HTTP AND DNS

HTTP...

Generally web is a set of co-operating clients and servers all using the same language HTTP.

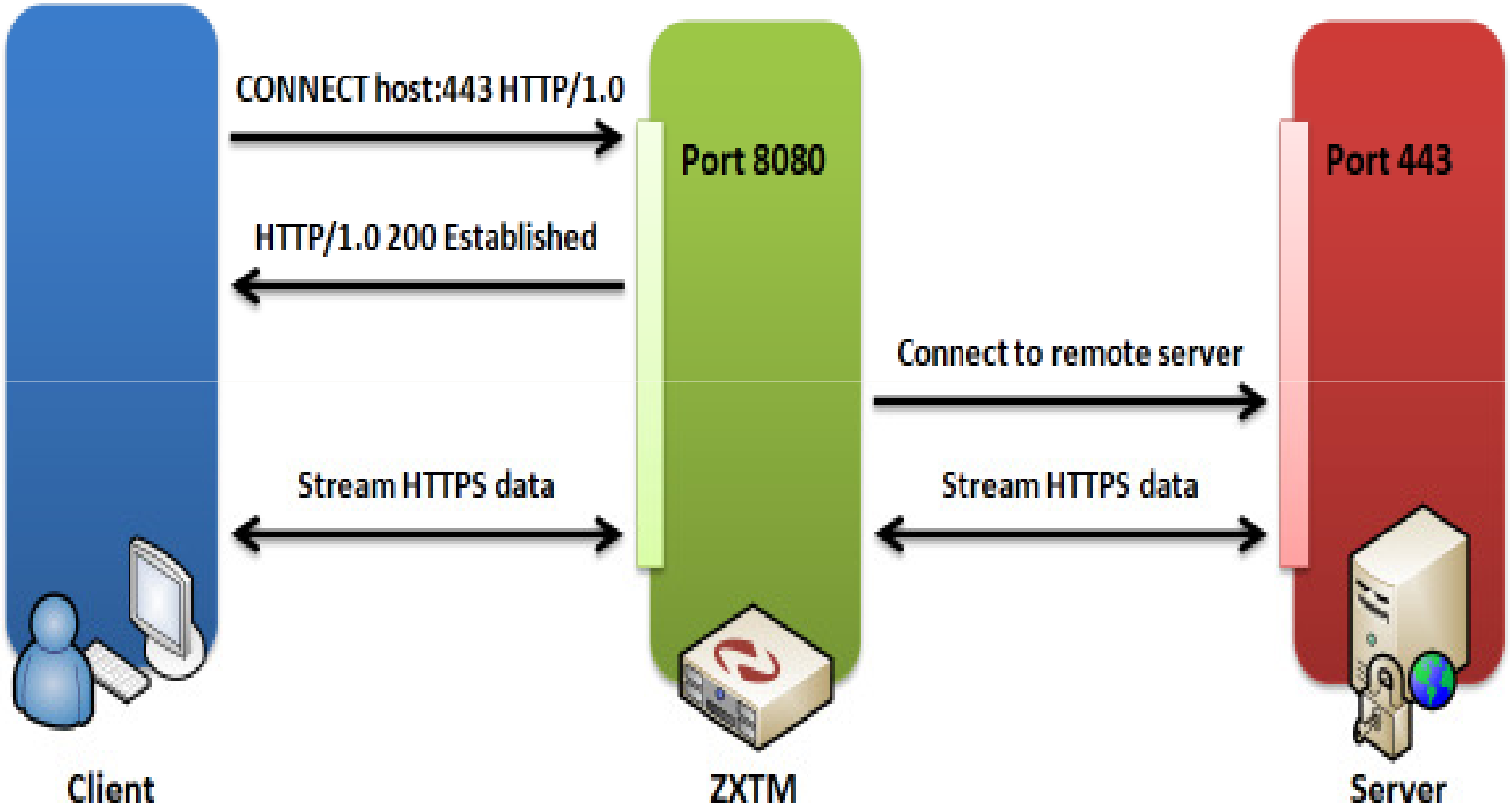
HTTP is an application protocol that is used to retrieve the web pages from the remote server.

It is similar to FTP, but it is simpler than FTP because it only uses the TCP connection.

The data transferred by protocol are either

1. **Plain text.**
2. **Cipher text.**
3. **Audio, images or any other internet accessible information.**

HTTP...


