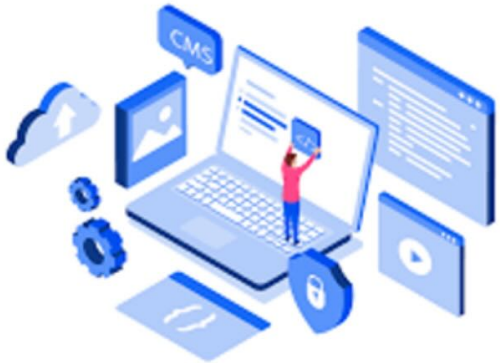
CSC-304
Web Programming

HTML , CSS , JavaScript , Git-hub based web development.



Section 3 of 8

CLIENT-SERVER MODEL

Client-Server Model

What is it?

The web is sometimes referred to as a client-server model of communications.

In the **client-server model**, there are two types of actors: clients and servers.

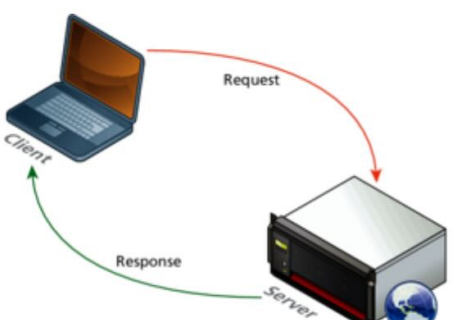
The **server** is a computer agent that is normally active 24 hours a day, 7 days a week (or simply 24/7), listening for queries from any client who make a request.

A **client** is a computer agent that makes requests and receives responses from the server, in the form of response codes, images, text files, and other data.

Request-Response Loop

Within the client-server model, the **request-response loop** is the most basic mechanism on the server for receiving requests and transmitting data in response.

The client initiates a **request** to a server and gets a **response** that could include some resource like an HTML file, an image or some other data.



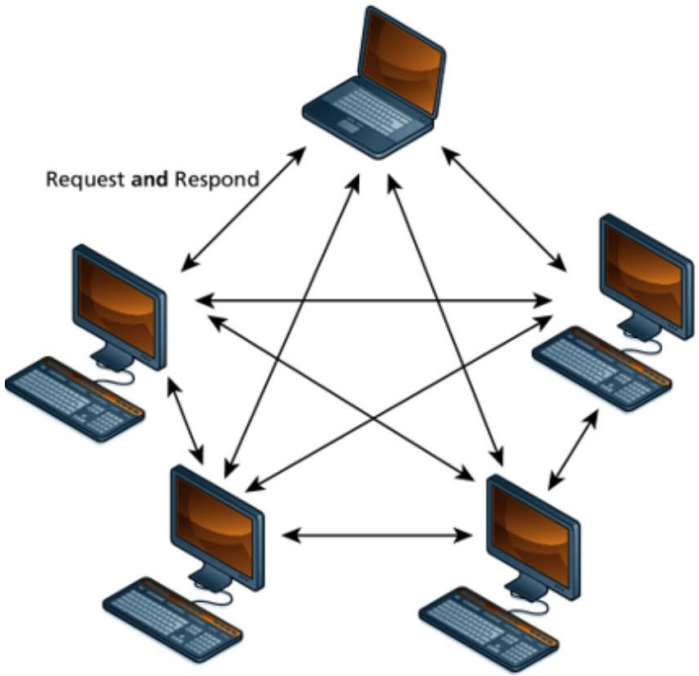
The Peer-to-Peer Alternative

Not actually illegal

In the **peer-to-peer model** where each computer is functionally identical, each node is able to send and receive directly with one another.

In such a model each peer acts as both a client and server able to upload and download information.

Peer-to-Peer Model



Server Types

A server is rarely just a single computer

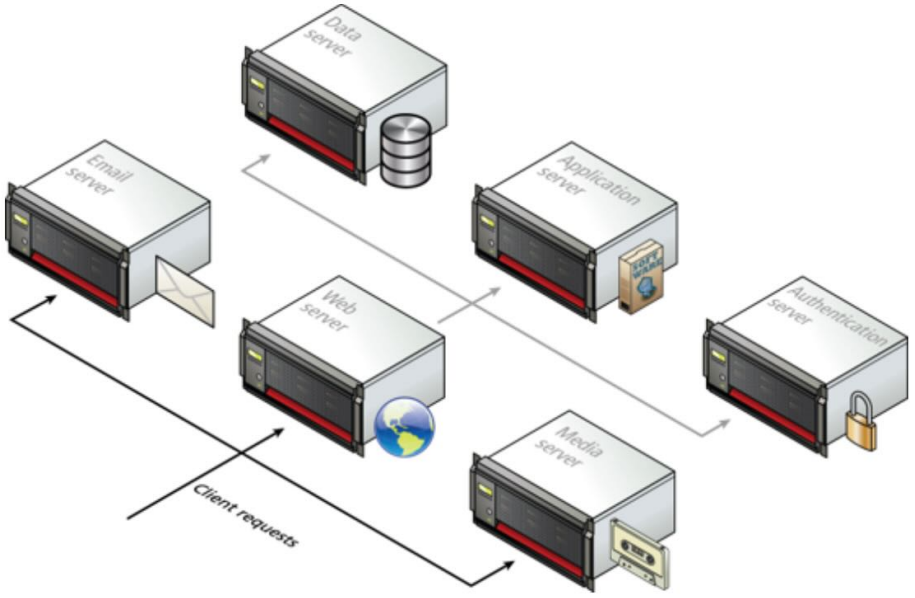
Earlier, the server was shown as a single machine, which is fine from a conceptual standpoint.

Clients make requests for resources from a URL; to the client, the server *is* a single machine.

However, most real-world web sites are typically not served from a single server machine, but by many servers.

It is common to split the functionality of a web site between several different types of server.

Server Types



Real-World Server Installations

Not only are there different types of servers, there is often replication of each of the different server types.

