

Apache Hudi - Release Guide (Pre Graduation)



Inactive; For bookkeeping purposes only

RETIRED

This page has been replaced since graduation. It's no longer accurate or maintained

Introduction

This release process document is based on [Apache Beam Release Guide](#) and [Apache Flink Release Guide](#)

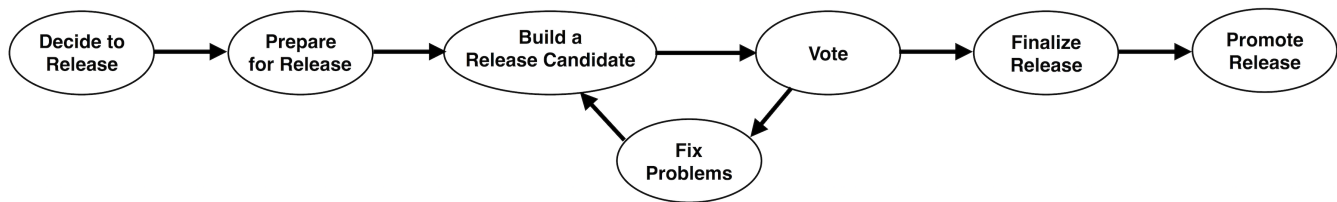
The Apache Hudi project periodically declares and publishes releases. A release is one or more packages of the project artifact(s) that are approved for general public distribution and use. They may come with various degrees of caveat regarding their perceived quality and potential for change, such as "alpha", "beta", "incubating", "stable", etc.

Hudi community treats releases with great importance. They are a public face of the project and most users interact with the project only through the releases. Releases are signed off by the entire Hudi community in a public vote.

Each release is executed by a *Release Manager*, who is selected among the Hudi PMC members. This document describes the process that the Release Manager follows to perform a release. Any changes to this process should be discussed and adopted on the dev@ mailing list.

Please remember that publishing software has legal consequences. This guide complements the foundation-wide [Product Release Policy](#) and [Release Distribution Policy](#).

Overview



The release process consists of several steps:

1. Decide to release
2. Prepare for the release
3. Build a release candidate
4. Vote on the release candidate
5. During vote process, run validation tests
6. If necessary, fix any issues and go back to step 3.
7. Finalize the release
8. Promote the release

Decide to release

Deciding to release and selecting a Release Manager is the first step of the release process. This is a consensus-based decision of the entire community.

Anybody can propose a release on the dev@ mailing list, giving a solid argument and nominating a committer as the Release Manager (including themselves). There's no formal process, no vote requirements, and no timing requirements. Any objections should be resolved by consensus before starting the release.

In general, the community prefers to have a rotating set of 3-5 Release Managers. Keeping a small core set of managers allows enough people to build expertise in this area and improve processes over time, without Release Managers needing to re-learn the processes for each release. That said, if you are a committer interested in serving the community in this way, please reach out to the community on the dev@ mailing list.

Checklist to proceed to the next step

1. Community agrees to release
2. Community selects a Release Manager

Prepare for the release

Before your first release, you should perform one-time configuration steps. This will set up your security keys for signing the release and access to various release repositories.

To prepare for each release, you should audit the project status in the JIRA issue tracker, and do the necessary bookkeeping. Finally, you should create a release branch from which individual release candidates will be built.

NOTE: If you are using [GitHub two-factor authentication](#) and haven't configure HTTPS access, please follow [the guide](#) to configure command line access.

One-time setup instructions

GPG Key

You need to have a GPG key to sign the release artifacts. Please be aware of the ASF-wide [release signing guidelines](#). If you don't have a GPG key associated with your Apache account, please create one according to the guidelines.

There are 2 ways to configure your GPG key for release, either using release automation script(which is recommended), or running all commands manually.

Use `preparation_before_release.sh` to setup GPG

- Script: [preparation_before_release.sh](#)
- Usage `./hudi/scripts/release/preparation_before_release.sh`
- Tasks included
 1. Help you create a new GPG key if you want.
 2. Configure git user.signingkey with chosen pubkey.
 3. Add chosen pubkey into [dev KEYS](#) and [release KEYS](#)
NOTES: Only PMC can write into [release repo](#).
 4. Start GPG agents.

