



This PDF is generated from authoritative online content, and is provided for convenience only. This PDF cannot be used for legal purposes. For authoritative understanding of what is and is not supported, always use the online content. To copy code samples, always use the online content.

Genesys Info Mart Private Edition Guide

GSP metrics and alerts

7/2/2022

Find the metrics GSP exposes and the alerts defined for GSP.

Related documentation:

-

Contents

- [1 Metrics](#)
- [2 Alerts](#)

Service	CRD or annotations?	Port	Endpoint/Selector	Metrics update interval
GSP	PodMonitor	9249	Endpoint: / Selector: matchLabels: app: {{ template "gsp.fullname" . }} where the value of <code>gsp.fullname</code> depends on deployment parameters such as Helm release name, <code>.Values.fullnameOverride</code> , and <code>.Values.nameOverride</code> .	30 seconds

See details about:

- GSP metrics
- GSP alerts

Metrics

GSP exposes some standard Apache Flink and Kafka metrics as well as Genesys-defined metrics, which are exposed via the Flink API. Therefore, all GSP metrics start with the prefix **flink_** but in some cases the values are calculated by GSP.

You can query Prometheus directly to see all the metrics Flink and the Flink Kafka connector expose through GSP.

- For full information about the standard Flink metrics, see the Apache Flink documentation.
- For full information about the Kafka metrics, see the Apache Kafka or Confluent Kafka documentation.

The following metrics are likely to be particularly useful. The naming convention is `_`. Genesys does not commit to maintain other currently available GSP metrics not documented on this page.

Metric and description	Metric details	Indicator of
flink_taskmanager_job_task_operator_errors_numInvalidRecords Number of invalid input records.	Unit: InvalidRecords Type: Gauge Label: Sample value: 0	Error
flink_jobmanager_numRunningJobs Number of running Flink jobs. If less than 1, there is a problem.	Unit: Jobs Type: Gauge Label: Sample value: 1	Error
flink_taskmanager_job_task_operator_user_errors_numOversizedMessages		Messages

Metric and description	Metric details	Indicator of
Number of messages exceeding the max.request.size Kafka option.	Type: Gauge Label: <ul style="list-style-type: none"> operator_name Sample value: 0	
Number of issues encountered, such as errors or warnings.	Unit: Type: Gauge Label: flink_taskmanager_job_task_operator_tenant_error_total <ul style="list-style-type: none"> operator_name tenant error Sample value:	Error
The last watermark received by this operator/task, in milliseconds since the Unix Epoch (00:00:00 UTC on 1 January 1970). Note: For operators/tasks with two inputs, this is the earlier of the last received watermarks.	Unit: milliseconds Type: Gauge Label: flink_taskmanager_job_task_operator_currentInputWatermark <ul style="list-style-type: none"> operator_name Sample value:	Latency
The last watermark this operator has emitted, in milliseconds since the Unix Epoch.	Unit: milliseconds Type: Gauge Label: flink_taskmanager_job_task_operator_currentOutputWatermark <ul style="list-style-type: none"> operator_name: Sink:_Agent_State_Facts Sink:_Interaction_Facts Sample value:	Latency
The maximum lag in terms of the number of records for any partition in this window. An increasing value over time is your best indication that the consumer group is not keeping up with the producers.	Unit: Type: Gauge Label: flink_taskmanager_job_task_operator_records_lag_max Sample value:	Latency
The average number of records consumed per second.	Unit: Type: Gauge Label: flink_taskmanager_job_task_operator_records_consumed_rate Sample value:	Traffic
Total number of EventCallCreated events GSP received since it started processing.	Unit: Type: Gauge Label: flink_taskmanager_job_task_operator_numCallsCreated	Traffic

Metric and description	Metric details	Indicator of
	Sample value:	
flink_taskmanager_job_task_operator_numCallsCreatedPerSecond Number of EventCallCreated events per second (CPS).	Unit: Type: Gauge Label: Sample value:	Traffic
flink_taskmanager_job_task_operator_numThreadsCreated Total number of CallThreads GSP received since it started processing.	Unit: Type: Gauge Label: Sample value:	Traffic
flink_taskmanager_job_task_operator_numCallThreadsCreatedPerSecond Number of CallThreads per second (CTHPS).	Unit: Type: Gauge Label: Sample value:	Traffic
flink_taskmanager_job_task_operator_numChainsProcessed Total number of EventOCSCChainStartProcessing events GSP received since it started processing.	Unit: Type: Gauge Label: Sample value:	Traffic
flink_taskmanager_job_task_operator_numChainsProcessedPerSecond Number of EventOCSCChainStartProcessing events per second (CPS).	Unit: Type: Gauge Label: Sample value:	Traffic
flink_(job task)manager_Status_JVM_CPU_Load The recent CPU usage for the JVM process. The value is a double in the [0.0,1.0] interval, where a value of 0.0 means that none of the CPUs were running threads from the JVM process, while a value of 1.0 means that all CPUs were actively running threads from the JVM 100% of the time during the recent period being observed. A negative value means usage data is not available. For more information, see https://docs.oracle.com/javase/7/docs/jre/api/management/extension/com/sun/management/OperatingSystemMXBean.html#getProcessCpuLoad() .	Unit: Type: Gauge Label: • pod Sample value:	Saturation
flink_(job task)manager_Status_JVM_Memory_Direct_TotalCapacity The total capacity of all buffers in the direct buffer pool.	Unit: bytes Type: Gauge Label: • pod Sample value:	Saturation
flink_(job task)manager_Status_JVM_Memory_Direct_MemoryUsed The amount of memory used by the JVM for the direct buffer pool.	Type: Gauge Label:	Saturation

Metric and description	Metric details	Indicator of
	<ul style="list-style-type: none"> pod Sample value:	
flink_(job task)manager_Status JVM Memory NonHeap_Max The maximum amount of non-heap memory that can be used for memory management.	Unit: bytes Type: Gauge Label: <ul style="list-style-type: none"> pod Sample value:	Saturation
flink_(job task)manager_Status JVM Memory NonHeap_Used The amount of non-heap memory currently used.	Unit: bytes Type: Gauge Label: <ul style="list-style-type: none"> pod Sample value:	Saturation
flink_(job task)manager_Status JVM Memory Heap_Max The maximum amount of heap memory that can be used for memory management.	Unit: bytes Type: Gauge Label: <ul style="list-style-type: none"> pod Sample value:	Saturation
flink_(job task)manager_Status JVM Memory Heap_Used The amount of heap memory currently used.	Unit: bytes Type: Gauge Label: <ul style="list-style-type: none"> pod Sample value:	Saturation

Alerts

The alerts are based on Flink and Kubernetes cluster metrics.

The following alerts are defined for GSP.

Alert	Severity	Description	Based on	Threshold
GspFlinkJobDown	Critical	Triggered when the GSP Flink job is not running (number of running jobs)	flink_jobmanager_numRunningJobs	For 5 minutes

Alert	Severity	Description	Based on	Threshold
		equals to 0 or metric is not available)		
GspOOMKilled	Critical	Triggered when a GSP pod is restarted because of OOMKilled	kube_pod_container_status_restarts_total	0
GspNoTmRegistered	Critical	Triggered when there are no registered TaskManagers (or metric not available)	flink_jobmanager_numRegisteredTaskManagers	For 5 minutes
GspUnknownPerson	High	Triggered when GSP encounters unknown person(s)	flink_taskmanager_job_to_skip_per_tenant_error_total	For 5 minutes