Fundamentals of Web Development

Third Edition by Randy Connolly and Ricardo Hoar

Chapter 1 - a

Introduction to Web Development

Pearson

Technical pre-requisites

1. Using Winrar or Winzip to **zip** and **unzip** multiple files or multiple folders.

2. Navigating a file system from a console:

   (cd ./dir , cd .. , cd ../dir …)

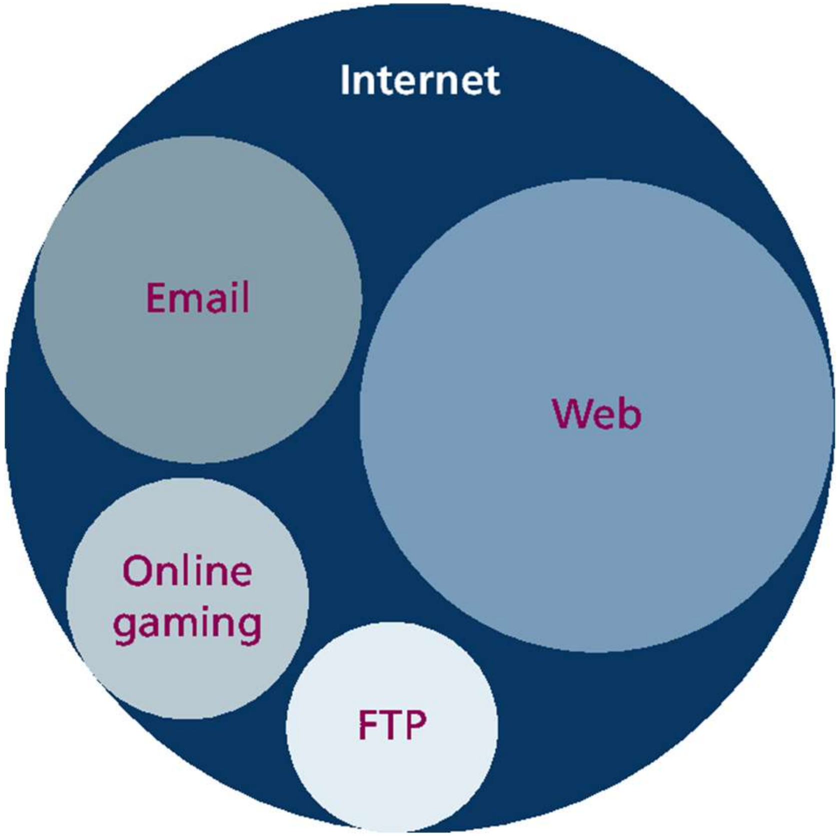
3. Programming:
   a. Conditionals (if..then..else)
   b. Loops (for, for each, while)
   c. Exception handling (try … catch)
   d. Objects: someObject.myFunction()

In this chapter you will learn . . .

- **The history of the Internet and World Wide Web**
- **Fundamental concepts that form the foundation of the Internet**

A Short History of the Internet

- Know the difference between "Internet" and "WWW"
- While this course is focused on the web, part of this chapter is also devoted to a broad understanding of that larger circle labeled the "Internet."
- **Protocols** define different kinds of interactions/services on the Internet

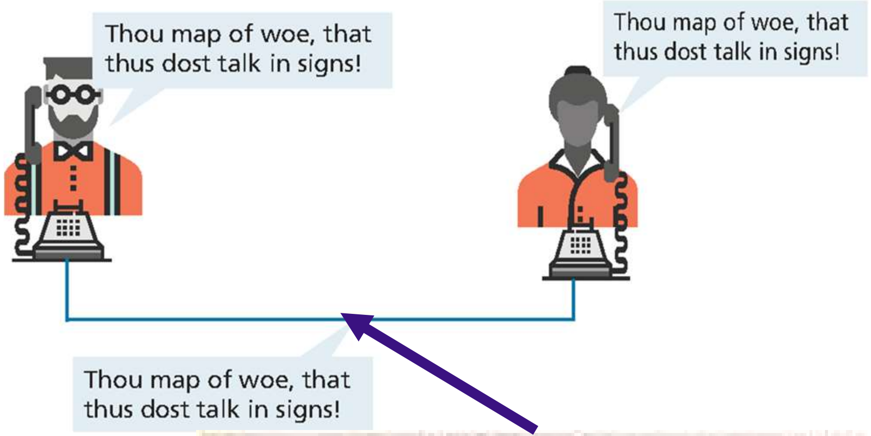


Pearson

Circuit Switched Networks

Circuit switching in early networking

- In the past, telephone calls were routed through operators who **physically connected** the caller and the receiver by connecting a wire to a switchboard to complete a circuit.
- Inefficient use of **bandwidth**
- Difficult to scale

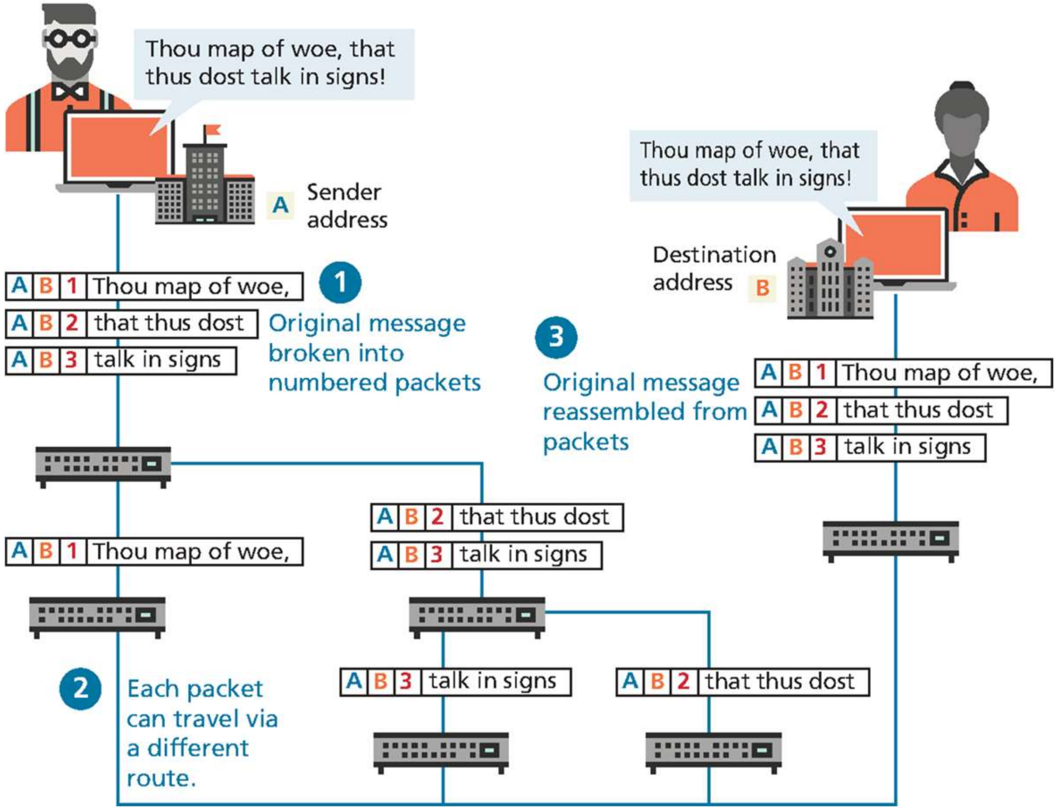

Thou map of woe, that thus dost talk in signs!
Thou map of woe, that thus dost talk in signs!
Thou map of woe, that thus dost talk in signs!

Packet Switched Networks

Packet switching came later, does not require a continuous connection

- 1960s ARPANET
- 1974 X.25
- 1979 USENET
- 1981 TCP/IP was introduced to unify disparate networks

On January 1, 1983, TCP/IP was adopted across all of ARPANET



The Birth of the Web

Pr. Tim Berners-Lee (Turing Award 2016) publishes the main features of the web we know today on 1992.

- A Uniform Resource Locator (**URL**) to uniquely identify a resource on the WWW.
- The Hypertext Transfer Protocol (**HTTP**) to describe how requests and responses operate
- A software program (**web server software**) that can respond to HTTP requests.
- Hypertext Markup Language (**HTML**) to publish documents.
- A program (**a browser**) that can make HTTP requests to URLs and that can display the HTML it receives.

Web Applications in Comparison to Desktop Applications (1 of 2)

Advantages

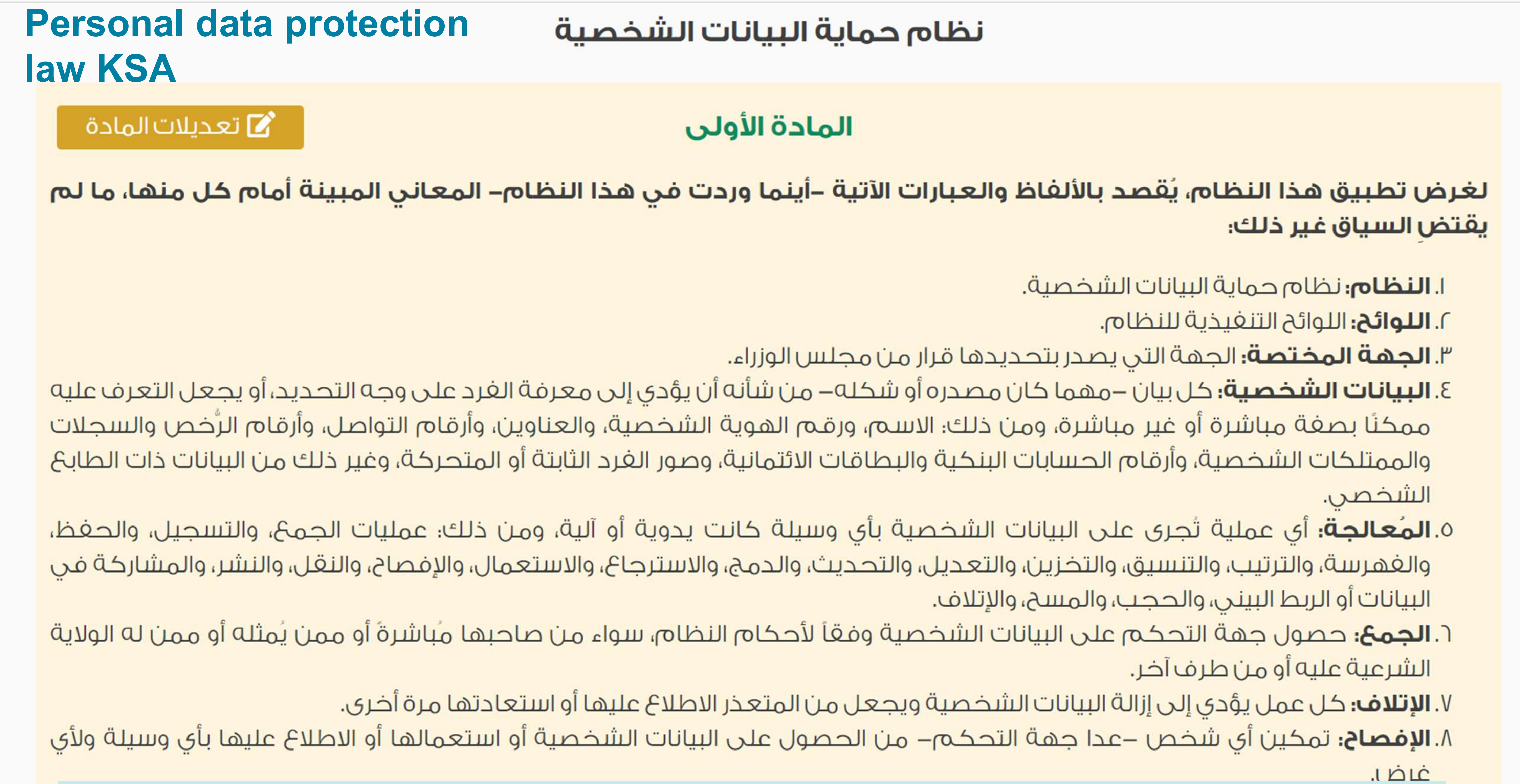
- They can be accessed from any Internet-enabled computer.
- They can be used with different operating systems and browser applications.
- It is easier to roll out program updates since only software on the server needs to be updated as opposed to every computer in the organization using the software.
- They have a centralized storage on the server, which means fewer security concerns about local storage (which is important for sensitive information such as healthcare data).

