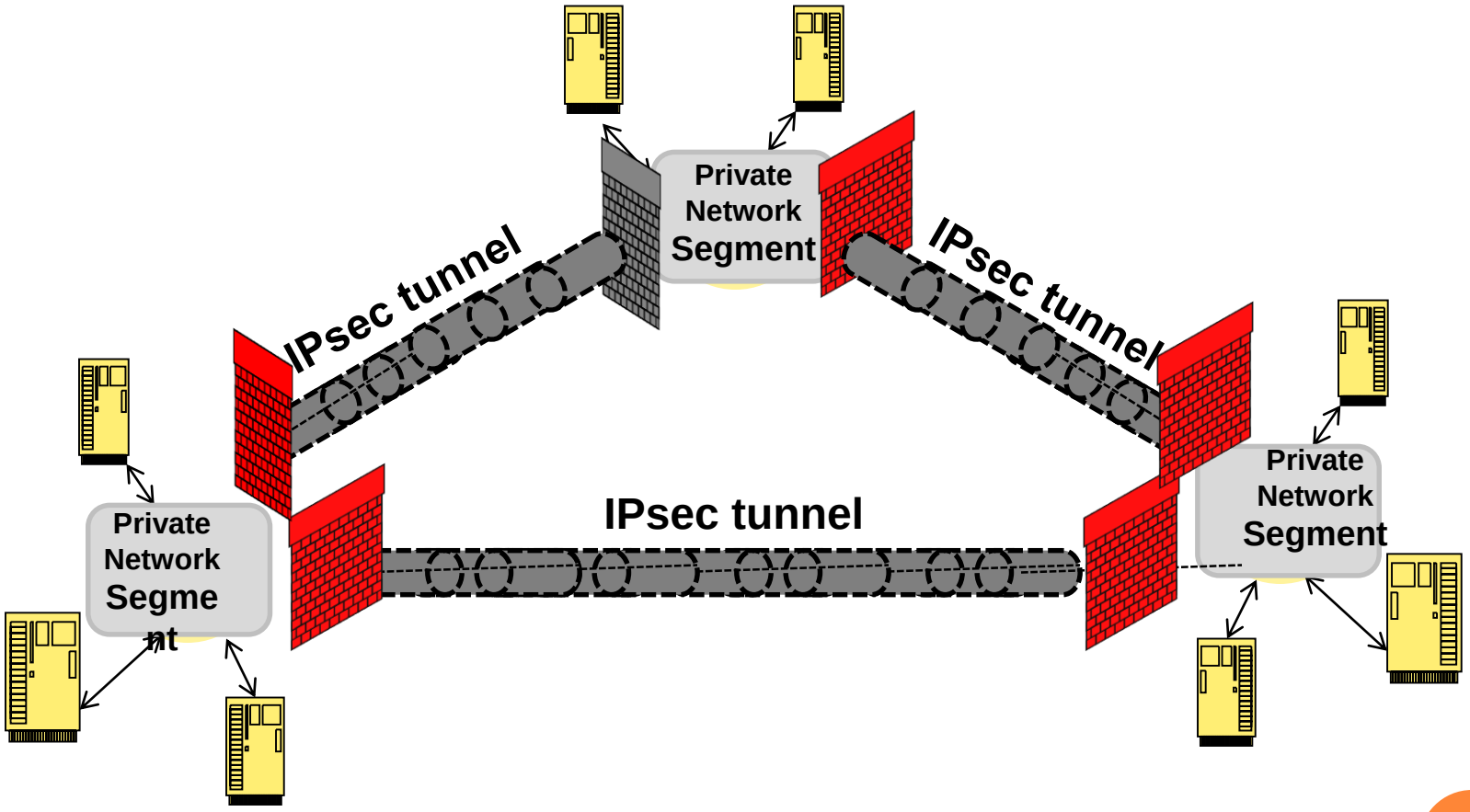



a) Interfaces




