

# Developing a monitoring system for Cloud-based distributed datacenters

*Tuesday, July 10, 2018 4:40 PM (20 minutes)*

Beside their increasing complexity and variety of provided resources and services, large data-centers nowadays often belong to a distributed network and need non-conventional monitoring tools. This contribution describes the implementation of a monitoring system able to provide active support for problem solving to the system administrators.

The key components are information collection and analysis. Information is acquired from multiple levels in order to allow algorithms to recognize a malfunction and suggest possible root-cause reducing service downtime.

The project has been developed using the Bari ReCaS data-center as testbed. The information is gathered from Zabbix, Openstack, HTCCondor as local monitoring system, cloud platform and batch system respectively.

Big Data solutions belonging to the Hadoop ecosystem have been selected: Flume and Kafka as transport layer and Spark as analysis component. Multiple tools have been used to store data, such as Hadoop Distributed File System, HBase and Neo4j. InfluxDB-Grafana and Elasticsearch-Kibana are used as visualization component. Event extraction, correlation and propagation algorithms have been also implemented using Artificial Intelligence and graph libraries to provide the root-cause feature. Results are forwarded to experts by email or Slack, using Riemann.

**Primary authors:** VINO, Gioacchino (Universita e INFN, Bari (IT)); ELIA, Domenico (INFN Bari); DONVITO, Giacinto (Universita e INFN, Bari (IT)); ANTONACCI, Marica

**Presenter:** VINO, Gioacchino (Universita e INFN, Bari (IT))

**Session Classification:** Posters

**Track Classification:** Track 8 –Networks and facilities