

# Continuous Integration

Apache Flink's is powered by [Azure pipelines](#) and currently hosted by Ververica/Alibaba.

## GitHub

### Repositories

AzureCI requires write access to the repository, but the Apache Foundation has certain restrictions who's able to access the public repository. Therefore, a workaround is provided that requires multiple Github repositories:

- [apache/flink](#): The actual Apache Flink repository. It's hosted under the Apache Github organization.
- [flink-ci](#) is a [Ververica](#)-owned Github organization that hosts multiple repositories relevant for Apache Flink's CI infrastructure
  - [flink-ci/flink-mirror](#): Mirror repository of `apache/flink`. The AzureCI account has access to this repository and can run CI builds on the corresponding code base.
  - [flink-ci/flink](#): Mirror repository of `apache/flink` for PR branch synchronization. PR CI runs don't have the CI secrets configured. Tests requiring secrets will only run on `master` and the release branches.
  - [flink-ci/git-repo-sync](#): Initial effort around synchronizing `apache/flink` and `flink-ci/flink-mirror`. This script is still in use to synchronize `master` and the release branches of the two most-recently published Flink versions. The actual script execution is hosted on Ververica-hosted machine.
  - [flink-ci/cibot](#): Repository for the CIBot which is another synchronization mechanism that's in charge of synchronizing branches of PRs that were opened on `apache/flink`. The bot is also in charge of updating the CI information of those PRs. The bot is hosted on a Ververica-hosted machine.
  - [flink-ci/pr-labeler](#): Repository for the PR labeler. The PR labeler is used to add Flink modules as labels to the PRs. A version of the `pr-labeler` is deployed onto a Ververica-hosted machine.

### Users

- [flinkbot](#): GitHub account that's used to run [flink-ci/cibot](#) and [flink-ci/pr-labeler](#). The credentials are currently held by the PMC (see contacts below).
  - The two deployed background processes use a classic GitHub personal token that has the scope `repo` enabled.

## AzureCI Infrastructure

Builds for the `apache/flink` repository run on [Ververica](#)-sponsored Azure account `apache-flink`. The running builds can be accessed through the [AzureCI web UI](#).

The jobs are defined in the [apache/flink:azure-pipelines.yml](#). Several templates also exist in the [apache/flink:tools/azure-pipelines/](#) directory.


### Azure CI Agents

Apache Flink's AzureCI has two different types of workers which are used for the CI runs (the agent that is used for running the job is usually logged in the first line of a job run's log):

- AzureCI hosts: AzureCI provides VMs to OpenSource projects
- Alibaba-hosted machines: Alibaba provides additional VMs to host Flink's CI builds.

## Future Plans

There are plans to migrate from AzureCI to Github Actions to provide a better integrated look-and-feel for developers (

 [FLINK-27075](#) - Migrate CI from Azure to Github Actions [OPEN](#) ). This would also enable us to get rid of `flink-ci/cibot` and `flink-ci/git-repo-sync`.

## Contacts

- [Robert Metzger](#) (PMC chair)
- [Chesnay Schepler](#) (PMC): one of the main maintainers of the CI infrastructure
- [Matthias Pohl](#) (PMC): one of the maintainers of the CI infrastructures
- [Jing Ge](#): Ververica contact person
- [Yang Wang](#) (PMC): Alibaba contact person who has access to dedicated AzureCI agents.

Defined environment variables are usually controlled by the PMC; if you require access to the original keys please reach out to the PMC. Please do not add additional keys without prior notice, nor pass them on to other people as they are for the most part credentials to sponsored resources.