

Proposed Standard

WD-countmethod-19981023

Internet Advertising Banner Counting Methodology

Working Draft *WD-countmethod-19981023*

Latest version:

<http://www.netgravity.com/standards/WD-countmethod.html>

This version:

<http://www.netgravity.com/standards/WD-countmethod-19981023.html>

Author:

Tom Shields <tshields@netgravity.com>

Status of this document

This is a Working Draft for review by IAB members and other interested parties. The IAB has NOT approved this draft for implementation; this draft is likely to change somewhat before approval. Anyone using this as a basis for implementation deserves what they get.

Abstract

A standard, practical methodology for counting Internet banner ad impressions and clicks is presented. The methodology is designed such that two compliant implementations will generate standard impression and standard click counts that differ by less than 5%. The methodology is implementable by content sites, content networks, ad networks, advertisers, agencies, and audit firms in order to achieve comparable counts.

Introduction

Internet advertisers often place similar ad buys across multiple web sites or advertising networks, and they require the ability to compare the results of those buys in order to evaluate the value they have received. Current Internet technology does not permit perfect accuracy in ad measurement, so ad delivery systems must use approximate methods of counting. Because there is no standard way to count, the performance numbers reported by different sites and networks are often measured using different methods, making comparisons between them invalid. Browser and proxy server issues often make even similar counting methods result in wildly differing numbers. Nevertheless, advertisers want reports with numbers that are comparable—advertisers want to compare apples to apples, not apples to oranges.

The counting methodology has the following design goals:

- **Comparable** - compliant implementations will generate numbers that differ by no more than 5%
- **Accurate** - count impressions and clicks as accurately as technically possible, while maintaining comparability
- **Implementable** - can be implemented with reasonable cost by sites, networks, advertisers, agencies, and audit firms
- **Practical** - takes into account the vagaries of the Internet, including browser bugs and non-standard caches
- **Efficient** - minimize latency and permit caching wherever possible, while maintaining accuracy
- **Simple** - this standard only addresses counting impressions and clicks for image banner ads

There are two basic methods for ad counting in use on the majority of the Internet today: ad requests (sometimes also called

ad HTML is requested. The ad download method counts an ad impression when the ad media (in this case, an image) is requested from the server. These two basic methods count at very different points in the communication channel between the browser and the server, and therefore produce different results. Most ad networks count using a variation of the ad download metric; it is often not technically possible for them to count ad requests because they have no access to the server that serves the pages containing their ads. Ad-supported web sites use variations on either metric. For the many sites that use variations of the ad request method, however, it should be technically possible for them to switch to an ad download metric, although it may require additional development and hardware resources. For this reason, the standard impression methodology described below is based on an ad download method.

The IAB MMTF has already published definitions of such terms as "ad requests" and "ad clicks" [\[IAB97\]](#). These definitions have been invaluable to ensure that companies using the defined concepts and metrics for advertising are consistent. That document also states: "It is important to note that for true comparability to exist, we need to define both the concepts and the metrics themselves as well as the methodology sites should use to generate those metrics."[\[IAB97\]](#) This document is an attempt to define a standard methodology, so the resulting reports can be truly comparable. This document also attempts to address the objections raised in the Q&A section titled "Why are images not a comparable measure?" [\[IAB97\]](#). In particular, by defining the methodology clearly, we hope to mitigate the effects of caching and other "environmental" factors on the comparability of the counts, resulting in a measure that is both implementable by ad networks, and potentially more accurate than ad requests.

There have been many incompletely defined terms used to describe measurement methodologies to date, including "impressions", "requests", "downloads", "insertions", "views", "exposures", etc. To distinguish this methodology from others, counts performed according to this method should be labelled "standard impressions" and "standard clicks".

For simplicity, this standard only addresses measuring ads that use or include clickable image media, including GIFs, animated GIFs, and JPGs. These ads constitute the vast majority of the advertising on the Internet today. No attempt is made in this document to define a methodology that can measure HTML ads such as banner forms (except to the extent that banner images within them can be measured), Java ads, Javascript ads, or embeddable media such as Shockwave; these may be defined in a later revision. Further, this standard does not attempt to account for client-side counting by offline browsers, or filtering of hits from non-human browsers such as spiders and robots. Finally, this standard does not define the user action required to measure a standard impression; in particular, it makes no distinction between a standard impression as a result of a user-initiated event or one resulting from a timed refresh.

Methodology

We will define an ad counter as a program that responds to browser requests related to advertising. For the purposes of this document, these requests will only include IMG SRC requests for ad media, and A HREF requests for ad clicks.