Web Applications in Comparison to Desktop Applications (2 of 2)

Disadvantages

- Requirement to have an active Internet connection
- Security concerns about sensitive private data being transmitted over the Internet.
- Concerns over the storage, licensing, and use of uploaded data.
- Problems with certain websites not having an identical appearance across all browsers.
- Restrictions on software from being installed and hardware from being accessed (like Adobe Flash on iOS).
- additional plugins might interfere with JavaScript, cookies, or advertisements.

Personal data protection law KSA



نظام حماية البيانات الشخصية

المادة الأولى

لغرض تطبيق هذا النظام، يُقصد بالألفاظ والعبارات الآتية -أينما وردت في هذا النظام- المعاني المبينة أمام كل منها، ما لم يقتضِ السياق غير ذلك:

١. **النظام:** نظام حماية البيانات الشخصية.

٢. **اللوائح:** اللوائح التنفيذية للنظام.

٣. **الجهة المختصة:** الجهة التي يصدر بتحديدها قرار من مجلس الوزراء.

٤. **البيانات الشخصية:** كل بيان -مهما كان مصدره أو شكله- من شأنه أن يؤدي إلى معرفة الفرد على وجه التحديد، أو يجعل التعرف عليه ممكنًا بصفة مباشرة أو غير مباشرة، ومن ذلك: الاسم، ورقم الهوية الشخصية، والعناوين، وأرقام التواصل، وأرقام الرُّخص والسجلات والممتلكات الشخصية، وأرقام الحسابات البنكية والبطاقات الائتمانية، وصور الفرد الثابتة أو المتحركة، وغير ذلك من البيانات ذات الطابع الشخصي.

٥. **المُعالجة:** أي عملية تُجرى على البيانات الشخصية بأي وسيلة كانت يدوية أو آلية، ومن ذلك: عمليات الجمع، والتسجيل، والحفظ، والفهرسة، والترتيب، والتنسيق، والتخزين، والتعديل، والتحديث، والدمج، والاسترجاع، والاستعمال، والإفصاح، والنقل، والنشر، والمشاركة في البيانات أو الربط البيني، والحجب، والمسح، والإتلاف.

٦. **الجمع:** حصول جهة التحكم على البيانات الشخصية وفقاً لأحكام النظام، سواء من صاحبها مُباشرةً أو ممن يُمثله أو ممن له الولاية الشرعية عليه أو من طرف آخر.

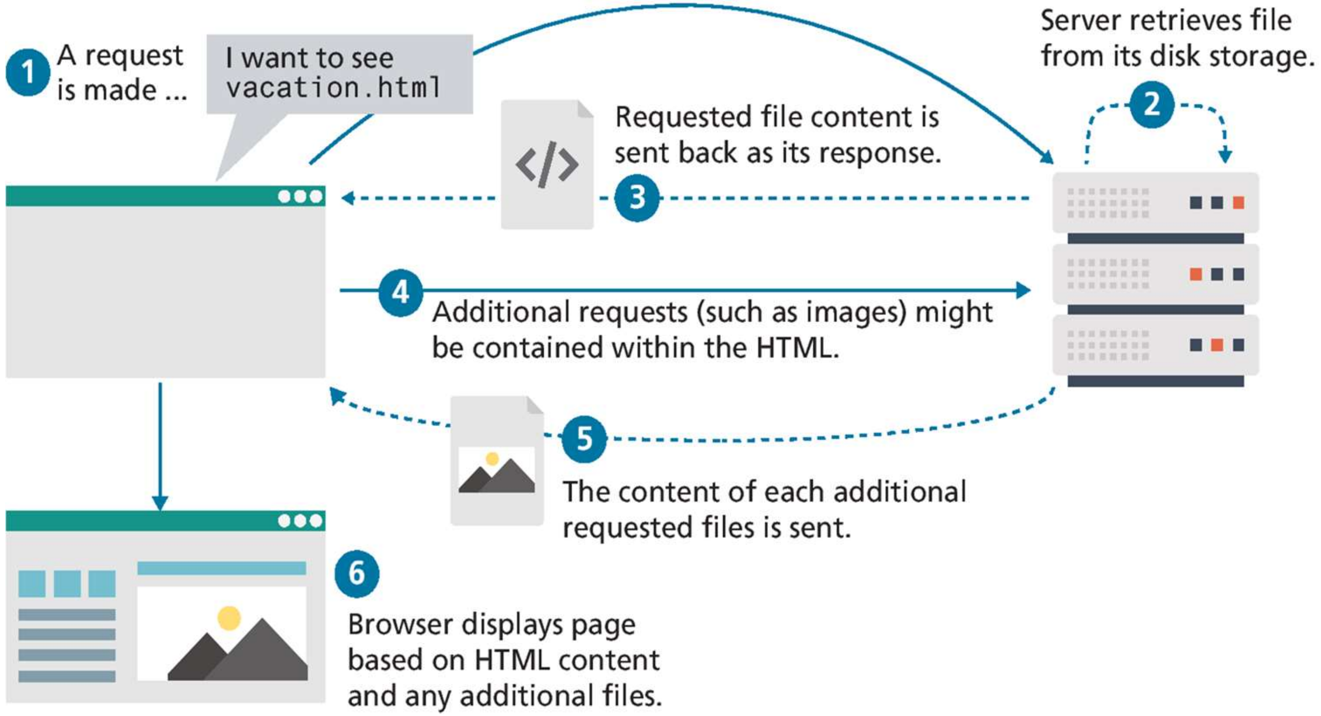
٧. **الإتلاف:** كل عمل يؤدي إلى إزالة البيانات الشخصية ويجعل من المتعذر الاطلاع عليها أو استعادتها مرة أخرى.

٨. **الإفصاح:** تمكين أي شخص -عدا جهة التحكم- من الحصول على البيانات الشخصية أو استعمالها أو الاطلاع عليها بأي وسيلة ولأي غرض.

https://laws.boe.gov.sa/boelaws/laws/lawdetails/b7cfae89-828e-4994-b167-adaa00e37188/1

From Static to Dynamic

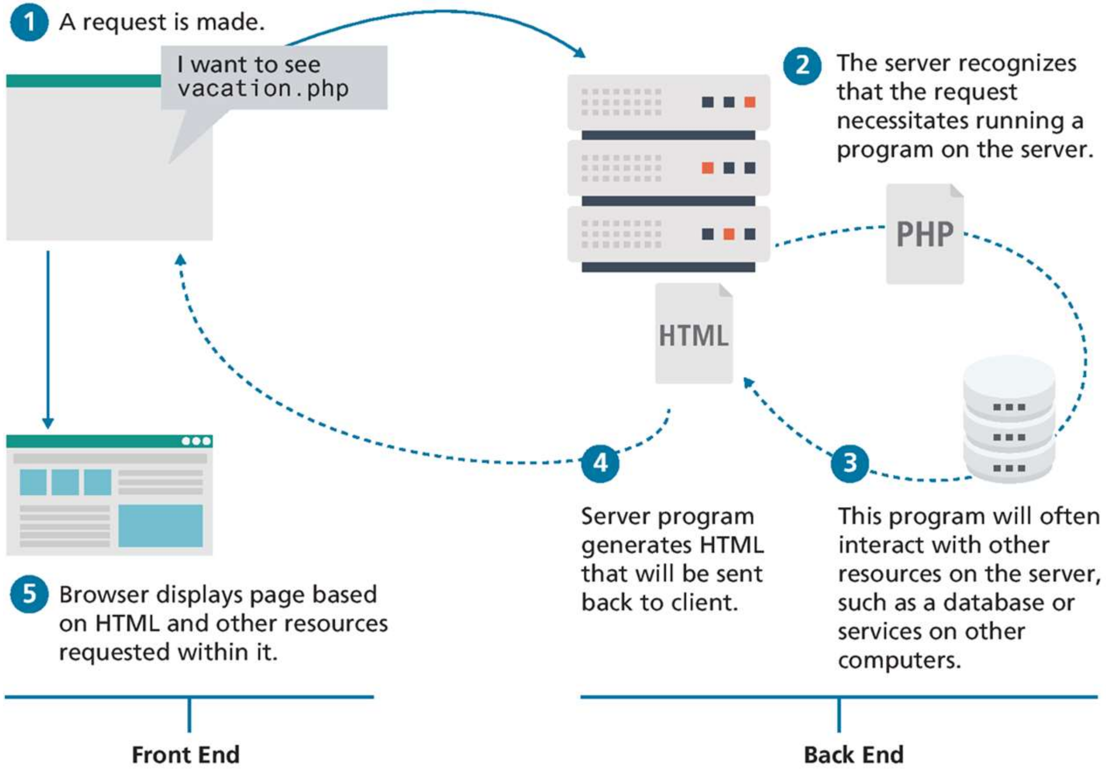
In the earliest days of the web, users could read the pages of a **static website**


1
A request is made ...
I want to see vacation.html
Server retrieves file from its disk storage.
2
Requested file content is sent back as its response.
3
4
Additional requests (such as images) might be contained within the HTML.
5
The content of each additional requested files is sent.
6
Browser displays page based on HTML content and any additional files.

Pearson

From Static to Dynamic ii

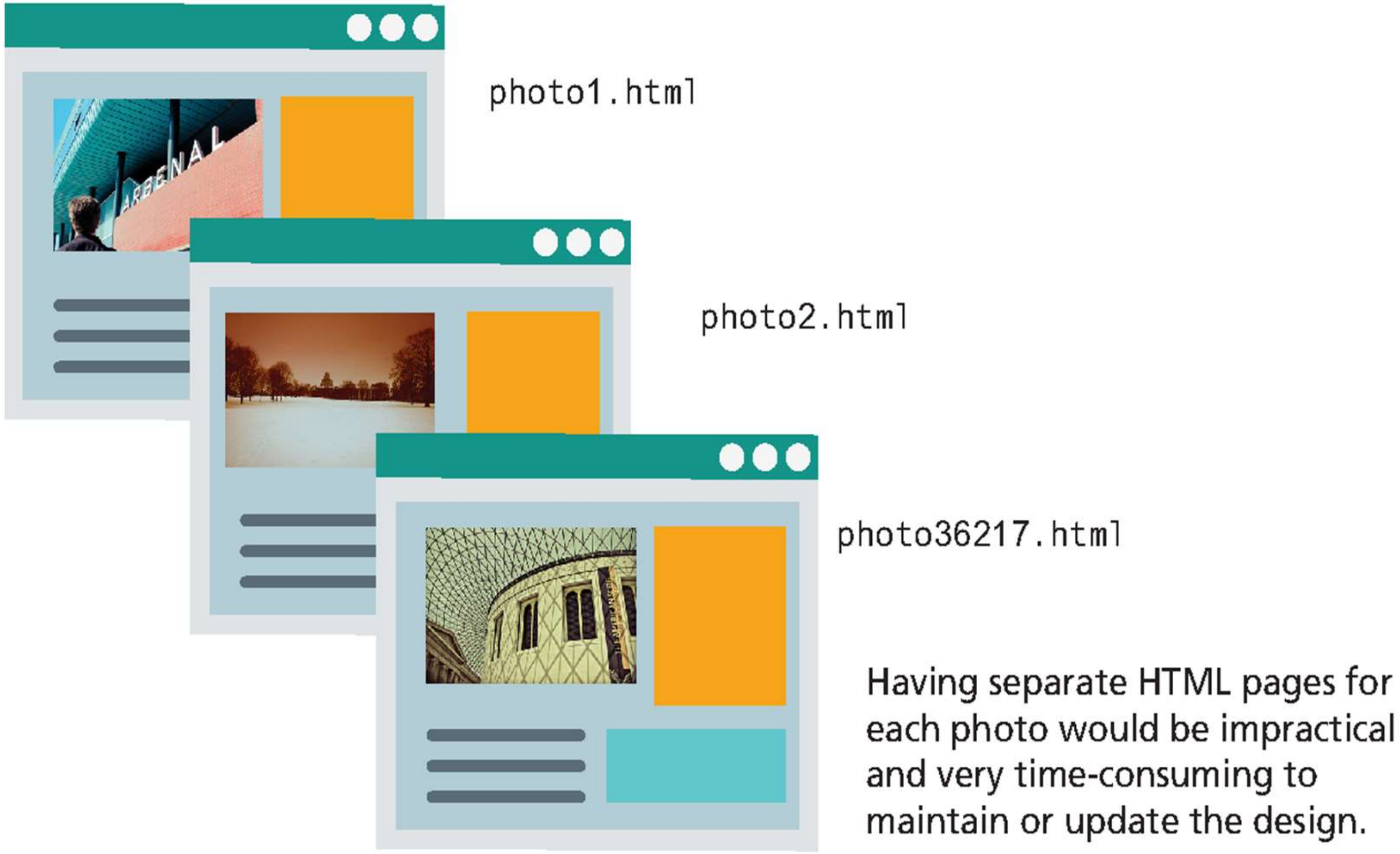
Later, programs running on web servers let websites generate content dynamically. This type of website is called a **dynamic server-side website**