CONNECT host:443 HTTP/1.0
Port 8080
Port 443
HTTP/1.0 200 Established
Connect to remote server
Stream HTTPS data
Stream HTTPS data
Client
ZXTM
Server

STEPS IN HTTP OPERATION...

- Client opens a TCP connection which is the end to end between client and server.

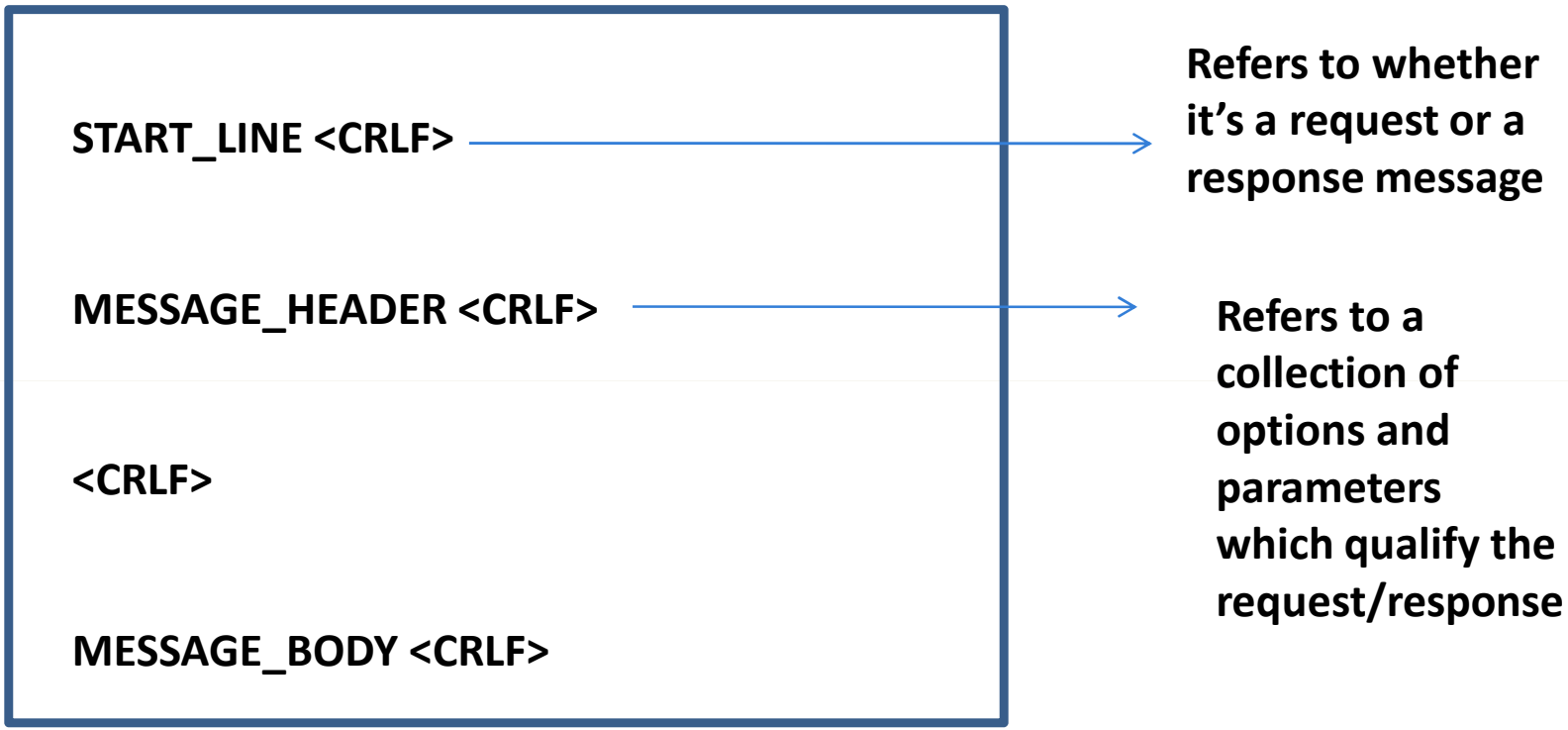
- Client gives a http request. Request has a specific command or a method named URL, MIME like message which has message like parameters, client information and some additional content information.

- When server gets the request, it tries to do the requested action and then it returns a HTTP response.

- Response has the status information, a success or an error code and a MIME like message which has the server information and body content.

- At last the TCP connection is closed.

