World Wide Web

Introduction to networking

Dr. Klára Pešková, Klara.Peskova@mff.cuni.cz
Department of Software and Computer Science Education

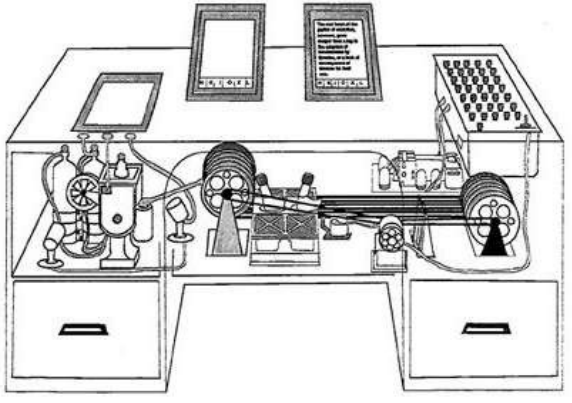
WWW – World Wide Web

- WWW is the most used Internet service
- Started as an experiment in CERN
- Nowadays…
  - Platform for information and data exchange
  - Environment for applications, that are accessible from anywhere
  - Social interactions
  - Shopping
  - Culture
  - Studying
  - Entertainment …

Ancient history

- 1945 – Vannevar Bush
  - Human brain works with associations
  - hypothetical electromechanical device Memex – "enlarged intimate supplement to one's memory", bookmark list of static microfilm pages
- ‘60s – Theodore Nelson first used the word hyper-text, i.e. text linked with associations
  - Project Xanadu
    - System for sharing information
    - Implemented as prototype

Differential Analyser, 1938, source: *University of Cambridge*



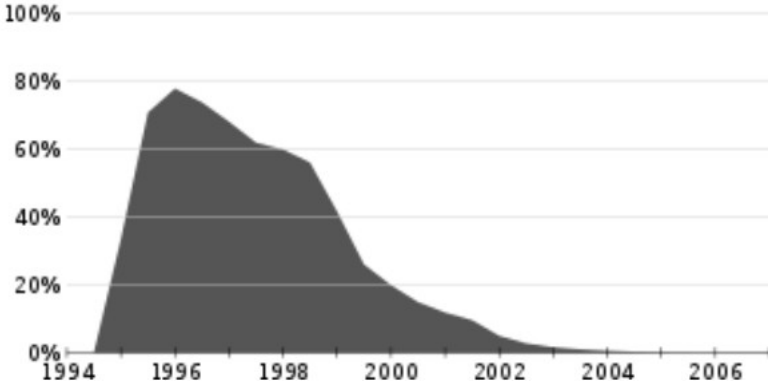
Memex, *„As We May Think“ (essay), The Atlantic magazine, 1945*

History (Middle Ages)

- Tim Berners-Lee – inventor of World Wide Web
  - 1980 – while in CERN he proposed a project based on the concept of hypertext to facilitate sharing and updating information among researchers
  - 1989 – he saw an opportunity to join hypertext with the Internet and implemented a server-client communication
  - He published the very first web page http://info.cern.ch/hypertext/WWW/TheProject.html
- NCSA Mosaic
  - 1993 – web browser by Marc Andreesen and Eric Bina
  - 1995 => Internet Explorer

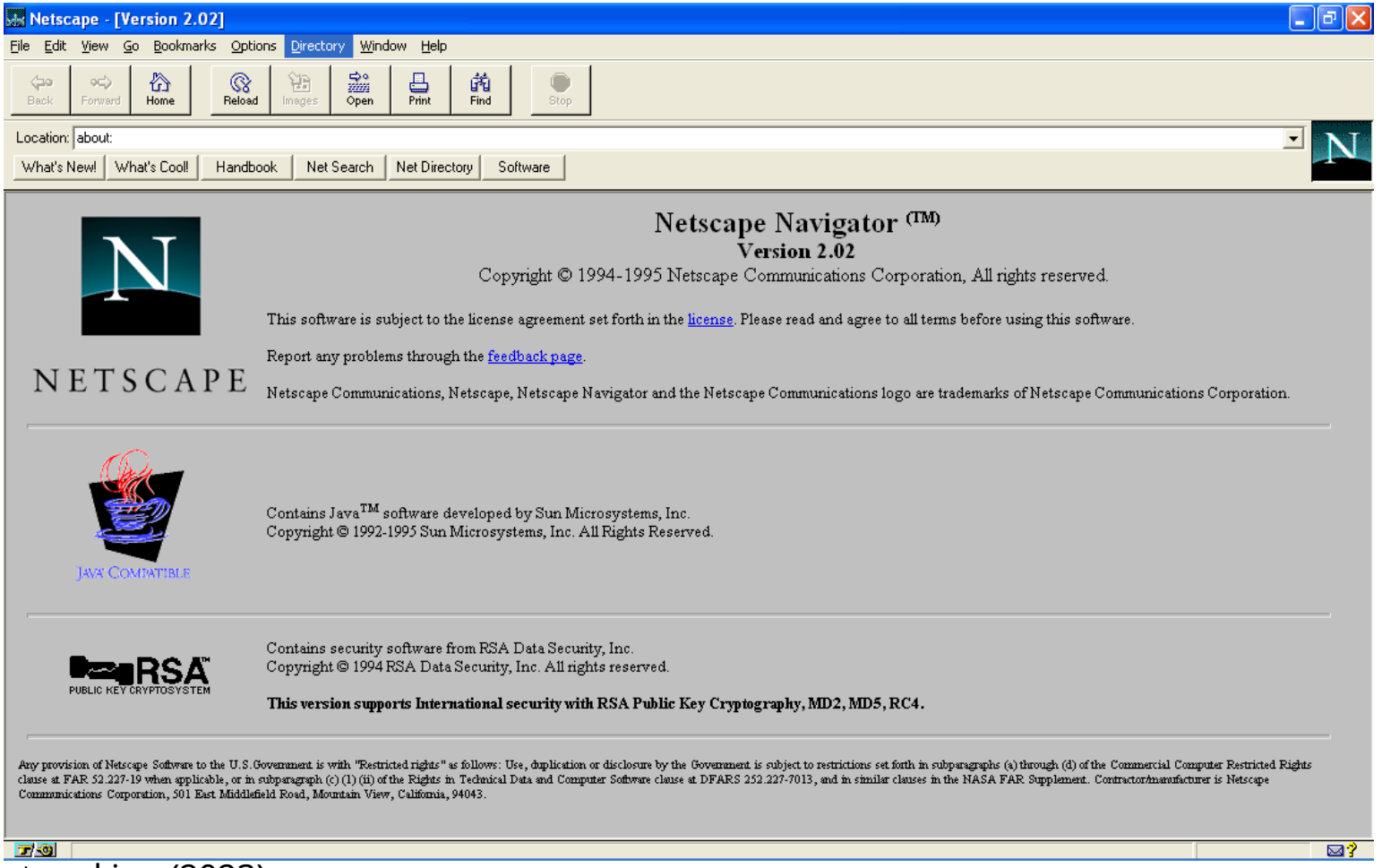
History (Modern age)

- 1996 – the war of browsers ☆
  - Internet Explorer vs. Netscape Navigator
- 1997 – HTML 4 published as a W3C Recommendation
- 2002 – The first idea of "Web 2.0"
  - The content of web is created by users
- 2004~2006 – introduction of AJAX applications
- 2010 – HTML5
  - Beginning of an end of Flash (Adobe will no longer update or maintain Flash after December 2020)

