# How the Web Works

Chapter 2

# Chapter 2

1. Internet Protocols
2. Domain Name System
3. Uniform Resource Locators
4. Hypertext Transfer Protocol
5. Web Browsers
6. Web Servers
7. Summary

# Chapter 2

1. Internet Protocols
2. Domain Name System
3. Uniform Resource Locators
4. Hypertext Transfer Protocol
5. Web Browsers
6. Web Servers
7. Summary

# Internet Protocols

## A Layered Architecture

TCP/IP.

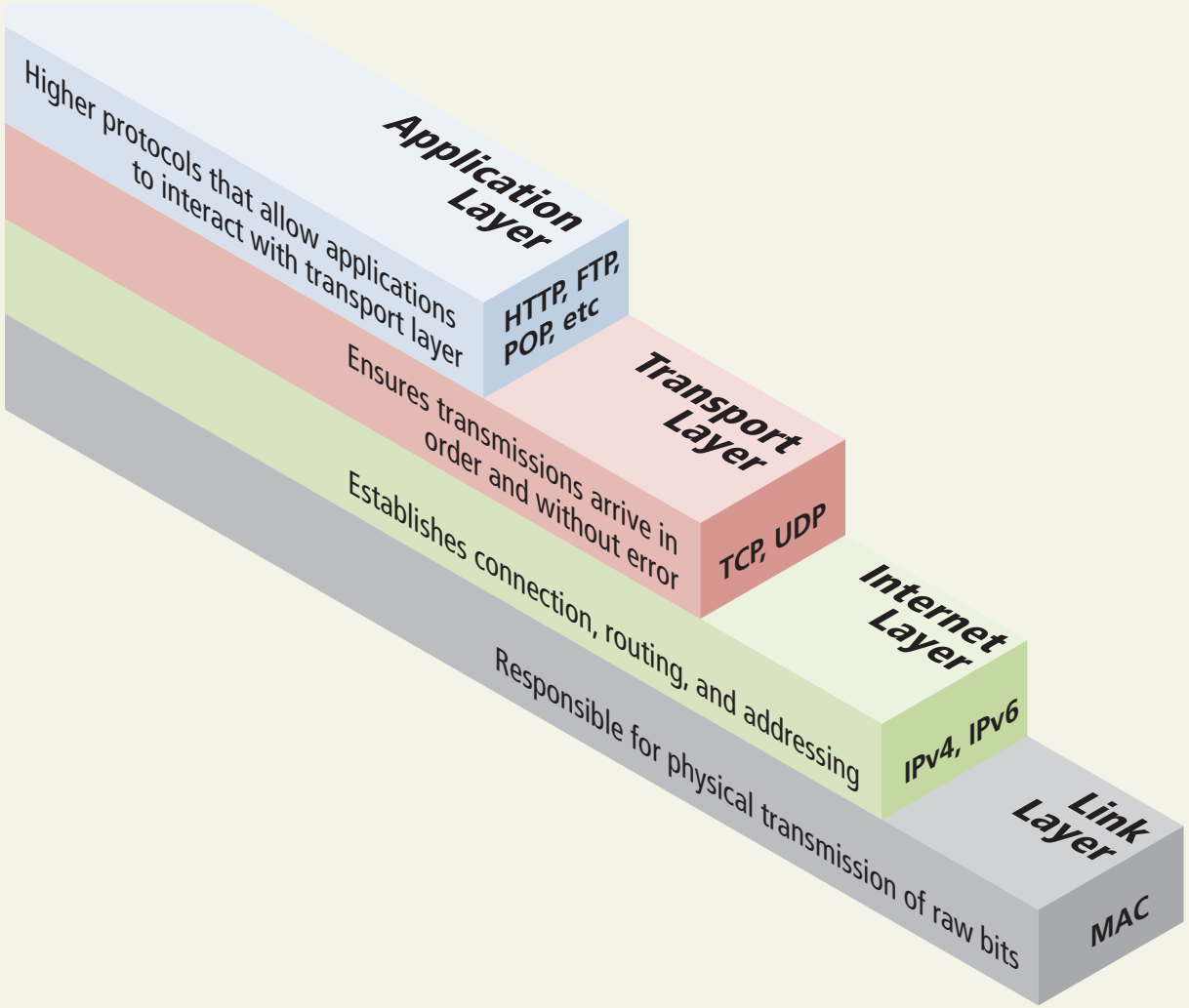
These protocols have been implemented in every operating system, and make fast web development possible.

Networking is it's own entire discipline.

Web developer needs general awareness of what the suite of Internet protocols does

# Internet Protocols

A Layered Architecture



# Internet Protocols

## Link Layer

- Responsible for
  - physical transmission of data across media (both wired and wireless) and
  - Establishing logical links.

  It handles issues like packet creation, transmission, reception, error detection, collisions, line sharing, and more.

  Much more to learn in Networking courses outside of web development.

# Internet Protocols

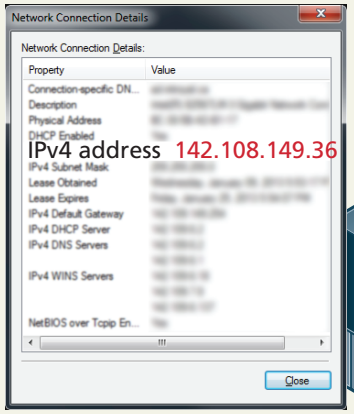
## Internet Layer

The Internet layer provides “best effort” communication.

