

# Programowanie aplikacji internetowych

WWW

# Protokół HTTP

HTTP (ang. *Hypertext Transfer Protocol*)

protokół przesyłania dokumentów hipertekstowych WWW.

Za pomocą protokołu HTTP przeglądarka wysyła żądanie do serwera i otrzymuje odpowiedź.

Najważniejsze cechy:

- Znormalizowany sposób komunikacji, umożliwia komunikowanie się komputerów niezależnie od platformy
- Ma formę transakcji, składających się z żądania (request) oraz odpowiedzi (response)
- Bezstanowy – nie przechowuje żadnych informacji o zrealizowanej transakcji; Stanowość zapewnia najczęściej mechanizm sesji kontrolowanej przez serwer, do czego wykorzystuje się mechanizm ciasteczek (cookies) zapisywanych przez klienta

# Protokół HTTP

## Żądanie HTTP ma postać:

Metoda Zasób Wersja-http

Nagłówek

Nagłówek

[pusta linia]

Zawartość (jeżeli potrzebna)

## Przykład:

GET / HTTP/1.1

Host: www.domena.com

Connection: keep-alive

# Protokół HTTP

## Odpowiedź HTTP ma postać:

Wersja-http Kod-odpowiedzi Uzasadnienie

Nagłówek

Nagłówek

[pusta linia]

Zawartość (jeżeli potrzebna)

## Przykład:

HTTP/1.1 200 OK.

Server: Apache

Content-type: text/html

[Dokument HTML]

# Protokół HTTP

Zdefiniowano 8 metod (rodzajów żądań) HTTP:

- GET – pobranie wskazanego przez URI zasobu; najczęściej wykorzystywany
- POST – przesłanie danych od klienta do serwera, np. danych formularza; dane przekazywane są do wskazanego zasobu, tj. skryptu CGI, PHP, ...
- HEAD – pobranie informacji o wskazanym zasobie
- PUT – umieszczenie zasobu (pliku) na serwerze
- DELETE – usunięcie zasobu z serwera
- OPTIONS – informacje o kanale komunikacyjnym
- TRACE – do celów diagnostycznych
- CONNECT – dotyczące tunelowania

# Protokół HTTP

Ważniejsze kody odpowiedzi serwera HTTP:

- 200 OK – serwer przesyła żądany dokument w zawartości
- 204 No content – serwer zrealizował żądanie, jednak nie wymaga ono przesłania zawartości
- 301 Moved permanently – zasób został przeniesiony
- 401 Unauthorized – dostęp do zasobu wymaga autoryzacji
- 403 Forbidden – serwer zrozumiał zapytanie, jednak dostęp do zasobu jest zabroniony
- 404 Not found – serwer nie odnalazł zasobu
- 500 Internal Server Error – serwer nie może zrealizować żądania z powodu błędów w oprogramowaniu (np. CGI)
- 501 Not implemented – metoda nie jest obsługiwana (np. PUT)
- 503 Service unavailable – usługa niedostępna (przeciążenie)

# Statyczne i dynamiczne strony WWW

## Statyczne strony WWW

- Dokumenty HTML umieszczone na serwerze; zmiana zawartości witryny WWW wymaga modyfikacji plików HTML; niewielkie obciążenie serwera

## Dynamiczne strony WWW

- Serwer ma dodatkowe moduły (np. interpreter PHP); Zamiast plików HTML na serwerze umieszczone są skrypty generujące dokumenty HTML (np. zamiast x.html –x.php)
- Zawartość dokumentu HTML zwracanego do przeglądarki może być uzależniona od wielu czynników, np. parametrów wybranych przez użytkownika w oknie przeglądarki, zawartości baz danych, odpowiedzi innych serwerów na zapytania itp.
- Większość serwisów do składowania informacji prezentowanych na witrynie wykorzystuje bazy danych

# Statyczne strony WWW

## Dokumenty HTML umieszczone na serwerze

- Konfiguracja serwera www określa folder, którego zawartość jest udostępniona przez http/www (żadne pliki spoza tego folderu nie mogą być przez serwer udostępnione), np.:

`/var/www/`

- Lokalizacja i nazwa pliku wynika wprost z żądania http (wyjątek – jeżeli w żądaniu jest nazwa folderu, serwer szuka w tym folderze pliku o nazwie domyślnej, zwykle index.html):

|                   |                         |
|-------------------|-------------------------|
| GET /             | - /var/www/index.html   |
| GET /style.css    | - /var/www/style.css    |
| GET /njiti/w.html | - /var/www/njiti/w.html |
| GET /img/logo.jpg | - /var/www/img/logo.jpg |

# Statyczne strony WWW

Dokumenty HTML umieszczone na serwerze

- ...
- Oprócz samego żądania http, serwer uwzględnia również nagłówki http (np. nagłówek "Host: domena.com" – jeżeli na serwerze jest kilka domen)
- Nagłówki odpowiedzi generuje serwer

# Dynamiczne strony WWW - PHP

Skrypty php umieszczone na serwerze

- Plik php zawiera html i instrukcje php
  - fragmenty html oraz dane "drukowane" przez php na standardowe wyjście – funkcja `echo()` – są wysyłane do przeglądarki
- Przeglądarka otrzymuje dokument html jak przy stronach www statycznych – nie potrzebuje nic dedykowanego dla php, żeby wyświetlać strony generowane w php

# Dynamiczne strony WWW - PHP

## Skrypty php umieszczone na serwerze

- Pliki php są jednocześnie plikami html, wszystkie zasady odnajdowania i udostępniania plików dotyczą też skryptów php
- Lokalizacja i nazwa pliku wynika wprost z żądania http (wyjątek – jeżeli w żądaniu jest nazwa folderu, serwer szuka w tym folderze pliku o nazwie domyślnej, zwykle index.php):

|                   |                         |
|-------------------|-------------------------|
| GET /             | - /var/www/index.php    |
| GET /style.css    | - /var/www/style.css    |
| GET /njiti/w.php  | - /var/www/njiti/w.php  |
| GET /img/logo.jpg | - /var/www/img/logo.jpg |

- Pliki skryptów php oraz wszelkie inne pliki potrzebne w serwisie (arkusze CSS, obrazki, czcionki, ...) koegzystują w tym samym folderze – pliki "statyczne" są wysyłane do przeglądarki, skrypty – uruchamiane przez interpreter PHP

# Dynamiczne strony WWW - PHP

Skrypty php umieszczone na serwerze

- ...
- Uruchamiany jest skrypt otrzymuje wszystkie informacje z żądania http – w tym żądany zasób oraz wszystkie nagłówki; Zasób z żądania może zawierać dane z formularza html, nagłówki mogą zawierać dane z formularza oraz ciasteczka, np.:  
`GET /find.php?str=xxx&options=strict`
- Interpreter udostępnia wszelkie dane z żądania http (nagłówki, dane formularzy, ciasteczka, ...) w postaci tablic, bardzo wygodnej dla programisty, np. `$_GET`, `$_COOKIES`
- Do nagłówków odpowiedzi serwera są dodawane nagłówki wygenerowane przez skrypt php, np. nowe wartości ciasteczek

# Dynamiczne strony WWW - PHP

## Skrypty php umieszczone na serwerze

- ...
- Aplikacje internetowe php, np. systemy CMS, zmieniają domyślny sposób działania serwera przez przekierowania; W serwerze Apache służy do tego moduł mod\_rewrite:

```
--- .htaccess ---  
RewriteEngine On  
RewriteCond %{REQUEST_FILENAME} -f [OR]  
RewriteRule ^ - [L]  
RewriteRule ^(*)$ /index.php?rq=$1 [QSA,NC,L]
```
- Niezależnie co będzie w żądaniu, wykonany zostanie skrypt index.php (dane z żądania zostaną przekazane do tego skryptu)

```
GET /studenci -> /index.php?rq=studenci
```

# Język HTML

HTML (ang. *Hypertext Markup Language*)

język znaczników hipertekstowych – obecnie dominujący język stron internetowych.

- Opisuje strukturę informacji zawartych w dokumencie (nagłówki, akapity listy, tabele, ...)
- Umożliwia określenie sposobu prezentacji poszczególnych elementów dokumentu (wyrównanie, obramowanie, odstępy, czcionka, kolor, ...)
- Umożliwia osadzanie w dokumencie dodatkowych elementów (grafika, animacje, aplety, formularze, ...)
- Umożliwia osadzanie w dokumencie skryptów, wpływających na zachowanie przeglądarki (JavaScript, VB Script, ...)