b) Split CPE

# Layer 3 VPN with firewalls



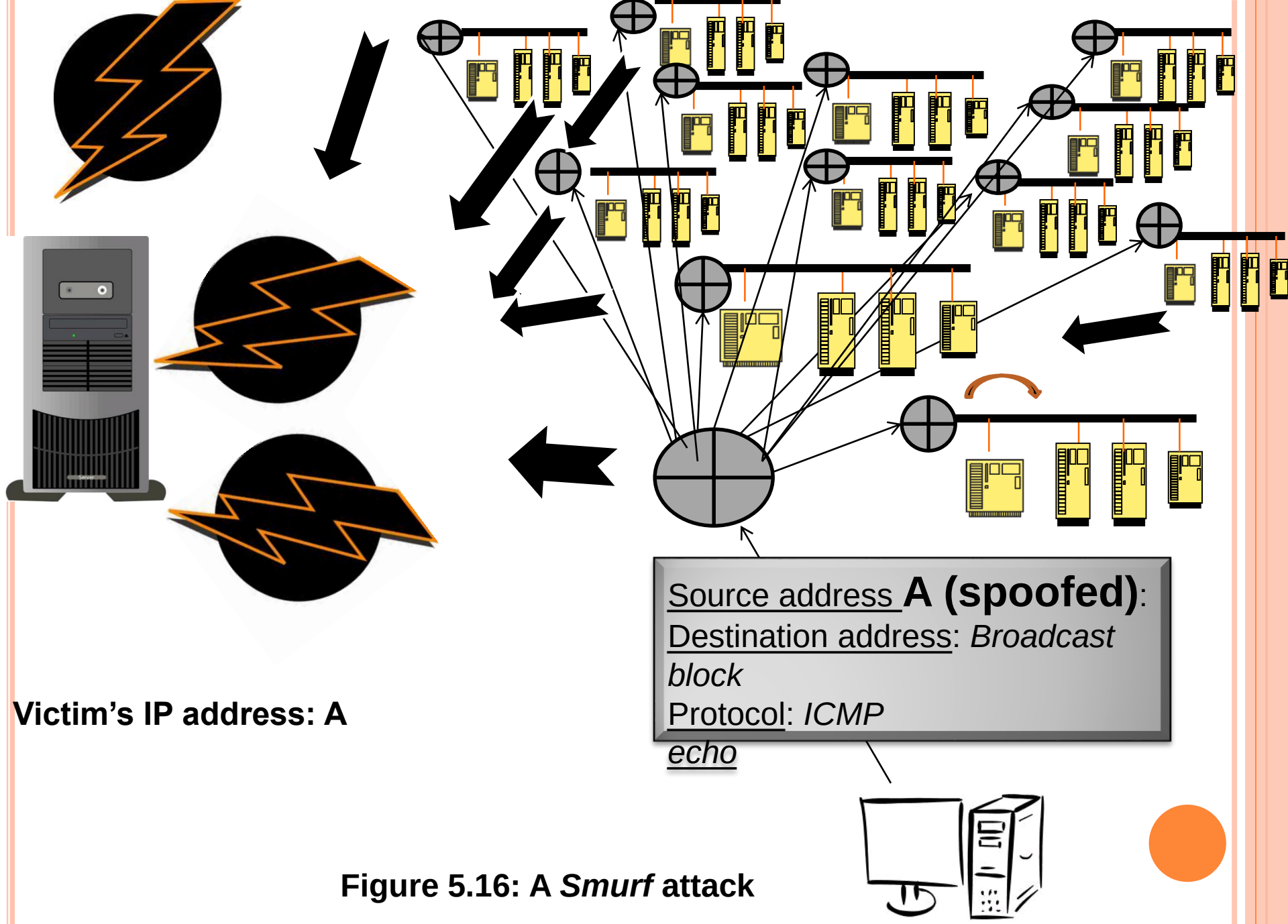
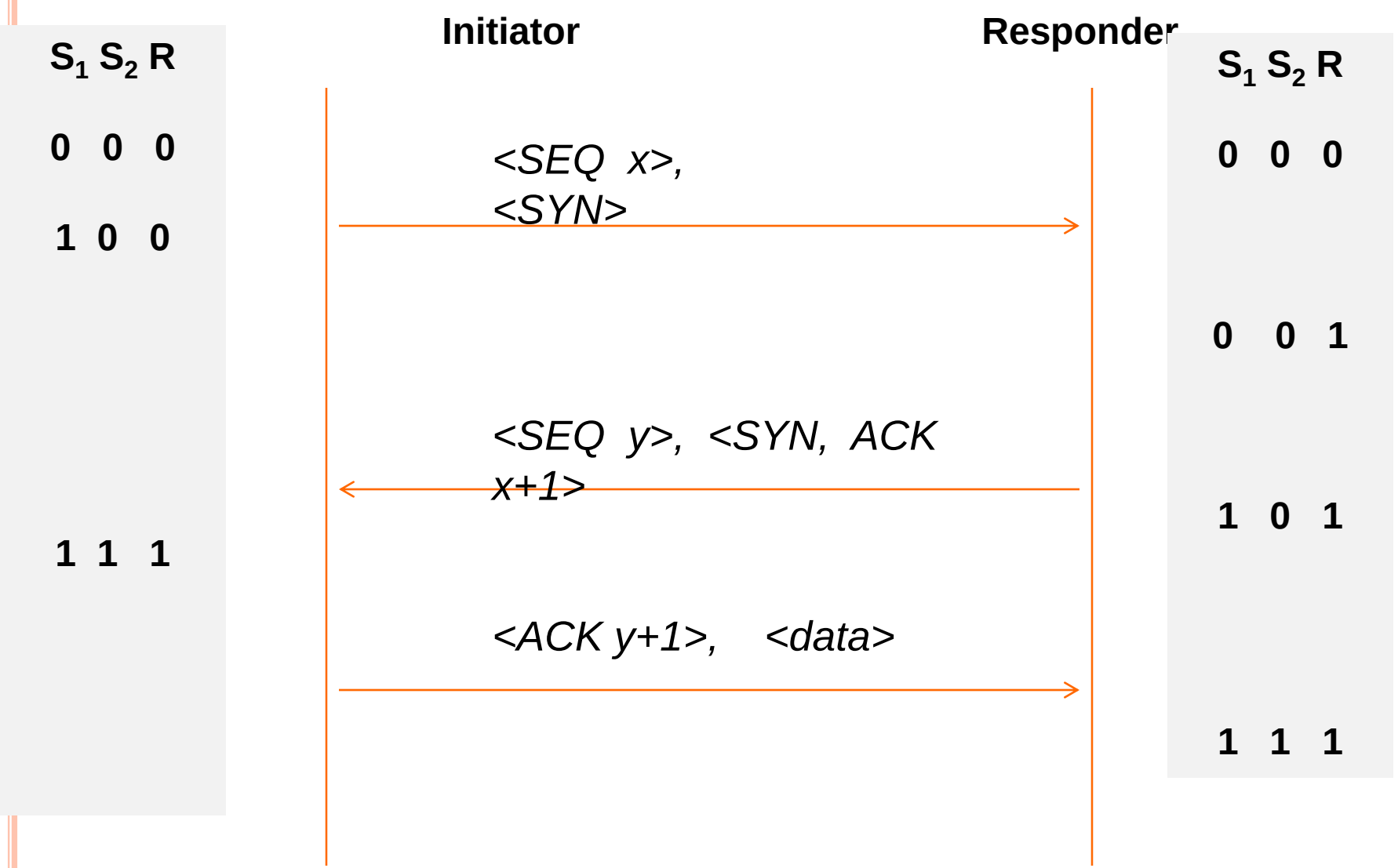
