

Proxv adaptativo

# Proxy Adaptativo para Protocolos Web Avanzados

Pablo Maximiliano Lulic

Universidad Nacional de Luján



2014



# Resumen

Proxy adaptativo

Pablo Maximiliano Lulic

Introduccio

Crecimient

Conceptos preliminare Protocolos TCP RTT

HTTPS Problemátic le la web

Proxy

Experimentos

Resultados

Conclusi

- 1 Introducción
- 2 Crecimiento de la Web
- 3 Conceptos preliminares
  - Protocolos
  - TCP
  - RTT
  - HTTP
  - HTTPS
- Problemática de la web
  - SPDY
- Proxy
- 6 Proxy Adaptativo
- Experimentos
- Resultados
- Onclusiones



## Introducción

Proxy adaptativo

Pablo Maximiliano Lulic

### Introducción

. . . . . . .

Conceptos preliminare

Protocolos TCP RTT HTTP

Problemátic de la web SPDY

Proxy

Proxy Adaptativ

Experimento

Canalusianas

Qué es la web?

- Desde el año 1995
- Repositorio distribuído
- Heterogénea
- Hipervínculos





### Cómo es la web?

Proxy adaptativo

Pablo Maximiliano

Introducción

Crecimiento

Conceptos preliminare:

Protoco TCP

RTT

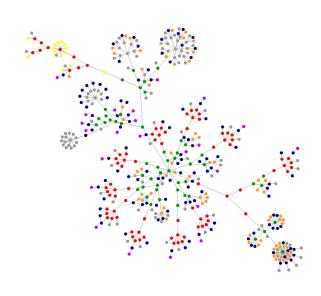
Problemát de la web

Proxy

Proxy Adaptativ

Experimentos

C . . . . l . . . . . . . . . .





### Cómo funciona la web?

Proxy adaptativo

Pablo Maximiliano

### Introducción

Crecimient

Conceptos preliminare

Protocolos

TCP

KII

Problemát

SPDY

Proxy

Proxy Adaptativ

Experimento

C = = = |......





### Cómo funciona la web?

Proxy adaptativo

Pablo Maximiliano Lulic

### Introducción

Crecimient

Conceptos preliminares Protocolos TCP

RTT HTTPS

Problemátic de la web SPDY

Proxy

Proxy Adaptative

Experimentos

Resultados

Canalusianas



### Intermediarios





# **Aportes**

Proxy adaptativo

Pablo Maximiliano Lulic

Introducción

Crecimiento

Conceptos preliminare Protocolos TCP RTT HTTP HTTPS

de la web

Provv

Proxy Adaptativo

Experimentos

Conclusione

### Aportes del trabajo realizado

- Estudio de la problemática de la web, respecto a su complejidad
- Estudio de los protocolos actuales y los que evolucionan a HTTP
- Propuesta y desarrollo de un proxy adaptativo que intenta mejorar la performance de la recuperación web
- Evaluación del proxy en un escenario real en donde se observan mejoras en la carga de los sitios
- Cliente SPDY



### La web de hace años

Proxy adaptativo

Pablo Maximiliano Lulic

Introducciói

Crecimiento

Conceptos preliminare

Protocolos TCP RTT HTTP HTTPS

Problemático de la web

Proxy

Proxy Adaptativ

Experimento

Conclusione

### Como eran los sitios hace 14 años?

- Sitios sencillos
- Poco contenido
- Tamaño aproximado de 70kb [WebsiteOptimization]<sup>1</sup>

 $<sup>^{1} \</sup>mathtt{http://www.websiteoptimization.com/speed/tweak/average-web-page}$ 



# Sitio de Amazon, año 2000 [WebArchive]<sup>2</sup>

Proxy adaptativo

Pablo Maximiliano Lulio

Introducciór

Crecimiento

Conceptos preliminare

Protocolos TCP

HTTP

Problemátic de la web

Proxy

Adaptativo

Experimentos

Conclusione





# La web hoy

Proxv adaptativo

Crecimiento de la Web

### Como son los sitios hoy?

- Sitios complejos
- Mucho contenido
- Tamaño aproximado de 1800kb (Junio 2014) [HTTPArchive]<sup>3</sup>