1 A request is made.
I want to see vacation.php
2 The server recognizes that the request necessitates running a program on the server.
PHP
HTML
3 This program will often interact with other resources on the server, such as a database or services on other computers.
4 Server program generates HTML that will be sent back to client.
5 Browser displays page based on HTML and other resources requested within it.
Front End
Back End

Web 2.0

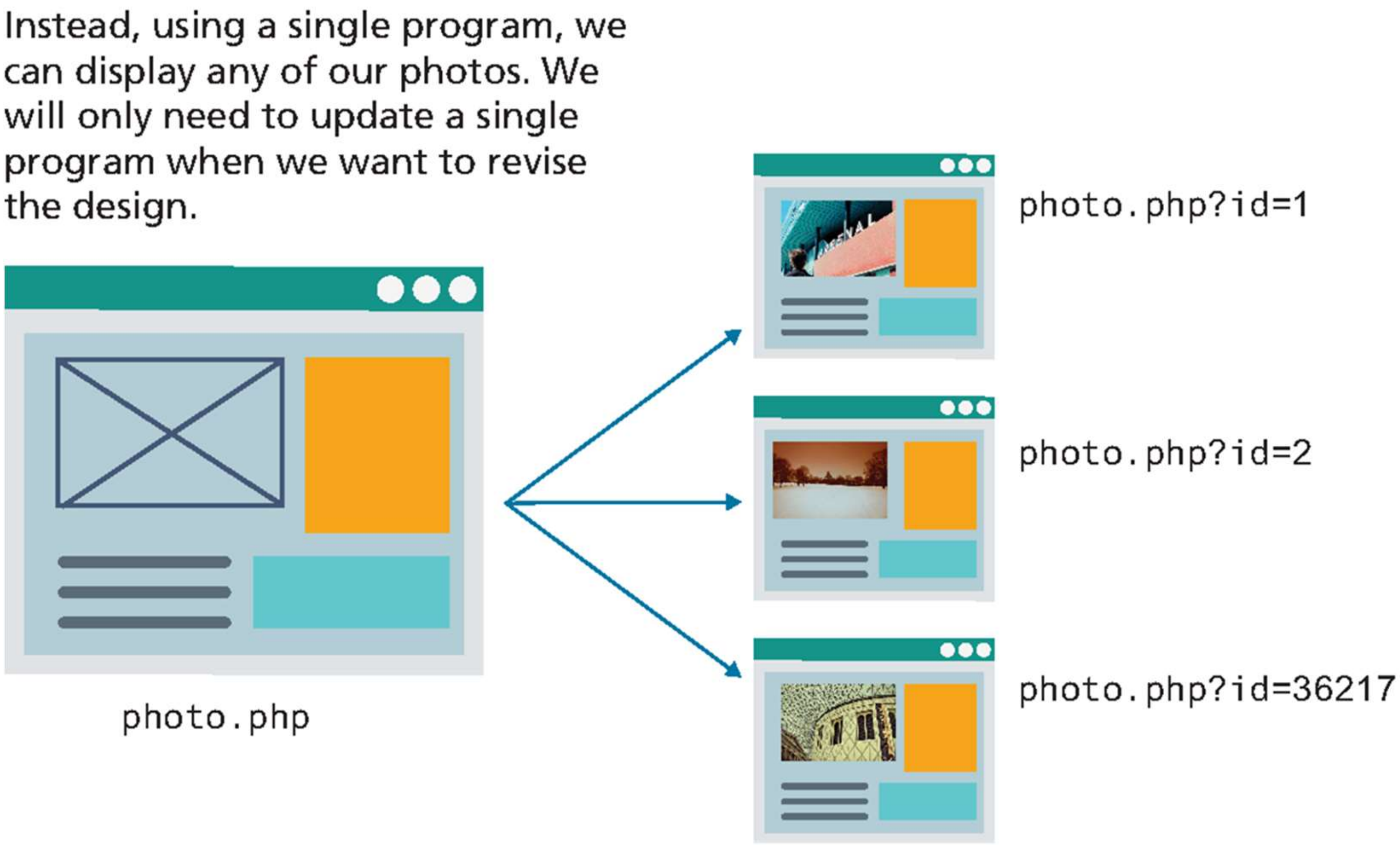
- **Web 2.0** refers to an interactive experience where users can contribute and consume web content, thus creating a more user-driven web experience.
- For software developers, Web 2.0 also refers to a change in the paradigm. Programming logic, which previously existed only on the server, began to migrate more and more to the browser, which requires learning **JavaScript**

Why are programs needed? (1 of 2)



Having separate HTML pages for each photo would be impractical and very time-consuming to maintain or update the design.

Why are programs needed? (2 of 2)


Instead, using a single program, we can display any of our photos. We will only need to update a single program when we want to revise the design.

The Client-Server Model

Client machines are the desktops, laptops, smart phones, and tablets you see everywhere.

Broad range of specifications regarding

- operating system,
- processing speed,
- screen size,
- available memory, and
- storage

Server machines host web applications, store user and program data, and perform security authorization tasks

Powerful machines to handle high traffic and bandwidth.

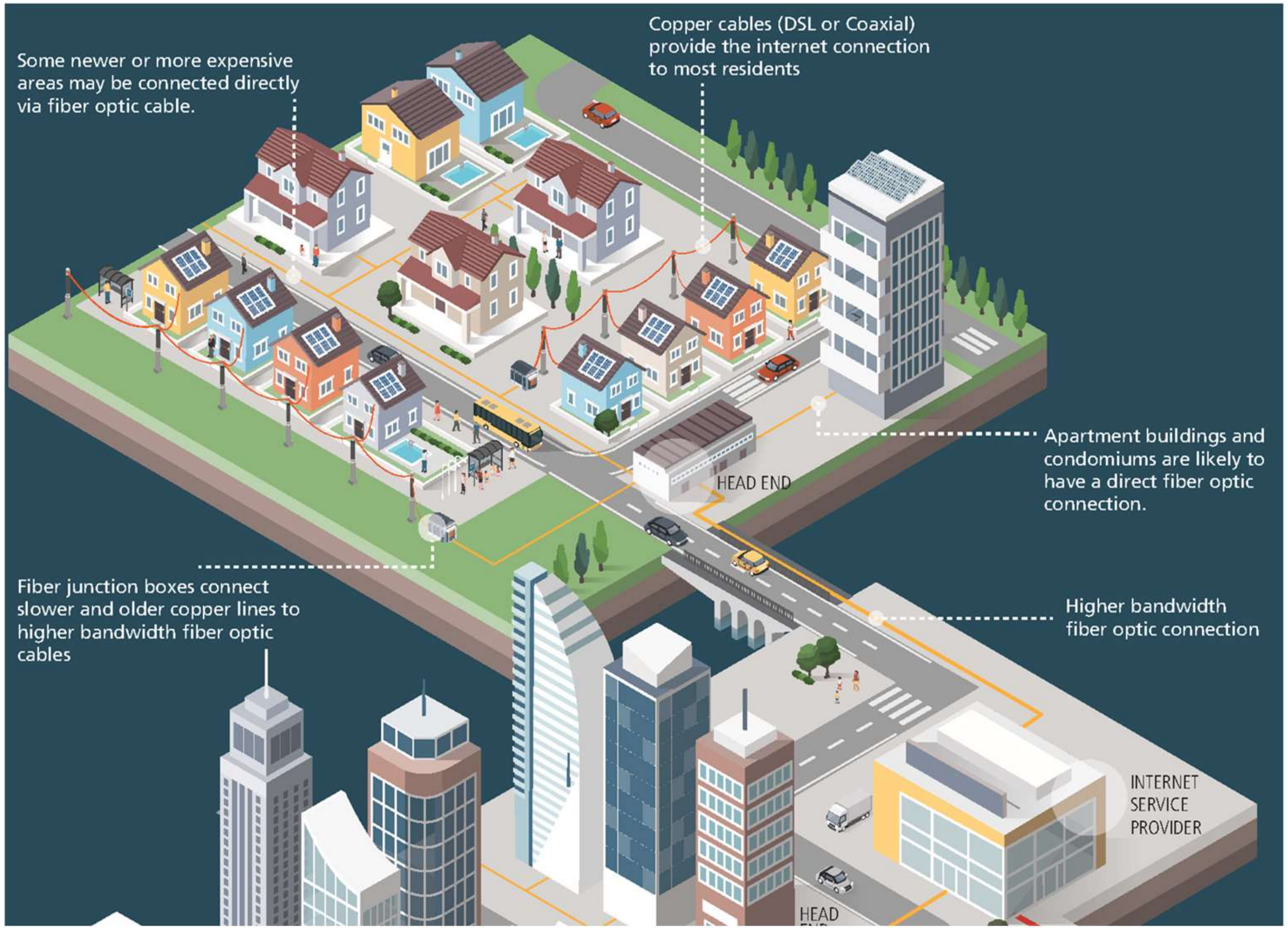
The essential characteristic of a server is that it is listening for requests, and upon getting one, responds with a message.

Where Is the Internet?

It is quite common for the Internet to be visually represented as a cloud

Actually implemented via millions of miles of copper (نحاس) wires and fiber optic cables connecting millions of server computers and other networked devices!

From the Home to the Ocean’s Edge



Some newer or more expensive areas may be connected directly via fiber optic cable.

