Увод у веб и интернет технологије

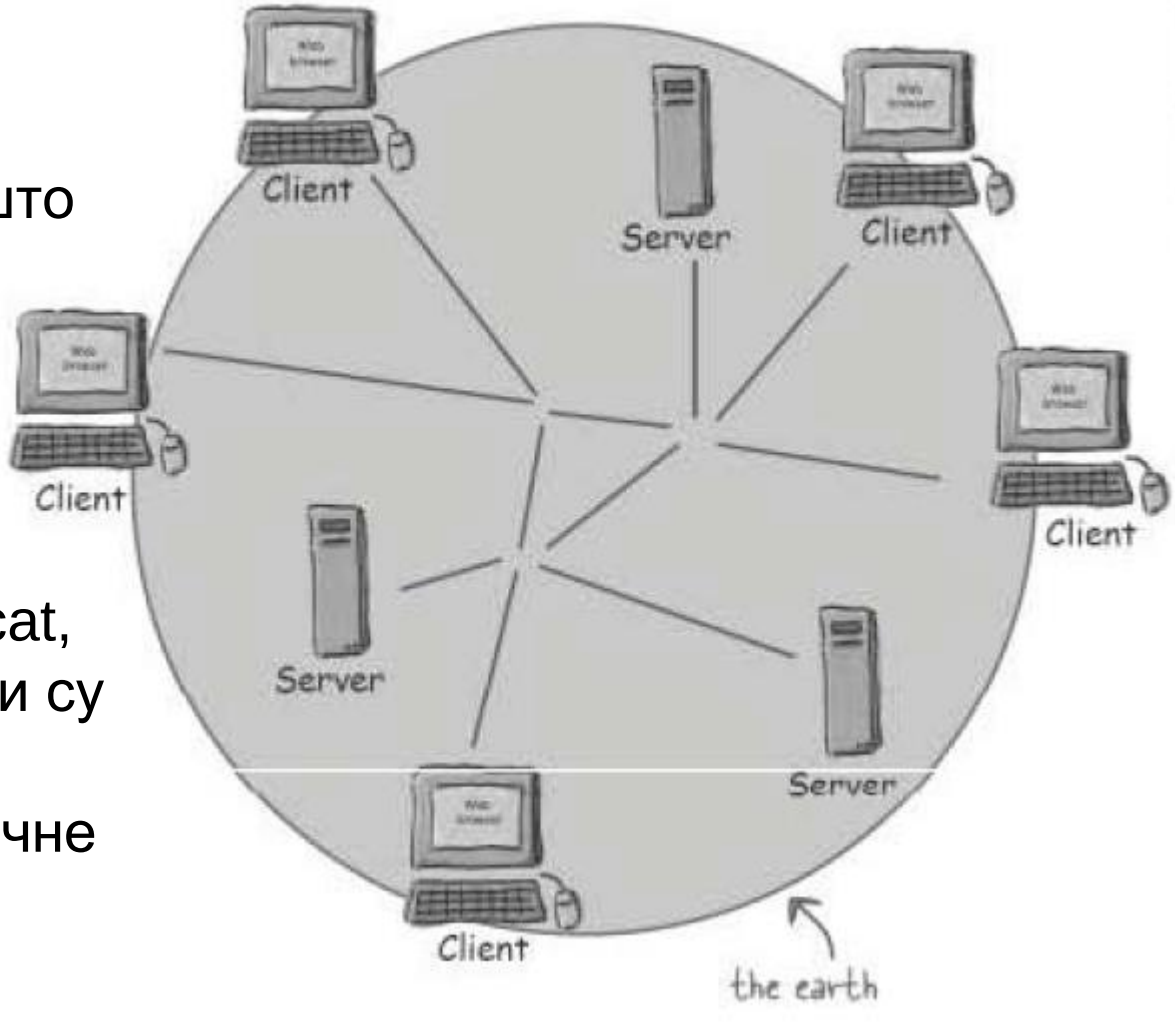
Математички факултет

vladaf@matf.bg.ac.rs

Веб сервери

Веб

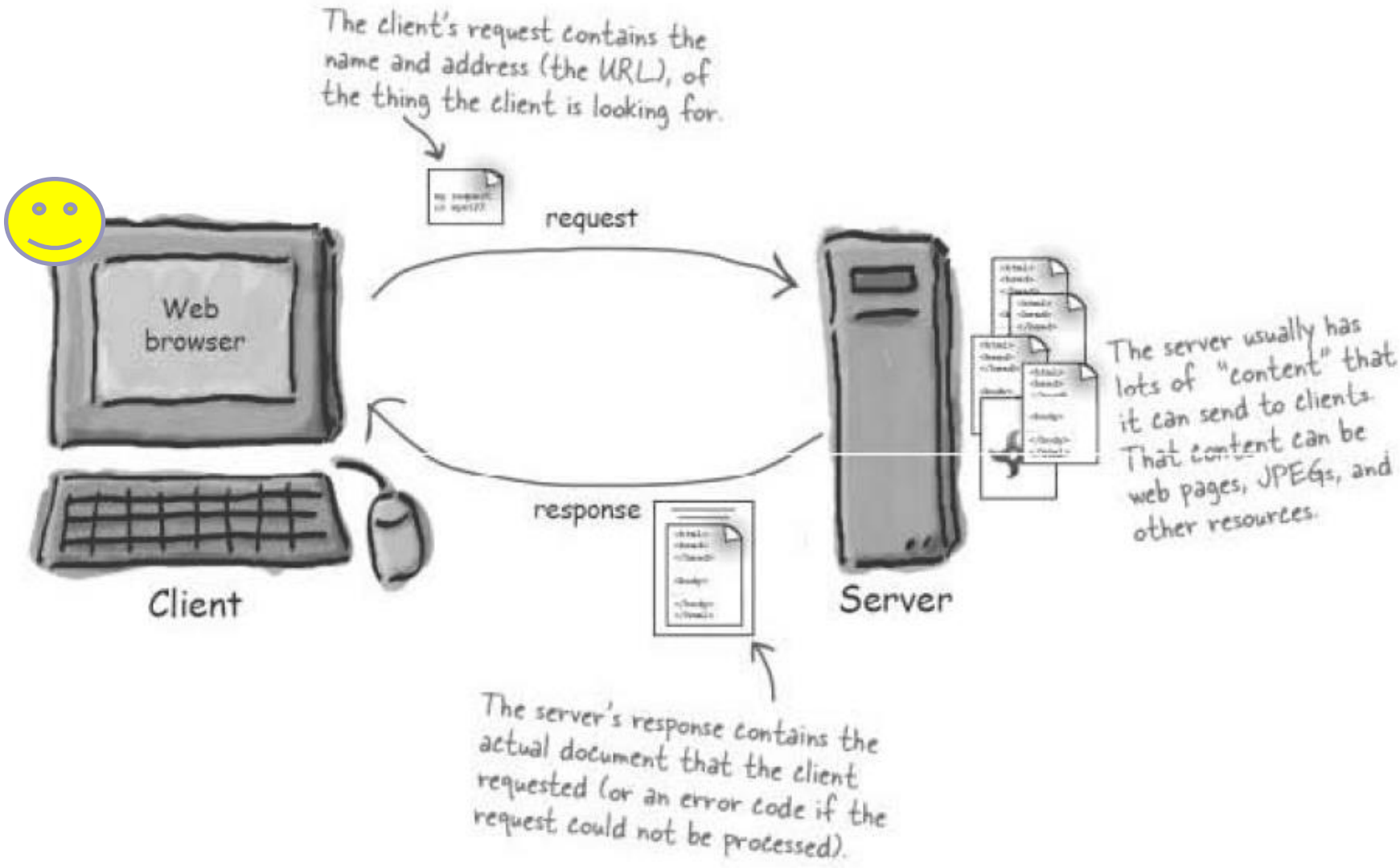
- Веб се састоји од огромног броја клијената са прегледачима (као што су Chrome, Mozzila, Yandex, Safari итд.) и од сервера (који користе веб сервере као што су Apache, JBoss, Tomcat, Microsoft IIS итд.) који су међусобно повезани кроз жичане и безжичне мреже.



Математички факултет

vladaf@matf.bg.ac.rs

Функционисање веб сервера



Математички факултет
vladaf@matf.bg.ac.rs

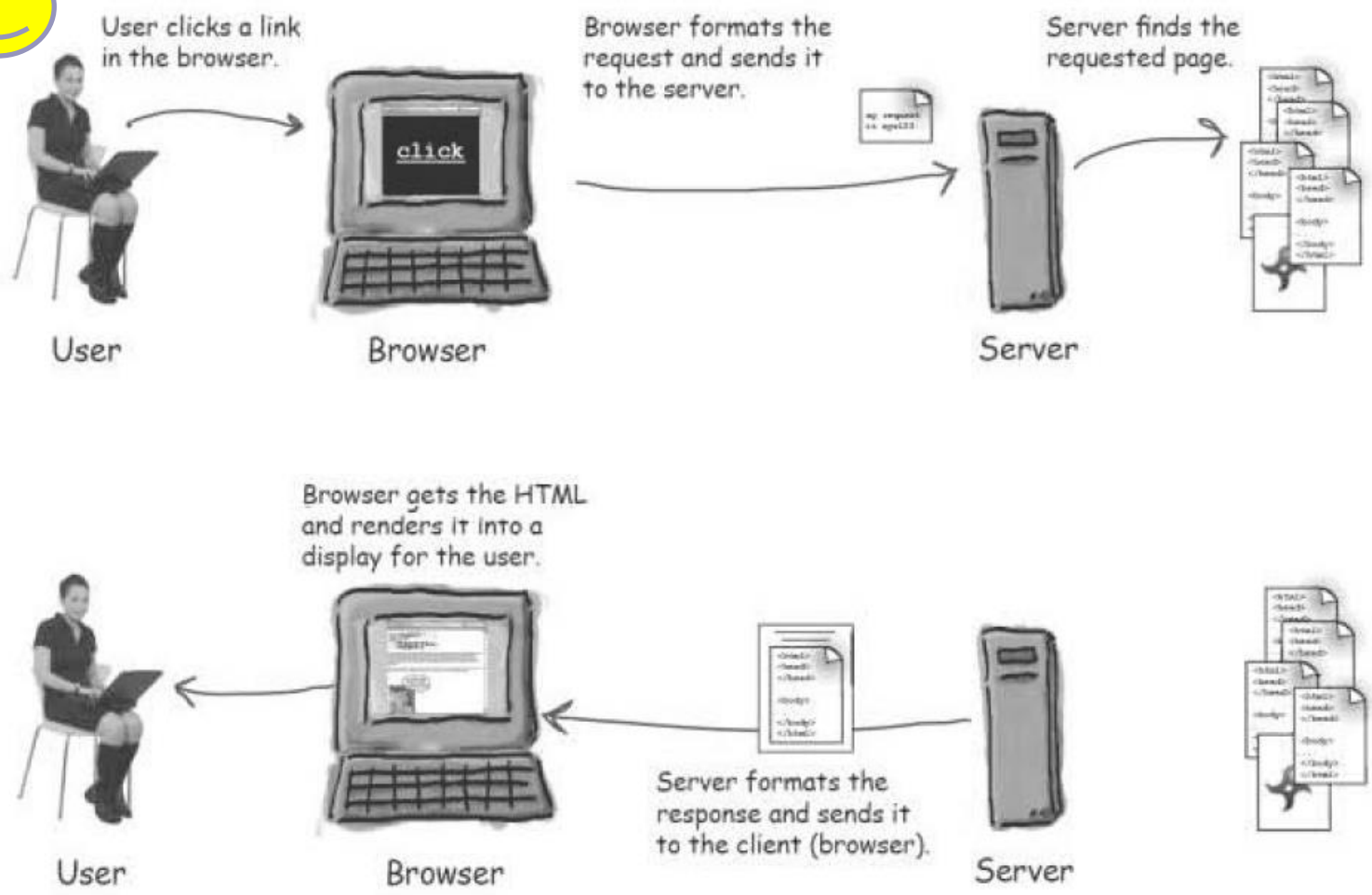
Функционисање веб сервера (2)

