

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

UMREŽAVANJE SADRŽAJA  
**1. DOMAĆA ZADAĆA**

ZAGREB, 15. listopada 2013.

## SADRŽAJ

UVOD.....	2
ZADATCI.....	3
1. Zadatak .....	3
2. Zadatak .....	5
3. Zadatak .....	7
4. Zadatak .....	10
5. Zadatak .....	11
6. Zadatak .....	12
7. Zadatak .....	15
8. Zadatak .....	17
9. Zadatak .....	19

## UVOD

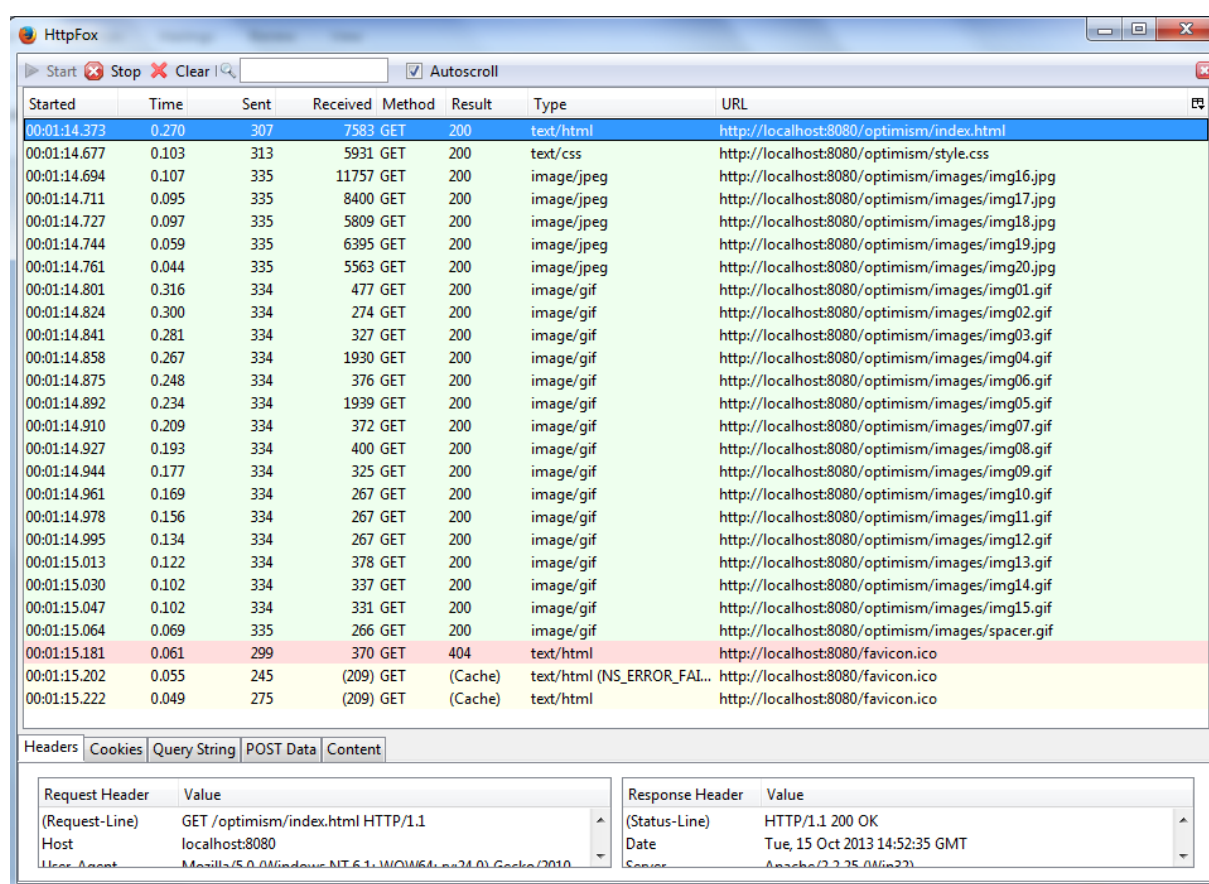
Zadatak ove domaće zadaće je upoznavanje s radom posredničkog poslužitelja weba s priručnim spremištem kao i sa svojstvima i sadržajem HTTP paketa koji se izmjenjuju između preglednika korisnika i poslužitelja weba. Ova domaća zadaća, dakle, omogućuje detaljniji uvid u komunikaciju poslužitelja i preglednika na kojoj počiva cijeli web kao i na metode i načine privremenog pohranjivanja podataka u priručna spremišta s ciljem brže i učinkovitije razmjene podataka (digitalnih resursa) između različitih entiteta korisnika i poslužitelja.

Za uspješnu realizaciju ove domaće zadaće potrebno je preuzeti i instalirati poslužitelj weba *Apache*, zatim posredničkog poslužitelja s priručnim spremištem *Squid* te web preglednik Firefox s ekstenzijama za analizu i modifikaciju HTTP paketa (*HTTPFox* i *Modify Headers*).

# ZADATCI

## 1. Zadatak

U prvom zadatku odabiremo opciju No Proxy u postavkama mreže za preglednik Firefox. Rezultat toga je isključivanje (nekorištenje) posredničkog poslužitelja s priručnim spremištem. Nakon otvaranja stranice <http://localhost:8080/optimism/index.html> pomoću dodatka HTTPFox uočavamo sljedeće HTTP pakete:



Started	Time	Sent	Received	Method	Result	Type	URL
00:01:14.373	0.270	307	7583	GET	200	text/html	http://localhost:8080/optimism/index.html
00:01:14.677	0.103	313	5931	GET	200	text/css	http://localhost:8080/optimism/style.css
00:01:14.694	0.107	335	11757	GET	200	image/jpeg	http://localhost:8080/optimism/images/img16.jpg
00:01:14.711	0.095	335	8400	GET	200	image/jpeg	http://localhost:8080/optimism/images/img17.jpg
00:01:14.727	0.097	335	5809	GET	200	image/jpeg	http://localhost:8080/optimism/images/img18.jpg
00:01:14.744	0.059	335	6395	GET	200	image/jpeg	http://localhost:8080/optimism/images/img19.jpg
00:01:14.761	0.044	335	5563	GET	200	image/jpeg	http://localhost:8080/optimism/images/img20.jpg
00:01:14.801	0.316	334	477	GET	200	image/gif	http://localhost:8080/optimism/images/img01.gif
00:01:14.824	0.300	334	274	GET	200	image/gif	http://localhost:8080/optimism/images/img02.gif
00:01:14.841	0.281	334	327	GET	200	image/gif	http://localhost:8080/optimism/images/img03.gif
00:01:14.858	0.267	334	1930	GET	200	image/gif	http://localhost:8080/optimism/images/img04.gif
00:01:14.875	0.248	334	376	GET	200	image/gif	http://localhost:8080/optimism/images/img06.gif
00:01:14.892	0.234	334	1939	GET	200	image/gif	http://localhost:8080/optimism/images/img05.gif
00:01:14.910	0.209	334	372	GET	200	image/gif	http://localhost:8080/optimism/images/img07.gif
00:01:14.927	0.193	334	400	GET	200	image/gif	http://localhost:8080/optimism/images/img08.gif
00:01:14.944	0.177	334	325	GET	200	image/gif	http://localhost:8080/optimism/images/img09.gif
00:01:14.961	0.169	334	267	GET	200	image/gif	http://localhost:8080/optimism/images/img10.gif
00:01:14.978	0.156	334	267	GET	200	image/gif	http://localhost:8080/optimism/images/img11.gif
00:01:14.995	0.134	334	267	GET	200	image/gif	http://localhost:8080/optimism/images/img12.gif
00:01:15.013	0.122	334	378	GET	200	image/gif	http://localhost:8080/optimism/images/img13.gif
00:01:15.030	0.102	334	337	GET	200	image/gif	http://localhost:8080/optimism/images/img14.gif
00:01:15.047	0.102	334	331	GET	200	image/gif	http://localhost:8080/optimism/images/img15.gif
00:01:15.064	0.069	335	266	GET	200	image/gif	http://localhost:8080/optimism/images/spacer.gif
00:01:15.181	0.061	299	370	GET	404	text/html	http://localhost:8080/favicon.ico
00:01:15.202	0.055	245	(209)	GET	(Cache)	text/html (NS_ERROR_FAL...	http://localhost:8080/favicon.ico
00:01:15.222	0.049	275	(209)	GET	(Cache)	text/html	http://localhost:8080/favicon.ico

Request Header		Response Header	
Request Header	Value	Response Header	Value
(Request-Line)	GET /optimism/index.html HTTP/1.1	(Status-Line)	HTTP/1.1 200 OK
Host	localhost:8080	Date	Tue, 15 Oct 2013 14:52:35 GMT
User-Agent	Mozilla/5.0 (Windows NT 6.1; WOW64; rv:24.0) Gecko/2010...	Server	Apache/2.2.25 (Ubuntu)

Slika 1 Prvi HTTP paketi dobiveni tijekom pristupanja stranici index.html

U svakom retku je naveden po jedan GET zahtjev i njegov odgovor (200 OK), dakle po dva paketa u jednom retku što je različito u odnosu na npr. Wireshark. Vrijednosti zaglavlja određenih paketa možemo pogledati u donjem dijelu prozora, a također nam se nude i opcije pregledavanja tzv. „kolačića“ (Cookies), zatim Query String, POST Data i Content.

Lako uočavamo da je poslano **26 zahtjeva** (GET zahtjevi), iako se vrlo vjerojatno samo prvih **23** odnose na cjelokupnu stranicu index.html, od strane preglednika prema poslužitelju. Prvo preglednik

šalje GET zahtjev za dohvaćanjem početne stranice *index.html*. Nakon uspješno primljenog odgovora, saznaje da mora dohvatiti još resursa (u *Contentu* prvog odgovora se nalazi samo „sirovi“ html s referencama na neke dodatne resurse koje treba naknadno dohvatiti) te šalje zahtjev za dohvaćanjem CSS datoteke u kojoj je opisan razmještaj HTML elemenata (naziv datoteke je *style.css*). Nakon uspješno primljenog odgovora, preglednik sada šalje veći niz GET zahtjeva za dohvaćanjem razno-raznih slika koje treba prikazati na zaslonu. Te slike su redom *img16.jpg*, *img17.jpg*, *img18.jpg* itd. (postoje i *gif* slike, *img01.gif*, *img02.gif* itd.).