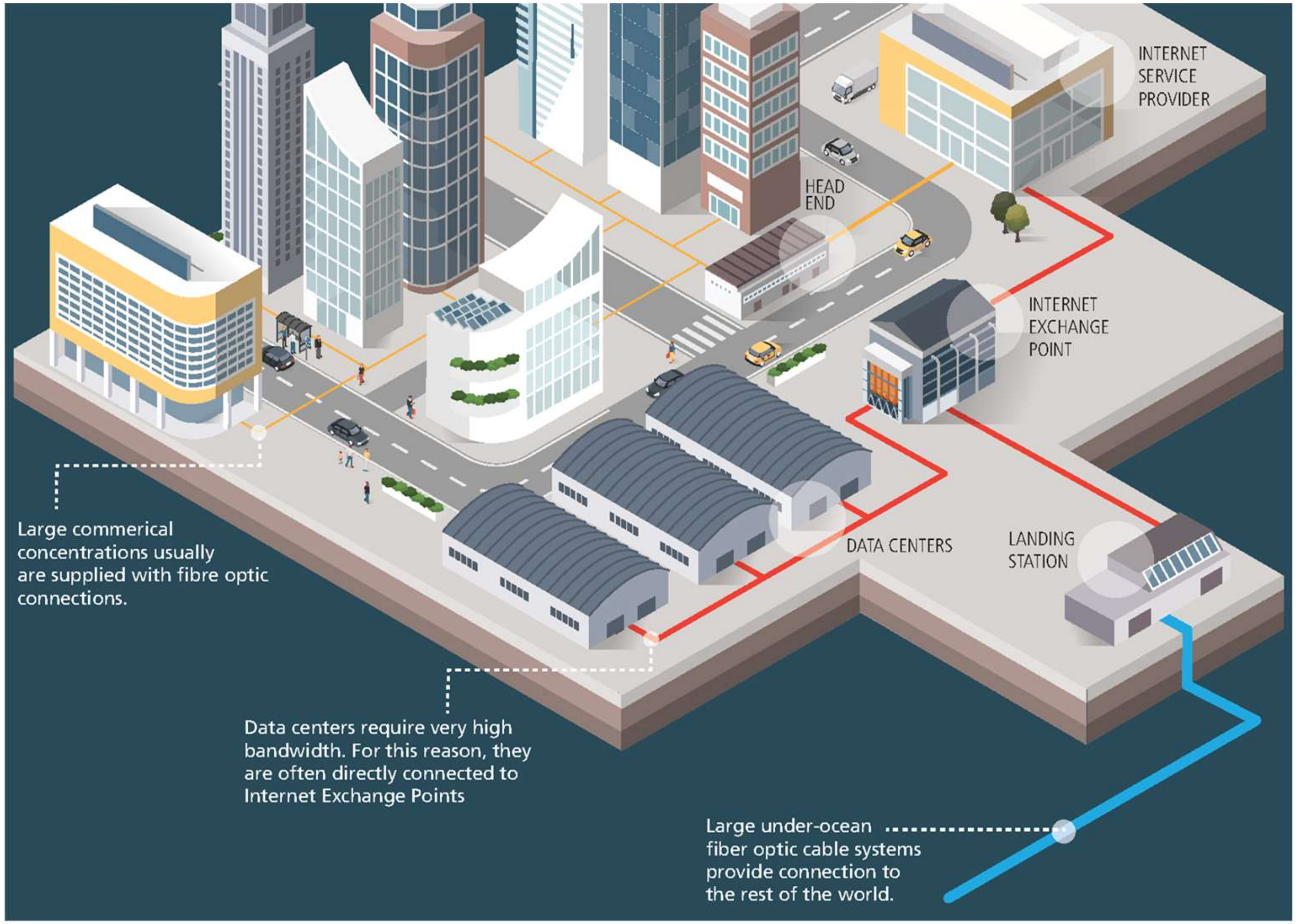
Copper cables (DSL or Coaxial) provide the internet connection to most residents

Apartment buildings and condominiums are likely to have a direct fiber optic connection.

Fiber junction boxes connect slower and older copper lines to higher bandwidth fiber optic cables

Higher bandwidth fiber optic connection

From the Home to the Ocean’s Edge (2)

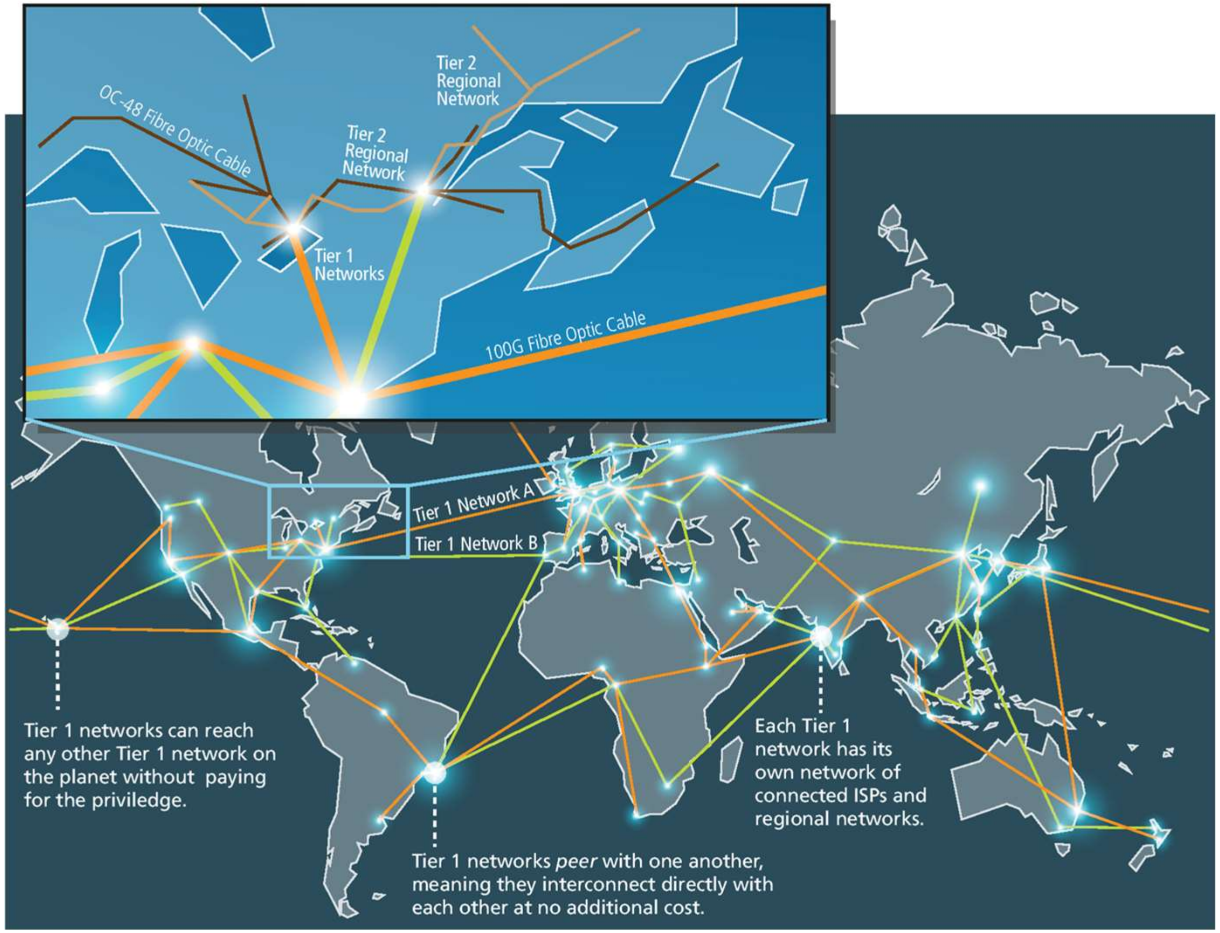


Large commerical concentrations usually are supplied with fibre optic connections.

Data centers require very high bandwidth. For this reason, they are often directly connected to Internet Exchange Points

Large under-ocean fiber optic cable systems provide connection to the rest of the world.

How the Internet Is Organized Today


Tier 2 Regional Network
OC-48 Fibre Optic Cable
Tier 2 Regional Network
Tier 1 Networks
100G Fibre Optic Cable
Tier 1 Network A
Tier 1 Network B
Tier 1 networks can reach any other Tier 1 network on the planet without paying for the priviledge.
Each Tier 1 network has its own network of connected ISPs and regional networks.
Tier 1 networks *peer* with one another, meaning they interconnect directly with each other at no additional cost.

When someone talks about the “Internet Backbone” they are talking about Tier 1 networks.

Tier 1 Networks make use of very fast fiber optic cable.

Regional networks have traditionally used less speedy infrastructure, though this is rapidly changing as prices of optical hardware decreases.

https://youtu.be/Gg1aFmsKQgk?t=188

Fundamentals of Web Development

Third Edition by Randy Connolly and Ricardo Hoar

Chapter 1 - b

How the Web Works

Pearson

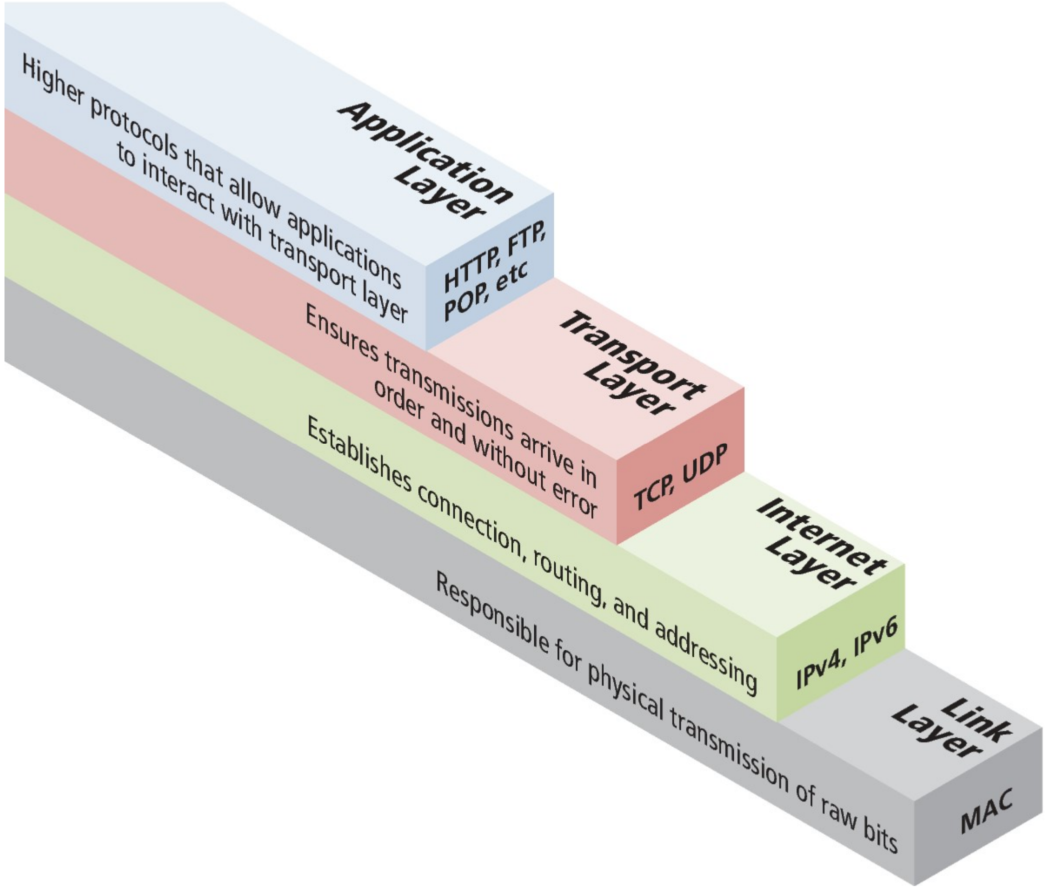
In this chapter you will learn . . .

- The fundamental protocols that make the web possible
- How the domain name system works
- What HTTP is
- How browsers and servers work to exchange and interpret HTML

Internet protocols

- A **protocol** is a set of rules that partners use when they communicate.
- TCP/IP is an essential internet protocol!
- These protocols have been implemented in every operating system and make fast web development possible. If web developers had to keep track of packet routing, transmission details, domain resolution, checksums, and more, it would be hard to build websites.

A Layered Architecture


Application Layer
Higher protocols that allow applications to interact with transport layer
HTTP, FTP, POP, etc
Transport Layer
Ensures transmissions arrive in order and without error
TCP, UDP
Internet Layer
Establishes connection, routing, and addressing
IPv4, IPv6
Link Layer
Responsible for physical transmission of raw bits
MAC

- The TCP/IP Internet protocols were originally abstracted as a four-layer stack
- Later abstractions subdivide it further into five or seven layers
- Since we focus on the top layer, we will use the earliest and simplest **four-layer network model**.

Link Layer

- The **link layer** is the lowest layer, responsible for both the physical transmission of data across media and establishing logical links.
- It handles issues like packet creation, transmission, reception, error detection, collisions, line sharing, and more.
- One term that is sometimes used in the Internet context is that of **MAC** (media access control) **addresses**.

