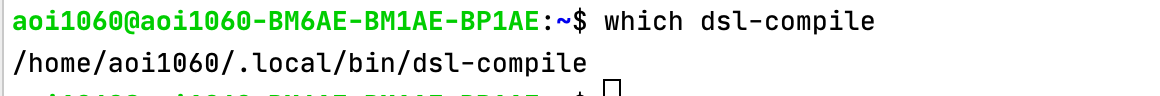
pipelines  
<https://www.kubeflow.org/docs/pipelines/>

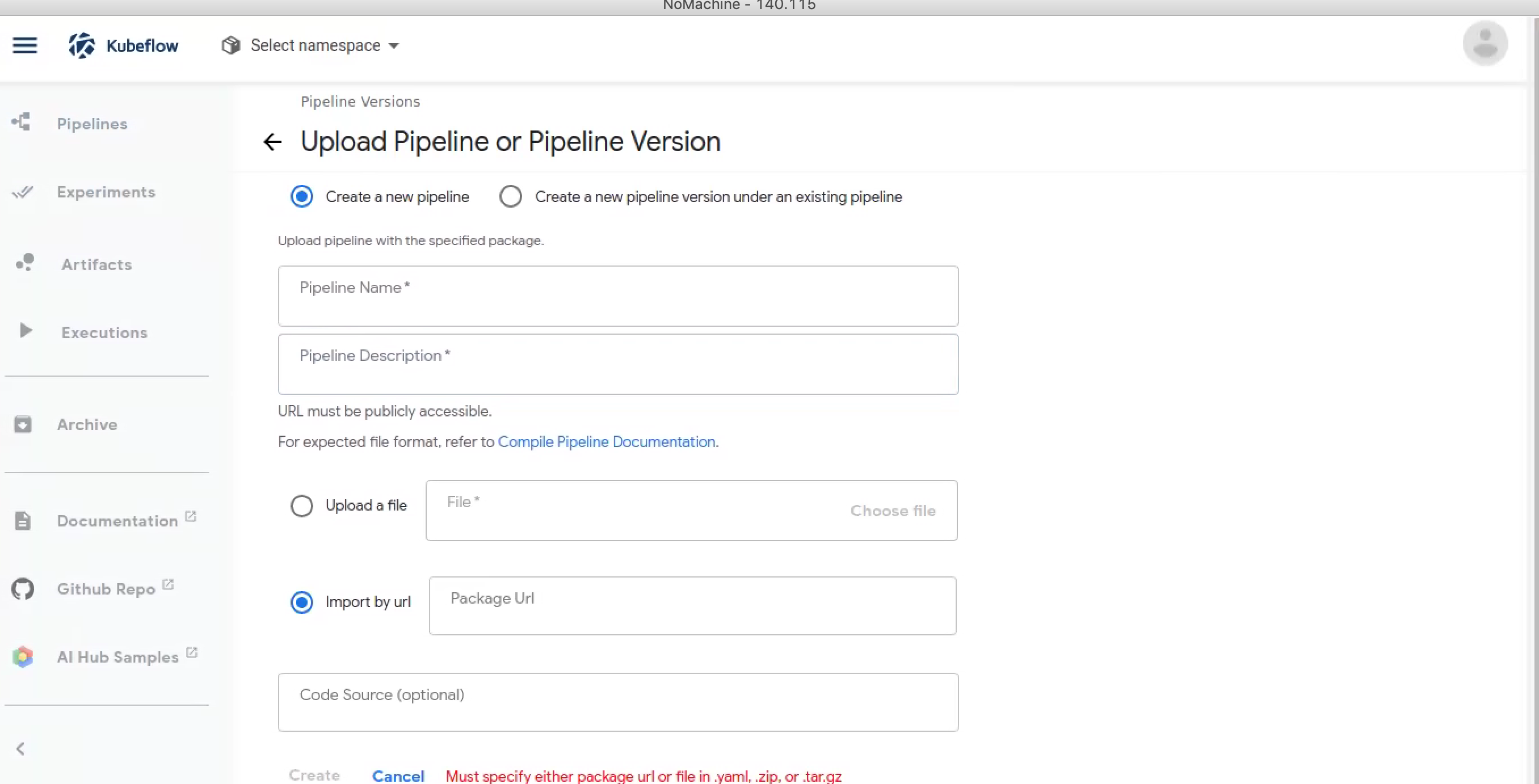
# 1、 安装pipelines SDK  
pip install https://storage.googleapis.com/ml-pipeline/release/0.1.20/kfp.tar.gz --upgrade  
# 2、 检查一下安装状态  
which dsl-compile