Each HTTP message has the general form….



Where CRLF refers to Carriage Return Line Feed.

HTTP REQUEST MESSAGE

This provides 3 things. They are :-

1.Operation to be done.

2. The webpage on which operation has to do.

3. HTTP version.

The Request line tells the type of request, resource and HTTP version.

The request type classifies the request message into one or more methods.

OPERATIONS	MEANINGS
Options	Request information about available options
GET	Have the document found in the URL
HEAD	Get meta data about the document in the URL
POST	To transmit the information to the server
PUT	Store document under the specified URL
DELETE	Remove the particular URL
TRACE	Loopback request message
CONNECT	Used by proxies

HEADERS

It is used to exchange information between client and server.

It has one or more number or header lines.

Every header line consists of the name of the header, a colon space and the header value.

UNIFORM RESOURCE IDENTIFIERS

The URL's that HTTP uses as addresses are one type of Uniform Resource Identifiers.

A URI is a character string which finds a resource. The resource can be anything which has an identity like a document, an image or a service.

A URI qualifies as an URL only if it is intended to be used to locate a resource.

TCP CONNECTIONS

The HTTP created a separate TCP connection for each data item retrieved from the server.

HTTP uses both non-persistent and persistent connections.

One TCP connection is established in a non-persistent connection to make a request and to get a response.

In persistent connections, subsequent request and response between the same client and server can be transmitted over the same connection.

CACHING

It is important to know how to effectively cache web pages in the internet.

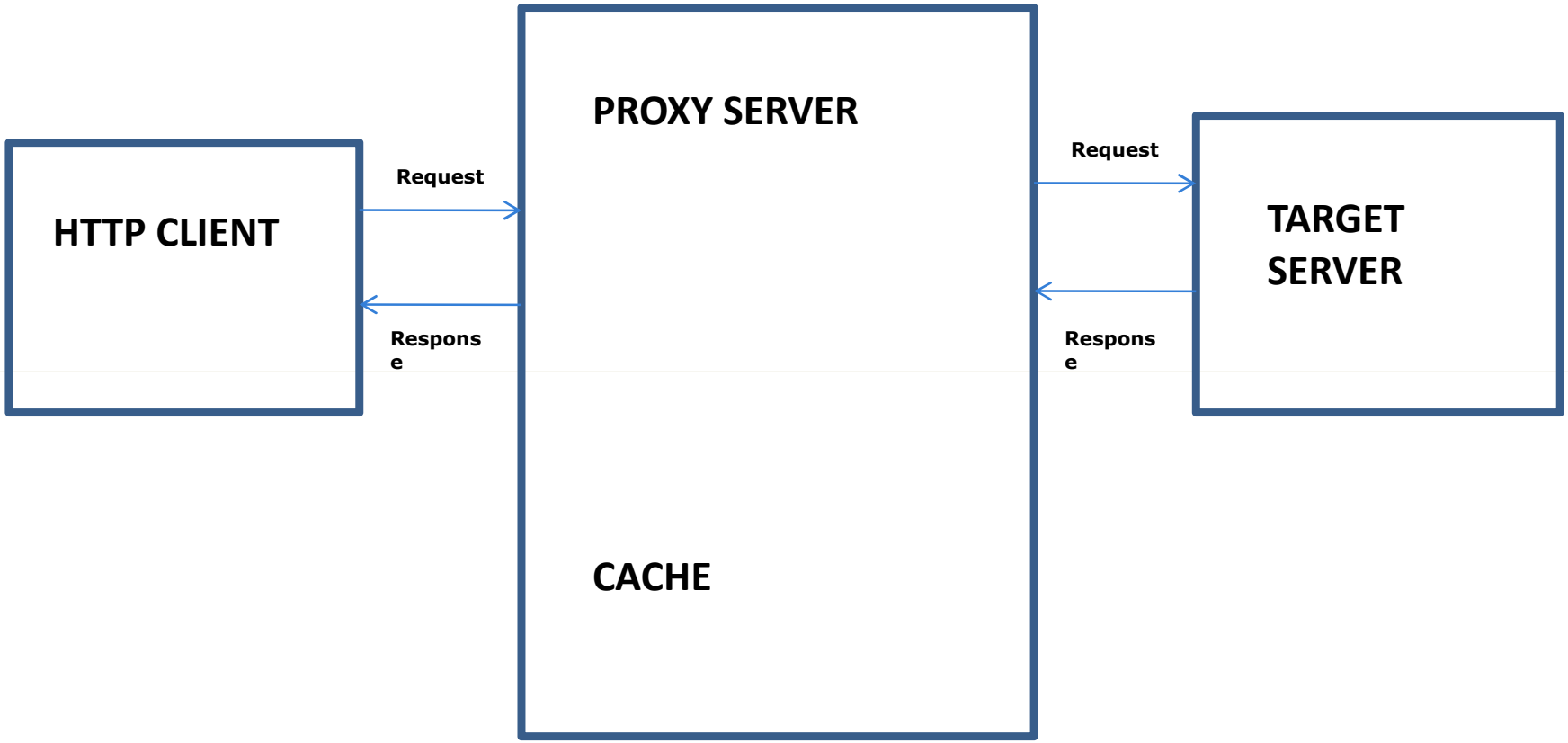
Caching has a lot of benefits. They are :-

