



Data Driven Success

Comparing Log Analytics Tools: Flowerfire's Sawmill vs. Google Analytics (GA)

In business, data is everything. Regardless of the products or services you sell or the systems you support, good data is critical to your bottom line. It helps you better understand your markets, your customers, and your competition. It becomes actionable intelligence that drives important decisions.

Your back-end data is just as crucial. Luckily, your technology is smart enough to realize this and diligent enough to handle the task. Application and system logs patiently capture and record detailed data on traffic, transactions, and trends. These logs can provide answers to key questions.

- **How are your systems performing?** Log data helps your teams monitor bandwidth, disk space, and memory usage, deploying additional technical resources where required. At the first sign of trouble, those in the know head to the logs looking for answers.
- **Who are your users and what information are they making important?** Log data tracks who visits your sites, where they come from, and how they're using your systems. Monitoring site traffic allows you to identify popular content and features and spot early trends and new opportunities.
- **Are your systems protected from intrusion and other technical threats?** The logs are the front line in the battle against a world of ever-evolving threats. Real-time data driving real-time alerts is the best strategy for keeping your systems secure and important information private.

These questions are just the start. Ultimately, the overwhelming volume of log data serves a single purpose: ensuring your technology investments are meeting the real-world needs of your customers and partners. Your IT organization does a valiant job converting log data into answers you can use to plan, protect, and grow your business. But are they using the best tools for the job?

This paper looks at two popular log analysis solutions, *Flowerfire's Sawmill* and *Google Analytics (GA)*. Sawmill, first released in 1997, has long been considered an essential tool by IT organizations of all sizes. While Sawmill's complete feature set and support for more than 850 log formats are important, this paper is mostly limited to the Sawmill website analytics solution.

Google's Analytics, launched in 2005, is an updated and improved version of *On Demand Urchin*, another longtime IT favorite. *Urchin* was purchased and subsequently re-launched as *Google Analytics* in 2005.



Your first challenge: Defining your need

The first step in choosing a solution is defining your need. Once you better understand how your systems and users produce and consume data you can choose a solution that delivers maximum benefit across your entire organization. **The decision has to be about more than just data: it has to fit the way your technology and your users work.**

COMPARING THE SOLUTIONS

While their feature sets might look similar, *Sawmill* and *Google Analytics* work very differently. While *Google Analytics* uses a traditional web-based client side logging architecture, *Sawmill* imports and processes log data locally, creating a real-time data stream from server log to the database and then to your reports.

This white paper will compare the two products on a number of features and services including:

- > [Processes](#)
- > [Viewing and Sharing Data](#)
- > [Real-time vs. Batch Update](#)
- > [Compatibility](#)
- > [Performance](#)
- > [Extensibility](#)
- > [Support](#)

[COMPARISON]: **Processes**

[Back to Features list](#)

SAWMILL

Sawmill imports and aggregates selected log data in its database. A sophisticated web interface displays a hierarchically-aggregated set of ‘tuned’ reports specific to that log data format and filters can be used to further refine the data displayed.

Data can be refreshed in real-time, and easy yet powerful zooming lets you quickly drill down to find the important piece of data you need.

Google Analytics

Google Analytics monitors the unique site code and begins a combination of client and server side logging. A dashboard aggregates popular data views and offers access to reports. Data is refreshed intermittently, with some data being refreshed frequently and other views not being updated for more than 24–48 hours.



[COMPARISON]: Viewing and Sharing Data

[Back to Features list](#)

Each solution offers similar reports, and both support some level of customization. Sawmill's real-time data that allows instantaneous and up-to-the-second statistical reports, based on data from your log files RIGHT NOW, expands reporting possibilities, while Google's default reports are tightly focused on website analytics.

Viewing SAWMILL Data

Sawmill uses reports to manage how data is viewed. Since *Sawmill* can stream real-time log data, these reports are more than snapshots, displaying live data as requested. Real-time updates can be configured or disabled as required.

Customizing SAWMILL Reports

Since default reports are based on log format, *Sawmill* already has a pretty good idea of what you're looking for. It builds the default report view based on this intelligence, so you can immediately focus on the data that's relevant to you. Use checkboxes to quickly add or hide data from a report, and update it at any time with a click of the mouse.

The powerful customize report capabilities allow you to edit default reports as well as add customized reports to the profile. Creating charts and graphs (including interactive pivot tables) is easy. You can also format individual data elements within the chart, choosing the most appropriate way to display the data.