Makes use of IP addresses

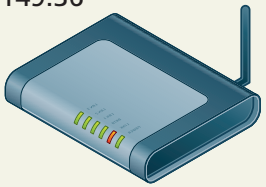
# Internet Protocols

## Internet Layer (IP)

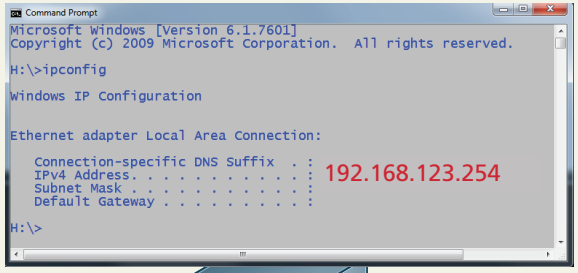


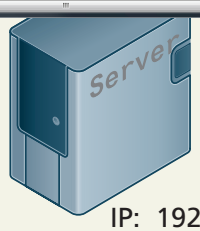

IP: 142.108.149.36

IP: 22.15.216.13



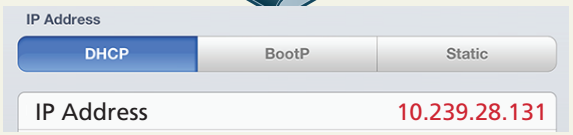
IP: 10.238.28.131

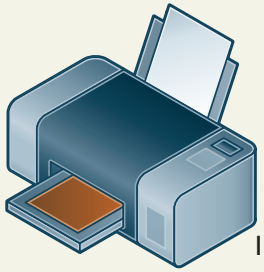



IP: 192.168.123.254

IP: 10.239.28.131

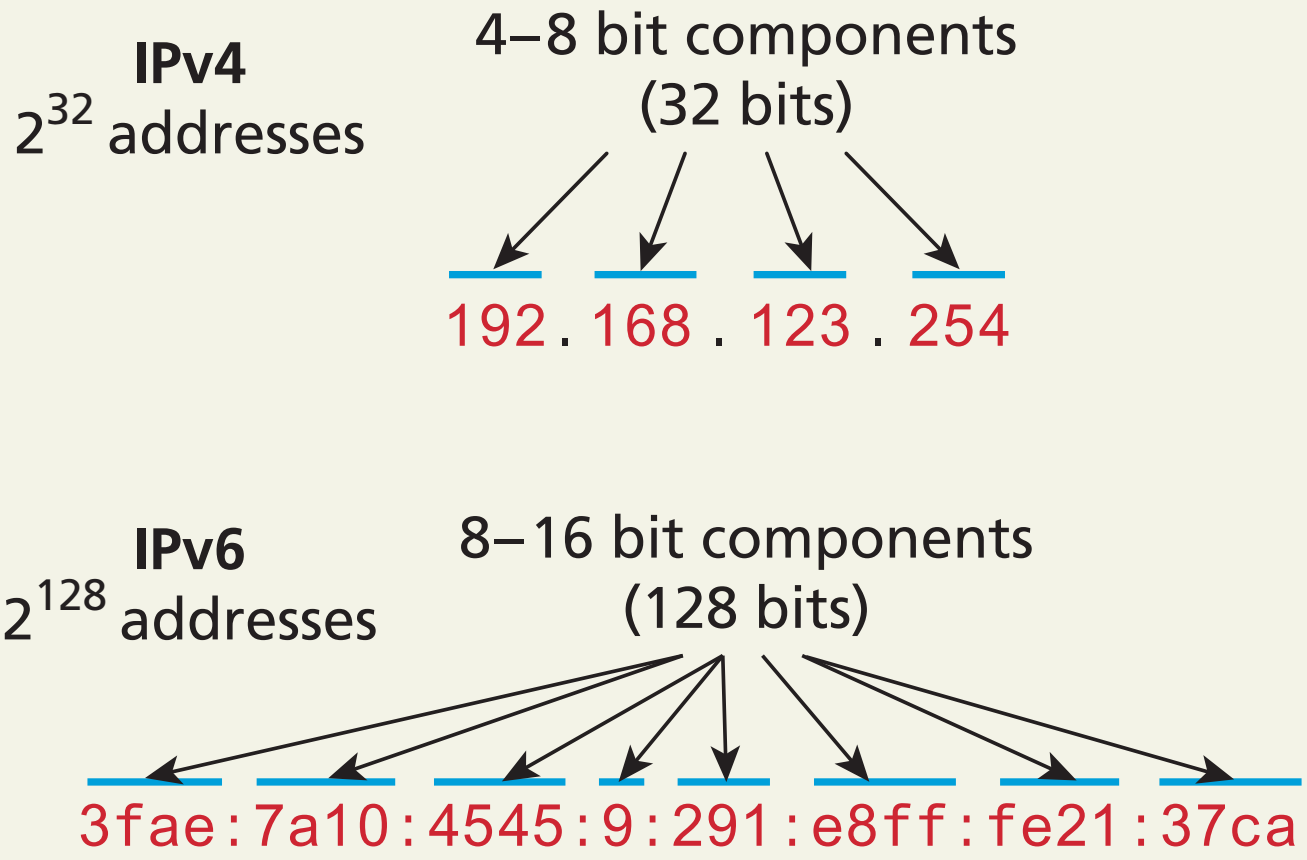


IP: 142.181.80.3

# Internet Protocols

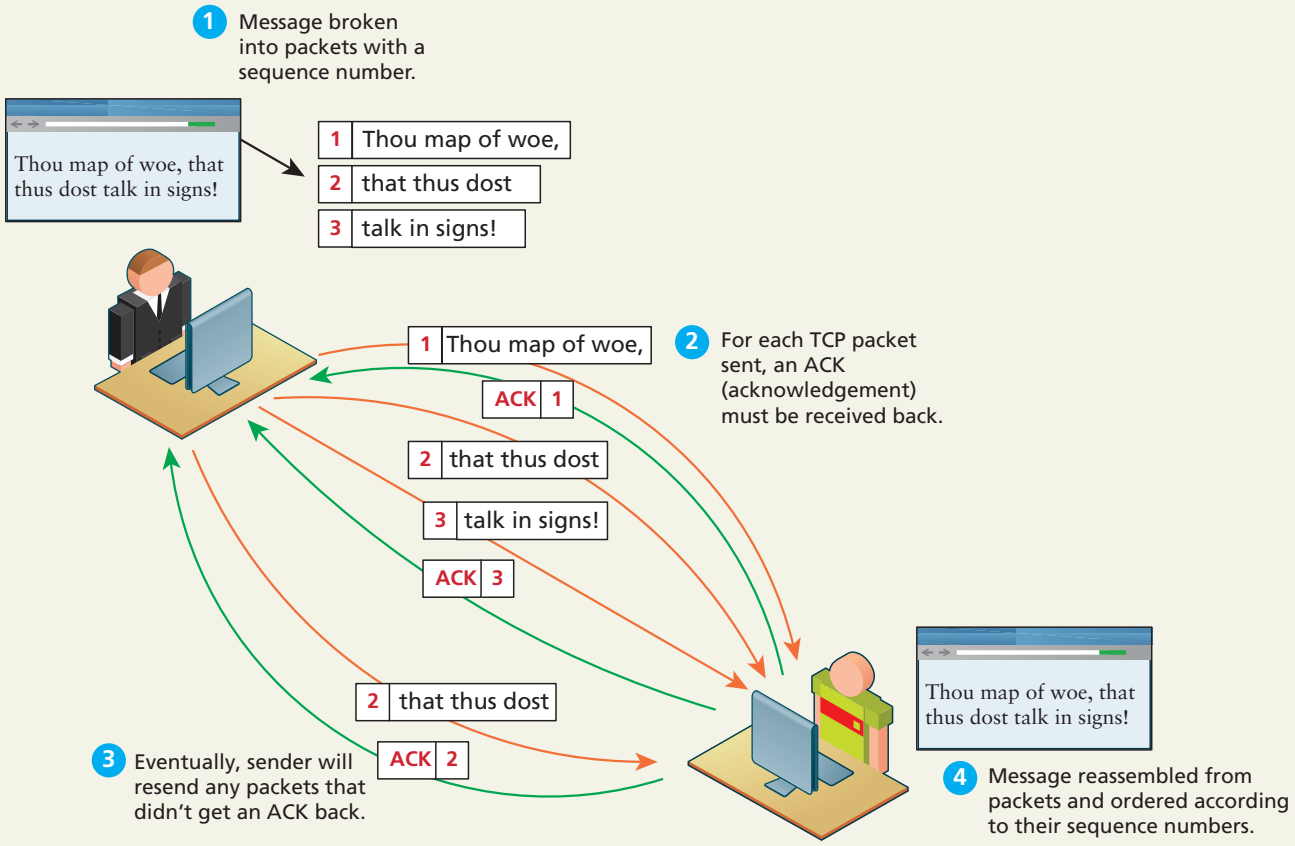
IP addresses



# Internet Protocols

## Transport Layer (TCP)

- Ensures transmissions arrive in order and without error



# Internet Protocols

## Application Layer

There are **many** application layer protocols. Web developers should be aware of :

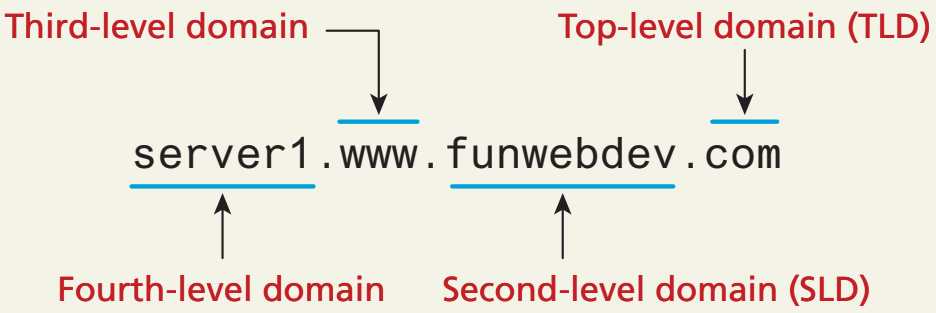
- **HTTP.** The Hypertext Transfer Protocol is used for web communication.
- **SSH.** The Secure Shell Protocol allows remote command-line connections to servers.
- **FTP**. The File Transfer Protocol is used for transferring files between computers.
- **POP/IMAP/SMTP**. Email-related protocols for transferring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

