Meu site Plone está lento. O que fazer?

Fabiano Weimar dos Santos
xiru@xiru.org

Roteiro

- Por que o Plone é Lento?
- Performance Tuning
- Dicas e Truques

Por que o Plone é Lento?

- Fato: qualquer software com grande quantidade de acessos simultâneos costuma ser lento.
- Não importa a tecnologia adotada!
- O que faz a diferença: atenção com a **infra-estrutura** e com a **qualidade**.

Fatos

- Quanto maior o tráfego de um portal, mais especializada deverá ser a solução.
- Frameworks genéricos geralmente não tem boa performance.
- Frameworks agilizam o desenvolvimento, mas costumam impactar na performance.

Performance Tuning

- Mesmo que você considere algo rápido, efetue medidas que avaliem a performance em cenários com centenas de usuários simultâneos.
- O fenômeno das "redes sociais" pode, sem aviso prévio, ser um DOS.

Não Chute!

Adote métricas.

Ferramentas

- Medir performance de websites não é uma tarefa trivial.
- Ferramentas dificilmente simulam adequadamente o comportamento das aplicações “reais”.
- Há muitas ferramentas de Benchmark
  - http://www.opensourcetesting.org/performance.php

Ferramentas

- Apache Benchmark - ab
  - Acompanha o Apache HTTP Server
  - Costuma ser a forma mais simples de medir a performance de um “request”
  - Não avalia o carregamento de css, javascript, imagens, etc.
  - Não leva em consideração a eficiência do cache dos navegadores.

Ferramentas

Siege - HTTP load testing and benchmarking utility

http://www.joedog.org/index/siege-home

Ferramentas

Flood - a profile-driven HTTP load tester

http://httpd.apache.org/test/flood/

É importante ter uma medida de quanto rápido é um site, com um determinado número de acessos simultâneos.

Dicas de Otimização de Performance

- Geralmente é difícil dar dicas "genéricas" de otimização de performance
- Em linhas gerais, quanto menos páginas o servidor processa, mais rápido é o site como um todo (cache)

Dicas de Otimização de Performance

- Nem todo site precisa de cache
  - Há situações onde fazer cache significa apenas mais processamento e, de fato, queda de performance (Youtube)
  - Desde que não exista processamento envolvido, nada é mais rápido do que servir conteúdo estático

Apache mod_rewrite

```
RewriteCond /PATH_WWW/DOMINIO/%{REQUEST_FILENAME} -f
RewriteCond %{REQUEST_FILENAME} !""
RewriteCond %{HTTP_COOKIE} !__ac=
RewriteCond %{HTTP:Authorization} !""
RewriteCond %{HTTP:If-None-Match} !""
RewriteRule ^/(.*) /DOMINIO/$1 [L]

RewriteCond /PATH_WWW/DOMINIO/%{REQUEST_FILENAME}/index.html -f
RewriteCond %{REQUEST_FILENAME} !""
RewriteCond %{HTTP_COOKIE} !__ac=
RewriteCond %{HTTP:Authorization} !""
RewriteCond %{HTTP:If-None-Match} !""
RewriteRule ^/(.*) /DOMINIO/$1 [L]
```

Apache mod_rewrite

- $ cd PATH_WWW
- $ wget -m -np http://DOMINIO

Cache Compartilhado

- Nem sempre servir apenas conteúdo estático é viável
- CMS geralmente não se preocupam em fazer “static deploy”; apenas publicam páginas dinâmicas
- A solução mais simples costuma ser adotar uma camada de webcache (Squid, Varnish, etc)

