

# ESTADÍSTICAS WEB

## ¿Qué es Estadísticas Web?

También conocida como analítica web o web analytics, es el proceso de analizar el comportamiento de los visitantes de un sitio web.

- Las estadísticas web son una herramienta importante para todo proyecto digital porque ayudan a saber el comportamiento de los usuarios y los buscadores.
- Una herramienta de análisis web es obligatoria en cualquier sitio web para poder realizar un seguimiento del tráfico y analizarlo.
- Aunque cada herramienta de análisis tiene sus propias características únicas, elegir la mejor herramienta en base a sus necesidades resulta de vital importancia.
- Con ellas, se puede ver qué contenidos son los más solicitados, desde qué sitios se recibe más tráfico, con qué palabras clave, etc. Hay páginas en Internet que ofrecen sistemas de estadísticas gratuitos muy buenos.
- Por otro lado, usar múltiples herramientas de análisis no tiene demasiado sentido si se escoge bien, y además afectará negativamente el rendimiento del sitio web, por lo que se recomienda que uses solamente 1 herramienta que ofrezca todas (o casi todas) las características que necesitas.

# ESTADÍSTICAS WEB

## CARACTERÍSTICAS

- Los servicios de estadísticas son sistemas para medir el tráfico de un sitio web. Sirven para contabilizar visitas, impresiones de páginas, usuarios únicos, etc. Así como monitorear el tipo de usuarios que visita, los caminos habituales en sus accesos, páginas de entrada y de salida, URLs desde las que se ha entrado al sitio web (referidos), etc.
- Los servicios de estadísticas utilizan diversos métodos para obtener la información de la que extraer los informes del tráfico: Algunos utilizan los **logs de accesos** del servidor y otros utilizan un **código Javascript** que se incorpora en las páginas web y que manda datos al servidor de estadísticas, con los que contabilizan los accesos.
- Lógicamente, los servicios que analizan los log, tienen que instalarse en el propio servidor para funcionar, algo que puede ser imposible para un usuario de un alojamiento web compartido y para un usuario de un servidor dedicado, difícil de instalar y configurar. Estos servicios de análisis de logs suelen estar incorporados entre los programas instalados en los servidores.
- Los servicios de estadísticas que funcionan incorporando un código Javascript son perfectamente fiables y sencillos de implementar en cualquier sitio web. El código Javascript necesario para realizar el conteo estadístico se suele colocar en la parte del código de la página. El navegador, la primera vez que accede al sitio web, tarda un poco más en terminar de cargar la página, pues se tiene que descargar el código Javascript. Sin embargo, en las siguientes páginas que solicita o en sucesivas visitas, ese código ya se encuentra en la caché del navegador, por lo que no se requiere que se vuelva a descargar, de modo que no se ralentiza la descarga de la página web.

# ESTADÍSTICAS WEB

## CARACTERÍSTICAS

- Los servicios de código abierto y comerciales, utilizan este método de inserción de un código Javascript, puesto que no se necesita instalar nada en el servidor donde está la página web, simplemente colocar el código en todas las páginas que se desean monitorear estadísticamente.

Existen multitud de servicios de estadísticas, muchos de ellos utilizan herramientas gratuitos o de pagos y otros de código abierto. Cada uno con su especificidad.

# ESTADÍSTICAS WEB

## ALGUNAS HERRAMIENTAS

**Google Analytics:** es una herramienta de analítica web de la empresa Google lanzada el 14 de noviembre de 2005. Ofrece información agrupada del tráfico que llega a los sitios web según la audiencia, la adquisición, el comportamiento y las conversiones que se llevan a cabo en el sitio web.

Además de hacer un seguimiento de los visitantes, también da recomendaciones para mejorar el sitio web. Para ese propósito, ofrece consejos para acelerar el sitio web, administrar contenido, conocer el comportamiento de los visitantes, etc.

**StatCounter:** es un sitio web de análisis de tráfico web iniciado en 1999. El acceso a los servicios básicos es gratuito y los servicios avanzados pueden costar entre \$5 y \$119 USD al mes. Tiene su sede en Dublín, Irlanda.

Ofrece un informe en tiempo real que muestra el tiempo total dedicado a cada página, el número de visitantes, por donde entran y salen, el contenido al que hacen clic...

Ofrece seguimiento gratuito para hasta 250 000 visitas por mes. Esto es más que suficiente para la mayoría de proyectos web, equivale a una media de unas 8 300 visitas diarias.

Ofrece contador invisible (no se muestra nada a los visitantes de tu web), la actividad de tus usuarios antes y después de visitar tu sitio web, mapas de calor, comparación entre motores de búsqueda, ver qué hacen los visitantes en tiempo real, rastrear sitios web HTTPS, incluso se pueden mostrar estadísticas del sitio web públicamente.

# ESTADÍSTICAS WEB

## ALGUNAS HERRAMIENTAS

**Piwik** Es ahora **Matomo** es una aplicación web libre y de código abierto, escrita por un equipo internacional de diseñadores que corre sobre un servidor web PHP/MySQL. Rastrea en tiempo real páginas vistas y visitas de un sitio web y muestra reportes de estos datos para su análisis. Hasta septiembre del 2015, Matomo fue usada por cerca de 900 mil sitios web y ha sido traducida a más de 45 idiomas. Fue conocida como Piwik hasta el 10 de enero de 2018. Piwik se lanzó a fines de 2007.

