

GENESYS

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 SIP Cluster Service metrics and alerts

Contents

- 1 Metrics
- 2 Alerts

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

Service	CRD or annotations?	Port	Endpoint/Selector	Metrics update interval
Voice SIP Cluster Service	Supports both CRD and annotations	11300	http://:11300/metrics	30 seconds

See details about:

- Voice SIP Cluster Service metrics
- Voice SIP Cluster Service alerts

Metrics

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

Metric and description	Metric details	Indicator of		
http_client_request_duration_s				
HTTP client time from request to response, measured in seconds.	Type: histogram Label: le, target_service_name Sample value:	Latency		
http_client_response_count	Unit: N/A			
Number of received HTTP client responses.	Type: counter Label: target_service_name Sample value:	Traffic		
kafka_producer_queue_depth	Unit: N/A			
Number of Kafka producer pending events.	Type: gauge Label: kafka_location Sample value:	Traffic		
kafka_producer_queue_age_seconds				
Age of the oldest producer pending event, measured in seconds.	Type: gauge Label: kafka_location Sample value:	Traffic		

Metric and description	Metric details	Indicator of
kafka_producer_error_total Number of Kafka producer errors.	Unit: N/A Type: counter Label: kafka_location Sample value:	Errors
log_output_bytes_total Total amount of log output in bytes.	Unit: bytes Type: counter Label: level, format, module Sample value:	Traffic
sipnode_requests_total Number of processed requests.	Unit: N/A Type: counter Label: tenant, request Sample value:	Traffic
sipnode_pending_requests_current Number of pending requests.	Unit: N/A rent Type: gauge Label: tenant, request Sample value:	Traffic
sipnode_requests_queue_size Number of postponed requests.	Unit: N/A Type: gauge Label: Sample value:	Saturation
sipnode_sips_request_duration_ Duration of the request processed by SIP Cluster Service, measured in seconds.	Seconds Type: histogram Label: le, tenant, request Sample value:	Traffic
sipnode_events_total Call events streamed to Redis.	Unit: N/A Type: counter Label: tenant, event Sample value:	Traffic
sipnode_ha_writes_total Number of HA writes to Redis.	Unit: N/A Type: counter Label: Sample value:	Traffic
sipnode_ha_reads_total Number of HA reads from Redis.	Unit: N/A Type: counter Label: Sample value:	Traffic
sipnode_monitoring_events_tot Number of monitoring events submitted to Kafka.	aUnit: N/A Type: counter Label: tenant Sample value:	Traffic
sipnode_redis_restored_calls_to	otUhit: N/A	Traffic

Metric and description	Metric details	Indicator of
Total number of restored calls from Redis cache.	Type: counter Label: Sample value:	
sipnode_sips_restarts_total Total number of SIP Server restarts.	Unit: N/A Type: counter Label: Sample value:	Errors
sipnode_sips_disconnects_total Total number of SIP Cluster Service disconnections from SIP Server.	Unit: N/A Type: counter Label: Sample value:	Errors
sipnode_redis_state Current Redis connection state.	Unit: N/A Type: gauge Label: redis_cluster_name Sample value:	Errors
sipnode_ors_tlib_latency_msec T-Library latency from Orchestration Service to SIP Cluster, measured in milliseconds.	Unit: milliseconds Type: histogram Label: le, ors Sample value:	Latency
sipnode_ors_health_check SIP Cluster Service to Orchestration Service health check.	Unit: N/A Type: gauge Label: Sample value:	Traffic
Service_version_info Displays the version of Voice SIP Cluster 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.	<pre>Unit: N/A Type: gauge Label: version Sample value: service_version_info{version="100.0.10000"1</pre>	006"}
sipnode_treatment_not_applied Number of unsuccessful treatments.	Unit: N/A Type: counter Label: tenant Sample value:	Errors
sipnode_default_routing_total Total number of default routed calls.	Unit: N/A Type: counter Label: tenant Sample value:	Traffic
sipnode_envoy_proxy_status Status of the Envoy proxy: -1 - error 0 - disconnected	Unit: N/A Type: gauge Label: Sample value: 1	Health

