Genesys Callback Private Edition Guide

No results metrics and alerts

8/4/2025

Contents

Find the metrics *No results* exposes and the alerts defined for *No results*.

Related documentation:

*

Service	CRD or annotations?	Port	Endpoint/Selector	Metrics update interval
No results	Supports both CRD (Service Monitor) and annotations	3050	/metrics	Real-time updates

See details about:

* *No results* metrics
* *No results* alerts

Metrics

GES exposes some default metrics such as CPU usage, memory usage, and the state of the Node.js runtime, as well as metrics coming directly from the GES API such as the number of created callbacks, call-in requests, and so on. These basic metrics are created as counters, which means that the values will monotonically increase over time from the beginning of a GES pod's lifespan. For more information about counters, see Metric Types in the Prometheus documentation.

You might see metrics documented on this page that you cannot find on the endpoint or – if they exist – they might have no value. These are alert-type metrics. This type of metric is set when the condition it tracks is first encountered. For example, if GES has never experienced a DNS failure since it started, then no GES_DNS_FAILURE alert has ever been generated and the GES_DNS_FAILURE metric would not yet exist. For more information, see Alerting.

You might see metrics with almost identical names, except for case (upper or lower). Metrics with names ending in _tolerance are simply thresholds and exist at the level at which an alert is triggered; they are not the same as the metric used for monitoring. For more information, see Alerting.

You can query Prometheus directly to see all the metrics that GES exposes. The following metrics are likely to be particularly useful. Genesys does not commit to maintain other currently available GES metrics not documented on this page.

Metric and description	Metric details	Indicator of
ges_callbacks_created	**Unit:** N/A	The number of callbacks booked in GES

Metric and description	Metric details	Indicator of
The number of callbacks booked in GES since the deployment went online.	**Type:** counter **Label:** tenant – The tenant for which the callback was booked. **Sample value:**	
ges_monitor_size The number of booked callbacks currently being monitored and managed in GES. This is a background task that both ensures that new callbacks are propagated to Redis and that callbacks are dispatched to ORS when appropriate. If this metric is consistently high, it might indicate issues with the GES deployment.	**Unit:** callback **Type:** gauge **Label:** type – The type of callback monitor. **Sample value:** 3	Latency related to starting scheduled callbacks
ges_push_notifications_sent The number of Push Notifications sent since the deployment went online. This tracks notifications that were both successfully and unsuccessfully dispatched.	**Unit:** N/A **Type:** counter **Label:** tenant – The tenant for which the Push Notification request was created. channel – The channel through which the Push Notification is delivered. Currently, this can be only Google FCM ("FCM"). result – Either "success" or "failure" based on whether the notification was successfully dispatched or not. **Sample value:** 3	How many Push Notifications that GES has dispatched
ges_http_requests_total The number of HTTP requests handled by GES since the deployment went online. This metric does not delineate between successful and unsuccessful requests.	**Unit:** N/A **Type:** counter **Label:** tenant – The tenant associated with the request. If no tenant can be identified, this defaults to "Unknown Tenant". path – The path of the request. If a private endpoint, then it is "Private API Endpoint". **Sample value:**	Overall GES activity and usage
ges_callin_created The total number of Click-to-Call-In requests handled since the GES deployment went online.	**Unit:** N/A **Type:** counter **Label:** tenant – The tenant for which the Click-to-Call-In request was booked. **Sample value:**	The number of Click-to-Call-In requests GES has received
ges_http_failed_requests_total The amount of failed (4XX/5XX) requests handled by GES since the deployment came online.	**Unit:** N/A **Type:** counter **Label:** tenant – The tenant associated with the request. If no tenant can be identified, this defaults to "Unknown Tenant". path – The path of the request. httpCode – The HTTP code associated with the result. **Sample value:**	Dependent on which HTTP codes you observe. Excessive 500 codes might indicate an issue with configuration or with GES itself. Excessive 400 errors might indicate malicious behavior.
ges_build_info Displays the version of GES that is	**Unit:** N/A **Type:** gauge	Software version