# Chapter 2

1. Internet Protocols
2. Domain Name System
3. Uniform Resource Locators
4. Hypertext Transfer Protocol
5. Web Browsers
6. Web Servers
7. Summary

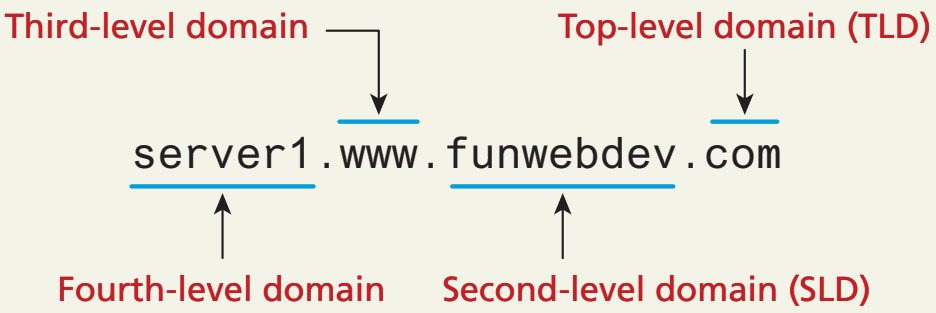
# Domain Name System

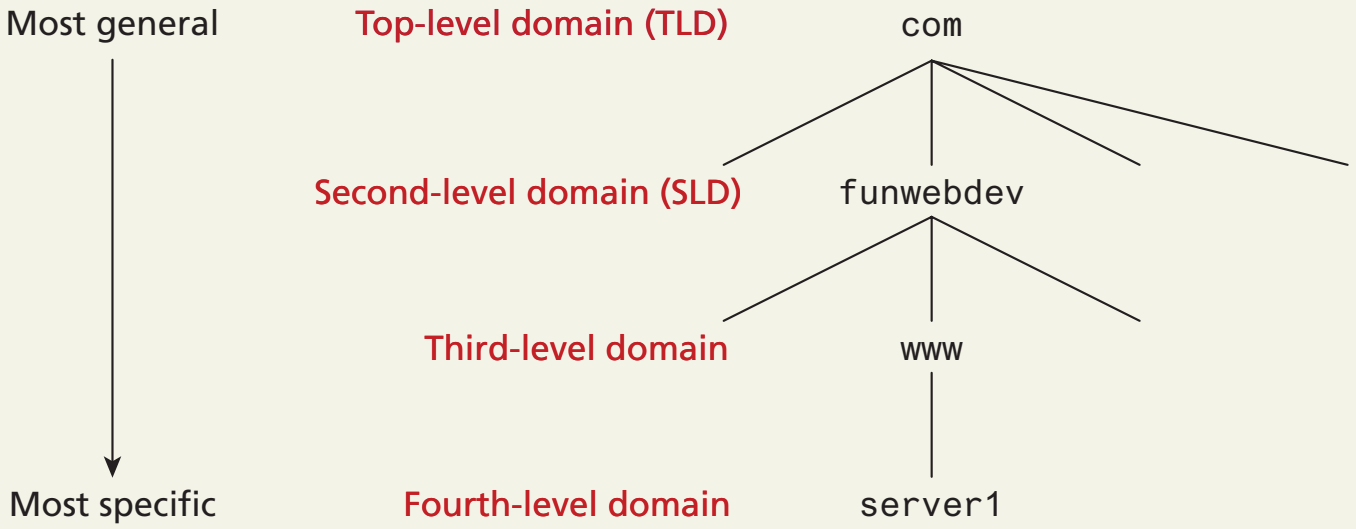
## Name Levels



# Domain Name System

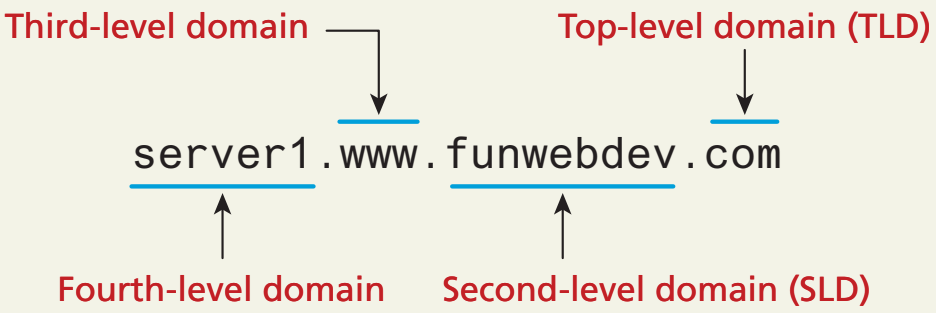
## Name Levels

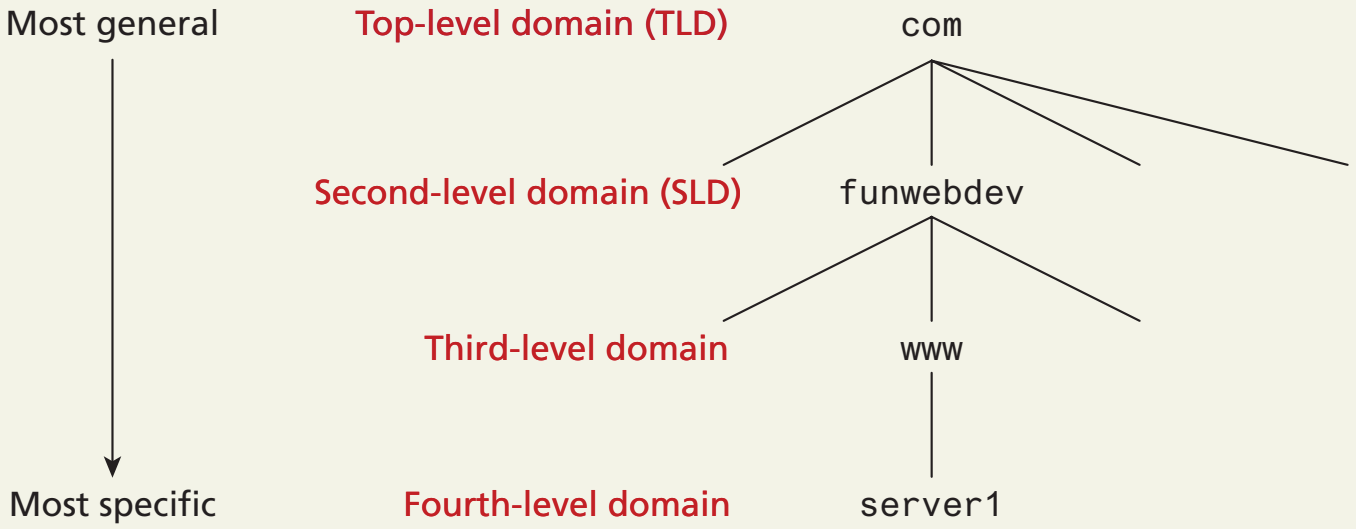


# Domain Name System

## Name Levels



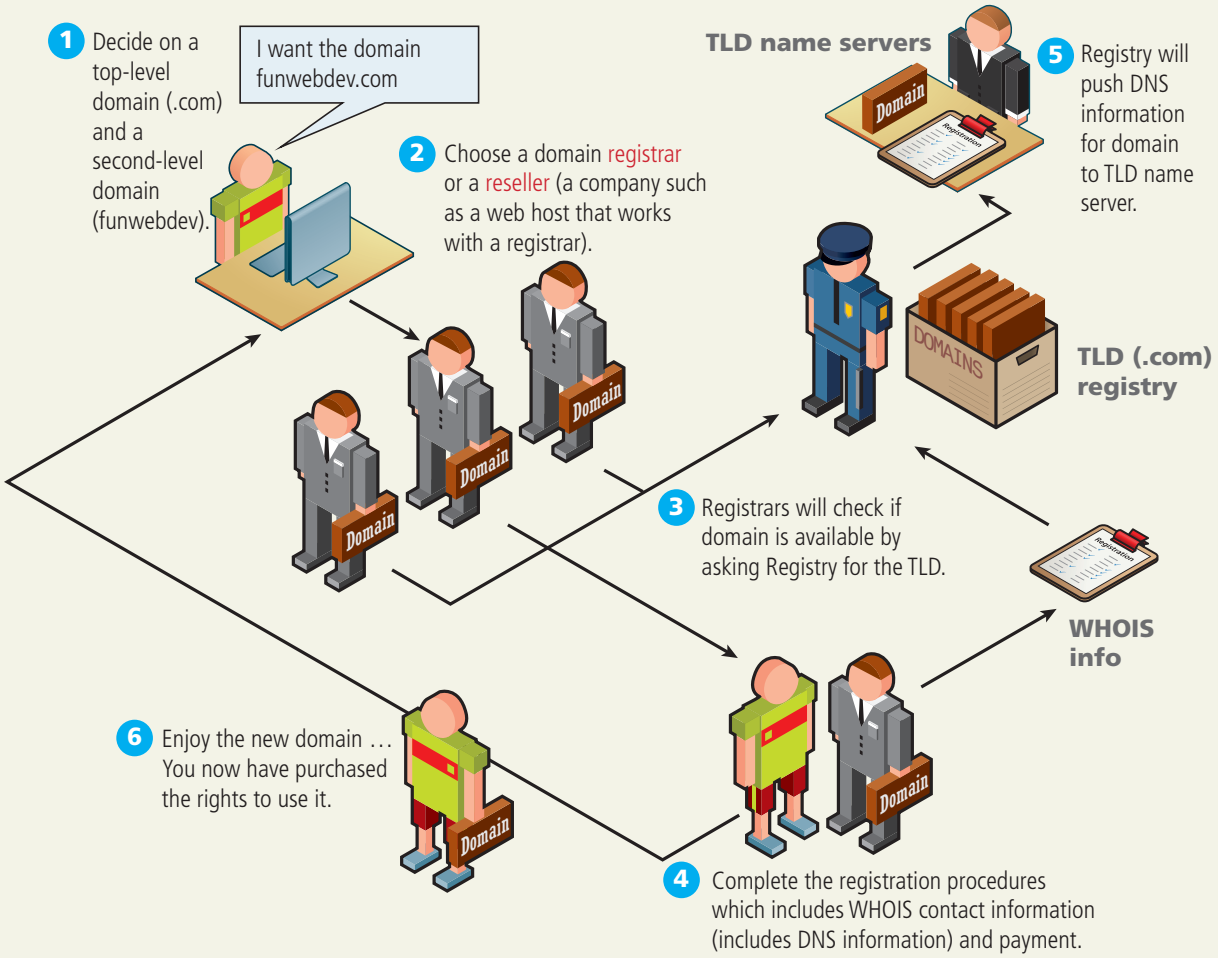
# Domain Name System

## Types of Top Level Domains

- Generic top-level domain (gTLD)
  - Unrestricted. TLDs include .com, .net, .org, and .info.
  - Sponsored. TLDs including .gov, .mil, .edu, and others.
  - New TLDs.
- Country code top-level domain (ccTLD)