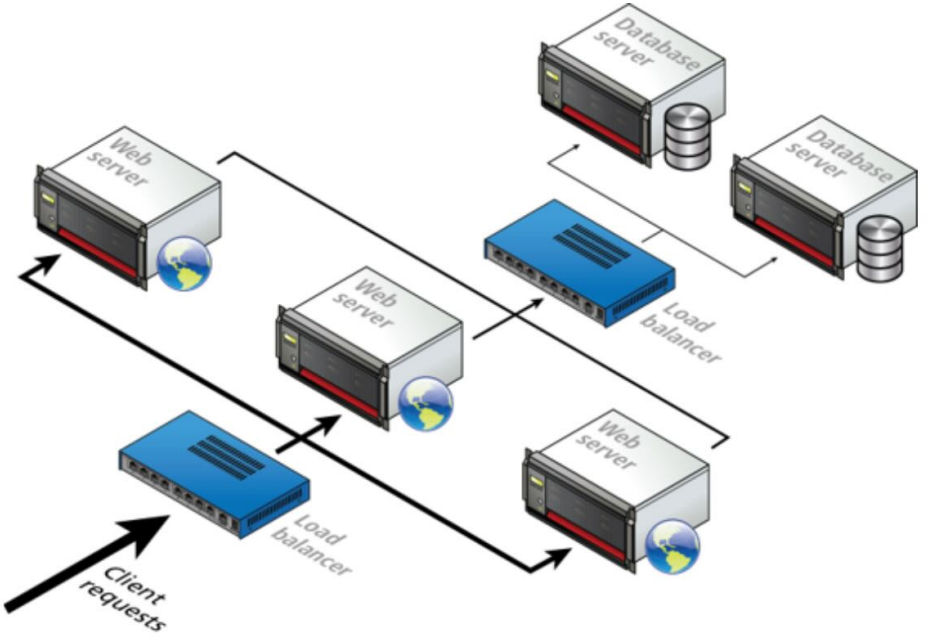
A busy site can receive thousands or even tens of thousands of requests a second; globally popular sites such as Facebook receive millions of requests a second.

Server Farms

Have no cows

A single web server that is also acting as an application or database server will be hard-pressed to handle more than a few hundred requests a second, so the usual strategy for busier sites is to use a **server farm**.

Server Farm


Database server
Database server
Web server
Web server
Web server
Load balancer
Load balancer
Client requests

Server Farms

The goal behind server farms is to distribute incoming requests between clusters of machines so that any given web or data server is not excessively overloaded.

Special routers called **load balancers** distribute incoming requests to available machines.

Server Farms

Even if a site can handle its load via a single server, it is not uncommon to still use a server farm because it provides **failover redundancy**.

That is, if the hardware fails in a single server, one of the replicated servers in the farm will maintain the site's availability.

Server Racks

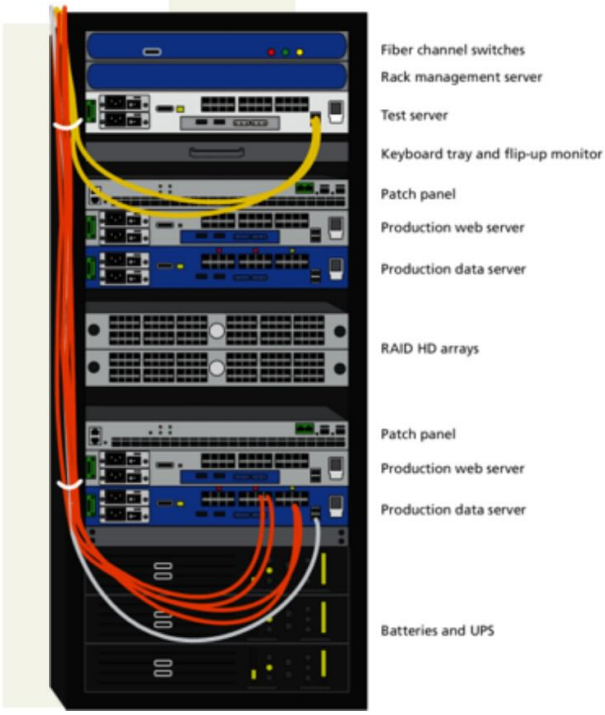
In a server farm, the computers do not look like the ones in your house.

Instead, these computers are more like the plates stacked in your kitchen cabinets.

That is, a farm will have its servers and hard drives stacked on top of each other in **server racks**.

A typical server farm will consist of many server racks, each containing many servers.

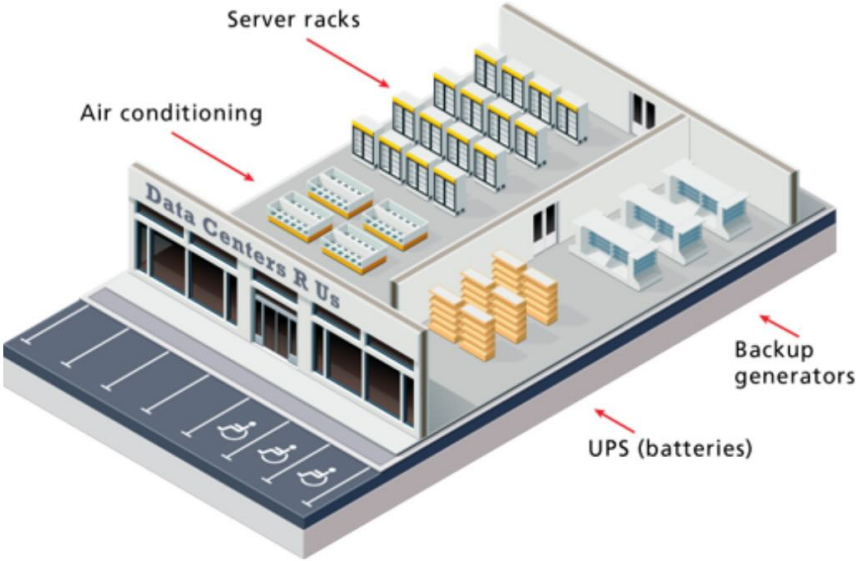
Server Rack


Fiber channel switches
Rack management server
Test server
Keyboard tray and flip-up monitor
Patch panel
Production web server
Production data server
RAID HD arrays
Patch panel
Production web server
Production data server
Batteries and UPS

Data Centers

Server farms are typically housed in special facilities called **data centers**.