Metric and description	Metric details	Indicator of
1 – connected		
sipnode_config_node_status	Unit: N/A	
Status of the config node connection:	Type: gauge	Health
0 – disconnected 1 – connected	Label: Sample value: 1	
sipnode_health_level		
Health level of the SIP node (SIP Cluster Service):	Unit: N/A	
-1 - fail	Type: gauge Label:	Traffic
0 - starting 1 - degraded	Sample value: 2	
2 - pass		
sipnode_call_state_health_chec	k ^{Unit:} N/A	
SIP Cluster Service to Call State Service	Type: gauge Label: memberid	Health
health check.	Sample value:	
sips_hastate	Herita NI/A	
Current HA state of SIP Server:	Unit: N/A	
0 - Unknown 1 - backup	Type: gauge Label: Sample value: 2	
2 – primary	Sample value. 2	
	Unit: N/A	
sips_calls	Type: gauge	Traffic
Current number of calls.	Label: Sample value:	
	Unit: N/A	
sips_call_rate	Type: gauge	Traffic
Call rate.	Label: Sample value:	
	Unit: N/A	
sips_cpu_usage_sips	Type: gauge	Saturation
SIP Server CPU usage.	Label: Sample value:	Saturation
	Unit: N/A	
sips_cpu_usage_main	Type: gauge	
SIP Server main thread CPU usage.	Label: Sample value:	Saturation
sips_cpu_usage_cm	Unit: N/A	Caburation
CPU usage of the call manager thread.	Type: gauge Label:	Saturation

Metric and description	Metric details	Indicator of
	Sample value:	
sips_calls_created Total number of created calls.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_abandoned_calls Total number of abandoned calls.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_rejected_calls Total number of rejected calls.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_dialogs_created Total number of created SIP dialogs.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_call_recording_failed Number of failed call recording sessions.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_urs_response_1_to_5_sec Number of URS responses from 1 to 5 seconds.	Unit: N/A Type: gauge Label: Sample value:	Latency
sips_urs_response_more_5_sec Number of URS responses more than 5 seconds.	Unit: N/A Type: gauge Label: Sample value:	Latency
sips_user_data_updates Number of UserData updates.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_routing_timeouts Number of routing timeouts.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_trequest_rate T-Requests rate.	Unit: N/A Type: gauge Label:	Traffic

Metric and description	Metric details	Indicator of
	Sample value:	
sips_treatment_rate TApplyTreatment requests rate.	Unit: N/A Type: gauage Label: Sample value:	Traffic
sips_userdata_rate UserData change rate.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_sips_memory_usage Memory usage of the SIP Server process.	Unit: N/A Type: gauge Label: Sample value:	Saturation
sips_stat_fetch_total Number of successful SIP Server statistic fetches.	Unit: N/A Type: counter Label: Sample value:	Other
sips_sip_response_time_ms SIP Server metric of response time, measured in milliseconds.	Unit: milliseconds Type: histogram Label: le Sample value:	Latency
sips_trunk_in_service Trunk devices that are in service.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Traffic
sips_trunk_ncallscreated Number of created calls per trunk.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Traffic
sips_trunk_noos_detected Number of trunks that are out of service.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Errors
sips_trunk_n4xx_received Number of received 4xx messages.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Errors
sips_trunk_n5xx_received Number of received 5xx messages.	Unit: N/A Type: gauge Label: device_name, tenant	Errors

Metric and description	Metric details	Indicator of
	Sample value:	
sips_trunk_n6xx_received Number of received 6xx messages.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Errors
sips_softswitch_in_service Softswitch devices that are in service.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Traffic
sips_softswitch_ncallscreated Number of created calls per softswitch device.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Traffic
sips_softswitch_noos_detected Number of softswitch devices that are out of service.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Errors
sips_softswitch_n4xx_received Number of received 4xx messages.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Errors
sips_softswitch_n5xx_received Number of received 5xx messages.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Errors
sips_softswitch_n6xx_received Number of received 6xx messages.	Unit: N/A Type: gauge Label: device_name, tenant Sample value:	Errors
sips_msml_in_service MSML devices that are in service.	Unit: N/A Type: gauge Label: device_name Sample value:	Traffic
sips_msml_ncallscreated Number of created calls per MSML device.	Unit: N/A Type: gauge Label: device_name Sample value:	Traffic
sips_msml_noos_detected Number of MSML devices that are out of service.	Unit: N/A Type: gauge Label: device_name	Errors

