Voice Microservices Private Edition Guide

Voice SIP Proxy Service metrics and alerts

4/5/2026

Contents

Find the metrics Voice SIP Proxy Service exposes and the alerts defined for Voice SIP Proxy Service.

Service	CRD or annotations?	Port	Endpoint/Selector	Metrics update interval
Voice SIP Proxy Service	Supports both CRD and annotations	11400	http://:11400/metrics	30 seconds

See details about:

- Voice SIP Proxy Service metrics
- Voice SIP Proxy Service alerts

Metrics

Voice SIP Proxy Service exposes Genesys-defined, SIP Proxy Service–specific metrics as well as some standard Kafka metrics. You can query Prometheus directly to see all the metrics that the SIP Proxy Service exposes. The following metrics are likely to be particularly useful. Genesys does not commit to maintain other currently available SIP Proxy Service metrics not documented on this page.

Metric and description	Metric details	Indicator of
sipproxy_requests_total Total number of received requests.	**Unit:** N/A **Type:** counter **Label:** method **Sample value:**	Traffic
sipproxy_rejected_requests_total The total number of rejected requests.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	Errors
sipproxy_requests_processed_self_total The total number of received requests that were processed by SIP Proxy itself.	**Unit:** N/A **Type:** counter **Label:** method **Sample value:**	Traffic
sipproxy_requests_forwarded_total The total number of forwarded requests.	**Unit:** N/A **Type:** counter **Label:** method, request_direction, sip_node_id **Sample value:**	Traffic
sipproxy_requests_sip_node_resolved_total	**Unit:** N/A	Errors

Metric and description	Metric details	Indicator of
Total count of sip-node reselection.	**Type:** counter **Label:** **Sample value:**	
sipproxy_responses_forwarded_total Total count of forwarded responses.	**Unit:** N/A **Type:** counter **Label:** method, sip_node_id, request_direction **Sample value:**	Traffic
sipproxy_response_latency SIP response latency.	**Unit:** **Type:** histogram **Label:** le, sip_node_id, request_direction, target, node_in_cache **Sample value:**	Latency
sipproxy_register_processed_total Total number of REGISTER requests that SIP Proxy received for processing.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	Traffic
sipproxy_register_rejected_total Total number of REGISTER requests for processing that were rejected.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	Errors
sipproxy_calls_per_second_count Current calculated calls per second.	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	Saturation
sipproxy_active_sip_nodes_count Current number of active SIP nodes.	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	
sipproxy_sip_nodes_count Current number of discovered SIP nodes.	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	
sipproxy_tenants_count Current count of discovered tenants.	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	
sipproxy_consul_record_processing_errors_count Current number of errors while processing records got from Consul.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	
sipproxy_consul_errors_count	**Unit:** N/A	

Metric and description	Metric details	Indicator of
Current number of Consul errors.	**Type:** counter **Label:** **Sample value:**	
sipproxy_sip_node_is_capacity_available Indicates whether SIP node has available capacity or not.	**Unit:** **Type:** gauge **Label:** sip_node_id **Sample value:**	
service_version_info Displays the version of Voice SIP Proxy Service that is 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.	**Unit:** N/A **Type:** gauge **Label:** version **Sample value:** service_version_info{version="100.0.1000006"} 1	
sipproxy_health_level Health level of the SIP Proxy node: -1 – fail 0 – starting 1 – degraded 2 – pass	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	
sipproxy_envoy_proxy_status Status of the Envoy proxy: -1 – error 0 – disconnected 1 – connected	**Unit:** N/A **Type:** gauge **Label:** **Sample value:** 1	
sipproxy_config_node_status Status of the Config node connection: 0 – disconnected 1 – connected	**Unit:** N/A **Type:** gauge **Label:** **Sample value:** 1	
sip_server_transactions_created_total Total number of created server transactions.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	Traffic
sip_client_transactions_created_total Total number of created client transactions.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	Traffic
sip_server_transactions_deleted_total Total number of deleted server transactions.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	Traffic
sip_client_transactions_deleted_total	**Unit:** N/A	Traffic

Metric and description	Metric details	Indicator of
Total number of deleted client transactions.	**Type:** counter **Label:** **Sample value:**	
sip_client_transactions_count Current number of client transactions.	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	Saturation
sip_server_transactions_count Current number of server transactions.	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	Saturation
sip_server_transactions_rejected_total Total number of server transactions rejected for internal reasons.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	Errors
sip_proxy_contexts_count Current number of active SIP Proxy forwarding contexts.	**Unit:** N/A **Type:** gauge **Label:** **Sample value:**	Saturation
sip_received_bytes_total Total traffic received, measured in bytes.	**Unit:** bytes **Type:** counter **Label:** transport **Sample value:**	Traffic
sip_sent_bytes_total Total traffic sent, measured in bytes.	**Unit:** bytes **Type:** counter **Label:** transport **Sample value:**	Traffic
sip_transport_errors_total Total number of transport errors.	**Unit:** N/A **Type:** counter **Label:** transport, address **Sample value:**	Errors
sip_stream_transport_wait_drain_total Total number of requests to wait for drain events on stream transports.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	
sip_stream_transport_flood_total Total number of flood events on the stream transports.	**Unit:** N/A **Type:** counter **Label:** **Sample value:**	
http_client_request_duration_seconds	**Unit:** seconds	Latency