 $<sup>^3</sup>$ http://httparchive.org



## Sitio de Amazon, año 2014

Proxy adaptativo

Pablo Maximiliano Lulic

Introducciór

Crecimiento de la Web

Conceptos

Protocolos

RTT HTTP

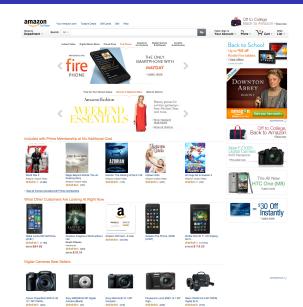
Problemátio de la web

Proxy

Proxy Adaptativ

Experimentos

Conclusiones





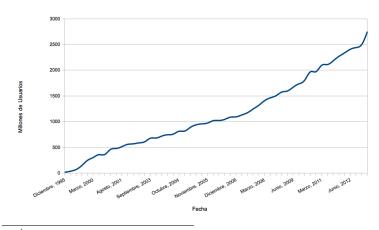
### Crecimiento de los usuarios

Proxv adaptativo

Crecimiento de la Web

## Aumento de los usuarios globales

(Información extraída de [IWS]<sup>4</sup>)



 $<sup>^{4} \</sup>mathtt{http://www.internetworldstats.com/emarketing.htm}$ 



## **Protocolos**

Proxy adaptativo

Pablo Maximiliano Lulic

Crecimiento

Conceptos preliminares Protocolos TCP RTT HTTP

Problemátic de la web SPDY

Proxy

Adaptativo

Experimentos

Conclusiones

### Protocolo

Un protocolo de comunicaciones define el formato y el orden de los mensajes intercambiados entre dos o más entidades que se comunican, así como las acciones que se toman en la transmisión y/o recepción de un mensaje u otro evento.

- Sintaxis
- Semántica
- Sincronización

Las tareas a realizar se dividen en capas



# TCP/IP

Proxy adaptativo

Pablo Maximiliano Lulic

Introducción

Crecimiento

Conceptos preliminare

Protocolos TCP RTT HTTP

HTTPS Problemátic de la web

Proxy

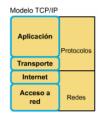
Proxy Adaptativ

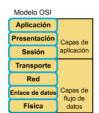
Experimento

Canalusianas

### TCP/IP

- Es el estándar de Internet
- OSI toma su experiencia
- Diseño en capas





# TCP

Proxy adaptativo

Pablo Maximiliano Lulic

Introduccio

de la Web

preliminares
Protocolos
TCP
RTT
HTTP

Problemátic de la web SPDY

Proxy

Evperimento

Experimento

Conclusione

### TCP (Transport Control Protocol)

Permite crear conexiones entre hosts para el envío de flujos de datos.

- Divide los datos en paquetes
- Control de recepción
- Tiempos de espera
- Multiplexación

## TCP

Proxy adaptativo

Pablo Maximiliano Lulic

Introducción

.....

Conceptos

preliminares
Protocolos

TCP

HTTP HTTPS

de la web

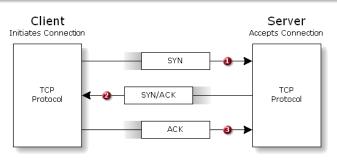
Proxy

Proxy Adaptative

Experimentos

Conclusioner

### Establecimiento de la conexión



Three Way Handshake

# **RTT**

Proxy adaptativo

Pablo Maximiliano Lulic

Introducción

Cracimianta

Conceptos preliminares Protocolos

TCP

RTT HTTP

Problemátic de la web SPDY

Proxy

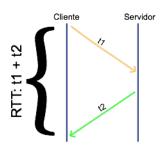
Proxy Adaptative

Experimentos

Conclusiones

### RTT (Round Time Trip)

Es el tiempo que demora un paquete en ir del emisor al receptor y volver al emisor nuevamente.