- Корисник преко веб прегледача шаље захтев за ресурсом
- Веб сервер прихвата захтев, проналази тражени ресурс и њега шаље кориснику
  - Ресурс може бити HTML страна, слика, PDF документ или нешто четврто – што год да је у питању, клијент захтева ресурс, а сервер шаље клијенту ресурс који је захтеван
  - У случају када нема захтеваног ресурса, генерише се грешка „404 Not found“
- Термин „сервер“ означава и сам рачунар (тј. хардвер) и програм који пердставља веб сервер (тј. софтвер)
  - Ако из контекста није јасно да ли се ради о хардверу или софтверу, то ће бити додатно истакнуто

Математички факултет

vladaf@matf.bg.ac.rs

Функционисање веб клијента



Математички факултет
vladaf@matf.bg.ac.rs

Функционисање веб клијента (2)

- Када се говори о клијенту, има се у виду корисник, али и веб прегледач – апликација коју корисник користи за сурфовање. Корисник преко веб прегледача шаље захтев за ресурсом
  - Дакле, прегледач је софтвер (нпр. Netscape, Chrome, Mozzila, Yandex, Safari, Edge, Opera и сл.) који комуницира са веб сервером. Осим послова комуникације, прегледач треба да интерпретира HTML код и да исцрта веб стране за корисника
- Ако се експлицитно не наведе другачије, надаље ће термин „клијент“ ће обухватити и софтвер (тј. прегледач) и човека (тј. корисника)
  - Другим речима, клијент је апликација - прегледач која обавља оно што корисник захтева да се уради

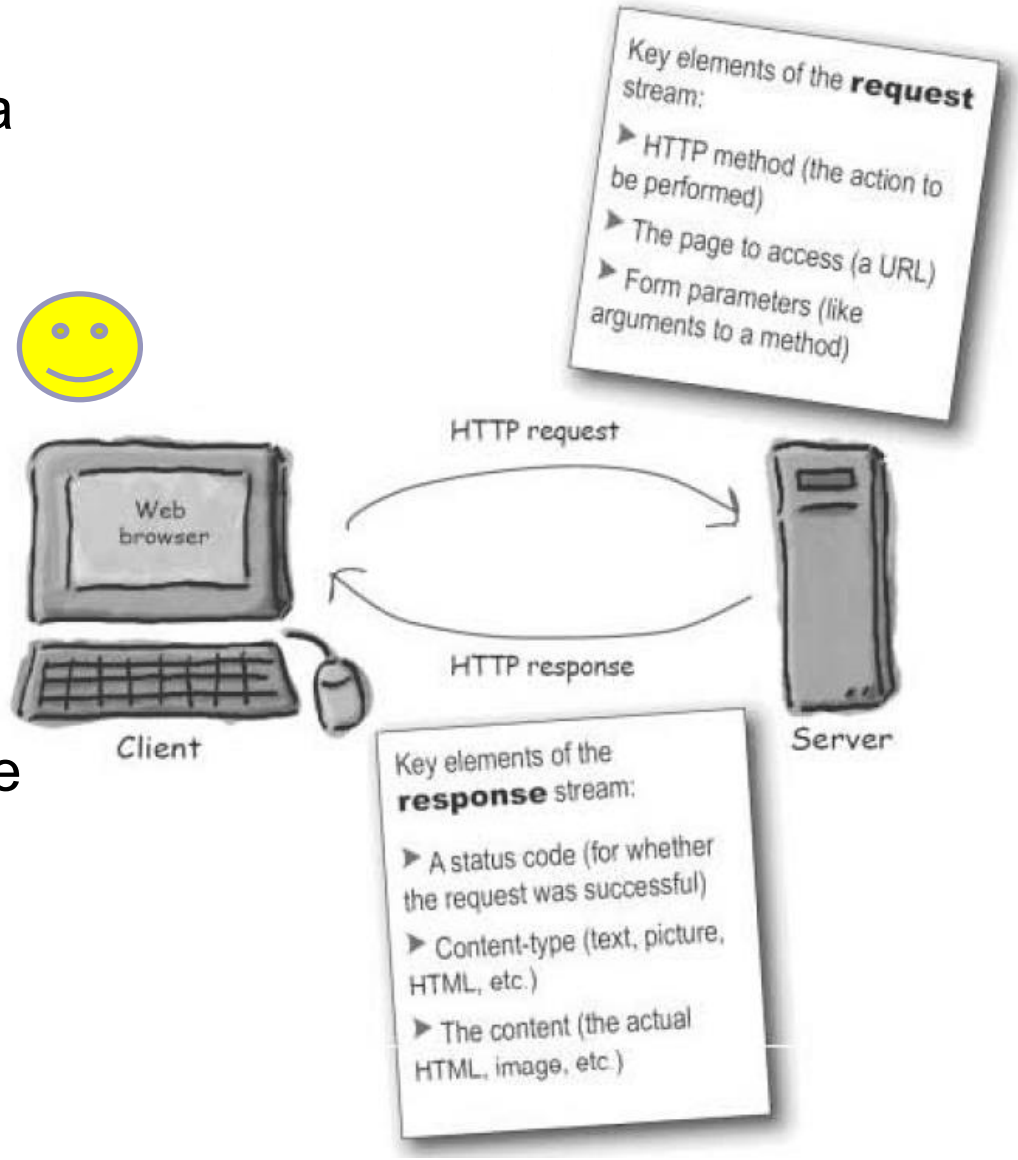
Математички факултет
vladaf@matf.bg.ac.rs

Клијент и сервер користе HTML и HTTP

- Када сервер шаље одговор на захтев, он прегледачу обично шаље садржај датог типа, тако да прегледач може да прикаже добијени одговор.
- Често је одговор који сервер шаље клијенту секвенца знакова који представљају документ у HTML формату. Тај HTML документ потом бива приказан од стране прегледача
  - Језик за означавање HTML (прецизније, HTML 5), је описан у претходним предавањима
- Највећи део конверзације између клијената и сервера се реализује коришћењем HTTP протокола
  - Клијент тада шаље HTTP захтев, а сервер одговара са HTTP одговором.
  - Када сервер пошаље HTML страну клијенту, он то ради коришћењем HTTP протокола.

HTTP протокол

- HTTP протокол се извршава преко TCP/IP протокола
- То је мрежни протокол са карактетистикама које се односе на веб, али он се ослања на TCP/IP протокол ради обезбеђења потпуног преноса захтева и одговора са једног места на друго
- Суштина HTTP конверзације је једноставна секвенца захтев/одговор: прегледач захтева а сервер одговара



Математички факултет
vladaf@matf.bg.ac.rs

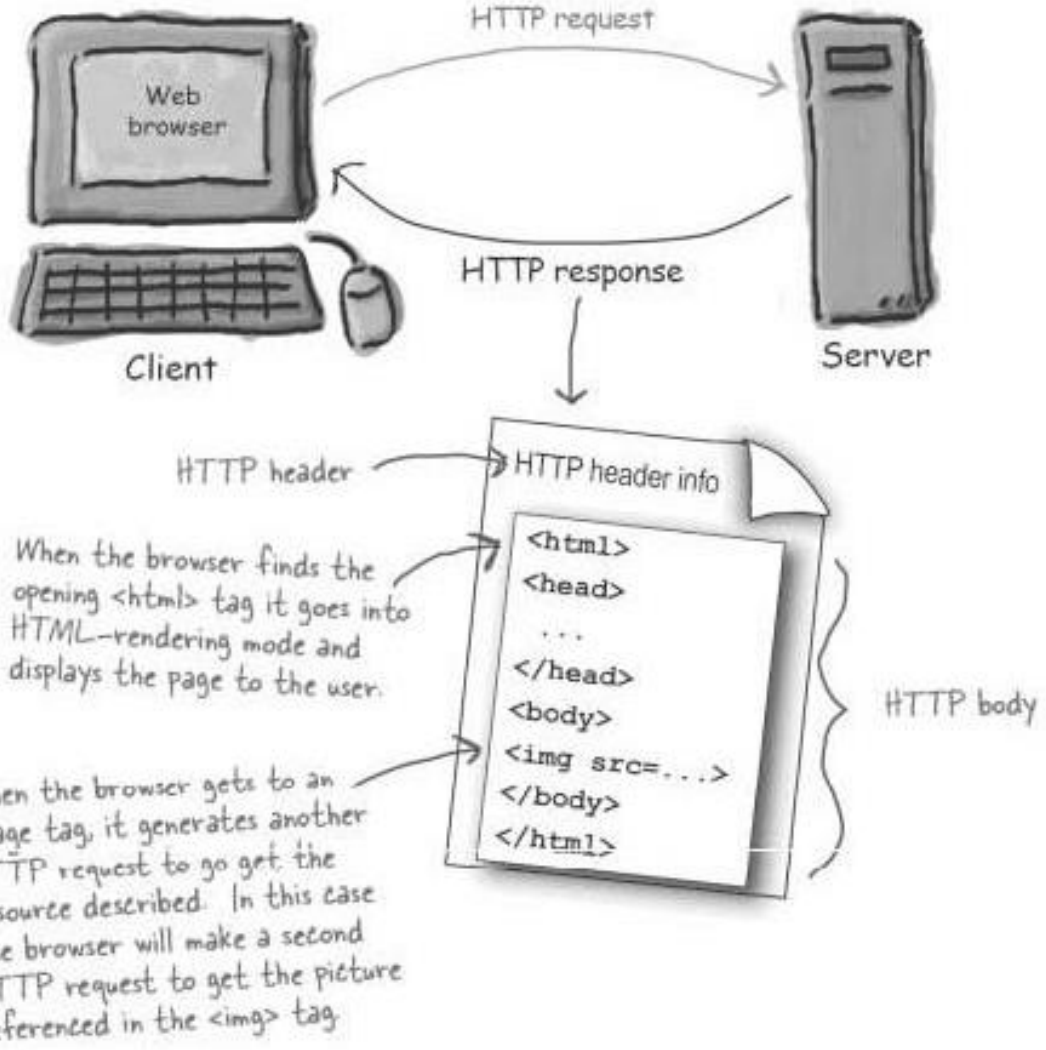
HTTP протокол (2)