Metric and description	Metric details	Indicator of
The time duration between the HTTP client request and the response, measured in seconds.	**Type:** histogram **Label:** le, target_service_name **Sample value:**	
http_client_response_count The number of HTTP client responses received.	**Unit:** N/A **Type:** counter **Label:** target_service_name, status **Sample value:**	Traffic

log_output_bytes_total
The total amount of log output, measured in bytes.
	Unit: bytes
	Type: counter
	Label: level, format, module
	Sample value:
	`log_output_bytes_total{level="info",format="txt",module="sipproxy_node@config-manager"} 3175`
	`log_output_bytes_total{level="info",format="txt",module="sipproxy_node@sipproxy-node"} 96`
	`log_output_bytes_total{level="info",format="txt",module="sipproxy_node@sipproxy@sip"} 181`
	`log_output_bytes_total{level="info",format="json",module="sipproxy_node@config-manager"} 4184`
	`log_output_bytes_total{level="info",format="json",module="sipproxy_node@sipproxy-node"} 135`
	`log_output_bytes_total{level="info",format="json",module="sipproxy_node@sipproxy@sip"} 259`	

kafka_consumer_recv_messages_total Number of messages received from Kafka.	**Unit:** **Type:** counter **Label:** **Sample value:**	Traffic
kafka_consumer_error_total Number of Kafka consumer errors.	**Unit:** **Type:** counter **Label:** **Sample value:**	Errors
kafka_consumer_latency Consumer latency is the time difference between when the message is produced and when the message is consumed. That is, the time when the consumer received the message minus the time when the producer produced the message.	**Unit:** **Type:** histogram **Label:** **Sample value:**	Latency
kafka_consumer_rebalance_total Number of Kafka consumer rebalance events.	**Unit:** **Type:** counter **Label:** **Sample value:**	
kafka_consumer_state Current state of the Kafka consumer.	**Unit:** **Type:** gauge **Label:** **Sample value:**	
kafka_producer__messages_total Number of messages received from	**Unit:** **Type:** counter	Traffic

Metric and description	Metric details	Indicator of
Kafka.	**Label:** **Sample value:**	
kafka_producer_queue_depth Number of Kafka producer pending events.	**Unit:** **Type:** gauge **Label:** kafka_location **Sample value:**	Saturation
kafka_producer_queue_age_seconds Age of the oldest producer pending event in seconds.	**Unit:** seconds **Type:** gauge **Label:** kafka_location **Sample value:**	
kafka_producer_error_total Number of Kafka producer errors.	**Unit:** **Type:** counter **Label:** kafka_location **Sample value:**	Errors
kafka_producer_state Current state of the Kafka producer.	**Unit:** **Type:** gauge **Label:** kafka_location **Sample value:**	
kafka_producer_biggest_event_size Biggest event size so far.	**Unit:** **Type:** gauge **Label:** kafka_location, topic **Sample value:** 231	
kafka_max_request_size Exposed config to compare with biggest event size.	**Unit:** **Type:** gauge **Label:** kafka_location **Sample value:** 1000000	
kafka_producer_dropped_event_number Number of dropped events.	**Unit:** **Type:** gauge **Label:** **Sample value:**	

Alerts

The following alerts are defined for Voice SIP Proxy Service.

Alert	Severity	Description	Based on	Threshold
Too many Kafka pending events	Critical	Too many Kafka producer pending events for pod {{ $labels.pod }}.	kafka_producer_queue_depth	Too many Kafka producer pending events for service {{

Alert	Severity	Description	Based on	Threshold
		This alert means there are issues with SIP REGISTER processing on this voice-sipproxy. Actions: • Make sure there are no issues with Kafka or with the {{ $labels.pod }} pod's CPU and network.		$labels.container }} (more than 100 in 5 minutes).
SIP server response time too high	Warning	Actions: • If the alarm is triggered for multiple sipproxy-nodes, make sure there are no issues on {{ $labels.sip_node_id }}. • If the alarm is triggered only for sipproxy-node {{ $labels.pod }}, check to see if there is an issue with the service related to the topic (CPU, memory, or network overload).	sipproxy_response_latency_bucket	SIP response latency for more than 95% of messages forwarded to {{ $labels.sip_node_id }} is more than 1 second for sipproxy-node {{ $labels.pod }}.
Pod status failed	Warning	Actions: • Restart the pod and check to see if there are any issues with the pod after restart.	kube_pod_status_phase	Pod {{ $labels.pod }} is in Failed state.
Pod status Unknown	Warning	Pod {{ $labels.pod }} is in Unknown state.	kube_pod_status_phase	Pod {{ $labels.pod }} is in Unknown state for 5 minutes.