# **HTTP**

Proxy adaptativo

Pablo Maximiliano Lulic

Introduccio

de la Web Conceptos preliminares

Protocolos TCP RTT HTTP HTTPS

Problemática de la web SPDY

Proxy

Adaptativo

Experimento

Conclusions

### HTTP (HyperText Transfer Protocol)

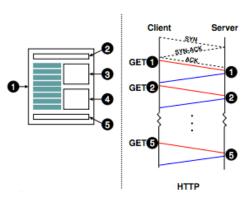
Es un protocolo de capa de aplicación que sirve para distribuir información.

- Primera versión en el año 1990 (0.9)
- Versión actual desde el año 1999 (1.1)
- 1 conexión por elemento
- Siempre la petición la inicia el cliente



# **HTTP**

Proxv adaptativo



Carga de un sitio completo [SPDYMobileWeb]<sup>5</sup>

 $<sup>^{5} {\</sup>sf Jeffrey\ Erman,\ Vijay\ Gopalakrishnan,\ Rittwik\ Jana,\ y\ K.K.\ Ramakrishnan.\ Towards\ a\ spdyier\ mobile\ web?}$ 



## **HTTPS**

Proxv adaptativo

### HTTPS (HyperText Transfer Protocol Secure)

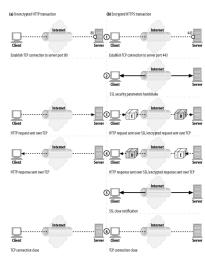
Es la versión "segura" de HTTP, añade a este protocolo una capa de cifrado utilizando SSL (Secure Sockets Layer).

- Datos protegidos
- Los datos viajan encriptados
- Necesita establecer una conexión segura



## **HTTPS**

Proxv adaptativo



## Conexión HTTPS [HTTPGuide]<sup>6</sup>

 $<sup>^6</sup>$  Gourley, Totty, Sayer, Aggarwal, y Reddy. HTTP: The Definitive Guide. OReilly Media, 2002.

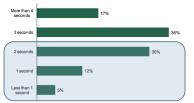


Proxv adaptativo

Problemática de la web

### Porqué es importante mejorar los tiempos de carga de las páginas?





#### Typically, how long are you willing to wait for a single web page to load before leaving the website?



# [Akamai]<sup>7</sup>

 $<sup>^{7}</sup>$  Forrester Consulting y Inc. Akamai Technologies. ecommerce web site performance today. Agosto 2009.



Proxy adaptativo

Pablo Maximiliano Lulic

Introducción

Conceptos preliminare Protocolos

TCP RTT HTTP HTTPS

Problemática de la web

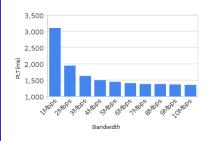
Proxy

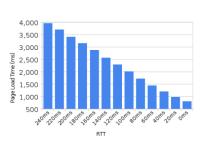
Proxy Adaptative

Experimentos

Conclusion

### Es importante el Ancho de Banda?





[Belsche]<sup>8</sup>

<sup>&</sup>lt;sup>8</sup>Mike Belshe. More bandwidth doesn?t matter (much). Agosto 2010.



Proxy adaptativo

Pablo Maximiliano Lulic

Introduccior

Cracimianta

Conceptos preliminare Protocolos TCP

TCP RTT HTTP HTTPS

Problemática de la web

Proxy

Proxy Adaptativ

Experimentos

Conclusiones

## Percepción de los usuarios [Grigorik]<sup>9</sup>

cepción del Usuario
antánea
raso pequeño perceptible
máquina se encuentra procesando
rracción en lo que se está realizando
tarea es abandonada

 $<sup>^9 \, {\</sup>rm Ilya}$  Grigorik. High Performance Browser Networking. OReilly Media, 2013.