Na kraju preglednik šalje GET zahtjev za dohvaćanjem resursa *favicon.ico*, ali prima odgovor 404 (Not Found) što ukazuje na to da resurs nije nađen, zatim šalje još dva GET zahtjeva čiji je rezultat privremeno pohranjivanje (za prvi od ta dva paketa se prikazuje *NS\_ERROR\_FAILURE*). Posljednja tri zahtjeva vjerojatno nisu povezana s izvornom stranicom *index.html* i njihovo porijeklo nije u potpunosti utvrđeno.

Ako promotrimo npr. prvi GET zahtjev kojeg šalje preglednik, uočavamo sljedeća zaglavlja:

- Request Line – sadrži HTTP metodu (u ovom slučaju GET), zatim URL resursa te verziju HTTP protokola
- Host – domena servera (u ovom slučaju *localhost:8080*), može biti i IP adresa
- User Agent – opis preglednika i neke opće informacije o operacijskom sustavu
- Accept – tipovi podataka (medija) i resursa koji se prihvataju
- Accept – Language – prihvaćeni (podržani) prirodni jezik
- Accept – Encoding – podržano kodiranje podataka
- Connection – omogućuje pošiljatelju da specificira određene opcije koje vrijede za konekciju uspostavljenu između njega i primatelja zahtjeva (to može biti poslužitelj najčešće)

*Keep-Alive* služi da se konekcija i dalje drži aktivnom (alive) i nakon primanja i obrade odgovora, dakle veza se neće terminirati između preglednika i poslužitelja čak i nakon zaprimanja odgovora i njegove potpune obrade. Ideja je da se više HTTP zahtjeva i odgovora šalje preko jedne TCP konekcije čime se postiže manje opterećenje mreže TCP paketima (ne mora se stalno uspostavljati TCP konekcija). *Keep-Alive* polje u odgovoru servera može sadržavati dodatne opcije vezane za održavanje veze (objašnjeno dolje).

Zaglavlja odgovora su sljedeća:

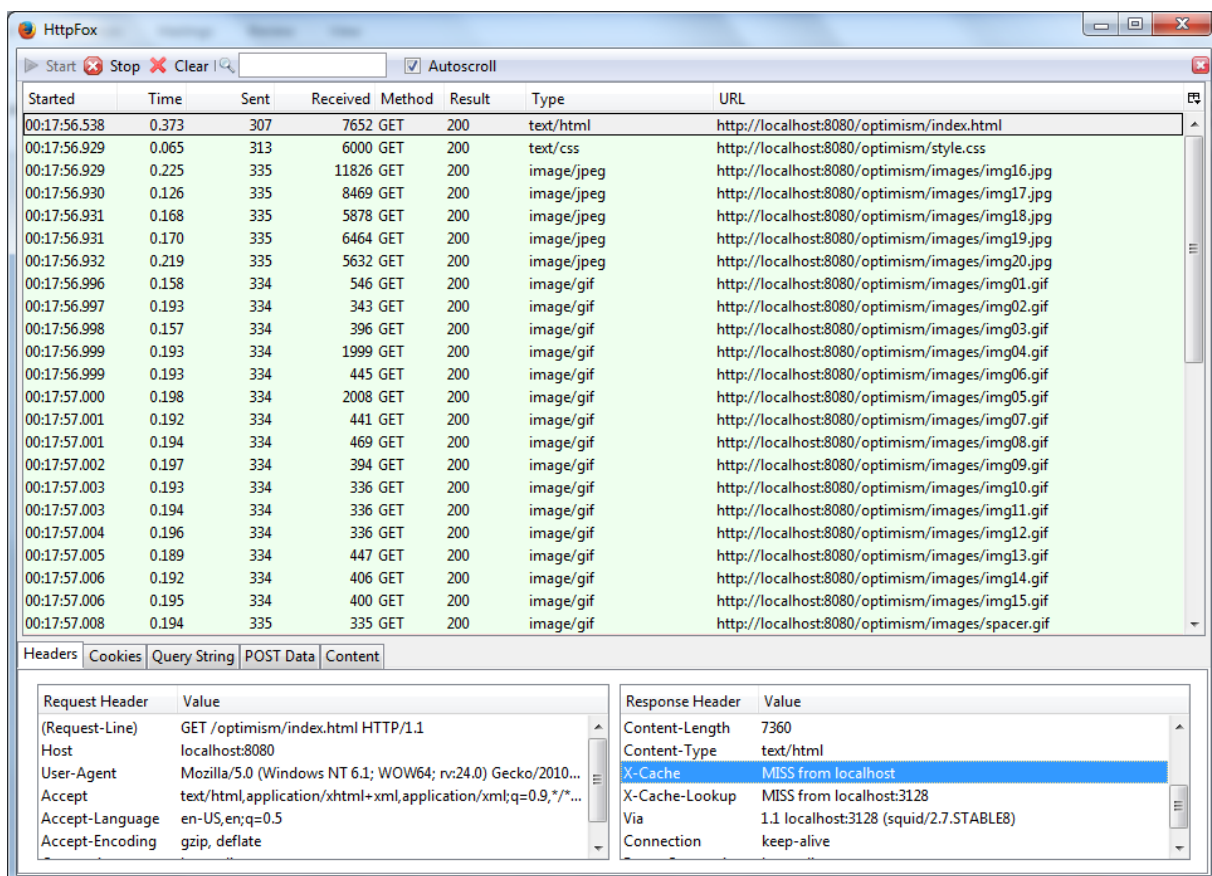
- Status-Line – sadrži status odgovora i verziju HTTP protokola
- Date – datum slanja odgovora
- Server – informacije o serveru koji šalje odgovor
- Last Modified – datum zadnje promjene varijante resursa koji se isporučuje u odgovoru
- Etag – sadrži trenutnu vrijednost oznake entiteta za zahtijevanu varijantu, može poslužiti za usporedbu s drugim entitetima istog digitalnog resursa
- Accept-Ranges – podržani raspon zahtjeva za određenim resursima
- Content-Length – duljina odgovora koji se isporučuje kao *content*

- Keep-Alive – dodatni parametri vezani uz opciju održavanja veze aktivnom u polju Connection (npr. server prihvata maksimalno 100 zahtjeva, a ako ne primi sljedeći zahtjev u roku od 5 sekundi terminira vezu, kao što je slučaj u prvom odgovoru)
- Connection – funkcija slična kao i kod zahtjeva (opisana gore)
- Content-Type – vrsta (tip) podataka odnosno digitalnog resursa čija se varijanta isporučuje

Etag polje se prvenstveno koristi za usporedbu *hash* vrijednosti originala i kopije nekog digitalnog resursa. Pomoću ovog *hasha* (i još drugih parametara) može se zaključiti je li kopija nekog resursa (obično se nalazi u priručnom spremištu) zastarjela u odnosu na original.

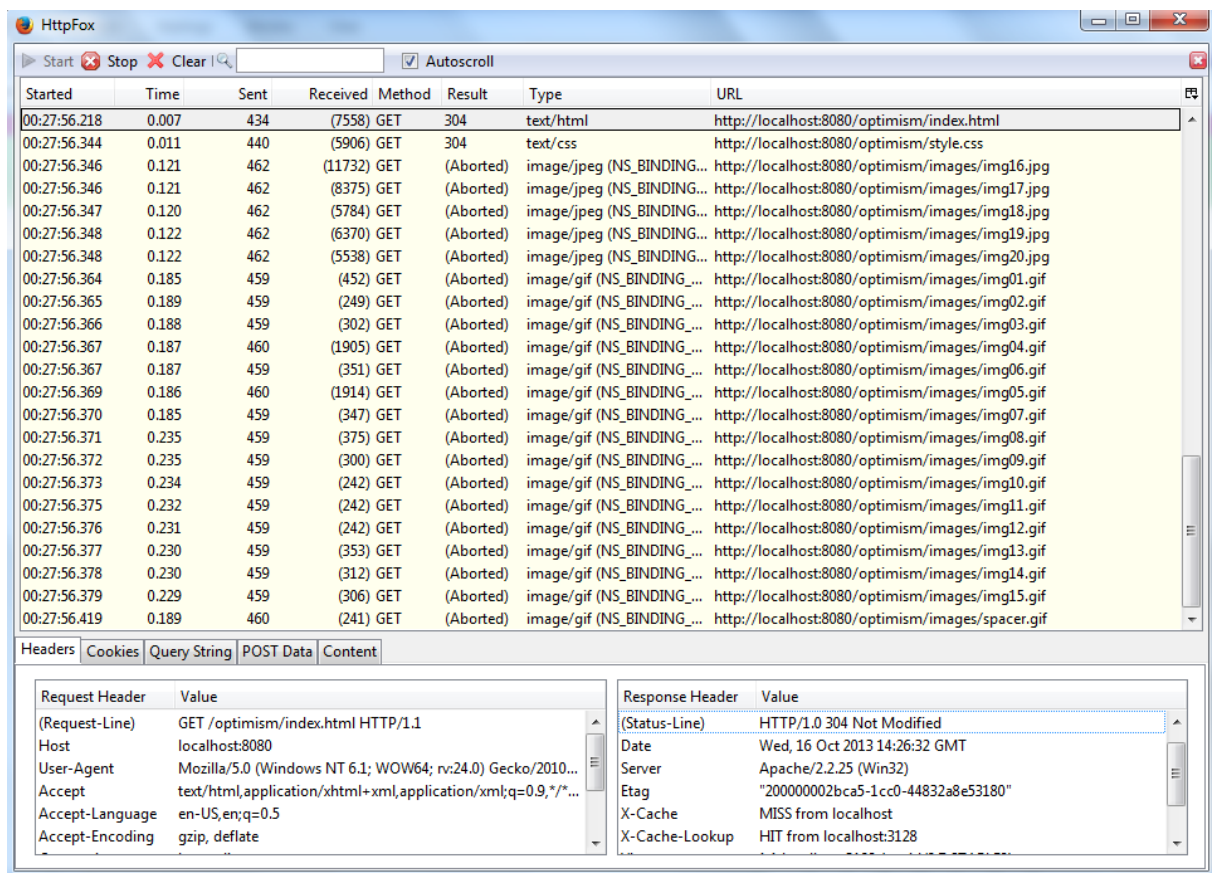
## 2. Zadatak

Budući da je priručno spremište *Firefox*a i *Squida* prazno, nakon pristupanja stranici `index.html` dohvaćaju se svi potrebni resursi kao i u zadatku 1. Dakle, za svaki GET zahtjev izvorni poslužitelj šalje kopiju resursa. Stoga preglednik ne navodi polja `Is-Modified-Since` i `If-None-Match` jer kopije tih resursa i njihove `Etag`ove još nema. Izvorni poslužitelj vraća tražene sadržaje (kopije resursa), a uočavamo i polje `X-Cache` čija je vrijednost `MISS from localhost`, a također i polje `X-Cache-Lookup` s vrijednošću `MISS from localhost:3128`. Značenja tih polja će biti objašnjena u nastavku. Sljedeća slika prikazuje prvi poslani i primljeni HTTP paket s označenim poljem `X-Cache` u odgovoru poslužitelja.