Metric and description	Metric details	Indicator of
	Sample value:	
sips_msml_n4xx_received Number of received 4xx messages.	Unit: N/A Type: gauge Label: device_name Sample value:	Errors
sips_msml_n5xx_received Number of received 5xx messages.	Unit: N/A Type: gauge Label: device_name Sample value:	Errors
sips_msml_n6xx_received Number of received 6xx messages.	Unit: N/A Type: gauge Label: device_name Sample value:	Errors
sips_dp_state Dial Plan Service state: 0 - Out-Of-Service 1 - In-Service	Unit: N/A Type: gauge Label: Sample value: 1	Traffic
sips_dp_queue_size Size of the request queue to Dial Plan Service.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_dp_avg_queue_time Average queue time (msec) of requests to Dial Plan Service.	Unit: milliseconds Type: gauge Label: Sample value:	Latency
sips_dp_connections Number of connections to Dial Plan Service per URL.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_dp_active_connections Number of active connections to Dial Plan Service.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_dp_req_rate Request rate to Dial plan Service.	Unit: N/A Type: gauge Label: Sample value:	Traffic
sips_dp_400_errors Dial Plan Service 400 type of errors.	Unit: N/A Type: gauge	Errors

Metric and description	Metric details	Indicator of
	Label: Sample value:	
sips_dp_404_errors Dial Plan Service 404 type of errors.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_dp_4xx_errors Dial Plan Service 4xx type of errors.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_dp_500_errors Dial Plan Service 500 type of errors.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_dp_501_errors Dial Plan Service 501 type of errors.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_dp_5xx_errors Dial Plan Service 5xx type of errors.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_dp_timeouts Dial Plan Service timeouts.	Unit: N/A Type: gauge Label: Sample value:	Errors
sips_dp_average_response_late Dial Plan Service average response latency.	Type: gauge Label: Sample value:	Latency
sips_sipproxy_in_service SIP Proxy Service state: 0 - Out-Of-Service 1 - In-Service	Unit: N/A Type: gauge Label: Sample value: 1	Traffic
trunk_config_synced_count Number of trunks synchronized with SIP Server.	Unit: N/A Type: gauge Label: Sample value:	
trunk_config_cached_count	Unit: N/A	

Metric and description	Metric details	Indicator of
Number of trunks obtained from the config node.	Type: gauge Label: Sample value:	
trunk_config_cfg_node_error_cd	uupit: N/A	
Number of failed attempts to read from the config node.	Type: counter Label: Sample value:	
trunk_config_tlib_connection	Unit: N/A	
Number of trunks with the T-Library connection.	Type: gauge Label: Sample value:	

Alerts

The following alerts are defined for Voice SIP Cluster Service.

Alert	Severity	Description	Based on	Threshold
Too many Kafka pending events	Critical	Too many Kafka producer pending events for pod {{ \$labels.pod }}. Actions: • Ensure there are no issues with Kafka, {{ \$labels.pod }} pod's CPU, and network.	kafka_producer_que	Too many Kafka producer pending events for service {{ u\$labpl\$,container }} (more than 100 in 5 minutes).
Dial Plan node is overloaded	Critical	Dial Plan node is overloaded as the response latency increases. Actions: Check that the inbound call rate to SIP Server is not too high. Check the Dial Plan node CPU and memory usage.	sips_dp_average_res	Dial Plan node is overloaded as the response latency pionsease (gore than 1000).

Alert	Severity	Description	Based on	Threshold
		 Check the network connection between SIP Server and Dial Plan nodes. 		
Dial Plan Queue Increase	Critical	Because Dial Plan requests are huge in size or there is a connection issue with the Dial Plan node, the processing queue size increases in size. Actions: Check SIP Server inbound call rate. Check the connection between SIP Server and the Dial Plan node.	sips_dp_queue_size	The processing queue size is greater than 10 requests for 1 minute.
SIP Proxy overloaded	Critical	SIP Proxy is overloaded. Actions: Check SIP Proxy nodes for CPU and memory usage. If SIP Proxy nodes have acceptable CPU and memory usage, then check for errors or a "hang-up" state which could delay SIP Proxy in forwarding. Check the SBC side for network	sips_sip_response_ti sips_sip_response_ti	Response time is greater than 20 m@illisesgnds for 1 m@inusecount