  - TLDs include .us , .ca , .uk , and .au.
  - Internationalized Domain Names
- arpa

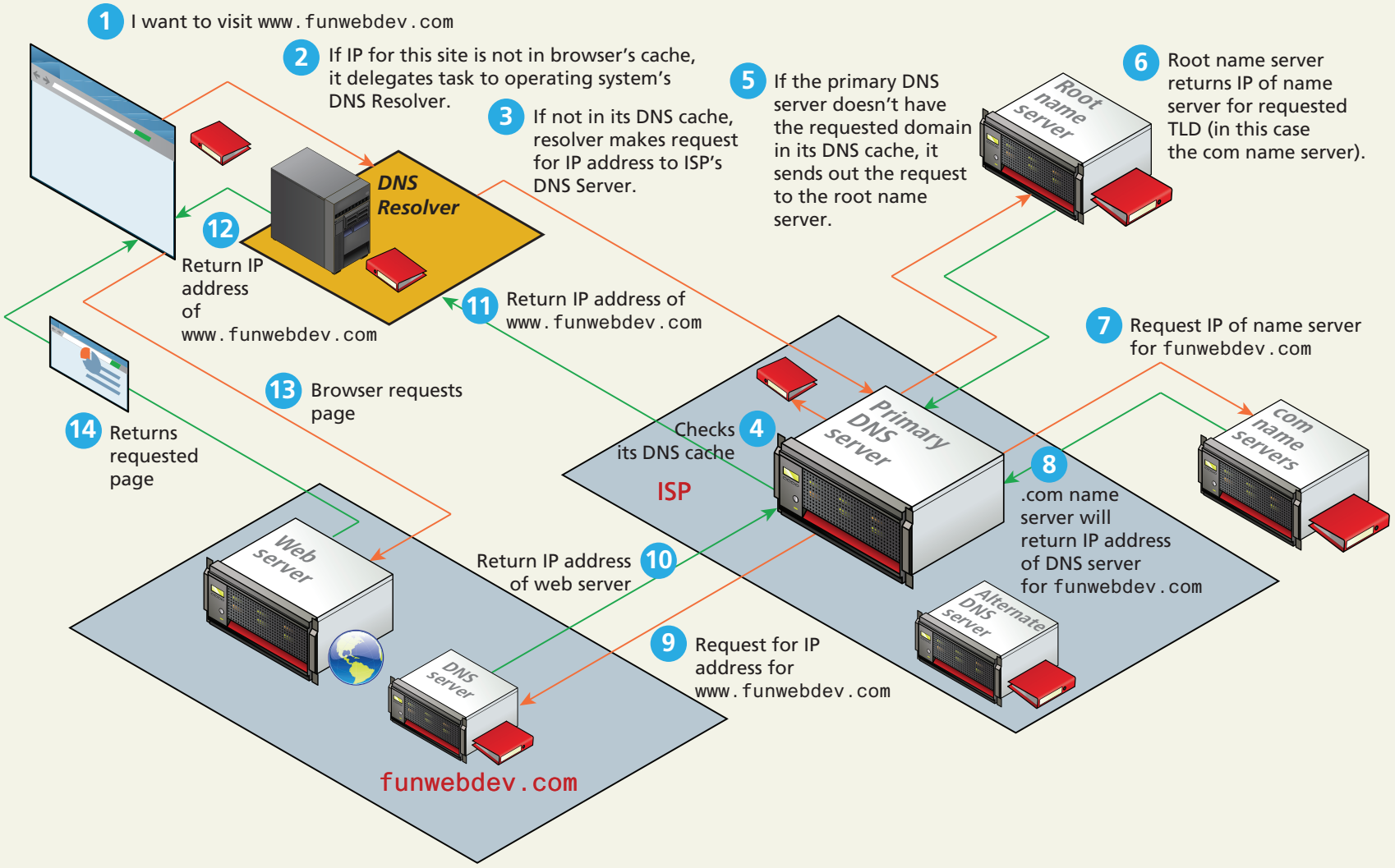
# Domain Name System

## Name Registration



# Domain Name System

## Address Resolution



# Chapter 2

1 Internet Protocols

2 Domain Name System

3 Uniform Resource Locators

4 Hypertext Transfer Protocol

5 Web Browsers

6 Web Servers

7 Summary

# Uniform Resource Locators

Overview

# Uniform Resource Locators

Protocol

Recall that in Section 2.1, we listed several application layer protocols on the TCP/IP stack. FTP, SSH, HTTP, POP, IMAP, DNS, …

Requesting

- **ftp**://example.com/abc.txt → sends out an FTP request on port 21, while
- **http**://example.com/abc.txt → transmits an HTTP request on port 80.