Slika 2 HTTP paket s označenim X-Cache poljem

Ako pak ponovimo postupak pristupanja istoj toj stranici, dobit ćemo sljedeći prikaz u dodatku HTTPFox:



Slika 3 Prikaz u dodatku HTTPFox nakon ponovnog pristupanja stranici index.html

Sada se u zahtjevu kojeg šalje preglednik pojavljuje polje `If-Modified-Since` (datum i vrijeme) i `If-None-Match` (Etag resursa) čija je uloga objašnjena u nastavku. Ako analiziramo odgovor koje šalje poslužitelj, vidimo da se vraća statusni kod `304 Not Modified` što ukazuje na to da se traženi resurs nije promijenio od datuma navedenog u polju `If-Modified-Since` (hash kod je isti i kao Etag poslan u polju `If-None-Match` što također ukazuje na istovjetnost kopije i originala) te se promijenilo polje `X-Cache-Lookup` u odgovoru poslužitelja te ono sada ima vrijednost `HIT from localhost:3128` što ukazuje na to da je preglednik dobio podatke iz svog priručnog spremišta (lokalno spremište Firefox) te izvorni poslužitelj nije morao ponovno slati kopiju resursa.

Polje `If-Modified-Since` služi za slanje uvjetnog GET zahtjeva od klijenta prema poslužitelju, a vrijednost tog polja je datum i vrijeme i ako je traženi resurs promijenio vrijednost od tog vremena, poslužitelj isporučuje kopiju resursa, u suprotnom vraća odgovor `304 Not Modified`. Dakle, time je omogućeno provjeravanje istovjetnosti kopije resursa i njegovog originala.

Polje `If-None-Match` se također koristi u uvjetnom GET zahtjevu i pomoću tog polja klijent može poslati `Etag` resursa kojeg želi ponovno dohvatiti (ili provjeriti je li promijenjen), a kojeg je već dohvatio u prošlosti i dobio njegov `Etag`. Po primitku tog zahtjeva, poslužitelj uspoređuje primljeni `Etag` s oznakom izvornog resursa i ako je došlo do promjene (oznake se ne poklapaju), šalje kopiju traženog resursa, u suprotnom šalje odgovor `304 Not Modified`.

Polje `X-Cache` u odgovoru definira je li se kopija traženog resursa nalazi u priručnom spremištu posredničkog poslužitelja i je li isporučena od njega (vrijednost `HIT` ...) ili nije isporučena od njega (vrijednost `MISS`). Dakle, definira se dostupnost kopije nekog resursa iz perspektive posredničkog poslužitelja. Posrednički poslužitelj `squid` definira ovo polje zaglavlja.

Polje `X-Cache-Lookup` čuva rezultate pretraživanja posebne tablice pohranjivanja prije nego što se provjeravaju pravila za ažuriranje kopije resursa. Ovo polje se također nalazi u odgovoru kojeg šalje `squid` i može pobliže odrediti razlog zbog kojeg se pojavio `MISS` u polju `X-Cache`. Naime, ovo polje na govori što bi posredničko priručno spremište učinilo kad bi se zatražila određena kopija resursa od njega, ali se u stvarnosti traži od nekog drugog izvora. Promotrimo sljedeći primjer: ako u zaglavlju odgovora nađemo sljedeća polja

```
X-Cache: MISS from localhost
```

```
X-Cache-Lookup: HIT from localhost:3128
```

to znači da naš posrednički poslužitelj (`Squid`) nije isporučio kopiju traženog resursa, ali kad bi morao isporučiti, mogao bi jer ima tu kopiju kod sebe. Ova kombinacija vrijednosti polja `X-Cache` i `X-Cache-Lookup` najčešće ukazuje na to da je preglednik dohvatio resurs iz svog lokalnog spremišta.

U slučaju ponovnog pristupa stranici `index.html`, priručno spremište je kontaktiralo originalnog poslužitelja samo kako bi provjerilo valjanost traženog resursa. Originalni poslužitelj ne šalje ponovno kopije, već Firefox nakon završenog procesa revalidacije vuče sadržaj iz svog spremišta.

### 3. Zadatak

Način postavljanja vremena valjanosti za JPG slike i stranicu `index.html` bit će objašnjen u 4. zadatku.

Nakon brisanja priručnih spremnika preglednika i posredničkog poslužitelja, dobivamo isti rezultat i prikaz u dodatku `HTTPFox` kao i u prvom zadatku i na početku drugog zadatka. Za svaki potrební resurs se šalje GET zahtjev, a poslužitelj odgovara s `200 OK` i kopijom traženog resursa jer preglednik i posrednik inicijalno nemaju kopije tih resursa i njihove `Etagove`.

Pojavljuje se polje `Expires` u zaglavlju odgovora koji nose kopije resursa na koje je stavljeno vrijeme valjanosti, a čija je vrijednost datum i vrijeme kada valjanost resursa istječe i ono je jednako 5 minuta nakon pristupa tom resursu (također nalazimo i polje `Cache-Control` s vrijednošću `max-`



age=300 što znači da je maksimalna valjanost kopije resursa 300 sekundi, odnosno upravo 5 minuta).

Nakon ponovnog pristupa stranici nakon 5 minuta, dobivamo odgovore 304 Not Modified i za one datoteke koje imaju vrijeme valjanosti 5 minuta i one koje nemaju što je vidljivo na sljedeće dvije slike (prva slika prikazuje odgovor na zahtjev za dohvaćanjem stranice index.html za čije je vrijeme valjanosti postavljeno 5 minuta, a druga prikazuje odgovor za jednu gif sliku koja nema vrijeme valjanosti).

Started	Time	Sent	Received	Method	Result	Type	URL
00:06:00.369	0.388	466	1748	POST	200	application...	http://ocsp.thawte.com/
00:06:00.488	0.412	466	1748	POST	200	application...	http://ocsp.thawte.com/
00:08:32.596	0.008	434	(7619)	GET	304	text/html	http://127.0.0.1/index.html
00:08:32.669	0.041	440	(5906)	GET	304	text/css	http://127.0.0.1/ism/style.css
00:08:32.670	0.100	462	(11732)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img16.jpg
00:08:32.671	0.132	462	(8375)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img17.jpg
00:08:32.672	0.131	462	(5784)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img18.jpg
00:08:32.673	0.133	462	(6370)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img19.jpg
00:08:32.673	0.133	462	(5538)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img20.jpg
00:08:32.719	0.135	459	(452)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img01.gif
00:08:32.721	0.165	459	(249)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img02.gif
00:08:32.722	0.164	459	(302)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img03.gif
00:08:32.723	0.163	460	(1905)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img04.gif
00:08:32.725	0.162	459	(351)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img06.gif

Request Header	Value	Response Header	Value
(Request-Line)	GET /optimism/index.html HTTP/1.1	(Status-Line)	HTTP/1.0 304 Not Modified
Host	127.0.0.1:8080	Date	Sun, 20 Oct 2013 13:31:12 GMT
User-Agent	Mozilla/5.0 (Windows NT 6.1; WOW64; rv:11.0) Gecko:20100101 Firefox/11.0	Server	Apache/2.2.25 (Win32)
Accept	text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8	Etag	"200000002bca5-1ccf-4e8de9a64e1a..."
Accept-Language	en-US,en;q=0.5	Expires	Sun, 20 Oct 2013 13:36:12 GMT
Accept-Encoding	gzip, deflate	Cache-Control	max-age=300

Slika 4 Statusni kod odgovora za html resurs koji ima vrijeme valjanosti

Started	Time	Sent	Received	Method	Result	Type	URL
00:06:00.369	0.388	466	1748	POST	200	application...	http://ocsp.thawte.com/
00:06:00.488	0.412	466	1748	POST	200	application...	http://ocsp.thawte.com/
00:08:32.596	0.008	434	(7619)	GET	304	text/html	http://127.0.0.1/index.html
00:08:32.669	0.041	440	(5906)	GET	304	text/css	http://127.0.0.1/ism/style.css
00:08:32.670	0.100	462	(11732)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img16.jpg
00:08:32.671	0.132	462	(8375)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img17.jpg
00:08:32.672	0.131	462	(5784)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img18.jpg
00:08:32.673	0.133	462	(6370)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img19.jpg
00:08:32.673	0.133	462	(5538)	GET	(Aborted)	image/jpeg...	http://127.0.0.1/img20.jpg
00:08:32.719	0.135	459	(452)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img01.gif
00:08:32.721	0.165	459	(249)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img02.gif
00:08:32.722	0.164	459	(302)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img03.gif
00:08:32.723	0.163	460	(1905)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img04.gif
00:08:32.725	0.162	459	(351)	GET	(Aborted)	image/gif (...)	http://127.0.0.1/img06.gif