Sharing SAWMILL Reports

Sawmill's granular access control allows you to manage access to reporting features, letting you restrict how users access report features. Users can be restricted from adding or editing reports based on role as required. Reports can also be emailed to users outside of *Sawmill* with an easy to use interface, and reports can also be automatically sent to users at set intervals. In addition, *Sawmill Reports* can be downloaded and exported as PDF and CSV.

Using SAWMILL Filters

Sawmill's report filters make finding relevant data even easier by restricting the display of data that doesn't meet predefined requirements or contain specified values. For example, you could choose to only see data on traffic from a specific domain, or focus on views and downloads for a specific piece of site content.

Multiple filters can also be used simultaneously. Once the data is filtered, you can quickly zoom in on specific results. The dynamic data filtering continues as you drill down, making your pursuit of that one important metric even faster. And *Sawmill's* optimized import process guarantees your browsing stays quick and efficient.



Viewing Google Analytics Data

Google also uses reports to manage how data is viewed. Reports contain standard metrics data organized into dimensions. A dimension is essentially a report column, which in *Google Analytics* maps to an attribute of a site visitor, such as country of origin, browser in use, or the name of the campaign that brought the user to the site.

Customizing Google Analytics Reports

Google Analytics is built to analyze web traffic, so default reports are built around that task. The default dashboard displays common reports like page views, visits by source, and visitors (new vs. returning). This dashboard can be rebuilt as required, adding other reports and customized views of existing data.

Google Analytics also supports customized reports. Most of that customization involves removing data from existing reports, allowing you to create reports that highlight the data that is important to you and other users. However, *GA's* ability to expand reports to include more or a richer set of data is either clumsy or unavailable.

Sharing Google Analytics Reports

Custom reports are automatically available to users with access to that *Google Analytics* profile, and you can also set custom permissions that manage user report access. *Google Analytics* also allows you to quickly share reports with other *GA* team members and other non-*GA Google* users, and you can also schedule reports to be automatically sent to users at set intervals. *Google* reports can be easily shared with other *Google* users. *Google Analytics Reports* can be downloaded and exported as PDF, XML, CSV, and TSV files. Exporting report data is subject to limits set by the application.

Using Google Analytics Filters

Google Analytics report filters simplify data display by restricting the display of data that doesn't meet predefined requirements or contain specified values. For example, you may decide you'd only like to see visitors using a specific browser or coming from a specific referring URL. Search and replace filters can also be used to simplify the display of data like long unfriendly URLs.

[COMPARISON]: Real-Time Versus Batch Update

[Back to Features list](#)

The basic reporting capabilities of *Sawmill* and *Google Analytics* are similar, with one very important difference. *Sawmill* gives users direct access to real-time data, while *Google Analytics* users must wait on batch updates to refresh their data. *Google* refreshes some data every couple of hours, while some information is updated only once a day.



Does it matter?

There are lots of situations where batch data is sufficient, and *Google Analytics*' popularity with industry leaders of all sizes is proof of that. A historical view of last week's unique site visitors does not require real-time reporting. Even a summary of top search keywords can be one or two days old and not lose any relevance.

However, batch data is not suitable for more advanced metrics. Sites that rely on current events or other special promotions to drive traffic can't wait on day-old information. Real-time data allows those sites to capitalize on opportunities and prepare for challenges in ways that are beyond the reach of intermittent batch updates.

What about real-time Alerts?

Your IT staff is already using real-time alerts to monitor your technology. They know that problems almost always reliably show up in the logs, and real-time alerts guarantee they have the time and information needed to respond. Real-time reporting can add similar value to your website analytics.

- Monitor traffic from sites like *Twitter* and other 'real-time' feeds
- Provide instant feedback on the popularity of new content or features
- Watch for data spikes that might indicate trouble

Obviously, as you move away from web site statistics, real-time reporting becomes even more critical. Intrusion detection, server load balancing, and most maintenance and support services require real-time reporting and alerts. Many other essential business applications, including inventory and billing, also benefit greatly from real-time reporting and alert capabilities like those provided by *Sawmill*.

Even if you're just looking to monitor site traffic and generate historical reporting, the potential of real-time data remains powerful. As you choose a solution, having real-time capabilities should be considered a must have—waiting hours to access your own data is unacceptable.

[COMPARISON]: **Compatibility**

[Back to Features list](#)

SAWMILL

Sawmill runs on *Windows*, *Mac*, and *UNIX*. It can also run in CGI mode. Database support includes *MySQL*, *MS-SQL* and *Oracle ODBC*.