# Uniform Resource Locators

## Domain

- The domain identifies the server from which we are requesting resources.
- Since the DNS system is case insensitive, this part of the URL is case insensitive.
- Alternatively, an IP address can be used for the domain

# Uniform Resource Locators

## Port

- The optional port attribute allows us to specify connections to ports other than the defaults

- Add a colon after the domain, then specify an integer port number.

# Uniform Resource Locators

## Path

Familiar concept to anyone who has ever used a computer file system.

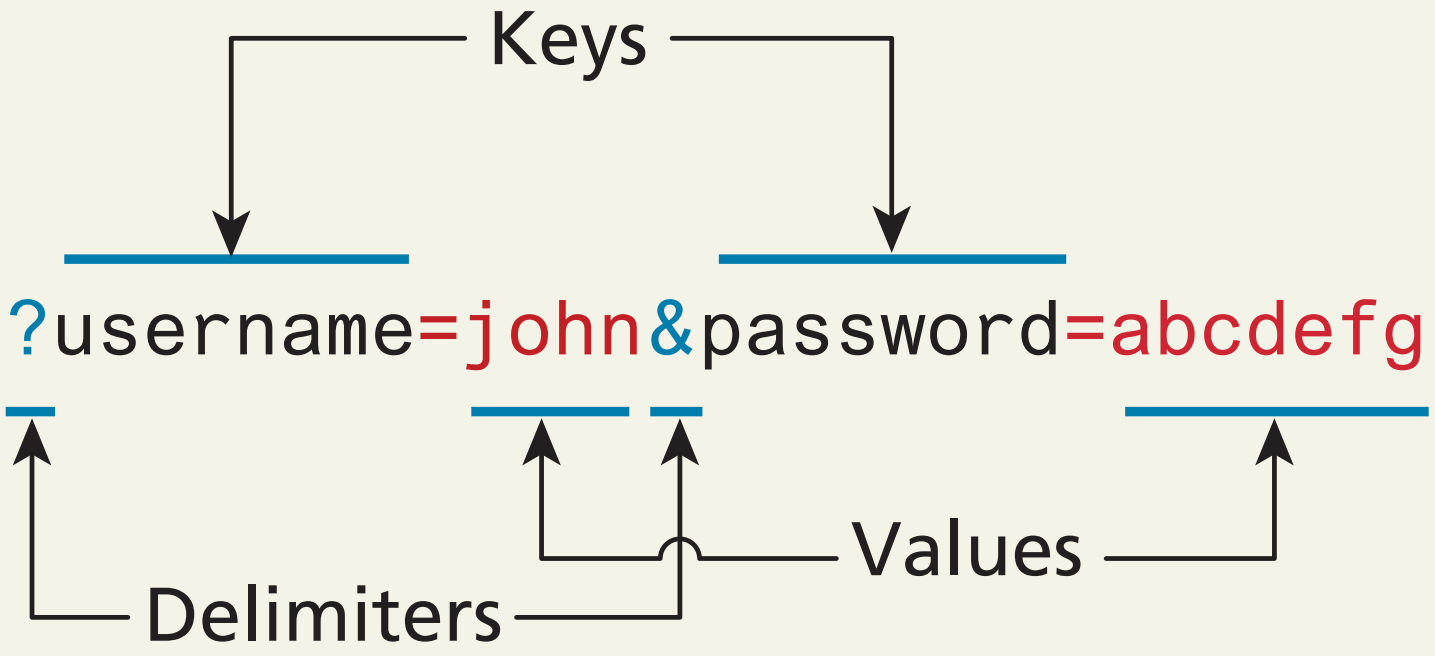
The root of a web server corresponds to a folder somewhere on that server.

- On many Linux servers that path is /var/www/html/
- On Windows IIS machines it is often /inetpub/wwwroot/

The path is optional. However, when requesting a folder or the top-level page, the web server will decide which file to send you.

# Uniform Resource Locators

## Query String



# Uniform Resource Locators

## Fragment

A way of requesting a portion of a page.

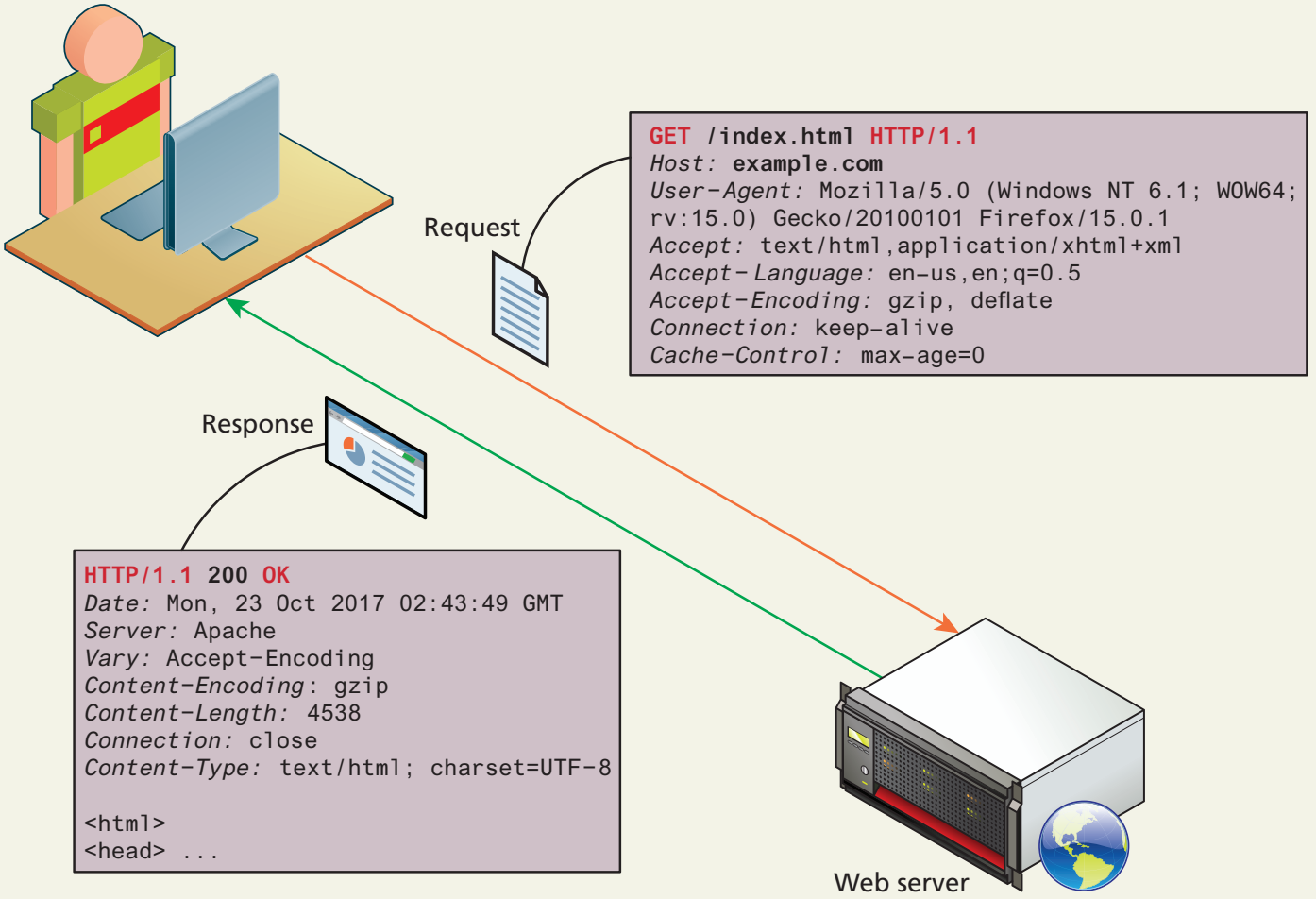
- Browsers will see the fragment in the URL, seek out the tag anchor in the HTML, and scroll the website to it.