Run all commands manually

- Get more entropy for creating a GPG key
 - `sudo apt-get install rng-tools`
 - `sudo rngd -r /dev/urandom`
- Create a GPG key
 - `gpg --full-generate-key`
- Determine your Apache GPG Key and Key ID, as follows:
 - `gpg --list-keys`
- This will list your GPG keys. One of these should reflect your Apache account, for example:
 - -----
 - pub 2048R/935D191 2019-08-29
 - uid Anonymous Anonymous <anonymous@apache.org>
 - sub 2048R/CD4C59FD 2019-08-29

Here, the key ID is the 8-digit hex string in the pub line: 845E6689 or more than 8-digit hex string like 623E08E06DB376684FB9599A3F5953147903948A. Now, add your Apache GPG key to the Hudi's KEYS file both in [dev](#) and [release](#) repositories at [dist.apache.org](#). Follow the instructions listed at the top of these files. (Note: Only PMC members have write access to the release repository. If you end up getting 403 errors ask on the mailing list for assistance.)

- Configure git to use this key when signing code by giving it your key ID, as follows:
 - `git config --global user.signingkey CD4C59FD, or git config --global user.signingkey 623E08E06DB376684FB9599A3F5953147903948A`
 - You may drop the `--global` option if you'd prefer to use this key for the current repository only.
- Start GPG agent in order to unlock your GPG key
 - `eval $(gpg-agent --daemon --no-grab --write-env-file $HOME/.gpg-agent-info)`
 - `export GPG_TTY=$(tty)`
 - `export GPG_AGENT_INFO`

Access to Apache Nexus repository

Configure access to the [Apache Nexus repository](#), which enables final deployment of releases to the Maven Central Repository.

1. You log in with your Apache account.
2. Confirm you have appropriate access by finding `org.apache.hudi` under Staging Profiles.
3. Navigate to your Profile (top right dropdown menu of the page).
4. Choose User Token from the dropdown, then click Access User Token. Copy a snippet of the Maven XML configuration block.
5. Insert this snippet twice into your global Maven settings.xml file, typically `$(HOME)/.m2/settings.xml`. The end result should look like this, where `TOKEN_NAME` and `TOKEN_PASSWORD` are your secret tokens:

```
<settings>
  <servers>
    <server>
      <id>apache.releases.https</id>
      <username>TOKEN_NAME</username>
      <password>TOKEN_PASSWORD</password>
    </server>
    <server>
      <id>apache.snapshots.https</id>
      <username>TOKEN_NAME</username>
      <password>TOKEN_PASSWORD</password>
    </server>
  </servers>
</settings>
```

Submit your GPG public key into MIT PGP Public Key Server

In order to make yourself have the right permission to stage java artifacts in Apache Nexus staging repository, please submit your GPG public key into [MIT PGP Public Key Server](#). If you have problem in submit your GPG key to the key server in browse, please try run

```
gpg --keyserver hkp://pool.sks-keyservers.net --send-keys ${PUBLIC_KEY} and verify via gpg --keyserver
hkp://pool.sks-keyservers.net --recv-keys ${PUBLIC_KEY}
```

also send public key to ubuntu server via

```
gpg --keyserver hkp://keyserver.ubuntu.com --send-keys ${PUBLIC_KEY} # send public key to ubuntu server
gpg --keyserver hkp://keyserver.ubuntu.com --recv-keys ${PUBLIC_KEY} # verify
```

would also refer to [stackoverflow](#).

Create a new version in JIRA

When contributors resolve an issue in JIRA, they are tagging it with a release that will contain their changes. With the release currently underway, new issues should be resolved against a subsequent future release. Therefore, you should create a release item for this subsequent release, as follows:

Attention: Only PMC has permission to perform this. If you are not a PMC, please ask for help in dev@ mailing list.