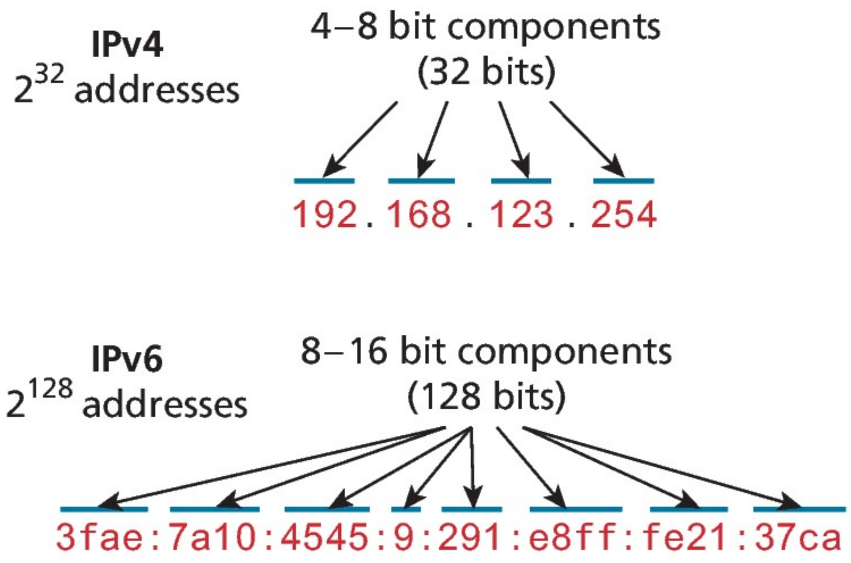
Internet Layer

- The **Internet layer** (sometimes also called the IP Layer) routes packets between communication partners across networks.
- It provides “best effort” communication. It sends out a message to its destination but expects no reply and provides no guarantee the message will arrive intact, or at all.
- The Internet uses the **Internet Protocol (IP) addresses**, which are numeric codes that uniquely identify destinations on the Internet.
- Every device connected to the Internet has such an **IP address**.

IP Addresses (cont)

There are two types of IP addresses: IPv4 and IPv6.

- In **IPv4**, four 8-bit integers separated by . encode the address.
- **IPv6** uses eight 16-bit integers (*each 4-bit integer is displayed in hexadecimal for readability*) and has over a billion billion times the number in IPv4



Port Address Translation (PAT)

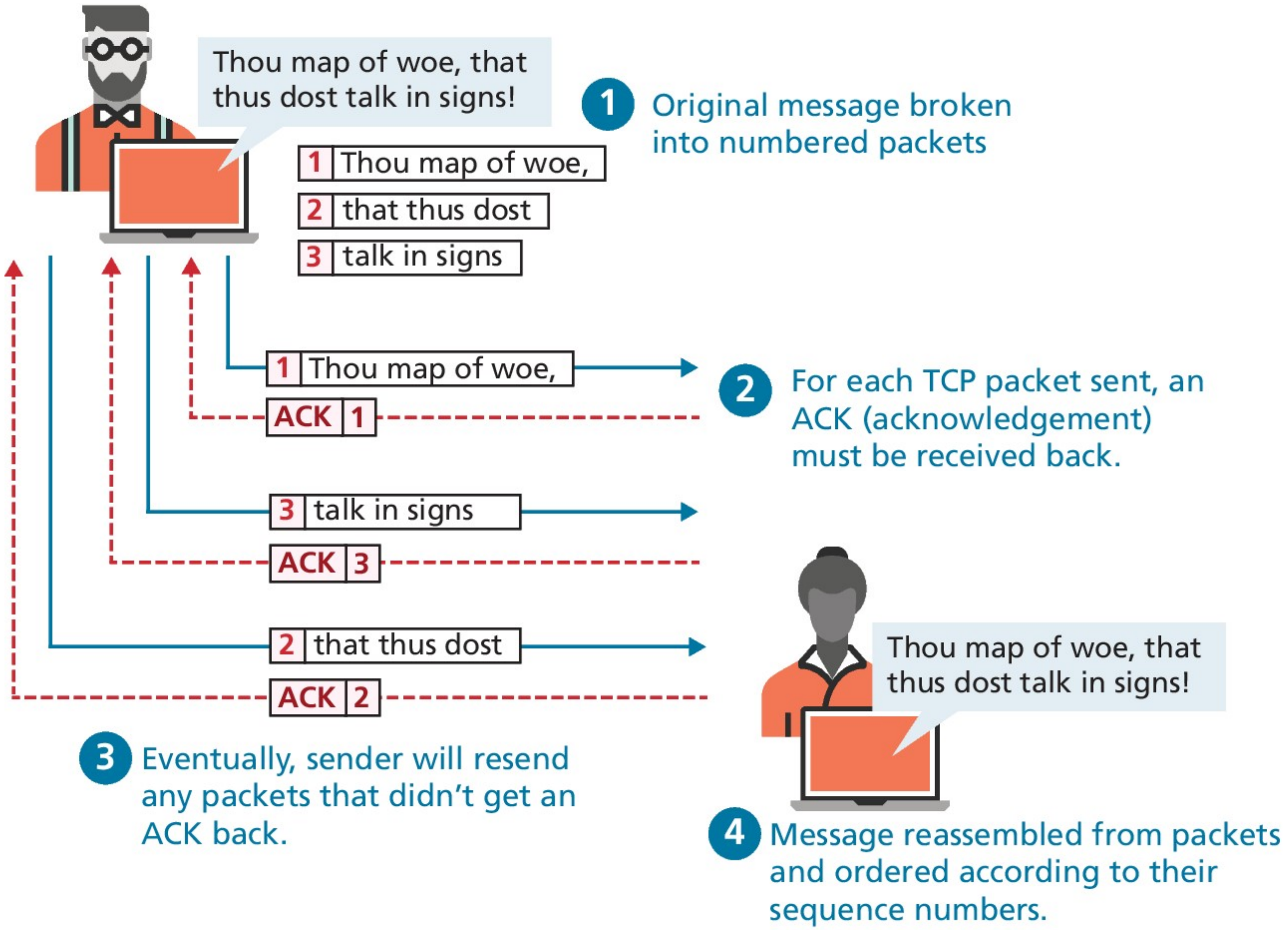
The IPv4 address space was depleted (almost filled) in 2011, but the number of computers connected to the Internet continued to grow.

- **Port Address Translation (PAT)**, allows multiple, unrelated networks to make use of the same IP address
- When you join a wireless network in a coffee shop, home, office or university, it is quite likely you are making use of PAT.
- For future growth, IPv6 will be necessary.

Transport Layer

- The **transport layer** ensures transmissions arrive in order and without error.
- First, the data is broken into **packets** formatted according to the **Transmission Control Protocol (TCP)**.
  - Each data packet has a header that includes a sequence number, so the receiver can put the original message back in order
  - Each packet acknowledges its successful arrival back to the sender (ACK).
  - In the event of a lost packet (since no ACK arrived for that packet the packet will be retransmitted.
- This means you have **a *guarantee*** that messages sent will arrive and will be in order.

Transport Layer (example)



Thou map of woe, that thus dost talk in signs!
1 Original message broken into numbered packets
1 Thou map of woe,
2 that thus dost
3 talk in signs
1 Thou map of woe,
ACK 1
2 For each TCP packet sent, an ACK (acknowledgement) must be received back.
3 talk in signs
ACK 3
2 that thus dost
ACK 2
Thou map of woe, that thus dost talk in signs!
3 Eventually, sender will resend any packets that didn't get an ACK back.
4 Message reassembled from packets and ordered according to their sequence numbers.

User Datagram Protocol (UDP)

PROTIP

Sometimes we do not want guaranteed transmission of packets.

Consider a live multicast of a soccer game, for example. Millions of subscribers may be streaming the game, and the broadcaster can’t afford to track and retransmit every lost packet. A small loss of data in the feed is acceptable, and the customers will still see the game.

An Internet protocol called **User Datagram Protocol (UDP)** is used in these scenarios in lieu of TCP. Other examples of UDP services include Voice Over IP (VoIP), many online games, and Domain Name System (DNS).

Application Layer

- The **application layer** is the level of protocols familiar to most web developers.
- Application layer protocols implement process-to-process communication.
- There are many application layer protocols. A few that are useful to web developers include:

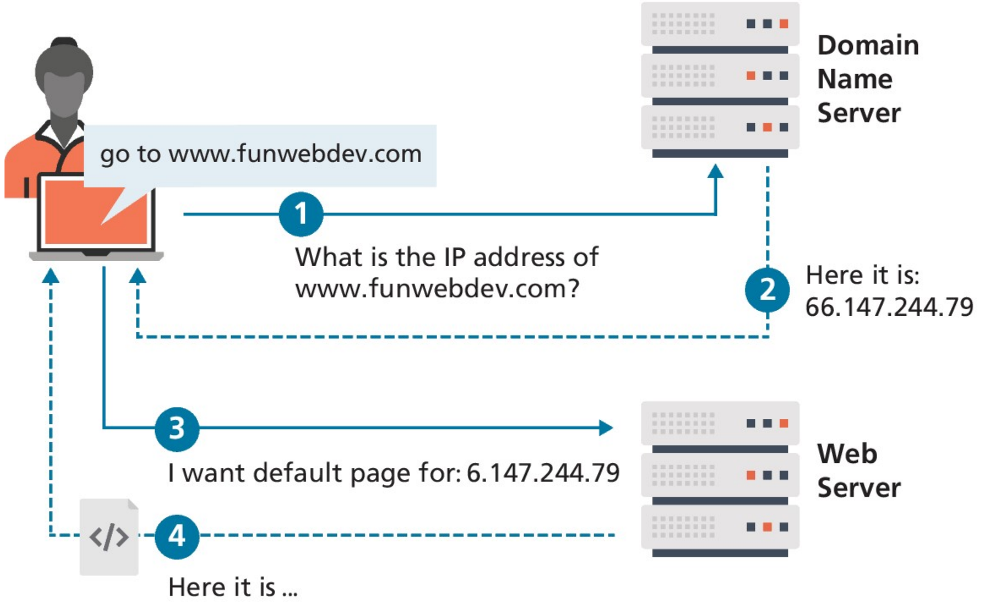
- **HTTP.** The Hypertext Transfer Protocol is used for **web** communication.
- **SSH**. The Secure Shell Protocol allows remote command-line connections to servers.
- **FTP**. The File Transfer Protocol is used for transferring files between computers.
- **POP/IMAP/SMTP.** Email-related protocols for transferring and storing email.
- **DNS.** The Domain Name System protocol used for resolving domain names to IP addresses.