- Сам HTTP протокол је описан IETF документом RFC 2616
- Веб сервери Apache, JBoss, Tomcat, Microsoft IIS и сл. су примери сервера који обрађују HTTP захтеве
- Прегледачи Netscape, Chrome, Mozzila, Yandex, Safari, Edge, Opera и сл. обезбеђују кориснику да генерише HTTP захтев, упути га према серверу и на адекватан начин прикаже HTTP одговор који добије од сервера
-
- Карактеристике HTTP протокола:
  1. HTTP не одржава конекцију (connectionless)
  2. HTTP је независтан од медијума (media independent)
  3. HTTP не подржава стања (stateless)

Математички факултет
vladaf@matf.bg.ac.rs

HTML је део HTTP одговора

- HTTP одговор може садржавати HTML
- HTTP додаје информације о заглављу на почетак садржаја који се враћа као одговор, какав год садржај био у питању
- Прегледач користи информације из заглавља као помоћ у процесирању HTML садржаја
- Дакле, HTML се може посматрати као садржај уметнут у HTTP одговор



Математички факултет
vladaf@matf.bg.ac.rs

HTTP метод

- HTTP захтев садржи назив метода у свом заглављу
- Назив метода говори серверу о каквој се врсти захтева ради и како ће бити форматиран остатак поруке
- HTTP протокол подржава следеће методе:
  - GET - користи се за преузимање информација са датог сервера на основу дате адресе. Захтеви који користе метод GET треба само да прибављају податке, а никако не треба да их мењају
  - HEAD - је врло сличан GET методу, са тим што се тело поруке не враћа клијенту (враћа се само статусна линија и заглавље). Метод се може користити ради утврђивања да ли је линк измењен у односу на претходно стање - измењено стање се тестира упоређивањем информација послатих у заглављу захтева са информацијама из заглавља генерисаног одговора
  - POST - се користи за захтев да се пошаљу подаци HTTP серверу коришћењем HTML форме

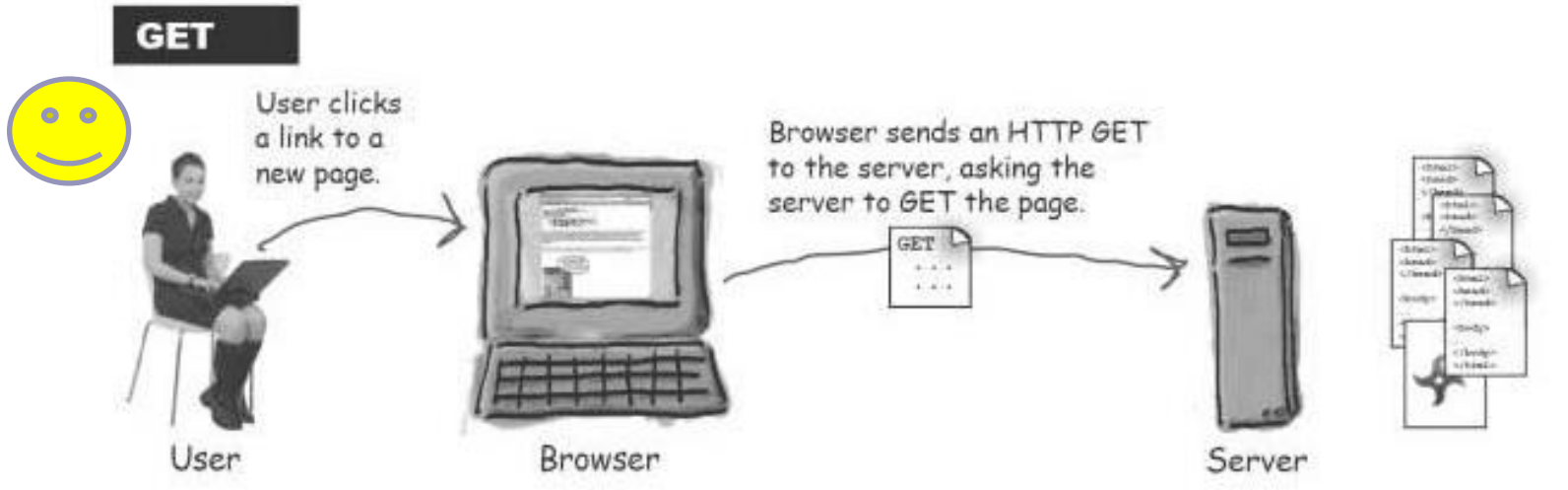
Математички факултет
vladaf@matf.bg.ac.rs

HTTP метод (2)

- PUT - користи се за захтев HTTP серверу да се подаци послати у оквиру захтева сместе на месту наведеног ресурса
- DELETE - користи се за захтев серверу да се уклони наведени ресурс
- TRACE – извршава тестирање повратне поруке дуж путање којом се захтев креће према циљном ресурсу
- OPTIONS – описује опције комуникације за циљни ресурс
- CONNECT – обезбеђује тунелску комуникацију према серверу одређеним са датом адресом

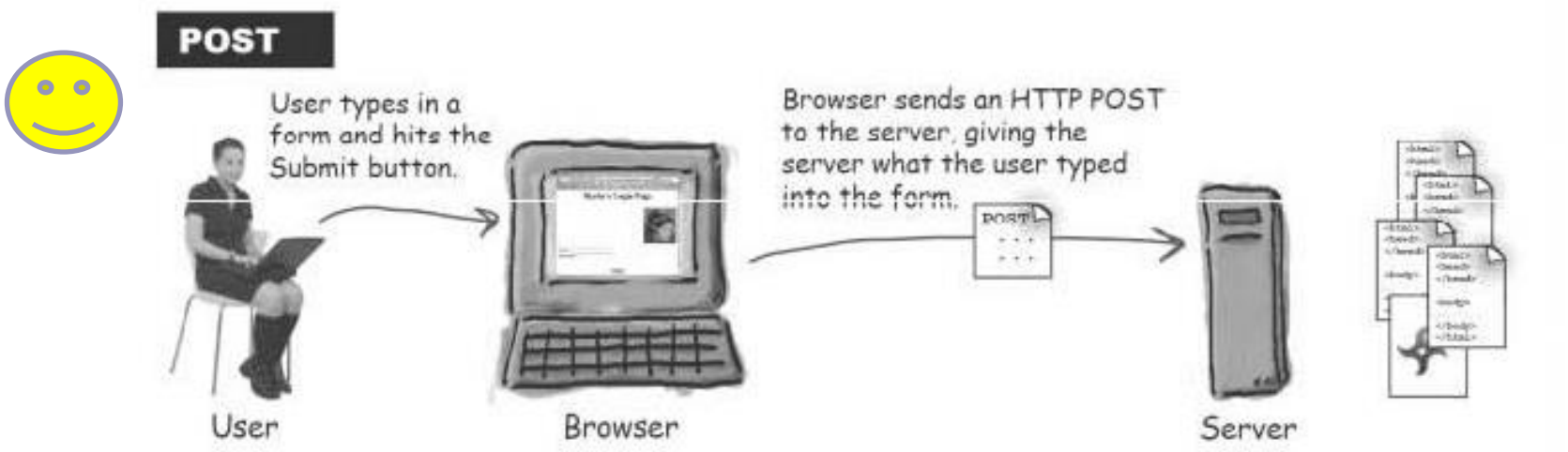
Методи GET и POST

- Метод GET је најједноставнији HTTP метод
- Метод GET тражи од сервера да прибави ресурс и да га врати позиваоцу
- Ресурс може бити HTML страна, PDF документ, JPG слика ...
- Сврха метода GET је да се добије ресурс од сервера



Методи GET и POST (2)

- Метод POST је моћнији од метода GET
- Коришћењем овог метода може се захтевати нешто од и истовремно слати податке на сервер (па сервер може процесирати приспеле податке)



Методи GET и POST (3)

- Подаци се могу слати на сервер и помоћу метода GET и помоћу метода POST
  - Укупан број знакова који се помоћу могу послати метода GET је много мањи од броја знакова који се могу послати преко метода POST, а то горње ограничење зависи од врсте веб сервера и прегледача
  - Подаци који се шаљу коришћењем метода GET се налепљују на адресу у адресној линији прегледача, па је све што се тим путем шаље на сервер директно видљиво кориснику (и самим тим подложније манипулацији)
  - Корисник не може поставити маркер на страницу где је садржај форме прослеђен методом POST, а може ако је за прослеђивање коришћен метод GET

Математички факултет

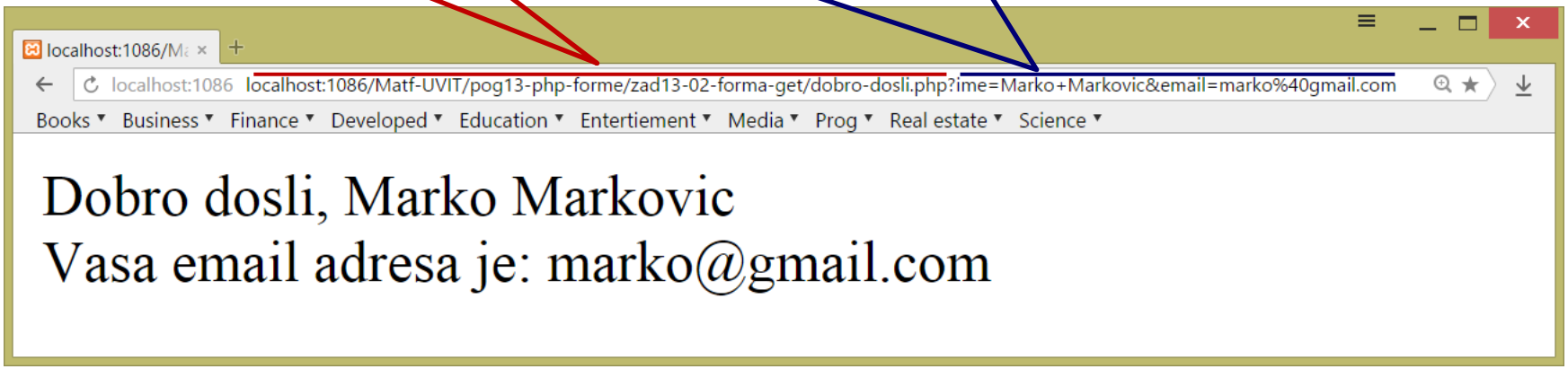
vladaf@matf.bg.ac.rs

Методи GET и POST (4)

- Илустрација слања података на сервер и помоћу метода GET

Оригинални URL без додатних параметара

Симбол ? раздваја путању од параметара, а симбол & раздваја параметре. Параметри се дефинишу у облику име = вредност



Анатомија GET захтева

The Request line.