Metric and description	Metric details	Indicator of
currently running. In the case of this metric, the labels provide the important information. The metric value is always 1 and does not provide any information.	**Label:** version – The version of GES that you are running in your deployment. **Sample value:** ges_build_info{version="100.0.000.0000.build.69.rev.d07b89146"} 1	
GES_HEALTH The overall health of the GES deployment; this is a composite of the connection statuses of GES and downstream services. Values are: 1 – healthy 0 – unhealthy If a value is not exported, assume that GES is healthy (unless the /metrics endpoint can't be reached).	**Unit:** N/A **Type:** gauge **Label:** version – The GES version that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:** 1	The overall health of the GES deployment and connections
GWS_CONFIG_STATUS The status of the connection to the GWS Configuration Service. Values are: 1 – healthy 0 – unhealthy	**Unit:** N/A **Type:** gauge **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:** 1	Health of the connection to the GWS Configuration Service
GWS_ENV_STATUS The status of the connection to the GWS Environment Service. Values are: 1 – healthy 0 – unhealthy	**Unit:** N/A **Type:** gauge **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:** 1	Health of the connection to the GWS Environment Service
GWS_AUTH_STATUS The status of the connection to the Genesys Authentication Service. Values are: 1 – healthy 0 – unhealthy	**Unit:** N/A **Type:** gauge **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:** 1	Health of the connection to the GWS Authentication Service
ALL_URS_DOWN A flag that raises when connections to both the primary and secondary URS components are unhealthy. Values are: 1 – healthy 0 – unhealthy If the metric is not being exposed, assume that the value is 0 and that URS connections are in an unhealthy state.	**Unit:** N/A **Type:** gauge **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:** 1	Health of the connection from GES to URS
REDIS_CONNECTION	**Unit:** N/A	Health of the connection to Redis

Metric and description	Metric details	Indicator of
Monitors the health of the connection between GES and its own Redis instance. Values are: 1 – healthy 0 – unhealthy Because GES is so dependent on Redis, you might have trouble confirming – with metrics – when Redis is actually down (GES might not respond to the /metrics query).	**Type:** gauge **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:** 1	
ORS_REDIS_STATUS Monitors the health of the connection between GES and the ORS Redis instance. Values are: 1 – healthy 0 – unhealthy	**Unit:** N/A **Type:** gauge **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:** 1	Health of the connection to ORS Redis.
RBAC_CREATE_VQ_PROXY_ERROR The number of times GES has encountered issues when managing virtual queue proxy objects. When a callback service (also called a virtual queue, or VQ) is added to GES using the CALLBACK_SETTINGS data table in Designer, GES automatically creates a script object for line-of-business segmentation (see Line of Business segmentation). When the callback service (VQ) is removed from the CALLBACK_SETTINGS data table, GES automatically deletes the script object.	**Unit:** N/A **Type:** counter **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:**	The ability of GES to create or delete the script objects.
LOGGING_FAILURE The number of times GES has encountered issues writing logs to standard output (stdout).	**Unit:** N/A **Type:** counter **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:**	Typically indicates some sort of issue with the Kubernetes pod or the host
UNCAUGHT_EXCEPTION The number of times GES has encountered an uncaught exception while running.	**Unit:** N/A **Type:** counter **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:**	There is no specific problem that this metric indicates. Check the logs for more information.
GES_DNS_FAILURE The number of times GES has encountered a failure in performing DNS resolution.	**Unit:** N/A **Type:** counter **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment.	Certain configuration values such as the location of GWS, Redis, Postgres, or ORS might be incorrect