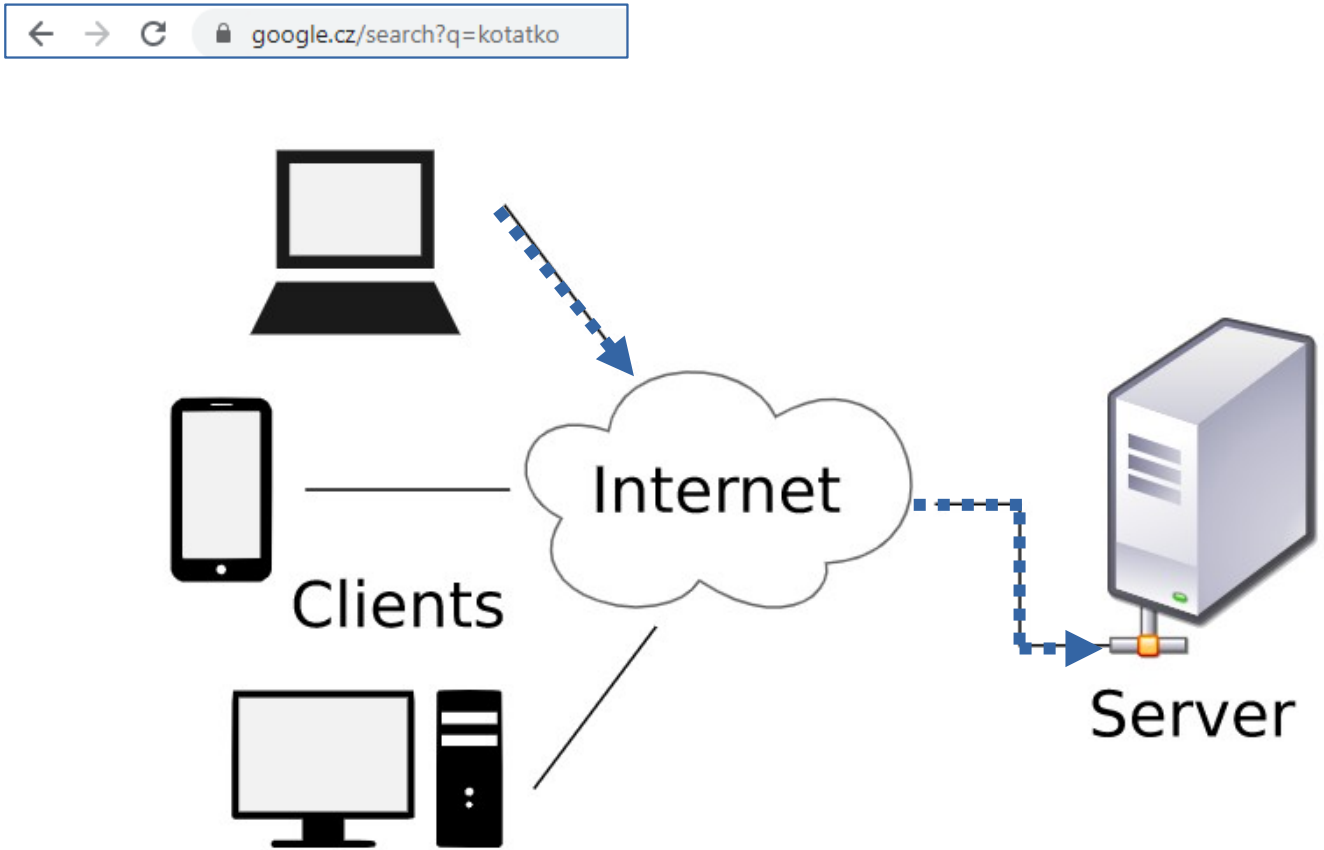


Usage share of Netscape Navigator

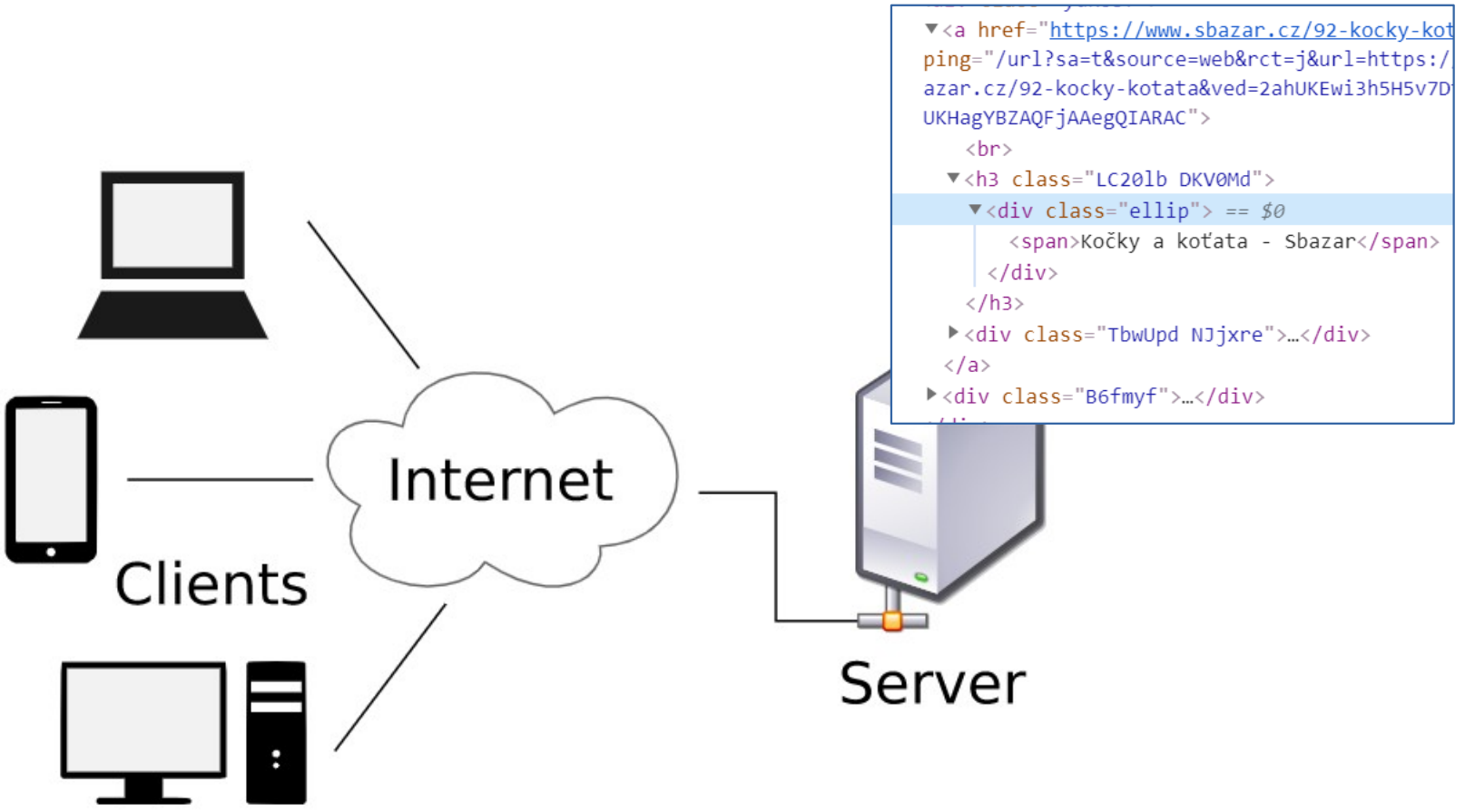
Netscape Navigator



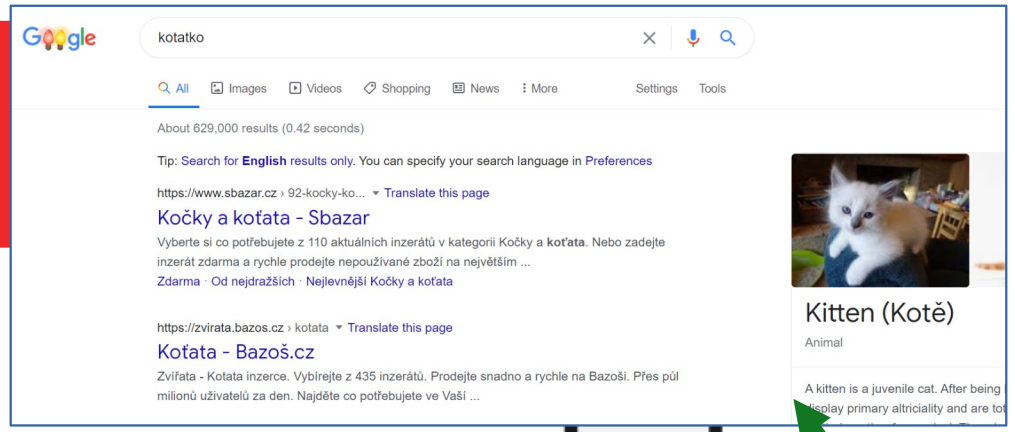
Displaying a web page – client-server architecture

