



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.

Voice Microservices Private Edition Guide

[Voice RQ Service metrics and alerts](#)

Contents

- 1 Metrics
- 2 Alerts

Find the metrics Voice RQ Service exposes and the alerts defined for Voice RQ Service.

Service	CRD or annotations?	Port	Endpoint/Selector	Metrics update interval
Voice RQ Service	Supports both CRD and annotations	12000	http://:12000/metrics	30 seconds

See details about:

- Voice RQ Service metrics
- Voice RQ Service alerts

Metrics

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

Metric and description	Metric details	Indicator of
rqnode_clients Number of clients connected.	Unit: N/A Type: gauge Label: Sample value:	Traffic
rqnode_streams Number of active streams present.	Unit: N/A Type: gauge Label: Sample value:	Traffic
rqnode_xreads Number of XREAD requests received.	Unit: N/A Type: counter Label: Sample value:	Traffic
rqnode_xadds Number of XADD requests received.	Unit: N/A Type: counter Label: Sample value:	Traffic
rqnode_redis_state Current Redis connection state.	Unit: N/A Type: gauge	Errors

Metric and description	Metric details	Indicator of
	Label: Sample value:	
rqnode_redis_disconnects The number of Redis disconnects that occurred for the RQ node.	Unit: Type: counter Label: Sample value:	Errors
rqnode_consul_leader_error Number of errors received from Consul during the leadership process.	Unit: N/A Type: counter Label: Sample value:	Errors
rqnode_active_master Service master role is active.	Unit: N/A Type: gauge Label: Sample value:	Saturation
rqnode_active_backup Service backup role is active.	Unit: N/A Type: gauge Label: Sample value:	Saturation
rqnode_read_latency RQ latency; that is, the time duration between when an event is added to Redis and when it's read via XREAD.	Unit: Type: histogram Label: le, healthcheck Sample value:	Latency
rqnode_add_latency RQ latency; that is, the time duration between when a message is received and when it's added to the list.	Unit: Type: histogram Label: le, healthcheck Sample value:	Latency
rqnode_redis_latency Latency caused by Redis read/write.	Unit: Type: histogram Label: le Sample value:	Latency

Alerts

The following alerts are defined for Voice RQ Service.

Alert	Severity	Description	Based on	Threshold
Number of Redis streams is too high	Warning	Too many active sessions. Actions:	rqnode_streams	More than 10000 active streams running.

Alert	Severity	Description	Based on	Threshold
		<ul style="list-style-type: none"> Check whether the horizontal pod autoscaler has triggered and if the maximum number of pods has reached. Check the number of voice, digital, and callback calls in the system. 		
Redis disconnected for 5 minutes	Warning	<p>Redis is not available for the pod {{ \$labels.pod }}.</p> <p>Actions:</p> <ul style="list-style-type: none"> If the alarm is triggered for multiple services, make sure there are no issues with Redis, restart Redis. If the alarm is triggered only for the pod {{ \$labels.pod }}, check to see if there is any issue with the pod. 	redis_state	Redis is not available for the pod {{ \$labels.pod }} for 5 minutes.
Redis disconnected for 10 minutes	Critical	<p>Redis is not available for the pod {{ \$labels.pod }}.</p> <p>Actions:</p> <ul style="list-style-type: none"> If the alarm is triggered for multiple services, make sure there are no issues with Redis. 	redis_state	Redis is not available for the pod {{ \$labels.pod }} for 10 minutes.

