Task Boundaries

DAG tasks should be designed such that they are:

- Atomic and have a single purpose
- Maximize parallelism
- Make failure states obvious

Every task in your dag should perform **only one job.**

> "Write programs that do one thing and do it well." - Ken Thompson's Unix Philosophy

Benefits of Task Boundaries

- Re-visitable: Task boundaries are useful for you if you revisit a pipeline you wrote after a 6 month absence. You'll have a much easier time understanding how it works and the lineage of the data if the boundaries between tasks are clear and well defined. This is true in the code itself, and within the Airflow UI.
- Tasks that do just one thing are often more easily parallelized. This parallelization can offer a significant speedup in the execution of our DAGs.