A page can be taken from nearby cache and is displayed faster than it has ot be taken from across the world.

Having a cache intercept and satisfy the request minimizes the load on the server.

It is mostly implemented using a proxy server.

PROXY SERVERS



STEPS...

1. First, a HTTP client makes a request to proxy server.
2. It checks the cache to see if it has a corresponding response.
3. If there is no response, it forwards the request to the target server.
4. Target server issues a response. It is forwarded through a proxy server.
5. Proxy server now holds a copy of the incoming responses for future response to other clients.
6. Also, it sends a response to a HTTP client.

DNS...

Generally, addresses are not exactly user friendly. So a unique name is given to every host in a network.

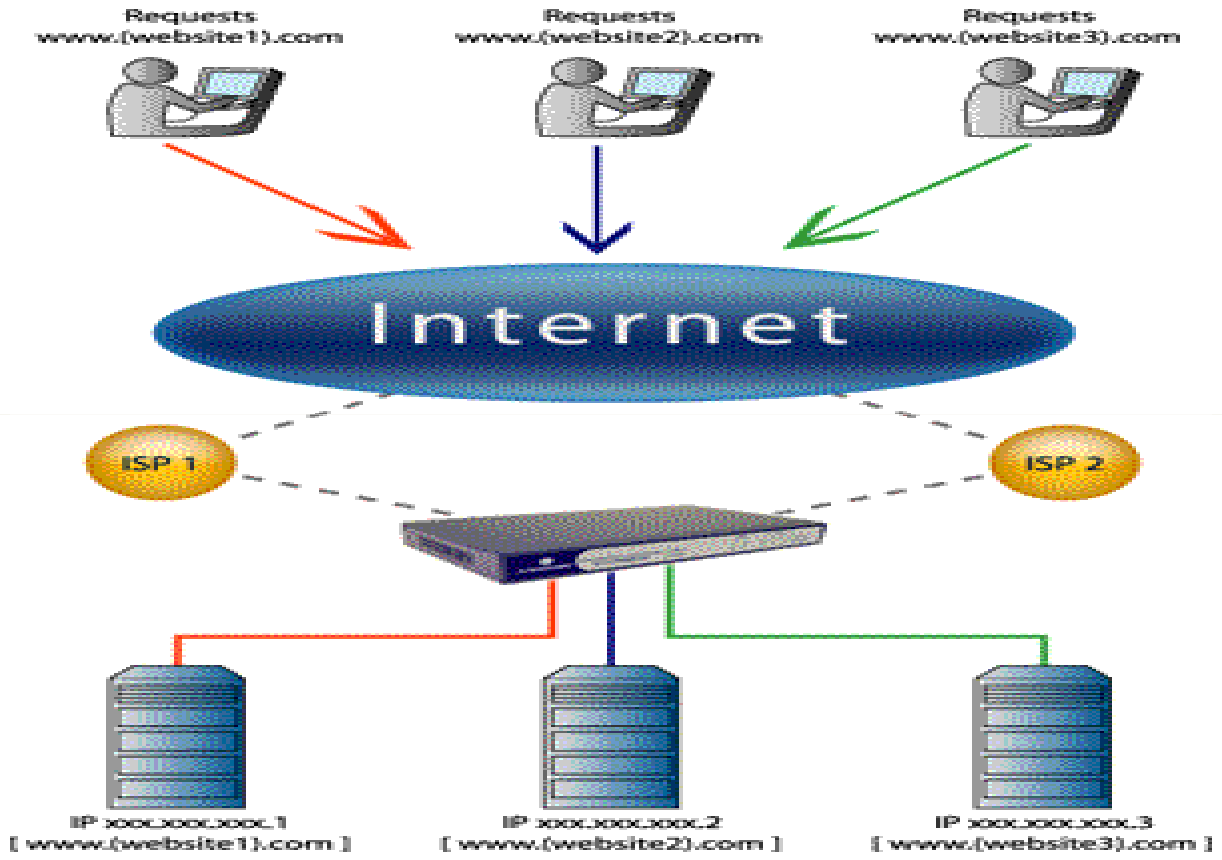
Host names are different from host addresses in 2 ways namely

