# HEAL Documentation

HEAL Software Inc.

# Telemetry Ingestion and GenAI Forensic Analysis

## Table of Contents

# Version History

| Doc Version | Date | Changes Done |
|---|---|---|
| 1.0 | 07 October 2025 | Initial Release |

# Introduction

This document explains the steps to configure the application to collect minute-level process telemetry and also provides the GenAI-driven forensics across the observability stack.

It defines the centralized data model, retention controls, operational safeguards, and integration touchpoints required to standardize collection and analysis within OpenSearch. To establish a single source of interfaces that accelerate troubleshooting, support proactive investigations, and remain configurable for evolving operational needs.

The document’s configuration enables the following actions:

* Collects **ps -ef** with per-process CPU% and MEM% every minute.

* Stores data in OpenSearch with a default 1-day retention (configurable for longer).

# Prerequisites

Ensure the following prerequisites are met:

* Linux hosts with ps, sudo, and a minute scheduler.

* Cluster-wide time sync is enabled.

* OpenSearch is reachable with ISM permissions and sufficient capacity.

* Consul, RabbitMQ, Redis, and HAProxy access with network/firewall is allowed.

* DB migration access – Running_Process_Details artifact is available, and DT/AppD connector credentials.

# Configuration

Follow the configuration steps below:

1. Before running migrations, remove any existing rows for relevant agents from the table **instance_command_mapping.**

2. Migrate the database with the shared SQL files.

3. Update the following Consul keys:

```
service/datareceiver/cache/worker/thread/queue/size 800
service/datareceiver/cache/worker/thread/size 100
service/datareceiver/opensearch/connection/io/reactor/size 2
service/datareceiver/redis/connection/timeout/secs 5
service/datareceiver/redis/socket/timeout/secs 30
service/datareceiver/rmq/queue/max/size 10000
service/datareceiver/sink/rmq/data/interval/milliseconds 5000
```

4. Update the lines below in **haproxy.cfg** and reload HAProxy to apply changes.

```
acl url_dr_http7 path_beg -i /external-data
use_backend data_receiver_http_backend if url_dr_http7
```

5. Place the provided **Running_Process_Details** script into the Forensic Agent identifier folder.

6. Update the **instance_command_mapping table** by setting **status = 1** (based on agent identifiers) to enable **process running metrics collection** and **store the data in OpenSearch**.

7. OpenSearch index naming for process running metrics is **heal_raw_external_data_${year.month.date}**.

8. Update the ISM policy **ism_heal_external_data_indexes** by changing **min_index_age** from **180d** to **7d**.

> **Note**:
>
> * Approximate OpenSearch CPU overhead for processing process running metrics: 40%
>
> * Approximate OpenSearch memory overhead: 5%
>
> * With min_index_age set to 7d, approximately 856 GB is required for one week of data.
>
> * With min_index_age set to 180 days, storage is approximately 5.1 GB per hour, so 122 GB per day, and 22 TB for 180 days.
>
> * Observed with a process line output of approximately 5,000 lines.