Cache Compartilhado

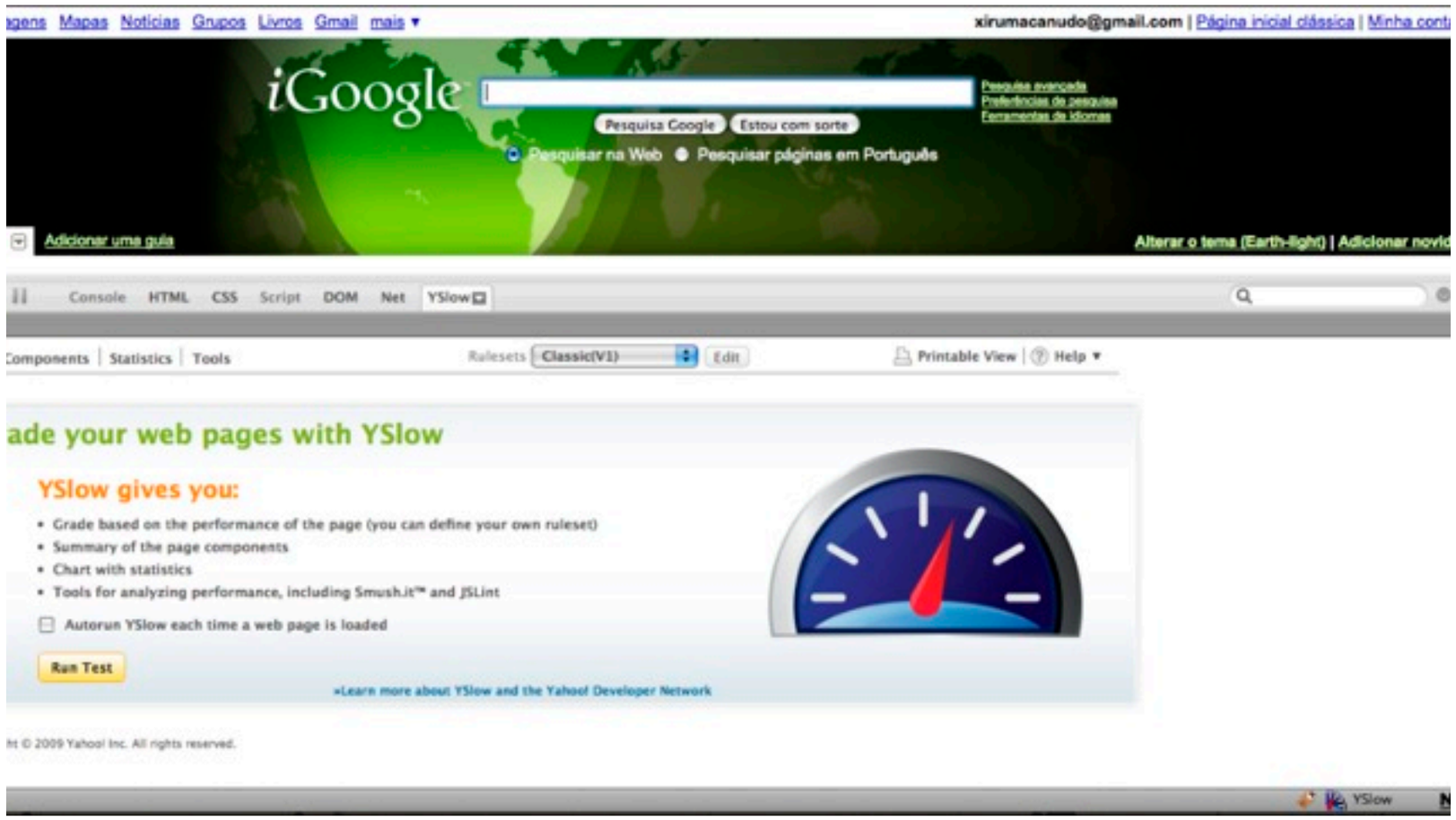
- Use o proxy para barrar tráfego indesejado
  - Bots de Indexação (Googlebot, msnbot, Yahoo Slurp, etc) podem significar até 35% do tráfego
  - Tráfego de bots é bastante custoso, pois não tem caráter repetitivo, desconsidera a relevância e visita o conteúdo "em profundidade"