Metric and description	Metric details	Indicator of
	Sample value:	
GWS_INCORRECT_CLIENT_CREDENTIALS The number of times that authentication on GWS has failed because the client credentials that were supplied were incorrect.	**Unit:** N/A **Type:** counter **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:**	Incorrect client credentials are being supplied to GWS. Check that correct credentials have been made available in the secret.
NEXUS_ACCESS_FAILURE The number of times the GES deployment has failed to contact Nexus. This is only relevant if you use the Push Notification feature.	**Unit:** N/A **Type:** counter **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:**	Indicates issues with the Nexus deployment or the connection from GES to Nexus.
CB_SUBMIT_FAILED The number of times that GES has failed to submit a callback to ORS.	**Unit:** N/A **Type:** counter **Label:** version – The version of GES that you are running in your deployment. host – The hostname associated with the GES deployment. **Sample value:**	There might be issues with the ORS deployment. GES could also be supplying an incorrect URL to ORS. Change the GES_URL environment variable to fix the latter issue.

Alerts

The following alerts are defined for *No results*.

Alert	Severity	Description	Based on	Threshold
GES_UP	Critical	Fires when fewer than two GES pods have been up for the last 15 minutes.		Triggered when fewer than two GES pods are up for 15 consecutive minutes.
GES_CPU_USAGE	Info	GES has high CPU usage for 1 minute.	ges_process_cpu_seconds_total	Triggered when the average CPU usage (measured by ges_process_cpu_seconds_total) is greater than 90% for 1 minute.
GES_MEMORY_USAGE	Info	GES has high memory usage for	ges_nodejs_heap_space_size_used_bytes, ges_nodejs_heap_space_size_available_bytes	Triggered when, memory usage

Alert	Severity	Description	Based on	Threshold
		a period of 90 seconds.		(measured as a ratio of Used Heap Space vs Available Heap Space) is above 80% for a 90-second interval.
GES-NODE-JS-DELAY-WARNING	Warning	Triggers if the base NodeJS event loop becomes excessive. This indicates significant resource and performance issues with the deployment.	application_ccecp_nodejs_eventloop_lag_seconds	Triggered when the event loop is greater than 5 milliseconds for a period exceeding 5 minutes.
GES_NOT_READY_CRITICAL	Critical	GES pods are not in the Ready state. Indicative of issues with the Redis connection or other problems with the Helm deployment.	kube_pod_container_status_ready	Triggered when more than 50% of GES pods have not been in a Ready state for 5 minutes.
GES_NOT_READY_WARNING	Warning	GES pods are not in the Ready state. Indicative of issues with the Redis connection or other problems with the Helm deployment.	kube_pod_container_status_ready	Triggered when 25% (or more) of GES pods have not been in a Ready state for 10 minutes.
GES_PODS_RESTART	Critical	GES pods have been excessively crashing and restarting.	kube_pod_container_status_restarts_total	Triggered when there have been more than five pod restarts in the past 15 minutes.
GES_HEALTH	Critical	One or more downstream components (PostGres, Config Server, GWS, ORS) are down. **Note:** Because GES goes into a crash loop when Redis is down, this does not fire when	GES_HEALTH	Triggered when any component is down for any length of time.

Alert	Severity	Description	Based on	Threshold
		Redis is down.		
GES_ORS_REDIS_DOWN	Critical	Connection to ORS_REDIS is down.	ORS_REDIS_STATUS	Triggered when the ORS_REDIS connection is down for 5 consecutive minutes.
GES_GWS_AUTH_DOWN	Warning	Connection to the Genesys Authentication Service is down.	GWS_AUTH_STATUS	Triggered when the connection to the Genesys Authentication Service is down for 5 minutes.
GES_GWS_ENVIRONMENT_DOWN	Warning	Connection to the GWS Environment Service is down.	GWS_ENV_STATUS	Triggered when the connection to the GWS Environment Service is down.
GES_GWS_CONFIG_DOWN	Warning	Connection to the GWS Configuration Service is down.	GWS_CONFIG_STATUS	Triggered when the connection to the GWS Configuration Service is down.
GES_GWS_SERVER_ERROR	Warning	GES has encountered server or connection errors with GWS.	GWS_SERVER_ERROR	Triggered when there has been a GWS server error in the past 5 minutes.
GES_HTTP_400_POD	Info	An individual GES pod is returning excessive HTTP 400 results.	ges_http_failed_requests_total, http_400_tolerance	Triggered when two or more HTTP 400 results are returned from a pod within a 5-minute period.
GES_HTTP_404_POD	Info	An individual GES pod is returning excessive HTTP 404 results.	ges_http_failed_requests_total, http_404_tolerance	Triggered when two or more HTTP 404 results are returned from a pod within a 5-minute period.