Hypothetical Data Center



Data Centers

Where are they?

To prevent the potential for site down times, most large web sites will exist in mirrored data centers in different parts of the country, or even world.

As a consequence, the costs for multiple redundant data centers are quite high, and only larger web companies can afford to create and manage their own.

Most web companies will instead lease space from a third-party data center.

Commercial Web Hosting

It is also common for the reverse to be true – that is, a single server machine may host multiple sites.

Large commercial web hosting companies such as GoDaddy, Blue Host, Dreamhost, and others will typically host hundreds or even thousands of sites on a single machine (or mirrored on several servers).

Section 4 of 8

WHERE IS THE INTERNET?

Is the Internet a Cloud?

The Internet is often visually represented as a cloud, which is perhaps an apt way to think about the Internet given the importance of light and magnetic pulses to its operation.

Is the Internet a Cloud?

No

It is important to recognize that our global network of networks does not work using magical water vapor, but is implemented via

millions of kilometers of copper wires and fiber optic cables, as well as via

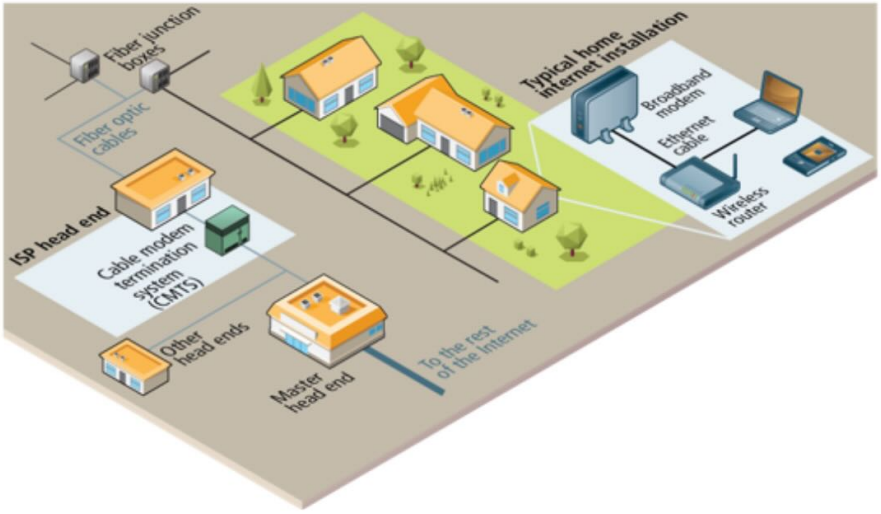
hundreds of thousands of server computers

and probably an equal number of routers, switches, and other networked devices,

along with many thousands of air conditioning units and specially-constructed server rooms and buildings.

From the Computer to the Local Provider

Our main experience of the hardware component of the Internet is that which we experience in our homes.



In the House

The **broadband modem** (also called a cable modem or DSL modem) is a bridge between the network hardware outside the house (typically controlled by a phone or cable company) and the network hardware inside the house.

These devices are often supplied by the ISP.

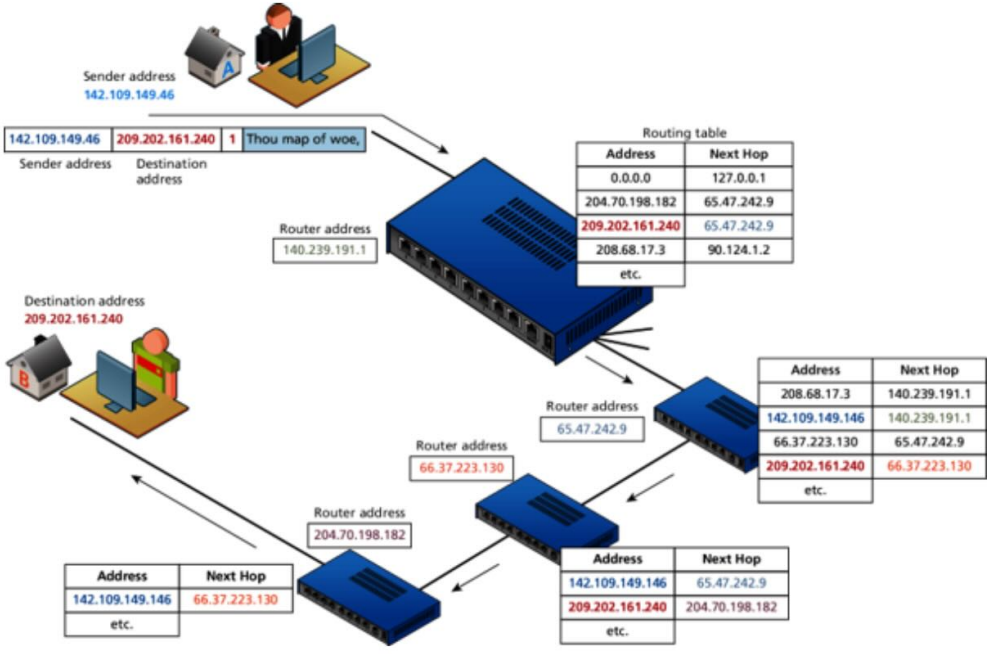
Routers

The wireless router is perhaps the most visible manifestation of the Internet in one's home, in that it is a device we typically need to purchase and install.

Routers are in fact one of the most important and ubiquitous hardware devices that makes the Internet work.

At its simplest, a **router** is a hardware device that forwards data packets from one network to another network.