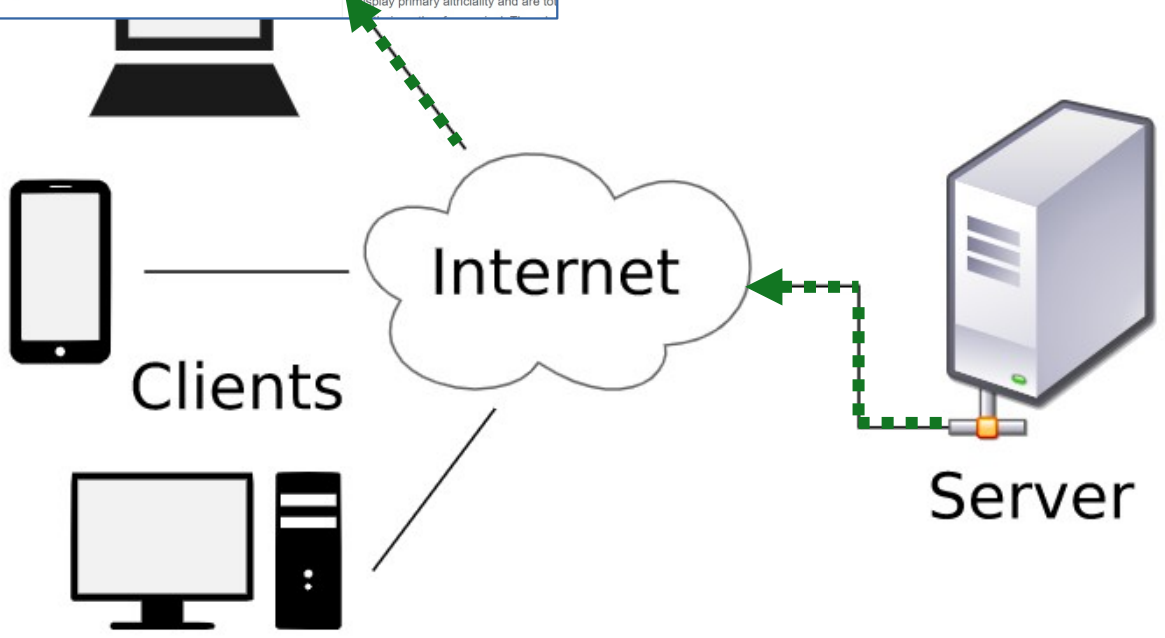


Displaying a web page – client-server architecture

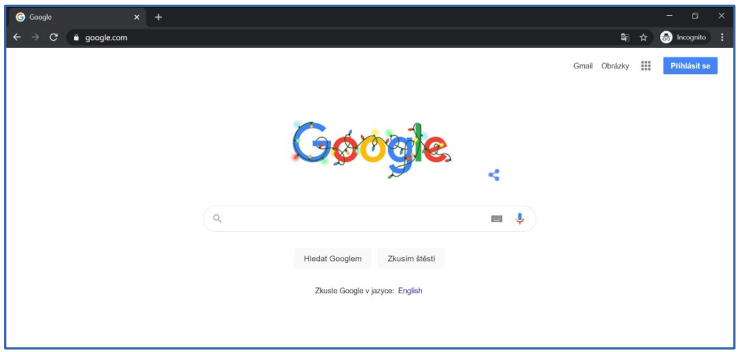


client-server





Web page




```
<!DOCTYPE HTML>
<html>
<head>
  <title>Google</title>
  <link rel="stylesheet"
        type="text/css"
        href="styles.css">
</head>

<body>
Text contents...

<a href="http://domain.cz">
  odkaz
</a>

<img src="image.png">

<video>
<source src="movie.mp4"
  type="video/mp4">
</video>
```

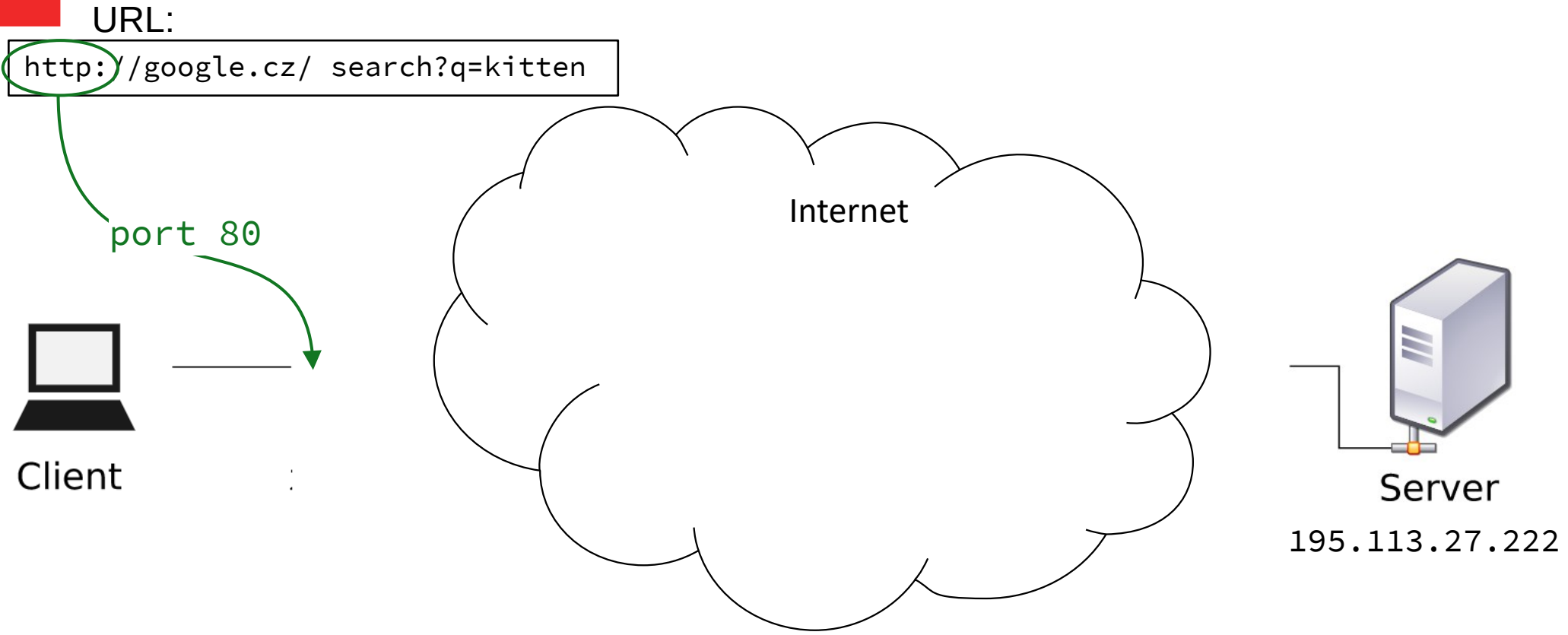
styles.css

```
body {
    Font: 12pt Calibri;
}
p {
    Margin: 10px;
}
```