1. In [JIRA](#), navigate to [Hudi > Administration > Versions](#).
2. Add a new release. Choose the next minor version number after the version currently underway, select the release cut date (today's date) as the Start Date, and choose Add.
3. At the end of the release, go to the same page and mark the recently released version as released. Use the ... menu and choose Release.

Triage release-blocking issues in JIRA

There could be outstanding release-blocking issues, which should be triaged before proceeding to build a release candidate. We track them by assigning a specific Fix version field even before the issue resolved.

The list of release-blocking issues is available at the [version status page](#). Triage each unresolved issue with one of the following resolutions:

For all JIRA issues:

- If the issue has been resolved and JIRA was not updated, resolve it accordingly.

For JIRA issues with type "Bug" or labeled "flaky":

- If the issue is a known continuously failing test, it is not acceptable to defer this until the next release. Please work with the Hudi community to resolve the issue.
- If the issue is a known flaky test, make an attempt to delegate a fix. However, if the issue may take too long to fix (to the discretion of the release manager):
 - Delegate manual testing of the flaky issue to ensure no release blocking issues.
 - Update the Fix Version field to the version of the next release. Please consider discussing this with stakeholders and the dev@ mailing list, as appropriate.

For all other JIRA issues:

- If the issue has not been resolved and it is acceptable to defer this until the next release, update the Fix Version field to the new version you just created. Please consider discussing this with stakeholders and the dev@ mailing list, as appropriate.
- If the issue has not been resolved and it is not acceptable to release until it is fixed, the release cannot proceed. Instead, work with the Hudi community to resolve the issue.

If there is a bug found in the RC creation process/tools, those issues should be considered high priority and fixed in 7 days.

Review Release Notes in JIRA

JIRA automatically generates Release Notes based on the Fix Version field applied to issues. Release Notes are intended for Hudi users (not Hudi committers/contributors). You should ensure that Release Notes are informative and useful.

Open the release notes from the [version status page](#) by choosing the release underway and clicking Release Notes.

You should verify that the issues listed automatically by JIRA are appropriate to appear in the Release Notes. Specifically, issues should:

- Be appropriately classified as Bug, New Feature, Improvement, etc.
- Represent noteworthy user-facing changes, such as new functionality, backward-incompatible API changes, or performance improvements.
- Have occurred since the previous release; an issue that was introduced and fixed between releases should not appear in the Release Notes.
- Have an issue title that makes sense when read on its own.

Verify that a Release Build Works

Run `mvn -Prelease clean install` to ensure that the build processes are in good shape.

Create a release branch in apache/hudi repository

Attention: Only committer has permission to create release branch in apache/hudi.

Release candidates are built from a release branch. As a final step in preparation for the release, you should create the release branch, push it to the Apache code repository, and update version information on the original branch.

Export Some Environment variables in the terminal where you are running the release scripts

- `export RELEASE_VERSION=<RELEASE_VERSION_TO_BE_PUBLISHED>`
- `export NEXT_VERSION=<NEW_VERSION_IN_MASTER>`
- `export RELEASE_BRANCH=release-<RELEASE_VERSION_TO_BE_PUBLISHED>`
- `export RC_NUM=<release_candidate_num_starting_from_1>`

Use `cut_release_branch.sh` to cut a release branch

- Script: [cut_release_branch.sh](#)

Usage

```
# Cut a release branch
Cd scripts && ./release/cut_release_branch.sh \
--release=${RELEASE_VERSION} \
--next_release=${NEXT_VERSION} \
--rc_num=${RC_NUM}
# Show help page
./hudi/scripts/release/cut_release_branch.sh -h
```

Checklist to proceed to the next step

- Release Manager's GPG key is published to [dist.apache.org](#)
- Release Manager's GPG key is configured in git configuration
- Release Manager has org.apache.hudi listed under Staging Profiles in Nexus
- Release Manager's Nexus User Token is configured in settings.xml
- JIRA release item for the subsequent release has been created
- All test failures from branch verification have associated JIRA issues
- There are no release blocking JIRA issues
- Release branch has been created
- Release Notes have been audited and added to [RELEASE_NOTES.md](#)

Build a release candidate

The core of the release process is the build-vote-fix cycle. Each cycle produces one release candidate. The Release Manager repeats this cycle until the community approves one release candidate, which is then finalized.