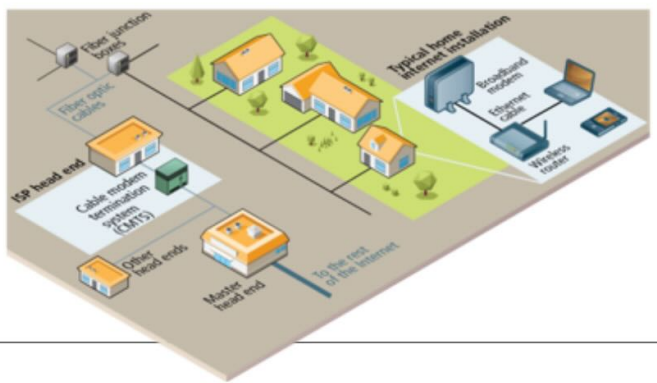
Routers and Routing Tables



Out of the House

Once we leave the confines of our own homes, the hardware of the Internet becomes much murkier.

In the illustration, the various neighborhood broadband cables (which are typically using copper, aluminum, or other metals) are aggregated and connected to fiber optic cable via fiber connection boxes.



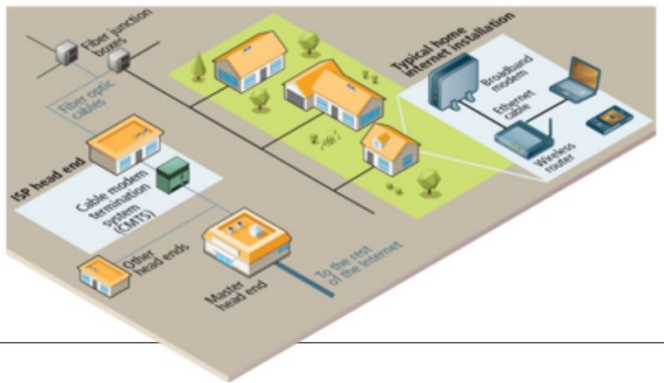
Fiber Optic Cable

Fiber optic cable (or simply optical fiber) is a glass-based wire that transmits light and has significantly greater bandwidth and speed in comparison to metal wires.

In some cities (or large buildings), you may have fiber optic cable going directly into individual buildings; in such a case the fiber junction box will reside in the building.

To the Provider

These fiber optic cables eventually make their way to an ISP's **head-end**, which is a facility that may contain a **cable modem termination system** (CMTS) or a digital subscriber line access multiplexer (DSLAM) in a DSL-based system.



From the Local Provider to the Ocean

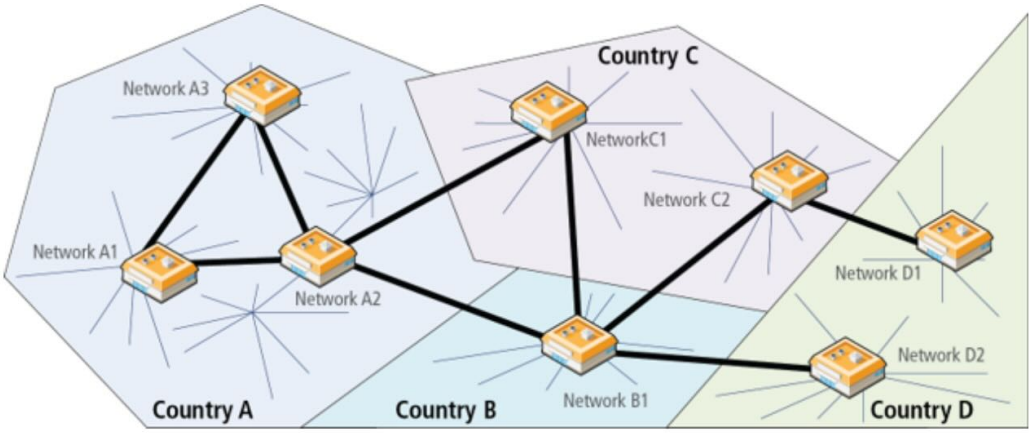
Eventually your ISP has to pass on your requests for Internet packets to other networks.

This intermediate step typically involves one or more regional network hubs.

Your ISP may have a large national network with optical fiber connecting most of the main cities in the country.

Some countries have multiple national or regional networks, each with their own optical network.