They are of some variable name and mnemonic.

Do not have information which help the network to locate the host.

Some of the basic terminology include Namespace, Resolution mechanism, and Name Server.

DNS...


Requests
www.{website1}.com
Requests
www.{website2}.com
Requests
www.{website3}.com
Internet
ISP 1
ISP 2
IP xxx.xxx.xxx.1
[www.{website1}.com]
IP xxx.xxx.xxx.2
[www.{website2}.com]
IP xxx.xxx.xxx.3
[www.{website3}.com]

DNS...

Namespace

defines a set of possible names. Can be of 2 types namely flat or hierarchial namespace.

Flat namespace is assigned to an address. It is a sequence of characters without a structure. Ie. Cannot be divided into components.

Hierarchial namespace can be divided into components. Section 1 deals with organizational nature, Section 2 defines the name and Section 3 specifies the departments and so on.

DNS...

Resolution mechanism

results a corresponding value when it is invoked with a name.

Name server

is a specific implementation of a resolution mechanism that is available and that can be queried by sending a message.

TASK OF DNS...

There is a need of a directory service which converts the host names to IP addresses. It is the main task of a DNS.

DNS is a distributed database provided in a hierarchy of DNS servers.

DNS servers are often UNIX machines running the Berkley internet Name Domain Software (BIND).

It is an application layer protocol which allows the host to query the distributed database.

The DNS protocol runs over the UDP and it uses the port 53 such as HTTP, SMTPO and FTP in order to convert it to the user supplied host names to IP addresses