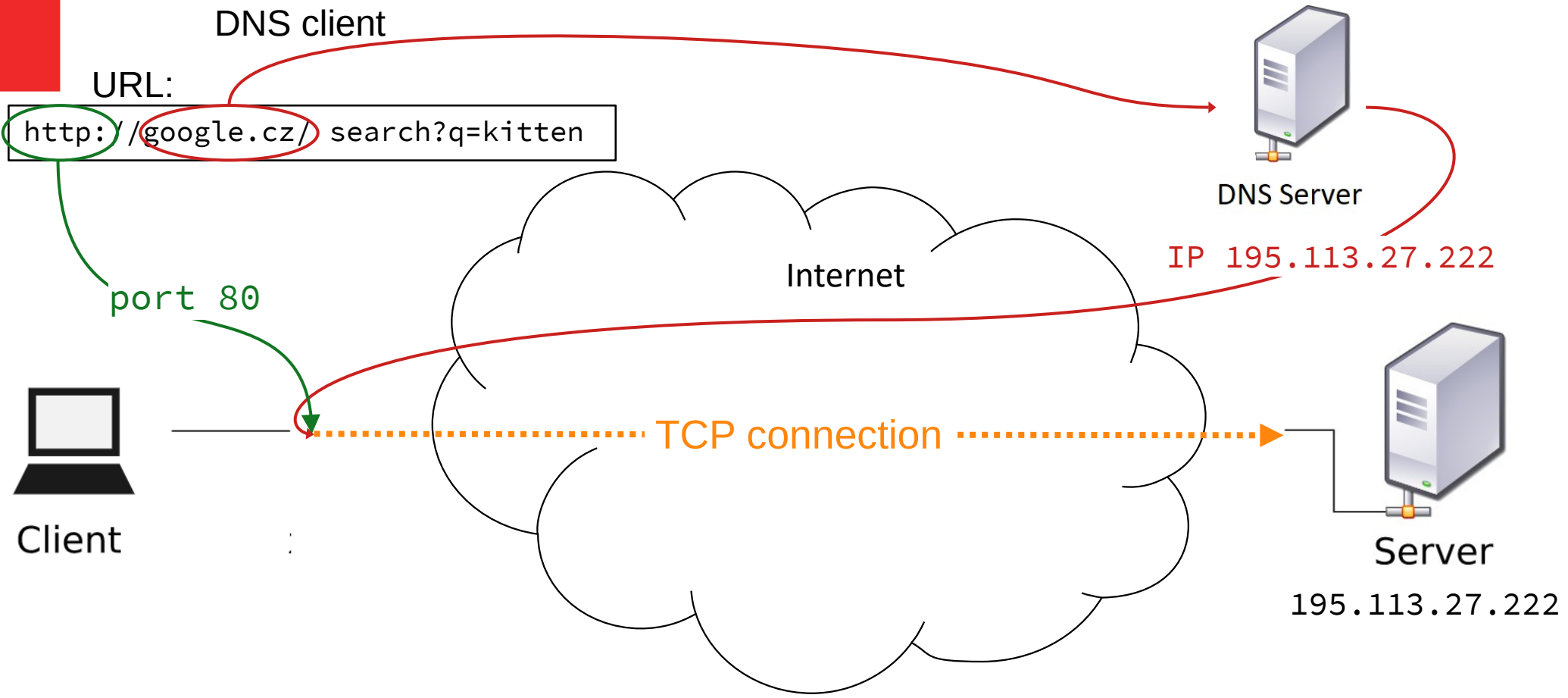
image.png

movie.mp4

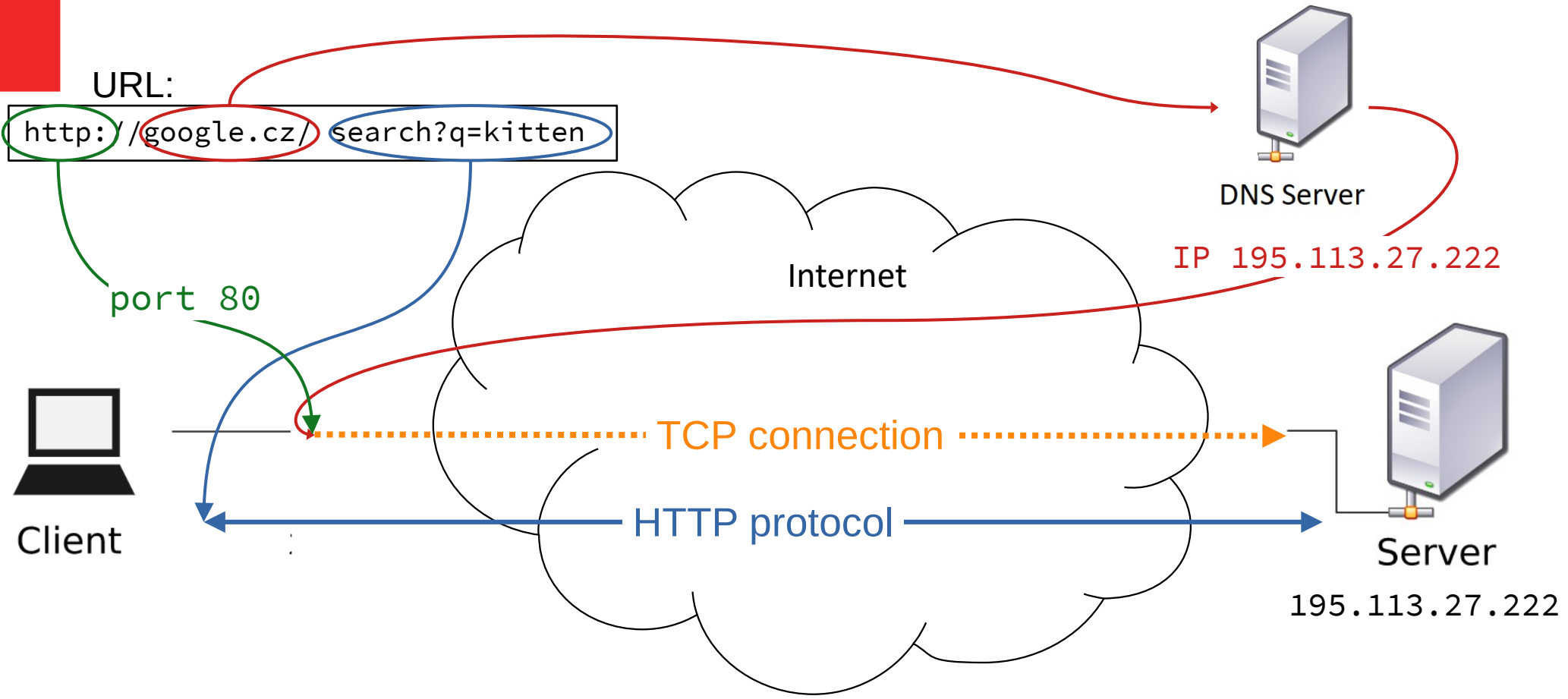
Requesting a web page – more details



Requesting a web page – more details



Requesting a web page – more details



URL (web site address)

- Uniform Resource Identifier (URI)
  - Identification string of the following format:

    `<schema>:<hierarchical_part>?<query>#fragment`
  - `query` and `fragment` parameters are optional
  - e.g. of URIs that are not URLs:

    `urn:isbn:0-476-27557-4`, `mailto:klara.peskova@mff.cuni.cz`
- Uniform Resource Locator
  - URI that describes a location of a resource

    ```
    protocol://username:password@domain:port/path
        ?parameter=value&parameter2=value2#element_id
    ```
  - e.g.: `http://is.cuni.cz/stud/schedule/teacher.php?teacherID=13270`

HTTP – Hyper-text transfer protocol

- Simple text-based protocol
- Uses TCP channel
- Designed for data retrieval
  - Originally only plain text data
  - Extended to support any data type and encoding (MIME)
- HTTP communication protocol
  - user (client) sends a HTTP Request
    - Specifying the details of the requested content
  - server replies with HTTP Response
    - (usually) containing the requested data