The HTTP Method.

The path to the resource on the web server.

In a GET request, parameters (if there are any) are appended to the first part of the request URL, starting with a "?". Parameters are separated with an ampersand "&".

The protocol version that the web browser is requesting.

```
GET /select/selectBeerTaste.jsp?color=dark&taste=malty HTTP/1.1
```

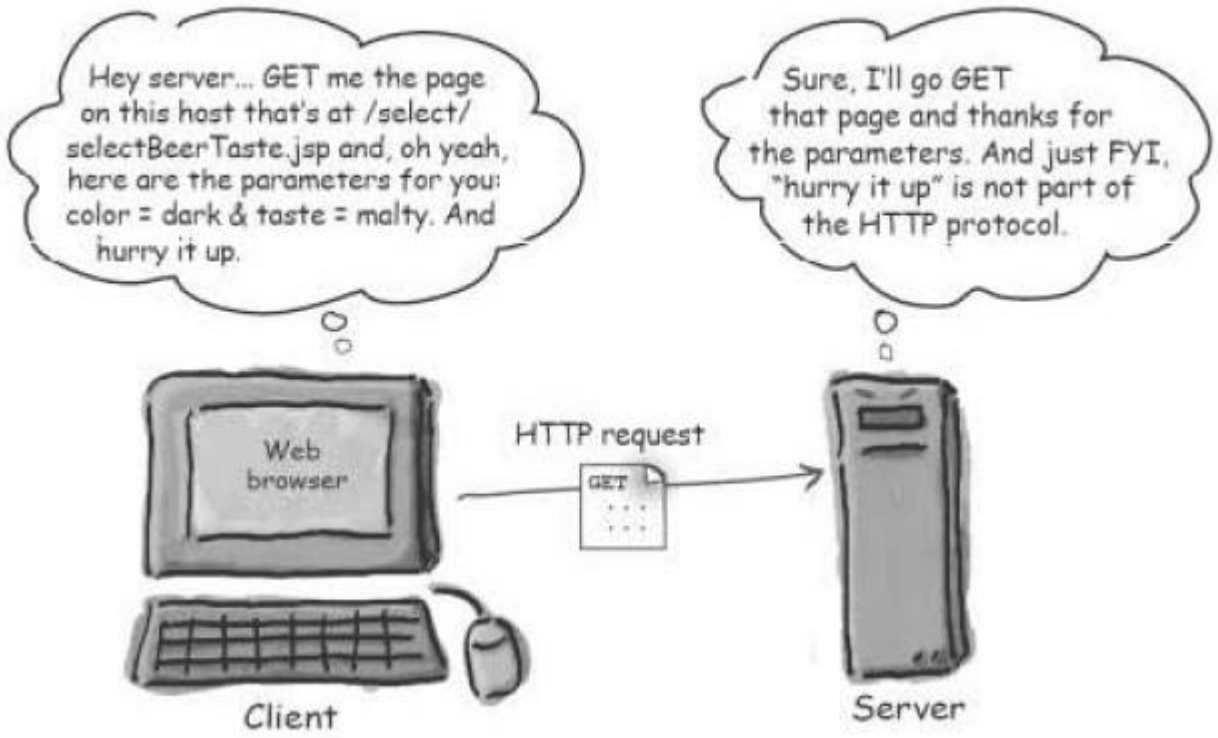
The Request headers.

```
Host: www.wickedlysmart.com
User-Agent: Mozilla/5.0 (Macintosh; U; PPC Mac OS X Mach-O; en-US; rv:1.4) Gecko/
20030624 Netscape/7.1
Accept: text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/
plain;q=0.8,video/x-mng,image/png,image/jpeg,image/gif;q=0.2,*/*;q=0.1
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
```

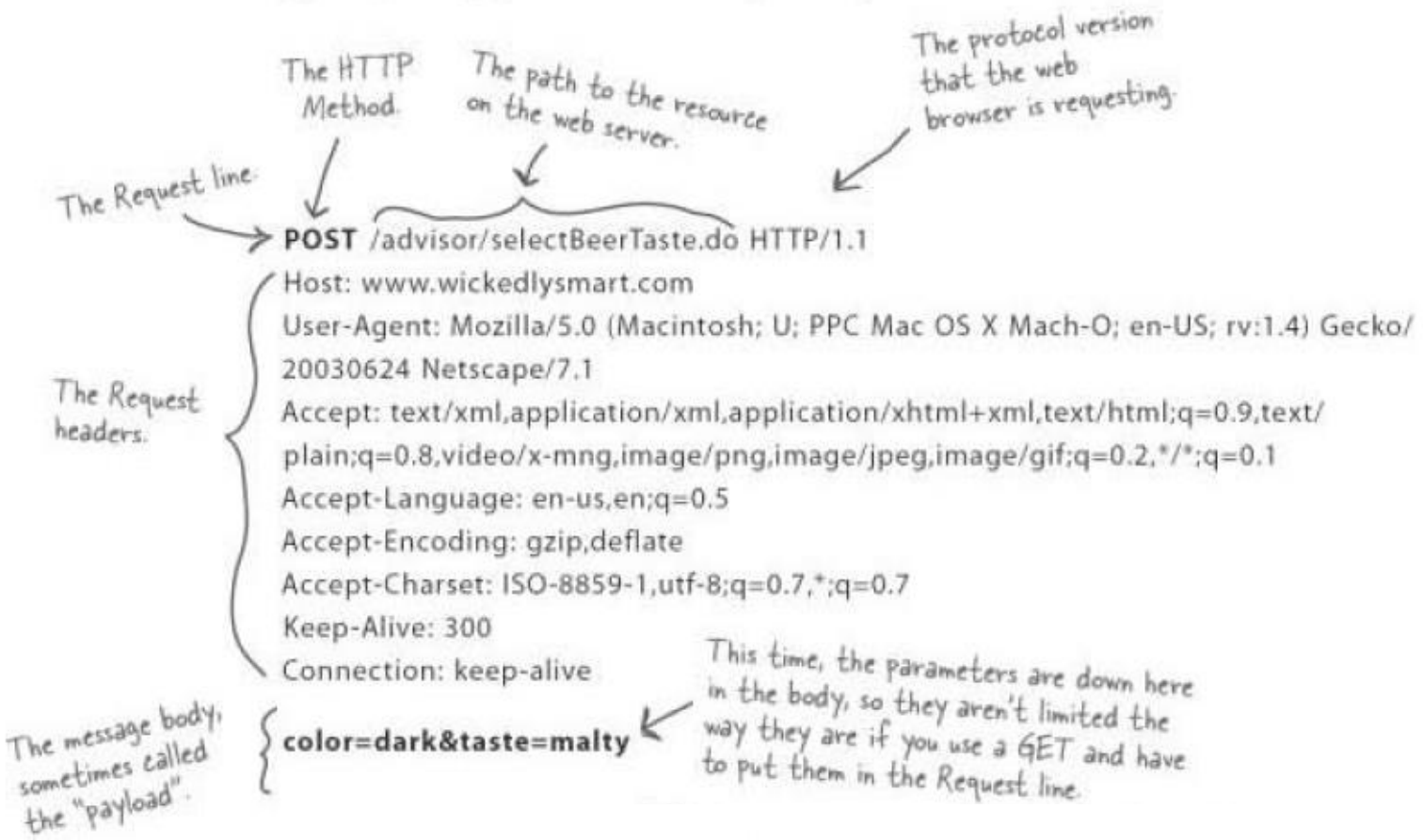
Математички факултет

vladaf@matf.bg.ac.rs

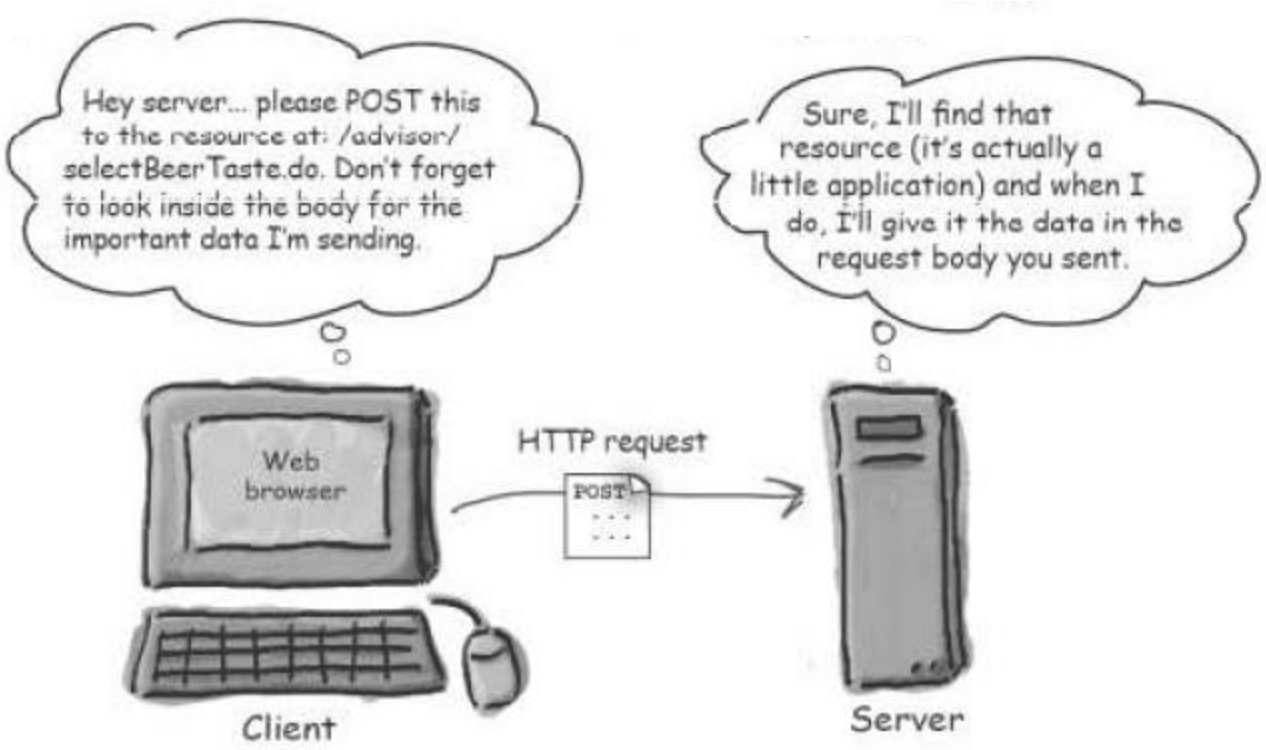
Анатомија GET захтева (2)