Connecting different networks within and between countries



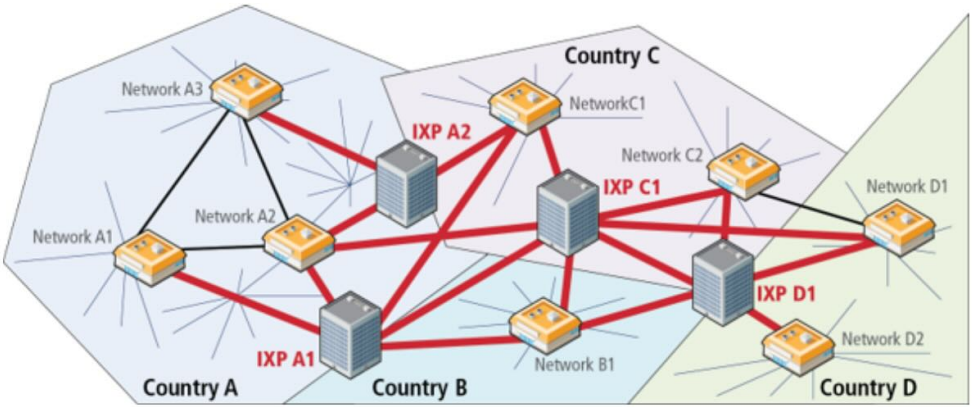
Internet Exchange Points

Connecting different networks

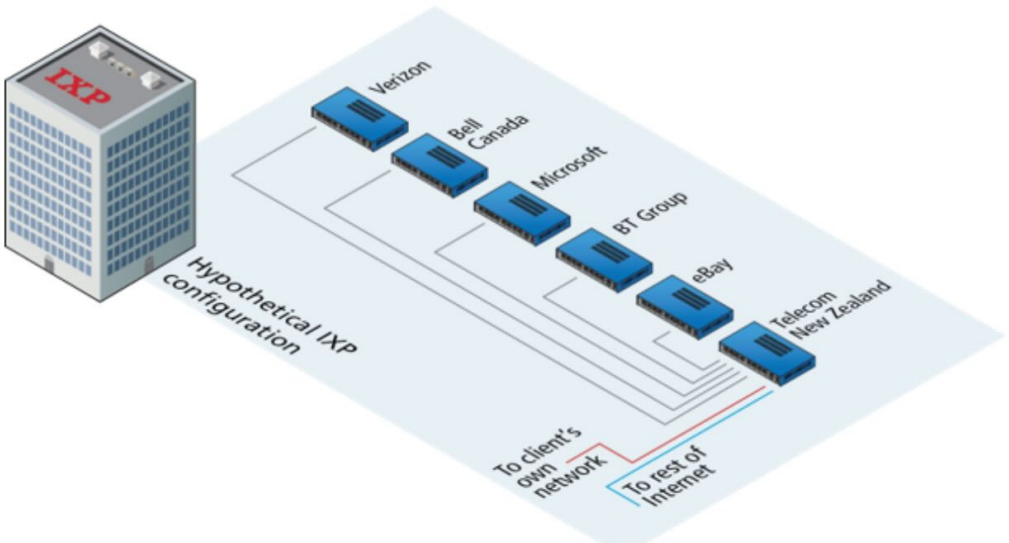
This type of network configuration began to change in the 2000s, as more and more networks began to interconnect with each other using an **Internet Exchange Point** (**IX** or **IXP**).

These IXPs allow different ISPs to **peer** with one another (that is, interconnect) in a shared facility, thereby improving performance for each partner in the peer relationship.

National and regional networks using Internet Exchange Points



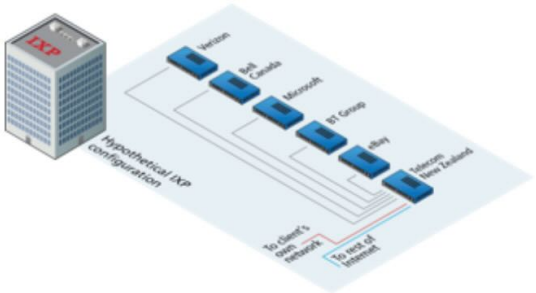
Sample Internet Exchange Point



IXPs

Not just for large networks

Different networks connect not only to other networks within an IXP, but now large web sites such as Microsoft and FaceBook are also connecting to multiple other networks simultaneously as a way of improving the performance of their sites.



Real IXPs

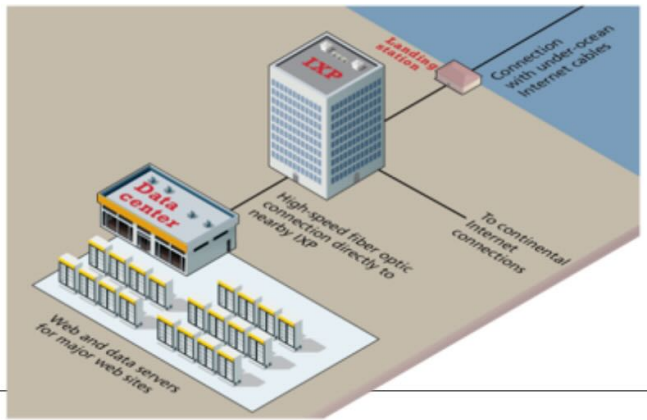
Real IXPs, such as at Palo Alto (PAIX), Amsterdam (AMS-IX), Frankfurt (CE-CIX), London (LINX), allow many hundreds of networks and companies to interconnect and have throughput of over 1000 gigabits per second.

The scale of peering in these IXPs is way beyond that shown in the diagram (which shows peering with only five others); companies within these IXPs use large routers from Cisco and Brocade that have hundreds of ports allowing hundreds of simultaneous peering relationships.

IXPs and Data Centers

In recent years, major web companies have joined the network companies in making use of IXPs.

As shown in the diagram, this sometimes involves mirroring a site's infrastructure (i.e., web and data servers) in a data center located near the IXP.

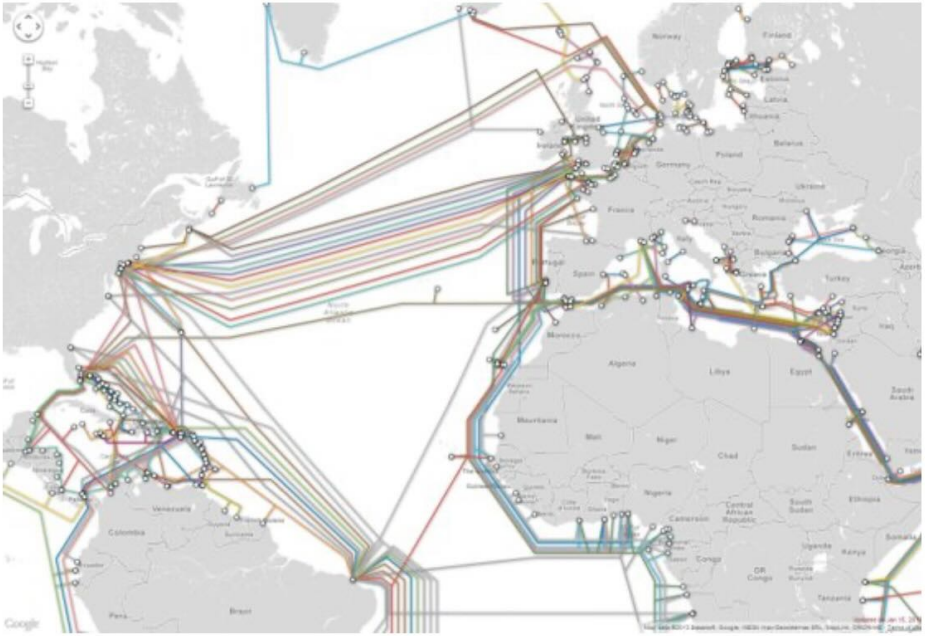


Across the Oceans

Eventually, international Internet communication will need to travel underwater.

The amount of undersea fiber optic cable is quite staggering and is growing yearly.

