

#### POLITECHNIKA WARSZAWSKA WYDZIAŁ ELEKTRYCZNY

Instytut Elektrotechniki Teoretycznej i Systemów Informacyjno-Pomiarowych Zakład Elektrotechniki Teoretycznej i Informatyki Stosowanej

# PRACA DYPLOMOWA INŻYNIERSKA

na kierunku Informatyka w specjalności: Inżynieria oprogramowania

Wykorzystanie protokołu HTTP/2 do budowy szybkiej aplikacji internetowej

Piotr Szklanko

nr albumu 244145

promotor mgr inż. Bartosz Chaber

Warszawa, 2017

# Wykorzystanie protokołu HTTP/2.0 do budowy szybkiej aplikacji internetowej

#### Streszczenie

Praca składa się z krótkiego wstępu jasno i wyczerpująco opisującego oraz uzasadniającego cel pracy, trzech rozdziałów (2-4) zawierających opis istniejących podobnych rozwiązań, komponentów rozpatrywanychjako kandydaci do tworzonego systemu i wreszcie zagadnień wydajności wirtualnych rozwiązań. Piąty rozdział to opis środowiska obejmujący opis konfiguracji środowiska oraz przykładowe ćwiczenia laboratoryjne. Ostatni rozdział pracy to opis możliwości dalszego rozwoju projektu.

Słowa kluczowe: praca dyplomowa, LaTeX, jakość

#### THESIS TITLE

#### Abstract

This thesis presents a novel way of using a novel algorithm to solve complex problems of filter design. In the first chapter the fundamentals of filter design are presented. The second chapter describes an original algorithm invented by the authors. Is is based on evolution strategy, but uses an original method of filter description similar to artificial neural network. In the third chapter the implementation of the algorithm in C programming language is presented. The fifth chapter contains results of tests which prove high efficiency and enormous accuracy of the program. Finally some posibilities of further development of the invented algoriths are proposed.

**Keywords:** thesis, LaTeX, quality

#### POLITECHNIKA WARSZAWSKA WYDZIAŁ ELEKTRYCZNY

#### **OŚWIADCZENIE**

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa inżynierska pt. Wykorzystanie protokołu HTTP/2 do budowy szybkiej aplikacji internetowej:

- została napisana przeze mnie samodzielnie,
- nie narusza niczyich praw autorskich,
- nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam, że przedłożona do obrony praca dyplomowa nie była wcześniej podstawą postępowania związanego z uzyskaniem dyplomu lub tytułu zawodowego w uczelni wyższej. Jestem świadom, że praca zawiera również rezultaty stanowiące własności intelektualne Politechniki Warszawskiej, które nie mogą być udostępniane innym osobom i instytucjom bez zgody Władz Wydziału Elektrycznego.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Piotr Szklanko
----------------

# Spis treści

1	Wst	zęp	1
<b>2</b>	HT	$\mathrm{TP/2}$	3
	2.1	Historia	3
	2.2	Protokół binarny	3
	2.3	Multiplexing	4
	2.4	Prioretyzacja	5
	2.5	Server push	5
	2.6	Kompresja nagłówków	6
3	Buc	lowa aplikacji	8
	3.1	node-spdy – konfiguracja serwera	8
	3.2	Server Push	8
4	Tes	$\mathbf{ty}$	10
	4.1	Chrome DevTools	10
	4.2	Testy wstępne	12
	4.3	Porównanie prędkości protokołu HTTP/2 i HTTP/1.1	14
	4.4	Porównanie prędkości przy bezpiecznym połączeniu HTTP/1.1	14
	4.5	Porównanie prędkości dla HTTP/1.1 z włączononym CACHE	17
	4.6	Porównanie prędkości przy wykorzystaniu Server Push	18
5	Wn	ioski	19
Bi	blios	grafia	20