# Chapter 2

1. Internet Protocols
2. Domain Name System
3. Uniform Resource Locators
4. Hypertext Transfer Protocol
5. Web Browsers
6. Web Servers
7. Summary

# Hypertext Transfer Protocol

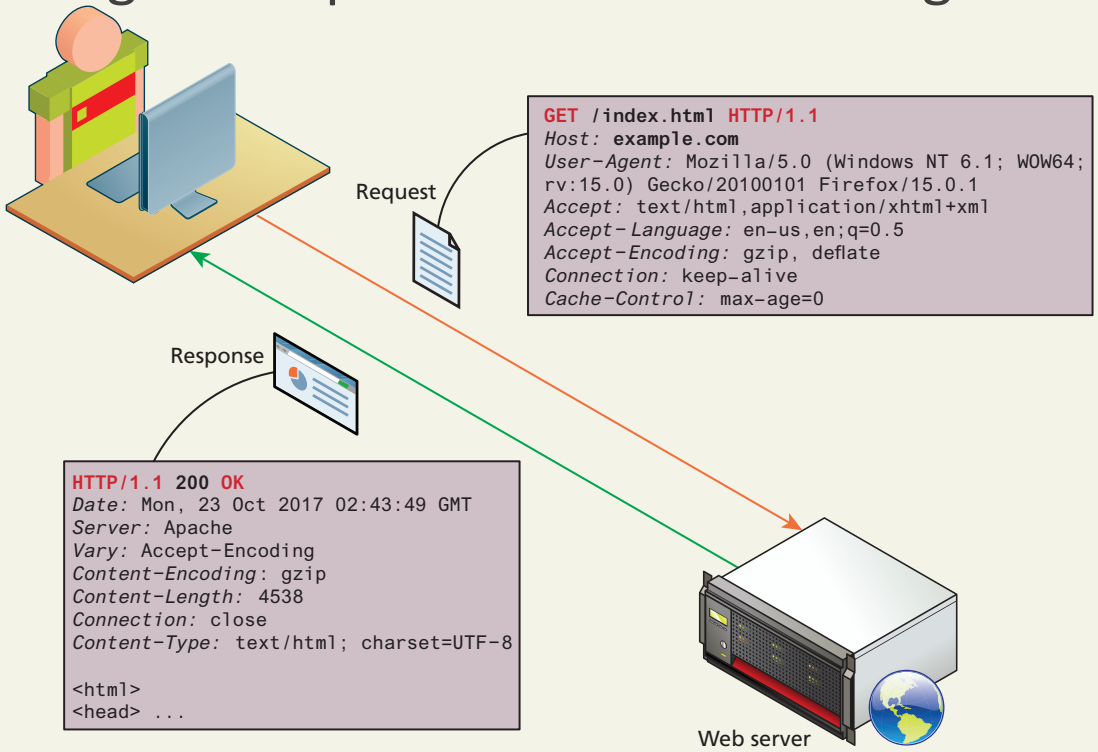
## Headers



# Hypertext Transfer Protocol

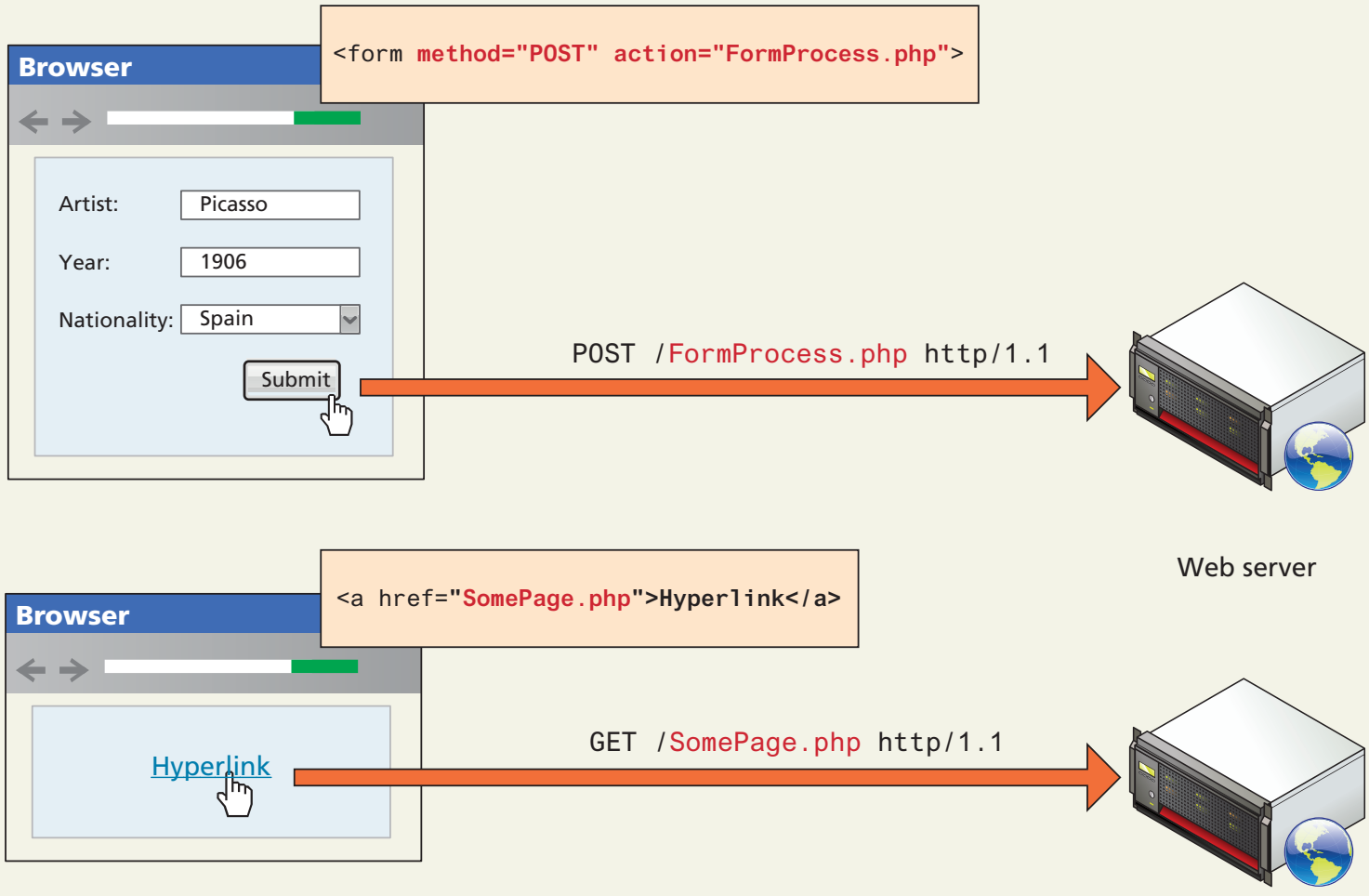
## Headers

- **Request headers** include data about the client machine.
- **Response headers** have information about the server answering the request and the data being sent



# Hypertext Transfer Protocol

## Request Methods



# Hypertext Transfer Protocol

## Response Codes

- 2## codes are for successful responses,
- 3## are for redirection-related responses,
- 4## codes are **client** errors,
- 5## codes are **server** errors.

