# Log Collection and How It Works, in Detail

January 5, 2021
by SentinelOne

Log collection is the process of collecting log entries from many different sources in an organization and bringing them all to a single place. Why would this be a good thing to do?

It all comes down to knowledge. Logs are ubiquitous in a tech organization since many different kinds of processes generate them. Because of that, your logs contain data about your whole system. Through log collection, you can leverage all that data, uncovering useful patterns in it that you can turn into valuable knowledge.

In this post, we'll offer you a guide on log collection so you can understand more about what it is and what you and your organization can gain from it.

## Log Collection: What It Is and Why Care About It

The log collection definition you've just read, though succinct, brings further questions:

- How do you actually collect all of those log entries? Is this the work of a human being, or is there some automation involved?
- What are the many different sources we've talked about? How many types of logging are there after all?
- How is it possible to extract knowledge from log files? Aren't they used mostly for troubleshooting?

We'll now answer the above questions and more.

### Defining Log Collection

Let's go deeper into the definition you've just seen by answering some of the questions it entails. First of all, what about the sources we've mentioned?

As a software engineer, you're probably used to working with application logging (i.e., logs generated by applications). You've probably both implemented logging in the applications you've worked with and read through log entries to troubleshoot problems when the said applications misbehaved in production. But logging doesn't restrict itself to applications. Operating systems also generate logs, and so do web servers such as Apache or IIS. Database systems—of the relational kind or otherwise—also generate logs.

Logs can also differ in regards to their actual medium. Though we talk a lot about "log files," processes can—and do—send their log events to a multitude of different destinations. Text files are indeed one of the most popular destinations for log events. However, logs can be sent to a database table, an email, or even a cloud service, among many other possible destinations.

Log collection is all about bringing logs from all of those different sources and in those different formats into a single location. There are many challenges involved in performing log collection, especially when you take into account the realities of today's tech scenario, such as the cloud and microservices. Because of that, leveraging automation through the use of specialized tooling is essential.

### Why Should You Care About Log Collection?

Why is log collection worthwhile? Simply stated, putting all of the logs your organization generates into a single place gives you the ability to do useful things with them. What kinds of useful things? Glad you asked.

> **Simply stated, putting all of the logs your organization generates into a single place gives you the ability to do useful things with them.**

#### Log Collection Makes Troubleshooting Faster and Easier…

The first great capability when adopting log collection is a fast search feature. You see, one of the problems of log files is that they tend to get really big, which makes searching through them a giant pain. Using tried and true approaches like grep and regex might help you here, but they might not be enough when it comes to truly massive amounts of data. The specialized tooling you'd use for log collection often includes powerful and fast search capabilities.

Another common feature in the log collection process is parsing. You need this because logs from different sources might greatly vary in the formats they use. This could include time-related differences (e.g., date and time formats and conflicting timezones), different names for log levels, and even different approaches to the recording of the log message itself (i.e., text-based versus structured logging).

So, you can think of log collection—the process and related tools—as an amplifier for your logging approach. All of the goodies you get by adopting log collection tools can greatly improve your logging approach, making troubleshooting and incident management faster and easier.

#### …But It Doesn't Stop There

Using logging as a troubleshooting mechanism is already a great thing in itself. However, it's possible to take it a step further and use your logging approach in an active and predictive—rather than simply reactive—way. You do that by venturing into the realm of log analytics, and log collection is an integral part of getting there.

Log analytics means analyzing your log data so you can detect interesting patterns in them. The knowledge you extract from your log data can be useful in many different ways. For starters, log analytics tools that count with real-time monitoring can help you improve your incident response approach, fixing problems before they have time to wreak much havoc. Security is another crucial area that can get a boost from log analytics: the trends you observe can help your organization detect past incidents or even fix security vulnerabilities before they're exploited.

But why stop there? What if you could predict and prevent problems before they happen? With log analytics, you might be able to do just that, by comparing current trends with how things looked before and during past incidents.

Log analytics can also play a crucial role in business planning and decision making. It can help you correlate seemingly unrelated events and obtain insights that wouldn't be available otherwise. You can then use those insights to help you when making decisions—for instance, as evidence when making the business case for an investment or purchase.

### How to Perform Log Collection in Practice

With the what and why out of the way, it's time for us to answer the question that's left: how do you perform log collection?

#### Specialized Tooling Is Essential

As mentioned several times, you do log collection with the assistance of specialized tooling. This might sound somewhat counterintuitive if you just take into account the process name. After all, log collection just sounds like copying files over to a single location.

As we've seen through the post, log collection involves a whole lot more than just that. The real value that you can extract from specialized log collection tools lies in the additional features they provide. So, that's why you need dedicated, specialized tools to perform log collections. A bunch of scripts can take you only so far. That is, they can help you copy log files over to a different place and then stop there.

#### What Does Log Collection Capture?

What kind of event should you capture when doing log collection? The more, the better.

The value you get from log collection—and what comes after it, like log analysis—derives from the fact that logs are ubiquitous. All kinds of processes in an IT system generate logs, which means the accumulated knowledge you get from your logs represents a unique opportunity to see and analyze everything.

> **The more you capture, the more opportunities you have for extracting useful insights.**

That's why it's so important to capture pretty much every type of log: application logs, logs from web servers (both the software and hardware varieties), logs from the database and operating systems, and the list could go on. The more you capture, the more opportunities you have for extracting useful insights.

## Log Collection: Logging Can Be Much More Than Incident Management

Logging can be much more than a helper in troubleshooting and incident response. Unfortunately, many professionals and organizations fail to realize that potential, often due to a lack of knowledge.

In this post, we've talked about log collection. You've learned that log collection is the process of moving all of your logs from many different sources to a single location, making them easily searchable, among many other benefits. Through the use of log collection—and what it facilitates, like log analysis—you can take your logging approach to the next level. Instead of just putting out fires, you might be able to prevent them from happening in the first place.

After learning about log collection, your next step should be looking at some of the tools at your disposal. We invite you to take a look at Scalyr's offering, a full-fledged log management solution that will help you not only collect and aggregate your logs but also perform fast searches and analyses on them. And all of that gets coupled with great visualization capabilities.

The knowledge inside your logs is like a hidden treasure, but don't just keep it hidden. Give Scalyr a try today.

---

**Like this article? Follow us on LinkedIn, Twitter, YouTube or Facebook to see the content we post.**

### Read more about Cyber Security

- Zipkin Tutorial: Get Started Easily With Distributed Tracing

©2022 SentinelOne. All Rights Reserved.