

Project-based Data Collection

The parameters of data collection currently can be configured only centrally, which means that the settings are applicable for all projects with no exceptions.

Cloud services can be offered to several consumers, who use them in different ways for different purposes. The data, gathered by Ceilometer, can also be used for different purposes from billing to monitoring. The multiple areas of processing the collected data and the several projects, which use the cloud services, introduce the need for configuring the data collection mechanism in a project specific way. It means that for instance, if the cpu utilization data is needed in every 10 seconds for one of the projects and is enough in every 2 minutes for another project, it is not needed to collect the samples in every 10 seconds for both projects. The project specific measurement parameters would also decrease the load on Ceilometer, by making it possible to store only as much information for a project, as necessary.

The data collection and processing is controlled by a chain of mechanisms called pipeline. The pipeline related configuration is now stored in the pipeline.yaml configuration file. To make the settings of a pipeline project specific, the pipeline sources definition should be extended with an additional section, called 'projects'. This section should contain a list of project_ids, for which the pipeline definition is applicable. For backward compatibility support, if the projects block is missing, it means that the configuration of that pipeline is applicable for all projects. For the new HW related counters, where the project mapping is mostly missing, another differentiator will be needed to provide a more flexible configuration, than the currently existing one.

Sample pipeline configuration:

```
---
sources:
  - name: meter_source
    interval: 600
    meters:
      - "*"
    projects:
      - project_id_1
      - project_id_2
    sinks:
      - meter_sink
  - name: cpu_source
    interval: 600
    meters:
      - "cpu"
    projects:
      - project_id_1
      - project_id_2
    sinks:
      - cpu_sink
```

```

- name: meter_source_60
  interval: 60
  meters:
    - "*"
  projects:
    - project_id_3
  sinks:
    - meter_sink

sinks:
- name: meter_sink
  transformers:
  publishers:
    - rpc://
- name: cpu_sink
  transformers:
    - name: "rate_of_change"
      parameters:
        target:
          name: "cpu_util"
          unit: "%"
          type: "gauge"
          scale: "100.0 / (10**9 * (resource_metadata.cpu_number or 1))"
  publishers:
    - rpc://

```

By making a change only in this configuration file, the cloud administrator is still the only authorized user, who can set these project specific settings.