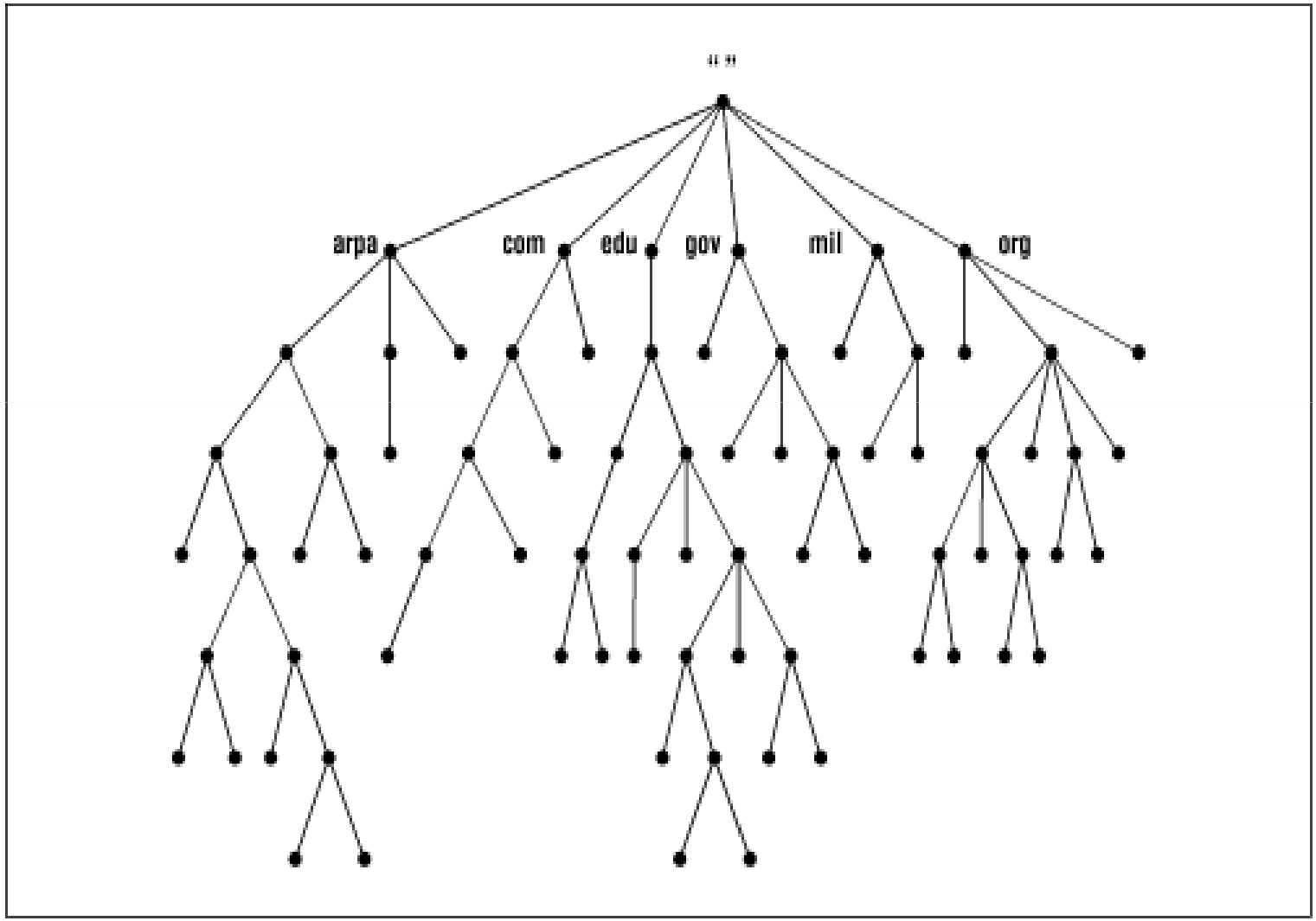
DOMAIN HIERARCHY

DNS implements a hierarchial namespace for internet objects. DNS names are processed right to left and it uses period as a separator. Eg. Domain name for a host is puma.cs.princeton.edu.

The hierarchy is not very wide at the first level. There are domains for each country plus 6 domains. The big 6 domains are your edu, com, gov, mil, org and net.

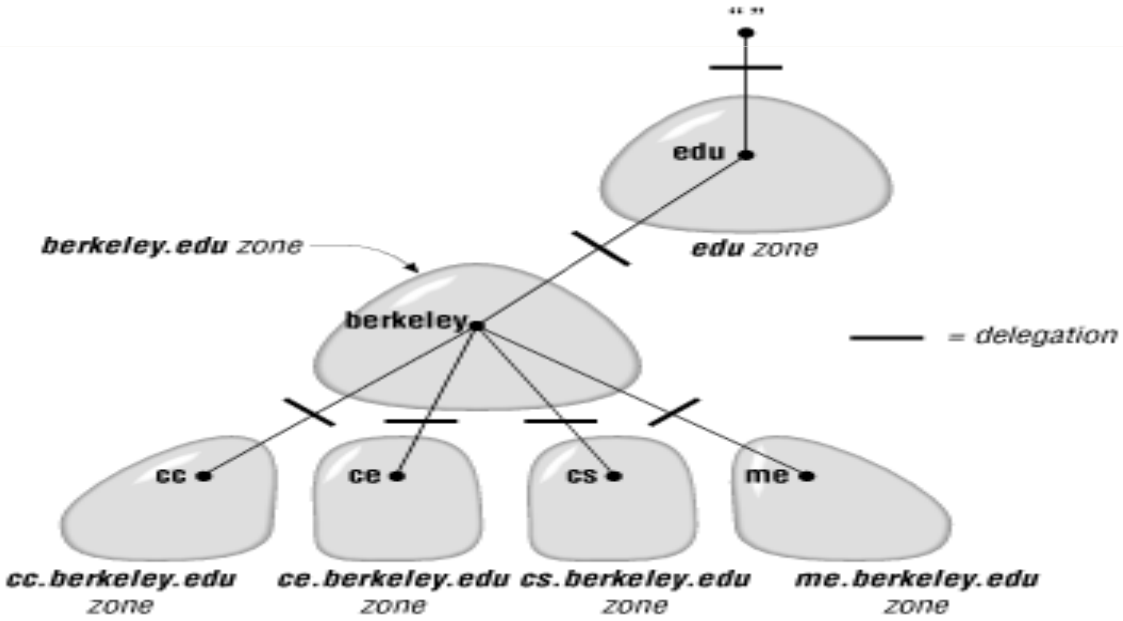
These domains are based in the United States.