HTTP

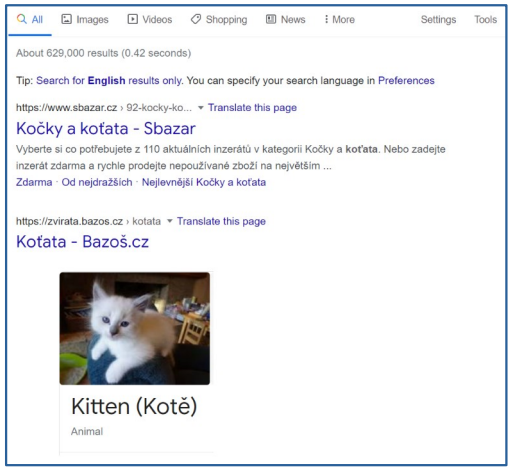
Client

TCP connection

Server

HTTP request

Headers (client's request), cookies, form data send by POST method



HTTP response

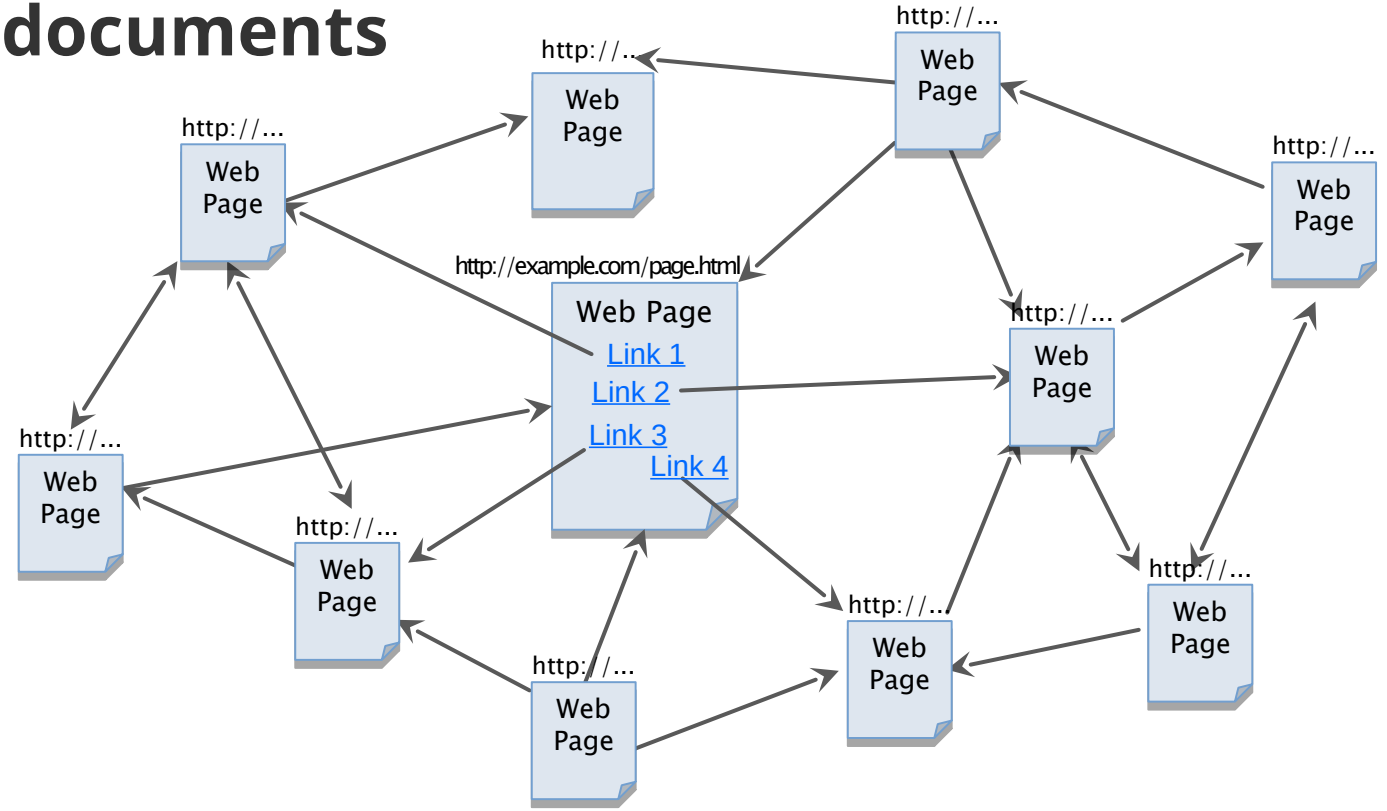
Headers specifying the response
+ the content that user requested (e.g. HTML file)

Static web page vs web application

- Static web page – server returns the requested document
  - WWW as Web of Documents
    - URLs are unique global identifiers of documents
    - Documents are linked by hyperlinks
  - HTML is a format for representing documents published on WWW

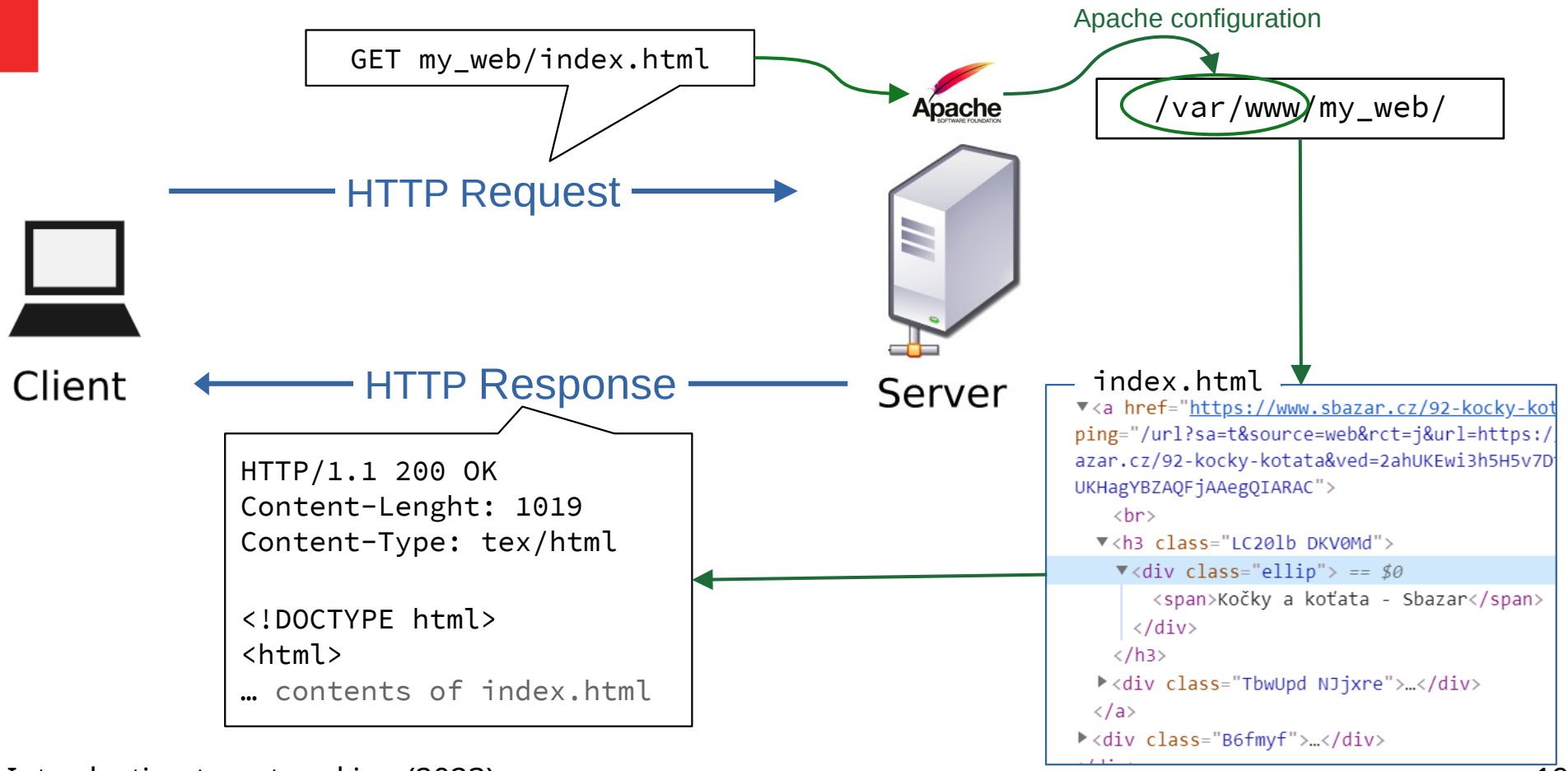
- Web applications – server generates the document dynamically based on user's request
  - HTML and CSS – application's UI presentation format

Web of documents



- Static web pages – HTML and CSS
- To view a static web page, you can use your browser, server is needed to publish the page on the internet