Set up a few environment variables to simplify Maven commands that follow. This identifies the release candidate being built. Start with RC_NUM equal to 1 and increment it for each candidate. Also, As Hudi is in incubation phase, all artifacts must include “-incubating” in their names

- a. `git checkout ${RELEASE_BRANCH}`
- b. Run `mvn version` to set the proper rc number in all artifacts
 - i. `mvn versions:set -DnewVersion=${RELEASE_VERSION}-incubating-rc${RC_NUM}`
- c. Run Unit tests and ensure they succeed
 - i. `mvn test -DskipITs=true`
- d. Run Integration Tests and ensure they succeed
 - i. `mvn verify -DskipUTs=true`
- e. Commit and push this change to RELEASE branch
 - i. `git commit -am "Bumping release candidate number ${RC_NUM}"`
 - ii. `git push origin ${RELEASE_BRANCH}`
- f. Generate Source Release:
 - i. This will create the tarball under `hudi/src_release` directory
 - i. `cd hudi/src_release`
 - ii. `gpg --verify hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM}.src.tgz.asc hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM}.src.tgz`
 1. You should verify that the signature is good
 - iii. `tar -zxvf hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM}.src.tgz && cd hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM} && mvn clean package -DskipTests`
 - iv. If they pass, delete the repository we got from the tar-ball
 1. `cd ../ && rm -rf hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM}`
 - i. `git checkout ${RELEASE_BRANCH}`
 - ii. `cd scripts && ./release/create_source_release.sh`
 - iii. Verify Source release is signed and buildable
- g. Create tag
 - i. `git tag -s ${RELEASE_VERSION}-incubating-rc${RC_NUM} "${RELEASE_VERSION}"`
 - ii. if apache repo is origin.
 1. `git push origin ${RELEASE_VERSION}-incubating-rc${RC_NUM}`
- h. Deploy maven artifacts and verify
 - i. This will deploy jar artifacts to the [Apache Nexus Repository](https://repository.apache.org/), which is the staging area for deploying jars to Maven Central.
 - ii. Review all staged artifacts (<https://repository.apache.org/>). They should contain all relevant parts for each module, including pom.xml, jar, test jar, source, test source, javadoc, etc. Carefully review any new artifacts.
 - i. `git checkout ${RELEASE_BRANCH}`
 - ii. This will deploy 2.11 artifacts to repository.apache.org
 1. `cd scripts && ./release/deploy_staging_jars.sh --scala_version=2.11`
 - iii. This will deploy 2.12 artifacts to repository.apache.org.
 1. `./release/deploy_staging_jars.sh --scala_version=2.12`
 - iv. Review all staged artifacts to ensure it contains both 2.11 and 2.12 artifacts, mainly `hudi-spark-bundle-2.11/2.12`, `hudi-spark-2.11/2.12`, `hudi-utilities-bundle_2.11/2.12` and `hudi-utilities_2.11/2.12`.
 - v. Close the staging repository on Apache Nexus. When prompted for a description, enter “Apache Hudi (incubating), version \${RELEASE_VERSION}-incubating, release candidate \${RC_NUM}”.
- i. Stage source releases on dist.apache.org
 - i. If you have not already, check out the Hudi section of the dev incubator repository on dist.apache.org via Subversion. In a fresh directory
 - ii. if you would not checkout, please try `svn checkout https://dist.apache.org/repos/dist/dev/incubator/hudi` again.
 1. `svn checkout https://dist.apache.org/repos/dist/dev/incubator/hudi --depth=immediates`
 - iii. Make a directory for the new release:
 1. `mkdir hudi/hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM}`
 - iv. Copy Hudi source distributions, hashes, and GPG signature:
 1. `mv <hudi-dir>/src_release/* hudi/hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM}`
 - v. Add and commit all the files.
 1. `cd hudi`
 2. `svn add hudi-${RELEASE_VERSION}-incubating-rc${RC_NUM}`
 3. `svn commit`
 - vi. Verify that files are [present](#)

vii. Run Verification Script to ensure the source release is sane

```
1. cd scripts && ./release/validate_staged_release.sh --release=${RELEASE_VERSION} --rc_num=${RC_NUM} --verbose
```