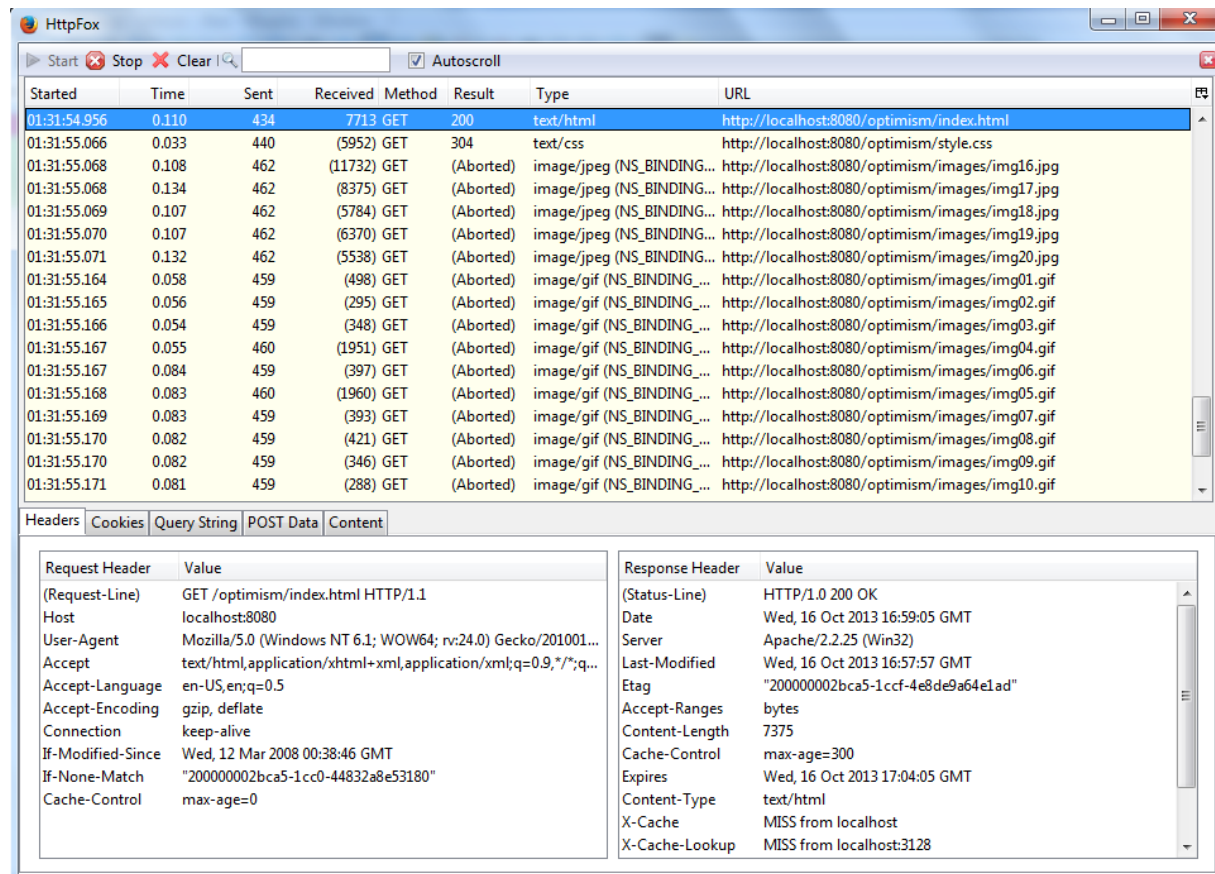
  

Request Header	Value	Response Header	Value
(Request-Line)	GET /optimism/images/img02.gif HTTP/1.1	(Status-Line)	HTTP/1.0 304 Not Modified
Host	127.0.0.1:8080	Date	Sun, 20 Oct 2013 13:31:12 GMT
User-Agent	Mozilla/5.0 (Windows NT 6.1; WOW64; rv:11.0) Gecko:20100101 Firefox/11.0	Server	Apache/2.2.25 (Win32)
Accept	image/png,image/jpeg;q=0.8,*/*;q=0.5	Etag	"100000002bcaa-33-4483055aeed00"
Accept-Language	en-US,en;q=0.5	X-Cache	MISS from localhost
Accept-Encoding	gzip, deflate	X-Cache-Location	HIT from localhost:3128

Slika 5 Statusni kod odgovora za gif sliku koja nema vrijeme valjanosti

Ako pak promijenimo izvorni html kod stranice `index.html` u direktoriju `optimism`, prilikom ponovnog pristupanja stranici dobiva se odgovor statusa 200 OK za prvi resurs (html kod) te se u sadržaju poruke isporučuje kopija tog resursa. Za ostale resurse koji se nisu mijenjali dobivamo odgovor 304 Not Modified.

Sljedeća slika prikazuje dobivene pakete gdje je samo prvi odgovor 200 OK.



Started	Time	Sent	Received	Method	Result	Type	URL
01:31:54.956	0.110	434	7713	GET	200	text/html	http://localhost:8080/optimism/index.html
01:31:55.066	0.033	440	(5952)	GET	304	text/css	http://localhost:8080/optimism/style.css
01:31:55.068	0.108	462	(11732)	GET	(Aborted)	image/jpeg (NS_BINDING_...	http://localhost:8080/optimism/images/img16.jpg
01:31:55.068	0.134	462	(8375)	GET	(Aborted)	image/jpeg (NS_BINDING_...	http://localhost:8080/optimism/images/img17.jpg
01:31:55.069	0.107	462	(5784)	GET	(Aborted)	image/jpeg (NS_BINDING_...	http://localhost:8080/optimism/images/img18.jpg
01:31:55.070	0.107	462	(6370)	GET	(Aborted)	image/jpeg (NS_BINDING_...	http://localhost:8080/optimism/images/img19.jpg
01:31:55.071	0.132	462	(5538)	GET	(Aborted)	image/jpeg (NS_BINDING_...	http://localhost:8080/optimism/images/img20.jpg
01:31:55.164	0.058	459	(498)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img01.gif
01:31:55.165	0.056	459	(295)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img02.gif
01:31:55.166	0.054	459	(348)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img03.gif
01:31:55.167	0.055	460	(1951)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img04.gif
01:31:55.167	0.084	459	(397)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img06.gif
01:31:55.168	0.083	460	(1960)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img05.gif
01:31:55.169	0.083	459	(393)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img07.gif
01:31:55.170	0.082	459	(421)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img08.gif
01:31:55.170	0.082	459	(346)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img09.gif
01:31:55.171	0.081	459	(288)	GET	(Aborted)	image/gif (NS_BINDING_...	http://localhost:8080/optimism/images/img10.gif

Request Header	Value	Response Header	Value
(Request-Line)	GET /optimism/index.html HTTP/1.1	(Status-Line)	HTTP/1.0 200 OK
Host	localhost:8080	Date	Wed, 16 Oct 2013 16:59:05 GMT
User-Agent	Mozilla/5.0 (Windows NT 6.1; WOW64; rv:24.0) Gecko/201001...	Server	Apache/2.2.25 (Win32)
Accept	text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=...	Last-Modified	Wed, 16 Oct 2013 16:57:57 GMT
Accept-Language	en-US,en;q=0.5	Etag	"200000002bca5-1ccf-4e8de9a64e1ad"
Accept-Encoding	gzip, deflate	Accept-Ranges	bytes
Connection	keep-alive	Content-Length	7375
If-Modified-Since	Wed, 12 Mar 2008 00:38:46 GMT	Cache-Control	max-age=300
If-None-Match	"200000002bca5-1ccf-44832a8e53180"	Expires	Wed, 16 Oct 2013 17:04:05 GMT
Cache-Control	max-age=0	Content-Type	text/html
		X-Cache	MISS from localhost
		X-Cache-Lookup	MISS from localhost:3128

Slika 6 HTTP odgovor 200 OK s novom kopijom resursa

U slučaju isteka valjanosti nekog resursa preglednik ili posrednik moraju kontaktirati originalnog poslužitelja s ciljem utvrđivanja istovjetnosti kopije u priručnom spremištu i originala kod poslužitelja. Rezultati pokazuju da se slično radi i za kopije resursa koji nemaju vrijeme valjanosti. U odgovoru za resurse koji imaju vrijeme valjanosti se nalazi i polje `Expires` koje naznačuje da valjanost tog resursa istječe za 5 min. Ako pak ručno promijenimo vrijednost resursa na poslužitelju, odgovor nosi statusni kod 200 OK i u tijelu poruke se isporučuje ažurirana kopija resursa. Tako preglednik može znati da je sadržaj resursa promijenjen.

U prvom slučaju (dok su sva priručna spremišta prazna i prvi se put dohvaća resurs) priručno spremište kontaktira originalni poslužitelj kako bi dobilo potrebne kopije resursa. U drugom slučaju (nakon isteka 5 minuta, prazno priručno spremište posredničkog poslužitelja i postavljeno vrijeme valjanosti slika, html-a i css-a) kontaktira se izvorni poslužitelj kao bi se ustvrdilo je li promijenjena vrijednost izvornog resursa i je li produžena valjanost kopije resursa. Na kraju, kada je izmijenjen izvorni kod stranice `index.html`, kontaktira se poslužitelj za utvrđivanjem promjene vrijednosti i

isteka valjanosti kopije resursa te poslužitelj vraća odgovor 200 OK i novi html kod za resurs `index.html`. dakle, Expires direktiva ne zahtijeva obavezno dohvaćanje nove kopije s izvornog poslužitelja nakon isteka valjanosti, već je dovoljno poslati uvjetni GET zahtjev i poslužitelj može odgovoriti s 304 Not Modified ako taj resurs nije mijenjan i može se onda uzeti kopija iz priručnog spremišta. Ako je resurs mijenjan, obavezno se šalje nova kopija s izvornog poslužitelja.

U 3. Zadatku je brisano priručno spremište posredničkog poslužitelja svaki put kada je ponovno pristupano stranici `index.html` i zbog toga je vrijednost polja `X-Cache` i `X-Cache-Lookup` bila `MISS from localhost ...` U slučaju da to priručno spremište nije brisano, vrijednost prvog polja ostaje ista, a vrijednost polja `X-Cache-Lookup` je sada `HIT from localhost:3128` jer se kopija resursa dohvaća iz lokalnog spremišta.

## 4. Zadatak<sup>1</sup>

Podешavanje vremena valjanosti datoteka koje predstavljaju slike se obavlja modificiranjem konfiguracijske datoteke poslužitelja Apache (ta datoteka je `httpd.conf`). Potrebno je prvo dodati sljedeće linije koda na početak datoteke:

```
LoadModule expires_module modules/mod_expires.so
LoadModule headers_module modules/mod_headers.so
```

U tim modulima se nalaze potrebne funkcije koje omogućuju postavljanje vremena aktivnosti. Sljedeći korak je pronalaženje onog dijela konfiguracijske datoteke koji se odnosi na korijenski direktorij poslužitelja Apache (to je `htdocs`). Taj dio počinje na sljedeći način:

```
<Directory "C:/Program Files (x86)/Apache Software
Foundation/Apache2.2/htdocs">
```

...