Bloqueio de Bots (Squid)

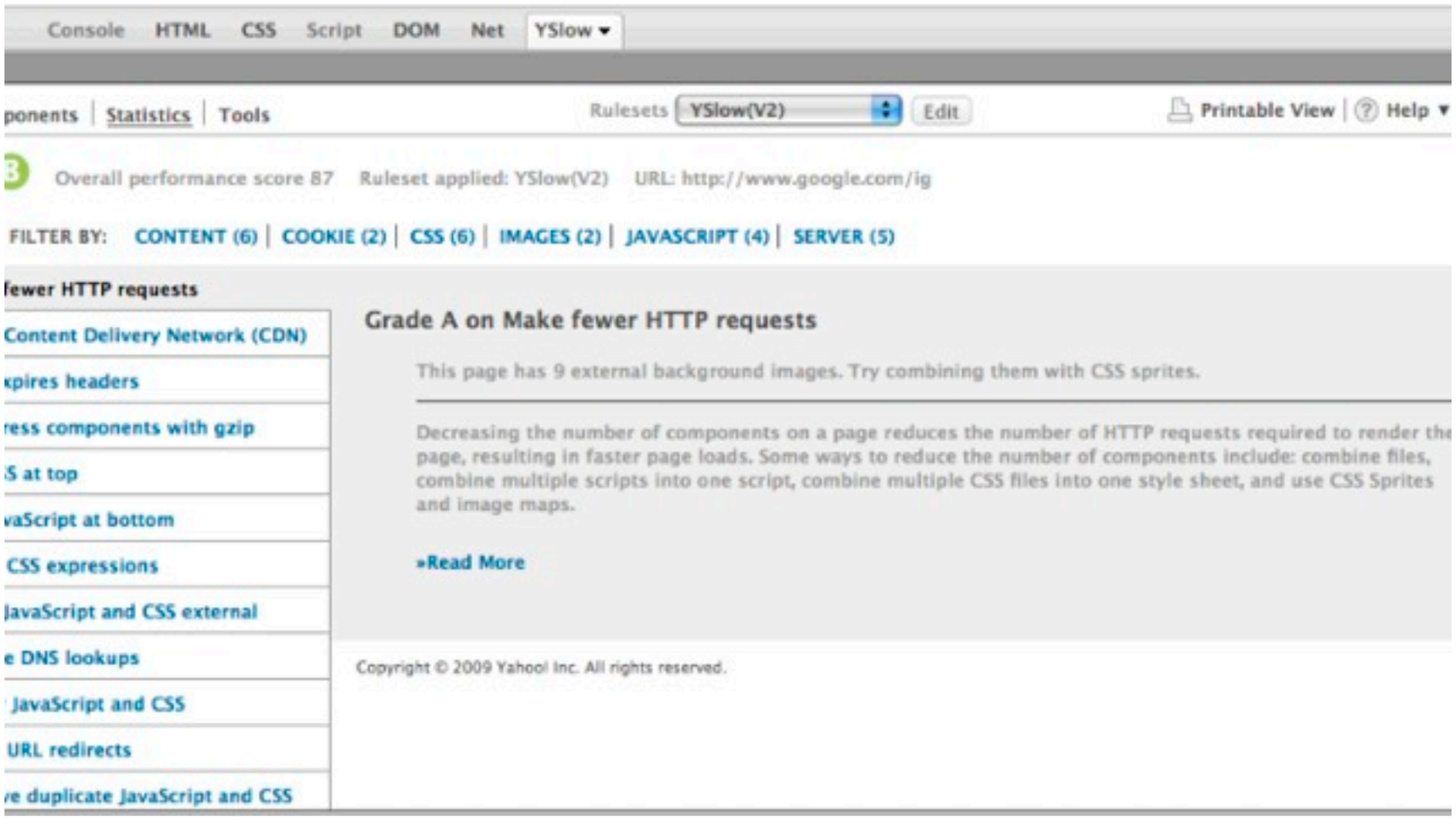
```
acl badrobot browser -i Twiceler

acl badrobot browser -i Yeti

acl badrobot browser -i Daumoa

http_access deny badrobot

acl bot browser -i bot

acl bot browser -i crawler

acl bot browser -i Slurp

acl horario_comercial time MTWHF 06:00-23:00

http_access deny horario_comercial bot
```