**Matomo** muestra informes con respecto a la ubicación geográfica de las visitas, origen de las visitas (es decir, si procedían de un sitio web, directamente o de alguna otra manera), las capacidades técnicas de los visitantes (navegador, tamaño de pantalla, sistema operativo, etc.), lo que la visitantes hicieron (páginas internas visitadas, acciones que tomaron, cómo dejaron el sitio web), el tiempo de visitas y más.

**AWStats** es una herramienta open source de informes de análisis web, apta para analizar datos de servicios de Internet como un servidor web, streaming, mail y FTP. AWstats analiza los archivos de log del servidor, y basándose en ellos produce informes HTML. Los datos son presentados visualmente en informes de tablas y gráficos de barra. Pueden crearse informes estáticos mediante una interfaz de línea de comando, y se pueden obtener informes on-demand a través de un navegador web, gracias a un programa CGI.

**AWStats** soporta la mayoría de los formatos de archivos log de servidor web conocidos, entre ellos Apache (formato de log NCSA combinado/XLF/ELF o formato común/CLFt), WebStar, IIS (formato de log del W3C) y muchos otros formatos comunes de Internet. Los desarrolladores pueden contribuir con el proyecto AWStats a través de SourceForge.

# ESTADÍSTICAS WEB

## ALGUNAS HERRAMIENTAS

**Webalizer** es un software de análisis de registros web, que genera páginas web de análisis, a partir de registros de acceso y uso.

Es una de las herramientas de administración de servidores web más utilizadas. Fue iniciado por Bradford L. Barrett en 1997.

**GoAccess** es una aplicación de análisis web de código abierto(open source) para sistemas operativos similares a Unix. La aplicación tiene una interfaz de usuario basada en texto y una aplicación web. GoAccess puede proporcionar análisis en tiempo real al monitorear continuamente los registros del servidor web.

Es una aplicación para mostrar gráficamente la información agregada sobre los logs de accesos. Funciona por consola, donde se le pueden pasar diversos parámetros. También existe un archivo de configuración donde se puede establecer todo lo necesario para operar. En este sentido, la herramienta es muy flexible, se puede adaptar al formato de logs que se tenga definido.

Es un visor interactivo y analizador de registros web en tiempo real que se ejecuta en una terminal o a través de un navegador.

Proporciona estadísticas HTTP rápidas y valiosas para administradores de sistemas que requieren un informe visual del servidor sobre la marcha.

# ESTADÍSTICAS WEB

## ALGUNAS HERRAMIENTAS

### **GoAccess**

Como resultado genera un reporte HTML bastante atractivo y al mismo tiempo puede generar reporte JSON y CSV, donde se puede ver gráficamente el contenido de los accesslogs analizados. Vale destacar que se puede configurar para analizar uno o varios archivos de logs.

**GoAccess Es la herramienta que se está utilizando actualmente para la estadísticas web de infomed en sustitución de Piwik**

# ESTADÍSTICAS WEB

## HERRAMIENTA GoAccess en INFOMED

La estadística web de los sitios del hosting de infomed se están generando con la herramienta GoAccess. La misma se logra a través de un conjunto de scripts programado en Bash que ejecutan el demonio **goaccess** pasando como parámetro archivos de logs. A este servidor donde se generan las estadísticas se le llama **Servidor Goaccess**.

Para agregar un nuevo sitio a esta estadística, hoy se ejecutan de forma manual una serie de pasos:

1 - En el servidor donde se hostea el sitio que se agregará la estadística.

1.1 El servidor de publicación ya sea Apache, Nginx u otro debe estar generando los archivos de logs en un formato adecuado.

1.2 Los archivos de logs deben cumplir con la sintaxis correcta de nombre del archivo>(\*access.log-\$(date '+%Y%m%d')\*.gz).

1.3 Los archivos de logs deben estar rotando de forma diaria con los nombres y sintaxis del paso 1.2.

1.4 Debe estar instalado y configurado el servidor rsync. Que permita acceso a los logs salvado desde el servidor goaccess a través del cliente rsync.

2 - En el servidor Goaccess se realizan los siguientes procesos auxiliandose del crontab sistema:

2.1 Copia de logs a través de rsync desde los diferentes servidores Web Hosting: Un script que ejecuta 3 veces al día.

2.2 Añadir el nuevo sitio al registro de sitios de estadísticas que se nombra: sitios\_goaccess\_estadistica\_infomed.txt y que tiene el formato:

**sitio log cuanto servidor formato**

www.scielo.sld.cu,scielo-access.log,1,e2b10,1

www.sld.cu,access-nuevoportal.sld.cu.log,1,sv48,2

uvs.sld.cu,uvs.sld.cu\_access.log,1,sv10,1

webmail.sld.cu,webmail.sld.cu\_access.log,1,webmail,1

2.3 Se ejecuta durante las 24 horas del día los scripts: goaccess\_estadistica\_infomed\_param\_ma.sh, crea\_html\_estadistica\_web.sh y crea\_html\_estadistica\_web\_json.sh

# ESTADÍSTICAS WEB

## HERRAMIENTA GoAccess en INFOMED

La estadística se genera a partir del análisis de los archivos de logs diarios en forma incremental y persistente obteniendo como resultado un reporte mensual y otro anual en formatos html json y csv.

**Para ver Resultados: <http://sv50.red.sld.cu/>**