Unutar te oznake (elementa) potrebno je dodati sljedeće linije koda:

```
ExpiresActive On
ExpiresByType text/html "access plus 5 minutes"
ExpiresByType image/jpeg "access plus 5 minutes"
```

```
#Ako želimo postaviti vrijeme valjanosti i za gif datoteke, onda treba
#upisati ExpiresByType image/gif "access plus 5 minutes"
```

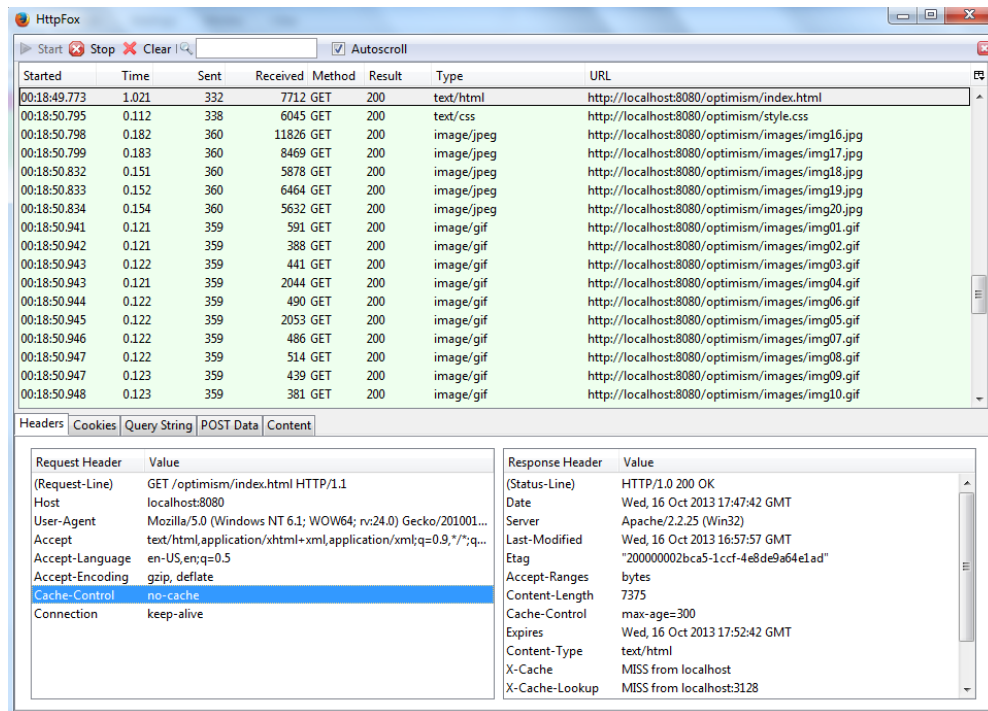
Time se definira vrijeme valjanosti za jpeg slike, te za html datoteku, i to vrijeme iznosi 5 minuta od trenutka pristupa izvornom resursu.

---

<sup>1</sup> <http://www.websiteoptimization.com/secrets/advanced/caching-example.html>

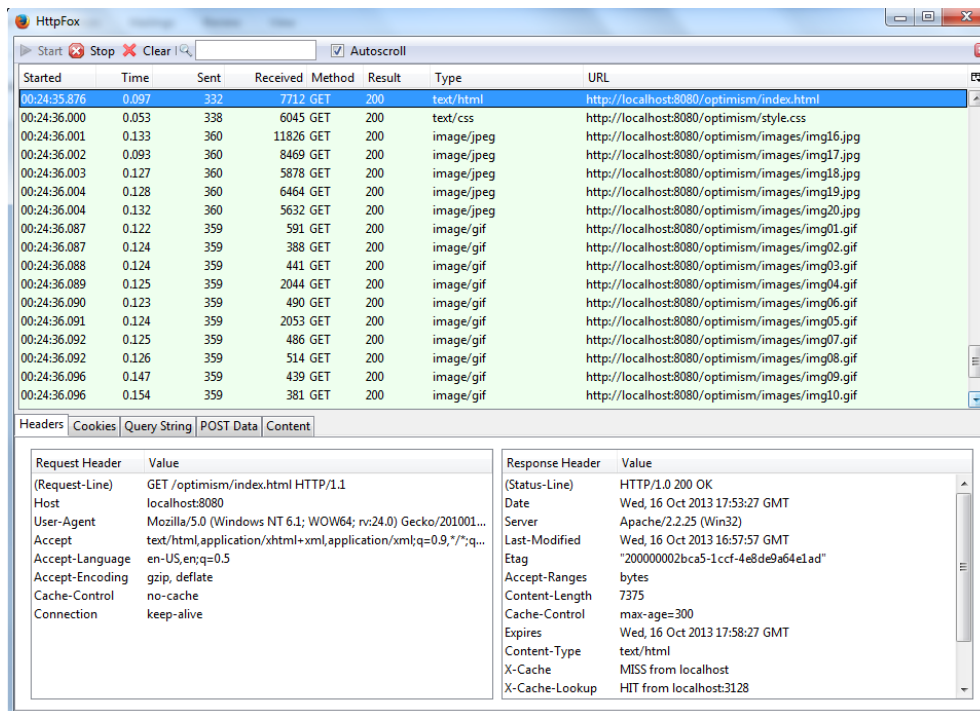
## 5. Zadatak

Nakon brisanja lokalnog spremnika preglednika, priručnog spremnika posrednika i postavljanja zaglavlja `Cache-Control: no-cache` te ponovnog pristupanja stranici `index.html`, dobivamo gotovi identičan rezultat kao i u prvom zadatku (šalje se GET zahtjev za svaki potrební resurs te se vraća odgovor 200 OK s kopijama resursa). To je prikazano na sljedećoj slici.



Slika 7 Označeno polje `Cache-Control` s vrijednošću `no-cache` u zaglavljima klijentovih zahtjeva

Nakon brisanja lokalnog spremnika preglednika i ponovnog pristupanja stranici dobivamo sljedeći prikaz:



Started	Time	Sent	Received	Method	Result	Type	URL
00:24:35.876	0.097	332	7712	GET	200	text/html	http://localhost:8080/optimism/index.html
00:24:36.000	0.053	338	6045	GET	200	text/css	http://localhost:8080/optimism/style.css
00:24:36.001	0.133	360	11826	GET	200	image/jpeg	http://localhost:8080/optimism/images/img16.jpg
00:24:36.002	0.093	360	8469	GET	200	image/jpeg	http://localhost:8080/optimism/images/img17.jpg
00:24:36.003	0.127	360	5878	GET	200	image/jpeg	http://localhost:8080/optimism/images/img18.jpg
00:24:36.004	0.128	360	6464	GET	200	image/jpeg	http://localhost:8080/optimism/images/img19.jpg
00:24:36.004	0.132	360	5632	GET	200	image/jpeg	http://localhost:8080/optimism/images/img20.jpg
00:24:36.087	0.122	359	591	GET	200	image/gif	http://localhost:8080/optimism/images/img01.gif
00:24:36.087	0.124	359	388	GET	200	image/gif	http://localhost:8080/optimism/images/img02.gif
00:24:36.088	0.124	359	441	GET	200	image/gif	http://localhost:8080/optimism/images/img03.gif
00:24:36.089	0.125	359	2044	GET	200	image/gif	http://localhost:8080/optimism/images/img04.gif
00:24:36.090	0.123	359	490	GET	200	image/gif	http://localhost:8080/optimism/images/img06.gif
00:24:36.091	0.124	359	2053	GET	200	image/gif	http://localhost:8080/optimism/images/img05.gif
00:24:36.092	0.125	359	486	GET	200	image/gif	http://localhost:8080/optimism/images/img07.gif
00:24:36.092	0.126	359	514	GET	200	image/gif	http://localhost:8080/optimism/images/img08.gif
00:24:36.096	0.147	359	439	GET	200	image/gif	http://localhost:8080/optimism/images/img09.gif
00:24:36.096	0.154	359	381	GET	200	image/gif	http://localhost:8080/optimism/images/img10.gif

Request Header	Value	Response Header	Value
(Request-Line)	GET /optimism/index.html HTTP/1.1	(Status-Line)	HTTP/1.0 200 OK
Host	localhost:8080	Date	Wed, 16 Oct 2013 17:53:27 GMT
User-Agent	Mozilla/5.0 (Windows NT 6.1; WOW64; rv:24.0) Gecko/201001...	Server	Apache/2.2.25 (Win32)
Accept	text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=...	Last-Modified	Wed, 16 Oct 2013 16:57:57 GMT
Accept-Language	en-US,en;q=0.5	Etag	"20000002bca5-1ccf-4e8de9a64e1ad"
Accept-Encoding	gzip, deflate	Accept-Ranges	bytes
Cache-Control	no-cache	Content-Length	7375
Connection	keep-alive	Cache-Control	max-age=300
		Expires	Wed, 16 Oct 2013 17:58:27 GMT
		Content-Type	text/html
		X-Cache	MISS from localhost
		X-Cache-Lookup	HIT from localhost:3128

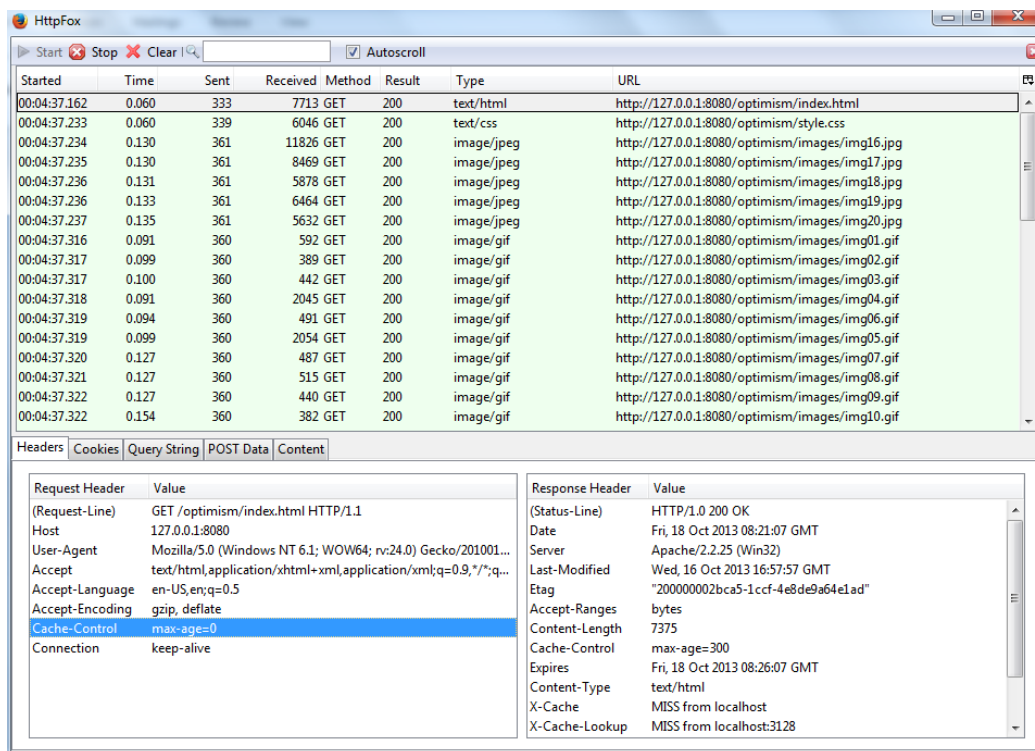
Slika 8 Ponovni pristup stranici index.html

Dakle, ponovno se šalju 200 OK odgovori s kopijama resursa, ali polje X-Cache-Lookup sada ima vrijednost HIT from localhost:3128.