Cache Compartilhado

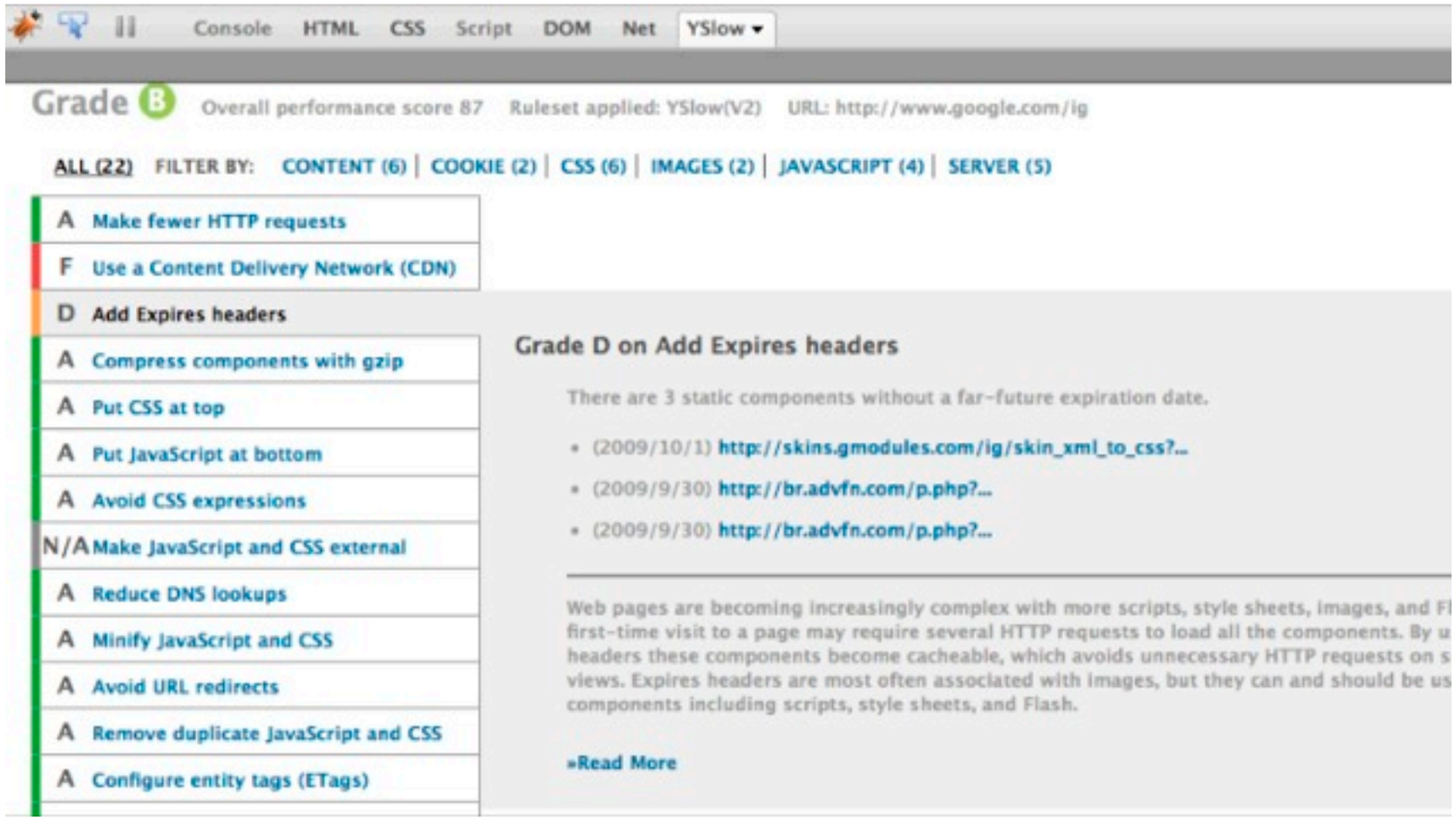
- Para que o cache compartilhado de páginas dinâmicas seja efetivo é importante que as páginas sejam geradas com cabeçalhos HTTP adequados
- Há extensões do Firefox que permitem a análise facilitada de “headers” HTTP, como a “web developer,” firebug, YSlow (minha predileta)


