Time Based Release Plan

We plan to move to a time-based release cadence for Apache Kafka. This document highlights the motivation for a time-based release and discusses how time-based releases will work for Apache Kafka in detail.

Motivation

The benefits of moving to a time-based release are:

1. A quicker feedback cycle and users can benefit from features shipped quicker
2. Predictability for contributors and users:
    a. Developers and reviewers can decide in advance what release they are aiming for with specific features.
    b. If a feature misses a release we have a good idea of when it will show up.
    c. Users know when to expect their features
3. Transparency - There will be a published cut-off date (AKA feature freeze) for the release and people will know about it in advance. Hopefully, this will remove the contention around which features make it.
4. Quality - we've seen issues pop up in release candidates due to last-minute features that didn't have proper time to bake in. More time between feature freeze and release will let us test more, document more, and resolve more issues.

Since nothing is ever perfect, there will be some downsides:

1. Most notably, features that miss the feature-freeze date for a release will have to wait a few months for the next release. Features will reach users faster overall as per benefit #1, but individual features that just miss the cut will lose out
2. More releases a year mean that being a committer is more work - release management is still some headache and we'll have more of those. Hopefully, we'll get better at it. Also, the committer list is growing and hopefully, it will be less than a once-a-year effort for each committer.
3. For users, figuring out which release to use and having frequent new releases to upgrade to may be a bit confusing.
4. Frequent releases mean we need to do bugfix releases for older branches. Right now we only do bugfix releases to the latest release.

We decided to experiment with time-based releases and see if the benefits for us as a community exceed the drawbacks. We will regularly iterate to improve our release process to have a community of happy developers and users as well as regular high-quality releases.

What will the Apache Kafka development process look like?

The goal is to have **3 releases a year**.

We are planning to make a release every 4 months.

A month before the release date, the release manager will cut branches and also publish (preferably on the wiki) a list of features that will be included in the release (these will typically be KIPs, but not always). We will leave another week for "minor" features to get in (see below for definitions), but at this point, we will start efforts to stabilize the release branch and contribute mostly to tests and fixes. Two weeks after branch cutting, we will announce code-freeze and start rolling out RCs, after which only fixes for blocking bugs will be merged.

Definitions

- Major Feature - Feature that takes more than 2 weeks to develop and stabilize. Requires more than one PR to complete the feature.
- Minor Feature - Feature that takes less than 2 weeks to develop and stabilize. Requires mostly one PR to complete the feature.
- Stabilization - This phase would involve writing incremental system tests for features and fixing any bugs identified in the checked-in feature.
- Feature Freeze - Major features should have only stabilization remaining. Minor features should have a PR in progress that can be checked in by a week. A release branch will be cut at this point.
- Code Freeze - Development is stopped. Blocker bugs will be fixed after the code freeze. The first RC will be created at this point.

For time-based releases, we will strictly ensure that a release happens on a given date. For example, in the first release, we will decide to have a release by the middle of October and we will stick to it. We will drop features that we think will not make it into the release at the time of feature freeze and also avoid taking on new features in the release branch. Trunk development can continue as usual and those features will be in the following release. Ideally, we would have started stabilization around this time. About two weeks before the release date, we would call for code freeze and code checkins will be allowed only if any blocker bugs are identified. In a rare scenario, we could end up with a feature that passed the feature freeze bar but still fails to complete on time. Such features will also be dropped from the release at the end to meet the release deadline.

What happens if features don't complete?

Features tend to be of different complexity. Some of them can be implemented within a single release while some span multiple releases. With time-based releases, we would need to ensure that ongoing features do not affect the stability of the release. There are a couple of options –

1. Ensure that every feature is broken down into testable units and only testable units get checked in. This means that a good set of unit test and system tests are written for sub-tasks before they are checked in. This will ensure that `trunk is in a relatively stable state to release at any point of time.
2. Use feature branches. Create branches from trunk that are focused on a specific feature. The feature developers ensure that the branch is in sync with `trunk` from time to time. Once there is high level of confidence on the stability, it can be merged into `trunk` . This approach has the additional overhead of branching and performing merges from time to time.

In practice, the right approach would be a mix of both ‘a’ and ‘b’. The feature developers need to make this call depending on the complexity of the feature.

What are the gaps that we need to focus on?

There are some gaps in the development process and testing infrastructure that we need to close to ensure that we can successfully do a time based release. This is going to be a long term effort but will simplify time based releases as we make more progress in closing them.

1. Testing coverage with check ins. Code reviews in Apache Kafka need to ensure we have good unit test coverage and enforce system tests be written along with the check ins (when applicable). Testing should not be at the end. We should also encourage documentation be accompanied with check ins where applicable.
2. Compatibility testing for AK. We have system tests for verifying the behaviour of AK rolling upgrades. These tests help detect unexpected incompatible protocol and ZK structure changes, but they are not sufficient to ensure that an upgrade from an AK trunk commit to another AK trunk commit works correctly (LinkedIn reported one such bug just before 0.10.0.0 was released). In addition, we don't have anything that verifies API source/binary compatibility (KAFKA-1880 includes a WIP PR for this). We would need to extend our suite of tests to provide the required coverage.

What Is Our EOL Policy?

Given 3 releases a year and the fact that no one upgrades three times a year, we propose making sure (by testing!) that rolling upgrade can be done from each release in the past year (i.e. last 3 releases) to the latest version.

We will also attempt, as a community to do bugfix releases as needed for the last 3 releases.

Who Manages The Releases?

As usual, a committer shall volunteer. If no committer volunteers, we'll cancel a release due to lack of interest.

What About Version Numbers?

Since the 1.0.0 release in Oct 2017, our current versioning protocol is the following:

1. We will use three digits for the version: *major.minor.bug-fix.* The first digit would indicate the major revision (starting at 1), and second indicating minor revision, and the last one number indicating the bug-fix revision.
2. When preparing RCs of the release we will still suffix it after the release version number. I.e. *major.minor.bug-fix [rc_number]*

Feature releases will be a minor release by default (i.e. we will only bump up the minor revision digit) unless:

1. We change the message format in Kafka.
2. We break compatibility (i.e. remove deprecated public methods after a reasonable period).
3. We make major version changes on our dependent projects or when we drop an old Java / Scala support version.
4. We do something totally amazing (exactly once?) and decide to release as a new major version milestone.

In those cases, we will bump the major revision digit.

Schedule

The proposed schedule for Apache Kafka is shown below. We will do a release in October and align the releases from January in 2017. We will follow a 4 month schedule for Apache Kafka releases.

Oct 2016

Jan 2017

May 2017

Sept 2017