A valid standard impression may only be counted when an ad counter receives and responds to a request for an ad image from a browser. This image request must be the result of an IMG tag in the page HTML. In response, the ad counter must return a location redirect, specifying the location of a file or other program that will deliver the image media. The location redirect must take the following form:

```
302 Moved Temporarily
Location: http://www.site.com/ad.gif
Pragma: no-cache
Cache-Control: no-cache
[Set-Cookie: A=A]
```

A valid standard click may only be recorded when an ad counter receives and responds to a click request from a browser. This click request must be the result of a user clicking on an Anchor tag in the page HTML. In response, the ad counter must return a location redirect, specifying the location of the destination for the ad. The location redirect must take the following form:

```
302 Moved Temporarily
Location: http://www.advertiser.com/index.html
```

The response in both cases must be a location redirect - implementations which respond with a valid page or a status code other than 302 will NOT count a standard impression or standard click. The response may NOT contain the Last-Modified header, as it may encourage caching. Other headers may be used if desired. Note that the Set-Cookie header MUST be used if the originating site does not ensure the IMG SRC URL is unique across page requests.

The URLs that are used to make the IMG SRC and A HREF requests may take the form of any valid URL (see [RFC1738](#)). If the Set-Cookie header is not desired for perceived privacy reasons, the IMG SRC URL MUST be unique for each page request by a single browser, in order to prevent browser caching. The URLs that are then redirected to may be any valid URL. In many cases, the ad counter functionality will be included in a more full-featured ad server, which chooses the appropriate image, or determines the correct ad destination, and may require additional parameters as part of each URL.

Commentary

This counting methodology is both accurate and makes efficient use of Internet resources. Conceptually, a small control request must go "end-to-end" from the browser to the origin server, in order to ensure accurate counting (and, possibly, to select the correct ad). The ad media, however, may come from a cache as close to the browser as possible.

The methodology requires the ability to defeat caching on a location redirect, in order to count accurately and efficiently. There are several ways that caching can occur, most commonly either in the browser or in an intermediate proxy server. There are also several mechanisms for defeating caching, including response headers, and URL construction techniques. These and other issues are discussed in this section.

The mechanism chosen here to defeat proxy caching is to use the headers "Pragma", and "Cache-Control". These methods should defeat most proxy caches that incorrectly consider a 302 to be cacheable. In addition, the omission of the "Last-Modified" header should help prevent caching. The "Set-Cookie" header also prevents proxy caches from storing documents, but it has social implications that make its use often undesirable, therefore it is optional if the originating site ensures that the IMG SRC URL is unique across page requests.

To increase accuracy at some expense of simplicity, this standard requires either the IMG SRC URL to be unique across page requests by a single browser, or the Set-Cookie header in the IMG response. These are the only known consistent methods of defeating browser caching of images. If cookies are allowed, the Set-Cookie header must simply be present, it may set any name-value pair, and may be any kind of cookie. If cookies are undesirable, then the IMG SRC URL must be constructed to be unique across page requests.

One simple method for ensuring IMG SRC URL uniqueness is to insert the current time with seconds, or a sufficiently large random number, in the IMG SRC URL as the page is delivered to the browser. Other methods involve client-side scripting using Javascript, or server side includes. It is not sufficient to ensure that IMG SRC URLs on different pages are different, because a single browser reloading the same page should generate multiple standard impressions. Note that the server side modifications occur at the originating site that delivers the page, which may not be the site hosting the ad counter. Practically, many ad server systems require this unique identifier as part of both the IMG SRC and the A HREF URLs, to link the image served with the appropriate clickthrough without using cookies.

There have been reports of browser bugs causing incompatibilities with this counting methodology. The most common problem is that in many versions of Netscape, animated GIFs set to rotate continuously will only rotate once, and then stop. This problem has been fixed in the latest versions, and should become less important as the browser population upgrades. Some versions of Netscape may also continuously re-request animated GIFs if the "Expires" header is used. For this reason, we do not include the "Expires" header as part of the cache-defeating mechanism.

The methodology has been designed to handle multiple cascading ad counters. For example, a small site may accept local advertising, but send the rest of its inventory to a larger ad network. The network, in turn, may accept advertising from a large advertiser who serves their own ads, or uses a third party ad server. In this case there are at least three ad counters that will be involved in each ad delivery. This methodology has been designed so that the standard impression and standard

click numbers produced by each counter should be within 5%

Many sites are already performing ad measurement according to their own methodology, and may have contracts with advertisers based on that method. The implementation period for adoption of this standard will need to include time to benchmark the variances which may occur between different ad counters.