# Hypertext Transfer Protocol

(Some) Response Codes

200: OK

301: Moved Permanently

304: Not Modified

307: Temporary redirect

400: Bad Request

401: Unauthorized

404: Not found

414: Request URI too long

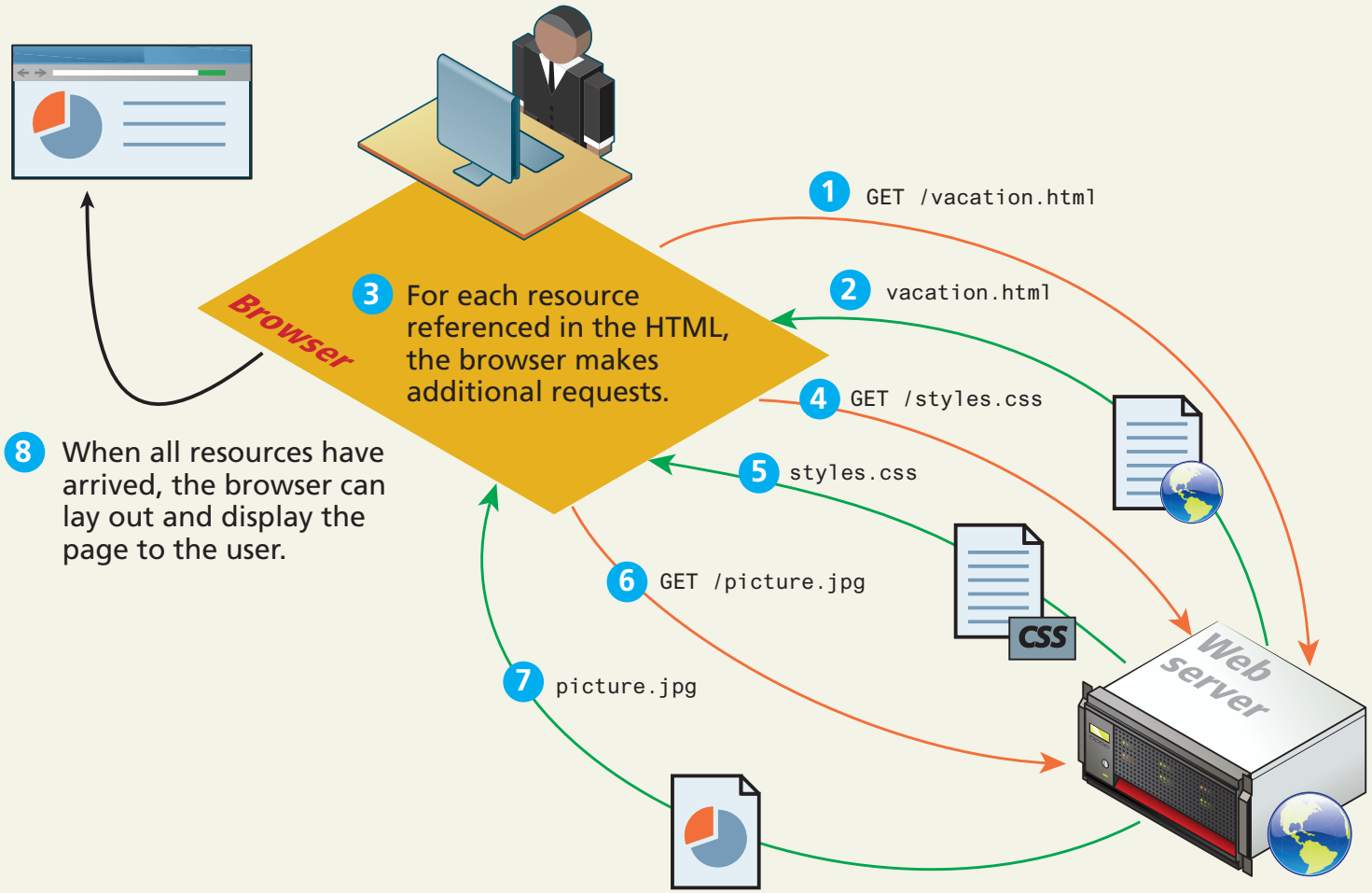
500: Internal server error

# Chapter 2

1. Internet Protocols
2. Domain Name System
3. Uniform Resource Locators
4. Hypertext Transfer Protocol
5. Web Browsers
6. Web Servers
7. Summary