U prvom slučaju spremište kontaktira poslužitelj kako bi dobavilo potrebne kopije resursa, u drugom slučaju ponovno kontaktira poslužitelja kako bi dobavilo potrebne kopije resursa jer je neposredno prije toga ispražnjeno kao što je zadano u zadatku, i opet dobiva kopije resursa od originalnog poslužitelja. Budući da se pomoću dodatka *Modify Headers* dodaju polja u zaglavlje onih paketa koje šalje preglednik, a ne poslužitelj, onda je potrebno analizirati učinak `no-cache` direktive kada se šalje s klijentske strane. Kada klijent uključi tu direktivu, onda bi sva priručna spremišta na putu prema originalnom poslužitelju trebala uvijek zatražiti novu i valjanu kopiju, osobito ako je valjanost traženog resursa istekla, a poslužitelji ne bi smjeli koristiti svoja kopije u priručnim spremištima kada vraćaju odgovor (zbog toga i u drugom slučaju kopije dolaze od originalnog poslužitelja, a ne od posrednika). Također, vrijednost `no-cache` u slučaju kada se koristi od strane servera, a ne klijenta, ne zabranjuje pregledniku i posredniku privremeno pohranjivanje, već upozorava da bi se kod ponovnog prikazivanja tog resursa prvo trebalo kontaktirati izvornog poslužitelja da se provjeri je li resurs još valjan.

## 6. Zadatak

Nakon dodavanja zaglavlja `Cache-Control: max-age=0` i pristupanja stranici index.html, dobivamo sljedeći prikaz u HTTPFox-u:

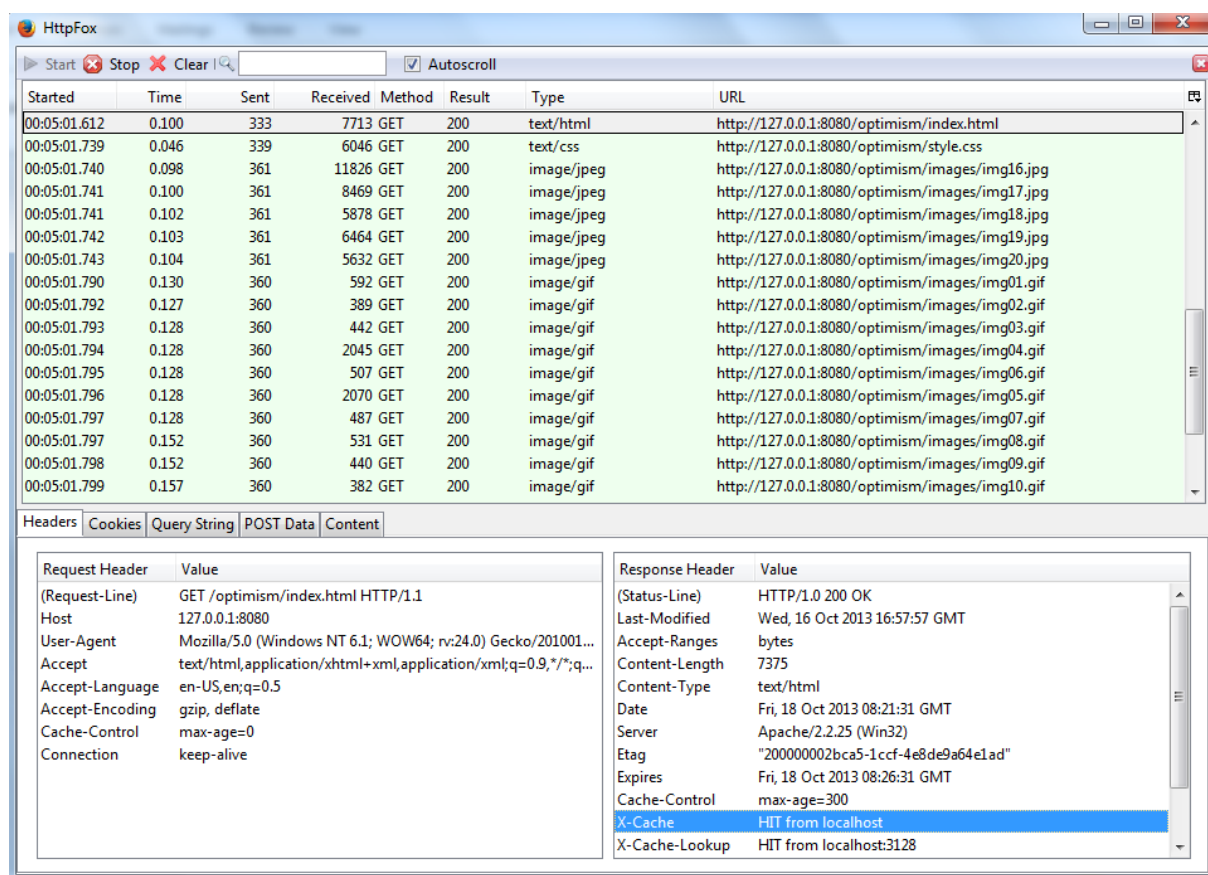


Slika 9 Označeno polje Cache-Control s vrijednošću max-age=0

U odgovoru kojeg primamo od poslužitelja možemo vidjeti da je posrednik squid u poljima X-Cache i X-Cache-Lookup postavio vrijednosti MISS. U zahtjevima preglednika nema polja If-Modified-Since i If-None-Match što znači da se resursi dohvaćaju po prvi put i kopije se isporučuju u tijelu odgovora 200 OK kojeg šalje originalni poslužitelj.

Ako pak obrišemo lokalno spremište Firefoxa kao što je zadano u zadatku te ponovno pristupimo istoj stranici, dobivamo sljedeći rezultat:





Slika 10 Ponovni pristup stranici index.html, polja X-Cache i X-Cache-Lookup imaju vrijednost HIT

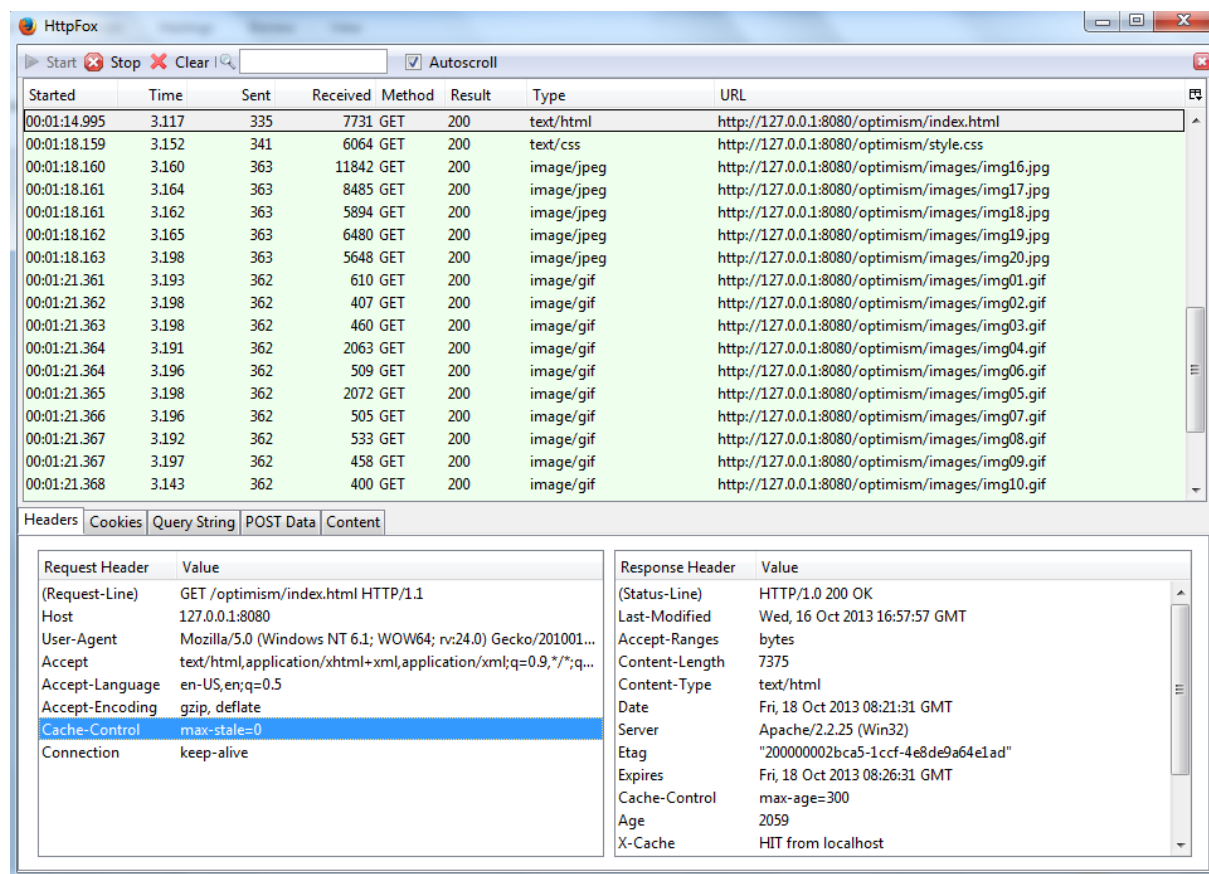
Budući da Firefox nema potrebne kopije resursa u svom lokalnom spremištu, te kopije se moraju ponovno isporučiti bilo od originalnog poslužitelja ili posredničkog poslužitelja. Budući da polja X-Cache i X-Cache-Lookup sada imaju vrijednost HIT, zaključujemo da je posrednički poslužitelj iz svog priručnog spremišta isporučio kopije resursa.