agens Mapas Notícias Grupos Livros Gmail mais ▾
xirumacanudo@gmail.com | Página inicial clássica | Minha cont
iGoogle
Pesquisa Google
Estou com sorte
Pesquisar na Web
Pesquisar páginas em Português
Pesquisa avançada
Preferências de pesquisa
Ferramentas de idiomas
Adicionar uma guia
Alterar o tema (Earth-light) | Adicionar novid
Console HTML CSS Script DOM Net YSlow
Components | Statistics | Tools
Rulesets Classic(V1) Edit
Printable View | Help ▾
ade your web pages with YSlow
YSlow gives you:
• Grade based on the performance of the page (you can define your own ruleset)
• Summary of the page components
• Chart with statistics
• Tools for analyzing performance, including SmushIt™ and JSLint
Autorun YSlow each time a web page is loaded
Run Test
»Learn more about YSlow and the Yahoo! Developer Network
ht © 2009 Yahoo! Inc. All rights reserved.
YSlow

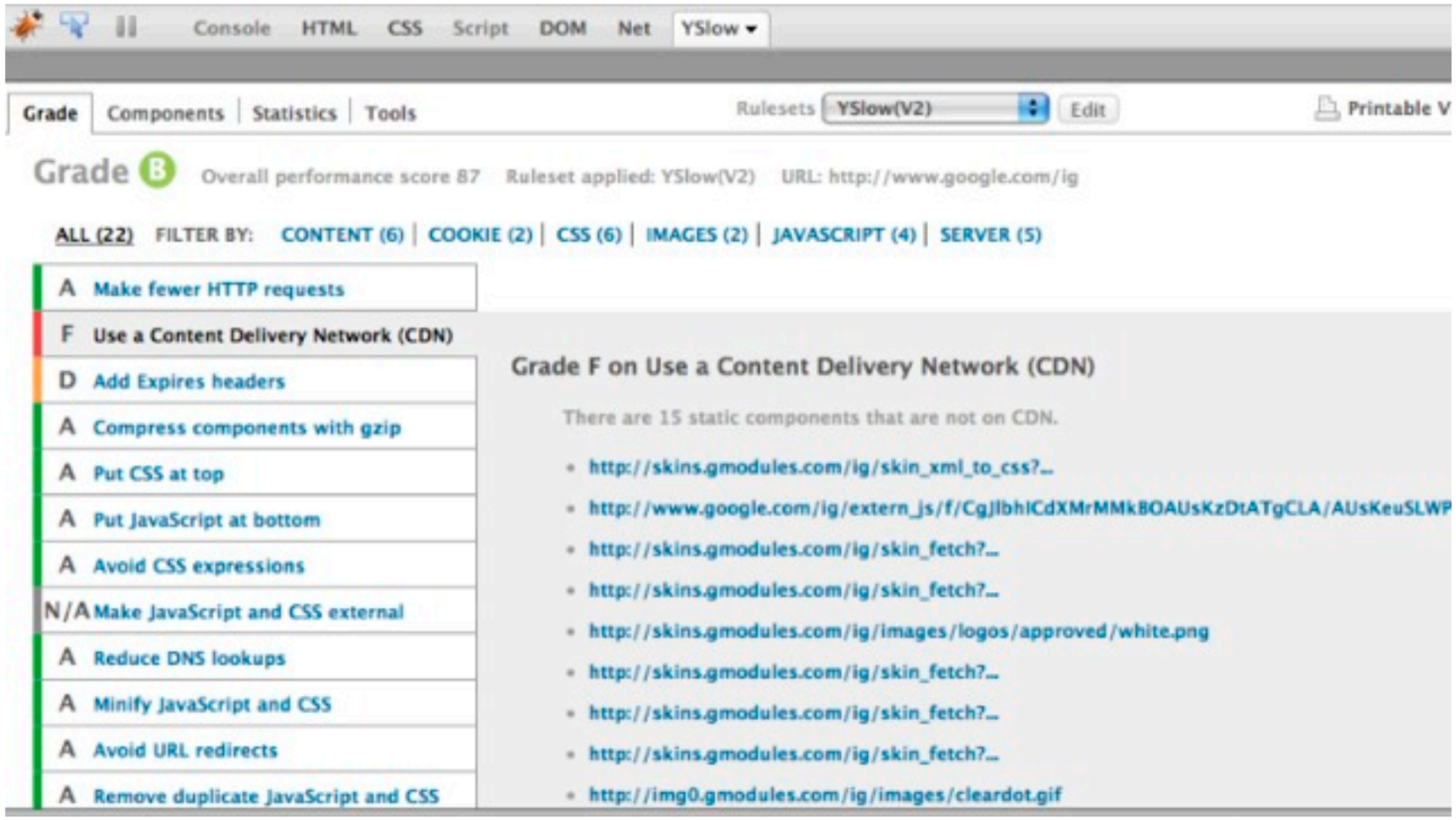
YSlow



