Server Log Files in a Nutshell

graylog.org/post/server-log-files-in-a-nutshell

Servers take a lot of requests daily, we know that...We also know that the server responds instantly. But who makes the request? What do they want, and what exactly are they looking for? Where do these visitors come from? How often they are making a request: once a month, once a day, every minute?

You can find answers to these and potentially a lot more questions in the server log file.

What is a server log file?

A server log file is a simple text document that contains all activities of a specific server in a given period of time (e.g.,one day). It is automatically created and maintained by the server, and it can provide you with a detailed insight into how, when, and by whom your website or the application was accessed.

Most servers generate CLF (Common Log Format) files, or raw log files, which are rather crude and difficult to understand for the most part. They also tend to become complex and comprehensive in a short amount of time, which is why the server automatically erases them after a designated period of time. For example, some servers are set up to delete log files after a certain amount of time (every 30 days), others are set up to delete them after the files goes beyond a certain size.

CLF files are files in which each line represents one request. Therefore, if the visitor clicks the link and gets an HTML file with three additional images on it, four lines of text will appear in the log file.

Each line contains an abundance of information, which may include:

- IP address and identity of the device making a request
- Name, location, and size of the requested file
- Time and date of the request, and the request method
- The referred webpage
- HTTP status code (if a file was not found, for example)

The log files are available only to administrative personnel, disabling general users from accessing them. Because these files provide an abundance of information, they are excellent tools to help you improve website administration efficiency, increase overall traffic to your site, or even fine-tune your SEO efforts.

WHY DO YOU NEED SERVER LOGS?

Server logs are an essential piece of data because they contain information that cannot be found anywhere else. For example, server errors and user access records along with a host of additional data. Having the ability to review these logs gives you the ability to determine what caused an issue on your server. You can also use them to pinpoint a potential security issue. Without server logs, you are left without a way to determine what is going on with your servers.

Why Should You Centralize your Server Logs?

If you have two more servers active, maintaining your server logs in a centralized location offers a number of key advantages, including.

- Saving time: Accessing all data in one place is much more time-effective than accessing all of your servers locally.
- Troubleshooting: Regular insight into centralized logs gives you a “before and after” picture of the status of your servers. Determining what went wrong and correcting the mistake quickly and easily is at the heart of optimum network administration.
- Reducing the possibility of losing data: If an individual server is down, there is no way to access its data locally. With a centralized hub, you can still access the data.
- Improving your network security: Reviewing server logs in a centralized place lets you easily and effectively pinpoint any unusual behavior (for example, an excessive number of login attempts in a short period of time) and adjust your network security accordingly.

Types of Log Files

Even though most websites rely on the CLF for their server log files, there are other formats as well.

Access Log

Access logs record information about which HTML files and graphic images are being requested from your server. Access logs tell you the number of visitors as well as their origins (did they come from the .com, .edu, or the .gov site, for example). Also, you can get usage patterns and information about which page visitors requested the most.

Agent Log

Agent logs can record and provide you with information about which web clients made requests on your server.

Referrer Log

Referrer logs record information about the URL the visitor was on immediately before moving onto your web page and making a request on your server. This type of log is extremely useful when you want to pinpoint the origin of the requests on your web server as well as which web pages are referring traffic to your server.

Error Log

Error logs record and provide you with information about a server's failed requests. Essentially, it automatically generates an error message as soon as someone tries to access a nonexistent file on your server.