Undersea fiber optic lines (courtesy TeleGeography)



Section 5 of 8

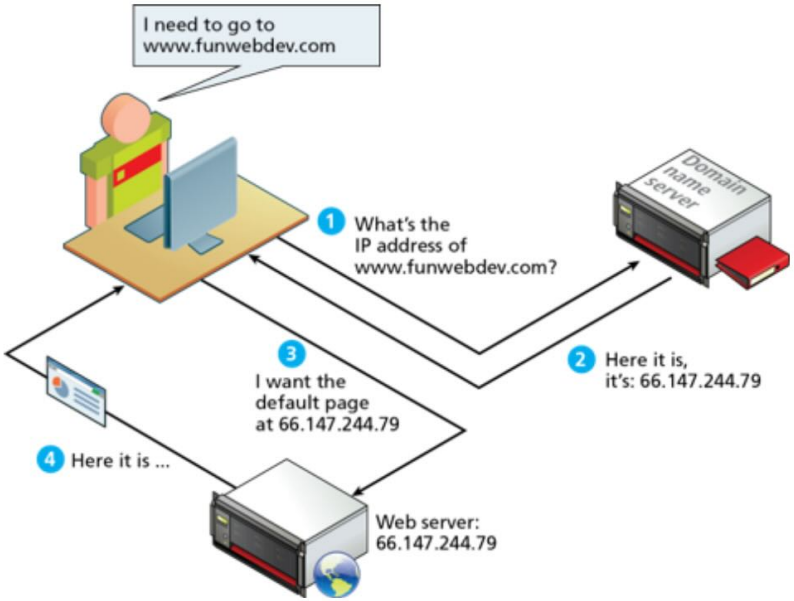
DOMAIN NAME SYSTEM (DNS)

Domain Name System

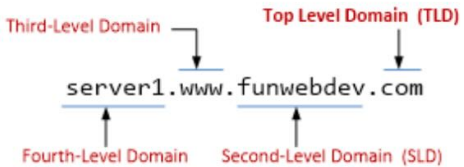
Why do we need it?

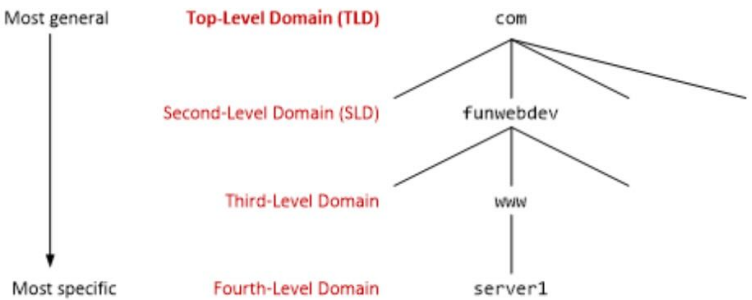
As elegant as IP addresses may be, human beings do not enjoy having to recall long strings of numbers. Instead of IP addresses, we use the **Domain Name System** (**DNS**)

DNS Overview



Domain Levels





Types of TLDs

Generic top-level domains (gTLD)

Country code top-level domain (ccTLD)

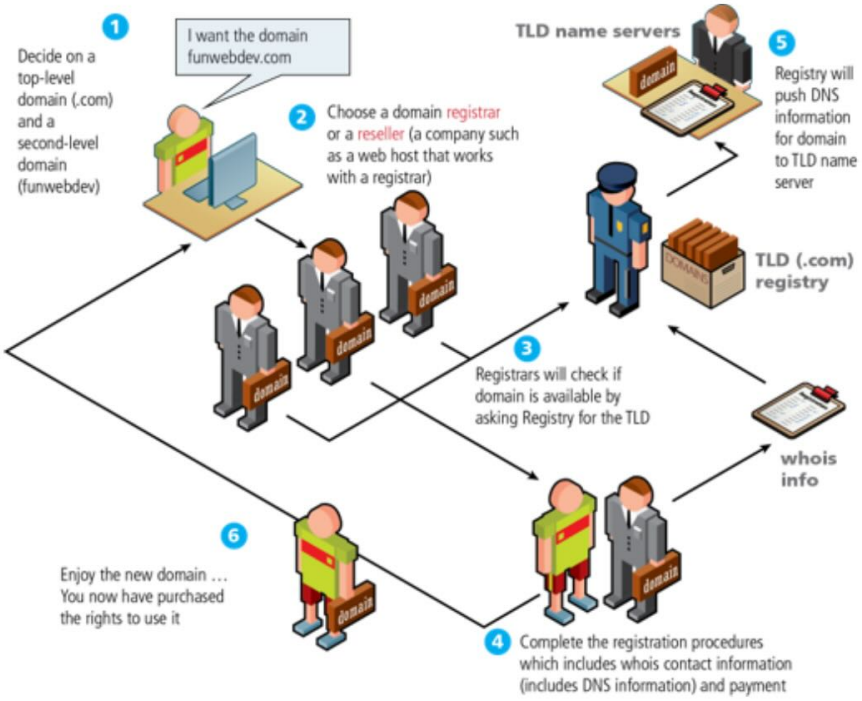
Name Registration

How are domain names assigned?

Special organizations or companies called **domain name registrars** manage the registration of domain names.

These domain name registrars are given permission to do so by the appropriate generic top-level domain (gTLD) registry and/or a country code top-level domain (ccTLD) registry.