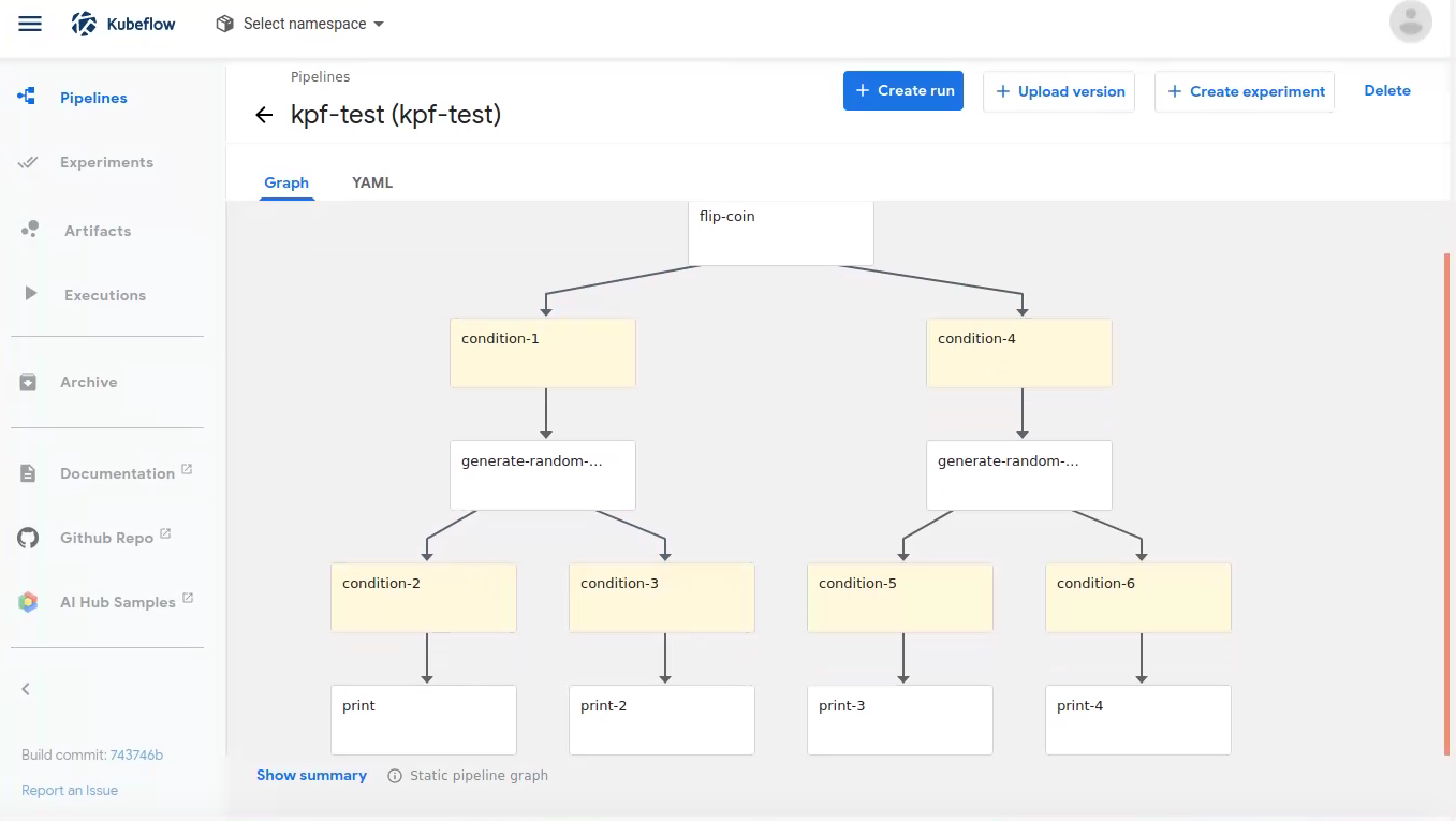
  
# 3、 编译需要上传的pipeline.py文件  
dsl-compile --py