# Wstęp

Moim celem jest przeprowadzenie testów protokołu HTTP w najnowszej wersji – HTTP/2. Obecnie powszechnie stosowana jest wersja 1.1, która została wprowadzona w roku 1999. Jednakże szybki rozwój technologii internetowych sprawia, że wprowadzony osiemnaście lat temu protokół przestaje powoli spełniać swoje zadanie. Obecnie bez wykorzystania takich środków jak:

- pamię cache przeglądarki, dzięki której nie musimy przesyłać wszystkich plików naszej aplikacji do użytkownika, który korzysta z niej kolejny raz. Wysyłamy jedynie to, co się zmieniło,
- 2. wielu połączeń TCP wiele połączeń TCP oznacza straty związane z czasem wymaganym do nawiązania połączenia. Jest to sczególnie widoczne przy pobieraniu małych zasobów, gdzie czas nawiązania połączenia jest duży w stosunku do czasu wykonania samego zapytania. Niestety w przypadku HTTP/1.1 jest to jedyny sposób na jednoczesne przesyłanie wielu zasobów,
- 3. łączenia plików sposób na ograniczenie liczby połączeń. Dzięki wykorzystaniu narzędzi takich jak webpack możemy ograniczyć liczbę połączeń TCP poprzez łączenie wielu plików danego typu w jeden duży plik. Na przykład gdy mamy wiele plików z arkuszami styli możemy je połączyć w jeden. Jest to tylko mała część możliwości tego pakietu, zainteresowanych odsyłam do strony internetowej[2].

nie jest możlwie stworzenie rozbudowanej aplikacji, która działałaby w sposób satysfakcjonujący użytkownika. Gdyby zmiany, które ma wprowadzić protokół HTTP/2 faktycznie pozwalały zapomnieć o wspomnianych środkach, to życie wielu developerów stałoby się dużo łatwiejsze. Dzięki temu mogliby się oni ten czas poświęcić na rozwój aplikacji.

Za pomocą własnoręcznie stworzonej aplikacji chcę przekonać się, czy wprowadzone funkcje faktycznie mają tak ogromny wpływ na szybkość komunikacji pomiędzy klientem i serwerem.

Swoją aplikację stworzyłem wykorzystując zestaw oprogramowania MEAN – MongoDB, Express.js, Angular i Node.js.

- MongoDB baza danych NoSQL,
- Express.js framework Node.js do tworzenia aplikacji sieciowych od strony serwera. Udostępnia on wiele metod ułatwiających obsługę zapytań HTTP, routing zapytań, renderowanie widoków HTML,
- Angular framework JavaScript służący do budowy dynamicznych aplikacji internetowych od strony użytkownika,
- Node.js środowisko uruchumieniowe języka JavaScript, które pozwala wystartować serwer.

Zdecydowałem się na to rozwiązanie z kilku powodów:

- po przejrzeniu dostępnych w sieci informacji doszedłem do wniosku, że implementacja protokołu HTTP/2 jest najlepiej opisana oraz wspierana przez środowisko związane z JavaScriptem,
- dobra znajomość języka JavaScript oraz jednoczesna chęć rozwoju umiejętności tworzenia aplikacji w tym języku,
- chęć poszerzenia wiedzy dotyczącej budowania aplikacji internetowych za pomocą technologii javascriptowych,
- nie ukrywam, że znaczący wpływ na moją decyzję miała również popularność języka JavaScript na rynku pracy.

Dodatkowo, poza narzędziami składającymi się na zestaw MEAN, wykorzystałem następujące biblioteki:

- 1. node-spdy zewnętrzny moduł do node.js, który umożliwia tworzenie serwerów wspierających HTTP/2. Jest on kompatybilny z biblioteką Express.js, którą wykorzystuję w swoim projekcie. Dzięki temu modułowi możemy zaimplementować serwer HTTP/2 wraz z Server Push. Pomimo, że nie jest to oficjalny moduł node.js, to trwające obecnie prace, które mają na celu wdrożenie HTTP/2 oficjalnie do node.js bazują na tej bibliotece,
- 2. mongoose ułatwia modelowanie danych mongoDB, walidację oraz pisanie logiki biznesowej.