U prvom slučaju (lokalno spremište Firefoxa i priručno spremište Squida je prazno) kontaktira se krajnji poslužitelj i isporučuje se sadržaj od poslužitelja Apache preko Squida do Firefoxa. U drugom slučaju preglednik opet šalje zahtjev za dohvaćanjem kopija resursa jer ih trenutno nema, ali se sada te kopije isporučuju od strane posredničkog poslužitelja, a poznavajući svojstvo direktive max-age=0, posrednik je trebao prvo revalidirati svoje kopije prije nego ih je isporučio (dakle i tu se kontaktira originalni poslužitelj, ali samo za revalidaciju). Nažalost, komunikaciju između posrednika i originalnog poslužitelja ne možemo pregledati pomoću HTTPFoxa stoga trebamo dobro poznavati svojstva direktive max-age=0.

Direktiva max-age=0 ima „blaži“ učinak u odnosu na direktivu no-cache, odnosno kada tu direktivu preglednik uključi u zaglavlje zahtjeva kojeg šalje prema originalnom poslužitelju, onda sva priručna spremišta na putu do tog originalnog poslužitelja mogu isporučiti kopiju resursa kada se ponovno zatraži, ali se mora prvo validirati njezina valjanost. no-cache s druge strane nalaže svima na putu da moraju dohvatiti nove kopije, a ne isporučivati one koje se već nalaze u njihovim priručnim spremištima (zato je u 5. zadatku odgovor primljen na strani preglednika nakon ponovnog pristupanja stranici imao X-Cache: MISS from localhost).

## 7. Zadatak

Nakon brisanja lokalnog spremišta i gašenja poslužitelja Apache, dobivamo sljedeći prikaz u HTTPFox-u tijekom pristupanja stranici index.html:

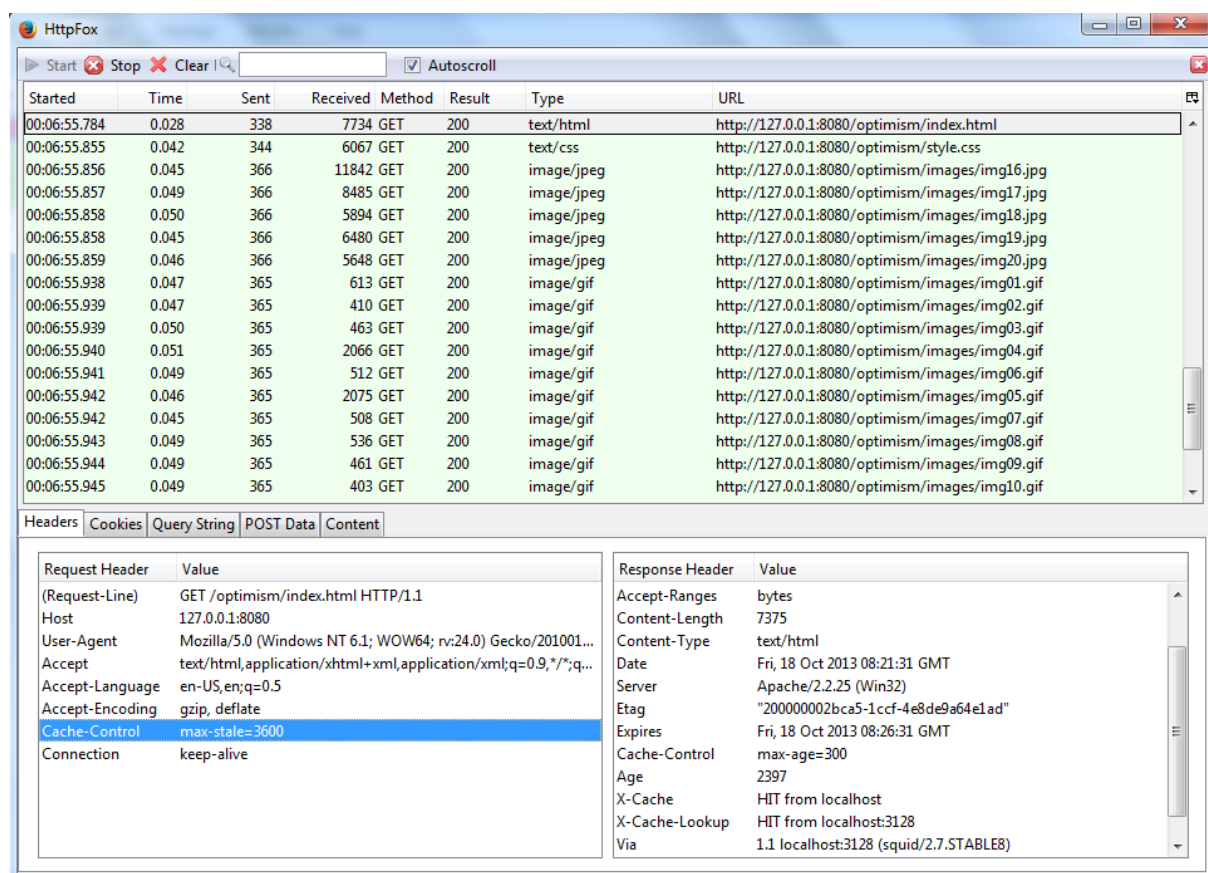


Slika 11 HTTP paket s poljem Cache-Control: max-stale=0

Vidimo da preglednik šalje zahtjev za dohvaćanjem potrebnih resursa te se vraća odgovor 200 OK, ali on dolazi od posredničkog poslužitelja Squid jer vrijednost polja X-Cache i X-Cache-Lookup je jednako HIT from localhost.

Ako obrišemo lokalno spremište, postavimo vrijednost `max-stale=3600` i ponovno pristupimo stranici, dobivamo sljedeći prikaz:





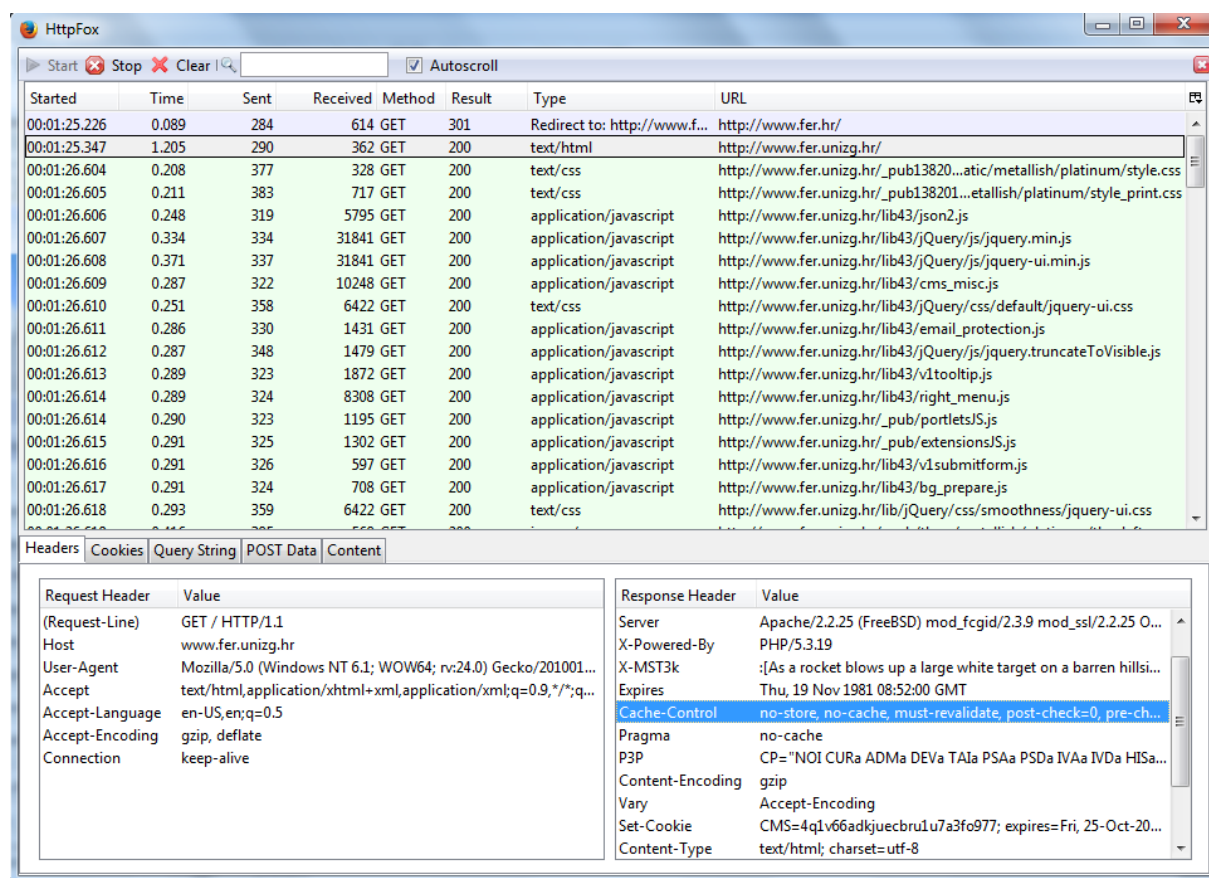
Slika 12 HTTP paket s poljem Cache-Control: max-stale=3600

Rezultat je praktički isti. Max-stale opcija ne iziskuje od posredničkog poslužitelja revalidaciju određenog resursa kod izvornog poslužitelja, stoga u ovom slučaju squid neće revalidirati kopije koje ima kod sebe, već će ih direktno poslati pregledniku kad to bude trebalo. Pomoću ove direktive preglednik može pristupiti resursu kojemu je vrijeme valjanosti već isteklo, ali ne više od onog vremena specificiranog pomoću direktive max-stale.

U ovom zadatku je učitavanje stranica bilo nešto sporije jer se kopije dohvaćaju preko posrednika.