condition.py --output test.tgz



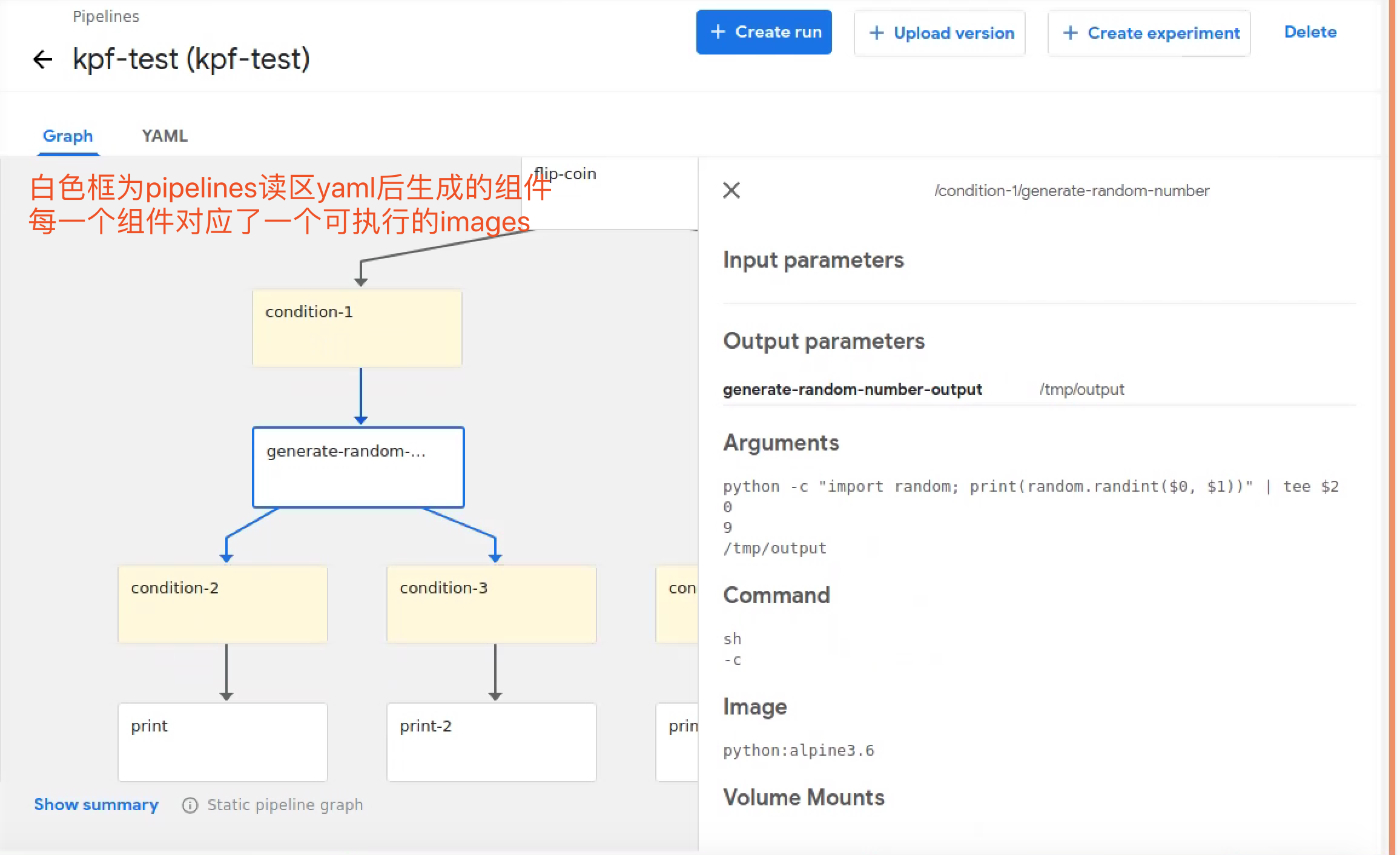
# 4、上传.tgz 