Alert	Severity	Description	Based on	Threshold
		delays.		
SIP Node HealthCheck Fail	Critical	SIP Node health level fails for pod { { \$labels.pod } }. Actions: • Check for failure of dependent services (Redis/ Kafka/SIP Proxy/GVP/Dial Plan). • Check for Envoy proxy failure, then restart the pod.	sipnode_health_leve	SIP Node health level fails for pod {{ \$labels.pod }} for 5 minutes.
Kafka not available	Critical	Kafka is not available for pod {{ \$labels.pod }}. Actions: If the alarm is triggered for multiple services, ensure there are no issues with Kafka. Restart Kafka. If the alarm is triggered only for pod {{ \$labels.pod }}, check if there is an issue with the pod.	kafka_producer_state	Kafka is not available for pod {{ \$labels.pod }} for 5 minutes.
Pod Status Error	Warning	Actions: • Restart the pod. Check if there are any issues with the pod after restart.	kube_pod_status_pha	Pod {{ \$labels.pod }} is in Failed, Unknown, or Pending state.
Pod Status NotReady	Warning	Pod {{ \$labels.pod }} is in NotReady	kube_pod_status_rea	Pod {{ \$labels.pod ady} is in NotReady

Alert	Severity	Description	Based on	Threshold
		state. Actions: Restart the pod. Check if there are any issues with the pod after restart.		state for 5 minutes.
Container Restarted Repeatedly	Critical	Container { { \$labels.container } } was repeatedly restarted. Actions: Check if the new version of the image was deployed. Check for issues with the Kubernetes cluster.	kube_pod_container_	Container {{ \$labels.container }} was restarted 5 ទុធឥយទ្ធាខន្ធ total within 15 minutes.
Ready Pods below 60%	Critical	The number of statefulset { { \$labels.statefulset } } pods in the Ready state has dropped below 60%. Actions: Check if the new version of the image was deployed. Check for issues with the Kubernetes cluster.	kube statefulset sta	For the last 5 minutes, fewer than 60% of the currently available tsatefulses feady, tsatefulses feady, tsatefulses feady the Ready state.
Pods scaled up greater than 80%	Critical	The current number of replicas is more than 80% of the maximum number of replicas. Actions:	kube_hpa_status_cur kube_hpa_spec_max	For 5 consecutive minutes, the number of replicas is more than 80% replicas of replicas of replicas.

Alert	Severity	Description	Based on	Threshold
		 Check if max replicas must be modified based on load. 		
Pods less than Min Replicas	Critical	The current number of replicas is less than the minimum replicas that should be available. This might be because Kubernetes cannot deploy a new pod or pods are failing to be active/ready. Actions: If all services have the same issue, then check Kubernetes nodes and Consul health. If the issue is only with the SIP Cluster service, then check pod logs or the deployment manifest/helm errors.	kube_hpa_status_cu kube_hpa_spec_min_	For 5 consecutive minutes, the number of replicas is less than the replicas replicas replicas replicas available.
Pod CPU greater than 80%	Critical	Critical CPU load for pod {{ \$labels.pod }}. Actions: Check whether the horizontal pod autoscaler has triggered and if the maximum number of pods has been reached. Check Grafana for abnormal	container_cpu_usage container_spec_cpu_	Container { { \$labels.container }} CPU usage = \$666466 859% for PSTIMMutes.

Alert	Severity	Description	Based on	Threshold
		Collect the service logs for pod {{ \$labels.pod }}; 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 if the maximum number of pods has been reached. Check Grafana for abnormal load. Collect the service logs for pod {{ \$labels.pod }}; raise an investigation ticket.	container_cpu_usage container_spec_cpu_	Container { { \$labels.container }} CPU usage = SECORDS 185% for -PSTIMMutes.
Pod memory greater than 80%	Critical	Critical memory usage for pod { { \$labels.pod } }. Actions: • Check whether the horizontal pod autoscaler has triggered and if the maximum number of pods has been reached. • Check Grafana	container_memory_v kube_pod_container	Container {{ \$labels.container }} memory usage Wexcledated at the container _resniffcesequests_m

Alert	Severity	Description	Based on	Threshold
		for abnormal load. Restart the service for pod {{ \$labels.pod }}.		
Pod memory greater than 65%	Warning	High memory usage for pod {{ \$labels.pod }}. Actions: • 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 for pod {{ \$labels.pod }}; raise an investigation ticket.	container_memory_v kube_pod_container	Container {{ \$labels.container }} memory usage verking sed 3% for _rssn#{Gesequests_m
Redis not available	Critical	Redis is not available for pod {{ \$labels.pod }}. Actions: If the alarm is triggered for multiple services, ensure there are no issues with Redis. Restart Redis. If the alarm is triggered only for pod {{ \$labels.pod }},	redis_state	Redis is not available for pod {{ \$labels.pod }} for 5 consecutive minutes.

