Analyze Jaeger trace data

Introduced 2.3

The Trace analytics functionality in the OpenSearch Observability plugin now supports Jaeger trace data. If you use OpenSearch as the backend for Jaeger trace data, you can use the Trace analytics built-in analysis capabilities. This provides support for OpenTelemetry (OTEL) formatted trace data.

When you perform trace analytics, you can select from two data sources:

- **Data Prepper** - Data ingested into OpenSearch through Data Prepper.
- **Jaeger** - Trace data stored within OpenSearch as its backend.

If you currently store your Jaeger trace data in OpenSearch, you can now use the capabilities built into Trace Analytics to analyze the error rates and latencies You can also filter the traces and look into the span details of a trace to pinpoint any service issues.

When you ingest Jaeger data into OpenSearch, it gets stored in a different index than the OTA-generated index that gets created when you run data through the Data Prepper. You can indicate which data source on which you want to perform trace analytics with the data source selector in the Dashboards.

Jaeger trace data that you can analyze includes span data, as well as service and operation endpoint data. Jaeger span data analysis requires some configuration.

Each time you ingest data for Jaeger, it creates a separate index for that day. The Dashboards will show the current index that has a mapping.

To learn more about Jaeger data tracing, see the Jaeger open source documentation.

Data Ingestion Requirements

If you want to see errors in your trace data, you need to set the following ElasticSearch flag to true prior to data ingestion: `--es.tags-as-fields.all=true`.

Jaeger data that is ingested for OpenSearch needs to have the flag set for errors. If data is not ingested in this format it will not work for trace analytics with OpenSearch.

About Data ingestion with Jaeger indexes

Trace analytics for non-Jaeger data use OTEL indexes with the naming conventions `otel-v1-apm-span-*` or `otel-v1-apm-service-map*`.

Jaeger indexes follow the naming conventions `jaeger-span-*` or `jaeger-service-*`.

need to confirm. this was provided in the spec and eng tickets. need more info to add it to the docs. Jaeger and OTEL indexes have different field names. Therefore when you run trace analytics, you'll need to create different queries and components depending on which index type you are using. (.note) -->

Step 1: Set up OpenSearch and OpenSearch Dashboards

You need to set up a local instance with Docker.

Run the following command to set up OpenSearch with Docker:

```
docker run --rm -it --name=opensearch -e "ES_JAVA_OPTS=-Xms2g -Xmx2g" -p 9200:9200 -p 9300:9300 -e "discovery.ty
```

Run the following command to set up OpenSearch Dashboards with Docker:

```
Need info here
```

Step 2: Set up Jaeger

To deploy Jaeger to run all the backend components in the same process, you need to get either the Jaeger binary or the Docker image from Download Jaeger.

The required flags are:

- `--link=opensearch` – Link to the OpenSearch container.
- `SPAN_STORAGE_TYPE=elasticsearch` – Defines the ElasticSearch storage type to store the Jaeger traces.
- `ES_TAGS_AS_FIELDS_ALL=true` – Sets OpenSearch mapping to ElasticSearch tags in Jaeger traces.

> Jaeger currently provides ElasticSearch as the storage type.

Run the following command to deploy Jaeger.

```
docker run --rm -it --link=opensearch --name=jaeger -e SPAN_STORAGE_TYPE=elasticsearch -e ES_SERVER_URLS=http:/
```

Upon success, you should be able to get to the Jaeger from http://localhost:16686.

Step 3: Simulate trace data

With Jaeger running, now you can create traces to verify they are stored with the OpenSearch instance.

To verify trace storage with OpenSearch, run the following command.

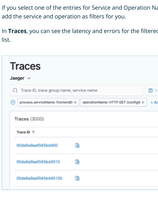
```
docker run --rm --link jaeger -e JAEGER_AGENT_HOST=jaeger -e JAEGER_AGENT_PORT=6831 -p 8000-8083:8000-8083 jaeg
```

Use trace analytics in OpenSearch Dashboards

To analyze your Jaeger trace data in the Dashboards, you need to set up Trace Analytics first. To get started, see Get started with Trace Analytics.

Data sources

You can specify either Data Prepper or Jaeger as your data source when you perform trace analytics. From the OpenSearch Dashboards, go to **Observability > Trace Analytics** and select Jaeger.

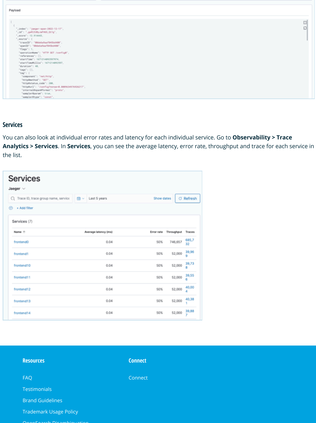


Dashboards views

After you select Jaeger for the data source, you can view all of your indexed data in **Dashboard** view including **Error rate** and **Throughput**.

Error rate

You can view the Trace error count over time in the Dashboard, and also view the top five combinations of services and operations that have a non-zero error rate.



Throughput

With **Throughput** selected, you can see the throughput of traces on Jaeger indexes that are coming in over time.

You can select an individual Trace from **Top 5 Service and Operation Latency** list and view the detailed trace data.

You can also see the combinations of services and operations that have the highest latency.

If you select one of the entries for Service and Operation Name and go to the **Traces** column to select a trace, it will add the service and operation as filters for you.

In **Traces**, you can see the latency and errors for the filtered service and operation for each individual Trace ID in the list.

If you select an individual Trace ID, you can see more detailed information about the trace, such as time spent by the service and each span for the service and operation. You can also view the payload that you get from the index in JSON format.

Services

You can also look at individual error rates and latency for each individual service. Go to **Observability > Trace Analytics > Services**. In **Services**, you can see the average latency, error rate, throughput and trace for each service in the list.