Domain Name System

- As elegant as IP addresses may be, human beings do not enjoy having to remember long strings of numbers.
- Even as far back as the days of ARPANET, researchers assigned **domain names** to IP addresses
- In those early days, the number of Internet hosts was small, so a list of a domains and associated IP addresses could be downloaded as needed as a **hosts** file.
- As the number of computers on the Internet grew, this **hosts** file had to be replaced with a better, more scalable, and distributed system. This system is called the **Domain Name System (DNS)**

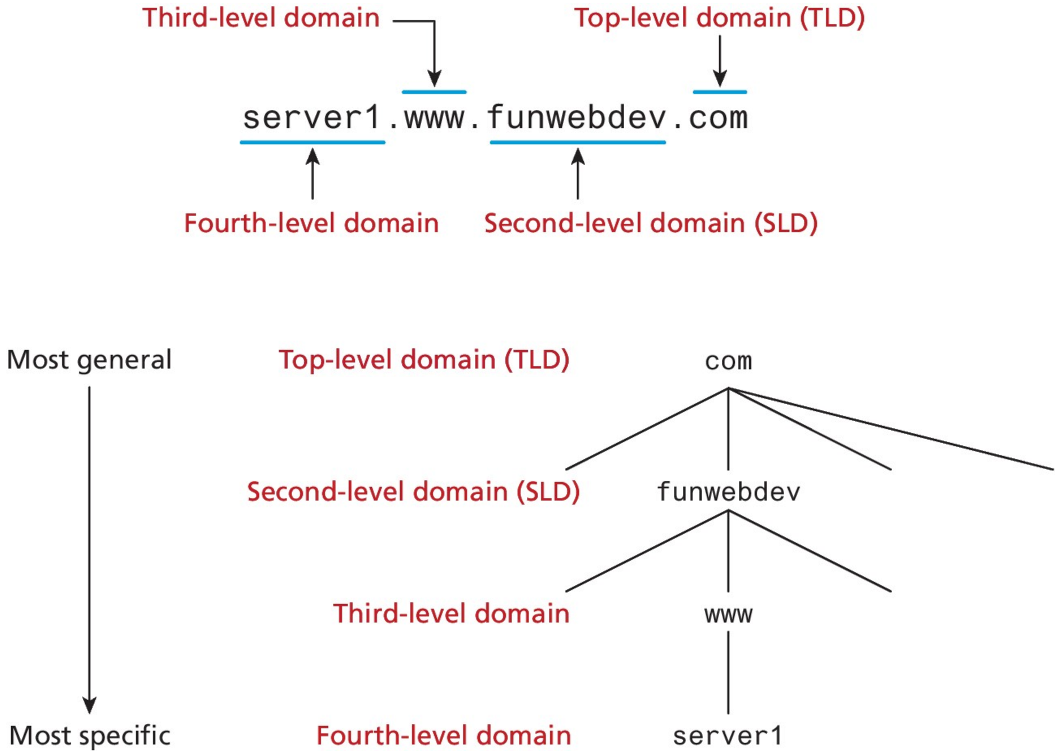
DNS Overview

- The DNS system maps resolves domain names to IP addresses.
- By separating the domain name of a server from its IP address, a site can move to a different host without changing its name.
- Since the entire request-response cycle can take less than a second, it is easy to forget that DNS requests are happening in all.


go to www.funwebdev.com
1
What is the IP address of www.funwebdev.com?
Domain Name Server
2
Here it is: 66.147.244.79
3
I want default page for: 6.147.244.79
Web Server
</>
4
Here it is ...

Name Levels

- A domain name can be broken down into several parts, which describe a hierarchy.
- All domain names have at least a **top-level domain (TLD)** name and a **second-level domain (SLD)** name



Name Levels (Top Level)

The rightmost portion of the domain name is called **the top-level domain**.

TLDs limited to two broad categories, plus a third reserved for other use.

1. **Generic top-level domain (gTLD)**
2. **Country code top-level domain (ccTLD)**
3. **.arpa** (used for reverse DNS lookups)

Generic top-level domain (gTLD)

Generic top-level domains (gTLD) include the famous .com and .org , There are 3 subtypes of gTLD.

- **Unrestricted.** TLDs include **.com**, **.net**, **.org**, and **.info**.
- **Sponsored.** TLDs including **.gov**, **.mil**, **.edu**, and others
- **New.** Starting in June 2012, ICANN invited companies to launch new TLDs in order to provide more choice. Since then over 1000 new TLD have been created including .art, .cash, .cool, .jobs, .tax and so on

Country code top-level domain

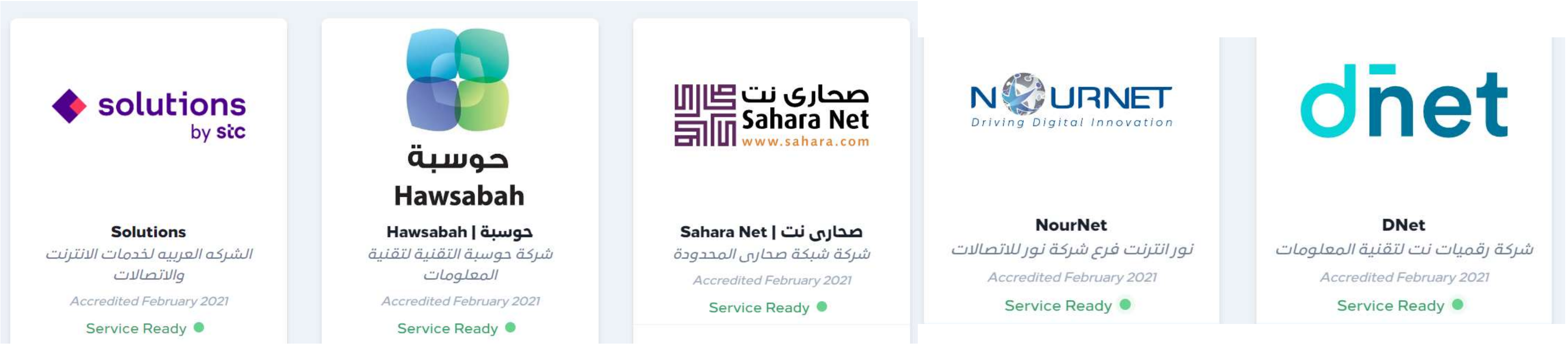
Country code top-level domain (ccTLD) are under the control of the countries which they represent, which is why each is administered differently.

- In the United Kingdom, for example, businesses must register subdomains to **co.uk** rather than second-level domains directly whereas in Canada, **.ca** domains can be obtained by any person, company, or organization living or doing business in Canada.
- **Internationalized top-level domain name (IDN)** allows domains to use non-ascii characters and has been deployed since 2009. There are over 9 million IDN domains

Name Registration

- Q: How then are domain names assigned?
- A: 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.
- The non-profit **Internet Corporation for Assigned Names and Numbers (ICANN)**—oversees the management of top-level domains, accredits registrars, and coordinates other aspects of DNS.
- Examples: domain.com, godaddy, hostgator, bluehost …

Domain Name Registrars in KSA



Solutions
الشركه العربيه لخدمات الانترنت والاتصالات
Accredited February 2021
Service Ready ●

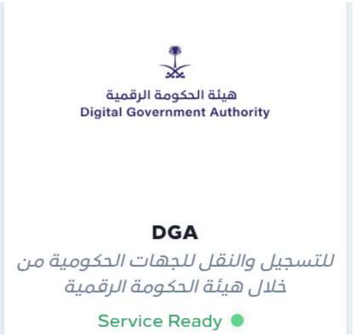
Hawsabah | حوسبة
شركة حوسبة التقنية لتقنية المعلومات
Accredited February 2021
Service Ready ●

Sahara Net | صحارى نت
شركة شبكة صحارى المحدودة
Accredited February 2021
Service Ready ●

NourNet
نور انترنت فرع شركة نور للاتصالات
Accredited February 2021
Service Ready ●

DNet
شركة رقميات نت لتقنية المعلومات
Accredited February 2021
Service Ready ●

Fig. list of registrars in KSA [https://nic.sa/en/registrars/]



DGA
للتسجيل والنقل للجهات الحكومية من خلال هيئة الحكومة الرقمية
Service Ready ●

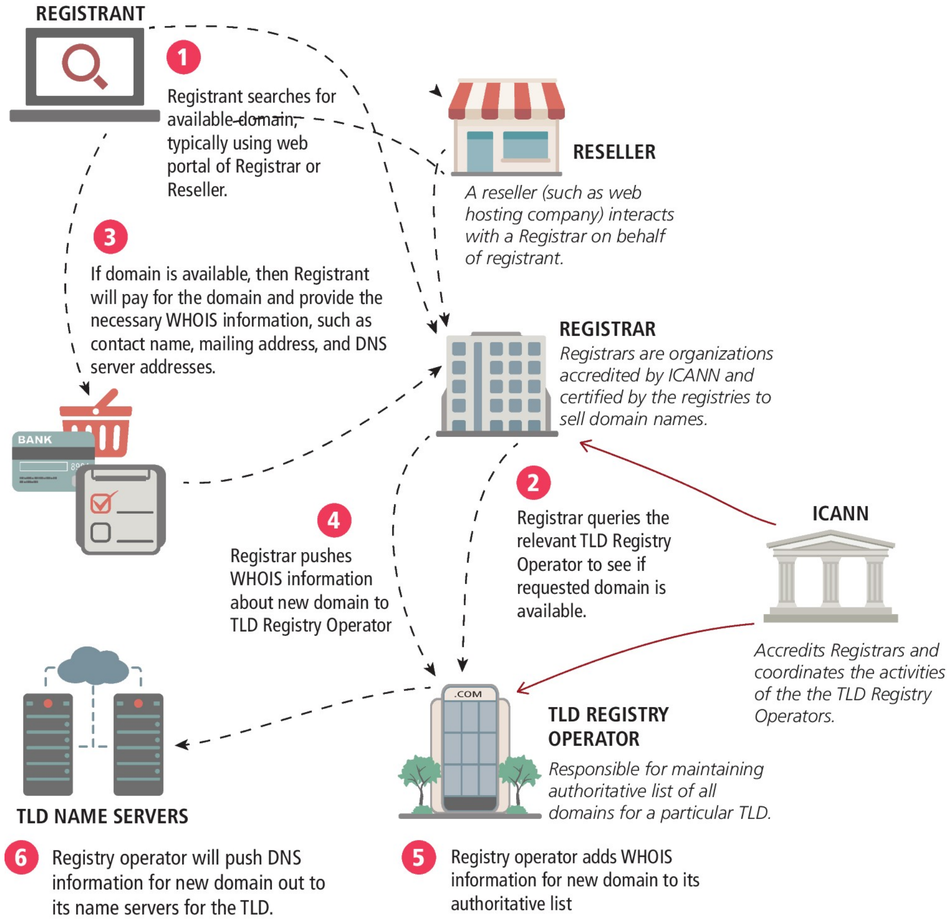
Fig. Registrar of governmental websites
ex: ***moi.gov.sa***

“Domain registrations under (.sa) started in 1995; while it started in 2010 for (.السعودية)

In February 2021, accredited domain registrars for Saudi domain names have begun providing services to registrants Individuals and Non-Government entities, and for Government entities the registration services are provided by Digital Government Authority (DGA)” [https://www.citc.gov.sa/]

Domain name registration process

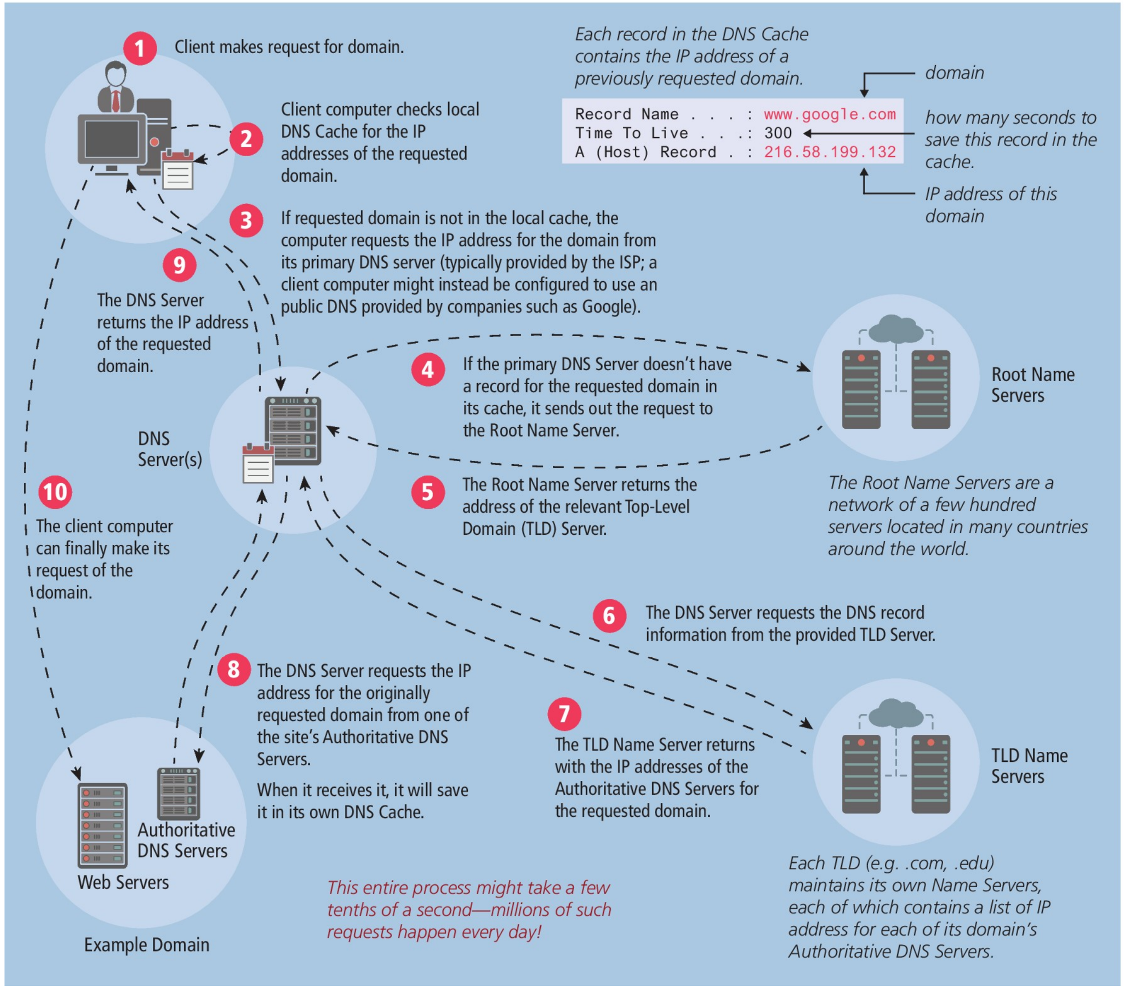
1. Registrant searches for domain, typically using web portal of Registrar or Reseller.
2. Registrar queries the relevant TLD Registry Operator to see if requested domain is available.
3. If domain is available, then Registrant will pay for the domain and provide the necessary WHOIS information
4. Registrar pushes WHOIS information about new domain to TLD Registry Operator
5. Registry operator adds WHOIS information for new domain to its authoritative list
6. Registry operator will push DNS information for new domain out to its name servers for the TLD.