文件/yaml档案

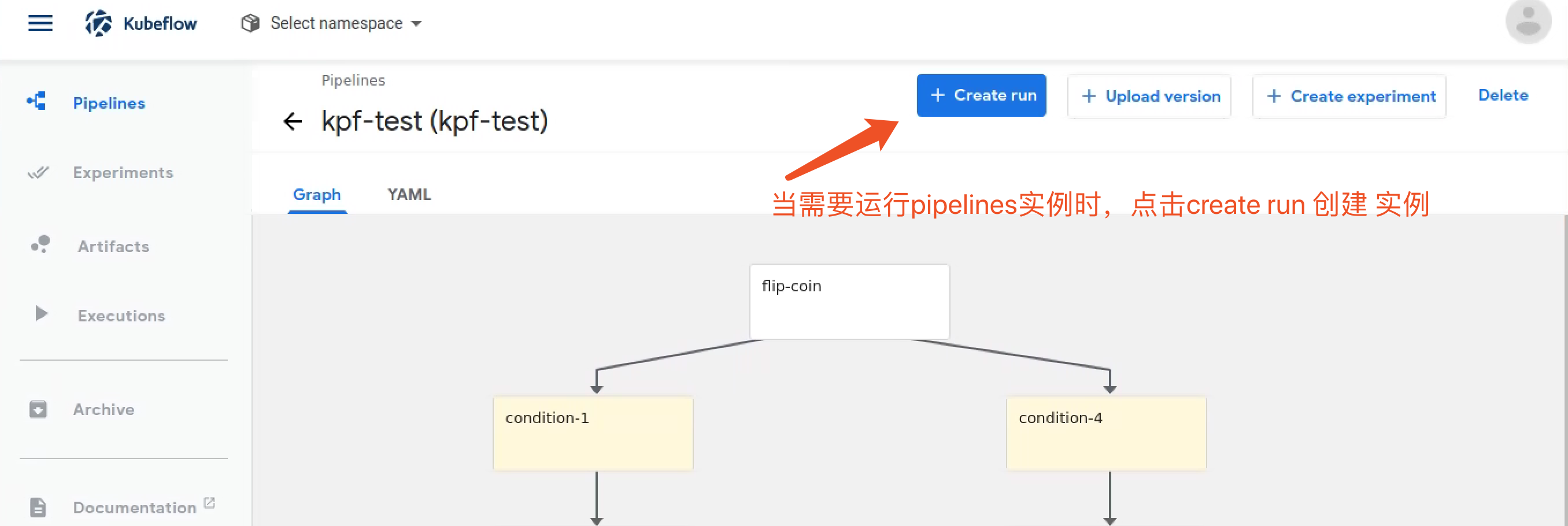
通过前端上传 zip、tgz文件，后端会解析成 Argo 需要的 YAML 文件。