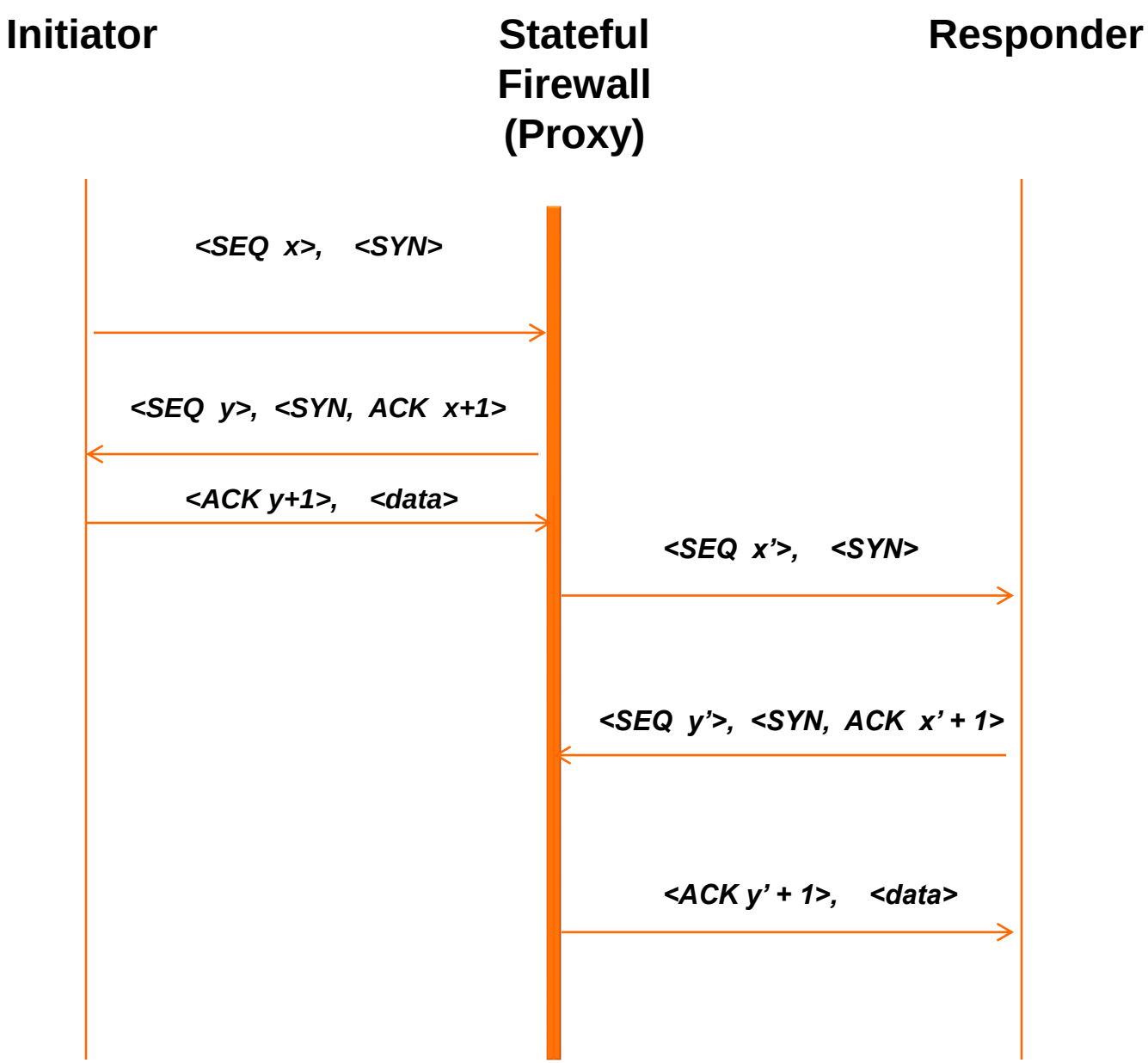