REGISTRANT
1
Registrant searches for available domain, typically using web portal of Registrar or Reseller.
RESELLER
A reseller (such as web hosting company) interacts with a Registrar on behalf of registrant.
3
If domain is available, then Registrant will pay for the domain and provide the necessary WHOIS information, such as contact name, mailing address, and DNS server addresses.
REGISTRAR
Registrars are organizations accredited by ICANN and certified by the registries to sell domain names.
BANK
2
Registrar queries the relevant TLD Registry Operator to see if requested domain is available.
ICANN
4
Registrar pushes WHOIS information about new domain to TLD Registry Operator
Accredits Registrars and coordinates the activities of the the TLD Registry Operators.
.COM
TLD REGISTRY OPERATOR
Responsible for maintaining authoritative list of all domains for a particular TLD.
TLD NAME SERVERS
6
Registry operator will push DNS information for new domain out to its name servers for the TLD.
5
Registry operator adds WHOIS information for new domain to its authoritative list

Pearson

Address Resolution

1. Client makes request for domain
2. Client computer checks local DNS Cache
3. If requested domain is not in the local cache, the computer requests the IP address for the domain from its primary DNS server
4. If the primary DNS Server doesn't have a record for the requested domain in its cache, it sends out the request to the Root Name Server
5. The Root Name Server returns the address of the relevant Top-Level Domain (TLD) Server.



1 Client makes request for domain.

2 Client computer checks local DNS Cache for the IP addresses of the requested domain.

Each record in the DNS Cache contains the IP address of a previously requested domain.

```
Record Name . . . : www.google.com
Time To Live . . .: 300
A (Host) Record . : 216.58.199.132
```

domain

how many seconds to save this record in the cache.

IP address of this domain

3 If requested domain is not in the local cache, the computer requests the IP address for the domain from its primary DNS server (typically provided by the ISP; a client computer might instead be configured to use an public DNS provided by companies such as Google).

4 If the primary DNS Server doesn't have a record for the requested domain in its cache, it sends out the request to the Root Name Server.

5 The Root Name Server returns the address of the relevant Top-Level Domain (TLD) Server.

The Root Name Servers are a network of a few hundred servers located in many countries around the world.

6 The DNS Server requests the DNS record information from the provided TLD Server.

7 The TLD Name Server returns with the IP addresses of the Authoritative DNS Servers for the requested domain.

Each TLD (e.g. .com, .edu) maintains its own Name Servers, each of which contains a list of IP address for each of its domain's Authoritative DNS Servers.

8 The DNS Server requests the IP address for the originally requested domain from one of the site's Authoritative DNS Servers.

When it receives it, it will save it in its own DNS Cache.

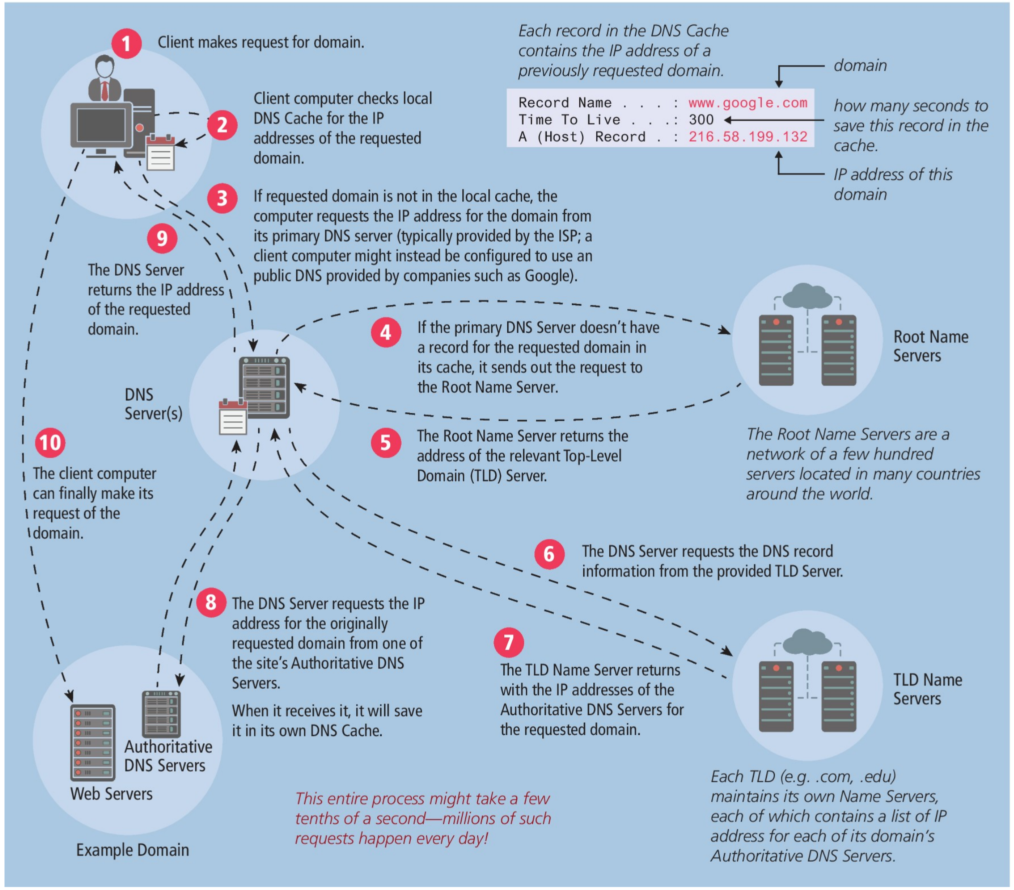
9 The DNS Server returns the IP address of the requested domain.

10 The client computer can finally make its request of the domain.

This entire process might take a few tenths of a second—millions of such requests happen every day!

Address Resolution (cont)

6. The DNS Server requests the DNS record information from the provided TLD Server.
7. The TLD Name Server returns with the IP addresses of the Authoritative DNS Servers for the requested domain.
8. The DNS Server requests the IP address for the originally requested domain from one of the site’s Authoritative DNS Servers.
9. The DNS Server returns the IP address of the requested domain.
10. The client computer can finally make its request of the domain.



1 Client makes request for domain.

2 Client computer checks local DNS Cache for the IP addresses of the requested domain.

Each record in the DNS Cache contains the IP address of a previously requested domain.

```
Record Name . . . : www.google.com
Time To Live . . .: 300
A (Host) Record . : 216.58.199.132
```

domain

how many seconds to save this record in the cache.

IP address of this domain

3 If requested domain is not in the local cache, the computer requests the IP address for the domain from its primary DNS server (typically provided by the ISP; a client computer might instead be configured to use an public DNS provided by companies such as Google).

9 The DNS Server returns the IP address of the requested domain.

DNS Server(s)

4 If the primary DNS Server doesn’t have a record for the requested domain in its cache, it sends out the request to the Root Name Server.

Root Name Servers

5 The Root Name Server returns the address of the relevant Top-Level Domain (TLD) Server.

The Root Name Servers are a network of a few hundred servers located in many countries around the world.

10 The client computer can finally make its request of the domain.

6 The DNS Server requests the DNS record information from the provided TLD Server.

8 The DNS Server requests the IP address for the originally requested domain from one of the site’s Authoritative DNS Servers.

When it receives it, it will save it in its own DNS Cache.

7 The TLD Name Server returns with the IP addresses of the Authoritative DNS Servers for the requested domain.

TLD Name Servers

Authoritative DNS Servers

Web Servers

Example Domain

This entire process might take a few tenths of a second—millions of such requests happen every day!

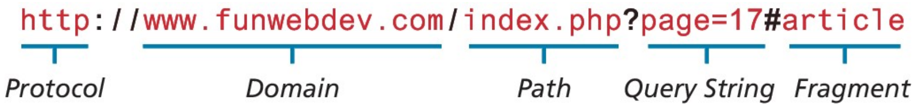
Each TLD (e.g. .com, .edu) maintains its own Name Servers, each of which contains a list of IP address for each of its domain’s Authoritative DNS Servers.

Uniform Resource Locators

Uniform Resource Locators (URL) allow clients to request particular resources (files) from the server.

URL’s consist of two required components:

1. the protocol used to connect and
2. the domain (or IP address) to connect to.



http://www.funwebdev.com/index.php?page=17#article

Protocol Domain Path Query String Fragment

Port (URL)

- A **port** is a type of software connection point used by the underlying TCP/IP protocol and the connecting computer.
- Although the port attribute is not commonly used in production sites, it can be used to route requests to a test server, to perform a stress test, or even to circumvent Internet filters.
- If no port is specified, the protocol determines which port to use. For instance, port 80 is the default port for web-related HTTP requests.
- Syntax is to add a colon after the domain, then specify an integer port number. **http://funwebdev.com:8080/** would connect on port 8080

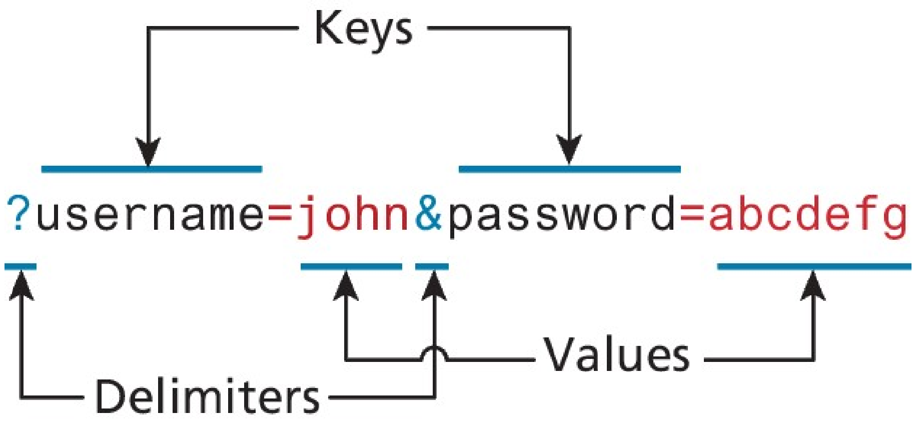
Path (URL)

- The **path** is an important concept to anyone who has ever used a computer file system.
- The root of a web server corresponds to a folder somewhere on that server.On many Linux servers that path is **/var/www/html/** or something similar (for Windows it is often **/inetpub/wwwroot/**).
- The path is optional. However, when requesting a folder or the top-level page of a domain, the web server will decide which file to send you.