# HTTP/2

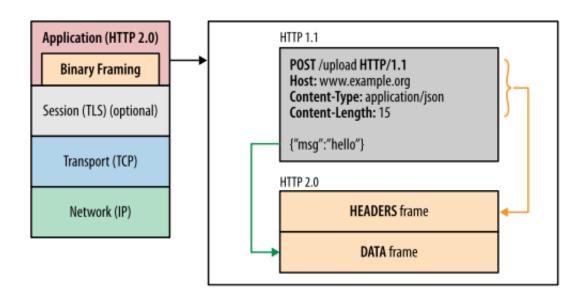
#### 2.1 Historia

Pracę nad zmianami w protokole zapoczątkowała w 2009 roku firma Google ze swoim projektem SPDY. Zdecydowali się oni na stworzenie protokołu, który miał usprawnić działanie aplikacji oraz stron internetowych rozwiązując ograniczenia nałożone przez HTTP/1.1. Z biegiem czasu coraz więcej przeglądarek oraz stron internetowych, zarówno tych dużych jak i tych małych, zaczęło wspierać SPDY, co zainteresowało osoby pracujące nad protokołem HTTP. Zdecydowali się oni wykorzystać dokumentację protokołu SPDY jako początek prac nad własnym protokołem – HTTP/2. Od tego momentu aż do roku 2015, kiedy to standard HTTP/2 został oficjalnie zaakceptowany (TODOodnośnik RFC 7540 i może 7541), projekty były rozwijane równolegle. SPDY było wykorzystywane do testów nowych funkcjonalności, które miały zostać wprowadzone do nowego protokołu HTTP. Niedługo po oficjalnym zaakceptowaniu HTTP/2 ogłoszono, że SPDY nie będzie dalej wspierane.

W kilku poniższych akapitach postaram się przybliżyć zmiany, które zostały wprowadzone do protokołu HTTP.

#### 2.2 Protokół binarny

Kluczową zmianą, która determinuje brak wstecznej kompatybilności z HTTP/1.1, jest przejście na kodowanie binarne przesyłanych wiadomości. Przykładowa ramka widoczna jest na rysunku 2.1 Jest to rozwiązanie dużo bardziej kompaktowe i łatwiejsze w implementacji, niż przesyłanie zwykłego tekstu. Dzięki temu zabiegowi w ramach jednego połączenia TCP z serwerem może zostać utworzonych wiele dwukierunkowych strumieni danych przesy-



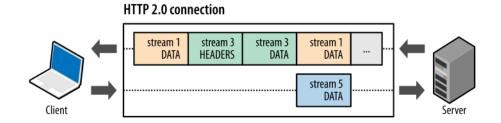
Rysunek 2.1: Schemat ramki protokołu HTTP/2

łających wiadomości HTTP. Taka wiadomość to w rzeczywistości zapytanie od klienta lub odpowiedź serwera składające się z ramek. Każda ramka natomiast musi przynajmniej posiadać nagłówek z informacją, do którego strumienia danych należy. Kodowanie binarne nie ma wpływu na składnie zawartości ramki – wszystkie nagłówki czy zapytania HTTP/1.1 pozostawiono bez zmian.

#### 2.3 Multiplexing

W poprzedniej wersji protokołu, pomimo, że istniała możliwość przesyłania wielu zapytań w ramach jednego połączenia, nie można było wykonywać ich równolegle. Każde zapytanie musiało być rozpatrywane i odesłane przez serwer do klienta zgodnie z kolejnością nadania, co powodowało efekt HOL (head-of-line blocking ODNOSNIK DO JAKIEGOŚ ŹRÓDŁA?). Aby wykonywać zapytania równolegle należało utworzyć kilka zapytań TCP, co obciąża serwer oraz jest czasochłonne. Protokuł HTTP/2 umożlwia przesyłanie oraz odbieranie wielu wiadomości jednocześnie, co pokazuje schemat na rysunku 2.2. Są one rozbijane na pojedyncze ramki, przesyłane, a następnie odczytywane i składane z powrotem w całość po stronie odbiorcy. Dzięki temu nie jest już konieczne uciekanie się do takich zabiegów jak:

• scalanie plików (WEBPACK),



Rysunek 2.2: Schemat wykorzystania multiplexingu w HTTP/2

- wykorzystywanie spritów,
- domain sharding (DOCZYTAC).

