

Adding a new component to Apache Bigtop bigdata stack

- [1. Introduction](#)
- [2. Hard expectations](#)
- [3. Soft expectations](#)

1. Introduction

In the last a couple of years we have seen a rapid influx of new components getting added to the Apache bigdata stack. It is becoming more crucial to outline clean guidelines on how to add new components to the mix. Bigtop motto is "Debian of Big Data" as such we are trying to be as inclusive as possible. However, certain constraints exist and have to be addressed accordingly. While we are trying to provide as full list of such requirements as possible, the list provided below might not be complete.

Bigtop stack introduces the notion of a component maintainer (see MAINTAINERS.txt top level file) and thus, it is expected that a front line of support for certain areas would be provided by certain individuals. However, the long-term maintenance costs are expected to be shared among all the members of our community. Hence it is important for us to make sure that all members are comfortable with all the projects that are getting added at least at a very basic level.

2. Hard expectations

This is a list of requirement that we don't want to deviate from unless there is a really major shift in Bigtop's charter. Projects that violate at least one of the following would have to go through community review and PMC approval on a case-by-case basis:

1. Code is expected to be Licensed under Apache License, Version 2.0 (and their dependencies are expected to be compatible with this license)
2. There's an active and interested contributor ready to make the necessary patches for inclusion
3. The project is expected to integrate well into "big data management software distribution based on Apache Hadoop". In other words, it's not a one-off, and has multiple integration points with the rest of our stack.
4. The project is expected to be unavailable (at least the desired version) from major Linux distributions (Debian, Ubuntu, RedHat, SuSE). In other words, we don't want to duplicate the effort,, which is already done elsewhere unless there's a very strong reason to do so.
5. The project is expected to be compatible with all of the supported platforms where Bigtop stack is officially supported by this community
6. Patches are expected to be added to the trunk first before adding to any released branches. In rear cases, Bigtop can patch a component in flight
 - a. to change the order of dependencies resolution ie to guarantee local artifacts are picked first
 - b. outright broken component's release build ([HADOOP-11489](#))
 - c. some other special conditions, considered by the community
7. The following is expected to be provided with each patch adding a new project to Bigtop distribution:
 - a. packaging code and packaging tests
 - b. deployment code
 - c. smoke testing code
8. The contribution passes project's CI and builds with the standard Bigtop toolchain

3. Soft expectations

Violating any of these expectations will require an explicit explanation attached to the proposal. They are flexible, but it doesn't mean that they can be disregarded willy-nilly.

1. Project provides test artifacts that go beyond our basic smoke-testing requirement (integration testing)
2. Project is an Apache Software Foundations project (top-level or incubating).
3. The project produces at least one executable system artifact. There are exceptions to the rule like Apache Tez, which is a library.