## 8. Zadatak

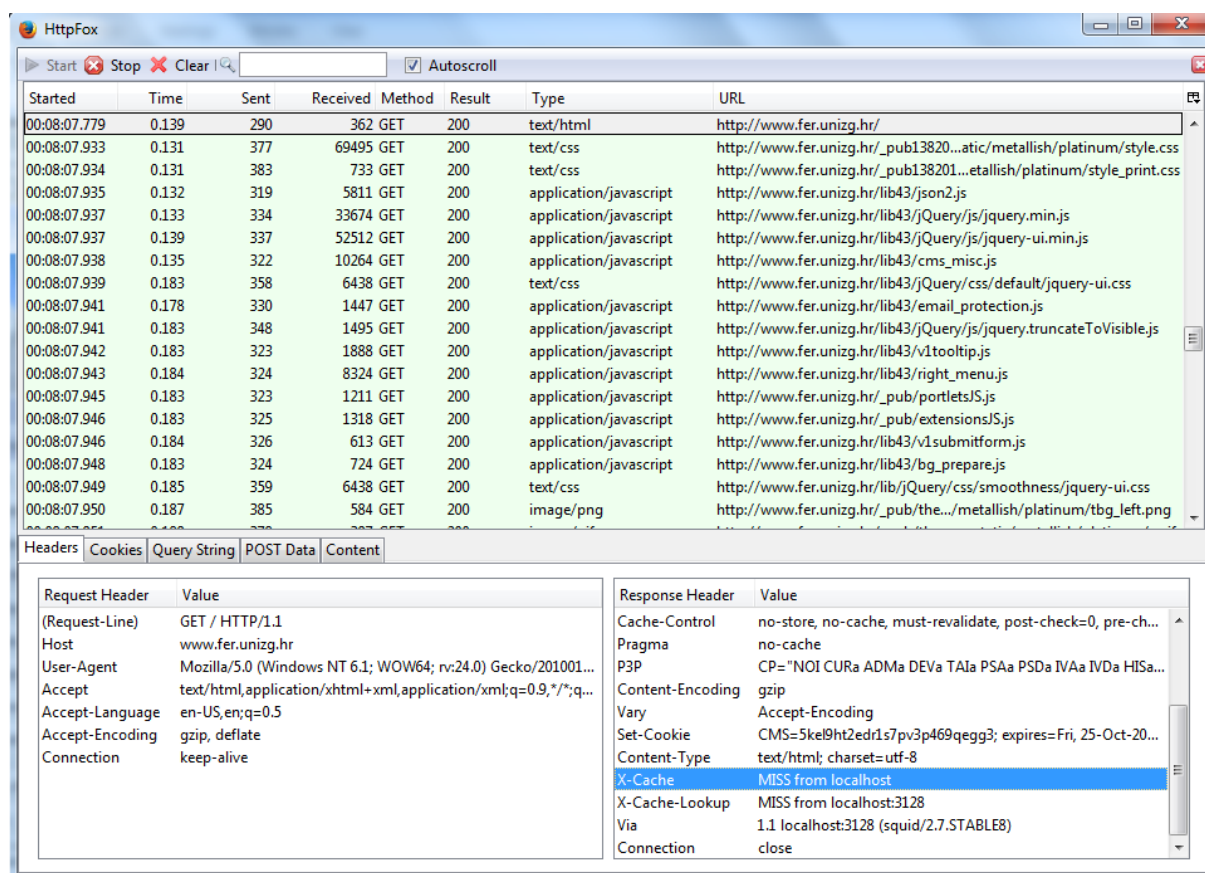
Nakon pristupanja stranici [www.fer.hr](http://www.fer.hr) dobivam sljedeći prikaz:



Slika 13 Paketi primljeni i poslani nakon pristupanja stranici fer.hr

Prvo se šalje GET zahtjev za dohvaćanjem početne stranice s adrese [www.fer.hr](http://www.fer.hr), ali se preusmjerava na stranicu [www.fer.unizg.hr](http://www.fer.unizg.hr) te preglednik šalje GET zahtjev sada na tu stranicu. Redom se dobivaju 200 OK odgovori s kopijama potrebnih resursa, od html-a, css-a, javascript-a, slika... Na slici je naznačena vrijednost polja `Cache-Control` u jednom od odgovora (`no-store` direktiva nalaže pregledniku da ne pohranjuje namjerno kopiju resursa na tvrdi disk te da pokuša ukloniti kopiju iz priručnog spremišta čim prije, `no-cache` je objašnjeno u zadatku 5, `must-revalidate` direktiva striktno nalaže pregledniku i posredniku da moraju provjeriti valjanost resursa prije ponovnog prikazivanja, `post-check` direktiva definira interval nakon kojeg preglednik ponovno dohvaća kopiju resursa iz priručnog spremišta i u pozadini traži ažuriranu verziju, `pre-check` definira interval nakon kojeg preglednik traži ažuriranu kopiju resursa i onu staru tretira kao zastarjelu; sve te direktive zapravo zabranjuju privremeno pohranjivanje određenog resursa ). Većina odgovora ima samo vrijednost `max-age=...` Polja `X-Cache` i `X-Cache-Lookup` imaju vrijednost `MISS from localhost` za svaki odgovor jer se sve dohvaća od izvornog poslužitelja.

Ako pak obrišemo lokalno spremište Firefoxa i ponovno pristupimo toj adresi, dobivamo sljedeći prikaz:



Slika 14 Ponovni pristup stranici fer.hr

Zbog strogih pravila vezanih uz privremeno pohranjivanje html koda od početne FER-ove stranice, vidimo da i sada polja X-Cache i X-Cache-Lookup imaju vrijednost MISS from localhost što znači da se kopija tog resursa opet dohvaćala s izvornog poslužitelja, a ne od strane posrednika.

Za neke druge resurse (npr. slike) dobiva se vrijednost HIT from localhost u poljima X-Cache i X-Cache-Lookup jer za njih nisu postavljena stroga pravila.

Pravila za upravljanje priručnom memorijom bi se možda mogla poboljšati neznatno. Moguće je postaviti stroža pravila za privremeno pohranjivanje u polje Cache-Control za neke druge resurse, npr. javascript i neke slike za koje se utvrdi da bi se mogle češće mijenjati na izvornom poslužitelju. Iako html kod definira kostur čitave stranice i sadrži reference na ostale resurse pa je logično postaviti strožu kontrolu na njega, možda bi bilo dobro takvu kontrolu pohranjivanja postaviti i na neke resurse kojima bi preglednik mogao pristupiti bez prvotnog dohvaćanja html koda.

## 9. Zadatak

Logovi poslužitelja Squid se nalaze u direktoriju `var/logs`. Postoje četiri datoteke u tom direktoriju: `access`, `cache`, `store` i `squid.pid`. Prve tri datoteke su tekstualnog tipa. Prva datoteka navodi zahtjeve koji su prolazili kroz posrednički poslužitelj te jesu li dohvaćeni iz priručnog spremišta ili od originalnog servera. Primjerice, tijekom prvog pristupanja stranici `fer.hr` u 8. zadatku dodane su sljedeće linije:

```
1382087567.887      86 127.0.0.1 TCP_MISS/301 778 GET
http://www.fer.hr/ - DIRECT/161.53.72.119 text/html

1382087569.196     1249 127.0.0.1 TCP_MISS/200 59744 GET
http://www.fer.unizg.hr/ - DIRECT/161.53.72.120 text/html
```

TCP\_MISS ukazuje na to da se kopija resursa ne može dohvatiti iz priručnog spremišta te da se mora preuzeti od strane originalnog poslužitelja. Za resurs čija se kopija može naći u priručnom spremištu, ispisuje se TCP\_MEM\_HIT ili samo TCP\_HIT.

Datoteka `cache` prikazuje neke općenite informacije i akcije koje su poduzete u određenom trenutku. Primjer:

```
2013/10/15 16:13:38| Starting Squid Cache version 2.7.STABLE8 for i686-pc-
winnt...

2013/10/15 16:13:38| Running on Windows Vista

2013/10/15 16:13:38| Process ID 2552

2013/10/15 16:13:38| With 2048 file descriptors available

2013/10/15 16:13:38| With 512 CRT stdio descriptors available

2013/10/15 16:13:38| Windows sockets initialized

2013/10/15 16:13:38| Using select for the IO loop

2013/10/15 16:13:38| Performing DNS Tests...

2013/10/15 16:13:38| Successful DNS name lookup tests...

2013/10/15 16:13:38| DNS Socket created at 0.0.0.0, port 50341, FD 4

2013/10/15 16:13:38| Adding DHCP nameserver 192.168.1.1 from Registry

...
```

Datoteka `store` prikazuje oznake i kodove pridružene pojedinim resursima koji su privremeno pohranjeni u priručnom spremištu posredničkog poslužitelja. Primjer (tri gif slike iz direktorija `optimism`):

```
1382086557.390 RELEASE -1 FFFFFFFF 171601DEEB1B819F051B8FAFCD26B65F 504
1382086557      -1      -1 text/html 1157/1157 GET
http://127.0.0.1:8080/optimism/images/img04.gif
```

```

1382086557.390 RELEASE -1 FFFFFFFF 5E0A92E15B5F76FECFBB76D416233F77 504
1382086557 -1 -1 text/html 1157/1157 GET
http://127.0.0.1:8080/optimism/images/img11.gif

1382086557.390 RELEASE -1 FFFFFFFF 4D7C08ABDA3FC11F7D4EAAA8F620A6CA 504
1382086557 -1 -1 text/html 1157/1157 GET
http://127.0.0.1:8080/optimism/images/img01.gif

```

Poslužitelj Apache ima log datoteku naziva `install` koja se nalazi u direktoriju `logs` i koja sadrži neke općenite informacije vezane uz pokretanje poslužitelja, modifikaciju konfiguracijske datoteke itd. Primjer (domena servera, administrator, port, neke akcije...):

```

Installing Apache HTTP Server 2.x with

DomainName =      somenet.com

ServerName =      www.somenet.com

ServerAdmin =     x.x@fer.hr

ServerPort =      8080

ServerSslPort =   443

ServerRoot =      C:/Program Files (x86)/Apache Software
Foundation/Apache2.2

Rewrote C:/Program Files (x86)/Apache Software
Foundation/Apache2.2/conf/original/extra/httpd-autoindex.conf.in

to C:/Program Files (x86)/Apache Software
Foundation/Apache2.2/conf/original/extra/httpd-autoindex.conf

Successfully removed C:\Program Files (x86)\Apache Software
Foundation\Apache2.2\conf\original\extra\httpd-autoindex.conf.in

```

Pored gornjeg loga, Apache sadrži i tzv. `Access Log` u kojem se bilježe svi zahtjevi procesirani od strane poslužitelja (vrijeme zaprimanja zahtjeva, IP adresa klijenta koji je poslao zahtjev, veličina objekta koji se vraća klijentu, status odgovora itd.), zatim `Error Log` u kojem su zabilježene pogreške generirane tijekom rada poslužitelja (tip pogreške, vrijeme i datum poruke, sadržaj pogreške itd.), postoje još neke pomoćne log datoteke (npr. `script log`).