Анатомија POST захтева

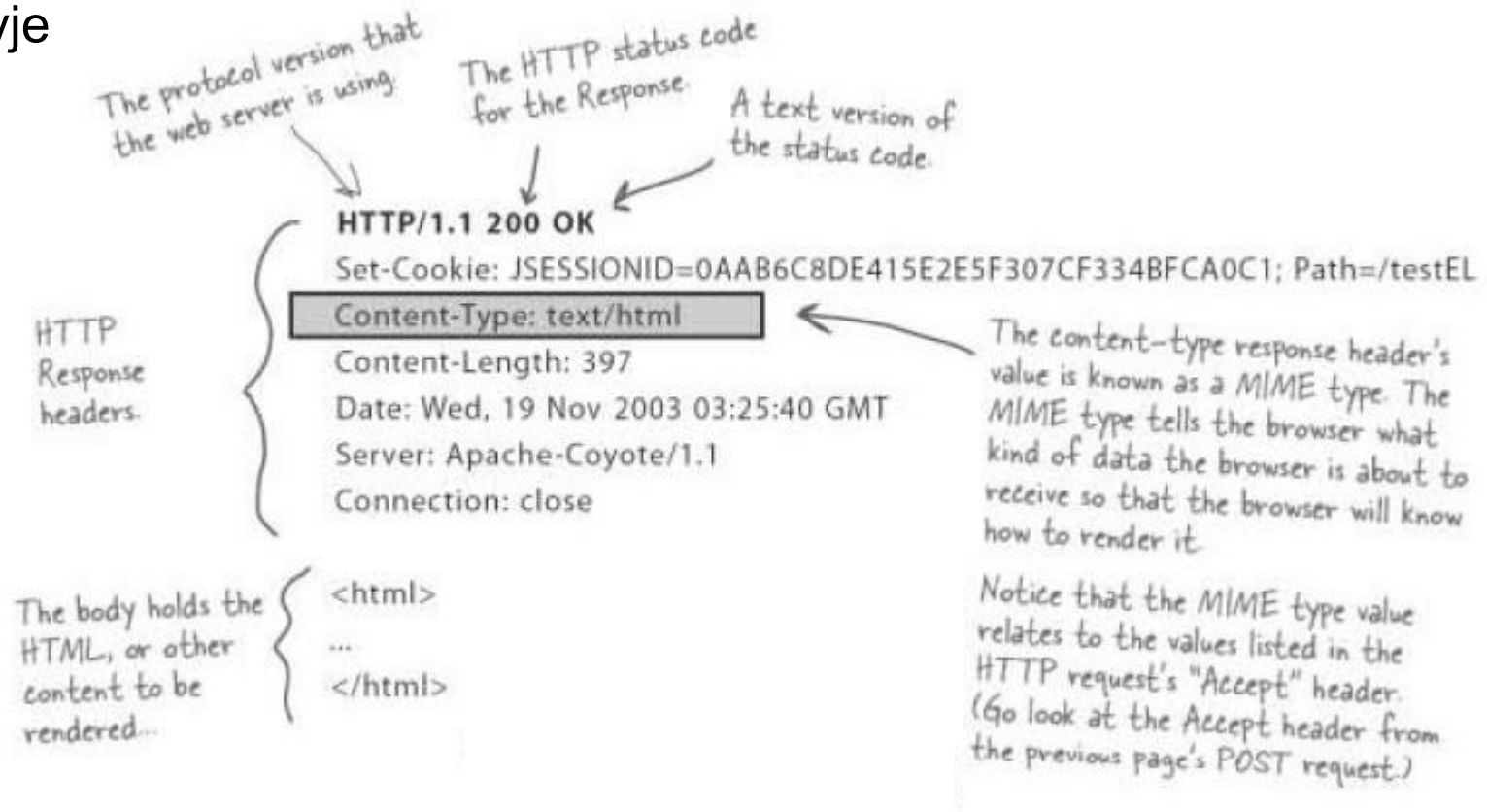


Анатомија POST захтева (2)



Анатомија HTTP одговора

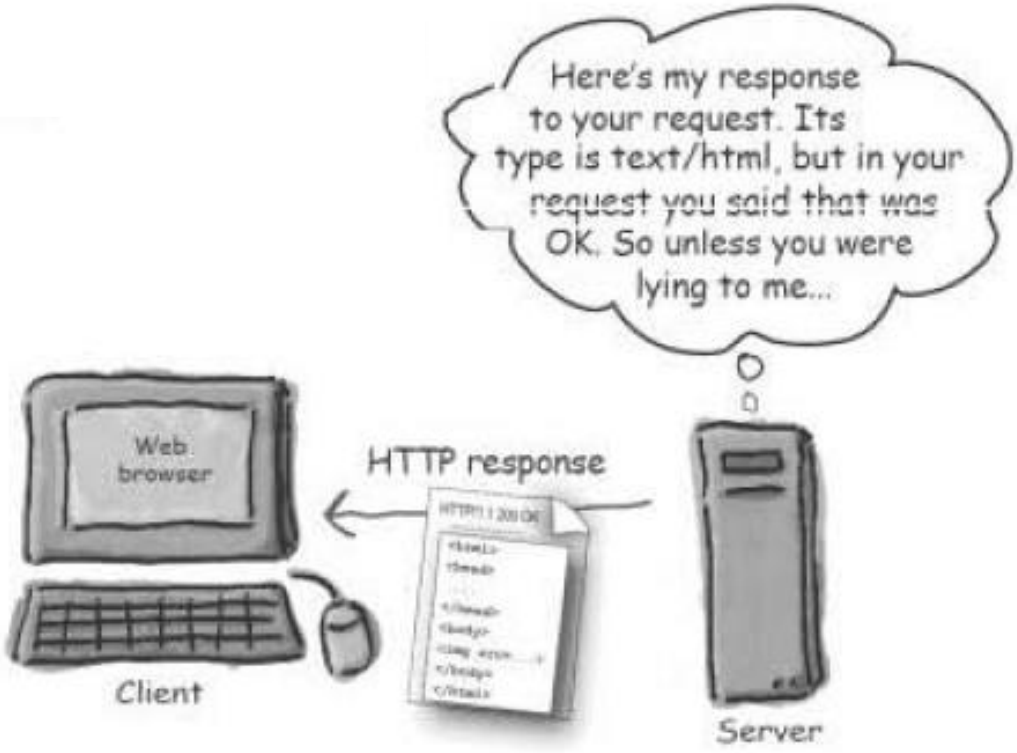
- HTTP одговор садржи заглавље и тело
  - Информације у заглављу говоре прегледачу који је протокол коришћен, да ли је захтев био успешан и која врста садржаја се налази у телу захтева, а тело садржи сам садржај који прегледач приказује



Математички факултет

vladaf@matf.bg.ac.rs

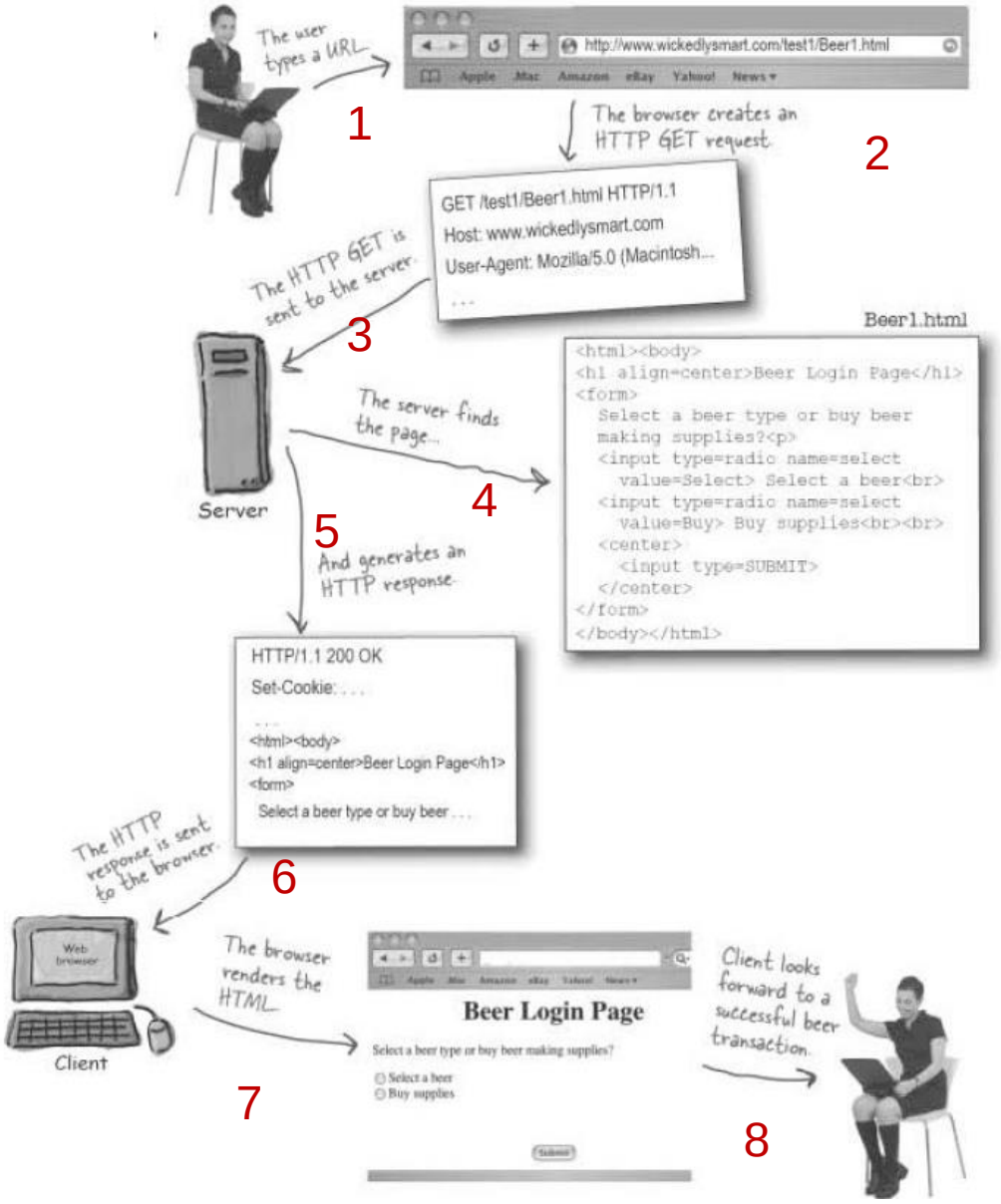
Анатомија HTTP одговора (2)



Математички факултет

vladaf@matf.bg.ac.rs

HTTP захтев и HTTP одговор



Математички факултет
vladaf@matf.bg.ac.rs

Статусни кодови HTTP одговора

- Статусни код HTTP одговора је троцифрен број. Прва цифра статусног кода HTTP одговора специфицира о којој се од пет класа одговора ради
- Статусни кодови су прошириви и HTTP клијенти нису обавезни да разумеју значење свих статусних кодова
- Побројани статусни кодови су део HTTP/1.1 стандарда (документ RFC 7231), осим ако се не истакне да је другачије

Математички факултет

vladaf@matf.bg.ac.rs

Статусни кодови HTTP одговора (2)

- 1xx: Information – Захтев је примљен и процес се наставља

1xx: Information

Message	Description
100 Continue	Only a part of the request has been received by the server, but as long as it has not been rejected, the client should continue with the request.
101 Switching Protocols	The server switches protocol.