Web server – serving a static web page

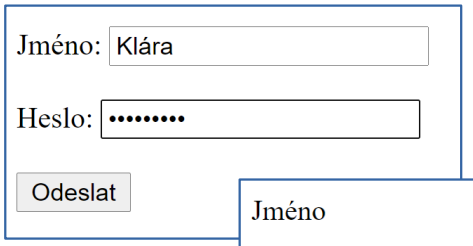


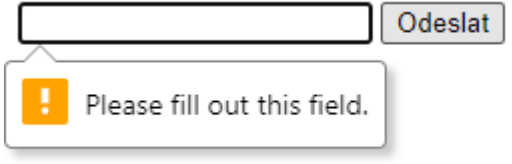
Web application

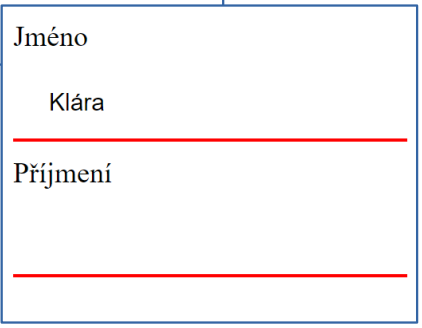
- Interactive elements
  - Hyperlinks – a way to change an application view
  - Form controls – user input data
    - Can be submitted and processed at the server
- Dynamic web pages
  - Server-side
    - The page is generated on demand
  - Client-side
    - Script running inside a browser can handle user events and modify the loaded HTML document

HTML Form

- Component of web page composed of UI controls
  - Text fields of various types (input)
  - Select boxes
  - Radio buttons
  - Check boxes
  - File uploads
  - Buttons
- Form components are a user-friendly version of hyperlinks
  - Sends user data and receives new web page
  - HTTP method can be selected (GET, POST)
- Data are sent when a submit button is hit



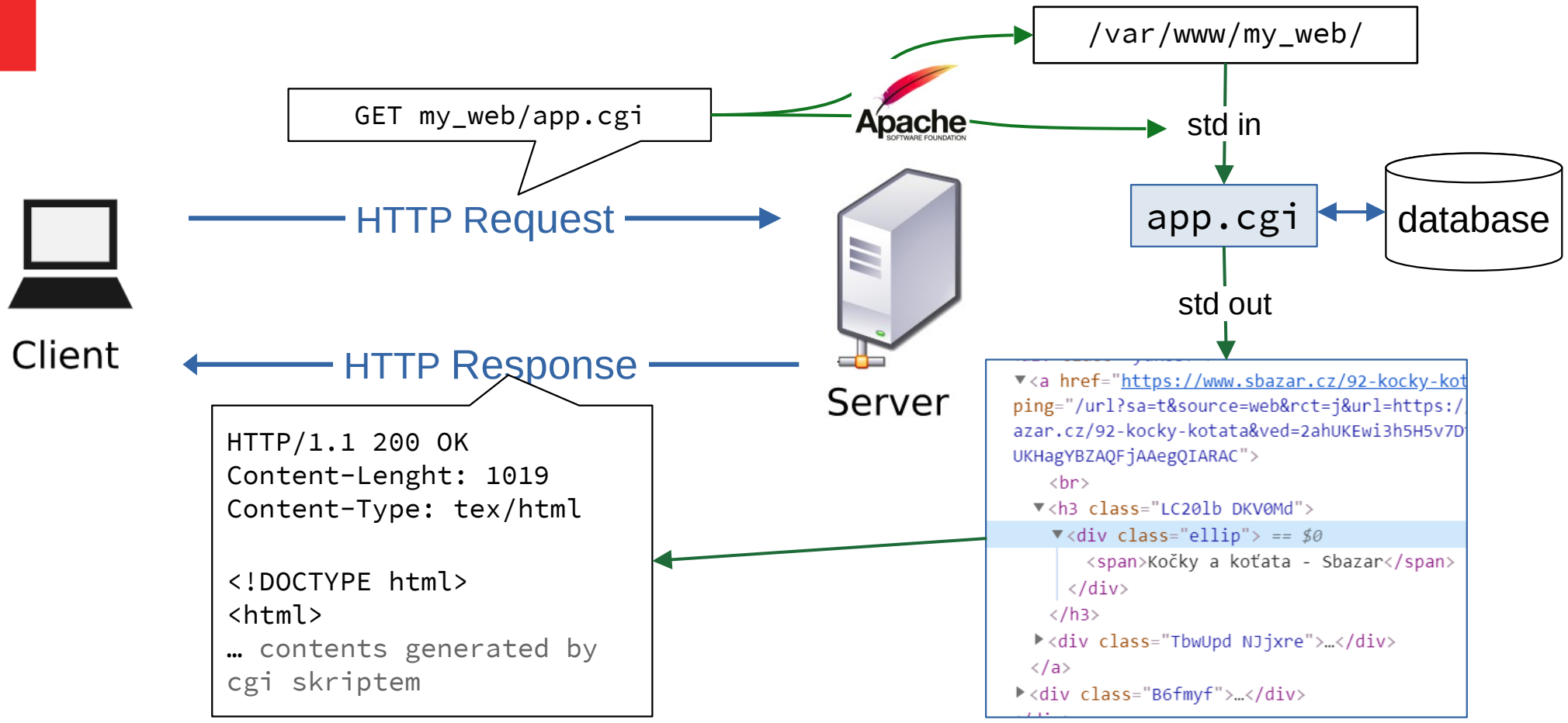




Web server – serving dynamic web page

- Common Gateway Interface (CGI)
  - One of the first standards for generating dynamic web content
  - Interface specification for web servers to execute programs like console applications that generate web pages dynamically
  - Interface specification only
    - The console application may be written in any programming language
    - Important information and headers are set as environment variables, POST body is directed to std. input
    - Response is taken from the std. output

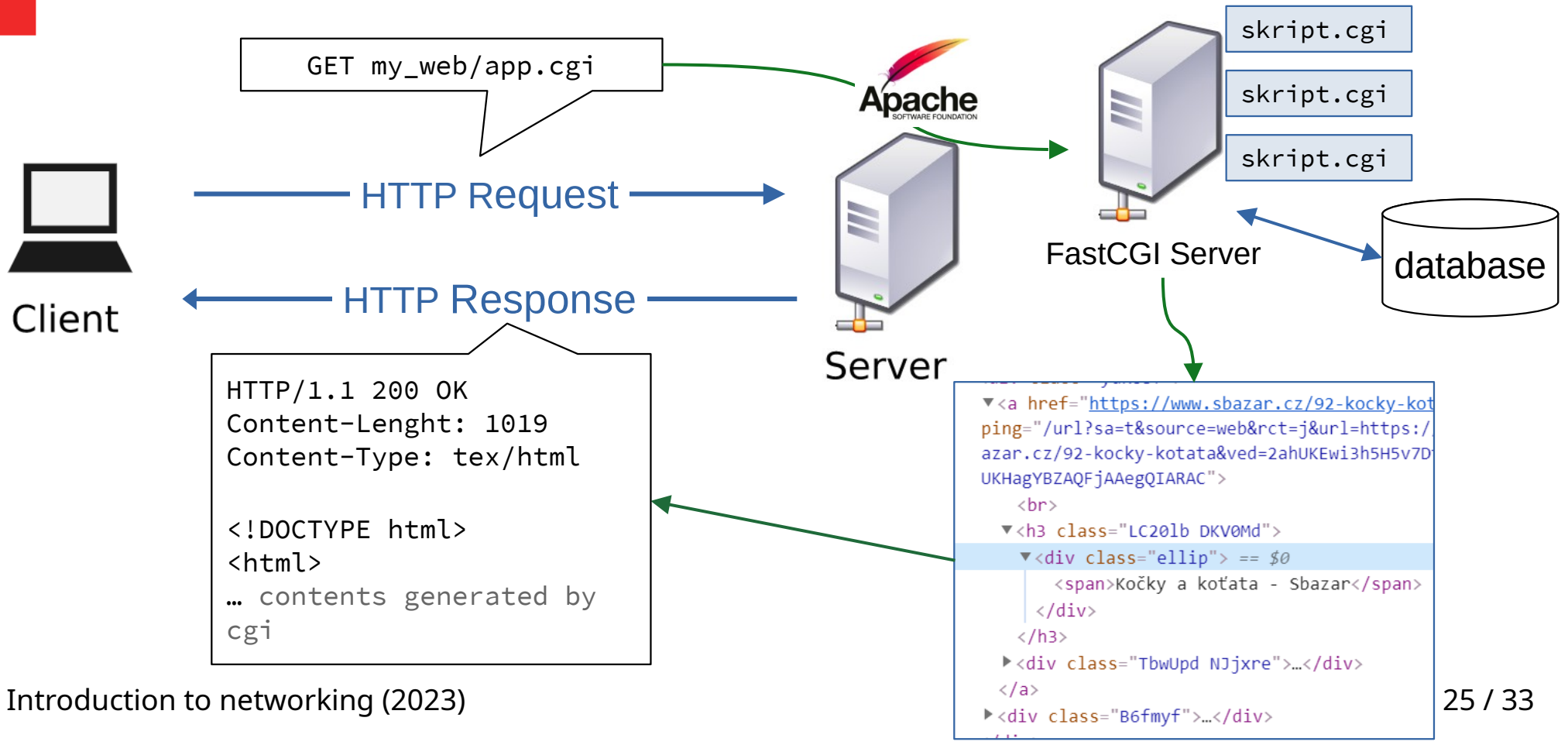
Web server – serving dynamic web page using CGI