Proxy adaptativo

Pablo Maximiliano Lulic

Introducció

Conceptos preliminare Protocolos

TCP RTT HTTP HTTPS

Problemática de la web

Proxy

Adaptativo

Experimentos

Conclusi

Percepción de los usuarios [Grigorik]<sup>9</sup>

Retardo	Percepción del Usuario
0 - 100ms	Instantánea
100 - 300ms	Retraso pequeño perceptible
300 - 1000ms	La máquina se encuentra procesando
$1000+\ \mathrm{ms}$	Distracción en lo que se está realizando
$10000+\ \text{ms}$	La tarea es abandonada

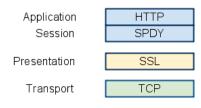
# Para que la experiencia sea positiva 250ms de retardo promedio

 $<sup>^9</sup>$ llya Grigorik. High Performance Browser Networking. OReilly Media, 2013.

Proxv adaptativo

### SPDY (SPeeDY)

Es un protocolo experimental desarrollado por Google a mediados del 2009, cuya meta principal es reducir los tiempos de carga de los sitios enfocándose en las limitaciones de HTTP.



Protocolo SPDY

Proxy adaptativo

Pablo Maximiliano Lulio

Introduccio

Conceptos preliminare Protocolos TCP RTT

HTTPS HTTPS Problemátic de la web

Provv

Proxy Adaptative

Experimento

Conclusione

### Problemas de HTTP

- Por cada conexión nueva se necesita establecer una conexión TCP (RTTs adicionales)
- Slow Start de TCP
- Límite de conexiones a un servidor
- Domain sharding

Proxy adaptativo

### Qué ofrece SPDY?

- 1 conexión TCP
- Multiplexación
- Priorización
- Compresión de headers
- Server Push
- Server Hint



Proxy adaptativo

Pablo Maximiliano Lulic

Introduccion

Crecimiento

Protocolos
TCP

RTT HTTP HTTPS

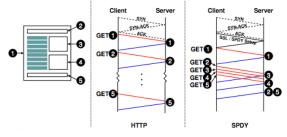
Problemátic de la web

Proxy

Proxy Adaptative

Experimentos

C----



Carga de un sitio completo [SPDYMobileWeb]



## Definición

Proxy adaptativo

Pablo Maximiliano Lulio

Introducciói

Conceptos preliminare Protocolos

TCP RTT HTTP

Problemátio de la web

Proxy

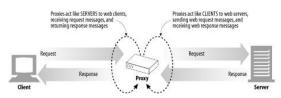
Proxy Adaptative

Experimentos

Conclusiones

### Proxy

Son intermediarios entre el cliente y el servidor. Actúan enviando los mensajes del cliente al servidor y viceversa.



Funcionamiento [HTTPGuide]

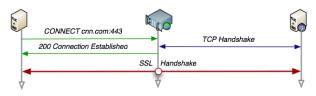


# Proxies y HTTPS

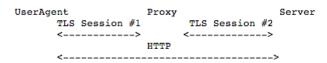
### Proxv adaptativo

Proxy

### Como se conecta HTTPS?



### Túnel HTTPS [Igvita]<sup>10</sup>



Proxy de confianza (MITM) [TrustedProxy]<sup>11</sup>

 $<sup>^{10} {\</sup>rm https://www.igvita.com/2011/12/01/\ web-vpn-secure-proxies-with-spdy-chrome}$ 

 $<sup>^{11} {\</sup>rm http://www.ietf.org/id/\ draft-loreto-httpbis-trusted-proxy20-01.txt}$ 



## Utilización

Proxy adaptativo

Pablo Maximiliano Lulic

Introduccio

Conceptos preliminare Protocolos TCP

RTT HTTP HTTPS

de la web

Proxy

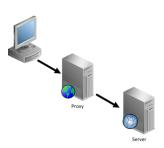
Experimento

Experimentos

Conclusiones

### Usos de un proxy

- NAT/enmascaramiento
- Filtro
- Cortafuegos
- Web caché \*
- Reverso
- Content-router
- Transcoder
- Anonimizador