Query String (URL)

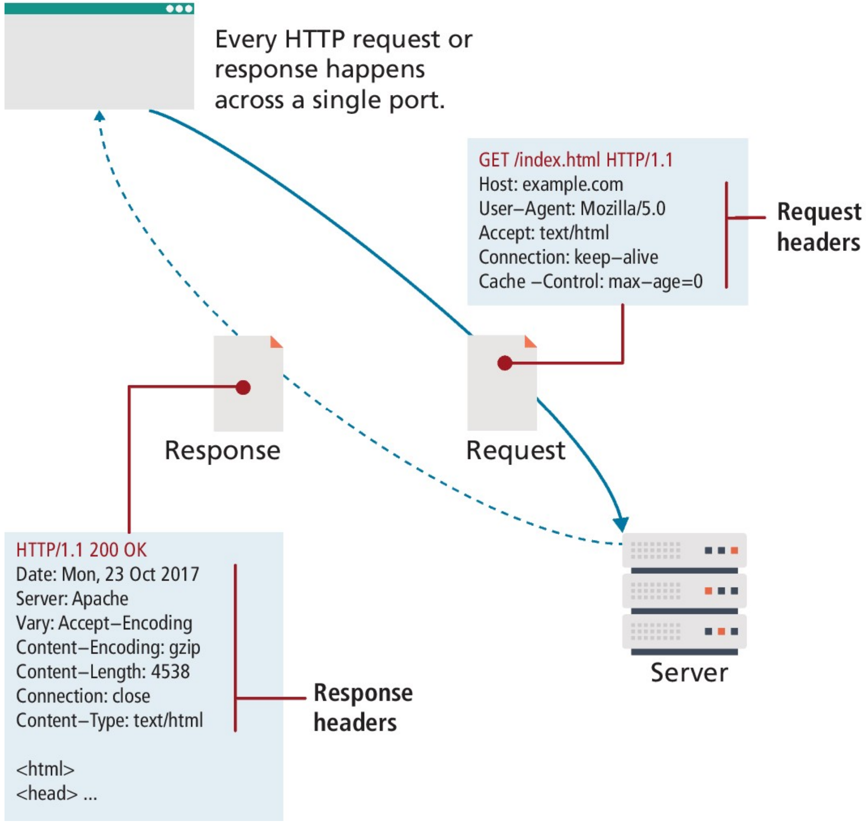
Query strings will be covered in depth when we learn more about HTML forms and server-side programming. They are a critical way of passing information, such as user form input, from the client to the server.

- In URLs, they are encoded as key-value pairs delimited by & symbols and preceded by the ? Symbol



Hypertext Transfer Protocol

- **HTTP** is an essential part of the web.
- HTTP establishes a TCP connection on port 80 (by default). The server waits for the request, and then responds with a
  - Headers,
  - Response code,
  - an optional message (which can include files)



Every HTTP request or response happens across a single port.

```
GET /index.html HTTP/1.1
Host: example.com
User–Agent: Mozilla/5.0
Accept: text/html
Connection: keep–alive
Cache –Control: max–age=0
```

```
HTTP/1.1 200 OK
Date: Mon, 23 Oct 2017
Server: Apache
Vary: Accept–Encoding
Content–Encoding: gzip
Content–Length: 4538
Connection: close
Content–Type: text/html

<html>
<head> ...
```

HTTP Headers

Headers are sent in the request from the client and received in the response from the server. Headers are one of the most powerful aspects of HTTP.

- **Request headers** include data about the client machine
  - Host, User-Agent, Cache settings and more
- **Response headers** have information about the server answering the request and the data being sent
  - Server, Last Modified, Content Type, Encoding,

Pearson

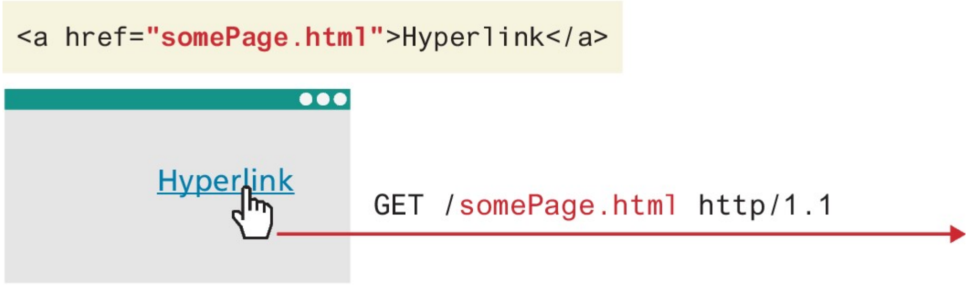
HTTP Request Methods

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

- Other HTTP verbs such as CONNECT, TRACE, and OPTIONS are less commonly used and are not covered.

GET Request

- The most common type of HTTP request is the **GET request**.
- One is asking for a resource located at a specified URL to be retrieved.
- Whenever you click on a link, type-in a URL in your browser, or click on a bookmark, you are *usually* making a GET request.
- Data can be transmitted through a GET request, with a query string

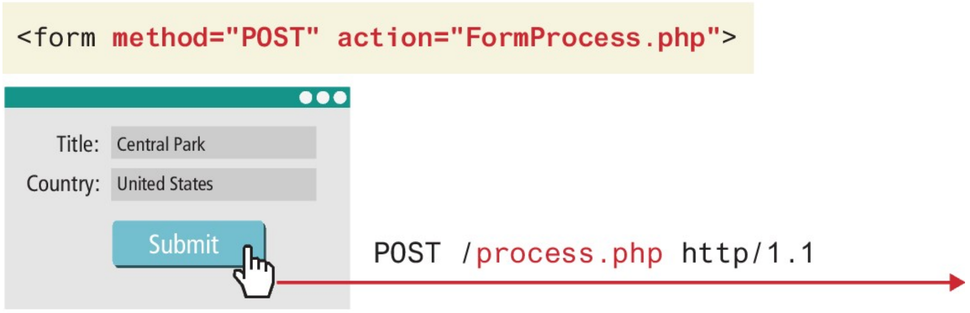


```
<a href="somePage.html">Hyperlink</a>
```

```
GET /somePage.html http/1.1
```

POST Request

- The other common request method is the **POST request**.
- This method is normally used to transmit data to the server using an HTML form



```
<form method="POST" action="FormProcess.php">
```

Title: Central Park

Country: United States

```
POST /process.php http/1.1
```

Response Codes

- **Response codes** are integer values returned by the server as part of the response header.
- These codes describe the state of the request, including whether it was successful, had errors, requires permission, and more.
- The codes use the first digit to indicate the category of response.
  - 2## codes are for successful responses,
  - 3## are for redirection-related responses,
  - 4## codes are client errors.
  - 5## codes are server errors.

HTTP Response Codes

Code	Description
200: OK	The request was successful.
301: Moved Permanently	Tells the client that the requested resource has permanently moved.
304: Not Modified	If the client requested a resource with appropriate Cache-Control headers, the response might say that the resource on the server is no newer than the one in the client cache.
401: Unauthorized	Some web resources are protected and require the user to provide credentials to access the resource.
404: Not found	404 codes are one of the only ones known to web users. Many browsers will display an HTML page with the 404 code to them when the requested resource was not found.
414: Request URI too long	A 414 response code likely means too much data is likely trying to be submitted via the URL.
500: Internal server error	This error provides almost no information to the client except to say the server has encountered an error.

Pearson

Web Browsers (Fetching a web page)

- Seeing a single web page is facilitated by the 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
- The algorithms within browsers to download, parse, layout, fetch assets, and create the final interactive page for the user are commonly referred to collectively as the *rendering* of the page

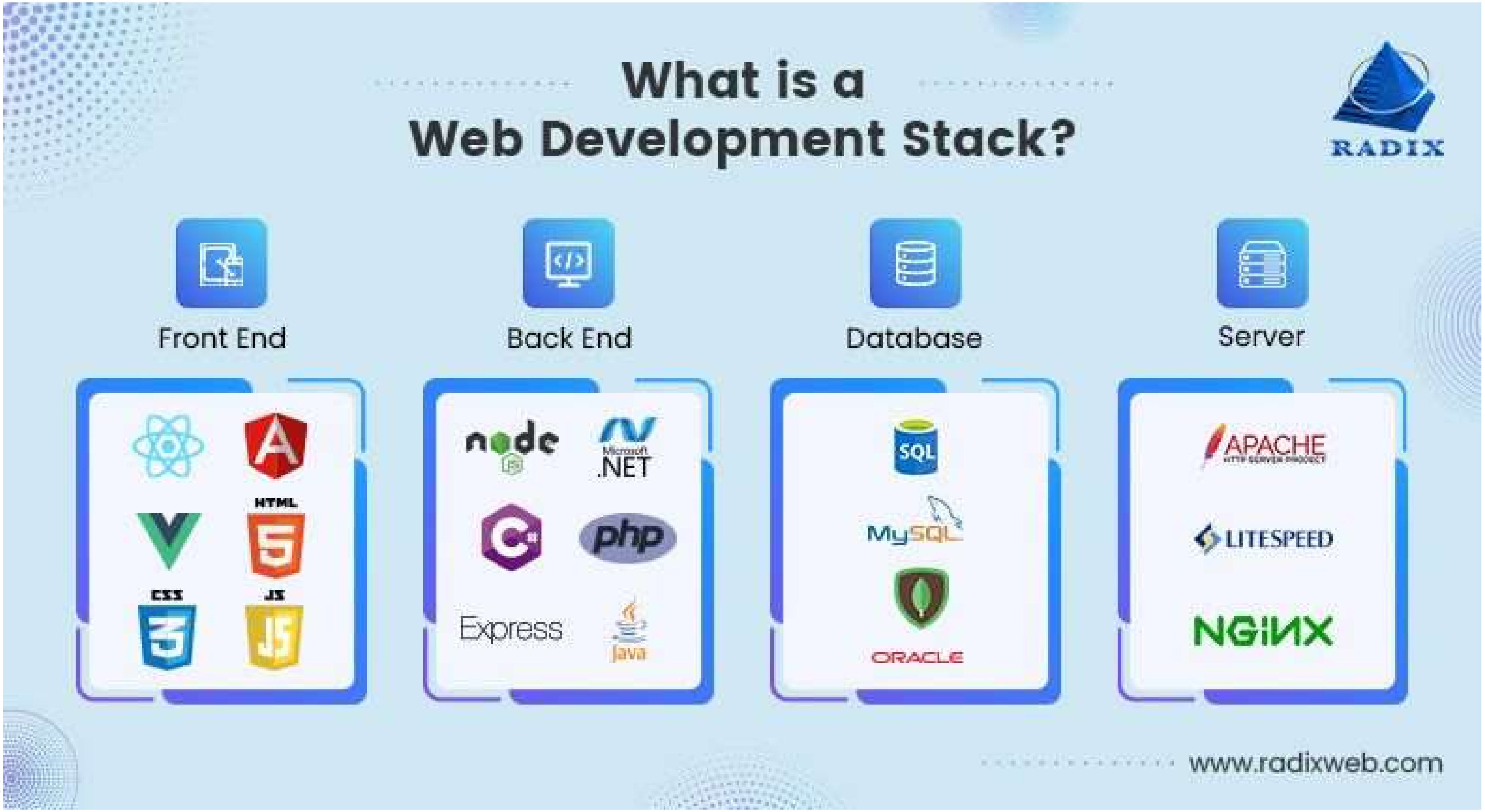
Web Servers

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

- Real-world websites typically have many web servers configured together in web farms.

- Regardless of the physical characteristics of the server, one must choose an application stack to run a website. This **application stack** will include
  - an operating system,
  - web server software,
  - a database,
  - a backend programming language and runtime,
  - a set of languages/frameworors/libraries for the front-end.

Web Development stacks



END