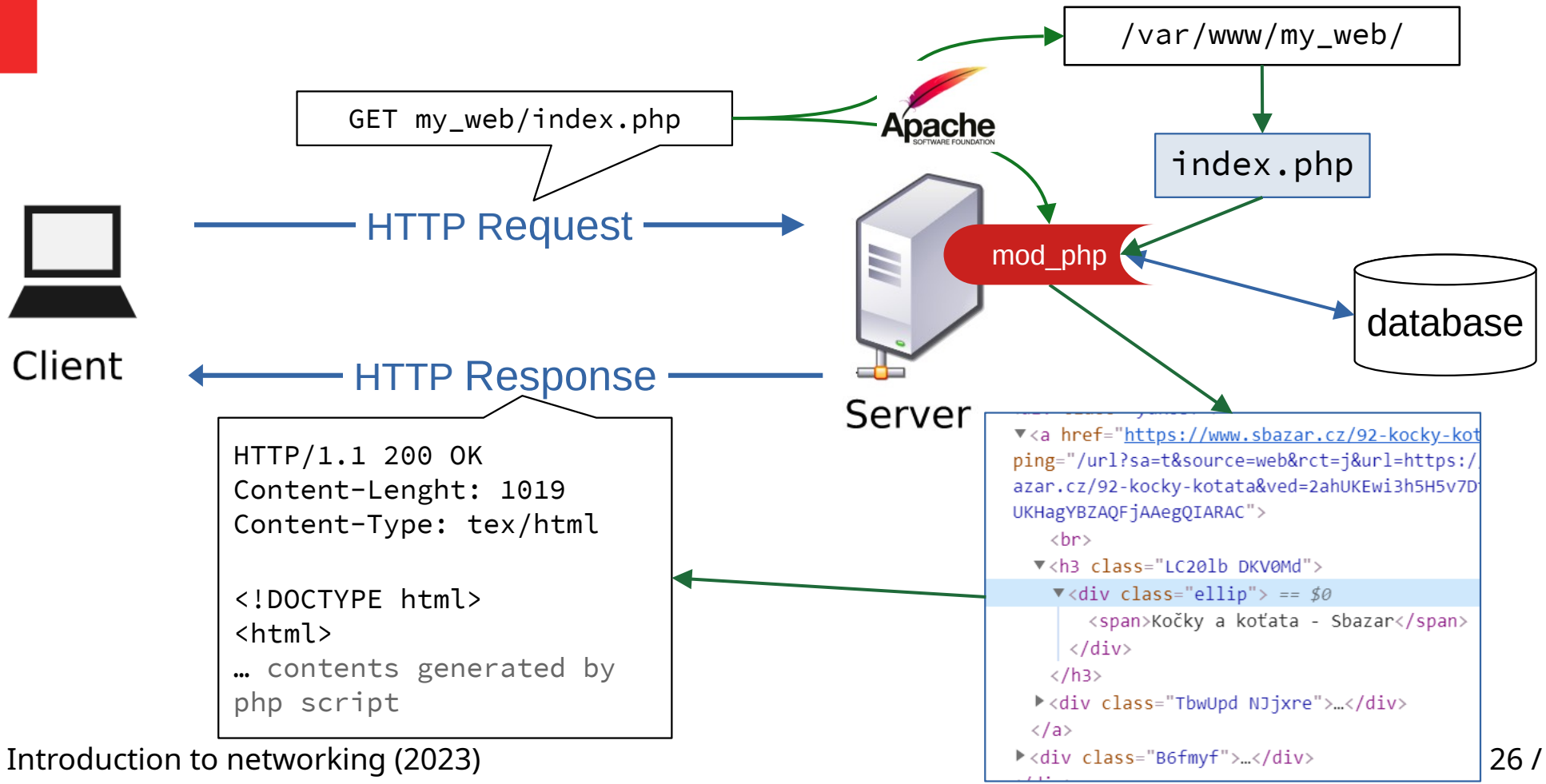
Web server – serving dynamic web page – three approaches

- FastCGI
  - CGI processes are already started and waiting for requests
  - More effective interface
- Scripting language module
  - A special module is integrated into HTTP server
  - The module interprets script files and serves their outputs
- Application with embedded HTTP server (as a library)
  - Dedicated web server designed for the specific application

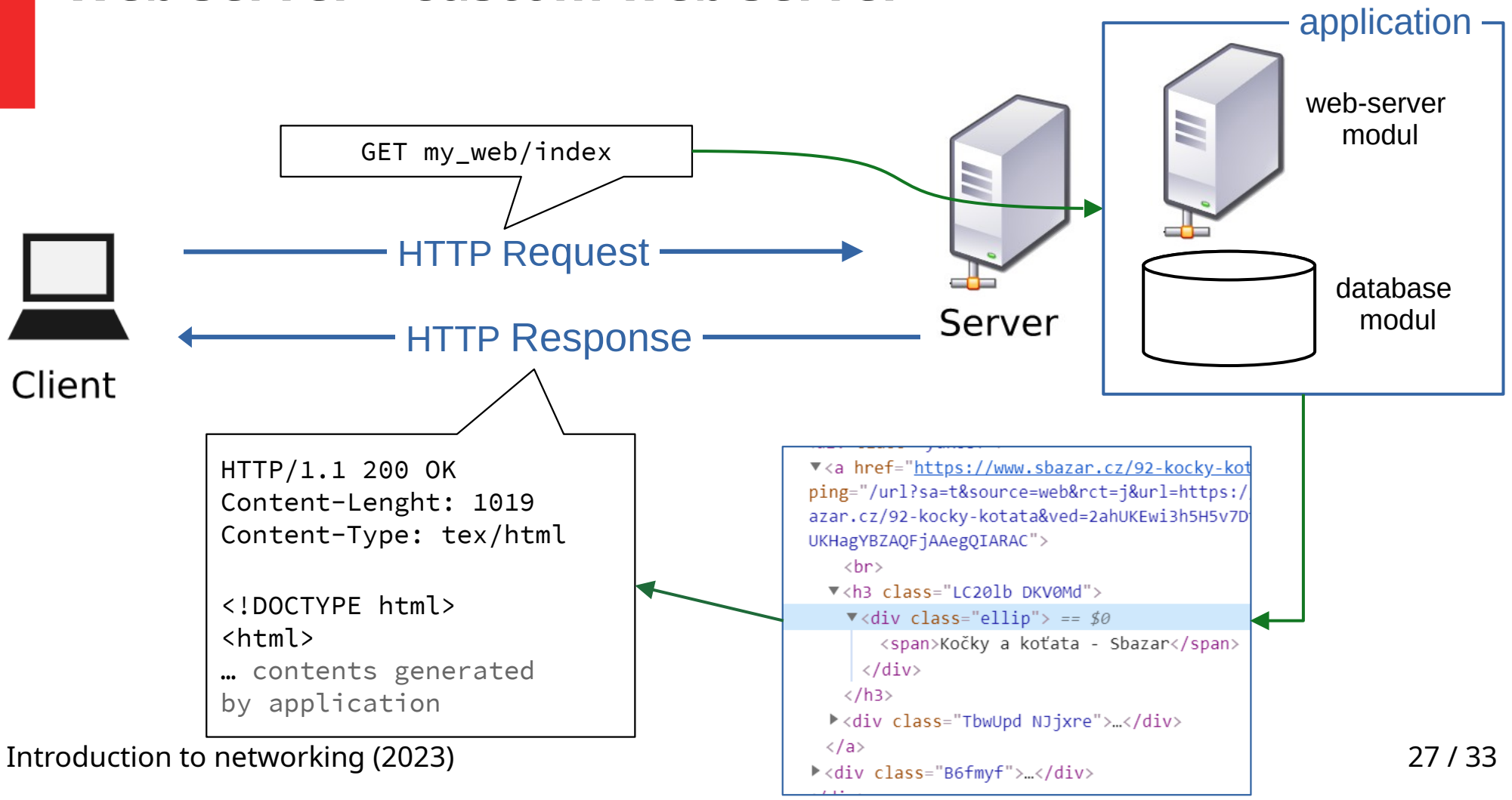
Web server – FastCGI



Web server – scripting language module (PHP)