DOMAIN HIERARCHY



NAME SERVERS...

The data has to be distributed in between many computers called DNS servers.
A single name server could have the full DNS database and respond to all queries about it.
Root is left alone and many sub trees are created as there are first level nodes.



ZONE...

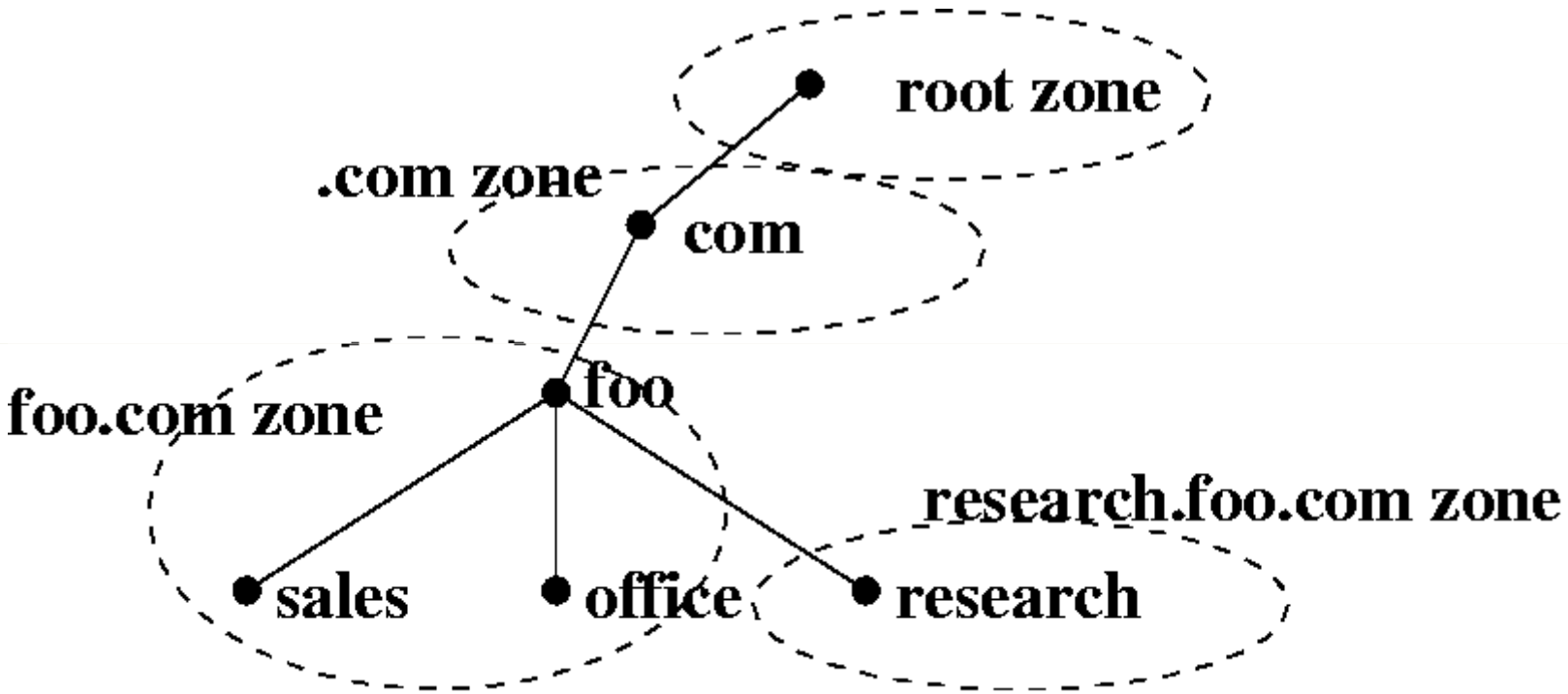
To avoid the problem associated with having only a single source of information, the DNS name is split into a non-overlapping zone.

Each zone has a part of a tree and it also has some name servers which contain authoritative data about the zone.

A zone will have one primary name server. And sometimes a server for a zone can be located outside the zone to increase its reliability.

Root server is a type of server whose zone has the entire tree. It does not store any information about the domains.

ZONE...


root zone
.com zone
com
foo
foo.com zone
research.foo.com zone
sales
office
research

RESOURCE RECORD

It is a name to value binding or a 5 tuple that contains some fields like

Name & Value fields refer the name and value of resource record.