# Proxy Adaptativo

Proxy adaptativo

Pablo Maximiliano Lulic

Introduccioi

Conceptos preliminare: Protocolos TCP RTT

HTTPS <sup>P</sup>roblemática de la web

Proxy

Proxy Adaptativo

Experimentos

Conclusiones

Se propone, desarrolla y evalúa un proxy con la siguiente funcionalidad:

- Selecciona de los protocolos disponibles (HTTP, HTTPS, SPDY), el óptimo para recuperar el recurso
- Caché
- Análisis de los sitios
- Man In The Middle
- Cliente SPDY







# Proxy Adaptativo

Proxv adaptativo

Proxy Adaptativo

### Porque es Adaptativo?

La selección del protocolo la realiza en base a información almacenada de recuperaciones previas. Se van analizando los obietos y los sitios a medida que el proxy se encuentra en funcionamiento.





# Árbol de decisión

Proxv adaptativo

Proxy Adaptativo

### Análisis de datos

- Protocolos habilitados.
- RTT al servidor
- Tamaño del objeto
- Cantidad de peticiones que va a generar





# Árbol de decisión

Proxy adaptativo

Pablo Maximiliano Lulic

Introducción

Crecimiento de la Web

Conceptos preliminare: Protocolos TCP

RTT HTTP

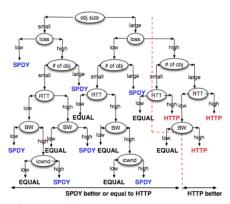
Problemátic de la web

Proxy

Proxy Adaptativo

Experimentos

Conclusiones



Árbol de decisión [HowSpeedy]<sup>12</sup>

 $<sup>^{12}\</sup>mbox{Wang, Balasubramanian, Krishnamurthy y Wetherall. How speedy is spdy? Abril 2014.}$ 



# Datos y metodología

Proxy adaptativo

Pablo Maximiliano Lulic

Introduccio

Crecimiento de la Web

Conceptos
preliminar
Protocolos
TCP
RTT
HTTP
HTTPS

Problemátic de la web SPDY

Proxy

Proxy Adaptativ

 ${\sf Experimentos}$ 

Conclusion

### **Datos**

- Sitios del ranking de Alexa<sup>13</sup>
- HTTP y HTTPS
- Tiempo onLoad

### Metodología

- Comparar proxy adaptativo vs proxy convencional
- Caché deshabilitado
- Fases: HTTP y HTTPS
- 3 repeticiones de cada fase con cada proxy

<sup>13</sup> http://www.alexa.com/topsites



# Fases del experimento

Proxy adaptativo

Pablo Maximiliano

Introducción

Crecimiento

Conceptos preliminares

Protocolos TCP

HTTP

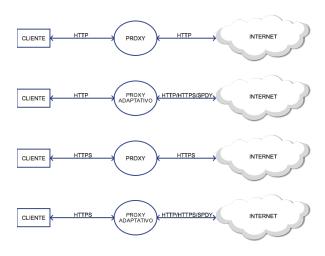
Problemátic de la web

Proxy

Proxy Adaptative

 ${\sf Experimentos}$ 

\_\_\_\_\_\_





Proxv adaptativo

Resultados

Mejoras en los tiempos de carga de los sitios (sólo los que soportan todos los protocolos):

### **HTTP**

5 % de mejora promedio con picos de hasta 17 %

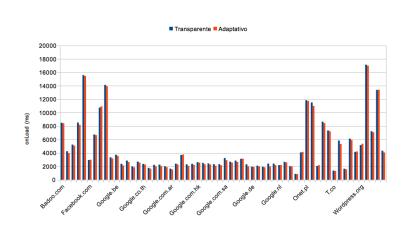
### **HTTPS**

7 % de mejora promedio con picos de hasta 24 %



Proxv adaptativo

Resultados



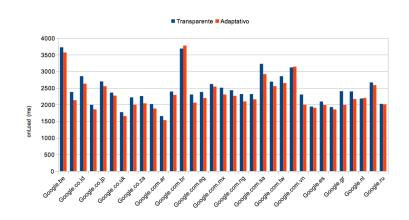
Sitios con todos los protocolos soportados, HTTP