Web server – custom web server



Client – dynamic page

- JavaScript
  - Developed by Brendan Eich in Netscape as a scripting language for web browser (early ‘90s)
    - c-like syntax
    - Dynamic, weakly typed
    - Object-oriented with prototyping and functional aspects
  - Named after Java for marketing reasons

Client – dynamic page

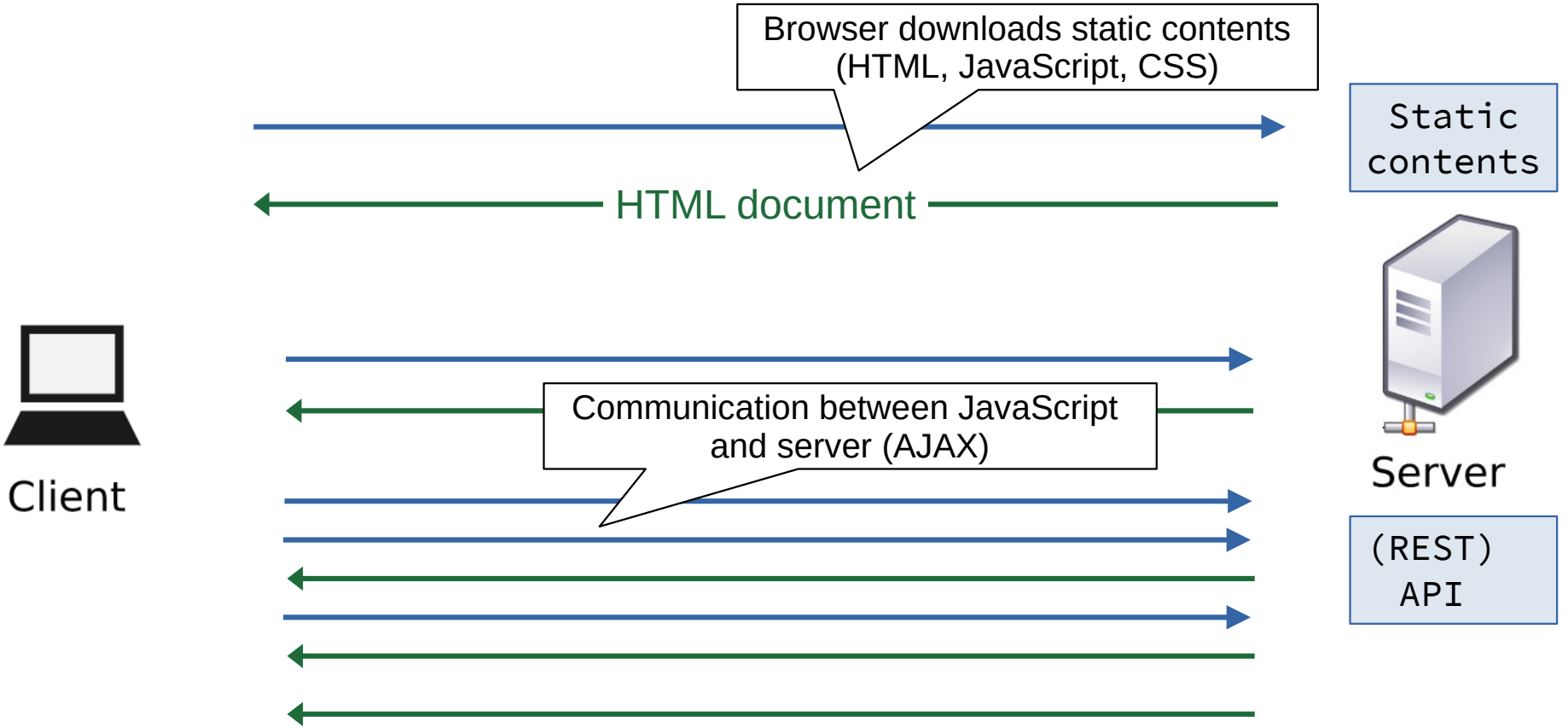
- JavaScript in a browser
  - Sandboxed, limited interactions with user
  - DOM (Document Object Model) API
    - Allows manipulating loaded HTML document and CSS
    - Modifications are immediately visualized
    - Event model for handling user actions
  - Other APIs
    - Browser window control, history, location (URL)
    - Animations, 3D graphics, sounds
    - Data storage (inside browser)
    - Networking (HTTP requests, WebRTC)

AJAX (/ˈeɪdʒæks/) – Asynchronous JavaScript and XML

- A technique that combines three technologies
  - JavaScript
  - Asynchronous HTTP client API integrated in browser
  - XML or other semi-structured data format

- Script invokes HTTP transfer – without the need to refresh the web page
- Requests are asynchronous, response data are processes using callbacks

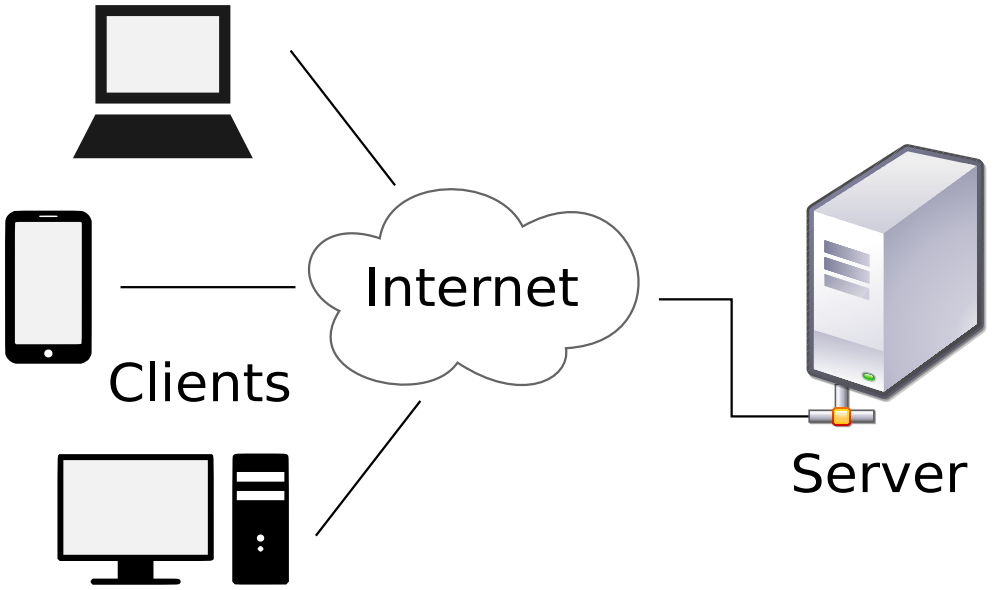
Web application – architecture



Web technologies

Client

- HTML
- CSS
- JavaScript
- Flash



Server

- PHP, Python, ...
  - Sending e-mails
  - Processing forms data
- database

- JavaScript (AJAX - Asynchronous JavaScript and XML)
- ~~Flash~~

Test yourself

- Explain how a web page gets to your browser.
- What is a URL? What parts does it consist of? What is their meaning?
- Explain a communication between client and server. What protocol do they use to communicate?
- What is the difference between static and dynamic web page?
- What are the two approaches to create a dynamic web page?
- Which web technologies run in a client and which are server-based?
- What is the most important feature of applications using AJAX?