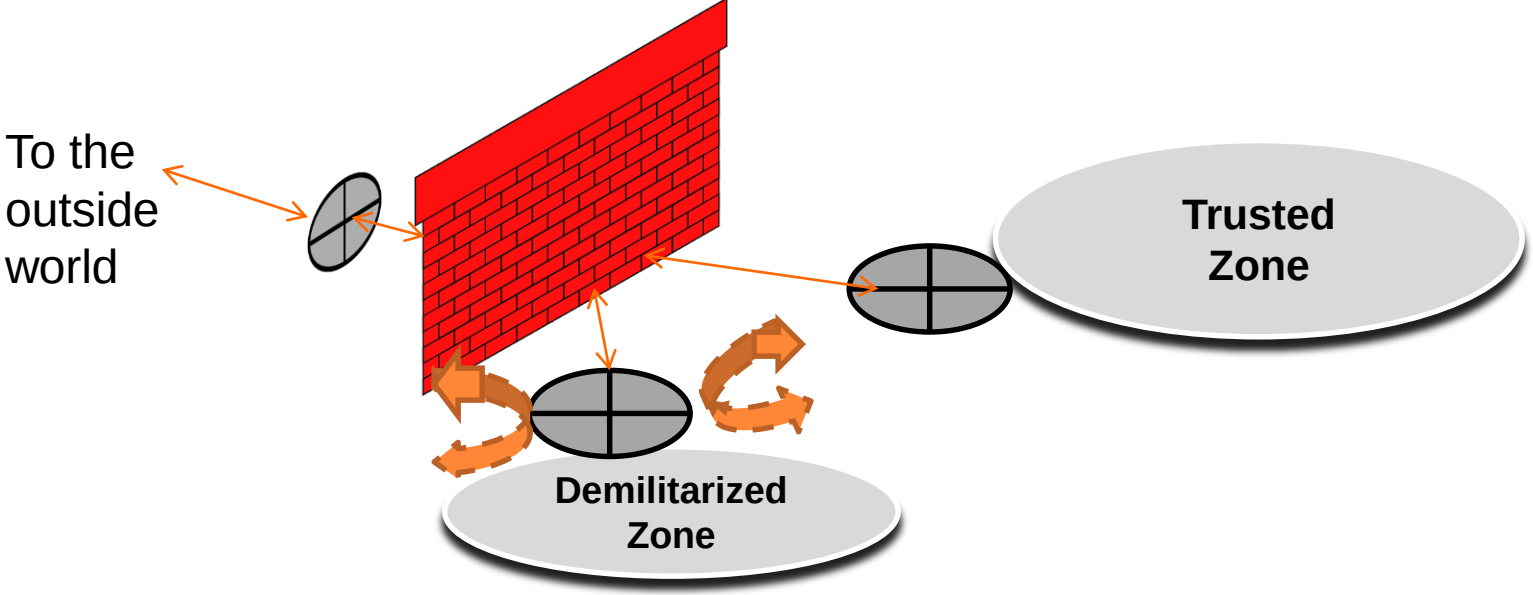

**Figure 5.16: A *Smurf* attack**



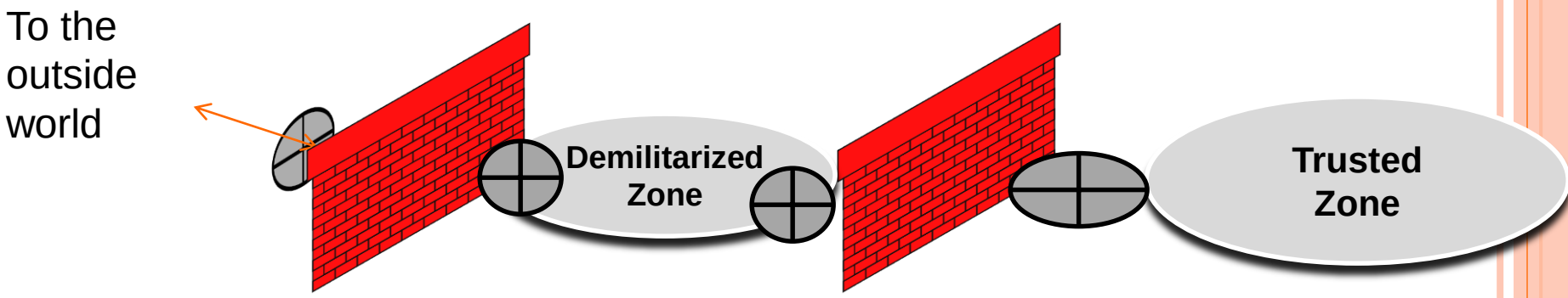

**TCP Connection Establishment (after RFC 675)**




**Stateful firewall (an example of a TCP connection establishment)**




a) **With a single firewall**




b) **With two firewalls**

## Network Zoning