HTTPArchive analysis of the usage of getMatchedCSSRules

meade@chromium.org

Analysis (2017_03_15)

I ran a BigQuery over the HTTPArchive data collected on 15 March, 2017 (2017_03_15_chrome_requests_bodies). The full data can be viewed in this spreadsheet.

Results:

- There were 277 unique domains.
- The total number of returned pages was 3325.
- Blogger.com was the most commonly appearing domain that used getMatchedCSSProperties when filtering by javascript content, followed by teads.tv (1502 and 1234 appearances respectively). Both of these domains used code that would break if getMatchedCSSProperties was removed.
- The next two largest domains were blueiconic.net and evernote.com, with 173 and 100 appearances respectively. Neither contained code that would break.

Summary table

Domain	count	breaks?
blogger.com	1502	
cdn.teads.tv	1234	yes
cdn.blueconic.net	173	no
static.evernote.com	100	no
all other domains	316	various

note: random sampling of the other domains suggests most are copies of Evernote's JS, <u>this polyfill</u>, or this <u>bootstrap-tokenizer</u> library, which won't break if getMatchedCSSRules is removed.

Actions taken

I submitted a CL to remove the usage in Blogger, which was pushed out at the start of April, causing the usage of getMatchedCSSRules to drop from 0.07% to 0.0036%.

We also contacted teads.tv's developer relations team, and they also removed their usage in mid-April.

Analysis (2017_08_01)

I re-ran the query using data from August 1, 2017 (2017_08_01_chrome_requests_bodies). The full data can be viewed in this spreadsheet.

Results:

- The total number of unique domains has increased to 294.
- The total number of returned pages has decreased to 656.
- blogger.com and teads.tv have almost disappeared as domains using getMatchedCSSRules (blogger.com appeared 6 times, potentially due to cached URLs hitting old endpoints?).
- As expected, the largest users are now blueiconic.net, and evernote.com (236 and 72 respectively), however neither domain's code will not break if getMatchedCSSRules is removed.
- A new user (siemens.com) has appeared (10 appearances), but they also will not break (they appear to be importing a polyfill).
- No other domains had 10 or more appearances.
- When random sampling of small domains I still didn't come across any that would break.

Summary table

Domain	count	breaks?
cdn.blueconic.net	236	no
static.evernote.com	72	no
blogger.com		yes
all other domains	351	various

BigQuery query used

```
SELECT bodies.page, bodies.url
  FROM [httparchive:har.2017_03_15_chrome_requests_bodies]
  AS bodies

JOIN (
  SELECT
    page, url,
        JSON_EXTRACT(payload, '$._contentType') AS contentType
  FROM [httparchive:har.2017_03_15_chrome_requests]
) AS requests ON requests.url=bodies.url AND requests.page=bodies.page
```

```
WHERE (requests.contentType CONTAINS 'javascript')
AND REGEXP_MATCH(body, 'getMatchedCSSRules')
```