Alert	Severity	Description	Based on	Threshold
		<p>Redis, and then restart Redis.</p> <ul style="list-style-type: none"> If the alarm is triggered only for the pod <code>{}{{\$labels.pod}}</code>, check to see if there is any issue with the pod. 		
Pod failed	Warning	<p>Pod <code>{}{{\$labels.pod}}</code> failed.</p> <p>Actions:</p> <ul style="list-style-type: none"> One of the containers in the pod has entered a Failed state. Check the Kibana logs for the reason. 	<code>kube_pod_status_phase</code>	Pod <code>{}{{\$labels.pod}}</code> is in Failed state.
Pod Unknown state	Warning	<p>Pod <code>{}{{\$labels.pod}}</code> in Unknown state.</p> <p>Actions:</p> <ul style="list-style-type: none"> If the alarm is triggered for multiple services, make sure there are no issues with the Kubernetes cluster. If the alarm is triggered only for the pod <code>{}{{\$labels.pod}}</code>, check whether the image is correct and if the container is starting up. 	<code>kube_pod_status_phase</code>	Pod <code>{}{{\$labels.pod}}</code> in Unknown state for 5 minutes.
Pod Pending state	Warning	Pod <code>{}{{\$labels.pod}}</code> is in the Pending state.	<code>kube_pod_status_phase</code>	Pod <code>{}{{\$labels.pod}}</code> is in the Pending state for 5 minutes.

Alert	Severity	Description	Based on	Threshold
		<p>Actions:</p> <ul style="list-style-type: none"> • If the alarm is triggered for multiple services, make sure the Kubernetes nodes where the pod is running are alive in the cluster. • If the alarm is triggered only for the pod {{ \$labels.pod }}, check the health of the pod. 		
Pod not ready for 10 minutes	Critical	<p>Pod {{ \$labels.pod }} in NotReady state.</p> <p>Actions:</p> <ul style="list-style-type: none"> • If this alarm is triggered, check whether the CPU is available for the pods. • Check whether the port of the pod is running and serving the request. 	kube_pod_status_ready	Pod {{ \$labels.pod }} in NotReady state for 10 minutes.
Container restored repeatedly	Critical	<p>Container {{ \$labels.container }} was repeatedly restarted.</p> <p>Actions:</p> <ul style="list-style-type: none"> • One of the containers in the pod has entered a failed state. Check the Kibana logs for 	kube_pod_container_status_restarts_total	Container {{ \$labels.container }} was restarted 5 or more times within 15 minutes.

Alert	Severity	Description	Based on	Threshold
		the reason.		
Pod memory greater than 65%	Warning	<p>High memory usage for pod {{ \$labels.pod }}.</p> <p>Actions:</p> <ul style="list-style-type: none"> Check whether the horizontal pod autoscaler has triggered and if the maximum number of pods has been reached. Check Grafana for abnormal load. Collect the service logs; raise an investigation ticket. 		Container {{ \$labels.container }}.memory_usage_exceeded_65%_for_5_minutes container_memory_working_set_bytes kube_pod_container_resource_requests_memory_bytes
Pod memory greater than 80%	Critical	<p>Critical memory usage for pod {{ \$labels.pod }}.</p> <p>Actions:</p> <ul style="list-style-type: none"> Check whether the horizontal pod autoscaler has triggered and if the maximum number of pods has been reached. Check Grafana for abnormal load. Restart the service. Collect the service logs; raise an investigation ticket. 		Container {{ \$labels.container }}.memory_usage_exceeded_80%_for_5_minutes container_memory_working_set_bytes kube_pod_container_resource_requests_memory_bytes

Alert	Severity	Description	Based on	Threshold
Pod CPU greater than 65%	Warning	<p>High CPU load for pod {{ \$labels.pod }}.</p> <p>Actions:</p> <ul style="list-style-type: none"> Check whether the horizontal pod autoscaler has triggered and if the maximum number of pods has been reached. Check Grafana for abnormal load. Collect the service logs; raise an investigation ticket 		Container {{ \$labels.container }} CPU usage exceeded 65% for 5 minutes.
Pod CPU greater than 80%	Critical	<p>Critical CPU load for pod {{ \$labels.pod }}.</p> <p>Actions:</p> <ul style="list-style-type: none"> Check whether the horizontal pod autoscaler has triggered and if the maximum number of pods has been reached. Check Grafana for abnormal load. Restart the service. Collect the service logs; raise an investigation ticket. 		Container {{ \$labels.container }} CPU usage exceeded 80% for 5 minutes.