Математички факултет

vladaf@matf.bg.ac.rs

Статусни кодови HTTP одговора (3)

- 2xx: Successful – захтев је успешно примљен, схваћен и прихваћен

2xx: Successful

Message	Description
200 OK	The request is OK.
201 Created	The request is complete, and a new resource is created .
202 Accepted	The request is accepted for processing, but the processing is not complete.
203 Non-authoritative Information	The information in the entity header is from a local or third-party copy, not from the original server.
204 No Content	A status code and a header are given in the response, but there is no entity-body in the reply.
205 Reset Content	The browser should clear the form used for this transaction for additional input.
206 Partial Content	The server is returning partial data of the size requested. Used in response to a request specifying a *Range* header. The server must specify the range included in the response with the *Content-Range* header.

Статусни кодови HTTP одговора (4)

- 3xx: Redirection – морају се предузети додатне акције да би се комплетирао захтев

3xx: Redirection

Message	Description
300 Multiple Choices	A link list. The user can select a link and go to that location. Maximum five addresses .
301 Moved Permanently	The requested page has moved to a new url .
302 Found	The requested page has moved temporarily to a new url .
303 See Other	The requested page can be found under a different url .
304 Not Modified	This is the response code to an *If-Modified-Since* or *If-None-Match* header, where the URL has not been modified since the specified date.
305 Use Proxy	The requested URL must be accessed through the proxy mentioned in the *Location* header.
306 *Unused*	This code was used in a previous version. It is no longer used, but the code is reserved.
307 Temporary Redirect	The requested page has moved temporarily to a new url.

Математички факултет

vladaf@matf.bg.ac.rs

Статусни кодови HTTP одговора (5)

- 4xx: Client Error – захтев садржи некоректну синтаксу или не може бити испуњен

4xx: Client Error

Message	Description
400 Bad Request	The server did not understand the request.
401 Unauthorized	The requested page needs a username and a password.
402 Payment Required	*You can not use this code yet.*
403 Forbidden	Access is forbidden to the requested page.
404 Not Found	The server can not find the requested page.
405 Method Not Allowed	The method specified in the request is not allowed.
406 Not Acceptable	The server can only generate a response that is not accepted by the client.
407 Proxy Authentication Required	You must authenticate with a proxy server before this request can be served.
408 Request Timeout	The request took longer than the server was prepared to wait.
409 Conflict	The request could not be completed because of a conflict.
410 Gone	The requested page is no longer available .
411 Length Required	The "Content-Length" is not defined. The server will not accept the request without it .
412 Precondition Failed	The pre condition given in the request evaluated to false by the server.
413 Request Entity Too Large	The server will not accept the request, because the request entity is too large.
414 Request-url Too Long	The server will not accept the request, because the url is too long. Occurs when you convert a "post" request to a "get" request with a long query information .
415 Unsupported Media Type	The server will not accept the request, because the mediatype is not supported .
416 Requested Range Not Satisfiable	The requested byte range is not available and is out of bounds.
417 Expectation Failed	The expectation given in an Expect request-header field could not be met by this server.

Статусни кодови HTTP одговора (6)

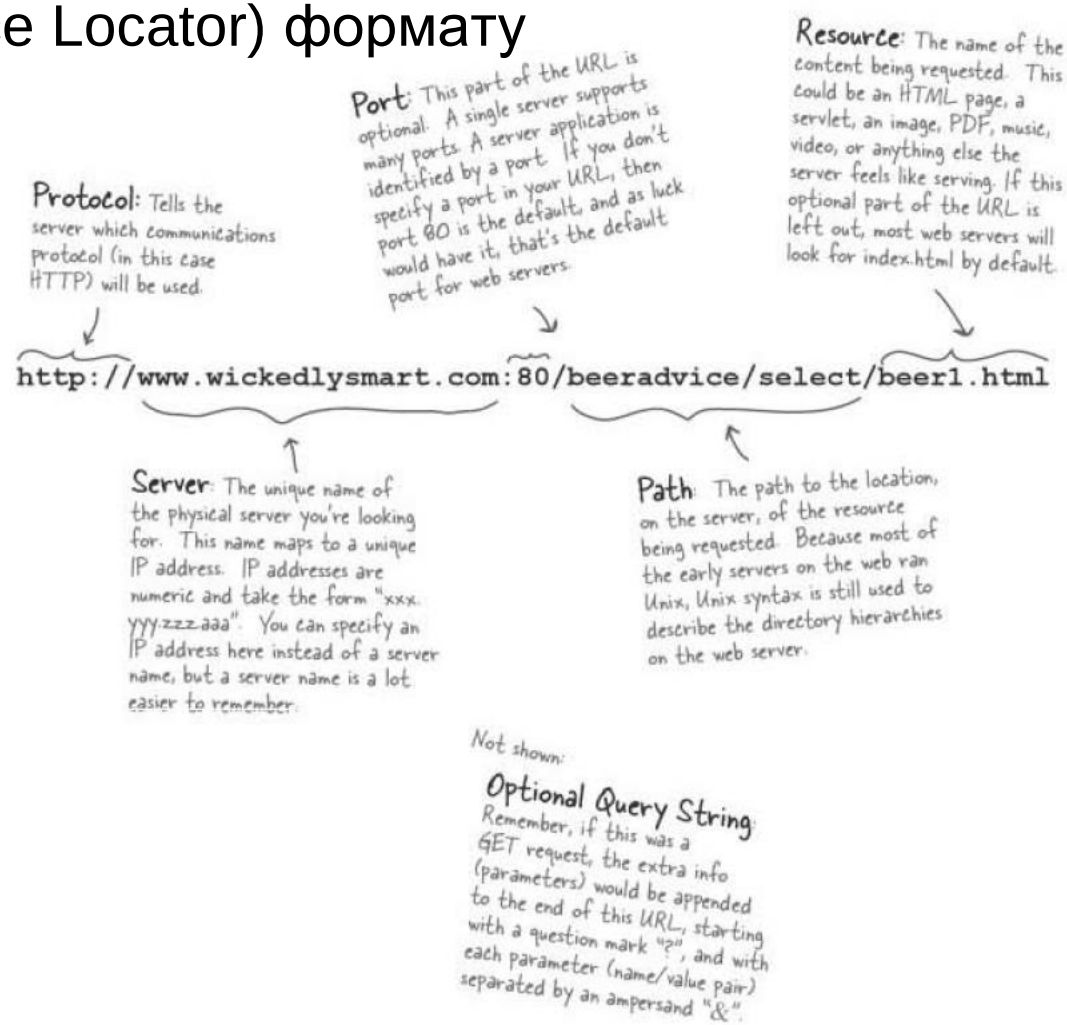
- 5xx: Server Error – сервер није успео да испуни наизглед ваљан захтев

5xx: Server Error

Message	Description
500 Internal Server Error	The request was not completed. The server met an unexpected condition.
501 Not Implemented	The request was not completed. The server did not support the functionality required.
502 Bad Gateway	The request was not completed. The server received an invalid response from the upstream server.
503 Service Unavailable	The request was not completed. The server is temporarily overloading or down.
504 Gateway Timeout	The gateway has timed out.
505 HTTP Version Not Supported	The server does not support the "http protocol" version.

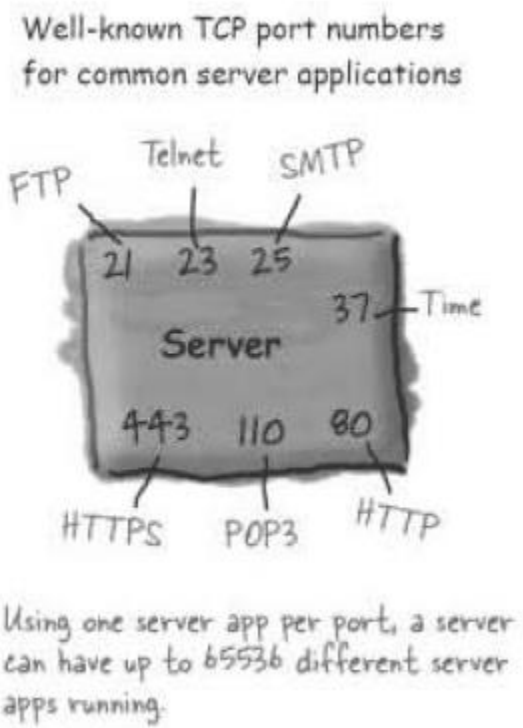
URL

- Сваки ресурс на вебу има своју јединствену адресу, у URL (Uniform Resource Locator) формату



TCP порт

- Различите апликације користе различите портове
  - HTTP сервер подразумевано користи порт 80
  - Telnet сервер подразумевано користи порт 23
  - FTP сервер подразумевано користи порт 21
  - POP3 порт 110, а SMTP порт 25
  - Сервер за време порт 37
- Порт представља једнозначни идентификатор – број између 0 и 65535
- Порт представља логичку везу између конкретног софтвера и хардвера на ком се тај софтвер извршава
- Порт не представља место на које се прикључује уређај, већ број који означава приступну тачку за комуникацију са сервером