Alert	Severity	Description	Based on	Threshold
GES_HTTP_500_POD	Info	An individual GES pod is returning excessive HTTP 500 results.	ges_http_failed_requests_total, http_500_tolerance	Triggered when two or more HTTP 500 results are returned from a pod within a 5-minute period.
GES_HTTP_401_POD	Info	An individual GES pod is returning excessive HTTP 401 results.	ges_http_failed_requests_total, http_401_tolerance	Triggered when two or more HTTP 401 results are returned from a pod within a 5-minute period.
GES_SLOW_HTTP_RESPONSE_TIME	Warning	Fired if the average response time for incoming requests begins to lag.	ges_http_request_duration_seconds_sum, ges_http_request_duration_seconds_count	Triggered when the average response time for incoming requests is above 1.5 seconds for a sustained period of 15 minutes.
GES_RBAC_CREATE_VQ_PROXY_ERROR	Info	Fires if there are issues with GES managing VQ Proxy Objects.	RBAC_CREATE_VQ_PROXY_ERROR, rbac_create_vq_proxy_error_tolerance	Triggered when there are at least 1000 instances of issues managing VQ Proxy objects within a 10-minute period.
GES_LOGGING_FAILURE	Warning	GES has failed to write a message to the log.	LOGGING_FAILURE	Triggered when there are any failures writing to the logs. Silenced after 1 minute.
GES_UNCAUGHT_EXCEPTION	Warning	There has been an uncaught exception within GES.	UNCAUGHT_EXCEPTION	Triggered when GES encounters any uncaught exceptions. Silenced after 1 minute.

Alert	Severity	Description	Based on	Threshold
GES_INVALID_CONTENT_LENGTH	Info	Fires if GES encounters any incoming requests that have exceeded the maximum content length of 10mb on the internal port and 500kb for the external, public-facing port.	INVALID_CONTENT_LENGTH, invalid_content_length_tolerance	Triggered when one instance of a message with an invalid length is received. Silenced after 2 minutes.
GES_DNS_FAILURE	Warning	A GES pod has encountered difficulty resolving DNS requests.	DNS_FAILURE	Triggered when GES encounters any DNS failures within the last 30 minutes.
GES_CB_TTL_LIMIT_REACHED	Info	GES is throttling callbacks for a specific tenant.	CB_TTL_LIMIT_REACHED	Triggered when GES has started throttling callbacks within the past 2 minutes.
GES_CB_ENQUEUE_LIMIT_REACHED	Info	GES is throttling callbacks to a given phone number.	CB_ENQUEUE_LIMIT_REACHED	Triggered when GES has begun throttling callbacks to a given number within the past 2 minutes.
GES_CB_SUBMIT_FAILED	Info	GES has failed to submit a callback to ORS.	CB_SUBMIT_FAILED	Triggered when GES has failed to submit a callback to ORS in the past 2 minutes for any reason.
GES_GWS_INCORRECT_CLIENT_CREDENTIALS	Warning	The GWS client credentials provided to GES are incorrect.	GWS_INCORRECT_CLIENT_CREDENTIALS	Triggered when GWS has had any issue with the GES client credentials in the last 5 minutes.

Alert	Severity	Description	Based on	Threshold
GES_NEXUS_ACCESS_FAILURE WARNING		GES has been having difficulties contacting Nexus. This alert is only relevant for customers who leverage the Push Notification feature in Genesys Callback.	NEXUS_ACCESS_FAILURE	Triggered when GES has failed to connect or communicate with Nexus more than 30 times over the last hour.