

Staff Data Engineer

At Hydrosat we use Dagster on a k8s cluster to execute a broad variety of data processing and machine learning jobs. This assignment is to create a simplified version of our setup and day-to-day tasks.

In particular, we would like you to deploy Dagster to a k8s cluster and materialize at least one Dagster “asset”. The asset(s) shall replicate a simplified version of assets we are dealing with at Hydrosat. It shall have daily partitions, where each partition depends on the partition of the preceding day. Further, it shall read asset inputs from and write asset outputs to a configurable S3 bucket (or an equivalent of the latter).

The asset logic should account for the following aspects:

- **Asset inputs**
 - A square/rectangular bounding box, which acts as the processing extent for any geospatial operation
 - Multiple “fields”, represented as polygons, which intersect with the bounding box and have different planting dates
- **Processing**
 - The asset should download or simulate any data of choice and process it within the extent of the bounding box. It shall provide some output values obtained on the field extent for each of the fields, starting at the field’s planting date.
- **Asset output**
 - The asset output should be one or multiple files containing the output values per field per daily partition
- **Complication**
 - Assume that some field data arrives late, e.g. because they were entered late in the system. This means that the asset’s processing status has reached timepoint t , but the field should have been processed at timepoint t_2 . How to handle this situation without reprocessing the entire bounding box?

Notes:

- You can use any k8s deployment you like
- Using any IaC tool is considered a plus. At Hydrosat, we use Terraform.
- You can use AI tools. Please disclose which AI tools you used.
- You are expected to spend no more than 8 hours on the assignment

- The end result should be a public git repository, that would allow us to reproduce the cluster setup and asset materialization.