Математички факултет
vladaf@matf.bg.ac.rs

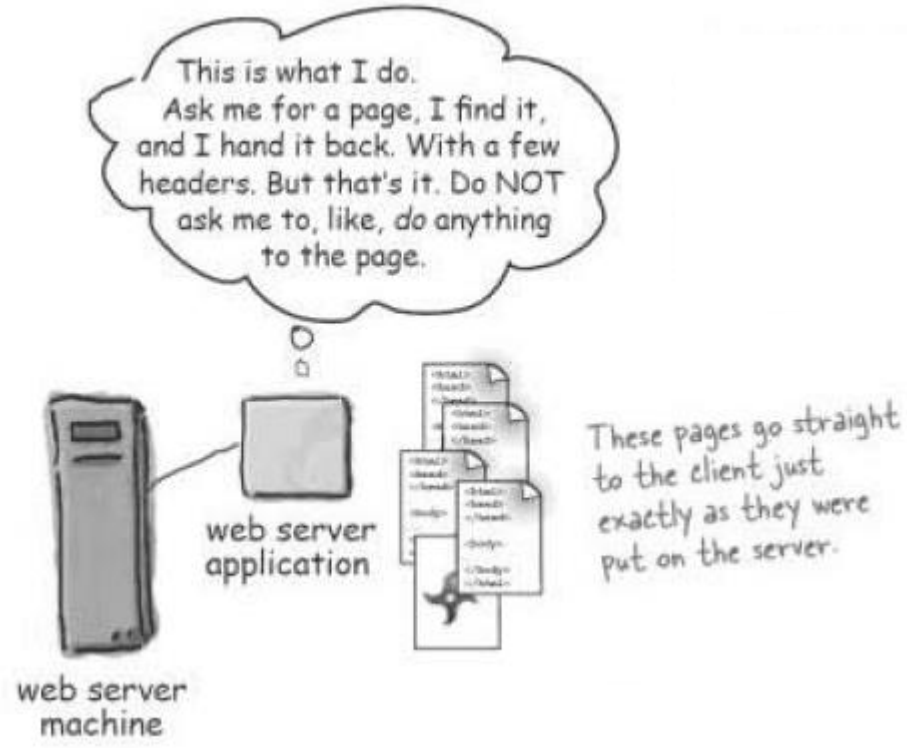
TCP порт (2)

- Ако не би било броја порта, онда сервер не би знао са којом клијентском апликацијом комуницира, нити по ком се протоклу та комуникација реализује
  - Могло би се догодити, на пример, да прегледач, уместо да комуницира са веб сервером, покушава успостави комуникацију са сервером електронске поште
- TCP портови између 0 и 1023 су резервисани за познате сервисе и не препоручује се да се ти портови користе за „нове“ серверске програме
- Ако се извршавају серверски програми на рачунарској мрежи компаније, тада са администратором система треба проверити који су портови слободни, а који заузети

Математички факултет
vladaf@matf.bg.ac.rs

Статичке веб стране

- Статичка веб страна се налази у директоријуму на веб серверу
  - Веб сервер такву страну само пронађе и проследи клијенту, баш онакву каква је на серверу
  - Сваки од клијената добија потпуно исти садржај као одговор

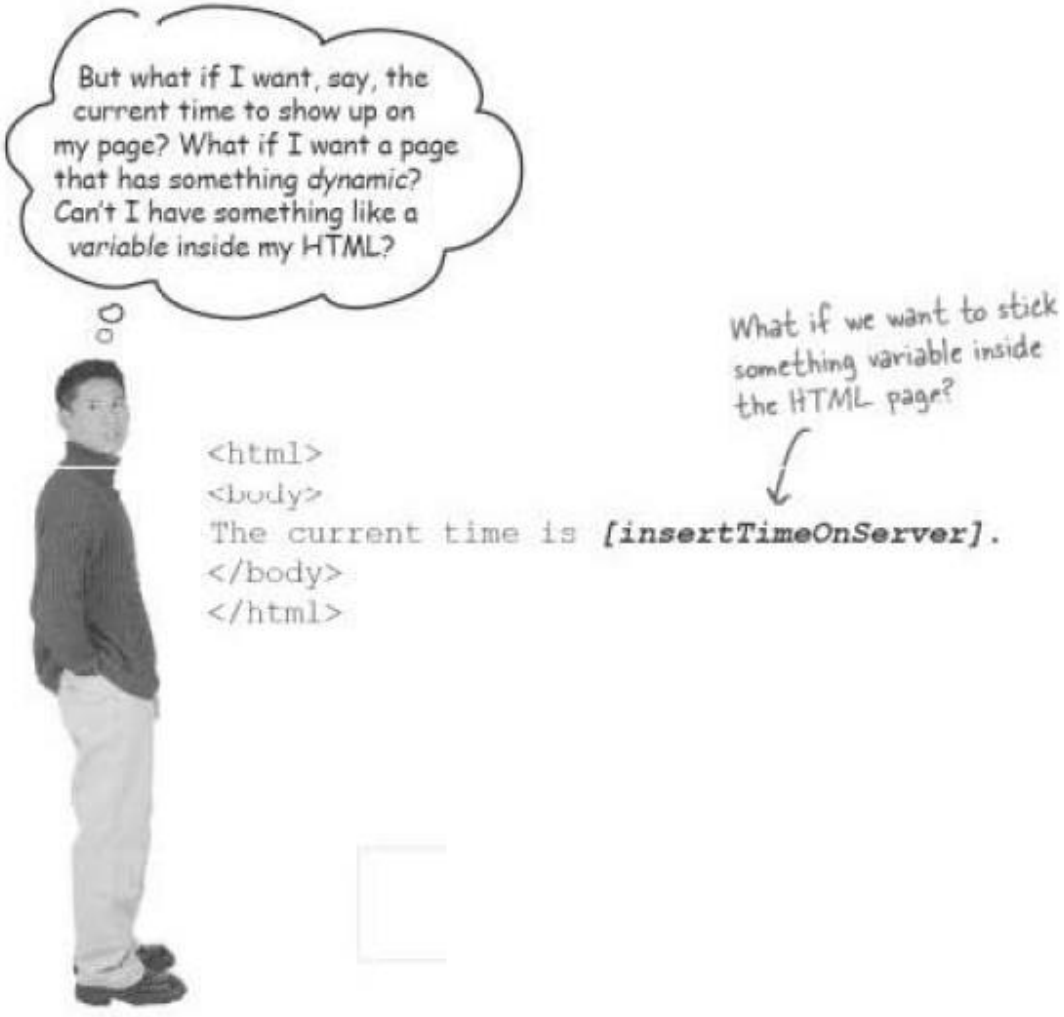


Математички факултет

vladaf@matf.bg.ac.rs

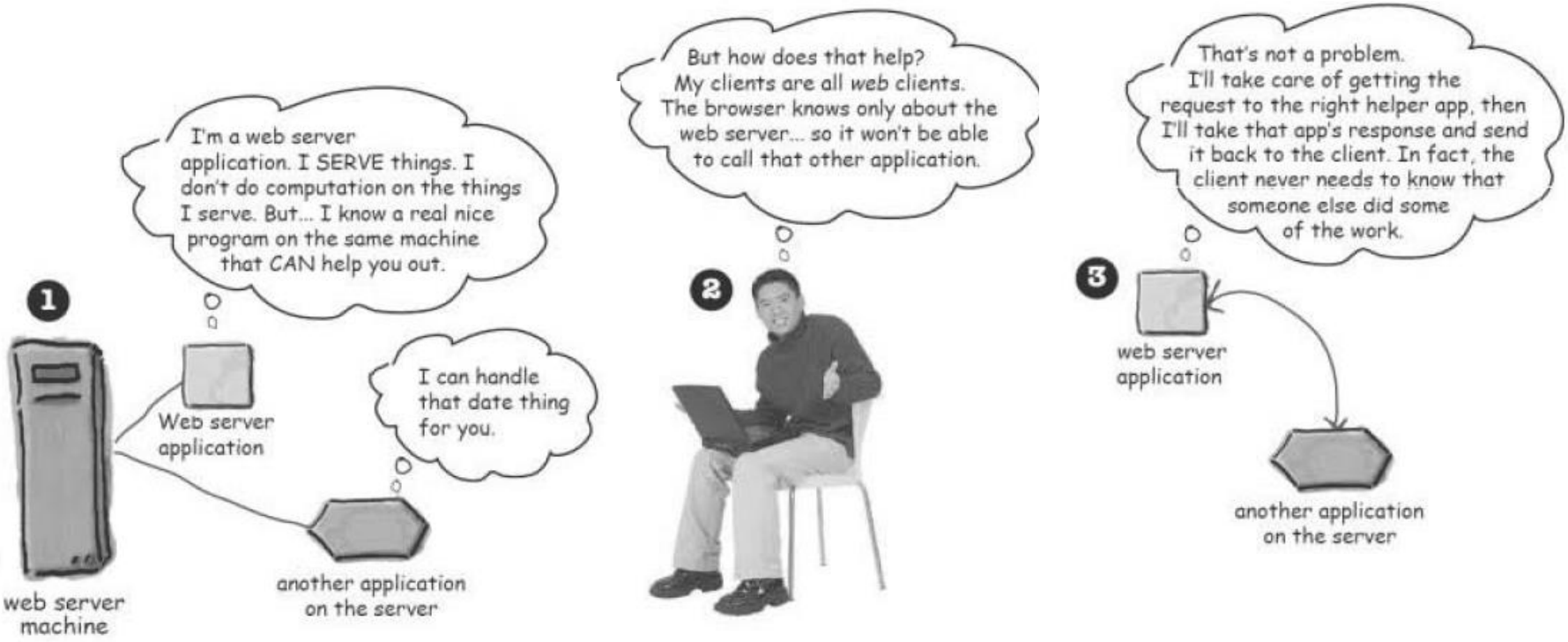
Статичке веб стране (2)

- Шта радити када треба обезбедити да се веб страна динамички мења?