Checklist to proceed to the next step

1. Maven artifacts deployed to the staging repository of repository.apache.org
2. Source distribution deployed to the dev repository of dist.apache.org and validated

Vote on the release candidate

Apache Hudi needs rounds of voting

- Voting on dev@hudi group
- ~~Voting on general@incubator group~~

Once you have built and individually reviewed the release candidate, please share it for the community-wide review. Please review foundation-wide [voting guidelines](#) for more information.

Start the review-and-vote thread on the dev@ mailing list. Here's an email template; please adjust as you see fit.

From: Release Manager

To: dev@hudi.apache.org or general@incubator.apache.org

Subject: [VOTE] Release 1.2.3, release candidate #3

Hi everyone,

Please review and vote on the release candidate #3 for the version 1.2.3, as follows:

[] +1, Approve the release

[] -1, Do not approve the release (please provide specific comments)

The complete staging area is available for your review, which includes:

* JIRA release notes [1],

* the official Apache source release and binary convenience releases to be deployed to dist.apache.org [2], which are signed with the key with fingerprint FFFFFFFF [3],

* all artifacts to be deployed to the Maven Central Repository [4],

* source code tag "release-1.2.3-rc3" [5],

The vote will be open for at least 72 hours. It is adopted by majority approval, with at least 3 PMC affirmative votes.

Thanks,

Release Manager

[1] link

[2] link

[3] <https://dist.apache.org/repos/dist/release/incubator/hudi/KEYS>

[4] link

[5] link

[6] link

If there are any issues found in the release candidate, reply on the vote thread to cancel the vote, and there's no need to wait 72 hours if any issues found. Proceed to the Fix Issues step below and address the problem. However, some issues don't require cancellation. For example, if an issue is found in the website, just correct it on the spot and the vote can continue as-is.

If there are no issues, reply on the vote thread to close the voting. Then, tally the votes in a separate email. Here's an email template; please adjust as you see fit.

From: Release Manager

To: dev@hudi.apache.org

Subject: [RESULT] [VOTE] Release 1.2.3, release candidate #3

I'm happy to announce that we have unanimously approved this release.

There are XXX approving votes, XXX of which are binding:

- * approver 1
- * approver 2
- * approver 3
- * approver 4

There are no disapproving votes.

Thanks everyone!

Please look at previous examples in previous releases. For example : Please see examples here : [voting in dev](#) , [voting in general](#) and [result of voting](#)

Checklist to proceed to the finalization step

1. Community votes to release the proposed candidate, with at least three approving PMC votes

Fix any issues

Any issues identified during the community review and vote should be fixed in this step.

Code changes should be proposed as standard pull requests to the master branch and reviewed using the normal contributing process. Then, relevant changes should be cherry-picked into the release branch. The cherry-pick commits should then be proposed as the pull requests against the release branch, again reviewed and merged using the normal contributing process.

Once all issues have been resolved, you should go back and build a new release candidate with these changes.

Checklist to proceed to the next step

1. Issues identified during vote have been resolved, with fixes committed to the release branch.

Finalize the release

Once the release candidate has been reviewed and approved by the community, the release should be finalized. This involves the final deployment of the release candidate to the release repositories, merging of the website changes, etc.

Current follow the below steps to finalize the release.

