

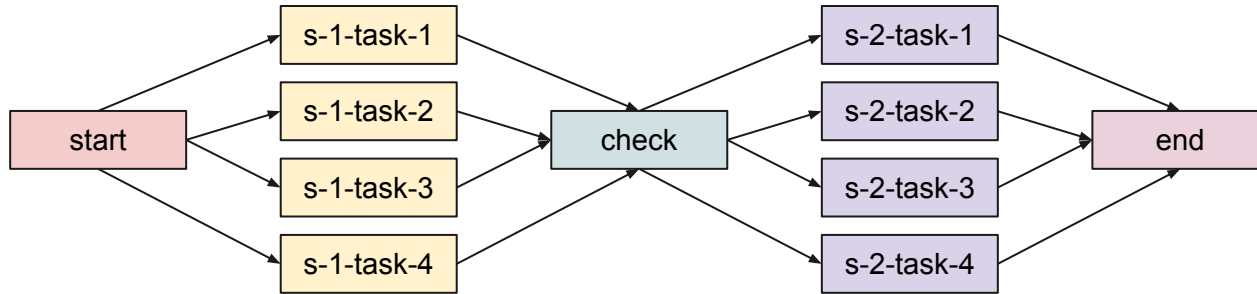


Minimising Repetitive Patterns With SubDAGs

How to make your dags clearer

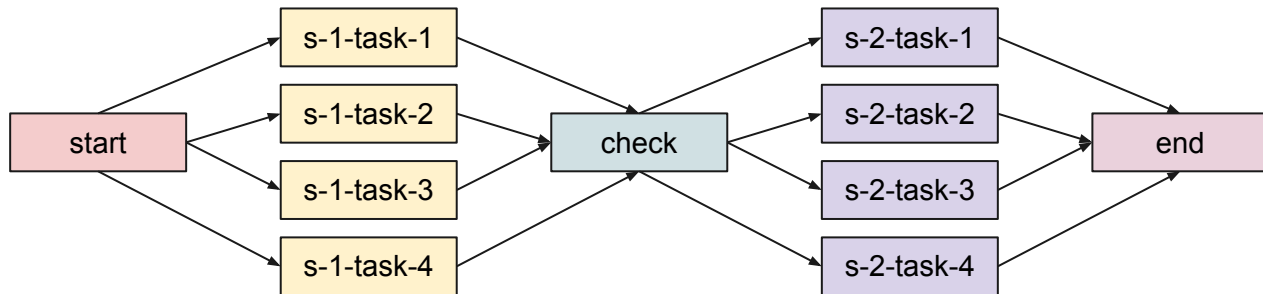
Use Case

Let's suppose we have a DAG like this:



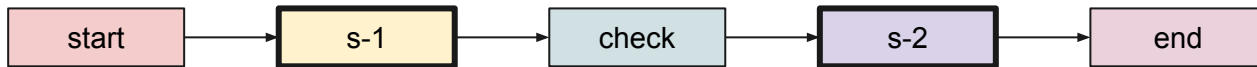
Use Case

As you can see we have many parallel tasks which could be combined in order to make the DAG clearer.



Use Case

By using subDAGs you can end up with the following DAG:



All the tasks are still part of the DAG but they are grouped into the two boxes with bold borders.

As we are going to see in the next lesson, Airflow UI allows you to zoom into each subDAG and sees which task are in it.



How to Create a SubDAG?

- In order to create a subDAG you have to use a factory function that returns a DAG Object (the subDAG in our case) and the SubDagOperator to attach the subDAG to the main DAG.
- The factory function returns an instantiated DAG with the associated tasks. This function should be in a different file from where you main DAG is defined.



Important Notes

- The main DAG manages all the subDAGs as normal tasks. They are going to follow the same dependencies you had before. Nothing change basically.
- Airflow UI only shows the main DAG. In order to see your subDAGs you will have to click on the related main DAG and then “zoom in” into the subDAGs from the “Graph View”.
- SubDags must be scheduled the same as their parent DAG.