Domain name registration process



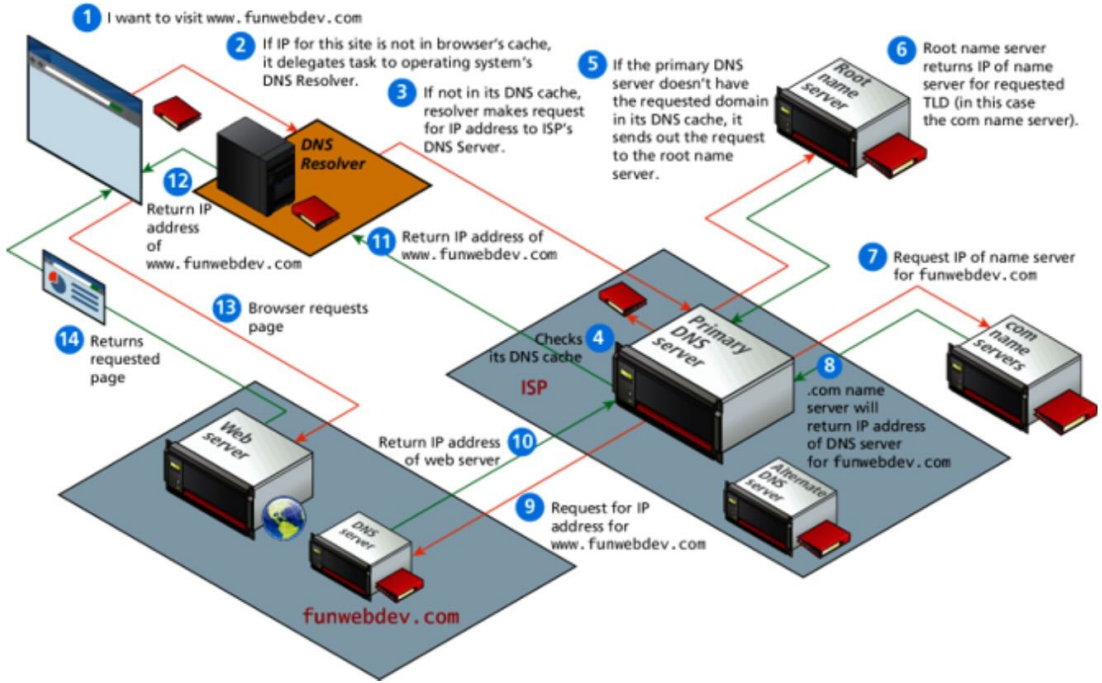
DNS Address Resolution

While domain names are certainly an easier way for users to reference a web site, eventually, your browser needs to know the IP address of the web site in order to request any resources from it.

The Domain Name System provides a mechanism for software to discover this numeric IP address.

This process is referred to here as **address resolution**.

Domain name address resolution process



Section 6 of 8

UNIFORM RESOURCE LOCATORS (URL)

URL Components

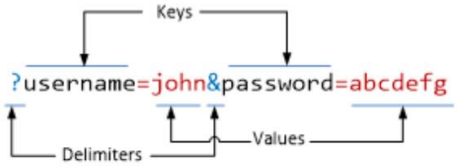
In order to allow clients to request particular resources from the server, a naming mechanism is required so that the client knows how to ask the server for the file.

For the web that naming mechanism is the **Uniform Resource Locator** (**URL**).

Query String

Query strings will be covered in depth when we learn more about HTML forms and server-side programming.

They are the way of passing information such as user form input from the client to the server. In URL's they are encoded as key-value pairs delimited by “&” symbols and preceded by the “?” symbol.



Section 7 of 8

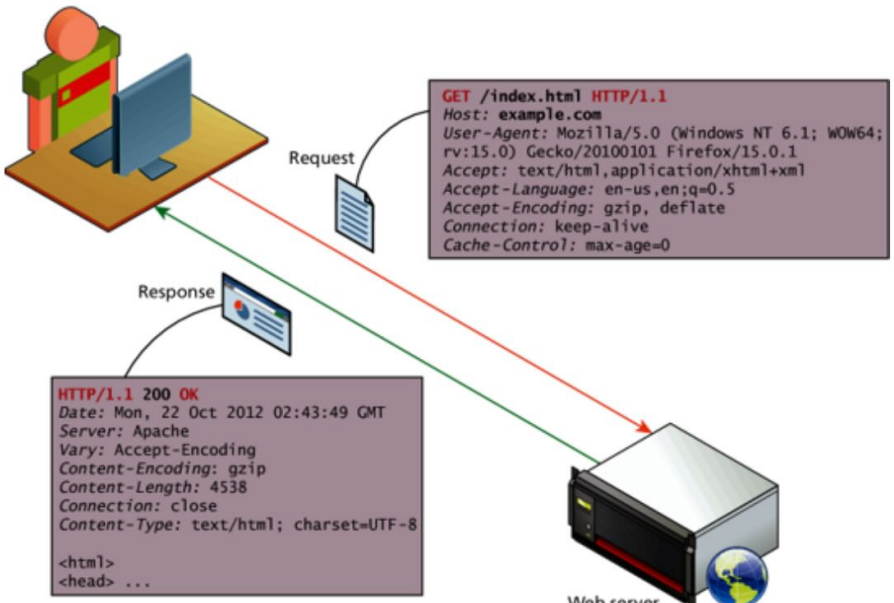
HYPERTEXT TRANSFER PROTOCOL (HTTP)

HTTP

The HTTP protocol establishes a TCP connection on port 80 (by default).

The server waits for the request, and then responds with a response code, headers and an optional message (which can include files).

HTTP


Request
GET /index.html HTTP/1.1
Host: example.com
User-Agent: Mozilla/5.0 (Windows NT 6.1; WOW64; rv:15.0) Gecko/20100101 Firefox/15.0.1
Accept: text/html,application/xhtml+xml
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip, deflate
Connection: keep-alive
Cache-Control: max-age=0
Response
HTTP/1.1 200 OK
Date: Mon, 22 Oct 2012 02:43:49 GMT
Server: Apache
Vary: Accept-Encoding
Content-Encoding: gzip
Content-Length: 4538
Connection: close
Content-Type: text/html; charset=UTF-8
<html>
<head> ...
Web server