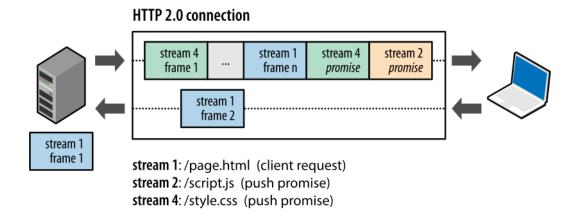
To wszystko sprawia, że aplikacje stają się szybsze oraz prostsze.

#### 2.4 Prioretyzacja

#### 2.5 Server push

Wykorzystując protokół HTTP/1.1 nie mamy możliwości otrzymania zasobu, o który nie poprosiliśmy wysyłając zapytanie. Powoduje to opóźnienia na przykład podczas ładowania strony internetowej. Zanim otrzymamy skrypty czy arkusze styli, które wykorzystuje nasza strona musi ona o nie poprosić. Zapytanie do serwera wysyłane jest gdy w kodzie pliku html napotkamy na taki kod (przykład z mojego projektu):

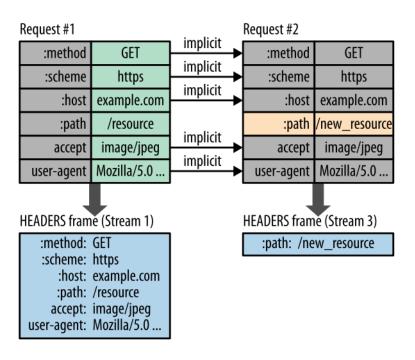
Takie rozwiązanie, chociaż w wielu przypadkach jest pożądane, tutaj jedynie spowalnia działanie aplikacji. Jeżeli mamy pewność, że użytkownik będzie potrzebował danych zasobów 2.3 możemy mu je od razu udostępnić, co zdecydowanie skraca czas ładowania aplikacji i dzięki temu unikam niechcianego efektu, gdy strona się załaduje, ale na przykład bez pliku zawierającego style, który jest dopiero przesyłany.



Rysunek 2.3: Schemat Server push HTTP/2

#### 2.6 Kompresja nagłówków

Kolejną ważną, chociaż pozornie niezauważalną zmianą jest kompresja nagłówków. Pomimo, że problem może wydawać się marginalny, to te kilka bajtów w każdym zapytaniu potrafi dość mocno wydłużyć czas komunikacji z serwerem. HTTP/2 wykorzystuje format kompresji zwany HPACK, który dzięki wykorzystaniu dwóch technik – kodowania oraz indeksowania, jest w stanie znacznie usprawnić komunikację ze względu na nagłówki. Po pierwsze kodowanie Huffmana znacznie kompresuje przesyłane dane. Dodatkowo klient oraz serwer przechowują listę widzianych wcześniej nagłówków i nie ma potrzeby wysyłania za każdym razem całego nagłówka. Pzesyłamy jedynie pola, które się zmieniły (patrz rysunek 2.4).