Динамичке веб стране

- Сам веб сервер опслужује само статичке стране, али се може користити посебна помоћна апликација, са којом комуницира веб сервер, а која креира динамички садржај



Математички факултет
vladaf@matf.bg.ac.rs

Динамичке веб стране (2)

- Динамички садржај може бити било шта: датум и време са сервера, списак датотека у директоријуму, случајно изабрана слика итд.
- Динамички садржај не постоји све док не стигне захтев
- По приспећу захтева, помоћна апликација „креира“ HTML а онда веб сервер тај HTML „спакује“ у одговор

Уместо статичног

```
<html>
<body>
The current time is
always 4:20 PM
on the server
</body>
</html>
```

Треба да се добије динамичан садржај

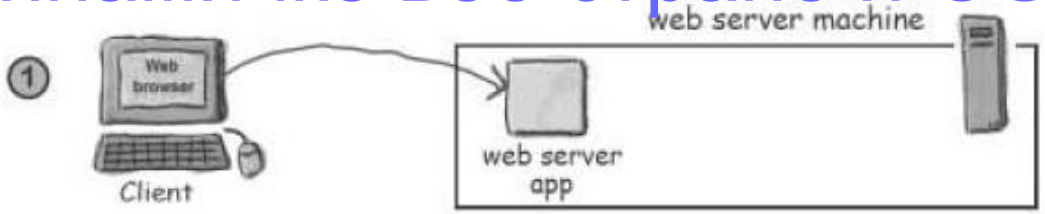
```
<html>
<body>
The current time is
[insertTimeOnServer]
on the server
</body>
</html>
```

Математички факултет
vladaf@matf.bg.ac.rs

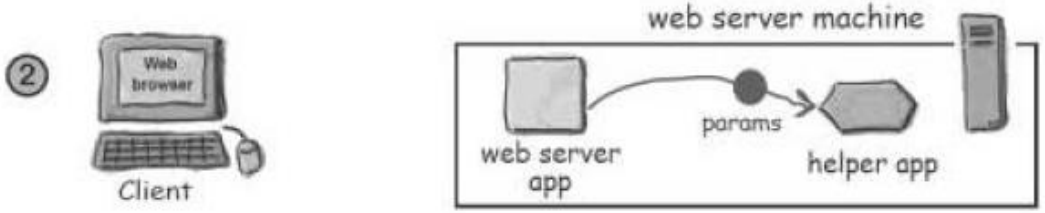
Динамичке веб стране (3)

- Када корисник проследи серверу податке са форме, тада је за процесирање прослеђених података (чување података у бази, ради генерисање одговора на основу података прослеђених уз захтев итд.) неопходно коришћење помоћне апликације
- Када сервер препозна да се захтев односи на помоћну апликацију, тада и прослеђене параметре проследи помоћној апликацији, па та апликација генерише одговор за који се потом проследи клијенту

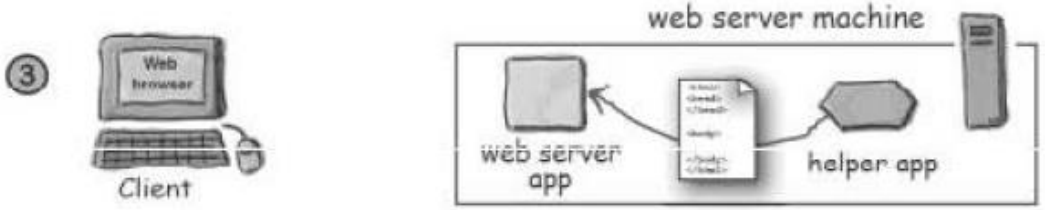
Динамичке веб стране и CGI



User clicks a link that has a URL to a CGI program instead of a static page.

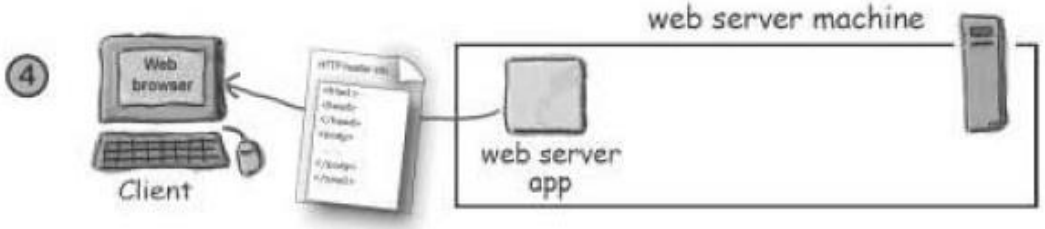


Web server application "sees" that the request is for a helper program, so the web server launches and runs the program. The web server app sends along any parameters from a GET or POST.



The helper app constructs the brand new page (that has the current date inserted) and sends the HTML back to the server.

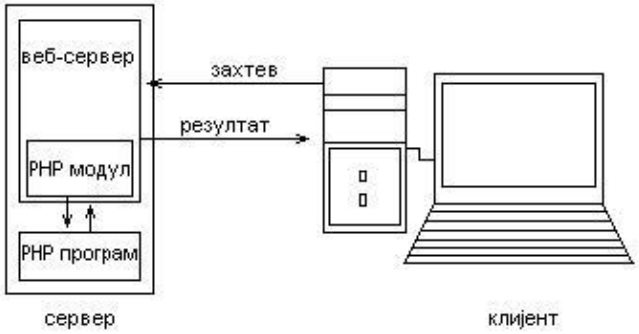
As far as the web server is concerned, the HTML from the helper app is a static page.



The helper application is shut down, and the client gets back an HTML page that has the current date as part of its now-static content.

Динамичке веб стране и PHP

- PHP интерпретатор може радити по CGI принципу
  - Интерпретатор постоји као екстерна апликација
  - Та екстерна апликација се позива да изврши дату скрипту (дати PHP програм) сваки пут кад буде захтевана од неког корисника
- PHP интерпретатор може бити инсталиран и као модул веб сервиса
  - Интерпретатор је ближе повезан са веб сервисом
  - Интерпретатор је на тај начин увек учитан у меморију, па се не мора позивати спољашњи програм
  - Овакав приступ пружа знатно већу брзину извршавања



Математички факултет
vladaf@matf.bg.ac.rs

Динамичке веб стране и PHP (2)

- Уобичајен сценарио по ком се извршавају PHP скрипте
  1. клијент (корисник Интернета који користи неки прегледач) захтева PHP страницу са сервера (рачунара)
  2. сервер прослеђује захтев веб серверу (програм на серверу)
  3. веб сервер препознаје да се тражи PHP датотека
  4. веб сервер не шаље његов садржај захеваног ресурса клијенту, него га извршава као програм помоћу PHP модула
  5. излазни текст програма (стандардни излаз тј. резулатат рада PHP модула) се шаље клијенту као резултат захтева
  6. прегледач препознаје врсту резултата (HTML код, слика, PDF садржај, архива итд.)
  7. прегледач приказује резултат клијенту на одговарајући начин

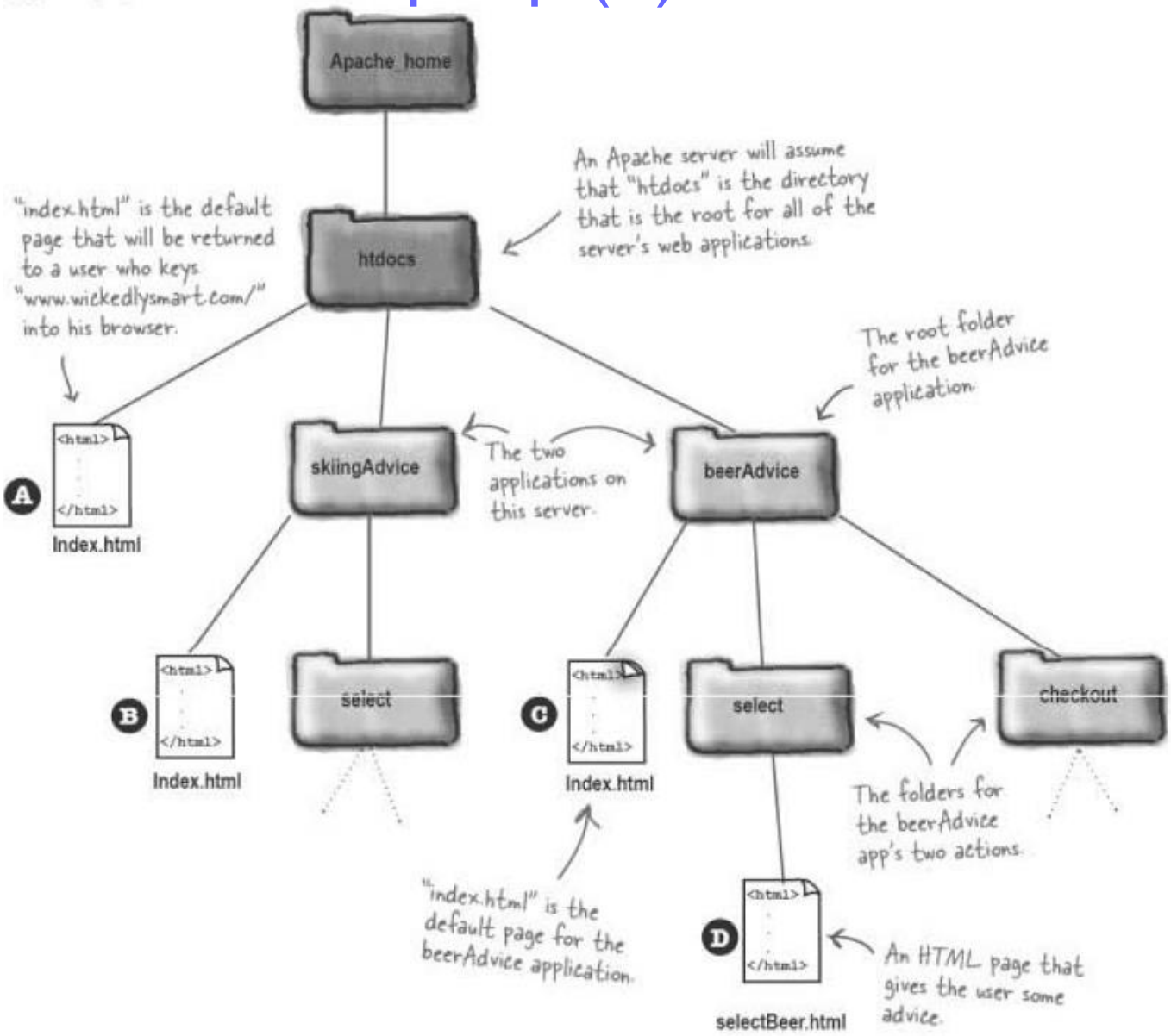
Apache веб сервер

- У оквиру овог курса биће коришћен Apache веб сервер
  - Веб сервер Apache је највише коришћен софтвер за веб сервер на свету
    - Заснован је NCSA HTTPd серверу
    - Развој Apache веб сервера је почео 1995, кад је заустављен развој NCSA кода
    - Apache је имао кључну улогу у иницијалном развоју веба, а од априла 1996. је најпопуларнији веб сервер
    - У новембру 2015 је процењено да Apache опслужује 50% активних веб сајтова и 37% најбољих сервера у свим доменима
  - Apache веб сервер је софтвер отвореног кода/слободан софтвер под лиценцом Apache License, који развија и одржава заједница Apache Software Foundation
  - Најчешће се користи на Unix-оликим системима, али и на eComStation, Microsoft Windows, NetWare, OpenVMS, OS/2 и TPF.

Математички факултет

vladaf@matf.bg.ac.rs

Apache веб сервер (2)



Математички факултет

vladaf@matf.bg.ac.rs

Захвалница

Делови материјала ове презентације су преузети из:

- Скрипте из предмета Увод у веб и интернет програмирање на Математичком факултету, аутор проф. др Филип Марић
- Скрипте из предмета Информатика на Универзитету Milano Bicocca, аутор Mirko Cesarini
- Књиге Head First PHP & MySQL, аутори Lynn Beighley и Michael Morrison, издавач O'Reilly, 2009.
- Књиге Head First Servlets and JSP, аутори Bryan Basham, Kathy Sierra и Bert Bates, издавач O'Reilly, 2008.