Improved Pilot Logging in DIRAC

J. Martyniak, F. Stagni & S. Fayer
lcg-site-admin@imperial.ac.uk

Introduction

- DIRAC is a widely used framework for distributed computing. It works by building a layer between the users and the resources, offering a common interface to a number of heterogeneous providers.
- DIRAC uses pilot jobs to check and configure the worker-node environment before fetching a user payload.
- The log messages generated by these pilot jobs ("pilot logs") are crucial for diagnosing problems in the infrastructure.
  - Retention policies for these logs vary by technology and resource provider, with some providers not storing them at all.
  - Collecting pilot logs in a reliable, resource independent manner was identified as a high priority issue for many DIRAC communities.
- The aim of the project is develop a highly configurable remote pilot logging system.

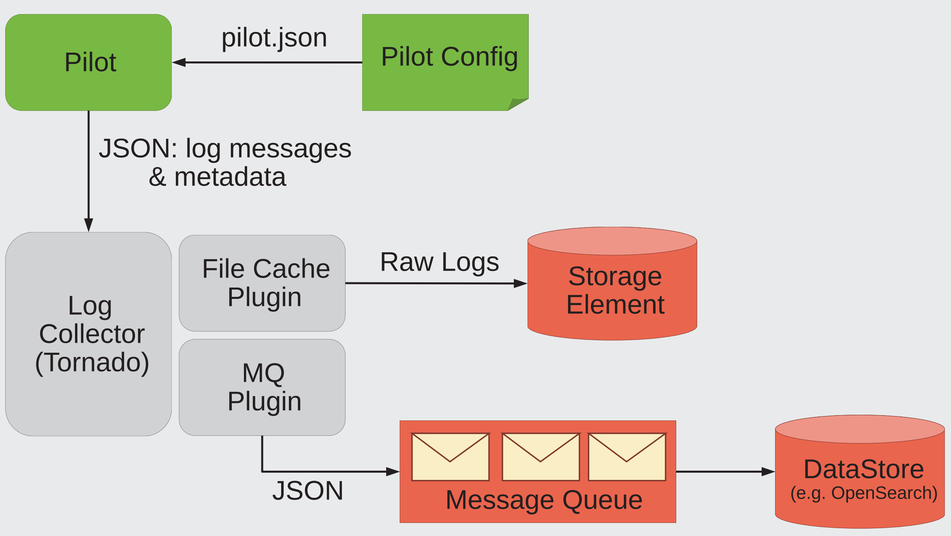
Requirements

- Preserve the current, "classic" pilot logger functionality.
- Build a configurable remote logging system which can operate in parallel with the original implementation.
- It should be possible to activate the remote logging on a resource-by-resource basis.
- The collected logs should be stored for further analysis. The storage system should be plugin-based so that different DIRAC instances can pick the most appropriate method for their use-case.

Implementation

- Pilots send logs to a central log collector service based on a Tornado Web Server.
- The logging level of messages sent is configurable.
- A selectable plugin stores the messages to external, permanent storage. This can be a Grid storage element or an MQ service and can be configured independently for each community.
- For fault tolerance we plan to add support for running multiple instances of the collector service.

System architecture



Status and future work

- Implemented and ready for installation on pre-production servers.
- The next step is to run a performance test at a larger scale.
- Integrate the new system with the existing DIRAC log browser.
- Design and implement an OpenSearch dashboard for analysing pilot logs.

https://www.github.com/DIRACGrid