Alert	Severity	Description	Based on	Threshold
		Actions: • Restart the pod and check to see if there are any issues with the pod after restart.		
Pod status Pending	Warning	Pod {{ $labels.pod }} is in Pending state. Actions: • Restart the pod and check to see if there are any issues with the pod after restart.	kube_pod_status_phase	Pod {{ $labels.pod }} is in Pending state for 5 minutes.
Pod status NotReady	Critical	Pod {{ $labels.pod }} is in NotReady state. Actions: • Restart the pod and check to see if there are any issues with the pod after restart.	kube_pod_status_ready	Pod {{ $labels.pod }} is in NotReady state for 5 minutes.
Container restarted repeatedly	Critical	Container {{ $labels.container }} was repeatedly restarted. Actions: • Check to see if a new version of the image was deployed. Also check for issues with the Kubernetes cluster.	kube_pod_container_status_restarts_total	Container {{ $labels.container }} was restarted 5 or more times within 15 minutes.
No sip-nodes available for 2 minutes	Critical	No sip-nodes are available for the pod {{ $labels.pod }}.	sipproxy_active_sip_nodes_count	No sip-nodes are available for the pod {{ $labels.pod }} for 2 minutes.

Alert	Severity	Description	Based on	Threshold
		Actions: • If the alarm is triggered for multiple services, make sure there are no issues with sip-nodes. • If the alarm is triggered only for pod {{ $labels.pod }}, check to see if there is any issues with the pod.		
sip-node capacity limit reached	Warning	The sip-node {{ $labels.sip_node_id }} hit capacity limit on {{ $labels.pod }}. Actions: • If alarm is triggered for multiple services make sure there is no issues with sip-node {{ $labels.sip_node_id }}. • If alarm is triggered only for pod {{ $labels.pod }} check if there is any issue with the pod	sipproxy_sip_node_is_capacity_available	The sip-node {{ $labels.sip_node_id }} hit capacity limit on {{ $labels.pod }} for 3 consecutive minutes.
Pod CPU greater than 80%	Critical	Critical CPU load for pod {{ $labels.pod }}. Actions: • Check whether the horizontal pod autoscaler has triggered	container_cpu_usage_seconds_total container_spec_cpu_period	Container {{ $labels.container }} CPU usage exceeded 80% for 5 minutes.

Alert	Severity	Description	Based on	Threshold
		and the maximum number of pods has been reached. • Check Grafana for abnormal load. • Collect the service logs for pod {{ $labels.pod }} and raise an investigation ticket.		
Pod CPU greater than 65%	Warning	High CPU load for pod {{ $labels.pod }}. Actions: • Check whether the horizontal pod autoscaler has triggered and the maximum number of pods has been reached. • Check Grafana for abnormal load. • Collect the service logs for pod {{ $labels.pod }} and raise an investigation ticket.	container_cpu_usage_seconds_total, container_spec_cpu_period	Container {{ $labels.container }} CPU usage exceeded 65% for 5 minutes.
Pod memory greater than 80%	Critical	Critical memory usage for pod {{ $labels.pod }}. Actions: • Check whether the horizontal pod autoscaler	container_memory_working_set_bytes, kube_pod_container_resource_requests_memory_bytes	Container {{ $labels.container }} memory usage exceeded 80% for 5 minutes

Alert	Severity	Description	Based on	Threshold
		has triggered and the maximum number of pods has been reached. • Check Grafana for abnormal load. • Restart the service for pod {{ $labels.pod }}.		
Pod memory greater than 65%	Warning	Pod {{ $labels.pod }} has high memory usage. Actions: • Check whether the horizontal pod autoscaler has triggered and the maximum number of pods has been reached. • Check Grafana for abnormal load. • Collect the service logs for pod {{ $labels.pod }} and raise an investigation ticket	container_memory_working_set_bytes kube_pod_container_resource_requests_memory_bytes	Container {{ $labels.container }} memory usage exceeded 65% for 5 minutes.
Config node fail	Warning	The request to the config node failed. Action: • Check if there is any problem with pod {{ $labels.pod }} and config node.	http_client_response_count	Requests to the config node fail for 5 consecutive minutes.