The HTTP 1.1 standard [\[RFC2068\]](#) recommends that servers responding to a request with a status 302 (location redirect) include an entity body that contains a short hypertext note with a hyperlink to the new URL. For maximum efficiency, and because nearly all browsers will automatically follow a 302, this method does not encourage an entity body in the 302 response. A "Content-Length: 0" header may be included, so browsers do not wait for the entity body to be transmitted, but it is not clear that this will actually improve performance, so this is not required.

As noted in the methodology, standard impressions may only be counted when an ad counter receives a request for an ad image from a browser, even though the measurement might be slightly more accurate if the counting were performed after the redirect was known to have transmitted successfully. This is much more difficult to implement with today's web servers, however, and would probably result in fewer sites counting this way. Also, because ad counters typically deal with requests in milliseconds, and the redirect is a relatively small response, the likelihood of successful transmission is high.

Some counting methods attempt to filter out impressions caused by non-human requests, such as spiders, robots, and pre-fetching browsers. However, most spiders and robots don't request images, so few standard impressions should be actually recorded. Also, requiring counting agents to filter against a potentially constantly changing list of known robots and spiders violates the simplicity constraint. Finally, robot activity can often be controlled with other standard mechanisms [\[KOST98\]](#).

Example

This example uses a hypothetical ad counter that is implemented as a CGI to a standard web server. The following is a sample page containing an ad that will be counted by the ad counter. Note the UTC parameter in the image URL - this is an example of how to make the image URL unique across browser requests.

```
<HTML><HEAD><TITLE>Ad Tester</TITLE></HEAD><BODY>
<H1>Ad Tester</H1>
<A HREF="http://ad.counter.com/cgi-bin/adcounter.cgi?DEST=http%3A%2F%2Fwww.advertiser.com%2Findex.html">
<IMG SRC="http://ad.counter.com/cgi-bin/adcounter.cgi?IMAGE=http%3A%2F%2Fwww.site.com%2Fad.gif&
UTC=908229511"></A>
</BODY></HTML>
```

When this page is downloaded by a browser, the browser will make a subsequent request to "adcounter.cgi?IMAGE" (assuming images are turned on). This hypothetical adcounter.cgi program would retrieve the correct image URL (in this case, from the query string), record a standard impression, and issue the following response:

```
Status: 302 Moved Temporarily
Location: http://www.site.com/ad.gif
Pragma: no-cache
Cache-Control: no-cache
```

The browser will receive this response, and automatically retrieve "ad.gif" and display it. Note that "ad.gif" may in fact be retrieved from an intermediate proxy cache, or even the browser's own cache. However, because the redirect has been marked non-cacheable, the next browser to retrieve the page will also make a call through to adcounter.cgi. Note also that had the IMG SRC URL not contained the UTC parameter, the response would have required a Set-Cookie header, as well.

When the user clicks on the ad, the browser will make a request to "adcounter.cgi?DEST". In this case, the ad counter will retrieve the appropriate destination URL (in this case, from the query string again), record a standard click, and issue the following response:

```
Status: 302 Moved Temporarily
Location: http://www.advertiser.com/index.html
Pragma: no-cache
```

Again, the browser will receive the response and retrieve the proper destination page. Because the location redirect defeats caching, click counting is not subject to cache discrepancies.

Note that multiple ad counters may be involved in counting standard impressions or standard clicks, in particular when a third party ad server is delivering advertising on behalf of an advertiser. To ensure comparability, it is important for all ad counters involved to adopt the same methodology. The first ad counter, for example, may redirect to another ad counter with a response like this:

```
Status: 302 Moved Temporarily
Location: http://www.othercounter.com/cgi-bin/adcounter.cgi?IMAGE=http%3A%2F%2Fwww.site.com%2Fad.gif&
UTC=908229511
Pragma: no-cache
Cache-Control: no-cache
```

Acknowledgements

Thanks to the MMTF for their invaluable work on Metrics and Methodology [\[IAB97\]](#), and to the IAB for providing the infrastructure to refine and agree upon standards proposals such as this one. Thanks in particular to Paul Hart from CNET, and Jim Jones from Discovery Online, for their technical work on standards proposals leading up to this one. Thanks also to Mike Griffiths from Matchlogic, and David Zinman from AdKnowledge, for their specific feedback on this proposal.

References

[IAB97]

IAB Media Measurement Task Force, [Metrics and Methodology for Internet Advertising](#), June 1997.

[RFC2068]

R. Fielding, et al, [Hypertext Transfer Protocol -- HTTP/1.1](#), January 1997

[RFC1738]

T. Berners-Lee, L. Masinter, [Uniform Resource Locators \(URL\)](#), December 1994

[KOST98]

Martijn Koster <m.koster@webcrawler.com>, [Robots Exclusion](#)

\$Id: count.html,v 1.7 1998/10/23 17:47:02 tshields Exp \$