1. change the version from `${RELEASE_VERSION}-incubating-rc${RC_NUM}` to `${RELEASE_VERSION}-incubating` against release branch, use command `mvn versions:set -DnewVersion=${RELEASE_VERSION}-incubating`, e.g. change `0.5.1-incubating-rc1` to `0.5.1-incubating`.
2. Repeat the steps from Generate Source Release(f) to Stage source releases on [dist.apache.org](https://dist.apache.org/repos/dist/release/incubator/hudi)(i). **Note that make sure remove the -rc\${RC_NUM} suffix when repeat the above steps. and please also verify the steps.**
3. One more step is to deploy source code to release dist. <https://dist.apache.org/repos/dist/release/incubator/hudi>.
 - a. `svn checkout https://dist.apache.org/repos/dist/release/incubator/hudi --depth=immediates`, if you would not checkout, please try `svn checkout https://dist.apache.org/repos/dist/release/incubator/hudi` again.
 - b. Make a directory for the new release:
 - i. `mkdir hudi/hudi-${RELEASE_VERSION}-incubating`
 - c. Copy Hudi source distributions, hashes, and GPG signature:
 - i. `mv <hudi-dir>/src_release/* hudi/hudi-${RELEASE_VERSION}-incubating`
 - d. Add and commit all the files.

- i. `cd hudi`
 - ii. `svn add hudi-${RELEASE_VERSION}-incubating`
 - iii. `svn commit`
- e. Verify that files are [present](#)
4. Use the Apache Nexus repository to release the staged binary artifacts to the Maven Central repository. In the Staging Repositories section, find the relevant release candidate `org.apache.hudi-xxx` entry and click Release. Drop all other release candidates that are not being released.
5. In Jira, go to Releases <Release Version> and ensure that all Jiras for the release are 'Closed' state, if not transition all 'Resolved' jiras to 'Closed'.
6. Finalize the Release in Jira by providing the release date.
7. Update [DOAP file](#) in the root of the project via sending a PR like [this one](#).

Steps to cut doc version and update website.

1. Follow the [instructions](#) for cutting a doc for this new release.
2. [Build the site locally](#) and ensure the new doc version is available as intended.
3. Update site using [instructions](#)

Promote the release

Once the release has been finalized, the last step of the process is to promote the release within the project and beyond. Please wait for 24h after finalizing the release in accordance with the [ASF release policy](#).

Apache mailing lists

Announce on the `dev@` mailing list that the release has been finished.

Announce on the release on the `user@` mailing list, listing major improvements and contributions.

Announce the release on the announce@apache.org , general@incubator.apache.org mailing lists.

Considering that `announce@` ML has restrictions on what is published, we can follow this email template:

From: Release Manager

To: announce@hudi.apache.org

Subject: [ANNOUNCE] Apache Hudi (incubating) <VERSION> released

The Apache Hudi(incubating) team is pleased to announce the release of Apache Hudi incubating <VERSION>.

Apache Hudi (incubating) (pronounced Hoodie) stands for Hadoop Upserts Deletes and Incrementals. Apache Hudi (incubating) manages storage of large analytical datasets on DFS (Cloud stores, HDFS or any Hadoop FileSystem compatible storage) and provides the ability to query them.

This release comes xxx months after xxx. It includes more than xxx resolved issues, comprising of a few new features as well as general improvements and bug-fixes. It includes support for xxx, xxx, xxx, and many more bug fixes and improvements.

You can start using it in Maven by simply updating your dependency to:

```
<dependency>
<groupId>org.apache.hudi</groupId>
<artifactId>hudi-core</artifactId>
<version><VERSION>-incubating</version>
</dependency>
```

If you'd like to download the source release, you can find it here:

<https://github.com/apache/incubator-hudi/releases/tag/release-<VERSION>-incubating>

You can read more about the release (including release notes) here:

https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12322822&version=<JIRA_VERSION>

We welcome your help and feedback. For more information on how to report problems, and to get involved, visit the project website at:

<http://hudi.apache.org/>

Thanks to everyone involved!

XXX

Recordkeeping

Use reporter.apache.org to seed the information about the release into future project reports.

Social media

Tweet, post on Facebook, LinkedIn, and other platforms. Ask other contributors to do the same.

Improve the process

It is important that we improve the release processes over time. Once you've finished the release, please take a step back and look what areas of this process can be improved. Perhaps some part of the process can be simplified. Perhaps parts of this guide can be clarified.

If we have specific ideas, please start a discussion on the dev@ mailing list and/or propose a pull request to update this guide. Thanks!