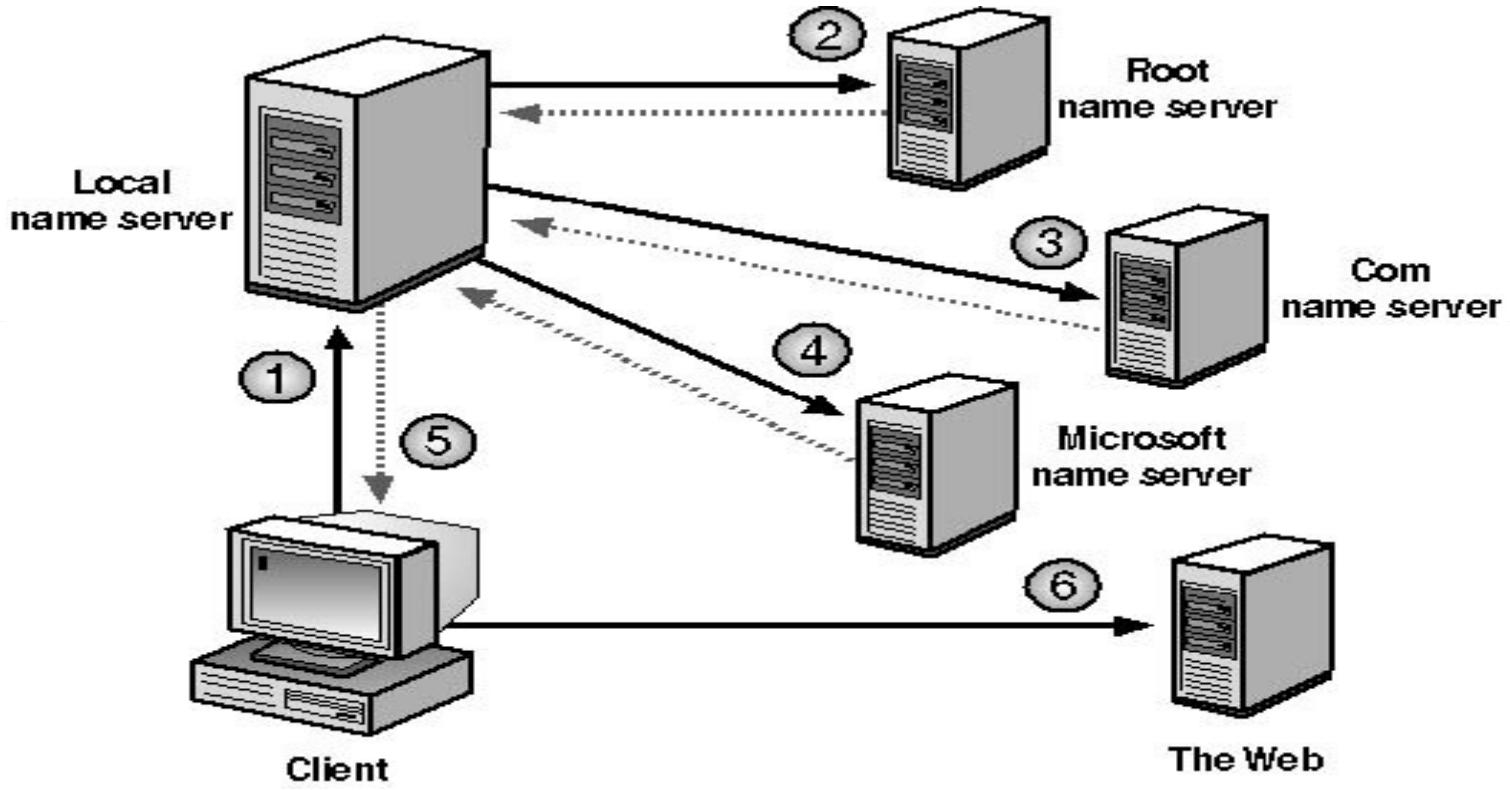
Type field says how the value should be inetrpreted.

NAME RESOLUTION

Generally, resolution is a process of a process of mapping a name to an address to name is known as name address resolution.

Resolver is a host which mapping calls a DNS client. The process of resolving a name actually involves a client query to the local server and it queries the remote servers on original clients behalf.

NAME RESOLUTION



THANK YOU!!