Google Analytics

Since all data is processed and stored on *Google* hardware, only a web browser is required to use *Google Analytics*.



[COMPARISON]: Performance

[Back to Features list](#)

SAWMILL

Sawmill's performance is based partially on the size of the log files being processed. There are two processes: data import and report generation. Multi-processor support means *Sawmill* processing can be virtualized, delivering an exponential increase in speed. Once the database is built, refreshing report views with live data is quick.

Google Analytics

Refresh rates for reports, when new data is available, is fairly quick. The page-tagging has been reported to slow down the actual web pages. Your website may show a slowdown once you implement *GA*.

[COMPARISON]: Extensibility

[Back to Features list](#)

The best solutions are always adding value. Both *Sawmill* and Google Analytics can be expanded beyond their default feature set, through the addition of other services or the use of the API (application programming interface).

SAWMILL

Designed and developed as a universal log tool, *Sawmill* is already integrated with 850 log formats for a wide range of hardware and software applications and appliances. Recent releases have expanded and improved *Sawmill's* web traffic analysis capabilities.

Sawmill's decision to add the *ClickStream* site traffic plug-in and the *GeoLite City* geographic overlay means *Sawmill* can now match nearly all of the client-side analysis feature set solutions like *Google Analytics* deliver.

The Power of SAWMILL's Salang

A good programmer can take your implementation of *Sawmill* even farther. Developers can use *Salang*, *Sawmill's* friendly and powerful extensibility language, to completely customize the web and report interface. This means that the user interface, the report interface, and even the log parsers can be modified and improved as your organization sees fit.

Google Analytics

When *Google* purchased *Urchin*, they knew integrating it with other popular *Google* marketing services was essential. As a result, *Google Analytics* works seamlessly with other popular applications like *AdWords* and *AdSense*. As *Google* expands that suite, those services should also plug into *Google Analytics*.



Data Export API

Google Analytics also offers an API. The Data Export API can be used to develop client applications that receive data from *GA*, as well as refine the returned data results via query. The API is currently still in beta, and only enable read-only access to your *GA* data. The *Google* API also places strict limits on the amount of data that can be shared between *GA* and your client application.

[COMPARISON]: Support

[Back to Features list](#)

Making the most of your analytics data can get complicated. Where can your organization go when you need help solving a technical issue or implementing a great new idea? Where can frustrated users turn when nothing seems to make any sense?

SAWMILL

Sawmill has made documentation available for both administrators and end users. The documentation is embedded right into the web app, so it's only a click away. *Sawmill's* long popularity with IT professionals also means there are forums of dedicated users and experts ready to offer you informal assistance.

Sawmill also offers official technical support. Forum and email support are free for all users, with Premium Support allowing live telephone access to technical support staff available for a yearly fee. *Sawmill* also offers fee-based consulting expertise to help you plan implementations, extend *Sawmill* services, or solve other issues.

Google Analytics

Google typically offers very little support for its products, even those aimed at the enterprise. This is true for *Google Analytics*, with no official support other than solid documentation and user forums (with answers that are generally helpful but never definitive). Luckily the widespread popularity of *Google Analytics* means there is a huge community of other *GA* users available for advice, even if it doesn't qualify as true technical support. If you're looking for more specialized support, *Google Analytics* consultants are also available for hire.

SUMMARY

If you have an IT organization of any size, some version of *Sawmill* is probably already hard at work somewhere, fearlessly parsing log data and getting the job done. It's a big tool with lots of power that far exceeds the narrow task of website analytics. So why not choose the more nimble *Google* solution, built by the very people who reinvented the web itself?

If *Google Analytics* offered functionality and value that couldn't be found elsewhere, this would be the smart choice. But the *Sawmill* solution matches *Google* nearly stride for stride, and web site analytics are just one of the application's many strengths. In the end, it is *Google* that fails to innovate, with no support for real-time data or real advanced customization. Lacking that edge, choosing *Sawmill* as your web analytics solution has become easier than ever.



Company profile: **Flowerfire**

Flowerfire, Inc., founded in 1996 and based in Santa Cruz, California, is a software company which provides log analysis and reporting solutions. Flowerfire's primary product is Sawmill, a universal log analysis tool which can analyze and generate reports from over 800 log formats, including log files from all popular web servers, media servers, security applications, mail servers, spam filters, firewalls, proxys, gateways, and FTP servers. In addition to Sawmill, Flowerfire also provides consulting services for log analysis and reporting solutions. Flowerfire is privately held, family-friendly, and employee owned and controlled.