Rysunek 2.4: Kompresja nagłówków w  $\mathrm{HTTP}/2$ 

# Budowa aplikacji

#### 3.1 node-spdy – konfiguracja serwera

#### 3.2 Server Push

```
// sciezki uproszczone dla przejrzystosci kodu
   app.get('/push', function(req, res) {
     fs.readFile('/push.html', function read(err, data) {
       if (err) {
4
5
         throw err;
6
7
       content = data;
       var stream = res.push('/push.html', {
8
         status:\ 200\,,\ //\ optional
9
         method: 'GET', // optional
10
         request: { accept: '*/*' },
11
         response: { 'content-type': 'text/html; charset=UTF-8' }
12
13
       })
       stream.on('error', function() {})
14
15
       stream.end(content)
16
       res.write(content)
17
     })
18
     fs.readFile('/jquery.js', function read(err, data) {
19
20
       if (err) {
21
         throw err;
22
23
       content = data;
```

```
var stream1 = res.push('/libs/jquery/dist/jquery.js', {
24
          status: 200, // optional
25
          method: 'GET', // optional
request: { accept: '*/*' },
26
27
          response: { 'content-type': 'application/javascript' }
28
        })
29
        stream1.on('error', function() {})
30
        stream1.end(content)
31
32
        res.end();
     })
33
34 })
```

### Testy

#### 4.1 Chrome DevTools

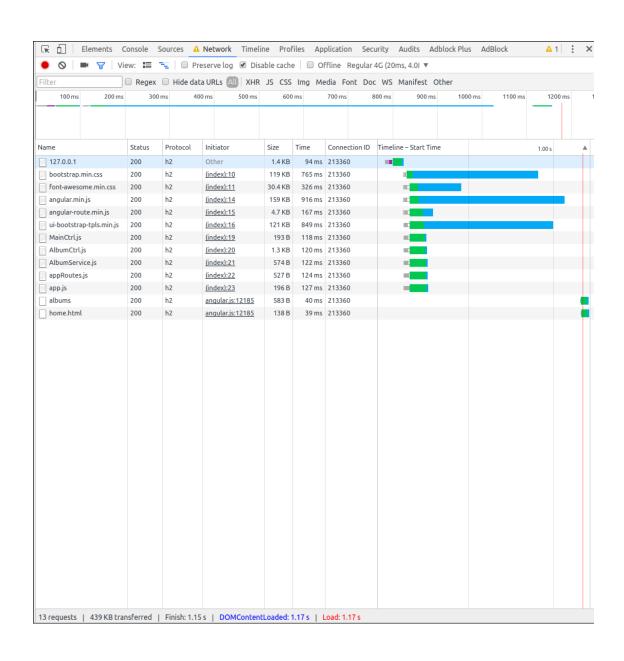
Do pomiaru prędkości oraz uzyskania innych ważnych informacji wykorzystałem narzędzie Chrome DevTools. Opiszę tutaj po krótce co i jak mierzyłem za pomocą tego oprogramowania.

Po uruchomieniu konsoli przeglądarki przechodzimy do zakładki Network i naszym oczom ukazuje się okno jak na rysunku 4.1.

Widoczne okno składa się z pięciu głównych elementów:

- paska kontroli umożliwia on między innymi edycję wyglądu panelu sieciowego,
- 2. paska filtrów pozwala na stworzenie reguł i wybór tylko tych pakietów, które nas interesują,
- 3. paska przeglądu ukazuje nam oś czasu, która daje nam obraz tego, jak przesyłane były pakiety danych,
- 4. tabeli zapytań zawiera szczegółowe informacje na temat każdego zapytania
- 5. podsumowania zawiera informacje o łącznej liczbie zapytań, przesłanych danych oraz czasie trwania.

W moich badaniach najczęściej korzystałem z informacji zawartych w tabeli zapytań, a szczególnie z przedstawionej w niej osi czasu. Po najechaniu kursorem na którykolwiek pasek na osi otrzymujemy szczegółowe informacje o czasie każdego z etapów zapytania jak na rysunku 4.2.



Rysunek 4.1: Wygląd okna Chrome DevTools



Rysunek 4.2: Szczegółowe informacje na temat czasu zapytania

#### 4.2 Testy wstępne

Przed przejściem do porównania prędkości działania obu wersji protokołu chciałem zaprezentować pierwsze efekty implementacji protokołu HTTP/2, które przedstawia rysunek 4.3.