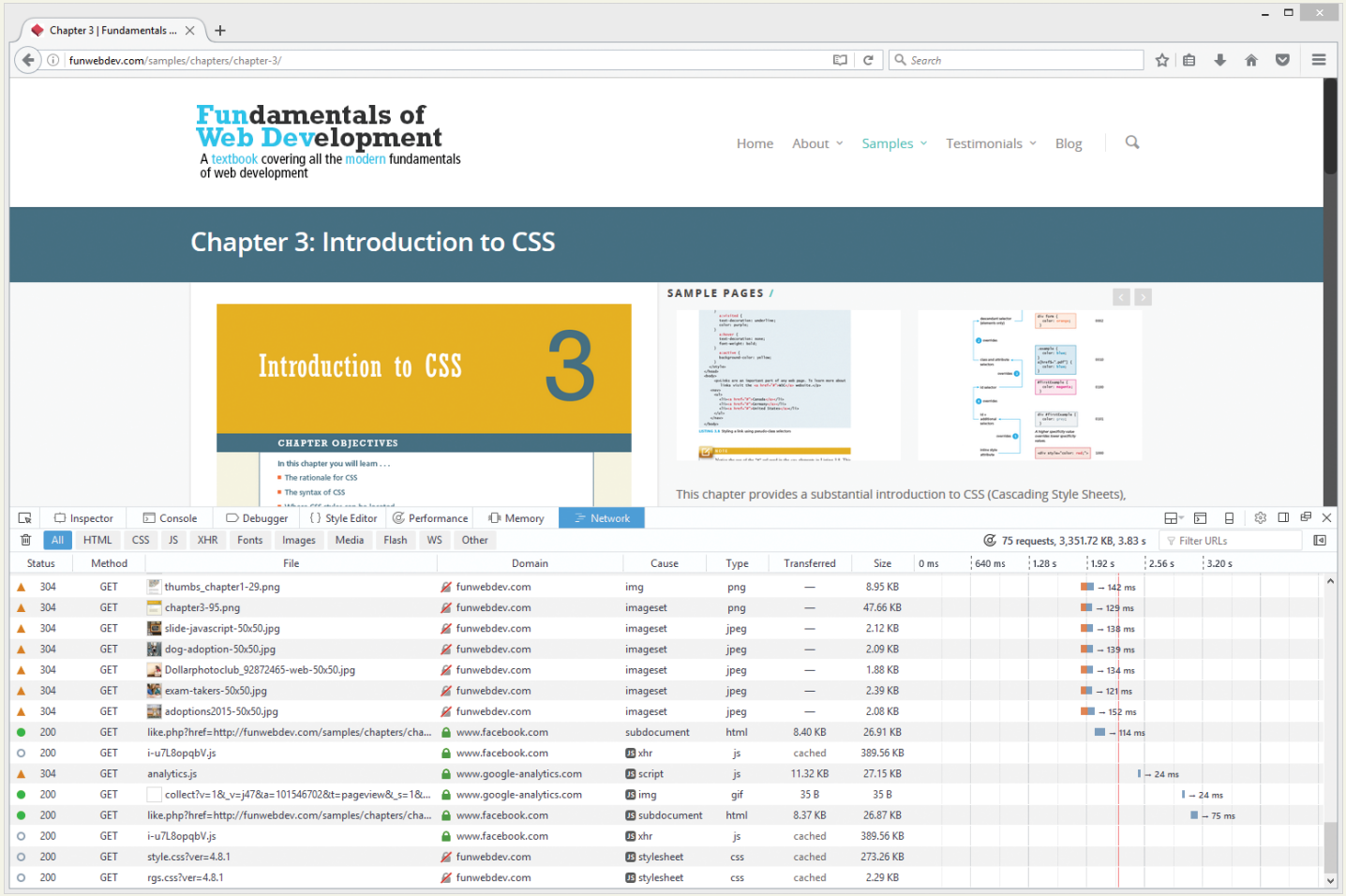
# Web Browsers

## Fetching a Web Page



# Web Browsers

## Fetching a Web Page – Load Times and Cascades



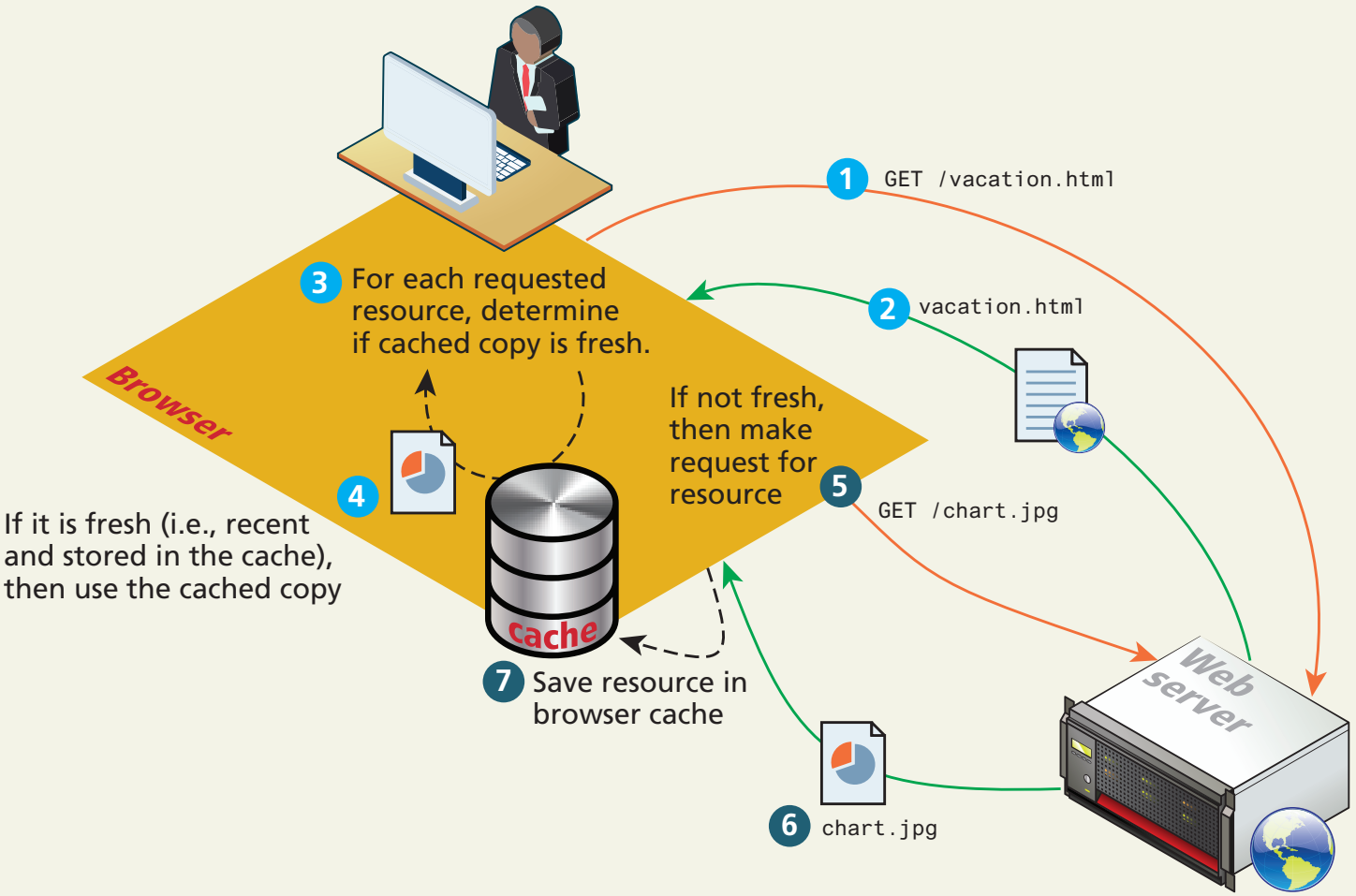
# Web Browsers

## Browser Rendering

- Interpreting the entire HTML markup together with the image and other assets into a grid of pixels for display within the browser window is called rendering the webpage.
- Implemented differently for each browser (Firefox, Chrome, Safari, Explorer, and Opera)

# Web Browsers

## Browser Caching



# Web Browsers

## Browser Features

- search engine integration,
- URL autocompletion,
- Form autocompletion,
- cloud caching of user history/bookmarks,
- phishing website detection,
- secure connection visualization,

and much more

# Web Browsers

## Browser Extensions

Can change what is shown to the end user. Newer challenge for web developers.

For developers, extensions like

- Firebug and
- Yslow

For the general public:

- Adblock
- Third Party Plugins

# Chapter 2

1 Internet Protocols

2 Domain Name System

3 Uniform Resource Locators

4 Hypertext Transfer Protocol

5 Web Browsers

6 Web Servers

7 Summary

# Web Servers

Operating Systems

- A **web server** is nothing more than a computer that responds to HTTP requests.
- Real-world web servers are often more powerful than your own desktop computer
- Webservers must choose an **application stack** to run a website. This application stack  will include an
  - operating system,
  - web server software,
  - a database,
  - and a scripting language for dynamic requests

# Web Servers

## Application Stacks

We will rely on the LAMP software stack ,which refers to

- L inux operating system,
- A pache web server,
- M ySQL database, and
- P HP scripting language

Other stacks include WAMP, WISA, MEAN, …

# Web Servers

## Operating Systems

- Linux
- Windows

# Web Servers

## Web Server Software

- Apache
- Nginx
- IIS

# Web Servers

## Database Software

- MySQL
- PostgreSQL
- Sqlite
- Oracle
- IBM DB2
- Microsoft SQL Server
- MongoDB

# Web Servers

## Scripting Software

- PHP
- ASP.NET
- Python
- Node.js
- …

# Chapter 2

1. Internet Protocols
2. Domain Name System
3. Uniform Resource Locators
4. Hypertext Transfer Protocol
5. Web Browsers
6. Web Servers
7. Summary

# Summary

## Key Terms

address resolution
Apache
Application stack
application layer
country code top-level domain (ccTLD)
DNS resolver
DNS server
domain names
domain name registrars
Domain Name System (DNS)
FTP
four-layer network model
generic top-level domain (gTLD)
GET request
HTTP
Internet Corporation for Assigned Names and Numbers (ICANN)
Internet Assigned Numbers Authority (IANA)
internationalized top-level domain name (IDN)
Internet layer
Internet Protocol (IP)
IP address
IPv4
IPv6
LAMP software stack
link layer
MAC addresses
MEAN software stack
packet
protocol
port
POST request
protocol
request
request headers
response codes
response headers
reverse DNS lookups
root name server
second-level domain
SFTP

# Summary

## Key Terms (Continued)

SSH

subdomain

TCP/IP (Transmission Control Protocol/ Internet Protocol)

transport layer

Transmission Control Protocol (TCP)

top-level domain (TLD)

TLD name server

User Datagram Protocol (UDP)

Uniform Resource Locator (URL)

web server

WISA software stack

# Questions?