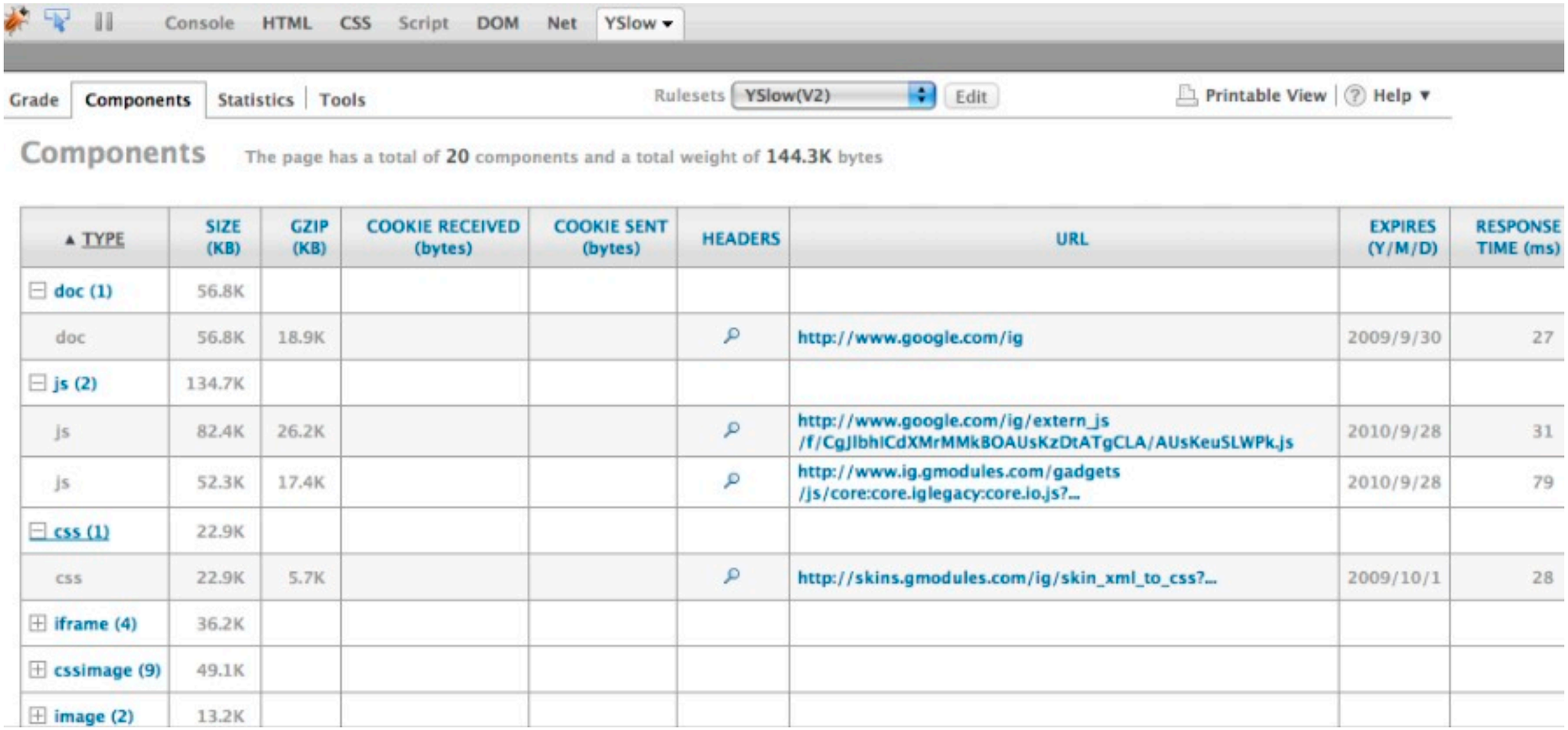
Fewer HTTP requests



Add Expires headers



Use a Content Delivery Network (CDN)



YSlow - Components

Grade | Components | **Statistics** | Tools Rulesets YSlow(V2) Edit Printable View | Help ▾

Statistics

The page has a total of **20** HTTP requests and a total weight of **144.3K** bytes with empty cache

WEIGHT GRAPHS

Empty Cache

HTTP Requests - 20

Total Weight - 144.3K

1	HTML/Text	18.9K
2	JavaScript File	43.7K
1	Stylesheet File	5.7K
4	IFrame	13.4K
9	CSS Image	49.1K
2	Image	12.2K
1	Favicon	1.1K

Primed Cache

HTTP Requests - 8

Total Weight - 50.3K

1	HTML/Text	18.9K
1	Stylesheet File	5.7K
4	IFrame	13.4K
2	Image	12.2K

YSlow - Statistics

Apache mod_expires

ExpiresActive On

ExpiresDefault "access plus 5 minutes"

ExpiresByType image/gif "access plus 1 day"

ExpiresByType image/jpeg "access plus 1 day"

ExpiresByType image/png "access plus 1 day"

ExpiresByType text/css "access plus 1 day"

ExpiresByType application/x-javascript "access plus 1 day"

Dicas de Otimização de Performance

- Cuidado com dependências remotas
  - browser -> application server -> webservice -> sgbd -> storage...
- Num cluster, o maior custo costuma não ser apenas processador, mas sim a latência de rede

Dicas de Otimização de Performance

- Evite acordar muitos objetos no SGBD
- Evite conexões com SGBD
- Evite conexões LDAP
- Evite conexões HTTP (RSS, SOAP)
- Não faça conexões sem timeout

Não tente advinhar porque seu site está lento: use um profiler!

Dicas de Otimização de Performance

- Use NTP para sincronizar relógios dos proxies, servidores de aplicação e banco de dados
- Evite gerar páginas dinâmicas com headers que expiram muito rápido
- Falta de sincronia de relógios pode acarretar a geração de headers “no passado”

Obrigado!

Fabiano Weimar dos Santos
xiru@xiru.org