Rysunek 4.3: Dowód działania protokołu HTTP/2

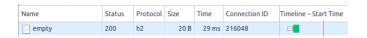
Widzimy tutaj dwie rzeczy, które powinny nas zainteresować. W sekcji 'Protocol' oznaczonej na rysunku 4.3 numerem 1 widnieje napis h2 przy każdym zapytaniu. Jest to informacja, że do komunikacji z serwerem wykorzystana została najnowsza wersja protokołu HTTP. Dodatkowo w sekcji 'Connection ID' (na rysunku 4.3 jest to numer 2) widzimy, że wszystkie zapytania zostały wykonane z wykorzystaniem tego samego połączenia TCP. Nie mogłoby to mieć miejsca, gdybyśmy wykorzystali HTTP/1.1, co pokazuje rysunek 4.4.



Rysunek 4.4: Przykładowe linia czasu dla zapytań wykonanych w HTTP/1.1

Na rysunku widzimy, że protokół z jakiego korzystamy to HTTP/1.1, a wszystkie zapytania, które są wykonywane równolegle przesyłane są w ramach różnych połączeń TCP. Świadczą o tym wartości w kolumnie 'Connection ID'.

Jeszcze jedną ciekawą rzeczą są wielkości nagłówków. Dzięki przejściu na kodowanie binarne nagłówki w protokole HTTP/2 są o wiele mniejsze niż te w jego starszym bracie. W swojej pracy stworzyłem możliwość wysłania zapytania, które w odpowiedzi dostaje odpowiedź w postaci samego nagłówka. Dzięki temu można porównać wielkość nagłówków HTTP na konkretnym przykładzie, Wyniki przedstawiaja rysunki 4.5 i 4.6.



Rysunek 4.5: Pusta odpowiedź w HTTP/2



Rysunek 4.6: Pusta odpowiedź w HTTP/1.1

Od razu widać, że nagłówki w najnowszej wersji protokołu uległy znacznej kompresji.

# 4.3 Porównanie prędkości protokołu HTTP/2 i HTTP/1.1

W ramach tego testu sprawdzałem prędkość ładowania wszystkich zasobów strony z wyłączonym CACHE przeglądarki oraz z ustawioną prędkością Regular 4G (20ms, 4.0 Mb/s, 3.0 Mb/s). Dane z wykorzystaniem protokołu HTTP/1.1 są przesyłane za pomocą niezabezpieczonego połączenia. HTTP/2 do utworzenia połączenia wymaga zabepieczenia SSL.

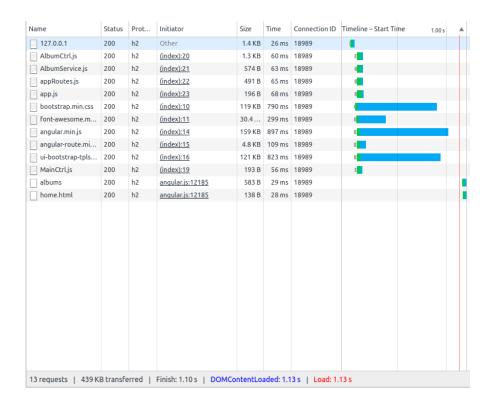
Próba	HTTP/1.1	HTTP/2
1	1.23s	1.26s
2	1.49s	1.12s
3	1.28s	1.39s
4	1.20s	1.09s
5	1.22s	1.09s
6	1.25s	1.26s
7	1.21s	1.12s
8	1.25s	1.11s
9	1.21s	1.16s
10	1.26s	1.10s
ŚREDNIA	1.26s	1.17s

Tabela ukazuje wyniki wykonanych pomiarów. Jak widać protokół HTTP/2 przyśpiesza czas ładowania strony o około 15W dużej mierze jest to czas zaoszczędzony na nawiązywanie nowych połączeń TCP. Sporą oszczędność utrzymujemy też dzięki zmniejszeniu objętości nagłówków odpowiedzi.

Na rysunkach 4.7 oraz 4.8 widoczny jest przykładowy wynik testu. W sekcji podsumowania widoczna jest całkowita objętość przesłanych danych. Są to odpowiednio 443 KB i439 KB dla protokołu HTTP/1.1 i HTTP/2, Możemy też porównać objętości oraz czasy przesyłania poszczególnych plików.

