The SeaDataCloud Monitoring Platform

Angelos Iykiardopoulos, HCMR (Greece), angelo@hcmr.*gr*
Themis zamani, GRNET (Greece), themis@admin.grnet.gr
Kostas koumantaros, GRNET (Greece), kkoum@admin.grnet.gr
Kostas Kagkelidis, GRNET (Greece), kkoum@admin.grnet.gr

SeaDataCloud (SDC) is a standardized system for managing the large and diverse data sets collected by the oceanographic fleets and the automatic observation systems. The SeaDataCloud platform currently includes national oceanographic data centres of 35 countries, active in data collection. The whole platform operates a unique virtual data management system providing integrated data sets of standardized quality on-line. SDC operates a number of distributed services that range from dataset acquisition, management, replication and delivery to the end users. Through the SeaDataCloud (SDC) project the infrastructure is extended in order to efficiently store, replicate and deliver the required datasets by utilizing the EUDAT CDI. As a research infrastructure, SeaDataNet contributes to build research excellence in Europe.

In order to fulfil that it needs to be constantly monitored for the availability and rehabilitee of the whole system as well as for each service separately. In the framework of SeaDataNet a monitoring system was developed based on Nagios as monitoring engine. On top of this a monitoring portal developed form HCMR in order to perform all necessary operations to perform the appropriate availability calculations and provide reporting as well as access to service administrators to the monitoring platform.

In the framework of SDC the whole platform is going to be upgraded adopting the ARGO monitoring system as monitoring platform.

The ARGO Monitoring service is comprised of the following building blocks - components. All these components are deployed and configured to monitor SeaDataCloud:

- The messaging service which is the layover layer between the ARGO monitoring engine and the other components. ARGO Messaging Service is a real-time messaging service that allows you to send and receive messages between independent applications.
- The monitoring engine that executes the service checks against the infrastructure and delivers the metric data (probe check results) to a Messaging network (messaging).
- The POEM service which is used in order to define checks (probes) and associate them to service types.
- The connectors is a collection of libraries that periodically (usually once per day) connect to sources of truth (the monitoring service can connect to multiple external Configuration Management Databases and Service Catalogues) and deliver the information to the analytics engine in a predefined format. The various data sources established in ARGO infrastructure are mainly data from a configuration database (ex: topology, downtimes), but there's also support for fetching alternative topology via various VO feeds, weights information for various sources and POEM metric profiles.
- The analytics engine collects the metric data from the Messaging Network. It is responsible for computing the availability and reliability (A/R) of services using the metric results that are collected from monitoring engine(s), and information from the various sources of truth (ex. topology, downtimes). Results (status and A/R) are passed onto a fast, reliable and distributed data store. It supports stream processing in real time. Monitoring results flow through the

messaging service, to the streaming layer (in parallel to the HDFS). The streaming layer is used in order to push raw metric results to the metric result store and to compute status results and push them to the status store in real-time.

- The Web API provides the Serving Layer of ARGO that delivers all computed status and A/R results via a programmatic interface. It is comprised of a high performance and scalable data store and a multi-tenant REST HTTP API.
- The Web UI which is the main interface for the users. It is used in order to represent the status and A/R results graphically and gives the ability to any given user to drill down from the availability of a given resource down to the actual metric results that were recorded and contributed to the computed figures.
- If there is a problem with a service an alert notification is sent via the notification engine. Based on the real-time streaming layer, the alerting functionality is introduced to the ARGO monitoring engine that goes beyond infrastructure monitoring. This component analyses the monitoring results and sends notification based on a set of predefined rules.