Data extraction with the /timeserieswindow endpoint

You want to extract past or streaming time series data that has been sent to Splunk Infrastructure Monitoring. You want to extract “raw” data (that is, the metrics and their values), as well as data that has been processed by Splunk analytics.

The /timeserieswindow endpoint in the Splunk API outputs raw metric data in JSON format. It is good to use to:

- Export raw data (no analytics applied), for a specific past time range, using a default rollup and resolution
- The advantages of /timeserieswindow are:
  - This method uses a very simple API, so it's easy to use.
  - It returns a standard JSON format

However, there are a number of requirements and limitations to be aware of. When using the /timeserieswindow endpoint for data extraction:

- You can't export streaming data.
- You can export data only to JSON format, which generally means you can make use of this API only by incorporating it into a script that either parses the JSON data or pumps it elsewhere.
- You can't specify a rollup type; the default rollup for the type of metric being exported will be used.
- You can’t specify a non-default resolution; only one of the resolutions at which Splunk retains data can be used: 1000 (1s), 60000 (1m), 300000 (5m), or 3600000 (1h).
- You can only export metrics with no filters, analytics, etc. (Note that what you can extract is not exactly the same as what was submitted, as rollups will be applied.)
- You must express the start and start time using milliseconds since epoch, which is more cumbersome than specifying a relative time range or using a different format, such as UTC. (Note that epoch is often denoted in seconds, so be sure to multiply by 1000 to get the time in milliseconds.)
- There is a maximum number of datapoints that you can get back in a single query. While the maximum is quite high (currently approximately 50 million datapoints), a query could theoretically apply to more than 50 million datapoints. In this situation, an error occurs and no data is returned.

Example usage

In the following example, curl is used to extract data for the metric “jvm.cpu.load” from 3/13/17 13:15 to 3/13/17 13:20, at the default resolution (1000ms).

```bash
curl
   --header "X-SF-TOKEN: YOUR_ACCESS_TOKEN"
   --header "Content-Type: application/json"
   --request GET
   'https://api.signalfx.com/v1/timeserieswindow?query=sf_metric:"jvm.cpu.load"&startMs=1489410900000&endMs=1489411200000'
```

The information provided in Splunk Lantern is intended for informational and educational purposes only. All information is provided in good faith, however, Splunk disclaims any and all representations and warranties, express and implied, regarding the information provided, including without limitation any warranties and representations regarding the completeness, adequacy or accuracy of the information. You agree to take full responsibility for the results arising from the use of the information provided.
In this next example, the same data is being extracted, but at 5-minute resolution. If the metric data is sent to Splunk Infrastructure Monitoring more frequently than once every 5 minutes, the returned data is rolled up using the default rollup for the type of metric (gauge, counter, or cumulative counter). In this case, the average of the values received during every 5-minute period is returned.

```
curl  
  --header "X-SF-TOKEN: YOUR_ACCESS_TOKEN" 
  --header "Content-Type: application/json" 
  --request GET 
  'https://api.signalfx.com/v1/timeserieswindow?query=sf_metric:"jvm.cpu.load"&startMs=1489410900000&endMs=1489411205000&resolution=300000'
```

**Troubleshoot the /timeserieswindow endpoint**

When extracting data using /timeserieswindow, there are situations where expected data isn’t being returned, even though the request is syntactically correct.

### No data is being returned

There are two cases in which your request might return no data; that is, returned data looks like this:

```
{"data":{},"errors" : []}
```

If you see this error, check to make sure there is actually data in SignalFx in the specified timeframe.

- If there is no data, choose a different time frame.
- If there is data, the likely reason for the error is that there are more than 5,000 metric time series in the timeframe. In this case, Splunk Infrastructure Monitoring has chosen a subset of the total data which happens to include only null values. For troubleshooting this issue, see “Only a subset of data is being returned,” below.

### Only a subset of data is being returned

If you are asking for a given metric and set of dimensions, /timeserieswindow will look for all of the time series that matches that query across all time “first”. If the number of matching time series includes more than 5,000 time series, it will return a subset of the total time series, with no regard to whether there is data in the timeframe you’re asking for.

One way to work around this issue is to add “sf_isActive:true” as another filter in your query. This will return only the time series that are currently active (have reported at least 1 data point within the last 36 hours). This may or may not be appropriate depending on the nature of your data and how you are sending it in.

If using this filter won’t work for your situation, you need to break up the query to ensure that your response doesn’t contain more than 5,000 results. For example, suppose you were asking for "sf_metric:content.global" AND "type:billable" and seeing only a subset of results. You could break the query into two queries:
Break up your original query as granularly as necessary to ensure that the results will match no more than 5,000 time series.

Next steps

You might be interested in other processes associated with the Extracting data from Splunk Infrastructure Monitoring use case.