# 4.4 Porównanie prędkości przy bezpiecznym połączeniu HTTP/1.1

W tym teście dane przesyłane za pomocą protokołu  $\mathrm{HTTP}/1.1$  są zabepieczone SSL.



Rysunek 4.7: Test dla HTTP/2

lame	Status	Protocol	Initiator	Size	Time	Connection ID	Timeline – Start Time	
127.0.0.1	200	http/1.1	Other	1.6 KB	34 ms	21718		
bootstrap.min.css	200	http/1.1	(index):10	119	819 ms	21718		
font-awesome.m	200	http/1.1	(index):11	30.6	304 ms	21755		
angular.min.js	200	http/1.1	(index):14	159	910 ms	21764		
angular-route.mi	200	http/1.1	(index):15	4.9 KB	95 ms	21760	•	
ui-bootstrap-tpls	200	http/1.1	(index):16	121	834 ms	21768		
MainCtrl.js	200	http/1.1	(index):19	406 B	54 ms	21751	•	
AlbumCtrl.js	200	http/1.1	(index):20	1.5 KB	108 ms	21751	4	
AlbumService.js	200	http/1.1	(index):21	787 B	125 ms	21760	-	
appRoutes.js	200	http/1.1	(index):22	704 B	140 ms	21751	-	
app.js	200	http/1.1	(index):23	409 B	168 ms	21760		
albums	200	http/1.1	angular.js:12185	707 B	28 ms	21764		
home.html	200	http/1.1	angular.js:12185	334 B	36 ms	21768		
favicon.ico	200	http/1.1	Other	1.6 KB		21768		

Rysunek 4.8: Test dla  $\mathrm{HTTP}/1.1$ 

Próba	HTTP/1.1	HTTP/2
1	1.12s	1.07s
2	1.20s	1.09s
3	1.05s	1.07s
4	1.12s	1.09s
5	1.08s	1.06s
6	1.10s	1.11s
7	1.13s	1.16s
8	1.09s	1.15s
9	1.08s	1.12s
10	1.11s	1.10s
ŚREDNIA	1.108s	1.102s

Podczas tego testu otrzymałem zaskakujące wyniki. Całkowity czas ładowania się zasobów strony dla protokołu HTTP/1.1 z wykorzystaniem SSL jest krótszy, niż przy nieszyfrowanym połączeniu. Jest to zaskakujące, ponieważ szyfrowanie wymusza na nas podjęcie dodatkowych kroków na początku transmisji, które zabierają czas.

# 4.5 Porównanie prędkości dla HTTP/1.1 z włączononym CACHE

Próba	HTTP/1.1	HTTP/2
1	$410 \mathrm{ms}$	$434 \mathrm{ms}$
2	$421 \mathrm{ms}$	$383 \mathrm{ms}$
3	$420 \mathrm{ms}$	$427 \mathrm{ms}$
4	$374 \mathrm{ms}$	$411 \mathrm{ms}$
5	$356 \mathrm{ms}$	$418 \mathrm{ms}$
6	$387 \mathrm{ms}$	$455 \mathrm{ms}$
7	$313 \mathrm{ms}$	$457 \mathrm{ms}$
8	$517 \mathrm{ms}$	$396 \mathrm{ms}$
9	$468 \mathrm{ms}$	$431 \mathrm{ms}$
10	$344 \mathrm{ms}$	$432 \mathrm{ms}$
ŚREDNIA	$401 \mathrm{ms}$	424ms

### 4.6 Porównanie prędkości przy wykorzystaniu Server Push

Próba	HTTP/1.1	HTTP $/2$
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
ŚREDNIA		

# Wnioski

# Bibliografia

- [1] M. Belshe, BitGo, R. Peon, Google, Inc, M. Thomson, Ed. Mozilla, "Hypertext Transfer Protocol Version 2 (HTTP/2)", 2015. https://tools.ietf.org/html/rfc7540
- [2] webpack concept https://webpack.js.org/concepts/