Proxv adaptativo

Resultados



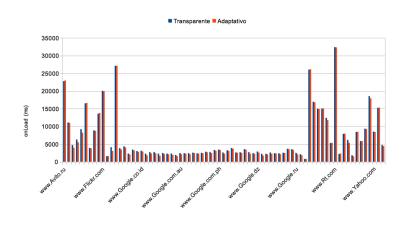


Sitios de Google, HTTP



Proxv adaptativo

Resultados



Sitios con todos los protocolos soportados, HTTPS



Proxy adaptativo

Pablo Maximiliano Lulic

Introducció

Cracimiant

Conceptos preliminare Protocolos TCP

RTT HTTP HTTPS

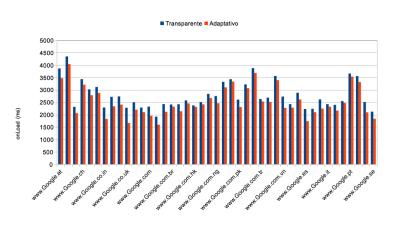
Problemátic de la web SPDY

Proxy

Adaptativo

Experimento Resultados

Conclusiones



Sitios de Google, HTTPS

Proxy adaptativo

Pablo Maximiliano Lulic

Introducció

Concepto preliminal Protocolos TCP RTT

RTT HTTP HTTPS

de la web SPDY

Proxy

Adaptativo

Resultados

Resultados

### Rendimiento del proxy adaptivo

**HTTP** 

Negativo	Bajo	Medio	Alto
16,5 %	31,5 %	45 %	7%

**HTTPS** 

111 11 5					
Negativo	Bajo	Medio	Alto		
16 %	36 %	32 %	16 %		

Rendimiento negativo: en el 90% de los casos las diferencias son mínimas (variaciones de RTT)



## Conclusiones

Proxy adaptativo

Pablo Maximiliano Lulic

Introducciór Crecimiento

Conceptos preliminare Protocolos TCP RTT HTTP HTTPS

de la web

Proxy

Experimentos

Conclusiones

- Aumento de las páginas, de 70kb a 1800kb en 14 años (2500 %). El crecimiento es exponencial
- El rendimiento de SPDY se encuentra directamente relacionado con el contexto en el que está operando
- Con el proxy adaptativo propuesto, se obtienen mejoras de 5 % y 7 % en el tiempo de carga de los sitios, para HTTP y HTTPS respectivamente
- Del ranking de Alexa (500 sitios) solo el 20 % soportan SPDY
- SPDY sentó las bases del próximo HTTP 2.0, es importante estudiar su rendimiento
- Ser MITM brinda la posibilidad de utilizar caché en HTTPS



## Conclusiones

Proxy adaptativo

Maximilian Lulic

Introduccio

C----

Conceptos preliminares Protocolos TCP RTT HTTP HTTPS

Problemátic de la web

Proxy

Experimento

Lxperimento

Conclusiones

### **Trabajos Futuros**

- Extraer el algoritmo e implementarlo en un cliente
- Proxy reverso para que servidores que no soporten SPDY puedan ofrecerlo igual
- Implementar un CDN (Content Delivery Networks) con proxies conectados a través de SPDY
- En un futuro, poder utilizar HTTP 2.0



Proxy adaptativo

Pablo Maximiliano Lulic

Introducción

Crecimiento

Conceptos

preliminare Protocolos

Protocolos TCP

RTT

HTTPS

Problemática de la web

Proxy

Proxy Adaptativo

Experimento:

Resultados

Conclusiones

# FIN