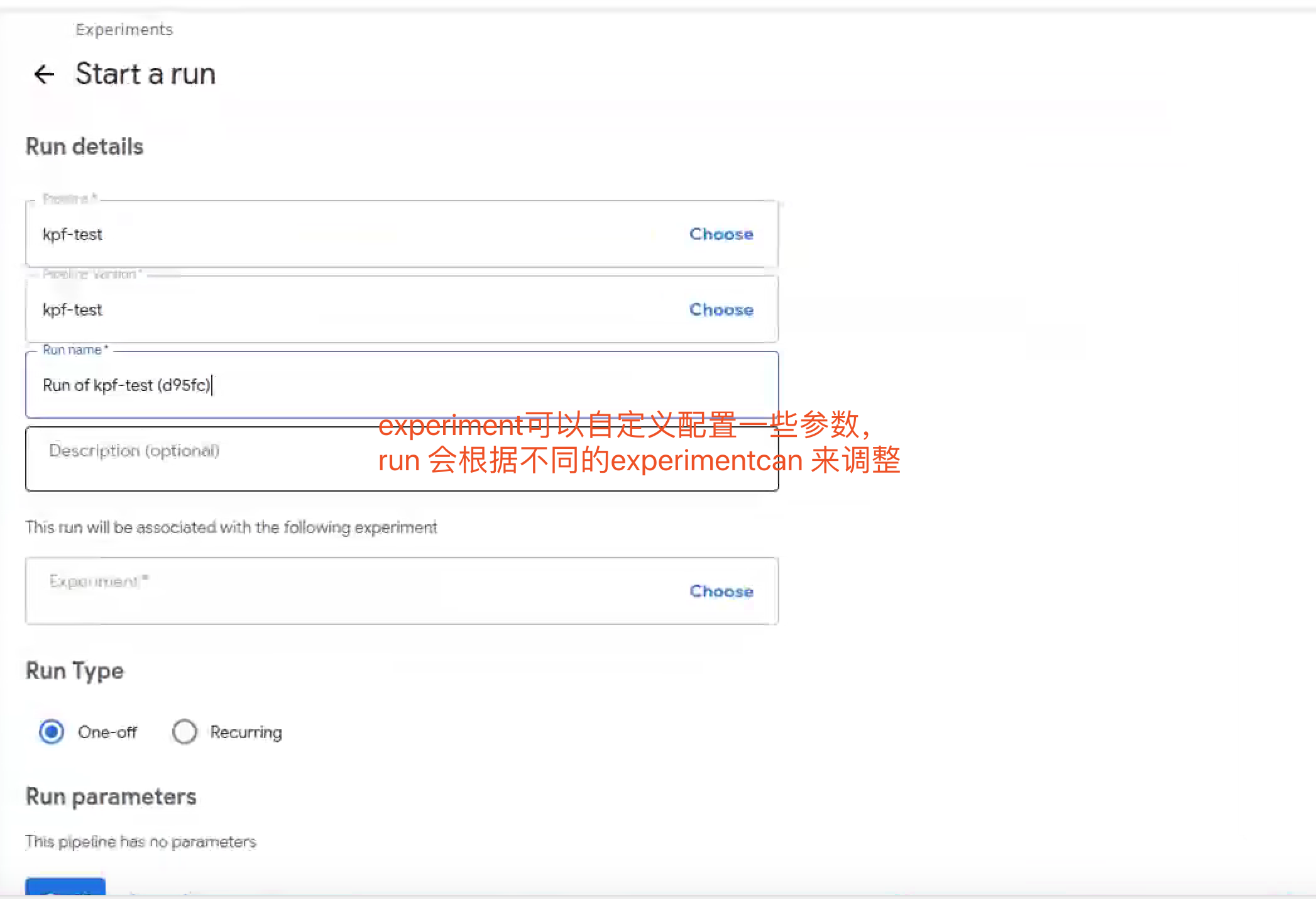


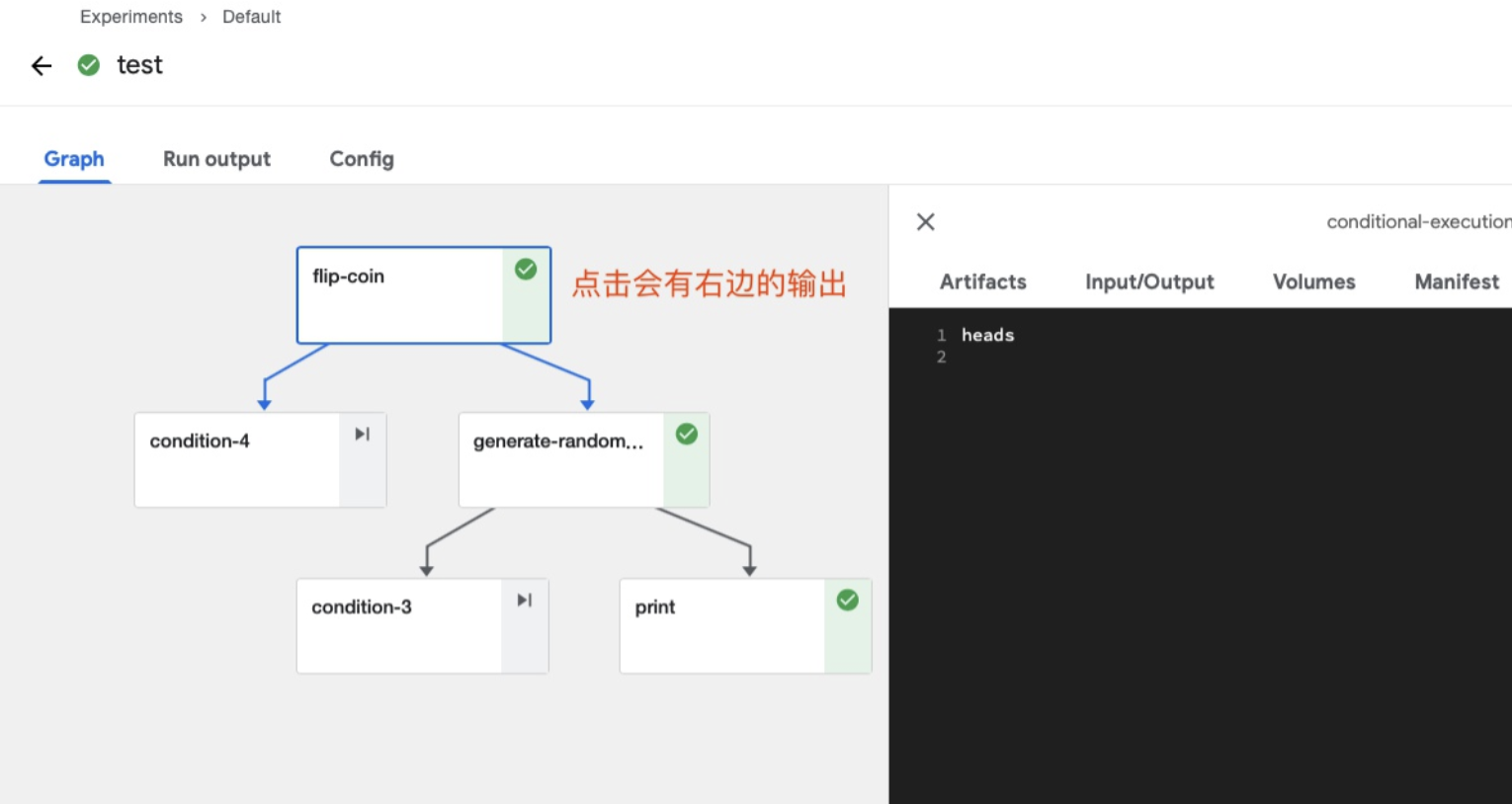


根据相应的档案生成一个 有向无环图（DAG）。其中的每一个节点就是Component（组件），而Component（组件）之间的连线代表它们之间的依赖关系









pipelines是有一个有向无环图，每一个节点上都是一个容器化的过程

每一个pipelines的组件都是一个独立的代码块，打包为Docker images。包换输入参数和输出。

组件处理真正的逻辑，比如预处理，数据清洗，模型训练等。每一个组件负责的功能不同，但有一个共同点，即组件都是以Docker镜像的方式被打包，以容器的方式被运行的。

Kubeflow Pipelines平台包括：

1. 能够运行和追踪实验的管理控制台
2. 能够执行多个机器学习步骤的工作流引擎（Argo）
3. 用来自定义工作流的SDK，目前只支持Python

而Kubeflow Pipelines的目标在于：

1. 端到端的任务编排: 支持编排和组织复杂的机器学习工作流，该工作流可以被直接触发，定时触发，也可以由事件触发，甚至可以实现由数据的变化触发
2. 简单的实验管理: 帮助数据科学家尝试众多的想法和框架，以及管理各种试验。并实现从实验到生产的轻松过渡。
3. 通过组件化方便重用: 通过重用Pipelines和组件快速创建端到端解决方案，无需每次从0开始的重新构建