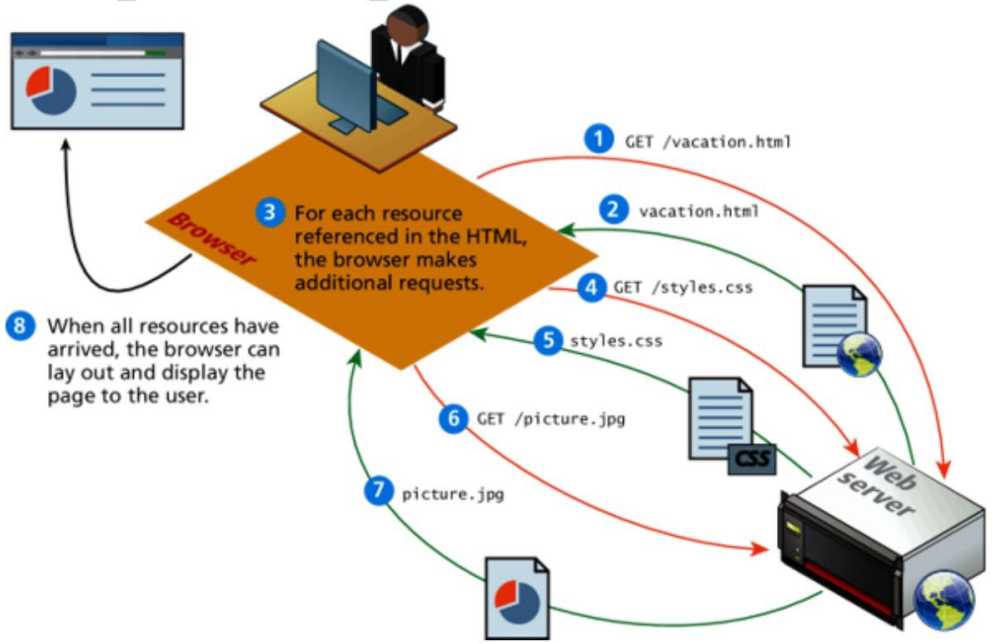
Web Requests

While we as web users might be tempted to think of an entire page being returned in a single HTTP response, this is not in fact what happens.

In reality the experience of seeing a single web page is facilitated by the client's browser which requests the initial HTML page, then parses the returned HTML to find all the resources referenced from within it, like images, style sheets and scripts.

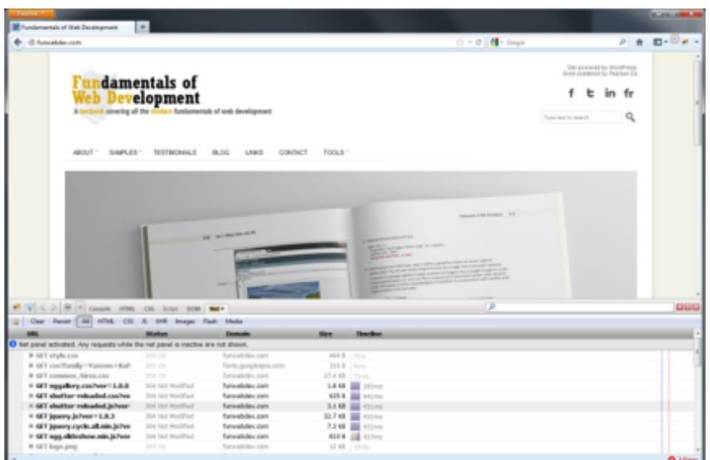
Only when all the files have been retrieved is the page fully loaded for the user

Browser parsing HTML and making subsequent requests



Browser Tools for HTTP

Modern browsers provide the developer with tools that can help us understand the HTTP traffic for a given page.



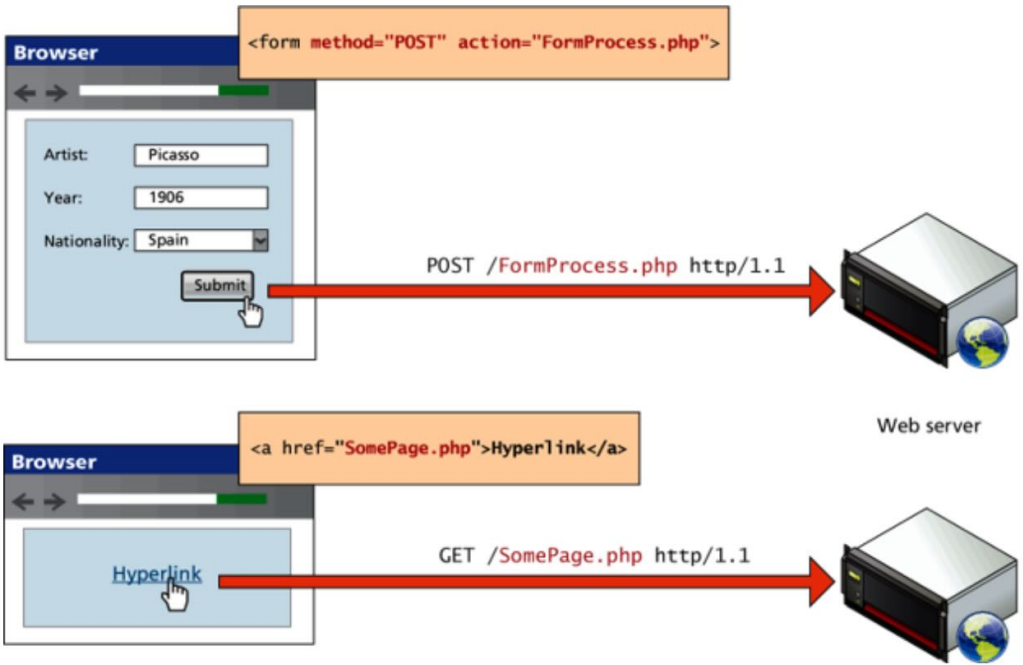
HTTP Request Methods

The HTTP protocol defines several different types of requests, each with a different intent and characteristics.

The most common requests are the GET and POST request, along with the HEAD request.

Other requests, such as PUT, DELETE, CONNECT, TRACE and OPTIONS are seldom used, and are not covered here.

GET versus POST requests



Section 8 of 8

WEB SERVERS

Web Servers

A **web server** is, at a fundamental level, nothing more than a computer that responds to HTTP requests.

Web Stack

Regardless of the physical characteristics of the server, one must choose an application stack to run a website.

This stack will include an operating system, web server software, a database and a scripting language to process dynamic requests.

LAMP Software Stack

Throughout this textbook we will rely on the **LAMP software stack**, which refers to the Linux operating system, Apache web server, MySQL database, and PHP scripting language

WISA software stack

Many corporations, for instance, make use of the Microsoft **WISA software stack**, which refers to Windows operating system, IIS web server, SQL Server database, and the ASP.NET server-side development technologies.