Alert	Severity	Description	Based on	Threshold
		check if there is an issue with the pod.		
Too many Kafka producer errors	Critical	Kafka responds with errors at pod {{ \$labels.pod }}. Actions: • For pod {{ \$labels.pod }}, ensure there are no issues with Kafka.	kafka_producer_erro	More than 100 errors for 5 consecutive ruminutes.
SIP Server main thread consuming more than 65% CPU for 5 mins	Warning	Main thread consumes too much CPU. Actions: • Collect SIP Server Main thread logs; that is, log files without index in the file name (appname_date.l files). Raise an investigation ticket.	sips_cpu_usage_mai	Main thread consumes too much CPU (more than 65% for 5 nconsecutive minutes).
Calls activity drop	Warning	A noticeable reduction in the number of active calls on a specific SIP Server and no new calls are arriving for processing. Actions: If a problematic SIP Server is primary, do a switchover, and then restart the former primary server. If a problematic SIP Server is	sips_calls, sips_calls_created	The absolute value of active calls on a specific SIP Server dropped by more than 30 calls in 2 minutes and no new calls are arriving at the SIP Server for processing.

Alert	Severity	Description	Based on	Threshold
		backup, restart the backup server. Collect SIP Server Main thread logs; that is, log files without index in the file name (appname_date.l files). Raise an investigation ticket.	og	
Dial Plan Node Down	Critical	No Dial Plan nodes are reachable from SIP Server and all connections to Dial Plan nodes are down. Actions: Check the network connection between SIP Server and the Dial Plan node host. Check the Dial Plan node cPU and memory usage.	sips_dp_active_conn	All connections to Dial Plan nodes are eelগুপান্ত
Dialplan Node problem	Warning	Dial Plan node rejects requests with an error or it doesn't respond to requests and requests are timed out. Actions: Check the network connection between SIP Server and the Dial Plan host. Check that Dial Plan nodes are running.	sips_dp_timeouts	During 1 minute, the Dial Plan node rejects more than 5 requests with an error or more than 5 requests time out because the Dial Plan node fails to respond.

Alert	Severity	Description	Based on	Threshold
Routing timeout counter growth	Warning	The trigger detects that routing timeouts are increasing. Actions: Check the URS_RESPONSE_stat value. If it's increasing, then investigate why URS doesn't respond to SIP Server in time. Check SIPS-to-URS network connectivity.	MORE5SEC sips_routing_timeout	The absolute value of NROUTINGTIMEOUTS on a specific SIP Server increased by more than 20 in 2 minutes.
SIP trunk is out of service	Critical	SIP trunk is out of service. Actions: • For Primary and Secondary trunks: • Troubleshoot SIP Serverto-SBC network connection. Collect network stats and escalate to the Network team to resolve network issues, if necessary. • Troubleshoot the SBC. For Inter-SIP Server trunks: troubleshoot the SIP Server-to-	sips_trunk_in_service	SIP trunk is out of service for more than 1 minute.

Alert	Severity	Description	Based on	Threshold
		SIP Server network connection. Collect network stats and escalate to the Network team to resolve network issues, if necessary.		
Media service is out of service	Critical	Media service is out of service. Actions: Troubleshoot the SIP Serverto-Resource Manager (RM) network connection. Collect network stats and escalate to the Network team to resolve network issues, if necessary. Troubleshoot RM, consider RM restart. After 5 minutes, redirect traffic to another site.	sips_msml_in_service	Media service is out of service for more than 1 minute.
SIP softswitch is out of service	Critical	Actions: • Troubleshoot the SIP Serverto-SBC network connection. Collect network stats and escalate to the Network team to resolve network issues,	sips_softswitch_in_se	SIP softswitch is out of service. ervice

Alert	Severity	Description	Based on	Threshold
		if necessary.Troubleshoot the SBC.		
SIP Proxy is out of service	Critical	Actions: • Troubleshoot the SIP Serverto-SIP Proxy nodes network connections. Collect network stats and escalate to the Network team to resolve network issues, if necessary. • Troubleshoot SIP Proxy nodes.	sips_sipproxy_in_ser	SIP Proxy is out of vice service.